Delft University of Technology

Free Energy Principle Based Precision Modulation for Robot Attention
Towards brain inspired robot intelligence

Anil Meera, A.

DOI

Publication date
2023

Document Version
Final published version

Citation (APA)

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Free Energy Principle based Precision Modulation for Robot Attention

Towards Brain Inspired Robot Intelligence

Ajith Anil Meera

Free Energy Principle Based Precision Modulation for Robot Attention

Towards brain inspired robot intelligence

Ajith ANIL MEERA

Free Energy Principle Based Precision Modulation for Robot Attention

Towards brain inspired robot intelligence

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen;
Chair of the Board of Doctorates
to be defended publicly on
Friday 10th of March 2023 at 10:00 o'clock

by

Ajith ANIL MEERA

Master of Science in Mechanical Engineering,
Delft University of Technology, The Netherlands,
born in Neyyattinkara, Kerala, India.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Prof.dr.ir. M. Wisse,	Delft University of Technology, promotor
Dr. P. Mohajerin Esfahani,	Delft University of Technology, copromotor

Independent members:

Prof. K.J. Friston,	University College London, UK
Prof. dr. ir. C.L. Buckley,	University of Sussex, UK
Prof. dr. G.C.H.E de Croon,	Delft University of Technology
Dr. T. Verbeelen,	Ghent University, Belgium
Prof.dr.ir. J. Hellendoorn,	Delft University of Technology, reserve member.

Other members:

Dr. ir. P.L. Lanillos Pradas,	Radbound University, The Netherlands

This work was done at the Department of Cognitive Robotics at TU Delft. Dr. ir. Pablo Lanillos has contributed to Chapter 2.

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Dream, dream, dream. Dreams transform into thoughts, and thoughts result in action.

A.P.J Abdul Kalam, The Missile Man of India

Contents

Summary

The potential impact of a grand unified theory of the brain on the robotics community might be immense, as it might hold the key to the general artificial intelligence. Such a theory might make revolutionary leaps in robot intelligence by improving the quality of our lives. The last two decades have witnessed the rise of one such brain theory - the free energy principle (FEP) - that seems to be successful in explaining a large body of cognitive functions. The tremendous amount of research centering FEP is a testament to its popularity within the neuroscience community. This raises two important questions: i) since biological systems are fundamentally different from robots, will FEP be useful in solving real robotics problems? ii) if so, will it outperform classical robot algorithms? To answer these questions, this thesis takes a step in the direction of applying FEP on three class of robotics challenges, with a special focus on Unmanned Aerial Vehicle (UAV): i) action, ii) perception and iii) active perception. This thesis demonstrates the usefulness of FEP in solving these challenges, and shows that FEP is particularly beneficial in dealing with colored (non-white [1]) noise during estimation (perception) when compared to classical methods, marking the utility of FEP not only in neuroscience, but also in robotics. With these results, this thesis aims to contribute to the rise of FEP as a unified theory of robot intelligence.

Problem/challenge and the proposed solution: It is well known that the modern robots are incapable of dealing with high uncertainties in the world, when compared to the robustness of the human brain that seamlessly deals with uncertainties on a day-to-day basis. This indicates a clear distinction between the status of robot intelligence and brain's intelligence, pointing to the need for a brain inspired robot algorithm that is robust against uncertainties. One of the contributing factor to the brain's capability to deal with uncertainties is its attention (the capacity to only process the relevant signals and adapt this modulation depending on the task). The recent advancements in neuroscience has resulted in the use of precision (inverse covariance) modulation as a means to model the brain's attention for uncertainty resolution, motivating the use of such a scheme for the development of a robot cognitive model. With an ambitious goal of stepping towards the brain-inspired robot intelligence, this thesis proposes an FEP based robot cognitive model. Drawing inspiration from the neuroscience literature, chapter 2 introduces a novel robot cognitive model. The core idea behind the model is the use of precision (inverse covariance) modulation as the means for robot attention. Four types of precision modulations are proposed in this thesis: precision modelling, precision learning, precision optimization and precision fluctuation. This thesis focuses on developing the mathematical framework for each type of precision modulation, followed by an extensive evaluation in simulation and on real robot data to demonstrate the advantages of the model. At the end, this thesis aims to use FEP to create a robot cognitive model with uncertainty resolution capabilities, that solves practical robotics challenges, such as learning and control, while outperforming classical robot algorithms.

Precision modelling: This category of precision modulation uses the prior knowledge about the structure of uncertainty (or noise) in the measurements to improve the quality of estimation and control. We solve two problems here: i) state and input estimation under colored noise (perception) and ii) multi robot navigation in formation (action). State and input estimation under colored noise is still an open problem in control systems. In chapter 3, a state and input observer design for Linear Time Invariant (LTI) systems under colored noise is proposed using FEP. The observer uses precision modelling to attend to the most informative measurement derivatives during inference which results in a better quality estimation when compared to the state of the art observers. In chapter 4, an experimental design with a quadrotor hovering under wind conditions is introduced to provide the proof of concept for the observer design. The results show the superior performance of the observer on real data, confirming the usefulness of the observer to solve real robot problems. In chapter 9, precision modelling is used to generate different robot behaviors during an obstacle avoidance task with multi robot navigation in formation. These results show that precision modelling using FEP is advantageous to solve robotics problems. The FEP based observer design in chapter 3 outperforms the state observer benchmarks under colored noise, answering positively to the second question of this thesis.

Precision learning: This category of precision modulation uses an FEP based scheme to learn different kinds of precision (inverse covariance) matrices from data, aiding a higher quality estimation. We solve two perception problems here: i) system identification under colored noise, which is an open challenge in control systems, and ii) noise smoothness learning. In chapter 6, a system identification tool for LTI systems under colored noise is proposed using FEP. The algorithm uses precision learning to attend towards learning the most uncertain parameters of the system. The safe operation of the algorithm on robot is motivated using the proof of convergence for the parameter estimator in chapter 7. In chapter 8, the estimator is tested on quadrotor flight data to show its usefulness for real robot problems. In chapter 5, a novel online noise smoothness estimator is proposed, based on precision learning. Through rigorous testing in simulation and on quadrotor flight data, the joint state and noise smoothness observer (DEMs) is shown to outperform classical methods in estimation, highlighting the advantages of precision learning in robotics.

Precision optimization: This category of precision modulation uses precision optimization (or uncertainty minimization) as a prime objective for active perception problems. It directly contributes to the use of robot attention for uncertainty minimization of robots during information gathering problems like mapping, data acquisition, target search etc. In chapter 2, the information seeking robot behavior based on uncertainty minimization is recognized as precision optimization. These planning algorithms are shown to give rise to pure exploratory strategies. Informative Path Planning problem for target search using a UAV is solved using precision optimization.

Precision fluctuations: This category of precision modulation uses fluctuations in precision as a core driving mechanism for active perception problems. In chapter 2, a rhythmic precision modulated attention model for active perception problems is proposed. The advantages of temporal scheduling of action and perception, similar to that of eye saccades and information gathering during a visual search is discussed. The advantages of using precision fluctuations to trade-off between exploration and exploitation during a system identification task is also discussed.

In summary, this thesis proposes a precision modulation based robot attention model using FEP. The advantages of such a model in the context of robotic challenges are explored with a special focus on UAVs using three class of problems: action, perception and active perception. Through extensive mathematical formulations, simulations and experimental analysis, the robot attention model is shown to be advantageous for estimation, control and planning. Despite the differences between a biological system and a robotic system, a unified brain theory could provide competent tools to solve robotic problems, even outperforming the classical methods for estimation under colored noise. This answers the two main questions addressed by the thesis.

Samenvatting

Een allesomvattende theorie over de werking van de hersenen kan potentiëel een enorme invloed hebben op de robotica, omdat het de sleutel zou kunnen zijn tot de zogenaamde "sterke"(= echt menselijke) kunstmatige intelligentie. Een dergelijke theorie zou revolutionaire sprongen maken in robotintelligentie en zou de kwaliteit van ons leven verbeteren. De laatste twee decennia zijn we getuige geweest van de opkomst van zo'n hersentheorie - het Free Energy Principle (FEP) - die succesvol lijkt te zijn in het verklaren van een groot aantal cognitieve functies. De enorme hoeveelheid onderzoek waarin FEP centraal staat, is een bewijs van zijn populariteit binnen de neurowetenschappelijke gemeenschap. Dit roept twee belangrijke vragen op: i) zal FEP, aangezien biologische systemen fundamenteel verschillen van robots, nuttig zijn bij het oplossen van echte robotica-problemen? ii) zo ja, zal het beter presteren dan klassieke robotalgoritmen? Om deze vragen te beantwoorden, zet dit proefschrift een stap in de richting van het toepassen van FEP op drie klassen van robotica-uitdagingen, met een speciale focus op Unmanned Aerial Vehicles (UAV): i) actie, ii) perceptie en iii) actieve perceptie. Dit proefschrift demonstreert het nut van FEP bij het oplossen van deze uitdagingen, en laat zien dat FEP bijzonder gunstig is bij het omgaan met gekleurde (niet-wit [1]) ruis tijdens schatting (perceptie) in vergelijking met klassieke methoden, wat het nut van FEP niet alleen in de neurowetenschappen, maar ook in robotica aangeeft. Met deze resultaten wil dit proefschrift bijdragen aan de opkomst van FEP als een algemene theorie voor robotintelligentie.

Probleem/ uitdaging en de voorgestelde oplossing: Het is algemeen bekend dat de moderne robots niet in staat zijn om te gaan met grote onzekerheden in de wereld, in vergelijking met de robuustheid van het menselijk brein dat probleemloos omgaat met onzekerheden van dag tot dag. Dit geeft een duidelijk onderscheid aan tussen de status van robotintelligentie en de intelligentie van de hersenen, wat wijst op de behoefte aan een op de hersenen geïnspireerd robotalgoritme dat robuust is tegen onzekerheden. Het vermogen van de hersenen om met onzekerheden om te gaan, verwijst rechtstreeks naar de aandachtsmodellen. De recente vooruitgang in de neurowetenschappen heeft geresulteerd in het gebruik van precisiemodulatie (inverse covariantie) als een middel om de aandacht van de hersenen te modelleren voor het oplossen van onzekerheid, wat het gebruik van een dergelijk schema voor de ontwikkeling van een robotaandachtsmodel motiveert. Met een ambitieus doel om de stap naar de door de hersenen geïnspireerde robotintelligentie te maken, stelt dit proefschrift een op FEP gebaseerd robotaandachtsmodel voor. Geïnspireerd door de neurowetenschappelijke literatuur, introduceert hoofdstuk 2 een nieuw robotaandachtsmodel. Het kernidee achter het model is het gebruik van precisie (inverse covariantie) modulatie als middel voor robotaandacht. In dit proefschrift worden vier soorten precisiemodulaties voorgesteld: precisiemodellering, precisieleren, precisie-optimalisatie en precisiefluctuatie. Dit proefschrift richt zich op het ontwikkelen van het wiskundige raamwerk voor elk type precisiemodulatie, gevolgd door een uitgebreide evaluatie in simulatie en op echte robotgegevens om de voordelen van het model aan te tonen. Het doel

van dit proefschrift is om FEP te gebruiken om een robotaandachtsmodel te creëren voor het oplossen van onzekerheid, dat praktische robotica-uitdagingen oplost, terwijl het de klassieke robotalgoritmen overtreft.

Precisiemodellering: Deze categorie precisiemodulatie gebruikt de voorkennis over de structuur en grootte van precisie (of onzekerheden) in de gegevens om de kwaliteit van schatting en controle te verbeteren. We lossen hier twee problemen op: i) schatting van toestand en invoer onder gekleurde ruis (perceptie) en ii) navigatie met meerdere robots in formatie (actie). Toestand- en ingangsschatting onder gekleurde ruis is nog steeds een open probleem in besturingssystemen. In hoofdstuk 3 wordt een toestand- en input-observer ontworpen voor lineaire tijdinvariante (LTI) systemen onder gekleurde ruis met behulp van FEP. De observer maakt gebruik van precisiemodellering om aandacht te geven aan de meest informatieve metingen tijdens inferentie, wat resulteert in een schatting van betere kwaliteit in vergelijking met state-of-the-art observers. In hoofdstuk 4 wordt dit nieuwe observer ontwerp getoetst op data van een quadrotor die in wind vliegt. De resultaten tonen de superieure prestaties van de observer, wat het nut van het Free Energy Principle bevestigt om echte robotproblemen op te lossen. In hoofdstuk 9 wordt precisiemodellering gebruikt om verschillende robotgedragingen te genereren tijdens een taak om obstakels te vermijden met multi-robotnavigatie in formatie. Ook deze resultaten laten zien dat precisiemodellering met FEP voordelig is om robotica-problemen op te lossen. Het op FEP gebaseerde observer ontwerp in hoofdstuk 3 presteert beter dan de benchmarks onder gekleurde ruis, wat een positief antwoord geeft op de tweede vraag van dit proefschrift.

Precisieleren: Deze categorie precisiemodulatie gebruikt een op FEP gebaseerd schema om verschillende soorten precisiematrices (inverse covariantie) uit gegevens te leren, wat bijdraagt aan een schatting van hogere kwaliteit. We lossen hier twee perceptieproblemen op: i) systeemidentificatie onder gekleurde ruis, wat een open uitdaging is in regelsystemen, en ii) het leren van de hoeveelheid autocorrelatie (smoothness) van ruis. In hoofdstuk 6 wordt een systeemidentificatietool voor LTI-systemen onder gekleurde ruis voorgesteld met behulp van FEP. Het algoritme gebruikt precisieleren om de meest onzekere parameters van het systeem te leren. De veilige werking van het algoritme op de robot wordt gemotiveerd met behulp van het bewijs van convergentie voor de parameterschatter in hoofdstuk 7. In hoofdstuk 8 wordt de schatter getest op quadrotorvluchtgegevens om zijn bruikbaarheid voor echte robotproblemen aan te tonen. In hoofdstuk 5 wordt een nieuwe online schatter voor ruis autocorrelatie voorgesteld, gebaseerd op precisieleren. Door rigoreuze tests in simulatie en op quadrotor-vluchtgegevens, is aangetoond dat de gecombineerde observer voor zowel systeemtoestand als ruis eigenschappen (DEM's) beter presteert dan klassieke methoden bij het schatten.

Precisie-optimalisatie: Deze categorie van precisiemodulatie gebruikt precisie optimalisatie (of onzekerheidsminimalisatie) als een hoofddoel voor actieve waarnemingsproblemen. Het draagt direct bij aan het gebruik van robot-aandacht voor het minimaliseren van onzekerheid van robots tijdens het verzamelen van informatie, zoals mapping, data-acquisitie, het zoeken naar doelen enz. Hoofdstuk 2 gebruikt precisie-optimalisatie om informatiezoekend robotgedrag op basis van onzekerheidsminimalisatie te creëren. Deze planningsalgoritmen blijken aanleiding te geven tot pure verkenningsstrategieën. Het informatieve padplanningsprobleem voor het zoeken naar doelen met behulp van een UAV wordt opgelost met behulp van precisie-optimalisatie.

Precisiefluctuaties: Deze categorie van precisiemodulatie gebruikt precisiefluctuaties als een centraal aandrijfmechanisme voor actieve waarnemingsproblemen. In hoofdstuk 2 wordt een ritmisch precisie-gemoduleerd aandachtsmodel voor actieve waarnemings-problemen voorgesteld. De voordelen van temporele planning van actie en waarneming, vergelijkbaar met de oogsaccades en het verzamelen van informatie tijdens een visuele zoekopdracht, worden besproken. De voordelen van het gebruik van precisiefluctuaties voor de afweging tussen exploratie en exploitatie tijdens een systeemidentificatietaak worden ook besproken.

Samenvattend stelt dit proefschrift een op precisiemodulatie gebaseerd robotaandachts-model voor dat gebruik maakt van FEP. De voordelen van een dergelijk model in de context van robotische uitdagingen worden onderzocht met een speciale focus op UAV's met behulp van drie klassen van problemen: actie, perceptie en actieve perceptie. Door middel van uitgebreide wiskundige formuleringen, simulaties en experimentele analyse, is aangetoond dat het robotattentiemodel voordelig is voor schatting, controle en planning. Ondanks de verschillen tussen een biologisch systeem en een robotsysteem, zou een algemene theorie over de werking van de hersenen kunnen leiden tot competente hulpmiddelen om robotproblemen op te lossen (hoofdvraag 1), en zelfs beter presteren dan de klassieke methoden voor schatting onder gekleurde ruis (hoofdvraag 2).

1

1

INTRODUCTION

1

T<small>HIS</small> chapter provides the general introduction to the thesis by providing the necessary neuroscience background information, motivating the method selections, defining the clear scope of the thesis within the robotics context, and stating the core technical contributions of the thesis chapters - all focused towards proposing a free energy principle based attention model for robots.

1.1 Background & Motivation

The quest for a grand unified theory of the brain has been long standing in the neuroscience domain. This section aims to provide the motivation and background behind using one such theory to move towards robot intelligence.

1.1.1 From brain intelligence to robot intelligence

Robots have long been used in highly controlled factory environments. Decades of robotics research has begun to move them to human friendly environments, to even sharing a living or working space together. Service robots, floor cleaning robots, shared control in cars, surveillance drones etc. are some of the examples of this sharing. However, the modern-day robots are highly susceptible to uncertainties in the real world that they fail too often than desirable. Drones randomly crashing into trees in wind, humanoids falling down due to obstacles, perception algorithms misclassifying a dog with a cat on a cloudy day are common examples of failure under uncertainties, making these robots seem senseless when compared to the robustness of the human brain. This has long motivated the robotics researchers into looking for a brain inspired robot algorithm (embodied intelligence) to deal with uncertainties in the real world, including algorithms like reinforcement learning, artificial neural networks, and swarm intelligence. However, none of these algorithms have succeeded in providing a unified account for robot intelligence under uncertainties. Unless the uncertainty handling capabilities of robots are significantly improved, the dream of a world where robots and humans could coexist with trust and reliability will not be fulfilled. Therefore, a brain inspired robot algorithm that can deal with uncertainties could make significant impact to the quality of our lives, motivating the core theme of this thesis.

Although a unified brain theory could be a potential candidate for the unified robot intelligence, the fact that the brain and a robot are two different entities - one a biological system and the other a mechatronic system - poses a serious challenge. Therefore, it is important to test the theory with regards to real robotics problems to show that it is indeed useful in the robotics context. This wouldn't be sufficient for the theory to be popularly adopted within the community. The theory should enable robots to be smarter than they currently are. In this thesis, we move a step closer by applying one of the most successful brain theories called Free Energy Principle (FEP) on real robotics problems, and showing that it outperforms classical robotics algorithms. **The two research questions addressed by the thesis are: i) Is FEP useful in solving real robotics problems in estimation and control? ii) Does FEP outperform classical robot algorithms?**

> **Research questions**
>
> Is FEP useful in solving real robotics problems in estimation and control? If so, does it outperform classical robot algorithms?

Within the context of uncertainty, the neuroscience literature points to two interconnected terminologies: i) precision and ii) attention. Precision or the inverse covariance matrix, denotes the level of noise or uncertainty in the world that directly enters the sensory system through measurements. In other words, precision is the inverse of the covariance matrix denoting the measurement noise. More the precision, less the noise. Precision can also be seen as the level of confidence that the brain has about its estimates of the model parameters of the world. More the parameter precision, less uncertain is the brain about its parameter estimates. Therefore, precision is directly connected to uncertainty. Attention is the brain's mechanism that identifies which part of the sensory measurement to focus on. The brain attends to the most noisy or uncertain measurements of the world. In other words, uncertainty grabs the brain's attention. This intertwining of precision and attention is leveraged in this thesis for the development of an FEP based robot attention model using precision. Therefore, building a robot attention model is directly linked to the uncertainty resolution power of robot algorithms. The following sections will motivate the use of FEP and attention models for this purpose.

1.1.2 Free energy principle

This section will briefly place FEP within the computational neuroscience literature. The detailed explanation for the relevant aspects of FEP will appear in the subsequent chapters. Very briefly, according to FEP, any self organising system that is in equilibrium with its environment must minimize an information theoretic objective called free energy [2], effectively minimizing its sensory surprisal. Since surprisal is a direct consequence of uncertainty, FEP provides a seamless way to model the uncertainty handling capability of the brain, which is highly relevant to the development of robot intelligence.

Recent advancements in neuroscience points to the dominant role of probabilistic inference in the brain's perception process [3]. According to the Bayesian brain hypothesis [4], the human perceptual computations are Bayes optimal, in the sense that the brain uses its sensory measurements to infer the causal dynamic model of the world in a Bayes optimal way. These models are then used by the brain to constantly make predictions about the world. This idea is fundamental to hierarchical predictive coding [5] where the brain's prediction errors propagate up the hierarchy, while its expectations propagate down the hierarchy. FEP builds on the insights from Bayesian brain hypothesis [6, 7] and predictive coding [8, 9] to naturally accommodate action and perception within a single framework, thereby proposing a unified theory of the brain. The explanatory power of FEP to describe the cognitive neuroscience functions like human visual exploration [10], saccadic eye movements for visual search [11], embodied cognition [12], schizophrenia [13], mirror neurons [14], illusory perception [15], dreaming consciousness [16], action selection [17] etc bolsters the role of FEP as a grand unified theory of the brain. This motivates the selection of FEP as the central pillar of inspiration from neuroscience to develop a brain inspired robot attention model for this thesis.

1.1.3 Attention

Attention is a central theme in the cognitive processing that determines which events from the environment are preferentially processed [18]. It is the process of optimizing synaptic gain to represent the precision (or uncertainty) of sensory information (prediction error) during inference [19], and has been widely used to model the eye saccades for information gathering [20, 21]. Moreover, the role of precision control within the context of attention has been shown to be central towards body representations in the brain [22]. Empirical evidence from neuroscience pointing to the role of neuromodulators that encode precision in accordance with FEP [17, 23, 24] motivates the use of FEP for attention modelling in this thesis. Since emulating the brain's attention on robots lies at the core of uncertainty resolution power of robot intelligence, the development of such a model is expected to be highly influential within the robotics community.

1.1.4 FEP in robotics

The robotics community has already identified the potential of FEP. The large body of research that uses FEP to solve crucial cognitive robotics problems [25, 26] is a testament to the emergence of FEP as a potential candidate for the brain-inspired robot intelligence. Robot manipulator arm control [27, 28], robot modulation and stacking [29], dyadic robot imitative interactions [30], humanoid body perception and action [31, 32], human robot interaction [33], SLAM [34], path planning [35], simulated robot arm learning to reproduce handwritten letters [36], reinforcement learning [37], message passing for NARMAX identification [38], factor graphs for optimal control [39] and signal processing [40], blind image quality assessment [41] are some examples of these applications. The similarities of FEP with reinforcement learning [42], neural networks [43, 44], Kalman Filtering [45, 46], PID control [47], optimal control [48, 49] and active learning [50] further guides the quest for a brain inspired robot algorithm towards FEP. This motivates the use of FEP to provide a unified theory for robot cognition. The next section deals with clearly defining the scope of this thesis within the robotics context.

1.2 Scope of the thesis

The robot attention model proposed in this thesis is wide and ambitious in its scope. However, three main class of problems were selected to demonstrate the usefulness of this model in real robotics problems: action, perception and active perception. Action refers to the general robot control problems, whereas perception refers to the learning and estimation problems. Active perception problem involves control and learning in the loop. Within action, two demonstrative example problems were selected, owing to its potential impact to the field: i) reactive path planning and ii) multi robot navigation in formation. Within perception, three demonstrative example problems in state space continuous formulations with colored noise were selected: i) state and input estimation, ii) noise estimation and iii) system identification. The main motivation for this selection is the seamless noise color handling capabilities of FEP for the estimation under colored noise for state space systems, which is still an open challenge in the control domain. For active perception, the IPP problem for target search was selected. The ideas proposed in this section is applicable to a wide range of information gathering problems that involve an

information seeking or uncertainty reduction based robot behavior. Although the proposed robot attention model can be applied to a large class of robots, this thesis uses UAVs as the prime robot application for all algorithms. The core motivation for this selection is the wide impact of UAVs for exploration and target search problems, and the growing demand for delivery drone algorithms that are robust against windy flight conditions. Figure 1.1 shows the robot applications (or problems) considered within the scope of this thesis. The next section will deal with delineating the core contributions of this thesis towards solving these problems.
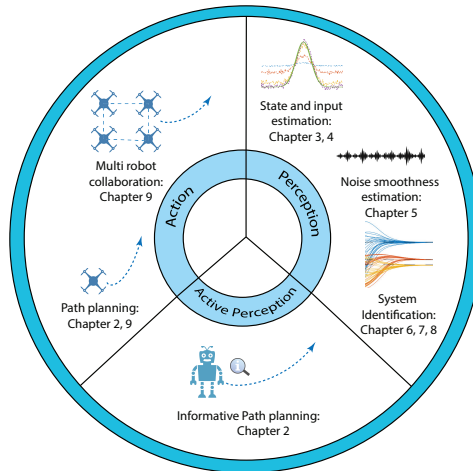
Fig 1.1: The scope of the thesis within the robotics context, all focused on UAVs as the prime robot application. Three main classes of problems are considered: action, perception and active perception.

1.3 Core contributions

The core contributions of this thesis are focused towards addressing the two main research questions posed in Section 1.1.1, both in light of developing an FEP based precision modulated attention model for robots, that can deal with high uncertainties from the real world. The overarching problem that is dealt here is the poor uncertainty handling capabilities of the robot algorithms. The solution is to develop a brain inspired robot attention model using FEP, to test it on real robotics problems to demonstrate its utility, and to show that it can outperform the classical methods.

> **Core contributions**
>
> The central contribution of this thesis is the introduction of a precision modulation based robot attention model using FEP for uncertainty resolution. Four types of precision modulations are proposed: i) precision modelling, ii) precision learning, iii) precision optimization and iii) precision fluctuation.

The central contribution of this thesis is the introduction of a precision modulation based

attention model for robots using FEP, as outlined in Chapter 2. Four types of precision modulations for robot attention are proposed in this thesis: i) precision modelling, ii) precision learning, iii) precision optimization and iii) precision fluctuation. The advantage of each model is examined in light of pertinent problems in robotics, with a special focus towards applications to UAVs. The theoretical framework is first formulated and then evaluated under rigorous simulation experiments, followed by experimental validation on real drone data. Figure 1.2 gives a quick overview of the thesis structure and the general storyline that holds all the chapters together under the hood of a precision based robot attention using FEP.
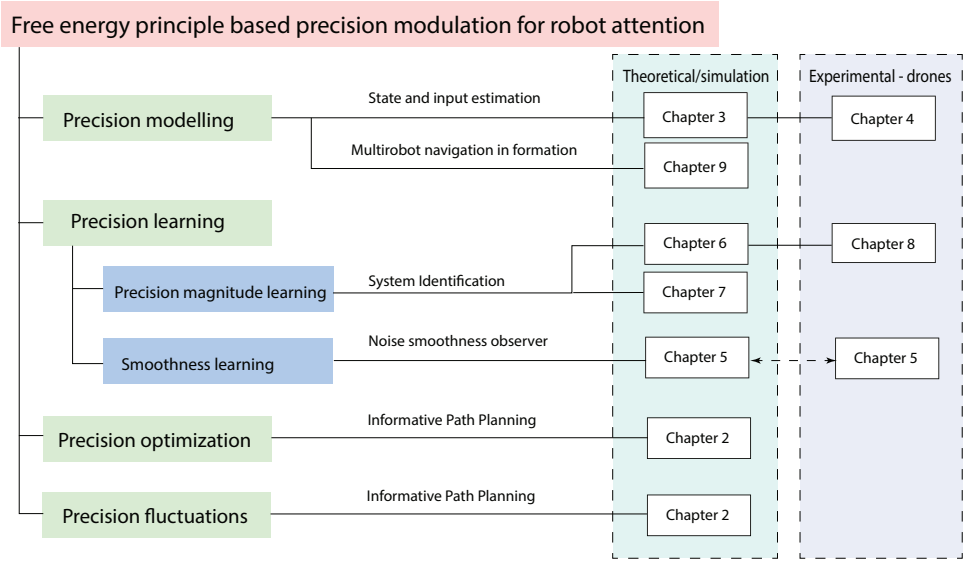
Fig 1.2: The chapter overview connecting all the chapters under the common hood of precision modulation for robot attention using FEP. Each chapter solves a relevant robotics problem using precision modulation.

1.3.1 Precision modelling

This relates to the user defined modelling of the structure of the precision matrix such that the robot attention is manipulated or weighed within the free energy scheme. The advantages of precision modelling for robotics is demonstrated by solving two practical robot problems: i) state and input estimation of an LTI system with colored noise and ii) multi robot navigation in formation. The core contributions of this thesis towards precision modelling for robot attention are:

1. This subcategory uses **precision modelling to handle uncertainties that appear as colored noise in a linear state space system** of the form:

$$\dot{x} = Ax + Bu + w, \ y = Cx + z, \tag{1.1}$$

where u is the input, y is the output, w and z are the process and measurement noises, and A, B and C are the system matrices. The chapter 3 contributes to the design of a state and input observer for LTI systems with colored noise. This is done by modelling the generalized noise precision matrix $\tilde{\Pi}^z = E[\tilde{z}\tilde{z}^T]^{-1}$, where $\tilde{z} = [z \ \dot{z} \ \ddot{z},...]^T$, using a smoothness matrix S such that $\tilde{\Pi}^z = S \otimes \Pi^z$, where Π^z is the measurement noise precision matrix (refer Chapter 3 for details). The modelling of the specific structure of the precision matrix (using the S matrix) to handle the colored noise aids the observer to attend to the most informative noise derivative information. This enables a better state and input estimation, as shown in Chapter 3. The chapter 4 introduces an experimental design to prove the validity of the observer design in chapter 3. It involves a quadcopter hovering in wind, generating the colored noise. The result shows that precision modelling to handle colored noise during the state and input estimation improves the estimation.

2. This subcategory uses a **precision based attention scheme to manipulate the behavior of a group of agents navigating in formation**. The chapter 9 introduces an active inference framework for the multi robot navigation in formation and demonstrates the use of precision modelling for changing the behavior of a group of agents navigating in formation. By modelling the precision associated with the prediction errors for goal directed behavior (Π^g), for obstacle avoidance (Π^o) and for formation (Π^f), the agents can alternate between a splitting and merging strategy or an obstacle avoidance strategy when facing an obstacle while navigating in formation.

1.3.2 Precision learning

The precision learning within the attention scheme is composed of two parts (Figure 1.2): i) learning the noise precision magnitude and ii) learning the noise smoothness (color). Its advantages are demonstrated within the context of two problems: i) system identification of LTI systems with colored noise and ii) noise smoothness learning. The core contributions of this thesis towards precision learning for robot attention are:

1. This subcategory uses **precision learning to handle the uncertainties that appear as colored noise in an LTI system** of the form given in Equation 1.1. The chapter 6 introduces a system identification tool for the estimation of an LTI under colored noise. This is done by learning the parameters of the system (A,B and C) while learning the noise precision (Π^z and Π^w) and the posterior precision of parameters. The estimator is supported for safe use on robots with the mathematical proof of convergence in chapter 7. The chapter 8 provides the experimental validation of the algorithm for the model learning of a quadrotor flying in wind. All three chapters together demonstrates the advantages of using an FEP based precision learning model for robot attention in the context of the estimation problem.

2. This subcategory uses **precision learning to learn the structure of uncertainties that appear as colored noise in an LTI system** of the form given in Equation 1.1. The chapter 5 introduces a novel online noise smoothness estimation algorithm for LTI systems with colored noise. The generalized noise precision matrix $\tilde{\Pi}^z = S \otimes \Pi^z$ has the smoothness matrix S that is parameterised using a smoothness parameter s

as:

$$S(s^2) = \begin{bmatrix} 1 & 0 & -\frac{1}{2s^2} & .. \\ 0 & \frac{1}{2s^2} & 0 & .. \\ -\frac{1}{2s^2} & 0 & \frac{3}{4s^4} & .. \\ .. & .. & .. & .. \end{bmatrix}_{(p+1)\times(p+1)}^{-1} , \tag{1.2}$$

where p is the number of derivatives of z under consideration. The chapter 5 provides the theoretical account for the online estimation of this parameter s, and applies it on the quadrotor flight data to show its effectiveness on practical robotics problems. The main advantage of the Gaussian convoluted white noise assumption for the colored noise is the easy parameterization of S matrix with just one parameter s. This also helps the colored noise modelling in the Chapters 3-8. The mathematical details of the derivation of Equation 1.2 can be found in [51].

1.3.3 Precision optimization

Precision optimization pertains to the use of precision as the objective function for uncertainty resolution in robots. Since the brain is efficient at uncertainty minimization, it is an obvious choice to use precision (inverse covariance) as the objective function for robots. A large class of robotics problems that seek to minimize uncertainty like SLAM, IPP etc falls under this category. The core contribution of this thesis towards using precision optimization for robot attention is:

1. This section uses **precision optimization to minimize the map uncertainty of a robot performing path planning for an information gathering task**. The section 2.5.3.2 presents the use of the map precision (the trace of inverse covariance matrix of an information map modelled using a Gaussian process) to develop a pure exploration based target search algorithm. This idea extends to a large class of information seeking robot behavior like mapping, exploration and search problems.

1.3.4 Precision fluctuation

Precision fluctuation pertains to the use of fluctuations in precision to change the robot attention. The core contributions of this thesis towards using precision fluctuation for robot attention are:

1. This subsection **introduces a rhythmic precision modulated attention model to solve the active perception problems like IPP for uncertainty resolution and information gathering**. The chapter 2 introduces a neurorobotic model of attention that proposes a rhythmically modulated precision to switch between action and perception, similar to how the eyes switch between saccades and information gathering during a visual search.

2. This subsection uses **precision fluctuation to trade-off between exploration-exploitation during system identification**. The section 2.5.2.4 proposes the implications of precision fluctuation (prior precision on parameters A, B and C) to balance between learning from data and adhering to the prior parameters, in the context of estimation for an LTI system.

In summary, this thesis explores a detailed account of the advantages of a precision modulation based robot attention model using FEP. With a hope that this research will take us a step closer to the grand theory for robot intelligence, the detailed chapters follow.

1.4 Free energy and ELBO

The FEP literature [2] treats the Evidence Lower Bound (ELBO) from Variational Bayesian methods [52] as the free energy objective. Here, the free energy is maximized the same way the ELBO term is maximized during inference (explained in detail in Section 6.5). The variational treatment of this objective results in the usage of the term "variational free energy" in Chapters 3-8, where the variational free energy is maximized during inference. However, the recent developments in literature uses the minimization of free energy (negative of the ELBO term), which is also used in Chapter 9. The reader is advised not to be confused with "free energy maximization" and "free energy minimization" throughout the thesis, as both free energy objectives are the same (one being the negative of the other).

1.5 Statement of contributions

This thesis primarily includes the verbatim copy of published peer reviewed papers, some of which were done in collaboration with other individuals and universities. This section lists down the author contributions for each chapter in this thesis.

1. As a first author, Ajith contributed to the robotics section of Chapter 2. Filip proposed the rhythmic precision model for the brain and Ajith translated it to robotics context. The robotics section is a compilation of Ajith's previous results from Chapter 3, 5, 6, 7, 8, and [53, 54], addressing the proposed neurorobotic attention model. Filip and Noor contributed to the neuroscience section. Pablo is behind the reclaiming saliency idea. Karl and Thomas contributed through manuscript editing and feedback. All authors contributed to the manuscript writing.

2. Ajith contributed to the experimental design, conceptualization and writing of Chapter 4. Under his supervision, Dennis conducted the experiment and Fred processed it to provide experimental confirmation for the previous simulation results in Chapter 3. Ajith and Fred contributed to the writing.

3. As the first author, Ajith was the main contributor to the Chapters 3, 5, 6, 7, 8 and 9. Martijn supervised the work and provided feedbacks to the draft.

4. Chapters 2, 3, 4, 5, 6 and 7 are a verbatim copy of peer reviewed papers published during the PhD. The paper contained in Chapter 8 is still under review. Chapter 9 contains the preliminary results of an upcoming paper. Parts of Chapter 3 and 6 were used to contribute to the paper [25].

2

Reclaiming Saliency: Rhythmic precision modulated active perception

Using inspirations from the neuroscience literature, this chapter proposes a neurorobot attention model - the backbone of this thesis. The utility of the model is demonstrated using pertinent problems in robotics like state and input estimation, system identification, noise estimation and informative path planning. The detailed mathematical account of the state and input observer used in Section 2.5.2.2 is provided in Chapter 3, and its robot implementation is provide in Chapter 4. The DEM based system identification tool used in Section 2.5.2.3, 2.5.2.4 and 2.5.2.5 is provided in Chapter 6 and 7, and its robot implementation is provided in Chapter 8. The Chapter 5 builds on the precision learning scheme proposed in this chapter for online noise smoothness estimation. The Chapter 9 builds on the precision modulation proposed in this chapter for the formation control of drones. The details of the simulation setup used for Section 2.5.3 are given in [53].

2.1 Abstract

Computational models of visual attention in artificial intelligence and robotics have been inspired by the concept of a saliency map. These models account for the mutual information between the (current) visual information and its estimated causes. However, they fail to consider the circular causality between perception and action. In other words, they do not consider where to sample next, given current beliefs. Here, we reclaim salience as an active inference process that relies on two basic principles: uncertainty minimisation and rhythmic scheduling. For this, we make a distinction between attention and salience. Briefly, we associate attention with precision control, i.e., the confidence with which beliefs can be updated given sampled sensory data, and salience with uncertainty minimisation that underwrites the selection of future sensory data. Using this, we propose a new account of attention based on rhythmic precision-modulation and discuss its potential in robotics, providing numerical experiments that showcase its advantages for state and noise estimation, system identification and action selection for informative path planning.

2.2 Introduction

Attention is a fundamental cognitive ability that determines which events from the environment, and the body, are preferentially processed [18]. For example, the motor system directs the visual sensory stream by orienting the fovea centralis (i.e., the retinal region of highest visual acuity) towards points of interest within the visual scene. Thus, the confidence with which the causes of sampled visual information are inferred is constrained by the physical structure of the eye – and eye movements are necessary to minimise uncertainty about visual percepts [56]. In neuroscience, this can be attributed to two distinct, but highly interdependent attentional processes: (*i*) attentional gain mechanisms reliant on estimating the sensory precision of current data [19, 57], and (*ii*) attentional salience that involves actively engaging with the sensorium to sample appropriate future data [20, 58]. Here we refer to perceptual-related salience, i.e., processing of low-level visual information [59]. Put simply, we formalise the fundamental difference between attention — as optimising perceptual processing — and salience as optimising the sampling of what is processed. This highlights the dynamic, circular nature with which biological agents acquire, and process, sensory information.

Understanding the computational mechanisms that undergird these two attentional phenomena is pertinent for deploying apt models of (visual) perception in artificial agents [60–62] and robots [63–66]. Previous computational models of visual attention, used in artificial intelligence and robotics, have been inspired (and limited) by the feature integration theory proposed by [67] and the concept of a saliency map [18, 68, 69]. Briefly, a saliency map is a static two-dimensional 'image' that encodes stimulus relevance, e.g., the importance of particular region. These maps are then used to isolate relevant information for control (e.g., to direct foveation of the maximum valued region). Accordingly, computational models reliant on this formulation do not consider the circular-dependence between action selection and cue relevance – and simply use these static saliency maps to guide action.

In this article, we adopt a first principles account to disambiguate the computational mechanisms that underpin attention and salience [20] and provide a new account of attention. Specifically, our formulation can be effectively implemented for robotic systems

and facilitates both state-estimation and action selection. For this, we associate attention with precision control, i.e., the confidence with which beliefs can be updated given (current) sampled sensory data. Salience is associated with uncertainty minimisation that influences the selection of future sensory data. This formulation speaks to a computational distinction between action selection (i.e., where to look next) and visual sampling (i.e., what information is being processed). Importantly, recent evidence demonstrates the rhythmic nature of these processes via a theta-cycle coupling that fluctuates between high and low precision—as unpacked in Sec 2.3. From a robotics perspective, resolving uncertainty about states of affair speaks to a form of Bayesian optimality, in which decisions are made to maximise expected information gain [70–72]. The duality between attention and salience is important for resolving uncertainty and enabling active perception. Significantly, it addresses an important challenge for defining autonomous robotics systems that can balance optimally between data assimilation (i.e., confidently perceiving current observations) and exploratory behaviour to maximise information gain [73].

In what follows, we review the neuroscience of attention and salience (Sec. 2.3) to develop a novel (computational) account of attention based on precision-modulation that underwrites perception and action (Sec. 2.4). Next, we face-validate our formulation within a robotics context using numerical experiments (Sec. 2.5). The robotics implementation instantiates a free energy principle (FEP) approach to information processing [2]. This allows us to modulate the (appropriate) precision parameters to solve relevant robotics challenges in perception and control; namely, state-estimation (Sec. 2.5.2.2), system identification (Sec. 2.5.2.3), planning (Sec. 2.5.3), and active perception (Sec. 2.5.3.3). We conclude with a discussion of the requisite steps for instantiating a full-fledged computational model of precision-modulated attention – and its implications in a robotics setting.

2.3 ATTENTION AND SALIENCE IN NEUROSCIENCE

Our interactions with the world are guided by efficient gathering and processing of sensory information. The quality of these acquired sensory data is reflected in attentional resources that select sensations which influence our beliefs about the (current and future) states of affairs [58, 74]. This selection is often related to gain control, i.e., an increase of neural spikes when an object is attended to. However, gain control only accounts for half the story because we can only attend to those objects that are within our visual field. Accordingly, if a salient object is outside the centre of our visual field, we orient the fovea to points of interest. This involves two separate, but often conflated, processes: attention and salience – where the former relates to processing current visual data, and the latter to ensuring the agent samples salient data in the future [20]. That these two processes are strongly coupled is exemplified by the pre-motor theory of attention [75], which highlights the close relationship between overt saccadic sampling of the visual field and the covert deployment of attention in the absence of eye movements. Specifically, it posits that covert attention[1] is realised via processes that are generated by particular eye movements but inhibits the action itself. In this sense, it does not distinguish between covert and overt[2] types of attention.

---

[1] Covert attention is where saccadic eye movements do not occur.
[2] Overt attention deals with how an agent tracks the object with eye movements

From a first principles (Bayesian) account, it is necessary to separate between attention and salience because they speak to different optimisation processes. Explicitly, attention as a precision-dependent (neural) gain control mechanism that facilitates optimisation of the *current* sampled sensory data [19, 76]. Conversely, salience is associated with selection of *future* data that reduces uncertainty [20, 77, 78]. Put simply, it is possible to optimise attention in the absence of eye movements and active vision, whereas salience is necessary to optimise the deployment of eye movements. In what follows, we formalise this distinction with a particular focus on visual attention [79], and discuss recent findings that speak to a rhythmic coupling that underwrites periodic deployment of gain control and saccades, via modulation of distinct precision parameters.

2.3.1 Attention as neural gain control

Neural gain control can be regarded as an amplifier of neural communication during attention tasks [80, 81]. Computationally, this is analogous to modulating a precision term, or the inverse temperature parameter [19, 82]. For this reason, we refer to precision and gain control interchangeably. An increase in gain amplifies the postsynaptic responses of neurons to their pre-synaptic input. Thus, gain control rests on synaptic modulation that can emphasise — or preferentially select — a particular type of sensory data. From a Bayesian perspective [83–85], this speaks to the confidence with which beliefs can be updated given sampled sensory data (i.e., optimal state estimation) – under a generative model [85, 86]. For example, affording high precision to certain sensory inputs would lead to confident Bayesian belief updating. However, low precision reduces the influence of sensory input by attenuating the precision of the likelihood, relative to a prior belief, and current observations would do little to resolve ensuing uncertainty. Thus, sampled visual data (from different areas) can be predicted with varying levels of precision, where attention accentuates sensory precision. The deployment of precision or attention is influenced by competition between stimuli (i.e., which sensory data to sample) and prior beliefs. Interestingly, casting attention as precision or, equivalently, synaptic gain offers a coherency between biased competition [76], predictive coding [84] and generic active inference schemes [19, 85, 87, 88].

Naturally, gain control is accompanied by neuronal variability, i.e., sharpened neural responses for the same task over time. Consistent with gain control, these fluctuations in neural responses across trials can be explained by precision engineered message passing [89] via (*i*) normalization models [90, 91], (*ii*) temperature parameter manipulation [19, 82, 85, 92, 93], or (*iii*) introduction of (conjugate hyper-)priors that are either pre-specified [94, 95] or optimised using uninformed priors [96, 97]. Recently, these approaches have been used to simulate attention by accentuating predictions about a given visual stimulus [19, 90, 91]. For example, normalization models propose that every neuronal response is normalized within its neuronal ensemble (i.e., the surrounding neuronal responses) [98, 99]. Thus, to amplify the neuronal response of particular neuron, the neuronal pool has to be inhibited such that that particular neuron has a sharper evoked response [100]. Importantly, these (superficially distinct) formulations simulate similar functions using different procedures to accentuate responses over a particular neuronal pool for a given neuron or a group of neurons. This introduces shifts in precision to produce attentional gain and the precision of neuronal encoding.

2

2.3.2 SALIENCE AS UNCERTAINTY MINIMISATION

In the neurosciences, (visual) salience refers to the 'significance' of particular objects in the environment. Salience often implicates the superior colliculus, a region that encodes eye movements [101]. This makes intuitive sense, as the superior colliculus plays a role in generation of eye movements – being an integral part of the brainstem oculomotor network [102] – and salient objects provide information that is best resolved in the centre of the visual field, thus motivating eye movements to that location. For this reason, our understanding of salience is a quintessentially action-driving phenomenon [20]. Mathematically, salience has been defined as Bayesian surprise [18, 103], intrinsic motivation [104], and subsequently, epistemic value under active inference [85, 105]. Active inference – a Bayesian account of perception and action [106, 107] – stipulates that action selection is determined by uncertainty minimisation. Formally, uncertainty minimisation speaks to minimisation of an expected free energy functional over future trajectories [72, 107]. This action selection objective can be decomposed into epistemic and extrinsic value, where the former pertains to exploratory drives that encourage resolution of uncertainty by sampling salient observations, e.g., only checking one's watch when one does not know the time. However, after checking the watch there is little epistemic value in looking at it again. Generally, the tendency to seek out new locations – once uncertainty has been resolved at the current fixation point – is called inhibition of return [108].

From an active inference perspective, this phenomenon is prevalent because a recent action has already resolved the uncertainty about the time and checking again would offer nothing more in terms of information gain [20]. Accordingly, salience involves seeking sensory data that have a predictable, uncertainty reducing, effect on current beliefs about states of affairs in the world [85, 105]. Thus salience contends with beliefs about data that must be acquired and the precision of beliefs about policies (i.e., action trajectories) that dictate it. Formally, this emerges from the imperative to maximise the amount of information gained regarding beliefs, from observing the environment. Happily, prior studies have made the connection between eye movements, salience, and precision manipulation [87, 109, 110]. This connection emerges from planning strategies that allow the agent to minimise uncertainty by garnering the right kind of data.

Next, we consider recent findings on how the coupling of these two mechanisms, attention and salience, may be realised in the brain.

2.3.3 RHYTHMIC COUPLING OF ATTENTION AND SALIENCE

To illustrate the coupling between attention and salience, we turn to a recent rhythmic theory of attention. The theory proposes that coupling of saccades, during sampling of visual information, happens at neuronal and behavioural theta oscillations; a frequency of 3-8Hz [111, 112]. This frequency simultaneously allows for: (*i*) a systematic integration of visual samples with action, and (*ii*) a temporal schedule to disengage and search the environment for more relevant information.

Given that gain control is related to increased sensory precision, we can accordingly relate saccadic eye movements to the decreased precision. This introduces saccadic suppression, a phenomenon that decreases visual gain during eye movements [110]. This phenomenon was described by Helmholtz who observed that externally initiated eye movements (e.g., when oneself gently presses a side of an eye) eludes the saccadic suppression

2

that accompanies normal eye movements – and we see the world shift, because optic flow is not attenuated [113]. An interesting consequence of this is that, as eye movements happen periodically [114, 115], there must be a periodic switch between high and low sensory precision, with high precision (or enhanced gain) during fixations and low precision (or suppressed gain) during saccades. Interestingly, it has been shown that rather than having action resetting the neural periodicity, it is better understood as something that aligns within an already existing rhythm [116, 117]. Additionally, the rhythmicity of higher and lower fidelity of sensory sampling has been shown to fluctuate rhythmically around 3Hz [118], suggesting that action emerges rhythmically when visual precision is low [117], triggering salience.

Building upon this, we hypothesise that theta rhythms generated in the fronto-parietal network [119–121] couples saccades with saccadic suppression causing the switches between visual sampling and saccadic shifting. This introduces a diachronic aspect to the belief updating process [122–124]; i.e., sequential fluctuations between attending to current data (perception) and seeking new data (action). This supports empirical findings that both eye movements [125] and filtering irrelevant information [121, 126, 127] are initiated in this cortical network. Interestingly, both eye movements and visual filtering then propagate to sub-cortical regions, i.e., the superior colliculus—for saliency map composition [101]—and the thalamus—for gain control [88, 128], respectively. Furthermore, this is consistent with recent findings that the periodicity of neural responses are important for understanding the relation of motor responses and sensory information – i.e., perception-action coupling [115]. Importantly, theta rhythms also speak to the speed (i.e., the temporal schedule) with which visual information is sampled from the environment [120, 129–131]. Meaning visual information is not sampled continuously, as our visual experiences would suggest, but rather it is made of successive discrete samples [132, 133].

The prefrontal theta rhythm has been associated with working memory (WM), a process that holds compressed information about the previously observed stimuli, in the sense that measured power in this frequency range using electroencephalography increases during tasks that place demands on WM [134–140]. The implication is that the neural processes that underwrite WM may depend upon temporal cycles with periods similar to that of perceptual sampling. Importantly, this cognitive process is influenced by how salient a particular stimulus was [141–143]. Moreover, WM has been implicated with attentional mechanisms [139, 144–147]. This is aligned with our account where we illustrate a rhythmic coupling between salience and attention.

In summary, the computations that underwrite attention and active vision are coupled and exhibit circular causality. Briefly, selective attention and sensory attenuation optimize the processing of sensory samples and which particular visual percepts are inferred. In turn, this determines appropriateness of future eye movements (or actions) and shapes which prior stimuli are encoded into the agent's working memory. Interestingly, the close functional (and computational) link between the two mechanisms endorses the pre-motor theory of attention.
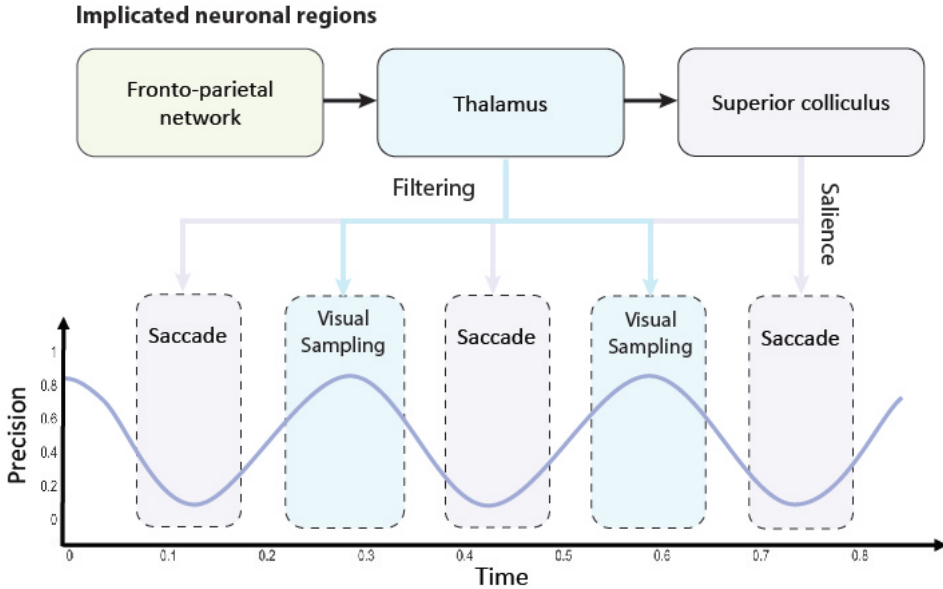
Fig 2.1: A graphical illustration of the precision-modulated account of perception and action. Salience and attention are computed based upon beliefs (assumed to be) encoded in parts of the fronto-parietal network and realised in distinct brain regions: superior colliculus (SC) for perception as inference and thalamus for planning as inference, respectively. To deploy attentional processes efficiently, these two mechanisms have to be aligned, which is done rhythmically, hypothetically in theta frequency. This coupling enables the saccadic suppression phenomenon through fluctuations in precision (on an arbitrary scale). When precision is low (i.e., the trough of the theta rhythm), the saccade emerges. Note that there might be distinct processes inhibiting the action (e.g., covert attention), and (despite a decline in precision) saccades might not emerge in every theta cycle. On the other hand, high precision facilitates confident inferences about the causes of visual data. Under this account, thalamus is used for initiating gain control (or visual sampling in general) by providing stronger sensory input, while superior colliculus dictates next saccades, that lead to most informative fixation positions.

2.4 Proposed precision-modulated account of attention and salience

Here, we introduce our precision-modulated account of perception and action. A graphical illustration is provided in Figure 2.1. For this, we turn to attention and salient action selection which have their roots in biological processes relevant for acquiring task-relevant information. Under an active inference account, this attention influences (posterior) state estimation and can be associated with increased precision of belief updating and gain control—described in Sec. 2.3.1. Furthermore, this is distinct from salience despite interdependent neuronal composition and computations.

Further alignment between the two constructs can be revealed by considering the temporal scheduling between movement (i.e., action) and perception for uncertainty res-

2

olution [20]. We postulate that this perception-action coupling is best understood as a periodic fluctuation between minimising uncertainty and precision control. Subsequently, action is deployed to reduce uncertainty. Such an alignment specifies what stimulus is selected and under what level of precision it is processed. . [20] hypothesise that action alignment with precision is due to the eye structure that provides precise information in the fovea and requires the agent to foveate the most informative stimulus. We extend this by considering the periodic deployment of gain control with saccades [111, 116–118, 148].

Accordingly, our formulation defines attention as precision control and salience as uncertainty minimisation supported by discrete sampling of visual information at a theta rhythm. This synchronises perception and action together in an oscillatory fashion [117]. Importantly, a Bayesian formulation of this can be realised as precision manipulation over particular model parameters. We reserve further details for Sec. 2.5.

Summary Based upon our review, we propose a precision-modulated account of attention and salience, emphasising the diachronic realisation of action and perception. In the following sections, we investigate a realisation of this model for a robotic system.

2.5 Precision-based attention for Robotics

The previous section introduced a conceptual account to explain the computational mechanisms that undergird attention based on neuroscience findings. We focused on reclaiming saliency as an active process that relies on neural gain control, uncertainty minimisation and structured scheduling. Here, we describe how we can mathematically realise some of these mechanisms in the context of well-known challenges in robotics. Enabling robots with this type of attention may be crucial to filter the sensory signals and internal variables that are relevant to estimate the robot/world state and complete any task. More importantly, the active component of salience (i.e., behaviour) is essential to interact with the world—as argued in active perception approaches [73].

We revisit the standard view of attention in robotics by introducing sensory precision (inverse variance) as the driving mechanism for modulating both perception and action [89, 109]. Although saliency was originally described to underwrite behaviour, most models used in robotics, strongly biased by computer vision approaches, focus on computing the most relevant region of an image [69]—mainly computing human fixation maps— relegating action to a secondary process. Illustratively, state-of-the-art deep learning saliency models—as shown in the MIT saliency benchmark [149]—do not have the action as an output. Conversely, the active perception approach properly defines the action as an essential process of active sensing to gather the relevant information. Our proposed model, based on precision modulated action and perception coupling (*i*) place attention as essential for state-estimation and system identification and (*ii*) and reclaims saliency as a driver for information-seeking behaviour, as proposed in early works [68], but goes beyond human fixation maps for both improving the model of the environment (exploration) and solving the task (exploitation).

In what follows, we highlight the key role of precision by reviewing relevant brain-inspired attention models deployed in robotics (Sec. 2.5.1). We propose precision-modulated attentional mechanisms for robots in three contexts - perception (Sec. 2.5.2), action (Sec. 2.5.3) and active perception (Sec. 2.5.3.3). The precision-modulated perception is formalised for a robotics setting; via (*i*) state estimation (i.e., estimating the hidden states of a dynamic

Table 2.1: Robotics applications and their precision realisations.

Task	Application	Precision manipulation	Section
Perception	State & input estimation	Noise precision modelling $\tilde{\Pi}$	2.5.2.2
	System Identification	Posterior parameter precision learning Π^θ	2.5.2.3
	Exploration-exploitation in learning	Prior parameter precision modelling P^θ	2.5.2.4
	Noise estimation	Noise precision learning $\tilde{\Pi}$	2.5.2.5
Action	Informative Path Planning (IPP)	Precision optimisation (of map)	2.5.3.2
Active perception	IPP with action-perception cycle	Precision modulation	2.5.3.3

system from sensory signals – Sec. 2.5.2.2), and (*ii*) system identification (i.e., estimating the parameters of the dynamic system from sensory signals – Sec. 2.5.2.3). Next, we show that precision-modulated action can be realised through precision optimisation (planning future actions – Sec. 2.5.3.2) and discuss practical considerations for coupling with precision-modulated perception (precision based active perception - Sec. 2.5.3.3). Table 2.1 summarises our proposed precision manipulations to solve relevant problems in robot perception and action. Table 2.2 provides the definitions of precision within our mechanism.

Table 2.2: Precision parameters that are manipulated in Sec.2.5.2

Term	Symbol	Definition
Sensory precision	Π^z	Inverse covariance of sensory noise $\mathbf{z}$ (Eqn. 2.1).
Prior parameter precision	P^θ	The robot's confidence on its prior parameters η^θ.
Noise precision	$\tilde{\Pi}$	The inverse covariance of all noises (Eqn. 2.5).
Posterior parameter precision	Π^θ	The robot's confidence on its parameter estimates.

2.5.1 Previous brain-inspired attention models in robotics

Brain-inspired attention has been mainly addressed in robotics from a 'passive' visual saliency perspective, e.g., which pixels of the image are the most relevant. This saliency map is then generally used to foveate the most salient region. This approach was strongly influenced by early computational models of visual attention [18, 68]. The first models deployed in robots were bottom-up, where the sensory input was transformed into an array of values that represents the importance (or salience) of each cue. Thus, the robot was able to identify which region of the scene has to look at, independently of the task performed—see [69] for a review on visual saliency. These models have also been useful for acquiring meaningful visual features in applications, such as object recognition [150, 151], localisation, mapping and navigation [63, 152, 153]. Saliency computation was usually employed as a helper for the selection of the relevant characteristics of the environment to be encoded. Thus, reducing the information needed to process.

More refined methods of visual attention employed top-down modulation, where the context, task or goal bias the relevance of the visual input. These methods were used, for instance, to identify humans using motion patterns [154, 155]. A few works also focused on object/target search applications, where top-down and bottom-up saliency attention

2

were used to find objects or people in a search and rescue scenario [156].

Attention has also been considered in human-robot interaction and social robotics applications [65], mainly for scene or task understanding [157–159], and gaze estimation [160] and generation [66]. For instance, computing where the human is looking at and where the robot should look at or which object should be grasped. Furthermore, multi-sensory and 3D saliency computation has also been investigated [161]. Finally, more complex attention behaviours, particularly designed for social robotics and based on human non-verbal communication, such as joint attention, have also been addressed. Here the robot and the human share the attention of one object through meaningful saccades, i.e., head/eye movements [66, 162, 163].

Although attention mechanisms have been widely investigated in robotics, specially to model visual cognition [64, 158], the majority of the works have treated attention as an extra feature that can help the visual processing, instead of a crucial component needed for the proper functioning of the cognitive abilities of the robot [164]. Furthermore, these methods had the tendency to leave the action generation out of the attention process. One of the reasons for not including saliency computation, in robotic systems, is that the majority of the models only output 'human-fixation map' predictions, given a static image. Saliency computation introduces extra computational complexity, which can be finessed by visual segmentation algorithms (e.g., line detectors in autonomous navigation). However, it does not resolve uncertainty nor select actions that maximise information gain in the future. In essence, the incomplete view of attention models that output human-fixation maps has arguably obscured the huge potential of neuroscience-inspired attentional mechanisms for robotics.

Our proposed model of attention, based on precision modulation, abandons the current robotics narrow view of attention and saliency by explicitly modelling attention within state estimation, learning and control. Thus, placing attentional processes at the core of the robot computation and not as an extra add-on. In the following sections, we describe the realisation of our precision-based attention formulation in robotics using common practical applications as the backbone motif.

2.5.2 Precision-modulated perception

We formalise precision-modulated perception from a first principles Bayesian perspective – explicitly the free energy principle approach proposed by [109]. Practically, this entails optimising precision parameters over (particular) model parameters.

Through numerical examples show how our model is able to perform accurate state estimation [165] and stable parameter learning [166, 167]. To illustrate the approach, we first introduce a dynamic system modelled as a linear state space system in robotics (Sec. 2.5.2.1)—we used this formulation in all our numerical experiments. We briefly review the formal terminologies for a robotics context to appropriately situate our precision-based mechanism for perception. Explicitly, we introduce: precision modelling (by adapting a known form of the precision matrix), precision learning (by learning the full precision matrix), and precision optimisation (use precision as an objective function during learning). As a reminder, precision modelling is associated with (instantaneous) gain control and precision learning (at slower time scales) is associated with optimising that control.

2.5.2.1 Precision for state space models

A linear dynamic system can be modelled using the following state space equations (boldface notation denotes components of the real system and non-boldface notation its estimates):

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{w}, \qquad \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{z}. \tag{2.1}$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are constant matrices defining the system parameters, $\mathbf{x} \in \mathbb{R}^n$ is the system state (usually an unobserved variable), $\mathbf{u} \in \mathbb{R}^r$ is the input or control actions, $\mathbf{y} \in \mathbb{R}^m$ is the output or the sensory measurements, $\mathbf{w} \in \mathbb{R}^n$ is the process noise with precision $\mathbf{\Pi}^{\mathbf{w}}$ (or inverse variance $\mathbf{\Sigma}^{\mathbf{w}-1}$), and $\mathbf{z} \in \mathbb{R}^m$ is the measurement noise with precision $\mathbf{\Pi}^{\mathbf{z}}$.

For instance, we can describe a mass-spring damper system (depicted in Fig. 2.2b) using state space equations. A mass ($m = 1.4kg$) is attached to a spring with elasticity constant ($k = 0.8N/m$), and a damper with a damping coefficient ($b = 0.4Ns/m$). When a force ($u(t) = e^{-0.25(t-12)^2}$) is applied on the mass, it displaces x from its equilibrium point. The linear dynamics of this system is given by:

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u, \qquad y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix}. \tag{2.2}$$

Note that Eq. (2.2) is equivalent to Eq. (2.1) with parameters $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0, \frac{1}{m} \end{bmatrix}^T$ and $\mathbf{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}$, and state $\mathbf{x} = \begin{bmatrix} x, \dot{x} \end{bmatrix}^T$.

Now we introduce attention as precision modulation assuming that the robotic goal is to minimise the prediction error [46, 109, 168], i.e., to refine its model of the environment and perform accurate state estimation, given the information available. In other words, the robot has to estimate $\mathbf{x}$ and $\mathbf{u}$ from input prior η^u with a prior precision of P^u, given the measurements $\mathbf{y}$, parameters $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and noise precision $\mathbf{\Pi}^{\mathbf{w}}$ and $\mathbf{\Pi}^{\mathbf{z}}$. Formally, the prediction error $\tilde{\epsilon}$ of the sensory measurements $\tilde{\epsilon}^y$, control input reference $\tilde{\epsilon}^u$ and state $\tilde{\epsilon}^x$ are:

$$\tilde{\epsilon} = \begin{bmatrix} \tilde{\epsilon}^y \\ \tilde{\epsilon}^u \\ \tilde{\epsilon}^x \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{y}} - \tilde{\mathbf{C}}\tilde{x} \\ \tilde{u} - \tilde{\eta}^u \\ D^x \tilde{x} - \tilde{\mathbf{A}}\tilde{x} - \tilde{\mathbf{B}}\tilde{u} \end{bmatrix} \quad \begin{cases} \text{sensory prediction error} \\ \text{control input prediction error} \\ \text{state prediction error} \end{cases} \tag{2.3}$$

Note that $\tilde{\epsilon}^y = \tilde{\mathbf{y}} - \tilde{\mathbf{C}}\tilde{x}$ is the difference between the observed measurement and the predicted sensory input given the state[3]. Here D^x performs the (block) derivative operation, which is equivalent to shifting up all the components in generalised coordinates by one block.

---

[3] The tilde over the variable refers to the generalised coordinates, i.e., the variable includes all temporal derivatives. Thus, $\tilde{\epsilon}$ is the combined prediction error of outputs, inputs and states. For example, the generalised output $\tilde{y}$ is given by $\tilde{\mathbf{y}} = [y, y', y'' ...]^T$, where the prime operator denotes the derivatives. We use generalised coordinates [169] for achieving accurate state and input estimation during the presence of (coloured) noise by modelling the time dependent quantities (x, v, y, w, z) in generalised coordinates. This involves keeping track of the evolution of the trajectory of the probability distributions of states, instead of just their point estimates. Here the coloured noise w and z are modelled as a white noise convoluted with a Gaussian kernel. The use of generalised coordinates has recently shown to outperform classical approaches under coloured noise on real quadrotor flight [165]

We can estimate the state and input using the Dynamic Expectation Maximisation (DEM) algorithm [46, 170] that optimises a free energy variational bound $\mathcal{F}$ to be tractable[4]. This is:

$$X = \begin{bmatrix} \tilde{x} \\ \tilde{u} \end{bmatrix} = \arg\max_X \mathcal{F} = \arg\max_X -\frac{1}{2}\tilde{\epsilon}^T \tilde{\Pi}\tilde{\epsilon} \tag{2.4}$$

Crucially, $\tilde{\Pi}$ is the generalised noise precision that modulates the contribution of each prediction error to the estimation of the state and the computation of the action. Thus, $\tilde{\Pi}$ is equivalent to attentional gain. For instance, we can model the precision matrix to attend to the most informative signal derivatives in $\tilde{\mathbf{y}}$. Concisely, the precision $\tilde{\Pi}$ has the following form:

$$\tilde{\Pi} = \begin{bmatrix} S \otimes \Pi^z & 0 & 0 \\ 0 & S \otimes P^u & 0 \\ 0 & 0 & S \otimes \Pi^w \end{bmatrix}, \tag{2.5}$$

where S is the smoothness matrix. In Sec. 2.5.2.2, we show that modelling the precision matrix $\tilde{\Pi}$ using the S matrix improves the estimation quality.

The full free energy functional (time integral of free energy $\bar{\mathcal{F}} = \int \mathcal{F}\,dt$ at optimal precision) that the robot optimises to perform state-estimation and system identification is described in Eq. (2.6)—for readability we omitted the details of the derivation of this cost function, and we refer to [97] for further details.

$$\bar{\mathcal{F}} = -\frac{1}{2}\sum_t \Big[\underbrace{\tilde{\epsilon}^{yT}\tilde{\Pi}^z\tilde{\epsilon}^y + \tilde{\epsilon}^{uT}P^{\tilde{u}}\tilde{\epsilon}^u + \tilde{\epsilon}^{xT}\tilde{\Pi}^w\tilde{\epsilon}^x}_{\text{precision weighed prediction error}} \Big] - \frac{1}{2}\Big[\underbrace{\epsilon^{\theta T}P^\theta\epsilon^\theta + \epsilon^{\lambda T}P^\lambda\epsilon^\lambda}_{\text{prior precision weighed prediction error of }\theta\text{ and }\lambda} \Big]$$

$$+ \underbrace{\frac{1}{2}n_t\ln|\Sigma^X|}_{\text{state and input entropy}} + \underbrace{\frac{1}{2}n^t\big[\ln|\tilde{\Pi}^z| + \ln|P^{\tilde{v}}| + \ln|\tilde{\Pi}^w|\big]}_{\text{noise entropy}} + \underbrace{\frac{1}{2}\ln|\Sigma^\theta P^\theta|}_{\text{parameter entropy}} + \underbrace{\frac{1}{2}\ln|\Sigma^\lambda P^\lambda|}_{\text{hyperparameter entropy}} \tag{2.6}$$

Here $\epsilon^\theta = \theta - \eta^\theta$, $\epsilon^\lambda = \lambda - \eta^\lambda$ are the prediction errors of parameters and hyper-parameters[5]. $\bar{\mathcal{F}}$ consist of two main components: i) precision weighed prediction errors and ii) precision-based entropy. The dominant role of precision — in the free energy objective –- is reflected in how modulating these precision parameters can have a profound influence perception and behaviour. The theoretical guarantees for stable estimation [167], and its application on real robots [25] make this formulation very appealing to robotic systems.

Note that we can manipulate three kinds of precision within the state space formulation: i) prior precision ($P^{\tilde{u}}, P^\theta, P^\lambda$), ii) conditional precision on estimates ($\Pi^X, \Pi^\theta, \Pi^\lambda$) and iii) noise precision (Π^z, Π^w). Therefore, to learn the correct parameter values θ, we i) learn the parameter precision Π^θ, ii) model the prior parameter precision P^θ, and iii) learn the noise precision Π^w and Π^z (parameterised using λ).

---

[4]Note that this expression of the variational free energy is using the Laplace and mean-field approximations commonly used in the FEP literature

[5]System identification involves the estimation of system parameters (denoted by θ, e.g., vectorised $\mathbf{A}$), given $\mathbf{y}, \mathbf{u}$, by starting from a parameter prior of η^θ with prior precision P^θ, and a prior on noise hyper-parameter η^λ with a prior precision of P^λ. Note that we parametrise noise precision (Π^w and Π^z) using $\lambda \in \mathbb{R}^{2\times 1} = \begin{bmatrix} \lambda^z \\ \lambda^w \end{bmatrix}$ as an exponential relation (e.g., $\Pi^w(\lambda^w) = \exp(\lambda^w)I^{n\times n}$).

2.5.2.2 State and input estimation

State estimation is the process of estimating the unobserved states of a real system from (noisy) measurements. Here, we show how we can achieve accurate estimation through precision modulation in a linear time invariant system under the influence of coloured noise [46]. State estimation in the presence of coloured noise is inherently challenging, owing to the non-white nature of the noise, which is often ignored in conventional approaches, such as the Kalman Filter [171].

Figure 2.2 summarises a numerical example that shows how one can use precision modulation to focus on the less noisy derivatives (lower derivatives) of measurements, relative to imprecise higher derivatives. Thus, enabling the robot to use the most informative data for state and input estimation, while discarding imprecise input. Figure 2.2b depicts the mass-spring damper system used. The numerical results show that the quality of the estimation increases as the embedding ordering increases but the lack of information in the higher order derivatives of the sensory input do not affect the final performance due to the precision modulation. The higher order derivatives (Fig. 2.2a) are less precise than the lower derivatives, thereby reflecting the loss of information in higher derivatives. The state and input estimation was performed using the optimisation framework described in the previous section. The quality of estimation is shown in Fig. 2.2c, where the input estimation using six derivatives (blue curve) is closer to the real input (yellow curve) than when compared to the estimation using only one derivative (red curve). The quality of the estimation reports the sum of squared error (SSE) in the estimation of states and inputs with respect to the embedding order (number of signal derivatives considered).
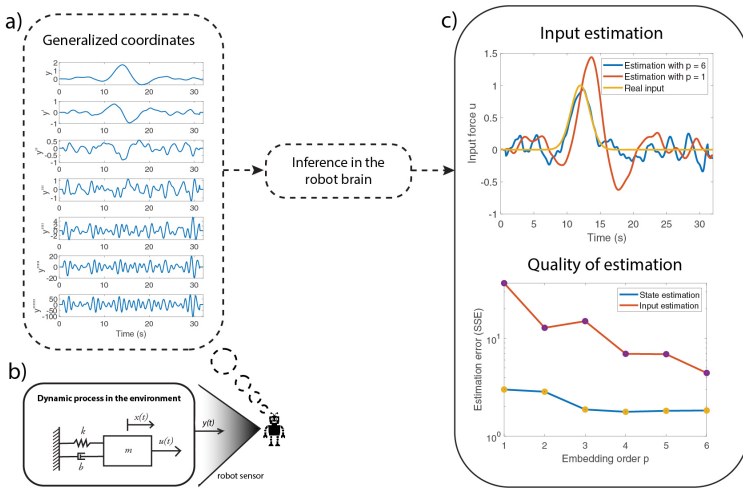
Fig 2.2: An illustration of an attention mechanism for state and input estimation of a system (shown in B). The quality of the estimation improves (C) as the embedding order (number of derivatives) of generalized coordinates are increased (A). However, the imprecise information in the higher order derivatives of the sensory input **y** does not affect the final performance of the observer because of attentional selection, which selectively weighs the importance afforded to each derivative, in the free energy optimization scheme.

2

To obtain accurate state estimation by optimising the precision parameters, we recall that the precision weights the prediction errors. From Eq. (2.3), the structural form of $\tilde{\Pi}$ is mainly dictated by the smoothness matrix S, which establishes the interdependence between the components of the variable expressed in generalised coordinates (e.g., the dependence between $\mathbf{y}$, $\mathbf{y}'$ and $\mathbf{y}''$ in $\tilde{\mathbf{y}}$). For instance, the S matrix for a Gaussian kernel is as follows [172]:

$$
S = \begin{bmatrix}
\frac{35}{16} & 0 & \frac{35}{8}s^2 & 0 & \frac{7}{4}s^4 & 0 & \frac{1}{6}s^6 \\
0 & \frac{35}{4}s^2 & 0 & 7s^4 & 0 & s^6 & 0 \\
\frac{35}{8}s^2 & 0 & \frac{77}{4}s^4 & 0 & \frac{19}{2}s^6 & 0 & s^8 \\
0 & 7s^4 & 0 & 8s^6 & 0 & \frac{4}{3}s^8 & 0 \\
\frac{7}{4}s^4 & 0 & \frac{19}{2}s^6 & 0 & \frac{17}{3}s^8 & 0 & \frac{2}{3}s^{10} \\
0 & s^6 & 0 & \frac{4}{3}s^8 & 0 & \frac{4}{15}s^{10} & 0 \\
\frac{1}{6}s^6 & 0 & s^8 & 0 & \frac{2}{3}s^{10} & 0 & \frac{4}{45}s^{12}
\end{bmatrix},
\tag{2.7}
$$

where s is the kernel width of the Gaussian filter that is assumed to be responsible for serial correlations in measurement or state noise. Here, the order of generalised coordinates (number of derivatives under consideration) is taken as six ($S \in \mathbb{R}^{7 \times 7}$). For practical robotics applications, the measurement frequency is high, resulting in $0 < s < 1$. It can be observed that the diagonal elements of S decreases because $s < 1$, resulting in a higher attention (or weighting) on the prediction errors from the lower derivatives when compared to the higher derivatives. The higher the noise colour (i.e., s increases), the higher the weight given to the higher state derivatives (last diagonal elements of S increases). This reflects the fact that smooth fluctuations have more information content in their higher derivatives. Having established the potential importance of precision weighting in state estimation, we now turn to the estimation (i.e., learning) of precision in any given context.

2.5.2.3 System identification

This section shows how to optimise system identification by means of precision learning [97, 167]. Specifically, we show how to fuse prior knowledge about the dynamic model with the data to recover unknown parameters of the system through an attention mechanism. This involves the learning of the 1) parameters and 2) noise precisions. Our model 'turns' the attention to the least precise parameters and uses the data to update those parameters to increase their precision. Hence, allowing faster parameter learning.

For the sake of clarity, we use again the mass-spring-damper system as the driving example (Sec. 2.5.2.1). We formalise system identification as evaluating the unknown parameters k, m and b, given the input $\mathbf{u}$, the output $\mathbf{y}$, and the general form of the linear system in Eq. (2.2).

Figure 2.3 depicts the process of learning unknown parameters (dotted boxes denote the processes inside the robot brain). The robot measures its position $x(t)$ using its sensors (e.g., vision or range sensor). We assume that the robot has observed the behaviour of a mass-spring-damper system before or a model is provided by the expert designer. However, some of the parameters are unknown. The robot can reuse the prior learned model of the system to relearn the new system. This can be realised by setting a high prior precision on the known parameters and a low prior precision on the unknown parameters. By means of precision learning, the robot uses the sensory signals to learn the parameter precision Π^θ,
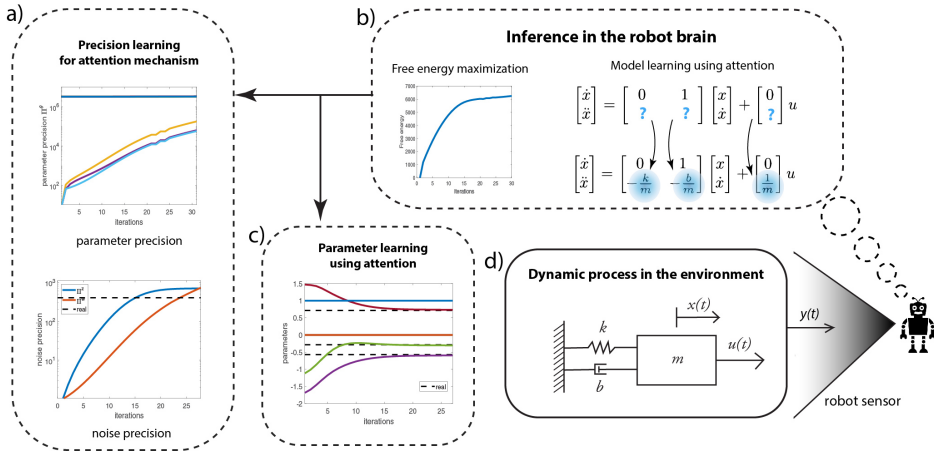
Fig 2.3: The schematic of the robot's attention mechanism for learning the least precise parameters of a given generative model of a mass-spring-damper system (shown in D). (A) Learning the conditional precision on parameters and the noise precision. (B) The free energy optimization helping to identify the unknown system parameters. (C) The parameter learning.

thereby improving the confidence in the parameter estimates θ. This directs the robot's attention towards the refinement of the parameters with least precision as they are the most uncertain. The requisite parameter learning proceeds by the gradient ascent of the free energy functional given in Eq. (2.6). The parameter precision learning proceeds by tracking the negative curvature of $\bar{\mathcal{F}}$ as $\Pi^\theta = -\frac{\partial^2 \bar{\mathcal{F}}}{\partial \theta^2}$ [97].

The learning process – by means of variational free energy optimisation (maximisation) – is shown in Fig. 2.3b. The learning involves two parallel processes: precision learning (Fig. 2.3a), and parameter learning (Fig. 2.3c). Precision learning comprises of parameter precision learning (top graph) – i.e., identifying the precision of an approximate posterior density for the parameters being estimated – and noise precision learning (bottom graph). The high prior precision on the known system parameters (0 and 1), and low prior precision on the unknown system parameters ($-\frac{k}{m}, -\frac{b}{m}$ and $\frac{1}{m}$, highlighted in blue) directs attention towards learning the unknown parameters and their precision. Note that in Fig. 2.3a, the precision on the three unknown parameters start from a low prior precision of $P^\theta = 1$ and increase with each iteration, whereas the precision of known parameters (0 and 1) remains a constant (3.3×10^6). The noise precisions are learned simultaneously, which starts from a low prior precision of $P^{\lambda^w} = P^{\lambda^z} = 1$ and finally converges to the true noise precision (dotted black line). Both precisions are used to learn the three parameters of the system (Fig. 2.3b), which starts from randomly selected values within the range [-2,2] and finally converges to the true parameter values of the system ($\theta_3 = -\frac{k}{m} = -0.5714$, $\theta_4 = -\frac{b}{m} = -0.2857$ and $\theta_6 = \frac{1}{m} = 0.7143$), denoted by black dotted lines. From an attentional perspective, the lower plot in (Fig. 2.3a) is particularly significant here. This is because the robot discovers the data are more informative than initially assumed, thereby leading to an increase in its estimate of the precision of the data-generating process. This means that the robot is not only using the data to optimise its beliefs about states and parameters (system identification), it is also

using these data to optimise the way in which it assimilates these data.

In summary, precision-based attention, in the form of precision learning, helps the robot to accurately learn unknown parameters by fusing prior knowledge with new incoming data (sensory measurements), and attending to the least precise parameters.

2.5.2.4 Precision-modulated exploration and exploitation in system identification

Exploration and exploitation in the parameter space can be advantageous to robots during system identification. Precision-based attention—here the prior precision—allows a graceful balance between the two, mediated by the prior precision[6]. A very high prior precision encourages exploitation and biases the robot towards believing its priors, while a low prior precision encourages exploration and makes the robot sensitive to new information.
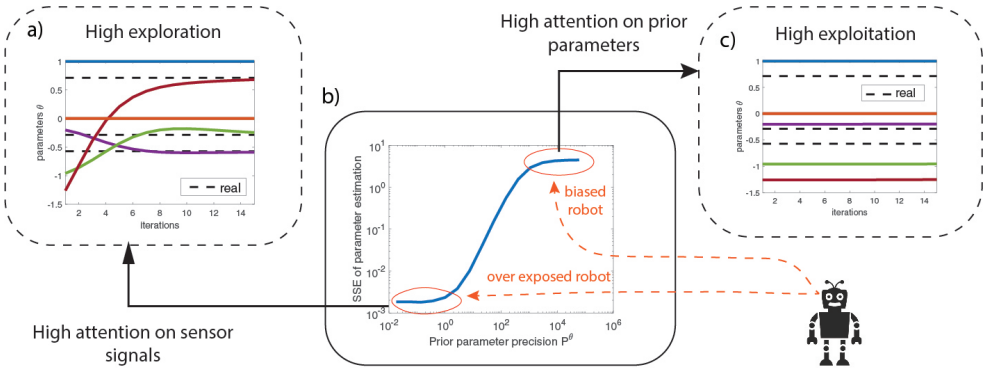
Fig 2.4: (A) Lower P^θ gives a high exploration strategy across the parameter space. (B) Precision-based attention allows exploration and exploitation balanced model learning mediated by the prior precisions on the parameters P^θ. (C) The higher the P^θ, the higher the attention on prior parameters η^θ and the lower the attention on the sensory signals while learning.

We use again the mass-spring-damper system example but with a different prior parameter precision P^θ. The prior parameters are initialised at random and learned using optimisation. Figure 2.4b shows the increase in parameter estimation error (SSE) as the prior parameter precision P^θ increases until it finally saturates. The bottom left region (circled in red) indicates the region where the prior precision is low, encouraging exploration with high attention on the sensory signals for learning the model. This region over-exposes the robot to its sensory signals by neglecting the prior parameters. The top right region (circled in red) indicates the biased robot where the prior precision is high, encouraging the robot to exploit its prior beliefs by retaining high attention on prior parameters. This regime biases the robot into being confident about its priors and disregarding new information from the sensory signals. Between those extreme regimes (blue curve) the prior precision balances the exploration-exploitation trade-off. Figure 2.4a describes how increased attention to

---

[6]Note that here we are using exploration and exploration not in terms of behaviour but for parameter learning. Exploration means adapting the parameter to a different (unexplored) value and exploitation means keeping that value

sensory signals helped the robot to recover from poor initial estimates of parameter values and converge towards the correct values (dotted black line). Conversely, in Fig. 2.4c, high attention on prior parameters did not help the robot to learn the correct parameter values.

These results establish that prior precision modelling allows balanced exploration and exploitation of parameter space during system identification. Although the results show that an over-exposed robot provides better parameter learning, we show – in the next section – that this is not always be the case.

2.5.2.5 Noise estimation
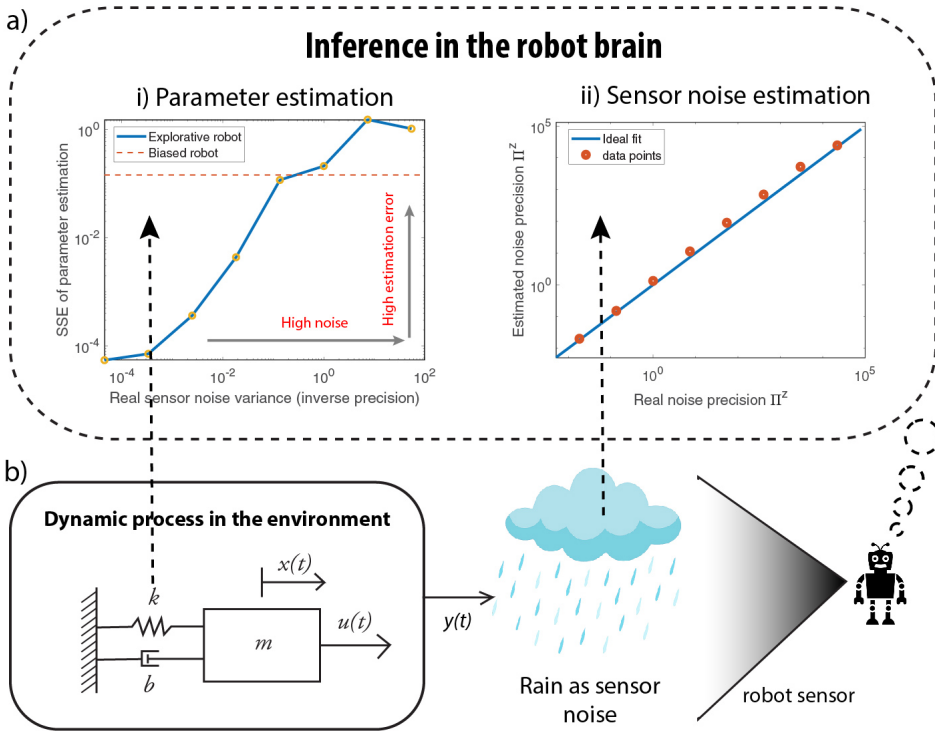
Fig 2.5: Simulations demonstrating how a biased robot could be advantageous, especially while learning in a highly noisy environment (shown in b). (a i) As the sensor noise increases, the quality of parameter estimation deteriorates to a point where an explorative robot generates higher parameter estimation errors than when compared to the biased robot that relies on its prior parameters. (a ii) However, the sensor noise estimation is accurate even for high noise environments, demonstrating the success of the attention mechanism using the noise precision.

In real-world applications, sensory measurements are often highly noisy and unpredictable. Furthermore, the robot does not have access to the noise levels. Thus, it needs to learn the noise precision (Π^z) for accurate estimation and robust control. Precision-based attention enables this learning. In what follows, we show how one can estimate Π^z using noise precision learning and that biasing the robot to prior beliefs can be advantageous in highly noisy environments.

Consider again the mass-spring-damper system in Figure 2.5b, where heavy rainfall/snow corrupts visual sensory signals. We evaluate the parameter estimation error under different noise conditions, using different levels of noise variances (inverse precision). For an over-exposed robot (only attending to sensory measurements), left plot of Fig. 2.5a, the estimation error increases as the noise strength increases, to a point where the error surpasses the error from a prior-biased robot. This shows that a robot, confident in its prior model, assigns low attention to sensory signals and outperforms an over-exposed robot that assigns high attention to sensory signals, in a highly noisy environment. The right plot of Fig. 2.5a shows the quality of noise precision learning for an over-exposed robot. It can be seen that all the data points in red lie close to the blue line, indicating that the estimated noise precision is close to the real noise precision. Therefore, the robot is capable of recovering the correct sensory noise levels even when the environment is extremely noisy, where accurate parameter estimation is difficult.

These numerical results show that attention mechanism — by means of noise precision learning — allows the estimation of the noise levels in the environment and thereby protects against over-fitting or overconfident parameter estimation.

Summary. We have shown how precision-based attention—through precision modelling and learning— yields to accurate robot state estimation, parameter identification and sensory noise estimation. In the next section, we discuss how action is generated in this framework.

2.5.3 Precision-modulated action

Selecting the optimal sequence of actions to fulfil a task is essential for robotics [173]. One of the most prominent challenges is to ensure robust behaviour given the uncertainty emerging from a highly complex and dynamic real world, where the robots have to operate on. A proper attention system should provide action plans that resolve uncertainty and maximise information gain. For instance, it may minimise the information entropy, thereby encouraging repeated sensory measurements (observations) on high uncertainty sensory information.

Salience, which in neuroscience is sometimes identified as Bayesian surprise (i.e., divergence between prior and posterior), describes which information is relevant to process. We go one step further by defining the saliency map as the epistemic value of a particular action [174]. Thus, the (expected) divergence now becomes the mutual information under a particular action or plan. This makes the saliency map more sophisticated because it is an explicit measure of the reduction in uncertainty or mutual information associated with a particular action (i.e., active sampling), and more pragmatic because it tells you where to sample data next, given current Bayesian beliefs.

We first describe a precision representation usually used in information gathering problems and then how to directly generate action plans through precision optimisation. Afterwards, we discuss the realisation of the full-fledged model presented in the neuroscience section for active perception. We use the informative path planning (IPP) problem, described in Fig. 2.6, as an illustrative example to drive intuitions.
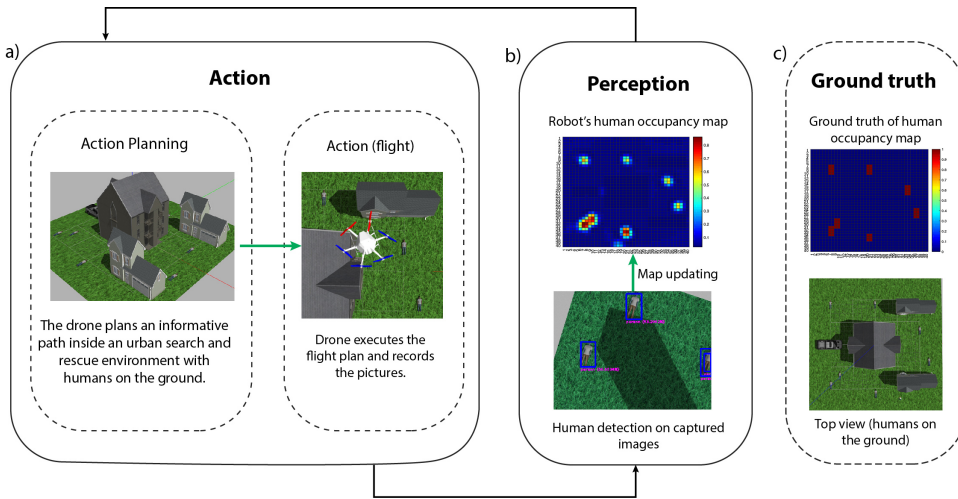
Fig 2.6: IPP problem for localizing human victims in an urban search and rescue scenario [53]. (a) Action: a UAV, in a realistic simulation environment, plans a finite look-ahead path to minimize the uncertainty of its human occupancy map (e.g., modeled as a Gaussian process) of the world. The planned path is then executed, during which the UAV flies and captures images at a constant measurement frequency. (b) Perception: after the data acquisition is complete, a human detection algorithm is executed to detect all the humans on the images. These detections are then fused into the UAV's human location map. The cycle is repeated until the uncertainty of the map is completely resolved (this usually implies enough area coverage and repeated measurements on uncertain locations). The ground truth of the human occupancy map and the UAV belief is shown in (b,c) respectively. The final map approaches the ground truth and all the seven humans on the ground are correctly detected.

2.5.3.1 PRECISION MAPS AS SALIENCY

One of the popular approaches in information gathering problems is to model the information map as a distribution (e.g., using Gaussian processes [175]). This is widely used in applications, such as a target search, coverage and navigation. The robot keeps track of an occupancy map and the associated uncertainty map (covariance matrix or inverse precision). While the occupancy map records the presence of the target on the map, the uncertainty map records the quality of those observations. The goal of the robot is to learn the distribution using some learning algorithm [176]. A popular strategy is to plan the robot path such that it minimises the uncertainty of the map in future [177]. In Sec. 2.5.3.2, we will show how we can use the map precision to perform active perception, i.e., optimise the robot path for maximal information gain. Optimising the map precision drives the robot towards an exploratory behaviour.

2.5.3.2 PRECISION OPTIMISATION FOR ACTION PLANNING

To introduce precision-based saliency we use an exemplary application of search and rescue. The goal is to find all humans using an unmanned air vehicle (UAV) [53, 156, 178, 179]. We use precision for two purposes: i) precision optimisation for action planning (plan flight path) and ii) precision learning for map refinement. In contrast to previous models of
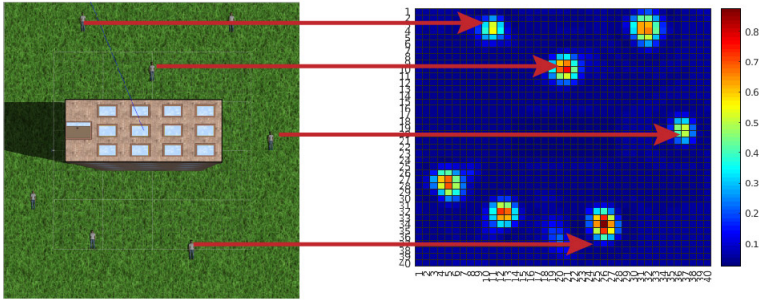
Fig 2.7: Finding humans with unmanned air vehicles (UAVs): an informative path planning (IPP) approach [54]. The simulation environment on the left consists of a tall building at the centre, surrounded by seven humans lying on the floor. The goal of the UAV is to compute the action sequence that allows maximum information gathering, i.e., the humans location uncertainty is minimised. On the right is the final occupancy map coloured with the probability of finding a human at that location. It can be observed that all humans on the simulation environment were correctly detected by the robot.
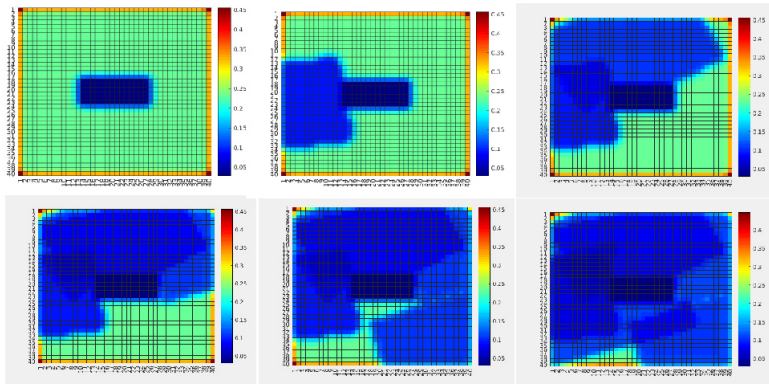
Fig 2.8: Variance map of the probability distribution of people location (Fig. 2.7) – inverse precision of human occupancy map. The plot sequence shows the reduction of map uncertainty (inverse precision) after measurements [54].

action selection within active inference in robotics [25, 32] here precision explicitly drives the agent behaviour. Figure 2.7 describes the scenario in simulation. The seven human targets on the ground are correctly identified by the UAV. We can formalise the solution as the UAV actions (next flight path) that minimise the future uncertainties of the human occupancy map. In our precision-based attention scheme, this objective is equivalent to maximising the posterior precision of the map. Figure 2.8 shows the reduction in map uncertainty after subsequent assimilation of the measurements (camera images from the UAV, processed by a human detector). The map (and precision) is learned using a recursive Kalman Filter by fusing the human detector outcome onto the map (and precision). The

algorithm drives the UAV towards the least explored regions in the environment, defined by the precision map.
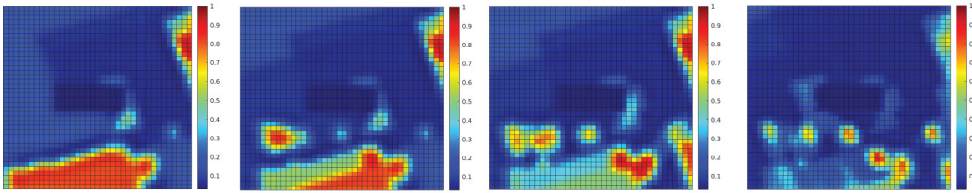
Fig 2.9: The human occupancy map (probability to find humans at every location of the environment) at four time instances during the UAV flight showing ambiguity resolution. The ambiguity arising from imprecise sensor measurements (false positive) is resolved through repeated measurements at the same location. The plot sequence shows how the assimilation of the measurements updates the probability of the people being in each location of the map [53].

Furthermore, Fig. 2.9 shows an example of uncertainty resolution under false positives. In this case, human targets are moved to the bottom half of the map. The first measurement provides a wrong human detection with high uncertainty. However, after repeated measurements at the same location in the map the algorithm was capable of resolving this ambiguity, to finally learn the correct ground truth map. Hence, the sought behaviour is to take actions that encourage repeated measurements at uncertain locations for reducing uncertainty.

Although the IPP example illustrates how to generate control actions through precision optimisation, the task, by construction, is constrained to explicitly reduce uncertainty. This is similar to the description of visual search described in [11], where the location was chosen maximise information gain. Information gain (i.e., the Bayesian surprise expected following an action) is a key part of the expected free energy functional that underwrite action selection in active inference. In brief, expected free energy can be decomposed into two parts the first corresponds to the information gain above (a.k.a., epistemic value or affordance). The second corresponds to the expected log evidence or marginal likelihood of sensory samples (a.k.a., pragmatic value). When this likelihood is read as a prior preference, it contextualises the imperative to reduce uncertainty by including a goal-directed, imperative. For example, in the search paradigm above, we could have formulated the problem in terms of reducing uncertainty about whether each location was occupied by a human or not. We could have then equipped the agent with prior preferences for observing humans.

In principle, this would have produced searching behaviour until uncertainty had been resolved about the scene; after which, the robot would seek out humans; simply because, these are its preferred outcomes. In thinking about how this kind of neuroscience inspired or biomimetic approach could be implemented in robotics, one has to consider carefully, the precision afforded sensory inputs (i.e., the likelihood of sensory data, given its latent causes) – and how this changes during robotic flight and periods of data gathering. This brings us back to the precision modulation and the temporal scheduling of searching and securing data. In the final section, we conclude with a brief discussion of how this might be implemented in future applications.

2.5.3.3 Precision-based active perception
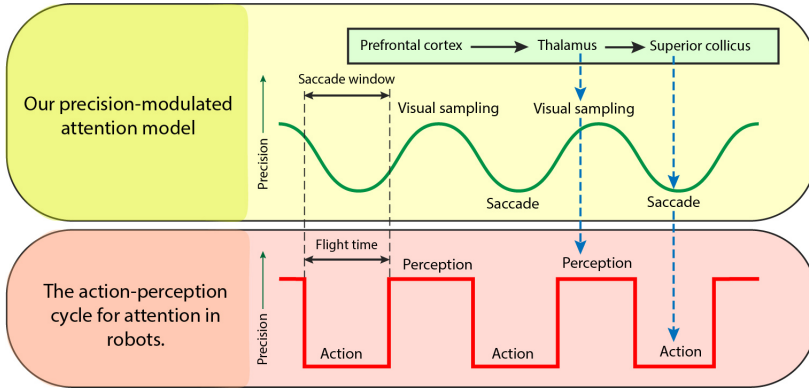
Fig 2.10: Precision-modulated attention model adapted to the action-perception loop in robotics. Each cycle consists of two steps: 1) action (planning and execution of a finite-time look ahead of the robot path for data collection) and 2) perception (learning using the collected data). This scheduling, using a finite time look-ahead plan, is quite common in real applications and of particular importance when processing is computationally expensive, e.g., slow rate of classification, non-scalable data fusion algorithms, Exponential planners, etc. However, the benefits of incorporating 'optimal' scheduled loop driven by precision should be further studied.

In this section, we discuss the realisation of a biomimetic brain-inspired model in relation to existing solutions in robotics in the context of path-planning. Figure 2.10 compares our proposed precision-modulated attention model—from Fig. 2.1—with the action-perception loop widely used in robotics. By analogy with eye saccades to the next visual sample, the UAV flies (action) over the environment to assimilate sensory data for an informed scene construction (perception). Once the flight time of the UAV is exhausted (similar to saccade window of the eye), the action is complete, after which the map is updated, and the next flight path is planned.

In standard applications of active inference, the information gain is supplemented with expected log preferences to provide a complete expected free energy functional [72]. This accommodates the two kinds of uncertainty that actions and choices typically reduce. The first kind of uncertainty is inherent in unknowns in the environment. This is the information gain we have focused on above. The second kind of uncertainty corresponds to expected surprise, where surprise rests upon a priori expected or preferred outcomes. As noted above, equipping robots with both epistemic and pragmatic aspects to their action selection or planning could produce realistic and useful behaviour that automatically resolves the exploration-exploitation dilemma. This follows because the expected free energy contains the optical mixture of epistemic (information-seeking) and pragmatic (i.e., preference seeking) components. Usually, after a period of exploration, the preference seeking components predominate because uncertainty has been resolved. Although expected free energy provides a fairly universal objective function for sentient behaviour, it does not specify how to deploy behaviour and sensory processing optimally. This brings us to the precision modulation model, inspired by neuroscientific considerations of attention

and salience.

Hence, there are key differences between biological and robotic implementations of the search behaviour. First, the use of oscillatory precision to modulate visual sampling and movement cycles, as opposed to arbitrary discrete action and perception steps currently used in robotics. Second, precision modulation influences both state estimation and action following the same uncertainty reduction principle. Importantly, our salience formulation speaks to selecting future data that reduces this uncertainty. For instance, we have shown—in the information gathering IPP example described in the previous subsection—that by optimising precision we also optimise behaviour.

We argue the potential need and the advantages of realising precision based temporal scheduling, as described the our brain-inspired model, for two practically relevant test cases: (*i*) learning dynamic models and (*ii*) information seeking applications.

In Section 2.5.2.4, we have shown how the exploration-exploitation trade-off can be mediated by the prior parameter precision during learning. However, the accuracy-precision curve (Fig. 2.4b) is often practically unavailable due to unknown true parameters values, challenging the modelling of prior precision. An alternative would be to use a precision based temporal scheduling mechanism to alternate between exploration and exploitation by means of a varying P^θ (similar to Fig. 2.10) during learning, such that system identification is neither biased nor over exposed to sensory measurements. In Fig. 2.5a, we showed how noise levels influence estimation accuracy, and how biasing the robot by modelling P^θ can be beneficial for highly noisy environments. A precision based temporal scheduling mechanism by means of a varying P^θ could provide a balanced solution between a biased robot (that exploits its model) and an exploratory one.

Furthermore, temporal scheduling, in the same way that eye saccades are generated, can be adapted for information gathering applications, such as target search, simultaneous localization and mapping, environment monitoring, etc. For instance, introducing precision-modulation scheduling for solving the IPP, and scheduling perception (map learning) and action (UAV flight). Precision modulation will switch between action and perception: when the precision is high, perception occurs (c.f., visual sampling), and when the precision is low, action occurs (c.f., eye movements). This switch, which is often implemented in the robotics literature using a budget for flight time, will be now dictated by precision dynamics.

In short, we have sketched the basis for a future realisation of precision-based active perception, where the robot computes the actions to minimise the expected uncertainty. While most attentional mechanisms in robotics are limited to providing a 'saliency' map highlighting the most relevant features, our attention mechanism proposes a general scheduling mechanism with action in the loop with perception, both driven by precision.

2.6 Concluding remarks

We have considered attention and salience as two distinct processes that rest upon oscillatory precision control processes. Accordingly, they require particular temporal considerations: attention to reliably estimate latent states from current sensory data and salience for uncertainty reduction regarding future data samples. This formulation addresses visual search from a first principles (Bayesian) account of how these mechanisms might manifest -– and the circular causality that undergirds them via a rhythmic theta-coupling. Crucially,

we have revisited the definition of salience from the visual neurosciences; where it is read as Bayesian surprise (i.e., the Kullback Leibler divergence between prior and posterior beliefs). We took this one step further and defined salience as the expected Bayesian surprise (i.e., epistemic value) of a particular action (e.g., sampling this set of data) [72, 180]. Formulating salience as the expected divergence renders it the mutual information under a particular action (or action trajectory) [71], and highlights its role in encoding working memory [21]. For brevity, our narrative was centred around visual attention and its realisation via eye movements. However, this model does not strictly need to be limited to visual information processing, because it addresses sensorimotor and auditory processing in general. This means it explains how action and perception can be coupled in other sensory modalities. For instance, [116] showed that visual information is coupled with finger movements at a theta rhythm.

The point of contact with the robotics use of salience emerges because the co-variation between a particular parameterisation and the inputs is a measure of the mutual information between the data and its estimated causes. In this sense, both definitions of salience reflect the mutual information – or information about a particular representation of a (latent) cause – afforded by an observation or consequence. However, our formulation is more sophisticated. Briefly, because it is an explicit measure of the reduction in uncertainty (i.e., mutual information) associated with a particular action (i.e., active sampling) and specifies where to sample data next, given current Bayesian beliefs. These processes (attention and salience) are a consequence of precision of beliefs over distinct model parameters. Explicitly, attention contends with precision over the causes of (current) outcomes and salience contends with beliefs about the data that has to be acquired and precision over beliefs about actions that dictate it. Since both processes can be linked via precision manipulation, the crucial thing is the precision that differentiates whether the agent acquires new information (under high precision) or resolves uncertainty by moving (low precision).

The focus of this work has been to illustrate the importance of optimising precision at various places in generative models used for data assimilation, system identification and active sensing. A key point – implicit in these demonstrations – rests upon the mean field approximation used in all applications. Crucially, this means that getting the precision right matters, because updating posterior estimates of states, parameters and precisions all depend upon each other. This may be particularly prescient for making the most sense of samples that maximises information gain. In other words, although attention and salience are separable optimisation processes, they depend upon each other during active sensing. This was the focus of our final numerical studies of action planning.

To face-validate our formulation, we evaluated precision-modulated attentional processes in the robotic domain. We presented numerical examples to show how precision manipulation underwrites accurate state and noise estimation (e.g., selecting relevant information), as well as allowing system identification (e.g., learning unknown parameters of the dynamics). We also showed how one can use precision-based optimisation to solve interesting problems; like the informative path planning in search and rescue scenarios. Thus, in contrast to previous uses of attention in robotics, we placed attention and saliency as integral processes for efficient gathering and processing of sensory information. Accordingly, 'attention' is not only about filtering the current flow of information from the

sensors but performing those actions that minimise expected uncertainty. Still, the full potential of our proposal has yet to be realised, as the precision-based attention should be able to account for prior preferences beyond the IPP problem (e.g., localising people using UAVs). Finally, we briefly considered the realisation of temporal scheduling for information gathering tasks, opening up interesting lines of research to provide robots with biologically plausible attention.

2

3

STATE AND INPUT OBSERVER DESIGN

*This chapter uses **precision modelling** to develop a state and input observer that can handle uncertainties entering an LTI system in the form of colored noise. This chapter develops the mathematical foundation of the DEM based observer design used in Section 2.5.2.2. The robot implementation of the observer is provided in Chapter 4. The observer is further extended to perform joint state and noise smoothness estimation in Chapter 5.*

3.1 Abstract

The free energy principle from neuroscience provides a biologically plausible solution to the brain's inference mechanism. This paper reformulates this theory to design a brain-inspired state and input estimator for a linear time-invariant state space system with colored noise. This reformulation for linear systems bridges the gap between the neuroscientific theory and control theory, therefore opening up the possibility of evaluating it under the hood of standard control approaches. Through rigorous simulations under colored noises, the observer is shown to outperform Kalman Filter and Unknown Input Observer with minimal error in state and input estimation. It is tested against a wide range of scenarios and the proof of concept is demonstrated by applying it on a real system.

3.2 Introduction

The design of state and input observer is vital towards the development of advanced industrial controllers, especially in the field of fault detection and correction [181]. Although a wide range of observer designs have been proposed for linear time invariant (LTI) systems, they assume the noises to be temporally uncorrelated (white) [182, 183]. This assumption is often violated in practice [184], resulting in a sub-optimal estimation [185]. An interesting approach called the Free Energy Principle (FEP) [2] is emerging from neuroscience, which leverages the noise smoothness in brain signals for perception. FEP demonstrates a potential to solve the problem of sub-optimal estimation. Therefore, it is imperative to reformulate it into an observer design and rigorously test it on real control system problems.

According to FEP proposed by Karl Friston, the brain's inference mechanism is a gradient descent over its free energy, where free energy is the information-theoretic measure that bounds the brain's sensory surprisal [43]. FEP emerges as a unified brain theory [186] by providing a mathematical description of brain functions [187], unifying action and perception [50], connecting physiological constructs like memory, attention, value, reinforcement and salience [187], and remaining consistent with Freudian ideas[188]. Similarities of FEP with reinforcement learning [42], neural networks [43, 44] and active learning [50] open up possibilities for biologically plausible control system techniques.

Dynamic Expectation Maximization (DEM) [170] is an interesting variant of FEP, that performs estimation on the brain's states, parameters and hyperparameters. Unlike point estimators like Kalman Filter (KF), DEM models the states in generalized coordinates and tracks the evolution of its trajectory. This renders DEM with a capability to gracefully handle coloured noise [170]. Considering that most noises in nature like wind are the product of dynamic systems - and hence coloured - it is imperative for filters to incorporate coloured noise. Moreover, the optimality of KF is affected as the noise turns coloured [189]. Many approaches have been proposed to modify KF to handle these colored noises: state augmentation and measurement differencing. State augmentation methods transform the system equations with colored noise into an equivalent higher order system with white noise before estimation [190]. However, the computations involved are sometimes ill-conditioned, and hence, other approaches like measurement differencing [191] were proposed. In this approach, the signal is whitened by subtracting the colored part from it. A detailed summary of these methods is provided in [192] and [193]. However, these methods model only causal noises and do not perform input estimation. Although few

methods have been proposed for simultaneous state and input estimation, they do not handle colored noise [183, 194]. This paper aims to bridge these research gaps.

Although, FEP based tools like active inference have found applications in robotics [27, 168], there is a gap in literature for the applications of DEM [170] in control theory, which can be attributed to the relatively formidable mathematics combined with a lack of rigorous testing of DEM for LTI systems. This paper aims to bridge this gap by simplifying DEM for an LTI system and introducing a state observer design. The core contributions of the paper are:

1. the formulation of a brain inspired linear observer design for LTI system, based on DEM.

2. the extensive evaluation of the observer in simulation and its validation using a real LTI system.

3.3 Problem Statement

Consider the linear plant dynamics given in Equation 3.1 where A, B and C are constant system matrices, $\mathbf{x} \in \mathbb{R}^n$ is the hidden state, $\mathbf{v} \in \mathbb{R}^r$ is the input and $\mathbf{y} \in \mathbb{R}^m$ is the output.

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{v} + \mathbf{w}, \qquad \mathbf{y} = C\mathbf{x} + \mathbf{z}. \tag{3.1}$$

Here $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$ represent the process and measurement noise respectively. Variables of the plant are denoted in boldface, while its estimates are denoted in non-boldface. The noises in this paper are generated through the convolution of white noise with a Gaussian filter.

Two problems are considered in this paper. The first one is an observer design with known inputs $\mathbf{v}$, and the second one is with unknown inputs for which both x and v are to be estimated. We show that DEM outperforms KF and Unknown Input Observer (UIO) [183] for these problems respectively.

3.4 Preliminaries

To lay the foundation of our observer design, this section introduces the key concepts and terminologies behind DEM.

3.4.1 Generative model

The key idea behind DEM to deal with coloured noise is to model the trajectory of the states using generalized coordinates. The states are expressed in generalized coordinates using its higher order derivatives as $\tilde{x} = [x \; x' \; x'' \; ...]^T$. The variables in generalized coordinates are denoted by a tilde, and its components (higher derivatives) are denoted by primes. The evolution of states is written as:

$$\begin{aligned} x' &= Ax + Bv + w & y &= Cx + z \\ x'' &= Ax' + Bv' + w' & \dot{y} &= Cx' + z' \\ &... & &... \end{aligned} \tag{3.2}$$

The coloured noises are modelled to be analytic such that the covariance of noise deriva-tives $\tilde{z} = [z, z', z'', ...]^T$ and $\tilde{w} = [w, w', w'', ...]^T$ are well defined. The generative model representing the system is compactly written as:

$$\dot{\tilde{x}} = D_x \tilde{x} = \tilde{A}\tilde{x} + \tilde{B}\tilde{v} + \tilde{w} \qquad \tilde{y} = \tilde{C}\tilde{x} + \tilde{z} \tag{3.3}$$

3

where $D_x = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & . & . & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}_{(p+1)\times(p+1)} \otimes I_{n\times n}$

performs derivative operation, equivalent to shifting up all components in generalized coordinates by one block. Similar definition holds for D_v (appears later) with size $r(d+1) \times r(d+1)$, where p and d are the order of generalized motion of states and inputs respectively. Here, $\tilde{A} = I_{p+1} \otimes A$, $\tilde{B} = I_{p+1} \otimes B$ and $\tilde{C} = I_{p+1} \otimes C$, where $\otimes$ is the Kronecker tensor product.

3.4.2 Colored noise

DEM uses generalized coordinates, which models a correlation between noise deriva-tives through the temporal precision matrix S (inverse of covariance matrix) [170]. The correlation is assumed to be due to a Gaussian filter with S given by:

$$S(\sigma^2) = \begin{bmatrix} 1 & 0 & -\frac{1}{2\sigma^2} & .. \\ 0 & \frac{1}{2\sigma^2} & 0 & .. \\ -\frac{1}{2\sigma^2} & 0 & \frac{3}{4\sigma^4} & .. \\ .. & .. & .. & .. \end{bmatrix}^{-1}_{(p+1)\times(p+1)} \tag{3.4}$$

where σ^2 is the variance of Gaussian filter, with σ denoting the noise smoothness. While $\sigma^2 = 0$ denotes white noise, non-zero σ^2 denotes coloured noise. The covariance between noise derivatives increases exponentially with the order of noise derivatives. Simulations show that derivatives above 6 can be neglected [170]. The generalized noise precision matrices are given by $\tilde{\Pi}_w = S(\sigma^2) \otimes \Pi_w$, $\tilde{\Pi}_z = S(\sigma^2) \otimes \Pi_z$ and $\tilde{\Pi}_v = S(\sigma^2) \otimes \Pi_v$, where Π_w, Π_z and Π_v are the inverse noise covariances.

3.4.3 Generalized motion of output and noises

The generalized motion of output $\tilde{y}$ is practically not accessible for sensors. Moreover, most of the sensors like encoders operate with discrete measurements, unlike biophysical systems like the brain. Therefore, as a pre-processing step for estimation, $\tilde{y}$ should be computed in discrete domain. Given the p temporal derivatives $\tilde{y}$ at time t, the p output sequence surrounding $\mathbf{y}$ can be approximated using Taylor series as [170]:

$$\hat{\mathbf{y}} = \begin{bmatrix} .. \\ \mathbf{y}(t-dt) \\ \mathbf{y}(t) \\ \mathbf{y}(t+dt) \\ .. \end{bmatrix} = (E \otimes I_m)\tilde{\mathbf{y}}, \ E_{ij} = \frac{[(i-\frac{p+1}{2})dt]^{j-1}}{(j-1)!}, \tag{3.5}$$

where $i, j = 1, 2, ... p + 1$ and $\hat{y}$ has the size $m(p + 1) \times 1$. Therefore, generalized motion of output $\tilde{\mathbf{y}}$ at time t is:

$$\tilde{\mathbf{y}} = (E^{-1} \otimes I_m)\hat{\mathbf{y}}. \tag{3.6}$$

Using $\tilde{\mathbf{y}}$ embeds more temporal information about the plant into the data in the form of conditional trajectories, with the disadvantage of a time latency of $\frac{p}{2} dt$ in estimation. The next section employs this generalized output along with the generative model for observer designs.

3.5 STATE AND INPUT OBSERVER DESIGN

This section delineates the main contribution of the paper through two observer designs: with (i) known inputs and (ii) unknown inputs, for a general LTI system, and formulates the stability condition.

3.5.1 OBSERVER DESIGN WITH UNKNOWN INPUTS

According to DEM, estimation is performed through a gradient ascent over the variational energy [1] (variational component of free energy) with respect to hidden states and inputs [170]. With the mean-field approximation and Laplace approximation in place [43], the variational energy at time t is given by:

$$V(t) = -\frac{1}{2}\tilde{\epsilon}^T \tilde{\Pi} \tilde{\epsilon}, \tag{3.7}$$

where $\tilde{\Pi} = diag(\tilde{\Pi}_z, \tilde{\Pi}_v, \tilde{\Pi}_w)$ is the generalized noise precision matrix, $\tilde{\epsilon}$ is the prediction error. The prediction error is formulated as:

$$\tilde{\epsilon}(t) = \begin{bmatrix} \tilde{\epsilon}^y \\ \tilde{\epsilon}^v \\ \tilde{\epsilon}^x \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{y}} - \tilde{C}\tilde{x} \\ \tilde{v} - \tilde{\eta} \\ D_x \tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix}, \tag{3.8}$$

where $\tilde{\eta}$ is the prior belief of the inputs, which is crucial for observer design with unknown input. During fault detection in industrial automation processes, the expected inputs under normal conditions are known. This prior knowledge can be incorporated into the observer using $\tilde{\eta}$ with higher precision $\tilde{\Pi}_v$. For a completely unknown $\tilde{\eta}$, precision $\tilde{\Pi}_v$ should be set very low. Taking derivative of Equation 3.7 with respect to $X = [\tilde{x}, \tilde{v}]^T$ yields:

$$V(t)_X = -\tilde{\epsilon}_X^T \tilde{\Pi} \tilde{\epsilon}, \tag{3.9}$$

where

$$\tilde{\epsilon}_X(t) = \begin{bmatrix} \tilde{\epsilon}_{\tilde{x}}^y & \tilde{\epsilon}_{\tilde{v}}^y \\ \tilde{\epsilon}_{\tilde{x}}^v & \tilde{\epsilon}_{\tilde{v}}^v \\ \tilde{\epsilon}_{\tilde{x}}^x & \tilde{\epsilon}_{\tilde{v}}^x \end{bmatrix} = \begin{bmatrix} -\tilde{C} & O \\ O & I \\ D_x - \tilde{A} & -\tilde{B} \end{bmatrix}. \tag{3.10}$$

The update equation after *free-form* approximation [170] is expressed as a gradient ascent over the variational energy:

$$\dot{X} = \begin{bmatrix} k_x I_{n(p+1)} & O \\ O & k_v I_{r(d+1)} \end{bmatrix} V(t)_X + \begin{bmatrix} D_x & O \\ O & D_v \end{bmatrix} X, \tag{3.11}$$

---

[1] maximization of the ELBO term (Section 1.4)

where k_x and k_v are the learning rates for state and input update respectively. Substituting Equations 3.8, 3.9 and 3.10 in 3.11 upon simplification yields a linear observer design formulated as:

$$\dot{X} = \begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{v}} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{v} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{y}} \\ -\tilde{\eta} \end{bmatrix} \; and \; Y = X, \tag{3.12}$$

where Y is the output of the observer and,

$$A_1 = \begin{bmatrix} D_x & O \end{bmatrix} + k_x \begin{bmatrix} \tilde{C}^T \\ O \end{bmatrix} \begin{bmatrix} \tilde{\Pi}_z & O \\ O & \tilde{\Pi}_v \end{bmatrix} \begin{bmatrix} -\tilde{C} & O \\ O & I \end{bmatrix}$$
$$-k_x(D_x - \tilde{A})^T \tilde{\Pi}_w \begin{bmatrix} D_x - \tilde{A} & -\tilde{B} \end{bmatrix},$$

$$A_2 = \begin{bmatrix} O & D_v \end{bmatrix} + k_v \begin{bmatrix} O & -I \end{bmatrix} \begin{bmatrix} \tilde{\Pi}_z & O \\ O & \tilde{\Pi}_v \end{bmatrix} \begin{bmatrix} -\tilde{C} & O \\ O & I \end{bmatrix}$$
$$+k_v \tilde{B}^T \tilde{\Pi}_w \begin{bmatrix} D_x - \tilde{A} & -\tilde{B} \end{bmatrix},$$

$$B_1 = k_x \begin{bmatrix} \tilde{C}^T & O \end{bmatrix} \begin{bmatrix} \tilde{\Pi}_z & O \\ O & \tilde{\Pi}_v \end{bmatrix}, \; B_2 = k_v \begin{bmatrix} O & -I \end{bmatrix} \begin{bmatrix} \tilde{\Pi}_z & O \\ O & \tilde{\Pi}_v \end{bmatrix}$$

It can be observed from Equation 3.12 that the inputs and states can be recovered by using the output of plant $\tilde{\mathbf{y}}$ and prior for inputs $\tilde{\eta}$ as inputs to the observer. An exact algebraic discretization is used to numerically integrate the observer for estimation with time-sampled output $\mathbf{y}$.

3.5.2 Observer design with known inputs
DEM can be reformulated for state estimation with known inputs. When the inputs are known, the update rule for $\tilde{x}$ in Equation 3.12 can be rearranged as:

$$\dot{\tilde{x}} = \begin{matrix} \left[D_x - k_x \tilde{C}^T \tilde{\Pi}_z \tilde{C} - k_x(D_x - \tilde{A})^T \tilde{\Pi}_w(D_x - \tilde{A}) \right] \tilde{x} + \\ k_x \begin{bmatrix} \tilde{C}^T \tilde{\Pi}_z & (D_x - \tilde{A})^T \tilde{\Pi}_w \tilde{B} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{y}} \\ \tilde{v} \end{bmatrix} \end{matrix} \tag{3.14}$$

The input to the linear state observer with known system inputs is the combined vector of generalized output and input of the system $[\tilde{\mathbf{y}}, \tilde{v}]^T$.

3.5.3 Stability condition for observer with unknown inputs
This section explicates the stability conditions for the observer design presented in Section 3.5.1. The error between observer's estimates and the ideal values from the plant is given

$$A_3 = \begin{bmatrix} D_x - k_x(D_x - \tilde{A})^T \tilde{\Pi}_w(D_x - \tilde{A}) - k_x \tilde{C}^T \tilde{\Pi}_z \tilde{C} & k_x(D_x - \tilde{A})^T \tilde{\Pi}_w \tilde{B} \\ k_v \tilde{B}^T \tilde{\Pi}_w(D_x - \tilde{A}) & D_v - k_v \tilde{B}^T \tilde{\Pi}_w \tilde{B} - k_v \tilde{\Pi}_v \end{bmatrix}$$

$$A_4 = \begin{bmatrix} (-I + k_x(D_x - \tilde{A})^T \tilde{\Pi}_w)(D_x - \tilde{A}) & \tilde{B} - k_x(D_x - \tilde{A})^T \tilde{\Pi}_w \tilde{B} \\ -k_v \tilde{B}^T \tilde{\Pi}_w(D_x - \tilde{A}) & k_v \tilde{B}^T \tilde{\Pi}_w \tilde{B} + k_v \tilde{\Pi}_v \end{bmatrix}, \; A_5 = \begin{bmatrix} I & -k_x \tilde{C}^T \tilde{\Pi}_z & O \\ O & O & -k_v \tilde{\Pi}_v \end{bmatrix}$$

$$\tag{3.13}$$

by $\tilde{e} = \begin{bmatrix} \tilde{e}^x \\ \tilde{e}^v \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{x}} - \tilde{x} \\ \tilde{\mathbf{v}} - \tilde{v} \end{bmatrix}$. Taking the derivative of $\tilde{e}$ and substituting Equations 3.3 and 3.12 into it after simplification yields the error dynamics of the observer as:

$$\dot{\tilde{e}} = \begin{bmatrix} \dot{\tilde{e}}^x \\ \dot{\tilde{e}}^v \end{bmatrix} = A_3 \begin{bmatrix} \tilde{e}^x \\ \tilde{e}^v \end{bmatrix} \tilde{e} + A_4 \begin{bmatrix} \tilde{x} \\ \tilde{v} \end{bmatrix} + A_5 \begin{bmatrix} \tilde{w} \\ \tilde{z} \\ \tilde{\eta} \end{bmatrix}, \tag{3.15}$$

where A_3, A_4 and A_5 are given by Equation 3.13. Assuming that the states, inputs and noises for the system are bounded for a stable system under consideration, the error dynamics stabilizes if all the eigen values of A_3 have negative real part. Therefore, selecting the learning rates k_x and k_v such that A_3 represents a stable system matrix acts as the design criteria for a stable observer. Section 3.7.3 deals with simulating the observer for a wide range of stable systems. The next section delineates the components of the observer design that contribute to its functioning.
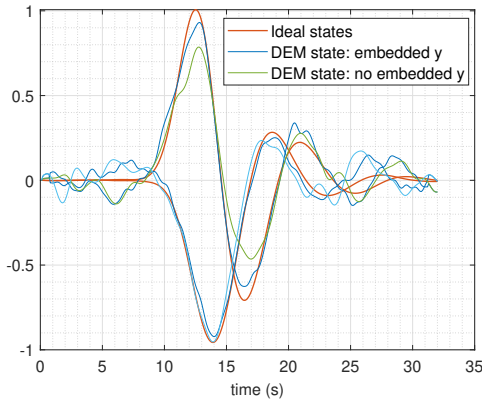
Fig 3.1: State estimation with and without embedding generalized motion to the output. The blue curve uses higher derivatives of y for state estimation, while the green curve doesn't. The blue is a better approximation of red than the green.

3.6 Components of observer design

This section elucidates the relevance of main components of the observer by testing it on an LTI system given by:

$$A = \begin{bmatrix} -0.25 & 1.00 \\ -0.50 & -0.25 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C = \begin{bmatrix} 0.125 & 0.1633 \\ 0.125 & 0.0676 \\ 0.125 & -0.0676 \\ 0.125 & -0.1633 \end{bmatrix} \tag{3.16}$$

and quantifying the performance using sum of squared error (SSE) in state and input estimation. Parameters from [170] are adopted for a convenient result comparison. Unless

mentioned otherwise, the same system will be used throughout the paper with an input of $v = e^{-0.25(t-12)^2}$, noise precisions $\Pi_w = e^8 I_{n \times n}$ and $\Pi_z = e^8 I_{m \times m}$, noise smoothness parameters (Equation 3.4) $\sigma = 0.5$, order of generalized motion $p = 6$, $d = 2$ and sampling time $dt = 0.1s$.

3.6.1 Generalized motion of output

The advantage of using generalized motion of the output **y** as discussed in Section 3.4.3 is illustrated in Figure 3.1 (for the system given above), which shows a better state estimation when generalized motion for **y** is used. Quantitatively, it corresponds to a drop in SSE for state estimation from *5.11* to *2.2* when generalized motion of output is used, thereby demonstrating its relevance.
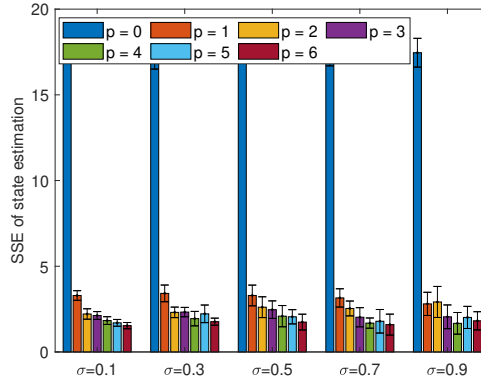
Fig 3.2: Embedding higher levels for the generalized motion of states decreases the state estimation error. The behaviour is robust against noise smoothness σ.

3.6.2 Embedding order of states and input

The embedding order represents the number of derivatives used in the generalized coordinates of the states and inputs. This section demonstrates the advantage of using generalized motion of states. Figure 3.2 and 3.3 shows the increasing accuracy of state and input estimation with increasing embedding order p for the states for the system given in Section 3.6. The experiments were repeated for different noise smoothness σ to demonstrate the robustness of the method. In summary, a higher order p ensures better accuracy in state estimation. Similar results can be achieved for higher order embedding d of input on state and input estimation.

3.6.3 Accuracy V/S Complexity

Free energy can be defined as complexity minus accuracy [50], where complexity is the measure of closeness of the estimates to the prior information, while accuracy is the measure of sensory surprisal or the sensory prediction error. A higher Π_v favours a higher weightage for priors during inference, while a lower Π_v favours a higher weightage for incoming data. Figure 3.4 demonstrates DEM's input estimation for a constant inaccurate
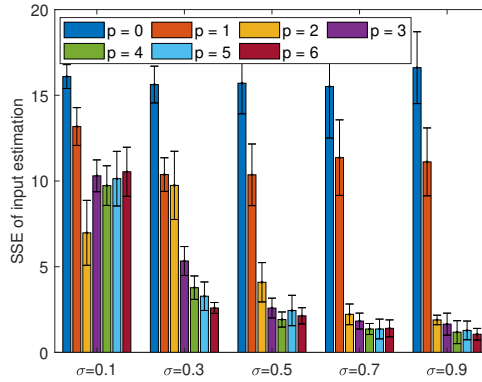
Fig 3.3: Higher order of generalized states increases the accuracy of input estimation for different noise smoothness σ. Input estimation is found to be less accurate for lower σ.

prior of 0.5 and varying Π_v. It shows that setting a low value for Π_v results in accurate input estimation, even with inaccurate prior η. Since accuracy is the focus of the paper, low value of Π_v will be used throughout.
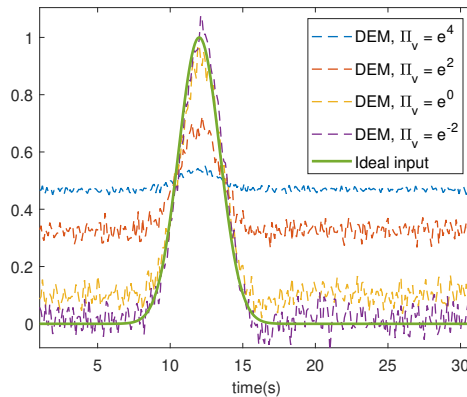
Fig 3.4: Setting low Π_v results in accurate input estimation even with wrong priors. An inaccurate constant prior of 0.5 was used with $\sigma = 0.01, \Pi_w = \Pi_z = e^{12}$.

3.6.4 LEARNING RATES

The inference process of brain is modelled as a gradient ascent over the variational energy. Since DEM is a fixed-form approximation to the ensemble density that tracks the trajectory of conditional mode, an increase in learning rate results in an exponentially faster convergence of this trajectory to the conditional mode, thereby maximizing variational energy [2] [170]. This results in a more accurate state and input estimation with increasing

---

[2]maximization of the ELBO term (Section 1.4)

learning rate, which is demonstrated in Figure 3.5. It can be observed that a higher learning rate results in lower error in state estimation. Similar conclusion can be drawn for input estimation. Since discretization of the observer is exact, higher learning rates does not entail numerical instability as long as the selection criteria given in Section 3.5.3 is met.
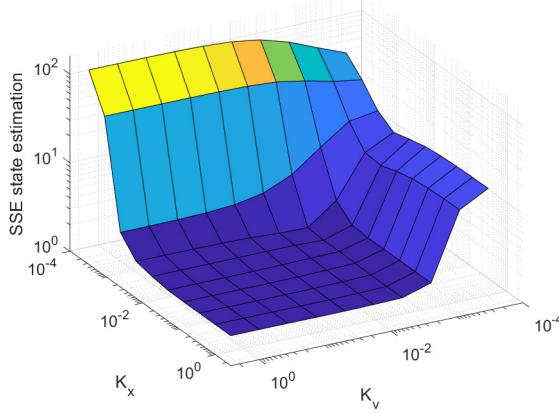
Fig 3.5: Higher learning rates result in better state estimation.

3.6.5 Equivalence of DEM with Kalman Filter

In case of white noise, the state estimation of DEM given in Equation 3.14 (without generalized motion) reduces to:

$$\dot{x} = k_x(-A^T\Pi_x A - C^T\Pi_y C)x - k_x(A^T\Pi_x)Bv + k_x C^T\Pi_y y. \tag{3.17}$$

This is very similar in form with the KF state update equation (with respect to variables x, v and y):

$$\dot{x} = (A - PC^T\Pi_y C)x + Bv + P(C^T\Pi_y)y, \tag{3.18}$$

where P is the error covariance that solves the Riccati equation $\dot{P} = AP + PA^T + \Pi_w^{-1} - PC^T\Pi_z^T CP^T$. If inputs are not modelled ($v = 0$ and $\Pi_x << \Pi_y$) and the learning rate k_x is tuned close to P ($k_x = P$), both the update equations 3.17 and 3.18 simplifies such that they behave exactly the same. Figure 3.6 shows the coinciding state estimation for DEM and KF for the same system given in Section 3.6 with $\sigma = 10^{-9}$, $dt = 0.5s$ and $\Pi_x = e^8 I_{2\times2}$ (during data generation). A small precision of $\Pi_x = e^0 I_{2\times2}$ is used while estimating with unmodelled v. Therefore, it can be concluded that DEM behaves like KF when noises are white and inputs are unmodelled. The next section deals with proving rigorously through simulations that when the noises are colored, DEM moves away from KF and outperforms KF and UIO.

Fig 3.6: State estimation of KF and DEM coincides when noises are white and inputs are not modelled. Here Π_x is assumed to be $e^0 I_{2\times 2}$ during estimation. $\sigma = 10^{-9}$ and $dt = 0.5s$.

3.7 BENCHMARK THE OBSERVER

This section aims to establish the superior performance of DEM against other benchmarks through a series of rigorous simulations under varied conditions.

3.7.1 BENCHMARK STATE OBSERVER

This section benchmarks the observer against KF and Unknown Input Observer (UIO) [183] for the system given in Equation 3.16. This version of UIO was selected owing to its capability to simultaneously estimate state and input without delays for a discrete LTI system. Correct noise smoothness value is key towards best inference. Figures 3.7a and 3.7b demonstrate the performance of DEM with unknown inputs for different levels of noise smoothness (with $\Pi_w = \Pi_z = e^{10}$) and noise variances (with $\sigma = 0.5$) respectively, where $DEM(\sigma = 0)$ assumes an unknown infinitely rough noise during estimation while $DEM(\sigma = \sigma)$ uses the known σ. DEM with known and unknown noise smoothness consistently outperforms KF for state estimation for varied noise smoothness and variance.

Given the correct noise smoothness and the system in Equation 3.16, 10 simulations each were used to generate Figure 3.7c with parameters from Section 3.6 and $dt = 0.02s$. It shows that DEM outperforms KF with lower SSE in state estimation, both for known and unknown input. Moreover, the performance of KF worsens as the noise smoothness increases. This can be attributed to the failed assumption of uncorrelated nature of the noise that KF is based upon. Modelling the states and inputs in the generalized coordinates embeds higher order motion during inference, motivating the use of DEM for state estimation while dealing with coloured noises.

3.7.2 BENCHMARK INPUT OBSERVER

This section benchmarks unknown input estimation given by Equation 3.12 against UIO. Figure 3.8 shows the successful input estimation by DEM and UIO for the system given in Equation 3.16 with $dt = 0.02s$. Input estimate from DEM is more similar to the ideal input when compared to the estimation from UIO. The mean SSE for 10 such simulations each for
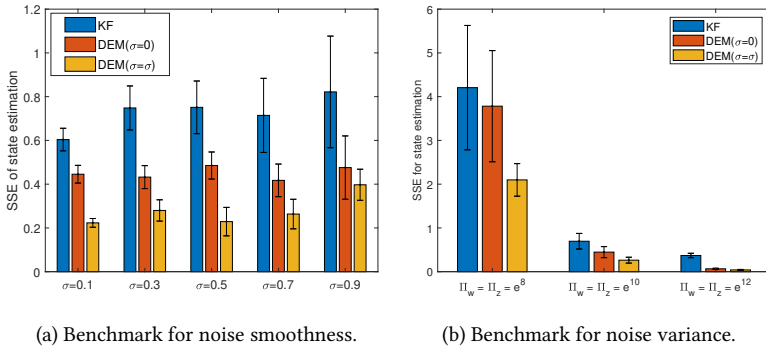
(a) Benchmark for noise smoothness.

(b) Benchmark for noise variance.
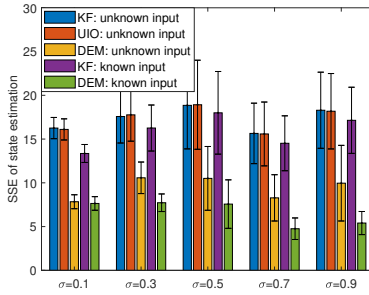
(c) Benchmark against other observers.

Fig 3.7: Benchmarking observer for different noise smoothness and noise variance against KF and UIO.
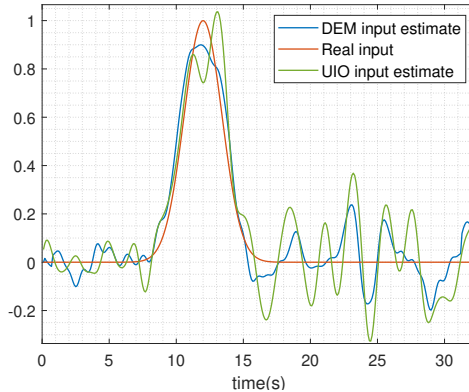
Fig 3.8: Gaussian bump input estimated by DEM and UIO. DEM provides a better input estimate than UIO.

different noise smoothness is plotted with one standard deviation as shown in Figure 3.9.

It can be observed that DEM outperforms UIO for input estimation under highly coloured noises. Better performance of DEM can be attributed to the tracking of trajectories of inputs through generalized coordinates, unlike just the point estimates by UIO.
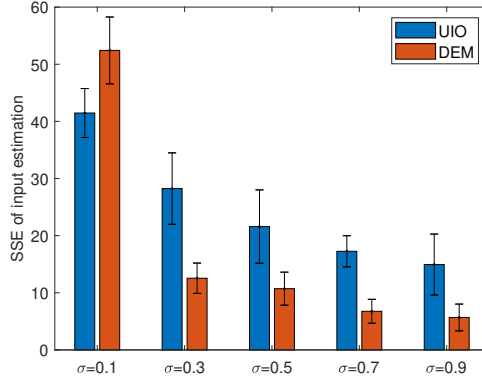
Fig 3.9: DEM outperforming UIO for input estimation under high noise smoothness.

3.7.3 Fully randomized benchmarking

This section establishes the superior performance of DEM by generalizing it over a wide range of systems with different input signals. System matrices A, B and C were randomized for values in the range [-1,1] with 2 state variables and randomly selected number of outputs ranging between 1 and 6. Noise variances $\Pi_w = e^6 I_{n \times n}$ and $\Pi_z = e^6 I_{m \times m}$ were used with a noise smoothness of 0.5. To preclude high SSE in state estimation, only stable A matrices were chosen . 100 simulations each for three different input signals - Gaussian $(e^{-0.25(t-12)^2})$, sinusoidal $(\sin 0.25t)$ and ramp $(t/32)$ - were used for a performance comparison. Figure 3.10 shows all the instances with known inputs. Most of the samples occupy the lower diagonal half of Figure 3.10, indicating that DEM outperforms KF in most cases with coloured noise by reducing the sum of squared error (SSE) in state estimation for a wide range of randomly generated LTI systems for different known inputs. In summary, DEM outperforms other benchmarks in state and input estimation for randomised noise smoothness, noise variance, systems and input signals. The next section demonstrates the applicability of the observer on real systems.
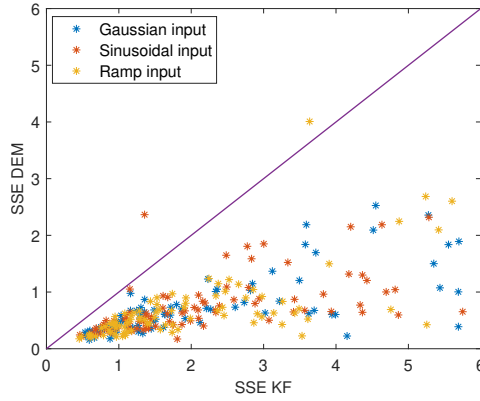
Fig 3.10: DEM outperforms KF for state estimation with known inputs. Coloured noises and three inputs were applied on 300 randomly generated systems to plot the estimation error.

3.8 PROOF OF CONCEPT ON REAL SYSTEM

This section aims to provide a proof of concept for the observer design with the help of data acquired from a maxon DC motor. The linear state space equation of a DC motor is:

$$\dot{x} = \begin{bmatrix} \dot{\theta} \\ \ddot{\theta} \\ \dot{i} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{b}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \\ i \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix} V$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \\ i \end{bmatrix}, \tag{3.19}$$

where θ is the angular position of the rotating motor knob, i the current, $b = 3.74 \times 10^{-5} Nms$ the damping ratio, $J = 2.69 \times 10^{-4} kgm^2$ the moment of inertia of the rotor, $K = 0.0472 Vrad^{-1}s^{-1}$ the electromotive force constant, $L = 8.4 \times 10^{-4} H$ the inductance, and $R = 9.24\Omega$ the resistance. The input to the system is voltage, while the output is the angular position of the knob.

The output data acquired for varying input was used for inference by the proposed observer. The results of the state and input estimation is shown in Figure 3.11 and 3.12 respectively. It can be observed that DEM estimates for current and voltage closely resemble the measurements respectively. The SSE for state estimation is then computed. DEM, has a lower SSE for state estimation of 8809, when compared to 14800 and 2.6×10^6 as that of UIO and KF respectively (overlapping plots were omitted for visibility). Therefore, DEM can effectively be used for state and input estimation on real systems with noises. MATLAB codes are available at $www.github.com/ajitham123/DEM\_observer$.
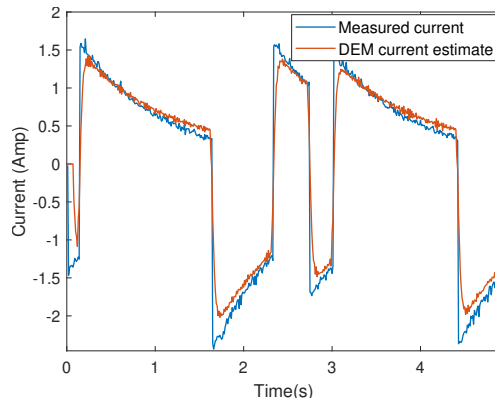
Fig 3.11: Current estimation using DEM in red is similar to the measured motor current in blue. Slight deviation is due to the imperfect motor parameters.
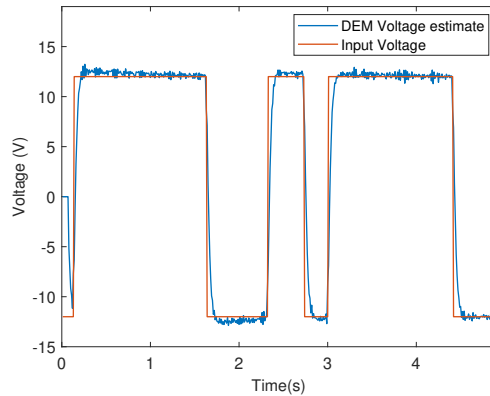
Fig 3.12: The input voltage estimation of DEM in blue is similar to the measured voltage in red, demonstrating that the observer works on real systems.

3.9 Conclusions and future work

The paper introduced a brain inspired state and input observer based on DEM that can be applied to general LTI systems. The observer leverages the information contained within coloured noises through generalized coordinates. Extensive simulations with coloured noises show that the observer outperform state-of-the-art observers through minimal state and input estimation errors. The observer was generalized for a wide range of system parameters, noise smoothness, noise variance and input signals. Finally, the proof of concept was provided through state and input estimation on a real system, demonstrating its real life performance. The main drawback of the observer is the poor input estimation for low noise smoothness. Future work will concentrate on extending the observer for complex hierarchical systems.

ACKNOWLEDGEMENT

3

4

STATE AND INPUT ESTIMATION OF A QUADROTOR IN WIND

*This chapter uses **precision modelling** for the state and input estimation of a quadrotor hovering in wind. This chapter provides the experimental confirmation on a quadrotor for the state and input observer developed in Chapter 3, and used in 2.5.2.2. The observer is further extended in Chapter 5 for the joint estimation of states and noise smoothness.*

4.1 Abstract

The free energy principle from neuroscience provides a brain-inspired perception scheme through a data-driven model learning algorithm called Dynamic Expectation Maximization (DEM). This paper aims at introducing an experimental design to provide the first experimental confirmation of the usefulness of DEM as a state and input estimator for real robots. Through a series of quadcopter flight experiments under unmodelled wind dynamics, we prove that DEM can leverage the information from colored noise for accurate state and input estimation through the use of generalized coordinates. We demonstrate the superior performance of DEM for state estimation under colored noise with respect to other benchmarks like State Augmentation, SMIKF and Kalman Filtering through its minimal estimation error. We demonstrate the similarities in the performance of DEM and Unknown Input Observer (UIO) for input estimation. The paper concludes by showing the influence of prior beliefs in shaping the accuracy-complexity trade-off during DEM's estimation.

4.2 Introduction

The widespread use of unmanned aerial vehicles (UAV) as delivery drones has increased the need for robust state and input estimators, mainly owing to its safety during uncertain events such as strong wind. We take a step in this direction by evaluating the usefulness of an approach from neuroscience to handle the wind during estimation.

In literature, a wide range of approaches have been used for the state estimation of linear time invariant (LTI) systems. However, most of them assume the noise to be white [195], which is often a wrong assumption in practice [196]. For example, Kalman filter (KF) [197] ensures optimality when the noises are white [198], but it is suboptimal when the noises are colored. An interesting approach from neuroscience called the Free Energy Principle (FEP) uses the concept called generalized coordinates that can leverage the noise derivative information in the brain signals for perception. The FEP based perception scheme called Dynamic Expectation Maximization (DEM) [170] was recently reformulated into a state and input observer for LTI systems with colored noise, and was shown to outperform the KF in simulation [46]. In this paper, we aim to provide the experimental validation of the DEM observer for a quadrotor under wind conditions using the setup given in Figure 4.1. The main contributions of the paper are:

1. Introduce an experimental design with real robots to provide the proof of concept for DEM as a state and input observer.

2. Provide the first experimental confirmation for the advantage of generalized coordinates in handling colored noise during state and input estimation on robots.

3. Demonstrate the influence of prior beliefs in shaping the accuracy-complexity trade-off during estimation.
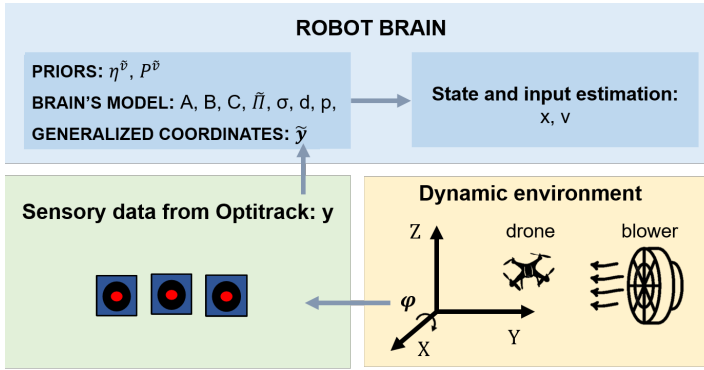
Fig 4.1: The schematic representation of our experimental setup for the DEM's state and input estimation using a quadrotor.

4.3 Related work

This section introduces the interdisciplinary nature of FEP, connecting neuroscience and robotics.

4.3.1 Neuroscience

FEP emerges from neuroscience as a unified theory of the brain which posits that all biological systems resist their natural tendency to disorder by minimizing their free energy [199], where free energy is an information theoretic measure that bounds sensory surprisal. FEP provides a mathematical formalism for the brain related functions [187], unifies action and perception [50], connects memory and attention [187] and explains Freudian ideas [188]. The brain inspired nature of FEP has already attracted roboticists to apply it to build intelligent agents [25]. A few of them includes the body perception of humanoid robots [200], control of manipulator robot [27], system identification of a quadrotor [166, 167], SLAM [34], PID controller [47] etc. With this work we aim to assess the performance of DEM for state and input estimation of a quadrotor under wind conditions.

4.3.2 Robotics and control systems

In control systems literature, numerous approaches are used to deal with colored noise during state estimation. State Augmentation (SA) assumes the colored process noise as an auto-regressive (AR) noise and augments the state space equation to transform it into an equivalent system influenced by white noise [196]. The Measurement Differencing [201] approach deals with handling colored measurement noise. Second Moment Information Kalman filter (SMIKF) [202] extends KF for coloured noise by incorporating the temporal correlations of the AR noise into the prior covariance calculation of KF. In the fault detection literature, many observers have been developed for input estimation, like the Unknown Input Observer (UIO) [183]. However, none of these methods perform simultaneous state and input estimation under colored noise other than DEM [46].

In robotics, different approaches are used for state estimation of quadrotors under wind conditions. The most common approach (Dryden wind model) is to treat wind as a

colored noise shaped by a filter acting on the white noise. Another approach is to model the wind dynamics and estimate wind velocity using complex nonlinear models [203]. Using additional cameras for accurate state estimation is another solution [204]. Our approach differs from these methods as we treat the wind noises as colored and use the information in the noise derivatives for accurate state and input estimation.

4.4 PROBLEM STATEMENT

Consider the plant dynamics given in Equation 4.1, where A, B and C are constant system matrices, $\mathbf{x} \in \mathbb{R}^n$ is the hidden state, $\mathbf{v} \in \mathbb{R}^r$ is the input and $\mathbf{y} \in \mathbb{R}^m$ is the output.

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{v} + \mathbf{w}, \qquad \mathbf{y} = C\mathbf{x} + \mathbf{z}. \tag{4.1}$$

Here $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$ are temporally correlated (colored) and represent the process and measurement noise respectively. The noises are assumed to be the result of the convolution of a Gaussian kernel on a white noise signal. The goal of the DEM observer is to simultaneously estimate $\mathbf{x}$ and $\mathbf{v}$, when the noises are colored (or non-white). The goal of this paper is to design an experimental setup for a real robot that can be used to validate the DEM observer and its usefulness in the presence of colored noise.

4.5 PRELIMINARIES

This section introduces the DEM observer fundamentals.

4.5.1 FREE ENERGY PRINCIPLE

Fundamentally based on Bayesian Inference, FEP estimates the posterior probability $p(\vartheta/y) = p(\vartheta, y)/\int p(\vartheta, y)d\vartheta$, where ϑ is the component to be estimated and y is the measurement [43]. The presence of an intractable integral motivates the use of a variational density $q(\vartheta)$, called the recognition density that approximates the posterior as $q(\vartheta) \approx p(\vartheta/y)$. This approximation is achieved by minimizing the Kullback-Leibler (KL) divergence of the distributions given by $KL(q(\vartheta)\|p(\vartheta/y)) = \langle \ln q(\vartheta) \rangle_{q(\vartheta)} - \langle \ln p(\vartheta/y) \rangle_{q(\vartheta)}$, where $\langle . \rangle_{q(\vartheta)}$ represents the expectation over $q(\vartheta)$. Upon simplification using $p(\vartheta/y) = p(\vartheta, y)/p(y)$, it can be rewritten as [199]:

$$\ln p(y) = F + KL(q(\vartheta)\|p(\vartheta|y)), \tag{4.2}$$

where $F = \langle \ln p(\vartheta, y) \rangle_{q(\vartheta)} - \langle \ln q(\vartheta) \rangle_{q(\vartheta)}$ is the free energy. Since $\ln p(y)$ is independent of ϑ, minimization of the KL divergence involves the maximization of free energy [1]. This is the fundamental idea behind using free energy as the proxy for brain's inference through the minimization of its sensory surprisal [199]. DEM uses this mathematical framework, in conjunction with the use of generalized coordinates to provide a hierarchical brain model [44]. We will be using a reformulated version of DEM given in [46] for this work.

4.5.2 GENERATIVE MODEL

The key concept that differentiates DEM from other methods is its use of generalized coordinates for noise color handling. This is done by keeping track of the trajectory

---

[1]maximization of the ELBO term (Section 1.4)

of all time-varying quantities (instead of only its point estimates) through a vector of derivatives. The state vector in generalized coordinates are written using a tilde operator as $\tilde{x} = [x \; x' \; x'' \; ....]^T$ where the dash operator represents the derivatives. Since the noises are colored, the higher derivatives of the system model can be written as [170]:

$$
\begin{aligned}
x' &= Ax + Bv + w && y = Cx + z \\
x'' &= Ax' + Bv' + w' && y' = Cx' + z' \\
&\; ... && \; ...
\end{aligned}
\tag{4.3}
$$

which can be compactly written as:

$$
\dot{\tilde{x}} = D^x \tilde{x} = \tilde{A}\tilde{x} + \tilde{B}\tilde{v} + \tilde{w} \qquad \tilde{y} = \tilde{C}\tilde{x} + \tilde{z}
\tag{4.4}
$$

where $D^x = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & 0 & 1 \\ & & & & 0 \end{bmatrix}_{(p+1)\times(p+1)} \otimes I_{n\times n}$.

Here, D^x represents the shift matrix, which performs the derivative operation on the generalized state vector. Similarly, D^v performs the same operation on inputs and has size $r(d+1) \times r(d+1)$. p and d represent the embedding order for the hidden states and the inputs respectively, indicating the number of derivatives used. The generalized system matrices are given by $\tilde{A} = I_{p+1} \otimes A$, $\tilde{B} = I_{p+1} \otimes B$, $\tilde{C} = I_{p+1} \otimes C$, where I denotes the identity matrix and $\otimes$ the Kronecker tensor product. The generalized output $\tilde{y}$ is calculated from the discrete measurements $\hat{y} = \begin{bmatrix} \vdots \\ y(t-dt) \\ y(t) \\ y(t+dt) \\ ... \end{bmatrix}_{m(p+1)}$ using the methodology in [170], resulting in

a latency of $\frac{p}{2}dt$ during online estimation, which is negligible for the large sampling rate (120Hz) used in this paper.

4.5.3 Noise modeling

The use of generalized coordinates helps to model the noise color through the temporal precision matrix of the noise derivatives. In DEM, the noise is assumed to be the result of a white noise signal that has been convoluted using a Gaussian filter of the form: $K(t) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{1}{2}(\frac{t}{\sigma})^2)$. This provides an easy computation of the covariance of the noise derivatives using the temporal precision matrix S [170]:

$$
S(\sigma^2) = \begin{bmatrix}
1 & 0 & -\frac{1}{2\sigma^2} & .. \\
0 & \frac{1}{2\sigma^2} & 0 & .. \\
-\frac{1}{2\sigma^2} & 0 & \frac{3}{4\sigma^4} & .. \\
.. & .. & .. & ..
\end{bmatrix}^{-1}_{(p+1)\times(p+1)}
\tag{4.5}
$$

σ is close to zero for white noise, while $\sigma > 0$ for colored noise. The generalized noise precision matrix can be written using S as $\tilde{\Pi} = diag(\tilde{\Pi}^z, P^{\tilde{v}}, \tilde{\Pi}^w)$, where $\tilde{\Pi}^z = S \otimes \Pi^z$, $\tilde{\Pi}^w = S \otimes \Pi^w$, and $P^{\tilde{v}} = S \otimes P^v$. Here Π^w and Π^z are the noise precisions (inverse covariance), and P^v is the prior precision on inputs.

4.5.4 State and Input Observer

The DEM observer in [46] simultaneously estimates the generalized state and input vector $X = \begin{bmatrix} \tilde{x} \\ \tilde{v} \end{bmatrix}$ through the gradient ascend over its variational free energy $V(t)$:

$$\dot{X} = kV(t)_X + D^X X, \tag{4.6}$$

where k is the learning rate, $V(t)_X$ is the gradient of $V(t)$ with respect to X and $D^X = \begin{bmatrix} D^x & O \\ O & D^v \end{bmatrix}$. Using the Laplace approximation [205], simplifies $V(t)$ as the precision weighted prediction error, $V(t) = -\frac{1}{2}\tilde{\epsilon}^T \tilde{\Pi}\tilde{\epsilon}$, where $\tilde{\epsilon}$ is the prediction error given by:

$$\tilde{\epsilon} = \begin{bmatrix} \tilde{y} - \tilde{C}\tilde{x} \\ \tilde{v} - \eta^{\tilde{v}} \\ D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix} \tag{4.7}$$

Here $\eta^{\tilde{v}}$ denotes the prior on the input. Therefore, $V(t)_X = -\tilde{\epsilon}_X^T \tilde{\Pi}\tilde{\epsilon}$, where $\tilde{\epsilon}_X$ is given by:

$$\tilde{\epsilon}_X = \begin{bmatrix} -\tilde{C} & O \\ O & I \\ D^x - \tilde{A} & -\tilde{B} \end{bmatrix}. \tag{4.8}$$

Substituting these results to Equation 4.6 upon simplification yields the DEM state and input observer of [46]:

$$\dot{X} = \begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{v}} \end{bmatrix} = A_1 \begin{bmatrix} \tilde{x} \\ \tilde{v} \end{bmatrix} + B_1 \begin{bmatrix} \tilde{\mathbf{y}} \\ -\tilde{\eta} \end{bmatrix} \text{ and } Y = X, \tag{4.9}$$

where Y is the output of the observer, $A_1 = D^X - kA_2$,

$$A_2 = \begin{bmatrix} \tilde{C}^T\tilde{\Pi}^z\tilde{C} + (D^A)^T\tilde{\Pi}^w D^A & -(D^A)^T\tilde{\Pi}^w\tilde{B} \\ -\tilde{B}^T\tilde{\Pi}^w D^A & \tilde{P}^v + \tilde{B}^T\tilde{\Pi}^w\tilde{B} \end{bmatrix}, \tag{4.10}$$

$B_1 = -\begin{bmatrix} -\tilde{C}^T\tilde{\Pi}^z & O \\ O & \tilde{P}^v \end{bmatrix}$ and $D^A = D^x - \tilde{A}$.

This observer was proved to outperform the KF for state estimation on LTI systems with colored noise in simulation [46]. We will use an exact discretization of this observer throughout the paper to provide the experimental validation on a real robot.

4.5.5 Uncertainty in state and input estimation

DEM provides a means to compute the uncertainty in estimation through the precision of estimates given by the negative curvature of variational free energy [97]:

$$\Pi^X = -V(t)_{XX} = \tilde{\epsilon}_X^T\tilde{\Pi}\tilde{\epsilon}_X = A_2, \tag{4.11}$$

where A_2 is given in Equation 4.10. Therefore, the precision of DEM's state and input estimates is independent of time, and is given by $\Pi^{\tilde{x}\tilde{x}} = \tilde{C}^T\tilde{\Pi}^z\tilde{C} + (D^A)^T\tilde{\Pi}^w D^A$ and $\Pi^{\tilde{v}\tilde{v}} = \tilde{P}^v + \tilde{B}^T\tilde{\Pi}^w\tilde{B}$ respectively.

4.6 EXPERIMENTAL DESIGN

The distinctive feature of DEM that enables it to handle colored noise (to outperform a KF for state estimation) is its use of generalized coordinates [46, 170]. This section aims at designing an experimental setup (as simple as possible) for real robots that can leverage this property and provide a proof of concept for our DEM-based state and input observer design for LTI systems with colored noise [46].

4.6.1 EXPERIMENTAL SETUP

Our experimental setup consist of a quadrotor (Parrot AR.drone 2.0) hovering in wind produced by a blower in a controlled lab, as shown in Figure 4.2. The blower induces wind in the negative y direction, against the hovering quadcopter. We use an OptiTrack motion capture system to record the position and orientation of the quadcopter. The PID controller tries to resist the wind to hover the quadrotor at the given position $(0\,m, 0\,m, 1\,m)$ and orientation $(0°, 0°, 0°)$, using the onboard sensor data. A total of 9 hovering experiments were performed - 4 experiments without wind (blower off) and 4 experiments with wind (blower on). The final experiment was used to tune all the benchmark observers and will not be used for benchmarking. Each experiment lasted $10\,s$ with $dt = 0.0083\,s$. The Optitrack pose and PID control signals were recorded for offline evaluations.
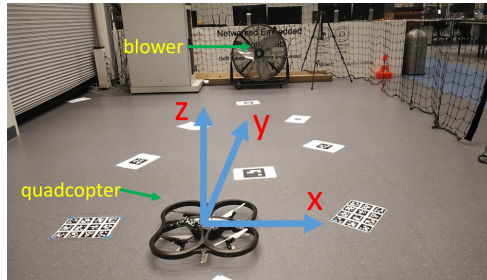
Fig 4.2: The controlled lab environment for the experimental setup with the quadrotor and the blower.

Since wind is the result of another (unmodelled) dynamic system, we hypothesize that the introduction of wind influences the quadrotor dynamics and acts as the source of colored noise to the system. The experimental design enables us to control the level of noise color entering the system by controlling the blower for its wind speed and direction. We hypothesize that our DEM observer can leverage on the information contained in the colored noise by keeping track of the higher derivatives of states and inputs through the generalized coordinates.

4.6.2 QUADROTOR MODEL SELECTION

The quadrotor model selection was performed to accommodate the influence of wind using minimum number of states, resulting in a controllable and observable LTI system. Since the wind flows in negative y direction, it influences the roll angle (ϕ around x-axis) and the roll angular velocity ($\dot{\phi}$) the most. Therefore, we only consider states $x = \begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix}$. The model involving these states is based on the one provided in [206]. By assuming small angles, ϕ

and $\dot{\phi}$ can be decoupled from the other system dynamics. Linearizing these states around hovering conditions gives:

$$
\begin{bmatrix} \dot{\phi} \\ \ddot{\phi} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{c_{B\phi}}{I_{xx}} & -\frac{c_{B\phi}}{I_{xx}} & -\frac{c_{B\phi}}{I_{xx}} & \frac{c_{B\phi}}{I_{xx}} \end{bmatrix} \begin{bmatrix} pwm_1 \\ pwm_2 \\ pwm_3 \\ pwm_4 \end{bmatrix}
$$
$$
y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix}
$$

(4.12)

Here pwm_i is the Pulse Width Modulation signal provided to the ith motor by the controller for stable hovering. I_{xx} is the quadcopter's moment of inertia around the x-axis. It's value is identified using the bifilar pendulum experiment and equals $3.4 \cdot 10^{-3} kg\,m^2$. $c_{B\phi}$ is the thrust coefficient that models the relation between the PWM values and the thrust generated by the quadcopter rotors. It's value is obtained by averaging the results of several static thrust tests and equals $1.274 \cdot 10^{-3} N\,m$.

We normalize the input pwm signals using $v = \frac{v - mean(v)}{max(v) - min(v)}$ and use the same factor to multiply the B matrix, such that the system dynamics are unaltered. See [207] for more details regarding the model derivation, system identification procedure and experimental setup.

Since we use an accurate measurement system, Π^z is very high for all experiments. However, the presence of colored process noise w through wind makes $\Pi^w << \Pi^z$. Π^w is further influenced by the modelling errors during linearization as the wind aggressively drives the quadrotor away from its equilibrium.

4.7 Results and analysis

This section aims to investigate the validity of the assumptions in our experimental design and to compare the performance of DEM observer against other benchmarks.

4.7.1 Validity of Laplace Approximation

The DEM framework approximates the probability densities of $p(\tilde{y})$ and $p(\tilde{x}/\tilde{v})$ to be Gaussian in nature, centred around their mean predictions ($\tilde{C}\tilde{x}$ and $\tilde{A}\tilde{x} + \tilde{B}\tilde{v}$, respectively) with the same precision as that of the noises ($\tilde{\Pi}^z$ and $\tilde{\Pi}^w$):

$$
p(\tilde{y}) = \frac{1}{\sqrt{(2\pi)^{m(p+1)}|\tilde{\Sigma}^z|}} e^{-\frac{1}{2}(\tilde{y} - \tilde{C}\tilde{x})^T \tilde{\Pi}^z (\tilde{y} - \tilde{C}\tilde{x})},
$$
$$
p(\tilde{x}/\tilde{v}) = \frac{1}{\sqrt{(2\pi)^{n(p+1)}|\tilde{\Sigma}^w|}} e^{-\frac{1}{2}\tilde{\epsilon}^{xT} \tilde{\Pi}^w \tilde{\epsilon}^x},
$$

(4.13)

where $\tilde{\epsilon}^x = D^x \tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v}$. The validity of this approximation on our experimental design was investigated by plotting the process noise histograms for both without wind and with wind conditions (for 400 data points each) and is shown in Figure 4.3a and 4.3b respectively. Similar trend holds for measurement noise as well. The strong Gaussian fit indicates the validity of Laplace approximation for our experimental design.
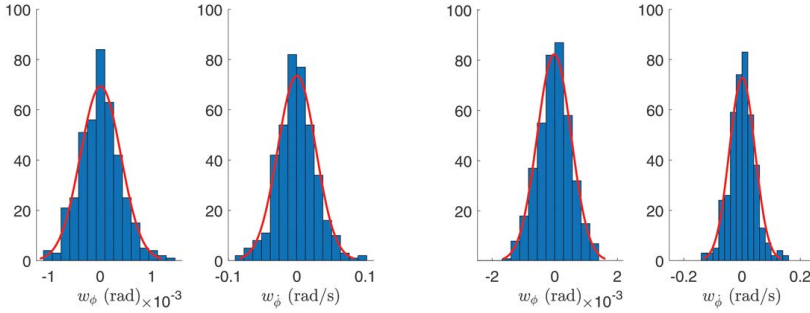
(a) Histograms of the process noise w_ϕ and $w_{\dot\phi}$ with a Gaussian fit for no wind conditions.

(b) Histograms of the process noise w_ϕ and $w_{\dot\phi}$ with a Gaussian fit under wind conditions.
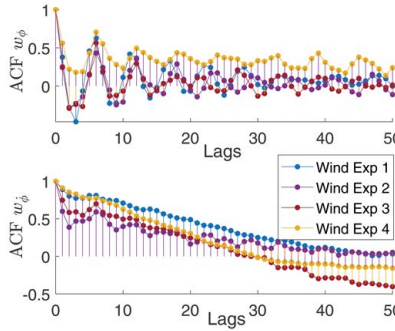
(c) The auto-correlations for the process noises w_ϕ and $w_{\dot\phi}$ under wind conditions.

Fig 4.3: The properties of process noise of our experiment. The wind introduces a colored Gaussian distributed disturbance to the system.

4.7.2 INFLUENCE OF WIND ON STATES AND PROCESS NOISE

In this section we validate the direct influence of wind on the states and process noise. Table 4.1 demonstrates a higher standard deviation for windy conditions than for non-windy conditions. A similar trend can be observed from the width of histograms in Figure 4.3a and 4.3b, indicating that our experimental design can control the noise generation.

4.7.3 CONFIRMATION OF NOISE COLOR

In this section we confirm that our experimental design generates colored process noise. Figure 4.3c shows the sample auto-correlation of the process noise of all experiments (with wind). There is stronger autocorrelation for $w_{\dot\phi}$ than for w_ϕ, because ϕ is observed. The auto-correlation is different from that expected from a white noise signal where the auto-correlation immediately drops to 0 after zero lag. This confirms the presence of strong noise color (time-correlated noise) in data.

	ϕ (rad)	$\dot{\phi}$ (rad/s)	w_ϕ (rad)	$w_{\dot{\phi}}$ (rad/s)
Without wind	0.00855	0.0544	0.000416	0.0284
With wind	0.0460	0.260	0.000937	0.0607

Table 4.1: The standard deviations of the states, ϕ and $\dot{\phi}$, and the process noises, w_ϕ and $w_{\dot{\phi}}$, for experiments with and without wind.

4.7.4 Estimator settings for benchmarking

We aim to benchmark the state estimation against KF, SMIKF and SA, and the input estimation against UIO, for a total of 8 experiments (4 with and 4 without wind). All methods use the same data **y** and initial condition $x(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. DEM was set with learning rate $k = 1$, and the order of generalized motion of states and inputs to $p = 6$ and $d = 2$ respectively. The SMIKF implementation could only accommodate a first order AR model, while the SA implementation uses a 6th order AR model, similar to the 6th order derivatives (p) of DEM. The noise precision Π^w was calculated for each experiment, while $\Pi^z = 8.1 \cdot 10^{-9}$ was calculated from static drone data. The 9th experiment was used to tune the noise smoothness to $\sigma = 0.006$, which was used for all experiments. The computational complexities of SMIKF, SA and DEM are theoretically higher than KF.

4.7.5 State estimation - benchmarking

In this section, we compare the performance of DEM with the aforementioned benchmarks for state estimation with known inputs. Figure 4.4a shows the state estimates of all benchmarks for an experiment with wind (zoomed for visualization). Although most benchmarks follow the general trend of the measured states (in blue), DEM performs the best. KF shows an inferior performance due to its incapability of dealing with colored noise. We use the sum of squared errors (SSE) between the estimate of $\dot{\phi}$ and its measurement as the metric to denote the quality of state estimation. The average SSE of all 4 experiments (with and without wind separately) for all benchmarks are shown in Figure 4.4b. DEM outperforms other benchmarks in state estimation under wind conditions with minimal SSE, demonstrating that it is a competitive state estimator.

4.7.6 Role of Generalized Coordinates

One of the main strengths of DEM - the capability to deal with colored noise - comes from the use of generalized coordinates. In this section, we demonstrate the usefulness of generalized coordinates in state estimation on experimental data. The mean (and standard deviations of) SSE of state estimation for all experiments (with wind) for varying orders of generalized motion p is shown in Figure 4.4c. The exponential decrease in SSE is consistent with the results from [46] on large simulated data, and indicates the importance of generalized coordinates in accurate state estimation in the presence of colored noise.

4.7.7 State estimation as free energy maximization

The fundamental idea behind state estimation using DEM is the gradient ascent over the variational free energy manifold [2]. In this section, we visually demonstrate that DEM's

---

[2] maximization of the ELBO term (Section 1.4)
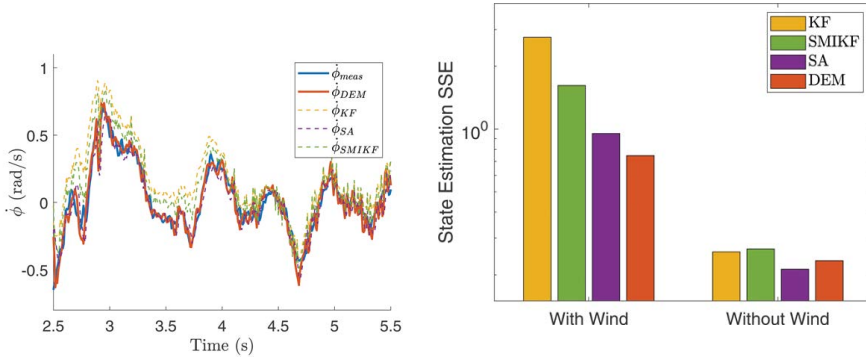
(a) State estimation benchmarks with wind.   (b) Average SSE of all flights.
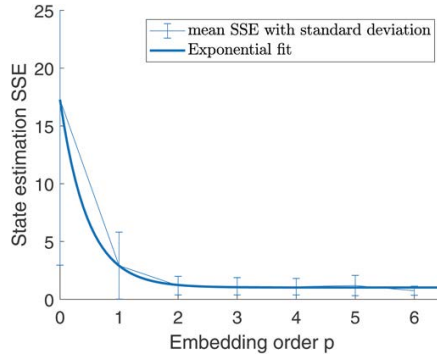
(c) Average SSE drops exponentially with p.

Fig 4.4: DEM outperforms other benchmarks with minimal SSE in state estimation for a quadrotor flying under windy conditions. The performance of DEM improves exponentially with higher orders of generalized motion p, for a quadrotor flying under windy conditions, highlighting the importance of generalized coordinates in the presence of colored noise.

state estimates for flight experiment maximize $V(t)$. Figure 4.5 shows that the DEM state estimate is on top of the $V(t)$ curve at each time instance.

4.7.8 INPUT ESTIMATION - BENCHMARKING

In this section, we aim to demonstrate our DEM observer's capability to estimate inputs in real robot application and benchmark it against an input observer (UIO) from control systems. We use the same settings as given in Section 4.7.4, except for providing the input priors for pwm_1 with a wrong value of $\eta^{pwm_1} = 0.5$ with a low precision of $P^{pwm_1} = 1$ to encourage exploration and $\Pi^w = e^3 I_2$. We use $C = I$ for this section to meet the observability requirements of our benchmark (UIO). Both DEM and UIO estimated the first pwm signal and the result is shown in Figure 4.6. Both DEM and UIO followed the trend of measured inputs (in blue).

The coinciding input estimates for DEM and UIO demonstrate that both estimators behave the same. The estimation was repeated for all experiments and the SSE for input
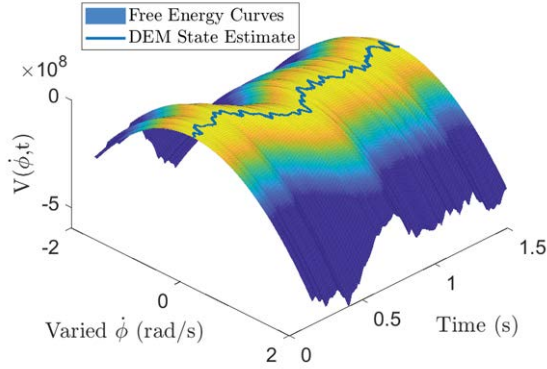
Fig 4.5: The DEM state estimate (blue curve) lies on top of the variational free energy surface, indicating that the DEM observer maximized $V(t)$. Plot zoomed for visualization.
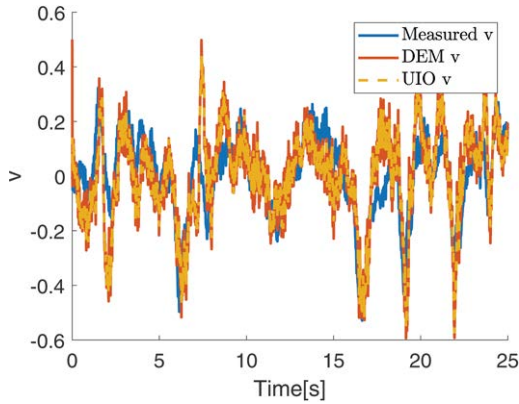
Fig 4.6: DEM's input estimation coinciding with that of UIO.

estimation is shown in Figure 4.7. This confirms the similarity in performance of UIO and DEM for input estimation in the presence of colored noise.

4.7.9 Accuracy v/s complexity

The inherent capability of DEM to balance between estimation accuracy and complexity is mediated by the priors η^v and P^v [46]. Here, accuracy is the measure of closeness of estimates to the real measurement, and complexity is the measure of closeness to the priors. This section aims at demonstrating this balance for simultaneous state and input estimation on quadrotor data. This section follows the same settings as Section 4.7.4 with $\Pi^w = e^3 I_2$. Simultaneous state and input estimation was performed using wrong input prior $\eta^v = 1$ for varying prior precisions P^v, and the resulting input estimation is shown in Figure 4.8, along with the measured input (in blue). As P^v is relaxed, the input estimation moves away
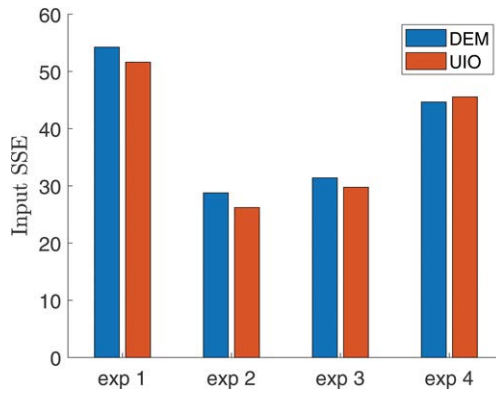
Fig 4.7: Similar performance of DEM and UIO for input estimation indicated by similar SSE in input estimation.
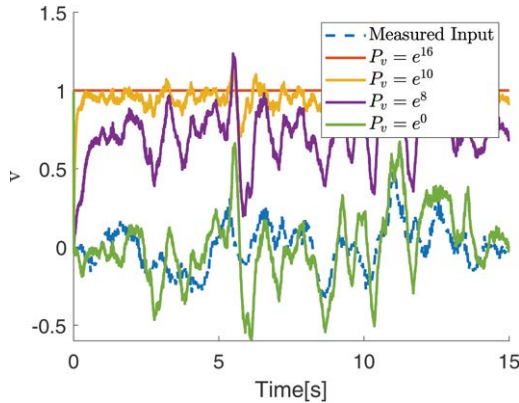
Fig 4.8: The input estimates moving from a wrong prior of $\eta^v = 1$ to the measured input (in blue), mediated by the prior precision P_v.

from the wrong prior η^v and moves closer to the correct inputs. The shift from wrong priors to the correct measurements, mediated by P^v can be seen as a balance (trade-off) between complexity and accuracy. Figure 4.9 demonstrates this balance for all experiments with windy conditions, both for state and input estimation. The increasing SSE for higher P^v indicates the shift from high accuracy with low complexity region to the low accuracy with high complexity region. This trade-off is useful mainly in industrial fault detection systems where any major deviations from the prior (known) inputs could be detected and isolated during runtime. DEM's inherent capability to balance accuracy and complexity is an added advantage when compared to other input estimators in literature like UIO.
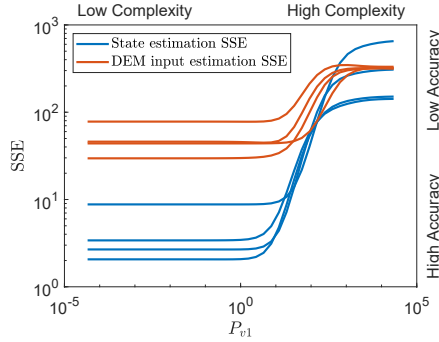
Fig 4.9: The SSE plot of state and input estimation demonstrating DEM's accuracy-complexity tradeoff. The SSE moves from a region of low complexity and high accuracy to a region of high complexity and low accuracy as the prior precision P^v is increased.

4.8 CONCLUSIONS AND FUTURE WORK

The FEP based perception scheme called DEM, has recently been reformulated into a simulataneous state and input observer for LTI systems under colored noise. With this paper, we propose an experimental design to validate the DEM observer on real robots. Through a series of quadrotor experiments under wind conditions, we show that the DEM based observer outperforms other benchmarks like KF, SMIKF and SA for state estimation with minimum estimation errors. We show that DEM's input estimation shows similar performance compared to classical input observers like UIO. With this paper, we provide the first experimental validation for the use of generalized coordinates to deal with colored noise during state and input estimation on real robots. We further demonstrate the unique capability of DEM to balance between accuracy and complexity during state and input estimation. The main challenge of the DEM observer is the need to know the noise precision and noise smoothness *a priori*. We intend to extend this work for simultaneous noise precision and smoothness estimation in future.

4.9 ACKNOWLEDGEMENT

5

Noise Smoothness Estimation

*This section uses **precision learning** to learn the noise precision matrix (parametrised by noise smoothness s, as given in Equation 2.7) using FEP. This chapter extends the DEM based observer design in Chapter 3 for a joint state and noise smoothness estimation, and provides an experimental confirmation using the robot data from Chapter 4.*

5.1 Abstract

The free energy principle (FEP) from neuroscience provides a framework called active inference for the joint estimation and control of state space systems, subjected to colored noise. However, the active inference community has been challenged with the critical task of manually tuning the noise smoothness parameter. To solve this problem, we introduce a novel online noise smoothness estimator based on the idea of free energy principle. We mathematically show that our estimator can converge to the free energy optimum during smoothness estimation. Using this formulation, we introduce a joint state and noise smoothness observer design called DEMs. Through rigorous simulations, we show that DEMs outperforms state-of-the-art state observers with least state estimation error. Finally, we provide a proof of concept for DEMs by applying it on a real life robotics problem - state estimation of a quadrotor hovering in wind, demonstrating its practical use.

5.2 Introduction

The rising demand for autonomous drone delivery systems has increased the need for accurate state observers that are robust against uncertain events like strong wind currents. These unmodelled wind currents induce colored noise to the system, hindering the safe operation of drones. We take a step in this direction by using the ideas from computational neuroscience to introduce a novel state and noise smoothness observer design for linear systems with colored noise.

The classical linear estimators like Kalman Filter (KF) assumes the noises to be white. This assumption is often violated in practice, resulting in a sub-optimal estimation [208]. Many adaptations on KF have been introduced to overcome this challenge, including Second Moment Information Kalman filter (SMIKF) [202], State Augmentation (SA) [196], Measurement Differencing (MD) [201] etc. Dynamic Expectation Maximization (DEM) [170], based on Free Energy Principle (FEP) [2] from neuroscience has recently been used to design state and input observers [46] that has shown to outperform the classical methods both in simulation and in real robot experiments [165]. However, DEM requires the prior knowledge of the noise smoothness parameter for the accurate state estimation. To solve this problem, we introduce a novel online noise smoothness estimator based on FEP, for linear systems with colored noise. The core contributions of the paper include:

1. introduction of an online smoothness estimator for the state estimation of linear systems under colored noise,

2. extensive evaluation of the estimator in simulation and its validation on a real robot (quadrotor flight) data.

5.3 Related work

This section highlights the interdisciplinary nature of FEP with related works in neuroscience and robotics literature.

5.3.1 Neuroscience

According to FEP, all biological systems resist their natural tendency to disorder by minimizing an information theoretic measure called free energy [2], which bounds its sensory

surprisal. FEP emerges as a unified theory of the brain by providing a mathematical formalism for brain functions [187], unifying action and perception [50], explaining Freudian ideas [188], and connecting memory and attention [187]. The work closest to our proposed idea is the Generalized filtering [169] that uses FEP for noise smoothness estimation during the inversion of dynamic models of the brain (fMRI data) [209]. We extend this idea into robotics to design an online state and noise smoothness observer for applications like quadrotor flights with wind as colored noise.

5.3.2 Robotics and control

Numerous approaches for state estimation under colored noise exists in the control systems literature [208]. SA models the colored process noise in a state space system as autoregressive (AR) noise. The system is then transformed into an equivalent augmented system influenced by white noise [196]. Another approach is to incorporate the temporal correlations of the AR noise into the prior covariance calculation of KF, resulting in an extended KF called SMIKF [202]. MD [201] approach deals with handling colored measurement noise. However, the white noise assumption is prevalent in robotics for the state estimation of a quadrotor [210], which might not be effective in outdoor windy conditions [165]. Our work fills this research gap by providing an online noise smoothness estimator.

The brain inspired nature of FEP has already inspired the development of intelligent agents [25] – body perception of humanoid robots [200], estimation and control of manipulator robot [211], system identification of a quadrotor [166], SLAM [34], PID controller [47], KF [45, 165] etc. These active inference applications can employ our noise estimator for better estimation and control of robots during colored noise.

5.4 Problem statement

Consider the linear plant dynamics given in Equation 5.1 where A, B and C are constant system matrices, $\mathbf{x} \in \mathbb{R}^n$ is the hidden state, $\mathbf{v} \in \mathbb{R}^r$ is the input and $\mathbf{y} \in \mathbb{R}^m$ is the output.

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{v} + \mathbf{w}, \qquad \mathbf{y} = C\mathbf{x} + \mathbf{z}. \tag{5.1}$$

Here $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$ represent the process and measurement noise with noise precision (inverse covariance) Π^w and Π^z respectively. Variables of the plant are denoted in boldface, while its estimates are denoted in non-boldface. In this paper, the white noise is convoluted with a Gaussian filter of kernel width s to generate the colored noise.

The problem considered in this paper is the state (x) and noise smoothness (s) observer design (DEMs) for the linear system given in Equation 5.1, subjected to colored noise. We show that our observer outperforms state-of-the-art state observers, both in simulation (Section 5.9) and on real robot data (Section 5.10).

5.5 Noise color modelling

The two key concepts behind the success of DEM in handling the colored noise are i) the use of generalized coordinates and ii) the noise precision modelling. This section aims to elaborate on these theoretical concepts.

5.5.1 Generalized coordinates

Generalized coordinates is a vector representation of the trajectory of a time varying quantity (x, v, y) using a collection of its higher order derivatives. For example, the state vector in generalized coordinates is written using a tilde operator as $\tilde{x} = [x \; x' \; x'' \; ....]^T$, where the dash operator represents the derivatives. The key advantage is the capability to track the trajectory of states, unlike the classical estimators that track only the point estimates. This provides additional data for DEM during estimation, resulting in its superior performance during state estimation under colored noise. Since the noises are colored, the higher derivatives of the system model can be written as [170]:

$$
\begin{aligned}
x' &= Ax + Bv + w & y &= Cx + z \\
x'' &= Ax' + Bv' + w' & y' &= Cx' + z' \\
&... & &...
\end{aligned}
\tag{5.2}
$$

which can be compactly written as:

$$
\dot{\tilde{x}} = D^x \tilde{x} = \tilde{A}\tilde{x} + \tilde{B}\tilde{v} + \tilde{w} \qquad \tilde{y} = \tilde{C}\tilde{x} + \tilde{z}
\tag{5.3}
$$

where $D^x = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}_{(p+1)\times(p+1)} \otimes I_{n\times n}$.

Here, D^x represents the shift matrix, which performs the derivative operation on the generalized state vector. The embedding order representing the number of derivatives of hidden states and inputs used is denoted by p and d respectively. The generalized system matrices are given by $\tilde{A} = I_{p+1} \otimes A, \quad \tilde{B} = I_{p+1} \otimes B, \quad \tilde{C} = I_{p+1} \otimes C$, where I denotes the identity matrix and $\otimes$ the Kronecker tensor product.

5.5.2 Noise precision modelling

The second key concept is the modelling of generalized noise precision (inverse covariance) matrix $\tilde{\Pi}$. Since the noises are assumed to be Gaussian convoluted white noise, the covariance matrix embedding the relation between noise derivatives take a specific structure [170]. The smoothness matrix defining this relation for $p = 6$ is calculated as [46]:

$$
S = \begin{bmatrix}
\frac{35}{16} & 0 & \frac{35}{8}s^2 & 0 & \frac{7}{4}s^4 & 0 & \frac{1}{6}s^6 \\
0 & \frac{35}{4}s^2 & 0 & 7s^4 & 0 & s^6 & 0 \\
\frac{35}{8}s^2 & 0 & \frac{77}{4}s^4 & 0 & \frac{19}{2}s^6 & 0 & s^8 \\
0 & 7s^4 & 0 & 8s^6 & 0 & \frac{4}{3}s^8 & 0 \\
\frac{7}{4}s^4 & 0 & \frac{19}{2}s^6 & 0 & \frac{17}{3}s^8 & 0 & \frac{2}{3}s^{10} \\
0 & s^6 & 0 & \frac{4}{3}s^8 & 0 & \frac{4}{15}s^{10} & 0 \\
\frac{1}{6}s^6 & 0 & s^8 & 0 & \frac{2}{3}s^{10} & 0 & \frac{4}{45}s^{12}
\end{bmatrix},
\tag{5.4}
$$

where s is the kernel width of the Gaussian filter. $s \approx 0$ is high frequency white noise. Since $s < 1$ second for practical cases (sensors have high sampling rate), the first elements in S matrix are higher than the last ones, implying a higher correlation between the first noise derivatives (more smooth) than the last derivatives (less smooth). The combined (generalized) noise precision matrix can be written using the S matrix as:

$$
\tilde{\Pi} = \begin{bmatrix} S \otimes \Pi^z & 0 \\ 0 & S \otimes \Pi^w \end{bmatrix}.
\tag{5.5}
$$

With the key concepts in place, the next section derives the free energy formulations that are necessary for the observer design in Section 5.7.

5.6 Free energy optimization

FEP uses Bayesian Inference to estimate the posterior probability $p(\vartheta/\mathbf{y}) = p(\vartheta, \mathbf{y}/\int p(\vartheta, \mathbf{y})d\vartheta$, where ϑ is the component to be estimated ($\vartheta = \{\tilde{x}, s\}$), and $\mathbf{y}$ is the measurement [43]. The presence of an intractable integral motivates the use of a variational density $q(\vartheta)$, called the recognition density that approximates the posterior as $q(\vartheta) \approx p(\vartheta/\mathbf{y})$. This approximation is achieved by minimizing the Kullback-Leibler (KL) divergence of the distributions given by $KL(q(\vartheta)\|p(\vartheta/\mathbf{y})) = \langle \ln q(\vartheta) \rangle_{q(\vartheta)} - \langle \ln p(\vartheta/\mathbf{y}) \rangle_{q(\vartheta)}$, where $\langle . \rangle_{q(\vartheta)}$ represents the expectation over $q(\vartheta)$. Upon simplification using $p(\vartheta/\mathbf{y}) = p(\vartheta, \mathbf{y})/p(\mathbf{y})$, it reduces to [2]:

$$\ln p(y) = F + KL(q(\vartheta)\|p(\vartheta|\mathbf{y})), \tag{5.6}$$

where $F = \langle \ln p(\vartheta, \mathbf{y}) \rangle_{q(\vartheta)} - \langle \ln q(\vartheta) \rangle_{q(\vartheta)}$ is the free energy. The minimization of KL divergence results in the maximization of free energy[1], as $\ln p(\mathbf{y})$ is independent of ϑ. This is the core idea behind using free energy minimization as a proxy for brain's inference, thereby minimizing the brain's sensory surprisal [2].

We use this idea from free energy principle for the joint observer design for $\tilde{x}$ and s through two fundamental assumptions about $q(\vartheta) = q(\tilde{x}, s)$: i) Mean field assumption [170] that facilitates a conditional independence between the subdensities, $q(\vartheta) = q(\tilde{x})q(s)$, and ii) Laplace assumption [205] that facilitates the use of Gaussian distributions with mean μ and variance Σ over these subdensities, $q(\tilde{x}) = \mathcal{N}(\tilde{x} : \mu^{\tilde{x}}, \Sigma^{\tilde{x}})$ and $q(s) = \mathcal{N}(s : \mu^s, \Sigma^s)$. We refer to [97] for an elaborate read on similar simplifications. Under these assumptions, F reduces to the sum of precision weighted prediction errors and the information entropy as:

$$F = -\frac{1}{2}\tilde{\epsilon}^T \tilde{\Pi} \tilde{\epsilon} + \frac{1}{2}\ln|\tilde{\Pi}| - \frac{1}{2}\epsilon^{sT}\Pi^s\epsilon^s + \frac{1}{2}\ln|\Pi^s|, \tag{5.7}$$

where $\tilde{\epsilon}$ is the combined prediction error for outputs and states, and ϵ^s is the prediction error for s, given by:

$$\tilde{\epsilon} = \begin{bmatrix} \tilde{\mathbf{y}} - \tilde{C}\tilde{x} \\ D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix}, \text{ and } \epsilon^s = s - \eta^s. \tag{5.8}$$

Here η^s and Π^s are the prior smoothness and its prior precision (confidence). For this work, we use a low prior $\eta^s \approx 0$ with a low precision $\Pi^s = 1$ (practically, $0 < s < 1$). Our observer design follows a constant precision approach where Π^s is kept constant throughout the estimation of s, unlike the precision update approach of DEM [97, 170]. This reduces the number of update rules and simplifies F from Equation 5.7 to:

$$F = -\frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon} + \frac{1}{2}\ln|\tilde{\Pi}| - \frac{1}{2}s^2. \tag{5.9}$$

The last term in Equation 5.9 is the novel term that we have introduced for optimizing smoothness, and doesn't appear in FEP literature. Using this, we propose an online noise smoothness estimation algorithm which estimates s through the gradient ascend

---

[1]maximization of the ELBO term (Section 1.4)

(maximization[2]) of F, where $\frac{\partial F}{\partial s}|_{s=s_o} = 0$ and $\frac{\partial^2 F}{\partial s^2}|_{s=s_o} < 0$, with s_o being the smoothness value that maximizes F. The free energy gradients necessary for this scheme are obtained by differentiating Equation 5.9:

$$
\begin{aligned}
\frac{\partial F}{\partial s} &= -\frac{1}{2}\tilde{\epsilon}^T \frac{\partial \tilde{\Pi}}{\partial s}\tilde{\epsilon} + \frac{1}{2}\frac{\partial \ln|\tilde{\Pi}|}{\partial s} - s \\
\frac{\partial^2 F}{\partial s^2} &= -\frac{1}{2}\tilde{\epsilon}^T \frac{\partial^2 \tilde{\Pi}}{\partial s^2}\tilde{\epsilon} + \frac{1}{2}\frac{\partial^2 \ln|\tilde{\Pi}|}{\partial s^2} - 1,
\end{aligned}
\tag{5.10}
$$

where the gradients of $\ln|\tilde{\Pi}|$ can be computed as (refer Appendix 5.12.1) :

$$
\frac{\partial \ln|\tilde{\Pi}|}{\partial s} = 42(n+m)\frac{1}{s}, \quad \frac{\partial^2 \ln|\tilde{\Pi}|}{\partial s^2} = -42(n+m)\frac{1}{s^2}.
\tag{5.11}
$$

The usage of a gradient ascent scheme on the free energy curve for the estimation of s is motivated by the proof for the existence of a unique maximum for F under practical bounds as follows.

Proposition 5.6.1. *The free energy F defined by Equation 5.9 has a unique maximum with respect to noise smoothness s, under the practical range of noise smoothness $(0 < s < 1)$.*

Proof. Consider all the smoothness values of s with zero free energy gradients $(\frac{\partial F}{\partial s}|_{s=s_o} = 0)$. Substituting Equation 5.11 in 5.10 and using $\frac{\partial F}{\partial s}|_{s=s_o} = 0$ yields the condition satisfied by all maximum and minimum points:

$$
\frac{1}{2}(\tilde{\epsilon}^T \tilde{\Pi}_s \tilde{\epsilon})|_{s=s_o} = \frac{21(n+m)}{s_o} - s_o
\tag{5.12}
$$

where we use the shorthand $\tilde{\Pi}_s = \frac{\partial \tilde{\Pi}}{\partial s}$. Since $0 < s < 1$, we have from Equation 5.12 that:

$$
(\tilde{\epsilon}^T \tilde{\Pi}_s \tilde{\epsilon})|_{s=s_o} > 0.
\tag{5.13}
$$

The proof for the existence of a unique maximum is complete if we prove that $\frac{\partial^2 F}{\partial s^2}|_{s=s_o} < 0$, for all s_o satisfying Equation 5.13. The curvature of F at $s = s_o$ is calculated from Equation 5.10 using Equation 5.11 as:

$$
F_{ss}|_{s=s_o} = -\frac{1}{2}\left((\tilde{\epsilon}^T \tilde{\Pi}_{ss} \tilde{\epsilon})|_{s=s_o} + 42(n+m)\frac{1}{s_o^2} + 1 \right)
\tag{5.14}
$$

Since $\tilde{\Pi} > 0$, from definition $\tilde{\epsilon}^T \tilde{\Pi}\tilde{\epsilon} > 0$, and since $(\tilde{\epsilon}^T \tilde{\Pi}_s \tilde{\epsilon})|_{s=s_o} > 0$, we can conclude that $(\tilde{\epsilon}^T \tilde{\Pi}_{ss} \tilde{\epsilon})|_{s=s_o} > 0$, even though $\tilde{\Pi}_s \nsucc 0$ and $\tilde{\Pi}_{ss} \nsucc 0$ (refer Appendix 5.12.2 for numerical analysis). From Equation 5.14, $(\tilde{\epsilon}^T \tilde{\Pi}_{ss} \tilde{\epsilon})|_{s=s_o} > 0 \implies F_{ss}|_{s=s_o} < 0$, completing the proof for the existence of a unique maximum of free energy at $s = s_o$. □

---

[2]maximization of the ELBO term (Section 1.4)

5.7 Observer design

This section aims to introduce a novel observer design (DEMs) for the joint state and noise smoothness estimation of a linear system with colored noise. We formulate the noise smoothness estimator from the previous section (gradient ascend on F), using the Newton-Gauss update scheme:

$$
\begin{aligned}
s(t+dt) &= s(t) + ds, \\
ds &= (e^{F_{ss}|_{s=s(t)}dt} - 1)(F_{ss}|_{s=s(t)})^{-1}F_s|_{s=s(t)},
\end{aligned}
\tag{5.15}
$$

where $s(t)$ is the smoothness at time t, and ds is the smoothness increment for a time increment of dt. We combine this observer design with the standard DEM observer design for state estimation [46], where the update equation in the continuous time is given by:

$$
\dot{\tilde{x}} = A_1\tilde{x} + B_1\begin{bmatrix}\tilde{\mathbf{y}} \\ \tilde{\mathbf{v}}\end{bmatrix}
\tag{5.16}
$$

where $A_1 = [D^x - k^x\tilde{C}^T\tilde{\Pi}^z\tilde{C} - k^x(D^x - \tilde{A})^T\tilde{\Pi}^w(D^x - \tilde{A})]$, $B_1 = k^x\left[\tilde{C}^T\tilde{\Pi}^z \quad (D^x - \tilde{A})^T\tilde{\Pi}^w\tilde{B}\right]$, and k^x is the learning rate which is set to 1 throughout this paper. Since Equation 5.16 is a linear differential equation, an exact algebraic discretization can be performed for the observer as:

$$
\tilde{x}(t+dt) = e^{A_1 dt}\tilde{x}(t) + A_1^{-1}(e^{A_1 dt} - I)B_1\begin{bmatrix}\tilde{\mathbf{y}}(t) \\ \tilde{\mathbf{v}}(t)\end{bmatrix}
\tag{5.17}
$$

Equations 5.15 and 5.17 together complete our observer design. Note that A_1 and B_1 are nonlinear functions of s because of the presence of $\tilde{\Pi}^w$ and $\tilde{\Pi}^z$ in it. Moreover, the state updates enter Equation 5.15 through the F terms. Therefore, the update equations of state and noise smoothness observers are coupled. Since this heavily complicates the stability proof of the joint estimator, we leave it for future research.

5.8 Working example

This section aims to provide a working example in simulation to show the capabilities of our observer design. We use simulation data at different s levels to show that DEMs can accurately estimate $\tilde{x}$ and s.

5.8.1 Simulation settings

A random system with $A = \begin{bmatrix} 0.0484 & 0.7535 \\ -0.7617 & -0.2187 \end{bmatrix}$, $B = \begin{bmatrix} 0.3604 \\ 0.0776 \end{bmatrix}$, and $C = \begin{bmatrix} 0.2265 & -0.4786 \\ 0.4066 & -0.2641 \\ 0.3871 & 0.3817 \\ -0.1630 & -0.9290 \end{bmatrix}$ was used to generate the synthetic data for a total time of $T = 32s$ with increments $dt = 0.1s$, and a Gaussian bump input $v = e^{-0.25(t-12)^2}$. The colored noise was generated using $\Pi^w = e^6 I_2$ and $\Pi^z = e^6 I_4$. This simulation setting will be used throughout the paper, unless mentioned otherwise. We generate eight such time series data using different levels of noise smoothness s, ranging from 0.1 to 0.8 and use it for the analysis in this section.
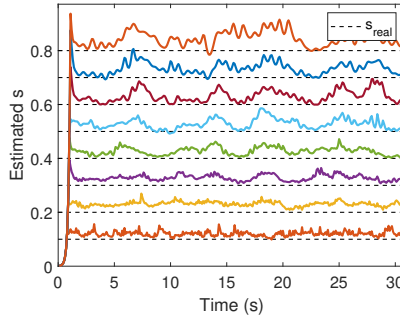
Fig 5.1: The maximization of F successfully estimates s for 8 simulations with different s_{real}. The colored solid lines represent the online estimation of s, whereas the dotted lines represent s_{real}. The estimation starts from $\eta^s = 0.001$ at time $t = 0$ for all simulations and converges close to s_{real} with a bias, within a few samples.
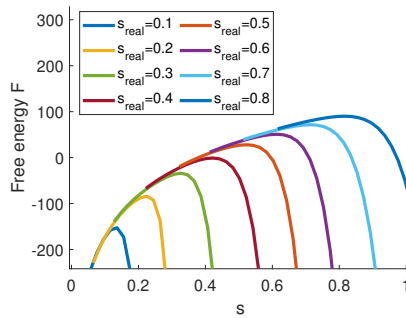
Fig 5.2: Free energy vs s plot for 8 different simulations (with 8 different real s values) from Figure 5.1 at $t = 5s$. All 8 free energy curves show a clear maximum around the real s values. This shows the effectiveness of our noise smoothness observer.

5.8.2 Test example

Figure 5.1 shows the results of our noise smoothness estimator for all eight simulations. All simulations start with the prior $\eta^s = 0.001$ and quickly stabilises around the correct smoothness value (s_{real} in dashed black), showing the success of our estimator for a range of noise smoothness values. Figure 5.2 shows the free energy vs s curve at $t = 5s$ for all eight simulations. The clear peaks of the free energy curve around the correct noise smoothness value (s_{real}) shows that free energy could be used as the objective function for noise estimation for the operational ranges of s. The importance of estimating the correct s is shown in Figure 5.3, where the minimum state estimation error is achieved when s_{real} is known. Therefore, Figure 5.3, 5.1 and 5.2, together demonstrates the validity of our observer design in simulation. In the next section, we will benchmark our observer against the state-of-the-art observers.
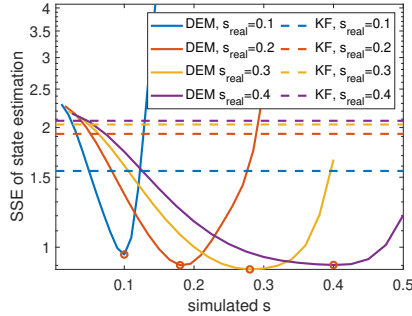
Fig 5.3: The sum of squared error (SSE) in state estimation of DEM deteriorates when the noise smoothness s used is different from the real noise smoothness s_{real}. The solid and dotted lines denote the SSE of DEM and KF for different simulated s. The SSE for DEM takes a minimum when $s \approx s_{real}$. For lower s (0.1 for example), KF outperforms DEM if s is not close to s_{real}, emphasizing the importance of an online noise smoothness observer.

5.9 Benchmarking

This section aims to benchmark the performance of our smoothness estimator for a state estimation problem. Through rigorous simulations, we show that our observer provides competitive performance during high colored noise cases.

5.9.1 Embedding order of states

In this section, we use rigorous simulations to show that our observer design can enable state estimation under a wide range of noises – at different embedding orders and smoothness levels. We manipulate on the dimension and component values of the S matrix in Equation 5.4 through different p and s values, under the same simulation setup described in Section 5.8.1 with $dt = 0.05s$. The size of S matrix increases with increasing p, whereas the components inside it increases with increasing s. Figure 5.4 shows the results of state estimation using 150 experiments (5 randomly generated noises each for five s values and six p values). The estimation error decreases with increasing p for different noise smoothness values, highlighting the importance of using higher order generalized coordination during estimation. This shows the applicability of our observer for a wide range of noise smoothness, embedding orders and noises.

5.9.2 Benchmark state observer

In this section, we benchmark our observer against other state-of-the-art observers like KF, SA and SMIKF, to show its competitiveness. 50 time series data (10 each for 5 smoothness values with $dt = 0.05s$) were generated using the simulation setup in Section 5.8.1 and the SSE in state estimation was computed for KF, SA, SMIFK and DEMs. The SMIKF and SA implementation accommodated an AR model of order 1 and 6 respectively for the noise modelling, whereas the DEM implementation used an embedding order of $p = 6$ for states and $d = 2$ for inputs. Figure 5.5 shows the results, clearly indicating the superior performance of DEMs with minimum error in state estimation for higher s. DEMs
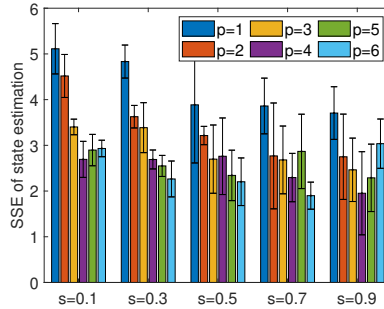
Fig 5.4: The error in state estimation decreases as the embedding order of states p increases, for a range of noise smoothness s. This shows that our smoothness estimation aids an accurate state estimation till an embedding order of $p = 4$ for a wide range of s.

outperforms other observers for a wide range of s values. However, for low noise color ($s = 0.1$), SA and SMIKF outperforms DEMs. In all cases, DEMs outperforms KF in the presence of colored noise.
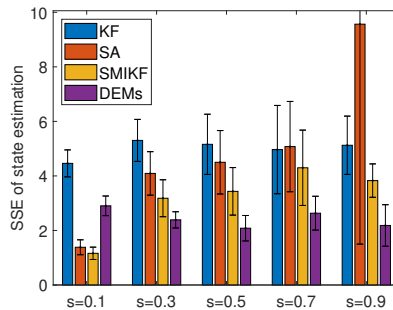
Fig 5.5: DEMs outperform KF, SA and SMIKF with minimal estimation error during state estimation under high colored noise ($s > 0.1$). For low colored noise ($s = 0.1$), DEMs outperforms KF, but SA and SMIKF performs better. Solutions of SA is unstable for higher s.

5.10 Proof of concept - quadrotor flight

This section aims to provide a proof of concept for our observer design by employing it for the state estimation of a real quadrotor flying under wind conditions. We use the experimental design from [165] to obtain the quadrotor flight data. The experiment consist of a quadrotor hovering at a fixed location, under the strong influence of wind generated by a blower. The linearized quadrotor model relating the input motor signals to the output roll angle (ϕ) of the quadcopter, without accounting for the wind dynamics is given by [165]:

$$\begin{bmatrix} \dot{\phi} \\ \ddot{\phi} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{c_{B\phi}}{I_{xx}} & -\frac{c_{B\phi}}{I_{xx}} & -\frac{c_{B\phi}}{I_{xx}} & \frac{c_{B\phi}}{I_{xx}} \end{bmatrix} \begin{bmatrix} pwm_1 \\ pwm_2 \\ pwm_3 \\ pwm_4 \end{bmatrix},$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix},$$

(5.18)

where pwm_i is the Pulse Width Modulation signal provided to the ith motor by the controller for stable hovering, $I_{xx} = 3.4 \cdot 10^{-3} kg\,m^2$ is the quadcopter's moment of inertia around the x-axis, and $c_{B\phi} = 1.274 \cdot 10^{-3} Nm$ is the thrust coefficient that models the relation between the PWM values and the thrust generated by the quadcopter rotors. ϕ was recorded using the Optitrack system, and was used for the state estimation for a time sequence of $T = 15s$ with $dt = 0.0083s$. The influence of wind dynamics on the quadrotor states (ϕ and $\dot{\phi}$) is unmodelled in Equation 5.18. Therefore, the wind dynamics induces strong colored noise (w) in the data [166]. The higher process noise ($\Pi^w = e^4$), and a lower measurement noise ($\Pi^z = e^{10}$) were used to represent high unmodelled wind noise and low Optitrack noise respectively. The embedding order of $p = 2$ and $d = 2$ were used to capture the noise color.
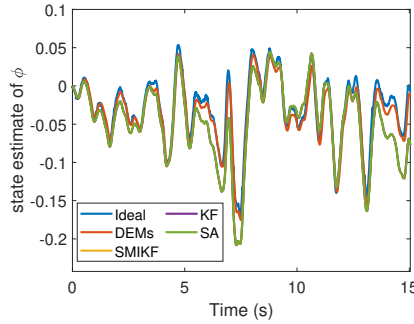
Fig 5.6: DEMs outperforms other benchmarks in state estimation on the quadrotor flight data where it hovers under the influence of wind, introducing colored noise into the system. DEMs (in red) is closer to the ground truth Optitrack measurement (ideal in blue), when compared to other benchmarks. KF, SA and SMIKF shows coinciding estimation plots.

Figure 5.6 shows the superior state estimation capabilities of DEMs. DEMs (in red) is closer to the ground truth (in blue) when compared to other benchmarks. KF, SMIKF and SA have coinciding state estimation curves. Figure 5.7 shows the free energy vs s curve at different time instances t of the quadrotor flight, showing a clear maximum, similar to the simulation results in Figure 5.2, validating the practical application of our estimator.
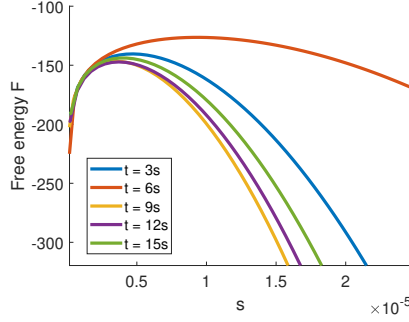
Fig 5.7: The free energy vs *s* plot for different time instances during the quadrotor flight (data from Figure 5.6). The curves show a clear maximum, similar to the simulation results from Figure 5.2. This provides an experimental validation for using the gradient ascend over free energy for smoothness estimation.

5.11 Conclusion

A novel observer (DEMs) for the joint state and noise smoothness estimation of linear systems with colored noise was introduced. Through rigorous simulations, DEMs was shown to outperform the benchmarks like KF, SMIKF and SA in state estimation under colored noise with minimum estimation error. The observer was face validated by applying it on a practical robotics application - the state estimation of a quadrotor hovering in unmodelled wind conditions, to show that DEMs is a competitive observer. The main limitation of this work is the absence of a stability proof for the joint state and noise smoothness observer, which can be the focus of future research. The estimator can be extended for a confident smoothness estimation with precision updates for Π^s. It can also be extended to solve the general active inference problem for the estimation and control of nonlinear systems with colored noise.

Acknowledgment

5.12 Appendix

5.12.1 Gradients of $\ln|\tilde{\Pi}|$

The log determinant of generalized precision can be calculated using Equations 5.4 and 5.5 as:

$$
\begin{aligned}
\ln|\tilde{\Pi}| &= \ln|S \otimes \Pi^z| + \ln|S \otimes \Pi^w| \\
&= \ln(|S|^m|\Pi^z|^{(p+1)}) + \ln(|S|^n|\Pi^w|^{(p+1)}) \\
&= (p+1)(\ln|\Pi^z| + \ln|\Pi^w|) + (n+m)\ln|S|.
\end{aligned}
\tag{5.19}
$$

The derivative of $\ln|\tilde{\Pi}|$ with respect to s becomes:

$$\frac{\partial \ln|\tilde{\Pi}|}{\partial s} = (n+m)\frac{\partial \ln|S|}{\partial s}, \quad \frac{\partial^2 \ln|\tilde{\Pi}|}{\partial s^2} = (n+m)\frac{\partial^2 \ln|S|}{\partial s^2}. \tag{5.20}$$

From Equation 5.4, $|S| = \frac{512}{6075}s^{42}$, resulting in $\frac{\partial \ln|S|}{\partial s} = \frac{42}{s}$ and $\frac{\partial^2 \ln|S|}{\partial s^2} = -\frac{42}{s^2}$. This simplifies Equation 5.20 to:

$$\frac{\partial \ln|\tilde{\Pi}|}{\partial s} = 42(n+m)\frac{1}{s}, \quad \frac{\partial^2 \ln|\tilde{\Pi}|}{\partial s^2} = -42(n+m)\frac{1}{s^2}. \tag{5.21}$$
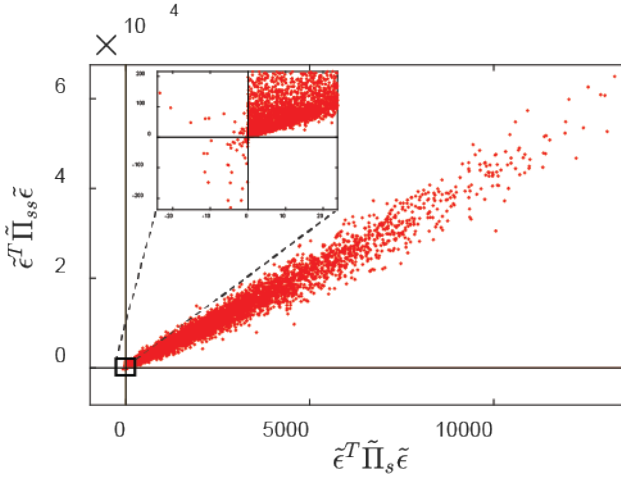
Fig 5.8: The plot demonstrating that the function $\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon}$ is mostly monotonically increasing with respect to s in the domain $(0,1]$ for $|\tilde{\epsilon}| < 1$. Moreover, when $\tilde{\epsilon}^T\tilde{\Pi}_s\tilde{\epsilon} > 0$, $\tilde{\epsilon}^T\tilde{\Pi}_{ss}\tilde{\epsilon} < 0$, since there are no data points on the fourth quadrant as shown in the zoomed plot.

5.12.2 Numerical analysis on the nature of $\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon}$

We recorded the first two gradients of the polynomial $\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon}$ with respect to s for 20,000 combinations of randomly sampled $\tilde{\epsilon}$ and s such that $|\tilde{\epsilon}| < 1$ and $s \in (0,1]$. From the results shown in Figure 5.8, the data points predominantly lie on the first quadrant, suggesting that the function has positive gradients, which is a sign of monotonically increasing function. The absence of any points on the fourth quadrant motivates the conclusion: if $\tilde{\epsilon}^T\tilde{\Pi}_s\tilde{\epsilon} > 0$ then $\tilde{\epsilon}^T\tilde{\Pi}_{ss}\tilde{\epsilon} > 0$. The results remain the same for different norm lengths of $\tilde{\epsilon}$.

6

Dynamic Expectation Maximization

*This chapter uses **precision learning** for the system identification of an LTI system under colored noise. It provides the detailed mathematical foundation of the system identification tool used in Sections 2.5.2.3, 2.5.2.4 and 2.5.2.5. The Chapter 7 provides the convergence proof of the parameter estimator introduced in this chapter, followed by the Chapter 8 that provides the experimental confirmation of the estimator on real robot data from Chapter 4.*

6.1 Abstract

The free energy principle from neuroscience has recently gained traction as one of the most prominent brain theories that can emulate the brain's perception and action in a bio-inspired manner. This renders the theory with the potential to hold the key for general artificial intelligence. Leveraging this potential, this paper aims to bridge the gap between neuroscience and robotics by reformulating an FEP-based inference scheme—Dynamic Expectation Maximization—into an algorithm that can perform simultaneous state, input, parameter, and noise hyperparameter estimation of any stable linear state space system subjected to colored noises. The resulting estimator was proved to be of the form of an augmented coupled linear estimator. Using this mathematical formulation, we proved that the estimation steps have theoretical guarantees of convergence. The algorithm was rigorously tested in simulation on a wide variety of linear systems with colored noises. The paper concludes by demonstrating the superior performance of DEM for parameter estimation under colored noise in simulation, when compared to the state-of-the-art estimators like Sub Space method, Prediction Error Minimization (PEM), and Expectation Maximization (EM) algorithm. These results contribute to the applicability of DEM as a robust learning algorithm for safe robotic applications.

6.2 Introduction

The free energy principle (FEP) has emerged from neuroscience as a unifying theory of the brain [2] and has begun to guide the search for a brain-inspired learning algorithm for robots. Many attempts have been made in this direction, including the state and input observer [46, 165], the adaptive controller for robot manipulators [27, 212, 213], the body perception and action scheme for humanoid robots [200], the robot navigation of ground robots [34] etc. However, the design of a parameter estimation algorithm for linear systems with colored noise remains unexplored. Since the design of an accurate parameter estimator for dynamic systems sits at the core of control systems and robotics, the reformulation of FEP into a brain-inspired estimation algorithm has an influential impact on the industry and applied robotics.

A wide range of estimators have been proposed in the literature for linear time-invariant (LTI) systems [214], including the blind system identification [215–217]. However, most of them assume the noises to be temporally uncorrelated (white noise), which is often violated in practice. This results in biased estimation for the least-square (LS)-based methods [218], and an inaccurate convergence for the iterative methods [219]. Although many attempts have been made to solve this problem, mainly through bias compensation methods [220, 221], none of them perform state, input, parameter, and noise estimation for systems with colored noises [222]. The only method that does it is the Dynamic Expectation Maximization (DEM) [170] algorithm, which uses FEP to invert a highly nonlinear and hierarchical brain model from sensory data. However, the disconnect between neuroscience and control system literature hinders the wide use of this method for practical robotics applications. Although FEP-based tools have already been applied to practical robotics applications [27, 34, 46, 165, 166, 200, 212, 213, 223], there is a gap in the literature on the applications of DEM owing to the mathematical formidability of the theory and its lack of formalism in the control systems domain. Therefore, it is important to reformulate DEM

for the control systems audience. While DEM from computational neuroscience focuses on emulating the brain's perception through the hierarchical abstraction of a number of interacting non-linear dynamic systems, our work focuses on reformulating it into a blind system-identification algorithm for an LTI system with colored noise, which is a well-known challenge in robotics. In this attempt, we keep all the brain-related approximations intact, thereby aiming for a biologically plausible parameter estimation algorithm.

According to an FEP proposed by Karl Friston, the brain's inference mechanism is a gradient ascent over its free energy, where free energy is the information-theoretic measure that bounds the brain's sensory surprisal [43]. FEP emerges as a unified brain theory [186] by providing a mathematical description of brain functions [187]; unifying action and perception [50]; connecting physiological constructs like memory, attention, value, reinforcement, and salience [187]; and remaining consistent with Freudian ideas[188]. Similarities of FEP with reinforcement learning [42], neural networks [43, 44], PID controller [47], Kalman Filter [46] and active learning [50] open up possibilities for biologically plausible parameter estimation algorithms. Although FEP emerged as a brain theory, the recent works have pushed the boundaries towards systems that survive over time, such as social and cultural dynamics. Notable works include the variational approach to culture [224], collective intelligence [225], cumulative cultural evolution [226], etc.

Numerous methods have been proposed based on the FEP framework. Predictive coding [8] models perception through a hierarchy of dynamical systems [44] with the brain's priors at the top, minimizing the prediction error at each level of the hierarchy. Bayesian message passing algorithms [227, 228] use similar ideas for belief propagation. Active inference [50, 229] uses FEP to model the brain's action and perception under one framework. On the perception side, there are two main type of methods to deal with dynamic systems: variational filtering and generalized filtering. Variational filtering [170, 230] uses mean-field approximation (conditional independence between densities), whereas generalized filtering [169] does not. DEM [170] is a type of variational filtering that uses a Laplace approximation [205] (a fixed-form assumption for the conditional density of variables), whereas [230] uses ensemble dynamics to model the free form of the conditional densities. We focus on DEM for this work.

DEM is an FEP-based variational inference algorithm that models the brain's inference process as a maximization of its free energy[1] for state, input, parameter, and noise estimation from data. Although the method shows high similarity to the variational inference [52], the key difference is in the use of generalized coordinates, which enables DEM to track the evolution of the trajectory of states instead of just the point estimates. This renders DEM with the capability of gracefully handling colored noises, a feature that conventional point estimators such as the Kalman Filter (KF) lacks. The modeling of noise color as analytic using generalized coordinates results in an improved state estimation under colored noise for LTI systems [46, 165] and for nonlinear filtering [231], which directly improves the parameter estimation accuracy, making DEM a topic of interest [166]. This work directly impacts various sub-domains of robotics community: input estimation to the industrial automation community for fault detection systems, state estimation to the control systems community, parameter estimation to the system identification community, and hyperparameter estimation to the signal processing community.

---

[1]maximization of the ELBO term (Section 1.4)

With this paper, we aim to present DEM to the robotics audience as a robot learning algorithm for the blind system identification of LTI systems with colored noise. We elaborate on various components of DEM that are most relevant to the robotics community: 1) the derivation of the free energy objectives from Bayesian principles, 2) the modeling of colored noise using generalized coordinates, and 3) the simplification of update rules for LTI systems with colored noise. We reformulate DEM for LTI systems into a form that is widely used in the robotics domain and use this mathematical formulations to prove that the estimation steps of DEM have theoretical guarantees of convergence [167]. In our prior work, we have discussed the stability conditions of our DEM-based linear state and input observer design [46]. This convergence guarantees and stability criteria are essential for robot safety while in operation and is of high relevance to the robotics community. Through extensive simulations on a range of random systems, we show that DEM is a competitive estimator when compared to other benchmarks in the control systems domain. The core contributions of the paper are:

1. Reformulating DEM into an estimation algorithm for LTI systems with colored noise (Section 6.13).

2. Proving that the estimator has theoretical guarantees of convergence for the estimation steps (Section 6.15).

3. Proving through rigorous simulation that DEM outperforms the state-of-the-art system identification methods for parameter estimation under colored noise (Section 6.17).

6.3 Problem Statement

Consider the linear plant dynamics given in Equation (6.1), where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are constant system matrices, $\mathbf{x} \in \mathbb{R}^n$ is the hidden state, $\mathbf{v} \in \mathbb{R}^r$ is the input and $\mathbf{y} \in \mathbb{R}^m$ is the output.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} + \mathbf{w}, \qquad \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{z}. \qquad (6.1)$$

Here, $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$ represent the process and measurement noise, respectively. The notations of the plant are denoted in boldface, whereas its estimates are denoted in nonboldface letters. The noises in this paper are generated through the convolution of white noise with a Gaussian kernel. The use of colored noise is motivated by the fact that in robotics, the unmodelled dynamics and the non-linearity errors can enter the plant dynamics through the noise terms, thereby violating the white noise assumption in practice [166].

The goal of this paper is to reformulate DEM for an LTI system such that given the output of the system $\mathbf{y}$, the estimator computes the associated states x, inputs v, parameter θ containing the matrices A, B and C, hyperparameters λ that model the noise precision ($\Pi = e^\lambda$), and the uncertainties of all its estimates ($\Sigma^x, \Sigma^v, \Sigma^\theta, \Sigma^\lambda$), with the help of the prior (learned) knowledge encoded in the robot brain, such that the estimate best predicts the data. The schematic of the proposed robot brain's inference process is given in Figure 6.1.
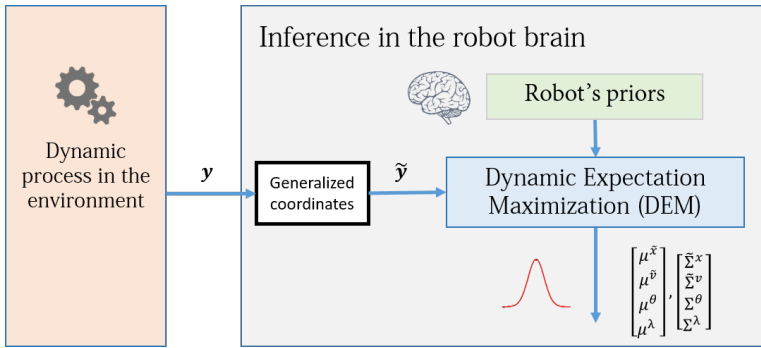
Fig 6.1: A simple block diagram of the robot brain's inference process using DEM. It uses the measurement data **y** generated from the environment (also called generative process). DEM enables the direct fusion of the prior information into the inference process. The concept of generalized coordinates will be detailed in Section 6.4.1.

6.4 Preliminaries

To reformulate DEM into an estimation algorithm for LTI systems, this section introduces the key concepts and terminologies that are familiar to the FEP audience.

6.4.1 Generative Model

The generative model (plant model) is the robot brain's estimate of the generative process in the environment that generated data. The robot brain infers this model via model evidencing from the measurement data. The key idea behind DEM to deal with colored noise is to model the trajectory of the time-varying components (states, for example) using generalized coordinates. The use of generalized coordinates is new to the control systems literature and is different from the familiar definition in classical mechanics. In mechanics, it is the minimum number of independent coordinates that define the configuration of a system, whereas in DEM, it is the vector defining the motion of a point using its higher derivatives. In DEM, the emphasis is on tracking the trajectories of states, inputs and outputs instead of their point estimates. The states are expressed in generalized coordinates using its higher-order derivatives, i.e., $\tilde{x} = [x \; x' \; x'' \; ...]^T$. The generalized state vector $\tilde{x}$ with an order of generalized motion of p will have $p+1$ terms, with the copy of the state vector as the first term, followed by its p derivatives. The variables in generalized coordinates are denoted by a tilde, and its components (higher derivatives) are denoted by primes. The evolution of states is written as:

$$\begin{aligned} x' &= Ax + Bv + w & y &= Cx + z \\ x'' &= Ax' + Bv' + w' & \dot{y} &= Cx' + z' \\ &... & &... \end{aligned} \qquad (6.2)$$

The colored noises are modeled such that the covariance of noise derivatives $\tilde{z} = [z, z', z'', ...]^T$ and $\tilde{w} = [w, w', w'', ...]^T$ are well defined (to be explained in Section 6.4.3). The generative

model representing the system is compactly written as:

$$\dot{\tilde{x}} = D^x \tilde{x} = \tilde{A}\tilde{x} + \tilde{B}\tilde{v} + \tilde{w} \qquad \tilde{y} = \tilde{C}\tilde{x} + \tilde{z} \tag{6.3}$$

where $D^x = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & 0 & 1 \\ & & & & 0 \end{bmatrix}_{(p+1)\times(p+1)} \otimes I_{n\times n}$, performs the block derivative operation, equivalent

to shifting up all components in generalized coordinates by one block. A similar definition holds for D^v (appears later) with size $r(d+1) \times r(d+1)$, where p and d are the order of generalized motion of states and inputs, respectively. Here, $\tilde{A} = I_{p+1} \otimes A$, $\tilde{B} = I_{p+1} \otimes B$, and $\tilde{C} = I_{p+1} \otimes C$, where $\otimes$ is the Kronecker tensor product.

6.4.2 Parameters and Hyperparameters
To simplify the parameter estimation steps of DEM for LTI systems (Section 6.14.2) and to facilitate the convergence proof (Section 6.15), we introduce an alternative generative model which is the direct reformulation of Equation (6.3) given by:

$$\dot{\tilde{x}} = M\theta + \tilde{w}, \quad \tilde{y} = N\theta + \tilde{z}, \quad \theta = \begin{bmatrix} vec(A^T) \\ vec(B^T) \\ vec(C^T) \end{bmatrix} \tag{6.4}$$

where

$$M = \begin{bmatrix} I_n \otimes x^T & I_n \otimes v^T & I_n \otimes O_{1\times m} \\ I_n \otimes x'^T & I_n \otimes v'^T & I_n \otimes O_{1\times m} \\ \dots & \dots & \dots \end{bmatrix}, N = \begin{bmatrix} I_n \otimes O_{1\times n} & I_n \otimes O_{1\times r} & I_m \otimes x^T \\ I_n \otimes O_{1\times n} & I_n \otimes O_{1\times r} & I_m \otimes x'^T \\ \dots & \dots & \dots \end{bmatrix}. \tag{6.5}$$

Throughout the paper, θ represents the set of parameters, whereas $\lambda = \begin{bmatrix} \lambda^z \\ \lambda^w \end{bmatrix}$ represents the hyperparameters that define the precision matrices (inverse covariance matrices) of the observation noise and the process noise. For noise modelling, we parametrize the noise precisions using an exponential relation with the hyperparameters, given by [170]:

$$\Pi^w(\lambda^w) = e^{\lambda^w}\Omega^w, \quad \Pi^z(\lambda^z) = e^{\lambda^z}\Omega^z, \tag{6.6}$$

where Ω^w and Ω^z represent constant matrices encoding how different noises are correlated. Here Π^w and Π^z are the inverse covariances $\{(\Sigma^w)^{-1}, (\Sigma^z)^{-1}\}$ or precisions of the noises. This parameterization ensures the selection of positive definite noise precision matrices through hyperparameter updates. We assume zero cross-correlation between w and z. We also assume that θ and λ are time-invariant.

6.4.3 Colored noise
The next step towards handling the colored noise is to embed the noise correlation between different components in the generalized noises $\tilde{w}$ and $\tilde{z}$ given in Equation (6.3). DEM uses generalized coordinates, which models a correlation between noise derivatives through the temporal precision matrix S (inverse of covariance matrix). The generalized noise

correlation is assumed to be due to a Gaussian filter, where S can be calculated as [170]:

$$
S(\sigma^2) = \begin{bmatrix} 1 & 0 & -\frac{1}{2\sigma^2} & .. \\ 0 & \frac{1}{2\sigma^2} & 0 & .. \\ -\frac{1}{2\sigma^2} & 0 & \frac{3}{4\sigma^4} & .. \\ .. & .. & .. & .. \end{bmatrix}^{-1}_{(p+1)\times(p+1)}
\tag{6.7}
$$

where σ (from the Gaussian kernel) denotes the noise smoothness level. While $\sigma^2 \to 0$ denotes white noise, nonzero σ^2 denotes colored noise. The covariance between noise derivatives increases exponentially with the order of noise derivatives. Simulations show that derivatives above 6 can be neglected [170]. The generalized noise precision matrices are given by $\tilde{\Pi}^w = S(\sigma^2) \otimes \Pi^w$ and $\tilde{\Pi}^z = S(\sigma^2) \otimes \Pi^z$, where Π^w and Π^z are the inverse noise covariances.

6.4.4 Generalized Motion of the Outputs And Noises

The generalized motion of output $\tilde{y}$ is practically expensive and inaccessible to robotics systems. Moreover, most sensors such as encoders operate with discrete measurements. Therefore, as a pre-processing step for estimation (refer to Figure 6.1), $\tilde{y}$ should be computed from discrete measurements. The goal is to first express the discrete measurements as a function of output derivatives and then invert the function to compute the derivatives from the discrete measurements. Given the p temporal derivatives $\tilde{y}$ at time t, the p output sequence surrounding y can be approximated using Taylor series as follows [170]:

$$
\hat{y} = \begin{bmatrix} .. \\ y(t-dt) \\ y(t) \\ y(t+dt) \\ .. \end{bmatrix} = (Y \otimes I_m)\tilde{y}, \ Y_{ij} = \frac{\left[\left(i - ceil\left(\frac{p+1}{2}\right)\right)dt\right]^{j-1}}{(j-1)!},
\tag{6.8}
$$

where $i, j = 1, 2, \ldots p+1$, $ceil(.)$ is the smallest integer function and $\hat{y}$ has the size $m(p+1) \times 1$. Therefore, generalized motion of output $\tilde{y}$ at time t is:

$$
\tilde{y} = (Y^{-1} \otimes I_m)\hat{y}.
\tag{6.9}
$$

Using $\tilde{y}$ embeds more temporal information about the plant into the data in the form of conditional trajectories, with the disadvantage of a time latency of $\frac{p}{2}dt$ in estimation. For robotic systems with high sampling rates, this latency in estimation is negligible [165, 166]. The next section employs this generalized output along with the generative model for observer designs.

6.4.5 Notations and conventions

Throughout the paper, the superscript notation will be used to represent the quantity being referred to, and the subscript notation will be used to represent the derivative operation. For example, $\tilde{\Pi}^w_\lambda$ represents the derivative of the generalized precision of the process noise w with respect to the hyperparameters λ. The tilde operator ($\tilde{x}$) is used to represent a quantity in its generalized coordinates, whereas the bar operator ($\bar{F}$) is used to represent the time integral of a quantity. All the probability distributions will be represented by the $p(.)$ notation, whereas its expectation will be represented by the $\langle p(.) \rangle$ notation.

6.5 Free Energy Objectives

With the preliminaries in place, we can build up towards the complete DEM algorithm. Eventually, in Sections 6.9 and 6.13, we will see that it is an optimization algorithm that finds the best estimates for states, inputs, parameters, and hyperparameters for given measurement data. This result is achieved by optimizing two objective functions, which are the core objectives under the entire Free Energy Principle: the free energy and the free action [2]. Here we derive and simplify this objective.

The derivation starts from the fundamentals of Variational Inference (VI) [52]. In VI, the posterior distribution $p(\vartheta/y)$ of parameter ϑ, given the measurement y, is expressed as:

$$p(\vartheta/y) = \frac{p(y/\vartheta)p(\vartheta)}{p(y)} = \frac{p(\vartheta, y)}{p(y)} = \frac{p(\vartheta, y)}{\int p(\vartheta, y)d\vartheta}. \tag{6.10}$$

However, the marginalization over ϑ to calculate $p(y)$ is often intractable because the search space of ϑ is large. A widely used technique is to introduce a variational distribution $q(\vartheta)$ known as the recognition density, which acts as an approximate representation of the posterior distribution with $q(\vartheta) \approx p(\vartheta/y)$. A common method used among variational Bayes algorithms is to minimize the Kullback–Leibler (KL) divergence between both the distributions, defined as:

$$KL(q(\vartheta)\|p(\vartheta/y)) = \langle \ln q(\vartheta)\rangle_{q(\vartheta)} - \langle \ln p(\vartheta/y)\rangle_{q(\vartheta)}, \tag{6.11}$$

where $\langle.\rangle_{q(\vartheta)}$ represents the expectation over $q(\vartheta)$. Substituting Equation (6.10) in Equation (6.11) and using $\int q(\vartheta)d\vartheta = 1$ yields:

$$KL(q(\vartheta)\|p(\vartheta/y)) = \langle \ln q(\vartheta)\rangle_{q(\vartheta)} - \langle \ln p(\vartheta, y)\rangle_{q(\vartheta)} + \ln p(y). \tag{6.12}$$

The rearrangement of terms yield:

$$\ln p(y) = \langle \ln p(\vartheta, y)\rangle_{q(\vartheta)} - \langle \ln q(\vartheta)\rangle_{q(\vartheta)} + KL(q(\vartheta)\|p(\vartheta/y)),$$
$$= \underbrace{\langle U(\vartheta, y)\rangle_{q(\vartheta)} + H(\vartheta)_{q(\vartheta)}}_{\text{free energy}} + KL(q(\vartheta)\|p(\vartheta/y)), \tag{6.13}$$

where $\ln p(y)$ is the log-evidence, $U(\vartheta, y) = \ln p(\vartheta, y)$ is the internal energy and $H(\vartheta)_{q(\vartheta)} = -\langle \ln q(\vartheta)\rangle_{q(\vartheta)}$ is the entropy of the density. The free energy term in Equation (6.13) is defined as the sum of an energy term and its entropy. It acts as the lower bound on the log-evidence because the KL divergence term $KL(q(\vartheta)\|p(\vartheta/y))$ is always positive. The maximization of free energy minimizes the divergence term in Equation (6.13) because the log-evidence is independent of $q(\vartheta)$, thereby rendering the variational density $q(\vartheta)$ as a close approximation of $p(\vartheta/y)$. Therefore, the difficult evaluation of an intractable integral term in Equation (6.10) is converted into a much simpler optimization problem of maximizing the free energy[2]. This reduces the problem of inference to a direct optimization of its free energy objectives and is the fundamental idea behind variation inference. The free energy term in Equation (6.13) is equivalent to the Evidence Lower Bound (ELBO) [52].

---

[2]maximization of the ELBO term (Section 1.4)

It can be simplified as:

$$F = \langle U(\vartheta, y)\rangle_{q(\vartheta)} + H(\vartheta)_{q(\vartheta)} = \int q(\vartheta)U(\vartheta, y)d\vartheta - \int q(\vartheta)\ln q(\vartheta)d\vartheta$$
$$= \int q(\vartheta)\ln p(\vartheta, y)d\vartheta - \int q(\vartheta)\ln q(\vartheta)d\vartheta. \tag{6.14}$$

However, when the parameter set to be estimated includes both the time-variant and the time-invariant parameters, the free action is used as the objective function to be maximized. The free action is defined as the time integral of the free energy and is given by:

$$\bar{F} = \bar{V}(\vartheta) + \bar{H}(\vartheta) = \int \langle U(\vartheta, y)\rangle_{q(\vartheta)}dt + \int H(\vartheta)_{q(\vartheta)}dt, \tag{6.15}$$

where $V(\vartheta) = \langle U(\vartheta, y)\rangle_{q(\vartheta)}$ is called the variational free energy (VFE).

The next sections will deal with two main assumptions used in DEM to simplify the free energy objectives, namely the Laplace approximation and the mean-field assumption. The aim is to derive the simplified free energy objectives for an LTI system under these assumptions.

6.6 LAPLACE APPROXIMATION

The first common approach to simplify the free energy objective is to assume the variational density $q(\vartheta)$ to be Gaussian in nature with variational parameters ϑ and Σ^θ as its mode and covariance, respectively, [205]. Here, the inverse of Σ^ϑ (denoted by Π^ϑ and known as the conditional precision), represents the confidence in estimation. The recognition density takes the following form:

$$q(\vartheta) = \mathcal{N}(\vartheta : \mu^\vartheta, \Sigma^\vartheta) = \frac{1}{\sqrt{(2\pi)^n|\Sigma^\vartheta|}}e^{-\frac{1}{2}(\vartheta-\mu^\vartheta)^T\Pi^\vartheta(\vartheta-\mu^\vartheta)}. \tag{6.16}$$

There are two main advantages with this approximation:

1. It simplifies the internal energy expression $U(\vartheta, y)$,

2. It facilitates an easy computation of the conditional precision Π^ϑ (derived in Section 6.8.2) as the negative curvature of the internal energy at it's mode μ^ϑ.

Therefore, the main aim of this section is to simplify the expression for internal energy $U(\vartheta, y)$ using the Laplace approximation, for an LTI system with its states, inputs, and outputs expressed in generalized coordinates.

The internal energy in Equation (6.13) can be expressed as the sum of log-likelihood and prior terms as:

$$U(\vartheta, y) = \ln p(\vartheta, y) = \ln \underbrace{p(y/\vartheta)}_{\substack{\text{generative} \\ \text{model}}} + \ln \underbrace{p(\vartheta)}_{\text{prior}}. \tag{6.17}$$

The parameter set ϑ includes two types of parameters:

1. the states and inputs, which are time-varying and therefore expressed in generalized coordinates,

2. the parameters and hyperparameters, which are time-invariant and not expressed in generalized coordinates.

Equation (6.17) can be simplified by assuming the conditional independence of $\tilde{x}$ and $\tilde{v}$ with θ and λ. This factorization separates the deterministic quantities from the stochastic ones, thereby providing a separation of temporal scales. This is one of the core ideas behind DEM, which will be detailed in Section 6.7. With the redefinition of $\vartheta = \{\tilde{x}, \tilde{v}, \theta, \lambda\}$, Equation (6.17) simplifies to:

$$
\begin{aligned}
U(\vartheta, y) &= \ln p(y/\tilde{x}, \tilde{v}, \theta, \lambda) + \ln p(\tilde{x}, \tilde{v}, \theta, \lambda) \\
&= \ln p(y/\tilde{x}, \tilde{v}, \theta, \lambda) + \ln p(\tilde{x}/\tilde{v}, \theta, \lambda) + \ln p(\tilde{v}) + \ln p(\theta) + \ln p(\lambda).
\end{aligned}
\tag{6.18}
$$

DEM combines the new sensory information y coming from the environment with the robot brain's priors (refer to Figure 6.1) in a Bayesian fashion, through the internal energy $U(\vartheta, y)$ expression given in Equation (6.18). The next sections will deal with simplifying $U(\vartheta, y)$ and its action $\bar{U}$ by first modeling the probability distributions for the generative model and the priors.

6.6.1 Generative Model

The probability distribution $p(y/\tilde{x}, \tilde{v}, \theta, \lambda)$, given in Equation (6.18), represents the generative model that predicts the output from the current parameter estimates. This probability can be assumed to be Gaussian-distributed, centered around the model's output prediction $\tilde{C}\tilde{x}$ (from Equation (6.3)), with the same uncertainty as that of the measurement noise $\tilde{\Sigma}^z$. The distribution $p(y/\vartheta)$ becomes:

$$
p(\tilde{y}/\vartheta) = \frac{1}{\sqrt{(2\pi)^{m(p+1)}|\tilde{\Sigma}^z|}} e^{-\frac{1}{2}(\tilde{y}-\tilde{C}\tilde{x})^T \tilde{\Pi}^z (\tilde{y}-\tilde{C}\tilde{x})}.
\tag{6.19}
$$

Since the robot cannot directly measure $\tilde{x}$ in Equation (6.3), we track the motion of the generalized states through the approximation $\dot{\tilde{x}} = \tilde{x}' = D^x \tilde{x}$. The prediction for motion is $\tilde{A}\tilde{x} + \tilde{B}\tilde{v}$ with an uncertainty of $\tilde{\Sigma}^w$. The Gaussian distribution becomes:

$$
p(\tilde{x}'/\tilde{x}, \tilde{v}, \vartheta, \lambda) = \frac{1}{\sqrt{(2\pi)^{n(p+1)}|\tilde{\Sigma}^w|}} e^{-\frac{1}{2}(D^x \tilde{x}-\tilde{A}\tilde{x}-\tilde{B}\tilde{v})^T \tilde{\Pi}^w (D^x \tilde{x}-\tilde{A}\tilde{x}-\tilde{B}\tilde{v})}.
\tag{6.20}
$$

6.6.2 Prior Distributions

The remaining distributions $p(\tilde{v})$, $p(\theta)$ and $p(\lambda)$ are the priors of the robot brain that can be transferred from prior (learned) experiences to the inference process. Similar to the previous section, a Gaussian prior is placed over the inputs $p(\tilde{v}) = \mathcal{N}(\tilde{v} : \eta^{\tilde{v}}, L^{\tilde{v}})$ as well, with mean $\eta^{\tilde{v}}$ and prior covariance $L^{\tilde{v}} = (P^{\tilde{v}})^{-1}$, as:

$$
p(\tilde{v}) = \frac{1}{\sqrt{(2\pi)^{r(d+1)}|L^{\tilde{v}}|}} e^{-\frac{1}{2}(\tilde{v}-\eta^{\tilde{v}})^T P^{\tilde{v}} (\tilde{v}-\eta^{\tilde{v}})}.
\tag{6.21}
$$

The prior distribution of parameters $\theta \in \mathbb{R}^l$ is assumed to be a Gaussian-centred around the prior parameter value η^θ, with the prior covariance $L^\theta = (P^\theta)^{-1}$:

$$
p(\theta) = \mathcal{N}(\theta : \eta^\theta, L^\theta) = \frac{1}{\sqrt{(2\pi)^l |L^\theta|}} e^{-\frac{1}{2}(\theta-\eta^\theta)^T P^\theta (\theta-\eta^\theta)}.
\tag{6.22}
$$

Similarly, a Gaussian prior is placed over the hyperparameters $\lambda \in \mathbb{R}^2$:

$$p(\lambda) = \mathcal{N}(\lambda : \eta^\lambda, L^\lambda) = \frac{1}{\sqrt{(2\pi)^2|L^\lambda|}} e^{-\frac{1}{2}(\lambda-\eta^\lambda)^T P^\lambda(\lambda-\eta^\lambda)}. \tag{6.23}$$

A higher value of $P^{\tilde{v}}$, P^θ and P^λ represents the robot's higher confidence in its prior estimates $\eta^{\tilde{v}}$, η^θ, and η^λ, respectively.

6.6.3 Simplification of the Internal Energy Action $\bar{U}$

This section aims at using the distributions from the previous sections to simplify $\bar{U}$. The logarithm of a Gaussian prior after dropping constants takes the general form:

$$\ln p(\theta) = \ln \mathcal{N}(\theta : \eta^\theta, L^\theta) = -\frac{1}{2}(\theta-\eta^\theta)^T(L^\theta)^{-1}(\theta-\eta^\theta) - \frac{1}{2}\ln|L^\theta| \tag{6.24}$$

Therefore, substituting Equations (6.19)-(6.23) in Equation (6.18) and simplifying it using the prediction error terms $\epsilon^{\tilde{x}} = D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v}$, $\epsilon^{\tilde{y}} = \tilde{y} - \tilde{C}\tilde{x}$, $\epsilon^{\tilde{v}} = \tilde{v} - \eta^{\tilde{v}}$ $\epsilon^\theta = \theta - \eta^\theta$, and $\epsilon^\lambda = \lambda - \eta^\lambda$, after dropping constants, yields:

$$U(\vartheta, y) = -\frac{1}{2}\epsilon^{\theta T}P^\theta\epsilon^\theta + \frac{1}{2}\ln|P^\theta| - \frac{1}{2}\epsilon^{\lambda T}P^\lambda\epsilon^\lambda + \frac{1}{2}\ln|P^\lambda| - \frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon} + \frac{1}{2}\ln|\tilde{\Pi}|, \tag{6.25}$$

where $\tilde{\epsilon} = \begin{bmatrix} \epsilon^{\tilde{y}} \\ \epsilon^{\tilde{v}} \\ \epsilon^{\tilde{x}} \end{bmatrix} = \begin{bmatrix} \tilde{y} - \tilde{C}\tilde{x} \\ \tilde{v} - \eta^{\tilde{v}} \\ D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix}$, and $\tilde{\Pi} = diag(\tilde{\Pi}^z, P^{\tilde{v}}, \tilde{\Pi}^w)$. Here $diag(.,.)$ is the block

diagonal operator. Grouping the internal energy terms of the temporal and nontemporal parameters yields:

$$U(\vartheta, y) = U(t) + U(\theta) + U(\lambda), \tag{6.26}$$

where

$$U(t) = -\frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon} + \frac{1}{2}\ln|\tilde{\Pi}|. \tag{6.27}$$

Summing up the internal energy of all the temporal parameters over time yields the action of internal energy as follows:

$$\begin{aligned} \bar{U} &= U(\theta) + U(\lambda) + \sum_t U(t) \\ &= -\frac{1}{2}\epsilon^{\theta T}P^\theta\epsilon^\theta + \frac{1}{2}\ln|P^\theta| - \frac{1}{2}\epsilon^{\lambda T}P^\lambda\epsilon^\lambda + \frac{1}{2}\ln|P^\lambda| - \sum_t \frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon} + \frac{1}{2}\sum_t\ln|\tilde{\Pi}|. \end{aligned} \tag{6.28}$$

It can be observed from Equation (6.28) that the robot's priors $(\eta^{\tilde{v}}, P^{\tilde{v}}, \eta^\theta, P^\theta, \eta^\lambda, P^\lambda)$ enter the free energy objective through the internal energy term. Intuitively, this can be seen as the direct influence of the robot's prior beliefs on the inference process through the mismatches in the robot's predictions. These weighed prediction errors drive the robot's desire to maintain an equilibrium between its internal model and the generative process in the environment. A large mismatch between the robot's predictions and the data results in a large prediction error, which gets precision-weighted and enters the free energy objective through its internal energy.

6.7 Mean-Field Approximation

The second widely used assumption for the simplification of free energy objectives is the factorization of parameters into independent subdensities for the recognition density [170], given by:

$$q(\vartheta) = \prod_i q(\theta^i) = q(\tilde{x})q(\tilde{v})q(\theta)q(\lambda). \tag{6.29}$$

This approximation assumes the conditional independence between subdensities (states and parameters, for example). The subdensities are assumed to interact with each other only through the mean-field quantities. The strong biological plausibility of this approximation in terms of biological brain's inference process is delineated in [44]. The main advantage of this approximation is the simplification of $\bar{V}$ in Equation (6.15). However, the mathematical proof for this simplification is missing from the DEM literature [170]. Therefore, this section aims to fill this gap by deriving these proofs by delineating all the intermediate assumptions.

6.7.1 Simplification of the Entropy Action $\bar{H}$

The entropy of the density in Equation (6.14), given by $H(\vartheta)_{q(\vartheta)} = -\langle \ln q(\vartheta) \rangle_{q(\vartheta)}$, can be simplified for all parameters as:

$$H(\vartheta)_{q(\vartheta)} = -\int \int \int \int q(\tilde{x}, \tilde{v}, \theta, \lambda) \ln q(\tilde{x}, \tilde{v}, \theta, \lambda) d\tilde{x} d\tilde{v} d\theta d\lambda \tag{6.30}$$

Substituting the mean-field assumption given in Equation (6.29) in Equation (6.30) and using the property of normalized recognition densities $\int q(\theta^i)d\theta^i = 1$ yields:

$$H(\vartheta)_{q(\vartheta)} = H(\theta) + H(\lambda) + H(t), \tag{6.31}$$

where $H(\theta) = -\langle \ln q(\theta) \rangle_{q(\theta)}$, $H(\lambda) = -\langle \ln q(\lambda) \rangle_{q(\lambda)}$ and $H(t) = -\langle \ln q(\tilde{x}) \rangle_{q(\tilde{x})} - \langle \ln q(\tilde{v}) \rangle_{q(\tilde{v})}$. We place the Laplace approximation over the marginals of the recognition densities of all parameters as:

$$
\begin{aligned}
q(\theta) &= \mathcal{N}(\theta : \mu^\theta, \Sigma^\theta) = \frac{1}{\sqrt{(2\pi)^l |\Sigma^\theta|}} e^{-\frac{1}{2}(\theta - \mu^\theta)^T \Pi^\theta (\theta - \mu^\theta)} \\
q(\lambda) &= \mathcal{N}(\lambda : \mu^\lambda, \Sigma^\lambda) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma^\lambda|}} e^{-\frac{1}{2}(\lambda - \mu^\lambda)^T \Pi^\lambda (\lambda - \mu^\lambda)} \\
q(\tilde{x}) &= \mathcal{N}(\tilde{x} : \mu^{\tilde{x}}, \Sigma^{\tilde{x}}) = \frac{1}{\sqrt{(2\pi)^{n(p+1)} |\Sigma^{\tilde{x}}|}} e^{-\frac{1}{2}(\tilde{x} - \mu^{\tilde{x}})^T \Pi^{\tilde{x}} (\tilde{x} - \mu^{\tilde{x}})} \\
q(\tilde{v}) &= \mathcal{N}(\tilde{v} : \mu^{\tilde{v}}, \Sigma^{\tilde{v}}) = \frac{1}{\sqrt{(2\pi)^{r(d+1)} |\Sigma^{\tilde{v}}|}} e^{-\frac{1}{2}(\tilde{v} - \mu^{\tilde{v}})^T \Pi^{\tilde{v}} (\tilde{v} - \mu^{\tilde{v}})}.
\end{aligned}
\tag{6.32}
$$

The recognition densities given in Equation (6.32) might look similar to the priors distributions given in Equation (6.21)-(6.23), mainly due to the common Gaussian distribution. The prior distributions are centered around the prior mean and prior covariances, whereas the recognition density is centered around the mean μ^i and conditional covariance Σ^i of the i-th parameter set.

The action of entropy can be calculated by substituting Equation (6.32) in Equation (6.31) and summing up the entropy terms of time dependent parameters with respect to time. Upon dropping the constant terms, it simplifies to:

$$\bar{H} = H(\theta) + H(\lambda) + \sum_t H(t) = \frac{1}{2}\ln|\Sigma^\theta| + \frac{1}{2}\ln|\Sigma^\lambda| + \frac{1}{2}\sum_t \ln|\Sigma^{\tilde{x}}| + \frac{1}{2}\sum_t \ln|\Sigma^{\tilde{v}}|, \qquad (6.33)$$

Equation (6.33) shows how the uncertainty in the robot's inference directly enters the objective $\bar{F}$, through $\bar{H}$.

6.7.2 Mean-Field Terms

Given the simplified expressions for $\bar{U}$ and $\bar{H}$, the next step towards finding $\bar{F}$ in Equation (6.15) is to evaluate $\bar{V}$ given by:

$$\bar{V} = \int \langle U(y, \vartheta) \rangle_{q(\vartheta)} dt \qquad (6.34)$$

$U(y, \vartheta)$ can be simplified using the second-degree Taylor series expansion near the mean $\mu^\vartheta = \{\mu^{\tilde{x}}, \mu^{\tilde{v}}, \mu^\theta, \mu^\lambda\}$ as:

$$U(y, \tilde{x}, \tilde{v}, \theta, \lambda) = U(y, \mu^\vartheta) + \sum_{i=1}^{4} U(y, \mu^\vartheta)_{\theta^i}(\theta^i - \mu^i) + \sum_{i=1}^{4}\sum_{j=1}^{4}(\theta^i - \mu^i)^T U(y, \mu^\vartheta)_{\theta^i\theta^j}(\theta^j - \mu^j),$$
$$(6.35)$$

where we use the shorthand $U(y, \mu^\vartheta) = U(y, \vartheta)|_{\vartheta=\mu^\vartheta}$ and $U(y, \mu^\vartheta)_\vartheta = U(y, \vartheta)_\vartheta|_{\vartheta=\mu^\vartheta}$. This approximation of the internal energy has nontrivial implications in terms of the biological brain's decision-making process [43]. The second order approximation is justified because the Laplace and mean-field approximations entail an internal energy that is quadratic in $\tilde{x}$, $\tilde{v}$, and θ, as given in Equation (6.25), thereby reducing all its higher-order derivatives to zero. Moreover, the higher derivatives of $U(y, \mu^\vartheta)$ with respect to λ, reduce to zero because of the assumptions made in Section 6.10. By differentiating Equation (6.25) at the mean μ^i, $U(y, \mu^\vartheta)_{\theta^i}$ can be found to be all zeroes, which upon substitution in Equation (6.35) simplifies to:

$$U(y, \vartheta) = U(y, \mu^\vartheta) + \sum_{i=1}^{4}\sum_{j=1}^{4}(\theta^i - \mu^i)^T U(y, \mu^\vartheta)_{\theta^i\theta^j}(\theta^j - \mu^j). \qquad (6.36)$$

Substituting it in Equation (6.34), upon simplification yields:

$$\bar{V} = \bar{U}(y, \mu^\vartheta) + \int W dt \qquad (6.37)$$

where

$$W = \frac{1}{2}\sum_{i,j=1}^{4}\langle(\theta^i-\mu^i)^T U(y,\mu^\vartheta)_{\theta^i\theta^j}(\theta^j-\mu^j)\rangle_{q(\vartheta^i)q(\vartheta^j)}$$

$$= \frac{1}{2}\sum_{i,j=1}^{4}\langle tr\Big((\theta^i-\mu^i)^T U(y,\mu^\vartheta)_{\theta^i\theta^j}(\theta^j-\mu^j)\Big)\rangle_{q(\vartheta^i)q(\vartheta^j)}$$

$$= \frac{1}{2}\sum_{i,j=1}^{4}\langle tr\Big((\theta^j-\mu^j)(\theta^i-\mu^i)^T U(y,\mu^\vartheta)_{\theta^i\theta^j}\Big)\rangle_{q(\vartheta^i)q(\vartheta^j)}$$

$$= \frac{1}{2}\sum_{i,j=1}^{4} tr\Big[\langle(\theta^j-\mu^j)(\theta^i-\mu^i)^T\rangle_{q(\vartheta^i)q(\vartheta^j)}U(y,\mu^\vartheta)_{\theta^i\theta^j}\Big] = \frac{1}{2}\sum_{i,j=1}^{4} tr\Big[\Sigma^{ij}U(y,\mu^\vartheta)_{\theta^i\theta^j}\Big]$$

Since the parameters ϑ^i are assumed to be independent of each other, the covariance between them drops to zero, resulting in:

$$\begin{aligned}\bar{V} &= \bar{U}(y,\mu^\vartheta) + \frac{1}{2}\int\sum_{i=1}^{4} tr\Big[\Sigma^i U(y,\mu^\vartheta)_{\theta^i\theta^i}\Big]dt\\ &= \bar{U}(y,\mu^\vartheta) + \int\Big[W^{\tilde{x}}+W^{\tilde{v}}+W^{\theta}+W^{\lambda}\Big]dt,\end{aligned} \tag{6.38}$$

where

$$W^{\vartheta^i} = \frac{1}{2}tr\Big[\Sigma^i U(y,\mu^\vartheta)_{\theta^i\theta^i}\Big] \tag{6.39}$$

is defined as the mean-field term. Equation (6.38) shows that $\bar{V}$ is the sum of internal energy action and the mean field terms. The simplification of $\bar{V}$ is one of the main advantages of using mean-field approximation. However, this approximation can be relaxed to build Generalized Filtering [169, 209], which is mainly relevant to nonlinear identification. It involves the modeling of parameters and hyperparameters in generalized coordinates (together with states) for online system identification. However, in this work, we take a simpler approach.

6.8 Simplified Free Energy Objectives
This section aims at simplifying the free energy objectives using the results from Sections 6.6 and 6.7.

6.8.1 Simplification of Free Action
$\bar{F}$ is simplified by substituting Equations (6.38), (6.28), and (6.33) into Equation (6.15), yielding:

$$\begin{aligned}\bar{F} =& -\frac{1}{2}(\epsilon^{\theta T}P^\theta\epsilon^\theta)|_{\vartheta=\mu^\vartheta} + \frac{1}{2}\ln|P^\theta| - \frac{1}{2}(\epsilon^{\lambda T}P^\lambda\epsilon^\lambda)|_{\vartheta=\mu^\vartheta} + \frac{1}{2}\ln|P^\lambda| - \sum_t\frac{1}{2}(\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon})|_{\vartheta=\mu^\vartheta}\\ &+ \frac{1}{2}\sum_t(\ln|\tilde{\Pi}|)|_{\vartheta=\mu^\vartheta} + \frac{1}{2}\sum_t(\ln|\Sigma^{\tilde{x}}| + \ln|\Sigma^{\tilde{v}}|) + \frac{1}{2}\ln|\Sigma^\theta| + \frac{1}{2}\ln|\Sigma^\lambda|\\ &+ \frac{1}{2}\sum_t tr\Big[\Sigma^{\tilde{x}}U(y,\mu^\vartheta)_{\tilde{x}\tilde{x}} + \Sigma^{\tilde{v}}U(y,\mu^\vartheta)_{\tilde{v}\tilde{v}} + \Sigma^\theta U(y,\mu^\vartheta)_{\theta\theta} + \Sigma^\lambda U(y,\mu^\vartheta)_{\lambda\lambda}\Big].\end{aligned} \tag{6.40}$$

We highlight three important terms in Equation (6.40): the combined prediction error of (generalized) outputs, inputs, and states $E = \frac{1}{2}(\tilde{\epsilon}^T \tilde{\Pi}\tilde{\epsilon})|_{\vartheta = \mu^\vartheta}$, the log determinant of noise precision $G = (\ln|\tilde{\Pi}|)|_{\vartheta = \mu^\vartheta}$, and the mean field term $W^{\vartheta^i} = \frac{1}{2}tr\left[\Sigma^i U(y,\mu^\vartheta)_{\theta^i\theta^i}\right]$. These terms will be used rigorously in the rest of the document.

6.8.2 Simplification of the Parameter Precisions

One of the main advantages of Laplace approximation is the simple evaluation of the covariance associated with the parameter estimation. This is achieved by setting the gradient of the free action with respect to individual parameter covariance as zero. The free action gradients with respect to covariance of paramaters ϑ^i is given by:

$$\frac{\partial \bar{F}}{\partial \Sigma^i} = \frac{\partial}{\partial \Sigma^i}\left(\frac{1}{2}\ln|\Sigma^i| + \frac{1}{2}\sum_t tr(\Sigma^i U(y,\mu^\vartheta)_{\vartheta^i\vartheta^i})\right) = \frac{1}{2}\left((\Sigma^i)^{-1} + \bar{U}(y,\mu^\vartheta)_{\vartheta^i\vartheta^i}\right). \quad (6.41)$$

The optimal parameter covariance is the covariance that maximizes the free action with zero gradients. Assuming that the parameter covariances are time-invariant, and equating the gradients to zero yields $\Pi^i = -\bar{U}(y,\mu^\vartheta)_{\vartheta^i\vartheta^i}$, resulting in:

$$\Pi^{\tilde{x}} = -U(y,\mu^\vartheta)_{\tilde{x}\tilde{x}}, \ \Pi^{\tilde{v}} = -U(y,\mu^\vartheta)_{\tilde{v}\tilde{v}}, \ \Pi^\theta = -\bar{U}(y,\mu^\vartheta)_{\theta\theta}, \ \Pi^\lambda = -\bar{U}(y,\mu^\vartheta)_{\lambda\lambda} \quad (6.42)$$

From Equation (6.42), it is clear that the precision of parameters can be estimated just by evaluating the negative curvature of the internal energy at the conditional mode. These precision values denote the confidence of the estimator in the parameter estimate. Ideally, the parameter precision increases as the estimation process proceeds. Intuitively, the robot is more confident about its estimates (higher precision) when its predictions on the sensory data show the least variance.

6.8.3 Free Action at Optimal Precision

Substituting Equation (6.42) into Equation (6.39) at optimal precisions results in constant mean field terms. Therefore, the mean field terms in the free action given in Equation (6.40), reduces to a constant under the optimal precision given by Equation (6.42). Therefore, the free action at optimal precision for an LTI system reduces to:

$$\bar{F} = -\frac{1}{2}\sum_t \left[\underbrace{(\tilde{\mathbf{y}} - \tilde{C}\tilde{x})^T \tilde{\Pi}^z(\tilde{\mathbf{y}} - \tilde{C}\tilde{x})\Big|_{\vartheta=\mu^\vartheta}}_{\text{prediction error of outputs}} + \underbrace{(\tilde{v} - \eta^{\tilde{v}})^T P^{\tilde{v}}(\tilde{v} - \eta^{\tilde{v}})\Big|_{\vartheta=\mu^\vartheta}}_{\text{prediction error of inputs}}\right]$$

$$-\frac{1}{2}\underbrace{(\theta - \eta^\theta)^T P^\theta(\theta - \eta^\theta)\Big|_{\vartheta=\mu^\vartheta}}_{\text{prediction error of parameters}} - \frac{1}{2}\underbrace{(\lambda - \eta^\lambda)^T P^\lambda(\lambda - \eta^\lambda)\Big|_{\vartheta=\mu^\vartheta}}_{\text{prediction error of hyperparameters}}$$

$$-\frac{1}{2}\sum_t \underbrace{(D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v})^T \tilde{\Pi}^w(D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v})\Big|_{\vartheta=\mu^\vartheta}}_{\text{prediction error of (generalized) states}} + \underbrace{\frac{1}{2}n_t\ln|\Sigma^X|}_{\text{state and input entropy}} \quad (6.43)$$

$$+\frac{1}{2}n^t \underbrace{\left[\ln|\tilde{\Pi}^z| + \ln|P^{\tilde{v}}| + \ln|\tilde{\Pi}^w|\right]\Big|_{\vartheta=\mu^\vartheta}}_{\text{noise entropy}} + \underbrace{\frac{1}{2}\ln|\Sigma^\theta P^\theta|}_{\text{parameter entropy}} + \underbrace{\frac{1}{2}\ln|\Sigma^\lambda P^\lambda|}_{\text{hyperparameter entropy}}$$

Note that the free action is not a function of the latent variables ($\vartheta = \{\tilde{x}, \tilde{v}, \theta, \lambda\}$) but the sufficient statistics ($\mu^{\vartheta}, \Sigma^{\vartheta}$) of the approximate posterior. For example, the weighted output prediction error term of $\bar{F}$ is $(\tilde{\mathbf{y}} - \tilde{C}\tilde{x})^T \tilde{\Pi}^z (\tilde{\mathbf{y}} - \tilde{C}\tilde{x})\big|_{\vartheta=\mu^{\vartheta}} = (\tilde{\mathbf{y}} - \mu^{\tilde{C}}\mu^{\tilde{x}})^T \tilde{\Pi}^z \big|_{\lambda^z=\mu^{\lambda z}} (\tilde{\mathbf{y}} - \mu^{\tilde{C}}\mu^{\tilde{x}})$, where $\big|_{\vartheta=\mu^{\vartheta}}$ denotes the evaluation at $\vartheta = \mu^{\vartheta}, (\tilde{x} = \mu^{\tilde{x}}, \tilde{v} = \mu^{\tilde{v}}, \theta = \mu^{\theta}, \lambda = \mu^{\lambda})$. We regroup the time-dependent components into one variable $X = \begin{bmatrix} \mu^{\tilde{x}} \\ \mu^{\tilde{v}} \end{bmatrix}$, and use capital letters for the mean estimates of time-independent components ($\Theta = \mu^{\theta}, \Lambda = \mu^{\lambda}$). Note that X, Θ, and Λ are part of the generative model and are the mean estimate of the components of plant dynamics.

6.8.4 Equivalence with the EM Algorithm

One of the most popular approaches to solve the Expectation Maximization (EM) algorithm for state space models is to use the Maximum Likelihood Estimation (MLE), where the objective function is the log-likelihood of data, given by [232]:

$$
\ln L = -\frac{1}{2} \sum_t (x_t - Ax_{t-1} - Bu_{t-1})^T \Pi^w (x_t - Ax_{t-1} - Bu_{t-1}) - \frac{n_t}{2} \ln |\Sigma^w|
$$
$$
- \frac{n_t}{2} \ln |\Sigma^z| - \frac{1}{2} \ln |\Sigma^x| - \frac{1}{2}(x_0 - \mu)^T \Pi^x (x_0 - \mu) - \frac{1}{2} \sum_t (y_t - Cx_t)^T \Pi^z (y_t - Cx_t).
$$

(6.44)

Comparing the objective functions of EM and DEM given by Equation (6.44) and Equation (6.43), respectively, DEM is equivalent to EM when:

- the mean field terms are neglected,

- the generalized motion is not considered, and

- the robot's priors on ϑ are not considered.

Therefore, DEM can be considered as a generalized version of the EM algorithm with additional capabilities.

The free action at optimal precision (Equation (6.43)) is the sum of prediction errors (PE) and entropy of generalized states, parameters, noise, and hyperparameters and is independent of mean-field terms. Although the mean-field terms turns into a constant at optimal precision, their gradients do not. This property is leveraged in Section 6.9 to account for the uncertainties in the parameter estimation during state estimation and vice versa, through the gradient of mean-field terms.

6.9 Update Rules for Estimation

The Free Energy objectives (Section 6.5), together with the two simplifying approximations (Sections 6.6 and 6.7), will be combined with an optimization procedure to form the ultimate DEM algorithm. The optimization procedure itself consists of the following two sets of update rules:

1. a gradient ascent over its free action $\bar{F}$ for the time invariant parameters θ and λ,

2. a gradient ascent over its free energy F for the time varying parameters X,

where F and $\bar{F}$ are related through $F = \frac{\partial \bar{F}}{\partial t}$. The core idea is that the time varying parameters $\tilde{x}$ and $\tilde{v}$ can be estimated online from the robot's instantaneous free energy, whereas the time-invariant parameters θ and λ can be estimated from its free action after observing a sequence of data. Accordingly, the update rules for both the gradient ascends are given by

$$\frac{\partial \mu^i}{\partial a} = k^i \frac{\partial \bar{F}}{\partial \mu^i}, \tag{6.45}$$

for the ath parameter update of the time invariant parameters and

$$\frac{\partial \mu^i}{\partial t} = D\mu^i + k^i \frac{\partial F}{\partial \mu^i}, \tag{6.46}$$

for the time=varying parameters, where k^i is the learning rate. The presence of $D\mu^i$ term in Equation (6.46) differentiates the update rule from the general gradient ascent equation used in machine learning. This is to accommodate the boundary condition that when F is maximized, $\frac{\partial F}{\partial \mu^i} = 0$ and $\dot{\mu}^i = D\mu^i$. In other words, when the free energy is maximized, the motion of the generalized states becomes their generalized motion [233]. However, the update equations for the time-invariant parameters, θ and λ, do not require the $D\mu^i$ term. Therefore, the update equations at t^{th} time, a^{th} parameter update step, and b^{th} hyperparameter update step, after regrouping the (generalized) states and inputs as $X = \begin{bmatrix} \mu^{\tilde{x}} \\ \mu^{\tilde{v}} \end{bmatrix}$, is given by:

$$\frac{\partial X}{\partial t} = DX + k^X \frac{\partial F}{\partial X}, \quad \frac{\partial \Theta}{\partial a} = k^\Theta \frac{\partial \bar{F}}{\partial \Theta}, \quad \frac{\partial \Lambda}{\partial b} = k^\Lambda \frac{\partial \bar{F}}{\partial \Lambda}, \tag{6.47}$$

where $D = diag(D^x, D^v)$. Note that the gradient update rules are written not on the latent variables ($\tilde{x}, \tilde{v}, \theta$ and λ), but on their mean estimates (X, Θ and Λ). Since the update rules should be implemented in the discrete domain for robotics applications, Equation (6.47) is discretized under local linearization with the corresponding Jacobians as $J^X = D + k^X \frac{\partial^2 F}{\partial X^2}$, $J^\Theta = k^\Theta \frac{\partial^2 \bar{F}}{\partial \Theta^2}$, and $J^\Lambda = k^\Lambda \frac{\partial^2 \bar{F}}{\partial \Lambda^2}$. The (generalized) state and input update at time t, parameter update at step a and hyperparameter update at step b are given by:

$$\begin{aligned}
X_{t+\Delta t} &= X_t + \left(e^{J^X \Delta t} - I \right) (J^X)^{-1} \frac{\partial X}{\partial t}, \\
\Theta_{a+1} &= \Theta_a + \left(e^{J^\Theta} - I \right) (J^\Theta)^{-1} \frac{\partial \Theta}{\partial a}, \\
\Lambda_{b+1} &= \Lambda_b + \left(e^{J^\Lambda} - I \right) (J^\Lambda)^{-1} \frac{\partial \Lambda}{\partial b}.
\end{aligned} \tag{6.48}$$

Equation (6.48) shows that the update rules are dependent only on the gradients and curvatures of the free energy objectives.

6.9.1 The DEM Algorithm

The DEM algorithm is an iterative model inversion algorithm that uses Equations (6.48) and (6.42) to perform estimation on causal dynamic systems. It can be expressed using three steps:

1. D step: (generalized) state and input estimation,

2. E step: parameter estimation,

3. M step: noise hyperparameter estimation,

a nomenclature that is similar to the EM algorithm. Figure 6.2 shows an intuitive block diagram that demonstrates the inference process of DEM as a coupled dynamics between D, E, and M steps. The data from the environment and the robot brain's prior distributions are used to infer the generative process. The pseudocode given in Algorithm 1 demonstrates how DEM performs estimation using only the gradient and curvatures of the free energy objectives. The next sections will focus on deriving the algebraic expressions for these quantities.
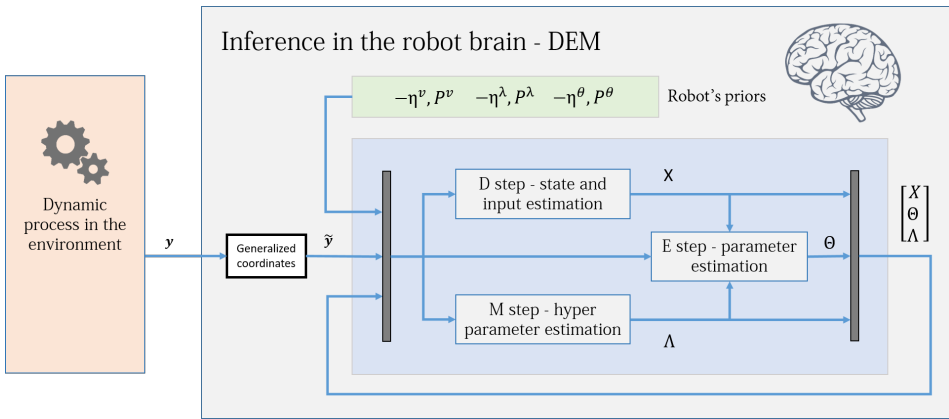
Fig 6.2: The DEM algorithm is represented using three coupled steps: D, E, and M steps. The algorithm combines the data from the environment with the robot's prior beliefs to infer the states, inputs, parameters and hyperparameters of the system. For each parameter update in the E step, the D step updates the (generalized) states and inputs for all times instances, and the M step iterates until hyperparameter convergence, as demonstrated in Algorithm 1. The dynamic process is the generative model in Section 6.4.1, the priors are the distributions given in Section 6.6.2 and the generalized coordinates block is defined in Section 6.4.4. Section 6.14 will elaborate on the D, E, and M blocks.

6.9.2 Updated Equations for Estimation

The free energy gradients in Equation (6.47) can be evaluated by differentiating Equation (6.40) with μ^i. Substituting the resulting expression in Equation (6.47), upon simplifi-

Algorithm 1 Dynamics Expectation Maximization

Data: Time series data of output **y**

Result: $X, \Theta, \Lambda, \Pi^{\tilde{x}, \tilde{v}}, \Pi^{\theta}, \Pi^{\lambda}$

$a = 0$

while Θ *not converged* **do**

   for $t = 0{:}\Delta t{:}T$ **do**   ▷ D step

$$\frac{\partial X}{\partial t} = DX + k^X \frac{\partial F}{\partial X}$$

$$J^X = D + k^X \frac{\partial^2 F}{\partial X^2}$$

$$X_{t+\Delta t} = X_t + (e^{J^X \Delta t} - I)(J^X)^{-1}\frac{\partial X}{\partial t}$$

   end

   while Λ *not converged* **do**   ▷ M step

$$\frac{\partial \Lambda}{\partial b} = k^{\Lambda}\frac{\partial \bar{F}}{\partial \Lambda}$$

$$J^{\Lambda} = k^{\Lambda}\frac{\partial^2 \bar{F}}{\partial \Lambda^2}$$

$$\Lambda \leftarrow \Lambda + (e^{J^{\Lambda}} - I)(J^{\Lambda})^{-1}\frac{\partial \Lambda}{\partial b}$$

   end

   $\bar{F}_a \leftarrow$ Equation (6.43)   ▷ $\bar{F}$ at optimal precision

   if $\bar{F}_a > \bar{F}_{a-1}$ **then**   ▷ update Θ if $\bar{F}$ increased

$$\frac{\partial \Theta}{\partial a} = k^{\Theta}\frac{\partial \bar{F}}{\partial \Theta}$$

$$J^{\Theta} = k^{\Theta}\frac{\partial^2 \bar{F}}{\partial \Theta^2}$$

$$\Theta \leftarrow \Theta + (e^{J^{\Theta}} - I)(J^{\Theta})^{-1}\frac{\partial \Theta}{\partial a}; \qquad\qquad\qquad ▷ \text{E step}$$

   end

   $a{+}{+}$

   $\Pi^{\tilde{x}, \tilde{v}} = -U(y, \mu^{\vartheta})\begin{bmatrix}\tilde{x}\\\tilde{v}\end{bmatrix}\begin{bmatrix}\tilde{x}\\\tilde{v}\end{bmatrix};$ ▷ update generalized state precision

   $\Pi^{\theta} = -\bar{U}(y, \mu^{\vartheta})_{\theta\theta};$ ▷ parameter precision

   $\Pi^{\lambda} = -\bar{U}(y, \mu^{\vartheta})_{\lambda\lambda};$ ▷ hyperparameter precision

end

cation yields:

$$\dot{X} = DX + k^X\left[-E_X + W_X^X + W_X^{\Theta} + W_X^{\Lambda}\right]$$

$$\frac{\partial \Theta}{\partial a} = k^{\Theta}\left[-P^{\theta}\epsilon^{\theta}\Big|_{\vartheta=\mu^{\vartheta}} + \sum_t\left(-E_{\Theta} + W_{\Theta}^X + W_{\Theta}^{\Theta} + W_{\Theta}^{\Lambda}\right)\right] \tag{6.49}$$

$$\frac{\partial \Lambda}{\partial b} = k^{\Lambda}\left[-P^{\lambda}\epsilon^{\lambda}\Big|_{\vartheta=\mu^{\vartheta}} + \sum_t(-E_{\Lambda} + W_{\Lambda}^X + W_{\Lambda}^{\Theta} + W_{\Lambda}^{\Lambda} + G_{\Lambda})\right],$$

where $E_{\mu^i} = \frac{\partial E}{\partial \mu^i} = \frac{1}{2}\frac{\partial(\tilde{\varepsilon}^T \tilde{\Pi}\tilde{\varepsilon})\big|_{\vartheta=\mu^{\vartheta}}}{\partial \mu^i}$ is the gradient of the prediction error with respect to μ^i, $W_{\mu^j}^{\mu^i} = \frac{\partial W^{\mu^i}}{\partial \mu^j} = \frac{1}{2}\frac{\partial}{\partial \mu^j}tr\left[\Sigma^i U(y, \mu^{\vartheta})_{\mu^i\mu^i}\right]$ is the gradient of the mean field term of μ^i with respect to μ^j, and $G_{\Lambda} = \frac{1}{2}\frac{\partial}{\partial \Lambda}\ln|\tilde{\Pi}|\big|_{\vartheta=\mu^{\vartheta}}$ is the gradient of the log-determinant of the precision matrix

with respect to Λ. From Equation (6.49), the Jacobians that are required for the updates given in Equation (6.48) can be evaluated as:

$$J^X = D + k^X \left[-E_{XX} + W_{XX}^X + W_{XX}^\Theta + W_{XX}^\Lambda \right]$$

$$J^\Theta = k^\Theta \left[-P^\theta + \sum_t \left(-E_{\Theta\Theta} + W_{\Theta\Theta}^X + W_{\Theta\Theta}^\Theta + W_{\Theta\Theta}^\Lambda \right) \right] \tag{6.50}$$

$$J^\Lambda = k^\Lambda \left[-P^\lambda + \sum_t \left(-E_{\Lambda\Lambda} + W_{\Lambda\Lambda}^X + W_{\Lambda\Lambda}^\Theta + W_{\Lambda\Lambda}^\Lambda + G_{\Lambda\Lambda} \right) \right]$$

The Jacobians are employed in different loops corresponding to the D, E, and M steps in the Algorithm 1.

6.9.3 Update Equation for Precision of Estimates

The uncertainty in estimation is represented by the inverse of precision matrices Π^{ϑ^i}. Differentiating Equations (6.25) and (6.28) twice and substituting it into Equation (6.42) upon simplification yield:

$$\Pi^{\tilde{x},\tilde{v}} = -U(y,\mu^\vartheta)_{\left[\begin{smallmatrix}\tilde{x}\\\tilde{v}\end{smallmatrix}\right]\left[\begin{smallmatrix}\tilde{x}\\\tilde{v}\end{smallmatrix}\right]} = E_{XX}$$

$$\Pi^\theta = -\bar{U}(y,\mu^\vartheta)_{\theta\theta} = P^\theta + \sum_t E_{\Theta\Theta} \tag{6.51}$$

$$\Pi^\lambda = -\bar{U}(y,\mu^\vartheta)_{\lambda\lambda} = P^\lambda + \sum_t \left(E_{\Lambda\Lambda} - G_{\Lambda\Lambda} \right))$$

The only unknowns in the DEM update equations given by Equations (6.49), (6.50), and (6.51) are the gradients and curvatures of E, W, and G. Sections 6.10–6.12 will deal with evaluating the simplified algebraic expressions for these gradients and curvatures. Using these simplifications, Section 6.13 will proceed towards expanding Algorithm 1.

6.10 Gradients of (Log Determinant of) Precision

This section aims at evaluating the gradients of the log determinant of noise precision $(G_\Lambda, G_{\Lambda\Lambda})$ that are required for the hyperparameter update rules of the DEM algorithm. The precision matrix for hyperparameter estimation is modeled as:

$$\tilde{\Pi} = diag(\tilde{\Pi}^z, P^{\tilde{v}}, \tilde{\Pi}^w) = diag(e^{\lambda^z}(S \otimes \Omega^z), P^{\tilde{v}}, e^{\lambda^w}(S \otimes \Omega^w)), \tag{6.52}$$

where S is the noise smoothness matrix given by Equation (6.7) and Ω^z, Ω^w are the constant matrices given in Equation (6.6). Here the precision matrix is parametrized using $\lambda = \left[\begin{smallmatrix}\lambda^z\\\lambda^w\end{smallmatrix}\right]$, which is the only unknown in Equation (6.52). Therefore, the log determinant of precision and its gradients can be written as:

$$G = \ln|\tilde{\Pi}|\Big|_{\vartheta=\mu^\vartheta} \,,\ G_\Lambda = \frac{1}{2} \begin{bmatrix} tr(\tilde{\Pi}_{\lambda^z}\tilde{\Pi}^{-1}) \\ tr(\tilde{\Pi}_{\lambda^w}\tilde{\Pi}^{-1}) \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta} \,,$$

$$G_{\Lambda^i\Lambda^j} = \frac{1}{2} tr(\tilde{\Pi}_{\lambda^i\lambda^j}\tilde{\Pi}^{-1} - \tilde{\Pi}_{\lambda^i}\tilde{\Pi}^{-1}\tilde{\Pi}_{\lambda^j}\tilde{\Pi}^{-1})\Big|_{\vartheta=\mu^\vartheta} \,, \tag{6.53}$$

where λ^i is the ith element in λ. The gradients of precision in Equation (6.53) can be evaluated by differentiating Equation (6.52) as:

$$\tilde{\Pi}_{\lambda^z} = diag(e^{\lambda^z}(S \otimes \Omega^z), O, O) = diag(\tilde{\Pi}^z, O, O),$$
$$\tilde{\Pi}_{\lambda^w} = diag(O, O, e^{\lambda^w}(S \otimes \Omega^z)) = diag(O, O, \tilde{\Pi}^z), \qquad (6.54)$$
$$\tilde{\Pi}_{\lambda^z \lambda^w} = O, \ \tilde{\Pi}_{\lambda^w \lambda^z} = O.$$

Substituting Equations (6.54) and (6.52) in Equation (6.53) yields:

$$G_\Lambda = \frac{1}{2} \begin{bmatrix} tr(I_{n^{\tilde{\Pi}^z}}) \\ tr(I_{n^{\tilde{\Pi}^w}}) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} n^{\tilde{\Pi}^z} \\ n^{\tilde{\Pi}^w} \end{bmatrix}, \qquad (6.55)$$

where $I_{n^{\tilde{\Pi}^z}}$ is the identity matrix of size $n^{\tilde{\Pi}^z}$, which is the size of the $\tilde{\Pi}^z$ matrix. Π^w and Π^z are modeled to have an exponential relation with λ, so that any updates on λ would result in positive semi-definite precision matrices. However, this relation entails an infinitely differentiable precision matrix with respect to λ, increasing the computational complexity of the algorithm. Therefore, an approximation is made by forcefully setting $\tilde{\Pi}_{\lambda^z \lambda^z} = \tilde{\Pi}_{\lambda^w \lambda^w} = O$, while maintaining the exponential relation between $\tilde{\Pi}$ and λ, thereby ensuring that the optimization process proceeds along the correct gradients $\tilde{\Pi}_\lambda$. Together with Equation (6.54), this approximation results in $\tilde{\Pi}_{\lambda\lambda} = O$. This assumption has two direct consequences:

- It simplifies all the update rules given in Equations (6.49) and (6.50),

- It simplifies the precision update rule for hyperparameters given in Equation (6.51).

A direct consequence of this approximation is in the simplification of $G_{\Lambda\Lambda}$ in Equation (6.53), expressed as:

$$G_{\Lambda\Lambda} = -\frac{1}{2} \begin{bmatrix} tr(\tilde{\Pi}_{\lambda^z}\tilde{\Pi}^{-1}\tilde{\Pi}_{\lambda^z}\tilde{\Pi}^{-1}) & tr(\tilde{\Pi}_{\lambda^z}\tilde{\Pi}^{-1}\tilde{\Pi}_{\lambda^w}\tilde{\Pi}^{-1}) \\ tr(\tilde{\Pi}_{\lambda^w}\tilde{\Pi}^{-1}\tilde{\Pi}_{\lambda^z}\tilde{\Pi}^{-1}) & tr(\tilde{\Pi}_{\lambda^w}\tilde{\Pi}^{-1}\tilde{\Pi}_{\lambda^w}\tilde{\Pi}^{-1}) \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta}, \qquad (6.56)$$

which upon substitution of Equations (6.52) and Equation (6.54) yields:

$$G_{\Lambda\Lambda} = -\frac{1}{2} \begin{bmatrix} tr(I_{n^{\tilde{\Pi}^z}}) & O \\ O & tr(I_{n^{\tilde{\Pi}^w}}) \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} n^{\tilde{\Pi}^z} & O \\ O & n^{\tilde{\Pi}^w} \end{bmatrix}, \qquad (6.57)$$

where $n^{\tilde{\Pi}^z} = m(p+1)$ and $n^{\tilde{\Pi}^w} = n(p+1)$ are the sizes of $\tilde{\Pi}^z$ and $\tilde{\Pi}^w$, respectively. From Equations (6.55) and (6.57), G_Λ and $G_{\Lambda\Lambda}$ are constants, and can be pre-computed in Algorithm 1.

6.11 Gradients of Prediction Error

As opposed to the result above, the gradients of the prediction errors are *not* constant, as is shown in this section.

6.11.1 Gradients of Prediction Error Along (Generalized) States

The error in prediction of (generalized) outputs, inputs and states is represented together by $\tilde{\epsilon}$, which makes up the precision weighted prediction error defined by $E = \frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon}\big|_{\vartheta=\mu^\vartheta}$, where $\tilde{\Pi} = diag(\tilde{\Pi}^z, P^{\tilde{v}}, \tilde{\Pi}^w)$. The error and its gradient are:

$$\tilde{\epsilon} = \begin{bmatrix} \tilde{\mathbf{y}} - \tilde{C}\tilde{x} \\ \tilde{v} - \eta^{\tilde{v}} \\ D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix}, \tilde{\epsilon}_X\big|_{\vartheta=\mu^\vartheta} = \begin{bmatrix} -\tilde{C} & O \\ O & I \\ D^x - \tilde{A} & -\tilde{B} \end{bmatrix}. \tag{6.58}$$

The gradient of prediction error with respect to X can be simplified as:

$$E_X = \tilde{\epsilon}_X^T\tilde{\Pi}\tilde{\epsilon}\big|_{\vartheta=\mu^\vartheta} = A_1 X + B_1 \begin{bmatrix} \tilde{\mathbf{y}} \\ -\eta^{\tilde{v}} \end{bmatrix}, \tag{6.59}$$

where $A_1 = \begin{bmatrix} \tilde{C}^T\tilde{\Pi}^z\tilde{C} + (D^x - \tilde{A})^T\tilde{\Pi}^w(D^x - \tilde{A}) & -(D^x - \tilde{A})^T\tilde{\Pi}^w\tilde{B} \\ -\tilde{B}^T\tilde{\Pi}^w(D^x - \tilde{A}) & P^{\tilde{v}} + \tilde{B}^T\tilde{\Pi}^w\tilde{B} \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta}$ and

$B_1 = \begin{bmatrix} -\tilde{C}^T\tilde{\Pi}^z & O \\ O & P^{\tilde{v}} \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta}$. Since E_X is linear in X, differentiating Equation (6.59) with respect to X yields a simple expression for curvature E_{XX} as:

$$E_{XX} = \tilde{\epsilon}_X^T\tilde{\Pi}\tilde{\epsilon}_X\big|_{\vartheta=\mu^\vartheta} = \begin{bmatrix} \tilde{C}^T\tilde{\Pi}^z\tilde{C} + (D^x - \tilde{A})^T\tilde{\Pi}^w(D^x - \tilde{A}) & -(D^x - \tilde{A})^T\tilde{\Pi}^w\tilde{B} \\ -\tilde{B}^T\tilde{\Pi}^w(D^x - \tilde{A}) & P^{\tilde{v}} + \tilde{B}^T\tilde{\Pi}^w\tilde{B} \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta}. \tag{6.60}$$

6.11.2 Gradients of Prediction Error Along Parameters

To evaluate the gradients of prediction error along the parameters Θ, the reformulated definition of $\tilde{\epsilon}$ is used:

$$\tilde{\epsilon} = \begin{bmatrix} \tilde{\mathbf{y}} - N\theta \\ \tilde{v} - \eta^{\tilde{v}} \\ D^x\tilde{x} - M\theta \end{bmatrix}, \tilde{\epsilon}_\Theta = (\tilde{\epsilon}_\theta)|_{\theta=\mu^\theta} = \begin{bmatrix} -N \\ O \\ -M \end{bmatrix}, \tag{6.61}$$

where M and N are given in Equation (6.5). This is to ensure that the variable Θ can be separated out of the expression for E_Θ such that it is linear in Θ as follows:

$$\begin{aligned} E_\Theta = (\tilde{\epsilon}_\theta^T\tilde{\Pi}\tilde{\epsilon})|_{\vartheta=\mu^\vartheta} &= -\begin{bmatrix} N^T & O & M^T \end{bmatrix}\tilde{\Pi}\begin{bmatrix} \tilde{\mathbf{y}} - N\theta \\ \tilde{v} - \tilde{\eta} \\ D^x\tilde{x} - M\theta \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta} \\ &= (N^T\tilde{\Pi}^z N + M^T\tilde{\Pi}^w M)\big|_{\vartheta=\mu^\vartheta}\,\Theta - \begin{bmatrix} N^T\tilde{\Pi}^z & M^T\tilde{\Pi}^w D^x \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta}\begin{bmatrix} \tilde{\mathbf{y}} \\ \mu^{\tilde{x}} \end{bmatrix} \\ &= A_2\Theta - B_2\begin{bmatrix} \tilde{\mathbf{y}} \\ \mu^{\tilde{x}} \end{bmatrix}. \end{aligned} \tag{6.62}$$

Since E_Θ is linear in Θ, differentiating Equation (6.62) with respect to Θ yields a simple expression for $E_{\Theta\Theta}$ as:

$$E_{\Theta\Theta} = (\tilde{\epsilon}_\theta^T\tilde{\Pi}\tilde{\epsilon}_\theta)\big|_{\vartheta=\mu^\vartheta} = \begin{bmatrix} N^T\tilde{\Pi}^z N + M^T\tilde{\Pi}^w M \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta} = A_2. \tag{6.63}$$

6.11.3 Gradients of Prediction Error Along Hyperparameters

The gradients of prediction error along the hyperparameters Λ is simpler, and is given by $E_\Lambda = \frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}_\lambda\tilde{\epsilon}\Big|_{\vartheta=\mu^\vartheta}$, $E_{\Lambda\Lambda} = \frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}_{\lambda\lambda}\tilde{\epsilon}\Big|_{\vartheta=\mu^\vartheta}$, which upon using $\tilde{\Pi}_{\lambda\lambda} = 0$, gives:

$$E_\Lambda = \frac{1}{2}tr(\tilde{\Pi}_\lambda\tilde{\epsilon}\tilde{\epsilon}^T)\Big|_{\vartheta=\mu^\vartheta} = \frac{1}{2}\begin{bmatrix} tr(\tilde{\Pi}^z(\tilde{\mathbf{y}}-\tilde{C}\tilde{x})(\tilde{\mathbf{y}}-\tilde{C}\tilde{x})^T) \\ tr(\tilde{\Pi}^w(D^x\tilde{x}-\tilde{A}\tilde{x}-\tilde{B}\tilde{u})(D^x\tilde{x}-\tilde{A}\tilde{x}-\tilde{B}\tilde{u})^T) \end{bmatrix}\Bigg|_{\vartheta=\mu^\vartheta} ,$$

$$E_{\Lambda\Lambda} = \frac{1}{2}\tilde{\epsilon}^T\tilde{\Pi}_{\lambda\lambda}\tilde{\epsilon}\Big|_{\vartheta=\mu^\vartheta} = 0. \tag{6.64}$$

In summary, Equations (6.59), (6.62), and (6.64) represent the gradients of the prediction error term, whereas Equations (6.60), (6.63), and (6.64) represent its curvatures. The next section will deal with evaluating the analytic expressions for all the gradients and curvatures of the mean field term.

6.12 Gradients of Mean Field Terms

This section aims to derive the analytic expressions for the mean field terms and their gradients: $W_{\mu^j}^{\mu^i}$ and $W_{\mu^j\mu^j}^{\mu^i}$, $\forall \mu \in \{X, \Theta, \Lambda\}$.

6.12.1 Gradients of mean field terms along hyperparameters

In this section, we prove that all the gradients and curvatures of W^Λ (namely $W_{\mu^i}^\Lambda$ and $W_{\mu^i\mu^i}^\Lambda$) are zeroes. The mean field term for hyperparameters Λ can be expressed as:

$$W^\Lambda = \frac{1}{2}tr\Big[\Sigma^\lambda U(y,\mu^\vartheta)_{\lambda\lambda}\Big]. \tag{6.65}$$

To compute the gradients of W^Λ, we need the curvature of internal energy with respect to λ. This can be evaluated by first differentiating Equation (6.25) with respect to λ and evaluating it at $\vartheta = \mu^\vartheta$, which yields:

$$U(y,\mu^\vartheta)_\lambda = U(y,\vartheta)_\lambda\Big|_{\vartheta=\mu^\vartheta} = -P^\lambda\epsilon^\lambda\Big|_{\vartheta=\mu^\vartheta} - \frac{1}{2}(\tilde{\epsilon}^T\tilde{\Pi}_\lambda\tilde{\epsilon})\Big|_{\vartheta=\mu^\vartheta} + G_\Lambda, \tag{6.66}$$

where G_Λ is given by Equation (6.55). Upon further differentiation, we get:

$$U(y,\mu^\vartheta)_{\lambda\lambda} = -P^\lambda - \frac{1}{2}(\tilde{\epsilon}^T\tilde{\Pi}_{\lambda\lambda}\tilde{\epsilon})\Big|_{\vartheta=\mu^\vartheta} + G_{\Lambda\Lambda}. \tag{6.67}$$

The assumption of $\tilde{\Pi}_{\lambda\lambda} = 0$ applied to Equation (6.67) yields:

$$U(y,\mu^\vartheta)_{\lambda\lambda} = -P^\lambda + G_{\Lambda\Lambda} \tag{6.68}$$

which contains only constants. Therefore, the assumption of $\tilde{\Pi}_{\lambda\lambda} = 0$ reduces all the gradients and curvatures of mean field terms of Λ to zeros:

$$W_{\mu^i}^\Lambda = \frac{1}{2}tr\Big[\Sigma^\lambda U(y,\mu^\vartheta)_{\lambda\lambda\mu^i}\Big] = 0, \;\; W_{\mu^i\mu^i}^\Lambda = 0. \tag{6.69}$$

Since the internal energy given in Equation (6.25) is quadratic in ϑ^i, and since $\tilde{\Pi}_{\lambda\lambda} = 0$, all the gradients and curvatures of the mean field term of ϑ^i with respect to itself are zeros:

$$W_{\mu^i}^{\mu^i} = \frac{1}{2}tr\Big[\Sigma^{\vartheta^i} U(y,\mu^\vartheta)_{\vartheta^i\vartheta^i\mu^i}\Big] = 0, \;\; W_{\mu^i\mu^i}^{\mu^i} = 0. \tag{6.70}$$

6.12.2 Gradients of Mean Field Terms Along Generalized States

The mean-field term of the combined generalized states X can be expressed as:

$$W^X = \frac{1}{2} tr \left[\Sigma^X U(y, \mu^\vartheta)_{XX} \right]. \tag{6.71}$$

The curvature of internal energy with respect to X can be calculated by differentiating Equation (6.25) with respect to X twice, resulting in $U(y, \mu^\vartheta)_{XX} = -\frac{1}{2} \tilde{\epsilon}_X^T \tilde{\Pi} \tilde{\epsilon}_X$. Substituting it in Equation (6.71) upon differentiation with Θ yields:

$$\begin{aligned}
W^X_{\Theta^i} &= -\frac{1}{2} tr(\Sigma^X \tilde{\epsilon}_{X\theta^i}^T \tilde{\Pi} \tilde{\epsilon}_X) \Big|_{\vartheta=\mu^\vartheta} = -\frac{1}{2} tr \left(\Sigma^X \begin{bmatrix} -\tilde{C}_{\theta^i}^T & O & -\tilde{A}_{\theta^i}^T \\ O & O & -\tilde{B}_{\theta^i}^T \end{bmatrix} \tilde{\Pi} \begin{bmatrix} -\tilde{C} & O \\ O & I \\ D^x - \tilde{A} & -\tilde{B} \end{bmatrix} \right) \Big|_{\vartheta=\mu^\vartheta} \\
&= -\frac{1}{2} tr \left[\Sigma^X \begin{bmatrix} \tilde{C}_{\theta^i}^T \tilde{\Pi}^z \tilde{C} - \tilde{A}_{\theta^i}^T \tilde{\Pi}^w (D^x - \tilde{A}) & \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{B} \\ -\tilde{B}_{\theta^i}^T \tilde{\Pi}^w (D^x - \tilde{A}) & \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{B} \end{bmatrix} \right] \Big|_{\vartheta=\mu^\vartheta} ,
\end{aligned} \tag{6.72}$$

where the gradient of the mean field with respect to Θ is given by $W^X_\Theta = \begin{bmatrix} W^X_{\Theta^1} & W^X_{\Theta^2} & ... & W^X_{\Theta^i} \end{bmatrix}^T$. Similarly, the elements of the curvature matrix of the mean field term with respect to Θ is given by:

$$W^X_{\Theta^i \Theta^j} = -\frac{1}{2} tr(\Sigma^X \tilde{\epsilon}_{X\theta^i}^T \tilde{\Pi} \tilde{\epsilon}_{X\theta^j}) \Big|_{\vartheta=\mu^\vartheta} = -\frac{1}{2} tr \left[\Sigma^X \begin{bmatrix} \tilde{C}_{\theta^i}^T \tilde{\Pi}^z \tilde{C}_{\theta^j} + \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{A}_{\theta^j} & \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{B}_{\theta^j} \\ \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{A}_{\theta^j} & \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{B}_{\theta^j} \end{bmatrix} \right] \Big|_{\vartheta=\mu^\vartheta} . \tag{6.73}$$

The gradient and curvature of the mean field term of X with respect to Λ can be evaluated as:

$$\begin{aligned}
W^X_\Lambda &= -\frac{1}{2} tr(\Sigma^X \tilde{\epsilon}_X^T \tilde{\Pi}_\lambda \tilde{\epsilon}_X) \Big|_{\vartheta=\mu^\vartheta} = -\frac{1}{2} tr(\tilde{\Pi}_\lambda \tilde{\epsilon}_X \Sigma^X \tilde{\epsilon}_X^T) \Big|_{\vartheta=\mu^\vartheta} \\
&= -\frac{1}{2} \begin{bmatrix} tr(\tilde{\Pi}^z \tilde{C} \Sigma^x \tilde{C}^T) \\ tr(\tilde{\Pi}^w \begin{bmatrix} (D^x - \tilde{A}), & -\tilde{B} \end{bmatrix} \Sigma^X \begin{bmatrix} (D^x-\tilde{A})^T \\ -\tilde{B}^T \end{bmatrix}) \end{bmatrix} \Big|_{\vartheta=\mu^\vartheta} \\
W^X_{\Lambda\Lambda} &= -\frac{1}{2} tr(\tilde{\Pi}_{\lambda\lambda} \tilde{\epsilon}_X \Sigma^X \tilde{\epsilon}_X^T) \Big|_{\vartheta=\mu^\vartheta} = 0,
\end{aligned} \tag{6.74}$$

where Σ^x is a component of $\Sigma^X = \begin{bmatrix} \Sigma^{\tilde{x}} & \Sigma^{\tilde{x}\tilde{v}} \\ \Sigma^{\tilde{x}\tilde{v}} & \Sigma^{\tilde{v}} \end{bmatrix}$. Here the curvature $W^X_{\Lambda\Lambda}$ vanishes due to the assumption that $\tilde{\Pi}_{\lambda\lambda} = 0$.

6.12.3 Gradients of Mean Field Terms Along Parameters

The mean-field term of the parameters Θ can be expressed as

$$W^\Theta = \frac{1}{2} tr \left[\Sigma^\theta U(y, \mu^\vartheta)_{\theta\theta} \right] = \frac{1}{2} tr \left[\Sigma^\theta (-P^\theta - \tilde{\epsilon}_\theta^T \tilde{\Pi} \tilde{\epsilon}_\theta) \right] \Big|_{\vartheta=\mu^\vartheta} . \tag{6.75}$$

Differentiating Equation (6.75) with X and substituting Equation (6.61) in it yields the gradient as:

$$W^\Theta_X = -\frac{1}{2} \begin{bmatrix} tr \left(\Sigma^\theta \left(N_{X^1}^T \tilde{\Pi}^z N + M_{X^1}^T \tilde{\Pi}^w M \right) \right) \\ tr \left(\Sigma^\theta \left(N_{X^2}^T \tilde{\Pi}^z N + M_{X^2}^T \tilde{\Pi}^w M \right) \right) \\ ... \end{bmatrix} \Big|_{\vartheta=\mu^\vartheta} , \tag{6.76}$$

and the elements of the curvature matrix as:

$$W_{X^i X^j}^{\Theta} = -\frac{1}{2} tr\Big(\Sigma^\theta \big(N_{X^i}^T \tilde\Pi^z N_{X^j} + M_{X^i}^T \tilde\Pi^w M_{X^j}^T\big)\Big)\Big|_{\vartheta=\mu^\vartheta} . \qquad (6.77)$$

Differentiating Equation (6.75) with respect to λ yields the gradient and curvature as:

$$
\begin{aligned}
W_{\Lambda}^{\Theta} &= -\frac{1}{2} tr(\Sigma^\theta \tilde\epsilon_\theta^T \tilde\Pi_\lambda \tilde\epsilon_\theta)\Big|_{\vartheta=\mu^\vartheta} = -\frac{1}{2} tr(\tilde\Pi_\lambda \tilde\epsilon_\theta \Sigma^\theta \tilde\epsilon_\theta^T)\Big|_{\vartheta=\mu^\vartheta} , \\
&= -\frac{1}{2}\begin{bmatrix} tr(\tilde\Pi_{\lambda^z} \tilde\epsilon_\theta \Sigma^\theta \tilde\epsilon_\theta^T) \\ tr(\tilde\Pi_{\lambda^w} \tilde\epsilon_\theta \Sigma^\theta \tilde\epsilon_\theta^T) \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta} = -\frac{1}{2}\begin{bmatrix} tr(\tilde\Pi^z N \Sigma^\theta N^T) \\ tr(\tilde\Pi^w M \Sigma^\theta M^T) \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta} \\
W_{\Lambda\Lambda}^{\Theta} &= -\frac{1}{2} tr(\tilde\Pi_{\lambda\lambda} \tilde\epsilon_\theta \Sigma^\theta \tilde\epsilon_\theta^T)\Big|_{\vartheta=\mu^\vartheta} = 0.
\end{aligned}
\qquad (6.78)
$$

Here, $W_{\Lambda\Lambda}^{\Theta}$ vanishes due to the assumption that $\tilde\Pi_{\lambda\lambda} = 0$.

6.13 The Complete DEM Algorithm

By combining the gradients found from Sections 6.10, 6.11, and 6.12 with the Algorithm 1, we can finalize the full DEM algorithm so that it can iteratively compute the estimates and the associated precisions from data.

6.13.1 DEM Estimates

The main equations that are required to perform the update rules of DEM given in Equation (6.48) can be summarized as:

$$
\begin{aligned}
\dot{X} &= DX + k^X(-E_X + W_X^\Theta),\ J^X = D + k^X(-E_{XX} + W_{XX}^\Theta) \\
\frac{\partial\Theta}{\partial a} &= k^\Theta\Big[-P^\theta\epsilon^\theta\Big|_{\vartheta=\mu^\vartheta} + \sum_t\Big(-E_\Theta + W_\Theta^X\Big)\Big],\ J^\Theta = k^\Theta\Big[-P^\theta + \sum_t\Big(-E_{\Theta\Theta} + W_{\Theta\Theta}^X\Big)\Big] \\
\frac{\partial\Lambda}{\partial b} &= k^\Lambda\Big[-P^\lambda\epsilon^\lambda\Big|_{\vartheta=\mu^\vartheta} + \sum_t(-E_\Lambda + W_\Lambda^X + W_\Lambda^\Theta + G_\Lambda)\Big],\ J^\Lambda = k^\Lambda(-P^\lambda + n^t G_{\Lambda\Lambda})
\end{aligned}
\qquad (6.79)
$$

where $E_X, E_{XX}, E_\Theta, E_{\Theta\Theta}, E_\Lambda, W_X^\Theta, W_{XX}^\Theta, W_\Theta^X, W_{\Theta\Theta}^X, W_\Lambda^X, W_\Lambda^\Theta, G_\Lambda, and\ G_{\Lambda\Lambda}$ are given by Equations (6.59), (6.60), (6.62), (6.63), (6.64), (6.76), (6.77), (6.72), (6.73), (6.74), (6.78), (6.55), and (6.57), respectively. The hyperparameter update rule can be further simplified to reduce the computational complexity as:

$$\frac{\partial\Lambda}{\partial b} = -k^\Lambda P^\lambda \epsilon^\lambda\Big|_{\vartheta=\mu^\vartheta} + \frac{k^\Lambda n^t}{2}\begin{bmatrix} n^{\tilde\Pi^z} \\ n^{\tilde\Pi^w} \end{bmatrix} - \frac{k^\Lambda}{2}\begin{bmatrix} tr(\tilde\Pi^z A_3) \\ tr(\tilde\Pi^w B_3) \end{bmatrix}\Big|_{\vartheta=\mu^\vartheta} \qquad (6.80)$$

where

$$
\begin{aligned}
A_3 &= \sum_t \Big((\tilde{\mathbf{y}} - \tilde{C}\tilde{x})(\tilde{\mathbf{y}} - \tilde{C}\tilde{x})^T + N\Sigma^\theta N^T + \tilde{C}\Sigma^{\tilde{x}\tilde{x}}\tilde{C}^T\Big) \\
B_3 &= \sum_t \Big((D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{u})(D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{u})^T + M\Sigma^\theta M^T + \big[(D^x-\tilde{A}),\ -\tilde{B}\big]\Sigma^X\begin{bmatrix}(D^x-\tilde{A})^T \\ -\tilde{B}^T\end{bmatrix}\Big).
\end{aligned}
$$

Substituting Equation (6.57) to the expression for J^Λ in Equation (6.79) yields:

$$J^\Lambda = -k^\Lambda(P^\lambda + \frac{n^t}{2} \begin{bmatrix} n^{\tilde{\Pi}^z} & O \\ O & n^{\tilde{\Pi}^w} \end{bmatrix}), \tag{6.81}$$

which is independent of Λ. This reduces the algorithm's computational complexity, as J^Λ can now be pre-computed.

6.13.2 Precision of Estimates

This section simplifies the precision for DEM's estimates for an LTI system. The confidence in the estimate of (generalized) states and inputs can be simplified using Equations (6.51) and (6.60) as:

$$\Pi^{\tilde{x},\tilde{v}} = E_{XX} = \begin{bmatrix} \tilde{C}^T\tilde{\Pi}^z\tilde{C} + (D^x - \tilde{A})^T\tilde{\Pi}^w(D^x - \tilde{A}) & -(D^x - \tilde{A})^T\tilde{\Pi}^w\tilde{B} \\ -\tilde{B}^T\tilde{\Pi}^w(D^x - \tilde{A}) & P^{\tilde{v}} + \tilde{B}^T\tilde{\Pi}^w\tilde{B} \end{bmatrix}\bigg|_{\vartheta=\mu^\vartheta}. \tag{6.82}$$

From Equation (6.82), the precisions for state and input estimation are $\Pi^{\tilde{x}\tilde{x}} = \tilde{C}^T\tilde{\Pi}^z\tilde{C} + (D^x - \tilde{A})^T\tilde{\Pi}^w(D^x - \tilde{A})$ and $\Pi^{\tilde{v}\tilde{v}} = P^{\tilde{v}} + \tilde{B}^T\tilde{\Pi}^w\tilde{B}$, respectively. The cross-correlation between the (generalized) states and inputs are given by $-\tilde{B}^T\tilde{\Pi}^w(D^x - \tilde{A})$. Since $\Pi^{\tilde{x},\tilde{v}}$ is independent of X, it can be updated outside the D step.

Combining the results of Equations (6.51) and (6.63) yields the precision of parameter estimates Π^θ, which is independent of Θ, as:

$$\Pi^\theta = P^\theta + \sum_t E_{\Theta\Theta} = P^\theta + \sum_t \left[N^T\tilde{\Pi}^z N + M^T\tilde{\Pi}^w M\right]\bigg|_{\vartheta=\mu^\vartheta}. \tag{6.83}$$

From Equations (6.51), (6.64), and (6.57), the precision of hyperparameter estimation is:

$$\Pi^\lambda = P^\lambda + \sum_t \left(E_{\Lambda\Lambda} - G_{\Lambda\Lambda}\right) = P^\lambda + \frac{n^t}{2}diag(n^{\tilde{\Pi}^z}, n^{\tilde{\Pi}^w}), \tag{6.84}$$

which is a constant and hence is never updated in the algorithm. In conclusion, the estimation using Equation (6.79), along with the precision of these estimates given by Equations (6.82), (6.83), and (6.84) completely define the DEM algorithm for an LTI system with colored noises. The complete DEM algorithm is given in Algorithm 2.

Algorithm 2 Dynamics Expectation Maximization

Data: Time series data of output $\mathbf{y}$
Result: $(X, \Theta, \Lambda), (\Pi^{\tilde{x}, \tilde{v}}, \Pi^{\theta}, \Pi^{\lambda})$
$a = 0$
  Evaluate $\tilde{\mathbf{y}}$ from $\mathbf{y}$ ▷ Equation (6.9)
  Precompute gradients $\tilde{A}_{\theta^i}, \tilde{B}_{\theta^i}, \tilde{C}_{\theta^i}, N_{X^i}, M_{X^i}$
  Set all priors: $\eta^{\tilde{v}}, P^{\tilde{v}}, \eta^{\theta}, P^{\theta}, \eta^{\lambda}, P^{\lambda}$
  Initialize $\Theta, \Lambda, \Pi^X, \Pi^\theta, \Pi^\lambda$ with their priors
  $G_\Lambda, G_{\Lambda\Lambda} \leftarrow$ Equation (6.55) and (6.57)
  $\Pi^\lambda \leftarrow P^\lambda - n^t G_{\Lambda\Lambda}$
  $J^\Lambda \leftarrow k^\Lambda \left(-P^\lambda + n^t G_{\Lambda\Lambda} \right)$
  while Θ *not converged* **do**   ▷ E step loop
    Set $E_\Theta, E_{\Theta\Theta}, E_1$ to zero
    $W_{XX}^\Theta \leftarrow find W_{XX}^\Theta(N_X, M_X, \tilde{\Pi}, \Pi^\theta)$;   ▷ Equation (6.77)
    $E_{XX} \leftarrow find E_{XX}(A, B, C, \tilde{\Pi})$;   ▷ Equation (6.60)
    $J^X \leftarrow D + k^X(-E_{XX} + W_{XX}^\Theta)$
    for $t = 0:\Delta t:T$ **do**   ▷ D step loop
      // Perform D step
      $W_X^\Theta \leftarrow$ Equation (6.76), $E_X \leftarrow$ Equation (6.59)
      $\frac{\partial X}{\partial t} \leftarrow DX + k^X(-E_X + W_X^\Theta)$
      $X_{t+\Delta t} \leftarrow X_t + (e^{J^X \Delta t} - I)(J^X)^{-1} \frac{\partial X}{\partial t}$
      // Accumulate terms for E step
      $M, N \leftarrow$ Equation (6.5);   ▷ use new X
      $A_2, B_2 \leftarrow$ Equation (6.62)
      $E_\Theta \leftarrow E_\Theta + A_2\Theta - B_2 \begin{bmatrix} \tilde{\mathbf{y}} \\ \mu^{\tilde{x}} \end{bmatrix}$
      $E_{\Theta\Theta} \leftarrow E_{\Theta\Theta} + A_2$
      // Accumulate terms for M step
      $\tilde{\epsilon}, \tilde{\epsilon}_X, \tilde{\epsilon}_\Theta \leftarrow$ Equation (6.58), (6.61)
      $E_1 \leftarrow E_1 + \tilde{\epsilon}\tilde{\epsilon}^T + \tilde{\epsilon}_X \Sigma^X \tilde{\epsilon}_X^T + \tilde{\epsilon}_\Theta \Sigma^\theta \tilde{\epsilon}_\Theta^T$
    end
    while Λ *not converged* **do**   ▷ M step loop
      $\frac{\partial \Lambda}{\partial b} \leftarrow k^\Lambda \left(-P^\lambda \epsilon^\lambda + G_\Lambda - \frac{1}{2} \begin{bmatrix} tr(\tilde{\Pi}_{\lambda z} E_1) \\ tr(\tilde{\Pi}_{\lambda w} E_1) \end{bmatrix} \right)$
      $\Lambda \leftarrow \Lambda + (e^{J^\Lambda} - I)(J^\Lambda)^{-1} \frac{\partial \Lambda}{\partial b}$
      $\tilde{\Pi} \leftarrow$ Equation (6.52) ;   ▷ update with new Λ
    end
    $\bar{F}_a \leftarrow find\bar{F}(X, \Pi^X, \Theta, \Pi^\theta, \Lambda, \Pi^\lambda)$;   ▷ Equation (6.43)
    if $\bar{F}_a > \bar{F}_{a-1}$ **then**   ▷ update Θ if $\bar{F}$ increased
      $W_\Theta^X, W_{\Theta\Theta}^X \leftarrow$ Equation (6.72), (6.73)
      $\frac{\partial \Theta}{\partial a} \leftarrow k^\Theta[-P^\theta \epsilon^\Theta - E_\Theta + n^t W_\Theta^X]$
      $J^\Theta \leftarrow k^\Theta[-P^\theta - E_{\Theta\Theta} + n^t W_{\Theta\Theta}^X]$
      $\Theta \leftarrow \Theta + (e^{J^\Theta} - I)(J^\Theta)^{-1} \frac{\partial \Theta}{\partial a}$;   ▷ E step
      $A, B, C \leftarrow$ Equation (6.4) ;   ▷ update with new Θ
    end
    $a{+}{+}$
    $\Pi^{\tilde{x}, \tilde{v}} = E_{XX}$;   ▷ update generalized state precision
    $\Pi^\theta = P^\theta + E_{\Theta\Theta}$;   ▷ update parameter precision
end

6.14 Translation into Simplified Mathematical Form

Although the pseudocode derived in the previous sections is sufficient to replicate the DEM algorithm for an LTI system with colored noise, it is not sufficient to analyze DEM using the standard control systems tools for stability checks, convergence, etc. Therefore, in this section we translate the algorithm into a simplified mathematical form that control engineers can easily analyze. The following subsections aim at converting the DEM updates into a coupled linear system.

6.14.1 State and Input Estimation as a Linear Observer

This section deals with reformulating the D step of DEM for an LTI system as a (generalized) state and input observer. Substituting Equation (6.59) in Equation (6.79) with a learning rate of $k^X = 1$ yields [46]:

$$\dot{X} = \begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{v}} \end{bmatrix} = (D - A_1)X - B_1 \begin{bmatrix} \tilde{\mathbf{y}} \\ -\eta^{\tilde{v}} \end{bmatrix} + W_X^\Theta. \tag{6.85}$$

We now aim to mathematically prove that the (generalized) state and input observer of DEM can be reduced into an augmented LTI system, for which an exact discretization can be performed. We proceed by simplifying the mean field terms in Equation (6.39) as:

$$\begin{aligned} W_X^\Theta &= -\frac{1}{2} \begin{bmatrix} tr\left(\Sigma^\theta(N_{X^1}^T \tilde{\Pi}^z N + M_{X^1}^T \tilde{\Pi}^w M)\right) \\ tr\left(\Sigma^\theta(N_{X^2}^T \tilde{\Pi}^z N + M_{X^2}^T \tilde{\Pi}^w M)\right) \\ ... \end{bmatrix} \Bigg|_{\vartheta = \mu^\vartheta} \\ &= -\frac{1}{2} \begin{bmatrix} vec(\Sigma^{\theta T})^T (I \otimes (N_{X^1}^T \tilde{\Pi}^z) vec(N)) + \\ vec(\Sigma^{\theta T})^T (I \otimes (M_{X^1}^T \tilde{\Pi}^w) vec(M)) \\ ... \end{bmatrix} \Bigg|_{\vartheta = \mu^\vartheta} \\ &= (Z_X^N vec(N) + Z_X^M vec(M)) \Big|_{\vartheta = \mu^\vartheta}, \end{aligned} \tag{6.86}$$

where,

$$Z_X^N = -\frac{1}{2}(I \otimes vec(\Sigma^{\theta T})^T) \begin{bmatrix} I \otimes (N_{X^1}^T \tilde{\Pi}^z) \\ I \otimes (N_{X^2}^T \tilde{\Pi}^z) \\ ... \end{bmatrix}, \; Z_X^M = -\frac{1}{2}(I \otimes vec(\Sigma^{\theta T})^T) \begin{bmatrix} I \otimes (M_{X^1}^T \tilde{\Pi}^w) \\ I \otimes (M_{X^2}^T \tilde{\Pi}^w) \\ ... \end{bmatrix}. \tag{6.87}$$

Since M and N can be obtained from linear transformation of X, $vec(M)$ and $vec(N)$ can be written as:

$$vec(M) = Z^M X \text{ and } vec(N) = Z^N X, \tag{6.88}$$

where Z^M and Z^N are matrices with elements 0 and 1. This leads to the mean field term being expressed as a linear transformation of X:

$$W_X^\Theta = (Z_X^N Z^N + Z_X^M Z^M)\Big|_{\vartheta = \mu^\vartheta} X. \tag{6.89}$$

Substituting Equation (6.89) into Equation (6.85) simplifies the observer as:

$$\begin{aligned} \dot{X} &= A_4 X + B_4 \begin{bmatrix} \tilde{\mathbf{y}} \\ -\eta^{\tilde{v}} \end{bmatrix}, \\ A_4 &= D - A_1 + (Z_X^N Z^N + Z_X^M Z^M)\Big|_{\vartheta = \mu^\vartheta} \text{ and } B_4 = -B_1. \end{aligned} \tag{6.90}$$

The (generalized) state and input observer given by Equation (6.90) is of the form of an augmented LTI system. Therefore, an exact discretization can be used to solve it without using the second order gradient J^X as given in Equation (6.48). This reduces the algorithm's computational complexity because E_{XX} and W_{XX}^Θ for J^X calculation are no longer necessary. Figure 6.3 shows the simplified control diagram of the observer. The stability condition of this observer (under known θ and λ) and its similarity with the Kalman Filter is discussed in our prior work [46]. To evaluate W_X^Θ, one could either use Equation (6.89) or Equation (6.76). Equation (6.90) is derived mainly for simplification and exact discretization.
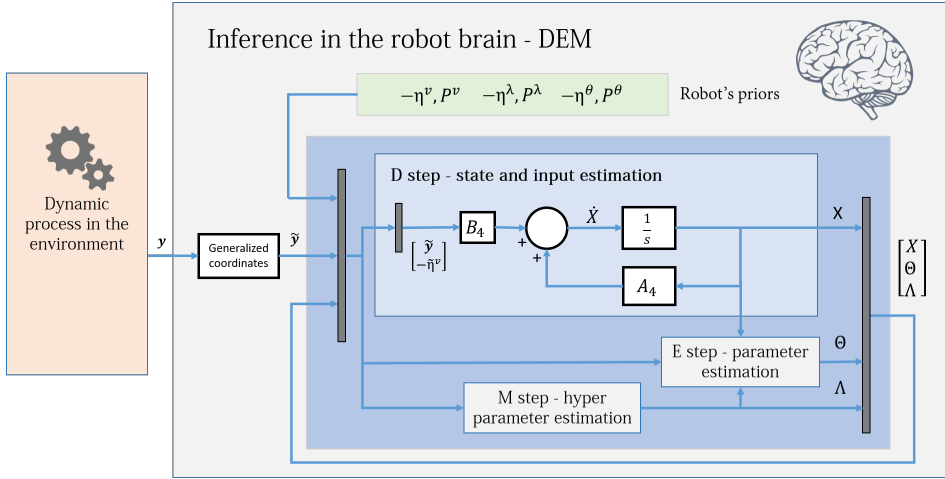
Fig 6.3: The DEM algorithm for an LTI system, with the D step simplified as an augmented LTI system given by Equation (6.90). The D-step block corresponds to the D-step loop in Algorithm 2 and operates at a different frequency from the E and M blocks.

6.14.2 Parameter Estimation—System Identification

This section aims to mathematically prove that the E step can be reduced to an augmented LTI system, for which an exact discretization can be performed. We proceed by first simplifying the parameter update equation given in Equation (6.79):

$$\frac{\partial \Theta}{\partial a} = -P^\theta(\Theta - \eta^\theta)\Big|_{\vartheta=\mu^\vartheta} + \sum_t (-E_\Theta + W_\Theta^X). \tag{6.91}$$

Grouping all $W_{\Theta i}^X$ using Equation (6.72) yields:

$$W_\Theta^X = -\frac{1}{2}\left(I \otimes vec(\Sigma^{XT})^T\right)\begin{bmatrix} I \otimes \tilde{\epsilon}_{X\theta 1}^T \tilde{\Pi} \\ I \otimes \tilde{\epsilon}_{X\theta 2}^T \tilde{\Pi} \\ \dots \end{bmatrix} vec(\tilde{\epsilon}_X)\Big|_{\vartheta=\mu^\vartheta} = Z_\theta^\epsilon vec(\tilde{\epsilon}_X)\Big|_{\vartheta=\mu^\vartheta} \tag{6.92}$$

where

$$\tilde{\epsilon}_{X\theta i}^T \tilde{\Pi} = -\begin{bmatrix} \tilde{C}_{\theta i}^T \tilde{\Pi}^z & O & \tilde{A}_{\theta i}^T \tilde{\Pi}^w \\ O & O & \tilde{B}_{\theta i}^T \tilde{\Pi}^w \end{bmatrix}, vec(\tilde{\epsilon}_X) = -vec\begin{bmatrix} \tilde{C} & O \\ O & O \\ \tilde{A} & \tilde{B} \end{bmatrix} + vec\begin{bmatrix} O & O \\ O & I \\ D^x & O \end{bmatrix} = -Z^\theta \theta + Z^I. \tag{6.93}$$

Here Z^θ and Z^I are constant matrices with elements 0 and 1. Substituting $vec(\tilde{\epsilon}_X)$ from Equation (6.93) in Equation (6.92) yields:

$$W_\Theta^X = -(Z_\theta^\epsilon Z^\theta)|_{\vartheta=\mu^\vartheta}\Theta + (Z_\theta^\epsilon Z^I)|_{\vartheta=\mu^\vartheta}. \tag{6.94}$$

Substituting Equations (6.94) and (6.62) in Equation (6.91), simplifies the parameter update equation to

$$\frac{\partial\Theta}{\partial a} = A_5\Theta + B_5,$$
$$A_5 = -[P^\theta + n^t(Z_\theta^\epsilon Z^\theta)|_{\vartheta=\mu^\vartheta} + \sum_t A_2], \; B_5 = P^\theta\eta^\theta + n^t(Z_\theta^\epsilon Z^I)|_{\vartheta=\mu^\vartheta} + \sum_t B_2\begin{bmatrix}\tilde{\mathbf{y}}\\\mu^{\tilde{x}}\end{bmatrix}, \tag{6.95}$$

where A_2 and B_2 are given in Equation (6.62). Equation (6.95) is a linear differential equation in Θ for which an exact discretization can be computed. For each Θ update in Algorithm 2, A_2 and B_2 are also updated, consequently updating A_5 and B_5. Therefore, Equation (6.95) is equivalent to a linear time-*varying* system. Figure 6.4 shows the simplified parameter estimation step of the robot brain. To evaluate W_Θ^X, one could use either Equation (6.72) or Equation (6.94). Equation (6.94) was derived mainly for the exact discretization and for the convergence proof in Section 6.15.
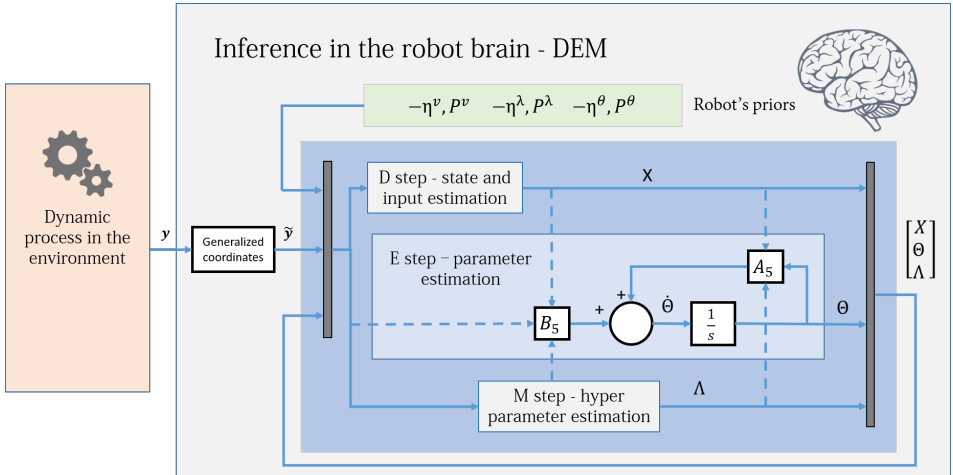
Fig 6.4: The DEM algorithm for an LTI system, with the E step simplified as an augmented LTI system given by Equation (6.95). The E-step block corresponds to the E-step outer loop in Algorithm 2 and operates at a different frequency when compared to the D and M blocks. The dotted lines illustrate the flow of variables from other blocks and demonstrate the coupled nature of D, E, and M steps. This diagram is illustrative and should not be confused with a control diagram.

6.14.3 HYPERPARAMETER UPDATE

The update equation in Equation (6.80) can be simplified as:

$$
\begin{aligned}
\frac{\partial \Lambda}{\partial b} &= k^\Lambda \left[-P^\lambda \epsilon^\lambda + \frac{n^t}{2} \begin{bmatrix} n^{\tilde{\Pi}z} \\ n^{\tilde{\Pi}w} \end{bmatrix} - \frac{1}{2} \begin{bmatrix} e^{\lambda z} tr((S \otimes \Omega^z)A_3) \\ e^{\lambda w} tr((S \otimes \Omega^w)B_3) \end{bmatrix} \right] \Big|_{\vartheta = \mu^\vartheta} \\
&= -\frac{k^\Lambda}{2} \begin{bmatrix} tr((S \otimes \Omega^z)A_3) & O \\ O & tr((S \otimes \Omega^w)B_3) \end{bmatrix} e^\Lambda - (k^\Lambda P^\lambda)\Lambda + (k^\Lambda P^\lambda \eta^\lambda + \frac{k^\Lambda n^t}{2} \begin{bmatrix} n^{\tilde{\Pi}z} \\ n^{\tilde{\Pi}w} \end{bmatrix}) \quad (6.96) \\
&= a_1 e^\Lambda + a_2 \Lambda + a_3,
\end{aligned}
$$

where a_1, a_2, and a_3 are constants that are independent of Λ. Since Equation (6.96) is *nonlinear* in Λ, an approximate discretization like the conventional Gauss–Newton update scheme given in Equation (6.48) should be used for the M step. In summary, the D and E steps follow an exact discretization, whereas the M step follows an approximate discretization.

6.15 CONVERGENCE PROOF FOR PARAMETER AND HYPERPARAMETER ESTIMATION

In robotics, it is important that learning algorithms provide a stable solution, especially when robot safety during operation is a concern. Therefore, a proof of convergence for DEM is important for its widespread use in robotics as a learning algorithm. However, the DEM literature lacks any such mathematical proof of convergence for the estimator. Therefore, this section aims at providing one for the parameter and hyperparameter estimation step on LTI systems.

Since the update equation given by Equation (6.95) is a linear differential equation, proving that $A_5 \prec O$ is sufficient to prove that Θ converges to a stable solution. Substituting the expression for A_2 from Equation (6.62) to the A_5 in Equation (6.95), yields:

$$
A_5 = -[P^\theta + n^t Z^\epsilon_\theta Z^\theta + \sum_t (N^T \tilde{\Pi}^z N + M^T \tilde{\Pi}^w M)]\Big|_{\vartheta = \mu^\vartheta} . \quad (6.97)
$$

Since the prior precision matrix can be chosen to be positive definite, $P^\theta \succ O$. It is straightforward to note from the expression for A_2 in Equation (6.62) that $\sum_t A_2 \succ O$, because $\tilde{\Pi}^z \succ O, \tilde{\Pi}^w \succ O \implies N^T \tilde{\Pi}^z N \succ O$, and $M^T \tilde{\Pi}^w M \succ O$. Therefore, the proof of convergence is complete if we prove that $Z^\epsilon_\theta Z^\theta \succ O$. Simplifying the expressions for Z^ϵ_θ and Z^θ from Equations (6.92) and (6.93), after some nontrivial linear algebra [167], yields:

$$
Z^\epsilon_\theta Z^\theta = \frac{1}{2} \frac{\partial \tilde{\theta}}{\partial \theta}^T Z_1 Z_2 \frac{\partial \tilde{\theta}}{\partial \theta} \quad (6.98)
$$

where

$$
Z_1 = diag(\tilde{\Pi}^w \otimes I, \tilde{\Pi}^w \otimes I, \tilde{\Pi}^z \otimes I), \quad \frac{\partial \tilde{\theta}}{\partial \theta} = diag(vec\tilde{A}^T_{vecA^T}, vec\tilde{B}^T_{vecB^T}, vec\tilde{C}^T_{vecC^T}),
$$

$$
Z_2 = \begin{bmatrix} I \otimes \Sigma^{\tilde{x}\tilde{x}T} & I \otimes \Sigma^{\tilde{v}\tilde{x}T} & O \\ I \otimes \Sigma^{\tilde{x}\tilde{v}T} & I \otimes \Sigma^{\tilde{v}\tilde{v}T} & O \\ O & O & I \otimes \Sigma^{\tilde{x}\tilde{x}T} \end{bmatrix}.
$$

It is straightforward from Equation (6.98) that $Z_\theta^\epsilon Z^\theta > O$ because $Z_1 > O$, and $Z_2 > O$. Combining all the results from this section, $P^\theta > O$, $Z_\theta^\epsilon Z^\theta > O$ and $\sum_t A_2 > O$, $\implies A_5 < O$. This completes the proof that the parameter estimation step of DEM converges for an LTI system. Similarly, from Equation (6.81), $J^\Lambda < O$ proves the convergence of hyperparameter estimation step. For a detailed account of the linear algebra behind the proof of convergence, readers may refer to [167].

6.16 A Demonstrative Example

This section aims to provide the proof of concept for DEM through simulation for the estimation of an LTI system with colored noise. Since the algorithm can find an infinite number of solutions for a black box estimation of $\tilde{x}$, $\tilde{v}$, θ and λ from **y**, a black box estimation is not ideal as a demonstrative example. Therefore, we restrict this section to the joint estimation of x, A, B, Π^w, and Π^z from known **y** and C.

6.16.1 Generative Model

A stable LTI system of the form Equation (6.3) was selected, with randomly generated parameters $\theta^i \in [-1, 1]$ having

$$A = \begin{bmatrix} 0.0484 & 0.7535 \\ -0.7617 & -0.2187 \end{bmatrix}, B = \begin{bmatrix} 0.3604 \\ 0.0776 \end{bmatrix}, C = \begin{bmatrix} 0.2265 & -0.4786 \\ 0.4066 & -0.2641 \\ 0.3871 & 0.3817 \\ -0.1630 & -0.9290 \end{bmatrix}.$$

A Gaussian bump input signal of $v = e^{-0.25(t-12)^2}$ was centered around $t = 12s$ and sampled at $dt = 0.1s$ till $T = 32s$ was used. The colored noise was generated with a smoothness value of $\sigma = 0.5$ for the Gaussian kernel. The noise precisions were $\Pi^w = e^8 I_2$ and $\Pi^z = e^8 I_4$, making $\lambda^z = \lambda^w = 8$. The embedding order of the generalized motion of states and inputs were $p = 6$ and $d = 2$, respectively.

6.16.2 Priors for Estimation

As discussed in Section 6.6.2, three prior distributions are necessary for the algorithm. Since the inputs are known, the input prior η^v is initialized with the known input v, and a tight prior precision of $P^v = e^{32} I_1$ is used to restrict any changes in v. Similarly, since the parameter C is known, the corresponding prior parameters in η^θ are initialized with C, with tight priors of $P^{\theta^i} = e^{32}$. The prior parameters η^{θ^i} for the unknown A and B matrices are randomly sampled from the range of [-2,2], and a low prior precision of $P^{\theta^i} = e^6$ is used to encourage exploratory behavior. In summary, $\eta^\theta = \begin{bmatrix} vec(rand(2,2)^T)^T & vec(rand(2,1)^T)^T & vec(C^T)^T \end{bmatrix}^T$ and $P^\theta = diag(e^6 I_4, e^6 I_2, e^{32} I_8)$. Since the hyperparameters are unknown, their priors were set to zero $\lambda = [0\ 0]^T$, with a prior precision of $P^\lambda = e^3 I_2$ to encourage exploration.

6.16.3 Results of Estimation

The data **y** generated from the system in Section 6.16.1 was used to run the DEM algorithm given in Algorithm 2. Figure 6.5a demonstrates the successful state estimation of the

algorithm. The results of parameter estimation (A and B) are shown in Figure 6.5b. The updates began from randomly selected priors η^θ, marked by red circles, to finally converge. Table 6.1 shows that the DEM's estimate of A and B are close to the real values.

Table 6.1: DEM's estimate of A and B converges to real value.

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Real	0.048	0.753	-0.761	-0.218	0.360	0.077
Estimate	0.034	0.714	-0.769	-0.219	0.333	0.098

This confirms that the parameter estimation can converge close to the real parameters, even when η^θ is randomly selected from the range $\eta^\theta \in [-2, 2]$ that is double the size of the real parameter range $\theta^i \in [-1, 1]$. Figure 6.5c shows the successful hyperparameter convergence close to $\lambda^z = \lambda^w = 8$.
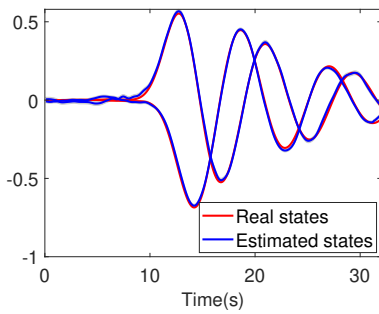
(a) State estimation.
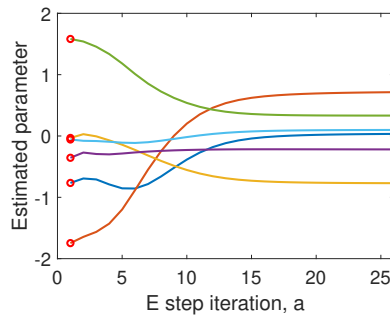
(b) Parameter estimation.
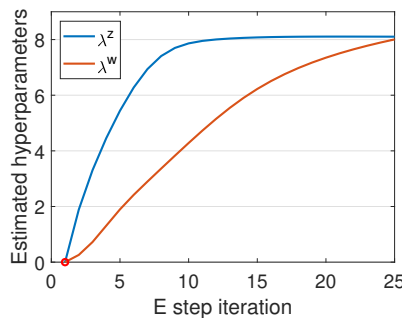
(c) Hyperparameter estimation.

Fig 6.5: The results of DEM's estimation process. (a) The estimated states in blue closely resembles the real states in red. (b) The parameter estimation starts from randomly selected η^θ, marked by red circles and converges with each E step iteration a. (c) Both the hyperparameters start from $\eta^\lambda = 0$, and converge close to the correct value of 8.

DEM's confidence on its estimates increase with the E step iterations, as can be seen from Figure 6.6a, which demonstrates an increase in parameter precision Π^θ. A similar trend can be observed for Π^X. However, Π^λ remains a constant during the entire algorithm, as proved in Section 6.13.2. The key idea behind DEM's inference is the maximization of free energy objectives. Read together, Figure 6.5 and 6.6 demonstrates that DEM successfully estimates $\tilde{x}$, θ and λ, with increasing confidence on its estimates as the estimation proceeds by maximizing $\bar{F}$ from Equation (6.40). In summary, DEM can be used for the joint estimation of states, parameters and hyperparameters of an LTI system, subjected to colored noise.

(a) Parameter precision Π^θ.

(b) Free action $\bar{F}(a) - \bar{F}(0)$.
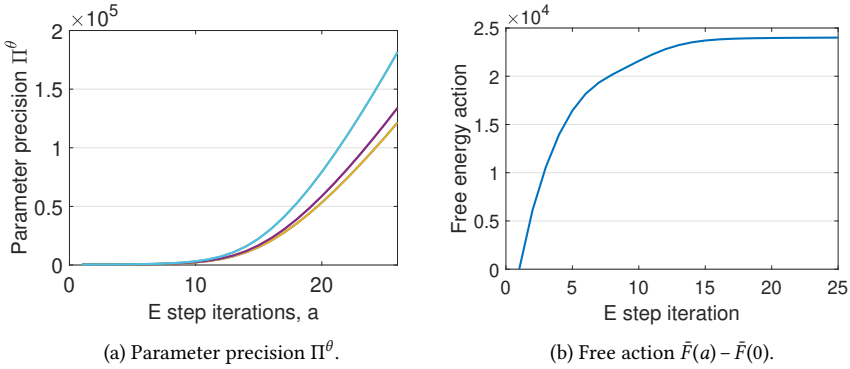
Fig 6.6: Maximization of $\bar{F}$ improves the confidence on estimates.

6.17 Benchmarking
This section deals with benchmarking DEM against the state-of-the-art parameter estimation methods such as Expectation Maximization (EM), Subspace method (SS), and Prediction Error Minimization (PEM), for black-box estimation (fully unknown x, θ and λ).

6.17.1 Evaluation Metric for Parameter Estimation
For the black box identification, with completely unknown x, θ, and λ, there are infinite solutions with accurate input–output mapping. However, for LTI systems, there exists a unique transformation for identical systems. We use the companion canonical form to check the validity of parameter estimation by transforming both the real and the estimated parameters into their companion canonical form and then using the (square of) Euclidean distance between them as the sum of squared error (SSE) in parameter estimation. This evaluation metric will be used for parameter estimation in the next section.

6.17.2 Simulation Setup
A total of 500 (5×100) different randomly generated stable systems were used with five different noise smoothness values for parameter estimation. All systems were selected with same number of parameters $n^\theta = 14$ ($n = 2, m = 4$ and $r = 1$), with each $\theta^i \in [-1, 1]$, while ensuring that A matrix is stable. All the noises were generated with the precision of e^6 ($\Pi^w = e^6 I_{2\times2}, \Pi^z = e^6 I_{4\times4}$), with the embedding orders of states and inputs as $p = 6$ and

d = 2. A Gaussian bump of $v = e^{-.25*(t-12)^2}$ was used as the input signal with $dt = 0.5s$ and $T = 32s$. The prior parameter was randomly initialized such that all $\eta^{\theta^i} \in [-2, 2]$ with a tight prior precision of $P^\theta = e^4 I_{14 \times 14}$. Both the hyperparameter priors η^{λ^i} were set to zero, with a prior precision of $P^\lambda = e^{-4} I_{2 \times 2}$.

The System Identification toolbox from MATLAB was used for SS ($n4sid()$) and PEM methods. The solution of SS was used to initialize PEM. An implementation of EM algorithm for state space models was written in MATLAB based on [234]. $n4sid()$ is inherently designed to handle colored noise, whereas the implemented EM algorithm is not. The code for the DEM algorithm will be openly available at: $https://github.com/ajitham123/DEM\_LTI$.

6.17.3 RESULTS

The results shown in Figure 6.7 demonstrate the superior performance of DEM in comparison with EM, PEM, and SS, with minimum SSE during parameter estimation across different noise smoothness. Additionally, EM and PEM exploded occasionally (< 5% times), resulting in outliers in SSE, which were removed for better visualization. DEM demonstrated a consistent performance without generating any such outliers or exploding solutions, which could be explained by DEM's convergence guarantees for parameter estimation under colored noise [167], as proved in Section 6.15. In summary, DEM is a competitive parameter estimator for LTI systems with colored noise.
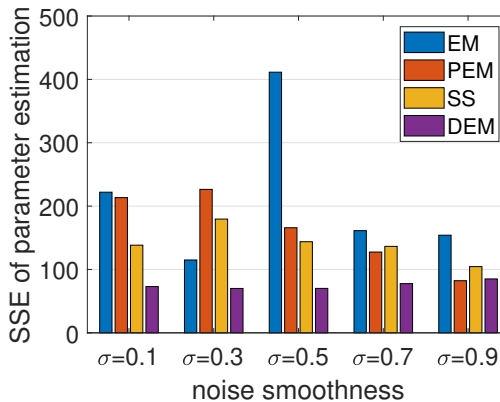
Fig 6.7: The sum of all SSE of Θ for 100 random systems each, for 5 different noise smoothnesses. DEM outperforms EM, PEM, and SS with minimum SSE for parameter estimation under colored noise.

6.18 DISCUSSION

The quest for a brain-inspired learning algorithm for robots has culminated in the free energy principle that postulates biological brain's perception as an optimization over its free energy objectives. FEP is of prime importance to robotics because of the use of generalized coordinates that enables it to gracefully handle colored noises. Colored noises appear in real robotics systems through the unmodeled dynamics and the non-linearity errors in

the model, thereby providing an advantage for DEM during estimation when compared to other estimators. An example could be the unmodeled wind disturbances acting on an unmanned aerial vehicle while in flight, or the non linearity errors in the dynamic model of a robotic manipulator arm involved in a pick and place operation. The scope of this work spans across the blind system identification of such linear dynamic systems with colored noise.

The fundamental difference between this work and the prior work is in the reformulation of DEM for an LTI system. While DEM from computational neuroscience focuses on emulating the biological brain's perception through the hierarchical abstraction of a number of non-linear dynamic systems that interact with each other, our work focuses on reducing this method into an algorithm for the system identification of an LTI system with colored noise, which is a well-known problem in robotics. This reformulation enables the standard analysis for convergence, stability and unbiased estimation, which is an essential analysis in practical robotics. It also enables DEM to be compared with other existing estimation algorithms in a control systems domain. The widespread use of DEM in robotics necessitates these mathematical analyses, especially when concerning the stable and safe operation of robots in industry and during human–robot interaction.

An algorithm with proved convergence for estimation is preferred for safe robotic applications. Therefore, one of the main contribution of this work was the reduction of the estimation algorithm into a coupled augmented system to prove the convergence of parameter and hyperparameter estimation steps. This work also demonstrated the successful applicability of DEM for the estimation of a randomly selected LTI system. Furthermore, we showed through rigorous simulations on a wide range of randomly generated LTI systems that DEM is a competitive algorithm for system identification under colored noise, thereby widening the scope of DEM to a large number of LTI systems in robotics.

One of the main drawbacks of the algorithm is its higher computational complexity when compared to the estimation algorithms that do not keep track of the trajectory of states. Therefore, future work can focus on the online estimation using DEM with reduced computational load. Future work can also focus on extending this algorithm for linear time varying systems to deal with robots with changing system parameters while in operation—a delivery drone dropping deliveries in mid-flight, for example. From a practical robotics point of view, DEM's parameter estimation module can be directly applied to a wide range of robots such as quadrotors, robotic arms, wheeled robots, etc. for black-box system identification, the input estimation module can be employed for fault-detection systems, and the hyperparameter estimation module can be used for online noise estimation for robust control. DEM can also be extended with a control loop for active inference to perform simultaneous perception and action on robots. This would result in the development of cognitive robots that can learn the generative model in the environment by interacting with it and actively seeking new information (active learning) for uncertainty resolution. This would influence multiple domains in robotics such as human–robot interaction for task learning, swarm robotics for collective learning and distributed control, informative path planning of aerial robots for environment monitoring, etc. The development of such brain-inspired autonomous agents sits at the core of cognitive robotics research. In summary, DEM has a huge potential to be the bioinspired learning

algorithm for future robots.

6.19 Conclusion

The free energy principle from neuroscience has a great potential to be one of the most prominent frameworks for learning and control for the autonomous systems in future. Therefore, this paper converted the FEP-based inference scheme called DEM into a joint state, input, parameter, and hyperparameter estimation algorithm for LTI systems with colored noise. We derived the mathematical framework of DEM for LTI systems to prove that the resulting estimator is a combination of linear estimators that are coupled. We provided the proof of convergence for the estimation steps. Through rigorous simulations on randomly generated linear systems with colored noise at varying smoothness levels, we demonstrated that the DEM algorithm outperforms EM, PEM, and SS methods for parameter estimation with minimal estimation error. In light of the potential for DEM to solve the parameter estimation problem, the future research will aim at applying DEM to a quadcopter flying in wind.

6.20 Acknowledgments

6

7

CONVERGENCE PROOF FOR DEM

This chapter provides the convergence proof for the DEM based parameter estimator introduced in Chapter 6. The Chapter 8 provides the experimental confirmation of this estimator on real robot data.

7.1 Abstract

The free energy principle from neuroscience provides an efficient data-driven framework called the Dynamic Expectation Maximization (DEM), to learn the generative model in the environment. DEM's growing potential to be the brain-inspired learning algorithm for robots demands a mathematically rigorous analysis using the standard control system tools. Therefore, this paper derives the mathematical proof of convergence for its parameter estimator for linear state space systems, subjected to colored noise. We show that the free energy based parameter learning converges to a stable solution for linear systems. The paper concludes by providing a proof of concept through simulation for a wide range of spring damper systems.

7.2 Introduction

The free energy principle (FEP) models the brain's perception and action as a gradient ascend over its free energy objective [2]. The action side of FEP, known as active inference [109], has already been applied to real robots including ground robots for SLAM [34], humanoid robots for body perception [200] and manipulator robots for pick and place operation [27]. Similarities with standard control technique like PID was also analyzed [47]. One of the variants of FEP, the Dynamic Expectation Maximization (DEM) [170], provides a model inversion framework for perception and system identification. DEM's distinctive feature lies in its capability to gracefully handle colored noise through the use of generalized coordinates [44], thereby rendering it with the potential to be the learning algorithm for robots. DEM was reformulated as a linear state and input observer under colored noise [46] and was validated for quadrotor flights [165]. A DEM based linear parameter estimator for system identification was developed by [97] and was applied for the perception of quadrotor in wind [166]. Since an estimator with convergence guarantees is preferred for safe robotics applications, we aim at paving way to DEM's practical use by mathematically analyzing it for its convergence properties. Moreover, it is of interest to the active inference community to develop active learning and control strategies with stability guarantees. The presence of generalized coordinates, mean field terms and brain priors complicates the convergence proof and makes it different from other estimators like Expectation Maximization [232]. The goal of this paper is: 1) to show that DEM has convergence guarantees for linear systems with colored noise, and 2) to show that it can be applied to control system problems like the estimation of a mass-spring-damper system.

7.3 Preliminaries

Consider the linear plant dynamics (generative process) given in Eq. 7.1, where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are constant system matrices, $\mathbf{x} \in \mathbb{R}^n$ is the hidden state, $\mathbf{v} \in \mathbb{R}^r$ is the input and $\mathbf{y} \in \mathbb{R}^m$ is the output.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} + \mathbf{w}, \qquad \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{z}. \tag{7.1}$$

Here $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$ represent the process and measurement noise respectively. The notations of the plant are denoted in boldface, whereas its estimates are denoted in non-boldface letters. Since the brain has no access to the plant dynamics except through the

sensory measurements **y**, it maintains the copy of an approximate model of the generative process called the generative model. The noise color assumption (convolution of white noise with a Gaussian kernel) facilitates the differentiated form of the generative model as [170]:

$$
\begin{aligned}
x' &= Ax + B\upsilon + w & y &= Cx + z \\
x'' &= Ax' + B\upsilon' + w' & \dot{y} &= Cx' + z' \\
\dots & & \dots &
\end{aligned}
\tag{7.2}
$$

One of the key technique behind DEM to model the colored noise is to express the time varying components in generalized coordinates, denoted by a tilde operator. The colored noises can be expressed in generalized coordinates using their higher derivatives as $\tilde{z} = [z, z', z'', \dots]^T$ and $\tilde{w} = [w, w', w'', \dots]^T$. The generative model in Eq. 7.2 can be compactly written as [170]:

$$
\dot{\tilde{x}} = D^x \tilde{x} = \tilde{A}\tilde{x} + \tilde{B}\tilde{\upsilon} + \tilde{w} \qquad \tilde{y} = \tilde{C}\tilde{x} + \tilde{z}
\tag{7.3}
$$

where $D^x = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & . & . & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}_{(p+1)\times(p+1)} \otimes I_{n\times n}$, $\tilde{A} = I_{p+1} \otimes A$, $\tilde{B} = I_{p+1} \otimes B$ and $\tilde{C} = I_{p+1} \otimes C$.

Here $\otimes$ is the Kronecker tensor product. To facilitate the convergence proof later in the paper, we introduce a redefinition for Eq. 7.3 with all parameters grouped to the right side as θ:

$$
\dot{\tilde{x}} = M\theta + \tilde{w}, \; \tilde{y} = N\theta + \tilde{z}, \; \theta = \begin{bmatrix} vec(A^T) \\ vec(B^T) \\ vec(C^T) \end{bmatrix},
\tag{7.4}
$$

where

$$
M = \begin{bmatrix} I_n \otimes x^T & I_n \otimes \upsilon^T & I_n \otimes O_{1\times m} \\ I_n \otimes x'^T & I_n \otimes \upsilon'^T & I_n \otimes O_{1\times m} \\ \dots & \dots & \dots \end{bmatrix}, N = \begin{bmatrix} I_n \otimes O_{1\times n} & I_n \otimes O_{1\times r} & I_m \otimes x^T \\ I_n \otimes O_{1\times n} & I_n \otimes O_{1\times r} & I_m \otimes x'^T \\ \dots & \dots & \dots \end{bmatrix}.
\tag{7.5}
$$

The goal of this paper is to mathematically prove that the DEM's estimate for θ converges while maximizing the free energy objective[1].

7.4 PARAMETER LEARNING AS FREE ENERGY OPTIMIZATION

DEM postulates the parameter learning algorithm as the gradient ascend over the free energy action, which is the time integral of free energy $\bar{F} = \int F dt$. The parameter update equation can be expressed as the gradient [97, 170]:

$$
\frac{\partial \theta}{\partial a} = k^\theta \frac{\partial \bar{F}}{\partial \theta} = -P^\theta(\theta - \eta^\theta) + \sum_t (-E_\theta + W_\theta^X),
\tag{7.6}
$$

---

[1] maximization of the ELBO term (Section 1.4)

where k^θ is the learning rate, $E_\theta = \frac{\partial E}{\partial \theta}$ is the gradient of precision weighed prediction error, $W_{\theta^i}^X = \frac{\partial W^X}{\partial \theta}$ is the gradient of state mean field term, η^θ is the prior parameters and P^θ is the prior parameter precision. Subscripts will be used for the derivative operator. E_θ for an LTI system can be simplified as:

$$E_\theta = \tilde{\epsilon}_\theta^T \tilde{\Pi} \tilde{\epsilon}, \text{ where } \tilde{\epsilon} = \begin{bmatrix} \tilde{\mathbf{y}} - N\theta \\ \tilde{v} - \tilde{\eta}^v \\ D^x \tilde{x} - M\theta \end{bmatrix} \text{ and } \tilde{\epsilon}_\theta = \begin{bmatrix} -N \\ O \\ -M \end{bmatrix} \tag{7.7}$$

are the prediction error and its gradient. Here $\tilde{\eta}^v$ is the prior on inputs with prior precision $\tilde{P}^v$, $\tilde{\Pi} = diag(\tilde{\Pi}^z, \tilde{P}^v, \tilde{\Pi}^w)$ is the generalized noise precision with Π^z and Π^w being the precisions (inverse covariance) for measurement and process noise. Here $diag()$ represents the block diagonal operation. Similarly, W_θ^X for an LTI system can be written as [97, 170]:

$$W_{\theta^i}^X = -\frac{1}{2} tr(\Sigma^X \tilde{\epsilon}_{X\theta^i}^T \tilde{\Pi} \tilde{\epsilon}_X), \ \tilde{\epsilon} = \begin{bmatrix} \tilde{\mathbf{y}} - \tilde{C}\tilde{x} \\ \tilde{v} - \tilde{\eta}^v \\ D^x \tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix}, \ \tilde{\epsilon}_X = \begin{bmatrix} -\tilde{C} & O \\ O & I \\ D^x - \tilde{A} & -\tilde{B} \end{bmatrix}. \tag{7.8}$$

7.5 Proof of convergence for parameter estimator

If E_θ and W_θ^X can be expressed as linear in θ, in the form $E_\theta = E_1\theta + E_2$ and $W_\theta^X = W_1\theta + W_2$, Eq. 7.6 can be rewritten as:

$$\frac{\partial \theta}{\partial a} = -\left[P^\theta + \sum_t (E_1 - W_1) \right]\theta + \left[P^\theta \eta^\theta + \sum_t (-E_2 + W_2) \right]. \tag{7.9}$$

The differential equation given by Eq. 7.9 is of the form of a linear state space equation ($\dot{\theta} = A^\theta \theta + B^\theta.1$). From the basics of control theory, the solutions of this equation converges exponentially (stabilise) if $A^\theta = -[P^\theta + \sum_t (E_1 - W_1)]$ is negative definite (negative eigen values). This section aims to prove this result.

Lemma 7.5.1. *If $A, B > O$, then $A + B > O$.*

As per Lemma 7.5.1, the positive definiteness of $P^\theta - \sum_t W_1 + \sum_t E_1$ can be proved by proving the positive definiteness of the individual terms P^θ, $-W_1$ and E_1. We know by definition that the prior parameter precision P^θ is positive definite. We now proceed to prove that $E_1 \geq O$. Upon simplification of Eq. 7.7, E_θ can be written as $E_\theta = E_1\theta + E_2$, where:

$$E_1 = N^T \tilde{\Pi}^z N + M^T \tilde{\Pi}^w M \text{ and } E_2 = -\begin{bmatrix} N^T \tilde{\Pi}^z & M^T \tilde{\Pi}^w D \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{y}} \\ \tilde{x} \end{bmatrix}. \tag{7.10}$$

Lemma 7.5.2. *If $A \geq O$, then $B^T A B \geq O$.*

Proof. By definition, if $A \geq O$, there exists a square root $A^{\frac{1}{2}} \geq O$. Therefore, $x^T(B^T A B)x = x^T(B^T A^{\frac{1}{2}} A^{\frac{1}{2}} B)x = (A^{\frac{1}{2}} Bx)^T(A^{\frac{1}{2}} Bx) \geq 0, \implies B^T A B \geq O$. □

Since $\tilde{\Pi}^z > O$ and $\tilde{\Pi}^w > O$ by definition, from Lemma 7.5.1 and 7.5.2, $E_1 = N^T \tilde{\Pi}^z N + M^T \tilde{\Pi}^w M \geq O$. Therefore, E_1 is proved to be positive semi-definite.

The final term under consideration is W_1. The rest of this section aims to prove that $W_1 \prec O$, which will conclude the entire convergence proof of parameter estimation. We rewrite the mean field term for parameter θ^i from Eq. 7.8 as:

$$
\begin{aligned}
W_{\theta^i}^X &= -\frac{1}{2} tr(\Sigma^X \tilde{\epsilon}_{X\theta^i}^T \tilde{\Pi} \tilde{\epsilon}_X), \\
&= -\frac{1}{2} tr\left[\begin{bmatrix} \Sigma^{\tilde{x}\tilde{x}} & \Sigma^{\tilde{x}\tilde{v}} \\ \Sigma^{\tilde{v}\tilde{x}} & \Sigma^{\tilde{v}\tilde{v}} \end{bmatrix} \begin{bmatrix} \tilde{C}_{\theta^i}^T \tilde{\Pi}^z \tilde{C} - \tilde{A}_{\theta^i}^T \tilde{\Pi}^w (D - \tilde{A}) & \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{B} \\ -\tilde{B}_{\theta^i}^T \tilde{\Pi}^w (D - \tilde{A}) & \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{B} \end{bmatrix} \right] \\
&= -\frac{1}{2} tr\left[\begin{bmatrix} \Sigma^{\tilde{x}\tilde{x}} & \Sigma^{\tilde{x}\tilde{v}} \\ \Sigma^{\tilde{v}\tilde{x}} & \Sigma^{\tilde{v}\tilde{v}} \end{bmatrix} \begin{bmatrix} \tilde{C}_{\theta^i}^T \tilde{\Pi}^z \tilde{C} + \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{A} & \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{B} \\ \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{A} & \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{B} \end{bmatrix} \right] \\
&\quad -\frac{1}{2} tr\left[\begin{bmatrix} \Sigma^{\tilde{x}\tilde{x}} & \Sigma^{\tilde{x}\tilde{v}} \\ \Sigma^{\tilde{v}\tilde{x}} & \Sigma^{\tilde{v}\tilde{v}} \end{bmatrix} \begin{bmatrix} -\tilde{A}_{\theta^i}^T \tilde{\Pi}^w D & O \\ -\tilde{B}_{\theta^i}^T \tilde{\Pi}^w D & O \end{bmatrix} \right].
\end{aligned}
\tag{7.11}
$$

Since the second trace term in Eq. 7.11 is independent of θ^i, it is lumped into the $W_2^{\theta^i}$ term. Equation 7.11 is further simplified as:

$$
\begin{aligned}
W_{\theta^i}^X = -\frac{1}{2} \Big[& tr(\Sigma^{\tilde{x}\tilde{x}} \tilde{C}_{\theta^i}^T \tilde{\Pi}^z \tilde{C}) + tr(\Sigma^{\tilde{x}\tilde{x}} \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{A}) + tr(\Sigma^{\tilde{x}\tilde{v}} \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{A}) \\
& + tr(\Sigma^{\tilde{v}\tilde{x}} \tilde{A}_{\theta^i}^T \tilde{\Pi}^w \tilde{B}) + tr(\Sigma^{\tilde{v}\tilde{v}} \tilde{B}_{\theta^i}^T \tilde{\Pi}^w \tilde{B}) \Big] + W_2^{\theta^i}
\end{aligned}
\tag{7.12}
$$

We aim to separate θ out so that the mean field term can be expressed in the form $W_\theta^X = W_1 \theta + W_2$. We proceed by first introducing the transpose of the generalized parameter matrices $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ to Eq. 7.12 and then moving them out of the trace terms.

Lemma 7.5.3. *If A, B, C and D are matrices, then $tr(ABCD) = tr(C^T B^T A^T D^T)$*

Proof. $tr(ABCD) = tr((ABCD)^T) = tr(D^T C^T B^T A^T) = tr(C^T B^T A^T D^T)$. $\qquad\square$

Lemma 7.5.4. *If A, B and C are matrices, then $tr(ABC) = vec(A^T)^T (I \otimes B) vec(C)$.*

Applying Lemma 7.5.3 throughout Eq. 7.12 results in:

$$
\begin{aligned}
W_{\theta^i}^X = -\frac{1}{2} \Big[& tr(\tilde{\Pi}^{zT} \tilde{C}_{\theta^i} \Sigma^{\tilde{x}\tilde{x}T} \tilde{C}^T) + tr(\tilde{\Pi}^{wT} \tilde{A}_{\theta^i} \Sigma^{\tilde{x}\tilde{x}T} \tilde{A}^T) + tr(\tilde{\Pi}^{wT} \tilde{B}_{\theta^i} \Sigma^{\tilde{x}\tilde{v}T} \tilde{A}^T) \\
& + tr(\tilde{\Pi}^{wT} \tilde{A}_{\theta^i} \Sigma^{\tilde{v}\tilde{x}T} \tilde{B}^T) + tr(\tilde{\Pi}^{wT} \tilde{B}_{\theta^i} \Sigma^{\tilde{v}\tilde{v}T} \tilde{B}^T) \Big] + W_2^{\theta^i},
\end{aligned}
\tag{7.13}
$$

which upon further expansion using Lemma 7.5.4 and grouping yields:

$$
\begin{aligned}
W_{\theta^i}^X = -\frac{1}{2} \Bigg[& \Big(vec(\tilde{A}_{\theta^i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{x}\tilde{x}T}) + vec(\tilde{B}_{\theta^i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{x}\tilde{v}T}) \Big) vec(\tilde{A}^T) \\
& + \Big(vec(\tilde{A}_{\theta^i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{v}\tilde{x}T}) + vec(\tilde{B}_{\theta^i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{v}\tilde{v}T}) \Big) vec(\tilde{B}^T) \\
& + \Big(vec(\tilde{C}_{\theta^i}^T \tilde{\Pi}^z)^T (I \otimes \Sigma^{\tilde{x}\tilde{x}T}) \Big) vec(\tilde{C}^T) \Bigg] + W_2^{\theta^i}.
\end{aligned}
\tag{7.14}
$$

We have now separated all the generalized parameters out of the trace terms in their vector forms. These vectors can be grouped such that the mean field term is linear with respect to the generalized parameter vector $\tilde{\theta} = \begin{bmatrix} vec(\tilde{A}^T) \\ vec(\tilde{B}^T) \\ vec(\tilde{C}^T) \end{bmatrix}$ as:

$$W_{\theta i}^X = -\frac{1}{2}\Bigg[vec(\tilde{A}_{\theta i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{x}\tilde{x}T}) + vec(\tilde{B}_{\theta i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{x}\tilde{v}T}),$$
$$vec(\tilde{A}_{\theta i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{v}\tilde{x}T}) + vec(\tilde{B}_{\theta i}^T \tilde{\Pi}^w)^T (I \otimes \Sigma^{\tilde{v}\tilde{v}T}), \tag{7.15}$$
$$vec(\tilde{C}_{\theta i}^T \tilde{\Pi}^z)^T (I \otimes \Sigma^{\tilde{x}\tilde{x}T}) \Bigg]\tilde{\theta} + W_2^{\theta^i}.$$

Lemma 7.5.5. *If A and B are matrices, then $vec(AB)^T = vec(A)^T(B \otimes I)$.*

We use Lemma 7.5.5 to further simplify Eq. 7.15 as:

$$W_{\theta i}^X = -\frac{1}{2}\Bigg[vec(\tilde{A}_{\theta i}^T)^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{x}\tilde{x}T}) + vec(\tilde{B}_{\theta i}^T)^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{x}\tilde{v}T}),$$
$$vec(\tilde{A}_{\theta i}^T)^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{v}\tilde{x}T}) + vec(\tilde{B}_{\theta i}^T)^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{v}\tilde{v}T}), \tag{7.16}$$
$$vec(\tilde{C}_{\theta i}^T)^T (\tilde{\Pi}^z \otimes I)(I \otimes \Sigma^{\tilde{x}\tilde{x}T}) \Bigg]\tilde{\theta} + W_2^{\theta^i}.$$

Since the parameters A, B and C are independent of each other, their derivatives with respect to each other are zeros, resulting in $vec(\tilde{A}_{\theta i}^T) = O, \forall \theta^i \in \{B, C\}$, $vec(\tilde{B}_{\theta i}^T) = O, \forall \theta^i \in \{A, C\}$ and $vec(\tilde{C}_{\theta i}^T) = O, \forall \theta^i \in \{A, B\}$. This simplifies the expression for $W_{\theta i}^X$ in Eq. 7.16. The total mean field term W_θ^X can be computed by vertically stacking the individual mean field contributions $W_{\theta i}^X$ from each parameter θ^i as:

$$W_\theta^X = -\frac{1}{2}W_3\tilde{\theta} + W_2, \tag{7.17}$$

where $W_3 = \begin{bmatrix} W_4 & O \\ O & W_{5,} \end{bmatrix}$ with $W_5 = vec(\tilde{C}^T)_{vecC^T}^T (\tilde{\Pi}^z \otimes I)(I \otimes \Sigma^{\tilde{x}\tilde{x}T})$ and

$$W_4 = \begin{bmatrix} vec(\tilde{A}^T)_{vecA^T}^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{x}\tilde{x}T}) & vec(\tilde{A}^T)_{vecA^T}^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{v}\tilde{x}T}) \\ vec(\tilde{B}^T)_{vecB^T}^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{x}\tilde{v}T}) & vec(\tilde{B}^T)_{vecB^T}^T (\tilde{\Pi}^w \otimes I)(I \otimes \Sigma^{\tilde{v}\tilde{v}T}) \end{bmatrix}.$$

W_3 can be simplified as:

$$W_3 = \frac{\partial \tilde{\theta}}{\partial \theta}^T \begin{bmatrix} \tilde{\Pi}^w \otimes I & O & O \\ O & \tilde{\Pi}^w \otimes I & O \\ O & O & \tilde{\Pi}^z \otimes I \end{bmatrix} \begin{bmatrix} I \otimes \Sigma^{\tilde{x}\tilde{x}T} & I \otimes \Sigma^{\tilde{v}\tilde{x}T} & O \\ I \otimes \Sigma^{\tilde{x}\tilde{v}T} & I \otimes \Sigma^{\tilde{v}\tilde{v}T} & O \\ O & O & I \otimes \Sigma^{\tilde{x}\tilde{x}T} \end{bmatrix}, \tag{7.18}$$

where $\frac{\partial \tilde{\theta}}{\partial \theta} = diag(vec\tilde{A}^T_{vecA^T}, vec\tilde{B}^T_{vecB^T}, vec\tilde{C}^T_{vecC^T})$. Since the generalized parameter vector $\tilde{\theta}$ is linear in parameter vector θ, we can write:

$$\tilde{\theta} = \frac{\partial \tilde{\theta}}{\partial \theta} \theta = \begin{bmatrix} vec\tilde{A}^T_{vecA^T} & O & O \\ O & vec\tilde{B}^T_{vecB^T} & O \\ O & O & vec\tilde{C}^T_{vecC^T} \end{bmatrix} \theta. \tag{7.19}$$

Substituting Eq. 7.18 and 7.19 in Eq. 7.17 yields:

$$W^X_\theta = W_1 \theta + W_2,$$

$$W_1 = -\frac{1}{2} \frac{\partial \tilde{\theta}}{\partial \theta}^T \begin{bmatrix} \tilde{\Pi}^w \otimes I & O & O \\ O & \tilde{\Pi}^w \otimes I & O \\ O & O & \tilde{\Pi}^z \otimes I \end{bmatrix} \begin{bmatrix} I \otimes \Sigma^{\tilde{x}\tilde{x}T} & I \otimes \Sigma^{\tilde{v}\tilde{x}T} & O \\ I \otimes \Sigma^{\tilde{x}\tilde{v}T} & I \otimes \Sigma^{\tilde{v}\tilde{v}T} & O \\ O & O & I \otimes \Sigma^{\tilde{x}\tilde{x}T} \end{bmatrix} \frac{\partial \tilde{\theta}}{\partial \theta}. \tag{7.20}$$

Therefore, the mean field term W^X_θ is linear in θ. For the parameter estimator to provide a converging solution, we need to prove that $W_1 \prec O$. Lemma 7.5.2 could be applied to the expression for W_1 to prove that $W_1 \prec O$ if:

$$W_6 = \begin{bmatrix} \tilde{\Pi}^w \otimes I & O & O \\ O & \tilde{\Pi}^w \otimes I & O \\ O & O & \tilde{\Pi}^z \otimes I \end{bmatrix} \begin{bmatrix} I \otimes \Sigma^{\tilde{x}\tilde{x}T} & I \otimes \Sigma^{\tilde{v}\tilde{x}T} & O \\ I \otimes \Sigma^{\tilde{x}\tilde{v}T} & I \otimes \Sigma^{\tilde{v}\tilde{v}T} & O \\ O & O & I \otimes \Sigma^{\tilde{x}\tilde{x}T} \end{bmatrix} \succ O \tag{7.21}$$

Lemma 7.5.6. *If $A, B \succeq O$ and A is invertible, then $AB \succeq O$.*

Proof. $AB = A^{\frac{1}{2}}(A^{\frac{1}{2}}BA^{\frac{1}{2}})A^{-\frac{1}{2}}$, implies AB and $A^{\frac{1}{2}}BA^{\frac{1}{2}}$ are similar matrices, sharing all eigen values. Using lemma 7.5.2, since $B \succeq O$, $A^{\frac{1}{2}}BA^{\frac{1}{2}} \succeq O \implies AB \succeq O$. □

Using lemma 7.5.6 it is straightforward to see that $W_6 \succeq O$ because: $\tilde{\Pi}^z \succ O, \tilde{\Pi}^w \succ O, \implies \tilde{\Pi}^z \otimes I \succ O$ and $\tilde{\Pi}^w \otimes I \succ O$, $I \otimes \Sigma^X \succ O$. Therefore, $W_1 \preceq O$. This completes the proof that the parameter estimation of DEM converges for an LTI system with colored noise.

7.6 Proof of concept: mass-spring-damper system

This section aims at providing a proof of concept for the convergence of DEM's parameter estimator, through realistic simulations. A mass-spring-damper system with mass $m = 1.4kg$, spring constant $k = 0.8N/m$ and damping coefficient $b = 0.4Ns/m$, is considered in the state space form given by:

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} v, \ y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix}. \tag{7.22}$$

A Gaussian bump input $v = e^{-0.25(t-12)^2}$, centred around 12s and sampled at $dt = 0.1s$ for $T = 32s$ was used. To generate the colored noise, the white noise ($\Pi^w = e^6 I_2$ and $\Pi^z = e^6$) was convoluted using a Gaussian kernel with a width of $\sigma = 0.5s$. A partially known system with unknown $\theta_3 = -\frac{k}{m}$, $\theta_4 = -\frac{b}{m}$ and $\theta_6 = \frac{1}{m}$ was considered. Using the output **y** generated from the spring damper system, parameter estimation was performed using DEM for 25 experiments with different η^θ. The parameter priors η^θ for unknown parameters were
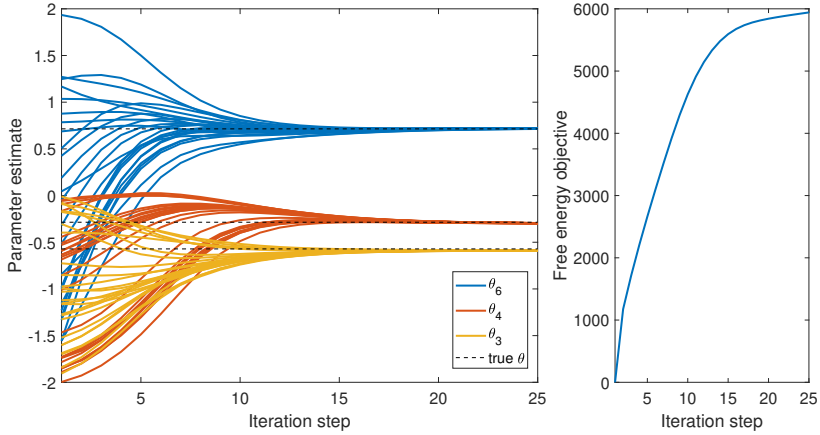
Fig 7.1: The parameter estimates of DEM converges to the correct value of $\theta_3 = -\frac{k}{m} = -0.5714$, $\theta_4 = -\frac{b}{m} = -0.2857$ and $\theta_6 = \frac{1}{m} = 0.7143$, marked in black, for a set of 25 experiments, despite being initialized by randomly sampled priors such that $\eta^{\theta^i} \in [-2, 2]$ and that the prior A matrix is stable. The parameter estimation proceeds by maximizing the free energy objective as shown on the right (sample realization).

randomly sampled from [-2,2] such that the resulting prior A matrix is stable. A low prior precision ($P^{\theta_i} = e^{-4}$) was used for known parameters, and a high precision ($P^{\theta_i} = e^{32}$) was used for unknown parameters. The order of generalized motion of $p = 6$ and $d = 2$ were used for the states and inputs respectively. The result for DEM's parameter estimation is shown in Fig. 7.1. Despite being initialized by random wrong priors, DEM's parameter estimates exponentially converges to the correct values, by maximizing the free energy objective.

Next, we proceed to show that the estimate converges for a wide range of systems. The same experiment was repeated for 25 different randomly selected stable mass-spring-damper systems. Although the convergence applies to unstable systems, sampling was restricted to stable systems within the range [-1,1] ($\theta_3, \theta_4 \in [-1, 0]$ and $\theta_6 \in [0, 1]$) for better visualization. DEM was initialized with the same priors for all experiments ($\eta^{\theta_6} = 2$, $\eta^{\theta_4} = -1$ and $\eta^{\theta_3} = -2$). Figure 7.2 shows that DEM is capable of providing converging solutions for a wide range of stable spring-damper systems, that are influenced by colored noise. Note that the numerical analysis is restricted to the dynamics of spring damper systems for demonstrative purposes, and can be extended to other systems. In summary, DEM can provide converging parameter estimates for linear systems with colored noise, by maximizing the free energy objective.
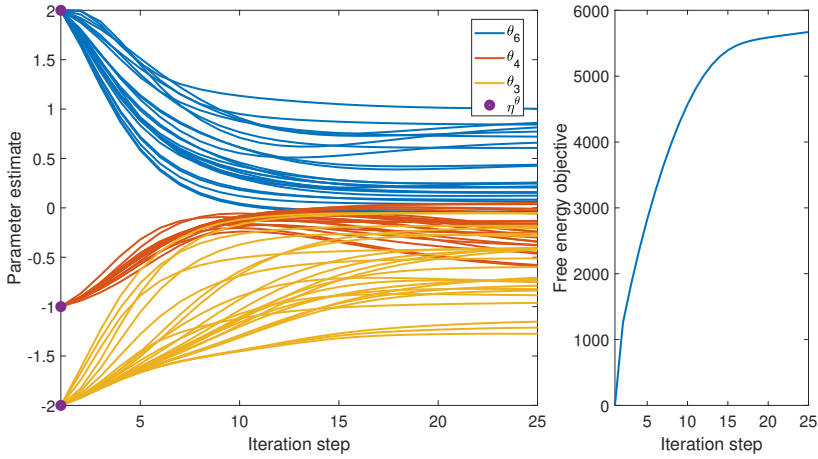
Fig 7.2: DEM's parameter estimates for 25 different randomly sampled stable mass-spring-damper systems. The estimates for all the experiments started from the same prior of $\eta^{\theta_6} = 2$, $\eta^{\theta_4} = -1$ and $\eta^{\theta_3} = -2$, and converged, while maximizing the free energy objective. Therefore, the estimator converges for a wide range of systems.

7.7 CONCLUSION

DEM has the potential to be a bioinspired learning algorithm for future robots, due to its capability to robustly handle colored noise. Its superior performance in state estimation under colored noise was proven by [46] and was experimentally validated by [165]. In this paper, we derived a mathematical proof of convergence for DEM's parameter estimator, applied to linear systems with colored noise. We proved that a perception scheme based on the gradient ascend of the free energy action, provides a converging solution. Since a convergence proof is mandatory for the safe and reliable application of DEM on real robots, this work widens its applicability in robotics. The applicability of DEM for real control system problem was demonstrated through rigorous simulations on the estimation problem for mass-spring-damper systems. The future research will focus on the conditions for unbiased estimation and on applying DEM to real robots.

8

DEM for Quadrotor in Wind

*This chapter uses **precision learning** for the robot model learning of a quadrotor in wind. It provides the experimental confirmation for the system identification tool developed in Chapter 6, and used in Section 2.5.2.3.*

This chapter is a verbatim copy of the under-review paper [166] 📄 *Ajith Anil Meera and Martijn Wisse. "Dynamic Expectation Maximization for the System Identification of a Quadrotor in Wind." IEEE Robotics and Automation Letters (2022), (under review).*

8.1 ABSTRACT

S ystem identification under colored noise is still an open challenge in robotics. In this paper, we introduce a brain inspired system identification method for the output prediction of a quadrotor hovering in wind. Using real quadrotor flight experiments, we show that our approach outperforms state of the art system identification methods like Subspace Identification, Expectation Maximization and Prediction Error Minimization with least output prediction error. With this work, we provide the first experimental confirmation outside of simulations, for the use of Dynamic Expectation Maximization as a robot model learner. Based on the empirical results, we propose an extended algorithm for the model order selection during the blind identification of linear systems.

8.2 INTRODUCTION

The uncertainty handling capabilities of the human brain has been inspiring roboticists to search for brain inspired algorithms for robot perception. The recent advancements in computational neuroscience has culminated in the Free Energy Principle (FEP) that models the brain's perception and action under one optimization scheme [2]. Fundamentally rooted in Bayesian Inference, FEP emerges as a brain theory that can learn hierarchical causal dynamic models from limited data under uncertainties. In light of these developments, we aim to bridge the gap between FEP and robotics by providing the first experimental proof of concept for an FEP based parameter estimator called Dynamic Expectation Maximization (DEM) [170] for robot model learning.

We introduce the idea of leveraging the information content in the noise - emerging from the unmodelled system dynamics (wind) and linearization error - for accurate output predictions. We propose a DEM based model learning scheme that models this noise (prediction error) as colored and improves the estimation using Generalized Coordinates (GC). Fig. 8.1 shows our proposed perception scheme applied to a quadrotor hovering under unmodelled wind conditions. The possible applications of this approach is wide and include the handling of unmodelled wind dynamics on a delivery drone, linearization errors of an industrial manipulator robot, friction dynamics of a skid steer ground robot in unknown terrains like martian surface. The core contributions of the paper include:

1. introduce a brain inspired output prediction scheme for a quadrotor hovering in wind,

2. provide the first experimental confirmation for the use of DEM and GC for robot model learning,

3. introduce a model order estimation algorithm (Algorithm 3) for the blind identification of linear systems.
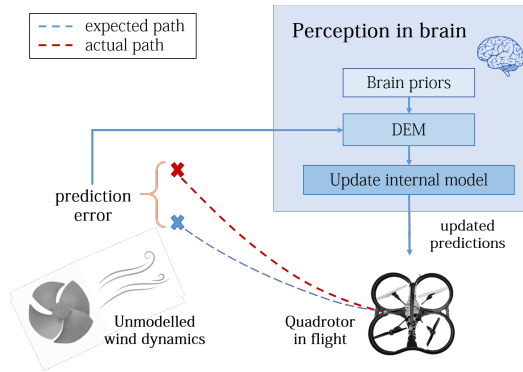
Fig 8.1: The proposed brain inspired prediction scheme. The unmodelled wind and the linearization errors in the robot brain's internal model manifests as the sensory surprisal (output prediction errors). The free energy principle (DEM) drives the robot brain to update its internal model of the world (quadrotor dynamics) for uncertainty resolution, resulting in better output predictions.

8.3 Related work

In this section, we capture the multidisciplinary nature of FEP literature, connecting neuroscience and robotics.

8.3.1 Cognitive neuroscience

FEP emerges from neuroscience as a unifying brain theory [2] that explains the brain functions under a single framework - free energy optimization. According to FEP, every self organising system that is in equilibrium with the environment should minimize its free energy. This drives a biological agent into minimizing its sensory surprisal for uncertainty resolution while interacting with the environment. This is done in two ways - through perception (learning) and action (active inference). Perception involves learning the generative process in the environment for accurate predictions, whereas active inference involves acting on the environment to suppress sensory surprisal. Numerous methods have been developed around these ideas to explain brain functions including predictive coding [8], hierarchical brain models [44], active inference [109], DEM [170] etc. The biological plausibility of FEP rests in its capability to provide a mathematical description of brain functions [187], to unify action and perception [50], to connect physiological constructs like memory, attention, value, reinforcement and salience [187], to explain active vision [133], while remaining consistent with Freudian ideas[188]. Similarities of FEP with reinforcement learning [42], neural networks [43, 44], Kalman Filtering [46], PID control [47] and active learning [50] further guides the quest for a brain inspired robot learning algorithm towards FEP as the unified robot learning algorithm.

8.3.2 Robotics

Recent applications of FEP in robotics include the body perception of humanoid robots [200], adaptive control for robot manipulators [27], robot navigation and mapping (SLAM) [34] etc. Simultaneous state and input observer designs for linear time invariant (LTI)

systems with colored noise was developed [46] and applied to quadrotors [165]. On the learning side, DEM was developed into a system identification method [97] with theoretical convergence guarantees [167]. In this paper we provide the proof of concept for DEM by learning a quadrotor model for output prediction.

8.3.3 System identification
In control systems, output predictions can be done using system identification, which is a mature field [235] with methods like Subspace Identification (SS), Expectation Maximization (EM), Prediction Error Minimization (PEM). However, most of them consider the noises to be white, which is often a wrong assumption in practice and results in biased estimation for least square based methods [218] and inaccurate convergence for the iterative methods [219]. Although various bias compensation methods have been proposed to solve this problem [220, 221], none of them perform simultaneous state, input, parameter and noise hyperparameter estimation, except for DEM. Therefore, DEM is of importance to the research community and requires experimental validation on real robots. With this paper, we aim to fill this research gap.

8.4 Problem statement
Consider the linearized plant dynamics given in Equation 8.1 where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are constant system matrices, with hidden state $\mathbf{x} \in \mathbb{R}^n$, input $\mathbf{v} \in \mathbb{R}^r$ and output $\mathbf{y} \in \mathbb{R}^m$.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} + \mathbf{w}, \qquad \mathbf{y} = C\mathbf{x} + \mathbf{z}. \tag{8.1}$$

Here $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^m$ represent the process and measurement noise respectively. In this paper, we consider a special case where the system is a hovering quadrotor. Variables of the plant are denoted in boldface, while its estimates are denoted in non-boldface. The noises are assumed to be colored such that it was generated by the convolution of white noise with a Gaussian filter. The unmodelled wind dynamics and the linearization errors enter the system through w, making it colored. The problems considered in the paper are: 1) learn an LTI model to accurately predict the output y from the inputs v, and 2) learn the order of the system n for black box identification.

8.5 Preliminaries
To lay the foundations of our prediction scheme, this section introduces the key concepts behind DEM.

8.5.1 Generalized coordinates
The key concept that differentiates DEM from other methods is its use of GC. GC is a relatively new concept in robotics and shouldn't be confused with the definition in multi-body dynamics. GC enables the estimator to gracefully handle colored noise by modelling the trajectory (instead of point estimates) of all the time dependent components like states, inputs, outputs and noises using their higher order derivatives, thereby providing additional information for model learning. For example, the states in GC is given by $\tilde{x} = [x \ x' \ x'' \ ...]^T$. The variables in generalized coordinates are denoted by a tilde, and their components

(higher derivatives) are denoted by primes. In Section 8.6.7 we will demonstrate the usefulness of GC in providing additional information for robot model learning.

8.5.2 GENERATIVE MODEL

The generative model denotes the robot brain's internal model of the generative process in the environment that is responsible for data generation. Since the time dependent components of the generative model are differentiable and because the noises are coloured, the evolution of states of an LTI system (generative process) in Equation 8.1 can be extended as:

$$
\begin{aligned}
x' &= Ax + Bv + w & y &= Cx + z \\
x'' &= Ax' + Bv' + w' & y' &= Cx' + z' \\
&\dots & &\dots
\end{aligned}
\tag{8.2}
$$

which can be compactly written as:

$$
\tilde{x}' = D^x \tilde{x} = \tilde{A}\tilde{x} + \tilde{B}\tilde{v} + \tilde{w} \qquad \tilde{y} = \tilde{C}\tilde{x} + \tilde{z}
\tag{8.3}
$$

where $D^x = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & 0 & 1 \\ & & & & 0 \end{bmatrix}_{(p+1)\times(p+1)} \otimes I_{n\times n}$

performs derivative operation, equivalent to shifting up all components in generalized coordinates by one block. p and d are the order of generalized motion of states and inputs respectively. Here, $\tilde{A} = I_{p+1} \otimes A$, $\tilde{B} = I_{p+1} \otimes B$ and $\tilde{C} = I_{p+1} \otimes C$, where $\otimes$ is the Kronecker tensor product. The generalized motion of output $\tilde{y}$ are computed from the discrete measurements $\mathbf{y}$ [46].

8.5.3 COLORED NOISE MODELING

The colored noises are analytic such that the covariance of noise derivatives $\tilde{z} = [z, z', z'', \dots]^T$ and $\tilde{w} = [w, w', w'', \dots]^T$ are well defined. The correlation between noise derivatives are represented using the temporal precision matrix S (inverse of covariance matrix). Since the correlation is assumed to be due to a Gaussian filter, S becomes [170]:

$$
S(\sigma^2) = \begin{bmatrix} 1 & 0 & -\frac{1}{2\sigma^2} & \cdot\cdot \\ 0 & \frac{1}{2\sigma^2} & 0 & \cdot\cdot \\ -\frac{1}{2\sigma^2} & 0 & \frac{3}{4\sigma^4} & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{bmatrix}^{-1}_{(p+1)\times(p+1)}
\tag{8.4}
$$

where σ^2 is the variance of Gaussian filter, with σ denoting the noise smoothness. While $\sigma^2 = 0$ denotes white noise, non-zero σ^2 denotes colored noise. The generalized noise precision matrices are given by $\tilde{\Pi}^w = S(\sigma^2) \otimes \Pi^w$ and $\tilde{\Pi}^z = S(\sigma^2) \otimes \Pi^z$, where Π^w and Π^z are the inverse noise covariances.

8

8.5.4 PARAMETERS AND NOISE HYPERPARAMETERS

The parameter θ is the vectorized A, B, C matrices, $\theta = \begin{bmatrix} vec(A^T)^T & vec(B^T)^T & vec(C^T)^T \end{bmatrix}^T$, and the noise hyperparameter $\lambda = \begin{bmatrix} \lambda^z & \lambda^w \end{bmatrix}^T$ models the noise precision:

$$\Pi^w(\lambda^w) = e^{\lambda^w} \Omega^w, \ \Pi^z(\lambda^z) = e^{\lambda^z} \Omega^z, \tag{8.5}$$

where Ω^w and Ω^z represent constant matrices encoding how different noises are correlated. We use Ω^w and Ω^z as identity matrices for this work. Parameter and hyperparameter estimation entails the estimation of θ and λ respectively.

8.5.5 PRIORS OF THE BRAIN

DEM enables the transfer of prior knowledge through Gaussian prior distributions for inputs, parameters and hyperparameters, centred around η as $p(\tilde{v}) = \mathcal{N}(\eta^{\tilde{v}}, P^{\tilde{v}})$, $p(\theta) = \mathcal{N}(\eta^\theta, P^\theta)$ and $p(\lambda) = \mathcal{N}(\eta^\lambda, P^\lambda)$ respectively. The mean η acts as the starting point for the model learning on new data and the precision P shapes the confidence on these priors. P controls the robot brain's exploration-exploitation trade off during learning - lower P favours exploration, whereas higher P favours exploitation. We will exhaustively use this idea to pass known information to the algorithm (for example, known inputs) through η with high P.

8.5.6 PERCEPTION AS BAYESIAN INFERENCE

The biological brain's perception is modelled as a Bayesian Inference which involves the computation of the posterior probability density $p(\vartheta/y)$ of parameter θ, given the sensory measurement y [170]. Since it involves the computation of an intractable integral $p(\vartheta/y) = p(\vartheta, y)/\int p(\vartheta, y)d\vartheta$, a variational density $q(\vartheta)$ called the recognition density is defined to closely approximate the posterior as $q(\vartheta) \approx p(\vartheta/y)$. This approximation is achieved by minimizing the Kullback-Leibler (KL) divergence of the distributions given by $KL(q(\vartheta)\|p(\vartheta/y)) = \langle \ln q(\vartheta) \rangle_{q(\vartheta)} - \langle \ln p(\vartheta/y) \rangle_{q(\vartheta)}$, where $\langle . \rangle_{q(\vartheta)}$ represents the expectation over $q(\vartheta)$. Upon simplification using $p(\vartheta/y) = p(\vartheta, y)/p(y)$, it becomes:

$$\ln p(y) = \underbrace{\langle \ln p(\vartheta, y) \rangle_{q(\vartheta)} - \langle \ln q(\vartheta) \rangle_{q(\vartheta)}}_{\text{free energy}} + KL(q(\vartheta)\|p(\vartheta/y)), \tag{8.6}$$

where $\ln p(y)$ is called the log-evidence. Since $\ln p(y)$ is independent of ϑ, the minimization of KL divergence for inference results in the maximization of free energy[1]. This is the core idea behind using free energy as a proxy for perception through Bayesian Inference.

8.5.7 FREE ENERGY OBJECTIVES

Two types of free energy objectives are used by DEM for perception: 1) free energy F for the estimation of time varying components ($X = \begin{bmatrix} \tilde{x} \\ \tilde{v} \end{bmatrix}$) and 2) free energy action $\bar{F} = \int F dt$ for the estimation of time invariant components (θ and λ). The free energy F emerges from

---

[1]maximization of the ELBO term (Section 1.4)

(a) $\bar{F}^o$ with respect to parameters A and B.

(b) Perception as maximization of $\bar{F}^o$.

(c) Top view of Fig. 8.2b.
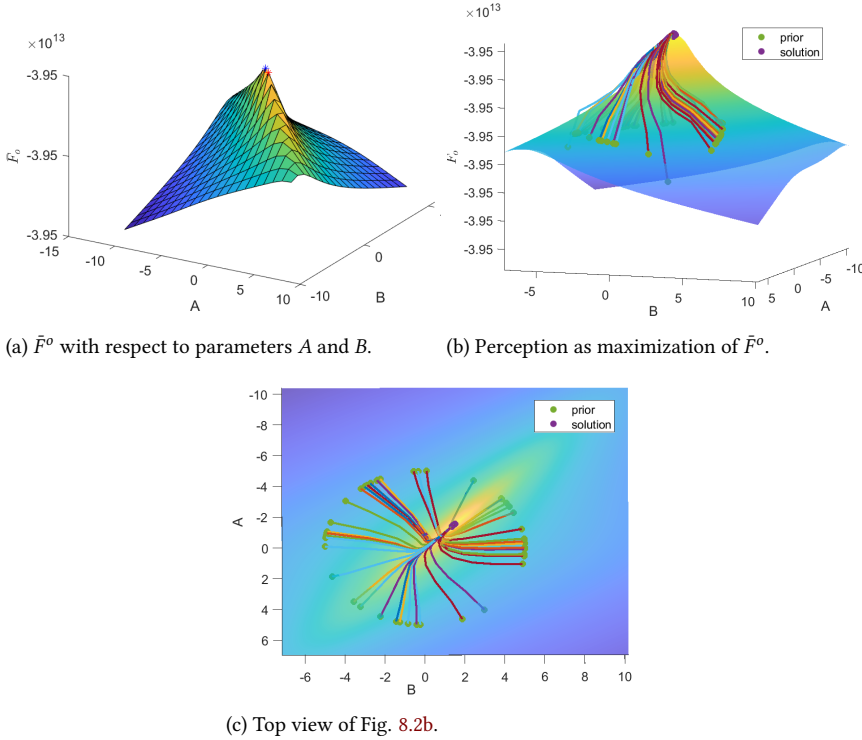
Fig 8.2: (a) The shape of the free energy manifold $\bar{F}^o$ with respect to parameters A and B (both chosen as scalars for vizualization) changes with each E step iteration i, because of the interdependence between $\tilde{x}$ and θ. Gradient ascend over $\bar{F}^o$ at each E step sharpens the peak around the real parameters, where $\bar{F}^o$ is the maximum. (b) Visualization of perception as a gradient ascend over free energy objective. 50 randomly sampled η^θ (green dots) lying on a circle climb up the free energy curve to converge to the same parameters (magenta dot) that coincides with the peak of one of the realizations of the free energy curve. (c) Top view of Fig. 8.2b.

Bayesian statistics (Variational Inference) as an upper bound on surprise [2], and can be written as the sum of its internal energy U, mean field term W and the entropy term H as:

$$F = U + W + H. \tag{8.7}$$

After Laplace approximation and mean-field approximation, U, W and H for an LTI system can be simplified as [97, 170]:

$$
\begin{aligned}
U &= -\frac{1}{2}\tilde{\epsilon}^T \tilde{\Pi} \tilde{\epsilon} - \frac{1}{2}\epsilon^{\theta T} P^\theta \epsilon^\theta - \frac{1}{2}\epsilon^{\lambda T} P^\lambda \epsilon^\lambda + \frac{1}{2}\ln|\tilde{\Pi}| + \frac{1}{2}\ln|P^\theta| + \frac{1}{2}\ln|P^\lambda|, \\
W &= tr(\Sigma^{\tilde{x}} U_{\tilde{x}\tilde{x}} + \Sigma^{\tilde{v}} U_{\tilde{v}\tilde{v}} + \Sigma^\theta U_{\theta\theta} + \Sigma^\lambda U_{\lambda\lambda}), \\
H &= \frac{1}{2}\ln|\Sigma^\theta| + \frac{1}{2}\ln|\Sigma^\lambda| + \frac{1}{2}\ln|\Sigma^{\tilde{x}}| + \frac{1}{2}\ln|\Sigma^{\tilde{v}}|,
\end{aligned}
\tag{8.8}
$$

where $\tilde{\epsilon} = \begin{bmatrix} \tilde{\mathbf{y}} - \tilde{C}\tilde{x} \\ \tilde{v} - \tilde{\eta}^v \\ D^x\tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v} \end{bmatrix}$, $\epsilon^\theta = \boldsymbol{\theta} - \theta$ and $\epsilon^\lambda = \boldsymbol{\lambda} - \lambda$ are the prediction error for components

in generalized coordinates, parameters and hyperparameters respectively. The prediction errors are precision weighed with the generalized precision matrix $\tilde{\Pi} = diag(\tilde{\Pi}^z, P^{\tilde{v}}, \tilde{\Pi}^w)$, where $diag(.)$ is the block diagonal operation. Here $\Sigma^{\tilde{x}}$, $\Sigma^{\tilde{v}}$, Σ^θ and Σ^λ are the covariance matrices denoting the uncertainty in the estimation of states, inputs, parameters and hyperparameters respectively. The free energy action $\bar{F}$ can be written as [97]:

$$\bar{F} = -\frac{1}{2}\epsilon^{\theta T}P^\theta\epsilon^\theta - \frac{1}{2}\epsilon^{\lambda T}P^\lambda\epsilon^\lambda + \frac{1}{2}\sum_{\vartheta^i}\sum_t W^{\vartheta^i} + \frac{1}{2}\sum_t\left(-\tilde{\epsilon}^T\tilde{\Pi}\tilde{\epsilon} + \ln|\tilde{\Pi}| + \ln|\tilde{\Sigma}^x| + \ln|\tilde{\Sigma}^v|\right)$$
$$+ \frac{1}{2}\ln|P^\theta| + \frac{1}{2}\ln|P^\lambda| + \frac{1}{2}\ln|\Sigma^\theta| + \frac{1}{2}\ln|\Sigma^\lambda|,$$

$$(8.9)$$

where $W^{\vartheta^i} = tr(\Sigma^{\vartheta^i}U_{\vartheta^i\vartheta^i})$ is the mean field term of $\vartheta^i \in \{\tilde{x}, \tilde{v}, \theta, \lambda\}$. $\bar{F}$ can be seen as a generalized objective for Expectation Maximization (EM) algorithm with additional capabilities to handle colored noise. Removing generalized coordinates, brain's priors and the mean-field terms equates the objective functions of EM and DEM.

8.5.8 Perception as free energy optimization

DEM models the brain's inference process probabilistically through the estimation of two main components: the mean estimate and the uncertainty (inverse precision) in estimation.
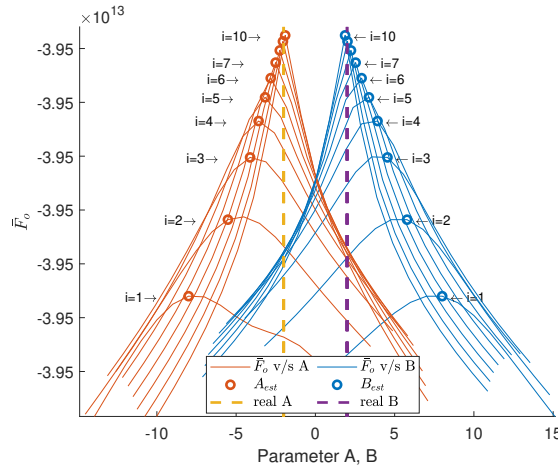
Fig 8.3: Perception as the maximization of $\bar{F}^o$. The parameter estimates (circles) start at $i = 1$ with wrong priors ($\eta^A = -8, \eta^B = 8$) and converges to the correct parameters (yellow and violet lines) at $A = -2$ and $B = 2$. The $\bar{F}^o$ curves (red and blue) are the cross sections of the manifold similar to the one in Fig. 8.2a at different E step iteration i. The peak of these curves rise and narrows around the correct parameter until convergence. The increasing curvature of these peaks during learning is indicative of increasing confidence in estimation Π^θ as given in Equation 8.12.

The mean estimate is computed through a gradient ascend on the free energy manifold. Accordingly, the update equation at time t, a^{th} parameter update and b^{th} hyperparameter update can be written as [170]:

$$\frac{\partial X}{\partial t} = DX + k^X \frac{\partial F}{\partial X}, \quad \frac{\partial \theta}{\partial a} = k^\theta \frac{\partial \bar{F}}{\partial \theta}, \quad \frac{\partial \lambda}{\partial b} = k^\lambda \frac{\partial \bar{F}}{\partial \lambda}, \tag{8.10}$$

where k^X, k^θ and k^λ are the learning rates. $\bar{F}$ is maximized with respect to the estimation uncertainty Σ^{ϑ^i} when the first gradient is zero and the second gradient is negative definite.

$$\frac{\partial \bar{F}}{\partial \Sigma^{\vartheta^i}} = \frac{1}{2} \frac{\partial}{\partial \Sigma^{\vartheta^i}} \left(\ln|\Sigma^{\vartheta^i}| + \sum_t tr(\Sigma^{\vartheta^i} U_{\vartheta^i \vartheta^i}) \right) = \frac{1}{2} \left((\Sigma^{\vartheta^i})^{-1} + \bar{U}_{\vartheta^i \vartheta^i} \right),$$

$$\frac{\partial^2 \bar{F}}{(\partial \Sigma^{\vartheta^i})^2} = -\frac{1}{2} (\Sigma^{\vartheta^i})^{-2} < O. \tag{8.11}$$

Forcing the first gradient to zero yields the optimal precision (inverse covariance) of estimates as:

$$\Pi^{\tilde{x}} = -U_{\tilde{x}\tilde{x}}, \ \Pi^{\tilde{v}} = -U_{\tilde{v}\tilde{v}}, \ \Pi^\theta = -\bar{U}_{\theta\theta}, \ \Pi^\lambda = -\bar{U}_{\lambda\lambda} \tag{8.12}$$

Note that $\bar{U}$ and U are used for time independent and time dependent ϑ^i respectively. Therefore, the mean estimates and the uncertainty in their estimation can be obtained from Equations 8.10 and 8.12, only by using the first two gradients of the energy terms $(F, \bar{F}, U, \bar{U})$. Substituting Equation 8.12 in 8.9 eliminates the mean field terms and simplifies $\bar{F}$ for an LTI system at optimal precision as the sum of weighted prediction errors and entropy [97]:

$$\bar{F}^o = \frac{1}{2} n^t \underbrace{\left[\ln|\tilde{\Pi}^z| + \ln|\tilde{P}^v| + \ln|\tilde{\Pi}^w| \right]}_{\text{noise entropy}} + \underbrace{\frac{1}{2} n_t \ln|\tilde{\Sigma}^X|}_{\text{state and input entropy}} + \underbrace{\frac{1}{2} \ln|\Sigma^\theta P^\theta|}_{\text{parameter entropy}} + \underbrace{\frac{1}{2} \ln|\Sigma^\lambda P^\lambda|}_{\text{hyperparameter entropy}}$$

$$-\frac{1}{2} \sum_t \left[\underbrace{(\tilde{\mathbf{y}} - \tilde{C}\tilde{x})^T \tilde{\Pi}^z (\tilde{\mathbf{y}} - \tilde{C}\tilde{x})}_{\text{prediction error of outputs}} + \underbrace{(\tilde{v} - \tilde{\eta}^v)^T \tilde{P}^v (\tilde{v} - \tilde{\eta}^v)}_{\text{prediction error of inputs}} \right] - \frac{1}{2} \underbrace{(\theta - \eta^\theta)^T P^\theta (\theta - \eta^\theta)}_{\text{prediction error of parameters}}$$

$$-\frac{1}{2} \sum_t \left[\underbrace{(D^x \tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v})^T \tilde{\Pi}^w (D^x \tilde{x} - \tilde{A}\tilde{x} - \tilde{B}\tilde{v})}_{\text{prediction error of states}} \right] - \frac{1}{2} \underbrace{(\lambda - \eta^\lambda)^T P^\lambda (\lambda - \eta^\lambda)}_{\text{prediction error of hyperparameters}}$$

$$\tag{8.13}$$

Maximizing $\bar{F}$ for perception is equivalent to minimizing the prediction error, while maximizing the uncertainty in estimation through the entropy terms. This acts like regularization, preventing the brain from overfitting the model to the data, making it an ideal objective function for robot learning.

8.5.9 Dynamic Expectation Maximization

DEM postulates the brain's perception as a gradient ascend of its free energy objectives using three steps [170]:

- D step - state ($\tilde{x}$) and input ($\tilde{v}$) estimation,

Table 8.1: LINEARIZED QUADCOPTER MODELS of DIFFERENT ORDER.

	Order	x	A	B				C	v
System 1	1	$\dot{\phi}$	0	$\begin{bmatrix} 0.3748 & -0.3748 & -0.3748 & 0.3748 \end{bmatrix}$				I_1	$\begin{bmatrix} pwm^1 \\ pwm^2 \\ pwm^3 \\ pwm^4 \end{bmatrix}$
System 2	2	$\begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.3748 & -0.3748 & -0.3748 & 0.3748 \end{bmatrix}$				I_2	$\begin{bmatrix} pwm^1 \\ pwm^2 \\ pwm^3 \\ pwm^4 \end{bmatrix}$
System 3	3	$\begin{bmatrix} \dot{y} \\ \phi \\ \dot{\phi} \end{bmatrix}$	$\begin{bmatrix} 0 & -9.81 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.3748 & -0.3748 & -0.3748 & 0.3748 \end{bmatrix}$				I_3	$\begin{bmatrix} pwm^1 \\ pwm^2 \\ pwm^3 \\ pwm^4 \end{bmatrix}$
System 4	4	$\begin{bmatrix} y \\ \dot{y} \\ \phi \\ \dot{\phi} \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -9.81 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0.3748 & -0.3748 & -0.3748 & 0.3748 \end{bmatrix}$				I_4	$\begin{bmatrix} pwm^1 \\ pwm^2 \\ pwm^3 \\ pwm^4 \end{bmatrix}$

- E step - parameter (θ) estimation and

- M step - hyperparameter (λ) estimation.

DEM results in Gaussian probability distributions with its mean as the estimate and its standard deviation as the uncertainty in estimation. The D steps follows the gradient ascend over F given in Equation 8.7 to estimate $\tilde{x}, \tilde{v}, \Pi^{\tilde{x}}$ and $\Pi^{\tilde{v}}$. The E and M steps follows the gradient ascend over $\bar{F}$ given in Equation 8.9 to estimate θ, Π^{θ} and λ, Π^{λ} respectively. We use the reformulated version of DEM for LTI systems from our previous work [97] for rest of the paper. As an illustrative example, Fig. 8.2 demonstrates DEM's parameter learning procedure, whereas Fig. 8.3 demonstrates the evolution of the cross sections of $\bar{F}^o$ manifold.
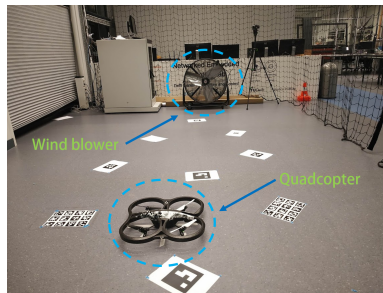

Fig 8.4: The quadcopter and the wind blower in the lab environment.

8.6 Experimental results and analysis

This section aims to provide a proof of concept for DEM using the data from a real quadrotor hovering in turbulent wind conditions, using the simplest linear model possible.

8.6.1 Experimental setup

A Parrot AR drone 2.0 was used to hover under turbulent wind conditions in a controlled lab environment [165], as shown in Fig. 8.4. A PID controller tries to resist the turbulent wind to hover the quadrotor at position $(0m, 0m, 1m)$. A simple linear state space model (Equation 8.1) that maps four rotor PWM signals of the quadrotor to its roll angle ϕ and roll angular velocity $\dot{\phi}$ is constructed (System 2 in Table 8.1). A linear model is used so that some linearization error is generated and contributes to the *colored* process noise. The model doesn't account for the wind disturbances, further contributing to the noise color. The optitrack motion capture system directly observes the internal states (x) of the quadcopter through measurements y with a precision of measurement noise Π^z. The model is derived from [207] after linearization around the equilibrium point. The key idea behind this experimental design is to generate colored process noise w in the quadrotor system (with unmodelled wind dynamics) using turbulent wind.
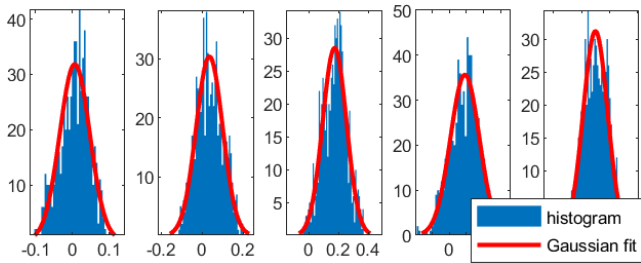
Fig 8.5: The histograms of process noise $w^{\dot{\phi}}$ for all five flight experiments follow a Gaussian distribution.

8.6.2 Data preparation

We consider five quadrotor flights under different wind conditions (wind speed and blower orientation), all for a duration of 850 time steps each with $dt = 0.0083s$. Although different wind conditions might induce different levels of noise color, it doesn't influence our final result. We split each time series data into two parts: training data (700 time steps $\simeq 80\%$) and test data (150 time steps $\simeq 20\%$). The training data is used to learn the model, whereas the test data (unseen data) is used to test the performance of the learned model for output prediction. As a pre-processing step, the input pwm signal to the rotor was mean shifted to zero, and scaled down from high values using: $v = \frac{v - mean(v)}{max(v) - min(v)}$.

8.6.3 Noise color and Laplace approximation

We validate the two fundamental assumptions of DEM, the Laplace approximation and the noise color assumption, by using noise histogram and noise autocorrelation graph

(a) Output prediction.   (b) Parameter estimation.

(c) Maximize free energy action $\bar{F}(i) - \bar{F}(0)$.
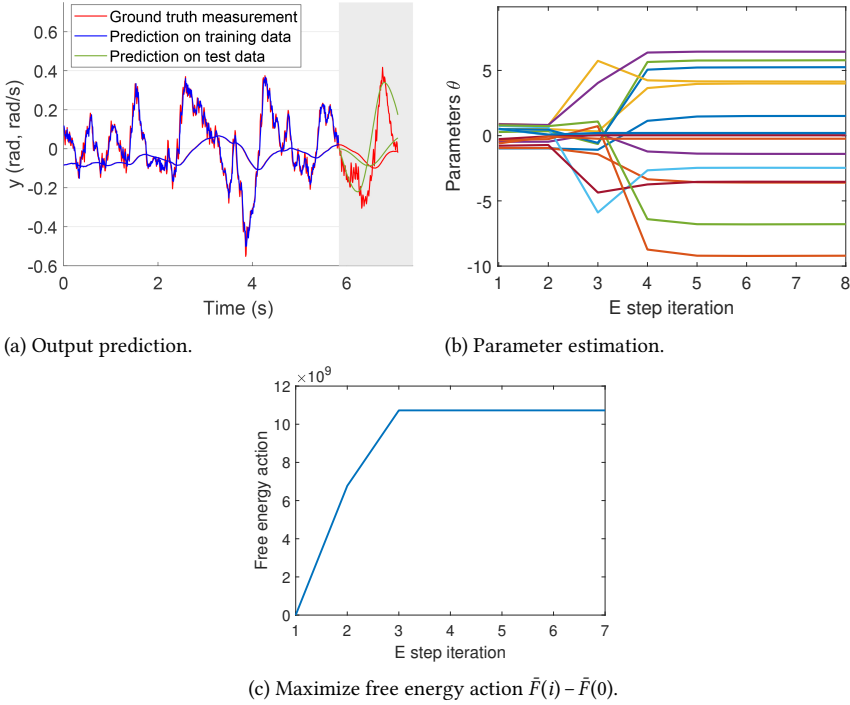
Fig 8.6: The robot brain's perception as a free energy optimization scheme (DEM) for the output prediction of a quadcopter hovering under wind conditions. (a) The coinciding blue and red curves demonstrate that DEM can accurately perform one step ahead output prediction on the training data (white background). The green curve following the trend of the red curve demonstrates that DEM can perform 150 step ahead output predictions on unseen data (grey background) using the learned model. (b) The parameter estimation step (E step) explores the parameter space to finally converge to a solution for A, B and C matrices. (c) Perception driven by the maximization of $\bar{F}$.

respectively on all experiments. Fig. 8.5 demonstrates that the histogram of process noise $w^{\dot{\phi}}$ is Gaussian in nature, confirming DEM's Laplace approximation. Fig. 8.7 demonstrates that the autocorrelation plot of $w^{\dot{\phi}}$ does not correspond to white noise where it should have been bounded within the confidence bounds for lags above 0. This confirms the presence of strong color in process noise, which was induced by unmodelled wind dynamics and linearization errors.

8.6.4 Algorithm settings for DEM

The parameter priors η^{θ} were randomly selected from [-1,1], and a moderate level of parameter precision ($P^{\theta} = e^4$) was set to encourage exploration in the parameter space, starting from random priors η^{θ}. A high observation noise hyperparameter ($\lambda^z = 20$) was used with high confidence ($P^{\lambda^z} = e^{25}$) to represent the accurate motion capture system measurements (optitrack). A low process noise hyperparameter ($\lambda^w = 3$) was used with

high confidence ($P^{\lambda^w} = e^{20}$) to represent high process noise emerging from wind and linearization errors. The noises were assumed to have a Gaussian temporal correlation with a noise smoothness of $s = dt$ for all the experiments. To handle colored noise, the generalized coordinate was used with an order of generalization for states and inputs as $p = 2$ and $d = 1$ respectively. DEM used the same settings to process all data.

8.6.5 Output prediction using DEM

The robot brain's perception of a quadrotor hovering in wind was emulated using the DEM algorithm. The quadrotor model was learned by maximizing $\bar{F}$ using the experiment 2 data and the result is shown in Fig. 8.6. The parameter estimation (E step) explores the parameter space and converges to a solution within few iterations (Fig. 8.6b), despite starting from wrong random priors (η^θ) in the range [-1,1]. The learned model is tested for one step ahead output prediction on the training data, and for output predictions until 150 step ahead on the test data. The coinciding predictions and measured output in Fig. 8.6a demonstrates DEM's successful model learning for output prediction, both on seen and unseen data. Fig. 8.6c shows the maximization of $\bar{F}$ during perception.

8.6.6 Metric for comparison

We measure the quality of output prediction using the Mean Squared Prediction Error (MSPE) for 150 step ahead predictions on unseen test data:

$$MSPE = \frac{1}{150} \sum_{i=T+1}^{T+150} (y_i - \hat{y}_i)^2, \tag{8.14}$$

where y_i is the measured output and $\hat{y}_i$ is the output prediction at time step i. A high quality perception algorithm will have the least MSPE when compared to other methods.
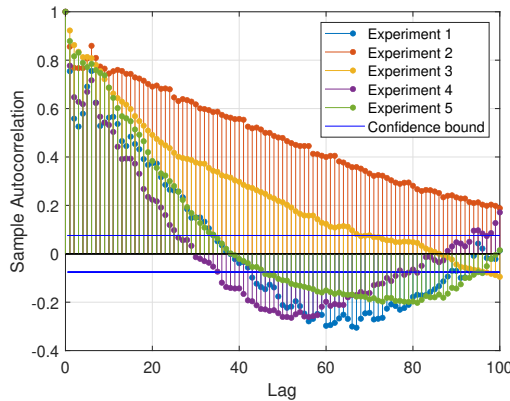
8

Fig 8.7: The autocorrelation plot of process noise ($w^{\dot{\phi}}$) for all five experiments doesn't drop within the confidence bound immediately after zero lag, confirming the presence of a range of noise color.

8.6.7 IMPORTANCE OF GENERALIZED COORDINATES

The key difference between DEM and other classical estimators is its capability to deal with colored noise using the generalized coordinates (GC). In this section, we show that the use of generalized coordinates improves the accuracy of output prediction of a quadrotor flying in wind. We repeat the same procedure in Section 8.6.5 for two different conditions: 1) output prediction with GC (p = 2) and 2) without GC (p = 1). Fig. 8.8 demonstrates that the use of GC provides a better output prediction than when no GC was used. MSPE was used to measure the quality of output prediction for all five flight experiments for both conditions (with and without GC), and the results are tabulated in Table 8.2. The results show a lower MSPE for DEM with GC when compared to DEM without using GC, revealing the importance of using GC for output prediction.

Table 8.2: INFLUENCE of GC on MSPE.

	expt 1	expt 2	expt 3	expt 4	expt 5	total
DEM without GC	0.1197	0.0640	0.1647	0.0518	0.0951	0.4953
DEM with GC	0.0521	0.0133	0.0583	0.0458	0.1172	**0.2867**

8.6.8 BENCHMARKING

In this section, we will show that DEM outperforms the classical system identification methods (SS, EM and PEM) with the least MSPE for five quadrotor flight experiments in wind. The System Identification toolbox from MATLAB was used for SS (*n4sid()*) and PEM (*pem()*) methods, whereas an EM algorithm implementation for state space model was written based on [234]. PEM was initialized using the solutions of SS. The data from experiment 2 was used to learn the state space model of the quadrotor for all methods. The grey area in Fig. 8.9a shows the results of output prediction on unseen test data, using the

8
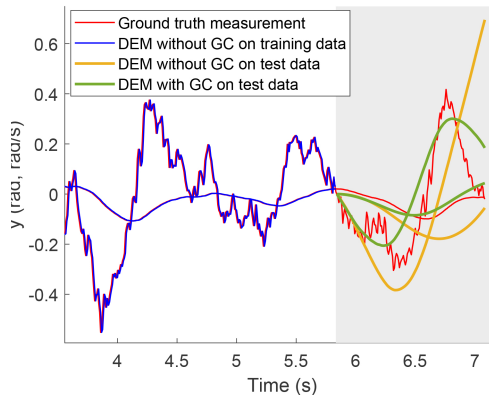
Fig 8.8: The output prediction of DEM on the test data improves when GC is used during model learning. The green curve follows the trend of the red curve better than the yellow curve.
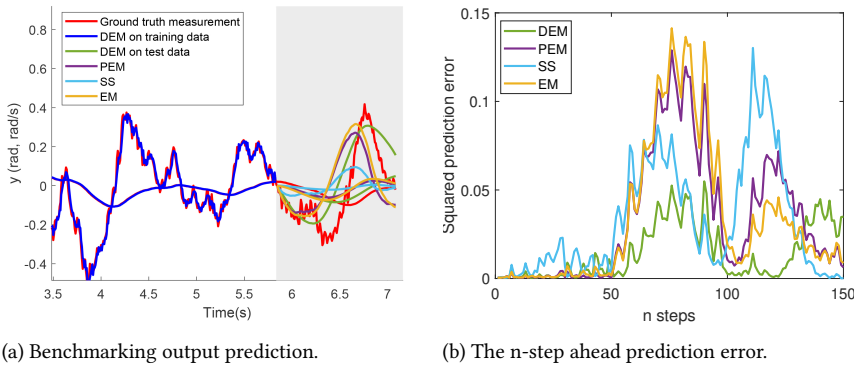
(a) Benchmarking output prediction.

(b) The n-step ahead prediction error.
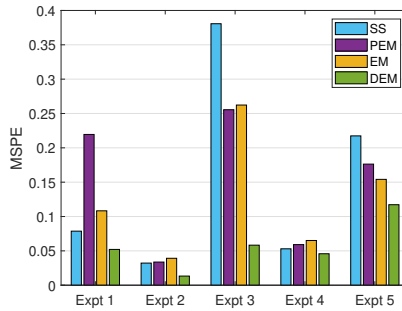
(c) MSPE for five quadrotor experiments.

Fig 8.9: The parameter learning using DEM is benchmarked against other system identification methods - PEM, SS and EM. (a) The output predictions of ϕ and $\dot{\phi}$ for experiment 2 on the unseen test data (in grey background). DEM (in green) best follows the trend of the ground truth measurement (in red) when compared to other methods. (b) A comparison of n-step ahead output prediction error (squared). DEM (green curve) has the lowest error, in comparison with PEM, SS and EM. (c) DEM outperforms other methods with the best quality output prediction by minimizing MSPE for all five experiments, thereby demonstrating that DEM is a very competitive algorithm.

model learned by the benchmarks. All predictions tend to follow the trend of the ground truth measurement (red curve). The prediction accuracy of different methods in Fig. 8.9a is visualized in Fig. 8.9b using the n step ahead squared prediction error. It can be observed that DEM outperforms other methods with the least prediction error on unseen data. MSPE was used as the evaluation metric to compare the performance of DEM with other methods on all five flight data, and the results are shown in Fig. 8.9c. DEM outperforms other methods for all five experiments with minimum MSPE on unseen test data.

8.6.9 Extended DEM for black-box estimation

The previous sections use the known inputs, outputs and model order for the output prediction, which differs from the biological brain's perception that do not have access to the real inputs and model order. Therefore, in this section we unleash the full capability

of DEM with unknown inputs, and then extend it by proposing a free energy objective based scheme to evaluate the model order for black box estimation. Since $\bar{F}$ is the sum of prediction errors for $\tilde{x}$, $\tilde{v}$, θ and λ and their entropies, it is intuitive for the correct model order (or above) to maximize $\bar{F}$. In this section, we test this hypothesis for quadrotor flights.

We consider the linearized model of the quadcopter dynamics given in [207] to derive 4 different LTI systems as given in Table 8.1. All systems are observable and controllable, and use motor pwm signals as the input. y, $\dot{y}$ ϕ and $\dot{\phi}$ were selected as states since they are the most influenced by the wind in y direction, thereby generating colored process noise in data. We use the same algorithm setup in Section 8.6.4 except for an additional constraint for unknown input ($\eta^v = 0$ with low precision $P^v = e^2$), to run DEM for all five experiments. The converged values of $\bar{F}$ for all experiments were recorded by assuming a model order of 1,2,3 and 4 for all systems. The average $\bar{F}$ of five experiments with different model orders for all four systems is shown in Fig. 8.10. $\bar{F}$ saturates when the model order matches the system order, proving that $\bar{F}$ is an indicator for model order selection. We use this idea to extend the original DEM algorithm for complete black box estimation as given in Algorithm 3. It generates an internal model via free energy maximization to estimate $\tilde{x}, \tilde{v}, \theta, \lambda$ and n^x that best explains the data.

Algorithm 3 Extended DEM - black box estimation

Initialize priors $\eta = \{\eta^v, P^v, \eta^\theta, P^\theta, \eta^\lambda, P^\lambda\}$
Initialize brain's model order $n^x = 0$
while $\bar{F}_b$ *not converged* **do**   ▷ model order
    Initialize $a \leftarrow 1$ and $\bar{F}_a \leftarrow -\infty$;
    $n^x \leftarrow n^x + 1$;   ▷ increment model order
    while θ *not converged* **do**
        for $t = 0$:Δt:T **do**
        | $\tilde{x}(t), \tilde{v}(t) \leftarrow$ D_STEP$(\tilde{y}(t), \theta, \lambda, n^x, \eta)$
        end
        while λ *not converged* **do**
        | $\lambda \leftarrow$ M_STEP$(\tilde{y}, \tilde{x}, \tilde{v}, \theta, \lambda, n^x, \eta)$
        end
        $\bar{F}_a \leftarrow$ Equation (8.13)   ▷ $\bar{F}$ at optimal precision
        if $\bar{F}_a > \bar{F}_{a-1}$ **then**   ▷ update θ if $\bar{F}$ increased
        | $\theta \leftarrow$ E_STEP$(\tilde{y}, \tilde{x}, \tilde{v}, \theta, \lambda, n^x, \eta)$
        end
        $a \leftarrow a + 1$
    end
    $\bar{F}_b = \bar{F}_{a-1}$
end
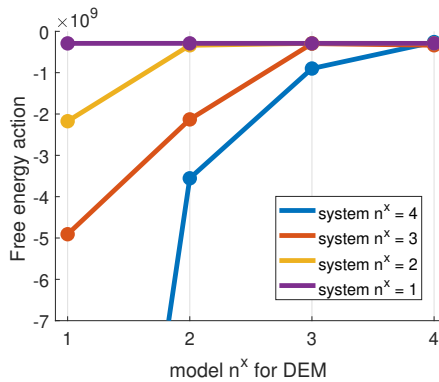
Fig 8.10: The average $\bar{F}$ of five experiments for different model orders. $\bar{F}$ saturates when the model order matches the real system order (4 for blue, 3 for red, 2 for yellow and 1 for violet).

8.7 Conclusion

System identification under colored noise is still an open challenge. We take a step towards developing a brain inspired model learning algorithm (DEM) for real robots. We introduced a DEM based system identification scheme under colored noise for a quadcopter flying in wind conditions. We demonstrated its superior performance through its minimum output prediction error, when compared to estimators like SS, PEM and EM. The usefulness of generalized coordinates in providing additional (derivative) information for estimation during unmodelled dynamics (wind) was demonstrated. Based on the results, the original DEM algorithm was extended for model order selection for complete black box estimation. The main disadvantage of DEM is that the noise smoothness needs to be known *apriori*, which can be addressed by the future research.

Acknowledgement

8

9

ACTIVE INFERENCE FOR MULTI ROBOT NAVIGATION IN FORMATION

*This chapter uses **precision modelling** to solve practical problems in multi robot navigation in formation, within the active inference framework. It builds on the idea of using precision to modulate robot behavior (from Chapter 2) in the context of multi robot problems.*

This chapter contains the preliminary results from an upcoming unpublished paper with the same title.

9.1 ABSTRACT

The last decade has witnessed the rise of active inference as a mathematical framework from neuroscience that describes the action selection procedure of the human brain as the minimization of an information theoretic objective called the free energy. In this paper, we introduce an active inference framework for multi robot navigation in formation problems, in a hope to move towards a brain inspired formation control algorithm for robot swarms. Through a series of simulations, we show that our active inference framework is effective in solving a number of multi robot navigation problems like goal directed navigation, static and dynamic obstacle avoidance and object transportation in formation. Finally, we show the utility of precision modulation within the active inference framework (as a mode of robot attention) in effectively solving multi robot problems like deadlock resolution in symmetric circular swapping of agents, splitting and merging of formation for obstacle avoidance and to negotiate through narrow gaps. With this framework, we contribute towards the development of active inference as the brain inspired swarm intelligence.

9.2 INTRODUCTION

Multi robot navigation is a well researched field in robotics [236] with a huge societal impact with applications including exploration [237], object transportation [238, 239] and pattern formation [240]. While most swarm intelligence researches have been inspired from the collective behavior of biological agent like ants, bees, birds and fishes [241], there is a gap in robotics literature for swarm intelligence from the cognitive neuroscience side [225]. With this work, we aim to take a step in this direction by using one of the most prominent neuroscience theory of the brain - the Free Energy Principle (FEP) [2] - to solve swarm robotics problems.

The recent developments in cognitive neuroscience have culminated in FEP, emerging as a unified theory of the brain [2]. The success of FEP in explaining a wide range of cognitive functions of the brain motivates its selection as the principled methodology for the development of a brain-inspired swarm intelligence algorithm. Based on FEP, active inference provides a mathematical framework to model the brain's action selection process as the minimization of an information theoretic measure called free energy [229]. The free energy minimization is a direct consequence of modelling the brain as a Bayesian engine that tries to do Bayesian inference using its sensory measurements [187]. The key idea behind active inference is that each agent in equilibrium with its environment tries to take those actions that would minimize its sensory surprisal, or in other words, the precision weighted prediction errors (and uncertainty). In the context of robot navigation, this is a plausible way to model the agent behavior because surprise minimization is central towards enabling robots with the capability to handle uncertainties (avoiding dynamic obstacles for example). This motivates the use of active inference for the design of our brain inspired multi robot navigation algorithm.

Owing to its wide popularity within the neuroscience community, active inference has been widely used to solve robotics challenges [25] like pick and place tasks [242], self-adaptation on humanoid robots [32], simultaneous localization and planning [34] etc. Our approach is different from other active inference based goal-directed obstacle avoidance algorithms mainly because of the use of a continuous time reactive navigation

scheme, instead of the discrete active inference solved as POMDPs [243] or the use of neural networks [35]. The novel and distinctive feature of our framework is the capability to solve multi robot navigation in formation problems in robotics using continuous time active inference [109]. The core contributions of the paper include:

1. introduce a continuous time active inference framework to solve the multi robot navigation in formation problem (Algorithm 4, Section 9.4, 9.5).

2. demonstrate the usefulness of precision modulated attention for multi robot navigation in formation problems (Section 9.6).

In Section 9.3, we introduce a continuous time active inference based reactive planner, which is used to solve the goal directed obstacle avoidance problem in Section 9.4, and the formation control problem in Section 9.5. The utility of precision modulation as a means to robot attention for multi robot problem is shown in Section 9.6.

9.3 Active inference based reactive planner

In this section, we introduce a general active inference framework for a reactive agent that takes control actions $a = \begin{bmatrix} a^x \\ a^y \end{bmatrix}$ in a 2D planar environment (x and y axis) such that it minimises its free energy. This section distinguishes the generative model from the generative process and builds a reactive planner. The generative process is part of the world where the actions are executed, while the agent's internal model of the generative process is called the generative model. The planner is part of the agent and is responsible for generating the control actions. The free energy is the only objective that shapes the agent's behavior. Here the planner and the generative model are part of the agent, while the generative process is part of the world. The agent can sense the world through measurements and can change the world by taking actions. This section works towards building the fundamentals of our active inference based reactive planner given in Algorithm 4.

9.3.1 Free energy

Active inference relies on action selection based on the minimization of an objective called the free energy, which derives from the Bayes rule, following the same routine as that of variational inference [52]. The generic form of the free energy objective used in this work can be seen as a sum of two terms: i) the precision weighted prediction error and ii) the entropy term. The combined free energy is written as [2]:

$$F = \frac{1}{2} \epsilon^T \Pi \epsilon - \ln |\Pi|, \tag{9.1}$$

where ϵ is the agent's prediction error and Π is the precision or inverse covariance associated with the prediction error. Here, ϵ is the difference between the sensory measurements and the agent's predictions about it, written as $\epsilon = measurement - prediction$. In active inference, predictions can be seen as the prior preferences - reference trajectory, control goal etc. In the later sections, different variations of this error will be used to encode different agent behaviors. Any deviation of the sensory measurements from the predictions will reflect in the free energy term. Therefore, taking those actions that minimizes F would

drive the agent towards the resolution of sensory surprisal, which is the core idea behind active inference. For continuous time problems, this is realised by doing gradient descent on the free energy surface for action selection. Since, we only deal with constant precision throughout this paper, the entropy term is constant and therefore dropped from F. Next sections will detail on how action influences the world through a generative process and how an agent can make predictions about the sensory measurements using a generative model and a planner.

9.3.2 Generative process

The dynamic system or process in the environment that is responsible for the data generation behind the sensory measurements is called the generative process. We use a simple linear time invariant state space representation to model the agent dynamics in the planar environment, given by:

$$
\begin{aligned}
\dot{X} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} &= -AX + U \\
&= -\begin{bmatrix} a^x & 0 \\ 0 & a^y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u^x \\ u^y \end{bmatrix},
\end{aligned}
\tag{9.2}
$$

where $X = \begin{bmatrix} x \\ y \end{bmatrix}$ is the combined position vector with x, y coordinates and U is the combined action vector of the agent. This equation determines the agent's dynamics, given the control actions. The exact discretization of the differential equation along x at time $t + dt$ can be written as:

$$
\begin{aligned}
x(t + dt) &= e^{-a^x dt} x(t) - \frac{1}{a^x}(e^{-a^x dt} - 1)u^x(t), \\
y(t + dt) &= e^{-a^y dt} y(t) - \frac{1}{a^y}(e^{-a^y dt} - 1)u^y(t).
\end{aligned}
\tag{9.3}
$$

This completes the generative process of an agent moving in a 2D planar environment. Since the equations along the x and y axes are similar in form, only those along x will be derived in the following sections.

9

9.3.3 Generative model

The agent tries to model the generative process by maintaining a generative model, which might not be the same as the generative process. We use the generative model that is same as that of the generative process, except for the knowledge of the parameters a^x and a^y, which are assumed to be 1. The generative model, maintaining the same dynamics as that of Equation 9.2 is given by:

$$
\begin{aligned}
\dot{x} &= -x + u^x, \\
x(t + dt) &= e^{-dt} x(t) - (e^{-dt} - 1)u^x(t).
\end{aligned}
\tag{9.4}
$$

This model will be used by the planner in the next section.

9.3.4 Reactive planner

We use a reactive planner based on the gradient descent of free energy, given by [109]:

$$\frac{du^x}{dt} = -k^u \frac{\partial F}{\partial u^x}, \tag{9.5}$$

where k^u is the learning rate. Since the action influences the free energy through the position, the planner is rewritten as:

$$\frac{du^x}{dt} = -k^u \frac{\partial F}{\partial x} \frac{\partial x}{\partial u^x} - k^u \frac{\partial F}{\partial y} \frac{\partial y}{\partial u^x}. \tag{9.6}$$

Since y is independent of u^x, $\frac{\partial y}{\partial u^x} = 0$. For simplicity, we assume that $\frac{\partial \dot{x}}{\partial u^x} = 0$. This yields $\frac{\partial x}{\partial u^x} = 1$ after differentiating the generative model given in Equation 9.4. From these results, the planner simplifies to:

$$\frac{du^x}{dt} = -k^u \frac{\partial F}{\partial x}. \tag{9.7}$$

We use Newton Gauss update scheme to evaluate discrete actions in steps:

$$u^x(t + \Delta t) = u^x(t) + \Delta u^x,$$
$$u^y(t + \Delta t) = u^y(t) + \Delta u^y,$$
$$\Delta u^x = (e^{-k^u \frac{\partial^2 F}{\partial x^2}|_t \Delta t} - 1)(\frac{\partial^2 F}{\partial x^2}|_t)^{-1} \frac{\partial F}{\partial x}|_t,$$
$$\Delta u^y = (e^{-k^u \frac{\partial^2 F}{\partial y^2}|_t \Delta t} - 1)(\frac{\partial^2 F}{\partial y^2}|_t)^{-1} \frac{\partial F}{\partial y}|_t. \tag{9.8}$$

These update rules plans the next control action such that the free energy of the agent is minimized via a gradient descent. The algorithmic form of our active inference based reactive planner is given in Algorithm 4. The next sections will use this reactive planner to perform goal directed obstacle avoidance and formation control.

9.4 Goal directed obstacle avoidance

This section aims to use the reactive planner in Algorithm 4 to embed goal directed obstacle avoidance behavior in agents, driven by free energy.

9.4.1 Goal directed behavior

A goal directed behavior can be embedded in an agent using active inference by using the goal position as the agent's prior, within the free energy objective [109]. Any deviation of the agent's position from its priors reflects in the free energy and drives the agent towards performing a goal directed behavior, similar to an attractor dynamics. Therefore, a goal directed behavior is a natural consequence of free energy minimization, resulting in actions that would drive the agent towards the least suprising state - the goal position. The component of free energy that drives the goal directed behavior is given by:

$$F^g = \frac{1}{2}(x - x^g)\Pi^g(x - x^g) + \frac{1}{2}(y - y^g)\Pi^g(y - y^g), \tag{9.9}$$

where Π^g is the precision associated with the goal directed behavior.

Algorithm 4 Active inference based reactive planner

for $t = 0 : \Delta t : T$ **do**

    Compute $F, \frac{\partial F}{\partial x}|_t, \frac{\partial^2 F}{\partial x^2}|_t$

    // *Compute next action*

    $\Delta u^x = (e^{-k^u \frac{\partial^2 F}{\partial x^2}|_t \Delta t} - 1)(\frac{\partial^2 F}{\partial x^2}|_t)^{-1} \frac{\partial F}{\partial x}|_t$

    $\Delta u^y = (e^{-k^u \frac{\partial^2 F}{\partial y^2}|_t \Delta t} - 1)(\frac{\partial^2 F}{\partial y^2}|_t)^{-1} \frac{\partial F}{\partial y}|_t$

    $u^x(t + \Delta t) = u^x(t) + \Delta u^x$

    $u^y(t + \Delta t) = u^y(t) + \Delta u^y$

    // *Take those actions in the world*

    $x(t + \Delta t) = e^{-a^x \Delta t} x(t) - \frac{1}{a^x}(e^{-a^x \Delta t} - 1) u^x(t + \Delta t)$

    $y(t + \Delta t) = e^{-a^y \Delta t} y(t) - \frac{1}{a^y}(e^{-a^y \Delta t} - 1) u^y(t \Delta t)$

end

9.4.2 Static obstacle avoidance

In this section we introduce an obstacle avoidance scheme into the active inference framework, and demonstrate its applicability for a number of drones performing goal direction behavior by avoiding all collisions. We combine the idea of Artificial Potential Field (APF) [244] with active inference by using the zero collision cost in the APF as the agent's expectation or prior. Any deviation from a zero collision cost in the APF would drive the agent towards an obstacle avoidance behavior. To ensure a smooth free energy landscape, we use a special form of APF that takes a Gaussian function form. The component of free energy that drives the obstacle avoidance behavior is defined as:

$$
\begin{aligned}
F^o &= \frac{1}{2} \sum_{j=1}^{m} \left[e^{-\frac{1}{2\sigma^o}((x - o_j^x)^2 + (y - o_j^y)^2)} - 0 \right] \Pi^o \\
&\qquad \left[e^{-\frac{1}{2\sigma^o}((x - o_j^x)^2 + (y - o_j^y)^2)} - 0 \right] \\
&= \frac{1}{2} \Pi^o \sum_{j=1}^{m} \left[e^{-\frac{1}{2\sigma^o}((x - o_j^x)^2 + (y - o_j^y)^2)} \right]^2,
\end{aligned}
\tag{9.10}
$$

where Π^o is the precision associated with the obstacle avoidance, σ^o is the hyperparameter of APF defining the required level of obstacle avoidance, o_j^x is the x position of obstacle j and m is the total number of obstacles. When the agent is sufficiently away from the obstacle, APF imposes zero collision cost.

9.4.3 Dynamic obstacle avoidance

In this section, we propose a position based dynamic obstacle avoidance strategy within the active inference framework, similar to the static obstacle avoidance in the previous section. The motivation behind not using a velocity based dynamic obstacle avoidance strategy is its redundancy because of the use of the gradients of free energy within the

active inference framework, which implicitly handles the velocity. The use of two gradients of F in Algorithm 4 helps in embedding the dynamic obstacle behavior into the active inference framework. The component of free energy that drives the obstacle avoidance behavior is defined as:

$$F^{oa} = \frac{1}{2}\Pi^{oa}\sum_{j=2}^{n}\left[e^{-\frac{1}{2\sigma^o}\|p-p^j\|^2}\right]^2,\tag{9.11}$$

where p is the position vector of the agent, p^j is the position vector of agent j and Π^{oa} is the precision associated with obstacle avoidance with other agents.
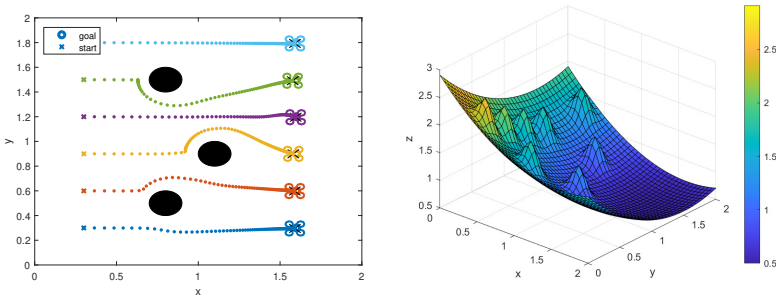
(a) Obstacle avoidance.

(b) 3D view of free energy curve
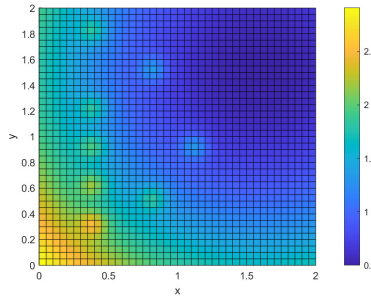
(c) Top view of free energy curve.

Fig 9.1: (a) The goal directed behavior of six active inference agents, successfully avoiding obstacles (in black) by going down their individual free energy curves for action selection. (b) 3D view of the free energy curve of agent 5 (green in Figure 9.1a) showing the peaks for all three static obstacles and five dynamic obstacles. (c) Top view of Figure 9.1b showing the minimum value at the goal position (1.6,1.5). The agent is pulled towards the goal while repelling the obstacles.

9.4.4 Goal directed obstacle avoidance

In this section, we show the navigation behavior of a group of active inference agents moving towards their individual goals in an obstacle filled environment. This is done by using Algorithm 4, by following a gradient descent on their free energy curves given by:

$$F = F^g + F^o + F^{oa}.\tag{9.12}$$

The free energy gradients in Algorithm 4 are computed using symbolic differentiation apriori. Figure 9.1 shows six active inference agents using the free energy curve to do obstacle avoidance during a goal directed behavior. Figures 9.1b and 9.1c shows the free energy curve of agent 5 (in green in Figure 9.1a). All agents successfully reaches the goal without any collisions with static and dynamic obstacles, showing the usefulness of the algorithm.

We take a step further by simulating the agent behavior in a challenging scenario that demands close contact between neighbouring agents. Figure 9.2 shows the asymmetric circular swapping of positions of 7 active inference agents, each following the gradients of their own free energy curve. Each agent avoids all other dynamic obstacles on their way towards the goal, without any collisions. This shows the usefulness of our framework to solve dynamic obstacle avoidance problems in robotics.
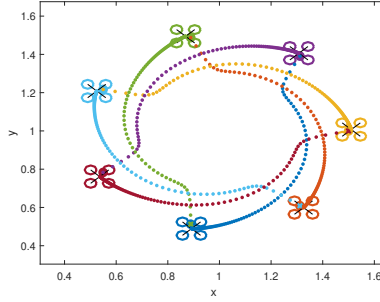
Fig 9.2: The final position and the path followed by a group of agents performing asymmetric circular swapping of positions using active inference. The agents initially move towards the centre. However, beyond a particular distance, the dynamic obstacles are too close that all the agents decide to move in a circular manner to negotiate towards their goals. This shows that active inference is effective in performing dynamic obstacle avoidance without any collisions.

9.5 Formation control

In this section we introduce a formation control scheme into the active inference framework for a group of agents to do multi robot navigation in formation tasks.

Some collaborative tasks in robotics like object transportation demands that all the agents maintains their formation throughout navigation so that the structural integrity of the transported object is maintained. We develop an active inference framework for these problems by shaping the free energy with expectations on inter agent distances. Each agent holds a prior expectation about the Euclidean distance it has to maintain with all other agents in the formation. Any deviation from this prior over distances will result in a corrective behavior by the agents, forcing them to remain in formation. The component of the free energy that drives agents towards formation control is defined as:

$$F^f = \frac{1}{2} \Pi^f \sum_{i=1}^{n} \sum_{j=i+1}^{n} \left(\|p^j - p^i\| - \|pf^j - pf^i\| \right)^2, \tag{9.13}$$

Fig 9.3: A group of five agents navigating in a circular formation towards their circular and linear shifted goal positions. The centroid of the formation marked in black follows a straight line between their start and goal centroid, indicating an unbroken formation throughout the navigation.

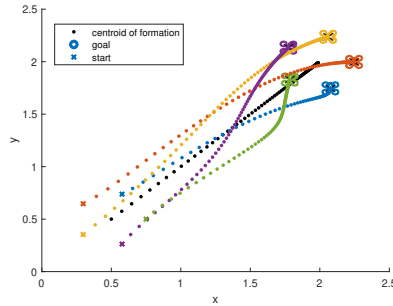where p^i and p^j is the position vector of agent i and j respectively, pf^j and pf^i are the initial position of the agents making a formation that is to be maintained, $\|.\|$ is the Euclidean distance, n is the number of agents, and Π^f is the precision associated with the formation. In summary, a group of agents that has to perform a goal directed behavior, by avoiding obstacles and maintaining the initial formation should use all three contributing free energy components defined as:

$$F = F^g + F^o + F^{oa} + F^f. \tag{9.14}$$

This free energy and its gradients are directly used in Algorithm 4, completing the reactive planner design.

We use a test case scenario to demonstrate the functionality of active inference to solve the multi robot object transportation task. We simulate such a scenario in Figure 9.3 where a group of agents tries to maintain the formation during a goal directed behavior towards the circular and linear shifted goal positions. The centroid of the formation, indicated in black follows a straight line between the centroid of the start and goal formations, indicating that the formation was maintained throughout the navigation. This shows the usefulness of active inference for multi robot collaboration problems.

9.6 PRECISION MODULATION FOR ROBOT ATTENTION

Precision modulation is central to attention models in neuroscience [19, 55, 245]. This section aims to show the utility of precision modulation in the context of robot attention for practically relevant multi robot problems like deadlock resolution, splitting and merging for obstacle avoidance and escape manoeuvre through narrow gaps.

9.6.1 PRECISION MODULATION FOR DEADLOCK RESOLUTION

The symmetric circular swapping of agents is a popular problem in multi agent systems that results in deadlocks. During deadlocks, the agents come face to face, resulting in a wastage of time while negotiating their way towards a deadlock resolution. In this section, we show the usefulness of precision modulation for deadlock resolution during multi robot
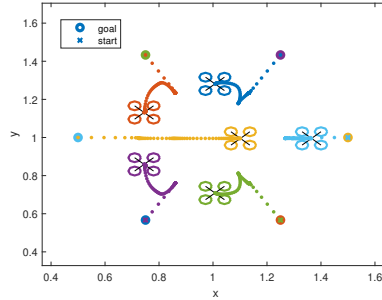
Fig 9.4: The symmetric circular swapping of agents resulting in a deadlock. The agents did not reach their goals within the stipulated time as most of it was wasted in coming face to face with other agents in a deadlock.
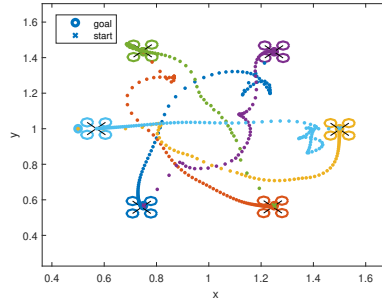
Fig 9.5: The use of precision modulation for deadlock resolution of agents in a symmetric circular swapping problem. Each agent uses a randomly selected goal precision $\Pi^g \in [0, 3]$ which determines the level of urgency of each agent to reach the goal. The agents with high Π^g chooses the easiest path towards the goal, while the agents with a low Π^g gives way to the agents in an urgency, thereby resolving the deadlock. The green agent is in urgency, while the yellow agent is in low urgency and gives way to other agents.

9

navigation. A challenging deadlock scenario is simulated in Figure 9.4 where an even number of agents try symmetric circular swapping and is unable to reach the goal in time.

We propose the use of precision modulation for the deadlock resolution, where each agent randomly picks a goal precision Π^g depending on their preference towards urgency (or the required attention level) to reach the goal, and communicates this information with other agents. The agents with a high Π^g are surprised more when they are far from the goal, when compared to the agents with low Π^g, This directly reflects in their velocity through the free energy term. Figure 9.5 shows the deadlock resolution through precision modulation, where some agents (green for example) act like in a sense of urgency, while other agents (yellow) gives way to other agents, thereby resolving the deadlock in Figure 9.4 and saving the time to reach the goal.
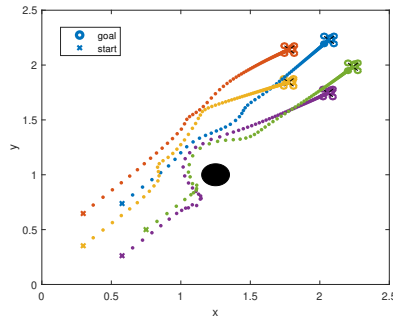
Fig 9.6: The obstacle avoidance behavior of a group of agents navigating in a circular formation during an object transportation task. A higher precision for formation control $\Pi^f = 5$ was used to attend more towards the formation control than towards the goal directed behavior.

9.6.2 Precision modulation for obstacle avoidance

This section demonstrates the role of precision modulation in splitting and merging behavior during obstacle avoidance tasks in multi robot navigation problems. We demonstrate a test case example where a group of agents navigating in formation shows two different behaviors just through precision modulation: i) obstacle avoidance by maintaining the formation and ii) obstacle avoidance by splitting and merging the formation. More the precision on formation control, the agent attends to keeping the formation. Figure 9.6 shows the goal directed behavior of a group of agents in formation, avoiding the obstacle in an object transportation task. The same setup results in an obstacle avoidance through splitting and merging when precision is modulated by changing Π^f from 5 to 2, as shown in Figure 9.7. The agent now attends more towards reaching the goal than towards keeping the formation. This demonstrates that precision modulation can be used to model emerging behavior during the robot navigation in formation.

9.6.3 Precision modulation for escape manoeuvre

This section aims to show the utility of precision modulation in tasks like escape manoeuvre of robots in formation where the formation has to operate in tight spaces. Figure 9.8a shows one such scenario where a group of robots should escape from a room through a window that is narrower than the formation size. These challenging scenarios results in splitting and merging of formation being the most effective solution for an escape maneuver. Through simulations, we show that the active inference agents can show such emerging behavior in formation without any hand coding for the expected behavior. A low Π^f enables such a behavior as shown in Figure 9.8b where the formation is temporarily split and each agent escapes in a narrow line through the gap, until all agents reach the other side and regroups to move towards their goal positions. This shows the usefulness of precision modulation in obtaining emerging behavior on robots moving in formation.

9
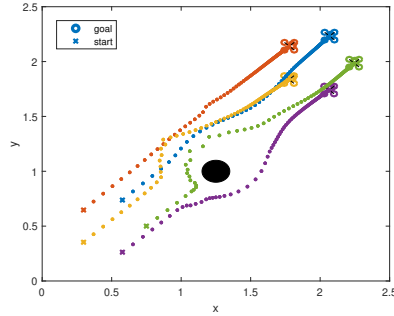
Fig 9.7: The exact same setting as that of Figure 9.6, except for a relaxed precision on formation control $\Pi^f = 2$. The violet agent autonomously decided to temporarily split from the formation and merge later. This behavior was not hard coded, but was an emergent behavior of the active inference framework, as a consequence of precision modulation. With a decreased Π^f, the agents now attend less to the formation control.
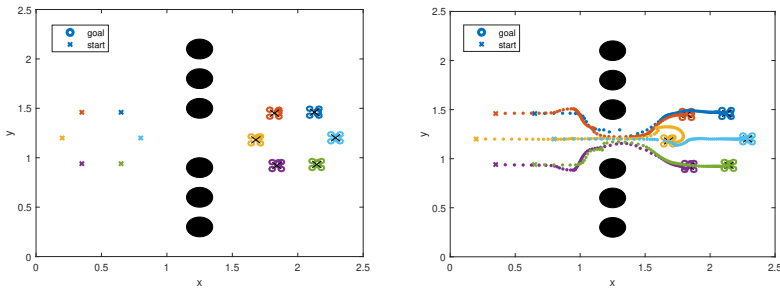
(a) Tight space scenario.

(b) Paths taken by the agents.
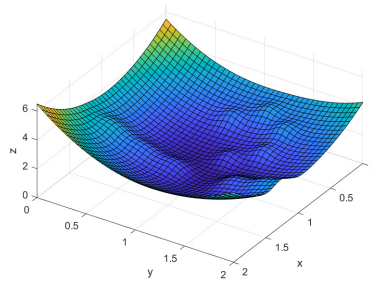
(c) Free energy curve of green agent.

Fig 9.8: (b) The active inference agents negotiating through a tight space (given in Figure 9.8a) to move towards their goal positions by splitting and merging the formation. This behavior emerges from lowering the formation precision Π^f as the formation approaches the obstacle wall, showing the usefulness of precision modulation in formation control. (c) The free energy curve of the green agent shows the obstacle wall and five other dynamic obstacles to be avoided. The agent takes actions to move down the free energy curve to reach its goal.

9.7 Conclusion

In this paper, we introduced an active inference framework to solve the multi robot naviga-
tion in formation problem. Through multiple simulations, we showed the effectiveness
of our algorithm in solving challenging problems in swarm robotics. We showed the use
of precision modulation within the active inference framework for deadlock resolution,
obstacle avoidance by splitting and merging in formation and escape manoeuvre through
narrow spaces. The main advantage of the algorithm is that it is real time and benefits from
the reactive nature that helps towards the online planning in formation control amongst
dynamic obstacles. However, a reactive planner might not be the best strategy for problems
that demands a finite look ahead. This could be solved by extending the algorithm using
expected free energy as the objective for path planning. The future research can focus on
implementing the algorithm on real robots and can include comparison against benchmarks.
The same framework could be used to design autonomous driving agents for road traffic
environments.

9

10

CONCLUSIONS AND FUTURE WORK

This thesis proposed a precision modulation based robot attention model using FEP - one of the most prominent brain theories from neuroscience. Four precision modulations were identified for robot attention to handle uncertainties: precision modelling, precision learning, precision optimization and precision fluctuation. The advantages of the model was explored within the scope of three class of robot problems: action, perception and active perception. Each chapter followed the development of the mathematical framework, extensive testing of the method in simulation with benchmarks, and its experimental validation on real robot data. These results were used together to answer the two main research questions of the thesis, showing that FEP is indeed useful to solve real robotics problems and that it can outperform classical robot algorithms, especially for estimation under colored noise. This chapter concludes the thesis by outlining the future research directions and ideas for extending this work.

10.1 CONCLUSIONS

This section lists the high-level conclusion of the research done in this thesis under each precision modulation method proposed in Chapter 2.

Precision modelling: In chapter 3, a state and input observer design for LTI systems under colored noise was proposed to seamlessly handle colored noise. Precision modelling was used to attend to the most informative noise derivatives during estimation. The resulting observer was shown to outperform classical linear observers in state and input estimation under colored noise. In chapter 4, an experimental setup was designed to test the effectiveness of the observer on real quadrotor flight data. The results demonstrated that the observer is highly competitive when compared to the state of the art methods, when applied to real robot data. In chapter 9, an active inference algorithm was introduced for multi robot navigation in formation. Using this algorithm, the precision modelling was used to attend to the desired robot behavior, where different strategies emerge during multi robot navigation in formation, when faced with an obstacle.

Precision learning: In chapter 6, a system identification tool for LTI systems under colored noise was developed through precision learning. The resulting estimator was proved to outperform classical estimators in simulation. The safe operation capabilities

of the algorithm on robots were motivated using a convergence proof for the parameter estimator in chapter 7. In chapter 8, the quadrotor hovering data was used to show that the estimator is competitive when applied to real robot data. In chapter 5, a novel online noise smoothness estimator was designed through precision (hyperparameter) learning. A joint state and noise smoothness estimator called DEMs was designed. DEMs was shown to outperform classical state estimation methods both using simulation data and real quadrotor data.

Precision optimization: In chapter 2, a general class of information gathering problems - especially for applications like target search, mapping, exploration etc - that relied on uncertainty reduction as the objective were identified as precision optimization. The path planner was shown to result in pure exploration strategies. Its application was shown on a sample IPP problem aimed at searching for human targets in an urban search and rescue.

Precision fluctuation: In chapter 2, a rhythmic precision modulated neurorobotic attention model was proposed to solve the active perception problems. The advantages of temporal scheduling of action and perception was shown within the context of the IPP problem, inspired from the rhythmic fluctuation between the eye saccades and the visual sampling of the eye on a reading task. The potential benefits of fluctuating the prior precision during system identification was highlighted in the context of an exploration-exploration tradeoff.

Combining all the results, the thesis answers the two main research questions. FEP is indeed useful to solve real robotics problems in estimation, control and planning. FEP based methods developed in the thesis outperform classical estimation algorithms in estimation accuracy, when the system is subjected to colored noise.

This thesis proposed a precision modulation based robot attention mechanism for uncertainty resolution in light of three pertinent sub-problems in robotics: action, perception and active perception. However, this proposition has a wider scope and can be expanded to accommodate a large class of robotics problems. The sub-problems addressed were limited in number and do not represent the full scale of the applicability of the model. For example, discrete problems that are typically solved using reinforcement leaning etc were not considered in this thesis. In the context of these problems, identifying an explicit precision term that needs to be modulated for attention is also a challenge. Although the neuroscience literature points to precision for attention modelling, this might not be feasible for all robotics problems because of the non existence of such a term. Therefore, more research should be done to identify the exact scope of this model, and to mathematically define alternatives for precision manipulation under these contexts.

10

10.2 Future work

This section lists the potential future research directions and further ideas to extend this work.

1. The distinguishing feature of DEM when compared to other methods is the use of generalized coordinates to model the colored noise in the system. This thesis considered only linear systems for estimation. However, the true potential of generalized coordinates could be exploited by evaluating DEM for nonlinear systems

using standard control system tools and techniques. This is an important research direction since a large part of the real world systems are non-linear in nature.

2. The usefulness of generalized coordinates for robot control still remains unexplored. Further research in this direction could result in better controller designs.

3. The practical use of the DEM based state and input observer design in Chapter 3 and 4 is for the delivery drones flying under strong wind conditions, which is of strong economic value to the industry. For this to be practical, the sensory measurements should rely on the internal signals of the drone, unlike the controlled lab experimental setup where the accurate Optitrack sensory data is available. Future research could focus on making this transition of the algorithm from the lab setup to real outdoor drone experiments.

4. The system identification method was not explored in the direction of identifiability, frequency domain analysis, robustness and persistent excitation. This research is essential for the widespread practical use of DEM as a system identification tool.

5. A stability proof for the DEMs observer developed in chapter 5 could be derived. The observer could be extended by learning the precision of the noise smoothness estimates, resulting in a more confident smoothness estimation and possibly a more accurate state estimation. The observer could also be extended by learning the magnitude of the noise precision online using the gradient ascent on the free energy curve. An accurate noise estimation is important for the widespread deployment of robots on highly uncertain environments.

6. The rhythmic precision modulated neurorobotic attention model was tested on the IPP model without using an explicit precision term, but by using the most informative future flight plan (path length). Further research in this direction to identify an explicit precision term that determines when to perform action and perception could make the robot search algorithm perform the search like a brain. This could help the roboticists to design a better search and rescue algorithm.

7. The precision optimization for IPP problem results in a pure exploration strategy, which is not favourable when there are constraints to be met. For example, obstacle avoidance, formation control etc. A Bayesian optimal way to do target search would be to use the full free energy as the objective function instead of just the precision, by encoding the constraints as prior preferences. This would result in an IPP algorithm that optimally balances between prior preferences and uncertainty resolution through exploration, thereby contributing towards the rise of an efficient search algorithm.

8. The preliminary results of the IPP problem points to the hypothesis that the free energy objective is the same as that of Bayesian optimization with Upper Confidence Bound (UCB) as the acquisition function. More research in this direction could be useful to test the mathematical and experimental validity of this hypothesis.

10

10.3 Discussion and insights

One of the main technical contributions of this thesis is the demonstration of the utility of generalized coordinates in robotics, especially for the estimation under colored noise. This thesis could contribute to the widespread use of generalized coordinates in solving real robot problems in future. Under colored noise, DEM outperforms the benchmarks in state estimation, system identification and noise estimation. Since building the robot algorithms that can perform tasks robustly in a highly noisy environment lies at the core of robotics, DEM could be highly impactful in the survival of the robot in a highly uncertain environment. The brain inspired nature of the mathematical framework further motivates its role in the rise of FEP as the central pillar of solving the general robot intelligence problem. If FEP is indeed the grand theory of the brain, then using it for the development of robot intelligence could lead to better algorithms that can seamlessly handle uncertainties. At the same time, it is prone to the pitfalls of the brain, giving rise to agents that are highly effective and robust against uncertainties in the world for survival, but are biased towards its prior beliefs, are heavily influenced by wrong data and data sources, and are prone to hyperactivity, attention deficits, conspiracy theories, hallucinations, quick generalizations, wrong judgements, prejudices, autistic behavior etc. It is highlighted that the uncertainty handling capabilities of the brain are embedded within the attention model and plays a crucial role in the survival of the organism. However, this might not lead to the global objective performance of the agent in the real world because it is evident that humans are mostly leading a suboptimal life that are highly biased by their environments, data sources and the prior genetic makeup. Therefore, care should be taken such that the future intelligent agents can escape both the pitfalls of the human brain and the suboptimal robot life.

With a hope that this thesis contributes to the rise of FEP as the grand unification theory for robot intelligence, this thesis is concluded.

10

BIBLIOGRAPHY

REFERENCES

[1] Peter Hänggi and Peter Jung. Colored noise in dynamical systems. *Advances in chemical physics*, 89:239–326, 2007.

[2] Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127, 2010.

[3] Daniel M Wolpert, Zoubin Ghahramani, and Michael I Jordan. An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882, 1995.

[4] David C Knill and Alexandre Pouget. The bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12):712–719, 2004.

[5] Rajesh PN Rao and Dana H Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1):79–87, 1999.

[6] Dirk De Ridder, Sven Vanneste, and Walter Freeman. The bayesian brain: phantom percepts resolve sensory uncertainty. *Neuroscience & Biobehavioral Reviews*, 44:4–15, 2014.

[7] Karl Friston. The history of the future of the bayesian brain. *NeuroImage*, 62(2):1230–1233, 2012.

[8] Karl Friston and Stefan Kiebel. Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society B: Biological sciences*, 364(1521):1211–1221, 2009.

[9] Michael W Spratling. A review of predictive coding algorithms. *Brain and cognition*, 112:92–97, 2017.

[10] M Berk Mirza, Rick A Adams, Christoph Mathys, and Karl J Friston. Human visual exploration reduces uncertainty about the sensed world. *PloS one*, 13(1):e0190429, 2018.

[11] Karl Friston, Rick Adams, Laurent Perrinet, and Michael Breakspear. Perceptions as hypotheses: saccades as experiments. *Frontiers in psychology*, 3:151, 2012.

[12] David Rudrauf, Daniel Bennequin, Isabela Granic, Gregory Landini, Karl Friston, and Kenneth Williford. A mathematical model of embodied consciousness. *Journal of theoretical biology*, 428:106–131, 2017.

[13] Rick A Adams, Laurent U Perrinet, and Karl Friston. Smooth pursuit and visual occlusion: active inference and oculomotor control in schizophrenia. *PloS one*, 7(10):e47502, 2012.

[14] James M Kilner, Karl J Friston, and Chris D Frith. Predictive coding: an account of the mirror neuron system. *Cognitive processing*, 8(3):159–166, 2007.

[15] Harriet Brown and Karl J Friston. Free-energy and illusions: the cornsweet effect. *Frontiers in psychology*, 3:43, 2012.

[16] J Allan Hobson and Karl J Friston. Waking and dreaming consciousness: neurobiological and functional considerations. *Progress in neurobiology*, 98(1):82–98, 2012.

[17] Karl J Friston, Tamara Shiner, Thomas FitzGerald, Joseph M Galea, Rick Adams, Harriet Brown, Raymond J Dolan, Rosalyn Moran, Klaas Enno Stephan, and Sven Bestmann. Dopamine, affordance and active inference. *PLoS computational biology*, 8(1):e1002327, 2012.

[18] Laurent Itti and Christof Koch. Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.

[19] Harriet Feldman and Karl Friston. Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*, 4:215, 2010.

[20] Thomas Parr and Karl J Friston. Attention or salience? *Current opinion in psychology*, 29:1–5, 2019.

[21] Thomas Parr and Karl J Friston. Working memory, attention, and salience in active inference. *Scientific reports*, 7(1):1–21, 2017.

[22] Jakub Limanowski. Precision control for a flexible body representation. *Neuroscience and Biobehavioral Reviews*, 134:104401, 2022.

[23] Rosalyn J Moran, Pablo Campo, Mkael Symmonds, Klaas E Stephan, Raymond J Dolan, and Karl J Friston. Free energy, precision and learning: the role of cholinergic neuromodulation. *Journal of Neuroscience*, 33(19):8227–8236, 2013.

[24] Simone Vossel, Christoph Mathys, Jean Daunizeau, Markus Bauer, Jon Driver, Karl J Friston, and Klaas E Stephan. Spatial attention, precision, and bayesian inference: a study of saccadic response speed. *Cerebral cortex*, 24(6):1436–1450, 2014.

[25] Pablo Lanillos, Cristian Meo, Corrado Pezzato, Ajith Anil Meera, Mohamed Baioumy, Wataru Ohata, Alexander Tschantz, Beren Millidge, Martijn Wisse, Christopher L Buckley, et al. Active inference in robotics and artificial agents: Survey and challenges. *arXiv preprint arXiv:2112.01871*, 2021.

[26] Alejandra Ciria, Guido Schillaci, Giovanni Pezzulo, Verena V Hafner, and Bruno Lara. Predictive processing in cognitive robotics: a review. *Neural Computation*, 33(5):1402–1432, 2021.

[27] Corrado Pezzato, Riccardo Ferrari, and Carlos Hernández Corbato. A novel adaptive controller for robot manipulators based on active inference. *IEEE Robotics and Automation Letters*, 5(2):2973–2980, 2020.

[28] Cristian Meo and Pablo Lanillos. Multimodal vae active inference controller. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2693–2699. IEEE, 2021.

[29] Jeffrey Frederic QueiSSer, Minju Jung, Takazumi Matsumoto, and Jun Tani. Emergence of content-agnostic information processing by a robot using active inference, visual attention, working memory, and planning. *Neural Computation*, 33(9):2353–2407, 2021.

[30] Nadine Wirkuttis and Jun Tani. Leading or following? dyadic robot imitative interaction using the active inference framework. *IEEE Robotics and Automation Letters*, 6(3):6024–6031, 2021.

[31] Cansu Sancaktar, Marcel AJ van Gerven, and Pablo Lanillos. End-to-end pixel-based deep active inference for body perception and action. In *2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, pages 1–8. IEEE, 2020.

[32] Guillermo Oliver, Pablo Lanillos, and Gordon Cheng. An empirical study of active inference on a humanoid robot. *IEEE Transactions on Cognitive and Developmental Systems*, 2021.

[33] Takato Horii and Yukie Nagai. Active inference through energy minimization in multimodal affective human–robot interaction. *Frontiers in Robotics and AI*, 8, 2021.

[34] Ozan Çatal, Tim Verbelen, Toon Van de Maele, Bart Dhoedt, and Adam Safron. Robot navigation as hierarchical active inference. *Neural Networks*, 142:192–204, 2021.

[35] Takazumi Matsumoto, Wataru Ohata, Fabien CY Benureau, and Jun Tani. Goal-directed planning and goal understanding by extended active inference: Evaluation through simulated and physical robot experiments. *Entropy*, 24(4):469, 2022.

[36] Louis Annabi, Alexandre Pitti, and Mathias Quoy. Bidirectional interaction between visual and motor generative models using predictive coding and active inference. *Neural Networks*, 143:638–656, 2021.

[37] Alexander Tschantz, Beren Millidge, Anil K Seth, and Christopher L Buckley. Reinforcement learning through active inference. *arXiv preprint arXiv:2002.12636*, 2020.

[38] WouterM Kouw, Albert Podusenko, MagnusT Koudahl, and Maarten Schoukens. Variational message passing for online polynomial narmax identification. In *2022 American Control Conference (ACC)*, pages 2755–2760. IEEE, 2022.

[39] Thijs van de Laar, Ayça Özçelikkale, and Henk Wymeersch. Application of the free energy principle to estimation and control. *IEEE Transactions on Signal Processing*, 69:4234–4244, 2021.

[40] Marco Cox, Thijs van de Laar, and Bert de Vries. A factor graph approach to automated design of bayesian signal processing algorithms. *International Journal of Approximate Reasoning*, 104:185–204, 2019.

[41] Ke Gu, Guangtao Zhai, Xiaokang Yang, and Wenjun Zhang. Using free energy principle for blind image quality assessment. *IEEE Transactions on Multimedia*, 17(1):50–63, 2014.

[42] Karl J Friston, Jean Daunizeau, and Stefan J Kiebel. Reinforcement learning or active inference? *PloS one*, 4(7):e6421, 2009.

[43] Christopher L Buckley, Chang Sub Kim, Simon McGregor, and Anil K Seth. The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology*, 81:55–79, 2017.

[44] Karl Friston. Hierarchical models in the brain. *PLoS computational biology*, 4(11):e1000211, 2008.

[45] Manuel Baltieri and Takuya Isomura. Kalman filters as the steady-state solution of gradient descent on variational free energy. *arXiv preprint arXiv:2111.10530*, 2021.

[46] Ajith Anil Meera and Martijn Wisse. Free energy principle based state and input observer design for linear systems with colored noise. In *2020 American Control Conference (ACC)*, pages 5052–5058. IEEE, 2020.

[47] Manuel Baltieri and Christopher L Buckley. Pid control as a process of active inference with linear generative models. *Entropy*, 21(3):257, 2019.

[48] Manuel Baltieri. A bayesian perspective on classical control. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2020.

[49] Manuel Baltieri and Christopher L Buckley. On kalman-bucy filters, linear quadratic control and active inference. *arXiv preprint arXiv:2005.06269*, 2020.

[50] Karl J Friston, Jean Daunizeau, James Kilner, and Stefan J Kiebel. Action and behavior: a free-energy formulation. *Biological cybernetics*, 102(3):227–260, 2010.

[51] Sherin Grimbergen. The state space formulation of active inference: Towards brain-inspired robot control. *Masters thesis, TU Delft*, 2019.

[52] David M Blei, Alp Kucukelbir, and Jon D McAuliffe. Variational inference: A review for statisticians. *Journal of the American statistical Association*, 112(518):859–877, 2017.

[53] Ajith Anil Meera, Marija Popović, Alexander Millane, and Roland Siegwart. Obstacle-aware adaptive informative path planning for uav-based target search. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 718–724. IEEE, 2019.

[54] Ajith Anil Meera. Informative path planning for search and rescue using a uav. *Maters thesis, TU Delft*, 2018.

[55] Ajith Anil Meera, Filip Novicky, Thomas Parr, Karl Friston, Pablo Lanillos, and Noor Sajid. Reclaiming saliency: Rhythmic precision-modulated action and perception. *Frontiers in Neurorobotics*, 16, 2022.

[56] PK Ahnelt. The photoreceptor mosaic. *Eye*, 12(3):531–540, 1998.

[57] Scott Cheng-Hsin Yang, Mate Lengyel, and Daniel M Wolpert. Active sensing in the categorization of visual patterns. *Elife*, 5:e12215, 2016.

[58] Máté Lengyel, Scott Cheng-Hsin Yang, and Daniel M Wolpert. Active sensing in the categorization of visual patterns. *eLife*, 2016.

[59] Valerio Santangelo. Forced to remember: when memory is biased by salient information. *Behavioural brain research*, 283:1–10, 2015.

[60] P Christiaan Klink, Pia Jentgens, and Jeannette AM Lorteije. Priority maps explain the roles of value, attention, and salience in goal-oriented behavior. *Journal of Neuroscience*, 34(42):13867–13869, 2014.

[61] Sajad Mousavi, Michael Schukat, Enda Howley, Ali Borji, and Nasser Mozayani. Learning to predict where to look in interactive environments using deep recurrent q-learning. *arXiv preprint arXiv:1612.05753*, 2016.

[62] Akanksha Atrey, Kaleigh Clary, and David Jensen. Exploratory not explanatory: Counterfactual analysis of saliency maps for deep reinforcement learning. *arXiv preprint arXiv:1912.05743*, 2019.

[63] Simone Frintrop and Patric Jensfelt. Attentional landmarks and active gaze control for visual slam. *IEEE Transactions on Robotics*, 24(5):1054–1065, 2008.

[64] Momotaz Begum and Fakhri Karray. Visual attention for robotic cognition: A survey. *IEEE Transactions on Autonomous Mental Development*, 3(1):92–105, 2010.

[65] Joao Filipe Ferreira and Jorge Dias. Attentional mechanisms for socially interactive robots–a survey. *IEEE Transactions on Autonomous Mental Development*, 6(2):110–125, 2014.

[66] Pablo Lanillos, João Filipe Ferreira, and Jorge Dias. Designing an artificial attention system for social robots. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4171–4178. IEEE, 2015.

[67] Anne M Treisman and Garry Gelade. A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136, 1980.

[68] John K Tsotsos, Scan M Culhane, Winky Yan Kei Wai, Yuzhong Lai, Neal Davis, and Fernando Nuflo. Modeling visual attention via selective tuning. *Artificial intelligence*, 78(1-2):507–545, 1995.

[69] Ali Borji and Laurent Itti. State-of-the-art in visual attention modeling. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):185–207, 2012.

[70] Dennis V Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956.

[71] Karl Friston, Lancelot Da Costa, Danijar Hafner, Casper Hesp, and Thomas Parr. Sophisticated inference. *Neural Computation*, 33(3):713–763, 2021.

[72] Noor Sajid, Lancelot Da Costa, Thomas Parr, and Karl Friston. Active inference, bayesian optimal design, and expected utility. *arXiv preprint arXiv:2110.04074*, 2021.

[73] Ruzena Bajcsy, Yiannis Aloimonos, and John K Tsotsos. Revisiting active perception. *Autonomous Robots*, 42(2):177–196, 2018.

[74] Scott Cheng-Hsin Yang, Daniel M Wolpert, and Máté Lengyel. Theoretical perspectives on active sensing. *Current opinion in behavioral sciences*, 11:100–108, 2016.

[75] Giacomo Rizzolatti, Lucia Riggio, Isabella Dascola, and Carlo Umiltá. Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25(1):31–40, 1987.

[76] Robert Desimone. Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences*, 93(24):13494–13499, 1996.

[77] M. Berk Mirza, Rick A. Adams, Christoph D. Mathys, and Karl J. Friston. Scene construction, visual foraging, and active inference. *Frontiers in Computational Neuroscience*, 10, 2016.

[78] Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. Active inference and epistemic value. *Cognitive Neuroscience*, 6(4):187–214, 2015. PMID: 25689102.

[79] Nancy Kanwisher and Ewa Wojciulik. Visual attention: insights from brain imaging. *Nature reviews neuroscience*, 1(2):91–100, 2000.

[80] Eran Eldar, Jonathan D Cohen, and Yael Niv. The effects of neural gain on attention and learning. *Nature neuroscience*, 16(8):1146–1153, 2013.

[81] John H Reynolds, Tatiana Pasternak, and Robert Desimone. Attention increases sensitivity of v4 neurons. *Neuron*, 26(3):703–714, 2000.

[82] Thomas Parr and Karl J Friston. Uncertainty, epistemics and active inference. *Journal of The Royal Society Interface*, 14(136):20170376, 2017.

[83] Rajesh PN Rao. Bayesian inference and attentional modulation in the visual cortex. *Neuroreport*, 16(16):1843–1848, 2005.

[84] Michael W Spratling. Predictive coding as a model of biased competition in visual attention. *Vision research*, 48(12):1391–1408, 2008.

[85] Thomas Parr, David A Benrimoh, Peter Vincent, and Karl J Friston. Precision and false perceptual inference. *Frontiers in integrative neuroscience*, 12:39, 2018.

[86] Louise Whiteley and Maneesh Sahani. Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *Journal of vision*, 8(3):2–2, 2008.

[87] Harriet Brown, Rick A Adams, Isabel Parees, Mark Edwards, and Karl Friston. Active inference, sensory attenuation and illusions. *Cognitive processing*, 14(4):411–427, 2013.

[88] Ryota Kanai, Yutaka Komura, Stewart Shipp, and Karl Friston. Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668):20140169, 2015.

[89] Andy Clark. The many faces of precision (replies to commentaries on "whatever next? neural prediction, situated agents, and the future of cognitive science"). *Frontiers in Psychology*, 4, 2013.

[90] John H Reynolds and David J Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009.

[91] Douglas A Ruff and Marlene R Cohen. Stimulus dependence of correlated variability across cortical areas. *Journal of Neuroscience*, 36(28):7546–7556, 2016.

[92] M Berk Mirza, Rick A Adams, Karl Friston, and Thomas Parr. Introducing a bayesian model of selective attention based on active inference. *Scientific reports*, 9(1):1–22, 2019.

[93] Thomas Parr, Andrew W Corcoran, Karl J Friston, and Jakob Hohwy. Perceptual awareness and active inference. *Neuroscience of consciousness*, 2019(1):niz012, 2019.

[94] Noor Sajid, Francesco Faccio, Lancelot Da Costa, Thomas Parr, Jürgen Schmidhuber, and Karl Friston. Bayesian brains and the r\'enyi divergence. *arXiv preprint arXiv:2107.05438*, 2021.

[95] Noor Sajid, Karl J Friston, Justyna O Ekert, Cathy J Price, and David W Green. Neuromodulatory control and language recovery in bilingual aphasia: An active inference approach. *Behavioral Sciences*, 10(10):161, 2020.

[96] Karl J Friston, Lee Harrison, and Will Penny. Dynamic causal modelling. *Neuroimage*, 19(4):1273–1302, 2003.

[97] Ajith Anil Meera and Martijn Wisse. Dynamic expectation maximization algorithm for estimation of linear systems with colored noise. *Entropy*, 23(10):1306, 2021.

[98] David J Heeger. Normalization of cell responses in cat striate cortex. *Visual neuroscience*, 9(2):181–197, 1992.

[99] Kenway Louie and Paul W Glimcher. Normalization principles in computational neuroscience. In *Oxford Research Encyclopedia of Neuroscience*. 2019.

[100] Taylor W Schmitz and John Duncan. Normalization and the cholinergic microcircuit: a unified basis for attention. *Trends in cognitive sciences*, 22(5):422–437, 2018.

[101] Brian J White, David J Berg, Janis Y Kan, Robert A Marino, Laurent Itti, and Douglas P Munoz. Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nature communications*, 8(1):1–9, 2017.

[102] MICHAEL S Raybourn and EDWARD L Keller. Colliculoreticular organization in primate oculomotor system. *Journal of Neurophysiology*, 40(4):861–878, 1977.

[103] Laurent Itti and Pierre Baldi. Bayesian surprise attracts human attention. *Vision research*, 49(10):1295–1306, 2009.

[104] Pierre-Yves Oudeyer and Frederic Kaplan. What is intrinsic motivation? a typology of computational approaches. *Frontiers in neurorobotics*, 1:6, 2009.

[105] M Berk Mirza, Rick A Adams, Christoph D Mathys, and Karl J Friston. Scene construction, visual foraging, and active inference. *Frontiers in computational neuroscience*, 10:56, 2016.

[106] Karl Friston, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. Active inference: a process theory. *Neural computation*, 29(1):1–49, 2017.

[107] Lancelot Da Costa, Thomas Parr, Noor Sajid, Sebastijan Veselic, Victorita Neacsu, and Karl Friston. Active inference on discrete state-spaces: a synthesis. *Journal of Mathematical Psychology*, 99:102447, 2020.

[108] Raymond M Klein. Inhibition of return. *Trends in cognitive sciences*, 4(4):138–147, 2000.

[109] Karl Friston, Jérémie Mattout, and James Kilner. Action understanding and active inference. *Biological cybernetics*, 104(1):137–160, 2011.

[110] Frederic Crevecoeur and Konrad P Kording. Saccadic suppression as a perceptual consequence of efficient sensorimotor estimation. *Elife*, 6:e25073, 2017.

[111] Ian C Fiebelkorn and Sabine Kastner. A rhythmic theory of attention. *Trends in cognitive sciences*, 23(2):87–101, 2019.

[112] Ian C Fiebelkorn and Sabine Kastner. Spike timing in the attention network predicts behavioral outcome prior to target selection. *Neuron*, 109(1):177–188, 2021.

[113] H von Helmholtz. Handbook of physiological optics southall. *JPC The Optical Society of America*, 3, 1925.

[114] Michele Rucci, Ehud Ahissar, and David Burr. Temporal coding of visual space. *Trends in cognitive sciences*, 22(10):883–895, 2018.

[115] Alessandro Benedetto, Maria Concetta Morrone, and Alice Tomassini. The common rhythm of action and perception. *Journal of cognitive neuroscience*, 32(2):187–200, 2020.

[116] Alice Tomassini, Luca Ambrogioni, W Pieter Medendorp, and Eric Maris. Theta oscillations locked to intended actions rhythmically modulate perception. *Elife*, 6:e25618, 2017.

[117] Hinze Hogendoorn. Voluntary saccadic eye movements ride the attentional rhythm. *Journal of cognitive neuroscience*, 28(10):1625–1635, 2016.

[118] Alessandro Benedetto and Maria Concetta Morrone. Saccadic suppression is embedded within extended oscillatory modulation of sensitivity. *Journal of Neuroscience*, 37(13):3661–3670, 2017.

[119] Ian C Fiebelkorn, Mark A Pinsk, and Sabine Kastner. A dynamic interplay within the frontoparietal network underlies rhythmic spatial attention. *Neuron*, 99(4):842–853, 2018.

[120] Randolph F Helfrich, Ian C Fiebelkorn, Sara M Szczepanski, Jack J Lin, Josef Parvizi, Robert T Knight, and Sabine Kastner. Neural mechanisms of sustained attention are rhythmic. *Neuron*, 99(4):854–865, 2018.

[121] Ian C Fiebelkorn and Sabine Kastner. Functional specialization in the attention network. *Annual review of psychology*, 71:221–249, 2020.

[122] Karl J. Friston, Thomas Parr, Yan Yufik, Noor Sajid, Catherine J. Price, and Emma Holmes. Generative models, linguistic communication and active inference. *Neuroscience & Biobehavioral Reviews*, 118:42–64, 2020.

[123] Thomas Parr and Giovanni Pezzulo. Understanding, explanation, and active inference. *Frontiers in Systems Neuroscience*, 15, 2021.

[124] Noor Sajid, Emma Holmes, Lancelot Da Costa, Cathy Price, and Karl Friston. A mixed generative model of auditory word repetition. *bioRxiv*, 2022.

[125] Marc A Sommer and Robert H Wurtz. Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, 444(7117):374–377, 2006.

[126] Miho Nakajima, L Ian Schmitt, and Michael M Halassa. Prefrontal cortex regulates sensory filtering through a basal ganglia-to-thalamus pathway. *Neuron*, 103(3):445–458, 2019.

[127] Jessica M Phillips, Niranjan A Kambi, and Yuri B Saalmann. A subcortical pathway for rapid, goal-driven, attentional filtering. *Trends in neurosciences*, 39(2):49–51, 2016.

[128] Ian C Fiebelkorn, Mark A Pinsk, and Sabine Kastner. The mediodorsal pulvinar coordinates the macaque fronto-parietal network during rhythmic spatial attention. *Nature communications*, 10(1):1–15, 2019.

[129] Niko A Busch and Rufin VanRullen. Spontaneous eeg oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, 107(37):16048–16053, 2010.

[130] Laura Dugué, Philippe Marque, and Rufin VanRullen. Theta oscillations modulate attentional search performance periodically. *Journal of cognitive neuroscience*, 27(5):945–958, 2015.

[131] Laura Dugué, Mariel Roberts, and Marisa Carrasco. Attention reorients periodically. *Current Biology*, 26(12):1595–1601, 2016.

[132] Rufin VanRullen. Perceptual rhythms. *Stevens Handbook of Experimental Psychology*, 1, 2016.

[133] Thomas Parr, Noor Sajid, Lancelot Da Costa, M Berk Mirza, and Karl J Friston. Generative models for active vision. *Frontiers in Neurorobotics*, 15:34, 2021.

[134] Liang-Tien Hsieh and Charan Ranganath. Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage*, 85:721–729, 2014.

[135] Ulrich Pomper and Ulrich Ansorge. Theta-rhythmic oscillation of working memory performance. *Psychological Science*, page 09567976211013045, 2021.

[136] Aneta Brzezicka, Jan Kamiński, Chrystal M Reed, Jeffrey M Chung, Adam N Mamelak, and Ueli Rutishauser. Working memory load-related theta power decreases in dorsolateral prefrontal cortex predict individual differences in performance. *Journal of cognitive neuroscience*, 31(9):1290–1307, 2019.

[137] Elio Balestrieri, Luca Ronconi, and David Melcher. Shared resources between visual attention and visual working memory are allocated through rhythmic sampling. *European Journal of Neuroscience*, 2021.

[138] Nikolai Axmacher, Melanie M Henseler, Ole Jensen, Ilona Weinreich, Christian E Elger, and Juergen Fell. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, 107(7):3228–3233, 2010.

[139] Benjamin Peters, Jochen Kaiser, Benjamin Rahm, and Christoph Bledowski. Object-based attention prioritizes working memory contents at a theta rhythm. *Journal of Experimental Psychology: General*, 2020.

[140] Moritz Köster, Holger Finger, Sebastian Graetz, Maren Kater, and Thomas Gruber. Theta-gamma coupling binds visual perceptual features in an associative memory task. *Scientific reports*, 8(1):1–9, 2018.

[141] Valerio Santangelo and Emiliano Macaluso. Visual salience improves spatial working memory via enhanced parieto-temporal functional connectivity. *Journal of Neuroscience*, 33(9):4110–4117, 2013.

[142] Valerio Santangelo, Simona Arianna Di Francesco, Serena Mastroberardino, and Emiliano Macaluso. Parietal cortex integrates contextual and saliency signals during the encoding of natural scenes in working memory. *Human Brain Mapping*, 36(12):5003–5017, 2015.

[143] Michael S Fine and Brandon S Minnery. Visual salience affects performance in a working memory task. *Journal of Neuroscience*, 29(25):8016–8021, 2009.

[144] Adam Gazzaley and Anna C Nobre. Top-down modulation: bridging selective attention and working memory. *Trends in cognitive sciences*, 16(2):129–135, 2012.

[145] Eric I Knudsen. Fundamental components of attention. *Annu. Rev. Neurosci.*, 30:57–78, 2007.

[146] Klaus Oberauer. Working memory and attention–a conceptual analysis and review. *Journal of cognition*, 2(1), 2019.

[147] Matthew F Panichello and Timothy J Buschman. Shared mechanisms underlie the control of working memory and attention. *Nature*, 592(7855):601–605, 2021.

[148] Ryohei Nakayama and Isamu Motoyoshi. Attention periodically binds visual features as single events depending on neural oscillations phase-locked to action. *Journal of Neuroscience*, 39(21):4153–4161, 2019.

[149] Zoya Bylinskii, Tilke Judd, Ali Borji, Laurent Itti, Frédo Durand, Aude Oliva, and Antonio Torralba. Mit saliency benchmark. http://saliency.mit.edu/, 2019.

[150] Francesco Orabona, Giorgio Metta, and Giulio Sandini. Object-based visual attention: a model for a behaving robot. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, pages 89–89. IEEE, 2005.

[151] Simone Frintrop. *VOCUS: A visual attention system for object detection and goal-directed search*, volume 3899. Springer, 2006.

[152] Ayoung Kim and Ryan M Eustice. Real-time visual slam for autonomous underwater hull inspection using visual saliency. *IEEE Transactions on Robotics*, 29(3):719–733, 2013.

[153] Richard Roberts, Duy-Nguyen Ta, Julian Straub, Kyel Ok, and Frank Dellaert. Saliency detection and model-based tracking: a two part vision system for small robot navigation in forested environment. In *Unmanned Systems Technology XIV*, volume 8387, page 83870S. International Society for Optics and Photonics, 2012.

[154] Nicholas J Butko, Lingyun Zhang, Garrison W Cottrell, and Javier R Movellan. Visual saliency model for robot cameras. In *2008 IEEE International Conference on Robotics and Automation*, pages 2398–2403. IEEE, 2008.

[155] Jan Morén, ALEŠ UDE, Ansgar Koene, and Gordon Cheng. Biologically based top-down attention modulation for humanoid interactions. *International Journal of Humanoid Robotics*, 5(01):3–24, 2008.

[156] Amir Rasouli, Pablo Lanillos, Gordon Cheng, and John K Tsotsos. Attention-based active visual search for mobile robots. *Autonomous Robots*, 44(2):131–146, 2020.

[157] Ales Ude, Valentin Wyart, Li-Heng Lin, and Gordon Cheng. Distributed visual attention on a humanoid robot. In *5th IEEE-RAS International Conference on Humanoid Robots, 2005.*, pages 381–386. IEEE, 2005.

[158] Danica Kragic, Mårten Björkman, Henrik I Christensen, and Jan-Olof Eklundh. Vision for robotic object manipulation in domestic settings. *Robotics and autonomous Systems*, 52(1):85–100, 2005.

[159] Pablo Lanillos, Emmanuel Dean-Leon, and Gordon Cheng. Yielding self-perception in robots through sensorimotor contingencies. *IEEE Transactions on Cognitive and Developmental Systems*, 9(2):100–112, 2016.

[160] Aaron P Shon, David B Grimes, Chris L Baker, Matthew W Hoffman, Shengli Zhou, and Rajesh PN Rao. Probabilistic gaze imitation and saliency learning in a robotic head. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 2865–2870. IEEE, 2005.

[161] Pablo Lanillos, Joao Filipe Ferreira, and Jorge Dias. Multisensory 3d saliency for artficial attention systems. 2015.

[162] Frederic Kaplan and Verena V Hafner. The challenges of joint attention. *Interaction Studies*, 7(2):135–169, 2006.

[163] Yukie Nagai, Koh Hosoda, Akio Morita, and Minoru Asada. A constructive model for the development of joint attention. *Connection Science*, 15(4):211–229, 2003.

[164] Pablo Lanillos and Gordon Cheng. Active attention applications in robotics. In *International Workshop on Active vision, Attention, and Learning, IEEE Developmental Learning and Epigenetic Robotics (ICDL-Epirob)*. IEEE, Sep 2018.

[165] Fred Bos, Ajith Anil Meera, Dennis Benders, and Martijn Wisse. Free energy principle for state and input estimation of a quadcopter flying in wind. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 5389–5395. IEEE, 2022.

[166] Ajith Anil Meera and Martijn Wisse. A brain inspired learning algorithm for the perception of a quadrotor in wind. *arXiv preprint arXiv:2109.11971*, 2021.

[167] Ajith Anil Meera and Martijn Wisse. On the convergence of dem's linear parameter estimator. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, pages 692–700, Cham, 2021. Springer International Publishing.

[168] Pablo Lanillos and Gordon Cheng. Adaptive robot body learning and estimation through predictive coding. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4083–4090. IEEE, 2018.

[169] Karl Friston, Klaas Stephan, Baojuan Li, and Jean Daunizeau. Generalised filtering. *Mathematical Problems in Engineering*, 2010, 2010.

[170] Karl J Friston, N Trujillo-Barreto, and Jean Daunizeau. DEM: a variational treatment of dynamic systems. *Neuroimage*, 41(3):849–885, 2008.

[171] Greg Welch, Gary Bishop, et al. An introduction to the kalman filter. 1995.

[172] Ajith Anil Meera and Martijn Wisse. Free energy principle for the noise smoothness estimation of linear systems with colored noise. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 1888–1893, 2022.

[173] Steven M LaValle. *Planning algorithms*. Cambridge university press, 2006.

[174] Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. Active inference and epistemic value. *Cognitive neuroscience*, 6(4):187–214, 2015.

[175] Gregory Hitz, Enric Galceran, Marie-Ève Garneau, François Pomerleau, and Roland Siegwart. Adaptive continuous-space informative path planning for online environmental monitoring. *Journal of Field Robotics*, 34(8):1427–1449, 2017.

[176] Roman Marchant and Fabio Ramos. Bayesian optimisation for informative continuous path planning. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6136–6143. IEEE, 2014.

[177] Marija Popović, Teresa Vidal-Calleja, Gregory Hitz, Inkyu Sa, Roland Siegwart, and Juan Nieto. Multiresolution mapping and informative path planning for uav-based terrain monitoring. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1382–1388. IEEE, 2017.

[178] Pablo Lanillos. *Minimum time search of moving targets in uncertain environments*. PhD thesis, PhD thesis, 2013.

[179] Pablo Lanillos, Seng Keat Gan, Eva Besada-Portas, Gonzalo Pajares, and Salah Sukkarieh. Multi-uav target search using decentralized gradient-based negotiation with expected observation. *Information Sciences*, 282:92–110, 2014.

[180] Karl J Friston, Marco Lin, Christopher D Frith, Giovanni Pezzulo, J Allan Hobson, and Sasha Ondobaka. Active inference, curiosity and insight. *Neural computation*, 29(10):2633–2683, 2017.

[181] Jie Chen, Ron J Patton, and Hong-Yue Zhang. Design of unknown input observers and robust fault detection filters. *International Journal of control*, 63(1):85–105, 1996.

[182] Mohamed Darouach, Michel Zasadzinski, A Bassong Onana, and Samuel Nowakowski. Kalman filtering with unknown inputs via optimal state estimation of singular systems. *International journal of systems science*, 26(10):2015–2028, 1995.

[183] Behnam Allahverdi Charandabi and Horacio J Marquez. Observer design for discrete-time linear systems with unknown disturbances. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 2563–2568. IEEE, 2012.

[184] Heiner Kuhlmann. Kalman-filtering with coloured measurement noise for deformation analysis. In *Proceedings, 11th FIG Symposium on Deformation Measurements, Santorini, Greece*, 2003.

[185] Kedong Wang, Yong Li, and Chris Rizos. Practical approaches to kalman filtering with time-correlated measurement errors. *IEEE Transactions on Aerospace and Electronic Systems*, 48(2):1669–1681, 2012.

[186] Jakob Hohwy. *The predictive mind*. Oxford University Press, 2013.

[187] Karl Friston. The free-energy principle: a rough guide to the brain? *Trends in cognitive sciences*, 13(7):293–301, 2009.

[188] Robin L Carhart-Harris and Karl J Friston. The default-mode, ego-functions and free-energy: a neurobiological account of freudian ideas. *Brain*, 133(4):1265–1283, 2010.

[189] Lili Ma, Huiran Wang, and Jinguang Chen. Analysis of kalman filter with correlated noises under different dependence. *JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE*, 7(5):1147–1154, 2010.

[190] A Bryson and D Johansen. Linear filtering for time-varying systems using measurements containing colored noise. *IEEE Transactions on Automatic Control*, 10(1):4–10, 1965.

[191] AE Bryson Jr and LJ Henrikson. Estimation using sampled data containing sequentially correlated noise. *Journal of Spacecraft and Rockets*, 5(6):662–665, 1968.

[192] Dan Simon. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[193] John L Crassidis and John L Junkins. *Optimal estimation of dynamic systems*. CRC press, 2011.

[194] Sze Zheng Yong, Minghui Zhu, and Emilio Frazzoli. Simultaneous input and state estimation for linear time-varying continuous-time stochastic systems. *IEEE Transactions on Automatic Control*, 62(5):2531–2538, 2016.

[195] Agnes Cohen Jalfon and Yoram Halevi. On partially augmented observers for systems with coloured noises. *International Journal of Control*, 57(2):335–349, 1993.

[196] A Bryson and D Johansen. Linear filtering for time-varying systems using measurements containing colored noise. *IEEE Transactions on Automatic Control*, 10(1):4–10, 1965.

[197] Bruce P Gibbs. *Advanced Kalman filtering, least-squares and modeling: a practical handbook*. John Wiley & Sons, 2011.

[198] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[199] Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138, 2010.

[200] Guillermo Oliver, Pablo Lanillos, and Gordon Cheng. Active inference body perception and action for humanoid robots. *arXiv preprint arXiv:1906.03022*, 2019.

[201] AE Bryson Jr and LJ Henrikson. Estimation using sampled data containing sequentially correlated noise. *Journal of Spacecraft and Rockets*, 5(6):662–665, 1968.

[202] Zebo Zhou, Jin Wu, Yong Li, Chen Fu, and Hassen Fourati. Critical issues on kalman filter with colored and correlated system noises. *Asian Journal of Control*, 19(6):1905–1919, 2017.

[203] Steven Waslander and Carlos Wang. Wind disturbance estimation and rejection for quadrotor position control. In *AIAA Infotech@ Aerospace conference and AIAA unmanned... Unlimited conference*, page 1983, 2009.

[204] Dinuka Abeywardena, Zhan Wang, Gamini Dissanayake, Steven L Waslander, and Sarath Kodagoda. Model-aided state estimation for quadrotor micro air vehicles amidst wind disturbances. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4813–4818. IEEE, 2014.

[205] Karl Friston, Jérémie Mattout, Nelson Trujillo-Barreto, John Ashburner, and Will Penny. Variational free energy and the laplace approximation. *Neuroimage*, 34(1):220–234, 2007.

[206] S. Bouabdallah, P. Murrieri, and R. Siegwart. Design and control of an indoor micro quadrotor. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, volume 5, pages 4393–4398 Vol.5, 2004.

[207] Dennis Benders. Ar.drone 2.0 state estimation using dynamic expectation maximization. Master's thesis, Delft University of Technology, 2020.

[208] John L Crassidis and John L Junkins. *Optimal estimation of dynamic systems*. Chapman and Hall/CRC, 2004.

[209] Baojuan Li, Jean Daunizeau, Klaas E Stephan, Will Penny, Dewen Hu, and Karl Friston. Generalised filtering and stochastic dcm for fmri. *neuroimage*, 58(2):442–457, 2011.

[210] Jing-Jing Xiong and En-Hui Zheng. Optimal kalman filter for state estimation of a quadrotor uav. *Optik*, 126(21):2862–2868, 2015.

[211] Mohamed Baioumy, Paul Duckworth, Bruno Lacerda, and Nick Hawes. Active inference for integrated state-estimation, control, and learning. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4665–4671. IEEE, 2021.

[212] Corrado Pezzato, Carlos Hernandez, and Martijn Wisse. Active inference and behavior trees for reactive action planning and execution in robotics. *arXiv preprint arXiv:2011.09756*, 2020.

[213] Mohamed Baioumy, Paul Duckworth, Bruno Lacerda, and Nick Hawes. Active inference for integrated state-estimation, control, and learning. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4665–4671. IEEE, 2021.

[214] Ljung Lennart. System identification: theory for the user. *PTR Prentice Hall, Upper Saddle River, NJ*, 28, 1999.

[215] L-Q Zhang, A Cichocki, and S Amari. Kalman filter and state-space approach to blind deconvolution. In *Neural Networks for Signal Processing X. Proceedings of the 2000 IEEE Signal Processing Society Workshop (Cat. No. 00TH8501)*, volume 1, pages 425–434. IEEE, 2000.

[216] Runzhe Han, Christian Bohn, and Georg Bauer. Blind identification of state-space models in physical coordinates. *arXiv preprint arXiv:2108.08498*, 2021.

[217] Karim Abed-Meraim, Wanzhi Qiu, and Yingbo Hua. Blind system identification. *Proceedings of the IEEE*, 85(8):1310–1322, 1997.

[218] Yong Zhang. Unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods. *Mathematical and Computer Modelling*, 53(9-10):1810–1819, 2011.

[219] Xinggao Liu and Jing Lu. Least squares based iterative identification for a class of multirate systems. *Automatica*, 46(3):549–554, 2010.

[220] Wei Xing Zheng. On a least-squares-based algorithm for identification of stochastic linear systems. *IEEE Transactions on Signal Processing*, 46(6):1631–1638, 1998.

[221] Yong Zhang and Guimei Cui. Bias compensation methods for stochastic systems with colored noise. *Applied Mathematical Modelling*, 35(4):1709–1716, 2011.

[222] Ting Cui, Feiyan Chen, Feng Ding, and Jie Sheng. Combined estimation of the parameters and states for a multivariable state-space system in presence of colored noise. *International Journal of Adaptive Control and Signal Processing*, 34(5):590–613, 2020.

[223] Thijs van de Laar, Ayça Özçelikkale, and Henk Wymeersch. Application of the free energy principle to estimation and control. *arXiv preprint arXiv:1910.09823*, 2019.

[224] Samuel PL Veissière, Axel Constant, Maxwell JD Ramstead, Karl J Friston, and Laurence J Kirmayer. Thinking through other minds: A variational approach to cognition and culture. *Behavioral and brain sciences*, 43, 2020.

[225] Rafael Kaufmann, Pranav Gupta, and Jacob Taylor. An active inference model of collective intelligence. *arXiv preprint arXiv:2104.01066*, 2021.

[226] Elena Miu, Ned Gulley, Kevin N Laland, and Luke Rendell. Innovation and cumulative culture through tweaks and leaps in online programming contests. *Nature communications*, 9(1):1–8, 2018.

[227] Thomas Parr, Dimitrije Markovic, Stefan J Kiebel, and Karl J Friston. Neuronal message passing using mean-field, bethe, and marginal approximations. *Scientific reports*, 9(1):1–18, 2019.

[228] Thijs W van de Laar and Bert de Vries. Simulating active inference processes by message passing. *Frontiers in Robotics and AI*, 6:20, 2019.

[229] Karl Friston, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, Giovanni Pezzulo, et al. Active inference and learning. *Neuroscience & Biobehavioral Reviews*, 68:862–879, 2016.

[230] Karl J Friston. Variational filtering. *NeuroImage*, 41(3):747–766, 2008.

[231] Bhashyam Balaji and Karl Friston. Bayesian state estimation using generalized coordinates. In *Signal Processing, Sensor Fusion, and Target Recognition XX*, volume 8050, page 80501Y. International Society for Optics and Photonics, 2011.

[232] Wolfgang Mader, Yannick Linke, Malenka Mader, Linda Sommerlade, Jens Timmer, and Björn Schelter. A numerically efficient implementation of the expectation maximization algorithm for state space models. *Applied Mathematics and Computation*, 241:222–232, 2014.

[233] Karl Friston. A free energy principle for a particular physics. *arXiv preprint arXiv:1906.10184*, 2019.

[234] F Javier Cara, Jesús Juan, and Enrique Alarcón. Using the em algorithm to estimate the state space model for omax. *In practice*, 1000(1J):3, 2014.

[235] Lennart Ljung. System identification. In *Signal analysis and prediction*, pages 163–173. Springer, 1998.

[236] Soon-Jo Chung, Aditya Avinash Paranjape, Philip Dames, Shaojie Shen, and Vijay Kumar. A survey on aerial swarm robotics. *IEEE Transactions on Robotics*, 34(4):837–855, 2018.

[237] Wolfram Burgard, Mark Moors, Cyrill Stachniss, and Frank E Schneider. Coordinated multi-robot exploration. *IEEE Transactions on robotics*, 21(3):376–386, 2005.

[238] Javier Alonso-Mora, Stuart Baker, and Daniela Rus. Multi-robot formation control and object transport in dynamic environments via constrained optimization. *The International Journal of Robotics Research*, 36(9):1000–1021, 2017.

[239] Nathan Michael, Jonathan Fink, and Vijay Kumar. Cooperative manipulation and transportation with aerial robots. *Autonomous Robots*, 30(1):73–86, 2011.

[240] Alex Kushleyev, Daniel Mellinger, Caitlin Powers, and Vijay Kumar. Towards a swarm of agile micro quadrotors. *Autonomous Robots*, 35(4):287–300, 2013.

[241] Ying Tan and Zhong-yang Zheng. Research advance in swarm robotics. *Defence Technology*, 9(1):18–39, 2013.

[242] Cristian Meo, Giovanni Franzese, Corrado Pezzato, Max Spahn, and Pablo Lanillos. Adaptation through prediction: multisensory active inference torque control. *arXiv preprint arXiv:2112.06752*, 2021.

[243] Raphael Kaplan and Karl J Friston. Planning and navigation as active inference. *Biological cybernetics*, 112(4):323–343, 2018.

[244] Charles W Warren. Global path planning using artificial potential fields. In *1989 IEEE International Conference on Robotics and Automation*, pages 316–317. IEEE Computer Society, 1989.

[245] Harriet Brown, Karl Friston, and Sven Bestmann. Active inference, attention, and motor preparation. *Frontiers in psychology*, 2:218, 2011.

GLOSSARY

AR auto regressive.

DEM Dynamic Expectation Maximization.

EM Expectation maximization.

FEP Free energy principle.

GC Generalized coordinates.

GP Gaussian process.

IPP Informative Path Planning.

KF Kalman filter.

KL Kullback–Leibler divergence.

LTI Linear time invariant.

MF Mean field.

PE Prediction error.

PWM Pulse width modulation.

SA State augmentation.

SC Superior colliculus.

SMIKF Second-moment information based Kalman Filtering.

SS Subspace identification.

SSE Sum of squared error.

UAV Unmanned aerial vehicle.

UIO Unknown Input Observer.

WM Working memory.

Acknowledgments

I don't know when it all started, the moment when I decided to choose science and technology as my passion, to grow up to be a researcher in pursuit of scientific truths. The dream had always been brewing within me for way too long, to as long as I could remember. Was it the science teachers that ignited the curiosity in me early on? Or was it the works and words of great scientists that lived before me that sparked the fire? I know not. Countless people who stood by my side and helped me get through, to the point where I am right now. Among them, I would thank my parents first - Anil and Meera - who trusted in me and my skills, and supported me throughout my life with all the resources and freedom that I need. Despite the absence of any doctors or engineers in the entire family line, despite their humble beginnings, they have placed my education as their top priority, since my childhood. Words could never acknowledge it all.

Secondly, my deepest gratitude to Martijn Wisse who trusted in my skills and gave me the opportunity to do a PhD under his guidance. Your warm and accommodating personality has always put me at ease during our meetings. Your feedback to my work has always been sharp, direct and constructive to a point that I could just follow it with ease right away. Thank you for believing in my skills and giving me ample freedom to shape my research the way I like it - I truly appreciate this. Thanks to my promoter Peyman Mohajerin Esfahani who has an amazing skill to jump straight into the weakest part of a work, who has always asked the right questions, that made me think critically about the mathematical formalism in this thesis.

I have had the luck of interacting with the one and only Karl Friston, the genius mind behind FEP. Despite being the true definition of an academic celebrity, you have been very humble in your interactions with me, educating me by patiently responding to all my queries. I feel lucky to finally write a paper with you. Thanks to my co-authors, Filip Novicky, Pablo Lanillos and Noor Sajid. It was really great collaborating with you. The neuroscience section of this thesis is heavily dependent on your ideas and feedback. I have learned a lot from you all during our weekly meetings, and during the writing phase of the paper. Corrado Pezzato and Mohamed Baioumy, it was a pleasure working with you guys towards the active inference review paper [25].

I have had the opportunity to supervise a few talented masters students. Dennis Benders, the data you collected from the drone experiments have been used in chapters 4, 5 and 8. Thank you for going all the way to New York on my behalf to present the work at ICRA, despite your busy schedule with the PhD work. Fred Bos, your data analysis has contributed to the chapter 4. Thank you for helping me with the last minute graph generation in the middle of your vacation. Luuk van Roessel, I have learned a lot about discrete active inference from you during our 3-4 hour long meetings.

Sherin Grimbergen, you helped me parse the works of Friston during the first few months of my PhD. These interactions have definitely motivated and helped me to get started with my thesis. Charel van Hoof, I still remember the discussions we used to have

about trying to decipher Friston's work. These discussions have driven up my curiosity on FEP! Marija Popovic and Alex Millane, the IPP section of chapter 2 is an extension of my masters thesis at ETH Zurich, guided by you both. Thanks for being the coolest supervisors I could as for! Alvaro Serra Gomez, thanks for helping us setup the drones to fly in the lab.

Linda van der Spaa, Tim de Bruin, Hongpeng Zhou, Carlos Celemin Paez, Giovanni Franzese and Jihong Zhu, thanks for all the countless philosophical, social and scientific discussions we had, sharing an office space together for years. You all were a constant company for me throughout. Hai Zhu, Bruno Brito, Tasos Tsolakis, Alvaro Serra, Padmaja Kulkarni and Maximilian Kronmüller and to all the colleagues with whom I casually meet at the department events and social room gatherings, your presence has been a positive addition to my experience here in the department. Hanneke and Rosanne, thanks for helping me with all the administrative stuff that made my conference visits as smooth as it can get.

Ashwin George and Allen George, thanks for listening to me while I vent out each time about reviewer 2 comments. You guys have been listening to my sorrows for 11 years now, from India to Europe. Aleix Pinardell Pons, you have been a constant company for me throughout my PhD phase, through the ups and the downs, even when I was contemplating over the convergence proof for months. Thanks for always being there for me. Sparsh Sharma, Ajinkya Krishnakumar, Mihir Tandon and Pratik Mishra, thanks for letting me crash into your apartment when I moved back to Delft for my PhD and went homeless right away for months. I really enjoyed our time together. Sparsh, you have been a great company for me here in Delft! Jacquez Aoustin, thanks for being my swimming buddy and the most chaotic friend anyone could ask for! Thanks to Aravind S.P for proofreading the thesis report.

Last but definitely not the least, a special thanks to two teachers who greatly influenced my life. Ramachandran Nair, the school principle who saw me enter his school at age 3 and leave at age 15, who was the greatest source of inspiration for me to work hard, dream big and get whatever I desire. Sudheer A.P, who let me and a bunch of bachelors students into his robotics lab to start a student group called RIG, from which all my passion towards robotics took shape. Thanks for being the constant source of support for all my works.

A special thanks to the US Embassy for scheduling me on a visa interview for February 2023, to attend ICRA in May 2022. A delay of just one year is truly remarkable, and you won't be forgotten. A big thanks to the Mexican immigration for giving me a three day experience on trauma at the Cancun airport, on my way to attend CDC 2022.

Ajith Anil Meera
Delft, September 2022

Curriculum Vitæ
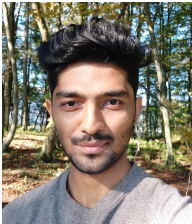
Ajith Anil Meera

04/03/1994   Date of birth in Neyyattinkara, Kerala, India

Education

2011-2015   B.Tech in Mechanical Engineering,
National Institute of Technology Calicut, India,
Thesis: Squid, Salp and Jellyfish inspired Underwater Robot.

2016-2018   M.Sc (cum-laude and honours degree),
Mechanical Engineering, TU Delft, The Netherlands,
Thesis: Informative Path Planning for Search and Rescue.

01/2018- 08/2018   Exchange student for Masters thesis,
Autonomous Systems Lab, ETH Zurich, Switzerland.

2018-2022   PhD in Cognitive Robotics
TU Delft, The Netherlands,
Thesis: Free Energy Principle for Robot Attention.

Awards

2015   Third prize for the development of wall climbing robot [link],
FESTO's International Productivity Contest, Germany.

2018   IDEA League Scholarship.

Experience

2011 - 2015   Co-founder, Robotics Interest Group (RIG-NITC) [link],
National Institute of Technology Calicut, India.

2015 - 2016   Graduate Engineering Trainee,
PepsiCo India Pvt Ltd, Gurgaon, India.

List of Publications

Journals

1. **Ajith Anil Meera**, Filip Novicky, Thomas Parr, Karl Friston, Pablo Lanillos, and Noor Sajid. "Reclaiming saliency: rhythmic precision-modulated action and perception." Frontiers in Neurorobotics 16 (2022).

2. Pablo Lanillos, Cristian Meo, Corrado Pezzato, **Ajith Anil Meera**, Mohamed Baioumy, Wataru Ohata, Alexander Tschantz et al. "Active inference in robotics and artificial agents: Survey and challenges." arXiv preprint arXiv:2112.01871 (2021).

3. **Ajith Anil Meera**, and Martijn Wisse. "Dynamic expectation maximization algorithm for estimation of linear systems with colored noise." Entropy 23, no. 10 (2021): 1306.

4. **Ajith Anil Meera**, and Martijn Wisse. "Dynamic Expectation Maximization for the System Identification of a Quadrotor in Wind." *Under review*.

Conferences

1. **Ajith Anil Meera** and Martijn Wisse, "Free Energy Principle for the Noise Smoothness Estimation of Linear Systems with Colored Noise," 2022 IEEE 61st Conference on Decision and Control (CDC), Cancun, Mexico, 2022, pp. 1888-1893, doi: 10.1109/CDC51059.2022.9992717.

2. Fred Bos, **Ajith Anil Meera**, Dennis Benders, and Martijn Wisse. "Free energy principle for state and input estimation of a quadcopter flying in wind." In 2022 International Conference on Robotics and Automation (ICRA), pp. 5389-5395. IEEE, 2022.

3. **Ajith Anil Meera**, and Martijn Wisse. "Free energy principle based state and input observer design for linear systems with colored noise." In 2020 American Control Conference (ACC), pp. 5052-5058. IEEE, 2020.

4. **Ajith Anil Meera**, Marija Popović, Alexander Millane, and Roland Siegwart. "Obstacle-aware adaptive informative path planning for uav-based target search." In 2019 International Conference on Robotics and Automation (ICRA), pp. 718-724. IEEE, 2019.

Workshops

1. **Ajith Anil Meera**, and Martijn Wisse. "On the convergence of DEM's linear parameter estimator." In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 692-700. Springer, Cham, 2021.

2. **Ajith Anil Meera**, Filip Novicky, Thomas Parr, Karl Friston, Pablo Lanillos, and Noor Sajid. "Reclaiming saliency: rhythmic precision-modulated action and perception." In Third International Workshop on Active Inference (2022).

PATENTS

1. A.P Sudheer, **Ajith Anil Meera**, Ashwin George, Basil George, Hari Prasanth and Jofred Joseph. "Radially Contracting Toroidal Pneumatic Artificial Muscle", Application number: 202041047095, Year: 2020. [link]

PREVIOUS PUBLICATIONS

1. **Ajith Anil Meera**, and Attadappa Puthanveetil Sudheer. "Design optimization of the biomimetic undulating fin of a knife fish robot." Journal of Automation, Mobile Robotics and Intelligent Systems (2016): 26-33.

2. **Ajith Anil Meera**, and Attadappa Puthanveetil Sudheer. "Mathematical Modelling and Feasibility Study of the Biomimetic Undulating Fin of a Knife Fish Robot." Journal of Automation Mobile Robotics and Intelligent Systems 10 (2016).

3. **Ajith, A.M**, Moitra, S. and Sudheer, A.P., 2015, July. Parameter optimization and experimentation of the undulating fin of a knife fish robot. In Proceedings of the 2015 Conference on Advances In Robotics (pp. 1-6).

4. **Ajith, A. M**, K. S. Sachin, and A. P. Sudheer. "Design, fabrication and analysis of a bio-inspired tuna Fish Robot." In Proceedings of the 2015 Conference on Advances In Robotics, pp. 1-6. 2015.