An Efficient Variable Step-Size Proportionate Affine Projection Algorithm

Constantin Paleologu[1], *Jacob Benesty*[2], *Felix Albu*[1], *and Silviu Ciochină*[1]

[1] University Politehnica of Bucharest, Romania
{pale, silviu, felix}@comm.pub.ro
[2] INRS-EMT, Universite du Quebec, Montreal, Canada
benesty@emt.inrs.ca

Outline

- Introduction
- Proportionate-type APAs (PAPAs)
- Variable Step-Size Memory PAPA (VSS-MPAPA)
- Simulation Results

Introduction

- **echo cancellation**

→ an *adaptive filter* identifies the echo path

➔ *system identification* problem

- **specific problems**

→ the echo path can be extremely long

→ the double-talk situation

- **requirements**

→ fast convergence ➔ the *affine projection algorithm* (APA)

→ robustness to double-talk ➔ *variable step-size* (VSS) versions

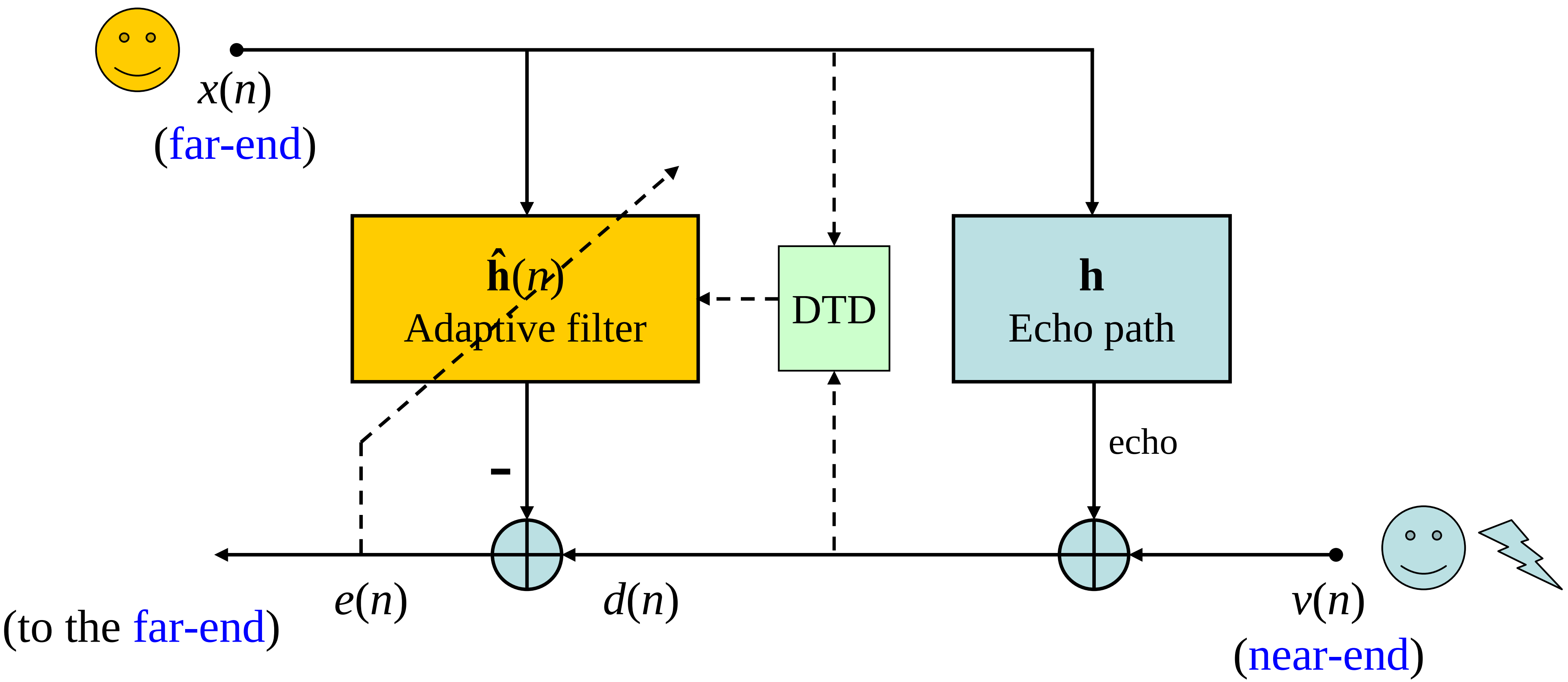
- **hints**

→ the echo paths are *sparse* in nature ➔ *proportionate* APAs

(a small percentage of the impulse response components have a significant magnitude while the rest are zero or small)

(adjust the adaptation step-size in proportion to the magnitude of the estimated filter coefficient)

Echo Cancellation



- **Goals**

- to ***recover*** the near-end signal from the error of the adaptive filter

➔ to ***cancel*** the echo

➔ to ***identify*** the echo path with the adaptive filter.

Proportionate-type APAs (PAPAs)

- **APA**

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{X}(n)\right]^{-1}\mathbf{e}(n)$$

step-size parameter

where $\mathbf{d}(n) = [d(n), d(n-1), \ldots, d(n-p+1)]^T$

projection order

$\mathbf{X}(n) = [\mathbf{x}(n), \mathbf{x}(n-1), \ldots, \mathbf{x}(n-p+1)]$

$\mathbf{x}(n-l) = [x(n-l), x(n-l-1), \ldots, x(n-l-L+1)]^T$

$l = 0, 1, \ldots, p-1$

adaptive filter length

- **PAPA**

$\mathbf{P}(n)$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu\mathbf{G}(n-1)\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n)\right]^{-1}\mathbf{e}(n)$$

$$\mathbf{G}(n-1) = \mathrm{diag}\left[g_0(n-1) \quad g_1(n-1) \quad \ldots \quad g_{L-1}(n-1)\right]$$

Proportionate-type APAs (PAPAs)

- **Classical PAPA**

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu \mathbf{P}(n)\left[\mathbf{X}^T(n)\mathbf{P}(n)\right]^{-1}\mathbf{e}(n)$$

Hadamard product

$$\mathbf{P}(n) = \left[\mathbf{g}(n-1)\odot\mathbf{x}(n) \quad \ldots \quad \mathbf{g}(n-1)\odot\mathbf{x}(n-p+1)\right]$$

$$\mathbf{g}(n-1) = \left[g_0(n-1) \quad g_1(n-1) \quad \ldots \quad g_{L-1}(n-1)\right]^T$$

- **“Memory” PAPA (MPAPA)** $\mathbf{M}(n)$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu \mathbf{P}'(n)\left[\mathbf{X}^T(n)\mathbf{P}'(n)\right]^{-1}\mathbf{e}(n)$$

$$\mathbf{P}'(n) = \left[\mathbf{g}(n-1)\odot\mathbf{x}(n) \quad \ldots \quad \mathbf{g}(n-p)\odot\mathbf{x}(n-p+1)\right]$$

$$\mathbf{P}'(n) = \left[\mathbf{g}(n-1)\odot\mathbf{x}(n) \quad \mathbf{P}'_{:,1:p-1}(n-1)\right]$$

$$\begin{bmatrix} \bullet & \bullet \\ \bullet & \mathbf{M}_{1:p-1,1:p-1}(n-1) \end{bmatrix}$$

Proportionate-type APAs (PAPAs)

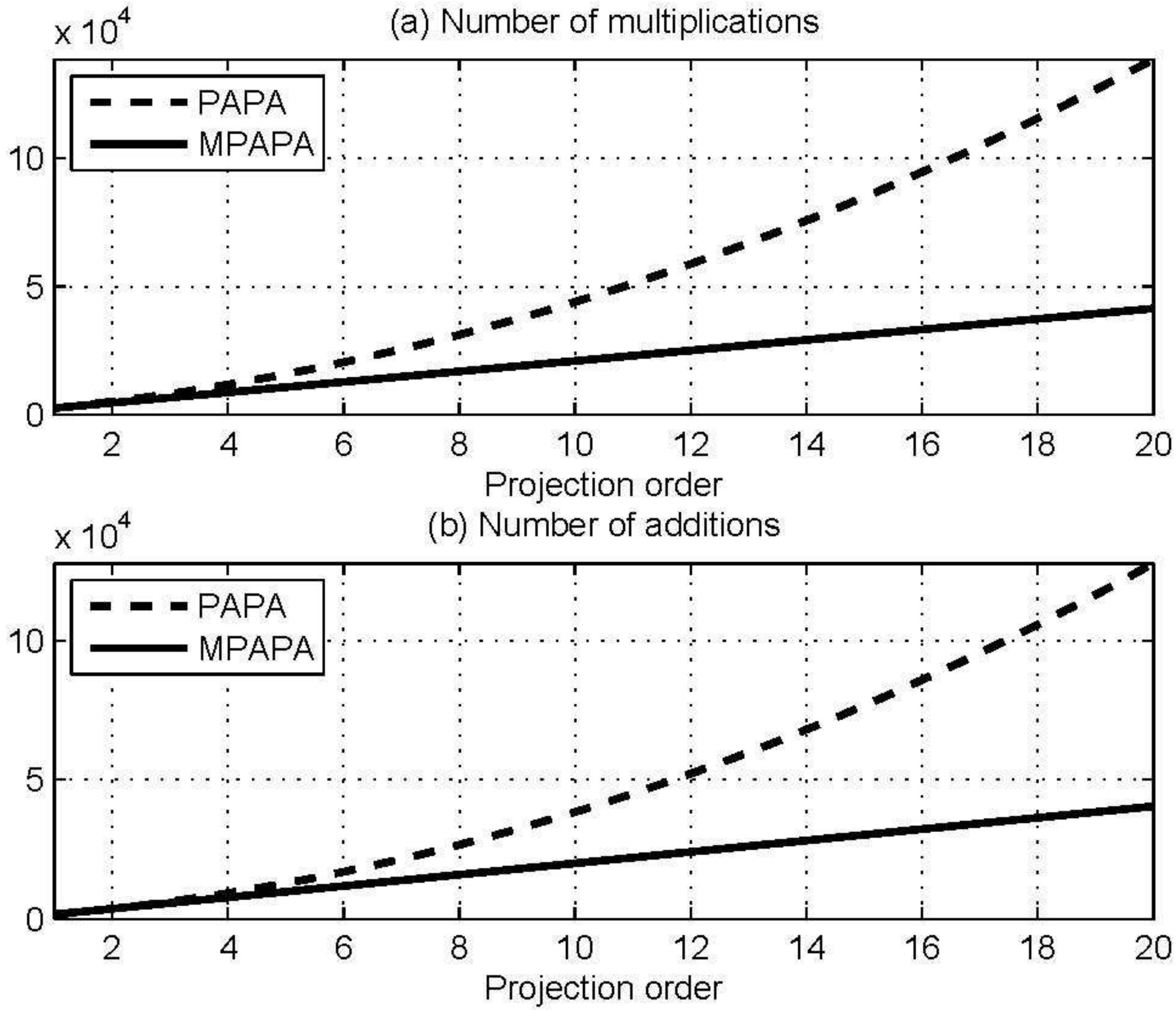


Fig. 1. (a) Number of multiplications of the classical PAPA and MPAPA, as a function of the projection order. (b) Number of additions of the classical PAPA and MPAPA, as a function of the projection order. The length of the adaptive filter is $L = 512$.

Variable Step-Size MPAPA

- **MPAPA update**

$$\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\mu\mathbf{P}'(n)\left[\mathbf{X}^T(n)\mathbf{P}'(n)\right]^{-1}\mathbf{e}(n)$$

- **step-size parameter**

→ *large* values ➔ fast convergence rate and tracking

→ *small* values ➔ low misadjustment and double-talk robustness

conflicting requirements ➔ Variable Step-Size MPAPA (VSS-MPAPA)

$$\boxed{\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\mathbf{P}'(n)\left[\mathbf{X}^T(n)\mathbf{P}'(n)\right]^{-1}\boldsymbol{\mu}(n)\mathbf{e}(n)}$$

$$\boldsymbol{\mu}(n)=\mathrm{diag}\left\{\mu_0(n),\mu_1(n),\ldots,\mu_{p-1}(n)\right\}$$

$$\mu_0(n)=\mu_1(n)=\ldots=\mu_{p-1}(n)=\mu$$ ➔ Fixed step-size MPAPA

Variable Step-Size MPAPA

$$\begin{cases} \boldsymbol{\varepsilon}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n) & \rightarrow \textit{a posteriori} \text{ error vector} \\ \mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1) & \rightarrow \textit{a priori} \text{ error vector} \end{cases}$$

$$\boldsymbol{\varepsilon}(n) = \left[\mathbf{I}_p - \boldsymbol{\mu}(n)\right]\mathbf{e}(n)$$

$$\cancel{\boldsymbol{\varepsilon}(n) = \mathbf{0}_{p\times 1}} \quad \Rightarrow \quad \boldsymbol{\mu}(n) = \mathbf{I}_p \quad \Rightarrow \text{ MPAPA with } \mu = 1$$

$$\boxed{\boldsymbol{\varepsilon}(n) = \mathbf{v}(n)}$$

where $\mathbf{v}(n) = [v(n), v(n-1), \ldots, v(n-p+1)]^T$

$$\varepsilon_{l+1}(n) = \left[1 - \mu_l(n)\right]e_{l+1}(n) = v(n-l) \qquad l = 0,1,\ldots,p-1$$

$$E\left\{\varepsilon_{l+1}^2(n)\right\} = E\left\{v^2(n-l)\right\} \quad \Rightarrow \quad \mu_l(n) = 1 - \sqrt{\frac{E\left\{v^2(n-l)\right\}}{E\left\{e_{l+1}^2(n)\right\}}}$$

Variable Step-Size MPAPA

$$\boxed{\mu_l(n) = 1 - \frac{\hat{\sigma}_v(n-l)}{\hat{\sigma}_{e_{l+1}}(n)}}$$

?

√ $\hat{\sigma}^2_{e_{l+1}}(n) = \lambda \hat{\sigma}^2_{e_{l+1}}(n-1) + (1-\lambda) e^2_{l+1}(n)$

$\lambda = 1 - 1/(KL)$, with $K > 1$

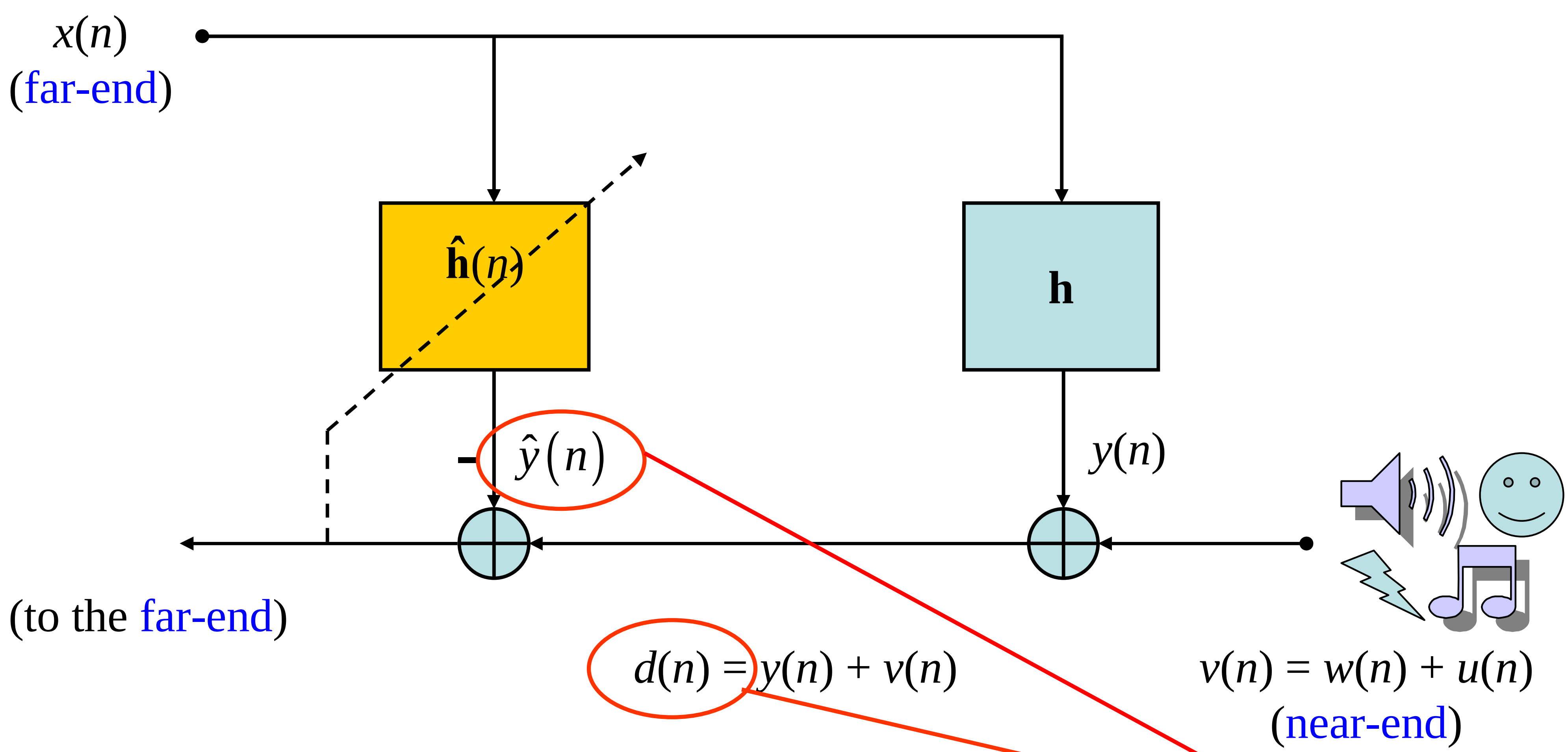
! near-end signal = background noise + near-end speech

$v(n) = w(n) + u(n)$ (*double-talk scenario*)

$$\hat{\sigma}^2_v(n) = \hat{\sigma}^2_w(n) + \hat{\sigma}^2_u(n)$$

near-end speech power estimate

Problem: non-stationary character of the speech signal



available !

$$\begin{cases} E\{d^2(n)\} = E\{y^2(n)\} + E\{v^2(n)\} \\ E\{y^2(n)\} \cong E\{\hat{y}^2(n)\} \end{cases}$$

→ assuming that the adaptive filter has converged to a certain degree

$$E\{v^2(n)\} \cong E\{d^2(n)\} - E\{\hat{y}^2(n)\} \quad \Rightarrow \quad \boxed{\hat{\sigma}_v^2(n) \cong \hat{\sigma}_d^2(n) - \hat{\sigma}_{\hat{y}}^2(n)}$$

Variable Step-Size MPAPA

$$\boxed{\mu_l(n) = 1 - \frac{\hat{\sigma}_v(n-l)}{\hat{\sigma}_{e_{l+1}}(n)}}$$

?

√ $\hat{\sigma}^2_{e_{l+1}}(n) = \lambda \hat{\sigma}^2_{e_{l+1}}(n-1) + (1-\lambda) e^2_{l+1}(n)$

$\lambda = 1 - 1/(KL)$, with $K > 1$

! near-end signal = background noise + near-end speech

$v(n) = w(n) + u(n)$ (*double-talk scenario*)

$$\hat{\sigma}^2_v(n) = \hat{\sigma}^2_w(n) + \hat{\sigma}^2_u(n)$$

near-end speech power estimate

Problem: non-stationary character of the speech signal

$$\boxed{\mu_l(n) = 1 - \frac{\sqrt{\hat{\sigma}^2_d(n-l) - \hat{\sigma}^2_{\hat{y}}(n-l)}}{\hat{\sigma}_{e_{l+1}}(n)}}$$

Simulation Results

- **conditions**
  - → network echo cancellation, L = 512
  - → $x(n)$ = white Gaussian noise or a speech sequence
  - → $w(n)$ = independent white Gaussian noise signal (SNR = 30dB)
  - → normalized misalignment (dB) = $20\log_{10}(\|\mathbf{h}-\hat{\mathbf{h}}(n)\|/\|\mathbf{h}\|)$
  - → echo-return loss enhancement (ERLE)
- **algorithms for comparisons** (p = 4)
  - → memory improved PAPA (MIPAPA)

$$g_k(n-1)=\frac{1-\alpha}{2L}+(1+\alpha)\frac{\left|\hat{h}_k(n-1)\right|}{2\sum_{i=0}^{L-1}\left|\hat{h}_i(n-1)\right|+\varepsilon} \qquad k=0,1,\ldots,L-1$$

$$-1\le\alpha<1$$

  - → VSS-APA
  - → VSS-MIPAPA

- single-talk scenario

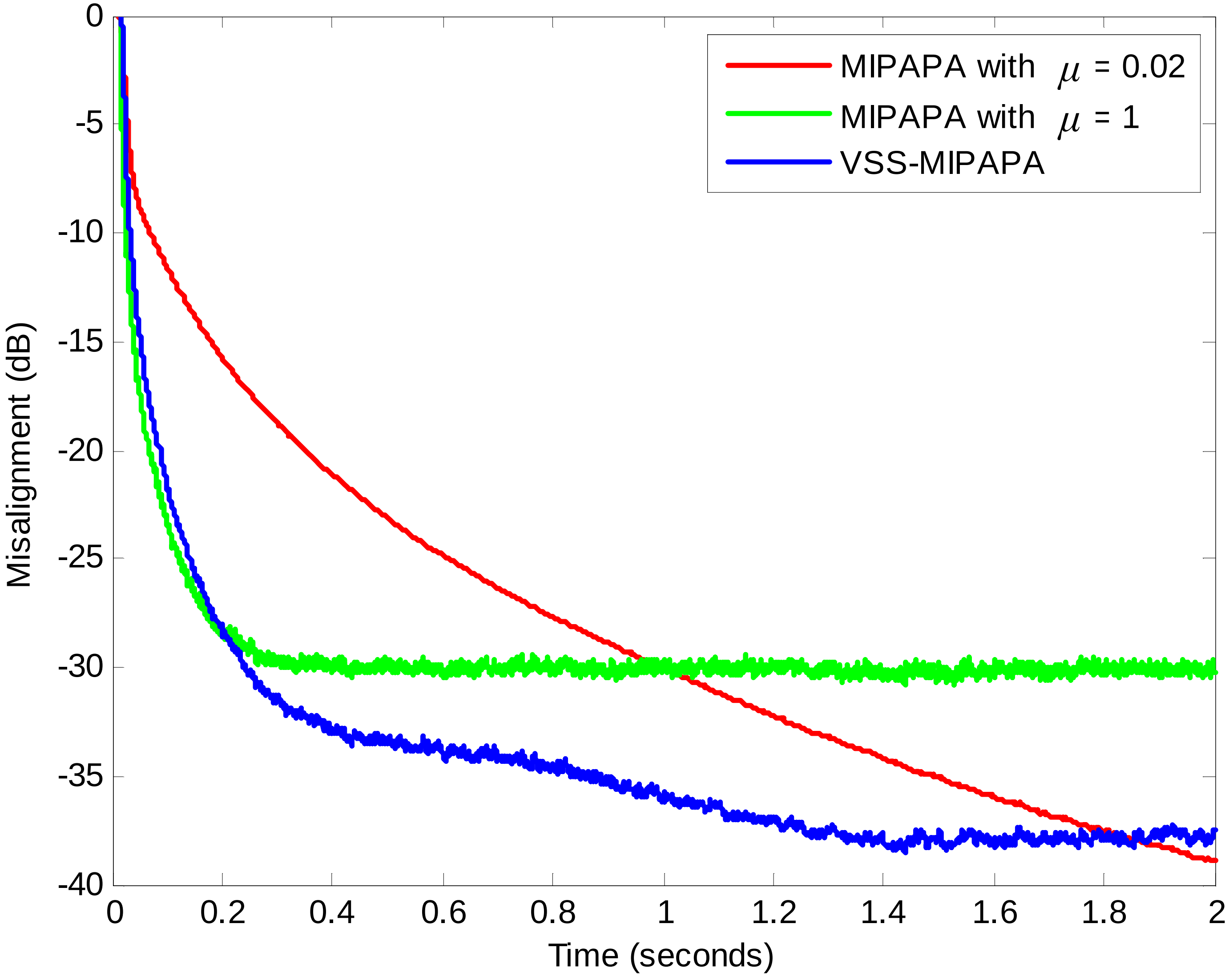


Fig. 2. Misalignment of the MIPAPA (μ = 1 and μ = 0.02) and VSS-MIPAPA; P = 4 and white Gaussian input signal.

- single-talk scenario + echo path changes

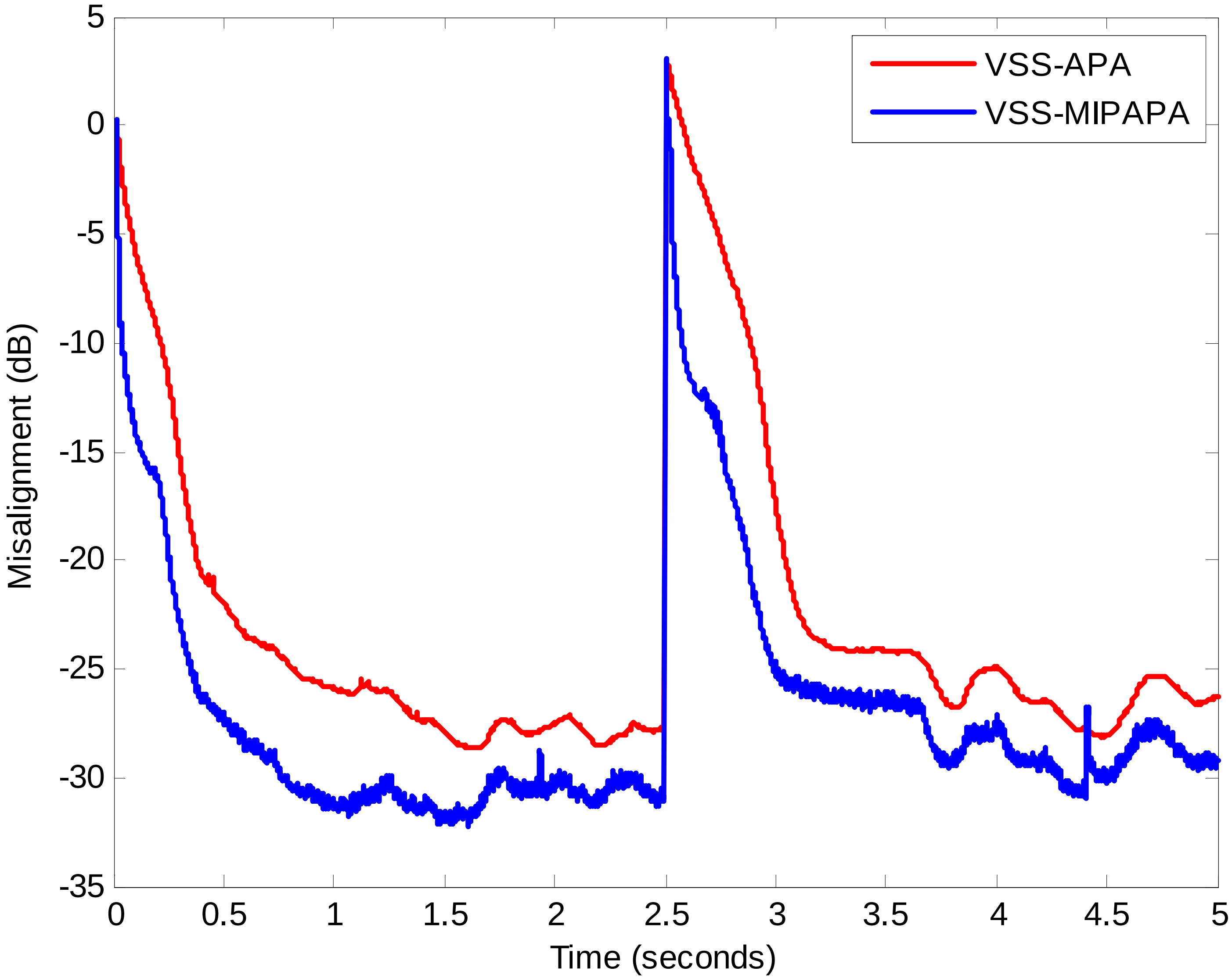


Fig. 3. Misalignment of the VSS-APA and VSS-MIPAPA; P = 4 and speech input signal. The impulse response changes at time 2.5 seconds.

- double-talk scenario

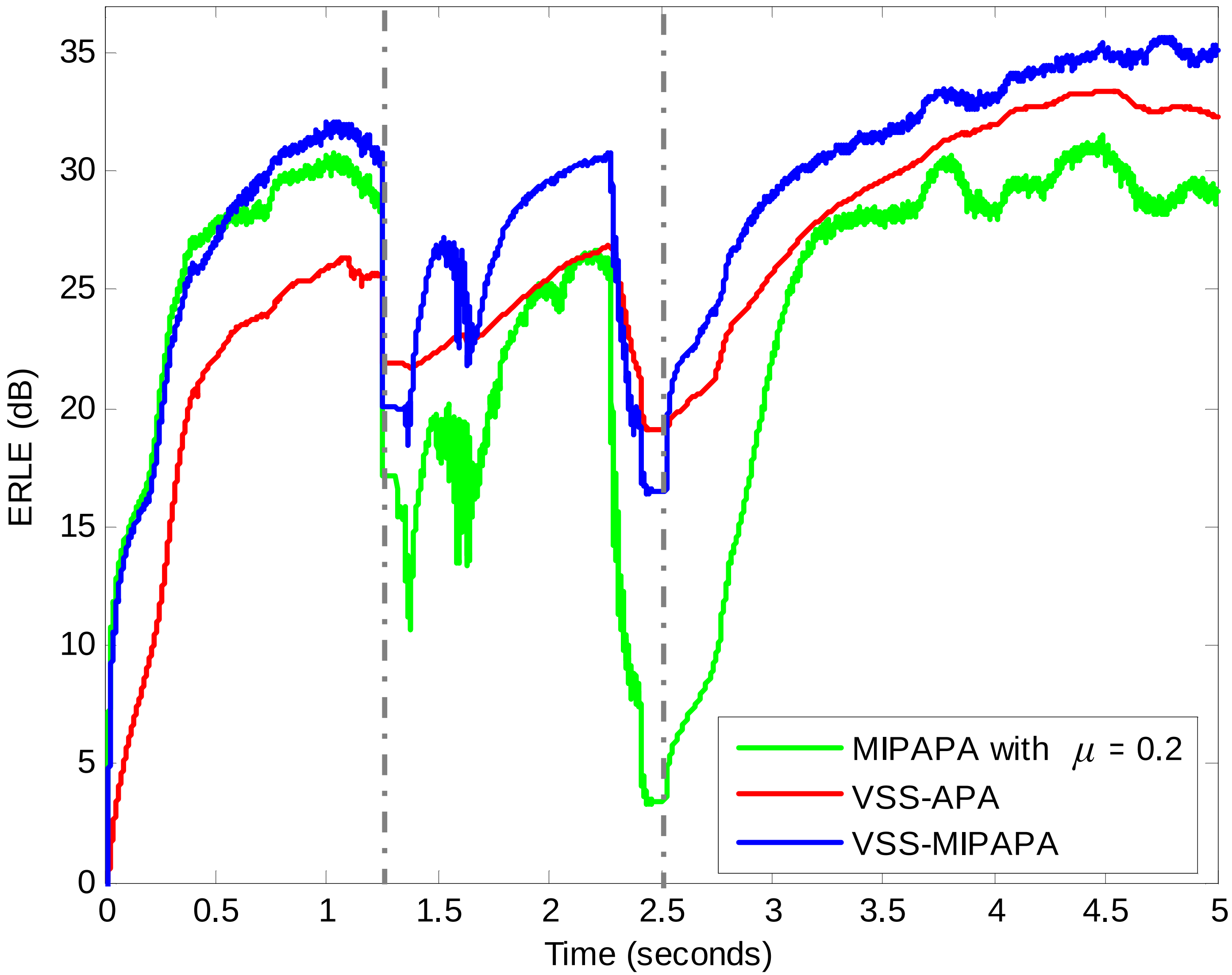


Fig. 4. ERLE of the MIPAPA (μ = 0.2), VSS-APA, and VSS-MIPAPA; P = 4 and speech input signal. Double talk appears between times 1.25 and 2.5 seconds; a Geigel DTD is used.

Conclusions

- a VSS-MPAPA was developed in the context of echo cancellation
- it is more computationally efficient as compared to the classical PAPAs.
- the VSS formula does not require any additional parameters from the environment (i.e., non-parametric)
- it is more robust to near-end signal variations like double-talk

Thank you!