Alternating Least-Squares-Based Microphone Array Parameter Estimation for A Single-Source Reverberant and Noisy Acoustic Scenario

Changheng Li and Richard C. Hendriks

PRELIMINARIES

Motivation

- Spatial filtering techniques depend on microphone array parameters: source RTF, source PSDs, late reverberation PSDs and ambient noise PSDs.
- Existing estimators either assume some of these parameters known or suffer from high computational cost.
- A joint estimator at a low complexity is in demand.

Signal model

- For a single source reverberant and noisy scenario (STFT domain),

$$\mathbf{y}(l,k) = \mathbf{a}(l,k)s(l,k) + \mathbf{r}(l,k) + \mathbf{v}(l,k).$$

- The noisy covariance matrix (omitting frequency bin index):

$$\mathbf{\Phi}_{\mathbf{y}}(l) = \phi_s(l)\,\mathbf{a}\mathbf{a}^H + \underbrace{\phi_\gamma(l)\mathbf{\Gamma}}_{\text{Late reverb}} + \underbrace{\phi_v(l)\mathbf{\Psi}}_{\text{Ambient noise}}.$$

- $\mathbf{a}$ is constant for N sequential time frames, $\mathbf{\Gamma}$ and $\mathbf{\Psi}$ are time-invariant known spatial coherence matrices.
- Goal: estimate $\mathbf{a}$, $\phi_s(l)$, $\phi_\gamma(l)$ and $\phi_v(l)$ using $\hat{\mathbf{\Phi}}_{\mathbf{y}}(l)$, given $\mathbf{\Gamma}$ and $\mathbf{\Psi}$ for $l = 1, \cdots, N$.

JALS

Minimize the model mismatch error using the sum of Frobenius norms:

$$\underset{\mathbf{a},\phi_s(l),\phi_\gamma(l),\phi_v(l)}{\arg\min} \sum_{l=1}^{N} \left\| \hat{\mathbf{P}}_{\mathbf{y}}(l) - \phi_s(l)\,\mathbf{a}\mathbf{a}^H - \phi_\gamma(l)\,\hat{\mathbf{\Gamma}} - \phi_v(l)\,\hat{\mathbf{\Psi}} \right\|_F^2.$$

Reparameterization to avoid ambiguous issue: $\tilde{\mathbf{a}} = \frac{\mathbf{a}}{\sqrt{\mathbf{a}^H\mathbf{a}}}$ and $\tilde{\phi}_s(l) = \phi_s(l)\mathbf{a}^H\mathbf{a}$.

Alternating estimation

For given $\tilde{\mathbf{a}}$, estimate the PSDs $\tilde{\phi} = \{\tilde{\phi}_s(l), \tilde{\phi}_\gamma(l), \tilde{\phi}_v(l)\}$ by $\tilde{\phi}(l) = \tilde{\mathbf{\Phi}}_{\tilde{\mathbf{a}}}^{-1}\tilde{\mathbf{b}}(l)$, with

$$\tilde{\mathbf{\Phi}}_{\tilde{\mathbf{a}}} = \begin{bmatrix} 1 & \tilde{\mathbf{a}}^H\hat{\mathbf{\Gamma}}\tilde{\mathbf{a}} & \tilde{\mathbf{a}}^H\hat{\mathbf{\Psi}}\tilde{\mathbf{a}} \\ \tilde{\mathbf{a}}^H\hat{\mathbf{\Gamma}}\tilde{\mathbf{a}} & \mathrm{tr}\left\{\hat{\mathbf{\Gamma}}^H\hat{\mathbf{\Gamma}}\right\} & \mathrm{tr}\left\{\hat{\mathbf{\Gamma}}^H\hat{\mathbf{\Psi}}\right\} \\ \tilde{\mathbf{a}}^H\hat{\mathbf{\Psi}}\tilde{\mathbf{a}} & \mathrm{tr}\left\{\hat{\mathbf{\Gamma}}^H\hat{\mathbf{\Psi}}\right\} & \mathrm{tr}\left\{\hat{\mathbf{\Psi}}^H\hat{\mathbf{\Psi}}\right\} \end{bmatrix}, \tilde{\mathbf{b}}(l) = \begin{bmatrix} \tilde{\mathbf{a}}^H\hat{\mathbf{P}}_{\mathbf{y}}(l)\tilde{\mathbf{a}} \\ \mathrm{tr}\left\{\hat{\mathbf{\Gamma}}^H\hat{\mathbf{P}}_{\mathbf{y}}(l)\right\} \\ \mathrm{tr}\left\{\hat{\mathbf{\Psi}}^H\hat{\mathbf{P}}_{\mathbf{y}}(l)\right\} \end{bmatrix}.$$

For given $\tilde{\phi}$, $\hat{\tilde{\mathbf{a}}}$ = principal eigenvector of $\sum_{l=1}^{N}\hat{\tilde{\phi}}_s(l)\left[\hat{\mathbf{P}}_{\mathbf{y}}(l) - \hat{\phi}_\gamma(l)\,\hat{\mathbf{\Gamma}} - \hat{\phi}_v(l)\,\hat{\mathbf{\Psi}}\right]$.

Robust PSDs constraints

Upper bounds:

$$\tilde{\phi}_s(l) \le \min_m \left\{ \frac{\hat{\mathbf{P}}_{\mathbf{y}_{m,m}}(l)}{\left|\hat{\tilde{a}}_m\right|^2} \right\},\ \phi_\gamma(l) \le \min_m \left\{ \frac{\hat{\mathbf{P}}_{\mathbf{y}_{m,m}}(l)}{\hat{\mathbf{\Gamma}}_{m,m}} \right\},\ \phi_v(l) \le \min_m \left\{ \frac{\hat{\mathbf{P}}_{\mathbf{y}_{m,m}}(l)}{\hat{\mathbf{\Psi}}_{m,m}} \right\}.$$

Lower bounds (avoid negative estimates):

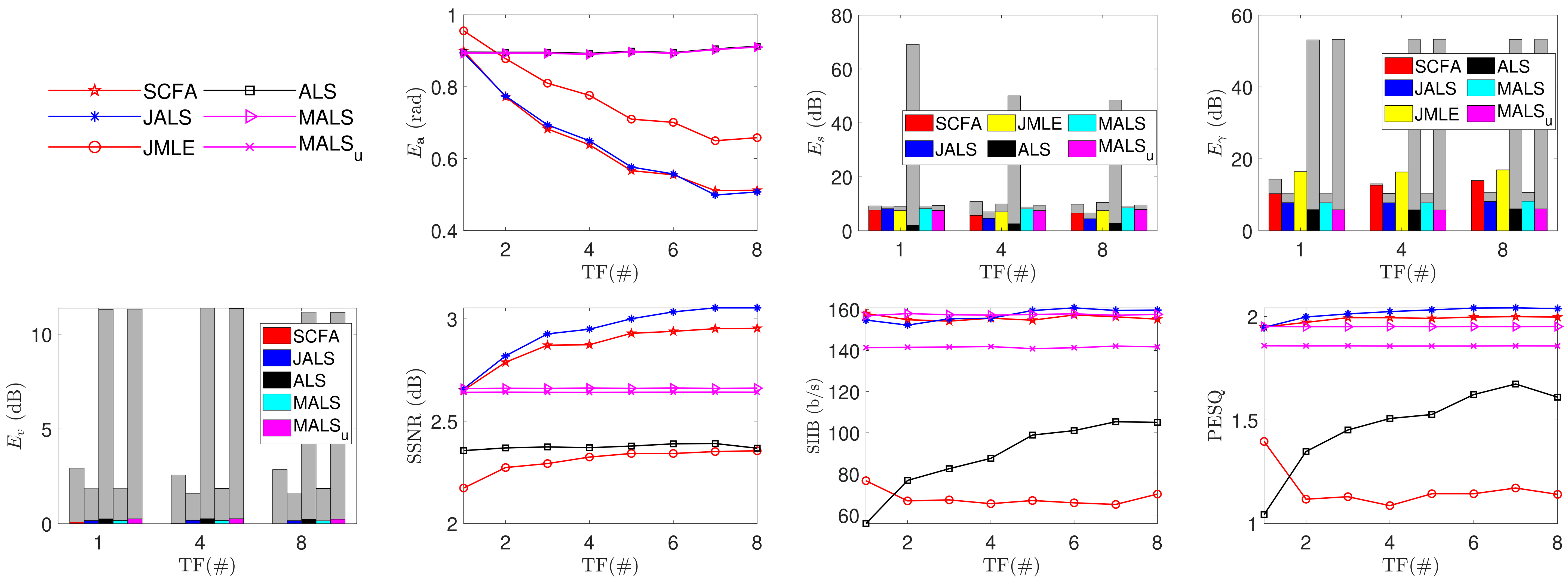
e.g. if $\hat{\phi}_\gamma(l) > 0$ while $\hat{\tilde{\phi}}_s(t) \le 0$, we lower bound $\tilde{\phi}_s(l)$ by

$$\hat{\phi}_\gamma(l) \min_{t=1,\cdots,l} \frac{\hat{\tilde{\phi}}_s(t)}{\hat{\phi}_\gamma(t)} \le \tilde{\phi}_s(l).$$

EXPERIMENTS

$$E_{\mathbf{a}} = \frac{1}{L(K/2+1)}\sum_{l=1}^{L}\sum_{k=1}^{K/2+1} \mathrm{acos}\left(\frac{\left|\mathbf{a}^H(l,k)\hat{\mathbf{a}}(l,k)\right|}{\|\mathbf{a}^H(l,k)\|_2\|\hat{\mathbf{a}}(l,k)\|_2}\right)\ (\mathrm{rad}),\ E_i = 10\log_{10}\left(\phi_i/\hat{\phi}_i\right)$$

method	SCFA[1]	ALS[2]	MALS	JMLE	JALS
Normalized run time	154.65	6.27	5.7	1.66	1



[1] A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, and J. Jensen, "Robust joint estimation of multimicrophone signal model parameters," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 7, pp. 1136–1150, 2019.

[2] M. Tammen, S. Doclo, and I. Kodrasi, "Joint Estimation of RETF Vector and Power Spectral Densities for Speech Enhancement Based on Alternating Least Squares," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 795–799.

TU Delft Delft University of Technology

Signal Processing Systems (SPS) Group