Complex Noise Analysis of DMT

Georg Tauböck, *Member, IEEE*

Abstract—In this paper, we consider discrete multitone (DMT) or baseband orthogonal frequency-division multiplexing (OFDM) modulation and perform a detailed noise analysis which takes into account dependencies and power (variance) differences of real and imaginary part after the complex-valued discrete Fourier transform (DFT). The derivation is based on the so-called *pseudocovariance matrix* of a complex random vector, which was introduced by Neeser and Massey (1993). We show that the relevant pseudocovariance matrix is *not* the zero matrix in general, in contrast to *pass-band* OFDM, for which it can be proven (under certain assumptions) that all occurring pseudocovariance matrices are vanishing. We show that for *colored* noise *rotated rectangular* symbol constellations are more appropriate than the common quadratic quadrature amplitude modulation (QAM) symbol constellations with respect to capacity and symbol error probability, and we derive formulas for the rotation angles and constellation sizes/densities. Finally, we extend the results to a multitransceiver [multiple-input–multiple-output (MIMO)] scenario, for which we assume a very general noise model at the inputs of the receivers, allowing *correlations* between the noise signals of different receivers. This requires the introduction of *pseudocross-covariance matrices* of complex random vectors, which are the important objects (together with cross-covariance matrices) in the MIMO situation.

Index Terms—Complex random vector, discrete multitone (DMT), multiple-input–multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), proper, pseudo-covariance matrix, rotationally (in)variant.

I. INTRODUCTION

IN this paper, we consider discrete multitone (DMT) modulation as it is used for asymmetric digital subscriber line (ADSL) and very high-speed digital subscriber line (VDSL) transmission, cf., for example, [8]–[13], and as it is described in, e.g., [5] and [29]. We mainly focus on the receiver part, especially on the noise characteristics disturbing the transmission.

First, observe (cf., also, Fig. 1) that all operations in the receiver before the input of the decision device are linear operations. Hence, under the assumption of *additive* noise, it is sufficient to study the noise characteristics in the absence of a transmitted signal.

Throughout this paper, the noise at the input of the receiver is modeled as a discrete-time, real-valued (due to baseband signalling), wide-sense stationary (not necessarily Gaussian) random process $\mathbf{z} = [z(n)]_{n=-\infty,\ldots,+\infty}$, i.e., we are given

the mean[1] $\mu_{\mathbf{z}}(n) = \mathcal{E}\{z(n)\} = \mu_{\mathbf{z}}$ and the autocorrelation function $R_{\mathbf{z}}(m+n,m) = \mathcal{E}\{z(m+n)z^*(m)\} = R_{\mathbf{z}}(n)$, where the superscript * denotes complex conjugation. Complex conjugation is of course redundant for real-valued random processes but since we are also dealing with complex-valued random processes later on, we write it here for completeness. Note that this model includes the practically more relevant assumption of *colored noise*.[2]

In the following, we analyze the first and second moments of the noise at the input of the frequency–domain equalizer (FDE). From this, we can easily derive the moments at the input of the decision device.

The first part of the receiver transforms the random process at the input into a sequence of real random vectors. Each of these vectors has a first- and second-order description of its statistical properties, i.e., a mean vector and a covariance matrix. Due to the stationarity of the input random process, these mean vectors and covariance matrices are time independent. Furthermore, they do not depend on (time) shifts of the serial to parallel conversion. In the next part of the receiver, each random vector of the sequence is passed through a discrete Fourier transform (DFT) which in turn produces a sequence of random vectors, this time complex valued. Note that for complex random vectors, mean vector and covariance matrix are not sufficient for a complete first- and second-order description of their statistical properties; one also needs the *pseudocovariance* matrix (or *complementary* covariance matrix [16]). The pseudo-covariance matrix $\mathbf{P}_{\mathbf{w}} = [P_{\mathbf{w}}(k,l)]_{k,l=0,\ldots,N-1}$ of a complex random vector $\mathbf{w} = [w(k)]_{k=0,\ldots,N-1}$ is defined as $P_{\mathbf{w}}(k,l) = \mathcal{E}\{(w(k) - \mu_{\mathbf{w}}(k))(w(l) - \mu_{\mathbf{w}}(l))\}$, whereas the covariance matrix $\mathbf{C}_{\mathbf{w}} = [C_{\mathbf{w}}(k,l)]_{k,l=0,\ldots,N-1}$ is defined as $C_{\mathbf{w}}(k,l) = \mathcal{E}\{(w(k) - \mu_{\mathbf{w}}(k))(w(l) - \mu_{\mathbf{w}}(l))^*\}$. Again, all mean vectors, covariance matrices, and pseudocovariance matrices are the same and do not depend on previous (time) shifts of the serial-to-parallel conversion. For other literature that deals with or is at least related to the pseudocovariance matrix we refer to [1]–[4], [14], [15], [20]–[29], [31]–[33], and [35].

We proceed with the following definition.

Definition 1: A random vector is called *rotationally invariant* (see [14]) or *circularly symmetric* (see [16]) or *proper* (see [18]) if its pseudocovariance matrix vanishes.

As it is shown in [18], for a rotationally invariant random vector, its real and imaginary part vectors have[3] identical

---

[1]Note that this process is usually zero-mean in our application. However, there is the possibility that there are other applications where a discrete-time, real-valued, wide-sense stationary random process is passed (blockwise) through a DFT that is **not** zero-mean. We can still apply our analysis to such situations.

[2]Crosstalk is colored noise and the time–domain equalizer transforms white noise into colored noise as well.

[3]In fact, these properties can be equivalently used for the definition of rotational invariance. Wooding [34] and Goodman [6] were apparently the first to deal with random vectors satisfying these conditions.
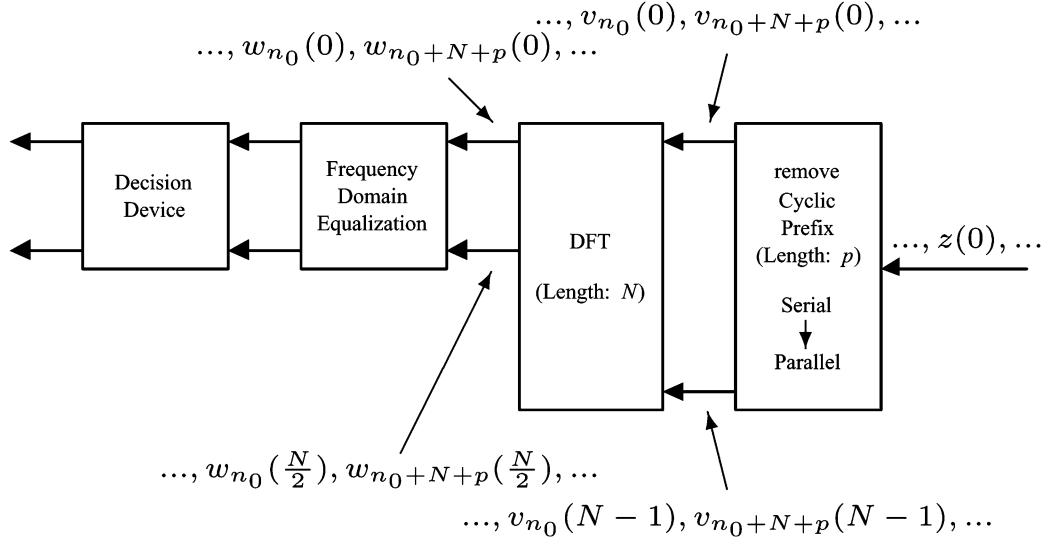
Fig. 1. Part of a DMT receiver.

(auto)covariance matrices and a skew-symmetric [7] cross-covariance matrix [16]. As an immediate consequence, real and imaginary parts of an element of this random vector have identical variances and are uncorrelated. Using the theory of proper random vectors developed in [18], it can be easily seen that the random vectors at the output of the DFT are rotationally *variant* except for the case when the input random vectors are constant with probability 1, i.e., roughly speaking, they are deterministic.[4] This suggests that at the output of the DFT and also at the input of the decision device, real and imaginary parts at certain frequencies have different variances and/or are correlated in general.

Remark: In the case of **passband** orthogonal frequency-division multiplexing (OFDM), the situation is different. At the input of the receiver, the demodulation of the signal (passband-to-baseband conversion) requires the calculation of an *analytic signal*. It is shown in [3] that the analytic signal of any stationary signal is rotationally invariant. It follows that all considered random vectors are rotationally invariant as well.

II. FIRST AND SECOND MOMENTS

To verify this conjecture about correlations and variance differences, we calculate mean vector, covariance matrix, and pseudocovariance matrix of the complex random vector $\mathbf{w}_{n_0} = [w_{n_0}(0) \ldots w_{n_0}(N-1)]^T$ (the superscript T denotes transposition without complex conjugation) at the output of the DFT analytically. Throughout this paper, N is assumed to be even. Due to the Hermitian symmetry satisfied by the DMT subcarrier symbols [5], the DMT receiver can ignore the subcarriers $N/2 + 1, \ldots, N$ (cf. Fig. 1). The same Hermitian symmetry is also satisfied by the noise samples at the DFT output, i.e., $w_{n_0}(k) = w_{n_0}^*(N-k)$, $k = 1, \ldots, (N/2) - 1$, and[5] $\Im\{w_{n_0}(0)\} = \Im\{w_{n_0}(N/2)\} = 0$.

Let us recall (cf. Fig. 1) that the real (DFT-)input random vector $\mathbf{v}_{n_0} = [v_{n_0}(0) \ldots v_{n_0}(N-1)]^T$, in which

$v_{n_0}(n) = z(n_0 + n)$, $n = 0, \ldots, N - 1$, is part of the discrete-time, real-valued, wide-sense stationary random process $\mathbf{z} = [z(n)]_{n=-\infty,\ldots,+\infty}$ with mean $\mu_{\mathbf{z}}$ and autocorrelation function $R_{\mathbf{z}}(n)$. Applying the DFT to $\mathbf{v}_{n_0}$, we obtain

$$w_{n_0}(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} v_{n_0}(n) e^{-j(2\pi/N)nk}$$
$$= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} z(n_0 + n) e^{-j(2\pi/N)nk},$$
$$k = 0, \ldots, N - 1$$

and

$$\mu_{\mathbf{w}}(k) = \mathcal{E}\{w_{n_0}(k)\} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mu_{\mathbf{z}} e^{-j(2\pi/N)nk}$$
$$= \begin{cases} \sqrt{N}\mu_{\mathbf{z}}, & k = 0 \\ 0, & k = 1, \ldots, N - 1 \end{cases}. \tag{1}$$

Note that the mean vector $[\mu_{\mathbf{w}}(0) \ldots \mu_{\mathbf{w}}(N-1)]^T$ is real and does not depend on n_0. The elements of covariance and pseudocovariance matrix of $\mathbf{w}_{n_0}$ are given by[6] $C_{\mathbf{w}}(k,l) = \mathcal{E}\{w_{n_0}(k)w_{n_0}^*(l)\} - \mu_{\mathbf{w}}(k)\mu_{\mathbf{w}}(l)$ and $P_{\mathbf{w}}(k,l) = \mathcal{E}\{w_{n_0}(k)w_{n_0}(l)\} - \mu_{\mathbf{w}}(k)\mu_{\mathbf{w}}(l)$, respectively, and can be written as

$$C_{\mathbf{w}}(k,l) = Q_{\mathbf{w}}(k,-l) - \mu_{\mathbf{w}}(k)\mu_{\mathbf{w}}(l) \tag{2a}$$
$$P_{\mathbf{w}}(k,l) = Q_{\mathbf{w}}(k,l) - \mu_{\mathbf{w}}(k)\mu_{\mathbf{w}}(l) \tag{2b}$$

with the function

$$Q_{\mathbf{w}}(k,l) = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \mathcal{E}\{z(n_0+n)z(n_0+m)\}$$
$$\times e^{-j(2\pi/N)(nk+ml)}$$
$$= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} R_{\mathbf{z}}(n-m) e^{-j(2\pi/N)(nk+ml)}$$

that in turn depends on the autocorrelation function $R_{\mathbf{z}}(n)$. Note that $Q_{\mathbf{w}}(k,l)$ is periodic both in k and l with period N, whereas $R_{\mathbf{z}}(n)$ is not periodic in general.

---

[4]That is the only situation when a *real* random vector happens to be rotationally invariant. In this case, both covariance and pseudocovariance matrix are the zero matrix.

[5]$\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote real and imaginary parts, respectively.

[6]The subsequent equations show that they do not depend on n_0.

The next step is to simplify the expression for $Q_{\mathbf{w}}(k,l)$. The idea is to reorder the terms of the double sum, so that only one sum remains (after some calculations). Proceeding[7] as in Section VII, we obtain

$$Q_{\mathbf{w}}(k,l) = \frac{1}{N} R_{\mathbf{z}}(0) \sum_{t=0}^{N-1} e^{-\jmath(2\pi/N)t(k+l)} + \frac{1}{N} \sum_{s=1}^{N-1} R_{\mathbf{z}}(s)$$
$$\times \sum_{t=0}^{N-1-s} \left(e^{-\jmath(2\pi/N)(sk+tk+tl)} \right.$$
$$\left. + e^{-\jmath(2\pi/N)(sl+tk+tl)} \right). \quad (3)$$

Using the formula

$$\sum_{t=0}^{M-1} a^t = \begin{cases} M, & a = 1 \\ \dfrac{1-a^M}{1-a}, & a \neq 1 \end{cases}$$

equation (3) further simplifies to (4), shown at the bottom of the page (cf. Section VII).

III. FREQUENCY-DEPENDENT NOISE ANALYSIS

In this section, we are interested in the first- and second-order statistics of the noise at certain fixed frequencies. More specifically, we specialize the results of Section II to obtain the noise characteristics of the complex random variables $w_{n_0}(k) = \Re\{w_{n_0}(k)\} + \jmath\Im\{w_{n_0}(k)\}$, $k = 0, \ldots, N-1$, where the frequency index k is considered as a fixed parameter. In particular, we investigate the covariance matrices

$$\mathbf{C}_{\hat{\mathbf{w}}_{n_0}(k)} = \mathcal{E}\left\{ \left(\hat{\mathbf{w}}_{n_0}(k) - \mathcal{E}\{\hat{\mathbf{w}}_{n_0}(k)\} \right) \right.$$
$$\left. \times \left(\hat{\mathbf{w}}_{n_0}(k) - \mathcal{E}\{\hat{\mathbf{w}}_{n_0}(k)\} \right)^T \right\}$$

and their properties of the real 2-D random vectors $\hat{\mathbf{w}}_{n_0}(k) = [\Re\{w_{n_0}(k)\}\Im\{w_{n_0}(k)\}]^T$ for $k = 0, \ldots, N-1$. Our analysis is based on Theorem 2.

Theorem 2: Let x denote a complex random variable with mean $\mu_x = \mathcal{E}\{x\}$, covariance $C_x = \mathcal{E}\{(x - \mu_x)(x - \mu_x)^*\}$, and pseudocovariance[8] $P_x = \mathcal{E}\{(x - \mu_x)(x - \mu_x)\} = r_x e^{\jmath\varphi_x}$, $r_x, \varphi_x \in \mathbb{R}$, respectively. Then, the covariance matrix of the 2-D real random vector $\hat{\mathbf{x}} = [\Re\{x\}\Im\{x\}]^T$ is given by

$$\mathbf{C}_{\hat{\mathbf{x}}} = \frac{1}{2}\begin{bmatrix} C_x + r_x\cos\varphi_x & r_x\sin\varphi_x \\ r_x\sin\varphi_x & C_x - r_x\cos\varphi_x \end{bmatrix} \quad (5)$$

[7]For the general case (MIMO case), this calculation is carried out in full detail in Section VII.

[8]For rotationally invariant variables, φ_x can be any number.

and has an eigenvalue decomposition

$$\mathbf{C}_{\hat{\mathbf{x}}} = \mathbf{U} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \mathbf{U}^T \quad (6)$$

with

$$\lambda_1 = \frac{C_x + r_x}{2} \quad \lambda_2 = \frac{C_x - r_x}{2} \quad (7)$$

and

$$\mathbf{U} = \begin{bmatrix} \cos\dfrac{\varphi_x}{2} & -\sin\dfrac{\varphi_x}{2} \\ \sin\dfrac{\varphi_x}{2} & \cos\dfrac{\varphi_x}{2} \end{bmatrix}. \quad (8)$$

Proof: For (5), see [18] or [29]. For (6) with (7) and (8), observe that

$$\mathbf{C}_{\hat{\mathbf{x}}} = \frac{1}{2}\begin{bmatrix} C_x & 0 \\ 0 & C_x \end{bmatrix} + \frac{r_x}{2}\underbrace{\begin{bmatrix} \cos\varphi_x & \sin\varphi_x \\ \sin\varphi_x & -\cos\varphi_x \end{bmatrix}}_{\mathbf{U}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\mathbf{U}^T}.$$

$\square$

Note that we do *not* require $r_x \geq 0$, i.e., we explicitly *allow* ambiguities of φ_x in this "relaxed" polar coordinates representation of P_x.

Next, we intend to apply Theorem 2 to the complex random variable $w_{n_0}(k)$. Specializing (2) to $k = l$, while inserting (1), we obtain the equation shown at the bottom of the next page, and, furthermore, (9), shown at the bottom of the next page. Hence, the eigenvalues and eigenvector matrices of the eigenvalue decompositions

$$\mathbf{C}_{\hat{\mathbf{w}}_{n_0}(k)} = \mathbf{U}(k)\begin{bmatrix} \lambda_1(k) & 0 \\ 0 & \lambda_2(k) \end{bmatrix}\mathbf{U}(k)^T$$

are found for $k = 1, \ldots, (N/2) - 1$ as

$$\lambda_1(k) = \frac{R_{\mathbf{z}}(0)}{2} + \sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \frac{s}{N}\right)\cos\left(\frac{2\pi}{N}ks\right)$$
$$- \frac{1}{N\sin\left(\dfrac{2\pi}{N}k\right)}\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\sin\left(\frac{2\pi}{N}ks\right) \quad (10a)$$

$$\lambda_2(k) = \frac{R_{\mathbf{z}}(0)}{2} + \sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \frac{s}{N}\right)\cos\left(\frac{2\pi}{N}ks\right)$$
$$+ \frac{1}{N\sin\left(\dfrac{2\pi}{N}k\right)}\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\sin\left(\frac{2\pi}{N}ks\right) \quad (10b)$$

$$Q_{\mathbf{w}}(k,l) = \begin{cases} R_{\mathbf{z}}(0) + 2\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \dfrac{s}{N}\right)\cos\left(\dfrac{2\pi}{N}ks\right), & k+l = 0 \text{ or } k+l = N \\ -\dfrac{e^{\jmath(2\pi/2N)(k+l)}}{N\sin\left(\dfrac{2\pi}{2N}(k+l)\right)}\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(\sin\left(\dfrac{2\pi}{N}ks\right) + \sin\left(\dfrac{2\pi}{N}ls\right)\right), & k+l \neq 0 \text{ and } k+l \neq N. \end{cases} \quad (4)$$

and

$$\mathbf{U}(k) = \begin{bmatrix} \cos\left(\dfrac{\pi}{N}k\right) & -\sin\left(\dfrac{\pi}{N}k\right) \\ \sin\left(\dfrac{\pi}{N}k\right) & \cos\left(\dfrac{\pi}{N}k\right) \end{bmatrix} \tag{11}$$

respectively. We also conclude from Theorem 2, especially from (5), that the noise variances (and thus the noise powers) for real and imaginary parts at certain frequencies are *different* in general. This immediately follows from the observation that $r_{w_{n_0}(k)} \cos(\varphi_{w_{n_0}(k)})$ is nonvanishing in general. Due to the properties of the DFT of a real vector (cf., for example, [30]), this is not surprising for $k = 0$ and $k = (N/2)$. Theorem 2 together with (9) implies that this also happens for $k \neq 0$ and $k \neq (N/2)$ and that there are *statistical dependencies* between real and imaginary parts of the noise samples, since $r_{w_{n_0}(k)} \sin(\varphi_{w_{n_0}(k)})$ is nonvanishing in general as well. It is important to observe that there is neither a correlation nor a variance difference between real and imaginary parts at a certain frequency k if and only if $r_{w_{n_0}(k)} = 0$. For $k \neq 0$ and $k \neq (N/2)$ [cf. (9a)], $r_{w_{n_0}(k)}$ can be rewritten as

$$r_{w_{n_0}(k)} = -\frac{2}{N\sin\left(\dfrac{2\pi}{N}k\right)} \sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\sin\left(\frac{2\pi}{N}ks\right)$$

$$= -\frac{2}{N\sin\left(\dfrac{2\pi}{N}k\right)} \Im\left\{\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)e^{j(2\pi/N)ks}\right\}$$

and, using the properties of the DFT (cf., for example, [30]), we are now able to state Theorem 3.

Theorem 3: Correlations or variance differences between real and imaginary parts of the noise at the output of the DFT do *not* occur at *any* frequency $k = 1, \ldots, (N/2) - 1$ if and only if the autocorrelation function $R_{\mathbf{z}}(n)$ of the real-valued input noise process $\mathbf{z}$ satisfies the relation

$$R_{\mathbf{z}}(n) = R_{\mathbf{z}}(N - n) \qquad \forall\, n = 1, \ldots, N-1. \tag{12}$$

Observing [cf. (2b), (1), and (4)] that $\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\sin((2\pi/N)ks) = 0$ for $k = 1, \ldots, (N/2) - 1$ also implies $P_{\mathbf{w}}(k, l) = 0$ for $k, l = 1, \ldots, (N/2) - 1$ and we obtain the even stronger theorem.

Theorem 4: The complex random noise vector $[w_{n_0}(1)\ldots w_{n_0}((N/2) - 1)]^T$ at the output of the DFT is rotationally invariant if and only if the autocorrelation function $R_{\mathbf{z}}(n)$ of the real-valued input noise process $\mathbf{z}$ satisfies the relation

$$R_{\mathbf{z}}(n) = R_{\mathbf{z}}(N - n) \qquad \forall\, n = 1, \ldots, N-1. \tag{13}$$

Note that the same theorems hold essentially for the noise at the input of the decision device, since rotational invariance is invariant under linear transformations [18], [29].

Theorem 3 and 4 tell us that correlations and/or variance differences occur only in the presence of *colored* noise. It is obvious that, in practical systems, the conditions (12) and (13) on

$$C_{w_{n_0}(k)} = \begin{cases} R_{\mathbf{z}}(0) + 2\sum\limits_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \dfrac{s}{N}\right) - N\mu_{\mathbf{z}}^2, & k = 0 \\[2ex] R_{\mathbf{z}}(0) + 2\sum\limits_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \frac{s}{N}\right)\cos\left(\dfrac{2\pi}{N}ks\right), & k \neq 0 \end{cases}$$

$$P_{w_{n_0}(k)} = \begin{cases} R_{\mathbf{z}}(0) + 2\sum\limits_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \dfrac{s}{N}\right) - N\mu_{\mathbf{z}}^2, & k = 0 \\[2ex] R_{\mathbf{z}}(0) + 2\sum\limits_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \dfrac{s}{N}\right)(-1)^s, & k = \dfrac{N}{2} \\[2ex] -\dfrac{2e^{j(2\pi/N)k}}{N\sin\left(\dfrac{2\pi}{N}k\right)}\sum\limits_{s=1}^{N-1} R_{\mathbf{z}}(s)\sin\left(\dfrac{2\pi}{N}ks\right), & k \neq 0 \text{ and } k \neq \dfrac{N}{2} \end{cases}$$

$$r_{w_{n_0}(k)} = \begin{cases} R_{\mathbf{z}}(0) + 2\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \dfrac{s}{N}\right) - N\mu_{\mathbf{z}}^2, & k = 0 \\[2ex] R_{\mathbf{z}}(0) + 2\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\left(1 - \dfrac{s}{N}\right)(-1)^s, & k = \dfrac{N}{2} \\[2ex] -\dfrac{2}{N\sin\left(\dfrac{2\pi}{N}k\right)}\sum_{s=1}^{N-1} R_{\mathbf{z}}(s)\sin\left(\dfrac{2\pi}{N}ks\right), & k \neq 0 \text{ and } k \neq \dfrac{N}{2} \end{cases} \tag{9a}$$

$$\varphi_{w_{n_0}(k)} = \begin{cases} 0, & k = 0 \text{ or } k = \dfrac{N}{2} \\[2ex] \dfrac{2\pi}{N}k, & k \neq 0 \text{ and } k \neq \dfrac{N}{2}. \end{cases} \tag{9b}$$

the autocorrelation function[9] are never fulfilled, i.e., there will be (at least) one $n \in \{1, \ldots, N-1\}$ with $R_{\mathbf{z}}(n) \neq R_{\mathbf{z}}(N-n)$, so one has to be aware that the noise powers for real and imaginary parts are different.

IV. CONSEQUENCES AND ASYMPTOTIC ANALYSIS OF THE NOISE CHARACTERISTICS

Consider again a common DMT system as described in [5] or [29]. In the transmitter, the bit stream is mapped onto quadratic quadrature amplitude modulation (QAM) constellations of size and density that depend on the noise power on the corresponding carrier/frequency, e.g., according to the water filling solution. However, this implicitly assumes that the noise power of real and imaginary parts are equal. As we showed in Section III, this is not the case in general. In order to be better adapted to the discovered noise characteristics, one should use *rotated (nonsquare) rectangular* complex symbol alphabets instead, i.e., symbol constellations of the form

$$\left\{ (t_1 V_1 + j t_2 V_2) e^{j\theta} : t_1 \in \{\pm 1, \pm 3, \ldots, \pm (M_1 - 1)\} \right.$$
$$\left. t_2 \in \{\pm 1, \pm 3, \ldots, \pm (M_2 - 1)\} \right\}$$

where θ is the rotation angle, V_1 and V_2 are certain gain factors, and M_1 and M_2 are the (even) numbers of signal points in the two orthogonal directions, respectively. The rotation angles can be obtained from the orthonormal eigenvector matrices[10] [cf. (11)], or directly from Theorem 2 together with (9b) and the rotations introduced by the frequency–domain equalizer (cf., also, Fig. 1), whereas size and density are determined by the eigenvalues [cf. (10)] and the scalings introduced by the frequency–domain equalizer (cf., also, Fig. 1).

Theorem 5: The angles of the noise rotations at the output of the DFT

$$\phi(k) = \begin{cases} \dfrac{\pi}{N} k, & k = 0, \ldots, \dfrac{N}{2} - 1 \\ 0, & k = \dfrac{N}{2} \end{cases}$$

are *independent* of the mean $\mu_{\mathbf{z}}$ and of the autocorrelation function $R_{\mathbf{z}}(n)$ of the noise at the input of the receiver.

To get a better picture of the deviations from common quadratic QAM constellations for the (ideal) rotated rectangular constellations, we have to look at the relative differences

$$d(k) = \frac{\lambda_2(k) - \lambda_1(k)}{\lambda_1(k) + \lambda_2(k)} \tag{14}$$

which can be viewed as a measure for the eccentricity of the noise ellipses and in turn for the *nonsquareness* of the rotated rectangular constellations. Note that the scaling factor that is

[9]An inspection of (12) shows that this condition is equivalent to the requirement that the autocorrelation function shifted to the left by $(N/2)$ is an even function within the interval $[-((N/2)-1), +((N/2)-1)]$. This must not be mixed up with the fact that any **nonshifted** autocorrelation function is an even function.

[10]These matrices are rotation matrices and describe the noise rotations.
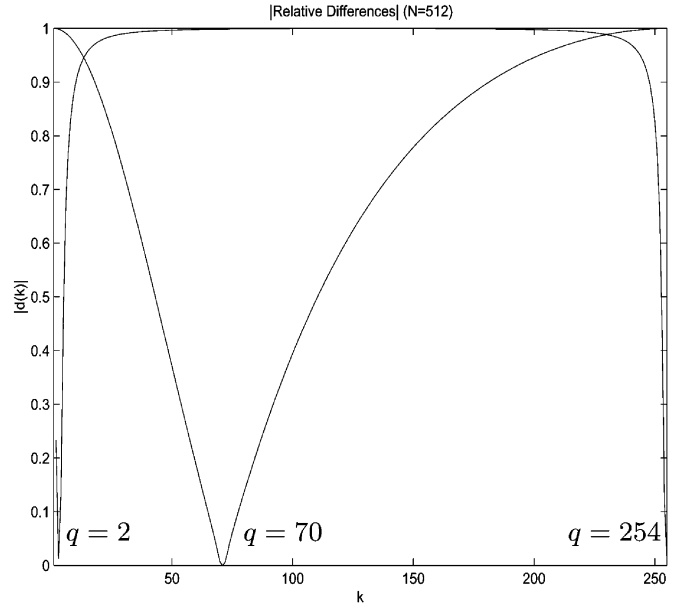

Fig. 2. Relative differences. The noise is zero mean filtered white Gaussian noise with variance 1.

introduced by the frequency–domain equalizer is canceled out in the normalized expression (14). From (10), we obtain

$$\lambda_1(k) + \lambda_2(k) = R_{\mathbf{z}}(0) + 2 \sum_{s=1}^{N-1} R_{\mathbf{z}}(s) \left(1 - \frac{s}{N}\right) \cos\left(\frac{2\pi}{N} ks\right).$$

We omit an analytic treatment of the relative differences. In order to interpret the relative differences, consider the following two examples.

Example 1: This example is artificially constructed and it merely shows the existence of extreme cases. The noise at the receiver input is zero-mean filtered white Gaussian noise with filter impulse response (q is a fixed parameter)

$$g_q(n) = \begin{cases} \left(1 - \cos\dfrac{2\pi n}{N}\right) \cos\dfrac{2\pi q n}{N}, & n = 0, \ldots, N-1 \\ 0, & n \neq 0, \ldots, N-1 \end{cases}.$$

From Fig. 2, one can see that significant differences between ideal rotated rectangular constellations and the quadratic QAM constellations used may occur. Note that if $|d(k)| = 1$ for a certain frequency k [this is the maximum value $|d(k)|$ can take, in this case, $\lambda_1(k) = 0$ or $\lambda_2(k) = 0$], the optimum rotated rectangular constellation at this frequency k collapses into a constellation whose symbols lie on a straight line.

Example 2: In the following, we assume that the random noise process at the input of the receiver is zero mean and has an autocorrelation function

$$R_{\mathbf{z}}(n) = \begin{cases} 1, & n = 0 \\ \dfrac{1}{2}, & n = -1 \text{ or } n = 1 \\ 0, & n \neq -1 \text{ and } n \neq 0 \text{ and } n \neq 1. \end{cases}$$

The corresponding relative (eigenvalue) difference at frequency $(N/2) - 1$ is given by

$$d\left(\frac{N}{2} - 1\right) = \frac{1}{N - (N - 1)\cos\left(\frac{2\pi}{N}\right)}$$

which tends to 1, if N tends to infinity. This shows that, in general, an increasing DFT-length N does *not* decrease this effect of having correlations and/or variance differences of real and imaginary parts. On the contrary, there are frequencies [dependent on N; here shown for frequency $(N/2) - 1$] for which nonsquare symbol constellations are becoming more and more appropriate. The bigger N becomes, the more dramatic the achievable gain is by using a nonsquare symbol constellation at frequency $(N/2) - 1$ instead of a common quadratic QAM constellation.

V. CAPACITY LOSS

In order to use rotated rectangular constellations, one has to modify the existing bit-loading/mapping algorithms, such that they take into account the powers of the real and imaginary parts of the noise. Note that the noise rotation angles (at the output of the DFT) do not depend on the actual noise characteristics (see Theorem 5) and do not need to be estimated during transmission. The additional noise rotations introduced by the frequency–domain equalizer are estimated anyway but have to be taken into account both at the transmitter and receiver. On the other hand, if one sticks to the conventional quadratic QAM constellations, one has to accept significant increases in symbol error probability (at least if there is no appropriate coding; cf. Section VI) and decreased capacity.

In case of Gaussian distributed noise, a specialization of the results of [28] (cf., also, [29]) yields the following (approximate) expression for the capacity loss (bit/channel use)

$$\Delta C \approx -\frac{1}{2(N + p)} \sum_{k=1}^{N/2-1} \log_2\left(1 - \frac{(\lambda_1(k) - \lambda_2(k))^2}{L^2 \left|H\left(e^{j2\pi k/N}\right)\right|^4}\right)$$

where $H(z)$ denotes the channel transfer function, and p and L are the length of the cyclic prefix and the water level (*water filling*), respectively. Note that L depends on the signal power. We note that this capacity loss is the loss in capacity that occurs if it is erroneously believed that the pseudocovariance matrix of the noise is vanishing. According to the *water filling* rule, this capacity loss is caused by a nonoptimum power allocation, since different powers of real and imaginary parts of the noise imply different powers of real and imaginary parts of the signal in general, which is not the case if the pseudocovariance matrix of the noise vector is the zero matrix. We use this capacity loss for an approximate expression of the rate loss that has to be accepted if quadratic QAM constellations are used instead of rotated rectangular (in general, nonquadratic) QAM constellations.

We performed some simulations in order to demonstrate the capacity loss. We used the parameters of a real ADSL scenario, i.e., a DFT-length of $N = 512$, a subcarrier spacing of 4312.5 Hz, and a transmit power of 100 mW. The transfer functions of the loops were obtained by measurements of
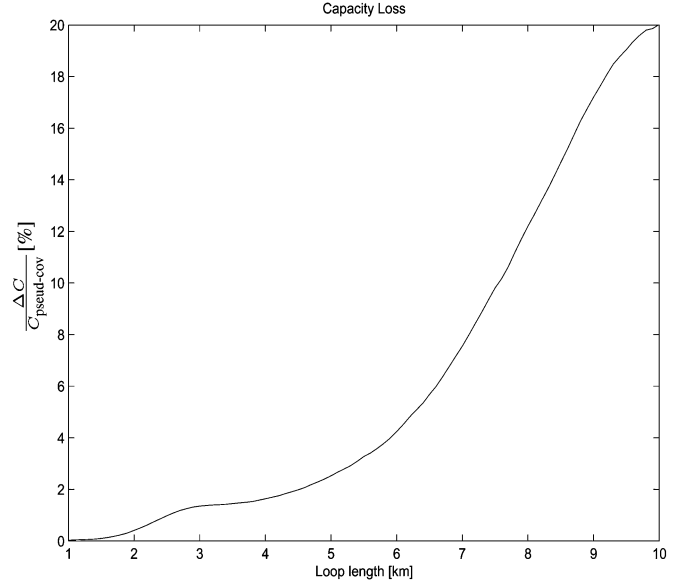

Fig. 3. Normalized capacity loss (with respect to the capacity $C_{\text{pseud-cov}}$ that assumes that the pseudocovariances of the complex noise variables at the input of the decision device are utilized) in terms of loop length.

Austrian cables. The noise model was as follows. We assumed two additive, statistically independent noise components. One was the typical noise environment in a cable bundle including crosstalk and background noise. The other was stationary narrowband interference with a bandwidth of 10 kHz, center frequency of 1.07 MHz, and 0 dBm power. The normalized capacity loss depending on the loop length can be seen in Fig. 3. Obviously, the capacity loss is not severe for short- and medium-length cables.

VI. SYMBOL ERROR PROBABILITY AND OPTIMIZED BIT LOADING

In this section, we are again interested in the gain achieved by rotated (nonsquare) rectangular constellations instead of quadratic QAM constellations. However, in contrast to Section V, where *capacity* was used to express this gain, we now use *uncoded symbol error probability* [and signal-to-noise ratio (SNR) gain] as a performance measure. Note that this measure can be regarded as the other extreme: While capacity results implicitly assume an optimum (not easily implementable) coding strategy, we do not take into account coding here at all.

Furthermore, we also address the issue of designing the optimum (rotated) rectangular constellations and present explicit formulas for the number of constellation points and gain factors in the respective directions.

We start our analysis with a basic property of rectangular constellations. Let us assume that the following rectangular constellation:

$$\left\{t_1 V_1(k) + j t_2 V_2(k) : t_1 \in \{\pm 1, \pm 3, \ldots, \pm(M_1(k) - 1)\}\right.$$
$$\left. t_2 \in \{\pm 1, \pm 3, \ldots, \pm(M_2(k) - 1)\}\right\}$$
$$(15)$$

is assigned to the kth subcarrier, where $V_1(k)$ and $V_2(k)$ are certain gain factors and $M_1(k)$ and $M_2(k)$ are the (even) numbers of signal points in real and imaginary part direction, respectively. Note that the total number of signal points is $M_{\text{all}}(k) = M_1(k)M_2(k)$. If these signal points are chosen with equal probabilities during transmission, the average signal power in real and imaginary part direction is calculated as ($i = 1, 2$)

$$
\begin{aligned}
S_i(k) &= \frac{2}{M_i(k)} \sum_{m=1}^{M_i(k)/2} (2m-1)^2 (V_i(k))^2 \\
&= \frac{2(V_i(k))^2}{M_i(k)} \sum_{m=1}^{M_i(k)/2} \left(4m^2 - 4m + 1\right) \\
&= \frac{1}{3}(V_i(k))^2 \left((M_i(k))^2 - 1\right)
\end{aligned}
\tag{16}
$$

where $\sum_{m=1}^{M} m = (M(M+1)/2)$ and $\sum_{m=1}^{M} m^2 = M(M+1)(2M+1)/6$ were used. If the number of signal points $M_i(k)$ is not too small, we can approximate the obtained signal power by

$$
S_i(k) \approx \frac{1}{3}(V_i(k))^2 (M_i(k))^2.
\tag{17}
$$

Next, we assume that a certain *conventional* bit-loading algorithm assigns a square QAM constellation to the kth subcarrier. This implies that the gain factors and the number of signal points in real and imaginary part direction, respectively, are equal [see (15)], i.e.,

$$
\begin{aligned}
V(k) &= V_1(k) = V_2(k) \\
M(k) &= M_1(k) = M_2(k).
\end{aligned}
$$

Note that the total number of signal points is $M_{\text{all}}(k) = (M(k))^2$. Specializing (16) and (17), the average signal power (in both real and imaginary part direction together) is obtained as

$$
\begin{aligned}
S_{\text{all}}(k) &= S_1(k) + S_2(k) = \frac{2}{3}(V(k))^2 (M_{\text{all}}(k) - 1) \\
&\approx \frac{2}{3}(V(k))^2 M_{\text{all}}(k)
\end{aligned}
\tag{18}
$$

the approximation being good, provided that the number of signal points $M_{\text{all}}(k)$ is not too small. At the input of the decision device, the eigenvalues of the covariance matrices of real and imaginary parts of the noise at certain frequencies are given by (see, also, [5] or [29])

$$
\begin{aligned}
\nu_1(k) &= \left| H\left(e^{j2\pi k/N}\right) \right|^{-2} \lambda_1(k) \\
\nu_2(k) &= \left| H\left(e^{j2\pi k/N}\right) \right|^{-2} \lambda_2(k).
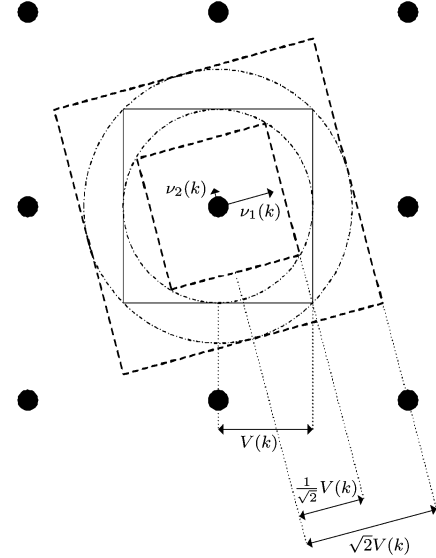\end{aligned}
\tag{19}
$$

Fig. 4. Decision region of interior point in square QAM constellation (bounded by solid line) and lower and upper bound areas (bounded by dashed lines). The edges of these lower and upper bound areas are chosen to be parallel to the eigenvectors of the noise and are margined by the inscribe circle and circum circle of the square decision region. This is the reason for the factors $1/\sqrt{2}$ and $\sqrt{2}$.

In the following, we will assume $\nu_1(k) > \nu_2(k)$ without loss of generality. In the presence of Gaussian noise, the symbol error probabilities can be approximated by (see Fig. 4)

$$
\begin{aligned}
P_{\text{Square QAM}}(k) &\approx 2Q\left(\sqrt{\frac{(V(k))^2}{\nu_1(k)}}\right) + 2Q\left(\sqrt{\frac{(V(k))^2}{\nu_2(k)}}\right) \\
&\quad - 4Q\left(\sqrt{\frac{(V(k))^2}{\nu_1(k)}}\right) Q\left(\sqrt{\frac{(V(k))^2}{\nu_2(k)}}\right) \\
&\approx 2Q\left(\sqrt{\frac{(V(k))^2}{\nu_1(k)}}\right)
\end{aligned}
\tag{20}
$$

where the Q-function is defined as

$$
Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt.
$$

Note that the first approximation in (20) originates from the fact that only interior points of the QAM constellation are considered and that the noise rotations are *not* taken into account, i.e., it is (erroneously) assumed that the noise rotation angles are equal to 0, $\pi/2$, $-\pi/2$, or π. Observe that the noise rotations at the input of the decision device (in contrast to the noise rotations at the output of the DFT) are not only determined by Theorem 5, but also depend on $\arg\left\{H\left(e^{j2\pi k/N}\right)\right\}$. The second approximation is due to the fact that the function $Q\left(\sqrt{\cdot}\right)$ is a strictly monotone decreasing function having values close to zero for the arguments under consideration, or—to say it differently—the expression on the right-hand side of (20) is dominated by the first term. We can see from Fig. 4 that the decision region is always contained in another quadratic area whose edges are parallel to the "noise," or more precisely, to the eigenvectors of the noise.

At the same time, the decision region always contains such a quadratic area. This enables us to determine a lower and an upper bound for the symbol error probability, which are valid for all possible rotation angles. To see this, consider the probability of a correct decision, which can be calculated by integrating the noise probability distribution over the area bounded by the solid line as illustrated in Fig. 4. This probability can be obviously upper and lower bounded by integrating the noise probability distribution over the areas bounded by the dashed lines (cf. Fig. 4) and these bounds can be expressed in terms of the Q-function and are independent of the noise rotation angles. Noting that the symbol error probability is one minus the probability of a correct decision, these bounds immediately imply the following bounds for the symbol error probabilities:

$$2Q\left(\sqrt{2\cdot\frac{(V(k))^2}{\nu_1(k)}}\right) \lesssim P_{\text{Square QAM}}(k)$$

$$\lesssim 2Q\left(\sqrt{\frac{1}{2}\cdot\frac{(V(k))^2}{\nu_1(k)}}\right)$$

which are valid for all rotation angles. Again, these inequalities are only approximations, as indicated by the special "less or equal than" symbol, because boundary points of the QAM constellation are not incorporated in the formulas and some Q-function expressions are approximated by the zero value. However, we want to emphasize that these approximations are quite good and do not affect the validity of the inequalities significantly.

Inserting (18), we can summarize the results obtained for the symbol error probability for a square QAM constellation under the presence of rotationally variant noise as follows:

$$P_{\text{Square QAM}}(k) \approx 2Q\left(\sqrt{c\cdot\frac{3S_{\text{all}}(k)}{2M_{\text{all}}(k)\nu_1(k)}}\right) \quad (21)$$

where $c = 1$ is an approximation, which neglects the noise rotations (but not the power differences), and $c = 2$ and $c = 1/2$ yield lower and upper bounds, respectively, which are valid for all possible rotation angles.

We continue our analysis by considering (optimized) rotated rectangular constellations. First, observe that we can restrict ourselves to the case where the noise eigenvectors are parallel to the constellation edges, because at the decision point in Fig. 1, i.e., after frequency–domain equalization, the net noise rotations, the desired orientations of the transmitter's rotated rectangular constellations, and the rotations of the decision regions are all given by $\theta(k) = \phi(k) - \arg\{H\left(e^{j2\pi k/N}\right)\}$ (cf. [5] or [29]). $\phi(k)$ is *a priori* known from Theorem 5 and $H\left(e^{j2\pi k/N}\right)$ has to be estimated anyway in order to guarantee that the frequency–domain equalizer works properly. We will calculate the reduced symbol error probability that occurs if one applies an optimized (with respect to symbol error probability) rotated rectangular constellation which has the same number $M_{\text{all}}(k)$ of signal points, i.e., it serves the same data rate and uses the same signal power $S_{\text{all}}(k)$.

Similarly to the square QAM case, the symbol error probabilities can be approximated by

$$P_{\text{Rot Rect}}(k) \approx 2Q\left(\sqrt{\frac{(V_1(k))^2}{\nu_1(k)}}\right) + 2Q\left(\sqrt{\frac{(V_2(k))^2}{\nu_2(k)}}\right)$$

$$\approx 2Q\left(\sqrt{\frac{3S_1(k)}{(M_1(k))^2\nu_1(k)}}\right)$$

$$+ 2Q\left(\sqrt{\frac{3S_2(k)}{(M_2(k))^2\nu_2(k)}}\right) \quad (22)$$

where (17) was inserted. As already mentioned, the approximations are valid for all rotation angles and neglect only boundary points of the constellation and some small values of the Q-function. According to (17), we also assume that the number of signal points is not too small. In order to determine the optimum constellation parameters, we have to minimize the function

$$F\left(S_1(k), S_2(k), M_1(k), M_2(k)\right)$$

$$= 2Q\left(\sqrt{\frac{3S_1(k)}{(M_1(k))^2\nu_1(k)}}\right) + 2Q\left(\sqrt{\frac{3S_2(k)}{(M_2(k))^2\nu_2(k)}}\right)$$

under the side constraints

$$\begin{aligned} S_1(k) + S_2(k) &= S_{\text{all}}(k) \\ M_1(k)M_2(k) &= M_{\text{all}}(k) \\ S_i(k) &> 0, \qquad i = 1, 2 \\ M_i(k) &> 0 \text{ and even}, \qquad i = 1, 2. \end{aligned} \quad (23)$$

This minimization is carried out in the Appendix (cf., also, [5]). It is shown that an (optimum with respect to symbol error probability) bit-loading algorithm that takes into account power differences and statistical dependencies between real and imaginary parts of the noise distributes the same signal power onto the two axes of the rotated rectangular constellations, i.e.,

$$S_1(k) = S_2(k) = \frac{S_{\text{all}}(k)}{2}. \quad (24)$$

Furthermore, the optimum numbers of signal points in the two directions are derived as

$$M_1(k) = \sqrt{M_{\text{all}}(k)}\left(\frac{\nu_2(k)}{\nu_1(k)}\right)^{1/4} = \sqrt{M_{\text{all}}(k)}\left(\frac{\lambda_2(k)}{\lambda_1(k)}\right)^{1/4}$$

$$M_2(k) = \sqrt{M_{\text{all}}(k)}\left(\frac{\nu_1(k)}{\nu_2(k)}\right)^{1/4} = \sqrt{M_{\text{all}}(k)}\left(\frac{\lambda_1(k)}{\lambda_2(k)}\right)^{1/4}$$

which, of course, have to be rounded to even numbers. Using (16), the gain factors $V_i(k)$ are determined as well. We also want to emphasize that the main result of the Appendix that the solution obtained is indeed a minimum, can be just as well omitted from a practical point of view, because we will show in the following that the symbol error probability is substantially reduced when rotated rectangular constellations with the obtained parameters are used. This is obviously sufficient for

application purposes. On the other hand, it is shown that rotated rectangular constellations with these parameters are really the optimum rotated rectangular constellations with respect to symbol error probability. Finally, the symbol error probabilities (22) are calculated as

$$P_{\text{Rot Rect}}(k) \approx 4Q\left(\sqrt{\frac{3S_{\text{all}}(k)}{2M_{\text{all}}(k)\sqrt{\nu_1(k)\nu_2(k)}}}\right). \quad (25)$$

Result (24) seems to be somewhat surprising given that the *water filling* principle for parallel channels would suggest allocating more power to channels with smaller noise power. If the kth frequency–domain equalizer output is rotated in such a way that the noise eigenvectors coincide with the real and imaginary axes, then the additive noise along the two axes becomes uncorrelated (for Gaussian distributed noise independent), with variances $\nu_1(k)$ and $\nu_2(k)$. In other words, we have two independent parallel channels in our situation. Suppose now we have $S_1(k) \neq S_2(k)$ satisfying (23). In order to get the same symbol error probability on both channels [cf. (22)], we have to choose $M_1(k)$ and $M_2(k)$ satisfying (23), such that

$$\frac{3S_1(k)}{(M_1(k))^2\,\nu_1(k)} = \frac{3S_2(k)}{(M_2(k))^2\,\nu_2(k)}.$$

This yields

$$M_1(k) = \sqrt{M_{\text{all}}(k)}\left(\frac{S_1(k)}{S_2(k)}\frac{\nu_2(k)}{\nu_1(k)}\right)^{1/4}$$

and an overall symbol error probability

$$P'_{\text{Rot Rect}}(k) \approx 4Q\left(\sqrt{\frac{3\sqrt{S_1(k)S_2(k)}}{M_{\text{all}}(k)\sqrt{\nu_1(k)\nu_2(k)}}}\right)$$

which is greater than or equal to (25) owing to the relation between geometric and arithmetic mean. While the choice (24) is indeed better in terms of symbol error probability, it is not the best choice in terms of capacity.

From (21) and (25), we obtain[11] (approximate) SNR gains of

$$G(k) = \frac{\text{SNR}_{\text{Rot Rect}}(k)}{\text{SNR}_{\text{Square QAM}}(k)} \approx \frac{1}{c}\sqrt{\frac{\nu_1(k)}{\nu_2(k)}} = \frac{1}{c}\sqrt{\frac{\lambda_1(k)}{\lambda_2(k)}}$$

where c has the same meaning as in (21) and is explained below equation (21). Finally, we can express the gains in terms of the relative (eigenvalue) differences as

$$G(k) \approx \frac{1}{c}\sqrt{\frac{1+|d(k)|}{1-|d(k)|}}.$$

We want to emphasize that in our example of Section V the modulus of the relative differences is close to 1 for almost all frequencies (see Fig. 5), so that the overall SNR gain is very high, without much influence from the value of c (cf. Fig. 6).

[11]The SNR is defined via $P \sim Q\left(\sqrt{\text{SNR}}\right)$.
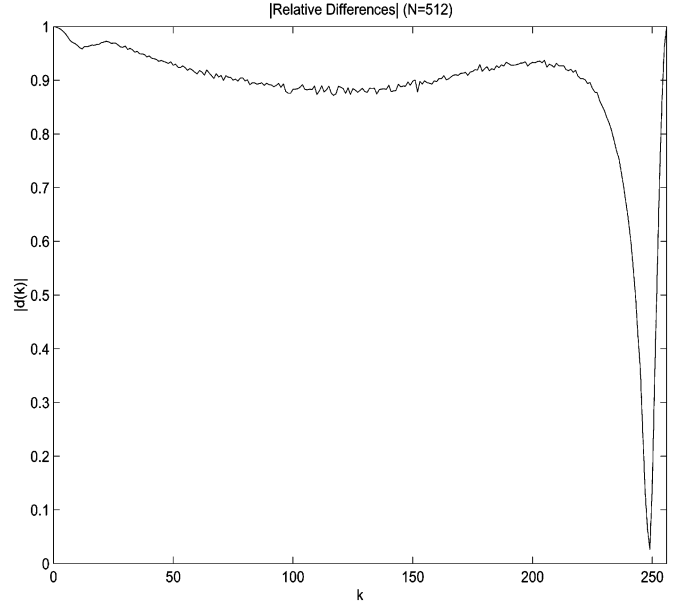
Fig. 5. Relative differences. Real ADSL scenario with narrowband interference.
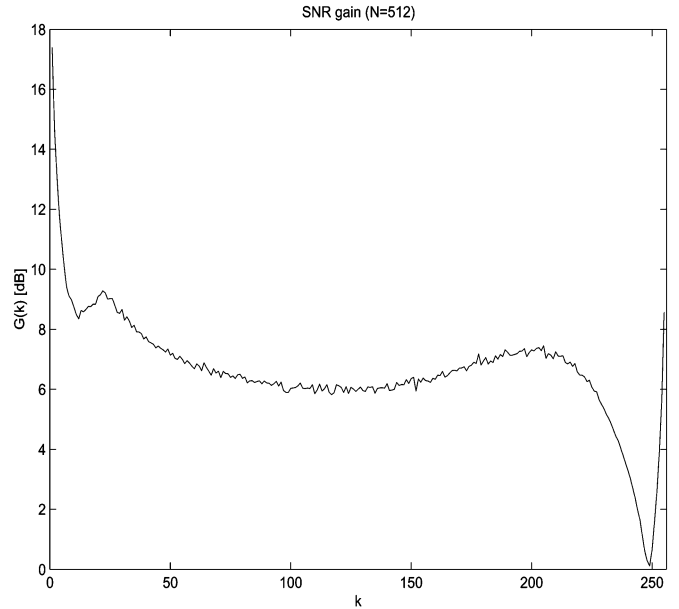
Fig. 6. SNR gains ($c = 1$). Real ADSL scenario with narrowband interference.

The SNR gains are independent of the channel transfer function and of the signal power and thus of the loop length, which is not the case for the capacity loss. Furthermore, the previous example shows that the use of rotated rectangular constellations is much more effective in reducing the (uncoded) symbol error probability than in increasing capacity. Note that statements about capacity always assume an optimum coding strategy which is not usually applicable in practice. For practical en-/decoders, the overall gain will be somewhere in-between. It depends on the ability of the code to use the more reliable transmission in one direction (corresponding to one eigenvector) to correct the more frequent errors in the other

Fig. 7. Multiple DMT receivers (relevant part).

direction (corresponding to the other eigenvector). The effort required to adapt the coding strategy is much higher than for implementing rotated rectangular constellations.

VII. EXTENSION TO THE MIMO CASE

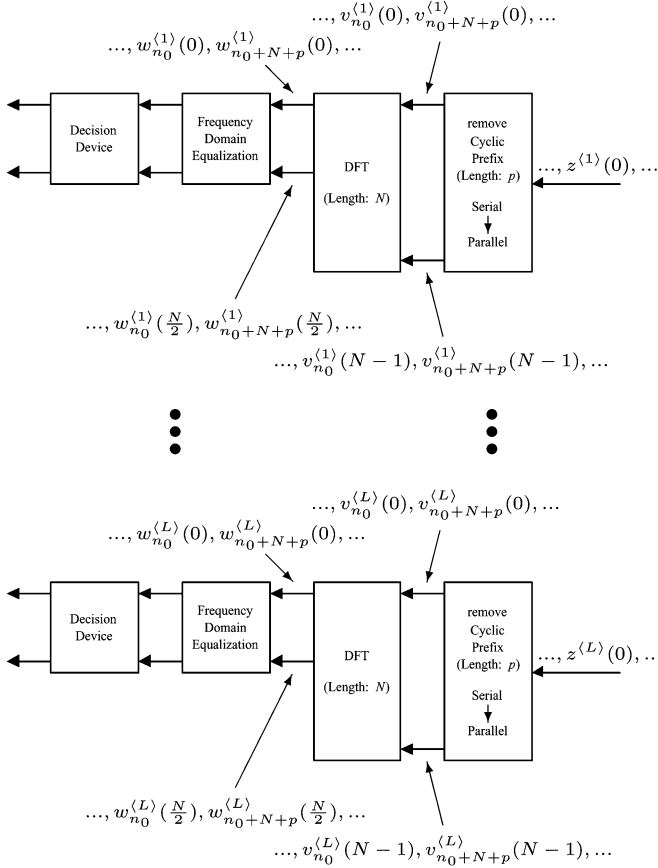In this section, we extend the results to a multiple-input–multiple-output (MIMO) scenario, i.e., to the situation where simultaneous transmission over several loops in a cable bundle takes place. We assume that all loops under consideration are equipped with a DMT modulation scheme having the same design parameters, in particular, the same (inverse) DFT lengths (denoted—similar to the single-loop case—by N) and the same lengths of the cyclic prefixes (denoted—again, similar to the single-loop case—by p). Furthermore, we assume that all transmitters and receivers are synchronized, i.e., all transmitters and receivers use the same clock and sampling rate. The number of considered loops is denoted by L and we will use a superscript bracket notation $\langle l \rangle$, $l = 1, \ldots, L$, to distinguish between signals on different loops (cf., also, Fig. 7), which illustrates the multiple receiver scenario.

The noise model is naturally generalized to the MIMO case. At each receiver input, there is a noise signal

$$\mathbf{z}^{\langle l \rangle} = \left[z^{\langle l \rangle}(n) \right]_{n=-\infty,\ldots,\infty}$$

where $l = 1, \ldots, L$ denotes the receiver number. These signals are modeled as discrete-time, real-valued (due to baseband signalling), (pairwise) jointly wide-sense stationary (not necessarily Gaussian) random processes with given means and cross-correlation (autocorrelation) functions [16]

$$\mu_{\mathbf{z}^{\langle l \rangle}}(n) = \mathcal{E}\left\{ z^{\langle l \rangle}(n) \right\} = \mu_{\mathbf{z}^{\langle l \rangle}} \tag{26a}$$

$$R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(m+n, m) = \mathcal{E}\left\{ z^{\langle l_1 \rangle}(m+n) z^{\langle l_2 \rangle *}(m) \right\}$$
$$= R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(n) \tag{26b}$$

where $l, l_1, l_2 = 1, \ldots, L$ and $n, m = -\infty, \ldots, \infty$. Note that this model incorporates dependencies between different ($l_1 \neq l_2$) noise signals $\mathbf{z}^{\langle l_1 \rangle}$ and $\mathbf{z}^{\langle l_2 \rangle}$ via the cross-correlation function $\mathbf{R}_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}$. Consider, e.g., the case when there is one dominant noise source in a cable bundle that influences all of its loops. Of course, there is also the possibility that the disturbances are independent of each other or one can think of hybrid situations, where a noise source only influences some nearby loops but not loops that are far away. Our model includes all cases mentioned so that we can assume that (26) covers a wide range of possible (colored) noise environments in a cable bundle.

In order to cope with the MIMO situation, we have to introduce the following notion of *cross-covariance matrix* and *pseudocross-covariance matrix*. Let $\mathbf{x}$ and $\mathbf{y}$ denote two complex random vectors with mean vectors $\boldsymbol{\mu}_{\mathbf{x}}$ and $\boldsymbol{\mu}_{\mathbf{y}}$, respectively. We will call the matrix

$$\mathbf{C}_{\mathbf{x},\mathbf{y}} = \mathcal{E}\{(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^H\}$$

the *cross-covariance matrix* and the matrix

$$\mathbf{P}_{\mathbf{x},\mathbf{y}} = \mathcal{E}\{(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^T\}$$

the *pseudocross-covariance matrix* of the random vectors $\mathbf{x}$ and $\mathbf{y}$. The superscript ^{H} denotes Hermitian transposition. According to Fig. 7, let $\mathbf{v}_{n_0}^{\langle l \rangle} = [z^{\langle l \rangle}(n_0) \ldots z^{\langle l \rangle}(n_0 + N - 1)]^T \in \mathbb{R}^N$ denote the noise vectors at the inputs of the DFTs and let $\mathbf{w}_{n_0}^{\langle l \rangle} = [w_{n_0}^{\langle l \rangle}(0) \ldots w_{n_0}^{\langle l \rangle}(N - 1)]^T$ denote the noise vectors at the outputs of the DFTs. In order to fully characterize the first and (joint) second moments of $\mathbf{w}_{n_0}^{\langle l \rangle}$, $l = 1, \ldots, L$, the mean vectors and both

$$\mathbf{C}_{\mathbf{w}_{n_0}^{\langle l_1 \rangle}, \mathbf{w}_{n_0}^{\langle l_2 \rangle}} \text{ and } \mathbf{P}_{\mathbf{w}_{n_0}^{\langle l_1 \rangle}, \mathbf{w}_{n_0}^{\langle l_2 \rangle}} \qquad \forall l_1, l_2 = 1, \ldots, L \tag{27}$$

must be specified [19].

Similarly to Section II, the kth elements of the mean vectors (do not depend on n_0) are calculated as

$$\mu_{\mathbf{w}^{\langle l \rangle}}(k) = \begin{cases} \sqrt{N} \mu_{\mathbf{z}^{\langle l \rangle}}, & k = 0 \\ 0, & k = 1, \ldots, N - 1 \end{cases} \tag{28}$$

and we can write the (k_1, k_2)th elements of (27), again no dependence on n_0, as

$$C_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2) = Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, -k_2)$$
$$- \mu_{\mathbf{w}^{\langle l_1 \rangle}}(k_1) \mu_{\mathbf{w}^{\langle l_2 \rangle}}(k_2) \tag{29a}$$
$$P_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2) = Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$$
$$- \mu_{\mathbf{w}^{\langle l_1 \rangle}}(k_1) \mu_{\mathbf{w}^{\langle l_2 \rangle}}(k_2) \tag{29b}$$
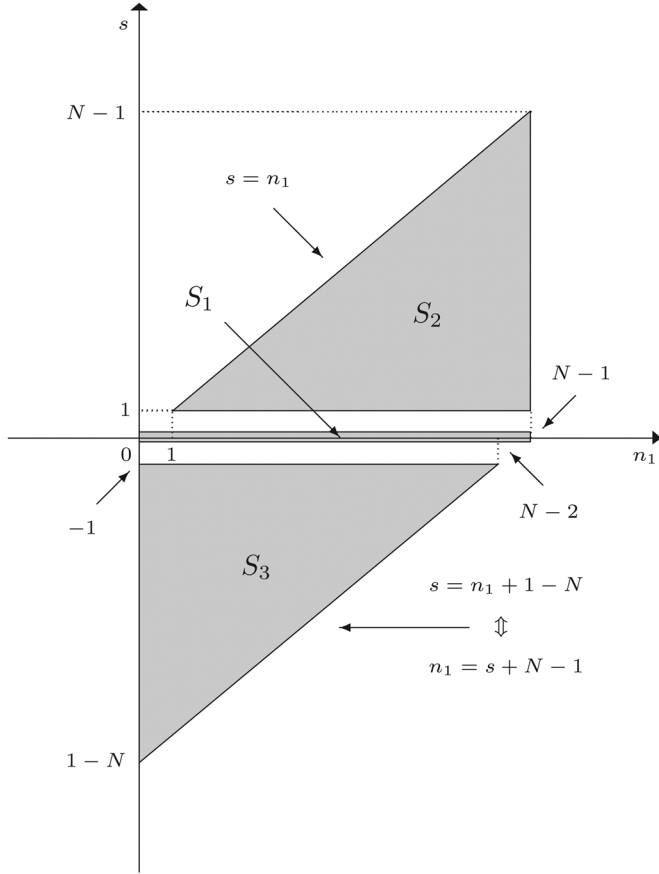
Fig. 8. Summation set in (30).

where $k_1, k_2 = 0, \ldots, N-1$ with the function

$$Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$$
$$= \frac{1}{N} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} \mathcal{E}\{z^{\langle l_1 \rangle}(n_0 + n_1) z^{\langle l_2 \rangle}(n_0 + n_2)\}$$
$$\times e^{-\jmath(2\pi/N)(n_1 k_1 + n_2 k_2)}$$
$$= \frac{1}{N} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(n_1 - n_2) e^{-\jmath(2\pi/N)(n_1 k_1 + n_2 k_2)}.$$

The next step is to simplify the expression for $Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$. The idea is to reorder the terms of the double sum, so that only one sum remains (after some calculations). We have

$$Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$$
$$= \frac{1}{N} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(n_1 - n_2) e^{-\jmath(2\pi/N)(n_1 k_1 + n_2 k_2)}$$
$$= \frac{1}{N} \sum_{n_1=0}^{N-1} \sum_{s=n_1+1-N}^{n_1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) e^{-\jmath(2\pi/N)(n_1 k_1 + n_1 k_2 - s k_2)}$$

$$(30)$$

where the index change $s = n_1 - n_2$ has been performed, such that the summation over n_2 is replaced by a summation over s. Next, we interchange the two sums. Due to the dependence of n_1 in the inner sum, we have to investigate the summation over

(n_1, s) in some more detail. Fig. 8 shows the effective pairs that are used in the two sums. They are denoted by the sets S_1, S_2, and S_3. Hence

$$Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$$
$$= \frac{1}{N} \sum_{n_1=0}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(0) e^{-\jmath(2\pi/N)n_1(k_1 + k_2)} \quad (S_1)$$
$$+ \frac{1}{N} \sum_{s=1}^{N-1} \sum_{n_1=s}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s)$$
$$\times e^{-\jmath(2\pi/N)(n_1 k_1 + n_1 k_2 - s k_2)} \quad (S_2)$$
$$+ \frac{1}{N} \sum_{s=1-N}^{-1} \sum_{n_1=0}^{s+N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s)$$
$$\times e^{-\jmath(2\pi/N)(n_1 k_1 + n_1 k_2 - s k_2)} \quad (S_3)$$

$$(31)$$

and, furthermore

$$Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$$
$$= \frac{1}{N} \sum_{n_1=0}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(0) e^{-\jmath(2\pi/N)n_1(k_1 + k_2)}$$
$$+ \frac{1}{N} \sum_{s=1}^{N-1} \sum_{t=0}^{N-1-s} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) e^{-\jmath(2\pi/N)(t k_1 + s k_1 + t k_2)}$$
$$+ \frac{1}{N} \sum_{s=1}^{N-1} \sum_{n_1=0}^{N-1-s} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(-s) e^{-\jmath(2\pi/N)(n_1 k_1 + n_1 k_2 + s k_2)}$$

where the index change $t = n_1 - s$ has been performed for term (S_2) in (31), such that its summation over n_1 is replaced by a summation over t, and in term (S_3) of (31), s has been replaced by $-s$. Writing t instead of n_1, we obtain

$$Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)$$
$$= \frac{1}{N} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(0) \sum_{t=0}^{N-1} e^{-\jmath(2\pi/N)t(k_1 + k_2)}$$
$$+ \frac{1}{N} \sum_{s=1}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) \sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)(s k_1 + t k_1 + t k_2)}$$
$$+ \frac{1}{N} \sum_{s=1}^{N-1} R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(-s) \sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)(s k_2 + t k_1 + t k_2)}.$$

$$(32)$$

Using

$$\sum_{t=0}^{M-1} a^t = \begin{cases} M, & a = 1 \\ \dfrac{1 - a^M}{1 - a}, & a \neq 1 \end{cases} \quad (33)$$

equation (32) further simplifies to

$$Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2) = R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(0) + \sum_{s=1}^{N-1} \left(1 - \frac{s}{N}\right)$$
$$\times \left(R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) e^{-\jmath(2\pi/N)s k_1} \right.$$
$$\left. + R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(-s) e^{\jmath(2\pi/N)s k_1} \right)$$

$$(34)$$

for $k_1 + k_2 = 0$ or $k_1 + k_2 = N$, and

$$
Q_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(k_1, k_2)
$$
$$
= -\frac{\jmath e^{\jmath(2\pi/2N)(k_1+k_2)}}{2N \sin\left(\left(\frac{2\pi}{2N}\right)(k_1+k_2)\right)}
$$
$$
\times \sum_{s=1}^{N-1} \Bigg(R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) \Big(e^{-\jmath(2\pi/N)sk_1} - e^{\jmath(2\pi/N)sk_2} \Big)
$$
$$
+ R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(-s)
$$
$$
\times \Big(e^{-\jmath(2\pi/N)sk_2} - e^{\jmath(2\pi/N)sk_1} \Big) \Bigg) \tag{35}
$$

otherwise, as we will show in the following. Suppose we have

$$
k_1 + k_2 = 0 \text{ or } k_1 + k_2 = N. \tag{36}
$$

Then, according to (33)

$$
\sum_{t=0}^{N-1} e^{-\jmath(2\pi/N)t(k_1+k_2)} = N
$$
$$
\sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)(sk_1+tk_1+tk_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_1} \sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)t(k_1+k_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_1}(N-s)
$$
$$
\sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)(sk_2+tk_1+tk_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_2} \sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)t(k_1+k_2)}
$$
$$
= e^{+\jmath(2\pi/N)sk_1}(N-s)
$$

where in the last equation assumption (36) was used twice. This applied to (32) implies (34). Suppose now we have

$$
k_1 + k_2 \neq 0 \text{ and } k_1 + k_2 \neq N.
$$

Then, according to (33)

$$
\sum_{t=0}^{N-1} e^{-\jmath(2\pi/N)t(k_1+k_2)} = \frac{1 - \overbrace{e^{-\jmath(2\pi/N)N(k_1+k_2)}}^{1}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}} = 0
$$

$$
\sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)(sk_1+tk_1+tk_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_1} \sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)t(k_1+k_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_1} \frac{1 - e^{-\jmath(2\pi/N)(N-s)(k_1+k_2)}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}}
$$
$$
= e^{-\jmath(2\pi/N)sk_1} \frac{1 - e^{\jmath(2\pi/N)s(k_1+k_2)}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}}
$$
$$
= \frac{e^{-\jmath(2\pi/N)sk_1} - e^{\jmath(2\pi/N)sk_2}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}}
$$

$$
\sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)(sk_2+tk_1+tk_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_2} \sum_{t=0}^{N-1-s} e^{-\jmath(2\pi/N)t(k_1+k_2)}
$$
$$
= e^{-\jmath(2\pi/N)sk_2} \frac{1 - e^{-\jmath(2\pi/N)(N-s)(k_1+k_2)}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}}
$$
$$
= e^{-\jmath(2\pi/N)sk_2} \frac{1 - e^{\jmath(2\pi/N)s(k_1+k_2)}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}}
$$
$$
= \frac{e^{-\jmath(2\pi/N)sk_2} - e^{\jmath(2\pi/N)sk_1}}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}}
$$

which applied to (32) implies (35), if one observes the equation shown at the bottom of the page. As already mentioned, (34) and (35) together with (28) and (29) fully specify the first and second moments of the noise at the outputs of the DFTs. However, if we are not interested in correlations between different frequencies (subcarriers), we can specialize[12] (29) to $k = k_1 = k_2$, while applying (28), (34), and (35), and we obtain

$$
C_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(0,0) = P_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}(0,0)
$$
$$
= R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(0) + \sum_{s=1}^{N-1} \left(1 - \frac{s}{N}\right)
$$
$$
\times \left(R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) + R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(-s) \right)
$$
$$
- N \mu_{\mathbf{z}^{\langle l_1 \rangle}} \mu_{\mathbf{z}^{\langle l_1 \rangle}}
$$
$$
C_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}\left(\frac{N}{2}, \frac{N}{2}\right) = P_{\mathbf{w}^{\langle l_1 \rangle}, \mathbf{w}^{\langle l_2 \rangle}}\left(\frac{N}{2}, \frac{N}{2}\right)
$$
$$
= R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(0) + \sum_{s=1}^{N-1} \left(1 - \frac{s}{N}\right)(-1)^s
$$
$$
\times \left(R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(s) + R_{\mathbf{z}^{\langle l_1 \rangle}, \mathbf{z}^{\langle l_2 \rangle}}(-s) \right)
$$

[12]This is then an extension of the results of Section III to the MIMO case.

$$
\frac{1}{1 - e^{-\jmath(2\pi/N)(k_1+k_2)}} = \frac{1}{e^{-\jmath(2\pi/2N)(k_1+k_2)}\left(e^{\jmath(2\pi/2N)(k_1+k_2)} - e^{-\jmath(2\pi/2N)(k_1+k_2)}\right)}
$$
$$
= \frac{-\jmath e^{\jmath(2\pi/2N)(k_1+k_2)}2\jmath}{2\left(e^{\jmath(2\pi/2N)(k_1+k_2)} - e^{-\jmath(2\pi/2N)(k_1+k_2)}\right)}
$$
$$
= -\frac{\jmath e^{\jmath(2\pi/2N)(k_1+k_2)}}{2\sin\left(\left(\frac{2\pi}{2N}\right)(k_1+k_2)\right)}.
$$

and, for $k \neq 0$ and $k \neq (N/2)$

$$C_{\mathbf{w}\langle l_1 \rangle, \mathbf{w}\langle l_2 \rangle}(k,k) = R_{\mathbf{z}\langle l_1 \rangle, \mathbf{z}\langle l_2 \rangle}(0) + \sum_{s=1}^{N-1} \left(1 - \frac{s}{N}\right)$$
$$\times \left(R_{\mathbf{z}\langle l_1 \rangle, \mathbf{z}\langle l_2 \rangle}(s) e^{-\jmath(2\pi/N)sk} \right.$$
$$\left. + R_{\mathbf{z}\langle l_1 \rangle, \mathbf{z}\langle l_2 \rangle}(-s) e^{\jmath(2\pi/N)sk} \right)$$

$$P_{\mathbf{w}\langle l_1 \rangle, \mathbf{w}\langle l_2 \rangle}(k,k) = -\frac{e^{\jmath(2\pi/N)k}}{N \sin\left(\frac{2\pi}{N}k\right)} \sum_{s=1}^{N-1} \sin\left(\frac{2\pi}{N}sk\right)$$
$$\times \left(R_{\mathbf{z}\langle l_1 \rangle, \mathbf{z}\langle l_2 \rangle}(s) + R_{\mathbf{z}\langle l_1 \rangle, \mathbf{z}\langle l_2 \rangle}(-s) \right).$$

Note that we have expressed the first and second moments of the noise at the outputs of the DFTs in terms of the specifying parameters of the noise at the inputs of the receivers. Any conventional algorithm that makes use of statistical properties of the noise at the outputs of the DFTs (or, equivalently, at the inputs of the decision devices) can be potentially improved when knowledge given by the pseudocross-covariance matrices is exploited. Consider, e.g., the calculation of (MIMO) time–domain equalizer coefficients: Any minimum mean square error (MMSE) approach that takes into account nonvanishing pseudo(cross)-covariance matrices will have a better performance than a conventional MMSE algorithm that ignores the information available from the pseudo(cross)-covariance matrices. We also want to emphasize that the overall performance can again be improved if rotated rectangular constellations are used instead of the conventional quadratic constellations. We cannot state the optimum design parameters explicitly, since they depend on the usually applied MIMO algorithms as well. Consider, e.g., a linear MIMO equalizer that processes the outputs of the DFTs. It obviously transforms the noise (pseudo)(cross)-covariance matrices into other (pseudo)(cross)-covariance matrices, but this transformation can be expressed by means of matrix multiplications that depend on the equalizer parameters. Hence, we can again calculate covariances and pseudocovariances of the complex noise variables at the input of the decision devices, so that we can apply Theorem 2 to obtain the noise rotations and the noise variances after the rotations. From these, the optimum constellation parameters can be derived as in Section VI. For theoretical capacity results that apply to general MIMO channels and involve pseudo(cross)-covariance matrices, we refer to the literature [26], [28], [29].

VIII. CONCLUSION

We considered a DMT transmission system and analyzed the noise at the output of the DFT and at input of the decision device. We showed that the noise vector at the output of the DFT (and, consequently, at input of the decision device) is rotationally variant in general. We calculated the corresponding covariance matrix and pseudocovariance matrix, which were then specialized in order to obtain the noise variances of real and imaginary parts and to obtain the correlations between real and imaginary parts for a fixed frequency/subcarrier. Via eigenvalue decompositions, we were able to determine the eccentric-

ities and the rotations of the noise ellipses. It turned out that the rotation angles are *independent* of the actual noise characteristics. They only depend on the number of the considered subcarrier. We presented a simple criterion for the occurrence of dependencies and/or variance differences which only requires the knowledge of the autocorrelation function of the random noise process at the input of the receiver. According to this criterion, different noise variances and correlations of real and imaginary parts do not occur in the presence of white noise (at the input of the receiver). For colored noise, they do occur, and one has to use rotated rectangular constellations instead of the common (square) QAM constellations. Otherwise, one has to accept a capacity loss and increased symbol error probability. We calculated both quantities and found that the impact on capacity is not very substantial due to the high SNR in wireline transmission. On the other hand, the loss measured by (uncoded) symbol error probability can be quite large (sometimes even dramatic), so that we can expect enough benefit to afford the additional effort required for implementation. We showed how to modify the existing bit-loading algorithms in order to obtain the optimum constellation parameters. Furthermore, we provided simulation plots, which confirmed the theoretical results and demonstrated the relevance of this work. Finally, we extended our results to the MIMO case, for which a very general noise model at the inputs of the receivers was assumed, allowing correlations between the noise signals of different receivers. In order to cope with this scenario we had to extend the notion of pseudocovariance matrix to the notion of pseudocross-covariance matrix, so that we were again able to derive analytic expressions for the first and second moments of the noise at the outputs of the DFTs.

APPENDIX

We have to minimize the function

$$F\left(S_1(k), S_2(k), M_1(k), M_2(k)\right)$$
$$= 2Q\left(\sqrt{\frac{3S_1(k)}{(M_1(k))^2 \nu_1(k)}}\right) + 2Q\left(\sqrt{\frac{3S_2(k)}{(M_2(k))^2 \nu_2(k)}}\right) \tag{37}$$

subject to

$$S_1(k) + S_2(k) = S_{\text{all}}(k)$$
$$M_1(k)M_2(k) = M_{\text{all}}(k)$$
$$S_i(k) > 0, \qquad i = 1,2$$
$$M_i(k) > 0 \text{ and even}, \qquad i = 1,2. \tag{38}$$

Let us consider the function

$$f(x,y) = 2Q\left(\sqrt{xy}\right) + 2Q\left(\sqrt{(x_0 - x)\frac{y_0}{y}}\right)$$

and observe that

$$F\left(S_1(k), S_{\text{all}}(k) - S_1(k), M_1(k), \frac{M_{\text{all}}(k)}{M_1(k)}\right) = f(x,y)$$

if

$$x = S_1(k)$$
$$y = \frac{3}{(M_1(k))^2 \nu_1(k)}$$
$$x_0 = S_{\text{all}}(k)$$
$$y_0 = \frac{9}{(M_{\text{all}}(k))^2 \nu_1(k)\nu_2(k)} \qquad (39)$$

such that we can find the minimum of $F(\cdot, \cdot, \cdot, \cdot)$ (under the required side constraints) by minimizing $f(\cdot, \cdot)$. Note that y is regarded as a continuous, real-valued parameter during the minimization process; this conflicts with the requirement that $M_1(k)$ is an even number. However, we will stick to this simplification and will fulfill the requirement by rounding to an even number after having found the "continuous minimum." Using

$$\frac{d}{dx}\left(Q\left(\sqrt{x}\right)\right) = -\frac{1}{2\sqrt{2\pi}}\frac{e^{-x/2}}{\sqrt{x}}$$

we obtain

$$\frac{\partial}{\partial x}f(x,y)$$
$$= -\frac{1}{\sqrt{2\pi}}\left(\frac{e^{-xy/2}}{\sqrt{xy}}y - \frac{e^{-((x_0-x)y_0/y)/2}}{\sqrt{(x_0-x)\frac{y_0}{y}}}\frac{y_0}{y}\right)$$
$$= \frac{1}{\sqrt{2\pi}}\left(\frac{e^{-(x_0-x)y_0/(2y)}}{\sqrt{(x_0-x)\frac{y}{y_0}}} - e^{-xy/2}\sqrt{\frac{y}{x}}\right)$$

$$\frac{\partial}{\partial y}f(x,y)$$
$$= -\frac{1}{\sqrt{2\pi}}\left(\frac{e^{-xy/2}}{\sqrt{xy}}x - \frac{e^{-((x_0-x)y_0/y)/2}}{\sqrt{(x_0-x)\frac{y_0}{y}}}(x_0-x)\frac{y_0}{y^2}\right)$$
$$= \frac{1}{\sqrt{2\pi}}\left(e^{-(x_0-x)y_0/(2y)}\sqrt{(x_0-x)\frac{y_0}{y^3}} - e^{-xy/2}\sqrt{\frac{x}{y}}\right).$$

Setting

$$\frac{\partial}{\partial x}f(x,y) = f_x(x,y) = 0$$
$$\frac{\partial}{\partial y}f(x,y) = f_y(x,y) = 0 \qquad (40)$$

yields

$$\frac{e^{-xy/2}}{\sqrt{xy}}\frac{\sqrt{(x_0-x)\frac{y_0}{y}}}{e^{-((x_0-x)y_0/y)/2}} = \frac{y_0}{y^2}$$
$$\frac{e^{-xy/2}}{\sqrt{xy}}\frac{\sqrt{(x_0-x)\frac{y_0}{y}}}{e^{-((x_0-x)y_0/y)/2}} = \frac{x_0-x}{x}\frac{y_0}{y^2} \qquad (41)$$

which implies

$$\frac{x_0 - x}{x} = 1$$

and, furthermore

$$x = \frac{x_0}{2}. \qquad (42)$$

Substituting this back into (41), we obtain the following for y:

$$\frac{y_0}{y}e^{-(x_0/2)(y_0/y)} = ye^{-(x_0/2)y} \qquad (43)$$

which has the solution

$$y = \sqrt{y_0}. \qquad (44)$$

Note that it can happen that this solution of (43) is not the unique solution. The question whether it is unique or not depends on the parameters x_0 and y_0. To see this, observe that (43) can be written as

$$g_{x_0}\left(\frac{y_0}{y}\right) = g_{x_0}(y)$$

with the function

$$g_{x_0}(y) = ye^{-(x_0/2)y}$$

which is *not one-to-one* (also called *not injective*) in general. However, we will not consider possible other solutions. Instead, we will show that the solution $(x,y) = (x_0/2, \sqrt{y_0})$ is indeed a (relative) minimum (and not, e.g., a maximum of $f(x,y)$). Note that (40) is only a necessary condition for (x,y) being a relative minimum. For a sufficient condition, we consider the 2×2 matrix consisting of the partial derivatives of second order evaluated at $(x,y) = (x_0/2, \sqrt{y_0})$, i.e.,

$$\mathbf{S}\left(\frac{x_0}{2}, \sqrt{y_0}\right) = \begin{bmatrix} f_{xx}\left(\frac{x_0}{2}, \sqrt{y_0}\right) & f_{xy}\left(\frac{x_0}{2}, \sqrt{y_0}\right) \\ f_{yx}\left(\frac{x_0}{2}, \sqrt{y_0}\right) & f_{yy}\left(\frac{x_0}{2}, \sqrt{y_0}\right) \end{bmatrix}.$$

It is lengthy but straightforward to compute these partial derivatives as

$$f_{xx}\left(\frac{x_0}{2}, \sqrt{y_0}\right)$$
$$= \frac{1}{\sqrt{2\pi}}e^{-(x_0/4)\sqrt{y_0}}\sqrt{\left(\frac{2}{x_0}\right)^3\sqrt{y_0}}\left(\frac{x_0}{2}\sqrt{y_0}+1\right)$$
$$f_{xy}\left(\frac{x_0}{2}, \sqrt{y_0}\right) = f_{yx}\left(\frac{x_0}{2}, \sqrt{y_0}\right)$$
$$= \frac{1}{\sqrt{2\pi}}e^{-(x_0/4)\sqrt{y_0}}\sqrt{\frac{2}{x_0}\frac{1}{\sqrt{y_0}}}\left(\frac{x_0}{2}\sqrt{y_0}-1\right)$$
$$f_{yy}\left(\frac{x_0}{2}, \sqrt{y_0}\right)$$
$$= \frac{1}{\sqrt{2\pi}}e^{-(x_0/4)\sqrt{y_0}}\sqrt{\frac{x_0}{2}\left(\frac{1}{\sqrt{y_0}}\right)^3}\left(\frac{x_0}{2}\sqrt{y_0}-1\right) \qquad (45)$$

such that the determinant of $\mathbf{S}\left(x_0/2, \sqrt{y_0}\right)$ is obtained as

$$\det\left(\mathbf{S}\left(\frac{x_0}{2}, \sqrt{y_0}\right)\right) = \frac{1}{\pi} e^{-(x_0/2)\sqrt{y_0}} \left(1 - \frac{2}{x_0}\frac{1}{\sqrt{y_0}}\right). \quad (46)$$

It is a well-known result of *differential calculus* (see, e.g., [17]) that $\mathbf{S}\left(x_0/2, \sqrt{y_0}\right)$ being positive definite is sufficient for $(x, y) = \left(x_0/2, \sqrt{y_0}\right)$ being a relative minimum of $f(x, y)$. Note (cf., also, [17]) that $\mathbf{S}\left(x_0/2, \sqrt{y_0}\right)$ being positive definite is equivalent to

$$f_{xx}\left(\frac{x_0}{2}, \sqrt{y_0}\right) > 0 \quad (47)$$

and

$$\det\left(\mathbf{S}\left(\frac{x_0}{2}, \sqrt{y_0}\right)\right) > 0. \quad (48)$$

Comparing (47) with (45), we conclude that the first condition (47) is always fulfilled. According to (46), (48) is equivalent to

$$\frac{x_0}{2}\sqrt{y_0} > 1$$

which then [see (39)] translates to

$$\frac{3S_{\text{all}}(k)}{2M_{\text{all}}(k)\sqrt{\nu_1(k)\nu_2(k)}} > 1.$$

It is clear that this condition is usually fulfilled in practice [cf. (25)]. Note that we are operating in a range (especially in wireline transmission), where the (symbol) error probabilities are small.

Substituting (39) back into (42) and (44) and making use of the side constraints (38), we finally obtain the minimum of (37) under the side constraints (38) as

$$S_1(k) = S_2(k) = \frac{S_{\text{all}}(k)}{2}$$

and, applying (19)

$$M_1(k) = \sqrt{M_{\text{all}}(k)}\left(\frac{\nu_2(k)}{\nu_1(k)}\right)^{1/4} = \sqrt{M_{\text{all}}(k)}\left(\frac{\lambda_2(k)}{\lambda_1(k)}\right)^{1/4}$$

$$M_2(k) = \sqrt{M_{\text{all}}(k)}\left(\frac{\nu_1(k)}{\nu_2(k)}\right)^{1/4} = \sqrt{M_{\text{all}}(k)}\left(\frac{\lambda_1(k)}{\lambda_2(k)}\right)^{1/4}$$

which, of course, have to be rounded to even numbers.

ACKNOWLEDGMENT

REFERENCES

[1] A. van den Bos, "The multivariate complex normal distribution-a generalization," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 537–539, Mar. 1995.

[2] A. van den Bos, "The real-complex normal distribution," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1670–1672, Jul. 1998.

[3] W. Brown and R. Crane, "Conjugate linear filtering," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 4, pp. 462–465, Jul. 1969.

[4] S. Buzzi, M. Loops, and A. M. Tulino, "A new class of muliuser CDMA receivers based on the minimum mean-output-energy strategy," in *Proc. IEEE Int. Symp. Inf. Theory*, Naples, Italy, Jun. 2000, pp. 355–355.

[5] R. F. H. Fischer, "Mehrkanal- und Mehrträgerverfahren für die Schnelle Digitale Übertragung im Ortsanschlussleitungsnetz," Ph.D. dissertation, Inst. Inf. Transmission, Univ. Erlangen-Nuremberg, Erlangen, Germany, 1996.

[6] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution," *Ann. Math. Statist.*, vol. 34, pp. 152–176, 1963.

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[8] *Asymmetric Digital Subscriber Line ADSL Transceivers*, G.992.1, ITU-T, 1999.

[9] *Splitterless Asymmetric Digital Subscriber Line ADSL Transceivers*, G.992.2, ITU-T, 1999.

[10] *Very High Speed Digital Subscriber Line Foundation*, G.993.1, ITU-T, 2001.

[11] *Asymmetric Digital Subscriber Line ADSL Transceivers 2 (ADSL2)*, G.992.3, ITU-T, 2002.

[12] *Splitterless Asymmetric Digital Subscriber Line ADSL Transceivers 2 (Splitterless ADSL2)*, G.992.4, ITU-T, 2002.

[13] *Asymmetric Digital Subscriber Line ADSL Transceivers—Extended Bandwidth ADSL2 (ADSL2plus)*, G.992.5, ITU-T, 2003.

[14] A. Lampe, R. Schober, and W. H. Gerstacker, "Iterative multiuser detection for complex modulation schemes," in *Proc. IEEE Int. Symp. Inf. Theory*, Washington, DC, 2001, pp. 33–33.

[15] A. Lampe, "Multiuser detection and channel estimation for DS-CDMA systems," Ph.D. dissertation, Inst. Inf. Transmission, Univ. Erlangen-Nuremberg, Erlangen, Germany, 2003.

[16] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, 2nd ed. Norwell, MA: Kluwer, 1994.

[17] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 2000.

[18] F. D. Neeser, "Communication theory and coding for channels with intersymbol interference," Ph.D. dissertation, Signal Inf. Process. Lab., ETH Zurich, Zurich, Switzerland, 1993.

[19] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[20] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[21] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[22] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.

[23] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 714–725, Mar. 2003.

[24] P. J. Schreier and L. L. Scharf, "Stochastic time-frequency analysis using the analytic signal: Why the complementary distribution matters," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3071–3079, Dec. 2003.

[25] P. J. Schreier, L. L. Scharf, and C. T. Mullis, "Detection and estimation of improper complex random signals," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.

[26] G. Tauböck, "Rotationally variant complex channels," in *Proc. 23rd Symp. Inf. Theory Benelux*, Louvain-la-Neuve, Belgium, May 2002, pp. 261–268.

[27] G. Tauböck, "Noise analysis of DMT," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Francisco, CA, Dec. 2003, vol. 4, pp. 2136–2140.

[28] G. Tauböck, "On the maximum entropy theorem for complex random vectors," in *Proc. IEEE Int. Symp. Inf. Theory*, Chicago, IL, Jun./Jul. 2004, p. 41.

[29] G. Tauböck, "Wireline multiple-input/multiple-output systems," Ph.D. dissertation, Inst. Commun. Radio-Frequency Eng., Vienna Univ. Technology, Vienna, Austria, 2005.

[30] A. W. M. Van Den Enden and N. A. M. Verhoeckx, *Discrete-Time Signal Processing, an Introduction*.   Englewood Cliffs, NJ: Prentice-Hall, 1989.

[31] C. Windpassinger, "Detection and precoding for multiple input multiple output channels," Ph.D. dissertation, Inst. Inf. Transmission, Univ. Erlangen-Nuremberg, Erlangen, Germany, 2004.

[32] M. Witzke, S. Bäro, and J. Hagenauer, "Iterative detection of generalized coded MIMO signals using a widely linear detector," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Francisco, CA, Dec. 2003, vol. 4, pp. 1821–1825.

[33] M. Witzke, "Linear and widely linear filtering applied to iterative detection of generalized MIMO signals," *Ann. Telecommun.*, vol. 60, no. 1–2, pp. 147–168, Feb. 2005.

[34] R. A. Wooding, "The multivariate distribution of complex normal variables," *Biometrika*, vol. 43, pp. 212–215, 1956.

[35] Y. C. Yoon and H. Leib, "Maximizing SNR in improper complex noise and applications in CDMA," *IEEE Commun. Lett.*, vol. 1, no. 1, pp. 5–8, Jan. 1997.

Georg Tauböck (S'01–M'07) received the first diploma degree in mathematics (with highest honors) and the Dipl.-Ing. and Dr.techn. degrees (with highest honors) in electrical engineering from Vienna University of Technology, Vienna, Austria, in 1998, 1999, and 2005, respectively. He also received the diploma for studying Violoncello at the Conservatory of Vienna, Vienna, Austria, in October 2000.

From December 1999 to November 2005, he was with the ftw. Telecommunications Research Center Vienna, Austria, and since November 2005, he has been with the Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology. From July 2003 to July 2004, he actively participated in the European Union research program (FP5) project Flexible Convergence of Wireless Standards and Services (FLOWS). His research interests include wireline and wireless communications, signal processing, and information theory.