

Performance Analysis of Channel Extrapolation in FDD Massive MIMO Systems

François Rottenberg, *Member, IEEE*, Thomas Choi, *Student Member, IEEE*,
Peng Luo, *Student Member, IEEE*, Jianzhong Zhang, *Fellow, IEEE*
and Andreas F. Molisch, *Fellow, IEEE*

Abstract

Channel estimation for the downlink of frequency division duplex (FDD) massive MIMO systems is well known to generate a large overhead as the amount of training generally scales with the number of transmit antennas in a MIMO system. In this paper, we consider the solution of extrapolating the channel frequency response from uplink pilot estimates to the downlink frequency band, which completely removes the training overhead. We first show that conventional estimators fail to achieve reasonable accuracy. We propose instead to use high-resolution channel estimation. We derive theoretical lower bounds (LB) for the mean squared error (MSE) of the extrapolated channel. Assuming that the paths are well separated, the LB is simplified in an expression that gives considerable physical insight. It is then shown that the MSE is inversely proportional to the number of receive antennas while the extrapolation performance penalty scales with the square of the ratio of the frequency offset and the training bandwidth. The channel extrapolation performance is validated through numeric simulations and experimental measurements taken in an anechoic chamber. Our main conclusion is that channel extrapolation is a viable solution for FDD massive MIMO systems if accurate system calibration is performed and favorable propagation conditions are present.

The work was partly supported by NSF under project ECCS-1731694 and a gift from Samsung America. The work of F. Rottenberg was also partly supported by the Belgian American Educational Foundation (B.A.E.F.). The material in this paper has been partially submitted at IEEE Globecom 2019 [1].

François Rottenberg, Thomas Choi, Peng Luo and Andreas F. Molisch are with the Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, USA (e-mail: {frottenb, choit, luop, molisch}@usc.edu). Jianzhong Zhang is with Samsung Research America, Richardson, TX, USA (e-mail: jianzhong.z@samsung.com).

I. INTRODUCTION

The deployment of massive multiple-input-multiple-output (MIMO) communications systems strongly rely on the acquisition of accurate channel state information (CSI) at base station (BS) [2]. Massive MIMO systems are typically characterized by a much larger number of antennas at the BS than the sum of the antennas at the user equipments (UEs). This implies that channel estimation is much less costly in the uplink (UL) than in the downlink (DL) [3]. In time division duplex (TDD) systems, the BS can efficiently perform DL channel estimation from UL pilot transmission from the UEs, since channel reciprocity holds as long as UL and DL transmission occurs within a coherence time of the channel, and *within the same frequency band*. However, in a frequency division duplex (FDD) scenario, reciprocity cannot be exploited as different bands, usually separated by more than a coherence bandwidth, are used in UL and DL. On the other hand, estimation of the channel by DL pilot transmission and feedback might result in a large overhead.

A variety of methods have been proposed to solve this dilemma, such as channel correlations in the spatial domain reflected in second-order statistics [4], [5], compression of the feedback [6], combinations thereof [7], or compressed sensing methods [8], each of which have their advantages and drawbacks. One of the most promising methods is channel extrapolation from the UL to the DL band as it completely removes the overhead. The extrapolation range of conventional extrapolators is very limited - typically to the order of one coherence bandwidth. To overcome this limit, [9] suggested estimation of the multipath components (MPCs) via high-resolution parameter estimation. Based on the extracted MPCs, extrapolation over large bandwidths can be achieved. However, the paper only considered the single-input-single-output (SISO) case, which resulted in a poor extrapolation performance. In [10], the authors extend the study to the MIMO case and channel extrapolation in space. They perform multiple measurements showing that they can reach a frequency extrapolation range larger than 5 times the coherence bandwidth. In [11], the authors present the so-called R2-F2 system to extract path parameters and infer CSI from a different training band. The authors show how to integrate the system into LTE cellular networks and use experimental measurements for validation. Their study restricted the frequency spacing between UL and DL band to be only 20-30 MHz. The authors in [12] compare different extrapolation algorithms. Their study shows that a super-resolution can outperform compressed sensing methods to perform channel extrapolation in frequency. In [13], only the

information about user angles is extracted from UL pilots using 2D unitary ESPRIT method [14]. Then, they perform directional training in the DL. In [15], the authors propose to train a neural network to perform the channel extrapolation in frequency. Their approach does not require the acquisition of the antenna array patterns through calibration but requires a large training dataset. In [16], the authors propose to acquire DL CSI through UL pilots in combination with a limited feedback from DL pilots.

Channel extrapolation in frequency also presents formal similarities to extrapolation in time. In contrast to frequency-domain extrapolation, channel prediction in time has been extensively investigated in the literature. A comprehensive review can be found in [17]. In [18], the authors proposed performance bounds for prediction in time of MIMO channels. They later extended their study to MIMO-OFDM channel estimation with interpolation and extrapolation being done both in time and frequency [19]. The authors make the observation that MIMO provides much longer prediction lengths than for SISO systems.

To provide understanding of the promise of low-overhead FDD massive MIMO systems, this paper investigates the performance of channel extrapolation in frequency. The main originality of our paper is that the analysis combines at the same time theoretical limits, numerical simulations and experimental validations. We believe this joint approach gives much insight on the complex problem at hand. In particular, most of the previous approaches were driven only numerically and/or experimentally but not theoretically. This paper extends our previous work [1] by: (i) considering a more general channel model taking into account the frequency dependence of the pattern; (ii) analyzing the limitations of conventional channel estimators in terms of channel extrapolation; (iii) deriving a closed-form expression of the extrapolation range; (iv) including a comparison of theoretical and experimental results. More specifically, our contributions can be summarized as follows. Section II first describes the considered transmission model.

Section III highlights the advantages of high-resolution channel estimation compared to conventional least squares (LS) and linear minimum mean squared (LMMSE) channel estimation. More specifically, we show that it results in potential large signal-to-noise (SNR) gains and extrapolation range.

In Section IV, we formulate a theoretical lower bound (LB) on the mean squared error (MSE) of the extrapolated channel, using a similar methodology as in [19]. The proposed LB differs from [19] as it takes into account elevation angles, the frequency dependence of the pattern, and the influence of the training pulse. Furthermore, a simplified LB is also proposed, giving more

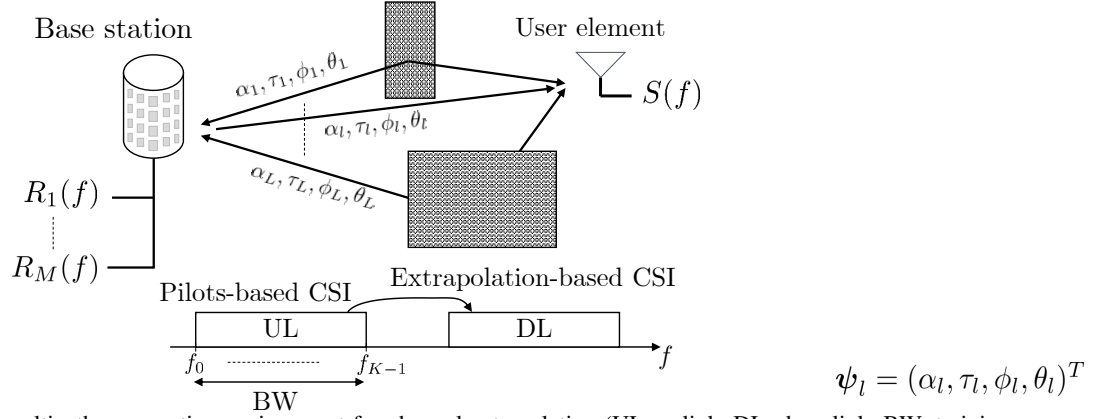


Fig. 1. Massive MIMO multipath propagation environment for channel extrapolation (UL: uplink, DL: downlink, BW: training bandwidth, CSI: channel state information).

physical intuition on the extrapolation range and the SNR gain that can be expected in practice. The validity of the LB relies on the strong assumption that the paths are “well separated”, which makes particular sense in a massive MIMO context. Under this condition, we show that the MSE of the extrapolated channel frequency response is inversely proportional to the number of receive antennas while the extrapolation performance penalty scales with the square of the ratio of the frequency offset from the carrier frequency and the training bandwidth.

In Section V, we analyze the performance of the theoretical LB by numerical simulations using a 3GPP channel model showing that the LB can be reached by practical high-resolution parameter extraction algorithms such as the space-alternating generalized expectation-maximization (SAGE) algorithm [20]. Our results show the very limited extrapolation range that is achieved in SISO systems, while much more promising results are obtained in the MIMO setting as the paths can be more easily separated in the delay-angle domain. We also show that the cylindrical array manufactured in our lab reaches similar extrapolation performance as a synthetic rectangular planar array.

Finally, in Section VI, using a cylindrical array and an omnidirectional transmit antenna, we perform channel measurements in an anechoic chamber on a 400 MHz bandwidth at a carrier frequency of 3.5 GHz. We perform high-resolution channel parameter estimation on 40 MHz of the total 400 MHz band used as the “training bandwidth”. Then, we evaluate the channel extrapolation performance by comparing the channel extrapolation relying on the extracted paths to the directly measured “ground truth” channels in the remaining 360 MHz. These measurements allow to evaluate the influence of calibration and channel modelling errors and demonstrate the

¹The “well separated” assumption will be properly formalized in Section IV.

accuracy of the theoretical LB previously derived. In particular, the simplified LB derived from Section IV is shown to provide accurate results even though its expression is very simple to evaluate and does not require any *a priori* information on the path parameters and the antenna array pattern.

Notations: Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively (resp.). Superscripts $*$, T and H stand for conjugate, transpose and Hermitian transpose operators. The symbols j , $\mathbb{E}(\cdot)$, $\Im(\cdot)$ and $\Re(\cdot)$ denote the imaginary unit, expectation, imaginary and real parts, respectively. The norm $\|\mathbf{A}\|$ is the Frobenius norm and δ_n is the Kronecker delta. The $\text{diag}(\cdot)$ operator applied to a vector returns a diagonal matrix whose k -th diagonal entry is equal to the k -th entry of the argument vector.

II. SIMO TRANSMISSION MODEL

We consider FDD massive MIMO scenarios where each user has a single-antenna and transmits an UL training sequence that is orthogonal to those of the other users. Thus, the estimations for different users become independent and in particular, the extrapolation in frequency of the SIMO channel of each user can be treated independently. Moreover, to highlight the gain of having antenna arrays, the SISO case is studied in parallel.

Let us consider the transmission of a baseband equivalent pulse of bandwidth (BW) *a priori* known by the receiver and whose Fourier transform is denoted by $S(f)$, as depicted in Fig. 1. We assume that the channel is quasi-static, *i.e.*, constant for the duration of the transmission. M denotes the number of antennas of the receive array. Antennas at transmit and receive sides are assumed to be vertically polarized. We also assume that the propagation channel is composed of L specular paths, where each path is completely characterized by its deterministic parameters: complex gain $\alpha_l = \Re(\alpha_l) + j\Im(\alpha_l)$, delay τ_l , azimuth angle ϕ_l and elevation angle θ_l . We denote by $a_m(\phi, \theta, f)$ the pattern of antenna m evaluated in the direction (ϕ, θ) and at frequency f . Note that the frequency dependence of the array pattern cannot generally be omitted for our scenario, depending on the considered training and extrapolated frequency range. More specifically, the frequency selectivity of $a_m(\phi, \theta, f)$ comes from two contributions: firstly the frequency dependence of each individual antenna pattern and secondly the frequency dependent phase shift across the antenna array elements (beam squinting).² The receiver obtains

²This dependence is often neglected in the literature when the ratio of the dimension of the array to the speed of light is much smaller than the inverse of the bandwidth of the signal.

K frequency samples and the baseband complex received sample at antenna m and baseband frequency f_k can be expressed as

$$R_m(f_k) = \sum_{l=1}^L \alpha_l a_m(\phi_l, \theta_l, f_k) e^{-j2\pi f_k \tau_l} S(f_k) + W_m(f_k), \quad (1)$$

where $W_m(f_k)$ is zero mean additive complex circularly symmetric Gaussian noise. We assume that the noise samples are white, *i.e.*, $\mathbb{E}(W_m(f_k)W_m^*(f_{k'})) = \sigma_w^2 \delta_{m-m'} \delta_{k-k'}$. Note that the model in (1) straightforwardly applies to OFDM systems.

The SISO case can be seen as a special case of the SIMO case where each ray is completely characterized by its complex gain $\tilde{\alpha}_l$ and its delay τ_l while the information on the angles of arrival is lost. In the following, the index “ m ” will be omitted when the single-antenna case is considered. Equation (1) becomes

$$R(f_k) = \sum_{l=1}^L \tilde{\alpha}_l e^{-j2\pi f_k \tau_l} \tilde{S}(f_k) + W(f_k). \quad (2)$$

Note that the SISO model in (2) relies on the underlying assumption that the receive antenna pattern $a(\phi, \theta, f)$ is frequency-angle separable, *i.e.*, $a(\phi, \theta, f) = b(\phi, \theta)G(f)$. This allows to take the angle dependence of the pattern $b(\phi, \theta)$ into account in by the channel complex gains $\tilde{\alpha}_l = \alpha_l b(\phi_l, \theta_l)$ while the remaining frequency dependence $G(f)$ can be taken into account by the pulse $\tilde{S}(f_k) = S(f_k)G(f_k)$. Furthermore, both the models in (1) and (2) rely on a similar assumption at the transmit side: the transmit antenna has a frequency-angle separable pattern so that its dependence in the angles of departure can be accounted for in the coefficients α_l while the frequency dependence can be accounted for in the transmit pulse $S(f)$.

We finally note that a number of straightforward generalizations of this model can be made: (i) V and H polarizations can be taken into account by representing the path amplitudes as 2×1 vectors and the array patterns as 2×2 polarimetric matrices. (ii) multiple antenna elements at the UEs can also be taken into account similarly; (iii) scatterers in the nearfield can be described by replacing the plane wave model of each path by a spherical wave model, where the wavefront curvature of each path is now an additional parameter of the model. However, for ease of exposition we use the simplified models of (1) and (2) in the remainder of this paper.

III. CHANNEL EXTRAPOLATION

Let us define $\boldsymbol{\psi} = (\boldsymbol{\psi}_1^T, \dots, \boldsymbol{\psi}_L^T)^T \in \mathbb{R}^{5L \times 1}$ and $\boldsymbol{\psi}_l = (\tau_l, \phi_l, \theta_l, \Re(\alpha_l), \Im(\alpha_l))^T \in \mathbb{R}^{5 \times 1}$ as the vectors containing all real path parameters and the real parameters of each path respectively. We

define the channel frequency response evaluated at frequency f and antenna m as

$$H_m(f) \triangleq \sum_{l=1}^L \alpha_l a_m(\phi_l, \theta_l, f) e^{-j2\pi f \tau_l}. \quad (3)$$

and the MSE of an estimate $\hat{H}_m(f)$ of $H_m(f)$ as

$$\text{MSE}(f, \boldsymbol{\psi}) \triangleq \mathbb{E} |\hat{H}_m(f) - H_m(f)|^2, \quad (4)$$

where the expectation is taken over the noise realizations for deterministic path parameters $\boldsymbol{\psi}$. In the following, we review two conventional low resolution channel estimators: LS and LMMSE estimator. We will highlight that these estimations relying on a linear combination of the received frequency samples results in an extrapolation range limited by the coherence bandwidth of the channel. In the light of its limitations, we detail our motivations for going towards high-resolution channel estimation.

A. Conventional Low-Resolution Estimation

1) *LS estimator*: LS estimators perform a simple per-antenna estimation at each frequency point as

$$\hat{H}_m^{\text{LS}}(f_k) = \frac{R_m(f_k)}{S(f_k)} = H_m(f_k) + \frac{W_m(f_k)}{S(f_k)}.$$

We can easily see that the LS estimator is unbiased and is only limited by additive noise. Since the noise samples $w_m[n]$ have variance σ_w^2 , we can write

$$\text{MSE}_{\text{LS}}(f_k, \boldsymbol{\psi}) \triangleq \mathbb{E} |\hat{H}_m^{\text{LS}}(f_k) - H_m(f_k)|^2 = \frac{\sigma_w^2}{|S(f_k)|^2}.$$

We define the total training energy as $E_T = \sum_{k=0}^{K-1} |S(f_k)|^2$ and the SNR as $\text{SNR} \triangleq E_T / \sigma_w^2$. If the training signal is chosen so that the energy is uniformly distributed across the frequency bins, *i.e.*, $|S(f_k)|^2 = \frac{E_T}{K}$, the MSE becomes constant across frequency and

$$\text{MSE}_{\text{LS}}(f_k, \boldsymbol{\psi}) = \frac{\sigma_w^2 K}{E_T} = \frac{K}{\text{SNR}}. \quad (5)$$

Let us assume that the received samples are uniformly spaced in frequency and their frequency spacing is lower than the inverse of the maximal delay spread of the channel $\tau_{\max}$ so that no aliasing is present. Then, the Whittaker-Shannon formula can be used to interpolate every frequency points in the training bandwidth. Furthermore, if the frequency samples are spaced more closely than $1/\tau_{\max}$, their correlation might be leveraged to improve the MSE. This gain is

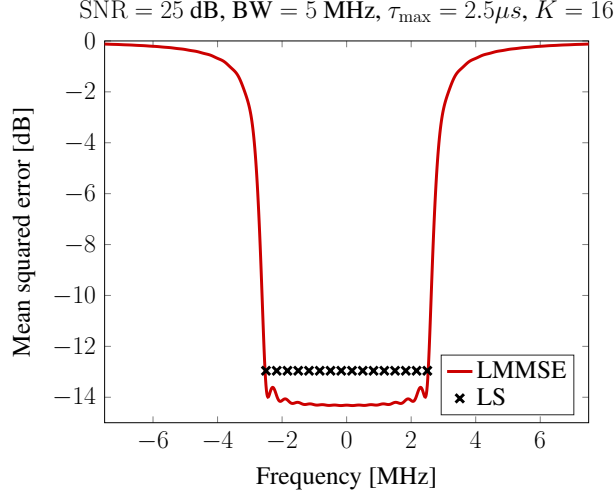


Fig. 2. Channel extrapolation methods based on a linear combination of received frequency samples such as LMMSE estimators give a very limited extrapolation range.

properly taken into account by the LMMSE estimator described below. This gain is also similar to a multicarrier system that would convert its pilot-based frequency channel estimates to the time domain, truncate the obtained impulse response to significant coefficients and finally convert it back to frequency [21].

2) *LMMSE estimator*: Based on the LS frequency estimates, simple linear extrapolation methods can be used. However, most of these methods would have a very limited extrapolation range, on the order of the coherence bandwidth of the channel. To demonstrate this, we will study the performance of the LMMSE estimator [22] in terms of channel extrapolation. The LMMSE frequency estimate at frequency f is simply given by

$$\hat{H}_m^{\text{LMMSE}}(f) = \mathbf{c}_{\text{LS}}^H(f) \mathbf{C}_{\text{LS}}^{-1} \hat{\mathbf{h}}_{\text{LS},m},$$

where $\hat{\mathbf{h}}_{\text{LS},m} = (\hat{H}_m^{\text{LS}}(f_0), \dots, \hat{H}_m^{\text{LS}}(f_{K-1}))^T$, $\mathbf{C}_{\text{LS}} = \mathbb{E}(\hat{\mathbf{h}}_{\text{LS},m} \hat{\mathbf{h}}_{\text{LS},m}^H)$ and $\mathbf{c}_{\text{LS}}^H(f) = \mathbb{E}(H_m(f) \hat{\mathbf{h}}_{\text{LS},m}^H)$.

The MSE is then given by

$$\text{MSE}_{\text{LMMSE}}(f) = \mathbb{E}(|H_m(f)|^2) - \mathbf{c}_{\text{LS}}^H(f) \mathbf{C}_{\text{LS}}^{-1} \mathbf{c}_{\text{LS}}(f),$$

where the expectation is taken here with respect to the statistics of both the noise and the parameters ψ . An implementation challenge of the LMMSE estimator is that it requires the knowledge of the autocorrelation function of the channel frequency response

$$C_{H,m}(\Delta f) \triangleq \mathbb{E}(H_m(f + \Delta f) H_m(f)^*). \quad (6)$$

This function depends on the joint distribution of the path parameters ψ . To highlight the limitations of the LMMSE estimator in terms of extrapolation, let us consider a simple optimistic

example where the receiver would have perfect knowledge of the noise variance σ_w^2 and the frequency autocorrelation of the channel $C_{H,m}(\Delta f)$. As in [22], we assume that the paths gains and delays are i.i.d. For the sake of simplicity, we here assume a frequency independent and isotropic array pattern so that $a_m(\phi, \theta, f) = a_m(\phi, \theta)$ and $|a_m(\phi, \theta)| = 1$. Furthermore, the delay of each path τ_l has a uniform distribution in $[0, \tau_{\max}]$ and the conditional distribution of the variable $\alpha_l; \tau_l$ is a zero-mean complex circularly symmetric Gaussian with variance $p(\tau_l) = e^{-\tau_l/\tau_{\text{rms}}}$. This gives (see Appendix VIII-A)

$$C_{H,m}(\Delta f) = \frac{1}{1 - e^{-\tau_{\max}/\tau_{\text{rms}}}} \frac{1 - e^{-\tau_{\max}/\tau_{\text{rms}} - j2\pi\Delta f\tau_{\max}}}{1 + j2\pi\Delta f\tau_{\text{rms}}}.$$

Fig. 2 shows the performance of the LS and LMMSE estimators. The training signal was chosen to have uniform energy distribution across frequency. The training bandwidth is set to $\text{BW} = 5$ MHz and the spacing between each frequency point is set to $0.8/\tau_{\max}$, which is slightly faster than the Nyquist rate $1/\tau_{\max}$. The delay parameters are fixed to $\tau_{\max} = 2.5\mu\text{s}$ and $\tau_{\text{rms}} = 0.5\mu\text{s}$. By definition, the LS estimator is restricted to the training bandwidth. On the other hand, the extrapolation range of the LMMSE estimator is extremely limited and only slightly improves the range of the LS estimator. The extrapolation range is on the order of the coherence bandwidth or $1/\tau_{\max} = 400$ kHz. Furthermore, we here assumed an optimistic case where the receiver has perfect knowledge of $C_{H,m}(\Delta f)$. This is generally not the case and we do not make this assumption in the following derivations.

B. High-Resolution Estimation

In the light of the previous discussions, we are interested in finding better channel estimators, which could extrapolate frequency way beyond one coherence bandwidth, to address typical FDD massive MIMO systems. For instance, let us take a simple example where the channel only has a few well separated specular multipath components, *i.e.*, $L \ll K$. Then, an intuitive reasoning suggests to go for high-resolution estimation of the L different path parameters directly [9]. There are two main motivations for this: **SNR gain** and **extrapolation**. The **SNR gain** would come from two sources. First, the receiver only has to estimate L complex coefficients rather than K , resulting in a potential SNR gain of $\frac{K}{L}$ with respect to LS estimation. One should note that this gain is stronger than the potential gain of using a frequency correlation filter as in [22]. It does not come from simply assuming that the channel impulse response has a finite length KT inducing correlation in the frequency domain but from its sparsity in the delay-angle domain.

Secondly, the received signal at each antenna can be coherently combined to jointly estimate and separate all path parameters instead of performing per-antenna independent channel estimation as in the LS case, which results in a potential total SNR gain of a factor $\frac{MK}{L}$ with respect to LS estimation. One should here note that an improved LMMSE estimators leveraging correlation in the spatial domain through the antennas could also achieve a similar gain.

By definition, low-resolution estimators are restricted to the bandwidth occupied by the training signal. However, high-resolution estimates of the path parameters allow for simple channel **extrapolation** in frequency possibly very far from the center frequency of the training band. If we denote by $\hat{\tau}_l$, $\hat{\phi}_l$, $\hat{\theta}_l$ and $\hat{\alpha}_l$ the high-resolution estimates of τ_l , ϕ_l , θ_l and α_l respectively, we can write the expression of the extrapolated channel as

$$\hat{H}_m(f, \hat{\psi}) = \sum_{l=1}^L \hat{\alpha}_l a_m(\hat{\phi}_l, \hat{\theta}_l, f) e^{-j2\pi f \hat{\tau}_l}. \quad (7)$$

Of course, intuitive reasoning tells us that the extrapolated channel will suffer from the estimation errors on the path parameters, especially as the extrapolation range becomes large. We also assume here that the parameters of the MPCs are independent of frequency. This is well fulfilled in most practical situations [23], since the range over which these parameters change is on the order of 10% of the carrier frequency, which is much larger than the extrapolation range we can usually obtain, see Section V.

IV. PERFORMANCE ANALYSIS

This section aims at theoretically formalizing and analyzing the two potential gains of high-resolution channel estimation previously described: SNR gain and extrapolation factor. To do this, we will in a first step derive the Fisher information matrix of the estimated path parameters. The second step will consist of deriving a lower bound on the MSE of the extrapolated channel. In a third step, a simplified LB will be derived giving much more physical intuition. Finally, the bounds for the MIMO case will be particularized to the SISO case.

A. Fisher Information Matrix

Let us define the vector $\mathbf{r} \in \mathbb{C}^{KM \times 1}$ as containing all received samples for all antennas and observation samples. Given the independence of the noise samples, the log-likelihood of vector $\mathbf{r}$ becomes

$$L(\mathbf{r}; \psi) = \sum_{k=0}^{K-1} \sum_{m=1}^M L(R_m(f_k); \psi).$$

The elements of the Fisher information matrix $\mathbf{I}_\psi \in \mathbb{R}^{5L \times 5L}$ can be obtained from the log-likelihood function as [24]

$$\begin{aligned} [\mathbf{I}_\psi]_{u,v} &= -\mathbb{E} \left(\frac{\partial^2 L(\mathbf{r}; \psi)}{\partial \psi_u \partial \psi_v} \right) \\ &= -\sum_{k=0}^{K-1} \sum_{m=1}^M \mathbb{E} \left(\frac{\partial^2 L(R_m(f_k); \psi)}{\partial \psi_u \partial \psi_v} \right), \end{aligned} \quad (8)$$

where the expectation is taken over the noise distribution. Since the random variable $R_m(f_k); \psi$ follows a circularly symmetric complex normal distribution with variance σ_w^2 and mean $\mu_{m,k} \triangleq \sum_{l=1}^L \alpha_l a_m(\phi_l, \theta_l, f_k) S(f_k) e^{-j2\pi f_k \tau_l}$, equation (8) can be rewritten as

$$[\mathbf{I}_\psi]_{u,v} = \frac{2}{\sigma_w^2} \sum_{k=0}^{K-1} \sum_{m=1}^M \Re \left\{ \frac{\partial \mu_{m,k}^*}{\partial \psi_u} \frac{\partial \mu_{m,k}}{\partial \psi_v} \right\}. \quad (9)$$

The full $5L \times 5L$ Fisher information matrix $\mathbf{I}_\psi$ can be partitioned into L^2 submatrices $\mathbf{I}_{\psi_l, \psi_{l'}} \in \mathbb{R}^{5 \times 5}$ as

$$\mathbf{I}_\psi = \frac{2}{\sigma_w^2} \begin{pmatrix} \mathbf{I}_{\psi_1, \psi_1} & \cdots & \mathbf{I}_{\psi_1, \psi_L} \\ \vdots & \ddots & \vdots \\ \mathbf{I}_{\psi_L, \psi_1} & \cdots & \mathbf{I}_{\psi_L, \psi_L} \end{pmatrix}, \quad (10)$$

and the structure of the submatrices $\mathbf{I}_{\psi_l, \psi_{l'}}$ is given by

$$\mathbf{I}_{\psi_l, \psi_{l'}} = \begin{pmatrix} I_{\tau_l \tau_{l'}} & I_{\tau_l \phi_{l'}} & I_{\tau_l \theta_{l'}} & I_{\tau_l \alpha_{l'}^R} & I_{\tau_l \alpha_{l'}^I} \\ I_{\phi_l \phi_{l'}} & I_{\phi_l \phi_{l'}} & I_{\phi_l \theta_{l'}} & I_{\phi_l \alpha_{l'}^R} & I_{\phi_l \alpha_{l'}^I} \\ I_{\theta_l \theta_{l'}} & I_{\phi_l \theta_{l'}} & I_{\theta_l \theta_{l'}} & I_{\theta_l \alpha_{l'}^R} & I_{\theta_l \alpha_{l'}^I} \\ I_{\tau_l \alpha_{l'}^R} & I_{\phi_l \alpha_{l'}^R} & I_{\theta_l \alpha_{l'}^R} & I_{\alpha_l^R \alpha_{l'}^R} & I_{\alpha_l^R \alpha_{l'}^I} \\ I_{\tau_l \alpha_{l'}^I} & I_{\phi_l \alpha_{l'}^I} & I_{\theta_l \alpha_{l'}^I} & I_{\alpha_l^R \alpha_{l'}^I} & I_{\alpha_l^I \alpha_{l'}^I} \end{pmatrix}.$$

Defining $\dot{a}_{m,\phi}(\phi, \theta, f) \triangleq \frac{da_m(\phi, \theta, f)}{d\phi}$ and $\dot{a}_{m,\theta}(\phi, \theta, f) \triangleq \frac{da_m(\phi, \theta, f)}{d\theta}$, we can write the partial derivatives appearing in (9) as

$$\begin{aligned} \frac{d\mu_{m,k}}{d\tau_l} &= \alpha_l a_m(\phi_l, \theta_l, f_k) (-j2\pi f_k) S(f_k) e^{-j2\pi f_k \tau_l} \\ \frac{d\mu_{m,k}}{d\phi_l} &= \alpha_l \dot{a}_{m,\phi}(\phi, \theta, f_k) S(f_k) e^{-j2\pi f_k \tau_l} \\ \frac{d\mu_{m,k}}{d\theta_l} &= \alpha_l \dot{a}_{m,\theta}(\phi, \theta, f_k) S(f_k) e^{-j2\pi f_k \tau_l} \\ \frac{d\mu_{m,k}}{d\alpha_l^R} &= a_m(\phi_l, \theta_l, f_k) S(f_k) e^{-j2\pi f_k \tau_l} \\ \frac{d\mu_{m,k}}{d\alpha_l^I} &= j a_m(\phi_l, \theta_l, f_k) S(f_k) e^{-j2\pi f_k \tau_l}. \end{aligned}$$

Inserting these partial derivatives in (9) and for a specific array pattern $a_m(\phi, \theta, f)$, the Fisher information matrix in (10) can be easily constructed. In the following, we will make the following assumption.

(As1): the Fisher information matrix $\mathbf{I}_\psi$ is nonsingular.

In practice, a rank deficiency of $\mathbf{I}_\psi$ could arise if two rays, or more, become extremely close in delay and angle, which would cause the determinant of $\mathbf{I}_\psi$ to go to zero. A solution in this case can be to replace the two correlated rays with one ray whose amplitude is the sum of the amplitudes of the components. It is intuitive that this operation will not cause a large information loss if the rays are close enough. However, care must be taken not to merge paths whose distance in the delay domain is comparable to or larger than the inverse of the extrapolation range.

B. Lower Bound on the MSE of the Extrapolated Channel

Let us denote by $\hat{\psi} \in \mathbb{R}^{5L \times 1}$ an unbiased estimator of ψ with covariance matrix

$$\mathbf{C}_{\hat{\psi}} = \mathbb{E} \left(\left(\psi - \hat{\psi} \right) \left(\psi - \hat{\psi} \right)^T \right),$$

where the expectation is taken over the noise distribution. The Cramer-Rao lower bound (CRLB) tells us that the matrix $\mathbf{C}_{\hat{\psi}} - \mathbf{I}_\psi^{-1}$ is positive semidefinite, which implies that $\mathbf{g}^H \mathbf{C}_{\hat{\psi}} \mathbf{g} \geq \mathbf{g}^H \mathbf{I}_\psi^{-1} \mathbf{g}$ for every vector $\mathbf{g} \in \mathbb{C}^{5L \times 1}$. If vector $\mathbf{g}$ is chosen as an all zero vector except a one at u -th entry, we get a LB for the variance of the estimated parameter ψ_u . The CRLB only provides a LB on the variance of the estimated parameters while we are interested on the variance on the error of the extrapolated channel defined in (4). To obtain a performance limit, we would like to lower bound the MSE by a certain quantity $\text{LB}_m(f, \hat{\psi})$ so that

$$\text{MSE}_m(f, \hat{\psi}) \geq \text{LB}_m(f, \psi),$$

where $\text{LB}_m(f, \psi)$ would only depend on deterministic parameters. The following theorem gives a closed-form expression of the LB on the extrapolated channel as a function of the path parameters ψ and the extrapolated frequency f .

Theorem 1. *Under (As1), the LB on the MSE of the extrapolation error $\text{LB}_m(f, \psi)$ for any unbiased estimator $\hat{H}_m(f, \psi)$ can be expressed as*

$$\text{LB}_m(f, \psi) \triangleq \mathbf{g}_{m,f,\psi}^H \mathbf{I}_\psi^{-1} \mathbf{g}_{m,f,\psi},$$

with the following vector definitions

$$\begin{aligned}
\mathbf{g}_{m,f,\psi} &= (\mathbf{g}_{m,f,\psi_1}^T, \dots, \mathbf{g}_{m,f,\psi_L}^T)^T \\
\mathbf{g}_{m,f,\psi_l} &= (g_{m,f,\tau_l}, g_{m,f,\phi_l}, g_{m,f,\theta_l}, g_{m,f,\alpha_l^I}, g_{m,f,\alpha_l^R})^T \\
g_{m,f,\tau_l} &= (-j2\pi f)\alpha_l a_m(\phi_l, \theta_l, f)e^{-j2\pi f\tau_l} \\
g_{m,f,\phi_l} &= \alpha_l \dot{a}_{m,\phi}(\phi_l, \theta_l, f)e^{-j2\pi f\tau_l} \\
g_{m,f,\theta_l} &= \alpha_l \dot{a}_{m,\theta}(\phi_l, \theta_l, f)e^{-j2\pi f\tau_l} \\
g_{m,f,\alpha_l^R} &= a_m(\phi_l, \theta_l, f)e^{-j2\pi f\tau_l} \\
g_{m,f,\alpha_l^I} &= ja_m(\phi_l, \theta_l, f)e^{-j2\pi f\tau_l}.
\end{aligned}$$

Proof. The proof is given in Appendix VIII-B and relies on the application of the CRLB formula for transformation of parameters [24]. $\square$

C. Separated Rays

The LB of Theorem 1 is in closed-form, which allows to easily evaluate it numerically. However, it requires the inversion of the Fisher information matrix and does not provide much intuition on the exact SNR gain and extrapolation range that we can expect. To further characterize and try to gain more insight, let us introduce the set of assumptions (As2) – (As4).

(As2): the array pattern is non frequency selective, *i.e.*, $a_m(\phi, \theta, f) = a_m(\phi, \theta)$.

This assumption makes sense if the antenna patterns are flat in the considered band and if the ratio of the dimension of the array to the speed of light is much smaller than that the inverse of the extrapolation range. In the remaining part of this section, we assume that (As2) holds, and we drop the frequency dependence of the array. We define the following vectors in order to introduce assumptions (As3) – (As4)

$$\begin{aligned}
\mathbf{s}_l &\triangleq \left(S(f_0)e^{-j2\pi f_0\tau_l} \quad \dots \quad S(f_{K-1})e^{-j2\pi f_{K-1}\tau_l} \right)^T \\
\dot{\mathbf{s}}_l &\triangleq -j2\pi(f_0 S(f_0)e^{-j2\pi f_0\tau_l}, \dots, f_{K-1} S(f_{K-1})e^{-j2\pi f_{K-1}\tau_l})^T \\
\mathbf{a}_l &\triangleq \left(a_1(\phi_l, \theta_l) \quad \dots \quad a_M(\phi_l, \theta_l) \right)^T \in \mathbb{C}^{M \times 1} \\
\dot{\mathbf{a}}_{l,\phi} &\triangleq \left(\dot{a}_{1,\phi}(\phi_l, \theta_l) \quad \dots \quad \dot{a}_{M,\phi}(\phi_l, \theta_l) \right)^T \in \mathbb{C}^{M \times 1} \\
\dot{\mathbf{a}}_{l,\theta} &\triangleq \left(\dot{a}_{1,\theta}(\phi_l, \theta_l) \quad \dots \quad \dot{a}_{M,\theta}(\phi_l, \theta_l) \right)^T \in \mathbb{C}^{M \times 1}.
\end{aligned}$$

(As3): separation of the L specular rays in delay, azimuth angle and/or elevation angle. We assume that, for each pair of rays l, l' ($l \neq l'$), at least one of the following two relationships is verified:

(1) Separation in delay:

$$\mathbf{s}_l^H \mathbf{s}_{l'} = \dot{\mathbf{s}}_l^H \dot{\mathbf{s}}_{l'} = \dot{\mathbf{s}}_l^H \mathbf{s}_{l'} = 0. \quad (11)$$

(2) Separation in azimuth and/or elevation angle:

$$\mathbf{a}_l^H \mathbf{a}_{l'} = \dot{\mathbf{a}}_{l,\theta}^H \dot{\mathbf{a}}_{l,\theta} = \dot{\mathbf{a}}_{l,\phi}^H \dot{\mathbf{a}}_{l,\phi} = \dot{\mathbf{a}}_{l,\theta}^H \dot{\mathbf{a}}_l = \dot{\mathbf{a}}_{l,\phi}^H \dot{\mathbf{a}}_l = \dot{\mathbf{a}}_{l,\phi}^H \dot{\mathbf{a}}_{l',\theta} = 0.$$

The assumption (As3) is a strong assumption, whose accuracy will typically depend on different parameters. The specular paths will generally become more separated in delay as the bandwidth of $S(f)$ increases, inducing higher resolution in delay. Similarly, the separation in azimuth and elevation will be improved as the number of antenna elements M is increased. More generally, the validity of (As3) will depend on the training signal $S(f_k)$, on the array pattern $a_m(\phi, \theta)$ and on the extrapolation range.

(As4): the training signal $S(f)$ has a symmetric energy distribution $|S(f)|^2 = |S(-f)|^2$ implying that

$$\dot{\mathbf{s}}_l^H \mathbf{s}_l = (j2\pi) \sum_{k=0}^{K-1} f_k |S(f_k)|^2 = 0, \quad l = 1, \dots, L. \quad (12)$$

Furthermore, the array pattern $a_m(\phi, \theta)$ satisfies the following symmetry condition

$$\dot{\mathbf{a}}_{l,\phi}^H \dot{\mathbf{a}}_l = \dot{\mathbf{a}}_{l,\theta}^H \dot{\mathbf{a}}_l = 0, \quad l = 1, \dots, L.$$

The symmetric condition on the array pattern is generally satisfied for symmetric array. For instance, it is trivial to check that the condition is fulfilled if each antenna element has an isotropic pattern according to (16) later studied in Section V. The following corollary gives a particularization of the LB of Theorem 1 under additional assumptions (As2) – (As4) and for the MSE averaged over the receive antennas, *i.e.*,

$$\text{MSE}(f, \hat{\psi}) \triangleq \frac{1}{M} \sum_{m=1}^M \text{MSE}_m(f, \hat{\psi}).$$

Corollary 1. Under (As2) – (As4), the expression of the LB of Theorem 1 averaged over the receive antennas simplifies to

$$\begin{aligned} \text{LB}(f, \boldsymbol{\psi}) &\triangleq \frac{1}{M} \sum_{m=1}^M \text{LB}_m(f, \boldsymbol{\psi}) \\ &= \underbrace{\frac{1}{\text{SNR}} \frac{L}{M}}_{\text{SNR gain}} \left(\underbrace{2}_{\text{Loss factor}} + \underbrace{\frac{1}{2} \left(\frac{f}{\sigma_F} \right)^2}_{\text{Extrapolation penalty}} \right), \end{aligned} \quad (13)$$

where σ_F^2 is the mean squared bandwidth of the transmit signal

$$\sigma_F^2 \triangleq \frac{\|\dot{\mathbf{s}}\|^2}{(2\pi)^2 \|\mathbf{s}\|^2} = \frac{\sum_{k=0}^{K-1} f_k^2 |S(f_k)|^2}{\sum_{k=0}^{K-1} |S(f_k)|^2}.$$

Proof. The proof is given in Appendix VIII-C. □

By adding some assumptions, the LB proposed in Theorem 1 can be greatly simplified and Corollary 1 provides much insight into the physical meaning of the different terms of the LB. We can clearly identify the two main advantages of high-resolution channel estimation. As expected, a SNR gain of a factor $\frac{MK}{L}$ can be observed with respect to the LS estimation performance that we derived in (5). This gain comes from two contributions: the array gain M and the estimation of only L channel paths instead of K as in the LS case. However, a loss factor of 2 appears, coming from the penalty of estimating the real and imaginary gains, the azimuth and the elevations angles of each path. Secondly, the channel can be extrapolated in frequency at the cost of a MSE penalty that quadratically scales with the ratio f/σ_F where the denominator indicates that the extrapolation range can be quantified in multiples of the signal bandwidth.

It is interesting to see that the dependence of $\text{LB}(f, \boldsymbol{\psi})$ on the path parameters $\boldsymbol{\psi}$ vanishes. This is in part explained by the fact that each path is well separated, which cancels the interdependence between different paths. Additionally, the channel frequency response is evaluated in the direction of the incoming specular waves, canceling the dependence on the parameters of each path as well as the dependence on the array pattern.

Based on the simplified LB, we can find a closed-form expression of the extrapolation range. We define the γ extrapolation range as the frequency beyond which the extrapolation performance falls γ times below that of the conventional LS estimator given in (5). Inserting the expressions of (5) and (13) in the equation $\text{LB}(f, \boldsymbol{\psi}) = \gamma \text{MSE}_{\text{LS}}$, we easily find

$$f_{\text{Extrapol-}\gamma} = 2\sigma_F \sqrt{\frac{MK\gamma}{2L} - 1}. \quad (14)$$

Note that this definition is independent on the SNR while the parameters K , L and σ_F are related to the training bandwidth BW. Depending on the system requirements, the value of γ can be chosen differently and the actual useful extrapolation range can be larger or smaller.

D. Single-Input-Single-Output

The specialization of the above results to the SISO case is straightforward. As the angles of arrival are not resolved, the Fisher information matrix becomes a $3L \times 3L$ matrix. To simplify the LB, we introduce the following adaptation of (As3) – (As4) to the SISO case:

(As3'): separation of the L specular rays in delay. We assume that, for each pair of rays l, l' ($l \neq l'$), the condition (11) is verified.

(As4'): the training signal $S(f)$ has a symmetric energy distribution $|S(f)|^2 = |S(-f)|^2$ implying that (12) holds.

Corollary 2. *Under (As1), the LB on the channel extrapolation error for any unbiased estimator $\hat{H}(f, \hat{\psi})$ in the SISO case is*

$$\text{LB}_{\text{SISO}}(f, \psi) \triangleq \mathbf{g}_{f, \psi}^H \mathbf{I}_{\psi, \text{SISO}}^{-1} \mathbf{g}_{f, \psi},$$

where $\mathbf{g}_{f, \psi} = (\mathbf{g}_{f, \psi_1}^T, \dots, \mathbf{g}_{f, \psi_L}^T)^T$ and $\mathbf{g}_{f, \psi_l} = (-j2\pi f \alpha_l, 1, j)^T e^{-j2\pi f \tau_l}$. Under (As3') – (As4'), the LB simplifies to

$$\text{LB}_{\text{SISO}}(f, \psi) = \underbrace{\frac{L}{\text{SNR}}}_{\text{SNR gain}} \left(\underbrace{1}_{\text{Loss factor}} + \underbrace{\frac{1}{2} \left(\frac{f}{\sigma_F} \right)^2}_{\text{Extrapolation penalty}} \right)$$

end the γ extrapolation range becomes

$$f_{\text{Extrapol-}\gamma, \text{SISO}} = 2\sigma_F \sqrt{\frac{K\gamma}{L2} - \frac{1}{2}}. \quad (15)$$

Proof. The proof is omitted as it follows the same methodology as the proofs of Theorem 1 and Corollary 1. $\square$

As could be expected, the only SNR gain now comes from estimating L coefficients rather than K while the array gain has vanished. One can note that the loss factor is only one versus two in the SIMO case as the azimuth and elevation angles of each path are not estimated. The main difference of the SISO case with the SIMO case is the fact that far fewer observations of the channel are available, especially compared to a massive MIMO scenario with a large M . This not only eliminates the array gain but also makes $\mathbf{I}_{\psi, \text{SISO}}$ more ill-conditioned as the rays

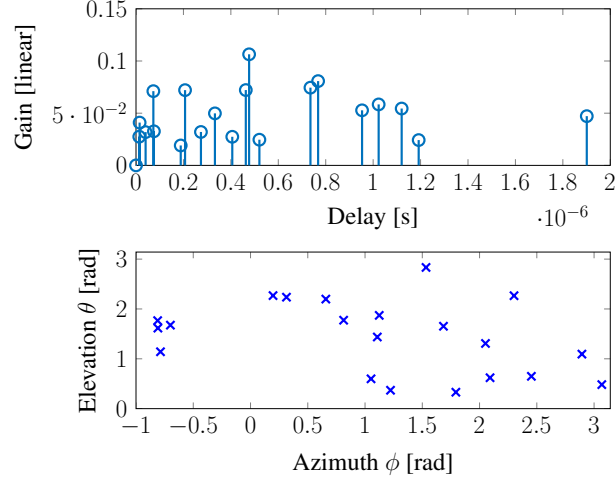


Fig. 3. Generated set of parameters $(\alpha_l, \tau_l, \phi_l, \theta_l)$ for $l = 1, \dots, L$ with $L = 21$ and following 3GPP 3D-UMa NLOS model. The sum of gains is normalized to one.

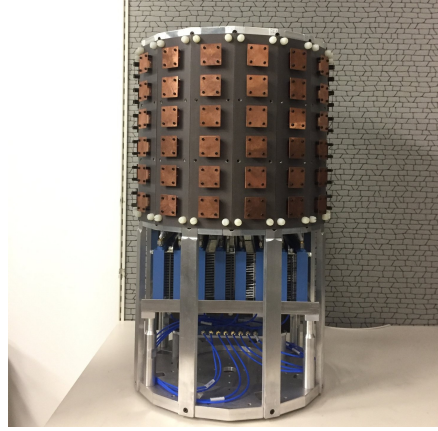


Fig. 4. Cylindrical array manufactured in our lab. Each ring/row of the array has 16 antenna elements. Top and bottom rings are composed of dummy/terminated elements while the four rings in the middle contain active elements, giving a total of $M = 4 \times 16 = 64$ elements.

can only be separated in the delay domain. As a result, (As3') is a stronger assumption than (As3) and might only be valid for a small number of rays L and/or a very large bandwidth. These factors tend to strongly limit the potential gains of high-resolution channel estimation in SISO systems [9].

V. NUMERICAL VALIDATION

This section aims at assessing the accuracy of the theoretical LB of the extrapolated channel through simulations. The training pulse shape $S(f)$ is chosen to have uniform energy distribution over the K frequency points f_k , which are uniformly spaced across the training bandwidth BW with spacing $1/\tau_{\max}$, *i.e.*, $f_k = (k - \frac{K-1}{2}) / \tau_{\max}$ for $k = 0, \dots, K-1$ and $K = \text{BW}\tau_{\max} + 1$.

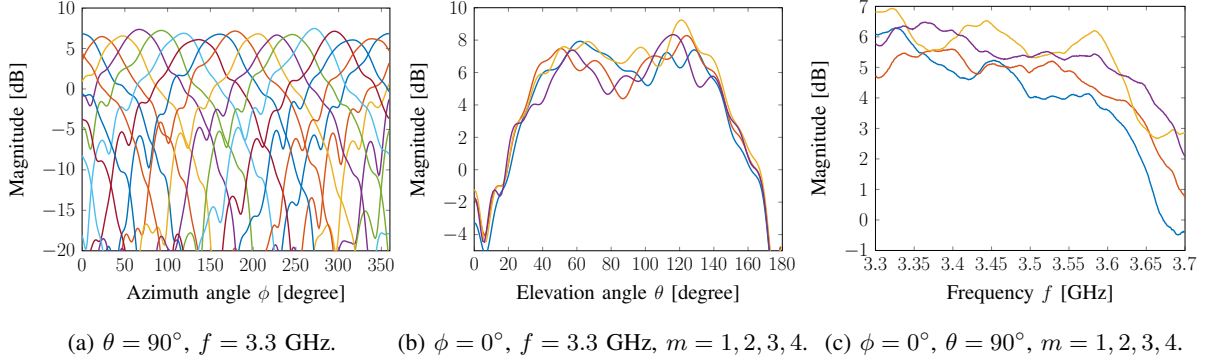


Fig. 5. Cylindrical array pattern $a_m(\phi, \theta, f)$. Normalization is arbitrary. (a) Azimuth cut of one ring/row of 16 elements. (b) Elevation cut of front column of 4 elements. (c) Frequency dependence of front column of 4 elements.

This assumption is consistent with, *e.g.*, the Zadoff-Chu training sequences in LTE and NR. We set $\tau_{\max} = 2.5\mu s$. In the following we consider a SNR of 25 dB. We recall that our definition of the SNR is the ratio of the total signal variance to the per-frequency tone noise variance, i.e., $\text{SNR} = E_T/\sigma_w^2$. This implies that the per-frequency tone SNR is K times smaller.

The performance in the figures is shown as a function of frequency normalized to the training bandwidth BW, as we expect from Corollary 1 that the extrapolation range scales accordingly. In the legend of the figures, the full LB refers to the LB of Theorem 1 averaged over the receive antennas and the simplified LB refers to the expression of Corollary 1. In the SISO case, the full and simplified LB curves refer to the corresponding expressions in Corollary 2. The conventional LS estimation performance derived (5) will be considered as a benchmark.

The MPC parameters were generated by the QuaDRiGa toolbox [25] according to the 3D-UMa NLOS model defined by 3GPP TR 36.873 v12.5.0 specifications [26]. The same set of parameters was used for all simulations and is depicted in Fig. 3. One can see that some rays are very closely spaced in delay and angle. We took on purpose a non line of sight scenario to consider a more challenging case as all paths need to be resolved to properly model the channel instead of only a few in a line of sight case.

The center frequency f_c is set to 3.5 GHz. We consider two types of array at receive side:

- A synthetic rectangular planar array with an inter-antenna element spacing of $\lambda_c/2$ where λ_c is the center wavelength. The antenna elements have an isotropic pattern so that the

pattern of each element becomes only a phase shift

$$a_m(\phi, \theta, f) = e^{-j2\pi \frac{f_c + f}{c} \mathbf{r}_m \cdot \hat{\mathbf{e}}(\phi, \theta)}, \quad (16)$$

where $\hat{\mathbf{e}}(\phi, \theta)$ is a unit vector in $\mathbb{R}^3$ pointing in the direction of the incoming ray l and the position of the m -th receive array element is denoted by $\mathbf{r}_m \in \mathbb{R}^3$ with respect to an reference point. The reference point is chosen to ensure that $\sum_m \mathbf{r}_m = \mathbf{0}$. Three rectangular array geometries are considered: $M = 8$ (4 Horiz. $\times$ 2 Vert.), $M = 16$ (4 Horiz. $\times$ 4 Vert.) and $M = 64$ (8 Horiz. $\times$ 8 Vert.).

- A cylindrical array manufactured in our lab shown in Fig. 4. Each ring/row of the array has 16 antenna elements. Top and bottom rings are composed of dummy/terminated elements while the four rings in the middle contain active elements, giving a total of $M = 4 \times 16 = 64$ active elements. Each active element has one vertically and one horizontally polarized port with high cross polarization discrimination ratio. We only used the vertically polarized ports in this study. In the following, we will consider different subsets of the total array: $M = 8$ contains only one ring with one out of two elements, $M = 16$ contains one ring of elements and $M = 32$ contains two rings of elements. The pattern of each array element $a_m(\phi, \theta, f)$ was obtained by careful calibration in our anechoic chamber over the band $f \in [3.3, 3.7]$ GHz, for an azimuth range $\phi \in [0, 2\pi]$ and elevation range $\theta \in [0, \pi]$.³ Azimuth, elevation and frequency cuts of the array are shown in Fig. 5. Note that the pattern shows frequency variations of about 6 dB on the 400 MHz bandwidth.

A. SAGE Performance versus Theoretical LB

For high resolution parameter extraction, we extend the SAGE algorithm introduced in [20] to extract elevation angles. The algorithm works in the frequency domain taking into account the frequency dependence of the pattern according to (1). The performance of the algorithm is averaged over 100 noise realizations. A delay step size of $\frac{1}{50\text{BW}}$ Hz and an angular grid size of 1 degree are used as parameters of the SAGE grid search. For the parameters $M = 16$, SNR = 25 dB and BW = 20 MHz, Fig. 6 shows that the SAGE-based parameter extraction and channel extrapolation based on (7) can approach the performance of the theoretical LB. This

³During calibration, we sampled $a_m(\phi, \theta, f)$ with a 1 MHz frequency resolution and 5 degrees angular resolution. We obtain the full pattern $a_m(\phi, \theta, f)$ by linear interpolation in frequency and through the effective aperture distribution function (EADF) in angle [27]. The EADF also allows simple evaluation of the angular derivatives of the pattern required for the LB computation.

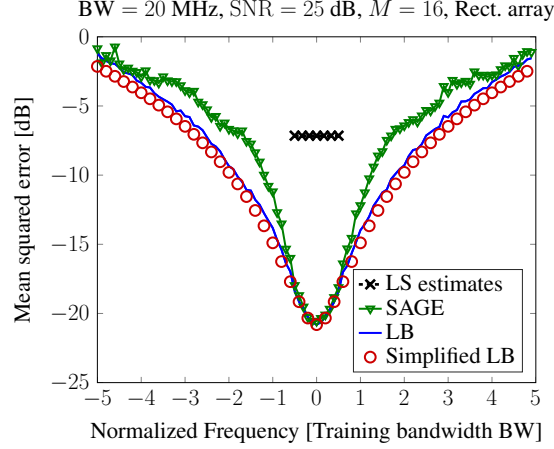


Fig. 6. The SAGE algorithm can approach the performance of the full LB (Theorem 1) and the simplified LB (Corollary 1).

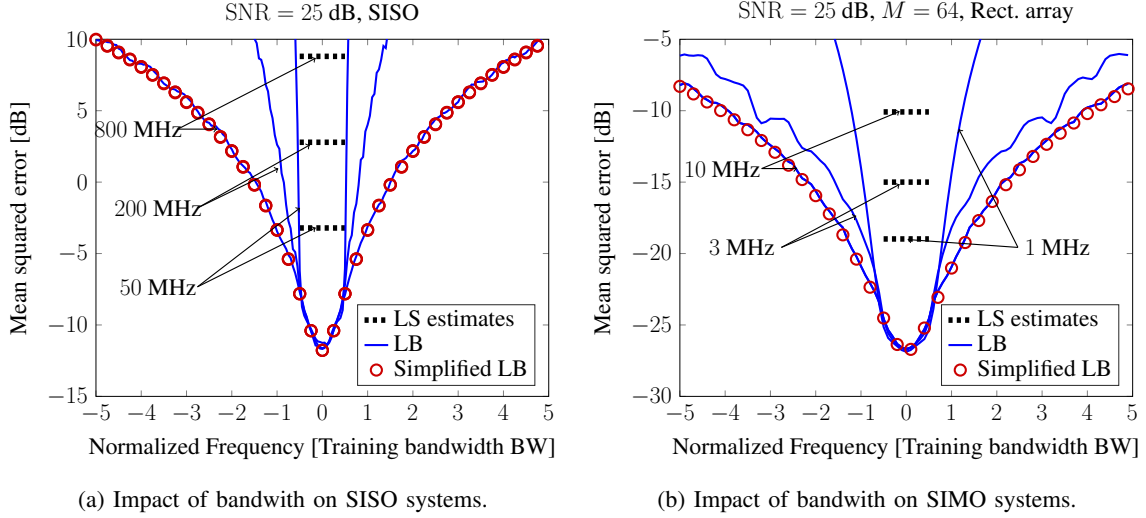


Fig. 7. Impact of the bandwidth on the extrapolation performance. As bandwidth increases, the full LB converges to the simplified LB meaning that (As3) holds.

implies that the LB gives a good indicator of the achievable MSE. Furthermore, we can see in the figure that the LB performs very close to the simplified one, implying that the paths are well separated. As expected, high resolution parameter extraction provides a large SNR gain of a factor $\frac{MK}{L} \approx 46 \approx 16$ dB with respect to the LS estimator. Moreover, the LS performance is very poor, as expected according to (5). Note that our definition of the SNR is not the per-frequency bin SNR, which is $K = 51 \approx 17$ dB times smaller.

B. Influence of the Bandwidth

Fig. 7 (a) and (b) show the impact of the bandwidth in the SISO and the $M = 64$ rectangular array cases respectively. As the bandwidth increases, the receiver has a larger resolution in time and it can better resolve the different paths in the delay domain. As the bandwidth increases, (As3') becomes more valid and the full LB converges to the simplified LB. The gap between the full and simplified LB can be seen as an indicator of the separability of the path parameters. The big difference of the SIMO case with the SISO case is that the extrapolation becomes possible for a much smaller training bandwidth: 10 MHz versus 800 MHz.⁴ This is explained by the fact that the paths can be separated in the delay-angle domain instead of just the delay domain. Moreover the SIMO system achieves a SNR gain of a factor $M/2 = 64/2 \approx 15$ dB.⁵

In the light of these limitations, we can conclude that the price to pay for channel extrapolation in SISO is to have a very large bandwidth at disposal and/or a channel that exhibits few well separated paths in delay. These observations tends to strongly limit the applicability of extrapolation for conventional SISO communication systems.

As opposed to high-resolution channel estimation, increasing the bandwidth with a fixed SNR is detrimental to conventional LS estimation as the number of frequency bins K to estimate becomes larger, as shown in (5). Another way to view this is that the energy is more spread out in frequency and leads to a lower per-frequency bin SNR.

C. Influence of the Number of Antennas and the Array Type

Fig. 8 (a) depicts the extrapolation performance given a fixed training bandwidth of 20 MHz, with different number of antenna elements. Both the rectangular and cylindrical arrays are considered. The same effect previously described in terms of bandwidth occurs in terms of antenna numbers, *i.e.*, as the number of antennas increases, the resolution in the angle domain increases and the full LB converges to the simplified LB. These observations imply that the separability of the rays can be achieved by trading bandwidth against number of antennas. The synthetic rectangular and the real cylindrical arrays seem to achieve similar performance. In the SISO case, no extrapolation is possible given that the paths cannot be well separated using only

⁴Note that a number of our system model assumptions starts to be violated for a bandwidth of 800 MHz at 3.5 GHz carrier frequency. Such limitations are not taken into account in the above figure for ease of interpretation.

⁵The factor 1/2 stands for the loss factor of estimating the azimuth and elevation parameters of each path in the SIMO case.

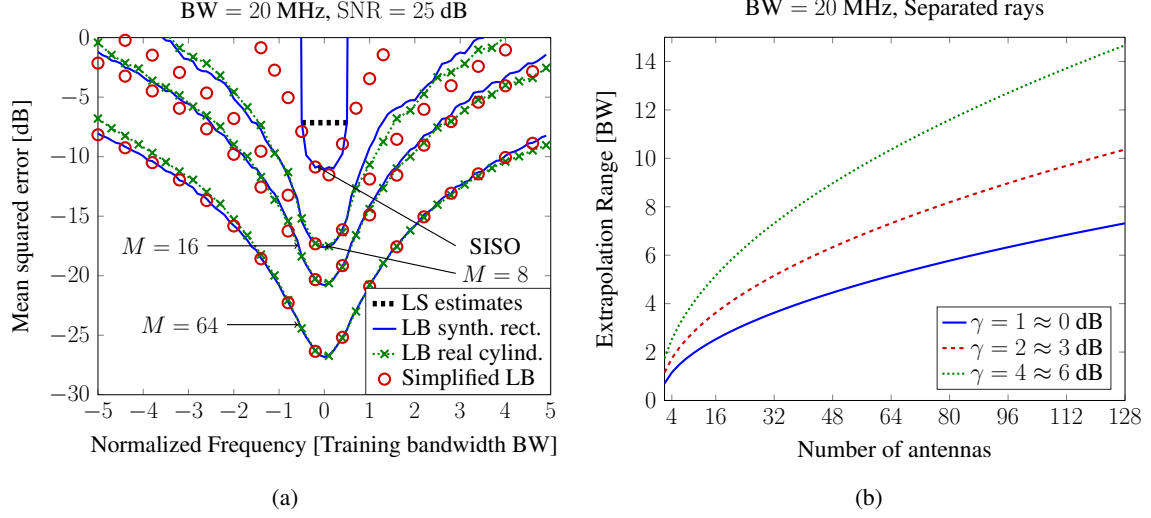


Fig. 8. (a) As number of antennas increases, the LB reaches the simplified LB while achieving an additional SNR gain. Synthetic rectangular and real cylindrical arrays exhibit very similar performance. (b) Extrapolation range $f_{\text{Extrapol}-\gamma}$, normalized in the training bandwidth BW, as a function of the number of antennas.

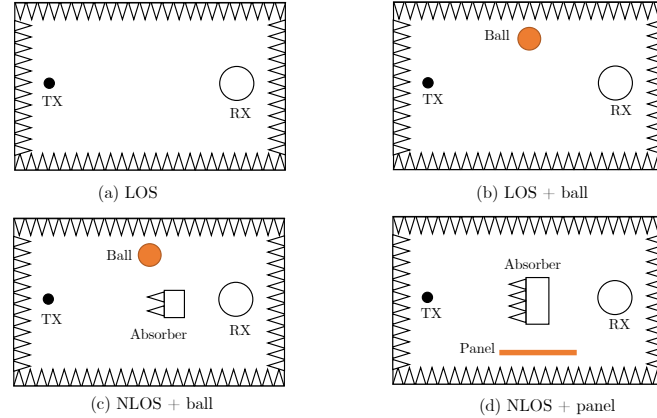


Fig. 9. Four scenarios investigated in the anechoic chamber for the experimental validation. Transmit antenna and receive array are spaced 5 meters apart. (a) is only line-of-sight (LOS). (b) is LOS + an aluminium ball. (c) is non-line-of-sight (NLOS) + aluminium ball. (d) is NLOS + aluminium panel.

the delay domain. The simplified LB of Corollary 1 is very close to the full LB of Theorem 1 as soon as the array has at least 8 antennas, even though it does not depend on the array pattern and the path parameters. In Fig. 8 (b), the extrapolation range $f_{\text{Extrapol}-\gamma}$ given in (14) is plotted as a function of the number of antennas. As a reminder, the formula assumes that the paths are well separated and is properly defined just before (14).

VI. EXPERIMENTAL VALIDATION

This section presents the validation of the extrapolation performance through experimental measurements for a series of scenarios. There were multiple motivations for doing this validation in an anechoic chamber. Firstly, using a vector network analyzer (VNA), we can obtain almost noise free measurements of the channel frequency response. In practice, we obtain the ground truth $H_m(f)$ on a 400 MHz bandwidth ranging from 3.3 GHz to 3.7 GHz with a 1 MHz frequency spacing and for all elements $m = 1, \dots, M$. We estimated the per-frequency tone SNR to be on the order of 50 dB within the chamber, which we will refer to as almost noise-free in the following.⁶ Secondly, we can easily investigate the impact of additive noise by simply adding it by post-processing the raw measurements according to a certain SNR figure. Furthermore, we have the freedom to design measurement scenarios using point reflectors or planar reflectors (aluminium balls or panels), at known locations, from which the path parameters can be inferred for verification of the results. More specifically, we consider the four scenarios of Fig. 9.

The downside of using the anechoic chamber is mostly related to its dimension being relatively small, with 5 meters between the transmitter and the receiver. This does not allow to model typical propagation delay spreads. Furthermore, the far field assumption might not hold as ball/panel reflections and diffraction around the absorbers take place very close to the receiver. Moreover, the reflector surface was not completely smooth. This implies that the receiver might see the effect of diffuse multipath components rather than specular components as assumed in our system model in (1).

In the experiments, we used an omni-directional antenna at the transmit side with vertical polarization and a high cross polarization discrimination ratio ($> 20\text{dB}$). At the receive side, we used the cylindrical array that was described in Section V. For extracting the path parameters, we used the SAGE algorithm using the same configuration as described in Section V-A. The extrapolation is then simply performed by plugging the parameters estimates in (7).

In the following experiments, we define the VNA measurements $H_m(f)$ on the 400 MHz band as the ground truth. We normalize $H_m(f)$ to have unit variance across antennas and frequency points. To extract the path parameters, we use a training bandwidth of $\text{BW} = 40$ MHz, ranging from 3.3 GHz to 3.34 GHz, which corresponds to the $K = 41$ lower frequency tones (frequency spacing of 1 MHz) and a center frequency $f_c = 3.32$ GHz. We validate the

⁶Practical systems are usually working in a much lower SNR regime.

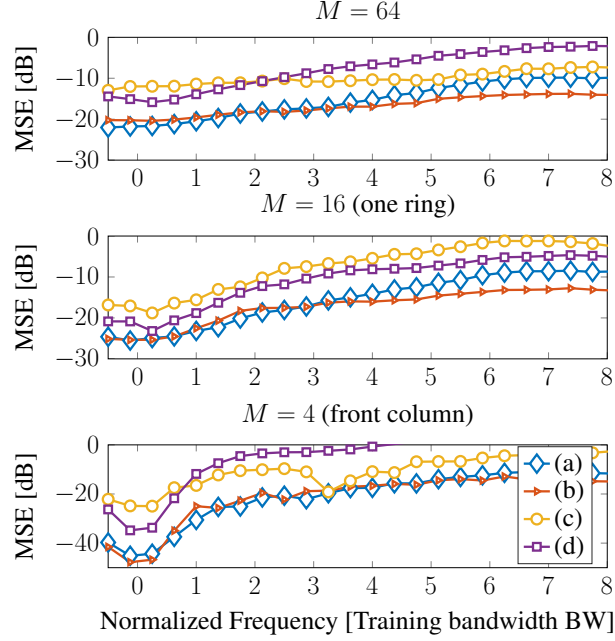


Fig. 10. Channel extrapolation performance for the four scenarios of Fig. 9. Path parameters are extracted running SAGE ($L = 5$) on the raw VNA measurements and a training bandwidth $BW = 40$ MHz. Different number of antennas of the array are considered.

extrapolation performance using the ground truth on the remaining 360 MHz, which corresponds to an extrapolation range of 9 times the training bandwidth BW .

A. Almost Noise-Free Performance

Using the raw VNA measurements, we run SAGE to extract parameters on a 40 MHz training band and evaluate the extrapolation performance for the four scenarios of Fig. 9. The result is shown in Fig. 10. As studied in Section. IV, the channel extrapolation performance is theoretically only limited by the additive noise power and the separability of the multipath components. Since the VNA measurements are almost noise free, one could expect a very high extrapolation performance, especially for a large number of antennas. Unfortunately, in practice, additive noise and path separability are not the only limiting factors as channel modeling and calibration errors are also present, as discussed above.

In the LOS scenarios (a) and (b), we can see that the extrapolation performance always increases as the number of antennas M decreases. We explain it by the fact that very few reflections are present. Even in scenario (b), the reflected path coefficient has a power of about 25dB lower than the main LOS path. Hence, it is not necessary to have a high resolution in

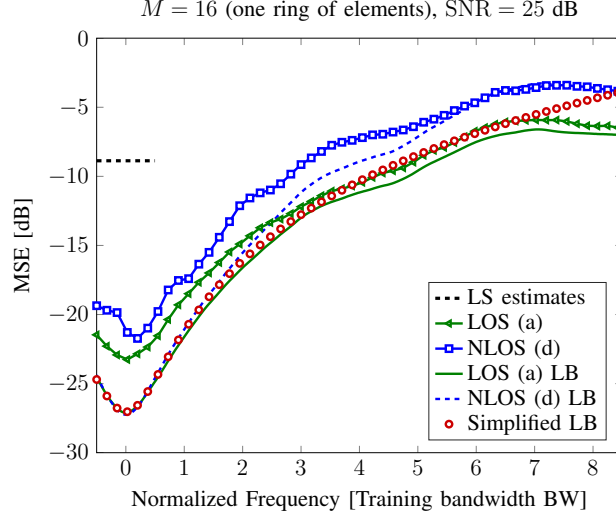


Fig. 11. Channel extrapolation performance based on SAGE extracted parameters for scenarios LOS (a) and NLOS (d) of Fig. 9 and a training bandwidth $BW = 40$ MHz. The VNA measurements are impacted by synthetic additive noise.

space to separate multipath components. Using a lower number of antennas actually reduces the effect of calibration errors as the system model in (1) is less constrained.

In the NLOS scenarios (c) and (d), the extrapolation performance as a function of the number antennas shows a different behavior. We should here emphasize that, in (c) and (d), the reflecting ball and panels were at the same height as the transmit and receive antennas. This implies that elevation angles of arrival are relatively similar and close to $\theta = 90^\circ$ at the receiver. On the other hand, the reflection on the ball, the reflection on the panel and the diffraction around the absorber clearly have different azimuth angles of arrival. These remarks first explain why the performance in the case $M = 16$ is better than in the case $M = 64$. Indeed, using only one ring of antennas of the array is enough to separate the incoming waves in azimuth. Adding three other rings of antennas ($M = 64$) gives more resolution in elevation but is not required in our scenarios and only leads to more calibration and/or modeling inaccuracy. Secondly, these remarks are in accordance with the case $M = 4$ using only the front column of antennas of the array. Using a column of the array gives high resolution in elevation and very poor resolution in azimuth. This explains why the extrapolation performance degrades very quickly as the extrapolation range increases as compared to the $M = 16$ case.

B. Performance under Additive Noise

In practice, we expect much more additive noise to be present so that the system would be more likely to be in a noise limited regime. To study such scenario, we synthetically add noise

on the VNA measurements based on the model in (1) with $S(f_k) = 1$ and $\text{SNR} = 25$ dB. The result, averaged over 100 noise realizations, is shown in Fig. 11 for scenarios LOS (a) and NLOS (d). The channel extrapolation performance relying on SAGE extracted parameters is compared to the LS estimator performance derived in (5) and the theoretical simplified LB of Corollary 1. The full LB of Theorem 1 averaged over the antennas is also plotted. Note that the full LB evaluation requires the ground truth path parameters ψ , which can only be approximated for real experiments. To avoid this and still get an approximation of the full LB, we assumed that the ground truth ψ can be well approximated by the SAGE extracted parameters $\hat{\psi}$ on the raw measurements without added noise. We can see in the figures that the SAGE performance in both LOS and NLOS cases approach the theoretical LB. In the training band, the gap is larger due to potential calibration and/or channel modeling errors while the gap decreases as the extrapolation range increases, meaning that the performance enters a noise-limited regime.

VII. CONCLUSIONS

This paper investigated the frequency channel extrapolation performance for FDD massive MIMO systems. We demonstrated the gain of applying high-resolution channel estimation as compared to conventional low-resolution estimators. Theoretical LBs for the MSE of the extrapolated channel were derived and validated through numerical simulations and experimental measurements. In particular, we showed that a simplified LB, obtained as a special case when paths are well separated, is very useful. It gives simplified yet accurate insight on the massive MIMO extrapolation performance without requiring the knowledge of the antenna patterns and the path parameters.

In conclusion, we demonstrated that, under a good calibration of the BS and favorable propagation conditions, channel extrapolation is a viable solution to deploy FDD massive MIMO systems and completely removing the DL training overhead. Our future studies will include performance assessment of extensive outdoor measurements. In particular, the impact of calibration errors and channel modeling errors such as, *e.g.*, diffuse multipath components, will require further investigation.

VIII. APPENDIX

A. Frequency Autocorrelation of the Channel

Using the definitions of $C_{H,m}(\Delta f)$ in (6), of $H_m(f)$ in (3) and the fact that the path parameters are assumed i.i.d, we can write

$$\begin{aligned} C_{H,m}(\Delta f) &= L\mathbb{E}(|\alpha_l|^2 |a_m(\phi_l, \theta_l)|^2 e^{-j2\pi\Delta f\tau_l}) \\ &= LC\mathbb{E}(|\alpha_l|^2 e^{-j2\pi\Delta f\tau_l}), \end{aligned}$$

where we additionally used the fact that the array pattern was assumed frequency independent and isotropic, *i.e.*, $|a_m(\phi, \theta)| = C$ where C a normalization constant. We now use the fact that τ_l is uniformly distributed in $[0, \tau_{\max}]$ and that the conditional variable $\alpha_l; \tau_l$ has variance $p(\tau_l) = e^{-\tau_l/\tau_{\text{rms}}}$

$$\begin{aligned} C_{H,m}(\Delta f) &= \frac{LC}{\tau_{\max}} \int_0^{\tau_{\max}} p(\tau_l) e^{-j2\pi\Delta f\tau_l} d\tau_l \\ &= \frac{LC\tau_{\text{rms}}}{\tau_{\max}} \frac{1 - e^{-\tau_{\max}/\tau_{\text{rms}} - j2\pi\Delta f\tau_{\max}}}{1 + j2\pi\Delta f\tau_{\text{rms}}}. \end{aligned}$$

We set $C = \frac{\tau_{\max}}{L\tau_{\text{rms}}} \frac{1}{1 - e^{-\tau_{\max}/\tau_{\text{rms}}}}$ to fix $\mathbb{E}(|H_m(f)|^2) = 1$.

B. Proof of Theorem 1

The extrapolated channel frequency response $H_m(f, \psi)$ is a non linear function of the path parameters ψ . A straightforward application of the CRLB for transformation of parameters [24] gives the following bound for the extrapolation error $\text{MSE}_m(f, \hat{\psi})$

$$\text{LB}_m(f, \psi) = \left(\frac{\partial H_m(f, \psi)}{\partial \psi} \right)^H \mathbf{I}_{\psi}^{-1} \frac{\partial H_m(f, \psi)}{\partial \psi}.$$

where

$$\left[\frac{\partial H_m(f, \psi)}{\partial \psi} \right]_u = \frac{dH_m(f, \psi)}{d\psi_u}.$$

Using the vector definitions introduced in Theorem 1, we can write

$$\mathbf{g}_{m,f,\psi} \triangleq \frac{\partial H_m(f, \psi)}{\partial \psi},$$

which concludes the proof.

C. Proof of Corollary 1

Using (9), we can compute the different elements of the full Fisher information matrix given in (10). In the following, we use the notations $\|\mathbf{s}_l\|^2 = \|\mathbf{s}\|^2$ and $\|\dot{\mathbf{s}}_l\|^2 = \|\dot{\mathbf{s}}\|^2$ given that the dependence in the path index vanishes.

First, using (As3), we can show that the off-diagonal blocks of $\mathbf{I}_\psi$ vanish, *i.e.*, $\mathbf{I}_{\psi_l, \psi_{l'}} = \mathbf{0}$ for $l \neq l'$. Indeed, for the diagonal elements of $\mathbf{I}_{\psi_l, \psi_{l'}}$, we find that

$$\begin{aligned} I_{\tau_l \tau_{l'}} &= \Re(\alpha_l^* \alpha_{l'} \mathbf{a}_l^H \mathbf{a}_{l'} \dot{\mathbf{s}}_l^H \dot{\mathbf{s}}_{l'}) = |\alpha_l|^2 \|\mathbf{a}_l\|^2 \|\dot{\mathbf{s}}\|^2 \delta_{l-l'} \\ I_{\phi_l \phi_{l'}} &= \Re(\alpha_l^* \alpha_{l'} \dot{\mathbf{a}}_{l, \phi}^H \dot{\mathbf{a}}_{l', \phi} \mathbf{s}_l^H \mathbf{s}_{l'}) = |\alpha_l|^2 \|\dot{\mathbf{a}}_{l, \phi}\|^2 \|\mathbf{s}\|^2 \delta_{l-l'} \\ I_{\theta_l \theta_{l'}} &= \Re(\alpha_l^* \alpha_{l'} \dot{\mathbf{a}}_{l, \theta}^H \dot{\mathbf{a}}_{l', \theta} \mathbf{s}_l^H \mathbf{s}_{l'}) = |\alpha_l|^2 \|\dot{\mathbf{a}}_{l, \theta}\|^2 \|\mathbf{s}\|^2 \delta_{l-l'} \\ I_{\alpha_l^R \alpha_{l'}^R} &= I_{\alpha_l^I \alpha_{l'}^I} = \Re(\mathbf{a}_l^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}) = \|\mathbf{a}_l\|^2 \|\mathbf{s}\|^2 \delta_{l-l'}. \end{aligned}$$

Still using (As3), we find the same results for the off-diagonal elements of $\mathbf{I}_{\psi_l, \psi_{l'}}$, $l \neq l'$. Actually, using (As4), we find that the result also holds when $l = l'$ for the following elements

$$\begin{aligned} I_{\tau_l \phi_{l'}} &= -\Re(\alpha_l^* \alpha_{l'} \mathbf{a}_l^H \dot{\mathbf{a}}_{l', \phi} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\tau_l \theta_{l'}} &= -\Re(\alpha_l^* \alpha_{l'} \mathbf{a}_l^H \dot{\mathbf{a}}_{l', \theta} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\tau_l \alpha_{l'}^R} &= -\Re(\alpha_l^* \mathbf{a}_l^H \mathbf{a}_{l'} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\tau_l \alpha_{l'}^I} &= \Im(\alpha_l^* \mathbf{a}_l^H \mathbf{a}_{l'} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\phi_l \alpha_{l'}^R} &= \Re(\alpha_l^* \dot{\mathbf{a}}_{l, \phi}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\phi_l \alpha_{l'}^I} &= -\Im(\alpha_l^* \dot{\mathbf{a}}_{l, \phi}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\theta_l \alpha_{l'}^R} &= \Re(\alpha_l^* \dot{\mathbf{a}}_{l, \theta}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\theta_l \alpha_{l'}^I} &= -\Im(\alpha_l^* \dot{\mathbf{a}}_{l, \theta}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}) = 0 \\ I_{\alpha_l^R \alpha_{l'}^I} &= -\Im(\mathbf{a}_l^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}) = 0. \end{aligned}$$

One can further check that, under (As3), the elements $I_{\phi_l \theta_{l'}}$ vanish for $l \neq l'$. However, even under (As4), $I_{\phi_l \theta_{l'}}$ does not vanish for $l = l'$, *i.e.*,

$$I_{\phi_l \theta_{l'}} = \Re(\alpha_l^* \alpha_{l'} \dot{\mathbf{a}}_{l, \phi}^H \dot{\mathbf{a}}_{l, \theta} \mathbf{s}_l^H \mathbf{s}_{l'}) = |\alpha_l|^2 \|\mathbf{s}\|^2 \Re(\dot{\mathbf{a}}_{l, \phi}^H \dot{\mathbf{a}}_{l, \theta}) \delta_{l-l'}.$$

Taking into account the above simplifications, the full Fisher matrix $\mathbf{I}_\psi$ becomes block diagonal and each block on its diagonal is itself block diagonal

$$\mathbf{I}_\psi = \frac{2}{\sigma_w^2} \begin{pmatrix} \mathbf{I}_{\psi_1, \psi_1} & & \\ & \ddots & \\ & & \mathbf{I}_{\psi_L, \psi_L} \end{pmatrix}, \quad \mathbf{I}_{\psi_l, \psi_l} = \begin{pmatrix} I_{\tau_l \tau_l} & & & \\ & I_{\phi_l \phi_l} & I_{\phi_l \theta_l} & \\ & I_{\phi_l \theta_l} & I_{\theta_l \theta_l} & \\ & & & I_{\alpha_l^R \alpha_l^R} \\ & & & & I_{\alpha_l^I \alpha_l^I} \end{pmatrix}.$$

Using the fact the inverse of a block diagonal matrix is a block diagonal matrix with the inverse of the original blocks on its diagonal, the LB of Theorem 1 averaged over the receive antennas becomes

$$\begin{aligned} \frac{1}{M} \sum_{m=1}^M \widehat{\text{LB}}_m(f, \psi) &= \frac{1}{M} \sum_{m=1}^M \mathbf{g}_{m,f,\psi}^H \mathbf{I}_\psi^{-1} \mathbf{g}_{m,f,\psi} \\ &= \frac{\sigma_w^2}{2M} \sum_{m=1}^M \sum_{l=1}^L \mathbf{g}_{m,f,\psi_l}^H \mathbf{I}_{\psi_l, \psi_l}^{-1} \mathbf{g}_{m,f,\psi_l} \\ &= \frac{\sigma_w^2}{2M} \sum_{m=1}^M \sum_{l=1}^L \left[\frac{|g_{m,f,\tau_l}|^2}{I_{\tau_l \tau_l}} + \frac{|g_{m,f,\alpha_l^R}|^2}{I_{\alpha_l^R \alpha_l^R}} + \frac{|g_{m,f,\alpha_l^I}|^2}{I_{\alpha_l^I \alpha_l^I}} \right. \\ &\quad \left. + \begin{pmatrix} g_{m,f,\phi_l}^* & g_{m,f,\theta_l}^* \end{pmatrix} \begin{pmatrix} I_{\phi_l \phi_l} & I_{\phi_l \theta_l} \\ I_{\phi_l \theta_l} & I_{\theta_l \theta_l} \end{pmatrix}^{-1} \begin{pmatrix} g_{m,f,\phi_l} \\ g_{m,f,\theta_l} \end{pmatrix} \right]. \end{aligned} \quad (17)$$

After some computations, we find that

$$\begin{aligned} \sum_{m=1}^M \sum_{l=1}^L \frac{|g_{m,f,\alpha_l^R}|^2}{I_{\alpha_l^R \alpha_l^R}} + \frac{|g_{m,f,\alpha_l^I}|^2}{I_{\alpha_l^I \alpha_l^I}} &= \frac{2L}{\|\mathbf{s}\|^2} \\ \sum_{m,l} \begin{pmatrix} g_{m,f,\phi_l}^* & g_{m,f,\theta_l}^* \end{pmatrix} \begin{pmatrix} I_{\phi_l \phi_l} & I_{\phi_l \theta_l} \\ I_{\phi_l \theta_l} & I_{\theta_l \theta_l} \end{pmatrix}^{-1} \begin{pmatrix} g_{m,f,\phi_l} \\ g_{m,f,\theta_l} \end{pmatrix} &= \frac{2L}{\|\mathbf{s}\|^2} \\ \sum_{m=1}^M \sum_{l=1}^L \frac{|g_{m,f,\tau_l}|^2}{I_{\tau_l \tau_l}} &= \frac{L(2\pi f)^2}{\|\dot{\mathbf{s}}\|^2} = \frac{L}{\|\mathbf{s}\|^2} \frac{f^2}{\sigma_F^2}, \end{aligned}$$

where $\sigma_F^2 = \frac{\|\dot{\mathbf{s}}\|^2}{(2\pi)^2 \|\mathbf{s}\|^2}$. Inserting the result of these last equations into (17) and using the definition $\text{SNR} \triangleq \frac{\|\mathbf{s}\|^2}{\sigma_w^2}$, we find the result of Corollary 1.

REFERENCES

- [1] F. Rottenberg, R. Wang, J. Zhang, and A. F. Molisch, "Channel Extrapolation in FDD Massive MIMO: Theoretical Analysis and Numerical Validation," *arXiv preprint arXiv:1902.06844*, 2019.

- [2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.
- [3] E. Björnson, J. Hoydis, L. Sanguinetti *et al.*, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [4] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing - The Large-Scale Array Regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, Oct 2013.
- [5] M. Barzegar Khalilsarai, S. Haghighatshoar, X. Yi, and G. Caire, "FDD Massive MIMO via UL/DL Channel Covariance Extrapolation and Active Channel Sparsification," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 121–135, Jan 2019.
- [6] "IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput," *IEEE Std 802.11n-2009*, pp. 1–565, Oct 2009.
- [7] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable Rates of FDD Massive MIMO Systems With Spatial Channel Correlation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2868–2882, May 2015.
- [8] X. Rao and V. K. N. Lau, "Distributed Compressive CSIT Estimation and Feedback for FDD Multi-User Massive MIMO Systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, June 2014.
- [9] M. Pun, A. F. Molisch, P. Orlik, and A. Okazaki, "Super-Resolution Blind Channel Modeling," in *2011 IEEE International Conference on Communications (ICC)*, June 2011, pp. 1–5.
- [10] N. Jalden, H. Asplund, and J. Medbo, "Channel extrapolation based on wideband MIMO measurements," in *2012 6th European Conference on Antennas and Propagation (EUCAP)*, March 2012, pp. 442–446.
- [11] D. Vasisht, S. Kumar, H. Rahul, and D. Katabi, "Eliminating Channel Feedback in Next-Generation Cellular Networks," in *Proceedings of the 2016 ACM SIGCOMM Conference*, ser. SIGCOMM '16. New York, NY, USA: ACM, 2016, pp. 398–411.
- [12] W. Yang, L. Chen, and Y. E. Liu, "Super-Resolution for Achieving Frequency Division Duplex (FDD) Channel Reciprocity," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 1–5.
- [13] X. Zhang, L. Zhong, and A. Sabharwal, "Directional Training for FDD Massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5183–5197, Aug 2018.
- [14] M. Haardt, M. D. Zoltowski, C. P. Mathews, and J. Nosssek, "2D unitary ESPRIT for efficient 2D parameter estimation," in *1995 International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, May 1995, pp. 2096–2099 vol.3.
- [15] M. Arnold, S. Dörner, S. Cammerer, S. Yan, J. Hoydis, and S. t. Brink, "Enabling FDD Massive MIMO through Deep Learning-based Channel Prediction," *arXiv preprint arXiv:1901.03664*, 2019.
- [16] U. Ugurlu, R. Wichman, C. B. Ribeiro, and C. Wijting, "A Multipath Extraction-Based CSI Acquisition Method for FDD Cellular Networks With Massive Antenna Arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2940–2953, April 2016.
- [17] A. Duel-Hallen, "Fading Channel Prediction for Mobile Radio Adaptive Transmission Systems," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec 2007.
- [18] T. Svantesson and A. L. Swindlehurst, "A performance bound for prediction of MIMO channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 520–529, Feb 2006.
- [19] M. D. Larsen, A. L. Swindlehurst, and T. Svantesson, "Performance Bounds for MIMO-OFDM Channel Estimation," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1901–1916, May 2009.
- [20] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, March 1999.

- [21] F. Rottenberg, F. Horlin, E. Kofidis, and J. Louveaux, "Generalized optimal pilot allocation for channel estimation in multicarrier systems," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2016, pp. 1–5.
- [22] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, July 1998.
- [23] A. F. Molisch, "Ultra-Wide-Band Propagation Channels," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 353–371, Feb 2009.
- [24] S. M. Kay, "Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1)," *PTR Prentice-Hall, Englewood Cliffs*, 1993.
- [25] 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, June 2014.
- [26] "3GPP TR 36.873 v12.5.0," Tech. Rep., 2017.
- [27] A. Richter, "Estimation of radio channel parameters: Models and algorithms." ISLE, 2005.