Benefits of Coding in a Noncoherent Massive SIMO System

Brian Knott, Mainak Chowdhury, Alexandros Manolakos and Andrea J. Goldsmith
Department of Electrical Engineering, Stanford University, Stanford, CA
Email: {knottb, mainakch, amanolak, andreag}@stanford.edu

Abstract—We consider one single antenna transmitter communicating with a receiver with a large number of antennas. Motivated by the optimal noncoherent detector in a Rayleigh fading channel, we propose a noncoherent energy-based communication scheme that does not require knowledge of instantaneous CSI (channel state information) at either the transmitter or the receiver; it uses only the statistics of the channel and noise. We explore the impact of coding to reduce the number of antennas needed for this system to achieve a given performance target. In particular, random coding error exponents for this system are used to determine tradeoff curves between the number of antennas and the blocklengths associated with a guaranteed performance target. However, since random codes have exponentially increasing decoding complexity with increasing blocklength, we also consider a simplified codebook design that has significantly lower encoding and decoding complexity. Simulations suggest that for small blocklengths, the performance of this simplified codebook is competitive with random coding constructions.

Index Terms—Massive MIMO, Noncoherent communication, Error exponents, Coding

I. INTRODUCTION

Large antenna systems are likely to be a fundamental part of future cellular standards [1], [2]. This is not only due to the favorable beamforming properties of large antenna arrays, but also to the fact that in time-varying channels, large antenna arrays offer diversity-combining benefits. While some of these properties are conditioned on the assumption of accurate channel state information [3], there has also been work which examines imperfect channel state information or investigates system-level performance benefits that can be obtained from long-term or outdated information [4], [5], [6], [7]. A surprising outcome of this line of work is the fact that a long term statistical characterization of the channels is often sufficient for reaping the diversity-combining or beamforming benefits from such systems. For example, in [8] it was shown that a simple noncoherent energy-based SIMO system achieves the same scaling of achievable rates as a coherent system with an asymptotically large number of antennas, and [7] presents a constellation design which takes into account imperfect knowledge of large-scale statistics at the transmitter and receiver. This motivates the study of noncoherent large antenna systems considering the potential of these systems to simplify both the physical layer and higher layer designs.

In this work we consider a noncoherent SIMO communication system and investigate coding across blocklengths for a large but finite number of receive antennas. We assume fixed rate transmission and consider achievable error exponents with increasing blocklength. Through random codebook and structured codebook constructions, we investigate the potential of coding to reduce the number of antennas needed to achieve the same error exponent. The presented random coding bounds suggest that a given performance target can be achieved with significantly fewer antennas as blocklength increases. In addition, motivated by encoding and decoding complexity requirements, we consider a simplified structured codebook design. While the error exponent achievable with this scheme is provably strictly inferior to that of random codebooks for large blocklengths, the performance for a small blocklength is competitive with the random coding performance.

The study of Gallager type random coding bounds [9], and of low complexity encoding and decoding schemes achieving a positive error exponent, has a rich history [10], [11]. While random coding achieves a positive error exponent [9], the search for low complexity decodable codebooks which achieve linear error exponents with respect to the blocklength remains an ongoing line of research [12], [13]. In particular, code constructions based on serial concatenated codes can achieve good error exponents with low complexity in certain channels [14]. In this work, we investigate how much coding can help us reduce the number of receive antennas needed in a noncoherent SIMO communication system.

The rest of the paper is organized as follows. We present the system model in Section II, followed by the random coding bound in Section III. The simplified codebook design is described in Section IV along with an upper bound on the achievable error exponents. We conclude with simulation results for both a random codebook and an optimized structured codebook in Section V.

II. SYSTEM MODEL

We consider a single user transmitter with one antenna communicating with a receiver with $n$ antennas over a Rayleigh fading channel with an average power constraint of 1 unit (Fig. 1). We assume the fading has a coherence time of one time slot and that the transmitter codes over $T$ time slots using $2^T$ codewords, i.e., it uses a blocklength of $T$ and a rate of 1 bit per time slot. Moreover, we consider an energy-based noncoherent communication scheme, initially presented in [7], [15], in which the receiver calculates only the average energy across its antennas in each time slot. Since only the average received energy is measured at each time slot (and the channel statistics are complex Gaussian), we consider the codebook to
be made of entries from positive reals ($\mathbf{R}_+$). Specifically, if the actual channel between the transmitter and the receiver antennas at time slot $i$ is $h^{(i)}$ the system can be written as
\[ \hat{y}^{(i)} = h^{(i)}x_i + \nu^{(i)}, \]
where $\hat{y}^{(i)} \in \mathbb{C}^n$ is the output of the channel, $h^{(i)} \in \mathbb{C}^n$ is the channel vector ($h_i^{(i)}$ assumed to be $\mathcal{CN}(0, \sigma^2)$), $x_i$ is the input to the channel and $\nu^{(i)}$ is the receiver noise, with $\nu_i^{(i)} \sim \mathcal{CN}(0, \sigma^2)$.

The transmitter chooses a codebook $\mathcal{C}$ of size $2^T$, and the decoder bases its decision on $y$ such that
\[ y_i = \|\hat{y}^{(i)}\|^2/n, \quad \forall i. \]

In particular, we consider the following:

A. Encoder

If $x \in \mathcal{C}$ corresponds to the message that the transmitter wants to send in a block of $T$ slots, the transmitter sends $x_i$ at slot $i$, $1 \leq i \leq T$.

B. Decoder

For every time slot $i$, the receiver looks at the average energy across all antennas, $y_i = \|\hat{y}^{(i)}\|^2/n$. Based on $T$ observations $y = (y_1, y_2, \ldots, y_T)$, it produces an estimate $\hat{x}(y)$. We declare an error iff $x \neq \hat{x}(y)$ when $x$ is transmitted and denote the probability of this event by
\[ P_{\text{error}}(x) \triangleq \text{Prob}(x \neq \hat{x}(y)). \]

To minimize the Bit Error Rate (BER), the decoder should be a maximum likelihood (ML) decoder. In that case
\[ \hat{x}(y) = \arg \max_{x \in \mathcal{C}} \log(p(y|x)) = \arg \max_{x \in \mathcal{C}} \left[ \frac{1}{T} \sum_{i=1}^{T} \log(p(y_i|x_i)) \right], \]
where, for our particular system model [15],
\[ p(y_i|x_i) \triangleq \frac{\gamma^{n-1} e^{-\gamma y_i}}{c^n n^{2n} \Gamma(n)}, \]
with $c_n = \frac{\sigma^2 + n^2 \sigma^2}{2n}$. Even though the ML decoder minimizes the BER, in this work we approximate its performance by considering the following decoder [15]:
\[ \hat{x}(y) = \arg \min_{x \in \mathcal{C}} d(y|x) = \arg \min_{x \in \mathcal{C}} \prod_{i=1}^{T} d(y_i|x_i), \tag{1} \]
where
\[ d(y|x) \triangleq e^{-nD(y|x)} = \prod_{i} e^{-nD(y_i|x_i)}, \]
with
\[ D(y_i|x) = \sup_{\theta > 0} \left( \theta y_i - E_{y_i|x_i}[e^{\theta y_i}] \right), \tag{2} \]
and $D(y|x) \triangleq \sum_i D(y_i|x_i)$, where $\theta$ in (2) is an auxiliary variable over which the supremum is taken. The above decoder is approximately the ML decoder for a large number of antennas, mainly due to the fact that for any $\epsilon > 0$,
\[ \lim_{n \to \infty} \frac{-\log(\text{Prob}(y_i \in [z_i, z_i + \epsilon]))}{n} = D(z_i|x_i), \]
where $\text{Prob}(y_i \in [z_i, z_i + \epsilon])$ is the probability of the received energy in the $i^{th}$ slot lying in an interval around $z_i$. If $z_i = \sigma^2/2 + \sigma^2$, $D(z_i|x_i) = 0$, suggesting that $y_i$ concentrates around $\sigma^2/2 + \sigma^2$ as $n$ becomes larger and larger. Note that unlike $p(y|x)$, $D(y|x)$ is independent of the number of antennas $n$. This will be used in the simplified codebook design in Section IV.

III. RANDOM CODING ERROR EXPONENT

A. General bounds

In this section, we provide a random coding bound for the error exponent with increasing block length $T$ for a fixed but finite number of receive antennas over the channel discussed in Section II.

Consider a codebook generated randomly according to a given distribution $q(x)$ satisfying an average power constraint (i.e., the second moment of the distribution is less than 1). Our analysis depends on the choice of $q(x)$, and in this work we do not optimize over all possible distributions (this question is investigated more in the extended version of this work [8]). We look at the probability of error averaged over all codebooks randomly generated according to the distribution $q(x)$, which yields
\[ \mathbb{E}[P_{\text{error}}(x)] \leq \mathbb{E}_x \left[ \mathbb{E}_{y|x} \left[ \mathbb{E}[P_{\text{error}}(x)] \right] \right] \]
\[ \leq \mathbb{E}_x \left[ \mathbb{E}_{y|x} \left[ \min \left( \mathbb{E} \left[ \sum_{z \neq x, z \in \mathcal{C}} P_{\text{error}}(x \rightarrow z) \right], 1 \right] \right] \right], \]
where $P_{\text{error}}(x \rightarrow z)$ refers to the probability of mistaking $z$ for another codeword $x$ in the codebook $\mathcal{C}$ when $x$ is the transmitted codeword. In the above, the subscripts refer to the variables with which the expectation is taken. Using the fact that $\min(1, b) \leq b^\rho$ for all $0 < \rho < 1$, we get that
\[ \mathbb{E}[P_{\text{error}}(x)] \]
\[ \leq \mathbb{E}_x \left[ \mathbb{E}_{y|x} \left[ \min \left( \mathbb{E} \left[ \sum_{y \neq x} P_{\text{error}}(x \rightarrow y) \right], 1 \right] \right] \right], \]
\[ \leq \mathbb{E}_x \left[ \mathbb{E}_{y|x} \left[ \left( \mathbb{E} \left[ \sum_{z \neq x} P_{\text{error}}(x \rightarrow z) \right] \right)^\rho \right] \right], \]
In this notation, $T$ have lower storage and decoding complexity. In particular, we
In this section we explore alternate coding constructions that
over, the size of the codebook also scales as exponential in
where $f$ is an indicator function, and (a) follows from the fact that
It should be noted that the above approach is the same if
$2^T - 1 = E_x \left[ \left( \int_x q(z) \left( \frac{d(y|x)}{d(y|z)} \right)^s dz \right)^\rho \right]$
\[ \leq (2^T - 1)^{\rho} E_x \left[ \left( \int_x q(z) \left( \frac{d(y|x)}{d(y|z)} \right)^s dz \right)^\rho \right] \]
\forall s > 0
\[ = (2^T - 1)^{\rho} \cdot E_x \left[ \int_x \int_y q(y)p(y|x) \left( \int_x q(z) \left( \frac{d(y|x)}{d(y|z)} \right)^s dz \right)^\rho dydx \right] \leq e^{-TE}, \]
where the last step uses the fact that
\[ p(y|x) = \prod_{i=1}^{T} p(y_i|x_i), \quad q(x) = \prod_{i=1}^{T} q(x_i) \]
and $d(y|x) = \prod_{i=1}^{T} d(y_i|x_i)$. In the above $1$ refers to the indicator function, and (a) follows from the fact that
\[ 1_{z:d(y|z)\leq d(y|x)} \leq \left( \frac{d(y|x)}{d(y|z)} \right)^s \]
for all $s > 0$. The error exponent $E$ can then be seen to be the following:
\[ E \triangleq \max_{s>0} \max_{0 \leq \rho \leq 1} \left( -\log \left( \int_x q(x) \int_y p(y|x) f(x,y)^\rho dy \right) - \rho \log 2 \right), \]
where $f(x,y) = \int_z q(z) \left( \frac{d(y|x)}{d(y|z)} \right)^s dz$.

It should be noted that the above approach is the same if we were using $p(y|x)$ instead of $d(y|x)$. Moreover, for large $n$, using either function leads to the same $E$.

IV. SIMPLIFIED CODEBOOK DESIGN

The random coding construction outlined above has an exponential decoding complexity in the blocklength $T$. Moreover, the size of the codebook also scales as exponential in $T$. In this section we explore alternate coding constructions that have lower storage and decoding complexity. In particular, we consider codebooks of the following form

$C \triangleq \{ x : x_i = a_{|S|}, \forall i \in S, x_j = 0, \forall j \notin S, S \in [n] \}$.

In this notation, $x_i$ stands for the $i^{th}$ coordinate of $x$, $[n] = \{1, 2, \ldots, n\}$, $x \in \mathbb{R}_+^T$, $a \in \mathbb{R}_+^T$ with $a_0 \triangleq 0$ (note that coordinates start from position 1). As an illustration, for $T = 4$, we have

$C = \{ (0, 0, 0, 0), (a_1, 0, 0, 0), (0, a_1, 0, 0), (0, 0, a_1, 0), (0, 0, 0, a_1), (a_2, a_2, 0, 0), (a_2, 0, a_2, 0), (a_2, 0, 0, a_2), (0, a_2, a_2, 0), (0, a_2, 0, a_2), (0, 0, a_2, a_2), (a_3, a_3, a_3, 0), (a_3, 0, a_3, a_3), (a_3, a_3, 0, a_3), (0, a_3, a_3, 0), (a_4, a_4, a_4, a_4) \}$.

Observe that this codebook is completely specified by $T$ parameters, which means that its storage complexity is linear in $T$, i.e., only these $T$ real valued numbers are needed to generate any codeword from this constellation.

A. Encoding

Given that the encoder sends codeword $x$, at time slot $j$, the transmitter sends out $x_j$, i.e., the $j^{th}$ coordinate.

B. Decoding

The decoder is the same as the one defined in Section II-B. However, computing the minimizer of $d(y|x)$ is now greatly simplified due to the codebook structure. In particular, we see that

$\arg \min_{x \in \mathcal{C}} d(y|x) = \arg \min_{x \in \mathcal{C}} \arg \min_{x_j=0, \forall j \notin S, |S|=i} d(y_j|x_j)$.

The inner minimization is over $\binom{T}{i}$ codewords, but only some of the codewords are candidates for the minimizer of $d(y|x)$. These candidate points can be determined just by sorting $y$, with the largest $i$ components of $y$ corresponding to the non-zero components of the candidate point $x$. Each such candidate point will have the best likelihood among all inner $\binom{T}{i}$ possible codewords since $D(y_j|x_j)$ is monotonically increasing with $|y_j - \sigma_j^x y_j^2 - \sigma^2|$.

The inner minimizer can be computed just by sorting $y$ and the outer minimizer requires computing at most $T$ distances. Thus, the decoding complexity is $\Theta(T^2 + T \log T)$.

C. Optimizing the simplified codebook

The aim of our optimization procedure is to find $a \in \mathbb{R}_+^T$ such that the resulting codebook $C$ maximizes the minimum distance $\alpha(a)$ as defined below.

Definition 1. For a given constellation $\mathcal{C}$ parametrized by $a$, we define the function $\alpha(a)$ as the maximum size of the constant-D contours in a constellation before they begin to overlap.

We can find $\alpha(a)$ by solving the optimization problem:

$maximize_{\alpha, \alpha} \alpha$
subject to $\{ z : D(z|s) \leq \alpha \} \cap \{ z : D(z|t) \leq \alpha \} = \emptyset$
$\forall s \neq t; s, t \in \mathcal{C}.$

The optimization variables are $\alpha, \alpha$. The above problem is
nonconvex in general, but we can solve it by solving a sequence of feasibility problems. To determine the feasibility of a given \(\alpha\)-value for a particular choice of \(a\), we consider the following problem for all combinations of points \(\{s,t\}\):

\[
\begin{align*}
\text{minimize}_z & \quad 0 \\
\text{subject to} & \quad D(z|s) \leq \alpha \\
& \quad D(z|t) \leq \alpha.
\end{align*}
\]

(5)

The optimization variable is \(z\). If this problem is feasible for any \(s, t \in \mathcal{C}, s \neq t\), then the current \(\alpha\)-value does not satisfy the constraint \(\{z : D(z|t) \leq \alpha\} \cap \{z : D(z|s) \leq \alpha\} = \emptyset\) for all \(s \neq t\). Therefore, the chosen \(\alpha\)-value becomes an upper bound on \(\alpha(a)\). To solve this feasibility problem, we consider its dual:

\[
\begin{align*}
\text{maximize}_{\lambda_1, \lambda_2} & \quad \inf_z (\lambda_1 (D(z|s) - \alpha) + \lambda_2 (D(z|t) - \alpha)) \\
\text{subject to} & \quad \lambda_1, \lambda_2 \geq 0.
\end{align*}
\]

where the optimization variables are \(\lambda_1, \lambda_2\). If the intersection problem is infeasible, this dual problem will be unbounded. Otherwise, the infimum in the objective function will cause it to be non-positive, forcing the maximum objective value over non-negative \(\lambda\)'s to be exactly zero. Therefore, to show feasibility, it is sufficient to find any \(\lambda_1, \lambda_2\) for which the infimum is positive. This lets us solve the optimization problem by using a bisection search on \(\alpha\) together with solving multiple feasibility problems (5). Because of the symmetry of the simplified codebook, only \(O(T^2)\) pairs need to be considered to calculate \(\alpha(a)\) rather than \(O(2^T)\).

Now that we can specify \(\alpha(a)\) for any constellation \(\mathcal{C}\) by solving (4), we optimize \(\alpha(a)\) across all possible \(a\) by solving the following problem:

\[
\begin{align*}
\text{maximize}_a & \quad \alpha(a) \\
\text{subject to} & \quad \sum_{i=1}^T \frac{1}{2^T} \alpha_i^2 \leq T.
\end{align*}
\]

(6)

This is a non-convex problem, and can only be sub-optimally solved using global optimization routines.

D. Performance analysis

We show that although the simplified codebook construction has good decoding complexity and reasonable storage, we do lose out on the linearity of the error exponent with respect to the blocklength \(T\). In particular, we will show that the error exponent is at most \(\Theta(\sqrt{T})\) with respect to the blocklength. In particular we have the following lemma.

**Lemma 1.** The error exponent of the simplified codebook construction scales at most like \(\Theta(\sqrt{T})\).

**Proof.** Recall that the simplified codebook construction has a rate of 1 bit per time slot. To show that the achieved error exponent with this construction scales at most like \(\Theta(\sqrt{T})\), we consider a slightly different codebook construction with a better minimum distance but lower rate than 1 bit per time slot. Since the error exponent of the codebook is a monotonically increasing function of the minimum distance associated with the codebook, we get that the error exponent of this construction is an upper bound on what is achievable using the simplified codebook construction.

In particular, consider a reduced-rate codebook construction with only the codewords corresponding to \(T/2\) nonzeros (with \(T\) even), and the other codewords set to all zeros. For this codebook, assuming equiprobable signaling, the average power constraint is

\[
1 \leq \frac{T \alpha^2}{2T} \left( \frac{T}{T/2} \right) \leq T,
\]

since there are still \(2^T\) codewords, with \(\left(\frac{T}{T/2}\right)\) of these having \(\frac{T}{2}\) non-zeros entries equal to \(\alpha \sqrt{T}/2\), and the rest are 0. We see that since \(\left(\frac{T}{T/2}\right) \approx k \frac{2}{\sqrt{T}}\) for some constant \(k\) and for a large \(T\), eventually, for \(T\) large enough, \(\alpha^2 \frac{T}{2} \leq \frac{k}{2} \sqrt{T}\), which is sublinear in \(T\). The minimum distance of this modified code construction can be seen to be at most equal to \(\alpha^2 \frac{T}{2}\) as for \(z\) large enough the error exponent (2) associated with a large \(n\) is linear in \(z\) [15].

Thus, although the decoding complexity for the codebook construction outlined in the previous section is low, it does not exhibit the linear increase in the error exponents with increasing \(T\). However, Section V compares the performance of this construction with the random coding bounds for small blocklengths \(T\) and we see that it can achieve a higher error exponent compared to random coding constructions for small blocklengths.

V. Numerical performance

In this section we study the relation between the number of antennas and the blocklength needed to achieve a specific target performance. To do so, in Subsection V-A we plot the error exponents as a function of \(n\) and \(T\) for the random codebook construction and compare them with those achieved with the simplified codebook construction. In the error exponent of the random coding we use as \(q(x)\) a truncated geometric distribution. Subsection V-B focuses on the BER performance when using the optimized simplified codebook construction for small \(T\) and \(n\) and shows the decrease in the number of antennas needed to achieve a BER target with increasing \(T\).
T gives a better error exponent than a system with double baseline system to a system with twice as large a blocklength than from adding more antennas. For example, going from the gains from moving to larger blocklengths is generally higher = figures: First, at SNR = \( \frac{\log(P_{error})}{n} \), for SNR={−3,0,5} dB, where \( \log(P_{error}) \) is the bound for the baseline system with \( n_0 = 20, T_0 = 2 \).

### A. Tradeoffs between \( n \) and \( T \)

Fig. 3 demonstrates the tradeoff between the number of antennas and the blocklength required in systems with a given BER performance for low and high SNR (SNR={−3,0,5} dB). Specifically, using as a baseline the BER of a system with \( n_0 = 20 \) antennas and \( T_0 = 2 \) blocklength, denoted as \( P_{error}^0 \), we plot the factor increase in the error exponent, i.e., \( \frac{\log(P_{error})}{\log(P_{error}^0)} \), as a function of \( \frac{n}{n_0} \) and \( \frac{T}{T_0} \) of a system with \( n \) antennas, blocklength \( T \) and BER \( P_{error} \). Two important qualitative conclusions can be made by observing these figures: First, at SNR={0,5}, the incremental performance gains from moving to larger blocklengths is generally higher than from adding more antennas. For example, going from the baseline system to a system with twice as large a blocklength \( T \) gives a better error exponent than a system with double the number of antennas. Second, at low SNR values, e.g., SNR= −3, both the number of antennas and the blocklength seem to have equal contribution.

Fig. 4 presents the error exponent of the random codebook as a function of the number of antennas. Specifically, we see that for the range of antennas \( n \) we are interested in, e.g., \( n < 500 \), the gains of adding more antennas is more, per antenna added, and then it becomes constant for \( n > 10000 \). Also, we observe that increasing the SNR provides diminishing returns on the error exponent, which is expected since, at high SNR, the error exponent is mostly affected by the fading of the channel and not the additive noise.

### B. Small \( n \) and/or \( T \) behavior

In this section we compare the achievable error exponents for both the random codebook and the simplified codebook design. Specifically, Fig. 5 shows that, compared to the random
codebook design, the simplified codebook achieves better error exponents with increasing number of antennas for small values of $T$ for SNR = $\{-3, 0, 5\}$ dB. Also, observe that at high SNR, the simplified constructions perform better than the random codebook constructions over a larger range of blocklengths.

Fig. 6a, 6b present Monte Carlo simulation plots for the BER performance of the simplified constellation design. We see that increasing the blocklength provides a non-trivial improvement in the BER performance for a small number of antennas at the receiver even by using very low $T$.

Finally, Fig. 7 presents the percentage increase of the minimum distance for the optimized simplified codebook design, for each SNR = $\{-3, 0\}$ dB, where each curve is normalized by the value of the minimum distance when $T = 1$ (the minimum distance is equal to the error exponent normalized by the number of receiver antennas when the latter is asymptotically large).

VI. CONCLUSION

We investigate the impact of coding in a massive antenna array deployment without instantaneous channel state information. We consider a single transmitter sending information to a multiple antenna receiver and consider the error exponents achievable with finite constellations and coding across finite blocklengths. Our random coding bounds suggest that with increasing blocklengths, we can significantly reduce the number of antennas needed to meet a certain performance target. Furthermore, for small blocklengths, we propose a structured codebook design which is competitive with (and under some conditions, even exceeds) the performance with random codes.

ACKNOWLEDGEMENTS

This work was supported by a 3Com Corporation Stanford Graduate Fellowship (SGF), an Arvanitidis SGF in Memory of William K. Linvill, an A.G. Leventis Foundation Scholarship, NSF grant 1320628 and ONR grant N000141210063.

REFERENCES