A Framework for Testing Identifiability of Bayesian Models of Perception

Luigi Acerbi, Wei Ji Ma & Sethu Vijayakumar
{luigi.acerbi,weijima}@nyu.edu, sethu.vijayakumar@ed.ac.uk

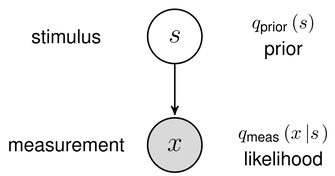
NIPS — Neural Information Processing Systems

Motivation

- Humans behave close to Bayesian observers in many psychophysical tasks.
- Bayesian Decision Theory (BDT) proposed as a *process model* of perception: correspondence between elements of the theory and representations in the brain.
- Several studies recover subjects' *internal representations* (priors, likelihoods, loss function) from behavioral data.
- But… **can we trust the recovered representations?**

Problem of Degeneracy

Consider the simple sensory estimation model:



Optimal estimate minimizes expected loss:

$$s^*(x) = \arg\min_{\hat{s}} \int \underbrace{q_{\text{meas}}(x|s)}_{\times\phi_1(s)} \underbrace{q_{\text{prior}}(s)}_{\times\phi_2(s)} \underbrace{\mathcal{L}(s,\hat{s})}_{\times\phi_3(s)} \, ds$$

Solution is **invariant** for $\prod_{i=1}^{3} \phi_i(s) = c \in \mathbb{R}$.

$\Longrightarrow$ Inverse mapping from observed behavior to internal representations is not unique [1].

Critical non-identifiability?

Degeneracy may be prevented by **theoretical** and **experimental** constraints. But role of various constraints not understood.

We need a *framework*.

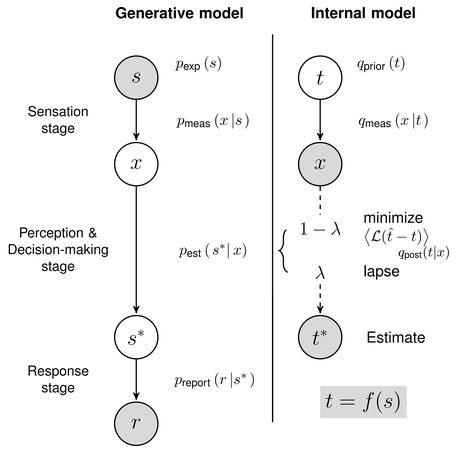
Framework

We propose a framework to study a priori identifiability of Bayesian models of sensory estimation. The framework consists of:

1. A unified observer model class that is both **general** and **efficient**.
2. A method to estimate identifiability between observer models in the class.

Bayesian Observer Model

General Bayesian observer model for (1-D) sensorimotor estimation tasks.



1. **Sensation stage:**
   Stimulus $s \longrightarrow$ noisy sensory measurement x.
2. **Perception stage:**
   Measurement x + prior $\longrightarrow$ posterior.
3. **Decision-making stage:**
   Minimization of the expected loss $\longrightarrow$ optimal estimate s^* (+ random lapses).
4. **Response stage:**
   Estimate s^* + report noise $\longrightarrow$ response r.

Model Details

- Sensory mapping $f(s) = \log\left[1 + (s/s_0)^d\right]$.
- $p_{\text{meas}}(x|s)$ parametrized as a unimodal mixture of 2 Gaussians with given mean $f(s)$, SD σ, skewness γ and kurtosis κ.
- $q_{\text{meas}}(x|t)$ as above with $\tilde{\sigma}, \tilde{\gamma}, \tilde{\kappa}$.
- $\mathcal{L}(\hat{t}-t)$ as above with $\sigma_\ell, \gamma_\ell, \kappa_\ell$.
- $q_{\text{prior}}(t) \sim \sum_m w_m \mathcal{N}\left(t|\mu_m, a^2\right)$.
- $\Longrightarrow$ Expected loss is a mixture of Gaussians (fast computation of optimal estimate [2]).
- Lapse $\sim q_{\text{prior}}(t)$ with lapse rate λ.
- Report noise $\sim \mathcal{N}\left(r|s^*, \rho_0^2 + \rho_1^2 s^{*2}\right)$.

Single observer model parametrized by $\boldsymbol{\theta}$:

$$\boldsymbol{\theta} = \left(\sigma, \gamma, \kappa, s_0, d, \tilde{\sigma}, \tilde{\gamma}, \tilde{\kappa}, \sigma_\ell, \gamma_\ell, \kappa_\ell, \{w_m\}_{m=1}^{M}, \lambda, \rho_0, \rho_1\right)$$

Mapping Identifiability

1. Reference true model $\boldsymbol{\theta}^*$, candidate model $\boldsymbol{\theta}$.
2. For each stimulus s_i in the experiment, compute distribution of responses:

$$p_{\text{resp}}(r|s_i;\boldsymbol{\theta}) = \int p_{\text{report}}(r|s^*) \times p_{\text{est}}(s^*|x) p_{\text{meas}}(x|s_i) \, ds^* \, dx.$$

3. Average support for $\boldsymbol{\theta}$ under true model $\boldsymbol{\theta}^*$:

$$\langle \log \Pr(\mathcal{D}|\boldsymbol{\theta}) \rangle_{\Pr(\mathcal{D}|\boldsymbol{\theta}^*)} = \text{const} \ldots - N_{\text{tr}} \sum_i P_i \cdot D_{\text{KL}}\left(p_{\text{resp}}(r|s_i;\boldsymbol{\theta}^*) || p_{\text{resp}}(r|s_i;\boldsymbol{\theta})\right)$$

   N_{tr} # of trials, P_i freq. of stimulus s_i.
4. Approximate posterior density of $\boldsymbol{\theta}$ under $\boldsymbol{\theta}^*$:

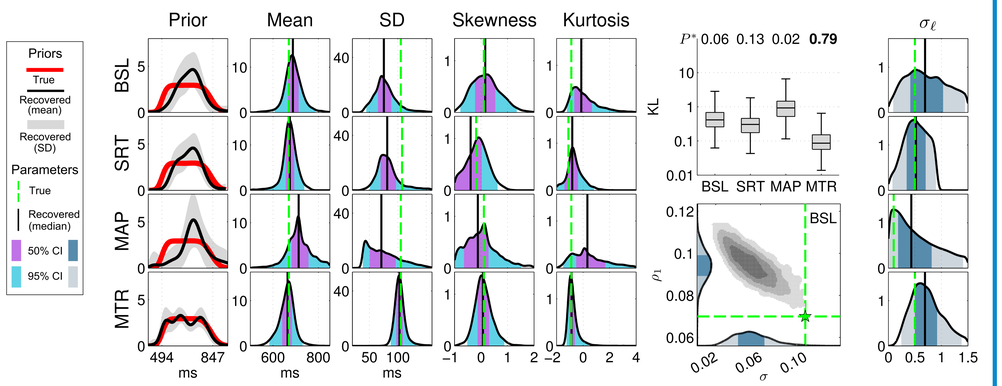
$$\mathcal{E}\left(\boldsymbol{\theta}|\boldsymbol{\theta}^*\right) \propto e^{\langle \log \Pr(\mathcal{D}|\boldsymbol{\theta}) \rangle}$$

5. Map identifiability by sampling from $\mathcal{E}\left(\boldsymbol{\theta}|\boldsymbol{\theta}^*\right)$.

Acknowledgments

We thank Jonathan Pillow, Paolo Puggioni, Peggy Seriès, and three anonymous reviewers for useful comments.
This research was partially supported by the EPSRC/BBSRC DTC studentship grant to LA. WJM is supported by award number R01EY020958 from the National Eye Institute and W911NF-12-1-0262 from the Army Research Office. SV is supported by Microsoft Research, Royal Academy of Engineering and EU FP7 programs.

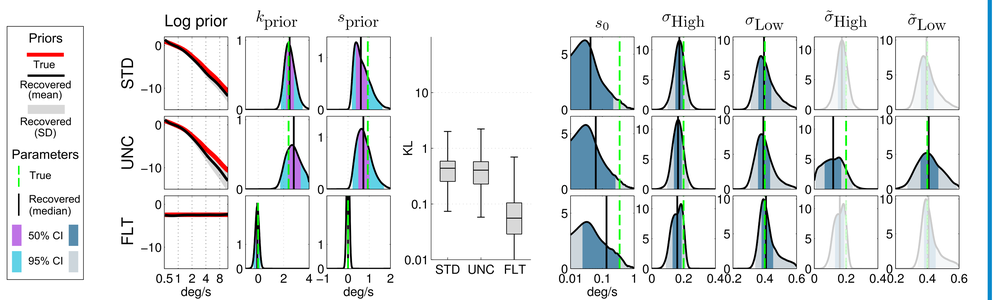
Case Study: Temporal Context and Interval Timing



1. BSL: Baseline (both 'short' and 'long' sessions).
2. SRT: Only 'short' session.
3. MAP: Experiment with narrow loss function.
4. MTR: Known motor noise.

Case study based on a time interval reproduction task [3]. We compare capability of recovering observer's priors q_{prior} and other parameters in four distinct setups. **Major issues of identifiability.**

Case Study: Slow-Speed Prior in Speed Perception



1. STD: Standard setup.
2. UNC: Uncoupled internal noise parameters.
3. FLT: Flat speed prior.

Case study based on a speed estimation task [4]. We compare capability of recovering observer's slow-speed priors (or lack thereof) and other noise parameters. **No issues of identifiability.**

References

[1] Mamassian, P. & Landy, M. S. (2010) It's that time again. *Nat Neurosci* 13, 914–916.
[2] Acerbi, L., Vijayakumar, S., & Wolpert, D. M. (2014) On the origins of suboptimality in human probabilistic inference. *PLoS Comput Biol* 10.
[3] Jazayeri, M. & Shadlen, M. N. (2010) Temporal context calibrates interval timing. *Nat Neurosci* 13, 1020–1026.
[4] Stocker, A. A. & Simoncelli, E. P. (2006) Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci* 9, 578–585.

Conclusions

- Applications of our framework:
  - Study robustness (or alternative explanations) of experimental findings.
  - Rank candidate experimental designs based on the identifiability of parameters of interest.