

Approximate Gram-Matrix Interpolation for Wideband Massive MU-MIMO Systems

Charles Jeon, Zequn Li, and Christoph Studer

Abstract—Linear and non-linear data-detection and precoding algorithms for wideband massive multi-user (MU) multiple-input multiple-output (MIMO) wireless systems that rely on orthogonal frequency-division multiplexing (OFDM) or single-carrier frequency-division multiple access (SC-FDMA) often require computation of the Gram matrix for each active subcarrier. Computing the Gram matrix for each active subcarrier, however, results in excessively high computational complexity. In this paper, we propose novel, approximate algorithms that significantly reduce the complexity of Gram-matrix computation by simultaneously exploiting channel hardening and correlation across subcarriers. We show analytically that a small fraction of Gram-matrix computations in combination with approximate interpolation schemes are sufficient to achieve near-optimal error-rate performance at low computational complexity in massive MU-MIMO systems. We furthermore demonstrate that our approximate interpolation algorithms are more robust against channel-estimation errors than exact Gram-matrix interpolation algorithms that require high computational complexity.

Index Terms—Equalization, Gram matrix, interpolation, massive MU-MIMO, orthogonal frequency-division multiplexing (OFDM), precoding, single-carrier frequency-division multiple access (SC-FDMA).

I. INTRODUCTION

MASSIVE multi-user (MU) multiple-input multiple-output (MIMO) is a key technology component in fifth-generation (5G) wireless systems [1], [2]. The idea of massive MU-MIMO is to equip the base-station (BS) with hundreds of antenna elements while serving tens of user equipments (UEs) in the same time-frequency resource by means of spatial multiplexing. Such large antenna arrays enable fine-grained beamforming in the uplink (UEs transmit to the BS) and in the downlink (BS transmits to the UEs), which offers superior spectral efficiency compared to traditional, small-scale MIMO technology that use only a few antennas at the BS.

In the uplink, linear data-detection algorithms that rely on linear minimum-mean square error (MMSE) equalization or

zero-forcing (ZF) are known to achieve near-optimal error-rate performance in realistic massive MU-MIMO systems with a finite number of transmit antennas [1], [3], [4]. Non-linear data-detection algorithms [5]–[7] have recently been shown to outperform linear methods in systems where the number of UEs is comparable to the number of BS antennas. Most of these linear and non-linear data-detection algorithms entail high computational complexity, often dominated by the computation of the *Gram matrix* $\mathbf{G} = \mathbf{H}^H \mathbf{H}$ [8], [9]. Here, $\mathbf{H} \in \mathbb{C}^{B \times U}$ is the (uplink) channel matrix, B is the number of BS antennas, and U is the number of (single-antenna) users. The computational complexity is orders-of-magnitude higher in wideband systems that use orthogonal frequency-division multiplexing (OFDM) or single-carrier frequency-division multiple access (SC-FDMA), in which a Gram matrix must be computed for each active subcarrier (i.e., subcarriers used for pilots or data transmission) [8]. For example, Gram matrix computation requires more than $2 \times$ higher complexity than data detection for a 128 BS antenna 16 UE antenna MU-MIMO system [9, Tbl. 4.2]. In the massive MU-MIMO downlink, precoding is necessary to focus the transmit energy towards the UEs and to mitigate multi-user interference [1]. In wideband systems, the complexity of linear precoding algorithms is—analogously to the uplink—typically dominated by Gram matrix computation on the active subcarriers.

While some data-detection and precoding algorithms have been proposed that avoid the computation of the Gram matrix altogether (see, e.g., [10]–[12]), these methods do not allow the re-use of intermediate results in time-division duplexing (TDD) systems. Specifically, the Gram matrix and its inverse cannot be re-used in the uplink (for equalization) and downlink (for precoding), which would significantly lower the computational complexity. Hence, such algorithms inevitably perform redundant computations during data-detection and precoding, which leads to inefficient transceiver designs.

A. Interpolation-Based Matrix Computations

In practical wideband communication systems, e.g., building upon IEEE 802.11n [13] and 3GPP-LTE [14], the channel's delay spread is often substantially smaller than the number of active subcarriers. Hence, the channel coefficients are correlated across subcarriers. This property can be exploited to reduce the computational complexity of commonly-used matrix computations required in multi-antenna systems. More specifically, the papers [15]–[18] avoid a brute-force approach in traditional, small-scale, and point-to-point MIMO-OFDM systems by using exact interpolation-based algorithms for

Copyright © 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

C. Jeon was with School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, and is now with Apple Inc., San Diego, CA; e-mail: cj339@cornell.edu

Z. Li was a visiting researcher at the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY and is now with Google LLC; e-mail: li-zq13@mails.tsinghua.edu.cn

C. Studer is with the School of Electrical and Computer Engineering at Cornell Tech, New York and Cornell University, Ithaca, NY, USA; e-mail: studer@cornell.edu; website: vip.ece.cornell.edu

A simulation framework to reproduce the results of this paper will be made available on GitHub after (possible) acceptance.

matrix inversion and QR factorization. While a few hardware designs [19], [20] have demonstrated the efficacy of such exact interpolation methods in small-scale MIMO systems, their complexity does not scale well to wideband massive MU-MIMO systems with hundreds of BS antennas, tens of users, and thousands of subcarriers. In fact, 3GPP specifications on New Radio (NR) access technology shows that the number of active subcarriers is 3300 or 6600 in Release 15 [21], [22]. In addition, the impact of imperfect channel-state information (CSI) and antenna correlation on such exact, interpolation-based matrix computation algorithms is routinely ignored, but significantly affects their performance in practical scenarios (see Section VI for a detailed discussion).

B. Contributions

Inspired by exact, interpolation-based matrix computation algorithms put forward in [15]–[18] for small-scale wideband MIMO systems, we propose novel algorithms for *approximate* Gram matrix computation in massive MU-MIMO systems. We start by establishing the minimum number of Gram matrix base-points that are required for exact interpolation. We then show that channel-hardening in massive MU-MIMO provides a new opportunity for *approximate* interpolation schemes that achieve near-exact error-rate performance, even with strong undersampling in the frequency domain. We provide analytical results that characterize the approximation errors of the proposed interpolation methods depending on the channel's delay spread and the antenna configuration. Furthermore, we derive exact mean-squared error (MSE) expressions of our approximate interpolation algorithms for imperfect CSI and BS-side antenna correlation. We also characterize the trade-offs between computational complexity and error-rate performance in realistic massive MU-MIMO-OFDM systems, and we demonstrate the robustness of our approximate interpolation methods under realistic channel conditions with imperfect CSI and BS-antenna correlation.

C. Relevant Prior Art

Data detection and precoding for small-scale, single- and multi-carrier MIMO systems is a well-studied topic; see e.g. [4], [23]–[26] and the references therein. Data detection and precoding algorithms for massive MIMO systems have been proposed in, e.g., [5], [6], [8], [27]–[29], which leverage the fact that the Gram matrix is diagonally dominant [30]. However, all of these results (i) do not exploit specific properties of massive MU-MIMO systems and (ii) ignore the fact that time-division duplexing (TDD)-based systems must perform data-detection *and* precoding, and hence, can potentially re-use intermediate results (such as the Gram matrix) to reduce complexity. In contrast, our methods exploit the specifics of massive MU-MIMO, namely channel hardening, and enable a re-use of the computations carried out in the uplink for downlink precoding.

The report [31] proposed an approximate interpolation-based ZF-based equalizer for a wideband massive MU-MIMO testbed. Therein, the authors interpolate the *inverse* of the Gram matrix. While simulation results in [31] show that the proposed method works well in practice, no theoretical results

have been provided. In contrast, we use approximate methods to interpolate the Gram matrix, and we provide analytical expressions that provide a solid foundation of approximate interpolation methods in massive MU-MIMO systems.

D. Notation

Lowercase and uppercase boldface letters stand column vectors and matrices, respectively. The transpose, Hermitian, and pseudo-inverse of the matrix $\mathbf{A}$ are denoted by $\mathbf{A}^T$, $\mathbf{A}^H$, and $\mathbf{A}^\dagger$, respectively. We use $[\mathbf{A}]_{m,n}$ to represent the m th row and n th column entry of the matrix $\mathbf{A}$. Sets are designated by uppercase calligraphic letters, and $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. The complex conjugate of $a \in \mathbb{C}$ is a^* . The indicator function is defined as $\mathbb{I}(a)$, where $\mathbb{I}(a) = 1$ if a is true and 0 otherwise. A complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

E. Paper Outline

The rest of the paper is organized as follows. Section II introduces the necessary prerequisites. Section III proposes our exact and approximate interpolation-based Gram matrix computation algorithms. Section IV and Section V provide an analysis of the approximation error and complexity for the proposed approximate interpolation methods, respectively. Section VI shows numerical results. We conclude in Section VII.

II. PREREQUISITES

We start by summarizing the considered wideband massive MU-MIMO system and channel model. We then outline computationally-efficient ways for linear data detection and precoding that make use of the Gram matrix.

A. System Model

Without loss of generality, we focus on the uplink¹ of a wideband massive MU-MIMO system with B base-station antennas, U single-antenna UEs (with $U \ll B$), and W subcarriers. For each active subcarrier $\omega \in \Omega$ with Ω containing the indices of the active (data and pilot) subcarriers, we model the received frequency-domain (FD) signal as follows:

$$\mathbf{y}_\omega = \mathbf{H}_\omega \mathbf{s}_\omega + \mathbf{n}_\omega. \quad (1)$$

Here, $\mathbf{H}_\omega \in \mathbb{C}^{B \times U}$ is the FD channel matrix, $\mathbf{s}_\omega \in \mathbb{C}^U$ is the transmit vector, and $\mathbf{n}_\omega \in \mathbb{C}^B$ models additive noise. The FD input-output relation in (1) is able to model both OFDM and SC-FDMA systems. For OFDM systems, the entries of the transmit vector $\mathbf{s}_\omega$ are taken from a discrete constellation set $\mathcal{O}$ (e.g., 16-QAM); in SC-FDMA systems, the constellation points are assigned in the time-domain and the resulting vectors are transformed into the FD to obtain the transmit vectors $\mathbf{s}_\omega$. See, e.g., [8], for details on SC-FDMA transmission.

¹By assuming channel reciprocity [1], our results directly apply to the downlink in which $\mathbf{H}^T$ is the downlink channel matrix. Once all of the Gram matrices have been computed, they can be re-used in the downlink for linear (e.g., zero-forcing or Wiener filter) precoding. See Section II-C for the details.

B. Wideband Channel Model

In wideband MIMO multicarrier systems, the FD channel matrices $\mathbf{H}_\omega$, $\omega = 0, \dots, W-1$, are directly related to the time-domain (TD) matrices $\hat{\mathbf{H}}_\ell \in \mathbb{C}^{B \times U}$, where $\ell = 0, \dots, W-1$ are the channel “taps.” We first introduce the model used to characterize antenna correlation at the BS side² which occurs in the TD. Specifically, we use the standard correlation model from [4] and express the ℓ th TD channel matrix as follows:

$$\hat{\mathbf{H}}_\ell = \mathbf{R}^{1/2} \hat{\mathbf{H}}_\ell^{\text{uncor}}. \quad (2)$$

Here, $\hat{\mathbf{H}}_\ell^{\text{uncor}}$ represents an uncorrelated TD channel matrix and $\mathbf{R} \in \mathbb{C}^{B \times B}$ is a correlation matrix that contains ones on the main diagonal and $\delta \in \mathbb{R}$ on the off-diagonals. We allow δ to be either real positive or negative as long as $\delta^2 \leq 1$. We rewrite the correlation matrix as $\mathbf{R} = (1 - \delta)\mathbf{I}_B + \delta\mathbf{1}_B$, where $\mathbf{I}_B$ and $\mathbf{1}_B$ is the $B \times B$ identity and all-ones matrix, respectively. We note that in the absence of receive-side correlation, i.e., $\delta = 0$, we have $\mathbf{R} = \mathbf{I}_B$ and $\hat{\mathbf{H}}_\ell = \hat{\mathbf{H}}_\ell^{\text{uncor}}$.

In order to take into account the practically-relevant case of imperfect CSI at the BS, we assume that the FD channel matrices $\mathbf{H}_\omega$, $\omega = 0, \dots, W-1$, are obtained from the TD matrices $\hat{\mathbf{H}}_\ell \in \mathbb{C}^{B \times U}$, $\ell = 0, \dots, W-1$, via the discrete Fourier transform [32] as follows:³

$$\mathbf{H}_\omega = \sum_{\ell=0}^{L-1} \hat{\mathbf{H}}_\ell \exp\left(-\frac{j2\pi\omega\ell}{W}\right) + \sigma \mathbf{E}_\omega. \quad (3)$$

Here, the matrix $\mathbf{E}_\omega \in \mathbb{C}^{B \times U}$ models the channel estimation error on subcarrier ω and the parameter $\sigma \in \mathbb{R}^+$ determines the intensity of channel-estimation errors; $\sigma = 0$ corresponds to the case for perfect CSI. We assume that the entries of the matrix $\mathbf{E}_\omega$ are i.i.d. (across entries and subcarriers) circularly-symmetric complex Gaussian with unit variance. Equation (3) relies on the assumption that at most $L \leq W$ of the first channel taps are non-zero (or dominant) and the remaining ones are zero (or insignificant), i.e., $\hat{\mathbf{H}}_\ell = \mathbf{0}_{B \times U}$ for $\ell = L, \dots, W-1$. In practical OFDM and SC-FDMA systems, the maximum number of non-zero channel taps should not exceed the cyclic prefix length (assuming perfect synchronization). Hence, it holds that $L \ll W$ for many practical scenarios and for most standards, such as IEEE 802.11n [13] or 3GPP-LTE [14].⁴

C. Linear and Non-linear Data Detection and Precoding

In the massive MU-MIMO uplink, linear and non-linear data detection methods were shown to achieve near-optimal error-rate performance [1], [6]. For linear minimum mean-square error (MMSE) equalization with complex transmit constellations, one first computes the Gram matrix $\mathbf{G}_\omega = \mathbf{H}_\omega^H \mathbf{H}_\omega$

²In massive MU-MIMO systems, the UEs signals are likely uncorrelated as they are spatially well-separated over potentially large cells or UE scheduling avoids correlated UEs; in contrast, the antennas at the BS are typically confined to a small area, which increases the potential for receive-side correlation.

³One could improve the channel estimates $\mathbf{H}_\omega$ by exploiting the fact that only the first L taps of $\mathbf{H}_\omega$ are active in the TD. An analysis of such channel estimation algorithms is left for future work.

⁴We note that the cyclic prefix length is designed to be no smaller than largest effective delay spread observed under typical operation conditions. Thus, although the exact value of L may be unknown at the receiver, it can be set to the cyclic prefix length.

and then, computes an estimate of the transmit vector as $\hat{\mathbf{s}}_\omega = (\mathbf{G}_\omega + \mathbf{I}_{E_s}^{N_0})^{-1} \mathbf{H}_\omega^H \mathbf{y}_\omega$, where N_0 and E_s stand for the noise variance and average energy per transmit symbol, respectively. For non-linear data detectors, such as the one in [6], one can operate directly on the Gram matrix $\mathbf{G}_\omega$ and the matched filter $\mathbf{H}_\omega^H \mathbf{y}_\omega$ without any performance loss [33].

For such algorithms, a direct computation of the Gram matrix $\mathbf{G}_\omega$ for every subcarrier results in excessively high complexity. In fact, even by exploiting symmetries, $2BU^2$ real-valued multiplications are required, which is more than 16,000 multiplications per subcarrier for a system with 128 BS antennas and 8 UEs. Furthermore, the hardware design for massive MU-MIMO data detection in [8] confirms this observation and shows that Gram-matrix computation dominates the overall hardware complexity and power consumption by at least $2\times$.

In the downlink, ZF of Wiener filter precoding are most commonly used [1]. For example, ZF precoding computes $\mathbf{x}_\omega = \mathbf{H}_\omega^H \mathbf{G}_\omega^{-1} \mathbf{s}_\omega$, where $\mathbf{x}_\omega$ is the B -dimensional transmit signal and $\mathbf{s}_\omega$ the data vector. If the Gram matrix $\mathbf{G}_\omega$ has been precomputed for equalization in the uplink phase, then it can be re-used for ZF precoding in the downlink to minimize recurrent operations. Hence, to minimize the overall complexity of equalization and precoding, efficient ways to compute the Gram matrix $\mathbf{G}_\omega$ on all active subcarriers $\omega \in \Omega$ are required.

III. INTERPOLATION-BASED GRAM MATRIX COMPUTATION

We now discuss exact and approximate interpolation-based methods for low-complexity Gram matrix computation. Note that exact Gram matrix interpolation assumes perfect CSI, i.e., we assume $\sigma = 0$. In Section IV, we will relax the perfect-CSI assumption and analyze the performance of exact and approximate interpolation methods with imperfect CSI.

As a result of (3), the Gram matrices in the FD are given by

$$\mathbf{G}_\omega = \sum_{\ell=0}^{L-1} \sum_{\ell'=0}^{L-1} \hat{\mathbf{H}}_\ell^H \hat{\mathbf{H}}_{\ell'} \exp\left(\frac{j2\pi\omega(\ell - \ell')}{W}\right) \quad (4)$$

for $w = 0, 1, \dots, W-1$. Given the FD channel matrices $\mathbf{H}_\omega$ for all active subcarriers $\omega \in \Omega$, a straightforward “brute-force” approach simply computes $\mathbf{G}_\omega = \mathbf{H}_\omega^H \mathbf{H}_\omega$ for each active subcarrier $\omega \in \Omega$. In order to reduce the complexity of such a brute-force approach, we next discuss exact and approximate Gram-matrix interpolation methods that take advantage of the facts that (i) the channel matrices (and hence, the Gram matrices) are correlated across subcarriers if $L < W$ and (ii) massive MU-MIMO benefits from channel hardening [1], [2].

A. Exact Gram-Matrix Interpolation

The Gram matrix $\mathbf{G}_\omega$ in (4) is a Laurent polynomial matrix in the variable $x_\omega = \exp(j2\pi\omega/W)$; we refer the reader to [18] for more details on Laurent polynomial matrices. Hence, we can establish the following result for exact Gram-matrix interpolation; a short proof is given in Appendix A.

Lemma 1. *The Gram matrices $\mathbf{G}_\omega$ in (4) for all subcarriers $\omega = 0, \dots, W$ are fully determined by $2L - 1$ distinct and non-zero Gram-matrix base-points.*

Consequently, one can interpolate all of the Gram matrices *exactly* from only $2L - 1$ distinct and non-zero Gram-matrix base-points that have been computed explicitly. To perform exact interpolation, we first define a set of base points $\mathcal{P} \subset \Omega$ that contains $|\mathcal{P}| \geq 2L - 1$ distinct subcarrier indices. We denote the k th base-point index as p_k , where $k = 0, \dots, |\mathcal{P}| - 1$, and the set of all base-point indices as $\mathcal{P} = \{p_0, \dots, p_{|\mathcal{P}|-1}\}$. For each subcarrier index in the base-point set $\mathcal{P}$, we then explicitly compute $|\mathcal{P}| \geq 2L - 1$ Gram matrices $\mathbf{G}_\omega = \mathbf{H}_\omega^H \mathbf{H}_\omega$, $\omega \in \mathcal{P}$, and perform entry-wise interpolation for the gram matrices $\mathbf{G}_\omega$ on all remaining active subcarriers $\omega \in \Omega \setminus \mathcal{P}$.

The exact interpolation procedure for each entry is as follows. For a fixed entry (m, n) , we define the vector $\mathbf{g}_\mathcal{P} \in \mathbb{C}^{|\mathcal{P}|}$, which is constructed from the (m, n) entries $[\mathbf{G}_\omega]_{m,n}$ taken from base-points $\omega \in \mathcal{P}$, i.e., $\mathbf{g}_\mathcal{P} = [[\mathbf{G}_{p_0}]_{m,n} \cdots [\mathbf{G}_{p_{|\mathcal{P}|-1}}]_{m,n}]^T$. Then, the vector $\mathbf{g}_{\Omega \setminus \mathcal{P}} \in \mathbb{C}^{|\Omega| - |\mathcal{P}|}$ that contains the entry (m, n) for all remaining Gram matrices $\mathbf{G}_\omega$, $\omega \in \Omega \setminus \mathcal{P}$, is given by

$$\mathbf{g}_{\Omega \setminus \mathcal{P}} = \mathbf{F}_{\Omega \setminus \mathcal{P}, L} (\mathbf{F}_{\mathcal{P}, L}^\dagger \mathbf{g}_\mathcal{P}). \quad (5)$$

Here, $\mathbf{F}_{\mathcal{P}, L}$ represents a $|\mathcal{P}| \times (2L - 1)$ matrix where we take the $|\mathcal{P}|$ rows indexed by $\mathcal{P}$ and the first L and last $L - 1$ columns from the W -point discrete Fourier transform (DFT) matrix; the entries of the DFT matrix are defined as $[\mathbf{F}]_{m,n} = \frac{1}{\sqrt{W}} \exp(-j2\pi(m-1)(n-1))$. Similarly, $\mathbf{F}_{\Omega \setminus \mathcal{P}, L}$ represents a $(|\Omega| - |\mathcal{P}|) \times (2L - 1)$ matrix where we take $|\Omega| - |\mathcal{P}|$ rows indexed by $\Omega \setminus \mathcal{P}$ and the first L and last $L - 1$ columns from the W -point DFT matrix. In words, the exact interpolation method in (5) first computes the $2L - 1$ TD Gram-matrix entries and then, transforms these elements into the frequency domain via the DFT. See [15]–[18] for additional details on other *exact* interpolation methods developed for MIMO systems.

Although the method in (5) is able to *exactly* interpolate the Gram matrix across all W subcarriers, it is in many situations not practical due to the high complexity of computing $\mathbf{F}_{\mathcal{P}, L}^\dagger \mathbf{g}_\mathcal{P}$. If, however, one can sample the base points uniformly over all W subcarriers, the complexity of matrix inversion can be reduced significantly. Unfortunately, this approach is often infeasible in practice due to the presence of guard-band constraints in OFDM-based or SC-FDMA-based standards [13], [14]. Another issue of exact interpolation methods, such as the ones in [15]–[18] and ours in (5), is that they generally assume perfect CSI and no spatial correlation. As we will show in Section VI-C, imperfect CSI results in poor interpolation performance—this is due to the fact that the matrix $\mathbf{F}_{\mathcal{P}, L}$ is typically ill-conditioned, especially when sampling Gram-matrices close to the minimum number of $2L - 1$ base points.

We next propose two approximate interpolation schemes that not only require lower complexity than a brute-force approach or exact interpolation in (5), but also approach the performance of a brute-force approach in massive MU-MIMO systems and are robust to channel-estimation errors.

B. Approximate Gram-Matrix Interpolation

We consider the following two approximate Gram-matrix interpolation methods illustrated in Fig. 1.

0th Order Interpolation: We select a set of $|\mathcal{P}|$ distinct base-points with $\mathcal{P} = \{p_0, \dots, p_{|\mathcal{P}|-1}\} \subset \Omega$. We explicitly compute



Fig. 1. Illustration of 0th and 1st order interpolation for the entry $[\mathbf{G}_\omega]_{1,1}$ across subcarriers. We explicitly compute the Gram matrix for every fifth subcarrier index ω and interpolate the remaining matrices. We assume a 128 BS antenna, 8 user massive MU-MIMO system with $W = 2048$ subcarriers, a delay spread of $L = 144$, and perfect CSI without BS-antenna correlation.

$\mathbf{G}_p = \mathbf{H}_p^H \mathbf{H}_p$ on these base points and perform 0th order (or nearest-neighbor) interpolation for the remaining subcarriers in the set $\Omega \setminus \mathcal{P}$ according to:

$$\tilde{\mathbf{G}}_\omega = \mathbf{G}_p, \quad p = \arg \min_{\tilde{p} \in \mathcal{P}} |\tilde{p} - \omega|, \quad \forall \omega \in \Omega \setminus \mathcal{P}. \quad (6)$$

In words, we set the interpolated Gram matrix $\tilde{\mathbf{G}}_\omega$ equal to the nearest Gram matrix that has been computed explicitly for one of the neighboring base points.

1st Order Interpolation: Analogously to the 0th order interpolation method, we explicitly compute $\mathbf{G}_p = \mathbf{H}_p^H \mathbf{H}_p$ on a selected set of base-points $p \in \mathcal{P}$. Then, for each target subcarrier $\omega \in \Omega \setminus \mathcal{P}$ we pick two nearest base-points p_k and p_{k+1} , i.e., $p_k \leq \omega \leq p_{k+1}$, and perform entry-wise linear interpolation according to

$$\tilde{\mathbf{G}}_\omega = \lambda_\omega \mathbf{G}_{p_k} + (1 - \lambda_\omega) \mathbf{G}_{p_{k+1}}, \quad \omega \in \Omega \setminus \mathcal{P}, \quad (7)$$

where $\lambda_\omega = (p_{k+1} - \omega) / (p_{k+1} - p_k)$ and $p_k < p_{k+1}$.

IV. ANALYSIS OF APPROXIMATION ERROR

We now analyze the approximation error associated with the approximate interpolation schemes from Section III-B. We use $\mathbf{G}_\omega$ to represent the Gram matrices that have been computed exactly and $\tilde{\mathbf{G}}_\omega$ to represent the Gram matrices that are obtained via approximate interpolation. Evidently, the exact interpolation scheme in Section III-A entails no approximation error.

A. Mean-Square-Error of Approximate Interpolation

We study the mean-squared error (MSE) on each entry (m, n) for the ω -th subcarrier, which we define as follows:

$$o\text{-MSE}_\omega^{(m,n)} \triangleq \mathbb{E} \left[|[\tilde{\mathbf{G}}_\omega]_{m,n} - [\mathbf{G}_\omega]_{m,n}|^2 \right]. \quad (8)$$

Here, o represents the order of interpolation, i.e., we have either $o = 0$ or $o = 1$. Our results make extensive use of the scaled Fejér kernel [34] given by

$$f_L(\phi) = L^{-2} \frac{1 - \cos(L\phi)}{1 - \cos(\phi)} = L^{-2} \frac{\sin^2(L\phi/2)}{\sin^2(\phi/2)} \quad (9)$$

and rely on the following key properties of this kernel; the proof is given in Appendix B.

Lemma 2. *The scaled Fejér kernel (9) is non-negative, bounded from above by one, and monotonically decreasing in ϕ for $\phi \in [0, 2\pi/L]$ with $L > 1$.*

B. MSE of 0th Order Interpolation

The following result precisely characterizes the MSE of 0th order interpolation for imperfect CSI as in (3) and BS-antenna correlation as in (2); the proof is given in Appendix C.

Theorem 1. *Let the entries of the TD matrices $\hat{\mathbf{H}}_\ell^{\text{uncor}}$, $\ell = 1, \dots, L$, be distributed $\mathcal{CN}(0, 1/(BL))$ per complex entry. Assume that the off-diagonal entries of the receive correlation matrix are δ , the variance of the channel estimation error σ , and $p \in \mathcal{P}$ is the closest base point to the target subcarrier ω . Then, for any (m, n) entry of the Gram matrix $\mathbf{G}_\omega$, the MSE for the 0th order interpolation method in (6) is given by*

$$0\text{-MSE}_\omega = \varepsilon_{\text{CSI}} + \frac{2}{B}(1 + \varepsilon_{\text{cor}}) \left(1 - f_L \left(\frac{2\pi}{W}(p - \omega) \right) \right), \quad (10)$$

where we use the definitions

$$\varepsilon_{\text{CSI}} = 2\sigma^2(2 + B\sigma^2) \quad \text{and} \quad \varepsilon_{\text{cor}} = \delta^2(B - 1).$$

From this result, we see that, as the number of BS antennas B increases, ε_{CSI} increases quadratically with respect to σ^2 . For perfect CSI, i.e., $\sigma = 0$ so $\varepsilon_{\text{CSI}} = 0$, the MSE for 0th order interpolation decreases with an increasing number of BS antennas B as $\frac{\partial}{\partial B} 0\text{-MSE}_\omega < 0$, if $\delta^2 < 1$. Also, we note in the case for non-zero correlation, i.e., $\delta \neq 0$, the MSE for 0th order interpolation is amplified (compared to that with no correlation) by a factor of $1 + \delta^2(B - 1)$. Furthermore, we observe that the MSE is independent of the entry of the Gram matrix (i.e., the MSE is identical for the diagonal as well as off-diagonal entries); this is a consequence of the i.i.d. assumption of the TD channel matrices $\hat{\mathbf{H}}_\ell^{\text{uncor}}$.

To gain additional insight on 0th order interpolation in the large BS-antenna limit, we have the following result.

Corollary 1. *Assume the conditions in Theorem 1 are met, and let $\sigma^2 \rightarrow 0$. Then, as $B \rightarrow \infty$, the MSE of 0th order interpolation is given by*

$$\lim_{B \rightarrow \infty} 0\text{-MSE}_\omega = 2\delta^2 \left(1 - f_L \left(\frac{2\pi}{W}(p - \omega) \right) \right).$$

Corollary 1 demonstrates that in the large-BS antenna limit, the MSE of 0th order interpolation is zero across all subcarriers if and only if the BS antennas are uncorrelated, i.e., $\delta = 0$. For $\delta \neq 0$, the MSE depends on the distance between the nearest base-point and the target subcarrier.

While the MSE expression in (10) is exact, it does not provide much intuition. We define the following quantity that enables us to further analyze the MSE in (10).

Definition 1. *The maximum distance between any subcarrier ω and the nearest base point is given by:*

$$d_{\max} = \max_{k \in \mathcal{P}} \left\lfloor \frac{p_{k+1} - p_k}{2} \right\rfloor. \quad (11)$$

With the maximum distance $d_{\max}$ for a given set of base points $\mathcal{P}$, Corollary 2 shows that the 0th order approximation can be bounded from above using simple analytic expressions; the proof is given in Appendix D.

Corollary 2. *Let $d_{\max}$ be the maximum distance in (11) and assume the conditions in Theorem 1 are met. Then, the maximum MSE of 0th order interpolation over all active subcarriers $\omega \in \Omega$ is bounded by:*

$$\max_{\omega \in \Omega} \{0\text{-MSE}_\omega\} \leq \begin{cases} \varepsilon_{\text{CSI}} + \frac{2}{B}(1 + \varepsilon_{\text{cor}}), & d_{\max} \geq \frac{W}{L} \\ \varepsilon_{\text{CSI}} + \frac{2(1 + \varepsilon_{\text{cor}})}{B} \left(1 - f_L \left(\frac{2\pi}{W} d_{\max} \right) \right), & d_{\max} < \frac{W}{L}. \end{cases}$$

Corollary 2 implies that regardless of small or large maximum distance $d_{\max}$, the MSE given by the 0th order approximation always decreases with the number of BS antennas B if $\delta = 0$ and $\sigma^2 = 0$ (see also Theorem 1). In addition, if the distance between the interpolated subcarrier index ω and its closest base point is sufficiently small, i.e., $d_{\max} < W/L$, then we obtain a sharper upper bound on the MSE than $\varepsilon_{\text{CSI}} + \frac{2}{B}(1 + \varepsilon_{\text{cor}})$. In a scenario with a large delay spread L , Corollary 2 reveals that one requires finer-spaced base points for 0th order interpolation in order to keep the approximation error strictly smaller than $\varepsilon_{\text{CSI}} + \frac{2}{B}(1 + \varepsilon_{\text{cor}})$. Since the maximum error is mainly determined by $d_{\max}$, a good strategy for selecting base points with 0th order approximation is uniformly spacing them in the set of active subcarriers Ω .

C. MSE of 1st Order Interpolation

We now present the approximation error analysis of 1st order interpolation. The following result characterizes the MSE of 1st order interpolation; the proof is given in Appendix E.

Theorem 2. *Let the entries of the TD matrices $\hat{\mathbf{H}}_\ell^{\text{uncor}}$, $\ell = 1, \dots, L$, be distributed $\mathcal{CN}(0, 1/(BL))$ per complex entry. Assume that the off-diagonal of the receive correlation matrix are δ , the variance of the channel estimation error σ across all subcarriers $\omega = 0, \dots$, and $p \in \mathcal{P}$ is the closest base point to the target subcarrier ω . Then, for any (m, n) -th entry of the Gram matrix $\mathbf{G}_\omega$, the MSE for the 1st order interpolation method in (6) is given by*

$$\begin{aligned} 1\text{-MSE}_\omega &= \varepsilon_{\text{CSI}}(1 - \lambda_\omega(1 - \lambda_\omega)) \\ &+ \frac{2}{B}(1 + \varepsilon_{\text{cor}})(1 - \lambda_\omega(1 - \lambda_\omega) + \lambda_\omega(1 - \lambda_\omega)f_L(\theta) \\ &- (1 - \lambda_\omega)f_L(\lambda_\omega\theta) - \lambda_\omega f_L((1 - \lambda_\omega)\theta)), \end{aligned} \quad (12)$$

where $\theta = \frac{2\pi}{W}(p_{k+1} - p_k)$ and $\lambda_\omega = (p_{k+1} - \omega)/(p_{k+1} - p_k)$.

Similarly to 0th order interpolation, we observe that the MSE of 1st order interpolation is independent of the entry (m, n) and impacted by CSI errors and receive correlation (see Section IV-B for detailed discussion). The result shown next in Corollary 3 reveals that if the spacing between the two base-points p_k and p_{k+1} defined as $d_k = p_{k+1} - p_k$ is sufficiently small, then the 1st order interpolation strictly

outperforms 0th order interpolation, i.e., $1\text{-MSE}_\omega < 0\text{-MSE}_\omega$ for all $\omega \in (p_k, p_{k+1})$; the proof is given in Appendix F.

Corollary 3. *Let d_k denote the spacing between two base-points p_k and p_{k+1} , and assume the conditions in Theorem 1 and Theorem 2 hold. If $d_k \leq W/(3L)$, then*

$$1\text{-MSE}_\omega \leq 0\text{-MSE}_\omega, \quad \text{for all } \omega \in (p_k, p_{k+1}), \quad (13)$$

which holds with equality if and only if $L = 1$ and $\varepsilon_{\text{CSI}} = 0$.

We note that the condition $d_k < W/(3L)$ is not sharp; Appendix F outlines the details on how it can be sharpened. Furthermore, given that d_k is significantly larger than $W/(3L)$, we can construct situations for which 0th order interpolation *outperforms* 1st order interpolation. Note that for $L = 1$, the FD channel is flat (i.e., $\mathbf{G}_\omega$ is constant for all ω) and hence, 1st and 0th order interpolation have the same MSE.

In summary, we observe that for both approximate interpolation methods, the MSE can be lowered by increasing the number of BS antennas B , assuming that the channel estimation error ε_{CSI} decreases with B . In the large-antenna limit $B \rightarrow \infty$ with perfect CSI and no BS-antenna correlation, the MSE vanishes, which is an immediate consequence of channel hardening in massive MU-MIMO systems [1]. Furthermore, 1st order interpolation generally outperforms 0th order interpolation for a sufficiently small minimum spacing between adjacent base points, i.e., for $d_k \leq W/(3L)$.

V. COMPLEXITY ANALYSIS

We next compare the computational complexity of the four studied Gram-matrix computation algorithms: brute-force computation, exact interpolation, 0th order interpolation, and 1st order interpolation. We measure the computational complexity by counting the number of real-valued multiplications.⁵

A. Brute-Force Computation

We start by deriving the total computational complexity required by the brute-force (BF) method. We only compute the upper triangular part of $\mathbf{G}_\omega$ (since the matrix is Hermitian). Each off-diagonal entry requires B complex-valued multiplications, which corresponds to $4B$ real-valued multiplications; each diagonal entry requires only $2B$ real-valued multiplications. Hence, the computational complexity of computing $\mathbf{G}_\omega$ using the BF method is

$$C_{\text{BF}} = |\Omega| \left(4B \frac{U(U-1)}{2} + 2BU \right) = 2|\Omega|BU^2 \quad (14)$$

for a total number of $|\Omega|$ active subcarriers.

B. Exact Interpolation

We now derive the computational complexity of exact interpolation as discussed in Section III-A. Exact interpolation requires a BF computation of the Gram matrix at each of the

$|\mathcal{P}|$ base points. We will use the $|\mathcal{P}|$ precomputed base points of $\mathbf{G}_\omega$ to interpolate the remaining $|\Omega| - |\mathcal{P}|$ Gram matrices.

We will assume that the base points and the assumed channel delay spread L are fixed *a-priori* so that $\mathbf{F}_{\Omega \setminus \mathcal{P}, L} \mathbf{F}_{\mathcal{P}, L}^\dagger$ in (5) can be precomputed and stored. We emphasize that this approach does not include the computational complexity of computing the interpolation matrix itself, which favors this interpolation scheme from a complexity perspective. In fact, we only need to multiply the precomputed interpolation matrix $\mathbf{F}_{\Omega \setminus \mathcal{P}, L} \mathbf{F}_{\mathcal{P}, L}^\dagger$ with the vector $\mathbf{g}_\mathcal{P}$, which requires $4(|\Omega| - |\mathcal{P}|)|\mathcal{P}|$ real-valued multiplications. Hence, the total computational complexity of exact interpolation is:

$$\begin{aligned} C_{\text{Exact}} &= \frac{|\mathcal{P}|}{|\Omega|} C_{\text{BF}} + 4(|\Omega| - |\mathcal{P}|)|\mathcal{P}| \frac{U(U+1)}{2} \\ &= 2|\mathcal{P}|(|\Omega| - |\mathcal{P}| + B)U^2 + 2|\mathcal{P}|(|\Omega| - |\mathcal{P}|)U. \end{aligned} \quad (15)$$

We note that if the number of users U is large and the number of base points is similar to the number of BS antennas, i.e., $|\mathcal{P}| \simeq B$, then the BF method in (14) and exact interpolation (15) exhibit similar complexity. We also observe that the complexity of exact interpolation (15) is lower than that of the BF method (14) if $|\mathcal{P}| < (1+U)^{-1}BU$. Since the use of $|\mathcal{P}| \geq 2L - 1$ distinct base points guarantees exact interpolation (assuming perfect CSI), we observe that exact interpolation has lower complexity than the BF method if L is (approximately) smaller than $B/2$.

C. 0th Order Interpolation

The computational complexity of the 0th order interpolation method is given by

$$C_{\text{0th}} = 2|\mathcal{P}|BU^2, \quad (16)$$

as we only need to compute the Gram matrices on all the base points. We note that since typically $|\mathcal{P}| \ll |\Omega|$ the savings (in terms of real-valued multiplications) are significant compared to the BF approach and exact interpolation, but does so at the cost of approximation errors (cf. Section VI-D).

D. 1st Order Interpolation

The computational complexity of the 1st order interpolation is given by

$$\begin{aligned} C_{\text{1st}} &= C_{\text{0th}} + 4(|\Omega| - |\mathcal{P}|) \frac{U(U+1)}{2} \\ &= 2|\mathcal{P}|BU^2 + 2(|\Omega| - |\mathcal{P}|)U(U+1), \end{aligned} \quad (17)$$

where we assume that the interpolation weight λ_ω was pre-computed. We note that the linear interpolation stage for each subcarrier $\omega \in \Omega \setminus \mathcal{P}$ requires four real-valued multiplications. By comparing (16) to (17), we observe that the complexity of 1st order interpolation always exceeds the complexity of the 0th order method, but the complexity is significantly lower than that of the BF method as we generally have $|\mathcal{P}| \ll |\Omega|$.

VI. NUMERICAL RESULTS

We now study the MSE, the error-rate performance, and the computational complexity of the proposed Gram-matrix

⁵We assume that a complex-valued multiplication requires four real-valued multiplications; computation of the squared magnitude of a complex number is assumed to require two real-valued multiplications.

interpolation schemes. We consider a MU-MIMO-OFDM system with 128 BS antennas and with 8 single-antenna users. We assume a total of $W = 2048$ subcarriers, with $|\Omega| = 1200$ active subcarriers, similar to that used in 3GPP LTE [14]. Unless stated otherwise, we assume that the entries of the TD channel matrices are i.i.d. circularly-symmetric complex Gaussian with variance $1/(BL)$ and we consider 16-QAM transmission (with Gray mapping). We use a linear MMSE equalizer for data detection; see Section II-C. For situations with imperfect CSI, we consider pilot-based maximum-likelihood (ML) channel estimation with a single orthogonal pilot sequence of length U with the same transmit power as the data symbols.

A. Complexity Comparison

We now assess the complexity of the various Gram-matrix computation methods in comparison to the overall complexity required for linear MMSE-based data detection, which includes Gram-matrix and matched-filter computation as well as matrix inversion for each active subcarrier. The results shown here are for a 128×8 (the notation represents $B \times U$) massive MU-OFDM-MIMO system with 1200 active subcarriers and a delay spread of $L = 144$.

Fig. 2 compares the complexity of Gram matrix computation for four different methods, brute-force, exact, 0th-, and 1st-order interpolation methods for $|\mathcal{P}| = \{0.25|\Omega|, 0.5|\Omega|, 0.75|\Omega|, |\Omega|\}$. The solid part of the bar plot shows the complexity of Gram matrix computation; the fenced part corresponds to the remaining complexity required for data detection (including matched-filter computation and a matrix inversion for each active subcarrier). The percentage values indicate the relative complexity of Gram-matrix computation compared to the total complexity required for data detection. We assume an implicit Cholesky-based matrix inversion for data detection [35]. As demonstrated in [26], Gram matrix computation requires most of the complexity, as it scales quadratically in the number of BS antennas.

We see that the exact interpolation method results in high complexity in the considered system (see Section V-B for exact details when exact interpolation achieves lower complexity than a BF approach). We also see that the proposed 0th and 1st order approximation methods both achieve significant complexity reductions. For $|\mathcal{P}| = 0.25|\Omega|$, the proposed methods requires less than half the complexity of a BF approach. As we will show in Section VI-C, the proposed approximate interpolation methods will exhibit similar error-rate performance as that the BF approach (see Figs. 4 and 5), but do so at fraction of the computational complexity.

B. MSE of Approximate Interpolation

Fig. 3 compares the MSE of 0th and 1st order interpolation as proposed in Section III-B. Note that the BF method and exact interpolation have an MSE of zero and hence, we exclude these results. We select two base points at subcarriers 500 and 600 and one target point at subcarrier 512, and compare the MSE for different numbers of BS antennas and under ideal and non-ideal scenarios. In the ideal scenario, we assume perfect CSI and no BS-antenna correlation, whereas in the non-ideal

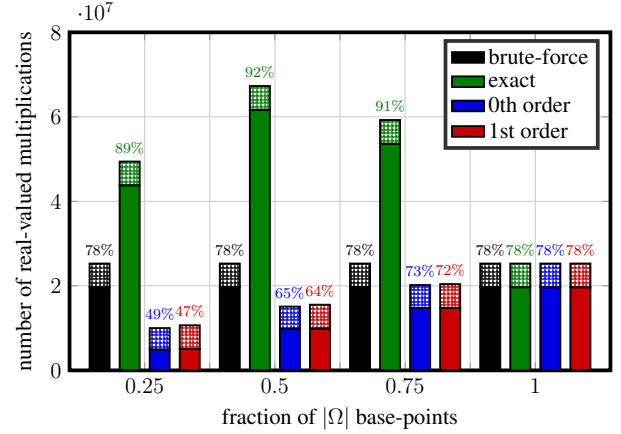


Fig. 2. We compare the complexity of various Gram matrix computation methods in comparison to the complexity of data detection for a 128×8 MU-OFDM-MIMO system with $|\Omega| = 1200$, $L = 144$, and four different sets of base-points $|\mathcal{P}| = \{0.25|\Omega|, 0.5|\Omega|, 0.75|\Omega|, |\Omega|\}$. The percentage values shown in the above bar plots show the relative percentage of Gram-matrix computation compared to the total complexity required for data detection.

scenario we assume channel-estimation at $\text{SNR} = 25$ dB across all subcarriers with the signal-to-noise ratio (SNR) defined by $\text{SNR} = U/(B\sigma^2)$ and a BS-antenna correlation of $\delta = 0.1$. In order to assess the approximation error with respect to different channel delay spreads, we set $L \in \{36, 72, 144\}$. The resulting MSE is shown in Fig. 3. Note that the MSE for both 0th and 1st are independent of the entry (as predicted by Theorems 1 and 2); hence, we consider the average MSE across all entries.

We observe that the 1st order interpolation method achieves a lower MSE than that given by 0th order interpolation, where the performance gap increases with larger delay spreads L ; this caused by the fact that for small delay spreads L , the channel is more smooth across subcarriers. For larger delay spreads L , 1st order interpolation captures the faster-changing behavior of the Gram matrix, whereas the 0th order interpolation ignores such changes. We also see that the MSE degrades in the non-ideal scenario, even if we increase the number of BS antennas; this behavior is reflected in our analytical results. Finally, we see that the simulated MSE matches perfectly our theoretical results in Theorems 1 and 2.

C. Error-rate Performance

We now compare the error-rate performance of the proposed Gram-matrix computation schemes. We simulate the bit-error rate (BER) for a MU-MIMO-OFDM system for a different number of base-points $|\mathcal{P}|$ and for perfect as well as imperfect CSI. We also investigate the impact of a more realistic channel model. For all results, we simulate three different numbers of base-points $|\mathcal{P}| = L/4$, $|\mathcal{P}| = 2L - 1$, and $|\mathcal{P}| = 4L$, and select equally-spaced base points. Figures 4(a) and 4(b) show BER simulation results for an i.i.d. Rayleigh fading scenario with perfect and imperfect CSI, respectively. Figure 4(c) shows BER simulation results for the QuaDRiGa channel model⁶ with

⁶We simulate a square antenna array with a non-line-of-sight scenario with a 2 GHz carrier frequency, 20 MHz bandwidth, and 200 m distance between BS antenna and the users. Our algorithms assume $L = 144$ but the true delay spread is slightly smaller.

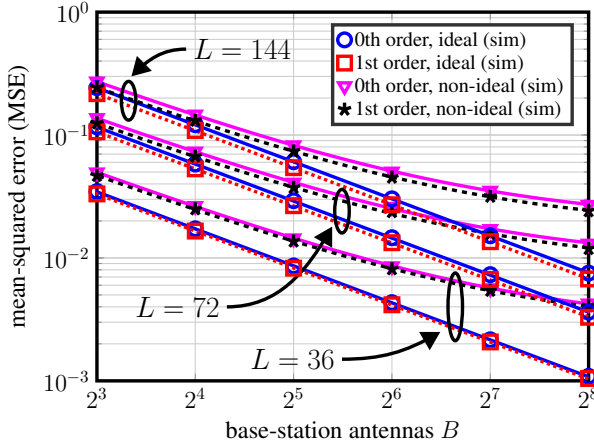


Fig. 3. MSE of 0th and 1st order interpolation for an entry of $\mathbf{G}_{512}$ using two base points $\mathbf{G}_{500}$ and $\mathbf{G}_{600}$ for three different delay spreads $L \in \{36, 72, 144\}$ for the ideal and non-ideal scenario. The markers represent simulation results whereas the lines represent our approximation-error analysis. Evidently, our theory matches perfectly with the simulated values.

imperfect CSI [36]. We note that QuaDRiGa channel model includes a path-loss model for each user.

Figure 4(a) shows that exact interpolation for $|\mathcal{P}| \geq 2L - 1$ base points provides identical results as the BF method (up to machine precision) for a system with perfect CSI. For $|\mathcal{P}| = L/4$ base points, the proposed 0th and 1st order interpolation exhibit an error floor; this performance loss can be mitigated substantially by increasing the number of base points to $|\mathcal{P}| = 2L - 1$. By setting $|\mathcal{P}| = 4L < 0.5|\Omega|$, the both the 0th and 1st order interpolation methods exhibit virtually no BER performance loss.

Figure 4(b) shows the situation for imperfect CSI (with channel estimation). We observe that the performance of the BF method and that of exact interpolation are no longer equal. In fact, for $|\mathcal{P}| = 2L - 1$ and $|\mathcal{P}| = 4L$ base points, exact interpolation exhibits a significant error floor. The reason is due to the fact that the interpolation matrix is ill-conditioned, which results in significant noise enhancement. Although the error floor is decreased for $|\mathcal{P}| = 4L$, a floor remains at 10^{-3} BER. In contrast, the error floor of 0th order and 1st order interpolation for $|\mathcal{P}| = 2L - 1$ and $|\mathcal{P}| = 4L$ base points is well-below 10^{-4} BER and hence, the proposed approximate interpolation schemes are more resilient to scenarios with imperfect CSI than exact interpolation.

Figure 4(c) shows the BER performance for the QuaDRiGa channel model with imperfect CSI. We observe that all considered interpolation methods achieve a lower error floor than that given in Fig. 4(b) for $|\mathcal{P}| = L/4$; this is due to the fact that the effective delay spread for the considered channel is smaller than $L = 144$ (which is assumed in the algorithms). Once again, we observe a BER floor of exact interpolation for all considered numbers of base points. In summary, the proposed approximate interpolation methods are more robust in practical scenarios than the exact interpolation method.

D. Performance/Complexity Trade-off

We now investigate the BER performance vs. computational complexity trade-off for the proposed approximate interpolation methods with imperfect CSI. We use the complexity C_{BF} of the BF method in (14) as our baseline, and we compare it to that of the proposed 0th and 1st order interpolation methods in (16) and (17), respectively. We vary the number of base points $|\mathcal{P}|$ from L to D and simulate the minimum SNR required for the linear MMSE equalizer to achieve 10^{-3} BER.

Figure 5 shows the trade-off results for 0th and 1st order interpolation. For a fixed fraction of the complexity of C_{BF} , we observe that the 1st order interpolation method always outperforms the 0th order interpolation method. Hence, Fig. 5 clearly reveals that the additional complexity required by linear interpolation is beneficial when jointly considering performance and complexity. In addition, we see that the 1st order interpolation method approaches the SNR performance of the BF method by 1 dB with only 45% of the complexity.

VII. CONCLUSIONS

We have studied the performance of exact and approximate interpolation-based Gram matrix computation for wideband massive MU-MIMO-OFDM systems. Instead of performing a brute-force (BF) computation of the Gram matrix for all subcarriers or using exact interpolation schemes, we have proposed two simple, yet efficient approximate interpolation methods. We have demonstrated that channel hardening in massive MU-MIMO enables the proposed 0th and 1st order interpolation schemes to perform close to that of an exact BF computation at only a fraction of the computational complexity. In addition, the proposed approximate interpolation methods are more robust to channel-estimation errors and receive-side antenna correlation than exact Gram-matrix interpolation.

There are many opportunities for future work. We expect the use of higher-order approximate Gram-matrix interpolation schemes to perform better at an increase in complexity. An analysis of such methods is left for future work. Our results also indicate that the broad range of existing exact interpolation schemes for small-scale, point-to-point MIMO systems, e.g., matrix inversion and QR decomposition [15], [16] can be made more robust and less complex in massive MU-MIMO systems if combined with approximate, low-order interpolation schemes. In fact, the recent result in [31] for approximate interpolation of matrix inversion demonstrates this claim in a massive MU-MIMO scenario via simulations—it is an open problem to analyze the performance of such interpolation methods. In addition, it would be practically relevant to develop theory for more realistic correlation models, such as the one in [37].

ACKNOWLEDGMENTS

The authors would like to thank M. Wu and C. Dick from Xilinx, Inc. for insightful discussions on interpolation-based matrix computations. The work of C. Jeon and C. Studer was supported in part by Xilinx Inc., and by the US National Science Foundation (NSF) under grants ECCS-1408006, CCF-1535897, CAREER CCF-1652065, CNS-1717559, and EECS-1824379.

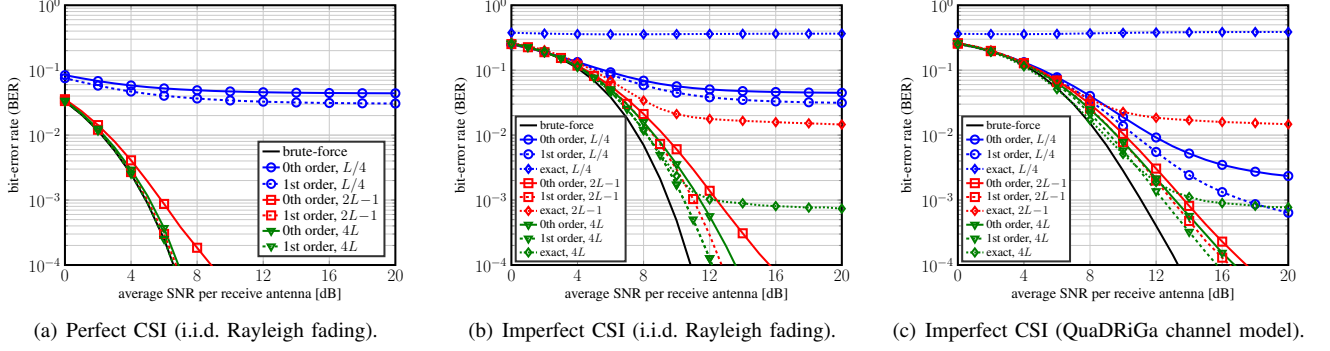


Fig. 4. Uncoded bit error-rate (BER) comparison in a 128 BS antenna, 8 (single-antenna) user, wideband massive MU-MIMO-OFDM system. The values next to the legend entries correspond to the number of base points $|\mathcal{P}|$. The proposed approximate interpolation schemes (0th and 1st order interpolation) outperform exact interpolation for scenarios with imperfect CSI and approach the performance of the exact brute-force method for a small number of base points.

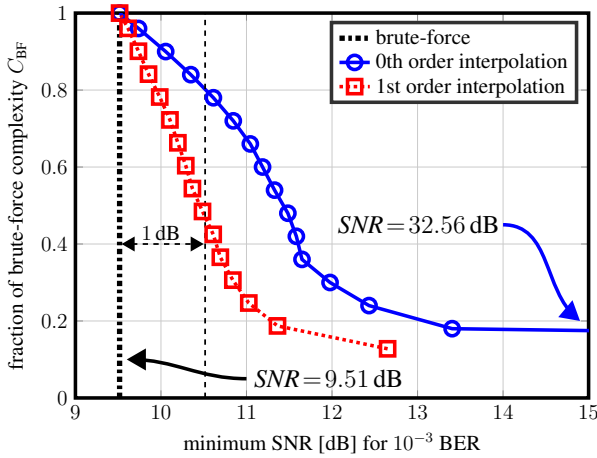


Fig. 5. Trade-off between SNR performance and computational complexity for 0th and 1st order interpolation in an i.i.d. Rayleigh fading channel with imperfect CSI. Both approximate interpolation methods approach the performance of an exact brute-force approach at a fraction of the complexity.

APPENDIX A PROOF OF LEMMA 1

Since all of the possible values in the exponent of (4), i.e., $\ell - \ell'$, are integers ranging from $-(L-1)$ to $L-1$, the Gram matrix $\mathbf{G}_\omega$ in (4) is a polynomial with degree no larger than $2L-1$. Consequently, the Gram matrices $\mathbf{G}_\omega$ in (4) for all subcarriers $\omega = 0, \dots, W$ are fully determined from $2L-1$ distinct and non-zero Gram-matrix base-points.

APPENDIX B PROOF OF LEMMA 2

Evidently, (9) is non-negative, i.e., $f_L(\phi) \geq 0$. To show that $f_L(\phi) \leq 1$, we use the fact that the Fejér kernel $L f_L(\phi) = L^{-1} \frac{1 - \cos(L\phi)}{1 - \cos(\phi)}$ is upper bounded by L [34, Eq. 1.2.24]. Now, we show that $f_L(\phi)$ is monotonically decreasing in $\phi \in [0, 2\pi/L]$ for $L > 1$. We start by defining an auxiliary function $g(\phi) = \sin(L\phi/2)/\sin(\phi/2)$, so that $f_L(\phi) = g_L^2(\phi)/L^2$. For $L > 1$, $\phi \in (\pi/L, 2\pi/L]$, $\sin(L\phi/2)$ and $\sin(\phi/2)$ are monotonically decreasing and increasing respectively, and hence, $g(\phi)$ and

$f_L(\phi)$, are monotonically decreasing. For $\phi \in [0, \pi/L]$, the derivative of $g(\phi)$ with respect to ϕ is given by:

$$\frac{dg(\phi)}{d\phi} = \frac{(L \cos(L\phi/2) \sin(\phi/2) - \cos(\phi/2) \sin(L\phi/2))}{2 \sin^2(\phi/2)},$$

and we have that

$$\frac{L \cos(L\phi/2) \sin(\phi/2)}{\cos(\phi/2) \sin(L\phi/2)} = \frac{L \tan(\phi/2)}{\tan(L\phi/2)} < 1.$$

Hence, $g'(\phi) < 0$, and therefore, $f_L(\phi)$ is monotonically decreasing in $\phi \in [0, 2\pi/L]$.

APPENDIX C PROOF OF THEOREM 1

Suppose we use $\mathbf{G}_{p_k}$ at base point p_k to approximate $\mathbf{G}_\omega$ at the target subcarrier index ω . Hence, the MSE in (8) is given by the following expression:

$$0\text{-MSE}_\omega^{(m,n)} = \mathbb{E} \left[|[\tilde{\mathbf{G}}_\omega]_{m,n} - [\mathbf{G}_\omega]_{m,n}|^2 \right]. \quad (18)$$

We will obtain an analytical expression for (18) with imperfect CSI and the BS-antenna correlation model introduced in (3) and (2), respectively. We start with expressing the channel matrix $\mathbf{H}_\omega$ from (3) as

$$\mathbf{H}_\omega \stackrel{(a)}{=} \sum_{\ell=0}^{W-1} \left(\hat{\mathbf{H}}_\ell \mathbb{I}(\ell < L) + \frac{\sigma}{\sqrt{W}} \tilde{\mathbf{E}}_\ell \right) \exp \left(-\frac{j2\pi\omega\ell}{W} \right), \quad (19)$$

where (a) follows from noting that the DFT is orthogonal (with normalization constant) with each entries of $\tilde{\mathbf{E}}_\ell$ distributed $\mathcal{CN}(0, 1)$. Hence, the TD channel matrix under imperfect CSI is given as $\hat{\mathbf{H}}_\omega^\sigma = \hat{\mathbf{H}}_\omega \mathbb{I}(\ell < L) + \frac{\sigma}{\sqrt{W}} \tilde{\mathbf{E}}_\ell$.

Now, use the correlation model introduced in (2). We note that the $B \times B$ BS correlation matrix $\mathbf{R} = (1 - \delta)\mathbf{I}_B + \delta\mathbf{1}_B$ can be expressed by $(\alpha\mathbf{I}_B + \beta\mathbf{1}_B)^2 = \mathbf{R}$ where $\alpha = \sqrt{1 - \delta}$, and $\beta = \frac{1}{B}(\sqrt{1 + (B-1)\delta} - \alpha)$. Since the matrix $(\alpha\mathbf{I}_B + \beta\mathbf{1}_B)$ is symmetric, we note that the $\hat{\mathbf{H}}_\ell$ is expressed as $\hat{\mathbf{H}}_\ell = (\alpha\mathbf{I}_B + \beta\mathbf{1}_B)\hat{\mathbf{H}}_\ell^{\text{uncor}}$ so that

$$[\hat{\mathbf{H}}_\ell]_{m,n} = \alpha[\hat{\mathbf{H}}_\ell^{\text{uncor}}]_{m,n} + \beta \sum_{b=1}^B [\hat{\mathbf{H}}_\ell^{\text{uncor}}]_{b,n} \quad (20)$$

For our derivation of the MSE, we will utilize the following auxiliary function:

$$A_{\ell\ell'} = \exp\left(j\frac{2\pi p_k}{W}(\ell - \ell')\right) - \exp\left(j\frac{2\pi\omega}{W}(\ell - \ell')\right). \quad (21)$$

By substituting $A_{\ell\ell'}$ into $[\tilde{\mathbf{G}}_\omega]_{m,n} - [\mathbf{G}_\omega]_{m,n}$ in (18), we obtain the following expression for the 0th order MSE:

$$\begin{aligned} 0\text{-MSE}_\omega^{(m,n)} &= \mathbb{E}\left[\left|\sum_{\ell,\ell'=0}^{W-1}\sum_{b=1}^B[\hat{\mathbf{H}}_\ell^\sigma]^*_{b,m}[\hat{\mathbf{H}}_{\ell'}^\sigma]_{b,n}A_{\ell\ell'}\right|^2\right] \\ &= \sum_{\ell_1,\ell_2=0}^{W-1}\sum_{\ell_3,\ell_4=0}^{W-1}\sum_{b_1,b_2=1}^B\left(A_{\ell_1\ell_2}^*A_{\ell_3\ell_4}\right. \\ &\quad \left.\mathbb{E}[[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_1,m}[\hat{\mathbf{H}}_{\ell_2}^\sigma]_{b_1,n}^*[\hat{\mathbf{H}}_{\ell_3}^\sigma]_{b_2,m}[\hat{\mathbf{H}}_{\ell_4}^\sigma]_{b_2,n}]\right) \\ &\stackrel{(a)}{=} \sum_{\substack{\ell_1,\ell_2=0 \\ \ell_1 \neq \ell_2}}^{W-1} |A_{\ell_1\ell_2}|^2 \\ &\quad \sum_{b_1,b_2=1}^B \mathbb{E}[[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_2,m}^*[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_1,m}]\mathbb{E}[[\hat{\mathbf{H}}_{\ell_2}^\sigma]_{b_1,n}^*[\hat{\mathbf{H}}_{\ell_2}^\sigma]_{b_2,n}] \quad (22) \end{aligned}$$

where (a) follows from $A_{\ell_1\ell_2} = 0$ if $\ell_1 = \ell_2$, and independence and the zero-mean assumption on the TD channel for $\ell_1 \neq \ell_2$, which enforces $\ell_1 = \ell_3$ and $\ell_2 = \ell_4$ for (22). By inspection of (22), we observe that $0\text{-MSE}_\omega^{(m,n)}$ is independent of m and n and thus, the MSE of the off-diagonal and diagonal entries are equal. We simplify (22) for imperfect CSI and the BS-antenna correlation model. We first note that

$$\begin{aligned} \mathbb{E}[[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_2,m}^*[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_1,m}] &= \frac{\sigma^2}{W}\mathbb{I}(b_1 = b_2) \\ &+ \frac{1}{BL}(\alpha^2\mathbb{I}(b_1 = b_2) + 2\alpha\beta + \beta^2B)\mathbb{I}(\ell_1 < L) \\ &= \frac{\sigma^2}{W}\mathbb{I}(b_1 = b_2) + \frac{1}{BL}(\alpha^2\mathbb{I}(b_1 = b_2) + \delta)\mathbb{I}(\ell_1 < L), \end{aligned}$$

where the last step is obtained by noting that $2\alpha\beta + \beta^2B = \delta$. Therefore, the inner sum is evaluated by

$$\begin{aligned} &\sum_{b_1,b_2=1}^B \mathbb{E}[[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_2,m}^*[\hat{\mathbf{H}}_{\ell_1}^\sigma]_{b_1,m}]\mathbb{E}[[\hat{\mathbf{H}}_{\ell_2}^\sigma]_{b_1,n}^*[\hat{\mathbf{H}}_{\ell_2}^\sigma]_{b_2,n}] \\ &= \frac{B\sigma^4}{W^2} + \frac{\sigma^2}{LW}(\alpha^2 + \delta)(\mathbb{I}(\ell_1 < L) + \mathbb{I}(\ell_2 < L)) \\ &\quad + \frac{1}{BL^2}(\alpha^2(\alpha^2 + 2\delta) + B\delta^2)\mathbb{I}(\ell_1 < L)\mathbb{I}(\ell_2 < L). \quad (23) \end{aligned}$$

Now, we simplify (22) using the results from (23) with the fact that $\alpha = \sqrt{1 - \delta}$ and $\sum_{\ell_1=0}^{W-1}|A_{\ell_1\ell_2}|^2 = 2W$ so that $\sum_{\ell_1,\ell_2=0}^{W-1}|A_{\ell_1\ell_2}|^2 = 2W^2$. Hence,

$$\begin{aligned} 0\text{-MSE}_\omega^{(m,n)} &= 2B\sigma^4 + 4\sigma^2 + \frac{1 + \delta^2(B-1)}{BL^2} \sum_{\substack{\ell_1,\ell_2=0 \\ \ell_1 \neq \ell_2}}^{L-1} |A_{\ell_1\ell_2}|^2 \\ &= 2B\sigma^4 + 4\sigma^2 + \frac{2(1 + \delta^2(B-1))}{B} \left(1 - \sum_{\ell=0}^{L-1} \sum_{\Delta\ell=-\ell}^{\ell} \frac{e^{j\theta\Delta\ell}}{L^2}\right) \\ &\stackrel{(a)}{=} 2\sigma^2(2 + B\sigma^2) + \frac{2}{B}(1 + \delta^2(B-1))(1 - f_L(\theta)), \end{aligned}$$

where (a) comes from the definition of Fejér kernel [34]. Note that we defined the shorthand variable $\theta = \frac{2\pi}{W}(p_k - \omega)$.

The proof can be generalized to per-UE large-scale fading by expressing the channel matrix as $\bar{\mathbf{H}}_\omega = \mathbf{H}_\omega \mathbf{D}$, where $\mathbf{H}_\omega$ was defined in (19) and the diagonal matrix $\mathbf{D}$ contains the large-scale fading coefficients for the UEs on the main diagonal. In addition, the proof can be generalized to receive-side correlation matrices $\mathbf{R}$ by rewriting $\hat{\mathbf{H}}_\ell$ in (20) with $\hat{\mathbf{H}}_\ell = \sqrt{\mathbf{R}}\hat{\mathbf{H}}_\ell^{\text{uncor}}$. A corresponding analysis is left for future work.

For the case of other correlation models, e.g., the one in [37], one can first compute the Cholesky decomposition of $\mathbf{R}$ to compute $\hat{\mathbf{H}}_\ell$ and proceed with the analysis as in (20). We note that the Cholesky decomposition of the exponential correlation model proposed in [37] can be computed in closed form. We leave the details for future work.

APPENDIX D PROOF OF COROLLARY 2

Since $f_L(\phi)$ is non-negative, it is obvious that $\max_{\omega \in \Omega} \{0\text{-MSE}_\omega\} \leq \varepsilon_{\text{CSI}} + 2(1 + \varepsilon_{\text{cor}})/B$ for all $\omega \in \Omega$. The equality is satisfied if $(p - \omega)/W = b/L$ for some integer $b > 0$ so that $f_L(\frac{2\pi}{W}(p - \omega)) = f_L(\frac{b2\pi}{L}) = 0$. We note that this can only happen if $d_{\text{max}} \geq W/L$, where d_{max} is the maximum distance between any target subcarrier ω point and its nearest base point; this is due to the fact that $\frac{2\pi}{W}(p - \omega) \leq \frac{2\pi}{W}d_{\text{max}} < \frac{2\pi}{L}$.

Assume $d_{\text{max}} < W/L$. Then, by Lemma 2, the maximum MSE of 0th order interpolation is given by:

$$\begin{aligned} &\max_{\omega \in \Omega} \{0\text{-MSE}_\omega\} \\ &= \varepsilon_{\text{CSI}} + \frac{2}{B}(1 + \varepsilon_{\text{cor}}) \left(1 - \min_{\omega \in \Omega \setminus \mathcal{P}} f_L\left(\frac{2\pi}{W}(p_k - \omega)\right)\right) \\ &= \varepsilon_{\text{CSI}} + \frac{2}{B}(1 + \varepsilon_{\text{cor}}) \left(1 - f_L\left(\frac{2\pi}{W}d_{\text{max}}\right)\right). \end{aligned}$$

APPENDIX E PROOF OF THEOREM 2

The proof is similar to that of Theorem 1 in Appendix C. We start by defining the following auxiliary function

$$\begin{aligned} Q_{\ell\ell'} &= \lambda_\omega \exp\left(j\frac{2\pi p_k}{W}\Delta\ell\right) + (1 - \lambda_\omega) \exp\left(j\frac{2\pi p_{k+1}}{W}\Delta\ell\right) \\ &\quad - \exp\left(j\frac{2\pi\omega}{W}\Delta\ell\right), \end{aligned}$$

where we introduced the variable $\Delta\ell = \ell - \ell'$. The result is obtained by substituting $Q_{\ell\ell'}$ in place of $R_{\ell\ell'}$ at (22) in Appendix C. Note that $\sum_{\ell_1=0}^{W-1}|Q_{\ell_1\ell_2}|^2 = 2(1 - \lambda_\omega(1 - \lambda_\omega))W$ which shows that $\sum_{\ell_1,\ell_2=0}^{W-1}|Q_{\ell_1\ell_2}|^2 = 2(1 - \lambda_\omega(1 - \lambda_\omega))W^2$.

APPENDIX F PROOF OF COROLLARY 3

Without loss of generality, we will assume that the target subcarrier index ω is closer to p_{k+1} so that $\lambda_\omega \in [0, 0.5]$. We will assume that $p_k < \omega < p_{k+1}$ so that $d_k = p_{k+1} - p_k >$

0. Using the results from Appendix C and Appendix E, the difference of 1-MSE_ω and 0-MSE_ω is given by:

$$1\text{-MSE}_\omega - 0\text{-MSE}_\omega = -\varepsilon_{\text{CSI}}\lambda_\omega(1 - \lambda_\omega) + \frac{1 + \varepsilon_{\text{cor}}}{BL^2} \sum_{\ell_1, \ell_2=0}^{L-1} (|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2).$$

Without loss of generality, we assume that $\ell_1 - \ell_2 > 0$ since $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2$ is even and is 0 if $\ell_1 = \ell_2$. We simplify the term $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2$ by denoting $\phi = \theta(\ell_1 - \ell_2) > \theta$, where $\theta = \frac{2\pi}{W}d_k$ and expand the expression $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2$ as follows:

$$\begin{aligned} |Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2 &\stackrel{(a)}{=} 2\lambda_\omega((1 - \lambda_\omega)(\cos(\phi) - 1) \\ &\quad - (\cos((1 - \lambda_\omega)\phi) - \cos(\lambda_\omega\phi))) \\ &\stackrel{(b)}{=} 2\lambda_\omega \sin(\phi/2)(\sin((1 - 2\lambda_\omega)\phi/2) - (1 - \lambda_\omega)\sin(\phi/2)). \end{aligned} \quad (24)$$

Here, (a) follows from the definition of $Q_{\ell_1\ell_2}$ and $R_{\ell_1\ell_2}$ and (b) is a results from simplifying the expression $\cos((1 - \lambda_\omega)\phi) - \cos(\lambda_\omega\phi) = -2\sin((1 - 2\lambda_\omega)\phi/2)\sin(\phi/2)$.

We first note that since $\varepsilon_{\text{CSI}} = 2\sigma^2(2 + B\sigma^2) \geq 0$ and by (24), $1\text{-MSE}_\omega = 0\text{-MSE}_\omega$ if $\varepsilon_{\text{CSI}} = 0$, and $L = 1$ or $\lambda_\omega = 0$. This behavior can be explained intuitively because when $\varepsilon_{\text{CSI}} = 0$ and $L = 1$, then the channel is flat across all subcarriers, and hence $1\text{-MSE}_\omega = 0\text{-MSE}_\omega$.

Hence, we now show that $1\text{-MSE}_\omega < 0\text{-MSE}_\omega$ for $L > 1$ and $\lambda_\omega \neq 0$ by showing that $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2 < 0$. First note that $\ell_1 - \ell_2 < L$ and if $d_k < W/(3L)$, then $\theta < \phi < L\theta = \frac{2\pi}{W}Ld_k < \frac{2\pi}{3}$. Hence, $\sin(\phi/2) > 0$. Therefore, showing that (24) is negative is equivalent to:

$$g(\lambda_\omega) = \frac{\sin((1 - 2\lambda_\omega)\phi/2)}{1 - \lambda_\omega} < \sin(\phi/2). \quad (25)$$

We now prove (25) by noting that $g(\lambda_\omega) = \sin(\phi/2)$ if $\lambda_\omega = 0$ and $g'(\lambda_\omega) < 0$ for all $\lambda_\omega \in (0, 0.5]$ so $g(\lambda_\omega)$ is monotonically decreasing in $(0, 0.5]$. The proof is straightforward by:

$$g'(\lambda_\omega) = \frac{-\phi \cos((1 - 2\lambda_\omega)\phi/2)}{1 - \lambda_\omega} + \frac{\sin((1 - 2\lambda_\omega)\phi/2)}{(1 - \lambda_\omega)^2}, \quad (26)$$

and, hence, $g'(\lambda_\omega) < 0$ in (26) can be expressed as:

$$\tan((1 - 2\lambda_\omega)\phi/2) < (1 - \lambda_\omega)\phi. \quad (27)$$

To show (27), we introduce the shorthand notation $\gamma = 1 - 2\lambda_\omega \in [0, 1]$. With the new notation γ , the proof is straightforward by:

$$\begin{aligned} \tan((1 - 2\lambda_\omega)\phi/2) &\stackrel{(a)}{\leq} \gamma \tan(\phi/2) \stackrel{(b)}{\leq} \gamma\phi < (\gamma + \lambda_\omega)\phi \\ &= (1 - \lambda_\omega)\phi, \end{aligned}$$

where (a) follows from the convexity of $\tan(x)$ in $x \in [0, \pi/2]$ and (b) follows from $\tan(\phi/2) < \phi$ for all $\phi \in (0, 2\pi/3]$. Since $g'(\lambda_\omega) < 0$ for all $\lambda_\omega \in (0, 0.5]$, $g(\lambda_\omega)$ is monotonically decreasing and thus, from (24), it follows that $|Q_{\ell_1\ell_2}|^2 - |R_{\ell_1\ell_2}|^2 < 0$ for $L > 1$.

We conclude by noting that a sharper upper bound on d_k can be obtained by directly computing the bounds for $1\text{-MSE}_\omega -$

0-MSE_ω , i.e.,

$$1\text{-MSE}_\omega - 0\text{-MSE}_\omega = -\varepsilon_{\text{CSI}}\lambda_\omega(1 - \lambda_\omega) + \frac{2\lambda_\omega(1 + \varepsilon_{\text{cor}})}{B} \times (f_L(\lambda_\omega\theta) - f_L((1 - \lambda_\omega)\theta) - (1 - \lambda_\omega)(1 - f_L(\theta))) < 0,$$

for all $\lambda_\omega \in (0, 0.5]$, but we leave an analysis of such refined bounds for future work.

REFERENCES

- [1] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [2] T. L. Marzetta, "Non-cooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Comm.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [3] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: How many antennas do we need?" in *Proc. Allerton Conf. Commun., Contr., Comput.*, Sept. 2011, pp. 545–550.
- [4] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge Univ. Press, 2003.
- [5] T. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 847–860, Oct. 2014.
- [6] C. Jeon, R. Ghods, A. Maleki, and C. Studer, "Optimality of large MIMO detection via approximate message passing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 1227–1231.
- [7] C. Jeon, A. Maleki, and C. Studer, "On the performance of mismatched data detection in large MIMO systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 180–184.
- [8] M. Wu, B. Yin, G. Wang, C. Dick, J. Cavallaro, and C. Studer, "Large-scale MIMO detection for 3GPP LTE: Algorithm and FPGA implementation," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 916–929, Oct. 2014.
- [9] B. Yin, "Low complexity detection and precoding for massive MIMO systems: Algorithm, architecture, and application," Ph.D. dissertation, Rice University, Dec. 2014.
- [10] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "Conjugate gradient-based soft-output detection and precoding in massive MIMO systems," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2014, pp. 3696–3701.
- [11] Z. Wu, C. Zhang, Y. Xue, S. Xu, and X. You, "Efficient architecture for soft-output massive MIMO detection with Gauss-Seidel method," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2016, pp. 1886–1889.
- [12] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer, "FPGA design of a coordinate descent data detector for large-scale MU-MIMO," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2016, pp. 1894–1897.
- [13] *IEEE Draft Standard; Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications; Amendment 4: Enhancements for Higher Throughput*, P802.11n/D3.0, Sep. 2007.
- [14] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation," 3rd Generation Partnership Project (3GPP), TS 36.211, Jan. 2016. [Online]. Available: <http://www.3gpp.org/ftp/Specs/html-info/36211.htm>
- [15] D. Cescato, M. Borgmann, H. Bölcskei, J. Hansen, and A. Burg, "Interpolation-based QR decomposition in MIMO-OFDM systems," in *Proc. IEEE Int. Workshop Signal Process. Advances Wireless Commun. (SPAWC)*, June 2005, pp. 945–949.
- [16] M. Borgmann and H. Bölcskei, "Interpolation-based efficient matrix inversion for MIMO-OFDM receivers," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, Nov. 2004, pp. 1941–1947 Vol.2.
- [17] D. Cescato and H. Bölcskei, "Algorithms for interpolation-based QR decomposition in MIMO-OFDM systems," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1719–1733, Apr. 2011.
- [18] —, "QR decomposition of Laurent polynomial matrices sampled on the unit circle," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4754–4761, Sep. 2010.
- [19] L. W. Chai, P. L. Chiu, and Y. H. Huang, "Reduced-complexity interpolation-based QR decomposition using partial layer mapping," in *Proc. IEEE Int. Symp. Circuits and Syst. (ISCAS)*, May 2011, pp. 2381–2384.

- [20] P. L. Chiu, L. Z. Huang, L. W. Chai, and Y. H. Huang, "Interpolation-based QR decomposition and channel estimation processor for MIMO-OFDM system," *IEEE Trans. Circuits Syst. I*, vol. 58, no. 5, pp. 1129–1141, May 2011.
- [21] 3GPP, "Study on new radio (NR) access technology," 3rd Generation Partnership Project (3GPP), Technical report (TR) 38.912, Jul. 2018, version 15.0.0. [Online]. Available: <http://www.3gpp.org/DynaReport/38912.htm>
- [22] J. Jeon, "NR wide bandwidth operations," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 42–46, Mar. 2018.
- [23] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 239–241, Jun. 2002.
- [24] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1051–1064, May 2002.
- [25] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.
- [26] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, 2011.
- [27] M. Čirkić and E. Larsson, "SUMIS: Near-optimal soft-in soft-out MIMO detection with low and fixed complexity," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3084–3097, Jun. 2014.
- [28] J. W. Choi, B. Lee, B. Shim, and I. Kang, "Low complexity detection and precoding for massive MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, April 2013, pp. 2857–2861.
- [29] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.
- [30] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [31] C. Dessel, E. Björnson, S. Kashyap, E. G. Larsson, C. Mollén, L. Liu, S. Malkowsky, H. Prabhu, Y. Liu, J. Vieira, O. Edfors, E. Karipidis, F. Dielacher, and D. V. Pop, "Distributed and centralized baseband processing algorithms, architectures, and platforms," MAMMOET, Tech. Rep. ICT-619086-D3.2, 2016.
- [32] H. Bölcskei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.
- [33] C. Jeon, K. Li, J. R. Cavallaro, and C. Studer, "On the achievable rates of decentralized equalization in massive MU-MIMO systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1102–1106.
- [34] P. L. Butzer and R. J. Nessel, *Fourier analysis and approximation*. Academic Press New York, 1971.
- [35] M. Wu, B. Yin, K. Li, C. Dick, J. R. Cavallaro, and C. Studer, "Implicit vs. explicit approximate matrix inversion for wideband massive MU-MIMO data detection," *J. Signal Process. Syst.*, vol. 90, no. 10, pp. 1311–1328, Oct. 2018.
- [36] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.
- [37] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sep. 2001.