Lab-STICC

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

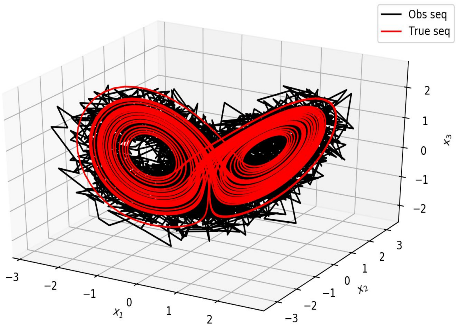
Learning Chaotic and Stochastic Dynamics from Noisy and Partial Observation using Variational Deep Learning

Duong Nguyen, Said Ouala, Noura Dridi, Lucas Drumetz, and Ronan Fablet

IMT Atlantique, Lab-STICC, Brest, France

Melody meeting, Oct 13, 2020

1. MOTIVATION

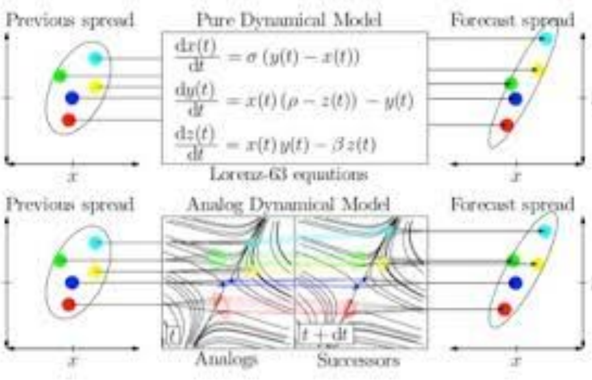


Model

$$\frac{\mathrm{d}\mathbf{x}_t}{dt} = f(\mathbf{x}_t) \quad \text{(ODE)}$$

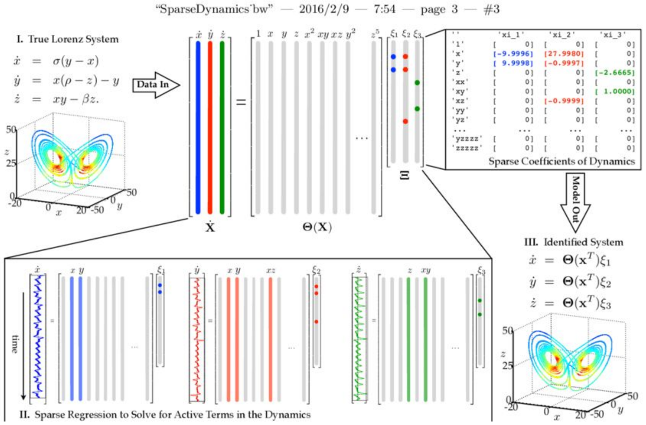
$$\frac{\mathrm{d}\mathbf{x}_t}{dt} = f(\mathbf{x}_t) + g(\mathbf{x}_t)\mathrm{dW}_t \quad \text{(SDE)}$$

1. MOTIVATION

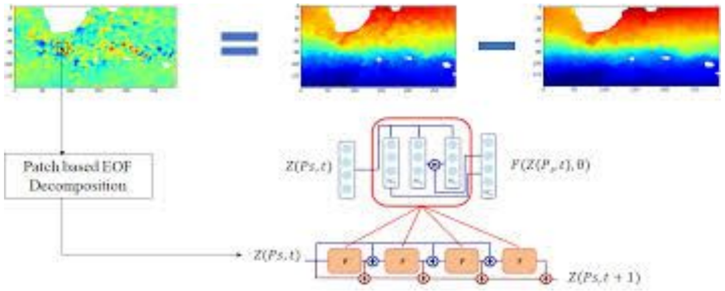


Source: Tandeo et el. 2015.

Dynamical systems identification under ideal conditions



Source: Brunton et al, 2016.



Source: Ouala et al, 2018.

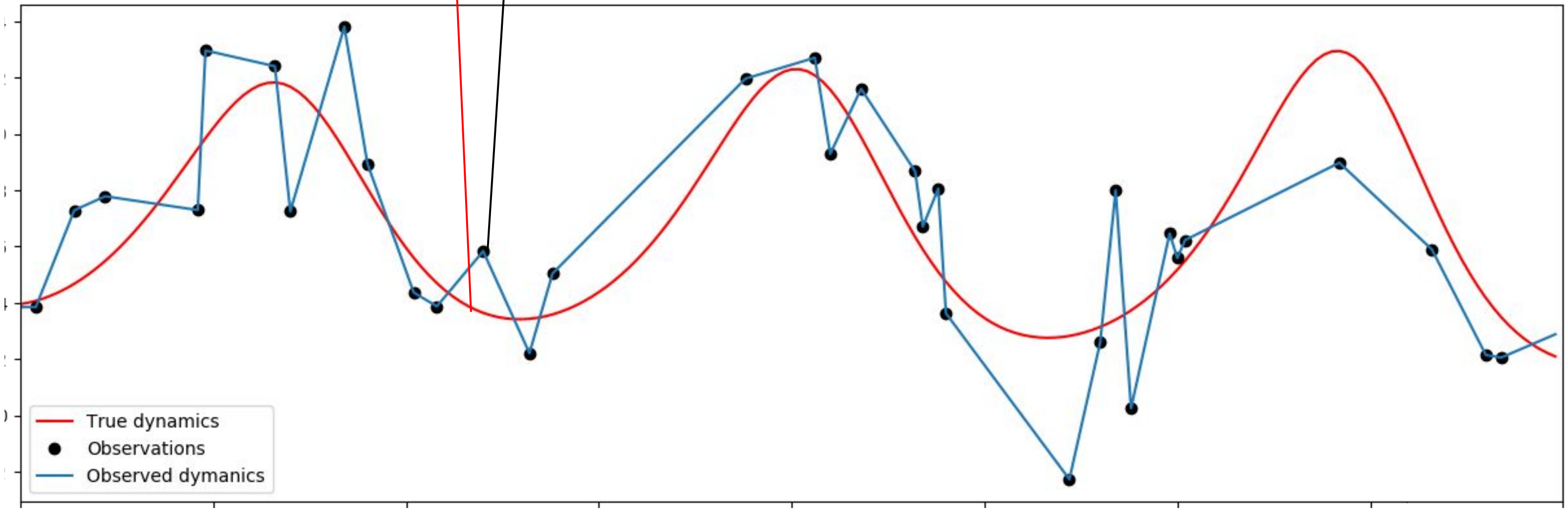
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Lab-STICC

1. MOTIVATION

Dealing with noisy and irregularly-sampled training data ?

$$\mathbf{x}_{t+n} = \mathcal{F}^n\left(\mathbf{x}_t\right) + \boldsymbol{\omega}_{t+n}$$
$$\mathbf{y}_t = \Phi_t\left(\mathcal{H}\left(\mathbf{x}_t\right) + \varepsilon_t\right)$$



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

2. PROPOSED FRAMEWORK

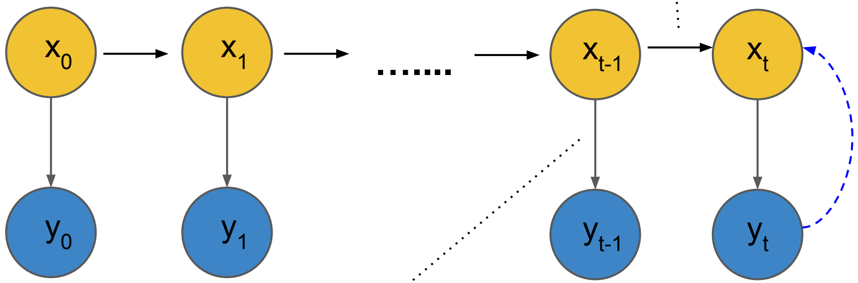
$$\boxed{\begin{aligned} \mathbf{x}_{t+n} &= \mathcal{F}^n(\mathbf{x}_t) + \boldsymbol{\omega}_{t+n} \\ \mathbf{y}_t &= \Phi_t(\mathcal{H}(\mathbf{x}_t) + \varepsilon_t) \end{aligned}}$$

$$p_\theta(\mathbf{x}_{1:T}) = p_\theta(\mathbf{x}_1) \prod_{t=1}^{T-1} p_\theta(\mathbf{x}_{t+1}|\mathbf{x}_t)$$

Dynamical model (NN model)

True states

Noisy and partial observation



Inference model (Data assimilation)

$$q(\mathbf{x}_{1:T}|\mathbf{y}_{1:T}) = \prod_{t=1}^{T} q(\mathbf{x}_t|\mathbf{y}_{1:T})$$

Observation model ($\mathcal{H}$)

$$p_\theta(\mathbf{y}_{1:T}|\mathbf{x}_{1:T}) = \prod_{t=1}^{T} p_\theta(\mathbf{y}_t|\mathbf{x}_t)$$

$$\ln p_\theta(\mathbf{y})$$

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Lab-STICC

$$p_\theta(\mathbf{x}_{t+1}|\mathbf{x}_t)$$

Data-assimilation-based ODE Net (DAODEN)

$$\frac{\mathrm{d}\mathbf{x}_t}{dt} = f(\mathbf{x}_t)$$

$\mathcal{F}^n(\mathbf{x}_t)$: Runge-Kutta 4

$$\begin{aligned} p_{\boldsymbol{\theta}}(\mathbf{x}_{t+n}|\mathbf{x}_{t+n}) &= \mathcal{N}(\boldsymbol{\mu}_{t+n}, \boldsymbol{\Sigma}_{t+n}) \\ \boldsymbol{\mu}_{t+n} &= \mathcal{F}^n(\mathbf{x}_t) \\ \boldsymbol{\Sigma}_{t+n} &= MLP(\mathbf{x}_t, \boldsymbol{\mu}_{t+n}) \end{aligned}$$

Data-assimilation-based Dynamical System Identification Net (DADIN)

$$\frac{\mathrm{d}\mathbf{x}_t}{dt} = f(\mathbf{x}_t) + g(\mathbf{x}_t)\mathrm{d}\mathrm{W}_t$$

$\mathcal{F}^n(\mathbf{x}_t)$: Euler-Maruyama

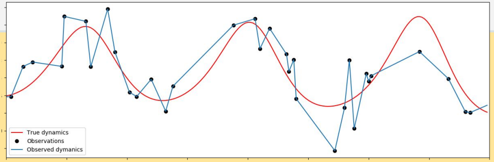
$$\begin{aligned} p_{\boldsymbol{\theta}}(\mathbf{x}_{t+n}|\mathbf{x}_{t+n}) &= \mathcal{N}(\boldsymbol{\mu}_{t+n}, \boldsymbol{\Sigma}_{t+n}) \\ \boldsymbol{\mu}_{t+n} &= \frac{1}{K}\sum_{s=1}^{K} \mathcal{F}^n(\mathbf{x}_t^{(s)}) \\ \boldsymbol{\Sigma}_{t+n} &= \frac{1}{K-1}\mathbf{E}_{t+n}\mathbf{E}_{t+n}^T \\ \mathbf{E}_{t+1}[s,:] &= \mathbf{x}_{t+n} - \boldsymbol{\mu}_{t+n} \end{aligned}$$

2. PROPOSED FRAMEWORK

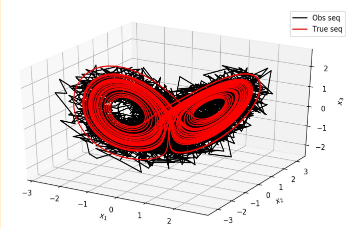
BiNN + RK4/Euler-Maruyama [*]

Dynamical model (NN model)

$$p_\theta(\mathbf{x}_{t+1}|\mathbf{x}_t)$$

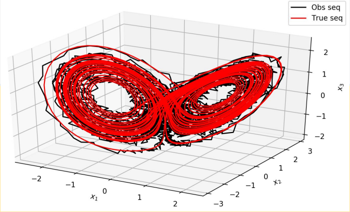


Noisy and partial data



Inference model (data assimilation)

$$\mathbf{x}_t$$



Inferred states

EnKS (BiNN-EnKS)[**]

LSTM-VariationalAutoEncoder (BiNN-LSTM_VAE)

----

[*] Ronan Fablet, Said Ouala, and Cedric Herzet, “Bilinear residual Neural Network for the identification and forecasting of dynamical systems,” arXiv:1712.07003 [physics], Dec. 2017, arXiv: 1712.07003.
[**] D. Nguyen, S. Ouala, L. Drumetz, and R. Fablet, “Em-like learning chaotic dynamics from noisy and partial observations,” SciRate , Mar. 2019. [Online]. Available: https://scirate.com/arxiv/1903.10335

3. EXPERIMENTS & RESULTS

Table 1: Short-term forecasting error and very-long-term forecasting topology of data-driven models learned on noisy Lorenz-63 data.

- **e_4**: prediction error after 4 timesteps
- **$\pi_{0.5}$**: first time when the prediction error reaches 0.5 times the std of the signal.
- **rec**: reconstruction error
- **λ_1**: first Lyapunov exponent

(**r** = std_{noise}/std_{signal})

Model		r			
		8.5%	16.7%	33.3%	66.7%
AnDA	e_4	0.351±0.184	0.777±0.350	1.683±0.724	3.682±1.346
	rec	0.416±0.019	0.941±0.037	2.134±0.076	4.876±0.168
	$\pi_{0.5}$	0.820±0.480	0.380±0.172	0.249±0.174	0.104±0.116
	λ_1	26.517±7.665	27.146±42.927	76.267±28.150	127.047±0.881
SINDy	e_4	0.068±0.052	0.149±0.106	0.311±0.196	0.694±0.441
	$\pi_{0.5}$	0.490±0.261	0.165±0.085	0.077±0.049	0.034±0.034
	λ_1	0.898±0.008	0.840±0.035	0.840±0.035	nan±nan
BiNN	e_4	0.045±0.030	0.119±0.085	0.283±0.185	0.684±0.408
	$\pi_{0.5}$	3.608±1.364	2.053±0.666	0.975±0.488	0.308±0.125
	λ_1	0.900±0.011	0.868±0.010	0.122±0.208	-0.422±0.047
Latent-ODE	e_4	0.051±0.027	0.062±0.034	*0.065±0.042*	0.213±0.084
	$\pi_{0.5}$	2.504±1.332	2.336±1.472	2.852±1.352	2.118±1.129
	λ_1	0.892±0.018	0.877±0.018	0.885±0.015	0.675±0.027
BiNN_EnKS	e_4	**0.019±0.016**	**0.024±0.023**	**0.037±0.024**	0.276±0.160
	rec	0.323±0.024	0.431±0.042	0.598±0.093	1.531±0.332
	$\pi_{0.5}$	2.807±1.128	3.004±1.355	*2.996±1.641*	*2.081±1.214*
	λ_1	0.856±0.031	0.869±0.024	0.826±0.065	0.868±0.014
DAODEN_full	e_4	*0.023±0.015*	*0.027±0.016*	0.072±0.045	*0.187±0.127*
	rec	**0.178±0.050**	*0.258±0.066*	**0.469±0.168**	*1.003±0.380*
	$\pi_{0.5}$	*3.533±1.139*	**3.496±1.215**	**3.426±1.512**	1.897±0.918
	λ_1	0.869±0.036	0.858±0.028	0.881±0.024	0.884±0.013
DADIN_full	e_4	0.061±0.045	0.056±0.032	0.109±0.067	**0.108±0.077**
	rec	*0.235±0.034*	0.350±0.091	*0.589±0.170*	**0.976±0.357**
	$\pi_{0.5}$	*3.066±1.444*	**3.805±1.579**	2.362±0.969	**2.253±0.914**
	λ_1	0.885±0.015	0.880±0.018	0.867±0.012	0.799±0.016

3. EXPERIMENTS & RESULTS

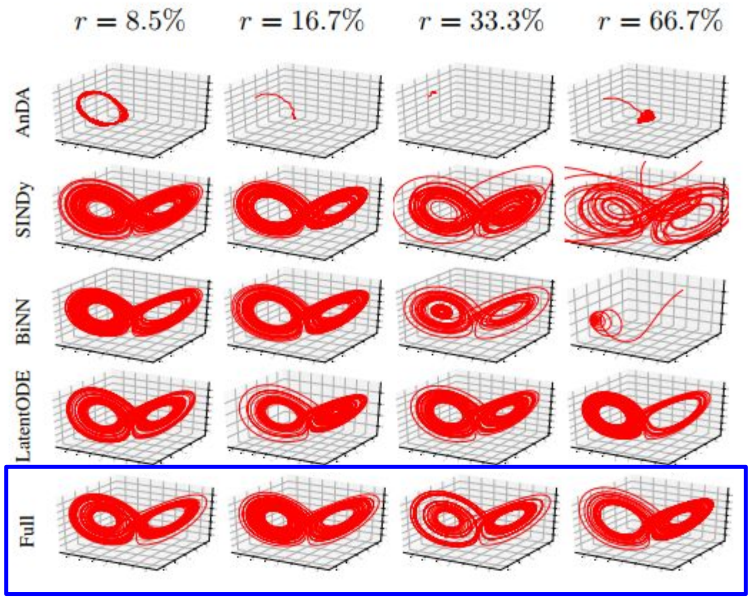


Fig. 1: Attractors generated by models trained on noisy data.

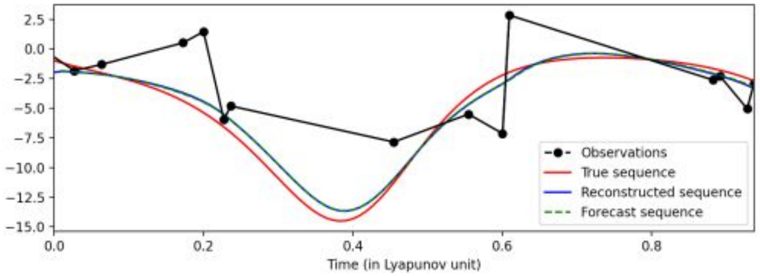


Fig. 2: An example of the the first dimension of the L63 system reconstructed by the inference module of our model. The observations are noisy (r = 33%) and irregularly sampled with a missing rate of 87.5%

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Lab-STICC

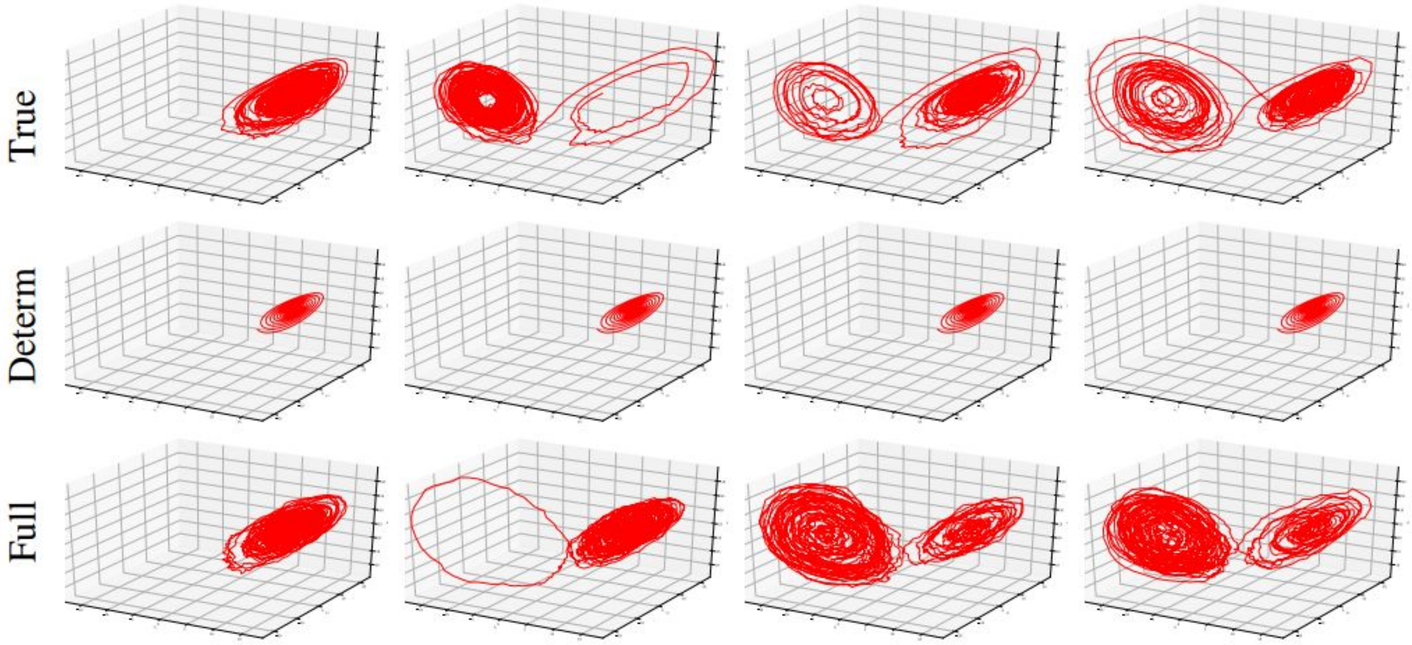


Fig. 3: Learning stochastic dynamics.
Because our model uses a probabilistic formulation, it can take into account stochastic variabilities, model errors and reconstruction uncertainties

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Lab-STICC

4. CONCLUSIONS & PERSPECTIVES

Conclusions:

- The **combination of data assimilation and neural network** is a very promising approach for dynamical systems identification.
- Inference schemes help models become much more **robust to noise** and can **deal with partial/irregular sampling**.
- The proposed framework is **general**, we can use **any suitable dynamical model/inference model.**
- It can capture stochasticity both in the observation and in the dynamics.

Perspectives:

- High dimensional states?
- **Some components of the hidden states are never observed?**