On the Statistics of Spectral Amplitudes After Variance Reduction by Temporal Cepstrum Smoothing and Cepstral Nulling

Timo Gerkmann and Rainer Martin, *Senior Member, IEEE*

***Abstract*—In this paper we derive the signal power bias that arises when spectral amplitudes are smoothed by reducing their variance in the cepstral domain (often referred to as *cepstral smoothing*) and develop a power bias compensation method. We show that if χ-distributed spectral amplitudes are smoothed in the cepstral domain, the resulting smoothed spectral amplitudes are also approximately χ-distributed but with more degrees of freedom and less signal power. The key finding for the proposed power bias compensation method is that the degrees of freedom of χ-distributed spectral amplitudes are directly related to their average cepstral variance. Furthermore, this work gives new insights into the statistics of the cepstral coefficients derived from χ-distributed spectral amplitudes using tapered spectral analysis windows. We derive explicit expressions for the variance and covariance of correlated χ-distributed spectral amplitudes and the resulting cepstral coefficients, parameterized by the degrees of freedom. The results in this work allow for a cepstral smoothing of spectral quantities without affecting their signal power. As we assume the parameterized χ-distribution for the spectral amplitudes, the results hold for Gaussian, super-Gaussian, and sub-Gaussian distributed complex spectral coefficients. The proposed bias compensation method is computationally inexpensive and shown to work very well for white and colored signals, as well as for rectangular and tapered spectral analysis windows.**

***Index Terms*—Cepstral analysis, time-frequency analysis, smoothing methods, bias compensation.**

I. Introduction

In many applications of statistical signal processing, a variance reduction of spectral quantities derived from time domain signals, such as the periodogram, is required [1], [2]. If a spectral quantity P is χ^2-distributed

$$p(P) = \frac{1}{\Gamma(\mu)} \left(\frac{\mu}{\sigma^2}\right)^{\mu} P^{\mu-1} \exp\!\left(-\frac{\mu}{\sigma^2} P\right), \qquad (1)$$

with 2μ degrees of freedom, mean $\mathrm{E}\{P\} = \sigma^2$, and variance $\mathrm{var}\{P\} = \sigma^4/\mu$, it is well known that a moving average smoothing of P over time and/or frequency results in an approximately χ^2-distributed random variable with the same mean and an increase in the degrees of freedom that goes along with the decreased variance [3], [4]. The χ^2-distribution holds exactly if the averaged values of P are uncorrelated. A drawback of smoothing in the frequency domain is that the temporal and/or frequency resolution is reduced. In speech processing this may not be desired as the temporal smoothing smears speech onsets and frequency smoothing reduces the resolution of speech harmonics. It has been recently shown that reducing the variance of spectral quantities in the cepstral domain outperforms a smoothing in the spectral domain because specific characteristics of speech signals can be taken into account [5], [6]. In the cepstral domain, speech is mainly represented by the lower cepstral coefficients that represent the spectral envelope, and a peak in the upper cepstral coefficients that represents the fundamental frequency and its harmonics [7]. Therefore, a variance reduction can be applied to the remaining cepstral coefficients without distorting the speech signal. In general, a cepstral variance reduction (CVR) can be achieved by either selectively smoothing cepstral coefficients over time (temporal cepstrum smoothing – TCS) [5], [6], or by setting those cepstral coefficients to zero that are below a certain variance threshold (cepstral nulling – CN) [8], [9]. A comprehensive analysis of TCS is given in [10].

However, the application of an unbiased smoothing process in the cepstral domain leads to a bias in the spectral domain: the CVR does not only change the variance of a χ^2-distributed spectral random variable P, but also its mean $\mathrm{E}\{P\} = \sigma^2$. If $P = |S|^2$ is the periodogram of a complex-valued zero-mean variable S for instance, changing the mean of the periodogram $\mathrm{E}\{|S|^2\}$ changes the signal power of S. As this is usually an undesired side-effect of CVR, in this work we present a framework to compensate for the bias in signal power. Further, we show that the distribution of spectral amplitudes after CVR can be well approximated by a χ-distribution. As opposed to the bias correction in [10], the bias correction proposed here also holds for spectrally correlated spectral coefficients, is computationally far less expensive, and is applicable to both TCS and CN.

After the definition of the cepstrum in Section II, we discuss the statistical properties of the log-periodogram and of cepstral coefficients in Section III for several spectral analysis windows. In Section IV we show how the degrees of freedom after CVR can be determined and how the signal power bias can be compensated. This procedure is summarized in Algorithm 1. In Section V we discuss the mean of the cepstral coefficients. In Section VI we apply the proposed bias compensation in a practical scenario. Section VII concludes this paper.

The authors are with the Institute of Communication Acoustics (IKA), Ruhr-Universität Bochum, 44780 Bochum, Germany (e-mail: timo.gerkmann@rub.de; rainer.martin@rub.de).

II. Definition of Cepstral Coefficients

We consider the cepstral coefficients derived from the discrete short-time Fourier transform $S_k(l)$ of a discrete time domain signal $s(n)$, where n is the discrete time index, k is the discrete frequency index, and l is the segment index. After segmentation the time domain signal is weighted with a window w_n and transformed into the discrete Fourier domain, as

$$S_k(l) = \sum_{n=0}^{N-1} w_n \, s(l\,L + n) \, e^{-\jmath 2\pi k n/N} \,, \tag{2}$$

where L is the number of samples between segments, and N is the segment size. The inverse discrete Fourier transform of the logarithm of the periodogram yields the cepstral coefficients

$$s_q(l) = \frac{1}{N} \sum_{k=0}^{N-1} \log\left(|S_k(l)\,|^2\right) e^{\jmath 2\pi k q/N} \,, \tag{3}$$

where q is the cepstral index, also known as the *quefrency* index [11]. As the log-periodogram is real-valued, the cepstrum is symmetric with respect to $q = N/2$. Therefore, in the following we will only discuss the lower symmetric part $q \in \{0, 1, \ldots, N/2\}$. For improved readability, we drop the segment index l wherever possible.

III. Statistical Properties of the log-Periodogram and Cepstral Coefficients

It is well known that for a Gaussian time domain signal $s(n)$, the spectral coefficients S_k are complex Gaussian distributed and the spectral amplitudes $|S_k|$ are Rayleigh distributed, *i.e.* χ-distributed with two degrees of freedom for $k \in \{1, \ldots, N/2-1, N/2+1, \ldots, N-1\}$, and with one degree of freedom at $k \in \{0, N/2\}$. The χ-distribution is given by

$$p(|S_k|) = \frac{2}{\Gamma(\mu)} \left(\frac{\mu}{\sigma_{s,k}^2}\right)^{\mu} |S_k|^{2\mu-1} \exp\!\left(-\frac{\mu}{\sigma_{s,k}^2}|S_k|^2\right) \,, \tag{4}$$

where 2μ are the degrees of freedom, $\sigma_{s,k}^2 = \mathrm{E}\{|S_k|^2\}$, and $\Gamma(\cdot)$ is the complete gamma function [12, (8.31)]. The distribution of the periodogram $P_k = |S_k|^2$ is then found to be the χ^2-distribution [13],

$$p(P_k) = \frac{1}{\Gamma(\mu)} \left(\frac{\mu}{\sigma_{s,k}^2}\right)^{\mu} P_k^{\mu-1} \exp\!\left(-\frac{\mu}{\sigma_{s,k}^2} P_k\right) \,. \tag{5}$$

If the time domain signal $s(n)$ is not Gaussian distributed, the complex spectral coefficients are asymptotically Gaussian distributed for large N [14]. However, for segment sizes used in common speech processing frameworks, it can be shown that the complex spectral coefficients of speech signals are super-Gaussian distributed [15], [16]. In fact, choosing $\mu < 1$ in (4) may yield a better fit to the distribution of speech spectral amplitudes than a Rayleigh distribution ($\mu = 1$) [17], [18]. Results in this paper are derived for arbitrary values of μ, and thus hold for complex Gaussian distributed spectral coefficients S_k ($\mu = 1$), complex super-Gaussian distributed spectral coefficients $\mu < 1$ and complex sub-Gaussian distributed spectral coefficients $\mu > 1$. In a practical scenario, μ should be chosen so that (4) fits the empirical distribution of the spectral amplitudes of the considered signal. However, we show in this work that μ can also be estimated from the empirical variance of cepstral coefficients (cf. Algorithm 1, step 1).

To compute the variance of the cepstral coefficients we first derive the variance of the log-periodogram,

$$\mathrm{var}\{\log(P_k)\} = \mathrm{E}\Big\{\left(\log(P_k)\right)^2\Big\} - \left(\mathrm{E}\{\log(P_k)\}\right)^2 . \tag{6}$$

With (5) and [12, (4.352.1)], the expected value of the log-periodogram can be derived as

$$\mathrm{E}\{\log P_k\} = \psi(\mu) - \log(\mu) + \log\left(\sigma_{s,k}^2\right) \,, \tag{7}$$

where $\psi(\cdot)$ is the psi-function [12, (8.360)]. The first term on the right hand side of (6) can be derived using [12, (4.358.2)]

$$\mathrm{E}\Big\{\left(\log P_k\right)^2\Big\} = \left(\psi(\mu) - \log(\mu) + \log\left(\sigma_{s,k}^2\right)\right)^2 + \zeta(2,\mu) \,, \tag{8}$$

where $\zeta(\cdot,\cdot)$ is Riemann's zeta-function [12, (9.521.1)]. With (6), (7), and (8) the variance of the log-periodogram κ_0 results in

$$\kappa_0 = \mathrm{var}\{\log P_k\} = \zeta(2,\mu) \,. \tag{9}$$

This is a generalization of the results in [19], where the variance of the log-periodogram was derived for the special case of Gaussian-distributed random variables, *i.e.* $\mu = 1$.

As shown in Appendix A, the covariance of the cepstral coefficients can be obtained by taking a two dimensional discrete Fourier transform of the covariance of the log-periodogram as

$$\mathrm{cov}\{s_{q_1}, s_{q_2}\} = \frac{1}{N^2} \sum_{k_2=0}^{N-1} \sum_{k_1=0}^{N-1} \mathrm{cov}\{\log(P_{k_1}), \log(P_{k_2})\} \, e^{\jmath \frac{2\pi}{N} q_1 k_1} e^{-\jmath \frac{2\pi}{N} q_2 k_2} \,, \tag{10}$$

where $k_1, k_2 \in \{0, \ldots, N-1\}$ are frequency indices, and $q_1, q_2 \in \{0, \cdots, N/2\}$ are quefrency indices. For large N, we may neglect the fact that at $k \in \{0, N/2\}$ the variance $\mathrm{var}\{\log P_k\} = \zeta(2, \frac{\mu}{2})$ is larger than for $k \in \{1, \ldots, N/2-1, N/2+1, \ldots, N-1\}$ where $\mathrm{var}\{\log P_k\} = \zeta(2,\mu) = \kappa_0$. If the frequency bins are uncorrelated, *i.e.* $\mathrm{cov}\{\log P_{k_1}, \log P_{k_2}\} = 0$ for $k_1 \neq k_2$, the covariance of the cepstral coefficients results in

$$\mathrm{cov}\{s_{q_1}, s_{q_2}\} = \begin{cases} \frac{1}{N}\kappa_0 & , q_1 = q_2, q_1 \in \left\{1, ..., \frac{N}{2} - 1\right\} \\ \frac{2}{N}\kappa_0 & , q_1 = q_2, q_1 \in \left\{0, \frac{N}{2}\right\} \\ 0 & , q_1 \neq q_2 \end{cases} \,, \tag{11}$$

with κ_0 defined in (9). Note that a tapered spectral analysis window w_n in (2) results in a correlation of adjacent frequency bins. Since in (11) uncorrelated frequency bins are assumed, this result holds only for rectangular spectral analysis windows. Tapered spectral analysis windows and correlated spectral coefficients are treated in the following.

A. Correlated Spectral Coefficients

While in [19] and (11) only rectangular spectral analysis windows w_n were considered, we now discuss the statistics of the log-periodogram and cepstral coefficients for tapered spectral analysis windows as used in many speech processing algorithms.

While for uncorrelated spectral coefficients we have μ degrees of freedom for $k \in \{0, N/2\}$ and 2μ degrees of freedom for $k \in \{1, \ldots, N/2-1, N/2+1, \ldots, N-1\}$, the correlation introduced by a tapered spectral analysis window results in a reduction of the degrees of freedom, and thus a higher variance for the log-periodogram bins adjacent to $k = 0$ and $k = N/2$. As for large N this hardly affects the cepstral coefficients, the effect is insignificant here. A derivation of the log-spectral variances is given in [20] for the special case $\mu = 1$ and different spectral analysis windows w_n.

However, the correlation of adjacent frequency coefficients

$$\rho_m^2 = \frac{|\mathrm{E}\{S_k S_{k+m}^*\}|^2}{\mathrm{E}\{|S_k|^2\}\mathrm{E}\{|S_{k+m}|^2\}} \tag{12}$$

greatly affects the variance of cepstral coefficients.

The resulting covariance of the logarithm of two periodogram bins

$$\kappa_m = \mathrm{cov}\{\log(P_k), \log(P_{k+m})\}$$

is derived below. As shown in Appendix B, for large N and a given κ_m, the covariance of cepstral coefficients s_{q_1} and s_{q_2} for correlated data results in

$$\mathrm{cov}\{s_{q_1}, s_{q_2}\} = \begin{cases} \frac{1}{N}\left(\kappa_0 + 2\sum_{m=1}^{M}\kappa_m \cos\left(m\frac{2\pi}{N}q_1\right)\right) \\ \qquad\qquad , \text{ for } q_1 = q_2, q_1 \in \left\{1, ..., \frac{N}{2}-1\right\} \\ \frac{2}{N}\left(\kappa_0 + 2\sum_{m=1}^{M}\kappa_m \cos\left(m\frac{2\pi}{N}q_1\right)\right) \\ \qquad\qquad , \text{ for } q_1 = q_2, q_1 \in \left\{0, \frac{N}{2}\right\} \\ 0 \qquad\qquad , \text{ for } q_1 \neq q_2 \end{cases}, \tag{13}$$

where M denotes the number of non-zero covariance values κ_m. It can be seen that cepstral coefficients are asymptotically uncorrelated for large N, even if log-periodogram bins are correlated. The cepstral variance is given as the diagonal of the covariance matrix, as

$$\mathrm{var}\{s_q\} = \begin{cases} \frac{2}{N}\left(\zeta(2,\mu) + 2\sum_{m=1}^{M}\kappa_m \cos\left(m\frac{2\pi}{N}q\right)\right) & , q \in \left\{0, \frac{N}{2}\right\} \\ \frac{1}{N}\left(\zeta(2,\mu) + 2\sum_{m=1}^{M}\kappa_m \cos\left(m\frac{2\pi}{N}q\right)\right) & , \text{else} \end{cases}. \tag{14}$$

To derive the covariance of two log-periodogram bins κ_m, we extend the χ^2-distribution (5) to the bivariate χ^2-distribution of two correlated periodogram bins $P_{k_1} = |S_k|^2$ and $P_{k_2} = |S_{k+m}|^2$ with the correlation ρ_m defined in (12). This distribution can be found *e.g.* in [21, Theorem 2.1], as

$$p(P_{k_1}, P_{k_2}) = \frac{P_{k_1}^{\mu-1} P_{k_2}^{\mu-1}}{2^{2\mu+1}\sqrt{\pi}\,\Gamma(\mu)\,(1-\rho_m^2)^{\mu}}\, e^{-\frac{P_{k_1}+P_{k_2}}{2(1-\rho_m^2)}} \sum_{n=0}^{\infty}(1+(-1)^n)\left(\frac{\rho_m}{1-\rho_m^2}\right)^n \frac{\Gamma\left(\frac{n+1}{2}\right)}{n!\,\Gamma\left(\frac{n}{2}+\mu\right)} P_{k_1}^{\frac{n}{2}} P_{k_2}^{\frac{n}{2}}. \tag{15}$$

Note that the infinite sum in (15) can also be expressed in terms of the hypergeometric function [22]. With $k_1 = k$, $k_2 = k+m$, (15), [12, (4.352.1)] and [12, (3.381.4)] we find

$$\begin{aligned} \kappa_m &= \mathrm{cov}\{\log(P_k), \log(P_{k+m})\} \\ &= \mathrm{E}\{\log(P_k)\log(P_{k+m})\} - \mathrm{E}\{\log(P_k)\}\mathrm{E}\{\log(P_{k+m})\} \\ &= \sum_{n=0}^{\infty} A_{n,\mu,\rho_m} (B_{n,\mu,\rho_m})^2 - \left(\sum_{n=0}^{\infty} A_{n,\mu,\rho_m} B_{n,\mu,\rho_m}\right)^2, \end{aligned} \tag{16}$$

where

$$A_{n,\mu,\rho_m} = \frac{(1-\rho_m^2)^{\mu}}{2\sqrt{\pi}\,\Gamma(\mu)}(1+(-1)^n)\,2^n\rho_m^n\frac{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}+\mu\right)}{n!}, \tag{17}$$

$$B_{n,\mu,\rho_m} = \psi\left(\mu+\frac{n}{2}\right) + \log\left(2\left(1-\rho_m^2\right)\right), \tag{18}$$

and ρ_m as defined in (12). This is a generalization of the results in [23, (6)] and [23, (20)] where the covariance is given for the special cases $\mu = 1$ and $\mu = 1/2$, respectively. With (16), the covariance κ_m of log-periodogram bins, and thus the covariance of cepstral coefficients (13), can be determined.

From above derivations we see that the covariance of cepstral coefficients depends only on the degrees of freedom 2μ of χ-distributed spectral amplitudes and the correlation between spectral coefficients ρ_m. Specifically, the covariance of the cepstral coefficients is independent of the signal power, the spectral shape, and the segment index l. In Appendix C we show that for a Hann window and $\sigma_{s,k-1}^2 \approx \sigma_{s,k}^2 \approx \sigma_{s,k+1}^2$, the normalized correlation results in $\rho_1 = 2/3$ and $\rho_2 = 1/6$. Hence, for a Hann window and $\mu = 1$ we have $\kappa_1 = 0.507$ and $\kappa_2 = 0.028$.

The cepstral variance for $\mu = 1$ and the rectangular window ($\kappa_m = 0, m \in \{1, ..., M\}$) or the Hann window ($\kappa_1 = 0.507, \kappa_2 = 0.028, \kappa_m = 0, m \in \{3, ..., M\}$) are compared in Fig. 1, where we also show empirical data. It is obvious that (14) provides an excellent fit for both the rectangular and Hann window. As the additional cosine-terms in (13) and (14) have zero mean with respect to q, the quefrency average of the cepstral variance $\overline{\mathrm{var}\{s_q\}}$ for arbitrary spectral correlation equals the cepstral variance for a rectangular window and is thus independent of the chosen analysis window w_n:

$$\overline{\mathrm{var}\{s_q\}} = \frac{1}{N}\sum_{q=0}^{N/2}\nu_q \mathrm{var}\{s_q\} = \zeta(2,\mu)/N, \tag{19}$$

with

$$\nu_q = \begin{cases} 1/2 & , q \in \{0, N/2\} \\ 2 & , \text{else} \end{cases}. \tag{20}$$

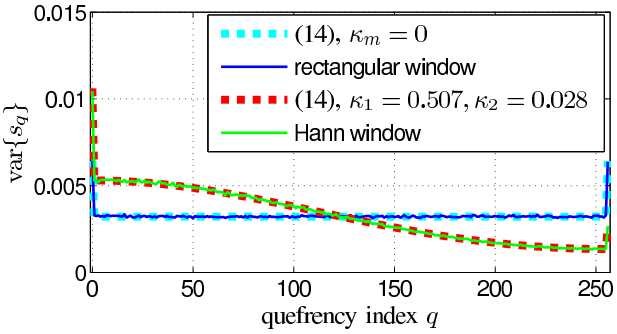


Fig. 1. The cepstral variance for a pink Gaussian time-domain signal analyzed with a non-overlapping rectangular analysis window w_n in (2) and a Hann window with half-overlapping frames. The empirical variances are compared to the theoretical results in (14) with $\kappa_m = 0, m \in \{1, ..., M\}$ for the rectangular window and $\kappa_1 = 0.507, \kappa_2 = 0.028, \kappa_m = 0, m \in \{3, ..., M\}$ for the Hann window. The sampling rate is $f_s = 16\,\text{kHz}$ and $N = 512$.

The coefficients ν_q account for the symmetry of the cepstrum and the different variances at the DC and Nyquist bin in (14). In this way the cosine terms in (14) cancel out and the modified average variance of the cepstral coefficients $\overline{\text{var}\{s_q\}}$ and the degrees of freedom 2μ are directly related.

IV. Statistical Properties after Cepstral Variance Reduction

In this section, we approximate the distribution of spectral amplitudes after CVR by the parametric χ-distribution. From experimental results in Section VI it can be seen that this approximation is valid. From (19) we see that CVR increases the parameter μ of the χ-distribution. Then, due to (7), changing μ also changes the spectral power $\sigma^2_{s,k}$. Hence, a variance reduction in the cepstral domain results in a bias in the spectral power that can now be accounted for. In the following, we denote parameters after CVR by a tilde. We will discuss CN and TCS separately.

In CVR by CN, a set of cepstral coefficients is set to zero [8], [9]. Then, the average variance after CVR can be determined as

$$\overline{\text{var}\{\widetilde{s_q}\}} = \frac{1}{N}\sum_{q=0}^{N/2} \nu_q \text{var}\{s_q\}\, b_q\,, \tag{21}$$

where $\widetilde{s_q}$ are the cepstral coefficients s_q after CVR, the indicator function $b_q \in \{0, 1\}$ sets certain cepstral coefficients to zero, and ν_q is defined as in (20).

For TCS, the cepstral coefficients are recursively smoothed over time with a quefrency dependent smoothing factor α_q

$$\widetilde{s_q}(l) = \alpha_q\, \widetilde{s_q}(l-1) + (1-\alpha_q)\, s_q(l)\,. \tag{22}$$

For uncorrelated successive signal segments, the average cepstral variance can be determined by

$$\overline{\text{var}\{\widetilde{s_q}\}} = \frac{1}{N}\sum_{q=0}^{N/2} \nu_q \text{var}\{s_q\}\, \frac{1-\alpha_q}{1+\alpha_q}\,, \tag{23}$$

which is also valid for Hann spectral analysis windows with 50% overlap. For higher signal segment correlation, the average variance after CVR, $\overline{\text{var}\{\widetilde{s_q}\}}$, can be measured offline for a fixed set of recursive smoothing constants α_q. For a given μ of the spectral amplitudes before CVR, the cepstral variance can be determined via (14) and thus the average cepstral variance after CVR, $\overline{\text{var}\{\widetilde{s_q}\}}$, via (21) or (23).

With a known average cepstral variance, the parameter $\widetilde{\mu}$ can be implicitly determined using

$$\zeta(2, \widetilde{\mu}) = N\overline{\text{var}\{\widetilde{s_q}\}}\,, \tag{24}$$

where $2\widetilde{\mu}$ are the degrees of freedom after CVR. In a practical application, the relation between $\widetilde{\mu}$ and $\zeta(2, \widetilde{\mu})$ can be stored in a table.

The spectral power bias $\sigma^2_{s,k}/\widetilde{\sigma^2_{s,k}}$ can then be determined using (7), as

$$\begin{aligned}\log\left(\sigma^2_{s,k}/\widetilde{\sigma^2_{s,k}}\right) = {} & \text{E}\{\log\left(|S_k|^2\right)\} - \psi\left(\mu\right) + \log\left(\mu\right) \\ & - \left(\text{E}\left\{\log\left(|\widetilde{S_k}|^2\right)\right\} - \psi\left(\widetilde{\mu}\right) + \log\left(\widetilde{\mu}\right)\right)\,.\end{aligned} \tag{25}$$

It can be seen from (25) that, if the zeroth cepstral coefficient is not changed, CVR changes the signal power for two reasons: first, the instantaneous spectral shape is changed, and secondly, the degrees of freedom are changed. While a signal power reduction due to the former is a desired property of CVR, a signal power reduction due to the latter is not desired and shall be compensated. The former is desired, because changing the spectral shape is actually the aim of CVR, *e.g.* eliminating babble burst in speech enhancement [5], [6]. A compensation of only the latter can be achieved by setting $\text{E}\{\log\left(|S_k|^2\right)\} = \text{E}\left\{\log\left(|\widetilde{S_k}|^2\right)\right\}$ in (25). We thus obtain the frequency independent factor

$$r^2 = \sigma^2_{s,k}/\widetilde{\sigma^2_{s,k}} = \frac{\mu}{\widetilde{\mu}}\, e^{\psi(\widetilde{\mu}) - \psi(\mu)} \tag{26}$$

that is applied to all spectral bins as

$$|\widehat{S_k}| = r\,|\widetilde{S_k}|\,. \tag{27}$$

Therefore, we obtain cepstrally-smoothed spectral amplitudes $|\widehat{S_k}|$ with reduced cepstral variance which are approximately χ-distributed (4) with $2\widetilde{\mu}$ degrees of freedom and have the correct signal power. The algorithm for computing unbiased signal power estimates after CVR is summarized in Algorithm 1. Note that the bias correction r depends only on μ and $\widetilde{\mu}$. For a fixed set of smoothing parameters α_q or b_q the bias correction r is thus fixed and independent of the segment index l.

V. Mean of the Cepstrum

In this section we derive the mean of the cepstral coefficients. We generalize the results of [19] and [8], [9], where $\mu = 1$ is assumed. Due to the linearity of the inverse discrete Fourier transform IDFT$\{\cdot\}$ and (7), the mean value of the cepstral coefficients defined by (3) is given by

$$\begin{aligned}\text{E}\{s_q\} = {} & \text{IDFT}\{\text{E}\{\log P_k\}\} \\ = {} & \text{IDFT}\{\log \sigma^2_{s,k}\} - \text{IDFT}\{\log \mu_k - \psi(\mu_k)\} \\ = {} & \text{IDFT}\{\log \sigma^2_{s,k}\} - \varepsilon_q\,.\end{aligned} \tag{28}$$

Therefore, even for white signals, when $\sigma_{s,k}^2$ is constant over frequency, the mean of the cepstral coefficients is not zero for $q > 0$ but $-\varepsilon_q$. When

$$\mu_k = \begin{cases} \mu/2 & , k \in \{0, N/2\} \\ \mu & , \text{else} \end{cases}$$

the deviation ε_q results in

$$\varepsilon_q = \mathrm{IDFT}\{\log \mu_k - \psi(\mu_k)\} = \begin{cases} \frac{N-2}{N}\left(\log\mu - \psi\left(\mu\right)\right) + \frac{2}{N}\left(\log\frac{\mu}{2} - \psi\left(\frac{\mu}{2}\right)\right) & , q = 0 \\ \frac{2}{N}\left(\log\frac{\mu}{2} - \psi\left(\frac{\mu}{2}\right)\right) - \frac{2}{N}\left(\log\mu - \psi\left(\mu\right)\right) & , q \text{ odd} \\ 0 & , q \text{ even} \end{cases} . \tag{29}$$

If $\mu_k = \mu$ is constant for all k, as assumed in [8], [9], the deviation results in

$$\varepsilon_q = \begin{cases} \log(\mu) - \psi(\mu) & , q = 0 \\ 0 & , \text{else} \end{cases} .$$

Because in the CVR method proposed in [8], [9] certain cepstral coefficients are set to zero, better performance is achieved when the cepstrum actually has zero mean for white signals. Such an alternative definition of the cepstrum is given by $\widehat{s_q} = s_q + \varepsilon_q$. However, since typically $\varepsilon_q^2 \ll \mathrm{var}\{s_q\}$ for $q > 0$, the influence of the mean bias ε_q given in (29) is of minor importance. For a temporal cepstrum smoothing, as proposed in [5], [6], zero-mean cepstral coefficients are neither assumed nor required.

VI. Experimental Results

In this section we show that Algorithm 1 successfully compensates for the signal power bias introduced by CVR. After providing results for a stationary colored signal, we also apply the bias compensation in a practical scenario, namely the *a priori* speech power estimation proposed in [6].

A. Stationary Colored Signal

Here we apply CVR to a stationary colored Gaussian distributed signal. In Fig. 2 and Fig. 3 we present the results for TCS using a rectangular and a Hann spectral analysis window in (2), respectively. In Fig. 4 and Fig. 5 we present the results for CN using a rectangular and a Hann spectral analysis window in (2), respectively. From the presented results, we see that CVR introduces a signal power bias, and that this bias is successfully compensated with Algorithm 1. Further, we compare the histograms of spectral amplitudes before and after CVR with and without a bias compensation to the derived probability density functions. It may be seen that the algorithm for estimating the degrees of freedom after CVR works well, as an excellent match for the histograms and the derived probability density functions may be observed.

Algorithm 1 Bias compensation for temporal cepstrum smoothing (TCS) and cepstral nulling (CN)

1: If unknown, determine the degrees of freedom 2μ using an empirical estimation of $\mathrm{var}\{s_q\}$ from representative data and (19):

$$\zeta\left(2, \mu\right) = N\overline{\mathrm{var}\{s_q\}} = \sum_{q=0}^{N/2} \nu_q \mathrm{var}\{s_q\} \ ,$$

with ν_q defined in (20).

2: Determine the correlation of neighboring $\log$-periodogram bins κ_m via (16):

$$\kappa_m = \sum_{n=0}^{\infty} A_{n,\mu,\rho_m} \left(B_{n,\mu,\rho_m}\right)^2 - \left(\sum_{n=0}^{\infty} A_{n,\mu,\rho_m} B_{n,\mu,\rho_m}\right)^2$$

with A, B, ρ_m defined in (17), (18), and (12).

3: Determine the cepstral variance before CVR (14):

$$\mathrm{var}\{s_q\} = \begin{cases} \frac{2}{N}\left(\zeta(2,\mu) + 2\sum_{m=1}^{M} \kappa_m \cos\left(m\frac{2\pi}{N}q\right)\right) & , q \in \left\{0, \frac{N}{2}\right\} \\ \frac{1}{N}\left(\zeta(2,\mu) + 2\sum_{m=1}^{M} \kappa_m \cos\left(m\frac{2\pi}{N}q\right)\right) & , \text{else.} \end{cases}$$

4: **for all** signal segments l **do**

5: **if** smoothing parameters b_q or α_q have changed **then**

6: Determine the average cepstral variance after CVR,

- in the case of CN (21):

$$\overline{\mathrm{var}\{\widetilde{s_q}\}} = \frac{1}{N}\sum_{q=0}^{N/2} \nu_q \mathrm{var}\{s_q\}\, b_q,$$

- in the case of TCS (23):

$$\overline{\mathrm{var}\{\widetilde{s_q}\}} = \frac{1}{N}\sum_{q=0}^{N/2} \nu_q \mathrm{var}\{s_q\}\, \frac{1-\alpha_q}{1+\alpha_q}.$$

7: Determine the degrees of freedom $2\widetilde{\mu}$ after CVR (24):

$$\zeta(2, \widetilde{\mu}) = N\overline{\mathrm{var}\{\widetilde{s_q}\}}\,.$$

8: Compute signal power bias (26):

$$r^2 = \sigma_{s,k}^2 / \widetilde{\sigma_{s,k}^2} = \frac{\mu}{\widetilde{\mu}}\, e^{\psi(\widetilde{\mu}) - \psi(\mu)}\,.$$

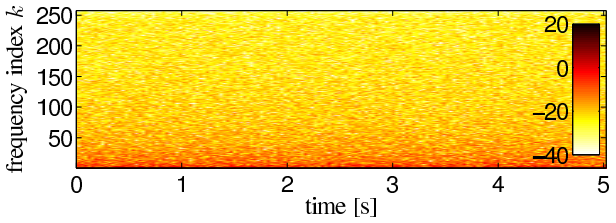
9: **end if**

10: Apply bias correction (27):

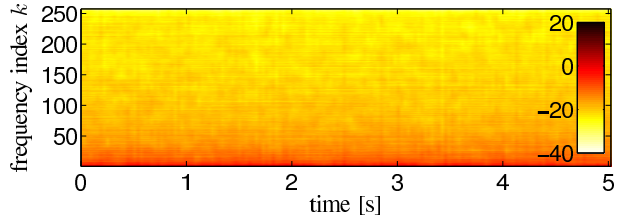
$$|\widehat{S_k}(l)| = r\,|\widetilde{S_k}(l)|\,.$$

11: **end for**

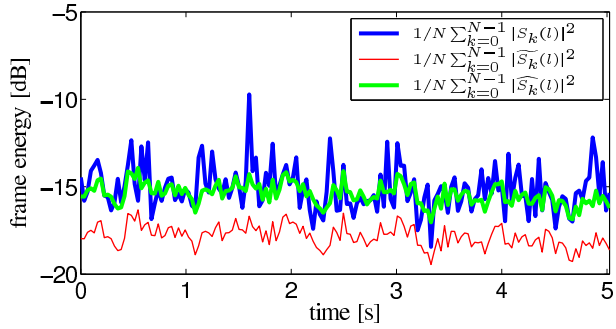
In a practical application, the relation between $\widetilde{\mu}$ and $\zeta(2, \widetilde{\mu}) = N\overline{\mathrm{var}\{\widetilde{s_q}\}}$ can be stored in a table.



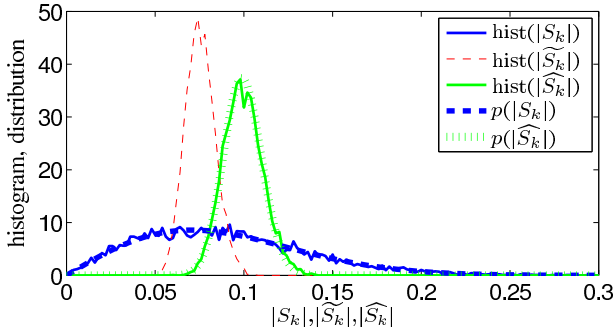
(a)



(b)

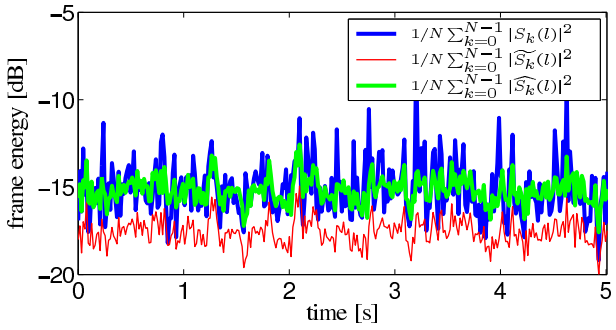


(c)

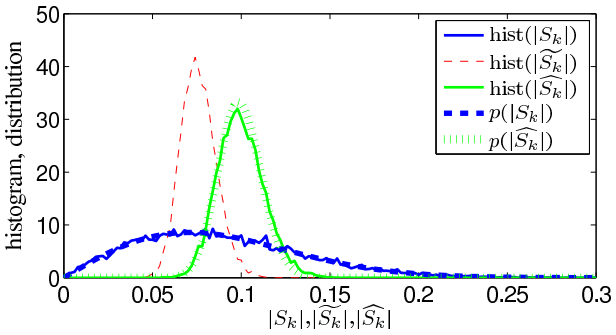


(d)

Fig. 2. CVR by TCS for a stationary pink Gaussian-distributed signal and non-overlapping rectangular spectral analysis windows w_n in (2). We use the same smoothing constants as in [6], but do not smooth the zeroth cepstral coefficient. We first present the spectrogram of the signal before CVR (a), then the spectrogram after CVR and bias correction (b). In (c) the signal segment energies before CVR, after CVR, and after CVR and bias correction are given. (d) compares the derived distributions to the histograms of the spectral amplitudes $|S_k|$ for $k = 111$ before CVR, $|\widetilde{S_k}|$ after CVR, and $|\widehat{S_k}|$ after CVR and bias correction. Here $N = 512$ and the sampling rate is $f_s = 16\,\text{kHz}$.



(a)



(b)

Fig. 3. CVR by TCS as in Fig. 2 but with half-overlapping Hann windows w_n in (2). In (a) the signal segment energies before CVR, after CVR, and after CVR and bias correction are given. (b) compares the derived distributions to the histograms of the corresponding spectral amplitudes.

B. A Priori Speech Power Estimation

Now, the bias compensation method is applied in a practical scenario, namely the TCS based *a priori* clean speech power estimation algorithm for speech enhancement proposed in [6]. There, a maximum likelihood estimation of the *a priori* clean speech power $|S_k(l)|^2$ is temporally smoothed in the cepstral domain to obtain the smoothed *a priori* speech power estimation $|\widetilde{S_k}(l)|^2$. The benefit of a smoothing in the cepstral domain is that in the cepstral domain, speech is very compactly represented, namely by few lower cepstral coefficients representing the speech spectral envelope and a maximum in the upper cepstrum representing the speech fundamental period. On the other hand, estimation errors that lead to spectral outliers change the fine structure of the spectrum, which is represented by the upper cepstral coefficients. Thus, to reduce spectral outliers without affecting the speech signal, only little smoothing is applied to the speech related cepstral coefficients and strong smoothing to the remaining cepstral coefficients. However, due to the CVR, the smoothed *a priori* speech power estimation $|\widetilde{S_k}(l)|^2$ will be biased as compared to $|S_k(l)|^2$. In Fig. 6 it may be seen that this bias can be successfully compensated with the proposed Algorithm 1, yielding the smoothed unbiased speech power estimate $|\widehat{S_k}(l)|^2$.

For the simulation we used nonoverlapping rectangular spectral analysis windows. We estimate the degrees of freedom before CVR, 2μ, by measuring the average cepstral variance $\overline{\text{var}\{s_q\}}$ and using the relation $\zeta(2,\mu) = N\overline{\text{var}\{s_q\}}$. We thus obtain $\mu = 0.37$. We use the same smoothing procedure as proposed in [6] but do not smooth the zeroth cepstral

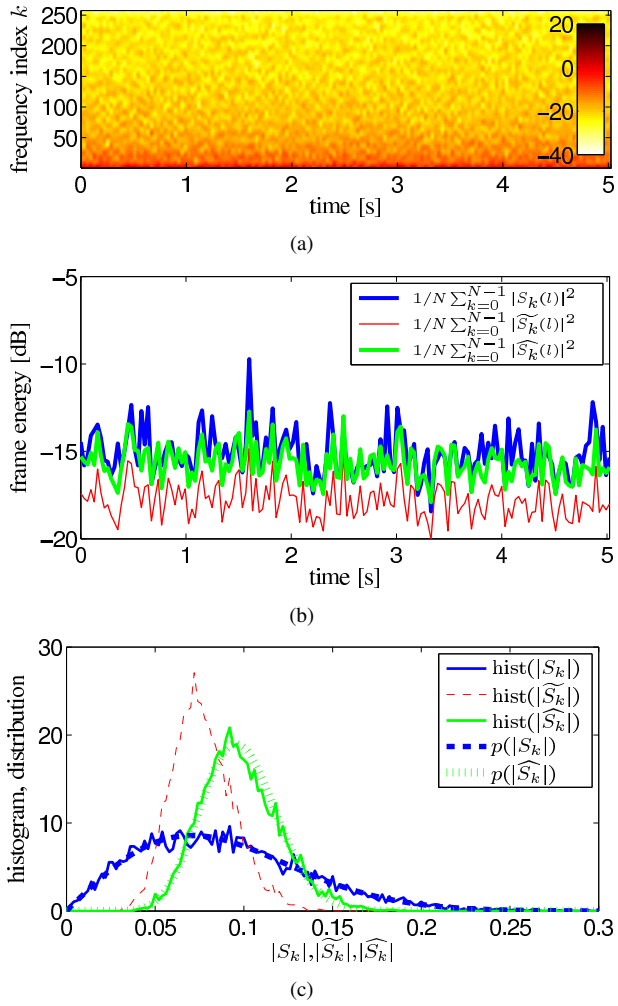


Fig. 4. CVR by CN for the signal in Fig. 2(a) using non-overlapping rectangular spectral analysis windows w_n in (2). Cepstral coefficients $q > N/16$ are set to zero. We first present the spectrogram after CN and bias correction (a). In (b) the signal segment energies before CVR, after CVR, and after CVR and bias correction are given. (c) compares the derived distributions to the histograms of the corresponding spectral amplitudes.

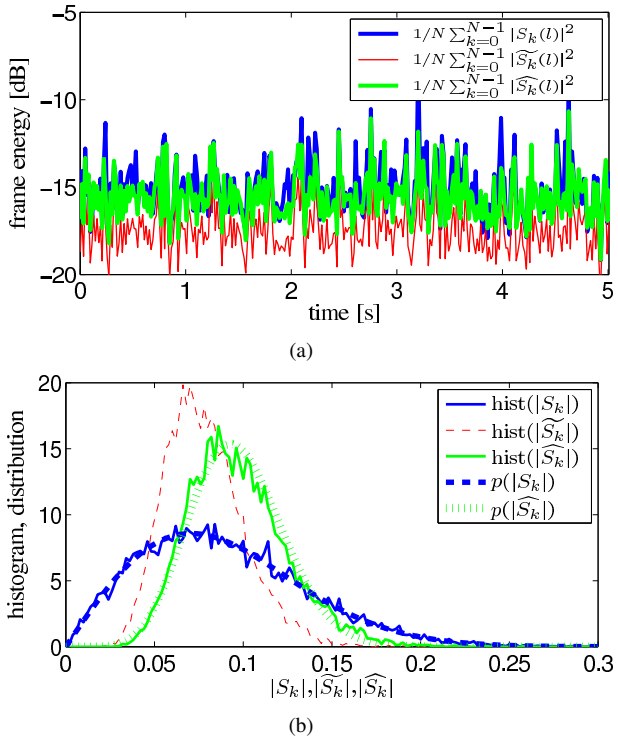


Fig. 5. CVR by CN as in Fig. 4 but with half-overlapping Hann windows w_n in (2). In (a) the signal segment energies before CVR, after CVR, and after CVR and bias correction are given. (b) compares the derived distributions to the histograms of the corresponding spectral amplitudes.

coefficient. As in [6] the smoothing constant α_q in (22) varies from signal segment to signal segment, a different bias r is introduced in each segment l. Note that the computational simplicity of Algorithm 1 allows for an individual computation of the signal power bias r in each signal segment l (steps 6-8 of Algorithm 1).

VII. Conclusion

We analyzed the effect of a cepstral variance reduction on the statistical properties of spectral amplitudes. We have shown that after a cepstral variance reduction the distribution of spectral amplitudes can be approximated by a χ-distribution with an increased number of degrees of freedom and decreased signal power. As a change in signal power is generally not desired, we propose a signal power bias correction based on the statistical properties of cepstral and spectral coefficients. The proposed bias correction results in a simple scaling of the spectral amplitudes and is fixed for a fixed set of cepstral variance reduction parameters. However, as the determination of the bias correction factor is computationally inexpensive, it can be computed on a segment-by-segment basis if the smoothing parameters change.

Appendix A
Derivation of (10)

In this appendix, we show that the covariance of the cepstral coefficients can be obtained by taking a two dimensional discrete Fourier transform of the covariance of the log-periodogram. With the definition of the cepstrum (3) we obtain

$$\begin{aligned}\operatorname{cov}\{s_{q_1}, s_{q_2}\} &= \mathrm{E}\{(s_{q_1} - \mathrm{E}\{s_{q_1}\})\,(s_{q_2} - \mathrm{E}\{s_{q_2}\})^*\} \\ &= \mathrm{E}\Bigg\{\frac{1}{N}\sum_{k_1=0}^{N-1}(\log P_{k_1} - \mathrm{E}\{\log P_{k_1}\})\,e^{\jmath\frac{2\pi}{N}k_1 q_1} \\ &\quad \cdot \frac{1}{N}\sum_{k_2=0}^{N-1}(\log P_{k_2} - \mathrm{E}\{\log P_{k_2}\})\,e^{-\jmath\frac{2\pi}{N}k_2 q_2}\Bigg\} \\ &= \frac{1}{N^2}\sum_{k_2=0}^{N-1}\sum_{k_1=0}^{N-1}\operatorname{cov}\{\log P_{k_1}, \log P_{k_2}\}\,e^{\jmath\frac{2\pi}{N}q_1 k_1}e^{-\jmath\frac{2\pi}{N}q_2 k_2}\,.\end{aligned} \tag{30}$$

Appendix B
Cepstral Covariance for Correlated Spectral Coefficients

In this appendix, we derive an explicit expression for the covariance of cepstral coefficients, when the log-periodogram bins are correlated. The derived results hold for large N as

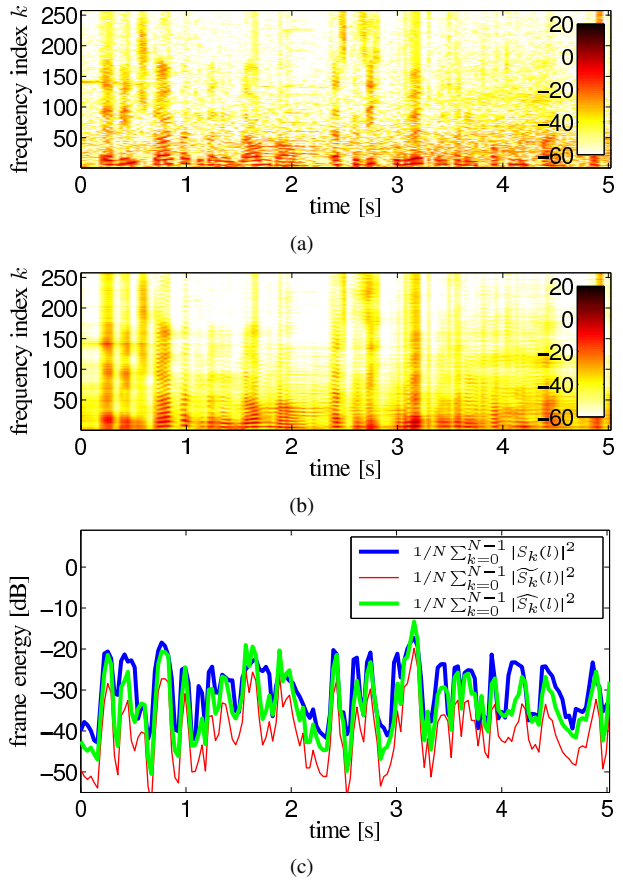


Fig. 6. Spectrogram of the *a priori* speech power estimation $|S_k(l)|^2$ before CVR (a) and $|\widetilde{S_k}(l)|^2$ after CVR and bias compensation (b). In (c) the signal segment energies before CVR, after CVR, and after CVR and bias compensation are given. The speech signal is disturbed by instationary traffic noise at a signal-to-noise ratio of 0 dB. The noise power estimation is done using the minimum statistics approach [1]. Here $N = 512$ and the sampling rate is $f_s = 16$ kHz.

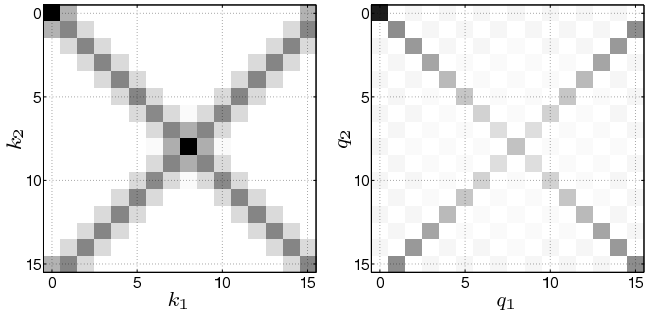


Fig. 7. The covariance matrix of the log-periodogram $\mathrm{cov}\{\log(P_{k_1}), \log(P_{k_2})\}$ (a) and the cepstral coefficients $\mathrm{cov}\{s_{q_1}, s_{q_2}\}$ (b). The periodogram bins are obtained from a computer generated white Gaussian time domain signal $s(n)$, a Hann window with 50% overlap, and $N = 16$.

usually used in speech enhancement applications. For large N, the covariance matrix of the log-periodogram can be approximated by a $N \times N$ symmetric Toeplitz matrix defined by the vector $[\kappa_0, \kappa_1, ..., \kappa_{N/2}, \kappa_{N/2+1}, \kappa_{N/2}, \kappa_{N/2-1}, ..., \kappa_1]$, where we neglect the fact that for $k \in \{0, N/2\}$ the variance of the log-periodogram is larger than κ_0, as we have less degrees of freedom than for $k \notin \{0, N/2\}$. The covariance of the cepstral coefficients is obtained by taking a two dimensional discrete Fourier transform, as presented in Appendix A. As in general the spectral covariance κ_m introduced by tapered spectral analysis windows rapidly decreases with increasing m, we assume that $\kappa_m = 0$ for $m > M$ and $M \ll N/2 + 1$. Then, the covariance of cepstral coefficients results in

$$
\begin{aligned}
&\mathrm{cov}\{s_{q_1}, s_{q_2}\} \\
&= \frac{1}{N^2} \sum_{k_2=0}^{N-1} \Bigg(2\kappa_0 \cos\left(\frac{2\pi}{N} q_1 k_2\right) \\
&\quad + \sum_{m=1}^{M} 2\kappa_m \Bigg(\cos\left(\frac{2\pi}{N} q_1 (k_2 - m)\right) \\
&\qquad + \cos\left(\frac{2\pi}{N} q_1 (k_2 + m)\right) \Bigg) \Bigg) e^{-\jmath \frac{2\pi}{N} q_2 k_2} \\
&= \frac{1}{N^2} \Bigg(\kappa_0 \left(\sum_{k_2=0}^{N-1} e^{\jmath \frac{2\pi}{N}(q_1 - q_2)k_2} + \sum_{k_2=0}^{N-1} e^{-\jmath \frac{2\pi}{N}(q_1 + q_2)k_2} \right) \\
&\quad + \sum_{m=1}^{M} \kappa_m \Bigg(e^{-\jmath \frac{2\pi}{N} q_1 m} \sum_{k_2=0}^{N-1} e^{\jmath \frac{2\pi}{N}(q_1 - q_2)k_2} \\
&\qquad + e^{\jmath \frac{2\pi}{N} q_1 m} \sum_{k_2=0}^{N-1} e^{-\jmath \frac{2\pi}{N}(q_1 + q_2)k_2} \\
&\qquad + e^{\jmath \frac{2\pi}{N} q_1 m} \sum_{k_2=0}^{N-1} e^{\jmath \frac{2\pi}{N}(q_1 - q_2)k_2} \\
&\qquad + e^{-\jmath \frac{2\pi}{N} q_1 m} \sum_{k_2=0}^{N-1} e^{-\jmath \frac{2\pi}{N}(q_1 + q_2)k_2} \Bigg) \Bigg) \\
&= \begin{cases} \frac{1}{N} \left(\kappa_0 + 2 \sum_{m=1}^{M} \kappa_m \cos\left(\frac{2\pi}{N} q_1 m\right) \right) \\ \qquad , \text{ for } (q_1 = q_2 \neq 0, N/2) \text{ OR } (q_1 + q_2 = N) \\ \frac{2}{N} \left(\kappa_0 + 2 \sum_{m=1}^{M} \kappa_m \cos\left(\frac{2\pi}{N} q_1 m\right) \right) \\ \qquad , \text{ for } q_1 = q_2 = 0, N/2 \\ 0 \qquad , \text{ else} \end{cases}
\end{aligned}
\tag{31}
$$

Note that (31) is the result for the full symmetric cepstrum $q \in \{0, ..., N-1\}$, while in (13) the solution for the lower symmetric part $q \in \{0, ..., N/2\}$ is given.

In Fig. 7 the covariance matrices of the log-periodogram and the cepstral coefficients are illustrated. There, the periodogram bins are obtained from a computer generated white Gaussian time domain signal. The spectral analysis (2) is obtained with a Hann spectral analysis window w_n and $N = 16$. As N is relatively small, a slight correlation may be observed in Fig. 7(b) when q_1 and q_2 are even, or q_1 and q_2 are odd. These correlations arise from the fact, that for $k \in \{0, N/2\}$ the variance of the log-periodogram is larger than κ_0, which we neglected in the derivation of (31). However, the resulting correlations drop with $1/N^2$ [19]. For segment sizes N, as usually used in speech processing, these correlations are insignificant, *i.e.* the cepstral coefficients are asymptotically uncorrelated for large N.

APPENDIX C
SPECTRAL CORRELATION FOR A HANN WINDOW

In this appendix we derive the correlations ρ_1 and ρ_2 for a Hann window, with ρ_m defined in (12).

The multiplication of the time domain signal with the window function w_n in (2) results in a convolution of the uncorrelated spectral coefficients U_k with the Fourier domain representation of the window function $W_k = \mathrm{DFT}\{w_n\}$, *i.e.* $S_k = U_k * W_k$, where the asterisk denotes convolution and $\mathrm{DFT}\{\cdot\}$ the discrete Fourier transform. For a normalized discrete Hann window we obtain the correlated frequency coefficients

$$S_k = -\sqrt{\frac{1}{6}}U_{k-1} + \sqrt{\frac{2}{3}}U_k - \sqrt{\frac{1}{6}}U_{k+1}\,. \tag{32}$$

Because U_k is spectrally uncorrelated, with $\sigma^2_{k-1} \approx \sigma^2_k \approx \sigma^2_{k+1}$ and $\mathrm{E}\{|U_k|^2\} = \sigma^2_k$ we have

$$\mathrm{E}\{|S_k|^2\} = \sigma^2_k\,. \tag{33}$$

For the covariances we obtain with (32)

$$\mathrm{E}\{S_k S^*_{k+1}\} = -\frac{1}{3}\sigma^2_k - \frac{1}{3}\sigma^2_{k+1}\,, \tag{34}$$

$$\mathrm{E}\{S_k S^*_{k+2}\} = -\frac{1}{6}\sigma^2_{k+1}\,. \tag{35}$$

Thus, with $\sigma^2_{k-1} \approx \sigma^2_k \approx \sigma^2_{k+1}$ and (12) we have $\rho_1 = 2/3$ and $\rho_2 = 1/6$.

REFERENCES

[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, 2001.

[2] T. Gerkmann, C. Breithaupt, and R. Martin, "Improved a posteriori speech presence probability estimation based on a likelihood ratio with fixed priors," *IEEE Trans. on Audio, Speech, and Language Proc.*, vol. 16, no. 5, pp. 910–919, July 2008.

[3] R. Martin and T. Lotter, "Optimal recursive smoothing of non-stationary periodograms," *Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, pp. 167–170, Sept. 2001.

[4] P. Vary and R. Martin, *Digital Speech Transmission: Enhancement, Coding And Error Concealment*. John Wiley & Sons, 2006.

[5] C. Breithaupt, T. Gerkmann, and R. Martin, "Cepstral smoothing of spectral filter gains for speech enhancement without musical noise," *IEEE Signal Proc. Letters*, vol. 14, no. 12, pp. 1036–1039, Dec. 2007.

[6] ——, "A novel a priori SNR estimation approach based on selective cepstro-temporal smoothing," *IEEE ICASSP*, pp. 4897–4900, Apr. 2008.

[7] A. M. Noll, "Cepstrum pitch estimation," *J. Acoust. Soc. Am.*, vol. 41, pp. 293–309, Feb. 1967.

[8] P. Stoica and N. Sandgren, "Smoothed nonparametric spectral estimation via cepstrum thresholding," *IEEE Signal Proc. Magazine*, vol. 23, no. 6, pp. 34–45, Nov. 2006.

[9] ——, "Total-variance reduction via thresholding: Application to cepstral analysis," *IEEE Trans. on Signal Proc.*, vol. 55, no. 1, pp. 66–72, Jan. 2007.

[10] D. Mauler, T. Gerkmann, and R. Martin, "An analysis of quefrency selective temporal smoothing of the cepstrum in speech enhancement," *Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Sept. 2008.

[11] B. P. Bogert, M. J. R. Healy, and J. W. Tukey, *The Quefrency Alanysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking*. Wiley, NY, 1963, ch. 15, pp. 209–243.

[12] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals Series and Products*, 6th ed., A. Jeffrey and D. Zwillinger, Ed. Academic Press, 2000.

[13] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1991.

[14] D. Brillinger, *Time Series: Data Analysis and Theory*. San Francisco, CA: Holden-Day, 1981.

[15] R. Martin, "Speech enhancement using MMSE short time spectral estimation with gamma distributed speech priors," *IEEE ICASSP*, vol. I, pp. 253–256, May 2002.

[16] ——, "Speech enhancement based on minimum mean-square error estimation and supergaussian priors," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 845–856, Sept. 2005.

[17] I. Andrianakis and P. R. White, "MMSE speech spectral amplitude estimators with chi and gamma speech priors," *IEEE ICASSP*, vol. III, pp. 1068–1071, 2006.

[18] C. Breithaupt, M. Krawczyk, and R. Martin, "Parameterized MMSE spectral magnitude estimation for the enhancement of noisy speech," *IEEE ICASSP*, pp. 4037–4040, Apr. 2008.

[19] Y. Ephraim and M. Rahim, "On second-order statistics and linear estimation of cepstral coefficients," *IEEE Trans. on Speech and Audio Proc.*, vol. 7, no. 2, pp. 162–176, Mar. 1999.

[20] A. H. Gray Jr., "Log spectra of Gaussian signals," *J. Acoust. Soc. Am.*, vol. 55, no. 5, pp. 1028–1033, May 1974.

[21] A. H. Joarder, "Moments of the product and ratio of two correlated chi-square variables," *SpringerLink Statistical Papers*, Nov. 2007.

[22] S. Nadarajah, "Comment on the paper by A. H. Joarder," *SpringerLink Statistical Papers*, Dec. 2007.

[23] Y. Ephraim and J. J. Roberts, "On second-order statistics of log-periodogram with correlated components," *IEEE Signal Proc. Letters*, vol. 12, no. 9, pp. 625–628, Sept. 2005.

Timo Gerkmann studied electrical engineering at the Universität Bremen, Bremen, Germany, and the Ruhr-Universität Bochum, Bochum, Germany. He received the Dipl.-Ing. degree from the Ruhr-Universität Bochum in 2004. He is currently pursuing the Dr.-Ing. degree at the Institute of Communication Acoustics, Ruhr-Universität Bochum.

From January 2005 to July 2005 he visited Siemens Corporate Research in Princeton, NJ, where he worked on artificial bandwidth extension. His main research interests are digital speech and audio processing, including single- and multichannel speech enhancement.

Rainer Martin (S'86-M'90-SM'01) received the Dipl.-Ing. and Dr.-Ing. degrees from RWTH Aachen University, Aachen, Germany, in 1988 and 1996, respectively, and the M.S.E.E. degree from Georgia Institute of Technology, Atlanta, GA, in 1989.

From 1996 to 2002, he was a Senior Research Engineer with the Institute of Communication Systems and Data Processing, RWTH Aachen University. From April 1998 to March 1999, he was on leave to the AT&T Speech and Image Processing Services Research Lab, Florham Park, N.J. From April 2002 until October 2003 he was a Professor of Digital Signal Processing at the Technische Universität Braunschweig, Braunschweig, Germany. Since October 2003 he has been a Professor of Information Technology and Communication Acoustics at Ruhr-Universität Bochum, Bochum, Germany, and since October 2007 Dean of the Electrical Engineering and Information Technology department. His research interests are signal processing for voice communication systems, hearing instruments, and human-machine interfaces. He is coauthor with P. Vary of *Digital Speech Transmission - Enhancement, Coding and Error Concealment* (Wiley, 2006) and coeditor with U. Heute and C. Antweiler of *Advances in Digital Speech Transmission* (Wiley, 2008). Dr. Martin served as an Associate Editor for the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING and is a member of the Audio and Electroacoustics Technical Committee of the IEEE Signal Processing Society.