Moment-closure approaches to statistical mechanics and inference in models of neural dynamics

Michael E. Rule[3], David Schnoerr[2], Matthias H. Hennig[1], Guido Sanguinetti[1]

[1]University of Edinburgh, Edinburgh UK; [2]Imperial College London, London UK; [3]University of Cambridge, Cambridge UK

1. From Autoregressive Point-Processes to Latent State-Space Models[1]

- *Point-Process Generalized Linear Models (PPGLMs)* are a popular statistical analysis tool for predicting Pr(spikes)
- Modeling stochastic effects: important for nonlinear recurrent systems
- Existing methods neglect recurrent stochastic effects of spiking
- **Goal:** *Build State-Space Model (SSM) representation of PPGLM*

Spiking as a Poisson point-process

$y(t) = \sum_{\tau \in \text{events}} \delta(t-\tau)$ — Emitted/observed spike train

$\lambda = \exp(H^\top y + \text{input})$ — Conditional intensity (firing rate)

$y \sim \text{Poisson}(\lambda \cdot dt)$ — Conditionally Poisson spiking

Langevin approximation:

- Consider spiking history process $h(\tau,t) = y(t-\tau)$
- History $h(\tau,t)$ is a delay line, with stochastic spiking input $y(t)$:

$$\partial_t h = \delta_{\tau=0} y(t) - \partial_\tau h$$

- Approximate Poisson spiking $\approx$ Gaussian with $\sigma^2 = \mu$:

$$\text{Poisson}(\lambda \cdot dt) \approx \lambda dt + \sqrt{\lambda} dW \Rightarrow$$
$$dh \approx (\delta_{\tau=0}\lambda - \partial_\tau h)\, dt + \delta_{\tau=0}\sqrt{\lambda} dW \qquad (1)$$

2. Model stochastic dynamics via moments

Model first two cumulants of Pr $h(\tau, t)$:

$$\mu = \langle h \rangle \qquad \partial_t \mu = \delta_{\tau=0}\langle\lambda\rangle - \partial_\tau \mu. \qquad (2)$$
$$\Sigma = \langle hh^\top \rangle - \mu\mu^\top \qquad \partial_t \Sigma = J\Sigma + \Sigma J^\top + Q \qquad (3)$$

(J: Jacobian of $\partial_t h$; Q: Poisson noise. Depend on μ, Σ!)

Evolution depends on higher moments; *Approximations* needed for closed equations (**moment closure**):

- *Log-normal:* Let $\ln\lambda \sim \mathcal{N}\left(\mu = H^\top\mu + \text{input}, \sigma^2 = H^\top \Sigma H\right)$

$$\langle\lambda\rangle \approx \exp\left(H^\top\mu + \text{input} + \tfrac{1}{2}H^\top\Sigma H\right) \qquad (4)$$

- *2nd-order[2]:* Taylor approximate $\lambda = f(a+\epsilon)$ as $\lambda \approx f(a) + \epsilon \cdot f'(a)$

$$\langle\lambda\rangle \approx \exp\left(H^\top\mu + \text{input}\right)\left(1 + \tfrac{1}{2}H^\top\Sigma H\right) \qquad (5)$$

- Moment-closure for Σ more involved, see Rule & Sanguinetti (2018).

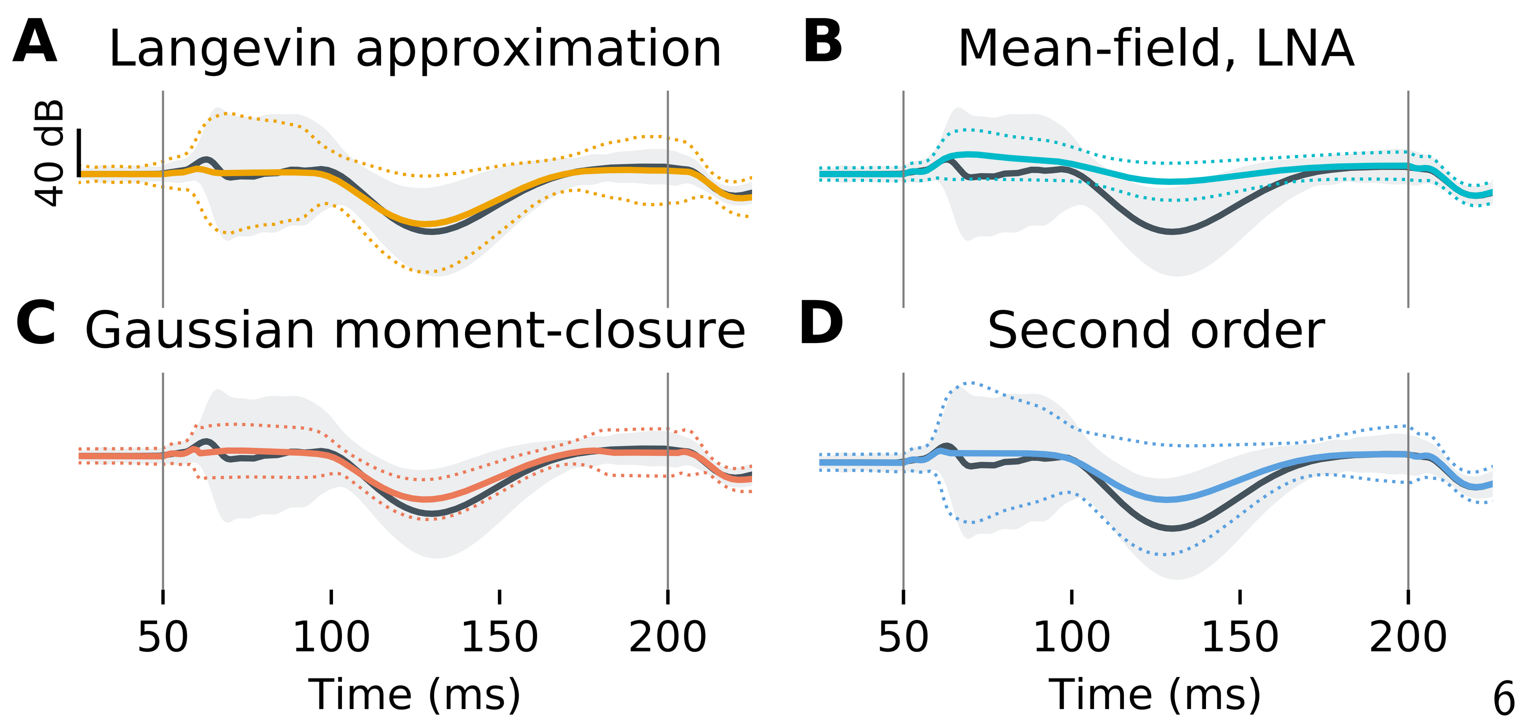

Figure 1. Moment closure captures slow timescales in the mean, and fast timescales in the variance. Four approximations of the mean (black) and 1σ (shaded) of the ln λ of a phasic bursting neuron model[3] driven by a .3 pA current (vertical lines). The Langevin approximation (**A**, Eq. 1) retains slow-timescale features. A mean-field model (**B**) misses the impact of fluctuations, which are better captured by the Log-Gaussian moment-closure (**C**, Eq. 4). The 2nd-order model (**D**, Eq. 5) better estimates variance.

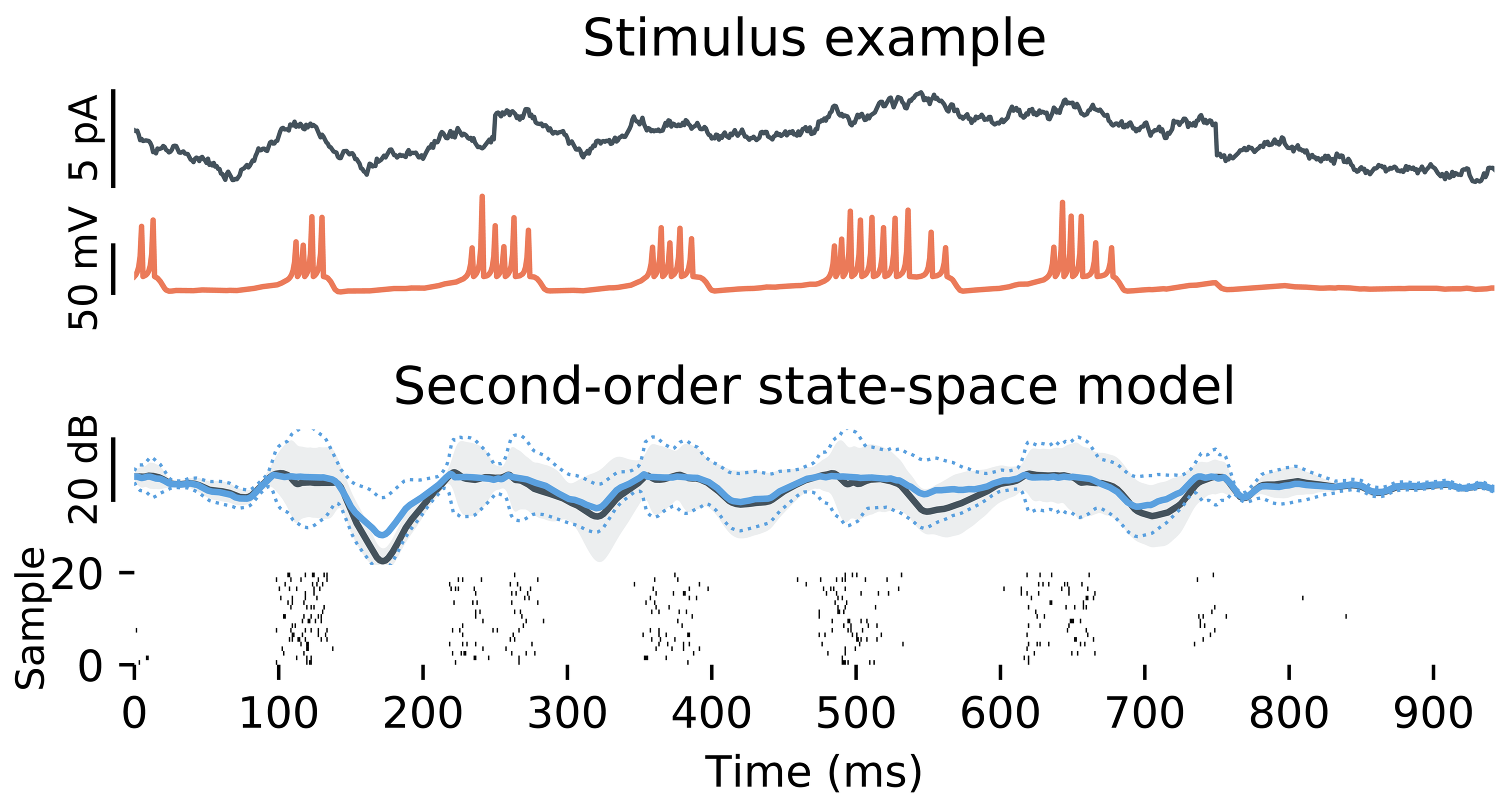

Figure 2. The moment-closure state-space model retains coarse-timescale characteristics of the original system. A 2nd-order moment closure (blue) yields a low-order model of the moments of the autoregressive PPGLM, which was trained to approximate a phasic-bursting Izhekevich neuron[3] (stimulus:black, voltage:red).

in summary

Using moment closure, we convert PPGLM neuronal models to low-order SSMs that track the distribution of paths in the process history. (*Code available at* `bit.ly/2Hj84cd`)

AR-PPGLM→SSM: *new ways to train*

» Bayesian filtering: infer moments & model likelihood from spikes

» Future: apply methods from training SSM to PPGLM?

[1] Rule ME et al. Autoregressive point-processes as latent state-space models: a moment-closure approach to fluctuations and autocorrelations. *Neural Computation*, 30(10):2757–2780, 2018.

[2] Ale A, et al. A general moment expansion method for stochastic kinetic models. *The Journal of chemical physics*, 138(17):174101, 2013.

[3] Weber AI et al. Capturing the dynamical repertoire of single neurons. *Neural computation*, 29(12):3260–3289, 2017.

3. Neural Field Models & Latent State Inference[4]

- *Neural field models* describe spatiotemporal population activity.
- 3-state model[5]: Quiescent/Active/Refractory
- Rich wave dynamics without inhibition (like retinal waves[6])
- **Goal:** *Infer neural-field states driving spatiotemporal spiking*

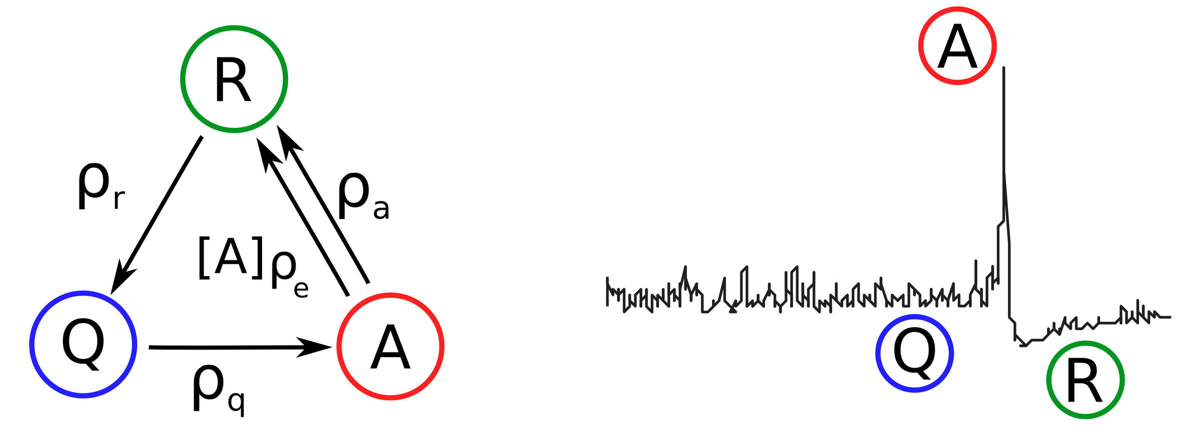

Figure 3. Quiescent-Active-Refractory model of neural waves. The Q-A-R states correspond to the Critical-Active-Stable retinal wave model in (6).

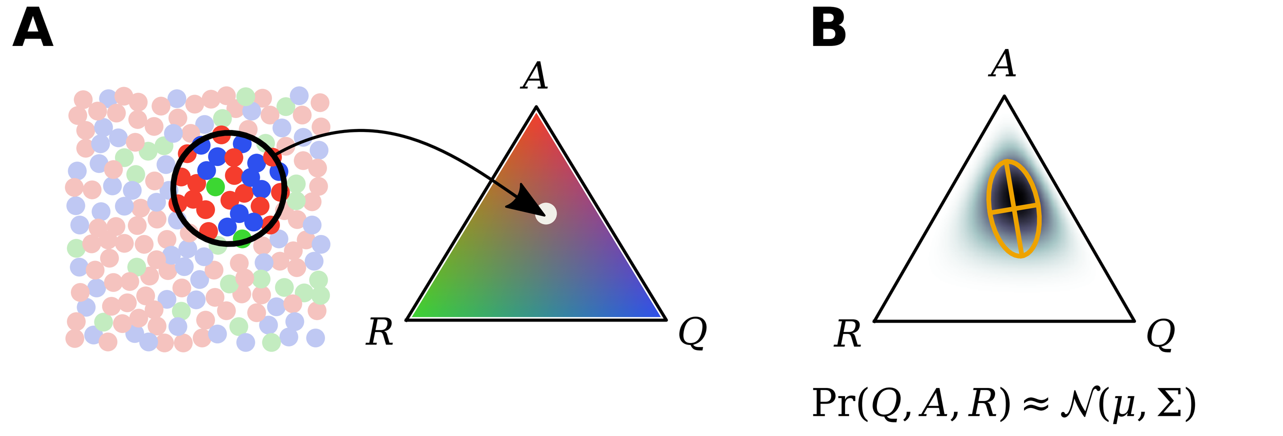

Figure 4. Population moments summarize neural states. (**A**) Local activity is summarized by the # of cells in each state; (**B**) Gaussian moment closure approximates the distribution of population states with μ & Σ (orange).

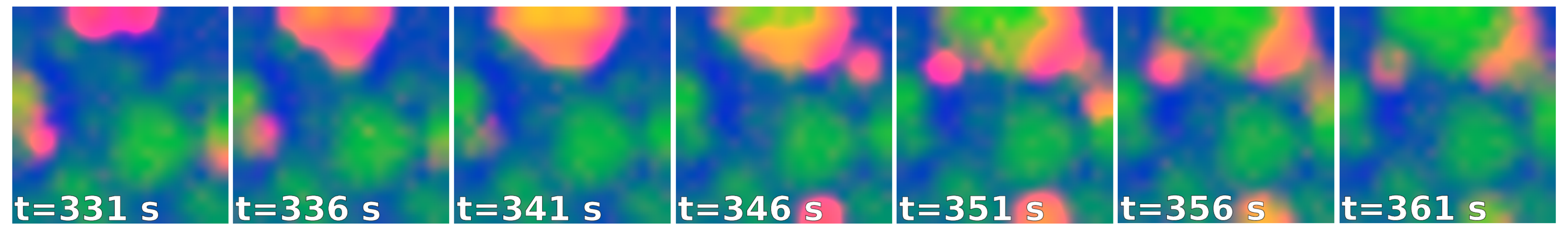

Figure 5. Simulated 3-state model exhibits wave patterns at diverse scales. Excitation of quiescent cells (blue) leads to propagating waves (red), which create localized refractory patches (green), recapitulating the critical wave activity studied in (6).

Moment equations (Gaussian moment closure)

Means: functions $\mu_Q(x)$, $\mu_A(x)$, $\mu_R(x)$ that depend on location x

- Evolve according to the rate cells enter and leave each state:

$$\partial_t \mu_Q(x) = r_{rq}(x) - r_{qa}(x)$$
$$\partial_t \mu_A(x) = r_{qa}(x) - r_{ar}(x)$$
$$\partial_t \mu_R(x) = r_{ar}(x) - r_{rq}(x)$$

- Transitions $r_{ar}(x) = \rho_a \mu_A(x)$ and $r_{rq}(x) = \rho_r \mu_R(x)$ are local & linear.
- *Nonlinear excito-excitatory interaction depends on nearby activity:*

$$r_{qa}(x) = \rho_q \mu_Q(x) + \rho_e \int k(x-x') \langle Q(x)A(x')\rangle \, dx',$$

  - $k(\Delta x)$ is a spatial convolution kernel $\propto \exp(-||\Delta x||^2/2\sigma_e^2)$
  - **Couples to 2nd moment** $\langle Q(x)A(x')\rangle$

Covariances:

- Functions $\Sigma_{Q,A}(x,x')$ over pairs of states and locations
- Deterministic evolution according to the Jacobian of the mean-field system, plus Poisson noise (similar form to Eq. 3; see (4) for details)

4. Bayesian filtering estimates latent neural-field states from spikes

Observation model: Spiking rate based on # active cells w. gain γ, bias β.

$$\lambda(x,t) = \gamma(x)A(x,t) + \beta(x)$$

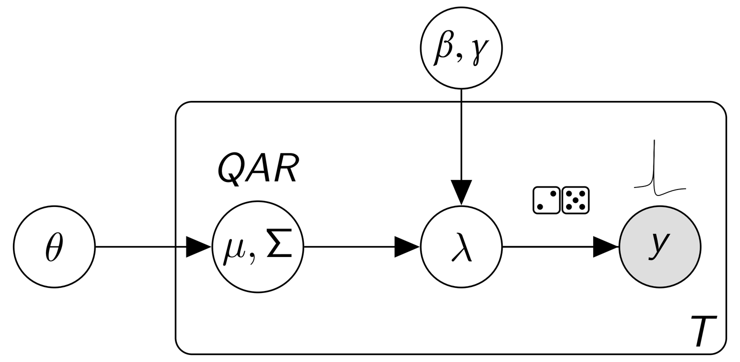

Figure 6. Hidden Markov model for neural fields. For all times $t \in T$, parameters θ dictate the evolution of a Gaussian-process model (μ, Σ) of neural fields. The observation model reflects how latent states impact spiking $y \sim \text{Poisson}(\lambda)$ observations.

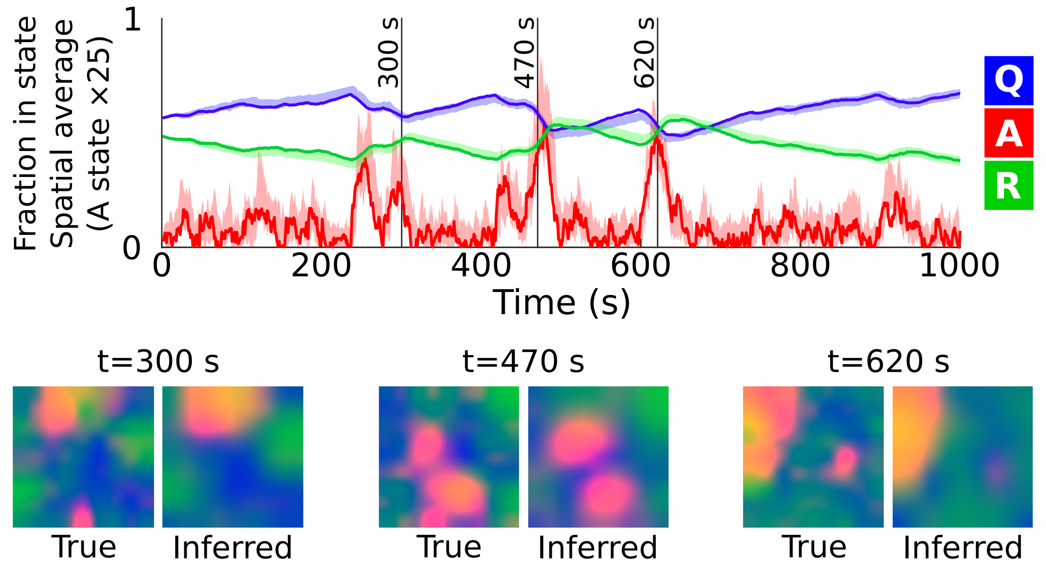

Figure 7. Bayesian filtering recovers latent states from spiking observations. (*top*) Average fraction of cells in each state over time. Solid lines: ground-truth; Shaded regions: 95% confidence interval (*A* state scaled by ×25). Colored plots (below) show the qualitative spatial organization is recovered. See (4) for details.

in summary

By modeling correlations in addition to mean-field dynamics, we construct a spatiotemporal Gaussian process model that evolves nonlinearly in time and can be used to infer latent neural-field states from spiking observations.

Applications:

» Integrate with spatiotemporal spiking datasets

» Future: Estimate parameters from data?

Funding provided by EPSRC EP/L027208/1 Large scale spatio-temporal point processes: novel machine learning methodologies and application to neural multi-electrode arrays.

[4] Rule ME, et al. Neural field models for latent state inference: Application to large-scale neuronal recordings. *bioRxiv*, 2019. doi: 10.1101/543769.

[5] Buice MA et al. Statistical mechanics of the neocortex. *Progress in Biophysics and Molecular Biology*, 99(2-3):53–86, feb 2009. ISSN 00796107. doi: 10.1016/j.pbiomolbio.2009.07.003.

[6] Hennig MH, et al. Early-stage waves in the retinal network emerge close to a critical state transition between local and global functional connectivity. *Journal of Neuroscience*, 29(4):1077–1086, 2009.