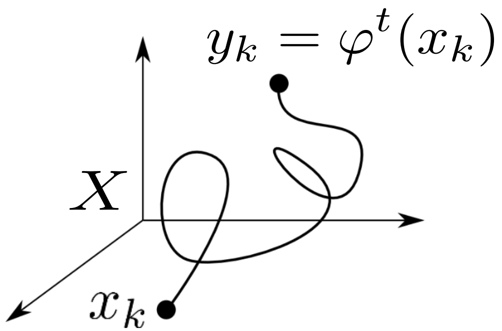
Dual identification methods in Koopman operator theory: from ODEs to PDEs

Alexandre Mauroy
(University of Namur, Belgium)

UNIVERSITÉ DE NAMUR

naxys
Namur Institute for Complex Systems

We aim at identifying a vector field from (possible low-sampled) data points generated by the dynamics



$$(x_k, \varphi^t(x_k)) \in X \times X$$

$$k = 1, \ldots, K$$

Find $F : X \to X$ such that

$$\dot{x} = F(x)$$

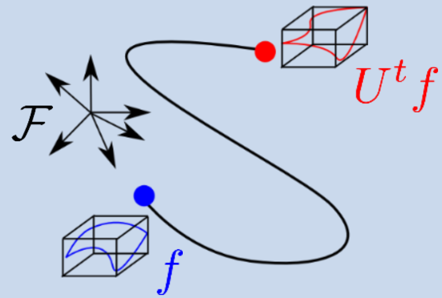
Constraint

No direct estimation of time derivatives
(e.g. low sampling rate, noise)

We can identify the linear Koopman operator instead of identifying the nonlinear system

Linear Koopman operator

Operator $U^t : \mathcal{F} \to \mathcal{F}$
acting on a functional space

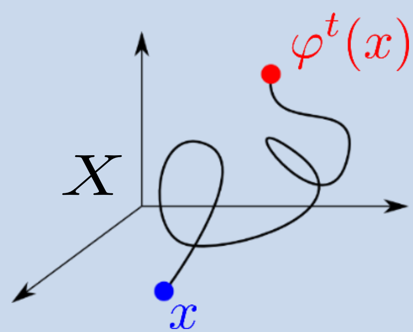


Linear identification
(infinite-dimensional)

LIFTING

Nonlinear system

Flow $\varphi^t : X \to X$
acting on the state space



Nonlinear identification
(finite-dimensional)

Outline

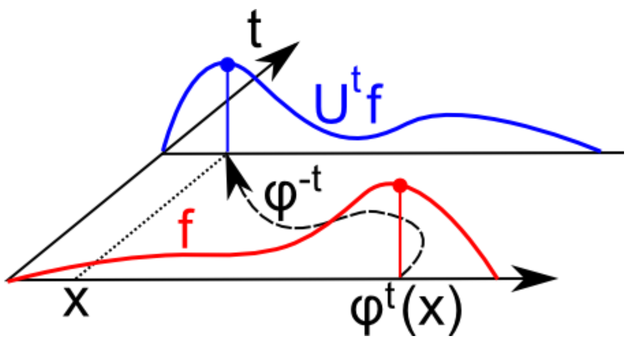
The Koopman operator in a nutshell

Two dual lifting methods
(joint work with J. Goncalves, U. Luxembourg)

Extension to PDEs
(preliminary work)

The Koopman operator in a nutshell

Semigroup of Koopman operators



$f \in \mathcal{F}$ is an observable

$$U^t f(x) = f \circ \varphi^t(x)$$

[Koopman 1931, Budisic, Mohr and Mezic 2012]

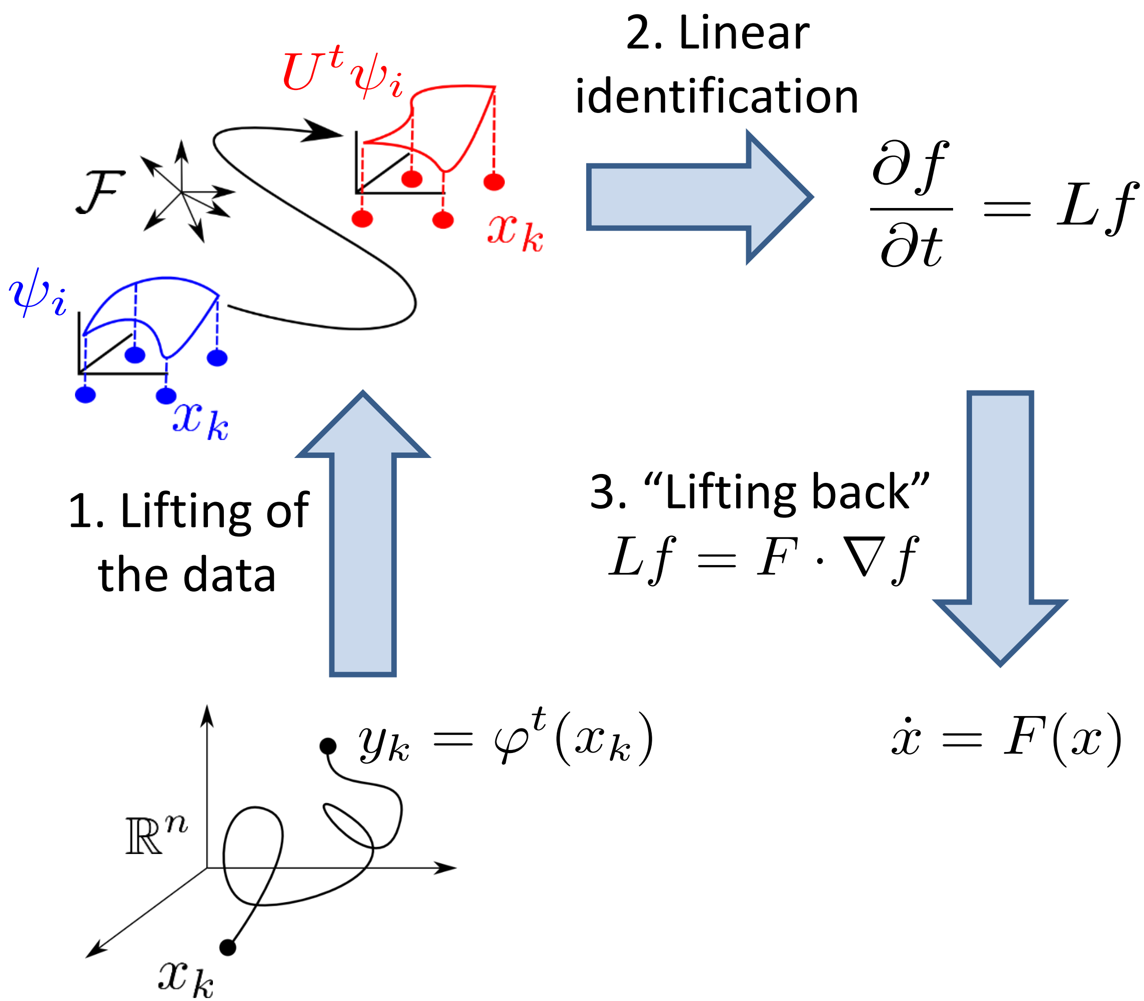
Strongly continuous semigroup

$$\lim_{t \downarrow 0} \|U^t f - f\| = 0 \quad \forall f \qquad (\mathcal{F} = C^0(X), L^2(X))$$

Infinitesimal generator of the Koopman operator:

$$Lf = \lim_{t \downarrow 0} \frac{U^t f - f}{t} \qquad \dot{x} = F(x) \Rightarrow Lf = F \cdot \nabla f$$

We estimate the vector field by identifying the infinitesimal generator in the lifted space



Outline

The Koopman operator in a nutshell

Two dual lifting methods
(joint work with J. Goncalves, U. Luxembourg)

Extension to PDEs
(preliminary work)

Our identification method relies on the Trotter-Kato approximation theorem

Trotter-Kato approximation theorem:

U^t, U^t_N strongly continuous semigroups with generators L, L_N

If $\|L_N f - Lf\| \to 0 \quad \forall f \in \mathcal{D}(L)$, then $\|U^t_N f - U^t f\| \to 0 \quad \forall f$

$L_N = P_N L P_N$

$U_N = e^{P_N L P_N t}$

P_N : orthogonal projection onto $\mathcal{F}_N = \mathrm{span}\{\psi_i\}_{i=1}^N$

$\mathcal{F}_\infty$ dense in $L^2(X)$

$\mathcal{F}_N \subseteq \mathcal{D}(L)$

$$\|e^{P_N L P_N t}\psi_i - P_N U^t \psi_i\|_{L^2} \to 0$$

matrix representation

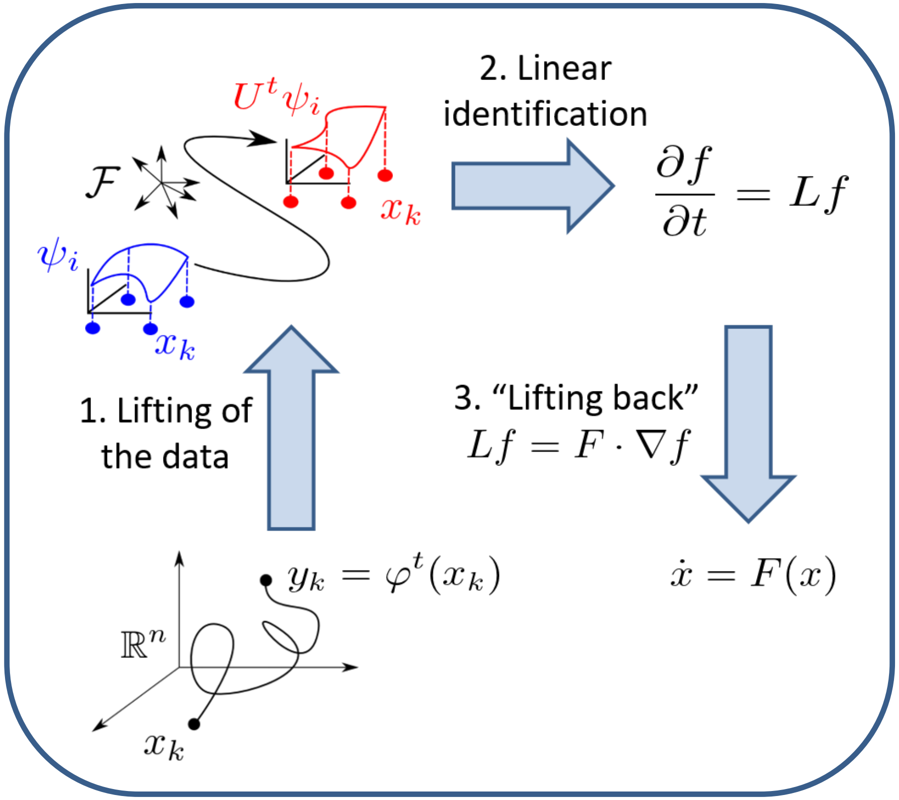
$$e^{\mathbf{L_N} t} - \mathbf{U_{\mathrm{EDMD}}} \to \mathbf{0}$$

⚠ Nyquist

matrix logarithm

$$\mathbf{L_N} - \frac{1}{t}\log \mathbf{U_{\mathrm{EDMD}}} \to \mathbf{0}$$

The method yields the expansion of the vector field in the basis of functions $\{\psi_i\}_{i=1}^N$



Step 3

Step 1

$$\boldsymbol{\Psi}_{\mathbf{X}} = \begin{pmatrix} \psi_1(x_1) & \cdots & \psi_N(x_1) \\ \vdots & & \vdots \\ \psi_1(x_K) & \cdots & \psi_N(x_K) \end{pmatrix}$$

$$\boldsymbol{\Psi}_{\mathbf{Y}} = \begin{pmatrix} \psi_1(y_1) & \cdots & \psi_N(y_1) \\ \vdots & & \vdots \\ \psi_1(y_K) & \cdots & \psi_N(y_K) \end{pmatrix}$$

Step 2

$$\mathbf{U}_{\mathrm{EDMD}} = \boldsymbol{\Psi}_{\mathbf{X}}^{\dagger}\boldsymbol{\Psi}_{\mathbf{Y}}$$

$$\mathbf{L_N} \approx \frac{1}{t}\log(\boldsymbol{\Psi}_{\mathbf{X}}^{\dagger}\boldsymbol{\Psi}_{\mathbf{Y}})$$

Assume:
$\psi_n = \mathrm{id}_j$
$F_j \in \mathcal{F}_N$

$$\begin{array}{cc} & n \\ \mathbf{L_N} = & \left[\quad \boxed{} \quad \right] \end{array}$$

$$\sum_{i=1}^{N} [\mathbf{L_N}]_{in}\psi_i = P_N L\mathrm{id}_j = P_N F_j = F_j$$

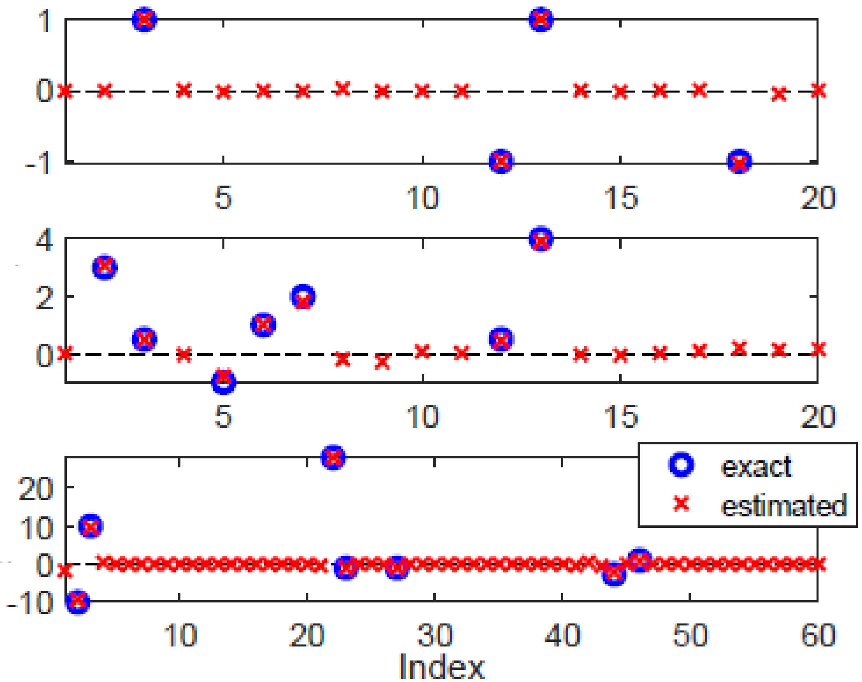
EDMD, *Williams et al., JNLS, 2015]*
[AM and Goncalves, CDC 2016]

The method is efficient to recover the vector field with low sampling rate

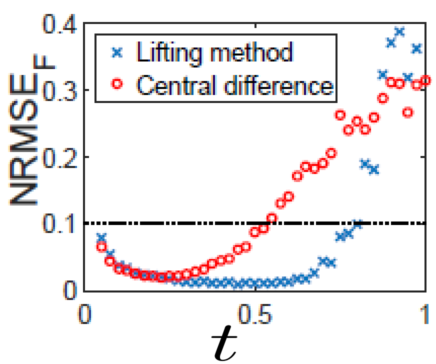
Van der Pol (limit cycle)

Unstable equilibrium

Lorenz (chaotic)



Lifting method vs central differences (Van der Pol)



Extensions

Inputs, stochastic

Under some conditions, we have convergence guarantees

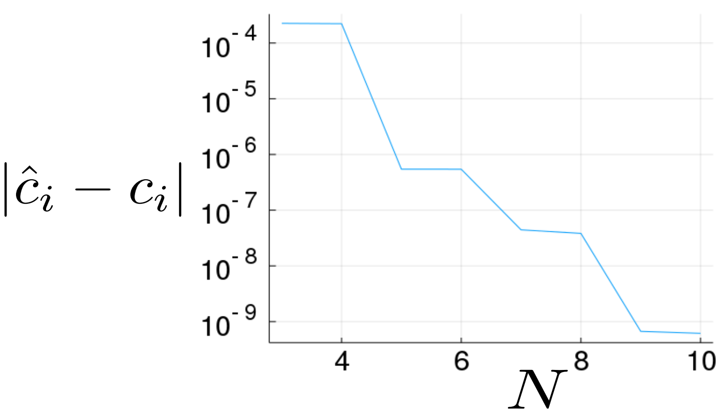
Theorem

Assumptions:

- $F = \sum_i c_i \psi_i$ polynomial (basis of monomials)
- invertible non-singular flow
- sample points uniformly randomly distributed in compact invariant set

Then the estimated vector field is given by $\hat{F} = \sum_i \hat{c}_i \psi_i$ with

$$\lim_{N \to \infty} \lim_{t \to 0} \lim_{K \to \infty} \hat{c}_i = c_i \qquad \text{(with probability one).}$$



Exponential convergence rate

[Kurdila and Bobade, Koopman theory and nonlinear approximation spaces, 2018]

Under some conditions,
we have convergence guarantees

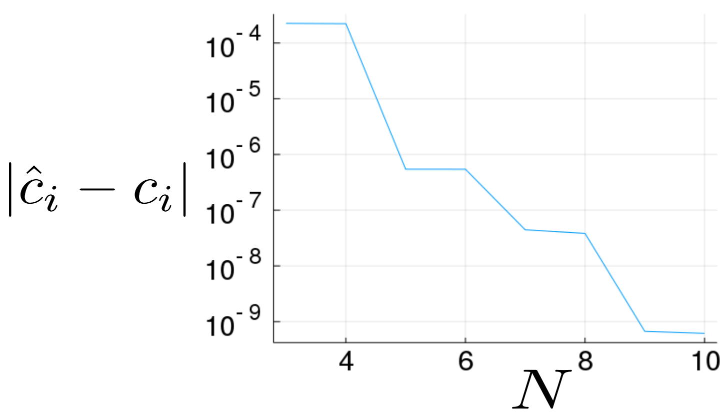
Theorem

Assumptions:

- ~~$F = \sum_i c_i \psi_i$ polynomial (basis of monomials)~~
- invertible non-singular flow
- sample points uniformly randomly distributed in compact invariant set

Then the estimated vector field is given by $\hat{F} = \sum_i \hat{c}_i \psi_i$ with

$$\lim_{N \to \infty} \lim_{t \to 0} \lim_{K \to \infty} \|F - \hat{F}\|_{L^2} = 0$$ (with probability one).



Exponential convergence rate

[Kurdila and Bobade, Koopman theory and nonlinear approximation spaces, 2018]

A dual method can also be obtained

Main method $(K \geq N)$

$$\mathbf{U}_{\text{EDMD}} = \mathbf{\Psi}_{\mathbf{X}}^{\dagger} \mathbf{\Psi}_{\mathbf{Y}}$$

$$\mathbf{U}_{\text{EDMD}} \in \mathbb{R}^{N \times N}$$

Function space $\mathcal{F}_N$

Dual method $(K \leq N)$

$$\mathbf{U}_{\text{DUAL}} = \mathbf{\Psi}_{\mathbf{X}} \mathbf{U}_{\text{EDMD}} \mathbf{\Psi}_{\mathbf{X}}^{\dagger} = \mathbf{\Psi}_{\mathbf{Y}} \mathbf{\Psi}_{\mathbf{X}}^{\dagger}$$

$$\mathbf{U}_{\text{DUAL}} \in \mathbb{R}^{K \times K}$$

$$\begin{pmatrix} U^t \psi_i(x_1) \\ \vdots \\ U^t \psi_i(x_K) \end{pmatrix} \approx \mathbf{U}_{\text{DUAL}} \begin{pmatrix} \psi_i(x_1) \\ \vdots \\ \psi_i(x_K) \end{pmatrix}$$

« Sample space » $\tilde{\mathcal{F}}_K$

$\tilde{P}_K U \psi_i \approx \tilde{U}_K \tilde{P}_K \psi_i$ $\quad$ $\tilde{P}_K$: projection onto $\tilde{\mathcal{F}}_K$ (piecewise-constant functions)

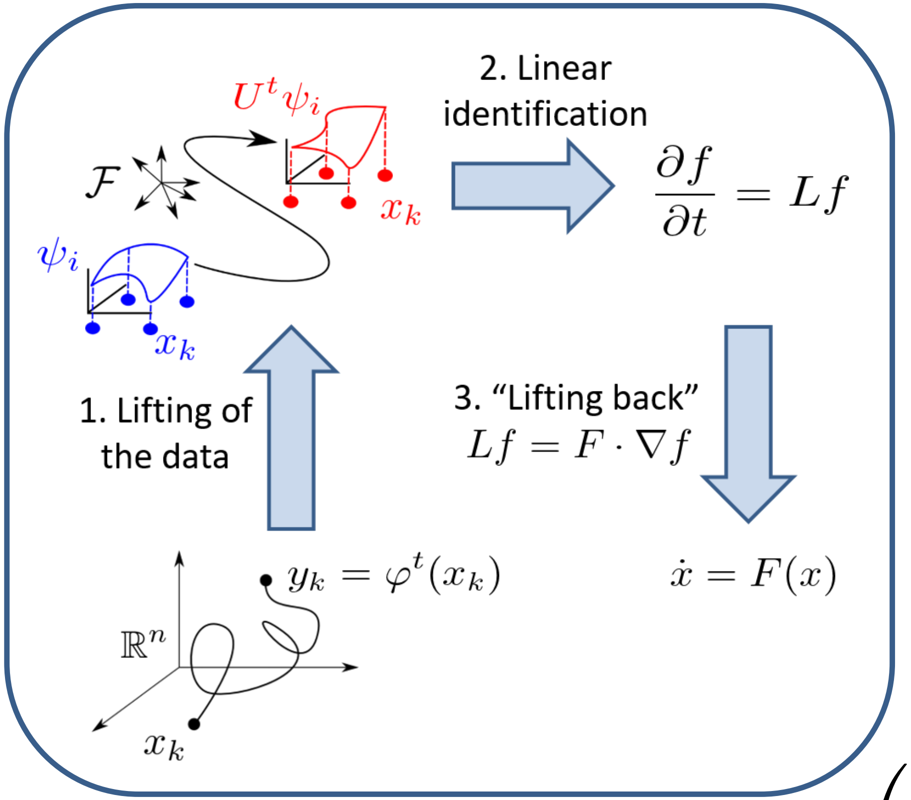
$\tilde{L}_K \tilde{P}_K \psi_i = \tilde{P}_K L \psi_i$ $\quad$ $\psi_i \in \mathcal{D}(L)$

$$\| e^{\tilde{L}_K t} \tilde{P}_K \psi_i - \tilde{P}_K U^t \psi_i \|_{L^2} \to 0 \quad \text{(Trotter-Kato theorem)}$$

$$\| e^{\tilde{\mathbf{L}}_{\mathbf{K}} t} - \mathbf{U}_{\text{DUAL}} \|_2 \to \mathbf{0}$$

$$\| \tilde{\mathbf{L}}_{\mathbf{K}} - \frac{1}{t} \log \mathbf{U}_{\text{DUAL}} \|_2 \to \mathbf{0}$$

The dual method yields the value of the vector field at the sample points



Step 1

Construct $\mathbf{\Psi_X}, \mathbf{\Psi_Y}$

Step 2

$$\mathbf{U}_{\text{DUAL}} = \mathbf{\Psi_Y \Psi_X^\dagger}$$

$$\tilde{\mathbf{L}}_{\mathbf{K}} \approx \frac{1}{t} \log(\mathbf{\Psi_Y \Psi_X^\dagger})$$

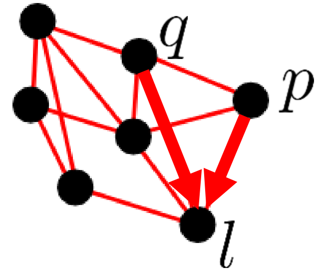
Step 3

$$\tilde{\mathbf{L}}_{\mathbf{K}} \begin{pmatrix} \mathrm{id}_j(x_1) \\ \vdots \\ \mathrm{id}_j(x_K) \end{pmatrix} \approx \begin{pmatrix} L\mathrm{id}_j(x_1) \\ \vdots \\ L\mathrm{id}_j(x_K) \end{pmatrix} = \begin{pmatrix} F_j(x_1) \\ \vdots \\ F_j(x_K) \end{pmatrix}$$

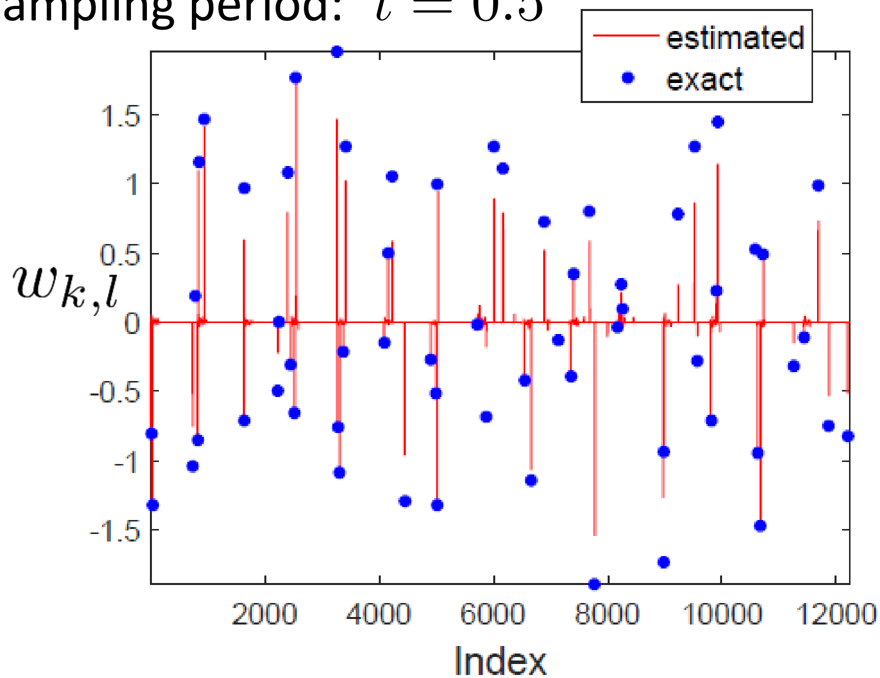
(+ regression, e.g. Lasso)

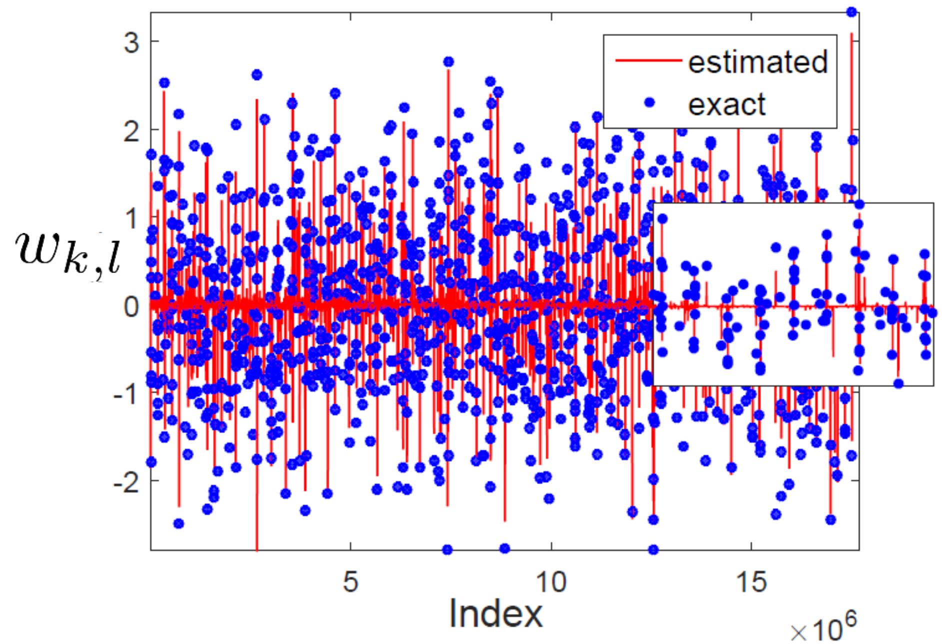
The method is efficient to reconstruct large networks with nonlinear dynamics

$$\dot{x}_l = \sum_{k=1}^{N_F} w_{k,l}\, h_k(\mathbf{x}) \qquad h_k \equiv x_p^a x_q^b \quad a+b \in \{1,2,3\}$$



Sampling period: $t = 0.5$





$n = 20$ states (nodes)

$N_F \approx 3.10^4$ coefficients

$K = 200$ data points

$n = 100$ states (nodes)

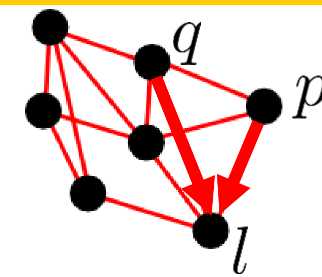
$N_F \approx 2.10^7$ coefficients

$K = 1000$ data points

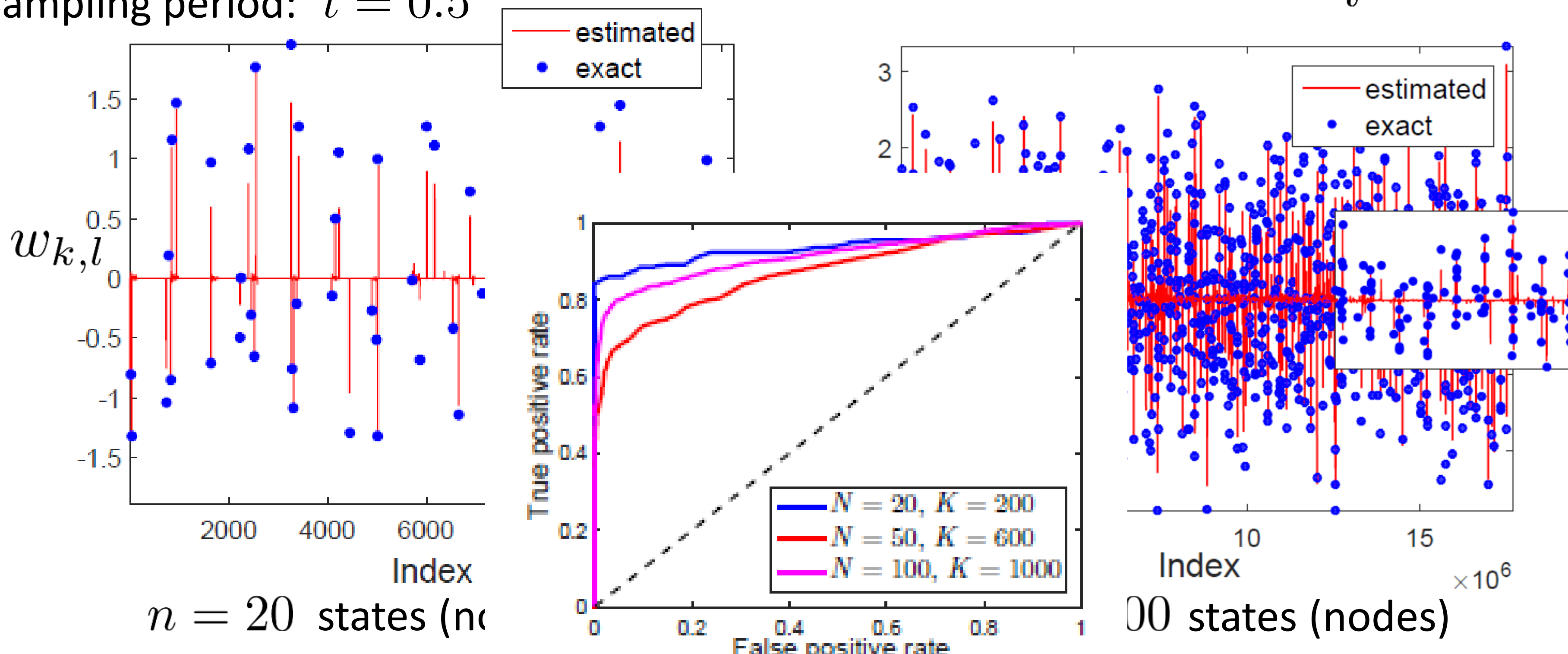
The method is efficient to reconstruct large networks with nonlinear dynamics

$$\dot{x}_l = \sum_{k=1}^{N_F} w_{k,l}\, h_k(\mathbf{x}) \qquad h_k \equiv x_p^a x_q^b$$

$$a + b \in \{1, 2, 3\}$$



Sampling period: $t = 0.5$



$n = 20$ states (nc

$N_F \approx 3.10^4$ coefficients

$K = 200$ data points

00 states (nodes)

$N_F \approx 2.10^7$ coefficients

$K = 1000$ data points

Under some conditions, we have convergence guarantees

Theorem

- invertible non-singular flow
- sample points uniformly randomly distributed over compact invariant set
- specific properties for the test functions ψ_i (e.g. Gaussian radial basis functions)

Then, the estimated vector field $\hat{F}$ satisfies

$$\lim_{t\to 0}\lim_{K\to}\lim_{N\to\infty}\frac{1}{K}\sum_{k=1}^{K}\|\hat{F}(x_k)-F(x_k)\|^2=0$$ with probability one.

Convergence rate

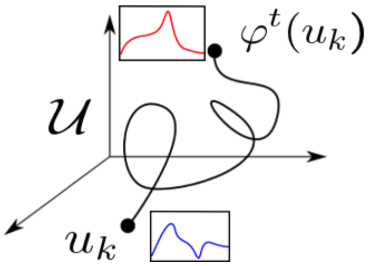
Outline

The Koopman operator in a nutshell

Two dual lifting methods
(joint work with J. Goncalves, U. Luxembourg)

Extension to PDEs
(preliminary work)

Can we use a similar method to identify PDEs?

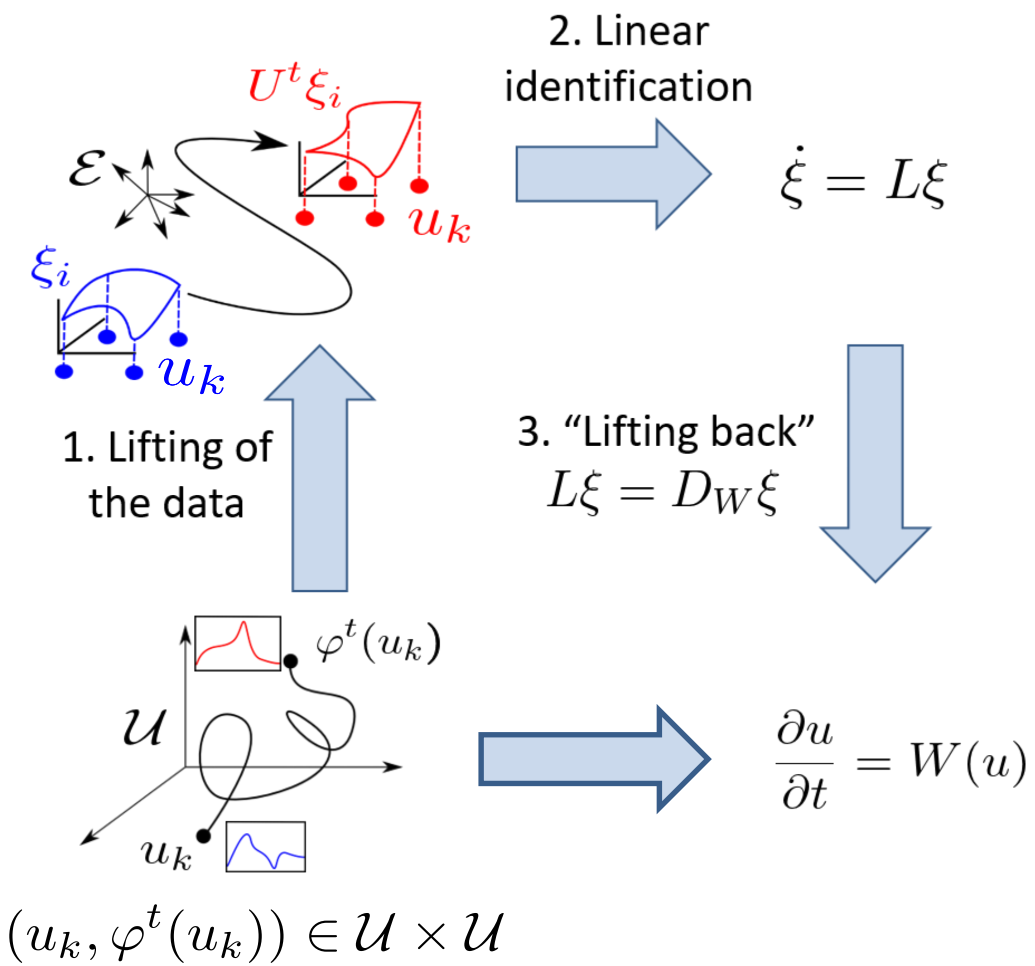


$$(u_k, \varphi^t(u_k)) \in \mathcal{U} \times \mathcal{U}$$

$$\frac{\partial u}{\partial t} = W(u)$$

[Rudy, Brunton, Proctor and Kutz, Data-driven discovery of PDEs, Science Advances, 2017]

Can we use a similar method to identify PDEs?



A semigroup of Koopman operators is defined in the case of nonlinear PDEs

[Mezic, Spectral Koopman operator methods in dynamical systems]
[Nakao, Proceedings of SICE, 2018] - Hiroya Nakao's talk

Nonlinear PDE

$$\frac{\partial u}{\partial t} = W(u) \qquad u : X \to \mathbb{R},\ u \in \mathcal{U} \qquad W : \mathcal{D}(W) \to \mathcal{U}$$

Nonlinear semigroup

$$\varphi^t : \mathcal{U} \to \mathcal{U}$$

(Nonlinear) observable functionals

$$\xi : \mathcal{U} \to \mathbb{C} \qquad \xi \in \mathcal{E}$$

Semigroup of Koopman operators

$$U^t : \mathcal{E} \to \mathcal{E} \qquad U^t \xi = \xi \circ \varphi^t$$

Infinitesimal generator $\xi \in \mathcal{D}(L)$

$$L\xi(u) = \lim_{t \downarrow 0} \frac{\xi(\varphi^t(u)) - \xi(u)}{t} = \left.\frac{\xi(\varphi^t(u))}{dt}\right|_{t=0} = D_{W(u)}\xi(u)$$

$$\triangleq \lim_{\lambda \to 0} \frac{\xi(u + \lambda W(u)) - \xi(u)}{\lambda}$$

(Gâteaux derivative)

? Strong continuity?

The infinitesimal generator applied to a specific functional provides information on the PDE

Basis of nonlinear functionals ξ_i $\quad\quad \mathrm{span}\{\xi_i\}_{i=1}^N = \mathcal{E}_N \subset \mathcal{E}$

$$\xi_1(u) = \langle w, u \rangle_{L^2(X)}$$

for some $w \in L^2(X)$

$$\xi_i(u) = \langle w, W_i(u) \rangle_{L^2(X)} \quad i > 1$$

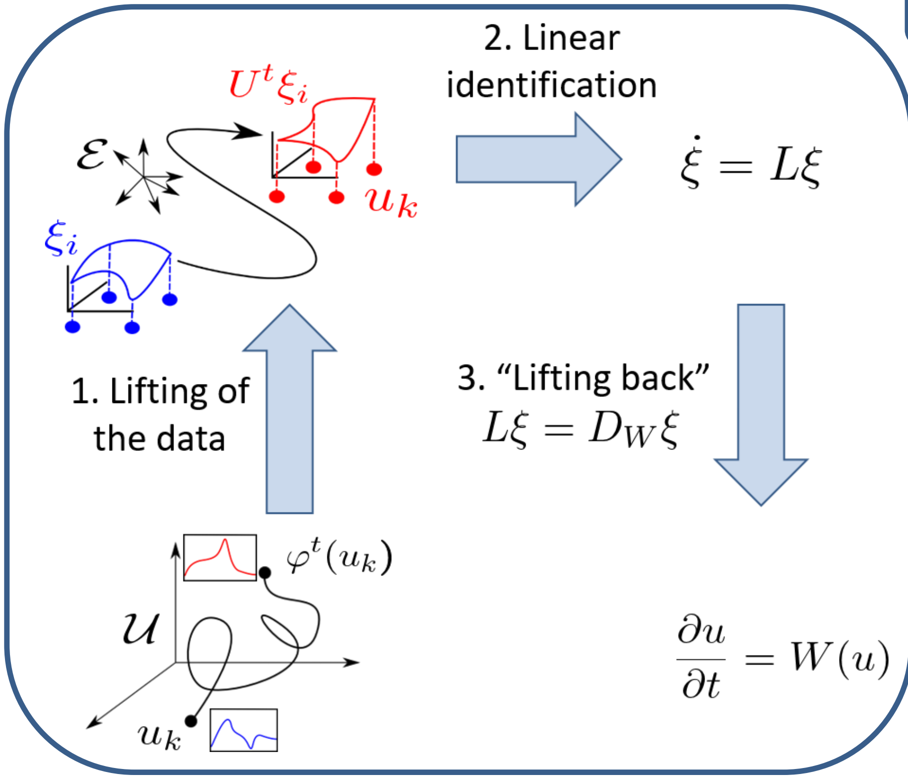
Action of the generator

$$\boxed{L\xi_1(u) = D_{W(u)} \langle w, u \rangle_{L^2(X)} = \langle w, W(u) \rangle_{L^2(X)}}$$

$$L\xi_1 = \sum_{i=1}^N c_i \xi_i \quad (\forall w) \quad \Leftrightarrow \langle w, W(u) \rangle_{L^2(X)} = \textstyle\sum_i c_i \langle w, W_i(u) \rangle_{L^2(X)} \quad (\forall w, \forall u)$$

$$\Leftrightarrow \quad W = \sum_{i=1}^N c_i \, W_i$$

We obtain a lifting method for identifying nonlinear PDEs



Step 1

$$\boldsymbol{\Xi}_{\mathbf{X}} = \begin{pmatrix} \xi_1(u_1) & \cdots & \xi_N(u_1) \\ \vdots & & \vdots \\ \xi_1(u_K) & \cdots & \xi_N(u_K) \end{pmatrix}$$

$$\boldsymbol{\Xi}_{\mathbf{Y}} = \begin{pmatrix} \xi_1(\varphi^t(u_1)) & \cdots & \xi_N(\varphi^t(u_1)) \\ \vdots & & \vdots \\ \xi_1(\varphi^t(u_K)) & \cdots & \xi_N(\varphi^t(u_K)) \end{pmatrix}$$

Step 2

$$\mathbf{L_N} \approx \frac{1}{t} \log(\boldsymbol{\Xi}_{\mathbf{X}}^{\dagger} \boldsymbol{\Xi}_{\mathbf{Y}})$$

Step 3

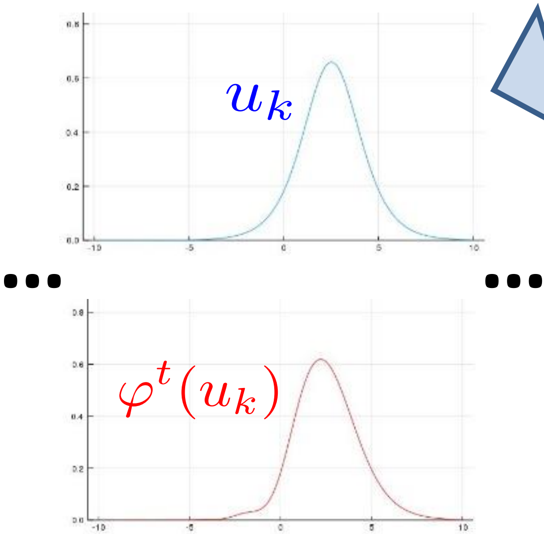
Assume $L\xi_1(\cdot) = \langle w, W(\cdot) \rangle \in \mathcal{E}_N$

$$\mathbf{L_N} = \begin{array}{c} \mathbf{1} \\ \left[\quad \right] \end{array}$$

$$L\xi_1 = \sum_{i=1}^{N} [\mathbf{L_N}]_{i1} \xi_i \quad \Rightarrow \quad W = \sum_{i=1}^{N} [\mathbf{L_N}]_{i1} W_i$$

Preliminary numerical results suggest that the method is efficient to identify nonlinear PDEs

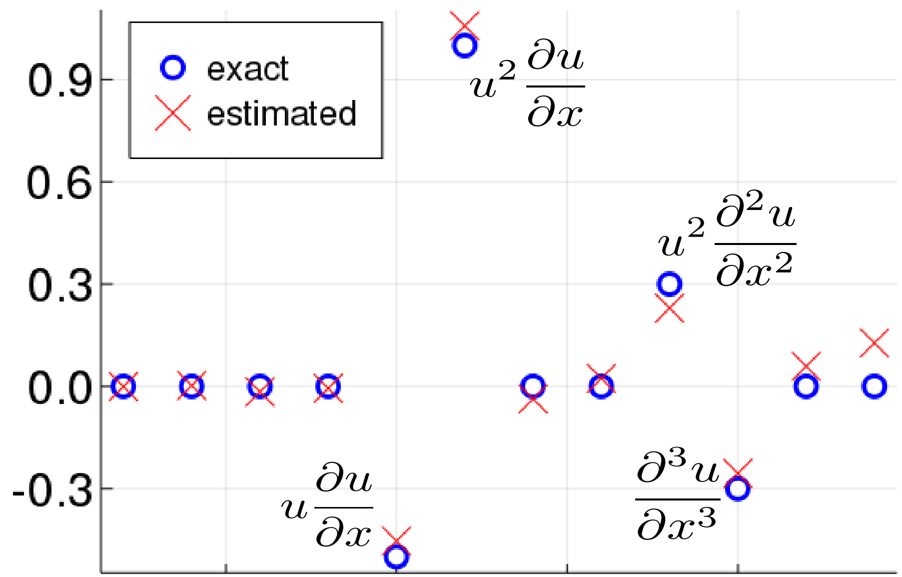
$$\frac{\partial u}{\partial t} = -0.5u\frac{\partial u}{\partial x} + u^2\frac{\partial u}{\partial x} - 0.3\frac{\partial^3 u}{\partial x^3} + 0.3u^2\frac{\partial^2 u}{\partial x^2}$$



$K = 20 \quad t = 0.75$

$$\xi_i(u_k) = \left\langle w, u_k^a \frac{\partial^b u_k}{\partial x^b} \right\rangle_{L^2(X)}$$

$a \in \{0, 1, 2\} \qquad b \in \{0, 1, 2, 3\}$



A not so smart idea: What about identifying the « Koopman operator of the Koopman operator »?

$$\mathcal{E} \xrightarrow{U^t_{U_\varphi}} \mathcal{E} \qquad \frac{\partial \xi}{\partial t}(u) = D_{F\cdot\nabla u}\xi(u)$$

$$\mathcal{U} \xrightarrow{U^t_\varphi} \mathcal{U} \qquad \frac{\partial u}{\partial t} = F \cdot \nabla u$$

$$X \xrightarrow{\varphi^t} X \qquad \dot{x} = F(x)$$

ξ linear functional

$$U^t_{U_\varphi}\xi(u) = \xi(U^t_\varphi u) = (U^t_\varphi)^*\xi(u)$$

dual method

➔ projection of the dual operator (PF)

$$\xi_i(u) = \int_X \delta(x - x_i)\, u(x)\, dx = u(x_i)$$

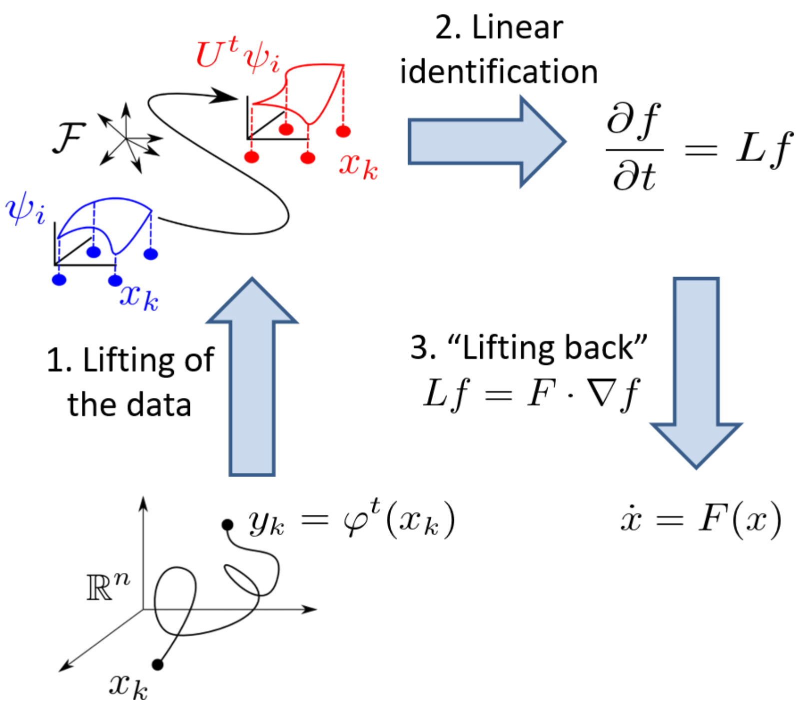
$$\Xi_{\mathbf{X}} = \begin{pmatrix} \xi_1(u_1) & \cdots & \xi_N(u_1) \\ \vdots & & \vdots \\ \xi_1(u_K) & \cdots & \xi_N(u_K) \end{pmatrix} = \begin{pmatrix} u_1(x_1) & \cdots & u_1(x_N) \\ \vdots & & \vdots \\ u_K(x_1) & \cdots & u_K(x_N) \end{pmatrix} = \Psi_X^T$$

sample functions ⇐ u ⇒ basis/test functions

basis functionals ⇐ ξ_i ⇒ sample points

$$\Xi_{\mathbf{X}}{}^{\dagger}\Xi_{\mathbf{Y}} = (\mathbf{\Psi}_{\mathbf{X}}^T)^{\dagger}\mathbf{\Psi}_{\mathbf{Y}}^T = \mathbf{U}_{\mathrm{DUAL}}^T$$

We have proposed a lifting method for system identification / parameter estimation



Two dual methods

Convergence results

Extension to PDEs

Perpsectives

Classic system input-output system identification (unobserved states)

Extension to PDEs (to be continued)

Dual identification methods in Koopman operator theory: from ODEs to PDEs

Alexandre Mauroy
(University of Namur, Belgium)

UNIVERSITÉ DE NAMUR

[AM and Goncalves, Koopman-based lifting techniques for nonlinear systems identification, arxiv]

Codes (Matlab, Julia) available online soon

naxys
Namur Institute for Complex Systems