Bounding the Number of Non-Zero Coefficients in Minimal Peak-to-Peak Gain Shaping Filters

Omer Tanovic[1] and Alexandre Megretski[1]

***Abstract*— In this paper we consider the task of designing linear time invariant filters which minimize maximal peak-to-peak gain subject to certain "spectral mask" and "no intersymbol interference" constraints. This problem (as well as many other similar questions) can be formalized as infinite dimensional L1 minimization subject to a single convex quadratic constraint. We show that, under the assumption of strict feasibility, every optimal solution corresponds to a finite unit sample response (FIR) filter. Furthermore, a constructive upper bound for the number of nonzero coefficients of the optimal filter is given. The results do not rely on a "restricted isometry" assumption, and potentially offer an alternative method of predicting the degree of sparsity of a solution of a convex quadratic program.**

I. NOTATION AND TERMINOLOGY

$\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{N}$ are the usual sets of complex, real, integer, and positive integer numbers, and $\mathbb{T} = \{z \in \mathbb{C} : \ |z| = 1\}$ is the unit circle in $\mathbb{C}$. For elements f, g of a (real) Hilbert space H, $(f,g)_H$ and $|f|_H$ denote the scalar product and the norm. ℓ is the real vector space of all functions $x : \ \mathbb{Z} \to \mathbb{R}$, interpreted as *discrete-time (DT) signals*, with $x[t]$ used for the value of x at $t \in \mathbb{Z}$. δ is the *unit sample* signal $\delta \in \ell$, defined by $\delta[0] = 1$, $\delta[t] = 0$ for $t \neq 0$. For $x \in \ell$, the *L1 norm* $\|x\|_1 \in [0, \infty]$, the *L-Infinity norm* $\|x\|_\infty \in [0, \infty]$, and the *total energy* $E(x) \in [0, \infty]$ are defined by

$$\|x\|_1 = \sum_t |x[t]|, \ \|x\|_\infty = \sup_t |x[t]|, \ E(x) = \sum_t |x[t]|^2.$$

ℓ^2 is the subset of finite energy signals from ℓ, treated as a Hilbert space, with the norm $|x| = E(x)^{1/2}$. L^2 is the real Hilbert space of all square integrable functions $f : \ \mathbb{T} \to \mathbb{C}$ satisfying the *real symmetry* condition $f(\bar{z}) \equiv \overline{f(z)}$ for all $z \in \mathbb{T}$, with the norm defined by

$$|f|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(e^{j\omega})|^2 d\omega.$$

The *Fourier transform* $\mathcal{F} : \ \ell^2 \to L^2$, where $\hat{x} = \mathcal{F}(x)$ is defined by

$$\hat{x}(z) = \sum_{t \in \mathbb{Z}} z^{-t} x[t],$$

is a norm-preserving bijection between ℓ^2 and L^2. For real Hilbert spaces X, Y, and a bounded linear operator $A : X \to Y$, $A' : Y \to X$ denotes adjoint operator of A.

[1] Omer Tanovic and Alexandre Megretski are with the Laboratory for Information and Decision Systems, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA {otanovic,ameg}@mit.edu

II. INTRODUCTION

Keeping the maximal time domain magnitude of the output reasonably small is a challenge in many signal processing tasks, especially those involving sharp spectral constraints. For example, the *ideal* linear time invariant (LTI) filters (e.g., the "ideal low-pass" filter with frequency response $H_b(\omega) = 1$ for $|\omega| < b$, $H_b(\omega) = 0$ for $b < |\omega| \leq \pi$, where $b \in (0, \pi)$) have infinite peak-to-peak gains, in the sense that a bounded input may generate an unbounded response. A real-life implementation of such filter is likely to be an approximation of the ideal case, with a finite peak-to-peak gain, but the trade-off between the sharpness of the frequency response, and maximal output magnitude, remains a limitation to be respected.

Since the peak-to-peak gain of an LTI filter equals the *L1 norm* of its unit sample response, minimizing the L1 norm is a natural objective in many linear filter optimization problems. There are many seemingly different problems in vast areas of engineering/applied math, that can be formulated and tackled in the framework of L1 optimization. Examples of such problems are [1] in control, [2] in machine learning, [3] in approximation theory, and well known series of papers [4]-[7] in areas of information theory and signal processing. Motivated by the knowledge that L1 norm minimization is a common approach to promoting sparsity of the optimal solutions, this paper aims to establish universal bounds on sparsity (interpreted as the number of non-zero coefficients) of the filters optimized with an L1 norm minimization objective.

The specific linear filter design problem in this paper is motivated by the setup shown in Figure 1, depicting the block diagram of a simplified digital communications transmitter circuit. Sequence of symbols $w = w[n]$, which are drawn from some compact set of complex numbers, is first upsampled by m (where $m > 1$ is a fixed integer) to produce signal $u = u[n]$, which is then fed into an LTI system $\mathbf{H}$. This system, called the shaping filter, should retain the original samples from w, and modify the inserted zero samples so that the output signal $v = v[n]$ has a lowpass characteristic. Therefore filter $\mathbf{H}$ should satisfy the *no inter-symbol interference* (no ISI) condition [8], and is modelled as a linear phase low-pass Nyquist-m filter [9]. Output v of $\mathbf{H}$ is first converted to analog domain, and then modulated onto a carrier to produce signal $s(t)$. This signal, before being sent to an antenna, is amplified by a power amplifier circuit $\mathbf{PA}$. Power efficiency, roughly defined as a ratio of transmitted signal energy to DC power of the PA, is one of the most important characteristics of a transmitter circuit, and there has been significant research effort in the area of power efficiency enhancement (e.g. [10]-[12]). This is even more important in modern mobile communication standards like *Long Term Evolution (LTE)*, which employ *Orthogonal Frequency Division Multiplexing (OFDM)* multicarrier modulation scheme and are prone to high *Peak-to-Average Power Ratio (PAPR)* [8]. High value of PAPR represents the main drawback of OFDM for the forthcoming wideband communication systems, and methods for its reduction have been studied extensively in the recent decade (e.g. [13]-[15]). One way of decreasing PAPR, and correspondignly enhancing power efficiency, is by decreasing the peak power of the transmitted signal, or equivalently by decreasing the maximal envelope amplitude $\|v\|_\infty$. This can be achieved by designing shaping filters with minimal peak-to-peak gains.

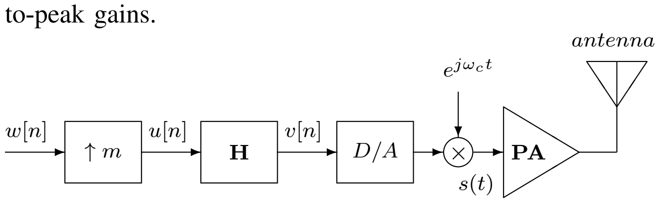

Fig. 1. Simplified block diagram of a transmitter circuit

In this paper we show that the problem of minimal peak-to-peak gain filter design can be formulated as an L1 minimization problem in an infinite dimensional Hilbert space, subject to a convex quadratic constraint. We show that, in the case of strictly feasible constraints, every optimal solution has a finite number of non-zero components, and therefore the corresponding optimal filter is a finite unit sample response (FIR) filter. Most importantly, we also give a constructive upper bound on the number of nonzero coefficients of the optimal filter.

The paper is organized as follows. In Section III, we introduce an "abstract" formulation of the L1 optimization problem under consideration, establish that a minimal output amplitude shaping filter design problem can be viewed as a special case of the abstract setup. In Section IV we provide standard background results regarding L1 optimization. In section V we use these results to formulate the main general statement of the paper: a bound on the number of non-zero coefficients in general L1 optimization. In section V— this bound is applied in the special case of minimizing output peak amplitude of a shaping filter.

III. PROBLEM FORMULATION: ABSTRACT SETUP AND APPLICATIONS

The optimization problems considered in this paper fall into the following general format.

Problem $\mathbb{P} = \mathbb{P}(X, e, Y, A, B, r)$: *given real Hilbert spaces* X, Y*, an orthonormal basis* $e = \{e_k\}_{k=1}^\infty$ *in* X*, a bounded linear operator* $A : X \to Y$*, an element* $B \in Y$*, and a positive real number* r*, minimize the functional* $J_e : X \to [0, \infty]$ *defined by*

$$J_e(x) = \sum_{k=1}^{\infty} |(e_k, x)_X|$$

on the set

$$\Omega_{A,B,r} = \{x \in X : |Ax - B| \leq r\}.$$

Here the word "minimize" means "find the arguments of minimum", i.e., vectors $v \in \Omega_{A,B,r}$ such that $J_e(v) \leq J_e(x)$ for all $x \in \Omega_{A,B,r}$, whenever such v do exist. More accurately, the paper is not really about "finding" arguments of minimum, but rather establishing certain properties of those.

One application context which leads to meaningful instances of problem $\mathbb{P}$ is that of *spectral shaping by upsampling*. Specifically, given an integer $m > 1$ and real numbers $\omega_0 \in (\pi/m, \pi)$ and $r > 0$, consider the task of designing a linear discrete time (DT) system S which converts input signals $u = u[\cdot] \in \ell^2$ to output signals $y = y[\cdot] \in \ell^2$ in such a way that

(a) for every $u \in \ell^2$ the output $y = S(u)$, satisfies $y[mk] = u[k]$ for all k;
(b) the relation between u and $y = S(u)$ is m-*shift invariant*, in the sense that $y_1[t] = y[t+m]$ for all $t \in \mathbb{Z}$ whenever $y_1 = S(u_1)$, $y = S(u)$, and $u_1[k] = u[k+1]$ for all $k \in \mathbb{Z}$;
(c) the relation between u and $y = S(u)$ has *finite peak-to-peak gain* γ_S, in the sense that

$$\gamma_S = \sup\left\{\frac{\|S(u)\|_\infty}{\|u\|_\infty} : \quad u \in \ell^2, u \neq 0\right\} < \infty;$$

(d) the spectrum of $h = S(\delta)$ is mostly limited to $[-\omega_0, \omega_0]$, in the sense that $E_0(h) \leq r^2$, where

$$E_0(h) = \frac{1}{\pi}\int_{\omega_0}^{\pi} |\hat{h}(e^{j\omega})|^2 d\omega, \quad \hat{h} = \mathcal{F}(h); \qquad (1)$$

(e) the peak-to-peak gain γ_S of S is to be made as small as possible.

It is now clear that system S can be seen as a series interconnection of an upsampling by m system, and a linear time invariant (LTI) system T, whose unit sample response $h = T(\delta)$ is completely characterized by conditions (a) and (c):

1) condition (a) means that $h[0] = 1$ and $h[km] = 0$ for all integer $k \neq 0$;
2) condition (c) means $\gamma_T = \|h\|_1 < \infty$.

LTI system T is in signal processing literature know as the linear phase Niquist-m filter [9].
Given these observations, the *spectral shaping by upsampling* task characterized by (a)-(e) can be viewed as a special case of the abstract problem $\mathbb{P}(X, e, Y, A, B, r)$, with X, e, Y, A, B, r defined as follows:

- $X = \{x \in \ell^2 : x[mk] = 0, x[k] = x[-k] \quad \forall\ k \in \mathbb{Z}\}$;

- $e = \{e_k\}_{k=1}^{\infty}$ is a basis in X formed by shifts of the unit sample sequence δ, for example,

$$e_k[t] = \frac{\delta[t-k] + \delta[t+k]}{\sqrt{2}}, \; k = 1, 2, \dots$$

- $Y = L^2$;
- $A: \; X \to Y$ is the restriction of $A_0: \; \ell^2 \to L^2$ to X, where A_0 maps $x \in \ell^2$ to

$$\hat{y}(z) = \begin{cases} \hat{x}(z), & z = e^{j\omega}, \; |\omega| \in [\omega_0, \pi], \\ 0, & z = e^{j\omega}, \; |\omega| < \omega_0, \end{cases}$$

and $\hat{x} = \mathcal{F}(x)$;
- $B = -A_0\delta$.

Indeed, in this case $h = x + \delta$, with x ranging over X, spans all admissible vectors h, and, in addition, the identities

$$|Ax - B|^2 = E_0(x+\delta), \quad \gamma_S = \|x+\delta\|_1 = 1 + J_e(x)$$

hold for all $x \in X$, $h = x + \delta$.

IV. L1 Optimization Background

In this section we provide some fairly standard background results regarding the general case of problem $\mathbb{P}$: the use of weak compactness to establish existence of optimum, and the use of Hahn-Banach theorem to establish necessary (and sufficient) conditions of optimality in convex optimization.

The first theorem extablishes existence of an argument of minimum in every feasible problem $\mathbb{P}(X, e, Y, A, B, r)$.

Theorem 1: An argument of minimum in problem $\mathbb{P}(X, e, Y, A, B, r)$ exists whenever the set $\Omega_{A,B,r}$ is not empty.

Proof.

Let $\gamma = \inf\{J_e(x) : \;\; X \in \Omega_{A,B,r}\}$. If $\gamma = \infty$, every $x \in \Omega_{A,B,r}$ is an argument of minimum of J_e on $\omega_{A,B,r}$. If $\gamma < \infty$, there exists a sequence $\{x_i\}_{i=0}^{\infty}$ of $x_i \in \Omega_{A,B,r}$ such that $J(x_i) \le \gamma + 1/i$. Since

$$|x|_X^2 = \sum_{k=1}^{\infty} |(e_k, x)|^2 \le \left(\sum_{k=1}^{\infty} |(e_k, x)| \right)^2 = J(x_i)^2,$$

$\{x_i\}_{i=0}^{\infty}$ is uniformly bounded in ℓ^2, and hence has a subsequence $\{\tilde{x}_j\}_{j=0}^{\infty}$, which converges weakly to a limit $v \in \ell^2$, in the sense that $(w, x_j)_X \to (w, v)$ as $j \to \infty$ for every $w \in \ell^2$. Then

$$|Av - B| = \sup_{|u|=1} (u, Av - B) \le \lim_{j \to \infty} \sup_{|u|=1} (u, Ax_i - B) \le r,$$

$$\|v\|_1 = \sum_i \lim_{j \to \infty} |(e_i, x_j)| \le \lim_{j \to \infty} \sum_i |(e_i, x_j)| = \gamma,$$

and hence v is an argument of minimum of J_e on $\omega_{A,B,r}$. □

For a real number a, let sgn(a) denote the set

$$\text{sgn}(a) = \begin{cases} \{a/|a|\}, & a \ne 0, \\ [-1, 1], & a = 0. \end{cases}$$

For an orthonormal basis $e = \{e_i\}_{i=1}^{\infty}$ in a real Hilbert space X, and an element $x \in X$, let $\sigma_e(x)$ denote the set

$$\sigma_e(x) = \{u \in X : \quad (e_i, u)_X \in \text{sgn}((e_i, x)_X) \; \forall \; i \in \mathbb{N}\}.$$

Note that $\sigma_e(x)$ is empty unless $(e_i, x) = 0$ for all but a finite number of indices i.

The following theorem establishes standard Lagrange multipliers-based necessary and sufficient conditions of optimality in a problem $\mathbb{P}(X, e, Y, A, B, r)$ with a finite maximal lower bound.

Theorem 2: For an instance of problem $\mathbb{P}(X, e, Y, A, B, r)$, assume that

(i) $\inf\{|Ax - B|_Y : \; x \in X\} < r$ (strict feasibility);
(ii) $|B|_Y > r$ ($x = 0$ is not an argument of minimum).

Then for every $v \in X$ the following conditions are equivalent:

(a) v is an argument of minimum in problem $\mathbb{P}(X, e, Y, A, B, r)$;
(b) $|B - Av| = r$, and there exists $\lambda \in (0, \infty)$ such that

$$\lambda r^{-1} A'(B - Av) \in \sigma_e(v). \qquad (2)$$

Proof (a)⇒(b).

Step 1. Consider the subset Λ of $\mathbb{R} \times \mathbb{R}$ defined by

$$\Lambda = \{(|Ax - B|_Y + a, \;\; J_e(x) + b) : \; x \in X, \;\; a, b > 0\}.$$

Since the functions$x \mapsto |Ax - B|_Y$ and $x \mapsto J_e(x)$ are convex on X, the set Λ is convex, too. Since, by assumption, the point $(r, J_e(v))$ is not in Λ, the Hahn-Banach theorem guarantees existence of a line separating Λ from $(r, J_e(v))$, which means existence of real numbers p, q such that $|p| + |q| > 0$ and $z_1 p + z_2 q \ge rp + J_e(v)q$ for all $(z_1, z_2) \in \Lambda$, or, equivalently,

$$(|Ax - B|_Y + a)p + (J_e(x) + b)q \ge rp + J_e(v)q$$

for all $x \in X$, $a, b \in (0, \infty)$. Taking into account the form of dependence on a and b, the inequality implies $p \ge 0$ and $q \ge 0$, and allows one to re-write the condition as

$$p(|Ax - B|_Y - r) + q(J_e(x) - J_e(v)) \ge 0 \; \forall \; x \in X. \quad (3)$$

Step 2. Note that (3) should imply both $p > 0$ and $q > 0$. Indeed, if $p = 0$ then $q > 0$, and (3) implies $J_e(x) \ge J_e(v)$ for *all* $x \in X$ such that $J_e(x) < \infty$, which, is clearly not true for $x = 0$, as, due to assumption (ii), $J_e(v) > 0$. Similarly, if $q = 0$ then $p > 0$, and (3) implies $|Ax - B|_Y \ge r$ for *all* $x \in X$ such that $J_e(x) < \infty$, which is not true because of assumption (i). Hence, $\lambda = p/q$ is a positive real number.

Step 3. Substituting $x = v$ into (3) (which is possible because assumption (i) implies $J_e(v) < \infty$) yields $p|Av - B|_Y \ge pr$ which, combined with $|Av - B| \le r$, yields $|Av - B| = r$.

Step 4. Substituting $x = v \pm tce_i$, where $c \in \mathbb{R}$, into (3), and using the directional differentiation identities

$$\lim_{t \to 0+} \frac{|A(v + tce_i) - B| - |Av - B|}{t} = c\frac{(e_i, A'(Av - B))}{|Av - B|},$$

$$\lim_{t\to 0+} \frac{J_e(v+tce_i)-J_e(v)}{t} = \begin{cases} c, & (e_i,v)>0,\\ -c, & (e_i,v)<0,\\ |c|, & (e_i,v)=0,\end{cases}$$

we conclude that

$$\lambda r^{-1}(e_i, A'(Av-B))_X \in \mathrm{sgn}((e_i,v)_X) \quad \forall\, i\in\mathbb{N},$$

which by definition is equivalent to (2).
□

Proof (b)⇒(a).
Using the inequalities

$$|f|_Y - |g|_Y \geq (f-g,g)_Y/|g|_Y \quad \forall\, g\neq 0,$$

$$J_e(x)-J_e(v) \geq (x-v,u)_X \quad \forall\, u\in\sigma_e(v),$$

we conclude that (2) implies

$$\lambda(|Ax-B|-|Av-B|)+J_e(x)-J_e(v)\geq 0 \quad \forall\, x\in X.$$

Since $|Ax-B|\leq r=|Av-B|$ for all $x\in\Omega_{A,B,r}$, this implies $J_e(x)\geq J_e(v)$ for all $x\in\Omega_{A,B,r}$, which proves that v is an argument of minimum in $\mathbb{P}(X,e,Y,A,B,r)$.
□

V. ESTIMATING THE NUMBER OF NON-ZERO ELEMENTS IN L1-OPTIMAL SOLUTION

For an orthonormal basis $e=\{e_k\}_{k=1}^{\infty}$ in a real Hilbert space X, and an element $x\in X$, let $N_e(x)\in\mathbb{N}\cup\{0,\infty\}$ denote the number of indices $k\in\mathbb{N}$ for which $(e_k,x)\neq 0$.

Theorem 3: Given real Hilbert spaces X, Y, an orthonormal basis $e=\{e_k\}_{k=1}^{\infty}$ in X, a bounded linear operator $A:\ X\to Y$, and an element $B\in Y$, let $r_0=\inf\{|Ax-B|:\ x\in X\}$. For $r>r_0$ let $\phi(r)$ be the minimum of J_e on $\Omega_{A,B,r}$. The function $\phi:\ (r_0,\infty)\to\mathbb{R}$ is convex, monotonically non-increasing, and, for every $r>r_0$, every argument of minimum v of J_e on $\Omega_{A,B,r}$ satisfies

$$N_e(v)^{1/2}\leq \|A\| \inf_{r_0<a<b\leq r}\frac{\phi(a)-\phi(b)}{b-a}. \tag{4}$$

Proof. Convexity of $\phi(\cdot)$ is implied by the convexity of the set Λ from the proof of Theorem 2. As was shown there, the minimizer v must satisfy the optimality condition $\lambda r^{-1}A'(B-Av)\in\sigma_e(v)$, where $-\lambda$ is the slope of the line separating Λ from $(r,\phi(r))$, and hence is such that $\phi(a)\geq\phi(r)+(r-a)\lambda$ for all $a>r_0$. By convexity of $\phi(\cdot)$, we have

$$\lambda\leq \inf_{r_0<a<b\leq r}\frac{\phi(a)-\phi(b)}{b-a}.$$

By the definition of $\sigma_e(\cdot)$, we have $N_e(v)\leq|u|^2$ for every $u\in\sigma_e(v)$. Hence

$$N(v)^{1/2}\leq\lambda r^{-1}|A'(B-Av)|\leq\lambda r^{-1}\|A\|r=\|A\|\lambda,$$

which implies (4).
□

In particular, for every $u\in X$ such that $|Au-B|=a<b=r$ we have $\phi(a)\leq J_e(u)$ and $\phi(r)\geq 0$, and Theorem 3 yields

$$N(v)^{1/2}\leq\frac{J_e(u)\|A\|}{r-|Au-B|}.$$

VI. MINIMAL PEAK-TO-PEAK GAIN SHAPING FILTER

In Sections II and III we introduced the problem of minimizing peak-to-peak gain of a linear phase low-pass Nyquist-m shaping filter, and argued that it leads to an instance of problem $\mathbb{P}$. Theorem 3 then implies that the optimal solution, i.e. the unit sample response of the filter, has finite number of non-zero samples, which means that the optimal filter is an FIR filter. In this section we show this for $m=2$, and compute an upper bound on the number of non-zero samples of the optimal filter.

Let $S\subset\mathbb{C}$ be a given compact set, such that for every $z\in S$, $-z$ and $\bar z$ are also in S. Let U be the subset of bounded sequences $u:\mathbb{Z}\to S\cup\{0\}$, such that $u[2t+1]=0$ and $u[2k]\neq 0$, for all $t\in\mathbb{Z}$.

Let $X=\{x\in\ell^2:\ x[2k]=0, x[k]=x[-k]\quad\forall\, k\in\mathbb{Z}\}$, be a separable Hilbert space with orthonormal basis $e=\{e_k\}_{k=1}^{\infty}$, defined by $e_k[t]=\frac{1}{\sqrt{2}}(\delta[t-2k+1]+\delta[t+2k-1])$, and let $Y=\{x\in\ell^2:\ x[k]=x[-k]\quad\forall\, k\in\mathbb{Z}\}$. Let $\hat Y$ be the separable Hilbert space of square integrable, 2π-periodic, real valued, and even functions, with a scalar product $(\cdot,\cdot)_{\hat Y}:\hat Y\times\hat Y\to\mathbb{R}$ defined by

$$(\hat f,\hat g)_{\hat Y}=\frac{1}{\pi}\int_0^{\pi}\hat f(\omega)\hat g(\omega)d\omega.$$

Let $\hat X=\{\hat f\in\hat Y:\ \hat f(\omega+\pi)=-\hat f(\omega)\}$. $\hat X$ is clearly a subspace of $\hat Y$. It is straightforward to check that the sequence $\{\hat e_k\}_{k=1}^{\infty}$, with $\hat e_k=\hat e_k(z)=\sqrt{2}\cos(2k-1)\omega$, is an orthonormal basis in $\hat X$. It can be easily seen that the Fourier transform as defined in Section I, when restricted to X and Y, is a norm-preserving bijection between X and $\hat X$, and Y and $\hat Y$, respectively.

Let T be a linear phase low-pass Nyquist-2 filter, as defined in Section II. The unit sample response h of T can be expressed as $h=x+\delta$, with $x=\{x_k\}_{k=1}^{\infty}\in X$, and $x_k=(e_k,x)_X$. This implies that the frequency response $\hat h$ of T can be expressed as

$$\mathcal{F}(h)=\hat h(e^{j\omega})=1+\sqrt{2}\sum_{k=1}^{\infty}x_k\cos{(2k-1)\omega}. \tag{5}$$

The spectrum of T should be mostly limited to $[-\omega_0,\omega_0]$, where $\omega_0\in(\pi/2,\pi)$, in the sense that

$$E_0(h)=\frac{1}{\pi}\int_{\omega_0}^{\pi}|\hat h(e^{j\omega})|^2d\omega\leq r^2, \tag{6}$$

where r is some small positive real number. This, and similar spectral mask conditions, are known as Adjacent Channel Power Ratio (ACPR) conditions [8].

The objective is to find h such that the supremum $\sup_{u\in U}\|h*u\|_\infty$ is minimized. Here $h*u$ denotes discrete-time convolution of h and u, i.e. $(h*u)[n]=\sum_{k=-\infty}^{+\infty}u[k]\cdot h[n-k]$. Due to the nature of set S, the above criterion function is equivalent to $\gamma_T=1+\sqrt{2}J_e(x)$, therefore the optimization problem becomes the one of minimizing $J_e(x)$.

Let $A:X\to\hat{Y}$ be the restriction of $A_0:Y\to\hat{Y}$ on X, where A_0 maps $x\in Y$ to $2\hat{\alpha}(z)\hat{x}(z)\in\hat{Y}$, with

$$\hat{\alpha}(z)=\begin{cases}1, & z=e^{j\omega},\ |\omega|\in[\omega_0,\pi],\\ 0, & z=e^{j\omega},\ |\omega|<\omega_0.\end{cases}$$

and $\hat{x}\in\hat{Y}$ is the Fourier transform of x. Also let $B=-A_0\delta=-\hat{\alpha}$. With $\hat{h}(z)$ as given in (5), the ACPR constraint in (6) can be re-written as

$$|Ax-B|\le r. \tag{7}$$

It clearly follows that the problem of finding minimal peak-to-peak gain linear-phase low-pass Nyquist-2 filter can be formulated as an instance of the optimization problem $\mathbb{P}$:

$$\begin{aligned}\min_{x\in X}\quad & J_e(x)\\ \text{s.t.}\quad & |Ax-B|\le r,\end{aligned} \tag{8}$$

where A and B are as defined above.

Theorem 3 then implies that the optimal x has finite number of non-zero elements, and hence the optimal T is an FIR filter. Let us compute an upper bound on the number of non-zero elements of the optimal solution of (8) using inequality (4) from Theorem 3.

First we prove the following lemma:

Lemma 4: Operator $Q=A'A:X\to X$ is a projection.

Proof. As earlier noted, operator $A:X\to\hat{Y}$ maps f to $2\hat{f}(z)\hat{\alpha}(z)$. Let $P:\hat{Y}\to\hat{X}$ be a projection, i.e. $(P\hat{f})(z)=\frac{1}{2}(\hat{f}(z)-\hat{f}(z\cdot e^{j\pi}))$, and $\mathcal{F}^{-1}:\hat{X}\to X$ restriction on $\hat{X}$ of the inverse Fourier transform. Therefore the adjoint $A':\hat{Y}\to X$ of A can be written as $A'=\mathcal{F}^{-1}P$. Now for any $f\in X$, the following sequence of mappings holds

$$f\overset{A}{\mapsto}2\hat{f}(z)\hat{\alpha}(z)\overset{P}{\mapsto}$$
$$\overset{P}{\mapsto}\hat{f}(z)\hat{\alpha}(z)-\hat{f}(z\cdot e^{j\pi})\hat{\alpha}(z\cdot e^{j\pi})=\hat{g}(z).$$

Since $f\in X$, it follows that

$$\hat{g}(z)=\hat{f}(z)[\hat{\alpha}(z)+\hat{\alpha}(z\cdot e^{j\pi})].$$

Thus $Qf=\mathcal{F}^{-1}(\hat{g})=g$. Now if we apply Q on g again, we get

$$g\overset{A}{\mapsto}2\hat{g}(z)\hat{\alpha}(z)\overset{P}{\mapsto}\hat{g}(z)\overset{\mathcal{F}^{-1}}{\mapsto}g,$$

since $\hat{g}(z)\hat{\alpha}(z)=\hat{f}(z)\hat{\alpha}(z)$ due to $\hat{\alpha}(z)^2=\hat{\alpha}(z)$, and $\hat{\alpha}(z)\hat{\alpha}(z\cdot e^{j\pi})=0$ for all $z=e^{j\omega},\omega\in\mathbb{R}$. Therefore $Q^2=Q$, i.e. operator Q is a projection. □

It immediately follows from Lemma 4 that $\|A\|=1$.

Let $u\in X$ such that

$$\hat{u}=\hat{u}(z)=\begin{cases}\frac{1}{\sqrt{2}}, & z=e^{j\omega},|\omega|\le\pi-\omega_0,\\ \frac{\sqrt{2}}{2\omega_0-\pi}(|\omega|-\frac{\pi}{2}), & z=e^{j\omega},\pi-\omega_0<|\omega|<\omega_0,\\ -\frac{1}{\sqrt{2}}, & z=e^{j\omega},\omega_0\le|\omega|\le\pi.\end{cases}$$

It is easy to see that $Au=B$, and componenets $(e_k,u)_X$ are given by

$$(e_k,u)_X=\frac{8}{\pi(2\omega_0-\pi)}\cdot\frac{\cos(2k-1)\omega_0+(-1)^k}{(2k-1)^2}.$$

Value of the functional $J_e(\cdot)$ evaluated at u is upper bounded by

$$J_e(u)\le\frac{16}{\pi(2\omega_0-\pi)}\cdot\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}=\frac{2\pi}{2\omega_0-\pi}.$$

With signal u as given above, Theorem 3 yields an upper bound on the number of non-zero coefficients of the optimal unit sample response h:

$$N(h)=2N(x)+1\le r^{-2}\cdot\frac{8\pi^2}{(2\omega_0-\pi)^2}+1.$$

VII. CONCLUSIONS

In this paper we showed that the problem of minimal peak-to-peak gain filter design can be formulated as an L1 minimization problem in an infinite dimensional Hilbert space, subject to a convex quadratic constraint. We showed that, in the case of strictly feasible constraints, every optimal solution has a finite number of non-zero components, and therefore the optimal filter is an FIR filter. Most importantly, we also gave a constructive upper bound on the number of nonzero coefficients of the optimal filter.

REFERENCES

[1] M.A. Dahleh and J., Jr. Pearson, "ℓ^1 -optimal feedback controllers for MIMO discrete-time systems," *Automatic Control, IEEE Transactions on* , vol. 32, no. 4, pp. 314-322, Apr 1987.

[2] R. Tibshirani, "Regression shrinkage and selection via the Lasso", *Journal of the Royal Statistical Society*, Series B. 58, pp. 267-288, 1996.

[3] H. Rauhut and R. Ward, Sparse Legendre expansions via ℓ_1-minimization, *Journal of Approximation Theory*, vol. 164, no. 5, pp. 517-533, May 2012.

[4] E.J. Candes and T. Tao, "Decoding by linear programming", *IEEE. Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203-4215, December 2005

[5] E.J. Candes, J. Romberg, and T. Tao, "Signal recovery from incomplete and inaccurate measurements", *Comm. Pure Appl. Math.*, 59:8, pp. 1207-1223, 2005.

[6] E.J. Candes and T. Tao, "Near-optimal signal recovery from random projections and universal encoding strategies", *IEEE. Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406-5425, December 2006

[7] E.J. Candes, "Compressed sensing", *IEEE. Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006

[8] J. G. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill, 2007

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

[10] S. L. Miller and R. J. O'Dea, "Peak power and bandwidth efficient linear modulation," *IEEE Trans. Comm.*, vol. 46, no. 12, pp. 1639-1648, Dec 1998.

[11] I. Kim, Y. Y. Woo, J. Kim, J. Moon, J. Kim, B. Kim, "High-Efficiency Hybrid EER Transmitter Using Optimized Power Amplifier," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 11, pp. 2582-2593, Nov. 2008.

[12] S. Boumaiza, "Advanced techniques for enhancing wireless RF transmitters' power efficiency", *Microelectronics, 2008. ICM 2008. International Conference on*, pp.68-73, 14-17 Dec. 2008.

[13] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Commun.*, vol. 52, pp. 5-65, March 2005

[14] S. H. Han and J. H. Lee, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," *IEEE Trans. Broad.*, vol. 54, no. 2, pp. 257-268, June 2008

[15] E.B. AL-Safadi and T. Y. AL-Naffouri, "Peak Reduction and CLipping Mitigation in OFDM by Augmented Compressive Sensing," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3834-3839, July 2012.