A Bayesian Approach for Online Recovery of Streaming Signals from Compressive Measurements

Uditha Lakmal Wijewardhana and Marian Codreanu

Geethu Joseph

SPC Lab, IISc

Recovery of Streaming Signal

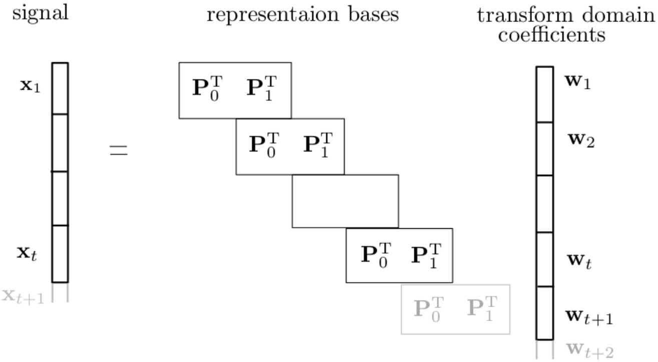
Recovery of a discrete-time streaming signal x from a time-varying linear measurement model

$$\boldsymbol{y}_t = \boldsymbol{\Phi}_t \boldsymbol{x}_t + \boldsymbol{e}_t, \qquad t = 1, 2, \ldots$$

- $\boldsymbol{x}_t = \begin{bmatrix} x(Nt - N + 1) & x(Nt - N + 2) & \ldots & x(Nt) \end{bmatrix} \in \mathbb{R}^N$
- $\boldsymbol{y}_t \in \mathbb{R}^M$, $M < N$
- $\boldsymbol{e}_t \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_M)$

Goal: Sequential recovery over short, shifting time intervals

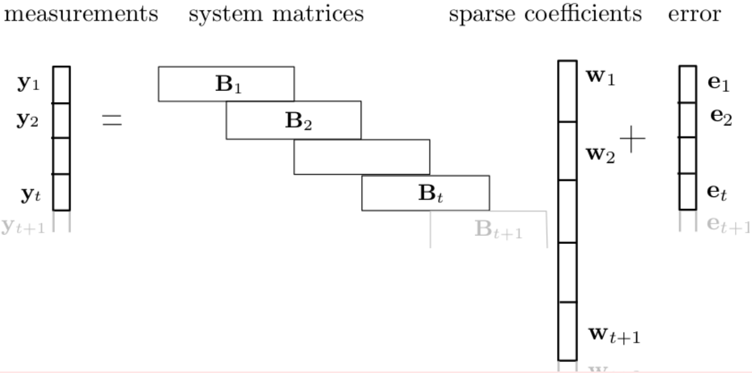
Sparsity: Lapped Orthogonal Transform



$$\boldsymbol{w}_t = \begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{t-1} \\ \boldsymbol{x}_t \end{bmatrix} \Leftrightarrow \boldsymbol{x}_t = \begin{bmatrix} \boldsymbol{P}_0^{\mathsf{T}} & \boldsymbol{P}_1^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_t \\ \boldsymbol{w}_{t+1} \end{bmatrix}$$

LOT matrix: $\boldsymbol{P}_0\boldsymbol{P}_0^{\mathsf{T}} + \boldsymbol{P}_0\boldsymbol{P}_1^{\mathsf{T}} = \boldsymbol{I}$ and $\boldsymbol{P}_0\boldsymbol{P}_1^{\mathsf{T}} = \boldsymbol{P}_1^{\mathsf{T}}\boldsymbol{P}_0 = \boldsymbol{0}$

System Model



$$\mathbf{x}_t = \begin{bmatrix} \mathbf{P}_0^\mathsf{T} & \mathbf{P}_1^\mathsf{T} \end{bmatrix} \begin{bmatrix} \mathbf{w}_t \\ \mathbf{w}_{t+1} \end{bmatrix} \implies \mathbf{y}_t = \mathbf{B}_t \begin{bmatrix} \mathbf{w}_t \\ \mathbf{w}_{t+1} \end{bmatrix} + \mathbf{e}_t, t = 1, 2, \ldots$$

Online Recovery

Optimal estimation: joint estimation of $\{\boldsymbol{w}_\tau\}_{\tau=1}^t$ from $\{\boldsymbol{y}_\tau\}_{\tau=1}^t$

$$\boldsymbol{y}_t = \boldsymbol{B}_t \begin{bmatrix} \boldsymbol{w}_t \\ \boldsymbol{w}_{t+1} \end{bmatrix} + \boldsymbol{e}_t, t = 1, 2, \ldots$$

Why online?

- Smaller reconstruction delay
- Low computational complexity
- Reduced memory demands

Approach: Sparse Bayesian Learning

- Easily accommodate the coupling among the measurements

Offline Sparse Bayesian Learning

System model

$$\underline{\boldsymbol{y}}_t = \underline{\boldsymbol{B}}_t \underline{\boldsymbol{w}}_t + \underline{\boldsymbol{e}}_t,$$

- $\underline{\boldsymbol{w}}_t = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \ldots & \boldsymbol{w}_{t+1} \end{bmatrix} \in \mathbb{R}^{N(t+1)}$
- $\underline{\boldsymbol{y}}_t, \underline{\boldsymbol{e}}_t \in \mathbb{R}^{Mt}$ and $\underline{\boldsymbol{B}} \in \mathbb{R}^{Mt \times N(t+1)}$

Two-stage hierarchal model: $\underline{\boldsymbol{\alpha}}_t \in \mathbb{R}_+^{N(t+1)}$ - precision parameter

$$p(\underline{\boldsymbol{y}}_t | \underline{\boldsymbol{w}}_t) = \mathcal{N}(\underline{\boldsymbol{y}}_t | \underline{\boldsymbol{B}}_t \underline{\boldsymbol{w}}_t, \sigma^2 \boldsymbol{I}_{Mt})$$

$$p(\underline{\boldsymbol{w}}_t | \underline{\boldsymbol{\alpha}}_t) = \prod_{\tau=1}^{t+1} \prod_{i=1}^{N} \mathcal{N}(w_{\tau,i} 0, \alpha_{\tau,i}^{-1})$$

$$p(\underline{\boldsymbol{\alpha}}_t | a, b) = \prod_{\tau=1}^{t+1} \prod_{i=1}^{N} \text{Gamma}(\alpha_{\tau,i} | a, b)$$

SBL Algorithm

Input: $\underline{\boldsymbol{y}}_t, \underline{\boldsymbol{B}}_t$ measurements, measurement matrix
Initialization: Initialize $\underline{\boldsymbol{\alpha}}_t$ precision parameter
Repeat

E-step

$$\underline{\boldsymbol{A}}_t^{-1} = \text{diag}(\underline{\boldsymbol{\alpha}}_t)$$
$$\boldsymbol{\Sigma}_{\underline{\boldsymbol{w}}}^t = \underline{\boldsymbol{A}}_t^{-1} - \underline{\boldsymbol{A}}_t^{-1}\underline{\boldsymbol{B}}_t^{\mathsf{T}}(\sigma^2 \boldsymbol{I}_{Mt} + \underline{\boldsymbol{B}}_t\underline{\boldsymbol{A}}_t^{-1}\underline{\boldsymbol{B}}_t)\underline{\boldsymbol{B}}_t\underline{\boldsymbol{A}}_t^{-1}$$ covariance update
$$\mu_{\underline{\boldsymbol{w}}}^t = \sigma^{-2}\boldsymbol{\Sigma}_{\underline{\boldsymbol{w}}}^t\underline{\boldsymbol{B}}_t^{\mathsf{T}}\underline{\boldsymbol{y}}_t$$ mean update

M-step

$$\underline{\boldsymbol{\alpha}}_{\tau,i} = f(\mu_{\tau,i}^t, \Sigma_{tau,ii}^t)$$ eq: $\underline{\boldsymbol{\alpha}}_{\tau,i} = \frac{1+2a}{(\mu_{\tau,i}^t)^2 + \Sigma_{\tau,ii}^t + 2b}$

Until $\underline{\boldsymbol{\alpha}}_t$ converges
Output: $\mu_{\underline{\boldsymbol{w}}}^t, \boldsymbol{\Sigma}_{\underline{\boldsymbol{w}}}^t, \underline{\boldsymbol{\alpha}}_{\tau,i}$

Recursive SBL Using d-block Banded Approximation

Consider the inverse of covariance update:

$$\underline{\boldsymbol{H}}_t = \underline{\boldsymbol{A}}_t + \sigma^{-2}\underline{\boldsymbol{B}}_t^{\mathsf{T}}\underline{\boldsymbol{B}}_t$$

$\underline{\boldsymbol{H}}_t$ is block tridiagonal matrix

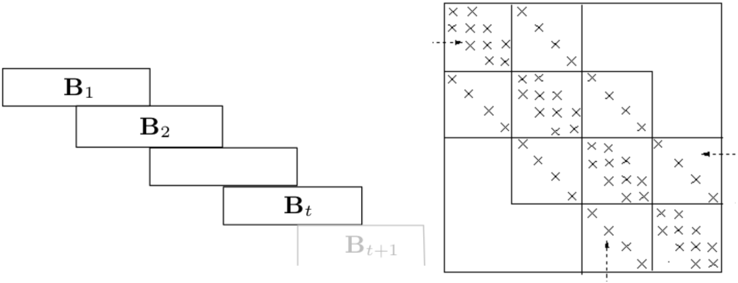


Figure: stucture of $\underline{\boldsymbol{B}}_t$ and $\underline{\boldsymbol{H}}_t$

Recursive SBL Using d-block Banded Approximation

Consider the inverse of covariance update:

$$\underline{\boldsymbol{H}}_t = \underline{\boldsymbol{A}}_t + \sigma^{-2} \underline{\boldsymbol{B}}_t^{\mathsf{T}} \underline{\boldsymbol{B}}_t$$

$\underline{\boldsymbol{H}}_t$ is block tridiagonal matrix

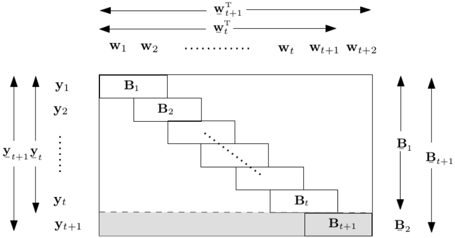
Result

Inverse of a banded matrix is band dominant matrix

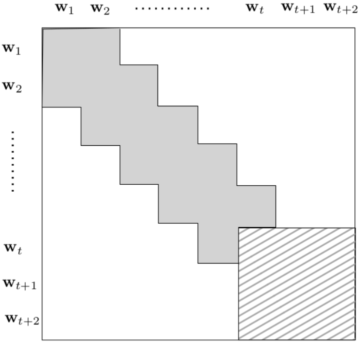
$$\boldsymbol{\Sigma}_{\underline{\boldsymbol{w}}}^{t} = \underline{\boldsymbol{H}}_t^{-1} \approx d\text{-block banded matrix}$$

d - design parameter

Approximation: Adding A New Measurement



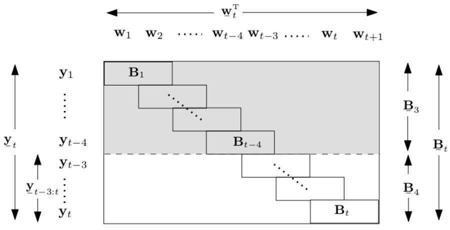
new sensing matrix



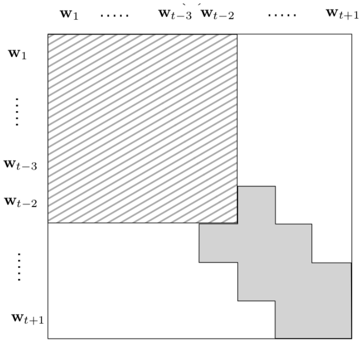
estimates to be recomputed

$p(\boldsymbol{w}_{1:t-d}|\boldsymbol{y}_{1:t+1}) \approx p(\boldsymbol{w}_{1:t-d}|\boldsymbol{y}_{1:t}) \implies$ estimates of $\boldsymbol{w}_{1:t-d}$ remain unchanged

Approximation: Removing Old Measurements



new sensing matrix



estimates to be recomputed

$p(\boldsymbol{w}_{t-d:t+1}|\boldsymbol{y}_{1:t}) \approx p(\boldsymbol{w}_{t-d:t+1}|\boldsymbol{y}_{t-2d-1:t}) \implies$ estimates of $\boldsymbol{w}_{1:t-d}$ remain unchanged

Effect of Approximation: $d = 1$ and $t = 5$

Adding $\boldsymbol{y}_6$

$\Downarrow$

estimates of $\boldsymbol{w}_{1:4}$ remain unchanged

Effect of Approximation: $d = 1$ and $t = 5$

Adding $\boldsymbol{y}_6$

$\Downarrow$

estimates of $\boldsymbol{w}_{1:4}$ remain unchanged

Removing $\boldsymbol{y}_1$

$\Downarrow$

estimate of $\boldsymbol{w}_{4:6}$ remain unchanged

Effect of Approximation: $d = 1$ and $t = 5$

Adding $\boldsymbol{y}_6$

$\Downarrow$

estimates of $\boldsymbol{w}_{1:4}$ remain unchanged

Removing $\boldsymbol{y}_1$

$\Downarrow$

estimate of $\boldsymbol{w}_{4:6}$ remain unchanged

$\boldsymbol{w}_4$ depends only on $\boldsymbol{y}_{2:5}$

Effect of Approximation: $d = 1$ and $t = 5$

Adding $\boldsymbol{y}_6$
$\Downarrow$
estimates of $\boldsymbol{w}_{1:4}$ remain unchanged

Removing $\boldsymbol{y}_1$
$\Downarrow$
estimate of $\boldsymbol{w}_{4:6}$ remain unchanged

$\boldsymbol{w}_4$ depends only on $\boldsymbol{y}_{2:5}$
$\Downarrow$
$\boldsymbol{w}_{t-d}$ depends only on $\tilde{\boldsymbol{y}}_t = \boldsymbol{y}_{t-2d-1:t}$

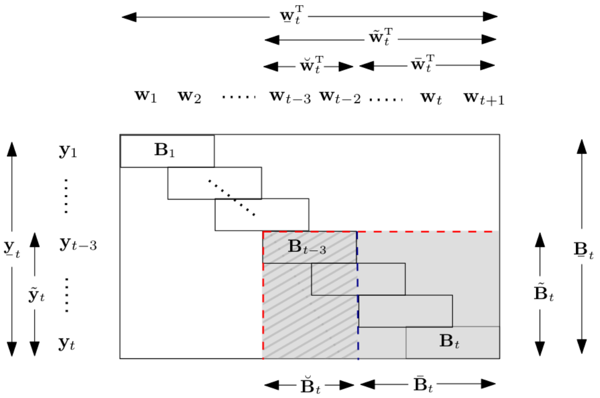
- To estimate $\boldsymbol{w}_{t-d}$ we consider only $\tilde{\boldsymbol{y}}_t$

Allows for processing over overlapping sliding windows

Sliding Processing Window

To estimate $\boldsymbol{w}_{t-d}$ we consider $\tilde{\boldsymbol{y}}_t = \tilde{\boldsymbol{B}}_t \tilde{\boldsymbol{w}}_t + \tilde{\boldsymbol{e}}_t$

- $\tilde{\boldsymbol{y}}_t = \boldsymbol{y}_{t-2d-1:t}$
- $\tilde{\boldsymbol{w}}_t = \boldsymbol{w}_{t-2d-1:t+1}$



$$\tilde{\mathbf{w}}_t = \begin{bmatrix} \breve{\mathbf{w}}_t^T & \bar{\mathbf{w}}_t^T \end{bmatrix}^T$$

$$\tilde{\mathbf{B}}_t = \begin{bmatrix} \breve{\mathbf{B}}_t & \bar{\mathbf{B}}_t \end{bmatrix}$$

$$\begin{aligned} \tilde{\mathbf{y}}_t &= \tilde{\mathbf{B}}_t \tilde{\mathbf{w}}_t + \tilde{\mathbf{e}}_t \\ &= \begin{bmatrix} \breve{\mathbf{B}}_t & \bar{\mathbf{B}}_t \end{bmatrix} \begin{bmatrix} \breve{\mathbf{w}}_t \\ \bar{\mathbf{w}}_t \end{bmatrix} + \tilde{\mathbf{e}}_t \\ &= \breve{\mathbf{B}}_t \breve{\mathbf{w}}_t + \bar{\mathbf{B}}_t \bar{\mathbf{w}}_t + \tilde{\mathbf{e}}_t \end{aligned}$$

Recursive SBL

- Model: $\tilde{\boldsymbol{y}}_t = \underbrace{\breve{\boldsymbol{B}}_t \breve{\boldsymbol{w}}_t}_{\text{does not depend on} \tilde{\boldsymbol{y}}_t} + \underbrace{\bar{\boldsymbol{B}}_t \bar{\boldsymbol{w}}_t}_{\text{depends on} \tilde{\boldsymbol{y}}_t} + \tilde{\boldsymbol{e}}_t$

Two stage hierarchical model

$$p(\bar{\boldsymbol{w}}_t | \bar{\boldsymbol{\alpha}}_t) = \prod_{\tau=1-d}^{t+1} \prod_{i=1}^{N} \mathcal{N}(w_{\tau,i} 0, \alpha_{\tau,i}^{-1})$$

$$p(\bar{\boldsymbol{\alpha}}_t | a, b) = \prod_{\tau=t-d}^{t+1} \prod_{i=1}^{N} \text{Gamma}(\alpha_{\tau,i} | a, b)$$

$$p(\tilde{\boldsymbol{y}}_t | \bar{\boldsymbol{w}}_t) = \mathcal{N}(\tilde{\boldsymbol{y}}_t | \breve{\boldsymbol{B}}_t \mathbb{E}(\breve{\boldsymbol{w}}_t | \bar{\boldsymbol{w}}_t) + \bar{\boldsymbol{B}}_t \bar{\boldsymbol{w}}_t, \\ \breve{\boldsymbol{B}}_t \text{cov}(\breve{\boldsymbol{w}}_t | \bar{\boldsymbol{w}}_t) \breve{\boldsymbol{B}}_t^{\mathsf{T}} + \sigma^2 \boldsymbol{I})$$

- $p(\breve{\boldsymbol{w}}_t | \bar{\boldsymbol{w}}_t)$: obtained from past estimates

Algorithm

Input: $\tilde{\boldsymbol{y}}_t$, $\tilde{\boldsymbol{B}}_t$, $\mu_{\tilde{\boldsymbol{w}}}^{t-1}$, $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{w}}}^{t-1}$ and $\bar{\boldsymbol{\alpha}}_{t-1}$ measurements and past estimates
Initialization:
Compute $p(\breve{\boldsymbol{w}}_t|\bar{\boldsymbol{w}}_t)$ using $\mu_{\underline{\boldsymbol{w}}}^{t-1}$, $\boldsymbol{\Sigma}_{\underline{\boldsymbol{w}}}^{t-1}$
Repeat

E-step

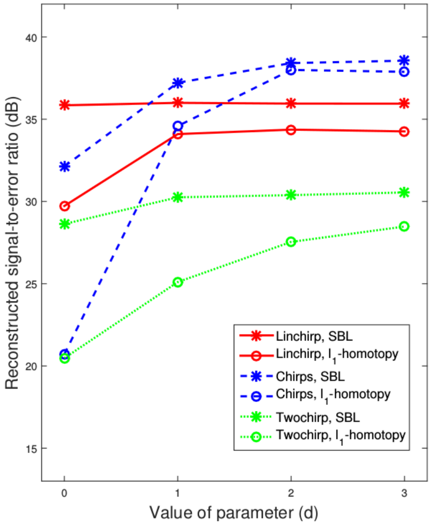
Compute $p(\tilde{\boldsymbol{y}}_t|\bar{\boldsymbol{w}}_t)$ using $p(\breve{\boldsymbol{w}}_t|\bar{\boldsymbol{w}}_t)$
Update $\mu_{\underline{\boldsymbol{w}}}^{t}$ and $\boldsymbol{\Sigma}_{\underline{\boldsymbol{w}}}^{t}$

M-step

Update $\bar{\alpha}_t$

Until convergence criteria is met
Output:$\mu_{\tilde{\boldsymbol{w}}}^{t}$, $\boldsymbol{\Sigma}_{\tilde{\boldsymbol{w}}}^{t}$ and $\bar{\boldsymbol{\alpha}}_t$

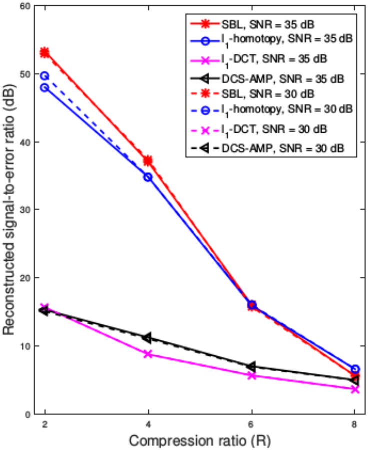
Numerical Results: Varying d

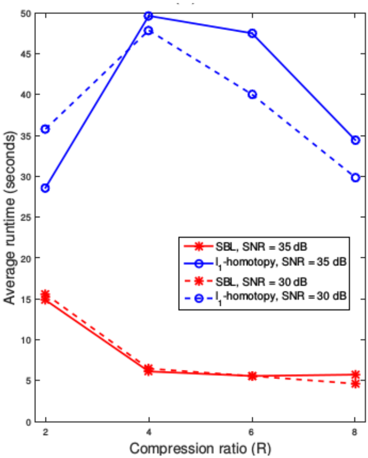


N	256
M	64
SNR	35dB
$\mathbf{\Phi}_t$	$\pm 1/\sqrt{M}$ with equal prob.

No significant improvement beyond $d = 1$

Numerical Results: Varying M





N	256
M	N/R
$\mathbf{\Phi}_t$	$\pm 1/\sqrt{M}$ with equal prob.

Faster than l_1 homotopy

Summary: Recovery Procedure

Step 1

Apply lapped orthogonal transform to get sparse representation: $\boldsymbol{x} \rightarrow \boldsymbol{w}$

Step 2

At time t use **recursive algorithm** to estimate $\boldsymbol{w}_{t-d}$ using $\boldsymbol{y}_{t-2d-1:t}$

- Sliding window processing
- SBL framework
- Utilizes previous estimates

Step 2

Reconstruct $\boldsymbol{x}_{t-d-1}$ as

$$\hat{\boldsymbol{x}}_{t-d-1} = \begin{bmatrix} \boldsymbol{P}_0^{\mathsf{T}} & \boldsymbol{P}_1^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_{t-d-1}(t-1) \\ \boldsymbol{w}_{t-d}(t) \end{bmatrix}$$