

Optimal sensor and actuator selection in distributed systems

Mihailo Jovanović

ee.usc.edu/mihailo

joint work with



Armin Zare

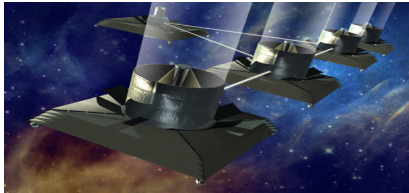


Neil Dhingra

IMA Sensor Location Workshop

Motivating applications

networks of dynamical systems



flexible wing aircraft



- **CHALLENGE:** sensor/actuator placement

Context

- RICH HISTORY

- ★ distributed parameter systems literature

John Burns' talk yesterday: outstanding overview!

- LESSONS LEARNED

- ★ importance of **problem formulation**
well-posedness; selection: context dependent
- ★ optimal estimation/control
much better tool for selection than **observability/controllability**
- ★ **difficult to solve**: nonconvex, computationally challenging

Context

- RICH HISTORY

- ★ distributed parameter systems literature

John Burns' talk yesterday: outstanding overview!

- LESSONS LEARNED

- ★ importance of **problem formulation**
well-posedness; selection: context dependent
- ★ optimal estimation/control
much better tool for selection than **observability/controllability**
- ★ **difficult to solve**: nonconvex, computationally challenging

- WHY NOW?

- ★ **applications**: networks, distributed sensor/actuator arrays
- ★ **optimization**: tremendous advances during the last decade

OBJECTIVE

select a subset of available sensors/actuators
that provides
“acceptable” degradation of estimation/control quality

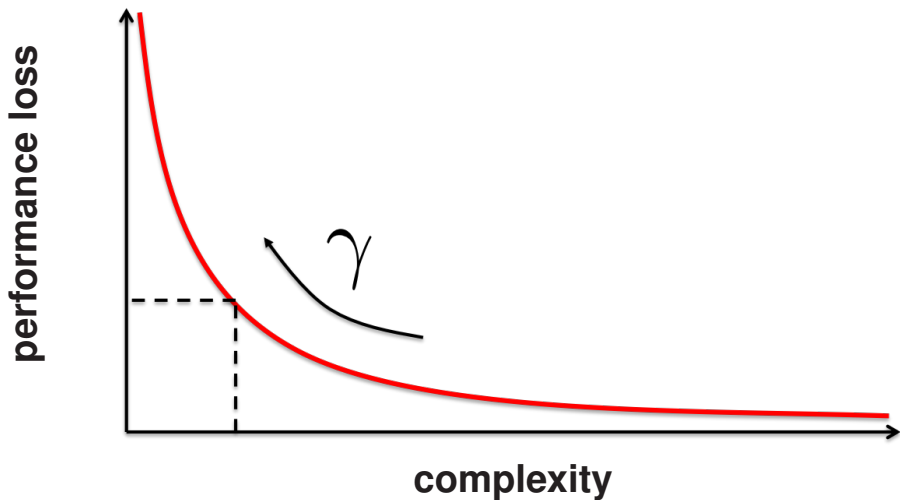
Selection via regularization

$$\begin{array}{ccc} \text{minimize} & J(K) & + \quad \gamma g(K) \\ & \downarrow & \downarrow \\ & \text{estimation/control} & \text{proxy for} \\ & \text{quality} & \text{selection} \end{array}$$

$\gamma > 0$ – performance vs “complexity” tradeoff

- TRADE-OFF CURVE

★ performance vs “complexity”



Minimum variance control problem

dynamics: $\dot{x} = Ax + B_1 d + B_2 u$

objective function: $J = \lim_{t \rightarrow \infty} \mathbf{E} (x^T(t) Q x(t) + u^T(t) R u(t))$

memoryless controller: $u = -F x$

Minimum variance control problem

dynamics: $\dot{x} = Ax + B_1 d + B_2 u$

objective function: $J = \lim_{t \rightarrow \infty} \mathbf{E} (x^T(t) Q x(t) + u^T(t) R u(t))$

memoryless controller: $u = -F x$

- CLOSED-LOOP VARIANCE AMPLIFICATION

J — **non-convex** function of F

No structural constraints

- SDP CHARACTERIZATION

$$\underset{X, F}{\text{minimize}} \quad \text{trace} \left((Q + F^T R F) X \right)$$

$$\begin{aligned} \text{subject to} \quad & (A - B_2 F) X + X (A - B_2 F)^T + B_1 B_1^T = 0 \\ & X \succ 0 \end{aligned}$$

No structural constraints

- SDP CHARACTERIZATION

$$\underset{X, F}{\text{minimize}} \quad \text{trace}((Q + F^T R F) X)$$

$$\text{subject to} \quad (A - B_2 F) X + X (A - B_2 F)^T + B_1 B_1^T = 0 \\ X \succ 0$$

★ **change of variables:** $FX = Y$

$$\underset{X, Y}{\text{minimize}} \quad \text{trace}(Q X) + \text{trace}(R Y X^{-1} Y^T)$$

$$\text{subject to} \quad (A X - B_2 Y) + (A X - B_2 Y)^T + B_1 B_1^T = 0 \\ X \succ 0$$

Schur complement \Rightarrow SDP characterization

- RICCATI-BASED-CHARACTERIZATION

globally optimal controller

$$A^T P + P A - P B_2 R^{-1} B_2^T P + Q = 0$$

$$F_c = R^{-1} B_2^T P$$

- STRUCTURAL CONSTRAINTS $F \in \mathcal{S}$

centralized

$$\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}$$

fully-decentralized

$$\begin{bmatrix} * & & & \\ & * & & \\ & & * & \\ & & & * \end{bmatrix}$$

localized

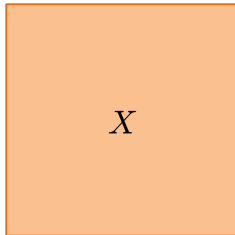
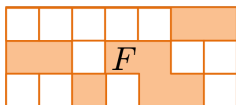
$$\begin{bmatrix} * & * & & \\ * & * & * & \\ & * & * & * \\ & & * & * \end{bmatrix}$$

GRAND CHALLENGE

convex characterization in the face of structural constraints

difficult to establish **relation between**

$\left\{ \begin{array}{c} \text{structural constraints} \\ \text{on } F \end{array} \right\}$ and $\left\{ \begin{array}{c} \text{structural constraints} \\ \text{on } X \text{ and } Y \end{array} \right\}$



$=$



Optimal actuator selection

- OBJECTIVE: identify **row-sparse** feedback gain

$$u = -Fx$$

minimize

$$J(F)$$

+

$$\gamma \sum_i \|e_i^T F\|_2$$

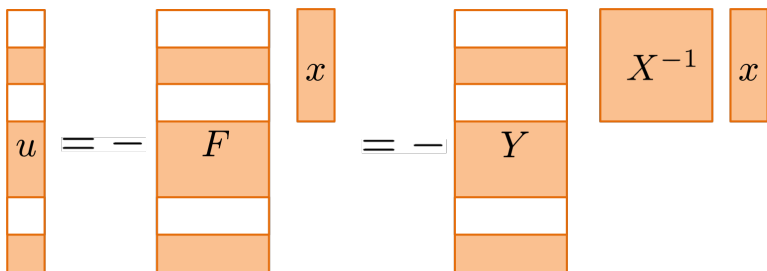


**variance
amplification**



**row-sparsity-promoting
penalty function**

- CHANGE OF VARIABLES: $Y := FX$
 - ★ **convex dependence** of J on X and Y
 - ★ **row-sparse structure preserved**



- OPTIMAL ACTUATOR SELECTION

- ★ admits SDP characterization

$$\begin{array}{ccc} \text{minimize} & J(X, Y) & + \quad \gamma \sum_i \|e_i^T Y\|_2 \\ & \downarrow & \downarrow \\ & \text{variance} & \text{row-sparsity-promoting} \\ & \text{amplification} & \text{penalty function} \end{array}$$

Polyak, Khlebnikov, Shcherbakov, ECC '13

Münz, Pfister, Wolfrum, IEEE TAC '14

Dhingra, Jovanović, Luo, CDC '14

Sensor selection: dual problem

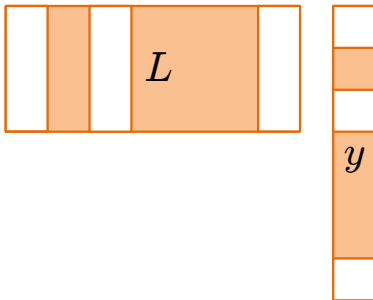
- KALMAN FILTER

- ★ minimum variance estimator

$$\dot{\hat{x}} = A \hat{x} + L(y - \hat{y}) + B d$$

$$\hat{y} = C \hat{x}$$

$$y = C x + w$$

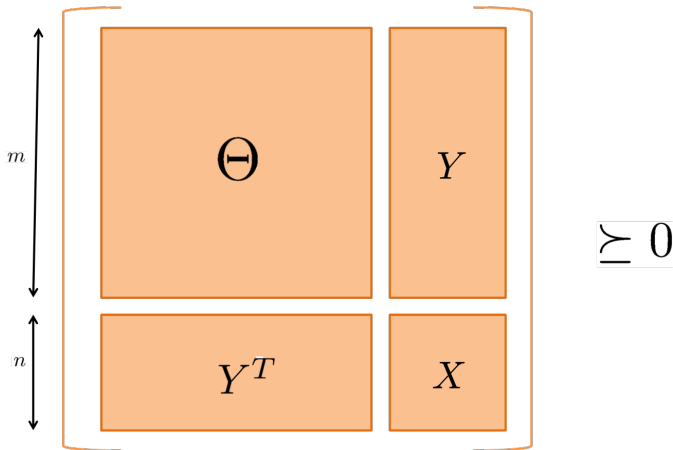


OBJECTIVE: minimize estimation error using a **few sensors**

- ★ proxy: **column sparsity** of Kalman gain L

Challenge: computational complexity

$$\text{trace}(R Y X^{-1} Y^T) = \text{trace}(R \Theta)$$



worst case complexity: $O((n+m)^6)$

Customized Algorithms

Actuator selection

$$\begin{aligned} & \underset{X, Y}{\text{minimize}} && J(X, Y) + \gamma g(Y) \\ & \text{subject to} && \mathcal{A}X - \mathcal{B}Y + W = 0 \\ & && X \succ 0 \end{aligned}$$

$$J(X, Y) := \text{trace} (Q X + R Y^T X^{-1} Y)$$

$$g(Y) := \sum_i \|e_i^T Y\|_2$$

$$\mathcal{A}X := A X + X A^T$$

$$\mathcal{B}Y := B_2 Y + Y^T B_2^T$$

$$W := B_1 B_1^T$$

Customized algorithms

- ALTERNATING DIRECTION METHOD OF MULTIPLIERS (ADMM)

Boyd et al., FnT in Machine Learning '11

- PROXIMAL GRADIENT ALGORITHM

Parikh & Boyd, FnT in Optimization '14

Two pillars

- AUGMENTED LAGRANGIAN

$$\mathcal{L}_{\rho}(X, Y; \Lambda) := J(X, Y) + \gamma g(Y) + \langle \Lambda, \mathcal{A}X - \mathcal{B}Y + W \rangle + \frac{\rho}{2} \|\mathcal{A}X - \mathcal{B}Y + W\|_F^2$$

Two pillars

- AUGMENTED LAGRANGIAN

$$\mathcal{L}_{\rho}(X, Y; \Lambda) := J(X, Y) + \gamma g(Y) + \langle \Lambda, \mathcal{A}X - \mathcal{B}Y + W \rangle + \frac{\rho}{2} \|\mathcal{A}X - \mathcal{B}Y + W\|_F^2$$

- PROXIMAL OPERATOR

$$\mathbf{prox}_{\mu g}(V) := \operatorname{argmin}_X g(X) + \frac{1}{2\mu} \|X - V\|_F^2$$

ADMM

$$X^{k+1} := \operatorname{argmin}_X \mathcal{L}_\rho(\textcolor{blue}{X}, Y^k; \Lambda^k)$$

$$Y^{k+1} := \operatorname{argmin}_Y \mathcal{L}_\rho(X^{k+1}, \textcolor{blue}{Y}; \Lambda^k)$$

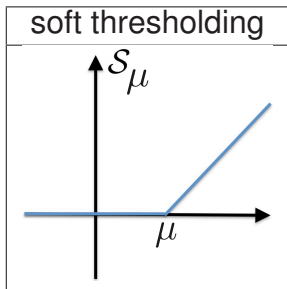
$$\Lambda^{k+1} := \Lambda^k + \rho (\mathcal{A} X^{k+1} - \mathcal{B} Y^{k+1} + W)$$

Y-update

$$\underset{\mathbf{Y}}{\text{minimize}} \quad \gamma \sum \| \mathbf{e}_i^T \mathbf{Y} \|_2 + \underbrace{\frac{\rho}{2} \| \mathcal{B} \mathbf{Y} - \mathbf{V} \|_F^2}_{h(\mathbf{Y})}$$

- GROUP LASSO

$$\mathbf{Y}^{j+1} = \text{prox}_{\gamma \alpha^j g}(\mathbf{Y}^j - \alpha^j \nabla h(\mathbf{Y}^j))$$



$$\mathbf{e}_i^T \mathbf{Y}^{j+1} = \mathcal{S}_{\gamma \alpha^j}(\mathbf{e}_i^T (\mathbf{Y}^j - \alpha^j \nabla h(\mathbf{Y}^j)))$$

complexity per inner iteration: $O(nm)$

X -update

$$\begin{aligned} & \underset{X}{\text{minimize}} \quad \text{trace} \left(X Q + X^{-1} Y^T R Y \right) + \frac{\rho}{2} \| \mathcal{A} X - U \|_F^2 \\ & \text{subject to} \quad X \succ 0 \end{aligned}$$

- CAN FORMULATE AS SDP

- ★ worst-case complexity $O(n^6)$

- PROJECTED NEWTON'S METHOD

- ★ use conjugate gradients to find the search direction
 - ★ project onto $\{X \mid X \succ 0\}$

worst-case complexity: $O(n^5)$

Dhingra, Jovanović, Luo, CDC '14

ADMM

- ★ difficult subproblems
- ★ slow overall convergence

ALTERNATIVE APPROACH

- ★ invertible \mathcal{A} : **avoid dualizing the linear constraint**

$$\mathcal{A}X - \mathcal{B}Y + W = 0$$

Elimination of X

- FOR INVERTIBLE \mathcal{A}

- ★ matrix A doesn't have e-values with equal positive and negative parts

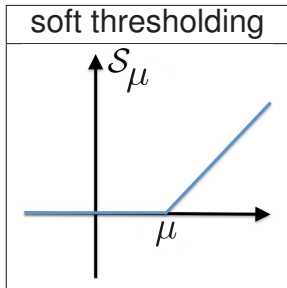
$$X(Y) = \mathcal{A}^{-1}(\mathcal{B}Y - W)$$

$$\begin{array}{ll} \underset{Y}{\text{minimize}} & J(Y) + \gamma g(Y) \\ \text{subject to} & X(Y) \succ 0 \end{array}$$

$$J(Y) := \text{trace} (Q X(Y) + R Y^T X^{-1}(Y) Y)$$

Proximal gradient method

$$Y^{k+1} := \text{prox}_{\gamma\alpha^k g}(Y^k - \alpha^k \nabla J(Y^k))$$



$$e_i^T Y^{k+1} = \mathcal{S}_{\gamma\alpha^k}(e_i^T(Y^k - \alpha^k \nabla J(Y^k)))$$

complexity: $O(\max(n^3, n^2m))$

- COMPLEXITY PER ITERATION

- ★ q backtracking steps: $O(q n^3)$

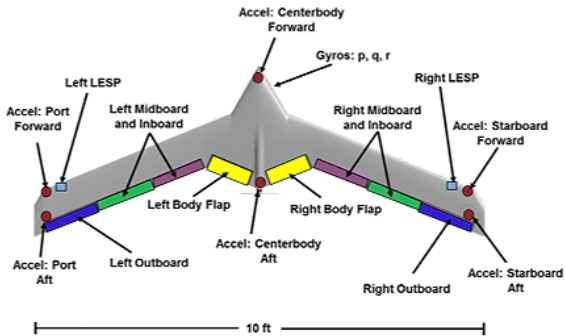
- STOPPING CRITERION

- ★ terminate when **relative and normalized residuals** are small

Goldstein, Studer, Baraniuk, arXiv:1411.3406

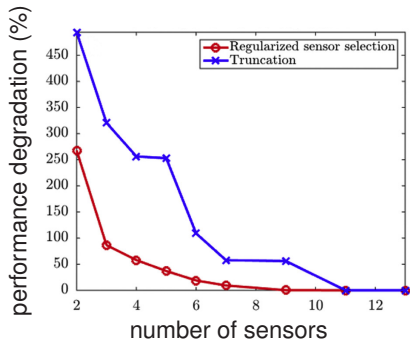
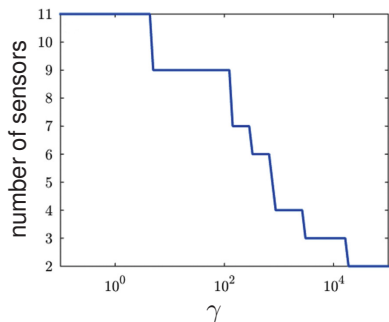
Examples

Flexible Wing Aircraft



- OBJECTIVE

- ★ detect aeroelastic instabilities



using half the sensors: degrades performance by $\approx 20\%$

Linearized Swift-Hohenberg equation

- PDE WITH SPATIALLY PERIODIC COEFFICIENTS

$$\partial_t \psi = -(\partial_{xx} + I)^2 \psi - c\psi + f \partial_x \psi + d + u$$

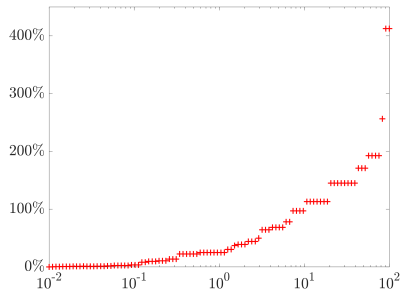
$$f(x) = \alpha \cos(\Omega x)$$

where

$$A = -(\partial_{xx} + I)^2 - cI + \alpha \cos(\Omega x) \partial_x$$

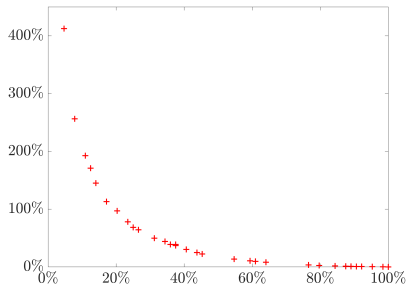
- $n = 64$; $c = -0.2$, $\alpha = 2$, $\Omega = 1.25$

$$(J - J_c)/J_c$$



γ

$$(J - J_c)/J_c$$



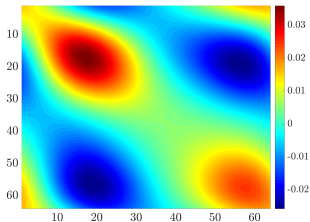
$\text{card}(F)/\text{card}(F_c)$

Structure of optimal controller

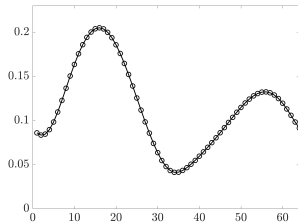
- $n = 64$; $c = -0.2$, $\alpha = 2$, $\Omega = 1.25$

feedback gain matrix

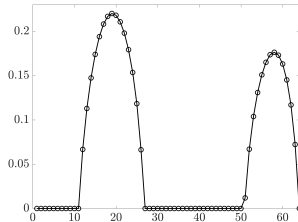
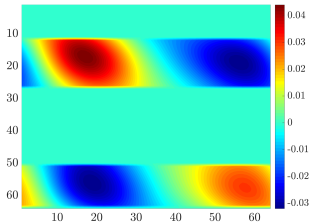
$\gamma = 0$



row norms

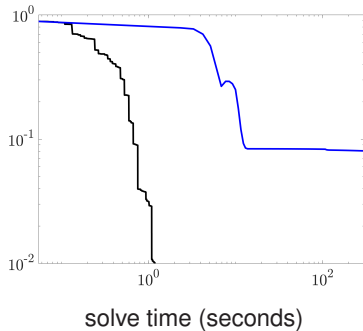
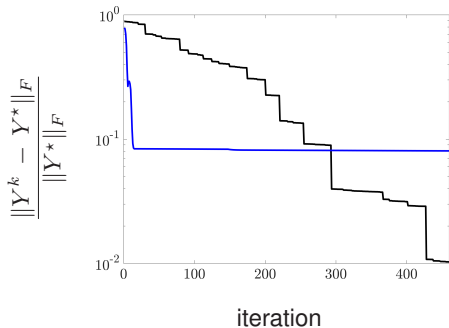


$\gamma = 0.675$



24.8% performance degradation

Comparison with ADMM



— Proximal gradient
— ADMM

Remarks

- CONVEX CHARACTERIZATION OF SENSOR/ACTUATOR SELECTION

Polyak, Khlebnikov, Shcherbakov, ECC '13

- ALTERNATING DIRECTION METHOD OF MULTIPLIERS

Dhingra, Jovanović, Luo, CDC '14

- PROXIMAL GRADIENT ALGORITHM

- ★ elimination of X
- ★ adaptive step-size selection

- RELATION TO MINIMUM ENERGY COVARIANCE COMPLETION PROBLEM

- ★ additional linear constraint on the covariance matrix X

Zare, Dhingra, Jovanović, Georgiou, CDC '17 (to appear)