Abstract—This paper studies noncoherent wideband systems with a single antenna transmitter and a multiple antenna receiver with many elements, under signaling with peak-to-average power ratio constraints. The analysis considers the scaling behavior of capacity and achievable rates by letting both the number of antennas and the bandwidth go to infinity jointly. In contrast to prior work on wideband single input single output (SISO) channels without a-priori channel state information, it is shown that a sufficiently large number of receive antennas can make up for the vanishingly small SNR at each antenna. In particular, it is shown that when bandwidth grows sufficiently slowly with the number of antennas, the capacity scaling with an increasing number of receive antennas is the same as the optimal coherent capacity scaling. If the bandwidth grows faster than a certain threshold, however, the additional bandwidth does not help because a finite transmit power is spread over an excessively large bandwidth.

Index Terms—Massive MIMO, wideband system, noncoherent Communications, Energy Receiver

I. INTRODUCTION

Large antenna arrays and high carrier frequencies look increasingly indispensable for future generation high-capacity cellular networks, and are being strongly considered for inclusion in mmWave-based 5G cellular standards [1], [2]. While such technologies may help in alleviating problems of network capacity, they bring their own set of challenges. In particular, a short channel coherence time associated with wavelengths on the order of millimeters, together with the large number of channel coefficients in multi-antenna transceivers, make it challenging to accurately track instantaneous receiver Channel State Information (CSI), as well as to feedback accurate CSI to the transmitters.

Channels without instantaneous CSI at the receiver and transmitter are referred to as noncoherent fading channels. In the wideband regime, for a fixed number of antennas, a typical strategy is to spread the available power across multiple frequency bands. However, as shown by Medard and Gallager [3], achievable total rates with “non-peaky” signals have to necessarily go to zero as the power per unit bandwidth becomes smaller and smaller. Characterizations of the low SNR noncoherent capacity for both single and multiple antennas have also been considered in [4]. These results show that noncoherent capacity with non-peaky signaling is a quadratic function of the signal to noise ratio rather than linear, and hence overspreading over a wide bandwidth is detrimental from a capacity viewpoint. On the other hand, implementation of a perfectly peaky signaling, such as the one described in [5], requires concentrating all the available power in a small “flash” interval of a very large instantaneous power. Hardware nonlinearities introduce significant distortion for signals of this nature; hence such signaling is not used in practice.

An alternate strategy when communicating with non-peaky signals is to not transmit in all the available bandwidth. In this case, Lozano and Porrat [6] show that the achievable rate is maximized at the critical bandwidth. Moreover, at the critical bandwidth, rate lies within a constant gap from the power-limited wideband capacity with perfect CSI. This means that with bandwidth above this value, one need not use all available bandwidth for optimal rate scaling.

The results obtained in [3], [5], [6] assume a fixed number of antennas, and do not necessarily extend to broadband systems with an asymptotically large numbers of antennas as in a massive Multi-Input Multi-Output (MIMO) communication system. Motivated by the emergence of massive MIMO systems [2], we consider a wideband massive Single-Input Multiple-Output (SIMO) non-coherent fading channel and study the effect of the joint scaling of bandwidth and the number of antennas on the achievable rates. We investigate a capacity scaling upper bound to SIMO systems with both the size of the large receive antenna arrays and the signal bandwidth growing asymptotically large simultaneously. We also provide a practical encoding scheme that achieves this upper bound.

Our analysis suggests a critical scaling of the bandwidth with the number of receive antennas. This scaling is characterized by the following:

- When the bandwidth is smaller than the critical value in a scaling law sense, the achievable rate is bandwidth-limited and grows with bandwidth. A practical strategy that achieves this scaling can be obtained using a multitone generalization of the narrowband scheme in [7], [8].
- When bandwidth scales faster than the critical value, the additional bandwidth does not help in increasing the achievable rates. The optimal achievable rates are obtained by restricting transmission to the critical value of bandwidth.

In traditional wideband analysis with a fixed number of
antennas there is a complete equivalence between the concepts of large bandwidth, small SNR, and capacity being power-limited. However, our results show that when the number of antennas grows much faster than bandwidth, there is a new type of wideband operating regime where a distinct type of bandwidth-limited scaling of capacity takes place. To the best of our knowledge, this regime has never been previously discussed in the literature.

This paper is structured as follows: Section II describes the system model and Section III summarizes the main results. Section IV presents the upper bound on the scaling law of the achieved rate and Section V contains the practical scheme that achieves this capacity scaling. We summarize our results and discuss future work in Section VI.

A. Notation

We use the bold notation \( \mathbf{a} \) to denote a vector and \( a_i \) to denote the \( i \)th coordinate of a vector \( \mathbf{a} \). We use \( a_n = \Theta(b_n) \) to denote that there exist constants \( c_1, c_2, n_0 > 0 \) such that for all \( n > n_0 \), \( c_1 b_n < a_n < c_2 b_n \), and \( a_n = o(b_n) \) to denote that \( \lim_{n \to \infty} \frac{a_n}{b_n} = 0 \). We use \( a_n \lesssim b_n \) to denote \( a_n \leq b_n + o(b_n) \).

II. SYSTEM MODELS

A. Channel

We consider a rich scattering, frequency selective, block fading, SIMO wideband channel with a single-antenna transmitter and \( n \) receiver antennas. We consider a coherence time of \( T_c \) seconds, coherence bandwidth of \( B \). Hertz and transmission over a bandwidth \( B \) Hertz in baseband. The noise power spectral density is \( \frac{N_0}{2} \) Watts/Hertz and the average transmit power is \( P \) Watts. We assume a noncoherent channel, i.e., the channel is unknown both at the transmitter and at the receiver. The transmitted signal during one independently-encoded transmission interval of duration \( T_c \) may be fully determined by a DFT with \( K = \lceil T_c B \rceil \) points. At each receive antenna, the received signal experiences frequency selective fading where each of these \( K \) points are separate “frequency bins” with a scalar fading gain. Typically, the quantity \( K/B = T_c > 1 \) and only a fraction \( \frac{K}{B} \) of these channel gains are uncorrelated. However, in this paper, we assume that the channel gains are independent both across frequency bins and across receiver antennas. As we point out later, our achievable scheme and one of our upper bounds are valid even for correlated channel gains.

The received signal for the \( i \)th antenna and the \( j \)th frequency band can be written as

\[
y_i^{(j)} = \sqrt{\frac{PT}{N_0}} h_i^{(j)} x_j + \nu_i^{(j)},
\]

where \( x_j \) is the transmitted symbol at the \( j \)th frequency band, the additive white Gaussian noise is \( \nu_i^{(j)} \sim \mathcal{CN}(0,1) \), \( h_i^{(j)} \sim \mathcal{CN}(0,1) \) corresponds to a Rayleigh fading distribution, and \( x \) satisfies the power constraint \( \mathbb{E}[\|x\|^2] \leq 1 \). Without loss of generality, we choose units such that \( \frac{PT}{N_0} = 1 \), giving

\[
y_i^{(j)} = h_i^{(j)} x_j + \nu_i^{(j)}.
\]

We define the quantity computed at the receiver

\[
y_j \triangleq \frac{\| \tilde{y}^{(j)} \|^2}{n},
\]

to be the average received power collected for frequency band \( j \). We define \( \tilde{Y} \) and \( \mathbf{H} \) to be the collections \( \{ \tilde{y}^{(j)} \}_{j=1}^K \) and \( \{ h^{(j)} \}_{j=1}^K \) respectively.

B. Signals

We assume that in each subband, we use signaling with a PAPR bounded above by \( S \in \mathbb{R}^+ \). This implies that if \( x_j \) is distributed according to \( p(x_j) \), then

\[
p(x_j) = 0 \text{ for all } |x_j| > S \sqrt{\mathbb{E}[|x|^2]}.
\]

Although in this work we derive our upper bounds based on this assumption (which is just a specific class of non-peaky signals as described in [3], [5], [6]), we believe that this assumption is not critical to our derivation and that the scaling laws of Theorem 1 can be generalized to all (not necessarily non-peaky) signaling methods.

C. Scaling behavior of bandwidth with number of antennas

We wish to analyze the scaling behavior of rate as \( n \to \infty \) and \( B \to \infty \) (with \( T_c, T_d \) fixed). Intuitively, increasing \( n \) increases the total energy captured by the receiver, allowing rate to grow, while \( B \to \infty \) spreads the same energy over a large bandwidth, thereby decreasing the SNR in each frequency bin. The typical wideband capacity with a fixed number of antennas is thus power-limited, but the typical massive SIMO system with fixed bandwidth is bandwidth-limited (due to a large total power collected across all the antennas for a fixed bandwidth). Hence, to study the joint scaling behavior, we allow for the following scaling between the number of receive antennas \( n \) and the available bandwidth \( B \), which is captured by the parameter \( \epsilon \) defined as

\[
\epsilon \triangleq \lim_{B,n \to \infty} \frac{\log(B)}{\log(n)}.
\]

As we will show in the subsequent sections, there exists a threshold \( \epsilon_{th} \) below which the system is bandwidth-limited (so that increasing bandwidth will increase the achievable rates). However, in case \( \epsilon > \epsilon_{th} \), the additional bandwidth does not increase capacity.

III. MAIN RESULT

Under the model in Section II, we have the following result.

Theorem 1. The capacity with finite PAPR in a noncoherent massive SIMO wideband channel with \( n \) antennas at the receiver and bandwidth \( B = n^\epsilon \) scales for large \( n \) as

\[
C(n) = \begin{cases} 
\Theta(n^\epsilon \log(n)) & \epsilon \in \left[0, \frac{1}{2}\right) \\
o \left(n^{\frac{\epsilon}{2} + \alpha}\right) & \epsilon > \frac{1}{2} \text{ for all } \alpha > 0.
\end{cases}
\]

Remark 1. We do not consider \( \epsilon = \frac{1}{2} \) although our arguments can be generalized to accommodate that.

Remark 2. We can achieve the optimal capacity scaling with simple energy-detection based encoders and decoders when \( \epsilon < \epsilon_{th} \triangleq \frac{1}{2} \). For \( \epsilon > \frac{1}{2} \), we can use a subset of the bandwidth.
The capacity with finite PAPR in a noncoherent massive SIMO wideband channel with \( n \) antennas at the receiver and bandwidth \( B = n' \), is upper bounded by

\[
\text{C}(n) \leq o(n^{\frac{1}{2}+\alpha}) \quad \text{for all } \alpha > 0.
\]

**Lemma 3.** For \( \epsilon > \epsilon_b = \frac{1}{2} \), the capacity with finite PAPR in a noncoherent massive SIMO wideband channel with \( n \) antennas at the receiver and bandwidth \( B = n' \), is upper bounded by

\[
\text{C}(n) \leq o(n^{\frac{1}{2}+\alpha}) \quad \text{for all } \alpha > 0.
\]

**Proof.** For the following, the PAPR constraints are implicitly assumed for \( p(x) \). For a fixed distribution \( p(x) \), the mutual information can be upper bounded as

\[
I \left( x; \hat{Y} \right) \overset{(a)}{=} I \left( x; y \right) = \max_{p(\mathbf{x}) : \sum_j |y_j|^2 p(x_j) dx_j = 1} E \left[ \log \left( \frac{p(y|x)}{p(y)} \right) \right] \\
\overset{(b)}{=} \max_{p(x)} \sum_{j=1}^{K} \max_{p(x_j)} \left[ y_j^T p(x) dx_j = p \right] E \left[ \log \left( \frac{p(y_j|d_j)}{p(y_j)} \right) \right] \\
\overset{(c)}{\leq} \max_{M \in \{1, \ldots, K \}} M f \left( \frac{1}{M} \right),
\]

where

\[
f(s) = \max_{p(t) : \int \int p(t) dt = s} E \left[ \log \left( \frac{p(y_j|t)}{p(y_j)} \right) \right].
\]

In the above (a) follows from the fact that \( y \) is a sufficient statistic for \( x \) given \( \hat{Y} \). Step (b) follows from the fact that the subbands experience i.i.d. fading and noise processes, so that the optimal rates can be achieved by first allocating power to each subband and then optimizing the distributions for each subband. Step (c) follows from the fact that any optimization problem of the form

\[
\min_{\alpha} \sum_{k=1}^{K} g(u_k) \\
\text{subject to } \sum_{k} u_k = 1 \\
u_k \geq 0 \text{ for all } k,
\]

for a differentiable \( g(\cdot) \) is solved by the following:

\[
\min_{M \in \{1, \ldots, K \}} M g \left( \frac{1}{M} \right).
\]

This may be seen by writing down the necessary KKT conditions for optimality of the optimization problem.

We now show that if \( K = n' \) for \( \epsilon > \frac{1}{2} \), then the maximizing \( M^* \) for the above expression cannot be \( \Theta \left( n^{\frac{1}{2}+\alpha} \right) \) for any \( \alpha > 0 \).
Let us compute \( f \left( \frac{1}{n^{\frac{3}{2} + \alpha}} \right) \) for some \( \alpha > 0 \). Defining \( x_j \equiv n^{\frac{3}{2} + \alpha} x_j \), we see that \( E \left[ |x_j|^2 \right] \leq \frac{1}{n^{\frac{3}{2} + \alpha}} \) is equivalent to \( E \left[ |x_j|^2 \right] \leq 1 \). Expressing \( p(y_j|x_j) \) in terms of \( p(y_j|x_j) \), we get that

\[
f \left( \frac{1}{n^{\frac{3}{2} + \alpha}} \right) = \max_{p(\hat{x}_j):E[|\hat{x}_j|^2] \leq 1} \mathbb{E} \left[ \log \left( \frac{p(y_j|\hat{x}_j)}{p(y_j)} \right) \right],
\]

where

\[
p(y|x) = \frac{y^{n-1} e^{-\frac{ny}{1 + \kappa(n)|x|^2}}}{(1 + \kappa(n)|x|^2)^n},
\]
as it follows from the statistics assumed in Section II, and \( p(y_j) = \int p(y_j|x_j)p(x_j)dx_j \), with \( \kappa(n) = \frac{1}{n^{\frac{3}{2} + \alpha}} \) and \( \Gamma(n) \) the Gamma function.

We now note that for any \( p(\tilde{x}_j) \) satisfying the PAPR constraint and the second moment condition \( E \left[ |\tilde{x}_j|^2 \right] = 1 \),

\[
\mathbb{E} \left[ \log \left( \frac{p(y_j|\tilde{x}_j)}{p(y_j)} \right) \right] \leq \mathbb{E} \left[ \log \left( 1 + \frac{1}{p(y_j|0)} \frac{\partial p(y_j|\tilde{x}_j)}{\partial \kappa(n)} \bigg|_{\kappa(n)=0} \kappa(n) \right) \right]
\]

\[
\leq E_{\tilde{x}_j} \left[ \frac{1}{p(y_j|0)} \frac{\partial p(y_j|\tilde{x}_j)}{\partial \kappa(n)} \bigg|_{\kappa(n)=0} \kappa(n) \right]
\]

\[
\leq \kappa(n)^2 n \mathbb{E}_{\tilde{x}_j} \left[ |\tilde{x}_j|^4 \right] \leq \kappa(n)^2 n S^4.
\]

(10)

In the above step (a) follows from a Taylor series expansion around \( \kappa(n) = 0 \), as well as from the fact that the support of \( |\tilde{x}_j| \) is bounded and that \( \kappa(n) \to 0 \) as \( n \to \infty \), step (b) follows from the fact that \( \log(1+x) \leq x \) for small \( x \) and step (c) follows by taking the expectation with respect to the distribution of \( p(y_j|x_j) \). Plugging in \( \kappa(n) = \frac{1}{n^{\frac{3}{2} + \alpha}} \), we get that

\[
\mathbb{E} \left[ \log \left( \frac{p(y_j|\tilde{x}_j)}{p(y_j)} \right) \right] \leq \Theta \left( \frac{1}{n^{2\alpha}} \right).
\]

(11)

This is a uniform bound on all \( p(\tilde{x}_j) \) satisfying the power and PAPR constraints and, hence, it follows that

\[
Mf \left( \frac{1}{M} \right) \leq \Theta \left( n^{\frac{3}{2} - \alpha} \right)
\]

(12)

for \( M = n^{\frac{3}{2} + \alpha} \). Thus, the maximizing \( M \) (let’s call this \( M^* \)) for the problem in (7) cannot be \( \Theta \left( n^{\frac{3}{2} + \alpha} \right) \). This implies \( M^* = o \left( n^{\frac{3}{2} + \alpha} \right) \) for all \( \alpha > 0 \), thus proving the lemma.

Figure 2a illustrates the effect of choosing \( M = \Theta \left( \frac{1}{n^{\frac{3}{2} + \alpha}} \right) \) on the detection of the transmitted \( x_j \) in each subband \( j \). We can see that, for a large \( n \), the average received energy, \( y_j \), concentrates in a region with width proportional to \( \frac{1}{n^{\frac{3}{2}}} \) around the energy of the signal and the noise \( |x_j|^2 + 1 \). When \( E \left[ |x_j|^2 \right] = \frac{1}{n^{\frac{3}{2}}} \), the distribution of the average received energy \( y_j \) in each subband \( j \) corresponding to distinct transmitted symbols have a significant overlap and we cannot distinguish \( x_j \) accurately at the multiple antenna receiver.

\[ \square \]
the same result with $N = o(n^{\frac{1}{2}})$.

**Theorem 4.** For any $\epsilon < \frac{1}{2}$, equal power and rate allocation in $N$ frequency bins of total bandwidth $B = \Theta(n^\alpha)$ achieve

$$\Theta(n^\alpha \log_2(n))$$

bits per transmission with vanishing probability of error with increasing $n$.

**Proof.** Let $\alpha \triangleq \frac{1}{2} - \epsilon$ for some $\alpha > 0$. Then $N = \Theta(n^{\frac{1}{2}-\alpha})$.

The transmitter chooses $C$ to be

$$C = \left\{0, \sqrt{2d}, \sqrt{4d}, \cdots, \sqrt{\frac{2}{N}}\right\},$$

where

$$d = \frac{1}{N((d-1) - 1)} = \frac{1}{(2N - 1)}.$$ (13)

The decoding regions are $\mathcal{I}_1 = (-\infty, 1+d]$, $\mathcal{I}_k = ((2k-1)d+1, (2k+1)d+1]$ for $2 \leq k \leq |C| - 1$, $\mathcal{I}_{|C|} = ((2|C|-1)d+1, \infty)$. Choosing $d = \Theta\left(\frac{1}{N}\right)$ for some $t$ such that $t > \frac{1}{2} - \alpha > 0$, we get that the rate $R$ should satisfy

$$R = \Theta\left(n^{\frac{1}{2}-\alpha} \log_2(1 + n^{t-(\frac{1}{2}-\alpha)\alpha})\right)$$

$$= \Theta\left(n^{\frac{1}{2}-\alpha} \log_2(n)\right).$$ (14)

We can also verify that $|C| = \Theta(n^{t-(\frac{1}{2}-\alpha)\alpha})$. We now show a vanishing probability of error if $0 < \alpha < \frac{1}{2}$ and $t < \frac{1}{2}$. We consider a union bound on the probability of error over all frequency bins. We have

$$P_{\text{error}} \leq \sum_{j=1}^{N} \sum_{k \in C} P_{\text{error},j}(k \neq k)$$

$$\leq n^{\frac{1}{2}-\alpha} n^{t-(\frac{1}{2}-\alpha)\alpha} \epsilon^{-n} (\min_k I_k(d)) \approx n^t \epsilon^{-n^{1-2\alpha}},$$ (15)

where $P_{\text{error},j}$ is the probability of error due to transmission in the $j$th subband, $(a)$ uses the following definition

$$I_k(d) \triangleq \lim_{n \to \infty} -\log\left(\frac{\text{Prob}\left(\frac{y^{(j)} \sqrt{2(t-1) + \epsilon_k} x_j^2}{n} - 2(k-1)d-1 > d\right)}{n}\right),$$ (16)

and Lemma 1 in [10], and $(b)$ follows from the fact that

$$\lim_{d \to 0} \frac{I_k(d)}{d^2} = \Theta(1)$$

for all $k$, as was shown in [10]. The right hand side in (15) goes to zero if $t < \frac{1}{2}$. Fig. 2b shows how values of $y^{(j)}$ are distributed around $|x_j|^2 + 1$. The figure suggests that if

$$E \left[|x_j|^2\right] < \frac{1}{n^{\frac{1}{2}-\alpha}},$$

the transmitted signals are sufficiently distinguishable at the receiver for a large enough $n$. Note that since we have used a union bound in (15) these achievable rates continue to hold even with channel correlations across frequency bins.

---

**VI. CONCLUSIONS**

In traditional wideband analysis, there is an equivalence between the concepts of bandwidth $B$ going to infinity, SNR going to zero, and capacity being power-limited. However, with the joint scaling of the number of receive antennas and the bandwidth, it is possible to have a new wideband regime that is bandwidth-limited, i.e., capacity increases with increasing bandwidth. In this regime, achievable rates scale with both the bandwidth and the number of receive antennas. When $B \leq o(n^{\frac{1}{2}})$ the system is bandwidth-limited, and spreading transmit power over different frequency bands helps the achievable rates. With $B \geq \Theta(n^{\frac{1}{2}+\alpha})$ for some $\alpha > 0$, we see the problem of overspreading previously reported for fixed $n$ causing the achievable rates to be less than those achieved using a $o\left(n^{\frac{1}{2}}\right)$ subset of the bandwidth.

We conjecture that our result (Theorem 1) holds not only with the PAPR constraint, but also for general signaling schemes (peaky or non-peaky). If true, this would highlight a qualitative difference in optimal signaling schemes for small antenna versus large antenna noncoherent wideband systems. Characterizing this difference is part of our future work.

**ACKNOWLEDGEMENTS**

This work was supported by a 3Com Corporation and an Alcatel-Lucent Stanford Graduate Fellowship, an A.G. Leventis Foundation Scholarship, NSF grant 1320628, NSF grant 1302336, ONR grant N000141210063, FPU12/01319 MECD.

**REFERENCES**


