AUTOMATIC SHRINKAGE TUNING ROBUST TO INPUT CORRELATION FOR SPARSITY-AWARE ADAPTIVE FILTERING

Kwangjin Jeong†, Masahiro Yukawa†‡, Masao Yamagishi*, and Isao Yamada*

† Dept. Electronics and Electrical Engineering, Keio University, JAPAN, ‡ Center for Advanced Intelligence Project, RIKEN, JAPAN,* Dept. Information and Communications Engineering, Tokyo Institute of Technology, JAPAN

Main Contributions

- **Automatic parameter tuning** for the APFBS algorithm is proposed.
- Utilizing time-averaged statistics
  ⟶ **Robustness to Input Correlation**

Model and Assumptions

Linear Adaptive Filtering Model

$$d_n = \boldsymbol{u}_n^{\mathsf T}\boldsymbol{h}^* + \epsilon_n \quad (n \in \mathbb{N}: \text{Time index})$$

Observable

- $d_n \in \mathbb{R}$: Output
- $\boldsymbol{u}_n \in \mathbb{R}^m$: Input

Unknown

- $\boldsymbol{h}^* \in \mathbb{R}^m$: Target system
- $\epsilon_n \in \mathbb{R}$: Additive noise

Assumptions

(A1) Noise has zero-mean: $E[\epsilon_n] = 0$

(A2) Input and noise are uncorrelated to each other:

$$E[\epsilon_n \boldsymbol{u}_n] = E[\epsilon_n]E[\boldsymbol{u}_n] \; (= \boldsymbol{0}) \text{ (by A1)}$$

(A3) The norm of the input does not change drastically:

$$E\left[\frac{\epsilon_n \boldsymbol{u}_n}{\|\boldsymbol{u}_n\|_2^2}\right] = \frac{E[\epsilon_n \boldsymbol{u}_n]}{E[\|\boldsymbol{u}_n\|_2^2]}$$

The APFBS Algorithm [1]

Step 1 (Gradient Descent)

$$\boldsymbol{g}_n = \boldsymbol{h}_n - \mu \frac{\boldsymbol{u}_n^{\mathsf T}\boldsymbol{h}_n - d_n}{\|\boldsymbol{u}_n\|_2^2}\boldsymbol{u}_n$$

↑ Step size

Step 2 (Soft Thresholding)

$$\boldsymbol{h}_{n+1} = \sum_{i=1}^{m} \operatorname{sgn}(g_{n,i}) \max\{|g_{n,i}| - \mu\lambda w_{n,i}, 0\}\boldsymbol{e}_i$$

$$= \boldsymbol{A}_n(\lambda)(\boldsymbol{g}_n - \mu\lambda\boldsymbol{v}_n)$$

Shrinkage Parameter (λ): Poor tuning ↓ Poor performance

Weighting coefficient

$$w_{n,i} = \frac{1}{|g_{n,i}| + \nu}$$

$$\boldsymbol{v}_n = \sum_{i=1}^{m} \operatorname{sgn}(g_{n,i}) w_{n,i} \boldsymbol{e}_i$$

$$[\boldsymbol{A}_n(\lambda)]_{ij} = \begin{cases} 1 & i = j \;\&\; |g_{n,i}| > \mu\lambda w_{n,i} \\ 0 & \text{Otherwise} \end{cases}$$

Proposed Shrinkage Tuning

Cost Formulation

$$J_n(\lambda) := (\mu\lambda)^2 \boldsymbol{v}_n^{\mathsf T}\boldsymbol{A}_n(\lambda)\boldsymbol{v}_n - \boldsymbol{g}_n^{\mathsf T}\boldsymbol{A}_n(\lambda)\boldsymbol{g}_n + 2c_n(\boldsymbol{g}_n - \mu\lambda\boldsymbol{v}_n)^{\mathsf T}\boldsymbol{A}_n(\lambda)(\hat{\boldsymbol{R}}_n\boldsymbol{g}_n - \hat{\boldsymbol{p}}_n)$$

$$\approx \|\boldsymbol{h}_{n+1} - \boldsymbol{h}^*\|_2^2 - C$$

System Mismatch ($\|\boldsymbol{h}_{n+1} - \boldsymbol{h}^*\|_2^2$), Constant ($C$)

Time-Averaged Statistics

$$\hat{\boldsymbol{R}}_n = \frac{1}{n}\sum_{k=1}^{n} \frac{\boldsymbol{u}_k\boldsymbol{u}_k^{\mathsf T}}{\|\boldsymbol{u}_k\|_2^2}$$

$$\hat{\boldsymbol{p}}_n = \frac{1}{n}\sum_{k=1}^{n} \frac{d_k}{\|\boldsymbol{u}_k\|_2^2}\boldsymbol{u}_k \approx \hat{\boldsymbol{R}}_n\boldsymbol{h}^* \; (\Leftarrow E[\epsilon_n] = 0)$$

Derivation of Shrinkage Tuning

- $\boldsymbol{A}_n(\lambda)$ takes discrete values.

⟶ $J_n(\lambda)$ **is piecewise quadratic**.

(1) Sort $\left\{0, \frac{|g_{n,1}|}{\mu w_{n,1}}, \frac{|g_{n,2}|}{\mu w_{n,2}}, \cdots, \frac{|g_{n,m}|}{\mu w_{n,m}}\right\}$ into $\{\rho_0, \rho_1, \cdots, \rho_m\}$ in nondecreasing order.

(2) Focus on each interval $[\rho_j, \rho_{j+1}]$. $(j \in \{0, 1, \cdots, m-1\})$

- Minimizer of each quadratic function

$$\lambda_{n,j} = \frac{c_n \boldsymbol{v}_n^{\mathsf T}\boldsymbol{A}_n(\lambda)(\hat{\boldsymbol{R}}_n\boldsymbol{g}_n - \hat{\boldsymbol{p}}_n)}{\mu\boldsymbol{v}_n^{\mathsf T}\boldsymbol{A}_n(\lambda)\boldsymbol{v}_n}$$

- Projection onto the interval $[\rho_j, \rho_{j+1}]$

$$\lambda_{n,j}^* = P_{[\rho_j,\rho_{j+1}]}(\lambda_{n,j}) = \begin{cases} \rho_j & \lambda_{n,j} < \rho_j \\ \lambda_{n,j} & \rho_j \le \lambda_{n,j} \le \rho_{j+1} \\ \rho_{j+1} & \lambda_{n,j} > \rho_{j+1} \end{cases}$$

- Choosing the global minimizer among $\lambda_{n,j}^*$ $(j \in \{0, 1, \cdots, m-1\})$

Overall Algorithm

Initialization: $\boldsymbol{h}_1$, step size μ

Step 1 of the APFBS algorithm

Step 1.5 Shrinkage Tuning

(i) Calculate $\hat{\boldsymbol{R}}_n$ and $\hat{\boldsymbol{p}}_n$.

(ii) Find piecewise minima $\lambda_{n,j}^*$ of $J_n(\lambda)$.

(iii) Among the piecewise minima, find the global minimum λ_n^* of $J_n(\lambda)$.

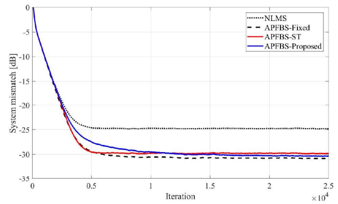
Step 2 of the APFBS algorithm (with λ_n^*)

Numerical Experiments

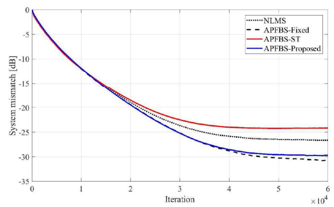
Conditions and Initialization

- Target system: an echo path [2]
- Noise: White Gaussian
- Input: White Gaussian, Colored (AR(1))
- $\boldsymbol{h}_1 = \boldsymbol{0}$, Step size: $\mu = 0.5$

Learning Curves

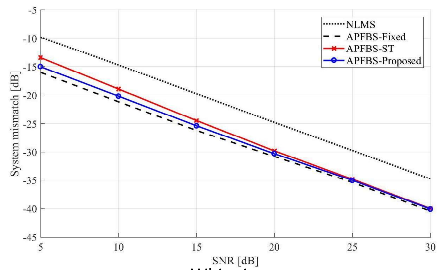


- For white input, **the proposed shrinkage tuning** converges slow, but it achieves **low system mismatch.**

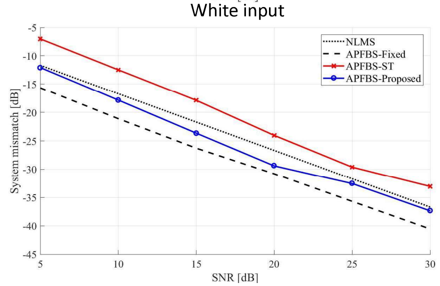


- For colored input, **the proposed shrinkage tuning** works closely to **the manual tuning.**

System Mismatch vs. SNR



White input



Colored input

- For white input, **the proposed shrinkage tuning** works closely to **the manual tuning**.
- For colored input, **the proposed shrinkage tuning is more robust** than **another automatic tuning** given in [3].

Summary

- An **automatic tuning** of a shrinkage parameter for the APFBS algorithm has been proposed.
- At each iteration, the minimizer of a piecewise quadratic cost is chosen as a shrinkage parameter.
- Utilizing **time-averaged statistics**, **the proposed shrinkage tuning achieves robust performance to input correlation**

References

[1] Y. Murakami, M. Yamagishi, M. Yukawa, and I. Yamada, "A sparse adaptive filtering using time-varying soft-thresholding techniques," in Proc. ICASSP 2010.

[2] ITU-T Recommendation G. 168, Digital Network Echo Cancellers, Int. Telecomm. Union, 2015.

[3] M. Yamagishi, M. Yukawa, and I. Yamada, "Automatic shrinkage tuning based on a system-mismatch estimate of sparsity-aware adaptive filtering," in Proc. ICASSP 2017.