Stochastic Iterative Algorithms for Signal Set Design for Gaussian Channels and Optimality of the L2 Signal Set

Yi Sun, *Member, IEEE*

***Abstract*—We propose a stochastic iteration approach to signal set design. Four practical stochastic iterative algorithms are proposed with respect to equal and average energy constraints and sequential and batch modes. By simulation, a new optimal signal set, the L2 signal set (consisting of a regular simplex set of three signals and some zero signals), is found under the strong simplex conjecture (SSC) condition (equal *a priori* probability and average energy constraint) at low signal-to-noise ratios (SNR). The optimality of the L1 signal set, the confirmation of the weak simplex conjecture, and two of Dunbridge's theorems are some of the results obtained by simulations. The influence of SNR and *a priori* probabilities on signal sets is also investigated via simulation. As an application to practical communication system design, the signal sets of eight two-dimensional (2-D) signals are studied by simulation under the SSC condition. Two signal sets better than 8-PSK are found. In the second part of this paper, optimal properties of the L2 signal set are analyzed under the SSC condition at low SNR's. The L2 signal set is proved to be uniquely optimal in 2-D space. The class of signal sets $S(M,K)$ (consisting of a regular simplex set of K signals and $M-K$ zero signals) is analyzed. It is shown that any of the signal sets $S(M,K)$ for $3 \le K \le M-1$ disproves the strong simplex conjecture for $M \ge 4$, and if $M \ge 7$, $S(M,2)$ (the L1 signal set) also disproves the strong simplex conjecture. It is proved that the L2 signal set is the unique optimal signal set in the class of signal sets $S(M,K)$ for all $M \ge 4$. Several results obtained by Steiner for all $M \ge 7$ are extended to all $M \ge 4$. Finally, we show that for $M \ge 7$, there exists an integer $K' < M$ such that any of the signal sets consisting of K signals equally spaced on a circle and $M-K$ zero signals, for $4 \le K \le K'$, also disprove the strong simplex conjecture.**

***Index Terms*— L1 signal set, M-ary communication, simplex conjecture, stochastic iteration.**

I. Introduction

The optimal selection of signal vectors embedded in white Gaussian noise for communications has been a research topic for many years and is not known in general. In 1948, it was conjectured that, with finite energy constraints but without constraint on dimension of signal space, the M optimal signal vectors are vertices of a regular simplex in $(M-1)$-dimensional signal space [2]. This conjecture is referred to as the strong simplex conjecture (SSC) when the signal vectors are constrained only by an average energy limitation, and as the weak simplex conjecture (WSC) when the signal vectors are equal-energy-constrained.

Under the assumption that signal vectors have equal energy, Balakrishnan proved [3] that the regular simplex is 1) optimal as λ goes to infinity, 2) optimal as λ goes to zero, and 3) locally optimal at all λ, where λ^2 is the signal-to-noise ratio (SNR). Dunbridge proved further [4] that under an average energy constraint the regular simplex is 1) the optimal signal set as λ goes to infinity and 2) a local extremum at all λ. For the case of $M = 2$, the regular simplex has been proved to be optimal at all λ for both the average and equal energy constraint. Dunbridge proved, under an average energy constraint, that the regular simplex with $M = 3$ is optimal as λ goes to zero. Work on the weak simplex conjecture in [5] was shown by Farber in [6] to prove this conjecture for $M < 5$.

In 1994, Steiner published several new results [1]. The major result is that a counterexample signal set, the L1 signal set, was found and disproved the strong simplex conjecture. It was proved that the L1 signal set is better than the regular simplex signal set at low SNR ratios for all $M \ge 7$. This leads to the result that, for all $M \ge 7$, there is no signal set of M signals which is optimal at all SNR's. Furthermore, the optimal signal set at low SNR's is not an equal energy set for any $M \ge 7$. The regular simplex is shown to be the unique signal set which maximizes the minimum distance between signals. It follows that a signal set which maximizes the minimum distance is not necessarily optimal. The L1 signal set is proved to be the optimal signal set in one-dimensional (1-D) space. Steiner's work changed our view on optimal signal set design.

It is clear that the optimal signal set can only be given in the sense that the dimension of signal space, the number of signals, *a priori* probabilities, the SNR, and the energy constraint of a signal set are given. However, all previous results were obtained only in some simple cases. It is desirable to establish a constructive method to find better signal sets under various conditions.

This paper presents the following work. In Section II, we propose four practical stochastic iterative algorithms as a constructive approach to signal set design. Simulation results are also reported in this section. In Section III, the optimal properties of the newly discovered L2 signal set are analyzed in the SSC conditions at low SNR's. Many other analytical

Manuscript received November 21, 1995; revised January 10, 1997. The material in this paper was presented in part at the 1996 International Symposium on Information Theory and its Applications, ISITA'96, Victoria, BC, Canada, September 17–20.

The author was with the Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with MIRL/Radiology, University of Utah, Salt Lake City, UT 84312 USA.

Publisher Item Identifier S 0018-9448(97)05324-8.

results are also given in this section. Proofs of analytical results are included in the Appendix.

II. Proposed Stochastic Iterative Algorithms

A. Problem Formulation

Given one of M signals $\boldsymbol{s}_i \in R^n$ for $i = 1, \cdots, M$ transmitted through a channel contaminated with additive white Gaussian noise. After transmission we receive a vector

$$\boldsymbol{y} = \boldsymbol{s}_i + \boldsymbol{n}, \qquad i = 1, \cdots, M \tag{1}$$

and wish to determine which of the M signals was transmitted, where $\boldsymbol{y}, \boldsymbol{n} \in R^n$, and $\boldsymbol{n}$ is a zero-mean Gaussian vector with covariance matrix equal to the n by n identity matrix. The optimal detector in terms of maximizing the average probability of correct detection chooses the signal $\boldsymbol{s}_i$ such that $f(\boldsymbol{y} \,|\, \boldsymbol{s}_i)p_i$ is maximized, where $f(\boldsymbol{y} \,|\, \boldsymbol{s}_i)$ denotes the probability density function of $\boldsymbol{y}$ conditioned on the transmission of $\boldsymbol{s}_i$, and p_i is the *a priori* probability of the ith signal. $f(\boldsymbol{y} \,|\, \boldsymbol{s}_i)p_i$ can be written as

$$f(\boldsymbol{y} \,|\, \boldsymbol{s}_i)p_i = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{s}_i)^T(\boldsymbol{y} - \boldsymbol{s}_i) + \log p_i\right). \tag{2}$$

By defining

$$e_i(\boldsymbol{y}) = \boldsymbol{y}^T \boldsymbol{s}_i - \frac{\boldsymbol{s}_i^T \boldsymbol{s}_i}{2} + \log p_i \tag{3}$$

we find that $f(\boldsymbol{y}|\boldsymbol{s}_i)p_i$ is maximized when the signal corresponding to the maximum of $e_i(\boldsymbol{y})$ is chosen. The average probability of correct detection is

$$\begin{aligned} P_d &= \sum_{i=1}^{M} p_i P_r\left(e_i(\boldsymbol{y}) = \max_{1 \le j \le M} e_j(\boldsymbol{y}) \mid \boldsymbol{s}_i\right) \\ &= \sum_{i=1}^{M} \frac{1}{(2\pi)^{n/2}} \int_{V_i} \exp\left(-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{y} + e_i(\boldsymbol{y})\right) d\boldsymbol{y} \end{aligned} \tag{4}$$

where $d\boldsymbol{y}$ denotes the volume integral and

$$V_i = \left\{\boldsymbol{y} \mid e_i(\boldsymbol{y}) = \max_{1 \le j \le M} e_j(\boldsymbol{y})\right\}, \qquad i = 1, 2, \cdots, M \tag{5}$$

are decision regions. The probability of correct detection can be rewritten as

$$\begin{aligned} P_d &= \frac{1}{(2\pi)^{n/2}} \int_{R^n} \exp\left(-\frac{1}{2}\boldsymbol{y}^T\boldsymbol{y} + \max_{1 \le i \le M} e_i(\boldsymbol{y})\right) d\boldsymbol{y} \\ &= E\left[\exp\left(\max_{1 \le i \le M} e_i(\boldsymbol{z})\right)\right] \\ &= \sum_{i=1}^{M} \int_{V_i} \exp(e_i(\boldsymbol{z})) f_{\boldsymbol{z}}(\boldsymbol{z})\, d\boldsymbol{z} \end{aligned} \tag{6}$$

where $\boldsymbol{z}$ is an n-dimensional Gaussian vector with zero mean and n by n identity covariance matrix. $f_{\boldsymbol{z}}(\boldsymbol{z})$ is the probability density function of $\boldsymbol{z}$ defined as

$$f_{\boldsymbol{z}}(\boldsymbol{z}) = \frac{1}{(2\pi)^{n/2}} \exp(-\frac{1}{2}\boldsymbol{z}^T\boldsymbol{z}).$$

The task of optimal signal set design is, given the number M of signals, dimension n of the signal space, *a priori* probabilities p_i's of signals, the SNR λ^2, and the energy constraint of a signal set, to find a signal set such that the correct detection probability P_d is maximized.

B. Definitions

A special case of the *a priori* probabilities is equally likely signals, which means that the *a priori* probability of every signal is identical to $1/M$. Most optimal signal set design results obtained previously, including the simplex conjectures and Steiner's results, are in this condition.

The energy constraint is usually imposed in the following ways:

1) Average energy constraint: the average energy of a signal set is defined as

$$\frac{1}{M}\sum_{i=1}^{M} \|\boldsymbol{s}_i\|^2 = \lambda^2 \tag{7}$$

where λ^2 is called the SNR of the signal set, and $\| \, \|$ denotes the l_2 norm.

2) Equal energy constraint: each signal has equal energy

$$\|\boldsymbol{s}_i\|^2 = \lambda^2, \qquad i = 1, \cdots, M. \tag{8}$$

In this paper, the condition that a signal set is equally likely and is average-energy-constrained is referred to as the strong simplex conjecture (SSC) condition. The condition that a signal set is equally likely and is equal-energy-constrained is referred to as the weak simplex conjecture (WSC) condition. We define the SSC and WSC conditions in this way because the strong and weak simplex conjectures were made in the SSC and WSC conditions, respectively.

The following two classes of signal sets are analyzed in this paper.

Definition 1 (The class of signal sets $E(M,K)$): For $M \ge 2$ and $2 \le K \le M$, given K two-dimensional unit vector $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K$, which are equally spaced on the unit circle. The signal set $E(M,K)$ consists of K signal vectors $\lambda\sqrt{M/K}\,\boldsymbol{v}_i$, $i = 1, \cdots, K$, and $M - K$ signals at the origin, i.e.,

$$E(M,K) = \{\lambda\sqrt{M/K}\boldsymbol{v}_1, \cdots, \lambda\sqrt{M/K}\boldsymbol{v}_K, \boldsymbol{0}, \cdots, \boldsymbol{0}\}. \quad \Box$$

Definition 2 (The class of signal sets $S(M,K)$): For $M \ge 2$ and $2 \le K \le M$, given K unit vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K$, which form a $(K-1)$-dimensional regular simplex with its center at the origin. The signal set $S(M,K)$ consists of K signals $\lambda\sqrt{M/K}\,\boldsymbol{v}_i$, $i = 1, \cdots, K$, and $M - K$ signals at the origin, i.e.,

$$S(M,K) = \{\lambda\sqrt{M/K}\boldsymbol{v}_1, \cdots, \lambda\sqrt{M/K}\boldsymbol{v}_K, \boldsymbol{0}, \cdots, \boldsymbol{0}\}. \quad \Box$$

The signal sets $E(M,K)$ (or $S(M,K)$) for all $M \ge 2$ and $2 \le K \le M$ have the same average energy λ^2 and satisfy the average energy constraint (7). Only for a fixed M, the signal sets $E(M,K)$ (or $S(M,K)$) for $2 \le K \le M$ have the same total energy $M\lambda^2$. $E(2,2)$ as well as $S(2,2)$ is the regular simplex set of two signals (i.e., the antipodal signal set). For $M \ge 3$, $E(M,2)$ as well as $S(M,2)$ is the L1 signal set. For

$M \geq 4$, $E(M,3)$ as well as $S(M,3)$ is called the L2 signal set in this paper for the same reason as those for naming the L1 signal set in [1]. $S(M,M)$ is the pure regular simplex set of M signals. In addition, all the signal sets $S(M,K)$ for $K=4,5,\cdots,M-1$ have a similar structure to those of the L1 and L2 signal sets. For the same reason as that for naming the L1 and L2 signal sets, we can call the signal set $S(M,4)$ the L3 signal set, $S(M,5)$ the L4 signal set, etc.

C. Proposed Stochastic Iterative Algorithms

Optimal signal sets are necessarily located at the local maximum points of the detection probability. From (6), the gradient of the detection probability P_d with respect to $\boldsymbol{s}_i$ is given by

$$\nabla_i P_d = \sum_{j=1}^{M} \int_{\Delta V_j} \exp(e_j(\boldsymbol{z})) f_{\boldsymbol{z}}(\boldsymbol{z})\, d\boldsymbol{z} + \int_{V_i} \exp(e_i(\boldsymbol{z}))(\boldsymbol{z}-\boldsymbol{s}_i) f_{\boldsymbol{z}}(\boldsymbol{z})\, d\boldsymbol{z}, \qquad i=1,2,\cdots,M \tag{9}$$

where ΔV_j is the boundary displacement of decision region V_j produced by the displacement of the ith signal from $\boldsymbol{s}_i$ to $\boldsymbol{s}_i + d\boldsymbol{s}_i$. The summation term is negligible compared to the integration term. We can establish a stochastic iterative algorithm that in stochastic sense, converges toward the direction of the gradient to a maximum point of the detection probability.

Let $\{\boldsymbol{z}(k)\}$ denote a sequence of independent and identically distributed (i.i.d.) Gaussian random vectors with zero mean and an n by n identity covariance matrix. A fundamental stochastic iterative algorithm is proposed as follows.

Algorithm 1 (Fundamental algorithm): Given $\boldsymbol{s}_i(0)$, $i=1,2,\cdots,M$. Assume

$$e_i(k) = \max_{1\leq j\leq M} e_j(k)$$

where

$$e_j(k) = \boldsymbol{z}^T(k)\boldsymbol{s}_j(k) - \frac{1}{2}\boldsymbol{s}_j^T(k)\boldsymbol{s}_j(k) + \log p_j, \qquad j=1,2,\cdots,M. \tag{10}$$

Compute

$$\Delta\boldsymbol{s}_i(k) = \exp(e_i(k))(\boldsymbol{z}(k) - \boldsymbol{s}_i(k)). \tag{11}$$

Update signal vectors by

$$\boldsymbol{s}_i(k+1) = \boldsymbol{s}_i(k) + \alpha(k)\Delta\boldsymbol{s}_i(k) \tag{12}$$

and $\boldsymbol{s}_j(k+1) = \boldsymbol{s}_j(k)$ for $j \neq i$, $j=1,2,\cdots,M$. □

We require that the sequence of updating coefficients $\{\alpha(k)\}$ satisfy: i) $\alpha(k) > 0$, ii)

$$\sum_{k=0}^{\infty} \alpha(k) = \infty$$

and iii)

$$\sum_{k=0}^{\infty} \alpha^2(k) < \infty.$$

We assume that the sequences of updating coefficients in all proposed algorithms in this paper satisfy these three conditions. The fundamental algorithm represents a stochastic dynamic system driven by the sequence $\{\boldsymbol{z}(k)\}$. In [13], the dynamic behavior of Algorithm 1 is analyzed by using the theory developed in [7]–[9].

To be practically useful, an algorithm must be energy-constrained. One of the necessary conditions for an optimal signal set under the average energy constraint is that the signal vectors must satisfy the following symmetric condition [4]:

$$\sum_{j=1}^{M} p_j \boldsymbol{s}_j = 0. \tag{13}$$

Algorithm 2 (Average energy constraint, sequential mode): Given initial signal vectors $\boldsymbol{s}_i(0)$, $i=1,2,\cdots,M$, which satisfy conditions (7) and (13). Assume

$$e_i(k) = \max_{1\leq j\leq M} e_j(k).$$

Compute $\Delta\boldsymbol{s}_i(k)$ by (11). Update the signal vectors by

$$\boldsymbol{s}_i(k+1) = \beta(\boldsymbol{s}_i(k) + \alpha(k)\Delta\boldsymbol{s}_i(k)) \tag{14}$$

and

$$\boldsymbol{s}_j(k+1) = \beta(\boldsymbol{s}_j(k) - \boldsymbol{c}), \qquad j \neq i,\ j=1,2,\cdots,M \tag{15}$$

where

$$\boldsymbol{c} = \frac{1}{1-p_i}\left(p_i\alpha(k)\Delta\boldsymbol{s}_i(k) + \sum_{j=1}^{M} p_j\boldsymbol{s}_j(k)\right) \tag{16}$$

and

$$\beta = \left(\frac{M\lambda^2}{\|\boldsymbol{s}_i(k)+\alpha(k)\Delta\boldsymbol{s}_i(k)\|^2 + \sum_{j=1, j\neq i}^{M}\|\boldsymbol{s}_j(k)-\boldsymbol{c}\|^2}\right)^{\frac{1}{2}}. \tag{17}$$

□

In Algorithm 2, when the ith-signal vector is adjusted by $\alpha(k)\Delta\boldsymbol{s}_i(k)$, each of the other $M-1$ signal vectors must be adjusted by $-p_i\alpha(k)\Delta\boldsymbol{s}_i(k)/(1-p_i)$ to keep the symmetric condition. In order to avoid cumulative quantization errors, we use (16) rather than $\boldsymbol{c} = p_i\alpha(k)\Delta\boldsymbol{s}_i(k)/(1-p_i)$ to obtain the displacement $\boldsymbol{c}$. For equally likely signals, $p_i = 1/M$ for $i=1,2,\cdots,M$. In this case, i.e., in the SSC condition, $\boldsymbol{c} = \alpha(k)\Delta\boldsymbol{s}_i(k)/(M-1)$.

Algorithm 3 (Equal energy constraint, sequential mode): Given initial signal vectors $\boldsymbol{s}_i(0)$, $i=1,2,\cdots,M$, which satisfy condition (8). Assume

$$e_i(k) = \max_{1\leq j\leq M} e_j(k).$$

Compute $\Delta\boldsymbol{s}_i(k)$ by (11). Update signal vectors by

$$\boldsymbol{s}_i(k+1) = \lambda\frac{\boldsymbol{s}_i(k)+\alpha(k)\Delta\boldsymbol{s}_i(k)}{\|\boldsymbol{s}_i(k)+\alpha(k)\Delta\boldsymbol{s}_i(k)\|} \tag{18}$$

and $\boldsymbol{s}_j(k+1) = \boldsymbol{s}_j(k)$ for $j \neq i$, $j=1,2,\cdots,M$. □

To speed up convergence in some cases, we need to establish batch algorithms. Given i.i.d. Gaussian vectors $\boldsymbol{z}_l$, $l=1,2,\cdots,L$, with zero mean and the $n \times n$ identity covariance matrix. Expression (9) can be approximated in the Monte Carlo sense as

$$\nabla_i P_d(k) = \sum_{\boldsymbol{z}_l \in V_i(k)} \exp(e_i(\boldsymbol{z}_l))\,(\boldsymbol{z}_l - \boldsymbol{s}_i(k)),\qquad i = 1, 2, \cdots, M \tag{19}$$

where $e_i(\boldsymbol{z}_l)$ is computed by (10) with $\boldsymbol{z}(k)$ replaced by $\boldsymbol{z}_l$, and

$$V_i(k) = \{\boldsymbol{z}_l \mid e_i(\boldsymbol{z}_l) = \max_{1\le j\le M} e_j(\boldsymbol{z}_l)\}.$$

In the following algorithms, if $V_i(k)$ is empty, $\nabla_i P_d(k) = 0$.

Algorithm 4 (Average energy constraint, batch mode): Given initial signal vectors $\boldsymbol{s}_i(0)$, $i = 1, 2, \cdots, M$, which satisfy conditions (7) and (13). Compute $\nabla_i P_d(k)$ for $i = 1, 2, \cdots, M$ by (19). Update signal vectors by

$$\boldsymbol{s}_i(k+1) = \beta(\boldsymbol{s}_i(k) + \alpha(k)\nabla_i P_d(k) - \boldsymbol{c}),\quad i = 1, 2, \cdots, M$$

where

$$\boldsymbol{c} = \sum_{i=1}^{M} p_i(\boldsymbol{s}_i(k) + \alpha(k)\nabla_i P_d(k))$$

and

$$\beta = \left(\frac{M\lambda^2}{\sum_{i=1}^{M} \|\boldsymbol{s}_i(k) + \alpha(k)\nabla_i P_d(k) - \boldsymbol{c}\|^2}\right)^{\frac{1}{2}}. \qquad \square$$

Algorithm 5 (Equal energy constraint, batch mode): Given initial signal vectors $\boldsymbol{s}_i(0)$, $i = 1, 2, \cdots, M$, which satisfy condition (8). Compute $\nabla_i P_d(k)$ for $i = 1, 2, \cdots, M$ by (19). Update the signal vectors by

$$\boldsymbol{s}_i(k+1) = \lambda \frac{\boldsymbol{s}_i(k) + \alpha(k)\nabla_i P_d(k)}{\|\boldsymbol{s}_i(k) + \alpha(k)\nabla_i P_d(k)\|},\quad \text{for } i = 1, 2, \cdots, M.$$

$\square$

D. Simulation Results

Simulations on the four practical algorithms are carried out. In simulations, all samples of n-dimensional Gaussian random vectors are generated by MATLAB Version 4.2c.1. The initial signal vectors are randomly set to the vectors in $[-0.5, 0.5]^n$ and then are scaled and translated (if necessary) to satisfy energy constraints and the symmetric condition. In all simulations, we chose $\alpha(k) = (2M/(k+1))\exp(\lambda^2/2)$ for $k = 0, 1, 2, \cdots$. The detection probability of a signal set is obtained by Monte Carlo simulation in which the same samples are used as those used to obtain the signal set in a stochastic iterative algorithm. In all figures showing a signal set, signals are normalized by dividing by λ.

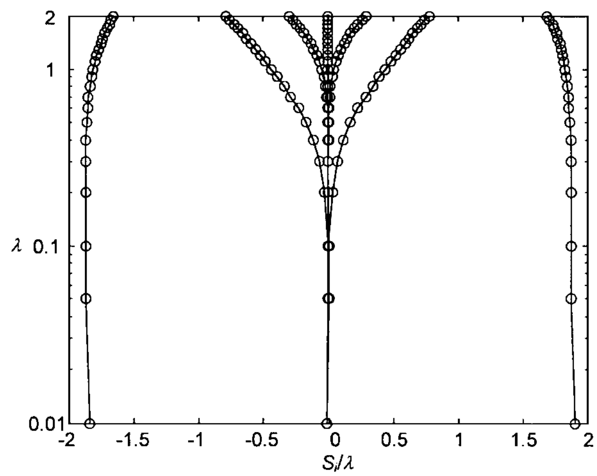

Fig. 1. 1-D signal sets of seven signals in the SSC condition. o's denote signal positions to which Algorithm 2 converges for fixed λ. As λ tends to zero, the signal set converges to the L1 signal set.

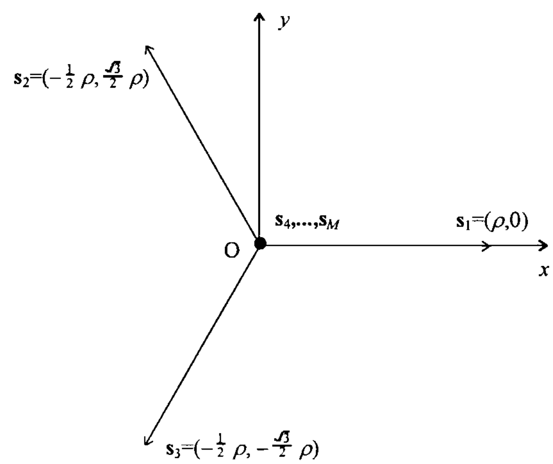

Fig. 2. The L2 signal set. $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, and $\boldsymbol{s}_3$ form a regular simplex set with signal length $\rho = \lambda\sqrt{M/3}$, and all others are zero.

1) The L1 Signal Set and the Influence of SNR on Signal Sets: Steiner proved in [1, Theorem 5] that in the SSC condition at low SNR's the L1 signal set is uniquely optimal in the class of 1-D signal sets. Fig. 1 shows simulation results for 1-D signal sets of seven signals in the SSC condition. Algorithm 2 is used. The number of iterations is 10^6 in obtaining one signal set. As shown in Fig. 1, as the SNR approaches zero, the signal set in the simulation approaches the L1 signal set. This verifies Steiner's theorem. From Fig. 1, one can observe how these signal sets change as SNR changes.

2) Discovery of the L2 Signal Set: Simulation results for two-, three-, and four-dimensional signal sets in the SSC condition are shown in Table I. Algorithm 2 is used. The number of iterations is 10^6 for each test. The purpose of the simulation is to inspect the optimal signal sets at low SNR's in higher dimensional space. For interpretation of these results, we note that each signal set satisfies the symmetric condition (13), which implies that the sum of the M equally likely signal vectors is equal to zero. This means that if three of the M signal vectors are nonzero and all others are zero, all the M signal vectors necessarily lie in a 2-D hyperplane

TABLE I
SIMULATION RESULTS OF SIGNAL SETS IN THE SSC CONDITION

n	M	λ	$\|\|s'_1\|\|$	$\|\|s'_2\|\|$	$\|\|s'_3\|\|$	$\|\|s'_4\|\|$	$\|\|s'_5\|\|$	$\|\|s'_6\|\|$	$\|\|s'_7\|\|$	set
2	3	1	1.0018	0.9958	1.0024					RS
2	3	10^{-3}	0.9990	0.9945	1.0065					RS
2	4	10^{-2}	1.1579	1.1512	1.1550	0.0026				L2
2	5	10^{-1}	1.2928	0.0008	1.2890	1.2912	0.0004			L2
2	6	10^{-1}	0.0009	0.0009	1.4149	0.0004	1.4160	1.4117		L2
2	7	10^{-1}	0.0010	1.5226	0.0013	0.0013	1.5319	0.0017	1.5280	L2
3	3	1	1.0165	0.9714	1.0114					RS
3	3	10^{-3}	0.9945	1.0127	0.9926					RS
3	4	10^{-1}	1.1415	1.1548	1.1676	0.0074				L2
3	5	5×10^{-2}	1.2787	0.0035	1.3075	1.2866	0.0027			L2
3	6	10^{-2}	1.4097	1.4131	0.0009	1.4199	0.0009	0.0009		L2
3	7	10^{-2}	0.0007	1.5250	1.5286	0.0008	1.5290	0.0008	0.0008	L2
4	3	10^{-1}	1.0142	0.9933	0.9923					RS
4	4	10^{-2}	1.1530	0.0031	1.1611	1.1500				L2
4	5	10^{-2}	1.2918	1.2840	1.2972	0.0055	0.0051			L2
4	6	5×10^{-4}	1.4206	1.4082	0.0019	1.4138	0.0019	0.0019		L2
4	7	10^{-4}	0.0021	1.5167	1.5305	1.5353	0.0021	0.0021	0.0021	L2

Note: (i) $s'_i = s_i/\lambda$; (ii) RS denotes a regular simplex set.

containing the origin. In terms of Table I, we can make the following observations.

1) When $M = 3$, at both high and low SNR's, Algorithm 2 converges to the regular simplex set that is proved to be optimal by Dunbridge [4]. The result is independent of dimension of signal space.
2) When $M \geq 4$ at low SNR's, Algorithm 2 always converges to the L2 signal set, which resulted in the discovery of the L2 signal set. The structure of the L2 signal set is shown in Fig. 2.

3) Confirmation of the Weak Simplex Conjecture: Table II shows simulation results in the WSC condition. Algorithm 5 is used. In the table, K is the total number of iterations and L is the number of samples used in each iteration. For a given signal set of M signals $\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_M$, the normalized signal covariance matrix $\boldsymbol{R}(\lambda_{ij})$ is defined by $\lambda_{ij} = \boldsymbol{s}_i^T \boldsymbol{s}_j/\lambda^2$. As is well known [10], the normalized signal covariance matrix of a regular simplex $\boldsymbol{R}(\gamma_{ij})$ has identical diagonal entries $\gamma_{ii} = 1$ and off-diagonal entries $\gamma_{ij} = -1/(M-1)$. In Table II, r is defined as relative difference of normalized signal covariance matrices between a signal set and the regular simplex

$$r = \sqrt{\sum_{i,j=1}^{M} (\lambda_{ij} - \gamma_{ij})^2 / \sum_{i,j=1}^{M} \gamma_{ij}^2}.$$

For $n = 2, 3, 4, 5, 6,$ and 7, simulations are done for $M = 2, \cdots, n+1$. For simplicity, only those results for $M = n+1$ with each given n are listed in Table II. In all simulation results including those that are not listed in the table, $r < 0.0500$. Hence, in all simulations, Algorithm 5 converges to regular simplex sets. In summary, the simulation results verify Farber's work [6] that the weak simplex conjecture is true for $M \leq 4$. And furthermore, all the results for $M = 5, 6, 7,$ and 8 suggest that the weak simplex conjecture is true, which has not been theoretically proved yet.

4) Influence of a priori Probabilities on Signal Sets: Fig. 3 shows simulation results for $n = 2$, $M = 4$, and $\lambda = 1$ under

TABLE II
CONFIRMATION OF THE WEAK SIMPLEX CONJECTURE

n	M	λ	K	L	r	set
2	3	1	40	25,000	0.0054	RS
2	3	10^{-2}	40	25,000	0.0053	RS
3	4	1	120	16,660	0.0159	RS
3	4	10^{-2}	60	16,660	0.0230	RS
4	5	1	80	25,000	0.0492	RS
4	5	10^{-2}	80	12,500	0.0470	RS
5	6	1	150	20,000	0.0494	RS
5	6	10^{-2}	100	10,000	0.0405	RS
6	7	1	200	30,000	0.0495	RS
6	7	10^{-2}	100	10,000	0.0241	RS
7	8	1	500	100,000	0.0396	RS
7	8	10^{-2}	50	50,000	0.0223	RS

the average energy constraint. Algorithm 2 is used. The total number of iterations is 10^6 in each operation to obtain one signal set. Fig. 4 shows simulation results for $n = 2$, $M = 4$, and $\lambda = 0.5$ under the equal energy constraint. Algorithm 5 is used. The total number of iterations is 25 000 in each operation and the number of samples used in each iteration is 100. In Figs. 3 and 4, the signal sets are normalized such that $\boldsymbol{s}_1$'s have zero phase angle. From Figs. 3 and 4 we can see that the larger the *a priori* probability of a signal, the larger the decision region of it. As shown in Fig. 3, under the average energy constraint, an unequally likely signal set has a structure completely different from that of an equally likely signal set.

5) Verification of Dunbridge Theorem: Dunbridge proved in [4] that under the equal energy constraint, if *a priori* probabilities have ordering $p_1 = \cdots = p_N > p_{N+1} \geq \cdots \geq p_M$, the optimal signal structure at vanishingly small SNR is that in which $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_N$ form an $(N-1)$-dimensional regular simplex. This theorem is verified in our simulation. Fig. 5 shows the simulation result in which $\lambda = 0.1$. In Fig. 5, the signal sets are normalized such that $\boldsymbol{s}_1$'s have the zero phase angle. As shown in the figure, when signals are equally likely, the signals are equally spaced, which is proved to be optimal

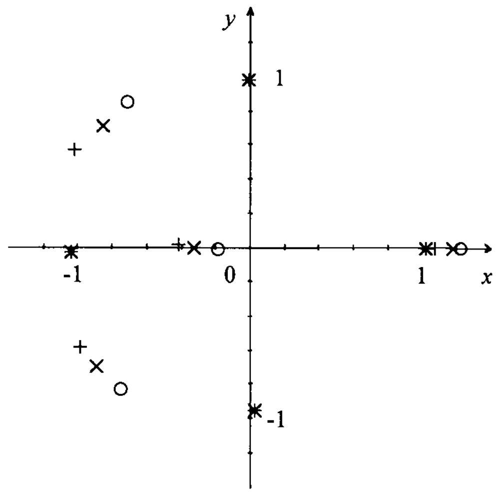

Fig. 3. Influence of *a priori* probabilities on signal sets under the average energy constraint. $\lambda = 1$. $*$: equally likely; o: $p_1 = 0.31$, $p_2 = p_3 = p_4 = 0.23$; $\times$: $p_1 = 0.37$, $p_2 = p_3 = p_4 = 0.21$; $+$: $p_1 = 0.43$, $p_2 = p_3 = p_4 = 0.19$.

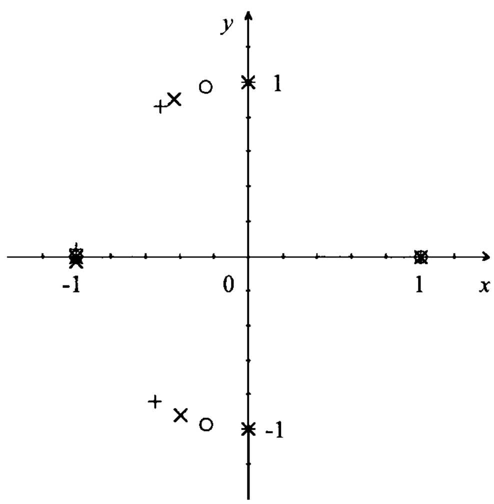

Fig. 4. Influence of *a priori* probabilities on signal sets under the equal energy constraint. $\lambda = 0.5$. $*$: equally likely; o: $p_1 = 0.28$, $p_2 = p_3 = p_4 = 0.24$; $\times$: $p_1 = 0.31$, $p_2 = p_3 = p_4 = 0.23$; $+$: $p_1 = 0.34$, $p_2 = p_3 = p_4 = 0.22$.

in 2-D space [10]. With $p_1 = p_2 = p_3 = 0.22$, $p_4 = 0.18$, and $p_5 = 0.16$, the signal set is that in which $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, and $\boldsymbol{s}_3$ form a 2-D regular simplex. As indicated by Dunbridge, the other signals of smaller *a priori* probabilities are not important and their positions can be arbitrary as the SNR is vanishingly small.

6) Example of Application to Practical Communication System Design: As an example of application of the proposed algorithms to practical communication system design, we consider the scenario of satellite communications where SNR's are small [12]. The phase-shift keying (PSK) modulation schemes are usually considered. One of the schemes is 8-PSK, the set of eight 2-D signals equally spaced on a circle. In the WSC condition, 8-PSK set is proved to be optimal in 2-D space [10], but is not necessarily optimal in the SSC condition. In this simulation, we intend to find better signal sets

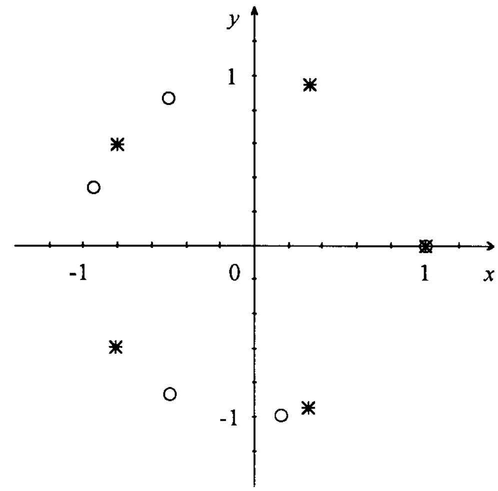

Fig. 5. Signal sets under the equal energy constraint. $\lambda = 0.1$. $*$: equally likely signal set in which signals are equally spaced. o: signal set with $p_1 = p_2 = p_3 = 0.22$, $p_4 = 0.18$, and $p_5 = 0.16$, in which $\boldsymbol{s}_1$, $\boldsymbol{s}_2$, and $\boldsymbol{s}_3$ form a 2-D regular simplex.

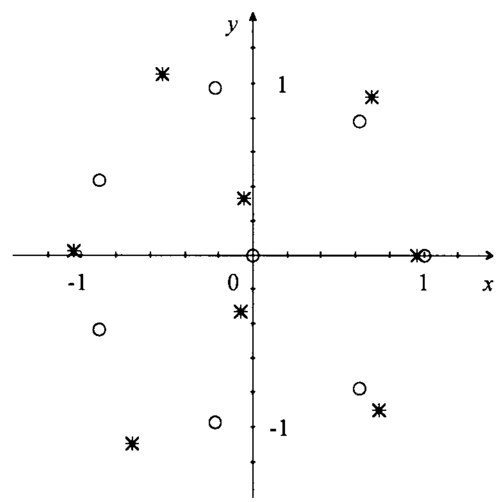

Fig. 6. Two signal sets of eight signals. o: the $E(8,7)$ signal set; $*$: the X signal set. They were found in the SSC condition with $\lambda = 2.51$ by Algorithm 2.

of eight signals in the SSC condition in 2-D space by using the proposed algorithms and comparing their error probabilities to that of the 8-PSK set.

By using Algorithm 2, in the SSC condition with $\lambda = 2.51$, two signal sets are found. One is the $E(8,7)$ signal set. The other is called the X signal set in this paper because of its constellation. The two signal sets are shown in Fig. 6. In the operation, Algorithm 2 sometimes converged to the $E(8,7)$ signal set and sometimes to the X signal set.

Although these two signal sets are found at $\lambda = 2.51$, we compare their error probabilities with that of the 8-PSK set in the range $\lambda \in [0, 8]$ that is common for satellite communications. The results are shown in Fig. 7. The error probability of the 8-PSK set can be computed by using formula in [10]

$$P_e = 1 - 2\int_0^{\infty} G(y)\,dy \int_{y\cot(\pi/M)-\lambda}^{\infty} G(x)\,dx$$

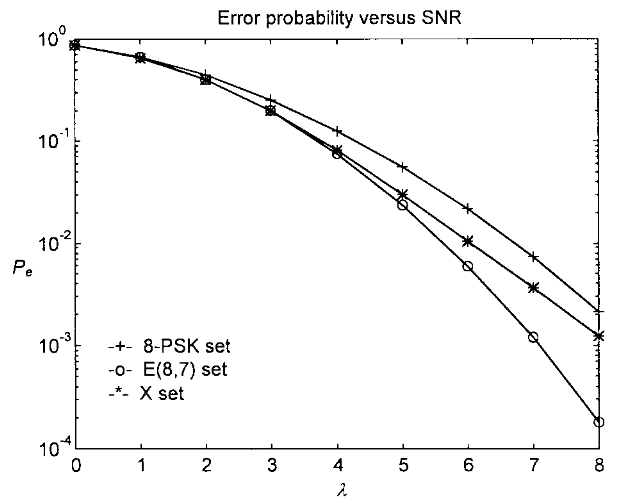

Fig. 7. Comparison of error probabilities of the 8-PSK set, the $E(8,7)$ signal set, and the X signal set.

where

$$G(x) = (1/\sqrt{2\pi})\exp(-x^2/2).$$

We can approximate it by

$$P_e = 1 - \sqrt{\frac{2}{\pi}} \sum_{i=0}^{k-1} \exp(-(i\Delta x)^2/2)\Phi(\lambda - i\Delta x \cot(\pi/M))\Delta x$$

for computation of numerical values where

$$\Phi(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{x} \exp(-t^2/2)\, dt.$$

In Fig. 7, $k = 50\,000$ and $\Delta x = 8/k$ are taken and MATLAB version 4.2c.1 is used to evaluate P_e, in which $\Phi(x)$ is calculated by the function $\mathrm{erf}(x)$ with $\Phi(x) = (1 + \mathrm{erf}(x/\sqrt{2}))/2$. The error probabilities of the $E(8,7)$ signal set and the X signal set are estimated by Monte Carlo approximation.

As we can see in Fig. 7, the $E(8,7)$ signal set has the smallest error probability in the computed range of SNR and the 8-PSK set has the largest error probability. Note that the error probability of a signal set is the average error probability over all eight signals. In many application cases, it is usually required that the largest error probability of a single signal in a signal set be smaller than a value. Since the $E(8,7)$ signal set and the X signal set are unequal-length sets, it is interesting to compare the largest and the smallest error probabilities of single signals among the three signal sets. We find [13] that the largest error probability of a signal in the X signal set is always above the error probability of the 8-PSK set. This implies that a signal in the X signal set has larger error probability than that of any signal in the 8-PSK set, though the average error probability of the X signal set is lower than that of the 8-PSK set. However, for a large range of $\lambda \in [3, 8]$, the error probabilities of all signals in the $E(8,7)$ signal set are all lower than those of the 8-PSK set. In summary, we can conclude that the $E(8,7)$ signal set is the best signal set among the three sets. As we reach this conclusion, we do not take account of possible complexity of practical implementation of the zero energy signal in the $E(8,7)$ signal set.

III. On the Optimality of the L2 Signal Set

A. Mean Width of Polytope and Detection Probability of Signal Set at Low SNR's

The polytope of a set of Mn-dimensional vectors $\{\boldsymbol{v}_i\}$ is the convex hull

$$C = \left\{ \boldsymbol{y} \,|\, \boldsymbol{y} = \sum_{i=1}^{M} \alpha_i \boldsymbol{v}_i,\ \sum_{i=1}^{M} \alpha_i = 1,\ \alpha_i \geq 0,\ i = 1, \cdots, M \right\}. \tag{20}$$

The mean width of the polytope is defined as

$$B = \frac{1}{w_n} \int_{\Omega_n} \max_{1 \leq i \leq M} \boldsymbol{y}_n^T \boldsymbol{v}_i \, d\boldsymbol{w} \tag{21}$$

where Ω_n is the surface of an n-dimensional unit sphere, $\boldsymbol{y}_n$ is a unit vector on Ω_n, $d\boldsymbol{w}$ is a surface element on Ω_n, and $w_n = 2\pi^{n/2}/\Gamma(n/2)$ is the surface area of the n-dimensional unit sphere. Γ is the Gamma function.

Lemma 1: Consider a set of Mn-dimensional vectors $\{\boldsymbol{v}_i\}$ that generate a polytope C. Among the M vectors, $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K$ are on the boundary of C and $\boldsymbol{v}_{K+1}, \cdots, \boldsymbol{v}_M$ are in the interior of C. Then the mean width of C depends only on $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K$ and is given by

$$B(\{\boldsymbol{v}_i\}) = \frac{1}{w_n} \int_{\Omega_n} \max_{1 \leq i \leq K} \left(\boldsymbol{y}_n^T \boldsymbol{v}_i\right) d\boldsymbol{w}. \tag{22}$$

□

Consider a signal set $\{\beta \boldsymbol{v}_i\}$ for $\beta > 0$ with a set of *a priori* probabilities $\{p_i\}$. By noticing that for any function $f(\boldsymbol{y})$

$$\int_{R^n} f(\boldsymbol{y})\, d\boldsymbol{y} = \int_0^{\infty} r^{n-1}\, dr \int_{\Omega_n} f(\boldsymbol{y})\, d\boldsymbol{w}$$

where $r^2 = ||\boldsymbol{y}||^2$, from (6) we can rewrite the detection probability of the signal set as

$$\begin{aligned} &P_d(\beta, \{\boldsymbol{v}_i\}, \{p_i\}) \\ &\quad = \frac{1}{(2\pi)^{n/2}} \int_0^{\infty} \exp\left(-\frac{r^2}{2}\right) r^{n-1}\, dr \\ &\quad\quad \times \int_{\Omega_n} \max_{1 \leq i \leq M} \left[p_i \exp\left(r\beta \boldsymbol{y}_n^T \boldsymbol{v}_i - \frac{\beta^2}{2} ||\boldsymbol{v}_i||^2 \right) \right] d\boldsymbol{w}. \end{aligned} \tag{23}$$

Theorem 1: For the n-dimensional signal set $\{\beta \boldsymbol{v}_i\}$ with a set of *a priori* probabilities $\{p_i\}$, the derivative of the detection probability with respect to β as $\beta \to 0$ is

$$\begin{aligned} P_d'(0, \{\boldsymbol{v}_i\}, \{p_i\}) &\equiv \left. \frac{\partial P_d(\beta, \{\boldsymbol{v}_i\}, \{p_i\})}{\partial \beta} \right|_{\beta=0} \\ &= \frac{\sqrt{2}\Gamma((n+1)/2)}{\Gamma(n/2)} B(\{p_i \boldsymbol{v}_i\}) \end{aligned} \tag{24}$$

where $B(\{p_i \boldsymbol{v}_i\})$ is the mean width of the polytope generated by $\{p_i \boldsymbol{v}_i\}$. If the signal set is equally likely

$$P_d'(0, \{\boldsymbol{v}_i\}) = \frac{\sqrt{2}\Gamma((n+1)/2)}{M\Gamma(n/2)} B(\{\boldsymbol{v}_i\}). \tag{25}$$

□

Corollary 1: For the signal set $\{\beta \boldsymbol{v}_i\}$ with *a priori* probability $\{p_i\}$, a necessary condition for optimality as $\beta \to 0$ is that the mean width of the polytope generated by $\{p_i \boldsymbol{v}_i\}$ be maximized. If the signal set is equally likely, it is necessary that the mean width of the polytope generated by $\{\boldsymbol{v}_i\}$ be maximized. □

Corollary 1 is applicable to average and equal energy constraints and applicable to any set of *a priori* probabilities. This extends the results in [3] and [4] for equally likely signal sets.

If

$$P'_d(0, \{\boldsymbol{v}_i\}, \{p_i\}) > P'_d(0, \{\boldsymbol{v}'_i\}, \{p'_i\})$$

following the theorem of real continuously differentiable function, there exists a neighborhood of $\beta \in [0, \delta)$ such that

$$P_d(\beta, \{\boldsymbol{v}_i\}, \{p_i\}) > P_d(\beta, \{\boldsymbol{v}'_i\}, \{p'_i\}).$$

In the paper, the following statements are always equivalent without further referring to each other: 1) $\{\boldsymbol{v}_i\}$ is better than $\{\boldsymbol{v}'_i\}$ at low SNR's, 2) there exists a neighborhood of $\beta \in [0, \delta)$ such that

$$P_d(\beta, \{\boldsymbol{v}_i\}, \{p_i\}) > P_d(\beta, \{\boldsymbol{v}'_i\}, \{p'_i\})$$

and 3)

$$P'_d(0, \{\boldsymbol{v}_i\}, \{p_i\}) > P'_d(0, \{\boldsymbol{v}'_i\}, \{p'_i\}).$$

Since $P'_d(0, \{\boldsymbol{v}_i\}, \{p_i\})$ is proportional to the mean width of the polytope generated by $\{\boldsymbol{v}_i p_i\}$, the properties of mean width can be used in the comparison.

Due to Theorem 1, if $p_i \boldsymbol{v}_i = p'_i \boldsymbol{v}'_i$ for $i = 1, \cdots, M$, then

$$P'_d(0, \{\boldsymbol{v}_i\}, \{p_i\}) = P'_d(0, \{\boldsymbol{v}'_i\}, \{p'_i\}).$$

Theorem 1 also implies

$$P'_d(0, \{r\boldsymbol{v}_i\}, \{p_i\}) = rP'_d(0, \{\boldsymbol{v}_i\}, \{p_i\}).$$

This means that the detection probability at low SNR's is proportional to squared root of the SNR. However, this is not true when the SNR is higher.

Under total energy constraint, one necessary condition for optimality is (13). In terms of (20), the origin is necessarily in the interior of the polytope generated by $\{p_i \boldsymbol{s}_i\}$. Hence, we have the following lemma.

Lemma 2: One of the necessary conditions for the optimal set of signal vectors $\{\boldsymbol{s}_i\}$ with *a priori* probabilities $\{p_i\}$ is that the origin is in the interior of the polytope generated by $\{p_i \boldsymbol{s}_i\}$. □

Lemma 3: The optimal set of signal vectors $\{\boldsymbol{s}_i\}$ with *a priori* probabilities $\{p_i\}$ at low SNR's has signals such that $p_i \boldsymbol{s}_i$ is either on the boundary of the polytope generated by $\{p_i \boldsymbol{s}_i\}$ or at the origin. For equally likely signal sets, the statement is true by replacing $p_i \boldsymbol{s}_i$ with $\boldsymbol{s}_i$.

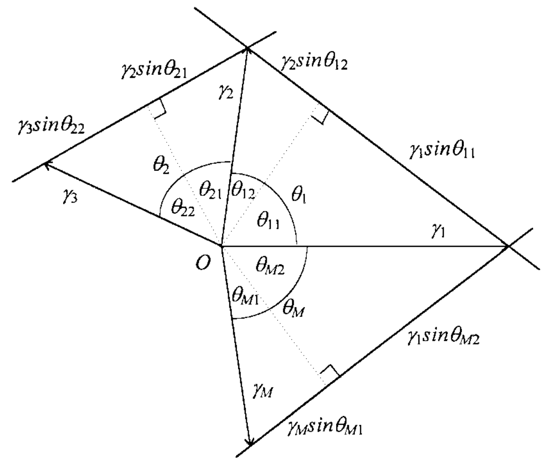

Fig. 8. Calculation of mean width of a polytope in 2-D space.

Corollary 2: The mean width of the polytope generated by the regular simplex of M unit vectors $\{\boldsymbol{v}_i\}$ is given by

$$B_S = \frac{M\sqrt{M(M-1)}\Gamma((M-1)/2)}{2\sqrt{2\pi}\Gamma(M/2)} E\left[\Phi^{M-2}\left(\frac{x}{\sqrt{2}}\right)\right]. \quad (26)$$

where x is the standard Gaussian random variable. □

We can prove the following theorem [13].

Theorem 2: Given a signal set $\{\beta \boldsymbol{v}_i\}$ and a set of *a priori* probabilities $\{p_i\}$. Let $P_d(\beta, V_j)$ denote the detection probability of the jth signal. If $p_j \boldsymbol{v}_j$ is in the interior of the polytope generated by $\{p_i \boldsymbol{v}_i\}$, then

$$P'_d(0, V_j) \equiv \left.\frac{\partial P_d(\beta, V_j)}{\partial \beta}\right|_{\beta=0} = 0. \quad (27)$$

B. The Class of Two-Dimensional Signal Sets

In this and the following sections, we assume that all discussed signal sets satisfy the SSC condition. In two-dimensional space, a polytope is a polygon (see Fig. 8). The mean width of a polygon is

$$B = \frac{1}{2\pi} \int_0^{2\pi} L(\theta)\, d\theta \quad (28)$$

where $L(\theta)$ is the radial distance at angle θ from the origin to the point where lines drawn perpendicular to the radial line first intersect the perimeter.

Theorem 3: Consider a set of M signal vectors $\{\boldsymbol{s}_i\}$ which generate a polygon C such that the M vectors are on the boundary of C and the origin is in the interior of C. Under the average energy constraint (7), the mean width of C is maximized if and only if the signals are equally spaced on the circle of radius λ. □

Lemma 4: The mean width of the polygon C generated by the signal set $E(M, K)$ is

$$B(E(M,K)) = \frac{\sqrt{MK}\lambda}{\pi} \sin\left(\frac{\pi}{K}\right), \qquad M \geq 2,\ 2 \leq K \leq M \quad (29)$$

and 1) for $M = 3$, $B(E(M,3)) > B(E(M,2))$; 2) for $M = 4$, $B(E(M,3)) > B(E(M,2)) = B(E(M,4))$; 3) for $M > 4$, $B(E(M,3)) > B(E(M,2)) = B(E(M,4)) > B(E(M,5)) > \cdots > B(E(M,M))$. □

Let the detection probability of the signal set $E(M,K)$ be denoted by $P_d(\lambda, E(M,K))$.

Theorem 4: For $M \geq 2$ and $2 \leq K \leq M$

$$P_d'(0, E(M,K)) \equiv \left.\frac{\partial P_d(\lambda, E(M,K))}{\partial \lambda}\right|_{\lambda=0} = \frac{1}{\sqrt{\pi M}} g(K), \tag{30}$$

where $g(K) = \sqrt{K/2}\sin\,(\pi/K)$, and 1) for $M = 3$, $P_d'\ (0, E(M,3)) > P_d'(0, E(M,2))$; 2) for $M = 4$, $P_d'(0, E(M,3)) > P_d'(0, E(M,2)) = P_d'(0, E(M,4))$; 3) for $M > 4$

$$\begin{aligned} P_d'(0,E(M,3)) &> P_d'(0,E(M,2)) \\ &= P_d'(0,E(M,4)) > P_d'(0,E(M,5)) > \\ &\cdots > P_d'(0,E(M,M)). \end{aligned}$$ □

To prove Theorem 4, in terms of (25) and (29) and by noticing $\Gamma(3/2) = \sqrt{\pi}/2$ and $\Gamma(1) = 1$, we obtain (30). By means of Lemma 4, the order of $P_d'(0, E(M,K))$'s is obtained.

Corollary 3: For $M \geq 4$, the L2 signal set is the unique optimal set in the class of signal sets $E(M,K)$.

Corollary 3 immediately follows Theorem 4. Theorem 4 also means that the L2 signal set is better than the L1 signal set in the SSC condition at low SNR's. Let $\sqrt{M/K}E(M,K)$ denote

$$\begin{aligned} &\sqrt{M/K}\{\lambda(M/K)\boldsymbol{v}_1, \cdots, \lambda\sqrt{M/K}\boldsymbol{v}_K, \mathbf{0}, \cdots, \mathbf{0}\} \\ &\quad = \{\lambda(M/K)\boldsymbol{v}_1, \cdots, \lambda(M/K)\boldsymbol{v}_1, \mathbf{0}, \cdots, \mathbf{0}\}. \end{aligned}$$

The following corollary can be proved [13].

Corollary 4: 1) For $M \geq 3$

$$P_d'(0, \sqrt{M/K}E(M,K)) = P_d'(0, E(K,K)).$$

2) If $M > K$

$$P_d'(0, \sqrt{M/K}E(M,M)) < P_d'(0, E(K,K)).$$ □

The function $g(K)$ for $K = 2, \cdots, 50$ is depicted in Fig. 9. Theorem 4 and Corollaries 3 and 4 can be interpreted from the figure.

Corollary 5: Signal set $E(M,4)$ disproves the strong simplex conjecture for all $M \geq 7$. □

In fact, except for different dimension, the signal set $E(M,4)$ and the L1 signal set have the same mean width of polytope. They both are unequal energy signal sets. So, the results in [1] created by the L1 signal set because of these characteristics can also be created by the signal set $E(M,4)$. Corollary 5 gives only one example.

Theorem 5: As $\lambda \to 0$, the L2 signal set is the unique optimal signal set in the class of two-dimensional signal sets in the SSC condition. □

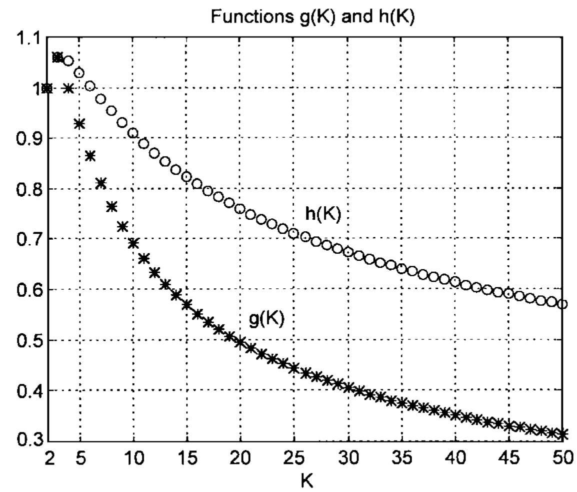

Fig. 9. Functions $g(K)$ and $h(K)$ for $K = 2, \cdots, 50$.

C. The Class of Signal Sets $S(M,K)$

For the class of signal sets $S(M,K)$ for $M \geq 2$ and $2 \leq K \leq M$, some of them were already analyzed. The regular simplex of two antipodal signals is proved to be optimal for $M = 2$ at all λ in both the SSC and WSC conditions. Dunbridge proved [4] that in the SSC condition $S(3,3)$ is optimal as λ goes to zero. Steiner proved [1] that in the SSC condition the L1 signal set is better than the regular simplex set $S(M,M)$ for all $M \geq 7$ as λ goes to zero. In the preceding section, we have proved that the L2 signal set is a unique optimal set in the class of two-dimensional signal sets in the SSC condition at low SNR's. However, the general comparison in this class of signal sets is unknown. It is interesting to find the optimal set in this class of signal sets. In this section, we will give the general comparison of the detection probability in the class of signal sets $S(M,K)$ for all $M \geq 4$ and $2 \leq K \leq M$, prove that the L2 signal set is a unique optimal signal set in the class of signal sets $S(M,K)$ for $M \geq 4$ and $2 \leq K \leq M$, and generate several other results.

The normalized inner product matrix of the signal set $S(M,K)$ is

$$\boldsymbol{R}_S = \begin{bmatrix} \frac{M}{K} & -\frac{M}{(K-1)K} & \cdots & -\frac{M}{(K-1)K} & 0 & \cdots & 0 \\ -\frac{M}{(K-1)K} & \frac{M}{K} & \cdots & -\frac{M}{(K-1)K} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -\frac{M}{(K-1)K} & -\frac{M}{(K-1)K} & \cdots & \frac{M}{K} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{31}$$

When $K = M$, it becomes the normalized inner product matrix of the regular simplex set of M signals. Let the detection probability of the signal set $S(M,K)$ be denoted by $P_d(\lambda, S(M,K))$.

Lemma 5: For $M \geq 2$ and $2 \leq K \leq M$, the derivative of the detection probability of the signal set $S(M,K)$ with respect to λ as $\lambda \to 0$ is

$$P'_d(0, S(M,K)) \equiv \left.\frac{\partial P_d(\lambda, S(M,K))}{\partial \lambda}\right|_{\lambda=0} = \sqrt{\frac{K}{M}} P'_d(0, S(K,K)) \quad (32)$$

□

Theorem 6: For the signal set $S(M,K)$ for $M \geq 2$ and $2 \leq K \leq M$

$$P'_d(0, S(M,K)) = \frac{1}{\sqrt{\pi M}} h(K) \quad (33)$$

where

$$h(K) = \frac{K\sqrt{K-1}}{2} E[\Phi^{K-2}(x)].$$ □

By means of Lemma 5 and (A1), we obtain the proof of Theorem 6.

The function $h(K)$ for $K = 2, \cdots, 50$ is depicted in Fig. 9. In the evaluation of $h(K)$, MATLAB version 4.2c.1 is used. The function $\Phi(x)$ is evaluated by the embedded subroutine $\mathrm{erf}(x)$ with $\Phi(x) = (1 + \mathrm{erf}(x/\sqrt{2}))/2$. Each interval in the summation that approximates the integral in $h(K)$ is sufficiently small and the total integration interval is large enough so that the total absolute error produced by the numerical evaluation of $h(K)$ is smaller than 10^{-4}. Fig. 9 shows that $h(K)$ is monotonously decreasing for $K \geq 3$. Fig. 9 suggests Propositions 1 and 2 and Corollaries 7 [13] and 8 as follows.

Proposition 1:

$$h(3) > h(4) > h(5) > h(6) > h(2) > h(7)$$

and $h(K)$ is monotonously decreasing for all $K \geq 7$. □

Corollary 6: For $K \geq 3$, if $M > K$

$$P'_d(0, \sqrt{M/K} S(M,M)) < P'_d(0, S(K,K)).$$ □

Based on Proposition 1, now we present one of the main results in this section, which gives the general comparison in the class of signal sets $S(M,K)$ for $M \geq 4$ and $2 \leq K \leq M$.

Proposition 2: 1) For $M = 4, 5, 6$, and $3 \leq K \leq M - 1$

$$P'_d(0, S(M,K)) > P'_d(0, S(M, K+1)) > P'_d(0, S(M,2)).$$

2) For $M = 7$

$$\begin{aligned} P'_d(0, S(M,3)) &> P'_d(0, S(M,4)) \\ &> P'_d(0, S(M,5)) > P'_d(0, S(M,6)) \\ &> P'_d(0, S(M,2)) > P'_d(0, S(M,7)). \end{aligned}$$

3) For $M > 7$

$$\begin{aligned} P'_d(0, S(M,3)) &> P'_d(0, S(M,4)) \\ &> P'_d(0, S(M,5)) > P'_d(0, S(M,6)) \\ &> P'_d(0, S(M,2)) > P'_d(0, S(M,7)) \\ &> \cdots > P'_d(0, S(M, M-1)) \\ &> P'_d(0, S(M,M)). \end{aligned}$$ □

Corollary 7: For $M \geq 4$, all the signal sets $S(M,K)$ for $3 \leq K \leq M - 1$ disprove the strong simplex conjecture, and $S(M,2)$, if $M \geq 7$, also disproves the strong simplex conjecture. □

In what follows, it is strictly shown that the L2 signal set is the unique optimal signal set in the class of signal sets $S(M,K)$ for $M \geq 4$ and $2 \leq K \leq M$ and that the strong simplex conjecture is false for $4 \leq M \leq 6$.

For $K = 2$ and 3, we can obtain the closed form of evaluation of $h(K)$ by noticing (A1): $h(2) = 1$ and $h(3) = 3\sqrt{2}/4$. This results in the following closed form of evaluation of $P'_d(0, S(M,K))$: 1) for the regular simplex set of two signals, $P'_d(0, S(2,2)) = 1/\sqrt{2\pi}$; 2) for the simplex set of three signals, $P'_d(0, S(3,3)) = \sqrt{6}/(4\sqrt{\pi})$; 3) for the L1 signal set, $P'_d(0, S(M,2)) = 1/\sqrt{\pi M}$ for $M \geq 3$. This result is also given by Steiner [1]; 4) for the L2 signal set, $P'_d(0, S(M,3)) = 3\sqrt{2}/(4/\sqrt{\pi M})$. Now, we present another main result in this section.

Theorem 7: For $M \geq 4$, there exist neighborhoods $\lambda \in [0, \delta_{M,K})$, $\delta_{M,K} > 0$, of λ, such that $P_d(\lambda, S(M,3))$ is strictly greater than $P_d(\lambda, S(M,2))$ and grater than $P_d(\lambda, S(M,K))$ for all $4 \leq K \leq M$. □

Corollary 8: For $M \geq 4$, the L2 signal set is the unique optimal signal set in the class of signal sets $S(M,K)$ for all $2 \leq K \leq M$. □

Corollary 8 immediately follows Theorem 7. Note that when $K = M$, $S(M,K)$ is the regular simplex set. Hence, now we have the following result.

Proposition 3: The strong simplex conjecture is false for $4 \leq M \leq 6$. □

The strong simplex conjecture is disproved by the L1 signal set for all $M \geq 7$ as shown in [1]. Corollary 8 means that the L2 signal set disproves the strong simplex conjecture for all $M \geq 4$. Proposition 3 updates the strong simplex conjecture to such a situation: the strong simplex conjecture is true for $M = 2$, true for $M = 3$ at low SNR's [4], and false for all $M \geq 4$.

By using the same method as in [1] for $M \geq 7$, Corollaries 9 and 10, that follow, can be proved [13], extending Steiner's results created by the L1 signal set for all $M \geq 7$ to all $M \geq 4$.

Corollary 9: For any fixed $M = 4, 5$, and 6, there is no signal set which is optimal at all SNR's. □

Corollary 10: For $M = 4, 5$, and 6 under the average energy constraint, the optimal solution as $\lambda \to 0$ is not an equal energy solution. □

According to Fig. 9, $h(K) \geq g(K)$ where the equality holds if and only if $K = 2$ and 3. In terms of (30) and (33), we obtain the following result.

Proposition 4: For $M \geq 2$ and $2 \leq K \leq M$

$$P'_d(0, S(M,K)) \geq P'_d(0, E(M,K))$$

where the equality holds if and only if $K = 2$ and 3. □

Finally, we present the following results [13].

Proposition 5: For $M \geq 7$, there exists an integer $K' < M$ such that

$$P'_d(0, E(M,K)) > P'_d(0, S(M,M))$$

for all $4 \leq K \leq K'$. □

Corollary 11: For $M \geq 7$, there exists an integer $K' < M$ such that all the signal sets $E(M, K)$ for $4 \leq K \leq K'$ disprove the strong simplex conjecture. □

Corollaries 7 and 11 show that there are many signal sets which can disprove the strong simplex conjecture for $M \geq 4$.

IV. Conclusion

In this paper, we propose four practically useful stochastic iterative algorithms for signal set design. The algorithms are suitable for operation in various conditions on the dimension of signal space, number of signals, *a priori* probabilities, SNR, and energy constraint. Simulation results reveal the optimality of the L2 signal set which results in the discovery of the L2 signal set, and verify the optimality of the L1 signal set, the weak simplex conjecture, and two Dunbridge's theorems. The influence of *a priori* probabilities and SNR's on signal sets is also observed by simulation. An example of application of the proposed algorithms in practical communication system design is demonstrated. All these simulation results show that the proposed algorithms are promising as a constructive approach to signal set design.

Based on the analysis of the mean width of the polytope generated by a signal set, the class of two-dimensional signal sets and the class of signal sets $S(M, K)$ are analyzed in the SSC condition at low SNR's. We give the order of the detection probabilities of the class of signal sets $E(M, K)$ and therefore prove that the L2 signal set is the unique optimal signal set in the class of two-dimensional signal sets for all $M \geq 4$. We also present the order of detection probabilities of the signal sets $S(M, K)$ for $M \geq 4$ and $2 \leq K \leq M$. This leads to the conclusion that all signal sets $S(M, K)$ for $M \geq 4$ and $3 \leq K \leq M - 1$ disprove the strong simplex conjecture, and $S(M, 2)$ (the L1 signal set), if $M \geq 7$, also disproves the strong simplex conjecture. We also prove that the L2 signal set is a unique optimal signal set in the class of signal sets $S(M, K)$ for all $M \geq 4$ and $2 \leq K \leq M$. These results also lead to extension of several Steiner's results for all $M \geq 7$ to all $M \geq 4$. The disproof of the strong simplex conjecture for $4 \leq M \leq 6$ updates the strong simplex conjecture in the situation that the strong simplex conjecture is true for $M = 2$, true for $M = 3$ at low SNR's, and false for all $M \geq 4$. We find that for $M \geq 7$, there are many signal sets $E(M, K)$ for some integer $K \leq M - 1$ that also disprove the strong simplex conjecture.

Simulation results show that if the number of samples used in one step is large, the batch algorithms may converge to local maxima. Nevertheless, the sequential algorithms always converge to global maxima. The proposed algorithms are sensitive at low SNR's. The convergent performance of the algorithms at high SNR's needs to be improved. Theoretical analysis of dynamic behavior of the proposed algorithms needs to be done in a more general condition under the two energy constraints. The continuous-time version of the stochastic dynamic system driven by white Gaussian process can be obtained in the same way as we derive the discrete-time version of the algorithms. We hope that by implementing the continuous-time version of the stochastic dynamic system with hardware, better signal sets can be more efficiently and more accurately found.

Since simulation results show that in the SSC condition at low SNR's the L2 signal set is the only signal set to which the proposed algorithms converge, and the L2 signal set is proved to be a unique optimal signal set in a large class of signal sets, we conjecture that the L2 signal set is the unique and globally optimal signal set without dimensional limitation at low SNR's. In order to prove this, based on what is proved in this paper, it is sufficient to prove that in the class of signal sets in which all signals of a set are located at the boundary of the polytope generated by the signal set, an unequal energy set is not better than an equal energy set (this is proved true in 2-D case in the paper).

Optimal design under a noncoherent assumption remains unsolved. Optimal signal set design for other kinds of noise channels are more difficult to analyze than Gaussian channels. Properly establishing some stochastic iterative algorithms for noncoherent channels and other type of channels is an interesting problem.

Appendix

Proof of Lemma 1

For any $j \in \{K+1, \cdots, M\}$, since $\boldsymbol{v}_j$ is in the interior of C, there are $\alpha_i > 0$ for $i = 1, \cdots, M$ and

$$\sum_{i=1}^{M} \alpha_i = 1$$

such that

$$\boldsymbol{v}_j = \sum_{i=1}^{M} \alpha_i \boldsymbol{v}_i.$$

For any unit vector $\boldsymbol{y}_n$ on the unit sphere

$$\boldsymbol{y}_n^T \boldsymbol{v}_j = (1/(1-\alpha_j)) \sum_{i=1, i\neq j}^{M} \alpha_i \boldsymbol{y}_n^T \boldsymbol{v}_i.$$

Assume $\boldsymbol{y}_n^T \boldsymbol{v}_l = \max_{1 \leq i \leq M, i \neq j} \boldsymbol{y}_n^T \boldsymbol{v}_i$.

$$\boldsymbol{y}_n^T \boldsymbol{v}_j \leq (1/(1-\alpha_j)) \sum_{i=1, i\neq j}^{M} \alpha_i \boldsymbol{y}_n^T \boldsymbol{v}_l = \boldsymbol{y}_n^T \boldsymbol{v}_l$$

where the equality holds if and only if

$$\boldsymbol{y}_n^T \boldsymbol{v}_1 = \boldsymbol{y}_n^T \boldsymbol{v}_2 = \cdots = \boldsymbol{y}_n^T \boldsymbol{v}_M.$$

It does not change the mean width of the polytope to exclude the trivial case. Therefore, we have

$$\boldsymbol{y}_n^T \boldsymbol{v}_j < \max_{1 \leq i \leq M, i \neq j} \boldsymbol{y}_n^T \boldsymbol{v}_i.$$

Hence, (22) is true. Q.E.D.

Proof of Theorem 1

Take the derivative on (23) with respect to β and take the limit $\beta \to 0$

$$\begin{aligned} &P_d'(0, \{\boldsymbol{v}_i\}, \{p_i\}) \\ &= \lim_{\beta\to 0} \left\{ \frac{1}{(2\pi)^{n/2}} \int_0^\infty \exp\left(-\frac{r^2}{2}\right) r^{n-1}\, dr \right. \\ &\quad \times \int_{\Omega_n} \max_{1\le i\le M} \left[p_i \exp\left(r\beta \boldsymbol{y}_n^T \boldsymbol{v}_i - \frac{\beta^2}{2}\|\boldsymbol{v}_i\|^2 \right) \right. \\ &\quad \left.\left. \times \left(r\boldsymbol{y}_n^T \boldsymbol{v}_i - \beta\|\boldsymbol{v}_i\|^2 \right) \right] d\boldsymbol{w} \right\} \\ &= \frac{1}{(2\pi)^{n/2}} \int_0^\infty \exp\left(-\frac{r^2}{2}\right) r^n\, dr \\ &\quad \times \int_{\Omega_n} \max_{1\le i\le M} \left(\boldsymbol{y}_n^T \boldsymbol{v}_i p_i \right) d\boldsymbol{w} \end{aligned}$$

which yields (24) in terms of (21) and the definition of the Gamma function [11]. Let $p_i = 1/M$, we obtain (25). Q.E.D.

Proof of Corollary 1

Since for any signal set of M signals the detection probability $P_d(\beta, \{\boldsymbol{v}_i\}, \{p_i\})$ as $\beta \to 0$ is equal to $P_d(0, \{\boldsymbol{v}_i\}, \{p_i\}) = 1/M$ and $P_d(\beta, \{\boldsymbol{v}_i\}, \{p_i\})$ is a continuously differentiable increasing function of $\beta \in [0, \infty)$, by means of the theorem of real continuously differentiable functions, the optimality requires that $P_d'(0, \{\boldsymbol{v}_i\}, \{p_i\})$ be maximized, which according to Theorem 1 requires that the mean width of the polytope generated by $\{p_i \boldsymbol{v}_i\}$, or by $\{\boldsymbol{v}_i\}$ if it is equally likely, be maximized. Q.E.D.

Proof of Lemma 3

Assume that signal set Ξ has $p_1\boldsymbol{s}_1, \cdots, p_K\boldsymbol{s}_K$ on the boundary of the polytope C_Ξ generated by $p_1\boldsymbol{s}_1, \cdots, p_K\boldsymbol{s}_K$, and $p_{K+1}\boldsymbol{s}_{K+1}, \cdots, p_M\boldsymbol{s}_M$ in the interior of C_Ξ but not at the origin. Lemma 2 shows that the origin is necessarily in the interior of C_Ξ. Due to Lemma 1, C_Ξ depends only on the signal vectors $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K$ and its mean width is

$$B_\Xi = (1/w_n) \int_{\Omega_n} \max_{1\le i\le K} \boldsymbol{y}_n^T \boldsymbol{s}_i p_i \, d\boldsymbol{w}.$$

Now, we produce a new signal set Ψ by taking $\boldsymbol{s}_i' = \gamma \boldsymbol{s}_i$ for $i = 1, \cdots, K$ and $\boldsymbol{s}_i' = \boldsymbol{0}$ for $i = K+1, \cdots, M$, where

$$\gamma = \sqrt{M\lambda^2 / \sum_{i=1}^K \|\boldsymbol{s}_i\|^2} > 1.$$

The two sets have the same total energy. However, the mean width of the polytope C_Ψ generated by $p_1\boldsymbol{s}_1', \cdots, p_M\boldsymbol{s}_M'$ is

$$B_\Psi = (1/w_n) \int_{\Omega_n} \max_{1\le i\le K} \boldsymbol{y}_n^T \boldsymbol{s}_i' p_i \, d\boldsymbol{w} = \gamma B_\Xi > B_\Xi.$$

Hence, due to Corollary 1, an optimal solution must have the signals such that $p_i\boldsymbol{s}_i$ is either on the boundary of the polytope generated by $\{p_i\boldsymbol{s}_i\}$ or at the origin. Q.E.D.

Proof of Corollary 2

Consider the regular simplex formed by $\{\lambda \boldsymbol{v}_i\}$ where $\boldsymbol{v}_i$'s are unit vectors. Its detection probability is [10, p. 162]

$$P_d(\lambda, S(M,M)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left[-\frac{\left(x - \lambda\sqrt{\frac{M}{M-1}}\right)^2}{2} \right] \times \Phi^{M-1}(x)\, dx, \qquad M \ge 2.$$

Taking a derivative with respect to λ and letting $\lambda \to 0$, we obtain

$$\begin{aligned} &P_d'(0, S(M,M)) \\ &= \sqrt{\frac{M}{2\pi(M-1)}} \int_{-\infty}^\infty x \exp\left(-\frac{x^2}{2}\right) \Phi^{M-1}(x)\, dx \\ &= \sqrt{\frac{M(M-1)}{2\pi}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp(-x^2) \Phi^{M-2}(x)\, dx \\ &= \frac{\sqrt{M(M-1)}}{2\sqrt{\pi}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{x^2}{2}\right) \Phi^{M-2}\left(\frac{x}{\sqrt{2}}\right) dx \\ &= \frac{\sqrt{M(M-1)}}{2\sqrt{\pi}} E\left[\Phi^{M-2}\left(\frac{x}{\sqrt{2}}\right) \right] \\ &= \frac{\sqrt{M(M-1)}}{2\sqrt{\pi}} \int_{-\infty}^\infty \int_{-\infty}^{\frac{x_1}{\sqrt{2}}} \cdots \int_{-\infty}^{\frac{x_1}{\sqrt{2}}} \frac{1}{(2\pi)^{\frac{M-1}{2}}} \\ &\quad \times \exp\left(-\frac{1}{2} \sum_{i=1}^{M-1} x_i^2 \right) dx_{M-1} \cdots dx_2\, dx_1. \end{aligned} \tag{A1}$$

By means of Theorem 1 and noticing $M = n + 1$, we obtain (26). Q.E.D.

Proof of Theorem 3

As shown in Fig. 8, assume that the signal vector $\boldsymbol{s}_i$ has magnitude γ_i, and the phase angle difference between two adjacent signal vectors is θ_i, $\theta_i = \theta_{i1} + \theta_{i2}$. The mean width of C is then

$$\begin{aligned} B &= \frac{1}{2\pi} \left(\int_{-\theta_{M2}}^{\theta_{11}} \gamma_1 \cos\theta\, d\theta + \int_{-\theta_{12}}^{\theta_{21}} \gamma_2 \cos\theta\, d\theta \right. \\ &\qquad \left. + \cdots + \int_{-\theta_{(M-1)2}}^{\theta_{M1}} \gamma_M \cos\theta\, d\theta \right) \\ &= \frac{1}{2\pi} [\gamma_1(\sin\theta_{M2} + \sin\theta_{11}) + \gamma_2(\sin\theta_{12} + \sin\theta_{21}) \\ &\qquad + \cdots + \gamma_M(\sin\theta_{(M-1)2} + \sin\theta_{M1})] \\ &= \frac{1}{2\pi} (\gamma_1\sin\theta_{11} + \gamma_2\sin\theta_{12} + \gamma_2\sin\theta_{21} + \gamma_3\sin\theta_{22} \\ &\qquad + \cdots + \gamma_M\sin\theta_{M1} + \gamma_1\sin\theta_{M2}) \\ &= \frac{P}{2\pi} \end{aligned} \tag{A2}$$

where

$$\begin{aligned} P &= \gamma_1\sin\theta_{11} + \gamma_2\sin\theta_{12} + \gamma_2\sin\theta_{21} + \gamma_3\sin\theta_{22} \\ &\quad + \cdots + \gamma_M\sin\theta_{M1} + \gamma_1\sin\theta_{M2} \\ &= \sqrt{\gamma_1^2 + \gamma_2^2 - 2\gamma_1\gamma_2\cos\theta_1} \\ &\quad + \sqrt{\gamma_2^2 + \gamma_3^2 - 2\gamma_2\gamma_3\cos\theta_2} \\ &\quad + \cdots + \sqrt{\gamma_M^2 + \gamma_1^2 - 2\gamma_M\gamma_1\cos\theta_M} \end{aligned}$$

$$P \le \sqrt{\left(\gamma_1^2+\gamma_2^2+\gamma_2^2+\gamma_3^2+\cdots+\gamma_M^2+\gamma_1^2\right)\left(\sin^2\theta_{11}+\sin^2\theta_{12}+\sin^2\theta_{21}+\sin^2\theta_{22}+\cdots+\sin^2\theta_{M1}+\sin^2\theta_{M2}\right)}$$
$$= \sqrt{2M}\lambda\sqrt{\sum_{i=1}^{M}(\sin^2\theta_{i1}+\sin^2\theta_{i2})} \quad \text{(A4)}$$

is obviously the perimeter of C. For any positive real numbers a_i and b_i, $i=1,\cdots,M$, due to Cauchy's inequality [11]

$$\sum_{i=1}^{M} a_i b_i \le \sqrt{\sum_{i=1}^{M} a_i^2 \sum_{i=1}^{M} b_i^2}$$

where the equality holds if and only if

$$\frac{a_1}{b_1}=\frac{a_2}{b_2}=\cdots=\frac{a_M}{b_M}. \quad \text{(A3)}$$

Notice that

$$\sum_{i=1}^{M}\gamma_i^2 = M\lambda^2.$$

From (A2) we have (A4) at the top of this page, where the equality holds if and only if

$$\frac{\gamma_1}{\sin\theta_{11}}=\frac{\gamma_2}{\sin\theta_{12}}=\frac{\gamma_2}{\sin\theta_{21}}=\frac{\gamma_3}{\sin\theta_{22}} =\frac{\gamma_3}{\sin\theta_{31}}=\cdots=\frac{\gamma_M}{\sin\theta_{M1}}=\frac{\gamma_1}{\sin\theta_{M2}}. \quad \text{(A5)}$$

Expression (A5) implies

$$\theta_{12}=\theta_{21},\ \theta_{22}=\theta_{31},\cdots,\theta_{M2}=\theta_{11} \quad \text{(A6)}$$

and yields

$$P \le 2\lambda\sqrt{M\sum_{i=1}^{M}\sin^2\theta_{i1}} \quad \text{(A7)}$$

where the equality holds if and only if

$$\frac{\gamma_1}{\sin\theta_{11}}=\frac{\gamma_2}{\sin\theta_{21}}=\frac{\gamma_3}{\sin\theta_{31}}=\cdots=\frac{\gamma_M}{\sin\theta_{M1}}.$$

Since $\sin^2 x$ is convex down for $\forall x\in(0,\pi)$, it is clear that for $\alpha_i\ge 0$ and $\sum_{i=1}^{M}\alpha_i=1$

$$\sum_{i=1}^{M}\alpha_i\sin^2(x_i)\le\sin^2\left(\sum_{i=1}^{M}\alpha_i x_i\right)$$

where the equality holds if and only if $x_1=x_2=\cdots=x_M$. Take $\alpha_i=1/M$ and $x_i=\theta_{i1}$, and notice $\sum_{i=1}^{M}\theta_{i1}=\pi$. We obtain

$$\sqrt{\sum_{i=1}^{M}\sin^2\theta_{i1}}\le\sqrt{M}\sin(\pi/M)$$

in which the equality holds if and only if $\theta_{i1}=\pi/M$. Therefore,

$$B\le\frac{M\lambda}{\pi}\sin\left(\frac{\pi}{M}\right) \quad \text{(A8)}$$

where the equality holds if and only if $\theta_{i1}=\theta_{i2}=\frac{\pi}{M}$ and $\gamma_i=\lambda$ for $i=1,\cdots,M$. Q.E.D.

Proof of Lemma 4

Since the zero signals of $E(M,K)$ are in the interior of the polygon generated by $E(M,K)$, due to Lemma 1, the mean width of C depends only on the nonzero signal vectors. In terms of Theorem 3, after replacing M in (A8) by K and λ by $\sqrt{M/K}\lambda$ for equality case, the mean width of C is obtained and is given by (29). Now we need to prove that $B(E(M,K))$'s have the order given in the lemma. This is equivalent to showing that the real function $f(x)=\sqrt{x}\sin(\pi/x)$ has the order of values for integers $x\ge 2$ in the form: $f(3)>f(2)=f(4)>f(5)>\cdots$. It is easy to verify $f(3)=3/2>\sqrt{2}=f(2)=f(4)$. If $f'(x)<0$ for $x\ge 3$, the proof is complete. In what follows, we prove this. The derivative of $f(x)$ is

$$f'(x)=\frac{1}{\sqrt{x}}\cos\left(\frac{\pi}{x}\right)\left(\frac{1}{2}\tan\left(\frac{\pi}{x}\right)-\frac{\pi}{x}\right).$$

We have $(1/\sqrt{x})\cos(\pi/x)>0$ for $x\in(2,\infty)$. It is clear that there exists a unique point $y_0\in(0,\pi/2)$ such that $(1/2)\tan(y_0)=y_0$, and $(1/2)\tan(y)<y$ for $\forall y\in(0,y_0)$ and $(1/2)\tan(y)>y$ for $\forall y\in(y_0,\pi/2)$. Hence, there exists a unique point $x_0=\pi/y_0\in(2,\infty)$ such that $f'(x_0)=0$, and $f'(x)>0$ for $\forall x\in(0,x_0)$ and $f'(x)<0$ for $\forall x\in(x_0,\infty)$. Since

$$f'(3)=\frac{3\sqrt{3}-2\pi}{12\sqrt{3}}<0, \qquad x_0\in(2,3)$$

and therefore $f'(x)<0$ for all $x\ge 3$. Q.E.D.

Proof of Corollary 5

According to Theorem 4, the signal set $E(M,4)$ and $E(M,2)$ (the L1 signal set) have the same derivative of detection probability with respect to λ as $\lambda\to 0$, $P_d'(0,E(M,4))=P_d'(0,E(M,2))$. Since Steiner [1] proved that $P_d'(0,E(M,2))>P_d'(0,S(M,M))$ for all $M\ge 7$, where $P_d'(0,S(M,M))$ denotes the derivative of detection probability of the regular simplex as $\lambda\to 0$, thus disproving the strong simplex conjecture, the signal set $E(M,4)$ also disproves the strong simplex conjecture for all $M\ge 7$. Q.E.D.

Proof of Theorem 5

According to Corollary 1, a necessary condition for optimality at low SNR's is that the mean width of the polytope be maximized. Since the probability of correct detection is an increasing function of λ, the solution will lie on the boundary where (7) is satisfied. Meanwhile, in terms of Lemma 3, an optimal solution must have signals either on the boundary of the polytope generated by the signal set or else at the origin which is necessarily in the interior of the polytope due to Lemma 2. If all signal vectors are located on the boundary of the polygon, Theorem 3 shows that the optimal solution is that in which all signals are equally spaced on the circle of radius λ. This signal set is included in the class of signal sets $E(M, K)$ as $K = M$. If the signal set has signal vectors $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K$ on the boundary of the polygon, and $\boldsymbol{s}_{K+1}, \cdots, \boldsymbol{s}_M$ at the origin, we only need to consider the optimal placement of $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K$. By applying Theorem 3 to $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K$, the optimal arrangement must be that in which $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K$ are equally spaced on the circle of radius $\sqrt{M/K}\lambda$. This signal set also belongs to the class of signal sets $E(M, K)$ as $2 \le K \le M - 1$. Hence, the optimal signal set that we are seeking must be in the class of signal sets $E(M, K)$. Corollary 3 shows that the L2 signal set is the unique optimal signal set in the class of signal sets $E(M, K)$. Therefore, the L2 signal set is the unique optimal signal set in the class of two-dimensional signal sets in the SSC condition at low SNR.s. Q.E.D.

Proof of Lemma 5

By paying attention to (25) and Lemma 1, we obtain

$$
\begin{aligned}
&P_d'(0, S(M, K)) \\
&\quad = \frac{\sqrt{2}\Gamma(K/2)}{M\Gamma((K-1)/2)} \\
&\qquad \times B\left(\left\{\sqrt{\frac{M}{K}}\boldsymbol{v}_1, \cdots, \sqrt{\frac{M}{K}}\boldsymbol{v}_K, \boldsymbol{0}, \cdots, \boldsymbol{0}\right\}\right) \\
&\quad = \frac{\sqrt{2}\Gamma(K/2)}{M\Gamma((K-1)/2)} B\left(\left\{\sqrt{\frac{M}{K}}\boldsymbol{v}_1, \cdots, \sqrt{\frac{M}{K}}\boldsymbol{v}_K\right\}\right) \\
&\quad = \sqrt{\frac{K}{M}} \frac{\sqrt{2}\Gamma(K/2)}{M\Gamma((K-1)/2)} B(\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K\}) \\
&\quad = \sqrt{\frac{K}{M}} P_d'(0, S(K, K)).
\end{aligned}
$$

Q.E.D.

Proof of Theorem 7

According to Theorem 6, it is equivalent to show that $h(3) > h(2)$ and $h(3) > h(M)$ for $M \ge 4$. Note that $h(M)$ is equal to $r = P_d'(0, S(M, M))/P_d'(0, S(M, 2))$ which is defined in [1]. Steiner proved [1] that $r < 1$ for $M \ge 7$. Equivalently, it is proved that $h(M) < 1$ for $M \ge 7$. Now, since $h(3) > 1 = h(2)$, we need only to prove that for $M = 4, 5$, and 6, $h(3) > h(M)$. By strict numerical evaluation, we obtain $h(3) = 3\sqrt{2}/4 \cong 1.0607$, $h(4) \cong 1.0534$, $h(5) \cong 1.0306$, and $h(6) \cong 1.0045$ with absolute errors smaller than 10^{-4}. Hence, $h(3)$ is greater than $h(4)$, $h(5)$, and $h(6)$. Q.E.D.

Acknowledgment

The author wishes to thank his advisor Prof. J. Kieffer for presenting the problem, the valuable discussions, and suggestions. The author also wishes to thank reviewers as well as the Associate Editor C. Georghiades for their helpful comments.

References

[1] M. Steiner, "The strong simplex conjecture is false," *IEEE Trans. Inform. Theory*, vol. 40, pp. 721–731, May 1994.

[2] T. M. Cover and B. Gopinath, *Open Problems in Communication and Computation*. New York: Springer-Verlag, 1987.

[3] A. V. Balakrishnan, "A contribution to the sphere-packing problem of communication theory," *J. Math. Anal. Appl.*, vol. 3, pp. 485–506, 1961.

[4] B. Dunbridge, "Asymmetric signal design for the coherent Gaussian channel," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 422–431, July 1967.

[5] H. J. Landau and D. Slepian, "On the optimality of the regular simplex code," *Bell Syst. Tech. J.*, vol. 45, pp. 1247–1272, Oct. 1966.

[6] S. M. Farber, "On the signal selection problem for phase coherent and incoherent communication channels," Ph.D. dissertation, Calif. Inst. Technol., Pasadena, 1968.

[7] L. Ljung, "Analysis of recursive stochastic algorithm," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 551–575, 1977.

[8] H. J. Kushner and D. S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*. New York: Springer-Verlag, 1978.

[9] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 1994.

[10] C. L. Weber, *Elements of Detection and Signal Design*. Berlin, Germany: Springer-Verlag, 1987.

[11] R. G. Bartle, *The Elements of Real Analysis*, 2nd ed. New York: Wiley, 1976.

[12] H. Stewart, "16-QAM modems in satellites," *Commun. Syst. Des.*, pp. 36–40, July 1995.

[13] Y. Sun, "Stochastic iterative algorithms for signal set design for Gaussian channels and the L2 signal set," Ph.D. dissertation, Univ. of Minn., Minneapolis, Jan. 1997.

[14] ——, "Seek the optimum signal sets for Gaussian noise channels by using stochastic iteration," presented in *1996 Int. Symp. on Information Theory and Its Applications, ISITA'96*, Victoria, B.C., Canada, Sept. 17–20, 1996.

[15] J. L. Massey, "All signal sets centered about the origin are optimal at low energy-to-noise ratios on the AWGN (additive white Gaussian noise) channel," in *Proc. IEEE Int. Symp. on Information Theory* (Ronneby, Sweden, June 21–24, 1976), p. 186.