A Recursive Least-Squares Algorithm Based on the Nearest Kronecker Product Decomposition

Camelia Elisei-Iliescu[1], *Constantin Paleologu*[1], *Jacob Benesty*[2], *and Silviu Ciochină*[1]

[1]Department of Telecommunications
University Politehnica of Bucharest, Romania
{pale, silviu}@comm.pub.ro

[2]INRS-EMT
University of Quebec, Montreal, Canada
benesty@emt.inrs.ca

Outline

- Introduction
- System Model
- RLS algorithm based on the nearest Kronecker product decomposition (RLS-NKP)
- Simulation Results
- Conclusions and Perspectives

Introduction

- **Recursive least-square (RLS) algorithm → frequently used in system identification problems**

  → this algorithm is computationally very complex

- **In this work → new approach to improve the efficiency of the RLS** → the impulse response decomposition based on the nearest Kronecker product

  → low-rank approximation

- **Target:** a high-dimension system identification problem

  ⇓

  low-dimension problems

→ RLS algorithm based on the nearest Kronecker product decomposition

System Model

Model

$$d(t) = \mathbf{h}^T \mathbf{x}(t) + w(t)$$

where $d(t)$ - desired signal
$w(t)$ - additive noise

→ $\mathbf{h}$ is the impulse response of the unknown system of length $L = L_1 L_2 (L_1 \geq L_2)$.

→ The impulse response can be decomposed as:

$$\mathbf{h} = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{s}_2^T & \ldots & \mathbf{s}_{L_2}^T \end{bmatrix}^T$$

where $\mathbf{s}_l, l = 1, 2, \ldots, L_2$ - short impulse responses (of length L_1)

→ ? $\mathbf{h}$ can be approximated as $\mathbf{h}_2 \otimes \mathbf{h}_1$

($\mathbf{h}_2$- length L_2, $\mathbf{h}_1$- length L_1)

→ The normalized misalignment:

$$\mathcal{M}(\mathbf{h}_1, \mathbf{h}_2) = \frac{\|\mathbf{h} - \mathbf{h}_2 \otimes \mathbf{h}_1\|_2}{\|\mathbf{h}\|_2}$$

System Model

→ We can reorganize the components of $\mathbf{h}$ into a matrix:

$$\mathbf{H} = \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \dots & \mathbf{s}_{L_2} \end{bmatrix}$$

$$\mathcal{M}(\mathbf{h}_1, \mathbf{h}_2) = \frac{\left\| \mathbf{H} - \mathbf{h}_1 \mathbf{h}_2^T \right\|_\mathrm{F}}{\|\mathbf{H}\|_\mathrm{F}}$$

→ The optimal values of $\mathbf{h}_1$ and $\mathbf{h}_2$ ⇨ minimization of $\mathcal{M}(\mathbf{h}_1, \mathbf{h}_2)$

→ Minimizing $\mathcal{M}(\mathbf{h}_1, \mathbf{h}_2)$ ⇔ finding the nearest rank-1 matrix to $\mathbf{H}$

$$\mathbf{H} = \mathbf{U}_1 \Sigma \mathbf{U}_2^T = \sum_{l=1}^{L_2} \sigma_l \mathbf{u}_{1,l} \mathbf{u}_{2,l}^T$$

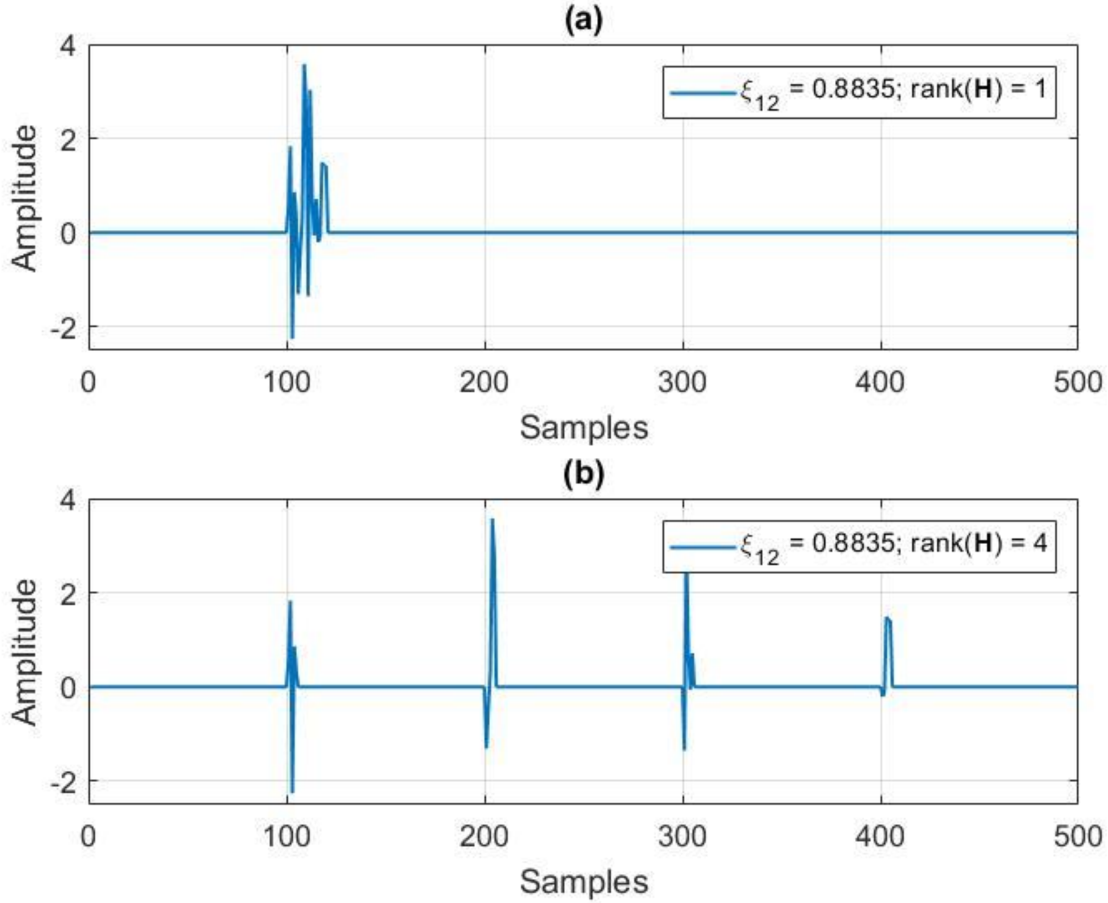
$$\mathrm{h} \approx \sum_{p=1}^{P} \mathbf{h}_{2,p} \otimes \mathbf{h}_{1,p}$$ where $P \leq L_2$

$$\overline{\mathbf{h}}(P) \approx \sum_{p=1}^{P} \overline{\mathbf{h}}_{2,p} \otimes \overline{\mathbf{h}}_{1,p} = \sum_{p=1}^{P} \sigma_p \mathbf{u}_{2,p} \otimes \mathbf{u}_{1,p}$$

L_2 L_1

System Model

$$\xi_{12}(\mathbf{h}) = \frac{L}{L-\sqrt{L}}\left(1 - \frac{\|\mathbf{h}\|_1}{\sqrt{L}\|\mathbf{h}\|_2}\right)$$



Impulse responses with $L = 500 (L_1 = 25, L_2 = 20)$:
(a) $\xi_{12} = 0.8835$ and $\text{rank}(\mathbf{H}) = 1$; (b) $\xi_{12} = 0.8835$ and $\text{rank}(\mathbf{H}) = 4$.

RLS algorithm based on the nearest Kronecker product decomposition

→ The goal is to estimate $\mathbf{h}$ with an adaptive filter $\hat{\mathbf{h}}(t)$

→ The error signal:

$$e(t) = d(t) - \hat{y}(t) = d(t) - \hat{\mathbf{h}}^T(t-1)\mathbf{x}(t)$$

→ We can decompose the adaptive filter:

$$\hat{\mathbf{h}}(t) = \sum_{p=1}^{P} \hat{\mathbf{h}}_{2,p}(t) \otimes \hat{\mathbf{h}}_{1,p}(t)$$

L_2 L_1

$$e(t) = d(t) - \sum_{p=1}^{P} \hat{\mathbf{h}}_{1,p}^T(t-1)\mathbf{x}_{2,p}(t) = d(t) - \underline{\hat{\mathbf{h}}}_1^T(t-1)\underline{\mathbf{x}}_2(t)$$

$$e(t) = d(t) - \sum_{p=1}^{P} \hat{\mathbf{h}}_{2,p}^T(t-1)\mathbf{x}_{1,p}(t) = d(t) - \underline{\hat{\mathbf{h}}}_2^T(t-1)\underline{\mathbf{x}}_1(t)$$

$$\mathbf{x}_{2,p} = \left[\hat{\mathbf{h}}_{2,p}(t-1) \otimes \mathbf{I}_{L_1}\right]^T \mathbf{x}(t) \qquad \underline{\hat{\mathbf{h}}}_1(t) = \left[\hat{\mathbf{h}}_{1,1}^T(t) \ \ \hat{\mathbf{h}}_{1,2}^T(t) \ \ \ldots \ \ \hat{\mathbf{h}}_{1,P}^T(t)\right]^T$$

$$\mathbf{x}_{1,p} = \left[\mathbf{I}_{L_2} \otimes \hat{\mathbf{h}}_{1,p}(t-1)\right]^T \mathbf{x}(t) \qquad \underline{\hat{\mathbf{h}}}_2(t) = \left[\hat{\mathbf{h}}_{2,1}^T(t) \ \ \hat{\mathbf{h}}_{2,2}^T(t) \ \ \ldots \ \ \hat{\mathbf{h}}_{2,P}^T(t)\right]^T$$

$$\underline{\mathbf{x}}_1(t) = \left[\mathbf{x}_{1,1}^T(t) \ \ \mathbf{x}_{1,2}^T(t) \ \ \ldots \mathbf{x}_{1,P}^T(t)\right]^T \qquad \underline{\mathbf{x}}_2(t) = \left[\mathbf{x}_{2,1}^T(t) \ \ \mathbf{x}_{2,2}^T(t) \ \ \ldots \mathbf{x}_{2,P}^T(t)\right]^T$$

RLS algorithm based on the nearest Kronecker product decomposition

→ The **cost functions**:

$$\mathcal{J}_{\underline{\hat{\mathbf{h}}}_2}\left[\underline{\hat{\mathbf{h}}}_1(t)\right] = \sum_{i=1}^{t} \lambda_1^{t-i}\left[d(i) - \underline{\hat{\mathbf{h}}}_1^T(t)\underline{\mathbf{x}}_2(i)\right]^2$$

$$\mathcal{J}_{\underline{\hat{\mathbf{h}}}_1}\left[\underline{\hat{\mathbf{h}}}_2(t)\right] = \sum_{i=1}^{t} \lambda_2^{t-i}\left[d(i) - \underline{\hat{\mathbf{h}}}_2^T(t)\underline{\mathbf{x}}_1(i)\right]^2$$

λ_1, λ_2 - forgetting factors

Normal equations

$$\underline{\mathbf{R}}_2(t)\underline{\hat{\mathbf{h}}}_1(t) = \underline{\mathbf{p}}_2(t)$$

$$\underline{\mathbf{R}}_1(t)\underline{\hat{\mathbf{h}}}_2(t) = \underline{\mathbf{p}}_1(t)$$

where $\underline{\mathbf{R}}_2(t) = \lambda_1\underline{\mathbf{R}}_2(t-1) + \underline{\mathbf{x}}_2(t)\underline{\mathbf{x}}_2^T(t)$ $\quad \underline{\mathbf{p}}_2(t) = \lambda_1\underline{\mathbf{p}}_2(t-1) + \underline{\mathbf{x}}_2(t)d(t)$

$\underline{\mathbf{R}}_1(t) = \lambda_2\underline{\mathbf{R}}_1(t-1) + \underline{\mathbf{x}}_1(t)\underline{\mathbf{x}}_1^T(t)$ $\quad \underline{\mathbf{p}}_1(t) = \lambda_2\underline{\mathbf{p}}_1(t-1) + \underline{\mathbf{x}}_1(t)d(t)$

→ **The RLS-NKP:**

$$\underline{\hat{\mathbf{h}}}_1(t) = \underline{\hat{\mathbf{h}}}_1(t-1) + \mathbf{k}_2(t)e(t)$$

$$\underline{\hat{\mathbf{h}}}_2(t) = \underline{\hat{\mathbf{h}}}_2(t-1) + \mathbf{k}_1(t)e(t)$$

RLS algorithm based on the nearest Kronecker product decomposition

→ The Kalman gain vectors:

$$\mathbf{k}_2(t) = \frac{\underline{\mathbf{R}}_2^{-1}(t-1)\underline{\mathbf{x}}_2(t)}{\lambda_1 + \underline{\mathbf{x}}_2^T(t)\underline{\mathbf{R}}_2^{-1}(t-1)\underline{\mathbf{x}}_2(t)}$$

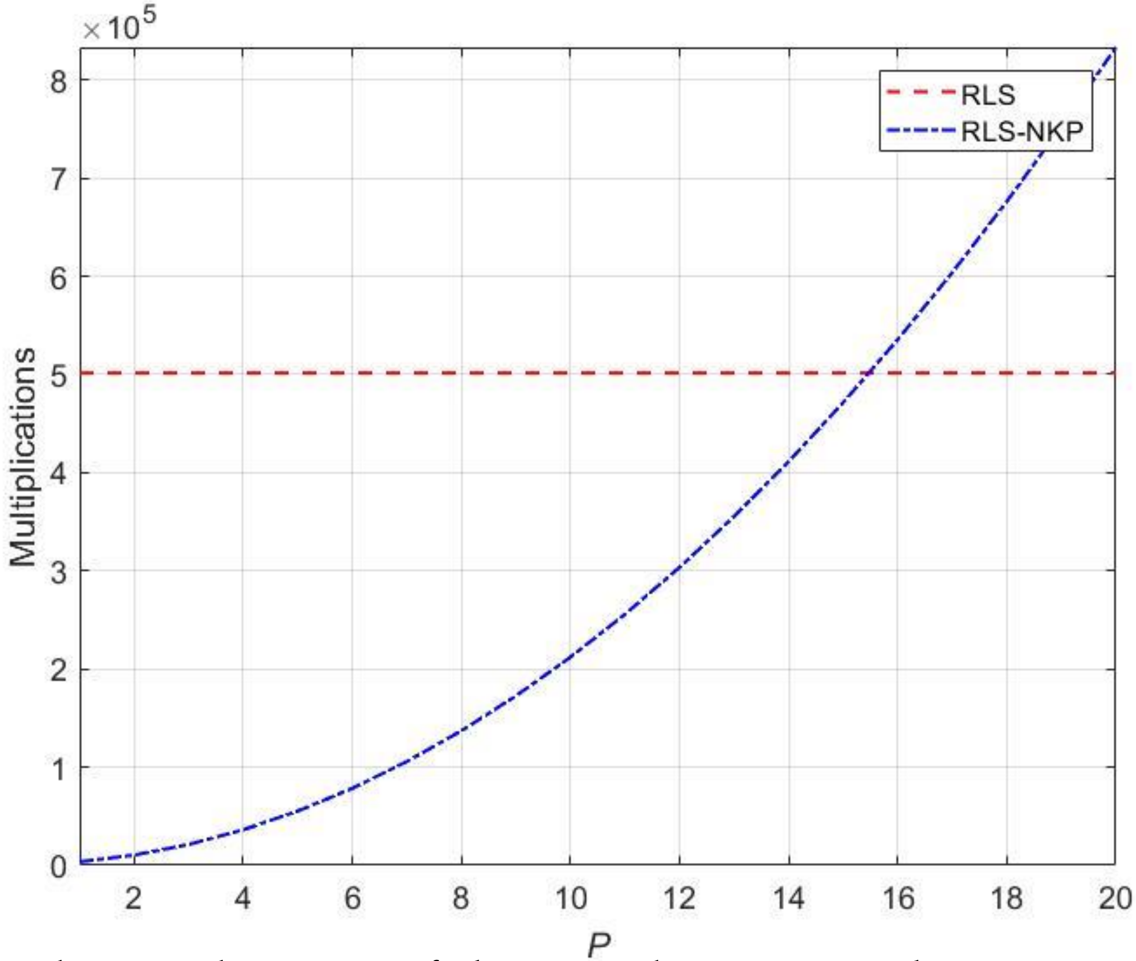
$$\mathbf{k}_1(t) = \frac{\underline{\mathbf{R}}_1^{-1}(t-1)\underline{\mathbf{x}}_1(t)}{\lambda_2 + \underline{\mathbf{x}}_1^T(t)\underline{\mathbf{R}}_1^{-1}(t-1)\underline{\mathbf{x}}_1(t)}$$

→ The updates of $\underline{\mathbf{R}}_1^{-1}(t)$ and $\underline{\mathbf{R}}_2^{-1}(t)$ (based on the inversion lemma):

$$\underline{\mathbf{R}}_2^{-1}(t) = \lambda_1^{-1}\left[\underline{\mathbf{R}}_2^{-1}(t-1) - \mathbf{k}_2(t)\underline{\mathbf{x}}_2^T(t)\underline{\mathbf{R}}_2^{-1}(t-1)\right]$$

$$\underline{\mathbf{R}}_1^{-1}(t) = \lambda_2^{-1}\left[\underline{\mathbf{R}}_1^{-1}(t-1) - \mathbf{k}_1(t)\underline{\mathbf{x}}_1^T(t)\underline{\mathbf{R}}_1^{-1}(t-1)\right]$$

RLS algorithm based on the nearest Kronecker product decomposition



Computational complexities of the regular RLS and RLS-NKP algorithms
($L = 500$, with $L_1 = 25$ and $L_2 = 20$)

Simulation Results

- **conditions**

→ **h** – echo paths from G168 Recommendation, random impulse responses (L = 500), and an acoustic echo path (L = 1024)

→ input signals – AR1(0.9) process/ speech sequence

→ additive noise $w(t)$ – WGN (SNR=20 dB)

→ measures of performance: normalized misalignment (NM)

$$\mathrm{NM[dB]} = 20\log_{10}\frac{\left\|\mathbf{h}-\hat{\mathbf{h}}(t)\right\|_2}{\|\mathbf{h}\|_2}$$

- **algorithms**

→ proposed RLS algorithm based on the nearest Kronecker product decomposition – RLS-NKP($\lambda_1 = 1 - 1/[K(PL_1)]$, $\lambda_2 = 1 - 1/[K(PL_2)]$)

→ regular RLS [$\lambda = 1 - 1/(KL)$, with $K > 1$]

→ RLS-DCD [Y. V. Zakharov, *Low-Complexity RLS using dichotomous coordinate descent iterations*, IEEE Trans. Signal Process., 2008]

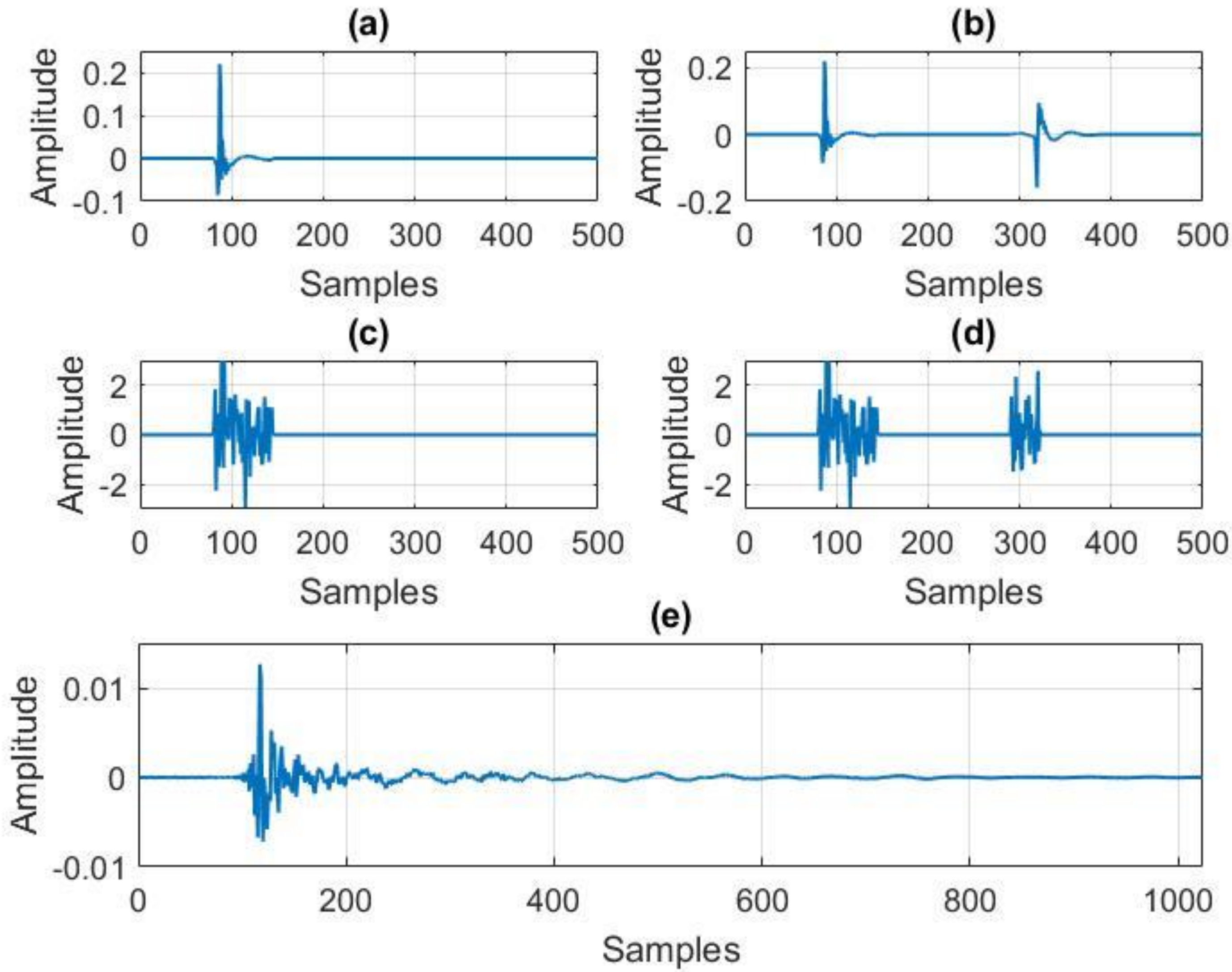


Fig. 1. Impulse responses used in the experiments: (a) $L = 500, \xi_{12} = 0.8957$, (b) $L = 500, \xi_{12} = 0.8080$, (c) $L = 500, \xi_{12} = 0.7549$, (d) $L = 500, \xi_{12} = 0.6867$, and (e) $L = 1024, \xi_{12} = 0.6880$.

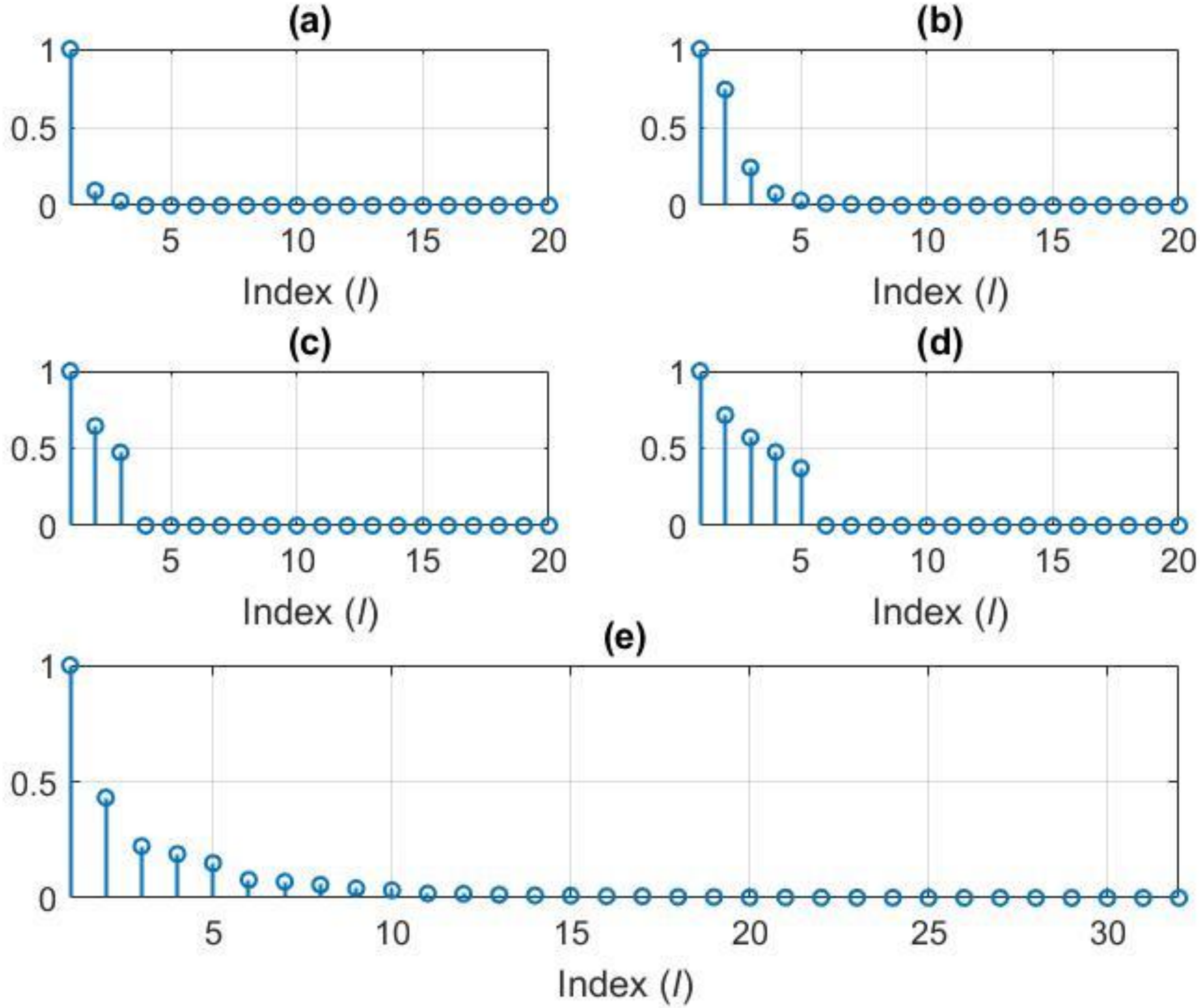


Fig. 2. Singular values (normalized with respect to the maximum one) of the matrix **H** for the corresponding impulse responses from Fig.1. The size of matrix **H** is $L_1 \times L_2$. (a)-(d) $L_1 = 25$ and $L_2 = 20$; (e) $L_1 = L_2 = 32$.

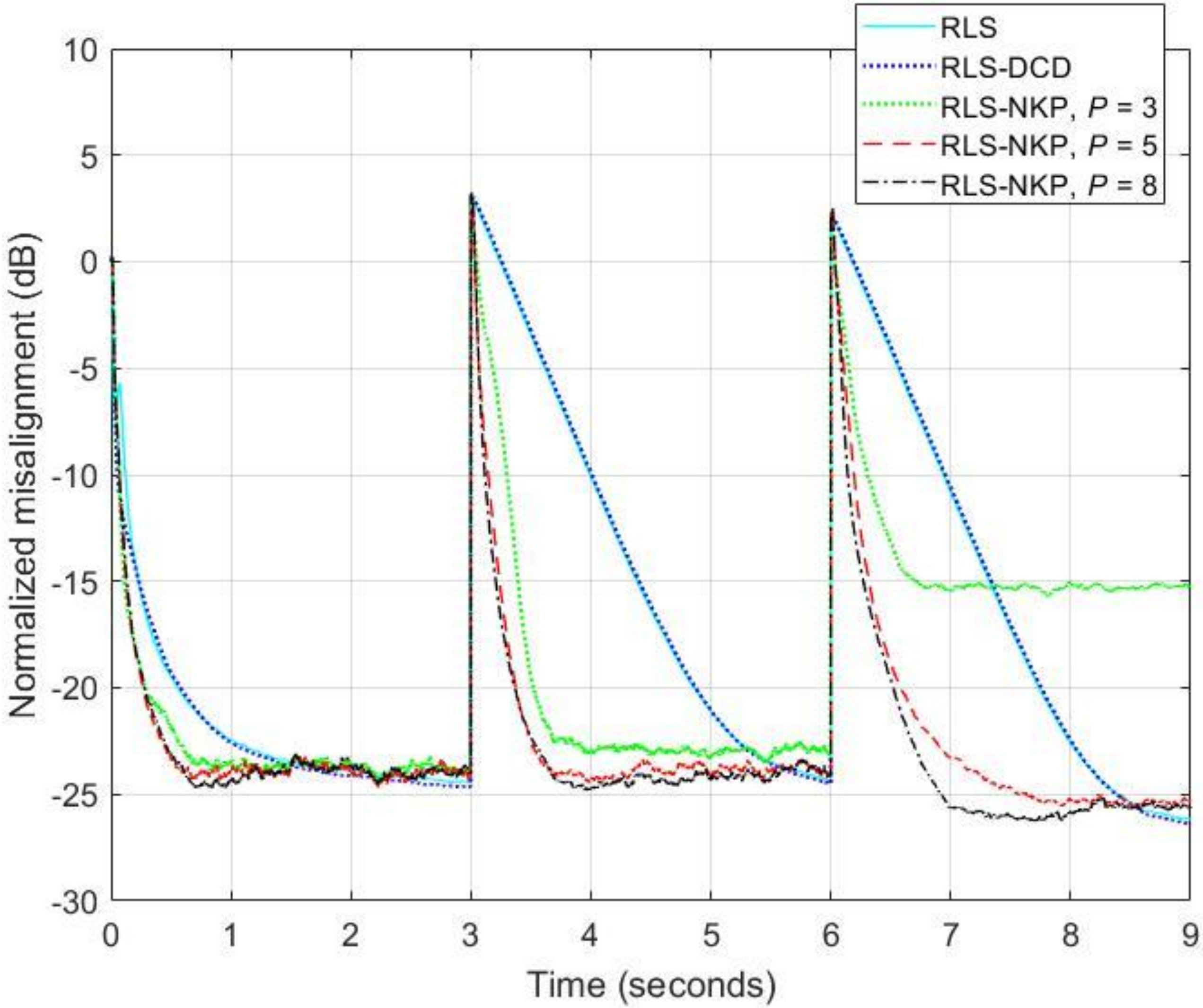


Fig. 3. Normalized misalignment of the regular RLS and RLS-DCD algorithms ($L = 500$), and RLS-NKP algorithm (using $L_1 = 25, L_2 = 20,$ and $\mathrm{P} < L_2$), for the identification of the impulse responses from Figs. 1(a) and (b). The input signal is an AR(1) process and the impulse response changes at times 3 and 6 seconds.

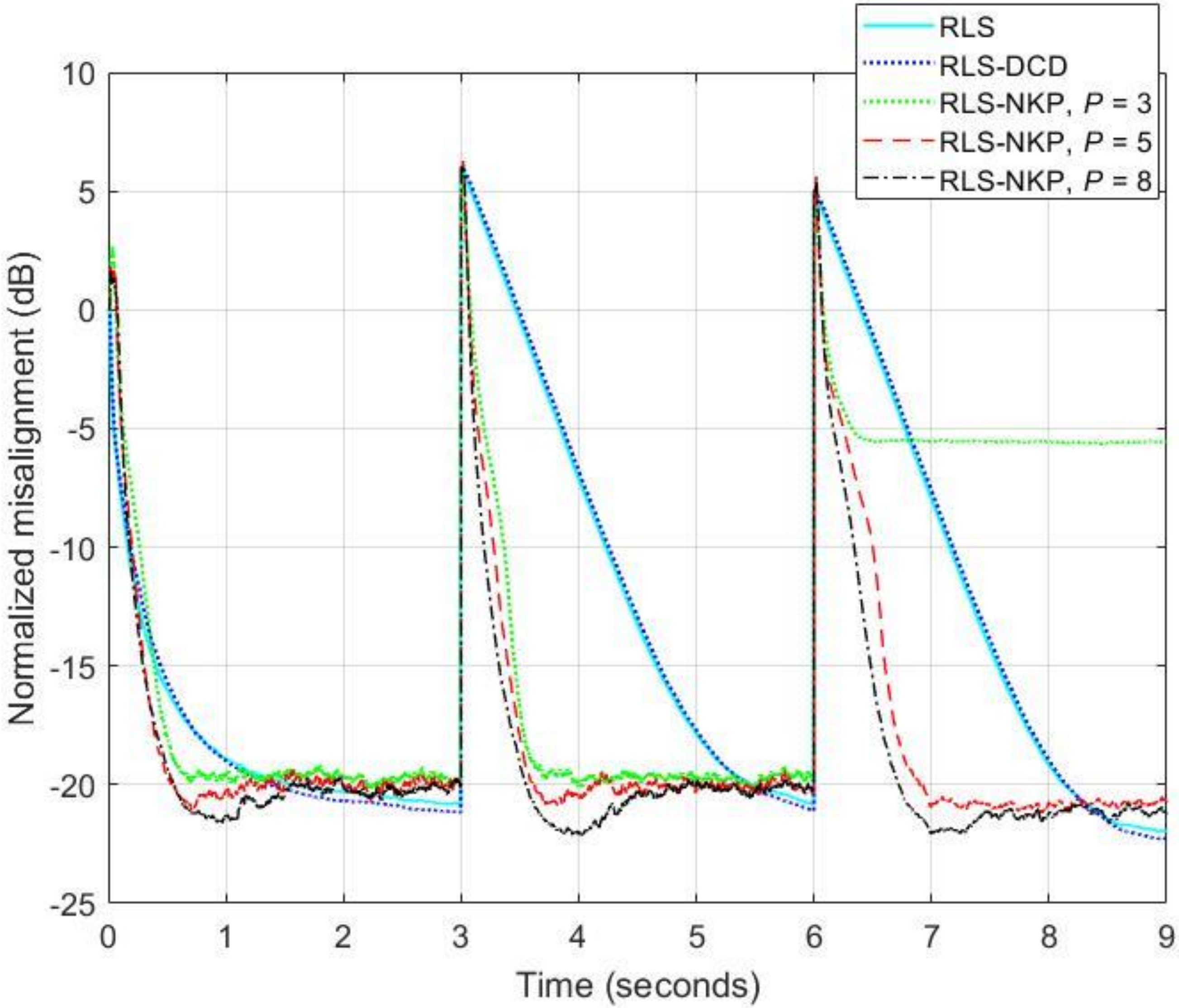


Fig. 4. Normalized misalignment of the regular RLS and RLS-DCD algorithms ($L = 500$), and RLS-NKP algorithm (using $L_1 = 25, L_2 = 20,$ and $P < L_2$), for the identification of the impulse responses from Figs. 1(c) and (d). The input signal is an AR(1) process and the impulse response changes at times 3 and 6 seconds.

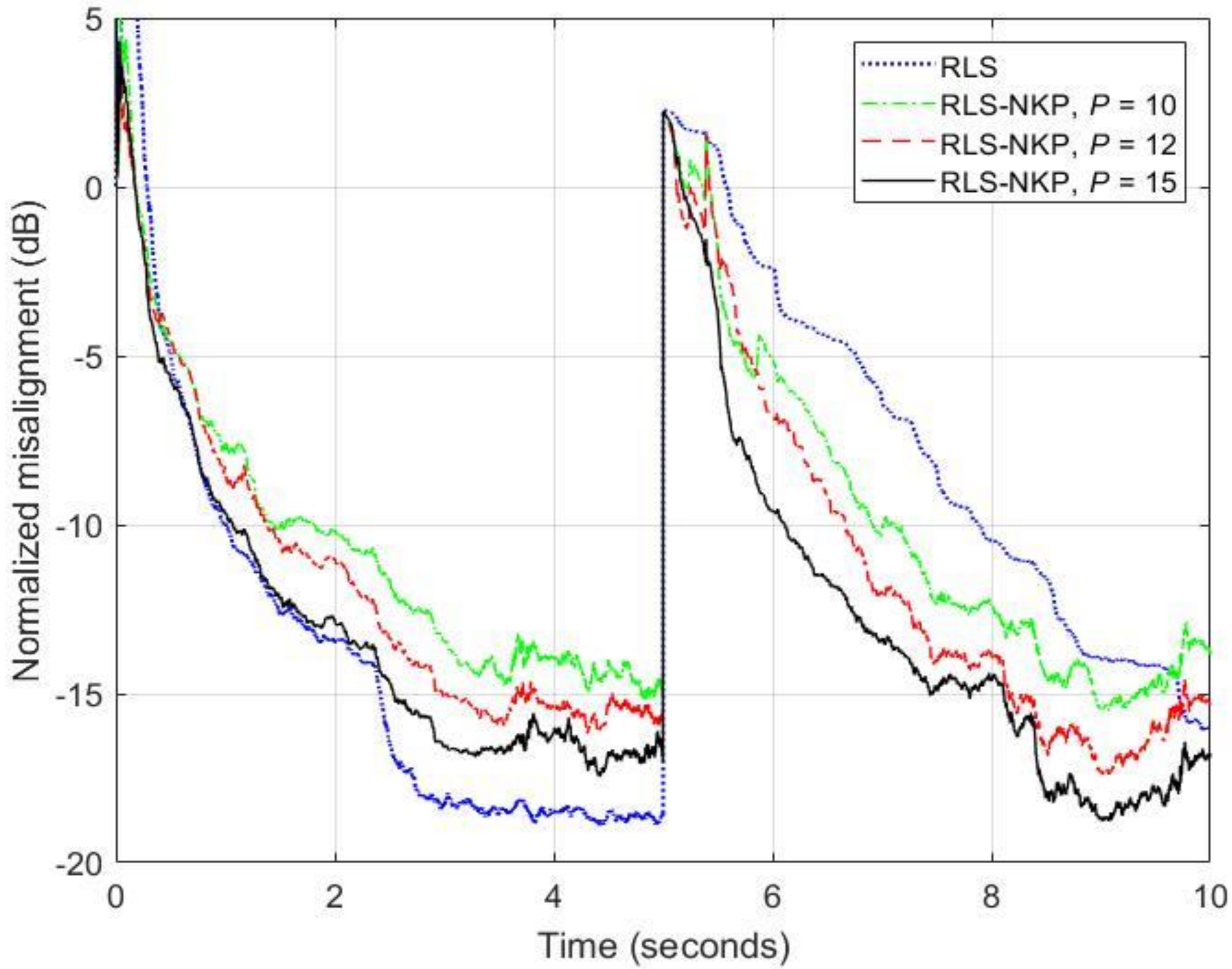


Fig. 5. Normalized misalignment of the RLS algorithm $(L = 1024)$ and RLS-NKP algorithm (using $L_1 = L_2 = 32,$ and P $< L_2$), for the identification of the impulse responses from Figs. 1(e) . The input signal is a speech sequence and the impulse response changes at time 5 seconds.

Conclusions and Perspectives

- We have proposed the RLS-NKP algorithm.
- Suitable for the identification of low-rank models, like the echo paths.
- The tracking capabilities of the of the RLS-NKP algorithm are better as compared to the conventional RLS algorithm.
- The computational complexity of the proposed algorithm could be much lower as compared to the RLS.

C. Elisei-Iliescu, C. Paleologu, J. Benesty, C. Stanciu, C. Anghel, and S. Ciochină, "Recursive Least-Squares Algorithms for the Identification of Low-Rank Systems," *IEEE/ACM Trans. Audio, Speech, Language Process.*, May 2019.

Thank you for your attention!