Inferring parameters of random networks from continuous-time trajectories

Alexander van Meegen[1,2], Tobias Kühn[1,2], Moritz Helias[1,2]

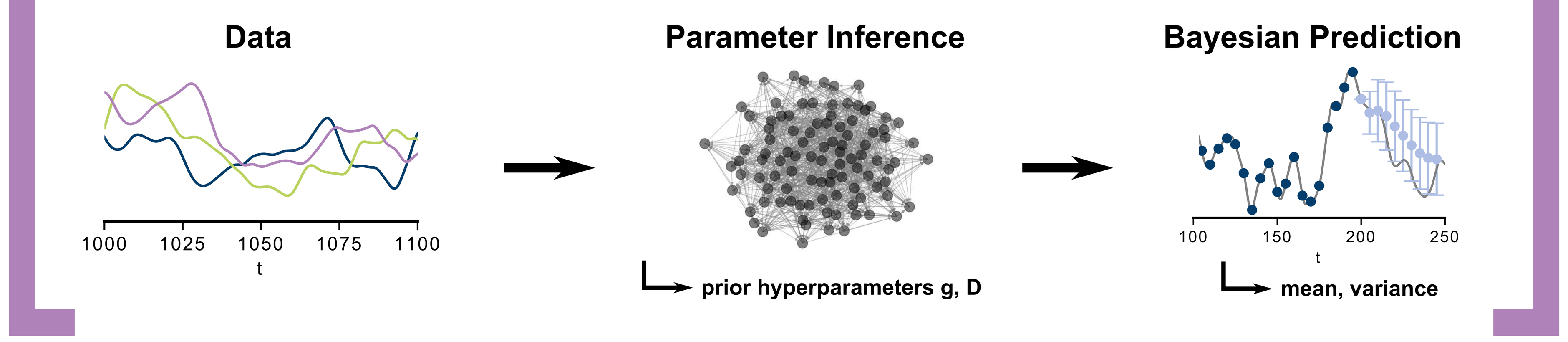
[1] Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
[2] Faculty 1, RWTH Aachen University, Aachen, Germany

Contact: `a.van.meegen@fz-juelich.de`

JÜLICH Forschungszentrum

JARA Jülich Aachen Research Alliance

In a nutshell:



Methods

Model

- random network of N nonlinearly interacting neurons:

$$\dot{x}_i(t) = -\nabla U(x_i(t)) + \frac{g}{\sqrt{N}} \sum_{j=1}^{N} J_{ij}\phi(x_j(t)) + \sqrt{2D}\xi_i(t)$$

- internal dynamics: overdamped motion in potential $U(x)$
- random network topology: independent J_{ij} distributed according to a standard normal distribution
- external input: independent Gaussian white noise processes $\xi_i(t)$ with zero mean and unit intensity
- for $\phi(x) = \tanh(x)$, $U(x) = \frac{1}{2}x^2$, $D = 0$, the model corresponds to the study by Sompolinsky, Crisanti, Sommers [1]

Large Deviation Theory

- long list of achievements since 1938 [2] but, with notable exceptions [3, 5, 6], rarely used in Neuroscience
- central notion: *rate function*

$$H(x) = -\lim_{N\to\infty} \frac{1}{N} \log P(x),$$

  i.e. the exponential part of the distribution $P(x)$
- canonical example [2]: xN heads in N tosses of a fair coin
  $\rightarrow P(x) = \binom{N}{xN} 2^{-N} = e^{-N[\log 2 + x\log x + (1-x)\log(1-x)] + o(N)}$
  $\rightarrow H(x) = \log 2 + x\log x + (1-x)\log(1-x)$

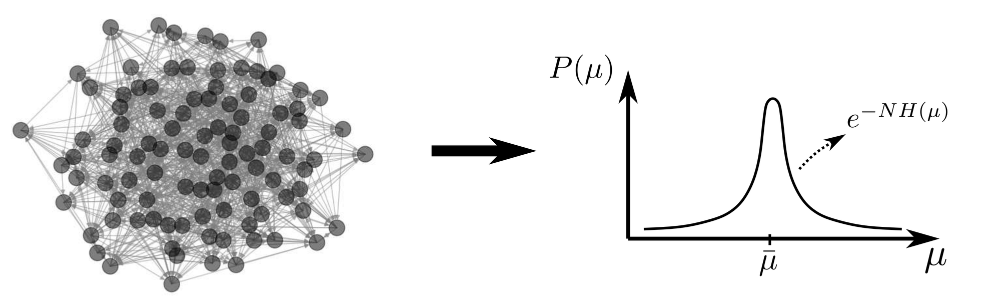
Rate Function of Empirical Measure

- we obtain the rate function of the empirical measure $\mu[x] := \frac{1}{N}\sum_{i=1}^{N} \delta[x_i - x]$:

$$H(\mu) = \int \mathcal{D}x\, \mu[x] \log \frac{\mu[x]}{\langle \delta[\dot{x} + \nabla U(x) - \eta]\rangle_\eta}, \quad \text{where}$$

$$C_\eta(t_1, t_2) = 2D\delta(t_1 - t_2) + g^2 \int \mathcal{D}x\, \mu[x] \phi(x(t_1))\phi(x(t_2)),$$

  which generalizes the rigorous result of Arous & Guionnet [3]



Parameter Inference

- data (rate profiles $x_i(t)$) is described by the corresponding empirical measure

$$\mu[x] = \frac{1}{N}\sum_{i=1}^{N} \delta[x_i - x]$$

- likelihood of the empirical measure given the (hyper-)parameters is

$$\log P(\mu|\, D, g) \approx -NH(\mu)$$

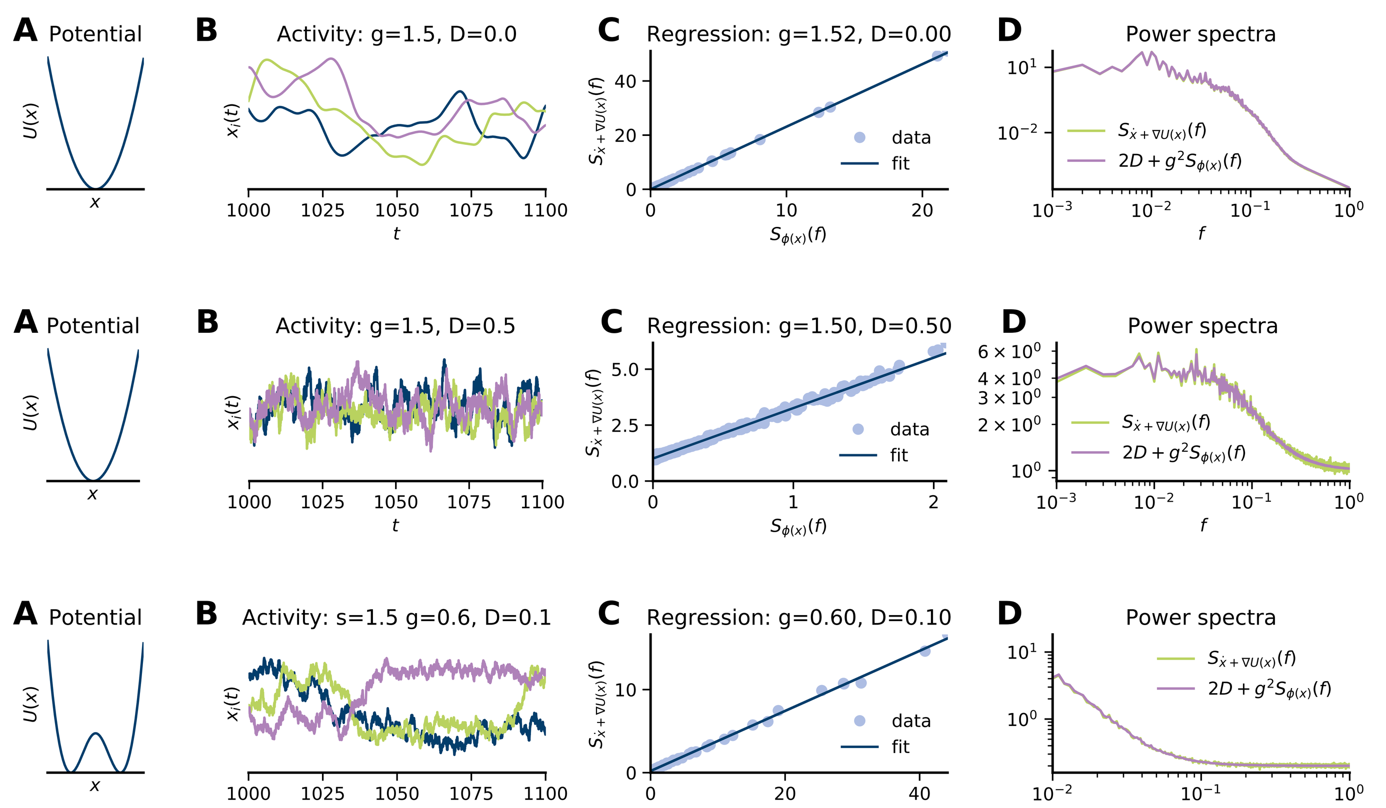
- for maximum likelihood parameter estimation, we need

$$\partial_{D,g} \log P(\mu|\, D, g) = 0$$

- this leads to the simple condition

$$S_{\dot{x}+\nabla U(x)}(f) = 2D + g^2 S_{\phi(x)}(f)$$

- g and D follow from a (non-negative) linear regression of this linear relation
- in figure: $U(x) = \frac{1}{2}x^2 - s\log(\cosh(x))$ with $s = 0$ unless specified and $\phi(x) = \tanh(x)$



Maximum likelihood parameter estimation for different potentials $U(x)$ and different strength of the external noise (**A**). Activity of three randomly chosen units (**B**), parameter estimation via non-negative linear regression (**C**), and the power spectra corresponding to the inferred parameters (**D**).

Activity Prediction

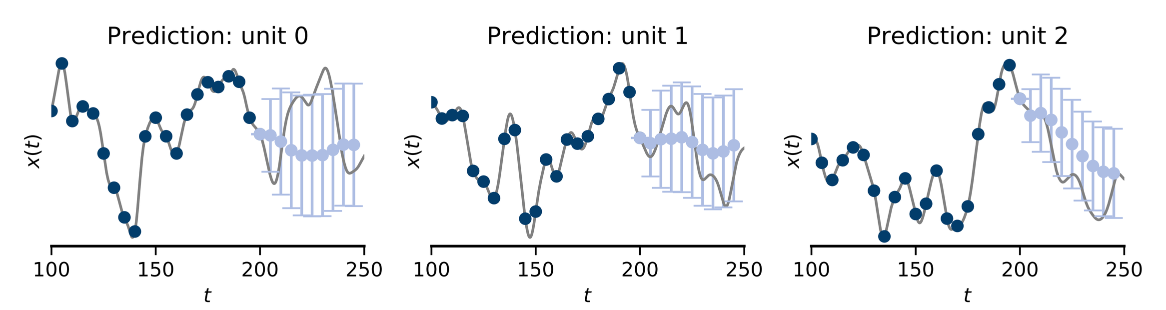
- for $U(x) = \frac{1}{2}x^2$, the most probable measure corresponds to a Gaussian Process with self-consistent statistics

$$x \sim \mathrm{GP}(0, C_x)$$

- for a Gaussian Process, missing datapoints $\hat{x} = x(\hat{t})$ and their variability $\sigma_{\hat{x}}$ can be predicted using [4]

$$\hat{x} = \mathbf{k}^{\mathrm{T}} \boldsymbol{K}^{-1}\mathbf{x},$$
$$\sigma_{\hat{x}}^2 = \kappa - \mathbf{k}^{\mathrm{T}} \boldsymbol{K}^{-1}\mathbf{k},$$

  where $K_{ij} = C_x(t_i, t_j)$, $k_i = C_x(t_i, \hat{t})$, and $\kappa = C_x(\hat{t}, \hat{t})$

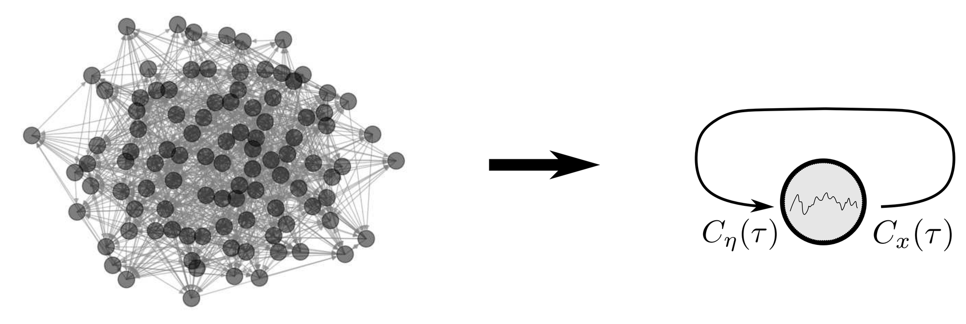


Predicting future activity of a trajectory (gray) for $g = 1.5$, $D = 0$, $s = 0$. Training data (dark blue) determines $\mathbf{x}$ and thus the prediction $\hat{x}$ (light blue). The parameters of the correlation function C_x of the Gaussian Process can be inferred very efficiently using the method outlined above; the remaining numerical effort is small because $\boldsymbol{K}$ is small.

Mean-Field Theory

Statistical Field Theory

- developed in spin glass theory [7], strongly based on functional integrals (path integrals) which are rarely treated rigorously
- considerable recent interest [8, 9, 10, 11, 12], comprehensive introduction in the *new book by Helias & Dahmen* [13]



Equivalence to Large Deviation Theory

- the rate function of the empirical measure takes the form of a Kullback-Leibler divergence:

$$H(\mu) = D_{\mathrm{KL}}(\mu[x] \,\|\, \langle \delta[\dot{x} + \nabla U(x) - \eta]\rangle_\eta)$$

- its unique minimum $\bar{\mu}$ is at

$$\bar{\mu}[x] = \langle \delta[\dot{x} + \nabla U(x) - \eta]\rangle_\eta, \quad \text{where}$$

$$C_\eta(t_1, t_2) = 2D\delta(t_1 - t_2) + g^2 \int \mathcal{D}x\, \bar{\mu}[x]\phi(x(t_1))\phi(x(t_2))$$

- this corresponds to the self-consistent stochastic dynamics

$$\dot{x}(t) = -\nabla U(x(t)) + \eta(t)$$

  obtained in [1] and extended to $D \neq 0$ in [10]

Outlook

- have shown for a highly nontrivial system that the path integral approach leads to mathematically sound results
- intensify / facilitate further dialogue between physics and mathematics communities
- obtain subexponential corrections to rate function, e.g. using a loop expansion, to account for finite size of networks and data
- explore the relation of Gaussian Processes and Artificial Neural Networks [14] further
- justify Gaussian Process prediction with replica calculation

References

[1] Sompolinsky H, Crisanti A, Sommers H (1988) *Phys Rev Lett* 61: 259.
[2] Varadhan SRS (2008) *Ann. Probab.* 36 (2): 397-419.
[3] Arous GB, Guionnet A (1995) *Probabil. Theory Related Fields* 102: 455.
[4] David JC MacKay. Information Theory, Inference, and Learning Algorithms. *Cambridge University Press*, 2003.
[5] Faugeras O, Maclaurin J (2014) *Comptes Rendus Mathematique* 352 (10): 841-846.
[6] Faugeras O, Maclaurin J, Tanré E (2019) *arXiv* 1901.10248.
[7] Sompolinsky H, Zippelius A (1982) *Phys Rev B* 25 (11): 6860-6875.
[8] Mastrogiuseppe F, Ostojic S (2018) *Neuron* 99 (3): 609-623.
[9] Dahmen D, Grün S, Diesmann M, Helias M (2019) *PNAS* 116 (26): 13051-13060.
[10] Schuecker J, Goedeke S, Helias M (2018) *Phys Rev X* 8 (4): 041029.
[11] van Meegen A, Lindner B (2018) *Phys Rev Lett* 121: 258302.
[12] Muscinelli SP, Gerstner W, Schwalger T (2019) *PLoS CB* 15 (6): e1007122.
[13] Helias M, Dahmen D. Statistical field theory for neural networks. *Springer Lecture Notes in Physics* (in press), 2020. arXiv:1901.10416.
[14] Williams CKI (1998) *Neural Comput* 10: 1203-1216.

Acknowledgments: This work was partly supported by the Helmholtz young investigator's group VH-NG-1028 and European Union Horizon 2020 grant 785907 (Human Brain Project SGA2).