Maximizing the Output Energy of a Linear Channel with a Time- and Amplitude-Limited Input

Michael L. Honig and Kenneth Steiglitz, *Fellow, IEEE*

Abstract—**The following problem is considered: Maximize the output energy of a linear time-invariant channel, given that the input signal is time and amplitude limited. It is shown that a necessary condition for an input u to be optimal, assuming a unity amplitude constraint, is that it satisfy the fixed-point equation $u = \text{sgn}\,[F(u)]$, where the functional F is the convolution of u with the autocorrelation function of the channel impulse response. It is also shown that all solutions to this equation for which $|u| = 1$ almost everywhere correspond to local maxima of the output energy. Iteratively recomputing u from the fixed-point equation leads to an algorithm for finding local optima. Numerical results are given for the cases where the transfer function is ideal lowpass, and has two poles. These results support the conjecture that in the ideal low-pass case the optimal input signal is a single square pulse. Often, several local optima are found by the iterative algorithm, and the global optimization problem appears to be computationally intractable. A generalization of the preceding fixed-point condition is also derived for the problem of maximally separating N outputs of a discrete-time, linear, time-invariant channel, assuming the inputs are constrained in time and amplitude.**

Index Terms—**Amplitude-constraint, maximum energy, signal design, fixed-point condition.**

I. Introduction

THE FOLLOWING signal design problem was posed by Wyner [1]: Maximize the output energy of a linear time-invariant channel, given that the input signal is time and amplitude limited. That is,

$$\max_{u(t)} E = \int_{-\infty}^{\infty} y^2(t)\, dt \qquad \text{(P1)}$$

subject to

$$|u(t)| \begin{cases} \leq 1, & \text{if } 0 \leq t \leq T, \\ = 0, & \text{otherwise}, \end{cases} \qquad \text{(C1)}$$

where

$$y(t) = \int_0^t h(t-s)u(s)\, ds,$$

and $h(t)$ is the channel impulse response, which is assumed to be in L_2. One motivation for studying this problem is the design of signals for channels whose inputs must be limited in amplitude, such as the magnetic recording channel [2]. Also, in some situations the input to the channel may be constrained by the dynamic range of the transmitter electronics, which also leads to a peak power, or maximum amplitude constraint. If $u(t)$ is a solution to (P1)–(C1), then the channel outputs corresponding to $\pm u(t)$ are maximally separated in L_2, making them relatively easy to distinguish at the receiver.

From Parseval's relation, the energy in (P1) can be rewritten as

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 |U(j\omega)|^2\, d\omega,$$

where $H(j\omega)$ and $U(j\omega)$ are the Fourier transforms of $h(t)$ and $u(t)$, respectively. Thus, the maximum energy criterion can be viewed as a measure of how well the input spectrum is matched to the channel spectrum. Problem (P1)–(C1) is therefore closely related to the study of the spectra of signals limited in both duration and amplitude [1].

Our main result is a necessary condition for optimality, which will be derived in two ways: first using the Pontryagin maximum principle, and second, using a Lagrange multiplier to solve (P1)–(C1) where the amplitude constraint (C1) is replaced by an L_p-norm constraint, and then letting p go to infinity. The condition is that the optimal solution must satisfy the fixed-point equation $u = \text{sgn}\,[F(u)]$, where F is the convolution of u with the autocorrelation function of the channel impulse response. We then show that fixed points represent local maxima of the channel output energy.

The fixed-point condition naturally suggests an iterative substitution algorithm, which is then applied to two interesting special cases: channels with two-pole transfer functions, and the ideal low-pass channel. In the two-pole case, the numerical results illustrate the spectrum-matching property of locally optimal solutions. In the ideal low-pass case, the numerical results support the conjecture that the optimal input is always a single pulse. This problem is one in a class of signal design problems whose global solution appears to be computationally intractable.

Finally, we consider a generalization of the problem (P1)–(C1) in discrete-time in which N inputs to the channel

Manuscript received November 5, 1990; revised May 22, 1991. This work was supported in part by NSF Grant MIP-8912100 and U.S. Army Research Office—Durham Contract DAAL03-89-K-0074. This work was presented in part at the 1990 International Symposium on Information Theory and Its Applications, Waikiki, HI, November 27–30, 1990.

M. L. Honig is with Bellcore, Room MRE 2L-343, 445 South Street, Morristown, NJ 07960-1910.

K. Steiglitz is with the Department of Computer Science, Princeton University, Princeton, NJ 08544.

IEEE Log Number 9105780.

0018-9448/92$03.00

are to be designed to maximize the minimum distance between channel outputs. We derive an optimality condition, analogous to the fixed-point condition, which must be satisfied by at least one globally optimal solution.

II. The Fixed-Point Condition via the Pontryagin Maximum Principle

In this section, we apply Pontryagin's maximum principle to a state-space formulation of the problem, and thereby derive a necessary condition for optimality. For the theory of the maximum principle, we follow [3], [4]. There is an important observation that makes it possible to carry out this plan: The particular cost criterion of maximizing the total channel output energy can be expressed as one of maximizing a state-variable at a single time instant. This follows from the assumption that the channel is linear. Throughout this section, we will assume that the channel can be represented by a finite-dimensional state-space model, which is equivalent to assuming that the channel transfer function is rational. We also assume that this transfer function is stable and causal.

The state equations that describe the channel dynamics are

$$\dot{x}(t) = Ax(t) + bu(t), \tag{1a}$$

where x is an n-dimensional state vector, u is a scalar input signal, and A and b are constant $n \times n$ and $n \times 1$ vectors, respectively. The scalar channel output y is assumed to be a linear function of the state variables, namely,

$$y(t) = c'x(t), \tag{1b}$$

where c is a fixed n-vector.

According to (P1), we wish to maximize

$$E = \int_0^\infty x'(\tau)Qx(\tau)\,d\tau, \tag{2}$$

where $Q = cc'$. Actually, the following discussion is valid for any symmetric, positive definite matrix Q. For $\tau \geq T$ the system is unforced, so we can write $x(\tau)$ in this range as

$$x(\tau) = e^{A(\tau - T)}x(T), \qquad \tau \geq T. \tag{3}$$

The (generalized) energy can therefore be written as

$$E = \int_0^T x'(\tau)Qx(\tau)\,d\tau + x'(T)Gx(T), \tag{4}$$

where the constant symmetric matrix G is given by

$$G = \int_0^\infty e^{A^T s}Qe^{As}\,ds. \tag{5}$$

A word about notation: We have used the n-dimensional state vector x in the usual way, but to develop the application of the maximum principle we will append an $(n+1)$th component x_{n+1} to represent the cost criterion. Sometimes, as before, we will want to use the vector x, and sometimes the $(n+1)$-dimensional vector with the extra component. We therefore introduce the $(n+1)$-dimensional vector

$$z = \begin{bmatrix} x \\ x_{n+1} \end{bmatrix}$$

to represent the new vector; of course $z_i = x_i$ for all $1 \leq i \leq n$.

The new component to represent the cost criterion is defined as

$$x_{n+1}(t) = \int_0^t x'(\tau)Qx(\tau)\,d\tau + x'(t)Gx(t). \tag{6}$$

This is done so that x_{n+1} at time T is equal to the energy we want to maximize; that is,

$$E = x_{n+1}(T). \tag{7}$$

The derivative is

$$\dot{x}_{n+1} = x'Qx + x'(A^TG + GA)x + b'(G + G^T)xu, \tag{8}$$

where we have used the state equations (1). From now on, we will not explicitly indicate dependence on t when this cannot cause confusion; x and u are functions of time, while Q, A, G, and b are not. A well known theorem [5, Theorem 6, p. 175] tells us that the first two terms in (8) cancel for all x; that is, G in (5) is the unique solution to the equation $Q + A^TG + GA = 0$. Thus, the added state equation is

$$\dot{x}_{n+1} = 2b'Gxu, \tag{9}$$

since G is symmetric. The governing equations for the $(n+1)$-dimensional system are therefore,

$$\dot{z} = f(z, u), \tag{10}$$

where f is the $(n+1)$-dimensional vector function having components

$$f_k(z, u) = [Ax]_k + [b]_k u, \qquad \text{for } k = 1, \cdots, n \tag{11a}$$

and

$$f_{n+1}(z, u) = 2b'Gxu. \tag{11b}$$

The next step in the application of the maximum principle is the generation of the adjoint variables. Again, we use an n-dimensional vector p to represent the variables adjoint to the original n-dimensional x; and introduce the $(n+1)$-dimensional vector

$$q = \begin{bmatrix} p \\ p_{n+1} \end{bmatrix}.$$

The defining equations for the adjoint system are

$$\dot{q}_i = -\sum_{k=1}^{n+1} q_k \frac{\partial f_k}{\partial z_i}, \qquad \text{for } i = 1, \cdots, n+1. \tag{12}$$

The boundary conditions for the adjoint system are determined by the cost function (see [3], [4]). More precisely, if the cost criterion is to maximize $d'z(T)$ where d is an $(n+1)$-vector, then

$$q(T) = -d. \tag{13}$$

The boundary conditions of the adjoint system are enforced at $t = T$—essentially the adjoint system runs backward in time.

Since we want to maximize $x_{n+1}(T)$, the fact that $d = [0\ 0\ 0 \cdots 0\ 1]'$ and (13) imply that $p_k(T) = 0$ when $k \le n$, and $p_{n+1}(T) = -1$. None of the functions $f_k(z, u)$ depends on x_{n+1}, so (12) yields $\dot{p}_{n+1}(t) = 0$, which implies that $p_{n+1}(t) = -1$ for $0 < t < T$.

We can now write the system dynamics in terms of the original n-dimensional state variable and its n-dimensional adjoint:

$$\dot{x} = Ax + bu, \tag{14a}$$

$$\dot{p} = -A^T p + 2Gbu, \tag{14b}$$

where (14b) is obtained by substituting (11) into (12). The component x_{n+1} does not appear in these equations, but simply determines the cost; and it has been shown that $p_{n+1} \equiv -1$. Furthermore, we have the $2n$ boundary conditions $x(0) = p(T) = \mathbf{0}$, which completely determine x and p.

The Hamiltonian of the problem is

$$\begin{aligned} H &= f'p \\ &= p'Ax - [b'(2Gx - p)]u. \end{aligned} \tag{15}$$

The maximum principle then states that a necessary condition for optimality is that u be chosen to minimize this for any values of the variables x, p, and t, which yields the condition

$$u = \operatorname{sgn}[b'(2Gx - p)]. \tag{FP-1}$$

This equation requires that an optimal u have the property that the sgn operation on the right-hand side reproduce the function u; we call this the *fixed-point condition*, and an equivalent form will be derived in the next section. Because both x and p in the bracketed quantity depend on the signal u, it appears quite difficult to obtain an analytical solution for u in general. Note that if $p = 2Gx$, then H, given values for x and p, is independent of u so that u from (FP-1) is ill-defined. The derivation in the next section implies that u must be zero when this occurs. It has been shown in [6], however, that there exists at least one global solution to (P1)–(C1) for which $|u| = 1$ almost everywhere in $[0, T]$, in which case the sgn operation in (FP-1) is well defined almost everywhere.

III. An Equivalent Condition via a Lagrange Multiplier

We now use a different argument, characterizing the channel by its impulse response, to arrive at a condition that will turn out to be equivalent to (FP-1). To enforce the amplitude constraint, we will constrain the L_p-norm of the input signal u and then let p approach infinity. It is convenient to restrict p to even values. Thus, letting $h(t)$ be the channel impulse response, we want to find a u that maximizes

$$L[u(t)] = \int_{-\infty}^{\infty} \left(\int_0^T h(t-\tau)u(\tau)\, d\tau \right)^2 dt - \lambda \int_0^T u^p(\tau)\, d\tau. \tag{16}$$

The Lagrange multiplier λ will be chosen to satisfy the constraint

$$\int_0^T u^p(\tau)\, d\tau = T. \tag{17}$$

This constraint is written with equality because if the L_p norm of u is less than T, it can be scaled up to satisfy (17), and that can only increase the channel output energy. As $p \to \infty$ this becomes equivalent to the condition $|u| = 1$ almost everywhere in $[0, T]$. A standard variational argument leads to the necessary condition

$$2\int_0^T \Phi(t-\tau)u(\tau)\, d\tau - \lambda p u^{p-1}(t) = 0, \tag{18}$$

where

$$\Phi(t-\tau) = \int_{-\infty}^{\infty} h(s-\tau)h(s-t)\, ds, \tag{19}$$

which is the autocorrelation function of the channel impulse response. The convolution of Φ with u is an important quantity, which we denote by $\Psi(t)$:

$$\Psi(t) = \int_0^T \Phi(t-\tau)u(\tau)\, d\tau. \tag{20}$$

Solving (18) for u yields

$$u(t) = \left(\frac{2\Psi(t)}{\lambda p} \right)^{\frac{1}{p-1}}. \tag{21}$$

For each value of p this yields a solution u that depends on λ, and λ is determined by (17). Substituting (21) in (17) gives

$$\frac{2}{\lambda p} \int_0^T \Psi(t)u(t)\, dt = T. \tag{22}$$

It is not hard to show that the integral in (22) is the channel output energy. Assuming u is optimal implies that

$$\lambda p = \frac{2E_p}{T}, \tag{23}$$

where E_p is the maximum channel output energy subject to the L_p constraint (17), for fixed p. It is shown in the appendix that $E_p \to E_\infty$ as $p \to \infty$, provided that

$$\sup_{t \in [0,T]} \int_0^T |\Phi(t-s)|\, ds < \infty,$$

where E_∞ is the maximum energy in (P1) subject to the amplitude constraint (C1). Therefore $\lambda p \to 2E_\infty/T$ as $p \to \infty$. Furthermore, for $h(\cdot) \in L_2$, as $p \to \infty$, any sequence of solutions to (21), $\{u_p\}$, lies in a compact set, and therefore has a convergent subsequence. Also, (21) implies that $u(t) = 0$ for any t such that $\Psi(t) = 0$. Since this holds for all p, we conclude, by letting $p \to \infty$ in (21), that there exists a solution to (P1)–(C1) that satisfies the fixed-point condition

$$u(t) = \operatorname{sgn}[\Psi(t)], \tag{FP-2}$$

where

$$\operatorname{sgn} x = \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ -1, & \text{if } x < 0. \end{cases} \tag{24}$$

Substituting (FP-2) into the expression for channel output energy gives

$$E_\infty = \int_0^T |\Psi(\tau)|\, d\tau. \tag{25}$$

It was pointed out to the authors by A. Friedman that the preceding argument shows only that there exists at least one solution to (P1)–(C1) that satisfies (FP-2). We now show that *any* solution to problem (P1)–(C1) must satisfy (FP-2) almost everywhere. The following argument is similar to that used in Lemma 1.2 of [7]. Namely, suppose that $u^*(t)$ is a solution to (P1)–(C1). Consider the problem

$$\max_{u(t)} \left\{ \tilde{E} = \int_{-\infty}^{\infty} \left(\int_0^T h(t-\tau) u(\tau)\, d\tau \right)^2 dt - \int_0^T [u(t) - u^*(t)]^2\, dt \right\} \tag{P2}$$

subject to the amplitude constraint (C1). Clearly, the maximum value of $\tilde{E}$ in this case must be equal to the maximum output energy E_∞, so that any solution to (P2)–(C1), namely $\tilde{u}(t)$, must satisfy

$$\int_0^T [\tilde{u}(t) - u^*(t)]^2\, dt = 0. \tag{26}$$

We can again write the Lagrangian for (P2), where (C1) is replaced by an L_p constraint, derive an associated optimality condition, and let $p \to \infty$. This leads to the following optimality condition for (P2)–(C1),

$$\tilde{u}(t) = \operatorname{sgn}\left(\frac{T}{E_\infty - \int_0^T \tilde{u}(t)[\tilde{u}(t) - u^*(t)]\, dt} \cdot [\Psi(t) - \tilde{u}(t) + u^*(t)] \right). \tag{27}$$

Now (26) implies that $\tilde{u}(t) = u^*(t)$ almost everywhere, so that this and the optimality condition (27) imply that $u^*(t)$ must satisfy (FP-2) almost everywhere.

It is interesting to observe analogies with problem (P1) where the input is constrained in energy, i.e., L_2-norm, rather than amplitude. In that case, the optimality condition analogous to (FP-2) is the eigenfunction equation

$$u(t) = \mu \Psi(t), \tag{28}$$

and the maximum channel output energy is

$$E_2 = \mu_{\max} \int_0^T \Psi^2(\tau)\, d\tau, \tag{29}$$

where $\mu_{\max}$ is the largest eigenvalue μ in (28). For finite T, the number of solutions to (28) is countably infinite in general. Whether or not this is also true for (FP-2) is unknown.

We now verify that the two conditions (FP-1) and (FP-2) are equivalent. From the forward system equation (14a) and the initial condition $\boldsymbol{x}(0) = 0$ we can write $\boldsymbol{x}$ explicitly as

$$\boldsymbol{x}(t) = \int_0^t e^{A(t-\tau)} \boldsymbol{b} u(\tau)\, d\tau. \tag{30}$$

Similarly, from the adjoint system equation (14b) and the fact that $\boldsymbol{p}(T) = \boldsymbol{0}$,

$$\boldsymbol{p}(t) = -\int_t^T e^{-A^T(\tau - t)} 2 G \boldsymbol{b} u(\tau)\, d\tau. \tag{31}$$

Substituting (30) and (31) in (FP-1) gives

$$u(t) = \operatorname{sgn}\left(\int_0^T \boldsymbol{b}' G e^{A|t-\tau|} \boldsymbol{b} u(\tau)\, d\tau \right). \tag{32}$$

Since the system output is $\boldsymbol{c}'\boldsymbol{x}$, where $\boldsymbol{c}$ is a fixed n-vector, the impulse response is

$$h(t) = \boldsymbol{c}' e^{At} \boldsymbol{b}. \tag{33}$$

The cost matrix $Q = \boldsymbol{c}\boldsymbol{c}'$, and it is then not hard to verify finally that

$$\boldsymbol{b}' G e^{A|t|} \boldsymbol{b} = \Phi(t). \tag{34}$$

IV. Local Behavior at a Fixed-Point

In this section, we study the local behavior of the channel output energy with respect to variations around solutions to (FP-2). The condition (FP-2) implies that $|u(t)|$ is either one or zero for all t. One approach is therefore to study the channel output energy as a function of the times at which $u(t)$ changes value (i.e., the switching times). We first compute the derivative of the channel output energy with respect to a switching time, and show that if the input satisfies (FP-2), then perturbing a single switching time decreases the output energy. We then examine the behavior of the channel output energy with respect to arbitrary variations around any solution to (FP-2), and thereby show that solutions to (FP-2) for which $|u(t)| = 1$ almost everywhere are local maxima.

In order to examine the behavior of the output energy with respect to switching times, we will assume that solutions to (FP-2) switch a finite number of times in $[0, T]$. This is supported by our numerical experience, but conditions on $h(t)$ under which this can be proved are not known. There is also the possibility that $u(t) = 0$ on a set of positive measure. However, if the channel transfer function $H(j\omega)$ is rational, then $\Psi(t)$ can be expressed as the sum of a finite number of complex exponentials, and therefore cannot vanish on any interval of length greater than zero. We therefore assume that $|u(t)| = 1$ almost everywhere, and characterize a particular u by its switching times.

An arbitrary candidate solution u with $(m-1)$ switching points τ_k in the open interval (0, 1) can be written as

$$u(t) = \sum_{k=0}^{m} \gamma_k S(t-\tau_k), \tag{35}$$

where $S(t)$ is the unit step function, $\gamma_k = (-1)^k$ if $k = 0$ or m, and $\gamma_k = 2(-1)^k$ otherwise. The first and last switching points are fixed; $\tau_0 = 0$ and $\tau_m = T$. We refer to such a solution as having m pulses.

We will need the derivative of u with respect to the location of a switching point τ_i, which in terms of a δ-function is

$$\frac{\partial u(t)}{\partial \tau_i} = -\gamma_i \delta(t-\tau_i). \tag{36}$$

The channel output energy is

$$E = \int_{-\infty}^{\infty} \left(\int_0^T h(t-\tau) u(\tau)\, d\tau \right)^2 dt. \tag{37}$$

The derivative of the energy with respect to a particular switching instant τ_i is therefore

$$\begin{aligned}\frac{\partial E}{\partial \tau_i} &= 2\int_{-\infty}^{\infty} \left(\int_0^T h(t-\tau)u(\tau)\, d\tau \right) \cdot \left(\int_0^T h(t-s) \frac{\partial u(s)}{\partial \tau_i} ds \right) dt \\ &= 2\int_{-\infty}^{\infty} \left(\int_0^T h(t-\tau)u(\tau)\, d\tau \right) \left(-\gamma_i h(t-\tau_i) \right) dt \\ &= -2\gamma_i \int_0^T \left(\int_{-\infty}^{\infty} h(t-\tau) h(t-\tau_i)\, dt \right) u(\tau)\, d\tau \\ &= -2\gamma_i \int_0^T \Phi(\tau-\tau_i) u(\tau)\, d\tau \\ &= -2\gamma_i \Psi(\tau_i). \end{aligned} \tag{38}$$

We now show that if $u(t)$ satisfies (FP-2), then E is a local maximum with respect to each switching time. Suppose that, without loss of generality, $\Psi(0) > 0$, so that $u(0) > 0$. Consider the first zero crossing of Ψ, that is, the first switching point τ_1. Since $\gamma_1 = -2$, (38) implies that $\partial E/\partial \tau_1$ changes from positive to negative, which shows that the energy is a local maximum with respect to τ_1. At the next switching point τ_2, Ψ goes from negative to positive, but $\gamma_2 = +2$, so that the energy is a local maximum with respect to τ_2, and so forth for all remaining switching times. In the singular case when Ψ is tangent to the time axis, we ignore the point of tangency as a potential switching point, which concludes the proof.

The preceding argument establishes that E is a local maximum with respect to an *individual* switching time; however, it does not show that E is a local maximum with respect to the set of switching times $\tau_1, \tau_2, \cdots, \tau_m$. A straightforward computation shows that the ijth element of the Hessian matrix in this case is

$$\frac{\partial^2 E}{\partial \tau_i \partial \tau_j} = 2\gamma_i\gamma_j \Phi(\tau_i - \tau_j) - \left(2\gamma_i \sum_{k=1}^{m} \gamma_k \Phi(\tau_i - \tau_k) \right) \delta_{ij}, \tag{39}$$

where $\delta_{ij} = 1$ if $i = j$, and is zero, otherwise. The output energy is a local maximum with respect to the set of switching times if this Hessian matrix is negative definite when $u(t)$ satisfies (FP-2). It is not obvious, however, that this is the case, and we therefore resort to an alternative argument to show that solutions to (FP-2) are local maxima in a different sense. We add that the Hessian matrix given by (39) can be useful if gradient search techniques are used to find solutions to (FP-2) [8, Section 8.5].

To show that solutions to (FP-2) for which $|u(t)| = 1$ almost everywhere for $t \in [0, T]$ correspond to local maxima of the output energy, we consider perturbations to a fixed point of the form

$$\tilde{u}(t) = \frac{u^*(t) + \epsilon\Delta(t)}{\| u^* + \epsilon\Delta \|_\infty}, \tag{40}$$

where $u^*(t)$ is any solution to (FP-2), $\Delta(t)$ is an arbitrary perturbation, ϵ is a small positive constant, and $\|f\|_\infty$ is the L_∞ norm of $f(\cdot)$. For convenience, we will assume that $\Delta(t)$ is continuous. (This is primarily to avoid the use of "essential" suprema and infima in what follows.) We wish to show that if ϵ is small enough, then $E(u^*) - E(\tilde{u}) > 0$, where $E(u)$ is the channel output energy as a functional of the input u. The following discussion applies to solutions to (FP-2) which may be zero on a set of positive measure.

We first write

$$\begin{aligned} E(u^*) - E(\tilde{u}) &= E_\infty - \frac{1}{\|u^* + \epsilon\Delta\|_\infty^2} \int_0^T \int_0^T \left[u^*(t) + \epsilon\Delta(t) \right] \Phi(t-s) \left[u^*(s) + \epsilon\Delta(s) \right] ds\, dt \\ &= \frac{1}{\|u^* + \epsilon\Delta\|_\infty^2} \left[\left[\|u^* + \epsilon\Delta\|_\infty^2 - 1 \right] E_\infty - 2\epsilon \int_0^T \Delta(t) \Psi^*(t)\, dt \right] + O(\epsilon^2), \end{aligned} \tag{41}$$

where $\Psi^*(t)$ denotes $\Psi(t)$ given by (20) in which u is replaced by u^*.

Define the sets

$$I = \{ t : u^*(t) > 0 \}, \qquad \bar{I} = \{ t : u^*(t) < 0 \},$$

and

$$I_0 = \{ t : u^*(t) = 0 \} \tag{42}$$

and the quantities

$$\begin{aligned} M_I &= \sup_{t \in I} \Delta(t), \qquad M_{\bar{I}} = -\inf_{t \in \bar{I}} \Delta(t), \\ M_{I_0} &= \sup_{t \in I_0} |\Delta(t)|. \end{aligned} \tag{43}$$

Then, we have that

$$\|u^* + \epsilon\Delta\|_\infty = \max\left(1 + \epsilon M_I, 1 + \epsilon M_{\bar{I}}, \epsilon M_{I_0}\right) = \max\left(1 + \epsilon M_I, 1 + \epsilon M_{\bar{I}}\right) \quad (44)$$

provided that ϵ is small enough so that $\epsilon\|\Delta\|_\infty < 1$ and $\epsilon M_{I_0} < \max(1 + \epsilon M_I, 1 + \epsilon M_{\bar{I}})$. We can now rewrite (41) as

$$E_\infty - E(\tilde{u}) = 2\epsilon \max(M_I, M_{\bar{I}}) E_\infty - 2\epsilon \int_0^T \Delta(t)\Psi^*(t)\,dt + O(\epsilon^2) = 2\epsilon \int_0^T \left(\max(M_I, M_{\bar{I}}) \operatorname{sgn}[\Psi^*(t)] - \Delta(t)\right)\Psi^*(t)\,dt + O(\epsilon^2), \quad (45)$$

where (25) has been used. From (FP-2) and the definitions of I, $\bar{I}$, and I_0, it follows that (45) can be rewritten as

$$E_\infty - E(\tilde{u}) = 2\epsilon\left(\int_I [\max(M_I, M_{\bar{I}}) - \Delta(t)]\Psi^*(t)\,dt - \int_{\bar{I}} [\max(M_I, M_{\bar{I}}) + \Delta(t)]\Psi^*(t)\,dt\right) + O(\epsilon^2). \quad (46)$$

The corresponding integral over the set I_0 is zero, since $\Psi^*(t) = 0$ for $t \in I_0$. By definition, for $t \in I$, $\max(M_I, M_{\bar{I}}) \geq M_I \geq \Delta(t)$. Also, since $u^*(t) = \operatorname{sgn}[\Psi^*(t)]$, $\Psi^*(t) > 0$ for $t \in I$, so that the first integral on the right must be nonnegative. Similarly, for $t \in \bar{I}$, $\max(M_I, M_{\bar{I}}) \geq M_{\bar{I}} \geq -\Delta(t)$, and $\Psi^*(t) < 0$, so that the second integral is less than or equal to zero, which implies that the term multiplying ϵ in (46) is nonnegative. Now the two integrals in (46) can both be zero, if and only if $\Delta(t) = M_I = M_{\bar{I}}$ for all $t \in I \cup \bar{I}$. If this is the case, then (40), (42), and (43) imply that $\tilde{u}(t) = u^*(t)$ for $t \in I \cup \bar{I}$. Consequently, we conclude that if there exists a solution to (FP-2) in which meas $I_0 > 0$, then there exist perturbations $\Delta(t)$ for which the local behavior of the output energy E is determined by higher-order terms in ϵ. However, if meas $I_0 = 0$, then the preceding argument implies that $E_\infty - E(\tilde{u}) > 0$ for sufficiently small ϵ.

V. A Fixed-Point Algorithm

The fixed-point condition (FP-2) suggests the following numerical algorithm for computing solutions to (P1)–(C1):

1) Choose an initial input u^0 with $|u^0| = 1$ in the interval $[0, T]$ and zero elsewhere, but otherwise arbitrary.
2) Compute a new input u^{i+1} from u^i by

$$u^{i+1}(t) = \operatorname{sgn}\left(\int_0^T \Phi(t-\tau)u^i(\tau)\,d\tau\right) \quad \text{(A)}$$

until the switching points of u^{i+1} and u^i are sufficiently close.

Of course, (FP-1) can also be used as the basis for this algorithm, in which case we would, given u, integrate the system equations (14) to get the required x and p. In practice, the algorithm is run several times from initial solutions u^0 that are chosen pseudorandomly. The method of choosing these is not critical, so long as a wide variety of different initial solutions is sampled. The particular algorithm used was the following: Choose a pseudorandom integer m uniformly between 1 and $m_{\max} = 10$. Then choose the switching points $\tau_0 = 0$, τ_i pseudorandomly and uniformly in the interval $[(i-1)/m, i/m)$, $i = 1, 2, \cdots, m-1$, and $\tau_m = 1$. The computationally intensive step is the determination of the zero crossings of the argument of the sgn function in (A), which requires repeated evaluations of the convolution integral.

One interpretation of (A) is that it is a gradient algorithm with a very large step-size, followed by a projection onto the feasibility region. Specifically, consider perturbing the input $u(t)$ by some function $\epsilon\tilde{u}(t)$, where ϵ is small. Letting $E[u(t)]$ denote the channel output energy as a functional of the input, it is easily shown that

$$\left.\frac{\partial E[u(t) + \epsilon\tilde{u}(t)]}{\partial\epsilon}\right|_{\epsilon=0} = \Psi(t).$$

Consequently, one version of a fixed step-size gradient algorithm is to let $u^{i+1}(t) = u^i(t) - \beta\Psi(t)$ for all t such that $|u^i(t) - \beta\Psi(t)| \leq 1$. If $|u^i(t) - \beta\Psi(t)| > 1$, then projecting onto the constraint region gives $u^{i+1}(t) = \operatorname{sgn}[u^i(t) - \beta\Psi(t)]$. For very large β, this algorithm becomes (A). Observe that if the gradient algorithm just described converges, then it also must converge to a solution of (FP-2).

VI. Numerical Results

A. Upper Bound and Normalization

It is easily shown using the Cauchy–Schwarz inequality that the channel output energy E is upper bounded by

$$E \leq T^2 \int_{-\infty}^{\infty} h^2(t)\,dt. \quad (47)$$

All energies reported in the numerical results are normalized by dividing by the upper bound in (47). It is assumed throughout that $T = 1$.

B. The Two-Pole Case

In this section, we consider the two-pole channel impulse response

$$h(t) = e^{-\sigma t}\cos(2\pi f t + \phi). \quad (48)$$

In this case the iterative algorithm described in Section IV can be made quite fast because the integrals required to compute Ψ and the energy can be evaluated in closed form in terms of elementary functions.

The first example is typical of the results obtained when the channel is bandpass. It corresponds to the parameters $\sigma = 1$, $f = 5$, and $\phi = 0$. The fixed-point algorithm was run 100 times, each with a random starting input u^0. A limit of

1000 iterations was set for each trial; if convergence was not obtained by then, the trial was abandoned. Out of 100 random starts, the program converged 57 times to a particular solution with 11 pulses, once to a solution with 10 pulses, and failed to converge within 1000 iterations for the remaining 42 trials. The two distinct solutions found are very similar, having normalized energy $E_0 = 0.29887$ and 0.29771, respectively. Fig. 1 shows the time waveform and spectrum of the best solution found.

Fig. 2 shows the best locally optimal solution obtained in a lowpass case, with the same parameters as before, except $f = 0.75$. This case turns out to be much less demanding computationally. A thousand random initial solutions were tried and the program converged within 2000 iterations every time. Again, two distinct local optima were found; one with 2 distinct pulses (900 times) and the other with 3 (100 times). In this case the two solutions are appreciably different, with normalized costs of 0.36964 and 0.24566, respectively. Fig. 3 shows the second-best solution.

We mention one property shared by all the local optima found numerically in the two-pole case: the switching points are symmetrically located in the interval $[0, T]$. That is, for every switching point at τ there is another at $T - \tau$. For example, the solution shown in Fig. 1 has the switching points

0	0.00000
1	0.05165
2	0.15123
3	0.25086
4	0.35052
5	0.45019
6	0.54987
7	0.64954
8	0.74919
9	0.84882
10	0.94840
11	1.00000.

We conjecture that all local optima have this property in the two-pole case. Counterexamples to this conjecture have been found in the ideal low-pass case, considered next.

C. The Ideal Low-pass Case

The result of extensive computation in the ideal lowpass case is that the single pulse is always optimal, which we leave as a second conjecture. It is easily verified that in this case $u(t) = 1$, $0 < t < T$, is always a solution to the fixed-point condition (FP-2). Table I shows a summary of the best and second-best locally optimal solutions obtained for bandwidths in the range $B = 0.1$ to 11, for the impulse response

$$h(t) = \sin(2\pi Bt)/(2\pi Bt). \tag{49}$$

As $B \to \infty$, this channel approaches the identity channel, so that the output energies corresponding to all unit-modulus input functions become equal. Consequently, the energies of the best and second-best local optima shown in Table I become very close for large B. The empirical probability of obtaining the best local optimum declines from 100% in the narrow-band cases ($B = 0.1$ and 0.5) to about 13% for the wide-band cases ($B = 10$ and 11). The second-best solutions have 2 pulses instead of one (except for the case $B = 2$, which has 3 pulses).

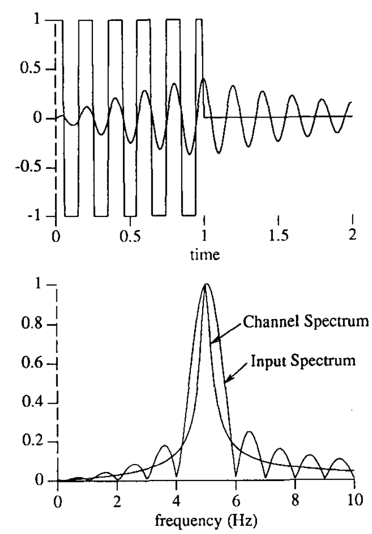

Fig. 1. Best locally optimal solution obtained for the two-pole, bandpass case with $\sigma = 1$, $f = 5$, and $\phi = 0$. Channel input and output time waveforms are shown at the top, and the channel and input spectra at the bottom. Spectra are normalized so that their peak values are unity.

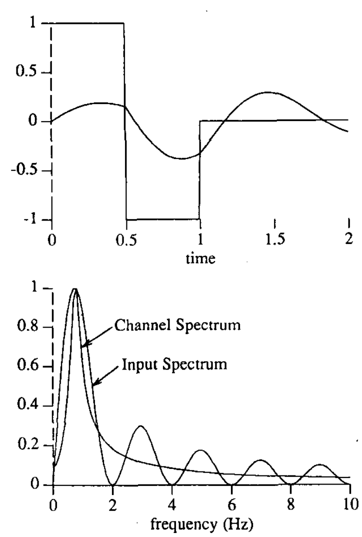

Fig. 2. Same as Fig. 1 except for the low-pass case with $f = 0.75$.

D. Comment on Energy Calculation

Some care was taken in the numerical implementation to check the energy calculation. For maximal computational efficiency, the formula

$$E = \int_0^T \Psi(\tau)u(\tau)\, d\tau$$

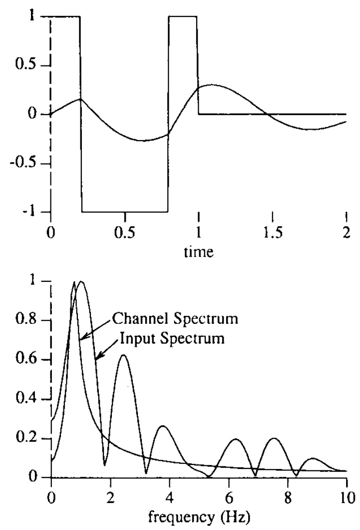

Fig. 3. Same as Fig. 2, but for the second-best locally optimal solution.

TABLE I
NORMALIZED MAXIMUM ENERGY OBTAINED FROM THE FIXED-POINT ALGORITHM, VS. BANDWIDTH B IN HZ., IN THE IDEAL LOW-PASS CASE*

B	max E	m	%	Second-Best E	m	%
0.1	0.9860	1	100	(none)		
0.5	0.7724	1	100	(none)		
1.0	0.4514	1	98†	(none)		
2.0	0.2375	1	32	0.2053	3	37
3.0	0.1611	1	20	0.1510	2	21
4.0	0.1218	1	16	0.1163	2	10
5.0	0.0980	1	16	0.0945	2	8
6.0	0.0819	1	16	0.0795	2	7
7.0	0.0704	1	15	0.0686	2	7
8.0	0.0617	1	14	0.0604	2	12
9.0	0.0549	1	14	0.0539	2	12
10.0	0.0495	1	13	0.0486	2	13
11.0	0.0450	1	13	0.0443	2	4

*For each bandwidth B, the table shows the best locally optimal energy obtained, the corresponding number of pulses m, and the number of times obtained out of 100. This is followed by the same information for the second-best locally optimal solution.

†Two cases failed to converge within the allowed number of iterations, which was 1000.

is best, because when combined with (35) it entails only $O(m)$ evaluations of the indefinite integral of Ψ, where m is the number of pulses in the solution. The energy can also be computed by numerical integration of the squared channel output in the time domain, or in the frequency domain using Parseval's relation. The time-domain numerical integration was used to check the energy calculation for the two-pole case, and all three methods were compared in the ideal low-pass case.

VII. DISCRETE-TIME PROBLEM

Given a discrete-time, linear, shift-invariant channel with impulse response $h[k]$, consider the problem

$$\max_{\{u[k]\}} E = \sum_{i=1}^{\infty} \left(\sum_{k=1}^{K} h[i-k]u[k] \right)^2 \qquad \text{(P3)}$$

subject to the constraint

$$|u[k]| \le 1, \qquad k = 1,\cdots,K, \quad \text{and} \quad u[k] = 0, \qquad \text{for } k < 1 \text{ and } k > K. \qquad \text{(C3)}$$

Here we assume that $h[k] = 0$ for $k < 0$ and $k > M-1$ (that is, the channel is FIR). It will be convenient to define the following vectors of channel inputs and outputs:

$$\boldsymbol{u} = [u[1] \quad u[2] \cdots u[K]], \qquad \text{(50a)}$$

$$\boldsymbol{y} = [y[1] \quad y[2] \cdots y[K+M-1]], \qquad \text{(50b)}$$

where M is the length of the impulse response $h[k]$. We can therefore write

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{u}, \qquad (51)$$

where $\boldsymbol{H}$ is the $(K+M-1) \times K$ matrix of impulse response coefficients that maps inputs to outputs. Note that $E = \boldsymbol{y}'\boldsymbol{y} = \boldsymbol{u}'\boldsymbol{H}'\boldsymbol{H}\boldsymbol{u}$.

Pontryagin's maximum principle, outlined in Section II, does not readily apply to discrete-time systems where the complete set of system equations are nonlinear [4]. This is true for the case considered when the channel state equations are augmented by a variable representing the cost function in (P3), as was done in Section II. However, we can again replace the maximum amplitude constraint by the l_p-constraint

$$\frac{1}{K}\sum_{k=1}^{K} |u[k]|^p \le 1, \qquad (52)$$

derive a necessary condition for an optimal input via the Lagrange multiplier technique, and let $p \to \infty$. This leads to the discrete-time fixed-point condition

$$\boldsymbol{u} = \operatorname{sgn}(\Phi \boldsymbol{u}), \qquad \text{(FP-3)}$$

where $\Phi = \boldsymbol{H}'\boldsymbol{H}$ and sgn x is again defined by (24). This condition suggests a discrete-time version of the continuous-time fixed-point algorithm (A), which can be used to search for local optima. Also in analogy with the continuous-time case, it can be shown that all solutions to (FP-3) for which each component of $\boldsymbol{u}$ is ± 1 correspond to local maxima of the output energy.

The discrete-time problem (P3)–(C3) has the following geometric interpretation. Any feasible input to the channel is a point contained in a cube in $\mathbb{R}^K$, with sides of length two, centered at the origin. Any channel output corresponding to a feasible input therefore lies within a parallelepiped $\boldsymbol{P}$ in $\mathbb{R}^{K+M-1}$, which is spanned by the column vectors of $\boldsymbol{H}$, and translated so that its center is at the origin. Problem (P3)–(C3) is therefore equivalent to finding a point in $\boldsymbol{P}$ farthest from the origin.

Note that $\nabla_{\boldsymbol{u}} E = \Phi \boldsymbol{u} = \boldsymbol{H}'\boldsymbol{y}$. That is, starting at any point in $\boldsymbol{P}$, moving incrementally in the direction of $\boldsymbol{H}'\boldsymbol{y}$ gives the largest increase in distance from the origin. If a point on the boundary of $\boldsymbol{P}$ is moved in the direction of $\boldsymbol{H}'\boldsymbol{y}$, then it moves outside of $\boldsymbol{P}$. Although written in terms of the input $\boldsymbol{u}$, (FP-3) simply states that an output point $\boldsymbol{y} \in \boldsymbol{P}$ is a local optimum if, after moving incrementally in the direction of the gradient and projecting onto $\boldsymbol{P}$, the result is again y.

Any solution y in the output space must be a vertex of $\boldsymbol{P}$. Otherwise, starting at a solution point, one could move along the boundary of $\boldsymbol{P}$ in a direction which increases the distance from the origin. This is because $\boldsymbol{P}$ is constructed from the intersection of hyperplanes, and it is always possible to move along a hyperplane away from the origin. Furthermore, the sphere of radius $\sqrt{E_{max}}$ centered at the origin, where $E_{\max} = \max E$ in (P3), can intersect $\boldsymbol{P}$ only at the vertices of $\boldsymbol{P}$ farthest from the origin. Since every vertex of $\boldsymbol{P}$ corresponds to a corner of the cube in which the input points must lie, any solution $\boldsymbol{u}$ to the discrete-time maximum energy problem satisfies $|u[k]| = 1$ for $k = 1,\cdots, K$. This implies that the global solution to (P3)–(C3) can be obtained by an exhaustive search through all 2^K vectors $\boldsymbol{u}$ for which $|u[k]| = 1$.

In general, given an arbitrary convex polytope, the problem of finding a vertex farthest from the origin is an NP-hard quadratic programming problem [9]. A related problem, which arises in the context of maximum-likelihood detection of data at the output of a multiuser channel, and shown in [10] to be NP-hard, is

$$\max_{\boldsymbol{u}} \boldsymbol{u}' A \boldsymbol{u} \tag{QP}$$

subject to

$$|u[k]| = 1, \qquad k = 1,\cdots, K, \tag{QPC}$$

where A can be any nonnegative definite $K \times K$ matrix with rational elements. This problem becomes (P3)–(C3) when $A = \Phi$, and (QPC) is replaced by (C3). The previous discussion shows, however, that any solution to (QP) subject to (C3) must satisfy the constraint (QPC). Consequently, (QP) subject to the constraint (C3) is also NP-hard. For the maximum energy problem considered here, $A = H'H$ is confined to the class of real-symmetric, Toeplitz matrices. Whether or not (QP)–(C3) remains NP-hard when A is constrained to be in this class is an open question [10].

A. Asymptotic Fixed-Point Condition

Consider, for the moment, (P3) where the input amplitude constraint is replaced by an input energy (l_2) constraint. A necessary condition for the optimal input is then the eigenvector equation

$$\boldsymbol{u} = \lambda \Phi \boldsymbol{u}. \tag{53}$$

Rather than restrict $u[k]$ to be zero outside the interval [1, K], we now assume that $u[k]$ is zero outside the interval $[-K, K]$. Note that (41) still holds, although H, and therefore Φ, must be modified accordingly.

We now consider the problem of maximizing the output energy per unit time (power) as $K \to \infty$. Since $h[k]$ has finite length M, the solution $u[k]$ for $|k| \le K - M$ is independent of the boundary condition $u[k] = 0$ for $|k| > K$. The "steady-state" solution that maximizes output power as $K \to \infty$ therefore satisfies

$$\sum_{k=-\infty}^{\infty} \phi_{i-k} u[k] = \lambda u[i], \tag{54}$$

for all i, where $\phi_i = \sum_{k=0}^{M-1} h[k]h[k+i]$. Solutions to (54) are sinusoidal sequences

$$u[k] = A \sin(k\omega + \theta), \tag{55}$$

where A and θ are arbitrary constants, and $\lambda = |H(e^{-j\omega})|^2$, the squared magnitude of the frequency response of the channel evaluated at ω. The solution to the maximum energy problem as $K \to \infty$ is simply the sinusoidal sequence $u[k]$ at the frequency ω for which $|H(e^{-j\omega})|$ assumes its maximum value.

Returning to (P3)–(C3), as $K \to \infty$, the condition (FP-3) becomes, in analogy with (54),

$$\operatorname{sgn}\left(\sum_{k=-\infty}^{\infty} \phi_{i-k} u[k]\right) = u[i]. \tag{FP-4}$$

As an example, consider the $1 - D$ channel for which $h[0] = 1$, $h[1] = -1$, and $h[k] = 0$, $k \neq 0, 1$. Then (FP-4) becomes

$$\operatorname{sgn}(2u[k] - u[k-1] - u[k+1]) = u[k]. \tag{56}$$

Assuming $|u[k]| = 1$ for all k, then solutions to (56) are all doubly infinite sequences containing elements 1 and -1 for which any three successive elements $u[k-1]$, $u[k]$, $u[k+1]$ cannot all be the same. It can be easily verified by inspection that the sequence $\cdots, 1, -1, 1, -1, \cdots$ maximizes the output power, and that other solutions to (56) may result in strictly less output power.

From this example, we see that there can be an uncountably infinite number of solutions to (FP-4) (the same applies to (54)). In contrast to solutions to (54), solutions to (FP-4) need not be periodic. The problem of further characterizing solutions to (FP-3) and (FP-4) remains open.

VIII. Optimality Condition for l_∞ / l_2 Signal Design With Many Inputs

Given a linear, shift-invariant channel with impulse response $h[k]$, suppose that we wish to find $N \ge 2$ time-limited inputs that maximize the minimum Euclidean distance between channel outputs. That is, we wish to find $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N$ to

$$\max \left\{ d_{\min} = \min_{i \neq j} \| H(\boldsymbol{u}_i - \boldsymbol{u}_j) \| \right\} \tag{P4}$$

subject to

$$|u_i[k]| \le 1, \qquad k = 1,\cdots, K, \qquad i = 1,\cdots, N, \tag{C4}$$

where $u_i[k]$ is the kth component of $\boldsymbol{u}_i$. This problem reduces to (P3)–(C3) when $N = 2$.

A further generalization of this signal design problem is to require the channel outputs to be separated in l_p norm, and the inputs to be constrained in l_q norm, for arbitrary p and

q. This l_q/l_p signal design problem is considered in [11] for the case $p = q = \infty$ (that is, the outputs are to be separated in amplitude, and the inputs are constrained in amplitude),[1] and in [12], for the case $p = q = 2$. The case considered here assumes $p = 2$ and $q = \infty$. Problem (P3)-(C3) is therefore the l_∞/l_2 signal design problem for $N = 2$.

The l_q/l_p signal design problem has the geometric interpretation of packing N points in some region in $\mathbb{R}^{K+M-1}$ to maximize minimum l_p-distance between the points. The l_2/l_2 problem considered in [12] is equivalent to packing N points in an ellipsoid to maximize minimum Euclidean distance. The l_∞/l_2 problem is equivalent to packing N points in the parallelopiped $\boldsymbol{P}$, defined earlier, again to maximize minimum Euclidean distance. Here we derive a local optimality condition for this problem, which is a generalization of the fixed point condition (FP-3).

We again replace the amplitude constraint in (C4) by the l_p-constraints

$$\frac{1}{K}\sum_{k=1}^{K} |u_i[k]|^p \le 1, \qquad i = 1,\cdots,N. \tag{57}$$

However, the objective function $d_{\min}$ is not a differentiable function of $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_N$, so that we cannot immediately derive a necessary condition for optimality. We can, however, replace $d_{\min}$ by another function, called the *potential function*, which is smooth, and closely approximates $d_{\min}$. This technique was used in [12] to obtain numerical solutions to the l_2/l_2 signal design problem via gradient search. For example, one candidate for the potential function is

$$f = -\frac{1}{W}\ln\left(\sum_{i\neq j} e^{-Wd_{ij}^2}\right), \tag{58}$$

where $d_{ij} = \|\boldsymbol{H}(\boldsymbol{u}_i - \boldsymbol{u}_j)\|$, and W is some large constant. For any signal set, as $W \to \infty$, f converges to $d_{\min}^2$.

For fixed K and N, define the Lagrangian as

$$L(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_N) = f - \sum_{i=1}^{N}\lambda_i\left(\sum_{k=1}^{K}|u_i[k]|^p\right), \tag{59}$$

where the λ_i's are chosen to satisfy the constraints (57). Taking the gradient of L with respect to $\boldsymbol{u}_i$ gives

$$\frac{\sum_{\substack{j=1\\ j\neq i}}^{N} e^{-Wd_{ij}^2}\Phi(\boldsymbol{u}_i - \boldsymbol{u}_j)}{\sum_{i\neq j} e^{-Wd_{ij}^2}} = \frac{\lambda_i p}{2}\boldsymbol{v}_i, \qquad i = 1,\cdots,N, \tag{60a}$$

where

$$\boldsymbol{v}_i' = \left[u_i^{p-1}[1] \quad u_i^{p-1}[2] \cdots u_i^{p-1}[K]\right], \tag{60b}$$

assuming p is even.

We now take limits as $W \to \infty$ and $p \to \infty$. The order in which these limits are taken is inconsequential. The left-hand side of (60a) can be rewritten as

$$\frac{e^{-Wd_{\min}^2}\sum_{\substack{j=1\\ j\neq i}} e^{-W(d_{ij}^2 - d_{\min}^2)}\Phi(\boldsymbol{u}_i - \boldsymbol{u}_j)}{e^{-Wd_{\min}^2}\sum_{i\neq k} e^{-W(d_{ij}^2 - d_{\min}^2)}} \to \frac{1}{M}\sum_{j\in J_i}\Phi(\boldsymbol{u}_i - \boldsymbol{u}_j) \qquad \text{as } W \to \infty, \tag{61}$$

where $J_i = \{j: \|\boldsymbol{H}(\boldsymbol{u}_i - \boldsymbol{u}_j)\| = d_{\min}\}$, and $M = \sum_{i=1}^{N}|J_i|$, where $|J_i|$ is the cardinality of J_i. Note that we can always assume that $|J_i| \ge 1$ for each i. Otherwise, there exists a point which is at distance $d > d_{\min}$ from all other points. This point can always be moved so that it is at distance $d_{\min}$ from some other point without decreasing $d_{\min}$.

The Lagrange multipliers, $\lambda_1,\cdots,\lambda_N$, are determined by taking the inner product of both sides of (60a) with $\boldsymbol{u}_i$, which after letting $W \to \infty$ gives

$$\sum_{j\in J_i}\boldsymbol{u}_i'\Phi(\boldsymbol{u}_i - \boldsymbol{u}_j) = \frac{\lambda_i p}{2}\|\boldsymbol{u}_i\|^p. \tag{62}$$

If the constraint (57) is satisfied with equality for particular i, then $\lambda_i p/2 = \sum_{j\in J_i}\boldsymbol{u}_i'\Phi(\boldsymbol{u}_i - \boldsymbol{u}_j)$. Otherwise, if strict inequality holds for particular i, then $\lambda_i = 0$, and as $W \to \infty$, (60a) becomes

$$\sum_{j\in J_i}(\boldsymbol{u}_i - \boldsymbol{u}_j) = 0, \tag{OC-1}$$

assuming that Φ is nonsingular. In this case, the sum of vectors from nearest neighbor points to $\boldsymbol{u}_i$ is therefore zero. This is intuitively satisfying, since the assumption that (57) is satisfied with strict inequality implies that $\boldsymbol{u}_i(\boldsymbol{y}_i)$ is in the interior of the input (output) region determined by (57). It therefore makes sense that there is no direction in which this point can be moved so that its minimum distance to other points increases, which is implied by (OC-1). On the other hand, when (57) is satisfied with equality, $\boldsymbol{u}_i$ is on the boundary of the constraint set, so that the distance between this point and its nearest neighbors can be increased by moving away from the constraint region. Note that summing both sides of (62) over i gives $\sum_{i=1}^{N'}\lambda_i = 2\gamma d_{\min}^2/p$, where N' is the number of points for which (57) is satisfied with equality, and γ is the number of pairs of inputs that are separated by $d_{\min}$. (When $N = 2$, $\gamma = 1$.)

The limit $p \to \infty$ can now be taken in much the same way as when $N = 2$, giving the condition

$$\operatorname{sgn}\left(\sum_{j\in J_i}\Phi(\boldsymbol{u}_i - \boldsymbol{u}_j)\right) = \boldsymbol{u}_i, \tag{OC-2}$$

for points $\boldsymbol{u}_i$ that satisfy the amplitude constraint (C4) with equality in at least one component (i.e., the l_∞-norm of $\boldsymbol{u}_i$, $\|\boldsymbol{u}_i\|_\infty = 1$). If $\|\boldsymbol{u}_i\|_\infty < 1$, then the condition (OC-1) applies. Observe that (OC-2) implies that if $\boldsymbol{u}_i$ satisfies (C4) with equality, then the components of $\boldsymbol{u}_i$ can only be 1, -1, or 0. The analogous condition for l_2/l_2 signal design, i.e.,

[1] A continuous-time channel is assumed throughout most of [11], which leads to the L_∞/L_∞ signal design problem.

when $p = 2$ in (57), is [12] (see also [13]),

$$\sum_{j \in J_i} \Phi(\boldsymbol{u}_i - \boldsymbol{u}_j) = \lambda \boldsymbol{u}_i, \tag{63}$$

for all $\boldsymbol{u}_i$ such that $\| \boldsymbol{u}_i \|^2 = K$, and is (OC-1) for all $\boldsymbol{u}_i$ such that $\| \boldsymbol{u}_i \|^2 < K$. The constant λ is selected to satisfy (57). This condition (63) has the geometric interpretation that the sum of the vectors from nearest neighbor points to $\boldsymbol{u}_i$ is colinear with $\Phi^{-1}\boldsymbol{u}_i$.

It is pointed out in [12] that the condition (63) is not a necessary condition for local optimality in the sense of (P4). That is, for given Φ and N it may be possible to construct a set of points $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N$ for which (63) is not satisfied for a particular point $\boldsymbol{u}_i$, but there is no direction in which $\boldsymbol{u}_i$ can be moved to increase the minimum distance to nearest neighbors. This is because the objective function $d_{\min}$ is not differentiable. However, it is shown in the appendix that there exists at least one (globally optimal) solution to (P4–C4) for which (OC-1) or (OC-2) is satisfied for every $\boldsymbol{u}_i$. (This proof also implies that there exists a global solution to the l_2/l_2 signal design problem for which every input satisfies (OC-1) or (63).) A numerical gradient search algorithm based on maximizing the potential function (58) for fixed W is used in [12] to design energy constrained signal sets, and can be easily modified to use the amplitude constraint (C4) instead.

IX. Conclusion

We have derived a necessary fixed-point condition for solutions to the problem of maximizing the output energy of a linear, time-invariant channel subject to a time- and amplitude-constrained input, and have shown that it leads to a practical numerical algorithm for finding locally optimal solutions. For the continuous-time problem the fixed-point condition was derived in two ways: using the Pontryagin maximum principle starting with a finite-dimensional state-space description of the channel, and using a Lagrange multiplier argument starting with the channel impulse response. We also showed that fixed points which are ± 1 almost everywhere must be local maxima.

The results presented here lead to a number of open questions, some of which are now listed.

1) Is it possible to characterize the solutions to (FP-1)–(FP-4) for any impulse response? How many solutions can there be?
2) What restrictions on the impulse response are necessary to guarantee that the solution to (FP-2) switches a finite number of times?
3) Are all solutions to (FP-2) for two-pole channels symmetric about $T/2$?
4) Is the single square pulse always the solution to (P1)–(C1) for the ideal lowpass channel?
5) Is the discrete-time problem (P3)–(C3) for an arbitrary impulse response NP-hard?

Another direction that can be explored is how to construct amplitude-limited signal sets that satisfy the optimality conditions (OC-1), (OC-2). A related problem is the asymptotic relationship between N and the maximum $d_{\min}$ for l_∞/l_2 signal design as $K \to \infty$. That is, for a fixed information rate $R = \log N/K$, how does the maximum $d_{\min}$ increase with K as $K \to \infty$?

An anonymous reviewer has pointed out that optimization of the transmitted pulse in digital communications must take into account the effect of intersymbol interference, which has been ignored here. A signal design problem related to those considered here, which does take intersymbol interference into account, is the following (see [14]): Assuming binary pulse amplitude modulation through a known dispersive channel, design an amplitude-limited transmitted pulse to maximize the minimum distance between sequences of received samples. Of course, other variations exist, and provide interesting possibilities for future investigation.

Acknowledgment

The authors thank A. Friedman and an anonymous reviewer for providing helpful comments and corrections to the original manuscript.

Appendix

We first show that $\lim_{p \to \infty} E_p = E_\infty$, where E_p is the maximum energy in (P1) given that the input satisfies the L_p constraint (17). Let u_p denote an optimal input. E_∞ is then the maximum energy when the input is constrained in amplitude (L_∞), and u_∞ is an optimal input in this case. Note that $E_p \geq E_\infty$ for any $p \geq 0$, since an input that satisfies the amplitude constraint also satisfies the L_p-constraint.

We first rewrite (18) and (23) as

$$\int_0^T \Phi(t-s) u_p(s)\, ds = \frac{E_p}{T} u_p^{p-1}(t) \tag{A.1}$$

for even p. Let $U_p = \sup_{t \in [0,T]} |u_p(t)|$. Then

$$\begin{aligned} \frac{E_p}{T} |u_p(t)|^{p-1} &= \left| \int_0^T \Phi(t-s) u_p(s)\, ds \right| \\ &\leq U_p \int_0^T |\Phi(t-s)|\, ds \\ &\leq U_p \kappa, \end{aligned} \tag{A.2}$$

where

$$\kappa = \sup_{t \in [0,T]} \int_0^T |\Phi(t-s)|\, ds. \tag{A.3}$$

Since this holds for all t, it follows that

$$\frac{E_p}{T} U_p^{p-1} \leq U_p \kappa \tag{A.4}$$

or

$$U_p \leq \left(\frac{\kappa T}{E_p} \right)^{\frac{1}{p-2}} \leq \left(\frac{\kappa T}{E_\infty} \right)^{\frac{1}{p-2}} = \alpha^{\frac{1}{p-2}}, \tag{A.5}$$

where α is a constant that depends only on T and $h(\cdot)$, and is finite provided that $\kappa < \infty$. It is easily shown that $\alpha \geq 1$. As $p \to \infty$, the maximum magnitude of the maximizing functions u_p therefore tends to one.

Since any optimal input u_p satisfies (A.5), E_p is less than the maximum energy at the output of the channel with impulse response $h(\cdot)$, subject to the constraint that the input is bounded in magnitude by $\alpha^{1/(p-2)}$ on $[0, T]$. Denote this energy by F_p. Note that the input $\alpha^{1/(p-2)}u_\infty$, where u_∞ is any solution to (P1)-(C1), achieves this maximum energy, so that

$$E_\infty \le E_p \le F_p = \alpha^{\frac{2}{p-2}} E_\infty, \tag{A.6}$$

and letting $p \to \infty$ shows that $E_p \to E_\infty$.

We now show that there exists at least one solution to (P4)-(C4), $u_1, \cdots, u_N$, that satisfies the conditions (OC-1), (OC-2). From (58),

$$\begin{aligned} f &= -\frac{1}{W} \ln\left(e^{-Wd_{\min}^2} \sum_{i \ne j} e^{-W(d_{ij} - d_{\min}^2)} \right) \\ &= d_{\min}^2 - \frac{1}{W} \ln\left(\sum_{i \ne j} e^{-W(d_{ij} - d_{\min}^2)} \right) \\ &\le d_{\min}^2, \end{aligned} \tag{A.7}$$

since $e^{-W(d_{ij} - d_{\min}^2)} = 1$ for at least one (i, j) pair. Note that (A.7) is independent of the input constraint.

For fixed p, W, and N, let $f_{\max}(p; W)$ be the maximum achievable value of f, assuming that the inputs satisfy the l_p-constraint (57), and let d_f be the corresponding minimum distance of any signal set which achieves $f_{\max}$. Also, let $d_{\max}(p) = \max d_{\min}$ over all signal sets, assuming the constraint (57), and let $f_d(p; W)$ be the value of the potential function (58) evaluated for any particular signal set that achieves $d_{\max}$. Then (A.7) implies that for all p and W,

$$f_d(p; W) \le f_{\max}(p; W) \le d_f^2(p) \le d_{\max}^2(p). \tag{A.8}$$

For any signal set that satisfies the l_p-constraint (57), we have that

$$|u_i[k]| \le K^{1/p}, \qquad i = 1, \cdots, N, \qquad k = 1, \cdots, K. \tag{A.9}$$

Clearly, all solutions to (P4) subject to constraint (A.9) can be generated by scaling solutions to (P4) with constraint (C4) by $K^{1/p}$. The maximized minimum distance, subject to (A.9), is therefore $K^{1/p}d_{\max}(\infty)$, where $d_{\max}(p)$ was defined earlier. The same scaling also applies when the objective function $d_{\min}$ is replaced by the potential function f. Since all signal sets that achieve $f_{\max}(p; W)$ also satisfy (A.9), we have from (A.8),

$$f_{\max}(p; W) \le K^{2/p} f_{\max}(\infty; W) \le K^{2/p} d_{\max}^2(\infty). \tag{A.10}$$

Finally, we use the fact that $f_{\max}(p; W) \ge f_{\max}(\infty; W)$, since any signal set that satisfies the constraint (57) for $0 < p < \infty$ also satisfies the amplitude constraint (C4) ($p = \infty$). Combining this with (A.8) and (A.10) gives

$$\begin{aligned} f_d(\infty; W) &\le f_{\max}(\infty; W) \le f_{\max}(p; W) \\ &\le K^{2/p} d_{\max}^2(\infty). \end{aligned} \tag{A.11}$$

Now $f_d(\infty; W) \to d_{\max}^2(\infty)$ as $W \to \infty$, so that letting both $p \to \infty$ and $W \to \infty$ shows that $f_{\max}(p; W) \to d_{\max}^2(\infty)$. Furthermore, signal sets that maximize f satisfy (OC-1), (OC-2) as $W \to \infty$ and $p \to \infty$, so that (OC-1), (OC-2) must be satisfied for some (perhaps not all) solutions to (P4)-(C4).

References

[1] A. Wyner, "Spectra of bounded functions," in *Open Problems in Communication and Computation*, T. M. Cover and B. Gopinath, Eds. New York: Springer-Verlag, 1987, pp. 46-48.

[2] J. M. Cioffi, W. L. Abbott, H. K. Thapar, C. M. Melas, and K. D. Fisher, "Adaptive equalization in magnetic-disk storage channels," *IEEE Commun. Mag.*, vol. 28, no. 2, pp. 14-29, Feb. 1990.

[3] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, D. E. Brown, transl. New York: Pergamon Press, Macmillan, 1964.

[4] L. I. Rozonoér, "L. S. Pontryagin maximum principle in the theory of optimum systems," (first of three parts in translation), *Automation Remote Contr.*, vol. 20, no. 10, pp. 1288-1301, 1959.

[5] R. Bellman, *Introduction to Matrix Analysis*. New York: McGraw-Hill, 1960.

[6] L. H. Ozarow, A. D. Wyner, and J. Ziv, "Achievable rates for a constrained Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 34, pp. 365-370, May 1988.

[7] A. Friedman, "Optimal control for parabolic variational inequalities," *SIAM J. Contr. Optimiz.*, vol. 25, no. 2, pp. 482-497, Mar. 1987.

[8] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming – Theory and Algorithms*. New York: John Wiley, 1979.

[9] M. R. Garey and D. S. Johnson, *Computers and Intractability*. San Francisco: W. H. Freeman, 1979.

[10] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, pp. 303-312, 1989.

[11] M. L. Honig, K. Steiglitz, B. Gopinath, and S. Boyd, "Bounds on maximum throughput for digital communications with finite-precision and amplitude constraints," *IEEE Trans. Inform. Theory*, vol. 36, pp. 472-484, May 1990.

[12] M. L. Honig, K. Steiglitz, and S. Norman, "Optimization of signal sets for partial response channels: Part I—Numerical techniques," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1327-1341, Sept. 1991.

[13] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "Optimization of two-dimensional signal constellations in the presence of Gaussian noise," *IEEE Trans. Commun.*, vol. COM-22, pp. 28-38, Jan. 1974.

[14] D. J. Hajela and M. L. Honig, "On finding maximally separated signals for digital communications," in *Open Problems in Communication and Computation*, T. M. Cover and B. Gopinath, Eds. New York: Springer-Verlag, 1987, pp. 92-99.