Stochastic model reduction: from nonlinear Galerkin to parametric inference

Fei Lu

Department of Mathematics, Johns Hopkins

Joint work with: Alexandre J. Chorin (UC Berkeley)
Kevin K. Lin (U. of Arizona)

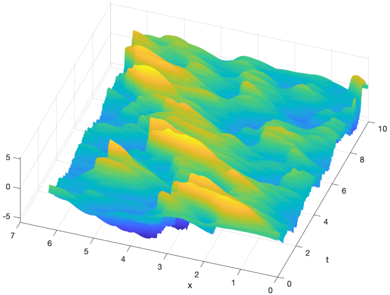
May 22, 2019
SIAM DS19, Snowbird

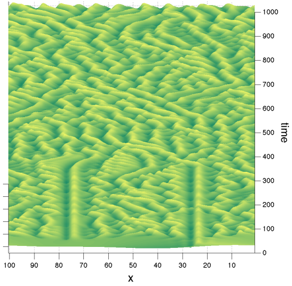
Consider dissipative PDEs in operator form:

$$v_t = \underbrace{Av}_{self-adjoint} + \underbrace{B(v)}_{nonlinear} + f,$$

Examples:

- Burgers $v_t = \nu v_{xx} - v v_x + f(x,t)$,
- Kuramoto-Sivashinsky: $v_t = -v_{xx} - \nu v_{xxxx} - v v_x$





To resolve the Eq. by Fourier-Galerkin (when periodic BC)

$$\frac{d}{dt}\widehat{v}_k = -q_k^\nu \widehat{v}_k + \frac{ik}{2} \sum_{|l|\leq N, |k-l|\leq N} \widehat{v}_l \widehat{v}_{k-l} + \widehat{f}_k(t),$$

Need: $N \gtrsim 5/\nu$ Fourier modes, $dt \sim 1/N$.
E.g. $\nu = 10^{-4}$: spatial grid= 5×10^4, time steps= $5T \times 10^4$

We are mainly interested in large scales, $K << N$.
Question: a reduced model for $(\widehat{v}_{1:K})$?

Reduce spatial dimension + Increase time step-size

Motivation: data assimilation in weather/climate prediction

High-dimensional Full system + **Discrete partial data** - - ▸ **Prediction**

$x' = f(x) + U(x, y),$
$y' = g(x, y).$

Observe only
$\{x(nh)\}_{n=1}^{N}.$

Forecast
$x(t), t \geq Nh.$

- HighD multiscale full chaotic/ergodic systems:
  - can only afford to resolve $x' = f(x)$ online
  - y: unresolved variables (subgrid-scales)
- Discrete noisy observations: missing i.c.
- Ensemble prediction: need many simulations

$x' = f(x) + U(x, y)$, $y' = g(x, y)$.
Data $\{x(nh)\}_{n=1}^{N}$

Objective: Develop a closed reduced model of x that

- captures key statistical + dynamical properties
- use it for online state estimation and prediction

[Approximate the stochastic process $(x(t), t > 0)$ in distribution.]

Various efforts in closure model reduction:

- Direct constructions:
  - non-linear/Petrov- Galerkin: $y(t) = F(x(t))$
  - Mori-Zwanzig formalism (memory)
    $\rightarrow$ statistical approximation by a non-Markov process
  - relaxation approximations
  - linear response / filtering / feedback control
  - ...

- Inference/Data-driven ROM
  - hypoellitpic SDEs, GLEs and SDDEs
  - discrete-time (time series) models
  - data-driven: POD, DMD, Kooperman operator
  - nonparametric inference
  - machine learning (NN's) ...

Inference-based model reduction

SDEs or time series – dynamical models

Differential system or discrete-time system?

$X' = f(X) + Z(t,\omega)$	$X_{n+1} = X_n + R_h(X_n) + Z_n$
informative	non-intrusive
Inference[1]	likelihood
Discretization[2]	error correction by data

[1] Brockwell, Sørensen, Pokern, Wiberg, Samson,...

[2] Milstein, Tretyakov, Talay, Mattingly, Stuart, Higham, ...

Discrete-time stochastic parametrization

NARMA(p, q) [Chorin-Lu (15)]

$$X_n = X_{n-1} + R_h(X_{n-1}) + Z_n,$$

$$Z_n = \Phi_n + \xi_n,$$

$$\Phi_n = \underbrace{\sum_{j=1}^{p} a_j X_{n-j} + \sum_{j=1}^{r}\sum_{i=1}^{s} b_{i,j} P_i(X_{n-j})}_{\text{Auto-Regression}} + \underbrace{\sum_{j=1}^{q} c_j \xi_{n-j}}_{\text{Moving Average}}$$

- $R_h(X_{n-1})$ from a numerical scheme for $x' \approx f(x)$
- Φ_n depends on the past
- NARMAX in system identification $Z_n = \Phi(Z, X) + \xi_n$,

Tasks:

Structure derivation: terms and orders (p, r, s, q) in Φ_n;

Parameter estimation: $a_j, b_{i,j}, c_j$, and σ. Conditional MLE

Model reduction for dissipative PDEs by parametric inference

- Kuramoto-Sivashinsky: $v_t = -v_{xx} - \nu v_{xxxx} - v v_x$
- Burgers: $v_t = \nu v_{xx} - v v_x + f(x,t)$,

Goal: a closed model for $(\widehat{v}_{1:K})$, $K = 2K_0 << N$.

$$\frac{d}{dt}\widehat{v}_k = -q_k^{\nu}\widehat{v}_k + \frac{ik}{2}\sum_{|l|\le K, |k-l|\le K} \widehat{v}_l \widehat{v}_{k-l} + \widehat{f}_k(t),$$

$$+ \frac{ik}{2}\sum_{|l|>K \text{ or } |k-l|>K} \widehat{v}_l \widehat{v}_{k-l}$$

View $(\widehat{v}_{1:K}) \sim x$, $(\widehat{v}_{k>K}) \sim y$:

$$x' = f(x) + U(x,y), \; y' = g(x,y).$$

TODO: represent the effects of high modes to the low modes

Derivation of a parametric form (KSE)

Let $v = u + w$. In operator form: $v_t = Av + B(v)$,

$$\frac{du}{dt} = PAu + PB(u) + [PB(u+w) - PB(u)]$$
$$\frac{dw}{dt} = QAw + QB(u+w)$$

Nonlinear Galerkin: approximate inertial manifold (IM)[1]

- $\frac{dw}{dt} \approx 0 \Rightarrow w \approx A^{-1}QB(u+w) \Rightarrow w \approx \psi(u)$
- Need: spectral gap condition ✓;
- $dim = (u) > K$: parametrization with time delay (Lu-Lin17)

[1] Foias, Constantin, Temam, Sell, Jolly, Kevrekidis, Titi et al (88-94)

Derivation of a parametric form (KSE)

Let $v = u + w$. In operator form: $v_t = Av + B(v)$,

$$\frac{du}{dt} = PAu + PB(u) + [PB(u+w) - PB(u)]$$
$$\frac{dw}{dt} = QAw + QB(u+w)$$

Nonlinear Galerkin: approximate inertial manifold (IM)[1]

- $\frac{dw}{dt} \approx 0 \Rightarrow w \approx A^{-1}QB(u+w) \Rightarrow w \approx \psi(u)$
- Need: spectral gap condition ✓;
- $dim = (u) > K$: parametrization with time delay (Lu-Lin17)

A time series (NARMA) model of the form

$$u_k^n = R^\delta(u_k^{n-1}) + g_k^n + \Phi_k^n,$$

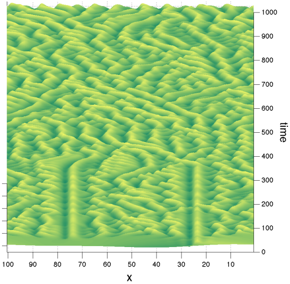
with $\Phi_k^n := \Phi_k^n(u^{n-p:n-1}, f^{n-p:n-1})$ in form of

$$\Phi_k^n = \sum_{j=1}^{p} c_{k,j}^v u_k^{n-j} + c_{k,j}^R R^\delta(u_k^{n-j}) + c_{k,j}^w \sum_{\substack{|k-l|\le K, K<|l|\le 2K \\ \text{or } |l|\le K, K<|k-l|\le 2K}} \widetilde{u}_l^{n-1}\widetilde{u}_{k-l}^{n-j}$$

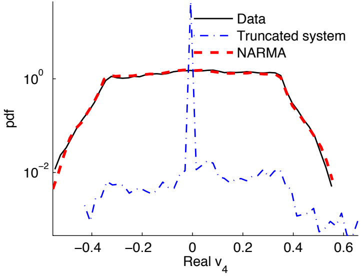
[1] Foias, Constantin, Temam, Sell, Jolly, Kevrekidis, Titi et al (88-94)

Test setting: $\nu = 3.43$
$N = 128$, $dt = 0.001$
Reduced model: $K = 5, \delta = 100dt$

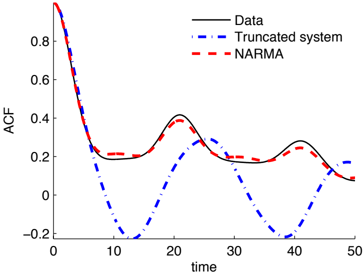
- 3 unstable modes
- 2 stable modes



Long-term statistics:



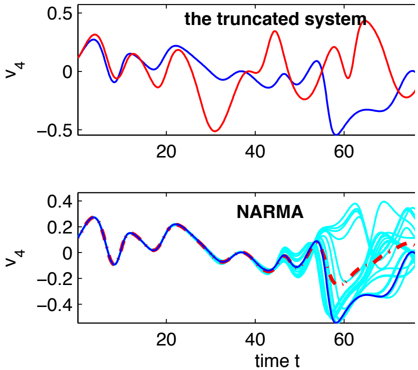
probability density function



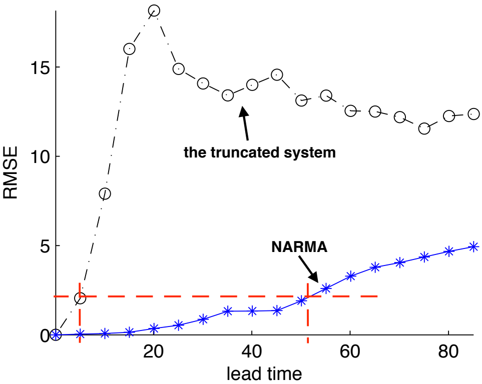
auto-correlation function

Prediction

A typical forecast:



RMSE of many forecasts:



Forecast time:

the truncated system: $T \approx 5$

the NARMA system: $T \approx \mathbf{50}$ ($\approx$ 2 Lyapunov time)

Derivation of a parametric form: stochastic Burgers

Let $v = u + w$. In operator form:

$$\frac{du}{dt} = PAu + PB(u) + Pf + [PB(u+w) - PB(u)]$$
$$\frac{dw}{dt} = QAw + QB(u+w)$$

- spectral gap: Burgers ? (likely not)
  $w(t)$ is not function of $u(t)$, but a functional of its path

Derivation of a parametric form: stochastic Burgers

Let $v = u + w$. In operator form:

$$\frac{du}{dt} = PAu + PB(u) + Pf + [PB(u+w) - PB(u)]$$
$$\frac{dw}{dt} = QAw + QB(u+w)$$

- spectral gap: Burgers ? (likely not)
  $w(t)$ is not function of $u(t)$, but a functional of its path

Integration instead:

$$w(t) = e^{-QAt}w(0) + \int_0^t e^{-QA(t-s)}[QB(u(s)+w(s))]ds$$
$$w^n \approx c_0 QB(u^n) + c_1 QB(u^{n-1}) + \cdots + c_p QB(u^{n-p})$$

Linear in parameter approximation:

$$PB(u+w) - PB(u) = P[(uw)_x + (u^2)_x]/2 \approx P[(uw)_x]/2 + noise$$
$$\approx \sum_{j=0}^{p} c_j P[(u^n QB(u^{n-j}))_x] + noise$$

A time series (NARMA) model of the form

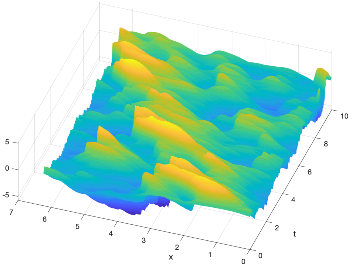
$$u_k^n = R^\delta(u_k^{n-1}) + f_k^n + g_k^n + \Phi_k^n,$$

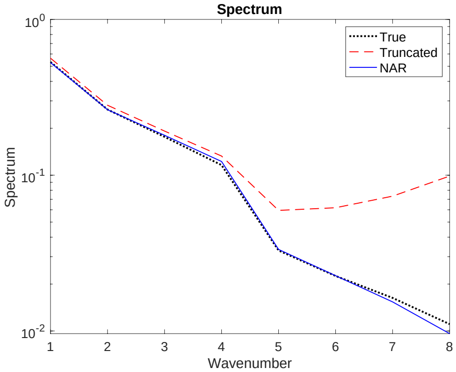
with $\Phi_k^n := \Phi_k^n(u^{n-p:n-1}, f^{n-p:n-1})$ in form of

$$\Phi_k^n = \sum_{j=1}^{p} c_{k,j}^v u_k^{n-j} + c_{k,j}^R R^\delta(u_k^{n-j}) + c_{k,j}^w \sum_{\substack{|k-l|\le K, K<|l|\le 2K \\ \text{or } |l|\le K, K<|k-l|\le 2K}} \widetilde{u}_l^{n-1} \widetilde{u}_{k-l}^{n-j}$$

Numerical tests:

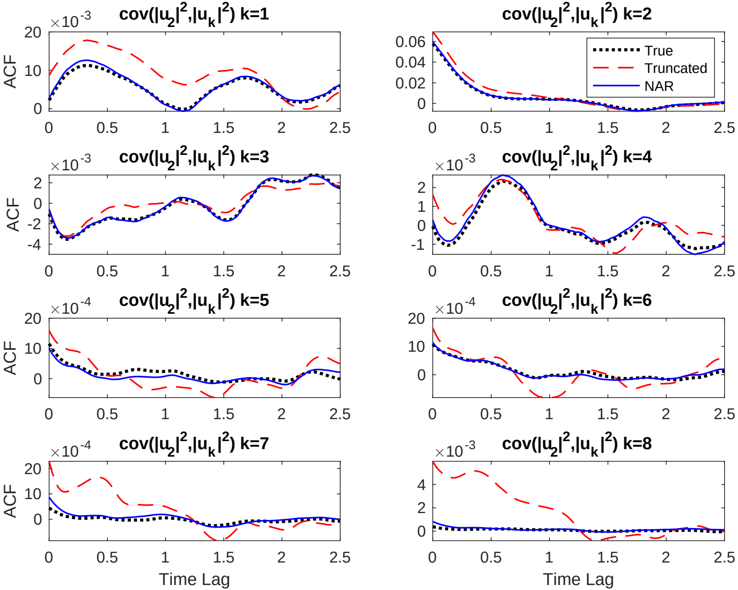
$\nu = 0.05$, $K_0 = 4 \rightarrow$ random shocks



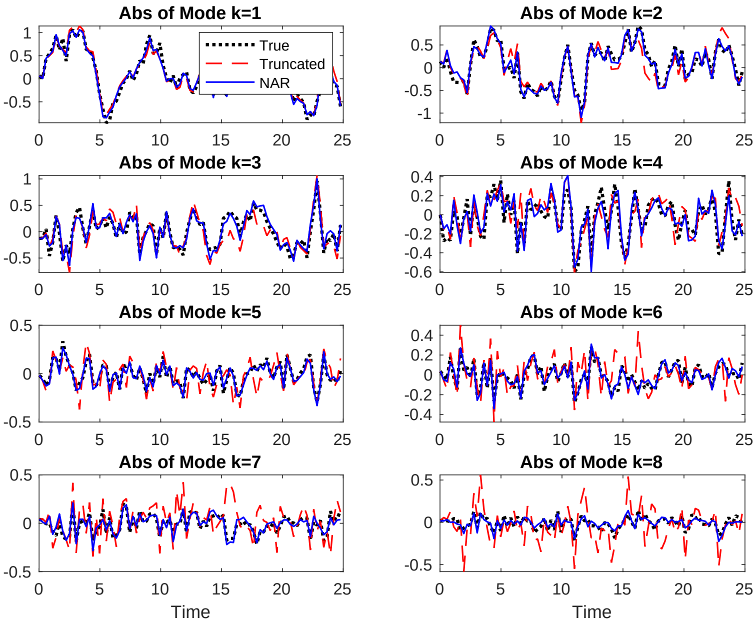


Energy spectrum

- Full model: $N = 128$, $dt = 0.005$
- Reduced model: $K = 8$, $\delta = 20dt$

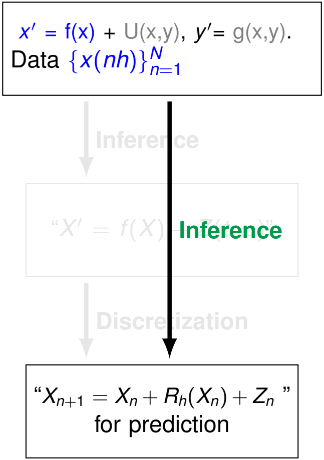


Cross-ACF of energy (4th moments!)



Trajectory prediction in response to force

Summary and ongoing work



Inference-based stochastic model reduction

- non-intrusive time series (**NARMA**)
- parametrize projections on path space

$\rightarrow$ Effective stochastic reduced model

Open problems:

- model reduction: model selection
- post-processing
- theoretical understanding of the approximation
  - distance between the two stochastic processes?

References

- Data-driven stochastic model reduction
  - Chorin-Lu: Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics. **PNAS 112 (2015)**, no. 32, 9804–9809.
  - Lu-Lin-Chorin: Comparison of continuous and discrete-time data-based modeling for hypoelliptic systems. **CAMCoS, 11 (2016)**, no. 8, 4227–4246.
  - Lu-Lin-Chorin: Data-based stochastic model reduction for the Kuramoto – Sivashinsky equation. **Physica D, 340 (2017)**, 46–57.
  - Lin-Lu: Data-driven model reduction, Wiener projections, and the Mori-Zwanzig formalism. preprint (2019)
- Data assimilation
  - Lu-Tu-Chorin: Accounting for model error from unresolved scales in EnKFs: improving the forecast model. **MWR, 340 (2017)**.

Thank you!

FL acknowledges supports from JHU, LBL, NSF