# IMPROVED SINGLE-CHANNEL NONSTATIONARY NOISE TRACKING BY AN OPTIMIZED MAP-BASED POSTPROCESSOR

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

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

Aleksej Chinaev[1], Reinhold Haeb-Umbach[1], Jalal Taghia[2], Rainer Martin[2]

[1] University of Paderborn, Department of Communications Engineering, 33098 Paderborn, Germany, {chinaev,haeb}@nt.uni-paderborn.de
[2] Ruhr-Universität Bochum, Institute of Communication Acoustics, 44780 Bochum, Germany, {jalal.taghia,rainer.martin}@rub.de

## Introduction

- An optimized Maximum A-Posteriori Based (MAP-B) noise power spectral density (PSD) estimator
- An extensive performance evaluation for nonstationary noise

## MAP-B Noise Tracker

- Idea
  - MAP-B estimate of noise variance in the presence of noise observations "corrupted" by speech [Chinaev, et al., 2012]
- Input
  - Noisy speech $Y_{k\ell} = X_{k\ell} + N_{k\ell}$
  - Speech power estimate of $\sigma^2_{X,k\ell} = E[|X_{k\ell}|^2]$
  - Smoothed SNR estimate

$$\hat{\zeta}_{k\ell} = \alpha_\zeta \cdot \hat{\zeta}_{k,\ell-1} + (1-\alpha_\zeta) \cdot \hat{\xi}_{k\ell}$$

- Output
  - MAP noise PSD estimate $\sigma^2_{N,k\ell} = E[|N_{k\ell}|^2]$

## MAP-B Optimization

- Bias reduction

$$\hat{\hat{\sigma}}^2_{N,k\ell} = \left(1 - \beta(\hat{\zeta}_{k\ell})\right) \cdot \tilde{\sigma}^2_{N,k\ell}$$

with a bias compensation factor $\beta(\hat{\zeta}_{k\ell})$ and a biased MAP-B estimate $\tilde{\sigma}^2_{N,k\ell}$

- Tracking bandwidth reduction at high SNR

$$\nu(\hat{\zeta}_{k\ell}) = \nu_0 + \frac{\Delta\nu}{\pi} \cdot \arctan\left(\hat{\zeta}_{k\ell}\right)$$

with a constant degrees of freedom $\nu_0 = 40$ and an adjustment range $\Delta\nu = 10$

## Speech Enhancement System

- MAP-B noise PSD estimator as a postprocessor

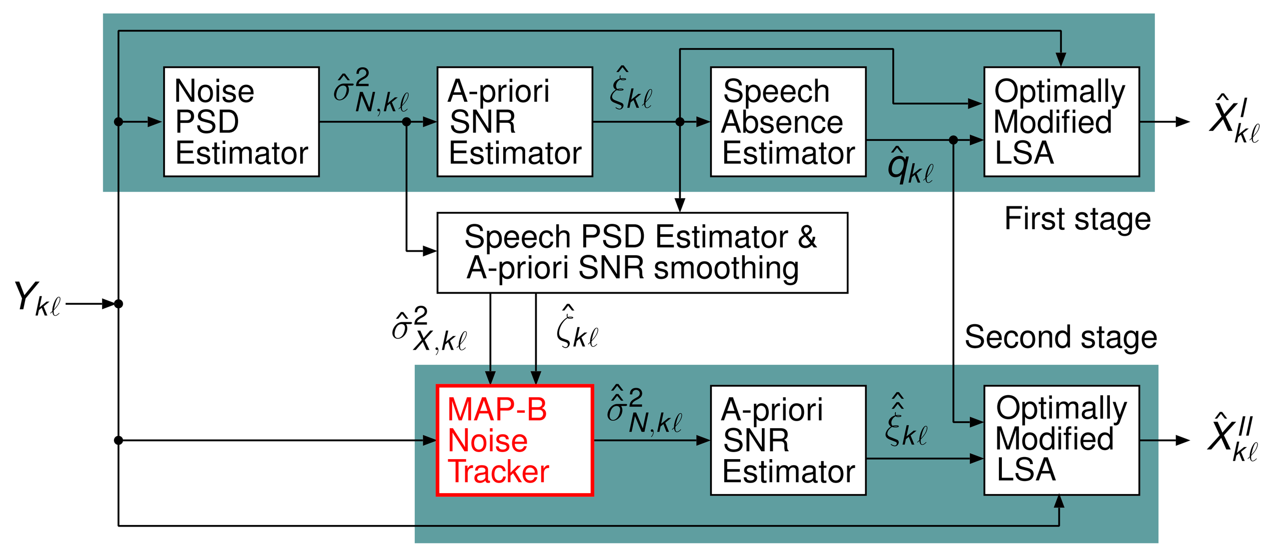



Fig. 1: Two-stage single-channel speech enhancement system

## Experimental Setup

- 3 nonstationary noise types
- 8 state-of-the-art noise PSD estimators in the first stage
- Evaluation of MAP-B gain in
  - noise PSD tracking performance
  - quality of estimated speech
- Overall performance measures
  - Estimation error reduction

$$\Delta\text{LogErr}_{\text{mean}} = \text{LogErr}^{(I)}_{\text{mean}} - \text{LogErr}^{(II)}_{\text{mean}}$$

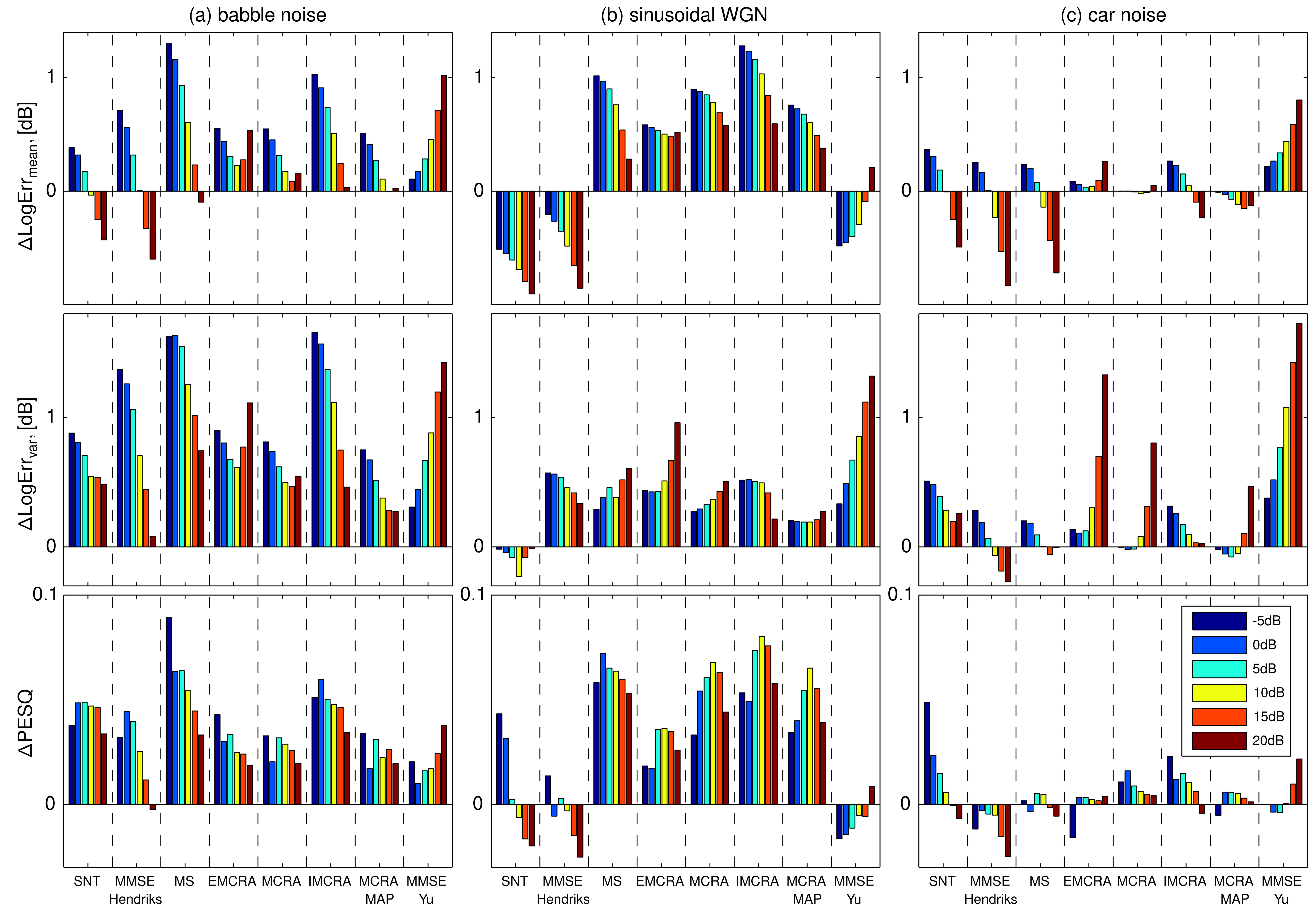
  - Estimator's variance reduction

$$\Delta\text{LogErr}_{\text{var}} = \text{LogErr}^{(I)}_{\text{var}} - \text{LogErr}^{(II)}_{\text{var}}$$

  - Speech quality improvement

$$\Delta\text{PESQ} = \text{PESQ}^{(II)} - \text{PESQ}^{(I)},$$

- For all performance measures
  - the larger values show better performance




Noise trackers used in the first stage:

- SNT - Subspace Noise Tracking, [Hendriks, 2008]
- MMSE-Hendriks, [Hendriks, 2010]
- MS: Minimum Statistics, [Martin, 2001]
- EMCRA: Enhanced MCRA, [Fan, 2007]
- MCRA: Minima Controlled Recursive Averaging, [Cohen, 2002]
- IMCRA: Improved MCRA, [Cohen, 2003]
- MCRA-MAP, [Kum, 2009]
- MMSE-Yu, [Yu, 2009]

## Experimental Results

- Reduction of noise estimation error variance for almost all setups
- Slight improvement in PESQ scores
- The more nonstationary the noise the larger the improvement

## Conclusions

- Second stage with MAP-based noise tracker vs. first stage
  - Improved noise tracking and speech quality in nonstationary noise
  - Computational effort roughly doubled
  - No additional latency

International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2013, Vancouver, Canada

Work supported by DFG under contract no. HA 3455/8-1