Reduced Rank Adaptive Filtering in Impulsive Noise Environments

Hamza SOURY (KAUST, Saudi Arabia)
Mohamed-Slim Alouini (KAUST, Saudi Arabia)
Karim Abed-Meraim (Polytech' Orléans, France)

April 19, 2015

Motivation

- ✓ Gaussian noise is widely used in communication systems
- ✓ In sensor networks and local spectrum sensing, the deployed noise is Generalized Gaussian (GG)
- ✓ In real life applications, sever perturbations may have an impulsive nature $\rightarrow$ Impulsive noise
- ✓ Impulsive noise can be modeled by an α-stable distribution

KAUST

α-stable distribution and Impulsive noise

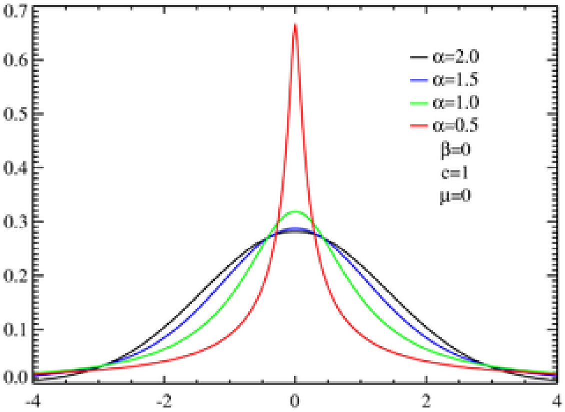

Figure: PDF of α stable distribution for different values of α and zero mean.

Adaptive Filtering

- ✓ From an observation vector (data) $\mathbf{r}$, one estimates the original signal $\mathbf{s}$
- ✓ The estimator $\hat{\mathbf{s}}$ is obtained by minimizing the error (square, L_1, L_p ...)
  - Square error: $e_2 = E\left[\|\hat{s} - s\|^2\right]$
  - L_1 Norm error: $e_1 = E\left[\|\hat{s} - s\|\right]$
  - L_p Norm error: $e_p = E\left[\|\hat{s} - s\|^p\right]$
- ✓ The linear filter is the most used (less complexity) $\hat{\mathbf{s}} = \mathbf{w}\ \mathbf{r}$
- ✓ However, full rank filter remains high complex filter
- ✓ Objectives:
  - Building a reduced rank adaptive filter in impulsive noise
  - Using different subspaces to get the reduced rank filter

System Model and Full Rank Solution

- ✓ Impulsive noise appears as impulses with high values and small probability
- ✓ The used family of impulsive noise is the α-stable family
- ✓ Dealing with impulsive noise requires the use of L_p norm rather than L_2
- ✓ The best choice of p is α
- ✓ The system model in this section is

$$\mathbf{r}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{n}(k),$$

where $\mathbf{r}(k)$ is an $N \times 1$ observed vector at time instant k, $\mathbf{s}(k)$ is the $M \times 1$ source data, $\mathbf{H}$ is the $N \times M$ channel matrix, and $\mathbf{n}(k)$ is a noise vector.

Objectives

- ✓ Estimate the source data given the observation vector using a filter **w**

$$\hat{s}_1(k) = \mathbf{w}^H \mathbf{r}(k)$$

- ✓ Minimize the error function $J_p(\mathbf{w})$

$$J_p(\mathbf{w}) = E\left[\|\hat{s}_1(k) - s_1(k)\|^p\right] = E\left[\|s_1(k) - \mathbf{w}^H \mathbf{r}(k)\|^p\right]$$

- ✓ Eliminate the impulses that degrades the signal
- ✓ Build a reduced rank subspace to get a reduced rank filter
- ✓ Test the filter to validate the model
- ✓ Analyze the performance of the system
- ✓ Analyze the convergence of the algorithm

Full Rank Solution Using L_2 Norm

- ✓ Using L_2, the filter is the usual Wiener filter given by

$$\mathbf{R}\mathbf{w}_{opt}^{N} = \mathbf{c},$$

where $\mathbf{c} = E\left[\mathbf{r}(k)s_1^*(k)\right]$ is the cross-correlation vector b/w $\mathbf{r}(k)$ and $s_1(k)$, and $\mathbf{R} = E\left[\mathbf{r}(k)\mathbf{r}^H(k)\right]$ is the covariance matrix of $\mathbf{r}(k)$

- ✓ An estimation $\mathbf{R}$ is needed due to indefinite variance of noise, variation on channel...
- ✓ In impulsive noise, the best estimator of $\mathbf{R}$ is the fixed point estimator

$$\mathbf{R} = \frac{N}{T}\sum_{k=1}^{T}\frac{\mathbf{r}(k)\mathbf{r}^H(k)}{\mathbf{r}^H(k)\mathbf{R}^{-1}\mathbf{r}(k)}$$

L_p Norm Solution

- ✓ The gradient of the error function

$$\nabla J_p(\mathbf{w}) = \frac{p}{2} E\left[csgn\left(\mathbf{w}^H \mathbf{r}(k) - s_1(k)\right) \|s_1(k) - \mathbf{w}^H \mathbf{r}(k)\|^{p-1} \mathbf{r}(k)\right]$$

- ✓ The steepest descent update formula

$$\begin{aligned} \mathbf{w}(k) &= \mathbf{w}(k-1) - \quad \mu \nabla J_p(\mathbf{w}(k-1)) \\ &\approx \mathbf{w}(k-1) + \quad \mu\, \mathbf{r}(k) csgn\left(s_1(k) - \mathbf{w}(k-1)^H \mathbf{r}(k)\right) \\ &\qquad \times \|s_1(k) - \mathbf{w}^H \mathbf{r}(k)\|^{p-1}. \end{aligned}$$

1 Full Rank Solution

2 **Signal Truncation**

3 Reduced Rank Filter

4 Simulation Results

5 Summary

Signal Truncation

- ✓ Big values of impulses degrades the received signal
- ✓ Degradation of $\mathbf{r}(k)$ causes a lose on system performance
- ✓ Truncation signal appears as a solution
- ✓ Need a good level of truncation $\Rightarrow$ What about using using quantile level?

Truncation by quantile level

- ✓ For given quantile q, δ_k is chosen so $q\%$ of data magnitude is less than δ_k
- ✓ The truncated data is given by

$$\hat{\mathbf{r}}_i(k) = \begin{cases} \mathbf{r}_i(k) & \text{If } \ |\mathbf{r}_i(k)| \leq \delta_k \\ \delta_k e^{j\ \text{angle}(\mathbf{r}_i(k))} & \text{If } \ |\mathbf{r}_i(k)| > \delta_k, \end{cases}$$

- ✓ A numerical relation between α and q is obtained by simulations

Reduced Rank Filter

- ✓ A D-dimensional subspace of $\mathbb{C}^N$, $\mathcal{S}^D$ should be defined
- ✓ The reduced rank filter is defined as

$$\mathbf{w}_{opt}^D = \arg\min_{\mathbf{w}\in\mathcal{S}^{\mathrm{D}}} \mathrm{J_p}(\mathbf{w}).$$

- ✓ The orthogonal projection of $\mathbf{w}_{opt}^D$ on $\mathcal{S}^D$ is $\mathbf{h}_{opt}^D$
- ✓ $\mathbf{h}_{opt}^D$ is defined as

$$\mathbf{h}_{opt}^D = \arg\min_{\mathbf{h}\in\mathbb{C}^{\mathrm{D}}} \mathrm{J_p}(\mathbf{Q}\mathbf{h})$$

where $\mathbf{Q}$ represents an orthonormal basis of $\mathcal{S}^D$

Successive Gradient

- ✓ Let $\nabla J_p^i(\mathbf{x})$ the successive application of the gradient of J_p to $\mathbf{x}$

$$\overbrace{\nabla J_p\left(\nabla J_p\left(\nabla J_p \ldots\right)\right)}^{i \text{ times}}(\mathbf{x})$$

- ✓ Successive gradient subspace

$$\mathcal{S}^D = \text{span}\left\{\mathbf{c}, \nabla J_p(\mathbf{c}), \ldots, \nabla J_p^{D-1}(\mathbf{c})\right\},$$

Krylov Subspace

✓ Krylov subspace, $\mathcal{K}^D(\mathbf{R}, \mathbf{c})$, inducted by $\mathbf{R}$ and $\mathbf{c}$

$$\mathcal{K}^D(\mathbf{R}, \mathbf{c}) = \text{span}\{\mathbf{c}, \mathbf{Rc}, \mathbf{R}^2\mathbf{c}, \ldots, \mathbf{R}^{D-1}\mathbf{c}\}.$$

$\rightarrow$ $\mathbf{R} = E\left[\hat{\mathbf{r}}(k)\hat{\mathbf{r}}^H(k)\right], \mathbf{c} = E\left[\hat{\mathbf{r}}(k)s_1^*(k)\right]$

✓ Orthonormal basis of $\mathcal{S}^D$ computed using Gram-Schmidt algorithm.

Impact of Solution Parameters on Filter Performance

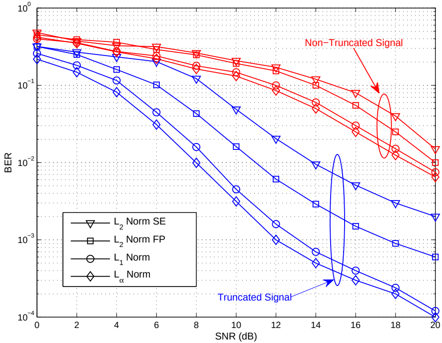

Figure: Impact of signal truncation, norm choice, and estimation method on the filter performance for $\alpha = 1.5$, $N = 32$, and $M = 6$.

- ✓ L_α has better performance than L_1 and L_2
- ✓ Fixed point method performs better than sample estimation
- ✓ Truncated signal model has less BER than non-truncated signal model (10^{-4} and 10^{-2} respectively)

Relation Between Quantile Order and Noise Factor

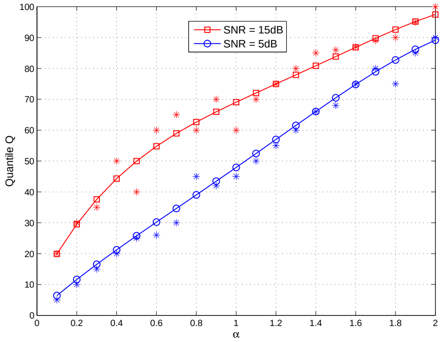

Figure: Best choice of quantile versus α ($N_w = 1000$)

- ✓ Increasing function b/w the quantile and α
- ✓ For close values of α to 2, the quantile approaches to 100%
- ✓ q is a function of α and the SNR, i.e. $q = f(\alpha, SNR)$
- ✓ An ad-hoc expression of f will be given in future works

Tracking Behavior for $D = 2,\ D = 4,\ D = M = 6$

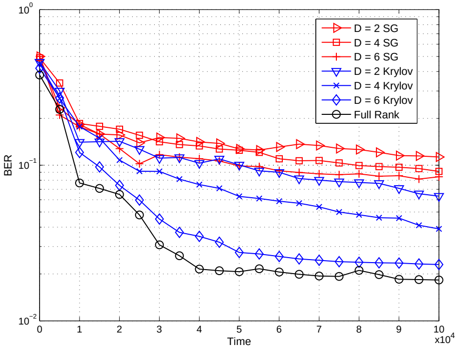

Figure: Tracking performance behavior for reduced rank filters using L_α norm for $\alpha = 1.2$ and $SNR = 15dB$.

- ✓ Krylov subspace has better performance than successive gradient
- ✓ The algorithm converges for low rank ($D = 2$)
- ✓ System performance degrades by reducing the filter rank
- ✓ A small gap b/w full dimension filter and low dimension $D = M$, 10^{-3} for BER$= 10^{-2}$

Conclusion

- ✓ Truncated signal performs better than non truncated signal in both full and reduced rank
- ✓ Even with small rank, the reduced rank filter performs near the full rank filter
- ✓ In subspace reduction, the Krylov subspace is the best subspace in filter reduction rank
- ✓ More analytical works are needed to choose the quantile and to prove the convergence of the filter

Thank you for your attention

Questions ?

Reduced Rank Adaptive Filtering in Impulsive Noise Environments

Hamza SOURY (KAUST, Saudi Arabia)
Mohamed-Slim Alouini (KAUST, Saudi Arabia)
Karim Abed-Meraim (Polytech' Orléans, France)

April 19, 2015