Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-20 (2018) 147–152

Real-Time Realization of a Family of Optimal Infinite-Memory Non-Causal Systems ⋆

Omer Tanovic * **Alexandre Megretski** *

* *Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: otanovic@mit.edu, ameg@mit.edu).*

Abstract: In this paper, we consider a problem of designing discrete-time systems which are optimal in frequency-weighted least squares sense subject to a maximal output amplitude constraint. It can be shown for such problems, in general, that the optimality conditions do not provide an explicit way of generating the optimal output as a real-time implementable transformation of the input, due to instability of the resulting dynamical equations and sequential nature in which criterion function is revealed over time. In this paper, we show that, under some mild assumptions, the optimal system has exponentially fading memory. We then propose a causal and stable finite-dimensional nonlinear system which, under an L1 dominance assumption about the equation coefficients, returns high-quality approximations to the optimal solution. The fading memory of the optimal system justifies the receding horizon assumption and suggests that such approach can serve as a cheaper alternative to standard MPC-based algorithms. The result is illustrated on a problem of minimizing peak-to-average-power ratio of a communication signal, stemming from power-efficient transceiver design in modern digital communication systems.

Keywords: Model approximation, non-causal systems, non-linear systems, model predictive control, linear optimal control, state space realization.

1. INTRODUCTION

Quadratic programs (QP) with box constraints (or inequality constraints in general) have over years attracted attention by researches from various communities. For instance, many physics and engineering problems can be formulated as QPs with inequality constraints, e.g. flow through a porous medium (see Lin and Cryer (1985)), soil moisture in hydrology (see Mead and Renaut (2010)), modeling of the ocean circulation (see Oreborn (1986)), support vector machines (see Osuna et al. (1997)), constrained linear quadratic optimal control (see Goodwin et al. (2005)), etc.

Various methods have been proposed for solving box constrained QP in a finite-dimensional setting: active set methods (see Moré and Toraldo (1989), Kunisch and Rendl (2003), Hungerländer and Rendl (2015)), gradient projection and conjugate gradients (see Bertsekas (1976); Dembo and Tulowitzki (1983)), Newton iteration (see Li and Swetits (1993)), primal-dual methods (see Pardalos et al. (1990)), etc. Such methods commonly rely on computer-aided optimization solvers and require non-negligible computation power, which makes them unfavorable in applications that have strict power budget. The infinite-dimensional bound-constraint quadratic programs are even more computationally demanding. An important instance is the infinite-horizon linear quadratic regulation problem (LQR) with bounded control. This problem is mostly addressed approximately, where model predictive control (MPC) has probably been the most popular method for approximately solving infinite-horizon constrained LQR. Such MPC schemes rely on replacing infinite-horizon with a receding (i.e. finite) one, where, in general, an easier finite-dimensional optimization problem is resolved at every time instance, see e.g., Sznaier and Damborg (1987); Chmielewski and Manousiouthakis (1996); Grieder et al. (2004); Stathopoulos et al. (2017).

Another instance of the infinite-dimensional setup emerges when one wants to design discrete-time systems which are optimal in the sense of some frequency-weighted least squares criterion subject to maximal output amplitude constraints. In particular, such optimization problems serve to represent a number of peak-to-average-power ratio (PAPR) reduction objectives which are of significant importance in modern communication systems, see e.g., Proakis and Salehi (2007); Reine and Zang (2013); Sochacki (2016). It is known for such optimization problems that, in general, the optimality conditions do not provide an explicit way of generating the optimal output as a real-time implementable transformation of the input. This is due to instability of the resulting dynamical equations as well as sequential nature in which criterion function is revealed over time (the quadratic functional to be minimized is commonly a function of the input signal, which comes sequentially over time). Therefore, at each time instance, the knowledge of the whole history of the

⋆ This work was supported by the National Science Foundation under award number 1743938.

10.1016/j.ifacol.2018.11.007

input signal should be known ahead of time in order to calculate the current sample of the optimal output signal. Due to difficulties in obtaining an explicit optimal solution, receding horizon optimization, i.e. model predictive control, appears to be a natural way of addressing these problems. Unfortunately, the high cost associated with MPC computations at every time step makes it unfavorable in power and time-sensitive applications such as those in signal processing for communication systems.

In this paper, we show that, under some mild assumptions, the optimal map uniquely defines a discrete-time system with exponentially fading memory. We then propose a real-time realizable algorithm which, under an L1 dominance assumption about the equation coefficients, returns high-quality approximations to the optimal solution. The algorithm exploits the optimality conditions and is realized as a causal and stable finite-memory nonlinear discrete-time system, and is allowed to look ahead at the input signal over a finite horizon. Fading memory of the optimal system justifies the finite horizon assumption and suggests that such approach can serve as a cheaper alternative to standard MPC-based algorithms, since it does not rely on resolving an optimization problem at every time instant. We illustrate this result on a problem of minimizing peak-to-average-power ratio of a communication signal, which pertains to power-efficient transceiver design in modern digital communication systems.

2. 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 $x : \ \mathbb{Z} \to \mathbb{C}$, interpreted as *discrete-time (DT) signals*, with $x(t)$ used for the value of x at $t \in \mathbb{Z}$. For $x \in \ell$, the *L1 norm* $\|x\|_1 \in [0, \infty]$, the *L2 norm* $\|x\| \in [0, \infty]$, and the *L-Infinity norm* $\|x\|_\infty \in [0, \infty]$ are defined by

$$\|x\|_1 = \sum_t |x(t)|, \ \|x\|_\infty = \sup_t |x(t)|, \ |x| = \left(\sum_t |x(t)|^2\right)^{\frac{1}{2}}.$$

ℓ^1 and ℓ^∞ are the subsets of absolutely summable and bounded signals from ℓ, respectively, treated as Banach spaces, with norms $||x||_1$ and $||x||_\infty$ as defined above. ℓ^2 is the subset of finite energy signals from ℓ, treated as a Hilbert space, with the norm $|x|$ defined above. $\{e_i\}_{i=-\infty}^{\infty}$ such that $e_i(t) = 1$ for $t = i$ and $e_i(t) = 0$ otherwise, is the standard orthonormal basis in the above defined spaces.

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 $X = \mathcal{F}(x)$ is defined by

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

is a norm-preserving bijection between ℓ^2 and L^2. In the rest of this paper, for simplicity, we will abuse notation by writing $X(\omega)$, instead of $X(e^{j\omega})$, to denote the value of $X = \mathcal{F}(x)$ at $z = e^{j\omega}$. For a bounded linear operator $A : \ell^2 \to \ell^2$, $A' : \ell^2 \to \ell^2$ denotes adjoint operator of A.

For a positive real number r, let function $\text{sat}_r : \mathbb{C} \to [-r, r]$ be defined by

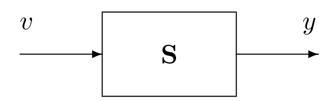
$$\text{sat}_r(\xi) = \begin{cases} \xi, & |\xi| < r \\ r\xi/|\xi|, & |\xi| \geq r \end{cases}.$$

Similarly, let $\text{Sat}_r : \ell^2 \to \ell^2$ be defined by

$$y = \text{Sat}_r(x) \Leftrightarrow (e_i, y) = \text{sat}_r((e_i, x)), \forall i \in \mathbb{Z}.$$

3. PROBLEM FORMULATION

We aim to optimize and implement efficiently discrete-time signal processing systems with scalar input v and scalar output y:



where the output $y = \mathbf{S}v$ is expected to be optimal, in the sense of minimizing a certain objective defined in terms of input v. Let $r > 0$, and let $\alpha : \mathbb{R} \to \mathbb{C}$ and $\beta : \mathbb{R} \to \mathbb{C}$ be trigonometric polynomials mapping $\omega \in \mathbb{R}$ to $\alpha(\omega) > 1$ and $\beta(\omega)$, respectively. For every discrete-time signal $v \in \ell^2$, the scalar signal $y = \mathbf{S}v \in \ell^2$ should have samples $|y(t)| \leq r$, and minimize the functional

$$J_{\alpha,\beta}(v, y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y(\omega)' \alpha(\omega) Y(\omega) d\omega - \frac{1}{\pi} \int_{-\pi}^{\pi} \text{Re}\{Y(\omega)' \beta(\omega) V(\omega)\} d\omega \quad (1)$$

where $V = V(\omega)$ and $Y = Y(\omega)$ are the Fourier transforms of v and y, respectively. Therefore, we are trying to solve the time-domain-value-constrained frequency-weighted least squares optimization problem

$$\min_y J_{\alpha,\beta}(v, y), \quad \text{subject to } ||y||_\infty \leq r. \quad (2)$$

Remark: In fact, it is sufficient to assume that $\alpha(\omega) \geq \epsilon$, for some $\epsilon > 0$, since by proper scaling of $J_{\alpha,\beta}(v, y)$ we could arrive at an optimization problem equivalent to (2).

It is clear that (2) is a convex infinite-dimensional quadratic problem with box constraints, which is feasible (see the appendix) and has a unique solution due to strict positivity of α. Let $\mathbf{T}_\alpha$ and $\mathbf{T}_\beta$ be the finite unit sample response LTI systems with frequency responses $\alpha(\omega)$ and $\beta(\omega)$, respectively. The necessary and sufficient conditions of optimality of (2) can be written as (the proof is omitted due to space constraints):

$$y = \text{Sat}_r(\mathbf{T}_\beta v + y - \mathbf{T}_\alpha y). \quad (3)$$

Moreover, with $w = \mathbf{T}_\beta v$ and $\mathbf{H} = \mathbf{T}_\alpha - \mathbf{I}$, the above optimality condition can be written as

$$y = \text{Sat}_r(w - \mathbf{H}y). \quad (4)$$

Let $H = H(\omega)$ and $h = h(t)$ be the frequency response and unit sample response of $\mathbf{H}$, respectively, and let T be the order of trigonometric polynomial $\alpha = \alpha(\omega)$. We have that $H(\omega) = \alpha(\omega) - 1 > 0$ for all $\omega \in [0, 2\pi)$, and therefore $h(-t) = h(t)$ for all $t \in \mathbb{Z}$. Furthermore, $h(t) = 0$ for all $|t| > T$. The optimal condition (4) can now be written sample-wise as

$$y(t) = \text{sat}_r \left(w(t) - \sum_{\tau=-T}^{T} h(\tau) y(t+\tau) \right). \tag{5}$$

Due to strict convexity of problem (2), and hence uniqueness of the optimal solution, equation (5) defines a system which maps input signal w into output signal y. We denote this system as $\mathbf{S}^*$ and refer to it as "the optimal system", in the rest of this paper. It is clear from (5) that, in general, the optimal system $\mathbf{S}^*$ is nonlinear and noncausal. Moreover, the optimality condition (5) is not attractive as a description of a real-time implementable system $\mathbf{S}$ mapping w to the optimal y. In the following sections, we show that, with adequate non-linear stability analysis and careful structuring, the optimal system $\mathbf{S}^*$ can be approximately realized in the form of a finite memory real-time signal processing algorithm.

4. MAIN RESULTS

4.1 Fading Memory Property of the Optimal Solution

If one hopes to find a finite memory system which approximates (in some sense) the optimal system $\mathbf{S}^*$ then it would be reasonable to assume that $\mathbf{S}^*$ posesses some type of near-finite memory. In fact, it has been shown that fading memory is a sufficient condition for any time-invariant causal operator to be approximated by a nonlinear moving-average operator, see Boyd and Chua (1985). Since $\mathbf{S}^*$ is noncausal, we first need to define fading memory for non-causal discrete-time systems. The following definition of fading memory is an intuitive one: a system has fading memory if two input signals that are close in the recent past and future, but not necessarily close in the remote past and future, yield present outputs which are close. This is a generalization of a standard definition of fading memory for causal systems, see e.g. Boyd and Chua (1985).

For every integer $T > 0$ let $P_T : \ell^2 \to \ell^2$ be defined by $(P_T w)(t) = \begin{cases} w(t), & |t| \leq T \\ 0, & |t| > T \end{cases}$.

Definition 1. A noncausal discrete-time system $\mathbf{S}$ has fading memory on ℓ^2 if for all $r > 0$

$$\lim_{T \to \infty} \sup_{\substack{w_1, w_2 \in \ell^2, \\ P_T w_1 = P_T w_2, \\ ||w_1 - w_2||_\infty \leq r}} |y_1(0) - y_2(0)| = 0, \tag{6}$$

where $y_1 = \mathbf{S} w_1$ and $y_2 = \mathbf{S} w_2$. Furthermore, if there exist $\gamma > 0$ and $\epsilon > 0$ such that

$$|y_1(0) - y_2(0)| \leq \gamma ||w_1 - w_2||_\infty e^{-\epsilon T}, \tag{7}$$

for any $w_1, w_2 \in \ell^2$ such that $P_T w_1 = P_T w_2$, we say that $\mathbf{S}$ has exponentially fading memory.

We are now ready to state the theorem which establishes fading memory property of the optimal system $\mathbf{S}^*$.

Theorem 2. Suppose that $\mathbf{H}$ has strictly positive frequency response. Then the optimal system $\mathbf{S}^*$ has exponentially fading memory.

Proof. Omitted due to space constraints.■

4.2 A Real-Time Algorithm for Sequentially Calculating High-Quality Approximations to Optimal Solution

In this section we propose an algorithm for sequentially obtaining high-quality approximations to the optimal solution of problem (2). We first define what is meant by "approximate" systems. Let $\epsilon > 0$. System $\mathbf{T} : \ell^2 \to \ell^2$ is an ϵ-approximation to system $\mathbf{S} : \ell^2 \to \ell^2$ if

$$|\mathbf{S}w(t) - \mathbf{T}w(t)| < \epsilon ||w||_\infty, \quad \forall t \in \mathbb{Z}, \forall w \in \ell^2.$$

In addition to the positivity and finite impulse response properties, we also assume that $\mathbf{H}$ has an "L1 dominance" property in the sense that $\sum_{t=-T}^{T} |h(t)| < 1$.

As before, we assume that the optimal system $\mathbf{S}^*$ maps signal $w \in \ell^2$ to $y \in \ell^2$, as defined by (5). For a given integer $m > T$, let matrices $\hat{A} \in \mathbb{R}^{(m+T)\times(m+T)}$ and $\hat{C} \in \mathbb{R}^{(m+T)}$ be defined by (8). In the following, for simplicity, we use the shorthand notation $h(t) = h_t$. The following theorem establishes that a certain finite-memory causal system $\hat{\mathbf{S}}$ is an ϵ-approximation to the optimal

$$\hat{A} = \begin{bmatrix}
0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
h_T & h_{T-1} & h_{T-2} & \dots & h_1 & h_0 & \dots & h_T & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & h_T & h_{T-1} & \dots & h_2 & h_1 & \dots & h_{T-1} & h_T & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & h_T & h_{T-1} & h_{T-2} & \dots & h_0 & h_1 & h_2 & \dots & h_T & 0 \\
0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & h_T & h_{T-1} & \dots & h_1 & h_0 & h_1 & \dots & h_{T-1} & h_T \\
\vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & h_T & h_{T-1} & \dots & h_1 & h_0 \\
0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & h_T & \dots & h_2 & h_1
\end{bmatrix}, \hat{C} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{8}$$

system $\mathbf{S}^*$, where $\epsilon = \rho^m$ for some $\rho \in (0,1)$ and positive integer $m > T$.

Theorem 3. Let the system $\hat{\mathbf{S}}$ mapping w to $\tilde{y}$ be defined by the following state space model

$$x(t+1) = \text{sat}_r(\hat{A}x(t) + \hat{w}(t)), \quad x(t_0) = 0,$$
$$\hat{y}(t) = \hat{C}x(t),$$

where matrices $\hat{A}$ and $\hat{C}$ are as defined in (8), $m \in \mathbb{Z}$, $m > T$, $x(t), \hat{w}(t) \in \mathbb{C}^{T+m}$, and

$$\hat{w}(t) = [0\ \dots\ 0\ w(t+1)\ w(t+2)\ \dots\ w(t+m+1)]^T$$

for all $t \geq t_0$. Let $P : \mathbb{R} \to \mathbb{R}$ be defined as

$$P(x) = \left(1 - \sum_{\tau=1}^{T} |h_\tau|\right) x^{T+1} - \sum_{\tau=0}^{T} |h_\tau| x^{T-\tau}.$$

Then

$$|y(t) - \hat{y}(t)| \leq ||w||_\infty \rho^m, \quad \forall t \geq t_0 + m,$$

where $\rho \in (0,1)$ such that $P(\rho) = 0$ and $P(x) > 0$ for all $x \in (\rho, 1)$.

Proof. Omitted due to space constraints.■

5. AN APPLICATION OF THEOREM 3.

5.1 Optimal peak-to-average-power ratio reduction

An important application which involves an optimization setup as defined in (2) is peak-to-average-power ratio (PAPR) reduction in digital communication systems. Power efficiency (PE) is one of the most important characteristics of a transmitter circuit, and in order to achieve high PE, the power amplifier (PA), that is used to amplify communication signals, must be operated close to its saturation level. Modern communication signals (e.g., LTE) have high PAPR and suffer from significant distortion when passed through a nonlinear PA operating close to saturation. This causes in-band and out of band spectral content which degrades spectral efficiency and error vector magnitude (EVM) specifications of communication system . The conventional solution to this problem is to back-off the operating point of the PA, which drastically decreases power efficiency of a transmitter. For example, high value of PAPR represents the main drawback of currently used signal standards (OFDM/LTE) in forthcoming wide-band communication systems.

A very simple and popular method for PAPR reduction is to intentionally clip and filter baseband signal before amplification (Ochiai and Imai (2002), Wang and Luo (2011), Zhu et al. (2013), see Fig. 1. Since large peaks occur with very low probability, clipping seems to be an effective method for PAPR reduction. The out-of-band spectral regrowth caused by clipping is mitigated by post filtering, which in turn might generate some peak regrowth. These clipping and filtering operations are denoted as the operation of system $\mathbf{S}$ in Fig. 1. The

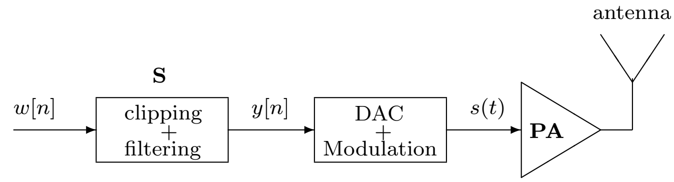


Fig. 1. Simplified block diagram of a typical transmitter circuit. A digital pulse shaping filter $\mathbf{S}$ is series interconnection of an upsampler and a shaping filter $\mathbf{F}$.

problem of designing optimal $\mathbf{S}$ can be posed as an infinite dimensional convex quadratic optimization problem with box constraints as defined in (2). Here a tradeoff has to be made between minimizing the output peak amplitude, satisfying the spectral mask condition, and satisfying the EVM specification. Namely, the spectral mask condition pertains to keeping the spectral content of $y = \mathbf{S}w$ mostly in a desired baseband frequency range. The EVM condition pertains to keeping $||y - w||$ small, where $||\cdot||$ is an appropriate metric/norm.

Now, a design of an optimal PAPR reduction system can be posed as

$$\min_y J(w,y), \quad \text{subject to } ||y||_\infty \leq r. \tag{9}$$

where $0 < r < ||w||_\infty$, and

$$J(w,y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} L(\omega)|Y(\omega)|^2 d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \gamma |Y(\omega) - W(\omega)|^2 d\omega \tag{10}$$

with $L(\omega) > 0$, for all $\omega \in \mathbb{R}$ and $\gamma > 0$. Clearly, the frequency weighting function $L(\omega)$ penalizes high frequency content of y and therefore the first term in (10) enforces spectral mask condition. The second summand in (10) serves to enforce closeness of optimal y to the true baseband signal w, and therefore parameter γ controls the tradeoff between spectral mask and EVM conditions. It can be seen that (9) is equivalent to (2), with $\alpha(\omega) = L(\omega) + \gamma$ and $\beta(\omega) = \gamma$, for all $\omega \in \mathbb{R}$. Therefore, the optimal system $\mathbf{S}$ defined by the solution of (9) has exponentially fading memory and can be approximated by a finite-memory causal system using the proposed algorithm from Theorem 3. provided that the "L1 dominance" condition is satisfied.

5.2 Numerical Experiment

In this subsection, we illustrate the use of the algorithm proposed in Theorem 3 in the design of an optimal PAPR reduction system, as described above. The frequency weighting function $L = L(\omega)$ is defined as a frequency response of a LTI system with unit sample response $l = l(t)$, such that $l(t) = l(-t)$ for all $t \in \mathbb{Z}$ and $l(t) = 0$ for $|t| > 27$. The frequency response of this filter, for $\omega \in (0,\pi)$, is depicted in Fig. 2. The trade-off parameter γ is set to $\gamma = 1.5$. Input (baseband) signal w is obtained by upsampling and spectrally shaping a uniformly distributed 64QAM sequence (Proakis and Salehi (2007)). Signal w is spectrally shaped so that most of its spectral content is distributed in the frequency range $(-\pi/2, \pi, 2)$ (on interval $(-\pi,\pi)$, and 2π-periodically extended otherwise).

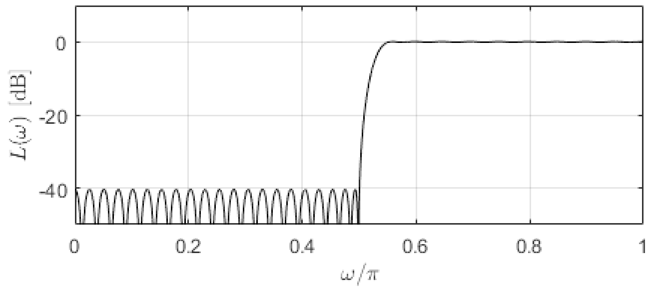


Fig. 2. Frequency weighting function $L(\omega)$ (i.e., frequency response of $l = l(t)$.

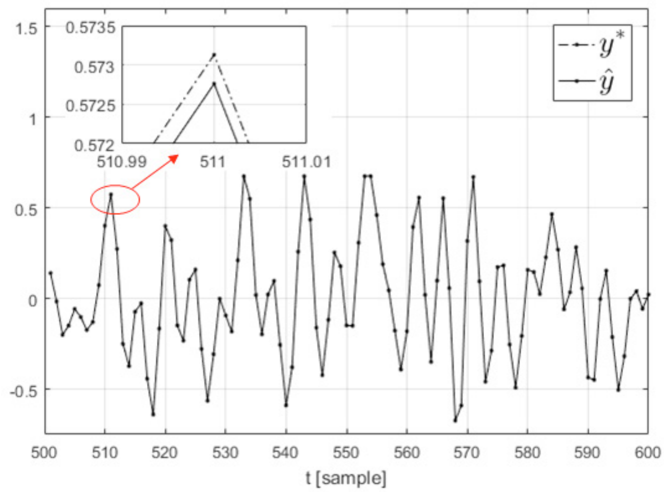


Fig. 3. Comparison of output signals of $\mathbf{S}$ obtained by the algorithm from Theorem 3 (solid) and a fixed point iteration in (4) (dash-dot).

Supremum bound on y as chosen as $r = 0.75||w||_\infty$. Unit sample response $l(t)$ and parameter γ were chosen such that the unit sample response $h = h(t)$ of $\mathbf{H} = \mathbf{I} - \mathbf{T}_\alpha$ satisfies $\sum_{t=-T}^{T} |h(t)| = 0.9 < 1$, where $T = 27$. It follows from Theorem 3. that $\rho = 0.96$. In order for system $\hat{\mathbf{S}}$ to be ϵ-approximation with $\epsilon = 10^{-3}$, the latency parameter m is chosen as $m = \lceil \log_\rho(\epsilon) \rceil = 170$. The output signal $\hat{y}$, which is calculated applying the proposed algorithm from Theorem 3, is compared to a suboptimal solution y^* to problem (2) obtained by running the fixed point iteration (4) on the whole length of input signal w. Signals $\hat{y}$ and y^* are compared in Fig. 3. As can be seen, the upper bound on error $\epsilon = \rho^m$, as stated by Theorem 3, is very conservative and much better error can be achieved (infact, $|y^*(t) - \hat{y}(t)| < 10^{-6}$ for almost all t used in simulation) .

6. CONCLUSION

In this paper, we consider a problem of designing discrete-time systems which are optimal in frequency-weighted least squares sense subject to a maximal output amplitude constraint. We have shown that the optimal system, correspoding to the optimal solution of such problem, has exponentially fading memory. An algorithm is proposed, based on time and value iterations of a carefully chosen causal and stable finite memory system, which, under some L1 dominance assumption, returns high quality approximations to the optimal solution. The result is illustrated on a problem of minimizing peak-to-average-power ratio of a communication signal, stemming from power-efficient transceiver design in modern digital communication systems. A problem of subsequent interest is the extension of the proposed algorithm to more general settings were contractivity of $\mathbf{H}$ in the sense of L1-norm does not hold.

REFERENCES

Bertsekas, D. (1976). On the Goldstein-Levitin-Polyak gradient projection method. *IEEE Transactions on Automatic Control*, 21(2), 174–184.

Boyd, S. and Chua, L. (1985). Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, 32(11), 1150–1161.

Chmielewski, D. and Manousiouthakis, V. (1996). On constrained infinite-time linear quadratic optimal control. *Systems & Control Letters*, 29(3), 121 – 129.

Dembo, R. and Tulowitzki, U. (1983). On the minimization of quadratic functions subject to box constraints. Technical report, Working Paper Series B #71, School of Organization and Management, Yale University, New Haven, Connecticut.

Goodwin, G., Seron, M., and de Doná, J. (2005). *Constrained Control and Estimation: An Optimisation Approach*. Communications and Control Engineering. Springer.

Grieder, P., Borrelli, F., Torrisi, F., and Morari, M. (2004). Computation of the constrained infinite time linear quadratic regulator. *Automatica*, 40(4), 701 – 708.

Hungerländer, P. and Rendl, F. (2015). A feasible active set method for strictly convex quadratic problems with simple bounds. *SIAM Journal on Optimization*, 25(3), 1633–1659.

Kunisch, K. and Rendl, F. (2003). An infeasible active set method for quadratic problems with simple bounds. *SIAM Journal on Optimization*, 14(1), 35–52.

Li, W. and Swetits, J. (1993). A Newton method for convex regression, data smoothing, and quadratic programming with bounded constraints. *SIAM Journal on Optimization*, 3(3), 466–488.

Lin, Y. and Cryer, C.W. (1985). An alternating direction implicit algorithm for the solution of linear complementarity problems arising from free boundary problems. *Applied Mathematics and Optimization*, 13(1), 1–17.

Mead, J.L. and Renaut, R.A. (2010). Least squares problems with inequality constraints as quadratic constraints. *Linear Algebra and its Applications*, 432(8), 1936 – 1949. Special issue devoted to the 15th ILAS Conference at Cancun, Mexico, June 16-20, 2008.

Moré, J.J. and Toraldo, G. (1989). Algorithms for bound constrained quadratic programming problems. *Numerische Mathematik*, 55(4), 377–400.

Ochiai, H. and Imai, H. (2002). Performance analysis of deliberately clipped OFDM signals. *IEEE Transactions on Communications*, 50(1), 89–101.

Oreborn, U. (1986). A direct method for sparse nonnegative least squares problems. Technical report, Department of Mathematics, Thesis, Linköping University, Linköping, Sweden.

Osuna, E., Freund, R., and Girosi, F. (1997). Support vector machines: Training and applications. Technical report, Cambridge, MA, USA.

Pardalos, P.M., Ye, Y., and Han, C.G. (1990). An interior-point algorithm for large-scale quadratic problems with box constraints. In A. Bensoussan and J.L. Lions (eds.), *Analysis and Optimization of Systes*, 413–422. Springer Berlin Heidelberg, Berlin, Heidelberg.
Proakis, J. and Salehi, M. (2007). *Digital Communications*. McGraw-Hill Education.
Reine, R. and Zang, Z. (2013). A quadratic programming approach in pulse shaping filter design to reducing PAPR in OFDM systems. In *2013 19th Asia-Pacific Conference on Communications (APCC)*, 572–576.
Sochacki, J.J. (2016). A preoptimized peak to average power ratio pulse shaping filter and its effect on system specifications. *IEEE Transactions on Microwave Theory and Techniques*, 64(7), 2137–2145.
Stathopoulos, G., Korda, M., and Jones, C.N. (2017). Solving the infinite-horizon constrained LQR problem using accelerated dual proximal methods. *IEEE Transactions on Automatic Control*, 62(4), 1752–1767.
Sznaier, M. and Damborg, M.J. (1987). Suboptimal control of linear systems with state and control inequality constraints. In *26th IEEE Conference on Decision and Control*, volume 26, 761–762.
Wang, Y.C. and Luo, Z.Q. (2011). Optimized iterative clipping and filtering for papr reduction of ofdm signals. *IEEE Transactions on Communications*, 59(1), 33–37.
Zhu, X., Pan, W., Li, H., and Tang, Y. (2013). Simplified approach to optimized iterative clipping and filtering for PAPR reduction of OFDM signals. *IEEE Transactions on Communications*, 61(5), 1891–1901.