Generation of Antipodal Random Vectors With Prescribed Non-Stationary 2-nd Order Statistics

Alberto Caprara, Fabio Furini, Andrea Lodi, Mauro Mangia, Riccardo Rovatti, and Gianluca Setti

***Abstract*—A Look-Up-Table-based method is proposed to generate random instances of an antipodal n-dimensional vector so that its 2-nd order statistics are as close as possible to a given specification. The method is based on linear optimization and exploits column-generation techniques to cope with the exponential complexity of the task. It yields a LUT whose storage requirements are only $\mathcal{O}(n^3)$ and thus are compatible with hardware implementation for non-negligible n. Applications are shown in the fields of Compressive Sensing and of Ultra Wide Band systems based on Direct Sequence – Code Division Multiple Acces.**

***Index Terms*—Compressed sensing, correlation, signal design, signal generators, spread spectrum.**

I. Introduction

CONTROLLING the 2-nd order statistics of a discrete-time random process means controlling how its energy/power is distributed in the signal space. It is a general problem with classical applications in telecommunications, radar, image processing and measurement engineering to name a few (see e.g., [1], [2], [3]).

In its easiest embodiment, the problem of attaining prescribed 2-nd order statistics poses no additional constraint on the process to generate x_j.

In this case, if the desired statistics are stationary, it may be given in terms of a correlation $\gamma_{j,k} = \mathbf{E}[x_j x_k] = \gamma_{0,k-j}$ but also in terms of the equivalent power spectrum $\Gamma(f) : \left[-\frac{1}{2}, \frac{1}{2}\right] \mapsto \mathbb{R}^+$ such that

$$\gamma_{j,k} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \Gamma(f) e^{2\pi \mathbf{i}(j-k)f} \mathrm{d}f$$

Exploiting this, one may synthesize a linear filter whose transfer function $H(f)$ is such that $|H(f)|^2 = \Gamma(f)$ and feed it with any unit-power white process to obtain an output process that satisfies the requirement.

If the desired statistics are non-stationary but we are interested in the samples of x_j within a time window (say for $j = 0, 1, \ldots, n-1$) we may note that the entries $\gamma_{j,k}$ for $j, k = 0, \ldots, n-1$ give a symmetric and positive semidefinite matrix such that $\gamma = \omega\omega^\top$ for some $n \times n$ matrix ω.

Hence, if $w = (w_0, \ldots, w_{n-1})^\top$ is any random vector whose correlation matrix is the identity, then ωw contains the samples x_j satisfying the requirement since $\mathbf{E}[(\omega w)(\omega w)^\top] = \omega \mathbf{E}[w w^\top]\omega^\top = \gamma$.

Yet, for obvious implementation reasons, most of contemporary information processing systems deal with discrete quantities and implicitly require that the generated process fits within that framework.

The above classical approaches, which are substantially based on linear processing, may still be applied whenever the number of quantization levels assigned to each x_j is sufficiently large. When the number of allowed levels decreases, quantization dominates and linear processing becomes much less effective.

What we address here is the extreme case in which x_j is forced to be an antipodal process, i.e., it may yield one of two opposite values.

When the 2nd-order requirement is stationary and we are interested in generating long waveforms, the problem may be tackled with existing methods [6], [7] that exploit a suitably designed linear filter in a feedback branch deciding the probability with which subsequent x_j are assigned one of the two possible values. The resulting spectra are approximations of the desired ones whose quality depends both on the target and on the algorithms used for feedback filter synthesis.

On the contrary, little can be found in the Literature to tackle the antipodal, non-stationary case when the process must be generated in a finite window. The most general available tool dates back to 1966 and is the *arcsine law* [4]. To our purposes and following an established path [3], [5][1], it can be stated that if $y \in \mathbb{R}^n$ is a zero mean Gaussian random vector whose entries

Manuscript received July 16, 2013; revised November 08, 2013 and November 08, 2013; accepted January 08, 2014. Date of publication January 27, 2014; date of current version February 26, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gesualdo Scutari.

A. Caprara, deceased, was with the Department of Electrical, Electronic, and Information Engineering (DEI), University of Bologna, 40123 Bologna, Italy.

F. Furini is with the Université Paris-Dauphine, 75775 Paris, France (e-mail: fabio.furini@dauphine.fr).

A. Lodi is with the Department of Electrical, Electronic, and Information Engineering (DEI), University of Bologna, 40123 Bologna, Italy (e-mail: andrea.lodi@unibo.it).

M. Mangia is with the Advanced Research Center on Electronic Systems for Information and Communication Technologies (ARCES), University of Bologna, 40123 Bologna, Italy (e-mail: mauro.mangia2@unibo.it).

R. Rovatti is with the Department of Electrical, Electronic, and Information Engineering (DEI), and also with the Advanced Research Center on Electronic Systems for Information and Communication Technologies (ARCES), University of Bologna, 40123 Bologna, Italy (e-mail: riccardo.rovatti@unibo.it).

G. Setti is with the Advanced Research Center on Electronic Systems for Information and Communication Technologies (ARCES), University of Bologna, 40123 Bologna, Italy, and also with the ENDIF, University of Ferrara, Ferrara, Italy (e-mail: gianluca.setti@unife.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2014.2302737

[1]The authors wish to thank one of the anonymous reviewers for bringing this technique to their attention.

are such that $\mathbf{E}[y_j^2] = 1$ for $j = 0, \ldots, n-1$ and $\mathbf{E}[y_j y_k] = g_{j,k}$ for $j, k = 0, \ldots, n-1$, and if $x_j = \operatorname{sign}(y_j)$ for $j = 0, \ldots, n-1$ then

$$\mathbf{E}[x_j x_k] = \frac{2}{\pi} \sin^{-1}(g_{j,k})$$

To have $\mathbf{E}[x_j x_k] = \gamma_{j,k}$, one may compute $g_{j,k} = \sin\left(\frac{\pi}{2}\gamma_{j,k}\right)$ and, whenever the resulting g is a nonnegative definite matrix, rely on standard generation of Gaussian multivariate to produce the needed x by clipping. We will indicate this methods as *Gaussian-based*.

This paper aims at improving on that tool and proposes a generator based on a randomly addressed digital LookUp-Table (LUT). The content of the LUT is computed by solving a large scale Linear Programming problem (LP) whose favorable properties allow attacking instances of non-negligible sizes.

The discussion is organized as follows. In Section II the problem is formally defined. In Section III we investigate how the antipodality constraint affects the possibility of reproducing an arbitrary γ and show that our method is more general than the Gaussian-based one. In Section IV the LP used to fill the LUT is described and its properties discussed. In Section V we detail the column-generation approach proposed to solve such a LP. Sections VI and VII are devoted to exemplifying the approach in two different applicative contexts. In both cases system-level performance improvement are achieved thanks to the possibility of generating antipodal vectors with a prescribed 2nd-order correlation. The last case also allows us to show that the method proposed here can improve the linear probability feedback approach when both are applicable.

II. Description of the Approach

The system we want to design is a generator of instances of an n-dimensional random vector x such that the correlation matrix $\gamma_x = \mathbf{E}[xx^\top]$ is as close as possible to a given matrix γ.

All the instances of x are constrained to be $x \in X^n$ with $X = \{-1, 1\}$. With this, we know in advance that the diagonal of γ_x contains only 1 since $\mathbf{E}[x_j^2] = 1$ independently of the vector statistics. The other entries of γ_x have to be matched to those of γ.

The general structure of the generator we propose is that of a simple digital Look-Up-Table (LUT) whose entries are strings of bits that can be mapped into vectors in X^n.

Each time an instance is needed, the LUT is addressed at random according to the probability assignment $p(x)$, i.e., so that the vector x appears at the output with probability $p(x)$.

With this, the generator, is completely defined by the content of the LUT and by the joint probability function $p : X^n \mapsto [0, 1]$.

In particular, we will see that the support

$$P = \operatorname{supp} p = \{x \in X^n \mid p(x) > 0\}$$

of the joint probability can be chosen so that its cardinality $|P|$ is much less than the 2^n possible instances of x but is limited to a number of elements that is $\mathcal{O}(n^2)$. Since each item in the LUT is an n-bit string, this makes the number of bits to be stored $\mathcal{O}(n^3)$, which remains compatible with a full hardware implementation imposing, for example, less than 1 Mbit of storage for n up to 120.

Then, the problem reduces to the determination of p starting from the desired correlation values in γ so that P has the least possible cardinalilty, i.e., as far as possible from the obvious 2^n upper bound.

This is done by defining $\delta = \gamma - \gamma_x$ as the matrix containing the deviations between the desired correlation and the correlation resulting from the choice of p and devising a design procedure based on the minimization of some norm $\|\cdot\|$ of such a deviation.

III. Properties and Limitations

We recalled in the Introduction that, starting from n independently and identically distributed random variables arranged in the vector w, one can easily generate instances of a vector x with a prescribed γ by means of a simple linear transformation.

Hence, without the constraint $x \in X^n$, the set of matrices γ for which the synthesis problem can be solved is that of $n \times n$ nonnegative definite matrices with unit diagonal entries. Let us indicate such a set as Γ_n^{NND}. For an $n \times n$ matrix γ, define $\gamma^{(k)}$ to be its $k \times k$ upper-left submatrix. The Sylvester's criterion tells us that Γ_n^{NND} is defined in the space of the parameters $\gamma_{j,k}$ with $0 < j < k < n$ by the polynomial inequalities

$$\det \gamma^{(k)} \geq 0 \quad \text{for } k = 1, \ldots, n \tag{1}$$

When $x \in X^n$, due to the discrete nature of the vector x, the correlation γ_x can be written as

$$\gamma_x = \mathbf{E}[xx^\top] = \sum_{x \in X^n} p(x) xx^\top = \sum_{x \in X^n} p(x) x^\times \tag{2}$$

where $x^\times = xx^\top$ remains implicitly defined. Since we have $p(x) \geq 0$ and $\sum_{x \in X^n} p(x) = 1$, (2) defines the convex hull of the 2^n points $x^\times$ for $x \in X^n$, i.e., a polytope that we will indicate as Γ_n^{Ant}. For example, if $n = 2$ then (2) reduces to

$$\gamma_x = \begin{pmatrix} 1 & p' - p'' \\ p' - p'' & 1 \end{pmatrix} \tag{3}$$

with $p', p'' \geq 0$ and $p' + p'' = 1$ that produces values of $\gamma_{x_{0,1}} = p' - p''$ in $[-1, 1]$.

Hence, the antipodality constraint, reduces the set of correlation matrices for which the synthesis problem can be solved to $\Gamma_n^{\mathrm{Ant}} \subseteq \Gamma_n^{\mathrm{NND}}$.

The Gaussian-based approach sketched in the Introduction relies on the generation of a multivariate zero-mean, unit variance Gaussian vector with correlations $g_{j,k} = \sin\left(\frac{\pi}{2}\gamma_{j,k}\right)$. Such an approach is successful only if the matrix g built from the matrix γ is nonnegative definite, i.e., it satisfies

$$\det g^{(k)} \geq 0 \quad \text{for } k = 1, \ldots, n \tag{4}$$

Indicate with Γ_n^{Gau} the set of matrices γ for which this happens.

If one wants to use Gaussian-based generation, the set of matrixes γ for which the synthesis problem can be solved is restricted to $\Gamma_n^{\mathrm{Gau}} \subseteq \Gamma_n^{\mathrm{Ant}} \subseteq \Gamma_n^{\mathrm{NND}}$.

It is easy to anticipate that, for n large enough, the above inclusions are strict. In fact, in the space of the parameters $\gamma_{j,k}$

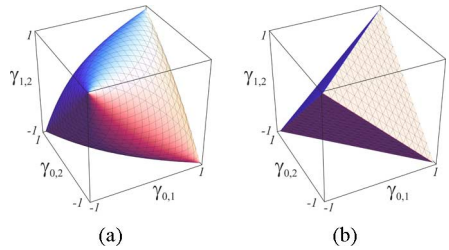


Fig. 1. Plot of Γ_3^{NND} (a) and Γ_3^{Ant} (b) in the parameter space $(\gamma_{0,1}, \gamma_{0,2}, \gamma_{1,2})$. Clearly $\Gamma_3^{\mathrm{Ant}} \subset \Gamma_3^{\mathrm{NND}}$.

with $0 < j < k < n$, Γ_n^{NND} is defined by a set of n-th degree polynomial inequalities (1), Γ_n^{Ant} is an n-dimensional polytope (2) defined by linear inequalities, and Γ^{Gau} is defined by a set of trascendental inequalities (4).

Trivially $\Gamma_1^{\mathrm{Gau}} = \Gamma_1^{\mathrm{Ant}} = \Gamma_1^{\mathrm{NND}}$. Moreover, $\Gamma_2^{\mathrm{Gau}} = \Gamma_2^{\mathrm{Ant}} = \Gamma_2^{\mathrm{NND}}$ since (1) requires $1 - \gamma_{0,1}^2 \geq 0$ while (3) implies $|\gamma_{0,1}| \leq 1$ and (4) implies $1 - \sin^2\left(\frac{\pi}{2}\gamma_{0,1}\right) \geq 0$.

For larger n, the difference between polynomial and linear inequalities comes into play and $\Gamma_3^{\mathrm{Ant}} \subset \Gamma_3^{\mathrm{NND}}$ as shown in Fig. 1 in which the two sets are plotted in the parameter space $(\gamma_{0,1}, \gamma_{0,2}, \gamma_{1,2})$. Still we have $\Gamma_3^{\mathrm{Gau}} = \Gamma_3^{\mathrm{Ant}}$.

Increasing n we get that $\Gamma_4^{\mathrm{Gau}} \subset \Gamma_4^{\mathrm{Ant}}$. In fact, we may take

$$\gamma = \begin{pmatrix} 1 & 3/10 & 3/10 & 3/5 \\ 3/10 & 1 & 3/10 & 3/5 \\ 3/10 & 3/10 & 1 & 3/5 \\ 3/5 & 3/5 & 3/5 & 1 \end{pmatrix}$$

for which we have

$$\gamma = \frac{1}{80}\left[\begin{pmatrix}+1\\-1\\-1\\+1\end{pmatrix}^{\times} + \begin{pmatrix}+1\\-1\\+1\\-1\end{pmatrix}^{\times} + \begin{pmatrix}+1\\+1\\-1\\-1\end{pmatrix}^{\times} + \begin{pmatrix}+1\\+1\\+1\\-1\end{pmatrix}^{\times}\right]$$
$$+\frac{13}{80}\left[\begin{pmatrix}+1\\-1\\-1\\-1\end{pmatrix}^{\times} + \begin{pmatrix}+1\\+1\\-1\\+1\end{pmatrix}^{\times} + \begin{pmatrix}+1\\-1\\+1\\+1\end{pmatrix}^{\times}\right] + \frac{37}{80}\begin{pmatrix}+1\\+1\\+1\\+1\end{pmatrix}^{\times}$$

To apply the Gaussian-based method we should set $g_{j,k} = \sin\left(\frac{\pi}{2}\gamma_{j,k}\right)$ for $j, k = 0, \ldots, n-1$. Yet, the resulting matrix features a minimum eigenvalue of ≈ -0.019 and thus is not nonnegative-definite.

Beyond this single case, one may, for example, scan all the 4×4 correlation matrices whose entries can be written as $\frac{k}{10}$ for $k = -10, \ldots, 10$ and find that 75480 of them are compatible with an antipodal process (i.e., belong to Γ_4^{Ant}) but cannot be obtained by applying the Gaussian-based method (i.e., they do not belong to Γ_4^{Gau}).

To proceed further, note that, if γ is the correlation matrix of $x = (x_0, \ldots, x_{n-1})^\top \in \mathbb{R}^n$, then $\gamma^{(n-1)}$ is the correlation of the subvector $(x_0, \ldots, x_{n-2})^\top \in \mathbb{R}^{n-1}$. Hence, if $\gamma \in \Gamma_n^*$ (where $*$ is any of "NND", "Ant", or "Gau") it must also be $\gamma^{(n-1)} \in \Gamma_{n-1}^*$. By reversing the implication we also get that if an $(n-1) \times (n-1)$ matrix exist not belonging to Γ_{n-1}^*, then there is a whole familiy of matrices not belonging to Γ_n^*.

Hence, if any of the inclusions $\Gamma_m^{\mathrm{Gau}} \subseteq \Gamma_n^{\mathrm{Ant}} \subseteq \Gamma_n^{\mathrm{NND}}$ is strict for a certain $\bar{n}$ then it is strict also for any $n \geq \bar{n}$.

In the light of this and of the above examples, for all relevant dimensionalities ($n \geq 4$), the antipodality constraint limits the ability of synthesizing second-order statistical properties but, within that limited scope, the method proposed here allows to tackle cases that would be otherwise unachievable by means of Gaussian-based generation.

As a final remark, since we base our approach on minimizing the deviation δ defined in Section II, target correlations $\gamma \in \Gamma_n^{\mathrm{NND}} \setminus \Gamma_n^{\mathrm{Ant}}$ can be still be approximated and from $\Gamma_n^{\mathrm{Gau}} \subset \Gamma_n^{\mathrm{Ant}} \subset \Gamma_n^{\mathrm{NND}}$ it is expected that the resulting approximation will be closer to the target than any Gaussian-based approximation.

IV. Small-Support Joint Probabilities With Prescribed Correlations

From the previous discussion we get that $\delta = \gamma - \gamma_x = \gamma - \sum_{x \in X^n} p(x) x^{\times}$.

We also choose to measure the quality of the solution we pursue with $\|\delta\| = \sum_{0 \leq j < k < n} |\delta_{j,k}|$ that exploits the fact that δ is symmetric (due to the intrinsic symmetry of γ and γ_x) and, most importantly, allows minimization by means of a LP problem.

In fact, we can split δ in its positive part by defining two auxiliary $n \times n$ matrices δ^+ (positive part of δ) and δ^- (negative part of δ) such that $\delta_{j,k}^+ = \max\{\delta_{j,k}, 0\}$ and $\delta_{j,k}^- = \max\{-\delta_{j,k}, 0\}$ for $j, k = 0, \ldots, n-1$. With this $\delta = \delta^+ - \delta^-$ while $\|\delta\| = \sum_{0 \leq j < k < n} \left(\delta_{j,k}^+ + \delta_{j,k}^-\right)$.

What we want is then encoded in the following minimization problem

$$\begin{aligned} \min \quad & \sum_{0 \leq j < k < n} \delta_{j,k}^+ + \delta_{j,k}^- \\ \text{s.t.} \quad & \delta^+ - \delta^- = \gamma - \sum_{x \in X^n} p(x) x^{\times} \\ & \sum_{x \in X^n} p(x) = 1 \\ & \delta^+ \geq 0 \\ & \delta^- \geq 0 \\ & p(x) \geq 0 \qquad \text{for } x \in X^n \end{aligned} \tag{5}$$

where matrix equalities and inequalities are meant to hold component-wise. Since the true degree of freedom is $\delta_{j,k} = \delta_{j,k}^+ - \delta_{j,k}^-$, the minimization of $\sum_{0 \leq j < k < n} \left(\delta_{j,k}^+ + \delta_{j,k}^-\right)$ ensures that at least one of $\delta_{j,k}^+$ and $\delta_{j,k}^-$ is zero for each j, k thus implying that δ^+ and δ^- are respectively the positive and negative part of δ as defined before. Clearly, when $\delta^+ = \delta^- = 0$ we have perfectly matched the target correlation γ_x.

Note that, since the number of independent non-diagonal entries in a symmetric $n \times n$ matrix is $N = \binom{n}{2}$, (5) contains $V = 2^n + 2N$ degrees of freedom (the values of $p(x)$ and the free entries of δ^+ and δ^-) constrained by V non-negativity constraints and $E = N + 1$ equality constraints (one enforcing probability normalization and N given by the matching the free entries of γ against those of γ_x).

The fact that (5) is a LP problem allows us to derive an upper bound on the minimum $|P|$, i.e., on the size of the smallest possible number of entries of the lookup table.

The problem is surely feasible since, for any given $\bar{x} \in X^n$ we may set $p(\bar{x}) = 1, p(x) = 0$ for $x \in X^n - \{\bar{x}\}$, and compute $\bar{\delta} = \gamma - \bar{x}^{\times}$.

Hence, the solution to (5) exists and can be found at one of the vertices of the polytope defined by the constraints, i.e., at a point defined by a number of equality constraints equal to the number of variables V.

Since $E < V$, an optimum exists such that at least $V - E$ out of the V non-negativity constraints are active, thus zeroing the corresponding variables. Among these, at most $2N$ are entries of either δ^+ or δ^- (N of which we already know are always null) leaving that at least $V - E - 2N$ values of $p(x)$ are null.

From all this, a solution to (5) exists corresponding to a lookup table with $|P| \leq E = N + 1$ entries thus substantiating the anticipated $\mathcal{O}(n^2)$ trend.

Further to that, from the very same polytope geometry, we get that optimal solutions with a larger support may exist as points on the facets between possibly multiple optimal vertices. Actually, this can be generalized by noting that if two different joint-probabilities $p'(x)$ and $p''(x)$ exist for which the deviation matrices are $\delta' = \gamma - \gamma_{x'}$ and $\delta'' = \gamma - \gamma_{x''}$ respectively then, for any $\epsilon \in [0, 1]$ we may set $p(x) = \epsilon p'(x) + (1 - \epsilon)p''(x)$ and obtain that $p(x)$ is a joint probability with support $P = P' \cup P''$, that causes a correlation matrix $\gamma_x = \epsilon\gamma_{x'} + (1 - \epsilon)\gamma_{x''}$ so that

$$\|\delta\| = \|\gamma - \epsilon\gamma_{x'} - (1 - \epsilon)\gamma_{x''}\| \leq \epsilon \|\delta'\| + (1 - \epsilon) \|\delta''\|$$

confirming that the "quality" of the new joint probability is analogous to that of the two generating joint probabilities.

V. Efficient Synthesis Procedures

The LP problem in (5) maps the task of designing the random vector generator into a precise mathematical procedure and highlights an important property (i.e., the fact that the cardinality of the lookup table can be $\mathcal{O}(n^2)$) that ensures the viability of the proposed approach.

Yet, it is a problem entailing a number of variables that is exponential in the size of the problem. Actually, though it goes out of the scope of this paper, it can be proved that solving (5) is NP-hard.

To tackle it for reasonable and useful values of n we should rely on slightly non-trivial Operation Research methods that can be better explained if we rewrite our problem in *standard form*, i.e., as

$$\begin{array}{ll} \min & z = cq \\ \text{s.t.} & Aq = r \\ & q \geq 0 \end{array} \tag{6}$$

where z is the real number to minimize that depends linearly on V non-negative variables collected in the column vector q through the coefficients contained in the V-dimensional row vector c. The same variables are also subject to E equality constraints that are described by the $E \times V$ matrix A and by the E-dimensional column vector r.

To put (5) in standard form, let us indicate with $[\![\cdot]\!]$ any operator that takes a symmetric matrix and rearranges the entries in its lower-left part (excluding the diagonal) in a column vector. Sort also the 2^n vectors $x \in X^n$ in any arbitrary order $x_0, \ldots, x_{2^n-1}$.

With this, set $q = \left([\![\delta^+]\!]^\top | [\![\delta^-]\!]^\top | p(x_0)| \ldots |p(x_{2^n-1})\right)^\top$, $r = (1| [\![\gamma]\!])^\top$ and $A = \bar{A}$, $c = \bar{c}$ with

$$\bar{c} = \left(1, \ldots, 1 \;\middle|\; 1, \ldots, 1 \;\middle|\; 0 \;\middle|\; \ldots \;\middle|\; 0 \right)$$

$$\bar{A} = \left(\begin{array}{c|c|c|c|c} 0, \ldots, 0 & 0, \ldots, 0 & 1 & \ldots & 1 \\ I_N & -I_N & [\![x_0 x_0^\top]\!] & \ldots & [\![x_{2^n-1} x_{2^n-1}^\top]\!] \end{array} \right)$$

where I_N is the $N \times N$ identity matrix.

The fact that the number of columns of $\bar{A}$ increases exponentially with n encourages us to look into *column generation* methods that directly follow from the property we already discussed on the support of the optimal $p(x)$.

In fact, from the fact that a solution of (6) exists for which at most E of the V entries of q are non-zero, we get that when the minimum is reached, only E of the columns of the matrix A are involved in the linear combination yielding r.

Hence, solving (6) boils down to selecting the *right* set of columns. In fact, once a set of E columns (usually called a *basis*) is selected and arranged as a square matrix A', and the corresponding coefficients in c are collected in the shorter vector c', when A' is nonsingular the non-zero entries of q can be found in the vector $q' = (A')^{-1} r$ and are such that the objective function amounts to $z = c'q' = c'(A')^{-1} r$.

To this, we may add the fundamental mechanism on which the most celebrated algorithm for solving LP problems (i.e., the simplex method) is based [8]: thanks to the linearity of the problem, optimality can be pursued in an iterative way by taking any candidate basis and possibly substituting its elements with new columns one at a time so that the corresponding value of z decreases at each substitution.

Hence, starting from any basis, our path to the minimum is made of two steps: (i) find a column that can enter the basis to reduce the value of z; (ii) find the column that can exit from the basis without impairing the reduction of z. If (i) does not find any suitable column then we are at the minimum and (6) is solved.

To rehearse the basics of this method, let us indicate with $\bar{A}(x)$ the column of $\bar{A}$ corresponding to a certain $x \in X^n$, i.e., $\bar{A}(x) = \begin{pmatrix} 1 \\ [\![x^\times]\!] \end{pmatrix}$ and with $\bar{c}(x)$ the corresponding coefficient in the row vector $\bar{c}$. The so-called *reduced cost* of the column corresponding to x is $\rho(x) = \bar{c}(x) - c'(A')^{-1}\bar{A}(x)$ and indicates the increase in the objective function that one obtains by changing the value of $p(x)$ from 0 to 1.

From this we get that if $\rho(x) \geq 0$ there is no advantage in setting $q(x) > 0$ and thus in introducing $\bar{A}(x)$ in the basis.

Actually, if $\rho(x) \geq 0$ for every $x \in X^n$ that currently has $q(x) = 0$ then, step (i) fails and the current solution is the optimal one since no new column can be substituted into the current basis to reduce z. On the contrary, any column $\bar{A}(x)$ corresponding to $\rho(x) < 0$ is a candidate entry in the basis.

TABLE I
COLUMN GENERATION METHOD TO SOLVE (5)

1: $p(x_j) \leftarrow 0 \quad \forall j = 1, \ldots, 2^n - 1$
2: $(A', c') \leftarrow \mathrm{LP}\left(\left(\bar{A}_0 \quad \ldots, \bar{A}_{2N}\right), (\bar{c}_0, \ldots, \bar{c}_{2N})^\top, r\right)$
3: **while** $c'(A')^{-1}r > \theta$ **do**
4: **if** $\exists j \geq 2N$ such that $-c'(A')^{-1}\bar{A}_j < 0$ **then**
5: $(A', c') \leftarrow \mathrm{LP}\left(\left(A'|\bar{A}_j\right), (c'|0), r\right)$
6: **else**
7: **exit while**
8: **end if**
9: **end while**
10: $q' \leftarrow (A')^{-1}r$
11: **for** $j \leftarrow 0, \ldots, (\#\text{columns of } A') - 1$ **do**
12: **if** A'_j is in the form $\begin{pmatrix} 1 \\ [\![x^\times]\!] \end{pmatrix}$ for some $x \in X^n$ **then**
13: $p(x) \leftarrow q'_j$
14: **end if**
15: **end for**

Once that a column $\bar{A}(x)$ corresponding to $\rho(x)$ is found, step (ii) above is implicitly obtained by solving (6) with $A = \left(A'|\bar{A}(x)\right)$ to obtain a solution in which the entry of q corresponding to the column that must exit the basis is set to 0.

To formalize the application of this method to our problem assume that the function $\mathrm{LP}(A, c, r)$ gives the matrix A' and the vector c' corresponding to the basis of the solution of (6), start from a set of columns that guarantees feasibility, and look among the rightmost columns of $\bar{A}$ that correspond to a null entry in the $\bar{c}$ vector for additional columns with a negative reduced cost.

By choosing a maximum tolerated error $\theta > 0$ we may lay down the procedure described in Table I, in which M_j indicates its j-th column of the generic matrix M.

Step 1 sets 0 as the default probability of each x. Step 2 finds a first feasible basis. Steps from 3 to 9 loop until the prescribed tolerance is not satisfied and a further column can be profitably added to the basis. Steps from 9 to 15 map the solution of (6) back into the probabilities $p(x)$ that we want to compute. Steps 2 and 5 entail the solution of a LP problem of size $\mathcal{O}(n^2)$ and can be tackled by standard means. On the contrary, step 4 entails a search that, in principle, scans all the 2^n possible candidate columns and must be considered carefully. More precisely, a trivial scan will be equivalent complexity-wise to solve the entire problem with 2^n columns, thus step 4 is generally formulated and solved as an optimization problem per se.

In fact, even assuming that finding a column with a negative reduced cost is a simple task (and it is not), the actual "quality" of a column is, in general, difficult to assess and columns that are introduced at a certain iteration may be discarded later and then re-admitted later-on in the pursue of optimality.

A possible and common approach to step 4 is to look for columns that promise to maximize the variation of the objective function if introduced in the basis, i.e., to minimize the reduced cost in the hope of minimizing the number of steps taken to update the starting basis to the optimal one. In mathematical terms, this translates into minimizing $\rho_j = -c'(A')^{-1}\bar{A}_j$ with respect to $j \geq 2N$. Since all the columns of $\bar{A}$ among which we are looking are of the kind $\begin{pmatrix} 1 \\ [\![x^\times]\!] \end{pmatrix}$ for some $x \in X^n$, this translates into the subsidiary optimization problem

$$\begin{aligned} \min \quad & \rho = -c'(A')^{-1}\begin{pmatrix} 1 \\ [\![x^\times]\!] \end{pmatrix} \\ \text{s.t.} \quad & x \in X^n \end{aligned} \tag{7}$$

that is a special case of a Binary Quadratic Problem (BQP) as the variables $x_0, \ldots, x_{n-1}$ can take only two values.

Since, in general, BQPs are NP-hard, there is little hope that a polynomial-time algorithm can be found for (7) though the same can be tackled by means of standard solvers for n up to few tens. One, well grounded technique is to recast it into a Binary Linear Programming (BLP) problem. This can be done by noting that the quadratic form in the objective function of (7) is a linear combination of the products $x_j x_k$.

To exploit that, one may then define n^2 alternative variables $y_{j,k}$ for $j, k = 0, \ldots, n-1$ with the additional constraints $y_{j,j} \in \{0, 1\}$ for any j, and $\max\{0, y_{j,j} + y_{k,k} - 1\} \leq y_{j,k} \leq \min\{y_{j,j}, y_{k,k}\}$ for $j \neq k$.

The constraints are such that $y_{j,k} = y_{j,j}y_{k,k}$ and thus, by setting $x_j = 2y_{j,j} - 1$ we obtain $x_j x_k = 4y_{j,k} - 2y_{j,j} - 2y_{k,k} + 1$. The latter equality allows the substitutions of a product with a linear combination and allows to express the objective function of (7) in linear terms.

Further to this linearization-based attack, heuristic approaches can be tried, among which we chose the evolutionary technique in [9] that is tailored to BQP problems and exhibits very good performance with larger scale problems. Indeed, note that for the method to be mathematically exact it is enough that at step 4 of the Algorithm in Table I, a column with negative reduced cost is selected, not necessarily the most negative one. In other words, in principle (7) must be solved to optimality only once to prove that no such column exists, provided a very effective heuristic for step 4 is available.

VI. APPLICATION TO COMPRESSIVE SENSING

Though the method we describe is completely general, it has a straightforward application in optimizing some Compressive Sensing (CS) architecture. CS concerns the possibility of exploiting dimensionality reduction techniques for signal acquisition to reduce the resources (hardware complexity/consumption, cost transmission, storage need, etc.) entailed by sensing operations.

In particular, CS is applied to signals modeled as n-dimensional random vectors s whose realizations can always be expressed as the linear combination of a number k of suitably chosen basis vectors, with k much smaller than n. In formulas, if Ψ is the $n \times n$ basis matrix, then $s = \Psi r$ where, for each instance of s, r has only at most $k \ll n$ non-zero entries. This assumption is called *sparsity*.

According to CS [10], instead of acquiring each component of s (for a total of n measurements), we project it along m vectors (with $k < m < n$) arranged as the rows of the so-called measurement matrix Φ so that the m-dimensional measurement vector is $t = \Phi s = \Phi\Psi r$. Though the matrix $\Theta = \Phi\Psi$ is $m \times n$ with $m < n$, sparsity allows to reconstruct r from t providing some assumptions are satisfied.

These assumptions are commonly fulfilled by using a random matrix Φ whose entries, for the sake of implementation simplicity, can be constrained to be either binary or antipodal.

When such a kind of matrix is used, reconstruction can be obtained with high probability by looking for the vector r that

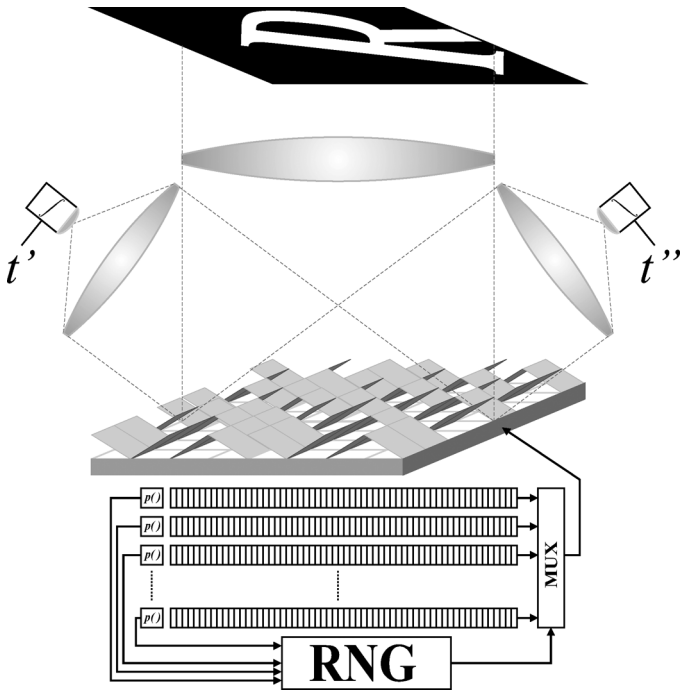


Fig. 2. Compressive sensing acquisition of an image using DMD driven by the proposed LUT-based generator of binary/antipodal vectors and one or two photon counting detectors integrating the incoming light. Measurements are taken as $t = t' - t''$ to exemplify antipodal correlation synthesis.

is compatible with measurements and has the least possible $\mathbb{L}_1$, i.e., solving

$$\begin{aligned} \min \quad & \sum_{j=0}^{n-1} |r_j| \\ \text{s.t.} \quad & \Theta r = t \end{aligned}$$

that can be easily recast into a LP problem with the same technique we used for our problem.

Further to that, it has been recently observed [12], [13] that signals, however sparse, do not usually distribute their energy evenly over the whole signal space but, on average, concentrate it along some subspaces. This localization can be exploited by keeping the classical random projection approach while maximizing the so called *rakeness*, i.e., the average amount of energy that the m projections are able to rake from the signal. The design flow (see, e.g., [13]) identifies a correlation profile that random projections should follow to achieve optimum performance as far as the reconstruction of r from t is concerned.

This is why, some of the many proposed hardware implementations of CS architectures (see e.g., [15], [16], [17]) can benefit from the generation of antipodal vectors with a prescribed correlation.

We may demonstrate this by simulating at functional level a slightly modified, toy version of the classical "one pixel camera" [14] in which an image is reflected by a digital micromirror device into two different photon counting devices simply integrating the incoming light as in Fig. 2 (actually making it a "two-pixel camera").

To exemplify our method we scale down the general system and consider the acquisition of simple black-and-white images by means of a 24×24 grid corresponding to 24×24 micromirrors.

For each image m integrations are performed, each corresponding to a randomized configuration of the mirrors. Since two counters are deployed we integrate the light coming from either kind of mirror cells and yield their difference thus ideally projecting the image along an antipodal vector in which the two possible mirror angles correspond to 1 and -1.

Original images represent small white printed numbers or letters on a black background with a gray dithering to make the curves smoother to the human eye. Numbers and letters are randomly rotated and offset from the center of the image. Although due to random rotations and offsets almost all pixels have a non-vanishing probability of being non-zero, a typical image contains only about 30 bright pixels, so that it can be considered sparse in the natural basis containing 576 vectors.

Random projections are generated by regularly partitioning the 24×24 grid into square blocks whose side is Q pixel, $Q = 2, 3, 4, 6, 8, 12$.

The statistics of each block are extracted from a training set of randomly generated images and used as the input of the design flow in [13] to compute the correlation profile that projections should have to optimize acquisition performance.

The method presented in this paper is then used to compile the $(24/Q)^2$ LUTs needed for the generation of the micromirror patterns, whose total size goes from approx 30 Kbit ($Q = 2$) to 7 Mbit ($Q = 12$).

The system is then simulated using other randomly generated testing images from which projections are taken and fed into a standard $\min -\mathbb{L}_1$ algorithm [10] to reconstruct the original image. Quality is assessed as the Average Reconstruction Signal-to-Noise Ratio (ARSNR), i.e., assuming that r is the vector unrolling of the original image and $\hat{r}$ is the result of the $\min -\mathbb{L}_1$ reconstrucion, the mean of $\|r - \hat{r}\|^2 / \|r\|^2$ over many acquisition trials.

As a reference case, we take the acquisitions of the same images by means of $m = 115$ measurements that are projections on Bernoulli sequences, i.e., sequences of independent antipodal random variables with a uniform probability assignment. The reference case results in an $\mathrm{ARSNR}^{\mathrm{ref}} = 47$ dB.

Starting from this, the performance of rakeness-based design is quantified as the minimum number, $m_{\min}$, of measurements for which $\mathrm{ARSNR} \geq \mathrm{ARSNR}^{\mathrm{ref}}$. Clearly, the lower the $m_{\min}$ the lower the resources needed to effectively acquire the given signals and the better the performance.

Fig. 3 shows how performance changes with Q, i.e., how much the ability of raking signal energy helps in reducing the resources needed to acquire it with respect to the reference case. Since adaptation to the signal depends on the ability of reproducing a given correlation profile by means of antipodal symbols, the same trend shows the effectiveness of the approach we propose here.

Fig. 4 shows some of the correlations γ_x that have been reproduced by the proposed method. For each Q, the pixels in the $Q \times Q$ block are numbered from 0 to $Q^2 - 1$ and their correlation is represented by a $Q^2 \times Q^2$ pseudo-image whose (j, k)-th pixel

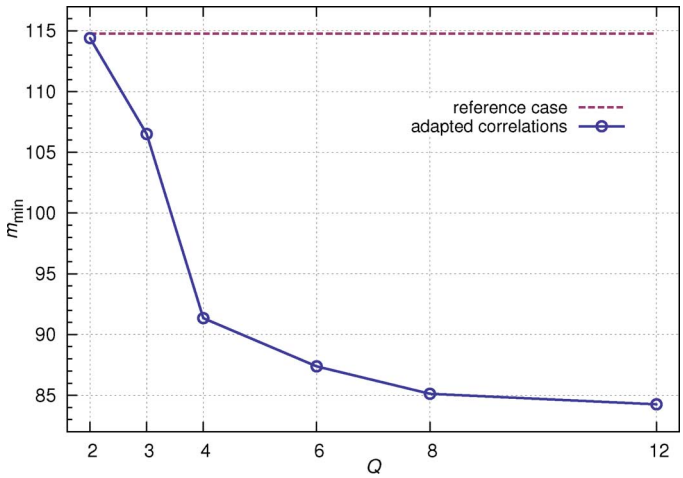


Fig. 3. The minimum number of measurements needed to ensure an ARSNR of 47 dB depending on the size of the blocks for which adaptation is sought. The increasing improvements with respect to the reference case (horizontal line) can be obtained thanks to the effectiveness of the method we propose in obtaining prescribed correlations using antipodal symbols.

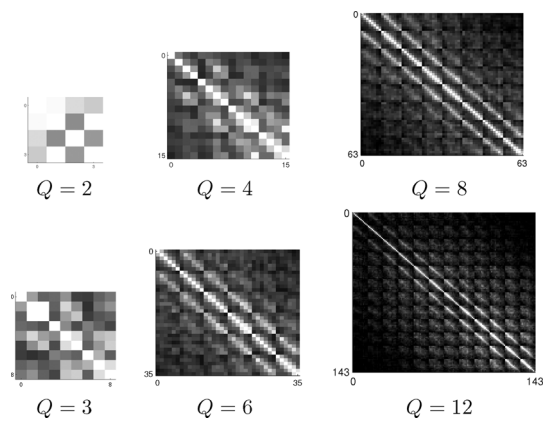


Fig. 4. Pictorial representation of some of the target correlations γ_x reproduced by the LUT-based generators in the sample application.

has a brightness proportional to the correlation value. Note that, with the exception of the diagonal in which all pixels are white since all correlations are 1, the absence of a Toeplitz structure indicates that the underlying process is non-stationary.

Improvements do not come for free since the time needed to find the LUTs content increases as Q increases. Our syntheses were carried out by means of a dedicated program written in C++ exploiting existing modules to implement the heuristic in [9], and the IBM CPLEX callable library to both solve (7) when the heuristic fails to find a suitable column, and to solve (6) whenever a column can be added. The program runs on dedicated quad-core AMD Opteron units with 32 GB of RAM and a 2.5 GHz clock frequency. With this, synthesis times went from fractions of a second for $Q = 2$ to 6×10^6 seconds (approx 73 days of offline computation) for $Q = 12$ (that implies $2^{144} \simeq 2.2 \times 10^{43}$ potential columns in the matrix A in (6)).

Note that, in this application, performance saturates and there is little need to go beyond $Q = 8$. The corresponding LUTs could be computed in approximately 3×10^5 seconds, i.e., 3.5 days of offline one-shot computation to yield a reduction in $m_{\min}$ of more than 25% with respect to the reference case.

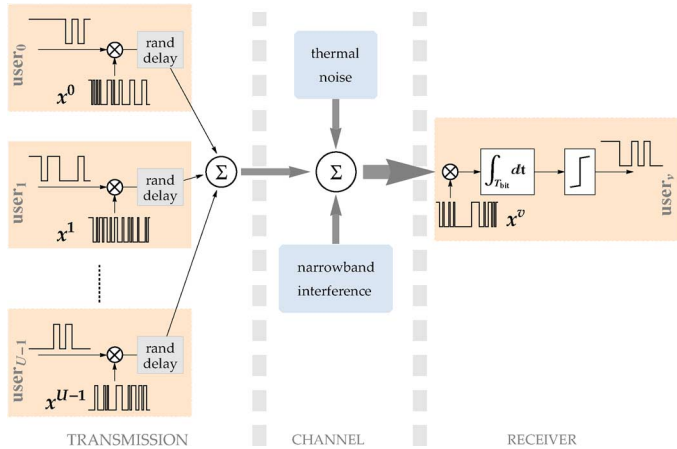


Fig. 5. Scheme of an asynchronous DS-CDMA system with multiple users using different spreading sequences and a receiver only one of them.

It is worthwhile to stress that, in any case, the computational effort is spent offline and only once for every class of signals we want to adapt to, while the generation of the random vectors requires a polynomial amount of resources.

VII. Application to Spectral Shaping in UWB Systems

Though its main purpose is the synthesis of non-stationary processes with given correlations, the method we propose can be used as a spectrum shaping procedure whenever the signal with a prescribed spectrum enters the system in windowed form.

When this happens, the fact that we rely on optimization, may result in increased performance with respect to existing approaches.

To exemplify this possibility we address the problem of designing an UWB multi-user asynchronous communication system based on DS-CDMA that must co-exist with a narrow-band service causing severe interference in part of the bandwidth. The baseband equivalent of the system is reported in Fig. 5. Each user encodes its information in an antipodal rectangular PAM with bit time $T_{\rm bit}$ and multiplies it by an antipodal spreading PAM waveform whose period is also $T_{\rm bit}$ but whose symbol time is $T_{\rm bit}/n$ with $n \gg 1$. The spreading signal of the u-th user ($u = 0, \ldots, U-1$) is identified by the symbols $x_j^u \in \{-1, 1\}$ for $j = 0, \ldots, n-1$.

Before reaching a receiver the transmitted signals experience different random delays and are corrupted by the narrow-band interference and by white thermal noise.

The receiver interested in decoding the v-th stream does so in the simplest possible way, i.e., by thresholding the output of a correlator acting as a filter matched with the signal of interest.

The presence of three kinds of disturbance reflects on the expression of the Bit-Error-Probability (BEP) suffered by the v-th receiver that is [18], [19]

$$\mathrm{BEP}^v = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty J_0\left(X_{\rm NB}^v \xi\right) \mathrm{sinc}\left(\xi\right) e^{-\xi^2 \left(X_{\rm MA}^v + X_{\rm Th}\right)} \mathrm{d}\xi \tag{8}$$

where J_0 is the Bessel function of the first-kind, $\mathrm{sinc}(\xi) = \sin(\xi)/\xi$, and the three parameters X^v_{NB}, X^v_{MA}, and X_{Th} account, respectively, for the narrow-band interference, for the multi-access interference, and for the thermal noise.

Assuming that E_{bit} is the bit energy, T_{bit} is the bit time, $P_{\mathrm{bit}} = E_{\mathrm{bit}}/T_{\mathrm{bit}}$ is the bit power, and that the narrow-band interference can be modeled as a sinusoidal tone with power P_{NB} and frequency f_{NB}, we have

$$X^v_{\mathrm{NB}} = n\sqrt{\frac{P_{\mathrm{NB}}}{P_{\mathrm{bit}}}}\left|\mathrm{sinc}\,(\zeta)\sum_{j=0}^{n-1} x^v_j e^{-2\pi \mathbf{i}\zeta j}\right| \tag{9}$$

where $\zeta = \frac{f_{\mathrm{NB}}T_{\mathrm{bit}}}{n}$.

Multiple access interference, is modeled with

$$X^v_{\mathrm{MA}} = \frac{1}{12N^3}\sum_{\substack{u=0\\u\neq v}}^{U-1}\sum_{\tau=-n+1}^{n-1} 2\Gamma^{u,v}(\tau)^2 + \Gamma^{u,v}(\tau)\Gamma^{u,v}(\tau+1) \tag{10}$$

where the partial cross-correlation function is defined as $\Gamma^{u,v}(\tau) = \sum_{j=0}^{n-\tau-1} x^u_j x^v_{j+\tau}$ if $\tau = 0,\ldots,n-1$, $\Gamma^{u,v}(\tau) = \Gamma^{v,u}(-\tau)$ if $\tau = -n+1,\ldots,-1$.

Thermal noise with two-sided spectral density equal to $N_0/2$ is accounted for by setting $X_{\mathrm{Th}} = \frac{N_0}{4E_{\mathrm{bit}}}$.

From (8), (9) and (10) one gets that performance depends on the antipodal symbols involved in spreading. Moreover, if one assumes that those symbols are taken from independent realizations of a stationary antipodal process, everything can be rewritten in terms of the autocorrelation of that process [20], thus translating performance optimization into a spectral shaping task.

As thoroughly discussed in [19], the design must administer the trade-off between the smoothly high-pass spectrum minimizing the average of X^v_{MA} even in presence of multipath propagation [20], [21], [22] and the spectrum minimizing X^v_{NB}, which clearly avoids putting power in a neighborhood of f_{NB}.

An example of a spectrum addressing such a trade-off when $f_{\mathrm{NB}}T_{\mathrm{bit}}n = 0.1$ and the NB interfering power is such that $P_{\mathrm{bit}}/P_{\mathrm{NB}} = -9$ dB is reported as a dashed line in Fig. 6. The notch around f_{NB} should, in principle, be as deep as possible while its width has been optimized in [19].

Such a profile is non-trivial per-se and the antipodality constraint forces us to consider approximations. The problem can be tackled by means of a Linear Probability Feedback (LPF) generator [7] yielding the approximation shown in Fig. 6 as the light gray continuous track.

The inverse Fourier transform of the desired spectrum can also be arranged in the matrix γ_x so that we may employ the method proposed in this paper to compute the content of a LUT from which we may extract the antipodal vectors identifying the spreading sequences of each user.

As an example, for $n = 64$, our procedure takes 1.4×10^5 s to select $\binom{64}{2}+1 = 2017$ vectors and their associated probabilities from $2^{64} \approx 1.8\times10^{18}$ potential columns.

The resulting spectrum is reported in Fig. 6 as the dark gray continuous track and is clearly a better approximation of the desired profile with respect to the one given by the LPF generator.

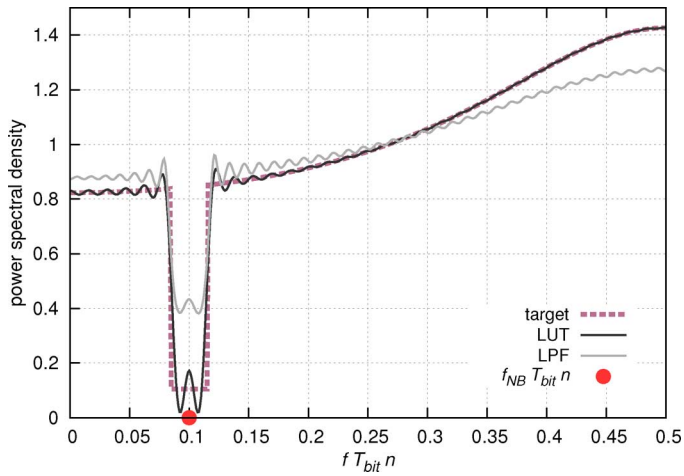


Fig. 6. A spectral profile coping with both multiple access interference and narrow-band interference (dashed line) and its approximation by means of an LPF generator and of a LUT-based generator.

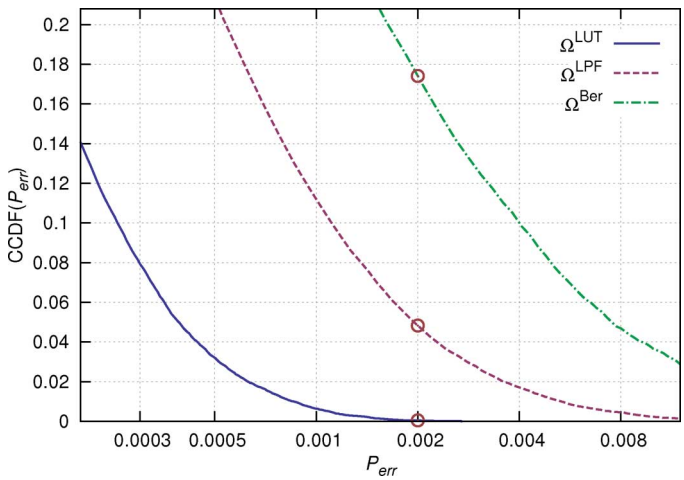


Fig. 7. The outage probability curves for systems in which spreading sequences are generated by an antipodal Bernoulli process, an LPF generator and a LUT-based generator.

To assess how much spectral shaping affects overall performance we simulate many sequence assignments and, without any loss of generality, focus on the 0-th receiver. Each assignment results in a different BEP^0 that can be computed with (8). Given a certain BEP level λ, the corresponding *outage* probability $\Omega(\lambda)$ is the probability that a sequence assignment yields a $\mathrm{BEP}^0 \geq \lambda$.

Outage probabilities depend on the policies for sequence assignment and if $\Omega'(\lambda) < \Omega''(\lambda)$ then the policy resulting in $\Omega'(\lambda)$ yields a better system than that resulting in $\Omega''(\lambda)$.

In the particular case addressed above and for $U = 6$ users, we simulate three possible policies: a purely random mechanism in which the x^u_j are independent antipodal Bernoulli random variables yielding a flat spectrum and resulting in $\Omega^{\mathrm{Ber}}(\lambda)$, a procedure based on a suitably designed LPF [19] yielding the first approximation in Fig. 6 and resulting in $\Omega^{\mathrm{LPF}}(\lambda)$, and a procedure based on the above computed LUT yielding the second approximation in Fig. 6 and resulting in $\Omega^{\mathrm{LUT}}(\lambda)$.

From the comparison of the three curves in Fig. 7 we get that spectrum shaping has a definite impact on performance and that the improvement due to the better approximation with the method we propose is unmistakeable. For example, a BEP of less than 2×10^{-3} can be guaranteed with a LUT-based sequence assignment in practically all cases ($\Omega^{\mathrm{LUT}}(0.002)\approx 0.047\%$) while it is exceeded in approximatively $\Omega^{\mathrm{LPF}}(0.002)\approx 4.8\%$ of the cases when using an LPF and in approximatively $\Omega^{\mathrm{Ber}}(0.002)\approx 17\%$ of the cases when no spectrum shaping is adopted.

VIII. Conclusion

We tackle the problem of generating n-dimensional random vectors with antipodal components whose 2nd-order correlation is as close as possible to a target one.

The design flow is formulated as a large-scale LP problem that can be solved by column generation methods and established heuristics.

The hardware-friendliness of two-value vectors allows the implementation of LUT-based generators whose storage requirement is $\mathcal{O}(n^3)$.

Examples are shown in the fields of Compressive Sensing and of UWB systems based on DS-CDMA.

Acknowledgment

To the great sadness of the other authors, Alberto Caprara prematurely passed away while working on this problem and after contributing, as it often happened, some powerful insights on its structure that will hopefully go well beyond what can be read here. Alberto will be missed by his coauthors of the present paper as an amazing scientist and great friend.

The authors would like to thank the anonymous reviewers for their thorough scan of the paper and for their punctual and constructive comments which greatly contributed to its improvement.

References

[1] F. Ochiai, "A novel trellis-shaping design with both peak and average power reduction for OFDM systems," *IEEE Trans. Commun.*, vol. 52, no. 11, pp. 1916–1926, Nov. 2004.

[2] Y. Wang, X. Dong, and I. J. Fair, "Spectrum shaping and NBI suppression in UWB communications," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1944–1952, May 2007.

[3] G. Jacovitti, A. Neri, and G. Scarano, "Texture synthesis-by-analysis with hard-limited Gaussian processes," *IEEE Trans. Image Process.*, vol. 7, no. 11, pp. 1615–1621, Nov. 1998.

[4] J. H. Van Vleck and D. Middleton, "The spectrum of clipped noise," *IEEE Proc.*, vol. 54, no. 1, pp. 2–19, Jan. 1966.

[5] P. Campisi, A. Neri, and G. Scarano, "Reduced complexity modeling and reproduction of colored textures," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 510–518, Mar. 2000.

[6] R. Rovatti, G. Mazzini, G. Setti, and S. Vitali, "Linear probability feedback processes," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2008, pp. 548–551.

[7] R. Rovatti, G. Mazzini, and G. Setti, "Memory-m antipodal processes: spectral analysis and synthesis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 1, pp. 156–167, Jan. 2009.

[8] K. G. Murty, *Linear Programming*. New York, NY, USA: Wiley, 1983.

[9] A. Lodi, K. Allemand, and T. M. Liebling, "An evolutionary heuristic for quadratic 0–1 programming," *Eur. J. Operat. Res.*, vol. 119, no. 3, pp. 662–670, Dec. 1999.

[10] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[11] M. Mangia, J. Haboba, R. Rovatti, and G. Setti, "Rakeness-based approach to compressed sensing of ECGs," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS) 2011*, 2011, pp. 424–427.

[12] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 5, pp. 1001–1014, May 2012.

[13] V. Cambareri, M. Mangia, R. Rovatti, and G. Setti, "A rakeness-based design flow for analog-to-information conversion by compressive sensing," presented at the IEEE Int. Symp. Circuits Syst. (ISCAS), 2013.

[14] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[15] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2007, pp. 1959–1962.

[16] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *IEEE J. of Sel. Topics inSignal Process.*, vol. 4, no. 2, pp. 375–391, Apr. 2010.

[17] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 443–459, Sep. 2012.

[18] A. Giorgetti, M. Chiani, and M. Z. Win, "The effect of narrowband interference on wideband wireless communication system," *IEEE Trans. Commun.*, vol. 53, no. 10, pp. 2139–2149, Oct. 2005.

[19] M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Spectral shaping of spreading sequences as a mean to address the trade-off between narrowband and multi-access interferences in UWB systems," in *Proc. Nonlinear Theory Its Appl., IEICE*, 2011, vol. 2, pp. 386–399, 4.

[20] G. Mazzini, R. Rovatti, and G. Setti, "Interference minimisation by autocorrelation shaping in asynchronous DS-CDMA systems: chaos-based spreading is nearly optimal," *Electron. Lett.*, vol. 35, pp. 1054–1055, 1999.

[21] R. Rovatti, G. Mazzini, and G. Setti, "A tensor approach to higher order expectations of quantized chaotic trajectories. I. General theory and specialization to piecewise affine Markov systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 47, no. 11, pp. 1571–1583, Nov. 2000.

[22] R. Rovatti, G. Mazzini, and G. Setti, "A tensor approach to higher order expectations of quantized chaotic trajectories. II. Application to chaos-based DS-CDMA in multipath environments," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 47, no. 11, pp. 1584–1596, Nov. 2000.

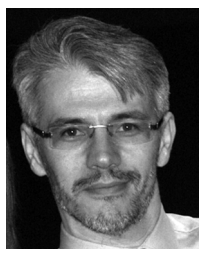
Alberto Caprara died tragically in April 2012 in a mountaineering accident. He was born in 1968, received the PhD in System Engineering from the University of Bologna in 1996, and since then he was a faculty member at the same University (Professor of Operations Research since 2005) until his death. Alberto Caprara was the recipient of several prestigious scientific prizes, including the "George B. Dantzig Dissertation Award" by the Institute for Operations Research and Management Science (NFORMS) for the best PhD thesis on applications of Operations Research in 1997. His main research interests were in Combinatorial Optimization, Mixed-Integer Linear Programming, Approximation Algorithms, Railways Applications and Computational Biology. He authored more than 100 scientific publications on top journals, books and conference proceedings.

Fabio Furini born in 1982, received his Bachelors and Masters degrees in Engineering and Industrial Management, and his Ph.D in Automation and Operations Research at the University of Bologna in 2004, 2007, and 2011 respectively. He has conducted research at prestigious universities in Europe and the United States such as the University of Colorado and Imperial College London. Currently, he is associate professor at Paris Dauphine University. His work is at the intersection of Mathematics, Economics, Information Technology, and Operations Research. He has used advanced analytical techniques to give optimal or near-optimal solutions to intricate decision-making problems. The focus of his research has been on achieving operational efficiencies, aiming to develop general software, generalizable insights and applications. In particular, he has dealt with techniques and algorithms for Mixed Integer Linear and Non-Linear Programming problems applied to challenging problems such as Aircraft Routing and Sequencing, and Train Timetabling. He has published extensively in leading international journals.

Andrea Lodi received the PhD in System Engineering from the University of Bologna in 2000 and he has been Herman Goldstine Fellow at the IBM Mathematical Sciences Department, NY in 2005–2006. He is full professor of Operations Research at DEI, University of Bologna since 2007. His main research interests are in Mixed-Integer Linear and Nonlinear Programming. He is author of more than 70 publications in the top journals of the field of Mathematical Programming. He serves as Associated Editor for several prestigious journals in the area and is currently network coordinator of the EU project FP7-PEOPLE-2012-ITN and ICT COST ACTION TD1207, and, since 2006, consultant of the IBM CPLEX research and development team.

Mauro Mangia received the B.S. and M.S. degree in electronic engineering from the University of Bologna, Italy, in 2004 and 2009 respectively and the Ph.D. degree in information technology from the same university. He is currently a post-doc researcher in the statistical signal processing group of the Advanced Research Center for Electronic Systems (ARCES) of the University of Bologna. In 2009 and 2012 he has been a visiting researcher at the École Polytechnique Fédérale de Lausanne (EPFL). His research interests are in non linear systems, compressed sensing, ultra wide band systems and system biology. He was recipient of the last Guillemin-Cauer best paper award (2013) and he also received the best student paper award at the 2011 IEEE International Symposium on Circuit and System.

Riccardo Rovatti born in 1969, he received the M.S. degree in electronic engineering and the Ph.D. degree in electronics, computer science, and telecommunications both from the University of Bologna, Italy, in 1992 and 1996, respectively. He is now a full professor of Electronics at the University of Bologna. He is the author of approximately 300 technical contributions to international conferences and journals and of two volumes. His research focuses on mathematical and applicative aspects of statistical signal processing and on the application of statistics to nonlinear dynamical systems. He received the 2004 IEEE CAS Society Darlington Award, the 2013 IEEE CAS Society Guillemin-Cauer Award, as well as the best paper award at ECCTD2005, and the best student paper award at EMCZurich2005 and at ISCAS2011. He was elected IEEE fellow in 2012 *for contributions to nonlinear and statistical signal processing applied to electronic systems*.



Gianluca Setti received a Dr. Eng. degree (with honors) in Electronic Engineering and a Ph.D. degree in Electronic Engineering and Computer Science from the University of Bologna in 1992 and in 1997. Since 1997 he has been with the School of Engineering at the University of Ferrara, Italy, where he is currently a Professor of Circuit Theory and Analog Electronics and is also a permanent faculty member of ARCES, University of Bologna. His research interests include nonlinear circuits, recurrent neural networks, implementation and application of chaotic circuits and systems, electromagnetic compatibility, and statistical signal processing.

He received the 2013 IEEE CAS Society Meritorious Service Award and was co-recipient of the 2004 IEEE CAS Society Darlington Award, of the 2013 IEEE CAS Society Guillemin-Cauer Award, as well as of the best paper award at ECCTD2005, and the best student paper award at EMCZurich2005 and at ISCAS2011.

He held several editorial positions and served, in particular, as the Editor-in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–PART II (2006–2007) and PART I (2008–2009).

He was the Technical Program Co-Chair ISCAS2007, ISCAS2008, ICECS2012, BioCAS2013 as well as the General Co-Chair of NOLTA2006.

He was Distinguished Lecturer of the IEEE CAS Society (2004–2005), a member of its Board of Governors (2005–2008), and he served as the 2010 President of CASS. He held several other volunteer positions for the IEEE and since 2013 is the first non North-American Vice President of the IEEE for Publication Services and Products.