Empirical models of spiking in neural populations

Jakob H. Macke, Lars Büsing,
John P. Cunningham, Byron M. Yu,
Krishna V. Shenoy and Maneesh Sahani

NIPS 2011
December 15, 2011
Characterizing neural function

- Single neuron recordings

![Image of neurons and graph](image.png)
Characterizing neural function

- Single neuron recordings

- Simultaneous recordings from populations of neurons

→ need for appropriate statistical models
Models of population activity: direct couplings or latent variables?

**model class I**

**generalized linear models (GLM)**

e.g. [Chornoboy et al., Biol Cybern, 1988]
[Paninski et al., Neuro Comp, 2004]
Models of population activity: direct couplings or latent variables?

model class I

generalized linear models (GLM)

direct couplings

model class II

latent dynamical system (DS)

dynamical system

external input  direct couplings  observed neurons

external input  dynamical system  observed neurons

e.g. [Chornoboy et al., Biol Cybern, 1988] [Paninski et al., Neuro Comp, 2004]

[Gat et al., 97, Network], [Yu et al., NIPS, 2006]
Statistical models offer insights into neural computation

**model class I**

- GLMs improve decoding of visual stimuli from retina recordings

[Diagram of coupled spiking model]

[Pillow et al.,Nature, 2008]
Statistical models offer insights into neural computation

**model class I**

- GLMs improve decoding of visual stimuli from retina recordings

- [Pillow et al., Nature, 2008]

**model class II**

- DS facilitate single trial analysis of cortical multi-electrode recordings

- [Chruchland et al, Nature Neurosci., 2010]

- Brain-computer interfaces

- [Wu et al., IEEE Tans. Neural Systems, 2000]
Statistical models offer insights into neural computation

**model class I**
- GLMs improve decoding of visual stimuli from retina recordings

![Coupled spiking model diagram](image)

[Chruchland et al., Nature Neurosci., 2010]

**model class II**
- DS facilitate single trial analysis of cortical multi-electrode recordings

![Brain-computer interfaces diagram](image)

[Chruchland et al., Nature Neurosci., 2010]

What are good models for cortical multi-electrode recordings?
Data: Multi-electrode recordings from primate cortex

- Data from Shenoy’s lab
- Array with 96 electrodes in motor / premotor areas
- 92 units = dimensions, 120 s of recordings, 10 ms binning
Model class I: Generalized linear models (GLMs)

- Instantaneous firing rate $\lambda_i(t)$ of neuron $i$

\[
\lambda_i(t) = \exp\left( \text{det. innovations} \cdot D \cdot \text{spike history} \right)
\]

- We used up to five different decaying exponentials as basis functions for the spike history:
  - Decaying time constants 0.1 - 80 ms

- We used $L_1$-regularization of the coupling matrix $D$ to avoid over-fitting

- e.g. [Chornoboy et al., Biol Cybern, 1988], [Paninski et al., Neuro Comp, 2004]
Model class I: Generalized linear models (GLMs)

- Instantaneous firing rate $\lambda_i(t)$ of neuron $i$

\[
\lambda_i(t) = \exp\left(\text{det. innovations} \ b(t) + \text{couplings} \cdot \text{spike history} \ s(t)\right)
\]

- Poisson observations

\[
y_i(t) \mid s(t) = \text{Poisson}(\lambda_i(t))
\]

- We used up to five different decaying exponentials as basis functions for the spike history:
  - time constants 0.1 - 80 ms

- We used $L_1$-regularization of the coupling matrix $D$ to avoid over-fitting

e.g. [Chornoboy et al., Biol Cybern, 1988], [Paninski et al., Neuro Comp, 2004]
Model class I: Generalized linear models (GLMs)

- Instantaneous firing rate $\lambda_i(t)$ of neuron $i$
  \[
  \lambda_i(t) = \exp([\overbrace{b(t)}^{\text{det. innovations}} + \overbrace{D \cdot s(t)}^{\text{spike history}}]_i)
  \]

- Poisson observations
  \[
  y_i(t) \mid s(t) = \text{Poisson}(\lambda_i(t))
  \]

- We used up to five different decaying exponentials as basis functions for the spike history:
  time constants 0.1 - 80 ms

E.g. [Chornoboy et al., Biol Cybern, 1988], [Paninski et al., Neuro Comp, 2004]
Model class I: Generalized linear models (GLMs)

- Instantaneous firing rate $\lambda_i(t)$ of neuron $i$

  $\lambda_i(t) = \exp(b(t) + D \cdot s(t))$

- Poisson observations

  $y_i(t) \mid s(t) = \text{Poisson}(\lambda_i(t))$

- We used up to five different decaying exponentials as basis functions for the spike history:
  time constants 0.1 - 80 ms

- We used $L_1$-regularization of the coupling matrix $D$ to avoid over-fitting

  e.g. [Chornoboy et al., Biol Cybern, 1988], [Paninski et al., Neuro Comp, 2004]
Model class II: Latent dynamical system (DS) models

- Latent variable $\mathbf{x}(t)$ evolves linearly, with Gaussian innovations

$$\mathbf{x}(t + 1) = A \mathbf{x}(t) + \mathbf{b}(t) + \eta(t)$$

- Gaussian linear dynamical system (GLDS):

$$\mathbf{y}(t) | \mathbf{x}(t) \sim N(C_x(t) + d, R)$$

- Poisson linear dynamical system (PLDS):

$$\mathbf{y}_i(t) | \mathbf{x}(t) \sim \text{Poisson}(\exp(C_i, x(t) + d_i + D_i s_i(t)))$$

- Parameters were learned by the EM algorithm

- We used a global Laplace approximation to the posterior for the PLDS
Model class II: Latent dynamical system (DS) models

- Latent variable $x(t)$ evolves linearly, with Gaussian innovations

$$x(t + 1) = Ax(t) + b(t) + \eta(t)$$

- Gaussian linear dynamical system (GLDS):

$$y(t) | x(t) \sim \mathcal{N}(Cx(t) + d, R)$$
Model class II: Latent dynamical system (DS) models

- Latent variable $x(t)$ evolves linearly, with Gaussian innovations

$$x(t+1) = A \underbrace{x(t)}_{\text{dynamics}} + b(t) + \eta(t)$$

- Gaussian linear dynamical system (GLDS):

$$y(t) \mid x(t) \sim \mathcal{N}(Cx(t) + d, R)$$

- Poisson linear dynamical system (PLDS):

$$y_i(t) \mid x(t) \sim \text{Poisson}(\exp(C_i \cdot x(t) + d_i + D_i s_i(t)))$$
Latent variable $\mathbf{x}(t)$ evolves linearly, with Gaussian innovations:

$$
\mathbf{x}(t + 1) = A \mathbf{x}(t) + \mathbf{b}(t) + \eta(t)
$$

- Gaussian linear dynamical system (GLDS):
  $$
  y(t) \mid \mathbf{x}(t) \sim \mathcal{N}(C\mathbf{x}(t) + \mathbf{d}, \mathbf{R})
  $$

- Poisson linear dynamical system (PLDS):
  $$
  y_i(t) \mid \mathbf{x}(t) \sim \text{Poisson}(\exp(C_i, \mathbf{x}(t) + \mathbf{d}_i + D_i\mathbf{s}_i(t)))
  $$

Parameters were learned by the EM algorithm.

We used a global Laplace approximation to the posterior for the PLDS.
Measuring performance: predicting neural firing from the population

- On test trials, for every neuron $i$:
  - Compute cross-prediction $\mathbb{E}[y_{i,:}|y_{\setminus i,:}]$ of neuron $y_{i,:}$ given all other neurons $y_{\setminus i,:}$
  - Compute MSE between prediction and true trajectory

- We report improvement over a constant predictor (i.e. higher = better)
Cross-prediction: Dynamical systems outperform GLMs

- GLM performance as a function of $L_1$-regularization parameter: controls zero-entries in coupling matrix
- PLDS with 5-dimensional state space
- For our recordings PLDS outperforms GLMs with different history filters
Cross-prediction: Poisson observations improve performance

- Poisson observation model improves performance over Gaussian models
- Single neuron history filters result in an slight increase in performance
Matching statistics of the data I: temporal cross-correlations

Latent dynamical systems (DS)

- Data: peaks at zero time lag and broad flanks
- DS: matches cross-correlations well
Matching statistics of the data I: temporal cross-correlations

**latent dynamical systems (DS)**

- GLDS
- PLDS
- PLDS 100ms
- recorded data

**generalized linear models (GLM)**

- GLM 10ms
- GLM 100ms
- GLM 150s
- recorded data

- **Data**: peaks at zero time lag and broad flanks
- **DS**: matches cross-correlations well
- **GLM**: difficulties to reproduce peak due to lack of instantaneous couplings
Matching statistics of the data II: population spike counts

- Gaussian DS: under-estimates large population events
- Poisson DS: good match to the data
- GLM: no regularization; GLM2: optimal regularization
How much variance of the data can we explain using cross-prediction?

% of explained variance

For small bins, more variance can be explained than would be possible if firing condition on latent state were Poisson. PLDS models can explain up to twice as much variance as a model with independent activity conditioned on PSTH.
How much variance of the data can we explain using cross-prediction?

% of explained variance

- For small bins, more variance can be explained than would be possible if firing conditioned on latent state were Poisson

![Graph showing the percentage of variance explained against bin size for different models, including PLDS models up to twice as much variance as PSTH predictor with Poisson activity.](image)
How much variance of the data can we explain using cross-prediction?

% of explained variance

Comparison with PSTH predictor

- For small bins, more variance can be explained than would be possible if firing conditioned on latent state were Poisson
- PLDS models can explain up to twice as much variance as a model with independent activity conditioned on PSTH
Summary & conclusions

- Dynamical systems match the investigated cortical recordings better than GLMs
  - Cross-prediction
  - Various statistics, e.g. cross-correlation structure

GLM modelling assumptions (direct couplings) might not be appropriate for this type of data:
- Multi-electrode arrays sample a neural population very sparsely
- Most inputs presumably come from outside the sampled population

Latent dynamical systems are arguably the more natural models for multi-electrode recording from cortex
Summary & conclusions

- Dynamical systems match the investigated cortical recordings better than GLMs
  - Cross-prediction
  - Various statistics, e.g. cross-correlation structure

- GLM modelling assumptions (direct couplings) might not be appropriate for this type of data:
  - Multi-electrode arrays sample a neural population very sparsely
  - Most inputs presumably come from outside the sampled population
Summary & conclusions

- Dynamical systems match the investigated cortical recordings better than GLMs
  - Cross-prediction
  - Various statistics, e.g. cross-correlation structure

- GLM modelling assumptions (direct couplings) might not be appropriate for this type of data:
  - Multi-electrode arrays sample a neural population very sparsely
  - Most inputs presumably come from outside the sampled population

- Latent dynamical systems are arguably the more natural models for multi-electrode recording from cortex
Acknowledgements

- **Jakob Macke** (Gatsby Unit, UCL)
- Maneesh Sahani (Gatsby Unit, UCL)
- Krishna Shenoy & lab (Stanford)
- John Cunningham (Cambridge)
- Byron Yu (Carnegie Mellon)

Funding:

- Gatsby Charitable Foundation
- EU-FP7 Marie Curie Fellowship
- DARPA: REPAIR N66001-10-C-2010
- EPSRC EP/H019472/1
- NIH CRCNS R01-NS054283
- NIH Pioneer 1DP1OD006409
Additional material
Cross-prediction: area under ROC

- Same qualitative results as measured by MSE
- Benefits of history filters for PLDS are more pronounced
Correlations in data are well reflected by DS models

**data bin at** 10 ms

**5-dimensional PLDS**