
BAYESIAN ACTIVE-PERCEPTION:
AN INFORMATIC VIEWPOINT

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

David Heavens MSc (Edinburgh) BSc (Surrey)
Department of Engineering
University of Leicester

March 2009

Bayesian active-perception: An informatic viewpoint

DAVID HEAVENS

Abstract

Localisation of places, prey and predators are usually of critical behavioural importance to an organism's survival. In this thesis, I conduct an investigation of localisation, considering specifically the inference of a target, event or the observer itself. I begin with an exploratory investigation into auditory localisation of a single sound source for a static (passive) observer. I evaluate the influence (sensitivity) of "cue" variables on localisation by the curvature of the location belief's Kullback-Leibler divergence. More generally, from this I observed a symbol grounding problem – corresponding one location to a data sample due to multiple locations mapping onto a single observed value. I demonstrate how action can support the grounding of symbols by breaking such symmetries (inference confusions) that exist in passive localisation. By considering the breaking of these symmetries, I go on to develop an information measure that generally selects the best localising action. This is the action expected to give the "next best view" for the system, hence removing ambiguities and uncertainties in inference with the greatest efficiency.

From these considerations, my main contribution is a general theoretical framework for selecting between actions during localisation and inference tasks according to an observer's representation. I illustrate this framework by using it to select head casts in localising binaural level cues for sound source localisation. Further illustration is through a learning problem, where I evaluate learning performance during directed and undirected selection of actions. This demonstrates how directed action is important in symbol grounding of the latent state space to the observation space. Because of its generality, my Bayesian-active perception framework may be used to derive novel domain specific action-selection and learning algorithms that optimise inference. It may also provide a principled account for existing action-selection algorithms (for instance in robotics) and specific animal behaviours as special cases.

Acknowledgements

First I would like to thank Tim Pearce, my supervisor for his help and understanding during the difficult birth of my thesis. Next I would like to thank Fernando Schlindwein, Postgraduate Tutor, for his great help in the administrative side of my PhD studies. My thanks also to my immediate family: my parents and little sister Sarah who helped me keep everything in context with their emotional support. But especially to my parents who without their help, both emotionally and materially, I would have been in no position to submit my thesis.

To all of my friends and family, who have continually helped keep things in perspective and provided much needed support during both my time in Leicester, and, finishing-up and writing-up at home. With special thanks to the now defunct *Coffee Meeting and Unofficial Support Group* which grew very large by the end, with special thanks to Gail Iles, Gautier Torricelli and Antonios Tzimzas. Thanks also to all of my Ultimate Frisbee team mates and opposing friends both past and present, in Cambridge and Leicester, to Edward Green especially. To Charlotte who helped me stay sane in Ipswich. And finally to my oldest friends, Robin, Michael and Mark for being there through thick and thin.

I would like to thank the EPSRC for providing funding during the course of my studies.

Contents

1	Introduction	1
1.1	Aims and motivations	1
1.2	Sound source localisation in biology	3
1.2.1	Auditory localisation systems neurophysiology	3
1.2.2	Auditory localisation modelling	8
1.2.3	Influence of motion upon sound localisation	11
1.3	Philosophy of active perception	16
1.4	Examples of active perception	18
1.4.1	Sensorimotor contingencies	18
1.4.2	Robotic mapping, SLAM/CML	22
1.4.3	Best localisation actions, Bayesian state estimation	23
1.5	Thesis structure	25
2	Bayesian approaches to passive sound source localisation	30
2.1	Introduction	30
2.2	Localisation model	35
2.2.1	Sound process	35
2.2.2	Cue processes	38
2.2.3	The joint distribution	42
2.2.4	Latent variable inference	42
2.2.5	Learning model parameters by expectation-maximisation (EM)	43
2.3	Results	46
2.3.1	Nix and Hohmann’s approach	48
2.3.2	Comparison of inferences	49
2.3.3	Analysis of inferences	53
2.4	The contribution of cues	57

2.5	Discussion	63
2.5.1	Findings	63
2.5.2	Limitations	65
2.5.3	Possible modifications	68
2.5.4	Summary	70
3	An information measure for optimal action selection	72
3.1	Introduction	72
3.2	Hypothesis of active perception	76
3.3	Sketch of Infomax	77
3.4	Properties of my hypothesis	83
3.5	Algorithmic complexity	88
3.6	Simulated head rotations for active sound source localisation	96
3.6.1	Analysis	103
3.6.2	An alternative stochastic strategy	106
3.7	Comparison with Infotaxis (Vergassola <i>et al.</i> , 2007a)	109
3.7.1	Analysis	114
3.8	Discussion	121
3.8.1	Findings	121
3.8.2	Limitations	123
3.8.3	Future work & summary	124
4	Learning a model for an active agent	126
4.1	Introduction	126
4.1.1	Hypothesis	129
4.2	Episodic learning	130
4.2.1	Learning requirements	130
4.2.2	Episodic maximum-likelihood	131
4.3	Systems background	133
4.4	Partially observable switched state space model	136
4.4.1	Background	136
4.4.2	System models algorithmic definition	137
4.5	Measures of expected knowledge	140
4.5.1	Adaptation of Porta <i>et al.</i> 's approach (Infomax)	141

4.5.2	Application of CRLB	141
4.5.3	Application of my approach from Chapter 3	143
4.6	Results	148
4.6.1	Comparison between informatic policies	150
4.6.2	A stochastic modification	157
4.6.3	Directed localising action is better than undirected actions	158
4.6.4	“Adult” learning of correspondence	163
4.7	Discussion	167
4.7.1	Findings	167
4.7.2	Limitations	170
4.7.3	Behavioural implications	171
5	Concluding remarks	173
5.1	Preamble	173
5.2	Sound source localisation	173
5.3	Optimal action selection	176
5.3.1	Future comparisons	182
5.4	Active learning	184
5.5	Final remarks	185
A	Analysis of the EM-algorithm	187
B	The variational responsibilities and updates	189
C	Nix and Hohmann (2006) ILD and IPD computations	192
D	The Kullback-Leibler divergence	194
E	Local weight analysis	197
F	Levels versus level disparities	204
F.1	The distribution of levels	204
F.2	The distribution of level disparities	208
G	Jensen’s inequality	210
H	Optimal statistical filtering	211

I	Information identities	213
J	My measure for a sequence	216
K	POMDP informatic action policy comparisons	219
L	Gaussian notation	222
M	Bayes law	223
N	Gaussian identities	224
O	Matrix identities	225
P	Trace equivalence	226
Q	Sampling from the Lorenz attractor	228
R	Episodic learning	230
	R.1 Learning requirements	230
	R.2 Episodic maximum-likelihood	231
	R.3 The dissimilarity between parameters	232
	R.4 The Lagrangian multiplier	234
S	Particle filtering	237
T	The RBPF model	240
U	Kalman filtering & smoothing	245
	U.1 The Kalman filter	245
	U.2 Kalman filter properties	246
	U.3 Kalman filtering	247
	U.3.1 The likelihood of a trajectory	252
	U.4 Kalman smoothing	253
V	Learning updates	259
	Bibliography	265

List of Figures

1.1	General depiction of the action-perception inference problem.	2
1.2	An individual with three similar independent sound sources.	5
1.3	Schematic of neural connectivity of the LSO.	6
1.4	The ILD pathway in the auditory midbrain.	7
1.5	The functional connectivity of the mustached bat’s auditory pathway.	7
1.6	A toy example of symmetry breaking using an idealised ILD example.	9
1.7	A toy example of symmetry breaking using an idealised ITD example.	10
1.8	Cartoon representation of the auditory motional experiments conducted by Klensch (1948), Jongkees and van de Veer (1958).	13
1.9	A toy instantiation of pinnae motion illustrating symmetry breaking.	14
2.1	The response of an LSO neuron to the loudness at either ear.	31
2.2	Graphical model indicating the dependencies between measurables, hidden vari- ables and location.	34
2.3	Figure 3 taken from Nix and Hohmann (2006).	39
2.4	Plots of the distribution of ILD for variation of signal-to-noise ratio.	40
2.5	Cartoon depicting the spatial distribution of sound source localisation experimental measurements.	47
2.6	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	50
2.7	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	50
2.8	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	51
2.9	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	51
2.10	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	52
2.11	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	52
2.12	An auditory location belief for my model contrasted to Nix and Hohmann’s model.	53
2.13	The uncertainty and accuracy of auditory location beliefs.	56

2.14	The uncertainty and accuracy of auditory location beliefs for the Nix and Hohmann (2006) approach.	57
2.15	Graphical models depicting the perturbations of interest for the dependencies in Figure 2.2.	58
2.16	Curvature analysis of the perturbations of interest from Figure 2.15.	62
2.17	Curvature analysis of the perturbations of interest from Figure 2.15 continued.	63
2.18	Distribution on the λ_1, λ_2 plane of the data likelihoods, illustrating the need for a joint distribution of $p(\lambda_1, \lambda_2 l)$	67
2.19	Learnt parameters for each location using my model in Figure 2.2.	68
2.20	The updated graphical model indicating the dependencies between measurables, hidden variables and location.	69
3.1	A depiction of the active-perception inference problem.	74
3.2	Graphical model of the action-perception model, repeated from Figure 3.1.	78
3.3	Cartoons of the consequence of “good” and “bad” localising actions.	80
3.4	A toy example of the best and worst selections of action for inference using an idealised ILD.	81
3.5	A toy example of the best and worst selections of action for inference using an idealised ITD.	82
3.6	Cartoons indicating a toy auditory head rotation problem.	96
3.7	Contours of egocentric location belief for a sequence of randomly selected actions.	97
3.8	Time course plots indicating that inference certainty increases in the same fashion as inference accuracy.	101
3.9	Time course plots of the change in inference certainty and inference accuracy.	102
3.10	Scatter plots indicating the correlation of accuracy with certainty, and the correlation of the equivalent changes.	102
3.11	Scatter plots indicating covariation of accuracy with certainty, and covariation of a one step look ahead accuracy with certainty.	104
3.12	The ROC illustrating performance of the information measures ordering of actions in all situations.	105
3.13	The ROC illustrating performance of the information measures ordering of actions in only the most uncertain and ambiguous situations.	105
3.14	Histograms of actions for the informatic policies.	106
3.15	The conditional histograms of actions for each informatic policy.	107

3.16	The ROC that a stochastic policy will favour taking the best action according to accuracy or certainty.	108
3.17	The ROC that a stochastic policy will favour taking the best action according to accuracy or certainty.	108
3.18	A single track for the agent using each of the Boundtaxis and Infotaxis policies. . .	115
3.19	Comparison between the performance of Infotaxis and Boundtaxis search times. . .	116
3.20	Comparison between the PDFs of the search times for Infotaxis and Boundtaxis. . .	116
3.21	Comparison between the residual entropy of the Infotaxis and Boundtaxis posterior beliefs versus the remaining search time.	117
3.22	Comparison between the PDFs of the search times for Infotaxis, Boundtaxis and a variety of other policies from Vergassola <i>et al.</i> (2007b).	117
3.23	The tracks of individual runs for the Boundtaxis policy.	118
3.24	The PDF of search times with the initial distance.	119
4.1	The graphical model indicating the action-perception dependencies of a sequence of sample and learn episodes.	130
4.2	Convergence of the episodic ML estimates.	133
4.3	The Lorenz attractor.	134
4.4	The sensitivity of the Lorenz attractor to small perturbations of the initial starting conditions.	134
4.5	Graphical models representing a generalised dynamic transition process and a switching dynamic transition process with similar measurement processes.	138
4.6	A visual interpretation of the measurement process for a contrived example. . . .	146
4.7	The likelihoods of the learning data sets and validation data sets for each of the informatic policies defined in Algorithms 4.2, 4.3 & 4.4 contrasted with a uniform random policy.	149
4.8	Figure indicating the episodic evolution, of each policy, for the learnt models' inference "certainties" and "accuracies".	150
4.9	The data trajectory, inferred trajectories and the learnt measurement process for the <i>a posteriori</i> Cramer-Rao lower bound (CRLB).	151
4.10	The data trajectory, inferred trajectories and the learnt measurement process for the adapted Infomax approach of Porta <i>et al.</i> (2003, 2005).	152
4.11	The data trajectory, inferred trajectories and the learnt measurement process for a policy based upon my approach from Chapter 3.	153

4.12	The data trajectory, inferred trajectories and the learnt measurement process for a uniform random policy.	154
4.13	The learning data sets likelihood across different weightings of uniform to greedy stochastic action selection.	158
4.14	The range of validation data sets' likelihood across different weightings of uniform to greedy stochastic action selection.	159
4.15	The data trajectory, inferred trajectories and the learnt measurement process. . .	160
4.16	Figure of the final learnt models' inference certainties and accuracies across the different weightings of uniform to greedy stochastic action selection.	161
4.17	Expansion of Figure 4.16 to indicate the episodic evolution of the certainties and accuracies.	161
4.18	The learning and validation likelihoods for relearning a model with a change to the system.	163
4.19	The KL-divergence of the learnt parameters with episodes after a change in the system's measurement process.	164
4.20	The data trajectory, inferred trajectories and the learnt measurement process for the initial system in Figure 4.18.	165
4.21	The data trajectory, inferred trajectories and the learnt measurement process after the change in the system in Figure 4.18.	166
4.22	Plot of the smoothness of the action selection policies; the different weightings of uniform to greedy stochastic action selection.	171
5.1	Cartoons of the consequence of "good" and "bad" localising actions, repeat of Figure 3.3.	176
D.1	Figure illustrating the impact that varying the mean offset between two Gaussians has upon the KL-divergence.	195
F.1	Graphical models depicting the relationship between the latent variables and the observable variables in a model of sound source localisation.	205
F.2	Plots of the distribution of sound levels with attenuation.	206
F.3	Plots of the distribution of sound levels with signal-to-noise ratio.	207
F.4	The negentropy of the distribution of levels for variation of attenuation and signal-to-noise ratio.	207
F.5	Plots of the distribution of ILD for variation of signal-to-noise ratio.	209

F.6	The negentropy of the ILD distributions for variation of attenuation and signal-to-noise ratio.	209
Q.1	The Lorenz attractor.	229
R.1	Convergence of episodic ML estimates compared with the ratio γ for initial divergent proposals.	236
S.1	Graphical model of the general filtering problem.	237

List of Tables

1.2	A summary of the state-of-the-art.	28
2.1	Table of the important quantities for the sound source localisation model in Figure 2.2.	36
2.2	Table of distributions for my models variables listed in Table 2.1 and Figure 2.2.	36
2.3	Table of parameters for the PDFs in Table 2.2.	36
2.4	Tabulation of the performances for my model compared to a description of its posterior belief for every sound source location.	55
2.5	Tabulation of the performances for Nix and Hohmann's model compared to a description of its posterior belief for every sound source location.	55
3.1	The tabulated performance of information policies for all situation and ambiguous situations.	104
3.2	The tabulated performance of stochastic policies derived from the behaviour of the information policies for all situation and ambiguous situations.	107

List of Algorithms

2.1	The variational EM algorithm for learning the parameters for the likelihood of my sound source localisation model.	45
2.2	Pseudocode used to generate inferences of sound source location.	49
2.3	My curvature analysis for my sound source localisation model.	61
3.1	The generic particle filtering algorithm.	91
3.2	The algorithm used by Porta <i>et al.</i> to calculate Infomax for the robotic situation as described in Porta <i>et al.</i> (2003, 2005).	92
3.3	An implementation of my information measure for the robotic situation described in Porta <i>et al.</i> (2003, 2005).	93
3.4	An alternative implementation of my information measure for a robotic situation described in Porta <i>et al.</i> (2003, 2005).	94
3.5	Calculation of Infomax for active sound source localisation using a particle filter defined in Algorithm 3.1.	99
3.6	An implementation of my information measure for active sound source localisation using a particle filter defined according to Algorithm 3.1.	100
3.7	The Infotaxis algorithm, implementing Vergassola <i>et al.</i> (2007a,b), for selecting the action that maximises the expected change in Shannon information.	111
3.8	The Boundtaxis algorithm for selecting the action that maximises a term proportional to the upper bound of the expected Shannon information.	114
4.1	A generic Rao-Blackwellized particle filtering algorithm.	139
4.2	Porta <i>et al.</i> 's Infomax calculation algorithm adapted to a partially observable switching SSM.	142
4.3	An implementation of the <i>a posteriori</i> CRLB adapted to a partially observable switching SSM.	144
4.4	An implementation of my information measure adapted to a partially observable switching SSM.	147

4.5	The protocol for learning episodically using a variety of informatic policies.	148
4.6	The protocol for learning episodically using a stochastic policy.	157
K.1	An Infomax algorithm for a POMDP.	220
K.2	An implementation of my information measure for a POMDP.	220
S.1	The generic particle filtering algorithm.	238
T.1	The Rao-Blackwellized particle filtering algorithm for a switching SSM problem.	243
T.2	The Rao-Blackwellized particle smoothing algorithm for a switching SSM problem.	244

Glossary of terms

Acoustic flow is the auditory equivalent of optic flow. It uses the dynamic variation of auditory variables to directly perform useful tasks, an example of which is *breaking control*.

Agent denotes the perceiving robot or simulated entity under evaluation.

Breaking control is the behavioural task of dynamically controlling motion according to the variation of an observable variable. It can imply a number of tasks such as, breaking to avoid a collision to safely land upon a surface, or to capture a target.

CCD denotes *charge coupled device*.

CF bats emit a *constant frequency* (CF) echolocating pulse of sound. This is in contrast to frequency modulating (FM) bats that modulate their echolocating pulses to form a descending frequency sweep over the pulses duration.

CML abbreviates *concurrent mapping and localisation*.

CN abbreviates the *cochlea nucleus*.

CRLB abbreviates the *Cramer-Rao lower bound*. This is the lower bound of the FIM.

Dichotic is typically used to denote the simultaneous presentation of sounds to either ear.

DNLL abbreviates the *dorsal nucleus of the lateral lemniscus*.

EM-algorithm abbreviates the *expectation maximisation algorithm*. It is an approach to ML learning for hidden variable problems by iteratively optimising a lower bound of the models likelihood.

FFT abbreviates the *fast Fourier transform*.

FIM abbreviates the *Fisher information matrix*.

GTM abbreviates the *generative topographic model*.

HMM abbreviates the *hidden Markov model*. This is a statistical model of a Markov process with an unobservable (latent) state. Typically this is used to refer to discrete states.

HRTF abbreviates the *head related transfer function*. This is a direction and sound frequency dependent gain function. It is used to normalise between individual physiology in behavioural hearing experiments.

IC abbreviates the *inferior colliculus*.

ILD is the *interaural level disparities*, and is associated like (ITD) with lateralisation of sound sources. ILD is defined as the difference in perceived loudness between either ear for a sound event. ILD is synonymous with interaural intensity disparities (IID) however it emphasises that it is a level (logarithmic) rather than intensity (linear) comparison.

Individual denotes a biological agent, either human or animal.

Infomax denotes information maximisation.

INLL abbreviates the *intermediate nucleus of the lateral lemniscus*.

ITD is the *interaural time disparities*, and is associated like (ILD) with lateralisation of sound sources. ITD is defined as the difference in time of arrival at either ear of a sound source. ITD is synonymous with interaural phase disparities (IPD) where IPDs emphasis the comparison of the two signals in the frequency domain to estimate ITD.

Latent state space denoted by the symbol L , is the hidden state of interest for the problem at hand. An instance on this space is denoted by l . An example of this (l) would be the target location, where L would represent all possible locations. This space L is however only a representation of the systems true latent space L^{true} .

Lateralisation is the task of inferring the lateral angle of a sound relative to the listener.

Learning data set is used to parameterise a model through an optimisation algorithm such as the EM-algorithm.

LSO abbreviates the *lateral superior olive*.

LWPR abbreviates *locally weighted projection regression*.

Measurement space denoted by the symbol X , is the space of possible observations for making inferences upon a *latent state space*. A measurement from this space is denoted by x . Examples of measurables can include raw sensory signals or filtered signals, such as ILD and ITD.

MCMC abbreviates *Markov chain Monte-Carlo*. It specifies a problem with random sampling for a Markov problem.

MGB abbreviates the *medial geniculate body*.

ML abbreviates a *maximum likelihood* optimisation algorithm. It represents the optimisation of a model using the likelihood that a model generated a *learning data set*.

MNTB abbreviates the *medial nucleus of the trapezoid body*.

NBV abbreviates the *next best view*.

NLL abbreviates the *nucleus of the lateral lemniscus*.

ODE abbreviates *ordinary differential equations*.

PDF abbreviates *probability density function*.

PF abbreviates *particle filtering*.

POMDP abbreviates *partially observable Markov decision processes*.

RBPF abbreviates *Rao-Blackwellized particle filtering*.

SLAM abbreviates *simultaneous localisation and mapping*.

SOC abbreviates the *superior olivary complex*.

SSM abbreviates the *state space model*. This is a HMM for continuous latent state variables, for example the Kalman filter is an SSM.

Symmetry breaking relates to the behavioural consequence of the mathematical problem of many possible *latent state space* point mapping to one point on an *observable space*. Meaning that there is symmetry between these points. If through some operation upon the process which produces this mapping a second observation can preclude a large portion of the previously possible latent state space points then such an operation can be described as symmetry breaking.

Validation data set is used as a fair test of performance for an optimised model. It is distinct from the learning data set, thus it is assumed to be identically distributed to the learning data set.

Variational EM-algorithm is a variant of the EM-algorithm which utilises variation Bayesian methods to treat the parameters as just another latent variable.

VNLL abbreviates the *ventral nucleus of the lateral lemniscus*.

VR abbreviates *virtual reality*.

Glossary of notation

i, j, k, n denote integers.

$i : j$ denotes an inclusive slice of integers equivalent to the set $[i, j]$. This is used to indicate via a subscript a sequence over a range, e.g. $x_{i:j}$.

$u, u_t, u_{t_1:t_2}$ indicates the notation for a variable u , the variable at a time t (u_t) and a trajectory of these variables between the times t_1 and t_2 ($u_{t_1:t_2}$).

$\mathbf{x}, \mathbf{z}, \mathbf{s}$ denote vectors.

\mathbf{x}, x denotes a measurement variable. In Chapter 2 the measurement variable takes the form

$\bar{\mathbf{x}}_{k,j,n}$ where j indexes the sound source locations for the n 'th segment of the received signal.

\mathbf{y} denotes a validation recording and has a notation similar to \mathbf{x} . This is specific to Chapter 2.

\mathbf{s} denotes the sound sources signal. This is specific to Chapter 2.

λ denotes the attenuation of a sound signal. This is specific to Chapter 2.

τ denotes the ITD, though in the text δt is sometimes used. This is specific to Chapter 2.

ν, σ, ω are used to denote the variance of variables.

Σ, Q are used to denote the covariance matrices of vector variables.

I denotes the identity matrix.

$0_{D_1}, 0_{D_1 \times D_2}$ denote respectively a vector of zeros with length D_1 , and a matrix of zeros with D_1 columns and D_2 rows.

Unitalicised capital letters denote a set. This set can form a space, in which case it is a complete set. If this is a space the associated lower case letters denote instantiations upon this space. Special cases are:

x, X represents respectively a measurement variable x and the space of possible measurements X , where $x \in X$.

l, L represents respectively a latent state variable l and the space of possible latent states L , where $l \in L$.

Probability distributions and information quantities have the following notation,

$P(i), P_i$ is used to represent a probability distribution on a discrete variable i .

$P(i|\dots)$ is used to represent a conditional discrete probability distribution.

$p(x)$ is used to represent a probability distribution on a continuous variable x .

$p(x|\dots)$ is used to represent a conditional continuous probability distribution.

$\phi(\mathbf{x}; \boldsymbol{\mu}, \sigma), \phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma)$ are used to denote a Gaussian distribution. This is for the variable \mathbf{x} of D_x dimensions, which is parameterised by the mean $\boldsymbol{\mu}$ and either a variance σ or a covariance matrix Σ . Which notationally is $\phi(\mathbf{x}; \boldsymbol{\mu}, \sigma) = \sqrt{\frac{1}{2\pi\sigma}}^{D_x} \exp\left\{-\frac{1}{2\sigma}(\mathbf{x} - \boldsymbol{\mu})^2\right\}$, or for a covariance matrix is $\phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma) = \sqrt{\frac{1}{(2\pi)^{D_x}|\Sigma|}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}$.

$I[L|X, a]$ represents the Shannon information of the space L , given the space X and the constraint a . Which is $I[L|X, a] = \int dl dx p(l, x|a) \log p(l|x, a)$.

$D_{\text{KL}}[p(X|\dots) || q(X|\dots)]$ represents the Kullback-Leibler (KL) divergence between the probability distributions $p(x|\dots)$ and $q(x|\dots)$. Which is defined notationally to be $D_{\text{KL}}[p(X|\dots) || q(X|\dots)] = \int dx p(x|\dots) \log \frac{p(x|\dots)}{q(x|\dots)}$.

$\mathbb{E}_{p(x,y|z)}[f(x, y, z)]$ denotes the expectation of the quantity $f(x, y, z)$ over the variables x and y given z . Which is defined as $\mathbb{E}_{p(x,y|z)}[f(x, y, z)] = \int dx dy p(x, y|z) f(x, y, z)$.

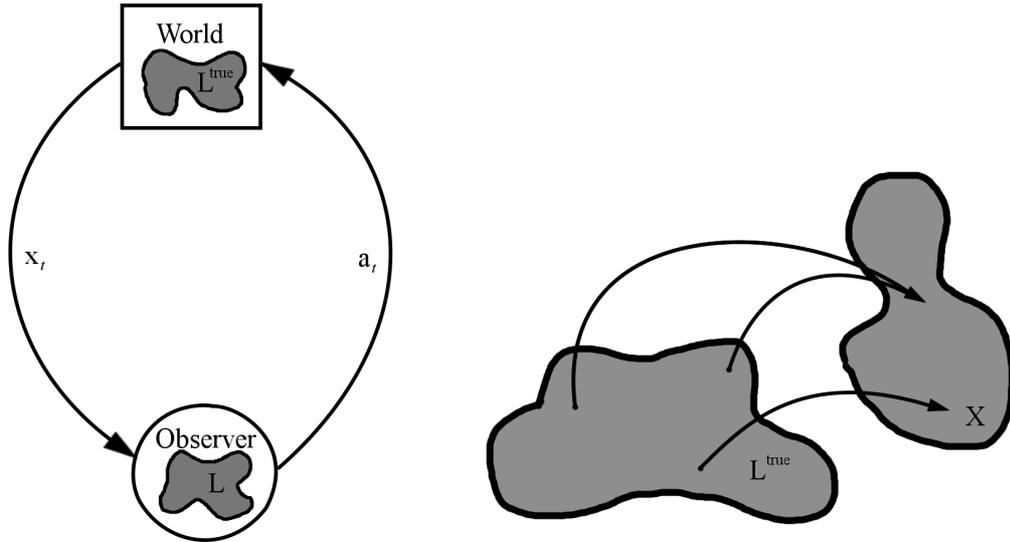
Chapter 1

Introduction

1.1 Aims and motivations

My principle research motivation is to shed light on the influence of action in perception. In this thesis I will consider an individual's perception during localisation. In fact, localisation is the unifying theme of this thesis. I will begin with the localisation of sound sources, and then I consider more generally how action influences the inference of latent state spaces. However, the original motivation for my research topic was the understanding of sound source localisation, specifically in context with the processes that echolocators use to navigate and hunt. This led to an exploratory investigation into the influences and coinfluences of the cues used in sound source localisation, as discussed in Chapter 2. The investigation indicated the difficulties in understanding perception of location as a purely passive problem, where the individual does not interact with its environment. The implication of the acoustic flow argument (Jenison, 1997, Muller and Schnitzler, 1999, 2000, 2001) is that it is not just instantaneous cues which lead to localisation, but also how these cues change. In passive localisation it is the instantaneously observed cues that define an individual's inferences. This causes passive localisation to be a deductive inference problem. There are however more encompassing arguments, where it is not only the change of these variables that leads to accurate localisation, but also the individual's knowledge of its own motion (Pettorossi *et al.*, 2005, Wallach, 1940, Lewald and Ehrenstein, 1998, Lewald and Karnath, 2001).

This fits firmly into the concept of embodiment in perception (Nagel *et al.*, 2005, Lenay *et al.*, 2001), where the joint interaction of the observer and its environment is a key factor, making the individual embodied within its environment. While reviewing the



(a) The observers model L of the world L^{true} . (b) Ambiguity in the observable X due to the mapping from L^{true} to X .

Figure 1.1: General depiction of the action-perception inference problem. Plot (a) is a cartoon depiction of the action-perception problem. The ‘world’ has a true hidden latent state space L^{true} which represents some of its characteristics which are useful for an ‘observer’ to know. The observer cannot perfectly know this true latent state space L^{true} so must approximate it with L . The ‘observer’ interacts with the ‘world’ by its actions $a_t \in A$ and its measurements $x_t \in X$. In practice this is an iterative process where an action a_t is taken and modifies the state of the world in a predictable manner. The new state $l_t^{\text{true}} \in L^{\text{true}}$ can be measured via the observable space X , which produces the measurement x_t . An example of 3 points on L^{true} mapping to X can be seen in plot (b). Of these three points on L^{true} , two are indicated as ambiguous as they both map to the same point upon X . The observer will similarly model a prediction of the transition from an old belief $l_{t-1} \in L$ to a current belief $l_t \in L$ due to an action a_t , and then refine this prediction conditional upon the measurement taken x_t .

literature I found there is a substantial body of work on the concept of action conveying useful localisation information (Noe, 2004, O'Regan and Noe, 2001, Pettorossi *et al.*, 2005, Lee *et al.*, 1992, Ashmead *et al.*, 1995, Jenison, 1997) (reviewed in Subsection 1.2.3 and Sections 1.3 & 1.4), and further that this active perception is important in an individual's perceptual development (Held and Hein, 1963). However, this leaves open the topic of how to best select actions in the light of predicted beliefs about the future – the sensor placement problem. Exceptions have addressed domains, such as problems similar to chemotaxis in turbulent flows (Vergassola *et al.*, 2007a) and robotic mapping and localisation (Fox *et al.*, 1998, Porta *et al.*, 2003). These apply an approach called Infomax, which minimises the entropy of the predicted *a posteriori* localisation belief. Infomax is a costly computation to compute. In Chapter 3, I present an alternative approach, which I show has more favourable scaling than Infomax for particular classes of problem. Arguments regarding the best localising actions typically fall under the problem of uncertainty, namely, how to best locate based upon our knowledge of the problem (*cf* Figure 1.1 (b)). Best locating actions also relate to our belief in a solution to the problem, and it is this issue that I address in Chapters 3 & 4.

1.2 Sound source localisation in biology

This section reviews sound perception in the context of sound source localisation. I start with an overview of the auditory system in Subsection 1.2.1. I focus upon bat neurophysiology, due to the preeminence of audition in bat perception (Altringham, 2001, Neuweiler, 2000, Suga, 1990). Following this I discuss various approaches in Subsection 1.2.2 to modelling the auditory perception problem. I then follow this in Subsection 1.2.3 with a review of the motional cues and the influence of active perception that contributes significantly to auditory localisation.

1.2.1 Auditory localisation systems neurophysiology

The problem of sound source localisation typically consists of locating a sound source (the target) in an environment that may have distracting sources also contributing to the sound field (Figure 1.2). When the observer is inactive and unmoving (passive observation) the principle sound cues are based upon the loudness and timing of features within the sound signals observed at either ear (Blauert, 1997, Hartmann, 1999). These cues are further categorised as either being perceived between the ears (binaural) or by one ear (monaural)

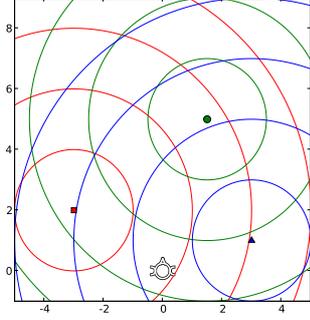
in nature (Blauert, 1997, Hartmann, 1999).

The monaural localisation cues are related to the attenuation of a sound's loudness with the distance from the animal to a target. Physiology interacts with the loudness to further attenuate the perceived loudness based upon the relative direction to the target from the observer (Kistler and Wightman, 1992, Wightman and Kistler, 1997). The methodology of applying such a directional attenuation is usually through head related transfer functions (HRTFs). Oldfield and Parker (1984) showed that this directional gain (attenuation through the HRTFs) can provide a correction to specific kinds of localisation error. However Aytekin *et al.* (2004) showed that comparing the HRTF of either ear implies the existence of binaural cues.

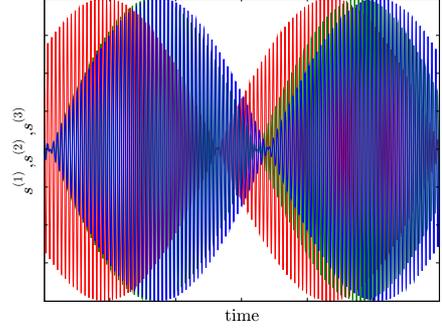
The binaural cues are principally considered as either interaural level disparities (ILD) or interaural time disparities (ITD) (Hartmann, 1999, Forsythe, 2002). In mammals, these binaural cues are typically associated with the superior olivary complex (SOC). The SOC comprises the lateral superior olive (LSO) and the medial superior olive (MSO).

The MSO is associated with sound localisation from temporal cues, such as interaural phase disparities and ITDs (Svirskis *et al.*, 2004, McAlpine and Grothe, 2003). However, there is ongoing debate as to the mechanism the MSO employs in estimating ITD (Svirskis *et al.*, 2004, McAlpine and Grothe, 2003). ITDs represent the difference in time of arrival for a recognisable feature in either ear. There are two disparate mechanisms proposed for the brain to compute these ITDs. The first is derived from the Jeffress neural delay model. This uses a bank of neurons to represent the different ITDs, akin to a bank of coincidence detectors (Pena and Konishi, 2002). The alternative is a phasic comparison between the signals, where the relative distribution of activities within the population provides localisation (McAlpine and Grothe, 2003, Joris *et al.*, 1998, Grothe, 2003). This is similar to computing a cross-correlation between the signals in the frequency domain.

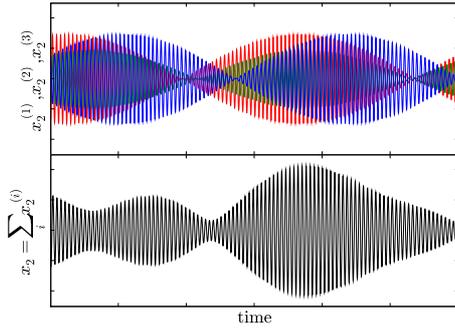
Many of the neurons in the LSO are characterised as excitatory-inhibitory (EI). This means that these neurons are excited by sound in the ipsilateral ear and suppressed by sound from the contralateral ear (Park *et al.*, 1997, Forsythe, 2002, Wang and Brown, 2006). The connectivity can be seen in Figure 1.3. The LSO is observed to encode intensity disparities between the ears (Park, 1998). Park *et al.* (1997) found the particular ILD that results in complete suppression is constant in a significant number of neurons and varies across the LSO population. This illustrates that the LSO has a population encoding strategy for ILD and is similar to the proposed IPD population mechanism for



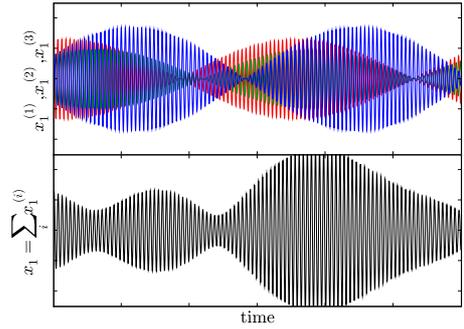
(a) Observer with three similar sound sources.



(b) The three example source signals, colour coded to the sources in plot (a).



(c) Signals received (top) and the combined signal observed at the left ear (bottom).



(d) Signals received (top) and the combined signal observed at the right ear (bottom).

Figure 1.2: An individual with three similar independent sound sources. Plot (a) depicts the propagation of the sound field for each of the sound sources by contours of the distance. This indicates the spreading of the sound field. Plot (b) shows the three source signals ($s^{(i)}$) corresponding by colour to the sound source locations in Plot (a). Plots (c) & (d) show the attenuated ($\lambda_j^{(i)}$) and delayed ($\tau_j^{(i)}$) signals (top subplots of plots (c) & (d)) and the combined signal (bottom subplots of c & d). The signals at either ear (j 'th ear) are constructed using the geometric attenuation from each source, from the i 'th source, which is $\lambda_j^{(i)} \propto \frac{1}{\text{distance to } i\text{'th ear}}$ and a delay operation D_τ of time τ for a delay of $\tau_j^{(i)} \propto \text{distance to } i\text{'th ear}$. The process mathematically is $x_j = \sum_i x_j^{(i)}$ where the contribution for each source $s^{(i)}$ each is $x_j^{(i)} = \lambda_j^{(i)} D_{\tau_j^{(i)}} s^{(i)}$.

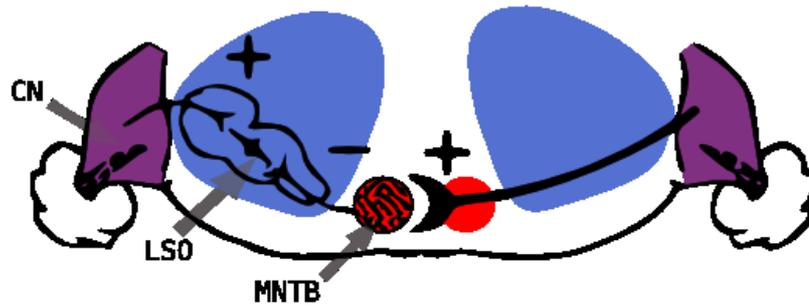


Figure 1.3: Schematic of neural connectivity of the lateral superior olive (LSO). The connectivity of the LSO showing glycinergic inhibitory (-) and excitatory (+) pathways in the Mexican free-tailed bat. Diagram adapted from Park (1998). Cochlea nucleus (CN) neurons have excitatory projections to ipsilateral LSO neurons and contralateral medial nucleus of the trapezoid body (MNTB) neurons. The MNTB neurons have inhibitory connections to the LSO. The superior olivary complex (SOC) is indicated surrounding the LSO.

ITDs (Grothe, 1994, 2003, McAlpine and Grothe, 2003).

The mid-brain encompasses the nucleus of the lateral lemniscus (NLL) and the inferior colliculus (IC). The IC has tuning for localisation cues such as ILD (Park, 1998, Harnischfeger *et al.*, 1985) and ITD (Fuzessery and Lohuis, 2003, Harnischfeger *et al.*, 1985). The LSO projects its tuning of ILD to IC neurons (Wenstrup *et al.*, 1985, Park and Pollak, 1993, Pollak *et al.*, 2002), Figure 1.4, although there is a disparity in activity profiles of ILD tuned neurons between the IC and LSO (Park and Pollak, 1993). It is unsurprising that Fuzessery and Pollak (1985) observed IC neurons tuned for direction. Relatedly, in the barn owl's neurophysiology, Pena *et al.* (2001) report a mapping similar to a Bayesian logical AND operation for combining the ITD and ILD to roughly map elevation and azimuth (Konishi, 2003). Bauer *et al.* (2002) observed the dorsal NLL (DNLL) as being selective to temporal spectra, and through its inhibition of the IC it can be seen to assist sound source selectivity (Burger and Pollak, 2001). Neurologically, this shows a form of classification used for filtering which is exhibited within the bat's auditory pathway. This can be seen as a part of the process linking the *what* with the *where* in the auditory system (Nelken *et al.*, 1999).

The auditory pathway ends at the auditory cortex (AC) where various cues are mapped, examples are:

- range (O'Neill and Suga, 1982, Jenison *et al.*, 1998), in the form of echo delay (time of flight),
- the fluttering of a target (Condon *et al.*, 1997),

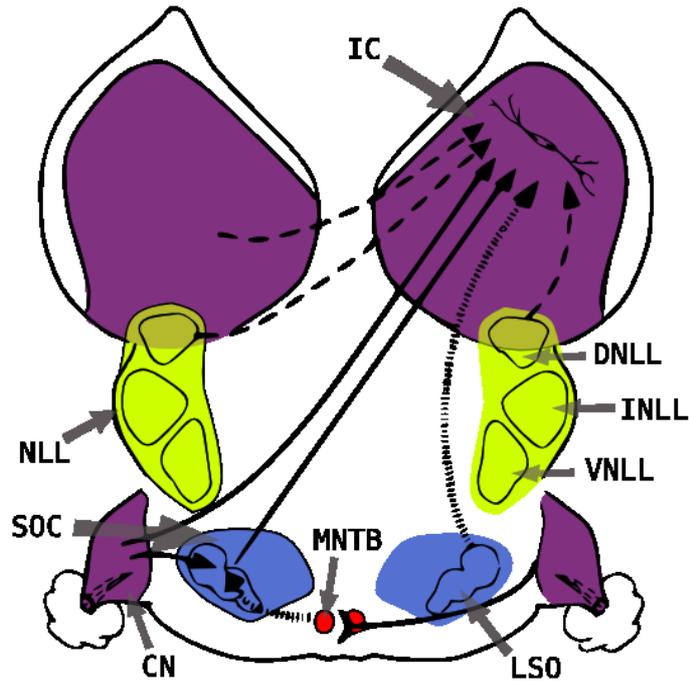


Figure 1.4: The interaural level disparity (ILD) pathway in the auditory midbrain. The connectivity of the inferior colliculus (IC), showing excitatory (—), glycinergic inhibitory (.....) and GABAergic inhibitory (---) connections in the Mexican free-tailed bat. Diagrams adapted from Pollak *et al.* (2002). The nucleus of the lateral lemniscus (NLL) is divided into three parts, the dorsal (DNLL), the ventral (VNLL) and the intermediate (INLL). Although the VNLL and INLL receive differing input from the cochlea nucleus (CN) and the medial nucleus of the trapezoid body (MNTB) there is a difference in the efficacy of these projections (Huffman and Covey, 1995). The DNLL receives binaural input from the LSO, the medial superior olive (MSO) and the CN (Yang *et al.*, 1996). It is observed to make an inhibitory projection to the IC (Burger and Pollak, 2001).

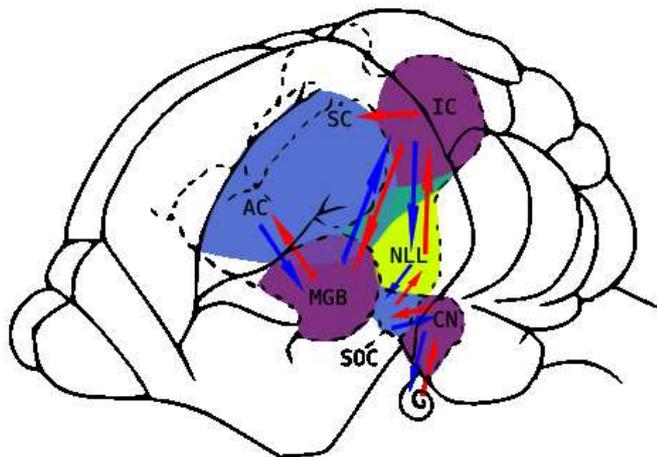


Figure 1.5: The functional connectivity of the mustached bat's auditory pathway. The ascending auditory pathway is represented by the red arrows and the descending (cortico-fugal) by the blue arrows. Diagram adapted from Suga and Ma (2003b).

- and a target's relative velocity (O'Neill, 2003).

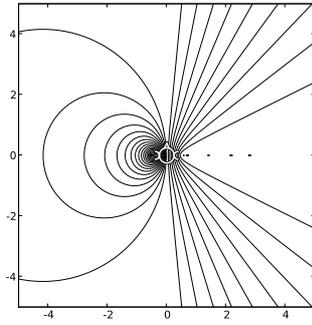
Many of these cues are also encoded at the mid-brain. Examples of this are the medial geniculate body (MGB) which is located in the thalamus (Olsen and Suga, 1991), and the cochlea nucleus (CN). Figure 1.5 shows the functional connections that make up the ascending and descending (corticofugal) auditory pathway.

In this subsection I have summarised the bat auditory system and the principle cues associated with it. I have also mentioned important aspects of human, avian and other species auditory tuning. This is to illustrate the auditory interactions available to the individual and provide neurophysiological context for the next subsection, through the biological grounding of the auditory cues.

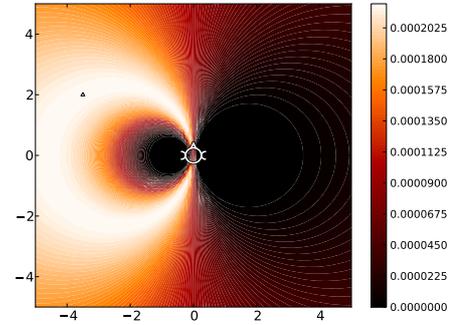
1.2.2 Auditory localisation modelling

The question of the properties of the cues themselves was considered by Zurek (1991), who derived distributions for the ILD and ITD cues. Zurek did so from the assumption that sound can be characterised as a Rayleigh intensity process, projected onto a logarithmic loudness domain. Similarly, Nix and Hohmann (2006) showed definitively that ILD and ITD sound processes are not normally distributed (Gaussian), but have higher order moments, e.g. non-zero skew (a Gaussian is symmetric and has zero skew) and non-normal Kurtosis (a Gaussian has a Kurtosis of 3). These moments complemented the lower order moments of the mean and standard deviation to convey enough information about location to localise a known sound in the presence of environmental noise. Konishi (2003) mentions the apparent evidence for the independence of ITD and ILD processing. This is principally because of the association of ITD and ILD coding for different coordinate axes of elevation and azimuth due to the barn owl's physiology. This independence was similarly assumed by Nix and Hohmann (2006) with results for localisation consistent with behavioural studies.

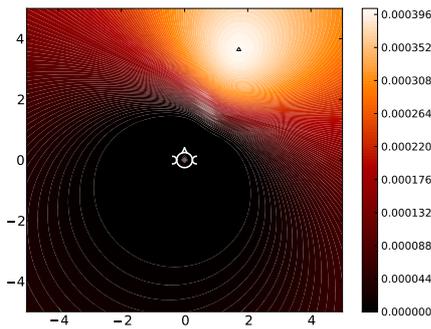
Theories relating specifically to the interaction of cues in sound localisation have typically developed from the duplex theory of sound. The duplex theory is related to the frequency properties of a sound signal. This is where higher frequencies convey greater resolution in time, making ITDs more precise and lower frequencies have less short time variation in loudness and so convey more precise ILDs. Cue distributions would hence depend upon the frequency characteristics of the sound itself (Zurek, 1991). Macpherson and Middlebrooks (2002) applied manipulations to the cues to evaluate the impact upon localisation (inference of azimuth or elevation separately) within the context of the duplex



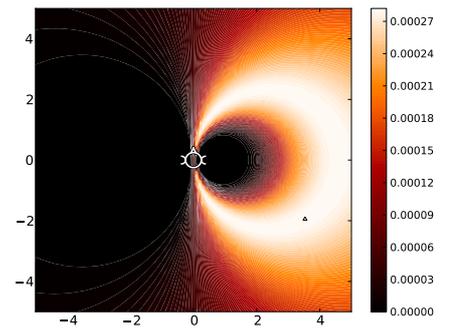
(a) Plot depicts the iso-clines for interaural level-disparity (left side) and interaural time-disparity (right side).



(b) The prior belief.

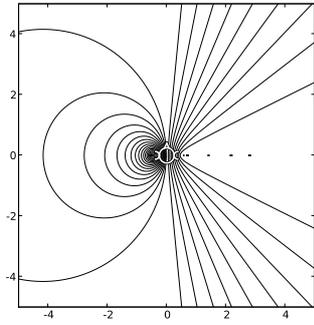


(c) The predicted posterior belief for the best action using the prior belief in plot (a).

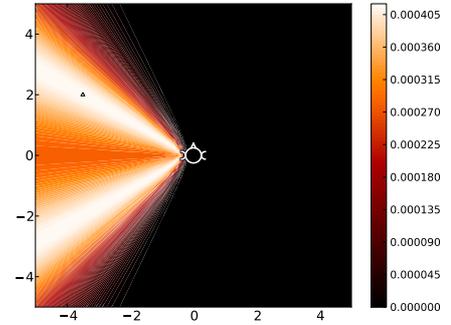


(d) The predicted posterior belief for the worst action using the prior belief in plot (a).

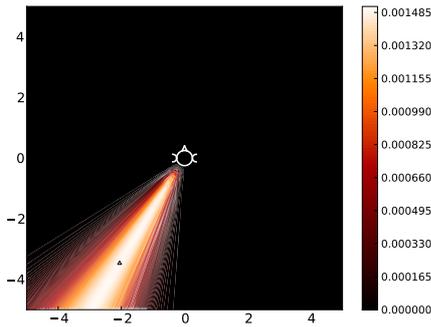
Figure 1.6: A toy example of symmetry breaking using an idealised interaural level disparity (ILD) example. The example depicts the best and worst choices of action using a free field model of ILD, the action is a head rotation. Using an idealised auditory example (plot (a)), the tori of confusion (Shinn-Cunningham *et al.*, 2000) can be seen in the iso-clines which indicate locations (ranges and directions) of equal level disparity and time disparity. Plot (b) depicts a prior belief in location where the source's position is indicated by a white filled triangle. The best and worst actions are selected according to the measure developed in Chapter 3. the expected best action (plot (c)) is clearly seen to collapse ambiguity. In contrast the expected worst action (plot (d)) retains much of the ambiguity from the prior. The units of the x and y axes are observer head widths with the observer indicated at the origin of both the x and y axes. All plots are egocentric to the agent at the origin.



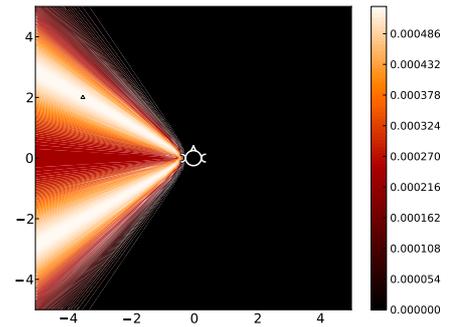
(a) Plot depicts the iso-clines for interaural level-disparity (left side) and interaural time-disparity (right side).



(b) The prior belief.



(c) The predicted posterior belief for the best action using the prior belief in plot (a).



(d) The predicted posterior belief for the worst action using the prior belief in plot (a).

Figure 1.7: A toy example of symmetry breaking using an idealised interaural time disparity (ITD) example. The example depicts the best and worst choices of action using a free field model of ITD, the action is a head rotation. Using an idealised auditory example (plot (a)), the tori of confusion (Shinn-Cunningham *et al.*, 2000) can be seen in the iso-clines which indicate locations (ranges and directions) of equal level disparity and time disparity. Plot (b) depicts a prior belief in location where the source's position is indicated by a white filled triangle. The best and worst actions are selected according to the measure developed in Chapter 3. The expected best action (plot (c)) is clearly seen to collapse ambiguity. In contrast the expected worst action (plot (d)) retains much of the ambiguity from the prior. The units of the x and y axes are observer head widths with the observer indicated at the origin of both the x and y axes. All plots are egocentric to the agent at the origin.

theory of sound localisation for broadband sounds. Similarly, Wenzel (1995) evaluated the contribution of ITD and level cues in the duplex theory but with head motion, finding that head movements helped to resolve location confusions considerably. This indicates that action plays the role of breaking symmetries in perception. These symmetries are inherent to the auditory cues, for example Figures 1.6 & 1.7. In the following subsection I review the influence of motion in sound localisation as this indicates the fundamental influence that action has upon auditory perception of location.

Pitfalls from these approaches (Nix and Hohmann, 2006, Zurek, 1991) are that:

1. the methodologies assume that the animal is given ILD and ITD variables,
2. an individual actually receives a sound wave at either ear,
3. ILD and ITD are internally computed quantities – these are referred to as hidden variables.

I tackle this problem in Chapter 2 by applying a graphical model to the problem, from which I generate a Bayesian solution to passive sound source localisation.

1.2.3 Influence of motion upon sound localisation

In this subsection I overview some of the classic behavioural work on auditory perception due to egomotion, and then review the more recent research, specifically considering the impact of egomotion upon perception of source location. This links the introduction of passive auditory localisation, in the previous two subsections, with the more general concepts of active perception and action selection explored in later sections of this literature review.

Blauert (1997) enumerates two broad and overlapping classes of motion roughly applicable to all sensory modalities:

- Reflexive being considered to be unconscious and spontaneous movement of the head towards the expected position of an auditory event which corresponds to the most probable position.
- Orienting being considered to be conscious searching movements where the goal is to assemble more information about an auditory event – the location becomes more definite during movement.

In both contexts the belief in where an event occurs should sharpen, becoming more certain and also more accurate (Thurlow and Runge, 1967). Blauert then lists three groups of questions that may be addressed:

1. Do head movement effects occur in natural hearing? If so, how are these movements instantiated?
2. For particular head movements, which attributes of ear input signals are available for interpretation?
3. What effects do head movements have on the predicted position of the auditory event?

Thurlow *et al.* (1967), Thurlow (1967) showed that the first head movement is towards an auditory event, which fits with the “reflexive” movement. A possible explanation could be due to those ILDs that correspond to forward directions having greater representation by neural units when compared with other directions (Park, 1998, Park and Pollak, 1993). It has the effect of bringing an object of interest into “view”, such that it can be located with greater accuracy. Further, Thurlow *et al.* observed that when permitted, subjects initiate prolonged and repeated head movements to determine the exact direction of a sound event – rotating and tipping head movements predominated. Wallach (1938) considered the mechanism of a rotating head movement to provide the angle of elevation for a sound event. The cues vary in a coherent manner according to direction. So if an agent knows the rotational shift and the function (of direction) that produces the change in its auditory cues, it may infer the elevation. This indicates that head movements coupled to the sensory consequences of these movements are useful for localisation.

Klensch (1948) considered how the interaction of the cues and head movements caused a sound’s location to be perceived. Klensch’s findings are summarised in Figure 1.8. It is important to note that the apparent location of the sound migrates due to the changes in the perceived sound coupled to the head’s motion. Figures 1.8 (b), (d) & (e) show the considerable impact a simple head motion has upon where a sound will be localised. Figure 1.8 (a) represents the control situation, and Figures 1.8 (a) & (c) indicate that head motion impacts localisation by perturbing sound cues. Similar interactions (Figure 1.8) were investigated by Jongkees and van de Veer (1958), Hofman *et al.* (2002), Wightman and Kistler (1999), Wallach (1938, 1939). Wallach (1938, 1939) used a slightly different methodology where an array of speakers were placed in an arc around the subject with a

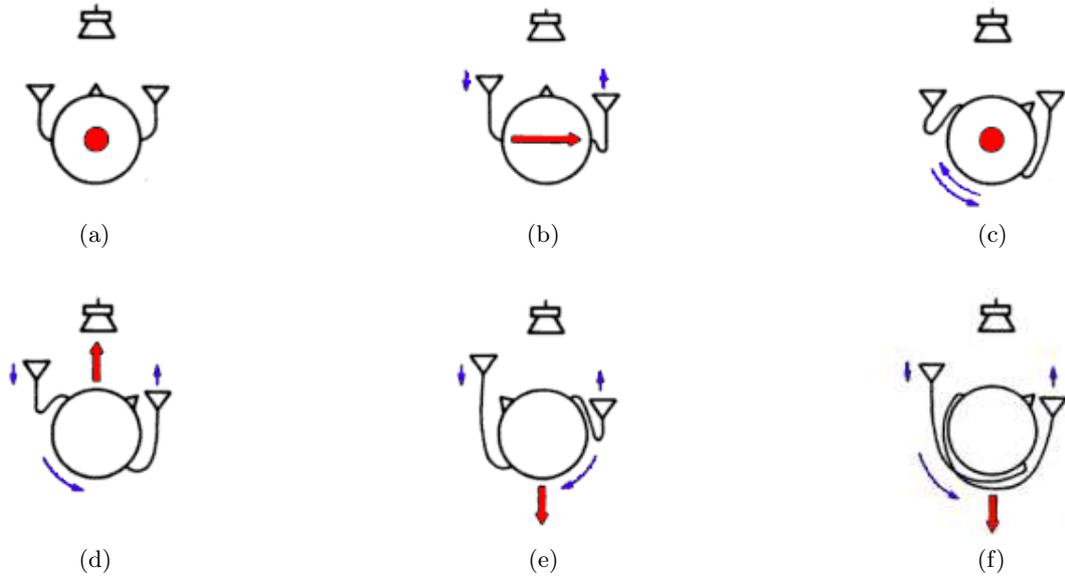


Figure 1.8: Cartoon representation of the auditory motional experiments conducted by Klensch (1948), Jongkees and van de Veer (1958). The bold red arrows indicate the perceived motion of a sound after an indicated head motion and/or motion of the hearing funnels (small blue arrows), a red dot indicates the perceived location of a sound in cases (a) & (c) in the centre of the head (*cf* when listening to sounds via headphones). Adapted from Blauert (1997).

rotary switch connected to the subject’s head. This caused a head movement to switch between which speaker was active. This took into account the HRTF cues that Klensch (1948) neglected through his use of hearing funnels. Wallach found that causing a sound to always stay at a constant angle to the subject, resulted in them reporting a migration of the sound event from its true location to a position above their head, *cf* Figure 1.8 (c). Conversely, the apparent location of the sound migrated behind the subject when a counter rotation in the angle of the sound’s movement was engineered relative to the subject’s head movement, *cf* Figure 1.8 (e) & (f). Finally, Wallach contrived to make the source appear at a particular angle and so changed the ratio at which the rotary switch switched.

Jenison (1997) argues for the sufficiency of acoustic information extraction for motion. This is through the use of acoustic cues (interaural time disparity, sound level and Doppler shift) as the observables for structural information (spectrum, loudness, target location and motion) and the hidden variables about a sound source. When the system is static it leaves some of the hidden variables insufficiently defined, however, motion by the observer corrects this. Hence, Jenison concluded that observer motion via a head rotation can disambiguate this uncertainty by essentially giving a second view of the scene. Handzel and Krishnaprasad (2002) took these ideas and tested them, by using binaurally placed acoustic sensors to computationally investigate (biomimetically) sound localisation. They

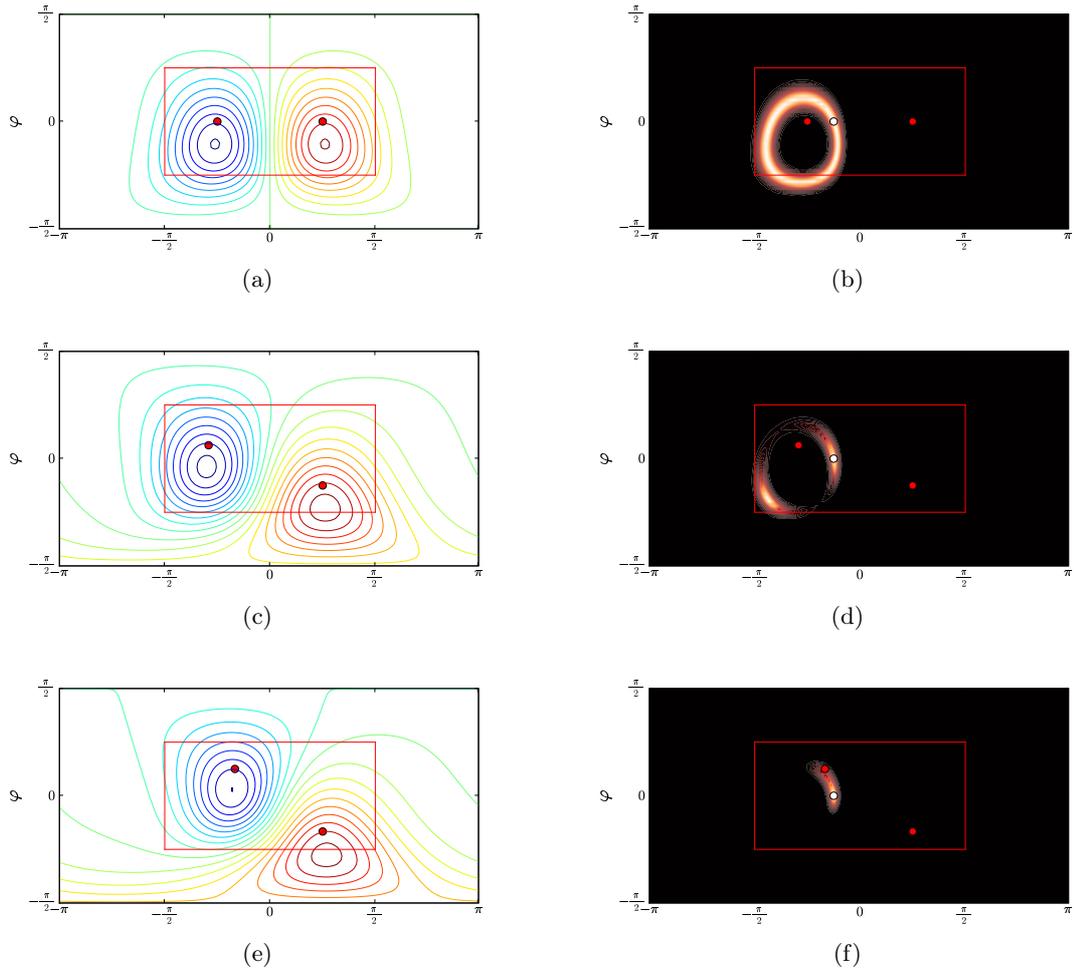


Figure 1.9: A toy instantiation of pinnae motion illustrating symmetry breaking. This example is inspired by Walker *et al.* (1998). Plots (a), (c) & (e) show the ILD with direction (elevation φ and azimuth θ) for a sequence of pinnae configurations. The ILDs are computed by the difference between a pair of basis functions with coordinates indicated by the red filled circles. The red rectangle indicates the forward quadrant of the observer. Plots (b), (d) & (f) show sequential inference of a target at a direction indicated by a white circle. The figure indicates the utility of combining information from a sequence of uncertain or ambiguous measurements, that combined, affect a change of view.

found that symmetry in static perception is broken by rotation of the apparatus (*cf* head rotation) relative to the sound source. This is similar to Jenison’s findings. Relatedly Muller and Schnitzler (1999) hypothesised that CF-bats evaluate time-variant cues (*cf* Jenison, 1997) defining this as acoustic flow – this is inspired by the parallels with optic flow (Lee *et al.*, 1992). Time variation is imposed by changes in these cues as the bat flies past its targets. Muller and Schnitzler found with respect to ambiguity, accuracy and detectability that such a methodology was sufficient for obstacle avoidance and other similar behavioural tasks.

Walker *et al.* (1998) took a similar approach but also evaluated more simple mapping concepts, as well as the dynamic motion cues of acoustic flow (Muller and Schnitzler,

1999, 2000, 2001). Firstly, in evaluating the concept of simply creating different views (perspectives), Walker *et al.* found that the inherent symmetry (Figure 1.9) in the system was broken by simply repositioning the pinnae (Jenison, 1997, Wallach, 1940, Klensch, 1948). Secondly, Walker *et al.* argued that the creation of dynamic cues, for example frequency and amplitude modulations, allows the individual to derive useful temporal cues for localisation. Muller and Schnitzler (1999, 2000, 2001) more generally argued this, defining this theory as acoustic flow.

In evaluating the integration of motional information with auditory cues we must also consider other sources of motional information, principally proprioception. In investigating the effect of whole-body rotation (about the earth-vertical axis) for the lateralisation of dichotic sound, Lewald and Karnath (2001) found a slight but significant influence of rotation upon sound lateralisation. This suggested a vestibular contribution to accurately track a static sound source during natural head and body rotations. Expanding this to include the effect of gaze, Goossens and van Opstal (1999) found the auditory system relies upon eye-head position for auditory guided behaviours. To summarise, the proprioceptive senses (Lewald *et al.*, 2000, Lewald and Karnath, 2000, Pettorossi *et al.*, 2005) play a role in sound localisation, as does the perceived direction of gravity (DiZio *et al.*, 2001), the orientation of the head related to inertial forces (Prieur *et al.*, 2005), and gaze direction (Getzmann, 2002, Lewald and Ehrenstein, 1998). This grounds the sense of motion and its impact on audition by use of proprioception.

As an example of this grounding, Wallach (1940) showed head movements are registered correctly in the cases where:

- position, tension and posture receptors of the neck muscles, and the cervical vertebrae provide no information,
- only the vestibular organ provides information,
- and only the sense of vision provides information.

An extension of this by Wightman and Kistler (1999) showed that only when the listener controlled a sound source's movement did the ambiguity in the position of an auditory event disappear. This leads to the wider question of a listener's interaction with the environment, which is applicable to all senses rather than just audition.

This section has discussed a range of the motional interactions with perception, including the senses of proprioception. This underlies more specifically that auditory localisation

is far more than simply the deduction of knowledge (about sound source localisation) from observed cues, as the other senses play a significant role (e.g. Goossens and van Opstal, 1999). This would indicate that there are wider issues to be considered:

- is this fundamental to perception?
- Is it simply motion or controlled motion that plays an important role?

The first issue is discussed in the next section from a philosophical vantage point. The second is addressed in Section 1.4. I also present a new quantity for selecting the best motion in Chapter 3.

1.3 Philosophy of active perception

Leading on from the motional auditory localisation at the end of the previous Section, I review the background philosophy that has partly led to the development of sensorimotor contingencies (Noe, 2004). These theories were developed to deal with more logic based formal rule theories (classicist) of cognition. The principle problems with these classicist views relate to the rules that ground the symbols used to describe an individual's precepts of its environment to the environment itself. This is the symbol grounding problem described by Searle (1981), Harnad (1990), Searle (1980). Further, there are arguments for representations without rules; the expertise problem (Aizawa, 1994).

Gibson (1978) defined an ecological approach to perception that considers the importance of understanding perception in terms of the tasks an individual must perform. In this context the individual's environment is considered to offer "affordances". One interpretation of these affordances are as possibilities for action. Further, the individual is structured both by nature (genetic) and nurture (development) to register these affordances in its environment. As an aside it is important to note that Gibson was critical of explanations that sought too detailed a view of the sensory systems. Ecological perception seeks an explanation, not in terms of what responses of receptors are to stimuli, but rather an explanation of how the stimulus carries relevant information.

Enactive perception regards behaviour as the "structural coupling" of the individual to its environment (Varela *et al.*, 1995). Structured coupling is a notion grounded in evolutionary theory (Varela *et al.*, 1995), this approach to perception is:

- that individuals interact in a complex manner with their environment,

- that these interactions affect both the internal structure of the individual but also the environment,
- that these structures (individual and environmental) coevolve over time.

These points lead to the conclusions that it is not possible to understand either the individual or its environment in isolation but in context to their interaction. This led Varela *et al.* to contend, that to understand perception, we need to know how the nervous system links the individual's sensory and motor surfaces (Noe, 2004).

Embodied cognition emphasises the formation role the environment plays in the development of cognitive processes (Vernon, 2008, Cowart, 2006). Cognitive processes are posited to develop from real time, goal directed interactions between an individual and its environment. The nature of these interactions influence further development and formation of cognitive capacities. Perceptually, embodiment is understood as the way an individual's sensorimotor capacities enable successful interaction with its environment (niche). Quick *et al.* (2000) define embodiment more formally as,

A system X is embodied in an environment E if perturbatory channels exist between the two. That is, X is embodied in E if for every time t at which both X and E exist, some subset of E 's possible states have the capacity to perturb X 's state, and some subset of X 's possible states have the capacity to perturb E 's state.

This describes generally all individuals that perceive information about an environment, and concurrently make decisions based upon these precepts. This creates a circular dependency between the individual and its environment.

O'Regan and Noe (2001) hold that perception is a function of both the senses and our ability to interact with them. This they define as sensorimotor contingencies. This can further be seen as taking aspects of both Gibson's ecological approach (Scholl and Simmons, 2001) and the enactive interactions with the organism and its environment. This is not without its detractors. Prinz (2006) who argued forcefully the relationship between the senses and action is causal rather than constitutive. Prinz argues that action causes change in the senses, rather than being fundamental to the process of sensing. However, O'Regan and Noe's sensorimotor approach is advantageous as it provides a clearly principled argument to account for perception. The experience of *seeing* occurs when the observer masters the governing laws of the requisite visiomotor contingencies.¹

¹Though in their comments van Gulick (2001), Schlesinger (2001), Gallese and Keyser (2001) argue

In this section I have discussed the philosophical underpinnings of how action interplays with observation to provide sensation. It gives a general explanation of active perception in the form of sensorimotor contingencies arguing the importance of an individual's ability to interact with its environment and sense the consequences of this interaction. The next section leads with a review of the behavioural work related to embodiment and sensorimotor contingencies. I then finish the next section and the literature by reviewing practical methods for selecting actions to aid perception (inference).

1.4 Examples of active perception

This section leads on from the theory of Noe (2004), O'Regan and Noe (2001) of sensorimotor contingencies. This is placed in context with a review of active perception and its development in localisation problems. Such localisation tasks include objects within the environment, the individual's position and pose in the environment, and the relation to tool use (Subsection 1.4.1). This is followed in Subsection 1.4.2 by a review of how best an individual can generally select a localising action in a range of domains. As, while the philosophical points made earlier discuss the importance of action in perception, they say little as to what actions will generally result in the most certain and unambiguous localisations and inferences.

1.4.1 Sensorimotor contingencies

O'Regan and Noe (2001), Noe (2004) argue that the perceived structure of an individual's reality (notions of body, environment, space, object and attribute) is a consequence of the interaction between the inputs and outputs of its brain. These define the individual's sensorimotor contingencies. Philipona *et al.* (2003) validated this statement by using a simulated organism to deduce, through only its input and output, the dimensionality of the phase space that its interactions with the environment engendered. Using differential geometry to define a subspace through the intersection of the environment and possible actions, this subspace represented the possible sensory inputs, thus allowing Philipona *et al.* to infer the dimensionality of the environment. Philipona *et al.* (2004) show it is possible, given access to sensor inputs and motor outputs, for an individual to algorithmically infer information about itself (physiology) and its world. Relatedly, Nagel *et al.*

that preventing internalised representations is unnecessarily limiting and does not appear to be grounded in evidence.

(2005) showed the integration of a belt based compass, giving a tactile measure of magnetic North, provided a subject sensory enhancement.

Grzywacz and Balboa (2002) suggest a mechanism that compares expected results of tasks to select consistently better representations of the relationships between input and output. This is grounded in the hypothesis that much of what we consider *a priori* about the environment, is deducible without prior knowledge, simply of the brain governing the linkages between the organisms inputs (senses) and outputs (actions) (O'Regan and Noe, 2001). Similarly, the findings in Blakemore *et al.* (2002), Droulez and Berthoz (1991), Blakemore *et al.* (2001) indicate the cerebellum plays a role in comparing the prediction of a movement's sensory consequence, to the actual consequence. This process allows the individual to deduce changes in the environment.

King *et al.* (2000) argue that the auditory system requires experience to localise sound. Neurophysiology studies of animals raised with abnormal sensory inputs, indicate the mapping of the auditory space is shaped during development by both visual and auditory experience (Knudsen and Brainard, 1991, King and Parsons, 1999, Rauschecker and Kniepert, 1994). The usefulness of this adaptation is limited. If for example one ear no longer provides input then it excludes binaural cues. This hampers an individual's sound localisation acuity, though, it does not eliminate the ability to localise (King *et al.*, 1988, Knudsen, 1985).² Experience induced plasticity, allows the auditory pathways to be adapted to individual physiology such as the size and shape of an individual's head and ears (Middlebrooks, 1999a, Carlile and Pralong, 1994, Middlebrooks, 1999b, Wenzel *et al.*, 1993). Wilmington *et al.* (1994) illustrate that binaural ability following corrective surgery exists to varying degrees in these tasks. This suggests there are different effects to abnormal early experience on different aspects of binaural hearing. Using ideas inspired by O'Regan and Noe (2001), Aytakin *et al.* (2008) have shown that a naive agent can learn the auditory space around it from its HRTFs.

The visiomotor system can influence and guide the calibration of sound source localisation in an individual's development (Knudsen and Knudsen, 1985, 1989, Held and Hein, 1963). Similarly there are other sensory influences upon the auditory system from somatosenses (Kanold and Young, 2001), proprioception (Alexeenko and Verderevskaya, 1976) and eye position (Groh *et al.*, 2001, Jay and Sparks, 1984). I next discuss the adaptive properties of the nervous system in active perception.

²This can be related back to the previous sections that discuss the importance of the various auditory cues, Sections 1.2.1 & 1.2.2.

While investigating the integration of haptic and visual information after visual adaptation, Rossetti *et al.* (1993) found evidence to support there being two mechanisms at work in sensory adaptation:

1. a short-term process that is involved in normal coordination to spatially align eye and hand systems,
2. a longer-term process that involves remapping spatial misalignments.

The former short-term mechanism may be employed to quickly optimise accuracy in a situation of misalignment, however complete adaptation must await the slower acting latter mechanism (Rossetti *et al.*, 1993). This was shown from a prismatic modification to the spatial alignment of eye and hand through finger pointing to visual targets. During adaptation the average latencies, between the end of an eye saccade to the beginning of a hand movement, more than doubled when compared to before prismatic modification. However, after adaptation these latencies had returned to the previous length of time. Illustrating that the subject's sensorimotor contingencies had fully adapted to the new configurations.

Using the definition of sensorimotor adaptation as a perceptual adaptation, whose effects depend upon the occurrence and nature of the performed motor actions, Bompas and O'Regan (2006) show that a sensorimotor adaptation can also occur for colour perception. They achieved this by introducing a new connection (sensorimotor contingency) between eye movements and colour changes. They found that motor activity, in the form of an eye saccade, allowed the experimental subjects to notice an environmental (e.g. chromatic) change. Skaff *et al.* (2002), Clark and O'Regan (2000) had illustrated this effect in modelling colour consistency using the non-uniformity of retinal sampling (Roorda and Williams, 1999). This shows the breadth of sensorimotor influence and the adaptability of perception due in part to this sensorimotor coinfluence.³

Rather than substituting one sense for another (y Rita *et al.*, 1969, y Rita, 2004, Arno *et al.*, 2001, Meijer, 1992), Nagel *et al.* (2005) illustrated the integration in human subjects of a completely alien sensorimodality, the perception of magnetic North through a haptic belt placed around the waist. They found that this new modality interacted with the vestibular system and led to half the subjects reporting a change in their sensory experience.

³A robotic example of this Olsson *et al.* (2006) used mutual information of signals from each unique pair of pixels to map their relative topology. Similarly to Pierce and Kuipers (1997) the actuator space was sampled to then extract the average effect, which in this context was the average optic-flow (Smith, 1997). Brenner *et al.* (2000) indicate that the encoding in contrast and motion neurons in the fly adapt to their input similarly to the grouping of sensors in Olsson *et al.* (2005a,b, 2004, 2006).

This, Nagel *et al.* concluded, indicated that new sensorimotor contingencies had been learnt, and further, integrated into the preexisting vestibular sensorimotor contingencies. Thus emphasising the utility of action-perception in learning, showing that not only small refinements, but also entire novel sensory modalities can be learnt.

So far I have reviewed the concept of sensorimotor contingencies as they relate to the influence of the senses on perception. However, sensorimotor contingencies also indicate adaption of the brain’s motive pathways as well as the sensory pathways. Hochberg *et al.* (2006) showed that a tetraplegic human subject, approximately 3 years after a major spinal injury, could after training operate a computer cursor utilising a neuromotor prosthesis implanted in his primary motor cortex. Further, through the use of this prosthesis, the subject could also perform rudimentary actions with a robotic arm and prosthetic hand. Alternative approaches have used non-invasive EEG sensors (Wolpaw *et al.*, 2000, Müller-Putza *et al.*, 2005, Mason *et al.*, 2004). However, typically these require constant concentration on the part of the user (Wolpaw and McFarland, 2004) unlike Hochberg *et al.*’s invasive method.

Relatedly, Bongard *et al.* (2006) showed that a robotic “star-fish” could accurately model its physiology through investigating its sensorimotor laws. The robot would experiment, continually testing a set of hypotheses of its physiology and continually refine these by performing actions calculated to maximally distinguish between these alternate hypotheses (Bongard and Lipson, 2005). Hence, the robot constantly tested its understanding of reality according to its environment.

A concept of self, where tools are viewed as extensions of an individual’s body, has some backing from neural studies. Umiltà *et al.* (2008) showed that the same neurons in the motor cortex that are activated (hence associated) with hand grasping, are active when grasping with a tool (a pair of pliers). This effect is not simply associated with the grasping motion, as these neurons are active even when the tool in question is a pair of “reverse pliers”.⁴ Neurons in the premotor cortex in monkeys may encode geometric positions in space relative to the animal (Graziano *et al.*, 1997). A subset of these neurons also appeared to encode the expectation of an object’s presence if it was occluded by darkness.

The review of the sensorimotor and related literature in this subsection shows the fluidity and flexibility of concepts of self with regard to tools, and the flexibility of the brain

⁴An implement that uses an opening of the fist to close the plier’s jaws.

in relating cause and effect for representation of the world. This is with the express consequence of extracting behaviourally useful information about the world from the senses. While this review shows the degree of coinfluence between the senses and action (sensorimotor), little is said behaviourally about how actions are typically selected. An exception in this subsection is Bongard *et al.* (2006). Bongard *et al.* had a robot model act, then contrast its hypotheses of the world with knowledge it extracted from its sensors. This leads to the question of how best to select an action for sensorimotor purposes. The obvious answer is to select one that gives the most knowledge for the least effort. In the next two subsections I discuss the state-of-the-art approaches to this action selection problem – the sensor placement problem.

1.4.2 Robotic mapping, SLAM/CML

In considering algorithmically the best localising actions, a large part of the literature pertains to robotic mapping. Robotic mapping is the task of a robot localising itself upon a map of possible positions. Thrun (2003) argues that robotic mapping can be categorised into two rough types of mapping: metric mapping, and topological mapping. Metric mapping tends to take the form of occupancy grids (Elfes, 1989, 1987). This is where the position is defined according to a grid which models the occupied and free space of the environment (Borenstein and Koren, 1991, Burgard *et al.*, 1999a,b, Yamauchi and Langley, 1997). Alternatives include mapping using sets of polyhedra to describe the environment (Chatila and Laumond, 1985). These deal with directly modelling the environment. Topological mapping typically represents the environment by using a list of significant places connected by arcs (Choset, 1996, Kuipers and Byun, 1981, Shatkay, 1999, Zimmer, 1996), though these rely upon knowledge of how to navigate between these places.

Thrun (2003) argues that mapping can be broken into two further categories:

1. a world centric map where the robot models explicitly its sensors to relate its map of the environment to the sensor’s recordings,
2. a robot centric map where the maps are described in the sensor’s measurement space, for example the raw storage of measurements for each location for comparison.

Smith and Cheeseman (1986) introduced a probabilistic framework for robotic mapping and localisation upon a growing map. The two terms for this are *simultaneous localisation and*

mapping (SLAM) (Durrant-Whyte *et al.*, 2003) and *concurrent mapping and localisation* (CML) (Thrun *et al.*, 2004). Typically SLAM and CML have implicitly assumed that ambiguities in belief will automatically disappear as a robot explores its environment. However, this is not always the case (Thrun, 2003). This eventually led to the development of more formal and explicit algorithmic methods for selecting actions. These algorithmic methods have the purpose of removing ambiguity through approaches such as Bayesian state estimation.

1.4.3 Best localisation actions, Bayesian state estimation

Exploration is typically greedy (Choset, 1996, Burgard *et al.*, 2000, Simmons *et al.*, 2000, Yamauchi and Beer, 1996), though it is also subject to safety constraints (Gonzalez-Banos and Latombe, 2002). Gonzalez-Banos and Latombe (2002) define a generalised framework that compliments SLAM and CML by using a next-best-view (NBV) algorithm to guide navigation. This NBV selects a robot motion, from among a group of *safe* candidate motions, according to the information gain of each candidate motion. The gain in information is defined according to an estimate of the area, or volume, of potentially visible unexplored space that a motion will bring into view (Briggs and Donald, 1994). This relates to the question of selecting an action to convey the most information about the robot’s environment. Gonzalez-Banos and Latombe defined this information in terms of the unknown geometric area that can potentially be brought into the robot’s view. This causes their implementation to be limited by the implementor’s prior assumptions, which define how the robot represents the geometry of its environment.

Likewise, Porta *et al.* (2005, 2003), Fox *et al.* (1998) also fall partly within this category of stochastic forward models known as Bayesian state estimation. Using robotic mapping of appearance with a stereo camera on a pan-tilt mounting attached to a robot frame, Porta *et al.* approximated the expected entropy of the *a posteriori* belief to select the NBV for robot self localisation.⁵ The lower this entropy, the more informative the corresponding movement can be expected to be. Such entropy minimisation techniques are referred to as Infomax – information maximisation to solve the sensor placement problem.

Porta *et al.*’s approach to Infomax uses particle filtering (importance sampling) where as Fox *et al.* (1998) used a full discretisation (occupancy grid) of the robot’s latent state space. Sujan and Dubowsky (2002) by contrast use such a measure for field-of-view selection in

⁵The selection of a best motor command according to which of a set of candidate motions minimise the entropy over all possible states in the latent state space given all possible observations.

integrating the state space of teams of mobile robots with mounted cameras. However, a limitation in the Infomax implementations of Porta *et al.* (2003, 2005) and Fox *et al.* (1998) is that the measurement space is approximated:

- Porta *et al.* construct a sample of the joint latent and measurement spaces to approximate the conditional Shannon information,
- Fox *et al.* similarly discretise the measurement space to approximate the conditional Shannon information.

I address this in Chapter 3, where I consider a new measure of the expected informativeness of an action that has useful properties.

In contrast to robotic localisation, Vergassola *et al.* (2007a) recreated moth like surges (forward motion into the wind) and casts (exploratory motion typically going cross wind) in simulated odour plumes (Murlis *et al.*, 1992, Mafra-Neto and Carde, 1994) using the reduction of entropy in the tracked posterior belief to select the best action (Fox *et al.*, 1998). Odour plumes are not just used by moths but also mosquitoes (Geier *et al.*, 1999) and water crabs which use them for hunting (Zimmer-Faust *et al.*, 1995, Weissburg and Zimmer-Faust, 1994). These species also exhibit casting and zigzagging upwind (up-flow) behaviour. While Vergassola *et al.* did not fully recreate the details of moth behaviour, their model did capture the coarse cast and surge behaviours exhibited across many species that use olfaction to search in turbulent environments.⁶

Similarly, forward models construct predictions of the consequence of a choice of action by simulating the consequence according to the robot’s sensorimotor model and belief in the model’s current state (Dearden and Demiris, 2005, Thrun, 2003) (*cf* Bongard *et al.*, 2006). Forward models however can be seen as describing all processes that use a predictive simulation stage to make a decision offering a generalisation of NBV by being applicable to almost any posed task. As such utilising Bayesian prediction, as I do in Chapters 3 & 4, falls within this category as stochastic forward models.

The work of Fox *et al.* (1998), Porta *et al.* (2003, 2005) and Vergassola *et al.* (2007a) shares a commonality of using statistical entropy to judge between choices of action. Although Fox *et al.* and Porta *et al.* are both robotic examples, in each the authors use Bayesian state estimation for the robot to locate itself upon a map.

In the field of radar an alternative measure of information to select the best action

⁶In a turbulent flow a body of gas is broken into intermittent patches of odour, *odour packets* (Balkovsky and Shraiman, 2002).

or manoeuvre is achieved using the *a posteriori* beliefs Fisher information matrix (FIM) (Passerieux and van Cappel, 1998, Cadre and Gauvrit, 1996, Logothetis *et al.*, 1998). In most problems the FIM is difficult to compute, so its complement, the *a posteriori* Cramer-Rao lower bound (CRLB) is often used instead (Helferty and Mudgett, 1993). Helferty and Mudgett (1993) compared both performance criterion, for computing optimal observer paths, for the problem of target localisation and tracking. They found the CRLB gave greater utility, as a weighted sum of its trace could be applied, allowing the discounting of some latent dimensions with respect to others. This allows a trade off between certainty in the target’s position against certainty in the corresponding target’s velocity. The CRLB is a covariance derived measure which requires an agent to make constraining assumptions in its representation of the world. For instance, in radar the latent state space is geometrically equivalent to the true latent state space.

The review of best localising action selection in this subsection illustrates the domains to which informatically driven action selection has been applied. Two principle quantities used to solve the sensor placement problem algorithmically, are the Shannon information (entropy) and the CRLB. In Chapter 3, I present a viable alternative to the Shannon information. Then in Chapter 4, I contrast this alternative, the Shannon information and the CRLB to inferring latent state in a partially observable sensor placement learning problem. To summarise, I have introduced two different state-of-the-art approaches to the sensor placement problem.

1.5 Thesis structure

In the light of this literature review, one of the aims of my thesis is to address the question of the best action selection in circumstances where the situation is uncertain, this is achieved in Chapter 3. Next, the question of how this impacts learning a viable model is addressed in Chapter 4. These Chapters evaluate the concept of how to select the best localising action and how this relates to learning.

While my work for this thesis initially started from the perspective of sound source localisation depicted in,

Chapter 2: Bayesian passive sound source localisation discusses a passive localiser and illustrates the difficulties that require prior knowledge for a viable model of the localisation process. I contrast my passive sound source localisation model with an application of the approach described in Nix and Hohmann (2006).

However, I found that the concepts I was developing were better expressed in a more generalised fashion. As the investigation indicated, there are difficulties in understanding perception of location where the agent does not interact with its environment. This is the treatment of perception as a purely passive problem. The implication of this leads to concepts which include acoustic flow, where it is not just instantaneous cues that lead to localisation but also how these cues change (Jenison, 1997, Muller and Schnitzler, 1999, 2000, 2001). More generally it is not only the change of these variables that lead to accurate localisation, but the individual's knowledge of what caused this change (Pettorossi *et al.*, 2005, Wallach, 1940, Lewald and Ehrenstein, 1998, Lewald and Karnath, 2001). Working from this I developed a new quantity, as an alternative to Infomax, for selecting the best localisation action in a general fashion.

Chapter 3: An information measure for optimal action selection develops and discusses a fully general framework to the problem of selecting the optimal localising action in active-perception. This illustrates a solution to the problem of selecting an action as an alternative to current approaches: Infomax and *a posteriori* CRLB.

Finally, I applied my framework to a more conceptual problem to illustrate the generality of my measure.

Chapter 4: Learning a model for an active agent applies the optimal action framework developed in Chapter 3 to a dynamic domain and learns a model concurrently with action and observing. This addresses the problem of active experimentation to indicate iteratively, the selected action is better than totally random action selection (body-babbling). I also contrast these methods with two other informatic action-selection approaches.

These three Chapters report the experimental work I conducted in the pursuit of my research.

I have approached dual aims in this thesis. My first aim was that of understanding from a theoretical point of view how the best localisation actions can be selected, with a focus upon the removal of inference uncertainty. While reviewing the literature I found a substantial body of work on the concept of action conveying useful localisation information (Noe, 2004, O'Regan and Noe, 2001, Pettorossi *et al.*, 2005, Lee *et al.*, 1992, Ashmead *et al.*, 1995, Jenison, 1997), reviewed in Subsections 1.2.3 & 1.4.1 and Section 1.3. In addition to this, the literature showed that active perception is important in an individual's

perceptual development (Held and Hein, 1963). However, these left open the topic of how to best select actions in the light of predicted beliefs about the future. How to select the most desirable outcome by applying Bayesian state prediction to predict the differing consequences of each action – one example for the best outcome, is that which has the least *a posteriori* uncertainty. I consider the problem of active localisation, which has a rich literature on selecting the expected best action, this is the sensor placement problem. Examples have addressed particular domains such as chemotaxis in turbulent flows (Vergassola *et al.*, 2007a) and robotic mapping (Fox *et al.*, 1998, Porta *et al.*, 2003) for Bayesian state estimation, reviewed in Subsection 1.4.3. The principle gaps in the literature for this problem domain are in terms of the scaling and the necessary approximating assumptions of the solutions. The necessity that the CRLB and Infomax quantities require the *a posteriori* belief to be approximated is the source of these issues. In Chapter 3, I present and contrast an alternative to Infomax. Then, in Chapter 4, I adapt and compare my alternative, Infomax and CRLB to a sensor placement learning problem – I shall relate this to some of the higher order concepts from the literature of sensorimotor contingencies.

My second aim was to consider how this ties into the more general problem of perception. This would make my work of interest to researchers in fields related to perception, including robotics and psychophysics, with specific possibilities including:

1. tracking a target in a cluttered environment (Moss and Surlykke, 2001), an emergent consequence could be in maintaining visibility (LaValle *et al.*, 1997),
2. pack hunting, where Spletzer and Taylor (2003) shared inference between agents. Other interesting questions could consider lossy channels of communication similar to vocalisations,
3. the tracking of stealthy targets, which is somewhat related to the problem of clutter (Ristic *et al.*, 2004b) when looking at the perspective of camouflage.

These aims were developed in the light of my exploratory foray into passive sound localisation.

Finally, Table 1.2 summarises the literature review in this chapter. In light of Table 1.2, my thesis fills a number of important limitations in the state-of-the-art. In Chapter 2 I present a Bayesian solution to the problem of passive sound source localisation. I compare and contrast my approach, in Chapter 2, with that of Nix and Hohmann (2006). Nix and Hohmann, like other state-of-the-art research into auditory perception, consider the

Field	Current state-of-the-art	Missing elements and limitations with the state-of-the-art
Radar and sonar	<i>Theoretical models (applied sometimes to radar data) & sensor network models</i> – typically use an <i>a posteriori</i> CRLB, to optimise the FIM, to select an optimal manoeuvre (or sensor placement). The state-of-the-art currently uses particle filters for tracking the latent state space. Latent state space prediction is used to estimate the <i>a posteriori</i> CRLB up to the modeller’s desired future time step.	A gap exists here for efficiently selecting actions where the state spaces are mixed, both continuous and discrete. The CRLB and Fisher information only apply to continuous latent state spaces.
Robotics which applies POMDP to SLAM / CML	<i>Robotic</i> – these typically use a discretisation of a latent state space for mapping. Current work is focused upon applying SLAM/CML to dynamic environments. The state-of-the-art currently uses particle filters to represent the POMDPs that define the model.	A gap exists here for a more efficient informatic action selection process than that of the Shannon information. Shannon information is best applied to discretised latent state spaces, e.g. occupancy grids or particle filtering.
Bayesian state estimation	<i>Theoretical models</i> – application of Infomax, <i>cf</i> robotics and radar and sonar to chemotaxis and other inference problem domains.	Infomax is a costly operation to compute, it also needs approximations of the entropy/Shannon information integral. There is a gap for an approach which is analytic under the latent predictions estimations.
Sensory-motor contingencies	<i>Theoretical</i> – using differential geometry and arbitrarily small actions to learn the relationships between the consequences of an action to the individual’s observables.	A gap exists here for applying an informatic policy to select the best action for learning the relationships between actions and consequences in a Bayesian fashion.
Neurophysiology and behavioural sound source localisation	<i>Neurophysiology & behavioural</i> – The state-of-the-art considers time disparity and level disparity. <i>Modelling (theoretical models)</i> – The state-of-the-art for passive sound source localisation is in Nix and Hohmann (2006). To localise a sound source, they use histograms to represent the level and time disparity distributions.	One gap for my work in Chapter 2 is the use of a pair of sound signals, explicitly retaining the two signals loudness rather than discarding these when computing the level disparity. A further gap is to define a parametric model to represent sound source localisation, as Nix and Hohmann (2006) used histogram distributions to capture the data’s variation.

Table 1.2: A summary of the state-of-the-art.

measured quantities as interaural level and time disparities (ILD and ITD). This however has the consequence of neglecting the absolute loudness of a sound. My approach in Chapter 2 addresses this limitation through the use of a graphical model. In Chapter 3 I present a new quantity that can be applied to solve the sensor placement problem. In this chapter I will describe the advantageous aspects of my approach. I will then compare and contrast my new quantity with Infomax – the current state-of-the-art in this domain. I compare my approach to Infomax in two ways:

1. I compare the algorithmic scaling of my approach to the Infomax algorithm of Porta *et al.* (2003, 2005),
2. and I also apply my approach to the domain of chemotaxis, through the problem described in Vergassola *et al.* (2007a,b).

In Chapter 4 I apply my measure, developed in Chapter 3, to a sensor placement learning problem. I do this to compare the performance of the different state-of-the-art informatic policies:

1. the application of Porta *et al.*'s Infomax,
2. and the application of the *a posteriori* CRLB (Hernandez, 2004, Martinez-Cantin *et al.*, 2007).

This summarises my contributions in this thesis.

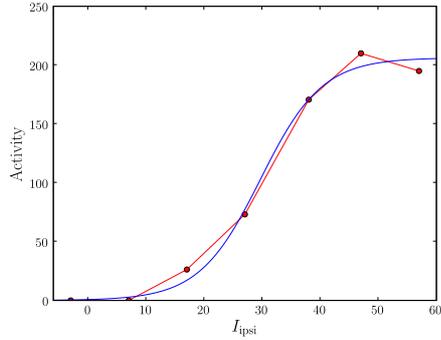
Chapter 2

Bayesian approaches to passive sound source localisation

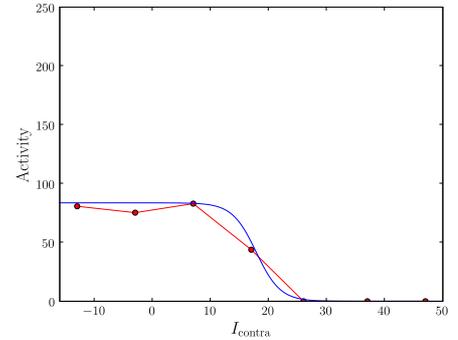
2.1 Introduction

At its simplest the problem of sound source localisation consists of locating a single static sound source, the target, in an environment that does not contribute to the sound field. When the agent is inactive and unmoving (passive observation) the principle sound cues are based upon the loudness and timing of features within the sound signals perceived at either ear (Blauert, 1997, Hartmann, 1999). These cues are further categorised as being perceived either between the ears (binaural) or by each ear individually (monaural) in nature (Hartmann, 1999). There is a co-influence between these binaural and monaural level and timing cues. It is the relationship between these and the source or targets location which is the subject of investigation in this Chapter.

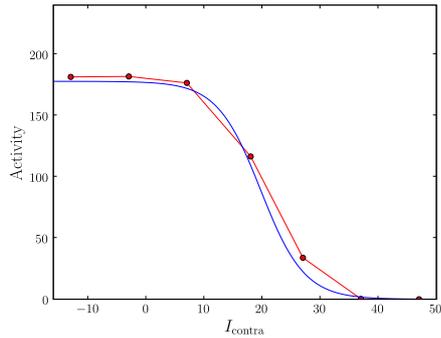
The monaural localisation cues are related to the attenuation of a sound source's loudness with the distance from the agent to the target. Physiology interacts with the loudness to further attenuate the perceived loudness based upon direction of the target relative to the individual (Kistler and Wightman, 1992). The methodology of applying such directional attenuation is usually through head related transfer functions (HRTFs). Oldfield and Parker (1984) showed that this directional gain or attenuation through the HRTFs can provide a correction to specific kinds of localisation error, for instance front-back confusions. The monaural cues are usually neglected in sound source localisation. Examples of this approach are Nix and Hohmann (2006), Zurek (1991). My methodology in this Chapter considers monaural cues as complimentary to binaural cues and forms part of my



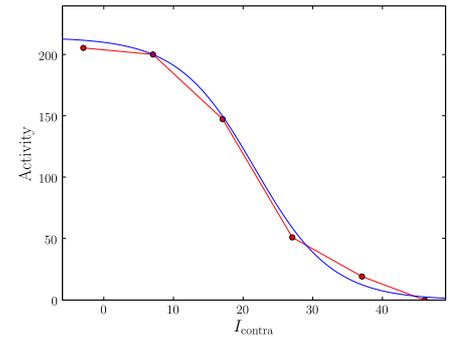
(a) The base activity where I_{contra} is 0 dB



(b) The activity for I_{ipsi} being 27 dB



(c) The activity for I_{ipsi} being 37 dB



(d) The activity for I_{ipsi} being 47 dB

Figure 2.1: The response of a lateral superior olive (LSO) neuron to the loudness at either ear. These plots depict the fitted functional activity of a single typical LSO unit. The neural response for a contralateral sound intensity of $I_{\text{contra}} = 0$ dB over the ipsilateral sound intensity, I_{ipsi} , is depicted in plot (a), the activities for three different I_{ipsi} , 27 dB, 37 dB and 47 dB respectively plots (b), (c) and (d). The neural responses are depicted with red dots. This shows a response curve that is more complicated than simple computation of level disparity ($I_{\text{ipsi}} - I_{\text{contra}}$). I have fitted the product of two sigmoid functions to the data presented in Park *et al.* (1997), this is marked with the blue line. This is a function of 6 parameters that can represent such LSO units. It is important to note that such a fitted functionality does not take into account the effect of physiology or timing effects – therefore neglecting the degree of correlation between either ears received sound signal.

argument.

The binaural cues are principally considered as either interaural level disparities (ILD) or interaural time disparities (ITD) (Hartmann, 1999, Forsythe, 2002). ITDs represent the difference in time of arrival for a recognisable feature in either ear. There are two disparate mechanisms proposed for the brain to compute ITDs. The first is developed from the Jeffress model of neural delays, using a bank of neurons to represent the different ITDs (Pena and Konishi, 2002). The second is a phasic comparison between the signals and is similar to comparison in the frequency domain (McAlpine and Grothe, 2003, Joris *et al.*, 1998). Phasic comparisons have an advantage in terms of speed even on computers (Frey and Jojic, 2001, 2003).

ILDs represent the difference in loudness (the logarithmic sound intensity) between the ears. The mechanism of computation in the brain is less controversial than that of ITDs. The accepted model consists of comparing an excitatory projection from one ear to an inhibitory projection from the other ear in a region of the brain called the lateral superior olive (LSO) (Forsythe, 2002, Wang and Brown, 2006). However, Park *et al.* (1997) have shown that ILD tuning in LSO neurons have a nonlinear dependence upon the absolute sound level from either ear (Figure 2.1). This gives the absolute sound level at either ear a role in computing the auditory systems ILD cues for localisation purposes.

The position of a sound event relative to the individual will determine their observed ITD, ILD and the monaural loudness. Shinn-Cunningham *et al.* (2000) discussed the localisation uncertainty of particular ITDs & ILDs and the dependence of these confusions with the target position. They defined this to be a “torus of confusion” as, in three dimensions, the uncertainty forms a toroid centred upon the individual. Further, Shinn-Cunningham *et al.* took a simple model of sound propagating in a free field and applied a corrective directional attenuation due to the HRTF. In their investigation, Shinn-Cunningham *et al.* investigated the binaural localisation cues alone, which ignored the monaural localisation cues. As a result their approach ignored the information about location conveyed by the absolute sound level. Most work on sound source localisation similarly considers only ILD and ITD, notably Zurek (1991), Nix and Hohmann (2006), Walker *et al.* (1998), Suga and Ma (2003a).

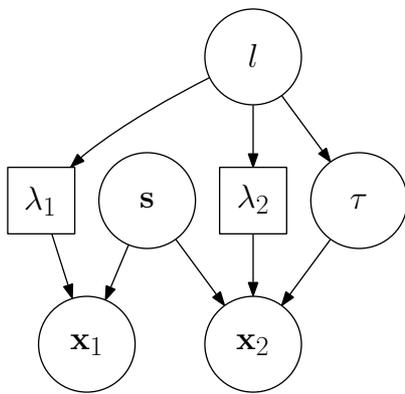
The co-dependence of the monaural level, ILD and ITD cues can lead to a fusion problem when conducting localisation inference. This presents the problem of how best to combine the information conveyed by a measurement of each of the cues about the targets

location to give a best estimate of location. Solutions to this problem have typically been inspired by biomimetics (Handzel and Krishnaprasad, 2002) which naively combine the cues by assuming conditional independence with the target’s location (naive Bayes). ITD is most easily computed by cross correlation based models and as such typically dominate ITD models (Wang and Brown, 2006). Alternatively, ITD can be computed in the frequency domain (Nix and Hohmann, 2006, Duda, 1997). In mammalian neurophysiology studies, level disparities are associated with excitatory-inhibitory (EI) neurons which exhibit a sigmoidal tuning curve with interaural level disparity (Forsythe, 2002). An example of this can be seen in Figure 2.1. Nix and Hohmann (2006) and Zurek (1991) respectively showed experimentally and analytically that for a given sound source location, both the distribution of measured levels and time disparities are both non-Gaussian. These distributions are also dependent upon the signal-to-noise ratio. Typical techniques for integrating these cues filter the inputs and mix the results in a biologically inspired manner (Wang and Brown, 2006). This, when naively computed, is equivalent to naive Bayes which is the approach taken by Nix and Hohmann (2006). I use the graphical model in Figure 2.2 to define in the next section a more complete model of sound source localisation. I contrast and compare this model with an implementation of Nix and Hohmann (2006) which represents the-state-of-the-art in the analysis of passive sound source localisation.

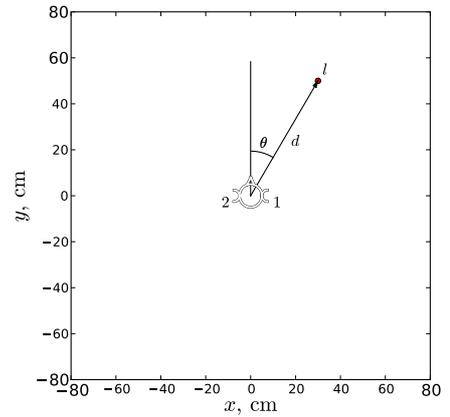
The structure of my model extends the sound perception component of the audio-visual inference model of Beal *et al.* (2003), Attias *et al.* (2001). The graphical model¹ in Figure 2.2 (a) depicts the dependence of location l upon the detected signals, \mathbf{x}_1 and \mathbf{x}_2 , through a latent variable time disparity variable τ and the attenuation parameters $\lambda_1(l)$ & $\lambda_2(l)$. This defines a model that considers the interactions between those quantities that are not directly observable (τ , λ_1 , λ_2 , l and \mathbf{s}) and those that are (\mathbf{x}_1 and \mathbf{x}_2).

To evaluate the influence of portions of the model that I define in the next section, I use a variation of local weight analysis (Cook, 1986, Zhu and Lee, 2001). Cook’s weight analysis considers the curvature of the likelihood of the model to small perturbations of its parameters or data points of the model. For instance, this allows consideration of the importance individual data points have for an inferred model (Zhu and Lee, 2001). I consider the curvature of the KL-divergence between the perturbed and unperturbed joint distributions. The purpose is to give a measure of the importance that a particular latent state variable or parameter has in inference tasks for each target location.

¹A graphical model represents visually the causal dependencies of quantities that a model uses to characterise a system. The nodes represent the quantities and the edges the causal dependencies.



(a)



(b)

Figure 2.2: Graphical model indicating the dependencies between measurables, hidden variables and location. The location dependent latent parameters are the attenuations λ_1 and λ_2 . The latent variables are, the relative time delay τ , the source signal \mathbf{s} , and the targets location l . The observable variables are the sound measured at either ear \mathbf{x}_1 and \mathbf{x}_2 . Cartoon (b) portrays the spatial relationship between agent at the origin and the location of a sound event (the target), at a bearing θ and distance d annotated by $l = \{d, \theta\}$. The structure of my model (a) extends in two ways the sound perception component of the audio-visual model in Beal *et al.* (2003), Attias *et al.* (2001), Hospedales and Vijayakumar (2006). The first is by accounting for the dependence of the attenuations, λ_1 and λ_2 , upon the sound sources location l . The second is due to Nix and Hohmann (2006) who found that the time disparity of the variance changes with sound source location. I make both the mean and variance of the time disparities likelihood dependent upon the sound sources location l . This is defined in Subsection 2.2.2.

However, it is not possible to analyse Nix and Hohmann’s model using a curvature analysis as their approach is non-parametric. The curvature analysis applies a perturbation to parameters and so is only applicable to parametric models. This illustrates a difficulty in understanding the processes being modelled in non-parametric approaches. To conduct a comparison with Nix and Hohmann’s approach I compare and contrast inference performance.

2.2 Localisation model

My sound localisation model, defined in Figure 2.2, takes as its starting point the seminal model of Beal *et al.* (2003) which was further developed by Hospedales and Vijayakumar (2006). These models are not formally for sound localisation, but targeted at real time audio-visual tracking. However, the audio component does provide a useful starting point for my development of a more complete sound localisation model. In Subsection 2.2.1 I restate the audio process from Beal *et al.*. Following on in Subsection 2.2.2, I extend this by relating the signal processes to level disparities in order to justify considering the attenuations as sufficient statistics for the distribution of ILDs. The structure of my model (Figure 2.2 (a)) extends in two ways the sound perception component of the audio-visual model in Beal *et al.* (2003), Attias *et al.* (2001), Hospedales and Vijayakumar (2006). The first is by accounting for the dependence of the attenuations, λ_1 and λ_2 , upon the sound source’s location l . The second is due to Nix and Hohmann (2006), who found that the time disparity of the variance changes with the sound source’s location l . Hence, I make both the mean and variance of the time disparities likelihood depend on l . The important quantities, the distributions and each distributions parameters are tabulated respectively in Tables 2.1, 2.2 & 2.3.

2.2.1 Sound process

In this subsection I discuss the construction of the data likelihood for the sound component from the audio-visual model of Beal *et al.* (2003), Attias *et al.* (2001), Hospedales *et al.* (2007). A binaural sound process can be constructed from a source signal (\mathbf{s}), which is measured at either sensor (\mathbf{x}_1 and \mathbf{x}_2). The measured signals are attenuated respectively by the factors λ_1 and λ_2 . Finally the signals each receive a time delay that can be considered instead as a relative delay, or, time disparity τ .

Firstly I assume that there is a source signal \mathbf{s} consisting of D_x samples, and is char-

quantity	description
\mathbf{x}_1	the measured sound signal vector at receiver 1 of D_x dimensions,
\mathbf{x}_2	the measured sound signal vector at receiver 2 of D_x dimensions,
\mathbf{s}	the source sound signal, which is a vector of D_x dimensions,
τ	the relative time delay (or time disparity) between the sound signals \mathbf{x}_1 and \mathbf{x}_2 ,
$\hat{\mathbf{D}}_\tau$	is an operator that applies a delay of τ ,
$\lambda_1(l)$ and $\lambda_2(l)$	the attenuation of the sources sound signal \mathbf{s} for the measured sound signals \mathbf{x}_1 and \mathbf{x}_2 at the location l ,
l	the sound sources location (the latent state),

Table 2.1: Table of important quantities for my sound source localisation model depicted in Figure 2.2. The variables \mathbf{x}_1 and \mathbf{x}_2 are observable. The variables \mathbf{s} , τ and l are hidden. The quantities λ_1 and λ_2 are parameters.

distribution	description
$p(\mathbf{s})$	the prior belief in the source signal \mathbf{s} ,
$p(\mathbf{x}_1 \mathbf{s}, \lambda_1(l))$	the likelihood of the measured signal \mathbf{x}_1 given the source signal \mathbf{s} ,
$p(\mathbf{x}_2 \mathbf{s}, \lambda_2(l), \tau)$	the likelihood of the measured signal \mathbf{x}_2 given the source signal \mathbf{s} ,
$p(\tau l)$	the likelihood of the time disparity τ given the source signals location l ,

Table 2.2: Table of distributions for my models variables listed in Table 2.1 and Figure 2.2.

distribution	parameter	description
$p(\mathbf{x}_1 \mathbf{s}, \lambda_1(l))$	ν_1	the variance of the noise process for the 1st receivers measured sound signal,
	\mathbf{w}_1	the GTM vector of weights used to construct $\lambda_1(l) = \mathbf{w}_1^T \boldsymbol{\psi}(l)$, $\boldsymbol{\psi}(l)$ is a vector of location dependent basis functions,
$p(\mathbf{x}_2 \mathbf{s}, \lambda_2(l), \tau)$	ν_2	the variance of the noise process for the 2nd receivers measured sound signal,
	\mathbf{w}_2	the GTM vector of weights used to construct $\lambda_2(l) = \mathbf{w}_2^T \boldsymbol{\psi}(l)$, $\boldsymbol{\psi}(l)$ is a vector of location dependent basis functions,
$p(\mathbf{s})$	η	the variance of the sound signal process,
$p(\tau l)$	$\omega_\tau(l)$	the variance of the time disparities belief for a given location l ,
	\mathbf{w}_ω	the GTM vector of weights used to construct $\omega_\tau(l) = \mathbf{w}_\omega^T \boldsymbol{\psi}(l)$, $\boldsymbol{\psi}(l)$ is a vector of location dependent basis functions,
	$\gamma_\tau(l)$	the expected time disparity for a given location l ,
	\mathbf{w}_τ	the GTM vector of weights used to construct $\gamma_\tau(l) = \mathbf{w}_\tau^T \boldsymbol{\psi}(l)$, $\boldsymbol{\psi}(l)$ is a vector of location dependent basis functions,

Table 2.3: Table of parameters for the PDFs in Table 2.2.

acterised as a Gaussian process²,

$$p(\mathbf{s}) = \phi(\mathbf{s}; \mathbf{0}, \eta). \quad (2.1)$$

While such an assumption is not generally accurate, in this case it is, because the source signal is Gaussian. This signal \mathbf{s} is measured simultaneously at two locations by two identical sensors, labelled 1 and 2. These sensors are assumed to make measurements according to the processes,

$$\mathbf{x}_1 = \lambda_1 \mathbf{s} + \sqrt{\nu_1} \mathbf{n}_{0,1}, \quad (2.2)$$

$$\mathbf{x}_2 = \lambda_2 \hat{\mathbf{D}}_\tau \mathbf{s} + \sqrt{\nu_2} \mathbf{n}_{0,1}, \quad (2.3)$$

where the source signals are attenuated by the respective factors λ_1 & λ_2 , and a relative delay τ by the operator $\hat{\mathbf{D}}_\tau$ Beal *et al.* (2003), Hospedales *et al.* (2007), Hospedales and Vijayakumar (2006). In practice both sensors receive a delayed signal, however only the relative delay is meaningful. Finally there is interference from zero mean additive noise of variance ν_i for the respective sensors. The likelihood of the measured signals \mathbf{x}_1 & \mathbf{x}_2 , given the source signal \mathbf{s} , are both Gaussian processes and characterised by the distributions

$$p(\mathbf{x}_1 | \mathbf{s}, \lambda_1) = \phi(\mathbf{x}_1; \lambda_1 \mathbf{s}, \nu_1), \quad (2.4)$$

$$p(\mathbf{x}_2 | \mathbf{s}, \lambda_2, \tau) = \phi(\mathbf{x}_2; \lambda_2 \hat{\mathbf{D}}_\tau \mathbf{s}, \nu_2). \quad (2.5)$$

Only the latent variables are stated in the probability density functions, with the parameters being neglected for brevity. After marginalising out the signal $\bar{\mathbf{s}}$ the resulting joint likelihood of the data \mathbf{x}_1 & \mathbf{x}_2 conditional upon λ_1 , λ_2 and τ is,

$$p(\mathbf{x}_1, \mathbf{x}_2 | \tau, \lambda_1, \lambda_2) = \sqrt{\frac{\nu_s}{2\pi\nu_1\nu_2\eta}}^{D_x} \exp \left\{ -\frac{1}{2\nu_1^2} (\nu_1 - \lambda_1^2 \nu_s) \mathbf{x}_1^T \mathbf{x}_1 - \frac{1}{2\nu_2^2} (\nu_2 - \lambda_2^2 \nu_s) \mathbf{x}_2^T \mathbf{x}_2 + \frac{\nu_s \lambda_1 \lambda_2}{\nu_1 \nu_2} \mathbf{x}_2^T \hat{\mathbf{D}}_\tau \mathbf{x}_1 \right\}, \quad (2.6)$$

where $\frac{1}{\nu_s} = \frac{1}{\eta} + \frac{\lambda_1^2}{\nu_1} + \frac{\lambda_2^2}{\nu_2}$. This distribution has two principle portions, the purely loudness

²Taking a Gaussian distributed variable \mathbf{x} of D_x dimensions to have the form

$$\phi(\mathbf{x}; \boldsymbol{\mu}, \sigma) = \sqrt{\frac{1}{2\pi\sigma}}^{D_x} \exp \left\{ -\frac{1}{2\sigma} (\mathbf{x} - \boldsymbol{\mu})^T (\mathbf{x} - \boldsymbol{\mu}) \right\},$$

where $\boldsymbol{\mu}$ represents the mean and σ represents the variance.

term,

$$\exp \left\{ -\frac{1}{2\nu_1^2} (\nu_1 - \lambda_1^2 \nu_s) \mathbf{x}_1^T \mathbf{x}_1 - \frac{1}{2\nu_2^2} (\nu_2 - \lambda_2^2 \nu_s) \mathbf{x}_2^T \mathbf{x}_2 \right\}, \quad (2.7)$$

and the time-disparity term,

$$\exp \left\{ \frac{\nu_s \lambda_1 \lambda_2}{\nu_1 \nu_2} \mathbf{x}_2^T \hat{\mathbf{D}}_\tau \mathbf{x}_1 \right\}, \quad (2.8)$$

which is the cross-correlation of \mathbf{x}_1 and \mathbf{x}_2 weighted by the attenuations. The cross-correlation term is of particular interest as it represents the degree of correlation between the received signals.

2.2.2 Cue processes

The two principle interaural variables that are considered as localising cues are ILD and ITD. Intrinsically, ITD is already part of my model as a hidden variable. However, I shall illustrate that ILD is characterised by the attenuations, λ_1 and λ_2 , and the signal to noise ratio $\frac{\eta}{\nu}$. I will use a common $\nu = \nu_1 = \nu_2$ and a time disparity τ of zero as this simplifies the resulting ILD distribution for the following illustration.

The distribution of interaural level disparities, ILD, can be computed from the joint likelihood (Equation 2.6) using the distribution to effectively cause a change of variable (x_1, x_2 to ILD),

$$p(\text{ILD}|x_1, x_2) = \delta \left(\text{ILD} - \log \left| \frac{x_1}{x_2} \right| \right). \quad (2.9)$$

Computing the expectation of Equation 2.9 over the distribution $p(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2|\lambda_1, \lambda_2, \tau, \nu_1, \nu_2, \eta)$ gives the ILD likelihood,

$$p(\text{ILD}|\lambda_1, \lambda_2, \nu, \eta) = \frac{1}{\pi} \sqrt{\frac{\nu_s}{\eta}} \left(\frac{1}{e^{\text{ILD}} + e^{-\text{ILD}} - \frac{\nu_s}{\nu} \left(\lambda_1 e^{\frac{\text{ILD}}{2}} + \lambda_2 e^{-\frac{\text{ILD}}{2}} \right)^2} + \frac{1}{e^{\text{ILD}} + e^{-\text{ILD}} - \frac{\nu_s}{\nu} \left(\lambda_1 e^{\frac{\text{ILD}}{2}} - \lambda_2 e^{-\frac{\text{ILD}}{2}} \right)^2} \right); \quad (2.10)$$

where the time disparity τ has been neglected and for simplicity $\nu_1 = \nu_2 = \nu$. This distribution of ILDs is depicted in Figure 2.4 to illustrate that the attenuations, λ_1 and λ_2 , relate to the measured ILDs observed in Nix and Hohmann (2006) in Figure 2.3. The contrast between Figure 2.4 plots (a) and (b) can be related to the contrast in Figure 2.3 between the cases of 15 degrees azimuth (the top pair of subplots) and 60 degrees azimuth

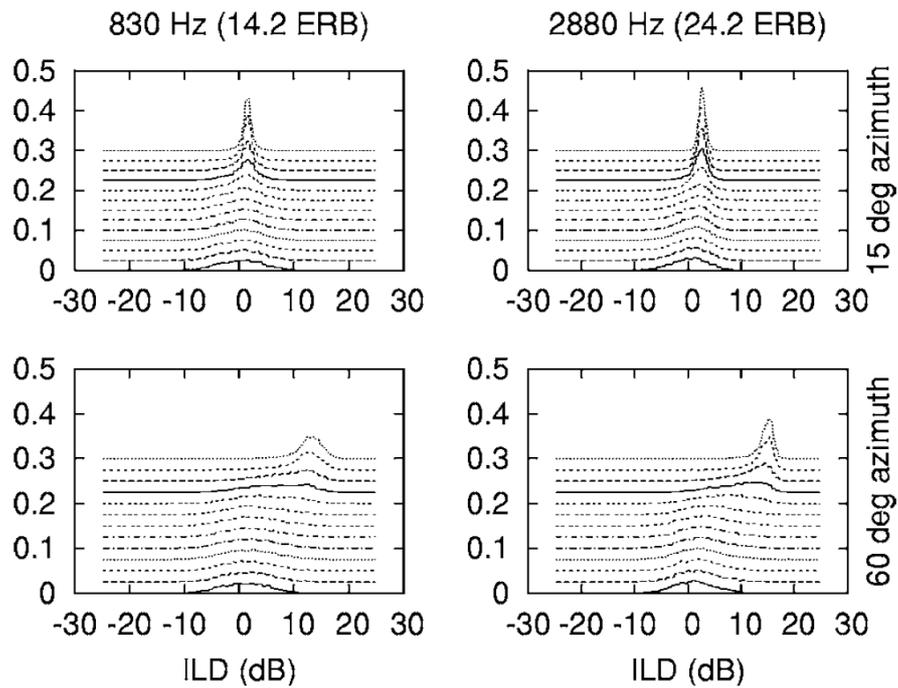
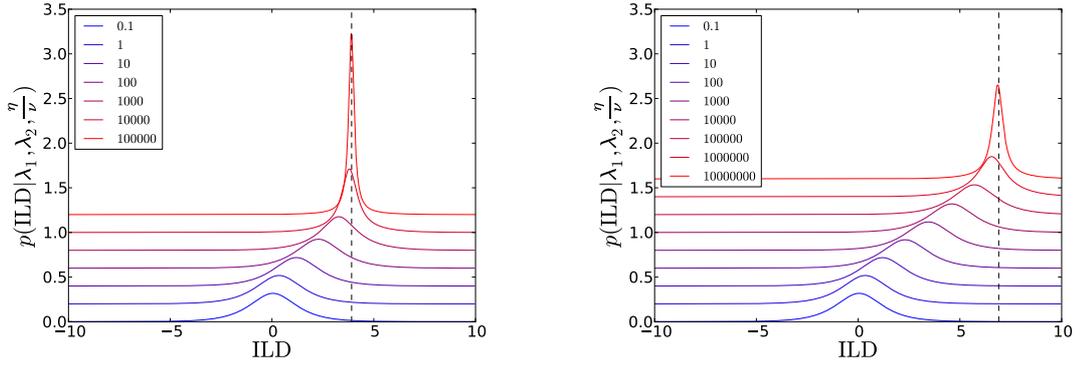


Figure 2.3: Figure 3 taken from Nix and Hohmann (2006). Histograms of interaural level disparity (ILD) values, i.e. the number of observations of a specific ILD normalised to the total count, for the speech target in cafeteria noise. The left panels are for a frequency of 830 Hz and the right panels are for 2.88 kHz. The upper panels are for 15° azimuth and the lower panels are for 60° azimuth 0° elevation, respectively. Each panel shows the distributions for the signal-to-noise ratio values of -5 , -2 , -1 , 0 , 1 , 2 , 3 , 5 , 10 , 15 , 20 , 30 dB and silence. The curves were shifted in order successively by 0.025 up the y-axis for clarity.



(a) Distribution of ILDs for $\lambda_2 = \frac{1}{50}\lambda_1$. The peak of the distribution, as $\frac{\eta}{\nu}$ decreases (indicated on the legend), corresponds to $\text{ILD} = \log \frac{\lambda_1}{\lambda_2} \approx 3.912$.

(b) Distribution of ILDs for $\lambda_2 = \frac{1}{1000}\lambda_1$. The peak of the distribution, as $\frac{\eta}{\nu}$ decreases (indicated on the legend), corresponds closely to $\text{ILD} = \log \frac{\lambda_1}{\lambda_2} \approx 6.908$.

Figure 2.4: Plots of the distribution of interaural level disparity (ILD) for variation of signal-to-noise ratio. Plotting the distribution $p(\text{ILD}|\lambda_1, \lambda_2, \eta, \nu)$ (Equation 2.10) with variation of the signal-to-noise ratio $\frac{\eta}{\nu}$, as indicated in the legend: plot (a) has $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{50}\lambda_1$, plot (b) has $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{1000}\lambda_1$. $\frac{\eta}{\nu}$ is roughly equivalent to the signal-to-noise ratio. This defines for simplicity in this figure that $\nu = \nu_1 = \nu_2$. This indicates the influence of the signal-to-noise ratio upon the distribution, a sharper peak suggests a greater correlation. This illustrates the dependence on signal-to-noise ratio of the ILD distributions, which are qualitatively similar to the plots in Figure 2.3 taken from Nix and Hohmann (2006). The difference between plots (a) and (b) illustrate the impact of the attenuations, λ_1 and λ_2 , on the distribution of level disparities ILD. This dependence of ILD upon λ_1 and λ_2 is starkly illustrated by the peak of the distribution, for decreasing $\frac{\eta}{\nu}$, corresponding closely to $\text{ILD} = \log \frac{\lambda_1}{\lambda_2}$. The curves were shifted in order successively by 0.2 up the y-axis for clarity.

(the bottom pair of subplots). At both frequencies (Figure 2.3) increasing the signal-to-noise ratio causes a sharpening and migration of the distribution of ILDs away from an ILD of 0 to a distribution centred at an ILD characteristic of the azimuth. Figure 2.4 shows the same sharpening and migration of $p(\text{ILD}|\lambda_1, \lambda_2, \frac{\eta}{\nu})$ with increasing signal-to-noise ratio $\frac{\eta}{\nu}$. Also, we can see that the $p(\text{ILD}|\lambda_1, \lambda_2, \frac{\eta}{\nu})$ with the sharpest peak³ is related to the ratio of attenuations $\frac{\lambda_1}{\lambda_2}$. This is indicative of the ILD being characterised by the attenuations, λ_1 and λ_2 . The attenuations convey the location dependence of the loudness cues.

Nix and Hohmann (2006) observed that both the mean and variance of the distribution of time disparities varied with a sound source’s location. I account for this using as a likelihood of the time disparity τ for a location l , a Gaussian distribution where the mean and variance are functions of location,

$$p(\tau|l) = \phi(\tau; \gamma_\tau(l), \omega_\tau(l)); \quad (2.11)$$

where the mean is $\gamma_\tau(l)$ and the variance is $\omega_\tau(l)$.

I choose to model the functions which link the source’s location l to the latent variables using a generative topographic mapping (GTM) (Bishop *et al.*, 1998a,b). This is because of its linearity with respect to learning and non-linearity with respect to inference. GTMs are a flexible method for representing a non-linear mapping between two spaces; it is relatively simple to integrate a GTM into Bayesian models in a principled fashion. This leads to the representation of the cues as the linear mixing of a previously selected set of functions. A vector of these functions $\boldsymbol{\psi}$ is constructed. While these functions may be non-linear with l , by adding a weight vector, this produces the mappings,

$$\lambda_1(l) = \mathbf{w}_1^T \boldsymbol{\psi}(l), \quad (2.12)$$

$$\lambda_2(l) = \mathbf{w}_2^T \boldsymbol{\psi}(l), \quad (2.13)$$

$$\gamma_\tau(l) = \mathbf{w}_\tau^T \boldsymbol{\psi}(l), \quad (2.14)$$

$$\omega_\tau(l) = \mathbf{w}_\omega^T \boldsymbol{\psi}(l), \quad (2.15)$$

which are linear with the weight vectors. This is functionally equivalent to an artificial neural network with a single layer (Bishop, 1995) and equivalent to charting a manifold in attenuations and time disparities using location as the point on the manifold. A useful aspect of the GTM model is due to this linearity with the weightings. This results in a

³The highest peak in Figure 2.4 corresponds to the highest signal-to-noise ratio $\frac{\eta}{\nu}$.

non-linear mapping for inference purposes but linear with its parameters for learning and analysis purposes.

2.2.3 The joint distribution

Combining my model of the cues with the sound process model of Beal *et al.* (2003) produces the joint distribution of measurable and latent variables

$$p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{s}, \tau | l) = \phi(\mathbf{x}_1; \lambda_1(l) \mathbf{s}, \nu) \phi(\mathbf{x}_2; \lambda_2(l) \hat{\mathbf{D}}_\tau \mathbf{s}, \nu) \phi(\mathbf{s}; 0, \eta) \times \phi(\tau; \gamma_\tau(l), \omega_\tau(l)). \quad (2.16)$$

Marginalising out the latent variables \mathbf{s} and τ of this joint belief leads to the likelihood of the measured variables \mathbf{x}_1 & \mathbf{x}_2 given the sound source's location l ,

$$p(\mathbf{x}_1, \mathbf{x}_2 | l) = \sqrt{\frac{\nu_{s|l}}{2\pi\nu_1\nu_2\eta}} \exp\left\{-\frac{1}{2\nu_1^2} (\nu_1 - \lambda_1^2(l) \nu_{s|l}) \mathbf{x}_1^T \mathbf{x}_1 - \frac{1}{2\nu_2^2} (\nu_2 - \lambda_2^2(l) \nu_{s|l}) \mathbf{x}_2^T \mathbf{x}_2\right\} \times \sum_{\tau} \left[\exp\left\{\frac{\lambda_1(l) \lambda_2(l) \nu_{s|l}}{\nu_1 \nu_2} \mathbf{x}_2^T \hat{\mathbf{D}}_\tau \mathbf{x}_1\right\} \times \phi(\tau; \gamma_\tau(l), \omega_\tau(l)) \right], \quad (2.17)$$

where $\frac{1}{\nu_{s|l}} = \frac{1}{\eta} + \frac{\lambda_1^2(l)}{\nu_1} + \frac{\lambda_2^2(l)}{\nu_2}$.

2.2.4 Latent variable inference

We can infer the latent variables \mathbf{s} , τ and l by the application of Bayes law to compute the posterior belief in the latent variables using the already discussed likelihoods. By computing the posterior of the source signal conditioned upon the data and latent variables,

$$p(\mathbf{s} | \mathbf{x}_1, \mathbf{x}_2, \tau, l) = \frac{p(\mathbf{s}, \mathbf{x}_1, \mathbf{x}_2 | \tau, l)}{p(\mathbf{x}_1, \mathbf{x}_2 | \tau, l)}, \quad (2.18)$$

$$= \frac{p(\mathbf{s}) p(\mathbf{x}_1 | \mathbf{s}, l) p(\mathbf{x}_2 | \mathbf{s}, \tau, l)}{p(\mathbf{x}_1, \mathbf{x}_2 | \tau, l)}, \quad (2.19)$$

$$= \phi(\mathbf{s}; \boldsymbol{\mu}_{\mathbf{s}|\tau, l}, \nu_{s|l}), \quad (2.20)$$

where $\frac{1}{\nu_{s|l}} = \frac{1}{\eta} + \frac{\lambda_1^2(l)}{\nu_1} + \frac{\lambda_2^2(l)}{\nu_2}$ and $\boldsymbol{\mu}_{\mathbf{s}|\tau, l} = \nu_{s|l} \left(\frac{\lambda_1(l)}{\nu_1} \mathbf{x}_1 + \frac{\lambda_2(l)}{\nu_2} \hat{\mathbf{D}}_\tau^T \mathbf{x}_2 \right)$, shows the source signal remains Gaussian when conditioned upon the evidence (\mathbf{x}_1 & \mathbf{x}_2).

The joint belief in the latent variables \mathbf{s} and τ over the data is a product of two

Gaussians and a third term,

$$p(\mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, l) = p(\mathbf{s} | \mathbf{x}_1, \mathbf{x}_2, \tau, l) p(\tau | \mathbf{x}_1, \mathbf{x}_2, l) \quad (2.21)$$

$$= \phi(\mathbf{s}; \boldsymbol{\mu}_{s|\tau, l}, \nu_{s|l}) p(\tau | \mathbf{x}_1, \mathbf{x}_2, l) \quad (2.22)$$

the likelihood of the source signal given evidence and the likelihood of the time disparities τ . This leads to the marginalised belief in τ of,

$$p(\tau | \mathbf{x}_1, \mathbf{x}_2, l) \propto p(\mathbf{x}_1, \mathbf{x}_2, \tau | l) \quad (2.23)$$

$$= p(\mathbf{x}_1, \mathbf{x}_2 | \tau, l) p(\tau | l) \quad (2.24)$$

$$\propto \exp \left\{ \frac{\lambda_1(l) \lambda_2(l) \nu_{s|l}}{\nu_1 \nu_2} \mathbf{x}_2^T \hat{\mathbf{D}}_\tau \mathbf{x}_1 \right\} \times \phi(\tau; \gamma_\tau(l), \omega_\tau(l)), \quad (2.25)$$

which is the product of a weighted cross correlation (equation 2.8) and moderated by a location, l , dependent prior likelihood of τ .

So given a uniform prior belief in location l , the posterior belief in l will simplify to,

$$p(l | \mathbf{x}_1, \mathbf{x}_2) \propto p(\mathbf{x}_1, \mathbf{x}_2 | l) p(l) \quad (2.26)$$

$$\propto \sum_{\tau} p(\mathbf{x}_1, \mathbf{x}_2 | \lambda_1(l), \lambda_2(l), \tau) \times \phi(\tau; \gamma_\tau(l), \omega_\tau(l)), \quad (2.27)$$

assuming a uniform prior $p(l)$. Thus, in marginalising the joint distribution I have constructed the posterior belief in location, given a stereo signal for passive sound source localisation tasks. These quantities are useful for location inference and learning by expectation-maximisation in the next subsection.

2.2.5 Learning model parameters by expectation-maximisation (EM)

Bayesian methods for learning maximise the likelihood of a data set for a parameterisation. However, in hidden variable problems an iterative approach is typically taken (MacKay, 2003, Hastie *et al.*, 2001, Bishop, 1995).

The expectation-maximisation (EM) algorithm starts with an initial guess for the parameters and maximises a bound upon the complete data likelihood to generate a new and “better guess” for the parameters (MacKay, 2003, Hastie *et al.*, 2001, Bishop, 1995). This is applied to a problem with J known locations l_j , and taking at each location l_j , N measure-

ments from a pair of sensors to produce a data set $\text{Dat} = \{\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, l_j : j \in [1, J], n \in [1, N]\}$.

The parameters that define the likelihood $p(\mathbf{x}_1, \mathbf{x}_2|l)$ are learnt using variational methods.

As the hidden variable \mathbf{s} is continuous, the problem is approached using the variational EM algorithm (Beal *et al.*, 2003, Ghahramani and Beal, 2000, Ghahramani and Hinton, 2000, Verbeek *et al.*, 2003). Assuming a uniform prior $p(l_j) = \frac{1}{J}$, the complete data log-likelihood is used to derive a lower-bound on the likelihood,

$$\mathcal{L}(\text{Dat}) = \sum_{n=1}^N \sum_{j=1}^J \log p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}|l_j) + \sum_{j=1}^J \log p(l_j), \quad (2.28)$$

$$\propto \sum_{n,j} \log \sum_{\tau} \int d\mathbf{s} p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau|l_j), \quad (2.29)$$

$$= \sum_{n,j} \log \sum_{\tau} \int d\mathbf{s} \frac{q_{j,n}(\mathbf{s}, \tau)}{q_{j,n}(\mathbf{s}, \tau)} \times p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau|l_j), \quad (2.30)$$

$$\geq \sum_{n,j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \log p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau|l_j) - \sum_{n,j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \log q_{j,n}(\mathbf{s}, \tau). \quad (2.31)$$

The inequality between Equation 2.30 and Equation 2.31 is due to Jensen's inequality, Appendix D, and is subject to the constraint that the arbitrary distribution q is a normalised probability density function, hence $\sum_{\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) = 1$. This provides a lower-bound upon the complete data log-likelihood for the optimal fitting problem.

The quantity $q_{j,n}(\mathbf{s}, \tau)$ is the proposal distribution, and as derived in Appendix B, is the joint posterior

$$q_{j,n}(\mathbf{s}, \tau) \equiv p^{\text{old}}(\mathbf{s}, \tau|l_j, \mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}), \quad (2.32)$$

for the initially proposed parameters $\lambda_{1|j}^{\text{old}}, \lambda_{2|j}^{\text{old}}, \gamma_{\tau|j}^{\text{old}}, \omega_{\tau|j}^{\text{old}}, \nu_1^{\text{old}}, \nu_2^{\text{old}}$ and η^{old} . The joint posterior factors according to the chain rule for probability⁴ are:

$$q_{j,n}(\mathbf{s}, \tau) = q_{\tau,j,n}(\mathbf{s}) q_{j,n}(\tau), \quad (2.33)$$

⁴A conditional joint distribution can always be factorised as $P(A, B|C) = P(A|B, C)P(B|C) = P(B|A, C)P(A|C)$. The choice of which factorisation to make is determined by its tractability and utility to the problem.

Algorithm 2.1 The variational EM algorithm for learning the parameters for the likelihood of my sound source localisation model. My model being defined according to the graphical model in Figure 2.2 (a) and described in Section 2.2.

Starting with a set of parameters labelled as old. For each location l_j construct an initial guess for $\lambda_{1|j}^{\text{old}}$, $\lambda_{2|j}^{\text{old}}$, $\gamma_{\tau|j}^{\text{old}}$ and $\omega_{\tau|j}^{\text{old}}$. Make an initial guess for ν_1^{old} , ν_2^{old} and η^{old} .

1. Compute the responsibilities as derived in Appendix B.

- Calculate for each location l_j the quantity $\nu_{s|j}^{\text{old}} = \frac{1}{\frac{1}{\eta^{\text{old}}} + \frac{(\lambda_{1|j}^{\text{old}})^2}{\nu_1^{\text{old}}} + \frac{(\lambda_{2|j}^{\text{old}})^2}{\nu_2^{\text{old}}}}$, $\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}}$

is not calculated, instead $\nu_{s|j}^{\text{old}} \left(\frac{\lambda_{1|j}^{\text{old}}}{\nu_1^{\text{old}}} \mathbf{x}_{1,j,n} + \frac{\lambda_{2|j}^{\text{old}}}{\nu_2^{\text{old}}} \hat{\mathbf{D}}_{\tau}^T \mathbf{x}_{2,j,n} \right)$ is inserted into the parameter updates.

- Compute for each of the N repeated measurements $\mathbf{x}_{1,j,n}$, $\mathbf{x}_{2,j,n}$, indexed by n , of each location l_j the distribution $q_{j,n}(\tau)$, see Equation 2.34.

2. Compute the parameter updates as stated in Appendix B for η , \mathbf{w}_1 , \mathbf{w}_2 , ν_1 , ν_2 , \mathbf{w}_{τ} , \mathbf{w}_{ω} and $\omega_{\tau|j}$.

3. If computing another iteration, set $\lambda_{1|j}^{\text{old}} = \lambda_{1|j}$, $\lambda_{2|j}^{\text{old}} = \lambda_{2|j}$, $\gamma_{\tau|j}^{\text{old}} = \gamma_{\tau|j}$, $\omega_{\tau|j}^{\text{old}} = \omega_{\tau|j}$, $\nu_1^{\text{old}} = \nu_1$, $\nu_2^{\text{old}} = \nu_2$ and $\eta^{\text{old}} = \eta$, then goto 1.

where $q_{\tau,j,n}(\mathbf{s}) = \phi(\mathbf{s}; \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}}, \nu_{s|j,n}^{\text{old}})$. The second term, which is independent of \mathbf{s} is,

$$q_{j,n}(\tau) \propto \exp \left\{ \frac{\lambda_{1|j}^{\text{old}} \lambda_{2|j}^{\text{old}} \nu_{s|j}^{\text{old}}}{\nu_1^{\text{old}} \nu_2^{\text{old}}} \mathbf{x}_{2,j,n}^T \hat{\mathbf{D}}_{\tau} \mathbf{x}_{1,j,n} \right\} \phi(\tau; \gamma_{\tau|j}^{\text{old}}, \omega_{\tau|j}^{\text{old}}). \quad (2.34)$$

Equations 2.33 & 2.34 are a restatement of Equations 2.21 & 2.25 from Subsection 2.2.4 for the previous best guess of the EM-algorithms parameters. Therefore, the portion of the bound that is dependent upon the new parametrisation is:

$$B_{\text{new}}^{\text{old}} \propto \sum_{j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \left\{ \log p(\mathbf{x}_{1,j,n} | \mathbf{s}, l_j) + \log p(\mathbf{x}_{2,j,n} | \mathbf{s}, \tau, l_j) + \log p(\tau | l_j) + \log p(\mathbf{s}) \right\}, \quad (2.35)$$

as $p(l_j) = \frac{1}{J} \forall j$, it is neglected.

I present the derivation of one EM-update in full. This is computed from the gradient of the bound as follows,

$$\frac{\partial}{\partial \eta^{-1}} B_{\text{new}}^{\text{old}} = \sum_{n,j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \frac{\partial}{\partial \eta^{-1}} \left\{ \frac{D_x}{2} \log \eta^{-1} - \frac{1}{2\eta} \mathbf{s}^T \mathbf{s} \right\}, \quad (2.36)$$

$$= \frac{1}{2} \sum_{n,j,\tau} q_{j,n}(\tau) \int d\mathbf{s} q_{\tau,j,n}(\mathbf{s}) (D_x \eta - \mathbf{s}^T \mathbf{s}), \quad (2.37)$$

$$= \frac{1}{2} \sum_{j,\tau} q_{j,n}(\tau) \left(D_x \eta - \left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} - D_x \nu_{s|j}^{\text{old}} \right), \quad (2.38)$$

which equating to zero and solving for η , results in the EM-update for the variance η of the source signal $\bar{\mathbf{s}}$ to give the update

$$\eta = \frac{1}{D_x J N} \sum_{n,j,\tau} q_{j,n}(\tau) \left(\left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} + D_x \nu_{s|j}^{\text{old}} \right), \quad (2.39)$$

where $\frac{1}{\nu_{s|j}^{\text{old}}} = \frac{1}{\eta^{\text{old}}} + \frac{(\lambda_{1|j}^{\text{old}})^2}{\nu_1^{\text{old}}} + \frac{(\lambda_{2|j}^{\text{old}})^2}{\nu_2^{\text{old}}}$ and $\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} = \nu_{s|j}^{\text{old}} \left(\frac{\lambda_{1|j}^{\text{old}}}{\nu_1^{\text{old}}} \mathbf{x}_{1,j,n} + \frac{\lambda_{2|j}^{\text{old}}}{\nu_2^{\text{old}}} \hat{\mathbf{D}}_\tau^T \mathbf{x}_{2,j,n} \right)$. The complete set of EM-updates are listed in Appendix B.

The cross-correlation computations $(\mathbf{x}_2^T \hat{\mathbf{D}}_\tau \mathbf{x}_1)$ are typically costly to compute and exist in the calculation of Equation 2.34 and every instance of $\left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}}$. However, using a fast Fourier transform to the Fourier domain (Frey and Jojic, 2001, 2003) expresses the convolution as a multiplication which scales as $O(D_x \log D_x)$, whereas directly computing the convolution scales as $O(D_x^2)$. An additional enhancement I use to reduce the computation time is to cache the quantities $\mathbf{x}_{1,j,n}^T \mathbf{x}_{1,j,n}$, $\mathbf{x}_{2,j,n}^T \mathbf{x}_{2,j,n}$ and $\mathbf{x}_{2,j,n}^T \hat{\mathbf{D}}_\tau \mathbf{x}_{1,j,n}$. This further reduces the calculation time for computing the responsibilities (the E-step) and the updates (the M-step). This defines the variational EM-updates in Algorithm 2.1 for learning the parameters of my model, the graphical model in Figure 2.2.

2.3 Results

In this section I present the results of comparing my model to that of Nix and Hohmann (2006). Both models are fit to a learning data set, with the graphed results constructed using a distinct validation data set.

I collected the data for a selection of locations upon the azimuth plane around a pair of microphones separated by 10 cm. This produced a set of stereo recordings for each location as indicated in Figure 2.5. 10 cm was selected in combination with the distributions of source positions in Figure 2.5 to provide experimental data over a range of locations proportionate to a humans reach relative to separation of the ears. The sound card used for collecting sound recordings was an M-Audio Audiophile 2496, this was due to its multichannel capabilities that were required for constructing the binaural signals.

The data set consisted of 60 location recordings, each of approximately 2 second duration using the default data encoding, RIFF little endian. These were each broken into

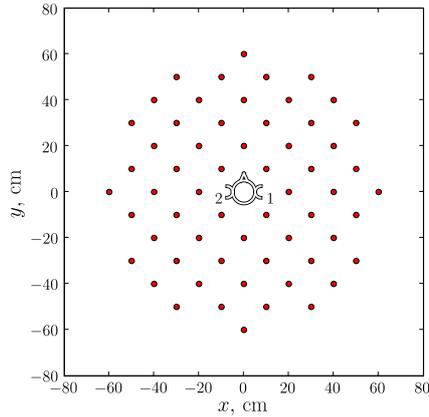


Figure 2.5: Cartoon depicting the spatial distribution of sound source localisation experimental measurements. The location of each microphone, labelled 1 or 2, is indicated by the label nearest to the cartoon heads ears near the origin.

200, 0.01s segments for use as the learning data set. The validation set was constructed through offsetting the segments of the learning data set by 0.005s to produce a different data set. To give the best range and resolution of sound loudness I used a 32 bit PCM encoding⁵ and likewise for the best resolution of timing features I used a sample rate of 44.1 kHz.

To have a sound source that was a random signal which exhibits time structure, I chose to use pink noise. This was generated by convolving Gaussian white noise with an exponential curve of 1kHz half life, which maintained the Gaussian source distribution. The source signal was encoded as a mono MPEG layer III (mp3), at a high bit-rate of 128kbits with a sample rate of 44.1kHz, so that the time structure could be resolved to the physical separation between sample points. This recording was presented through a speaker connected to an MP3 player (iRiver U10) at each selected location to construct the data sets. Due to technical constraints⁶ the source signal was randomly offset in time at each location, in effect the recording at each location was a different sound source with identical and stationary spectral characteristics. This produced independently and identically sampled data sets for the target location depicted in Figure 2.5 for a sound source with random time structure.

⁵Though there are logarithmic encodings available, consumer sound-cards seem to only support linear encoding.

⁶Having data that is sampled at 44.1kHz and different manual controls for starting the source playback and recording at each location.

2.3.1 Nix and Hohmann's approach

I construct a comparison between my approach and that of Nix and Hohmann (2006) on modelling passive sound localisation. I then compare the results of my model with the results of their model. The parameterisation is learnt using the learning data set, then each model is analysed with a validation data set.

Nix and Hohmann constructed location likelihoods from histograms of the ILD and IPD variables. The ILD and IPD variables were computed according to Appendix C.⁷ Each histogram had 50 bins on the range $-\pi$ to π for the phase difference and -25 to 25 dB level difference – They used a range of -50 to 50 dB level difference. I used the ILD range of -25 to 25 dB as the ILDs for my data fitted within this range. Both histograms were normalised and together form the joint likelihood

$$p(\text{ILD}, \text{IPD}|l_i) = p(\text{ILD}|l_i) p(\text{IPD}|l_i) \quad (2.40)$$

for the measured ILD and IPD given the location l_i which is indexed by $i \in [1, J]$. The posterior for each location is computed according to Bayes law for the n 'th segment taken from the j 'th location,

$$p(l_i|\text{ILD}_{j,n}, \text{IPD}_{j,n}) = \frac{p(\text{ILD}_{j,n}, \text{IPD}_{j,n}|l_i) p(l_i)}{\sum_i p(\text{ILD}_{j,n}, \text{IPD}_{j,n}|l_i) p(l_i)}, \quad (2.41)$$

where the prior $p(l_i)$ is uniform across all locations, e.g. $p(l_i) = \frac{1}{J}$.

In contrast, the individual posterior beliefs for my model are constructed from the validation set $\{\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n} : j \in [1, J], n \in [1, N_{\text{val}}]\}$. I compute a posterior belief according to,

$$p(l_i|\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}) = \frac{p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}|l_i) p(l_i)}{\sum_i p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}|l_i) p(l_i)}, \quad (2.42)$$

where the likelihood is,

$$p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}|l_i) = \sum_{\tau} p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}|\lambda_{1|i}, \lambda_{2|i}, \tau) \phi(\tau; \gamma_{\tau|i}, \omega_{\tau|i}), \quad (2.43)$$

and the prior $p(l_i) = \frac{1}{J}$ for all $i \in [1, J]$.

In keeping with Nix and Hohmann, the posterior beliefs were smoothed using a first-

⁷To summarise Appendix C, the only changes I made to the approach and model of Nix and Hohmann was in the computations of the ILD and IPD. These changes were to adapt their approach to become a valid comparison to my model using my data set. Specifically these were the use of one frequency channel, due to the nature of my data set, and using a segment length of 440 to correspond with the segment length used in my model.

Algorithm 2.2 Pseudocode used to generate inferences of sound source location. The measurement data Dat_j is taken at a location l_j , Dat_j is segmented into N_{val} segments denoted by $\text{Dat}_{j,n}$ for $n \in [1, N_{\text{val}}]$. The smoothed posterior is plotted for a selection of l_j 's in Figures 2.6-2.12 using the probability $p_{\text{smoo}}(l_i|\text{Dat}_j) = p_{\text{smoo}}(l_i|\text{Dat}_{j,0:N_{\text{val}}})$. For my model's likelihood this is $p_{\text{smoo}}(l_i|\mathbf{y}_{1,j}, \mathbf{y}_{2,j}) = p_{\text{smoo}}(l_i|\mathbf{y}_{1,j,0:N_{\text{val}}}, \mathbf{y}_{2,j,0:N_{\text{val}}})$ and for Nix and Hohmann's model's likelihood this is $p_{\text{smoo}}(l_i|\text{ILD}_j, \text{IPD}_j) = p_{\text{smoo}}(l_i|\text{ILD}_{j,0:N_{\text{val}}}, \text{IPD}_{j,0:N_{\text{val}}})$.

To construct a posterior belief in the location l_i with a uniform prior belief,

1. for each $j \in [1, J]$,
 - (a) for each $i \in [1, J]$,

$$p(l_i|\text{Dat}_{j,n}) = \frac{p(\text{Dat}_j|l_i)p(l_i)}{\sum_{i'} p(\text{Dat}_j|l_{i'})p(l_{i'})}.$$

The smoothed posterior belief is constructed iteratively using the process,

1. for each $j \in [1, J]$,
 - (a) for each $i \in [1, J]$,

$$p_{\text{smoo}}(l_i|\text{Dat}_{j,0:n}) = \alpha p(l_i|\text{Dat}_{j,n}) + (1 - \alpha) p_{\text{smoo}}(l_i|\text{Dat}_{j,0:n-1}).$$

Where $\alpha = \frac{\delta t}{\text{RC} + \delta t}$, δt is the segment duration of 10ms, and RC is the filters time constant of 100ms.

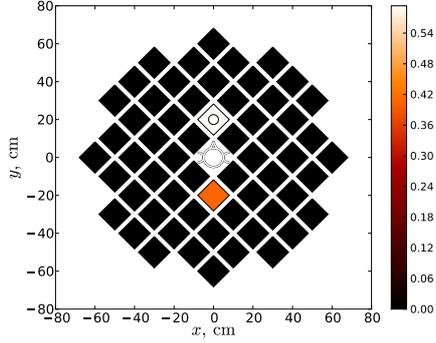
order, low-pass filter with a 100ms time constant.⁸ This is applied to the sets of posterior beliefs for both my model and Nix and Hohmann's model. There is a distinction between Nix and Hohmann's analysis compared to mine – They extracted the most probable direction from each of these sequentially smoothed posteriors to construct a histogram of estimated locations. In my comparison I shall use the posterior belief itself as this represents the current best estimate of target location for both models.

In order to make a comparison I smooth the sequential posterior beliefs with the same first-order, low-pass filter that Nix and Hohmann used. For my model and theirs these smoothed beliefs are compared, rather than histograms of the maximum *a posterior* estimates. The final smoothed beliefs are denoted for Nix and Hohmann's model as $p_{\text{smoo}}(l_i|\text{ILD}_j, \text{IPD}_j)$ and for my model $p_{\text{smoo}}(l_i|\mathbf{y}_{1,j}, \mathbf{y}_{2,j})$. This is summarised in Algorithm 2.2.

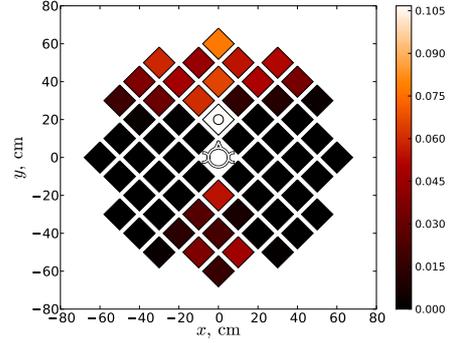
2.3.2 Comparison of inferences

A representative sample of the posterior location beliefs, for both my model and Nix and Hohmann's, is shown in Figures 2.6-2.12. These Figures are characteristic of the posterior

⁸For the sequence of 50 time segments constructed by offsetting the learning data set by 128 samples corresponding to $\sim 0.3s$.

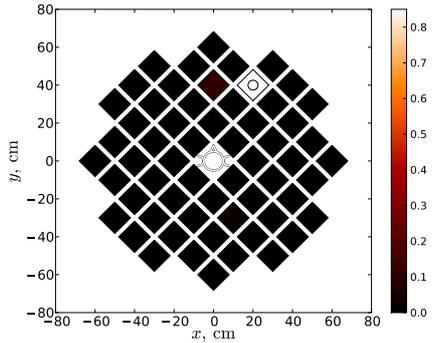


(a) Posterior probability, $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$.

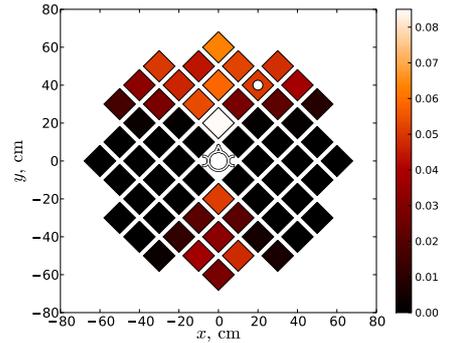


(b) Nix and Hohmann model's posterior probability, $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$.

Figure 2.6: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, an accurate most probable posterior belief with small front-back confusion in the posterior. Plot (b) depicts, for Nix and Hohmann's model, an accurate most probable posterior belief with an angular ambiguity and small front-back confusion in the posterior. The asymmetry in plot (b) is due to peculiarities in the data, as Nix and Hohmann use a non-parametric model – which uses the learning data set directly for making inferences.

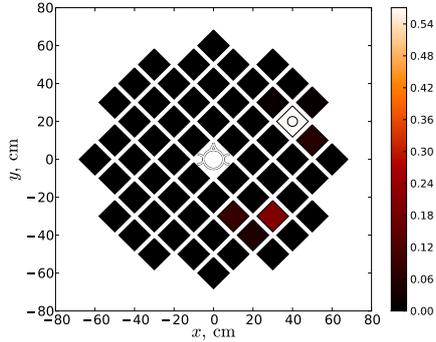


(a) Posterior probability, $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$.

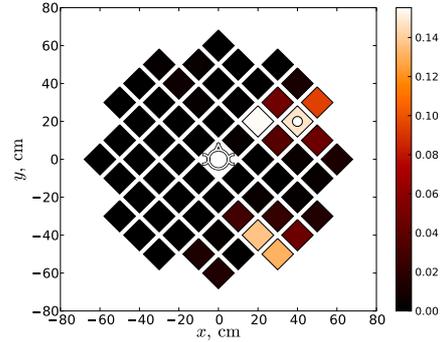


(b) Nix and Hohmann model's posterior probability, $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$.

Figure 2.7: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, an accurate most probable posterior belief with a mostly certain posterior. Plot (b) depicts, for Nix and Hohmann's model, small angular error in the most likely posterior belief with a small angular ambiguity and small front-back confusion in the posterior.

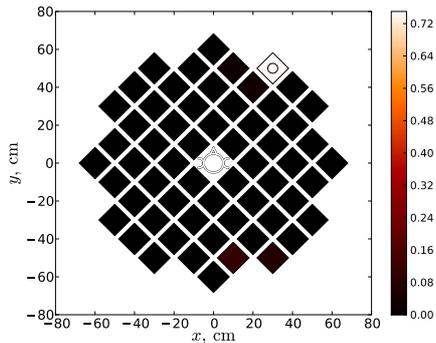


(a) Posterior probability, $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$.

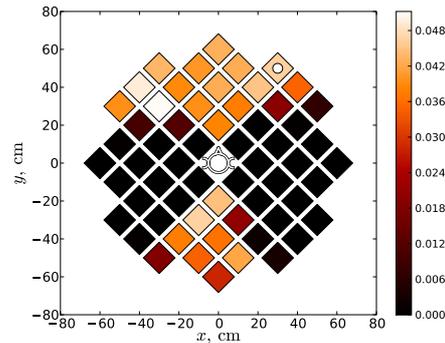


(b) Nix and Hohmann model's posterior probability, $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$.

Figure 2.8: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, an accurate most probable posterior belief with small front-back confusion in the posterior. Plot (b) depicts, for Nix and Hohmann's model, small angular error in the most likely posterior belief with a small angular ambiguity and a small front-back confusion in the posterior.

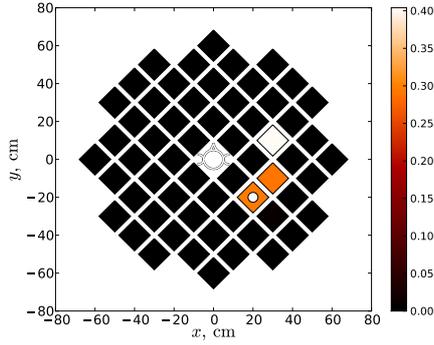


(a) Posterior probability, $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$.

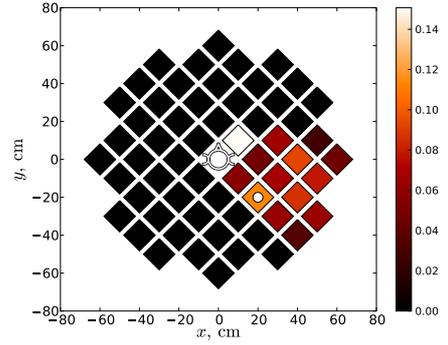


(b) Nix and Hohmann model's posterior probability, $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$.

Figure 2.9: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, an accurate most probable posterior belief with a mostly certain posterior. Plot (b) depicts, for Nix and Hohmann's model, a significant angular error is the most likely posterior belief with a significant angular ambiguity and front-back confusion in the posterior.

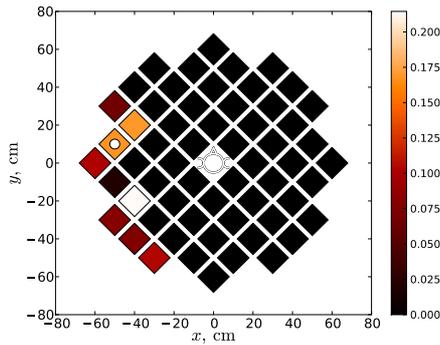


(a) Posterior probability, $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$.

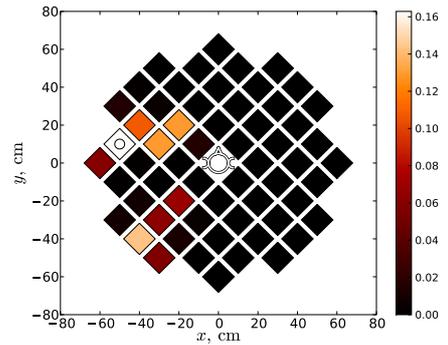


(b) Nix and Hohmann model's posterior probability, $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$.

Figure 2.10: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, a front-back error for the most probable posterior belief with a small front-back confusion in the posterior small angular ambiguity in the posterior. Plot (b) depicts, for Nix and Hohmann's model, front-back error for the most probable posterior belief with a small angular ambiguity and a front-back confusion in the posterior.



(a) Posterior probability, $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$.



(b) Nix and Hohmann model's posterior probability, $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$.

Figure 2.11: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smoo}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smoo}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, a small angular error in the most likely posterior belief with a small angular ambiguity and front-back confusion in the posterior. Plot (b) depicts, for Nix and Hohmann's model, an accurate most probable posterior belief with a small angular ambiguity and front-back confusion in the posterior.

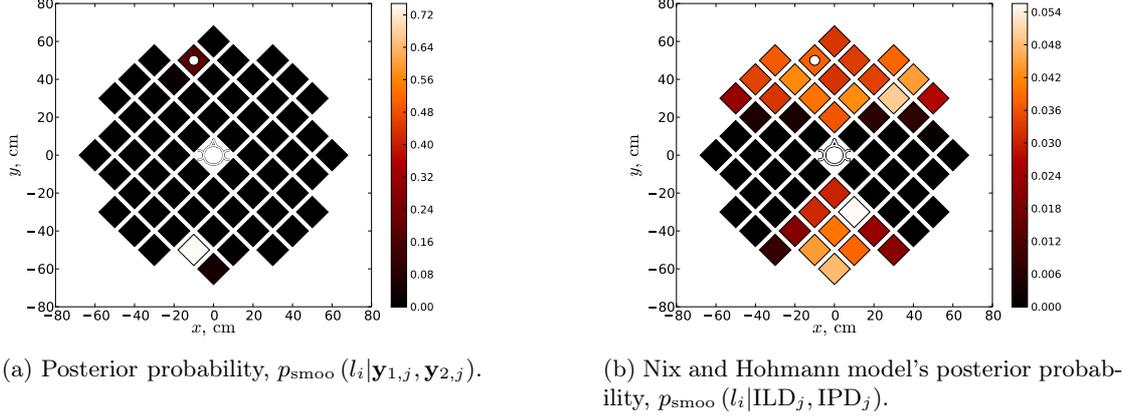


Figure 2.12: An auditory location belief for my model contrasted to Nix and Hohmann's model. The probability $p_{\text{smooth}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ (a) and comparative (Nix and Hohmann, 2006) posterior belief $p_{\text{smooth}}(l_i | \text{ILD}_j, \text{IPD}_j)$ (b) over the location. The true location is indicated by the white dot. Plot (a) depicts, for my model, a front-back error for the most probable posterior belief with a mostly certain posterior. Plot (b) depicts, for Nix and Hohmann's model, a front-back error for the most probable posterior belief with a significant angular ambiguity and front-back confusion in the posterior.

distributions for data sampled for all of the locations.

As can be seen from these plots there is a much greater certainty from my model in each instance when compared with that of Nix and Hohmann's. This could be caused by a number of factors. First that my model is significantly better than theirs, possibly due to the computation of the absolute sound level being discarded. Alternatively, I may have over simplifying assumptions in my model causing the likelihood to become over-fit by the learning algorithm (Hastie *et al.*, 2001, Bishop, 1995). This is a point that I shall discuss further in this chapters Discussion (Section 2.5).

2.3.3 Analysis of inferences

The smoothed posterior distributions in Figures 2.6-2.12 are characterised according to the accuracy and errors of the most probable location \hat{l}_j and also for the shape of the distribution itself. The most probable location for my model is defined as

$$\hat{l}_j = \arg \max_{l_i, i \in [1, J]} p_{\text{smooth}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j}), \quad (2.44)$$

where $p_{\text{smooth}}(l_i | \mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ is the smoothed posterior over the $n \in [1, N_{\text{val}}]$ indexed segments. Similarly, the same is the case for Nix and Hohmann's model,

$$\hat{l}_j = \arg \max_{l_i, i \in [1, J]} p_{\text{smooth}}(l_i | \text{ILD}_j, \text{IPD}_j), \quad (2.45)$$

where $p(l_i | \text{ILD}_j, \text{IPD}_j)$ is the smoothed posterior over the $n \in [1, N_{\text{val}}]$ indexed segments.

Firstly, the accuracy and errors for the Figures are characterised as,

1. an accurate most probable posterior belief, $\hat{l}_i = l_j$, Figures 2.6(a), 2.6(b), 2.7(a), 2.8(a), 2.9(a) & 2.11(b),
2. a small angular error (with a small ranging error) in the most probable posterior belief, roughly $\hat{l}_i \approx l_j$, Figures 2.7(b), 2.8(b) & 2.11(a),
3. a front-back error for the most probable posterior belief, roughly \hat{l}_i has an opposite sign along the y-axis and a similar location on the x-axis to l_j , Figures 2.10(a), 2.10(b), 2.12(a) & 2.12(b),
4. finally, a significant error, which is none of the previous errors, Figure 2.9(b).

Secondly, the posterior distributions shape is characterised for the Figures as,

1. almost total certainty in the posterior, roughly $p(\hat{l}_i | \dots) > 0.8$,
2. a mostly certain posterior, roughly $p(\hat{l}_i | \dots) > 0.65$, Figures 2.7(a), 2.9(a) & 2.12(a),
3. a significant angular ambiguity in the posterior, Figures 2.9(b) & 2.12(b),
4. an angular ambiguity in the posterior, Figure 2.6(b),
5. a small angular ambiguity in the posterior, Figures 2.7(b), 2.8(b), 2.10(a), 2.10(b), 2.11(a) & 2.11(b),
6. a front-back confusion in the posterior, Figures 2.9(b), 2.10(b), 2.11(a), 2.11(b) & 2.12(b),
7. a small front-back confusion in the posterior, Figures 2.6(a), 2.6(b), 2.7(b), 2.8(a), 2.8(b) & 2.10(a).

Both of these characterisations are tabulated for my model in Table 2.4, and for Nix and Hohmann's model in Table 2.5.

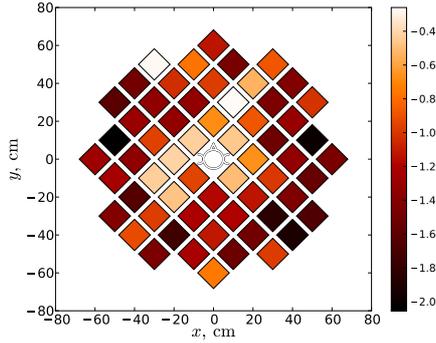
Looking at the tabulated results for my model in Table 2.4, we can see from the summed certainties (rightmost column \sum) that the posterior beliefs tend to be certain with any ambiguities also tending to be small; of the 60 locations, there are 16 unique certain cases and 38 unique accurate cases. The inference performance is very good; of the 60 locations 44 unique cases are accurate, 10 are slightly inaccurate and a few 6 have front-back errors.

	accurate	errors			Σ
		small angular	front-back	significant	
total certainty	2	0	0	0	2
certainty	14	1	2	0	17
large angular ambiguity	0	0	0	0	0
angular ambiguity	0	0	0	0	0
small angular ambiguity	14	9	2	0	25
front-back ambiguity	5	1	0	0	6
small front-back ambiguity	14	1	4	0	19
total inference performance	44	10	6	0	60

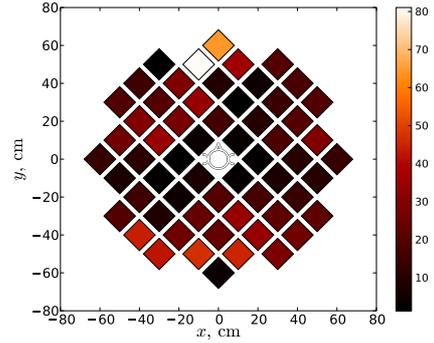
Table 2.4: Tabulation of the performances for my model compared to a description of its posterior belief for every sound source location. The columns represent the performances for my model. The rows represent the description of its posterior belief for every location. The last row *total inference performance* is not a sum of each column. This is because each inference is either accurate or makes one of a number of errors, but the posterior belief may have multiple descriptions, for instance it may have both angular and front-back ambiguity.

	accurate	error			Σ
		small angular	front-back	significant	
total certainty	0	0	0	0	0
certainty	0	0	0	0	0
large angular ambiguity	5	11	5	3	24
angular ambiguity	3	1	3	0	7
small angular ambiguity	17	6	6	0	29
front-back ambiguity	16	8	13	2	39
small front-back ambiguity	2	3	0	1	6
total inference performance	26	17	14	3	60

Table 2.5: Tabulation of the performances for Nix and Hohmann’s model compared to a description of its posterior belief for every sound source location. The columns represent the performances for my model. The rows represent the description of its posterior belief for every location. The last row *total inference performance* is not a sum of each column. This is because each inference is either accurate or makes one of a number of errors, but the posterior belief may have multiple descriptions, for instance it may have both angular and front-back ambiguity.



(a) The smoothed posterior Shannon information.



(b) The mean square error from the smoothed posterior belief.

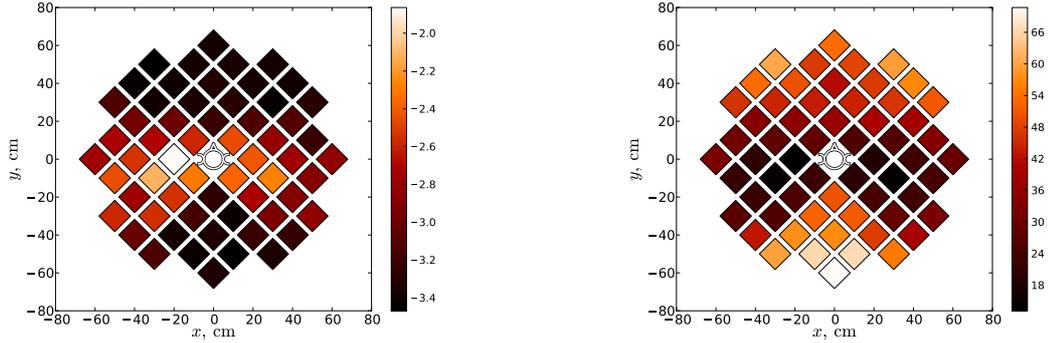
Figure 2.13: The uncertainty and accuracy of auditory location beliefs. The Shannon information, $I[L|\mathbf{y}_{1,j}, \mathbf{y}_{2,j}]$ for $l_i \in L$ in plot (a), representing belief’s uncertainty. The Cartesian mean square error of $p_{\text{smoo}}(l_i|\mathbf{y}_{1,j}, \mathbf{y}_{2,j})$ in plot (b), representing belief’s inverse accuracy. The correlation of the mean square error to the Shannon information is -0.446 . As such this provides a visual summary of ambiguity (a) and error (b) for Table 2.4.

The more certain inferences are associated with accurate inferences; within the 19 cases of certain inference there are 16 cases of inference accuracy.

A worrying problem is expressed in the cases where a few posterior certainties result in front-back errors, for example Figure 2.10(a). Similarly, in both cases of front-back error, the small angular ambiguity is coupled with a small front-back ambiguity, for example Figure 2.11(a). This indicates that the error is due to a much more probable posterior location causing a distraction from the actual location.

In contrast looking at the tabulated results for Nix and Hohmann’s model in Table 2.5, we can see from the cases of the summed certainties (rightmost column Σ) that the posterior beliefs tend to have ambiguities in the larger categories; of the 60 locations there are 47 unique cases that fit into both of the categories, *large angular ambiguity* and *front-back ambiguity*. While the inference performance is good; of the 60 locations 26 unique cases of being accurate, 17 of slight inaccuracy with 14 cases of front-back errors and a few (3 cases) of significant errors. However, large ambiguities are associated with inference errors; within 47 cases of ambiguity in the larger categories there are 30 cases of inference error.

For my approach the tabulated results in Table 2.4 are visually summarised by the mean square error (inverse of the accuracy) and the Shannon information (certainty) in Figure 2.13. For Nix and Hohmann’s approach the tabulated results in Table 2.5 are visually summarised by the mean square error (inverse of the accuracy) and the Shannon information (certainty) in Figure 2.14. Both Figure 2.14 and Tables 2.4 & 2.5 indicate a



(a) The smoothed posterior Shannon information.

(b) The mean square error from the smoothed posterior belief.

Figure 2.14: The uncertainty and accuracy of auditory location beliefs for the Nix and Hohmann (2006) approach. The Shannon information, $I[L|ILD_j, IPD_j]$ for $l_i \in L$ in plot (a), representing belief’s uncertainty. The Cartesian mean square error of the Nix and Hohmann (2006) comparison belief $p_{\text{smoo}}(l_i|ILD_j, IPD_j)$ in plot (b), representing belief’s inverse accuracy. The correlation of the mean square error to the Shannon information is -0.896 . As such this provides a visual summary of ambiguity (a) and error (b) for Table 2.5.

correlation between the mean square error and Shannon information, -0.886 for Nix and Hohmann’s approach and -0.446 for my approach. The lower degree of correlation for my approach appears to be due to the Shannon information being more noisy than that of Nix and Hohmann’s, respectively Figure 2.13 (a) compared to Figure 2.14 (a). Though in Nix and Hohmann’s case the regions to either side of the sensors display the greatest certainty and lowest error.

2.4 The contribution of cues

To consider the influence of the various latent cues upon the model I apply a variation of local weight analysis (Cook, 1986, Zhu and Lee, 2001) to the KL-divergence (Kullback and Leibler, 1951) of the model over the latent parameters (Appendix E). I use this to compare the two models Ω_a and Ω_b across both the inferred location and the hidden variables \bar{s} and τ . My use of local weight analysis considers the curvature of this objective function to a meta parameter r that operates upon an aspect of the model as indicated by Figure 2.15. I shall sketch the important parts of this approach, though the full details are described in Appendix E.

Local weight analysis makes use of the geometric curvature of the likelihood⁹ to consider the unit weighting of a part of a model to perturbations. This indicates the influence of

⁹In Zhu and Lee (2001) the curvature was taken for the EM bound of the model’s likelihood.

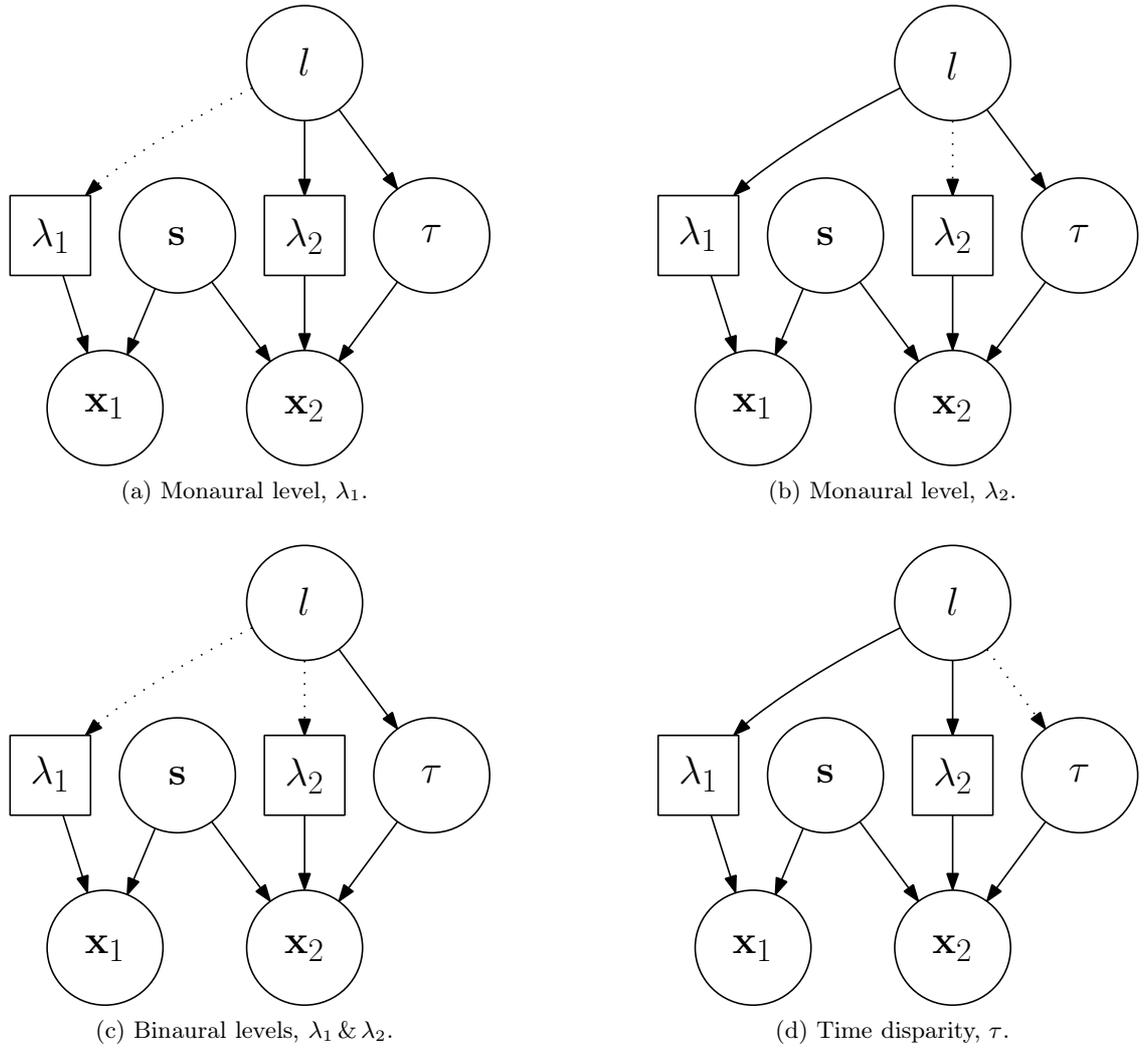


Figure 2.15: Graphical models depicting the perturbations of interest for the dependencies in Figure 2.2. The importance of these dependencies is evaluated by local weight analysis for the validation data set $\{\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n} : n \in [1, N_{\text{val}}]\}$ for each sampled location l_j indexed by j .

a part of the model's likelihood. The curvature for a curve, $c(\delta) = \{u(\delta), v(\delta)\}$ on the u, v -plane which is a function of δ , is defined as

$$C(\delta) = \frac{\left| \frac{\partial u}{\partial \delta} \frac{\partial^2 v}{\partial \delta^2} - \frac{\partial v}{\partial \delta} \frac{\partial^2 u}{\partial \delta^2} \right|}{\left(\left(\frac{\partial u}{\partial \delta} \right)^2 + \left(\frac{\partial v}{\partial \delta} \right)^2 \right)^{\frac{3}{2}}}. \quad (2.46)$$

I use as the curve c a quantity representing a functional measure of similarity of a probability belief to its unperturbed probability. The measure of similarity between the point belief of a model Ω_a and the perturbation Ω_b applied to itself is defined to be the KL-divergence of these two distributions. The perturbation is defined as a function of δ , $\Omega_b(\delta)$, which applies the analysis for the cases in Figure 2.15.

As I am considering the spatial dependence of the parameters (Figure 2.15), I consider the curvature is the line formed by $c(\delta) = \{\delta, \tilde{d}_j(\delta)\}$ – where, for a location l_j indexed by j , \tilde{d}_j represents the approximated measure of dissimilarity between Ω_a and $\Omega_b(\delta)$. The curvature of $c(\delta)$ for the perturbation $\Omega_b(\delta)$ applied to the set of parameters Ω is

$$C_j(\delta)|_{\Omega \leftarrow \Omega_b(\delta)} = \frac{\left| \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \Big|_{\Omega \leftarrow \Omega_b(\delta)} \right|}{\left(1 + \left(\frac{\partial \tilde{d}_j}{\partial \delta} \Big|_{\Omega \leftarrow \Omega_b(\delta)} \right)^2 \right)^{\frac{3}{2}}}. \quad (2.47)$$

To construct a measure of similarity \tilde{d}_j I start from the joint posterior belief in the hidden variables $p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega)$ and take the KL-divergence for the perturbation defined as $\Omega \leftarrow \Omega_b(\delta)$ to the unperturbed $\Omega \leftarrow \Omega_a$ which can be expressed as

$$D_{\text{KL}} [p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_a) || p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_b(\delta))]. \quad (2.48)$$

As I wish to analyse the spatial importance of these parameters I take the KL-divergence over this distribution for each true location l_j , indexed by j . I do this by taking the expectation of the KL-divergence in Equation 2.48 for the validation data sets¹⁰ sample distribution,

$$p(\mathbf{x}_1, \mathbf{x}_2 | l_j) \approx \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \delta(\mathbf{x}_1 - \mathbf{y}_{1,j,n}) \delta(\mathbf{x}_2 - \mathbf{y}_{2,j,n}). \quad (2.49)$$

Adding the quantity $\log \frac{p(\mathbf{x}_1, \mathbf{x}_2)}{p(\mathbf{x}_1, \mathbf{x}_2)}$ to Equation 2.48 and averaging over Equation 2.49 gives

¹⁰The validation measurements at each sampled location l_j are $\{\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n} : n \in [1, N_{\text{val}}]\}$.

an expectation of the spatial influence of a perturbation, which is

$$\tilde{d}_j \approx \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \log \frac{p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_a)}{p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b)}. \quad (2.50)$$

To compute the curvature $C_j(\delta)$ in Equation 2.47, requires the calculation of the first and second derivatives of \tilde{d}_j with respect to δ . As only Ω_b is dependent upon the perturbation by a factor δ then the derivatives of d_j simplify to

$$\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\Omega \leftarrow \Omega_b(\delta)} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b(\delta)), \quad (2.51)$$

and

$$\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\Omega \leftarrow \Omega_b(\delta)} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b(\delta)). \quad (2.52)$$

The cases of perturbation are $(\lambda_1 \leftarrow \delta \lambda_1)$, $(\lambda_2 \leftarrow \delta \lambda_2)$, $(\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta \lambda_2)$, $(\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2)$, $(\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2)$ and $(\gamma_\tau \leftarrow \delta \gamma_\tau)$.¹¹ The expected gradients for each of these perturbations are calculated in Appendix E. Algorithm 2.3 defines the algorithm to calculate the curvature for the perturbation $\lambda_1 \leftarrow \delta \lambda_1$, shown in Figure 2.2 (a).

The curvature is a measure of the sensitivity of the model's belief to variation of the parameter that the meta parameter δ affects. This allows me to show the importance of individual perturbations to each of the cases indicated in Figure 2.15 – this represents perturbations to the cues:

- λ_1 or λ_2 representing the monaural cues at either ear,
- λ_1 & λ_2 representing the average interaural loudness or the disparity in interaural

¹¹The joint posterior belief can be factorised by Bayes law to be,

$$p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) = p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a),$$

Where $p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a)$ is defined by Equation 2.27, $p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a)$ in Equation 2.25 and $p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a)$ in Equation 2.20. As $p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a)$ is a Gaussian, of mean $\boldsymbol{\mu}_{\mathbf{s} | \tau, j, i, n} = \nu_{\mathbf{s} | i} \left(\frac{\lambda_{1|i}}{\nu_1} \mathbf{y}_{1,j,n} + \frac{\lambda_{2|i}}{\nu_2} \hat{\mathbf{D}}_\tau^T \mathbf{y}_{2,j,n} \right)$ and the variance $\nu_{\mathbf{s} | i} = \frac{1}{\frac{1}{\eta} + \frac{\lambda_{1|i}^2}{\nu_1} + \frac{\lambda_{2|i}^2}{\nu_2}}$, the integrals over $\bar{\mathbf{s}}$ in Equations 2.51 & 2.52 are analytical.

Algorithm 2.3 My curvature analysis for my sound source localisation model. The algorithm used to compute the curvature of the KL-divergence between the joint belief and the perturbed joint belief. The algorithm is calculated for the case ($\lambda_1 = \delta\lambda_1$), Figure 2.15 (a). Similar computations are made for the other cases by substituting the relevant quantities for $h'(\tau)$ and $h''(\tau)$ from Appendix E. These quantities are the gradients under the expectations. This algorithm constructs the curvature for a set of $N_{\text{val}} = 50$ validation data points $\{\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n} : n \in [1, N_{\text{val}}]\}$ for each of the locations l_j indexed by $j \in [1, J]$.

for each $i \in [1, J]$,

$$\text{calculate and save } \nu_{s|i} = \frac{1}{\left(\frac{1}{\eta} + \frac{\lambda_1(l_i)^2}{\nu_1} + \frac{\lambda_2(l_i)^2}{\nu_2}\right)},$$

for each $j \in [1, J]$,

for each $n \in [1, N_{\text{val}}]$,

for each $i \in [1, J]$,

calculate

$$p(\tau|l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}) \propto \exp\left\{\frac{\lambda_1(l_i)\lambda_2(l_i)\nu_{s|i}}{\nu_1\nu_2}\mathbf{y}_{2,j,n}\hat{\mathbf{D}}_\tau\mathbf{y}_{1,j,n}\right\}\phi(\tau; \gamma_\tau(l_i), \omega_\tau(l_i)),$$

calculate

$$p(l_i|\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}) \propto \sum_{\tau} p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}|\lambda_1(l_i), \lambda_2(l_i), \tau)\phi(\tau; \gamma_\tau(l_i), \omega_\tau(l_i)),$$

normalise $p(\tau|l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n})$,

normalise $p(l_i|\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n})$,

set $\delta = 1$,

for each $j \in [1, J]$,

set $d'_j = 0$ and $d''_j = 0$,

for each $n \in [1, N_{\text{val}}]$,

set $g' = 0$ and $g'' = 0$,

for each $i \in [1, J]$,

$$h'(\tau) = \frac{\lambda_1(l_i)}{\nu_1}\left(\mathbf{y}_{1,j,n}^T\boldsymbol{\mu}_{s|\tau,j,i,n} - \delta\lambda_1(l_i)\left(\boldsymbol{\mu}_{s|\tau,j,i,n}^T\boldsymbol{\mu}_{s|\tau,j,i,n} + D_x\nu_{s|i}\right)\right),$$

$$h''(\tau) = \frac{\lambda_1(l_i)^2}{\nu_1}\left(\boldsymbol{\mu}_{s|\tau,j,i,n}^T\boldsymbol{\mu}_{s|\tau,j,i,n} + D_x\nu_{s|i}\right),$$

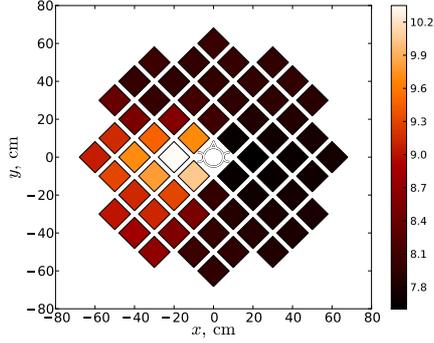
$$g' = p(l_i|\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n})\sum_{\tau} p(\tau|l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n})h'(\tau) + g',$$

$$g'' = p(l_i|\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n})\sum_{\tau} p(\tau|l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n})h''(\tau) + g'',$$

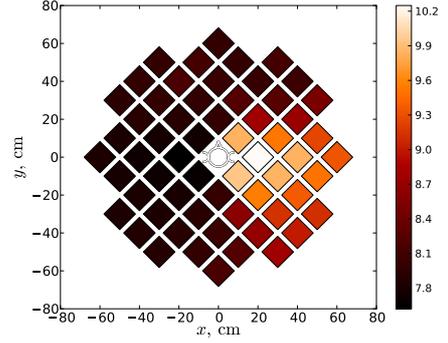
$$d'_j = g' + d'_j,$$

$$d''_j = g'' + d''_j,$$

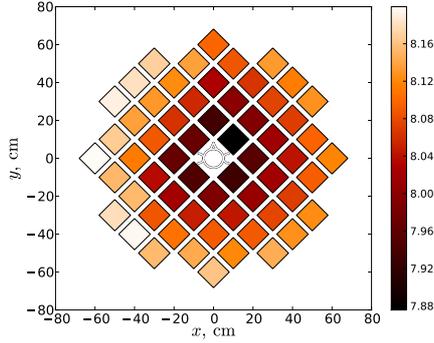
$$\text{calculate and save } C_j = \frac{|d''_j|}{(1+(d'_j)^2)^{\frac{\eta}{2}}},$$



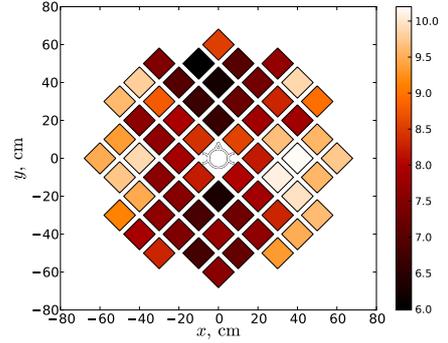
(a) Monaural level, ($\lambda_1 \leftarrow \delta\lambda_1$).



(b) Monaural level, ($\lambda_2 \leftarrow \delta\lambda_2$).



(c) Binaural levels, ($\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta\lambda_2$).



(d) Time disparity, ($\gamma_\tau \leftarrow \delta\gamma_\tau$).

Figure 2.16: Curvature analysis of the perturbations of interest from Figure 2.15. The average logarithmic curvature of the contribution for a sequence of measurements taken at each location. Top left ($\lambda_1 \leftarrow \delta\lambda_1$), top right ($\lambda_2 \leftarrow \delta\lambda_2$), bottom left ($\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta\lambda_2$) and bottom right ($\gamma_\tau \leftarrow \delta\gamma_\tau$).

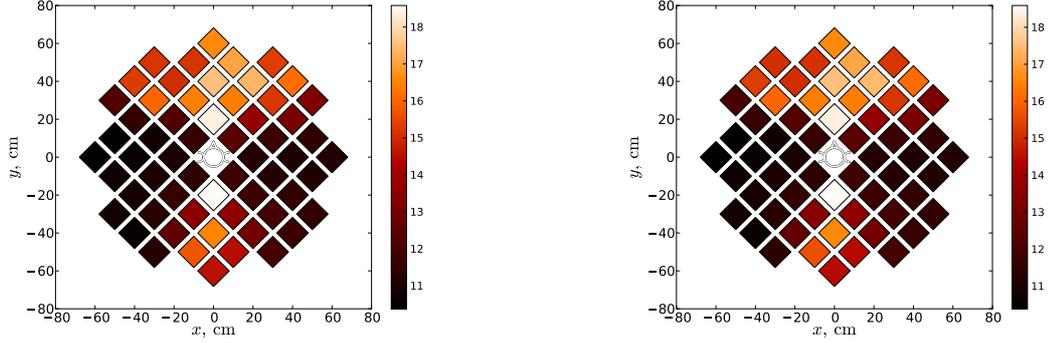
loudness

- and τ representing the interaural time disparity.

This cases are plotted in Figures 2.16 & 2.17.

Firstly, the curvatures of λ_1 and λ_2 (Figure 2.15 (a) and (b)) over space indicate that a perturbation will influence most the ipsilateral inferences in the near field (Figure 2.16 (a) and (b) plots respectively λ_1 and λ_2). The combined curvature of λ_1 and λ_2 (Figure 2.15 (c)) co-varying is greatest in the near field (Figure 2.16 (c)). These two results underline the fact that the influence of the hidden variable \bar{s} as a perturbation of the attenuations has the greatest influence, and hence curvature, when either of the λ_i 's are highest.

Secondly, the curvature of γ_τ depicted in Figure 2.15 (d) indicates that a change in γ_τ has the greatest impact upon both lateral regions (Figure 2.16 (d)). This is due to the nature of γ_τ influence – The Gaussian likelihood $p(\tau|l)$ acts as a windowing function that



(a) Level disparity, $(\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta^{-1}\lambda_2)$.

(b) Level disparity, $(\lambda_1 \leftarrow \delta^{-1}\lambda_1, \lambda_2 \leftarrow \delta\lambda_2)$.

Figure 2.17: Curvature analysis of the perturbations of interest from Figure 2.15 continued. The average logarithmic curvature of the contribution for a sequence of measurements taken at each location. Left $(\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta^{-1}\lambda_2)$ and right $(\lambda_1 \leftarrow \delta^{-1}\lambda_1, \lambda_2 \leftarrow \delta\lambda_2)$. Both curvatures seem to be identical.

moves away from the maximum of the cross-correlation term

$$\exp \left\{ \frac{\lambda_1(l) \lambda_2(l) \nu_{s|l}}{\nu_1 \nu_2} \mathbf{x}_2^T \hat{\mathbf{D}}_\tau \mathbf{x}_1 \right\} \quad (2.53)$$

from Equation 2.6. Next, the influence of increasing and decreasing the λ_i 's relative to one another (Figure 2.15 (c)) seems identical (Figure 2.17). These curvatures are dissimilar to those in Figure 2.16, the medial regions to both the agents front and back tend to have the greatest curvature reaching a minima in the lateral regions (Figure 2.17).

2.5 Discussion

2.5.1 Findings

I chose to compare my approach to that of Nix and Hohmann (2006) as their approach represents the-state-of-the-art in the analysis of passive sound source localisation. They define a non-parametric approach to naive Bayes sound source localisation published in the well read *Journal of The Acoustical Society of America*. I start my discussion by contrasting my model with the approach of Nix and Hohmann (2006), which similarly models passive sound source localisation. The model of Nix and Hohmann only considers the loudness processes as an ILD distribution. They showed that this ILD distribution is location and noise dependent in a characteristic fashion (Figure 2.3). Constructing an ILD likelihood for my model shows a qualitatively similar behaviour for noise dependence and location (through the attenuations, Figure 2.4).

My model considers the processes which cause the time and level cues as coinfluential and the influence of this coinfluence. In contrast these cues are treated as completely separable by Nix and Hohmann (2006) in a naive Bayes approach. My model assigns a high degree of importance to the interaction of the time disparity with the ILD as can be seen in Equation 2.6. This is in agreement with Park *et al.* (1996), which found that latency had a significant influence upon ILD encoding in LSO units. Similarly, for the sound level perspective, these results suggest that the time disparity component of my model associating time to location acts as a filter to consider correspondence of events in the sound field.

In the results section (Section 2.3) of this chapter I have presented a comparison of inference between my approach and Nix and Hohmann’s. This was through the smoothed posterior beliefs and relative performances of my approach to theirs. Both approaches construct location dependent likelihoods of the data. Their approach makes use of histograms of ILD and ITD through interaural phase disparities (IPD), whereas, my approach models the relationship between hidden variables using the graphical model in Figure 2.2.

My model considers the interactions between various latent quantities. This is in marked contrast to the approach by Nix and Hohmann (2006) where a non-parametric model is constructed leading only to the statement that the ILD and IPD play a role in the inferences. Figures 2.6-2.12 show that my approach has a much greater certainty when contrasted with that of Nix and Hohmann’s. This could be caused by my model retaining the absolute sound level. A qualification of the indicated correlation between error (mean square error) and certainty (Shannon information) from Tables 2.4 & 2.5 are respectively -0.446 for my approach (Figure 2.13) and -0.886 for Nix and Hohmann’s approach (Figure 2.14). This indicates another cause to my model’s much greater certainty than Nix and Hohmann’s approach. Though, an over simplification in my assumptions may have caused the likelihood to become over-fit (Hastie *et al.*, 2001, Bishop, 1995).

The analysis of curvature demonstrates that a perturbation applied to the monaural cues (single attenuations) will influence inference most in the contralateral region. That is because the magnitude of the attenuation is lowest when it is smoothest in Figure 2.19, hence a perturbation has the greatest impact. To consider the impact upon localisation the steepness of the attenuations λ_{ij} should be considered between neighbouring l_j ’s.

The curvature for the covariation of the attenuations has a discordance depending upon whether the variation is proportional (Figure 2.16 (c)) or inverse proportional (Fig-

ure 2.17 (a) & (b)). For the case of $(\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta\lambda_2)$, Figure 2.16 (c) depicts the importance of the same co-influence, *ie* that increasing or decreasing both attenuations in concert causes little impact. The curvature of the attenuations with one increasing and one decreasing causes a marked effect. It shows the importance of the ratio of attenuations to the medial regions both to the front and back, especially due to the magnitude of the curvature, Figures 2.16 & 2.17. This is in marked contrast to proportionate variation and like Macpherson and Middlebrooks (2002), little importance is assigned to monaural cues in lateral localisation.

We have seen how a perturbation to time disparity has the least impact upon the medial region as contrasted to the contralateral and ipsilateral regions, Figure 2.16 (d), which is surprising. My model treats the time disparity as a filter. The portion $p(\tau|l)$ acts as a window through the joint distribution and is equivalent to a convolved window to produce the likelihood. If the window hides the correlation then it reduces any consequent belief in that set of attenuations, resulting in a discordance. Hence a perturbation upon τ causes the greatest change where the correlation is lowest.

Not all Bayesian formulations for sound source localisation apply modelling tools such as graphical models to understand the rich interaction of the latent variables, Figure 2.2. This can leave a number of unexpected assumptions hidden due to being implicit within the framework. This is true of non-parametric methods such as Nix and Hohmann (2006), who construct likelihoods from histograms of data corresponding to known sound source locations. In contrast, graphical models provide great utility as a visual summary when considering the often complex dependencies within a model.

2.5.2 Limitations

There are a number of limitations to the approach I have taken in this Chapter.

Firstly, in learning the parameters for the likelihood $p(\mathbf{x}_1, \mathbf{x}_2|l)$ I have constructed the problem as a fitting problem, learning the distribution $p(\mathbf{x}_{1,l_a}, \mathbf{x}_{2,l_a}|l_a)$. Whereas in my analysis I use the likelihood for the inference of the sound source's location making the analysis a hidden variable problem, using the distribution $p(l_b|\mathbf{x}_{1,l_a}, \mathbf{x}_{2,l_a}) = \frac{p(\mathbf{x}_{1,l_a}, \mathbf{x}_{2,l_a}|l_b)p(l_b)}{\sum_{l_b} p(\mathbf{x}_{1,l_a}, \mathbf{x}_{2,l_a}|l_b)p(l_b)}$. In the analysis of models of passive sound source localisation such an approach is typical of the field (Nix and Hohmann, 2006). It is impractical to approach learning the locations as a hidden variable unless there is a further constraint upon it – The EM-algorithm in such a case converges to a solution where all locations are equally likely for every data point. The

choice of a prior has a large impact upon this, by providing a way to break this ambiguity. Handzel and Krishnaprasad (2002) make two suggestions for breaking such ambiguities:

1. by sensor placement or
2. head rotation.

These lead to concepts of active perception, how to choose a head rotation or other action and more generally to the sensor placement problem – I consider this problem in Chapter 3.

Secondly, related to the hidden variable problem is the lack of constraints upon the signals. In learning, this causes a problem of correspondence, which I define the *correspondence problem*.¹² With different types of signals and the co-influence of another sensory modality this can be overcome, for instance vision (Beal *et al.*, 2003, Hospedales *et al.*, 2007) which has an impact akin to the ventriloquist effect.¹³ Most recent work considers explicit spectral components (Oldfield and Parker, 1984), which can be applied through frequency channels in a naive Bayes manner. A more complete method could be used to optimally mix the time and level cues for different frequency channels.

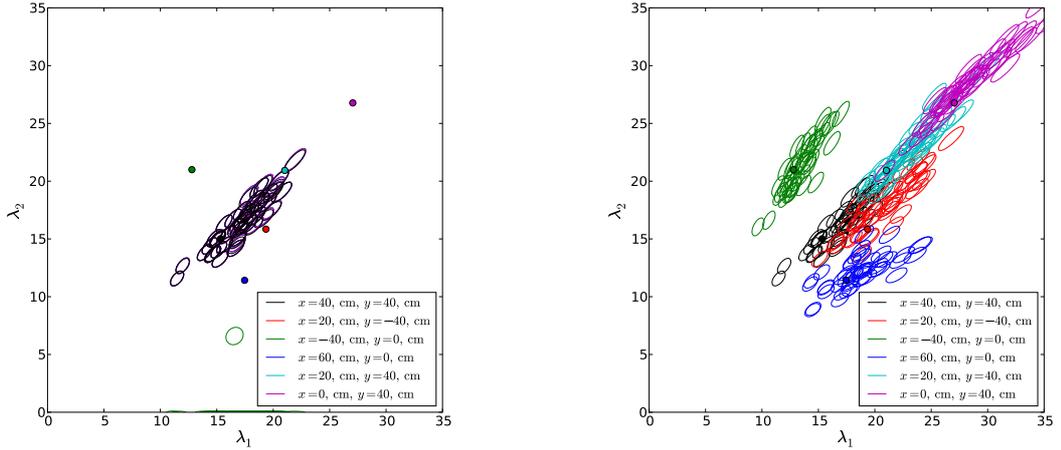
Finally, but partly related to the previous point is the treatment of the distribution of attenuations as parameters, which is equivalent to an attenuation likelihood of

$$p(\lambda_i|l) = \delta(\lambda_i - \lambda_i(l)), \quad (2.54)$$

a Dirac delta function. In Figure 2.18 I have plotted the likelihood $p(\mathbf{x}_{1,j}, \mathbf{x}_{2,j}|\lambda_1, \lambda_2, l_{j'})$ of the validation data for a selection of locations l_j spread over the λ_1, λ_2 -plane. Both plots in Figure 2.18 represent the EM-algorithms learnt parameters $\lambda_1(l_{j'})$ and $\lambda_2(l_{j'})$ as filled circles with the colours corresponding to the locations of the likelihoods as indicated in the legend. Figure 2.18 (a) indicates the confusion which can cause inference errors with a high posterior certainty. From Figure 2.18 (b) I hypothesise that the joint likelihood $p(\lambda_1, \lambda_2|l)$ can be constructed as a two dimensional Gaussian with a full covariance matrix, for example

¹²The problem of correspondence is a description of how the points on the hidden latent space (l_i) correspond to individual measurements ($\mathbf{x}_{1,j}, \mathbf{x}_{2,j}$). The difficulty comes when there is no prior constraint upon relating this latent space to individual measurements ($i \rightarrow j$) resulting in a belief $P(i|j)$ being uniform across a set of measurements. Thus without constraints such as prior knowledge it is difficult for a passive observer to learn that a measurement $\mathbf{x}_{1,j}, \mathbf{x}_{2,j}$ corresponds to a particular point on the latent space (l_i) rather than a different point $l_{i'}$ for all $i' \neq i$.

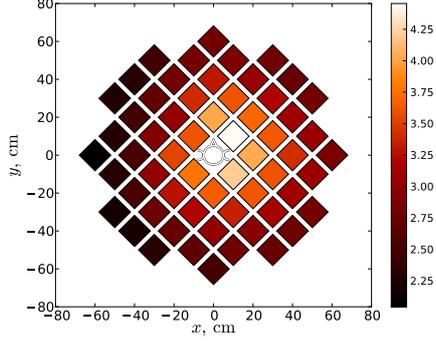
¹³Generally the ventriloquist effect is the subsumption of one sensory mode by another. In this framing it is the dominance of auditory localisation by visual speech cues using a dummy to confound localisation. Stratton (1897) (cited in Blauert (1997)) who found that if the visual field was flipped upside down then an auditory event was also inverted if it lay within the visual field. Interesting exceptions to this rule include Ewert (1930) (cited in Blauert (1997)) who discovered that using a lens system rather than a prism as by Stratton to distort the visual space that subjects could learn to form a direction despite a distorted visual field.



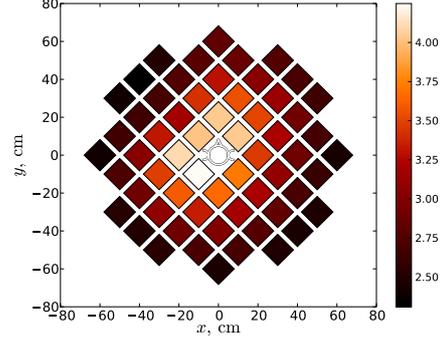
(a) Comparison of the likelihoods for data drawn from the same location $l_j = \{40\text{cm}, 40\text{cm}\}$ with different time disparity locations $l_{j'}$ for the belief $p(\tau|l_{j'})$ over which τ is marginalised.

(b) Comparison of the likelihoods for data drawn from the same location $l_{j'}$ as the belief $p(\tau|l_{j'})$ over which τ is marginalised.

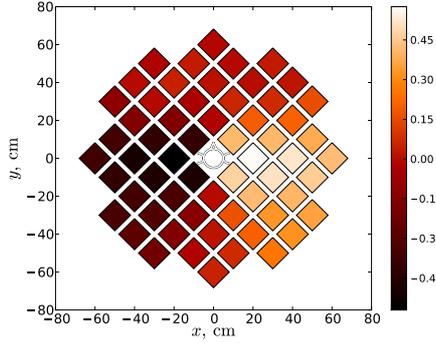
Figure 2.18: Distribution on the λ_1, λ_2 plane of the data likelihoods, illustrating the need for a joint distribution of $p(\lambda_1, \lambda_2|l)$. The distribution of each likelihood $p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}|l_{j'}, \lambda_1, \lambda_2)$ is illustrated with a line of the full width half maximum for the n 'th measurement at a location l_j , $\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}$. In both plots the legend indicates the location $l_{j'}$ which defines the distribution of time disparities $p(\tau|l_{j'})$ over which τ is marginalised. The joint belief being, $p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \tau|l_{j'}, \lambda_1, \lambda_2) = p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}|\tau, \lambda_1, \lambda_2) p(\tau|l_{j'})$, and from marginalising out τ gives, $p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}|l_{j'}, \lambda_1, \lambda_2) = \int d\tau p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \tau|l_{j'}, \lambda_1, \lambda_2)$. Plot (a) takes the data $\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}$ for the location $l_j = \{40\text{cm}, 40\text{cm}\}$ and compares its likelihood over the plane λ_1, λ_2 for the case of the correct location l_j ($j' = j$) and other incorrect locations $l_{j'}$ ($j' \neq j$). This indicates the confusion which causes inference errors in Figures 2.10-2.12 (a). Plot (b) takes the data $\mathbf{x}_{1,j',n}, \mathbf{x}_{2,j',n}$ at the locations $l_{j'}$ and compares these likelihoods over the plane of λ_1, λ_2 for each case in plot (a). We can see there is an overlap for a number of the source locations, which would suggest a greater degree of posterior uncertainty for each plot (a) in Figures 2.6-2.9.



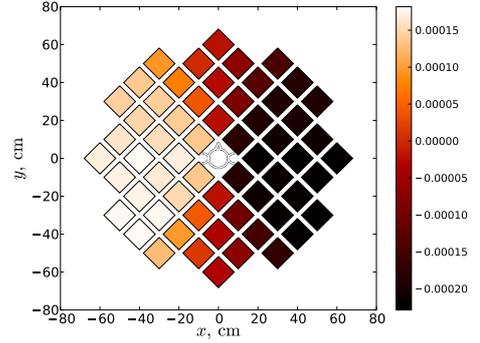
(a) The learnt attenuation $\log \lambda_1(l)$



(b) The learnt attenuation $\log \lambda_2(l)$



(c) The learnt attenuation ratio $\log \frac{\lambda_1(l)}{\lambda_2(l)}$



(d) The learnt time disparities $\gamma_\tau(l)$

Figure 2.19: Learnt parameters for each location using my model in Figure 2.2. The EM-algorithm's converged learned attenuations $\lambda_1(l)$ (a), $\lambda_2(l)$ (b), the ratio of these attenuations $\frac{\lambda_1(l)}{\lambda_2(l)}$ (c) and the time disparities $\gamma_\tau(l)$ (d).

$p(\lambda_1, \lambda_2|l) = \phi(\{\lambda_1, \lambda_2\}; \boldsymbol{\mu}_\lambda(l), \Sigma_\lambda(l))$. This would seem to represent the distribution and shape of the likelihoods on the λ_1, λ_2 -plane. The difficulty with such an extension is in estimating the marginal likelihood $p(\mathbf{x}_1, \mathbf{x}_2|l)$, as $p(\lambda_1, \lambda_2|l)$ is far broader in the λ_1, λ_2 -plane than $p(\mathbf{x}_1, \mathbf{x}_2|\lambda_1, \lambda_2, l)$ is for a given measurement.

2.5.3 Possible modifications

As seen in Figure 2.19 there is a smooth variation across the space of locations for both λ_i 's and the ratio $\frac{\lambda_1}{\lambda_2}$. The treatment of attenuation as a parameter, which is equivalent to a likelihood of a Dirac delta, could have been the cause of any over certainty in my results. I update the graphical model in Figure 2.2 (a), to account for the variation of the likelihood on the λ_1, λ_2 -plane in Figure 2.18, making λ_1 and λ_2 co-dependent. This causes the computation of the likelihood $p(\mathbf{x}_1, \mathbf{x}_2|l)$ to become a non-analytic integral

$$p(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2|l) = \int d\lambda_1 d\lambda_2 p(\lambda_1, \lambda_2|l) \sum_{\tau} p(\mathbf{x}_1, \mathbf{x}_2|\lambda_1, \lambda_2, \tau) p(\tau|l), \quad (2.55)$$

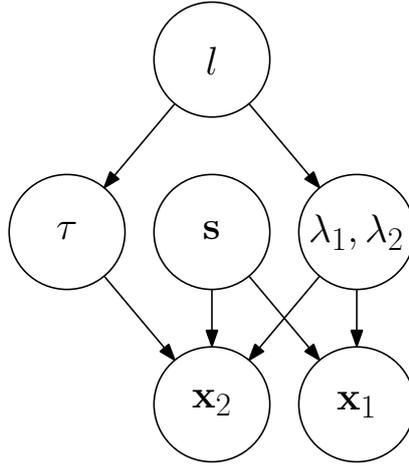


Figure 2.20: The updated graphical model indicating the dependencies between measurable, hidden variables and location. The latent variables are the attenuations λ_1 and λ_2 , and the relative time delay τ , the source signal \mathbf{s} and its location l . The observable variables are the sound measured at either ear \mathbf{x}_1 and \mathbf{x}_2 .

$$= \int d\lambda_1 d\lambda_2 p(\lambda_1, \lambda_2 | l) p(\mathbf{x}_1, \mathbf{x}_2 | l, \lambda_1, \lambda_2). \quad (2.56)$$

The integrals over λ_1 and λ_2 can be approached either by Gibbs sampling (MacKay, 2003) or grid based methods – both approaches will slow down calculation of the likelihood significantly. On the λ_1, λ_2 -plane, in Figure 2.18, for each location l the distribution of likelihoods $p(\mathbf{x}_1, \mathbf{x}_2 | \lambda_1, \lambda_2, l)$ is far broader than the individual likelihoods $p(\mathbf{x}_1, \mathbf{x}_2 | \lambda_1, \lambda_2, l)$. Therefore there will be a need for a large number of samples or a high grid resolution to adequately represent the attenuations joint likelihood $p(\lambda_1, \lambda_2 | l)$.

Finally, a more general model will treat the location and source loudness as hidden variables, though to do so and have the learnt space of locations have any geometric meaning may require an approach similar to Beal *et al.* (2003), Hospedales and Vijayakumar (2006), Hospedales *et al.* (2007) where location is defined as visual offset.¹⁴ A compelling alternative could utilise another method of associating measurement changes to a choice of action (Aytakin *et al.*, 2008, Noe, 2004, Philipona *et al.*, 2003, O’Regan and Noe, 2001, Handzel and Krishnaprasad, 2002, Pettorossi *et al.*, 2005), for instance head rotations.

¹⁴Beal *et al.* (2003), Hospedales and Vijayakumar (2006), Hospedales *et al.* (2007) constrain the problem such that there is little or no spatial ambiguity in the auditory portion of the problem, they considered only those portions of the world which are visible. These simplifications are for real time operation of the algorithms. Though the simplifications do directly constrain their models inferences,

- there is no auditory ambiguity between the forward and rear,
- it can only consider a limited range of attenuations and loudness’s.

2.5.4 Summary

I have conducted a Bayesian treatment of the spatial dependence of inference upon particular cues. I have compared the inference results of this treatment to the approach of Nix and Hohmann (2006). My principle contribution has been to develop a Bayesian treatment of the sound process for localisation in auditory cue fusion and how these cues influence the inference of a sound source's location.

I chose to compare my approach to that of Nix and Hohmann (2006) because their approach represents the-state-of-the-art in the analysis of passive sound source localisation. I have presented a comparison of inference, through the smoothed posterior beliefs and relative performances of my approach to theirs. This indicated my approach has better inference performance – consistently more certain and accurate, Tables 2.4 & 2.5.

I have also found that the attenuations are better represented as location dependent variables (as a joint distribution) and not as parameters. I have extended the work of Beal *et al.* (2003) concerning auditory localisation, by considering attenuation as location dependent and using a GTM to represent the relationship between spatial location and the hidden quantities (the attenuations and time disparities). I have applied a curvature analysis to indicate the influence of the cue processes as depicted in Figure 2.15. This methodology of curvature analysis can be applied to many different models to evaluate the consequence that an individual latent variable, parameter or collection of these has upon inference. However it is not applicable to non-parametric models. In this Chapter I have applied curvature analysis to my model of passive sound source localisation.

As a consequence of the curvature analysis we see that the influence of ILD and ITD upon sound source localisation is more complicated than usually assumed. The implication ties in with Park *et al.* (1997) that ILD tuning in LSO neurons of the auditory mid-brain have a non-linear dependence upon the loudness of perceived sound at either ear. Further Park *et al.* (1996) found ITD had an influence upon these LSO units in the mid-brain. A surprising feature of this confluence is that the curvature showed that each cue associated with a latent variable (Figure 2.15) had definite regions of the azimuth plane where each were most sensitive to perturbations. This suggests that there should be more confluences between cues in the tuning of neural units.

This work has considered passive localisation as a filtering problem. Patently it is not. The difficulty however is in considering the location as a hidden variable, due primarily to the problem of correspondence. Current work by Aytakin *et al.* (2008) suggests that

active observation could provide a framework for dealing with the correspondence problem. I approach this question of active observation and specifically how best to select these actions in the following Chapters.

Chapter 3

An information measure for optimal action selection

3.1 Introduction

A localising action is an action that takes a prior ambiguous belief in an inferred space such as self location upon a map, and will take an action that will reduce the consequent ambiguity in the belief after a measurement (posterior) by contrast to the prior belief. This leads to the question, *how are localising actions chosen efficiently when observations have led to ambiguous and conflicting inferences?*

There is little generality in our understanding of what makes different actions better or worse for such fast or reflexive localisation. Reflexive localisation is the fast or short time scale choice of action to rapidly increase an individual's knowledge of the world. Rapid inference in an uncertain world is critical for an individual's survival. In the context of sound source localisation a head rotation (Wallach, 1940, 1939) or pinnae movement (Walker *et al.*, 1998) is hypothesised to be used by an animal to apply a known effect upon the sensory cues that such an individual measures with its ears (Aytekin *et al.*, 2008). In this context a head rotation is literally the animal rotating its head within the sound field to cause a predictable change in the sensory cues it perceives (Wallach, 1940, 1939). Pinnae oscillations¹ are an oscillatory bobbing of the ears (the ears oscillating out of phase with each other) that cause a predictable change to the animal's perception of the perceived sound field (Walker *et al.*, 1998). This indicates the importance of controlled self motion (or at least known (Pettorossi *et al.*, 2005)) to an observer's accuracy and certainty in its

¹Pinnae oscillations are restricted to some genus' of echo-locating bats.

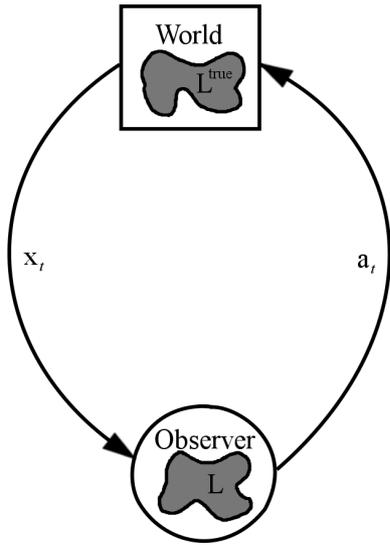
perceptions. However a typical limitation of such analyses of egocentric action perception is the specificity of any explanation – These are typically specific to the problem domain and do not form a fully general framework.²

In their work on robotics Cassandra *et al.* (1996) have shown that the expected entropy of belief in self localisation can be combined with other action selection methods, to concurrently generate a choice of action that actively reduces uncertainty in a filtered latent state space if it becomes too high. A general depiction of this problem can be seen in Figure 3.1. On average, taking a measurement will reduce inference ambiguity. Equivalently this can be stated as, the tracked *a posteriori* belief will be expected to become less ambiguous than its prior belief. Fox *et al.* (1998) continued to develop this and applied it as a method of reducing ambiguities for a robot localising itself upon a map (self localisation). Porta *et al.* (2003, 2005) further developed this concept by utilising an appearance based templating method (Beal *et al.*, 2003, Kristjansson *et al.*, 2004) for self localisation upon a previously constructed graph of positions representing the latent state space (similar to a self-organising map (Rojas, 1996)). The state of the art for this methodology is to select the optimal action when the ambiguity in an *a priori* belief has passed a certain threshold. This optimal action is performed as an alternative to other tasks which are typically exploitative of the individual’s knowledge rather than seeking to add to that knowledge (Porta *et al.*, 2003, 2005, Fox *et al.*, 1998, Cassandra *et al.*, 1996).

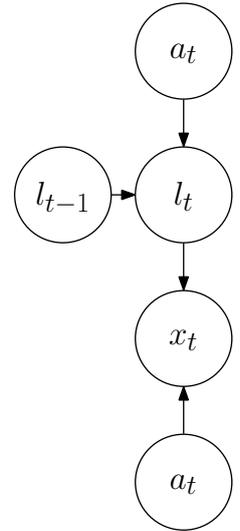
Vergassola *et al.* (2007a) have presented a similar approach but with the distinction that the information measure is the complete policy for action selection. They utilised a model of mating behaviour, where a male must locate and capture a female within a turbulent odour plume (Murlis *et al.*, 1992). Vergassola *et al.* define a method called Infotaxis which chooses the action expected to minimise the entropy in belief over the hidden latent state space of possible source positions. Infotaxis results in similar behaviour to the moth’s cast and surge behaviours (Baker, 1986, Vickers and Baker, 1994) and has a similar form to the best action selection framework of Cassandra *et al.* (1996), Fox *et al.* (1998). This suggests that an approach similar to Infotaxis, utilising statistical entropy in belief, can be generalised to construct a localising action selection framework (policy).

The approach of Vergassola *et al.* works in the tracking problem domain due to the existence of a smooth attractor of belief around the target. Put simply, the closer the agent

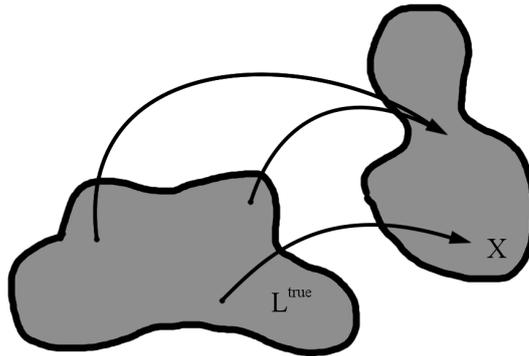
²The active models expressed in Walker *et al.* (1998), Muller and Schnitzler (1999, 2000, 2001) are limited to the sonar domains. Those models described in Baker (1986), Vickers and Baker (1994) are limited to Chemotaxis and olfactory domains. The model described in Olsson *et al.* (2006, 2005a), Smith (1997) is principally limited to visual problems.



(a) The observers model L of the world L^{true} .



(b) Graphical model of the Markov transitions.



(c) Ambiguity in the observable X due to the mapping from L^{true} to X . This can be dependent upon an action a_t .

Figure 3.1: A depiction of the active-perception inference problem. In plot (a), the ‘world’ has a true hidden latent state space L^{true} which represents some of its characteristics which are useful for an ‘observer’ to know. The observer cannot perfectly know this true latent state space L^{true} so must approximate it with L . The ‘observer’ interacts with the ‘world’ by its actions $a_t \in A$ and its measurements $x_t \in X$. In practice this is an iterative process where an action a_t is taken and modifies the state of the world in a predictable manner. I denote this change in plot (b) using a graphical model to depict the Markov transition, namely that a prior state l_{t-1} and a choice of action a_t causes a transition to a state l_t . The choice of a_t can also impact the measurement process. This graphical model captures the complete class of Markovian sensor placement problems. The new state l_t can be measured via the observable space X , which produces the measurement x_t . Plot (c) shows an example of 3 points on L^{true} mapping to X . Of these three selected points on L^{true} , two map to the same point upon X and are therefore ambiguous. The observer will similarly model a prediction of the transition from l_{t-1} to l_t due to a_t , and then refine this prediction conditional upon the taken measurement x_t also dependent on a_t .

gets to its target within the plume the more likely it is to detect a packet of pheromone and hence its certainty will increase until it reaches a maximum at its target. If the reverse were true then there would be no symmetry between the Infotaxis and Chemotaxis policies leading to the observer fleeing the target.

The work in Porta *et al.* (2003, 2005), Fox *et al.* (1998), Cassandra *et al.* (1996) and Vergassola *et al.* (2007a) provide an insight into more complex localisation tasks. For example, if a consequence of an action is to capture a target then the best localisation action naturally has the result of being exploitative. Hence, such informatic policies deal implicitly with the exploitation versus exploration dilemma for such classes of problem.

Information based approaches provide a set of tools for the analysis of reflexive localising actions. Examples of such actions include head rotations and pinnae movements³ of animals for auditory localisation. The advantage of these frameworks are the inherent flexibility and emergent consequence of the rules that can be elucidated when applied to particular problems. These tools are not limited to any particular sensory cues or modes of perception.

Infomax approaches are typically costly in time complexity. In this Chapter I introduce an alternative information measure with some useful properties. I will place it upon a concrete footing and relate it to the Shannon information and mutual information. I will show by application to a toy active sound source localisation problem that it will select a near optimal action. I then experimentally compare my method to a state of the art problem (Vergassola *et al.*, 2007a) where an information based policy was presented. This allows me to show the usefulness of my method to localisation problems when compared to approaches such as Porta *et al.* (2003, 2005), Fox *et al.* (1998), Cassandra *et al.* (1996), Vergassola *et al.* (2007a).

I have developed an alternative method to Shannon information and mutual information for solving the sensor placement problem satisfactorily. I consider the mathematical consequences of my solution in Section 3.4 and the algorithmic complexity in its estimation when contrasted to Shannon information (Section 3.5). I then compare it to Infomax in two problems:

- A toy perception problem related to the sound localisation problem from previous chapter in Section 3.6.
- A state-of-the-art problem taken from the literature (Vergassola *et al.*, 2007a) in

³Pinnae movements (oscillations) refer to the controlled movement of deformation of the auricle of the ear (pinnae) for a given task, in this context to aid auditory perception; for example Walker *et al.* (1998) provided an explanation of pinnae oscillation for echo-locating bats.

3.2 Hypothesis of active perception

With reference to Figure 3.1, there exists an observable space X which is assumed to be dependent upon a hidden latent state space L . Taking a prior belief on this hidden latent state space at the $(t - 1)$ th time as $p(l_{t-1})$ for all $l_{t-1} \in L_{t-1}$,⁴ what is the optimal approach to the selection of an action a_t ? Both the latent dynamics $p(l_t|l_{t-1}, a_t)$ and the measurement process $p(x_t|l_t, a_t)$ can be effected by an action a_t . The action a_t is selected so as to maximise our expected knowledge of the latent space L_t for a future measurement x_t .⁵

The current state of the art is described as Infomax (Porta *et al.*, 2003, 2005, Fox *et al.*, 1998, Cassandra *et al.*, 1996, Vergassola *et al.*, 2007a). It selects an action to maximise the conditional Shannon information,

$$I[L_t|X_t, a_t] = \int dl_t dx_t p(x_t, l_t|a_t) \log p(l_t|x_t, a_t), \quad (3.1)$$

which is the expectation of the Shannon information for the posterior belief. An agent will then take the action expected to give the most informative *a posteriori* belief $p(l_t|x_t, a_t)$.

These Infomax calculations can be slow and non-analytic. These calculations usually need approximation, for instance by approximating the posterior $p(l_t|x_t, a_t)$. In this chapter I describe a method for constructing a measure that scales more favourably, and does not need to approximate the posterior belief. There is a class of problem that I show has a lower algorithmic complexity to equivalent Infomax approaches.

Instead of looking directly at the posterior, maximising the expected posterior information (conditional Shannon information, Equation 3.1), I consider the action to maximally distinguish between the hypotheses of the latent space L_t for a measurement using the likelihood $p(x_t|l_t, a_t)$. The purpose is to construct the situation where the measurement will distinguish most between the hypotheses on the latent space. I do this through the KL-divergence (Cover and Thomas, 2006) between the likelihood of two latent hypotheses,

⁴A note on notation, the space Y_t for the t th time step is equivalent to Y .

⁵Typically this will tend to decrease the uncertainty from l_{t-1} to l_t with a measurement of x_t , however it is not possible to prove that this is the case, see Isard and Blake (1998).

l_t and l'_t ,

$$D_{\text{KL}} [p(X_t|l_t, a_t) || p(X_t|l'_t, a_t)] = \int dx_t p(X_t|l_t, a_t) \log \frac{p(X_t|l_t, a_t)}{p(X_t|l'_t, a_t)}. \quad (3.2)$$

This is a measure of the divergence between two distributions and as such it is not symmetric for l_t and l'_t . It is not a metric quantity. Therefore it does not define a distance between the two hypotheses. However, it does form a premetric quantity which can generate a topology on the space of hypotheses of l'_t given a hypothesis l_t and an action a_t .

To account for the distribution of latent predictions hypotheses ($p(l_t|a_t)$ and $p(l'_t|a_t)$), the expectation over the latent predictions of the KL-divergence is taken

$$B_{X_t|a_t} [L_t || L'_t] = \int dl_t dl'_t p(l_t|a_t) p(l'_t|a_t) D_{\text{KL}} [p(X_t|l_t, a_t) || p(X_t|l'_t, a_t)]. \quad (3.3)$$

This is the expected divergence between sections of the latent prediction. I shall show that maximising this measure gives the choice of action that will cause a measurement x_t to distinguish most between the latent state predictions.

3.3 Sketch of Infomax

In this section I sketch some interpretations of the Infomax approach to action selection through the use of various standard identities (Cover and Thomas, 2006 and Appendix I). There are two closely related Infomax measures, conditional Shannon information and mutual information. However, problems containing continuous measurement variables require approximation when computing either of these two Infomax measures. This can be approached using either sampling (Porta *et al.*, 2005), grid based methods or assuming the form of the marginal distribution (Vergassola *et al.*, 2007a). The choice of action is achieved by conducting a brute force argument maximisation over all of the possible actions.

The conditional Shannon information for the posterior belief $p(l_t|x_t, a_t)$ is the expected posterior Shannon information

$$I [L_t|X_t, a_t] = \mathbb{E}_{p(x_t|a_t)} [I [L_t|x_t, a_t]], \quad (3.4)$$

by Definition I.1. Where $I [L_t|x_t, a_t]$ is the Shannon information or negentropy of the posterior belief for a measurement x_t , the notation $\mathbb{E}_{p(x|\dots)} [f(x)]$ indicates the expectation of $f(x)$ over the distribution $p(x|\dots)$. The conditional Shannon information can be

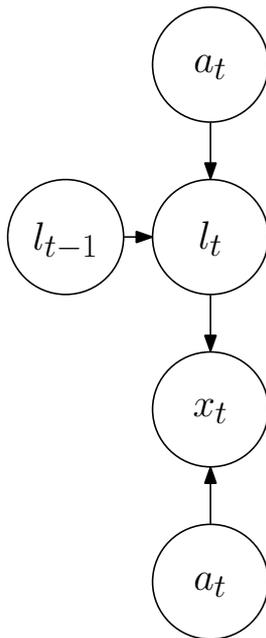


Figure 3.2: Graphical model of the action-perception model, repeated from Figure 3.1. For a prior belief in some latent state space L_{t-1} , $p(l_{t-1})$, is used to make a prediction $p(l_t|a_t)$ of the current belief l_t for a choice of action a_t using the likelihood $p(l_t|l_{t-1}, a_t)$. The measurement ($x_t \in X_t$) process is modelled using the likelihood $p(x_t|l_t, a_t)$ for a current belief l_t and choice of action a_t . This graphical model captures the complete class of Markovian sensor placement problems except for those cases where an action is constrained by the agent's pose – Pose can be modelled as part of the latent space l_t . A common action a_t is used for both the dynamics of the latent state, $p(l_t|l_{t-1}, a_t)$, and the choice of measurements, $p(x_t|l_t, a_t)$. I do this to simplify the notation as it avoids notational collisions when the choice of actions is from a discrete set.

separated by Bayes law into,

$$I [\mathbf{L}_t | \mathbf{X}_t, a_t] = I [\mathbf{L}_t | a_t] + I [\mathbf{X}_t | \mathbf{L}_t, a_t] - I [\mathbf{X}_t | a_t], \quad (3.5)$$

$$= I [\mathbf{L}_t | a_t] + I_{\text{MI}} [\mathbf{X}_t; \mathbf{L}_t | a_t], \quad (3.6)$$

by Theorem I.2 and Definition I.3. The conditional Shannon information is equivalent to the mutual information and the Shannon information of the predicted latent prior $p(l_t | a_t)$. There are two interpretations of maximising $I [\mathbf{L}_t | \mathbf{X}_t, a_t]$:

1. Maximising the expected predicted posterior Shannon information, or the average “order” in the posterior (Equation 3.4).
2. Maximising the mutual information between the latent and measurement variables, l_t and x_t subject to a penalisation of disorder in the predicted latent prior $p(l_t | a_t)$ (Equation 3.6).

The mutual information is an Infomax measure that is used to assist image classification and registration tasks (Pluim and Maintz, 2003, Maes *et al.*, 2003, Meyer *et al.*, 2006, Peng *et al.*, 2005).

Mutual information is a measure of the independence for the variables x_t and l_t given a choice of action a_t ,

$$I_{\text{MI}} [\mathbf{X}_t; \mathbf{L}_t | a_t] = I [\mathbf{X}_t | \mathbf{L}_t, a_t] - I [\mathbf{X}_t | a_t] \equiv I [\mathbf{L}_t | \mathbf{X}_t, a_t] - I [\mathbf{L}_t | a_t], \quad (3.7)$$

$$= D_{\text{KL}} [p(\mathbf{X}_t, \mathbf{L}_t | a_t) || p(\mathbf{X}_t | a_t) p(\mathbf{L}_t | a_t)], \quad (3.8)$$

which is Definition I.3. This is a measure that maximises the expected divergence between the distributions $p(\mathbf{X}_t, \mathbf{L}_t | a_t)$ and $p(\mathbf{X}_t | a_t) p(\mathbf{L}_t | a_t)$. What this means is that maximising the mutual information chooses an action a_t to maximise the co-dependence between x_t and l_t and so the average divergence between the posterior $p(l_t | x_t, a_t)$ and the prior $p(l_t | a_t)$ is maximised, as

$$I_{\text{MI}} [\mathbf{X}_t; \mathbf{L}_t | a_t] \equiv \mathbb{E}_{p(x_t | a_t)} [D_{\text{KL}} [p(\mathbf{L}_t | x_t, a_t) || p(\mathbf{L}_t | a_t)]], \quad (3.9)$$

which is Theorem I.4. This is a quantification of the expected information gain in taking a measurement x_t for a choice of action a_t between the predicted posterior $p(l_t | x_t, a_t)$ and the prior prediction $p(l_t | a_t)$.

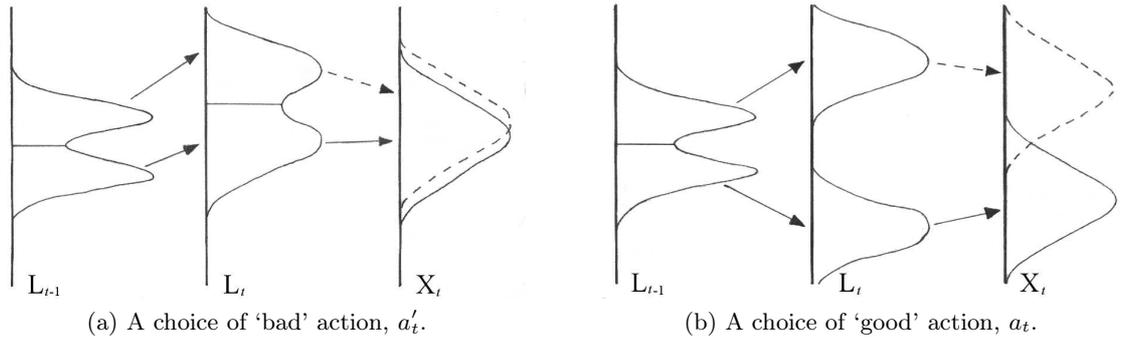


Figure 3.3: Cartoons of the consequence of “good” and “bad” localising actions. Plot (a) depicts a “good” action a_t which causes the two hypotheses upon L_{t-1} to be well separated in X_t . By contrast, in plot (b), a “bad” action a'_t does not lead to separation in X_t . It is important to note that this is only a cartoon and I have implicitly assumed that separation upon L_t leads to separation upon X_t , but this is not necessarily the case. It cannot be assumed that greater separation upon L_t leads to greater separation in X_t , as the distance between any set of hypotheses will be related to the projection of each according to $p(X_t|L_t)$. A measurement x_t for a choice of action a'_t (a) will find it hard to differentiate between either hypothesis upon L_t . However for a choice a_t (b) will find it much easier as the separation of hypotheses on L_t cause a similar separation on X_t .

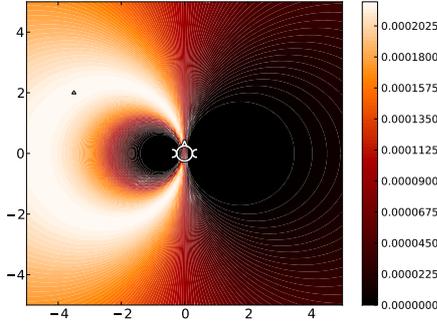
Maximising $I_{\text{MI}}[X_t; L_t|a_t]$ is also finding a tradeoff between minimising $I[X_t|a_t]$ and maximising $I[X_t|L_t, a_t]$, which intuitively is the trade off between

1. a smoother more distributed probability density function $p(x_t|a_t)$, and
2. a preference for sharper conditional distributions $p(x_t|l_t, a_t)$, over the latent prediction $p(l_t|a_t)$.

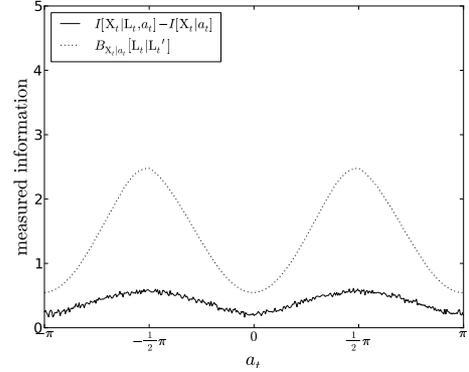
This indicates a preference for actions as sketched in Figure 3.3.

In non-linear models with a continuous measurement variable both Infomax quantities have a degree of difficulty being computed due to the necessity of calculating or estimating the marginal $p(x_t|a_t)$. This difficulty is dealt with in Vergassola *et al.* (2007a) by assuming a simplified marginal distribution $p(x_t|a_t)$. Whereas Porta *et al.* (2005) construct the Shannon information using a sample set to represent the space of $X \cup L$ states.

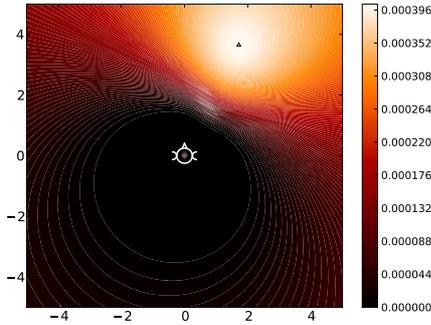
In most cases the Shannon information and the mutual information are not easily computable without careful approximation. These quantities form two related quantitative measures of knowledge that an agent can expect to extract for a choice of action. Instead, I apply a different measure to the problem of maximising an agent’s expected knowledge (my hypothesis, Section 3.2) which is related to the mutual information. A formal proof that maximising my method also maximises the Shannon information is not forthcoming, nor is it likely except in special cases, Figures 3.4 & 3.5. Although, I will show in the



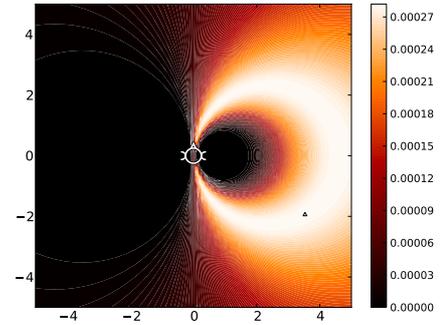
(a) The prior belief.



(b) The information measure contrasted to the bound using the prior belief in plot (a).

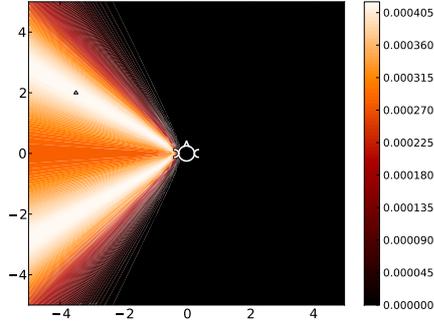


(c) The predicted best actions posterior belief using the prior belief in plot (a).

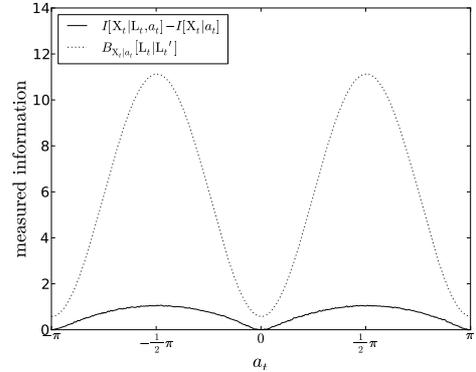


(d) The predicted worst actions posterior belief using the prior belief in plot (a).

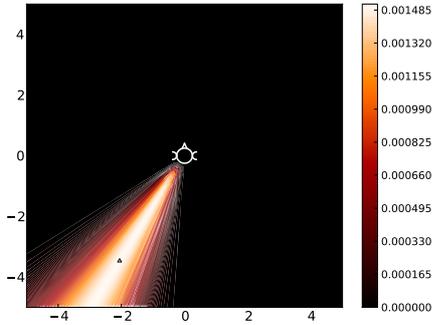
Figure 3.4: A toy example of the best and worst selections of action for inference using an idealised interaural level disparity (ILD). This depicts the selection of the best and worst localising actions using a free field auditory model. Plot (a) depicts a prior belief in location where the source's position is indicated by a white filled triangle. In plot (b) the two information measures ($I[X_t|L_t, a_t] - I[X_t|a_t]$ and its bound $B_{X_t|a_t}[L_t||L'_t]$) are computed across a range of possible head rotations a_t . The contribution to $I[L_t|X_t, a_t]$ of the dynamic component $I[L_t|a_t]$ is neglected as the evolution of l_t is deterministic. It can be clearly seen that both measures of action roughly agree which actions are good and bad. The noise in the calculated $I[X_t|L_t, a_t] - I[X_t|a_t]$ is due to $I[X_t|a_t]$ being approximated by Gibbs sampling. The integral $I[X_t|a_t] = \int dx_t p(x_t|a_t) \log p(x_t|a_t)$, using a sample set to represent the expectation of the log probability over a sample from the belief $p(x_t|a_t)$. This approximates, $p(x_t|a_t) \approx \frac{1}{N} \sum_i \delta(x_t - x_t^{(i)})$, where the $x_t^{(i)}$ sample is drawn by sampling first from the distribution $p(l_t|a_t)$ a set $l_t^{(i)}$, then for each $l_t^{(i)}$ sampling a corresponding $x_t^{(i)}$. This gives an approximation to the integral as $I[X_t|a_t] = \frac{1}{N} \sum_i \log p(x_t^{(i)}|a_t)$. The expected best action (plot (c)) is clearly seen to collapse ambiguity. In contrast the expected worst action (plot (d)) retains much of the ambiguity from the prior.



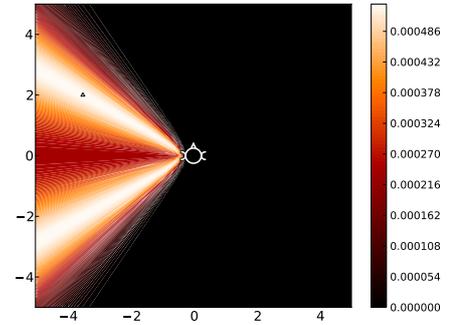
(a) The prior belief.



(b) The information measure contrasted to the bound using the prior belief in plot (a).



(c) The predicted best actions posterior belief using the prior belief in plot (a).



(d) The predicted worst actions posterior belief using the prior belief in plot (a).

Figure 3.5: A toy example of the best and worst selections of action for inference using an idealised interaural time disparity (ITD). This depicts the selection of the best and worst localising actions using a free field auditory model. Plot (a) depicts a prior belief in location where the source's position is indicated by a white filled triangle. In plot (b) the two information measures ($I[X_t|L_t, a_t] - I[X_t|a_t]$ and its bound $B_{X_t|a_t}[L_t||L'_t]$) are computed across a range of possible head rotations a_t . The contribution to $I[L_t|X_t, a_t]$ of the dynamic component $I[L_t|a_t]$ is neglected as the evolution of l_t is deterministic. It can be clearly seen that both measures of action roughly agree which actions are good and bad. The noise in the calculated $I[X_t|L_t, a_t] - I[X_t|a_t]$ is due to $I[X_t|a_t]$ being approximated by Gibbs sampling. The integral $I[X_t|a_t] = \int dx_t p(x_t|a_t) \log p(x_t|a_t)$, using a sample set to represent the expectation of the log probability over a sample from the belief $p(x_t|a_t)$. This approximates, $p(x_t|a_t) \approx \frac{1}{N} \sum_i \delta(x_t - x_t^{(i)})$, where the $x_t^{(i)}$ sample is drawn by sampling first from the distribution $p(l_t|a_t)$ a set $l_t^{(i)}$, then for each $l_t^{(i)}$ sampling a corresponding $x_t^{(i)}$. This gives an approximation to the integral as $I[X_t|a_t] = \frac{1}{N} \sum_i \log p(x_t^{(i)}|a_t)$. The expected best action (plot (c)) is clearly seen to collapse ambiguity. In contrast the expected worst action (plot (d)) retains much of the ambiguity from the prior.

following Section that my method will reduce the expected overlap between the predicted posterior latent beliefs.

3.4 Properties of my hypothesis

In this section I evaluate the utility of my informatic measure (Section 3.2) and its relationship to the Infomax quantities described in the previous section.

Theorem 3.1. *The measure*

$$B_{X_t|a_t} [L_t||L'_t] = \int dl_t dl'_t p(l_t|a_t) p(l'_t|a_t) D_{\text{KL}} [p(X_t|l_t, a_t) || p(X_t|l'_t, a_t)] \quad (3.10)$$

is the upper bound of the mutual information, $I_{\text{MI}} [X_t; L_t|a_t]$, ie

$$I_{\text{MI}} [X_t; L_t|a_t] \leq B_{X_t|a_t} [L_t||L'_t]. \quad (3.11)$$

Proof. Starting with,

$$I_{\text{MI}} [X_t; L_t|a_t] = \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t, l_t|a_t)}{p(x_t|a_t) p(l_t|a_t)}, \quad (3.12)$$

$$= \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t|l_t, a_t)}{\int dl'_t p(x_t, l'_t|a_t)}, \quad (3.13)$$

$$\leq \int dx_t dl_t dl'_t p(x_t, l_t|a_t) p(l'_t|a_t) \log \frac{p(x_t|l_t, a_t)}{p(x_t|l'_t, a_t)}, \quad (3.14)$$

$$= \int dl_t dl'_t p(l_t|a_t) p(l'_t|a_t) D_{\text{KL}} [p(X_t|l_t, a_t) || p(X_t|l'_t, a_t)]. \quad (3.15)$$

Equation 3.15 is $B_{X_t|a_t} [L_t||L'_t]$, and thus completes the proof.⁶ □

Definition 3.2. (Axiom) The quantity

$$D_{\text{KL}} [p(X) || q(X)] + D_{\text{KL}} [q(X) || p(X)] \quad (3.16)$$

is a symmetric measure of the dissimilarity between the two distributions $p(x)$ and $q(x)$, and was originally defined in Kullback and Leibler (1951).

Theorem 3.3. *The following equality holds,*

$$B_{X_t|a_t} [L_t||L'_t] = I_{\text{MI}} [X_t; L_t|a_t] + \varepsilon_B \quad (3.17)$$

⁶Equation 3.14 is the upper bound of Equation 3.13 by Jensen's inequality, Appendix G.

where $\varepsilon_B = D_{\text{KL}} [p(X_t|a_t) p(L_t|a_t) \| p(X_t, L_t|a_t)]$. ε_B quantifies how much greater than the mutual information the quantity $B_{X_t|a_t} [L_t \| L'_t]$ is.

Proof. From the definition

$$B_{X_t|a_t} [L_t \| L'_t] = \int dx_t dl_t dl'_t p(l_t|a_t) p(l'_t|a_t) p(x_t|l_t, a_t) \log \frac{p(x_t|l_t, a_t)}{p(x_t|l'_t, a_t)}, \quad (3.18)$$

$$\begin{aligned} &= \int dx_t dl_t p(x_t, l_t|a_t) \log p(x_t|l_t, a_t) \\ &\quad - \int dx_t dl_t p(x_t|a_t) p(l_t|a_t) \log p(x_t|l_t, a_t), \end{aligned} \quad (3.19)$$

$$\begin{aligned} &= \int dx_t dl_t p(x_t, l_t|a_t) \log p(x_t, l_t|a_t) \\ &\quad - \int dx_t dl_t p(x_t|a_t) p(l_t|a_t) \log p(x_t, l_t|a_t), \end{aligned} \quad (3.20)$$

$$\begin{aligned} &= \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t, l_t|a_t)}{p(x_t|a_t) p(l_t|a_t)} \\ &\quad - \int dx_t dl_t p(x_t|a_t) p(l_t|a_t) \log \frac{p(x_t, l_t|a_t)}{p(x_t|a_t) p(l_t|a_t)}, \end{aligned} \quad (3.21)$$

$$\begin{aligned} &= D_{\text{KL}} [p(X_t, L_t|a_t) \| p(X_t|a_t) p(L_t|a_t)] \\ &\quad + D_{\text{KL}} [p(X_t|a_t) p(L_t|a_t) \| p(X_t, L_t|a_t)], \end{aligned} \quad (3.22)$$

hence proving the identity, as $D_{\text{KL}} [p(X_t, L_t|a_t) \| p(X_t|a_t) p(L_t|a_t)] = I_{\text{MI}} [X_t; L_t|a_t]$.⁷ \square

Remark 3.4. Theorem 3.1 follows from Theorem 3.3 as the KL-divergence is always non-negative.

Remark 3.5. The mutual information is a measure of the statistical independence between the two variables x_t and l_t , where $I_{\text{MI}} [X_t; L_t|a_t] = 0$ indicates total independence of x_t and l_t .

Theorem 3.6. $D_{\text{KL}} [p(X) p(L) \| p(X, L)]$ is a measure of independence between the variables x and l , such that $D_{\text{KL}} [p(X) p(L) \| p(X, L)] = 0$ implies that $x \in X$ and $l \in L$ are independent.

Proof. For x and l to be independent implies that $p(x, l) = p(x) p(l)$ so,

$$D_{\text{KL}} [p(X) p(L) \| p(X, L) = p(X) p(L)] \equiv \int dx dl p(x) p(l) \log \frac{p(x) p(l)}{p(x) p(l)}, \quad (3.23)$$

$$= 0, \quad (3.24)$$

as $\lim_{b \rightarrow a} \log \frac{a}{b} \rightarrow 0$, thus completing the proof. \square

⁷Equation 3.20 is arrived at by the addition of $I[L_t|a_t] - I[L_t|a_t]$ to Equation 3.19. Equation 3.21 is arrived at by the addition of $I[L_t|a_t] + I[X_t|a_t] - I[L_t|a_t] - I[X_t|a_t]$ to Equation 3.20.

Remark 3.7. Conversely when $D_{\text{KL}} [p(\mathbf{X}_t|a_t) p(\mathbf{L}_t|a_t) \| p(\mathbf{X}_t, \mathbf{L}_t|a_t)] \neq 0$ it indicates that $p(x_t, l_t|a_t) \neq p(x_t|a_t) p(l_t|a_t)$.

Remark 3.8. For a selected a_t , $D_{\text{KL}} [p(\mathbf{X}_t|a_t) p(\mathbf{L}_t|a_t) \| p(\mathbf{X}_t, \mathbf{L}_t|a_t)]$ can similarly be seen as a measure of the independence of x_t and l_t similarly to $I_{\text{MI}} [\mathbf{X}_t; \mathbf{L}_t|a_t]$.

Theorem 3.9. *The bound $B_{\mathbf{X}_t|a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ can be re-expressed as the quantity*

$$B_{\mathbf{L}_t|a_t} [\mathbf{X}_t \| \mathbf{X}'_t] = \int dx_t dx'_t p(x_t|a_t) p(x'_t|a_t) D_{\text{KL}} [p(\mathbf{L}_t|x_t, a_t) \| p(\mathbf{L}_t|x'_t, a_t)]. \quad (3.25)$$

Proof. Starting from the definition

$$B_{\mathbf{X}_t|a_t} [\mathbf{L}_t \| \mathbf{L}'_t] = \int dl_t dl'_t p(l_t|a_t) p(l'_t|a_t) D_{\text{KL}} [p(\mathbf{X}_t|l_t, a_t) \| p(\mathbf{X}_t|l'_t, a_t)], \quad (3.26)$$

$$\begin{aligned} &= \int dx_t dl_t p(x_t, l_t|a_t) \log p(x_t|l_t, a_t) \\ &\quad - \int dx_t dl_t p(x_t|a_t) p(l_t|a_t) \log p(x_t|l_t, a_t), \end{aligned} \quad (3.27)$$

$$\begin{aligned} &= \int dx_t dl_t p(x_t, l_t|a_t) \log p(l_t|x_t, a_t) \\ &\quad - \int dx_t dl_t p(x_t|a_t) p(l_t|a_t) \log p(l_t|x_t, a_t), \end{aligned} \quad (3.28)$$

$$= \int dl_t dl'_t p(x_t|a_t) p(x'_t|a_t) D_{\text{KL}} [p(\mathbf{L}_t|x_t, a_t) \| p(\mathbf{L}_t|x'_t, a_t)] \quad (3.29)$$

which completes the proof.⁸ □

I interpret the measure $B_{\mathbf{X}_t|a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ as maximising the average divergence between the measurement likelihood of every pair of hypotheses. Using Theorem 3.9 shows that this is the same as maximising the average divergence between projected posterior beliefs. As a final statement from Theorem 3.6, I can state that like the mutual information $I_{\text{MI}} [\mathbf{X}_t; \mathbf{L}_t|a_t]$, $B_{\mathbf{X}_t|a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ is a measure of statistical independence between x_t and l_t for a given action a_t . However, $B_{\mathbf{X}_t|a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ differs from mutual information $I_{\text{MI}} [\mathbf{X}_t; \mathbf{L}_t|a_t]$ in the consequent preferences for an a_t .

As $B_{\mathbf{X}_t|a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ is closely related to the mutual information by Theorem 3.3, we can see it is a similar but distinct measure of the independence of the variables x_t and l_t for a choice of a_t . Mutual information has been used in a number of Infomax applications such as image registration (Pluim and Maintz, 2003, Maes *et al.*, 2003, Meyer *et al.*, 2006, Peng *et al.*, 2005).

⁸Equation 3.28 is arrived at by the addition of $I[\mathbf{L}_t|a_t] + I[\mathbf{X}_t|a_t] - I[\mathbf{L}_t, \mathbf{X}_t|a_t]$ to Equation 3.27.

Theorem 3.10. *That maximising $B_{X_t|a_t} [L_t||L'_t]$ will maximise the expected dissimilarity between every pair of predicted latent posterior beliefs $p(l_t|x_t, a_t)$ and $p(l_t|x'_t, a_t)$. Every possible pair of measurements are denoted as x_t and x'_t and have the respective weights $p(x_t|a_t)$ and $p(x'_t|a_t)$.*

Proof. Starting with Theorem 3.9,

$$2B_{X_t|a_t} [L_t||L'_t] = 2B_{L_t|a_t} [X_t||X'_t], \quad (3.30)$$

$$= 2 \int dx_t dx'_t p(x_t|a_t) p(x'_t|a_t) D_{\text{KL}} [p(L_t|x_t, a_t) || p(L_t|x'_t, a_t)], \quad (3.31)$$

$$= \int dx_t dx'_t p(x_t|a_t) p(x'_t|a_t) \left\{ D_{\text{KL}} [p(L_t|x_t, a_t) || p(L_t|x'_t, a_t)] \right. \\ \left. + D_{\text{KL}} [p(L_t|x'_t, a_t) || p(L_t|x_t, a_t)] \right\}, \quad (3.32)$$

as $\int dx_t p(x_t|a_t) f(x_t) = \int dx'_t p(x'_t|a_t) f(x'_t)$, Equation 3.32 leading to

$$B_{X_t|a_t} [L_t||L'_t] = \frac{1}{2} \int dx_t dx'_t p(x_t|a_t) p(x'_t|a_t) \left\{ D_{\text{KL}} [p(L_t|x_t, a_t) || p(L_t|x'_t, a_t)] \right. \\ \left. + D_{\text{KL}} [p(L_t|x'_t, a_t) || p(L_t|x_t, a_t)] \right\}, \quad (3.33)$$

which by Definition 3.2 is half the average dissimilarity between the posterior beliefs over every pair of measurements x_t and x'_t . Hence completing the proof. \square

Corollary 3.11. *That*

$$B_{X_t|a_t} [L_t||L'_t] = \frac{1}{2} \int dl_t dl'_t p(l_t|a_t) p(l'_t|a_t) \left\{ D_{\text{KL}} [p(X_t|l_t, a_t) || p(X_t|l'_t, a_t)] \right. \\ \left. + D_{\text{KL}} [p(X_t|l'_t, a_t) || p(X_t|l_t, a_t)] \right\}, \quad (3.34)$$

is the average dissimilarity (Definition 3.2) between measurement likelihood beliefs over every pair of latent hypotheses, l_t and l'_t .

Corollary 3.12. *The quantity*

$$\frac{1}{2} \int dx_t dx'_t p(x_t|a_t) p(x'_t|a_t) \left\{ D_{\text{KL}} [p(L_t|x_t, a_t) || p(L_t|x'_t, a_t)] + D_{\text{KL}} [p(L_t|x'_t, a_t) || p(L_t|x_t, a_t)] \right\} \quad (3.35)$$

can be defined instead as the expectation over the unique pairs of measurement $\{x_t, x'_t\}$ to

be

$$\int d\{x_t, x'_t\} p(\{x_t, x'_t\} | a_t) \left\{ D_{\text{KL}} [p(\mathbf{L}_t | x_t, a_t) \| p(\mathbf{L}_t | x'_t, a_t)] + D_{\text{KL}} [p(\mathbf{L}_t | x'_t, a_t) \| p(\mathbf{L}_t | x_t, a_t)] \right\} \quad (3.36)$$

where $p(\{x_t, x'_t\} | a_t) = \frac{1}{2} p(x_t | a_t) p(x'_t | a_t)$. Hence we can say that my measure is the average dissimilarity between every unique pair of predictions for the latent posterior belief.

Theorem 3.10 shows that the posterior bound $B_{\mathbf{L}_t | a_t} [\mathbf{X}_t \| \mathbf{X}'_t]$ may be re-expressed to be an average of the addition of the two KL-divergences,

$$D_{\text{KL}} [p(\mathbf{L}_t | x_t, a_t) \| p(\mathbf{L}_t | x'_t, a_t)] + D_{\text{KL}} [p(\mathbf{L}_t | x'_t, a_t) \| p(\mathbf{L}_t | x_t, a_t)]. \quad (3.37)$$

Therefore this is a symmetric premetric measure of the dissimilarity between the distributions $p(l_t | x_t, a_t)$ and $p(l_t | x'_t, a_t)$. The consequence of this and Theorem 3.9 is to express $B_{\mathbf{X}_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ as maximising the average dissimilarity between latent posterior beliefs, giving the a_t expected to distinguish most between the predicted future measurements.

The preferences of my measure $B_{\mathbf{X}_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ differs from that of Shannon information $I[\mathbf{L}_t | \mathbf{X}_t, a_t]$ and mutual information $I_{\text{MI}}[\mathbf{X}_t; \mathbf{L}_t | a_t]$. Intuitively $I[\mathbf{L}_t | \mathbf{X}_t, a_t]$ selects the action with least uncertainty. In the simplest unimodal case this is typically the action expected to produce the sharpest posterior $p(l_t | x_t, a_t)$ from a measurement. In contrast $B_{\mathbf{X}_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ will by Theorem 3.10 select the action expected to maximise the dissimilarity between each pair of predicted *a posteriori* beliefs, $p(l_t | x_t, a_t)$ and $p(l_t | x'_t, a_t)$ (averaged over the expected measurements x_t and x'_t). This is equivalent to selecting an action a_t to maximise the expected dissimilarity between each possible pair of posterior beliefs $p(l_t | x_t, a_t)$ and $p(l_t | x'_t, a_t)$, weighted according to each unique pair of measurements x_t and x'_t by a factor $\frac{1}{2} p(x_t | a_t) p(x'_t | a_t)$. Hence, while $I[\mathbf{L}_t | \mathbf{X}_t, a_t]$, $I_{\text{MI}}[\mathbf{X}_t; \mathbf{L}_t | a_t]$ and $B_{\mathbf{X}_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ will select a_t to increase our knowledge of the latent space \mathbf{L}_t , the actions will differ due to each measure's description of useful knowledge.

To summarise:

- Maximising $B_{\mathbf{X}_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$ selects an action a_t to maximise the statistical dependence between x_t and l_t . This is also equivalent to maximising over all possible pairs of expected measurements the consequent posterior beliefs in l_t (Theorem 3.10). This indicates the measure will select an action that most distinguishes between all possible posterior predictions.

- Maximising $I_{\text{MI}} [X_t; L_t|a_t]$ selects the action a_t (maximising the mutual information) that will maximise the dependence between the variables l_t and x_t , cf $B_{X_t|a_t} [L_t||L'_t]$. By Theorem 3.3 this relates as

$$B_{X_t|a_t} [L_t||L'_t] = I_{\text{MI}} [X_t; L_t|a_t] + D_{\text{KL}} [p(X_t|a_t) p(L_t|a_t) || p(X_t, L_t|a_t)] \quad (3.38)$$

where the second term on the right-hand-side of the equation is also a measure of the statistical dependence between x_t and l_t for a choice of action a_t .

- Maximising $I [L_t|X_t, a_t]$ is the maximisation of the expected latent posterior's Shannon information

$$I [L_t|X_t, a_t] = \mathbb{E}_{p(x_t|a_t)} [I [L_t|x_t, a_t]], \quad (3.39)$$

by Theorem I.1. By Theorems I.2 & I.3, this quantity is related to the mutual information as

$$I [L_t|X_t, a_t] = I [L_t|a_t] + I_{\text{MI}} [X_t; L_t|a_t]. \quad (3.40)$$

Due to Theorem 3.1, it is also related to my measure through the inequality

$$I [L_t|X_t, a_t] \leq I [L_t|a_t] + B_{X_t|a_t} [L_t||L'_t]. \quad (3.41)$$

From these properties I argue that $B_{X_t|a_t} [L_t||L'_t]$ is a reasonable substitute for Shannon information or the mutual information as an information quantity. If other quantities are being considered it can also be constructed as an upper bound. This allows it to include terms to penalise other features of the belief model (Figure 3.2), such as $I [L_t|a_t]$ which penalises higher entropy in the latent prediction $p(l_t|a_t)$.

My measure is the average dissimilarity between possible posterior beliefs. This is not necessarily the least entropy but the greatest dependence between measurable and latent variable (respectively x_t and l_t conditional upon a_t).

3.5 Algorithmic complexity

Any analysis of the time complexity of an informatic measure is by necessity dependent upon the agent's model of the environment. In this section I will contrast the application of my method to the problem defined in Porta *et al.* (2005), to which the authors applied an Infomax algorithm. This will illustrate the time complexity of my approach compared

to state-of-the-art Infomax solutions.

In Appendix K I have, for partially observable Markov decision processes (POMDPs), contrasted the time complexity of Infomax ($I[L_t|X_t, a_t]$) to my measure ($B_{X_t|a_t}[L_t||L'_t]$). Both of these informatic quantities scale equivalently. This scaling is linear with the number of latent states, measurement states and choices of action. However, for other problems my method can be defined to scale more favourably.

An example of such a problem is based on Gaussian processes. Fox *et al.* (1998) considered the problem of mobile robot localisation with sonar and radar sensors. Porta *et al.* (2005) applied Fox *et al.*'s Infomax method to a stereo vision robot. It is the algorithmic time complexity of Porta *et al.*'s algorithm that I shall contrast with my method. Key features of their agent's model of the environment and its sensors were:

1. As this is a visual problem, the dimensionality of any raw measurement is large. A PCA is applied to significantly reduce the dimensionality of any measurement calculations, making a measurement \mathbf{x} have D_x dimensions.
2. The union of the two complete sets representing the measurement and latent variables is approximated by a joint set $S \approx X \cup L$ which is used to construct a sample X_{a_t} and $p(\mathbf{x}|\mathbf{x}_{0:t-1}, a_{0:t}) \forall \mathbf{x} \in X_{a_t}$ for a latent prediction $p(l_t|\mathbf{x}_{0:t-1}, a_{0:t})$.
3. The latent prediction is constructed deterministically using the approach of Fox *et al.* (1998). This leads to the approximation

$$p(l_t|\mathbf{x}_{0:t-1}, a_{0:t}) \approx \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta\left(l_t - l_{t|t-1}^{(i)}(a_t)\right), \quad (3.42)$$

where $l_{t|t-1}^{(i)}(a_t) = g\left(l_{t-1|t-1}^{(i)}, a_t\right)$ from the forward process, and $g(\cdot, \cdot)$ represents the deterministic aspect of the latent state model $p(l_t|l_{t-1}, a_t)$.

4. A further approximation made use of a nearest-neighbour approximation for each point x in S . The likelihood of the latent point l is given by

$$p(\mathbf{x}|l, a_t) = \sum_{j=1}^J \lambda_j \phi(l; l_j, \sigma_j), \quad (3.43)$$

with the collection $\{\lambda_j, l_j, \sigma_j : j \in [1, J]\}$ corresponding to the measurement \mathbf{x} from the set S .

The general approach taken by Fox *et al.*, to estimate the conditional Shannon information,

is still state-of-the-art (Porta *et al.*, 2005, Thrun *et al.*, 2006). However, the details of a particular system imposes the specific extensions necessary for computation (e.g. Porta *et al.*, 2005 and also the next Chapter for a chaotic dynamical system).

In Porta *et al.* (2005), the author's construct $I[L_t|X_t, a_t]$ by starting with a sample set $S = \{(l^{(i)}, \mathbf{x}^{(i)}) : i \in [1, N_S]\}$. The implicit assumption is that the set S represents a uniform sample. This is sampled to construct action dependent subsets of unique $\mathbf{x}^{(i)}$'s to form a set \tilde{X}_{a_t} with a distribution $p(\mathbf{x}^{(i)}|a_t)$ being the sum of the weights for those samples $l_{t|t-1}^{(j)}(a_t)$ closest to $l^{(i)}$. This leads to the Shannon information being approximated as

$$I[L_t|X_t, a_t] \approx \sum_{\mathbf{x} \in \tilde{X}_{a_t}} \sum_{j=1}^N \pi_{t-1|t-1}^{(i)} p(\mathbf{x}|l_{t|t-1}^{(j)}(a_t), a_t) \log \frac{\pi_{t|t-1}^{(j)} p(\mathbf{x}|l_{t|t-1}^{(j)}(a_t), a_t)}{p(\mathbf{x}|a_t)}. \quad (3.44)$$

The marginal belief in a measurement is approximated as $p(\mathbf{x}|a_t) \approx \sum_{k=1}^K \pi_t^{(i_k)}$, with the i_k 's representing the particle indexes associated with the measurement \mathbf{x} .

Porta *et al.*'s algorithm is illustrated in Algorithm 3.2 and has a time complexity of $O(N_a N^2 J)$. This is because the size of X_{a_t} is at most N (the number of particles), and J is the number of nearest-neighbours to compute the sensor model's likelihood.

There are two methods for applying my method to Porta *et al.*'s problem. The first is by not using the nearest neighbour approach. If the likelihood of the PCA data can be represented as a Gaussian of the form

$$p(\mathbf{x}_t|l_t, a_t) = \phi(\mathbf{x}_t; \boldsymbol{\mu}(l_t, a_t), \Sigma(l_t, a_t)), \quad (3.45)$$

then the KL-divergence is both analytic and separable, Appendix D. To calculate my approach, using the identities $\boldsymbol{\mu}_{a_t}^{(i)} = \boldsymbol{\mu}(l_{t|t-1}^{(i)}(a_t))$ and $\Sigma_{a_t}^{(i)} = \Sigma(l_{t|t-1}^{(i)}(a_t))$ for compactness, leads to the identity

$$\begin{aligned} B_{X_t|a_t}[L_t||L'_t] &= \frac{1}{2} \left\{ \text{tr} \left[\left(\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \Sigma_{a_t}^{(i)-1} \right) \left(\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \Sigma_{a_t}^{(i)} \right) \right] - D_x \right. \\ &\quad \left. + \text{tr} \left[\left(\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \boldsymbol{\mu}_{a_t}^{(i)T} \boldsymbol{\mu}_{a_t}^{(i)} \right) \left(\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \Sigma_{a_t}^{(i)-1} \right) \right] \right. \\ &\quad \left. - 2 \left(\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \boldsymbol{\mu}_{a_t}^{(i)} \right)^T \left(\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \Sigma_{a_t}^{(i)-1} \boldsymbol{\mu}_{a_t}^{(i)} \right) \right. \\ &\quad \left. + \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \boldsymbol{\mu}_{a_t}^{(i)T} \Sigma_{a_t}^{(i)-1} \boldsymbol{\mu}_{a_t}^{(i)} \right\}. \quad (3.46) \end{aligned}$$

Algorithm 3.1 The generic particle filtering algorithm. This algorithm is a simplification of the examples from Isard and Blake (1998), Murphy and Russell (2001), Doucet *et al.* (2000).

1. Selection step,

- sample N samples from $l_{t-1|t-1}^{(i)}$ according to the importance distribution $\pi_{t-1|t-1}^{(i)}$ to obtain N random samples $l_{t-1|t-1}^{*(i)}$ approximating the distribution $p(l_t|\mathbf{x}_{0:t})$.

2. Sequential importance sampling step,

- for $i \in [1, N]$, sample

$$l_{t|t-1}^{(i)} \sim p\left(l_t | l_{0:t-1}^{*(i)}, \mathbf{x}_{0:t-1}, a_{0:t}\right),$$

- for $i \in [1, N]$, evaluate the importance of the proposed trajectory according to

$$\pi_{t|t}^{(i)} \propto p\left(x_t | l_{t|t-1}^{(i)}, \mathbf{x}_{0:t-1}, a_{0:t}\right),$$

where these weights are normalised, $\sum_{i=1}^N \pi_{t|t}^{(i)} = 1$.

3. Set $t = t + 1$, and goto 1.

To contrast the expectation over the latent posterior belief of a function $f(l)$ is

$$\mathbb{E}_{p(l_t|\mathbf{x}_{0:t}, a_{0:t})} [f(l_t)] = \sum_{i=1}^N \pi_{t|t}^{(i)} f\left(l_{t|t}^{(i)}\right),$$

similarly the expectation for the prediction is

$$\mathbb{E}_{p(l_t|\mathbf{x}_{0:t-1}, a_{0:t})} [f(l_t)] = \sum_{i=1}^N \pi_{t|t-1}^{(i)} f\left(l_{t|t-1}^{(i)}\right),$$

where $\pi_{t|t-1}^{(i)} = \frac{1}{N}$. The expectation for a prediction can be approximated deterministically as

$$\begin{aligned} p(l_t|\mathbf{x}_{0:t-1}, a_{0:t}) &= \mathbb{E}_{p(l_{t-1}|\mathbf{x}_{0:t-1}, a_{0:t-1})} [\delta(l_t - g(l_{t-1}, a_t))], \\ &\approx \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta\left(l_t - l_{t|t-1}^{(i)}(a_t)\right), \end{aligned}$$

for the dynamical process

$$l_t \sim g(l_{t-1}, a_t) + w_t$$

where w_t represents a noise term indicating the uncertainty in the system states evolution and $g(l_{t-1}, a_t)$ is the deterministic component of the system states evolution (Porta *et al.*, 2003, 2005, Fox *et al.*, 1998).

Algorithm 3.2 The algorithm used by Porta *et al.* to calculate Infomax for the robotic situation as described in Porta *et al.* (2003, 2005). The sensor model is constructed as using a nearest-neighbour based approach. This scales with a time complexity of $O(N_a N^2 J)$, where there are N particles representing the latent belief $p(l_t | \mathbf{x}_{0:t-1}, a_{0:t})$, J basis functions representing the nearest-neighbours, and N_a possible choices of action a_t . The storage scales as $O(N_S J)$ where N_S is the size of the set S , this represents number of the nearest-neighbours. The latent belief is $p(l_t | \mathbf{x}_{0:t-1}, a_{0:t}) \approx \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta(l_{t|t-1}^{(i)}(a_t) - l_t)$ and is constructed using Algorithm 3.1.

Calculation of Porta *et al.*'s method,

1. for each a_t ,

$$X_{a_t} = \emptyset$$

- (a) for each $i \in [1, N]$

select $(\mathbf{x}, l) \in S$ with minimum $\|l - l_{t|t-1}^{(i)}(a_t)\|$

- i. if $\mathbf{x} \in X_{a_t}$ then

$$p(\mathbf{x}|a_t) = p(\mathbf{x}|a_t) + \pi_{t-1|t-1}^{(i)}$$

- ii. else

$$X_{a_t} = X_{a_t} \cup \{\mathbf{x}\}$$

$$p(\mathbf{x}|a_t) = \pi_{t-1|t-1}^{(i)}$$

$$h = 0$$

- (a) for each $\mathbf{x} \in X_{a_t}$

- i. for each $i \in [1, N]$

$$g = \pi_{t-1|t-1}^{(i)} \sum_{j=1}^J \lambda_j \phi(l_{t|t-1}^{(i)}(a_t); l_j, \sigma_j)$$

$$h = h + g \log\left(\frac{g}{p(\mathbf{x}|a_t)}\right)$$

$$I[\mathbf{L}_t | X_t, a_t] = h$$

2. select the a_t that maximises $I[\mathbf{L}_t | X_t, a_t]$.
-

Algorithm 3.3 An implementation of my information measure (Equation 3.3) for the robotic situation described in Porta *et al.* (2003, 2005). The observable variables are collected together as \mathbf{x}_t , the hidden variables are collected as l_t , and the decision variables are collected as a_t . This is if the measurement model is Gaussian, $p(\mathbf{x}_t|l_t, a_t) = \phi(\mathbf{x}_t; \boldsymbol{\mu}(l_t), \Sigma(l_t))$. This scales as $O(N_a N D_x^2)$, where there are N particles representing the latent belief $p(l_t|\mathbf{x}_{0:t-1}, a_{0:t})$ for N_a possible choices of action a_t and D_x is the number of dimensions in the measurable x_t . The storage scales as $O(D_x^2)$. The latent belief is $p(l_t|\mathbf{x}_{0:t-1}, a_{0:t}) \approx \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta(l_{t|t-1}^{(i)}(a_t) - l_t)$ and is constructed using Algorithm 3.1.

Calculation of my method,

1. for each a_t ,

$$H_1 = 0_{D_x \times D_x}, H_2 = 0_{D_x \times D_x}, \mathbf{h}_3 = 0_{D_x}, H_4 = 0_{D_x \times D_x}, \mathbf{h}_5 = 0_{D_x}, h_6 = 0$$

- (a) for each $i \in [1, N]$

$$\Sigma^{-1} = \left[\Sigma \left(l_{t|t-1}^{(i)}(a_t), a_t \right) \right]^{-1}, \Sigma = \Sigma \left(l_{t|t-1}^{(i)}(a_t), a_t \right) \text{ and } \boldsymbol{\mu} = \boldsymbol{\mu} \left(l_{t|t-1}^{(i)}(a_t), a_t \right)$$

$$H_1 = H_1 + \pi_{t-1|t-1}^{(i)} \Sigma$$

$$H_2 = H_2 + \pi_{t-1|t-1}^{(i)} \Sigma^{-1}$$

$$\mathbf{h}_3 = \mathbf{h}_3 + \pi_{t-1|t-1}^{(i)} \boldsymbol{\mu}$$

$$H_4 = H_4 + \pi_{t-1|t-1}^{(i)} \boldsymbol{\mu} \boldsymbol{\mu}^T$$

$$\mathbf{h}_5 = \mathbf{h}_5 + \pi_{t-1|t-1}^{(i)} \Sigma^{-1} \boldsymbol{\mu}$$

$$h_6 = h_6 + \pi_{t-1|t-1}^{(i)} \boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}$$

$$B_{X_t|a_t} [L_t \| L'_t] = \frac{1}{2} (\text{tr} [H_2 H_1] - D_x + \text{tr} [H_4 H_2] - 2\mathbf{h}_5^T \mathbf{h}_5 + h_6)$$

2. select the a_t that maximises $B_{X_t|a_t} [L_t \| L'_t]$.
-

Algorithm 3.4 An alternative implementation of my information measure (Equation 3.3) for a robotic situation described in Porta *et al.* (2003, 2005). The sensor model is constructed as using a nearest-neighbour based approach. This scales as $O(N_a N^2 J)$, where there are N particles representing the latent belief $p(l_t | \mathbf{x}_{0:t-1}, a_{0:t})$, J basis functions representing the nearest-neighbours, and N_a possible choices of action a_t . The storage scales as $O(N_S J)$ where N_S is the size of the set S , this represents the number of the nearest-neighbours. The latent belief is $p(l_t | \mathbf{x}_{0:t-1}, a_{0:t}) \approx \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta(l_{t|t-1}^{(i)}(a_t) - l_t)$ and is constructed using Algorithm 3.1.

Calculation of my method,

1. for each a_t ,

$$X_{a_t} = \emptyset$$

- (a) for each $i \in [1, N]$

select $(\mathbf{x}, l) \in S$ with minimum $\|l - l_{t|t-1}^{(i)}(a_t)\|$

- i. if $\mathbf{x} \notin X_{a_t}$ then
 $X_{a_t} = X_{a_t} \cup \{\mathbf{x}\}$

$$h_1 = 0, h_2 = 0, h_3 = 0$$

- (a) for each $\mathbf{x} \in X_{a_t}$

- i. for each $i \in [1, N]$

$$g = \sum_{j=1}^J \lambda_j \phi(l_{t|t-1}^{(i)}(a_t); l_j, \sigma_j)$$

$$h_1 = h_1 + \pi_{t-1|t-1}^{(i)} g \log g$$

$$h_2 = h_2 + \pi_{t-1|t-1}^{(i)} g$$

$$h_3 = h_3 + \pi_{t-1|t-1}^{(i)} \log g$$

$$B_{X_t|a_t} [L_t \| L'_t] = h_1 - h_2 h_3$$

2. select the a_t that maximises $B_{X_t|a_t} [L_t \| L'_t]$.
-

This quantity can be calculated using Algorithm 3.3, and scales as $O(N_a N)$ for a selection of one from N_a actions with N particles. The second is by using the nearest-neighbour approach for calculating the likelihood $p(\mathbf{x} | l_t, a_t)$. This gives a solution of the form,

$$B_{X_t|a_t} [L_t \| L'_t] = \sum_{\mathbf{x} \in X_{a_t}} \left\{ \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} p(\mathbf{x} | l_{t|t-1}^{(i)}, a_t) \log p(\mathbf{x} | l_{t|t-1}^{(i)}, a_t) - \left[\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} p(\mathbf{x} | l_{t|t-1}^{(i)}, a_t) \right] \left[\sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \log p(\mathbf{x} | l_{t|t-1}^{(i)}, a_t) \right] \right\} \quad (3.47)$$

where $p(\mathbf{x} | l_t, a_t) = \sum_{j=1}^J \lambda_j \phi(l_t; l_j, \sigma_j)$, Equation 3.43. This quantity can be calculated using Algorithm 3.4. This approach scales with a time complexity of $O(N_a N^2 J)$ as the size of X_{a_t} is at most N (the number of particles), and J is the number of nearest-neighbours to compute the sensor model's likelihood. Like Porta *et al.*'s Infomax algorithm, Algorithm 3.4 can be modified such that it has a time complexity of $O(N_a N^2)$ by precom-

putting the sensor model and replacing the latent point $l_t^{(i)}(a_t)$ with the closest point in the training set, making $p(\mathbf{x}_t|l_t, a_t)$ fully precomputed (Porta *et al.*, 2005). If the number of principle components is favourably small then, for my method, there is an advantage in time complexity when comparing Algorithm 3.3 to Algorithm 3.4, these have respective time complexities of $O(N_a N D_x^2)$ and $O(N_a N^2)$ (or $O(N_a N^2 J)$).

More generally, if the likelihood $p(\mathbf{x}_t|l_t, a_t)$ produces an analytic KL-divergence, there is a short cut in computing $B_{\mathbf{X}_t|a_t}[\mathbf{L}_t||\mathbf{L}'_t]$. This is due to the structure of the measurement model $p(\mathbf{x}_t|l_t, a_t)$. For instance, if it is Gaussian or Poisson then the KL-divergence takes the form

$$D_{\text{KL}}[p(\mathbf{X}_t|l_t, a_t)||p(\mathbf{X}_t|l'_t, a_t)] = h_0(l_t, a_t) + \sum_{k=1}^K h_k(l_t, a_t) h_{K+k}(l'_t, a_t) + h_{2K+1}(l'_t, a_t) : \quad (3.48)$$

allowing my measure to be separated and optimised to give

$$B_{\mathbf{X}_t|a_t}[\mathbf{L}_t||\mathbf{L}'_t] = \sum_{i=1}^N \pi_{t|t-1}^{(i)} \left(h_0(l_{t|t-1}^{(i)}, a_t) + h_{2K+1}(l_{t|t-1}^{(i)}, a_t) \right) + \sum_{k=1}^K \left[\sum_{i=1}^N \pi_{t|t-1}^{(i)} h_k(l_{t|t-1}^{(i)}, a_t) \right] \left[\sum_{i=1}^N \pi_{t|t-1}^{(i)} h_{K+k}(l_{t|t-1}^{(i)}, a_t) \right], \quad (3.49)$$

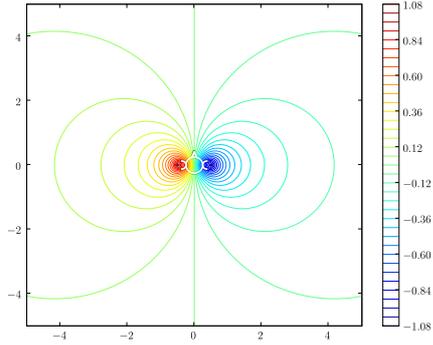
which scales as $O(N_a N K)$ for example Algorithm 3.3. For a Gaussian measurement model K is a polynomial of the dimensionality of \mathbf{x}_t as the sum over K represents 2 matrix multiplications. If the cost of computing the K variables is smaller than the cost of computing $N \times J$ nearest neighbours it makes an approach like Algorithm 3.3 viable. In contrast, if

$$D_{\text{KL}}[p(\mathbf{X}_t|l_t, a_t)||p(\mathbf{X}_t|l'_t, a_t)] = g_1(l_t, a_t) + g_2(l_t, l'_t, a_t) + g_3(l_t, a_t), \quad (3.50)$$

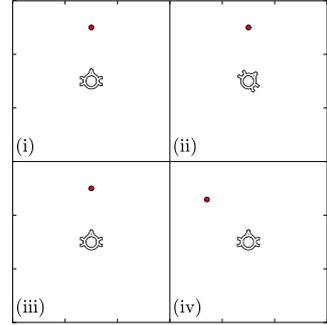
then

$$B_{\mathbf{X}_t|a_t}[\mathbf{L}_t||\mathbf{L}'_t] = \sum_{i=1}^N \pi_{t|t-1}^{(i)} \left(g_1(l_{t|t-1}^{(i)}, a_t) + g_3(l_{t|t-1}^{(i)}, a_t) \right) + \sum_{i=1}^N \sum_{i'=1}^N \pi_{t|t-1}^{(i)} \pi_{t|t-1}^{(i')} g_2(l_{t|t-1}^{(i)}, l_{t|t-1}^{(i')}, a_t) \quad (3.51)$$

which will scale as $O(N_a N^2)$ due to the double sum over i and i' . Therefore, if the KL-divergence of the measurement processes model is analytic then any computation of Equation 3.3 will scale as either $O(N_a N)$ or $O(N_a N^2)$. This indicates a potential com-



(a) ILD isoclines.



(b) Head rotations, (i) and (ii) are world centric, (iii) and (iv) are observer centric.

Figure 3.6: Cartoons indicating a toy auditory head rotation problem. A contour line plot (plot (a)) of the free-field interaural level disparity (ILD) with respect to the latent space, $\mathbf{l} = \{l_x, l_y\}$. A figure (plot (b)) indicating the shift in perspective caused by a head rotation and the consequent impact with respect to the latent space, \mathbf{l} , for the observer (plot (b,iii) to plot (b,iv)) and observing the observer (plot (b,i) to plot (b,ii)).

putational advantage when compared to sampling approaches, for instance Porta *et al.* (2005). Though this may be lost due to the addition of other penalisation terms such as $I[\mathbf{L}_t | \mathbf{x}_{0:t}, a_{0:t}]$ in the next section.

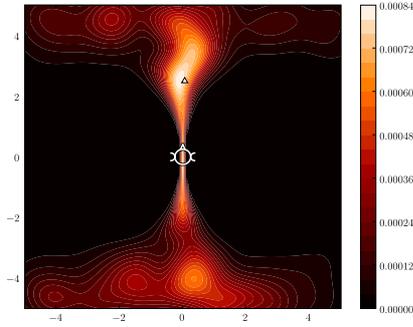
3.6 Simulated head rotations for active sound source localisation

In this section I apply my action selection framework to a concrete but simplified non-linear case. I approach this using a free-field simulation of sound propagation with an observable space of interaural level disparity⁹ (Figure 3.6 (a)), and with the possible actions only being head rotations (Pettorossi *et al.*, 2005). I selected this problem as my illustrative example due to my familiarity with sound perception and the ease with which non-practitioners can understand it. The tracked posterior belief is computed using a particle filter (Doucet *et al.*, 2001, Ristic *et al.*, 2004c). While this case is highly constrained, it is still a concrete example of the applicability of my framework to a localisation problem.

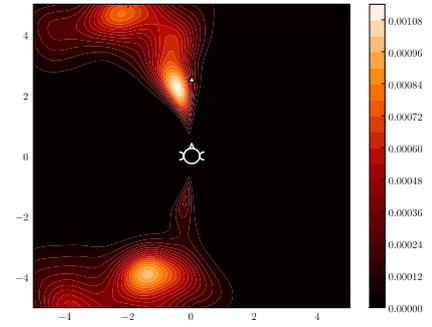
The model is constructed according to two distributions that are needed for the Bayesian filtering (Algorithm 3.1 and Appendix H). Firstly from the distribution,

$$p(\mathbf{l}_t | \mathbf{l}_{t-1}, a_t) = \phi(\mathbf{l}_t; R(a_t) \mathbf{l}_{t-1}, \sigma_l), \quad (3.52)$$

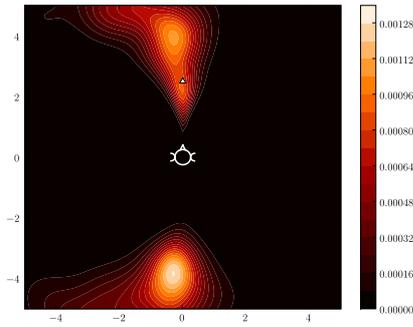
⁹Interaural level disparity is defined as the difference in perceived loudness between either ear for a sound event.



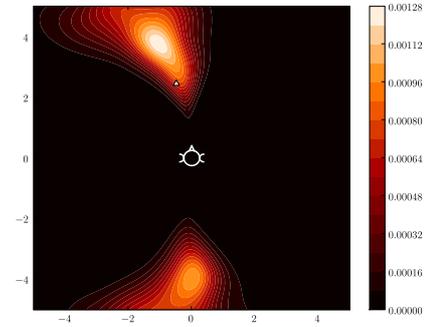
(a) $t = 1$.



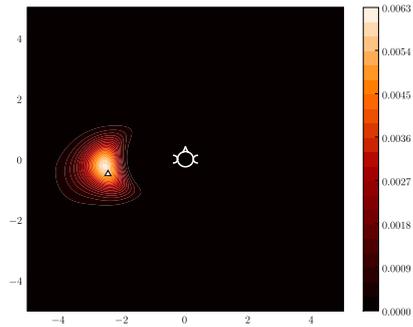
(b) $t = 2$.



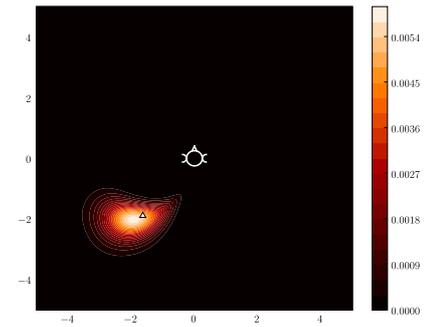
(c) $t = 3$.



(d) $t = 4$.



(e) $t = 5$.



(f) $t = 6$.

Figure 3.7: Contours of egocentric location belief for a sequence of randomly selected actions. The hue in the above figures indicates the belief $p(\mathbf{l}_t | x_{1:t}, a_{1:t})$. The true location is indicated by a triangle outlined with black and white filled. The initial prior $p(\mathbf{l}_0)$ is uniform.

where $R(a_t)$ is a rotation matrix defined by the action a_t (which is an angle) as depicted in Figure 3.6 (b) and a variance representing the model's uncertainty in the dynamics of $\sigma_l = \frac{1}{4}$. Secondly the distribution,

$$p(x_t|\mathbf{l}_t) = \phi(x_t; f(\mathbf{l}_t), \sigma_x), \quad (3.53)$$

where the variance represents uncertainty in the mapping $\sigma_x = \frac{1}{1000}$ and the mapping $f(\mathbf{l}_t)$ represents the calculated difference in loudness between the ears. The level disparity is represented as,

$$f(\mathbf{l}_t) = \log d\left(\mathbf{l}_t, \left[\frac{1}{2}w, 0\right]\right) - \log d\left(\mathbf{l}_t, \left[-\frac{1}{2}w, 0\right]\right), \quad (3.54)$$

for a unit head width ($w = 1$) depicted in Figure 3.6 (a). The function $d(\cdot, \cdot)$ is the Euclidian distance. I use these to define a particle filter according to Algorithm 3.1 using $N = 400$ particles.

My proposed information policy $B_{X_t|a_t}[\mathbf{L}_t||\mathbf{L}'_t]$ ($+I[\mathbf{L}_t|x_{0:t-1}, a_{0:t}]$) is constructed with an implementation of the Shannon information. These two policies are also compared to a uniform random policy, which acts as a minimum standard against which the other policies must be better. If a policy cannot outperform a random policy which has little computation cost then it has no utility.

The possible actions are selected from a set of actions, $a_t \in \left\{ \pi \left(\frac{j}{8} - 1 \right)^3 : j = 0, \dots, 8 \right\}$. This results in a bias toward little or no action if the likelihood of the action is uniformly distributed. So the j 'th action $a_t^{(j)} = \pi \left(\frac{j}{8} - 1 \right)^3$ is selected from one of the following three policies:

1. A random action policy where the likelihood of each action being selected is equal. An example trajectory of belief can be seen in Figure 3.7. This policy is included as a base comparison against which the other policies must perform better.
2. Through Infomax by selecting the j 'th action which maximises the Shannon information, $I[\mathbf{L}_t|X_t, a_t^{(j)}, x_{0:t-1}, a_{1:t-1}]$. This represents the Infomax comparison and is by necessity approximated, see Algorithm 3.5.
3. Through selecting the j 'th action which maximises my policy measure constructed as the upper bound to the Shannon information, $I[\mathbf{L}_t|a_t^{(j)}, x_{0:t-1}, a_{1:t-1}] + B_{X_t|a_t^{(j)}}[\mathbf{L}_t||\mathbf{L}'_t]$, see Algorithm 3.5.

Algorithm 3.5 Calculation of Infomax for active sound source localisation using a particle filter defined in Algorithm 3.1. Infomax is represented by the conditional Shannon information $I[\mathbf{L}_t|\mathbf{X}_t, x_{0:t-1}, a_{1:t}]$. This utilises Gibbs sampling for the information quantities $I[\mathbf{L}_t|x_{0:t-1}, a_{1:t}]$ and $I[\mathbf{X}_t|x_{0:t-1}, a_{1:t}]$. I use the latent prior belief $p(\mathbf{l}_{t-1}|x_{0:t-1}, a_{1:t-1}) = \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta(\mathbf{l}_{t-1|t-1}^{(i)} - \mathbf{l}_{t-1})$ and the latent prediction $p(\mathbf{l}_t|x_{0:t-1}, a_{1:t}) = \sum_{i=1}^N \pi_{t|t-1}^{(i)} \delta(\mathbf{l}_{t|t-1}^{(i)}(a_t) - \mathbf{l}_t)$, $\pi_{t|t-1}^{(i)} = \pi_{t-1|t-1}^{(i)}$ and $\mathbf{l}_{t|t-1}^{(i)}(a_t) \sim p(\mathbf{l}_t|\mathbf{l}_{t-1|t-1}^{(i)}, a_t)$. This algorithm scales as $O(N_a N^2)$.

To calculate the Shannon information,

1. for each a_t

(a) for each $i' \in [1, N]$

select an i^* according to the probabilities $\pi_{t|t-1}^{(i)}$

sample a point $x_{t|t-1}^{(j)}(a_t) \sim \phi(x_t; f(\mathbf{l}_{t|t-1}^{(i^*)}(a_t)), \sigma_x)$

$h_0 = 0$, $h_1 = 0$ and $h_2 = 0$

(a) for each $i' \in [1, N]$

$g_0 = 0$, $g_1 = 0$

i. for each $i \in [1, N]$

$g_2 = \phi(x_{t|t-1}^{(i')}(a_t); f(\mathbf{l}_{t|t-1}^{(i)}(a_t)), \sigma_x)$

$g_0 = g_0 + \pi_{t-1|t-1}^{(i)} \phi(\mathbf{l}_{t|t-1}^{(i')}(a_t); R(a_t)\mathbf{l}_{t-1|t-1}^{(i)}, \sigma_l)$

$h_1 = h_1 + \frac{\pi_{t|t-1}^{(i)}}{N} \log g_2$

$g_1 = g_1 + \pi_{t|t-1}^{(i)} g_2$

$h_0 = h_0 + \pi_{t|t-1}^{(j)} \log g_0$

$h_2 = h_2 + \frac{1}{N} g_1 \log g_1$

$I[\mathbf{L}_t|\mathbf{X}_t, a_t] = h_0 + h_1 - h_2$

2. select the a_t that maximises $I[\mathbf{L}_t|\mathbf{X}_t, a_t]$.

Algorithm 3.6 An implementation of my information measure (Equation 3.3) for active sound source localisation using a particle filter defined according to Algorithm 3.1. This utilises Gibbs sampling for the information quantity $I[\mathbf{L}_t|x_{0:t-1}, a_{1:t}]$. I use the latent prior belief $p(\mathbf{l}_{t-1}|x_{0:t-1}, a_{1:t-1}) = \sum_{i=1}^N \pi_{t-1|t-1}^{(i)} \delta(\mathbf{l}_{t-1|t-1}^{(i)} - \mathbf{l}_{t-1})$ and the latent prediction $p(\mathbf{l}_t|x_{0:t-1}, a_{1:t}) = \sum_{i=1}^N \pi_{t|t-1}^{(i)} \delta(\mathbf{l}_{t|t-1}^{(i)}(a_t) - \mathbf{l}_t)$, $\pi_{t|t-1}^{(i)} = \pi_{t-1|t-1}^{(i)}$ and $\mathbf{l}_{t|t-1}^{(i)}(a_t) \sim p(\mathbf{l}_t|\mathbf{l}_{t-1|t-1}^{(i)}, a_t)$. This algorithm scales as $O(N_a N^2)$, though if the quantity $I[\mathbf{L}_t|x_{0:t-1}, a_{1:t}]$ were neglected it would be equivalent to Algorithm 3.3 and hence scale as $O(N_a N)$.

Calculation of my method,

1. for each a_t ,

$$h_0 = 0, h_1 = 0, h_2 = 0, h_3 = 0, h_4 = 0, h_5 = 0, h_6 = 0$$

- (a) for each $i \in [1, N]$

$$h_0 = h_0 + \pi_{t|t-1}^{(i)} \log \left(\sum_{i'=1}^N \pi_{t-1|t-1}^{(i')} \phi \left(\mathbf{l}_{t|t-1}^{(i')} (a_t); R(a_t) \mathbf{l}_{t-1|t-1}^{(i)}, \sigma_l \right) \right)$$

$$h_1 = h_1 + \sigma \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right) \pi_{t|t-1}^{(i)}$$

$$h_2 = h_2 + \frac{1}{\sigma \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right)} \pi_{t|t-1}^{(i)}$$

$$h_3 = h_3 + f \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right) \pi_{t|t-1}^{(i)}$$

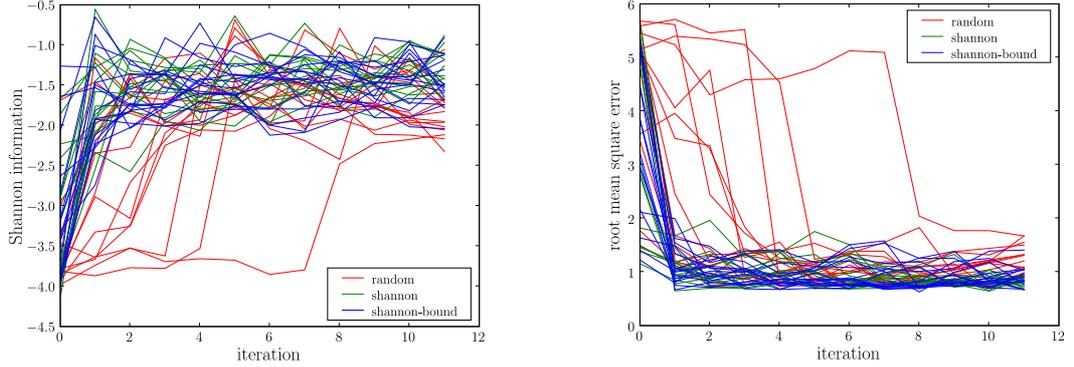
$$h_4 = h_4 + f \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right)^2 \pi_{t|t-1}^{(i)}$$

$$h_5 = h_5 + \frac{f \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right)}{\sigma \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right)} \pi_{t|t-1}^{(i)}$$

$$h_6 = h_6 + \frac{f \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right)^2}{\sigma \left(\mathbf{l}_{t|t-1}^{(i)} (a_t) \right)} \pi_{t|t-1}^{(i)}$$

$$I[\mathbf{L}_t|a_t] + B_{\mathbf{X}_t|a_t}[\mathbf{L}_t||\mathbf{L}'_t] = h_0 + \frac{1}{2} (h_1 h_2 - 1 + h_3 h_4 - 2h_4 h_5 + h_6)$$

2. select the a_t that maximises $I[\mathbf{L}_t|a_t] + B_{\mathbf{X}_t|a_t}[\mathbf{L}_t||\mathbf{L}'_t]$.
-

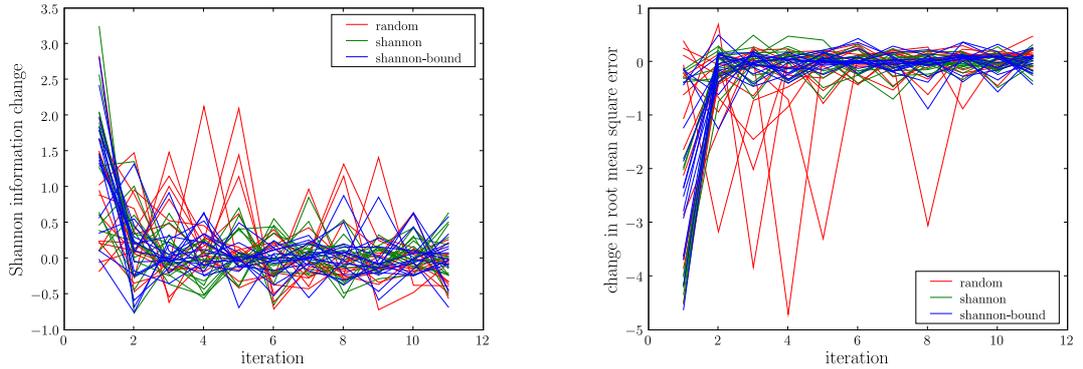


(a) The tracked posterior Shannon information, cf certainty. (b) The tracked posterior root mean square error, cf inverse accuracy.

Figure 3.8: Time course plots indicating that inference certainty increases in the same fashion as inference accuracy. Plots of the tracked measured Shannon information, plot (a), and the tracked root mean square error, plot (b), for multiple reruns of each action policy: random action selection, conditional Shannon information $I[\mathbf{L}_t | \mathbf{X}_t, x_{0:t-1}, a_{1:t}]$ and the bound upon the conditional Shannon information $I[\mathbf{L}_t | x_{0:t-1}, a_{1:t}] + B_{\mathbf{X}_t | a_t}[\mathbf{L}_t | \mathbf{L}'_t]$. This indicates that certainty (measured Shannon information) increases at the same rate as accuracy (root mean square error). The average correlation between the measured Shannon information and root mean square error for each policy, for the random action selection is -0.950 , the maximum conditional Shannon information $I[\mathbf{L}_t | \mathbf{X}_t, x_{0:t-1}, a_{1:t}]$ is -0.909 , and the bound upon this information content $I[\mathbf{L}_t | x_{0:t-1}, a_{1:t}] + B_{\mathbf{X}_t | a_t}[\mathbf{L}_t | \mathbf{L}'_t]$ is -0.914 . This shows that in this system, certainty (the measured Shannon information) is strongly correlated with accuracy (the inverse root mean square error).

These policies are evaluated by stepping one iteration forward through the model for each possible choice of action. I consider the behaviour of each of these informatic policies. Then construct stochastic alternatives to these from the behaviour. Finally, I evaluate the utility of the stochastic alternatives in the same way as the informatic policies.

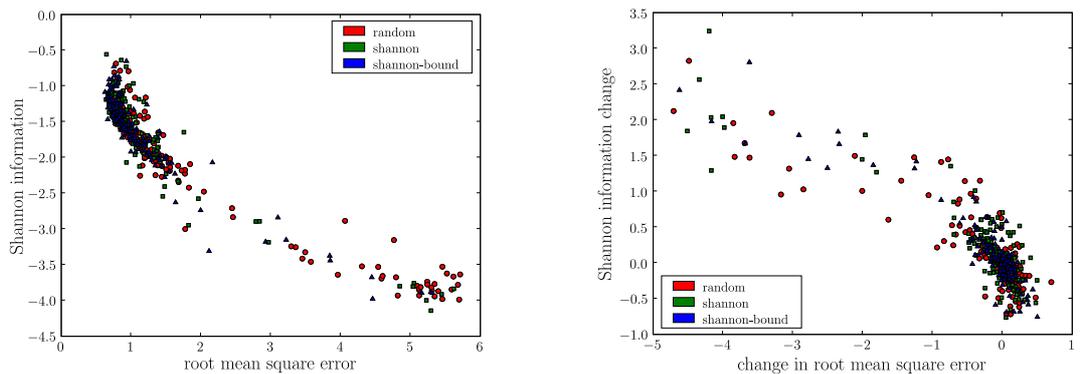
The process of evaluating each policy starts with the selection of a point $\mathbf{I}_{t-1}^{\text{true}}$ representing the true state of the world. For each action $a_t^{(j)}$ I simulate a step forward in time by generating a new true state as $\mathbf{I}_t^{\text{true}} = R(a_t^{(j)}) \mathbf{I}_{t-1}^{\text{true}}$. This in turn is used to generate a measurement from the cue space $x_t^{(j)} = f(\mathbf{I}_t^{\text{true}})$. I take the prior belief $p(\mathbf{I}_{t-1} | x_{0:t-1}, a_{1:t-1})$ and constructing the tracked posterior $p(\mathbf{I}_t | x_t^{(j)}, a_t^{(j)}, x_{0:t-1}, a_{1:t-1})$ to evaluate the measured Shannon information $I[\mathbf{L}_t | x_t^{(j)}, a_t^{(j)}, x_{0:t-1}, a_{1:t-1}]$ (a measure of certainty) and the root mean square error $\sqrt{\mathbb{E}_{p(\mathbf{I}_t | x_t^{(j)}, a_t^{(j)}, x_{0:t-1}, a_{1:t-1})} [(\mathbf{I}_t - \mathbf{I}_t^{\text{true}})^2]}$ (an inverse measure of accuracy). This allows a comparison for each possible action $a_t^{(j)}$ to evaluate if the policy's chosen action is in fact the best or simply acceptable according to these measures of certainty and accuracy.



(a) The change between iterations of the tracked posterior Shannon information, *cf* change in certainty.

(b) The change between iterations of the tracked posterior root mean square error, *cf* change in inverse accuracy.

Figure 3.9: Time course plots of the change in inference certainty and inference accuracy. Plots of the change in the tracked measured Shannon information plot (a) and change in the tracked root mean square error plot (b) of the different action policies with multiple reruns. Quantifying the number of times that the maximum change in certainty (the measured Shannon information) and inverse accuracy (the root mean square error) occurs for the first action a_1 of each policy is: $6/12$ for the random policy, $11/12$ for the conditional Shannon information (Infomax), and $11/12$ for the upper bound upon the conditional Shannon information.



(a) Correlation of the Shannon information (certainty) and the root mean square error (inverse accuracy).

(b) Correlation of the change in Shannon information and the change in root mean square error, *cf* change in certainty compared to the change in inverse accuracy.

Figure 3.10: Scatter plots indicating the correlation of accuracy with certainty, and the correlation of the equivalent changes. The tracked measured Shannon information represents the certainty. The tracked root mean square error represents the inverse accuracy. All policies result in a bulk of certainty (measured Shannon information) and accuracy (inverse root mean square error) and changes in these between iterations.

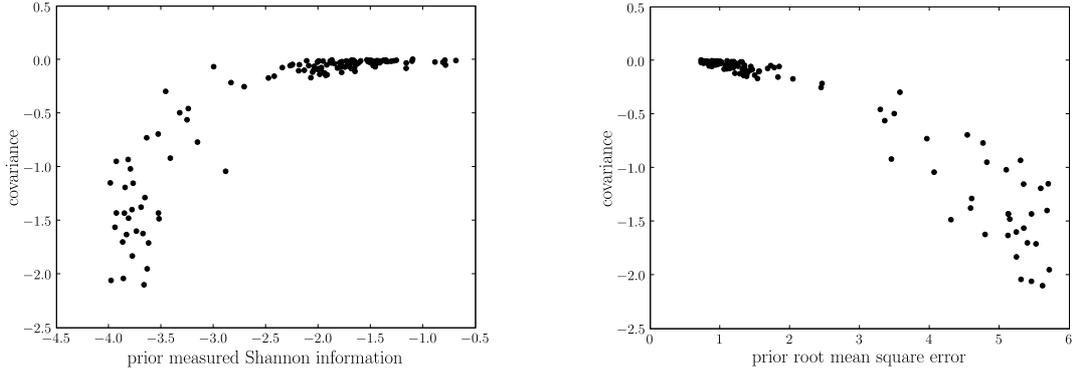
3.6.1 Analysis

It can be seen when considering the actual performance of each policy across multiple reruns (Figure 3.8) that there is a noisy bound upon the certainty (measured Shannon information) and accuracy (inverse root mean square error). Overall the random policy is the worst at guaranteeing an increase in the certainty and accuracy, Figure 3.9. This is to be expected as the random policy has no explicit mechanism to avoid the worst localising actions.

The correlation between accuracy and certainty is depicted in Figure 3.10. This shows a propensity for the information policies to hold tracked posterior beliefs in a region of high accuracy and certainty (Figure 3.8). The changes have a propensity to stabilise, see Figure 3.9. This indicates that the best actions cause a larger shift in both accuracy and certainty.

The information based policies each consistently collapse the uncertainty and consequently increases the accuracy of the tracked posterior for the first chosen action a_1 . This corresponds to a front-back ambiguous belief as can be seen in Figure 3.7 (a-d) (for instance in Figures 2.6-2.9). Hence, the information policies will “seek” to collapse any ambiguity in the latent space at the earliest opportunity, Figures 3.8 & 3.9. Out of a total of 12 runs for each policy the number of times the maximum change in certainty and accuracy occurs for the first action is 6 for the random policy, 11 for the conditional Shannon information, and 11 for the bound upon the conditional Shannon information. This indicates that when a gain in certainty and accuracy can be made, either information policy is consistent in selecting an action that provides such a gain, though not necessarily the best.

The importance of the prior accuracy and certainty upon the variability of the possible accuracies and certainties across the choices of action is indicated in Figure 3.11. The covariance between the single step look ahead accuracies and certainties is contrasted to the degree of prior certainty (Figure 3.11 (a)) and accuracy (Figure 3.11 (b)). The degree of covariance is dependent upon the level of prior ambiguity. The greater the prior ambiguity the greater the correlation is between the look ahead accuracies and certainties. Finally, when the prior ambiguity has receded these covariances approach zero. This suggests that once the tracked posterior has localised, the differences in projected certainties and accuracies over the possible actions are dominated by the noise in the data. The result of



(a) Relation of the covariance between the Shannon information and the root mean square error, to the corresponding time steps Shannon information of the prior belief (prior certainty).

(b) Relation of the covariance between the Shannon information and the root mean square error, to the corresponding time steps root mean square error of the prior belief (prior inverse accuracy).

Figure 3.11: Scatter plots indicating covariation of accuracy with certainty, and covariation of a one-step look ahead accuracy with certainty. The prior measured Shannon information $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}]$, in plot (a), represents the prior certainty. The prior root mean square error $\sqrt{\mathbb{E}_{p(\mathbf{l}_{t-1}|x_{0:t-1}, a_{1:t-1})}[(\mathbf{l}_{t-1} - \mathbf{l}_{t-1}^{\text{true}})^2]}$, in plot (b), represents the prior inverse accuracy. Both plots consider the covariance between the look ahead measured conditional Shannon information $I[L_t|x_t^{(j)}, a_t^{(j)}, x_{0:t-1}, a_{1:t-1}]$ and the root mean square error $\sqrt{\mathbb{E}_{p(\mathbf{l}_t|x_t^{(j)}, a_t^{(j)}, x_{0:t-1}, a_{1:t-1})}[(\mathbf{l}_t - \mathbf{l}_t^{\text{true}})^2]}$ using the tracked priors $p(\mathbf{l}_{t-1}|x_{0:t-1}, a_{1:t-1})$ from Figure 3.8 for comparison between actions.

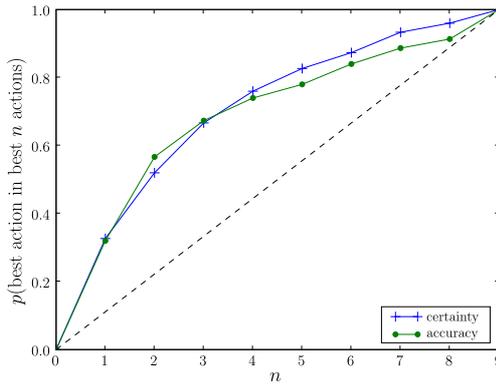
greater covariance between look ahead accuracy and certainty¹⁰ show that there is also a larger variance between the consequent ambiguities for the choices of action.

Similarly, the conditional Shannon information (Equation 3.5) or its bound (Equation 3.3) actually select for more accurate and certain future posterior beliefs $p(\mathbf{l}_t|x_{0:t}, a_{1:t})$ (Figure 3.12). This indicates that a choice of action according to either information policy will select better actions according to both accuracy and certainty.

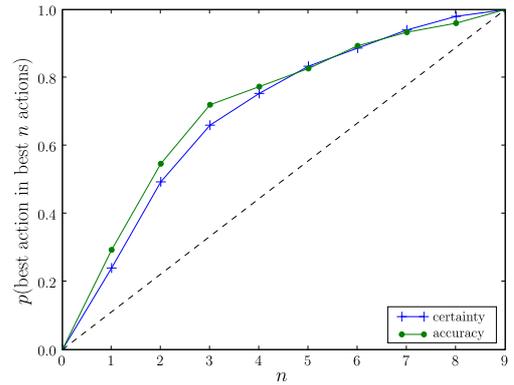
¹⁰The look ahead accuracy and certainty denote a simulated single step forward for the system to compare across the possible choices of action; essentially taking a single step in the simulation and rolling back the system to take a different action and again taking a single step in the simulation.

prior situations	policy	certainty	accuracy
all	$I[L_t x_{0:t-1}, a_{1:t}] + B_{X_t a_t}[L_t L'_t]$	6.79	6.79
only ambiguous	$I[L_t x_{0:t-1}, a_{1:t}] + B_{X_t a_t}[L_t L'_t]$	7.14	7.29
all	$I[L_t X_t, x_{0:t-1}, a_{1:t}]$	6.87	6.72
only ambiguous	$I[L_t X_t, x_{0:t-1}, a_{1:t}]$	6.88	6.82

Table 3.1: The tabulated performance of information policies for all situation and ambiguous situations. The area under each ROC in Figures 3.12 & 3.13 can be used to indicate the performance of each policy, for the selection of all prior cases (Figure 3.12) in contrast to only those cases that are ambiguous (Figure 3.13). A perfect action selection policy would be 8.5, in contrast the random policy would be 4.5.

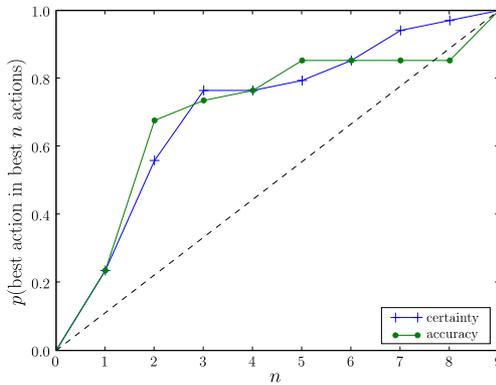


(a) The Shannon information.

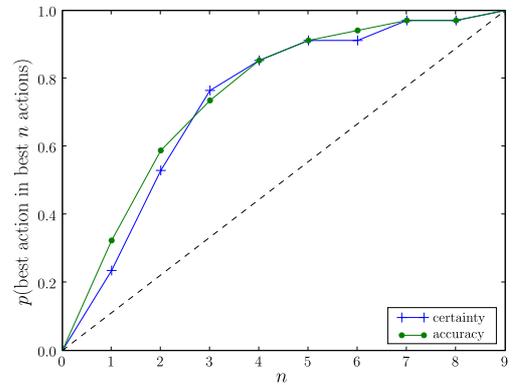


(b) The upper bound of the Shannon information.

Figure 3.12: The receiver operator curves (ROC) illustrating performance of the information measures ordering of actions in all situations. The ROCs depict when the predicted best action is amongst either information policies best n alternatives. These ROCs are for each policy and show the probability that the best action is taken, according to accuracy and certainty – this is for the conditional Shannon information in plot (a) and my approach in plot (b). The expected result of a uniform random policy is indicated by the black dashed line. The best action is judged by the one-step look ahead simulations. The certainty is proportional to the look ahead measured Shannon information. The accuracy is proportional to the look ahead inverse root mean square error.



(a) The Shannon information.



(b) The upper bound of the Shannon information.

Figure 3.13: The receiver operator curves (ROC) illustrating performance of the information measures ordering of actions in only the most uncertain and ambiguous situations. The ROC for the most ambiguous prior beliefs depict when the predicted best action is amongst either information policies best n alternatives. These ROCs are for each policy and show the probability that the best action is taken, according to accuracy and certainty – this is for the conditional Shannon information in plot (a) and my approach in plot (b). The expected result of a uniform random policy is indicated by the black dashed line. The best action is judged by the one-step look ahead simulations. The certainty is proportional to the look ahead measured Shannon information. The accuracy is proportional to the look ahead inverse root mean square error.

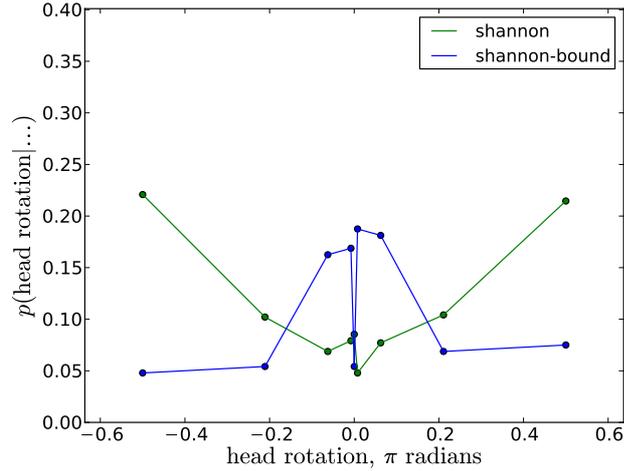


Figure 3.14: Histograms of the two informatic policies showing the selection of actions.

The performance for all cases of these two information policies in selecting the best action is indicated by the receiver operating curves (ROC) Figure 3.12. This indicates that although the best action is not guaranteed for either policy, it or the next best has a greater than even chance of selection. This is more pronounced (Table 3.1) when only the most ambiguous prior beliefs are considered, see Figure 3.13. In comparing Figure 3.13 (b) and Figure 3.13 (a) we can see that where the *a priori* belief is ambiguous the ROC for my policy (Algorithm 3.6) is smoother than the Shannon information policy (Algorithm 3.5). I account for this as the distribution of N samples in Algorithm 3.5 must represent a more complex $p(x_t|a_t^{(j)}, x_{0:t-1}, a_{1:t-1})$ distribution than when the prior is not ambiguous. This is especially true for the “worse” actions, Figure 3.13 (a).

3.6.2 An alternative stochastic strategy

In this subsection I consider an alternative stochastic strategy to the uniform random case. I define two strategies that broadly mimic the behaviour of the informatic strategies: the conditional Shannon information, and my measure $I[L_t|x_{0:t-1}, a_{1:t}] + B_{X_t|a_t}[L_t||L_t]$. The histograms of an action, a head rotation, for each of these policies is shown in Figure 3.14. The Infomax approach selects larger rotations with a higher probability, cf Figures 3.5 (b) & 3.6 (b) with the most likely rotations as $\pm\frac{\pi}{2}$ for an initial prior. However, from Figure 3.14, my measure seems to show a preference for smaller head rotations. Such a simple analysis does not seem to capture either informatic policies complexity.

In Figure 3.15 (a) we can see the distribution of prior certainties, the Shannon information of the prior belief $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}]$. For both policies there is a long tail of

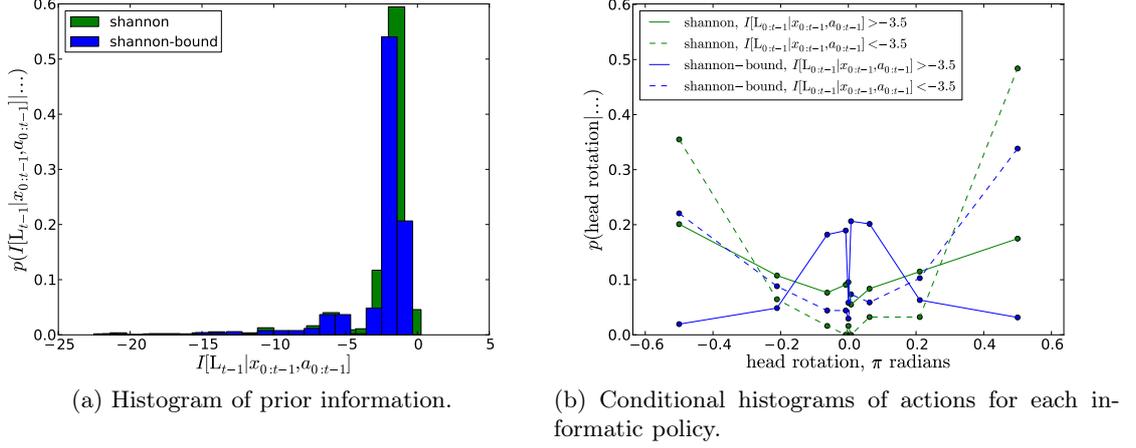


Figure 3.15: The conditional histograms of actions for each informatic policy. The conditional histograms in plot (b) are used to define alternative stochastic policies to each informatic policy. In plot (a), for both informatic policies, the histogram of information content of the prior beliefs $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}]$ shows a long tailed distribution with negative skew. The detail shows a second peak below $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}]$ of -5 . I segment the priors at -3.5 to give the conditional histograms of actions in plot (b).

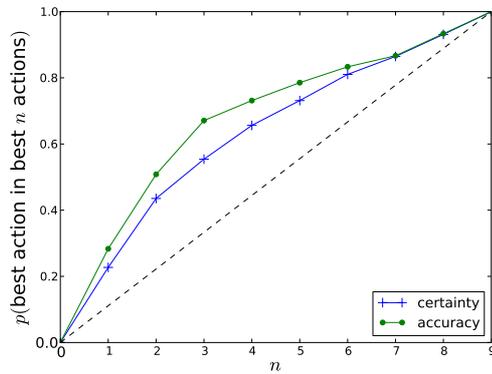
prior situations	policy	certainty	accuracy
all	stochastic $I[L_t x_{0:t-1}, a_{1:t}] + B_{X_t a_t}[L_t L'_t]$	4.35	4.16
only ambiguous	stochastic $I[L_t x_{0:t-1}, a_{1:t}] + B_{X_t a_t}[L_t L'_t]$	6.38	6.88
all	stochastic $I[L_t X_t, x_{0:t-1}, a_{1:t}]$	5.71	6.11
only ambiguous	stochastic $I[L_t X_t, x_{0:t-1}, a_{1:t}]$	7.19	7.62

Table 3.2: The tabulated performance of stochastic policies derived from the behaviour of the information policies for all situation and ambiguous situations. The area under each ROC in Figures 3.16 & 3.17 can be used to indicate the performance of either stochastic policy, for the selection of all prior cases (Figures 3.16 & 3.16 (a)) in contrast to only those cases that are ambiguous (Figures 3.16 & 3.16 (b)). A perfect action selection policy would be 8.5, in contrast the random policy would be 4.5.

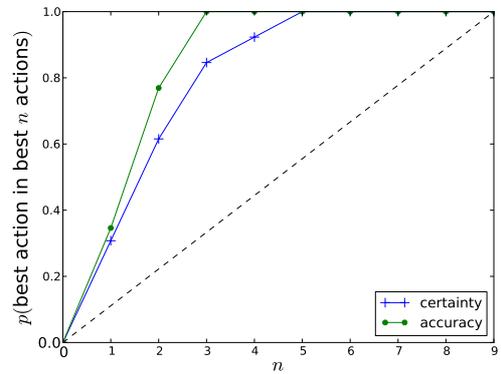
ambiguous prior beliefs for $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}] < -3.5$, with a substantial peak of unambiguous priors for $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}] > -3.5$. Constructing histograms of the informatic choice of actions conditional upon this cut at $I[L_{t-1}|x_{0:t-1}, a_{0:t-1}] = -3.5$ produces Figure 3.15 (b). We can see that larger ambiguities, for both informatic policies, leads to larger head rotations of $\pm\frac{\pi}{2}$. In contrast the distributions of actions for the less ambiguous prior situations of both policies retain the same shape, Figure 3.14.

I use the histograms in Figure 3.15 (b) to define two stochastic policies. Both stochastic strategies define a probability of selecting an action conditional upon the prior beliefs Shannon information $I[L_{t-1}|x_{0:t-1}, a_{1:t-1}]$. The resultant ROCs for using these stochastic policies can be seen in Figures 3.16 & 3.17.

The stochastic policy's ROC defined from the Shannon information in Figure 3.16 (a)

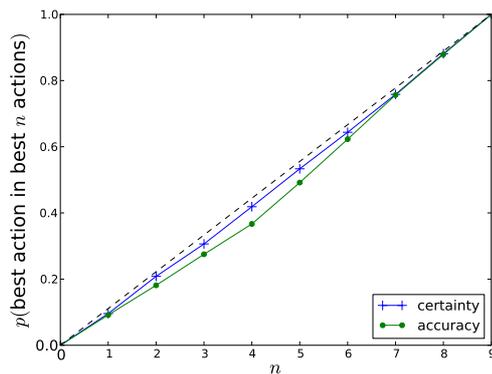


(a) All prior beliefs.

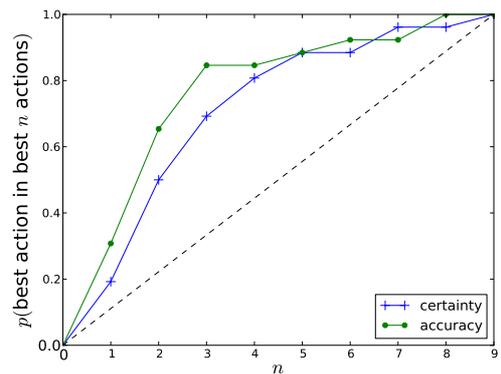


(b) Only ambiguous priors.

Figure 3.16: The ROC that a stochastic policy will favour taking the best action according to accuracy or certainty. The stochastic policy is derived from the action selection histograms of the Shannon information in Figure 3.15. Plot (a) depicts the ROC for all prior beliefs, and plot (b) depicts the ROC for only the most ambiguous prior beliefs. The expected result of a uniform random policy is indicated by the black dashed line.



(a) All prior beliefs.



(b) Only ambiguous priors.

Figure 3.17: The ROC that a stochastic policy will favour taking the best action according to accuracy or certainty. The stochastic policy is derived from the action selection histograms of the upper bound of the Shannon information in Figure 3.15. Plot (a) depicts the ROC for all prior beliefs, and plot (b) depicts the ROC for only the most ambiguous prior beliefs. The expected result of a uniform random policy is indicated by the black dashed line.

has an almost identical curve to that of the Shannon information's, Figure 3.12 (a). Further, in contrasting Table 3.2 and Figure 3.16 (b) to Table 3.1 and Figure 3.13 (a), we can see that the stochastic policy can be expected to outperform the Shannon information policy for the ambiguous priors.

The contrast between the ROCs using the stochastic policy defined from my approach, for ambiguous priors and all situations in Figure 3.17 is striking. It indicates that the stochastic policy does not take into account import factors of the prior belief that my approach does. This is illustrated best by comparing Figure 3.17 (a) to Figure 3.12 (b), as the stochastic policy for all situations is worse than the uniform random case.

Comparing Tables 3.2 & 3.1 we can see, for ambiguous prior situations, that stochastic policies have expected performances which approach or exceed the comparative informatic policies. However, in the wider context of all the situations, only the stochastic policy defined from the Shannon information policy achieves a similar expected performance (Table 3.2).

The cost of calculating a stochastic policy is far less than an informatic policy. This leads to the conclusion that a policy which captures the behaviour of an informatic policy can be an acceptable substitute, for instance $I[L_t|X_t, x_{0:t-1}, a_{1:t}]$ policy. However, comparing Figure 3.17 (a) to Figure 3.12 (b) illustrates the case where a stochastic policy has failed to capture the specific behaviour of the informatic policy. This is most catastrophic for the general case of all situations.

3.7 Comparison with Infotaxis (Vergassola *et al.*, 2007a)

Infotaxis (Vergassola *et al.*, 2007a) is applied to the chemotaxis in turbulent diffusive mediums. This is modelled as the use of discrete events to navigate towards a source, for instance a chemical source emitting a number of chemical packets. There is no smooth gradient of increasing chemical concentration that can be followed to the source. Thus there is no concentration gradient that can be followed using hill climbing techniques.

Infotaxis considers the expected change in information about a source's location ΔI between two time steps. The prior time steps information content does not change with a choice of action. Thus, this makes $\arg \max_{a_t} \mathbb{E}[\Delta I|a_t]$ equivalent to $\arg \max_{a_t} \mathbb{E}[I|a_t]$ the Shannon information of the expected posterior belief in source latent location.

The posterior belief in a source's location \mathbf{l}_t at time t is annotated as $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t})$ where $x_{0:t}$ represents the sequence of measured odour encounters along the traversed tra-

jectory $\mathbf{a}_{0:t}$. Two effects contribute to the posterior $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t})$. First that $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t}) = 0$, as the source has not been found. Secondly, the estimated posterior is modified by x_t observed detections. Therefore the expected change in Shannon information for the agent moving into one of the neighbouring cells \mathbf{a}_t is defined in Equation 3.58. There is neither drift nor uncertainty in the location of the agent. This leads to the posterior being proportional to the product of the likelihoods $p(x_{t'}|\mathbf{l}_{t'}, \mathbf{a}_{t'})$ for $t' \in [0, t]$, as a consequence $p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) \equiv p(\mathbf{l}_{t-1}|x_{0:t-1}, \mathbf{a}_{0:t-1})$. This posterior is

$$p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t}) = \frac{\exp\left\{-\int_0^t dt' R(\mathbf{a}_{t'}|\mathbf{l}_t)\right\} \prod_{t'=0}^t R(\mathbf{a}_{t'}|\mathbf{l}_t)^{x_{t'}}}{\int d\mathbf{l}'_t \exp\left\{-\int_0^t dt' R(\mathbf{a}_{t'}|\mathbf{l}'_t)\right\} \prod_{t'=0}^t R(\mathbf{a}_{t'}|\mathbf{l}'_t)^{x_{t'}}}, \quad (3.55)$$

this is taken from Vergassola *et al.* (2007a) where $R(\mathbf{a}_t|\mathbf{l}_t)$ is the rate of detections that an agent located at a_t can be expected to detect from a source at \mathbf{l}_t .¹¹ This can be computed in closed form, as a hidden Markov model (HMM), using

$$p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t}) = \frac{p(x_t|\mathbf{l}_t, \mathbf{a}_t) p(\mathbf{l}_{t-1} = \mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})}{\int d\mathbf{l}'_t p(x_t|\mathbf{l}'_t, \mathbf{a}_t) p(\mathbf{l}_{t-1} = \mathbf{l}'_t|x_{0:t-1}, \mathbf{a}_{0:t-1})}. \quad (3.57)$$

The likelihood $p(x_t|\mathbf{l}_t, \mathbf{a}_t)$ is a Poisson distribution taking the form $\frac{R(\mathbf{a}_t|\mathbf{l}_t)^{x_t} e^{-R(\mathbf{a}_t|\mathbf{l}_t)}}{x_t!}$.

The expected change in Shannon information for moving to one of the neighbouring points a_t (or standing still) is

$$\begin{aligned} \mathbb{E}[\Delta I|\mathbf{a}_t] &= p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{t-1}) [-I[\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]] \\ &\quad + (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{t-1})) \left[\sum_{x_t=0}^{\infty} \rho(x_t; \bar{x}(\mathbf{a}_t)) \mathbb{E}[\Delta I|\mathbf{a}_t, x_t] \right]. \end{aligned} \quad (3.58)$$

The first term represents the situation of the agent finding the source by moving into it, making the posterior belief collapse to a delta function and causing the Shannon information to become zero. We cannot, with any certainty, state how many detections the agent will make, instead this is estimated using a Poisson law $\rho(x; \bar{x}) = \frac{\bar{x}^x e^{-\bar{x}}}{x!}$ for independent detections over the time-step Δt . The expected number of detections is $\bar{x}(\mathbf{l}) = \Delta t \int d\mathbf{l}' p(\mathbf{l}'_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) R(\mathbf{l}'_t|\mathbf{l}_t)$, where $R(\mathbf{l}|\mathbf{l}_t)$ denotes the mean rate of detec-

¹¹The mean rate of detections for an agent located at $\mathbf{l} = \{l_x, l_y\}$ and a source located at $\mathbf{l}^{\text{source}} = \{l_x^{\text{source}}, l_y^{\text{source}}\}$ for a 2D grid (Vergassola *et al.*, 2007a,b) is,

$$R(\mathbf{l}|\mathbf{l}^{\text{source}}) = \frac{R}{\log\left(\frac{\Delta}{a}\right)} e^{\frac{(l_y^{\text{source}} - l_y)v}{2D}} K_0\left(\frac{|\mathbf{l} - \mathbf{l}^{\text{source}}|}{\lambda}\right); \quad \lambda = \sqrt{\frac{D\tau}{1 + \frac{v^2\tau}{4D}}} \quad (3.56)$$

where the parameters of this equation are τ is the particle lifetime, v is the wind speed along the y-axis, D is the diffusivity and a is the size of the searcher corresponding to one grid cell.

Algorithm 3.7 The Infotaxis algorithm, implementing Vergassola *et al.* (2007a,b), for selecting the action that maximises the expected change in Shannon information. The possible actions \mathbf{a}_t , at a time t , consist of the agent remaining stationary or moving to one of the adjacent grid cells. The algorithm takes a prior belief in the position of the source \mathbf{l}_t over the field of locations \mathbf{L}_t as $p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$ which has a Shannon information of I_t . This algorithm scales as $O(N_l N_x N_a)$, where N_a is the number of candidate actions, N_x is the assumed maximum possible number of detections, and N_l represents the number of locations in the set of possible source locations \mathbf{L}_t . The mean rate of detections $R(\mathbf{a}_t|\mathbf{l}_t)$ for the agents proposed location \mathbf{a}_t and the source locations \mathbf{l}_t can be precomputed for the field of source locations \mathbf{L}_t and each of the possible agent locations \mathbf{a}_t . This algorithm is elucidated from Vergassola *et al.* (2007a) and its supplementary material (Vergassola *et al.*, 2007b).

To select the expected most informative movement \mathbf{a}_t of an agent from a location \mathbf{a}_{t-1} .

1. for each \mathbf{a}_t ,

(a) calculate $\bar{x}(\mathbf{a}_t) = \sum_{\forall \mathbf{l}_t \in \mathbf{L}_t} p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) R(\mathbf{a}_t|\mathbf{l}_t)$

(b) for each $x_t \in [0, N_x]$,

calculate $p(x_t|\mathbf{a}_t) = \frac{\bar{x}(\mathbf{a}_t)^{x_t} e^{-\bar{x}(\mathbf{a}_t)}}{x_t!}$

initialise $Z = 0$,

i. for each $\mathbf{l}_t \in \mathbf{L}_t$,

calculate $p(\mathbf{l}_t, x_t|x_{0:t-1}, \mathbf{a}_{0:t}) \propto p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) e^{-R(\mathbf{a}_t|\mathbf{l}_t)} R(\mathbf{a}_t|\mathbf{l}_t)^{x_t}$

accumulate $Z = Z + p(\mathbf{l}_t, x_t|x_{0:t-1}, \mathbf{a}_{0:t})$

ii. for each $\mathbf{l}_t \in \mathbf{L}_t$,

calculate $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t}) = \frac{p(\mathbf{l}_t, x_t|x_{0:t-1}, \mathbf{a}_{0:t})}{Z}$

calculate $I[\mathbf{L}_t|\mathbf{X}_t, x_{0:t-1}, \mathbf{a}_{0:t}] = \sum_{\forall \mathbf{l}_t \in \mathbf{L}_t} p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t}) \log p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t})$,

calculate $\mathbb{E}[\Delta I|x_t, \mathbf{a}_t] = I[\mathbf{L}_t|\mathbf{X}_t, x_{0:t-1}, \mathbf{a}_{0:t}] - I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]$,

calculate the expected change in Shannon information

$$\begin{aligned} \mathbb{E}[\Delta I|\mathbf{a}_t] &= p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) [-I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]] + \\ &\quad (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) \sum_{x_t=0}^{N_x} p(x_t|\mathbf{a}_t) \mathbb{E}[\Delta I|x_t, \mathbf{a}_t], \end{aligned}$$

2. select the \mathbf{a}_t that maximises $\mathbb{E}[\Delta I|\mathbf{a}_t]$.

tions at position \mathbf{l} if the source is located at \mathbf{l}_t . This assumes that a particular number of detections x_t causes a change in information $\mathbb{E}[\Delta I|\mathbf{a}_t, x_t]$ between the probability fields $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t})$ and $p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$. This is the Infotaxis policy model of Vergassola *et al.*, and is expressed in Algorithm 3.7.

I convert $\mathbb{E}[\Delta I|\mathbf{a}_t]$ to the expected conditional Shannon information with the addition of $I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]$ (the prior Shannon information) making Equation 3.58 compatible with my notation in this chapter,

$$\begin{aligned} \overline{I[\mathbf{L}_t|X_t, x_{0:t-1}, \mathbf{a}_{0:t}]} &= p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) [-I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]] \\ &\quad + (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) \times \\ &\quad \left[\sum_{x_t=0}^{\infty} \rho(x_t; \bar{x}(\mathbf{a}_t)) I[\mathbf{L}_t|x_{0:t}, \mathbf{a}_{0:t}] \right], \end{aligned} \quad (3.59)$$

$$\begin{aligned} &= p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) [-I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]] \\ &\quad + (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) I[\mathbf{L}_t|X_t, x_{0:t-1}, \mathbf{a}_{0:t}]. \end{aligned} \quad (3.60)$$

The quantity $I[\mathbf{L}_t|X_t, x_{0:t-1}, \mathbf{a}_{0:t}] = \mathbb{E}_{p(x_t|x_{0:t-1}, \mathbf{a}_{0:t})} [I[\mathbf{L}_t|x_{0:t}, \mathbf{a}_{0:t}]]$ takes an approximation of the marginal probability for the number of detections as, $p(x_t|x_{0:t-1}, \mathbf{a}_{0:t}) \approx \rho(x_t; \bar{x}(\mathbf{a}_t))$ which is a unimodal distribution ($\bar{x}(\mathbf{a}_t) = \int d\mathbf{l}_t R(\mathbf{a}_t|\mathbf{l}_t) p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$). In contrast the true distribution is

$$p(x_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) = \int d\mathbf{l}_t p(x_t|\mathbf{l}_t, \mathbf{a}_t) p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}), \quad (3.61)$$

$$= \int d\mathbf{l}_t \rho(x_t; R(\mathbf{a}_t|\mathbf{l}_t)) p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}). \quad (3.62)$$

This is the approximation of Vergassola *et al.* for speeding the Infotaxis algorithm.

My measure, the quantity $B_{X_t|\mathbf{a}_t}[\mathbf{L}_t||\mathbf{L}'_t]$ is computed as the upper limit to $I[\mathbf{L}_t|X_t, x_{0:t-1}, \mathbf{a}_{0:t}] - I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t}]$, producing the bound

$$\begin{aligned} \overline{I[\mathbf{L}_t|X_t, x_{0:t-1}, \mathbf{a}_{0:t}]} &\leq p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) [-I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]] \\ &\quad + (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) \times \\ &\quad (I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t}] + B_{X_t|\mathbf{a}_t}[\mathbf{L}_t||\mathbf{L}'_t]), \end{aligned} \quad (3.63)$$

$$\begin{aligned} &= (1 - 2p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) I[\mathbf{L}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}] \\ &\quad + (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) B_{X_t|\mathbf{a}_t}[\mathbf{L}_t||\mathbf{L}'_t]. \end{aligned} \quad (3.64)$$

The likelihood of x_t detections at the agents location \mathbf{a}_t for a source at \mathbf{l}_t , $p(x_t|\mathbf{l}_t, \mathbf{a}_t)$ is governed by a Poisson law (where the expected number of detections is $R(\mathbf{a}_t|\mathbf{l}_t)$), the KL-divergence between the likelihood of two hypothesis' \mathbf{l}_t and \mathbf{l}'_t can be constructed analytically as

$$D_{\text{KL}} [p(X_t|\mathbf{l}_t, \mathbf{a}_t) \| p(X_t|\mathbf{l}'_t, \mathbf{a}_t)] = \sum_{x_t} p(x_t|\mathbf{l}_t, \mathbf{a}_t) \log \frac{p(x_t|\mathbf{l}_t, \mathbf{a}_t)}{p(x_t|\mathbf{l}'_t, \mathbf{a}_t)}, \quad (3.65)$$

$$= R(\mathbf{a}_t|\mathbf{l}_t) \left(\log \frac{R(\mathbf{a}_t|\mathbf{l}_t)}{R(\mathbf{a}_t|\mathbf{l}'_t)} - 1 \right) + R(\mathbf{a}_t|\mathbf{l}'_t). \quad (3.66)$$

Thus the bound of the mutual information is computed as,

$$B_{X_t|\mathbf{a}_t} [\mathbf{L}_t \| \mathbf{L}'_t] = \int d\mathbf{l}_t d\mathbf{l}'_t p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) p(\mathbf{l}'_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) D_{\text{KL}} [p(X_t|\mathbf{l}_t, \mathbf{a}_t) \| p(X_t|\mathbf{l}'_t, \mathbf{a}_t)], \quad (3.67)$$

$$= \int d\mathbf{l}_t p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) R(\mathbf{a}_t|\mathbf{l}_t) \log R(\mathbf{a}_t|\mathbf{l}_t) + \left[\int d\mathbf{l}_t p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) R(\mathbf{a}_t|\mathbf{l}_t) \right] \times \left[\int d\mathbf{l}'_t p(\mathbf{l}'_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) \log R(\mathbf{a}_t|\mathbf{l}'_t) \right]. \quad (3.68)$$

We can see that in none of these expectations over the prior prediction $p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$ is it necessary to estimate the posterior $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t})$ for any $x_t \in [0, \infty)$. The quantity Equation 3.64 (with Equation B.9 inserted for $B_{X_t|\mathbf{a}_t} [\mathbf{L}_t \| \mathbf{L}'_t]$) defines a policy model I call Boundtaxis which is expressed in Algorithm 3.8.

Algorithms 3.7 & 3.8 are used to compute respectively Infotaxis, and Boundtaxis. These two 'taxis algorithms can be seen to scale differently. For a grid of N_l possible source locations, N_a choices of action¹² and a maximum number of expected detections N_x , my Boundtaxis algorithm scales as $O(N_l N_a)$, whereas Infotaxis scales as $O(N_l N_x N_a)$. Hence Boundtaxis scales better than Infotaxis. Though, comparing the real time performance of the two algorithms, I run both algorithms with no detections, finding that the Infotaxis algorithm on average takes 10 times as long as Boundtaxis to calculate one choice of action. This is using $N_x = 20$, but as the Infotaxis algorithms time_complexity scales linearly with N_x , we can expect $N_x = 10$ to only take Infotaxis about 5 times, and $N_x = 5$ to take about $2\frac{1}{2}$ times as long to make a choice of action as Boundtaxis. N_x must be selected so as to represent the extent of the highest mean rate $R(\mathbf{a}|\mathbf{l})$'s for the Poisson likelihood. Another way to improve the computational speed of Infotaxis would be to choose an N_x dependent

¹²There are 5 choices of \mathbf{a}_t at any \mathbf{a}_{t-1} , the 4 neighbouring grid cells or remaining in the same cell.

Algorithm 3.8 The Boundtaxis algorithm for selecting the action that maximises a term proportional to the upper bound of the expected Shannon information. The possible actions \mathbf{a}_t , at a time t , consist of the cases of the agent remaining stationary or moving to one of the adjacent grid cells. The algorithm takes a prior belief in the position of the source \mathbf{l}_t over the field of locations L_t as $p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$ which has a Shannon information of I_t . This algorithm scales as $O(N_l N_a)$, where N_a is the number of candidate actions, and N_l represents the number of locations in the set of possible source locations L_t . The mean rate of detections $R(\mathbf{a}_t|\mathbf{l}_t)$ for the agents proposed location \mathbf{a}_t and the source locations \mathbf{l}_t can be precomputed for the field of source locations L_t and each of the possible agent locations \mathbf{a}_t .

To select the expected most informative movement \mathbf{a}_t of an agent from a location \mathbf{a}_{t-1} .

1. for each \mathbf{a}_t ,

initialise $h_1 = 0$, $h_2 = 0$ and $h_3 = 0$,

- (a) for each $\mathbf{l}_t \in L_t$,

accumulate $h_1 = h_1 + p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) R(\mathbf{a}_t|\mathbf{l}_t) \log R(\mathbf{a}_t|\mathbf{l}_t)$,

accumulate $h_2 = h_2 + p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) R(\mathbf{a}_t|\mathbf{l}_t)$

accumulate $h_3 = h_3 + p(\mathbf{l}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) \log R(\mathbf{a}_t|\mathbf{l}_t)$,

calculate the upper bound of the Infotaxis Shannon information

$$B_{\text{upper}}[\mathbf{a}_t] = (1 - 2p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) I[L_t|x_{0:t-1}, \mathbf{a}_{0:t-1}] + (1 - p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1})) (h_1 - h_2 h_3),$$

2. select the \mathbf{a}_t that maximises $B_{\text{upper}}[\mathbf{a}_t]$.
-

on $\bar{x}(\mathbf{a}_t)$.

3.7.1 Analysis

The simulation for this comparison Infotaxis (Vergassola *et al.*, 2007a) was based around the Infotaxis paper Vergassola *et al.* (2007a) and the supplementary material (Vergassola *et al.*, 2007b). Vergassola *et al.* derived the mean rate of detections $R(\mathbf{a}_t|\mathbf{l}_t)$ for a turbulent and diffusive packet model which is the basis of both the simulations and the policy models for Infotaxis (Algorithm 3.7) and my Boundtaxis (Algorithm 3.8). Boundtaxis uses the measure constructed in Equation 3.64 as the upperbound to the Infotaxis information change Equation 3.58.

A single track of the agent's path for my Boundtaxis policy model compared to Infotaxis is shown in Figure 3.18. The posterior belief $p(\mathbf{l}_t|x_{0:t}, \mathbf{a}_{0:t})$ is symmetric around the agent's starting point and the agent starts spiralling around it. In Figure 3.18 (a), Boundtaxis constructs a far wider spiral than Infotaxis then spirals inwards (Figure 3.18 (b)). This example uses the parameterisation for $R(\mathbf{a}_t|\mathbf{l}_t)$ used in Vergassola *et al.* (2007a) and applies

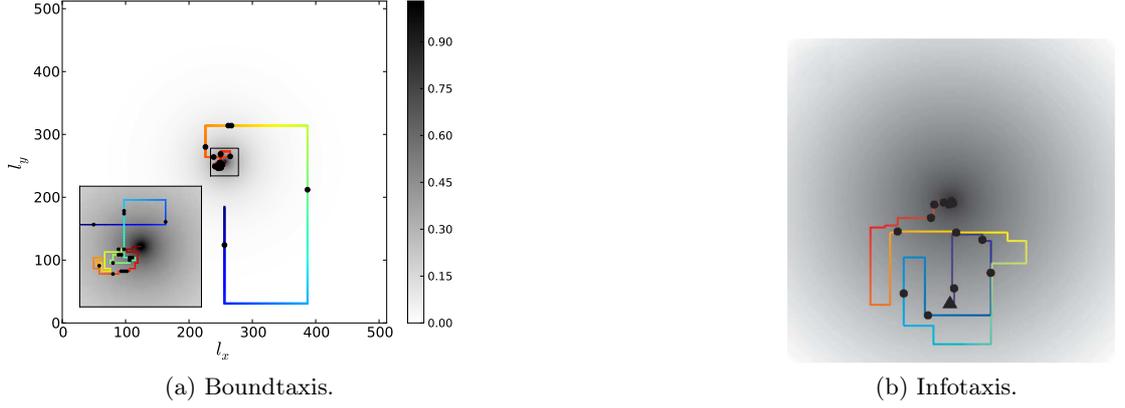


Figure 3.18: A single track for the agent using each of the Boundtaxis (a) and Infotaxis (b) policies. The time is indicated by a colour spectrum. The detections are indicated by filled black circles. The inset figure provides detail of the search for the immediate locale around the source. The grey scale depicts the expected rate of detection. Plot (b) is taken from Vergassola *et al.* (2007a), the axis and scale was not indicated in their paper.

to Figures 3.18-3.21.¹³

I compare the performance of Infotaxis (Vergassola *et al.*, 2007a) and Boundtaxis search times in Figure 3.19. The error bars indicate the standard deviation of the average search times. On average Boundtaxis has consistently faster search times than Infotaxis. There are two interesting and problematic features:

1. Boundtaxis has a significantly greater variance.
2. The trend for larger initial distances is unfavourable for those greater than 110 (the extent of the figures range).

The Boundtaxis policy has the trend that larger distances have increasing variance in the search times. However, comparing the distribution of search times between the two policy models (Figure 3.20) would suggest that an agent can expect to find its target (the source) quicker by using Boundtaxis than Infotaxis. The evolution of the posterior entropy with the remaining search time in Figure 3.21 shows clearly that Boundtaxis is not targeted at minimising the posterior entropy.

I compare Boundtaxis with the results from Vergassola *et al.* (2007b) in Figures 3.22 & 3.23. These figures use a different set of parameters to the simulations in Figure 3.18-3.21.¹⁴ Figure 3.23 is a comparison between the probability density functions (PDF) of

¹³An emission rate of $R = 1$, a particle lifetime $\tau = 2500$ and diffusivity of $D = 1$. The simulations are performed on a grid of 512^2 , with no wind and a starting position along one axis away from the source. Which equates to an average particle lifetime of $\sqrt{D\tau} = 50$.

¹⁴An emission rate of $R = 1$, a particle lifetime $\tau = 400$ and diffusivity of $D = 1$. The simulations are performed on a grid of 256^2 , with no wind and a starting position along one axis away from the source. Which equates to an average particle lifetime of $\sqrt{D\tau} = 520$.

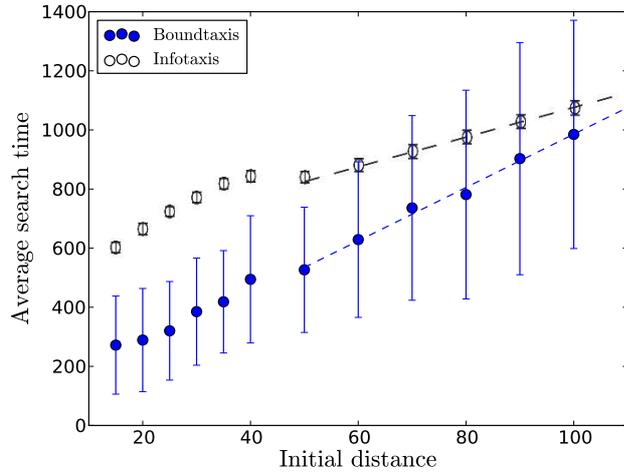


Figure 3.19: Comparison between the performance of Infotaxis and Boundtaxis search times. The error bars indicate the standard deviation of the average search times. While Boundtaxis has consistently faster search times than Infotaxis, there are two interesting and problematic features: (1) the Boundtaxis has a significantly greater variance, (2) the trend for larger initial distances is unfavourable for those greater than 110 (the extent of the figures range). The trend for Boundtaxis is that at larger distances the variance in the search times is almost double that of smaller distances. The Infotaxis data is taken directly from Vergassola *et al.* (2007a).

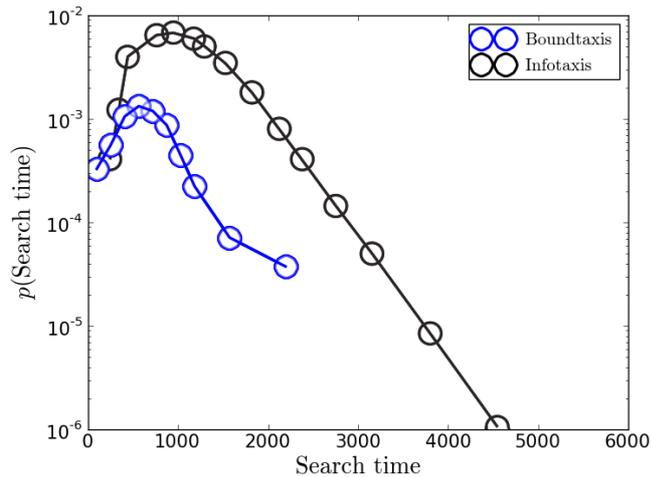


Figure 3.20: Comparison between the probability density functions (PDF) of the search times for Infotaxis and Boundtaxis. The PDF of search times for the Boundtaxis has a much sharper peak than for the Infotaxis policy. The peak of the PDF for the Boundtaxis policy also corresponds to a quicker search time than the peak search time for the Infotaxis policy. The Infotaxis data is taken directly from Vergassola *et al.* (2007a).

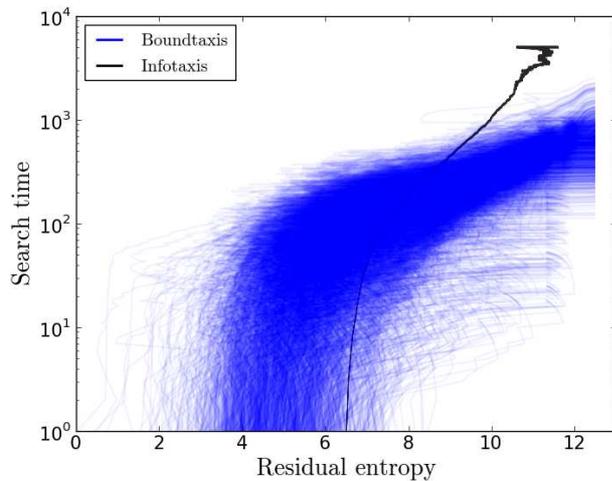


Figure 3.21: Comparison between the residual entropy of the Infotaxis and Boundtaxis posterior beliefs versus the remaining search time. Infotaxis data taken directly from Vergassola *et al.* (2007a).

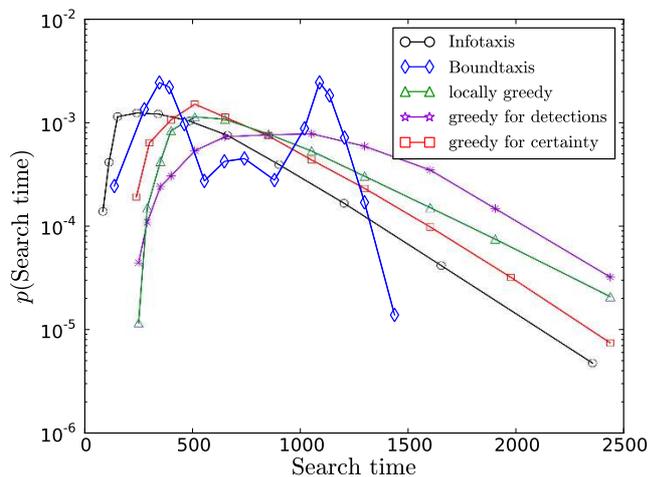
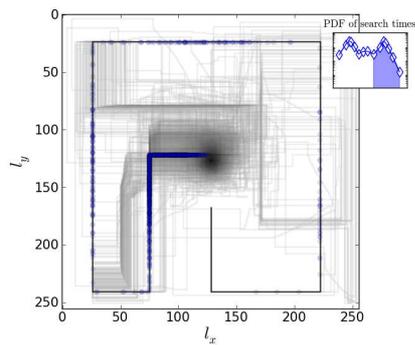
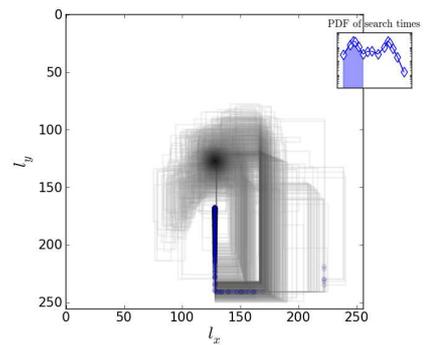


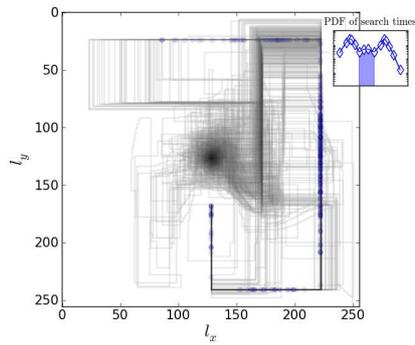
Figure 3.22: Comparison between the probability density functions (PDF) of the search times for Infotaxis, Boundtaxis and a variety of other policies (Vergassola *et al.*, 2007b). The PDF of search times for the Boundtaxis policy has three distinct peaks, this is in contrast to the other policies, which are all unimodal distributions. The other policies are described in the text. The data for Infotaxis and the other comparative policies are taken directly from Vergassola *et al.* (2007b).



(a) Longest range of search times.

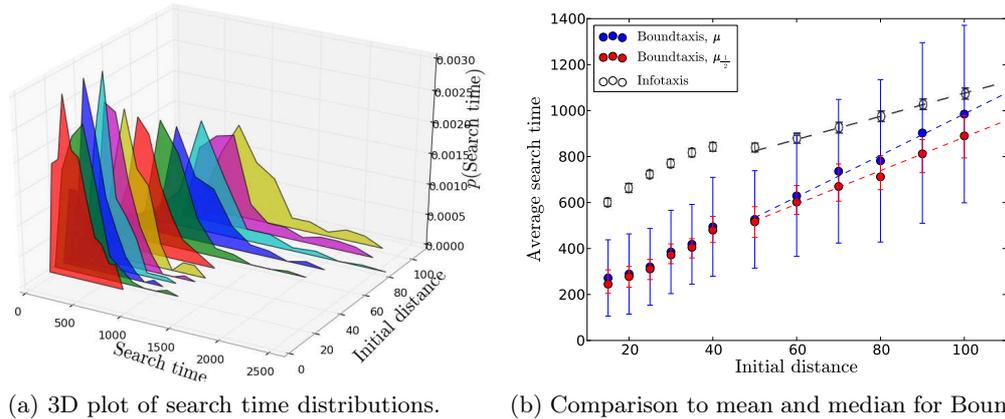


(b) Shortest range of search times.



(c) Midrange of search times.

Figure 3.23: The tracks (grey lines) of individual runs for the Boundtaxis policy. The first detection for each individual track is indicated by an opaque filled blue circle; hence a region of darker blue indicates an area of first detection for the segment of search times. The default behaviour of Boundtaxis seems to be highly exploratory on the field of possible source locations $\{l_x, l_y\} \in L_t$.



(a) 3D plot of search time distributions. (b) Comparison to mean and median for Boundtaxis.

Figure 3.24: The probability distribution function (PDF) of search times with the initial distance. This illustrates that with reducing concentration (increasing initial distances in Figure 3.20) the PDF of search times becomes multi-modal (Figure 3.22). Plot (a) depicts the distribution of search times for each initial distance. Plot (b) extends Figure 3.19 by plotting the data (in red) of the median $\mu_{\frac{1}{2}}$ with the error bars indicating the upper and lower quartiles: 50% of the distribution of search times are within the red error bars. The median in plot (b) indicates a more favourable scaling for the bulk of the search time distributions for Boundtaxis. Plots (a) and (b) explain the sharpness of aggregate PDF of search times in Figure 3.20.

the search times for Infotaxis, Boundtaxis and a variety of other policies (Vergassola *et al.*, 2007b). These other policies are:

- A *locally greedy* strategy where the agent moves towards locations with the highest estimated probability of the source location, $p(l_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$.
- A strategy *greedy for detections*, where the agent moves to maximise the probability of detection, e.g. $p(x_t|x_{0:t-1}, \mathbf{a}_{0:t})$.
- A strategy *greedy for certainty*, where the impact of finding the source is neglected, e.g. $I[L_t|X_t, x_{0:t-1}, \mathbf{a}_{0:t}]$.

The data for Infotaxis and the other comparative policies are taken directly from Vergassola *et al.* (2007b). The PDF of search times for the Boundtaxis policy has three distinct peaks, this is in contrast to the Infotaxis, *locally greedy*, *greedy for detections* and *greedy for certainty* policies which are all unimodal distributions. The three peaks of the Boundtaxis search times are correlated to regions on the latent space where the agent makes its first detection, Figure 3.23. So why does the Boundtaxis make such circuitous paths in Figure 3.23 (a)? In this case the Boundtaxis policy has already explored the perimeter of the latent space before the agent makes its first detection.

The difference between the shapes of the Boundtaxis PDFs of search times in Fig-

ures 3.20 & 3.22 is due to the contrasting parameterisations combined with the degree to which the Boundtaxis will explore the latent space. The particle lifetime of Figure 3.20 is $\sqrt{\tau D} = 50$ whereas for Figure 3.22 is $\sqrt{\tau D} = 20$. The agent’s search time PDFs for each of the different starting distances from Figure 3.19 can be seen in Figure 3.24 (a), which for the larger distances has a second cluster of search times. This shows that the PDF of search times for the larger starting distances has a multi-modal distribution where the second mode is an order of magnitude less likely.

The $B_{X_t|\mathbf{a}_t} [L_t||L'_t]$ quantity alone is similar to the *greedy for certainty*. Though while using $B_{X_t|\mathbf{a}_t} [L_t||L'_t]$ alone did find the source’s location, e.g. $p(\mathbf{l}_t = \mathbf{l}^{\text{true}}|x_{0:t}, \mathbf{a}_{0:t}) \approx 1$, the agent did not capture it. In contrast to this, adding the term which accounts for finding the source $-p(\mathbf{l}_t = \mathbf{a}_t|x_{0:t-1}, \mathbf{a}_{0:t-1}) I [L_t|x_{0:t-1}, \mathbf{a}_{0:t-1}]$, leads to no failures in source capture on any Boundtaxis run. Adding the term accounting for the source’s capture makes the Boundtaxis algorithm an upper bound of $\mathbb{E} [\Delta I|\mathbf{a}_t]$ rather than $I_{\text{MI}} [X_t; L_t|x_{0:t-1}, \mathbf{a}_{0:t}]$. Finally, the Boundtaxis algorithm’s probability of failure is less than 2 in 10^{-4} as it did not fail to capture the source in any of the computed runs.

To summarise:

- Boundtaxis has more variance in its distribution of search times than Infotaxis (Figure 3.19). Though the bulk of these search times (50% of runs) are significantly quicker than Infotaxis (Figure 3.24).
- Boundtaxis is typically quicker to calculate and scales more favourably, see Algorithms 3.7 & 3.8.
- Boundtaxis makes a larger initial spiral than Infotaxis (Figure 3.18). From Figures 3.19-3.21 and Figure 3.24 we can see Boundtaxis can be expected to outperform Infotaxis. Though for a different configuration depicted in Figure 3.22, this cannot be guaranteed.
- Boundtaxis is not reliant upon the assumption that the marginal $p(x_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$ is unimodal, which is an assumption of Infotaxis in Vergassola *et al.* (2007a,b).

3.8 Discussion

3.8.1 Findings

I have developed an alternative method to Infomax (Shannon information and mutual information) for solving the sensor placement problem satisfactorily. I have considered the mathematical consequences of my solution in Section 3.4 and the algorithmic complexity in its estimation when contrasted to a state-of-the-art approach to Infomax (Section 3.5). I then compare it to Infomax in two problems. The first is a toy perception problem related to the previous chapter (Section 3.6). The second is a state-of-the-art system from the literature (Vergassola *et al.*, 2007a) in Section 3.7.

In this Chapter I have developed and presented a tool to explain generally why and how perceptual ambiguity is collapsed by a choice of action. My work relates to the prior work of Fox *et al.* (1998) and Cassandra *et al.* (1996) by considering approximations to the more general localisation problem rather than a robot’s self localisation upon a map. This leads to a general method for distinguishing most between latent posterior predictions, and consequently to collapse localisation ambiguity in the sensor placement problem.

The group of domains, where my approach has the potential to be superior to Infomax, have the properties that the likelihood $p(x_t|l_t, a_t)$ is non-linear. As a consequence it is necessary in applying Infomax to approximate $I[L_t|X_t, a_t]$ and $I_{MI}[X_t; L_t|a_t]$. For instance, two approaches taken in the literature are:

1. approximating of $p(x_t|x_{0:t-1}, a_{0:t})$ as in Vergassola *et al.* (2007a)
2. or sampling the distribution $p(x_t, l_t|x_{0:t-1}, a_{0:t})$ as in Porta *et al.* (2005).

An example of such a domain is sound source localisation. As seen in Section 3.6, the target location (latent space) can typically be treated as a linear process and the mapping from location to sound level and time disparity are non-linear and many-to-one.

How is the Infotaxis of Vergassola *et al.* (2007a) related to the entropy action selection of Fox *et al.* (1998)? One assumption by both methodologies is that either the target or environment is static and unchanging. This is achieved by neglecting the uncertainty in latent dynamics. This was true of Fox *et al.* (1998) and Porta *et al.* (2005). The model of Vergassola *et al.* (2007a) also has uncertainty in updating the latent state space and only an action directed state change – hence, Vergassola *et al.* assumed deterministic dynamics. As well, the model in Vergassola *et al.* (2007a) assumes the agent knows where it is and that the source will remain static. In most Markov processes there is a diffusion process to

account for uncertainty in the latent state space dynamics. This is where the agent cannot be certain that its actions will succeed, or that the target or scene is static.

The conditional Shannon information is typically incomputable without some approximation Porta *et al.* (2005), Fox *et al.* (1998), Vergassola *et al.* (2007a). In developing my method I have presented a valid alternative to Shannon information (Infomax) for satisfactorily solving the problem of optimal action selection. My framework makes action selection easier and faster if the averaged quantities are analytic, making it less computationally costly when compared with Infomax approaches like Fox *et al.* (1998), Porta *et al.* (2005), Cassandra *et al.* (1996), Vergassola *et al.* (2007a).

In Section 3.3 I have shown that there exists an upper bound, upon the conditional Shannon information of the form,

$$I[\mathbf{L}_t | \mathbf{X}_t, a_t] \leq I[\mathbf{L}_t | a_t] + B_{\mathbf{X}_t | a_t}[\mathbf{L}_t || \mathbf{L}'_t], \quad (3.69)$$

and the mutual information,

$$I_{\text{MI}}[\mathbf{X}_t; \mathbf{L}_t | a_t] \leq B_{\mathbf{X}_t | a_t}[\mathbf{L}_t || \mathbf{L}'_t]. \quad (3.70)$$

This upper bound is fully quantified by Theorem 3.3 as,

$$B_{\mathbf{X}_t | a_t}[\mathbf{L}_t || \mathbf{L}'_t] = I_{\text{MI}}[\mathbf{X}_t; \mathbf{L}_t | a_t] + D_{\text{KL}}[p(\mathbf{X}_t | a_t) p(\mathbf{L}_t | a_t) || p(\mathbf{X}_t, \mathbf{L}_t | a_t)], \quad (3.71)$$

$$\begin{aligned} &= D_{\text{KL}}[p(\mathbf{X}_t, \mathbf{L}_t | a_t) || p(\mathbf{X}_t | a_t) p(\mathbf{L}_t | a_t)] \\ &\quad + D_{\text{KL}}[p(\mathbf{X}_t | a_t) p(\mathbf{L}_t | a_t) || p(\mathbf{X}_t, \mathbf{L}_t | a_t)], \end{aligned} \quad (3.72)$$

which is a measure of the statistical independence of the variables x_t and l_t for a choice of action a_t (Theorems 3.3 & 3.6). $B_{\mathbf{X}_t | a_t}[\mathbf{L}_t || \mathbf{L}'_t]$ is the averaged KL-divergence between the likelihoods of pairs of hypotheses in the measurable space,

$$B_{\mathbf{X}_t | a_t}[\mathbf{L}_t || \mathbf{L}'_t] = \int dl_t dl'_t p(l_t | a_t) p(l'_t | a_t) D_{\text{KL}}[p(\mathbf{X}_t | l_t, a_t) || p(\mathbf{X}_t | l'_t, a_t)]. \quad (3.73)$$

Thus, considering the contribution from the measurable \mathbf{X}_t , the upper bound is minimised when the expected overlap of projected posterior beliefs is minimised (Theorem 3.10).

In Sections 3.6 & 3.7 I have used my measure to evaluate the upper bound of the expected Shannon information as an informatic policy model. These two examples show

that my policy model is an adequate but distinct alternative to Infomax approaches.

3.8.2 Limitations

Vergassola *et al.* (2007a) postulate that these information terms are computed by an individual for localisation. However, particular action policies may be represented satisfactorily with a cheaper approximation. An example of this is the use of finite state machines (Thill and Pearce, 2007). Hence, leading to the question, why does a more expensive information calculation needs to be computed when a simpler method produces an equivalent result and similar choice of action?

For example, in the multimodal case, it is a reflex action of human subjects to seek to bring a sound event within view of their eyes (Blauert, 1997). This is in agreement with my hypothesis, as in this instance, the individual can only increase its localisation certainty with a simple and easy heuristic to select an action.

This is most useful to problems that can be represented behaviourally. Taking as an example, head rotations for auditory localisation, it is easy to argue that *a priori* beliefs can be classified (e.g. in Sections 2.3.2 & 2.3.3) and assigned a probability distribution for the candidate actions. An example of this approach can be seen in Subsection 3.6.2. I classified an *a priori* belief according to its Shannon information as either ambiguous or unambiguous. I used this to construct two stochastic policies according to the informatic policies in Section 3.6. The Infomax policy was well represented by this stochastic strategy, making the stochastic strategy a valid alternative. In contrast my framework had a more complex relationship with *a priori* belief than Infomax, therefore making the stochastic strategy significantly under perform my approach.

As mentioned in the introduction, moth behaviour can be explained using a similar information measure (Vergassola *et al.*, 2007a) or to have an equivalent choice of predicted action using a finite state machine. It is obvious from a neural perspective, that a finite state machine is far cheaper to construct than an entropy based measure. Therefore if a behavioural representation of a policy exists, to construct a decision process for selecting an action equivalent to an information model of policy decision, the agent should use this behavioural representation.

Infomax and related information policy models, such as my approach, can provide insight into more complex localisation tasks. As an example, if a consequence of an action is to capture a target then the best localisation action naturally has the result of being both

exploitative and exploratory. Though as reported in Section 3.7 the independence of x_t and l_t is not sufficient to move into the source. Adding a term to account for the agent capturing the source rectified this and leads to the source’s capture. Therefore with care, information based policies can deal implicitly with the exploitation versus exploration dilemma for both target localisation and capture. Specifically my method does not assume the structure of the marginal $p(x_t|x_{0:t-1}, a_{0:t})$, whereas Vergassola *et al.* imposes a unimodal distribution. In making this assumption they limit the applicability of Infotaxis to a subset of the source capture problems.

I have not considered directly the case of non-linear dynamics. This class of problem has the difficulty of accurately representing the projected posterior belief, though this is the class of problem dealt with in the next Chapter. This difficulty is especially acute for the more chaotic problems, as the causal relations which an observer has learnt may not be general or even accurate. Worse still, depending on the dynamics, in such a problem there is a requirement for a near infinite number of tracked hypotheses to adequately represent the projected belief in the latent state space. However, such problems can be approached if there is conditional structure in the latent state space. A technique to approach such a class of problem is Rao-Blackwellization. I use this approach in the next Chapter.

I purposefully kept the framework’s model (Figure 3.1) simple by neglecting a conditional relationship between the prior latent space and the choice of action. This relationship would be in the form of the observer’s pose constraining which actions are possible. This may require the inclusion of a risk term to prevent catastrophic actions being taken. One potential consequence of this is to have an observer learn its morphology by pruning possible morphologies in a directed manner similarly to Bongard *et al.* (2006), rather than through completely random (undirected) body babbling as Olsson *et al.* (2005a, 2006). A difficulty is that even a simple morphology would have a massive dimensionality making it an intractable problem if approached naively.

3.8.3 Future work & summary

This framework considers the problem of reflexive localisation actions. This complements the work on sensorimotor contingencies of Noe (2004). Sensorimotor contingencies consider the relationship between actions and sensory measurements to map the correlations of these as a latent space Noe (2004), O’Regan and Noe (2001), Bompas and O’Regan (2006), Philipona *et al.* (2003, 2004), Aytekin *et al.* (2008). This however neglects ambi-

guity, which exists in a large number of problems and can cause difficulty for any learning process. Hence an informatic policy model such as mine or Infomax, which by its nature, assumes an uncertain and noisy world will have an advantage when combined with a learning algorithm. Further, my model is useful for a class of problems such as POMDPs or likelihoods where the KL-divergence is analytic by scaling as well as, or better than, other informatic approaches. Examples of this for Infomax are approximating the marginal $p(x_t|x_{0:t-1}, a_{0:t})$ in Vergassola *et al.* (2007a), or assuming a sample of the joint latent and measurement variables $L \cup X$ in Porta *et al.* (2005).

The consequence of this for the animal is that there exists a informatic computation other than Infomax. Meaning that if an animal uses information to make reflexive localising actions, it needs to do so in as fast and generalisable a fashion as possible. My approach is favourable if the animal can represent its environment such that the KL-divergence is separable, Section 3.5. Using an informatic policy an animal can learn and act in a new and unknown environment, or learn to best utilise a new sense (Nagel *et al.*, 2005).¹⁵ While, if the animal’s sensorimotor contingencies do not change, then a heuristic strategy for selecting localising actions would be favourable, for instance Subsection 3.6.2.

A major assumption of my framework and also that of Porta *et al.* (2003, 2005), Vergassola *et al.* (2007a) is that an applicable model exists and has been suitably parameterised. The scenario of having a suitably parameterised model of the system, is typical of most research involving Infomax. However, an agent can use an informatic decision process to enhance its learning process. It is this issue which I consider in the next Chapter.

¹⁵Given the assumption that an information quantity is used to select an action, it is possible to test which measure of knowledge is used by defining a pathological case.

Chapter 4

Learning a model for an active agent

4.1 Introduction

In general, an animal in its natural environment can accurately and robustly perceive the state of the environment around itself. Examples of this are the individual's abilities to see objects and structures within the environment amongst clutter, to focus upon one voice among others, to smell one complex scent among many, and to relate information from these sensory modes to one another. This is in part due to an evident change in the view of the world that an animal imposes by its actions causing a predictable alteration to its view of the scene. An example of this is a head rotation to localise a sound which makes one ear of an individual perceive an increase in loudness and at the same time the other ear to perceive a decrease in loudness (see Chapter 3). Noe (2004) argued that an ability to affect¹ the scene an individual sees, hears and feels leads to what we term perceiving. This is related to the concept of being embodied within the world.

Noe (2004) also argued that an individual's entire understanding of the world is due to its grounding within the world, for instance:

1. an individual's understanding of cause and effect, and
2. the geometric relationship between sections of the retina.

This view is supported by the work of Held and Hein (1963) who showed that a kitten's sensory development was critically dependent upon its ability to direct its body's motion.

¹There are two ways to affect an observed scene:

1. To cause a change in the environment, to affect the scene.
2. To change the view of the world, to affect the observation.

The individual makes a prediction of the consequences of an action, and reinforces the accurate predictions.

O'Regan and Noe (2001) further developed this idea of embodiment by using differential geometry to show how an agent can infer the dimensionality of a contrived problem domain. This is the combined dimensionality of the agent's physiology and environment. Olsson *et al.* (2006, 2005a) investigated this concept more specifically with regards to robot visual perception. They used an information metric to relate a CCD camera's pixels to a visual map. They achieved this by using each pixel pair's mutual information with respect to time during random pans and tilts of the camera. These random pans and tilts are referred to as *body babbling*. Philipona *et al.* (2003, 2004) used a correlation based approach derived from differential geometry to infer the sensory configuration of the agent. The approach of Philipona *et al.* and O'Regan and Noe was extended into the auditory domain by Aytekin *et al.* (2008). Aytekin *et al.* constructed an approach to infer the relationships between location relative to the agent and the frequency domain – making the agent essentially learn the mapping of the latent state space to the measurable cues. These approaches do not consider the problem from a Bayesian perspective but more specifically as a Markov problem. A Bayesian perspective allows the explicit accounting for uncertainty and ambiguity and the relative weighting of competing hypotheses.

An assumption made by Olsson *et al.* (2006, 2005a) was that there exists a period in an individual's life that is specific to learning its own sensory configuration; this is the body babbling phase which is assumed to occur during childhood. An alternative argument that can be developed is that adult sensory systems have a relatively stable configuration. This configuration can be refined until radically different information is presented to it, for example by a change in physiology (Rossetti *et al.*, 1993, 1998, Welch *et al.*, 1993). Rossetti *et al.*, Welch *et al.* have definitively shown that there is significant plasticity in adult humans' sensory perception. They did this through the use of prisms, illustrating rapid visual adaptation, similar to the adaptation caused by Olsson *et al.*'s robotic body babbling. Moreover, Hofman *et al.* (1998) showed adaptation in the human auditory system's spatial tuning to a mould placed upon the subject's pinnae. Similarly, Mrsic-Flogel *et al.* (2001) suggested that neural units in the ferret's auditory cortex are tuned to the physiological peculiarities of an individual's ears.

Hence it is reasonable to posit that (in certain situations or contexts) individuals retain the ability to learn new sensory motor relationships throughout life. In adulthood

specifically, there are examples of adults learning to utilise new senses, for example magnetic perception within the body-modification community through insertion of rare earth magnets into the finger tips (Norton, 2006). Similarly Nagel *et al.* (2005) showed the integration of a belt based compass (giving a tactile measure of magnetic North) for a subject’s sensory enhancement. There are also examples in vision and audition being augmented to create a bistable state of adaptation for conflicting sensory mappings in Hofman *et al.* (1998), Welch *et al.* (1993). There is also the mapping between the motor system and the visual system in Koerdig and Wolpert (2004). Also, using prisms to modify an individual’s visual perception (Rossetti *et al.*, 1993, 1998, Welch *et al.*, 1993), an already utilised sense can be successfully perturbed. Hence, this suggests a general feature of perception. It appears possible to learn to perceive a new sense, and learn modifications to an existing sense, well into adulthood. I take this to mean that there is no special learning phase but more lifelong learning that can refine or radically alter an individual’s sensorimotor contingencies as required.

This Chapter expands upon the work in the previous Chapter 3, by investigating the consequence of information driven action selection to learning in a non-trivial dynamic system. I contrast this with two other measures from the literature which also seek to maximise the agent’s expected knowledge: an Infomax approach adapted from Porta *et al.* (2003, 2005), and an *a posteriori* Cramer Rao Lower Bound (CRLB) from the radar and sonar literature (Hernandez, 2004, Helferty and Mudgett, 1993). Details of each of these measures are described in Section 4.5.

I develop a system to investigate this by first defining the structure of the model I shall work with and the measure used to define the “goodness” of an action according to how much information a consequent measurement will convey. However, to select a set of actions requires a stable parameterisation of the model. I have approached this problem using a sequence of episodes. I define an episodic modification to maximum likelihood learning that allows the previous episode’s knowledge of a model’s parameters to be retained in constructing a new parameterisation.

I then describe in Section 4.3 the non-linear dynamical system (the Lorenz attractor) that I shall use. In Section 4.4 I define the specifics of the model I use to describe this dynamical system. The measure of expected goodness for a choice of action and its interpretation is discussed in Section 4.5. Section 4.6 has 3 parts:

1. the contrast between my approach from Chapter 3 and other informatic approaches,

2. analysis of the utility of localising actions for an agent,
3. then a re-learning of correspondence due to a continued change to the measurement process.

Finally I place these results in context in the discussion. This allows me to state that directed action can construct a better model which is more certain in inference and more likely than undirected action. Moreover directed action coupled with episodic learning allows for re-learning of correspondence.²

4.1.1 Hypothesis

I argue that a good localising choice of action, which is more certain in inference and more likely, will allow a better model to be learnt from a sampled data set than random and undirected actions. Accordingly, an active agent can learn a more optimal model of the world than an inactive agent (Chapter 2) or one that explores in an undirected manner. More strongly, it can be argued from Held and Hein (1963) that active and directed perception is a critical developmental requirement of learning to represent the world. However, what is the mechanism that directs directed perception? My argument is contingent upon the agent having an internalised measure of uncertainty, for example, the measure of knowledge I developed in Chapter 3. My measure constructed in Chapter 3 selects an action to maximise the expected disparity in predicted *a posteriori* beliefs. This leads to a measurement causing the most predicted hypotheses to be discarded due to a measurement. This is contrasted with two other informatic approaches for selecting an action in the sensor placement problem.

What I shall show is that directed actions, selected to increase an agent's expected knowledge of the system's state, is better for learning than undirected action. I will compare between the performance for different models of future knowledge. I then construct a stochastic variation of my method to create a graduated analysis between a uniform random policy and my approach developed in Chapter 3. I will accomplish this by using a linear switching state space model (SSM) (Ghahramani and Hinton, 2000) with the action defining which measure of the state space is observed. This choice of measure is defined as the action. However the correspondence of actions to the measures will be unknown to the observer and have to be relearnt. This Chapter applies the action selection method de-

²The problem of correspondence is a description of how to relate a hidden latent space to individual measurements.

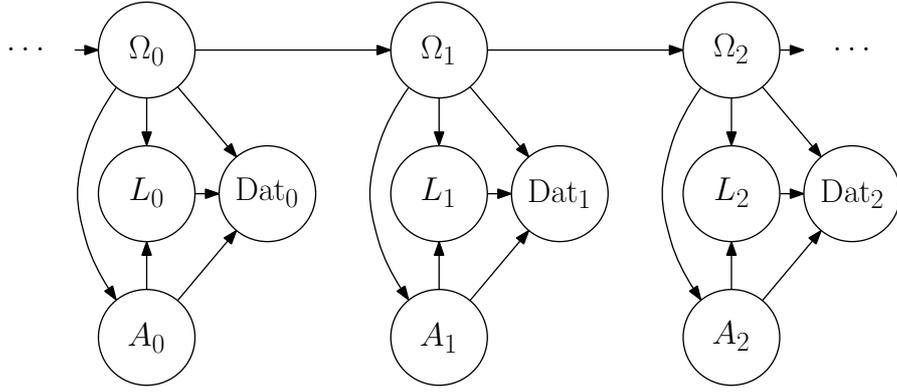


Figure 4.1: The graphical model indicating the action-perception dependencies of a sequence of sample and learn episodes. This figure represents a sequence of cliques that depend upon the n 'th episode's parameter Ω_n . Ω_n represents the agent's model of the system and is used to generate the set of actions A_n that interact with a belief in the latent state space L_n that constructs the set of measurements Dat_n . These are then used to construct the new parameterisation Ω_{n+1} using the arg max operation of Equation 4.1. Making $\Omega_{n+1} = \arg \max_{\Omega} f(\Omega, A_n, \text{Dat}_n, L_n, \Omega_n)$, where f is an objective function such as a measure of the parameter Ω 's likelihood or error. The episode is indicated by number in the subscripts.

veloped in Chapter 3 to a dynamic problem. This in part addresses the problem of an agent learning by active experimentation. This will illustrate the combination of exploration for new knowledge and exploitation of learnt knowledge applied to a learning problem.

4.2 Episodic learning

For an agent that is learning to represent its environment (the system) by exploration, there is the influence of causality. The agent needs a model of the system from which it can construct an informatic policy. Hence, a series of actions are selected according to a model to generate a corresponding series of measurements. Figure 4.1 shows the dependence for this influence between each episode's system model Ω , the choice of actions A and consequent measurements Dat .

4.2.1 Learning requirements

The requirements for a learning algorithm is to learn the parameters of a system which has a hidden latent space that can be constructed from the sequence of measurements $x_{0:T}$ that comprises the data set Dat . The data set is generated sequentially according to a set of actions $a_{0:T}$. As such it is dependent completely upon the parameter Ω due to the recursive selection of each action a_t using an expectation of knowledge, to then generate a corresponding measurement x_t .

To learn a new parameterisation Ω^{new} from this data set and action set I use maximum likelihood (ML), hence applying argument maximisation to a model’s log-likelihood,

$$\Omega^{\text{new}} = \arg \max_{\Omega} \mathcal{L}(\Omega; a_{0:T}, x_{0:T}), \quad (4.1)$$

where $\mathcal{L}(\Omega; \text{Dat}) = \log p(\text{Dat}|\Omega)$. Though any measure of goodness such as prediction error $e_t = (\mathbb{E}[x_t] - x_t)^2$, could be used rather than the likelihood.

Typically however, this neglects any prior parameterisation of θ which contains knowledge of the past data and action sets. This would be akin to an individual waking up each day with a sensory system that reconfigures its connections and weights to optimally represent only that day’s actions and measurement. However if all prior actions and data could be accounted for then there would be an episodic refinement to the parameter learning. An example of an episodic refinement using the expectation-maximisation (EM) algorithm is discussed in the next section.

4.2.2 Episodic maximum-likelihood

Episodic learning uses a prior parameter to help construct a current estimate for this parameter for a data set. To do this I utilise a Lagrangian based upon a measure of dissimilarity between these parameters, the prior parameter and a new estimate, to derive the learning rules. I use the EM-algorithm for this learning process due to its utility as a pseudo Bayesian learning method and its wide acceptance for constructing optimal parameters.

Retaining the previous episode’s optimal parameterisation gives a way of retaining knowledge of past data without needing to maintain a copy of the past data set. This can be a significant space saving for a small loss with respect to optimality. The degree of influence upon the current parameterisation is managed via the magnitude of the Lagrangian multipliers. This can be interpreted by contrast to standard EM. EM sequentially optimises a bound $B[\Omega; \Omega^{\text{old}}]$ upon the likelihood $\mathcal{L}(\text{Dat}|\Omega)$ of the data Dat for a parameterisation Ω over a set of latent variables L , which are related by,

$$\mathcal{L}(\text{Dat}|\Omega) = \log p(\text{Dat}|\Omega) \quad (4.2)$$

$$= \log \int dL p(\text{Dat}, L|\Omega) \quad (4.3)$$

$$\leq \int dL p(L|\text{Dat}, \Omega^{\text{old}}) \log \frac{p(\text{Dat}, L|\Omega)}{p(L|\text{Dat}, \Omega^{\text{old}})} \quad (4.4)$$

$$= B \left[\Omega; \Omega^{\text{old}} \right], \quad (4.5)$$

when the parameters have converged the bound will tend to the likelihood (Appendix A). The episodic EM modifies the bound by a Lagrangian multiplier γ_Ω and a measure of dissimilarity $d(\Omega; \bar{\Omega})$ between the parameterisation Ω and the prior parameter $\bar{\Omega}$ to become,

$$G \left[\Omega; \Omega^{\text{old}} | \bar{\Omega} \right] = \int dL p \left(L | \text{Dat}, \Omega^{\text{old}} \right) \log \frac{p(\text{Dat}, L | \Omega)}{p(L | \text{Dat}, \Omega^{\text{old}})} + \gamma_\Omega d(\Omega; \bar{\Omega}), \quad (4.6)$$

where the prior parameter is distinct from the old update for standard EM. This suggests that, if we assume a prior $p(\Omega) \triangleq \frac{1}{Z_{\bar{\Omega}}} e^{\gamma_\Omega d(\Omega; \bar{\Omega})}$ where $Z_{\bar{\Omega}} = \int d\Omega e^{\gamma_\Omega d(\Omega; \bar{\Omega})}$, then the joint belief of the parameterisation and the data is,

$$\mathcal{J}(\text{Dat}, \Omega) = \log p(\text{Dat}, \Omega) \quad (4.7)$$

$$= \log p(\text{Dat} | \Omega) + \gamma_\Omega d(\Omega; \bar{\Omega}) - Z_{\bar{\Omega}}. \quad (4.8)$$

Hence such a Lagrangian multiplier is equivalent to applying a prior belief to the likelihood of Ω (Equations 4.2 to 4.5) making any optimisation of $G \left[\Omega; \Omega^{\text{old}} | \bar{\Omega} \right]$ an optimisation of the joint belief $\mathcal{J}(\text{Dat}, \Omega)$.

The methodology of the EM-algorithm is to first construct a proposal distribution over the hidden latent variables L conditioned upon the data Dat and the prior parameters Ω^{old} , $p(L | \text{Dat}, \Omega^{\text{old}})$. This constructs the expectation or E-step of the EM-algorithm. The EM-algorithm finds the stationary point of the bounds Lagrangian $G \left[\Omega; \Omega^{\text{old}} | \bar{\Omega} \right]$ with respect to the parameters Ω , which maximises the joint belief $\mathcal{J}(\text{Dat}, \Omega)$. This constructs the new parameters and so constitutes the maximisation or M-step of the EM-algorithm. The EM-algorithm is conceptually a pseudo maximum likelihood where the expectation $p(L | \text{Dat}, \Omega^{\text{old}})$ is used in the maximisation of $p(\text{Dat}, L | \Omega)$ with respect to Ω and the prior defined by the Lagrangian multiplier $\gamma_\Omega d(\Omega, \bar{\Omega})$. In the following I use the gradients to construct analytical updates for the learning rules as stationary points to the episodic Lagrangian $G \left[\Omega; \Omega^{\text{old}} | \bar{\Omega} \right]$. This is somewhat like an iterative maximum a-posteriori (MAP) algorithm rather than an ML-algorithm. Though implicitly all ML algorithms assume a uniform prior for the parameters.

An episodic ML-algorithm optimisation for a unimodal Gaussian process is depicted in Figure 4.2. To emphasise the flexibility of the episodic process, a transition in the parameters at episode 80 illustrates the influence through γ of the weight assigned to a

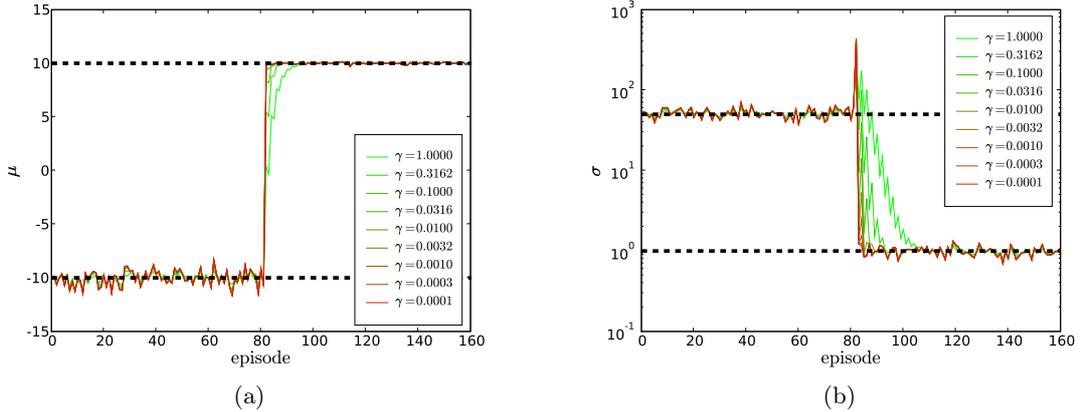


Figure 4.2: Convergence of the episodic maximum-likelihood (ML) estimates. The episodic ML estimates are compared according to the ratio γ which indicates the weighting given to the previous episodes parameters. At episode 80 there is a transition in the data from one parameter to another. The purpose of this transition is to indicate the varying speed with which the different γ 's converge to the new parameters. The true parameters are indicated with dashed lines. For $\gamma \rightarrow 1$ acts as a smoothed variational estimate for the parameters. Where each episode's sample size is $N = 100$ and $\gamma\Omega = \gamma N$.

past episode's parameters. The episodic ML-algorithm converges to the true parameters at an episodic rate dependent upon γ .

The advantage of this *episodic EM-algorithm* is that it may be applied to any model that is learnable using a likelihood based method, including expectation-maximisation based variational learning methods such as Ghahramani and Beal (2000), Ghahramani and Hinton (2000), Jordan *et al.* (1999), Verbeek *et al.* (2003), Beal *et al.* (2003), Hospedales and Vijayakumar (2006), Hospedales *et al.* (2007). It does however require that a suitable measure of dissimilarity exists for the particular parameter being learnt. I develop a few measures and consider the application of one to a ML-algorithm (Figure 4.2) in Appendix R.

4.3 Systems background

The seminal discovery of Lorenz that originated the field of chaos was originally motivated by the problem of weather prediction. Lorenz investigated a series of equations derived from thermal convection in the atmosphere (Ott, 2002). His demonstration that thermally driven convection could result in chaos raised the possibility that the atmosphere is chaotic. The dynamic equations for the Lorenz attractor are,

$$\frac{\partial x^{(1)}}{\partial t} = a(x^{(2)} - x^{(1)}), \quad (4.9)$$

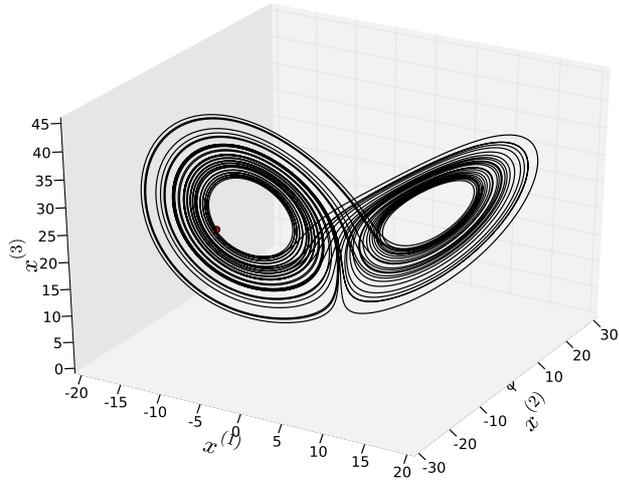


Figure 4.3: The Lorenz attractor. The parameters for the attractor are $a = 10$, $b = 25$ and $c = \frac{8}{3}$. The starting point is indicated by a red dot. Sampling this trajectory was implemented using the Python scientific package Scipy (using the included ODE solver).

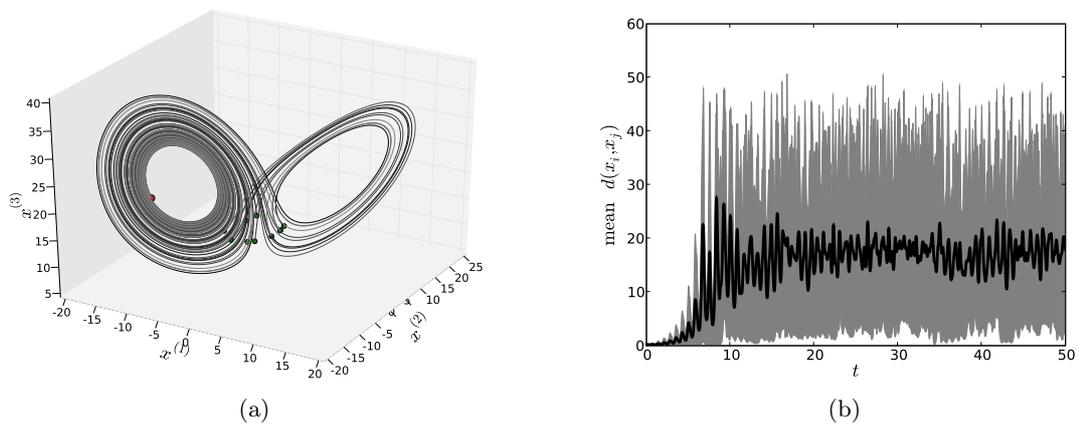


Figure 4.4: The sensitivity of the Lorenz attractor to small perturbations of the initial starting conditions. A small collection of similar starting conditions is indicated in plot (a) by a red dot. The trajectories from these initial points are also indicated in the same plot. Plot (b) shows this sensitivity more clearly in depicting the average distance between the trajectories in plot (a) by the black line and the limits of these distances are indicated by grey block in plot (b). Sampling these trajectories was implemented using the Python scientific package Scipy (using the included ODE solver).

$$\frac{\partial x^{(2)}}{\partial t} = (b - x^{(3)})x^{(1)} - x^{(2)}, \quad (4.10)$$

$$\frac{\partial x^{(3)}}{\partial t} = x^{(1)}x^{(2)} - cx^{(3)}, \quad (4.11)$$

where the state vector $\mathbf{x} = \{x^{(i)} : i = 1, 2, 3\}$. See Figure 4.3 for an example attractor for the parameters $a = 10$, $b = 25$ and $c = \frac{8}{3}$ implemented using the Python scientific package Scipy (using one of the included *ordinary differential equation* (ODE) solvers, see Appendix Q). The Lorenz attractor is the classic example of a chaotic dynamical system. It is for this reason that I choose to use it as the state dynamics for my problem.

The exponential sensitivity of chaotic trajectories means that as time progresses small errors in the solution grow exponentially with time. Thus, after some time, effects such as noise and computer rounding can cause significant divergence between a solution and what it would be in the absence of these effects. An example for the Lorenz attractor (implemented using Scipy) can be seen in Figure 4.4. The fact that chaos makes prediction past a certain time difficult, is relevant principally for making specific predictions as to what the future state of the system shall be. This is not such a problem for filtering (Grewal and Andrews, 2001), as the problem of uncertainty due to chaotic divergence is countered by conditioning upon a sequence of measurements.

The Lorenz attractor has a dynamic when converted to discrete time of,

$$\mathbf{x}_t = F(\mathbf{x}_{t-1}), \quad (4.12)$$

where $F(\cdot)$ is a non-linear process, making a naive application of a linear process inapplicable, e.g.,

$$\mathbf{x}_t \neq A\mathbf{x}_{t-1} + b. \quad (4.13)$$

However there are regions of the attractor (Figure 4.3) which can be said to be locally linear where

$$\mathbf{x}_t = A_k\mathbf{x}_{t-1} + b_k. \quad (4.14)$$

This leads to modelling the dynamics of \mathbf{x} with a hidden variable k that defines which A_k and b_k to use to evolve \mathbf{x} . The structure of this approximation can be defined from a Bayesian perspective and will be approached in the next Section.

4.4 Partially observable switched state space model

4.4.1 Background

Most common probability models for tracking time series are developed from discrete hidden Markov models (HMM) (Ghahramani, 1998) or stochastic linear dynamical systems which are also referred to as state-space models (SSM) (Sontag, 1999, Kalman, 1960). The latent state in HMMs is usually represented as a sequence of discrete random variables.³ This posterior probability of the state is constructed from transitioning the previous hidden state via a transition matrix and conditioning upon an observation. Knowing the state at any time makes the past, present and future observations statistically independent. This however depends upon complete knowledge of the present.

This conditional independence of the future and the past, given the present, is the property that defines Markov models. SSMs represent information about the past using a continuous hidden state vector. Like HMMs an SSM is a Markov model; conditioning upon the state vector makes the past, present and future observations statistically independent. The dependency between the current and previous state vector is specified via a dynamic equation and a noise model to represent a system.⁴

Most real processes cannot be described completely as either discrete or continuous processes. An example would be a manoeuvring aircraft (Ristic *et al.*, 2004c). This has discrete modes of manoeuvre, each with a dynamic that is approximately linear. Such processes however can be expressed as a switched SSM (Ghahramani and Hinton, 2000). This addresses those processes that can be construed as having locally linear dynamics.⁵ Switched SSMs are a generalisation of HMMs and SSMs in which a continuous dynamic process can transition in a discrete manner from one linear regime to another. Signal processing and other fields make wide use of such models (Ghahramani and Beal, 2000, Ghahramani and Hinton, 2000, Ristic *et al.*, 2004c). Though one difficulty is the explosion in computational complexity with the number of discrete states and the number of time steps.

³HMMs are similar conceptually to symbolic dynamics (Smale, 1967, Levi, 1981). Symbolic dynamics will represent a systems dynamical process using a set of discrete symbols representing the system's state entering a portion of the state space. HMMs represent a discrete state space using a tabulated posterior belief in the discrete states conditional upon a set of observables.

⁴When the dynamic equations are linear and the noise model Gaussian, the SSM is called a Kalman filter.

⁵In dynamics as discussed by Ott (2002) this is equivalent to the mixing of embedding and symbolic dynamics which can be used for prediction (Farmer and Sidorowich, 1987, Abarbanel *et al.*, 1990, Poggio and Girosi, 1990, Linsay, 1991), noise removal (Kostelich and Yorke, 1988) and process control (Ott *et al.*, 1990, Shinbrot *et al.*, 1990, Ditto *et al.*, 1990, Dressler and Nitsche, 1992, Ott, 2002).

As many continuous hidden Markov problems are not possible to solve accurately using either HMMs or SSMs alone, switching SSMs can be used to account for non-linear dynamics. However there is a problem with the number of required tracked states with each time-step (Murphy and Russell, 2001). Instead, approximating the belief in the model's state can be constructed using techniques such as importance sampling. The development of importance sampling methods allowed many previously intractable non-linear problems to be numerically modelled with relative ease. Importance sampling methods, when applied to Markov problems, are typically referred to as either Markov chain Monte Carlo (MCMC) or particle filtering (PF) (Isard and Blake, 1998, Ristic *et al.*, 2004c, Doucet *et al.*, 2001). PFs can represent distributions of arbitrary shape over a sequence of measurements. A PF model maintains the probability of the state space (be it discrete or continuous) by using a set of samples weighted by importance. Using this technique allows computational resources to be focused upon regions of the state space that the model deems to be more likely. See Appendix S for a more formal definition of a PF.

In high dimensional state spaces PFs can be inefficient (Doucet *et al.*, 2000, Murphy and Russell, 2001). A technique used to increase the efficiency of the particle filter is to reduce the size of the sampled state space by marginalising out some variables analytically. This takes advantage of the conditionality in the state space. The marginalised variables can then be constructed analytically using HMMs, SSMs or other filtering methods. Such a technique is referred to as Rao-Blackwellization and for my problem results in a Rao-Blackwellized particle filter (RBPF) (Murphy and Russell, 2001).

4.4.2 System models algorithmic definition

Using a state space L_t , I wish to compute the distribution $p(l_t|x_{0:t}, a_{0:t})$ where l_t is the latent state at a time t , $x_{0:t}$ is the sequence of measurements and $a_{0:t}$ is the sequence of actions. In general, the required integrals cannot be computed in closed form. Therefore I shall use an RBPF which approximates a part of the posterior using sequential importance sampling and calculate the rest analytically.

Suppose the state space L_t is partitioned into two sub-spaces K_t and Z_t , where the variable k_t represents a switching process and the variable \mathbf{z}_t represents a continuous linear process of D_z dimensions. There is also a further hidden variable j_t which is used to denote uncertainty in the choice of measure an action a_t elicits. Hence the problem of estimating $P(j_t|a_t)$ is the same as the correspondence problem discussed in Chapter 2. Then, by the

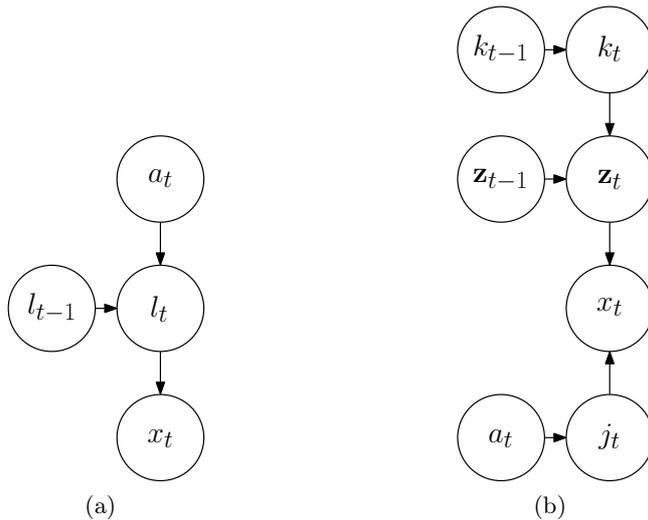


Figure 4.5: Graphical models representing a generalised dynamic transition process and a switching dynamic transition process with similar measurement processes. Plot (a) depicts the latent state variable l_t with the measurement process $p(x_t|l_t)$. Plot (b) depicts a switching Markov process with the switching variable k_t and a continuous variable \mathbf{z}_t with the measurement process $p(x_t|l_t = \{k_t, \mathbf{z}_t, j_t\}) \equiv p(x_t|\mathbf{z}_t, j_t)$. The node labelled j_t indicates our uncertainty as to which measure is actually conducted. The node labelled a_t in both plots represents the choice of measure in sampling x_t . If $l_t = \{k_t, \mathbf{z}_t, j_t\}$ then plot (b) is a special case of plot (a), cf Figure 3.2.

chain rule for probability I may write

$$p(\mathbf{z}_{0:t}, k_{0:t}, j_{0:t}|x_{0:t}, a_{0:t}) = p(\mathbf{z}_{0:t}|k_{0:t}, j_{0:t}, x_{0:t}) P(k_{0:t}, j_{0:t}|x_{0:t}, a_{0:t}), \quad (4.15)$$

if $p(\mathbf{z}_{0:t}|k_{0:t}, j_{0:t}, x_{0:t})$ can be updated efficiently and analytically then only the distribution $P(k_{0:t}, j_{0:t}|x_{0:t}, a_{0:t})$ need be importance sampled (Murphy and Russell, 2001). As this is sampling a smaller space, there is a need for fewer particles to represent the space of $\mathbf{K}_{0:t} \cup \mathbf{J}_{0:t}$ in contrast to $\mathbf{Z}_{0:t} \cup \mathbf{K}_{0:t} \cup \mathbf{J}_{0:t}$.

In a RBPF each particle $k_{0:t}^{(i)}, j_{0:t}^{(i)}$ is a sample from $P(k_{0:t}, j_{0:t}|x_{0:t}, a_{0:t})$. Each particle has an associated parametric representation of the distribution $p(\mathbf{z}_{0:t}|k_{0:t}, j_{0:t}, x_{0:t})$ denoted as $\alpha_t^{(i)}$. The samples from the space $\mathbf{K}_t \cup \mathbf{J}_t$ are updated using standard PF, then the parametric conditional distributions for z_t are updated using an exact filter. This general RBPF algorithm is expressed in Algorithm 4.1. A property of this algorithm is that at any moment in time the particle filtering component is smoothed. A consequence of this is that it is unnecessary to construct a backwards filter to generate a smoothed set of particles. However the $\alpha_t^{(i)}$'s do need to be smoothed to produce a new parameter set to represent the distribution $p(\mathbf{z}_{0:T}|k_{0:T}^{(i)}, j_{0:T}^{(i)}, x_{0:T})$. Smoothed distributions are useful for pseudo Bayesian learning via ML-algorithms such as the EM-algorithm. These distribu-

Algorithm 4.1 A generic Rao-Blackwellized particle filtering algorithm. A shorthand $s = \{j, k\}$ is used, this has the consequence that $S_t = \mathbf{J}_t \cup \mathbf{K}_t$, $s_t^{(i)} = \{j_t^{(i)}, k_t^{(i)}\}$ and $s_{0:t}^{(i)} = \{j_{0:t}^{(i)}, k_{0:t}^{(i)}\}$. If references to s_t are replaced with the full latent space $l_t = \{j_t, k_t, \mathbf{z}_t\}$ and step 3 is omitted, the result is a regular (non Rao-Blackwellized) particle filter. Notice that if only a filtering solution is desired then there is no need to store the full trajectories $s_{0:t}^{(i)}$ for each particle, just its most recent component, $s_t^{(i)}$, since $\alpha_t^{(i)}$ is updated online ($\alpha_t^{(i)}$ forms a parametric representation of $p(\mathbf{z}_t | s_{0:t}^{(i)}, x_{0:t})$). This algorithm is adapted from Murphy and Russell (2001) who in turn adapted it from Doucet *et al.* (2000).

1. Sequential importance sampling step,

- for $i \in [1, N]$, sample

$$s_t^{*(i)} \sim \hat{P}\left(s_t | s_{0:t-1}^{(i)}, x_{0:t-1}, a_t\right), \quad (4.16)$$

and set

$$s_{0:t}^{*(i)} \triangleq \left\{s_t^{*(i)}, s_{0:t-1}^{(i)}\right\}. \quad (4.17)$$

- for $i \in [1, N]$, evaluate the importance of the proposed trajectory according to,

$$\pi_t^{(i)} \propto p\left(x_t | s_{0:t}^{*(i)}, x_{0:t-1}\right) \frac{P\left(s_t^{*(i)} | s_{0:t-1}^{(i)}, x_{0:t-1}, a_t\right)}{\hat{P}\left(s_t^{*(i)} | s_{0:t-1}^{(i)}, x_{0:t-1}, a_t\right)}, \quad (4.18)$$

where these weights are normalised, $\sum_{i=1}^N \pi_t^{(i)} = 1$.

2. Selection step,

- resample N samples from $s_{0:t}^{*(i)}$ according to the importance distribution $\pi_t^{(i)}$ to obtain N random samples $s_{0:t}^{(i)}$ approximating the distribution $p(s_{0:t} | x_{0:t}, a_{0:t})$.

3. Exact step,

- update $\alpha_t^{(i)}$ given $\alpha_{t-1}^{(i)}$, $s_t^{(i)}$, $s_{t-1}^{(i)}$, x_t and a_t .
-

tions represent the current parameter's complete data posterior belief in the hidden latent state space, over which, the EM-algorithm constructs a pseudo maximum-likelihood. In Appendix V the EM-updates for the smoothed Kalman process is constructed variationally using the analytic portion of the RBPF derived in Appendix U. Appendix U constructs the filtered distribution $p(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, x_{0:t})$, followed by the smoothed distribution $p(\mathbf{z}_t | k_{0:T}^{(i)}, j_{0:T}^{(i)}, x_{0:T})$ and the smoothed cross distribution $p(\mathbf{z}_{t-1:t} | k_{0:T}^{(i)}, j_{0:T}^{(i)}, x_{0:T})$. The importance sampling terms $k_{0:T}^{(i)}$ and $j_{0:T}^{(i)}$ that are used to construct these distributions are stated in Appendix T.

4.5 Measures of expected knowledge

The prediction of expected knowledge requires the agent to construct prediction distributions for the latent variables $l_t = \{\mathbf{z}_t, j_t, k_t\}$. Hence, I shall construct the distribution $p(\mathbf{z}_t, j_t, k_t | x_{0:t-1}, a_{0:t})$. To derive this I first start with the joint distribution which by Bayes law factorises as

$$p(\mathbf{z}_t, j_{0:t}, k_{0:t} | x_{0:t-1}, a_{0:t}) = p(\mathbf{z}_t | j_{0:t}, k_{0:t}, x_{0:t-1}) P(j_{0:t}, k_{0:t} | x_{0:t-1}, a_{0:t}). \quad (4.19)$$

By Bayes law the distribution of $j_{0:t}$ and $k_{0:t}$ factorises as

$$P(j_{0:t}, k_{0:t} | x_{0:t-1}, a_{0:t}) = P(j_t | a_t) P(k_t | k_{t-1}) P(j_{0:t-1}, k_{0:t-1} | x_{0:t-1}, a_{0:t-1}), \quad (4.20)$$

and as this is a PF, with the definition that

$$P(j_{0:t-1}, k_{0:t-1} | x_{0:t-1}, a_{0:t-1}) = \frac{1}{N} \sum_{i=1}^N \delta_{j_{0:t-1}}^{j_{0:t-1}^{(i)}} \delta_{k_{0:t-1}}^{k_{0:t-1}^{(i)}} \equiv \frac{1}{N} \sum_{i=1}^N \prod_{t'=0}^{t-1} \delta_{j_{t'}}^{j_{t'}^{(i)}} \delta_{k_{t'}}^{k_{t'}^{(i)}}. \quad (4.21)$$

I use the last two equations to construct the prediction distribution

$$p(l_t = \{\mathbf{z}_t, j_t, k_t\} | x_{0:t-1}, a_{0:t}) = \sum_{j_{0:t-1}, k_{0:t-1}} p(\mathbf{z}_t, j_{0:t}, k_{0:t} | x_{0:t-1}, a_{0:t}), \quad (4.22)$$

$$= \sum_{j_{0:t-1}, k_{0:t-1}} p(\mathbf{z}_t | j_{0:t}, k_{0:t}, x_{0:t-1}) P(j_t | a_t) P(k_t | k_{t-1}) \times P(j_{0:t-1}, k_{0:t-1} | x_{0:t-1}, a_{0:t-1}), \quad (4.23)$$

$$= \frac{1}{N} P(j_t | a_t) \sum_{i=1}^N p(\mathbf{z}_t | j_t, k_t, j_{0:t-1}^{(i)}, k_{0:t-1}^{(i)}, x_{0:t-1}) \times P(k_t | k_{t-1}^{(i)}), \quad (4.24)$$

$$\approx \frac{1}{N} P(j_t|a_t) \sum_{i=1}^N p\left(\mathbf{z}_t | k_{0:t}^{*(i)}, j_{0:t-1}^{(i)}, x_{0:t-1}\right) \times \frac{P\left(k_t^{*(i)} | k_{t-1}^{(i)}\right)}{\hat{P}\left(k_t^{*(i)} | k_{t-1}^{(i)}\right)} \delta_{k_t}^{k_t^{*(i)}}, \quad (4.25)$$

where as can be seen from Algorithms 4.1 & T.1 that the sample $k_t^{*(i)}$ is taken from $\hat{P}\left(k_t | k_{t-1}^{(i)}\right)$, which accounts for the ratio of $\frac{P}{\hat{P}}$. These equations construct a prediction representing the belief $p(l_t | x_{0:t-1}, a_{0:t})$ and is used in the following subsection to estimate the amount of knowledge an action will give the agent.

4.5.1 Adaptation of Porta *et al.*'s approach (Infomax)

There are many methods for approximating the conditional Shannon information. In this subsection I shall use the approach discussed in Porta *et al.* (2003, 2005), as it is a recent addition to the literature. Using the assumptions from Porta *et al.* I construct a sample distribution S to represent the joint space of $A \cup L \cup X$, where A is the space of possible actions. The addition of actions to S is necessitated by the dependency illustrated in Figure 4.5 between the variables a_t and x_t . The sample distribution S is used to construct a set X_{a_t} , for each unique $x \in X_{a_t}$ construct the marginal distribution $p(x|a_t)$. Each $p(x|a_t)$ represents the sum of the weights for the particle that is best attributed to the pair x and a_t . This is used, as in Porta *et al.* (2003, 2005), to approximate the conditioned Shannon information

$$I[L_t | X_t, a_t] \approx \sum_{x \in X_{a_t}} \sum_{i=1}^N \sum_{j=1}^J \frac{1}{N} \frac{P\left(k_t^{*(i)} | k_{t-1}^{(i)}\right)}{\hat{P}\left(k_t^{*(i)} | k_{t-1}^{(i)}\right)} P(j|a_t) \phi\left(x; B_j \hat{\mathbf{z}}_{t|t-1}^{(i)}, B_j \Sigma_{t|t-1}^{(i)} B_j^T + \nu\right) \times \log\left(\frac{1}{N} \frac{P\left(k_t^{*(i)} | k_{t-1}^{(i)}\right)}{\hat{P}\left(k_t^{*(i)} | k_{t-1}^{(i)}\right)} P(j|a_t) \phi\left(x; B_j \hat{\mathbf{z}}_{t|t-1}^{(i)}, B_j \Sigma_{t|t-1}^{(i)} B_j^T + \nu\right)\right). \quad (4.26)$$

This is a costly Algorithm to compute, as within the summation is Gaussian computation. The Algorithm to compute this is expressed in Algorithm 4.2 and scales as $O(N_a N^2 J)$.

4.5.2 Application of CRLB

The CRLB is the expected error or variance in the estimate of a continuous latent state space,

$$\mathbb{E}_{p(x_t, \mathbf{z}_t | x_{0:t-1}, a_{0:t})} \left[(\mathbf{z}_t - \boldsymbol{\mu}_t(x_t))^2 \right] = \mathbb{E}_{p(\mathbf{z}_t | x_{0:t-1}, a_{0:t})} \left[\mathbf{z}_t^2 \right] - \mathbb{E}_{p(x_t | x_{0:t-1}, a_{0:t})} \left[\boldsymbol{\mu}_t^2(x_t) \right], \quad (4.27)$$

Algorithm 4.2 Porta *et al.*'s Infomax calculation algorithm adapted to a partially observable switching SSM. This scales with a time complexity of $O(N_a N^2 J)$, where there are N Rao-Blackwellized particles representing the latent belief $p(l_t = \{\mathbf{z}_t, k_t, j_t\} | x_{0:t-1}, a_{0:t})$, J choices of measure and N_a possible choices of action a_t . S is constructed from the previous episodes learnt distribution, as the individual trajectories upon the space Z . For every time t from 1 to T , each point of the RBPF trajectory is assigned to the corresponding measurement x_t and chosen action a_t . The storage scales as $O(N_S)$ where $N_S = NT$. The shorthand $r(i) = \frac{P(k_t^{*(i)} | k_{t-1}^{(i)})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)})}$ is used in the Algorithm.

Calculation of Porta *et al.*'s method adapted to a switching SSM,

1. for each $i \in [1, N]$

$$r(i) = \frac{P(k_t^{*(i)} | k_{t-1}^{(i)})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)})}$$

$$\hat{\mathbf{z}}_{t|t-1}^{(i)} = A_{k^{(i)}} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + b_{k^{(i)}}$$

$$\Sigma_{t|t-1}^{(i)} = A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{(i)T} + Q$$

2. for each a_t

$$X_{a_t} = \emptyset$$

- (a) for each $i \in [1, N]$

$$(\mathbf{z}, a, x) \in S \text{ for a minimum } \|\mathbf{z} - \mathbf{z}_t^{(i)}\| \text{ and } a_t = a$$

- i. if $x \in X_{a_t}$ then

$$p(x|a_t) = p(x|a_t) + \frac{1}{N} r(i)$$

- ii. else

$$p(x|a_t) = \frac{1}{N} r(i)$$

$$X_{a_t} = X_{a_t} \cup \{x\}$$

- (b) for each $x \in X_{a_t}$

$$h = 0$$

- i. for each $i \in [1, N]$

- A. for each $j \in [1, J]$

$$g = \frac{1}{N} r(i) P(j|a) \phi(x; B_j \hat{\mathbf{z}}_{t|t-1}^{(i)}, B_j \Sigma_{t|t-1}^{(i)} B_j^T + \nu)$$

$$h = h + g \log(g/p(x|a))$$

$$I[L_t | X_t, a_t] = h$$

3. select the a_t that maximises $I[L_t | X_t, a_t]$.
-

where

$$\boldsymbol{\mu}_t(x_t) = \mathbb{E}_{p(\mathbf{z}_t|x_{0:t}, a_{0:t})}[\mathbf{z}_t], \quad (4.28)$$

$$= \left(I_{D_z} - K_{t|j}^{(i)} B^{(j)} \right) \hat{\mathbf{z}}_{t|t-1}^{(i)} + B^{(j)} x_t. \quad (4.29)$$

Hernandez (2004), Martinez-Cantin *et al.* (2007) define a CRLB that utilises stochastic prediction of the predicted latent state space. As I am applying this to a chaotic dynamical system I only use one step forward along the Markov chain. Using Equation 4.27, we can construct the CRLB as the difference between the two quantities,

$$\mathbb{E}_{p(\mathbf{z}_t|x_{0:t-1}, a_{0:t})}[\mathbf{z}_t^2] = \frac{1}{N} \sum_{i=1}^N r^{(i)} \left(\hat{\mathbf{z}}_{t|t-1}^{(i)T} \hat{\mathbf{z}}_{t|t-1}^{(i)} + \text{tr} \left[\Sigma_{t|t-1}^{(i)} \right] \right) \quad (4.30)$$

and

$$\begin{aligned} \mathbb{E}_{p(x_t|x_{0:t-1}, a_{0:t})}[\boldsymbol{\mu}_t(x_t)^2] &= \frac{1}{N^3} \left\{ \sum_{i,i',j,j'} P(j|a_t) P(j'|a_t) \frac{P(k_t^{*(i)}|k_{t-1}^{(i)}) P(k_t^{*(i')}|k_{t-1}^{(i')})}{\hat{P}(k_t^{*(i)}|k_{t-1}^{(i)}) \hat{P}(k_t^{*(i')}|k_{t-1}^{(i')})} \times \right. \\ &\quad \sum_{i''} \left(\left(\left(I_{D_z} - K_{t|j}^{(i)} B^{(j)} \right) \hat{\mathbf{z}}_{t|t-1}^{(i)} + B^{(j)} \hat{x}_{t|t-1}^{*(i)} \right)^T \right. \\ &\quad \left. \left(\left(I_{D_z} - K_{t|j'}^{(i')} B^{(j')} \right) \hat{\mathbf{z}}_{t|t-1}^{(i')} + B^{(j')} \hat{x}_{t|t-1}^{*(i')} \right) \right. \\ &\quad \left. \left. + \text{tr} \left[B^{(j)} H_{t|t-1}^{*(i'')} B^{(j')T} \right] \right) \right\}, \quad (4.31) \end{aligned}$$

where $i, i' \in [1, N]$, $j, j' \in [1, J]$ and the Kalman gain for the i 'th particle is $K_{t|j}^{(i)} = \Sigma_{t|t-1}^{(i)} B^{(j)T} \left(B^{(j)} \Sigma_{t|t-1}^{(i)} B^{(j)T} + \nu \right)^{-1}$. These equations can be used to construct Algorithm 4.3, which scales as $O(N_a J (N + J))$.

4.5.3 Application of my approach from Chapter 3

The information measure of action is used to select an action such that the filtered posterior's variance is expected to be minimised. The measure takes the upper bound over the conditional Shannon information from Theorem 3.1 in the previous Chapter

$$I[Z_t, \mathbf{J}_t, \mathbf{K}_t | X_t, a_t, \dots] = I[Z_t, \mathbf{K}_t | \dots] + I[\mathbf{J}_t | a_t] + I_{\text{MI}}[X_t; Z_t, \mathbf{J}_t, \mathbf{K}_t | a_t, \dots], \quad (4.32)$$

$$\leq I[Z_t, \mathbf{K}_t | \dots] + I[\mathbf{J}_t | a_t] + B_{X_t|a_t} [Z_t, \mathbf{J}_t, \mathbf{K}_t | Z'_t, \mathbf{J}'_t, \mathbf{K}'_t], \quad (4.33)$$

Algorithm 4.3 An implementation of the *a posteriori* CRLB (Hernandez, 2004, Martinez-Cantin *et al.*, 2007) adapted to a partially observable switching SSM. This scales with a time complexity of $O(N_a J(N + J))$, where there are N Rao-Blackwellized particles representing the latent belief $p(l_t = \{k_t, j_t, z_t\} | x_{0:t-1}, a_{0:t})$, J choices of measure and N_a possible choices of action a_t . The shorthand $r(i) = \frac{P(k_t^{*(i)} | k_{t-1}^{(i)})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)})}$ is used in this Algorithm. The $\frac{P}{\hat{P}}$ quantity is used in Algorithm 4.1 to construct the prediction particles for $p(l_t | x_{0:t-1}, a_{0:t})$.

Calculation of the *a posteriori* CRLB,

$$\mathbf{g}_0 = 0_{D_z}, G_1 = 0_{D_z \times D_z}$$

1. for each $i \in [1, N]$

$$r(i) = \frac{P(k_t^{*(i)} | k_{t-1}^{(i)})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)})}$$

$$\hat{\mathbf{z}}_{t|t-1}^{(i)} = A_{k^{(i)}} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + b_{k^{(i)}}$$

$$\Sigma_{t|t-1}^{(i)} = A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{(i)T} + Q$$

$$\mathbf{g}_0 = \mathbf{g}_0 + r(i) \hat{\mathbf{z}}_{t|t-1}^{(i)}$$

$$G_1 = G_1 + r(i) \left(\hat{\mathbf{z}}_{t|t-1}^{(i)} \left(\hat{\mathbf{z}}_{t|t-1}^{(i)} \right)^T + \Sigma_{t|t-1}^{(i)} \right)$$

2. for each a_t

$$G_3 = 0_{D_z \times D_z}, \mathbf{g}_4 = 0_{D_z}, \mathbf{g}_5 = 0_{1 \times D_z}$$

- (a) for each $i \in [1, N]$

$$\hat{\mathbf{z}}_{t|t-1} = \hat{\mathbf{z}}_{t|t-1}^{(i)}$$

$$\Sigma_{t|t-1} = \Sigma_{t|t-1}^{(i)}$$

- i. for each $j \in [1, J]$

$$K = \Sigma_{t|t-1} B_j^T \left(\nu + B_j \Sigma_{t|t-1} B_j^T \right)^{-1}$$

$$G_3 = G_3 + P(j|a) \left(\nu + B_j G_1 B_j^T \right)$$

$$\mathbf{g}_4 = \mathbf{g}_4 + P(j|a) r(i) (I_{D_z} - K B_j) \hat{\mathbf{z}}_{t|t-1}$$

- (b) for each $j \in [1, J]$

$$\mathbf{g}_5 = \mathbf{g}_5 + P(j|a) B_j$$

$$\text{CRLB}(a_t) = \frac{1}{N} \text{tr}[G_1] - \frac{1}{N^3} \left(\frac{1}{\nu^2} \text{tr}[\mathbf{g}_5^T G_3 \mathbf{g}_5] + \mathbf{g}_4^T \mathbf{g}_4 + \frac{2}{\nu} h \mathbf{g}_4^T \mathbf{g}_5^T \mathbf{g}_5 g_0 \right)$$

3. select the a_t that maximises $\text{CRLB}(a_t)$.
-

where $L_t = Z_t \cup K_t \cup J_t$. As the contribution from $I[Z_t, K_t | \dots]$ is independent of a_t it is neglected. The contribution of J_t for Equation 4.32 is the Shannon information of J_t ,

$$I[J_t | a_t] = \sum_j P(j | a_t) \log P(j | a_t). \quad (4.34)$$

The bound portion which represents an easily computable contribution of X_t to Equation 4.32 is expressed fully as,

$$B_{X_t | a_t} [Z_t, J_t, K_t \| Z'_t, J'_t, K'_t] = \sum_{j_t, j'_t, k_t, k'_t} \int d\mathbf{z}_t d\mathbf{z}'_t p(\mathbf{z}_t, j_t, k_t | x_{0:t-1}, a_{0:t}) p(\mathbf{z}'_t, j'_t, k'_t | x_{0:t-1}, a_{0:t}) \times D_{KL} [p(X_t | \mathbf{z}_t, j_t) \| p(X_t | \mathbf{z}'_t, j'_t)]. \quad (4.35)$$

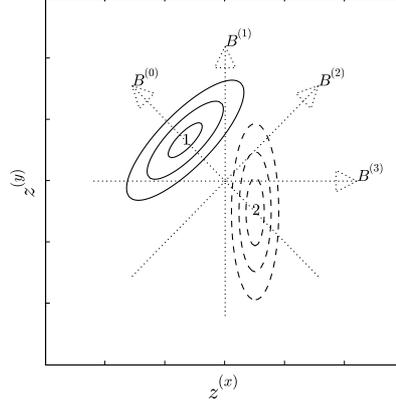
From the definition of the KL-divergence in Appendix D for one dimensional Gaussian measurement beliefs from Equation U.2,

$$p(x_t | \mathbf{z}_t, j_t) = \phi(x_t; B^{(j_t)} \mathbf{z}_t, \nu), \quad (4.36)$$

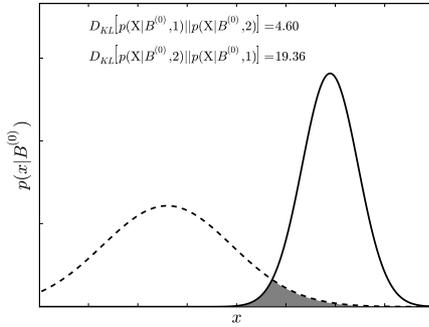
which gives

$$D_{KL} [p(X_t | \mathbf{z}_t, j_t) \| p(X_t | \mathbf{z}'_t, j'_t)] = \frac{1}{2\nu} \left(B^{(j_t)} \mathbf{z}_t - B^{(j'_t)} \mathbf{z}'_t \right)^T \left(B^{(j_t)} \mathbf{z}_t - B^{(j'_t)} \mathbf{z}'_t \right), \quad (4.37)$$

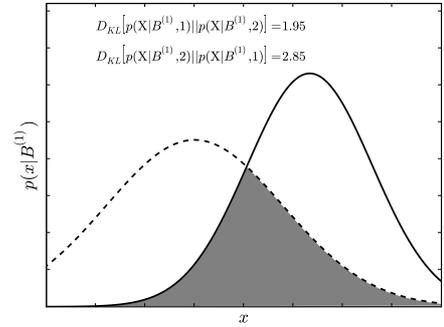
the $B^{(j)}$ mappings are interpreted as unit vector filters. An interpretation of this divergence is depicted in Figure 4.6, for a prior state belief $p(\mathbf{z} | \dots)$ comprised of a mixture of two Gaussians on the \mathbf{z} -plane, and the ability to choose a known mapping to the measurable x -line from \mathbf{z} is interpreted as measuring along this unit vector. As the correspondence of a to j is fully known in this contrived example it means that Figure 4.6 indicates that the bound $B_{X_t | a_t} [Z_t, S_t \| Z'_t, S'_t]$ shall select the mapping $B^{(j)}$ that has the greatest separation for the mappings onto the x -line for the two Gaussians. Figure 4.6 is for the special case where the correspondence between j and a is known and for simplicities sake is chosen to be $B^{(a)} = B(j)$ for $a = j$. The insight this contrived example gives to the more general problem is that the action that corresponds to the unit vector filter which gives the greatest separation upon X between the competing hypotheses as represented by the particles will be favoured.



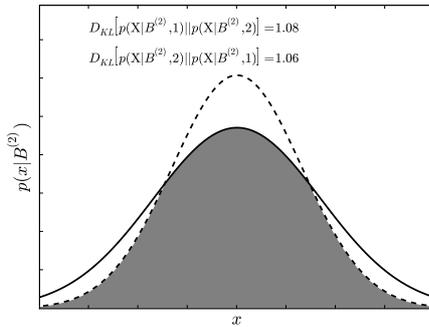
(a) The pair of distributions on the z plane, $p(z|1)$ and $p(z|2)$ are indicated respectively by solid and dashed lines. The $B^{(i)}$ mappings are indicated by a dotted arrow.



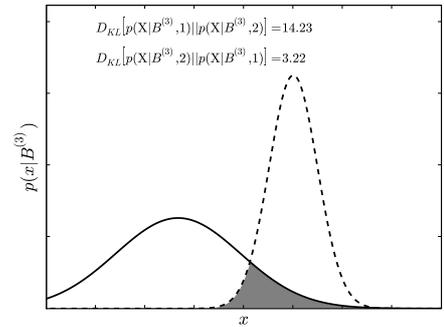
(b) The mapping $B^{(0)}$.



(c) The mapping $B^{(1)}$.



(d) The mapping $B^{(2)}$.



(e) The mapping $B^{(3)}$.

Figure 4.6: A visual interpretation of the measurement process for a contrived example. The figure indicates an interpretation of the measurement process $B^{(a)}$ as defined in Subsection 4.4.2 for a contrived example. Using a mixture of two equally likely Gaussians ($p(\mathbf{z}|i) = \phi(\mathbf{z}; \boldsymbol{\mu}_i, \Sigma_i)$) representing a pair of hypotheses of the z -plane as depicted in plot (a). Further in plot (a) a number of unit transpose vectors $B^{(j)}$ (e.g. $B^{(j)} B^{(j)T} = 1 \forall j$) representing a choice of measurement through the process $p(x|j, \mathbf{z}) = \delta(x - B^{(j)}\mathbf{z})$. In keeping with the notation in the text the shape of $B^{(j)}$ is the same as the transpose of \mathbf{z} . Plots (b) through (e) depict the overlap of $p(x|i, j) = \int dz p(x|j, \mathbf{z}) p(\mathbf{z}|i)$ and both KL-divergences for the hypotheses are indicated. This shows that the bound, *cf* Equations 4.35 & 4.37 which is in this case the average of the two KL-divergences, and will favour the mapping to the x -line of $B^{(0)}$ and $B^{(3)}$.

Algorithm 4.4 An implementation of my information measure (Equation 3.3) adapted to a partially observable switching SSM. This is an application of Chapter 3. This scales with a time complexity of $O(N + N_a J)$, where there are N Rao-Blackwellized particles representing the latent belief $p(l_t = \{\mathbf{z}_t, k_t, j_t\} | x_{0:t-1}, a_{0:t})$, J choices of measure and N_a possible choices of action a_t . The shorthand $r(i) = \frac{P(k_t^{*(i)} | k_{t-1}^{(i)})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)})}$ is used in the Algorithm.

Calculation of my method from Chapter 3,

$$g_1 = 0$$

$$\mathbf{g}_2 = 0_{D_z}, G_3 = 0_{D_z \times D_z}$$

1. for each $i \in [1, N]$

$$r(i) = \frac{P(k_t^{*(i)} | k_{t-1}^{(i)})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)})}$$

$$\hat{\mathbf{z}}_{t|t-1} = A_{k^{(i)}} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + b_{k^{(i)}}$$

$$\Sigma_{t|t-1}^{(i)} = A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{(i)T} + Q$$

$$g_1 = g_1 + r(i)$$

$$\mathbf{g}_2 = \mathbf{g}_2 + r(i) \hat{\mathbf{z}}_{t|t-1}$$

$$G_3 = G_3 + r(i) \left(\Sigma_{t|t-1} + \hat{\mathbf{z}}_{t|t-1} \hat{\mathbf{z}}_{t|t-1}^T \right)$$

2. for each a_t

$$\mathbf{g}_4 = 0_{1 \times D_z}, G_5 = 0_{D_z \times D_z}, g_6 = 0$$

- (a) for each $j \in [1, J]$

$$\mathbf{g}_4 = \mathbf{g}_4 + P(j|a_t) B_j$$

$$G_5 = G_5 + P(j|a_t) B_j^T B_j$$

$$g_6 = g_6 + P(j|a_t) \log P(j|a_t)$$

$$I[\mathbf{J}_t | a_t] + B_{X_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t] = \frac{1}{\nu N^2} (g_1 \text{tr}[G_5 G_3] - \mathbf{g}_2^T \mathbf{g}_4^T \mathbf{g}_4 \mathbf{g}_2) + g_6$$

3. select the a_t that maximises $I[\mathbf{J}_t | a_t] + B_{X_t | a_t} [\mathbf{L}_t \| \mathbf{L}'_t]$.
-

Algorithm 4.5 The protocol for learning episodically using a variety of informatic policies.

1. Generate a new trajectory of actions and measurements with a policy selecting actions according to one of Algorithms 4.2, 4.3 & 4.4. The latent state space is constructed using Algorithms 4.1 & T.1,
 2. optimally update the parameters based upon this data and the previous episode's parameters using the smoothing updates from Algorithm T.2 and the EM-updates defined in Appendix V until the likelihood converges or a set number of EM iterations is reached,
 3. return to 1.
-

Accordingly Equation 4.35 becomes,

$$\begin{aligned}
B_{X_t|a_t} [Z_t, J_t, K_t | Z'_t, J'_t, K'_t] &= \frac{1}{2\nu} \int dz_t dz'_t \sum_{j_t, k_t, j'_t, k'_t} p(\mathbf{z}_t, j_t, k_t | x_{0:t-1}, a_{0:t}) p(\mathbf{z}'_t, j'_t, k'_t | x_{0:t-1}, a_{0:t}) \times \\
&\quad \left(B^{(j_t)} \mathbf{z}_t - B^{(j'_t)} \mathbf{z}'_t \right)^T \left(B^{(j_t)} \mathbf{z}_t - B^{(j'_t)} \mathbf{z}'_t \right), \tag{4.38} \\
&= \frac{1}{2\nu N^2} \sum_{j, j'} P(j|a_t) P(j'|a_t) \sum_{i, i'} \frac{P(k_t^{*(i)} | k_{t-1}^{(i)}) P(k_t^{*(i')} | k_{t-1}^{(i')})}{\hat{P}(k_t^{*(i)} | k_{t-1}^{(i)}) \hat{P}(k_t^{*(i')} | k_{t-1}^{(i')})} \times \\
&\quad \left\{ \left(B^{(j)} \hat{\mathbf{z}}_{t|t-1}^{*(i)} - B^{(j')} \hat{\mathbf{z}}_{t|t-1}^{*(i')} \right)^T \left(B^{(j)} \hat{\mathbf{z}}_{t|t-1}^{*(i)} - B^{(j')} \hat{\mathbf{z}}_{t|t-1}^{*(i')} \right) + \right. \\
&\quad \left. \text{tr} \left[B^{(j)} \Sigma_{t|t-1}^{*(i)} B^{(j)T} \right] + \text{tr} \left[B^{(j')} \Sigma_{t|t-1}^{*(i')} B^{(j')T} \right] \right\}. \tag{4.39}
\end{aligned}$$

This measures the overlap between each and every pair of trajectories (including a trajectory to itself) after mapping to the measurable space X_t according to an action a_t via $B^{(j_t)}$ and the belief in the correspondence between j_t and a_t using $P(j_t|a_t)$. A simplified example of which can be seen in Figure 4.6. Algorithmically this is constructed as Algorithm 4.4, which scales as $O(N + N_a J)$.

4.6 Results

In Subsection 4.1.1, I pose the hypothesis that a good localising choice of action will allow a better model to be learnt from the data set rather than random and undirected actions. This is to test the hypothesis that an active agent can learn a more optimal model of its environment than an agent that explores in a random manner. The random policy is uniform and so represents undirected exploration through a choice of action.

I accomplish this by using a linear switching state space model (SSM) as discussed in Section 4.4. The action defines which measure (unit vector filter) of the state space is observed by an action. However, the correspondence of an action to the measures will be

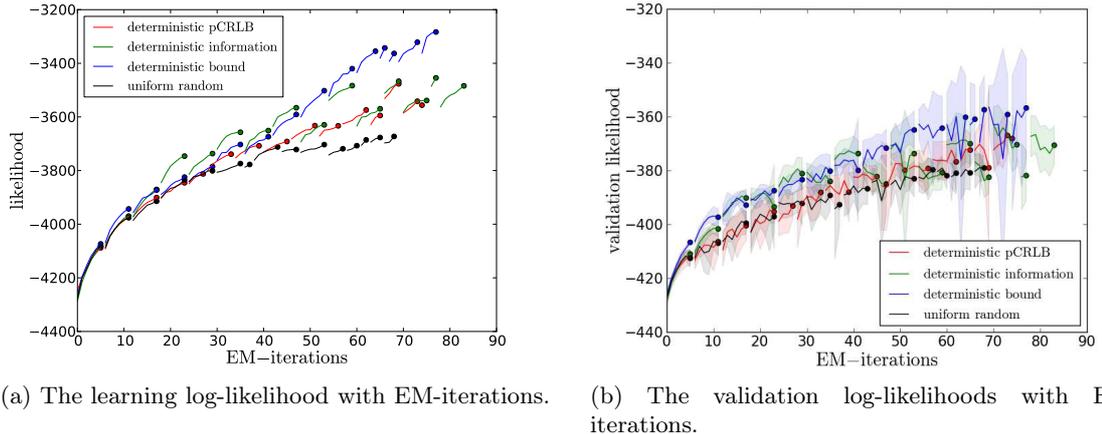
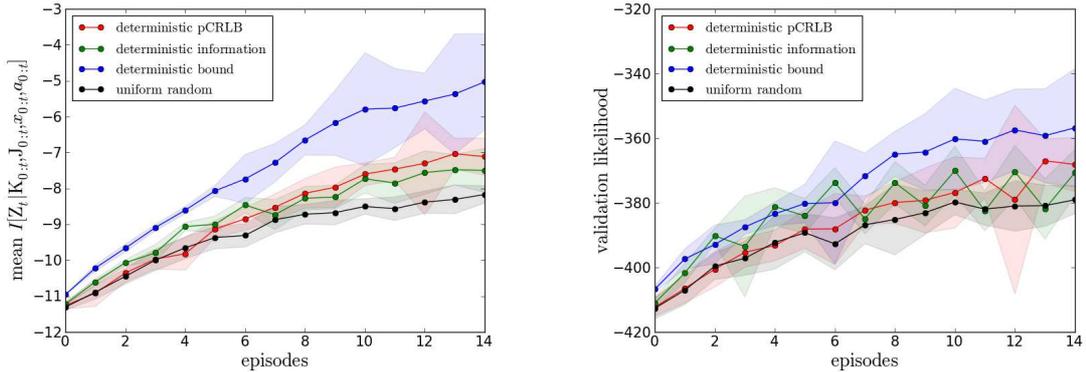


Figure 4.7: The likelihoods of the learning data sets and validation data sets for each of the informatic policies defined in Algorithms 4.2, 4.3 & 4.4 contrasted with a uniform random policy. The dots indicated in plot (a) represent the likelihood for an episode after the EM-algorithm has ceased to increase the likelihood. It is important to notice the upward trends in the likelihood of the selected parameterisation for each episode (with a dot indicating its termination). The distribution of the validation data set’s log-likelihoods, for each learnt model, can be seen in plot (b). Each line ending with a dot indicates the mean of the validation likelihoods over the validation data sets. The opaque blocks indicate the range of the standard deviation of the validation likelihoods for the collection of validation data sets. The deterministic upper bound can be seen to significantly outperform each of the other policies.

unknown and have to be learnt. The switching state space model used $K = 25$ states. Each of these states represent a distinct linear process of the form $z_t = A_k z_{t-1} + b_k$. 25 states were observed to be an adequate number to represent the dynamics of the attractor with a fully observed latent space (*cf* HMM). The particle filter between time steps (a time t to a time $t + 1$) stores $N = 40$ particles. A super sample of 80 particles is used to process each prediction of the latent state space. This was found to be an adequate tradeoff between increasing the number of particles N and the time complexity of the learning algorithm. In my development of this model I found that $N = 40$ particles was sufficient for tracking the latent state space.

The action policies derived in Section 4.5 for this switching SSM use varying measures of knowledge. The learning phase for each episode uses a fresh sample of $T = 500$ time steps from the Lorenz attractor – this is referred to as the learning data set. The graphs used for analysis have a set of 15 samples each of $T = 50$ time steps from the Lorenz attractor to validate the models constructed in the learning phase – this is referred to as the validation data set.

I shall show that an action selection policy with directed localising actions (Section 4.5)



(a) Learnt models’ “certainties” in their inferences for the learnt parameters for each episode.

(b) Learnt models’ “accuracies” in their inferences for the learnt parameters for each episode.

Figure 4.8: Figure indicating the episodic evolution, of each policy, for the learnt models’ inference “certainties” and “accuracies”. Plot (a) indicates the validation “certainties” as quantified by the conditional Shannon information $I[Z_t | K_{0:t}, J_{0:t}, x_{0:t}, a_{0:t}] = \int d\mathbf{z}_t \sum_{k_{0:t}, j_{0:t}} p(\mathbf{z}_t, k_{0:t}, j_{0:t} | x_{0:t}, a_{0:t}) \log p(\mathbf{z}_t | k_{0:t}, j_{0:t}, x_{0:t})$. Plot (b) indicates the validation “accuracies” as quantified by the validation likelihood, see Figure 4.7. The opaque blocks indicate the range of the standard deviation of the quantities for validation data sets. The deterministic upper bound can be seen to significantly outperform each of the other policies.

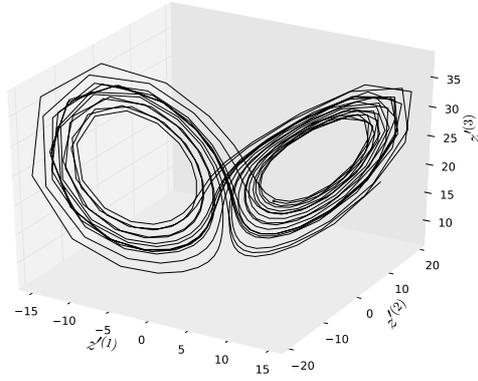
is better than unbiased random actions. In this section I analyse different action policies for an agent which has a choice of measure as an action. I then consider the case of a change of correspondence due to a change in the measurement processes $B^{(a)}$. This allows me to test my hypothesis that directed action can construct a better model than undirected action. Further, that directed action coupled with episodic learning allows for successful re-learning of a model of the system Ω .

4.6.1 Comparison between informatic policies

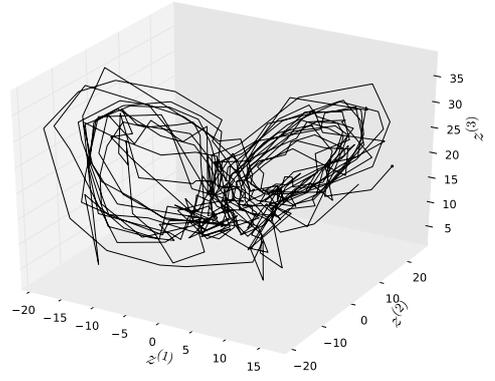
In this subsection I compare and contrast the performances of each policy described in Section 4.5. Further, these informatic policies are compared to a uniform random policy. The purpose of the random policy is to act as a base comparison against which the other policies must perform better.⁶ I approach this by constructing models of the Lorenz attractor’s dynamics according to Algorithm 4.5 for 15 episodes. Each episode consists of up to 5 EM-updates. The number of iterations of the EM-algorithm for each episode is low, because the aim is to compare the impact of each policy upon learning, as having the learning algorithm converge in the first episode would be counter productive.

As can be seen in Figure 4.7 both the CRLB and Infomax policies, respectively Al-

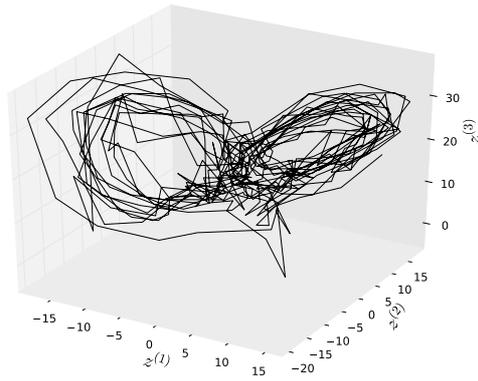
⁶A uniform random policy acts as a minimum standard against which other policies must be better. If a policy cannot outperform a random policy which has little computation cost then it has no utility.



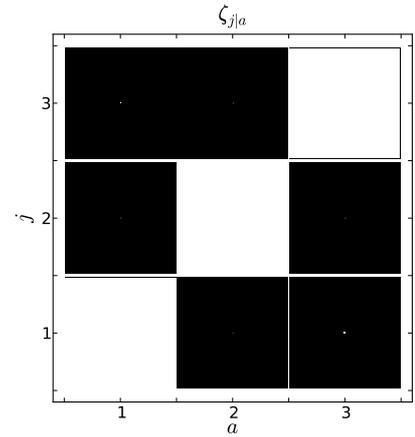
(a) A sample \mathbf{z}' trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.

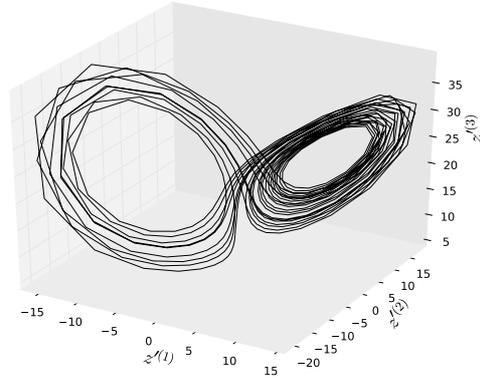


(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.

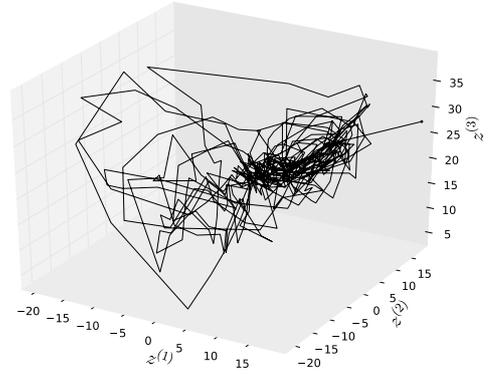


(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$) for an action a to a choice of measure j . Where the probabilities are represented by the area of the white square.

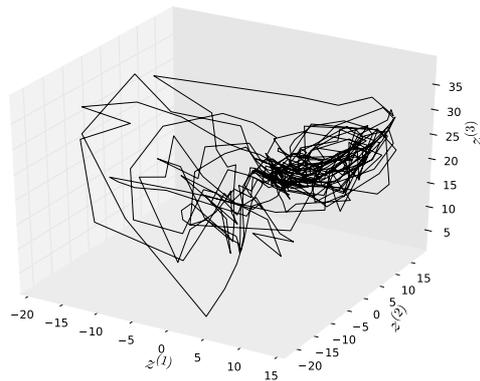
Figure 4.9: The data trajectory, inferred trajectories and the learnt measurement process for the *a posteriori* Cramer-Rao lower bound (CRLB). The filtered trajectory in plot (b) and the smoothed trajectory in plot (c) for the data set indicated in plot (a). The $B^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.



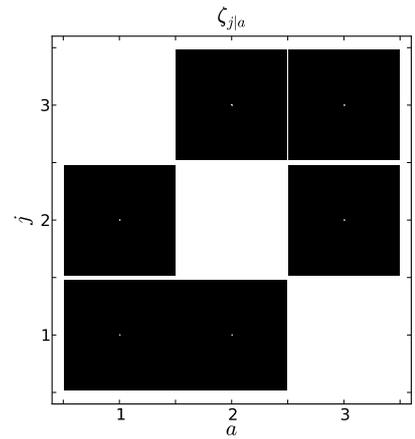
(a) A sample \mathbf{z}^l trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.

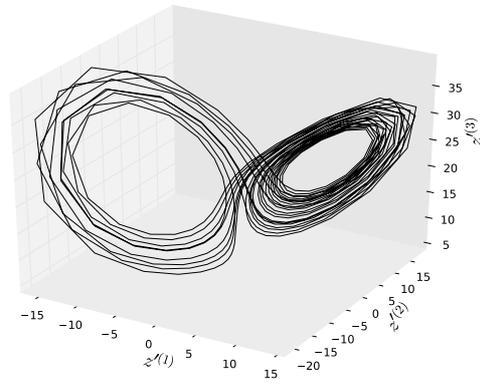


(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.

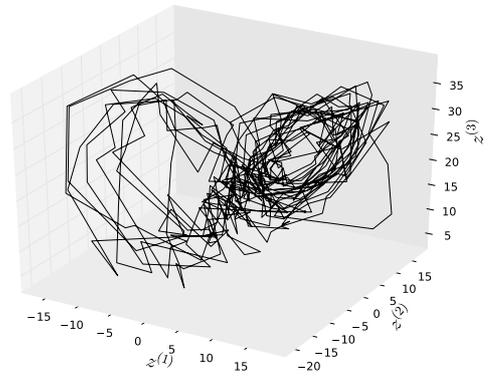


(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$) for an action a to a choice of measure j . Where the probabilities are represented by the area of the white square.

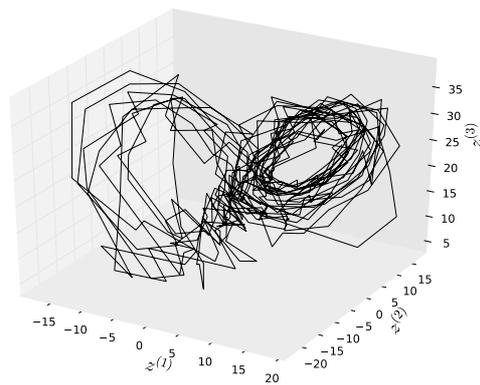
Figure 4.10: The data trajectory, inferred trajectories and the learnt measurement process for the adapted Infomax approach of Porta *et al.* (2003, 2005). The filtered trajectory in plot (b) and the smoothed trajectory in plot (c) for the data set indicated in plot (a). The $B^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.



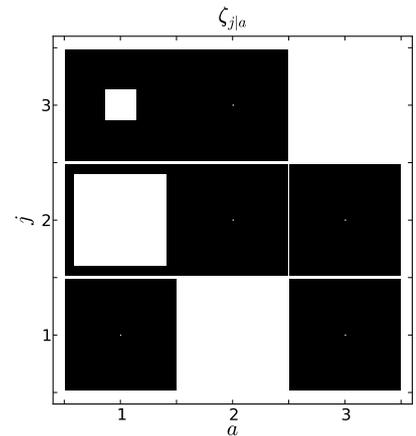
(a) A sample \mathbf{z}' trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.

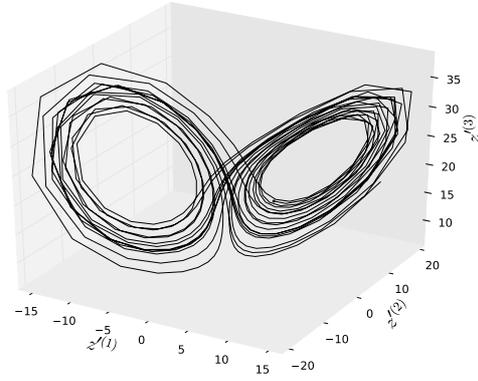


(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.

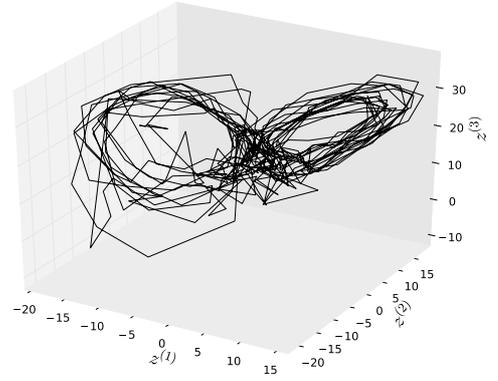


(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$) for an action a to a choice of measure j . Where the probabilities are represented by the area of the white square.

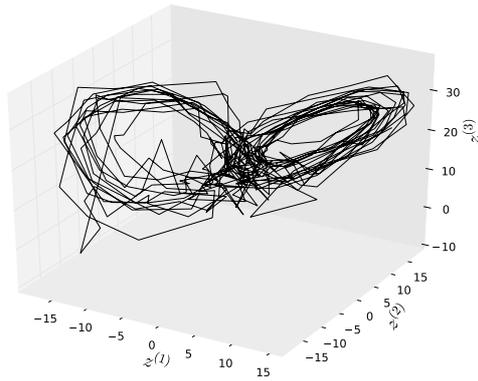
Figure 4.11: The data trajectory, inferred trajectories and the learnt measurement process for a policy based upon my approach from Chapter 3. The filtered trajectory in plot (b) and the smoothed trajectory in plot (c) for the data set indicated in plot (a). The $B^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.



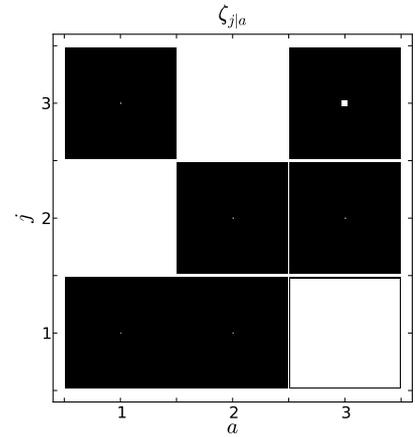
(a) A sample \mathbf{z}' trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.



(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.



(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$) for an action a to a choice of measure j . Where the probabilities are represented by the area of the white square.

Figure 4.12: The data trajectory, inferred trajectories and the learnt measurement process for a uniform random policy. The filtered trajectory in plot (b) and the smoothed trajectory in plot (c) for the data set indicated in plot (a). The $B'^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.

gorithms 4.2 & 4.3, are out performed by the upper bound (Algorithm 4.4). The learning data sets' log-likelihood for each informatic policy shows the Infomax policy to initially be higher for the EM-iterations 20 to 50. This though does not translate to better performance in later EM-iterations. Nor does this translate into a higher validation likelihood, Figure 4.7 (b). The CRLB and Infomax algorithms finish performing slightly better than a uniform random policy.

Figure 4.7 shows a far smoother incline for my approach, labelled as the deterministic bound, when contrasted to the Infomax policy. The cause of the zig-zagging in the mean validation likelihood, in Figure 4.8 (b), for Infomax can be attributed to the use of a sample S of the latent state space L , action A and measurements X .⁷ If a combination has not been explored, then it cannot be considered by the Infomax policy (Algorithm 4.2).

This shows a separation between my approach and the other informatic policies – my approach is cheaper to compute than both the CRLB and Infomax policies, with favourable Algorithmic scaling, Section 4.5. Also from Figure 4.8 (a), my policy (the upper bound) has a consistently higher mean Shannon information. The mean Shannon information equates to a higher certainty in inferences upon Z . There is also a trend for use of my policy to cause the learning algorithm to outperform the use of the other policies in accurately modelling the validation data sets. This is shown through the average validation likelihood in Figure 4.8 (b).

To verify, for each policy, that the episodic EM-algorithm can adequately represent the Lorenz attractor I present Figures 4.9-4.12. Each figure depicts in plot (a) a sample trajectory, in plots (b) & (c) the inferred trajectories and in plot (d) the belief table $P(j|a)$. The switched SSMs expected filtered Z state space is

$$\hat{\mathbf{z}}_{t|t} = \mathbb{E}_{p(\mathbf{z}_t|x_{0:t}, a_{0:t})} [\mathbf{z}_t], \quad (4.40)$$

$$= \int d\mathbf{z}_t \mathbf{z}_t \sum_{j_{0:t}, k_{0:t}} p(\mathbf{z}_t, j_{0:t}, k_{0:t} | x_{0:t}, a_{0:t}), \quad (4.41)$$

$$= \frac{1}{N} \sum_{i=1}^N \int d\mathbf{z}_t \mathbf{z}_t p(\mathbf{z}_t | j_{0:t}^{(i)}, k_{0:t}^{(i)}, x_{0:t}, a_{0:t}), \quad (4.42)$$

$$= \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}, \quad (4.43)$$

which are depicted in plot (b) for each figure. Similarly, for the smoothed case depicted in

⁷The set $S = A \cup L \cup X$ and is constructed from the previously learnt episodes trajectory for this joint space.

plot (c) of each figure, the expected smoothed Z state space is

$$\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}. \quad (4.44)$$

In plot (b) & (c) for Figures 4.9, 4.11 & 4.12, which correspond respectively to the CRLB, my approach and the random policies show that visually the trajectories $\hat{\mathbf{z}}_{t|t}$ and $\hat{\mathbf{z}}_{t|\tau}$ have a similar structure to the Lorenz attractor (each figure’s corresponding plot (a)). The exception is the Infomax policy’s trajectories in Figure 4.10. Figure 4.10 (b) & (c) shows that the episodic EM-algorithm using this policy, is unable to adequately track one of the Lorenz attractor’s loops. As already mentioned, the cause of the zig-zagging in the mean validation likelihood using the Infomax policy, can be attributed to the use of an inadequate sample S of the latent state space, action and measurements. Therefore, if a combination of a , \mathbf{z} and x have not been explored, and hence stored, then Algorithm 4.2 cannot consider such a combination. This also explains the contradiction in Figure 4.10 of a fully one-to-one association of a to j of the $P(j|a)$ table in plot (d), with the shape of the filtered and smoothed trajectories in plots (b) & (c).

Plot (d) of Figures 4.9-4.12 shows each policies learnt $P(j|a)$ table. These all exhibit a bias for a one-to-one association between j and a via the learnt $\zeta_{j|a}$ tables. However, my approach in Figure 4.11 (d) has a caveat, one of the directions is not queried uniquely for one particular action ($a = 1$). A probable explanation for this, not seriously affecting inference of $\hat{\mathbf{z}}_{t|t}$ and $\hat{\mathbf{z}}_{t|T}$, is that the dynamics cause the filtered posterior to be well mixed. This makes an inference on one dimension convey information about other dimensions. A further explanation is that there are some regions of the attractor where one dimension of \mathbf{z} has the same value as another dimension. This is what can allow the EM-algorithm to maintain any uncertainty in $P(j|a)$ for an action a . Considering the topologically accurate inferences when comparing Figures 4.11 (b) & (c) to Figure 4.11 (a), the general topology of the Lorenz attractor seems to have been learnt – this topology being the double loop in Z. The same also is true for Figures 4.9 & 4.12, indicating the system (the Lorenz attractor) has been adequately represented.

To summarise the analysis of this subsection, I have made a comparison between three informatic policies:

1. my approach from Chapter 3,
2. an Infomax approach which is an adaptation of Porta *et al.* (2003, 2005),

Algorithm 4.6 The protocol for learning episodically using a stochastic policy.

1. Generate a new trajectory of actions and measurements with a policy selecting actions according to $P(a_t|\dots) = \frac{e^{n \cdot C(a_t)}}{\sum_{a_t} e^{n \cdot C(a_t)}}$, where $C(a_t) = I[\mathbf{J}_t|a_t] + B_{\mathbf{X}_t|a_t}[\mathbf{Z}_t, \mathbf{J}_t, \mathbf{K}_t || \mathbf{Z}'_t, \mathbf{J}'_t, \mathbf{K}'_t]$, iteratively over time with Algorithms 4.1 & T.1,
 2. optimally update the parameters based upon this data and the previous episode's parameters using the smoothing updates from Algorithm T.2 and the EM-updates defined in Appendix V until the likelihood converges or a set number of EM iterations is reached,
 3. return to 1.
-

3. and the application of an *a posteriori* CRLB estimator.

In terms of accuracy and certainty, these are found to outperform a uniform random policy. However, it is clear that my approach retains some ambiguity in the measurement process Figure 4.11 (d). The Infomax policy though, has the filtered and smoothed trajectories (Figure 4.10 (b) & (c)) that represent worst the sample trajectory Figure 4.10 (a). In contrast, the policies other than Infomax have each represented the double loop of the Lorenz attractor.

4.6.2 A stochastic modification

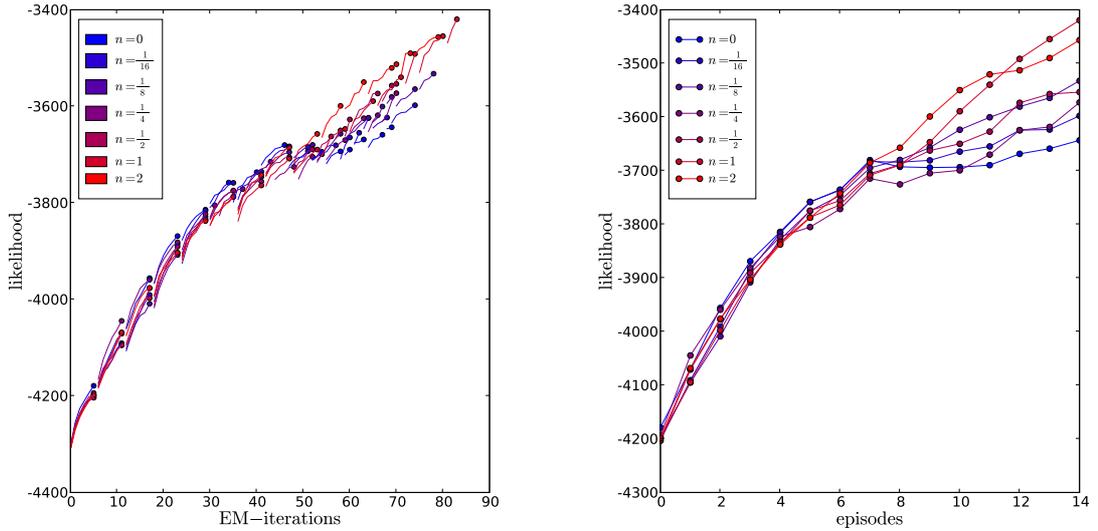
How do we consider a continuum of policies between an informatic policy, such as Algorithm 4.4, and the uniform random policy? This can be approached by treating the measure of information as an energy like quantity and varying the temperature of the distribution. Using the shorthand

$$C(a_t) \triangleq I[\mathbf{J}_t|a_t] + B_{\mathbf{X}_t|a_t}[\mathbf{Z}_t, \mathbf{S}_t || \mathbf{Z}'_t, \mathbf{S}'_t], \quad (4.45)$$

which is the components of Equation 4.32 that depend upon a_t . This allows for a transition between actions being informatively exploitative of knowledge and purely exploratory. The probability of choosing an action a_t is defined to be,

$$P(a_t|\dots) = \frac{e^{n \cdot C(a_t)}}{\sum_{a_t} e^{n \cdot C(a_t)}}, \quad (4.46)$$

where n is used to make the policy more or less greedy in terms of knowledge by respectively decreasing or increasing n – n being the temperature of the belief $P(a_t|\dots)$. In the next two subsections, this is used to learn episodically according to the policy defined in Subsection 4.5.3 which is described in Algorithm 4.6.



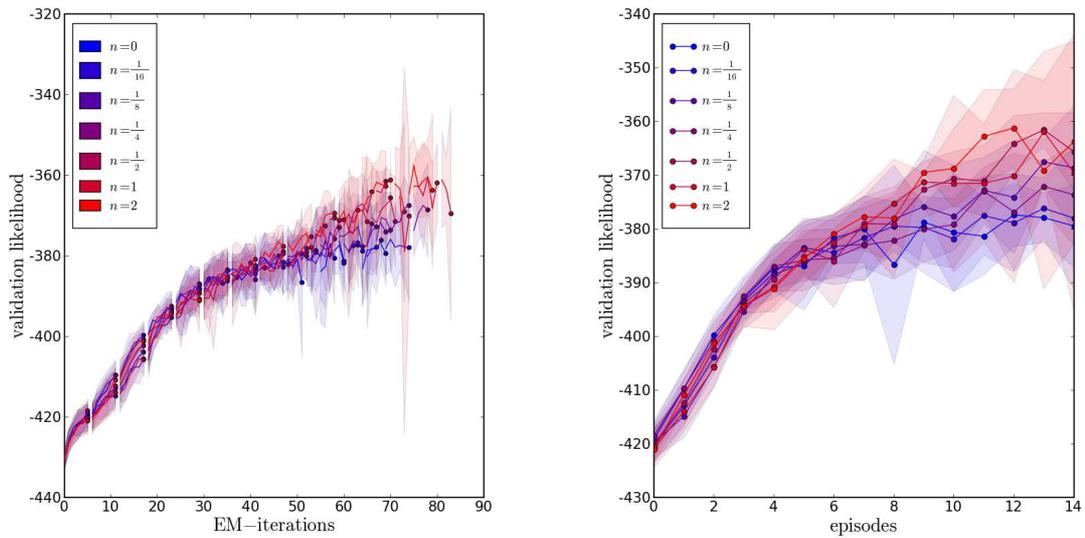
(a) The learning log-likelihood with EM-iterations (b) The learning log-likelihood with episode

Figure 4.13: The learning data set’s likelihood across different weightings of uniform to greedy stochastic action selection. The different weightings n for the quantity $C(a_t)$ to make a stochastic selection of a_t , Equation 4.46. $n = 0$ represents a completely random policy and $n = 2$ represents the greediest localising policy considered. In both plots (a) & (b) the filled dots indicate the likelihood for an episode after the EM-algorithm has reached either of its stopping conditions; these conditions are either five EM-iterations, or if the likelihood drops. It is important to notice the upward trends in the likelihood of the selected parameterisation for each episode (dots).

4.6.3 Directed localising action is better than undirected actions

In this subsection I analyse different action policies for an observer that can make a choice of measure, to consider the best choice of policy, be it entirely exploratory ($n = 0$) or mostly greedy ($n = 2$). As can be seen in Figure 4.13 (a) the log-likelihood in every episode always rises, though there can be a drop from the end of one episode to the start of the next episode due to each episode using a new data sample. Figure 4.13 (b) indicates that after episode 7 the learning data likelihoods for some n ’s drop. The validation likelihood does not always match this, however like Figure 4.13 (a) the validation likelihoods in Figure 4.14 (b) begin to diverge after episode 7. This indicates there is an initial learning phase where all n ’s are essentially the same, followed by a reinforcement phase where inference certainty appears to be more important for the model’s accuracy as measured by the validation likelihood.

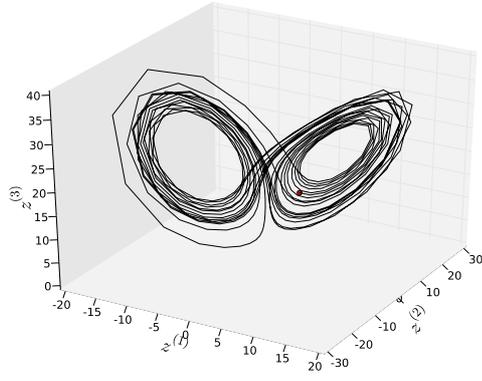
I use Figure 4.15 to show that the model and learning algorithms have made acceptable inferences for the system’s state space. For the filtered case depicted in Figure 4.15 (b) the actual state space is depicted in Figure 4.15 (a), the switched SSMs expected filtered



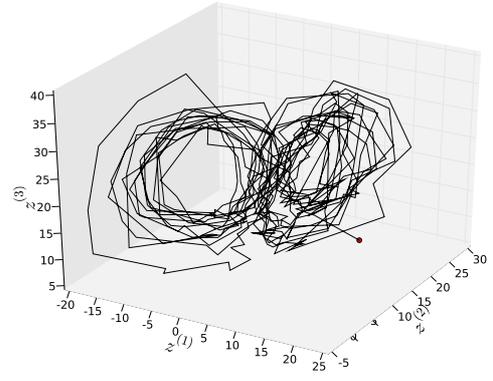
(a) The validation log-likelihoods with EM-iterations.

(b) The validation log-likelihoods with episode.

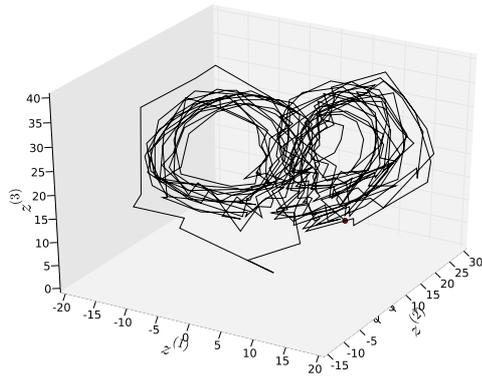
Figure 4.14: The range of validation data sets' likelihood across different weightings of uniform to greedy stochastic action selection. The different weightings n for the quantity $C(a_t)$ to make a stochastic selection of a_t , Equation 4.46. $n = 0$ represents a completely random policy and $n = 2$ represents the greediest localising policy considered. In plot (a) each line ending with a filled dot indicates the mean of the validation likelihoods over the validation data sets. In plot (b) the distribution of validation likelihoods at the end of each learning episode is indicated by filled dots. The coloured blocks indicate the range of the standard deviation of the n 's validation likelihoods with EM iteration for plot (a) and the end episode for plot (b).



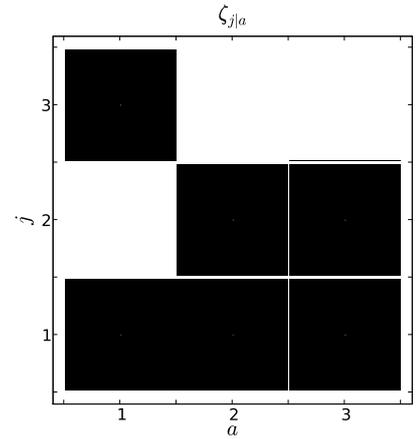
(a) A sample \mathbf{z}' trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.

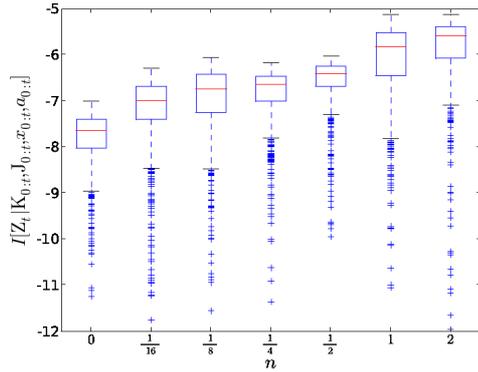


(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.

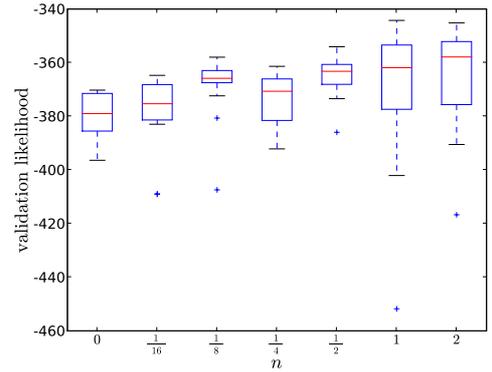


(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$ for an action a to a choice of measure j). Where the probabilities are represented by the area of the white square.

Figure 4.15: The data trajectory, inferred trajectories and the learnt measurement process. The filtered trajectory in plot (b) and smoothed trajectory in plot (c) for the data set indicated in plot (a) and the learnt model for $n = 2$. The $B^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.

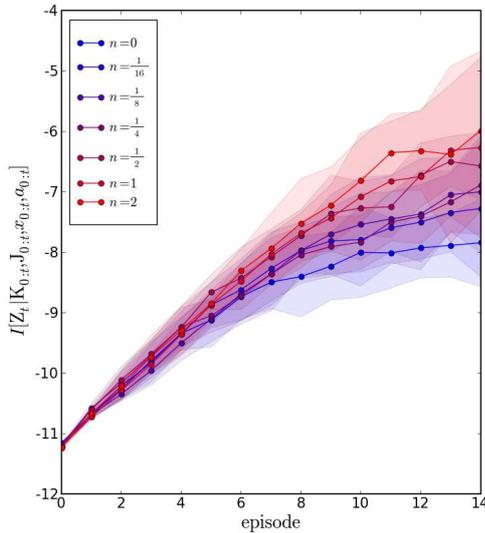


(a) Learnt models' "certainties" in their inferences for the learnt model of the final episode.

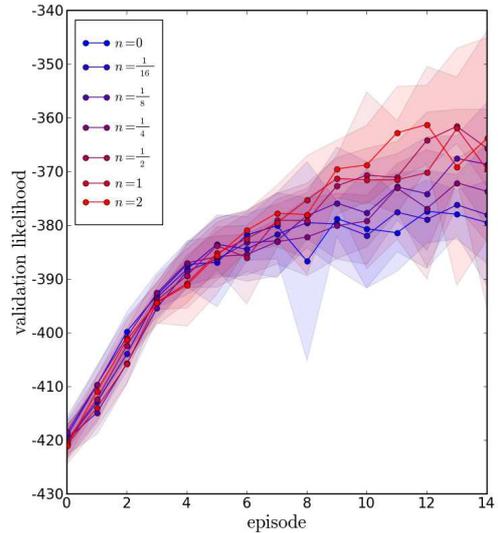


(b) Learnt models' "accuracies" in their inferences for the learnt model of the final episode.

Figure 4.16: Figure of the final learnt models' inference certainties and accuracies across the different weightings of uniform to greedy stochastic action selection. Box and whisker plots of the learnt model's validation inference certainties in plot (a) as measured by the conditional Shannon information $I[Z_t | J_{0:t}, K_{0:t}, x_{0:t}, a_{0:t}] = \int dz_t \sum_{s_{0:t}} p(\mathbf{z}_t, j_{0:t}, k_{0:t} | x_{0:t}, a_{0:t}) \log p(\mathbf{z}_t | j_{0:t}, k_{0:t}, x_{0:t})$. Box and whisker plots of the learnt model's accuracy in plot (b) as measured by the final iterations validation likelihoods for each n . The conditional Shannon information is related to the model's certainty and not its accuracy, it is not a measure of how well a model represents a system, but how certain the model is in its inferences. The distribution of validation likelihoods in plot (a) do indicate an unbiased estimate of how accurately each n 's learnt model represents the system from which these validation data sets were drawn.



(a) Learnt models' "certainties" in their inferences for the learnt parameters for each episode.



(b) Learnt models' "accuracies" in their inferences for the learnt parameters for each episode.

Figure 4.17: Expansion of Figure 4.16 to indicate the episodic evolution of the certainties and accuracies. The episodic evolution of certainties is depicted in plot (a) and accuracies are depicted in plot (b). The filled dots on both plots indicate the end of a learning episode. There is a trend for greater n to have both a higher "certainty" and "accuracy" as well as a greater variance (the final episode is shown in more detail in Figure 4.16).

Z state space is,

$$\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}, \quad (4.47)$$

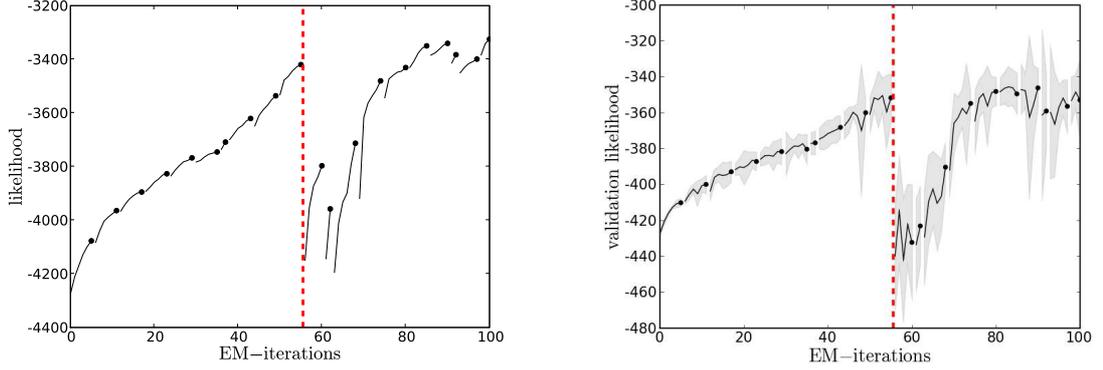
and similarly for the smoothed case depicted in Figure 4.15 (c), the expected smoothed Z state space is,

$$\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}. \quad (4.48)$$

Figure 4.15 (b) & (c) show that the trajectories $\hat{\mathbf{z}}_{t|t}$ and $\hat{\mathbf{z}}_{t|\tau}$ visually have a similar structure to the Lorenz attractor Figure 4.15 (a). The difficulty is that some of the directions are not queried properly, as the correspondence table in Figure 4.15 (d) assumes that two directions are identical. This is mostly due to the dynamics causing the filtered posterior to be well mixed, making an inference on one dimension convey information about other dimensions. However, this is contingent upon miss-corresponding inferences not being too misleading. Considering the topologically accurate inferences when comparing Figures 4.15 (b) & (c) to Figure 4.15 (a) this appears to be the case. This indicates the model has adequately learnt the system in spite of miss-corresponding some actions to measures in the correspondence table, Figure 4.15 (d). While Figure 4.15 shows the model has learnt to represent the system, it says little as to which n has constructed the best representation.

Next I consider which n is most accurate and certain. I take the conditional Shannon information $I[\mathbf{Z}_t | \mathbf{J}_{0:t}, \mathbf{K}_{0:t}, x_{0:t}, a_{0:t}]$ as my measure of an individual model's certainty in its inferred belief for \mathbf{z}_t . I use the validation likelihood as a measure of an individual model's accuracy in its representation of the system. The certainty and accuracy of the final learnt models for each n from Figure 4.16 shows a trend to more certainty and accuracy for larger n 's. The long tails in both plots in Figure 4.16 are accounted for by initialising the prior $p(\mathbf{z}_0)$ to be a Gaussian and covering the entire attractor on the space \mathbf{Z} , and $P(k_0)$ to be a uniform prior. These priors are chosen for validation data sets as the priors learnt for the learning data sets will be meaningless. I expand the comparison from Figure 4.16 to look back at the learnt models across the episodes in Figure 4.17, this shows that it takes 6 to 8 episodes for the accuracies and certainties to meaningfully diverge. This is likely to be an artifact of both the episodic learning algorithm (Section 4.2) and the policies being mostly exploratory. The divergence is likely due to the $n \neq 0$ policies becoming more greedy or exploitative of knowledge as $C(a_t)$ differs more between different choices of a_t , Equations 4.45 & 4.46.

To summarise the analysis of Figures 4.13-4.17, we see that the more uniformly random



(a) The learning log-likelihood with EM-iterations. (b) The validation log-likelihoods with EM-iterations.

Figure 4.18: The learning and validation likelihoods for relearning a model with a change to the system. This is for the weighting $n = \frac{1}{4}$ in $C(a_t)$ using the stochastic selection of a_t , Equation 4.46. The dots indicated in plot (a) represent the likelihood for an episode after the EM-algorithm has ceased to increase the likelihood. It is important to notice the upward trends in the likelihood of the selected parameterisation for each episode (with a dot indicating its termination). The range of validation data sets log-likelihoods for the corresponding iteration’s model can be seen in plot (b). Each line ending with a dot indicates the mean of the validation likelihoods over the validation data sets. The grey blocks indicate the range of the standard deviation of the validation likelihoods for the collection of validation sets. The dashed red line in both plots (a) & (b) indicates the shuffle of $B^{(a)}$ (learning other variables intact) with the consequence that the correspondence table $\zeta_{j|a}$ is no longer valid. The substantial difference in the learning versus re-learning phases can be accounted for by the initial policy being primarily exploratory (learning) versus exploratory re-evaluation of hypotheses (re-learning).

a policy is, the more inferior it tends to be to stochastic directed action selection for learning to localise – learning to track upon an unknown latent state space. This is due to the “greedy” policy of $n = 2$ being exploratory relative to how “certain” the measure $C(a_t)$ is in Equations 4.45 & 4.46. This lends further credence to my argument that the policy for all the n ’s are equivalently exploratory of the system until a reinforcement phase where certainty is more important for the model’s accuracy.

4.6.4 “Adult” learning of correspondence

In this subsection I analyse the effect of an imposed change to the measurement process as modelled by $p(x|\mathbf{z}, j)$. This change is to the sampling process as denoted by $B^{(a)}$, and takes the form of a shuffle. This causes the correspondence of j to a , to change due to the shift in $B^{(a)}$ relative to $B^{(j)}$ making the learnt correspondence table $P(j|a) = \zeta_{j|a}$ obsolete. This is a simple analysis of the mechanism of an agent adapting its model of its environment to a contrived change in physiology – it relates to the work in Hofman *et al.* (1998), Rossetti *et al.* (1993, 1998).

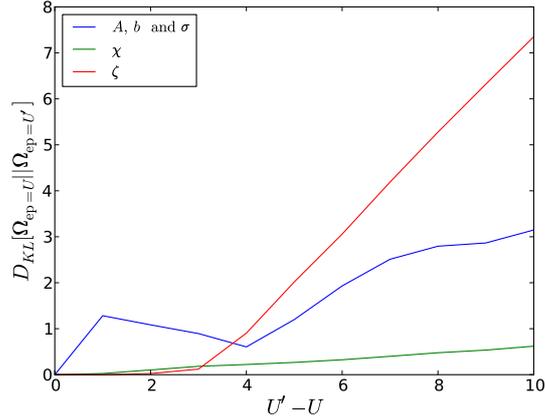
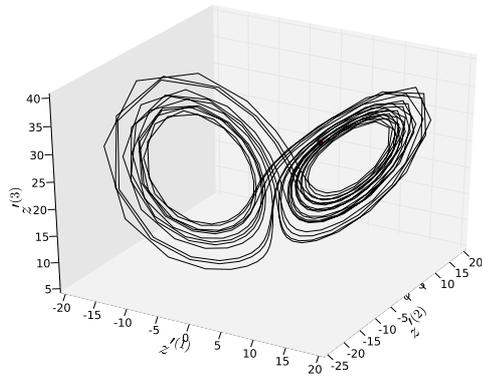


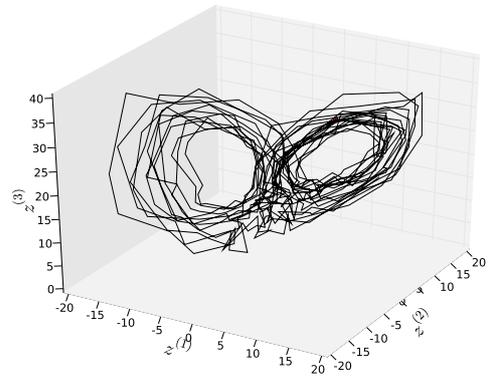
Figure 4.19: The KL-divergence of the learnt parameters with episodes after a change in the system’s measurement process. The KL-divergence of the learnt parameters after the change in $B^{(a)}$ to the learnt parameters before the change in $B^{(a)}$, *ie* $U = 10$ and U' varies from U to $2U$. $U' - U$ are the episodes since the change in $B^{(a)}$. $U' - U = 0$ corresponds to the red dashed line in Figure 4.18.

After the change of $B^{(a)}$ the model’s data and validation likelihoods fall dramatically (Figure 4.18). This is to be expected, as the system has been arbitrarily changed. This makes the agent’s model of the system incorrect and necessitates relearning of the affordances/contingencies. Initially the learning algorithm appears to optimise specifically for the learning set and not the validation sets in Figure 4.20. Though after this initial lack of learning generality, the model’s accuracy, as measured by the validation likelihood, returns to its previous high a few episodes after the change in $B^{(a)}$. This indicates that an equally good representation of the system has been relearnt. These changes can be associated in Figure 4.19 with an initial change to the learnt dynamics (parameterised by A , b and σ) not affecting the accuracy. This is followed by a change in correspondence after 3 episodes seeming to cause the most significant shift in the accuracy (validation likelihood in Figure 4.18 (b)).

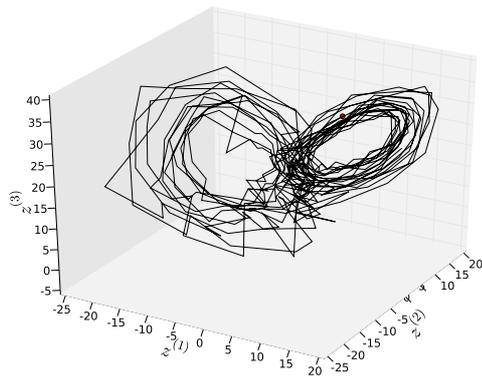
Using the filtered and smoothed expectations of z with respect to time t , and contrasting these with the sample from the Lorenz attractor, the end of the initial learning phase (upto the 10th episode) is depicted in Figure 4.20. The structure of the attractor has been learnt for both the initial $B^{(a)}$ (Figure 4.20) and after its transition to a new $B^{(a)}$ (Figure 4.21). The difficulty is that some of the directions are not queried properly, as the correspondence table in Figure 4.21 (d) assumes that two directions are identical. This is mostly due to the dynamics causing the filtered posterior being well mixed, making an inference on one dimension convey information about other dimensions. However, this is contingent upon miss-corresponding inferences not being too misleading. Considering the topologically



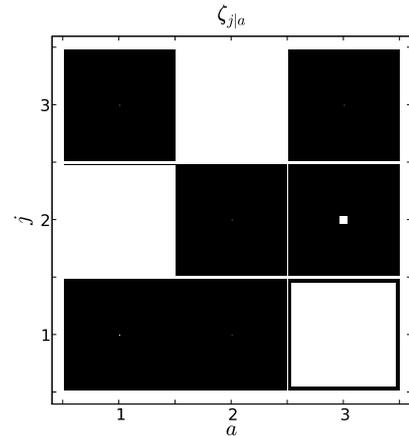
(a) A sample \mathbf{z}' trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.

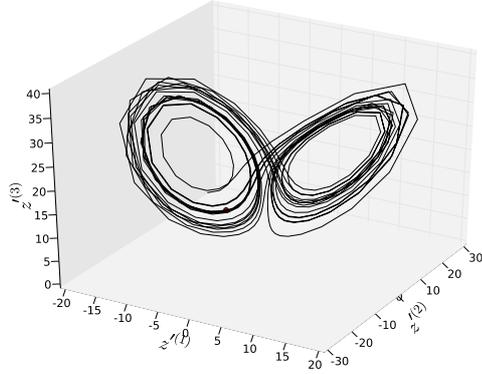


(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.

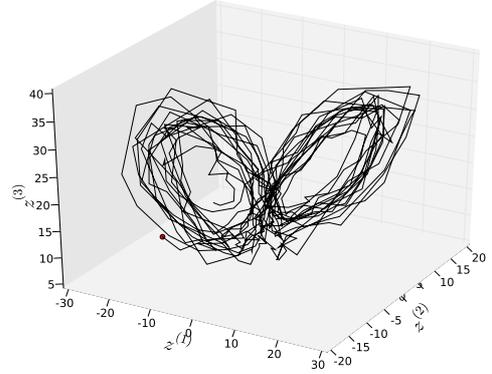


(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$) for an action a to a choice of measure j . Where the probabilities are represented by the area of the white square.

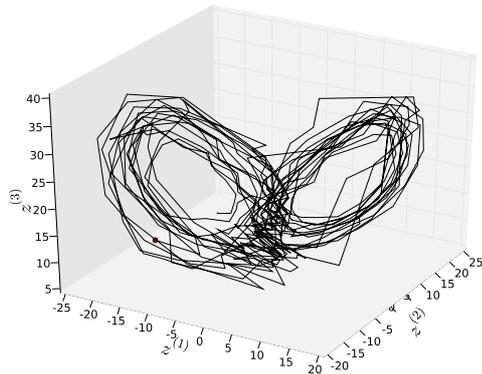
Figure 4.20: The data trajectory, inferred trajectories and the learnt measurement process for the initial system in Figure 4.18. The filtered trajectory in plot (b) and the smoothed trajectory in plot (c) for the data set indicated in plot (a). The $B^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.



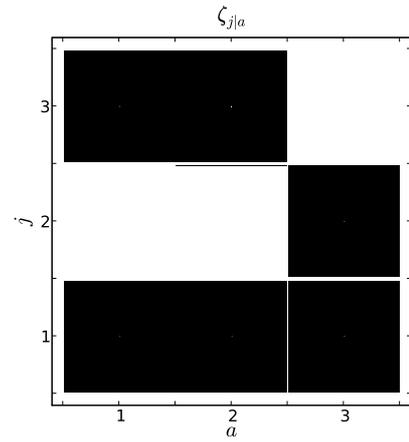
(a) A sample \mathbf{z}' trajectory.



(b) The mean of the filtered trajectories, $\hat{\mathbf{z}}_{t|t}$.



(c) The mean of the smoothed trajectories, $\hat{\mathbf{z}}_{t|T}$.



(d) The learnt correspondence table $\zeta_{j|a}$ (which represents the belief $P(j|a)$) for an action a to a choice of measure j . Where the probabilities are represented by the area of the white square.

Figure 4.21: The data trajectory, inferred trajectories and the learnt measurement process after the change in the system in Figure 4.18. The re-learnt model's trajectories for filtered expectation $\hat{\mathbf{z}}_{t|t}$ in plot (b) and smoothed expectation $\hat{\mathbf{z}}_{t|T}$ in plot (c) trajectories for the data set indicated in plot (a) for a learnt model of $n = \frac{1}{4}$. The $B^{(a)}$ for the measurement process of the data in plot (a) is distinct from the model's assumed $B^{(j)}$, as can be seen in the optimised correspondence belief table $\zeta_{j|a}$ plot (d) which represents the belief $P(j|a)$. The mean of the filtered trajectories represents $\hat{\mathbf{z}}_{t|t} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|t}^{(i)}$ and for the smoothed trajectories $\hat{\mathbf{z}}_{t|T} = \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{z}}_{t|T}^{(i)}$.

accurate inferences when comparing Figures 4.21 (b) & (c) to Figure 4.21 (a) this appears to be the case.

To summarise the analysis of Figures 4.18 to 4.21 in this subsection, I have shown that an episodic learning algorithm can adapt a model to a reconfiguring of the measurement process represented by $P(j|a)$ parameterised using $\zeta_{j|a}$. This is similar to the findings of Hofman *et al.* (1998), Rossetti *et al.* (1993, 1998) where a previously stable sensory configuration is arbitrarily modified (in the cases via the individual's physiology) and the individuals adapt to this. Finally comparing Figures 4.20 & 4.21 I have shown that an imposed change upon the measurement process results in a reconfiguration of the model's representation of the measurement process.

4.7 Discussion

4.7.1 Findings

To validate my hypothesis that a directed active agent can learn a better model, which is more certain and accurate in its inferences than an undirected active agent, I have used a RBPF to construct a Bayesian representation of a partially observable system. In the development of this Chapter I have found the following requirements for active perception:

1. A representable latent space.
2. A method of refining a learnt parameter over a number of episodes.
3. The observables are dependent upon a choice referred to in this thesis as an action.

This defines the requirements that a system or problem must have for learning active-perception. My instantiation of these respective conditions are:

1. To represent a partially observable non-linear dynamical process using a linear switching SSM, defined in Section 4.4.
2. An episode has a data set which is learnt optimally using an episodic EM-algorithm which I defined as a method to allow the EM-algorithm to bootstrap from a previous episode's learning (see Section 4.2).
3. The data set of an episode is generated with reference to an informatic policy, examples are defined in Section 4.5.

Putting these together I have constructed a useful model of observation systems.

In approaching the optimisation part of the problem and its consequent requirements for a solution, I have constructed an episodic learning algorithm. This episodic learning algorithm takes the form of a Lagrangian upon the EM updates and can be interpreted as a joint belief update according to a prior episode's learnt parameter. This episodic learning is especially important in problems with rare dynamic events which are unlikely to be seen in any one episode. Prior rare dynamic events will not be ignored by the episodic updates, but if too rare, these may be forgotten. This is especially important for policies that make use of previous experience, for instance my adaptation of Infomax from Porta *et al.* (2003, 2005).

In Subsection 4.6.1, I contrast three informatic policies:

1. my approach from Chapter 3,
2. an Infomax approach which adapted Porta *et al.* (2003, 2005)
3. and the application of an *a posteriori* CRLB estimator.

These policies were defined mathematically and algorithmically in Section 4.5. There is a clear difference between the results of these informatic policies. Firstly, my approach is cheaper to compute with favourable Algorithmic scaling, Section 4.5. Secondly, my policy has a consistently higher mean Shannon information, which I equate to a higher inference certainty. Thirdly, there is a trend for my policy to be the most accurate, which I equate to the validation likelihood. In terms of accuracy and certainty, all three informatic policies are found to outperform a uniform random policy. However, my approach retains some ambiguity in the measurement process Figure 4.11 (d). The Infomax policy though, has the filtered and smoothed trajectories (Figure 4.10 (b) & (c)) that represent worst the sample trajectory Figure 4.10 (a). In contrast, the policies other than Infomax have each represented the double loop of the Lorenz attractor.

In Subsection 4.6.3, I have shown that a choice of action that implicitly localises and explores when necessary, will allow a better model to be learnt from the data set than completely random and undirected actions. Hence, an active agent can learn a more optimal model of the world than an inactive agent (which can be seen in Chapter 2), or one that explores in an undirected manner (a policy defined in Section 4.5 by Equation 4.46 setting $n = 0$). More strongly it can be argued from Held and Hein (1963) that active and directed perception is a critical developmental requirement of learning to represent the

world. In this context a representation of an agent’s environment is better than another one due to having more certain inferences, which equates to a higher accuracy. This is an obvious but important definition of “better” for making inference based decisions. In addition, having greater inference certainty, this can result in a learnt model of the system that is better at representing the systems dynamics. This is especially important for a sensorimotor system where the former constraint, being more certain in inferences, is utilised to define what the best actions are. Both constraints are used in this Chapter to firstly infer the most certain inferences by selecting actions to increase certainty and then refining the model’s parameters to fit these more certain inferences to the data set.

What I show is a slightly stronger statement, that informatically directed action selection is better for learning than undirected or random action selection. Though both of these are better than the passive case where prior knowledge is required (Chapter 2). I have accomplished this by using a linear switching state space model (SSM) discussed in Section 4.4 where the action defines which unit vector measure of the state space is observed. However, the correspondence of an action to the measures is unknown and has to be inferred. I derived an action selection policy in Section 4.5 using the information measure I developed in Chapter 3. I compared this with two other widely accepted measures of knowledge from the literature. I then analysed different action policies for an agent for which an action is a choice of measure. I have shown that a directed localising policy which can explore when necessary is generally better for inferences. The more directed a policy, the greater are its expected certainty and expected accuracy (Figure 4.16) than an entirely random action selection (Subsection 4.6.3). This allows me to state that directed action can construct a generally better model than undirected action.

In my analysis of different stochastic action policies I showed that a policy which is informatically greedy is typically best. In contrasting Figure 4.8 to Figures 4.13 & 4.14 we see that the stochastic policy forms a range of roughly ordered performance from uniformly random policy up to the deterministic informatic policy (Algorithm 4.4). Performance is defined according to accuracy and certainty. In Subsection 4.6.3, these stochastic policies take on the aspect of the random policy when the information measure does not distinguish between actions. However, the non-stochastic policy in Figures 4.7 & 4.8 show a consistently higher validation likelihood than the stochastic policies for all episodes.

The uniform stochastic policy did however remove the ambiguity in the measurement process, Figure 4.11 (d) to Figure 4.12 (d). Subsection 4.6.3 shows that there is an initial

learning phase where each policy performs equivalently, followed by a reinforcement phase where the importance of inference certainty is more exaggerated. The consequence is to make all policies appear to “body babble” initially (Olsson *et al.*, 2005a, 2006), as can be seen most clearly in Figure 4.22, where a higher entropy indicates a smoother distribution. Though, as the episodes pass, for the policies where $n \neq 0$, the entropies in Figure 4.22 drift away from the purely exploration or uniformly random policy.

I then showed that a change of the measurement process had the effect that my episodic learning algorithm reconfigured the model’s measurement process to match the new environment. This is similar to re-learning to localise due to a contrived change in physiology as observed by Hofman *et al.* (1998), Rossetti *et al.* (1993, 1998) where a previously stable sensory configuration is arbitrarily modified (in those cases via the individual’s physiology) and the individuals adapt to this. This showed concretely that an active agent can learn refinements of the world when the world changes. Hence, I have shown that for a previously learnt model, an episodic learning algorithm can adapt this model to an arbitrary reconfiguring of the measurement process. This is without special alterations being made for either inference or learning. This is similar to Hofman *et al.*’s and Rossetti *et al.*’s findings that adult humans could adapt their sensory configuration to continue to localise events. This could suggest that humans have a meta representation of previous sensory experience against which we learn episodically.

4.7.2 Limitations

In discussing my findings it is important to note that a learnt latent space Z will not always correspond, in a one to one fashion, with the true latent state space Z' . This is obvious with respect to Figures 4.15 & 4.21 in Section 4.6, which, due to the structure of the respective correspondence tables $\zeta_{j|a}$, makes it impossible for Z to map perfectly to Z' . This is because some directions will never be queried properly if the correspondence table $\zeta_{j|a}$ assumes that some a ’s are identical, Figures 4.15 & 4.21 (d). In spite of this, those same figures show that the attractor’s structure has been learnt.

A difficulty with my episodic learning algorithm is that a miss-correspondence by the likelihood $P(j|a)$ occurs. This assumes that some actions have identical consequences, e.g. that $P(j|a) \approx \delta_{ja}$ is not the case, this is illustrated in Figures 4.15 & 4.21. A potential solution to this is to construct a Lagrangian multiplier which sets the constraint that the

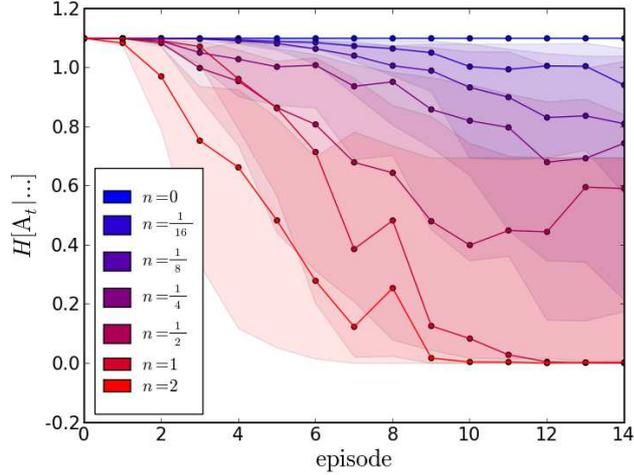


Figure 4.22: Plot of the smoothness of the action selection policies; the different weightings of uniform to greedy stochastic action selection. The initial validation action entropies for each episode as measured using $H[A_t|\dots] = -\sum_{a_t} P(a_t|\dots) \log P(a_t|\dots)$. The lower entropies indicate greater bias towards one action. According to definition $n = 0$ has no bias to any action. The figure indicates that all actions are equally likely over the first few episodes. An entropy of 0.0 indicates certainty in $P(a_t|\dots)$.

table $\zeta_{j|a}$, representing the likelihood $P(j|a)$, is orthonormal for a choice of action a .⁸

4.7.3 Behavioural implications

Though my findings do not show the completely passive case that Held and Hein (1963) consider, I have shown that random undirected action will only yield so much accuracy and inference certainty. The suggestion in Olsson *et al.* (2005a, 2006) is that an initial developmental phase of body babbling can be observed in even the greediest stochastic policy ($n = 2$) when it produces a uniform $P(a_t|\dots)$, Figure 4.22. This is due to the system model for initial episodes having very uncertain latent state space beliefs.

In contrast to Philipona *et al.* (2003, 2004), Aytekin *et al.* (2008) who used a correlation based approach to model learning I have used a pseudo Bayesian learning algorithm on a Bayesian model. One advantage of a correlation based approach is it will typically be cheaper to construct than a fully Bayes approach. Though this is sometimes at the cost of correctness as correlation does not indicate causation. This does not cause a difficulty

⁸For $\zeta_{j|a}$ to be orthogonal the following conditions must hold: firstly, for all a and a'

$$\sum_{j=1}^J \zeta_{j|a} \zeta_{j|a'} = 0. \quad (4.49)$$

Secondly, for all a

$$\sum_{j=1}^J \zeta_{j|a} \zeta_{j|a} = \varepsilon, \quad (4.50)$$

where $\varepsilon \neq 0$. The second condition can be strengthened to make $\zeta_{j|a}$ orthonormal by making $\varepsilon = 1$.

in arguing that an agent can relearn a sense when the relationship between Z' and Z has changed. When an individual relearns, there will be other senses grounding and hence restrain the reconfiguration of the dynamics.⁹

A future area of enquiry could develop a neural network for a connectionist based model of the world for an active agent. This would be especially useful where there are a large number of sensory channels, for instance:

1. in vision, for the rods and cones in the eye,
2. in audition, for the inner hair cells along the basilar membrane,
3. and in olfaction, for the genetic tuning to molecules.

Then all that would need to be learnt are the relationships between the responses of these filters and the action signals sent to the agent's actuators. This would make a generalisable active perceiver and would provide a model to investigate interesting questions of Nagel *et al.* (2005), Noe (2004), O'Regan and Noe (2001).

⁹From Beal *et al.* (2003), Hospedales and Vijayakumar (2006) we see an agent can simply learn the hyperplane of correlated cues, similarly to Philipona *et al.* (2003, 2004), where actions may be considered among these cues. However, this should include knowledge of the dynamic state of the system as in Hospedales *et al.* (2007) and this Chapter.

Chapter 5

Concluding remarks

5.1 Preamble

The overarching theme that ties this thesis together is that of localisation. Following the development of my research described in Section 1.5, I examined passive sound source localisation (Chapter 2) and then progressed to the selection of the best localising action in Chapters 3 & 4. For each Chapter, I review the context of my findings, I examine the limitation of my work and I suggest future directions for research into these problems.

5.2 Sound source localisation

In Chapter 2 “Bayesian passive sound source localisation” I discussed a passive localiser and illustrated the difficulty, that a passive model requires prior knowledge for viable localisation. In the context of my thesis the principle finding of Chapter 2 is essentially related to the symbol grounding problem (Harnad, 1990). In Chapter 2 I refer to this as the problem of correspondence, that a location corresponds to a particular measurement. I suggested that extra information is required rather than simply having a data set of readings and locations. While this does not show definitively that action is required to adequately learn a representation of location, it does indicate the limitations of passive localisation when placed in context with Wallach (1938, 1939, 1940), Jenison (1997), Klensch (1948), Aytekin *et al.* (2008).

In comparison to Nix and Hohmann (2006), who used a naive Bayes approach (*cf* Pena and Konishi, 2002), I defined a more complete model of the hidden variables dependencies. I compared the results of my model with the application of Nix and Hohmann’s. I found my approach gave both a higher certainty in inferences and greater accuracy (Tables 2.4 & 2.5).

My model's purpose was to capture the influence of the cue variables to the sound signals. In accordance with Park *et al.* (1996), a cursory inspection of my model from Chapter 2 revealed that signal latency had a marked influence upon sound level encoding. The spatial dependency of source location, from Section 2.4, upon the cues showed analytically that:

1. monaural loudness had the greatest influence on location inferences to the ipsilateral sensor,
2. absolute loudness had little if any influence upon location inferences and was generally swamped by the other cues (coinfluences),
3. time disparity however had the greatest influence upon non-medial location inferences. This was due to the windowing of time disparities, almost subordinating time disparity to the loudness terms,
4. level disparity had the greatest influence on the medial location inferences both ahead and to the rear of the apparatus.

Similar work has been done by Macpherson and Middlebrooks (2002) though it does not focus on the same issues as my research. I have principally explored sound source localisation, relating hidden variables in my model to sensory cues and evaluated the causal influence of these to localisation. Macpherson and Middlebrooks applied manipulations to the cues to evaluate the consequences for the inference of azimuth and elevation. Using the duplex theory of sound perception, they considered the localisation of broadband sounds. Similarly, the contribution of ITD and level cues in the case of head motion were considered by Wenzel (1995). Wenzel found that head movements helped considerably to resolve localisation confusions (inference ambiguities). This indicates the direction I took in my research, which was towards localisation as an active problem. A future development of my model in Chapter 2 could look at the frequency channels, allowing it to be related more fully to the duplex theory of sound perception.

The limitations of Chapter 2 stem from two sources, over-fitting and symbol grounding. Firstly, I argued that over-fitting is due to a number of factors:

1. Taking advantage of the knowledge that can be extracted from multiple frequency channels. This is similar to the problem of multisensory integration. It could be useful to apply an approach that integrates the different frequency channels using the concept of inference of structure Hospedales *et al.* (2007), Hospedales and Vijayakumar (2006). This was originally applied to the sensory integration of sound and

vision (Beal *et al.*, 2003, Hospedales *et al.*, 2007, Hospedales and Vijayakumar, 2006). However this could be problematic, as unlike the visual scene imposing a location (the latent state space), the signals in each channel would covary in a systematically similar fashion. This would still leave a problem of how to uniquely correspond one location to each measurement. Finally, it is important to note that the frequency decomposition of a sound signal, performed by the cochlea (Lyon, 1982, Slaney, 1988, Givelberg and Bunn, 2003), is preserved in almost all regions of the auditory pathway (Covey *et al.*, 1991, Suga and Tsuzuki, 1985, Wenstrup *et al.*, 1986). This could be used to mimic the results from Rayleigh’s Duplex theory of sound for different frequencies by the use of the multiple channels for inference.

2. Taking advantage of the attenuations for each location forming a distribution – making the attenuations variables rather than parameters. Further, the attenuations should form a joint distribution dependent on a sound source’s location (Figure 2.20).
3. I may have been using an unrealistic noise model. Hence, one might investigate using other distributions such as the Rayleigh distribution (Zurek, 1991) or the Pearson family of distributions. The Pearson family of distributions can represent parametrically the skew and kurtosis, which would allow the explicit representation of ILD and ITD processes (Nix and Hohmann, 2006) rather than constructing a model from the sound intensity processes.

Secondly, the problem of symbol grounding is due to fitting the data of known locations to sound samples rather than holding the locations as hidden variables. If target locations are hidden variables, we run up against the symbol grounding problem. By posing it as a sensorimotor problem (O’Regan and Noe, 2001), Aytakin *et al.* (2008) elegantly illustrated a method of solving this problem of symbol grounding. If we however consider the problem as lacking an active component, partly from the active motional theories of sound localisation (Subsection 1.2.3) and Noe (2004), then approaching the influence of action upon localisation becomes an important issue. This led me to consider in Chapter 3, the problem of active-perception and a novel method to solve the sensor placement problem under uncertainty.

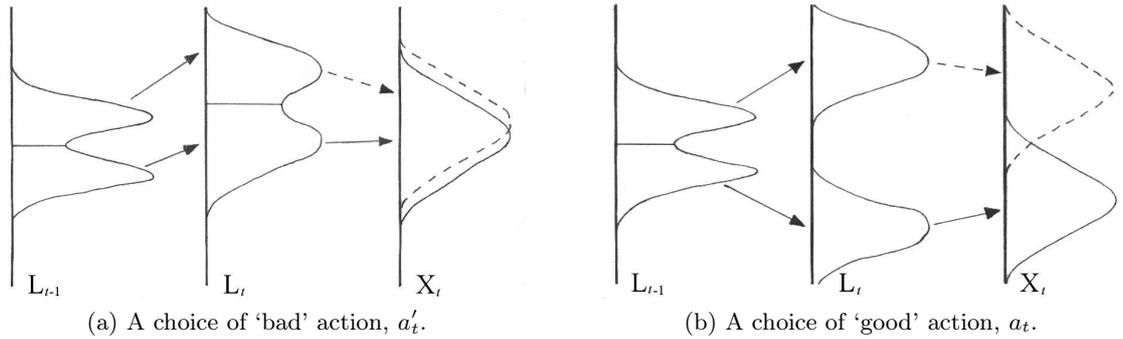


Figure 5.1: Cartoons of the consequence of “good” and “bad” localising actions, repeat of Figure 3.3. This depicts a “good” action a_t (right cartoon) which causes the two hypotheses upon L_{t-1} to be well separated in X_t . By contrast a “bad” action a'_t (left cartoon) does not lead to separation in X_t . It is important to note that this is only a cartoon and I have implicitly assumed that separation upon L_t leads to separation upon X_t , but this is not necessarily the case. It cannot be assumed that greater separation upon L_t leads to greater separation in X_t , as the distance between any set of hypotheses will be related to the projection of each according to $p(X_t|L_t)$. A measurement x_t for a choice of action a'_t plot (a) will find it hard to differentiate between either hypothesis upon L_t . However for a choice a_t plot (b) will find it much easier as the separation of hypotheses on L_t cause a similar separation on X_t .

5.3 Optimal action selection

In Chapter 3 “An information measure for optimal action selection” I developed and discussed a general information measure to the problem of selecting the optimal localising action in the sensor placement problem. This illustrated a solution to the problem of selecting an action to automatically give the next-best-view (NBV) (Gonzalez-Banos and Latombe, 2002) regardless of the measurement process or system dynamics.¹ Similarly Bongard *et al.* (2006) tracked hypotheses for robot physiology (body configuration), though my focus was upon Bayesian state space estimation for localisation. They selected an action to maximise the disagreement between candidate hypotheses (Bongard and Lipson, 2005). This had a similar effect to my measure (Chapter 3) which selected actions to maximally separate the latent hypotheses according to possible measurements, Theorem 3.10.

Principally, in using my measure for action selection, we find that the best and worst actions follow the cartoon in Figure 5.1. My measure is the average dissimilarity between possible posterior beliefs. This is not necessarily the least entropy but the greatest dependence between measurable and latent variable (respectively x_t and l_t conditional upon a_t). To repeat the summary, from Chapter 3, of my measure’s properties:

¹This concept of NBV is related to the use of stochastic forward models (Dearden and Demiris, 2005). These can be used to predict the sensory consequence of an agent’s actions, and are equivalent to the Bayesian prediction which I use extensively in Chapters 3 & 4. Though Gonzalez-Banos and Latombe (2002) construct the predicted latent state space through more deterministic methods.

- Maximising $B_{X_t|a_t} [L_t||L'_t]$ selects an action a_t to maximise the statistical dependence between x_t and l_t . This is also equivalent to maximising over all possible pairs of expected measurements the consequent posterior beliefs in l_t (Theorem 3.10). This indicates the measure will select an action that most distinguishes between all possible posterior predictions.
- Maximising $I_{\text{MI}} [X_t; L_t|a_t]$ selects the action a_t (maximising the mutual information) that will maximise the dependence between the variables l_t and x_t , cf $B_{X_t|a_t} [L_t||L'_t]$. By Theorem 3.3 this relates as

$$B_{X_t|a_t} [L_t||L'_t] = I_{\text{MI}} [X_t; L_t|a_t] + D_{\text{KL}} [p(X_t|a_t)p(L_t|a_t)||p(X_t, L_t|a_t)] \quad (5.1)$$

where the second term on the right-hand-side of the equation is also a measure of the statistical dependence between x_t and l_t for a choice of action a_t .

- Maximising $I [L_t|X_t, a_t]$ is the maximisation of the expected latent posterior's Shannon information

$$I [L_t|X_t, a_t] = \mathbb{E}_{p(x_t|a_t)} [I [L_t|x_t, a_t]], \quad (5.2)$$

by Theorem I.1. By Theorems I.2 & I.3, this quantity is related to the mutual information as

$$I [L_t|X_t, a_t] = I [L_t|a_t] + I_{\text{MI}} [X_t; L_t|a_t]. \quad (5.3)$$

Due to Theorem 3.1, it is also related to my measure through the inequality

$$I [L_t|X_t, a_t] \leq I [L_t|a_t] + B_{X_t|a_t} [L_t||L'_t]. \quad (5.4)$$

From these properties I argued that $B_{X_t|a_t} [L_t||L'_t]$ is a reasonable substitute for Shannon information, or the mutual information, as an information quantity.

If other quantities as well as $B_{X_t|a_t} [L_t||L'_t]$ are considered, my approach can also be constructed as an upper bound to the Shannon information. This allows it to include terms to penalise other features of the belief model, such as $I [L_t|a_t]$ which penalises higher entropy in the latent prediction $p(l_t|a_t)$. This point was illustrated experimentally in both comparisons in Chapter 3. The consequence of maximising the upper bound of the conditional Shannon information, is to select those actions that result in a measurement causing the greatest separation in hypotheses upon L_t according to the predicted *a posteriori*-

ori belief, Theorem 3.9. To summarise a bad action will generally result in a measurement being uninformative for inference, as the hypotheses are identical for any measurement (Figures 3.4 (d) & 3.5 (d)).

In Sections 3.6 & 3.7 I have used my measure to experimentally evaluate, through computer simulation, my approaches performance against Infomax in two localisation problems. In the first example, of head rotations in a free-field environment, I illustrated that:

- the target location can typically be treated as a linear process and the relationship between location to sound level, or time disparity, is non-linear and many-to-one.
- my approach slightly out performed Infomax in terms of certainty and accuracy.
- my approach and Infomax each had significantly different distributions of chosen actions. This made a general stochastic approximation untenable to represent the consequent policy of my approach. However, the reverse was the case for Infomax – the stochastic approximation had a similar performance.

In the second example, of detecting and capturing a source through chemotaxis in a turbulent medium (Vergassola *et al.*, 2007a), I illustrated that:

- using my approach I constructed an algorithm which I called Boundtaxis. Vergassola *et al.* had constructed an Infomax measure called Infotaxis.
- my Boundtaxis algorithm had a much greater variance in its distribution of search times than Infotaxis. Though the bulk of these search times (50% of runs) are significantly quicker than Infotaxis (Figure 3.24).
- algorithmically Boundtaxis scales more favourably than Vergassola *et al.*'s Infotaxis algorithm. Also my Boundtaxis algorithm was quicker to compute, by a factor of 20. Though this computation time is dependent upon the range of detections that Infotaxis is calculated for.
- my Boundtaxis algorithm will make a circuit of the latent space and then spiral in towards the source after a detection. Infotaxis by contrast will make a slowly increasing spiral from the agent's starting position (Vergassola *et al.*, 2007a). The behavioural advantages and disadvantages of either approach do however need further study.

- my Boundtaxis algorithm is not reliant upon the assumption that the marginal distribution $p(x_t|x_{0:t-1}, \mathbf{a}_{0:t-1})$ is unimodal, which is an assumption of Vergassola *et al.*'s Infotaxis.

These two comparisons have shown that my approach is a distinct and valid alternative to Infomax for solving the sensor placement problem.

The preferences of my measure $B_{X_t|a_t} [L_t||L'_t]$ differs from that of Shannon information $I [L_t|X_t, a_t]$ and mutual information $I_{\text{MI}} [X_t; L_t|a_t]$. Intuitively $I [L_t|X_t, a_t]$ selects the action with least uncertainty. In the simplest unimodal case this is typically the action expected to produce the sharpest posterior $p(l_t|x_t, a_t)$ from a measurement. In contrast $B_{X_t|a_t} [L_t||L'_t]$ will by Theorem 3.10 select the action expected to maximise the dissimilarity between each pair of predicted *a posteriori* beliefs, $p(l_t|x_t, a_t)$ and $p(l_t|x'_t, a_t)$.² This is equivalent to selecting an action a_t to maximise the expected dissimilarity between each possible pair of posterior beliefs $p(l_t|x_t, a_t)$ and $p(l_t|x'_t, a_t)$, weighted according to each unique pair of measurements x_t and x'_t by a factor $\frac{1}{2}p(x_t|a_t)p(x'_t|a_t)$. Hence, while $I [L_t|X_t, a_t]$, $I_{\text{MI}} [X_t; L_t|a_t]$ and $B_{X_t|a_t} [L_t||L'_t]$ will select a_t to increase our knowledge of the latent space L_t , the actions will differ due to each measure's description of useful knowledge.

The group of domains, where my approach has the potential to be superior to Infomax, has the property that the likelihood of a measurement $p(x_t|l_t, a_t)$ is non-linear. This in turn causes the *a posteriori* belief $p(l_t|x_t, a_t)$ to be non-linear. As a consequence, when applying Infomax, it is necessary to approximate the integral $I [L_t|X_t, a_t]$. For instance, two approaches taken in the literature are:

1. approximation of $p(x_t|x_{0:t-1}, a_{0:t})$ as in Vergassola *et al.* (2007a),
2. or sampling the distribution $p(x_t, l_t|x_{0:t-1}, a_{0:t})$ as in Porta *et al.* (2005).

The conditional Shannon information is typically incomputable without some approximation Porta *et al.* (2005), Fox *et al.* (1998), Vergassola *et al.* (2007a). In developing my method I have presented a valid alternative to Infomax for satisfactorily solving the problem of optimal action selection. My approach makes action selection easier and faster when the averaged quantities are analytic, making it less computationally costly when compared with Infomax approaches like Fox *et al.* (1998), Porta *et al.* (2005), Cassandra *et al.* (1996),

²The KL-divergence between each pair of *a posteriori* belief is averaged over the expected measurements x_t and x'_t as

$$B_{L_t|a_t} [X_t||X'_t] = \int dx_t dx'_t p(x_t|a_t) p(x'_t|a_t) D_{\text{KL}} [p(L_t|x_t, a_t) || p(L_t|x'_t, a_t)]. \quad (5.5)$$

Vergassola *et al.* (2007a). Also, in Appendix K I have shown that my framework scales as well as Infomax for partially observable Markov decision processes (POMDPs).

An important point regarding my approach is that even when the KL-divergence $D_{\text{KL}} [p(X_t|l_t, a_t) || p(X_t|l'_t, a_t)]$ is analytic, the *a posteriori* belief $p(l_t|x_{0:t}, a_{0:t})$ is not guaranteed to be analytic.³ Therefore, an average over an analytic KL-divergence is likely to be a more stable computation than an average over a sampled *a posteriori* belief. This makes the estimation of $B_{X_t|a_t} [L_t || L'_t]$ likely to be more stable than that of $I [L_t | X_t, a_t]$. However, this stability argument does not necessarily apply to those cases where the KL-divergence of the measurement likelihoods is non-analytic.

The work of Porta *et al.* (2003, 2005), Fox *et al.* (1998), Cassandra *et al.* (1996) and Vergassola *et al.* (2007a) provides insight into more complex localisation tasks. For instance if a consequence of an action is to capture a target then the best localisation action naturally has the result of being both exploitative and exploratory. Vergassola *et al.* (2007a) postulate that these information terms are computed by an individual for localisation. However, particular action policies may be represented satisfactorily with a cheaper approximation. An example of this is the use of finite state machines (Thill and Pearce, 2007). Hence, leading to the question, why does a more expensive information calculation need to be computed when a simpler method produces an equivalent result and similar choice of action?

This is most useful to problems that can be represented behaviourally:

1. For example, in the multimodal case, it is a reflex action of human subjects to seek to bring a sound event within view of their eyes (Blauert, 1997). This is in agreement with my hypothesis, as in this instance, the individual can only increase its localisation certainty with a simple and easy heuristic to select an action.
2. Taking as an example, head rotations for auditory localisation, it is possible to argue that *a priori* beliefs can be classified, as in Subsections 2.3.2 & 2.3.3, and assigned a probability distribution for the candidate actions. An example of this approach can be seen in Subsection 3.6.2. I classified an *a priori* belief according to its Shannon information as either ambiguous or unambiguous. Then I used this to construct two stochastic policies according to the informatic policies in Section 3.6. The Infomax

³Many problems use Gaussian processes for the likelihood of the continuous spaces. When the measurement space is continuous, this leads to the term $B_{X_t|a_t} [L_t || L'_t]$ averages being analytic. In contrast the term $I [L_t | X_t, a_t]$ averages is the *a posteriori* belief and is not necessarily so – this is due to a non-linear Gaussian likelihood, causing the posterior to be non-linear as well.

policy was well represented by this stochastic strategy, which makes the stochastic strategy a valid alternative. In contrast my approach $B_{X_t|a_t} [L_t||L'_t]$ had a more complex relationship with *a priori* belief than Infomax does – the stochastic strategy constructed from my approach significantly under performed my approach.

3. Moth behaviour can be explained as using either an Infomax measure (Vergassola *et al.*, 2007a) or to have an equivalent choice of predicted action using a finite state machine. It is obvious from a neural perspective, that it is far cheaper to compute an action from a finite state machine than to perform an argument-maximisation on an informatic quantity. Therefore if a behavioural decision process for selecting an action equivalent to an informatic policy exists, then the agent should use this behavioural representation and its policy.

The consequence of this for the animal is that there can exist a heuristic to give a cheaply calculated but good localising action.

A constraint for reflexive localising action selection is that any calculation needs to be fast and applicable to the individual’s senses. However, should the animal utilise an informatic method, like my approach or Infomax, it can learn and act in a new and unknown environment or learn to best utilise a new sense (Nagel *et al.*, 2005). This is the principle advantage of the informatic approaches over an approach using approximated heuristic.

Chapter 3 considered the problem of reflexive localisation actions. This complements the work on sensorimotor contingencies of Noe (2004), O’Regan and Noe (2001). Sensorimotor contingencies consider the relationship between actions and sensory measurements to map the correlations of these as a latent space Noe (2004), O’Regan and Noe (2001), Bompas and O’Regan (2006), Philipona *et al.* (2003, 2004), Aytakin *et al.* (2008). This however neglects ambiguity, that exists in a large number of problems, which can cause difficulty for any learning process. Hence an informatic policy model such as mine or Infomax, which by its nature assumes an uncertain and noisy world, will have an advantage when combined with a learning algorithm. Further, my approach is useful for a class of problem such as POMDPs or likelihoods, where the KL-divergence is analytic, by scaling as well as or better than other informatic approaches. Examples of this for Infomax approximate the marginal $p(x_t|x_{0:t-1}, a_{0:t})$ as in Vergassola *et al.* (2007a), or assuming a sample of the joint latent and measurement variables $L \cup X$ as in Porta *et al.* (2005).

5.3.1 Future comparisons

The most interesting aspect of Vergassola *et al.* (2007a) was linking turbulent chemotaxis to an informatically guided selection of action that resulted in qualitatively equivalent biological behaviours. This leads to interesting questions of what insights could be garnered for other perception tasks that have an active component and operate under uncertainty? Some possibilities include:

1. Tracking a target in a cluttered environment (Moss and Surlykke, 2001). An emergent consequence could be in maintaining visibility (LaValle *et al.*, 1997).
2. Pack hunting, where Spletzer and Taylor (2003) shared inference between agents. More interesting questions would consider lossy channels of communication similar to vocalisations.
3. Stealthy targets, which is somewhat related to the problem of clutter (Ristic *et al.*, 2004b) when looking at the perspective of camouflage.

Any investigation of such questions must consider multimodal and multiobject tracking under uncertainty – this should be posed in a unified manner (e.g. Hospedales *et al.*, 2007). Thus, there are examples where information guided active-perception could be applied to examine the behavioural consequences, as Vergassola *et al.* did with chemotaxis search in turbulent flows.

Next, I consider the analogues between Chapters 3 & 4, the Fisher information matrix (FIM) and the Cramer-Rao lower bound (CRLB) measures for selecting optimal manoeuvres in radar and sonar problems (Passerieux and van Cappel, 1998, Cadre and Gauvrit, 1996, Helferty and Mudgett, 1993, Logothetis *et al.*, 1998, Ristic *et al.*, 2004a). The result of my framework (Chapter 3) can be seen specifically in Figures 3.4 & 3.5 where most directional ambiguity will be removed through a head rotation. The next best action, if it were available in both cases, would be to approach or pass by the target. Similarly, this combination of turn and fly by is exhibited as a consequence of selecting manoeuvres according to FIM (Helferty and Mudgett, 1993). This indicates a potential commonality of the behaviour between the selection of informatic based actions or manoeuvres.

Passerieux and van Cappel (1998), Cadre and Gauvrit (1996), Helferty and Mudgett (1993), Logothetis *et al.* (1998), Ristic *et al.* (2004a) used the model's FIM which is the

higher dimensional representation of Fisher information

$$I_{\mathcal{F}}[\mathbf{L}|\mathbf{X}, a] = - \int dl dx p(x, l|a) \frac{\partial^2}{\partial x^2} \log p(l|x, a). \quad (5.6)$$

Fisher information ($I_{\mathcal{F}}$) can be seen as the expected sharpness in any belief of $l \in \mathbf{L}$ according to $x \in \mathbf{X}$. By contrast my work is related to the conditional Shannon information

$$I[\mathbf{L}|\mathbf{X}, a] = \int dl dx p(x, l|a) \log p(l|x, a). \quad (5.7)$$

Shannon information is an inverse measure of the smoothness of a probability distribution. A more complete investigation of the shared properties for identical models of these two information measures would be interesting. While there is likely to be great similarity, there is the prospect for subtle differences in the respective consequences for particular situations and problems. An example would be considering a time slice of future beliefs, similar to the application of CRLB in radar (Helferty and Mudgett, 1993). According to Theorem J.3 my measure becomes

$$B_{\mathbf{X}_{t:t+\delta t}|\mathbf{a}_{t:t+\delta t}}[\mathbf{L}_{t:t+\delta t}|\mathbf{L}'_{t:t+\delta t}] = \sum_{t'=t}^{t+\delta t} \int dl'_{t'} p(l'_{t'}|\mathbf{a}_{t:t'}) p(l'_{t'}|\mathbf{a}_{t:t'}) \times \\ D_{\text{KL}}[p(\mathbf{X}_{t'}|l'_{t'}, \mathbf{a}_{t'}) || p(\mathbf{X}_{t'}|l'_{t'}, \mathbf{a}_{t'})]. \quad (5.8)$$

Though, the Shannon information is not as easy to simplify, it remains as

$$I[\mathbf{L}_{t:t+\delta t}|\mathbf{X}_{t:t+\delta t}, \mathbf{a}_{t:t+\delta t}] = \int dl_{t:t+\delta t} dx_{t:t+\delta t} p(x_{t:t+\delta t}, l_{t:t+\delta t}|\mathbf{a}_{t:t+\delta t}) \log p(l_{t:t+\delta t}|x_{t:t+\delta t}, \mathbf{a}_{t:t+\delta t}). \quad (5.9)$$

This is the expectation of the complete data *a posteriori* belief. For any problem calculating the integral for $I[\mathbf{L}_{t:t+\delta t}|\mathbf{X}_{t:t+\delta t}, \mathbf{a}_{t:t+\delta t}]$ will be very difficult and require significant computation resources. In contrast, calculating $B_{\mathbf{X}_{t:t+\delta t}|\mathbf{a}_{t:t+\delta t}}[\mathbf{L}_{t:t+\delta t}|\mathbf{L}'_{t:t+\delta t}]$ simply needs a prediction of the latent state *a priori* belief $p(l_{t'}|\mathbf{a}_{t:t'})$, where $t \leq t' \leq t + \delta t$. A possible implementation could constrain this using NBV algorithm (Gonzalez-Banos and Latombe, 2002) with collections of possible action chains $\left\{ \mathbf{a}_{t:t+\delta t}^{(j)} \right\}_{j=1:N}$.⁴

⁴Also, from Helferty and Mudgett (1993) the CRLB measure allowed discounting of particular latent state space (\mathbf{L}) axes. For my approach, particular dimensions in \mathbf{L} could be marginalised to favour certainty in portions of this latent state space \mathbf{L} . Though this would complicate the computation of $B_{\mathbf{X}_t|\mathbf{a}_t}[\mathbf{L}_t|\mathbf{L}'_t]$.

This approach considers the problem of reflexive localisation actions. This complements the work on sensorimotor contingencies of Noe (2004). Sensorimotor contingencies consider the relationship between actions and sensory measurements to map the correlations of these as a latent space Noe (2004), O'Regan and Noe (2001), Bompas and O'Regan (2006), Philipona *et al.* (2003, 2004), Aytekin *et al.* (2008). This however neglects uncertainty and ambiguity, which exists in a large number of problems and can cause difficulty in any learning process. Hence an approach such as mine, which by nature assumes an uncertain and noisy world, should have an advantage when combined with a learning algorithm. Further, an informatic approach such as mine, as I have already shown, is useful as it appears to reduce latent space uncertainty.

5.4 Active learning

In Chapter 4 “Learning a model for an active agent” I applied my approach developed in Chapter 3 to a dynamic problem domain to learn a model semi-concurrently with action selection and observing. This addressed the problem of active experimentation, and indicated that selecting an action according to its expected knowledge is better than simply body-babbling (Olsson *et al.*, 2004, 2005a,b, 2006). My findings show a trend for the more directed policies to achieve higher certainties in inference and system model likelihoods after repeated iteration of learning and sampling episodes. This is due to the influence of actions allowing the agent to learn from more certain and accurate inferences. However, an interesting qualification was that the applied Infomax and CRLB policies constructed inference certainties and likelihoods that were only marginally better than the body-babbling policy.

A further finding was that when the measurement process is arbitrarily remapped the system can readily adjust to this remapping more quickly than it took to originally learn the unremapped system. This is similar in consequence to the bistable sensory adaptation observed by Welch *et al.* (1993), Hofman *et al.* (1998), a short period of adaptation followed by a return to prior performance (*cf* Figure 4.18).

Chapter 4 is limited to learning a predefined model's parameterisation for iterating between querying the model's belief in the system and refining its parameterisation through learning. Though this illustrates the utility in active-perception for adapting the correspondence (symbol grounding) of the model's state to observations by an active choice of what to measure, it does not fully ground my framework in the sensorimotor ideas of Noe

(2004), O'Regan and Noe (2001). This limitation does however elicit more general directions that I could take in future research. A significant potential area of research concerns those active perception problems where the agent holds no real prior knowledge about the structure of its environment. One solution is to apply a neural network model (Droulez and Berthoz, 1991) for retaining the contingencies between the sensory and motor spaces and using my approach, in Chapter 3, to automatically explore the sensorimotor contingencies to best effect. A further solution could model the sensorimotor contingencies using a locally weighted projection regression (LWPR). This is a locally linear non-parametric model that uses receptive weighting to maintain its locality (Vijayakumar and Schaal, 1998, 2000, Vijayakumar *et al.*, 2005). DeMarse *et al.* (2001) and DeMarse and Dockendorf (2005) illustrate this for biological neural networks, which are certainly capable of this. Specifically DeMarse and Dockendorf (2005) used rat neurons to operate an aircraft in a flight simulator. The implication of my thesis, from Chapters 3 & 4, is how to best select the actions to explore these contingencies efficiently.

5.5 Final remarks

Localisation of places, prey and predators are usually of critical behavioural importance to any organism and hence to its survival. In this thesis I have illustrated the difficulties of passive localisation (Chapter 2) and the advantages of being an active observer, specifically with a measure to select the action giving the NBV (Chapters 3 & 4). Though, not a true sensorimotor model, it does indicate that the choice of view (Walker *et al.*, 1998) and NBV (Gonzalez-Banos and Latombe, 2002) are important and influential for an individual in learning to represent its environment.

To think of action as merely a physical motion or manoeuvre is to ignore its usefulness as a metaphor for a much wider class of problems and processes. We can instead see an action as Gibson (1978) does as the offering of affordances by the environment – making our senses the channel by which the environment answers the questions asked by our actions. This leads to the idea of sensorimotor contingencies between actions (queries) and sensations (answers) (Noe, 2004, O'Regan and Noe, 2001). This is the structured coupling of action and perception related to enactive perception (Vergassola *et al.*, 2007a). This makes the distinction between the brain's representation of a tool and the hand that wields it irrelevant, as we should instead be considering the effect (Hochberg *et al.*, 2006, Umiltà *et al.*, 2008). Also, this makes irrelevant the distinction between how we view the

world and our feeling of presence within it. For example, there are distinctions between remote viewing, virtual reality and local reality (Lenggenhager *et al.*, 2007, Ehrsson, 2007, IJsselsteijn and Reiner, 2004).

Finally, one might consider O'Regan and Noe (2001) sensorimotor contingencies from the perspective of Ito's Lemma, as this would take into account uncertainty through the dynamic propagation of probability in continuous time. This would be an interesting direction to try to take (Philipona *et al.*, 2003, 2004, Aytakin *et al.*, 2008) for the sensorimotor hypothesis, as the analysis of the stochastic differential equations could provide a unique insight into the policy directives of action selection under sensorimotor relationships.

Appendix A

Analysis of the EM-algorithm

Expressing the bound, Equation B.9, in a more compact form, where,

$$q_j(\mathbf{s}, \tau) = p^{\text{old}}(\mathbf{s}, \tau | \mathbf{x}_{1,j}, \mathbf{x}_{2,j}, l_j), \quad (\text{A.1})$$

which takes the bound,

$$B_{\text{old}}^{\text{new}} = \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log p(\mathbf{x}_{1,j}, \mathbf{x}_{2,j}, \mathbf{s}, \tau | l_j) - \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log q^{\text{old}}(\mathbf{s}, \tau). \quad (\text{A.2})$$

So taking this representation and expanding,

$$B_{\text{old}}^{\text{new}} = \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log p(\mathbf{x}_{1,j}, \mathbf{x}_{2,j}, \mathbf{s}, \tau | l_j) - \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log q_j(\mathbf{s}, \tau), \quad (\text{A.3})$$

$$= \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log p(\mathbf{x}_{1,j}, \mathbf{x}_{2,j}, \mathbf{s}, \tau | l_j) + \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log \frac{q_j^{\text{new}}(\mathbf{s}, \tau)}{q_j(\mathbf{s}, \tau)}, \quad (\text{A.4})$$

$$= \sum_j \log p(\mathbf{x}_{1,j}, \mathbf{x}_{2,j} | l_j) + \sum_{j,\tau} \int d\mathbf{s} q_j(\mathbf{s}, \tau) \log \frac{q_j^{\text{new}}(\mathbf{s}, \tau)}{q_j(\mathbf{s}, \tau)}, \quad (\text{A.5})$$

$$= \mathcal{L}(\mathbf{x}_{1,j=1:J}, \mathbf{x}_{2,j=1:J} | l_{j=1:J}) - \sum_j D_{KL}[q_j(\mathbf{s}, \tau) || q_j^{\text{new}}(\mathbf{s}, \tau)]. \quad (\text{A.6})$$

As the KL-divergence between two distributions is always non-negative we see that,

$$B_{\text{old}}^{\text{new}} \leq \mathcal{L}(\mathbf{x}_{1,j=1:J}, \mathbf{x}_{2,j=1:J} | l_{j=1:J}). \quad (\text{A.7})$$

Thus both Equations 2.31 & A.7, requires that any selection of a new parameter is that the bound touches the likelihood. Thus each iteration of the EM-algorithm selects a new parameterisation (conditional upon the old) such that both the bound and the log-likelihood are maximised. This is typically iterated until convergence; when the sum of the divergence terms between the old latent posterior $q_j(\mathbf{s}, \tau)$ and the new latent posterior $q_j^{\text{new}}(\mathbf{s}, \tau)$ becomes zero.

Appendix B

The variational responsibilities and updates

I shall show that the optimal q is in fact the posterior of the hidden latent variables over the data,

$$q_{j,n}(\mathbf{s}, \tau) = p^{\text{old}}(\mathbf{s}, \tau | l_j, \mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}), \quad (\text{B.1})$$

where $p^{\text{old}}(\mathbf{s}, \tau | l_j, \mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n})$ represents the old parameterisations posterior belief. Hence the responsibilities have already been defined in Subsection 2.2.4. Firstly using the bound upon the likelihood,

$$B_{\text{new}}^{\text{old}} = \sum_{n,j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \log p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau | l_j) - \sum_{n,j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \log q_{j,n}(\mathbf{s}, \tau), \quad (\text{B.2})$$

now, I add a Lagrangian multiplier $\alpha_{j,n}$ for the set of constraints expressed by, $\sum_{\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) = 1 \forall j, n$. Thus the Lagrangian is,

$$G_{\text{old}}^{\text{old}} = B_{\text{old}}^{\text{old}} - \sum_{n,j} \alpha_{j,n} \left(1 - \sum_{\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \right). \quad (\text{B.3})$$

Taking the gradient with respect to $q_{j,n}$ of the Lagrangian and equating it to zero to find the Lagrangian stationary point with respect to $q_{j,n}(\mathbf{s}, \tau)$ gives,

$$q_{j,n}(\mathbf{s}, \tau) = \frac{p^{\text{old}}(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau | l_j)}{e^{1-\alpha_{j,n}}}. \quad (\text{B.4})$$

So as $q_{j,n}$ is normalised, according to the conditions $\sum_{\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) = 1 \forall j, n$, then the denominator is,

$$e^{1-\alpha_{j,n}} = \sum_{\tau} \int d\mathbf{s} p^{\text{old}}(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau | l_j), \quad (\text{B.5})$$

$$= p^{\text{old}}(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n} | l_j). \quad (\text{B.6})$$

Inserting this into Equation B.4 and applying Bayes law give the responsibilities,

$$q_{j,n}(\mathbf{s}, \tau) = \frac{p^{\text{old}}(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau | l_j)}{p^{\text{old}}(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n} | l_j)}, \quad (\text{B.7})$$

$$= p^{\text{old}}(\mathbf{s}, \tau | l_j, \mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}). \quad (\text{B.8})$$

Substituting $q_{j,n}$, Equation B.8, into the bound, Equation 2.31, produces,

$$\begin{aligned} B_{\text{new}}^{\text{old}} &= \sum_{j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \log p(\mathbf{x}_{1,j,n}, \mathbf{x}_{2,j,n}, \mathbf{s}, \tau | l_j) - \\ &\quad \sum_{j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \log q_{j,n}(\mathbf{s}, \tau). \end{aligned} \quad (\text{B.9})$$

Hence the responsibilities are the old parameterisations latent variable posterior beliefs defined in Subsection 2.2.4. Now the bound is in a form that can be optimised.

The derivatives of the bound for each parameter allows the EM-updates to be derived. I present the derivation of one EM-update in full. This is computed from the gradient of the bound as follows,

$$\frac{\partial}{\partial \eta^{-1}} B_{\text{new}}^{\text{old}} = \sum_{n,j,\tau} \int d\mathbf{s} q_{j,n}(\mathbf{s}, \tau) \frac{\partial}{\partial \eta^{-1}} \left\{ \frac{D_x}{2} \log \eta^{-1} - \frac{1}{2\eta} \mathbf{s}^T \mathbf{s} \right\}, \quad (\text{B.10})$$

$$= \frac{1}{2} \sum_{n,j,\tau} q_{j,n}(\tau) \int d\mathbf{s} q_{\tau,j,n}(\mathbf{s}) (D_x \eta - \mathbf{s}^T \mathbf{s}), \quad (\text{B.11})$$

$$= \frac{1}{2} \sum_{j,\tau} q_{j,n}(\tau) \left(D_x \eta - \left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} - D_x \nu_{s|j}^{\text{old}} \right), \quad (\text{B.12})$$

which equating to zero and rearranging, results in the EM-update, for the variance of the source signal,

$$\eta = \frac{1}{D_x J N} \sum_{n,j,\tau} q_{j,n}(\tau) \left(\left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} + D_x \nu_{s|j}^{\text{old}} \right), \quad (\text{B.13})$$

where $\frac{1}{\nu_{s|j}^{\text{old}}} = \frac{1}{\eta^{\text{old}}} + \frac{(\lambda_{1|j}^{\text{old}})^2}{\nu_1^{\text{old}}} + \frac{(\lambda_{2|j}^{\text{old}})^2}{\nu_2^{\text{old}}}$ and $\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} = \nu_{s|j}^{\text{old}} \left(\frac{\lambda_{1|j}^{\text{old}}}{\nu_1^{\text{old}}} \mathbf{x}_{1,j,n} + \frac{\lambda_{2|j}^{\text{old}}}{\nu_2^{\text{old}}} \hat{\mathbf{D}}_{\tau}^T \mathbf{x}_{2,j,n} \right)$. The

remaining list of these updates are,

$$\mathbf{w}_1 = \left[\sum_j \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T \left\{ \sum_{n,\tau} q_{j,n}(\tau) \left(\left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} + D_x \nu_{s|j}^{\text{old}} \right) \right\} \right]^{-1} \sum_j \boldsymbol{\psi}_j \sum_{n,\tau} q_{j,n}(\tau) \mathbf{x}_{1,j,n}^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}}, \quad (\text{B.14})$$

$$\mathbf{w}_2 = \left[\sum_j \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T \left\{ \sum_{n,\tau} q_{j,n}(\tau) \left(\left(\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} + D_x \nu_{s|j}^{\text{old}} \right) \right\} \right]^{-1} \sum_j \boldsymbol{\psi}_j \sum_{n,\tau} q_{j,n}(\tau) \mathbf{x}_{2,j,n}^T \hat{\mathbf{D}}_\tau \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}}, \quad (\text{B.15})$$

which causes $\lambda_{1|j} = \mathbf{w}_1^T \boldsymbol{\psi}_j$ and $\lambda_{2|j} = \mathbf{w}_2^T \boldsymbol{\psi}_j$,

$$\nu_1 = \frac{1}{D_x J N} \sum_{n,j,\tau} q_j(\tau) \left[\left(\mathbf{x}_{1,j,n} - \lambda_{1|j} \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \left(\mathbf{x}_{1,j,n} - \lambda_{1|j} \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right) + D_x \lambda_{1|j}^2 \nu_{s|j}^{\text{old}} \right], \quad (\text{B.16})$$

$$\nu_2 = \frac{1}{D_x J N} \sum_{n,j,\tau} q_{j,n}(\tau) \left[\left(\mathbf{x}_{2,j,n} - \lambda_{2|j} \hat{\mathbf{D}}_\tau \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right)^T \left(\mathbf{x}_{2,j,n} - \lambda_{2|j} \hat{\mathbf{D}}_\tau \boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} \right) + D_x \lambda_{2|j}^2 \nu_{s|j}^{\text{old}} \right], \quad (\text{B.17})$$

$$\mathbf{w}_\tau = \left[N \sum_j \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T \right]^{-1} \sum_j \boldsymbol{\psi}_j \sum_{n,\tau} q_{j,n}(\tau) \tau, \quad (\text{B.18})$$

which causes $\gamma_{\tau|j} = \mathbf{w}_\tau^T \boldsymbol{\psi}_j$,

$$\bar{\mathbf{w}}_\omega = \left[N \sum_j \boldsymbol{\psi}_j \boldsymbol{\psi}_j^T \right]^{-1} \sum_j \boldsymbol{\psi}_j \sum_{n,\tau} q_{j,n}(\tau) (\tau - \gamma_{\tau|j})^2, \quad (\text{B.19})$$

where $\frac{1}{\nu_{s|j}^{\text{old}}} = \frac{1}{\eta^{\text{old}}} + \frac{(\lambda_{1|j}^{\text{old}})^2}{\nu_1^{\text{old}}} + \frac{(\lambda_{2|j}^{\text{old}})^2}{\nu_2^{\text{old}}}$ and $\boldsymbol{\mu}_{s|\tau,j,n}^{\text{old}} = \nu_{s|j}^{\text{old}} \left(\frac{\lambda_{1|j}^{\text{old}}}{\nu_1^{\text{old}}} \mathbf{x}_{1,j,n} + \frac{\lambda_{2|j}^{\text{old}}}{\nu_2^{\text{old}}} \hat{\mathbf{D}}_\tau^T \mathbf{x}_{2,j,n} \right)$.

Appendix C

Nix and Hohmann (2006) ILD and IPD computations

The interaural level disparities (ILD) and interaural phase disparities (IPD) were computed by Nix and Hohmann (2006) using a windowed short-term fast Fourier transform (FFT) analysis. I specifically use only one frequency channel, this makes their model directly comparable to mine. The data set was constructed as segments of length 440 to correspond with the segment length used for my model, these were each zero padded up to a length of 512 for applying an FFT. Nix and Hohmann used segments of length 400 with zero padding. The segments each overlapped by 220 samples. This corresponds to a window duration of ~ 10 ms and a window time shift ~ 5 ms.

Each of these segments were multiplied by a Hann window and transformed using a FFT. These short-term FFT spectra of the right and left sensors, respectively sensors 1 and 2, are denominated $F_1(f, n)$ and $F_2(f, n)$ where f and n are respectively the frequency and time segment index.

A simplification that I apply is to consider only one frequency channel, due to the simplicity of my sound source and experimental setup by contrast to that of Nix and Hohmann. They constructed an analysis for far more complex signals, including speech, in the presence of background noise. Hence they used multiple frequency channels. In contrast my experimental setup used a sound of very simple time and intensity structure without frequency dependence. This means there is no necessity for multiple frequency channels; hence, I used a single channel.

To compute the ILD and IPD Nix and Hohmann made use of frequency averaging using

the squared magnitude spectrums and the complex-valued cross spectrum,

$$F_{11}(n) = \sum_f |F_1(f, n)|^2, \quad (\text{C.1})$$

$$F_{22}(n) = \sum_f |F_2(f, n)|^2, \quad (\text{C.2})$$

$$F_{12}(n) = \sum_f F_1(f, n) F_2(f, n)^*, \quad (\text{C.3})$$

where $F_2(f, n)^*$ indicates the complex conjugate of $F_2(f, n)$. The ILDs and IPDs were computed as,

$$\text{ILD}_n = 10 \log \left| \frac{F_{11}(n)}{F_{22}(n)} \right|, \quad (\text{C.4})$$

$$\text{IPD}_n = \arg F_{12}(n), \quad (\text{C.5})$$

for each segment indexed by n .

To summarise, the only changes I made to the approach and model of Nix and Hohmann was in the computations of the ILD and IPD. These changes were to adapt their approach to become a valid comparison to my model using my data set. Specifically these were the use of one frequency channel, due to the nature of my data set, and using a segment length of 440 to correspond with the segment length used in my model.

Appendix D

The Kullback-Leibler divergence

The Kullback-Leibler (KL) divergence between two probability density functions, p and q , over the space $x \in X$ is defined to be (Cover and Thomas, 2006, Kullback and Leibler, 1951),

$$D_{KL} [f||g] = \int dx f(x) \log \frac{f(x)}{g(x)}. \quad (\text{D.1})$$

it has the property $D_{KL} [f||g] \geq 0$, and where $f(x) = g(x) \forall x$ it follows that $D_{KL} [f||g] = 0$. As $D_{KL} [f||g] = \alpha$ does not imply that $D_{KL} [g||f] = \alpha$. The KL-divergence is not symmetric as can be seen,

$$D_{KL} [f||g] \neq D_{KL} [g||f], \quad (\text{D.2})$$

by,

$$\int du f(u) \log \frac{f(u)}{g(u)} \neq \int du g(u) \log \frac{g(u)}{f(u)}. \quad (\text{D.3})$$

To apply the KL-divergence to a pair of Gaussian distributions. Using,

$$D_{KL} [p||q] \triangleq \int dx p(x) \log \frac{p(x)}{q(x)}, \quad (\text{D.4})$$

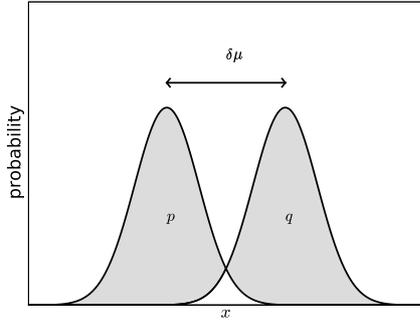
with $p(x) = \phi(x; \mu_p, \sigma_p)$ and $q(x) = \phi(x; \mu_q, \sigma_q)$. So,

$$D_{KL} [p||q] = -\frac{1}{2} \log(2\pi\sigma_p e) - \int dx p(x) \log q(x), \quad (\text{D.5})$$

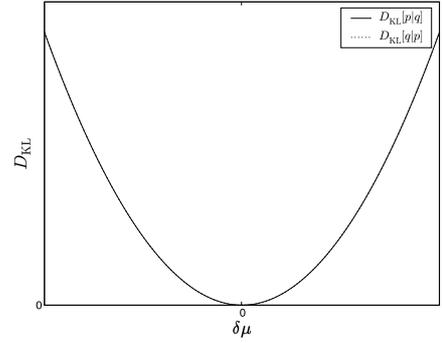
where the cross term simplifies to,

$$\int dx p(x) \log q(x) = \int dx \phi(x; \mu_p, \sigma_p) \log \phi(x; \mu_q, \sigma_q), \quad (\text{D.6})$$

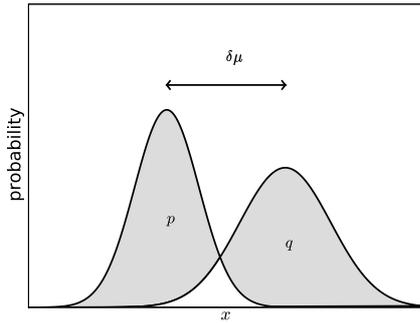
$$= \frac{1}{2} \int dx \phi(x; \mu_p, \sigma_p) \left\{ \log \frac{\sigma_q}{2\pi} - \sigma_q (x - \mu_q)^2 \right\}, \quad (\text{D.7})$$



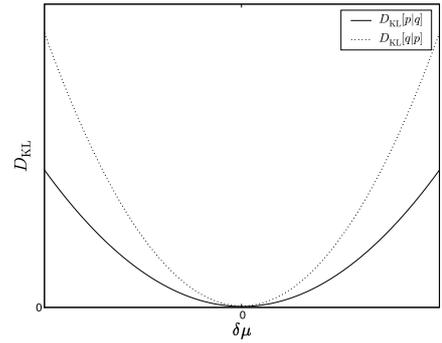
(a) Density plots for two Gaussians of identical variance.



(b) KL-divergences for identical variances.



(c) Density plots for two Gaussians of different variance, specifically.



(d) KL-divergences for different variances, specifically.

Figure D.1: Figure illustrating the impact that varying the mean offset between two Gaussians has upon the KL-divergence. The offset $\delta\mu$ is between the mean of the two Gaussian distribution. Two cases are considered, (i) identical variances $\sigma_p = \sigma_q$, plot (a) and (b) different variances $\sigma_p \neq \sigma_q$; specifically $\sigma_p < \sigma_q$, plot (b). Plot (c) depicts the case of identical variance which causes both divergences $D_{\text{KL}}[p||q]$ and $D_{\text{KL}}[q||p]$ to be identical across $\delta\mu$. In contrast, plot (d) depicts the case of differing variances which causes the divergences to diverge with increasing absolute offset $|\delta\mu|$. Plot (d) depicts the case that $D_{\text{KL}}[q||p] > D_{\text{KL}}[p||q] \forall \delta\mu$ iff $\sigma_p < \sigma_q$. As can be seen the KL-divergence is inversely related to the overlap of the two densities labelled p and q .

$$= -\frac{1}{2} \log(2\pi\sigma_q) - \frac{1}{2\sigma_q} \int dx \phi(x; \mu_p, \sigma_p) (x^2 - 2x\mu_q + \mu_q^2), \quad (\text{D.8})$$

$$= -\frac{1}{2} \log(2\pi\sigma_q) - \frac{1}{2} \frac{\sigma_p}{\sigma_q} - \frac{1}{2\sigma_q} (\mu_p - \mu_q)^2. \quad (\text{D.9})$$

So the divergence is,

$$D_{\text{KL}} [p||q] = \frac{1}{2} \left(\frac{\sigma_p - \sigma_q}{\sigma_q} \right) + \frac{1}{2\sigma_q} (\mu_p - \mu_q)^2 + \frac{1}{2} \log \frac{\sigma_p}{\sigma_q}, \quad (\text{D.10})$$

as $\mu_q \rightarrow \mu_p$ and $\sigma_q \rightarrow \sigma_p$ then the divergence $D_{\text{KL}} [p||q] \rightarrow 0$. As can be seen the polynomial term makes the divergence sensitive to large differences between μ_p and μ_q . This can be seen in Figure D.1.

The KL-divergence of two Gaussian distribution with full covariance matrices, represented by $\phi(\mathbf{x}; \boldsymbol{\mu}_1, \Sigma_1)$ with mean $\boldsymbol{\mu}_1$ and covariance Σ_1 and $\phi(\mathbf{x}; \boldsymbol{\mu}_2, \Sigma_2)$ with mean $\boldsymbol{\mu}_2$ and covariance Σ_2 , is

$$\begin{aligned} D_{KL} [\phi(\mathbf{x}; \boldsymbol{\mu}_1, \Sigma_1) || \phi(\mathbf{x}; \boldsymbol{\mu}_2, \Sigma_2)] &= \int dx \phi(\mathbf{x}; \boldsymbol{\mu}_1, \Sigma_1) \log \frac{\phi(\mathbf{x}; \boldsymbol{\mu}_1, \Sigma_1)}{\phi(\mathbf{x}; \boldsymbol{\mu}_2, \Sigma_2)}, \quad (\text{D.11}) \\ &= \frac{1}{2} \log \frac{|\Sigma_2|}{|\Sigma_1|} + \frac{1}{2} \text{tr} [\Sigma_2^{-1} (\Sigma_1 - \Sigma_2)] \\ &\quad + \frac{1}{2} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma_2^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2). \quad (\text{D.12}) \end{aligned}$$

Appendix E

Local weight analysis

Local weight analysis uses the curvature (from geometry) of the likelihood (or a bound upon it) to consider the unit weighting of a factor or portion of a model to perturbations. For a curve, $c(\delta) = \{u(\delta), v(\delta)\}$ on the u, v -plane which is a function of δ , the curvature is defined as,

$$C(\delta) = \frac{\left| \frac{\partial u}{\partial \delta} \frac{\partial^2 v}{\partial \delta^2} - \frac{\partial v}{\partial \delta} \frac{\partial^2 u}{\partial \delta^2} \right|}{\left(\left(\frac{\partial u}{\partial \delta} \right)^2 + \left(\frac{\partial v}{\partial \delta} \right)^2 \right)^{\frac{3}{2}}}. \quad (\text{E.1})$$

I use as the curve c a quantity representing a functional measure of similarity of a probability belief to its unperturbed probability. The measure of similarity between the point belief of a model Ω_a and the perturbation Ω_b applied to itself is defined to be the KL-divergence of these two distributions. The perturbation is defined as a function of δ , $\Omega_b(\delta)$, which applies the analysis for the cases in Figure 2.15.

As I am considering the spatial dependence of the parameters (Figure 2.15), I consider the curvature is the line formed by $c(\delta) = \{\delta, \tilde{d}_j(\delta)\}$ – where, for a location l_j indexed by j , \tilde{d}_j represents the approximated measure of dissimilarity between Ω_a and $\Omega_b(\delta)$. The curvature of $c(\delta)$ for the perturbation $\Omega_b(\delta)$ applied to the set of parameters Ω is

$$C_j(\delta)|_{\Omega \leftarrow \Omega_b(\delta)} = \frac{\left| \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \Big|_{\Omega \leftarrow \Omega_b(\delta)} \right|}{\left(1 + \left(\frac{\partial \tilde{d}_j}{\partial \delta} \Big|_{\Omega \leftarrow \Omega_b(\delta)} \right)^2 \right)^{\frac{3}{2}}}. \quad (\text{E.2})$$

for the perturbation $\Omega_b(r)$ applied to the parameters Ω . To construct a measure of similarity \tilde{d}_j I start from the joint posterior belief in the hidden variables $p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega)$ ¹ and

¹The distribution can be factorised by Bayes law as

$$p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega) = p(l_i | \mathbf{x}_1, \mathbf{x}_2, \Omega) p(\tau | l_i, \mathbf{x}_1, \mathbf{x}_2, \Omega) \times$$

take the KL-divergence for the perturbation defined as $\Omega \leftarrow \Omega_b(\delta)$ to the unperturbed $\Omega \leftarrow \Omega_a$ which can be expressed as

$$D_{\text{KL}} [p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_a) \| p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_b(\delta))]. \quad (\text{E.3})$$

As I wish to analyse the spatial importance of these parameters I take the KL-divergence over this distribution for each true location l_j , indexed by j ; I do this by taking the expectation of the KL-divergence in Equation E.3 for the validation data sets² sample distribution,

$$p(\mathbf{x}_1, \mathbf{x}_2 | l_j) = \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \delta(\mathbf{x}_1 - \mathbf{y}_{1,j,n}) \delta(\mathbf{x}_2 - \mathbf{y}_{2,j,n}). \quad (\text{E.4})$$

Putting this together gives an expectation of the spatial influence of a perturbation as

$$\tilde{d}_j = \mathbb{E}_{p(\mathbf{x}_1, \mathbf{x}_2 | l_j)} [D_{\text{KL}} [p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_a) \| p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_b)]], \quad (\text{E.5})$$

$$= \int d\mathbf{x}_1 d\mathbf{x}_2 p(\mathbf{x}_1, \mathbf{x}_2 | l_j) \sum_{i=1}^J \sum_{\tau} \int d\mathbf{s} p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_a) \times \log \left(\frac{p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_a)}{p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_b)} \times \frac{p(\mathbf{x}_1, \mathbf{x}_2)}{p(\mathbf{x}_1, \mathbf{x}_2)} \right), \quad (\text{E.6})$$

$$= \int d\mathbf{x}_1 d\mathbf{x}_2 p(\mathbf{x}_1, \mathbf{x}_2 | l_j) \sum_{i=1}^J \sum_{\tau} \int d\mathbf{s} p(l_i, \mathbf{s}, \tau | \mathbf{x}_1, \mathbf{x}_2, \Omega_a) \times \log \frac{p(\mathbf{x}_1, \mathbf{x}_2, l_i, \mathbf{s}, \tau | \Omega_a)}{p(\mathbf{x}_1, \mathbf{x}_2, l_i, \mathbf{s}, \tau | \Omega_b)}, \quad (\text{E.7})$$

$$\approx \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int d\mathbf{s} p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \log \frac{p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_a)}{p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b)}. \quad (\text{E.8})$$

To compute the curvature $C_j(\delta)$ in Equation E.2 I need to calculate the first and second derivatives of \tilde{d}_j with respect to δ . As only Ω_b is dependent upon the perturbation by a factor δ then the derivatives of \tilde{d}_j simplify to

$$\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\Omega \leftarrow \Omega_b(\delta)} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int d\mathbf{s} p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b(\delta)), \quad (\text{E.9})$$

$$p(\mathbf{s} | l_i, \tau, \mathbf{x}_1, \mathbf{x}_2, \Omega),$$

which are Equations 2.27, 2.25 & 2.20 respectively.

²The validation measurements are, for each sampled location l_j , $\{\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n} : n \in [1, N_{\text{val}}]\}$.

and

$$\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\Omega \leftarrow \Omega_b(\delta)} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b(\delta)). \quad (\text{E.10})$$

The curvature is a measure of the model's belief to perturbation of a parameter – through variation of the meta parameter δ . I am considering perturbations to the cues:

1. through variation of λ_1 or λ_2 , to represent the monaural cues at either ear ($\lambda_1 \leftarrow \delta \lambda_1$ or $\lambda_2 \leftarrow \delta \lambda_2$),
2. through variation of λ_1 & λ_2 , to represent the average loudness ($\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta \lambda_2$) or the disparity in interaural loudness (either $\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2$ or $\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2$),
3. and through variation of τ , to represent the interaural time disparity ($\gamma_\tau \leftarrow \delta \gamma_\tau$).

I calculate both of the gradients in Equations E.9 & E.10 for each of these cases. The joint posterior belief can be factorised by Bayes law to be

$$p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega) = p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega) p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega) \times p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega), \quad (\text{E.11})$$

where $p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega)$ is defined by Equation 2.27, $p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega)$ by Equation 2.25 and $p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega)$ by Equation 2.20. As the belief $p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a)$ is a Gaussian, of mean $\boldsymbol{\mu}_{\mathbf{s} | \tau, j, i, n} = \nu_{s|i} \left(\frac{\lambda_{1|i}}{\nu_1} \mathbf{y}_{1,j,n} + \frac{\lambda_{2|i}}{\nu_2} \hat{\mathbf{D}}_\tau^T \mathbf{y}_{2,j,n} \right)$ and the variance $\nu_{s|i} = \frac{1}{\frac{1}{\eta} + \frac{\lambda_{1|i}^2}{\nu_1} + \frac{\lambda_{2|i}^2}{\nu_2}}$, the integrals over \mathbf{s} in Equations E.9 & K.2 are analytic.

The expected gradients, of the KL-divergence of the joint belief, in Equations E.9 & E.10 for the cases ($\lambda_1 \leftarrow \delta \lambda_1$), ($\lambda_2 \leftarrow \delta \lambda_2$) and ($\gamma_\tau \leftarrow \delta \gamma_\tau$) are:

- for ($\lambda_1 \leftarrow \delta \lambda_1$),

$$\begin{aligned} \left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta \lambda_1} &= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta \lambda_1}, \quad (\text{E.12}) \\ &= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \end{aligned}$$

$$\begin{aligned}
& \int ds p(\mathbf{s}|l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \frac{\lambda_{1|i}}{\nu_1} (\mathbf{y}_{1,j,n}^T \mathbf{s} - \delta\lambda_{1|i} \mathbf{s}^T \mathbf{s}), \quad (\text{E.13}) \\
&= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \frac{\lambda_{1|i}}{\nu_1} \left(\mathbf{y}_{1,j,n}^T \boldsymbol{\mu}_{s|\tau,j,i,n} - \delta\lambda_{1|i} \left(\boldsymbol{\mu}_{s|\tau,j,i,n}^T \boldsymbol{\mu}_{s|\tau,j,i,n} + D_x \nu_{s|i} \right) \right), \quad (\text{E.14})
\end{aligned}$$

and

$$\begin{aligned}
\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_1 \leftarrow \delta\lambda_1} &= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \left. \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta\lambda_1}, \quad (\text{E.15})
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \int ds p(\mathbf{s}|l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \frac{\lambda_{1|i}^2}{\nu_1} \mathbf{s}^T \mathbf{s}, \quad (\text{E.16})
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \frac{\lambda_{1|i}^2}{\nu_1} \left(\boldsymbol{\mu}_{s|\tau,j,i,n}^T \boldsymbol{\mu}_{s|\tau,j,i,n} + D_x \nu_{s|i} \right), \quad (\text{E.17})
\end{aligned}$$

- for $(\lambda_2 \leftarrow \delta\lambda_2)$,

$$\begin{aligned}
\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_2 \leftarrow \delta\lambda_2} &= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \left. \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_2 \leftarrow \delta\lambda_2}, \quad (\text{E.18})
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \int ds p(\mathbf{s}|l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \frac{\lambda_{2|i}}{\nu_2} \left(\mathbf{y}_{2,j,n}^T \hat{\mathbf{D}}_{\tau} \mathbf{s} - \delta\lambda_{2|i} \mathbf{s}^T \mathbf{s} \right) (\text{E.19})
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \frac{\lambda_{2|i}}{\nu_2} \left(\mathbf{y}_{2,j,n}^T \hat{\mathbf{D}}_{\tau} \boldsymbol{\mu}_{s|\tau,j,i,n} - \delta\lambda_{2|i} \left(\boldsymbol{\mu}_{s|\tau,j,i,n}^T \boldsymbol{\mu}_{s|\tau,j,i,n} + D_x \nu_{s|i} \right) \right) (\text{E.20})
\end{aligned}$$

and

$$\begin{aligned}
\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_2 \leftarrow \delta\lambda_2} &= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \\
& \quad \left. \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_2 \leftarrow \delta\lambda_2}, \quad (\text{E.21})
\end{aligned}$$

$$= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \int ds p(\mathbf{s} | l_i, \tau, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \frac{\lambda_{2|i}^2}{\nu_2} \mathbf{s}^T \mathbf{s}, \quad (\text{E.22})$$

$$= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\lambda_{2|i}^2}{\nu_2} \left(\boldsymbol{\mu}_{\mathbf{s}|\tau,j,i,n}^T \boldsymbol{\mu}_{\mathbf{s}|\tau,j,i,n} + D_x \nu_{\mathbf{s}|i} \right), \quad (\text{E.23})$$

- and for $(\gamma_{\tau} \leftarrow \delta\gamma_{\tau})$,

$$\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\gamma_{\tau} \leftarrow \delta\gamma_{\tau}} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\gamma_{\tau} \leftarrow \delta\gamma_{\tau}}, \quad (\text{E.24})$$

$$= -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\gamma_{\tau|i}}{\omega_{\tau|i}} (\tau - \delta\gamma_{\tau|i}), \quad (\text{E.25})$$

and

$$\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\gamma_{\tau} \leftarrow \delta\gamma_{\tau}} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\gamma_{\tau} \leftarrow \delta\gamma_{\tau}}, \quad (\text{E.26})$$

$$= \frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J p(l_i | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \sum_{\tau} p(\tau | l_i, \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \frac{\gamma_{\tau|i}^2}{\omega_{\tau|i}}. \quad (\text{E.27})$$

I construct the gradients used to calculate the curvatures of the interaural attenuations from a combination of the gradients of both monaural attenuations. These gradients for the interaural attenuations are:

- for $(\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta\lambda_2)$ representing the average loudness,

$$\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta\lambda_2} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta\lambda_1, \lambda_2 \leftarrow \delta\lambda_2}, \quad (\text{E.28})$$

$$= \left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta \lambda_1} + \left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_2 \leftarrow \delta \lambda_2}, \quad (\text{E.29})$$

and

$$\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta \lambda_2} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta \lambda_2}, \quad (\text{E.30})$$

$$= \left. \frac{\partial}{\partial \delta} \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta \lambda_2}, \quad (\text{E.31})$$

$$= \left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_1 \leftarrow \delta \lambda_1} + \left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_2 \leftarrow \delta \lambda_2}, \quad (\text{E.32})$$

- for $(\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2)$ representing average interaural loudness,

$$\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2} \quad (\text{E.33})$$

$$= \left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta \lambda_1} - \frac{1}{\delta^2} \left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_2 \leftarrow \delta \lambda_2}, \quad (\text{E.34})$$

and

$$\left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2} \quad (\text{E.35})$$

$$= \left. \frac{\partial}{\partial \delta} \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta \lambda_1, \lambda_2 \leftarrow \delta^{-1} \lambda_2}, \quad (\text{E.36})$$

$$= \left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_1 \leftarrow \delta \lambda_1} + 2 \frac{1}{\delta^3} \left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_2 \leftarrow \delta \lambda_2} - \frac{1}{\delta^2} \left. \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \right|_{\lambda_2 \leftarrow \delta \lambda_2} \quad (\text{E.37})$$

- and for $(\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2)$ also representing average interaural loudness,

$$\left. \frac{\partial \tilde{d}_j}{\partial \delta} \right|_{\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int ds p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \left. \frac{\partial}{\partial \delta} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \right|_{\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2} \quad (\text{E.38})$$

$$= \frac{\partial \tilde{d}_j}{\partial \delta} \Big|_{\lambda_2 \leftarrow \delta \lambda_2} - \frac{1}{\delta^2} \frac{\partial \tilde{d}_j}{\partial \delta} \Big|_{\lambda_1 \leftarrow \delta \lambda_1}, \quad (\text{E.39})$$

and

$$\frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \Big|_{\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2} = -\frac{1}{N_{\text{val}}} \sum_{n=1}^{N_{\text{val}}} \sum_{i=1}^J \sum_{\tau} \int d\mathbf{s} p(l_i, \mathbf{s}, \tau | \mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, \Omega_a) \times \frac{\partial^2}{\partial \delta^2} \log p(\mathbf{y}_{1,j,n}, \mathbf{y}_{2,j,n}, l_i, \mathbf{s}, \tau | \Omega_b) \Big|_{\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2} \quad (\text{E.40})$$

$$= \frac{\partial}{\partial \delta} \frac{\partial \tilde{d}_j}{\partial \delta} \Big|_{\lambda_1 \leftarrow \delta^{-1} \lambda_1, \lambda_2 \leftarrow \delta \lambda_2}, \quad (\text{E.41})$$

$$= \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \Big|_{\lambda_2 \leftarrow \delta \lambda_2} + 2 \frac{1}{\delta^3} \frac{\partial \tilde{d}_j}{\partial \delta} \Big|_{\lambda_1 \leftarrow \delta \lambda_1} - \frac{1}{\delta^2} \frac{\partial^2 \tilde{d}_j}{\partial \delta^2} \Big|_{\lambda_1 \leftarrow \delta \lambda_1} \quad (\text{E.42})$$

This defines the gradients, used to construct the curvature $C_j(\delta)$, for each perturbation of the curve $c(\delta) = \{\delta, \tilde{d}_j(\delta)\}$.

Appendix F

Levels versus level disparities

The levels and level disparity models are, derived from the prior model in the previous Chapter on the contribution of cues, Equation 2.6 but with the noise processes of equivalent variance $\nu = \nu_1 = \nu_2$,

$$p(x_1, x_2 | \lambda_1, \lambda_2, \tau) = \sqrt{\frac{\nu_s}{2\pi\nu^2\eta}} \exp \left\{ -\frac{1}{2\nu^2} (\nu - \lambda_1^2\nu_s) x_1^2 - \frac{1}{2\nu^2} (\nu - \lambda_2^2\nu_s) x_2^2 + \frac{\lambda_1\lambda_2\nu_s}{\nu^2} x_1\hat{D}_\tau x_2 \right\}, \quad (\text{F.1})$$

where $\frac{1}{\nu_s} = \frac{1}{\eta} + \frac{\lambda_1^2}{\nu} + \frac{\lambda_2^2}{\nu}$. The graphical model in Figure 2.2 is modified to project onto two distinct logarithmic domains Figure F.1. Both projected distributions normalised probability density functions.

F.1 The distribution of levels

For the levels, lv_1, lv_2 , having likelihoods of,

$$p(lv_i | x_i) = \delta(lv_i - \log|x_i|), \quad (\text{F.2})$$

which results in,

$$p(lv_1, lv_2 | \lambda_1, \lambda_2, \nu, \eta) = \sqrt{\frac{\nu_s}{\pi\nu^2\eta}} e^{lv_1+lv_2} \exp \left\{ -\frac{\nu - \lambda_1^2\nu_s}{2\nu^2} e^{2\cdot lv_1} - \frac{\nu - \lambda_2^2\nu_s}{2\nu^2} e^{2\cdot lv_2} \right\} \times \left(\exp \left\{ \frac{\nu_s\lambda_1\lambda_2}{\nu^2} e^{lv_1+lv_2} \right\} + \exp \left\{ -\frac{\nu_s\lambda_1\lambda_2}{\nu^2} e^{lv_1+lv_2} \right\} \right), \quad (\text{F.3})$$

where $\frac{1}{\nu_s} = \frac{1}{\eta} + \frac{\lambda_1^2}{\nu} + \frac{\lambda_2^2}{\nu}$.

For the levels distribution, Equation F.3, in Figure F.2 we see that as we vary the at-

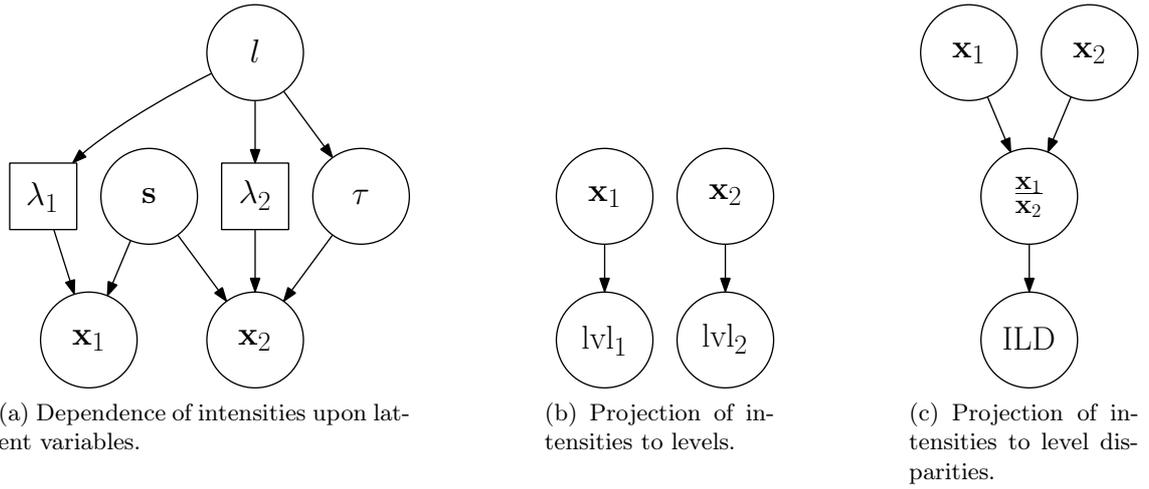


Figure F.1: Graphical models depicting the relationship between the latent variables and the observable variables in a model of sound source localisation. The latent variables of location l , time disparity τ and the sound source \mathbf{s} , and the attenuations λ_1 & λ_2 . The observable variables of sound intensity \mathbf{x}_1 & \mathbf{x}_2 (plot a), levels lvl_1 & lvl_2 (plot b) and level disparity ILD (plot c) in a model of sound source localisation.

attenuation λ_2 , the levels lvl_1 & lvl_2 become more correlated as λ_2 approaches λ_1 . Figure F.3 shows that increasing the signal-to-noise ratio causes the levels lvl_1 & lvl_2 become more correlated.

These two results are unsurprising, and are depicted more clearly using the negentropy (Shannon information) over lvl_1 & lvl_2 .¹ In Figure F.4 we see that the sharpness of the distributions increases with decreasing attenuation, similarly the same is true for increasing signal-to-noise ratios.²

¹In the case of a unimodal distribution negentropy is a direct measure of the distributions sharpness.

²A smaller r as defined in Equation F.3 indicates a greater contribution of the signal to the measured signals.

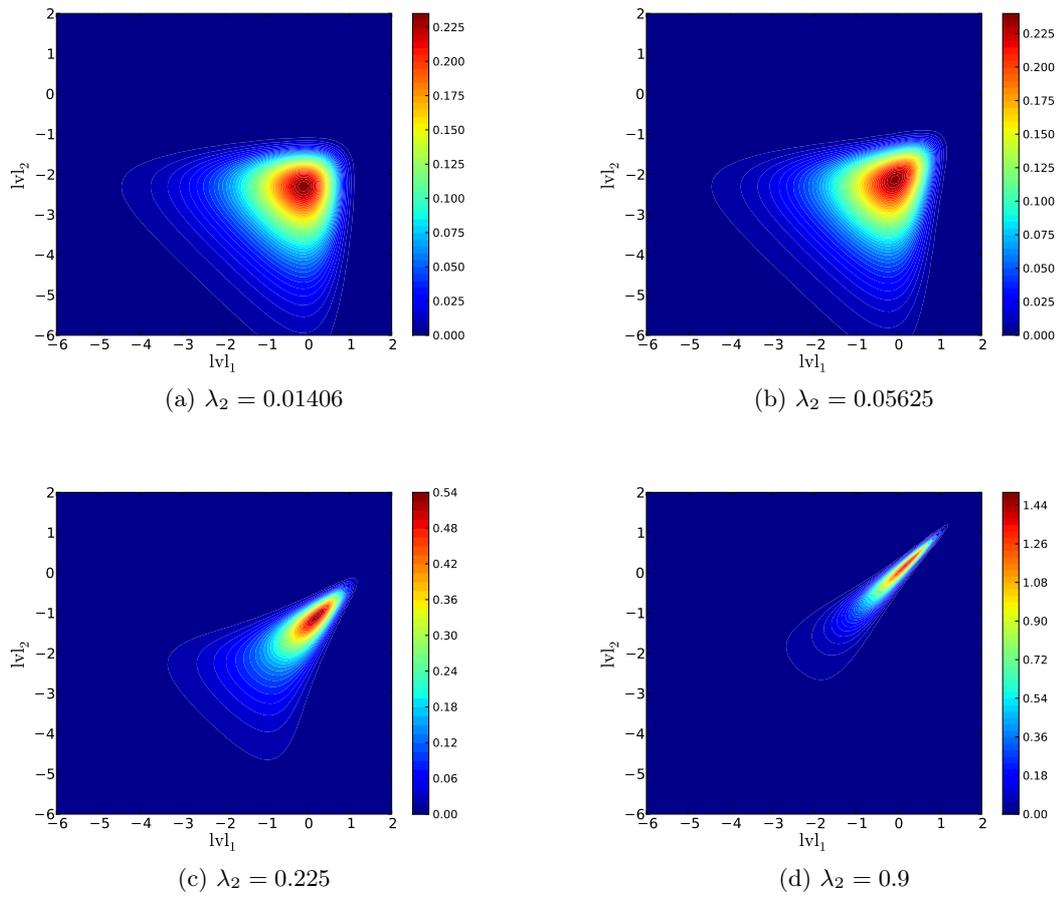


Figure F.2: Plots of the distribution of sound levels with attenuation. The distribution of sound levels $p(lvl_1, lvl_2 | \lambda_1, \lambda_2)$ is depicted for variation of λ_2 , where $\lambda_1 = 0.9$, $\nu = \frac{1}{100}$ and $\eta = 1$. This illustrates the importance of attenuation upon correlation.

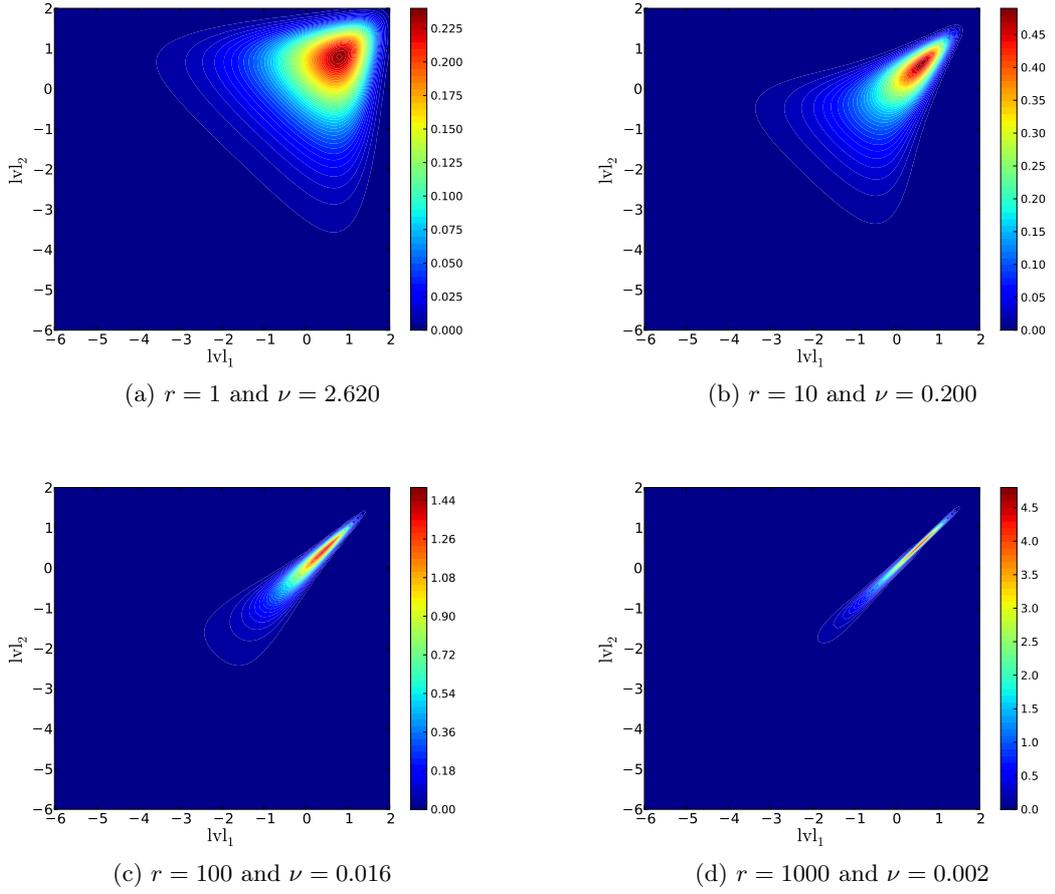


Figure F.3: Plots of the distribution of sound levels with signal-to-noise ratio. The distribution of levels $p(\text{lvl}_1, \text{lvl}_2 | \nu, \eta)$ with variation of the ratio between η and ν so as to keep the expected lvl_1 and lvl_2 constant, where $\lambda_1 = 0.9$ and $\lambda_2 = 0.9$. Indicating the influence of the signal-to-noise ratio upon correlation, a sharper diagonal indicates a greater correlation.

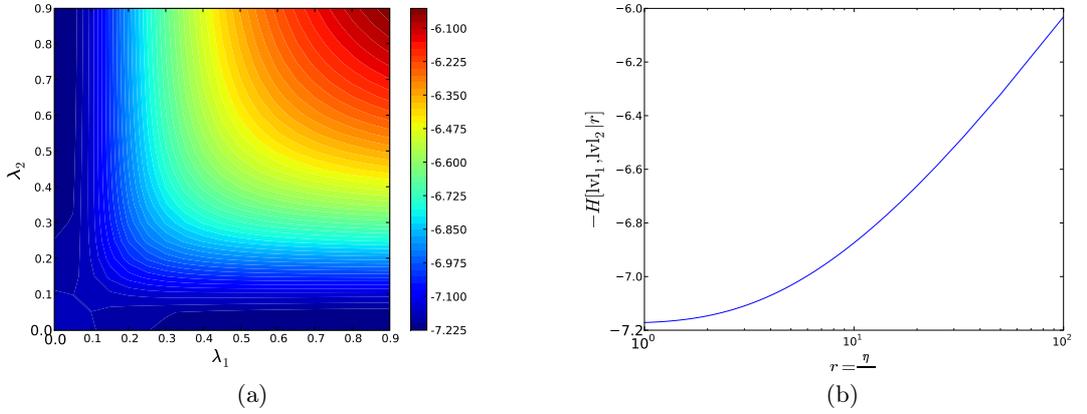


Figure F.4: The negentropy of the distribution of levels for variation of attenuation and signal-to-noise ratio. The negentropy, in (a) $-H[\text{lvl}_1, \text{lvl}_2 | \lambda_1, \lambda_2]$ over λ_1 and λ_2 where $\nu = \frac{1}{100}$ and $\eta = 1$, and in (b) $-H[\text{lvl}_1, \text{lvl}_2 | r]$ depict the influence of the signal-to-noise ratio $r = \frac{\eta}{\nu}$ where $\nu = \frac{1}{100}$, $\lambda_1 = 0.9$ and $\lambda_2 = 0.9$, for the distribution $p(\text{lvl}_1, \text{lvl}_2 | \lambda_1, \lambda_2, \eta, \nu)$.

F.2 The distribution of level disparities

Similarly the level disparity, ILD, has a likelihood,³

$$p(\text{ILD}|x_1, x_2) = \delta \left(\text{ILD} - \log \left| \frac{x_1}{x_2} \right| \right), \quad (\text{F.4})$$

which results in,

$$p(\text{ILD}|\lambda_1, \lambda_2, \nu, \eta) = \frac{1}{\pi} \sqrt{\frac{\nu_s}{\eta}} \left(\frac{1}{e^{\text{ILD}} + e^{-\text{ILD}} - \frac{\nu_s}{\nu} \left(\lambda_1 e^{\frac{\text{ILD}}{2}} + \lambda_2 e^{-\frac{\text{ILD}}{2}} \right)^2} + \frac{1}{e^{\text{ILD}} + e^{-\text{ILD}} - \frac{\nu_s}{\nu} \left(\lambda_1 e^{\frac{\text{ILD}}{2}} - \lambda_2 e^{-\frac{\text{ILD}}{2}} \right)^2} \right), \quad (\text{F.5})$$

with $\frac{1}{\nu_s} = \frac{1}{\eta} + \frac{\lambda_1^2}{\nu} + \frac{\lambda_2^2}{\nu}$.

For the level disparity distribution, Figure F.3 shows that increasing the signal-to-noise ratio causes the levels lvl_1 & lvl_2 become more correlated. The attenuation dependent term can be seen in the ν_s as the switch deciding correlation as $\frac{1}{\nu_s}$ increases then so does the correlation.

These two results are unsurprising, and are depicted more clearly using the negentropy (Shannon information) over lvl_1 & lvl_2 , in the case of a unimodal distribution it is a direct measure of the distributions sharpness. In Figure F.4 we see that the sharpness of the distributions increases with decreasing attenuation, similarly the same is true for increasing signal-to-noise ratios.

³The derivation was approached in parts, first by computing the ratio $p(\text{ratio}|x_1, x_2) = \delta \left(\text{ratio} - \left| \frac{x_1}{x_2} \right| \right)$ to give the distribution of $p(\text{ratio}|\lambda_1, \lambda_2, \nu, \eta)$, and then $p(\text{ILD}|\text{ratio}) = \delta(\text{ILD} - \log \text{ratio})$.

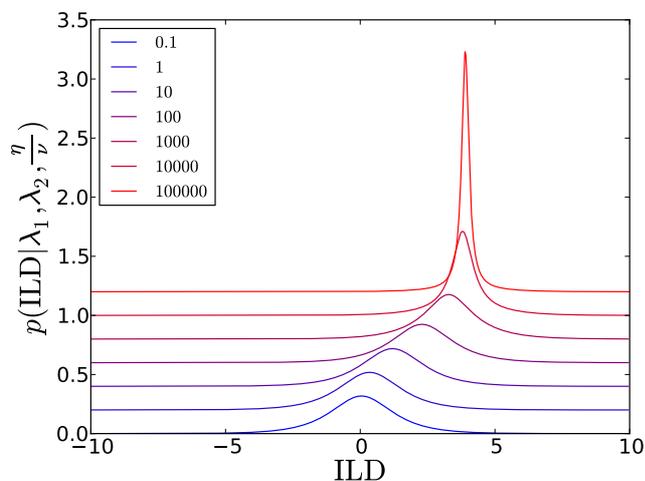


Figure F.5: Plots of the distribution of ILD for variation of signal-to-noise ratio. The distribution of interaural level disparities (ILD) $p(\text{ILD}|r)$ with variation of $r = \frac{\eta}{\nu}$ (indicated in the legend), for $\lambda_1 = 0.9$ and $\lambda_2 = \frac{1}{50}\lambda_1$. Indicating the influence of the signal-to-noise ratio upon the distribution, a sharper peak suggests a greater correlation.

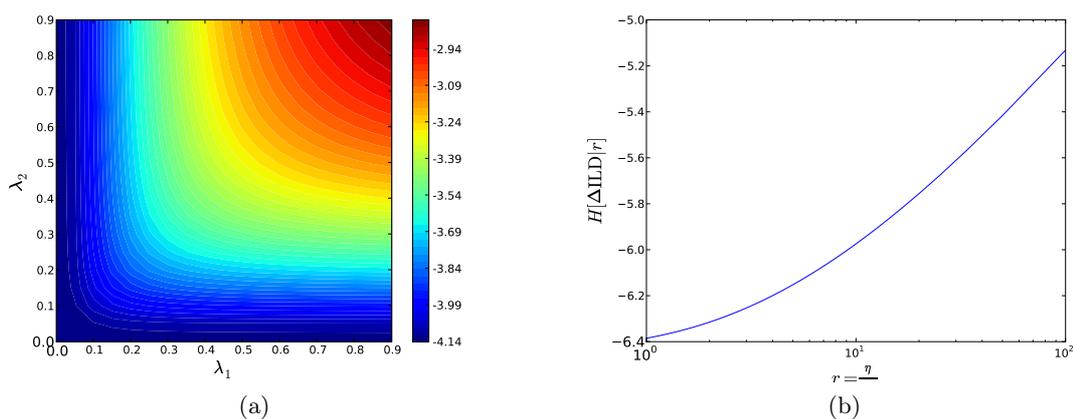


Figure F.6: The negentropy of the ILD distributions for variation of attenuation and signal-to-noise ratio. The negentropy, in (a) $-H[\text{ILD}|\lambda_1, \lambda_2]$ over λ_1 and λ_2 where $\nu = \frac{1}{100}$ and $\eta = 1$, and in (b) $-H[\text{ILD}|r]$ depicting the influence of the signal-to-noise ratio $r = \frac{\eta}{\nu}$ where $\nu = \frac{1}{100}$, $\lambda_1 = 0.9$ and $\lambda_2 = 0.9$, for the distribution $p(\text{ILD}|\lambda_1, \lambda_2, \eta, \nu)$.

Appendix G

Jensen's inequality

Jensen's inequality (Cover and Thomas, 2006): if f is a concave function, g is some function of a random variable x ,

$$\int dx p(x) f(g(x)) \leq f\left(\int dx p(x) g(x)\right), \quad (\text{G.1})$$

where $p(x)$ is a normalised probability distribution. The logarithm is such a concave function.

Appendix H

Optimal statistical filtering

The purpose of this section is to concretely define the model's structure and derive the conditional Shannon information for such a system.

Bayes law, for a prior belief A , $p(A)$, and a likelihood of B given knowledge of A , $p(B|A)$, states that the posterior belief in A given B is defined according to the relation,

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}, \quad (\text{H.1})$$

where $p(B)$ is the marginal of the product of the prior $p(A)$ and the likelihood $p(B|A)$,

$$p(B) = \int dA p(B|A) p(A). \quad (\text{H.2})$$

These are the principle tools for constructing the time evolution updates in belief or Markov relations.

The Markov relations are defined by two stages. Working from a prior belief $p(l_{t-1})$ and an arbitrary dynamical process characterised by, $p(l_t|l_{t-1}, a_t)$, a prediction in the latent space can be made,

$$p(l_t|a_t) \triangleq \int dl_{t-1} p(l_t|l_{t-1}, a_t) p(l_{t-1}). \quad (\text{H.3})$$

Next, working from this prediction and conditioning upon a measurement, x_t , using Bayes law the posterior belief in the latent space $l_t \in L_t$ given a measurement x_t and action a_t ,

$$p(l_t|x_t, a_t) \triangleq \frac{p(x_t|l_t, a_t) p(l_t|a_t)}{p(x_t|a_t)}, \quad (\text{H.4})$$

where,

$$p(x_t|a_t) \triangleq \int dl_t p(x_t|l_t, a_t) p(l_t|a_t). \quad (\text{H.5})$$

This defines the Markov relations and as such are entirely general.

Appendix I

Information identities

Definition I.1. The conditional Shannon information is the expectation of the posterior Shannon information,

$$I[\mathbf{L}_t|\mathbf{X}_t, a_t] = \mathbb{E}_{p(x_t|a_t)} [I[\mathbf{L}_t|x_t, a_t]]. \quad (\text{I.1})$$

Proof. That,

$$\mathbb{E}_{p(x_t|a_t)} [I[\mathbf{L}_t|x_t, a_t]] = \int dx_t p(x_t|a_t) \int dl_t p(l_t|x_t, a_t) \log p(l_t|x_t, a_t), \quad (\text{I.2})$$

$$= \int dx_t dl_t p(x_t, l_t|a_t) \log p(l_t|x_t, a_t), \quad (\text{I.3})$$

completing the proof as, $I[\mathbf{L}_t|\mathbf{X}_t, a_t] \triangleq \int dx_t dl_t p(x_t, l_t|a_t) \log p(l_t|x_t, a_t)$. □

Theorem I.2. Bayes law for information, making

$$I[\mathbf{L}_t|\mathbf{X}_t, a_t] = I[\mathbf{L}_t|a_t] + I[\mathbf{X}_t|\mathbf{L}_t, a_t] - I[\mathbf{X}_t|a_t]. \quad (\text{I.4})$$

Proof. Follows from Bayes law,

$$p(l_t|x_t, a_t) = \frac{p(x_t|l_t, a_t) p(l_t|a_t)}{p(x_t|a_t)}, \quad (\text{I.5})$$

where, by definition

$$I[\mathbf{L}_t|\mathbf{X}_t, a_t] = \int dx_t dl_t p(x_t, l_t|a_t) \log p(l_t|x_t, a_t), \quad (\text{I.6})$$

$$= \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t|l_t, a_t) p(l_t|a_t)}{p(x_t|a_t)}, \quad (\text{I.7})$$

$$= \int dl_t p(l_t|a_t) \log p(l_t|a_t) + \int dx_t dl_t p(x_t, l_t|a_t) \log p(x_t|l_t, a_t)$$

$$- \int dx_t p(x_t|a_t) \log p(x_t|a_t), \quad (\text{I.8})$$

$$\equiv I[\mathbf{L}_t|a_t] + I[\mathbf{X}_t|\mathbf{L}_t, a_t] - I[\mathbf{X}_t|a_t], \quad (\text{I.9})$$

completing the proof. \square

Definition I.3. The mutual information is the difference between two information terms,

$$I_{\text{MI}}[\mathbf{X}_t; \mathbf{L}_t|a_t] = I[\mathbf{X}_t|\mathbf{L}_t, a_t] - I[\mathbf{X}_t|a_t], \quad (\text{I.10})$$

$$\equiv I[\mathbf{L}_t|\mathbf{X}_t, a_t] - I[\mathbf{L}_t|a_t]. \quad (\text{I.11})$$

Proof. Starting with

$$I_{\text{MI}}[\mathbf{X}_t; \mathbf{L}_t|a_t] = \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t, l_t|a_t)}{p(x_t|a_t)p(l_t|a_t)}, \quad (\text{I.12})$$

$$= \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t|l_t, a_t)}{p(x_t|a_t)}, \quad (\text{I.13})$$

$$= \int dx_t dl_t p(x_t, l_t|a_t) \log p(x_t|l_t, a_t) - \int dx_t p(x_t|a_t) \log p(x_t|a_t), \quad (\text{I.14})$$

$$= I[\mathbf{X}_t|\mathbf{L}_t, a_t] - I[\mathbf{X}_t|a_t], \quad (\text{I.15})$$

proving the first equality of the definition. Next,

$$I[\mathbf{X}_t|\mathbf{L}_t, a_t] - I[\mathbf{X}_t|a_t] = \int dx_t dl_t p(x_t, l_t|a_t) \log p(x_t|l_t, a_t) - \int dx_t p(x_t|a_t) \log p(x_t|a_t), \quad (\text{I.16})$$

$$= \int dx_t dl_t p(x_t, l_t|a_t) \log \frac{p(x_t|a_t)p(l_t|x_t, a_t)}{p(l_t|a_t)} - \int dx_t p(x_t|a_t) \log p(x_t|a_t), \quad (\text{I.17})$$

$$= \int dx_t p(x_t|a_t) \log p(x_t|a_t) + \int dx_t dl_t p(x_t, l_t|a_t) \log p(l_t|x_t, a_t) - \int dl_t p(l_t|a_t) \log p(l_t|a_t) - \int dx_t p(x_t|a_t) \log p(x_t|a_t), \quad (\text{I.18})$$

$$= I[\mathbf{X}_t|a_t] + I[\mathbf{L}_t|\mathbf{X}_t, a_t] - I[\mathbf{L}_t|a_t] - I[\mathbf{X}_t|a_t], \quad (\text{I.19})$$

completing the proof. \square

Theorem I.4. *That*

$$I_{\text{MI}} [X_t; L_t | a_t] = \mathbb{E}_{p(x_t | a_t)} [D_{\text{KL}} [p(L_t | x_t, a_t) || p(L_t | a_t)]] . \quad (\text{I.20})$$

Proof. Starting with

$$\begin{aligned} \mathbb{E}_{p(x_t | a_t)} [D_{\text{KL}} [p(L_t | x_t, a_t) || p(L_t | a_t)]] &= \int dx_t p(x_t | a_t) \int dl_t p(l_t | x_t, a_t) \times \\ &\quad \log \frac{p(l_t | x_t, a_t)}{p(l_t | a_t)}, \end{aligned} \quad (\text{I.21})$$

$$\begin{aligned} &= \int dx_t dl_t p(x_t, l_t | a_t) \log p(l_t | x_t, a_t) \\ &\quad - \int dl_t p(l_t | a_t) \log p(l_t | a_t), \end{aligned} \quad (\text{I.22})$$

$$= I[L_t | X_t, a_t] - I[L_t | a_t]. \quad (\text{I.23})$$

□

Corollary I.5. $I_{\text{MI}} [X_t; L_t | a_t] = \mathbb{E}_{p(l_t | a_t)} [D_{\text{KL}} [p(X_t | l_t, a_t) || p(X_t | a_t)]]$ which follows by simply switching x_t and l_t for each other.

Appendix J

My measure for a sequence

Theorem J.1. *For a sequence pair of sequences $u_{n_1:n_2}$ and $v_{n_1:n_2}$, which are related by the conditional distribution $p(u_{n_1:n_2}|v_{n_1:n_2})$. Also subject to the constraint that $n_1 < n_2$. Then the following identity holds*

$$\int du_{n_1:n_2} p(u_{n_1:n_2}|v_{n_1:n_2}) \sum_{n'=n_1}^{n_2} f_{n'}(u_{n'}) = \sum_{n'=n_1}^{n_2} \int du_{n'} p(u_{n'}|v_{n_1:n_2}) f_{n'}(u_{n'}). \quad (\text{J.1})$$

Proof. As

$$\begin{aligned} \int du_{n_1:n_2} p(u_{n_1:n_2}|v_{n_1:n_2}) \sum_{n'=n_1}^{n_2} f_{n'}(u_{n'}) &= \int du_{n_1+1:n_2} p(u_{n_1+1:n_2}|v_{n_1:n_2}) \times \\ &\quad \sum_{n'=n_1+1}^{n_2} f_{n'}(u_{n'}) + \\ &\quad \int du_{n_1} p(u_{n_1}|v_{n_1:n_2}) f_{n_1}(u_{n_1}), \end{aligned} \quad (\text{J.2})$$

it follows that

$$\int du_{n_1:n_2} p(u_{n_1:n_2}|v_{n_1:n_2}) \sum_{n'=n_1}^{n_2} f_{n'}(u_{n'}) = \sum_{n'=n_1}^{n_2} \int du_{n'} p(u_{n'}|v_{n_1:n_2}) f_{n'}(u_{n'}). \quad (\text{J.3})$$

Hence proving the theorem. □

Corollary J.2. *If we replace n with t , u_n with x_t , v_t with a_t and l_t , and make the sum over the quantity $\log p(x_t|l'_t, a_t)$ then the identity in Theorem J.1 leads to the identity*

$$\int dx_{t:t+\delta t} p(x_{t:t+\delta t}|l_{t:t+\delta t}, a_{t:t+\delta t}) \sum_{t'=t}^{t+\delta t} \log p(x_{t'}|l'_{t'}, a_{t'}) = \sum_{t'=t}^{t+\delta t} \int dx_{t'} p(x_{t'}|l_{t:t+\delta t}, a_{t:t+\delta t}) \log p(x_{t'}|l'_{t'}, a_{t'}) \quad (\text{J.4})$$

If we set $l'_t = l_t$ this gives the identity,

$$D_{\text{KL}} [p(X_{t:t+\delta t} | l_{t:t+\delta t}, a_{t:t+\delta t}) \| p(X_{t:t+\delta t} | l'_{t:t+\delta t}, a_{t:t+\delta t})] = \sum_{t'=t}^{t+\delta t} D_{\text{KL}} [p(X_{t'} | l_{t'}, a_{t'}) \| p(X_{t'} | l'_{t'}, a_{t'})]. \quad (\text{J.5})$$

Theorem J.3. For a sequence of actions $a_{t:t+\delta t}$ over the time interval t to $t + \delta t$, the space of measurements $X_{t:t+\delta t}$ and for the latent hypotheses, $l_{t:t+\delta t}$ and $l'_{t:t+\delta t}$ gives the identity

$$B_{X_{t:t+\delta t} | a_{t:t+\delta t}} [L_{t:t+\delta t} \| L'_{t:t+\delta t}] = \sum_{t'=t}^{t+\delta t} \int dl_{t'} dl'_{t'} p(l_{t'} | a_{t:t'}) p(l'_{t'} | a_{t:t'}) \times D_{\text{KL}} [p(X_{t'} | l_{t'}, a_{t'}) \| p(X_{t'} | l'_{t'}, a_{t'})], \quad (\text{J.6})$$

$$\equiv \sum_{t'=t}^{t+\delta t} B_{X_{t'} | a_{t:t'}} [L_{t'} \| L'_{t'}]. \quad (\text{J.7})$$

Proof. According to the definition

$$B_{X_{t:t+\delta t} | a_{t:t+\delta t}} [L_{t:t+\delta t} \| L'_{t:t+\delta t}] = \int dl_{t:t+\delta t} dl'_{t:t+\delta t} p(l_{t:t+\delta t} | a_{t:t+\delta t}) p(l'_{t:t+\delta t} | a_{t:t+\delta t}) \times D_{\text{KL}} [p(X_{t:t+\delta t} | l_{t:t+\delta t}, a_{t:t+\delta t}) \| p(X_{t:t+\delta t} | l'_{t:t+\delta t}, a_{t:t+\delta t})].$$

Due to each measurement x_t being conditionally independent to all other measurements, the KL-divergence becomes a summation

$$D_{\text{KL}} [p(X_{t:t+\delta t} | l_{t:t+\delta t}, a_{t:t+\delta t}) \| p(X_{t:t+\delta t} | l'_{t:t+\delta t}, a_{t:t+\delta t})] = \sum_{t'=t}^{t+\delta t} D_{\text{KL}} [p(X_{t'} | l_{t'}, a_{t'}) \| p(X_{t'} | l'_{t'}, a_{t'})]. \quad (\text{J.8})$$

Putting this together with Corollary J.2 equates to

$$B_{X_{t:t+\delta t} | a_{t:t+\delta t}} [L_{t:t+\delta t} \| L'_{t:t+\delta t}] = \int dl_{t:t+\delta t} dl'_{t:t+\delta t} p(l_{t:t+\delta t} | a_{t:t+\delta t}) p(l'_{t:t+\delta t} | a_{t:t+\delta t}) \times \sum_{t'=t}^{t+\delta t} D_{\text{KL}} [p(X_{t'} | l_{t'}, a_{t'}) \| p(X_{t'} | l'_{t'}, a_{t'})], \quad (\text{J.9})$$

$$= \sum_{t'=t}^{t+\delta t} \int dl_{t'} dl'_{t'} p(l_{t'} | a_{t:t+\delta t}) p(l'_{t'} | a_{t:t+\delta t}) \times D_{\text{KL}} [p(X_{t'} | l_{t'}, a_{t'}) \| p(X_{t'} | l'_{t'}, a_{t'})]. \quad (\text{J.10})$$

As this is a Markov problem the prediction beliefs $p(l_t | a_{0:T}) \equiv p(l_t | a_{0:T})$ according to the principle that the current state l_t being independent of future states $l_{t+1:T}$. Finally using

the notation that

$$B_{\mathbf{X}_t|a_{0:t}} [\mathbf{L}_t|\mathbf{L}'_t] = \int dl_t dl'_t p(l_t|a_{0:t}) p(l'_t|a_{0:t}) D_{\text{KL}} [p(\mathbf{X}_t|l_t, a_t) \| p(\mathbf{X}_t|l'_t, a_t)], \quad (\text{J.11})$$

makes

$$B_{\mathbf{X}_{t:t+\delta t}|a_{t:t+\delta t}} [\mathbf{L}_{t:t+\delta t}|\mathbf{L}'_{t:t+\delta t}] = \sum_{t'=t}^{t+\delta t} \int dl_{t'} dl'_{t'} p(l_{t'}|a_{t:t'}) p(l'_{t'}|a_{t:t'}) \times \\ D_{\text{KL}} [p(\mathbf{X}_{t'}|l_{t'}, a_{t'}) \| p(\mathbf{X}_{t'}|l'_{t'}, a_{t'})], \quad (\text{J.12})$$

$$= \sum_{t'=t}^{t+\delta t} B_{\mathbf{X}_{t'}|a_{t:t'}} [\mathbf{L}_{t'}|\mathbf{L}'_{t'}]. \quad (\text{J.13})$$

This proves the theorem. □

Appendix K

POMDP informatic action policy comparisons

Starting with a POMDP where the latent (l_t) and measurement (x_t) variables are both discrete. The calculation for the Shannon information is,

$$I[\mathbf{L}_t | \mathbf{X}_t, a_t] = \sum_{l_t, x_t} p(l_t, x_t | a_t) \log p(l_t | x_t, a_t), \quad (\text{K.1})$$

$$\begin{aligned} &= \sum_{l_t} p(l_t | a_t) \log p(l_t | a_t) + \sum_{l_t} p(l_t | a_t) \sum_{x_t} p(x_t | l_t, a_t) \log p(x_t | l_t, a_t) - \\ &\quad \sum_{x_t} p(x_t | a_t) \log p(x_t | a_t), \end{aligned} \quad (\text{K.2})$$

which can be calculated algorithmically using Algorithm K.1. In contrast my measure is calculated according to,

$$B_{\mathbf{X}_t | a_t}[\mathbf{L}_t || \mathbf{L}'_t] = \sum_{l_t, l'_t} p(l_t | a_t) p(l'_t | a_t) \sum_{x_t} p(x_t | l_t, a_t) \log \frac{p(x_t | l_t, a_t)}{p(x_t | l'_t, a_t)}, \quad (\text{K.3})$$

$$\begin{aligned} &= \sum_{l_t} p(l_t | a_t) \sum_{x_t} p(x_t | l_t, a_t) \log p(x_t | l_t, a_t) - \\ &\quad \sum_{x_t} \left[\sum_{l_t} p(l_t | a_t) p(x_t | l_t, a_t) \right] \times \\ &\quad \left[\sum_{l_t} p(l_t | a_t) \log p(x_t | l_t, a_t) \right], \end{aligned} \quad (\text{K.4})$$

which can be calculated in Algorithm K.2. Both algorithms have a time complexity of $O(N_x N_l N_a)$; x_t can take one of N_x possible states, l_t can take one of N_l possible states, and there are N_a possible decision choices. In Fox *et al.* (1998) the authors reduced the

Algorithm K.1 An Infomax algorithm for a POMDP. The observable variables are collected together as x_t , the hidden variables are collected as l_t , and the decision variables are collected as a_t . This scales as $O(N_x N_l N_a)$; x_t can take one of N_x possible states, l_t can take one of N_l possible states, and there are N_a possible decision choices.

Calculation of Shannon information,

1. for each a_t ,
 - $h_1 = 0, h_2 = 0$
 - (a) for each x_t in X_t
 - $h_3 = 0$
 - i. for each l_t in L_t
 - $g = p(x_t|l_t, a_t) p(l_t|a_t)$
 - $h_1 = h_1 + g \log g$
 - $h_3 = h_3 + g$
 - $h_2 = h_2 + h_3 \log h_3$
 - $I[L_t|X_t, a_t] = h_1 - h_2$
 2. select the a_t that maximises $I[L_t|X_t, a_t]$.
-

Algorithm K.2 An implementation of my information measure (Equation 3.3) for a POMDP. The observable variables are collected together as x_t , the hidden variables are collected as l_t , and the decision variables are collected as a_t . This scales as $O(N_x N_l N_a)$; x_t can take one of N_x possible states, l_t can take one of N_l possible states, and there are N_a possible decision choices.

Calculation of my method,

1. for each a_t ,
 - $h_1 = 0, h_2 = 0$
 - (a) for each x_t in X_t
 - $h_3 = 0, h_4 = 0$
 - i. for each l_t in L_t
 - $g_0 = p(l_t|a_t)$
 - $g_1 = p(x_t|l_t, a_t) g_0$
 - $h_1 = h_1 + g_0 g_1 \log g_1$
 - $h_3 = h_3 + g_0 \log g_1$
 - $h_4 = h_4 + g_1$
 - $h_2 = h_4 h_3$
 - $B_{X_t|a_t}[L_t||L'_t] = h_1 - h_2$
 2. select the a_t that maximises $B_{X_t|a_t}[L_t||L'_t]$.
-

number of latent states to consider defined by $p(l_t|a_t) > \varepsilon$ where ε is small but greater than zero. This reduces the time complexity for both Algorithms K.1 & K.2 by reducing N_a . For POMDP both algorithms scale equivalently: for other problems my measure can be defined to scale more favourably.

Appendix L

Gaussian notation

A Gaussian is defined according to its first two moments; the mean $\boldsymbol{\mu}$ and variance σ . A Gaussian is defined for a space \mathbf{x} as

$$\phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma) = \sqrt{\frac{1}{(2\pi)^{D_x} |\Sigma|}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}, \quad (\text{L.1})$$

with a mean $\boldsymbol{\mu}$ and the covariance defined by Σ . If the covariance has the form $\Sigma = I_{D_x} \sigma$ then σ represents the variance of the distribution.

Appendix M

Bayes law

Bayes law, for a prior belief A , $p(A)$, and a likelihood of B given knowledge of A , $p(B|A)$, Bayes law states that the posterior belief in A given B is defined according to the relation,

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}, \quad (\text{M.1})$$

where $p(B)$ is the marginal of the product of the prior $p(A)$ and the likelihood $p(B|A)$,

$$p(B) = \int dA p(B|A)p(A). \quad (\text{M.2})$$

These are the principle tools for constructing the time evolution updates for Markov tracking.

Appendix N

Gaussian identities

Important relations are from the Matrix Cookbook (Petersen and Pedersen, 2008).

For a normally distributed random variable \mathbf{y} with mean $\boldsymbol{\mu}$ and covariance Σ

$$\phi(\mathbf{y}; \boldsymbol{\mu}, \Sigma) = \sqrt{\frac{1}{(2\pi)^{D_y} |\Sigma|}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\}. \quad (\text{N.1})$$

The following identities exist

$$\phi(\mathbf{x}; A\mathbf{y}, \Sigma) = \kappa(\mathbf{x}) \phi\left(\mathbf{y}; (A^T \Sigma A)^{-1} A^T \Sigma \mathbf{x}, (A^T \Sigma A)^{-1}\right), \quad (\text{N.2})$$

$$\phi(\mathbf{y}; \mathbf{a}, A) \phi(\mathbf{y}; \mathbf{b}, B) = \phi\left(\mathbf{y}; (A + B)^{-1} (A\mathbf{a} + B\mathbf{b}), (A + B)^{-1}\right). \quad (\text{N.3})$$

Also if $\mathbf{z} = \{\mathbf{x}, \mathbf{y}\}$, $\boldsymbol{\mu} = \{\boldsymbol{\mu}_x, \boldsymbol{\mu}_y\}$ and, $\Sigma = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$, then marginalising out \mathbf{y} gives

$$\mathbf{x} \sim \int d\mathbf{y} \phi(\mathbf{z}; \boldsymbol{\mu}, \Sigma) = \phi(\mathbf{x}; \boldsymbol{\mu}_x, \Sigma_{xx}), \quad (\text{N.4})$$

and conditioning \mathbf{x} upon \mathbf{y} gives

$$\mathbf{x}|\mathbf{y} \sim \frac{\phi(\mathbf{z}; \boldsymbol{\mu}, \Sigma)}{\phi(\mathbf{y}; \boldsymbol{\mu}_y, \Sigma_{yy})} = \phi\left(\mathbf{x}; \boldsymbol{\mu}_x - \Sigma_{xy} \Sigma_{yy}^{-1} (\boldsymbol{\mu}_y - \mathbf{y}), \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx}\right), \quad (\text{N.5})$$

as a consequence

$$\mathbb{E}_{p(\mathbf{x}|\mathbf{y}=\boldsymbol{\mu}_y)}[\mathbf{x}] = \mathbb{E}[\mathbf{x}]. \quad (\text{N.6})$$

Appendix O

Matrix identities

Important relations are from the Matrix Cookbook (Petersen and Pedersen, 2008). For a $d \times d$ matrix $P > 0$, a $k \times k$ matrix $R > 0$ and a $k \times d$ matrix B where $P > 0$ implies that $\mathbf{a}^T P \mathbf{a} > 0 \forall \mathbf{a}$ e.g. the positive eigenvectors. The following equalities hold,

$$(P^{-1} + B^T R^{-1} B)^{-1} = P - P B^T (B P B^T + R)^{-1} B P, \quad (\text{O.1})$$

$$(P^{-1} + B^T R^{-1} B)^{-1} B^T R^{-1} = P B^T (B P B^T + R)^{-1}. \quad (\text{O.2})$$

Appendix P

Trace equivalence

It is possible to show that

$$\mathbf{x}^T \mathbf{A} \mathbf{z} = \text{tr} [\mathbf{A} \mathbf{z} \mathbf{x}^T]. \quad (\text{P.1})$$

First by defining $\mathbf{y} = \mathbf{A} \mathbf{z}$, show that

$$\mathbf{x}^T \mathbf{y} = \text{tr} [\mathbf{y} \mathbf{x}^T]. \quad (\text{P.2})$$

The vectors are be defined as $\mathbf{y} = \{y_i : i = 1, \dots, N\}$ and $\mathbf{x} = \{x_i : i = 1, \dots, N\}$, which makes

$$x^T y \triangleq \sum_{i=1}^N x_i y_i. \quad (\text{P.3})$$

Defining $\mathbf{y} \mathbf{x}^T \triangleq M$, and given that

$$M \triangleq \begin{bmatrix} m_{11} & \cdots & m_{1N} \\ \vdots & \ddots & \vdots \\ m_{N1} & \cdots & m_{NN} \end{bmatrix}, \quad (\text{P.4})$$

then $m_{ij} = y_i x_j \forall i, j$, and as

$$\text{tr} [M] \triangleq \sum_{i=1}^N m_{ii}, \quad (\text{P.5})$$

then

$$\text{tr} [\mathbf{y} \mathbf{x}^T] = \sum_{i=1}^N y_i x_i, \quad (\text{P.6})$$

$$\equiv \mathbf{x}^T \mathbf{y}. \quad (\text{P.7})$$

Thus as it can be said that

$$\mathbf{x}^T \mathbf{y} = \text{tr} [\mathbf{y}\mathbf{x}^T], \quad (\text{P.8})$$

then as a corollary

$$\mathbf{x}^T \mathbf{A}\mathbf{z} = \text{tr} [\mathbf{A}\mathbf{z}\mathbf{x}^T]. \quad (\text{P.9})$$

What about linear operations, so $F(\mathbf{x})$ where $F(\mathbf{x} + \mathbf{y}) = F(\mathbf{x}) + F(\mathbf{y})$. With the behaviour

$$F(\mathbf{x}) \triangleq \begin{bmatrix} F(x_1) \\ \vdots \\ F(x_N) \end{bmatrix}, \quad (\text{P.10})$$

and

$$F(M) \triangleq \begin{bmatrix} F(m_{11}) & \cdots & F(m_{1N}) \\ \vdots & \ddots & \vdots \\ F(m_{N1}) & \cdots & F(m_{NN}) \end{bmatrix}. \quad (\text{P.11})$$

Then hence due to F being a linear operation the following are true

$$F(\mathbf{x}^T \mathbf{y}) = F\left(\sum_{i=1}^N x_i y_i\right), \quad (\text{P.12})$$

$$= \sum_{i=1}^N F(x_i y_i), \quad (\text{P.13})$$

and,

$$F(\text{tr} [\mathbf{y}\mathbf{x}^T]) = F\left(\sum_{i=1}^N y_i x_i\right), \quad (\text{P.14})$$

$$= \sum_{i=1}^N F(y_i x_i), \quad (\text{P.15})$$

thus the following identity is true,

$$F(\mathbf{x}^T \mathbf{y}) = \text{tr} [F(\mathbf{y}\mathbf{x}^T)]. \quad (\text{P.16})$$

Appendix Q

Sampling from the Lorenz attractor

The seminal discovery of Lorenz that originated the field of chaos was originally motivated by the problem of weather prediction. Lorenz investigated a series of equations derived from thermal convection in the atmosphere (Ott, 2002). His demonstration that thermally driven convection could result in chaos raised the possibility that the atmosphere is chaotic.

The Lorenz attractor is a series of ordinary differential equations

$$\frac{\partial x^{(1)}}{\partial t} = a(x^{(2)} - x^{(1)}), \quad (\text{Q.1})$$

$$\frac{\partial x^{(2)}}{\partial t} = (b - x^{(3)})x^{(1)} - x^{(2)}, \quad (\text{Q.2})$$

$$\frac{\partial x^{(3)}}{\partial t} = x^{(1)}x^{(2)} - cx^{(3)}, \quad (\text{Q.3})$$

where the state vector $\mathbf{x} = \{x^{(i)} : i = 1, 2, 3\}$. See Figure Q.1 for an example attractor for the parameters $a = 10$, $b = 25$ and $c = \frac{8}{3}$ implemented using the Python scientific package Scipy.

To use the Scipy ODE package I constructed a Python function returning the vector of derivatives

```
def foo(t,x):  
  
    A = 10.  
    B = 25.  
    C = 8./3.  
    dx = zeros([3])  
    dx[0] = -A*x[0]+A*x[1]  
    dx[1] = -x[0]*x[2]+B*x[0]-x[1]  
    dx[2] = x[0]*x[1]-C*x[2]  
    return dx
```

Then a function to sample the Lorenz attractor for this set of parameters ($a = 10$, $b =$

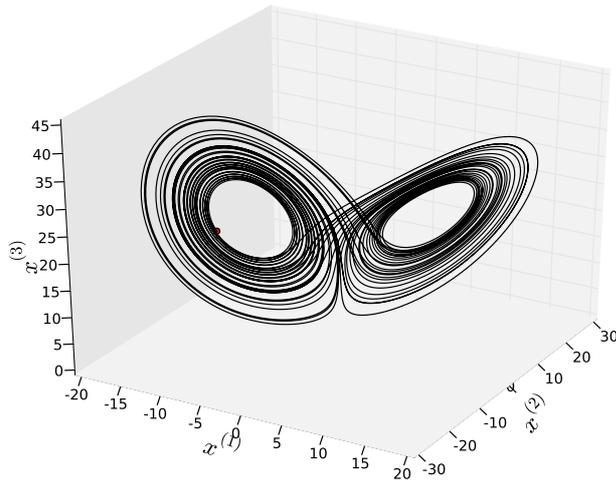


Figure Q.1: The Lorenz attractor. The attractor with the parameters $a = 10$, $b = 25$ and $c = \frac{8}{3}$. The starting point is indicated by a red dot. Sampling this trajectory was implemented using the Python scientific package Scipy (using the included ode solver).

25 and $c = \frac{8}{3}$). This makes use of the variable-coefficient ordinary differential (VODE) equation solver. This is wrapped in Python using the function

```
def sample_lorenz(x0,t0,dt,tEnd=50.):

    x = []
    t = []
    index = 0
    r = scipy.integrate.ode(foo).set_integrator('vode').set_initial_value(x0,t0)
    while r.successful() and r.t < tEnd:

        r.integrate(r.t+dt)
        x.append(r.y)
        t.append(r.t)
        index += 1

    return array(x,dtype='float64'),array(t,dtype='float64')
```

Finally, the usage of these functions to construct a Lorenz sample is

```
x0 = [0.,1.,1.]
dt = 0.05
tEnd = 50.
t0 = 0.
x,t = sample_lorenz(x0,t0,dt,tEnd)
```

This is from the start point $\mathbf{x}(t_0) = \{0, 1, 1\}$, for 50 time steps of length $\delta t = 0.05$.

Appendix R

Episodic learning

R.1 Learning requirements

The purpose is to learn the parameters of a system which has a hidden latent space that can be constructed from the sequence of measurements $x_{0:\tau}$ that comprises the data set Dat . The data set is generated sequentially according to a set of actions $a_{0:\tau}$. As such it is dependent completely upon the parameter Ω due to the recursive selection of each action a_t using a measure of goodness based upon $I [L_t | X_t, x_{0:t-1}, a_{0:t}, \Omega]$, to then generate a corresponding measurement x_t .

To learn a new parameterisation Ω^{new} from this data set and action set I use maximum likelihood (ML), hence applying argument maximisation to a models log-likelihood,

$$\Omega^{\text{new}} = \arg \max_{\Omega} \mathcal{L} [\Omega; a_{0:\tau}, x_{0:\tau}]. \quad (\text{R.1})$$

Though any measure of goodness such as prediction error $e_t = (\mathbb{E} [x_t] - x_t)^2$, could be used rather than the likelihood.

However typically this results in neglecting the prior parameterisation of Ω which contains knowledge of the past data and action sets. This would be akin to an individual waking up each day with a sensory system that reconfigures its connections and weights to optimally represent only that episodes actions and data. However if a prior actions and data could be accounted for then there would be an episodic refinement to the parameter learning. An example of an episodic refinement is discussed in the next section.

R.2 Episodic maximum-likelihood

Episodic learning uses a prior parameter to help construct a current estimate for this parameter with a data set. To do this I utilise a Lagrangian based upon a measure of dissimilarity between these parameters, the prior parameter and a new estimate, to derive the learning rules. I use the EM-algorithm for this learning process due to its utility as a pseudo Bayesian learning method and its wide acceptance for constructing optimal parameters.

Retaining the previous episodes optimal parameterisation gives a way of retaining knowledge of past data without needing to maintain a copy of the past data set. This can be a significant space saving for a small loss with respect to optimality. The degree of influence upon the current parameterisation is managed via the magnitude of the Lagrangian multipliers. This can be interpreted by contrast to standard EM. EM sequentially optimises a bound B upon the likelihood \mathcal{L} of the data Dat for a parameterisation Ω over a set of hidden variables H , which are related by,

$$\mathcal{L}(\text{Dat}|\Omega) = \log p(\text{Dat}|\Omega) \quad (\text{R.2})$$

$$= \log \int dH p(\text{Dat}, H|\Omega) \quad (\text{R.3})$$

$$\leq \int dH p(H|\text{Dat}, \Omega^{\text{old}}) \log \frac{p(\text{Dat}, H|\Omega)}{p(H|\text{Dat}, \Omega^{\text{old}})} \quad (\text{R.4})$$

$$= B[\Omega; \Omega^{\text{old}}], \quad (\text{R.5})$$

when the parameters have converged the bound will tend to the likelihood (Appendix A). The episodic EM modifies the bound by a Lagrangian multiplier γ_Ω and a measure of dissimilarity $d(\Omega; \bar{\Omega})$ between the parameterisation Ω and the prior parameter $\bar{\Omega}$, to be,

$$G[\Omega; \Omega^{\text{old}}|\bar{\Omega}] = \int dH p(H|\text{Dat}, \Omega^{\text{old}}) \log \frac{p(\text{Dat}, H|\Omega)}{p(H|\text{Dat}, \Omega^{\text{old}})} + \gamma_\Omega d(\Omega; \bar{\Omega}), \quad (\text{R.6})$$

where the prior parameter is distinct from the old update for the EM. This suggests that, if we assume a prior $p(\Omega) \triangleq \frac{1}{Z_{\bar{\Omega}}} e^{\gamma_\Omega d(\Omega; \bar{\Omega})}$ where $Z_{\bar{\Omega}} = \int d\Omega e^{\gamma_\Omega d(\Omega; \bar{\Omega})}$, then the joint belief of the parameterisation and the data is,

$$\mathcal{J}(\text{Dat}, \Omega) = \log p(\text{Dat}, \Omega) \quad (\text{R.7})$$

$$= \log p(\text{Dat}|\Omega) + \gamma_\Omega d(\Omega; \bar{\Omega}) - Z_{\bar{\Omega}}. \quad (\text{R.8})$$

Hence such a Lagrangian multiplier is equivalent to applying a prior belief in the estimation of Ω making any optimisation that of the joint belief.

The methodology of the EM algorithm is to first construct a proposal distribution over the hidden latent variables H conditioned upon the data Dat and the prior parameters Ω^{old} , $p(H|\text{Dat}, \Omega^{\text{old}})$; this constructs the expectation or E-step of the EM-algorithm. Then finding the stationary point of the bounds Lagrangian $G[\Omega; \Omega^{\text{old}}|\bar{\Omega}]$ with respect to the parameters Ω to iteratively maximise the joint $\mathcal{J}(\text{Dat}, \Omega)$; this constructs the new parameters and so constitutes the maximisation or M-step of the EM-algorithm. The EM-algorithm is conceptually a pseudo maximum likelihood where the expectation $p(H|\text{Dat}, \Omega^{\text{old}})$ is used in the maximisation of $p(\text{Dat}, H|\Omega)$ with respect to Ω and the prior defined by the Lagrangian multiplier $\gamma_{\Omega}d(\Omega, \bar{\Omega})$. In the following I use the gradients to construct analytical updates for the learning rules as stationary points to the episodic Lagrangian $G[\Omega; \Omega^{\text{old}}|\bar{\Omega}]$. This is somewhat like an iterative maximum a-posteriori (MAP) algorithm rather than an ML algorithm. Though implicitly all ML algorithms assume a uniform prior.

The advantage of this *episodic EM-algorithm* is that it may be applied to any model that is learnable using a likelihood based method, including expectation-maximisation based variational learning methods such as Ghahramani and Beal (2000), Ghahramani and Hinton (2000), Jordan *et al.* (1999), Verbeek *et al.* (2003), Beal *et al.* (2003), Hospedales and Vijayakumar (2006), Hospedales *et al.* (2007). It does however require that a suitable measure of dissimilarity exists for the particular parameter being learnt. Such measures are developed in the following subsection.

R.3 The dissimilarity between parameters

A suitable measures of dissimilarity for Bayesian parameters can be constructed from the KL-divergence between two probability distributions. After a bound $B[\Omega; \Omega^{\text{old}}]$ has been constructed a dissimilarity measure $d(\Omega, \bar{\Omega})$ between the current parameter θ and the prior episodes parameter $\bar{\Omega}$, where the Lagrangian is

$$G[\Omega; \Omega^{\text{old}}|\bar{\Omega}] = B[\Omega; \Omega^{\text{old}}] - \gamma_{\Omega}d(\Omega, \bar{\Omega}). \quad (\text{R.9})$$

So for example, for a discrete probability distribution $P(a|b) = \pi_{a|b}$ the dissimilarity being measured directly by the KL divergence

$$d(\pi_{a|b}, \bar{\pi}_{a|b}) = \sum_{a,b} \bar{\pi}_{a|b} \log \frac{\bar{\pi}_{a|b}}{\pi_{a|b}}. \quad (\text{R.10})$$

Taking the derivative of this divergence gives

$$\frac{\partial}{\partial \pi_{a|b}} d(\pi_{a|b}, \bar{\pi}_{a|b}) = -\frac{\bar{\pi}_{a|b}}{\pi_{a|b}}, \quad (\text{R.11})$$

which is suitable for maximum likelihood methods. A suitable measure of distance for Gaussian parameters is the KL divergence between two Gaussians $p(\mathbf{x}) = \phi(\mathbf{x}; \boldsymbol{\mu}_p, \Sigma_p)$ and $q(\mathbf{x}) = \phi(\mathbf{x}; \boldsymbol{\mu}_q, \Sigma_q)$ is

$$\int dx \phi(\mathbf{x}; \boldsymbol{\mu}_p, \Sigma_p) \log \frac{\phi(\mathbf{x}; \boldsymbol{\mu}_p, \Sigma_p)}{\phi(\mathbf{x}; \boldsymbol{\mu}_q, \Sigma_q)}, \quad (\text{R.12})$$

which analytically gives the measure of dissimilarity

$$d(p, q) = \frac{1}{2} \log \frac{|\Sigma_q|}{|\Sigma_p|} + \frac{1}{2} \text{tr} [\Sigma_q^{-1} (\Sigma_p - \Sigma_q)] + \frac{1}{2} (\boldsymbol{\mu}_p - \boldsymbol{\mu}_q)^T \Sigma_q^{-1} (\boldsymbol{\mu}_p - \boldsymbol{\mu}_q). \quad (\text{R.13})$$

The divergence between the various parameters in most models can be constructed by suitable selection of the Σ 's and $\boldsymbol{\mu}$'s. As an example the gradient of this divergence for a covariance matrix Σ_p by setting $\boldsymbol{\mu}_p = \boldsymbol{\mu}_q = \mathbf{0}_D$, gives

$$\frac{\partial}{\partial \Sigma_p^{-1}} d(p, q) = \frac{1}{2} (\Sigma_p - \Sigma_q), \quad (\text{R.14})$$

and for μ_p and by setting $\Sigma_p = \Sigma_q$, gives

$$\frac{\partial}{\partial \boldsymbol{\mu}_p} d(p, q) = \Sigma_p^{-1} (\boldsymbol{\mu}_q - \boldsymbol{\mu}_p), \quad (\text{R.15})$$

which is suitable for maximum likelihood methods for learning Gaussian problems. Therefore by a suitable selection of the variables most parameters may have an episodic Lagrangian applied to its learning algorithm.

R.4 The Lagrangian multiplier

How should γ_θ be interpreted? I consider this in the context of optimising a Gaussian. For the ML case of fitting a Gaussian distribution of mean μ and variance σ ,

$$p(x) = \phi(x; \mu, \sigma), \quad (\text{R.16})$$

to a data set $\text{Dat} = \{x_i : i = 1, \dots, N\}$ and the parameterisation $\Omega = \{\mu, \sigma\}$. The likelihood and its Lagrangian becomes

$$\mathcal{L}(\text{Dat}, \Omega) \propto \log p(\text{Dat}|\Omega) + \gamma_\Omega d(\Omega; \bar{\Omega}), \quad (\text{R.17})$$

$$= \sum_{i=1}^N \log \phi(x_i; \mu, \sigma) - \gamma_\theta \left(\frac{1}{2} \log \frac{\bar{\sigma}}{\sigma} + \frac{1}{2\bar{\sigma}} (\sigma + \bar{\sigma}) + \frac{1}{2\bar{\sigma}} (\mu + \bar{\mu})^2 \right). \quad (\text{R.18})$$

As can be seen the Lagrangian maintains a relationship between the mean μ and the variance σ .

To start I define the episodic Lagrangian multiplier to take the form $\gamma_\Omega = \gamma\alpha$. I extract α after holding $\gamma = 1$; hence γ can be interpreted as a ratio. Starting by extracting the ML mean

$$\frac{\partial \mathcal{L}}{\partial \mu} = - \sum_{i=1}^N \frac{1}{\sigma} (x_i - \mu) - \gamma_\Omega \frac{1}{\sigma} (\bar{\mu} - \mu). \quad (\text{R.19})$$

Equating the gradient of the likelihood to zero and rearranging to make μ the subject of the equation gives

$$\mu = \frac{1}{N + \gamma_\Omega} \left(\sum_{i=1}^N x_i + \gamma_\Omega \bar{\mu} \right). \quad (\text{R.20})$$

Using an independent but identically distributed data set for $\bar{\mu}$ which can be expressed as

$$\bar{\mu} = \frac{1}{\bar{N}} \sum_{i=N+1}^{N+\bar{N}} x_i, \quad (\text{R.21})$$

inserting $\gamma_\Omega = \gamma\alpha$ and holding $\gamma = 1$,

$$\mu = \frac{1}{N + \alpha} \left(\sum_{i=1}^N x_i + \alpha \frac{1}{\bar{N}} \sum_{i=N+1}^{N+\bar{N}} x_i \right) = \frac{1}{\bar{N} + N} \sum_{i=1}^{\bar{N}+N} x_i, \quad (\text{R.22})$$

from which I imply $\alpha = \bar{N}$. Accordingly the episodic update becomes

$$\mu = \frac{1}{N + \gamma\bar{N}} \left(\sum_{i=1}^N x_i + \gamma\bar{N}\bar{\mu} \right). \quad (\text{R.23})$$

Continuing for σ and setting $\gamma_\Omega = \gamma\bar{N}$ thus

$$\frac{\partial \mathcal{L}}{\partial \sigma^{-1}} = \frac{N}{2}\sigma - \frac{1}{2} \sum_{i=1}^N (x_i - \mu)^2 - \gamma_\Omega \frac{1}{2} (\bar{\sigma} - \sigma). \quad (\text{R.24})$$

Equating the gradient of the likelihood to zero and making σ the subject of the equation gives the update,

$$\sigma = \frac{1}{N + \gamma_\Omega} \left(\sum_{i=1}^N (x_i - \mu)^2 + \gamma_\Omega \bar{\sigma} \right). \quad (\text{R.25})$$

Using the same notation for Equation R.23 then I may say

$$\bar{\sigma} = \frac{1}{\bar{N}} \sum_{i=N+1}^{N+\bar{N}} (x_i - \mu)^2, \quad (\text{R.26})$$

and substituting in $\gamma_\Omega = \gamma\bar{N}$ and setting $\gamma = 1$,

$$\sigma = \frac{1}{N + \bar{N}} \left(\sum_{i=1}^N (x_i - \mu)^2 + \sum_{i=N+1}^{N+\bar{N}} (x_i - \bar{\mu})^2 \right) \neq \frac{1}{N + \bar{N}} \sum_{i=1}^{N+\bar{N}} (x_i - \mu)^2. \quad (\text{R.27})$$

This discrepancy with Equation R.23 is principally due to the variance σ being computed in terms of the mean μ . However as the number of episodes increases the estimated variance should converge to the true variance. The maximum-likelihood episodic update for the mean is,

$$\mu = \frac{1}{N + \gamma\bar{N}} \left(\sum_{i=1}^N x_i + \gamma\bar{N}\bar{\mu} \right), \quad (\text{R.28})$$

and the variance is

$$\sigma = \frac{1}{N + \gamma\bar{N}} \left(\sum_{i=1}^N (x_i - \mu)^2 + \gamma\bar{N}\bar{\sigma} \right), \quad (\text{R.29})$$

where \bar{N} is the size of the prior episodes data set and γ represents the weight assigned to the prior episodes parameterisation relative to the current episodes data set. This is illustrated in Figure R.1 with different γ ratios. The effect of the different ratios is best seen at episode 80 where the data sampling transitions to a new parameterisation. The closer γ is to 1 the greater the number of episodes the ML optimisation needs to approach the new parameters. This is however mitigated by $\gamma \rightarrow 1$ planning less relative emphasis

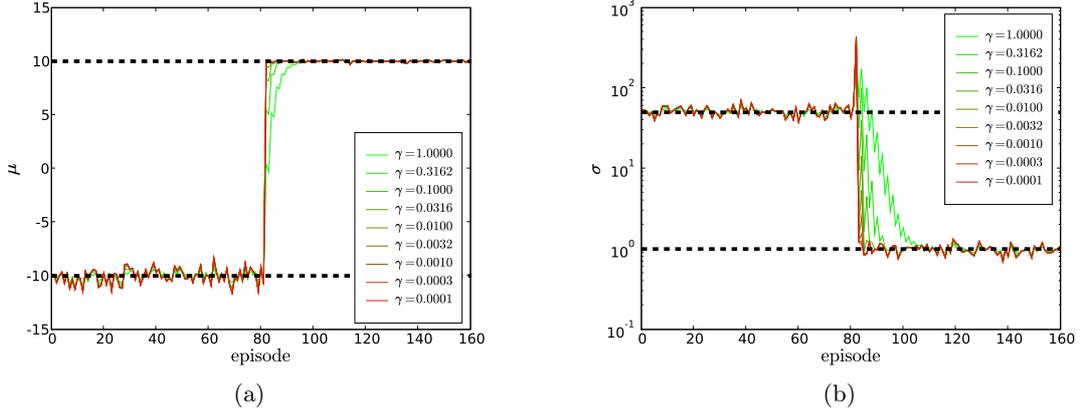


Figure R.1: Convergence of episodic maximum-likelihood (ML) estimates compared with the ratio γ for initial divergent proposals. At episode 80 there is a transition in the data from one parameter to another. The purpose of this transition is to indicate the varying speed with which the different γ 's converge to the new parameters. The true parameters are indicated with dashed lines. For $\gamma \rightarrow 1$ acts as a smoothed less variational estimate for the parameters. Where each episodes sample size is $N = 100$.

on the data and its estimates jump less.

In practical terms it is easiest to set $\bar{N} = N$ which results in γ simply being a ratio ($\gamma \geq 0$), where $\gamma = 1$ causes $\bar{\mu}$ to have the same importance as the data set $x_{1:N}$, where $\gamma < 1$ means the update holds the current episodes data to be more relevant than prior episodes. This is applied to the EM updates to set the individual Lagrangian multipliers to $\gamma_{\Omega_i} = \gamma \alpha_{\Omega_i}$ where each Ω_i represents a set of linked parameters such that γ is a ratio across all parameters. An example of which would be for a Gaussian mixture model where the belief in each mixture would have one $d(\Omega_i; \bar{\Omega}_i)$ Equation R.13 and the class weights would have one $d(\Omega_i; \bar{\Omega}_i)$ similar to Equation R.11.

Appendix S

Particle filtering

Tracking is a fundamental aspect of surveillance, guidance and obstacle avoidance systems, with the role being to determine the number, position and motion of any targets. The principle building block of a tracking system is the filter (a recursive state estimator). However an implicit building block of the filter is the model. In many domains the model's parametrisation is assumed (Ristic *et al.*, 2004c). A backward process whereby previous estimates of the state are typically used to modulate previous estimates with current measurements.

Using a state space L_t I wish to compute the distribution $p(L_t|x_{0:t})$ where L_t is the hidden state at a time t , $x_{0:t}$ is the sequence of measurements $\{x_i : i \in [0, t]\}$. In general the required integrals cannot be computed in a closed form. Therefore I shall use particle filtering which approximates the posterior using sequential importance sampling.

A particle filter can be defined according to the simple graphical model of Figure S.1. Firstly the prior belief $p(l_{t-1})$ is approximated using a mixture of Dirac deltas,

$$p(l_{t-1}) = \frac{1}{N} \sum_{i=1}^N \delta(l_{t-1} - l_{t-1}^{(i)}), \quad (\text{S.3})$$

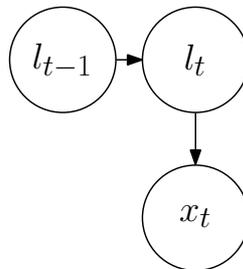


Figure S.1: Graphical model of the general filtering problem. The figure represents a generalised dynamic process $L_{t-1} \mapsto L_t$ with the measurement process $L_t \mapsto X_t$. The nodes represent the space that random variables will exist upon and the direction of the edges indicate causality within the model.

Algorithm S.1 The generic particle filtering algorithm. This algorithm is a simplification of the examples from Isard and Blake (1998), Murphy and Russell (2001), Doucet *et al.* (2000).

1. Sequential importance sampling step,

- for $i = 1, \dots, N$, sample

$$l_t^{*(i)} \sim p\left(l_t | l_{0:t-1}^{(i)}, x_{0:t-1}\right), \quad (\text{S.1})$$

- for $i = 1, \dots, N$, evaluate the importance of the proposed trajectory according to,

$$\pi_t^{(i)} \propto p\left(x_t | l_t^{*(i)}, x_{0:t-1}\right), \quad (\text{S.2})$$

where these weights are normalised, $\sum_{i=1}^N \pi_t^{(i)} = 1$.

2. Selection step,

- resample N samples from $l_t^{*(i)}$ according to the importance distribution $\pi_t^{(i)}$ to obtain N random samples $l_t^{(i)}$ approximating the distribution $p(L_t | x_{0:t})$.
-

where the set of $l_t^{(i)}$ represent the particles. This is used to feed forward through the stochastic dynamic process defined as $p(l_t | l_{t-1})$ to generate a latent space prediction,

$$p(l_t) = \int dl_{t-1} p(l_t | l_{t-1}) p(l_{t-1}), \quad (\text{S.4})$$

$$= \frac{1}{N} \sum_{i=1}^N p\left(l_t | l_{t-1}^{(i)}\right). \quad (\text{S.5})$$

A new proposal particle distribution for $p(l_t)$ is generated according to Equation S.5. This is however difficult to develop further, hence a set of sample properties are generated each denoted as $l_t^{*(i)}$. Such that for each particle sample a point $l_t^{*(i)} \sim p\left(l_t | l_{t-1}^{(i)}\right)$,¹ approximating the prediction distribution as

$$p(l_t) \approx \frac{1}{N} \sum_{i=1}^N \delta\left(l_t - l_t^{*(i)}\right). \quad (\text{S.6})$$

Each point in this set of points has a likelihood or weight $\pi_t^{(i)}$ assigned according to a measurement x_t ,

$$\pi_t^{(i)} \propto p\left(x_t | l_t^{*(i)}\right), \quad (\text{S.7})$$

where $\sum_{i=1}^N \pi_t^{(i)} = 1$. Finally a set of $l_t^{(i)}$'s are sampled according to the weights of each

¹Sampling an arbitrary distribution is relatively simple using sampling methods MacKay (2003).

particle $l_t^{*(i)}$ to generate the posterior belief,

$$p(l_t|x_t) = \frac{1}{N} \sum_{i=1}^N \delta(l_t - l_t^{(i)}). \quad (\text{S.8})$$

This process is summarised in Algorithm S.1 to define the generic particle filter.

Appendix T

The RBPF model

Assuming that trajectories representing the posterior filtered belief in $\mathbf{K}_{0:t}, \mathbf{J}_{0:t}$ are sampled according to Algorithm 4.1. This gives a joint posterior distribution of

$$P(k_{0:t}, j_{0:t} | \mathbf{x}_{0:t}, a_{0:t}) = \frac{1}{N} \sum_{i=1}^N \prod_{m=0}^t \delta_{k_m}^{k_m^{(i)}} \delta_{j_m}^{j_m^{(i)}}. \quad (\text{T.1})$$

Henceforth for brevity a shorthand $s = \{j, k\}$ is used where appropriate, this has the consequence that $\mathbf{S}_t = \mathbf{J}_t \cup \mathbf{K}_t$, $s_t^{(i)} = \{j_t^{(i)}, k_t^{(i)}\}$ and $s_{0:t}^{(i)} = \{j_{0:t}^{(i)}, k_{0:t}^{(i)}\}$. The state of the Kalman filter at a time t is denoted by z_t and remains unobserved. At every time t there is a vector of observations denoted by x_t which depend linearly upon the state vector with some additive Gaussian noise and a given choice of action a_t . I assume the following dynamic process,

$$p(\mathbf{z}_t | \mathbf{z}_{t-1}, k) = \phi(\mathbf{z}_t; A^{(k)} \mathbf{z}_{t-1} + \mathbf{b}^{(k)}, Q), \quad (\text{T.2})$$

the dynamics are governed by a Markov process; the state z_t is independent of all preceding states given the prior state z_{t-1} and k . This is only a Markov process for the RBPF in Figure 4.5 (b) if it is also conditioned upon a particles trajectory of $k_{0:t}^{(i)}$, as the dynamics are dependent upon $k_t^{(i)}$ as denoted by $A_t^{(i)} = A^{(k)}$ and $b_t^{(i)} = b^{(k)}$ for $k = k_t^{(i)}$. I assume the following measurement process,

$$p(\mathbf{x}_t | \mathbf{z}_t, j) = \phi(\mathbf{x}_t; B^{(j)} \mathbf{z}_t, R), \quad (\text{T.3})$$

where the measurement process is based upon a choice of measure j which is conditioned upon the i 'th particle through $j_{0:t}^{(i)}$, I use the shorthand of $B_t^{(i)} = B^{(j)}$ for $j = j_t^{(i)}$. The

measurement mapping $B^{(j)}$ is assumed to be fixed with the choice of j_t corresponding to a_t through the belief $P(j_t|a_t)$.

To generate the filtering terms a proposal distribution is generated for each particle (mathematical details in Appendix U.3) which has a Gaussian form,

$$p\left(\mathbf{z}_t | s_{0:t}^{(i)}, \mathbf{x}_{0:t-1}\right) = \phi\left(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t-1}^{(i)}, \Sigma_{t|t-1}^{(i)}\right), \quad (\text{T.4})$$

where the projected mean is

$$\hat{\mathbf{z}}_{t|t-1}^{(i)} = A_t^{(i)} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + \mathbf{b}_t^{(i)}, \quad (\text{T.5})$$

and the projected covariance is

$$\Sigma_{t|t-1}^{(i)} = A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{(i)T} + Q. \quad (\text{T.6})$$

These are used to construct the weights of the particles $p\left(\mathbf{x}_t | s_{0:t}^{(i)}, \mathbf{x}_{0:t-1}\right)$ by taking the expectation of Equation U.2 (mathematical details in Appendix U.3) to give the belief

$$p\left(\mathbf{x}_t | s_{0:t}^{(i)}, \mathbf{x}_{0:t-1}\right) = \phi\left(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{(i)}, H_{t|t-1}^{(i)}\right), \quad (\text{T.7})$$

where the expected measurement is

$$\hat{\mathbf{x}}_{t|t-1}^{(i)} = B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{(i)}, \quad (\text{T.8})$$

and its covariance is

$$H_{t|t-1}^{(i)} = R + B_t^{(i)} \Sigma_{t|t-1}^{(i)} B_t^{(i)T}. \quad (\text{T.9})$$

Conditioning upon a measurement (mathematical details in Appendix U.3) results in the terms for the filtered mean and covariance to give the belief

$$p\left(\mathbf{z}_t | s_{0:t}^{(i)}, \mathbf{x}_{0:t}\right) = \phi\left(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t}^{(i)}, \Sigma_{t|t}^{(i)}\right), \quad (\text{T.10})$$

where the filtered mean is

$$\hat{\mathbf{z}}_{t|t}^{(i)} = \hat{\mathbf{z}}_{t|t-1}^{(i)} + K_t^{(i)} \left(\mathbf{x}_t - B_t^{(i)} \hat{\mathbf{x}}_{t|t-1}^{(i)}\right), \quad (\text{T.11})$$

the filtered covariance is

$$\Sigma_{t|t}^{(i)} = \left(I - K_t^{(i)} B_t^{(i)} \right) \Sigma_{t|t-1}^{(i)} \left(I - K_t^{(i)} B_t^{(i)} \right)^T + K_t^{(i)} R K_t^{(i)T}, \quad (\text{T.12})$$

and the Kalman gain is

$$K_t^{(i)} = \Sigma_{t|t-1}^{(i)} B_t^{(i)T} \left(R + B_t^{(i)} \Sigma_{t|t-1}^{(i)} B_t^{(i)T} \right)^{-1}. \quad (\text{T.13})$$

This defines the stages necessary to compute the analytic component updates for my RBPF, which allows me to express Algorithm 4.1 more fully as Algorithm T.1.

The state space smoothing terms are conditioned upon the trajectories $s_{0:\tau}^{(i)} = \{k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}\}$ which represent the distribution $P(s_{0:\tau} | \mathbf{x}_{0:\tau}, a_{0:\tau})$ as

$$P(s_{0:\tau} | \mathbf{x}_{0:\tau}, a_{0:\tau}) = \frac{1}{N} \sum_{i=1}^N \prod_{t=0}^{\tau} \delta_{s_t}^{s_t^{(i)}}. \quad (\text{T.19})$$

The analytic smoothing terms (mathematical details in Appendix U.4) can be represented as

$$p(\mathbf{z}_t | s_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}) = \phi(\mathbf{z}_t; \hat{\mathbf{z}}_{t|\tau}^{(i)}, \Sigma_{t|\tau}^{(i)}), \quad (\text{T.20})$$

where the smoothed mean is

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = \hat{\mathbf{z}}_{t|t}^{(i)} + J_t^{(i)} \left(\hat{\mathbf{z}}_{t+1|\tau}^{(i)} - A_{t+1}^{(i)} \hat{\mathbf{z}}_{t|t}^{(i)} - \mathbf{b}_{t+1}^{(i)} \right), \quad (\text{T.21})$$

the smoothed covariance is

$$\Sigma_{t|\tau}^{(i)} = \Sigma_{t|t}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1|\tau}^{(i)} - \Sigma_{t+1|t}^{(i)} \right) J_t^{(i)T}, \quad (\text{T.22})$$

and the term

$$J_t^{(i)} = \Sigma_{t|t}^{(i)} A_{t+1}^{(i)T} \Sigma_{t+1|t}^{(i)-1}. \quad (\text{T.23})$$

This is however only half of the solution as the joint distribution

$$p(\mathbf{z}_{t-1:t} | s_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}) = \phi \left(\begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{z}_t \end{bmatrix}; \begin{bmatrix} \hat{\mathbf{z}}_{t-1|\tau}^{(i)} \\ \hat{\mathbf{z}}_{t|\tau}^{(i)} \end{bmatrix}, \begin{bmatrix} \Sigma_{t-1|\tau}^{(i)} & \Sigma_{t-1,t|\tau}^{(i)} \\ \Sigma_{t,t-1|\tau}^{(i)} & \Sigma_{t|\tau}^{(i)} \end{bmatrix} \right), \quad (\text{T.24})$$

Algorithm T.1 The Rao-Blackwellized particle filtering algorithm for a switching SSM problem. The shorthand $s = \{j, k\}$ is used, this has the consequence that $S_t = J_t \cup K_t$, $s_t^{(i)} = \{j_t^{(i)}, k_t^{(i)}\}$ and $s_{0:t}^{(i)} = \{j_{0:t}^{(i)}, k_{0:t}^{(i)}\}$. Notice that the full trajectories $s_{0:t}^{(i)}$ for each particle represents an instantiation of $P(s_{0:t}|\mathbf{x}_{0:t}, a_{0:t})$. This algorithm is adapted from Algorithm 4.1 which was in turn adapted from Murphy and Russell (2001).

1. Sequential importance sampling step,

- for $i \in [1, N]$, sample

$$s_t^{\star(i)} \sim \hat{P}\left(s_t | s_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}, a_t\right), \quad (\text{T.14})$$

and set $s_{0:t}^{\star(i)} \triangleq \{s_t^{\star(i)}, s_{0:t-1}^{(i)}\}$.

- for $i \in [1, N]$, evaluate the importance of the proposed trajectory according to,

$$\pi_t^{(i)} \propto \phi\left(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{\star(i)}, H_{t|t-1}^{\star(i)}\right) \frac{P\left(s_t^{\star(i)} | s_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}, a_t\right)}{\hat{P}\left(s_t^{\star(i)} | s_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}, a_t\right)}, \quad (\text{T.15})$$

where, $\hat{\mathbf{x}}_{t|t-1}^{\star(i)} = B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{\star(i)}$, $H_{t|t-1}^{\star(i)} = R + B_t^{(i)} \Sigma_{t|t-1}^{\star(i)} B_t^{(i)T}$, $\hat{\mathbf{z}}_{t|t-1}^{\star(i)} = A_t^{\star(i)} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + \mathbf{b}_t^{\star(i)}$ and, $\Sigma_{t|t-1}^{\star(i)} = A_t^{\star(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{\star(i)T} + Q$. Then normalise these weights such that, $\sum_{i=1}^N \pi_t^{(i)} = 1$.

2. Selection step,

- resample N samples from $s_{0:t}^{\star(i)}$ according to the importance distribution $\pi_t^{(i)}$ to obtain N random samples $s_{0:t}^{(i)}$ approximating the distribution $p(s_{0:t}|\mathbf{x}_{0:t}, a_{0:t})$.

3. Exact step,

- update the parameters with the selected $s_{0:t}^{(i)}$ using,

$$\hat{\mathbf{z}}_{t|t}^{(i)} = \hat{\mathbf{z}}_{t|t-1}^{(i)} + K_t^{(i)} \left(\mathbf{x}_t - B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{(i)}\right), \quad (\text{T.16})$$

$$\Sigma_{t|t}^{(i)} = \left(I_t - K_t^{(i)} B_t^{(i)}\right) \Sigma_{t|t-1}^{(i)} \left(I_t - K_t^{(i)} B_t^{(i)}\right)^T + K_t^{(i)} R K_t^{(i)T}, \quad (\text{T.17})$$

$$K_t^{(i)} \triangleq \Sigma_{t|t-1}^{(i)} B_t^{(i)T} \left(R + B_t^{(i)} \Sigma_{t|t-1}^{(i)} B_t^{(i)T}\right)^{-1}, \quad (\text{T.18})$$

where, $\hat{\mathbf{z}}_{t|t-1}^{(i)} = A_t^{(i)} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + \mathbf{b}_t^{(i)}$ and $\Sigma_{t|t-1}^{(i)} = A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{(i)T} + Q$.

Algorithm T.2 The Rao-Blackwellized particle smoothing algorithm for a switching SSM problem. $\Sigma_{t|t}^{(i)}$, $\Sigma_{t+1|t}^{(i)}$, $\mathbf{z}_{t|t}^{(i)}$ and $\mathbf{z}_{t+1|t}^{(i)}$ have been previously by filtering in Algorithm T.1. A shorthand $s = \{j, k\}$ is used, this has the consequence that $S_t = J_t \cup K_t$, $s_t^{(i)} = \{j_t^{(i)}, k_t^{(i)}\}$ and $s_{0:t}^{(i)} = \{j_{0:t}^{(i)}, k_{0:t}^{(i)}\}$. The full trajectories $s_{0:\tau}^{(i)}$ for each particle represents a smoothed instantiation of $P(s_{0:\tau}|\mathbf{x}_{0:\tau}, a_{0:\tau})$ from Algorithm T.1.

1. Exact step,

- update the parameters for the selected $s_{0:\tau}^{(i)}$ using,

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = \hat{\mathbf{z}}_{t|t}^{(i)} + J_t^{(i)} \left(\hat{\mathbf{z}}_{t+1|\tau}^{(i)} - A_{t+1}^{(i)} \hat{\mathbf{z}}_{t|t}^{(i)} - \mathbf{b}_{t+1}^{(i)} \right), \quad (\text{T.27})$$

$$\Sigma_{t|\tau}^{(i)} = \Sigma_{t|t}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1|\tau}^{(i)} - \Sigma_{t+1|t}^{(i)} \right) J_t^{(i)T}, \quad (\text{T.28})$$

$$\Sigma_{t,t-1|\tau}^{(i)} = \left(\Sigma_{t|\tau}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1|\tau}^{(i)} - \Sigma_{t+1|t}^{(i)} \right) J_t^{(i)T} \right) J_{t-1}^{(i)T}, \quad (\text{T.29})$$

where, $J_t^{(i)} \triangleq \Sigma_{t|t}^{(i)} A_{t+1}^{(i)T} \Sigma_{t+1|t}^{(i)-1}$.

where the cross-covariance is

$$\Sigma_{t,t-1|\tau}^{(i)} = \left(\Sigma_{t|\tau}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1,t|\tau}^{(i)} - A_{t+1}^{(i)} \Sigma_{t|t}^{(i)} \right) \right) J_{t-1}^{(i)T}, \quad (\text{T.25})$$

$\Sigma_{t,t-1|\tau}^{(i)} = \Sigma_{t-1,t|\tau}^{(i)T}$ as the covariance of $\begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{z}_t \end{bmatrix}$ is by definition positive definite. The backward smoother for the cross covariance term is initialised with

$$\Sigma_{\tau,\tau-1|\tau}^{(i)} = \left(I - K_\tau^{(i)} B_\tau^{(i)} \right) A_\tau^{(i)} \Sigma_{\tau-1|\tau-1}^{(i)}. \quad (\text{T.26})$$

These define the analytic component smoothing relations for the RBPF as expressed in Algorithm T.1 and are expressed in Algorithm T.2.

The smoothing distributions Equations T.20 & T.24 as dedicated in Algorithm T.2 are used to iteratively and optimally update the model parameters which I derive in Appendix V.

Appendix U

Kalman filtering & smoothing

U.1 The Kalman filter

I assume that trajectories for $K_{0:t}$ and $J_{0:t}$ have already been sampled (Algorithm 4.1), therefore algorithmically these are neglected in the following. Much of this appendix is a generalisation of typical Kalman filtering adapted from Welling (2008) to include a time dependent state dynamic and measurement processes.

The state of the Kalman filter is denoted by \mathbf{z}_t at a time t and remains unobserved. At every time t there is a vector of observations denoted by \mathbf{x}_t which depend linearly upon the state vector with some additive Gaussian noise and a given choice of action a_t . I assume the following dynamic process,

$$\mathbf{z}_t \sim \phi\left(\cdot; A^{(k)}\mathbf{z}_{t-1} + \mathbf{b}^{(k)}, Q\right), \quad (\text{U.1})$$

the dynamics are governed by a Markov process; the state \mathbf{z}_t is independent of all preceding states given the prior state \mathbf{z}_{t-1} . This is only a Markov process for the RBPF in Figure 4.5 (b) if it is also conditioned upon a particles trajectory $k_{0:t}^{(i)}$, as the dynamics are dependent upon $k_t^{(i)}$ as denoted by $A_t^{(i)} = A^{(k)}$ and $\mathbf{b}_t^{(i)} = \mathbf{b}^{(k)}$ for $k = k_t^{(i)}$.

$$\mathbf{x}_t \sim \phi\left(\cdot; B^{(j)}\mathbf{z}_t, R\right), \quad (\text{U.2})$$

where the measurement process is based upon a choice of measure j which is conditioned upon the i 'th particle through $j_{0:t}^{(i)}$, I use the shorthand of $B_t^{(i)} = B^{(j)}$ for $j = j_t^{(i)}$. This model can be seen as a mixture of conditionally independent factor analyses (FA) over time; at every instant there is a FA model where the factors depend upon the prior time

steps factors. Finally the initial state \mathbf{z}_0 is normally distributed according to,

$$\mathbf{z}_0 \sim \phi(\cdot; \boldsymbol{\mu}_0, \Sigma_0). \quad (\text{U.3})$$

Since the initial state is Gaussian, the evolution is linear and the noise from the measurements and state evolution are also Gaussian, this then implies that the state at later times will remain Gaussian.

U.2 Kalman filter properties

I want to be able to estimate the state and covariance of the state at any time t given a set of observations $\mathbf{x}_{0:\tau} = \{\mathbf{x}_0, \dots, \mathbf{x}_\tau\}$, a set of hypothesised measurements $j_{0:\tau}^{(i)} = \{j_0^{(i)}, \dots, j_\tau^{(i)}\}$ and the trajectory $k_{0:\tau}^{(i)} = \{k_0^{(i)}, \dots, k_\tau^{(i)}\}$. The actual choice of actions a_t are neglected in this Appendix for the purposes of brevity and that conditioning upon $J_{0:\tau}$ makes Z_t independent of a_t for $0 \leq t \leq \tau$. If τ is equal to the current time t this is a filter for the state, if τ is smaller than t then this is a prediction of the state, and finally, if τ is larger than t this is a smoothing of the state. The probability that describes these is,

$$p\left(\mathbf{z}_t | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}\right), \quad (\text{U.4})$$

as it conveys all of the information regarding \mathbf{z}_t at a time t given all the observations up to a time τ and a trajectory of states $k_{0:\tau}^{(i)}$ and $j_{0:\tau}^{(i)}$ upto a time τ . As this probability is Gaussian I need only calculate the mean and covariance (these being the sufficient statistics for such a distribution), denoted by,

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = \mathbb{E}_{p(\mathbf{z}_t | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} [\mathbf{z}_t], \quad (\text{U.5})$$

$$\Sigma_{t|\tau}^{(i)} = \mathbb{E}_{p(\mathbf{z}_t | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} \left[\tilde{\mathbf{z}}_{t|\tau}^{(i)} \tilde{\mathbf{z}}_{t|\tau}^{(i)T} \right], \quad (\text{U.6})$$

where $\tilde{\mathbf{z}}_{t|\tau}^{(i)} = \mathbf{z}_t - \hat{\mathbf{z}}_{t|\tau}^{(i)}$ is defined as the state prediction error. Note that these quantities depend upon the random variables $\mathbf{x}_{0:\tau}$ and hence are also random variables. We can see that the covariance P does not depend upon $\mathbf{x}_{0:\tau}$, thus Σ maybe considered a parameter in the following derivations. To prove this claim I show the correlation between the random variables $\tilde{\mathbf{z}}_{t|\tau}^{(i)}$ and $\mathbf{x}_{0:\tau}$ vanishes. For normally distributed random variables this implies independence.

The random variables $\tilde{\mathbf{z}}_{t|\tau}^{(i)} = \mathbf{z}_t - \hat{\mathbf{z}}_{t|\tau}^{(i)}$ and $\mathbf{x}_{0:\tau} = \{\mathbf{x}_0, \dots, \mathbf{x}_\tau\}$ are independent,

$$\begin{aligned} \mathbb{E}_{p(\mathbf{z}_t, \mathbf{x}_{0:\tau})} [\mathbf{x}_{0:\tau} \mathbf{z}_t^T] - \mathbb{E}_{p(\mathbf{x}_{0:\tau})} \left[\mathbf{x}_{0:\tau} \hat{\mathbf{z}}_{t|\tau}^{(i)T} \right] &= \int d\mathbf{z}_t d\mathbf{x}_{0:\tau} p(\mathbf{z}_t, \mathbf{x}_{0:\tau}) \mathbf{x}_{0:\tau} \mathbf{z}_t^T \\ &\quad - \int d\mathbf{x}_{0:\tau} p(\mathbf{x}_{0:\tau}) \mathbf{x}_{0:\tau} \times \\ &\quad \left[\int d\mathbf{z}_t p(\mathbf{z}_t | \mathbf{x}_{0:\tau}) \mathbf{z}_t \right]^T, \end{aligned} \quad (\text{U.7})$$

$$\begin{aligned} &= \int d\mathbf{z}_t d\mathbf{x}_{0:\tau} p(\mathbf{z}_t, \mathbf{x}_{0:\tau}) \mathbf{x}_{0:\tau} \mathbf{z}_t^T \\ &\quad - \int d\mathbf{z}_t d\mathbf{x}_{0:\tau} p(\mathbf{x}_{0:\tau}) \times \\ &\quad p(\mathbf{z}_t | \mathbf{x}_{0:\tau}) \mathbf{x}_{0:\tau} \mathbf{z}_t^T, \end{aligned} \quad (\text{U.8})$$

$$= 0, \text{ as } p(\mathbf{z}_t, \mathbf{x}_{0:\tau}) = p(\mathbf{z}_t | \mathbf{x}_{0:\tau}) p(\mathbf{x}_{0:\tau}) \quad (\text{U.9})$$

this implies,

$$\Sigma_{t_1, t_2}^{(i)} = \mathbb{E} \left[\tilde{\mathbf{z}}_{t_1}^{(i)} \tilde{\mathbf{z}}_{t_2}^{(i)T} \right], \quad (\text{U.10})$$

$$= \mathbb{E} \left[\mathbf{z}_{t_1} \mathbf{z}_{t_2}^T \right] - \mathbb{E} \left[\hat{\mathbf{z}}_{t_1}^{(i)} \hat{\mathbf{z}}_{t_2}^{(i)T} \right], \quad (\text{U.11})$$

another useful result is that the prediction measurement error $\boldsymbol{\varepsilon}_t = \mathbf{x}_t - B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{(i)}$ is independent of the past measurements $\mathbf{x}_{0:t-1}$. The predicted measurement error is also called the innovation, as it represents that part of the new measurement \mathbf{x}_t that cannot be predicted using knowledge of the $\mathbf{x}_{0:t-1}$ measurements as they are independent.

U.3 Kalman filtering

I shall now derive the Kalman filter update equations. First using,

$$p\left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t}\right) = \frac{p\left(\mathbf{x}_t | \mathbf{z}_t, j_t^{(i)}\right) p\left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right)}{p\left(\mathbf{x}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1}\right)}, \quad (\text{U.12})$$

where,

$$p\left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right) = \int d\mathbf{z}_{t-1} p\left(\mathbf{z}_t | \mathbf{z}_{t-1}, k_t^{(i)}\right) p\left(\mathbf{z}_{t-1} | k_{0:t-1}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right). \quad (\text{U.13})$$

The denominator in Equation U.12 is important as it defines the likelihood of the i 'th particle in Algorithm 4.1 but can only be calculated from marginalising \mathbf{z}_t out of the

numerator. This is calculated at the end of this section. The remaining densities are defined to be,

$$p(\mathbf{x}_t | \mathbf{z}_t, j) = \phi(\mathbf{x}_t; B^{(j)} \mathbf{z}_t, R), \quad (\text{U.14})$$

where for compactness I use the shorthand $B_t^{(i)} \triangleq B^{(j_t^{(i)})}$ where appropriate, and,

$$p(\mathbf{z}_t | \mathbf{x}_{t-1}, k) = \phi(\mathbf{z}_t; A^{(k)} \mathbf{z}_{t-1} + \mathbf{b}^{(k)}, Q), \quad (\text{U.15})$$

where I use for compactness the shorthands $A_t^{(i)} \triangleq A^{(k_t^{(i)})}$ and $\mathbf{b}_t^{(i)} \triangleq \mathbf{b}^{(k_t^{(i)})}$ where appropriate. Equations U.12 and U.13 may be interpreted as a reactive re-enforcement due to an observation following a switched drift and Gaussian diffusion between observations. Equation U.12 is simply the application of Bayes law to a posterior,

$$p(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}) = \phi(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t-1}^{(i)}, \Sigma_{t|t-1}^{(i)}), \quad (\text{U.16})$$

this is simply updating the belief in an unknown random variable by the inclusion of evidence, the effect is to decrease the variance *cf* uncertainty. The second equation, Equation U.13 evolves the hidden state from one instant to the next without the consideration of evidence, this has the effect of increasing the variance and hence the uncertainty. Together these equations express $p(\mathbf{z}_t | \mathbf{x}_{0:t}, k_{0:t}^{(i)})$ in terms of $p(\mathbf{z}_{t-1} | k_{0:t-1}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})$ and form a recursive updating of belief in the hidden state z_t . It is easy to verify that in the case of prediction, as only Equation U.13 remains,

$$p(\mathbf{z}_t | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}) = \int d\mathbf{z}_{t-1} p(\mathbf{z}_t | \mathbf{z}_{t-1}, k_t^{(i)}) p(\mathbf{z}_{t-1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}). \quad \forall \tau < t \quad (\text{U.17})$$

I shall deal with smoothing ($\tau > t$) in the next section. Though for filtering I need a prediction to apply evidence to Equation U.12. I calculate the projected mean $\hat{\mathbf{z}}_{t|t-1}^{(i)}$ and covariance $\Sigma_{t|t-1}^{(i)}$ of the probability density function (PDF) $p(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})$ in terms of the mean $\hat{\mathbf{z}}_{t-1|t-1}^{(i)}$ and covariance $\Sigma_{t-1|t-1}^{(i)}$ of the PDF $p(\mathbf{z}_{t-1} | k_{0:t-1}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})$. These two estimators determine the conditional densities completely as the densities are Gaussian.

To calculate the updates the mean $\hat{\mathbf{z}}_{t|t-1}^{(i)}$ and the covariance $\Sigma_{t|t-1}$, first I rearrange the joint belief by exploiting both forms of conditionality,

$$p(\mathbf{z}_{t-1:t} | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}) = p(\mathbf{z}_t | \mathbf{z}_{t-1}, k_t^{(i)}) \times$$

$$p\left(\mathbf{z}_{t-1} | k_{0:t-1}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right) \quad (\text{U.18})$$

$$= p\left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right) \times \quad (\text{U.19})$$

$$p\left(\mathbf{z}_{t-1} | \mathbf{z}_t, k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right), \quad (\text{U.20})$$

which means the equality holds,

$$\phi\left(\mathbf{z}_t; A_t^{(i)} \mathbf{z}_{t-1} + \mathbf{b}_t^{(i)}, Q\right) \phi\left(\mathbf{z}_{t-1}; \hat{\mathbf{z}}_{t-1|t-1}^{(i)}, \Sigma_{t-1|t-1}^{(i)}\right) = \phi\left(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t-1}^{(i)}, \Sigma_{t|t-1}^{(i)}\right) \phi\left(\mathbf{z}_{t-1}; \mathbf{e}_t^{(i)}, E_t^{(i)}\right), \quad (\text{U.21})$$

with the unknowns $\hat{\mathbf{z}}_{t|t-1}^{(i)}$, $\Sigma_{t|t-1}^{(i)}$, $\mathbf{e}_t^{(i)}$ and $E_t^{(i)}$. Using Equations N.2 & N.3 gives,

$$E_t^{(i)} = \left(A_t^{(i)T} Q^{-1} A_t^{(i)} + \Sigma_{t-1|t-1}^{(i)-1} \right)^{-1}, \quad (\text{U.22})$$

and

$$\mathbf{e}_t^{(i)} = E_t^{(i)} \left(A_t^{(i)T} Q^{-1} \left(\mathbf{b}_t^{(i)} - \mathbf{z}_t \right) + \Sigma_{t-1|t-1}^{(i)-1} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} \right). \quad (\text{U.23})$$

It follows that the projected covariance,

$$\Sigma_{t|t-1}^{(i)} = \left(Q^{-1} + Q^{-1} A_t^{(i)} E_t^{(i)} A_t^{(i)T} Q^{-1} \right)^{-1}, \quad (\text{U.24})$$

so by Equation O.1,

$$\Sigma_{t|t-1}^{(i)} = A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} A_t^{(i)T} + Q, \quad (\text{U.25})$$

by and the mean,

$$\begin{aligned} & \Sigma_{t|t-1}^{(i)-1} \hat{\mathbf{z}}_{t|t-1}^{(i)} \\ &= \\ & Q^{-1} \mathbf{b}_t^{(i)} + Q^{-1} A_t^{(i)} E_t^{(i)} \left(\Sigma_{t-1|t-1}^{(i)-1} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} - A_t^{(i)T} Q^{-1} \mathbf{b}_t^{(i)} \right), \end{aligned} \quad (\text{U.26})$$

by Equations O.1 & O.2 applying to the first and second right hand terms respectively leads to,

$$\Sigma_{t|t-1}^{(i)-1} \hat{\mathbf{z}}_{t|t-1}^{(i)} = \Sigma_{t|t-1}^{(i)-1} \mathbf{b}_t^{(i)} + \Sigma_{t|t-1}^{(i)-1} A_t^{(i)} \Sigma_{t-1|t-1}^{(i)} \Sigma_{t-1|t-1}^{(i)-1} \hat{\mathbf{z}}_{t-1|t-1}^{(i)}, \quad (\text{U.27})$$

cancelling the $\Sigma_{\cdot|\cdot}^{(i)}$ matrices gives the projected mean,

$$\hat{\mathbf{z}}_{t|t-1}^{(i)} = A_t^{(i)} \hat{\mathbf{z}}_{t-1|t-1}^{(i)} + \mathbf{b}_t^{(i)}. \quad (\text{U.28})$$

Together Equations U.25 & U.28 are used to calculate the projected belief for a trajectory and to calculate the i 'th trajectories likelihood.

Next to calculate $\hat{\mathbf{z}}_{t|t}^{(i)}$ and $\Sigma_{t|t}^{(i)}$ in terms of the one step prediction I use Equation U.12, rewritten to be the joint belief,

$$p\left(\mathbf{x}_t, \mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1}\right) = p\left(\mathbf{x}_t | \mathbf{z}_t, j_t^{(i)}\right) p\left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}\right), \quad (\text{U.29})$$

$$= p\left(\mathbf{x}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1}\right) \times p\left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t}\right), \quad (\text{U.30})$$

which leads to the equality

$$\begin{aligned} \phi\left(\mathbf{x}_t; B_t^{(i)} \mathbf{z}_t, R\right) \phi\left(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t-1}^{(i)}, \Sigma_{t|t-1}^{(i)}\right) \\ = \\ \phi\left(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{(i)}, H_{t|t-1}^{(i)}\right) \phi\left(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t}^{(i)}, \Sigma_{t|t}^{(i)}\right), \end{aligned} \quad (\text{U.31})$$

with

$$\Sigma_{t|t}^{(i)} = \left(B_t^{(i)T} R^{-1} B_t^{(i)} + \Sigma_{t|t-1}^{(i)-1} \right)^{-1}, \quad (\text{U.32})$$

which by application of Equation O.1 gives

$$\Sigma_{t|t}^{(i)} = \left(I - K_t^{(i)} B_t^{(i)} \right) \Sigma_{t|t-1}^{(i)}, \quad (\text{U.33})$$

where

$$K_t^{(i)} = \Sigma_{t|t-1}^{(i)} B_t^{(i)T} \left(B_t^{(i)} \Sigma_{t|t-1}^{(i)} B_t^{(i)T} + R \right)^{-1}, \quad (\text{U.34})$$

and

$$\hat{\mathbf{z}}_{t|t}^{(i)} = \Sigma_{t|t}^{(i)} \left(\Sigma_{t|t-1}^{(i)-1} \hat{\mathbf{z}}_{t|t-1}^{(i)} + B_t^{(i)T} R^{-1} \mathbf{x}_t \right) \quad (\text{U.35})$$

which by application of Equations O.1 & O.2 for the first and second components on the right hand side of the relation gives

$$\hat{\mathbf{z}}_{t|t}^{(i)} = \left(I - K_t^{(i)} B_t^{(i)} \right) \hat{\mathbf{z}}_{t|t-1}^{(i)} + K_t^{(i)} \mathbf{x}_t, \quad (\text{U.36})$$

$$\equiv \hat{\mathbf{z}}_{t|t-1}^{(i)} + K_t^{(i)} \left(\mathbf{x}_t - B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{(i)} \right). \quad (\text{U.37})$$

The likelihood is computed from $\hat{\mathbf{x}}_{t|t-1}^{(i)}$ and $H_{t|t-1}^{(i)}$ which come from the solution to Equations

tion U.31 as

$$H_{t|t-1}^{(i)} = \left(R^{-1} - K_t^{(i)T} \Sigma_{t|t}^{(i)-1} K_t^{(i)} \right)^{-1}, \quad (\text{U.38})$$

which is expanded using the updates then rearrange both $K_t^{(i)}$'s using the reverse of Equation O.2 this cancels $\Sigma_{t|t}^{(i)-1}$ then rearrange using the reverse of Equation O.1 to give the result

$$H_{t|t-1}^{(i)} = R + B_t^{(i)} \Sigma_{t|t-1}^{(i)} B_t^{(i)T}, \quad (\text{U.39})$$

and similarly for the mean,

$$\hat{\mathbf{x}}_{t|t-1}^{(i)} = H_{t|t-1}^{(i)} K_t^{(i)T} \Sigma_{t|t}^{(i)-1} \left(I - K_t^{(i)} B_t^{(i)} \right) \hat{\mathbf{z}}_{t|t-1}^{(i)}, \quad (\text{U.40})$$

after expanding $K_t^{(i)}$ and expanding using the $\Sigma_{t|t}^{(i)}$ update, these cancel to give

$$\hat{\mathbf{x}}_{t|t-1}^{(i)} = B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{(i)}. \quad (\text{U.41})$$

Therefore in summary these relations give the filtered mean and covariance,

$$\hat{\mathbf{z}}_{t|t}^{(i)} = \hat{\mathbf{z}}_{t|t-1}^{(i)} + K_t^{(i)} \left(\mathbf{x}_t - B_t^{(i)} \hat{\mathbf{x}}_{t|t-1}^{(i)} \right), \quad (\text{U.42})$$

$$\Sigma_{t|t}^{(i)} = \left(I - K_t^{(i)} B_t^{(i)} \right) \Sigma_{t|t-1}^{(i)}, \quad (\text{U.43})$$

$$K_t^{(i)} = \Sigma_{t|t-1}^{(i)} B_t^{(i)T} \left(R + B_t^{(i)} \Sigma_{t|t-1}^{(i)} B_t^{(i)T} \right)^{-1}, \quad (\text{U.44})$$

where $K_t^{(i)}$ is defined as the Kalman gain. The equations are initialised by setting $\hat{\mathbf{z}}_{0|-1}^{(i)} = \boldsymbol{\mu}_0$, and $\Sigma_{0|-1}^{(i)} = \Sigma_0$ from Equation U.3. These equations as well as Equations U.28 & U.25 characterise the parametric component of my RBPF which represents the continuous on-line estimate of the systems state. Note that the Kalman gain grows if the observables covariance R decreases, thus assigning more weight to the measurements residual (the difference between the predicted and actual measurement). Further if the covariance $\Sigma_{t|t-1}^{(i)}$ becomes smaller less emphasis is placed by the model upon the measurement residual $(\mathbf{x}_t - B_t^{(i)} \hat{\mathbf{z}}_{t|t-1}^{(i)})$. Numerically Equation U.43 is not ideal due to it being the difference of two positive definite matrices, which is not guaranteed to result in a positive definite matrix and in implementation can lead to numerical instabilities. This is simple to fix, noting that Equation U.44 implies

$$K_t^{(i)} R K_t^{(i)T} = \left(I - K_t^{(i)} B_t^{(i)} \right) \Sigma_{t|t-1}^{(i)} B_t^{(i)T} K_t^{(i)T}, \quad (\text{U.45})$$

then

$$\Sigma_{t|t}^{(i)} = \left(I - K_t^{(i)} B_t^{(i)} \right) \Sigma_{t|t-1}^{(i)} \left(I - K_t^{(i)} B_t^{(i)} \right)^T + K_t^{(i)} R K_t^{(i)T}, \quad (\text{U.46})$$

which is the sum of two positive definite matrices and hence is guaranteed to result in a positive definite matrix. Therefore Equations U.42 & U.46 define the analytic component represented by $\alpha_t^{(i)}$ of the RBPF in step 3 of Algorithm 4.1,

$$p \left(\mathbf{z}_t | k_{0:t} = k_{0:t}^{(i)}, j_{0:t} = j_{0:t}^{(i)}, \mathbf{x}_{0:t} \right) = \phi \left(\mathbf{z}_t; \hat{\mathbf{z}}_{t|t}^{(i)}, \Sigma_{t|t}^{(i)} \right). \quad (\text{U.47})$$

Next I consider the likelihood of a trajectory.

U.3.1 The likelihood of a trajectory

To construct the denominator of Equation U.12 and hence compute the generic RBPF weights as defined in Algorithm 4.1. I compute the likelihood of a trajectory $k_{0:t}^{(i)}, j_{0:t}^{(i)}$ given all prior evidence $\mathbf{x}_{0:t-1}$ for a measurement \mathbf{x}_t , which is computed from the marginalising z_t from the joint distribution of \mathbf{x}_t and \mathbf{z}_t ,

$$p \left(\mathbf{x}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1} \right) = \int d\mathbf{z}_t p \left(\mathbf{x}_t | \mathbf{z}_t, j_t^{(i)} \right) p \left(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1} \right), \quad (\text{U.48})$$

which is according to Equations N.2 & N.3 Gaussian. The mean $\hat{\mathbf{x}}_{t|t-1}^{(i)}$ and covariance $H_{t|t-1}^{(i)}$ of this Gaussian distribution $p \left(\mathbf{x}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1} \right)$ is computed in the filtering section. The relations defined as Equations U.41 & U.39 previously fully describe the calculation of importance in step 1 of Algorithm 4.1 for the RBPF, as

$$p \left(\mathbf{x}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t-1} \right) = \phi \left(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{(i)}, H_{t|t-1}^{(i)} \right). \quad (\text{U.49})$$

Next I define the likelihood of a filter, which is equivalent to the likelihood of the model. The likelihood can merely be constructed from the individual trajectory weights,

$$L \left(\mathbf{x}_{0:\tau}; a_{0:\tau} \right) = p \left(\mathbf{x}_{0:\tau} | a_{0:\tau} \right), \quad (\text{U.50})$$

$$= \prod_{t=0}^{\tau} p \left(\mathbf{x}_t | \mathbf{x}_{0:t-1}, a_{0:t} \right), \quad (\text{U.51})$$

where $p \left(\mathbf{x}_0 | \mathbf{x}_{0:-1}, a_0 \right) \triangleq p \left(\mathbf{x}_0 | a_0 \right)$ where $\mathbf{x}_{0:-1}$ can be interpreted as an empty set. Using

$$p \left(\mathbf{x}_t | \mathbf{x}_{0:t-1}, a_{0:t} \right) = \sum_{k_{0:t}, j_{0:t}} p \left(\mathbf{x}_t | k_{0:t}, j_{0:t}, \mathbf{x}_{0:t-1} \right) P \left(k_{0:t}, j_{0:t} | \mathbf{x}_{0:t-1}, a_{0:t} \right), \quad (\text{U.52})$$

$$= \sum_{k_{0:t}, j_{0:t}} p(\mathbf{x}_t | k_{0:t}, j_{0:t}, \mathbf{x}_{0:t-1}) \frac{1}{N} \sum_{i=1}^N \delta_{k_{0:t}}^{k_{0:t}^{*(i)}} \delta_{j_{0:t}}^{j_{0:t}^{*(i)}}, \quad (\text{U.53})$$

$$\equiv \frac{1}{N} \sum_{i=1}^N p(\mathbf{x}_t | k_{0:t}^{*(i)}, j_{0:t}^{*(i)}, \mathbf{x}_{0:t-1}), \quad (\text{U.54})$$

hence, the likelihood becomes

$$L(\mathbf{x}_{0:\tau}; a_{0:\tau}) = \prod_{t=0}^{\tau} \frac{1}{N} \sum_{i=1}^N p(\mathbf{x}_t | k_{0:t}^{*(i)}, j_{0:t}^{*(i)}, \mathbf{x}_{0:t-1}). \quad (\text{U.55})$$

This produces a log-likelihood of,

$$\mathcal{L}(\mathbf{x}_{0:\tau}; a_{0:\tau}) = \sum_{t=0}^{\tau} \log \left(\frac{1}{N} \sum_{i=1}^N p(\mathbf{x}_t | k_{0:t}^{*(i)}, j_{0:t}^{*(i)}, \mathbf{x}_{0:t-1}) \right), \quad (\text{U.56})$$

where $k_{0:t}^{*(i)}$ and $j_{0:t}^{*(i)}$ are the filtering constructs of Algorithm 4.1 and constructs the distributions,

$$p(\mathbf{x}_t | k_{0:t}^{*(i)}, j_{0:t}^{*(i)}, \mathbf{x}_{0:t-1}) = \phi(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}^{*(i)}, H_{t|t-1}^{*(i)}). \quad (\text{U.57})$$

In the next section I define the smoothing process which I will use for learning the models parameters.

U.4 Kalman smoothing

Next I shall solve the smoothing problem for this model. This will use future measurements for $\mathbf{x}_{0:\tau}$, e.g. $\tau > t$, to improve the estimates for the states $\mathbf{z}_t \in \mathbf{z}_{0:\tau}$. The resultant estimates will be less noisy and hence smoother. This however causes the estimation to be performed offline.

Starting with the smoothed joint posterior,

$$p(\mathbf{z}_{t:t+1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}) = p(\mathbf{z}_t | \mathbf{z}_{t+1}, k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}) \times p(\mathbf{z}_{t+1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}), \quad (\text{U.58})$$

$$= p(\mathbf{z}_t | \mathbf{z}_{t+1}, k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t}) \times p(\mathbf{z}_{t+1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}), \quad (\text{U.59})$$

$$= \frac{p(\mathbf{z}_{t+1} | \mathbf{z}_t, k_t^{(i)}) p(\mathbf{z}_t | k_{0:t}^{(i)}, j_{0:t}^{(i)}, \mathbf{x}_{0:t})}{p(\mathbf{z}_{t+1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} \times p(\mathbf{z}_{t+1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}), \quad (\text{U.60})$$

then taking advantage of the fact that all of these quantities are Gaussian then,

$$p\left(\mathbf{z}_{t:t+1} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}\right) = \phi\left(\begin{bmatrix} \mathbf{z}_t \\ \mathbf{z}_{t+1} \end{bmatrix}; \begin{bmatrix} \hat{\mathbf{z}}_{t|\tau}^{(i)} \\ \hat{\mathbf{z}}_{t+1|\tau}^{(i)} \end{bmatrix}, \begin{bmatrix} \Sigma_{t|\tau}^{(i)} & \Sigma_{t,t+1|\tau}^{(i)} \\ \Sigma_{t+1,t|\tau}^{(i)} & \Sigma_{t+1|\tau}^{(i)} \end{bmatrix}\right), \quad (\text{U.61})$$

if I use the substitution,

$$\begin{bmatrix} \Sigma_{t|\tau}^{(i)} & \Sigma_{t,t+1|\tau}^{(i)} \\ \Sigma_{t+1,t|\tau}^{(i)} & \Sigma_{t+1|\tau}^{(i)} \end{bmatrix} = \begin{bmatrix} F_t^{(i)} & F_{t,t+1}^{(i)} \\ F_{t+1,t}^{(i)} & F_{t+1}^{(i)} \end{bmatrix}^{-1}, \quad (\text{U.62})$$

then there are the following equalities hold as a result of Equation U.60 and Equations U.61 & U.62,

$$F_t^{(i)} = \Sigma_{t|t}^{(i)-1} + A_{t+1}^{(i)T} Q^{-1} A_{t+1}^{(i)}, \quad (\text{U.63})$$

$$F_{t+1}^{(i)} = Q^{-1} - \Sigma_{t+1|t}^{(i)-1} + \Sigma_{t+1|\tau}^{(i)-1}, \quad (\text{U.64})$$

$$F_{t,t+1}^{(i)} = -A_{t+1}^{(i)T} Q^{-1}, \quad (\text{U.65})$$

and

$$F_{t,t+1}^{(i)} \hat{\mathbf{x}}_{t+1|\tau}^{(i)} + F_t^{(i)} \hat{\mathbf{x}}_{t|\tau}^{(i)} = A_{t+1}^{(i)T} Q^{-1} \mathbf{b}_{t+1}^{(i)} + \Sigma_{t|t}^{(i)-1} \hat{\mathbf{x}}_{t|t}^{(i)}. \quad (\text{U.66})$$

Relating individually leads to

$$\begin{aligned} & \begin{bmatrix} \Sigma_{t|\tau}^{(i)} & \Sigma_{t,t+1|\tau}^{(i)} \\ \Sigma_{t+1,t|\tau}^{(i)} & \Sigma_{t+1|\tau}^{(i)} \end{bmatrix} \\ & \quad = \\ & \quad \begin{bmatrix} F_t^{(i)} & F_{t,t+1}^{(i)} \\ F_{t+1,t}^{(i)} & F_{t+1}^{(i)} \end{bmatrix}^{-1} \\ & \quad = \\ & \quad \begin{bmatrix} G_{t,i}^{-1} & -F_t^{(i)-1} F_{t,t+1}^{(i)} G_{t+1,i}^{-1} \\ -F_{t+1}^{(i)-1} F_{t+1,t}^{(i)} G_{t,i}^{-1} & G_{t+1,i}^{-1} \end{bmatrix} \end{aligned} \quad (\text{U.67})$$

where

$$G_{t,i} = F_t^{(i)} - F_{t,t+1}^{(i)} F_{t+1}^{(i)-1} F_{t+1,t}^{(i)}, \quad (\text{U.68})$$

and

$$G_{t+1,i} = F_{t+1}^{(i)} - F_{t+1,t}^{(i)} F_t^{(i)-1} F_{t,t+1}^{(i)}, \quad (\text{U.69})$$

these two maybe rearranged to give

$$\Sigma_{t|\tau}^{(i)} = F_t^{(i)-1} + F_t^{(i)-1} F_{t,t+1}^{(i)} G_{t+1,i}^{-1} F_{t+1,t}^{(i)} F_t^{(i)-1} \quad (\text{U.70})$$

$$\equiv G_{t,i}^{-1}, \quad (\text{U.71})$$

and

$$\Sigma_{t+1|\tau}^{(i)} = F_{t+1}^{(i)-1} + F_{t+1}^{(i)-1} F_{t+1,t}^{(i)} G_{t,i}^{-1} F_{t,t+1}^{(i)} F_{t+1}^{(i)-1}, \quad (\text{U.72})$$

$$\equiv G_{t+1,i}^{-1}. \quad (\text{U.73})$$

Firstly solving for

$$F_t^{(i)-1} = \left(\Sigma_{t|t}^{(i)-1} + A_{t+1}^{(i)T} Q^{-1} A_{t+1}^{(i)} \right)^{-1}, \quad (\text{U.74})$$

I expand this using Equation O.1 then substitute the update $\Sigma_{t+1|t}^{(i)}$ (Equation U.25) in giving the result

$$F_t^{(i)-1} = \Sigma_{t|t}^{(i)} - J_t^{(i)} \Sigma_{t+1|t}^{(i)} J_t^{(i)T}, \quad (\text{U.75})$$

where

$$J_t^{(i)} = \Sigma_{t|t}^{(i)} A_{t+1}^{(i)T} \Sigma_{t+1|t}^{(i)-1}. \quad (\text{U.76})$$

Secondly solving

$$F_t^{(i)-1} F_{t,t+1}^{(i)} = - \left(\Sigma_{t|t}^{(i)-1} + A_{t+1}^{(i)T} Q^{-1} A_{t+1}^{(i)} \right)^{-1} A_{t+1}^{(i)T} Q^{-1}, \quad (\text{U.77})$$

by first using Equation O.2 then replacing the trailing term with the projection $\Sigma_{t+1|t}^{(i)}$ then noting that the result is simply $-J_t^{(i)}$ therefore

$$F_t^{(i)-1} F_{t,t+1}^{(i)} = -J_t^{(i)}. \quad (\text{U.78})$$

These two terms insert into Equation U.71, which gives

$$\Sigma_{t|\tau}^{(i)} = F_t^{(i)-1} + F_t^{(i)-1} F_{t,t+1}^{(i)} G_{t+1,i}^{-1} F_{t+1,t}^{(i)} F_t^{(i)-1} \quad (\text{U.79})$$

noting that $F_{t,t+1}^{(i)} = F_{t+1,t}^{(i)T}$ and substituting $G_{t+1,i}^{-1} = \Sigma_{t+1|t}^{(i)}$ gives

$$\Sigma_{t|\tau}^{(i)} = \Sigma_{t|t}^{(i)} - J_t^{(i)} \Sigma_{t+1|t}^{(i)} J_t^{(i)T} + J_t^{(i)} \Sigma_{t+1|\tau}^{(i)} J_t^{(i)T}, \quad (\text{U.80})$$

$$\equiv \Sigma_{t|t}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1|\tau}^{(i)} - \Sigma_{t+1|t}^{(i)} \right) J_t^{(i)T}. \quad (\text{U.81})$$

The cross covariance update can be seen to be

$$\Sigma_{t+1,t|\tau}^{(i)} = -G_{t+1,i}^{-1} F_{t,t+1}^{(i)} F_t^{(i)-1}, \quad (\text{U.82})$$

substituting in Equation U.73 and the transpose of Equation U.81 gives

$$\Sigma_{t+1,t|\tau}^{(i)} = \Sigma_{t+1|\tau}^{(i)} J_t^{(i)T}, \quad (\text{U.83})$$

this gives

$$\Sigma_{t,t-1|\tau}^{(i)} = \Sigma_{t|\tau}^{(i)} J_{t-1}^{(i)T}, \quad (\text{U.84})$$

expanding these with Equation U.81 gives

$$\Sigma_{t,t-1|\tau}^{(i)} = \left(\Sigma_{t|\tau}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1|\tau}^{(i)} - \Sigma_{t+1|t}^{(i)} \right) J_t^{(i)T} \right) J_{t-1}^{(i)T}, \quad (\text{U.85})$$

substituting in Equation U.83 and using the identity $\Sigma_{t+1|t}^{(i)} J_t^{(i)T} = A_{t+1}^{(i)} \Sigma_{t|t}^{(i)}$ gives

$$\Sigma_{t,t-1|\tau}^{(i)} = \left(\Sigma_{t|\tau}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1,t|\tau}^{(i)} - A_{t+1}^{(i)} \Sigma_{t|t}^{(i)} \right) \right) J_{t-1}^{(i)T}, \quad (\text{U.86})$$

thus defining the backwards recursive update for the cross covariance terms. This recursion is initialised with

$$\Sigma_{\tau,\tau-1|\tau}^{(i)} = \Sigma_{\tau|\tau}^{(i)} J_{\tau-1}^{(i)T}, \quad (\text{U.87})$$

$$= \left(I - K_{\tau}^{(i)} B_{\tau}^{(i)} \right) \Sigma_{\tau|\tau-1}^{(i)} J_{\tau-1}^{(i)T}, \quad (\text{U.88})$$

$$= \left(I - K_{\tau}^{(i)} B_{\tau}^{(i)} \right) A_{\tau}^{(i)} \Sigma_{\tau-1|\tau-1}^{(i)}, \quad (\text{U.89})$$

where $\Sigma_{\tau|\tau}^{(i)}$ is expanded using Equation U.43 and $J_{\tau-1}^{(i)}$ and finally cancel the $\Sigma_{\tau|\tau-1}^{(i)}$ terms.

This gives the initialisation to the cross-covariance backwards recursion.

Finally constructing the mean from

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = -F_t^{(i)-1} F_{t,t+1}^{(i)} \hat{\mathbf{z}}_{t+1|\tau}^{(i)} + F_t^{(i)-1} \left(\Sigma_{t|t}^{(i)-1} \hat{\mathbf{z}}_{t|t}^{(i)} - A_{t+1}^{(i)T} Q^{-1} \mathbf{b}_{t+1}^{(i)} \right), \quad (\text{U.90})$$

substituting in Equation U.78 gives

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = J_t^{(i)} \hat{\mathbf{z}}_{t+1|\tau}^{(i)} + \left(\Sigma_{t|t}^{(i)} - J_t^{(i)} \Sigma_{t+1|\tau}^{(i)} J_t^{(i)T} \right) \left(\Sigma_{t|t}^{(i)-1} \hat{\mathbf{z}}_{t|t}^{(i)} - A_{t+1}^{(i)T} Q^{-1} \mathbf{b}_{t+1}^{(i)} \right), \quad (\text{U.91})$$

which rearranges to give after applying Equation O.2 the term multiplies $\mathbf{b}_{t+1}^{(i)}$ to give $J_t^{(i)}$ and expanding the $J_t^{(i)}$'s and substituting the projection $J_{t+1|t}^{(i)}$ to the term multiplying $\hat{\mathbf{z}}_{t|t}^{(i)}$ to give

$$\hat{\mathbf{z}}_{t|\tau} = J_t^{(i)} \hat{\mathbf{z}}_{t+1|\tau}^{(i)} - J_t^{(i)} \mathbf{b}_{t+1}^{(i)} + \left(\Sigma_{t|t}^{(i)} - \Sigma_{t|t}^{(i)} A_{t+1}^{(i)T} \Sigma_{t+1|\tau}^{(i)-1} A_{t+1}^{(i)} \Sigma_{t|t}^{(i)} \right) \Sigma_{t|t}^{(i)-1} \hat{\mathbf{z}}_{t|t}^{(i)}, \quad (\text{U.92})$$

removing those terms that cancel and substituting in $J_t^{(i)}$ in gives

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = J_t^{(i)} \hat{\mathbf{z}}_{t+1|\tau}^{(i)} - J_t^{(i)} \mathbf{b}_{t+1}^{(i)} + \left(I - J_t^{(i)} A_{t+1}^{(i)} \right) \hat{\mathbf{z}}_{t|t}^{(i)}, \quad (\text{U.93})$$

$$\equiv \hat{\mathbf{z}}_{t|t}^{(i)} + J_t^{(i)} \left(\hat{\mathbf{z}}_{t+1|\tau}^{(i)} - A_{t+1}^{(i)} \hat{\mathbf{z}}_{t|t}^{(i)} - \mathbf{b}_{t+1}^{(i)} \right). \quad (\text{U.94})$$

Thus giving the smoothed means update.

Equations U.94, U.81 & U.86 define the parametric smoothing for my RBPF. Therefore in summary the recursive relations for the smoothed mean,

$$\hat{\mathbf{z}}_{t|\tau}^{(i)} = \hat{\mathbf{z}}_{t|t}^{(i)} + J_t^{(i)} \left(\hat{\mathbf{z}}_{t+1|\tau}^{(i)} - A_{t+1}^{(i)} \hat{\mathbf{z}}_{t|t}^{(i)} - \mathbf{b}_{t+1}^{(i)} \right), \quad (\text{U.95})$$

and covariances are given by

$$\Sigma_{t|\tau}^{(i)} = \Sigma_{t|t}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1|\tau}^{(i)} - \Sigma_{t+1|t}^{(i)} \right) J_t^{(i)T}, \quad (\text{U.96})$$

$$\Sigma_{t,t-1|\tau}^{(i)} = \left(\Sigma_{t|\tau}^{(i)} + J_t^{(i)} \left(\Sigma_{t+1,t|\tau}^{(i)} - A_{t+1}^{(i)} \Sigma_{t|t}^{(i)} \right) \right) J_{t-1}^{(i)T}, \quad (\text{U.97})$$

where the cross covariance term is initialised by

$$\Sigma_{\tau,\tau-1|\tau}^{(i)} = \left(I - K_\tau^{(i)} B_\tau^{(i)} \right) A_\tau^{(i)} \Sigma_{\tau-1|\tau-1}^{(i)}, \quad (\text{U.98})$$

and

$$J_t^{(i)} = \Sigma_{t|t}^{(i)} A_{t+1}^{(i)T} \Sigma_{t+1|t}^{(i)-1}. \quad (\text{U.99})$$

If the (i) superscripts were neglected then these equations would define the Kalman smoothing relations. To estimate the state z_t given a state trajectory $k_{0:\tau}^{(i)}$ and a set of measure-

ments $\mathbf{x}_{0:\tau}$ where $t < \tau$, firstly a ‘forward’ process must be computed, then a ‘backward’ smoothing process is computed to construct the complete data posterior. This produces a better estimate of \mathbf{x}_t as the smoothing case takes advantage of measurements both before and after t , by contrast the filtering case only considers measurements upto t . So the joint smoothed belief is defined completely by Equations U.94, U.81 & U.86 to define the density,

$$p\left(\mathbf{z}_{t-1:t} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}\right) = \phi\left(\begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{z}_t \end{bmatrix}; \begin{bmatrix} \hat{\mathbf{z}}_{t-1|\tau}^{(i)} \\ \hat{\mathbf{z}}_{t|\tau}^{(i)} \end{bmatrix}, \begin{bmatrix} \Sigma_{t-1|\tau}^{(i)} & \Sigma_{t-1,t|\tau}^{(i)} \\ \Sigma_{t,t-1|\tau}^{(i)} & \Sigma_{t|\tau}^{(i)} \end{bmatrix}\right), \quad (\text{U.100})$$

where $\Sigma_{t,t-1|\tau}^{(i)} = \Sigma_{t-1,t|\tau}^{(i)T}$ as the covariance of $\begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{z}_t \end{bmatrix}$ is by definition positive definite.

Appendix V

Learning updates

I will now proceed to estimate the parameters of the RBPF model using the EM algorithm. Much of the state space model (SSM) updates are derived similarly Welling (2008). I consider the states \mathbf{z}_t and k_t as hidden variables while $\mathbf{x}_{0:\tau}$ are the observations. I assume I have observed one sequence of length τ . Starting from the joint belief of the complete data

$$p(\mathbf{z}_{0:\tau}, k_{0:\tau}, j_{0:\tau}, \mathbf{x}_{0:\tau} | a_{0:\tau}) = p(\mathbf{z}_0) P(k_0) \left[\prod_{t=1}^{\tau} p(\mathbf{z}_t | \mathbf{z}_{t-1}, k_t) P(k_t | k_{t-1}) \right] \times \left[\prod_{t=0}^{\tau} p(\mathbf{x}_t | \mathbf{z}_t, j_t) P(j_t | a_t) \right], \quad (\text{V.1})$$

$$= \phi(\mathbf{z}_0; \boldsymbol{\mu}_0, \Sigma_0) \pi_{k_0} \left[\prod_{t=1}^{\tau} \phi(\mathbf{z}_t; A^{(k_t)} \mathbf{z}_{t-1} + \mathbf{b}^{(k_t)}, Q) \chi_{k_t | k_{t-1}} \right] \times \left[\prod_{t=0}^{\tau} \phi(\mathbf{x}_t; B^{(j_t)} \mathbf{z}_t, R) \zeta_{j_t | a_t} \right]. \quad (\text{V.2})$$

The quantity to optimise is the bound upon the complete data log-likelihood

$$\begin{aligned} B[\Omega; \Omega^{\text{old}}] &= \int dz^\tau \sum_{j^\tau, k^\tau} p^{\text{old}}(\mathbf{z}_{0:\tau}, k_{0:\tau}, j_{0:\tau} | \mathbf{x}_{0:\tau}, a_{0:\tau}) \log p(\mathbf{z}_{0:\tau}, k_{0:\tau}, j_{0:\tau}, \mathbf{x}_{0:\tau} | a_{0:\tau}), \quad (\text{V.3}) \\ &= \sum_{i=1}^N \frac{1}{N} \int d\mathbf{z}_{0:\tau} p^{\text{old}}(\mathbf{z}_{0:\tau} | k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau}) \left\{ -\frac{1}{2} D_z(\tau+1) \log 2\pi - \frac{1}{2} D_x(\tau+1) \right. \\ &\quad - \frac{1}{2} \log |\Sigma| - \frac{1}{2} \tau \log |Q| - \frac{1}{2} (\tau+1) \log |R| \\ &\quad - \frac{1}{2} (\mathbf{z}_0 - \boldsymbol{\mu}_0)^T \Sigma^{-1} (\mathbf{z}_0 - \boldsymbol{\mu}_0) \\ &\quad - \frac{1}{2} \sum_{t=1}^{\tau} \left(\mathbf{z}_t - A^{(k_t^{(i)})} \mathbf{z}_{t-1} - \mathbf{b}^{(k_t^{(i)})} \right)^T Q^{-1} \left(\mathbf{z}_t - A^{(k_t^{(i)})} \mathbf{z}_{t-1} - \mathbf{b}^{(k_t^{(i)})} \right) \\ &\quad \left. - \frac{1}{2} \sum_{t=0}^{\tau} \left(\mathbf{x}_t - B^{(j_t^{(i)})} \mathbf{z}_t \right)^T R^{-1} \left(\mathbf{x}_t - B^{(j_t^{(i)})} \mathbf{z}_t \right) \right\} \end{aligned}$$

$$\left. + \log \pi_{k_0^{(i)}} + \sum_{t=1}^{\tau} \log \chi_{k_t^{(i)}|k_{t-1}^{(i)}} + \sum_{t=0}^{\tau} \log \zeta_{j_t^{(i)}|a_t} \right\}. \quad (\text{V.4})$$

Inspecting the bound reveals that the only sufficient statistics which need be calculated in the E-step are

$$\mathbb{E}_{p(\mathbf{z}_t|k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} [\mathbf{z}_t] = \hat{\mathbf{z}}_{t|\tau}^{(i)}, \quad (\text{V.5})$$

$$\mathbb{E}_{p(\mathbf{z}_t|k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} [\mathbf{z}_t \mathbf{z}_t^T] = \Sigma_{t|\tau}^{(i)} + \hat{\mathbf{z}}_{t|\tau}^{(i)} \hat{\mathbf{z}}_{t|\tau}^{(i)T}, \quad (\text{V.6})$$

$$\triangleq M_t^{(i)}, \quad (\text{V.7})$$

$$\mathbb{E}_{p(\mathbf{z}_t|k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} [\mathbf{z}_t \mathbf{z}_{t-1}^T] = \Sigma_{t,t-1|\tau}^{(i)} + \hat{\mathbf{z}}_{t|\tau}^{(i)} \hat{\mathbf{z}}_{t-1|\tau}^{(i)T}, \quad (\text{V.8})$$

$$\triangleq M_{t,t-1}^{(i)}, \quad (\text{V.9})$$

further, as I can say $\mathbf{x}^T \mathbf{y} \triangleq \text{tr} [\mathbf{y} \mathbf{x}^T]$ from Appendix P, then

$$\mathbb{E}_{p(\mathbf{z}_t|k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} [\mathbf{z}_t^T D \mathbf{z}_t] = \text{tr} \left[D \mathbb{E}_{p(\mathbf{z}_t|k_{0:\tau}^{(i)}, j_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau})} [\mathbf{z}_t \mathbf{z}_t^T] \right], \quad (\text{V.10})$$

$$= \text{tr} [D M_t^{(i)}], \quad (\text{V.11})$$

together these sufficient statistics determine an analytic solution to the integral in the bound. The later equations are easier to compute in this form than any other due to the notational conventions of python and Scipy which were the tools I used to implement these algorithms.

Starting by constructing the Lagrangian multipliers for the parameters defining the dynamical evolution for the continuous state. For $d(A^{(k)}, \bar{A}^{(k)})$, using the KL-divergence between two Gaussian measures, Equation R.13, and setting $\boldsymbol{\mu}_p = A^{(k)} \boldsymbol{\alpha}$ and $\boldsymbol{\mu}_q = \bar{A}^{(k)} \boldsymbol{\alpha}$ with $\boldsymbol{\alpha} \boldsymbol{\alpha}^T = I_{D_x}$ and $\Sigma_q = Q$ gives a relation similar to Equation R.15, which values the gradient of the Lagrangian,

$$\frac{\partial G}{\partial A^{(k)}} = -Q^{-1} \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(M_{t,t-1}^{(i)} - A^{(k)} M_{t-1}^{(i)} - \mathbf{b}^{(k)} \hat{\mathbf{z}}_{t-1|\tau}^{(i)T} \right) - \gamma_{\sigma} Q^{-1} \left(\bar{A}^{(k)} - A^{(k)} \right), \quad (\text{V.12})$$

which produces the relation

$$A^{(k)} \left[\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} M_{t-1}^{(i)} + \gamma_{\sigma} I_{D_x} \right] = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(M_{t,t-1}^{(i)} - \mathbf{b}^{(k)} \hat{\mathbf{z}}_{t-1|t}^{(i)T} \right) + \gamma_{\sigma} \bar{A}^{(k)}. \quad (\text{V.13})$$

Similarly for \mathbf{b}_k using Equation R.13, where inserting into Equation R.15 $\boldsymbol{\mu}_q = \mathbf{b}^{(k)}$, $\boldsymbol{\mu}_q = \bar{\mathbf{b}}^{(k)}$ and $\Sigma_q = Q$, the Lagrangian gradient is

$$\frac{\partial G}{\partial \mathbf{b}^{(k)}} = -Q^{-1} \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(\hat{\mathbf{z}}_{t|t}^{(i)} - A^{(k)} \hat{\mathbf{z}}_{t-1|t}^{(i)} - \mathbf{b}^{(k)} \right) - \gamma_{\sigma} Q^{-1} \left(\bar{\mathbf{b}}^{(k)} - \mathbf{b}^{(k)} \right), \quad (\text{V.14})$$

which produces the relation

$$\mathbf{b}^{(k)} \left[\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} + \gamma_{\sigma} \right] = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(\hat{\mathbf{x}}_{t|t}^{(i)} - A^{(k)} \hat{\mathbf{z}}_{t-1|t}^{(i)} \right) + \gamma_{\sigma} \bar{\mathbf{b}}^{(k)}. \quad (\text{V.15})$$

Then solving for A_k by inserting \mathbf{b}_k into Equation V.13 gives

$$\begin{aligned} A^{(k)} &= \left[\left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} M_{t,t-1}^{(i)} + \gamma_{\sigma} \bar{A}^{(k)} \right) \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} + \gamma_{\sigma} \right) - \right. \\ &\quad \left. \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \hat{\mathbf{x}}_{t|t}^{(i)} + \gamma_{\sigma} \bar{\mathbf{b}}^{(k)} \right) \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \hat{\mathbf{z}}_{t-1|t}^{(i)T} \right)^T \right] \\ &\quad \left[\left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} M_{t-1}^{(i)} + \gamma_{\sigma} I_{D_x} \right) \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} + \gamma_{\sigma} \right) - \right. \\ &\quad \left. \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \hat{\mathbf{z}}_{t-1|t}^{(i)} + \gamma_{\sigma} \bar{\mathbf{b}}^{(k)} \right) \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \hat{\mathbf{z}}_{t-1|t}^{(i)T} \right)^T \right]^{-1} \quad (\text{V.16}) \end{aligned}$$

and,

$$\mathbf{b}^{(k)} = \frac{1}{\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} + \gamma_{\sigma}} \left(\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(\hat{\mathbf{z}}_{t|t}^{(i)} - A^{(k)} \hat{\mathbf{z}}_{t-1|t}^{(i)} \right) + \gamma_{\sigma} \bar{\mathbf{b}}^{(k)} \right). \quad (\text{V.17})$$

These updates trivialise to the non-episodic EM updates when both $\gamma_{\sigma} = 0$. These updates are similar to a least square estimate of the dynamic process with a prior belief in the parameter.

Constructing the state evolution noise, where $\Sigma_q = \sigma I_{D_x}$ and $\Sigma_p = \bar{\sigma} I_{D_x}$, gives

$$\begin{aligned} \frac{\partial G}{\partial \sigma^{-1}} &= -\frac{1}{2N} \sum_{i=1}^N \int d\mathbf{z}_{0:\tau} p^{\text{old}} \left(\mathbf{z}_{0:\tau} | j_{0:\tau}^{(i)}, k_{0:\tau}^{(i)}, \mathbf{x}_{0:\tau} \right) \left\{ -\tau D_z \sigma + \sum_{t=1}^{\tau} \left(\mathbf{z}_t - A^{(k_t^{(i)})} \mathbf{z}_{t-1} - \mathbf{b}^{(k_t^{(i)})} \right)^2 \right\} \\ &\quad - \frac{\gamma_{\sigma}}{2} D_z (\bar{\sigma} - \sigma), \quad (\text{V.18}) \end{aligned}$$

which by Appendix P may be expressed as the trace of the expected cross product,

$$\begin{aligned}
\sigma D_z(\tau + \gamma_\sigma) &= \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \text{tr} \left[M_t^{(i)} - A^{(k_t^{(i)})} M_{t-1,t}^{(i)} - \mathbf{b}^{(k_t^{(i)})} \hat{\mathbf{z}}_{t|\tau}^{(i)T} - M_{t,t-1}^{(i)} A^{(k_t^{(i)})T} \right. \\
&\quad + A^{(k_t^{(i)})} M_{t-1}^{(i)} A^{(k_t^{(i)})T} + \mathbf{b}^{(k_t^{(i)})} \hat{\mathbf{z}}_{t-1|\tau}^{(i)T} A^{(k_t^{(i)})T} - \hat{\mathbf{z}}_{t|\tau}^{(i)} \mathbf{b}^{(k_t^{(i)})T} \\
&\quad \left. + A^{(k_t^{(i)})} \hat{\mathbf{z}}_{t-1|\tau}^{(i)} \mathbf{b}^{(k_t^{(i)})T} + \mathbf{b}^{(k_t^{(i)})} \mathbf{b}^{(k_t^{(i)})T} \right] + \gamma_\sigma D_z \bar{\sigma}. \quad (\text{V.19})
\end{aligned}$$

Using the previously derived state evolution equations to simplify this, so the identity

$$\begin{aligned}
&\sum_{i=1}^N \sum_{t=1}^{\tau} \left(\hat{\mathbf{z}}_{t|\tau}^{(i)} - A^{(k_t^{(i)})} \hat{\mathbf{z}}_{t-1|\tau}^{(i)} - \mathbf{b}^{(k_t^{(i)})} \right) \mathbf{b}^{(k_t^{(i)})T} \\
&= \\
&\sum_{k=1}^K \left[\sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(\hat{\mathbf{z}}_{t|\tau}^{(i)} - A^{(k)} \hat{\mathbf{z}}_{t-1|\tau}^{(i)} - \mathbf{b}^{(k)} \right) \right] \mathbf{b}^{(k)T} \quad (\text{V.20}) \\
&= \\
&N \gamma_\sigma \sum_{k=1}^K (\mathbf{b}^{(k)} - \bar{\mathbf{b}}^{(k)}) \mathbf{b}^{(k)T},
\end{aligned}$$

which is a consequence of Equation V.15, and similarly

$$\begin{aligned}
&\sum_{i=1}^N \sum_{t=1}^{\tau} \left(M_{t,t-1}^{(i)} - A^{(k_t^{(i)})} M_{t-1}^{(i)} - \mathbf{b}_{k_t^{(i)}} \hat{\mathbf{z}}_{t-1|\tau}^{(i)T} \right) A^{(k_t^{(i)})T} \\
&= \\
&\sum_{k=1}^K \left[\sum_{i=1}^N \sum_{t=1}^{\tau} \delta_k^{k_t^{(i)}} \left(M_{t,t-1}^{(i)} - A^{(k)} M_{t-1}^{(i)} - \mathbf{b}^{(k)} \hat{\mathbf{z}}_{t-1|\tau}^{(i)T} \right) \right] A^{(k)T} \quad (\text{V.21}) \\
&= \\
&N \gamma_\sigma \sum_{k=1}^K (A^{(k)} - \bar{A}^{(k)}) A^{(k)T},
\end{aligned}$$

which is a consequence of Equation V.13. Inserting these identities (Equations V.20 & V.21) into the update equation Equation V.19 gives

$$\begin{aligned}
\sigma &= \frac{1}{D_z(\tau + \gamma_\sigma)} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \text{tr} \left[M_t^{(i)} - A^{(k_t^{(i)})} M_{t-1,t}^{(i)} - \mathbf{b}^{(k_t^{(i)})} \hat{\mathbf{z}}_{t|\tau}^{(i)T} \right] \right. \\
&\quad + \gamma_\sigma \sum_{k=1}^K \text{tr} \left[(\bar{\mathbf{b}}^{(k)} - \mathbf{b}^{(k)}) \mathbf{b}^{(k)T} + (\bar{A}^{(k)} - A^{(k)}) A^{(k)} \right] \\
&\quad \left. + \gamma_\sigma D_z \bar{\sigma} \right\}. \quad (\text{V.22})
\end{aligned}$$

Thus the state evolutions noise is dependent upon both the prior episodes state evolution dynamics and noise.

To generate the measurement noise where $R = \nu I_{D_x}$ and the Lagrangian multiplier has

$\Sigma_p = \nu I_{D_x}$ and $\Sigma_q = \bar{\nu} I_{D_x}$, then, after computing the expectation of the gradient with respect to ν gives,

$$\begin{aligned} \frac{\partial G}{\partial \nu^{-1}} = & -\frac{1}{2N} \sum_{i=1}^N \left\{ -D_y (\tau + 1) \nu + \sum_{t=0}^{\tau} \text{tr} \left[\mathbf{x}_t \mathbf{x}_t^T - \mathbf{x}_t \hat{\mathbf{z}}_{t|\tau}^{(i)T} B^{(j_t^{(i)})T} \right. \right. \\ & \left. \left. - B^{(j_t^{(i)})} \hat{\mathbf{z}}_{t|\tau}^{(i)} \mathbf{x}_t^T + B^{(j_t^{(i)})} M_t^{(i)} B^{(j_t^{(i)})T} \right] \right\} \\ & - \frac{\gamma \nu}{2} D_x (\bar{\nu} - \nu). \end{aligned} \quad (\text{V.23})$$

Solving Equation V.23 with respect to ν gives the stationary point of G ,

$$\begin{aligned} \nu = & \frac{1}{D_x (\tau + 1 + \gamma \nu)} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{\tau} \text{tr} \left[\mathbf{x}_t \mathbf{x}_t^T - \mathbf{x}_t \hat{\mathbf{z}}_{t|\tau}^{(i)T} B^{(j_t^{(i)})T} \right. \right. \\ & \left. \left. - B^{(j_t^{(i)})} \hat{\mathbf{z}}_{t|\tau}^{(i)} \mathbf{x}_t^T + B^{(j_t^{(i)})} M_t^{(i)} B^{(j_t^{(i)})T} \right] \right. \\ & \left. + \gamma \nu D_x \bar{\nu} \right\}. \end{aligned} \quad (\text{V.24})$$

Hence Equation V.24 forms the episodic EM update ν , this is equivalent to the variance between the expected residuals for each trajectory and a weighted contribution from the prior episodes ν of $\bar{\nu}$.

To generate the symbolic dynamic transition matrix χ update I use a second Lagrangian to the episodic Lagrangian to give,

$$G \left[\chi; \Omega^{\text{old}} | \bar{\Omega} \right] \propto \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \log \chi_{k_t^{(i)} | k_{t-1}^{(i)}} + \gamma \chi d(\chi; \bar{\chi}) - \sum_{k'} \lambda_{k'}^{(\chi)} \left(\sum_k \chi_{k|k'} - 1 \right), \quad (\text{V.25})$$

Taking the gradient of this equation with respect to the elements of χ gives

$$\frac{\partial G}{\partial \chi_{l|m}} = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \frac{1}{\chi_{l|m}} \delta_l^{k_t^{(i)}} \delta_m^{k_{t-1}^{(i)}} - \lambda_m^{(\chi)} + \gamma \chi \frac{\bar{\chi}_{l|m}}{\chi_{l|m}}. \quad (\text{V.26})$$

Equating Equation V.30 to zero and rearranging to make the elements of χ the subject gives

$$\chi_{l|m} = \frac{1}{\lambda_m^{(\chi)}} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_l^{k_t^{(i)}} \delta_m^{k_{t-1}^{(i)}} + \gamma \chi \bar{\chi}_{l|m} \right\}. \quad (\text{V.27})$$

Finally using the condition $\sum_l \chi_{l|m} = 1$ to remove the Lagrangian multiplier λ_1 gives the

episodic EM update,

$$\chi_{l|m} = \frac{1}{\frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_m^{k_{t-1}^{(i)}} + \gamma_{\chi}} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^{\tau} \delta_l^{k_t^{(i)}} \delta_m^{k_{t-1}^{(i)}} + \gamma_{\chi} \bar{\chi}_{l|m} \right\}. \quad (\text{V.28})$$

Equation V.32 is simply the normalised tabulation of the switching transitions from time-step to time-step across the trajectories $k_{0:\tau}^{(i)}$ with the weighted addition of the prior episodes transitions $\bar{\chi}$.

To generate the belief matrix $\zeta_{j|a}$ for the believed measure j for action a update I use a second Lagrangian to the episodic Lagrangian to give

$$G \left[\zeta; \Omega^{\text{old}} | \bar{\Omega} \right] \propto \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{\tau} \log \zeta_{j_t^{(i)} | a_t} + \gamma_{\zeta} d(\zeta; \bar{\zeta}) - \sum_a \lambda_a^{(\zeta)} \left(\sum_j \chi_{j|a} - 1 \right), \quad (\text{V.29})$$

Taking the gradient of this equation with respect to the elements of χ gives

$$\frac{\partial G}{\partial \zeta_{l|m}} = \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{\tau} \frac{1}{\chi \zeta_{l|m}} \delta_l^{j_t^{(i)}} \delta_m^{a_t} - \lambda_m^{(\zeta)} + \gamma_{\zeta} \frac{\bar{\zeta}_{l|m}}{\zeta_{l|m}}. \quad (\text{V.30})$$

Equating Equation V.30 to zero and rearranging to make the elements of χ the subject gives

$$\zeta_{l|m} = \frac{1}{\lambda_m^{(\zeta)}} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{\tau} \delta_l^{j_t^{(i)}} \delta_m^{a_t} + \gamma_{\zeta} \bar{\zeta}_{l|m} \right\}. \quad (\text{V.31})$$

Finally using the condition $\sum_l \zeta_{l|m} = 1$ to remove the Lagrangian multiplier λ_1 gives the episodic EM update,

$$\zeta_{l|m} = \frac{1}{\frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{\tau} \delta_m^{a_t} + \gamma_{\zeta}} \left\{ \frac{1}{N} \sum_{i=1}^N \sum_{t=0}^{\tau} \delta_l^{j_t^{(i)}} \delta_m^{a_t} + \gamma_{\zeta} \bar{\zeta}_{l|m} \right\}. \quad (\text{V.32})$$

Equation V.32 is simply the normalised tabulation of the trajectories $j_{0:\tau}^{(i)}$ and the chosen actions $a_{0:\tau}$ with the weighted addition of the prior episodes correspondence $\bar{\zeta}$.

The episodic EM algorithms Lagrangian multiplier is interpreted in a similar context to that of the Gaussian distribution in the main text, Subsection R.4. So the dynamic processes γ 's are set to $\gamma_{\sigma} = \gamma_{\chi} = \gamma_D (\tau - 1)$ and the measurement and correspondences are set to $\gamma_{\nu} = \gamma_{\zeta} = \gamma_M \tau$.

Bibliography

- H. D. I. Abarbanel, R. Brown, and J. B. Kadtke. Prediction in chaotic nonlinear systems: Methods for time series with broadband fourier spectra. *Phys. Rev. A*, 41:1782–1807, 1990.
- K. Aizawa. Representations without rules, connectionism and the syntactic argument. *Synthese*, 101:465–492, 1994.
- N. Y. Alexeenko and N. N. Verderevskaya. Proprioceptive effects on evoked responses to sounds in the cat auditory cortex. *Experimental Brain Research*, 26:495–508, 1976.
- J. D. Altringham. *Bats; biology and behaviour*. Oxford University Press, 2001.
- P. Arno, A. Vanlierde, E. Streel, M.-C. Wanet-Defalque, S. Sanabria-Bohorquez, and C. Veraart. Auditory substitution of vision: pattern recognition by the blind. *Applied Cognitive Psychology*, 15:509–519, 2001.
- D. H. Ashmead, D. L. Davis, and A. Northington. Contribution of listeners' approaching motion to auditory distance perception. *Journal of Experimental Psychology*, 21(2): 239–256, 1995.
- H. Attias, L. Deng, A. Acero, and J. C. Platt. A new method for speech denoising and robust speech recognition using probabilistic models for clean speech and for noise. In *EUROSPEECH-2001*, pages 1903–1906, 2001.
- M. Aytikin, E. Grassi, M. Sahota, and C. F. Moss. The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation. *Journal of Acoustical Society of America*, 116:3594–3605, 2004.
- M. Aytikin, C. F. Moss, and J. Z. Simon. A sensorimotor approach to sound localization. *Neural Computation*, 20:603–635, 2008.

- T. Baker. *Mechanisms in Insect Olfaction*, chapter Pheromone-modulated movements of flying moths, page 39–48. Oxford University Press, 1986.
- E. Balkovsky and B. I. Shraiman. Olfactory search at high Reynolds number. *PNAS*, 99:12589–12593, 2002.
- E. E. Bauer, A. Klug, and G. D. Pollak. Spectral determination of responses to species-specific calls in the dorsal nucleus of the lateral lemniscus. *J. Neurophysiol.*, 88:1955–1967, 2002.
- M. J. Beal, N. Jojic, and H. Attias. A graphical model for audiovisual object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:828–836, 2003.
- C. M. Bishop. *Neural networks for pattern recognition*. Oxford, 1995.
- C. M. Bishop, M. Svensen, and C. K. I. Williams. Gtm: The generative topographic mapping. *Neural Computation*, 10:215–234, 1998a.
- C. M. Bishop, M. Svensen, and C. K. I. Williams. Developments of the generative topographic mapping. *Neurocomputing*, 21:203–224, 1998b.
- S.-J. Blakemore, C. D. Frith, and D. M. Wolpert. The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*, 12:1879–1884, 2001.
- S.-J. Blakemore, D. M. Wolpert, and C. D. Frith. Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6:237–242, 2002.
- J. Blauert. *Spatial hearing*. MIT-Press, revised edition edition, 1997.
- A. Bompas and K. J. O’Regan. More evidence for sensorimotor adaptation in color perception. *Journal of Vision*, 6:145–153, 2006.
- J. Bongard, V. Zykov, and H. Lipson. Resilient machines through continuous self-modeling. *Science*, 17:1118–1121, 2006.
- J. C. Bongard and H. Lipson. Nonlinear system identification using coevolution of models and tests. *IEEE Transactions on Evolutionary Computation*, 9:361–384, 2005.
- J. Borenstein and Y. Koren. The vector field histogram-fast obstacle avoidance for mobile robots. *Robotics and Automation*, 7:278–288, 1991.

- N. Brenner, W. Bialek, and R. de Ruyter van Steveninck. Adaptive rescaling maximizes information transmission. *Neuron*, 26:695–700, 2000.
- A. J. Briggs and B. R. Donald. Automatic sensor configuration for task-directed planning:. In *IEEE International Conference on Robotics and Automation*, 1994.
- W. Burgard, A. B. Cremersa, D. Fox, D. Hahnela, G. Lakemeyerc, D. Schulza, W. Steinera, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114:3–55, 1999a.
- W. Burgard, D. Fox, H. Jans, C. Matenar, and S. Thrun. Sonar-based mapping with mobile robots using em. In *Proceedings of the International Conference on Machine Learning*, 1999b.
- W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun. Collaborative multi-robot exploration. In *International Conference on Robotics and Automation*, 2000.
- R. M. Burger and G. D. Pollak. Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role in the processing of multiple sound sources in the inferior colliculus of bats. *J. Neurosci.*, 21:4830–4843, 2001.
- J. P. L. Cadre and H. Gauvrit. Optimization of the observer motion for bearings-only target motionanalysis. In *First Australian Data Fusion Symposium*, 1996.
- S. Carlile and D. Pralong. The location-dependent nature of perceptually salient features of the human head-related transfer functions. *Journal of the Acoustical Society of America*, 95:3445–3459, 1994.
- A. R. Cassandra, L. P. Kaelbling, and J. A. Kurien. Acting under uncertainty: discrete bayesian models for mobile-robot navigation. In *International Conference on Intelligent Robots and Systems*, 1996.
- R. Chatila and J. Laumond. Position referencing and consistent world modeling for mobile robots. In *International Conference on Robotics and Automation*, 1985.
- H. Choset. *Sensor based motion planning: the hierarchical generalised Voronoi graph*. PhD thesis, California Institute of Technology, 1996.
- J. J. Clark and J. K. O’Regan. A temporal-difference model of perceptual stability in color vision. In *15th International Conference on Pattern Recognition*, 2000.

- C. J. Condon, A. Galazyuk, K. R. White, and A. S. Feng. Neurons in the auditory cortex of the little brown bat exhibit selectivity for complex amplitude-modulated signals that mimic echoes from fluttering insects. *Auditory Neurosci.*, 3:269–287, 1997.
- R. D. Cook. Assessment of local influence. *Journal of the Royal Statistical Society: Series B*, 48:133–169, 1986.
- T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, second edition edition, 2006.
- E. Covey, M. Vater, and J. H. Casseday. Binaural properties of single units in the superior olivary complex of the mustached bat. *J. Neurophysiol.*, 66:1080–1094, 1991.
- M. Cowart. Embodied cognition. The internet encyclopedia of philosophy, 2006. URL <http://www.iep.utm.edu/e/embodcog.htm>.
- A. Dearden and Y. Demiris. Learning forward models for robots. In *International joint conference on Artificial Intelligence*, volume 19, pages 1440–1445, 2005.
- T. B. DeMarse and K. P. Dockendorf. Adaptive flight control with living neuronal networks on microelectrode arrays. In *International Joint Conference on Neural Networks*, 2005.
- T. B. DeMarse, D. A. Wagenaar, A. W. Blau, and S. M. Potter. The neurally controlled animat: Biological brains acting with simulated bodies. *Autonomous Robots*, 11:305–310, 2001.
- W. L. Ditto, S. N. Rauseo, and M. L. Spano. Experimental control of chaos. *Phys. Rev. Lett.*, 65:3211–3214, 1990.
- P. DiZio, R. Held, J. R. Lackner, B. Shinn-Cunningham, and N. Durlach. Gravitoinertial force magnitude and direction influence head-centric auditory localization. *Journal of Neurophysiology*, 85:2455–2460, 2001.
- A. Doucet, J. F. G. de Freitas, K. Murphy, and S. Russell. Rao blackwellised particle filtering for dynamic bayesian networks. In *Uncertainty in Artificial Intelligence*, 2000.
- A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer, 2001.
- U. Dressler and G. Nitsche. Controlling chaos using time delay coordinates. *Phys. Rev. Lett.*, 68:1–4, 1992.

- J. Droulez and A. Berthoz. A neural network model of sensoritopic maps with predictive short-term memory properties. *PNAS*, 88:9653–9657, 1991.
- O. Duda. *Binaural and spatial hearing in real and virtual environments*, chapter 3, pages 49–75. Lawrence Erlbaum Associates, 1997. Elevation dependence of the interaural transfer function.
- H. Durrant-Whyte, S. Majumder, S. Thrun, M. de Battista, and S. Scheduling. A bayesian algorithm for simultaneous localisation and map building. In *Robotics Research*, chapter A Bayesian Algorithm for Simultaneous Localisation and Map Building, pages 49–60. Springer, 2003.
- H. H. Ehrsson. The experimental induction of out-of-body experiences. *Science*, 24:1048, 2007.
- A. Elfes. Sonar-based real-world mapping and navigation. *Robotics and Automation*, 3: 249–265, 1987.
- A. Elfes. *Occupancy grids: a probabilistic framework for robot perception and navigation*. PhD thesis, Carnegie Mellon University, 1989.
- P. H. Ewert. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior. *Genetic Psychology Monographs*, 7:177–363, 1930.
- J. D. Farmer and J. J. Sidorowich. Predicting chaotic time series. *Phys. Rev. Lett.*, 59: 845–848, 1987.
- I. D. Forsythe. Auditory processing. In *Encyclopedia of Life Sciences*. John Wiley & Sons, 2002.
- D. Fox, W. Burgard, and S. Thrun. Active markov localization for mobile robots. *Robotics and Autonomous Systems*, 25:195–207, 1998.
- B. Frey and N. Jojic. Fast, large-scale transformation-invariant clustering. In *Advances in Neural Information Processing Systems*, number 14, 2001.
- B. J. Frey and N. Jojic. Transformation-invariant clustering using the em algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1:1–17, 2003.

- Z. M. Fuzessery and T. D. Lohuis. *Echolocation in bats and dolphins*, chapter Sensitivity to interaural time differences in the sound envelope in the inferior colliculus and auditory cortex of the pallid bat, pages 129–136. Chicago, 2003.
- Z. M. Fuzessery and G. D. Pollak. Determinants of sound location selectivity in bat inferior colliculus: a combined dichotic and free-field stimulation study. *Journal of Neurophysiology*, 54:757–781, 1985.
- V. Gallese and C. Keysers. Mirror neurons: A sensorimotor representation system. *Behavioral and Brain Sciences*, 24:983–984, 2001. Commenatry/O’Regan & Noe: A sensorimotor account of vision and visual consciousness.
- M. Geier, O. J. Bosch, and J. Boeckh. Influence of odour plume structure on upwind flight of mosquitoes towards hosts. *Journal of Experimental Biology*, 202:1639–1648, 1999.
- S. Getzmann. The effect of eye position and background noise on vertical sound localization. *Hearing Research*, 169:130–139, 2002.
- Z. Ghahramani. Learning dynamic bayesian networks. In *Adaptive Processing of Sequences and Data Structures*, pages 168–197. Springer-Verlag, 1998.
- Z. Ghahramani and M. J. Beal. Propagation algorithms for variational bayesian learning. In *Advances in Neural Information Processing Systems*, 13, 2000.
- Z. Ghahramani and G. E. Hinton. Variational learning for switched state-space models. *Neural Computation*, 12:831–864, 2000.
- J. J. Gibson. The ecological approach to the visual perception of pictures. *Leonardo*, 11: 227–235, 1978.
- E. Givelberg and J. Bunn. A comprehensive three-dimensional model of the cochlea. *J. Comput. Physics*, 191:377–391, 2003.
- H. H. Gonzalez-Banos and J.-C. Latombe. Navigation strategies for exploring indoor environments. *International Journal of Robotics Research*, 21:10–11, 2002.
- H. H. L. M. Goossens and A. J. van Opstal. Influence of head position on the spatial representation of acoustic targets. *Journal of Neurophysiology*, 81:2720–2736, 1999.
- M. S. A. Graziano, X. T. Hu, and C. G. Gross. Coding the locations of objects in the dark. *Science*, 11:239–241, 1997.

- M. S. Grewal and A. P. Andrews. *Kalman filtering: Theory and practice using MATLAB*. Wiley, 2001.
- J. M. Groh, A. S. Trause, A. M. Underhill, K. R. Clark, and S. Inati. Eye position influences auditory responses in primate inferior colliculus. *Neuron*, 29:509–518, 2001.
- B. Grothe. Interaction of excitation and inhibition in processing of pure tone and amplitude-modulated stimuli in the medial superior olive of the mustached bat. *J. Neurophysiol.*, 2:706–721, 1994.
- B. Grothe. New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience*, 4:540–550, 2003.
- N. M. Grzywacz and R. M. Balboa. A bayesian framework for sensory adaptation. *Neural Computation*, 14:543–559, 2002.
- A. A. Handzel and P. S. Krishnaprasad. Biometric sound-source localization. *IEEE Sensors Journal*, 2:607–616, 2002.
- S. Harnad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.
- G. Harnischfeger, G. Neuweiler, and P. Schlegel. Interaural time and intensity coding in the superior olivary complex and inferior colliculus of the echolocating bat molossus ater. *J. Neurophysiol.*, 53:89–109, 1985.
- W. Hartmann. How we localize sound. *Physics Today*, 52:24–29, 1999.
- T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2001.
- R. Held and A. Hein. Movement produced stimulation in the development of visually guided behaviour. *Journal of Comparative Physiology*, 56:872–876, 1963.
- J. P. Helferty and D. R. Mudgett. Optimal observer trajectories for bearings-only tracking by minimizing the trace of the cramer-rao lower bound. In *Conference on Decision and Control*, 1993.
- M. L. Hernandez. Optimal sensor trajectories in bearings-only tracking. In *Proceedings of the Seventh International Conference on Information Fusion*, pages 893–900, 2004.

- L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, , M. Saleh, A. H. Caplan, A. Branner, D. Chen, , R. D. Pen, and J. P. Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442(7099):164–171, 2006.
- P. M. Hofman, J. G. A. van Riswick, and A. J. van Opstal. Relearning sound localization with new ears. *Nature Neuroscience*, 1:417–421, 1998.
- P. M. Hofman, M. S. M. G. Valming, P. J. J. Termeer, , and A. J. van Opstal. A method to induce swapped binaural hearing. *Journal of Neuroscience Methods*, 113:167–179, 2002.
- T. Hospedales and S. Vijayakumar. Bayesian multisensory perception. Informatics EDI-INF-RR0829, Edinburgh University, 2006.
- T. M. Hospedales, J. J. Cartwright, and S. Vijayakumar. Structure inference for bayesian multisensory perception and tracking. In *International Joint Conferences on Artificial Intelligence*, pages 2122–2128, 2007.
- R. F. Huffman and E. Covey. Origin of ascending projections to the nuclei of the lateral lemniscus in the big brown bat, *ptesicus fuscus*. *J. Comp. Neurol.*, 357:532–545, 1995.
- W. IJsselstein and M. Reiner. On the importance of reliable real-time sensorimotor dependencies for establishing telepresence. In *PRESENCE*, 2004.
- M. Isard and A. Blake. Condensation - conditional density propagation for visual tracking. *International Journal of Machine Vision*, 29:2–28, 1998.
- M. F. Jay and D. L. Sparks. Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*, 309:345–347, 1984.
- R. L. Jenison. On acoustic information for motion. *Ecological Psychology*, 9:131–151, 1997.
- R. L. Jenison, R. A. Reale, J. E. Hind, and J. F. Brugge. Modeling of auditory spatial receptive fields with spherical approximation functions. *Journal of Neurophysiology*, 80:2645–2656, 1998.
- L. B. W. Jongkees and J. J. van de Veer. On directional sound localization in unilateral deafness and its explanation. *Acta Otolaryngol*, 49:119–131, 1958. from Blauert (1997).
- M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37:183–233, 1999.

- P. X. Joris, P. H. Smith, and T. C. T. Yin. Coincidence detection in the auditory system: 50 years after jeffress. *Neuron*, 21:1235–1238, 1998.
- R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering*, 82(Series D):35–45, 1960.
- P. O. Kanold and E. D. Young. Proprioceptive information from the pinna provides somatosensory input to cat dorsal cochlear nucleus. *Journal of Neuroscience*, 21:7848–7858, 2001.
- A. J. King and C. H. Parsons. Improved auditory spatial acuity in visually deprived ferrets. *European Journal of Neuroscience*, 11:3945–3956, 1999.
- A. J. King, M. E. Hutchings, D. R. Moore, and C. Blakemore. Developmental plasticity in the visual and auditory representations in the mammalian superior colliculus. *Nature*, 332:73–76, 1988.
- A. J. King, C. H. Parsons, and D. R. Moore. Plasticity in the neural coding of auditory space in the mammalian brain. *PNAS*, 97:11821–11828, 2000.
- D. J. Kistler and F. L. Wightman. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *Journal of Acoustical Society of America*, 91:1637–1647, 1992.
- H. Klensch. Beitrag zur frage der lokalisattion des schalles im raum [a contribution to the study of the localisation of sound in space]. *Pflugers Arch*, 250:492–500, 1948. from Blauert (1997).
- E. I. Knudsen. Experience alters the spatial tuning of auditory units in the optic tectum during a sensitive period in the barn owl. *Journal of Neuroscience*, 5:3094–3109, 1985.
- E. I. Knudsen and M. S. Brainard. Visual instruction of the neural map of auditory space in the developing optic tectum. *Science*, 253:85–87, 1991.
- E. I. Knudsen and P. F. Knudsen. Vision guides the adjustment of auditory localization in young barn owls. *Science*, 230:545–548, 1985.
- E. I. Knudsen and P. F. Knudsen. Vision calibrates sound localization in developing barn owls. *Journal of Neuroscience*, 9:3306–3313, 1989.

- K. Koerdig and D. Wolpert. Bayesian integration in sensorimotor learning. *Nature*, 427: 244–247, 2004.
- M. Konishi. Coding of auditory space. *Annual Review Neuroscience*, 26:31–55, 2003.
- E. J. Kostelich and J. A. Yorke. Noise reduction in dynamical systems. *Phys. Rev. A*, 38: 1649–1652, 1988.
- T. Kristjansson, H. Attias, and J. Hershey. *Computer Vision (ECCV 2004)*, volume 3024/2004, chapter Stereo Based 3D Tracking and Scene Learning, Employing Particle Filtering within EM, pages 546–559. Springer Berlin / Heidelberg, 2004.
- B. Kuipers and Y.-T. Byun. A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. *Robotics and Autonomous Systems*, 8:47–63, 1981.
- S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22:79–86, 1951.
- S. M. LaValle, H. H. Gonzalez-Banos, C. Becker, and J.-C. Latombe. Motion strategies for maintaining visibility of a moving target. In *IEEE International Conference on Robotics and Automation*, 1997.
- D. N. Lee, F. R. van der Weel, T. Hitchcock, E. Matejowsky, and J. D. Pettigrew. Common principle of guidance by echolocation and vision. *J. Comp Physiol. A*, 171:563–571, 1992.
- C. Lenay, O. Gapenne, and J. Stewart. The constitution of spatiality in relation to the lived body: a study based on prosthetic perception. In *Emergence and Development of Embodied Cognition*, 2001.
- B. Lenggenhager, T. Tadi, T. Metzinger, and O. Blanke. Video ergo sum: Manipulating bodily self-consciousness. *Science*, 24:1096–1099, 2007.
- M. Levi. *Qualitative analysis of the periodically forced relaxation oscillations*, volume 32. AMS Bookstore, 1981.
- J. Lewald and W. H. Ehrenstein. Influence of head-to-trunk position on sound lateralization. *Experimental Brain Research*, 121:230–238, 1998.
- J. Lewald and H.-O. Karnath. Vestibular influence on human auditory space perception. *Journal of Neurophysiology*, 84:1107–1111, 2000.

- J. Lewald and H.-O. Karnath. Sound lateralization during passive whole-body rotation. *European Journal of Neuroscience*, 13:2268–2272, 2001.
- J. Lewald, G. J. Dörrscheidt, and W. H. Ehrenstein. Sound localization with eccentric head position. *Behavioural Brain Research*, 108:105–125, 2000.
- P. S. Linsay. An efficient method of forecasting chaotic time series using linear interpolation. *Physics Letters A*, 153:6–7, 1991.
- A. Logothetis, A. Isaksson, and R. J. Evans. Comparison of suboptimal strategies for optimal own-ship maneuvers in bearings-only tracking. In *American Control Conference*, 1998.
- R. F. Lyon. A computational model of filtering, detection, and compression in the cochlea. In *IEEE Acoust. Speech, and Sig. Proc.*, pages 1282–1285, 1982.
- D. J. C. MacKay. *Information theory, inference, and learning algorithms*. Cambridge University Press, 2003.
- E. A. Macpherson and J. C. Middlebrooks. Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *Journal of Acoustical Society of America*, 111:2219–2236, 2002.
- F. Maes, D. Vandermeulen, and P. Suedens. Medical image registration using mutual information. In *Proceedings of IEEE*, volume 91, pages 1699–1722, 2003.
- A. Mafra-Neto and R. T. Carde. Fine-scale structure of pheromone plumes modulates upwind orientation of flying moths. *Nature*, 369:142–144, 1994.
- R. Martinez-Cantin, N. de Freitas, A. Doucet, and J. Castellanos. Active policy learning for robot planning and exploration under uncertainty. In *Robotics: Science and Systems*, 2007.
- S. G. Mason, R. Bohringer, J. F. Borisoff, and G. E. Birch. Real-time control of a video game with a direct brain-computer interface. *Journal of Clinical Neurophysiology*, 21: 404–408, 2004.
- D. McAlpine and B. Grothe. Sound localization and delay lines – do mammals fit the model? *TRENDS in Neurosciences*, 26:347–350, 2003.

- P. B. L. Meijer. An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39:112–121, 1992.
- C. R. Meyer, B. A. Moffat, K. Kuszpit, P. L. Bland, T. L. Chenevart, A. Rehemtulla, and B. D. Ross. A methodology for registration of a histological slide and in vivo mri olume based on optimizing mutual information. *Molecular Imaging*, 5:16–23, 2006.
- J. C. Middlebrooks. Individual differences in external-ear transfer functions reduced by scaling in frequency. *Journal of the Acoustical Society of America*, 106:1480–1492, 1999a.
- J. C. Middlebrooks. Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *Journal of the Acoustical Society of America*, 106:1493–1510, 1999b.
- C. F. Moss and A. Surlykke. Auditory scene analysis by echolocation in bats. *Journal of the Acoustical Society of America*, 110:2207–2226, 2001.
- T. D. Mrsic-Flogel, A. J. King, R. L. Jenison, and J. W. H. Schnupp. Listening through different ears alters spatial response fields in ferret primary auditory cortex. *Journal of Neurophysiology*, 86:1043–1046, 2001.
- R. Muller and H.-U. Schnitzler. Acoustic flow perception in cf-bats: properties of the available cues. *Journal of the Acoustical Society of America*, 105:2959–2966, 1999.
- R. Muller and H.-U. Schnitzler. Acoustic flow perception in cf-bats: extraction of parameters. *Journal of the Acoustical Society of America*, 108:1298–1307, 2000.
- R. Muller and H.-U. Schnitzler. *Computational models of auditory function*, chapter Computational assessment of an acoustic flow hypothesis for cf-bats, pages 141–150. IOS Press, 2001.
- J. Murlis, J. S. Elkinton, and R. T. Card? Odor plumes and how insects use them. *Annual Review of Entomology*, 37:505–532, 1992.
- K. Murphy and S. Russell. Rao-blackwellised particle filtering for dynamical bayesian networks. In *Sequential Monte Carlo Methods in Practice*. Springer, 2001.
- G. R. Müller-Putza, R. Scherera, G. Pfurtschellera, and R. Rupp. Eeg-based neuroprosthesis control: A step towards clinical practice. *Neuroscience Letters*, 382:169–174, 2005.

- S. K. Nagel, C. Carl, T. Kringe, and R. Martin. Beyond sensory substitution - learning the sixth sense. *Journal of Neural Engineering*, 2:R13–R26, 2005.
- I. Nelken, Y. Rotman, and O. B. Yosef. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature*, 397:154–157, 1999.
- G. Neuweiler. *The biology of bats*. Oxford University Press, 2000.
- J. Nix and V. Hohmann. Sound source localization in real sound fields based on empirical statistics of interaural parameters. *JASA*, 119:463–479, 2006.
- A. Noe. *Action in perception*. MIT Press, 2004.
- Q. Norton. A sixth sense for a wired world. Wired News Online, 7 June 2006. URL <http://www.wired.com/news/technology/0,71087-0.html>.
- S. R. Oldfield and S. P. Parker. Acuity of sound localisation: a topography of auditory space. ii. pinna cues absent. *Perception*, 13:601–617, 1984.
- J. F. Olsen and N. Suga. Combination-sensitive neurons in the medial geniculate body of the mustached bat: encoding of relative velocity information. *J. Neurophysiol.*, 65: 1254–1274, 1991.
- L. Olsson, C. L. Nehaniv, and D. Polani. Sensory channel grouping and structure from uninterpreted sensor data. In *2004 NASA/DoD Conference on Evolvable Hardware*, pages 24–26. IEEE Computer Society Press, 2004.
- L. Olsson, C. Nehaniv, and D. Polani. From unknown sensors and actuators to visually guided movement. In *Proceedings of Development and Learning*, pages 1–6, 2005a.
- L. Olsson, C. L. Nehaniv, and D. Polani. Discovering motion flow by temporal-informational correlations in sensors. In *Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, 2005b.
- L. Olsson, C. Nehaniv, and D. Polani. From unknown sensors and actuators to actions grounded in sensorimotor perceptions. *Connection Science*, 18:121–144, 2006.
- W. O’Neill. *Echolocation in bats and dolphins*, chapter Feature extraction in the Mustached Bat auditory cortex, pages 176–185. Chicago, 2003.
- W. E. O’Neill and N. Suga. Encoding of target range and its representation in the auditory cortex of the mustached bat. *J. Neurosci.*, 2:17–31, 1982.

- K. J. O'Regan and A. Noe. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24:939–1031, 2001.
- E. Ott. *Chaos in dynamical systems*. Cambridge, 2nd edition, 2002.
- E. Ott, C. Grebogi, and J. A. Yorke. Controlling chaos. *Phys. Rev. Lett.*, 64:1196–1199, 1990.
- T. J. Park. Iid sensitivity differs between two principle centers in the interaural intensity difference pathway: the lso and the ic. *Journal of Neurophysiology*, 79:2416–2431, 1998.
- T. J. Park and G. D. Pollak. Gaba shapes sensitivity to interaural intensity disparities in the mustache bat's inferior colliculus: implications for encoding sound location. *Journal of Neuroscience*, 13:2050–2067, 1993.
- T. J. Park, B. Grothe, G. D. Pollak, G. Schuller, and U. Koch. Neural delays shape selectivity to interaural intensity differences in the lateral superior olive. *Journal of Neuroscience*, 16:6554–6566, 1996.
- T. J. Park, P. Monsivais, and G. D. Pollak. Processing of interaural intensity differences in the lso: Role of interaural threshold differences. *Journal of Neurophysiology*, 77: 2863–2878, 1997.
- J. M. Passerieux and D. van Cappel. Optimal observer maneuver for bearings-only tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 34:777–788, 1998.
- J. L. Pena and M. Konishi. From postsynaptic potentials to spikes in the genesis of auditory spatial receptive fields. *The Journal of Neuroscience*, 22:5652–5658, 2002.
- J. L. Pena, S. Viete, K. Funabiki, K. Saberi, and M. Konishi. Cochlear and neural delays for coincidence detection in owls. *Journal of Neuroscience*, 21:9455–9459, 2001.
- H. Peng, F. Lond, and C. Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Learning*, 27:1226–1238, 2005.
- K. B. Petersen and M. S. Pedersen. The matrix cookbook, February 2008. URL <http://matrixcookbook.com>.

- V. E. Pettorossi, M. Brosch, R. Panichi, F. Botti, S. Grassi, and D. Troiani. Contribution of self-motion perception to acoustic target localization. *Acta Oto-Laryngologica*, 125: 524–528, 2005.
- D. Philipona, K. J. O’Regan, and J.-P. Nadal. Is there something out there? inferring space from sensorimotor dependencies. *Neural Computation*, 15:2029–2049, 2003.
- D. Philipona, K. J. O’Regan, J.-P. Nadal, and O. Coenen. Perception of the structure of the physical world using unknown multimodal sensors and effectors. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*. MIT Press, Cambridge, MA, 2004.
- D. Pierce and B. J. Kuipers. Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92:169–227, 1997.
- J. P. W. Pluim and J. B. A. Maintz. Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22:986–1004, 2003.
- T. Poggio and F. Girosi. Regularization algorithms for learning that are equivalent to multilayer networks. *Science*, 247:978–982, 1990.
- G. D. Pollak, R. M. Burger, T. J. Park, A. Klug, and E. E. Bauer. Roles of inhibition for transforming binaural properties in the brainstem auditory system. *Hearing Research*, 168:60–78, 2002.
- J. M. Porta, B. Terwijn, and B. Krose. Efficient entropy-based action selection for appearance-based robot localization. In *International conference on Robotics and Automation*, 2003.
- J. M. Porta, J. J. Verbeek, and B. J. A. Krose. Active appearance-based robot localization using stereo vision. *Autonomous Robots*, 18:59–80, 2005.
- J.-M. Prieur, C. Bourdin, J.-L. Vercher, F. Sarès, J. Blouin, and G. M. Gauthier. Accuracy of spatial localization depending on head posture in a perturbed gravito-inertial force field. *Experimental Brain Research*, 161:432–440, 2005.
- J. Prinz. Putting the brakes on enactive perception. *Psyche*, 12:1–19, 2006.
- T. Quick, K. Dautenhahn, C. L. Nehaniv, and G. Roberts. The essence of embodiment: A framework for understanding and exploiting structural coupling between system and environment. In *Computing Anticipatory Systems*, 2000.

- J. P. Rauschecker and U. Kniepert. Auditory localization behaviour in visually deprived cats. *European Journal of Neuroscience*, 6:149–160, 1994.
- B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter*, chapter Bearings-only tracking, pages 103–151. Artech House, 2004a.
- B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter*, chapter Detection and tracking of stealthy targets, pages 239–259. Artech House, 2004b.
- B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter*. Artech House, 2004c.
- R. Rojas. *Neural Networks – A Systematic Introduction*. Springer-Verlag, 1996.
- A. Roorda and D. R. Williams. The arrangement of the three cone classes in the living human eye. *Nature*, 397:520–522, 1999.
- Y. Rossetti, K. Koga, and T. Mano. Prismatic displacement of vision induces transient changes in the timing of eye hand coordination. *Perception and Psychophysics*, 54:355–364, 1993.
- Y. Rossetti, G. Rode, L. Pisella, A. Farnè, L. Li, D. Boisson, and M.-T. Perenin. Prism adaptation to a rightward optical deviation rehabilitates left hemispatial neglect. *Nature*, 395:166–169, 1998.
- M. Schlesinger. Reexamining visual cognition in human infants: On the necessity of representation. *Behavioral and Brain Sciences*, 24:1003–1004, 2001. Commentary/O’Regan & Noe: A sensorimotor account of vision and visual consciousness.
- B. J. Scholl and D. J. Simmons. Change blindness, gibson, and the sensorimotor theory of vision. *Behavioral and Brain Sciences*, 24:1004–1006, 2001. Commentary/O’Regan & Noe: A sensorimotor account of vision and visual consciousness.
- J. R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–457, 1980.
- J. R. Searle. *Mind Design*, chapter Minds, brains, and programs, pages 282–306. MIT Press, 1981.
- H. Shatkay. *Learning models for robot navigation*. PhD thesis, Brown University Providence, 1999.

- T. Shinbrot, E. Ott, C. Grebogi, and J. A. Yorke. Using chaos to direct trajectories to targets. *Phys. Rev. Lett.*, 65:3215–3218, 1990.
- B. G. Shinn-Cunningham, S. Santarelli, and N. Kopco. Tori of confusion: binaural localization cues for sources within reach of a listener. *Journal of the Acoustical Society of America*, 107:1627–1636, 2000.
- R. G. Simmons, D. Apfelbaum, W. Burgard, D. Fox, M. Moors, S. Thrun, and H. L. S. Younes. Coordination for multi-robot exploration and mapping. In *Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, 2000.
- S. Skaff, T. Arbel, and J. J. Clark. Active bayesian color constancy with non-uniform sensors. In *16th International Conference on Pattern Recognition*, 2002.
- M. Slaney. Lyon’s cochlear model. Technical Report 13, Apple Computer, 1988.
- S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.*, 73:747–817, 1967.
- R. C. Smith and P. Cheeseman. On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research*, 5:56–68, 1986.
- S. M. Smith. Reviews of optic flow, motion segmentation, edge finding and corner finding. Technical Report TR97SMS1, Oxford University, 1997. URL <http://users.fmrrib.ox.ac.uk/~steve/review/review/review.html>.
- E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Springer, second edition edition, 1999.
- J. R. Spletzer and C. J. Taylor. Dynamic sensor planning and control for optimally tracking targets. *International Journal of Robotics Research*, 22:7–20, 2003.
- G. M. Stratton. Vision without inversion of the retinal image. *Psychol. Rev.*, 4:341–360, 1897.
- N. Suga. Cortical computational maps for auditory imaging. *Neural Networks*, 3:3–21, 1990. Semi-complete introduction by a very important figure in the Neuro-biology of Microchiroptera echolocation. This appears to cover almost all that is known about the ascending system but for the work by Covey & Casaday, and a few others.

- N. Suga and X. Ma. Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Reviews Neuroscience*, 4:783–794, 2003a.
- N. Suga and X. Ma. Multiparametric corticofugal modulation and plasticity in the auditory system. *Nature Reviews Neurosci.*, 4:783–794, 2003b.
- N. Suga and K. Tsuzuki. Inhibition and level-tolerant frequency tuning in the auditory cortex of the mustached bat. *J. Neurophysiol.*, 53:1109–1145, 1985.
- V. A. Sujan and S. Dubowsky. Visually built task models for robot teams in unstructured environments. In *IEEE International Conference on Robotics and Automation*,, 2002.
- G. Svirskis, D. H. S. Vibhakar Kotak, and J. Rinzel. Sodium along with low-threshold potassium currents enhance coincidence detection of subthreshold noisy signals in mso neurons. *J. Neurophysiol.*, 91:2465–2473, 2004.
- S. Thill and T. C. Pearce. Understanding complex behaviors by analyzing optimized models: *C. elegans* gradient navigation. *HFSP Journal*, 1:263–273, 2007.
- S. Thrun. Learning occupancy grid maps with forward sensor models. *Autonomous Robots*, 15:111–127, 2003.
- S. Thrun, W. Burgard, and D. Fox. A probabilistic approach to concurrent mapping and localization for mobile robots. *Autonomous Robots*, 5:253–271, 2004.
- S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*, chapter 17, Exploration, pages 569–605. MIT Press, 2006.
- W. R. Thurlow. Head movements during sound localization. *Journal of the Acoustical Society of America*, 42:1347, 1967.
- W. R. Thurlow and P. S. Runge. Effect of induced head movements on localization of direction of sounds. *Journal of the Acoustical Society of America*, 42:480–488, 1967.
- W. R. Thurlow, J. W. Mangels, and P. S. Runge. Head movements during sound localization. *Journal of the Acoustical Society of America*, 42:489–493, 1967.
- M. A. Umilta, L. Escola, I. Intskirveli, F. Grammont, M. Rochat, F. Caruana, A. Jezzini, V. Gallese, and G. Rizzolatti. When pliers become fingers in the monkey motor system. *PNAS*, 105:2209–2213, 2008.

- R. van Gulick. Still room for representations. *Behavioral and Brain Sciences*, 24:1007–1008, 2001.
- Commenatry/O’Regan & Noe: A sensorimotor account of vision and visual consciousness
- Commenatry/O’Regan & Noe: A sensorimotor account of vision and visual consciousness
- Commenatry/O’Regan & Noe: A sensorimotor account of vision and visual consciousness, 2001.
- Commenatry/O’Regan & Noe: A sensorimotor account of vision and visual consciousness.
- F. J. Varela, E. Rosch, and E. Thompson. *Embodied Mind: Cognitive Science and Human Experience*. MIT Press, 1995.
- J. J. Verbeek, N. Vlassis, and J. R. J. Nunnink. A variational em algorithm for large-scale mixture modeling. In *8th Annual Conference of the Advanced School for Computing and Imaging*, 2003.
- M. Vergassola, E. Villermaux, and B. I. Shraiman. ‘infotaxis’ as a strategy for searching without gradients. *Nature*, 25:406–409, 2007a.
- M. Vergassola, E. Villermaux, and B. I. Shraiman. Supplementary materials for ‘infotaxis: searching without gradients’. *Nature*, 25:406–409, 2007b. URL <http://www.nature.com/nature/journal/v445/n7126/extref/nature05464-s1.pdf>. supplementary material.
- D. Vernon. Cognitive vision: The case for embodied perception. *Image and Vision Computing*, 26:127–140, 2008.
- N. Vickers and T. Baker. Reiterative responses to single strands of odor promote sustained upwind flight and odor source location by moths. *PNAS*, 9:5756–5760, 1994.
- S. Vijayakumar and S. Schaal. Local adaptive subspace regression. *Neural Processing Letters*, 7:139–149, 1998.
- S. Vijayakumar and S. Schaal. Lwpr: An $o(n)$ algorithm for incremental real time learning in high dimensional space. In *International Conference on Machine Learning*, pages 1079–1086, 2000.
- S. Vijayakumar, A. D’Souza, and S. Schaal. Incremental online learning in high dimensions. *Neural Computation*, 17:2602–2634, 2005.

- V. A. Walker, H. Peremans, and J. C. T. Hallam. One tone, two ears, three dimensions: A robotic investigation of pinnae movements used by rhinolophid and hipposiderid bats. *Journal of the Acoustical Society of America*, 104:569–579, 1998.
- H. Wallach. Uber die wahrnehmung der schallrichtung [on the perception of the direction of sound]. *Psychologische Forschung*, 22:238–266, 1938. from Blauert (1997).
- H. Wallach. On sound localization. *Journal of the Acoustical Society of America*, 10: 270–274, 1939.
- H. Wallach. The role of head movements and vestibular and visual cues in sound localization. *Experimental Psychology*, 27:339–368, 1940.
- D. Wang and G. J. Brown. *Computational auditory scene analysis: principles, algorithms, and applications*. John Wiley & Sons, 2006.
- M. J. Weissburg and R. K. Zimmer-Faust. Odor plumes and how blue crabs use them in finding prey. *Journal of Experimental Biology*, 197:349–375, 1994.
- R. B. Welch, B. Bridgeman, S. Anadr, and K. E. Browman. Alternating prism exposure causes adaptation and generalization to a novel displacement. *Perception and Psychophysics*, 54:195–204, 1993.
- M. Welling. The kalman filter. Classnotes, 2008.
- J. J. Wenstrup, L. S. Ross, and G. D. Pollak. A functional organization of binaural responses in the inferior colliculus. *Hearing Res.*, 17:191–195, 1985. Referenced in Pollak EtAL HearingRes 168 2002 60-78.
- J. J. Wenstrup, L. S. Ross, and G. D. Pollak. Binaural response organisation within a frequency-band representation of the inferior colliculus: implications for sound localization. *J. Neurosci.*, 6:962–973, 1986.
- E. M. Wenzel. The relative contribution of interaural time and magnitude cues to dynamic sound localisation. In *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, 1995.
- E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman. Localization using nonindividualized head-related transfer functions. *Journal of the Acoustical Society of America*, 94:111–123, 1993.

- F. L. Wightman and D. J. Kistler. Monaural sound localization revisited. *Journal of Acoustical Society of America*, 101:1050–1063, 1997.
- F. L. Wightman and D. J. Kistler. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *Journal of the Acoustical Society of America*, 105:2841–2853, 1999.
- D. Wilmington, L. Gray, and R. Jahrsdoerfer. Binaural processing after corrected congenital unilateral conductive hearing loss. *Hearing Research*, 74:99–114, 1994.
- J. R. Wolpaw and D. J. McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *PNAS*, 101:17849–17854, 2004.
- J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8:164–173, 2000.
- P. B. y Rita. Tactile sensory substitution studies. *Annals of the New York Academy of Sciences*, 1013:83–91, 2004.
- P. B. y Rita, C. C. Collins, F. A. Saunders, B. White, and L. Scadden. Vision substitution by tactile image projection. *Nature*, 221:963–964, 1969.
- B. Yamauchi and R. Beer. Spatial learning for navigation in dynamic environments. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 26:496–505, 1996.
- B. Yamauchi and P. Langley. Place recognition in dynamic environments. *Journal of Robotic Systems*, 14:107–120, 1997.
- L. Yang, Q. Liu, and G. D. Pollak. Afferent connections to the dorsal nucleus of the lateral lemniscus of the mustached bat: Evidence for two functional subdivisions. *J. Comp. Neurol.*, 373:575–592, 1996.
- H.-T. Zhu and S.-Y. Lee. Local influence for incomplete data models. *Journal of the Royal Statistical Society: Series B*, 63:111–126, 2001.
- U. R. Zimmer. Robust world-modelling and navigation in a real world. *Neurocomputing*, 13:247–260, 1996.

R. K. Zimmer-Faust, C. M. Finelli, N. D. Pentcheff, and D. S. Wethey. Odor plumes and animal navigation in turbulent water flow: a field study. *Biological Bulletin*, 188: 111–116, 1995.

P. M. Zurek. Probability distributions of interaural phase and level differences in binaural detection stimuli. *Journal of the Acoustical Society of America*, 90:1927–1932, 1991.