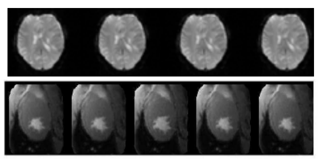
Kalman Filtered Compressed Sensing

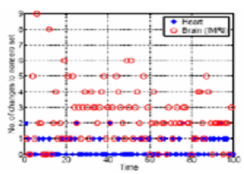
Namrata Vaswani, Iowa State University, namrata@iastate.edu

The Problem

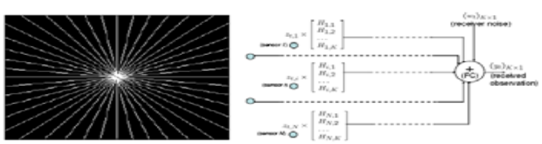
- **Causally** reconstruct time sequences of sparse signals with **slowly changing sparsity patterns** from a **limited number of noise-corrupted "incoherent" measurements**
- **Examples:**
  - making dynamic MRI reconstruction real-time
    - MRI measures a limited number of Fourier coefficients
    - Fourier "incoherent" w.r.t. sparsity basis of image (wavelet)
    - real-time capture & recons: needed for interventional radiology
  - "single-pixel" video camera with real-time display
  - real-time estimation of temperature or other random fields using the sensor network of [Haupt-Nowak06]



examples of sparse image sequences (brain, heart)



slowly changing sparsity pattern



(a) CT or MRI (projection) geometry [8]   (b) Sensing protocol of Haupt-Nowak [30]

obtaining incoherent measurements

Existing Work and Our Goal

- **Static version of the problem**
  - Compressed Sensing (CS) [Candes et al, Donoho], MP, OMP,...

$$\hat{x} = \arg\min_{\beta} ||\beta||_1 \ s.t. \ ||A'(y - A\beta)||_\infty < \lambda\sigma$$

- **The actual problem (recons. time series of sparse signals)**
  - CS at each time: causal (real-time), but large reconstruction error
  - Batch CS [Wakin et al (video), Gamper et al (MRI)]: non-causal
  - other related work: [Rozell et al'07], [Jung et al'08],
- **Our Goal:** a **causal** algorithm with **smallest possible reconstruction error**, for a given (small) number of measurements
- **Key Ideas:**
  - Sparsity pattern changes slowly with time: use estimated sparsity pattern from $t-1$ to improve CS reconstruction at t
  - Use prior model on nonzero coefficients values (if available)

Problem Formulation

- observation, $y_t : n \times 1$, unknown sparse vector, $x_t : m \times 1$, $n < m$
- at each time t, $y_t = Ax_t + w_t$, $w_t \sim \mathcal{N}(0, \sigma^2 I)$
- each x_t is s_t-sparse with support set T_t and $s_t = |T_t| << n < m$
- prior model on $(x_t)_{T_t}$: $(x_t)_{T_t} = (x_{t-1})_{T_t} + \mathcal{N}(0, \sigma^2_{sys} I_{s_t})$

Goal: Recursively obtain the "best" estimate of x_t from $y_1, y_2, \dots y_t$

Finding A Solution

- **If the sparsity pattern (support), T_t, is known at each t**
  - easy to compute the **restricted Least Squares (LS)** estimate:

$$(\hat{x}_t)_{T_t} = {A_{T_t}}^\dagger y_t, \ (\hat{x}_t)_{T_t^c} = 0$$

  - if prior model also available: Kalman filter (KF) on $(x_t)_{T_t}$
- **Our problem: the support, T_t, is unknown**
- **Solution strategy: use compressed sensing (CS) + thresholding**
  - option 1: CS on observation

$$\hat{T}_t = \text{threshold}(\text{CS}(y_t))$$

  - option 2: CS on LS error in observation computed using $\hat{T}_{t-1}$

$$\tilde{y}_t = y_t - A\hat{x}_{t,LS}, \ \hat{x}_{t,LS} := \text{restricted-LS}(y_t, \hat{T}_{t-1})$$

$$\hat{T}_t = \text{threshold}(CS(\tilde{y}_t) + \hat{x}_{t,LS})$$

  - option 3: replace LS by KF in option 2 if prior model available

Comparing CS with CS on LS (or KF) error

- **CS on observation**
  - uses CS to estimate x_t from $y_t = Ax_t + w_t$
  - x_t is $|T_t|$-sparse
- **CS on LS error (LSE) in observation**
  - uses CS to estimate $\beta_t = x_t - \hat{x}_{t,LS}$ from $\tilde{y}_t = A\beta_t + w_t$
  - β_t is only $|T_t \setminus \hat{T}_{t-1}|$-compressible (assumes $(\beta_t)_{T_{t-1}}$ is small)
- **If the sparsity pattern (T_t) changes slowly enough,**
  - LS estimation error, $(\beta_t)_{T_{t-1}}$, will be small
  - **CS on LSE has much smaller error than CS on observation** (CS error strongly depends on effective support size)
  - ongoing work: shown the above rigorously
- **CS on KF error (KFE) in observation**
  - analysis gets more complicated but **similar conclusions will hold**
  - KF estimation error will be smaller than LS error (assuming CS selects correct model)

Kalman Filtered Compressed Sensing (KF-CS)

Input: y_t, $\hat{T}_{t-1}$, $\hat{x}_{t-1}$, P_{t-1}, **Output:** $\hat{T}_t$, $\hat{x}_t$, P_t
Initialize: $T_0 = \phi$, $\hat{x}_0 = 0$, $P_0 = 0$. At each time $t > 0$, do

- **Kalman filter (KF) on $(x_t)_{\hat{T}_{t-1}}$ to compute KF error, $\tilde{y}_t$**

$$\hat{x}_{t,0} = \text{KF}(I, \sigma^2_{sys} I_{\hat{T}_{t-1}}, A_{\hat{T}_{t-1}}, \sigma^2 I)(\hat{x}_{t-1}, P_{t-1})$$

$$\tilde{y}_t = y_t - A\hat{x}_{t,0}$$

- **Compressed Sensing (CS) on KF error + thresholding**

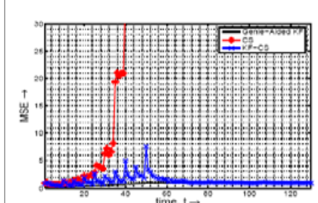
$$\hat{T}_t = \text{threshold}(\text{CS}(\tilde{y}_t) + \hat{x}_{t,0})$$

  - added set: $\hat{\Delta}_t = \hat{T}_t \setminus \hat{T}_{t-1}$, deleted set: $\hat{D}_t = \hat{T}_{t-1} \setminus \hat{T}_t$
  - zero out elements of deleted set from $\hat{x}_{t-1}$, P_{t-1}
- **Kalman filter (KF) on $(x_t)_{\hat{T}_t}$**

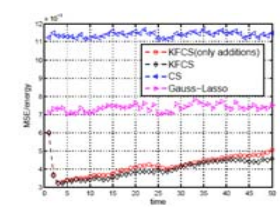
$$[\hat{x}_t, P_t] = \text{KF}(I, \sigma^2_{sys} I_{\hat{T}_t}, A_{\hat{T}_t}, \sigma^2 I)(\hat{x}_{t-1}, P_{t-1})$$

Simulation Results

- **Left plot: simulated data, average MSE over 100 simulations**
  - $m = 256$, $n = 72$, A = random Gaussian
  - s_t increased from 8 to 26,
  - but maximum number of new additions per unit time was 2
  - delay b/w two addition times was 5 time units
  - nonzero coefficients, $(x_t)_{T_t}$ followed random walk
- **Right plot: ongoing work on brain MRI reconstruction**



MSE plot for simulated data
m=256, n=72, A: rand. Gaussian



MRI recon [Qiu et al, submitted]
m=4096, n=2049, A = F_s W

Ongoing Work

- Shown that CS on LSE has lower error than CS if sparsity pattern changes slowly enough [Vaswani, submitted]
  - Ongoing work: similar analysis for CS on KFE
- Obtained sufficient conditions under which KF-CS approaches the genie-aided KF for large t [Vaswani, submitted]
- Working on KF-CS for dynamic MRI reconstruction [Qiu et al, submitted]