AMSI SUMMERRESEARCH SCHOLARSHIPS 2023–24

SET YOUR SIGHTS ON RESEARCH THIS SUMMER

Investigation of switching thresholds in the Switching Observer model for human perceptual estimation

Chih Yuan CHAN

Supervised by Dr Robert Cope

University of New England

Contents

Abstract

The human brain constantly receives and interprets sensory information from the environment to arrive at an estimate, or *percept*, of the true state of the world. This process is inherently uncertain due to the imperfect acuity of our senses and the presence of sensory noise or ambiguity. Bayesian inference offers a powerful framework to interpret how the brain processes this uncertainty by representing perception as an unconscious process of statistical inference. In this study, 3 separate models were fit to an experimental dataset for a dot motion direction estimation task to determine which model best fits the data. The first 'Basic Bayesian Observer' (BBO) model adopts a Bayesian inference approach. The second 'Switching Observer' (SO) model implements a heuristic approach to switch between sensory likelihood and prior, and the last 'Triple Switching Observer' (TSO) model implements a hybrid approach combining both the BBO and SO approaches.

The study found that whilst all three models fitted the data reasonably well, only the latter 2 models could account for the apparent inherent bias towards the prior mean in the data as reflected by bimodal estimate distributions. The study validated previous findings by others that for this visual estimation task at least, the human brain appears to implement heuristics to achieve near, but not perfect, optimality in perceptual estimates. Furthermore, the study demonstrates that the TSO model obtains the best results for data fit by a small margin. This suggests the possibility that the brain can implement both mathematically optimal and heuristic approaches, selecting one or the other based on some hidden utility objective.

Introduction

The human brain is constantly processing uncertainty during its perception of the environment through the senses of sight, hearing, smell and touch. Such uncertainty exists in proportion to the strength of the presented sensory stimuli and its accompanying complexity (Ma, Kording, & Goldreich 2023, pp.13-18). For example, visual uncertainty could present from, say, low levels of lighting or from the ambiguity of similar competing visual cues, such as an animal hiding against a camouflaging background. Furthermore, the sensory information being presented to the brain is itself constantly changing, and newly received sensory information requires the brain to continually update its perceptual estimates (Ma, Kording, & Goldreich 2023, p.115).

Bayesian inference provides a powerful and elegant mathematical framework to explore how the brain processes sensory uncertainty during perception. In this framework, the brain is thought to perform unconscious statistical inference to generate a best estimate of the environment using available sensory measurements and a prior internal model of the world (Girshick, Landy, & Simoncelli 2011). Uncertainty in sensory information is quantified in the brain as a probabilistic distribution from which an optimal estimate can be read out (Ma, Kording, & Goldreich 2023, p.12). Bayesian inference is also consistent with the idea of updating beliefs in response to new evidence; hence this framework can be conveniently adapted to explain how the brain updates perception as new sensory information is received (Kurt 2019, p.65). Since Bayesian inference leads to mathematically optimal estimates in the presence of prior knowledge, it may be surmised that these estimates represent what the brain *should* decide upon to maximise accuracy.

In this study, we conduct an independent study of an experimental dataset by Laquitaine and Gardner (2017) for a motion direction estimation task to determine if the data supports the notion that the human brain conducts visual perception according to Bayesian inference. Laquitaine and Gardner (2017) proposed 2 main models to fit the data from a dot motion direction estimation task: a 'Basic Bayesian Observer' (BBO) model implementing Bayesian inference and a 'Switching Observer' (SO) model implementing a heuristic approach to switch between the sensory likelihood and prior. Our study builds on this original research by conducting an independent review of the above models using a subset of the data, followed by an extension investigation discussed below.

Firstly, we independently fit the BBO and SO models to a subset of the original dataset (5 subjects) using a modified approach. Modelling outcomes are used to review the conclusions reached by Laquitaine and Gardner (2017). Secondly, we examine the possibility that the brain may deploy *both* mathematically optimal and heuristic approaches, switching between either to maximise an implicit utility objective, such as an optimal speed accuracy trade-off. We propose and implement a hybrid 'Triple Switching Observer' (TSO) model that combines the BBO with the SO using a mixing coefficient to fit the data. The mixing coefficient acts as a switching mechanism and represents a threshold at which either model is selected over the other - this is examined to identify any trends or patterns that might reveal clues to a hidden utility objective.

Statement of Authorship

The concept of this study is based on, and builds upon, original research by Laquitaine and Gardner (2017).

- Dr Robert Cope provided supervision in the form of guidance and feedback throughout the project, and also provided an example of the model fitting framework for the Basic Bayesian Observer model.
- Chih Yuan Chan was responsible for the detailed implementation of all models using Python code, data analysis, compilation and interpretation of modelling results and the production of this report.

1 Background

1.1 Experimental Dataset

The experimental dataset captures data from a study involving 12 subjects performing a dot motion direction estimation task (Laquitaine & Gardner 2017). Each instance of the motion direction estimation task is termed a *trial*, and an experimental *block* consists of approximately 200 trials. Each subject performed at least 5 *sessions*, each consisting of minimum 5 pseudo-randomly selected blocks. The dataset contains a total of 83213 rows of user entered direction estimates.

1.1.1 Experimental Sequence (Laquitaine & Gardner 2017)

The subject first fixates on a central dot before a group of moving dots appears on the screen with a pre-determined *motion coherence* - this represents the proportion of dots moving in the same direction with the remainder in random Brownian motion. The subject's task (Fig.1) is to turn a electronic paddle dial and confirm the estimated direction on a keyboard. Following confirmation, a feedback display of the true motion direction is displayed.

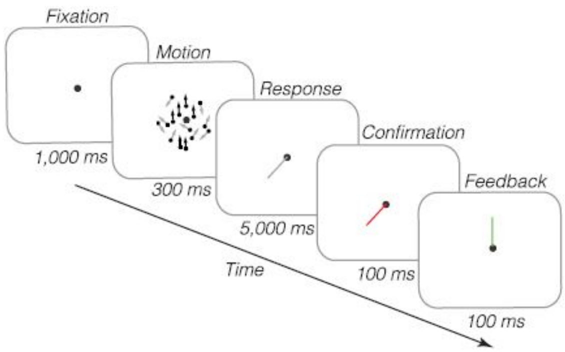

Figure 1: Sequence of motion direction estimating task (Source: Laquitaine & Gardner 2017)

1.1.2 Experimental Prior

The prior represents the motion distribution that the subject has associated with the dot motion, even before viewing the trial. This conditioning may occur by exposing the subject to many prior motion videos before the trials.

From Laquitaine and Gardner (2017), the prior is represented by a discretised Von Mises distribution with a fixed mean of 225 degrees, taking values from 5 to 355° in steps of 10°(36 directions). Priors for each trial block are randomly assigned from one of four distinct priors, each with a concentration width (standard deviation) of 10, 20, 40 and 80 degrees respectively (refer Figure 2). Each trial block therefore has the same mean but a different prior concentration. To manipulate the priors in each experimental block, true dot motion directions were randomly drawn from the assigned prior distributions. It is assumed that priors were learnt by the subjects before the commencement of each trial block.

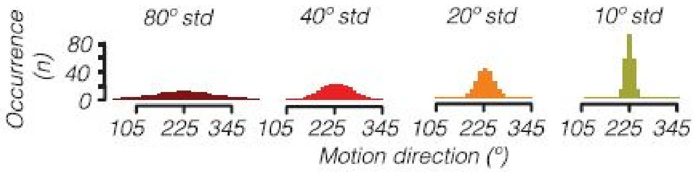

Figure 2: Experimental Priors (Source: Laquitaine & Gardner 2017)

1.1.3 Sensory Evidence

The sensory evidence is the visual information that is being received by the brain. The strength of the sensory evidence is reflected in visual acuity. To manipulate the strength of sensory evidence on a trial-by trial basis, the motion stimulus on each trial was displayed with a pseudo-randomly selected motion coherence of 6, 12 or 24% (Laquitaine & Gardner 2017).

2 Project methodology

This approach of this study was divided into the following steps:

1. **Exploratory Data Analysis**: Clean dataset and analyse data for trends using visual plots and statistical summaries
2. **Model Development - a) BBO / SO**: Independently develop and implement Gardner and Laquitaine (2017) Basic Bayesian and Switching Observer models in Python code, with modifications to the original research approach to account for computational constraints:
   - This study will explore a subset of the data (5 subjects: 1,2,7,10 and 12 only). Subjects were chosen to represent a range of log-likelihood results for the Laquitaine and Gardner (2017) SO modelling.
   - The relatively small effect of random motor lapses is ignored in our approach
   - Model parameter fitting is performed subject by subject per pairing of motion coherence and prior

   Model Development - b) TSO: Develop and implement new proposed hybrid Triple Switching Observer) model (in Python code)
3. **Sensitivity analyses**: Review effect on the model of varying different parameters to validate model fitting and compare to original published results (omitted here for brevity).
4. **Maximum Likelihood Estimation (MLE)**: Perform MLE inference on each model's parameters using standard numerical optimisation techniques (BFGS / CMA-ES)

5. **Assess goodness of fit**: Goodness of fit for each model will be assessed and compared using the Akaike Information Criterion (AIC)

6. **Switching thresholds**: Examine switching thresholds for the SO and TSO models to identify trends or patterns that could uncover an underlying utility objective in the brain

2.1 Mathematical Models for Visual Perception

The aim of this study is to form a mathematical model of the internal brain processes that can predict the subject's estimate response for the dot motion direction estimation task. 3 models are proposed (BBO/SO/TSO) with each model represented by a probability distribution over the subject's motion direction estimate response given the true motion direction, $P(\theta_r|\theta_{true})$.

The internal brain processes during perception are hidden and unknown. Hence, the parameters defining each model of the brain will need to be inferred from the following available information:

- θ_{true}: The true motion direction as recorded in the data.
- θ_r: The subject's motion direction estimate response as recorded in the data.

The following representations are adopted across all fitted models:

1. **Learnt prior** θ'_{true}: This is the prior learnt by the observer's brain. Like the true prior distribution (refer 2.1.3), it is formalised as a Von Mises distribution that is centered on the prior mean $\theta_{prior\ mean}$ (fixed at 225°), but has a fitted concentration width K_p :

$$\theta'_{true} \sim \mathcal{V}(\theta_{prior\ mean=225}, K_p) \tag{1}$$

2. **Sensory evidence distribution** θ_e - The sensory evidence distribution in the observer's brain; also a Von Mises distribution that is centred on θ_{true}, with a concentration width K_e:

$$\theta_e \sim \mathcal{V}(\theta_{true}, K_e) \tag{2}$$

K_p and K_e are the unknown model parameters representing the internal brain processes and these will need to be fitted to the alternative models (BBO/SO/TSO) per experimental subject, as follows:

- Four K_p parameters, one for each standard prior strength (10°, 20°, 40°, 80°)
- Three K_e parameters, one for each motion coherence (6%, 12%, 24%)

2.1.1 Basic Bayesian Observer (BBO) (Laquitaine & Gardner 2017)

Using Bayes theory, the BBO model multiplicatively combines the learnt prior (equation 1) and sensory likelihood to produce a posterior distribution of the subject's percept.

For each trial i, a particular true motion direction θ_{true} is displayed. Assume for the moment that the subject perceives the motion direction to be θ_{e_i}. Using the sensory evidence distribution θ_e (refer equation 2) in their brain, the observer forms a sensory likelihood $P(\theta_{e_i}|\theta_{true})$. Thus, given the learnt prior and the sensory likelihood, for any trial i the observer's perceived true motion direction is given by the posterior:

$$P(\theta'_{true}|\theta_{e_i}) = \frac{P(\theta_{e_i}|\theta_{true})P(\theta'_{true})}{P(\theta_{e_i})} \tag{3}$$

In the experiment, the observer outputs a single estimated motion direction or percept θ_p as their response, and this is taken to be the mode of the posterior or maximum a posteriori (MAP):

$$\theta_p = \underset{\theta'_{true}}{\text{argmax}}(P(\theta'_{true}|\theta_{e_i})) \tag{4}$$

Motor noise is present when the subject physically inputs the percept on the keyboard and this can be represented by a Von Mises distribution with fitted concentration width K_m, where K_m is fitted to each subject. The percept is then convolved with the motor noise to obtain the subject's response estimate:

$$P(theta_r|\theta_{true}) = \mathcal{V}(\theta_p, K_m) \tag{5}$$

For each trial i, θ_{e_i} in the subject's brain is unknown. Therefore, to obtain $P(\hat{\theta_r}|\theta_{true})$ we need marginalise out θ_{e_i} across all possible discrete values in the range of 0°to 360°:

$$P(\theta_r|\theta_{true}) = \sum_{\theta_{e_i}} P(\theta_r|\theta_{e_i}, \theta_{true})P(\theta_{e_i}|\theta_{true}) \tag{6}$$

Finally, since we have the actual recorded subject response and true motion direction from the dataset, θ_r and θ_{true} respectively, we can estimate the concentration parameters K_p, K_e and K_m (8 parameters in total per subject) using a maximum likelihood inference (refer section 3.3).

2.1.2 Switching Observer (SO) (Laquitaine & Gardner 2017)

The approach for developing the SO model is similar to the BBO, except that the posterior calculation is omitted. Instead, a switching mechanism is implemented to select the prior mean or sensory evidence based on their relative strengths, as follows:

- Probability that percept is drawn from the prior mean: $P_{prior} = \dfrac{K_p}{K_p + K_e}$
- Probability that percept is drawn from the sensory evidence: $P_e = 1 - K_p$

The observer's percept is then obtained as:

$$P(\theta_p|\theta_{true}) = p_e P(\theta_e|\theta_{true}) + p_{prior}\delta(\theta_p - \theta_{prior\ mean}) \tag{7}$$

where $\delta()$ is a delta function over the percepts θ_p that peaks at the prior mean $\theta_{prior\ mean}$.

Per the BBO, we convolve the percept with motor noise to obtain the subject's target estimate and marginalise out the unknown θ_{e_i} (refer equations 5 and 6), followed by fitting the resulting 8 concentration parameters (K_p, K_e, K_m), per subject, using maximum likelihood inference.

2.1.3 Triple Switching Observer (TSO)

This study introduces the TSO model, which is a hybrid combination of the BBO and SO models. The motivation of the TSO is the hypothesis that the brain switches between the BBO and the SO depending on some hidden utility objective, such as a speed accuracy tradeoff.

The TSO combines the BBO and SO using a switching mechanism similar to that for the SO, with P_{switch}: the probability of selecting the BBO over the SO. The oberver's percept is then obtained as:

$$P(\theta_p|\theta_{true}) = P_{switch}P_{BBO}(\theta_r|\theta_{true}) + (1 - P_{switch})P_{SO}(\theta_p|\theta_{true}) \tag{8}$$

which is then convolved with motor noise per BBO and SO approaches and marginalised for the unknown θ_{e_i} (refer equations 5 and 6).

For the TSO, the introduction of the additional P_{switch} parameter on top of the 8 concentration parameters K_p, K_e, K_m results in 9 parameters to be fitted in total, per subject.

2.2 Maximum Likelihood Estimation

With the subjects' predicted motion direction estimate response for each model, $P(\theta_r|\theta_{true})$, we have a likelihood function that we can maximise to estimate the unknown model parameters (K_e, K_p, K_m), for each subject. In this study, we will utilise the Scipy library "optimize.minimize" function (SciPy. (n.d.-a)) to perform maximum likelihood estimation, i.e. minimising the negative log likelihood, because this function is too complex to maximise analytically.

2.2.1 Optimisation algorithms

We employ 2 separate algorithms (BFGS and CMA-ES) for cross-verification purposes. BFGS uses first derivatives only and has proven good performance even for non-smooth optimizations. (SciPy n.d.). CMA-ES is an evolutionary algorithm designed for non-linear non-convex optimisation problems (Strock n.d.). The probability of finding a near optimum in an early stage of the optimization process with CMA-ES is very high (Galvan et al. 2003). The TSO hybrid formulation results in a highly complex likelihood space and requires the use of the CMA-ES from the outset.

2.3 Goodness of fit - Akaike Information Criterion (AIC)

To compare goodness of fit across all models, we use the Akaike Information Criterion (AIC). The AIC addresses overfitting by adding a penalty term for every free model parameter to the maximum log likelihood of the model, according to the formula (Ma, Kording, & Goldreich 2023, p.351):

$$AIC = -2(log\mathscr{L}^* - n_{parameters}) \tag{9}$$

3 Results

3.1 Exploratory Data Analysis

We produce subplots of Estimated Motion Direction Vs True Motion Direction for each of the 5 subjects (1, 2, 7, 10, 12). Each of the 12 subplots accounts for a combination of the 4 standard priors (grouped column-wise) and 3 motion coherences (grouped row-wise).

For each subplot, the veridical line represents 1 to 1 perfect correlation between the estimated direction with the true motion direction. Note that the directions are measured from the prior mean (marked 0 at the middle of each axis). Refer Figure 3 below for the reproduced Subject 1 subplots. Subject 10 subplots are attached in Figure A.2.

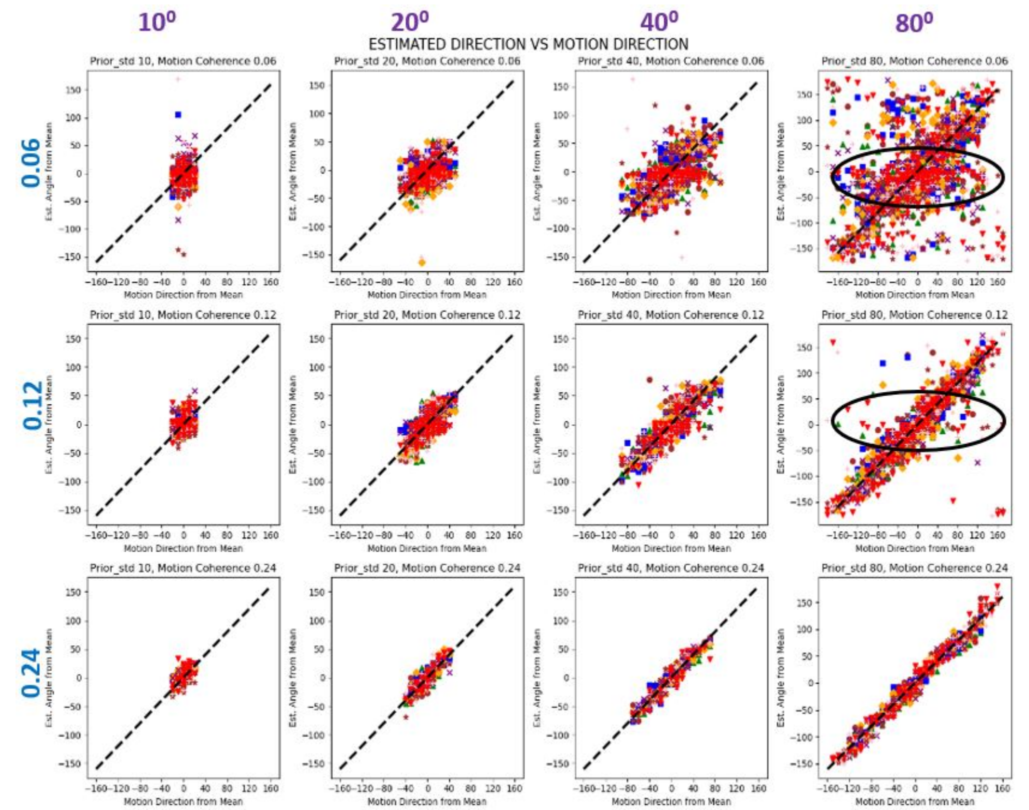

Figure 3: Estimated Motion Vs Motion Direction - Subject 1

Observer's motion direction estimates from a single trial, coloured by session number

We note the following salient aspects of the subplots that are generally characteristic of all 5 subjects:

- At the highest coherence of 24% the subject's estimates closely follow the veridical line - this makes sense since the sensory evidence is at its maximum.

- As the coherences reduce in strength, the variance in the subject estimates increases, as reflected in the higher vertical dispersion of points away from the veridical line.
- With weaker coherences combined with less-informative priors, a bunching of subject estimates along the prior mean can be observed (refer circled zones), reflecting a bias - this effect is most pronounced for the lowest motion coherence (6%) combined with the flattest prior (80° standard deviation). This suggests a greater defaulting of subject estimates to the prior mean as the strength of sensory evidence is reduced.

3.2 Fitted Models - Example plots

We plot each model superimposed against the subjects' estimate histograms, with motion direction estimates (Y axis) and true motion directions (X axis) plotted as differences from the prior mean. Refer to Figure A.1 for estimated model parameters. All plots generally reflect the data reasonably well. Refer to Appendix A.1 for example plots for subjects 1 and 10 (6% motion coherence and standard prior of 80°).

3.2.1 Basic Bayesian Observer Vs Switching Observer

Refer to Figure 4 and Figures A.3, A.4 (Subject 1) & A.7, A.8 (Subject 10) in Appendix A.3 for the discussion below. All BBO and SO model plots are superimposed over the subject estimate histograms for comparative purposes.

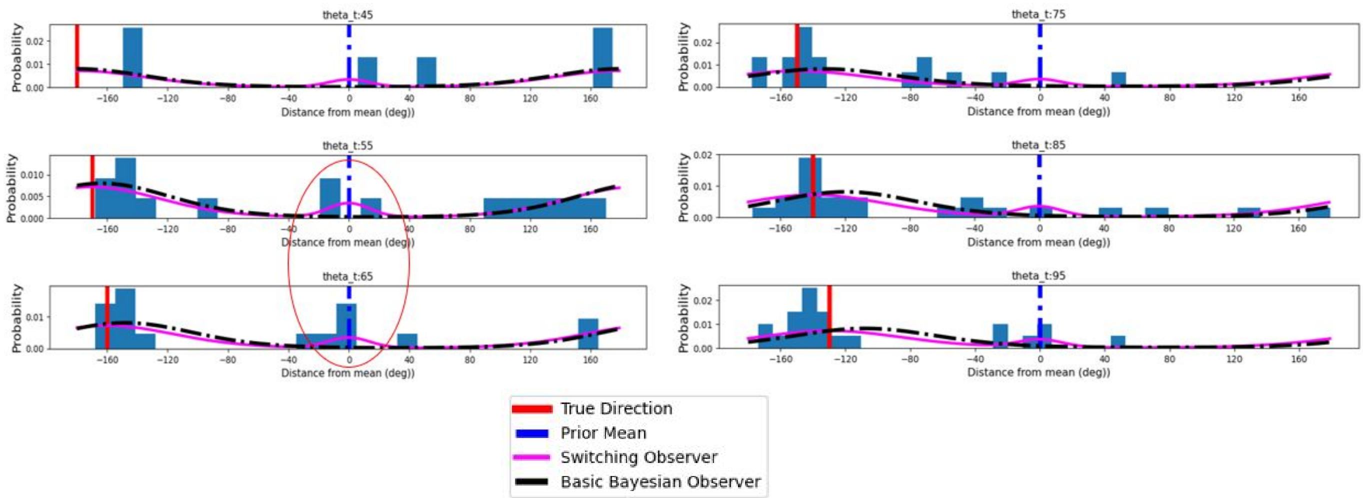

Figure 4: BBO Vs SO extracted subplots (45°to 90°)- Subject 1; Prior: 80 deg; Coherence: 6%
BBO and SO curves superimposed over histogram of observer's direction estimates

- **Bimodality**: Bimodality occurs in the observers' estimate distributions as the sensory evidence reduces in strength. It presents as 2 peaks in the data: one at the true motion direction and another at the prior mean (refer Figure 4 circled). This effect is more pronounced as the true motion direction increases from the prior mean. From the model plots, we see that the BBO model plots is unimodal and cannot account for the bimodal distribution that occurs in the Subject 1 estimates, such as at $\theta_{true} = 65°$ and 355° (refer Figure 4). However, the SO model is formulated to account for bimodality (refer section 2.1.2) - the delta 'spike' function in the model formulation presents as a distinct 'hump' centred on the prior mean. See below for Subject 10 discussion.

- **Dispersion**: The higher the dispersion or randomness in the observer estimates, the less distinct the clustering of estimates around the prior mean. This can be seen in the EDA subplots for Subject 10 (refer Figure A.2 for prior of 80°and coherence of 6%). The BBO model is able to capture this wide spread by having low concentrations in its modelling parameters. As bimodality is less pronounced in Subject 10 estimate distributions, the BBO and SO models are less differentiated.

3.2.2 Switching Observer Vs Triple Switching Observer

Refer to Figure 5 and Figures A.3, A.4 (Subject 1) & A.7, A.8 (Subject 10) in Appendix A.1 for the discussion below. All SO and TSO model plots are superimposed over the subject estimate histograms for comparative purposes.

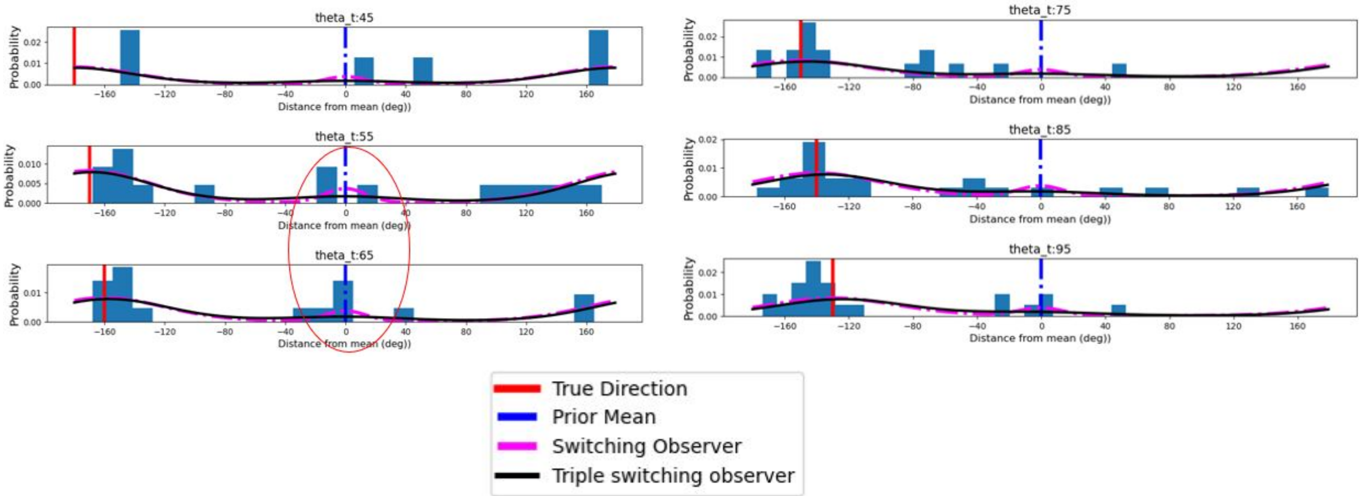

Figure 5: SO Vs TSO extracted subplots (45°to 90°)- Subject 1; Prior: 80 deg; Coherence: 6%
SO and TSO curves superimposed over histogram of observer's direction estimates

- **TSO P_{switch} coefficient**: The P_{switch} factor for the Subject 1 TSO model is 0.234, indicating a 23.4% probability of switching to the BBO model and a 76.6% probability to switch to the SO model. Since the SO is the dominant component in the TSO model, we expect the SO and TSO model plots to be very similar for Subject 1, which is the case. For Subject 10, P_{switch} is 0.97, indicating an almost 100% contribution from the BBO. This reflects that the BBO model is able to fit the data closely when there is greater dispersion and less distinct bimodality in the observer estimates (refer Figure A.2 for prior of 80°and coherence of 6%).

- **Bimodality**: P_{switch} for the TSO is less than 1 and this has the effect of reducing the 'hump' at the prior mean compared to the SO plot. However, the TSO plot line is still slightly lifted up locally in this region to capture some effect of this clustering (per circled zone in Figure 5). For Subject 10, the bimodality effects are less pronounced and the contribution of the SO model component to the TSO is insignificant.
- **Dispersion**: Due to the dispersion and reduced bimodality in observer estimates, the BBO, SO models and TSO models are almost identical.

3.3 Model Comparison - Goodness of fit

We evaluate the goodness of fit for each of the 3 models by examining the AIC scores for all 5 subjects(refer Figure 6 below). The lower the AIC score, the closer the proposed model reflects the experimental data (refer section 2.3). To obtain subject total AIC scores, the scores for each combination of prior and motion coherence pertaining to the subject are summed. As the score is aggregated per subject, the overall performance of a model may not reflect its performance on a specific prior and motion coherence pairing. Per previous discussion, the BBO model perform betters than the SO model for Subject 10 at a prior of 80°and motion coherence of 6%, but overall the SO outperforms the BBO for this subject. The following AIC results were obtained:

- TSO model has the lowest AIC indices out of all 3 models and fits all 5 subjects better than all competing models, followed by the SO model.
- The evidence for the TSO over SO (SO-TSO) is less marked than for the SO over BBO (BBO-SO).

	AIC				
Subject	**BBO**	**SO**	**TSO**	***BBO-SO***	***SO-TSO***
1	8069.45	7591.40	7202.20	478.05	389.20
2	9667.11	8423.83	8192.80	1243.29	231.03
7	8057.95	7405.35	7262.00	652.60	143.35
10	13854.59	13035.85	12921.80	818.75	114.05
12	13556.87	12856.05	12777.60	700.82	78.45

Figure 6: AIC evidence for the SO vs BBO & TSO vs SO

3.4 Discussion - Switching thresholds

The 'switching' mechanisms for the SO model and the TSO model are not equivalent. For the SO model, the switching coefficients P_{prior} and P_e select between the sensory evidence or the prior mean (refer section 2.1.2). For the TSO model, the switching coefficient P_{switch} selects between the component BBO and SO models (refer section 2.1.3).

- **SO model**:
  Refer to Figure 7 below for subplots of switching thresholds (P_{prior}) for all 5 subjects.
  Switching probabilities and patterns are highly subject specific.As prior strength decreases, subjects'

reliance on sensory evidence increases and the probability to switch to the prior reduces, and vice versa. A distinct threshold (minimum 40°) at which subjects rely exclusively on sensory evidence appears to exist for the majority of subjects, at a sufficiently high motion coherence (24%).

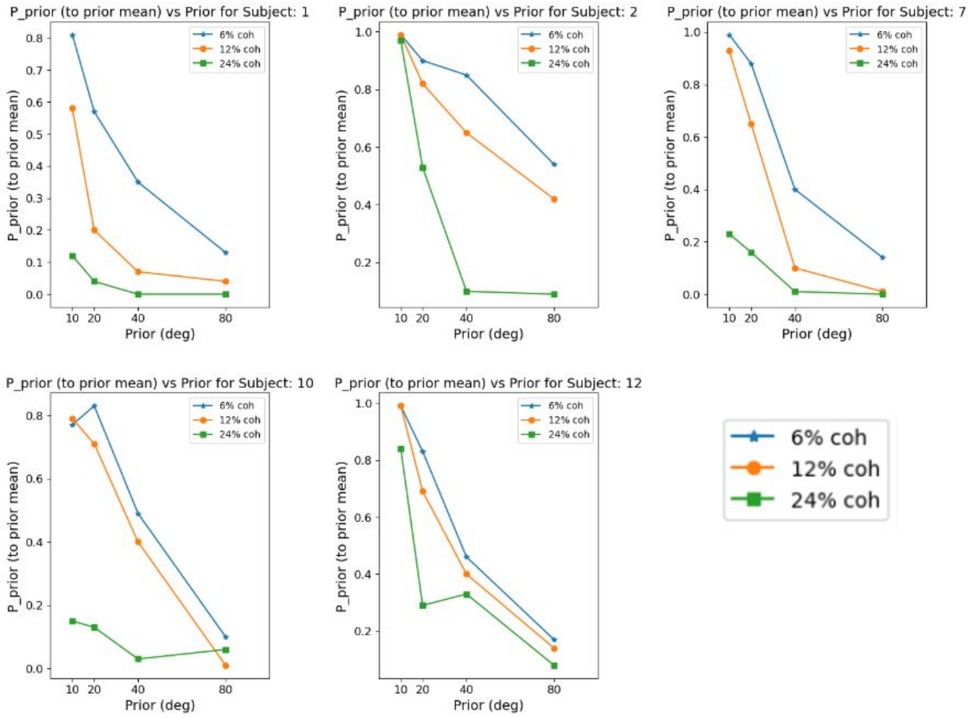

Figure 7: SO model switching thresholds - P_{prior} Vs Standard Prior
(Adapted from Laquitaine and Gardner 2017, supp. Figure S3a)

- **TSO model**:
  Refer to Figure 8 below for subplots of switching thresholds (P_{switch}) for all 5 subjects.
  All subjects appear to be relying on a combination of performing mathematical inference (represented by the BBO component) and heuristic switching (represented by the SO component). Subjects' switch probabilities are highly variable and unstable - there does not appear to be any generalising pattern or distinct thresholds for switching across all subjects. More work is required to uncover the correlation, if any, of switching probabilities with subject experimental features, such as reaction time. Slower reaction times could reflect more complex brain computations, such as performing inference to obtain a mathematically optimal estimate. Whereas, quicker response times could reflect defaulting to the prior or sensory likelihood.

4 Conclusions & future recommendations

In this study, 3 alternative mathematical models were fitted to an experimental dataset to determine if the human brain combines prior knowledge and sensory evidence into mathematically optimal perceptual estimates,

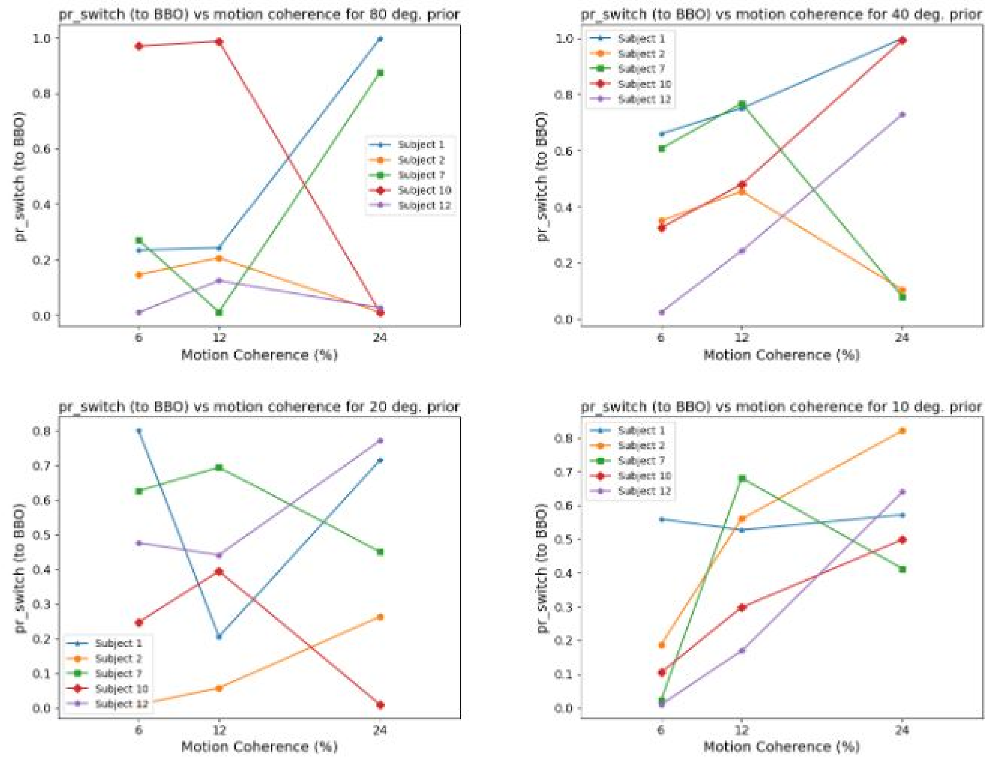

Figure 8: TSO model switching thresholds - P_{switch} Vs Motion Coherence

as prescribed by a Basic Bayesian Observer (BBO) model. The results of the study generally support the findings of Laquitaine and Gardner (2017) : that heuristic approaches, as represented by a Switching Observer (SO) model, fit the data better. However, our implementation of a new Triple Switching Observer (TSO) model provides some supporting evidence that the human brain may flexibly switch between both approaches without exclusively relying on one or the other. Generalising switching patterns for the different subjects that could uncover a hidden utility objective for the brain were not identified in this preliminary study. Notwithstanding, the correlation of switching probabilities with other experimental features, such as subject reaction times, has yet to be undertaken, and is recommended for future work.

5 Acknowledgement

The author would like to sincerely thank Dr Robert Cope for his patient and dedicated instruction, and the generous giving of his time to supervise this project.

6 References

Galvan, B., Griener, D., Periaux, J., Sefrioui, M. and Winter, G. (2003). Parallel Evolutionary Computation for Solving Complex CFD Optimization Problems : A Review and Some Nozzle Applications. Parallel Computational Fluid Dynamics, [online] pp.573–604.

Girshick, A.R., Landy, M.S. and Simoncelli, E.P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. Nature Neuroscience, 14(7), pp.926–932.

Kurt, W. (2019). Bayesian Statistics the Fun Way. San Francisco, California: No Starch Press.

Laquitaine, S. and Gardner, J.L. (2018). A Switching Observer for Human Perceptual Estimation. Neuron, 97(2), pp.462-474.e6.

Ma, W.J., Kording, K.P. and Goldreich, D. (2023). Bayesian Models of Perception and Action. 1st ed. Cambridge, Massachusetts: MIT Press.

SciPy (n.d.). scipy.optimize.minimize — SciPy v1.12.0 Manual. [online] Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.htmlid5 [Accessed 28 Feb. 2024].

Strock, R. (n.d.). cma-es: Covariance Matrix Adaptation Evolution Strategy (CMA-ES) implemented with TensorFlow. [online] PyPI. Available at: https://pypi.org/project/cma-es/ [Accessed 28 Feb. 2024].

A Appendices

A.1 Estimated model parameters

Model parameters were estimated per combination of prior and motion coherence for each subject. To obtain model parameters per subject, parameters were summed and averaged across combinations of prior and motion coherence.

	Basic Bayesian Observer								
Subjects / Parameters	**Ke_24**	**Ke_12**	**Ke_6**	**Kp_80**	**Kp_40**	**Kp_20**	**Kp_10**	**Km**	
1	36.7	29.1	17.0	8.6	2.4	8.5	30.4	33.7	
2	22.0	12.3	7.6	15.2	35.2	27.4	37.7	15.5	
7	20.7	39.2	0.7	0.5	4.2	20.4	22.4	17.2	
10	10.1	13.6	4.6	0.1	38.1	12.2	5.0	31.5	
12	26.9	21.8	7.4	2.0	30.3	21.4	47.4	27.9	
	Switching Observer								
Subjects / Parameters	**Ke_24**	**Ke_12**	**Ke_6**	**Kp_80**	**Kp_40**	**Kp_20**	**Kp_10**	**Km**	
1	352.8	14.1	4.9	0.2	1.1	9.4	31.4	37.7	
2	34.7	8.8	6.3	0.7	8.6	50.4	1294.3	29.9	
7	200.7	4.0	1.2	0.1	1.1	19.2	29.5	28.9	
10	281.7	1.7	1.1	5.6	4.0	17.4	19.1	15.8	
12	16.2	2.2	1.0	0.3	1.4	3.4	240.3	17.9	
	Triple Switching Observer								
Subjects / Parameters	**Ke_24**	**Ke_12**	**Ke_6**	**Kp_80**	**Kp_40**	**Kp_20**	**Kp_10**	**Km**	**P_switch**
1	185.6	57.2	28.9	16.3	10.6	3.2	39.5	66.8	0.61
2	22.0	10.7	3.8	4.0	7.7	76.5	38.5	32.3	0.26
7	192.2	32.2	0.9	2.0	2.2	18.2	11.4	42.0	0.46
10	178.0	1.5	1.0	6.6	5.4	8.9	4.3	39.8	0.44
12	27.2	2.2	0.9	3.2	1.2	2.1	29.0	28.3	0.31

Figure A.1: Estimated model parameters

A.2 Exploratory Data Analysis for Subject 10

NOTE: Subplots ordered by motion coherence (rows) & and priors (columns)

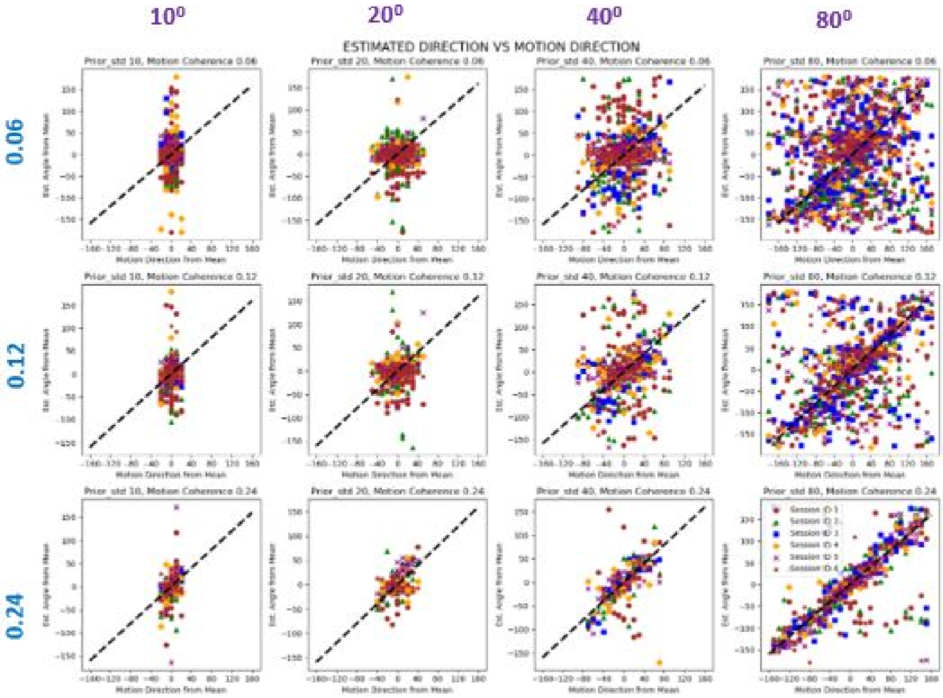

Figure A.2: Estimated Motion Direction Vs True Motion Direction - Subject 10
Observer's motion direction estimates from a single trial, grouped by session number

A.3 Example Plots - Subjects 1 & 10; Prior: 80°; Coherence: 6%

Model plots for BBO Vs SO and SO Vs TSO superimposed over subject estimate histograms.
Vertical red tick indicates true motion direction and blue tick indicates the prior mean.
All distances are measured from the prior mean = 225°.

- **Figure A.3 & A.4**: BBO Vs SO (Subject 1)
- **Figure A.5 & A.6**: SO VS TSO (Subject 1)
- **Figure A.7 & A.8**: BBO Vs SO (Subject 10)
- **Figure A.9 & A.10**: SO VS TSO (Subject 10)

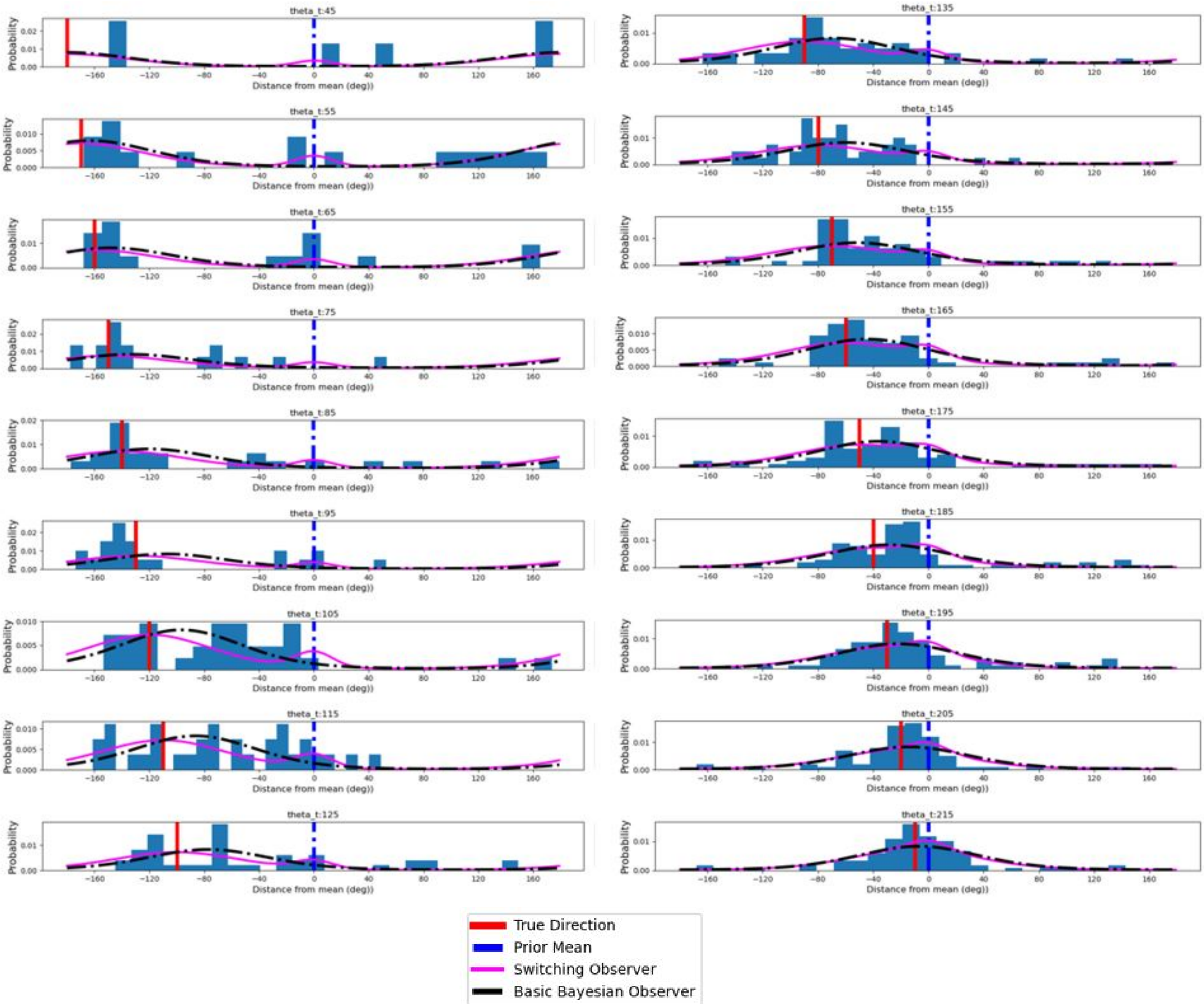

Figure A.3: BBO Vs SO (Subject: 1; Prior: 80°; Coherence: 6%)

True Direction: 45°to 215°(distance from prior mean)

BBO: $K_e = 52.7; K_p = 21.2; K_m = 1.77$

SO: $K_e = 1.84; K_p = 0.27; K_m = 17.26$

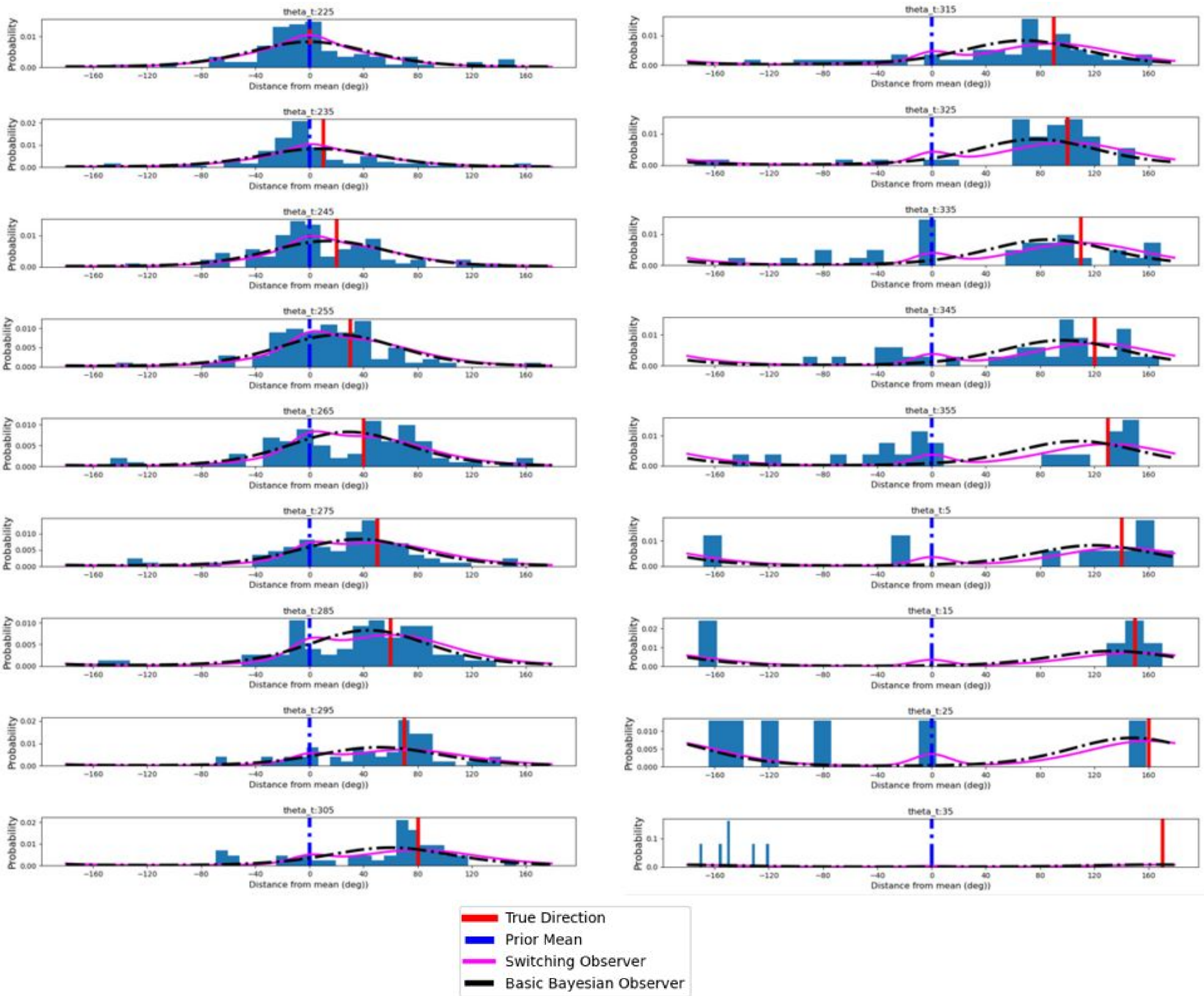

Figure A.4: BBO Vs SO (Subject: 1; Prior: 80°; Coherence: 6%)

True Direction: 225°to 35°(distance from prior mean)

BBO: $K_e = 52.7; K_p = 21.2; K_m = 1.77$

SO: $K_e = 1.84; K_p = 0.27; K_m = 17.26$

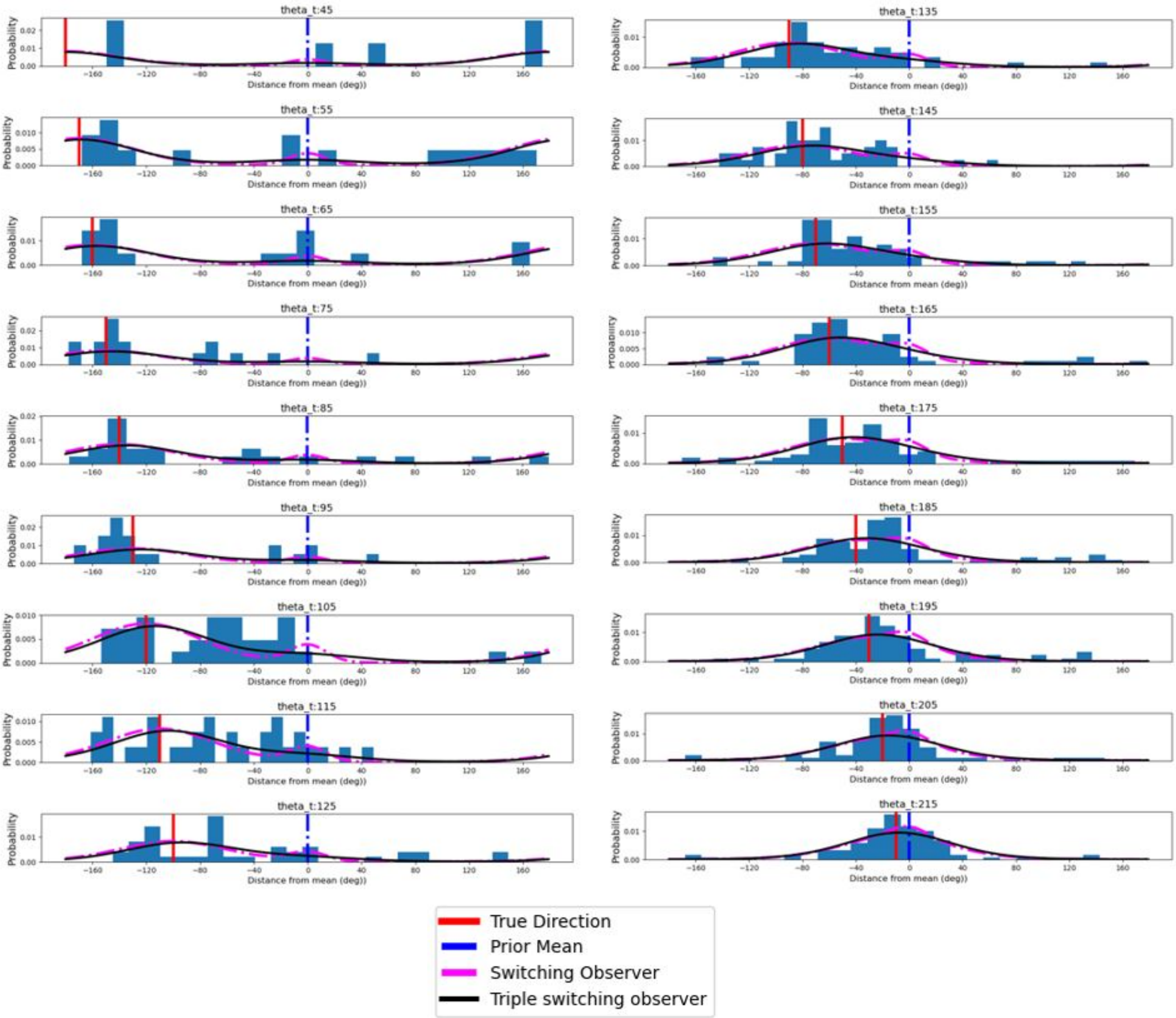

Figure A.5: SO Vs TSO (Subject: 1; Prior: 80°; Coherence: 6%)

True Direction: 45°to 215°(distance from prior mean)

SO: $K_e = 2.4; K_p = 0.4; K_m = 17.3$

TSO: $K_e = 107; K_p = 41.9; K_m = 2.22; P_{switch} = 0.234$

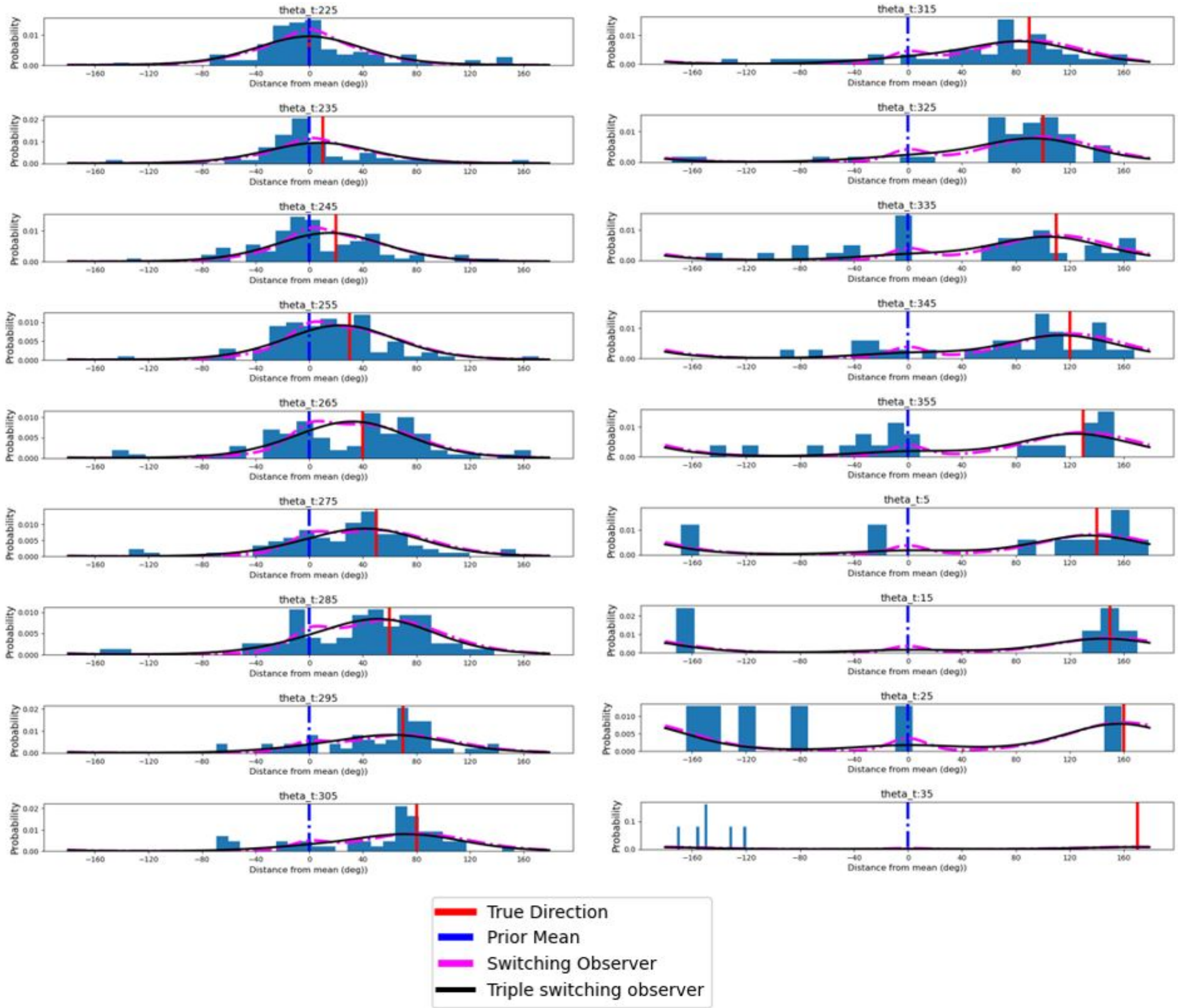

Figure A.6: SO Vs TSO (Subject: 1; Prior: 80°; Coherence: 6%)

True Direction: 225°to 35°(distance from prior mean)

SO: $K_e = 2.4; K_p = 0.4; K_m = 17.3$

TSO: $K_e = 107; K_p = 41.9; K_m = 2.22; P_{switch} = 0.234$

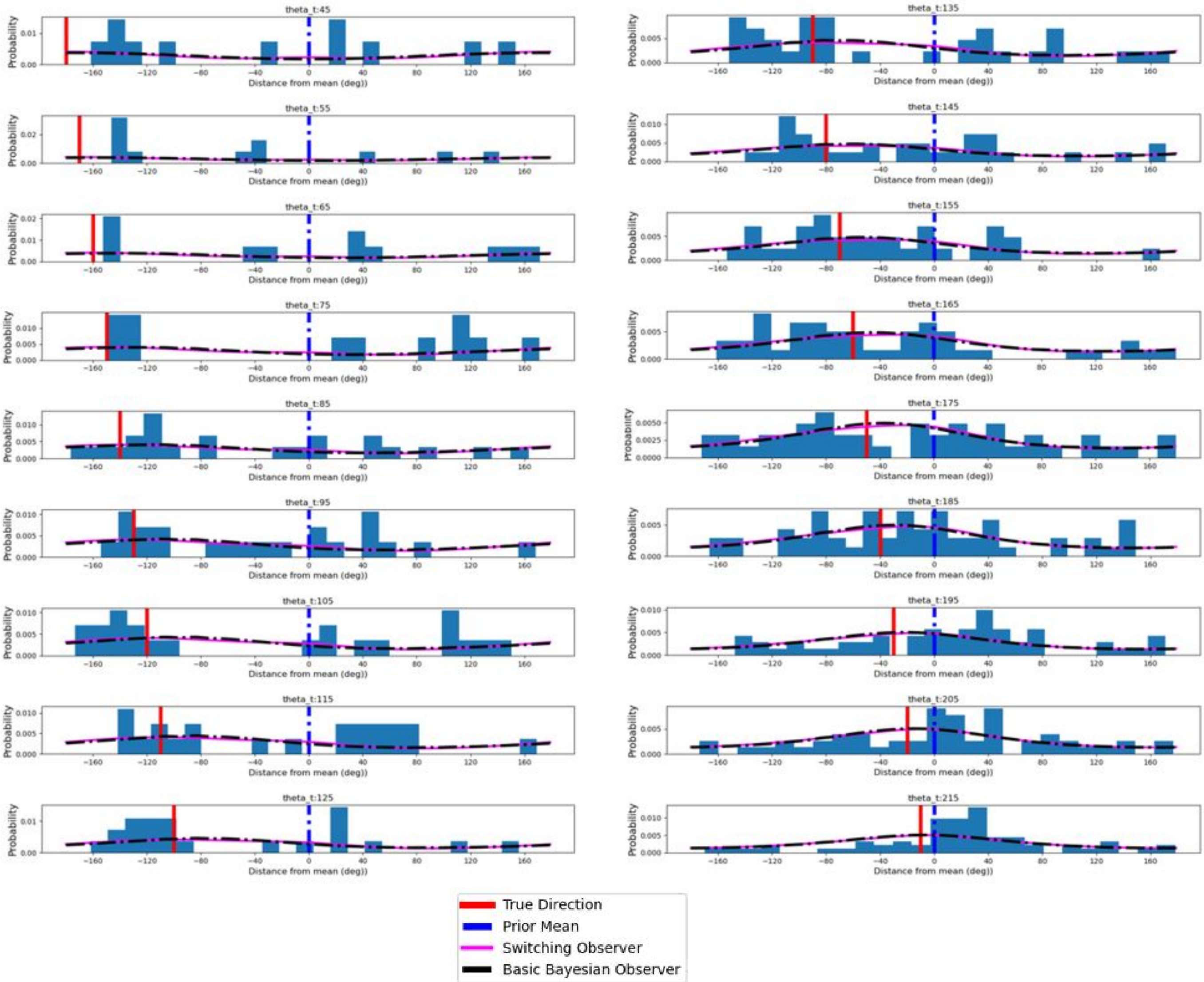

Figure A.7: BBO Vs SO (Subject: 10; Prior: 80°; Coherence: 6%)

- True Direction: 45°to 215°(distance from prior mean)

BBO: $K_e = 0.54; K_p = 0.08; K_m = 3.11$

SO: $K_e = 0.7; K_p = 0.07; K_m = 67.3$

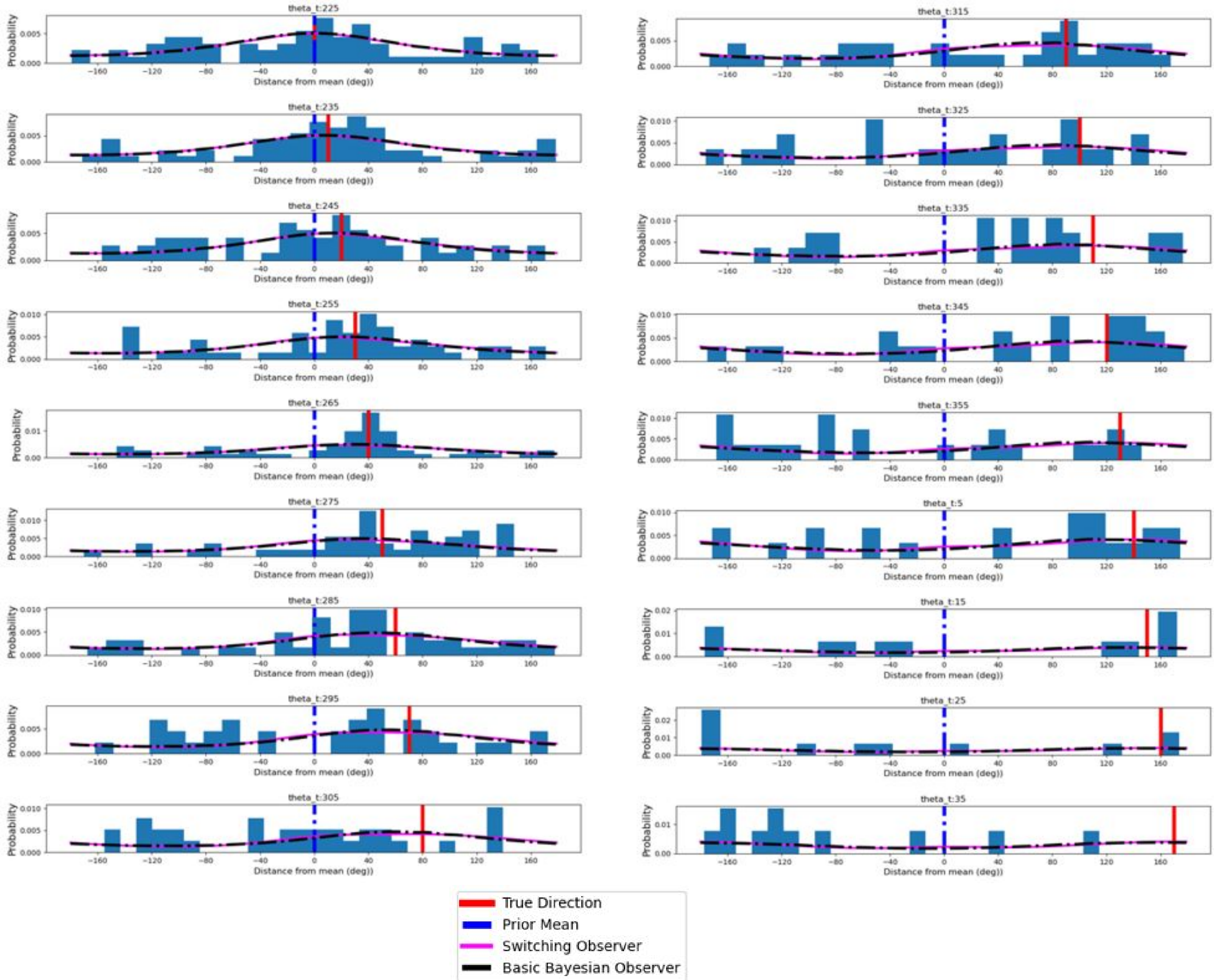

Figure A.8: BBO Vs SO (Subject: 10; Prior: 80°; Coherence: 6%)

- True Direction: 225°to 35°(distance from prior mean)

BBO: $K_e = 0.54; K_p = 0.08; K_m = 3.11$

SO: $K_e = 0.7; K_p = 0.07; K_m = 67.3$

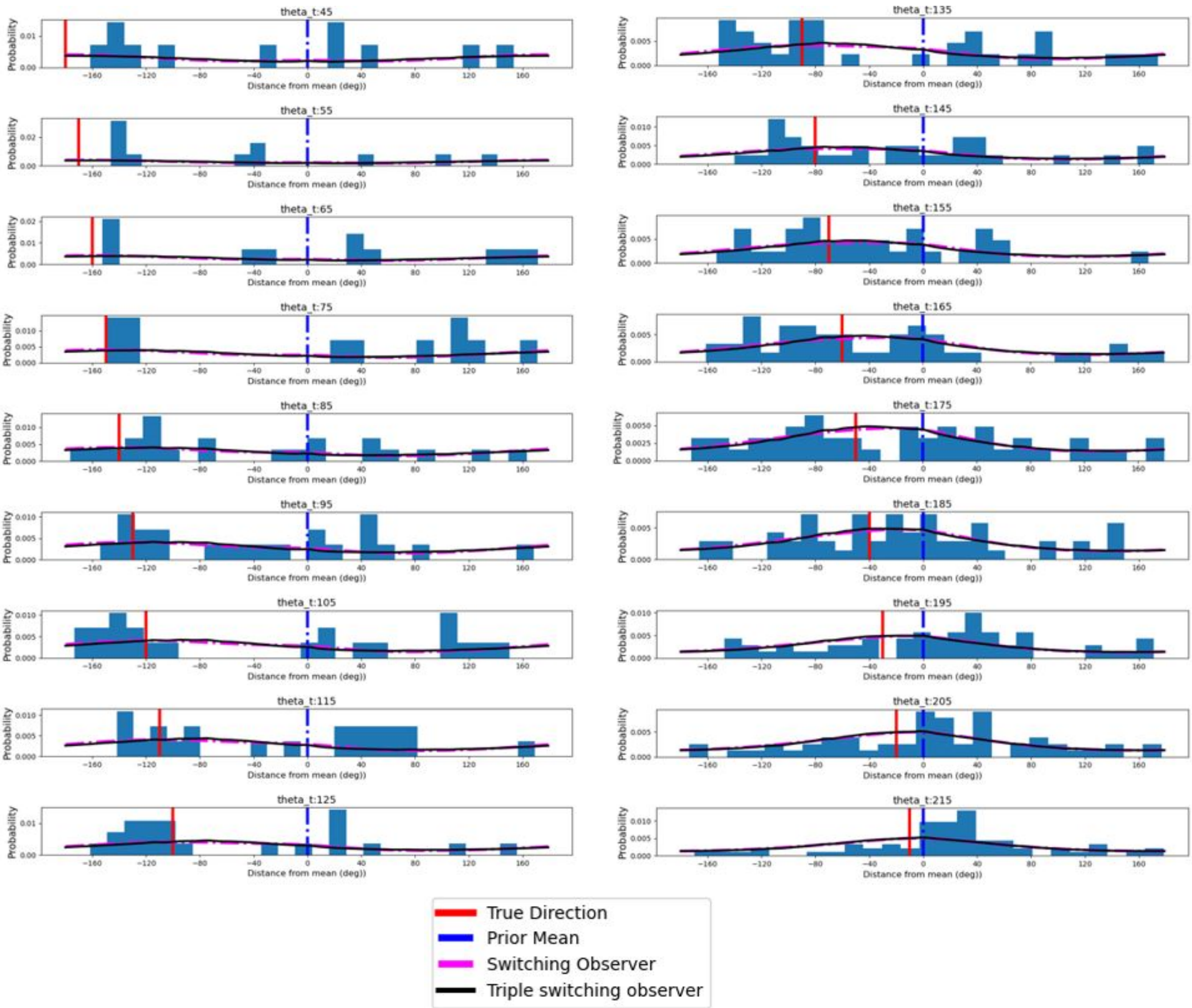

Figure A.9: SO Vs TSO (Subject: 10; Prior: 80°; Coherence: 6%)

- True Direction: 45°to 215°(distance from prior mean)

SO: $K_e = 0.7; K_p = 0.07; K_m = 3.11$

TSO: $K_e = 0.52; K_p = 0.08; K_m = 116; P_{switch} = 0.97$

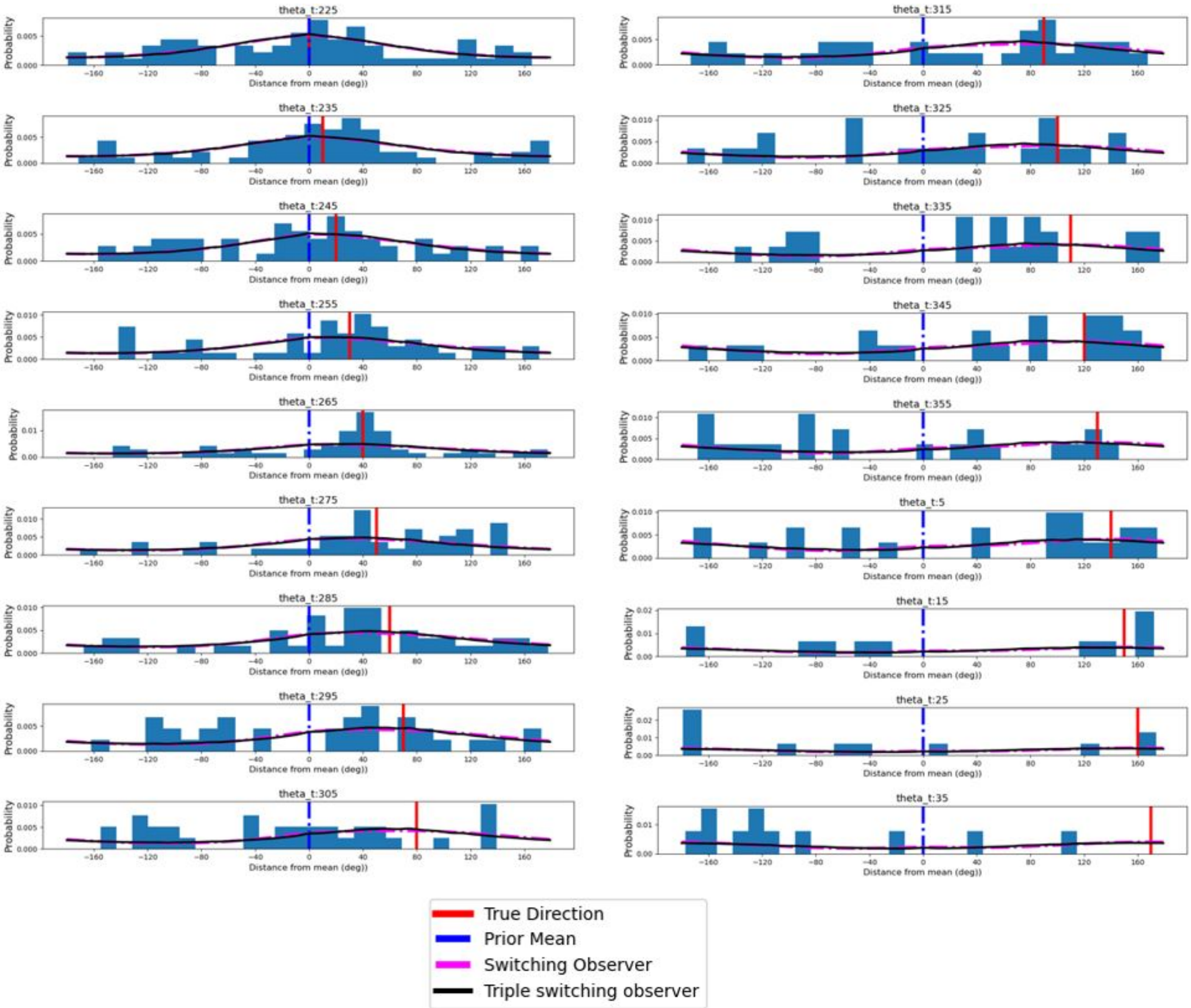

Figure A.10: SO Vs TSO (Subject: 10; Prior: 80°; Coherence: 6%)

- True Direction: 225°to 35°(distance from prior mean)

SO: $K_e = 0.7; K_p = 0.07; K_m = 3.11$

TSO: $K_e = 0.52; K_p = 0.08; K_m = 116; P_{switch} = 0.97$