ORTHOGONAL PULSE SHAPE DESIGN VIA SEMIDEFINITE PROGRAMMING

Timothy N. Davidson[*] *Zhi-Quan Luo and K. Max Wong*[†]

Communications Research Laboratory,
McMaster University,
Hamilton, Ontario, L8S 4K1, Canada.

ABSTRACT

In digital communications, orthogonal pulse shapes are often used to represent message symbols for transmission through a channel. The design of such pulse shapes is formulated as a convex semidefinite programming problem, from which a globally optimal pulse shape can be efficiently found using interior point methods. The formulation is used to design filters which achieve the minimal bandwidth for a given filter length, and the minimal filter length for a given bandwidth. The effectiveness of the method is demonstrated by the design of waveforms with substantially improved performance over the 'chip' waveforms specified in recent standards for digital mobile telecommunications.

1. INTRODUCTION

One of the fundamental operations in digital communications is the representation of a message symbol by an analog waveform for transmission through a channel. The most common techniques involve linear pulse amplitude modulation (PAM) of a self-orthogonal waveform (or approximation thereof). In conventional communication systems, the available analog PAM technology has tended to restrict orthogonal waveform design to a choice between a small set of waveforms (e.g., rectangular pulses, pulses with raised-cosine power spectra). However, the increasing deployment of baseband digital signal processors has extended the class of waveforms which can be easily implemented. In such a situation, the design of the waveform can be transformed to the design of an orthogonal multi-rate discrete-time finite impulse response (FIR) filter. Unlike conventional FIR filter design objectives, which can often be formulated as optimization problems with analytic or computationally efficient solutions (e.g., convex linear programming problems), the orthogonal FIR filter design problem has a translation orthogonality constraint which is not convex. Such non-convexity makes design algorithms for orthogonal FIR filters susceptible to being trapped by local minima. To overcome this difficulty, multiple starting point techniques or the branch-and-bound approach can be used, but they typically yield only limited success.

In this paper, we reformulate the design problem in terms of the autocorrelation sequence of the filter, based on the observation that many of the desirable properties of an orthogonal pulse shape are actually properties of its autocorrelation. In that case, the translation orthogonality constraints become linear and hence convex. Once the autocorrelation sequence has been designed, the

*Email: davidson@sunspot.crl.mcmaster.ca

†Currently on leave with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong.

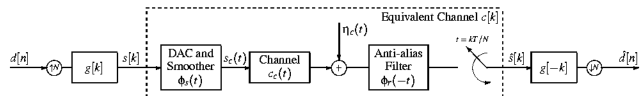


Fig. 1: A multi-rate DSP implementation of a baseband digital communications scheme.

transmission and reception filters can be extracted (non-uniquely) by spectral factorization. We use the Positive Real Lemma [3] to transform the infinite set of linear inequality constraints which is required to ensure factorizability [1] into a finite linear matrix inequality. The transformed problem is a convex semidefinite programming problem [4] whose global optimal solution can be found in an efficient manner using interior point methods. In addition we use a result in [3] to compute the minimum phase spectral factor directly. We point out that the state space transformation was first suggested in [5] in the context of conventional FIR filter design. Recently, it has been employed in the design of orthogonal energy compaction filters for signal compression [6].

In this paper, we show that apart from the design of orthogonal energy compaction filters [6], a number of additional important pulse shaping filter design problems can be cast as semidefinite programmes and hence efficiently solved. The problems considered here include finding filters achieving: a) the minimum bandwidth for a given filter length; and b) the minimum filter length required to achieve a certain bandwidth, where bandwidth is measured in terms of spectral energy concentration, or with respect to a spectral mask. The applicability of our techniques are demonstrated in several examples, including the design of: a) filters with improved performance over the root raised cosine filters; and b) 'chip waveforms' with improved performance over the chip waveform specified in IS95, the recent Code Division Multiple Access (CDMA) based mobile telephony interim standard. (Further details, some extensions and a more complete reference list are available in [7].)

2. BASEBAND PULSE AMPLITUDE MODULATION

Consider the multi-rate digital signal processing (DSP) implementation of a baseband digital communication scheme in Fig. 1. For notational convenience, we consider only real-valued systems, but the methods can be extended to the complex-valued case in a straightforward manner. The real data are waveform coded by PAM as $s_c(t) = \sum_n d[n]p(t - nT)$, where $p(t) = \sum_k g[k]\phi_s(t - kT/N)$, and the received signal is passed through the (approximately) matched filter $\sum_k g[-k]\phi_r(kT/N - t) \approx p(-t)$. Here $\phi_s(t)$ incorporates the digital to analog converter (DAC) characteristic and the smoothing filter, and $\phi_r(-t)$ is the 'anti-aliasing' filter at the receiver. Assuming that the filters $\phi_s(t)$ and $\phi_r(t)$ are sufficiently flat across the bandwidth of interest, a common design goal is to find a filter $g[k]$ which minimizes the spectral occupation of the communication scheme subject to the constraint that the filters are self-orthogonal at translations of integer multiples of N (see (1a) below). The orthogonality constraint ensures that there is no inter-symbol interference in a distortionless equivalent channel, and that the receiver filter neither amplifies nor correlates the white noise component of $\hat{s}[k]$. (Compensation for the filters $\phi_s(t)$ and $\phi_r(t)$ is discussed in [7].)

The spectral occupation of a communication scheme is usually measured in terms of its (time-averaged) power spectrum. For the simple case of stationary white data with zero mean and variance v_d, the (time-averaged) power spectrum of $s[k]$ is $S_s(e^{j2\pi f}) = v_d|G(e^{j2\pi f/N})|^2$, where $G(e^{j2\pi f})$ is the Fourier transform of $g[k]$. Due to the fact that $g[k]$ is FIR, it is only approximately bandlimited. One commonly used measure of the spectral occupation [1] is the 100γ% *energy bandwidth*, denoted B_γ, which is the smallest β such that $\int_0^\beta |G(e^{j2\pi f})|^2\,df \geq \gamma \int_0^{1/2} |G(e^{j2\pi f})|^2\,df$. For convenience, we will normalize the filter energy so that $\sum_{k=0}^{L-1} g[k]^2 = 2\int_0^{1/2} |G(e^{j2\pi f})|^2\,df = 1$.

3. A FEASIBILITY PROBLEM

In this section we introduce the fundamentals of our design framework by studying various formulations of the following simple feasibility problem for the filter in Fig. 1: *For a given γ, B, N and L, either find a filter $g[k]$ of length at most L with a 100γ% energy bandwidth less than or equal to B, or show that none exists.* This problem can be formulated directly in terms of the filter coefficients as follows:

Formulation 1 *Given γ, B, N and L, either find a filter $g[k]$ of length at most L such that*

$$\sum_{k=\ell N}^{L-1} g[k]g[k-N\ell] = \delta[\ell], \qquad \ell = 0,1,\ldots,\lfloor (L-1)/N \rfloor, \quad \text{(1a)}$$

$$\int_0^B |G(e^{j2\pi f})|^2\,df \geq \gamma/2, \quad \text{(1b)}$$

or show that none exists. Here $\lfloor x \rfloor$ denotes the greatest integer $\leq x$.

Unfortunately, both constraints in Formulation 1 are non-convex in the parameters $g[k]$. As a result, determining an answer to the above feasibility problem tends to be rather difficult. However, by using the autocorrelation of the filter $g[k]$,

$$r_g[m] \triangleq \sum_k g[k]g[k+m], \quad \text{(2)}$$

the feasibility problem can be re-cast:

Formulation 2 *Given γ, B, N and L, either find an autocorrelation sequence $r_g[m]$, $m = 1-L, 2-L, \ldots, L-1$, with $r_g[-m] = r_g[m]$, such that*

$$r_g[\ell N] = \delta[\ell], \qquad \text{for } \ell = 0,1,\ldots,\lfloor (L-1)/N \rfloor, \quad \text{(3a)}$$

$$\int_0^B R_g(e^{j2\pi f})\,df \geq \gamma/2, \quad \text{(3b)}$$

$$R_g(e^{j2\pi f}) \geq 0, \qquad \text{for all } f \in [0,1/2], \quad \text{(3c)}$$

where $R_g(e^{j2\pi f})$ is the discrete-time Fourier transform of $r_g[m]$, or show that none exists.

Formulation 2 is equivalent to Formulation 1 and since $r_g[-m] = r_g[m]$, they have the same number of variables. The additional constraint in (3c) is a necessary and sufficient condition for $r_g[m]$ to be factorizable in the form of (2). By noting that $R_g(e^{j2\pi f}) = r_g[0] + 2\sum_{m=1}^{L-1} r_g[m]\cos(2\pi m f)$, the constraints in (3b) and (3c) are clearly linear and hence Formulation 2 is a linearly constrained convex feasibility problem. However, (3c) is a semi-infinite constraint in that it must be satisfied for all values of $f \in [0,1/2]$. Such a constraint can be handled numerically (e.g., [2]), but may lead to overly conservative designs and can be rather awkward numerically.

Inspired by the work of Wu *et al.* [5] we now apply the following lemma, known as the Positive-Real Lemma (e.g., [3]), to transform the semi-infinite constraint in (3c) into a finite dimensional constraint with some auxiliary variables.

Lemma 1 (Positive-Real Lemma) *Let $H(z)$ be a real rational function with its poles (if any) inside the unit circle. Suppose that $H(\infty)$ is finite, and $H(z)$ admits a controllable and detectable state space realization $H(z) = d + \mathbf{c}(z\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b}$. Then $H(e^{j2\pi f}) + H(e^{-j2\pi f}) \geq 0$ for all $f \in \mathbb{R}$ if and only if there exists a real symmetric matrix $\boldsymbol{P}$ such that*

$$\boldsymbol{M}(\boldsymbol{P}) \triangleq \begin{bmatrix} \boldsymbol{P} - \boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A} & \mathbf{c}^T - \boldsymbol{A}^T\boldsymbol{P}\boldsymbol{b} \\ (\mathbf{c}^T - \boldsymbol{A}^T\boldsymbol{P}\boldsymbol{b})^T & 2d - \boldsymbol{b}^T\boldsymbol{P}\boldsymbol{b} \end{bmatrix} \geq 0. \quad \text{(4)}$$

Using a result of Lyapunov it can be shown that all symmetric matrices $\boldsymbol{P}$ satisfying (4) are positive semidefinite.

To apply Lemma 1, let $R_g(z) = H(z) + H(z^{-1})$, where

$$H(z) = \left[\begin{array}{cc|c} \mathbf{0} & \boldsymbol{I}_{L-2} & \mathbf{0} \\ 0 & \mathbf{0} & 1 \\ \hline r_g[L-1] & \tilde{\boldsymbol{r}}_g & 1/2 \end{array}\right], \quad \text{(5)}$$

$\tilde{\boldsymbol{r}}_g = [r_g[L-2], r_g[L-3], \ldots, r_g[1]]$, and we have used the common notation $\left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{b} \\ \hline \mathbf{c} & d \end{array}\right] \triangleq d + \mathbf{c}(z\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b}$. This realization is controllable and detectable. Formulation 2 can then be re-cast:

Formulation 3 *Given γ, B, N and L, either find $r_g[m]$, $m = 0,1,\ldots L-1$, and $\boldsymbol{P} = \boldsymbol{P}^T$ such that* (3a) *holds,*

$$\frac{1}{\pi}\sum_{m=1}^{L-1} r_g[m]\sin(2\pi m B)/m \geq \gamma/2 - B, \quad \text{(6)}$$

and (4) *holds for the realization in* (5)*, or show that none exist.*

Formulation 3 is a semidefinite feasibility problem (or linear matrix inequality system) and is equivalent to Formulations 1 and 2. It is convex and can be solved in an efficient manner using interior point methods [4].

A particular advantage of Formulation 3 is that it can be simply modified to produce the minimum phase spectral factor, implying that there is no need for further spectral factorization once we have obtained a feasible autocorrelation sequence $r_g[m]$. This is shown in the following lemma (collected from results in [3]).

Lemma 2 *Assume the same setting as Lemma 1 and that* (4) *holds for some $\boldsymbol{P} = \boldsymbol{P}^T$. Then there exists a minimal solution*

$\bar{\boldsymbol{P}}$ *to* (4)*; i.e.,* $\forall \boldsymbol{P} = \boldsymbol{P}^T$ *such that* (4) *holds,* $\boldsymbol{P} \geq \bar{\boldsymbol{P}}$*. Let* $d_w = \sqrt{2d - \boldsymbol{b}^T \bar{\boldsymbol{P}} \boldsymbol{b}}$*, and* $\boldsymbol{c}_w = (\boldsymbol{c}^T - \boldsymbol{A}^T \bar{\boldsymbol{P}} \boldsymbol{b})/d_w$*. Then* $W(z) = d_w + \boldsymbol{c}_w (z\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{b}$ *is the minimum phase spectral factor (up to a sign ambiguity) of* $H(z) + H(z^{-1})$*.*

For the realization of $H(z)$ in (5), and $\bar{\boldsymbol{P}} = \begin{bmatrix} \bar{\boldsymbol{P}}_{11} & \bar{\boldsymbol{p}}_{12} \\ \bar{\boldsymbol{p}}_{12}^T & \bar{p}_{22} \end{bmatrix}$ partitioned conformally with that realization, the minimum phase spectral factor of $R_g(z) = H(z) + H(z^{-1})$, and hence a filter which satisfies Formulation 3, is given by

$$g[k] = \begin{cases} \sqrt{1 - \bar{p}_{22}} & k = 0, \\ \left[\tilde{\boldsymbol{r}}_g - \bar{\boldsymbol{p}}_{12}^T\right]_{L-1-k} / g[0] & k = 1, 2, \ldots, L-2, \\ r_g[L-1]/g[0] & k = L-1, \end{cases}$$

where $[\boldsymbol{v}]_i$ denotes the ith element of a vector $\boldsymbol{v}$.

4. SOME ENERGY BANDWIDTH DESIGNS

A natural extension to the feasibility problem studied in the previous section is to search for the smallest bandwidth B such that there is a filter of length L with $B_\gamma \leq B$. Although a solution to this problem could be attempted using a non-convex formulation involving the filter $g[k]$, we now show that the problem can be solved in an efficient and structured manner using the convex feasibility problems in Section 3.

Problem 1 *For a given* N*,* L *and* γ*, find a filter achieving* $\min B$ *over* $r_g[m]$*,* $m = 0, 1, \ldots, L-1$*,* $\boldsymbol{P} = \boldsymbol{P}^T$ *and* B *subject to the constraints in Formulation 3.*

For a fixed value of B, Problem 1 is the convex feasibility problem in Formulation 3. Furthermore, that feasibility problem will yield a positive result for $B \geq B^*$ and a negative result for $B < B^*$, where B^* is the solution to Problem 1. Therefore, B^* can be found using a bisection search on B, where at each step we solve Formulation 3. The key to the efficiency of the method is the efficiency with which Formulation 3 can be solved. Once an autocorrelation sequence has been found for a value of $B = \hat{B}^*$ a corresponding filter can be found by several methods (i.e., spectral factorization). We will use the method suggested by Lemma 2. In order to find $\bar{\boldsymbol{P}}$, at the last step of the bisection method, we replace the semidefinite feasibility problem in Formulation 3 with the semidefinite optimization problem of minimizing trace$(\boldsymbol{P})$ subject to the constraints in Formulation 3. A property of many interior point methods for solving Formulation 3 and this minimization problem is that there is often very little difference in their computational cost. We demonstrate an application of Problem 1 in the following example.

Example 1 Here we design a filter to compete with a sampled and truncated implementation of the filter with a root raised cosine frequency response with roll-off factor $\alpha = 0.22$. (The same choice of α was made for the 'chip waveform' in the recent UMTS proposal for wideband CDMA mobile telecommunications.) We choose $N = 4$ and $L = 49$ so that the implemented filter is approximately orthogonal. That filter has a 99% energy bandwidth of $B_{0.99} \approx 0.13511$. By solving Problem 1 for the same values of N, L and γ, using `SeDuMi` [8], the minimum achievable $B_{0.99}$ was found to be ≈ 0.12914—a reduction of more than 4% over than of the root raised cosine filter. The power spectra of the two filters are plotted in Fig. 2. □

An alternative design is to find an orthogonal filter which minimizes the delay in the received data subject to a constraint on the energy bandwidth. That is:

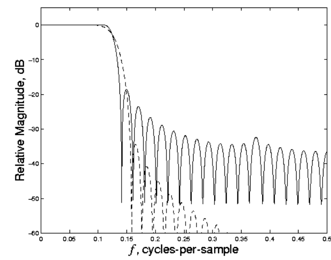


Fig. 2: Relative magnitude spectra (in decibels) of the designed (solid) and square-root raised cosine filters (dashed) in Example 1.

Problem 2 *For a given* γ*,* N *and* B*, find a filter achieving* $\min L$ *over* $r_g[m]$*,* $m = 0, 1, \ldots, L-1$*,* $\boldsymbol{P} = \boldsymbol{P}^T$ *and* $L \in \mathbb{Z}$ *subject to the constraints in Formulation 3.*

Once again we can use a bisection-based search in which Formulation 3 is solved at each step to find the minimal L. For example, the minimal length filter which achieves the same 99% energy bandwidth as the root raised cosine filter in Example 1 can be found in this way. It has $L = 31$ which represents a substantial reduction in delay and in computational requirements.

5. SOME SPECTRAL MASK DESIGNS

Although the percentage energy bandwidth is a convenient measure of spectral occupation, we lose control over the actual spectrum of the pulse shape. Since many communication standards are specified in terms of a spectral mask which the transmitted signal must satisfy, an arguably more satisfactory measure of spectral occupation would be to constrain the power spectrum to lie within a given spectral mask; i.e.,

$$M_\ell(f) \leq R_g(e^{j2\pi f}) \leq M_u(f), \qquad \text{for all } f \in [0, 1/2], \quad (7)$$

for some given mask $M_\ell(f)$ and $M_u(f)$. Filter design problems with mask constraints of the form in (7), but without the orthogonality constraint, have been formulated as semidefinite programming problems [2, 5]. (The mask constraint is *not* convex in $g[k]$ unless the filter is constrained to have linear phase [2, 5], which can lead to increased spectral occupation.) Although the mask constraints in (7) are semi-infinite, they are substantially 'softer' than that in (3c). If they are violated, a filter exists—it just fails to satisfy the mask. In contrast, violation of (3c) is catastrophic in the sense that no filter would exist. In practice, the mask constraint can be (conservatively) enforced using discretization techniques.

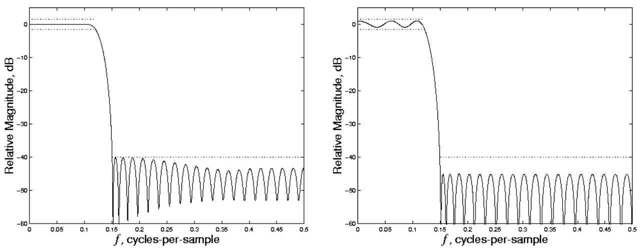
In many applications, filter masks are specified in terms of the relative magnitude of the power spectrum at different frequencies, usually on a logarithmic (decibel) scale. If we let $\rho_\ell(f)$ and $\rho_u(f)$ denote the lower and upper relative magnitude bounds in decibels, then a problem corresponding to Formulation 3 is:

Problem 3 *Given* $\rho_\ell(f)$*,* $\rho_u(f)$*,* N *and* L*, either find* $r_g[m]$*,* $m = 0, 1, \ldots, L-1$*,* $\boldsymbol{P} = \boldsymbol{P}^T$*, and* $\zeta > 0$ *such that* (3a) *and* (4) *hold and*

$$\zeta 10^{\rho_\ell(f)/10} \leq R_g(e^{j2\pi f}) \leq \zeta 10^{\rho_u(f)/10}, \text{ for all } f \in [0, 1/2], \quad (8)$$

or show that none exist.

Problem 3 is a semidefinite feasibility problem and can be used as the sub-problem in a bisection search for the minimal length filter satisfying a given mask, as we demonstrate below.



(a) Designed filter, $L=51$ (b) IS95 filter, $L=48$

Fig. 3: Power spectra of the filters in Example 2 with the magnitude bounds from the IS95 standard.

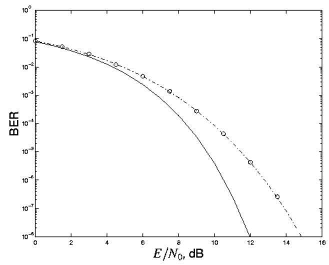


Fig. 4: Calculated (lines) and simulated (circles) bit error rates (BER) against signal-to-noise ratio for the filters in Example 2. Legend—Solid: designed filter; Dash-Dot: IS95 filter.

Example 2 Here we design a filter to compete with the filter specified for the chip waveform in IS95. The standard requires a filter with a ± 1.5dB ripple in the pass-band $f \in [0, f_p]$, and 40dB attenuation in the stop-band $f \in [f_s, 1/2]$, where $f_p \approx 0.12$ and $f_s \approx 0.15$. The filter chosen in the standard has linear phase, $N=4$, and $L=48$. Whilst that filter satisfies the spectral mask, it does not satisfy the orthogonality constraints. Hence, the IS95 filter can induce substantial ISI even when the physical channel is benign. Therefore, we seek the minimal length filter such that *both* the frequency response mask is satisfied *and* the filter is orthogonal. The (global) solution to this problem can be found using a bisection-based search in which Problem 3 is solved at each stage. (The analogous problem of minimizing trace($\boldsymbol{P}$) is solved at the final stage.) This resulted in a length $L=51$, so orthogonality is achieved for the price of a mild increase in filter length (from $L=48$). The frequency response of the designed filter is shown in Fig. 3 along with that of the IS95 filter. We observe that the equiripple characteristic associated with conventional linear phase filters satisfying such masks does not necessarily extend to the case of minimum phase orthogonal filters. To demonstrate the performance improvement due to the orthogonality of the designed filter, we calculated the bit error rates for binary transmission with energy E over an additive white Gaussian noise channel with noise variance $N_0/2$. The probability of error can be calculated analytically [7] and the results are plotted in Fig. 4. Note that for bit error rates of 10^{-3} and 10^{-6} the SNR gains of the designed filter are 1 dB and 2.4 dB respectively. □

In addition to the minimal length filter problem for a given mask studied in Example 2, a number of related problems can be solved in a similar way. For example, for several simple parameterizations of the mask, the 'smallest' mask such that a filter of a given length satisfies both the mask and the orthogonality constraints (and finding such a filter) can be found in this way.

In practice, spectral masks are enforced on the transmitted rather than the baseband signal. In this paper, we have neglected any nonlinear effects in the transmission path so that the mask can be applied directly to the filter. Adaptations of our approach for non-negligible nonlinearities are reserved for future work.

6. CONCLUSION

In this paper, we have shown that the design of orthogonal waveforms for pulse amplitude modulation can be formulated as a convex semidefinite programme and hence globally optimal waveforms can be found in an efficient manner. The formulation was motivated by the observation that the deployment of baseband digital signal processors removes many of the physical constraints in analogue waveform coding applications, and by a desire to exploit the resulting design freedom in an efficient manner. We demonstrated the effectiveness of our design technique by designing 'chip waveforms' with improved performance over that of those specified in recent standards for CDMA-based mobile telecommunications.

There are many other pulse shape characteristics which can be incorporated into the semidefinite programmes, including: a) compensation for the effects of the smoothing and anti-aliasing filters; and b) maximal robustness to timing error [7]. However, there are a few waveform characteristics which are important in some communications applications but are functions of the waveform itself, rather than its autocorrelation. For example, the strength of the cyclic autocorrelation coefficients of $s[k]$, and the magnitude of the envelope variation. An interesting direction for future work is to examine ways in which such characteristics can be incorporated into the current design framework.

7. REFERENCES

[1] A. Said and J. B. Anderson, "Bandwidth-efficient coded modulation with optimized linear partial-response signals", *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 701–713, Mar. 1998.

[2] S.-P. Wu, S. Boyd, and L. Vandenberghe, "FIR filter design via spectral factorization and convex optimization", in *Applied Computational Control, Signal Processing and Communications*, B. Datta, Ed. Birhauser, 1997.

[3] B. D. O. Anderson, K. L. Hitz, and N. D. Diem, "Recursive algorithm for spectral factorization", *IEEE Trans. on Circuits and Systems*, vol. CAS–21, no. 6, pp. 742–750, Nov. 1974.

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

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

[6] J. Tuqan and P. P. Vaidyanathan, "Globally optimal two channel FIR orthonormal filter banks adapted to the input signal statistics", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Seatle, WA, May 1998, vol. III, pp. 1353–1356.

[7] T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Spectrally-efficient orthogonal pulse shape design via semidefinite programming", in preparation.

[8] J. F. Sturm, "SeDuMi", Software package for MATLAB, CRL, McMaster University, May 1998.