Generation of antipodal random vectors with prescribed non-stationary 2-nd order statistics

Alberto Caprara*, Fabio Furini†, Andrea Lodi*, Mauro Mangia†, Riccardo Rovatti*,†, Gianluca Setti*‡
*DEI – University of Bologna, viale Risorgimento 2, Bologna, Italy †Université Paris-Dauphine, LAMSADE, Place du Maréchal de Lattre de Tassigny, CNRS, 75775 Paris CEDEX16, France ‡ARCES – University of Bologna, via Toffano 2/2, Bologna, Italy ¶ ENDIF – University of Ferrara, via Saragat 1, Ferrara, Italy

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 and exploits column-generation techniques to cope with the temporary 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], it can be stated that if \( y \in \mathbb{R}^n \) is a zero mean Gaussian random vector whose entries are such that \( E[y_j^2] = 1 \) for \( j = 0, \ldots, n - 1 \) and \( E[y_j y_k] = g_{j,k} \) for \( j, k = 0, \ldots, n - 1 \), and if \( x_j = \text{sign}(y_j) \) for \( j = 0, \ldots, n - 1 \) then

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

To have \( E[x_j x_k] = |\gamma_{j,k}| \), one may compute \( g_{j,k} = \sin((\frac{\pi}{2})\gamma_{j,k}) \) 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 Look-Up-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

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 \( \gamma_{j,k} = E[x_j x_k] = \gamma_{0,k-j} \) but also in terms of the equivalent power spectrum \( \Gamma(f) : [-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{R}^+ \) such that

\[
\gamma_{j,k} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \Gamma(f) e^{2\pi j(k-j)f} df
\]

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 \( j = 0, \ldots, n-1 \)) 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 \( \gamma = \omega \omega^T \) for some \( n \times n \) matrix \( \omega \).

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

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.

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 antipodality constraint affects the possibility of reproducing an arbitrary $\gamma$ 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 $\gamma_x = E[xx^\top]$ is as close as possible to a given matrix $\gamma$.

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 $\gamma_x$ contains only 1 since $E[x_j^2] = 1$ independently of the vector statistics. The other entries of $\gamma_x$ have to be matched to those of $\gamma$.

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

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

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 $O(n^2)$. Since each item in the LUT is an $n$-bit string, this makes the number of bits to be stored $O(n^3)$, which remains compatible with a full hardware implementation imposing, for example, less than 1Mbit of storage for $n$ up to 120.

Then, the problem reduces to the determination of $p$ starting from the desired correlation values in $\gamma$ so that $P$ has the least possible cardinality, 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 $\gamma$ by means of a simple linear transformation.

Hence, without the constraint $x \in X^n$, the set of matrices $\gamma$ 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 $\Gamma_n^{\text{NND}}$. For an $n \times n$ matrix $\gamma$, define $\gamma^{(k)}$ to be its $k \times k$ upper-left submatrix. The Sylvester’s criterion tells us that $\Gamma_n^{\text{NND}}$ is defined in the space of the parameters $\gamma_{j,k}$ with $0 < j < k < n$ by the polynomial inequalities

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

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

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

where $x^\top = x^\top$ 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 $x^\top x$ for $x \in X^n$, i.e., a polytope that we will indicate as $\Gamma_n^{\text{Ant}}$. For example, if $n = 2$ then (2) reduces to

$$\gamma_x = \left( \begin{array}{rr} 1 & p' - p'' \\ p' - p'' & 1 \end{array} \right)$$

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

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

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 $\gamma$ is nonnegative definite, i.e., it satisfies

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

Indicate with $\Gamma_n^{\text{Gau}}$ the set of matrices $\gamma$ for which this happens.

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

It is easy to anticipate that, for $n$ large enough, the above inclusions are strict. In fact, in the space of the parameters $\gamma_{j,k}$ with $0 < j < k < n$, $\Gamma_n^{\text{NND}}$ is defined by a set of $n$-th degree polynomial inequalities (1), $\Gamma_n^{\text{Ant}}$ is an $n$-dimensional polytope (2) defined by linear inequalities, and $\Gamma_n^{\text{Gau}}$ is defined by a set of transcendental inequalities (4).

Trivially $\Gamma_n^{\text{Gau}} = \Gamma_n^{\text{Ant}} = \Gamma_n^{\text{NND}}$. Moreover, $\Gamma_n^{\text{Gau}} = \Gamma_n^{\text{Ant}} = \Gamma_n^{\text{NND}}$ since (1) requires $1 - \gamma_{0,1} \geq 0$ 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 $\Gamma_n^{\text{Ant}} \subseteq \Gamma_n^{\text{NND}}$ as shown in Figure 1 in which the two sets are plotted in the parameter space $(\gamma_{0,1}, \gamma_{0,2}, \gamma_{1,2})$. Still we have $\Gamma_n^{\text{Gau}} = \Gamma_n^{\text{Ant}}$.

Increasing $n$ we get that $\Gamma_4^{\text{Gau}} \subseteq \Gamma_n^{\text{Ant}}$. In fact, we may take

$$\gamma = \left( \begin{array}{ccc} 1 & 3/10 & 3/10 \\ 3/10 & 1 & 3/10 \\ 3/10 & 3/10 & 1 \end{array} \right), \Gamma_3^{\text{Ant}} = \left( \begin{array}{ccc} 3/5 & 3/5 & 3/5 \\ 3/5 & 1 & 3/10 \\ 3/10 & 3/10 & 3/5 \end{array} \right)$$
for which we have

\[
\gamma = \frac{1}{80} \left[ \begin{array}{cccc}
+1^x & +1^x & +1^x & +1^x \\
-1 & -1 & -1 & +1 \\
+1 & +1 & -1 & -1 \\
1 & 0 & 1 & 0 \\
\end{array} \right] + \\
\frac{13}{80} \left[ \begin{array}{cccc}
+1^x & +1^x & +1^x & +1^x \\
-1 & +1 & -1 & -1 \\
1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 \\
\end{array} \right] + \\
\frac{37}{80} \left[ \begin{array}{cccc}
+1^x & +1^x & +1^x & +1^x \\
-1 & -1 & -1 & -1 \\
1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 \\
\end{array} \right].
\]

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 \( \approx -0.019 \) and thus is nonnegative-definite.

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

To proceed further, note that, if \( \gamma \) is the correlation matrix of \( x = (x_0, \ldots, x_n)^T \in \mathbb{R}^n \), then \( \gamma_{n-1} \) is the correlation of the subvector \( (x_0, \ldots, x_{n-1})^T \in \mathbb{R}^{n-1} \). 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 \( \Gamma_{n-1}^* \), then there is a whole family of matrices not belonging to \( \Gamma_n^* \).

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

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 \( \delta \) defined in Section II, target correlations \( \gamma \in \Gamma^\text{NND} \setminus \Gamma_n^\text{Ant} \) can be still be approximated and from \( \Gamma_m^\text{Gau} \subseteq \Gamma_n^\text{Ant} \subseteq \Gamma_n^\text{NND} \) we infer 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 \( \delta = \gamma - \gamma_x = \gamma - \sum_{x \in X^n} p(x) x^x \).

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 \( \delta \) is symmetric (due to the intrinsic symmetry of \( \gamma \) and \( \gamma_x \)) and, most importantly, allows minimization by means of a LP problem.

In fact, we can split \( \delta \) in its positive part by defining two auxiliary \( n \times n \) matrices \( \delta^+ \) (positive part of \( \delta \)) and \( \delta^- \) (negative part of \( \delta \)) 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 \( ||\delta|| = \sum_{0 \leq j < k < n} (\delta^+_{j,k} + \delta^-_{j,k}) \).

What we want is then encoded in the following minimization problem

\[
\min_{\delta} \sum_{0 \leq j < k < n} \delta^+_{j,k} + \delta^-_{j,k} \\
\text{s.t.} \quad x \in X^n \\
\delta^+ \geq 0 \\
\delta^- \geq 0 \\
p(x) \geq 0
\]

for \( x \in X^n \)

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 \( \sum_{0 \leq j < k < n} (\delta^+_{j,k} + \delta^-_{j,k}) \) ensures that at least one of \( \delta^+_{j,k} \) and \( \delta^-_{j,k} \) is zero for each \( j, k \) thus implying that \( \delta^+ \) and \( \delta^- \) are respectively the positive and negative part of \( \delta \) as defined before. Clearly, when \( \delta^+ = \delta^- = 0 \) we have perfectly matched the target correlation \( \gamma_x \).

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 \( \delta^+ \) and \( \delta^- \)) 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 \( \gamma \) against those of \( \gamma_x \)).

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 \( \tilde{x} \in X^n \) we may set \( p(\tilde{x}) = 1, \ p(x) = 0 \) for \( x \in X^n \setminus \{\tilde{x}\} \), and compute \( \tilde{\delta} = \tilde{\gamma} - \tilde{x}^x \).

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 $\delta^+$ or $\delta^-$ (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 $O(n^2)$ 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 $O(n^2)$) 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

$$\min \ z = cq$$

$$\text{s.t.} \ \ Aq = r$$

$$q \geq 0$$ (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 $q = ([\|\delta^+\| \ [\|\delta^-\| \ \|p(x_0)\| \ldots \|p(x_{2^n - 1})\|]^T$, $r = ([\|\gamma\|]^T$ and $A = \bar{A}$, $c = \bar{c}$ with

$$\bar{c} = (1, \ldots, 1 \ 1, \ldots, 1 \ 0 \ \ldots \ 0)$$

$$\bar{A} = \begin{pmatrix} 0, \ldots, 0 & 0, \ldots, 0 & 1 & \ldots & 1 \\ I_N & -I_N & x_0 x_0^T & \ldots & x_{2^n - 1} x_{2^n - 1}^T \end{pmatrix}$$

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., $\bar{A}(x) = \begin{pmatrix} 1 \\ [x x] \end{pmatrix}$ 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 $\rho(x) = c(x) - c'(A')^{-1} \bar{A}(x)$ 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.

Once that a column $\bar{A}(x)$ corresponding to $\rho(x)$ is found, step (ii) above is implicitly obtained by solving (6) with $A = (A'|\bar{A}(x))$ 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 LP($A, c, r$) gives the matrix $A'$ and
TABLE I
COLUMN GENERATION METHOD TO SOLVE (5)

1: \( p(x_j) \leftarrow 0 \) \( \forall j = 1, \ldots, 2^n - 1 \)
2: \( (A', c') \leftarrow \text{LP} \left( (A_0, \ldots, A_{2N}), (\bar{x}_0, \ldots, \bar{x}_{2N})^\top, r \right) \)
3: while \( c'(A')^{-1}r > \theta \) do
4: if \( \exists j \geq 2N \) such that \( -c'(A')^{-1}A_j < 0 \) then
5: \( (A', c') \leftarrow \text{LP} \left( (A|\bar{A}_j), (c'[0], r) \right) \)
6: else
7: exit while
8: end if
9: end while
10: \( q' \leftarrow (A')^{-1}r \)
11: for \( j = 0, \ldots, \text{(columns of } A') - 1 \) do
12: if \( A'_j \) is in the form \( \frac{1}{\|x\|} \) for some \( x \in X^n \) then
13: \( p(x) \leftarrow q'_j \)
14: end if
15: end for

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 \( x_jx_k \).

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 \( x_jx_k = 4y_{j,k} - 2y_{j,j} - 2y_{k,k} + 1 \). 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 COMPRESSION 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 \( \Psi \) 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 \( \Phi \) 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 \( \Phi \) 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 is compatible with measurements and has the least possible \( \|r\|_1 \), i.e., solving

...
by means of a $24 \times 24$ grid corresponding to $24 \times 24$ 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 $24 \times 24$ 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 $30\text{Kbit}$ ($Q=2$) to $7\text{Mbit}$ ($Q=12$).

The system is then simulated using other randomly generated testing images from which projections are taken and fed into a standard $\text{min}-L_1$ 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 $\min - L_1$ reconstruction, 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 ARSNR$_{\text{ref}} = 47$dB.

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

Figure 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.

Figure 4 shows some of the correlations $\gamma_{jk}$ 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 has a brightness proportional to the correlation value. Note that, with the exception of the diagonal in which
Fig. 3. The minimum number of measurements needed to ensure an ARSNR of 47dB 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 \( \gamma_x \) reproduced by the LUT-based generators in the sample application.

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

i.e., 3.5 days of offline one-shot computation to yield a reduction in \( \eta_{\text{min}} \) of more than 25% with respect to the reference case.

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 Figure 5. Each user encodes its information in an antipodal rectangular PAM with bit time \( T_{\text{bit}} \) and multiplies it by an antipodal spreading PAM waveform whose period is also \( T_{\text{bit}} \) but whose symbol time is \( T_{\text{bit}}/n \) 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]
where \( \zeta = \frac{f_{NB} T_{MA}}{T_{bit}} \).

Multiple access interference is modeled with

\[
X_{\text{MA}} = \frac{1}{12N^3} \sum_{u=0}^{U-1} \sum_{v=0}^{N-1} \sum_{\tau=-n+1}^{n-1} 2 \Gamma^{u,v}(\tau)^2 + \Gamma^{u,v}(\tau)\Gamma^{u,v}(\tau+1)
\]

where the partial cross-correlation function is defined as

\[
\Gamma^{u,v}(\tau) = \sum_{j=0}^{n-\tau-1} x_j^u x_{j+\tau}^v \quad \text{if } \tau = 0, \ldots, n-1, \quad \Gamma^{v,u}(\tau) = \Gamma^{v,u}(-\tau) \quad \text{if } \tau = -n+1, \ldots, -1.
\]

Thermal noise with two-sided spectral density equal to \( \frac{N_0}{2} \) is accounted for by setting \( X_{\text{Th}} = \frac{N_0}{4 T_{bit}} \).

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 \( X_{\text{MA}} \) even in presence of multipath propagation [20][22][? ] and the spectrum minimizing \( X_{\text{NB}} \), which clearly avoids putting power in a neighborhood of \( f_{\text{NB}} \).

An example of a spectrum addressing such a trade-off when \( f_{\text{NB}} T_{\text{bit}} / n = 0.1 \) and the NB interfering power is such that \( P_{\text{bit}} / P_{\text{NB}} = -9 \) dB is reported as a dashed line in Figure 6. The notch around \( f_{\text{NB}} \) 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 Figure 6 as the light gray continuous track.

The inverse Fourier transform of the desired spectrum can also be arranged in the matrix \( \gamma_x \) 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 \( 1.4 \times 10^5 \) s to select \( \binom{64}{1} + 1 = 2017 \) vectors and their associated probabilities from \( 2^{64} \approx 1.8 \times 10^{18} \) potential columns.

The resulting spectrum is reported in Figure 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.

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 that can be computed with (8). Given a certain BEP level \( \lambda \), the corresponding outage probability \( \Omega(\lambda) \) is the probability that a sequence assignment yields a BEP \( \lambda \).

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 \( U = 0 \) users, we simulate three possible policies: a purely random mechanism in which the \( x_j^u \) are independent antipodal Bernoulli random variables yielding a flat spectrum and resulting in
\( \Omega_{\text{LUT}}(\lambda) \), a procedure based on a suitably designed LPF \([19]\) yielding the first approximation in Figure 6 and resulting in \( \Omega_{\text{LPF}}^{\text{LUT}}(\lambda) \), and a procedure based on the above computed LUT yielding the second approximation in Figure 6 and resulting in \( \Omega_{\text{LUT}}^{\text{LUT}}(\lambda) \).

From the comparison of the three curves in Figure 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 unmistakable. For example, a BEP of less than \( 2 \times 10^{-3} \) can be guaranteed with a LUT-based sequence assignment in practically all cases \( (\Omega_{\text{LUT}}(0.002) \approx 0.047\%) \) while it is exceeded in approximatively \( \Omega_{\text{LPF}}(0.002) \approx 4.8\% \) of the cases when using an LPF and in approximatively \( \Omega_{\text{Ber}}(0.002) \approx 17\% \) 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 \( O(n^3) \).

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

REFERENCES


