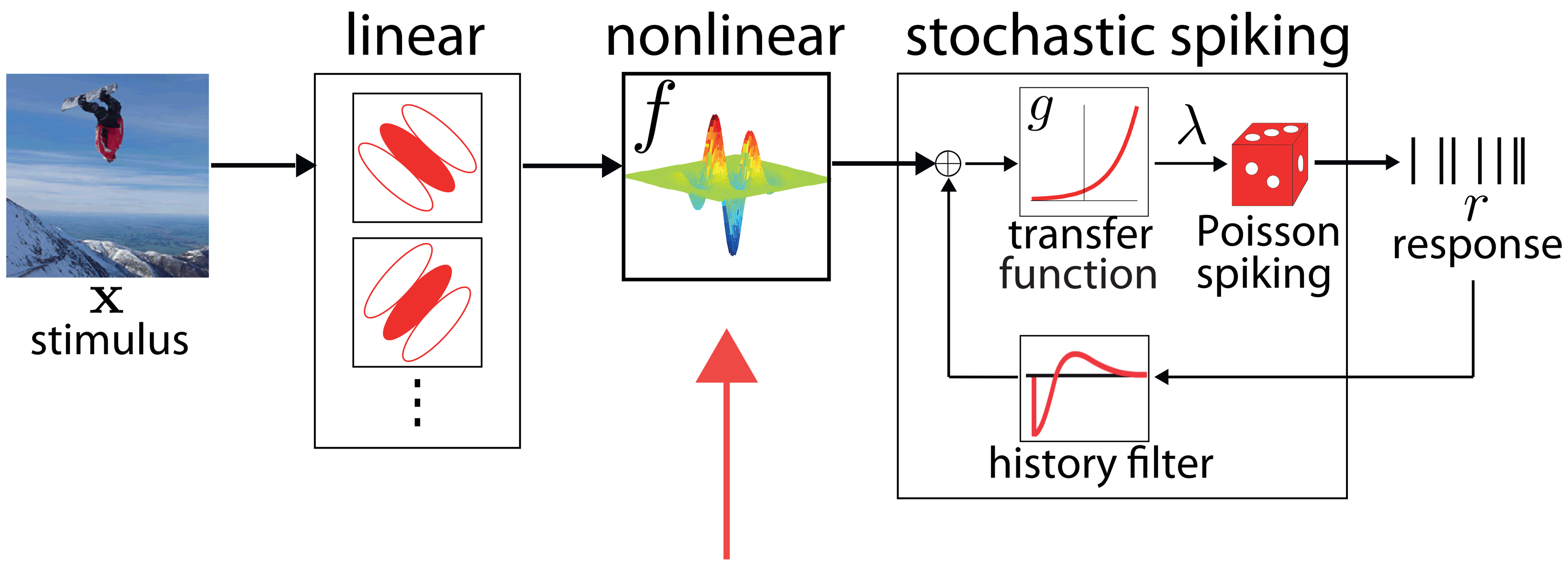
Active Learning of Neural Response Functions with Gaussian Processes

NIPS 2011 #T020

Mijung Park[1], Greg Horwitz[2], and Jonathan Pillow[3]

1: Electrical and Computer Engineering, The University of Texas at Austin, 2: Physiology and Biophysics, University of Washington,
3: Center for Perceptual Systems, Psychology and Neurobiology, The University of Texas at Austin

1. Neural characterization problem



Question: how to efficiently learn neural response nonlinearities in closed-loop experiments?

2. GP-Poisson encoding model

generative model

hyper-parameters θ

flat hyper-prior on $\theta = \{\mu_f, \rho, \tau\}$ (mean, variance, correlation scale)

nonlinearity f

prior

$$p(f|\theta) = \mathcal{N}(f|\mu_f, K)$$

$$k(\mathbf{x}_i, \mathbf{x}_j) = \rho \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{(2\tau)}\right)$$

$\mathbf{x}$ stimulus → r response

likelihood

$$p(r|\mathbf{x}, f) = \frac{1}{r!}\lambda^r e^{-\lambda}$$

$$\lambda = g(f(\mathbf{x}))$$

MAP inference

$$\mathbf{f}_{map} = \arg\max_{\mathbf{f}} \underbrace{\log p(\mathbf{f}|\mathbf{r}, \mathbf{X}, \theta)}$$

convex optimization

$$\mathbf{r}^T \log(g(\mathbf{f})) - \mathbf{1}^T g(\mathbf{f}) - \frac{1}{2}(\mathbf{f} - \mu_f)^T K^{-1}(\mathbf{f} - \mu_f) + c$$

3. Efficient evidence optimization for θ

- Set θ by maxizing the evidence

$$p(\mathbf{r}|\theta) = \int p(\mathbf{r}|\mathbf{f}) p(\mathbf{f}|\theta) d\mathbf{f}$$

Poisson likeli. Gaussian prior $\Longrightarrow$ intractable !

- Gaussian posterior by Laplace Approximation

$$p(\mathbf{f}|\mathbf{r}, \theta) \approx \mathcal{N}(\mathbf{f}|\mathbf{f}_{map}, \Lambda)$$

- Evidence at $\mathbf{f} = \mathbf{f}_{map}$

$$p(\mathbf{r}|\theta) \approx \frac{p(\mathbf{r}|X, \mathbf{f})\mathcal{N}(\mathbf{f}|\mu_f, K)}{\mathcal{N}(\mathbf{f}|\mathbf{f}_{map}, \Lambda)}$$

iterate!

- Numerical update of posterior & evidence for each θ: expensive!

numerical update evidence optimization

$\hat{\theta} \rightarrow$ $\boxed{\mathcal{N}(\mathbf{f}_{map}, \Lambda)}$ posterior $\rightarrow$ $\boxed{\mathbf{f}_{map} = \Lambda(H\mathbf{m} + K^{-1}\mu_f)}$

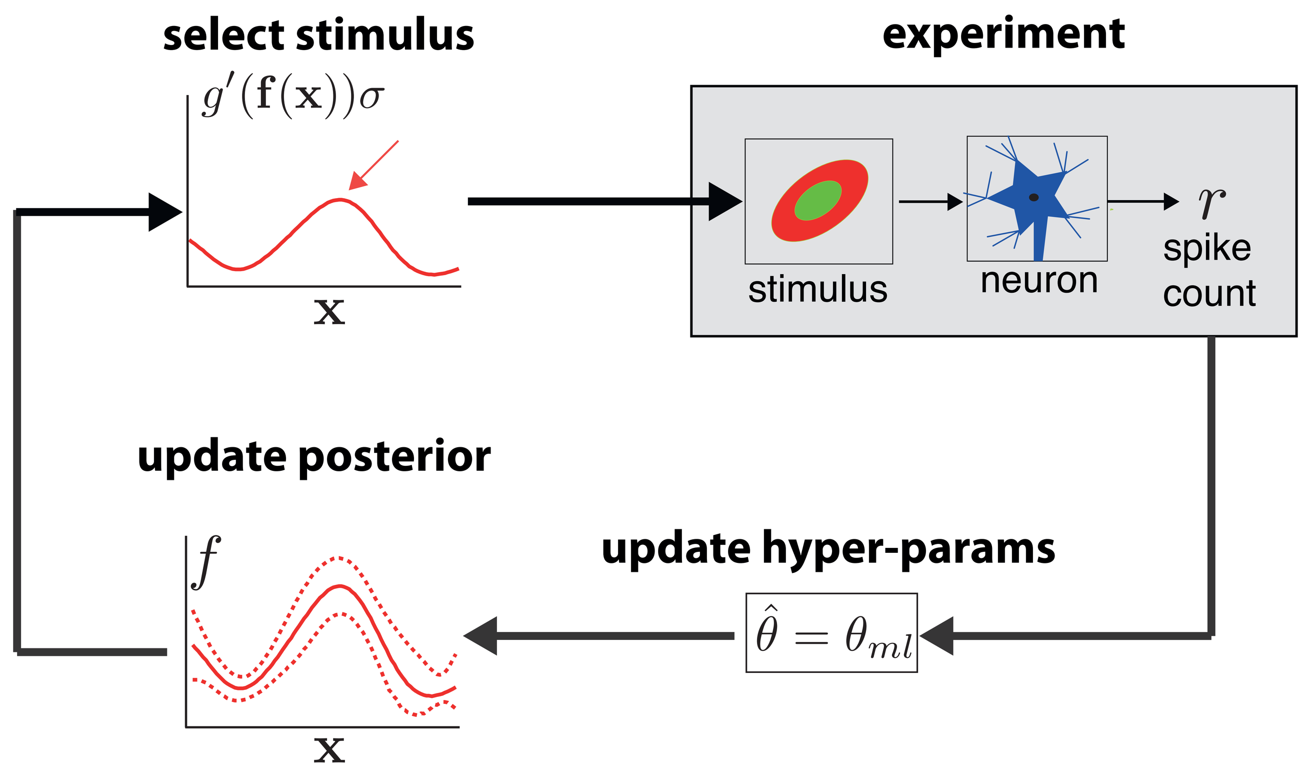
$\Downarrow$

$\mathcal{N}(\mathbf{m}, H^{-1})$ likelihood

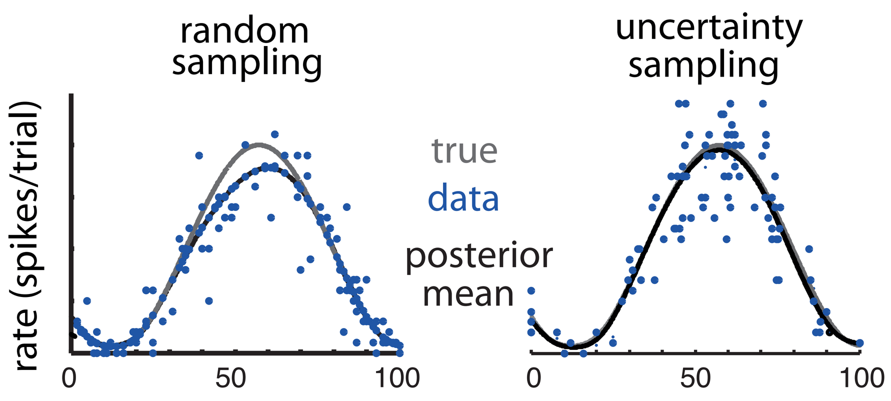
where $\Lambda = (H + K^{-1})^{-1}$

4. Active learning: uncertainty sampling

- select stimuli that minimize posterior uncertainty on $g(\mathbf{f})$
- delta method: $\mathbf{x}^* = \arg\max_{\mathbf{x}} g'(\mathbf{f}(\mathbf{x}))\sigma$ (posterior standard deviation on f)
- Adaptive stimulus selection

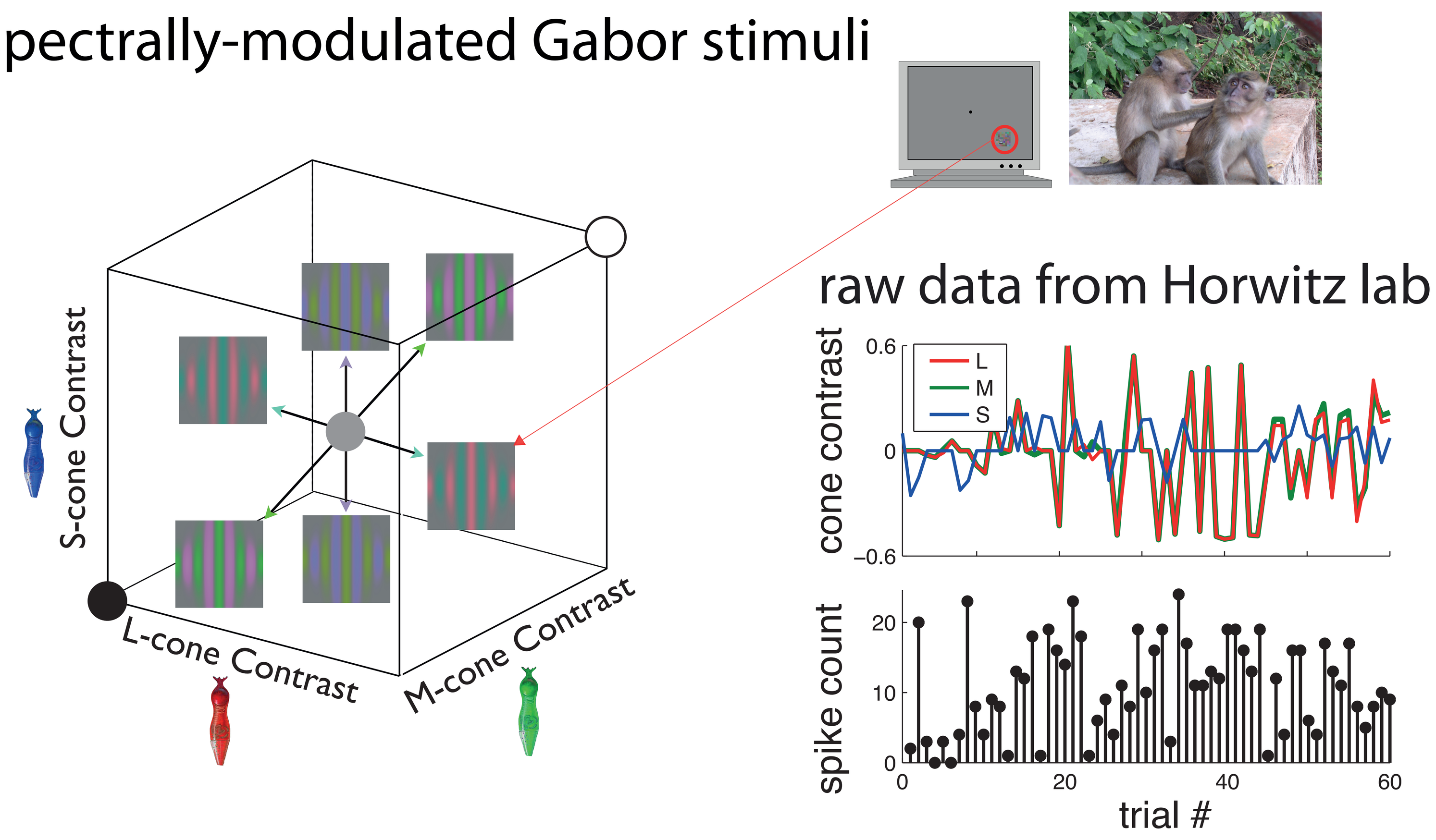


- 1D simulated example:

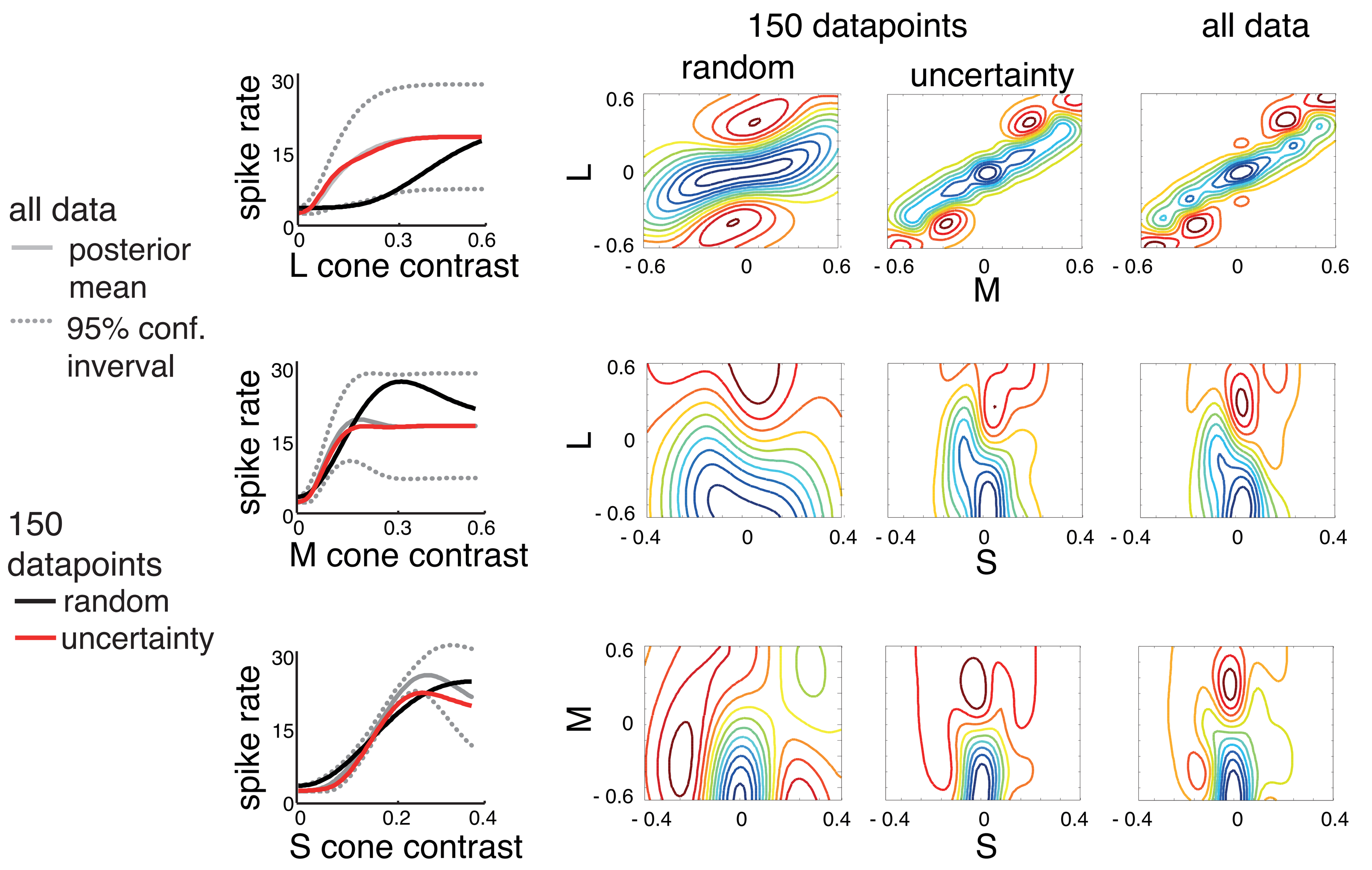


5. Applicaton to color tuning of a V1 complex cell

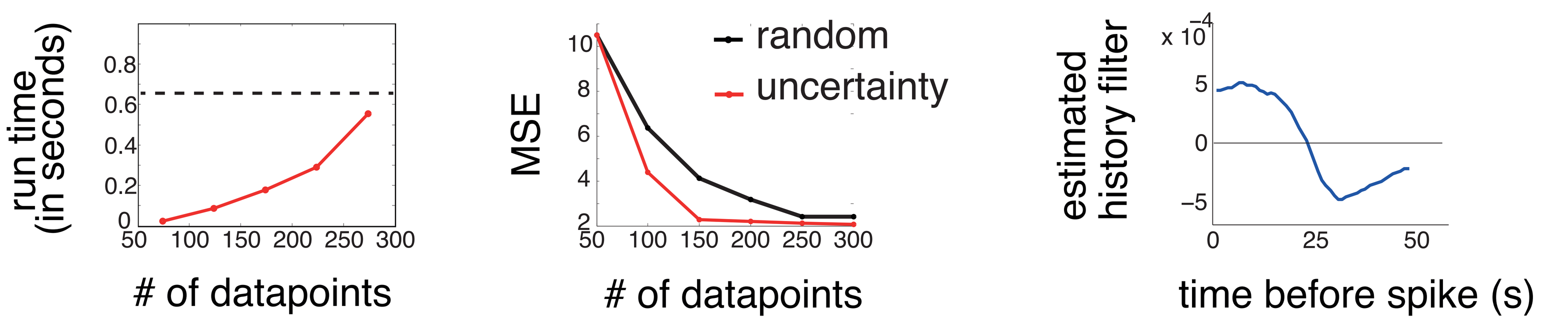
- color-tuned neurons in macaque V1
- spectrally-modulated Gabor stimuli



- 1D/2D slices of 3D color tuning function



- run time, performance measure, history effect



6. Conclusion

- flexible GP-Poisson model for neural response nonlinearities
- optimal design based on uncertainty sampling
- rapid learning of nonlinearities in closed-loop experiments