Efficient Design of Waveforms for Robust Pulse Amplitude Modulation

Timothy N. Davidson, *Member, IEEE*

Abstract—In this paper, a large and flexible set of computationally efficient algorithms is developed for the design of waveforms for pulse amplitude modulation that provide robust performance in the presence of uncertainties in the channel and noise models. Performance is measured either by a sensitivity function for threshold detection or by the mean square error of the data estimate. For uncertainties that are modeled as being deterministically bounded, robustness is measured in terms of the worst-case performance, and for uncertainties that are modeled statistically, robustness is measured in terms of the average performance. The algorithms allow efficient evaluation of the inherent tradeoffs between robustness, nominal performance, and spectral occupation in waveform design and are used to design "chip" waveforms with superior performance to those specified in recent standards for digital mobile telephony.

Index Terms—Chip waveform, code division multiaccess, multirate FIR digital filters, optimization methods, pulse amplitude modulation, robustness, signal design.

I. INTRODUCTION

IN DIGITAL communications, waveform coding is often performed by linear pulse amplitude modulation (PAM) of translated versions of a given waveform [1]. The choice of waveform critically impacts many system performance criteria and usually involves a compromise between spectral efficiency, robustness to expected channel imperfections (including noise and interference), system delay, and receiver complexity. In applications in which accurate channel and noise models are available (and satisfy certain assumptions), there are several established techniques by which a waveform can be designed [1]–[3]. For the special case of the additive white Gaussian noise (AWGN) channel, a root-Nyquist waveform [1], [3]–[9] (and references therein) is usually chosen. However, in some applications, particularly in the wireless area, the transmission environment may undergo substantial variations, and it might not be possible to obtain accurate channel and noise models. In that case, one ought to design a waveform that provides robust performance in the presence of this model uncertainty. Unfortunately, there are few design techniques that explicitly incorporate robustness to broad classes of uncertainty.[1] (Two candidate techniques [12], [13] are discussed at the end of this section.) Furthermore, it can be rather difficult to determine the extent to which the performance under the nominal channel model must be compromised in order to obtain sufficient robustness to uncertainties in that model.

The purpose of the present paper is to show that the design of a waveform that provides maximal robustness to an uncertain frequency-selective channel and an uncertain noise correlation can be formulated as a convex optimization problem from which an optimal filter can be efficiently obtained. An important implication of this result is that the inherent design tradeoffs between robustness to model uncertainties, performance under the nominal model, and spectral occupation can be efficiently evaluated. These tradeoffs are particularly important in applications in which spectral efficiency is required, but (adaptive) equalization of the (slowly-varying) channel is deemed to be too expensive. In the design examples, we will use these trade-off curves to select chip waveforms with substantially improved performance over those specified in the IS95 standard [14] and UMTS proposal [15] for code division multiple access (CDMA) digital mobile telephony.

In this paper, we consider digital signal processor (DSP)-based PAM schemes in which PAM is performed by a finite impulse response (FIR) filter [4]–[9], [16]. In such schemes, the nominal performance, robustness, and spectral occupation of the scheme can be measured by functions of the filter coefficients. Two measures of the performance of a PAM scheme are employed in this paper: a sensitivity function that leads directly to a bound on the probability of error for threshold detection and the mean square error of the data estimate (prior to detection). For uncertainties that are modeled as being deterministically bounded, robustness is measured in terms of the worst-case performance, and for uncertainties that are modeled statistically, robustness is measured in terms of the average performance. Spectral occupation is constrained by enforcing a relative spectral mask on the power spectrum of the filter output. In some cases, we will also constrain the relative power transmitted in given spectral bands. Although the derivations herein generate nominal performance and robustness objectives that confirm much of the established intuition, these criteria, and the spectral occupation criteria, involve nonconvex functions of the filter coefficients. Hence, any direct design algorithm for the optimal filter is complicated by the intricacies of dealing with potential local minima. The key observation in obtaining our

[1]For certain structured uncertainties, such as timing error [1], [8], [10]–[12], some design techniques are available.
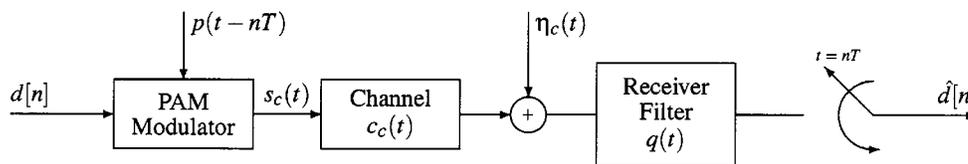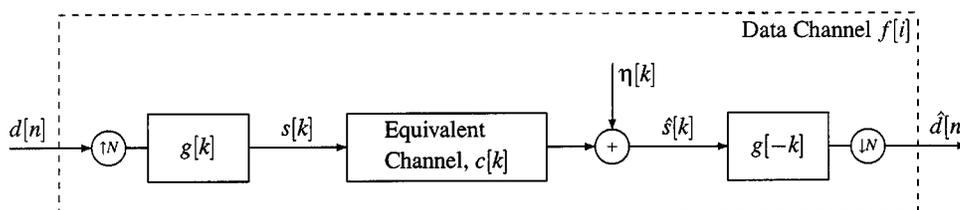
Fig. 1. Standard model for baseband PAM.

Fig. 2. Equivalent discrete-time model of baseband PAM [for real-valued $p(t)$ and $q(t) = p(-t)$].

efficient design algorithms is that the nominal performance, robustness, and spectral occupation can all be expressed as linear or convex quadratic functions of the autocorrelation coefficients of the filter. By reformulating the design criteria in terms of the autocorrelation coefficients of the filter, we obtain convex symmetric cone programs [17]–[19] that can be efficiently solved using interior point methods [19]–[21]. (The resulting design problems contain some of the previously obtained convex formulations of root-Nyquist filter design problems [8] as special cases. They are also related to some previous convex formulations of the design of general FIR filters [22], [23] and certain "signal-adapted" orthonormal multirate filterbanks [24]–[26]). Once an optimal autocorrelation function has been obtained, an optimal filter can be extracted (nonuniquely) by spectral factorization [23], [27]. A feature of the method presented in this paper is that the semi-infinite factorizability constraint is precisely transformed into a finite number of linear equality constraints on a (finite) positive semidefinite matrix.

As mentioned earlier, there are few PAM waveform design techniques that explicitly incorporate robustness to broad classes of model uncertainty. Verdú and Poor [13] considered a signal selection problem for uncertain channels and noise covariances that are deterministically bounded. Although that signal selection problem reduces to an eigenvalue problem, their technique does not lend itself to the inclusion of constraints on spectral occupation. In addition, they simply seek the optimal worst-case performance, without regard for performance under the nominal model. A design technique that, like the technique in the present paper, leads to a convex optimization problem (in that case a semi-infinite linear program) was introduced by Coleman [12]. However, his robust performance criteria are based on "eye-flattening" arguments (i.e., the derivative of the autocorrelation function of the waveform should be "small" in the neighborhood of the "zero-crossings"), and hence, his method is best suited to "small" structured uncertainties in the channel model. In this paper, we develop methods that provide robustness to more general perturbations in both the channel

and noise models and can easily incorporate constraints on nominal performance and spectral occupation.

II. BASEBAND PULSE AMPLITUDE MODULATION

Consider the standard model for baseband pulse amplitude modulation in Fig. 1. In that figure, the transmitted signal can be written as $s_c(t) = \sum_n d[n]p(t - nT)$, where

$d[n]$ nth data symbol;
$p(t)$ waveform;
T symbol interval.

If the channel and the noise correlation are known at both the transmitter and receiver (and satisfy certain assumptions), then $p(t)$ and the receiver filter $q(t)$ can be designed to jointly optimize certain performance criteria [1], [2]. (Block-based generalizations of PAM schemes [28], such as multicarrier modulation [29], are also of interest in that case.) However, if the channel and the noise correlation are not known and are deemed too costly to obtain, a standard approach [1, p. 41] is to choose $q(t) = p^*(-t)$ and to design $p(t)$ for a nominally AWGN channel. In this paper, we will improve that approach by providing efficient design algorithms for $p(t)$ that explicitly incorporate robustness to deviations from the AWGN assumption. For simplicity, we will consider only real-valued waveforms, but the methods can be extended to the complex-valued case in a straightforward manner.

If the scheme in Fig. 1 [with $p(t)$ being real-valued and $q(t) = p(-t)$] is implemented in baseband DSPs at the transmitter and receiver, then $p(t) = \sum_k g[k]\phi_s(t - kT/N)$, where $g[k]$ is the FIR filter which synthesizes $p(t)$, $\phi_s(t)$ is the impulse response of the smoothing filter in the digital-to-analog converter, and N is the oversampling rate. In that case, we can form the equivalent discrete-time model for baseband PAM shown in Fig. 2. In Fig. 2, the equivalent channel includes conversion to and from a continuous-time signal, carrier modulation and demodulation, and the physical frequency-selective (fading) channel (as in [8], for example). We will focus on scenarios in which the equivalent channel does not vary significantly (in time) over the duration

of the waveform. In that case, the received data estimate $\hat{d}[n]$ can be written as

$$\hat{d}[n] = f[0]d[n] + u[n] + \sum_k g[k - Nn]\eta[k] \qquad (1)$$

where $u[n] = \sum_{i \neq 0} f[i]d[n-i]$ is the intersymbol interference (ISI), $f[i] = \sum_k c[k]r_g[k - Ni]$ is the equivalent channel from the data perspective, $r_g[m] = \sum_k g[k]g[k + m]$ is the (deterministic) autocorrelation sequence for $g[k]$, and $\eta[k]$ models the additive noise. (Here, we have allowed finitely anticausal filtering for notational convenience.)

The spectral occupation of a PAM scheme is usually measured in terms of the (time-averaged) power spectrum of the transmitted signal. For stationary white data with zero mean and variance v_d, the power spectrum of $s_c(t)$ in Fig. 1 is $S_{s_c}(\Omega) = (v_d/T)|\Phi_s(\Omega)|^2|G(e^{j\Omega T/N})|^2$, where $\Phi_s(\Omega)$ and $G(e^{j\omega})$ are the (continuous-time) Fourier transform of $\phi_s(t)$ and the (discrete-time) Fourier transform of $g[k]$, respectively. (Here, we have used Ω and ω to represent [angular] frequency in the continuous-time and discrete-time settings, respectively.) In many communications standards, $S_{s_c}(\Omega)$ must satisfy a relative spectral mask of the form $M_{\ell,c}(\Omega) \leq S_{s_c}(\Omega)/\zeta_c \leq M_{u,c}(\Omega)$, where $M_{\ell,c}(\Omega)$ and $M_{u,c}(\Omega)$ are specified, and $\zeta_c > 0$ is a reference value. In some applications, the relative power transmitted in a given spectral band is also constrained, i.e., $\int_{\Omega_1}^{\Omega_2} S_{s_c}(\Omega)\, d\Omega \leq \rho \int_0^\infty S_{s_c}(\Omega)\, d\Omega$, for given frequencies Ω_1, $\Omega_1 \in [0, \infty)$ and a given factor ρ. Using the expression for $S_{s_c}(\Omega)$, we can transform these constraints into corresponding constraints on $g[k]$ (see, e.g., [8]): $M_\ell(e^{j\omega}) \leq |G(e^{j\omega})|^2/\zeta \leq M_u(e^{j\omega})$, and $P_g(f_1, f_2) = \int_{2\pi f_1}^{2\pi f_2} W_P(\omega)|G(e^{j\omega})|^2\, d\omega \leq \rho P_g(0, 1/2)$. The derivation of $M_\ell(e^{j\omega})$, $M_u(e^{j\omega})$, and $W_P(\omega)$ is particularly straightforward in scenarios in which $\phi_s(t)$ can be assumed to have ideal spectral characteristics: For $-\pi \leq \omega < \pi$, $M_m(e^{j\omega}) = M_{m,c}(\omega N/T)$, $m = \ell, u$, and $W_P(\omega) = 1$. For simplicity, we will make that assumption here. However, compensation for nonideal characteristics of $\phi_s(t)$ can easily be incorporated into the design methods presented herein. (See [8] for some examples of compensation in closely related design problems.) An observation that is a key element in our design method is that $R_g(e^{j\omega}) = |G(e^{j\omega})|^2$. Hence, the spectral occupation constraints involve linear functions of $r_g[m]$, but in general, quadratic functions of $g[k]$.

The fundamental performance criterion of a PAM scheme is the probability of error, which we now evaluate for the simple receiver in Fig. 2. We assume that the noise $\eta[k]$ is a zero-mean Gaussian random process, with (stochastic) autocorrelation $N_0 r_\eta[m]$, where $r_\eta[0] = 1$. If $g[k]$ is normalized so that it has unit energy ($r_g[0] = 1$), then for antipodal equally likely signaling with a transmitted signal energy per bit of E and threshold (sign) detection of $\hat{d}[n]$, the probability of error is[2]

$$P_e = \frac{1}{2^{L_f - 1}} \sum_{v=1}^{2^{L_f - 1}} \frac{1}{2}\, \mathrm{erfc}\left(\left(f[0] + \Delta^{(v)}\right)\sqrt{\frac{E}{N_0\xi}}\right) \qquad (2)$$

[2]This expression is a straightforward generalization of [1, (4.64)] to the case of possibly correlated noise. If $r_\eta[m] = \delta[m]$, then $\xi = 1$, and [1, (4.64)] is recovered.

where $\mathrm{erfc}(x) = (2/\sqrt{\pi}) \int_x^\infty \exp(-z^2)\, dz$ is the complement of the error function, $\xi = \sum_\ell r_g[\ell]r_\eta[\ell]$ denotes the noise amplification of the receiver filter, $L_f = \lfloor (2L_g + L_c - 3)/N \rfloor + 1$ is the length of $f[i]$, and $L_f - 1$ is equal to the number of interfering bits. Here

L_g length of $g[k]$;
L_c length of $c[k]$ (which may be unbounded in some cases);
$\lfloor x \rfloor$ greatest integer $\leq x$.

The term

$$\Delta^{(v)} = \sum_{i \neq 0} f[i]d^{(v)}[n - i] \qquad (3)$$

is the ISI generated by $d^{(v)}[n - i]$, $i \neq 0$, where $d^{(v)}[n - i]$ represents the vth combination of ± 1s as the $L_f - 1$ interfering bits. [Note that the symmetry of the data guarantees that for each v, there exists a v' such that $\Delta^{(v')} = -\Delta^{(v)}$.]

By using (2) as an objective, it is possible to formulate a design problem for a filter that minimizes the probability of error for a given channel or class of channels. However, the mere evaluation of that objective, let alone its application in an optimization routine, has a computational cost that is exponential in L_f. The purpose of the following two sections of the paper is to derive efficient design algorithms for pulse shaping filters that yield "small" probability of error with threshold detection. In Section V, we will derive efficient design algorithms for filters that yield minimal mean square error.

III. WORST-CASE SENSITIVITY

A natural approach to the design of a robust pulse shaping filter for threshold detection is to ensure that the intersymbol interference (ISI) term $u[n]$ in (1) is always "small" with respect to the gain of the desired symbol $f[0]$ and to ensure that the noise amplification ξ remains "close" to one. In this section, we will derive a bound on the worst-case $|u[n]|$ over a class of unknown but (deterministically) bounded channels, a bound on the worst-case $f[0]$, and a bound on the worst-case ξ over a class of unknown but bounded noise correlations. It will be shown that a filter that optimizes these bounds can be efficiently found from the solution of a convex optimization problem. It will also be shown that the designed filter directly minimizes an upper bound on the probability of error for threshold detection in (2).

A. Formulation

The first step in the derivation is to determine the worst-case value of $u[n] = \sum_{i \neq 0} f[i]d[n-i]$ over all combinations of data symbols $d[n - i]$. Using an instance of the Hölder inequality [30], we have that[3]

$$|u[n]| \leq \max_j |d[n-j]| \sum_{i \neq 0} |f[i]| = C_d \sum_{i \neq 0} |f[i]| \qquad (4)$$

where $C_d = \max_j |d[n - j]|$, and the bound is achieved by at least one combination of data symbols. The term $\sum_{i \neq 0} |f[i]|$ is

[3]Certain other instances of the Hölder inequality can also be used. For example, the Cauchy–Schwarz inequality is used in Section IV.

the "peak ISI," which was a key figure of merit in early equalization algorithms [1, p. 79].

The second step is to determine the worst-case $|u[n]|$ over a deterministically bounded set of channels. If the channel is genuinely unknown, then a neutral assertion is that the nominal channel is the ideal channel $\delta[k]$, where $\delta[k]$ is the Kronecker delta. (The case of a general nominal channel is discussed at the end of this section.) If $c_e[k] = c[k] - \delta[k]$ denotes the distorting component of the channel, then

$$f[i] = r_g[Ni] + \sum_\ell c_e[\ell + Ni] r_g[\ell]. \qquad (5)$$

The first term on the right-hand side of (5) is the ISI generated by the filter itself (the "self-ISI"), and the other term is the additional ISI generated by the distorting channel. By applying the triangle and Cauchy–Schwarz inequalities [30] to (4) and (5), we have that

$$\sum_{i \neq 0} |f[i]| \leq \sum_{i \neq 0} |r_g[Ni]| + \left(\sum_\ell r_g[\ell]^2 \right)^{1/2}$$
$$\cdot \sum_{i \neq 0} \left(\sum_{\ell=-L_g+1}^{L_g-1} c_e[\ell + Ni]^2 \right)^{1/2}$$
$$= \beta_g + \tilde{B}_g C_{nz} \qquad (6)$$

where $\beta_g = \sum_{i \neq 0} |r_g[Ni]| = 2 \sum_{i \geq 1} |r_g[Ni]|$ is the "peak self-ISI" (i.e., the peak ISI in an ideal channel). The term $\tilde{B}_g = (r_g[0]^2 + B_g^2)^{1/2}$, where

$$B_g = \left(2 \sum_{\ell \geq 1} r_g[\ell]^2 \right)^{1/2}$$

is a sensitivity coefficient for an unknown but bounded channel. The coefficients $C_i = (\sum_{\ell=-L_g+1}^{L_g-1} c_e[\ell + Ni]^2)^{1/2}$ determine the "size" of the error channel, and it is assumed that $C_{nz} = \sum_{i \neq 0} C_i$ is bounded. (Note that this assumption does not require that $c[k]$ is FIR.) Using similar analysis, it can be shown that the gain of the desired symbol is bounded below by

$$f[0] \geq r_g[0] - \tilde{B}_g C_0. \qquad (7)$$

If the noise correlation is genuinely unknown, then a neutral assertion is that the nominal noise is white. If we define $r_{\eta,e}[m] = r_\eta[m] - \delta[m]$ to be the error in that nominal model, then, in an analogous way, the worst-case value of ξ over the class of unknown but bounded noise correlations is

$$\xi \leq r_g[0] + \tilde{B}_g V \qquad (8)$$

where $V = (\sum_{m=-L_g+1}^{L_g-1} r_{\eta,e}[m]^2)^{1/2}$.

If the size of the uncertainties in the channel and noise models are known (i.e., if C_{nz}, C_0 and V are known), then one might seek to minimize an appropriate linear combination of the bounds in (6)–(8). Although that problem can be cast as a convex optimization problem, in many wireless applications, it is unlikely that C_{nz}, C_0, and V will be known at the transmitter. For such applications, a natural design approach is to search for

a filter for which the sensitivity coefficients to the worst-case channel and the worst-case noise correlation are both "small," and the worst-case gain of the desired symbol is "large," subject to a bound on the ISI in the nominal (ideal) channel, a normalization constraint on the gain of the desired symbol over the nominal channel, and a spectral mask constraint. Under the normalization constraint ($r_g[0] = 1$), minimizing B_g will minimize the sensitivity to both the worst-case channel distortion and the worst-case noise correlation and will also maximize the worst-case gain of the desired symbol. Therefore, the design problem can be formulated as follows: *For a relative spectral mask specified by $M_\ell(e^{j\omega})$ and $M_u(e^{j\omega})$, and for some $\epsilon > 0$, N and L_g, find a filter of length L_g achieving $\min_{g[k]} B_g^2$ subject to $\sum_{k=0}^{L_g-1} g[k]^2 = 1$, $\beta_g \leq \epsilon$, and $\zeta M_\ell(e^{j\omega}) \leq |G(e^{j\omega})|^2 \leq \zeta M_u(e^{j\omega})$ for all $\omega \in [0, \pi]$ and some $\zeta > 0$, or show that none exist.* Unfortunately, B_g^2 is a quartic function of $g[k]$, and the constraint on β_g and the lower bound constraint on the power spectrum generate nonconvex quadratic constraints on $g[k]$. Therefore, any direct design algorithm for the optimal $g[k]$ is complicated by the intricacies of dealing with potential local minima. Furthermore, the constraint on β_g is not smooth, and hence, algorithms based on analytic gradients cannot be directly applied. As a result, algorithms for the direct solution of this optimization problem can be rather computationally intensive.

In contrast, B_g^2 is a convex quadratic function of $r_g[m]$, and the spectral mask generates linear constraints on $r_g[m]$. Furthermore, the bound $\beta_g \leq \epsilon$ for some $\epsilon \geq 0$ can be rewritten as a set of linear constraints with additional variables $\mu_i \geq 0$, $1 \leq i \leq \lfloor (L_g - 1)/N \rfloor$ in the following standard manner: $-\mu_i \leq r_g[Ni] \leq \mu_i$ with $\sum_{i \geq 1} \mu_i \leq \epsilon/2$ (where we have exploited the symmetry of $r_g[m]$). To complete the reformulation of the design in terms of $r_g[m]$ instead of $g[k]$, we must add the additional constraint $R_g(e^{j\omega}) \geq 0$ for all $\omega \in [0, \pi]$, which is a necessary and sufficient condition for $r_g[m]$ to be factorizable in the form $r_g[m] = \sum_k g[k]g[k + m]$ (by the Féjer–Riesz theorem). By performing this reformulation, we obtain an optimization problem with a convex quadratic objective and linear constraints (often called a quadratic program). However, the constraint $R_g(e^{j\omega}) \geq 0$ generates an infinite number of linear constraints on $r_g[m]$ because it must be satisfied for all $\omega \in [0, \pi]$. Although that constraint can be handled using discretization techniques [24], such an approach may lead to overly conservative designs and can become rather awkward numerically. As an alternative, we can apply the positive real lemma [31] to transform this semi-infinite constraint into a finite set of linear equality constraints on a (symmetric) positive semidefinite (PSD) matrix. (See [8], [22], [25], and [26] for applications of the positive real lemma in other FIR filter design problems.) A version of the positive real lemma [33][4] states that $R_g(e^{j\omega}) \geq 0$ for all $\omega \in [0, \pi]$ if and only if there exists an $L_g \times L_g$ PSD matrix X (denoted $X \succeq 0$) with trace and off-diagonal sums satisfying

$$\sum_{k=1}^{L_g-m} [X]_{k, k+m} = r_g[m] \qquad \text{for } 0 \leq m \leq L_g - 1. \qquad (9)$$

[4]This is simply the dual of the version in [8] and [31] for an FIR system.

Using this result and the symmetry of $r_g[m]$, the design problem can be reformulated as the following problem.

Problem 1: Given $M_\ell(e^{j\omega})$, $M_u(e^{j\omega})$, ϵ, N, and L_g, find a filter of length L_g achieving $\min \theta$ over $r_g[m]$, $0 \le m \le L_g - 1$, $\boldsymbol{X} \succeq 0$, $\mu_i \ge 0$, $1 \le i \le \lfloor (L_g - 1)/N \rfloor$, $\zeta > 0$, and θ, subject to $r_g[0] = 1$

$$\sum_{m \ge 1} r_g[m]^2 \le \theta \tag{10}$$

$$-\mu_i \le r_g[Ni] \le \mu_i, \qquad \text{with } \sum_{i \ge 1} \mu_i \le \epsilon/2 \tag{11}$$

$$\zeta M_\ell(e^{j\omega}) \le R_g(e^{j\omega}) \le \zeta M_u(e^{j\omega}), \quad \text{for all } \omega \in [0, \pi] \tag{12}$$

and to the linear equality constraints in (9), or show that none exist.

Problem 1 consists of a linear objective, subject to linear equality (9) and inequality [(11) and (12)] constraints, a semidefiniteness constraint on $\boldsymbol{X}$, and the constraint in (10). Since (10) can be transformed [32] into a "rotated" second-order cone [18] constraint and an additional linear constraint, Problem 1 is a convex symmetric cone program [17]–[19], and the globally optimal autocorrelation sequence can be efficiently found via interior point methods [19]–[21]. Furthermore, infeasibility of Problem 1 (i.e., where the constraints cannot be satisfied by any autocorrelation sequence of the given length) can be reliably detected. For piecewise constant (and piecewise trigonometric polynomial) mask shapes, the infinite number of linear constraints in (12) can be transformed into a finite set of linear equality constraints on a finite set of PSD matrices [33] so that the mask constraint, like the constraint $R_g(e^{j\omega}) \ge 0$, can be precisely enforced in a finite manner. For other masks, we argue that (12) is usually less "critical" than $R_g(e^{j\omega}) \ge 0$ in the sense that a filter $g[k]$ with autocorrelation $r_g[m]$ may still exist, even if the mask is violated. Hence, discretization of (12) over a sufficiently fine grid, with an appropriate "tightening" of the mask, will often suffice. A "rule of thumb" is to choose $15L_g$ uniformly spaced discretization points [23], plus any "corner" frequencies of the masks. Once the optimal autocorrelation has been found by solving Problem 1, an optimal pulse-shaping filter can be found by spectral factorization (which can be performed in several different ways [23], [27]).

Pulse-shaping filter design using Problem 1 has a number of intuitively appealing interpretations. First, the quantity that is minimized, namely B_g^2, is the mean square difference between $r_g[m]$ (normalized so that $r_g[0] = 1$) and the "ideal" autocorrelation function, $\delta[m]$. Second, using Parseval's equality, we have that

$$2\pi B_g^2 = \int_{-\pi}^{\pi} \left(|G(e^{j\omega})|^2 - 1 \right)^2 d\omega \tag{13}$$

and hence, minimizing B_g is equivalent to making $|G(e^{j\omega})|$ as flat as possible (in a mean-square sense). These interpretations suggest natural weighted designs: Minimize $\sum_{m \ge 1} \psi_m r_g[m]^2$ for some non-negative weights ψ_m, or minimize $\int_{-\pi}^{\pi} W(e^{j\omega})(|G(e^{j\omega})|^2 - 1)^2 d\omega$, for some real, non-negative, weighting function $W(e^{j\omega})$. These weighted designs can also be expressed as convex cone programs. (That

for the former can be deduced later in from Problem 2, and that for the latter appeared in [34].) Third, the constraint $\beta_g \le \epsilon$ enforces pointwise frequency-flatness constraints on $|G(e^{j\omega})|$, as we now show: By generalizing Nyquist's first criterion for ISI-free transmission [1], [3], [35], [36] to filters designed using Problem 1, it can be shown (see Appendix A) that

$$\left| |G(e^{j\omega})|^2 - N \right| \le N\beta_g + \zeta^\star \sum_{k=1}^{N-1} M_u \left(e^{j(\omega - 2\pi k/N)} \right) \tag{14}$$

where $\zeta^\star$ is the optimal value of ζ from Problem 1, and that at the "folding frequency" $f_{\text{fold}} = 1/(2N)$

$$\left| \left| G\left(e^{j\pi/N} \right) \right|^2 - \frac{N}{2} \right|$$
$$\le \frac{N\beta_g}{2} + \frac{\zeta^\star}{2} \sum_{k=2}^{N-1} M_u \left(e^{j((1-2k)\pi/N)} \right). \tag{15}$$

If the upper spectral mask $M_u(e^{j\omega})$ has a constant "stop-band level" Δ_s for all $2\pi f_s \le \omega \le \pi$ and some $f_s < 1/N$ (the mask in Fig. 5 is an example), then (14) generates the following bound on the "pass-band ripple":

$$\left| |G(e^{j\omega})|^2 - N \right|$$
$$\le N\beta_g + (N-1)\zeta^\star \Delta_s, \qquad \text{for all } |\omega| \le 2\pi(1/N - f_s). \tag{16}$$

A fourth intuitively appealing interpretation of Problem 1 is derived from the fact (see Appendix B) that in scenarios of interest, i.e., where $r_g[0] - \beta_g - \tilde{B}_g(C_0 + C_{\text{nz}}) \ge 0$ so that the worst-case eye is open [1, p. 67]), the probability of error in (2) is upper bounded by

$$P_e \le \frac{1}{4} \text{erfc} \left((r_g[0] - \tilde{B}_g C_0) \sqrt{\frac{E}{N_0(r_g[0] + \tilde{B}_g V)}} \right)$$
$$+ \frac{1}{4} \text{erfc} \left((r_g[0] - \beta_g - \tilde{B}_g C) \sqrt{\frac{E}{N_0(r_g[0] + \tilde{B}_g V)}} \right) \tag{17}$$

where $C = C_0 + C_{\text{nz}}$. Since $\text{erfc}(x)$ is a monotonically decreasing function and the arguments of $\text{erfc}(\cdot)$ in (17) are decreasing functions of $\tilde{B}_g$ and β_g (when the bound is valid), the bound in (17) is a decreasing function of $\tilde{B}_g$ and β_g. Therefore, Problem 1, which seeks to minimize B_g (and, hence, $\tilde{B}_g$) subject to $\beta_g \le \epsilon$, $r_g[0] = 1$ and the spectral mask, provides direct control over the upper bound on the probability of error in (17). Finally, it is pointed out that an alternative design strategy to that pursued in Problem 1 might be to minimize the peak self-ISI β_g subject to an upper bound on the sensitivity coefficient B_g. This problem can also be cast as a convex symmetric cone program in $r_g[m]$ and, hence, efficiently solved.

Problem 1 can be extended in a straightforward manner to the case in which the channel and noise correlations are partially known, rather than being totally unknown. The analysis is almost identical—the difference being that $c_e[k] = c[k] - c_{\text{nom}}[k]$ and $r_{\eta, e}[m] = r_\eta[m] - r_{\eta, \text{nom}}[m]$, where $c_{\text{nom}}[k]$ and $r_{\eta, \text{nom}}[m]$ represent the nominal models for the channel
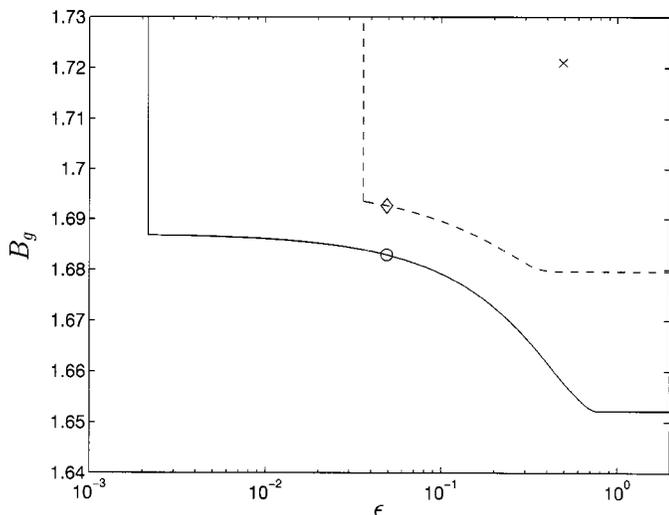
Fig. 3. Tradeoffs between the minimal value of B_g (linear scale) and ϵ (the bound on β_g, logarithmic scale) for the specified (solid) and achieved (dashed) spectral masks from the IS95 scheme (see Example 1). All points on or above the appropriate curve can be achieved, but no point below the curve can be achieved by a length 48 filter. Legend—$\times$: IS95 filter; $\circ$, $\diamond$: typical robust filters used to generate Figs. 4–6.

Fig. 4. Detail of the autocorrelation sequences for the designed ("$\circ$") and IS95 ("$\times$") filters from Example 1. For visual clarity, the sequences have been interpolated using an ideal ("sinc-function") interpolator.

and noise correlation, respectively. Once again, the design objective is to optimize the worst-case performance, subject to a bound on the nominal performance and the spectral occupation. Although the resulting design problem is also a convex cone program, the receiver structure in Fig. 2 no longer resembles the optimal receiver structure for the nominal channel.[5] Since the major application considered in this paper is chip waveform design for CDMA, and since it is unlikely in that application that the equivalent discrete-time channel (at N samples per chip) will be even partially known at the transmitter, we will focus on the case where the channel and noise correlations are genuinely unknown. Extensions of the principles of this paper to more general communication schemes that incorporate partial knowledge of the environment are currently being explored [38]. These schemes include both PAM-based schemes and block-PAM based schemes, such as discrete multitone modulation (DMT) [29] and its generalizations [28].

B. Application: Tradeoffs in Chip Waveform Design

We now show how Problem 1 can be used to efficiently evaluate some of the tradeoffs in the design of chip waveforms for CDMA-based digital mobile telephony. Once the tradeoffs have been evaluated, we will select chip waveforms with improved performance over those specified in the IS95 standard [14] and the UMTS proposal [15].

Example 1: The filter specified for the synthesis of the chip waveform in IS95 [14] has $N = 4$ and $L_g = 48$. While that filter satisfies the spectral mask specified in the standard, it has rather large values of β_g and B_g. To determine whether these values can be improved upon, Problem 1 was solved for a range of values of ϵ, subject to the IS95 spectral mask. (Each instance of Problem 1 was solved in under 20 s on a 400 MHz

---

[5]Optimal PAM receiver structures for more general nominal channels appear in [1]–[3] and [37]. The receiver structure in Fig. 2 is the optimal receiver for the AWGN channel when $\beta_g = 0$.
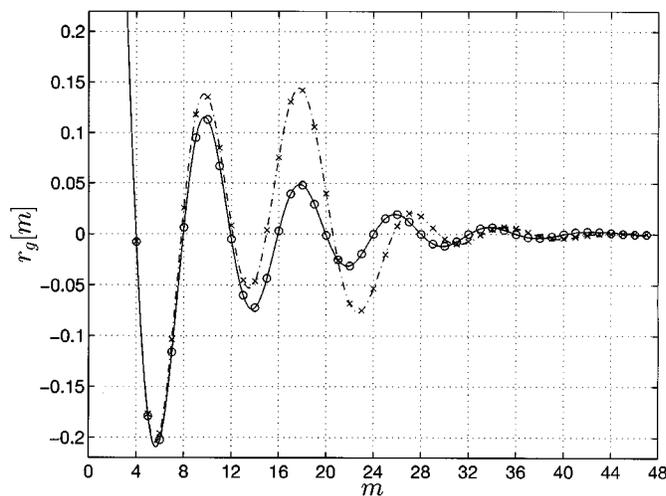
PENTIUM II workstation using a MATLAB-based general-purpose symmentric cone program solver called SeDuMi [32]. A MATLAB "m-file" that expresses Problem 1 in the input format required by SeDuMi is available from the author's web site at http://www.ece.mcmaster.ca/~davidson.) The resulting optimal values of B_g are plotted with a solid line in Fig. 3, from which it can be seen that a substantial reduction in β_g and B_g can be made, without violating the mask or increasing the filter length. The "floor effect" in Fig. 3 for large values of ϵ is due to the limit which the lowpass nature of the spectral mask (see Fig. 5) imposes on the achievable frequency flatness [see (13)]. The fact that β_g cannot be made arbitrarily small reinforces a previous result [8] that the shortest self-orthogonal filter for IS95 is of length $L_g = 51$.

The autocorrelation of a typical optimal filter for the IS95 mask (more precisely, one with $\epsilon = \beta_{IS95}/10$) is shown in Fig. 4, along with that of the IS95 filter. The improved "zero-crossing" behavior of the designed autocorrelation enforced by the constraint on β_g is evident from that figure, as are the smaller deviations from zero between the zero crossings induced by the minimization of B_g. The power spectra of the designed and IS95 filters are shown in Fig. 5, from which the improved frequency-flatness in the passband of the designed filter is clear. It is also clear from that figure that the IS95 filter satisfies the specified spectral mask by a considerable margin. The tradeoff between B_g and ϵ for the spectral mask *achieved* by the IS95 filter is shown by the dashed line in Fig. 3, and the power spectrum of a typical optimal filter for the achieved mask (again, one with $\epsilon = \beta_{IS95}/10$) is shown in Fig. 5(c). The autocorrelation of this optimal filter is very close to that of the optimal filter for the specified mask at the scale of Fig. 4 and has been omitted for clarity.

To demonstrate the improved performance of the robust filters, we simulated the "chip error rate" (CER) for transmission of binary chips over a slowly varying Rician-like channel with additive white Gaussian noise and sign detection of the chips at the receiver. The linear time-invariant "snap shots" of the channel were of length 33 and, hence, extend over eight chips. They were generated with $c[0] = 1$, with the remaining $c[k]$
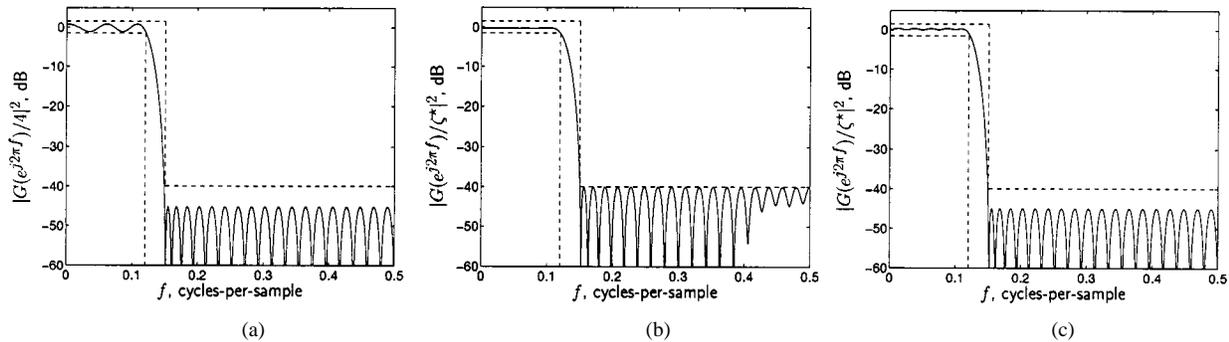
Fig. 5.   Relative power spectra (in decibels) of the filters in Example 1 with the IS95 spectral mask. Here, $\zeta^\star$ is the optimal value of ζ from Problem 1.
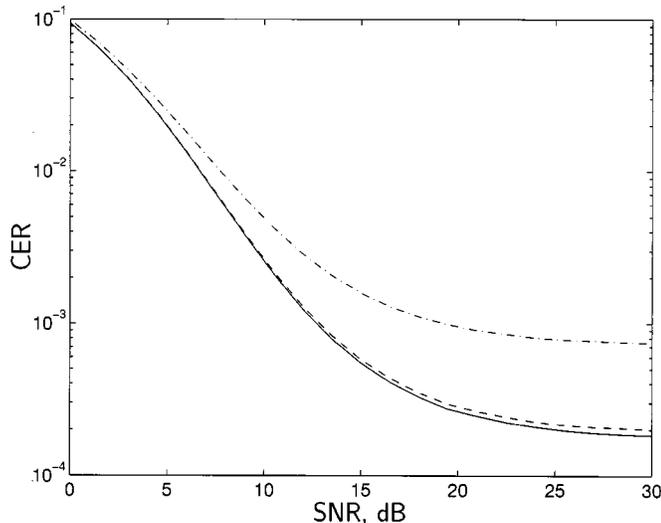

Fig. 6.   Simulated chip error rates (CER) against signal-to-noise ratio (SNR) for Example 1. Legend—Dash-dot: IS95; Solid: robust, specified mask; Dashed: robust, achieved mask.

being real, independent, and Gaussian with zero mean and standard deviation 0.05. (Such channels exhibit a wide variety of frequency selective effects.) The resulting CER curves, averaged over 100 000 channel realizations, are plotted in Fig. 6, from which the improved performance of the robust filters in Fig. 5(b) and (c) is evident. [The signal-to-noise-ratio (SNR) is defined as the ratio of the transmitted signal energy per bit to the receiver noise variance, i.e., E/N_0 in (2).]   □

In Example 1, we used Problem 1 to efficiently determine the tradeoff between robustness and self-ISI, subject to a spectral mask. While this is a key design tradeoff, one might also be interested in other tradeoffs. For instance, efficient use of the electromagnetic spectrum requires that the frequency spacing between spectrally adjacent communication schemes that operate independently of each other be kept as small as possible. The required frequency spacing depends on the shape of the spectral mask, the edge of the stopband or transition band, and on the sidelobe level. For a number of lowpass mask shapes, we can efficiently obtain the tradeoff between the sidelobe level and the band edge, subject to bounds on the robustness coefficient and the self-ISI, by using a bisection-based search for the feasibility boundary of a convex cone feasibility problem.[6] The feasibility

problem that is evaluated at each stage of the bisection search is a modified version of Problem 1 in which θ is fixed to the given bound. This bisection-based technique is similar to that recently outlined for self-orthogonal filters [8], and exploits the fact that the infeasibility of Problem 1 can be reliably detected. We now demonstrate the effectiveness of this technique by designing an improved chip waveform for the UMTS proposal [15].

Example 2: The UMTS proposal [15] specifies that a root-raised cosine (RRC) waveform[7] with a roll-off factor $\alpha = 0.22$ be used as the chip waveform. The power spectrum of an $N = 4$, $L_g = 49$ implementation of such a filter is illustrated in Fig. 7(a), along with a spectral mask chosen to tightly bound the spectrum. The $L_g = 49$ implementation was chosen because a fortuitous combination of sampling and truncation effects results in a rapid spectral decay. In fact, the spectral decay of the $L_g = 49$ filter is much faster than that of the filters of length 48, 50, 51, and 52. From a spectral efficiency perspective, this makes $L_g = 49$ a good choice amongst the RRC filters with $\alpha = 0.22$. However, a disadvantage of the RRC family of filters is that for a fixed filter length, both the transition-band edge, which is indicated by f_t in Fig. 8(a), and the height of the first sidelobe, which is indicated by D_t in Fig. 8(a), are determined by the roll-off factor α and, therefore, cannot be controlled independently. To determine whether the resulting tradeoff between D_t and f_t can be improved upon, without compromising the desirable values of β_g and B_g achieved by the RRC filters, we consider the following convex cone feasibility problem: *Given f_t, D_t, N, and L_g, find a filter satisfying the constraints of Problem 1 with $\epsilon = \beta_{\mathrm{UMTS}}$, $B_g \leq B_{\mathrm{UMTS}}$, and the parameterized spectral mask in Fig. 8(a), or show that none exist.* For a fixed f_t, a filter exists if $D_t \geq D_t^\star$, and none exists if $D_t < D_t^\star$, where $D_t^\star$ is the minimal sidelobe level. Hence, $D_t^\star$ can be efficiently found using a bisection search on D_t for the feasibility boundary of above problem. By repeating this process for different values of f_t, we obtain the tradeoff illustrated in Fig. 8(b), from which it is clear that a substantial improvement can be made over the RRC filter. (Each instance of the feasibility problem was evaluated in around 20 s.) The power spectrum of a filter with the minimum sidelobe level and the same transition-band edge as the RRC filter is plotted in Fig. 7(b), from which the reduction of around 6.8 dB in the maximum sidelobe level can be seen. Since the values of β_g and B_g of the optimal filter are constrained to be less than or equal to those of the RRC filter, one

---

[6]For certain special mask shapes, the tradeoff can be evaluated even more efficiently via convex optimization.
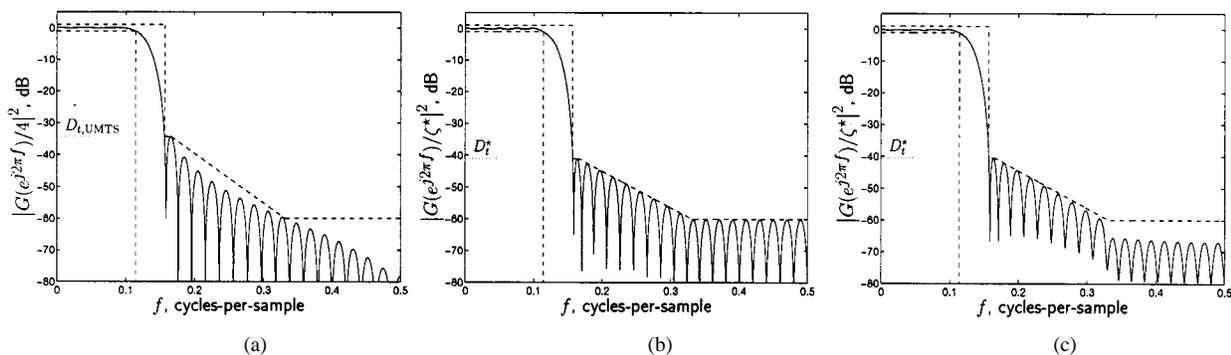
[7]That is, a waveform with an RRC magnitude spectrum.

Fig. 7. Relative power spectra (in decibels) for the filters in Example 2, along with their respective spectral masks. Note that in (b), $D_{t,\text{UMTS}} - D_t^\star \approx 6.8$ dB, and in (c), $D_{t,\text{UMTS}} - D_t^\star \approx 6.3$ dB.
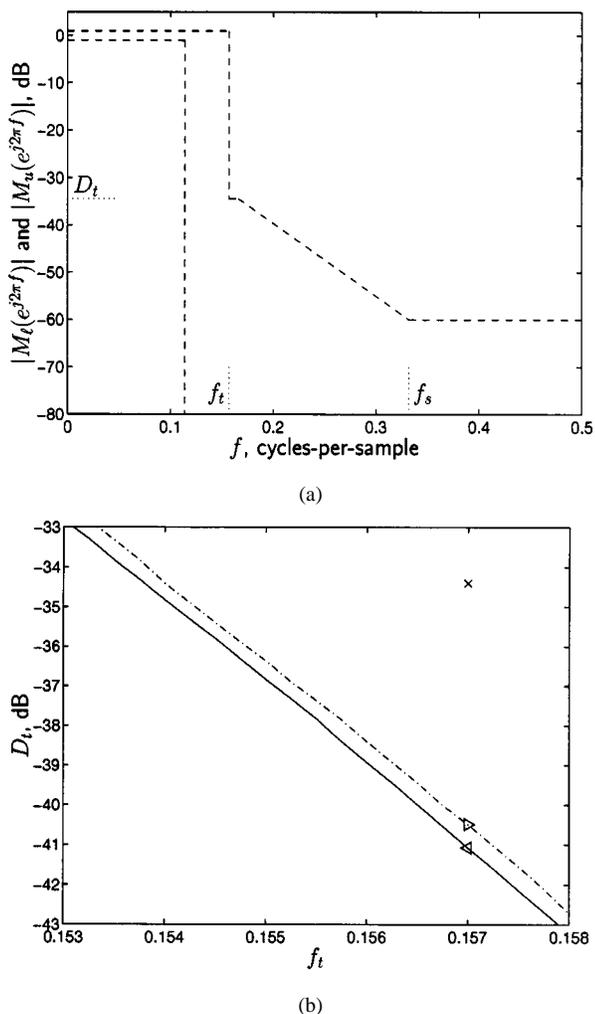

(a)

(b)

Fig. 8. (a) Parameterization of the spectral mask and (b) the tradeoffs between the minimum value of the sidelobe level D_t and the transition-band edge f_t for Example 2. Legend for (b)—Solid: direct tradeoff; Dash-dot: tradeoff with additional stopband energy constraint; $\times$: RRC filter; $\triangleleft$, $\triangleright$: optimal filters in Fig. 7(b) and (c), respectively.

would expect the CER performance of the optimal and RRC filters in slowly varying frequency-selective fading environments to be similar. Simulations results (which are not reported here) have shown that under the channel model used in Example 1, the CER curves for these optimal and RRC filters are almost indistinguishable at the scale of Fig. 6.

It is clear from Fig. 7(b) that the reduction in the sidelobe level of the optimal filter has been achieved at the expense of an increase in the power transmitted in the "stopband" $P_g(f_s, 1/2)$, where f_s is illustrated in Fig. 8(a). Fortunately, $P_g(f_1, f_2)$ is a linear function of $r_g[m]$, [8], and therefore, we can add the constraint that the designed filter transmits no more power in the stopband than the UMTS filter [i.e., $P_g(f_s, 1/2) \le P_{\text{UMTS}}(f_s, 1/2)$] while retaining the efficiency of the design technique. [The power transmitted by the designed filters in the "transition band," namely $P_g(f_t, f_s)$, is lower than that for the RRC filter.] The resulting tradeoff is also shown in Fig. 8(b). For the representative filter in Fig. 7(c), the maximum sidelobe level has been reduced by around 6.3 dB from that of the RRC filter. □

In addition to determining the "smallest" spectral mask that can be achieved without degradation in B_g or β_g, one might also be interested in determining how much "smaller" the spectral mask can be made if the value of β_g or B_g is allowed to degrade. Once again, this tradeoff can be efficiently evaluated using a sequence of convex cone feasibility problems, as we now illustrate.

Example 3: Let f_p and f_s denote the passband and stopband edges, respectively, of the IS95 spectral mask illustrated in Fig. 5. For a given value of ϵ, the smallest f_s such that there exists a filter that satisfies the mask and has $\beta_g \le \epsilon$ can be efficiently found by a bisection search on $f_s \in [f_p, 1/2]$ for the feasibility boundary of Problem 1. The resulting tradeoffs between the spectral occupation and the self-ISI are illustrated in Fig. 9 for both the specified and achieved masks from IS95. The tradeoffs are calculated with and without the additional convex constraint $B_g \le B_{\text{IS95}}$. The floor effect in Fig. 9 for large ϵ is due to the lower bound component of the mask and the constraint $B_g \le B_{\text{IS95}}$ (where it is present). The level at which a curve flattens for small ϵ is the smallest f_s, which can be achieved by a root-Nyquist filter that satisfies the relevant mask. (These levels are greater [8] than the stopband edge specified in IS95.) Note that the constraint on B_g is inactive for small ϵ because any filter that satisfies the frequency flatness constraints imposed by $\beta_g \le \epsilon$ and the spectral mask [see (14) and (16)] also satisfies the frequency flatness constraint imposed by $B_g \le B_{\text{IS95}}$ [see (13)]. For large ϵ, the constraint on B_g (when it is applied) restricts the amount by which the power spectrum can vary in the passband of the mask and, hence, limits the stopband edge that can be achieved. □
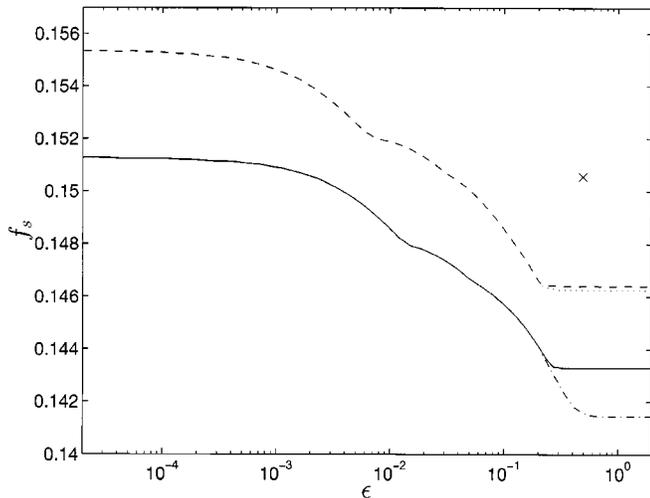
Fig. 9. Minimal stopband edge f_s against ϵ for the mask specified (solid, dash-dot) or achieved (dashed, dotted) in IS95. The solid and dashed curves include the constraint $B_g \leq B_{\mathrm{IS95}}$. The "$\times$" denotes the position of the IS95 filter.

IV. AVERAGE SENSITIVITY TO FIR CHANNELS

Although the design method proposed in Section III is applicable in a broad range of scenarios, it may be conservative in the sense of average performance in that it may provide robustness to channels and noise correlations that are unlikely to occur at the expense of sensitivity to those which are more likely to occur. If a statistical model for the channel and the noise correlation coefficients is known, or can be postulated, then an alternative design objective is to minimize an average sensitivity over the statistical model.

Using the analysis in Section III, a natural approach would be to attempt to make the expected worst-case ISI $\mathbb{E}\{\sum_{i\neq0}|f[i]|\}$ "small" with respect to the expected gain of the desired symbol $\mathbb{E}\{f[0]\}$, and to ensure that the expected noise amplification factor $\mathbb{E}\{\xi\}$ remains "close" to one. Here, $\mathbb{E}\{\cdot\}$ denotes expectation over the statistical model for the channel or the noise correlation coefficients, as appropriate. (We will model the channel coefficients and the noise correlation coefficients as being statistically independent.) Unfortunately, it appears to be difficult to incorporate the quantity $\mathbb{E}\{\sum_{i\neq0}|f[i]|\}$ into an optimization problem that can be efficiently solved. Since the goal of this paper is to obtain efficient design algorithms, we now make a slight modification to the analysis in Section III in order to obtain efficient design algorithms for filters that minimize the average sensitivity over a given statistical model.

An alternative representation of the worst-case data sequence can be obtained by applying the Cauchy–Schwarz inequality to (3), rather than the instance of the Hölder inequality that was used to obtain (4). In that case, we have that

$$u[n]^2 \leq C_{d_2}^2 \sum_{i\neq0} f[i]^2 \tag{18}$$

where $C_{d_2}^2 = \max_{d[k]} \sum_{i\neq0,\, i\in\mathrm{supp}\, f} d[n-i]^2$ and supp f denotes the support of f. (Note that for standard systems, we will need $f[i]$ and, therefore, $c[k]$, to be FIR to ensure that C_{d_2} is finite.) For a given constellation, the average value of this worst-case ISI over the statistical model of the channel is pro-

portional to $\mathbb{E}\{\sum_{i\neq0} f[i]^2\}$. By defining $\boldsymbol{r}_g$ such that $[\boldsymbol{r}_g]_m = r_g[m]$, $0 \leq m \leq L_g - 1$, and $\tilde{\boldsymbol{r}}_g = \boldsymbol{T}\boldsymbol{r}_g$, where

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{J}_{L_g-1} \\ 1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{L_g-1} \end{bmatrix} \tag{19}$$

with $\boldsymbol{I}_K$ being the $K \times K$ identity matrix, and $\boldsymbol{J}_K$ being the $K \times K$ matrix consisting of ones on the antidiagonal and zeros elsewhere, we can write $f[i] = \boldsymbol{c}_i^T \tilde{\boldsymbol{r}}_g$, where $[\boldsymbol{c}_i]_\ell = c[\ell + Ni]$. Hence, the expected value of the worst-case ISI is proportional to $\mathbb{E}\{\sum_{i\neq0} f[i]^2\} = \boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g$, where $\boldsymbol{Q}_c = \boldsymbol{T}^T \sum_{i\neq0} \mathbb{E}\{\boldsymbol{c}_i \boldsymbol{c}_i^T\} \boldsymbol{T}$. Similarly, the expected gain of the desired symbol is $\mathbb{E}\{f[0]\} = \boldsymbol{l}_c^T \boldsymbol{r}_g$, and the expected noise amplification factor is $\mathbb{E}\{\xi\} = \boldsymbol{l}_\eta^T \boldsymbol{r}_g$, where $\boldsymbol{l}_c = \boldsymbol{T}^T \mathbb{E}\{\boldsymbol{c}_0\}$, $\boldsymbol{l}_\eta = \boldsymbol{T}^T \mathbb{E}\{\tilde{\boldsymbol{r}}_\eta\}$, and $[\tilde{\boldsymbol{r}}_\eta]_m = r_\eta[m]$, $1 - L_g \leq m \leq L_g - 1$. Seeing as we wish to make both the averaged worst-case ISI and the averaged noise amplification terms small with respect to the average gain of the desired symbol, a natural optimization problem would be to minimize $\boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g + \lambda \boldsymbol{l}_\eta^T \boldsymbol{r}_g$ subject to $\boldsymbol{l}_c^T \boldsymbol{r}_g = 1$, the spectral mask, and the linear equality constraints (9) of the PSD matrix $\boldsymbol{X}$. Here, λ weights the contributions from the ISI and the noise. Although that problem can be cast as a convex cone program, our design framework is most suited to applications in which the average channel is the AWGN channel (because the receiver structure then resembles the optimal receiver structure for the average channel). In that case $\boldsymbol{l}_\eta^T \boldsymbol{r}_g = \boldsymbol{l}_c^T \boldsymbol{r}_g = r_g[0]$, and we obtain the following simplified convex cone program that is independent of λ.

Problem 2: Given $M_\ell(e^{j\omega})$, $M_u(e^{j\omega})$, $\boldsymbol{Q}_c = \boldsymbol{L}_c \boldsymbol{L}_c^T$, N, and L_g, find a filter of length L_g achieving $\min \theta_{\mathrm{av}}$ over $\boldsymbol{r}_g$, $\boldsymbol{X} \succeq \boldsymbol{0}$, $\zeta > 0$ and θ_{av}, subject to $r_g[0] = 1$, $\|\boldsymbol{L}_c^T \boldsymbol{r}_g\|_2^2 \leq \theta_{\mathrm{av}}$, the spectral mask in (12), and to the linear equality constraints in (9), or show that none exist.

Note that in Problem 2, there is no explicit differentiation between the ISI in the nominal channel and the ISI induced by deviations from that nominal channel, as there was in Problem 1. The weighting of these nominal and uncertainty-induced ISI terms is implicit in the statistical model of the channel via $\boldsymbol{Q}_c$.

While solutions to Problem 2 generate PAM schemes that perform well under the assumed statistical model for the channel, one may wish to modify that problem to ensure that the resulting PAM scheme performs well even if the statistical channel model is inaccurate. To do so, we let $\boldsymbol{Q}_c = \bar{\boldsymbol{Q}}_c + \boldsymbol{\Delta}_Q$, where $\bar{\boldsymbol{Q}}_c$ represents the matrix $\boldsymbol{Q}_c$ for the nominal statistical model, and the symmetric matrix $\boldsymbol{\Delta}_Q$ represents the uncertainty in $\boldsymbol{Q}_c$ due to inaccuracies in the statistical model. We model this uncertainty by constraining $\boldsymbol{\Delta}_Q$ to lie in the parameterized "admissible set" $\Omega_\upsilon = \{\boldsymbol{\Delta}_Q \,|\, \boldsymbol{\Delta}_Q = \boldsymbol{\Delta}_Q^T, -\upsilon\boldsymbol{I} \leq \boldsymbol{\Delta}_Q \leq \upsilon\boldsymbol{I}\}$, where υ represents the "size" of the uncertainty. The approach we will take here is to minimize the worst-case value of $\boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g$ over all matrices $\boldsymbol{\Delta}_Q$ in the admissible set, subject to a performance degradation under the nominal statistical model of at most $100\nu\%$. That problem can be simplified by observing that $\max_{\boldsymbol{\Delta}_Q \in \Omega_\upsilon} \boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g = \boldsymbol{r}_g^T \bar{\boldsymbol{Q}}_c \boldsymbol{r}_g + \upsilon \boldsymbol{r}_g^T \boldsymbol{r}_g$ and that the performance degradation constraint can be written as $\boldsymbol{r}_g^T \bar{\boldsymbol{Q}}_c \boldsymbol{r}_g \leq (1+\nu)\theta_{\mathrm{av}}^\star$, where $\theta_{\mathrm{av}}^\star$ is the solution to Problem 2 with $\boldsymbol{Q}_c = \bar{\boldsymbol{Q}}_c$. Therefore, the design problem reduces to the following convex cone program.
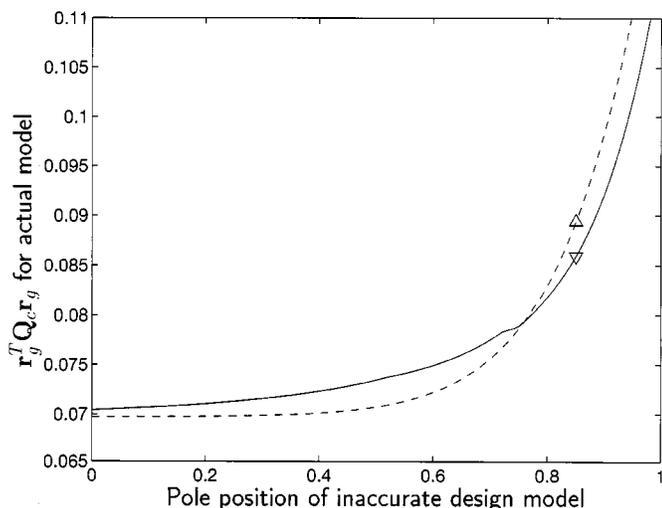
Fig. 10. Value of $\boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g$ for the actual statistical channel model against the pole position of the inaccurate statistical model used in the design of $\boldsymbol{r}_g$. (When the pole position is zero the design model is accurate.) Legend—Dashed: Original design using Problem 2; Solid: Modified design using Problem 3 with $v = 0.2$ and $\nu = 0.01$; $\triangle, \triangledown$: Designs for the inaccurate model used in Fig. 11. Note that for the actual statistical channel model the IS95 filter generates $\boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g = 0.0948$.
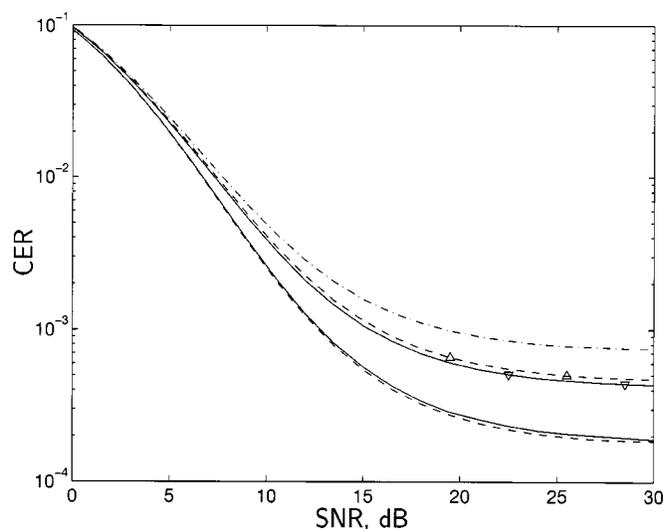

Fig. 11. Simulated chip error rates (CERs) against SNR for the accurate and inaccurate design models from Example 4. Legend—Dash-dot: IS95; Dashed: standard design using Problem 2; Solid: modified design using Problem 3 with $v = 0.2$ and $\nu = 0.01$. The triangles distinguish the case of the inaccurate design model from Fig. 10 from that of the accurate design model.

Problem 3: Given $M_\ell(e^{j\omega})$, $M_u(e^{j\omega})$, $\overline{\boldsymbol{Q}}_c = \overline{\boldsymbol{L}}_c \overline{\boldsymbol{L}}_c^T$, v, N L_g, $\theta_{\mathrm{av}}^\star$, and ν, find a filter of length L_g achieving $\min \theta_{\mathrm{av,rob}} + v\theta$ over $\boldsymbol{r}_g$, $\boldsymbol{X} \succeq \boldsymbol{0}$, $\zeta > 0$, $\theta_{\mathrm{av,rob}}$ and θ, subject to $r_g[0] = 1$, $\|\overline{\boldsymbol{L}}_c^T \boldsymbol{r}_g\|_2^2 \leq \theta_{\mathrm{av,rob}}$, $\|\boldsymbol{r}_g\|_2^2 \leq \theta$, $\theta_{\mathrm{av,rob}} \leq (1 + \nu)\theta_{\mathrm{av}}^\star$, the spectral mask in (12), and to the linear equality constraints in (9), or show that none exist.

Observe that as either v, which is the "size" of the uncertainty in $\boldsymbol{Q}_c$, or ν, which is the nominal performance degradation factor, approach zero, the solution to Problem 3 approaches that of Problem 2. We demonstrate the robustness of both design methods to inaccurate statistical channel models in the following example.

Example 4: As in Examples 1 and 3, we design chip waveforms for IS95 [14]. The actual statistical channel model in which the designs will be evaluated in the same as that used in the simulations in Example 1. That is, we have a real FIR channel of length 33 (i.e., eight chips) with $c[0] = 1$ and the remaining $c[k]$ being independent zero-mean Gaussian random variables with standard deviation 0.05. However, the statistical channel model postulated in the design has $c[0] = 1$ and the remaining $c[k]$ being from a first-order autoregressive process (of the same standard deviation) generated by passing an independent and identically distributed zero-mean Gaussian sequence with standard deviation 0.05 through the causal filter with transfer function $\sqrt{1 - z_p^2}/(1 - z_p z^{-1})$, which has a single real pole at $z = z_p$ and no zeros. (When $z_p = 0$, the postulated statistical model equals the actual statistical model.) In Fig. 10, the variation of $\boldsymbol{r}_g^T \boldsymbol{Q}_c \boldsymbol{r}_g$ in the actual statistical model is plotted against the pole position of the statistical model used in the design of $\boldsymbol{r}_g$ for both the original and modified designs (Problems 2 and 3, respectively). For visual clarity, we report results only for the specified IS95 mask—those for the achieved mask are qualitatively similar. It is evident from Fig. 10 that when the design model is substantially inaccurate, the modified design pro-

vides improved performance. The (small) performance degradation (controlled by ν) that the modified design incurs for reasonably accurate channel models (in order to obtain robustness) is also evident from the figure. The chip error rates (CERs) evaluated in the actual statistical channel model are shown in Fig. 11 for designs based on accurate and inaccurate statistical channel models. The improved robustness of the modified design and the (small) performance degradation it incurs for accurate channel models are also evident from these CER curves. $\quad\square$

An interesting connection between the design for deterministically bounded uncertainty in Section III and those for statistically modeled uncertainty in this section can be obtained by analyzing a limiting case of the modified design in Problem 3 when the mean of the nominal statistical model is the ideal channel. As the size of uncertainty v grows, the objective of Problem 3 is dominated by the term $\boldsymbol{r}_g^T \boldsymbol{r}_g = r_g[0]^2 + B_g^2/2$, and as the variance of the nominal statistical model approaches zero, $\boldsymbol{r}_g^T \overline{\boldsymbol{Q}}_c \boldsymbol{r}_g$ approaches $2\sum_{i \geq 1} r_g[Ni]^2$. In this limiting case, and with the normalization $r_g[0] = 1$, Problem 3 reduces to the minimization of B_g^2 subject to a bound on $\sum_{i \geq 1} r_g[Ni]^2$, the spectral mask, and the linear equality constraints on $\boldsymbol{X} \succeq \boldsymbol{0}$ in (9). This is similar to Problem 1—the difference being that the ISI in the nominal channel is constrained by (the square of) its two-norm rather than its one-norm. The difference is due to the different instances of the Hölder inequality used to determine the worst-case data sequence in (4) and (18). However, the resulting design tradeoffs are often qualitatively similar to those obtained in Section III, as we will illustrate in the next section.

V. MEAN SQUARE ERROR DESIGNS

In Sections III and IV, the sensitivity function was chosen with threshold (sign) detection in mind. However, in applications in which the filtered received signal is further processed before detection, the mean square error (MSE) may be a more appropriate sensitivity criterion. In this section, we show that

worst-case and average MSE design criteria also result in convex optimization problems in the autocorrelation sequence of the filter, and hence, pulse-shaping filters that provide robust performance from an MSE perspective can be efficiently obtained. The resulting design tradeoffs are often qualitatively similar to those obtained using the sensitivity functions in the previous two sections, as we will demonstrate in an example.

A. Worst-Case MSE

With reference to Fig. 2 and (1), the error in the data estimate (prior to detection) is

$$\hat{d}[n] - d[n] = \sum_i (f[i] - \delta[i])d[n-i] + \sum_k g[k - Nn]\eta[k].$$

For white data with zero mean and transmitted energy per symbol E and for noise that is independent of the data with autocorrelation $N_0 r_\eta[m]$ and $r_\eta[0] = 1$, the mean square error is

$$\text{MSE} = E \sum_i (f[i] - \delta[i])^2 + N_0 \xi \tag{20}$$

where $\xi = \sum_m r_\eta[m] r_g[m]$, as in Section III. For applications in which the channel environment is genuinely unknown, we make the neutral assertion that the nominal environment is the AWGN channel. Defining $c_e[k] = c[k] - \delta[k]$ and $r_{\eta,e}[m] = r_\eta[m] - \delta[m]$ and making the normalization $r_g[0] = 1$, we have that

$$\begin{aligned}
\text{MSE} = &E \sum_{i \neq 0} r_g[Ni]^2 \\
&+ 2E \sum_{i \neq 0} r_g[Ni] \sum_\ell c_e[\ell + Ni] r_g[\ell] \\
&+ E \sum_i \left(\sum_\ell c_e[\ell + Ni] r_g[\ell] \right)^2 \\
&+ N_0 \left(1 + \sum_m r_{\eta,e}[m] r_g[m] \right).
\end{aligned} \tag{21}$$

Here, the first term is the MSE due to the ISI in the ideal channel (the "self-induced" MSE), the second and third terms represent the MSE induced by the distorting channel, and the last term represents the MSE due to the noise. Applying the triangle and Cauchy–Schwarz inequalities to (21) [see Appendix C], we can bound the MSE by

$$\text{MSE} \leq E\left((\gamma_g + C_{\text{nz}}\tilde{B}_g)^2 + C_0\tilde{B}_g^2\right) + N_0(1 + \tilde{B}_g V) \tag{22}$$

where $\gamma_g = (2\sum_{i \geq 1} r_g[Ni]^2)^{1/2}$ is the root MSE due to ISI in the ideal channel. Here, $\tilde{B}_g$ measures the MSE sensitivity to modeling errors, and C_{nz}, C_0, and V measure the "size" of the channel and noise model errors, respectively. (These terms were defined in Section III and are assumed to be finite.)

If the signal-to-noise ratio (SNR) E/N_0 is constant and known, and if C_{nz}, C_0, and V are known, then one can formulate a convex cone program to efficiently obtain a filter that minimizes the right-hand side of (22), subject to normalization of the nominal gain of the desired symbol and the spectral mask constraints. However, in practice, the receiver noise level might not be known by the transmitter, and it is unlikely that the other
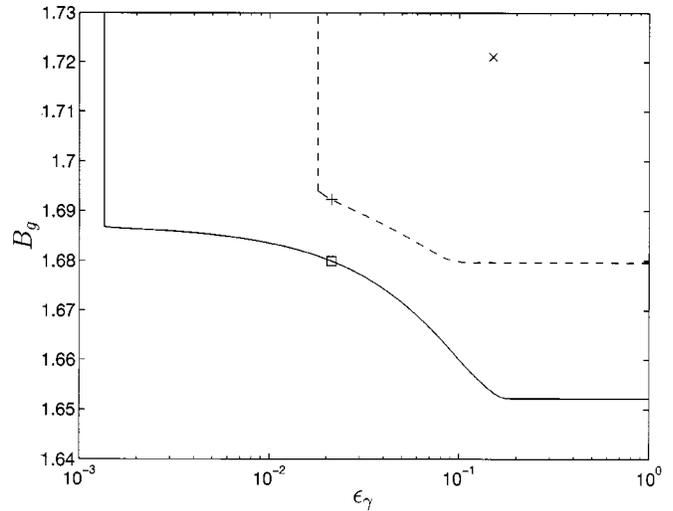

Fig. 12. Tradeoffs between the minimal value of B_g (linear scale) and ϵ_γ (the bound on γ_g, logarithmic scale) for the specified (solid) and achieved (dashed) spectral masks from the IS95 scheme (see Example 5). Legend—×: IS95 filter; □, +: typical robust filters used to generate Fig. 13.

terms will be known. An alternative approach would be to seek the minimal value of the bound in (22), subject to a bound on the self-induced MSE γ_g and the normalization $r_g[0] = 1$. Under those constraints, one only has to minimize $\tilde{B}_g^2$ in order to minimize the MSE sensitivity to modeling errors. Therefore, a filter that provides robust performance in an MSE sense in the presence of unknown but bounded channel and noise models can be efficiently obtained as the solution to the following convex cone program, which is independent of the SNR.

Problem 4: Given $M_\ell(e^{j\omega})$, $M_u(e^{j\omega})$, ϵ_γ, N, and L_g, find a filter of length L_g achieving $\min \theta_{\text{mse}}$ over $r_g[m]$, $0 \leq m \leq L_g - 1$, $\boldsymbol{X} \succeq \boldsymbol{0}$, $\zeta > 0$, and θ_{mse}, subject to $r_g[0] = 1$, $\sum_{m \geq 1} r_g[m]^2 \leq \theta_{\text{mse}}$, $(\sum_{i \geq 1} r_g[Ni]^2)^{1/2} \leq \epsilon_\gamma/\sqrt{2}$, the spectral mask in (12), and to the linear equality constraints in (9), or show that none exist.

Note that this problem is similar to Problem 1—the difference being that in Problem 4, the ISI in the nominal channel is bounded by its two-norm rather than its one-norm. (Problem 4 is equivalent to the limiting case of Problem 3 described at the end of Section IV.) As a result, the tradeoffs between robustness, nominal performance, and spectral occupation obtained for MSE performance are often qualitatively similar to those obtained for the performance measure derived in Section III, as we now illustrate. (Like the bound on β_g, the bound on γ_g places pointwise frequency-flatness constraints on $|G(e^{j\omega})|$; see Appendix A.)

Example 2: In this example, we determine the tradeoff between B_g and γ_g for both the spectral mask specified in IS95 and that achieved by the IS95 filter by solving Problem 4 for different values of ϵ_γ, with $L_g = 48$ and $N = 4$. The tradeoff is illustrated in Fig. 12. Observe the qualitative similarity to Fig. 3 and that it is once again clear that the IS95 filter is a substantial distance from providing an optimal tradeoff between B_g and γ_g. The CER curves for typical filters on these tradeoff curves (more precisely, those with $\epsilon_\gamma = \gamma_{\text{IS95}}/\sqrt{50}$) in the environment of Example 1 are qualitatively similar to those in Fig. 6, as shown in Fig. 13. □
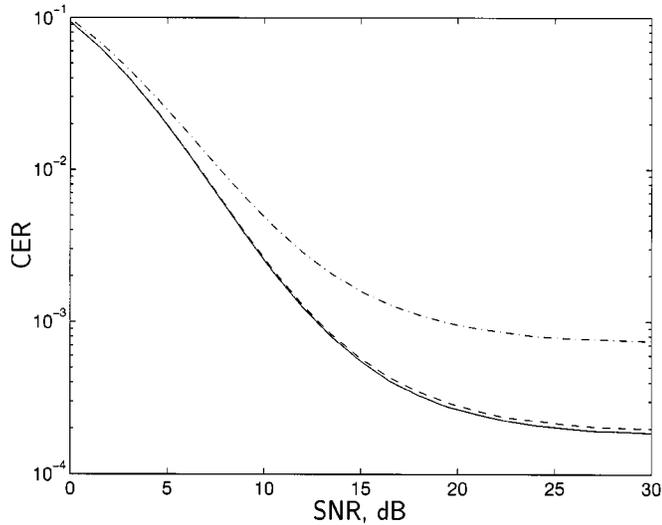
Fig. 13. Simulated CER against SNR for Example 5 for the IS95 filter (dash-dot) and the robust filters for the specified (solid) and achieved (dashed) spectral masks from Fig. 12.

B. Average MSE

If statistical models for the channel and the noise correlation coefficients are available, an alternative design objective to that in the previous section would be to minimize the average MSE over these models. By taking the expectation of (20) over the channel and the noise correlation coefficients (assuming that the channel and the noise correlation coefficients are independent), the average MSE is given by

$$
\mathbb{E}\{\text{MSE}\} = E \sum_{m, k} r_g[m] r_g[k] \sum_i \mathbb{E}\{c[m + Ni]c[k + Ni]\}
$$
$$
- 2E \sum_m r_g[m] \mathbb{E}\{c[m]\} + E
$$
$$
+ N_0 \sum_m r_g[m] \mathbb{E}\{r_\eta[m]\}
$$
$$
= E(\boldsymbol{r}_g^T \tilde{\boldsymbol{Q}}_c \boldsymbol{r}_g (N_0/E \, \boldsymbol{l}_\eta^T - 2\boldsymbol{l}_c^T)\boldsymbol{r}_g + 1)
$$

where we have used notation from Section IV and $\tilde{\boldsymbol{Q}}_c = \boldsymbol{Q}_c + \boldsymbol{T}^T \mathbb{E}\{\boldsymbol{c}_0 \boldsymbol{c}_0^T\}\boldsymbol{T}$. When the average channel is the AWGN channel, $\boldsymbol{l}_c^T \boldsymbol{r}_g = \boldsymbol{l}_\eta^T \boldsymbol{r}_g = r_g[0]$, and a filter that minimizes the average MSE, subject to a normalization of the expected gain of the desired symbol, can be efficiently found by solving the following convex cone program, which is independent of the SNR.

Problem 5: Given $M_\ell(e^{j\omega}), M_u(e^{j\omega}), \tilde{\boldsymbol{Q}}_c = \tilde{\boldsymbol{L}}_c \tilde{\boldsymbol{L}}_c^T$, N and L_g, find a filter of length L_g achieving $\min \theta_{\text{mse, av}}$ over $\boldsymbol{r}_g$, $\boldsymbol{X} \succeq \boldsymbol{0}, \zeta > 0$ and $\theta_{\text{mse, av}}$, subject to $r_g[0] = 1, \|\tilde{\boldsymbol{L}}_c^T \boldsymbol{r}_g\|_2^2 \leq \theta_{\text{mse, av}}$, the spectral mask in (12), and to the linear equality constraints in (9), or show that none exist.

This problem is similar to Problem 2, except that $\tilde{\boldsymbol{Q}}_c$ replaces $\boldsymbol{Q}_c$. Hence, robustness to inaccuracies in $\tilde{\boldsymbol{Q}}_c$ can be handled by modifying Problem 3. A special case of Problem 5 has previously been used to find an optimal PAM scheme in the presence of timing error with a known probability density function (see [8, ex. 6]).

VI. CONCLUSION

In this paper, it was shown that the inherent tradeoffs between robustness, nominal performance, and spectral occupation that occur in the design of waveforms for pulse amplitude modulation can be efficiently evaluated via convex optimization techniques. Several of these tradeoffs were calculated for the design of "chip" waveforms for the IS95 standard [14] and the UMTS proposal [15] for code division multiple access mobile telephony and were used to select waveforms with substantially improved performance over those specified in IS95 and UMTS. These improved waveforms are directly implementable in transceivers equipped with baseband digital signal processors and have the same implementation complexity as those in the standards.

The key to the efficiency of the methods described herein was to show that the nominal performance, robustness, and spectral occupation can all be effectively measured by linear or convex quadratic functions of the autocorrelation sequence of the filter that synthesizes the waveform. In this paper, performance was measured either by a sensitivity function for threshold detection or by the mean square error, and robustness was measured in terms of the worst-case performance over a bounded uncertainty set or the average performance over a statistically modeled uncertainty set. Spectral occupation was measured via a relative spectral mask and by constraints on the relative power transmitted in given spectral bands.

Since the intersection of convex sets is itself convex, many combinations of the problems considered herein (and others in [8] and [23]) can be solved using similar techniques. These techniques can also be generalized to the complex-valued case to provide efficient methods for the joint design of the "inphase" and "quadrature" pulse shapes (as distinct from the standard practice of separate design). In addition, attempts to extend the philosophy of the current work to more sophisticated communications schemes, such that those with (nominal or adaptive) equalization, or block-based PAM transmission [28], are underway [38].

In closing, it is pointed out that the goal of this work was to obtain a large and flexible set of efficient design algorithms for waveforms that provide robustness to uncertain but linear and time-invariant channels. The design of waveforms that provide robustness to channels with significant nonlinearities appears to be more complicated, even in simplified scenarios [39]. An interesting direction for further work is to examine ways in which robustness to uncertain time-varying nonlinear channels can be incorporated into the current design framework.

APPENDIX A
GENERALIZATION OF NYQUIST'S FIRST CRITERION

Using the definition of the Fourier transform, it can be shown [35], [36] that

$$
\sum_i r_g[Ni]e^{-j\omega Ni} = \frac{1}{N} \sum_{k=0}^{N-1} R_g\left(e^{j(\omega - 2\pi k/N)}\right).
$$

Since $r_g[m]$ is symmetric and $r_g[0] = 1$, we have that

$$R_g(e^{j\omega}) - N = 2N \sum_{i \geq 0} r_g[Ni] \cos(\omega Ni)$$
$$- \sum_{k=1}^{N-1} R_g \left(e^{j(\omega - 2\pi k/N)} \right). \quad (23)$$

Applying the triangle inequality and an instance of the Hölder inequality and using the fact that $R_g(e^{j\omega}) \geq 0$, we have that

$$|R_g(e^{j\omega}) - N| \leq 2N \sum_{i \geq 0} |r_g[Ni]| + \sum_{k=1}^{N-1} R_g \left(e^{j(\omega - 2\pi k/N)} \right)$$

and, hence, (14). Equation (15) can be obtained by using the fact that $R_g(e^{-j\omega}) = R_g(e^{j\omega})$ to rewrite (23) for $\omega = \pi/N$ as

$$2R_g(e^{j\pi/N}) - N$$
$$= 2N \sum_{i \geq 0} r_g[Ni] \cos(\pi i) - \sum_{k=1}^{N-1} R_g \left(e^{j((1-2k)\pi/N)} \right)$$

and then applying the triangle and Hölder inequalities. It is pointed out that an application of the triangle and Cauchy–Schwarz inequalities to (23) generates the bound

$$\left| |G(e^{j\omega})|^2 - N \right|$$
$$\leq N\gamma_g \sqrt{2\lfloor (L_g - 1)/N \rfloor} + \zeta^\star \sum_{k=1}^{N-1} M_u \left(e^{j(\omega - 2\pi k/N)} \right)$$

where $\gamma_g = (2 \sum_{i \geq 1} r_g[Ni]^2)^{1/2}$, and $\zeta^\star$ is the optimal value of ζ. The term γ_g is used in Section V.

APPENDIX B
DERIVATION OF P_e BOUND

Using the symmetry of $\Delta^{(v)}$ in (2) with respect to v and reordering the $\Delta^{(v)}$s so that $\Delta^{(\tilde{v}+2^{L_f-2})} = -\Delta^{(\tilde{v})}$, we have that

$$P_e = \frac{1}{2^{L_f-1}} \sum_{\tilde{v}=1}^{2^{L_f-2}} P_e^{(\tilde{v})} \quad (24)$$

where

$$P_e^{(\tilde{v})} = \frac{1}{2} \sum_{s=\pm 1} \text{erfc} \left(\left(f[0] + s \left| \Delta^{(\tilde{v})} \right| \right) \sqrt{\frac{E}{N_0 \xi}} \right).$$

Since $\text{erfc}(x)$ is a monotonically decreasing function of x

$$2P_e^{(\tilde{v})} \leq \text{erfc} \left(f[0] \sqrt{\frac{E}{N_0 \xi}} \right)$$
$$+ \text{erfc} \left(\left(f[0] - \left| \Delta^{(\tilde{v})} \right| \right) \sqrt{\frac{E}{N_0 \xi}} \right). \quad (25)$$

Furthermore, since $|d[n]| = 1$, we have that

$$\left| \Delta^{(\tilde{v})} \right| \leq \sum_{i \neq 0} |f[i]| \leq \beta_g + \tilde{B}_g C_{\text{nz}}. \quad (26)$$

Substituting the lower bound on $f[0]$ in (7) and the upper bound on $|\Delta^{(\tilde{v})}|$ in (26) into the right-hand side of (25), we have that

$$2P_e^{(\tilde{v})} \leq \text{erfc} \left((r_g[0] - \tilde{B}_g C_0) \sqrt{\frac{E}{N_0 \xi}} \right)$$
$$+ \text{erfc} \left((r_g[0] - \beta_g - \tilde{B}_g C) \sqrt{\frac{E}{N_0 \xi}} \right) \quad (27)$$

for all $\tilde{v}$, where $C = C_0 + C_{\text{nz}}$. If $r_g[0] - \beta_g - \tilde{B}_g C \geq 0$ (that is, if the worst-case eye is open [1, p. 67]) then by substituting the upper bound for ξ in (8) into (27), we have that

$$2P_e^{(\tilde{v})} \leq \text{erfc} \left((r_g[0] - \tilde{B}_g C_0) \sqrt{\frac{E}{N_0(r_g[0] + \tilde{B}_g V)}} \right)$$
$$+ \text{erfc} \left((r_g[0] - \beta_g - \tilde{B}_g C) \sqrt{\frac{E}{N_0(r_g[0] + \tilde{B}_g V)}} \right). \quad (28)$$

The expression in (17) is then obtained by substituting (28) into the expression for P_e in (24). Note that if the assumption that $r_g[0] - \beta_g - \tilde{B}_g C \geq 0$ is violated, then $(1/2)\text{erfc}((r_g[0] - \beta_g - \tilde{B}_g C)\sqrt{E/(N_0 \xi)}) \geq 1/2$. In that case, the contribution of this term to the bound on the overall P_e is decreased by making ξ larger, that is, by *decreasing* the SNR.

APPENDIX C
DERIVATION OF (22)

The noise term is dealt with in the same way as (8) so that we only consider the signal terms. Using the notation of Section V-A and applying the Cauchy–Schwarz inequality, we have that

$$\left(\sum_\ell c_e[\ell + Ni] r_g[\ell] \right)^2 \leq \tilde{B}_g^2 C_i^2 \quad (29)$$

and

$$\left| \sum_{i \neq 0} r_g[Ni] \sum_\ell c_e[\ell + Ni] r_g[\ell] \right|$$
$$\leq \gamma_g \left(\sum_i (c_e[\ell + Ni] r_g[\ell])^2 \right)^{1/2}$$
$$\leq \gamma_g \tilde{B}_g \left(\sum_{i \neq 0} C_i^2 \right)^{1/2}. \quad (30)$$

Taking the absolute value of the right-hand side of (21) and applying the triangle inequality and (29) and (30), MSE $\leq N_0 + E(\gamma_g^2 + 2\gamma_g \tilde{B}_g C_{\text{nz}} + \tilde{B}_g^2(C_{\text{nz}}^2 + C_0^2))$ and, hence, (22).

REFERENCES

[1] R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communications*. New York: McGraw-Hill, 1968.
[2] T. Berger and D. W. Tufts, "Optimum pulse amplitude modulation Part I: Transmitter-receiver design and bounds from information theory," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 196–208, Apr. 1967.

[3] J. G. Proakis, *Digital Communications*, 4th ed. Boston, MA: McGraw-Hill, 2001.

[4] K. H. Mueller, "A new approach to optimum pulse shaping in sampled systems using time-domain filtering," *Bell Syst. Tech. J.*, vol. 52, no. 5, pp. 723–729, May–June 1973.

[5] P. H. Halpern, "Optimal finite duration Nyquist signals," *IEEE Trans. Commun.*, vol. COMM-27, pp. 884–888, June 1979.

[6] P. R. Chevillat and G. Ungerboeck, "Optimum FIR transmitter and receiver filters for data transmission over band-limited channels," *IEEE Trans. Commun.*, vol. COMM-30, pp. 1909–1915, Aug. 1982.

[7] H. Samueli, "On the design FIR digital data transmission filters with arbitrary magnitude specifications," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1563–1567, Dec. 1991.

[8] T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Design of orthogonal pulse shapes for communications via semidefinite programming," *IEEE Trans. Signal Processing*, vol. 48, pp. 1433–1445, May 2000.

[9] M. Liu, C. J. Zarowski, and F. W. Fairman, "Factorable FIR Nyquist filters with least stopband energy under sidelobe level constraints," *IEEE Trans. Signal Processing*, vol. 48, pp. 1495–1498, May 2000.

[10] D. W. Tufts and T. Berger, "Optimum pulse amplitude modulation Part II: Inclusion of timing jitter," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 209–216, Apr. 1967.

[11] B. Farhang-Boroujeny and G. Mathew, "Nyquist filters with robust performance against timing jitter," *IEEE Trans. Signal Processing*, vol. 46, pp. 3427–3431, Dec. 1998.

[12] J. O. Coleman, "Linear-programming design of data-communication pulses tolerant of timing jitter or multipath," in *Proc. 5th Int. Conf. Wireless Commun.*, Calgary, AB, Canada, July 1993.

[13] S. Verdú and H. V. Poor, "Signal selection for robust matched filtering," *IEEE Trans. Commun.*, vol. COMM-31, pp. 667–670, May 1983.

[14] "Proposed EIA/TIA interim standard. Wideband spread spectrum digital cellular system dual-mode mobile station-base station compatibility standard," QUALCOMM, Inc., San Diego, CA, Tech. Rep. TR45.5, Apr. 1992.

[15] "Universal mobile telecommunications systems (UMTS); UMTS terrestrial radio access (UTRA); Concept evaluation," Euro. Telecommun. Stand. Inst., Sophia Antipolis, France, Tech. Rep. TR 101 146 version 3.0.0 (1997-12), Dec. 1997.

[16] A. Said and J. B. Anderson, "Bandwidth-efficient coded modulation with optimized linear partial-response signals," *IEEE Trans. Inform. Theory*, vol. 44, pp. 701–713, Mar. 1998.

[17] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 31, no. 1, pp. 49–95, Mar. 1996.

[18] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Lin. Algebra Applicat.*, vol. 284, no. 1–3, pp. 193–228, Nov. 1998.

[19] H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds., *Handbook of Semidefinite Programming. Theory, Algorithms, and Applications*. Dordrecht, The Netherlands: Kluwer, 2000.

[20] Y. Nesterov and A. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.

[21] Y. Ye, *Interior Point Algorithms: Theory and Analysis*. New York: Wiley, 1997.

[22] S.-P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via semidefinite programming and spectral factorization," in *Proc. IEEE Conf. Decision Contr.*, 1996, pp. 271–276.

[23] ——, "FIR filter design via spectral factorization and convex optimization," in *Applied and Computational Control, Signals, and Circuits*, B. Datta, Ed. Boston, MA: Birkhauser, May 1999, ch. 5.

[24] P. Moulin, M. Anitescu, K. O. Kortanek, and F. A. Potra, "The role of linear semi-infinite programming in signal adapted QMF bank design," *IEEE Trans. Signal Processing*, vol. 45, pp. 2160–2174, Sept. 1997.

[25] J. Tuqan and P. P. Vaidyanathan, "A state space approach to the design of globally optimal FIR energy compaction filters," *IEEE Trans. Signal Processing*, vol. 48, pp. 2822–2838, Oct. 2000.

[26] B. Dumitrescu and C. Popeea, "A low complexity SDP method for designing optimum compaction filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. I, Istanbul, Turkey, June 2000, pp. 516–519.

[27] T. N. T. Goodman, C. A. Micchelli, G. Rodriguez, and S. Seatzu, "Spectral factorization of Laurent polynomials," *Adv. Comput. Math.*, vol. 7, no. 4, pp. 429–454, 1997.

[28] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

[29] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[30] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[31] B. D. O. Anderson, K. L. Hitz, and N. D. Diem, "Recursive algorithm for spectral factorization," *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 742–750, Nov. 1974.

[32] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimiz. Methods Softw.*, vol. 11–12, pp. 625–653, 1999.

[33] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm, "Linear matrix inequality formulation of spectral mask constraints," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Salt Lake City, UT, May 2001.

[34] T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Design of robust pulse shaping filters via semidefinite programming," in *Proc. Sixth Can. Workshop Inform. Theory*, Kingston, ON, Canada, June 1999, pp. 71–74.

[35] F. Mintzer, "On half-band, third-band and Nth band FIR filters and their design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 734–738, Oct. 1982.

[36] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[37] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, Sept. 1985.

[38] J. Milanović, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Design of robust redundant precoding filter banks with zero-forcing equalizers for unknown frequency-selective channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. V, Istanbul, Turkey, June 2000, pp. 2761–2764.

[39] M. A. Landolsi and W. E. Stark, "DS-CDMA chip waveform design for minimal interference under bandwidth, phase and envelope constraints," *IEEE Trans. Commun.*, vol. 47, pp. 1737–1746, Nov. 1999.

Timothy N. Davidson (M'96) received the B.Eng. (Hons. I) degree in electronic engineering from The University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, Oxford, U.K., in 1995.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. His research interests are in signal processing, communications, and control, with current activity focused on signal processing for digital communication systems. He has held research positions at the Communications Research Laboratory, McMaster University, the Adaptive Signal Processing Laboratory, UWA, and the Australian Telecommunications Research Institute, Curtin University of Technology, Perth.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize (for "the most outstanding [UWA] graduand" in the pure and applied sciences) and the 1991 Rhodes Scholarship for Western Australia.