

# Root Nyquist Pulses with an Energy Criterion

Ritesh Sood

Department of Electrical & Computer Engineering  
University of California, Davis  
Email: rsood@ucdavis.edu

Hong Xiao

Department of Mathematics  
University of California, Davis  
Email: hxiao@ucdavis.edu

**Abstract**—The design of root-Nyquist pulses with the maximum possible energy concentration of the power spectrum is a fundamental problem in communications engineering. Since the problem is one of global optimization in nature, earlier attempts at a solution have been primarily numerical. Consequently, it is very likely that the proposed design procedures yield solutions that are only locally optimal. In this paper, a detailed analysis of the optimization problem is presented and it is shown that under certain conditions, a globally optimal solution is possible. The global optimization procedure is arrived at by focusing on dual, which is an eigenvalue minimization problem. Numerical results are presented comparing the pulses so obtained with the raised-cosine pulse and the “better than” raised-cosine pulse recently proposed in the literature.

## I. INTRODUCTION

The equivalent complex baseband signal in a digital communication system employing linear modulation is given by the equation

$$r(t) = \sum_{m=-\infty}^{\infty} d_m g(t - mT) + w(t), \quad (1)$$

where  $d_m$  is the  $m^{\text{th}}$  (complex) message symbol transmitted at the rate  $1/T$ ,  $g(t)$  is the overall channel impulse response, and  $w(t)$  is the complex additive white Gaussian noise (AWGN) process with one-sided power spectral density  $N_0$ . Denoting the transmitter side pulse signal by  $p(t)$  and its Fourier transform by  $P(f)$ , the impulse response of the physical channel by  $c(t)$ , and the receive filter by  $p'(t)$ , the overall channel impulse response is  $g(t) = (p \star c \star p')(t)$ . Since matched filtering is optimal for AWGN channels, we have that  $p'(t) = (p \star c)(-t)$ . An estimate  $\hat{d}_n$  of the transmitted symbol  $d_n$  is obtained by sampling  $r(t)$  at  $t = nT$ .

In order to eliminate inter-symbol interference (ISI), it is necessary and sufficient that  $g(t)$  satisfies the Nyquist condition [1] such that

$$g(t) = \begin{cases} 1 & t = 0, \\ 0 & t = mT, \quad m = \pm 1, \pm 2, \dots \end{cases} \quad (2)$$

We call a composite pulse  $g(t)$  that satisfies the above conditions a *Nyquist pulse*. Perhaps the most widely employed Nyquist pulse is the raised-cosine pulse. Denoting the normalized excess bandwidth parameter (*i.e.* the roll-off factor) by  $\alpha$ , the bandwidth  $W$  of the raised-cosine pulse is  $W = (1 + \alpha)/(2T)$ .

In the event that the physical channel is distortion-less, *i.e.*  $c(t) = \delta(t)$ , if the Fourier transform  $G$  of a Nyquist pulse  $g$

is such that  $G(f) \geq 0$  for all  $f$ , we may extract the transmit pulse  $p$  as  $P(f) = \sqrt{G(f)}$ . We call such a pulse  $p$  a *root Nyquist pulse*.

Nyquist showed that a sufficient condition for a pulse to satisfy (2) is that its spectrum has odd symmetry about  $1/(2T)$ . This implies that an infinite number of Nyquist pulses is possible. Recent research has focused on utilizing this freedom in the design of pulses that are robust to channel imperfections. In [2] a family of pulses is reported which are ISI-free with or without matched filtering. This concept is extended in [3], where a technique for the construction of root Nyquist filters through phase compensation has been proposed. Tan and Beaulieu [4] however showed that the aforementioned pulses are not better than the raised-cosine pulse in terms of robustness to timing jitter, and the noise margin at the sampling instance. These properties are essential to ensure reliable communication in a realistic setting. In [5], [6] analytic formulae for new Nyquist pulses are given which are better than the raised-cosine pulse with respect to the above properties.

In the above communication system model, the pulse  $p$  is assumed to be band-limited, and consequently has infinite support in the time-domain. In a practical implementation, the band-limited pulse  $p$  is truncated outside a finite interval  $[-T_p/2, T_p/2]$  to yield an approximation  $\tilde{p}$ . Let  $T_p/2T = K_p$ . The composite pulse  $\tilde{g}$  is then time limited in the interval  $[-2K_p T, 2K_p T]$  and is not necessarily orthogonal.

A small value of  $K_p$  is desirable for several reasons. For a given oversampling rate, the number of multiplication and addition operations that the receiver must perform per detected symbol is directly proportional to the length of the pulse. In ISI channels, the number of interfering neighbors is smaller when  $K_p$  is small, which simplifies the task of equalization.

The pulse length cannot be made arbitrarily small however, since this would lead to poor confinement of the transmitted signal spectrum. This follows from the Uncertainty Principle [7], [8] which states that a signal cannot have arbitrary concentrations in time and frequency domains simultaneously. For instance, suppose that the duration of the composite pulse is limited to one symbol interval, and the normalized excess bandwidth is  $\alpha$ . The maximum energy concentration of the power spectrum in this case is achieved by the zeroth order *prolate spheroidal wave function* (PSWF)  $\psi_0$  with the band-limit parameter  $\pi(1 + \alpha)/2$ . Since there are no side lobes, ISI is absent under ideal conditions. With  $\alpha = 1$ , the in band

energy concentration achieved by  $\psi_0$  is 98.1046% which is not sufficient for most applications. In order to achieve higher spectral confinement, it becomes necessary to increase  $K_p$  beyond 0.5. The Nyquist conditions (2) must now be imposed explicitly for ISI free reception.

The approach taken in this paper is to express the transmit pulse as a linear combination of the PSWFs and determine the optimal coefficients in the sense of maximizing the in-band energy concentration under the constraint of overall zero ISI. Similar approaches have been taken in [9], [10] previously. The resulting optimization problem is however non-convex, and in general admits multiple local maxima. In the previous studies, issues regarding the global optimality or uniqueness of the solution have not been addressed. It is very likely that the reported solutions are local optima and it is possible to improve upon them.

The foremost contribution of this paper is to lay down the conditions under which the problem admits global maxima. We accomplish this by focusing on the *dual* of the above optimization problem. Furthermore, when such conditions are satisfied, we show that the dual problem is an *unconstrained* eigenvalue minimization problem, a subject that has received considerable attention in the linear algebra literature [11], [12]. Our numerical results show that the dual problem is solved efficiently, typically within twenty iterations.

We conclude this section with a note on the notation used in the following sections. Uppercase symbols denote matrices, while lowercase boldface symbols are (column) vectors. Gradient vectors are written as row vectors. The  $i^{\text{th}}$  column of  $X$  is written as  $X_{:,i}$  and similarly,  $X_{i,:}$  is the  $i^{\text{th}}$  row. The vector-space of  $n \times n$  real, symmetric matrices is denoted by  $\mathcal{S}^n$ .

## II. ENERGY CONCENTRATED NYQUIST PULSES

Given the interval  $\mathcal{T} = [-T_p/2, T_p/2]$ , we denote by  $\mathcal{D}$  the linear space of square integrable functions time limited to  $\mathcal{T}$ . Let  $\mathcal{W}$  be the frequency interval  $[-W, W]$ , and denote by  $c = \pi W T_p$  the time-bandwidth product. The prolate spheroidal wave functions  $\psi_m$ ,  $m = 0, 1, \dots$  are parametrized by  $c$ , and form a basis of  $\mathcal{D}$ . Let  $\Psi_m$  denote the Fourier transform of  $\psi_m$ . The relevant properties of the PSWFs are summarized below.

- J1.  $\{\psi_m\}_{m=0}^{\infty}$  comprises an orthonormal basis for  $\mathcal{D}$ .
- J2. The functions  $\Psi_m$  are orthogonal in the interval  $\mathcal{W}$ , and the energy concentration in this interval is  $\int_{-W}^W |\Psi_m(f)|^2 df = \mu_m$ .
- J3.  $\psi_m$  is even when  $m$  is even and odd when  $m$  is odd.
- J4. The highest energy concentration in  $\mathcal{W}$  possible for any function in  $\mathcal{D}$  is  $\mu_0$ , which is achieved by  $\psi_0$ .
- J5. For each  $n \in \mathcal{Z}_+$ , among functions in  $\mathcal{D}$  which are orthogonal to the linear space spanned by  $\{\psi_0, \dots, \psi_{n-1}\}$ , the highest energy concentration in  $\mathcal{W}$  is achieved by  $\psi_n$ .

### A. A primal formulation

We are concerned with the design of the transmit pulse  $p$  which is time limited to  $\mathcal{T}$  and achieves the maximum energy

concentration in the frequency interval  $\mathcal{W}$ . Furthermore, it is required that the composite pulse  $g$  satisfies (2). Due to (J1) and (J3),  $p(t)$  and its Fourier transform  $P(f)$  are written as

$$p(t) = \sum_{m=1}^n x_m \psi_{2m-2}(t), \quad P(f) = \sum_{m=1}^n x_m \Psi_{2m-2}(f). \quad (3)$$

From (J1), we have that the net energy of the pulse is  $\|p\|^2 = \mathbf{x}^T \mathbf{x}$ , and the in-band energy  $\int_{-W}^W |P(f)|^2 df = \mathbf{x}^T M \mathbf{x}$  due to (J2). Here  $M$  is a diagonal matrix with elements  $M_{l,l} = \mu_{2l-2}$ . Since  $p$  is symmetric, the receive pulse is also  $p$  and the composite pulse is therefore given by the auto-correlation function

$$g(t) = \int_{-T_p/2}^{T_p/2} p(\tau)p(\tau+t) d\tau. \quad (4)$$

Substituting (3) in (4), we have for  $t = jT$ ,  $j \in \mathcal{Z}$

$$\begin{aligned} g(jT) &= \sum_{m=1}^n \sum_{l=1}^n x_m x_l \int_{-T_p/2}^{T_p/2} \psi_{2m-2}(\tau)\psi_{2l-2}(\tau+jT) d\tau \\ &= \mathbf{x}^T H_j \mathbf{x}. \end{aligned}$$

By defining the objective function  $\zeta : \mathcal{R}^n \rightarrow \mathcal{R}$  as

$$\zeta(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T M \mathbf{x},$$

the vector valued constraint function  $\mathbf{h} : \mathcal{R}^n \rightarrow \mathcal{R}^k$  as

$$h_j(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T H_j \mathbf{x}, \quad (5)$$

and the norm constraint

$$\mathbf{x}^T \mathbf{x} = 1, \quad (6)$$

we have the following optimization problem, which we call the *primal*.

$$(P) \quad \begin{array}{ll} \max_{\mathbf{x}} & \zeta(\mathbf{x}) \\ \text{s. t.} & \mathbf{h}(\mathbf{x}) = 0 \\ & \frac{1}{2}(\mathbf{x}^T \mathbf{x} - 1) = 0 \end{array}$$

We denote by  $[\boldsymbol{\lambda}^T \lambda_0]^T$  ( $\boldsymbol{\lambda} \in \mathcal{R}^k$ ,  $\lambda_0 \in \mathcal{R}$ ), the vector of Lagrange-multipliers associated with (P), and by  $l(\mathbf{x}, \boldsymbol{\lambda}, \lambda_0)$  the Lagrangian function defined as

$$\begin{aligned} l(\mathbf{x}, \boldsymbol{\lambda}, \lambda_0) &= \zeta(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{h}(\mathbf{x}) + \frac{1}{2} \lambda_0 (\mathbf{x}^T \mathbf{x} - 1) \\ &= \frac{1}{2} \mathbf{x}^T M \mathbf{x} + \sum_{j=1}^k \frac{1}{2} \lambda_j (\mathbf{x}^T H_j \mathbf{x}) + \frac{1}{2} \lambda_0 (\mathbf{x}^T \mathbf{x} - 1) \\ &= \frac{1}{2} \mathbf{x}^T \left[ M + \sum_{j=1}^k \lambda_j H_j \right] \mathbf{x} + \frac{1}{2} \lambda_0 (\mathbf{x}^T \mathbf{x} - 1). \end{aligned}$$

Introducing the matrix function  $A : \mathcal{R}^k \rightarrow \mathcal{S}^n$  defined as

$$A(\boldsymbol{\lambda}) = M + \sum_{j=1}^k \lambda_j H_j, \quad (7)$$

we have

$$l(\mathbf{x}, \boldsymbol{\lambda}, \lambda_0) = \frac{1}{2} \mathbf{x}^T A(\boldsymbol{\lambda}) \mathbf{x} + \frac{1}{2} \lambda_0 (\mathbf{x}^T \mathbf{x} - 1). \quad (8)$$

The gradient of  $l$  with respect to  $\mathbf{x}$  is

$$\nabla_{\mathbf{x}} l(\mathbf{x}, \boldsymbol{\lambda}, \lambda_0)^T = A(\boldsymbol{\lambda})\mathbf{x} + \lambda_0\mathbf{x}, \quad (9)$$

while that with respect to  $\lambda_j$  is

$$\nabla_{\lambda_j} l(\mathbf{x}, \boldsymbol{\lambda}, \lambda_0) = \frac{1}{2}\mathbf{x}^T H_j \mathbf{x}, \quad j = 1, \dots, k. \quad (10)$$

$$\nabla_{\lambda_0} l(\mathbf{x}, \boldsymbol{\lambda}, \lambda_0) = \frac{1}{2}(\mathbf{x}^T \mathbf{x} - 1). \quad (11)$$

### B. Local optimality conditions

In order to obtain the first-order necessary conditions for a local extremum, we equate (9)–(11) with zero. Equation (9) yields the eigenvalue equation

$$A(\boldsymbol{\lambda})\mathbf{x} = -\lambda_0\mathbf{x}, \quad (12)$$

while equations (10) and (11) restate the conditions for  $\mathbf{x}$  to be a feasible point of (P):

$$\frac{1}{2}\mathbf{x}^T H_j \mathbf{x} = 0 \quad j = 1, \dots, k \quad (13)$$

$$\frac{1}{2}\mathbf{x}^T \mathbf{x} = \frac{1}{2}. \quad (14)$$

Since  $A(\boldsymbol{\lambda}) \in \mathcal{S}^n$ , all its eigenvalues are real and the eigenvectors comprise an orthonormal basis in  $\mathcal{R}^n$ . Let  $\sigma_i, i = 1, \dots, \eta$  denote the distinct eigenvalues of  $A(\boldsymbol{\lambda})$ , indexed so that  $\sigma_1 > \dots > \sigma_\eta$ . Let the integers  $t_i$  denote the multiplicity of  $\sigma_i$  ( $i = 1, \dots, \eta$ ). Denoting by  $\mathcal{E}_i(\boldsymbol{\lambda})$  the eigenspace of  $\sigma_i$  and  $\mathbf{x}_i$  any unit eigenvector in this space (*i.e.*,  $\mathbf{x}_i \in \mathcal{E}_i(\boldsymbol{\lambda}), \|\mathbf{x}_i\| = 1$ ), we have that the pair  $(\mathbf{x}_i, \sigma_i)$  is a solution of (12)

$$A(\boldsymbol{\lambda})\mathbf{x}_i = \sigma_i\mathbf{x}_i$$

with the substitution  $\lambda_0 = -\sigma_i$ . Introducing the notation that  $L(\mathbf{x}_i, \boldsymbol{\lambda}, \sigma_i) = \nabla_{\mathbf{x}\mathbf{x}} l(\mathbf{x}_i, \boldsymbol{\lambda}, \sigma_i)$  which is the Hessian matrix of the Lagrangian function  $l$ , and differentiating (9) on both sides with respect to  $\mathbf{x}$ , we obtain

$$L(\mathbf{x}_i, \boldsymbol{\lambda}, \sigma_i) = A(\boldsymbol{\lambda}) - \sigma_i I_n,$$

where  $I_n$  is the identity matrix in  $\mathcal{R}^{n \times n}$ .

The gradient vectors to the constraint surface defined by the equations (13) at the point  $\mathbf{x}_i$  are the vectors  $\nabla_{\mathbf{x}} h_j(\mathbf{x}_i) = \mathbf{x}_i^T H_j$ ,  $j = 1, \dots, k$ . Let  $\mathcal{H}(\mathbf{x}_i)$  denote the sub-spaces

$$\mathcal{H}(\mathbf{x}_i) = \text{span}(\{H_j \mathbf{x}_i, j = 1, \dots, k\}),$$

the orthogonal complement of  $\mathcal{H}(\mathbf{x}_i)$  is the sub-space of first order feasible variations [13] to the constraint (13). Defining the matrix of gradient vectors  $\nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}_i)^T$  as

$$[\nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}_i)^T]_{:,j} = H_j \mathbf{x}_i, \quad (15)$$

we obtain the following expression for this sub-space:

$$\mathcal{H}(\mathbf{x}_i)^\perp = \{\mathbf{y} \in \mathcal{R}^k \mid [\nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}_i)]\mathbf{y} = 0\}.$$

The gradient vector of the constraint (14) is  $\mathbf{x}_i$ . If we denote by  $\mathcal{X}_i$  the one dimensional sub-space spanned by  $\mathbf{x}_i$ , then following the above discussion, the space of first-order feasible variations to the constraint (14) is  $\mathcal{X}_i^\perp$ .

The second order sufficiency conditions for  $\mathbf{x}_i$  to be a local maximum for (P) then require that  $L(\mathbf{x}_i, \boldsymbol{\lambda}, \sigma_i)$  is negative definite on  $\mathcal{V}(\mathbf{x}_i)$ , where

$$\mathcal{V}(\mathbf{x}_i) = \mathcal{H}(\mathbf{x}_i)^\perp \cap \mathcal{X}_i^\perp. \quad (16)$$

Suppose that for some  $\boldsymbol{\lambda}$ , the largest eigenvalue  $\sigma_1$  is simple, and the unique eigenvector  $\mathbf{x}_1$  (up to a sign) is feasible. Then it is easily verified that  $A(\boldsymbol{\lambda}) - \sigma_1 I$  is negative definite on  $\mathcal{X}_1^\perp$ . Consequently,  $L(\mathbf{x}_1, \boldsymbol{\lambda}, \sigma_1)$  is negative definite on  $\mathcal{V}(\mathbf{x}_1)$  due to (16), and  $\mathbf{x}_1$  is a strict locally maximizing vector for (P). Even when  $\sigma_1$  has multiplicity greater than one, given that the following condition is satisfied for  $i = 1$

$$\mathcal{H}(\mathbf{x}_i)^\perp \cap \mathcal{E}_i(\boldsymbol{\lambda}) = \{0\},$$

$\mathbf{x}_1$  is a strict local maximum.

A similar analysis for the smallest eigenvalue ( $i = \eta$ ) reveals that a corresponding eigenvector never satisfies the second order sufficiency conditions for a local maximum point. When  $1 < i < \eta$ , however, not much can be said *a priori*. The quantity  $A(\boldsymbol{\lambda}) - \sigma_i I$  on  $\mathcal{E}_m(\boldsymbol{\lambda})$  is positive definite for  $1 \leq m < i$ , negative definite for  $i < m \leq \eta$  and identically zero on  $\mathcal{E}_i(\boldsymbol{\lambda})$ . The definiteness of the restriction of the Hessian matrix on  $\mathcal{V}(\mathbf{x}_i)$  in these cases depends on the particular alignment and containment of  $\mathcal{H}(\mathbf{x}_i)^\perp$  with respect to the subspaces  $\mathcal{E}_m(\boldsymbol{\lambda})$ .

Suppose that the eigenvectors  $\mathbf{x}_l, \mathbf{x}_m$  satisfy the second order sufficiency conditions and that  $\sigma_l > \sigma_m$  are the corresponding eigenvalues. From the definition of  $A(\boldsymbol{\lambda})$  (see (7)) and the feasibility conditions (13), we have that

$$\sigma_i = \mathbf{x}_i^T A(\boldsymbol{\lambda})\mathbf{x}_i = \mathbf{x}_i^T M \mathbf{x}_i = 2\zeta(\mathbf{x}_i). \quad (17)$$

Thus among candidate vectors in the solution set of (P), we choose the one corresponding to the largest eigenvalue. In other words, the following two conditions are sufficient for a solution  $\mathbf{x}^*$  to have strict local optimality:

- C1. There is a  $\boldsymbol{\lambda}^* \in \mathcal{R}^k$  such that  $\mathbf{x}^*$  is an eigenvector with unity norm corresponding to the largest eigenvalue  $\sigma_1$  of  $A(\boldsymbol{\lambda}^*)$ , and  $\sigma_1$  is simple.
- C2.  $\mathbf{x}^*$  is feasible for (P):  $\mathbf{h}(\mathbf{x}^*) = 0$ .

### C. Existence of a globally optimal solution

The largest eigenvalue function on the space of real symmetric (complex hermitian) matrices  $\sigma_1 : \mathcal{S}^n \rightarrow \mathcal{R}$  has an important extremal characteristic: it is real-valued and *convex*. As in the present case where  $A$  itself depends on a parameter  $\boldsymbol{\lambda}$  in an affine manner, the convexity still holds when  $\sigma_1$  is seen as a function of  $\boldsymbol{\lambda}$ ,  $\sigma_1(\boldsymbol{\lambda}) = \sigma_1(A(\boldsymbol{\lambda}))$ . In the next section we will see that the convexity of  $\sigma_1$  naturally leads to defining it as the *dual* function, and the optimization problem dual to (P) is the *unconstrained minimization* of  $\sigma_1$ . Furthermore, when the conditions (C1) and (C2) hold, we will see that not only is  $\mathbf{x}^*$  a strict local maxima, but it is a *global* maximizing point for (P). For the dual problem in this case,  $\boldsymbol{\lambda}^*$  is the unique global minimizer of  $\sigma_1$ .

In a converse manner, if at the minimizing  $\boldsymbol{\lambda}^*$ ,  $\sigma_1(\boldsymbol{\lambda}^*)$  is simple, then (C2) automatically holds and  $\mathbf{x}_1(\boldsymbol{\lambda}^*)$  maximizes

$\zeta$  globally. In Section III-A we will see that the first and second derivatives of  $\sigma_1$  (*i.e.* the gradient vector  $\nabla_{\lambda}\sigma_1$  and the Hessian matrix  $\nabla_{\lambda\lambda}\sigma_1$ ) exist at the point  $\lambda$  if and only if  $\sigma_1(\lambda)$  is simple. The derivative information enables us to use quadratically convergent Newton's method for the minimization of  $\sigma_1$ . Due to space constraints, we confine our analysis to this particular case.

Extension to the case where in a neighborhood of the minimizing  $\lambda^*$ ,  $\sigma_1(\lambda^*)$  has multiplicity greater than one is presented in [14]. Here  $\nabla_{\lambda}\sigma_1$  is replaced by the *subdifferential*  $\partial\sigma_1(\lambda)$  and the optimality condition  $\nabla_{\lambda}\sigma_1(\lambda^*) = 0$  by the inclusion  $0 \in \partial\sigma_1(\lambda^*)$  [12]. It is shown in [14] that if  $\mathbf{x}^* \in \mathcal{E}_1(\lambda^*)$  is such that  $\mathbf{h}(\mathbf{x}^*)$  is an extreme point of the convex set  $\partial\sigma_1(\lambda^*)$ , then  $\mathbf{x}^*$  is a global maximizer of (P).

A third possibility exists, *i.e.*, for certain matrices  $A$  and  $H_j$ 's, it might be that for the  $\lambda^*$  which minimizes  $\sigma_1$ , there is no vector in the eigenspace of  $\sigma_1(\lambda^*)$  that is feasible for (P). When this is the case, in the terminology of mathematical programming [13] we say that there is a *duality gap* between the primal and dual problems. It can be shown that the maximum value of (P) equals (half) the magnitude of the second-largest, or subsequent eigenvalues. In the case that a duality gap does exist in the problem, it will be revealed during the course of minimizing  $\sigma_1$  [14]. A possible approach then is to iteratively solve the non-linear system of equations  $\mathbf{h}(\mathbf{x}_i(\lambda))$  in the unknown  $\lambda$ . At each step of the iteration,  $\mathbf{x}_i(\lambda)$  is an eigenvector corresponding to the  $i^{\text{th}}$  eigenvalue for  $1 < i < \eta$ .

### III. A DUAL FORMULATION

We first establish the convexity of the largest eigenvalue function. The following result is well-known [12] and included here for completeness.

*Lemma 1:* The largest eigenvalue function  $\sigma_1 : \mathcal{S}^n \rightarrow \mathcal{R}$  on the domain of real symmetric matrices  $\mathcal{S}^n$  is convex.

*Proof:* From the Rayleigh principle [15] for real symmetric (complex hermitian) matrices

$$\sigma_1(A) = \max_{\|\mathbf{x}\|=1} \mathbf{x}^T A \mathbf{x}.$$

Suppose that  $A_1, A_2 \in \mathcal{S}^n$  and  $0 \leq \theta \leq 1$ . By virtue of  $\mathcal{S}^n$  being a vector space, the convex combination  $\theta A_1 + (1 - \theta)A_2 \in \mathcal{S}^n$ .

$$\begin{aligned} \sigma_1(\theta A_1 + (1 - \theta)A_2) &= \max_{\|\mathbf{x}\|=1} \mathbf{x}^T (\theta A_1 + (1 - \theta)A_2) \mathbf{x} \\ &= \max_{\|\mathbf{x}\|=1} \theta \mathbf{x}^T A_1 \mathbf{x} + (1 - \theta) \mathbf{x}^T A_2 \mathbf{x} \\ &\leq \theta \left( \max_{\|\mathbf{x}\|=1} \mathbf{x}^T A_1 \mathbf{x} \right) \\ &\quad + (1 - \theta) \left( \max_{\|\mathbf{x}\|=1} \mathbf{x}^T A_2 \mathbf{x} \right) \\ &= \theta \sigma_1(A_1) + (1 - \theta) \sigma_1(A_2). \end{aligned}$$

Recall the definition of  $A(\lambda)$  (7). Since  $M, H_j$  are fixed,  $A(\lambda)$  is an affine function of its argument  $\lambda \in \mathcal{R}^k$  into  $\mathcal{S}^n$ . Therefore  $\sigma_1(A(\lambda))$  is the composition of an affine function on  $\mathcal{R}^k$  with a convex function on  $\mathcal{S}^n$ . It follows from the

theory of convex functions [16] that the composite function  $\sigma_1 : \mathcal{R}^k \rightarrow \mathcal{R}$  is convex.

We will now relate the optimal value  $\zeta^*$  of the primal optimization problem (P) and the largest eigenvalue  $\sigma_1(\lambda)$  of  $A(\lambda)$ . For any  $\lambda \in \mathcal{R}^k$

$$\begin{aligned} \zeta^* &= \max_{\|\mathbf{x}\|=1, \mathbf{h}(\mathbf{x})=0} \zeta(\mathbf{x}) \\ &= \max_{\|\mathbf{x}\|=1, \mathbf{h}(\mathbf{x})=0} \zeta(\mathbf{x}) + \lambda^T \mathbf{h}(\mathbf{x}) \\ &= \max_{\|\mathbf{x}\|=1, \mathbf{h}(\mathbf{x})=0} \frac{1}{2} \mathbf{x}^T A(\lambda) \mathbf{x} \\ &\leq \max_{\|\mathbf{x}\|=1} \frac{1}{2} \mathbf{x}^T A(\lambda) \mathbf{x} = \frac{1}{2} \sigma_1(\lambda). \end{aligned}$$

Therefore

$$\zeta^* \leq \inf_{\lambda \in \mathcal{R}^k} \frac{1}{2} \sigma_1(\lambda). \quad (18)$$

Suppose that there is a  $\lambda^*$  such that  $\mathbf{x}^*$  which maximizes  $\zeta$

$$\mathbf{x}^* \in \arg \max_{\|\mathbf{x}\|=1, \mathbf{h}(\mathbf{x})=0} \zeta(\mathbf{x}).$$

is also an eigenvector corresponding to  $\sigma_1(\lambda^*)$ . Then

$$\zeta^* = \zeta(\mathbf{x}^*) = \frac{1}{2} \mathbf{x}^{*T} A(\lambda^*) \mathbf{x}^* = \frac{1}{2} \sigma_1(\lambda^*). \quad (19)$$

Combining (18) and (19) we conclude that  $\lambda^*$  minimizes  $\sigma_1$  and the minimum value is  $\zeta^*$ . Equivalently,  $\lambda^*$  solves the following optimization problem

$$(D) \quad \min_{\lambda \in \mathcal{R}^k} \sigma_1(\lambda).$$

We call (D) the *dual* problem of (P).

Conversely, suppose that the pair  $(\bar{\mathbf{x}}, \bar{\lambda})$ ,  $\bar{\mathbf{x}} \in \mathcal{E}_1(\bar{\lambda})$  is such that  $\bar{\mathbf{x}}$  is feasible for (P), meaning  $\mathbf{h}(\bar{\mathbf{x}}) = 0$ . Then

$$\begin{aligned} \zeta(\bar{\mathbf{x}}) &= \frac{1}{2} \bar{\mathbf{x}}^T A(\bar{\lambda}) \bar{\mathbf{x}} \\ &= \max_{\|\mathbf{x}\|=1} \frac{1}{2} \mathbf{x}^T A(\bar{\lambda}) \mathbf{x} \\ &\geq \max_{\|\mathbf{x}\|=1, \mathbf{h}(\mathbf{x})=0} \frac{1}{2} \mathbf{x}^T A(\bar{\lambda}) \mathbf{x} \\ &= \zeta^*. \end{aligned} \quad (20)$$

But since  $\zeta(\bar{\mathbf{x}}) \leq \zeta^*$ , equality holds in the above equations throughout. Thus  $\zeta(\bar{\mathbf{x}}) = \zeta^*$  which means that

$$\begin{aligned} \bar{\mathbf{x}} &\in \arg \max_{\|\mathbf{x}\|=1, \mathbf{h}(\mathbf{x})=0} \zeta(\mathbf{x}), \quad \text{and} \\ \bar{\lambda} &\in \arg \min_{\lambda \in \mathcal{R}^k} \sigma_1(\lambda). \end{aligned}$$

#### A. Unconstrained minimization of $\sigma_1$

The eigenvalues  $\sigma_i(A)$  of a symmetric matrix  $A$  are Lipschitz-continuous functions of the matrix entries, and the Lipschitz-constant is one [17]. When  $A$  is a smooth function of a parameter  $\lambda$ , the smoothness and differentiability properties extend to  $\sigma_i(\lambda)$  as a function of  $\lambda$ . In case the  $i^{\text{th}}$  eigenvalue  $\sigma_i(\lambda)$  is simple (*i.e.* its multiplicity is one), it is

Fréchet differentiable and the gradient vector  $\nabla_{\lambda}\sigma_i(\lambda)$  exists. Furthermore, the corresponding eigenvector  $\mathbf{x}_i(\lambda)$  is uniquely defined and continuously differentiable in some neighborhood of  $\lambda$  [15]. As is true in the present case, that  $A(\lambda)$  is a twice continuously differentiable function of  $\lambda$ , then  $\sigma_i(\lambda)$  inherits the  $C^2$  smoothness of  $A$  and we can compute the Hessian matrix  $\nabla_{\lambda\lambda}\sigma_i(\lambda)$ .

By differentiating the equation

$$\begin{aligned} \frac{1}{2}\sigma_i(\lambda) &= \frac{1}{2}\mathbf{x}_i(\lambda)^T A(\lambda)\mathbf{x}_i(\lambda) \\ &= \zeta(\mathbf{x}_i(\lambda)) + \lambda^T \mathbf{h}(\mathbf{x}_i(\lambda)) \end{aligned}$$

on both sides with respect to  $\lambda$  and using the relation  $\mathbf{x}_i(\lambda)^T \mathbf{x}_i(\lambda) = 1$ , it is easily verified that the gradient of  $\sigma_i$  is

$$\nabla_{\lambda}\sigma_i(\lambda)^T = 2\mathbf{h}(\mathbf{x}_i(\lambda)). \quad (21)$$

In other words, the gradient vector of  $\sigma_i$  has a simple expression: it is twice the constraint function evaluated at the eigenvector  $\mathbf{x}_i(\lambda)$ .

It follows from the convexity of  $\sigma_1$  that the first order necessary condition for  $\lambda^*$  to belong to the solution set of (D),  $\nabla_{\lambda}\sigma_1(\lambda^*) = 0$ , is also sufficient. Equation (21) implies that when  $\lambda^*$  minimizes  $\sigma_1(\lambda)$ ,  $\mathbf{x}_1(\lambda^*)$  is feasible for (P). It follows from (20) that  $\mathbf{x}_1(\lambda^*)$  is a global maximizer for (P).

The uniqueness of  $\lambda^*$  as the minimization vector is determined by the definiteness of the Hessian matrix  $\nabla_{\lambda\lambda}\sigma_1(\lambda^*)$ . We give the following expression for the Hessian matrix without derivation

$$\nabla_{\lambda\lambda}\sigma_1(\lambda) = [\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}_1(\lambda))] X \Sigma^{\dagger} X^T [\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}_1(\lambda))]^T,$$

where the diagonal matrix  $\Sigma^{\dagger}$  is defined as

$$\Sigma_{l,l}^{\dagger} = \begin{cases} 0 & l = 1, \\ (\sigma_1 - \sigma_l)^{-1} & \text{otherwise.} \end{cases}$$

We impose a regularity assumption on  $\mathbf{x}_1(\lambda^*)$ , *i.e.* the matrix  $\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}_1(\lambda^*))^T$  (15) has full column-rank. The feasibility condition (13) implies that  $\mathbf{x}_1(\lambda^*) \in \mathcal{H}(\mathbf{x}_1(\lambda^*))^{\perp}$ . Thus  $\mathbf{x}_1(\lambda^*)$  is not in the range space of  $\nabla_{\mathbf{x}}\mathbf{h}(\mathbf{x}_1(\lambda^*))^T$ . On the other hand, the quadratic form  $\mathbf{y}^T (X \Sigma^{\dagger} X^T) \mathbf{y} = 0$  only if  $\mathbf{y} = \beta \mathbf{x}_1(\lambda^*)$  for some real  $\beta$ , and positive otherwise. It follows that the Hessian matrix above is positive definite.

We employ Newton's method for the unconstrained minimization of  $\sigma_1$ . The algorithm proceeds according to the outer iteration

$$\lambda^{\nu+1} = \lambda^{\nu} - \beta^{\nu} [\nabla_{\lambda\lambda}\sigma_1(\lambda^{\nu})]^{-1} \nabla_{\lambda}\sigma_1(\lambda^{\nu})^T,$$

where  $\nu \in \mathcal{Z}_+$  is the iteration index, and  $0 < \beta \leq 1$  is a step-size. Near the solution,  $\beta = 1$  gives quadratic convergence, while a smaller step-size in the initial phase of the algorithm ensures global convergence. A line search procedure that minimizes the objective function in the Newton direction  $-\nabla_{\lambda\lambda}\sigma_1(\lambda^{\nu})^{-1} \nabla_{\lambda}\sigma_1(\lambda^{\nu})^T$  constitutes the inner iteration [13].

## IV. NUMERICAL RESULTS

We select the number of basis functions as  $n = \lceil 2(1 + \alpha)K_p \rceil$ . This selection follows from [8] where it is shown that the dimension of the space of functions that are time limited in  $[-T_p/2, T_p/2]$  and *essentially* band limited in the interval  $[-W, W]$  is  $2WT_p$ . As is pointed out in Section I, when  $K_p$  is large, the composite truncated pulses  $\tilde{g}$  proposed in [6], [3] are very close to the original bandlimited pulse  $g$ . It is in the region of small  $K_p$  that the difference between the bandlimited pulse and its truncated version is substantial. In view of these observations, the results in this section are for  $K_p = 1.25$  throughout. This means that the number of constraints in (5) is three, and provided that the (normalized) timing offset at the receiver is bounded as  $|\Delta| \leq 0.5$ , a symbol experiences interference from at most two neighbors on either side.

Fig. 1 compares the in-band energy concentration of the raised cosine, "better than" raised cosine (btrc) [5], and psfw pulses. This is upper bound by  $\mu_0$ , achieved by the zeroth PSWF  $\psi_0$  with  $c = \pi(1 + \alpha)K_p$ . The raised-cosine pulse is more concentrated than the btrc and psfw pulses in that order. However, as is shown in Fig. 2, this advantage is achieved at the cost of loss of orthogonality and robustness to ISI. Fig. 2 shows the matched pulses and their tails when the excess bandwidth parameter is  $\alpha = 0.4$ . While the optimized pulse satisfies the Nyquist property, the btrc, and raised-cosine pulses have amplitudes at the location of the first neighboring symbols that are as large as 5% and 10% of the amplitude at the sampling instance.

Suppose that imperfect timing recovery at the receiver results in a (normalized) time offset  $\delta$  of the receivers clock. From (1) the  $n^{\text{th}}$  estimated symbol is

$$\hat{d}_n = d_n g(0, \delta) + \sum_{i=I_l}^{I_r} d_{(n+i)} g(-iT, \delta) + w_n,$$

where  $g(iT, \delta) = \int_{-T_p/2}^{T_p/2} p(\tau) p(\tau + [i + \delta]T) d\tau$ ,  $I_l = \lceil -2K_p - \delta \rceil$ , and  $I_r = \lfloor 2K_p + \delta \rfloor$ . It is assumed that the sampled noise  $w_n$  has variance  $N_0/2$  in each of the real and imaginary components, and that the message symbols have mean energy  $E_s$ , are mutually independent, and are independent of the additive noise. The resulting signal to noise and interference ratio (SINR) is thus given by

$$\rho(\delta) = \frac{E_s g(0, \delta)^2}{E_s \sum_{i=I_l}^{I_r} g(-iT, \delta)^2 + N_0}.$$

In a practical setting,  $\delta$  varies with time as the receiver acquires the timing signal and tracks it. We therefore assume that  $\delta$  is uniformly distributed in the interval  $[-\Delta, \Delta]$  and the mean SINR  $\rho = \frac{1}{2\Delta} \int_{-\Delta}^{\Delta} \rho(\delta) d\delta$ . The received mean SINR for various values of  $\Delta$ , and SNR =  $E_s/N_0$  is illustrated in Fig. 3. While Fig. 2 verifies that the psfw pulse completely eliminates ISI under ideal conditions, Fig. 3 reveals that it is the most robust to interference under non-ideal conditions.

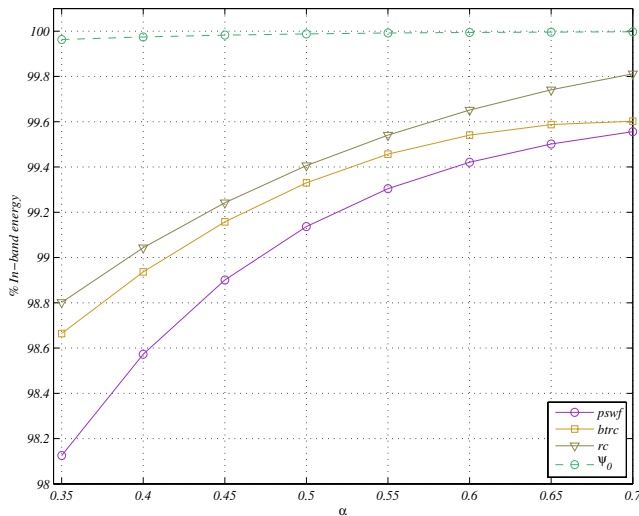


Fig. 1. In-band energy concentration against excess bandwidth parameter  $\alpha$ .

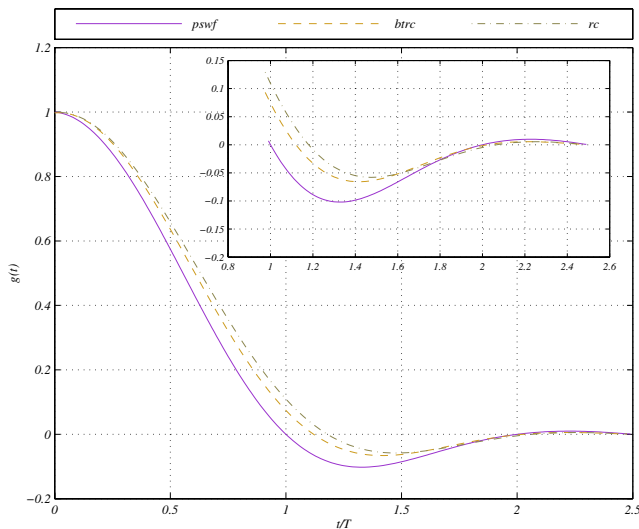


Fig. 2. Composite pulses and side-lobes (inset).

## V. CONCLUSIONS

A method for the construction of root Nyquist pulses which are optimal in the sense of energy concentration in the power spectrum is presented. In its original formulation, the problem is non-convex and admits multiple locally optimal solutions. It is shown in this paper that under certain conditions, solving the dual optimization problem yields a globally optimal solution to the primal. It is advantageous to solve the dual problem, since the dual objective function is easily evaluated, and is convex. While the dual objective function is non-smooth in general, when certain smoothness criteria are satisfied, the quadratically convergent Newton's method can be employed. Numerical results show that convergence is typically attained within twenty iterations. Design examples using the proposed method are presented, and the performance of the pulses so obtained is studied when channel conditions are non-ideal.

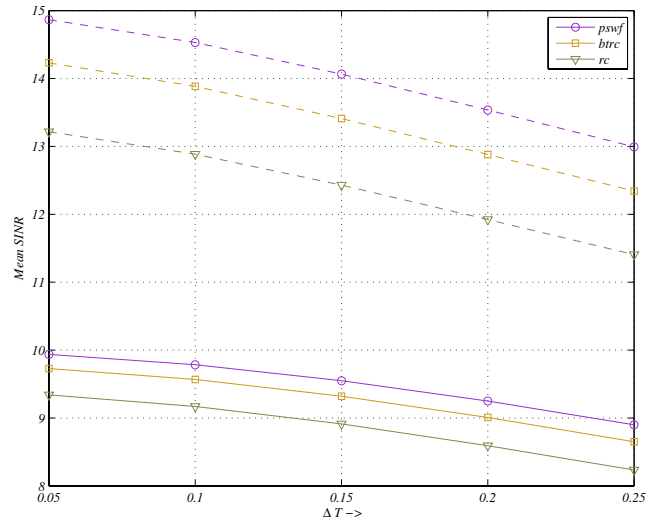


Fig. 3. Received SINR with SNR=10 dB (solid) and 15 dB (dashed).

Our experiments show that besides being optimal in terms of spectral energy concentration, the optimized pulses are more resilient to channel imperfections when compared to other well-known pulses.

## REFERENCES

- [1] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw Hill, 2000.
- [2] X. G. Xia, "A family of pulse shaping filters with ISI-free matched and unmatched properties," *IEEE Commun. Lett.*, pp. 1157–1158, Oct. 1997.
- [3] N. S. Alagha and P. Kabal, "Generalized raised-cosine filters," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 989–997, July 1999.
- [4] C. C. Tan and N. C. Beaulieu, "Transmission properties of conjugate-root pulse," *IEEE Trans. Commun.*, vol. 52, pp. 53–558, Apr. 2004.
- [5] N. C. Beaulieu, C. C. Tan, and M. O. Damen, "A "better than" Nyquist pulse," *IEEE Commun. Lett.*, vol. 5, pp. 367–368, Sept. 2001.
- [6] A. Assalini and A. M. Tonello, "Improved Nyquist pulses," *IEEE Commun. Lett.*, vol. 8, no. 2, pp. 87–89, Feb. 2004.
- [7] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - I," *Bell Syst. Tech. J.*, vol. 40, no. 1, pp. 43–63, Jan. 1961.
- [8] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - II," *Bell Syst. Tech. J.*, pp. 65–94, Jan. 1961.
- [9] P. H. Halpern, "Optimal finite duration Nyquist signals," *IEEE Trans. Commun.*, vol. COMM-27, pp. 884–888, June 1979.
- [10] 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.
- [11] M. L. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix," *SIAM J. Matrix Anal. Appl.*, vol. 9, no. 2, pp. 256–268, 1988. [Online]. Available: <http://locus.siam.org/SIMAX/volume-09/art.0609021.html>
- [12] —, "Large scale optimization of eigenvalues," *SIAM J. Optimization*, vol. 2, no. 1, pp. 88–120, 1992. [Online]. Available: <http://locus.siam.org/SIOPT/volume-02/art.0802007.html>
- [13] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, Mass.: Athena Scientific, 2003.
- [14] R. Sood and H. Xiao, "A global optimization procedure for the design of Nyquist pulses with maximum in-band energy," in preparation.
- [15] P. Lancaster, "On eigenvalues of matrices dependent on a parameter," *Numerische Mathematik*, vol. 6, pp. 377–387, 1964.
- [16] R. T. Rockafellar, *Convex Analysis*. NJ: Princeton University Press, 1996.
- [17] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*. Springer Verlag, 1982.