# Robust Regularized Least-Squares Beamforming Approach to Signal Estimation


**Mohamed Suliman, Tarig Ballal, and Tareq Y. Al-Naffouri**

King Abdullah University of Science and Technology (KAUST)

Emails:{mohamed.suliman, tarig.ahmed, tareq.alnaffouri}@kaust.edu.sa


## 1. Abstract


- We address the problem of robust adaptive beamforming of signals received by a linear array.
- The challenge associated with the beamforming problem is twofold. Firstly, the process requires the inversion of an ill-conditioned covariance matrix of the received signals. Secondly, the steering vector pertaining to the direction of arrival of the signal of interest is not known precisely.
- To tackle these two challenges, we manipulate the standard capon beamformer to a form where the beamformer output is obtained as a scaled version of the inner product of two vectors that are linearly related to the steering vector and the received signal snapshot. The linear operator, in both cases, is the square root of the covariance matrix.
- We proposed a new *regularized least-squares* (RLS) approach to estimate these two vectors and to provide robustness without any prior information.


## 2. Background

- Let us consider the linear model

$$\mathbf{r} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{1}$$

where
  - $\mathbf{A} \in \mathbb{C}^{m\times n}$ is a Hermitian matrix.
  - $\mathbf{v}$ is AWGN noise vector with unknown variance $\sigma_{\mathbf{v}}^2$.
- Estimating $\mathbf{x}$ using the *least-squares* (LS) leads to a solution that is very sensitive to perturbations in the data.
- To overcome this difficulty, *regularization methods* are frequently used. We are particularly interested in the RLS given by

$$\hat{\mathbf{x}}_{\text{RLS}} = (\mathbf{A}^H\mathbf{A} + \gamma I)^{-1}\mathbf{A}^H\mathbf{r}, \tag{2}$$

- Several methods have been proposed to select $\gamma$
  - L-curve.
  - generalized cross validation (GCV).
  - quasi-optimal.

## 3. Proposed Beamforming Approach

- The output of a beamformer for an array with $n_e$ elements, at time instant $t$, is

$$y_{\text{BF}}[t] = \mathbf{w}^H\mathbf{y}[t], \tag{3}$$

where:
  - $\mathbf{w} \in \mathbb{C}^{n_e}$ is the weighting coefficients vector.
  - $\mathbf{y}[t] \in \mathbb{C}^{n_e}$ is the array observations (snapshots) vector.
- For the Capon/MVDR beamformer, $\mathbf{w}$ is given by

$$\mathbf{w}_{\text{MVDR}} = \frac{\hat{\mathbf{C}}_{\mathbf{yy}}^{-1}\mathbf{a}}{\mathbf{a}^H\hat{\mathbf{C}}_{\mathbf{yy}}^{-1}\mathbf{a}}, \tag{4}$$

where:
  - $\mathbf{a}$: array *steering vector*.
  - $\hat{\mathbf{C}}_{\mathbf{yy}}$: sample covariance matrix of $\mathbf{y}$

$$\hat{\mathbf{C}}_{\mathbf{yy}} = \frac{1}{n_s}\sum_{t=1}^{n_s}\mathbf{y}[t]\mathbf{y}[t]^H. \tag{5}$$

- The difficulty with the MVDR beamformer is due to the ill-conditionedness of the matrix $\hat{\mathbf{C}}_{\mathbf{yy}}$ and the uncertainty in the steering vector $\mathbf{a}$.
- Based on (3) and (4), we can write

$$y_{\text{BF}}[t] = \frac{\mathbf{a}^H\hat{\mathbf{C}}_{\mathbf{yy}}^{-\frac{1}{2}}\hat{\mathbf{C}}_{\mathbf{yy}}^{-\frac{1}{2}}\mathbf{y}}{\mathbf{a}^H\hat{\mathbf{C}}_{\mathbf{yy}}^{-\frac{1}{2}}\hat{\mathbf{C}}_{\mathbf{yy}}^{-\frac{1}{2}}\mathbf{a}} = \frac{\mathbf{b}^H\mathbf{z}}{\mathbf{b}^H\mathbf{b}}, \tag{6}$$

where:
  - $\mathbf{b} \triangleq \hat{\mathbf{C}}_{\mathbf{yy}}^{-\frac{1}{2}}\mathbf{a}$ and $\mathbf{z} \triangleq \hat{\mathbf{C}}_{\mathbf{yy}}^{-\frac{1}{2}}\mathbf{y}$.
- $\mathbf{b}$ and $\mathbf{z}$ can be thought of as

$$\mathbf{a} = \hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}\mathbf{b}, \tag{7}$$

and

$$\mathbf{y} = \hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}\mathbf{z}. \tag{8}$$

- Since $\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}$ is ill-conditioned, direct inversion does not provide a viable solution.
- Given that $\mathbf{a}$ and $\mathbf{y}$ are noisy, we propose using a regularization algorithm to estimate $\mathbf{b}$ and $\mathbf{z}$ based on (7) and (8).
- Using (2) for $\mathbf{A} = \hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}$ and the *eigenvalue decomposition* (EVD) $\hat{\mathbf{C}}_{\mathbf{yy}} = \mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^H$, the beamformer output using RLS will take the form

$$y_{\text{BF-RLS}} = \frac{\mathbf{a}^H\mathbf{U}\left(\mathbf{\Sigma}^2 + \gamma_{\text{b}}\mathbf{I}\right)^{-1}\left(\mathbf{\Sigma}^2 + \gamma_{\text{z}}\mathbf{I}\right)^{-1}\mathbf{\Sigma}^2\mathbf{U}^H\mathbf{y}}{\mathbf{a}^H\mathbf{U}\left(\mathbf{\Sigma}^2 + \gamma_{\text{b}}\mathbf{I}\right)^{-2}\mathbf{\Sigma}^2\mathbf{U}^H\mathbf{a}}, \tag{9}$$

where:
  - $\gamma_{\text{b}}$ and $\gamma_{\text{z}}$ are the regularization parameters pertaining to the linear systems (7) and (8), respectively.
- Equation (9) suggests that the weighting coefficients are given by

$$\mathbf{w}_{\text{BF-RLS}} = \frac{\mathbf{a}^H\mathbf{U}\left(\mathbf{\Sigma}^2 + \gamma_{\text{b}}\mathbf{I}\right)^{-1}\left(\mathbf{\Sigma}^2 + \gamma_{\text{z}}\mathbf{I}\right)^{-1}\mathbf{\Sigma}^2\mathbf{U}^H}{\mathbf{a}^H\mathbf{U}\left(\mathbf{\Sigma}^2 + \gamma_{\text{b}}\mathbf{I}\right)^{-2}\mathbf{\Sigma}^2\mathbf{U}^H\mathbf{a}}. \tag{10}$$

- Existing regularization methods can be used to find $\gamma_{\text{b}}$ and $\gamma_{\text{z}}$ in (10).
- We will introduce a new regularization approach called MVDR constrained perturbation regularization approach (MVDR-COPRA) that is based on exploiting the eigenvalue structure of $\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}$ in order to find $\gamma_{\text{b}}$ and $\gamma_{\text{z}}$ in (10).
- To this end, we replace $\mathbf{A}$ in (1) by $\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}$ to obtain the model

$$\mathbf{r} = \hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}\mathbf{x} + \mathbf{v}. \tag{11}$$

## 4. Proposed MVDR-COPRA

- As a form of regularization, we allow a perturbation $\mathbf{\Delta}$ into $\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}$.
- This perturbation is aimed to improve the eigenvalue structure of $\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}$.
- To maintain the balance between improving the eigenvalue structure and maintaining the fidelity of the model in (11), we add the constraint $||\mathbf{\Delta}||_2 \le \lambda$, $\lambda \in \mathbb{R}^+$.
- Thus, (11) is modified to

$$\mathbf{r} \approx \left(\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}} + \mathbf{\Delta}\right)\mathbf{x} + \mathbf{v}. \tag{12}$$

- Assuming that we know the *best* choice of $\lambda$, we consider minimizing the worst-case residual function of (12)

$$\min_{\hat{\mathbf{x}}}\max_{\mathbf{\Delta}} ||\mathbf{r} - \left(\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}} + \mathbf{\Delta}\right)\hat{\mathbf{x}}||_2 \quad \text{subject to } ||\mathbf{\Delta}||_2 \le \lambda. \tag{13}$$

- It can be shown that (13) is equivalent to

$$\min_{\hat{\mathbf{x}}} ||\mathbf{r} - \hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}\hat{\mathbf{x}}||_2 + \lambda\,||\hat{\mathbf{x}}||_2. \tag{14}$$

- The solution to (14) is given by:

$$\hat{\mathbf{x}} = \left(\hat{\mathbf{C}}_{\mathbf{yy}} + \gamma\mathbf{I}\right)^{-1}\hat{\mathbf{C}}_{\mathbf{yy}}^{\frac{1}{2}}\mathbf{r}, \tag{15}$$

where $\gamma$ is obtained by solving

$$\mathcal{G}(\gamma) = \mathbf{y}^T\mathbf{U}\left(\mathbf{\Sigma}^2 - \lambda^2\mathbf{I}\right)\left(\mathbf{\Sigma}^2 + \gamma\mathbf{I}\right)^{-2}\mathbf{U}^T\mathbf{y} = 0. \tag{16}$$

- The solution requires knowledge of $\lambda$, which we do not know.
- By taking the expectation to (16) we can manipulate to get

$$\lambda_{\text{o}}^2 = \frac{\sigma_{\mathbf{v}}^2\,\text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\right) + \text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\mathbf{\Sigma}^2\mathbf{U}^H\mathbf{C}_{\mathbf{xx}}\mathbf{U}\right)}{\sigma_{\mathbf{v}}^2\,\text{tr}\left(\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\right) + \text{tr}\left(\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\mathbf{\Sigma}^2\mathbf{U}^H\mathbf{C}_{\mathbf{xx}}\mathbf{U}\right)}. \tag{17}$$

- Divide $\mathbf{\Sigma}$ into $n_1$ large and $n_2$ small eigenvalues.
- Write $\mathbf{\Sigma} = \begin{bmatrix}\mathbf{\Sigma}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma}_2\end{bmatrix}$ and $\mathbf{U} = [\mathbf{U}_1\ \mathbf{U}_2]$.
  - $\mathbf{\Sigma}_1 \in \mathbb{C}^{n_1\times n_1}$ (large eigenvalues).
  - $\mathbf{\Sigma}_2 \in \mathbb{C}^{n_2\times n_2}$ (small eigenvalues).
  - $\|\mathbf{\Sigma}_2\|^2 \ll \|\mathbf{\Sigma}_1\|^2$.
- Apply the partitioning to (17), with some manipulations and reasonable approximations to get

$$\lambda_{\text{o}}^2 \approx \frac{\text{tr}\left(\mathbf{\Sigma}_1^2\left(\mathbf{\Sigma}_1^2 + \gamma_{\text{o}}\mathbf{I}_1\right)^{-2}\left(\mathbf{\Sigma}_1^2 + \frac{n_1\sigma_{\mathbf{v}}^2}{\text{tr}(\mathbf{C}_{\mathbf{xx}})}\mathbf{I}_1\right)\right)}{\text{tr}\left(\left(\mathbf{\Sigma}_1^2 + \gamma_{\text{o}}\mathbf{I}_1\right)^{-2}\left(\mathbf{\Sigma}_1^2 + \frac{n_1\sigma_{\mathbf{v}}^2}{\text{tr}(\mathbf{C}_{\mathbf{xx}})}\mathbf{I}_1\right)\right) + \frac{n_2}{\gamma_{\text{o}}^2}\frac{n_1\sigma_{\mathbf{v}}^2}{\text{tr}(\mathbf{C}_{\mathbf{xx}})}}. \tag{18}$$

- Problem: $\lambda_{\text{o}}$ dependents on $\sigma_{\mathbf{v}}^2$ and $\mathbf{C}_{\mathbf{xx}}$ which are not known.
- We apply the MSE criterion to eliminate this dependency and to set $\lambda_{\text{o}}$ that minimizes the MSE approximately.

## 5. Minimizing the MSE

- The MSE for an estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$ can be defined as

$$\text{MSE} = \text{tr}\left\{\mathbb{E}\left((\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^H\right)\right\}, \tag{19}$$

- We manipulate the MSE to the form:

$$\text{MSE} = \sigma_{\mathbf{v}}^2\text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma\mathbf{I}\right)^{-2}\right) + \gamma^2\text{tr}\left(\left(\mathbf{\Sigma}^2 + \gamma\mathbf{I}\right)^{-2}\mathbf{U}^H\mathbf{C}_{\mathbf{xx}}\mathbf{U}\right). \tag{20}$$

- The first derivative can be obtained as

$$\frac{\partial\,(\text{MSE})}{\partial\,\gamma} = -2\sigma_{\mathbf{v}}^2\text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma\mathbf{I}\right)^{-3}\right) + 2\gamma\ \text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma\mathbf{I}\right)^{-3}\mathbf{U}^H\mathbf{C}_{\mathbf{xx}}\mathbf{U}\right) = 0. \tag{21}$$

- Solving (21) does not provide a closed-form expression for $\gamma_{\text{o}}$. By using an approximation we obtain

$$\gamma_{\text{o}} \approx \frac{n_e\sigma_{\mathbf{v}}^2}{\text{tr}\left(\mathbf{C}_{\mathbf{xx}}\right)}. \tag{22}$$

- From (18) we can replace the following

$$\frac{n_1\sigma_{\mathbf{v}}^2}{\text{tr}\left(\mathbf{C}_{\mathbf{xx}}\right)} \to \frac{n_1}{n_e}\gamma_{\text{o}}.$$

- Substitute this results in (18) and then substitute (18) in (16).
- MVDR-COPRA characteristic equation:

$$\begin{aligned} S\left(\gamma_{\text{o}}\right) &= \text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\mathbf{d}\mathbf{d}^H\right)\text{tr}\left(\left(\mathbf{\Sigma}_1^2 + \gamma_{\text{o}}\mathbf{I}_1\right)^{-2}\left(\beta\mathbf{\Sigma}_1^2 + \gamma_{\text{o}}\mathbf{I}_1\right)\right) \\ &- \text{tr}\left(\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\mathbf{d}\mathbf{d}^H\right)\text{tr}\left(\mathbf{\Sigma}_1^2\left(\mathbf{\Sigma}_1^2 + \gamma_{\text{o}}\mathbf{I}_1\right)^{-2}\left(\beta\mathbf{\Sigma}_1^2 + \gamma_{\text{o}}\mathbf{I}_1\right)\right) \\ &+ \frac{n_2}{\gamma_{\text{o}}}\text{tr}\left(\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \gamma_{\text{o}}\mathbf{I}\right)^{-2}\mathbf{d}\mathbf{d}^H\right) = 0, \end{aligned} \tag{23}$$

where $\mathbf{d} \triangleq \mathbf{U}^T\mathbf{y}$, and $\beta \triangleq \frac{n_e}{n_1}$.

## 6. Simulation Results

- Setup:
  - Uniform linear array with 10 elements placed at half of the wavelength of the signal of interest and two interfering signals.
  - The directions of arrival (DOA) for the signal of interest and the interference are generated from a uniform distribution in the interval $[-90^{\text{o}}, 90^{\text{o}}]$.
  - The steering vector $\mathbf{a}$ is calculated from the true DOA of the signal of interest plus a uniformly distributed error in the interval $[-5^{\text{o}}, 5^{\text{o}}]$.

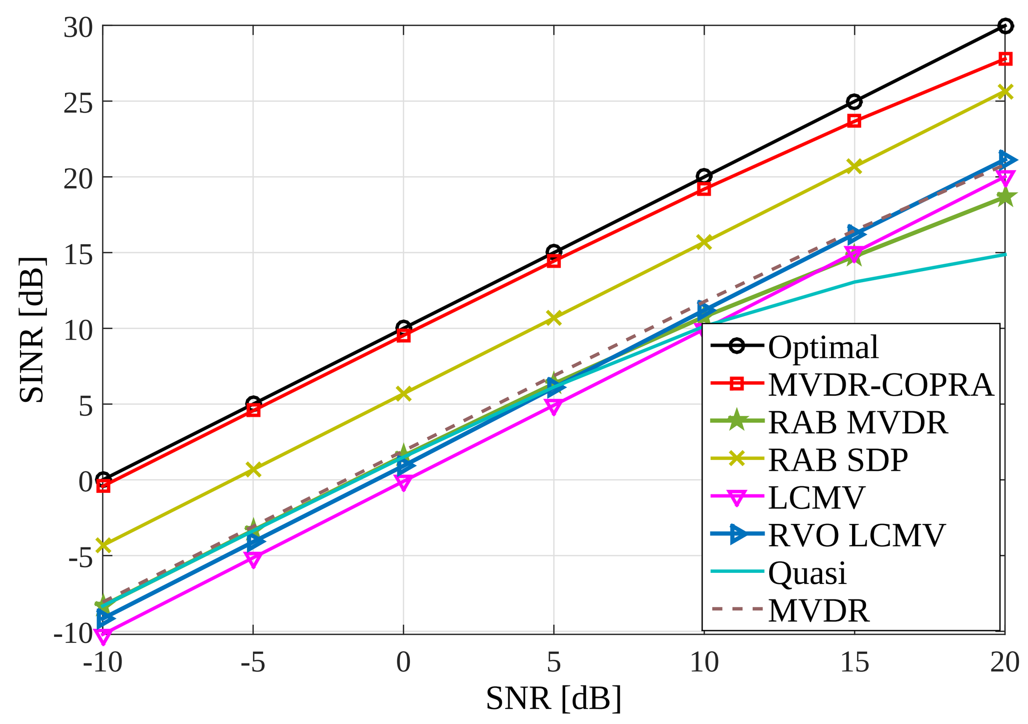


**Figure 1:** *Output SINR vs input SNR for* $n_s = 30$.

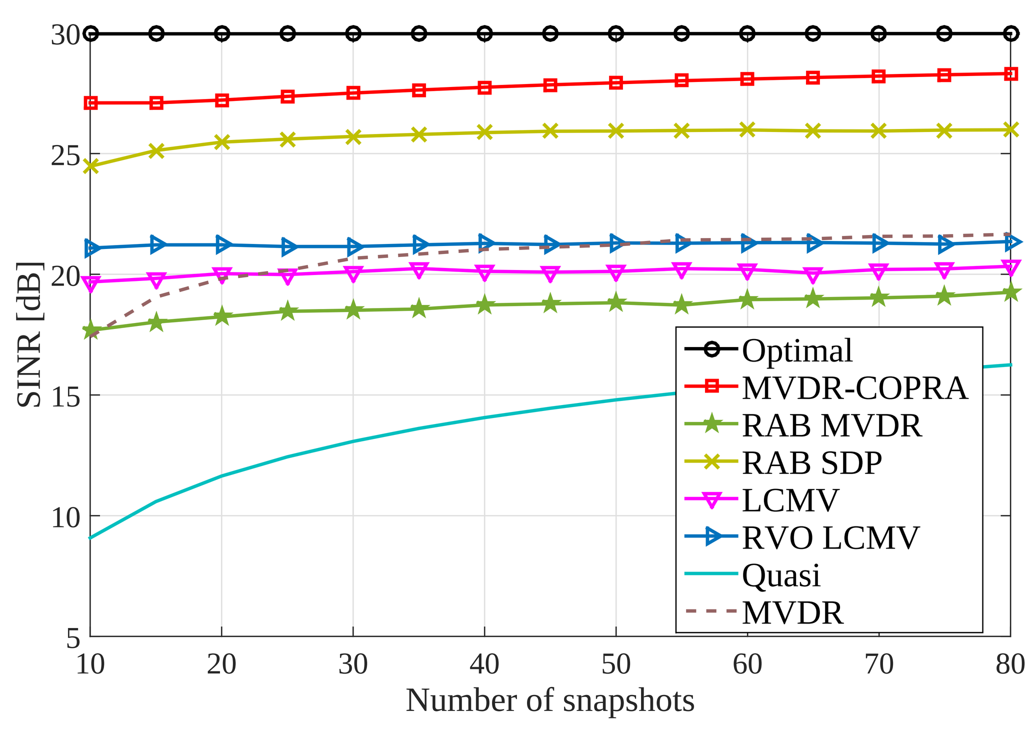


**Figure 2:** *Output SINR vs number of snapshots at SNR = 20 dB.*

## 7. Conclusions

- The robust MVDR beamforming is converted to a pair of linear estimation problems with ill-conditioned matrices and new regularization method is proposed to solve these problems.
- Simulations demonstrate that the proposed approach outperforms a number of benchmark methods in terms of SINR.