UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

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

Quality Analysis and Optimization of the MAP-based Noise Power Spectral Density Tracker

Aleksej Chinaev and Reinhold Haeb-Umbach
University of Paderborn, Paderborn, Germany; (chinaev,haeb)@nt.uni-paderborn.de

Introduction

- Maximum a-posteriori based (MAP-B) noise power spectral density (PSD) estimation algorithm
- Quality analysis of MAP-B estimator (Monte Carlo method): unbiasedness, tracking ability and sensitivity
- Optimization of MAP-B noise PSD tracker and its evaluation in a single-channel speech enhancement system

MAP-B Tracker and Simulation Framework

- MAP-B algorithm: noise PSD estimation $\hat{\sigma}^2_{N,l}$ of short-time Fourier transform (STFT) N_l at frame l and one frequency bin, given
  - Noisy speech observation $Y_l = X_l + N_l$
  - Estimate of clean speech PSD $\tilde{\sigma}^2_{X,l}$
- $\hat{\sigma}^2_{N,l} = \underset{\sigma^2_{N,l}}{\operatorname{argmax}}\, p(\sigma^2_{N,l}|Y_l)$ with $p(\sigma^2_{N,l}|Y_l) \propto p(Y_l|\sigma^2_{N,l}) \cdot p(\sigma^2_{N,l})$
  - A-priori PDF $p(\sigma^2_{N,l})$ is modeled as the Scale-inv-$\chi^2(\sigma^2_{N,l}; \nu_0, \lambda^2_l)$

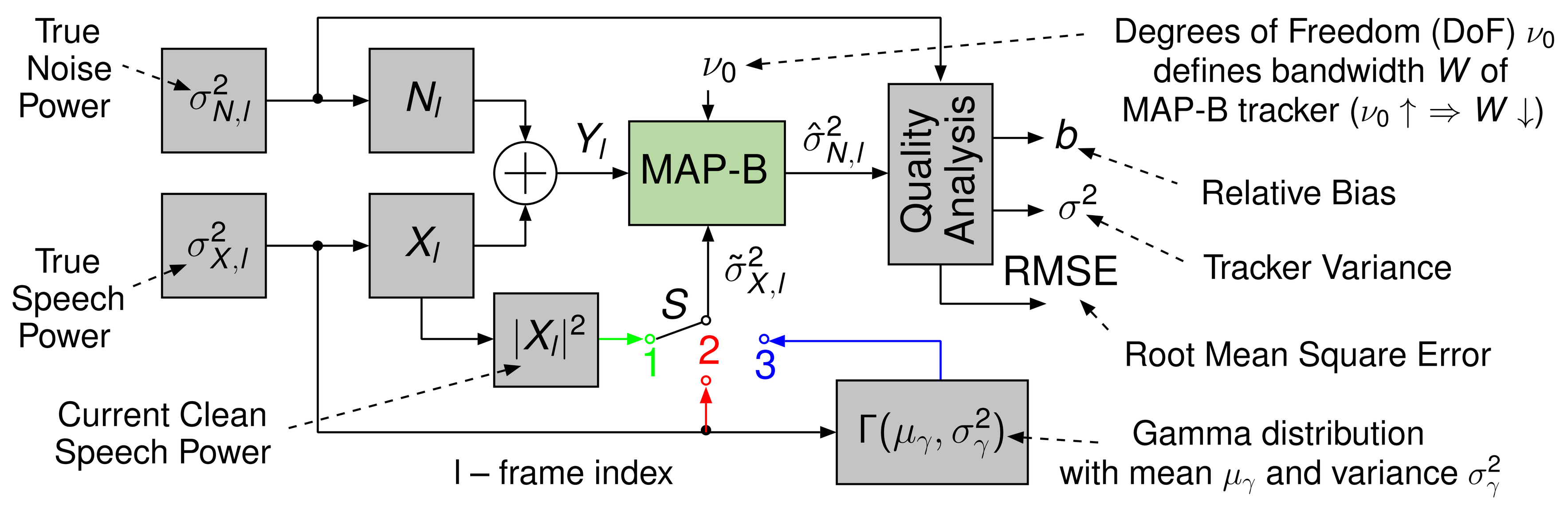

Figure 1: Simulation framework for quality analysis

Quality Analysis of MAP-B Estimator

- Switch $S = 1$ (unbiasedness): stationary noise $\sigma^2_{N,l} = \sigma^2_N$
  - Relative bias $b = \frac{E[\hat{\sigma}^2_{N,l}] - \sigma^2_N}{\sigma^2_N}$ is always positive, Fig.2 (a)
- Switch $S = 2$ (tracking ability): $\sigma^2_{N,l} = 1 + \sin^2(2\pi l/L)$:
  - RMSE grows both for small and for large ν_0 values, Fig.2 (b)
  - $\nu_0 = 40$ seems to be a good choice to minimize RMSE
- Switch $S = 3$ (sensitivity to $\tilde{\sigma}^2_{X,l}$): $\sigma^2_{N,l} = \sigma^2_N, \nu_0 = 40, \mu_\gamma = \sigma^2_{X,l}$
  - MAP-B estimator holds its estimation variance σ^2 at a low constant value for a wide range of σ^2_γ values, Fig.2 (c)

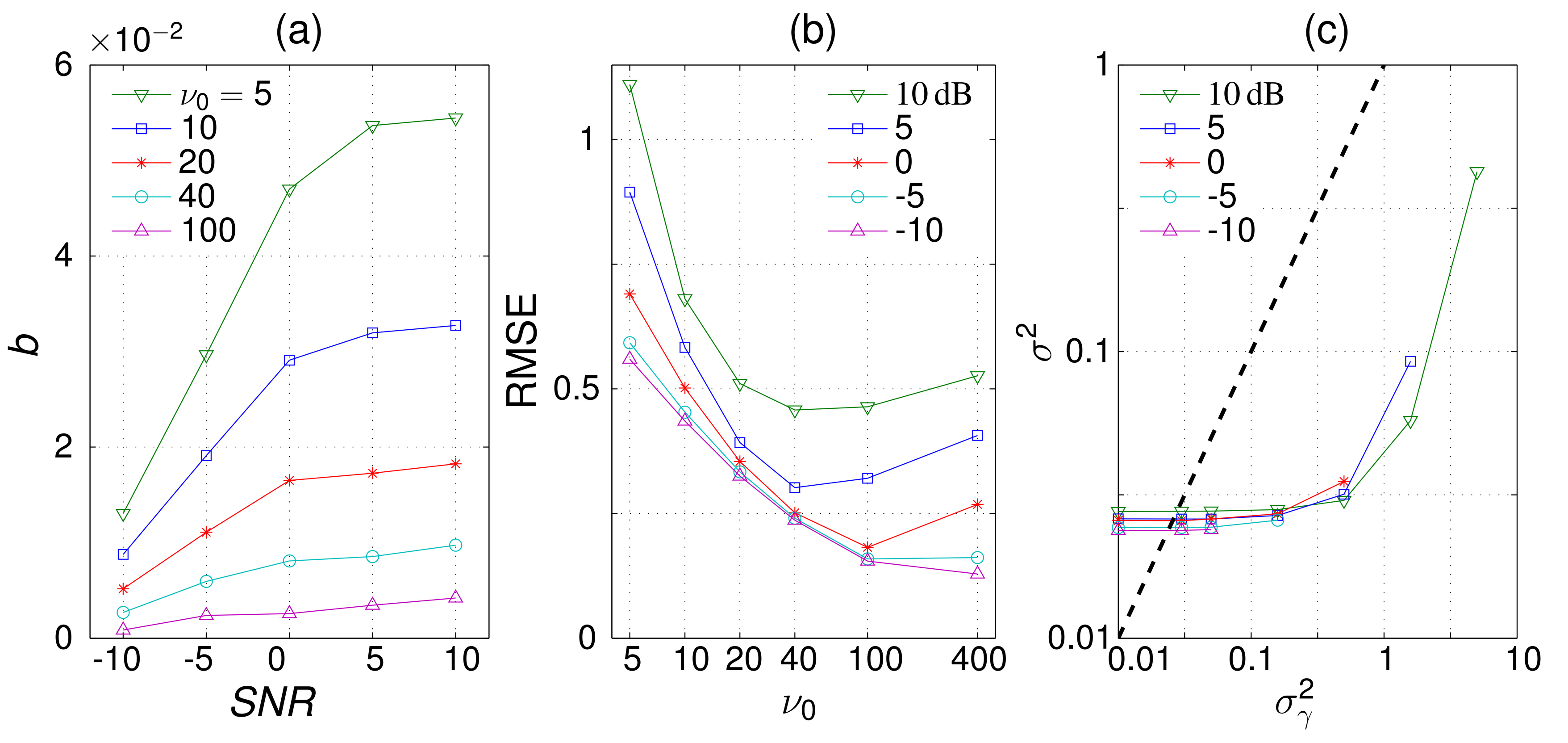

Figure 2: (a) Relative bias b over SNR for $\nu_0 \in \{5, 10, 20, 40, 100\}$, (b) RMSE over ν_0 for $SNR \in \{$-10, -5, 0, 5, 10$\}$dB and (c) Variance σ^2 over σ^2_γ for $SNR \in \{$-10, -5, 0, 5, 10$\}$dB. The black dashed line in (c) corresponds to the variance of the estimator based on the trivial estimation rule $\hat{\sigma}^2_Y = \tilde{\sigma}^2_X + \hat{\sigma}^2_N$, which results in $\sigma^2 = \sigma^2_\gamma$

Optimization of MAP-B Tracker

- According to the results of the quality analysis we suggest two improvements of the MAP-B tracker:
  - Bias reduction (unbiased estimate $\hat{\sigma}^2_{N,UN}$ from biased $\hat{\sigma}^2_N$)

$$\hat{\sigma}^2_{N,UN} = \left(1 - \beta_{max} \cdot \left(\frac{\arctan(\mathrm{SNR})}{\pi} + \frac{1}{2}\right)\right) \cdot \hat{\sigma}^2_N \quad (1)$$

  with a bias compensation factor $\beta_{max} = 0.01$, see Fig.2 (a)
  - Bandwidth adjustment (countering systematic errors in $\tilde{\sigma}^2_{X,l}$)

$$\nu_0(\mathrm{SNR}) = \nu_0 + \frac{\Delta\nu_0}{\pi} \cdot \arctan(\mathrm{SNR}) \quad (2)$$

  with $\nu_0 = 40$ and a DoF adjustment range $\Delta\nu_0 = 10$

Experimental Evaluation

- Smoothed estimation of a-priori SNR of decision-directed approach $\tilde{\zeta}_{k,l} = 0.7 \cdot \tilde{\zeta}_{k,l-1} + 0.3 \cdot \tilde{\xi}_{k,l}$ is used in (1) and (2)

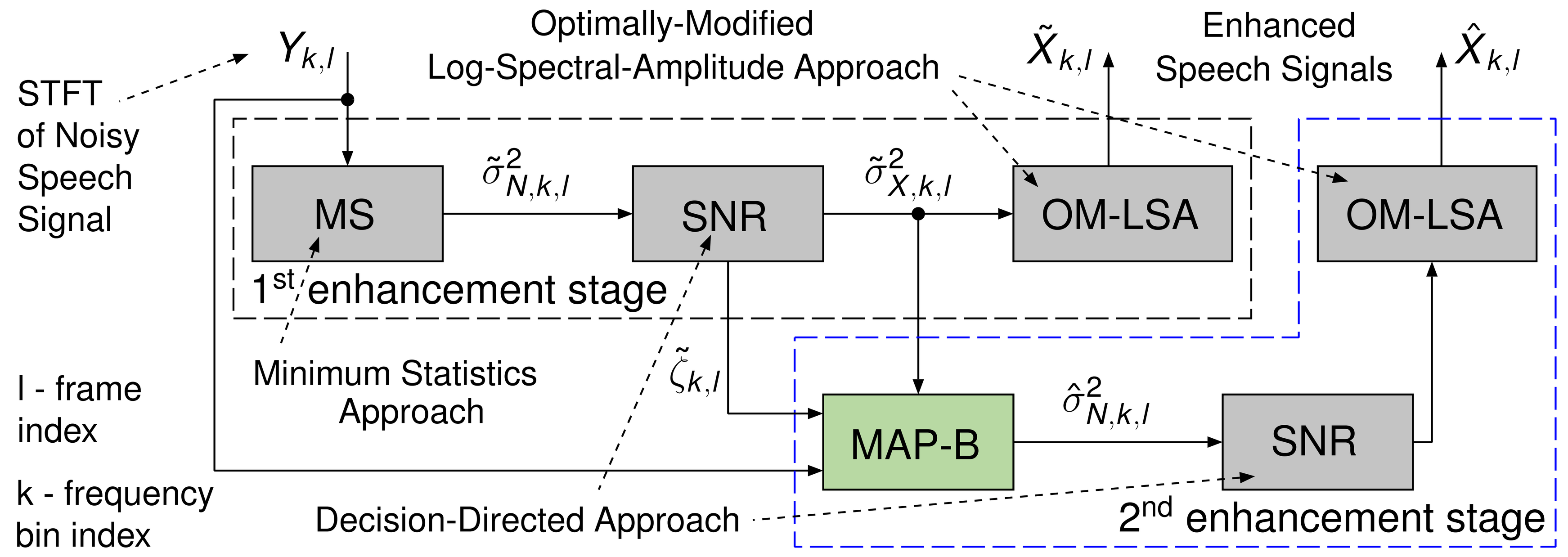

Figure 3: MAP-B postprocessor in a single-channel speech enhancement system

- The reduction of the log-spectral distance $R_{SD} = \tilde{S}_D - \hat{S}_D$ between the true and the estimated noise PSD is used with

$$\check{S}_D = \frac{1}{L}\sum_{l=1}^{L}\left[\frac{2}{K}\sum_{k=1}^{K/2} 10\log_{10}\frac{\sigma^2_{N,k,l}}{\check{\sigma}^2_{N,k,l}}\right]^{-\frac{1}{2}} \quad \text{for} \quad \check{} \in \{\tilde{},\hat{}\}$$

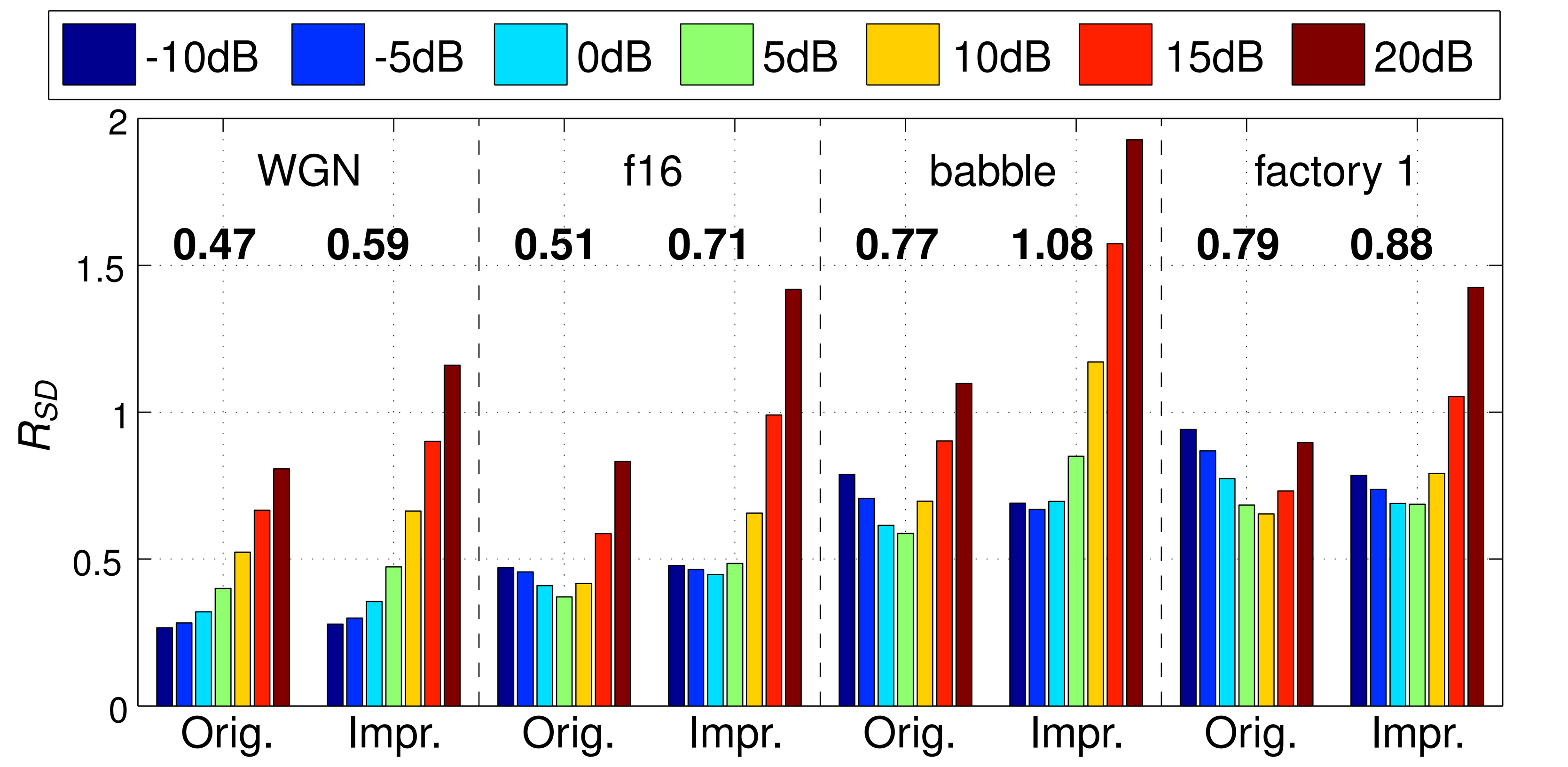

Figure 4: R_{SD} values calculated by using the original and the improved MAP-B estimates. Values in bold are R_{SD} values averaged over all considered SNRs

Conclusions

- Quality analysis reveals that the original MAP-B tracker:
  - slightly overestimates the noise power
  - keeps its low variance for a wide range of zero-mean errors of the input speech power
- Optimization of the MAP-B tracker by bias compensation and bandwidth adjustment leads to improved noise tracking