

Discrete-Phase Constant Envelope Precoding for Massive MIMO Systems

Mohammad Kazemi, *Student Member, IEEE*, Hassan Aghaeinia,
and Tolga M. Duman, *Fellow, IEEE*

Abstract

We consider downlink of a multiuser massive multiple-input multiple-output (MIMO) system and focus on reducing the hardware costs by using a single common power amplifier and separate phase shifters for antenna front-ends. In the previous literature, the use of analog phase shifters in this set-up has been considered. Here, we study the use of practical digital phase shifters which only support a limited set of discrete phases. Considering the sum of interference powers as a metric, we formulate the corresponding nonlinear discrete optimization problem and solve for the phases to be used during transmission. We devise a low-complexity algorithm which employs a trellis structure providing suboptimal but efficient and effective solutions. We demonstrate via examples that the proposed solutions have comparable performance to conventional analog phase shifter based algorithms. Furthermore, we prove that by utilizing discrete-phase constant envelope precoding (CEP), the interference can be made arbitrarily small by increasing the number of antennas. Therefore, the asymptotic gains promised by massive MIMO systems are preserved. We also obtain closed-form expressions for the rate loss due to errors in phase and amplitude of the phase shifters, for both low and high SNR regimes.

Index Terms

M. Kazemi and H. Aghaeinia are with the Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), 424 Hafez Ave, Tehran, Iran (e-mail: mohammadkaazemi@aut.ac.ir, aghaeini@aut.ac.ir).

T. M. Duman is with the Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, Ankara, 06800, Turkey (email: duman@ee.bilkent.edu.tr).

A preliminary version of this paper was presented in part in the 2016 Information Theory and Applications (ITA) Workshop, La Jolla, CA.

Constant envelope precoding, massive MIMO, multiuser MIMO, digital phase shifter, trellis structure, discrete optimization.

I. INTRODUCTION

The seminal work of [1] in 2010 has sparked a great interest in utilizing a large number of antennas at the base station (BS), a.k.a. massive multiple-input multiple-output (MIMO), in recent years. Numerous advantages of massive MIMO have made it one of the key technology enablers for the next generation of wireless communication systems [3], [17]. One of the main advantages of massive MIMO is that as the number of BS antennas goes to infinity, the performance of linear precoders approaches to that of optimal nonlinear ones [2], [18]. On the other hand, to implement the usual precoding algorithms, each antenna element needs a radio frequency (RF) power amplifier and a phase shifter. Therefore, by increasing the number of antennas, the number of RF elements, and hence the implementation costs, increase tremendously. One way to reduce this cost is to deploy a single common RF power amplifier for all antenna elements and a separate RF phase shifter for each one, leading to antenna outputs with equal amplitudes but different phases. Precoding techniques exploiting this scenario are called constant envelope precoding (CEP) which were first considered in the context of massive MIMO in [4].

In CEP, since antenna elements have the same output amplitude and use a single common power amplifier, a highly power efficient nonlinear amplifier can be used instead of a linear one [4]. Typically, a nonlinear power amplifier can be about 4 to 6 times more power efficient than a linear one [5]. Therefore, with CEP, higher power efficiency and lower cost can outweigh the slight performance degradation, especially in the massive MIMO regime.

In [6], considering constant envelope (CE) transmission, the authors derive achievable rates of a Gaussian broadcast channel for the special case of a single user. They extend their work to the case of multiuser precoding in [4]. They obtain the desired antenna phases by solving a nonlinear nonconvex continuous optimization problem with sum of squares of interuser interference terms as the objective function to be minimized. They prove that the performance of the resulting CEP algorithm is comparable to that of precoding under an average only total transmit power constraint (APC) in the massive MIMO regime. To reduce the complexity, they also provide an alternating minimization (AM) method by taking all the phases but one to be a constant in each

sub-iteration. That is, in each iteration, they solve N one-dimensional problems instead of an N -dimensional one. The authors extend their work to frequency selective channels in [7].

In [8], cross-entropy optimization (CEO) method is used to solve the CEP problem, and it is shown that the resulting performance is less sensitive to the selection of the initial phase guesses. In order to reduce the out-of-band emissions in the common nonlinear power amplifier, in [9], the authors present a CEP with small phase variations by limiting the difference of the phase angles transmitted in consecutive channel uses. They show that, by achieving low phase variations, the performance degrades by 3 dB with every doubling of the number of BS antennas.

Only analog phase shifters, which can support a continuous range of phase shifts, have been considered in the existing literature for CEP. However, in practice, due to advantages like increased immunity to noise on their voltage control lines, more uniform unit-to-unit performance and having flat phase over a wider bandwidth, mostly digital phase shifters are used [10]. Digital phase shifters provide a discrete set of phase states which are controlled by a string of binary digits; the higher the number of supported phase states, the higher the price. For instance, a 2-bit digital phase shifter supports only 4 phase states, namely, 0, 90, 180 and 270 degrees [10].

We show that the CEP design problem can be formulated as a nonlinear discrete optimization problem considering practical digital phase shifters. In general, discrete optimization problems are NP-hard, so they are usually solved using some suboptimal algorithms which in the worst-case scenarios (depending on the problem parameters) can have exponential time complexity [11]. Even in well-conditioned problems, when dimensions of the problem grow, as in the case of massive MIMO systems, their solution time is nontrivial and in most cases impractical. Therefore, new innovative approaches are necessary to perform precoding in an efficient manner.

The contributions of this paper are fourfold: 1) considering digital phase shifters, the precoder design problem is formulated as a nonlinear discrete optimization problem, and then, using a trellis structure, a low-complexity algorithm is developed to find the optimal phases to be utilized, 2) a robust version of the proposed algorithm against channel estimation errors is presented 3) it is shown that using CEP with discrete phases, interference powers at the users can be made arbitrarily small by taking the number of BS antennas to be sufficiently large, 4) the effects of phase and amplitude errors in practical phase shifters on the performance of CEP algorithms are analyzed.

Our main approach is the following. We reformulate the interference power objective function as a sum of real terms. Then, by appropriately defining trellis states, we solve the corresponding nonlinear discrete optimization problem using a trellis structure in a manner similar to the Viterbi algorithm. Our proposed algorithm is compared with the technique of nonlinear optimization by mesh adaptive direct search (NOMAD) which has almost the same performance as other conventional discrete optimization techniques but it is much faster [13]. We demonstrate that the proposed approach has three main advantages. First, it is a few orders of magnitude less complex (hence much faster). Secondly, its performance is independent of the initial starting point. Thirdly, contrary to the conventional algorithms, its computational complexity is independent of factors such as input data or channel coefficients. We also demonstrate via extensive numerical results that the newly proposed algorithm has comparable or even better performance than NOMAD based solutions, especially in the massive MIMO scenario.

We also provide some important analytical results. We prove that, similar to the case of continuous-phase CEP in [4], the user interference is guaranteed to be arbitrarily small in discrete-phase CEP by increasing the number of BS antennas. Closed-form expressions are also provided for the rate loss due to errors in the phase and the amplitude of the phase shifters for both low and high signal-to-noise ratio (SNR) regimes. The results show that by increasing the number of antennas, the effects of phase and amplitude errors in the phase shifters do not fade away.

The rest of the paper is organized as follows. In Sections II and III, the system model and main problem formulation are introduced, respectively. In Section IV, the proposed trellis structure is presented and employed to solve the corresponding interference power minimization problem. The effects of number of BS antennas on the user interference, and phase and amplitude errors on the data rates are analyzed in Sections V and VI, respectively. Numerical examples are presented in Section VII. Finally, the paper is concluded in Section VIII.

II. SYSTEM MODEL

We consider downlink transmission in a massive MIMO system and assume that there is no cooperation among users. The system model consists of K single antenna users and a BS with N antennas. We assume that the BS knows the channel vector to each user. In our set-up, the

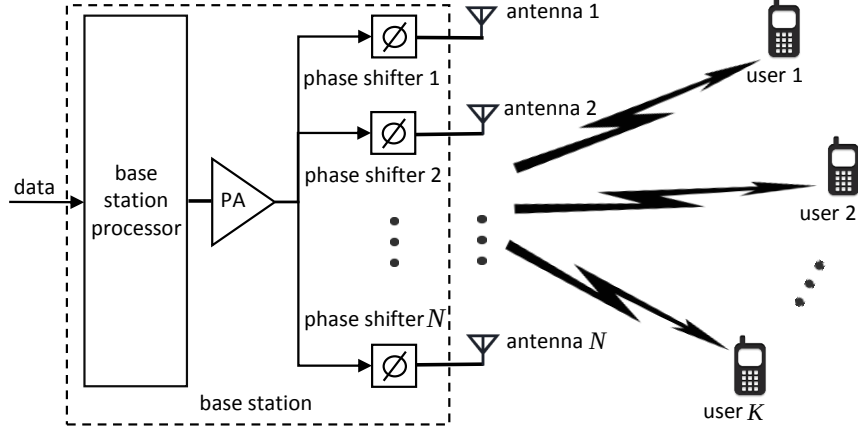


Fig. 1: System model showing downlink of a multiuser MIMO with the BS utilizing a single common power amplifier (PA) and separate phase shifters.

BS utilizes a single common RF power amplifier for all the antennas and a separate RF phase shifter for each one with the objective of CEP. The system model is illustrated in Fig. 1.

The received signal at the k th user, y_k , can be written as

$$y_k = \sqrt{\frac{P_T}{N}} \mathbf{h}_k \mathbf{q} + w_k, \quad (1)$$

where $\mathbf{q} \triangleq [e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_N}]^T$ is the vector of CE signals transmitted from the BS, and θ_i is the output phase of the i th antenna. In (1), w_k is a zero-mean complex circularly symmetric Gaussian noise with variance σ_w^2 , $\mathbf{h}_k$ is the channel vector to the k th user (k th row of the $K \times N$ channel matrix $\mathbf{H}$), and P_T is the total transmission power. We do not assume any specific fading channel model. While we provide numerical examples for Rayleigh fading channels, the proposed algorithm for CEP applies for any channel distribution, and the theoretical findings apply under very general conditions on the channel statistics specified in Sections V and VI.

III. PROBLEM FORMULATION

Denoting the desired symbol for the k th user by x_k , (1) can be divided into desired and interference parts as

$$y_k = \underbrace{\sqrt{E_k} x_k}_{\text{desired}} + \underbrace{\left(\sqrt{\frac{P_T}{N}} \mathbf{h}_k \mathbf{q} - \sqrt{E_k} x_k \right)}_{\text{interference}} + w_k, \quad (2)$$

where E_k denotes the symbol energy of the k th user. Therefore, by considering the sum of interference powers at the users as the main criterion, the desired phases can be obtained as

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \quad & \sum_{k=1}^K \left| \sqrt{\frac{P_T}{N}} \mathbf{h}_k \mathbf{q} - \sqrt{E_k} x_k \right|^2, \\ \text{s.t.} \quad & |q_n| = 1, \quad n = 1, 2, \dots, N, \end{aligned} \quad (3)$$

which is a nonlinear nonconvex optimization problem.

As can be seen in (3), a challenge in CEP is, in general, that the corresponding optimization problem is a function of both the channel coefficients and the user data, hence the precoding coefficients are functions of them as well. Therefore, even when the channel does not change, the optimization should be performed in each data transmission interval.

All the previous CEP related works (e.g., [4], [7]–[9]) have considered the use of analog phase shifters, and have solved the optimization problem (3) by using such techniques as AM [4] and CEO [8]. However, as mentioned earlier, digital phase shifters may be more preferable in practice due to factors such as increased immunity to noise on their voltage control lines [10]. Assuming digital phase shifters which support a discrete set of phase states, the optimization problem in (3) turns into a nonlinear discrete optimization problem, namely,

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \quad & \sum_{k=1}^K \left| \sqrt{\frac{P_T}{N}} \mathbf{h}_k \mathbf{q} - \sqrt{E_k} x_k \right|^2, \\ \text{s.t.} \quad & q_n \in Q_M, \quad n = 1, 2, \dots, N, \end{aligned} \quad (4)$$

where Q_M is a set of M unit amplitude symbols (e.g., phase shift keying (PSK) symbols).

Taking $q_n = \exp\left(\frac{2\pi m_n}{M}\right)$, $m_n \in \{1, \dots, M\}$, (4) turns into a nonlinear integer optimization problem and it can be solved by conventional integer programming methods such as branch-and-cut [16] and NOMAD. However, the worst-case complexity of such discrete (or integer) programming methods is exponential, and when the problem dimensions are large, their computational complexity is nontrivial, and in many cases, they become impractical even in well-conditioned problems [11]. Hence, they are not suitable options for the massive MIMO scenario in which N and consequently the dimension of the problem is very large.

To tackle this problem, we first expand the squares in the objective function of (4) and then take the summation over k . By ignoring the terms independent of $\mathbf{q}$, we have

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \quad & \frac{P_T}{N} \sum_{j=1}^N \sum_{i=1}^N q_j^* \mathbf{G}_{j,i} q_i - 2\sqrt{\frac{P_T}{N}} \sum_{j=1}^N \text{Re} \{ \mathbf{x}^H \mathbf{E} \mathbf{H}_j q_j \} \\ \text{s.t.} \quad & q_n \in Q_M, \quad n = 1, 2, \dots, N, \end{aligned} \quad (5)$$

where $\mathbf{H}_j$ is the j th column of the channel matrix $\mathbf{H}$, $\mathbf{G} \triangleq \mathbf{H}^H \mathbf{H}$, $\mathbf{x}$ is the $K \times 1$ vector of desired users' symbols, and $\mathbf{E}$ is a $K \times K$ diagonal matrix whose k th diagonal element is $\sqrt{E_k}$. Since $\mathbf{G}$ is a Hermitian matrix, we can write $\sum_{j=1}^N \sum_{i=1}^N q_j^* \mathbf{G}_{j,i} q_i = \sum_{j=1}^N \mathbf{G}_{j,j} |q_j|^2 + 2 \sum_{j=1}^N \sum_{i=1}^{j-1} \text{Re} \{ q_j^* \mathbf{G}_{j,i} q_i \}$. Substituting this into (5), we have

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \sum_{j=1}^N \left(\frac{P_T}{N} \mathbf{G}_{j,j} |q_j|^2 + 2 \text{Re} \left\{ \frac{P_T}{N} \sum_{i=1}^{j-1} q_j^* \mathbf{G}_{j,i} q_i - \sqrt{\frac{P_T}{N}} \mathbf{x}^H \mathbf{E} \mathbf{H}_j q_j \right\} \right) \\ \text{s.t. } q_n \in Q_M, \quad n = 1, 2, \dots, N. \end{aligned} \quad (6)$$

Since $|q_j|^2 = 1$, the first term becomes independent of $\mathbf{q}$, and can be ignored, so the final form of the discrete optimization problem becomes

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \sum_{j=1}^N \text{Re} \left\{ \sqrt{\frac{P_T}{N}} \sum_{i=1}^{j-1} q_j^* \mathbf{G}_{j,i} q_i - \mathbf{x}^H \mathbf{E} \mathbf{H}_j q_j \right\} \\ \text{s.t. } q_n \in Q_M, \quad n = 1, 2, \dots, N. \end{aligned} \quad (7)$$

In this form, the objective function is transformed into sum of N real terms in which the j th term is only a function of the first j variables. In the next section, we propose a trellis based optimization algorithm which exploits this property to provide a suboptimal but efficient and effective method to solve this discrete optimization problem.

IV. DISCRETE-PHASE CONSTANT ENVELOPE PRECODING

We reformulated the discrete-phase CEP problem as a sum of N real terms in which the j th term is only a function of first j variables in (7). However, solving for the optimal phases to be used is a very challenging task. For instance, the conventional discrete optimization methods are not suitable for real-time applications especially for the massive MIMO systems due to their high computational costs. In order to reduce the complexity and develop a practical design approach, we propose a low-complexity algorithm which exploits the special structure of the reformulated problem in (7). By appropriately defining trellis states, we introduce a low-complexity trellis-based solution to the main discrete optimization problem in (7), then, we analyze its computational complexity in the following two subsections.

A. Proposed Trellis-Based CEP

Since the objective function in (7) is the sum of N real terms in which the j th term is only a function of the first j variables, the problem can be solved using a multistage approach. A

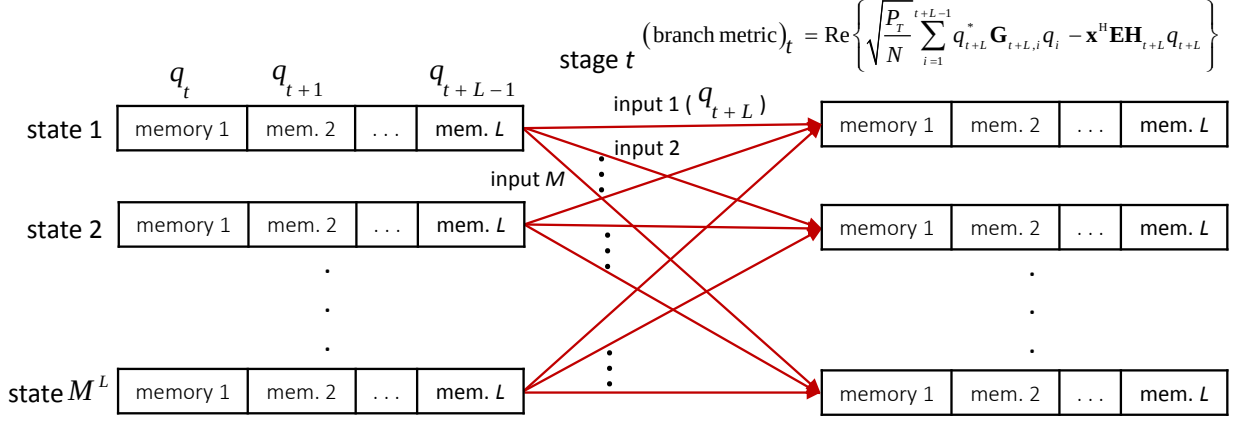


Fig. 2: Trellis structure of TB-CEP.

simple way is the greedy approach. That is, at the j th stage, we can select the symbol in the set Q_M (corresponding to the phases for the phase shifters) which minimizes the j th term of the objective function of (7) as the j th antenna output. However, we make the observation that if we keep the most likely set of variables at each stage instead of only one variable, and make the overall decision at the end of the algorithm's run, we can obtain improved results. We can do this by employing a trellis structure much like what the Viterbi algorithm utilizes in decoding or channel equalization. We call the resulting algorithm the trellis-based CEP (TB-CEP).

Define the L previous q_i variables as the memory, each taking on M possible values. Hence, we have a trellis with M^L states and each state has M outgoing branches with labels chosen from Q_M (e.g., set of M -PSK symbols). The trellis structure used in the proposed algorithm is presented in Fig. 2.

The proposed algorithm (TB-CEP) works as follows. At the first step, the initial L variables are taken as initial memory values and their M^L possible combinations form the trellis states. Initial metrics are determined by putting the values of each initial state in the first L terms of the objective function of (7), i.e., $\sum_{j=1}^L \text{Re} \left\{ \sqrt{\frac{P_T}{N}} \sum_{i=1}^{j-1} q_j^* \mathbf{G}_{j,i} q_i - \mathbf{x}^H \mathbf{E} \mathbf{H}_j q_j \right\}$. At the t th stage, the branch labels correspond to the $(t+L)$ th variable (q_{t+L}), and the branch metric is the $(t+L)$ th term of the objective function of (7) (i.e., $\text{Re} \left\{ \sqrt{\frac{P_T}{N}} \sum_{i=1}^{t+L-1} q_{t+L}^* \mathbf{G}_{t+L,i} q_i - \mathbf{x}^H \mathbf{E} \mathbf{H}_{t+L} q_{t+L} \right\}$). At each stage, the cumulative metric of branches are calculated by adding the branch metrics to the cumulative metric of their originating paths. Then, among the branches entering the same state,

Algorithm 1: Trellis-Based Constant Envelope Precoding (TB-CEP)

- 1: Calculate the initial metrics by applying the state values into the first L terms of the objective function of (7); i.e., $\sum_{j=1}^L \text{Re}\{\sqrt{\frac{P_T}{N}} \sum_{i=1}^{j-1} q_j^* \mathbf{G}_{j,i} q_i - \mathbf{x}^H \mathbf{E} \mathbf{H}_j q_j\}$
 - 2: **for** $j = L + 1 : N$ **do**
 - 3: Calculate each branch metric by applying its label and the path values associated with its originating state into the j th term of the objective function of (7); i.e., $\text{Re}\{\sqrt{\frac{P_T}{N}} \sum_{i=1}^{j-1} q_j^* \mathbf{G}_{j,i} q_i - \mathbf{x}^H \mathbf{E} \mathbf{H}_j q_j\}$
 - 4: Calculate the cumulative metric of each branch by adding its metric to the cumulative metric of its originating path
 - 5: Among branches entering the same state, choose the branch with the minimum cumulative metric and discard the other branches
 - 6: Update the cumulative path metrics
 - 7: **end for**
 - 8: Choose the path with the minimum cumulative metric
 - 9: $\mathbf{q}(L + 1 : N) \leftarrow$ branch labels of the selected path
 - 10: $\mathbf{q}(1 : L) \leftarrow$ values of the corresponding initial state
-

the branch with the least cumulative metric is kept and the others are discarded. The algorithm is terminated after $N - L$ stages and the path with the minimum cumulative metric is selected as the final decision. The labels on the selected path give the latter $N - L$ variables, and the initial state associated with the selected path gives the first L variables. The proposed algorithm is summarized in Algorithm 1. It should be noted that our proposed algorithm is not iterative and solves the optimization problem in a single run.

B. Computational Complexity Analysis

In our proposed algorithm, we have $N - L$ stages and M^L states, with M branches entering each state. Therefore, there are a total of $(N - L)M^{L+1}$ comparisons. Since the terms of (7) are used as the branch metrics, the numbers of additions and multiplications are both $O(N^2 K M^{L+1})$ for each channel use. In other words, the complexity is polynomial in the number of BS antennas which is significantly better than the alternatives which have exponential complexity [11]. We

further note that the complexity of the proposed algorithm is comparable to that of continuous-phase CEP and also to conventional linear precoders (without the CE constraint). For example, the numbers of additions and multiplications in zero forcing (ZF) precoding are $O(NK^2)$ for carrying out the pseudo-inverse in each coherence time interval of the channel, and $O(NK)$ for matrix vector product in each channel use [4]. The continuous-phase CEP has a complexity of $O(NKN_{iter})$ in each channel use, where N_{iter} is the number of iterations used in the optimization process [4].

Our numerical results in Section VII show that the proposed algorithm is a few orders of magnitude faster than the conventional discrete programming algorithms such as NOMAD in the massive MIMO scenario. Furthermore, unlike the conventional discrete programming algorithms, the complexity and running time of the proposed algorithm are fixed and independent of factors like user data and channel coefficients. Finally, the proposed algorithm does not need an initial point, on the contrary, the performance and speed of the conventional algorithms heavily depend on the choice of the initial point.

C. Robust CEP

In practice, the channel coefficients are estimated with some error. Hence, we need to focus on the effects of estimation errors on the performance of the proposed algorithm. Let the estimated channel matrix be $\hat{\mathbf{H}} \triangleq \mathbf{H} + \tilde{\mathbf{H}}$, where $\tilde{\mathbf{H}}$ denotes the error values. Assuming that $\tilde{\mathbf{H}}$ has zero mean entries and $\mathbf{C} \triangleq \mathbb{E}[\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}]$, using (3), a robust CEP problem can be presented as follows [19]:

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \quad & \sum_{k=1}^K \left| \sqrt{\frac{P_T}{N}} \hat{\mathbf{h}}_k \mathbf{q} - \sqrt{E_k} x_k \right|^2 + \frac{P_T}{N} \left\| \mathbf{C}^{\frac{1}{2}} \mathbf{q} \right\|_2^2, \\ \text{s.t.} \quad & |q_n| = 1, \quad n = 1, 2, \dots, N. \end{aligned} \quad (8)$$

Since $\mathbf{C}$ is a Hermitian matrix and $|q_n| = 1$, we have $\left\| \mathbf{C}^{\frac{1}{2}} \mathbf{q} \right\|_2^2 = \sum_{j=1}^N \sum_{i=1}^N q_j^* \mathbf{C}_{j,i} q_i = \sum_{j=1}^N \mathbf{C}_{j,j} + 2 \sum_{j=1}^N \sum_{i=1}^{j-1} \text{Re} \{ q_j^* \mathbf{C}_{j,i} q_i \}$. Therefore, in the same manner as Section III, using (8), the robust discrete-phase CEP problem can be formulated as follows:

$$\begin{aligned} \hat{\mathbf{q}} = \arg \min_{\mathbf{q}} \quad & \sum_{j=1}^N \text{Re} \left\{ \sqrt{\frac{P_T}{N}} \sum_{i=1}^{j-1} q_j^* (\mathbf{B}_{j,i} + \mathbf{C}_{j,i}) q_i - \mathbf{x}^H \mathbf{E} \hat{\mathbf{H}}_j q_j \right\} \\ \text{s.t.} \quad & q_n \in Q_M, \quad n = 1, 2, \dots, N, \end{aligned} \quad (9)$$

where $\mathbf{B} \triangleq \mathbb{E} [\hat{\mathbf{H}}^H \hat{\mathbf{H}}]$.

It can be seen that by substituting $\mathbf{G}$ with $\mathbf{B} + \mathbf{C}$ in the non-robust discrete-phase CEP problem (7), we get the above robust discrete-phase CEP problem. Therefore, the robust discrete-phase CEP problem can be solved using the proposed TB-CEP algorithm (summarized in Algorithm 1) by substituting $\mathbf{G}$ with $\mathbf{B} + \mathbf{C}$. We call the resulting algorithm the robust TB-CEP. For a more detailed analysis and management of channel uncertainties in MIMO systems, see [20] and the references therein.

V. ANALYSIS OF INTERFERENCE WITH INCREASING NUMBER OF BS ANTENNAS

In this section, we show that the interference terms at the users in CEP with discrete phases can be made arbitrarily small by taking N sufficiently large. This is proved to hold for the case of continuous-phase CEP in [4], hence we only state and prove the result as an extension to the case of digital phase shifters in Theorem 1. The general idea of the proof is as follows. Not considering noise and taking the channel vectors and the transmission power as constants, the number of possible received signal vectors at the users is equal to M^N , each one related to a specific set of antenna output phases. So by increasing the number of BS antennas, the number of possible received signal vectors increases. That is, since the transmission power is kept constant, the space of the received signal vectors becomes denser. As a result, under mild conditions on the channel gains, the probability that there exists a received signal vector (or equivalently a set of antenna output phases), whose distance to the vector of desired symbols is smaller than a certain value, increases. This probability can be shown to be nonzero, indicating that there is a solution with the desired precoding properties.

Theorem 1: Fix the number of users K . Considering the random channel coefficients matrices, $\{\mathbf{H}_N\}_{N=K}^\infty$, which satisfy $\lim_{N \rightarrow \infty} \frac{\mathbb{E}[h_{k,i} h_{r,i} h_{s,i}^* h_{t,i}^*]}{N} = 0, \forall k, r, s, t \in 1, 2, \dots, K$, for any data vector $\mathbf{x}$ chosen from a fixed finite alphabet χ and any $\Delta > 0$, there exists an integer $N'(\{\mathbf{H}_N\}, \chi, \Delta)$ such that with $N \geq N'(\{\mathbf{H}_N\}, \chi, \Delta)$, there exists a phase vector $\theta_N^x(\Delta) = [\theta_1^x(\Delta), \dots, \theta_N^x(\Delta)]^T$ chosen from a set of discrete phases which when transmitted results in the multiuser interference energy at each user being upper bounded by $2\Delta^2$. In other words, the user interference in discrete-phase CEP can be guaranteed to be arbitrarily small as $N \rightarrow \infty$.

Proof. The continuous-phase case of Theorem 1 is proved in [4]. Proof of the discrete case follows similar steps with some necessary modifications. The proof is based on showing that the vector of noiseless received signals at the users is asymptotically jointly normal which, in [4], is only proved to hold for the case of continuous phases. Since the proof in [4], except the part of joint normality, is independent of the phases being continuous or discrete, in order to avoid repetition of a long proof, here we only show that joint normality also holds for the case of discrete phases, and refer the reader to [4] for the rest of the proof. We need the following two results.

Result 1: (Lyapunov CLT) Let $\{X_n\}$ be a sequence of independent random variables. Let $E[X_n] = \mu_n$, $E[(X_n - \mu_n)^2] = \sigma_n^2$, and for some fixed $\xi > 0$, $E[|X_n - \mu_n|^{2+\xi}] = \beta_n$ exists for all n . Furthermore let $B_n \triangleq (\sum_{i=1}^n \beta_i)^{\frac{1}{2+\xi}}$, $C_n \triangleq (\sum_{i=1}^n \sigma_i^2)^{\frac{1}{2}}$. Then if $\lim_{n \rightarrow \infty} \frac{B_n}{C_n} = 0$ (to ensure that no one variable X_n dominates the summation), the cumulative distribution function (c.d.f.) of $Y_n = \frac{\sum_{i=1}^n (X_i - \mu_i)}{C_n}$ converges in the limit to the c.d.f. of a real Gaussian random variable with mean zero and unit variance [14].

Result 2: If every linear combination of a set of random variables has normal distribution, then they are jointly normal [12].

To complete the proof of Theorem 1, first using Result 1, we show that every linear combination of the noiseless received signals at the users, $z_k = \sqrt{\frac{P_T}{N}} \mathbf{h}_k \mathbf{q}$, has asymptotically a normal distribution. Then, using Result 2, we conclude that the noiseless received signals at the users are jointly normally distributed. We define the following linear combination of the noiseless received signals.

$$\zeta \triangleq \sum_{k=1}^K \lambda_k z_k = \sqrt{\frac{P_T}{N}} \sum_{k=1}^K \lambda_k \mathbf{h}_k \mathbf{q} = \sqrt{\frac{P_T}{N}} \sum_{i=1}^N e^{j\theta_i} \sum_{k=1}^K \lambda_k h_{k,i} = \sum_{i=1}^N \eta_i, \quad (10)$$

where λ_k is an arbitrary coefficient and $\eta_i \triangleq \sqrt{\frac{P_T}{N}} e^{j\theta_i} \sum_{k=1}^K \lambda_k h_{k,i}$. Since Theorem 1 is an existence theorem, only stating that such a discrete-phase precoder exists, we can assume the phase angles

to be independent of each other, and hence η_i 's become independent. Without loss of generality, we assume the output phases to be uniformly distributed, therefore for the mean and the variance of η_i , we have

$$\begin{aligned}\mu_i &\triangleq \mathbb{E}[\eta_i] = \sqrt{\frac{P_T}{N}} \mathbb{E}[e^{j\theta_i}] \sum_{k=1}^K \lambda_k h_{k,i} = 0, \\ \sigma_i^2 &\triangleq \mathbb{E}[|\eta_i - \mu_i|^2] = \mathbb{E}[|\eta_i|^2] = \frac{P_T}{N} \left| \sum_{k=1}^K \lambda_k h_{k,i} \right|^2.\end{aligned}\quad (11)$$

For $\xi = 2$, we have

$$\beta_i = \mathbb{E}[|\eta_i|^4] = \left(\frac{P_T}{N}\right)^2 \left| \sum_{k=1}^K \lambda_k h_{k,i} \right|^4, \quad (12)$$

which exists for all i . Now, using (11) and (12), we calculate B_N and C_N as follows.

$$B_N \triangleq \left(\sum_{i=1}^N \beta_i \right)^{\frac{1}{4}} = \left(\frac{P_T^2}{N} \sum_{k=1}^K \sum_{r=1}^K \sum_{s=1}^K \sum_{t=1}^K \lambda_k \lambda_r \lambda_s^* \lambda_t^* \frac{1}{N} \sum_{i=1}^N h_{k,i} h_{r,i} h_{s,i}^* h_{t,i}^* \right)^{\frac{1}{4}}, \quad (13)$$

$$C_N \triangleq \left(\sum_{i=1}^N \sigma_i^2 \right)^{\frac{1}{2}} = \left(\frac{P_T}{N} \sum_{i=1}^N \left| \sum_{k=1}^K \lambda_k h_{k,i} \right|^2 \right)^{\frac{1}{2}} = \left(P_T \sum_{k=1}^K \sum_{t=1}^K \lambda_k \lambda_t^* \frac{1}{N} \sum_{i=1}^N h_{k,i} h_{t,i}^* \right)^{\frac{1}{2}}. \quad (14)$$

Using the law of large numbers, the normalized summations of the product of channel coefficients in (13) and (14) asymptotically tend to their expected values for large N . Hence, we obtain

$$B_N \rightarrow \left(\frac{P_T^2}{N} \sum_{k=1}^K \sum_{r=1}^K \sum_{s=1}^K \sum_{t=1}^K \lambda_k \lambda_r \lambda_s^* \lambda_t^* \mathbb{E}[h_{k,i} h_{r,i} h_{s,i}^* h_{t,i}^*] \right)^{\frac{1}{4}}, \quad (15)$$

$$\begin{aligned}C_N &\rightarrow \left(P_T \sum_{k=1}^K \sum_{t=1}^K \lambda_k \lambda_t^* \mathbb{E}[h_{k,i} h_{t,i}^*] \right)^{\frac{1}{2}} \\ &= \left(P_T \left(\sum_{k=1}^K \sum_{\substack{t=1 \\ t \neq k}}^K \lambda_k \lambda_t^* |\mathbb{E}[h_{k,i}]|^2 + \sum_{k=1}^K |\lambda_k|^2 \mathbb{E}[|h_{k,i}|^2] \right) \right)^{\frac{1}{2}},\end{aligned}\quad (16)$$

where C_N is a nonzero constant independent of N (ignoring the trivial case of all-zero λ_k 's).

Using (15) and (16), we finally get

$$\lim_{N \rightarrow \infty} \frac{B_N}{C_N} = \frac{\left(P_T^2 \sum_{k=1}^K \sum_{r=1}^K \sum_{s=1}^K \sum_{t=1}^K \lambda_k \lambda_r \lambda_s^* \lambda_t^* \lim_{N \rightarrow \infty} \frac{\mathbb{E}[h_{k,i} h_{r,i} h_{s,i}^* h_{t,i}^*]}{N} \right)^{\frac{1}{4}}}{\left(P_T \left(|\mu|^2 \left(\left| \sum_{k=1}^K \lambda_k \right|^2 - \sum_{k=1}^K |\lambda_k|^2 \right) + \sigma^2 \sum_{k=1}^K |\lambda_k|^2 \right) \right)^{\frac{1}{2}}} = 0, \quad (17)$$

where we employed the assumption $\lim_{N \rightarrow \infty} \frac{\mathbb{E}[h_{k,i}h_{r,i}h_{s,i}^*h_{t,i}^*]}{N} = 0, \forall k, r, s, t \in 1, 2, \dots, K$. Note that this is a very mild constraint satisfied with many interesting fading channel models including Rayleigh and Rician fading scenarios. Hence, we conclude that the Lyapunov CLT holds for the case of the discrete phases. Therefore, using Result 2, the noiseless received signals at the users are jointly normal. The rest of the proof is the same as the continuous-phase case, so to avoid repetition, we refer the reader to [4]. $\blacksquare$

Theorem 1 basically says that by increasing N , the user interference fades away, and hence the users' interference links turn into noninterfering parallel links increasing the user data rates. This fact reinforces our choice of user interference as the design criterion in the precoding design problem which is discussed in Sections III and IV. More importantly, Theorem 1 implies that restricting the phases of the antenna elements to a limited set has no effect on the performance of the system when N is sufficiently large (as in the massive MIMO systems).

VI. RATE LOSS DUE TO PHASE AND AMPLITUDE ERRORS OF PHASE SHIFTERS

Phase shifters are not ideal in practice. A nonideal phase shifter changes the amplitude of the input signal, and also its output phase is not exactly the same as the set of predefined discrete phases. In a phase shifter, these deviations from the ideal case are called phase and amplitude errors, respectively. In this section, we analyze the effects of such errors on the performance of CEP systems. To achieve this goal, closed-form expressions for rate loss due to errors in the phase and the amplitude of the phase shifters are obtained for both low and high SNR regimes.

The following is shown to be an achievable rate for the k th user in CEP systems [4]

$$R_k(\boldsymbol{\theta}) = \log_2 \left(\frac{E_k}{I_k(\boldsymbol{\theta}) + \sigma_w^2} \right), \quad (18)$$

where $I_k(\boldsymbol{\theta}) \triangleq \mathbb{E}_{h,x} \left[\left| \sqrt{\frac{P_T}{N}} \sum_{i=1}^N h_{k,i} e^{j\theta_i} - \sqrt{E_k} x_k \right|^2 \right]$ is the average interference power at the k th user and $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_N]$ is the vector of BS output phases. Here, we employ this rate as a performance metric as well. We note that since phase shifter errors occur after the precoder optimization, they do not affect the optimization itself, however they impact the overall system performance.

Using (18), the following achievable rate can be obtained in the presence of phase and amplitude errors

$$R_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha}) = \log_2 \left(\frac{E_k}{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha}) + \sigma_w^2} \right), \quad (19)$$

where $I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha}) \triangleq \mathbb{E}_{h,x} \left[\left| \sqrt{\frac{P_T}{N}} \sum_{i=1}^N h_{k,i} (1 + \alpha_i) e^{j(\theta_i + \epsilon_i)} - \sqrt{E_k} x_k \right|^2 \right]$ is the average interference power at the k th user in the presence of phase and amplitude errors, and $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \dots, \epsilon_N]$ and $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_N]$ are the vectors of phase and amplitude errors, respectively. We define the rate loss of the k th user due to phase and amplitude errors as follows

$$(\text{RateLoss})_k \triangleq \mathbb{E}_{\epsilon, \alpha} [R_k(\boldsymbol{\theta}) - R_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})]. \quad (20)$$

Theorem 2: Let the phase shifters have zero mean amplitude and zero mean symmetric phase error vectors with covariance matrices of $\sigma_\epsilon^2 \mathbf{I}$ and $\sigma_\alpha^2 \mathbf{I}$, respectively. Then, in a CEP with $K \ll N$, the rate loss expressions of the k th user due to phase and amplitude errors are approximately equal to

$$(\text{RateLoss})_k^{\text{low SNR}} \approx \log_2(e) \frac{P_T}{\sigma_w^2} ((1 + \rho_k) + \sigma_\alpha^2 - (1 + \rho_k) \varphi_E^2(1)), \quad (21)$$

$$(\text{RateLoss})_k^{\text{high SNR}} \approx \log_2 \left(\frac{P_T}{\sigma_w^2} ((1 + \rho_k) + \sigma_\alpha^2 - (1 + \rho_k) \varphi_E^2(1)) \right), \quad (22)$$

where $\varphi_E(\cdot)$ denotes the characteristic function of the phase error and $\rho_k \triangleq \frac{E_k}{P_T}$. Note that we have $\sigma_w^2 \gg \mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})]$ and $\sigma_w^2 \ll \mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})]$ for the low and the high SNR regimes, respectively.

Proof. We make the following approximation: a good precoder minimizes the objective function well enough for $K \ll N$ to make the interference power at each user much smaller than the noise power, i.e., $I_k(\boldsymbol{\theta}) \ll \sigma_w^2$ (this is verified by our simulations of the CEP algorithms for large number of antennas). Therefore, the data rate (18) can be approximated as follows

$$R_k(\boldsymbol{\theta}) \approx \log_2 \left(\frac{E_k}{\sigma_w^2} \right). \quad (23)$$

First, we consider the case that phase and amplitude errors do not increase each user's interference power more than the noise power. That is, $\mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})] \ll \sigma_w^2$, or equivalently we operate in the low SNR regime. In this case, using first-order Taylor approximations

$\log_2(1 + \delta) \approx \delta \log_2(e)$ and $\frac{1}{1+\delta} \approx 1 - \delta$ ($|\delta| \ll 1$), and considering phase and amplitude errors, an approximation for the data rate (19) can be obtained as

$$\begin{aligned}
R_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha}) &= \log_2 \left(\frac{E_k}{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha}) + \sigma_w^2} \right) \\
&\approx \log_2 \left(\frac{E_k}{\sigma_w^2} \left(1 - \frac{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})}{\sigma_w^2} \right) \right) \\
&= \log_2 \left(\frac{E_k}{\sigma_w^2} \right) + \log_2 \left(1 - \frac{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})}{\sigma_w^2} \right) \\
&\approx \log_2 \left(\frac{E_k}{\sigma_w^2} \right) - \log_2(e) \frac{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})}{\sigma_w^2}.
\end{aligned} \tag{24}$$

Using (24) and (23), the approximate rate loss of the k th user in the low SNR regime due to phase and amplitude errors is

$$(\text{RateLoss})_k^{\text{low SNR}} \approx \frac{\log_2(e)}{\sigma_w^2} \mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})]. \tag{25}$$

The expected value of the interference power in the presence of phase and amplitude errors is obtained in the Appendix as follows

$$\mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})] \approx P_T ((1 + \rho_k) + \sigma_\alpha^2 - (1 + \rho_k) \varphi_E^2(1)). \tag{26}$$

Substituting (26) into (25) gives (21).

We next consider the case that phase and amplitude errors increase user interference powers more than the noise power. That is, $\mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})] \gg \sigma_w^2$, or equivalently the high SNR regime. In this case, considering phase and amplitude errors, the data rate can be approximated in a similar manner to (24) as follows

$$\begin{aligned}
R_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha}) &\approx \log_2 \left(\frac{E_k}{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})} \right) - \log_2(e) \frac{\sigma_w^2}{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})} \\
&\approx \log_2 \left(\frac{E_k}{\sigma_w^2} \right) + \log_2 \left(\frac{\sigma_w^2}{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})} \right).
\end{aligned} \tag{27}$$

Therefore, using (23) and (27), the approximate rate loss of the k th user in the high SNR regime due to phase and amplitude errors is

$$(\text{RateLoss})_k^{\text{high SNR}} \approx \mathbb{E}_{\epsilon, \alpha} \left[\log_2 \left(\frac{I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})}{\sigma_w^2} \right) \right] \approx \log_2 \left(\frac{\mathbb{E}_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})]}{\sigma_w^2} \right). \tag{28}$$

Finally, (22) is obtained by substituting (26) into (28). ■

It should be noted that we do not consider the channel coefficients and precoding weights to be independent of each other in the above arguments. Theorem 2 implies that by increasing N , the effects of phase and amplitude errors in the phase shifters do not fade away. One reason is that these errors occur independently of the precoding optimization in the antenna front-ends and are not accounted for in the optimization process. It should be noted that the result obtained in Theorem 2 is for general CEP and not restricted to the discrete case.

Corollary 1: For the case of phase errors with sufficiently small variances, the rate loss expressions (21) and (22) can be further approximated as

$$(\text{RateLoss})_k^{\text{low SNR}} \approx \log_2(e) \frac{P_T}{\sigma_w^2} (\sigma_\alpha^2 + (1 + \rho_k) \sigma_\epsilon^2), \quad (29)$$

$$(\text{RateLoss})_k^{\text{high SNR}} \approx \log_2 \left(\frac{P_T}{\sigma_w^2} (\sigma_\alpha^2 + (1 + \rho_k) \sigma_\epsilon^2) \right). \quad (30)$$

Proof. Assuming the phase error to be sufficiently small and knowing that it is zero mean, the second order Taylor approximation of its characteristics function becomes

$$\varphi_E(t) \triangleq \mathbb{E}_\epsilon [e^{jt\epsilon}] \approx 1 + jt\mathbb{E}_\epsilon [\epsilon] - \frac{1}{2}t^2\mathbb{E}_\epsilon [\epsilon^2] = 1 - \frac{1}{2}t^2\sigma_\epsilon^2. \quad (31)$$

Using (31) and knowing that $\sigma_\epsilon^2 \gg \sigma_\epsilon^4$, the rate loss expressions (21) and (22) can be further simplified as stated in (29) and (30), respectively. ■

Corollary 2: Assuming sufficiently small error variances, the normalized rate loss (defined as $(\text{RateLoss})_k/R_k(\boldsymbol{\theta})$) in the high SNR regime due to phase and amplitude errors of the phase shifters is

$$\frac{(\text{RateLoss})_k^{\text{high SNR}}}{R_k(\boldsymbol{\theta})} \approx 1 + \frac{\log_2 \left(\frac{1}{\rho_k} \sigma_\alpha^2 + \left(\frac{1}{\rho_k} + 1 \right) \sigma_\epsilon^2 \right)}{\log_2 \left(\frac{\rho_k P_T}{\sigma_w^2} \right)}. \quad (32)$$

Proof. The above expression is readily obtained using (23) and (30). ■

Corollary 2 shows that even the normalized rate loss due to phase and amplitude errors is independent of N and it does not fade away as N increases. According to Corollary 2, the normalized rate approaches one at very high transmit SNRs ($\frac{P_T}{\sigma_w^2}$), which makes CEP algorithms impractical in that regime. However, our simulations show that, in the high (but still practical)

SNR regime, the normalized rate loss of CEP is not too detrimental, and the proposed CEP solution remains effective.

VII. NUMERICAL EXAMPLES

First, we study the performance of the proposed trellis based discrete-phase CEP algorithm (TB-CEP) in various set-ups and compare it with other alternatives. Then, we investigate its complexity-performance trade-off. The first algorithm to be compared with is NOMAD which, as mentioned earlier, is a discrete optimization solver based on mesh adaptive direct search algorithm. The reason behind choosing this solver is that by comparing the results with those obtained via conventional solvers, we have observed that NOMAD has almost the same performance but it is much faster. The other algorithm is the continuous CEP (solving (3) using a quasi-Newton algorithm), as an upper bound on the performance.

We also note that even though we choose the sum of interference powers as our objective in the optimization, since the overall data rate is the final goal, we use the average rate per user as the main comparison criterion. Since the actual sum-rate of CEP is not available in the literature, we use an achievable data rate of CEP developed in [4] as stated in (18) for our numerical examples. In our simulations, we consider Rayleigh flat fading with independent zero mean and unit variance channel coefficients. The number of BS antennas N , the number of users K , the noise variance σ_w^2 and the transmit power P_T are set to 100, 10, 1 and 10 dBW, respectively, unless mentioned otherwise. In discrete-phase algorithms, possible values of the antenna outputs are from a set of M -PSK symbols. All the figures are obtained by averaging over 100 realizations of the channel.

As for the choice of the symbol energy, E_k , the optimal value can be obtained by maximizing (19) over E_k 's. However, to circumvent this exhaustive search, we take $E_k = P_T$, $k = 1, 2, \dots, K$, which is a near-optimal choice based on our extensive simulations.

Since solving the continuous CEP is less complex than the discrete CEP, an idea that comes to mind for the discrete CEP is to first solve the continuous CEP problem and then quantize the obtained phases to the desired set of predefined phases. We have tried this approach for many examples, and observed that there is a high total phase error especially for small constellation sizes and large number of antennas. As a result, due to the sensitivity of precoding coefficients

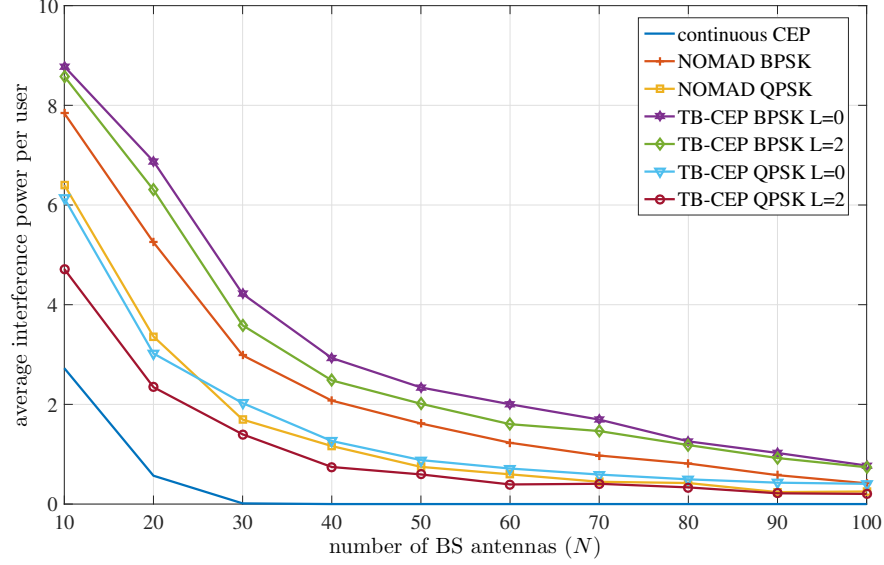


Fig. 3: Average interference power per user versus number of BS antennas ($K = 10$).

to phase errors, the overall performance degrades considerably. For instance, considering above system parameters with 4-ary phase shifters, the relative achievable rate of the quantized version of the continuous CEP with respect to the proposed TB-CEP with $L = 2$ decreases from 87% for $N = 10$ to 51% for $N = 100$. We do not consider this approach any further because of its significantly inferior performance.

Another way of solving discrete optimization problems such as discrete-phase CEP problem in (4) is to utilize semidefinite relaxation (SDR) method [21] which is presented for the case of M -PSK symbols in [22]. In order to study its effectiveness, we have performed simulations using this technique as well. We have observed that the SDR method outperforms the newly proposed TB-CEP algorithm when there are limited number of BS antennas (i.e., for regular MIMO systems), however, it has a significantly inferior performance in the massive MIMO regime. For instance, considering the above system parameters with 4-ary phase shifters, the relative achievable rate of the SDR method with respect to the proposed TB-CEP with $L = 2$ decreases from 120% for $N = 30$ to 54% for $N = 100$.

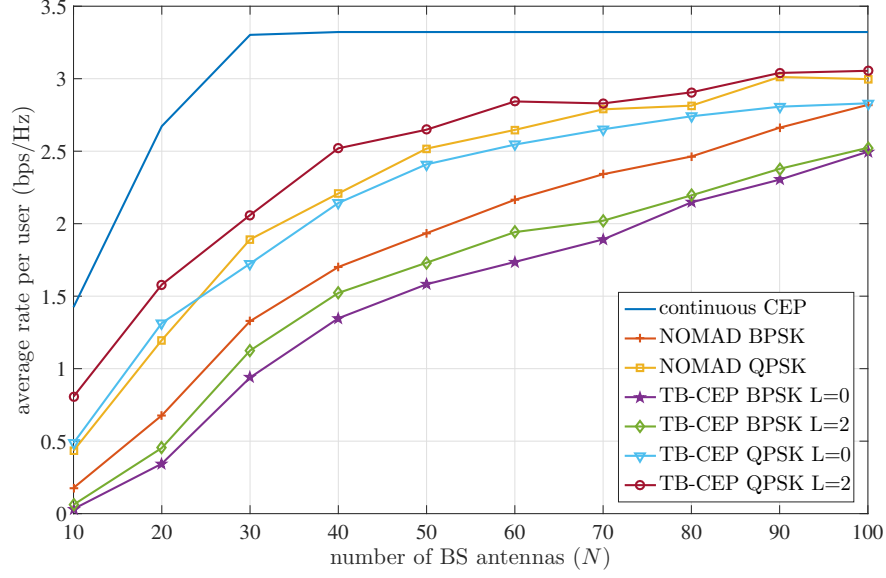


Fig. 4: Average rate per user versus number of BS antennas ($K = 10$).

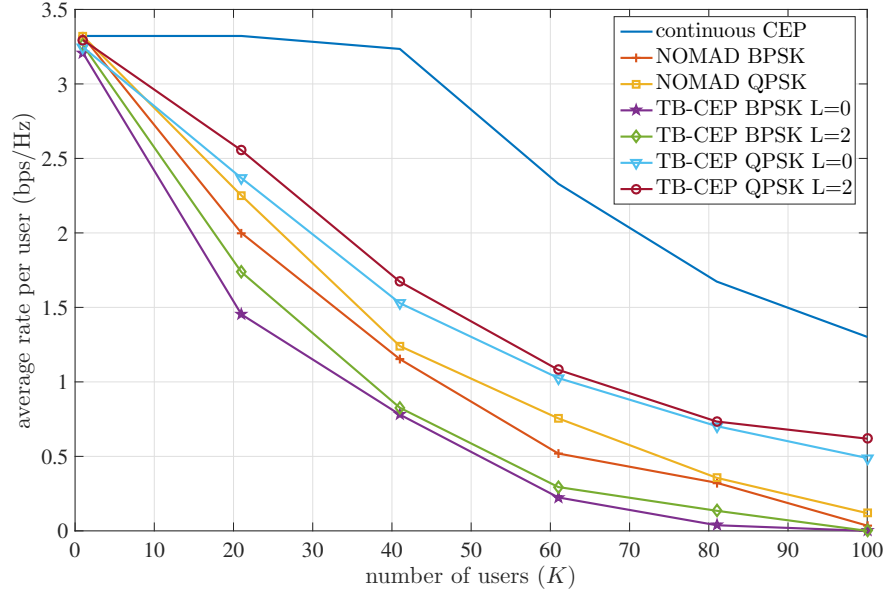


Fig. 5: Average rate per user versus number of users ($N = 100$).

A. System Performance

In Fig. 3, the average interference power per user versus the number of BS antennas is shown. It can be seen that the interference powers of all the algorithms decrease as N increases, which is in accordance with Theorem 1. Fig. 3 also gives us the minimum number of base station

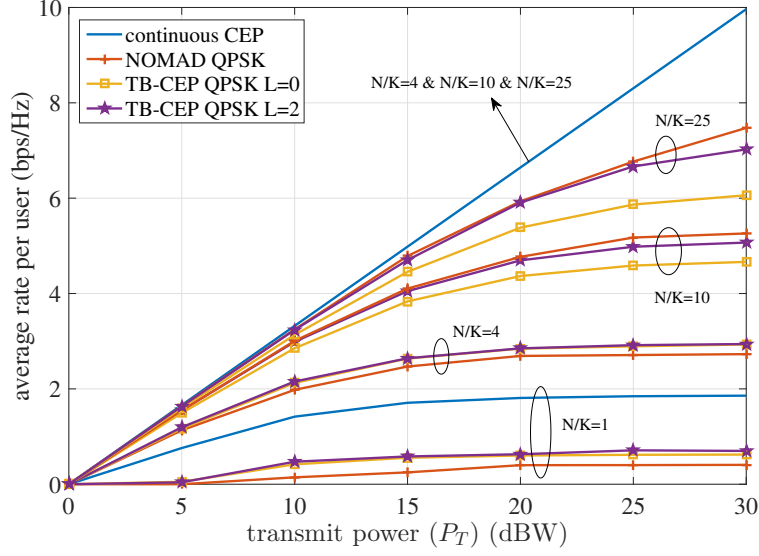


Fig. 6: Average rate per user versus transmit power.

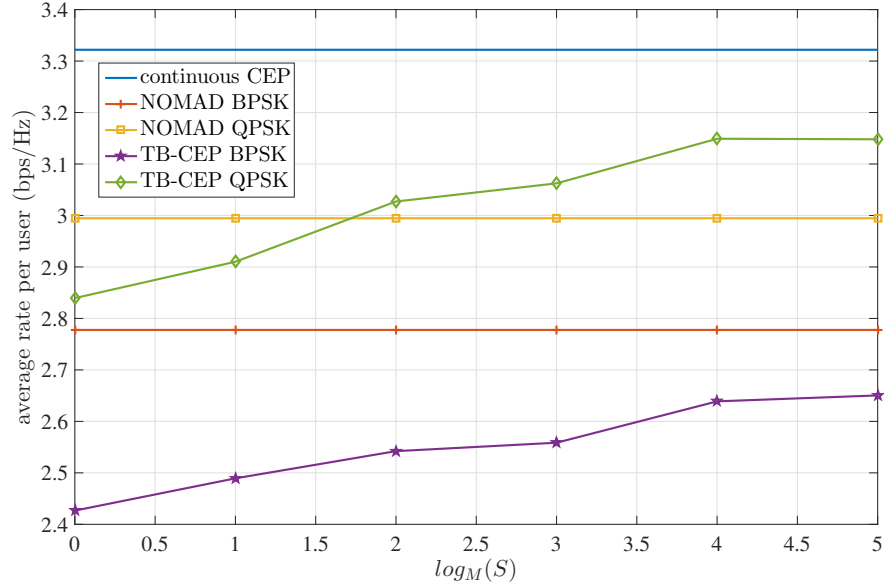


Fig. 7: Average rate per user versus number of trellis states.

antennas required for a certain interference level. For instance, the minimum number of BS antennas required for an average interference power per user of less than one (which is the same value as the noise power considered in the simulations) is 18, 36 and 44 for continuous CEP, TB-CEP QPSK with $L = 2$ and NOMAD QPSK, respectively.

The average rate per user versus the number of BS antennas is presented in Fig. 4. It can be

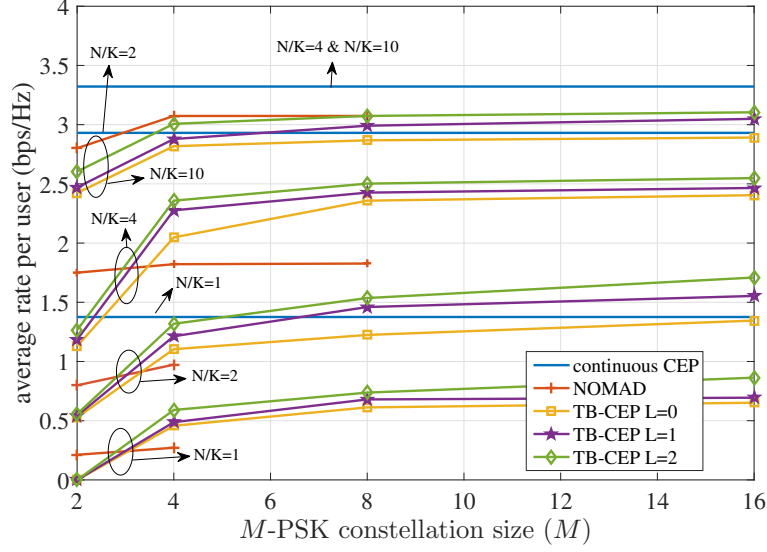


Fig. 8: Average rate per user versus M -PSK constellation size.

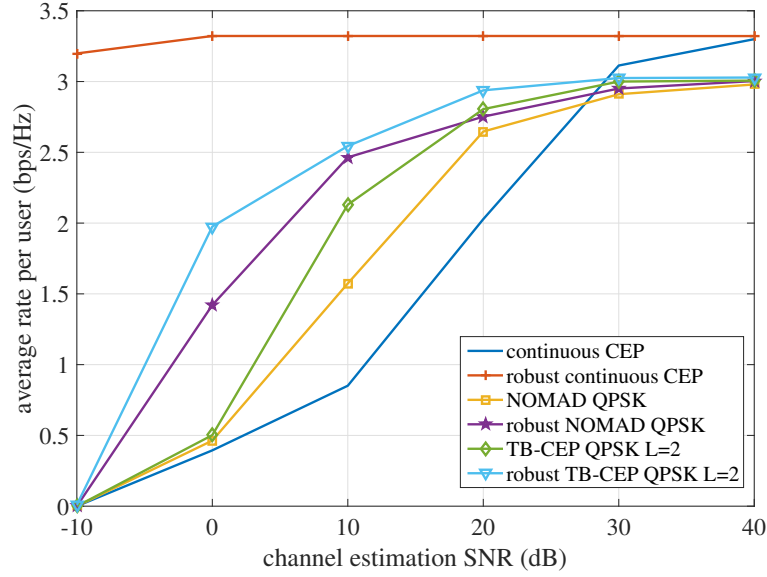


Fig. 9: Average rate per user versus channel estimation SNR.

seen that the rate increases until it saturates for all the algorithms as N increases. According to Fig. 4, for higher numbers of BS antennas (in the massive MIMO regime), the proposed trellis-based CEP (TB-CEP) has almost the same performance as the continuous CEP and surpasses NOMAD for the case of 4-ary phase shifters. For instance, for $N = 50$ with 4-ary phase shifters,

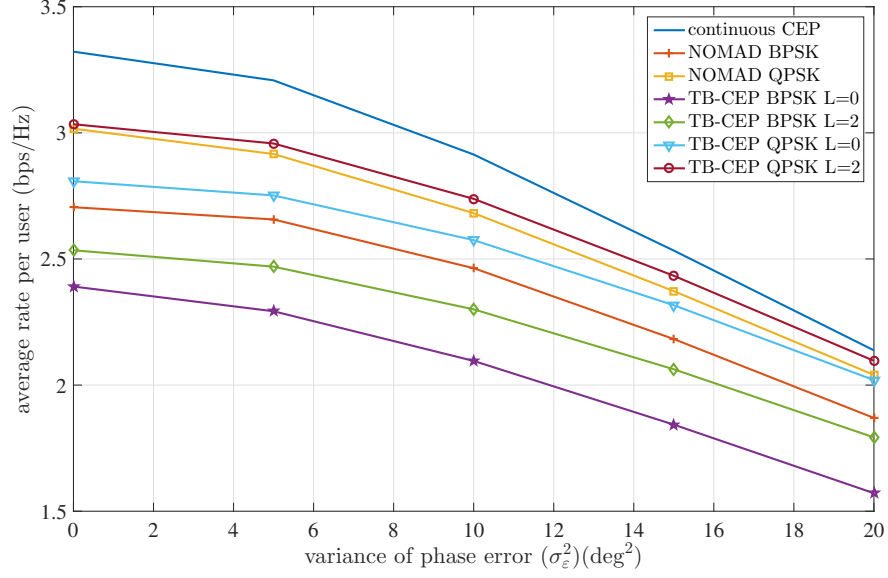


Fig. 10: Average rate per user versus phase error in RF phase shifters.

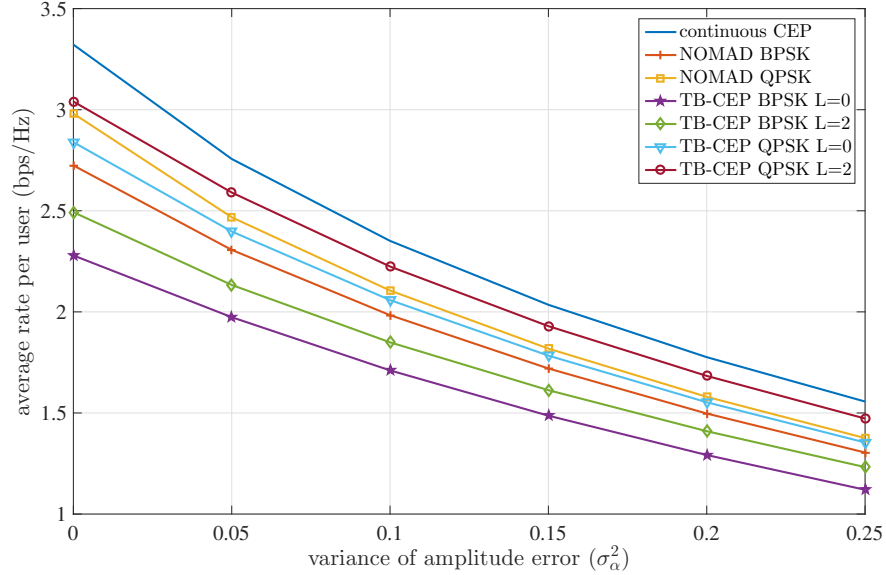


Fig. 11: Average rate per user versus amplitude error in RF phase shifters.

TB-CEP with $L = 0$ and $L = 2$ achieve about 73% and 80% of the performance of the continuous CEP, respectively. These percentages increase to about 85% and 92% for $N = 100$.

The average rate per user versus the number of users (K) is plotted in Fig. 5. As K increases, while N is kept constant, the average rate per user decreases. This is because of the fact that by increasing K , the per-user interuser interference increases. As it is shown in Fig. 5, the newly

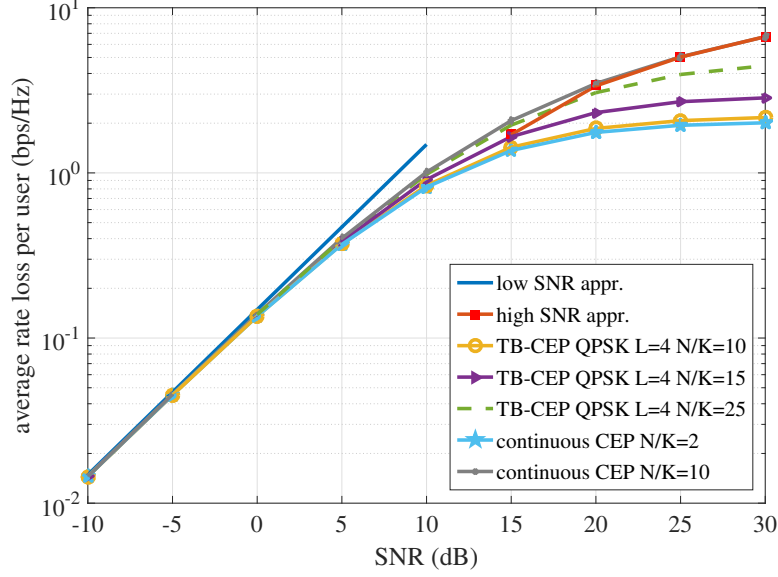


Fig. 12: Average rate loss per user due to phase and amplitude errors versus SNR.

proposed algorithm again offers a significantly better performance than NOMAD.

Since K and N are fixed in Fig. 4 and Fig. 5, respectively, we can also view these results as the average rate per user versus the ratio N/K ranging from 1 to 10 and decreased from 100 to 1 in Fig. 4 and Fig. 5, respectively.

In Fig. 6, the average rate per user is plotted as a function of the transmit power for various N/K ratios for $N = 100$. We observe that by increasing the transmit power, the performances of all the algorithms improve until they saturate at a certain point. Also, the performance of all the algorithms improves as the N/K ratio (degree of freedom per user) increases. For instance, 4-ary TB-CEP with $L = 2$ and $P_T = 10$ dBW for $N/K = 1$, $N/K = 4$ and $N/K = 10$ achieves about 33%, 65% and 90% of the rate offered by the continuous CEP, respectively.

In Fig. 7, the average rate per user versus the number of trellis states (S) is plotted, while keeping the constellation size constant. As mentioned earlier, the number of states used in the proposed algorithm is M^L . Clearly, the performance of the proposed algorithm is improved by increasing the number of states. For instance, in the case of 4-ary phase shifters, the rate relative to that of continuous CEP increases from 85% for $L = 0$ to 95% for $L = 5$, surpassing that of NOMAD with the same constellation size.

The average rate per user versus the constellation size is presented in Fig. 8 for various N/K ratios for $N = 100$. For NOMAD, only the range in which it gives acceptable results is plotted. We observe that, as the constellation size increases, the performances of all the algorithms improve until they rapidly saturate. For instance, for 4-ary phase shifters with $N/K = 10$, TB-CEP with $L = 0$, $L = 1$ and $L = 2$ achieve about 85%, 89% and 91% of the rate performance of the continuous CEP, respectively. Also, as expected, we obtain a better performance as N/K increases. For example, 4-ary TB-CEP with $L = 2$ achieves about 45%, 70% and 90% of the rate performance of the continuous CEP for $N/K = 2$, $N/K = 4$ and $N/K = 10$, respectively.

In Fig. 9, the average rate per user is plotted versus the channel estimation SNR, in order to study the effect of errors in channel estimation on the performance of the algorithms. The robust versions of the algorithms are based on the results in Subsection IV.C. We define the channel estimation SNR as $E[\|\mathbf{H}\|_F^2] / E[\|\tilde{\mathbf{H}}\|_F^2]$, where $\|\cdot\|_F$ is the Frobenius norm. We observe that the performance of all the precoding algorithms improve as the channel estimation SNR increases. The proposed robust TB-CEP has a better performance than TB-CEP especially in moderate values of the channel estimation SNRs.

In Fig. 10 and Fig. 11, the average rate per user is plotted versus the phase and amplitude error variances in the RF phase shifters, respectively. We clearly see that while the performance degrades for all the algorithms as the imperfections in the RF phase shifters increase, the proposed algorithm is fairly robust, in the sense that it shows the same level of degradation as the continuous-phase algorithm of [4].

Finally, as a verification of the result in Theorem 2, the average rate loss per user due to phase and amplitude errors versus SNR is plotted in Fig. 12 for $N = 100$. The phase and the amplitude error variances are assumed to be $\sigma_\varepsilon^2 = 5 \text{ deg}^2$ and $\sigma_\alpha^2 = 0.1$, respectively. It should be noted that the expressions in Theorem 2 are obtained under the assumption that a good CEP precoder minimizes the objective function well enough for $K \ll N$ to make the interference power at each user much smaller than the noise power. As it can be seen in Fig. 12, we obtain better approximations of the rate loss due to the phase and amplitude errors for both low and high SNR regimes as N/K increases (i.e., when the approximation is more accurate).

TABLE I: COMPLEXITY-PERFORMANCE TRADE-OFF of TB-CEP

	$L = 0$		$L = 1$		$L = 2$		$L = 3$	
$M = 2$	$t_c = 1\%$	$r_c = 71\%$	$t_c = 80\%$	$r_c = 72\%$	$t_c = 156\%$	$r_c = 76\%$	$t_c = 286\%$	$r_c = 78\%$
	$t_d = 0.5\%$	$r_d = 85\%$	$t_d = 39\%$	$r_d = 86\%$	$t_d = 76\%$	$r_d = 90\%$	$t_d = 139\%$	$r_d = 94\%$
$M = 4$	$t_c = 0.9\%$	$r_c = 86\%$	$t_c = 29\%$	$r_c = 87\%$	$t_c = 1147\%$	$r_c = 91\%$	$t_c = 4644\%$	$r_c = 93\%$
	$t_d = 0.003\%$	$r_d = 94\%$	$t_d = 1\%$	$r_d = 95\%$	$t_d = 4\%$	$r_d = 100\%$	$t_d = 17\%$	$r_d = 102\%$

B. Computational Complexity

The complexity-performance trade-off of the proposed algorithm is presented in Table I, where t_c and t_d are the relative computation times of TB-CEP with respect to continuous CEP [4] and NOMAD [13], respectively. Also, r_c and r_d are the relative achievable rates with respect to the continuous-phase CEP [4] and NOMAD [13], respectively.

According to Table I, by increasing the trellis memory (or equivalently the number of states), the relative performance of the proposed algorithm increases, with the cost of increasing its relative computational time. This trade-off can help us to select a practical trellis memory as a design parameter. For example, for $M = 2$ with $L = 0$, we achieve 85% of the rate of NOMAD, with only 0.5% of its computation time. For this specific example, the relative computation time increases while the relative performance does not change significantly with increasing L , making $L = 0$ a good choice for both cases of $M = 2$ and $M = 4$.

VIII. CONCLUSIONS

In this paper, we formulate and solve the discrete-phase CEP problem for massive MIMO systems. We utilize the sum of interference powers as the objective function and reformulate it as a sum of real terms. Then, by appropriately defining states, we propose a low-complexity trellis-based algorithm to solve the reformulated optimization problem. Next, we extend our algorithm to be robust against channel estimation errors. Based on extensive numerical results, we conclude that the performance of the proposed algorithm is comparable to the conventional ones (which have exponential complexity), and in some cases it even surpasses them, especially

in the massive MIMO scenario. Also, compared to the continuous-phase algorithm, as an upper-bound on performance, the proposed algorithm shows only a small degradation. Furthermore, the performance loss decreases as the number of antennas is increased. We also provide important analytical results. We prove that by increasing the number of BS antennas, user interference can be guaranteed to be arbitrarily small, and we derive closed-form expressions for rate loss due to errors in the phase and the amplitude of the phase shifters, for both low and high SNR regimes.

APPENDIX

Knowing that ε and α are zero mean with variances σ_ε^2 and σ_α^2 , respectively, the expected value of the interference power in the presence of phase and amplitude errors is

$$\begin{aligned}
& \mathbb{E}_{\varepsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})] \\
&= \mathbb{E}_{h, x, \varepsilon, \alpha} \left[\left| \sqrt{\frac{P_T}{N}} \sum_{i=1}^N h_{k,i} (1 + \alpha_i) e^{j(\theta_i + \varepsilon_i)} - \sqrt{E_k} x_k \right|^2 \right] \\
&= \mathbb{E}_{h, x, \varepsilon, \alpha} \left[\frac{P_T}{N} \sum_{i=1}^N (1 + \alpha_i)^2 |h_{k,i}|^2 + \frac{P_T}{N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N h_{k,i} h_{k,j}^* (1 + \alpha_i) (1 + \alpha_j) e^{j(\varepsilon_i - \varepsilon_j)} e^{j(\theta_i - \theta_j)} \right. \\
&\quad \left. - 2\sqrt{\frac{P_T E_k}{N}} \operatorname{Re} \left\{ \sum_{i=1}^N h_{k,i} (1 + \alpha_i) e^{j(\theta_i + \varepsilon_i)} x_k^* \right\} + E_k |x_k|^2 \right]. \tag{33}
\end{aligned}$$

Since $\mathbb{E}_{\varepsilon, \alpha} [f(\varepsilon, \alpha)] = \mathbb{E}_\varepsilon [\mathbb{E}_{\alpha|\varepsilon} [f(\varepsilon, \alpha)]]$, we have

$$\begin{aligned}
& \mathbb{E}_{\varepsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})] \\
&= \mathbb{E}_{h, x, \varepsilon} \left[\frac{P_T}{N} \sum_{i=1}^N (1 + \sigma_\alpha^2) |h_{k,i}|^2 + \frac{P_T}{N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N h_{k,i} h_{k,j}^* e^{j(\varepsilon_i - \varepsilon_j)} e^{j(\theta_i - \theta_j)} \right. \\
&\quad \left. - 2\sqrt{\frac{P_T E_k}{N}} \operatorname{Re} \left\{ \sum_{i=1}^N h_{k,i} e^{j(\theta_i + \varepsilon_i)} x_k^* \right\} + E_k |x_k|^2 \right] \\
&= \mathbb{E}_{h, x} \left[\varphi_E(-1) \varphi_E(1) \left| \sqrt{\frac{P_T}{N}} \mathbf{h}_k e^{j\boldsymbol{\theta}} - \sqrt{E_k} x_k \right|^2 + \frac{P_T}{N} (1 + \sigma_\alpha^2 - \varphi_E(-1) \varphi_E(1)) \sum_{i=1}^N |h_{k,i}|^2 \right. \\
&\quad \left. + (1 - \varphi_E(-1) \varphi_E(1)) E_k |x_k|^2 - 2\sqrt{\frac{P_T E_k}{N}} (\varphi_E(1) - \varphi_E(-1) \varphi_E(1)) \operatorname{Re} \{ \mathbf{h}_k e^{j\boldsymbol{\theta}} x_k^* \} \right], \tag{34}
\end{aligned}$$

where $\varphi_E(t) \triangleq \mathbb{E}_\varepsilon [e^{jt\varepsilon}]$ is the characteristic function of the random variable ε . We assume that the phase error is a zero mean symmetric random variable, i.e., positive and negative phase errors occur with the same probability, which is a practical assumption. The characteristic function of a zero mean symmetric random variable is a real-valued even function [15]. Therefore, assuming

the desired transmitted symbols have unit power and knowing that channel gains are zero mean random variables with unit variances, and that we have $\sqrt{\frac{P_T}{N}} \mathbf{h}_k e^{j\theta} \approx \sqrt{E_k} x_k$ for large number of antennas and $K \ll N$, (34) can be simplified as

$$E_{\epsilon, \alpha} [I_k(\boldsymbol{\theta} + \boldsymbol{\epsilon}, \boldsymbol{\alpha})] \approx P_T ((1 + \rho_k) + \sigma_\alpha^2 - (1 + \rho_k) \varphi_E^2(1)), \quad (35)$$

where $\rho_k \triangleq \frac{E_k}{P_T}$.

REFERENCES

- [1] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Trans. on Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.
- [2] T. L. Marzetta, "Massive MIMO: An Introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11-22, 2015.
- [3] E. Larsson, O. Edfors, F. Tufvesson and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186-195, Feb. 2014.
- [4] S. K. Mohammed and E. G. Larsson, "Per-Antenna Constant Envelope Precoding for Large Multi-User MIMO Systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059-1071, Mar. 2013.
- [5] S. C. Cripps, *RF Power Amplifiers for Wireless Communications.*, 2nd ed. Norwood, MA, USA: Artech House, 2006.
- [6] S. K. Mohammed and E. G. Larsson, "Single-User Beamforming in Large-Scale MISO Systems with Per-Antenna Constant-Envelope Constraints: The Doughnut Channel," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 11, pp. 3992-4005, Nov. 2012.
- [7] S. K. Mohammed and E. G. Larsson, "Constant-Envelope Multi-User Precoding for Frequency-Selective Massive MIMO Systems," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 547-550, Oct. 2013.
- [8] Jung-Chieh Chen, Chao-Kai Wen and Kai-Kit Wong, "Improved Constant Envelope Multiuser Precoding for Massive MIMO Systems," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1311-1314, Aug. 2014.
- [9] S. Mukherjee and S. K. Mohammed, "Constant-Envelope Precoding With Time-Variation Constraint on the Transmitted Phase Angles," *IEEE Wireless Commun. Lett.*, vol. 4, no. 2, pp. 221-224, Apr. 2015.
- [10] Microwaves101, "Phase Shifters," Oct. 2015; <http://www.microwaves101.com/encyclopedias/phase-shifters>
- [11] R. G. Parker and R. L. Rardin, *Discrete Optimization*. San Diego, CA, USA: Academic Press, 1988.
- [12] V. S. Varadarajan, "A useful convergence theorem," *Sankhyā: The Indian Journal of Statistics*, vol. 20, pp. 221-222, 1958.
- [13] S. Le Digabel, "Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm," *ACM Transactions on Mathematical Software (TOMS)*, vol. 37, no. 4, article 44, Feb. 2011.
- [14] P. Billingsley, *Probability and Measure.*, 4th ed. Hoboken, NJ, USA: John Wiley and Sons, 2012.
- [15] R. M. Dudley, *Real Analysis and Probability*. Cambridge, UK: Cambridge Univ. Press, 2002.

- [16] J.E. Mitchell, "Branch-and-cut algorithms for combinatorial optimization problems," *Handbook of Applied Optimization*, pp. 65-77, 2002.
- [17] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 742-758, Oct. 2014.
- [18] E. Björnson, E. G. Larsson and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE J. Sel. Topics Signal Process.*, vol. 54, no. 2, pp. 114-123, Feb. 2016.
- [19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge Univ. Press, 2004.
- [20] M. F. Hanif, L. N. Tran, A. Tölli and M. Juntti, "Computationally Efficient Robust Beamforming for SINR Balancing in Multicell Downlink With Applications to Large Antenna Array Systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1908-1920, June 2014.
- [21] Z. q. Luo, W. k. Ma, A. M. c. So, Y. Ye and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20-34, May 2010.
- [22] Wing-Kin Ma, Pak-Chung Ching and Zhi Ding, "Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2862-2872, Oct. 2004.