Minimax Game-Theoretic Approach to Multiscale H-infinity Optimal Filtering

Hamza Anwar and Quanyan Zhu

Electrical and Computer Engineering Department
New York University

Presentation by Jeffrey Pawlick

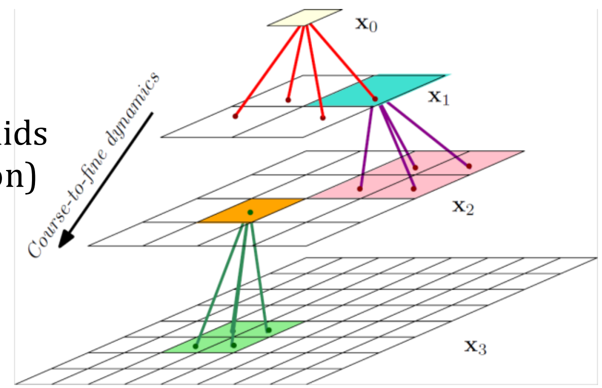
NYU | TANDON SCHOOL OF ENGINEERING

Outline

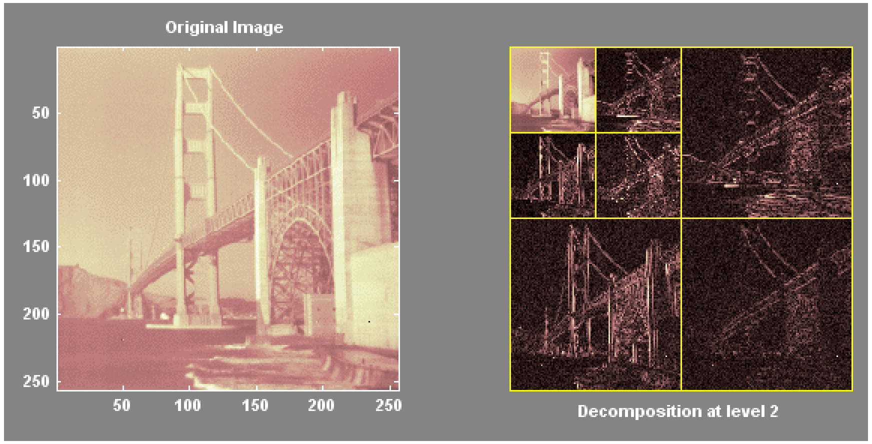
- Multiscale systems
- State-space dyadic tree model
- Estimator Design
- Multiscale H-infinity Filter
- Implementation Results
  - H-infinity vs Kalman Filter
  - Estimation Error
- Conclusion

Multiscale systems

Image pyramids
(compression)

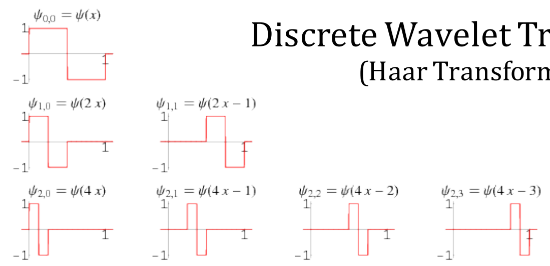


© Ruye Wang "Edge detection with image pyramid." Web Resource 2013-09-25. http://fourier.eng.hmc.edu/e161/lectures/canny/node2.html



© Sonka-Hlavac-Boyle: Image Processing, Analysis and Machine Vision, 3rd edition

Discrete Wavelet Transform
(Haar Transform)

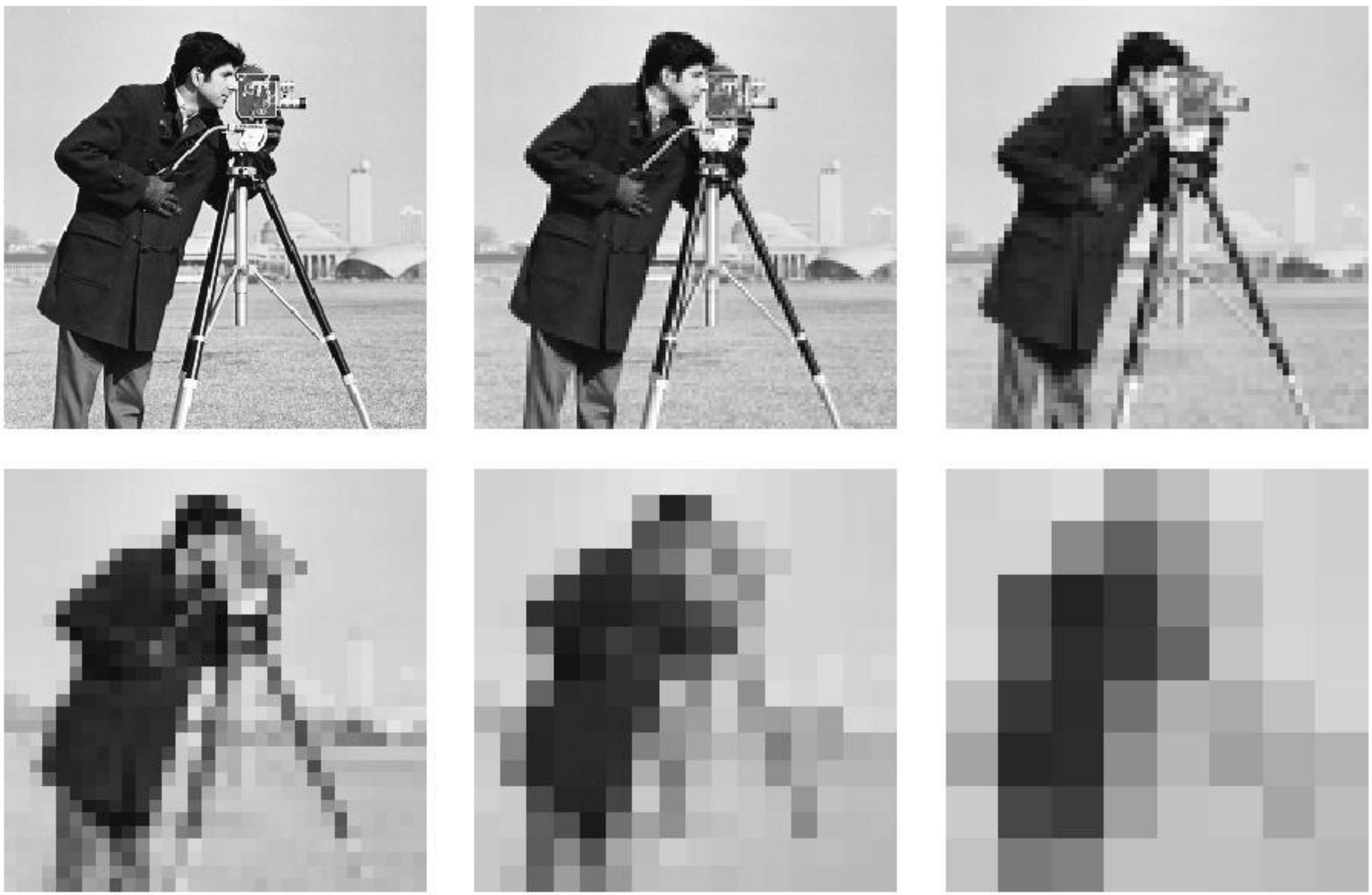


© Weisstein, Eric W. "Haar Function." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/HaarFunction.html

- Applications:
  - Data fusion in Demographics
  - Terrain Mapping using UAV
  - Video streaming/encoding
  - Efficiency in recursive parallelizable algorithms

For 2D images applications: just use ***quad*** instead of ***dyadic trees***

Multiresolution Signal Example

- 

State-Space Models on Dyadic Trees

Given a multiresolution representation of a signal, we relate levels of the tree through a state-space model

Forward System Dynamics:

$$x_{k+1}^{m\alpha} = A_{k+1}^{m\alpha} x_k^m + B_{k+1}^{m\alpha} w_{k+1}^{m\alpha}$$
$$x_{k+1}^{m\beta} = A_{k+1}^{m\beta} x_k^m + B_{k+1}^{m\beta} w_{k+1}^{m\beta}$$
$$y_k^m = C_k^m x_k^m + v_k^m$$

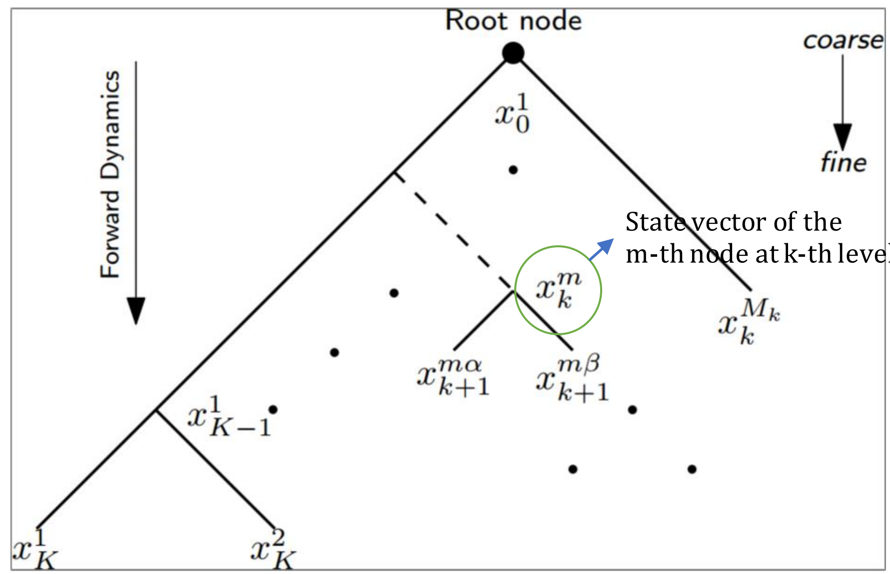
$$k = 0, 1, \dots, K; \ m = 1, 2, \dots, M_k$$

$$x_0^1 = x_0$$

$$w_k^m \sim N(0, I), \ v_k^m \sim N(0, R_k^m)$$

- Interpolation and higher resolution detail

Note: noise signals need to by l_2 bounded

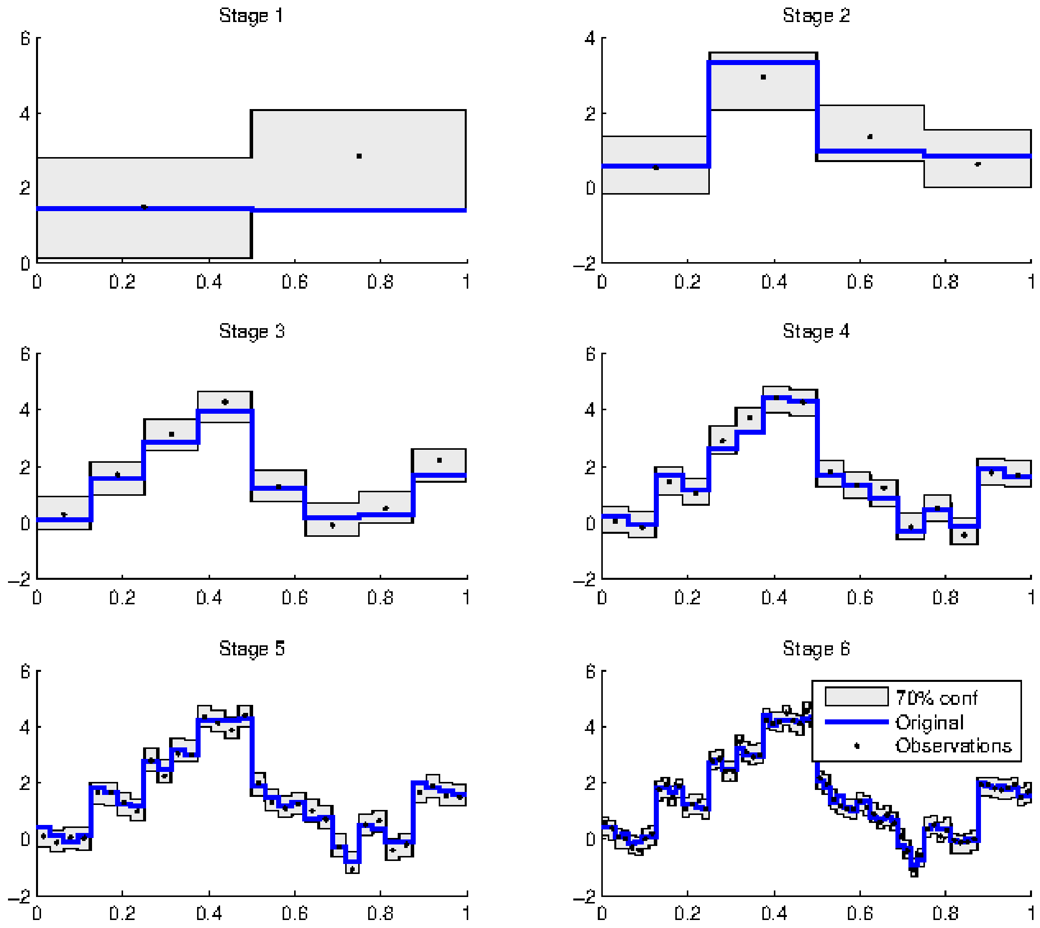


Depiction of the dyadic tree structure for a multiscale linear system

Example: Multiresolution Signal

Stage 0: $x_0^1 \sim N(0, p_0) \in \mathbf{R}$ $\quad A = B = C = 1$

$$w_k \sim N(0, 1/2^{K-k})$$
$$v_k \sim N(0, 1/(K-k))$$



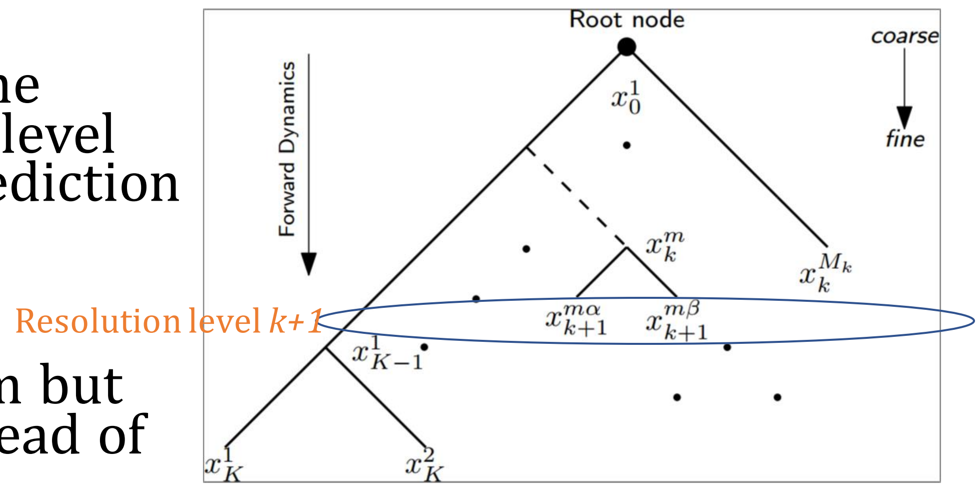
$$x_{k+1}^{m\alpha} = A_{k+1}^{m\alpha} x_k^m + B_{k+1}^{m\alpha} w_{k+1}^{m\alpha}$$
$$x_{k+1}^{m\beta} = A_{k+1}^{m\beta} x_k^m + B_{k+1}^{m\beta} w_{k+1}^{m\beta}$$
$$y_k^m = C_k^m x_k^m + v_k^m$$

Key features

- Coarse-to-fine recursion: Multiresolution synthesis of signals

$$x_{k+1}^{m\alpha} = A_{k+1}^{m\alpha} x_k^m + B_{k+1}^{m\alpha} w_{k+1}^{m\alpha}$$
$$x_{k+1}^{m\beta} = A_{k+1}^{m\beta} x_k^m + B_{k+1}^{m\beta} w_{k+1}^{m\beta}$$

- A resolution level captures the features of signals up to that level that are relevant for finer prediction downwards



- Similar to a dynamical system but state evolution is *spatial* instead of *temporal*

State-Estimation

H_∞ Estimator Design Problem for Multiscale Systems

Problem: Estimator Design

"Given exogenous noisy measurements at various resolution levels, design an optimal filter for state-estimation."

System:

$$
\begin{aligned}
x_{k+1}^{m\alpha} &= A_{k+1}^{m\alpha} x_k^m + B_{k+1}^{m\alpha} w_{k+1}^{m\alpha} \\
x_{k+1}^{m\beta} &= A_{k+1}^{m\beta} x_k^m + B_{k+1}^{m\beta} w_{k+1}^{m\beta} \\
y_k^m &= C_k^m x_k^m + v_k^m
\end{aligned}
\qquad (1)
$$

Estimate:

$$z_k^m = L_k^m x_k^m \qquad (2)$$

Given attenuation level, minimize the supremum of quadratic cost subject to (1)

$$J = \frac{\sum_{k=1}^{K-1} \sum_{m=1}^{M_k} ||z_k^m - \hat{z}_k^m||^2_{Q_k^m}}{||x_0 - \hat{x}_0||^2_{p_0^{-1}} + \sum_{k=1}^{K-1} \sum_{m=1}^{M_k} \{||w_{k+1}^{m\alpha}||_I^2 + ||w_{k+1}^{m\beta}||_I^2 + ||v_k^m||^2_{R_k^{m^{-1}}}\}};$$

$\sup J < 1/\gamma.$

Solution Methodology

- Equivalent Minimax problem

$$\min_{\hat{x}_k^m} \max_{(y_k^m, w_k^m, x_0)} J = \frac{1}{2} \sum_{k=1}^{K-1} \sum_{m=1}^{M_k} \{ ||x_k^m - \hat{x}_k^m||^2_{\bar{Q}_k^m} - \frac{1}{\gamma} (||w_{k+1}^{m\alpha}||_I^2$$

$$+ ||w_{k+1}^{m\beta}||_I^2 + ||y_k^m - C_k^m x_k^m||^2_{R_k^{m-1}}) \} - \frac{1}{2\gamma} ||x_0 - \hat{x}_0||^2_{p_0^{-1}}$$

subject to (1) and (2) where, $\bar{Q}_k^m = L_k^{m^T} Q_k^m L_k^m$

- First, solve for optimal x_0 and w_k by constructing Hamiltonian (Lagrange multipliers) and using Maximum Principle
- Solution assumption:

$$x_k^{m*} = \bar{x}_k^m + P_k^m \lambda_k^{m*}. \text{ where } \lambda_k^{m*} \text{ are the Lagrange multipliers}$$

Estimator Design

- This gives optimal values for x_0 and w_k

$$\begin{bmatrix} w_{k+1}^{m\alpha *} \\ w_{k+1}^{m\beta *} \end{bmatrix} = \begin{bmatrix} B_{k+1}^{m\alpha^T} & \\ & B_{k+1}^{m\beta^T} \end{bmatrix} \begin{bmatrix} \lambda_{k+1}^{m\alpha *} \\ \lambda_{k+1}^{m\beta *} \end{bmatrix}, \ \ x_0^* = \hat{x}_0 + p_0 \lambda_0^*.$$

- Simplifying the cost with above optimal values reduces the problem to

$$\min_{\hat{x}_k^m} \max_{y_k^m} J = \frac{1}{2} \sum_{k=1}^{K-1} \sum_{m=1}^{M_k} [||\bar{x}_k^m - \hat{x}_k^m||^2_{\bar{Q}_k^m} - \frac{1}{\gamma}(||y_k^m - C_k^m \bar{x}_k^m||^2_{R_k^{m^{-1}}})]$$

subject to two coupled equations for two children nodes

Coupling Terms in RHSs

$$\bar{x}_{k+1}^{m\alpha} - A_{k+1}^{m\alpha} \bar{x}_k^m - A_{k+1}^{m\alpha} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} [\gamma \bar{Q}_k^m (\bar{x}_k^m - \hat{x}_k^m) + C_k^{m^T} R_k^{m^{-1}} (y_k^m - C_k^m \bar{x}_k^m)] =$$

$$[-P_{k+1}^{m\alpha} \lambda_{k+1}^{m\alpha *} + A_{k+1}^{m\alpha} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} [A_{k+1}^{m\alpha^T} \lambda_{k+1}^{m\alpha *} + A_{k+1}^{m\beta^T} \lambda_{k+1}^{m\beta *}] + B_{k+1}^{m\alpha} B_{k+1}^{m\alpha^T} \lambda_{k+1}^{m\alpha *} = 0,$$

$$\bar{x}_{k+1}^{m\beta} - A_{k+1}^{m\beta} \bar{x}_k^m - A_{k+1}^{m\beta} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} [\gamma \bar{Q}_k^m (\bar{x}_k^m - \hat{x}_k^m) + C_k^{m^T} R_k^{m^{-1}} (y_k^m - C_k^m \bar{x}_k^m)] =$$

$$[-P_{k+1}^{m\beta} \lambda_{k+1}^{m\beta *} + A_{k+1}^{m\beta} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} [A_{k+1}^{m\alpha^T} \lambda_{k+1}^{m\alpha *} + A_{k+1}^{m\beta^T} \lambda_{k+1}^{m\beta *}] + B_{k+1}^{m\beta} B_{k+1}^{m\beta^T} \lambda_{k+1}^{m\beta *} = 0.$$

Estimator Design (contd.)

- Coupled equality constraints are expressed as a matrix equation

$$\min_{\hat{x}_k^m} \max_{y_k^m} J = \frac{1}{2} \sum_{k=1}^{K-1} \sum_{m=1}^{M_k} \left[||\bar{x}_k^m - \hat{x}_k^m||^2_{\bar{Q}_k^m} - \frac{1}{\gamma} (||y_k^m - C_k^m \bar{x}_k^m||^2_{R_k^{m^{-1}}}) \right]$$

$$\text{subject to} \quad \begin{bmatrix} P_{k+1}^{m\alpha} & \\ & P_{k+1}^{m\beta} \end{bmatrix} \begin{bmatrix} \lambda_{k+1}^{m\alpha *} \\ \lambda_{k+1}^{m\beta *} \end{bmatrix} = \begin{bmatrix} *_1 & *_2 \\ *_2^T & *_3 \end{bmatrix} \begin{bmatrix} \lambda_{k+1}^{m\alpha *} \\ \lambda_{k+1}^{m\beta *} \end{bmatrix}$$

- Must hold for arbitrary lambda, so we arrive at matrix equality constraint:

$$\begin{bmatrix} P_{k+1}^{m\alpha} - *_1 \\ P_{k+1}^{m\beta} - *_3 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

- This gives us a **two-person zero-sum game** having unique saddle-point equilibrium strategy

$$\implies \hat{x}_k^{m*} = \bar{x}_k^m, \; y_k^{m*} = C_k^m \bar{x}_k^m$$

Estimator Design (contd.)

- The optimal strategy for measurement noise follows directly

$$v_k^{m*} = y_k^{m*} - C_k^m \hat{x}_k^{m*} = C_k^m \bar{x}_k^m - C_k^m \bar{x}_k^{m*} = 0.$$

- Hence, the optimal H-infinity filter is given as:

$$\hat{z}_k^{m*} = L_k^m \hat{x}_k^{m*}; \quad k = 0 \ldots K,\ m = 1 \ldots M_k$$

where

$$\begin{aligned}
\hat{x}_{k+1}^{m\alpha*} &= A_{k+1}^{m\alpha} \hat{x}_k^m + K_{k+1}^{m\alpha}(y_k^m - C_k^m \hat{x}_k^{m*}),\\
\hat{x}_{k+1}^{m\beta*} &= A_{k+1}^{m\beta} \hat{x}_k^m + K_{k+1}^{m\beta}(y_k^m - C_k^m \hat{x}_k^{m*}),\\
K_{k+1}^{m\alpha} &= A_{k+1}^{m\alpha} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} C_k^{m^T} R_k^{m^{-1}},\\
K_{k+1}^{m\beta} &= A_{k+1}^{m\beta} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} C_k^{m^T} R_k^{m^{-1}}.
\end{aligned}$$

- And P_k^m is determined by a Joint Difference Riccati Equation

Theorem:

For noise attenuation $\gamma > 0$, an H_∞ filter for x_k^m exists if and only if there exists a stabilizing solution $P_k^m > 0\ \forall k, m$ to the coupled-pair of discrete-time Riccati equations:

$$P_{k+1}^{m\alpha} = A_{k+1}^{m\alpha} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} A_{k+1}^{m\alpha^T} + B_{k+1}^{m\alpha} B_{k+1}^{m\alpha^T},$$
$$P_{k+1}^{m\beta} = A_{k+1}^{m\beta} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} A_{k+1}^{m\beta^T} + B_{k+1}^{m\beta} B_{k+1}^{m\beta^T},$$
$$P_0 = p_0.$$

The H_∞ filter is given by,

$$\hat{x}_{k+1}^{m\alpha *} = A_{k+1}^{m\alpha} \hat{x}_k^m + K_{k+1}^{m\alpha}(y_k^m - C_k^m \hat{x}_k^{m*}),$$
$$\hat{x}_{k+1}^{m\beta *} = A_{k+1}^{m\beta} \hat{x}_k^m + K_{k+1}^{m\beta}(y_k^m - C_k^m \hat{x}_k^{m*}),$$
$$K_{k+1}^{m\alpha} = A_{k+1}^{m\alpha} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} C_k^{m^T} R_k^{m^{-1}},$$
$$K_{k+1}^{m\beta} = A_{k+1}^{m\beta} P_k^m (I - \gamma \bar{Q}_k^m P_k^m + C_k^{m^T} R_k^{m^{-1}} C_k^m P_k^m)^{-1} C_k^{m^T} R_k^{m^{-1}}.$$

- A special case is when $\gamma \to 0$ the filter reduces to standard Kalman filter.

Proof was just described earlier.
Constructed on the same lines as: Shen, Xuemin, and Li Deng. "Game theory approach to discrete H∞ filter design." Signal Processing, IEEE Transactions on 45.4 (1997): 1092-1095.

Current H-infinity Estimator [2]

- To take into account measurements at current stage
- For a simple linear system, the current H-infinity estimator is given as:

System:-

$$x_{k+1} = A_k x_k + B_k w_k$$
$$y_k = C_k x_k + v_k$$
$$z_k = L_k x_k$$

State estimate error:-

$$e_k = \hat{z}_{k|k} - L_k x_k$$

Note that $\mathbf{z_{k|k-1}}$ *is replaced by* $\mathbf{z_{k|k}}$ *here.*
The rest of cost function is the same

Filter:-

$$\hat{z}_k = L_k \hat{x}_k$$
$$\hat{x}_{k|k-1} = A_k \hat{x}_{k-1|k-1}$$
$$P_{k|k-1} = A_{k-1} P_{k-1|k-1} A_{k-1}^T + B_{k-1} B_{k-1}^T$$
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - C x_{k|k-1})$$
$$K_k = P_{k|k-1} C_k^T (R + C_k P_{k|k-1} C_k^T)^{-1}$$
$$P_{k|k} = P_{k|k-1} - P_{k|k-1} \begin{bmatrix} C_k^T & L_k^T \end{bmatrix} R_{e,k}^{-1} \begin{bmatrix} C_k \\ L_k \end{bmatrix} P_{k|k-1}$$
$$R_{e,k} = \begin{bmatrix} R & 0 \\ 0 & -I/\gamma \end{bmatrix} + \begin{bmatrix} C_k \\ L_k \end{bmatrix} P_{k|k-1} \begin{bmatrix} C_k^T & L_k^T \end{bmatrix}$$

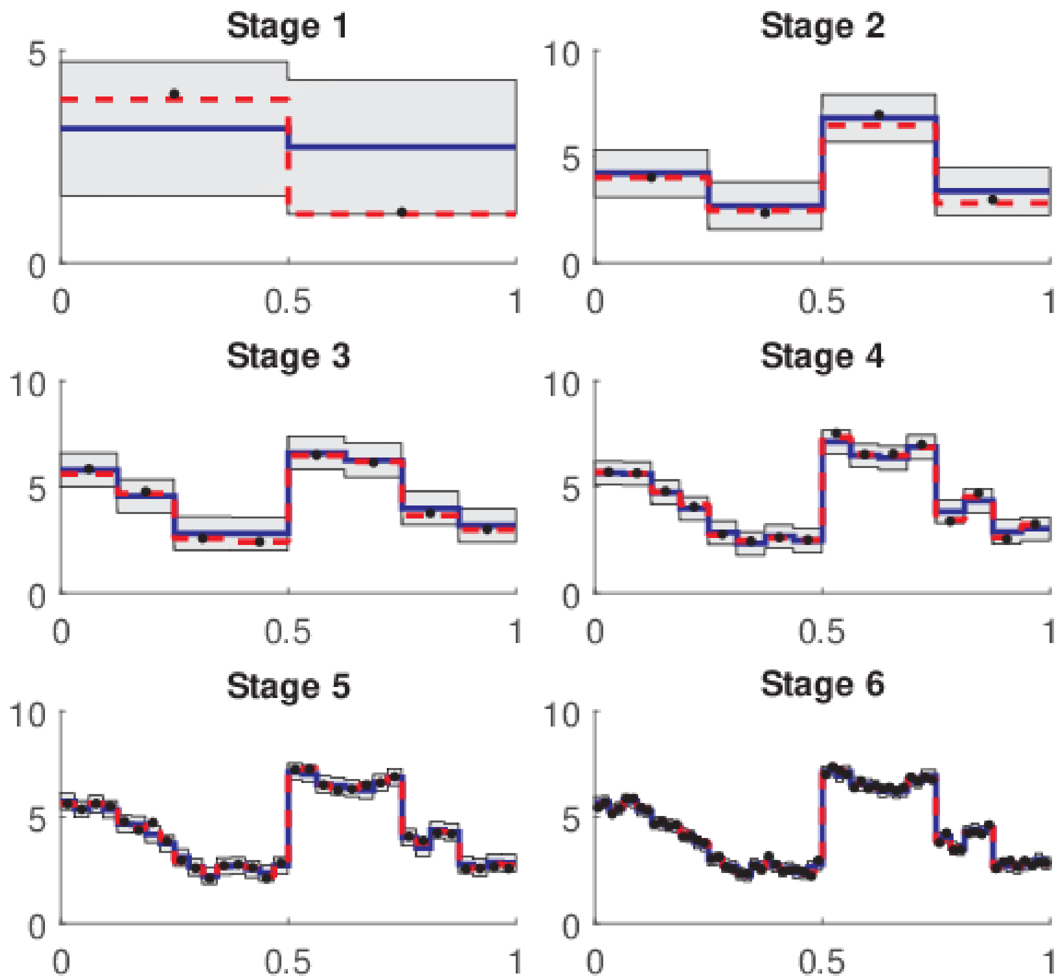
Combining the above result with predictor-based H-infinity design proposed earlier, gives the current-H-infinity-estimator – completely written in paper.
Note: For simplicity, we've avoided writing it down here.

[2] Green, Michael, and David J N Limebeer. *Linear Robust Control*. Courier Corporation, 2012.

Results and Comparison

- Kalman vs. H-infinity Filter
- Estimation error trend along stages

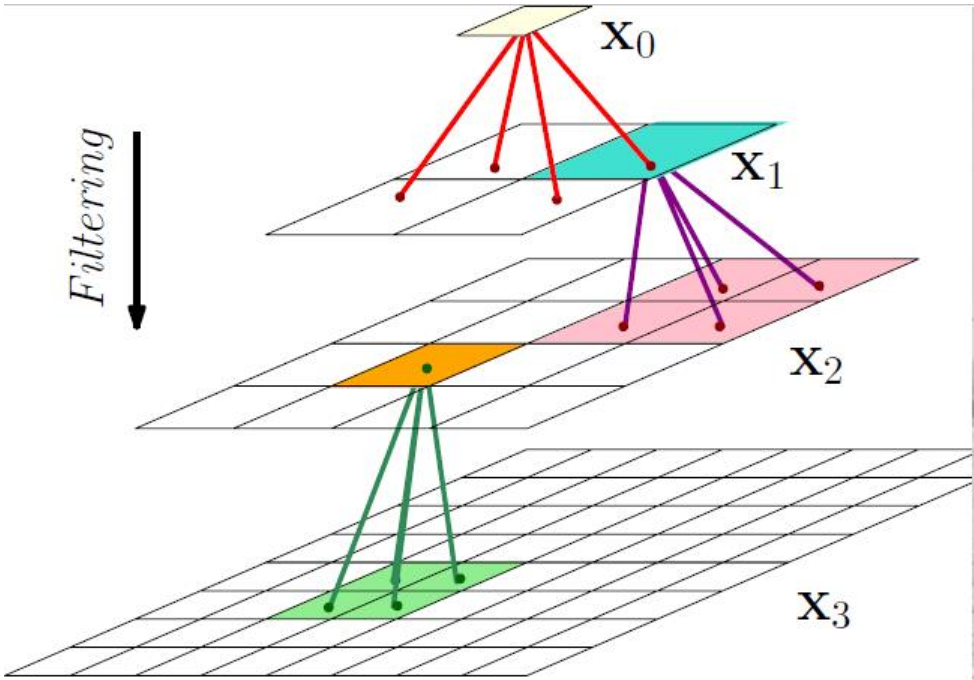
1-D Multiscale H-infinity filter



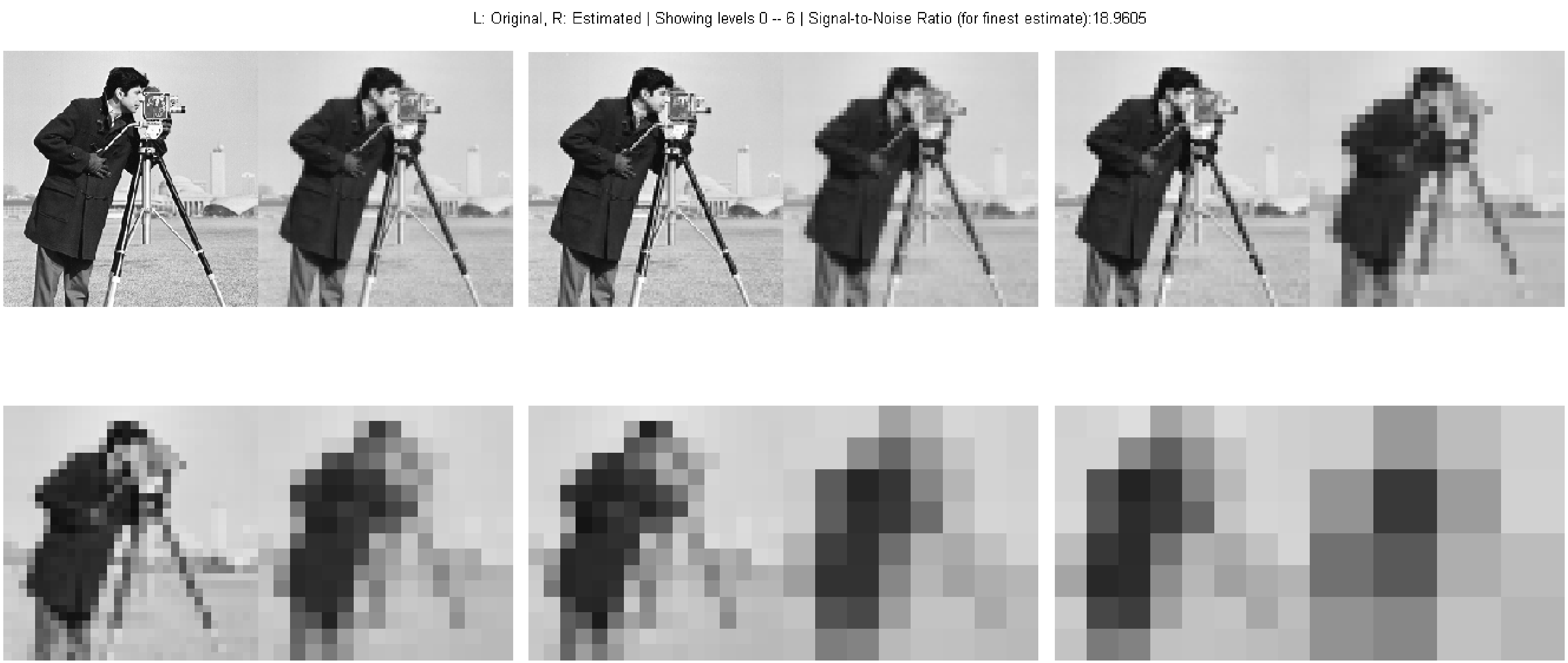
Gray 70% confidence bounds indicate the uncertainty in state transitions. Observations are indicated by black dots, while the original signal is in blue. State estimates recovered by proposed filter are indicated by red dotted line.

Working with Images: The Signal

Working with Images: The Model



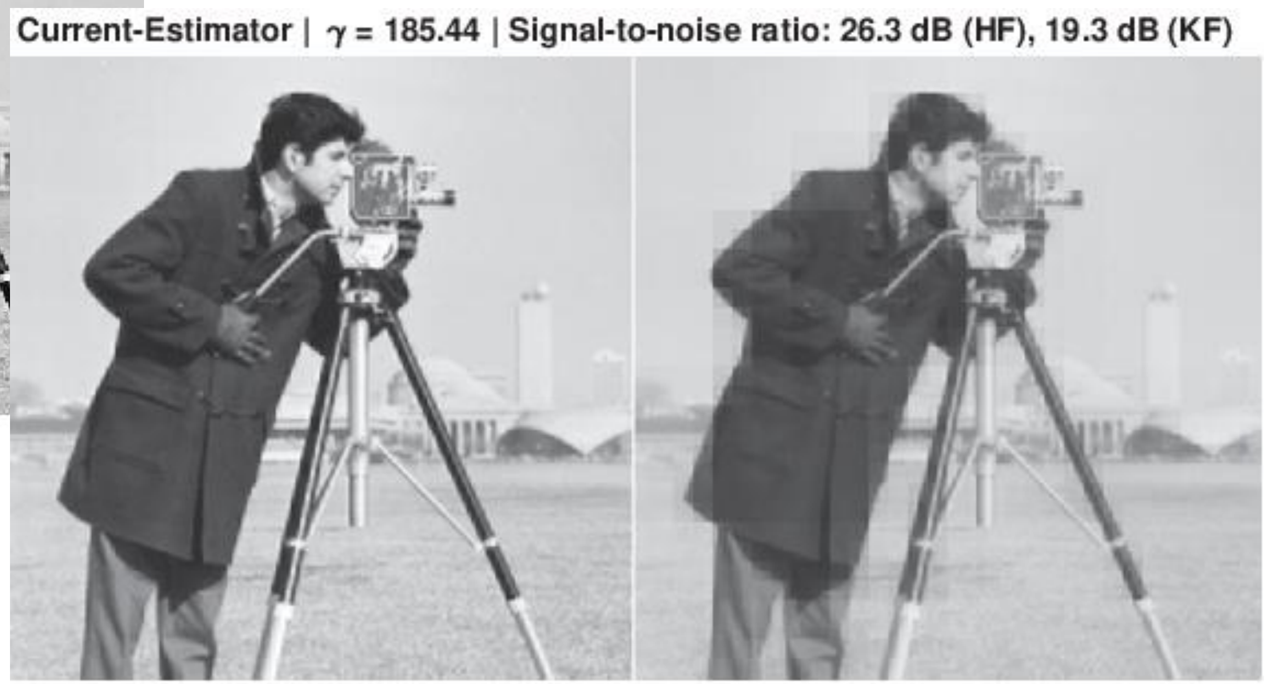
Working with Images: The Estimation



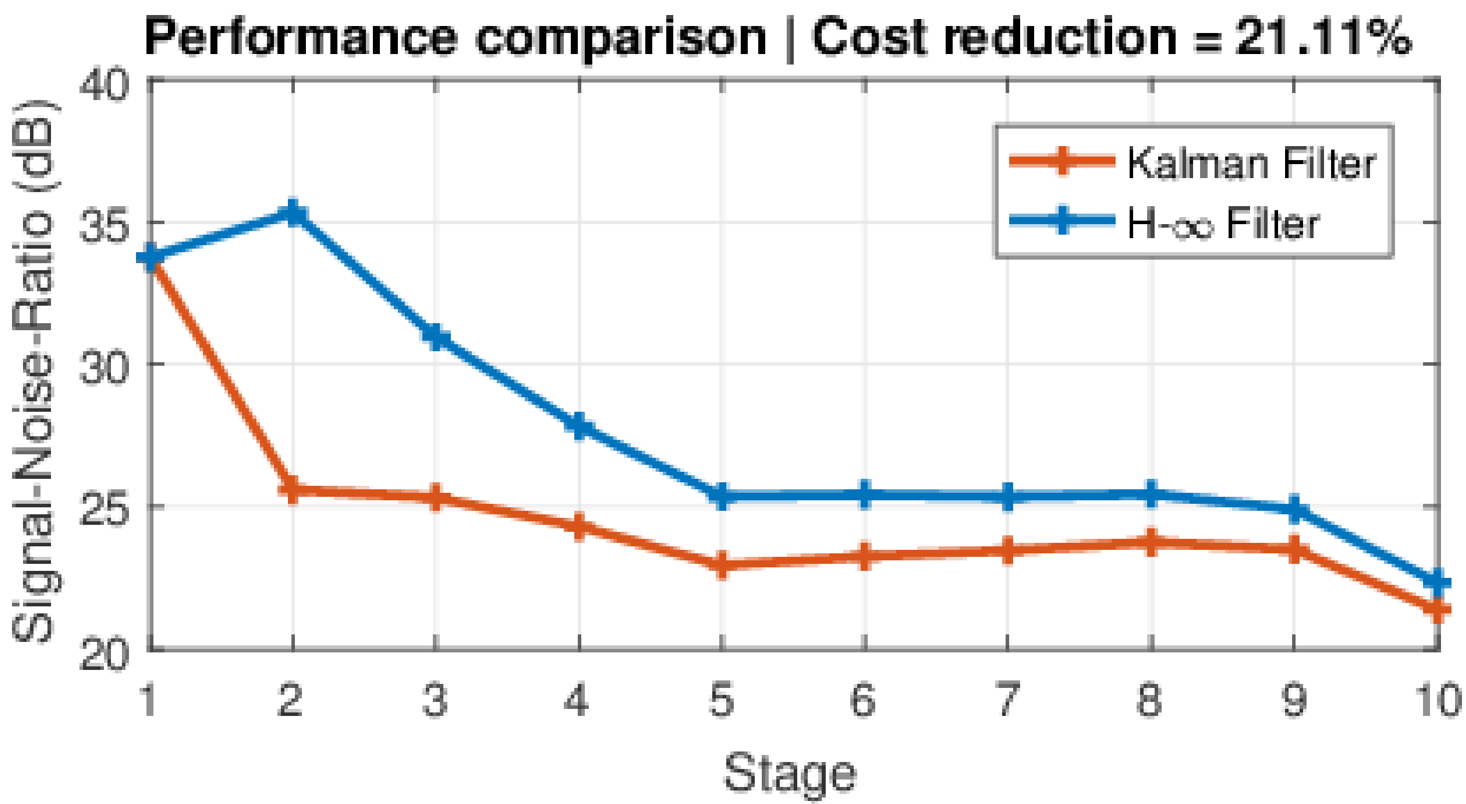
6 pairs showing Original (left) and Estimated (right) images of each of the 6 stages

Working with Images: Current Estimator

Original



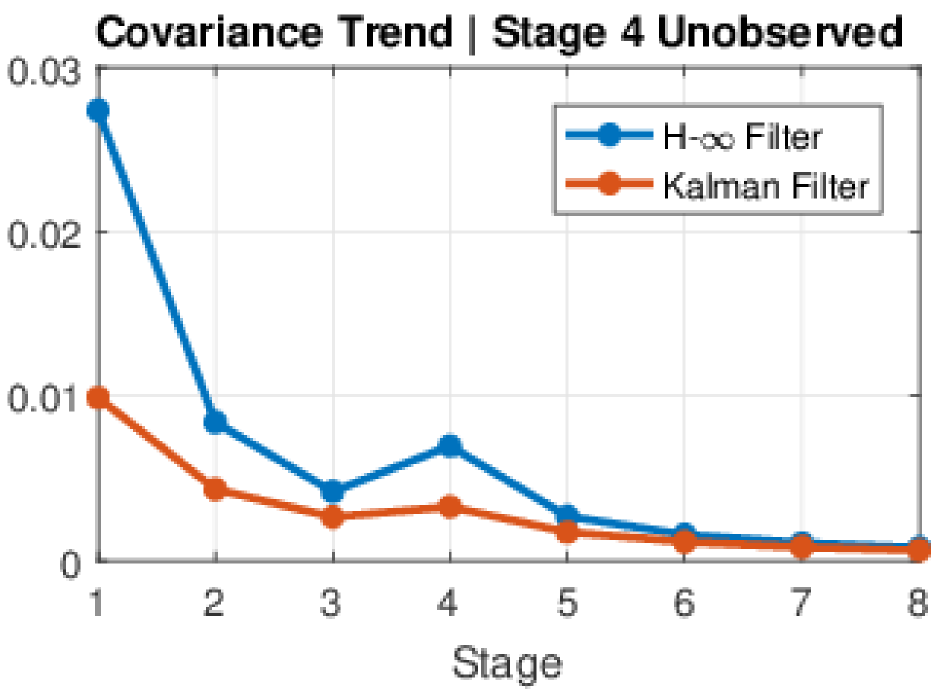
Comparison of SNR

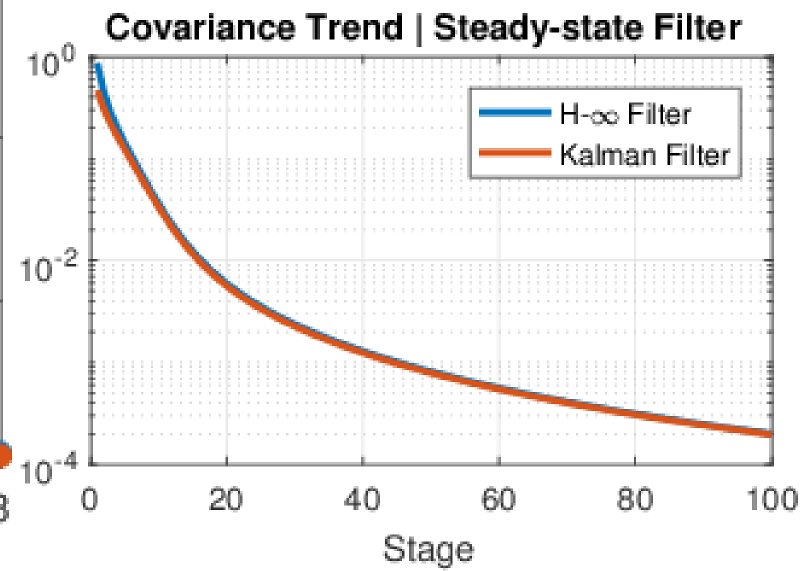


Working with Images: Missing Data

*Stage 4 was unobserved, hence the estimates (images on right) of stage 4 and 5 are identical

Covariance trend across stages





Conclusion

- Presented a minimax game-theoretic solution to multiscale optimal estimation problem.
  This solution avoids the need to separately solve smoothing and filtering problems which has been the classical approach in multiscale recursive estimation [3, 4].
- Reduction of 21% in the estimation cost observed by using H-infinity filter instead of Kalman filter.
  Estimation cost is calculated by accumulating squared differences between original and state estimates.
- High SNR value for H-infinity filter estimates in comparison to Kalman filter shown for all stages of multiresolution signal.
- *State-estimate-recovery* properties highlighted in the experiments exhibiting missing observations.
- Robust to worst-case additive exogenous noise.
- Experimental results corroborate theoretical findings; evaluations shown for 1-D and 2-D signal examples.

[3] Kenneth C Chou, Alan S Willsky, and Albert Benveniste, "Multiscale recursive estimation, data fusion, and regularization," IEEE Transactions on Automatic Control, vol. 39, no. 3, pp. 464–478, Mar 1994.
[4] Krishan M Nagpal and Pramod P Khargonekar, "Filtering and smoothing in an h setting," IEEE Transactions on Automatic Control, vol. 36, no. 2, pp. 152–166, 1991.