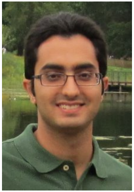
Stochastic Dynamical Modeling: Structured Matrix Completion of Partially Available Statistics

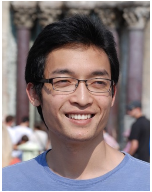
Mihailo Jovanović

www.umn.edu/∼mihailo

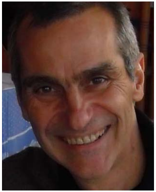
joint work with



Armin Zare



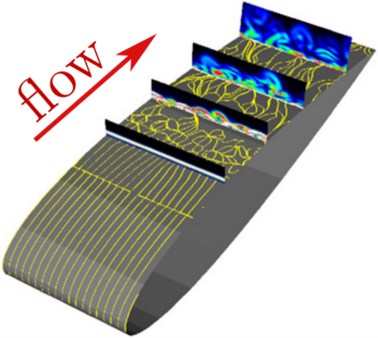
Yongxin Chen

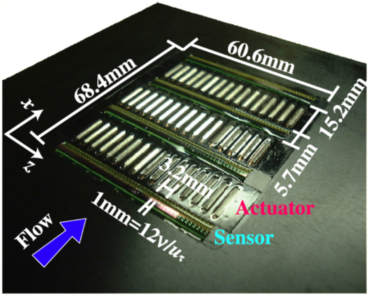


Tryphon Georgiou

IMA Workshop on Optimization and Parsimonious Modeling

Motivating application: flow control





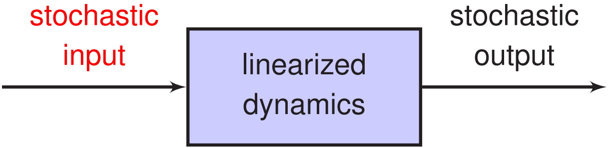
technology: shear-stress sensors; surface-deformation actuators

application: turbulence suppression; skin-friction drag reduction

challenge: distributed controller design for complex flow dynamics

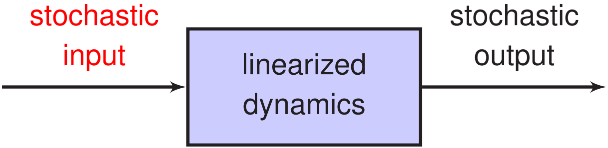
Control-oriented modeling

$$\begin{aligned} \dot{x} &= A\,x + {\color{red}B\,d} \\ y &= C\,x \end{aligned}$$



Control-oriented modeling

$$\begin{aligned}\dot{x} &= A\,x + B\,d\\ y &= C\,x\end{aligned}$$



- OBJECTIVE
  - ⋆ combine physics-based with data-driven modeling
  - ⋆ account for statistical signatures of dynamical systems using stochastically-forced linear models

- PROPOSED APPROACH
  - $\star$ view second-order statistics as data for an inverse problem

- KEY QUESTIONS
  - $\star$ Can we identify input statistics to reproduce available statistics?
  - $\star$ Can this be done by white in-time stochastic process?

- PROPOSED APPROACH
  - ⋆ view second-order statistics as data for an inverse problem

- KEY QUESTIONS
  - ⋆ Can we identify input statistics to reproduce available statistics?
  - ⋆ Can this be done by white in-time stochastic process?

OUR CONTRIBUTION

principled way of embedding statistics in control-oriented models

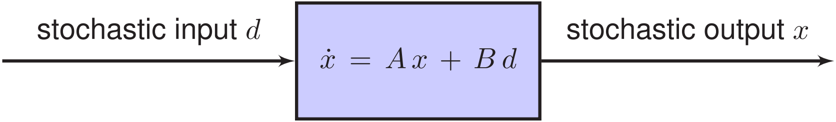
Response to stochastic inputs

stochastic input d $\longrightarrow$ $\boxed{\dot{x} \;=\; A\,x \,+\, B\,d}$ $\longrightarrow$ stochastic output x

- LYAPUNOV EQUATION
  - ⋆ propagates white correlation of d into colored statistics of x

$$A\,X \;+\; X A^* \;=\; -\,B\,W B^*$$

Response to stochastic inputs



- Lyapunov equation
  - ⋆ propagates white correlation of d into colored statistics of x

$$A\,X + X A^* = -\,B\,W B^*$$

  - ⋆ colored-in-time d

$$A\,X + X A^* = -\overbrace{(B\,H^* + H\,B^*)}^{Z}$$

white input: $H = (1/2)\,B\,W$

Georgiou, IEEE TAC '02

Lyapunov equation

discrete-time dynamics: $x_{t+1} \;=\; A\,x_t \;+\; B\,d_t$

white-in-time input: $\mathbf{E}\left(d_t\, d_\tau^*\right) \;=\; W\,\delta_{t\,-\,\tau}$

- LYAPUNOV EQUATION

$$
\begin{aligned}
X_{t+1} \;&:=\; \mathbf{E}\left(x_{t+1}\, x_{t+1}^*\right) \\
&=\; \mathbf{E}\Big(\left(A\,x_t \;+\; B\,d_t\right)\left(x_t^*A^* \;+\; d_t^*B^*\right)\Big) \\
&=\; A\,\mathbf{E}\left(x_t\, x_t^*\right)A^* \;+\; B\,\mathbf{E}\left(d_t\, d_t^*\right)B^* \\
&=\; A\,X_t\,A^* \;+\; B\,W\,B^*
\end{aligned}
$$

  ⋆ continuous-time version

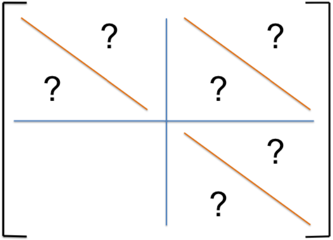
$$
\frac{\mathrm{d}\,X_t}{\mathrm{d}\,t} \;=\; A\,X_t \;+\; X_t\,A^* \;+\; B\,W\,B^*
$$

Outline

- STRUCTURED COVARIANCE COMPLETION PROBLEM
  - ⋆ embed available statistical features in control-oriented models
  - ⋆ complete unavailable data (via convex optimization)
- ALGORITHM
  - ⋆ Alternating Minimization Algorithm (AMA)
  - ⋆ works as proximal gradient on the dual problem
- CASE STUDY
  - ⋆ turbulent channel flow
  - ⋆ verification in linear stochastic simulations
- SUMMARY AND OUTLOOK

Problem setup

known elements of X



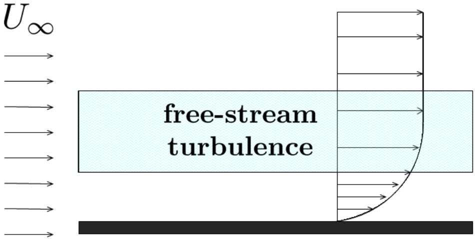
$$A\,X \;+\; X\,A^* \;=\; \underbrace{-\,(B\,H^* \;+\; H\,B^*)}_{Z}$$

- **PROBLEM DATA**
  - ⋆ system matrix A
  - ⋆ partially available entries of X

- **UNKNOWNS**
  - ⋆ missing entries of X
  - ⋆ disturbance dynamics Z $\begin{cases} \text{input matrix } B \\ \text{input power spectrum } H \end{cases}$

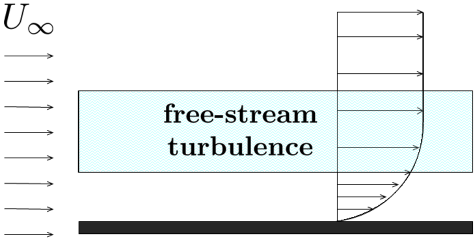
An example

- RESPONSE OF A BOUNDARY LAYER TO FREE-STREAM TURBULENCE



An example

- RESPONSE OF A BOUNDARY LAYER TO FREE-STREAM TURBULENCE



$$A\,X \;+\; X A^* \;=\; -\underbrace{(B\,H^* \;+\; H\,B^*)}_{Z}$$

number of input channels: limited by the rank of Z

Chen, Jovanović, Georgiou, IEEE CDC '13

Inverse problem

- CONVEX OPTIMIZATION PROBLEM

$$
\begin{array}{lll}
\underset{X,\,Z}{\text{minimize}} & -\log\det\left(X\right) \;+\; \gamma\,\|Z\|_* & \\
\text{subject to} & A\,X \;+\; X\,A^* \;+\; Z \;=\; 0 & \textbf{physics} \\
 & X_{ij} \;=\; G_{ij} \quad \text{for given } \; i,j & \textbf{available data}
\end{array}
$$

Inverse problem

- CONVEX OPTIMIZATION PROBLEM

$$\begin{array}{lll} \underset{X,\,Z}{\text{minimize}} & -\log\det(X) \;+\; \gamma\,\|Z\|_* & \\ \text{subject to} & A\,X \;+\; X\,A^* \;+\; Z \;=\; 0 & \textbf{physics} \\ & X_{ij} \;=\; G_{ij} \quad \text{for given } \; i,j & \textbf{available data} \end{array}$$

⋆ nuclear norm: proxy for rank minimization

$$\|Z\|_* \;:=\; \sum \sigma_i(Z)$$

Fazel, Boyd, Hindi, Recht, Parrilo, Candès, Chandrasekaran, . . .

Primal and dual problems

- PRIMAL

$$\begin{array}{ll} \underset{X,\,Z}{\text{minimize}} & -\log\det(X) \;+\; \gamma\,\|Z\|_* \\ \text{subject to} & \mathcal{A}\,X \;+\; \mathcal{B}\,Z \;-\; \mathcal{C} \;=\; 0 \end{array}$$

Primal and dual problems

- PRIMAL

$$\begin{array}{ll} \underset{X,\,Z}{\text{minimize}} & -\log\det(X) + \gamma\,\|Z\|_* \\ \text{subject to} & \mathcal{A}\,X + \mathcal{B}\,Z - \mathcal{C} = 0 \end{array}$$

- DUAL

$$\begin{array}{ll} \underset{Y_1,\,Y_2}{\text{maximize}} & \log\det(\mathcal{A}^\dagger\,Y) - \langle G, Y_2\rangle \\ \text{subject to} & \|Y_1\|_2 \le \gamma \end{array}$$

$$\mathcal{A}^\dagger \ - \ \text{adjoint of } \mathcal{A}; \qquad Y := \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

SDP characterization

$$Z = Z_+ - Z_-, \quad Z_+ \succeq 0, \quad Z_- \succeq 0$$

$$\begin{array}{ll} \underset{X,\, Z_+,\, Z_-}{\text{minimize}} & -\log\det(X) + \gamma \operatorname{trace}(Z_+ + Z_-) \\ \text{subject to} & \mathcal{A}\, X + \mathcal{B}\, Z - \mathcal{C} = 0 \\ & Z_+ \succeq 0, \quad Z_- \succeq 0 \end{array}$$

Customized algorithms

- ALTERNATING DIRECTION METHOD OF MULTIPLIERS (ADMM)

  Boyd et al., Found. Trends Mach. Learn. '11

- ALTERNATING MINIMIZATION ALGORITHM (AMA)

  Tseng, SIAM J. Control Optim. '91

Augmented Lagrangian

$$\mathcal{L}_{\rho}(X, Z; Y) = -\log\det(X) + \gamma \|Z\|_* + \langle Y, \mathcal{A}X + \mathcal{B}Z - \mathcal{C}\rangle$$
$$+ \frac{\rho}{2}\|\mathcal{A}X + \mathcal{B}Z - \mathcal{C}\|_F^2$$

Augmented Lagrangian

$$\mathcal{L}_\rho(X, Z; Y) \;=\; -\log\det(X) \;+\; \gamma\,\|Z\|_* \;+\; \langle Y, \mathcal{A}\,X \;+\; \mathcal{B}\,Z \;-\; \mathcal{C}\rangle$$
$$+\;\frac{\rho}{2}\,\|\mathcal{A}\,X \;+\; \mathcal{B}\,Z \;-\; \mathcal{C}\|_F^2$$

- METHOD OF MULTIPLIERS
  - minimizes $\mathcal{L}_\rho$ jointly over X and Z

$$\left(X^{k+1}, Z^{k+1}\right) \;:=\; \operatorname*{argmin}_{X, Z}\; \mathcal{L}_\rho(X, Z; Y^k)$$
$$Y^{k+1} \;:=\; Y^k \;+\; \rho\left(\mathcal{A}\,X^{k+1} \;+\; \mathcal{B}\,Z^{k+1} \;-\; \mathcal{C}\right)$$

ADMM vs AMA

- ADMM

$$
\begin{aligned}
X^{k+1} &:= \operatorname*{argmin}_{X} \; \mathcal{L}_{\rho}(X, Z^k; Y^k) \\
Z^{k+1} &:= \operatorname*{argmin}_{Z} \; \mathcal{L}_{\rho}(X^{k+1}, Z; Y^k) \\
Y^{k+1} &:= Y^k \; + \; \rho \left(\mathcal{A} X^{k+1} \; + \; \mathcal{B} Z^{k+1} \; - \; \mathcal{C} \right)
\end{aligned}
$$

ADMM vs AMA

- ADMM

$$
\begin{aligned}
X^{k+1} &:= \underset{X}{\operatorname{argmin}}\ \mathcal{L}_{\rho}(X, Z^k; Y^k) \\
Z^{k+1} &:= \underset{Z}{\operatorname{argmin}}\ \mathcal{L}_{\rho}(X^{k+1}, Z; Y^k) \\
Y^{k+1} &:= Y^k + \rho\left(\mathcal{A}\,X^{k+1} + \mathcal{B}\,Z^{k+1} - \mathcal{C}\right)
\end{aligned}
$$

- AMA

$$
\begin{aligned}
X^{k+1} &:= \underset{X}{\operatorname{argmin}}\ \mathcal{L}_{0}(X, Z^k; Y^k) \\
Z^{k+1} &:= \underset{Z}{\operatorname{argmin}}\ \mathcal{L}_{\rho}(X^{k+1}, Z; Y^k) \\
Y^{k+1} &:= Y^k + \rho_k\left(\mathcal{A}\,X^{k+1} + \mathcal{B}\,Z^{k+1} - \mathcal{C}\right)
\end{aligned}
$$

Z-update

$$\underset{Z}{\text{minimize}} \quad \gamma \, \|Z\|_* \; + \; \frac{\rho}{2} \, \|Z \; - \; V^k\|_F^2$$

$$\begin{aligned} V^k \quad &:= \quad -\left(\mathcal{A}_1 X^{k+1} \; + \; (1/\rho) \, Y_1^k\right) \\ &= \quad U \, \Sigma \, U^* \qquad\qquad\qquad\qquad \text{svd} \end{aligned}$$

Z-update

$$\underset{Z}{\text{minimize}} \quad \gamma \, \|Z\|_* \; + \; \frac{\rho}{2} \, \|Z \; - \; V^k\|_F^2$$

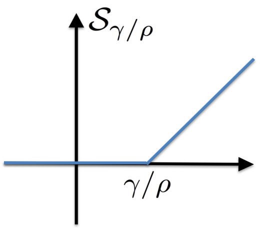
$$\begin{aligned} V^k \quad &:= \quad - \left(\mathcal{A}_1 X^{k+1} \; + \; (1/\rho) \, Y_1^k \right) \\ &= \quad U \, \Sigma \, U^* \end{aligned}$$

svd

singular value thresholding

$$Z^{k+1} \; = \; U \, \mathcal{S}_{\gamma/\rho}(\Sigma) \, U^*$$

soft-thresholding



complexity: $O(n^3)$

X-update in AMA

$$\underset{X}{\text{minimize}} \quad -\log\det\left(X\right) \;+\; \left\langle Y^k, \mathcal{A}\,X \right\rangle$$

explicit solution: $X^{k+1} \;=\; \left(\mathcal{A}^\dagger\, Y^k\right)^{-1}$

$\mathcal{A}^\dagger$ – adjoint of $\mathcal{A}$

complexity: $O(n^3)$

X-update in ADMM

$$\underset{X}{\text{minimize}} \quad -\log\det(X) \; + \; \frac{\rho}{2} \|\mathcal{A}\, X \; - \; U^k\|_F^2$$

optimality condition: $-X^{-1} \; + \; \rho\, \mathcal{A}^\dagger\big(\mathcal{A}\, X \; - \; U^k\big) \; = \; 0$

challenge: non-unitary $\mathcal{A}$

solution: proximal gradient algorithm

- Proximal algorithm
  - linearize $\dfrac{\rho}{2}\,\|\mathcal{A}\,X \;-\; U^k\|_F^2$ around X_i
  - add proximal term $\dfrac{\mu}{2}\,\|X \;-\; X_i\|_F^2$

optimality condition:

$$\boxed{\begin{aligned} \mu\,X \;-\; X^{-1} \;&=\; \left(\mu\,I - \rho\,\mathcal{A}^\dagger\mathcal{A}\right) X_i \;+\; \rho\,\mathcal{A}^\dagger\left(U^k\right) \\ &=\; V\,\Lambda\,V^* \end{aligned}}$$

- PROXIMAL ALGORITHM
  - ⋆ linearize $\frac{\rho}{2}\,\|\mathcal{A}\,X \;-\; U^k\|_F^2$ around X_i
  - ⋆ add proximal term $\frac{\mu}{2}\,\|X \;-\; X_i\|_F^2$

optimality condition:

$$\boxed{\begin{aligned} \mu\,X \;-\; X^{-1} \;&=\; \left(\mu\,I - \rho\,\mathcal{A}^\dagger\mathcal{A}\right)X_i \;+\; \rho\,\mathcal{A}^\dagger\left(U^k\right) \\ &=\; V\,\Lambda\,V^* \end{aligned}}$$

explicit solution: $X_{i+1} \;=\; V\,\mathrm{diag}\,(g)\,V^*$

$$g_j \;=\; \frac{\lambda_j}{2\mu} \;+\; \sqrt{\left(\frac{\lambda_j}{2\mu}\right)^2 \;+\; \frac{1}{\mu}}$$

complexity per iteration: $O(n^3)$

Y-update in AMA

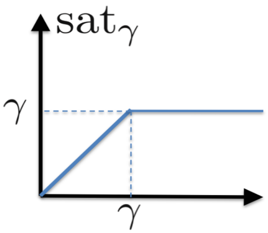
$$Y_1^{k+1} \;=\; \mathrm{sat}_\gamma\big(Y_1^k \,+\, \rho_k\,\mathcal{A}_1\,X^{k+1}\big) \qquad \longrightarrow \qquad \|Y_1\|_2 \;\leq\; \gamma$$

$$Y_2^{k+1} \;=\; Y_2^k \;+\; \rho_k\big(\mathcal{A}_2\,X^{k+1} \;-\; G\big)$$

Y-update in AMA

$$
\begin{aligned}
Y_1^{k+1} &= \text{sat}_\gamma\left(Y_1^k + \rho_k \, \mathcal{A}_1 \, X^{k+1}\right) \qquad \longrightarrow \qquad \|Y_1\|_2 \; \leq \; \gamma \\
Y_2^{k+1} &= Y_2^k + \rho_k\left(\mathcal{A}_2 \, X^{k+1} - G\right)
\end{aligned}
$$

saturation of singular values



saturation operator

$$
\text{sat}_\gamma(M) = M - \mathcal{S}_\gamma(M)
$$

Properties of AMA

- COVARIANCE COMPLETION VIA AMA
  - ⋆ proximal gradient on the dual problem
  - ⋆ sub-linear convergence with constant step-size

Properties of AMA

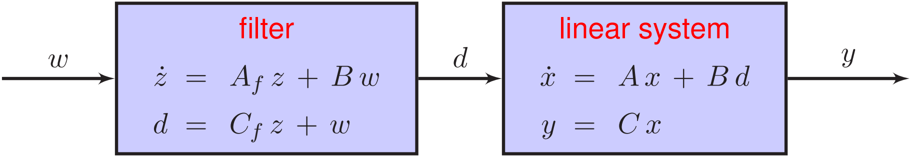
- Covariance completion via AMA
  - ⋆ proximal gradient on the dual problem
  - ⋆ sub-linear convergence with constant step-size

Step-size selection
  - ⋆ Barzilla-Borwein initialization followed by backtracking
  - ⋆ positive definiteness of X^{k+1}
  - ⋆ sufficient dual ascent

Dalal & Rajaratnam, arXiv:1405.3034

Zare, Chen, Jovanović, Georgiou, arXiv:1412.3399

Filter design



⋆ **white-in-time input**

$$\mathbf{E}\left(w(t_1)\, w^*(t_2)\right) \;=\; \Omega\, \delta(t_1 \,-\, t_2)$$

⋆ **filter dynamics**

$$\begin{aligned} A_f &= A \,+\, B\, C_f \\ C_f &= \left(H^* \,-\, \frac{1}{2}\, \Omega\, B^*\right) X^{-1} \end{aligned}$$

- LINEAR SYSTEM WITH FILTER

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} A & B\,C_f \\ 0 & A + B\,C_f \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} w$$

$$y = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}$$

- LINEAR SYSTEM WITH FILTER

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} A & B\,C_f \\ 0 & A + B\,C_f \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} w$$

$$y = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}$$

  - ⋆ coordinate transformation

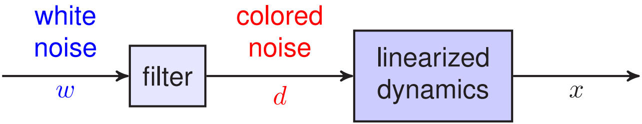
$$\begin{bmatrix} x \\ q \end{bmatrix} = \begin{bmatrix} I & 0 \\ -I & I \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}$$

  - ⋆ reduced-order representation

$$\begin{bmatrix} \dot{x} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} A + B\,C_f & B\,C_f \\ 0 & A \end{bmatrix} \begin{bmatrix} x \\ q \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} w$$

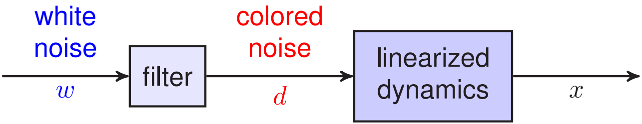
$$y = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ q \end{bmatrix}$$

Low-rank modification

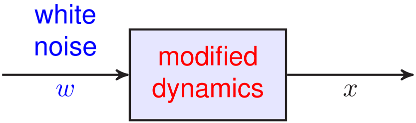


colored input: $\dot{x} \; = \; A\,x \; + \; B\,d$

Low-rank modification



colored input: $\dot{x} \;=\; A\,x \;+\; B\,d$



low-rank modification: $\dot{x} \;=\; (A \;+\; B\,C_f)\,x \;+\; B\,w$

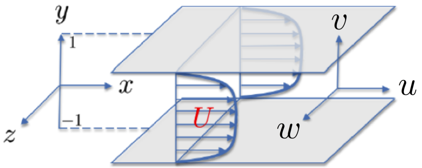
APPLICATION TO FLUIDS

please see Armin's poster for additional info

Turbulent channel flow

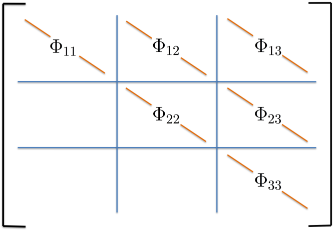
output covariance:

$$
\begin{aligned}
\Phi(\mathbf{k}) \;&:=\; \lim_{t \to \infty} \mathbf{E}\left(\mathbf{v}(t,\mathbf{k})\,\mathbf{v}^*(t,\mathbf{k})\right) \\
\mathbf{v} \;&=\; [\, u \;\; v \;\; w \,]^T \\
\mathbf{k} \;&-\; \text{horizontal wavenumbers}
\end{aligned}
$$



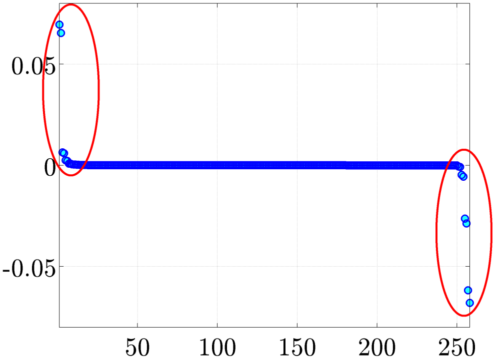
$$
A \;=\; \begin{bmatrix} A_{11} & 0 \\ A_{12} & A_{22} \end{bmatrix}
$$

known elements of $\Phi(\mathbf{k})$



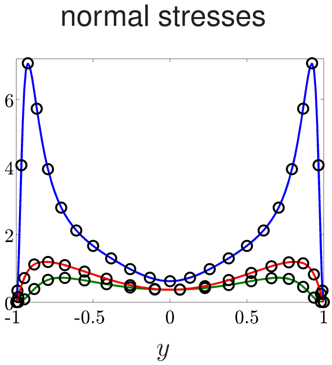
- KEY OBSERVATION
  - ⋆ white-in-time input: **too restrictive**

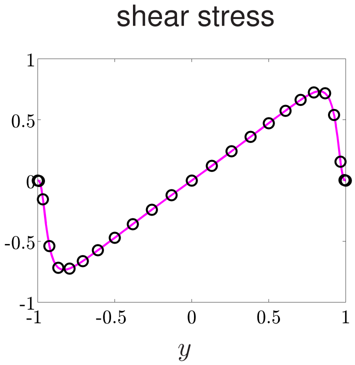
$$\lambda_i \left(A X_{\mathrm{ns}} \; + \; X_{\mathrm{ns}} A^* \right)$$



Jovanović & Georgiou, APS DFD '10

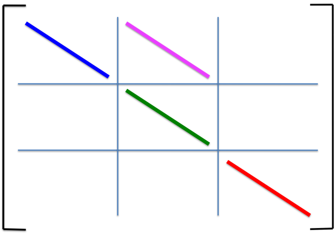
One-point correlations





Nonlinear simulations —

Solution to inverse problem ○



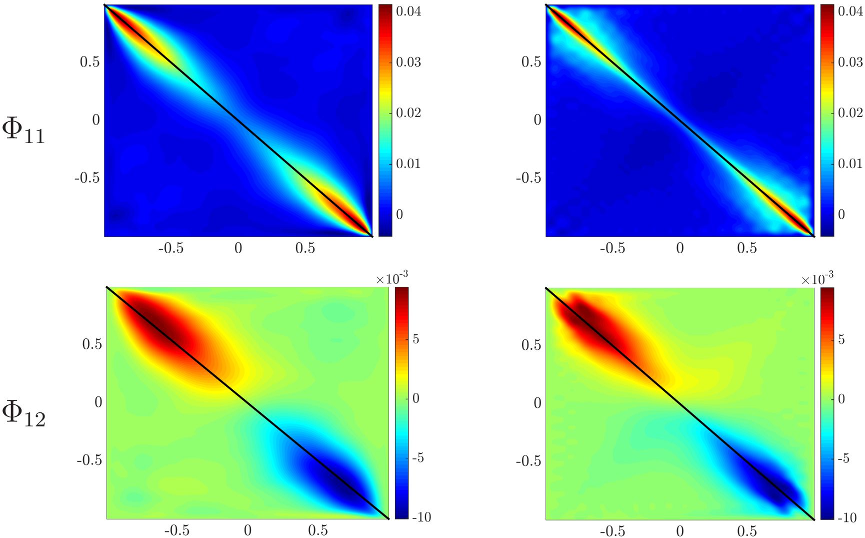
Importance of physics

- Covariance completion problem

$$
\begin{array}{lll}
\underset{X,Z}{\text{minimize}} & -\log\det(X) \,+\, \gamma\,\|Z\|_* & \\
\text{subject to} & A\,X \,+\, X\,A^* \,+\, Z \,=\, 0 & \textbf{physics} \\
 & (C\,X\,C^*)_{ij} \,=\, G_{ij} \quad \text{for given } \; i,j & \textbf{available data}
\end{array}
$$

Two-point correlations

nonlinear simulations | **covariance completion**



physics helps!

Challenges

- THEORETICAL
  - ⋆ conditions for exact recovery
  - ⋆ convergence rate of AMA with BB step-size initialization

- ALGORITHMIC
  - ⋆ alternative rank approximations
    (e.g., iterative re-weighting, matrix factorization)
  - ⋆ improving scalability

- APPLICATION
  - ⋆ development of turbulence closure models
  - ⋆ design of flow estimators/controllers

Summary

- CUSTOMIZED ALGORITHMS FOR COVARIANCE COMPLETION
  - ⋆ ADMM vs AMA
  - ⋆ AMA works as a proximal gradient on the dual problem
- THEORETICAL AND ALGORITHMIC DEVELOPMENTS
  - ⋆ *Chen, Jovanović, Georgiou, IEEE CDC '13*
  - ⋆ *Zare, Chen, Jovanović, Georgiou,* arXiv:1412.3399
- APPLICATION TO TURBULENT FLOWS
  - ⋆ *Zare, Jovanović, Georgiou, ACC '14*
  - ⋆ *Zare, Jovanović, Georgiou, 2014 Summer Program, CTR Stanford*