

COMPETIȚIA DE PROIECTE DE CERCETARE A ACADEMIEI OAMENILOR
DE ȘTIINȚĂ DIN ROMÂNIA “AOȘR-TEAMS-III”
EDIȚIA 2024-2025

Dezvoltarea de soluții iterative îmbunătățite bazate pe descompuneri tensoriale
pentru aplicații în domeniul acustic

RAPORT DE CERCETARE ȘTIINȚIFICĂ

Rezumatul și rezultatele proiectului

Scopul principal al acestui proiect de cercetare a constat în dezvoltarea unor soluții iterative îmbunătățite bazate pe descompuneri tensoriale, pentru aplicații din domeniul acustic. Rezultatul global a constat în obținerea și analiza unor algoritmi de filtrare adaptivă bazați pe descompuneri tensoriale de ordin superior, ce au fost validați în cadrul unor scenarii experimentale de compensare a ecoului acustic. Astfel, au fost îndeplinite toate obiectivele specifice prevăzute în planul de realizare asociat proiectului, după cum urmează:

1. Implementarea descompunerilor tensoriale de ordin superior în modelarea sistemelor utilizate pentru aplicații acustice.
2. Dezvoltarea unor algoritmi adaptivi bazați pe descompuneri tensoriale de ordin superior.
3. Îmbunătățirea performanțelor și studiul de convergență al algoritmilor adaptivi bazați pe descompuneri tensoriale de ordin superior.
4. Evaluarea performanțelor metodelor iterative utilizate în aplicații acustice utilizând simulări software.

Rezultatele din cadrul acestui proiect de cercetare au fost publicate în lucrările [1]–[8] din lista de referințe. Acestea constau în 4 articole publicate în reviste cotate **ISI-WOS**, dintre care un articol **Q1** și 3 articole **Q2**, precum și 4 lucrări ce apar în volumele unor conferințe indexate în baza de date internațională *IEEE Xplore*.

1 Introducere

În numeroase aplicații, filtrele adaptive de orice fel joacă un rol fundamental [9]. Practic, un astfel de sistem adaptiv constă într-un filtru digital cu coeficienți variabili în timp, care sunt actualizați pe baza unei metode iterative, cunoscută și sub numele de algoritm adaptiv. Alegerea unui anumit algoritm adaptiv pentru o aplicație specifică depinde de mai multe criterii de performanță, precum rata de convergență, acuratețea soluției, robustețea și eficiența computațională. Deși algoritmul de tipul *least-mean-square* (LMS) rămâne una dintre cele mai populare alegeri datorită simplității sale, performanța îmbunătățită a algoritmului *recursive least-squares* (RLS) și a filtrului Kalman justifică utilizarea lor pentru aplicațiile actuale.

Lungimea filtrului adaptiv (adică numărul de coeficienți) este un factor critic în implementarea practică a unui astfel de sistem. Cu cât filtrul este mai lung, cu atât implementarea/funcționarea lui devine mai dificilă. Acest aspect nu este asociat doar cu complexitatea computațională, va crește în mod natural odată cu numărul de coeficienți, ci și cu principalele criterii de performanță. Un filtru mai lung implică o rată de convergență și o urmărire mai lente, influențând totodată acuratețea soluției. Deoarece există multe aplicații care necesită filtre adaptive cu sute de coeficienți (de exemplu, compensarea ecoului, modelarea canalului, reducerea zgomotului etc.), gestionarea unor astfel de răspunsuri la impuls de lungime mare reprezintă o provocare importantă, mai ales în combinație cu algoritmi bazați pe RLS [10].

Lucrări recente au abordat această problemă concentrându-se pe tehnici de descompunere a răspunsului la impuls și urmăresc să reformuleze o problemă cu un spațiu mare de parametri (adică un număr mare de coeficienți ai filtrului) ca o combinație de probleme ”mai mici”, adică filtre mai scurte. În acest cadru, una dintre cele mai atractive soluții se bazează pe decompunerea răspunsului la impuls folosind cel mai apropiat produs Kronecker (*nearest Kronecker product* - NKP), care exploatează caracteristicile de rang scăzut ce caracterizează majoritatea sistemelor din lumea reală [11]–[14].

Baza abordării NKP a fost dezvoltată în contextul unei descompunerii de ordinul al doilea [11], [12], unde un răspuns la impuls de lungime $L = L_1 L_2$ (cu $L_2 \leq L_1$) este aproximat printr-o sumă de R produse Kronecker (cu $R \ll L_2$) între răspunsuri la impuls mai scurte, de lungimi L_2 și L_1 . De fapt, R este rangul aproximat al matricii $L_1 \times L_2$ care rezultă prin remodelarea răspunsului la impuls de lungime L . Se pare că o descompunere de ordin superior este de dorit, deoarece ar implica răspunsuri la impuls și mai scurte. Cu toate acestea, în acest caz, răspunsul la impuls ar fi remodelat ca un tensor de ordin superior, iar rangul unei astfel de structuri nu este întotdeauna ușor de gestionat. Soluția propusă în [13] evită această problemă și controlează rangul unui tensor de ordinul al treilea printr-un rang al unei matrici. Ca urmare, algoritmul bazat pe RLS care folosește o astfel de descompunere tensorială de ordinul al treilea (*third-order tensor* - TOT), și anume RLS-TOT [14], îl depășește pe omologul său dezvoltat anterior, care exploatează descompunerea NKP de ordinul al doilea [12]. O idee similară a dus la dezvoltarea filtrului Kalman bazat pe descompunerea TOT [1]. Abordarea poate fi extinsă și la o descompunere de ordinul al patrulea (FOT) [5].

Pe lângă criteriile de performanță menționate anterior (de exemplu convergență, urmărire și acuratețe), există un alt aspect important care trebuie luat în considerare atunci când sistemul funcționează în medii zgomotoase. În astfel de scenarii, robustețea algoritmului adaptiv devine o problemă critică. De exemplu, în compensarea ecouului, microfonul captează zgomotul de fundal, care poate fi puternic și nestaționar. Prin urmare, prezența sa ar putea influența semnificativ comportamentul adaptiv al filtrului, dacă algoritmul nu este echipat cu caracteristici bune de robustețe. De obicei, robustețea algoritmului este luată în considerare în faza de proiectare, prin includerea unor termeni specifici în funcția de cost de optimizare. Cu toate acestea, algoritmi regulatiți rezultă frecvent parametri suplimentari care nu sunt întotdeauna ușor de controlat în practică. În acest context, soluțiile propuse recent în [6] și [8] iau în considerare o funcție de cost care include atât contribuția zgomotului extern, cât și pe cea a incertitudinilor sistemului, iar varianțele acestor termeni sunt estimate într-un mod practic în cadrul algoritmilor.

În cadrul acestui proiect a fost urmarită metoda de descompunere a tensorilor de ordin superior și au fost proiectați algoritmi de filtrare adaptivă robusți la diverse categorii de perturbații [1]–[8]. Algoritmi rezultă actualizează în paralel și combină coeficienții a mai multe filtre adaptive de lungimi mult mai mici (în comparație cu lungimea răspunsului la impuls original), obținând astfel un câștig important atât în ceea ce privește reducerea complexității, cât și creșterea performanței. Caracteristica principală a acestei abordări este că rangul tensorului este controlat și limitat la o valoare cunoscută, fără a necesita tehnici de aproximare. Această metodă de descompunere este potrivită pentru identificarea răspunsurilor la impuls de lungime mare, deoarece operează cu structuri de date mai mici. De asemenea, tehnica de regularizare utilizată contribuie la creșterea robusteții algoritmilor în medii afectate de zgomot de nivel mare și/sau nestaționar.

2 Algoritmi de filtrare adaptivă bazați pe descompuneri tensoriale de ordin superior

În această secțiune este prezentat contextul din spatele descompunerii TOT, împreună cu cadrul și notația specifice. Să considerăm un sistem variabil în timp, caracterizat prin răspunsul la impuls de lungime L reprezentat prin vectorul:

$$\mathbf{h}(n) = [h_1(n) \quad h_2(n) \quad \cdots \quad h_L(n)]^\top, \quad (1)$$

unde n reprezintă indicele de timp discret, $h_l(n)$ este coeficientul al l -lea (cu $l = 1, 2, \dots, L$), iar superscriptul \top reprezintă transpoziția. Se presupune că lungimea poate fi factorizată ca $L = L_1 L_2$, cu $L_1 \geq L_2$. Aceasta poate să nu pară o presupunere trivială, dar în multe implementări cu precizie finită lungimea unui filtru digital este de obicei setată ca putere a lui doi (din motive practice/algoritmice), adică $L = 2^B$, unde B este un întreg pozitiv. Prin urmare, poate fi ușor factorizată ca $L = 2^{B_1} 2^{B_2}$, cu $B = B_1 + B_2$ și $B_1 \geq B_2$ (unde B_1 și B_2 sunt, de asemenea, numere întregi pozitive).

Folosind produsul Kronecker, răspunsul la impuls din (1) poate fi exprimat echivalent ca [11]

$$\mathbf{h}(n) = \sum_{i=1}^{L_2} \mathbf{h}_{2,i}(n) \otimes \mathbf{h}_{1,i}(n), \quad (2)$$

unde $\mathbf{h}_{2,i}(n)$ și $\mathbf{h}_{1,i}(n)$, cu $i = 1, 2, \dots, L_2$, sunt răspunsuri la impuls mai scurte, de lungimi L_2 și respectiv L_1 , în timp ce \otimes denotă produsul Kronecker [16]. Așa cum se arată în [11], coeficienții acestor răspunsuri la impuls mai

scurte pot fi evaluați folosind descompunerea în valori singulare (*singular value decomposition* - SVD) a matricii $\mathbf{H}(n)$, de dimensiune $L_1 \times L_2$, obținută prin rearanjarea elementelor din $\mathbf{h}(n)$. Cu toate acestea, în termeni de spațiu al parametrilor, nu există niciun motiv să se folosească (2) în loc de (1), deoarece primul implică $L_2(L_2 + L_1) = L_2^2 + L$ coeficienți, care depășesc lungimea L asociată cu (1). Totuși, pentru sistemele de rang mic, adică atunci când rangul matricii $\mathbf{H}(n)$ este mai mic decât L_2 , există un câștig potențial. Vom nota cu R rangul lui $\mathbf{H}(n)$, cu $R < L_2$. În acest caz, folosind descompunerea NKP [11], (2) devine

$$\mathbf{h}(n) = \sum_{r=1}^R \mathbf{h}_{2,r}(n) \otimes \mathbf{h}_{1,r}(n), \quad (3)$$

unde $\mathbf{h}_{2,r}(n)$ și $\mathbf{h}_{1,r}(n)$, cu $r = 1, 2, \dots, R$, sunt răspunsuri la impuls de lungimi L_2 și respectiv L_1 . Astfel, spațiul parametrilor constă în $R(L_2 + L_1)$ coeficienți, un număr care poate fi semnificativ mai mic decât L . De fapt, în multe scenarii practice (de exemplu, compensarea ecoului, reducerea zgomotului etc.), configurația este $R \ll L_2$ [11], [12], ceea ce conduce la avantaje importante.

Aparent, creșterea ordinului de descompunere ar aduce un beneficiu în termeni de reducere a spațiului parametrilor. În acest caz, lungimea L ar fi factorizată în mai mult de doi termeni, în timp ce mai mult de două răspunsuri la impuls (chiar mai scurte) ar fi combinate prin produsul Kronecker în descompunerea lui $\mathbf{h}(n)$, similar cu (2). Cu toate acestea, numărul de termeni din sumă nu ar mai fi legat de un rang matricial, ci de un rang tensorial, care reprezintă o problemă mai sensibilă. De exemplu, în timp ce rangul unei matrici este limitat de numărul coloanelor sale, rangul unui tensor ar putea depăși cea mai mare dimensiune a descompunerii sale.

Pentru a evita astfel de neajunsuri, soluția bazată pe TOT din [13] începe cu factorizarea $L = L_1 L_2$, dar ținând cont de faptul că $L_1 \gg L_2$, descompunând totodată $L_1 = L_{11} L_{12}$, cu $L_{11} \geq L_{12}$. Apoi, urmând același raționament privind rangul scăzut, descompunerea răspunsului la impuls conduce la [13]

$$\mathbf{h}(n) = \sum_{l=1}^{L_2} \sum_{p=1}^P \mathbf{h}_{2,l}(n) \otimes \mathbf{h}_{12,lp}(n) \otimes \mathbf{h}_{11,lp}(n), \quad (4)$$

cu $P < L_{12}$, iar răspunsurile la impuls mai scurte $\mathbf{h}_{2,l}(n)$, $\mathbf{h}_{12,lp}(n)$ și $\mathbf{h}_{11,lp}(n)$ au lungimile L_2 , L_{12} și respectiv L_{11} . Prin urmare, conform (4), spațiul parametrilor devine $L_2^2 + PL_{12}L_2 + PL_{11}L_2$, care poate fi mult mai mic decât L , cu presupunerile comune că $L_2 \ll L_{11}L_{12}$ și $P \ll L_{12}$ [13], [14].

Pentru a facilita dezvoltările viitoare, coeficienții răspunsurilor la impuls componente din (4) sunt grupați în următoarele structuri de date, care sunt specifice descompunerii TOT:

$$\underline{\mathbf{h}}_2(n) = \begin{bmatrix} \mathbf{h}_{2,1}(n) \\ \mathbf{h}_{2,2}(n) \\ \vdots \\ \mathbf{h}_{2,L_2}(n) \end{bmatrix}, \quad (5)$$

$$\bar{\mathbf{h}}_{12}(n) = \begin{bmatrix} \bar{\mathbf{h}}_{12,1}(n) \\ \bar{\mathbf{h}}_{12,2}(n) \\ \vdots \\ \bar{\mathbf{h}}_{12,L_2}(n) \end{bmatrix}, \quad \bar{\mathbf{h}}_{12,l}(n) = \begin{bmatrix} \mathbf{h}_{12,1}(n) \\ \mathbf{h}_{12,2}(n) \\ \vdots \\ \mathbf{h}_{12,P}(n) \end{bmatrix}, \quad (6)$$

$$\bar{\mathbf{h}}_{11}(n) = \begin{bmatrix} \bar{\mathbf{h}}_{11,1}(n) \\ \bar{\mathbf{h}}_{11,2}(n) \\ \vdots \\ \bar{\mathbf{h}}_{11,L_2}(n) \end{bmatrix}, \quad \bar{\mathbf{h}}_{11,l}(n) = \begin{bmatrix} \mathbf{h}_{11,1}(n) \\ \mathbf{h}_{11,2}(n) \\ \vdots \\ \mathbf{h}_{11,P}(n) \end{bmatrix}, \quad (7)$$

cu $l = 1, 2, \dots, L_2$. Trebuie observat de asemenea că termenii care compun sumele din (4) pot fi exprimați echivalent ca

$$\mathbf{h}_{2,l}(n) \otimes \mathbf{h}_{12,lp}(n) \otimes \mathbf{h}_{11,lp}(n) = [\mathbf{I}_{L_2} \otimes \mathbf{h}_{12,lp}(n) \otimes \mathbf{h}_{11,lp}(n)] \mathbf{h}_{2,l}(n) \quad (8)$$

$$= [\mathbf{h}_{2,l}(n) \otimes \mathbf{I}_{L_{12}} \otimes \mathbf{h}_{11,lp}(n)] \mathbf{h}_{12,lp}(n) \quad (9)$$

$$= [\mathbf{h}_{2,l}(n) \otimes \mathbf{h}_{12,lp}(n) \otimes \mathbf{I}_{L_{11}}] \mathbf{h}_{11,lp}(n), \quad (10)$$

unde \mathbf{I}_{L_2} , $\mathbf{I}_{L_{12}}$ și $\mathbf{I}_{L_{11}}$ reprezintă matricile identitate de dimensiuni $L_2 \times L_2$, $L_{12} \times L_{12}$ și respectiv $L_{11} \times L_{11}$. Aceste forme echivalente sunt utile pentru a extrage în continuare componentele individuale ale răspunsului global la impuls, adică ultimii termeni care apar în partea dreaptă a (8), (9) și respectiv (10). De fapt, aceste seturi de coeficienți compun răspunsurile la impuls din (5)–(7), adică $\underline{\mathbf{h}}_2(n)$, $\underline{\mathbf{h}}_{12}(n)$ și $\underline{\mathbf{h}}_{11}(n)$, care au lungimile L_2^2 , $PL_{12}L_2$ și respectiv $PL_{11}L_2$.

În concluzie, filtrele adaptive care exploatează descompunerea bazată pe TOT sunt concepute să estimeze aceste trei răspunsuri la impuls componente, în loc să modeleze direct răspunsul global la impuls de lungime L . Mai mult, coeficienții răspunsului la impuls estimat (de lungime totală) pot fi obținuți similar cu (4), așa cum este arătat în [1], [8]. De exemplu, RLS-TOT cu regularizare variabilă (VR-RLS-TOT) propus în [8] este rezumat în Tabelul 1.

3 Rezultate experimentale

Cadrul experimental ia în considerare un scenariu de compensare a ecoului, unde scopul principal este identificarea răspunsului la impuls al unei căi de ecou $\mathbf{h}(n)$, pentru a obține o replică a ecoului, care este scăzută ulterior din semnalul microfonului [10]. În acest scop, este selectat primul răspuns la impuls din Recomandarea ITU-T G168 [17], care constă într-un vector de 64 de coeficienți completat cu zerouri până la lungimea $L = 512$. Semnalul de intrare (cunoscut și ca *far-end* - îndepărtat) este o secvență de voce feminină, folosind o rată de eșantionare de 8 kHz. Ieșirea căii de ecou este coruptă de un zgomot de fundal considerat alb și gaussian (*white Gaussian noise* - WGN), cu un raport semnal-zgomot (*signal-to-noise ratio* - SNR) de 20 dB. Diferite alte tipuri de zgomot afectează de asemenea semnalul microfonului, așa cum va fi indicat în unele dintre experimentele viitoare, pentru a evalua robustețea algoritmilor. În câteva dintre experimente se folosește un scenariu de schimbare a căii de ecou, prin schimbarea semnului coeficienților lui $\mathbf{h}(n)$ în mijlocul simulării, cu scopul de a evalua capacitatea de urmărire a algoritmilor. Măsura performanței este dezalinierea normalizată, care evaluează "diferența" dintre răspunsul real la impuls, $\mathbf{h}(n)$, și estimatul său, $\hat{\mathbf{h}}(n)$. Aceasta se calculează (în dB) ca

$$\mathcal{M}(n) = 10 \log_{10} \frac{\|\mathbf{h}(n) - \hat{\mathbf{h}}(n)\|^2}{\|\mathbf{h}(n)\|^2}. \quad (11)$$

Un nivel mai scăzut al acestei măsuri de performanță indică o acuratețe mai bună a soluției/estimării, în timp ce o curbă de dezaliniere mai abruptă este asociată cu o rată de convergență și/sau urmărire mai rapidă a algoritmului.

Pe lângă VR-RLS-TOT propus, alți algoritmi sunt implicați în comparație în câteva experimente. În mod evident, este selectat corespondentul RLS-TOT din [14], care reprezintă versiunea neregularizată a VR-RLS-TOT. În plus, este utilizat și RLS-TOT regularizat (R-RLS-TOT) din [3], care folosește o tehnică de regularizare diferită, ce se bazează doar pe SNR. Toți acești algoritmi bazați pe TOT combină cele trei filtre adaptive mai scurte (cu lungimi L_2^2 , $PL_{12}L_2$ și $PL_{11}L_2$) și folosesc factorii asociați de uitare, λ_\star , cu $\star \in \{2, 12, 11\}$. Deoarece $L = 512 = 2^9$, cea mai naturală factorizare este $L_{11} = L_{12} = 16 = 2^4$ și $L_2 = 2$. În final, algoritmul RLS cu regularizare variabilă (VR-RLS) dezvoltat recent în [6] este implicat de asemenea în comparații. Folosește o tehnică de regularizare similară cu cea a VR-RLS-TOT, dar pentru algoritmul convențional RLS, folosind un singur filtru adaptiv de lungime mare (și un singur factor de uitare, λ), cu L coeficienți. Factorii de uitare ai tuturor acestor algoritmi bazați pe RLS sunt stabiliți folosind o regulă generală comună, ținând cont de lungimea filtrului [14]. Deoarece factorul de uitare este considerat un parametru legat de memorie, valoarea sa este asociată în mod natural cu lungimea filtrului, conform relației:

$$\text{forgetting factor} = 1 - \frac{1}{K \cdot (\text{filter length})}, \quad (12)$$

cu $K > 1$, unde valoarea lui K este considerată un număr întreg. Evident, cu cât acest număr întreg este mai mare, cu atât valoarea factorului de uitare este mai apropiată de unu. De asemenea, folosirea unui filtru RLS mai lung necesită o valoare mai mare a factorului de uitare. De exemplu, în cazul algoritmului VR-RLS [6], lungimea filtrului este L , astfel încât factorul său (unic) de uitare este evaluat ca $\lambda = 1 - 1/(KL)$. Este clar că această valoare este mai mare comparativ cu factorii de uitare ai algoritmilor bazați pe TOT, adică $\lambda_\star = 1 - 1/(KL_\star)$, cu $\star \in \{2, 12, 11\}$.

Înainte de comparația cu omologii săi, în primul experiment este important să analizăm performanța algoritmului VR-RLS-TOT propus, pentru valori diferite ale parametrului de descompunere P . Valoarea lui P este limitată de L_{12} , însă în practică este necesară o valoare mult mai mică. După cum este indicat în [13] și [14], valoarea lui P este de obicei între $L_{12}/8$ și $L_{12}/4$. Prin urmare, pentru configurația actuală de descompunere, evaluăm performanța

Tabelul 1 Algoritmul VR-RLS-TOT.

Factorizare: $L = L_{11}L_{12}L_2$, cu $L_2 \ll L_{11}L_{12}$ și $P \ll L_{12}$

Parametri: $0 < \lambda_\star < 1$, $\star \in \{2, 12, 11\}$, $\varepsilon > 0$

Inițializare:

$$\mathbf{R}_{12,11}(0) = \mathbf{O}_{L_2^2}, \quad \mathbf{R}_{2,11}(0) = \mathbf{O}_{PL_{12}L_2}, \quad \mathbf{R}_{2,12}(0) = \mathbf{O}_{PL_{11}L_2}$$

$$\sigma_{v_\star}^2(0) = 0, \quad \sigma_{w_\star}^2(0) = \zeta, \quad \zeta > 0$$

pentru $l = 1, 2, \dots, L_2$, $p = 1, 2, \dots, P$

$$\hat{\mathbf{h}}_{2,l}(0) = \begin{bmatrix} \epsilon \\ \mathbf{0}_{L_2-1} \end{bmatrix}, \quad \hat{\mathbf{h}}_{12,lp}(0) = \begin{bmatrix} \epsilon \\ \mathbf{0}_{L_{12}-1} \end{bmatrix}, \quad \hat{\mathbf{h}}_{11,lp}(0) = \begin{bmatrix} \epsilon \\ \mathbf{0}_{L_{11}-1} \end{bmatrix}, \quad 0 < \epsilon \leq 1$$

$$\implies \hat{\mathbf{H}}_{12,11}(0), \quad \hat{\mathbf{H}}_{2,11}(0), \quad \hat{\mathbf{H}}_{2,12}(0)$$

Iterații:

pentru $n = 1, 2, \dots$

$$\mathbf{x}_{12,11}(n) = \hat{\mathbf{H}}_{12,11}^\top(n-1)\mathbf{x}(n)$$

$$\mathbf{x}_{2,11}(n) = \hat{\mathbf{H}}_{2,11}^\top(n-1)\mathbf{x}(n)$$

$$\mathbf{x}_{2,12}(n) = \hat{\mathbf{H}}_{2,12}^\top(n-1)\mathbf{x}(n)$$

$$\mathbf{R}_{12,11}(n) = \lambda_2 \mathbf{R}_{12,11}(n-1) + \mathbf{x}_{12,11}(n)\mathbf{x}_{12,11}^\top(n)$$

$$\mathbf{R}_{2,11}(n) = \lambda_{12} \mathbf{R}_{2,11}(n-1) + \mathbf{x}_{2,11}(n)\mathbf{x}_{2,11}^\top(n)$$

$$\mathbf{R}_{2,12}(n) = \lambda_{11} \mathbf{R}_{2,12}(n-1) + \mathbf{x}_{2,12}(n)\mathbf{x}_{2,12}^\top(n)$$

$$e_2(n) = d(n) - \mathbf{x}_{12,11}^\top(n)\hat{\mathbf{h}}_2(n-1)$$

$$e_{12}(n) = d(n) - \mathbf{x}_{2,11}^\top(n)\hat{\mathbf{h}}_{12}(n-1)$$

$$e_{11}(n) = d(n) - \mathbf{x}_{2,12}^\top(n)\hat{\mathbf{h}}_{11}(n-1)$$

$$\sigma_{v_\star}^2(n) = \lambda_\star \sigma_{v_\star}^2(n-1) + (1 - \lambda_\star)e_\star^2(n), \quad \star \in \{2, 12, 11\}$$

$$\delta_2(n) = \frac{1}{L_2^2(1 - \lambda_2)} \cdot \frac{\sigma_{v_2}^2(n)}{\varepsilon + \sigma_{w_2}^2(n-1)}$$

$$\delta_{12}(n) = \frac{1}{PL_{12}L_2(1 - \lambda_{12})} \cdot \frac{\sigma_{v_{12}}^2(n)}{\varepsilon + \sigma_{w_{12}}^2(n-1)}$$

$$\delta_{11}(n) = \frac{1}{PL_{11}L_2(1 - \lambda_{11})} \cdot \frac{\sigma_{v_{11}}^2(n)}{\varepsilon + \sigma_{w_{11}}^2(n-1)}$$

$$\hat{\mathbf{h}}_2(n) = \hat{\mathbf{h}}_2(n-1) + [\mathbf{R}_{12,11}(n) + \delta_2(n)\mathbf{I}_{L_2^2}]^{-1} \mathbf{x}_{12,11}(n)e_2(n)$$

$$\hat{\mathbf{h}}_{12}(n) = \hat{\mathbf{h}}_{12}(n-1) + [\mathbf{R}_{2,11}(n) + \delta_{12}(n)\mathbf{I}_{PL_{12}L_2}]^{-1} \mathbf{x}_{2,11}(n)e_{12}(n)$$

$$\hat{\mathbf{h}}_{11}(n) = \hat{\mathbf{h}}_{11}(n-1) + [\mathbf{R}_{2,12}(n) + \delta_{11}(n)\mathbf{I}_{PL_{11}L_2}]^{-1} \mathbf{x}_{2,12}(n)e_{11}(n)$$

$$\implies \hat{\mathbf{h}}_{2,l}(n), \quad \hat{\mathbf{h}}_{12,lp}(n), \quad \hat{\mathbf{h}}_{11,lp}(n), \quad l = 1, 2, \dots, L_2, \quad p = 1, 2, \dots, P$$

$$\implies \hat{\mathbf{H}}_{12,11}(n), \quad \hat{\mathbf{H}}_{2,11}(n), \quad \hat{\mathbf{H}}_{2,12}(n)$$

$$\sigma_{w_2}^2(n) = \lambda_2 \sigma_{w_2}^2(n-1) + (1 - \lambda_2) \frac{\|\hat{\mathbf{h}}_2(n) - \hat{\mathbf{h}}_2(n-1)\|^2}{L_2^2}$$

$$\sigma_{w_{12}}^2(n) = \lambda_{12} \sigma_{w_{12}}^2(n-1) + (1 - \lambda_{12}) \frac{\|\hat{\mathbf{h}}_{12}(n) - \hat{\mathbf{h}}_{12}(n-1)\|^2}{PL_{12}L_2}$$

$$\sigma_{w_{11}}^2(n) = \lambda_{11} \sigma_{w_{11}}^2(n-1) + (1 - \lambda_{11}) \frac{\|\hat{\mathbf{h}}_{11}(n) - \hat{\mathbf{h}}_{11}(n-1)\|^2}{PL_{11}L_2}$$

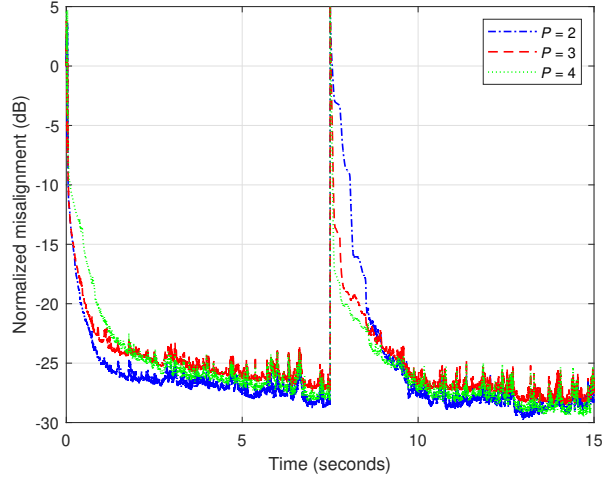


Figura 1: Dezalinierea normalizată a algoritmului VR-RLS-TOT folosind diferite valori ale parametrului de descompunere P . Factorii de uitare λ_* , cu $\star \in \{2, 12, 11\}$, sunt setați conform (12), folosind $K = 30$. Semnalul de intrare este o secvență de vorbire, iar calea de ecou se schimbă după 7,5 secunde.

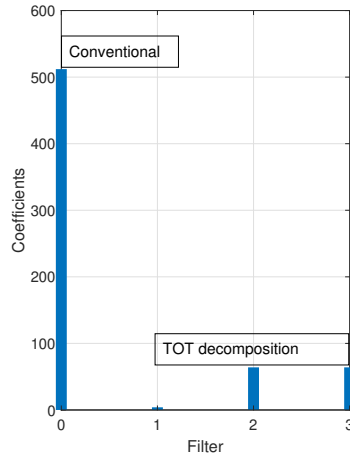


Figura 2: Spațiul parametrilor (adică numărul de coeficienți) ai soluției convenționale ($L = 512$) și ai descompunerii bazate pe TOT cu $P = 2$, $L_{11} = L_{12} = 16$ și $L_2 = 2$ (adică $L_2^2 = 4$, $PL_{12}L_2 = PL_{11}L_2 = 64$).

algoritmului VR-RLS-TOT folosind valorile $P = 2, 3$, și 4 , setând $K = 30$ pentru calcularea celor trei factori de uitare (adică λ_* , cu $\star \in \{2, 12, 11\}$) conform (12). Rezultatele sunt prezentate în Fig. 1, unde un scenariu de schimbare a căii de ecou este simulat la 7,5 secunde. În ceea ce privește acuratețea (adică nivelul de dezaliniere), putem observa o performanță similară pentru toate valorile lui P . O reacție de urmărire doar puțin mai bună se obține atunci când valoarea lui P crește. Totuși, acest lucru crește și complexitatea computațională. Prin urmare, folosirea valorii $P = 2$ reprezintă un compromis adecvat în ceea ce privește performanța și complexitatea. Această valoare va fi folosită suplimentar pentru toți algoritmi bazati pe TOT analizați în această secțiune.

Avantajul descompunerii bazate pe TOT se bazează pe utilizarea unui spațiu de parametri mult mai mic (adică un număr mult mai mic de coeficienți), comparativ cu abordarea convențională, care implică un singur filtru de lungime L . Acest lucru este susținut în Fig. 2, unde filtrul indicat cu "0" pe axa x reprezintă soluția clasică folosind $L = 512$ coeficienți. Aici, filtrele indicate cu "1", "2" și "3" sunt asociate descompunerii bazate pe TOT, cu lungimile $L_2^2 = 4$ și $PL_{12}L_2 = PL_{11}L_2 = 64$, care corespund configurației actuale de descompunere folosind $P = 2$. Este clar că, atunci când se folosește descompunerea bazată pe TOT, există o reducere importantă a spațiului parametrilor, ceea ce duce

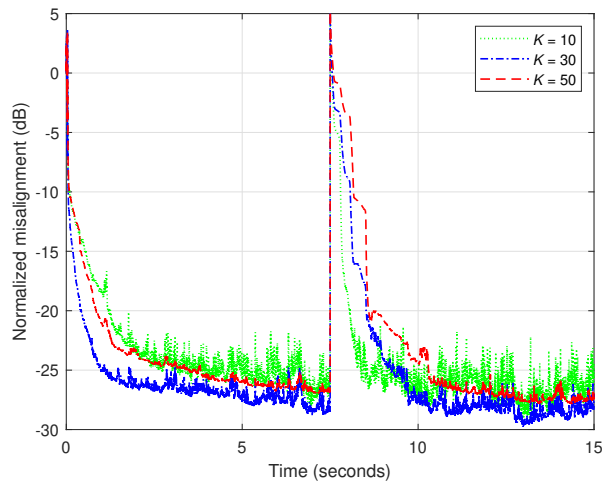


Figura 3: Dezalinierea normalizată a algoritmului VR-RLS-TOT folosind valori diferite ale lui K pentru evaluarea factorilor de uitare λ_* , cu $\star \in \{2, 12, 11\}$, conform (12). Parametrul de descompunere este setat la $P = 2$. Semnalul de intrare este o secvență de vorbire, iar calea de ecou se schimbă după 7,5 secunde.

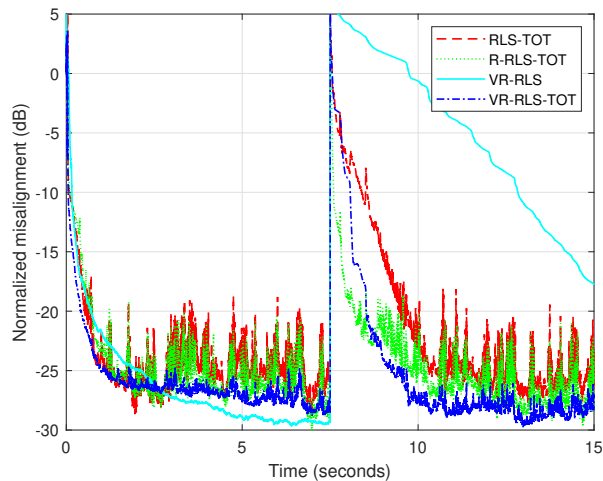


Figura 4: Dezalinierea normalizată a RLS-TOT [14], R-RLS-TOT [3], VR-RLS [6] și VR-RLS-TOT propus. Toți algoritmi folosesc $K = 30$ pentru evaluarea factorilor de uitare corespunzători, conform (12), în timp ce algoritmi bazați pe TOT folosesc parametrul de descompunere $P = 2$. Semnalul de intrare este o secvență de vorbire, iar calea de ecou se schimbă după 7,5 secunde.

la o complexitate aritmetică mai scăzută și la o performanță îmbunătățită comparativ cu abordarea convențională.

Factorul de uitare este un parametru important care controlează performanța oricărui filtru de tip RLS ponderat exponențial. O valoare mai mare a factorului de uitare duce la o acuratețe bună a soluției (adică o dezaliniere scăzută), dar cu o urmărire lentă. Pe de altă parte, reducerea valorii acestui parametru afectează acuratețea, îmbunătățind în același timp capacitatea de urmărire. VR-RLS-TOT propus folosește trei factori de uitare, fiecare asociat cu unul dintre filtrele componente. În Fig. 3, valorile lor sunt stabilite folosind valori diferite ale numărului întreg K . Așa cum era de așteptat, o valoare mai mică a lui K îmbunătățește urmărirea, dar crește și dezalinierea. Pentru un compromis mai bun în ceea ce privește aceste criterii de performanță, valoarea "mijlocie" $K = 30$ reprezintă o alegere favorabilă, care va fi folosită în următoarele experimente pentru toți algoritmi bazați pe RLS.

În continuare, algoritmul VR-RLS-TOT propus este comparat cu omologii săi bazați pe TOT din [14] și [3], adică

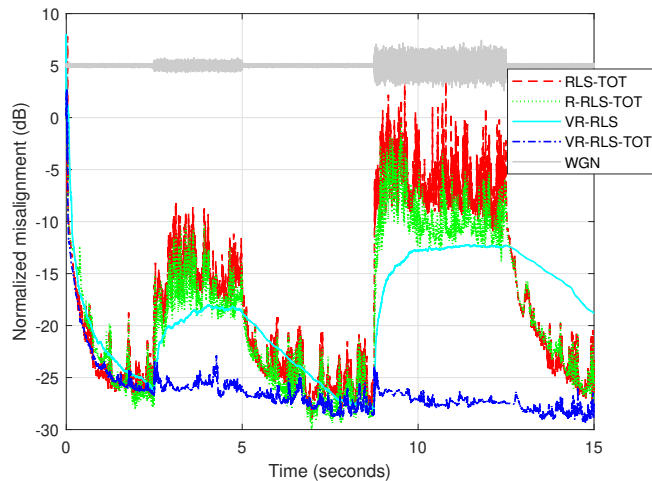


Figura 5: Dezalinierea normalizată a RLS-TOT [14], R-RLS-TOT [3], VR-RLS [6], și VR-RLS-TOT propus. Toți algoritmi folosesc $K = 30$ pentru evaluarea factorilor de uitare corespunzători, conform (12), în timp ce algoritmi bazați pe TOT folosesc parametrul de descompunere $P = 2$. Semnalul de intrare este o secvență de vorbire și două explozii de WGN sunt simulate în capătul *near-end* (apropiat), cu $\text{SNR} = 10$ dB and 0 dB.

algoritmi RLS-TOT și R-RLS-TOT. De asemenea, algoritmul VR-RLS din [6] este inclus în comparații. Factorii de uitare pentru toți algoritmi sunt setați folosind (12), cu $K = 30$. Parametrul de descompunere pentru algoritmi bazați pe TOT este $P = 2$, așa cum a fost menționat anterior. Rezultatele sunt prezentate în Fig. 4. În primul rând, se poate observa că VR-RLS-TOT propus este mai robust în ceea ce privește acuratețea, atingând un nivel de dezaliniere ușor mai scăzut comparativ cu ceilalți doi algoritmi bazați pe TOT, care sunt mai afectați de natura nestaționară a semnalului de intrare (adică secvența de vorbire). De asemenea, algoritmul VR-RLS-TOT are o urmărire bună atunci când calea de ecou se schimbă (după 7,5 secunde). În acest caz, putem observa o reacție semnificativ mai lentă a algoritmului VR-RLS, care funcționează cu un singur filtru adaptiv (de lungime mare), cu $L = 512$ coeficienți. Pe de altă parte, algoritmi bazați pe TOT combină estimările furnizate de trei filtre adaptive (mult) mai scurte, cu 4, 64 și respectiv 64 de coeficienți (așa cum este indicat în Fig. 2), ceea ce conduce la un comportament de urmărire îmbunătățit.

În continuare, caracteristicile de robustețe ale algoritmilor sunt testate în condiții de zgomot, când semnalul microfonului din *near-end* captează și alte semnale diferite, pe lângă semnalul de ecou și zgomotul de fundal. În primul rând, în Fig. 5 se iau în considerare două explozii de zgomot gaussian alb (WGN), cu $\text{SNR} = 10$ dB și 0 dB. Se poate observa că VR-RLS-TOT propus este foarte robust la astfel de perturbații, în timp ce ceilalți algoritmi sunt afectați semnificativ. R-RLS-TOT [3] are performanțe ușor mai bune comparativ cu RLS-TOT [14], în timp ce algoritmul VR-RLS [6] are o anumită latență legată de reacții/tranziții, din cauza filtrului de lungime mare care trebuie actualizat.

Un alt scenariu ia în considerare tipuri mai dificile de zgomot, comparativ cu exploziile WGN anterioare (care sunt secvențe staționare). În Fig. 6, sursa de perturbație (și natura sa) se schimbă, folosind două explozii de zgomot de autostradă și respectiv zgomot de motor, a doua fiind mai puternică. Toți algoritmi sunt mai afectați în aceste condiții, comparativ cu scenariul anterior. Cu toate acestea, VR-RLS-TOT propus rămâne excelent în ceea ce privește robustețea, fiind mai puțin afectat de prima explozie de zgomot de pe autostradă și recuperându-se destul de bine în cazul celei de-a doua explozii de zgomot de motor.

În final, se ia în considerare cel mai dificil scenariu de compensare a ecoului, care este așa-numitul caz de dublă vorbire, când ambii vorbitori (la capătul *far-end* și la capătul *near-end*) vorbesc în același timp [10]. În acest scenariu, vorbirea la capătul apropiat acționează ca un nivel mare de perturbare nestaționară pentru filtrul adaptiv. În Fig. 7 apar două perioade de dublă vorbire, cu intensități diferite, a doua fiind mai semnificativă. După cum putem observa, atât comportamentul RLS-TOT [14], cât și cel al R-RLS-TOT [3] sunt serios influențate în timpul vorbirii duble. Algoritmul VR-RLS [6] are o anumită robustețe în acest scenariu, fiind totuși influențat de latența inerentă cauzată de operarea cu un filtru de lungime mare ($L = 512$). Pe de altă parte, algoritmul propus VR-RLS-TOT își revine

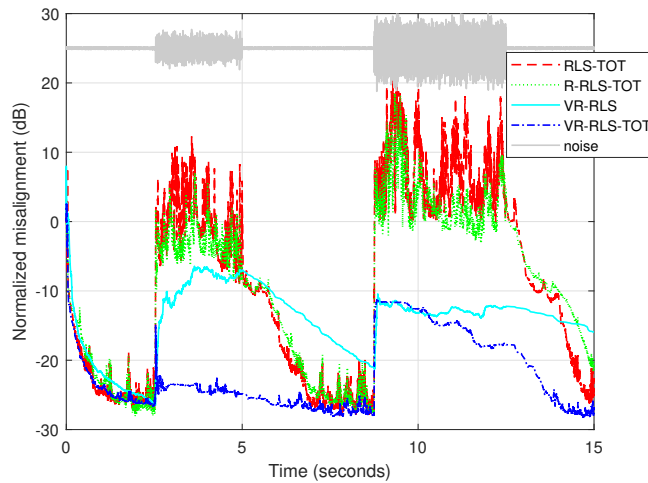


Figura 6: Dezalinierea normalizată a RLS-TOT [14], R-RLS-TOT [3], VR-RLS [6], și VR-RLS-TOT propus. Toți algoritmi folosesc $K = 30$ pentru evaluarea factorilor de uitare corespunzători, conform (12), în timp ce algoritmi bazați pe TOT folosesc parametrul de descompunere $P = 2$. Semnalul de intrare este o secvență de vorbire și două explozii de zgomote diferite sunt simulate în capătul *near-end*, adică un zgomot de autostradă și respectiv un zgomot de motor.

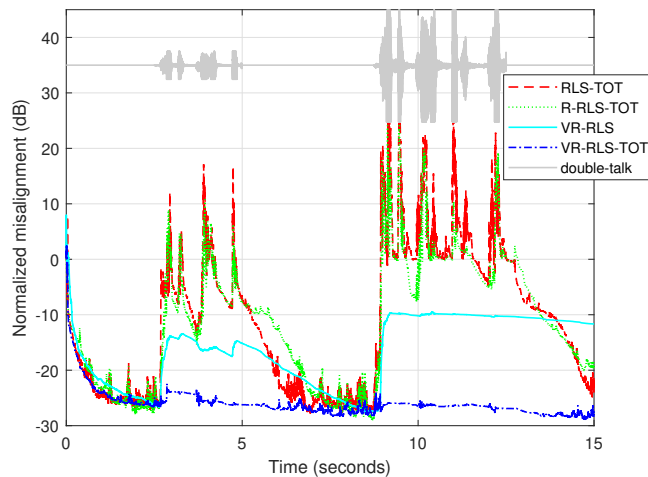


Figura 7: Dezalinierea normalizată a RLS-TOT [14], R-RLS-TOT [3], VR-RLS [6], și VR-RLS-TOT propus. Toți algoritmi folosesc $K = 30$ pentru evaluarea factorilor de uitare corespunzători, conform (12), în timp ce algoritmi bazați pe TOT folosesc parametrul de descompunere $P = 2$. Semnalul de intrare este o secvență de vorbire și un scenariu de dublă vorbire este simulat folosind două perioade de vorbire în capătul *near-end*, cu intensități diferite.

rapid la începutul primei perioade de dublă vorbire, fiind foarte robust în a doua perioadă de dublă vorbire (mai puternică). În concluzie, se obține un compromis mult mai bun în ceea ce privește principalele criterii de performanță, comparativ cu ceilalți algoritmi.

References

- [1] L. M. Dogariu, C. Paleologu, J. Benesty, and F. Albu, “Kalman filter using a third-order tensorial decomposition of the impulse response,” *Applied Sciences*, vol. 14, id. 4507, 17 pages, May 2024. [ISI-WOS Q1 la data publicării]
- [2] C. L. Stanciu, C. Anghel, I. D. Fîciu, C. Elisei-Iliescu, R. M. Udrea, and L. Stanciu, “On the regularization of recursive least-squares adaptive algorithms using line search methods,” *Electronics*, vol. 13, id. 1479, 21 pages, Apr. 2024. [ISI-WOS Q2 la data publicării]
- [3] R. Otopoleanu, L. M. Dogariu, C. L. Stanciu, C. Paleologu, J. Benesty, and S. Ciochină, “A data-reuse regularized recursive least-squares adaptive filtering algorithm,” in *Proc. IEEE International Symposium on Electronics and Telecommunications (ISETC)*, Nov. 2024, 4 pages.
- [4] C. L. Stanciu, C. Anghel, L. M. Dogariu, I. D. Fîciu, and L. Stanciu, “Regularized affine projection algorithm for stereophonic acoustic echo cancellation,” in *Proc. IEEE Conference Advanced Topics on Measurement and Simulation (ATOMS)*, Aug. 2024, pp. 323–326.
- [5] L.-M. Dogariu, R.-L. Costea, C. Paleologu, and J. Benesty, “A fourth-order tensorial Wiener filter using the conjugate gradient method,” *Symmetry*, vol. 16, id. 1433, 18 pages, Dec. 2024. [ISI-WOS Q2 la data publicării]
- [6] R. A. Otopoleanu, J. Benesty, C. Paleologu, C. L. Stanciu, L. M. Dogariu, and S. Ciochină, “A practical regularized recursive least-squares algorithm for robust system identification,” in *Proc. European Signal Processing Conference (EUSIPCO)*, Sept. 2025, pp. 1417-1421.
- [7] R. A. Otopoleanu, C. L. Stanciu, C. Paleologu, J. Benesty, L. M. Dogariu, and C. Anghel, “Robust recursive least-squares algorithm with the conjugate gradient and the data reuse approach,” in *Proc. IEEE International Symposium for Design and Technology in Electronic Packaging (SIITME)*, Oct. 2025, 4 pages.
- [8] R. A. Otopoleanu, C. Paleologu, J. Benesty, C. L. Stanciu, L.-M. Dogariu, and R.-L. Costea, “Robust regularized recursive least-squares algorithm based on third-order tensor decomposition,” *Algorithms*, 23 pages, accepted Dec. 2025. [ISI-WOS Q2 la data acceptării]
- [9] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Fifth Edition, Cham, Switzerland: Springer Nature, 2020.
- [10] E. Hänsler and G. Schmidt, *Acoustic Echo and Noise Control—A Practical Approach*. Hoboken, NJ, USA: Wiley, 2004.
- [11] C. Paleologu, J. Benesty, and S. Ciochină, “Linear system identification based on a Kronecker product decomposition,” *IEEE/ACM Trans. Audio, Speech, Language Processing*, vol. 26, pp. 1793–1808, Oct. 2018.
- [12] 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 Processing*, vol. 27, pp. 903–918, May 2019.
- [13] J. Benesty, C. Paleologu, and S. Ciochină, “Linear system identification based on a third-order tensor decomposition,” *IEEE Signal Processing Lett.*, vol. 30, pp. 503–507, May 2023.
- [14] C. Paleologu, J. Benesty, C. L. Stanciu, J. R. Jensen, M. G. Christensen, and S. Ciochină, “Recursive least-squares algorithm based on a third-order tensor decomposition for low-rank system identification,” *Signal Processing*, vol. 213, id. 109216, Dec. 2023.
- [15] R. Otopoleanu, C. Elisei-Iliescu, C. Paleologu, J. Benesty, and S. Ciochină, “Regularized RLS algorithm based on third-order tensor decomposition,” in *Proc. IEEE Conf. Advanced Topics Measurement Simulation (ATOMS)*, 2024, pp. 307–310.
- [16] C. F. Van Loan, “The ubiquitous Kronecker product,” *J. Computational Applied Mathematics*, vol. 123, pp. 85–100, Nov. 2000.
- [17] *Digital Network Echo Cancellers*, ITU-T Recommendations G.168, 2015. [<https://www.itu.int/rec/T-REC-G.168> (accessed on Nov 14th, 2025)]

Membrii echipei proiectului,

Laura-Maria Dogariu

Ionuț-Dorinel Fîciu

