Machine-learning continuously-monitored quantum systems

parameter estimation and beyond

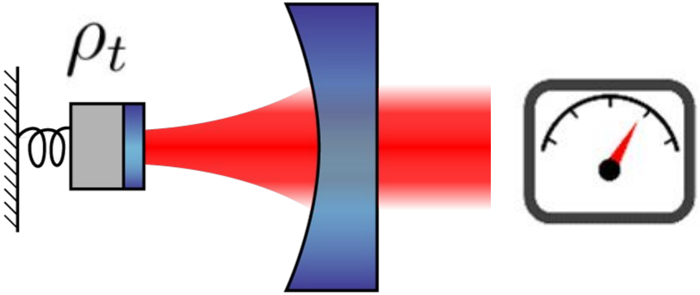
Matías Bilkis

Computer Vision Center

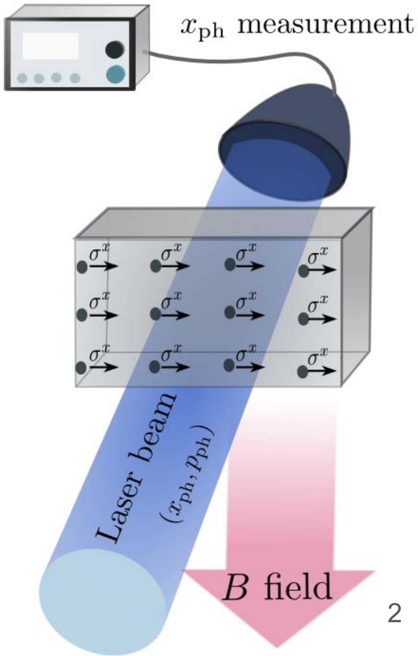
Joint work w/
John Calsamiglia
Giulio Gasbarri
Elisabet Roda

CMONqs

- **CMONqs = Continuously-monitored quantum systems** [Wiseman2009Book]



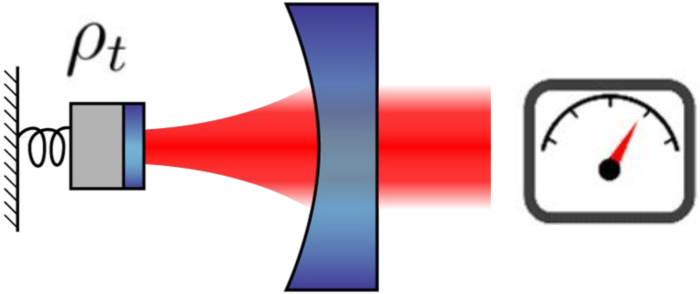
$$d\rho_t = dt\mathcal{L}_\theta[\rho_t]dt + \mathcal{H}[\eta \mathbf{c}]\rho_t \cdot \; d\mathbf{W}_t$$

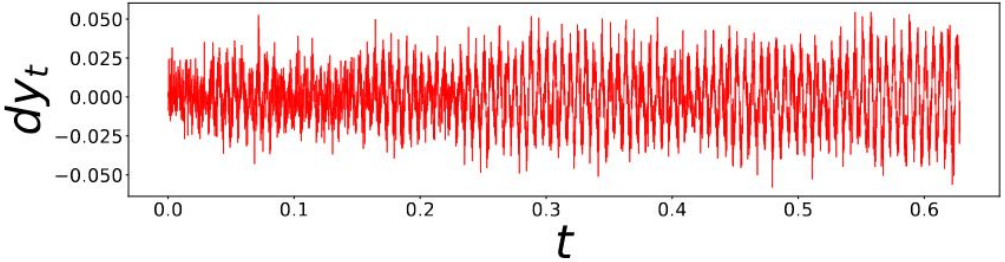


- **Physical systems:**
    - atomic sensors [Martínez2018Signal]
    - optomechanical cavities [Aspelmayer2010Cavity]

CMONqs

- **Measurement outcomes → noisy signal**

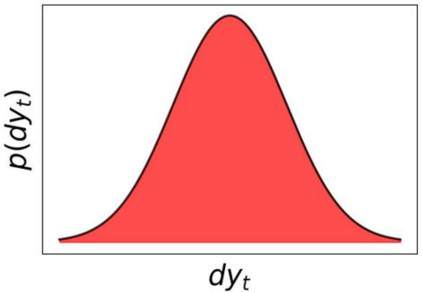


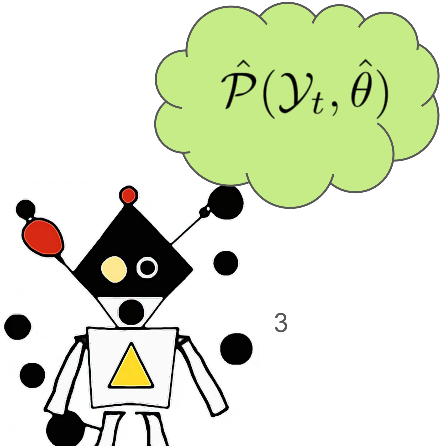


$$d\rho_t = dt\mathcal{L}_\theta[\rho_t]dt + \mathcal{H}[\eta\mathbf{c}]\rho_t \cdot \ d\mathbf{W}_t$$

- **Measurement model**

$$d\mathbf{W}_t = d\mathbf{y}_t - \mathbf{Tr}[\mathcal{H}_{\eta\mathbf{c}}[\rho_t]]dt$$

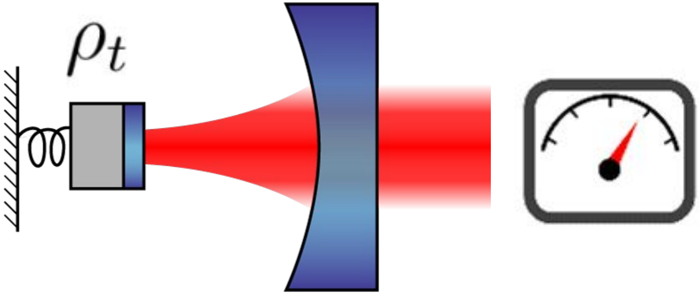


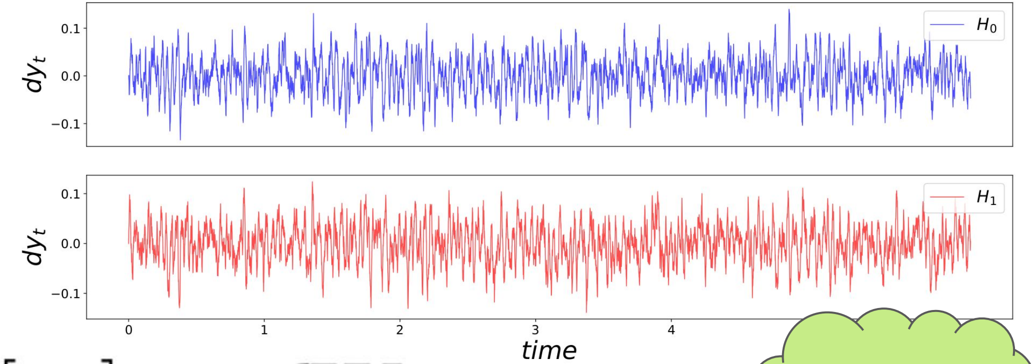


Statistical inference on CMONqs

(2307.14954)

- **Hypothesis testing quantum dynamics** $\rightarrow$ [Gasbarri2023Sequential]

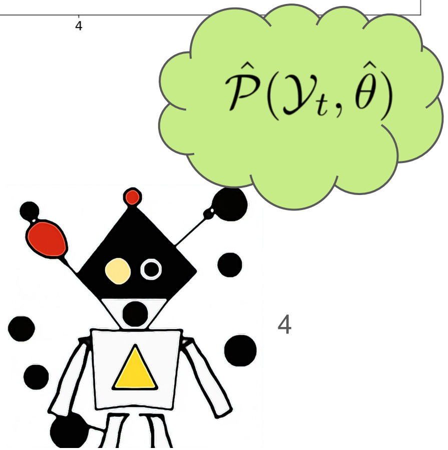




$$d\rho_t = dt\mathcal{L}_\theta[\rho_t]dt + \mathcal{H}[\eta \mathbf{c}]\rho_t \cdot \; d\mathbf{W}_t$$

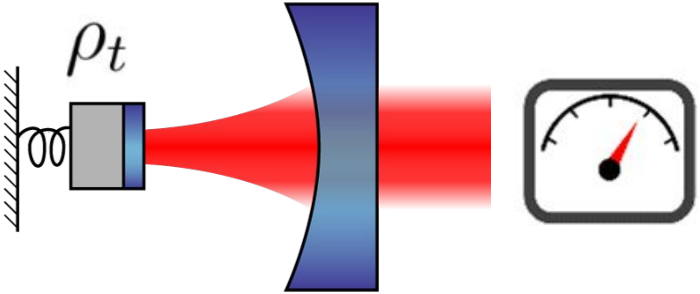
- **Measurement model**

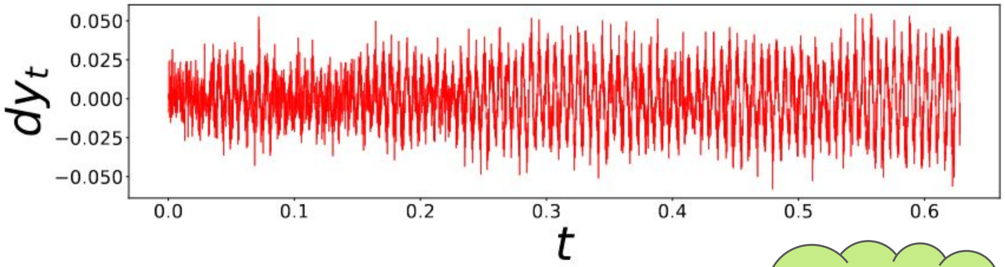
$$d\mathbf{W}_t = d\mathbf{y}_t - \mathbf{Tr}[\mathcal{H}_{\eta\mathbf{c}}[\rho_t]]dt$$



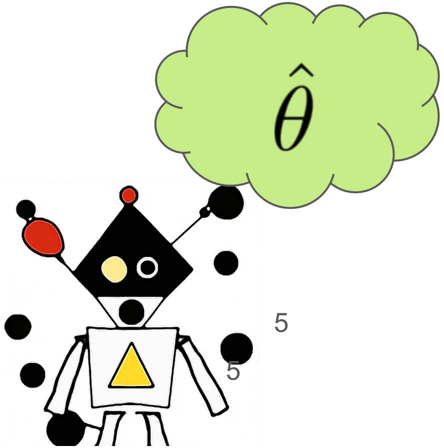
Statistical inference on QMONs

- **Today → parameter estimation** [ArXiv coming soon]





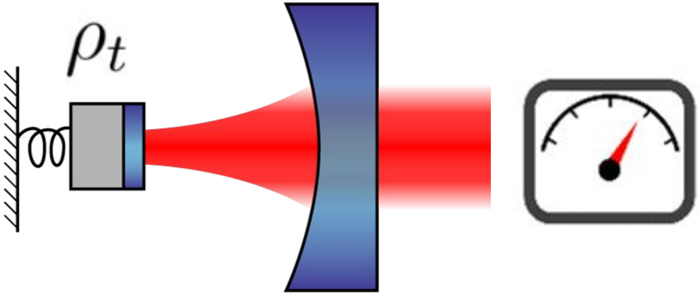
$$d\rho_t = dt\mathcal{L}_\theta[\rho_t]dt + \mathcal{H}[\eta\mathbf{c}]\rho_t \cdot \; d\mathbf{W}_t$$

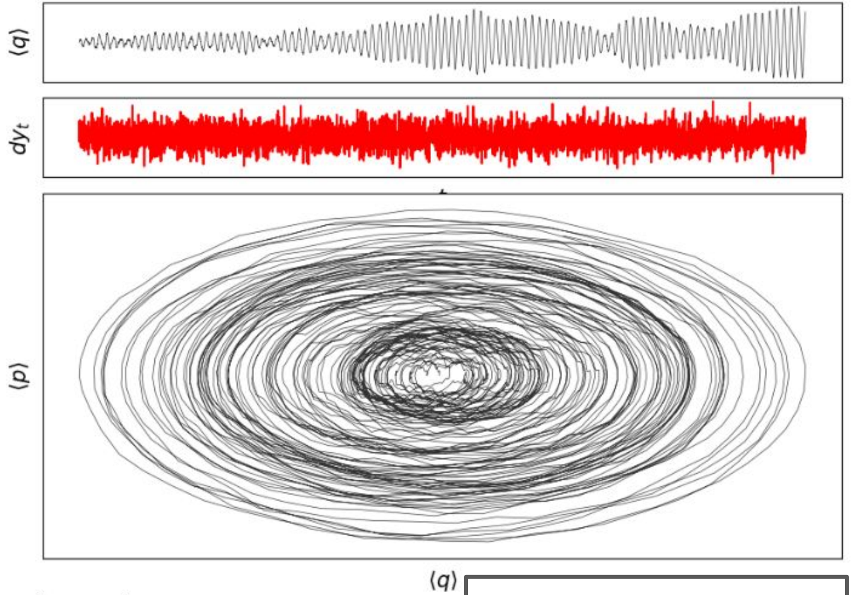


5

Extra assumption: Gaussianity

- **Mechanical-mode dynamics: damped harmonic oscillator +**





- **Linear SDE** (≡ Kalman filter)

$$d\bar{\boldsymbol{r}}_t = A\bar{\boldsymbol{r}}_t dt + \chi(\boldsymbol{\Sigma}_t) d\mathbf{W}_t$$

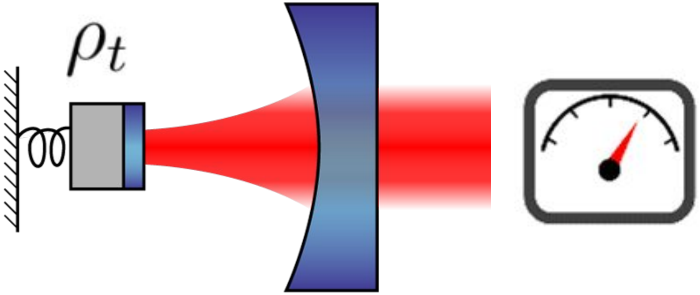
$$d\boldsymbol{\Sigma}_t = A\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_t A^T + D - \chi(\boldsymbol{\Sigma}_t)^T \chi(\boldsymbol{\Sigma}_t)$$

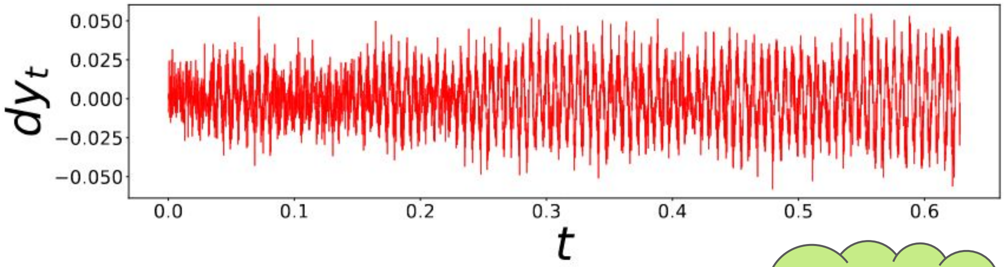
$$d\mathbf{y}_t = C\bar{\mathbf{r}}_t dt + d\mathbf{W}_t$$

$$A = \begin{pmatrix} -\gamma & \omega \\ -\omega & -\gamma \end{pmatrix}$$

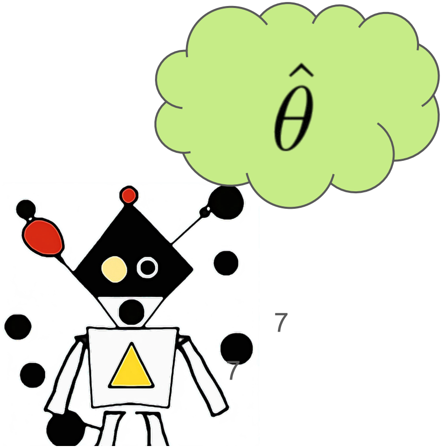
Statistical inference on QMONs

- **Today → parameter estimation** [ArXiv coming soon]





$$d\bar{\boldsymbol{r}}_t = A\bar{\boldsymbol{r}}_t dt + \chi(\boldsymbol{\Sigma}_t) d\mathbf{W}_t \qquad d\mathbf{y}_t = C\bar{\mathbf{r}}_t dt + d\mathbf{W}_t$$



- **Estimators**

$$A = \begin{pmatrix} -\gamma & \omega \\ -\omega & -\gamma \end{pmatrix}$$

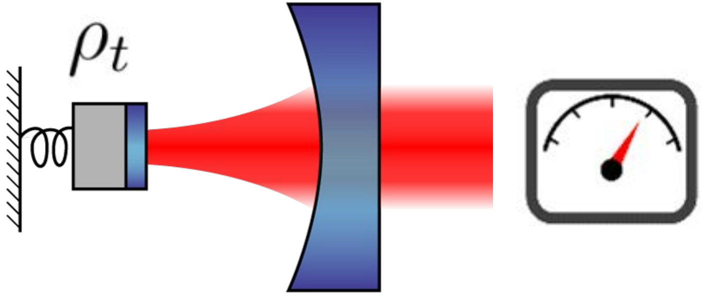
$$P(\mathcal{Y}_t|\theta) = e^{\lambda(\mathcal{Y}_t|\theta)} P_W(\mathcal{Y}_t)$$

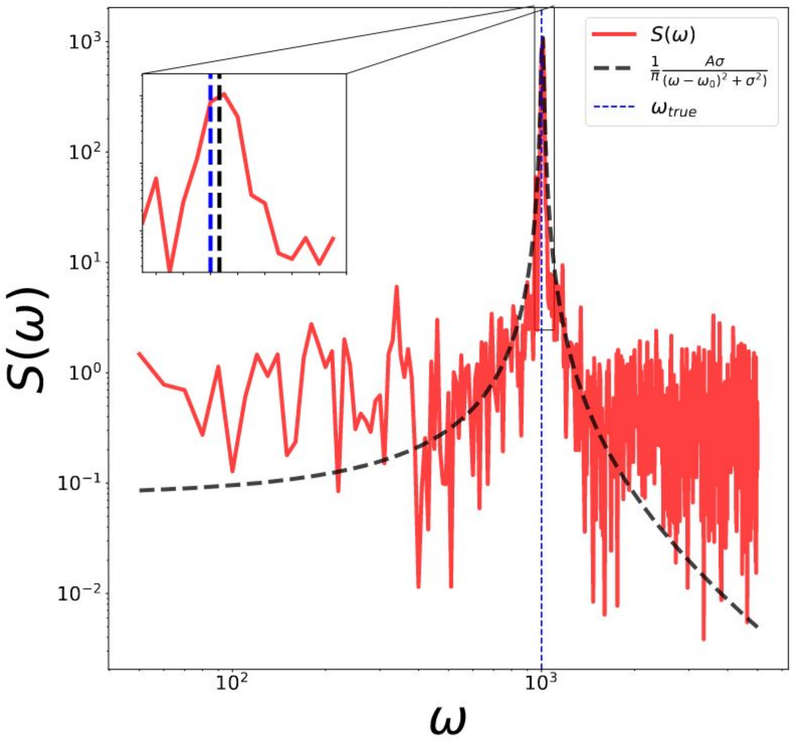
$$\mathrm{Var}[\hat{\theta}] \geq \frac{1}{N I(\theta)}$$

$$I_t(\theta) = \langle (\partial_\theta \lambda(\mathcal{Y}_t;\theta)^2 \rangle = \langle -\partial_\theta^2 \lambda(\mathcal{Y}_t,\theta) \rangle$$

Frequency estimation: Lorentzian

- Experimentally → Lorenztian fit on power spectrum





- Many quantum trajs

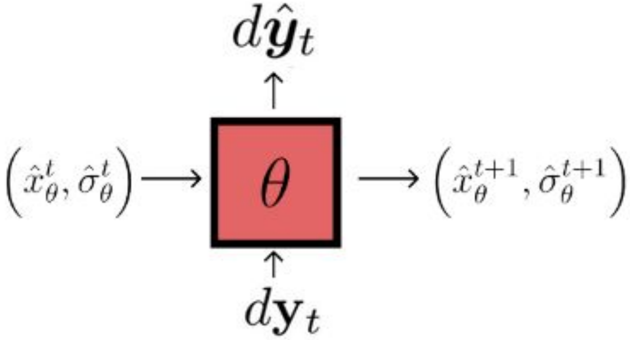
→ estimator's variance

Maximum likelihood & RNN

- **Predict next measurement outcome**

$$d\hat{\boldsymbol{y}}_t = C\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}} dt$$

$$d\mathbf{y}_t = C\bar{\mathbf{r}}_t dt + d\mathbf{W}_t$$

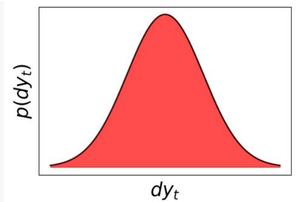


- **Hidden state propagation:**

$$d\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}} = [A_{\hat{\omega}} - \chi(\boldsymbol{\Sigma_t})C]\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}} dt + \chi(\boldsymbol{\Sigma_t}) d\boldsymbol{y}_t$$

$$d\bar{\boldsymbol{r}}_t = A\bar{\boldsymbol{r}}_t dt + \chi(\boldsymbol{\Sigma}_t) d\mathbf{W}_t$$

- **Cost-function = Log-likelihood** **Training = Maximum Likelihood**



$$\mathrm{C}(\hat{\omega}) = \sum_t |d\boldsymbol{y}_t - d\hat{\boldsymbol{y}}_t|^2 = -\log p(\boldsymbol{y}_t|\hat{\omega})$$

Frequency estimation: RNN

- **Predict next measurement outcome**

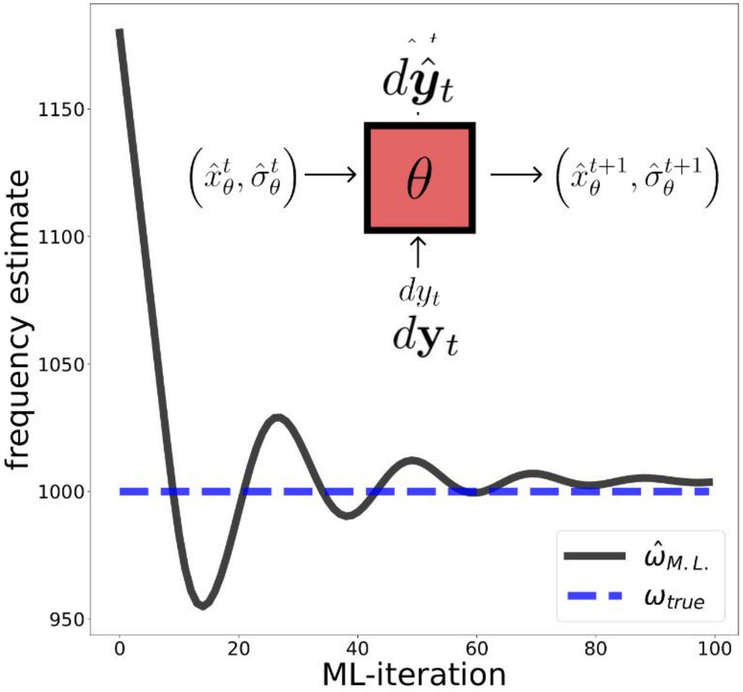
$$d\hat{\boldsymbol{y}}_t = C\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}}\, dt$$

$$d\mathbf{y}_t = C\bar{\mathbf{r}}_t\, dt + d\mathbf{W}_t$$

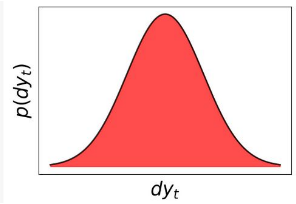
- **Hidden state propagation:**

$$d\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}} = [A_{\hat{\omega}} - \chi(\boldsymbol{\Sigma_t})C]\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}}\, dt + \chi(\boldsymbol{\Sigma_t})\, d\boldsymbol{y}_t$$

$$d\bar{\boldsymbol{r}}_t = A\bar{\boldsymbol{r}}_t\, dt + \chi(\boldsymbol{\Sigma}_t)\, d\mathbf{W}_t$$



- **Cost-function = Log-likelihood**

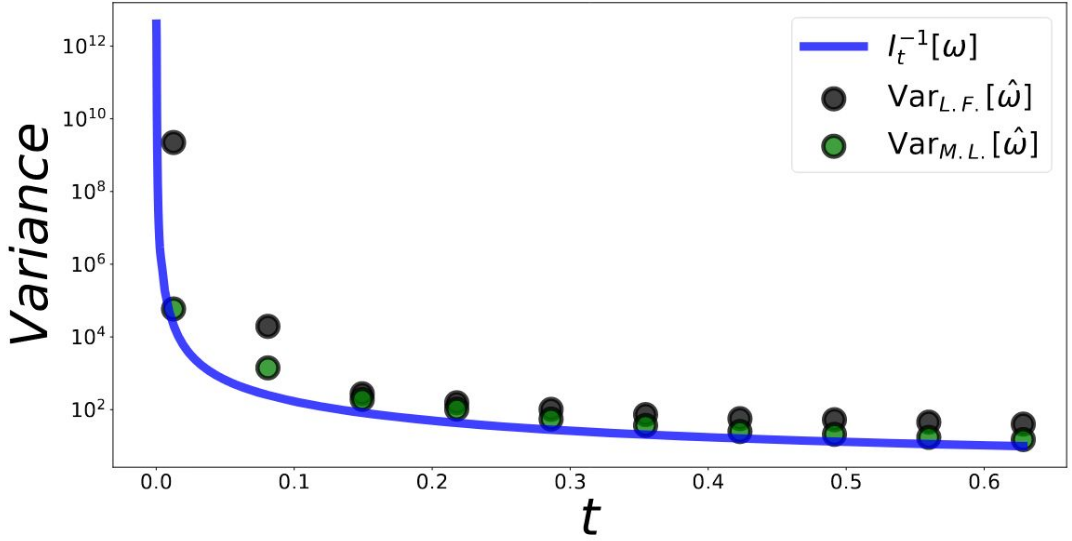


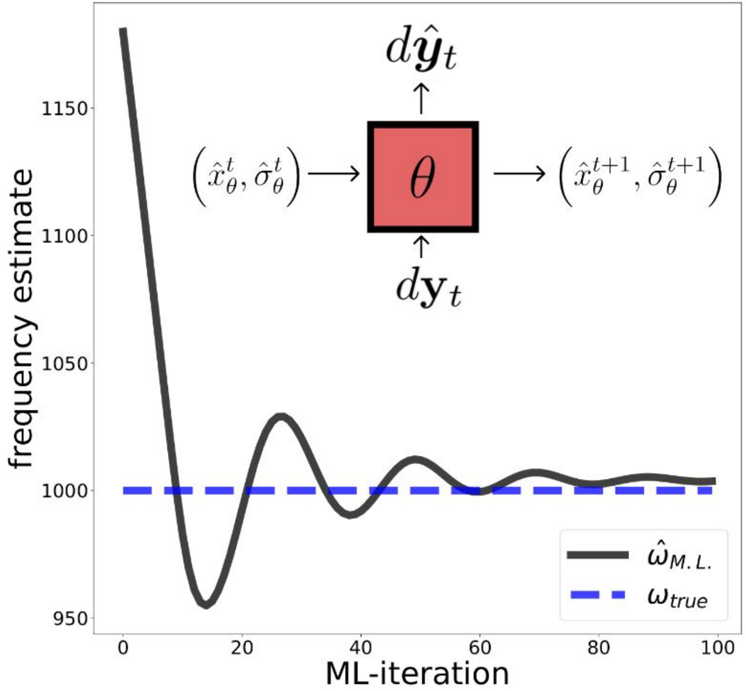
$$\mathrm{C}(\hat{\omega}) = \sum_t |d\boldsymbol{y}_t - d\hat{\boldsymbol{y}}_t|^2 = -\log p(\boldsymbol{y}_t|\hat{\omega})$$

Recap: frequency estimation

- **Recurrent module** (≡ Kalman filter)
- **Trained via maximum likelihood**

$$\hat{\theta} = \underset{\theta}{\text{ArgMax}}\ \lambda(\mathcal{Y}_t, \theta)$$

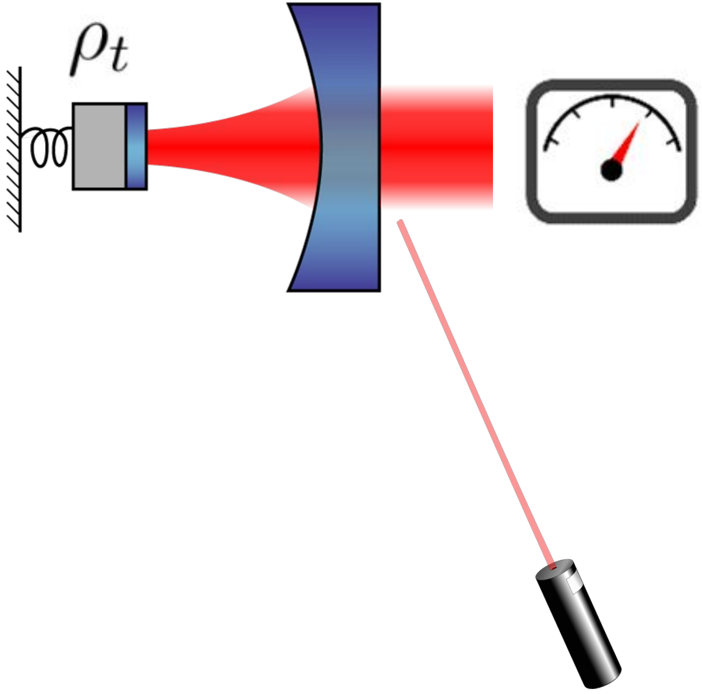


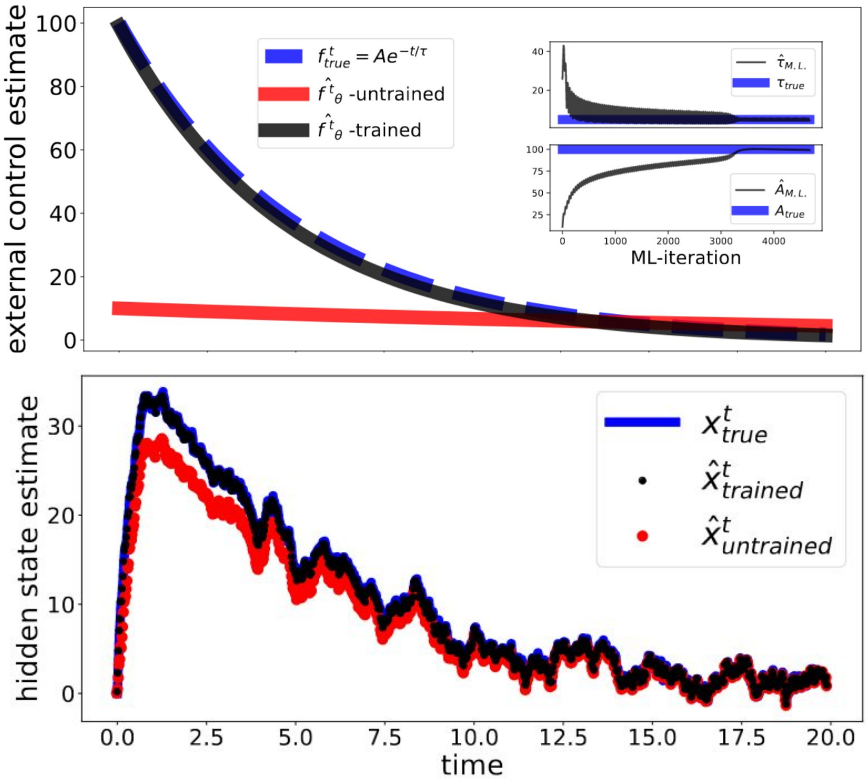


External-signals: force estimation

$$d\bar{\mathbf{r}}_t = \Big(A - \chi(\Sigma)\Big)\bar{\mathbf{r}}_t dt + \chi(\Sigma) d\mathbf{y}_t + f^t_\theta dt$$

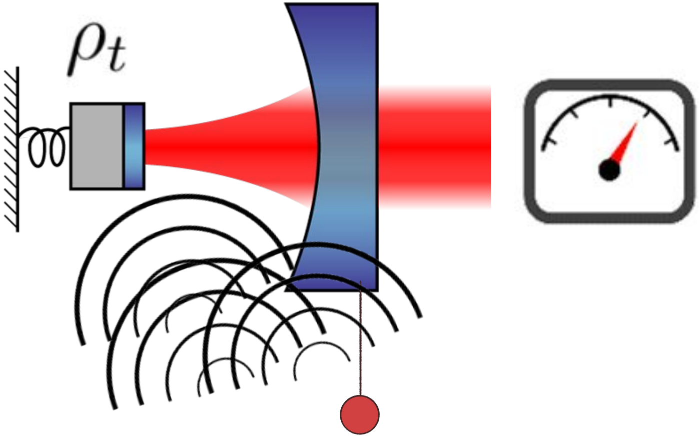
$$f(t) = Ae^{-t/\tau}$$

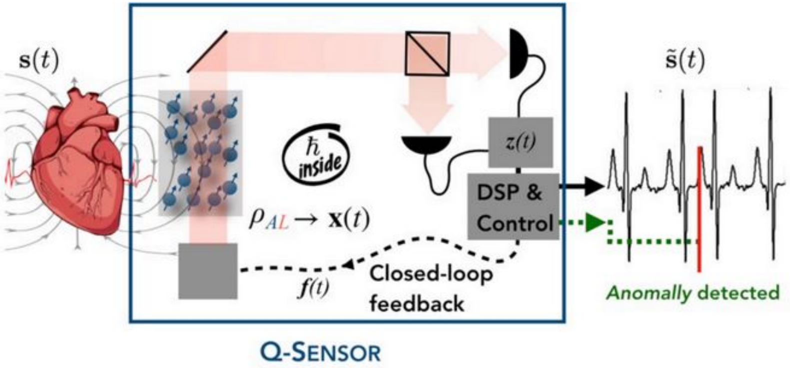




What now? Machine-learning dynamics

- Infer a <u>dynamical equation</u> for f(t) out of measurement record



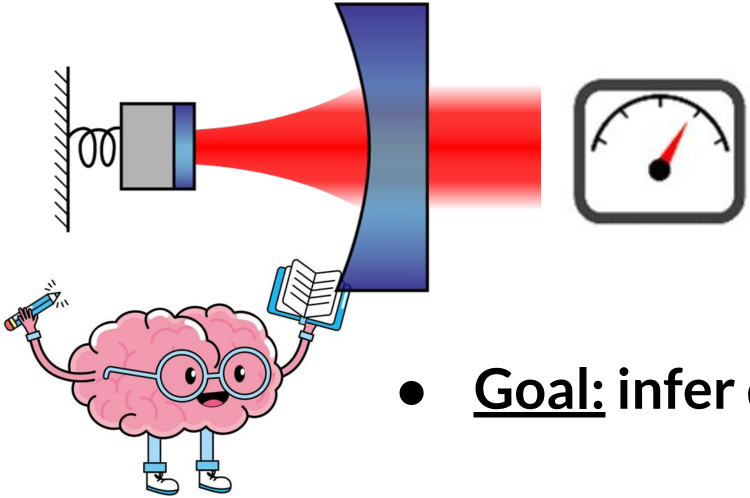


- <u>Quantum sensor:</u> measurement signal + function dictionary +Max.Lik.

→ infer signal's dynamics df = g(f,t) dt

External-signals: force estimation

- We consider FitzHugh–Nagumo model (models a neuron's activity)

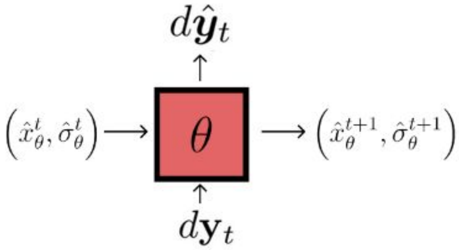


External force dynamics

$$\dot{v} = v - \frac{v^3}{3} - w + RI_{\text{ext}}$$

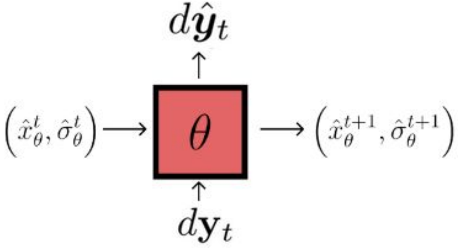
$$\tau\dot{w} = v + a - bw.$$

- **Goal:** infer dynamics from measurement outcomes only!



Frequency estimation: RNN

- **Predict next measurement outcome**



$$d\hat{\boldsymbol{y}}_t = C\hat{\bar{\boldsymbol{r}}}^t_{\hat{\omega}} dt$$

$$d\mathbf{y}_t = C\bar{\mathbf{r}}_t dt + d\mathbf{W}_t$$

- **Hidden state propagation:**

$$d\hat{\bar{\boldsymbol{r}}}^t_{\hat{v}_t} = [A - \chi(\boldsymbol{\Sigma_t})C]\hat{\bar{\boldsymbol{r}}}^t_{\hat{v}_t} dt + \chi(\boldsymbol{\Sigma_t}) d\boldsymbol{y}_t + \begin{pmatrix} 0 \\ \hat{v}_t \end{pmatrix} dt$$

$$d\bar{\boldsymbol{r}}_t = [A - \chi(\boldsymbol{\Sigma_t})C]\bar{\boldsymbol{r}}_t dt + \chi(\boldsymbol{\Sigma_t}) d\boldsymbol{y}_t + \begin{pmatrix} 0 \\ v_t \end{pmatrix} dt$$

- **Inferred dynamics**

$$\dot{\hat{v}} = \xi_0 g_0(\hat{v}_t, \hat{w}_t) + \xi_1 g_1(\hat{v}_t, \hat{w}_t) + \ldots + \xi_n g_n(\hat{v}_t, \hat{w}_t)$$

$$\dot{\hat{w}} = \xi_{n+1} g_0(\hat{v}_t, \hat{w}_t) + \quad (\ldots) \quad + \xi_N g_N(\hat{v}_t, \hat{w}_t)$$

COST

$$\mathrm{C}(\vec{\xi}) = \sum_t |d\boldsymbol{y}_t - d\hat{\boldsymbol{y}}_t|^2 + \lambda ||\vec{\xi}||_1$$

FHN-model: before training

$$\hat{\dot{v}} = \xi_0 g_0(\hat{v}_t, \hat{w}_t) + \xi_1 g_1(\hat{v}_t, \hat{w}_t) + \ldots + \xi_n g_n(\hat{v}_t, \hat{w}_t)$$

$$\hat{\dot{w}} = \xi_{n+1} g_0(\hat{v}_t, \hat{w}_t) + \quad (\ldots) \quad + \xi_N g_N(\hat{v}_t, \hat{w}_t)$$

```
### PARAMETERS IN THE SIMULATION
[params_force[0], K0, K1, K3]

[[0.8, 1.0],
 [0.9893689920528697, 0.11615082966737726],
 array([[ 1.        , -1.        ],
        [ 0.07341903, -0.0734248 ]]),
 array([[-0.33333333,  0.        ],
        [ 0.        ,  0.        ]])]
```

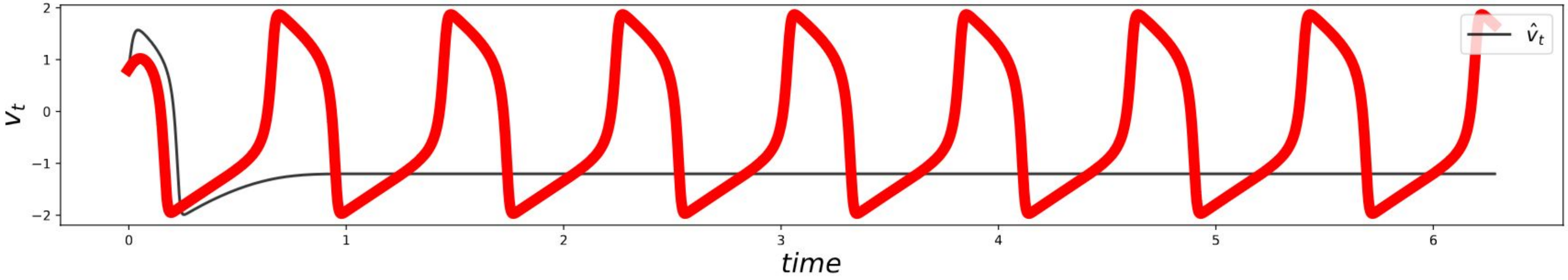
```
## INITIAL SEED MAXIMUM LIKELIHOOD
history["params"][0]

[tensor([0.9365, 1.1401]),
 tensor([0.9984, 0.1045]),
 tensor([[ 1.0007, -0.9887],
         [ 0.0810, -0.0856]]),
 tensor([[-0.3312,  0.0009],
         [ 0.0030,  0.0053]])]
```

$\vec{\xi}^{(0)}/\vec{\xi}^*$

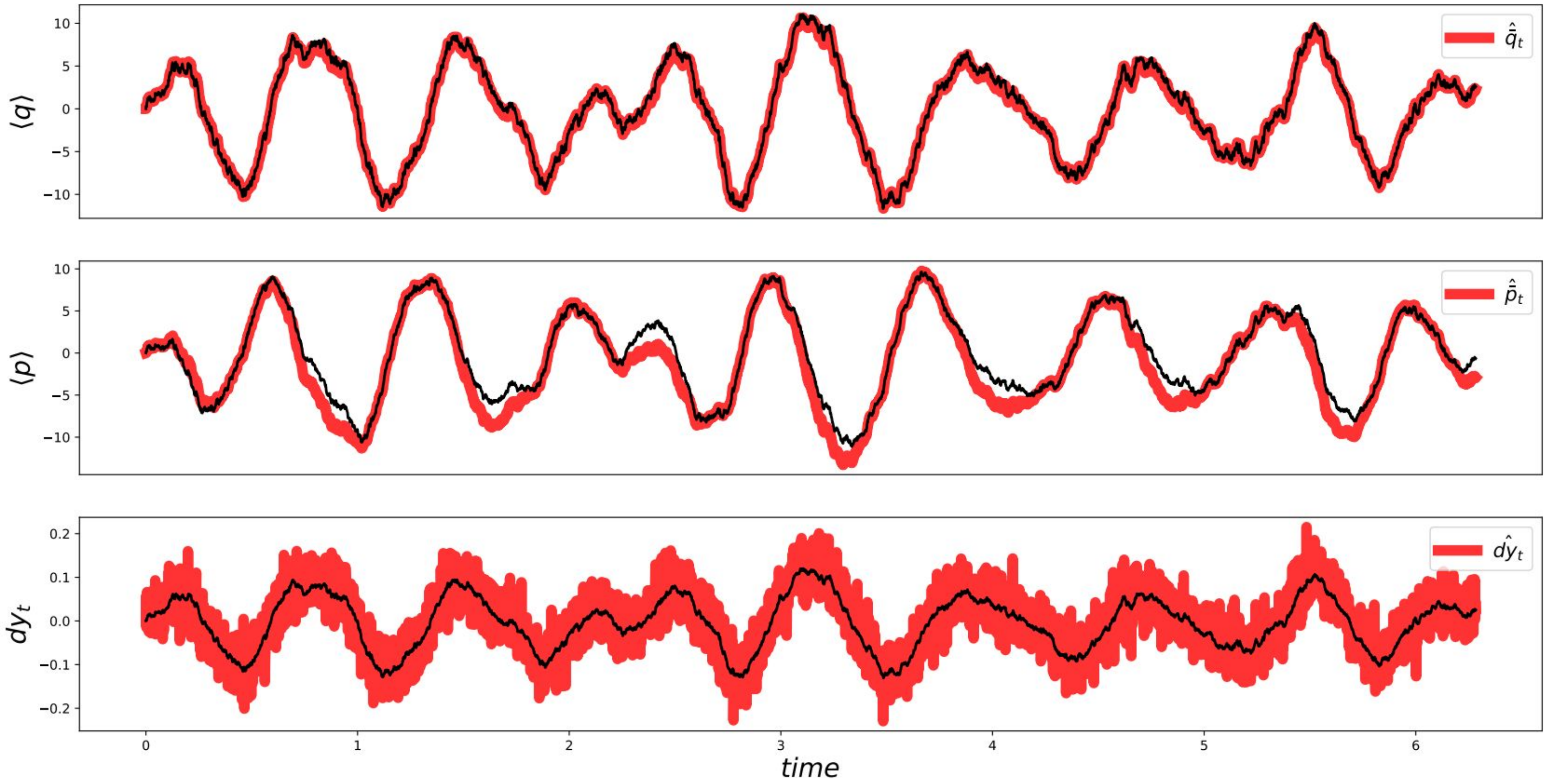
"discover" this out of data:

$$\dot{v} = v - \frac{v^3}{3} - w + RI_{\text{ext}}$$

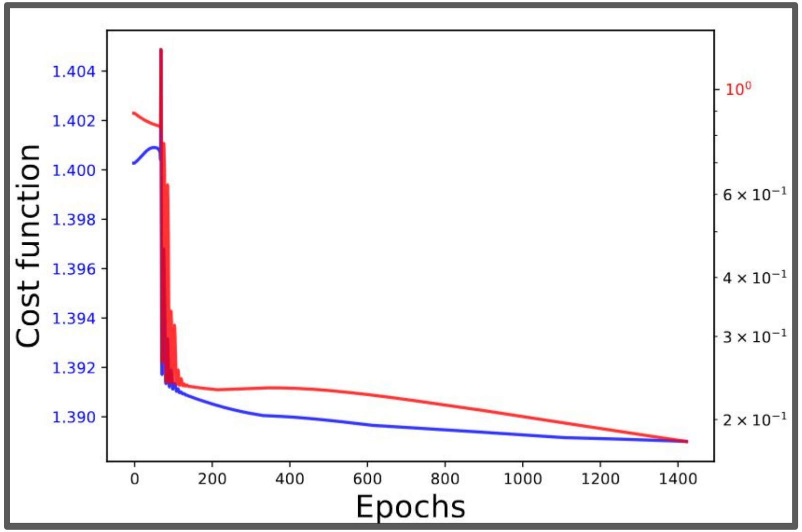
$$\tau\dot{w} = v + a - bw.$$



FHN-model: before training



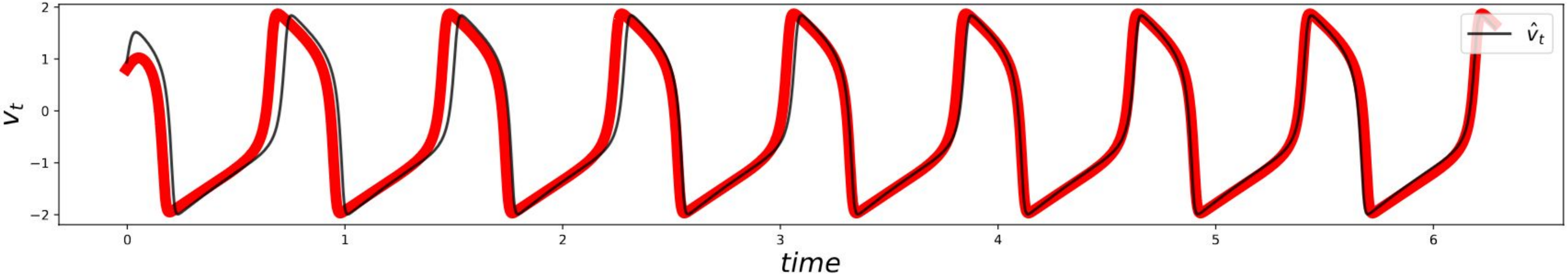
FHN-model: after training




```
### PARAMETERS IN THE SIMULATION
[params_force[0], K0, K1, K3]

[[0.8, 1.0],
 [0.9893689920528697, 0.11615082966737726],
 array([[ 1.        , -1.        ],
        [ 0.07341903, -0.0734248 ]]),
 array([[-0.33333333,  0.        ],
        [ 0.        ,  0.        ]])]
```

```
## BEST RESULTS MAXIMUM LIKELIHOOD
history["params"][best_ind]

[tensor([0.8353, 1.0331]),
 tensor([0.9973, 0.1081]),
 tensor([[ 0.9942, -0.9953],
         [ 0.0804, -0.0817]]),
 tensor([[-3.3775e-01, -2.6911e-03],
         [ 5.7242e-03, -3.1335e-04]])]
```

$\vec{\xi}^{(f)}/\vec{\xi}^*$

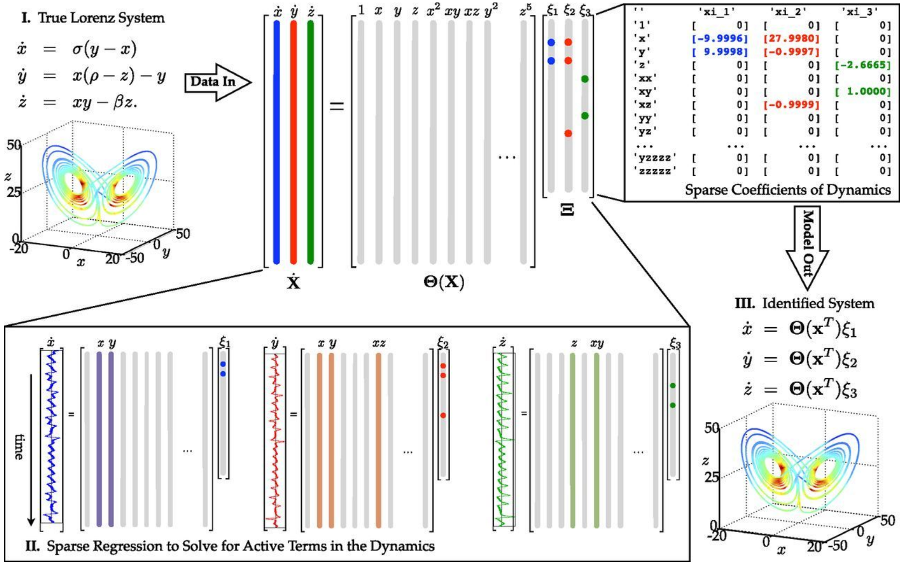


Conclusion

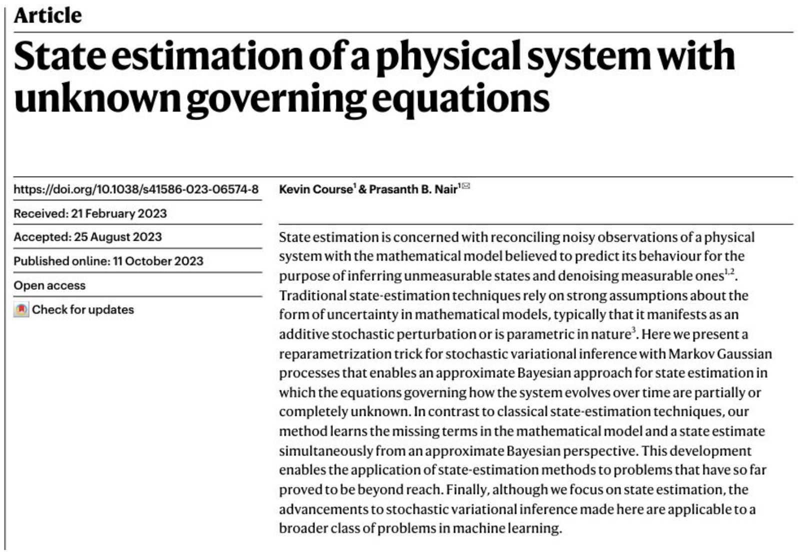
- **We study continuous-time quantum sensors**
- **What can we do with a <u>single trajectory</u>?**
- **Parameter estimation**
  Recurrent cell → maximum likelihood
- **External signals equation of motion discovery**

Related works

[Brunton 2015 Discovering]



[Kevin2023State]



Article

State estimation of a physical system with unknown governing equations

https://doi.org/10.1038/s41586-023-06574-8
Received: 21 February 2023
Accepted: 25 August 2023
Published online: 11 October 2023
Open access
Check for updates

Kevin Course[1] & Prasanth B. Nair[1]✉

State estimation is concerned with reconciling noisy observations of a physical system with the mathematical model believed to predict its behaviour for the purpose of inferring unmeasurable states and denoising measurable ones[1,2]. Traditional state-estimation techniques rely on strong assumptions about the form of uncertainty in mathematical models, typically that it manifests as an additive stochastic perturbation or is parametric in nature[3]. Here we present a reparametrization trick for stochastic variational inference with Markov Gaussian processes that enables an approximate Bayesian approach for state estimation in which the equations governing how the system evolves over time are partially or completely unknown. In contrast to classical state-estimation techniques, our method learns the missing terms in the mathematical model and a state estimate simultaneously from an approximate Bayesian perspective. This development enables the application of state-estimation methods to problems that have so far proved to be beyond reach. Finally, although we focus on state estimation, the advancements to stochastic variational inference made here are applicable to a broader class of problems in machine learning.

See also:
[Flurin2020Using]
[Choi2022Learning]
[Chen2018Neural]