Reference Tone Aided Transmission for Massive MIMO: Analog Beamforming without CSI

Vishnu V. Ratnam*, *Student Member, IEEE* and Andreas F. Molisch*, *Fellow, IEEE*
*Ming Hsieh Dept. of Electrical Engineering, University of Southern California
Los Angeles, California, USA. Email: {ratnam, molisch}@usc.edu

***Abstract*—This work proposes a novel transmission scheme, namely Reference Tone Aided Transmission (RTAT), that can enable communication in massive MIMO systems with a single analog-to-digital converter at the receiver. Unlike conventional low complexity MIMO transceivers, it can perform receive beamforming without explicit channel estimation or use of phase shifters, thereby leading to a significant reduction in channel estimation overhead and simpler initial access. In RTAT, a sinusoidal reference tone is transmitted along with the data signals. At each receive antenna, the reference tone is recovered, and multiplied by the received signals, to obtain a base-band signal whose inter-antenna phase shift has been compensated. The resulting low-pass signals from all the antennas are then added, emulating maximal ratio combining at the receiver with imperfect channel knowledge. In this work, a receiver architecture for RTAT is proposed, its performance is analyzed and is compared with that of fully digital beamforming. The capacity maximizing power allocation problem is also considered, and one possible initial access protocol for RTAT is suggested. Simulation results show that, in comparison to hybrid beamforming, RTAT suffers only a small loss in signal-to-noise ratio, while providing a significant reduction in the channel estimation overhead.**

I. Introduction

Despite their numerous benefits, especially at millimeter-wave (mm-wave) frequencies, massive Multiple-input-multiple-output (MIMO) systems [1], wherein the transmitter (TX) and/or receiver (RX) is equipped with a very large array of antenna elements, are hard to implement in practice. This is because, though the antenna elements are themselves cheap, the corresponding up/down-conversion chains - which include mixed signal components such as filters, mixers, analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) - are both expensive and power hungry. As a solution, two prospective technologies: hybrid beamforming [2], [3] and digital beamforming with low-resolution ADCs/DACs[1] [4], [5] have been proposed. The former approach uses analog hardware, such as phase-shifters, to connect a large antenna array to a small number of up/down-conversion chains, while the latter approach uses low bit-resolution ADCs in the down-conversion chains. These approaches rely on the premise that the analog hardware and low resolution ADCs are more power efficient and cost effective than the full complexity up/down-conversion chains.

A major challenge with these approaches is the estimation of channel parameters. In hybrid beamforming, multiple pilot transmissions are required to acquire the average channel state information (aCSI) [6]–[8], such as the directions of departure/arrival of the channel multi-path components (MPCs). As an illustration, for an exhaustive aCSI estimation procedure, a total of $O(M_{\rm tx}M_{\rm rx}/J)$ such pilot transmissions are required, where $M_{\rm tx}, M_{\rm rx}$ are the number of TX and RX antennas and J is the number of down-conversion chains at the RX. Similarly, digital beamforming with low-resolution ADCs may require long pilot sequences to compensate for the impact of quantization on channel estimation and timing and frequency synchronization [9], [10]. Such large pilot overheads may consume a significant portion of the time-frequency resources when channel estimation is required frequently, such as in fast time varying channels, e.g., vehicle-to-vehicular systems, or in channels with high blocking probabilities, e.g., at mm-wave frequencies [11]. They also increase system latency and make procedures such as initial access (IA) - wherein, a user equipment (UE) discovers a base-station (BS), synchronizes, and coordinates to initiate communication - very cumbersome [12], [13]. While some approaches can reduce this pilot overhead [8], [14], [15], they require the timing and frequency synchronization [16] to be performed without the RX beamforming gain, which may be difficult at the low signal-to-noise ratio (SNR) and high phase noise levels expected in mm-wave systems.

Therefore, in this work we propose Reference Tone Aided Transmission (RTAT), a novel scheme that enables communication in massive MIMO systems with reduced hardware and energy cost, while alleviating some of the aforementioned problems. In RTAT, the TX transmits a reference tone at a known frequency, along with the data. At each RX antenna, the received waveform corresponding to the reference signal is recovered, and multiplied to the received waveform corresponding to the data signals, in the analog domain. This results in a low-pass signal with the inter-antenna phase shift compensated. The outputs from each antenna are then summed up, low-pass filtered and sampled for data demodulation. This operation emulates maximal ratio combining (MRC) with imperfect channel state information at the RX (CSIR). In the architecture considered here, the RTAT RX uses only one down-conversion chain and can perform beamforming without requiring analog phase shifters - which may suffer large insertion losses [17], and without explicit channel estimation, thereby obviating the need for aCSI acquisition at the RX. It also exploits power from all the channel MPCs and is therefore resilient to shadow fading and blocking of MPCs. While not explored here, the received signals from the reference tone can

This work was supported by the National Science Foundation.

[1]Henceforth, we shall use the phrase 'low-resolution ADCs' to imply 'low-resolution ADCs' when at the RX and 'low-resolution DACs' when at the TX.

also be used for transmit beamforming on the reverse link. On the flip side, the RTAT transceivers, in their current suggested form, do not support multiple spatial data streams and can only be used for beamforming at one end of a communication link. They are therefore more suited for use at the UEs.

RTAT is inspired by frequency shifted reference transmission for single-input-single-output (SISO) ultra-wideband (UWB) systems [18], [19], where a similar product between data and reference signal leads to in-phase accumulation of the UWB data pulses, albeit with significant noise enhancement. Similarly, a reference tone is also used in some legacy mm-wave and optical networks [20] to compensate for carrier frequency offset and channel dispersion. However, its use for beamforming has not been explored before, to the best of our knowledge, so that RTAT constitutes a novel method for reduced-complexity massive MIMO. Furthermore, unlike frequency shift reference, RTAT does not involve bandwidth expansion and hence does not suffer from significant noise enhancement. Thus, the contributions of this paper are:

1) We propose RTAT, a novel transmission scheme for low-complexity massive MIMO systems, that does not require phase-shifters or explicit CSI estimation at the RX.
2) We propose an RX architecture for the RTAT scheme, and characterize its achievable throughput in a wide-band channel, for a single spatial data-stream.
3) We derive a near-optimal power allocation for data streams and propose an IA procedure for RTAT.
4) Simulations compare performance of the proposed RTAT scheme to hybrid beamforming.

Notation: scalars are represented by light-case letters; vectors by bold-case letters; and sets by calligraphic letters. Additionally, $\mathrm{j} = \sqrt{-1}$, $\mathbb{E}\{\}$ represents the expectation operator, c^* is the complex conjugate of a complex scalar c, $\mathbf{c}^\dagger$ is the Hermitian transpose of a complex vector $\mathbf{c}$ and $\mathrm{Re}\{\cdot\}$/$\mathrm{Im}\{\cdot\}$ refer to the real/imaginary component, respectively.

II. General Assumptions and System Model

We consider a point-to-point single-input-multiple-output (SIMO) system where the TX has a single antenna and the RX has $M \gg 1$ antennas. Note that even a point-to-point MIMO system where the TX transmits a single spatial data stream can be modeled as an equivalent SIMO system. The applicability of this model to the downlink of a cellular network is exemplified later in Section V. Both the TX and RX are assumed to have one up/down-conversion chains each. The TX transmits OFDM symbols with $2K+1$ sub-carriers, indexed as $\{-K, -K+1, ..., K-1, K\}$, respectively. A reference tone, i.e., a pure sinusoidal signal with a pre-determined frequency known both to the TX and RX, is transmitted on the 0-th subcarrier. The system also transmits data on $K-g$ lower and higher sub-carriers, represented by the index set $\mathcal{K} = \{-K, ..., -g-1, g+1, ..., K\}$. The sub-carriers $\{-g, ..., g\}$ are blanked, to act as a guard band between the reference and data sub-carriers. The value of g is determined by the carrier recovery process at RX, as shall be discussed later. For convenience, we also neglect phase noise at the TX.

The *complex equivalent* transmit signal for the 0-th symbol can then be expressed as:

$$s_{\mathrm{tx}}(t) = \sqrt{\frac{2}{T_{\mathrm{s}}}} \left[\sqrt{E_{\mathrm{r}}} + \sum_{k \in \mathcal{K}} x_k e^{\mathrm{j}2\pi f_k t} \right] e^{\mathrm{j}2\pi f_{\mathrm{c}} t} \tag{1}$$

for $-T_{\mathrm{cp}} \leq t \leq T_{\mathrm{s}}$, where E_{r} is the energy allocated to the reference tone, x_k is the data signal at the k-th OFDM sub-carrier, f_{c} is the carrier frequency, $f_k = k/T_{\mathrm{s}}$ represents the frequency offset of k-th sub-carrier from the reference signal and $T_{\mathrm{s}}, T_{\mathrm{cp}}$ are the symbol duration and the cyclic prefix duration, respectively. Here we define the *complex equivalent* signal such that the actual (real) transmit signal is given by $\mathrm{Re}\{s_{\mathrm{tx}}(t)\}$. For convenience, we also assume that f_{c} is a multiple of both $1/T_{\mathrm{s}}$ and $1/(T_{\mathrm{cp}}+T_{\mathrm{s}})$, which ensures that the reference tone has no phase transitions between symbols. We further assume the data vector $\mathbf{x}$ has independent zero mean entries. The total average transmit symbol energy is given by $E_{\mathrm{s}} = E_{\mathrm{r}} + \sum_{k \in \mathcal{K}} E_{\mathrm{d}}^{(k)}$, where $E_{\mathrm{d}}^{(k)} = \mathbb{E}\{|x_k|^2\}$ is the energy allocated to the k-th sub-carrier.[2]

The channel is assumed to have $L \ll M$ scatterers with the $M \times 1$ channel impulse response vector given as [11]:

$$\mathbf{h}(t) = \sum_{\ell=0}^{L-1} \alpha_\ell \mathbf{a}(\phi_\ell, \theta_\ell) \delta(t - \tau_\ell) \tag{2}$$

where ϕ_ℓ, θ_ℓ are the azimuth and elevation angles of arrival, α_ℓ is the complex gain and τ_ℓ is the delay, respectively, of the ℓ-th scatterer, $\delta(t)$ is the Dirac delta function, and $\mathbf{a}(\phi, \theta)$ is the RX array response vector for azimuth angle ϕ and elevation angle θ. To prevent inter symbol interference, we let the cyclic prefix be longer than the maximum channel delay: $T_{\mathrm{cp}} > \tau_{L-1}$. For ease of analysis, we further assume that the array response vectors for the scatterers are mutually orthogonal, i.e., $\mathbf{a}(\phi_i, \theta_i)^\dagger \mathbf{a}(\phi_j, \theta_j) = M$ if $i = j$ and 0 otherwise. This assumption is reasonable if the scatterers are well separated and $M \gg L$ [21]. To model a generic time varying channel, we assume that the MPC parameters $\{\alpha_\ell, \phi_\ell, \theta_\ell, \tau_\ell\}$ remain constant for at least a coherence time interval T_{coh}, and may/ may not change afterwards.

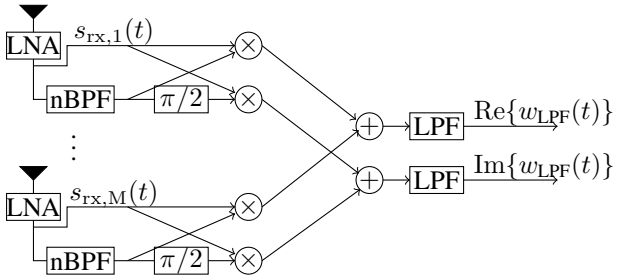
Ignoring any RX non-linear effects, the $M \times 1$ *complex equivalent* received waveform for $0 \leq t \leq T_{\mathrm{s}}$ can be expressed as:

$$\mathbf{s}_{\mathrm{rx}}(t) = \sum_{\ell=0}^{L-1} \alpha_\ell \mathbf{a}(\phi_\ell, \theta_\ell) s_{\mathrm{BB}}(t - \tau_\ell) e^{\mathrm{j}2\pi f_{\mathrm{c}}(t-\tau_\ell)} + \mathbf{n}(t), \tag{3}$$

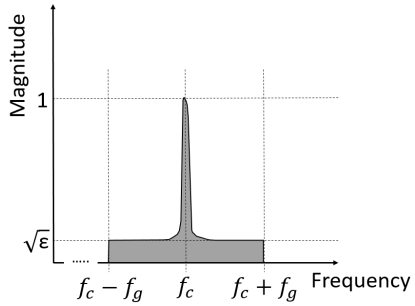
where $s_{\mathrm{BB}}(t) \triangleq s_{\mathrm{tx}}(t) e^{-\mathrm{j}2\pi f_c t}$ is the base-band transmit signal, and $\mathbf{n}(t)$ is the $M \times 1$ stationary, complex additive Gaussian noise process vector, with individual entries being circularly symmetric, independent and identically distributed (i.i.d.), and having a power spectral density: $\mathcal{S}_{\mathrm{n}}(f) = 2\mathrm{N}_0$ for all $f \geq 0$. Using a Fourier series expansion, the noise for $0 \leq t \leq T_{\mathrm{s}}$ can be expressed as:

$$\mathbf{n}(t) = \sum_{k=0}^{\infty} \widetilde{\mathbf{N}}[k] e^{\mathrm{j}2\pi f_k t} = \sum_{k=-f_{\mathrm{c}}T_{\mathrm{s}}}^{\infty} \mathbf{N}[k] e^{\mathrm{j}2\pi f_k t} e^{\mathrm{j}2\pi f_{\mathrm{c}} t}, \tag{4}$$

[2]For ease of notation, we do not consider the additional energy per symbol required to transmit the cyclic prefix viz., $E_{\mathrm{s}} T_{\mathrm{cp}}/T_{\mathrm{s}}$.



(a) Block Diagram



(b) nBPF frequency response

Fig. 1. Block diagram for a multi-antenna RTAT receiver and the frequency response of the narrow band pass filter.

where $f_k = k/T_{\mathrm{s}}$ as before, $\widetilde{\mathbf{N}}[k]$ is a $M \times 1$ vector with i.i.d. $\mathcal{CN}(0, 2\mathrm{N}_0/T_{\mathrm{s}})$ entries and we define $\mathbf{N}[k] = \widetilde{\mathbf{N}}[k + f_{\mathrm{c}}T_{\mathrm{s}}]$. Furthermore $\widetilde{\mathbf{N}}[\bar{k}]$ and $\widetilde{\mathbf{N}}[\ddot{k}]$ are mutually independent for $\bar{k} \neq \ddot{k}$. This expansion shall be helpful in the forthcoming analysis.

The received signal $\mathbf{s}_{\mathrm{rx,m}}(t)$ at each antenna m is then passed through a low noise amplifier (LNA) and multiplied by a narrow band-pass filtered version of itself, as depicted in Fig. 1a. The frequency response of the narrow band-pass filter (nBPF) used is as depicted in Fig. 1b, and it essentially recovers the reference tone. The $M \times 1$ *complex equivalent* output from the nBPFs at the M antennas for $0 \leq t \leq T_{\mathrm{s}}$ can be expressed as:

$$\hat{\mathbf{s}}_{\mathrm{rx}}(t) \;=\; \sum_{\ell=0}^{L-1} \sqrt{\frac{2E_{\mathrm{r}}}{T_{\mathrm{s}}}} \alpha_\ell \mathbf{a}(\phi_\ell, \theta_\ell) e^{\mathrm{j}2\pi f_{\mathrm{c}}(t-\tau_\ell)} + \hat{\mathbf{n}}(t), \quad (5)$$

where $\hat{\mathbf{n}}(t) = \sum_{k=-g}^{g} \hat{\mathbf{N}}[k] e^{\mathrm{j}2\pi(f_{\mathrm{c}}+f_k)t}$ with $\hat{\mathbf{N}}[k] = \sqrt{\epsilon}\mathbf{N}[k]$ for $k \neq 0$ and $\hat{\mathbf{N}}[0] = \mathbf{N}[0]$. By using a linear approximation for the output phase noise of a phase locked loop (PLL) [22], it can be shown that such a sharp nBPF can be closely approximated by a PLL followed by a variable gain amplifier. In such an implementation, the noise in (5) represents the PLL phase noise, f_g is the PLL loop bandwidth and g can be as low as 6 [20]. For brevity, we shall assume availability of such an nBPF block at each antenna, and shall discuss implementation details in future work.

The received signal $\mathrm{Re}\{\mathbf{s}_{\mathrm{rx,m}}(t)\}$ at each antenna m is then mixed with the recovered nBPF output $\mathrm{Re}\{\hat{\mathbf{s}}_{\mathrm{rx}}(t)\}$ and its quadrature counterpart $\mathrm{Im}\{\hat{\mathbf{s}}_{\mathrm{rx}}(t)\}$, as depicted in Fig. 1a. The mixed outputs are then summed up and low pass filtered (with cut-off frequency f_K) to obtain, for $0 \leq t \leq T_{\mathrm{s}}$:

$$\begin{aligned}
w_{\mathrm{LPF}}(t) &= \mathrm{LPF}\left\{\hat{\mathbf{s}}_{\mathrm{rx}}^{\dagger}(t)\mathrm{Re}\{\mathbf{s}_{\mathrm{rx}}(t)\}\right\} \\
&= \sum_{\ell=0}^{L-1}\Bigg[\frac{\sqrt{E_{\mathrm{r}}}}{T_{\mathrm{s}}} M|\alpha_\ell|^2\Big(\sqrt{E_{\mathrm{r}}} + \sum_{k\in\mathcal{K}} x_k e^{\mathrm{j}2\pi f_k(t-\tau_\ell)}\Big) \\
&+ \sum_{k=-K}^{K} \sqrt{\frac{E_{\mathrm{r}}}{2T_{\mathrm{s}}}} \alpha_\ell^* e^{\mathrm{j}2\pi f_{\mathrm{c}}\tau_\ell} \mathbf{a}(\phi_\ell,\theta_\ell)^{\dagger}\mathbf{N}[k] e^{\mathrm{j}2\pi f_k t} \\
&+ \sum_{\bar{k}=-g}^{g}\Bigg(\sqrt{\frac{E_{\mathrm{r}}}{2T_{\mathrm{s}}}}\alpha_\ell e^{-\mathrm{j}2\pi f_{\mathrm{c}}\tau_\ell}\hat{\mathbf{N}}[\bar{k}]^{\dagger}\mathbf{a}(\phi_\ell,\theta_\ell)e^{-\mathrm{j}2\pi f_{\bar{k}}t} \\
&+ \sum_{k=\bar{k}-K}^{\bar{k}+K} \frac{\alpha_\ell e^{-\mathrm{j}2\pi(f_{\mathrm{c}}+f_k)\tau_\ell}}{\sqrt{2T_{\mathrm{s}}}} x_k e^{\mathrm{j}2\pi(f_k-f_{\bar{k}})t}\hat{\mathbf{N}}[\bar{k}]^{\dagger}\mathbf{a}(\phi_\ell,\theta_\ell)\Bigg)\Bigg] \\
&+ \sum_{\bar{k}=-g}^{g}\sum_{k=\bar{k}-K}^{\bar{k}+K} \frac{e^{\mathrm{j}2\pi(f_k-f_{\bar{k}})t}}{2}\hat{\mathbf{N}}[\bar{k}]^{\dagger}\mathbf{N}[k], \qquad (6)
\end{aligned}$$

where we represent the outputs corresponding to the in-phase and quadrature phase components of $\hat{\mathbf{s}}_{\mathrm{rx}}(t)$ as the Re, Im components of $w_{\mathrm{LPF}}(t)$, we set $x_k = 0$ for $k \notin \mathcal{K}$, we use the orthogonality assumption for the array response vectors, and the Fourier series expansion of the noise terms. Finally, $w_{\mathrm{LPF}}(t)$ is sampled by an ADC at a sampling rate of $T_{\mathrm{s}}/(2K+1)$ samples/sec and conventional OFDM demodulation follows.[3] As shall be shown in following sections, $w_{\mathrm{LPF}}(t)$ attains a good SNR and so we shall henceforth assume perfect timing synchronization between TX and RX.

III. Analysis of the demodulation outputs

The OFDM demodulated output for the k-th subcarrier, for any $k \in \mathcal{K}$, can be expressed as:

$$Y_k = \frac{T_{\mathrm{s}}}{2K+1}\sum_{u=0}^{2K} w_{\mathrm{LPF}}\Big(\frac{uT_{\mathrm{s}}}{2K+1}\Big)e^{-\frac{\mathrm{j}2\pi ku}{2K+1}}. \quad (7)$$

We shall split Y_k as $Y_k = S_k + Z_k$ where S_k, referred to as the signal component, involves terms in (7) not containing the channel noise and Z_k, referred to as the noise component, containing the remaining terms. For simplicity, we consider a sub-optimal demodulation approach where Z_k is treated as noise.

A. Signal component analysis

From (6)–(7), the signal component of the OFDM demodulated output Y_k, for any $k \in \mathcal{K}$, can be expressed as:

$$\begin{aligned}
S_k &= \sum_{u=0}^{2K}\sum_{\ell=0}^{L-1}\sum_{\bar{k}\in\mathcal{K}} \frac{\sqrt{E_{\mathrm{r}}}M|\alpha_\ell|^2}{2K+1} x_{\bar{k}} e^{\mathrm{j}2\pi f_{\bar{k}}(\frac{uT_{\mathrm{s}}}{2K+1}-\tau_\ell)} e^{-\frac{\mathrm{j}2\pi ku}{2K+1}} \\
&= M\sqrt{E_{\mathrm{r}}}\Big[\sum_{\ell=0}^{L-1}|\alpha_\ell|^2 e^{-\mathrm{j}2\pi f_k\tau_\ell}\Big]x_k. \qquad (8)
\end{aligned}$$

Essentially, the RX utilizes the received signal vector corresponding to the reference tone as *weights* to combine the

[3]Since the reference tone-reference tone product in (6) produces a DC output, using a DC shift can negate its impact on the ADC dynamic range.

received signal vector corresponding to the data sub-carriers, i.e., it emulates MRC combining with imperfect CSIR. Since this operation takes place in the RF domain, the proposed RX does not rely on explicit channel estimation or phase-shifters for combining the received signals. However, as is evident from (8), the signal from the L multi-path components do not add up *in-phase*, at the demodulation outputs. This is due to the fact that the reference tone and the k-th data stream pass through slightly different channels owing to the difference in their modulating frequency. This leads to some amount of frequency selective fading, the impact of which is quantified in sections IV and VI.

B. Noise component analysis

From (6)–(7), the noise component of the OFDM demodulation output Y_k, for any $k \in \mathcal{K}$, can be expressed as:

$$
\begin{aligned}
Z_k &\overset{(1)}{=} \sum_{u=0}^{2K} \Bigg[\sum_{\ell=0}^{L-1} \Bigg(\sqrt{\frac{E_\mathrm{r}}{2T_\mathrm{s}}} \alpha_\ell^* e^{\mathrm{j}2\pi f_\mathrm{c} \tau_\ell} \mathbf{a}(\phi_\ell, \theta_\ell)^\dagger \mathbf{N}[k] e^{\frac{\mathrm{j}2\pi k u}{2K+1}} \\
&+ \sum_{\bar{k}=-g}^{g} \frac{\alpha_\ell e^{-\mathrm{j}2\pi (f_\mathrm{c}+f_{\bar{k}}+f_k)\tau_\ell}}{\sqrt{2T_\mathrm{s}}} x_{\bar{k}+k} e^{\frac{\mathrm{j}2\pi k u}{2K+1}} \hat{\mathbf{N}}[\bar{k}]^\dagger \mathbf{a}(\phi_\ell, \theta_\ell) \Bigg) \\
&+ \sum_{\bar{k}=-g}^{g} \frac{e^{\frac{\mathrm{j}2\pi k u}{2K+1}}}{2} \hat{\mathbf{N}}[\bar{k}]^\dagger \mathbf{N}[\bar{k}+k] \Bigg] \frac{T_\mathrm{s} e^{-\frac{\mathrm{j}2\pi k u}{2K+1}}}{2K+1} \\
&= \sum_{\ell=0}^{L-1} \Bigg(\sqrt{\frac{E_\mathrm{r} T_\mathrm{s}}{2}} \alpha_\ell^* e^{\mathrm{j}2\pi f_\mathrm{c} \tau_\ell} \mathbf{a}(\phi_\ell, \theta_\ell)^\dagger \mathbf{N}[k] \\
&+ \sum_{\bar{k}=-g}^{g} \frac{\alpha_\ell \sqrt{T_\mathrm{s}} e^{-\mathrm{j}2\pi (f_\mathrm{c}+f_{\bar{k}}+f_k)\tau_\ell}}{\sqrt{2}} x_{\bar{k}+k} \hat{\mathbf{N}}[\bar{k}]^\dagger \mathbf{a}(\phi_\ell, \theta_\ell) \Bigg) \\
&+ \sum_{\bar{k}=-g}^{g} \frac{T_\mathrm{s}}{2} \hat{\mathbf{N}}[\bar{k}]^\dagger \mathbf{N}[\bar{k}+k], \qquad (9)
\end{aligned}
$$

where $x_{\bar{k}} = 0$ for $\bar{k} \notin \mathcal{K}$, and $\overset{(1)}{=}$ follows by observing that the other terms in (6) do not contribute to the k-th demodulation output. As is evident, the noise consists of both: noise-noise and data-noise cross product terms. Given the transmit data vector $\mathbf{x}$ and channel impulse response $\mathbf{h}(t)$, the conditional mean of the noise components can be computed as:

$$\mathbb{E}\{Z_k | \mathbf{x}, \mathbf{h}\} = 0, \quad k \in \mathcal{K},$$

which follows from (4). Similarly, the conditional second order statistics of the noise components can be expressed as (detailed steps are given in Appendix A):

$$
\begin{aligned}
\mathcal{K}_{a,b}(\mathbf{x}, \mathbf{h}) &\triangleq \mathbb{E}\{Z_a Z_b^*\} \\
&= \beta_{0,0} M \mathrm{N}_0 E_\mathrm{r} \delta_{a,b} + \beta_{a,-b} M \mathrm{N}_0 \Bigg(x_a x_b^* (1-\epsilon) \\
&+ \sum_{k=-g}^{g} x_{k+a} x_{k+b}^* \epsilon \Bigg) + \delta_{a,b} M \Bigg((1-\epsilon)\mathrm{N}_0^2 + \sum_{k=-g}^{g} \epsilon N_0^2 \Bigg), \qquad (10) \\
\widetilde{\mathcal{K}}_{a,b}(\mathbf{x}, \mathbf{h}) &\triangleq \mathbb{E}\{Z_a Z_b\} = \sum_{(\bar{k},\ddot{k}) \in \mathcal{A}_a} \delta_{a,-b} \epsilon M \mathrm{N}_0^2, \qquad (11)
\end{aligned}
$$

where $\beta_{a,b}(\mathbf{h}) \triangleq \sum_\ell |\alpha_\ell|^2 e^{-\mathrm{j}2\pi(f_a+f_b)\tau_\ell}$, $\delta_{a,b}$ is the Kronecker delta function i.e., $\delta_{a,b} = 1$ if $a = b$, $\delta_{a,b} = 0$ otherwise and $\mathcal{A}_a = \{(\bar{k}, \ddot{k}) | \ddot{k} = \bar{k} + a, -g \le \bar{k}, \ddot{k} \le g\}$. Note that $\mathcal{A}_a = \{\}$ for $|a| > 2g$, i.e., the pseudo cross-covariance is 0 for most of the demodulated sub-carriers. Clearly, from (10)–(11) we see that the noise components at the OFDM outputs are mutually correlated and are further dependent on the data vector $\mathbf{x}$. Therefore, for optimal decoding, a joint estimation of the data vector $\mathbf{x}$ should be performed. Here we consider the simple approach, where the dependence of noise terms on $\mathbf{x}$ is not exploited. Under this assumption, the noise variances, averaged over the transmit data vector $\mathbf{x}$ reduce to:

$$
\begin{aligned}
\mathcal{K}_{a,b}(\mathbf{h}) &= \beta_{0,0} M \mathrm{N}_0 \delta_{a,b} \Bigg[E_\mathrm{r} + E_\mathrm{d}^{(a)} + \sum_{k \in \mathcal{B}_a} E_\mathrm{d}^{(k+a)} \epsilon \Bigg] \\
&\quad + \delta_{a,b} M \mathrm{N}_0^2 \big[1 + 2g\epsilon \big], \qquad (12a) \\
\widetilde{\mathcal{K}}_{a,b}(\mathbf{h}) &\approx 0, \qquad (12b)
\end{aligned}
$$

where we use the fact that $\mathbf{x}$ has zero mean, mutually independent entries, we define $\mathcal{B}_a = \{k \big| |a+k| > g, -g \le k \le g, k \ne 0\}$, and neglect the pseudo cross-covariance term in (11). We shall further approximate the noise components $\{Z_k | k \in \mathcal{K}\}$ to be jointly Gaussian distributed, which allows us to find a closed-form lower bound to the system capacity. The validity of this assumption is tested via simulations in Section VI.

IV. Performance Analysis

Using (8), (12) and the joint Gaussian assumption for $\{Z_k | k \in \mathcal{K}\}$, the effective SISO channel between the transmit data and the demodulation outputs can be expressed as:

$$Y_k = M \sqrt{E_\mathrm{r} E_\mathrm{d}^{(k)}} \beta_{k,0}(\mathbf{h}) \bar{x}_k + Z_k \qquad k \in \mathcal{K}, \qquad (13)$$

where, $\bar{x}_k \triangleq x_k / \sqrt{E_\mathrm{d}^{(k)}}$ and $Z_k \sim \mathcal{CN}(0, \mathcal{K}_{k,k}(\mathbf{h}))$. Note that (13) represents a parallel Gaussian channel. The SNR and instantaneous spectral efficiency (iSE), respectively, for the k-th data stream can then be expressed as:

$$
\begin{aligned}
\gamma_k(\mathbf{h}(t)) &= \frac{M E_\mathrm{r} E_\mathrm{d}^{(k)} |\beta_{k,0}(\mathbf{h})|^2}{\beta_{0,0} \mathrm{N}_0 \Big[E_\mathrm{r} + E_\mathrm{d}^{(k)} + \sum_{\bar{k} \in \mathcal{B}_k} E_\mathrm{d}^{(\bar{k}+k)} \epsilon \Big] + \mathrm{N}_0^2 \big[1 + 2g\epsilon \big]}, \qquad (14) \\
\mathrm{iSE}\big(\mathbf{h}(t)\big) &\triangleq \sum_{k \in \mathcal{K}} \frac{1}{2K+1} \log\big(1 + \gamma_k(\mathbf{h}(t))\big), \qquad (15)
\end{aligned}
$$

where we neglect the cyclic prefix overhead for convenience. Note that even though the RX does not explicitly estimate the array response vectors $\mathbf{a}(\phi_\ell, \theta_\ell)$ for the MPCs and does not use phase-shifters, we still observe a beamforming gain of M in $\gamma_k(\mathbf{h}(t))$. However, one drawback is that $\beta_{k,0}(\mathbf{h}) \le \beta_{0,0}(\mathbf{h})$, and is different for each k i.e., we encounter some frequency selective fading. The performance loss due to this fading behavior is studied via simulations in Section VI.

Note that the channel gain and noise variance in (13): $\{\beta_{k,0}(\mathbf{h}), \mathcal{K}_{k,k}(\mathbf{h}) | k \in \mathcal{K}\}$ need to be estimated at the RX, for data demodulation. This can be achieved by transmitting pilot symbols (and null pilots) on the data sub-carriers. Since the RX has a good beamforming gain (14), and $\{\beta_{k,0}(\mathbf{h}), \mathcal{K}_{k,k}(\mathbf{h}) | k \in \mathcal{K}\}$ are average channel parameters which change slowly with time, they can be tracked accurately with very low estimation overhead. If required, these parameters can further be fed back to the TX, via a feedback channel,

for adaptive modulation and power allocation. Since the pilots are used only to estimate the average SISO channel parameters and not the actual SIMO channel, the advantages of simplified IA and channel estimation are still applicable for the proposed scheme.

A. Power Allocation

In this subsection we aim to find the iSE maximizing power allocation to the data sub-carriers and reference tone. Since both the signal and noise variances in (13) are affected by the transmit powers in a non-linear way, finding the optimal power allocation is difficult. We shall therefore rely on the following approximate SNR expression for power allocation:

$$\hat{\gamma}_k(\mathbf{h}(t)) = \frac{M E_{\mathrm{r}} E_{\mathrm{d}}^{(k)} |\beta_{k,0}(\mathbf{h})|^2}{\beta_{0,0}\mathrm{N}_0\Big[E_{\mathrm{r}} + (E_{\mathrm{s}} - E_{\mathrm{r}})\frac{1+g\epsilon}{|\mathcal{K}|}\Big] + \mathrm{N}_0^2\big[1+2g\epsilon\big]}, \quad (16)$$

which is obtained by using $|\mathcal{B}_k| \approx g$ and by replacing $E_{\mathrm{d}}^{(k)}$ by $(E_{\mathrm{s}} - E_{\mathrm{r}})/|\mathcal{K}|$. For $\hat{\gamma}_k(\mathbf{h}(t))$ we have the following lemma.

Lemma 1. *For any feasible power allocation $\{E_{\mathrm{r}}, E_{\mathrm{d}}^{(k)} | k \in \mathcal{K}\}$, we have: $\hat{\gamma}_k(\mathbf{h}(t)) \leq \hat{\bar{\gamma}}_k(\mathbf{h}(t))$ for all $k \in \mathcal{K}$, where $\hat{\bar{\gamma}}_k(\mathbf{h}(t))$ is the approximate SNR with:*

$$\bar{E}_{\mathrm{r}} = \frac{-\chi + \sqrt{\chi^2 + \xi\chi E_{\mathrm{s}}}}{\xi}, \ \bar{E}_{\mathrm{d}}^{(k)} = \frac{E_{\mathrm{d}}^{(k)}(E_{\mathrm{s}} - \bar{E}_{\mathrm{r}})}{E_{\mathrm{s}} - E_{\mathrm{r}}}, \quad (17)$$

$\chi = \mathrm{N}_0^2\big[1+2g\epsilon\big] + \beta_{0,0}\mathrm{N}_0 E_{\mathrm{s}} \frac{1+g\epsilon}{|\mathcal{K}|}$ *and* $\xi = \beta_{0,0}\mathrm{N}_0(1 - \frac{1+g\epsilon}{|\mathcal{K}|})$.

Proof. For any power allocation $\{E_{\mathrm{r}}, E_{\mathrm{d}}^{(k)} | k \in \mathcal{K}\}$, consider the alternate power allocation $\{\tilde{E}_{\mathrm{r}}, \frac{E_{\mathrm{d}}^{(k)}(E_{\mathrm{s}} - \tilde{E}_{\mathrm{r}})}{E_{\mathrm{s}} - E_{\mathrm{r}}} | k \in \mathcal{K}\}$ for some $\tilde{E}_{\mathrm{r}} \in [0, E_{\mathrm{s}}]$. Then from (16), for any $k \in \mathcal{K}$ we have:

$$\hat{\tilde{\gamma}}_k(\mathbf{h}(t)) = \frac{M \tilde{E}_{\mathrm{r}} E_{\mathrm{d}}^{(k)} (E_{\mathrm{s}} - \tilde{E}_{\mathrm{r}}) |\beta_{k,0}(\mathbf{h})|^2}{(E_{\mathrm{s}} - E_{\mathrm{r}})(\xi \tilde{E}_{\mathrm{r}} + \chi)}, \quad (18)$$

where $\hat{\tilde{\gamma}}_k(\mathbf{h}(t))$ is the approximate SNR with the alternate power allocation, and ξ, χ are as defined in (17). By finding the second derivative of (18) with respect to $\tilde{E}_{\mathrm{r}}$, it can be verified that $\hat{\tilde{\gamma}}_k(\mathbf{h}(t))$ is a concave function of $\tilde{E}_{\mathrm{r}}$. Therefore, now equating the first derivative of (18) with respect to $\tilde{E}_{\mathrm{r}}$ to zero, we find that $\hat{\tilde{\gamma}}_k(\mathbf{h}(t))$ is maximized for:

$$\tilde{E}_{\mathrm{r}} = \frac{-\chi + \sqrt{\chi^2 + \xi\chi E_{\mathrm{s}}}}{\xi}, \quad (19)$$

where we skip the explicit steps for brevity. Since $\hat{\tilde{\gamma}}_k(\mathbf{h}(t)) = \hat{\gamma}_k(\mathbf{h}(t))$ for $\tilde{E}_{\mathrm{r}} = E_{\mathrm{r}}$, we have $\hat{\gamma}_k(\mathbf{h}(t)) \leq \hat{\bar{\gamma}}_k(\mathbf{h}(t))$. □

As a consequence of Lemma 1, the approximate data stream SNRs in (16) are maximized by using $E_{\mathrm{r}} = \bar{E}_{\mathrm{r}}$.[4] Now using $\hat{\gamma}_k(\mathbf{h}(t))$ instead of $\gamma_k(\mathbf{h}(t))$ in (15) with $\sum_k E_{\mathrm{d}}^{(k)} = E_{\mathrm{s}} - \bar{E}_{\mathrm{s}}$, a sub-optimal iSE maximizing power allocation for $\{E_{\mathrm{d}}^{(k)} | k \in \mathcal{K}\}$ can be obtained by the water-filling algorithm.

Note that for $|\mathcal{K}| \gg 1$, by ignoring the N_0^2 term and assuming $|\beta_{k,0}(\mathbf{h})|^2 \approx |\beta_{0,0}(\mathbf{h})|^2$ in (16), $\hat{\gamma}_k(\mathbf{h}(t))$ is same as the SNR for MRC combining with perfect CSIR. To show that such an N_0^2 term is also encountered in other receivers with noisy channel estimation, we next consider a fully digital M-antenna RX.

B. Digital beamforming with full resolution ADCs

For simplicity, we make the optimistic assumption that the fully digital RX has perfect knowledge of $\mathrm{span}\{\mathbf{a}(\phi_\ell, \theta_\ell) | 1 \leq \ell \leq L\}$, and therefore uses a $M \times L$ unitary digital pre-combiner $\mathbf{U}$ with $\mathrm{columnspan}\{\mathbf{U}\} = \mathrm{span}\{\mathbf{a}(\phi_\ell, \theta_\ell) | 1 \leq \ell \leq L\}$. Within each coherence time interval, the TX transmits one pilot symbol and RX uses the minimum-mean-square-error (MMSE) channel estimation procedure. We consider the use of sub-optimal non-parametric approach, where the estimation is independent for each coherence time and coherence bandwidth. Within a coherence bandwidth however, the channel is assumed to be frequency flat, allowing a joint estimation across the sub-carriers. We state without proof that under this setting, and using analysis similar to Section III, the effective SNR for a fully digital RX can be obtained as:

$$\gamma_k^{\mathrm{FD}}(\mathbf{h}(t)) = \frac{M E_{\mathrm{s}}^2 |\beta_{0,0}(\mathbf{h})|^2 / (2K+1)}{\beta_{0,0}\mathrm{N}_0 E_{\mathrm{s}}(1 + \frac{1}{K_{\mathrm{coh}}}) + \frac{L(2K+1)}{M K_{\mathrm{coh}}}\mathrm{N}_0^2}, \quad (20)$$

where K_{coh} represents the number of sub-carriers within a coherence bandwidth. Note the similarity of $\gamma_k^{\mathrm{FD}}(\mathbf{h}(t))$ to (16). A similar analytical comparison to hybrid beamforming and digital beamforming with low-resolution ADCs is beyond the scope of this paper, and shall be explored in future work.

V. A Sample Initial Access Protocol

To illustrate the IA advantages of RTAT, here we propose a sample IA protocol. We consider several multi-antenna BSs deployed over a coverage area, using either fully digital or hybrid beamforming. The BSs are assumed to periodically transmit a primary synchronization sequence (PSS), with a common reference tone. We assume an exhaustive beam-scanning procedure [12], where this PSS is sequentially transmitted along different azimuth directions by each BS. Since no beam-search is required at a UE with a RTAT RX, the PSS needs to be transmitted *only once* for each BS azimuth direction. While use of a common reference tone for the PSS may cause reference tone contamination (similar to pilot contamination), the interference can be avoided by using disjoint data sub-carriers to transmit each BS's PSS control information. The control information in the PSS includes the frequency for a second reference tone, which is exclusive and well-separated for each BS. Using this control information, the UE adapts its nBPF to the exclusive reference tone of a chosen BS. It also transmits a return packet to inform the BS the azimuth direction to use. Further control information and the actual data transmission happens using this exclusive tone, with the data streams at each BS occupying same bandwidth.[5] Note that there is no carrier frequency offset, and the symbol synchronization is easy since it is performed with the RX beamforming gain.

While a low PSS count per TX azimuth direction can also be achieved by using wider beams at the RX [13], this reduces

[4] In practice, the value of E_{r} may further be limited by spectral mask regulations, a discussion of which is beyond the scope of this paper.

[5] Note that $\mathcal{K}$ does not need to be symmetric about the reference tone.

the RX beamforming gain and the coverage radius. In contrast, RTAT retains the full RX beamforming gain.

VI. Simulation Results

For simulations, we assume a single-user MIMO system in downlink, having a multi-antenna BS and a UE with a half-wavelength spaced uniform linear array ($M = 32$). We also assume that the BS has aCSI knowledge (via some IA procedure), and it beamforms a single spatial data stream towards the UE, having an RTAT receiver architecture and one down-conversion chain. The BS transmits OFDM symbols with $T_{\rm s} = 2\mu s$, $T_{\rm cp} = 0.2\mu s$ and $f_{\rm c} = 30$ GHz. By including the transmit precoding beam into the channel, we consider a sample *effective* channel impulse response $\mathbf{h}(t)$ between the BS beam and UE with: $L = 3$, $\tau_\ell = 50|\ell - 1|$ ns, $\phi_\ell = (\ell-1)\pi/10$, $\alpha_\ell = (-1)^{\ell+1}\exp\{-|\phi_\ell|/\sigma_\phi\}$, $\sigma_\phi = \pi/10$ and $[a(\phi_\ell,\theta_\ell)]_m = \exp\{{\rm j}\pi\sin(\phi_\ell)\}$. We also assume perfect timing synchronization, and perfect knowledge of $\{\beta_k | k \in \mathcal{K}\}$ at the UE. For this $\mathbf{h}(t)$, the symbol error rate (SER) for (7), obtained by Monte-Carlo simulations, is compared to the analytical SER for the effective channel (13) in Fig. 2. The perfect match between the analytical and simulation results validates our analysis and the effective channel model in (13). Due to the frequency selective fading, we also observe that the SER changes with k.

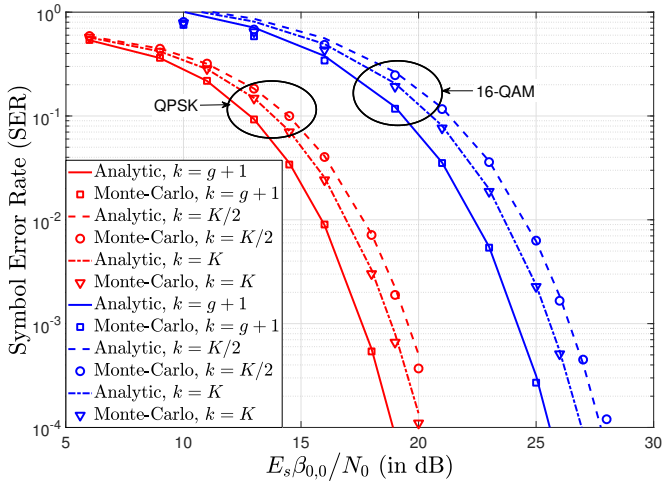


Fig. 2. SER for data streams $k = 1, K/2, K$ of an RTAT receiver with QAM modulation ($K = 50$, $E_{\rm r}$ from Lemma 1, $E_{\rm d}^{(k)} = (E_{\rm s} - E_{\rm r})/|\mathcal{K}|$, $K = 50$, $g = 6$, $\epsilon = 0.1$)

We next compare the iSE of RTAT (15) to the iSE of hybrid beamforming[6] with no phase noise. For channel estimation in hybrid beamforming, in each aCSI coherence time, we assume the UE sequentially scans the received signal at co-prime locations of the antenna array [8]. Using these received signals, the UE uses the nuclear norm minimization (NNM) algorithm [8] to estimate the RX correlation matrix $\sum_\ell |\alpha_\ell|^2 a(\phi_\ell,\theta_\ell)a(\phi_\ell,\theta_\ell)^\dagger$. It's largest eigenvector is then used as the analog beamformer for data reception. We also assume perfect knowledge of $\rm N_0$ and the exact ℓ_2-norm of error for NNM. Since there is very little work on performance

[6]Since UE has one RF chain, this is same as analog beamforming.

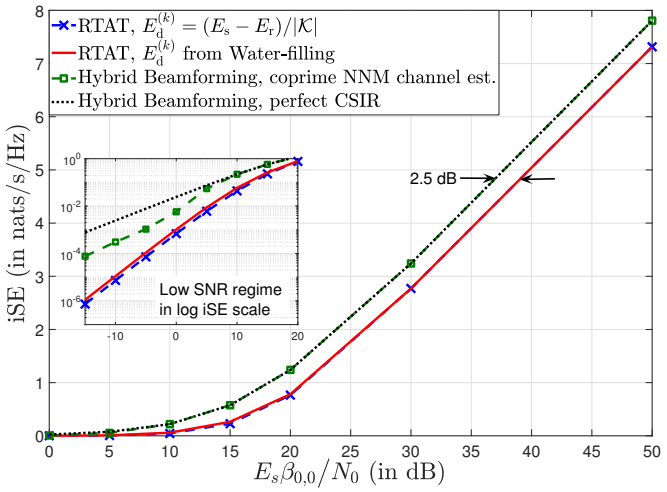


Fig. 3. Comparison of the iSE (without channel estimation overhead) of RTAT and hybrid beamforming (a) For RTAT, $E_{\rm r}$ is from Lemma 1, $K = 512$, $g = 6$ and $\epsilon = 0.5$ (b) For hybrid beamforming, we use $(5, 6)$ co-prime array and assume pilot symbols use only $1/10$-th of the sub-carriers (to reduce noise accumulation).

of low resolution ADCs with channel estimation in sparse wide-band channels [4], we do not include it in the comparison here. The results show that for $\beta_{0,0}E_{\rm s}/{\rm N}_0 \geq 15$ dB, the RTAT suffers from an SNR loss of ≈ 2.5 dB in comparison to hybrid beamforming. Note that since $E_{\rm s}$ is the *total* transmit power and $\beta_{0,0}$ includes the BS beamforming gain, having $\beta_{0,0}E_{\rm s}/{\rm N}_0 \geq 15$ dB is reasonable. Furthermore, performance of equal data power allocation is comparable to water-filling. However these results are subjective and depend on L. Larger values of L intensify the frequency selective fading of RTAT, thereby increasing the SNR loss. Therefore, RTAT is more suited to sparse channels with very few MPCs between the TX beam and RX. Note that the iSE curves in Fig. 3 do not include the channel estimation overhead. As an illustration, assuming that the aCSI coherence time is 10 ms and the BS has 100 antennas and performs exhaustive beam search, channel estimation consumes 2% and 22% of the time-frequency resources for RTAT and hybrid beamforming, respectively. Therefore RTAT can provide significant savings in estimation overhead at the cost of a small drop in SNR.

On the analog hardware front, the RTAT RX requires $2M$ mixers, M LNAs and M nBPF filters (implemented using PLLs), where as, hybrid beamforming requires M analog phase-shifters, and possibly M LNAs and M mixers [17]. A comparative study of the hardware cost or energy consumption of these designs is beyond the scope of this paper. Alternate, and possibly more cost-efficient, designs for an RTAT RX may also exist, which shall be explored in future work.

VII. Conclusion

In this work a novel new approach to reduce the hardware cost of massive MIMO systems, namely reference tone aided transmission is proposed. In RTAT the TX transmits a sinusoidal reference tone along with the data. The RX essentially recovers the reference tone at each antenna and uses it for down-converting the data signals to base-band, thereby,

compensating for the inter-antenna phase shift of each MPC. Note that this operation can be interpreted as *analog channel estimation* of the reference tone. Thus effectively, this work shows that channel estimation of a sinusoidal pilot/reference tone is sufficient to obtain a good RX beamforming gain for a wide-band system with a single spatial data-stream. Further, the analysis suggests that the effective RTAT channel between the TX data and RX can be interpreted as a parallel Gaussian channel with frequency selective fading, where the frequency selectivity arises due to frequency mismatch between the reference tone and data sub-carriers. It is also shown that the SNR expressions for RTAT and for a fully digital RX with non-parametric channel estimation are similar. Finally, the simulation results suggest that, in comparison to hybrid beamforming, RTAT suffers only a small loss in SNR, while providing significant savings in estimation overhead.

APPENDIX A

From (9), the conditional cross-covariance between the noise components at the a and b-th demodulation outputs can be computed as:

$$\begin{aligned}
\mathcal{K}_{a,b}(\mathbf{x},\mathbf{h}) &\overset{(1)}{=} \sum_{\ell_1,\ell_2}\Bigg(\frac{E_{\mathrm{r}}T_{\mathrm{s}}}{2}\alpha_{\ell_1}\alpha_{\ell_2}^{*}e^{\mathrm{j}2\pi f_{\mathrm{c}}(\tau_{\ell_1}-\tau_{\ell_2})}\mathbf{a}(\phi_{\ell_1},\theta_{\ell_1})^{\dagger} \\
&\quad\times\mathbb{E}_{\mathbf{N}}\{\mathbf{N}[a]\mathbf{N}[b]^{\dagger}\}\mathbf{a}(\phi_{\ell_2},\theta_{\ell_2})\Bigg) \\
&+\sum_{\ell_1,\ell_2}\Bigg(\sum_{\bar{k},\ddot{k}=-g}^{g}\frac{\alpha_{\ell_1}\alpha_{\ell_2}^{*}T_{\mathrm{s}}e^{-\mathrm{j}2\pi(f_{\mathrm{c}}+f_{\bar{k}}+f_a)\tau_{\ell_1}}e^{\mathrm{j}2\pi(f_{\mathrm{c}}+f_{\ddot{k}}+f_b)\tau_{\ell_2}}}{2} \\
&\quad\times x_{\bar{k}+a}x_{\ddot{k}+b}^{*}\mathbf{a}(\phi_{\ell_2},\theta_{\ell_2})^{\dagger}\mathbb{E}_{\mathbf{N}}\{\hat{\mathbf{N}}[\ddot{k}]\hat{\mathbf{N}}[\bar{k}]^{\dagger}\}\mathbf{a}(\phi_{\ell_1},\theta_{\ell_1})\Bigg) \\
&+\frac{T_{\mathrm{s}}^2}{4}\sum_{\bar{k}=-g}^{g}\sum_{\ddot{k}=-g}^{g}\mathbb{E}_{\mathbf{N}}\Big\{\hat{\mathbf{N}}[\bar{k}]^{\dagger}\mathbf{N}[\bar{k}+a]\mathbf{N}[\ddot{k}+b]^{\dagger}\hat{\mathbf{N}}[\ddot{k}]\Big\} \\
&\overset{(2)}{=}\sum_{\ell=1}^{L}|\alpha_\ell|^2 M\mathrm{N}_0\Bigg[E_{\mathrm{r}}\delta_{a,b}+e^{\mathrm{j}2\pi(f_b-f_a)\tau_\ell}\Bigg(x_a x_b^{*}(1-\epsilon) \\
&\quad+\sum_{k=-g}^{g}x_{k+a}x_{k+b}^{*}\epsilon\Bigg)\Bigg]+\delta_{a,b}M\Bigg((1-\epsilon)\mathrm{N}_0^2+\sum_{k=-g}^{g}\epsilon N_0^2\Bigg) \qquad (21)
\end{aligned}$$

where $\overset{(1)}{=}$ follows from the fact that odd moments of $\mathbf{N}[k]$ are 0; and $\overset{(2)}{=}$ follows from (4)–(5), by defining $\delta_{a,b}=1$ if $a=b$ and 0 otherwise, and the fact the array response vectors $\mathbf{a}(\phi_\ell,\theta_\ell)$ are mutually orthogonal. Using similar steps, the conditional pseudo cross-covariance between the noise components at the a and b-th demodulation outputs can be shown to be given by (11).

REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, pp. 186–195, February 2014.

[2] X. Zhang, A. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Transactions on Signal Processing*, vol. 53, pp. 4091–4103, Nov 2005.

[3] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, 2017.

[4] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 87–100, Jan 2017.

[5] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath, "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 2274–2287, April 2017.

[6] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 831–846, Oct 2014.

[7] V. V. Ratnam, A. F. Molisch, N. Rabeah, F. Alawwad, and H. Behairy, "Diversity versus training overhead trade-off for low complexity switched transceivers," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec 2016.

[8] S. Haghighatshoar and G. Caire, "Massive MIMO channel subspace estimation from low-dimensional projections," *IEEE Transactions on Signal Processing*, vol. 65, pp. 303–318, Jan 2017.

[9] A. Wadhwa and U. Madhow, "Near-coherent QPSK performance with coarse phase quantization: A feedback-based architecture for joint phase/frequency synchronization and demodulation," *IEEE Transactions on Signal Processing*, vol. 64, pp. 4432–4443, Sept 2016.

[10] T. C. Zhang, C. K. Wen, S. Jin, and T. Jiang, "Mixed-ADC massive MIMO detectors: Performance analysis and design optimization," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 7738–7752, Nov 2016.

[11] M. Akdeniz, Y. Liu, M. Samimi, S. Sun, S. Rangan, T. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 1164–1179, June 2014.

[12] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmwave cellular networks," *IEEE Communications Magazine*, vol. 54, pp. 40–47, November 2016.

[13] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, "Initial access in millimeter wave cellular systems," *IEEE Transactions on Wireless Communications*, vol. 15, pp. 7926–7940, Dec 2016.

[14] J. Lee, G. T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Transactions on Communications*, vol. 64, pp. 2370–2386, June 2016.

[15] H. Qiao and P. Pal, "Gridless line spectrum estimation and low-rank Toeplitz matrix compression using structured samplers: A regularization-free approach," *IEEE Transactions on Signal Processing*, vol. 65, pp. 2221–2236, May 2017.

[16] A. A. Nasir, S. Durrani, H. Mehrpouyan, S. D. Blostein, and R. A. Kennedy, "Timing and carrier synchronization in wireless communication systems: a survey and classification of research in the last 5 years," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, p. 180, Aug 2016.

[17] L. Kong, *Energy-Efficient 60GHz Phased-Array Design for Multi-Gb/s Communication Systems*. PhD thesis, EECS Department, University of California, Berkeley, Dec 2014. http://www2.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-191.html.

[18] D. L. Goeckel and Q. Zhang, "Slightly frequency-shifted reference ultra-wideband (UWB) radio," *IEEE Transactions on Communications*, vol. 55, pp. 508–519, March 2007.

[19] V. V. Ratnam, A. F. Molisch, A. Alasaad, F. Alawwad, and H. Behairy, "Bit and power allocation in QAM capable multi-differential frequency-shifted reference UWB radio," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, Dec 2017.

[20] Z. Liu, J. Y. Kim, D. S. Wu, D. J. Richardson, and R. Slavk, "Homodyne OFDM with optical injection locking for carrier recovery," *Journal of Lightwave Technology*, vol. 33, pp. 34–41, Jan 2015.

[21] O. El Ayach, R. Heath, S. Abu-Surra, S. Rajagopal, and Z. Pi, "The capacity optimality of beam steering in large millimeter wave MIMO systems," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 100–104, June 2012.

[22] L. Piazzo and P. Mandarini, "Analysis of phase noise effects in OFDM modems," *IEEE Transactions on Communications*, vol. 50, pp. 1696–1705, Oct 2002.