On Optimal Smoothing in Minimum Statistics Based Noise Tracking

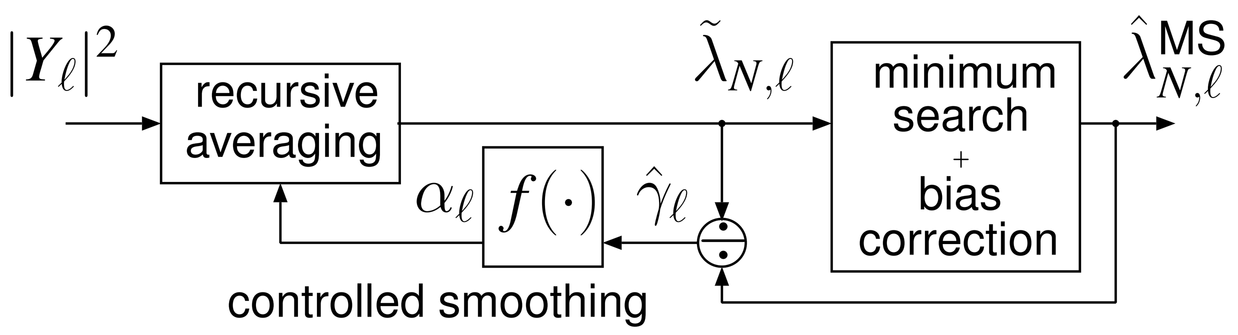
UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Aleksej Chinaev and Reinhold Haeb-Umbach, University of Paderborn, Germany, {chinaev, haeb}@nt.upb.de, http://nt.upb.de

Computer Science, Electrical Engineering and Mathematics
NT Communications Engineering
Prof. Dr.-Ing. Reinhold Häb-Umbach

Introduction

- Noise power spectral density (PSD) estimation is a key component of speech enhancement systems
- Minimum Statistics (MS) approach [1]



with the 1. order recursive averaging

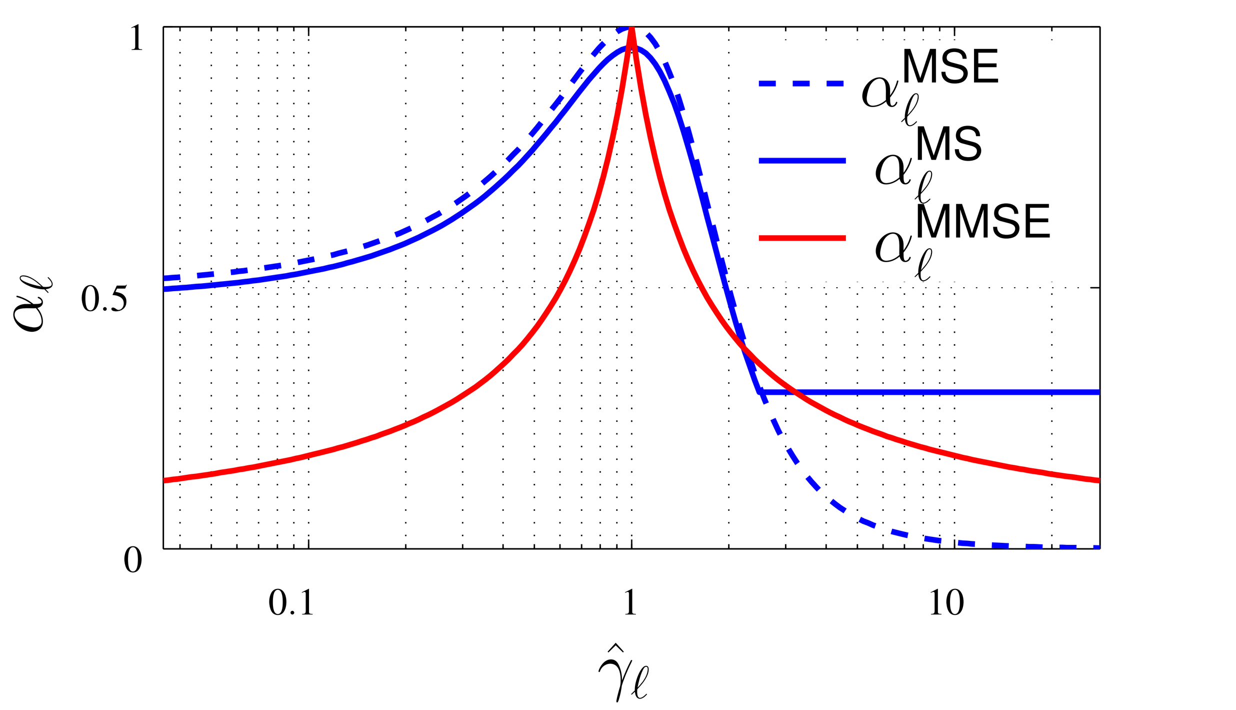
$$\tilde{\lambda}_{N,\ell} = \alpha_\ell \cdot \tilde{\lambda}_{N,\ell-1} + (1-\alpha_\ell)\cdot |Y_\ell|^2 \qquad (1)$$

controlled by smoothed *a posteriori* SNR

$$\hat{\gamma}_\ell = \frac{\tilde{\lambda}_{N,\ell-1}}{\hat{\lambda}^{\text{MS}}_{N,\ell-1}} \quad \text{via} \quad \alpha_\ell = f(\gamma_\ell).$$

- We propose a new control function $f(\cdot)$

Compare of control functions



- Properties of control functions $\alpha_\ell = f(\gamma_\ell)$

$\hat{\gamma}_l$	-	α_ℓ^{MSE}	α_ℓ^{MS}	$\alpha_\ell^{\text{MMSE}}$
< 1	tracking ability	☹	☹	☺
≈ 1	deadlock	☹	☺	☺
$\gg 1$	estimator variance	☹	☺	☺
	extra parameters	☺	☹	☺

MS control function [1]

- Likelihood in speech absence

$$p_{|Y_\ell|^2}(x;\lambda_{N,\ell}) = \frac{1}{\lambda_{N,\ell}}\cdot e^{-\frac{x}{\lambda_{N,\ell}}},\ x > 0$$

$\lambda_{N,\ell}$ - unknown fixed parameter

- Optimal smoothing in MSE sense for (1)

$$\alpha_\ell^{\text{MSE}} = \underset{\alpha}{\text{argmin}}\ E\left[\left(\tilde{\lambda}_{N,\ell} - \lambda_{N,\ell}\right)^2 \middle| \tilde{\lambda}_{N,\ell-1}\right]$$

$$\alpha_\ell^{\text{MSE}} = \frac{1}{1+(\hat{\gamma}_\ell - 1)^2}$$

- Heuristic mitigation of some drawbacks

$$\alpha_\ell^{\text{MS}} = \max\left(\alpha_{\max}\cdot\alpha_\ell^{\text{MSE}},\ \alpha_{\min}(\text{SNR})\right)$$

- α_ℓ^{MS} still not able to track $\lambda_{N,\ell}$ for $\hat{\gamma}_\ell < 1$
  - crucial for the following minimum search

Proposed control function

- Treat $\lambda_{N,\ell}$ as a random variable

$$p_{\lambda_{N,\ell}}(x;\nu_\ell,\tau_\ell^2) = \text{Scaled-Inv-}\chi^2\left(\nu_\ell,\tau_\ell^2\right)$$

- Update of hyperparameter τ_ℓ^2 in speech absence for MMSE point estimate

$$\tilde{\lambda}^{\text{MMSE}}_{N,\ell} = E\left[\lambda_N \,\middle|\, |Y_\ell|^2\right] = \frac{\nu_\ell}{\nu_\ell - 2}\cdot\tau_\ell^2$$

gives (1) with the smoothing parameter

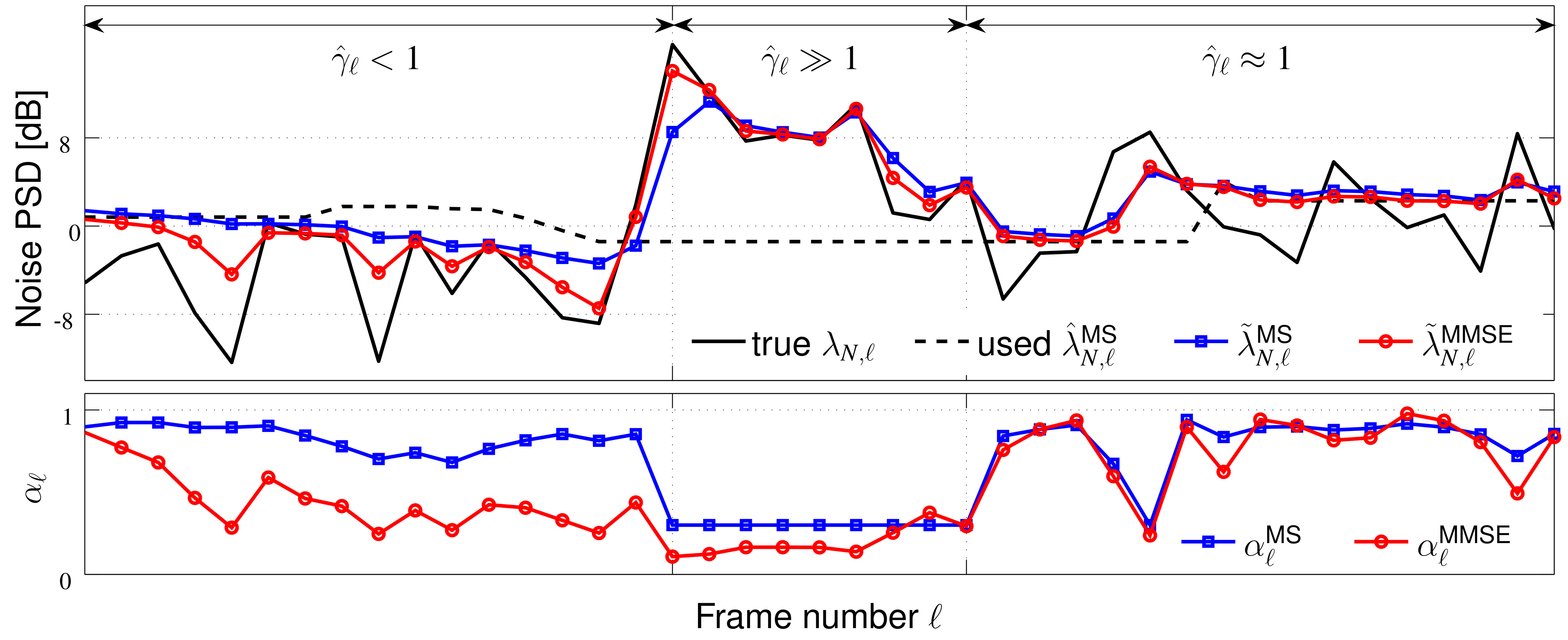
$$\alpha_\ell^{\text{MMSE}} = 1 - \frac{2}{\nu_{\ell-1}} \quad \text{for} \quad \nu_{\ell-1} > 2$$

- Let $d_\ell = |\ln(\hat{\gamma}_\ell)|$ be a measure of the degree of nonstationarity.
  Adjust smoothing parameter $\alpha_\ell^{\text{MMSE}}$ via

$$\nu_{\ell-1} = 2 + \frac{1}{d_\ell}$$

to depend on d_ℓ

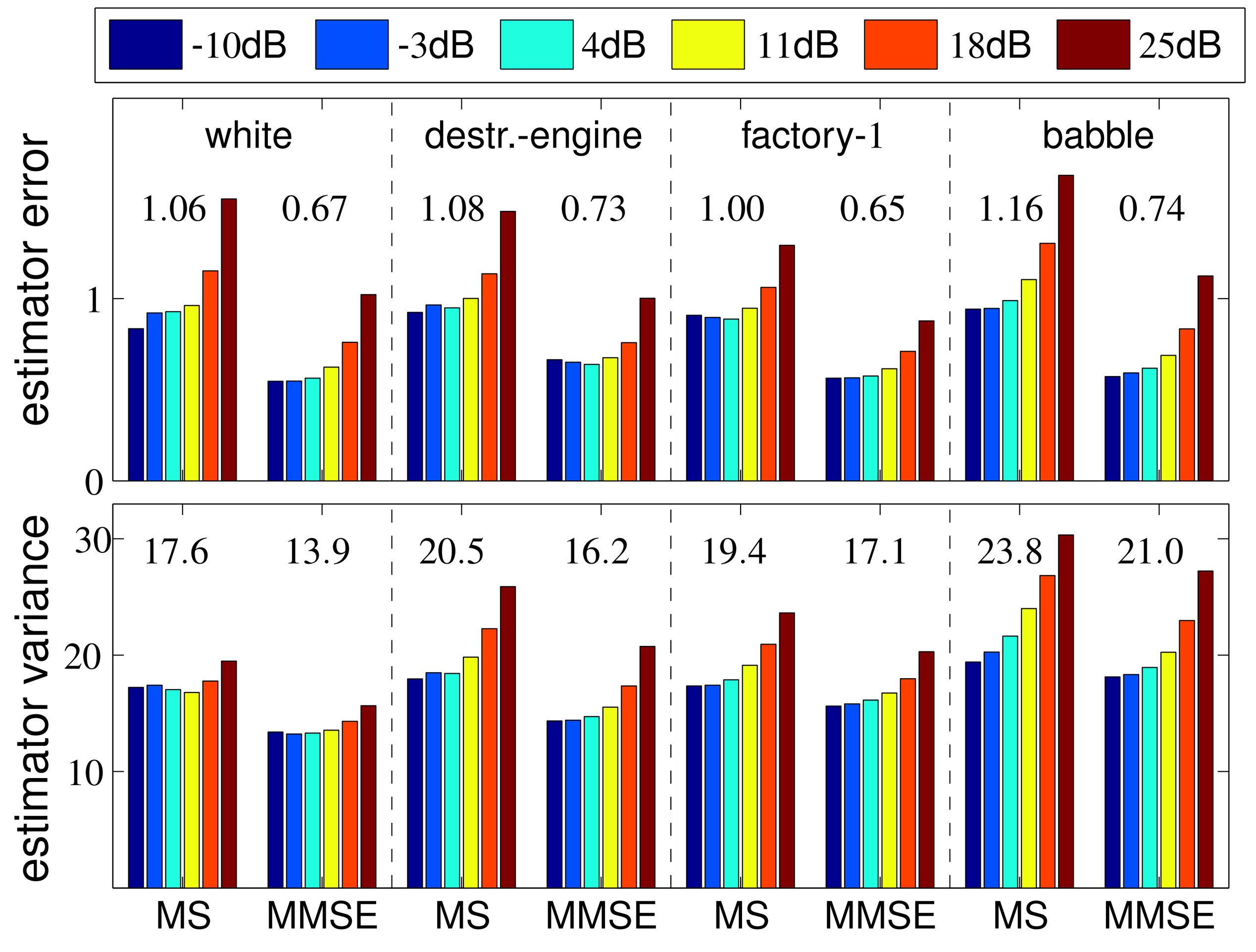
$$\alpha_\ell^{\text{MMSE}} = \frac{1}{1+2\cdot|\ln\hat{\gamma}_\ell|}$$



Noise PSD trajectories (top) and corresponding smoothing parameters (bottom)

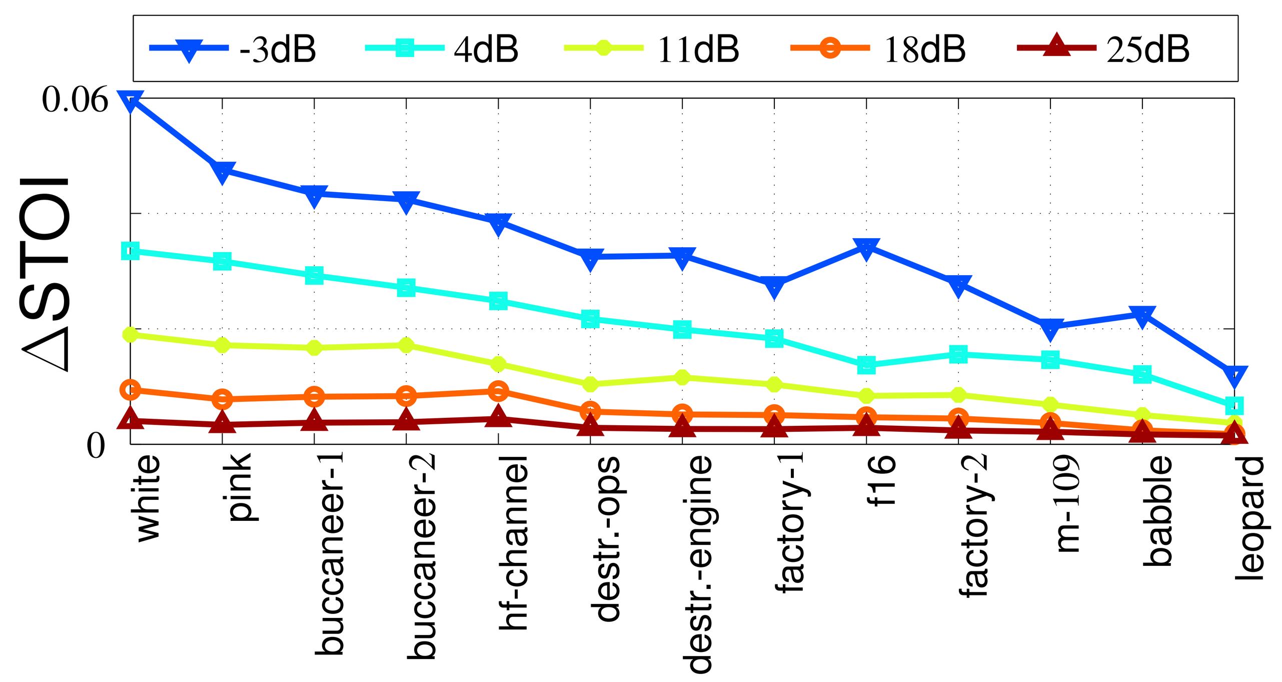
Experimental results

- Reduction of mean estimator error and estimator variance



Noise tracking performance

- Small improvement of speech quality



$\Delta\text{STOI} = \text{STOI}^{\text{MMSE}} - \text{STOI}^{\text{MS}}$

Conclusions

- We proposed the MMSE smoothing controlled by a degree of nonstationarity measure for the MS approach [1]
- Improved noise tracking and speech enhancement performance

Reference: [1] R. Martin, 'Noise power spectral density estimation based on optimal smoothing and minimum statistics', *IEEE Transactions on Speech and Audio Processing,* vol. 9, no. 5, pp. 504-512, July 2001.