Efficient Stochastic LASF codes for MIMO-OFDM Systems

Evan Mella
University of Cambridge, Digital Technology Group.
William Gates Building, JJ Thompson Avenue,
Cambridge, CB3 OFD, UK
Email: eem27@cam.ac.uk

Ian Wassell
University of Cambridge, Digital Technology Group.
William Gates Building, JJ Thompson Avenue,
Cambridge, CB3 OFD, UK
Email: ijw24@cam.ac.uk

***Abstract*— High rate MIMO-OFDM systems can be adapted to improve symbol error rate (SER) by employing effective Lattice Aided Space Frequency (LASF) codes. Such high rate systems boasts high parametric dimensions to an extent that analytical formulation becomes restrictive, and in most cases impossible. The design of LASF codes have recently been studied under gradient-based stochastic methods. In this paper a low latency and low complexity novel stochastic method to design LASF codes is proposed. This is achieved by simultaneously perturbing the LASF generator matrix G and, from the noisy empirical measurements of the objective function, estimating the gradient. Algorithm convergence, in terms of the number of iterations, is faster than any other known stochastic methods. In fact, it is five times faster than the gradient based schemes. Comparing with an average randomly generated LASF matrix $\mathbf{G}_r$, it is demonstrated that in quasi-static frequency selective channels the SER is observed to halve at a SNR of 30dB. While having significantly lower computational complexity, simulations confirm that the novel stochastic method achieves improved SER performance across the entire SNR range. Furthermore, latency is significantly reduced since each computation requires a single pair of noisy measurements of the SER per iteration irrespective of the dimension of the system.**

I. Introduction

High rate MIMO-OFDM systems can be adapted to improve symbol error rate (SER) by employing effective Lattice Aided Space Frequency (LASF) codes. Such high rate systems boasts high parametric dimensions to an extent that analytical formulation becomes restrictive, and in most cases impractical. The design of LASF codes have recently been studied under gradient-based stochastic methods. In this paper we propose a novel application of stochastic methods to design an optimal generator matrix $\mathbf{G}$ applicable across OFDM carriers. The generator matrix is designed to minimize the average SER at a given SNR. The method is based on simultaneous perturbation stochastic approximation. While having significantly lower computational complexity, simulations confirm that the novel stochastic method achieves improved SER performance across the entire SNR range. Furthermore, latency is significantly reduced since each computation requires a single pair of noisy measurements of the SER per iteration irrespective of the dimension of the system.

II. System Model and Problem formulation

A. Signal Model

Suppose $\Lambda \subseteq \mathbb{R}^n$ is an n-dimensional lattice defined as

$$\Lambda = \{\mathbf{c} = \mathbf{G}\mathbf{z} : \mathbf{z} \in \mathbb{Z}^n\} \tag{1}$$

where $\mathbf{G} \in \mathbb{R}^{n\times n}$ is the lattice generator matrix.

With a Shaping Region $S \in \mathbb{R}^n$ and a translation vector $\mathbf{u} \in \mathbb{R}^n$ [1], a lattice code specified by the generator matrix $\mathbf{G}$ is defined as

$$C = (\Lambda + \mathbf{u}) \cap S. \tag{2}$$

Note that, in this work we have assumed a spherical shaping region. A maximum of 1.6dB SNR gain has been reported in the literature.

Considering an OFDM based signal model with QAM modulation. For an l_{th} carrier we have

$$\mathbf{r}_l^c = \sqrt{\frac{\rho}{M}}\mathbf{H}_l^c\mathbf{c}_l^c + \mathbf{n}_l^c, \qquad 1 \leq l \leq L \tag{3}$$

where l denotes the carrier index, ρ is signal to noise ratio, L is the total number of carriers, $\mathbf{c}_l^c \in \mathbb{C}^{M\times 1}$ is the coded transmitted vector on the l-th carrier, $\mathbf{H}_l^c \in \mathbb{C}^{N\times M}$ is complex channel matrix for the carrier l. M is the number of transmit antennas.

A real coefficient channel matrix can equivalently be given as

$$\mathbf{H}_l = \begin{bmatrix} \Re\{\mathbf{H}_l\} & -\Im\{\mathbf{H}_l\} \\ \Im\{\mathbf{H}_l\} & \Re\{\mathbf{H}_l\} \end{bmatrix}, \tag{4}$$

which give rise to Real signal model as

$$\mathbf{r} = \sqrt{\frac{\rho}{M}}\mathbf{H}\mathbf{c} + \mathbf{n}, \tag{5}$$

where $\mathbf{c} \in \mathbb{R}^{2ML}$, and the global Frequency Domain channel matrix is given as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & 0 & \cdots & 0 & 0 \\ 0 & \mathbf{H}_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \mathbf{H}_{L-1} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{H}_L \end{bmatrix}, \quad (6)$$

where $\mathbf{H} \in \mathbb{R}^{2NL \times 2ML}$.

B. SER as the Objective Function

Recalling equation (5) and without loss of generality omitting the scaling factors and making appropriate substitution for $\mathbf{c}$ we have,

$$\mathbf{r} = \mathbf{HG}(\mathbf{z} + \mathbf{u}) + \mathbf{n}. \quad (7)$$

Now, we desire to minimize the average symbol error rate SER given $\mathbf{r}$, $\mathbf{H}$, $\mathbf{G}$, $\mathbf{z}$ and $\mathbf{u}$. This can be expressed mathematically as $\min \Theta(\mathbf{G})$. Where $\Theta(.)$ is the average symbol error rate defined as

$$\begin{aligned} \Theta(\mathbf{G}) &= \mathcal{E}\{\theta(\mathbf{G})\} \quad (8) \\ &= \int\int\int \theta(\mathbf{r}, \mathbf{H}, \mathbf{z}, \mathbf{G}) p(\mathbf{r}, \mathbf{z}, \mathbf{H}/\mathbf{G}) d\mathbf{r} d\mathbf{z} d\mathbf{H}. \end{aligned}$$

Where $\mathcal{E}\{.\}$ is the expectation operator, $p(\mathbf{r}, \mathbf{z}, \mathbf{H}/\mathbf{G})$ is the joint probability density function of $(\mathbf{r}, \mathbf{z}, \mathbf{H})$ given code generator matrix $\mathbf{G}$, and $\theta(.)$ is the empirical symbol error rate.

Optimization of $\Theta(.)$ is a minimization process if $\Theta(.)$ is a convex function, or positive definite if discrete. Furthermore, we restrict the total transmit power to L per OFDM block. This is achieved by ensuring that each $\mathbf{G}_l$ is appropriately scaled by dividing by its Frobenius norm. Therefore, while minimizing (8), this power constraint has to be satisfied simultaneously. This restriction on average total power can be expressed as

$$\begin{aligned} \sum_{l=1}^{L} \mathcal{E}[\|\mathbf{G}_l \mathbf{z}_l\|^2] &\leq P_T \\ \sum_{l=1}^{L} tr(\mathbf{G}_l \mathbf{z}_l \mathbf{z}_l^T \mathbf{G}_l^T) &\leq L \\ tr(\mathbf{c}\mathbf{c}^T) &\leq L, \quad (9) \end{aligned}$$

where P_T is the power in units of energy per OFDM-block transmission, and $\mathcal{E}[\mathbf{z}_l] = 1$. $tr(.)$ is *trace* of a matrix.

Minimization of the expression in (8) has no analytical solution, and moreover an optimal generator matrix $\mathbf{G}$ depends on the receiver structure. However, given the receiver scheme, $\mathbf{H}$, $\mathbf{z}$ and $\mathbf{r}$ statistics, an optimal $\mathbf{G}$ can be found by using stochastic methods. Recent stochastic approximation methods in [2], [3], and [4] are gradient based. A new gradient-free simultaneous perturbation stochastic approximation method, applied to OFDM-MIMO systems, is discussed next.

III. Simultaneous Perturbation Stochastic Approximation (SPSA)

A recursive procedure to iterate towards an optimal generator matrix $\mathbf{G}$ can be represented as

$$\mathbf{G}_{k+1} = \mathbf{G}_k - a_k \hat{\mathbf{g}}(\mathbf{G}_k), \quad (10)$$

a_k is a gain which satisfies $a_k > 0$, $\sum_1^\infty a_k = \infty$ and $\sum_1^\infty a_k < \infty$ for (10) to converge [5]. We need the estimate of the SER, $\hat{\mathbf{g}}(\mathbf{G}_k)$. In contrast to the recent work on stochastic code designs [3], [4], or the use of standard method of partial derivatives-which operates on every element individually, SPSA gradient approximation randomly perturbs all the elements of $\mathbf{G}$ simultaneously to obtain two measurements of $\Theta(.)$. Such that, each $\hat{\mathbf{g}}(\mathbf{G}_k)$ is computed as,

$$\hat{\mathbf{g}}(\mathbf{G}_k) = \frac{\Theta(\mathbf{G} + c_k \boldsymbol{\Delta}_k) - \Theta(\mathbf{G} - c_k \boldsymbol{\Delta}_k)}{2 c_k \boldsymbol{\Delta}_k}, \quad (11)$$

where, c_k is a function of γ, and $\boldsymbol{\Delta}_k$ is a $2ML \times 2ML$ matrix. Elements in $\boldsymbol{\Delta}_k$ are independent and randomly generated from a zero mean probability distribution. A typical choice is the use of Bernoulli ± 1 (i.e., the set (-1,1)) distribution each with probability 0.5. Normal and Uniform probability distributions are not allowed since they have infinite inverse moments. Note that, the division in (11) is an element-wise, Hadamard division.

The algorithm to implement this procedure is given in Algorithm 1 and the function $serr\_fxn$ in Algorithm 2. Other parametric inputs, such as α, γ, etc, in Algorithm 1 are obtained heuristically, by finding values that ensures convergence of the algorithm and also results in a fewer number of iteration per solution i.e., faster convergence.

Algorithm 1 SPSA Pseudocode

INPUT: K, A, c, α, γ.
Require: $\alpha, a, c > 0$,
$\mathbf{G} \leftarrow randn(n)$ ▷ Initialize, random nxn $\mathbf{G}$ matrix
for $k = 1$ to K **do**
$a_k \leftarrow a/(k + A)^\alpha$
$c_k \leftarrow c/(k + 1)^\gamma$
$\mathbf{G} \leftarrow \sqrt{n}\mathbf{G}/\|\mathbf{G}\|_F$ ▷ Normalize $\mathbf{G}$, $\|.\|_F$ if the Frobenius norm
$\boldsymbol{\Delta} \leftarrow 2round(rand(n)) - 1$
$\mathbf{G}_{p/m} \leftarrow \mathbf{G} \pm c_k \boldsymbol{\Delta}_\mathbf{k}$ ▷ Find $\mathbf{G}_p$ and $\mathbf{G}_m$ and normalize as above
$\theta_{p/m} \leftarrow serr\_fxn(\mathbf{G}_{p/m})$ ▷ Obtain avrg SER for $\mathbf{G}_p$ and $\mathbf{G}_m$
$\hat{\mathbf{g}} \leftarrow (\theta_p - \theta_m)./(2 c_k \boldsymbol{\Delta}_\mathbf{k})$
$\mathbf{G} \leftarrow \mathbf{G} - a_k \hat{\mathbf{g}}$
end for
OUTPUT: $\mathbf{G}$.

The Underlying feature of SPSA which contrasts it from other stochastic approximation methods (e.g., Finite Difference) is that it requires only two measurements per iteration regardless of dimension. This scheme distinguishes this work

Algorithm 2 serr_fxn; QAM and ZF detection

INPUT: $\mathbf{G}, SNR = 20dB$
Require: $TxDatageneration$
$D_{set} \leftarrow [-7,-5,-3,-1,1,3,5,7]$ ▷ 64-QAM
$P_s \leftarrow \sum D_{set}.^2/size(D_{set})$ ▷ 64-QAM sig power
$\mathbf{u} \leftarrow 0$ ▷ Set $u = 0$
$serr_t \leftarrow 0$ ▷ Initialize counter
$err_{runs} \leftarrow 500000$ ▷ Number of independent Txs
for $e_r = 1$ to err_{runs} **do**
$\mathbf{z} \leftarrow \mathbf{z}/\sqrt{2n_tP_s}$ ▷ Normlz $\mathbf{z} \in \Re^{2ntL}$;Pwr=.25 per Real Dim.
$\mathbf{r} \leftarrow \sqrt{snr}\mathbf{H}(\mathbf{G}\mathbf{z}+\mathbf{u}) + noise$ ▷ Received vector $\mathbf{r}$
$\mathbf{z}_{est} \leftarrow 2\sqrt{2P_s/snr}\mathbf{G}^{-1}\mathbf{H}^{-1}\mathbf{r}$ ▷ ZF Detection
$\mathbf{z}_{est} \leftarrow 2round(z_{est})$
$[number, ratio] \leftarrow symerr(z_0, z_{est})$ ▷ count errors
$serr_t \leftarrow serr_t + number$
end for
$empr_{SER} \leftarrow serr_t/(n_t \times err_{runs})$
OUTPUT: $empr_{SER}$

Simulation Parameters	Values
$Channel\ Model$	$2 \times 1\ SUI-3$
$Frames\ Transmitted$	$500,000$
$No.of carriers\ L$	256
α, β, γ	$1, .5, 1$
$Modulation$	$64-QAM$

TABLE I
SIMULATION PARAMETERS

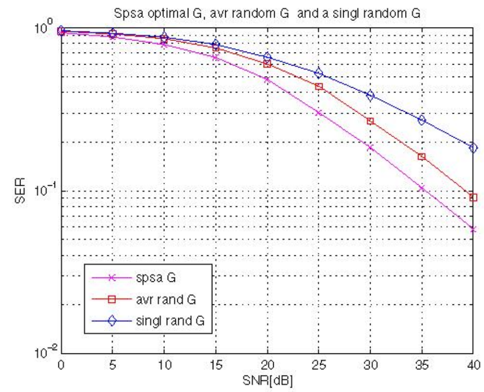

Fig. 1. SPSA-, average random- and a single random-**G**

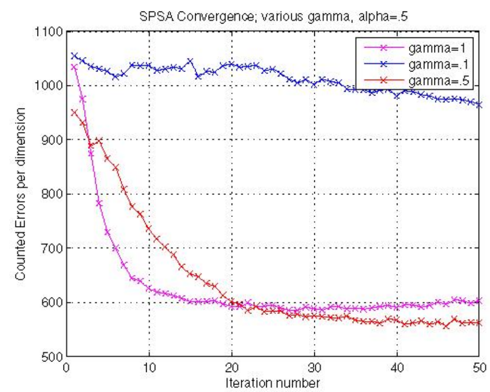

Fig. 2. The impact of the values of γ on the convergence of the SPSA Algorithm

from such work as [3], [4] which further requires explicit function-gradient relationship. Moreover, SPSA is particularly useful when considering high dimensions or a large band of frequency subcarriers over which LASF coding can be applied as it will be observed in the next section.

Performance results of the implementation of the SPSA algorithm across OFDM-MIMO systems are discussed next.

IV. Results and Discussions

Preliminary results which shows an improvement in SER performance are shown in Figure 1. Parameters for this implementation are given in Table I. With appropriate values for the parameters the algorithm as shown in Figure 2 converges. This figure illustrates the rate of convergence, and most importantly whether the algorithm converges, at all, given various values of γ. An optimal value, in this case, $\gamma = 1$ has been found heuristically.

V. CONCLUSIONS

Algorithms presented and implemented in this work are based on faster approximations of the gradient, which is obtained from noisy empirical measurements of the Symbol Error Probability. This optimization is aimed at achieving the best error rate performance at the lowest computational cost. Even with relatively large size of OFDM blocks, in terms of the number of carriers, computational complexity remains low. Although there may be global evolutionary search techniques such Genetic Algorithms and Simulated Annealing, their significantly higher requirements for resources such as power (computational complexity) and latency questions their suitability for on line applications limited in both power and size.

References

[1] H. D. M. C. G. Murugan, A. D. El Gamal, "A unified framework for treee search decoding: Rediscovering the sequential decoder," *IEEE Transactions on Information Theory,*, pp. 1483–1490, 2005.
[2] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat.*, pp. 400–407, 1951.
[3] I. Berenguer, "Advanced signal processing techniques for mimo communication systems," Ph.D. dissertation, University of Cambridge, 2005.
[4] J. W. X. Wang and M. Madihian, "Design of minimum error-rate cayley differential unitary space-time codes," *IEEE Journal on Selected Areas of Communication*, vol. 23, pp. 1779–1787, 2005.
[5] J. Spall, *Introduction to stochastic Search and Optimization: Estimation, Simulation and Control*, 2nd ed. J. Wiley and Sons, 2003.