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. This drastically reduces the downlink pilot overhead and completely removes the need for a feedback from the users. The price to pay is a degradation in the quality of the channel estimates, which reduces the downlink spectral efficiency. We first show that conventional estimators fail to achieve reasonable accuracy. We propose instead to use high-resolution channel estimation. We derive the Cramer-Rao lower bound (CRLB) of the mean squared error (MSE) of the extrapolated channel. Furthermore, a relationship between the imperfect channel state information (CSI) and the downlink user performance is derived. The extrapolation-based FDD massive MIMO performance is validated through numerical simulations and compared to a corresponding time division duplex (TDD) system. Considered figures of merit for extrapolation performance include channel MSE, beamforming efficiency, extrapolation range, spectral efficiency and uncoded symbol error rate. Our main conclusion is that channel extrapolation is a viable solution for FDD massive MIMO systems.

I. INTRODUCTION

The deployment of massive multiple-input-multiple-output (MIMO) communications systems strongly relies on the acquisition of accurate channel state information (CSI) at the 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 (see Fig. 1), 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], compression of the feedback based on deep learning [8], 3-D beamforming based on channel statistics [9], or compressed sensing methods [10], each of which involves some feedback from the users. 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 least squares (LS) and linear minimum mean squared (LMMSE) is very limited - typically to the order of one coherence bandwidth, as will be further shown in this work. To overcome this limit, [11] suggested estimation of the multipath components (MPCs) via high-resolution parameter estimation (HRPE). Based on the structure of the channel and the extracted MPCs, extrapolation over wide frequency range can be achieved. However, the paper only considered the single-input-single-output (SISO) case, which resulted in a poor extrapolation performance. Ref. [12] extends the setup to the MIMO case and the extrapolation to the spatial domain. Multiple measurements show that a frequency extrapolation range larger than 5 times the coherence bandwidth can be reached. Ref. [13] presents the so-called R2-F2 system to extract path parameters and extrapolate the channel in frequency. The paper shows how to integrate the system into LTE cellular networks and uses experimental measurements for validation. The study restricted the frequency spacing between UL and DL band to be only 20-30 MHz and did not study the mean squared error (MSE) of the extrapolated channel. Ref. [14] compares different extrapolation algorithms. This study shows that super-resolution can outperform compressed sensing methods for frequency channel extrapolation. In [15], information about user angles is extracted from UL pilots using 2D unitary ESPRIT, a subspace-based HRPE method [16]. Then, directional training is performed in the DL. Ref. [17] similarly proposes an angle-of-departure (AoD) adaptive subspace codebook to reduce channel feedback overhead. In [18], a hybrid statistical-instantaneous feedback mechanism where the users are separated into two classes of feedback design based on their channel covariance. Ref. [19] proposes to train a neural network to perform the channel extrapolation in frequency. This approach does not require the acquisition of the antenna
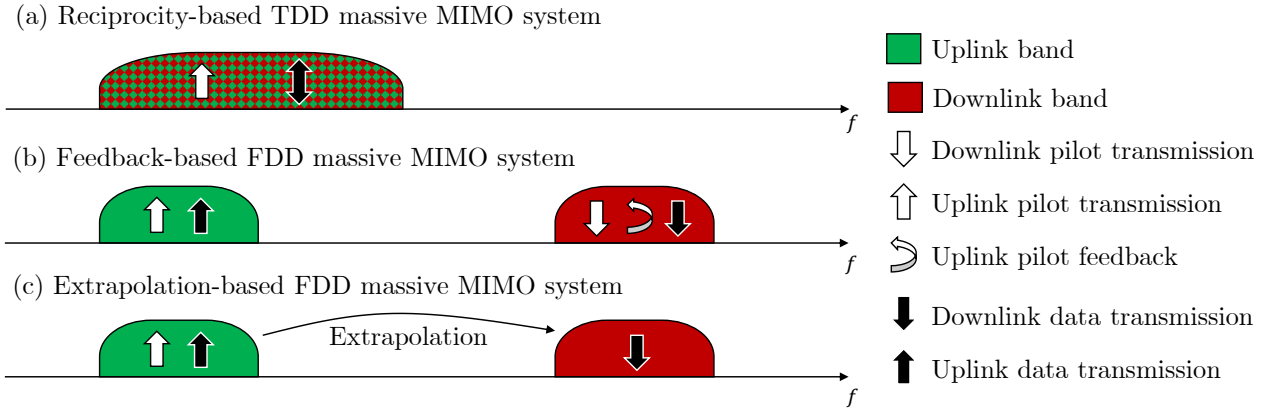
Fig. 1. (a) In TDD, the BS can rely on channel reciprocity to acquire full CSI from UL pilots as UL and DL links share the same band. (b) In FDD, UL and DL bands are disjoint and reciprocity does not hold. DL pilots and their feedback is generally used to acquire DL CSI at the base station. (c) In extrapolation-based FDD, the DL CSI is directly inferred from UL pilots.

array patterns through calibration but requires a large training dataset. Ref. [20] proposes to acquire DL CSI through UL pilots in combination with a limited feedback from DL pilots. In [21], channel extrapolation performance is experimentally evaluated, in terms of MSE of the extrapolated channel and beamforming efficiency.

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 [22]. In [23], 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 [24]. It is observed that MIMO provides much longer prediction lengths in time and frequency than for SISO systems.

To provide understanding of low-overhead FDD massive MIMO systems, this paper investigates the theoretical performance of channel extrapolation in frequency. The main originality of our paper is that it provides an in-depth theoretical study of the system performance as opposed to previous approaches, that were mostly validated through simulations and/or experiments. More specifically, we highlight the advantages of HRPE in terms of channel extrapolation as compared to conventional LS and LMMSE channel estimation. The channel MSE of both types of estimators is analytically studied. We derive the Cramer-Rao lower bound (CRLB) of the MSE, using a similar methodology as in [24]. The proposed CRLB differs from [24] by taking into account elevation angles, the frequency dependence of the pattern, and the influence of the training symbols. Furthermore, we propose a simplified CRLB, obtained under the assumption of well separated paths and giving more physical intuition about the frequency extrapolation range that can be expected in practice. The simplified CRLB shows 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. This paper furthermore studies analytically the downlink user performance under imperfect CSI, with emphasis on the induced beamforming power loss. Finally, extensive numerical evaluations validate the extrapolation-based FDD performance and carefully compares it to a corresponding TDD system. Various performance metrics are included such as channel MSE, beamforming efficiency, extrapolation range, spectral efficiency and uncoded symbol error rate.

The rest of this paper is structured as follows. Section II describes the channel model used in this work. Section III introduces the LS, the LMMSE and the high-resolution SAGE estimator. Section IV studies the theoretical performance of the previously introduced estimation algorithms. Section V numerically validates the performance of an extrapolation-based FDD massive MIMO system using a standardized 3GPP channel model. Finally, Section VI concludes the paper and appendixes contain the mathematical proof of previous sections.

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 $\jmath$, $\Im(.)$ and $\Re(.)$ denote the imaginary unit, imaginary and real parts, respectively. The expectation $\mathcal{E}[.]$ is taken over both the noise and channel statistics as opposed to $\mathbb{E}[.]$ which denotes the expectation with respect to the noise statistics only. The norm $\|.\|$ is the Frobenius norm and δ_n is the Kronecker delta. The $\mathrm{diag}(.)$ 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. CHANNEL MODEL

We consider a FDD massive MIMO scenario 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. For the sake of clarity and without loss of generality, we only consider one
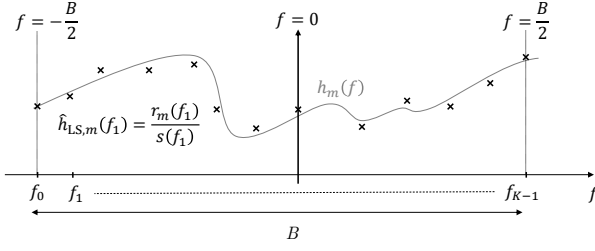
Fig. 2. The base station receives K pilot symbols scattered across the uplink training bandwidth B and at each antenna $m = 1, ..., M$. As a convention, the baseband frequency $f = 0$ corresponds to the passband uplink carrier frequency.
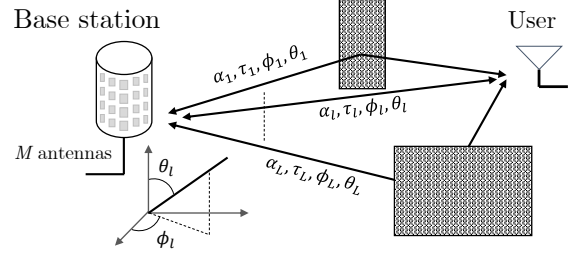

Fig. 3. Massive MIMO multipath propagation environment described by L components. Each multipath component is characterized by its complex gain α_l, its delay τ_l, its azimuth angle ϕ_l and elevation angle θ_l.

user in the following. We denote by M the number of BS antennas.

We consider the transmission of a single orthogonal frequency division multiplexing (OFDM) multicarrier symbol. Since pilot symbols are orthogonal to data symbols, we only consider the samples received at pilot subcarriers for clarity. As depicted in Fig. 2, the BS obtains a total of K pilot symbols scattered across frequency. The k-th transmitted pilot symbol is denoted by $s(f_k)$ with f_k being the baseband frequency of the pilot subcarrier. All pilot subcarriers are transmitted inside the uplink transmission band of width B, i.e., $f_k \in [-B/2, B/2]$, $k = 0, ..., K - 1$. As common, we assume that the channel is time-invariant between the transmission of the uplink pilots and the use of the (extrapolated) CSI, e.g., for downlink beamforming. In other words, the mobility of the environment should be low enough to ensure that the coherence time of the channel is larger than this delay. Note that this quasi-static assumption is better fulfilled than in most of the FDD works in the literature involving a feedback from the user. Indeed, these works assume that the channel remains time-invariant during the transmission of downlink pilots, feedback from the users and finally the use of the obtained CSI for downlink beamforming. The OFDM demodulated pilot symbol at antenna m and frequency f_k can be expressed as

$$r_m(f_k) = h_m(f_k)s(f_k) + w_m(f_k), \tag{1}$$

with $m = 1, ..., M$, $k = 0, ..., K - 1$ and where $h_m(f_k)$ is the channel frequency response at frequency f_k and antenna m. Samples $w_m(f_k)$ are zero mean additive complex circularly symmetric Gaussian noise of variance σ_w^2. We assume that the noise samples are uncorrelated, i.e., $\mathbb{E}\left(w_m(f_k)w_m^*(f_{k'})\right) = \sigma_w^2 \delta_{m-m'} \delta_{k-k'}$.

Furthermore, we assume that the propagation channel is composed of L specular paths, where each path is completely characterized by its deterministic parameters: complex gain $\alpha_l = \Re(\alpha_l) + \jmath\Im(\alpha_l)$, delay τ_l, azimuth angle ϕ_l and elevation angle θ_l, as depicted in Fig. 3. All antennas are assumed to be vertically polarized. Under these assumptions, the channel frequency response $h_m(f)$ can be expressed as

$$h_m(f) \triangleq \sum_{l=1}^{L} \alpha_l a_m(\phi_l, \theta_l, f)e^{-\jmath 2\pi f \tau_l}, \tag{2}$$

where $a_m(\phi, \theta, f)$ denotes 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, depending on the size of the band B and the targeted extrapolation frequency range. More specifically, the frequency selectivity of $a_m(\phi, \theta, f)$ comes from two contributions: (i) the frequency dependence of each individual antenna pattern and (ii) the frequency dependent phase shift across the antenna array elements (beam squinting). This 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.

Furthermore, a number of straightforward generalizations of the model in (2) 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 by considering their array pattern and angles of departures; (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 model of (1) in the remainder of this paper.

We finally note that a further requirement for using uplink CSI for downlink beamforming is reciprocity calibration [25]–[27], since upconverters and downconverters might have different transfer functions that have to be compensated for by suitable calibration. Since reciprocity calibration affects FDD and TDD systems in the same manner, we disregard it in the following derivations and simulations, i.e., we assume that it is perfect.

III. CHANNEL ESTIMATION AND EXTRAPOLATION

In this section, we first review conventional low-resolution channel estimators such as LS and LMMSE. We then explain the general concept of high-resolution channel estimation and we detail the principle of the SAGE algorithm. If the frequency of interest f of the channel estimate $\hat{h}_m(f)$ is inside the training band $f \in [-B/2, B/2]$, we refer to the estimation process as interpolation. Otherwise, if $f \notin [-B/2, B/2]$, we refer to it as extrapolation and f is also referred to as the extrapolation range. One can note that the interpolation performance corresponds to the channel estimation performance of a downlink TDD performance since uplink and downlink bands are shared in TDD mode.

A. Conventional Low-Resolution Estimation

As depicted in Fig. 2, LS estimators perform a simple per-antenna estimation at each pilot subcarrier as

$$\hat{h}_{\mathrm{LS},m}(f_k) = \frac{r_m(f_k)}{s(f_k)} = h_m(f_k) + \frac{w_m(f_k)}{s(f_k)}, \qquad (3)$$

for $k = 0, ..., K-1$. Based on the K channel estimates obtained at each pilot subcarrier, a linear method is generally used to obtain the channel at non-pilot subcarriers. We here propose to use a LMMSE estimator [28]. Denoting the vector containing the LS estimates by $\hat{\mathbf{h}}_{\mathrm{LS},m} \triangleq (\hat{h}_{\mathrm{LS},m}(f_0), \dots, \hat{h}_{\mathrm{LS},m}(f_{K-1}))^T$, the LMMSE estimate at frequency f is given by

$$\hat{h}_{\mathrm{LMMSE},m}(f) = \mathbf{p}_m^H(f)\hat{\mathbf{h}}_{\mathrm{LS},m}, \qquad (4)$$

where the vector of coefficient $\mathbf{p}_m^H(f)$ is obtained by minimizing the MSE of the estimate. Assuming that the complex channel $h_m(f_k)$ has a zero mean, this gives

$$\mathbf{p}_m^H(f) = \mathbf{c}_{\mathrm{LS},m}^H(f)\mathbf{C}_{\mathrm{LS},m}^{-1}$$
$$\mathbf{C}_{\mathrm{LS},m} \triangleq \mathcal{E}\left[\hat{\mathbf{h}}_{\mathrm{LS},m}\hat{\mathbf{h}}_{\mathrm{LS},m}^H\right], \ \mathbf{c}_{\mathrm{LS},m}^H(f) \triangleq \mathcal{E}\left[h_m(f)\hat{\mathbf{h}}_{\mathrm{LS},m}^H\right],$$

where the expectation $\mathcal{E}[.]$ is taken over both the noise and channel statistics as opposed to $\mathbb{E}[.]$ which denotes the expectation with respect to the noise statistics only. We can further write

$$[\mathbf{c}_{\mathrm{LS},m}^H(f)]_k = C_{h,m}(f, f_k)$$
$$[\mathbf{C}_{\mathrm{LS},m}]_{k,k'} = C_{h,m}(f_k, f_{k'}) + \frac{\sigma_w^2}{|s(f_k)|^2}\delta_{k-k'},$$

where

$$C_{h,m}(f, f') \triangleq \mathcal{E}\left(h_m(f)h_m(f')^*\right) \qquad (5)$$

is the autocorrelation function of the channel frequency response. An implementation challenge of the LMMSE estimator is the computation of $C_{h,m}(f, f')$. This function depends on the joint distribution of the path parameters. In the following, we approximate the computation of $C_{h,m}(f, f')$ by assuming that, as in [28], the paths gains and delays are i.i.d. We also assume a frequency independent pattern[1] $a_m(\phi, \theta, f) = a_m(\phi, \theta)$ and isotropic array pattern $|a_m(\phi, \theta)|^2 = 1$. Furthermore, the delay of each path τ_l has a uniform distribution in $[0, \tau_{\max}]$ while the complex path gain α_l has a uniform power across delay. This gives

$$C_{h,m}(f, f') = \frac{L}{\tau_{\max}}\mathcal{E}\left[|\alpha_l|^2\right]\int_0^{\tau_{\max}} e^{-j2\pi\Delta f\tau}d\tau$$
$$= P_h e^{-j\pi\Delta f\tau_{\max}}\operatorname{sinc}(\pi\Delta f\tau_{\max}),$$

where $\Delta f = f - f'$ and $P_h = C_{h,m}(0)$ is the averaged channel power. The LMMSE estimator performs relatively well in-band as long as the pilot spacing is smaller than $1/\tau_{\max}$, in accordance with the Nyquist sampling theorem. However, its extrapolation performance degrades quickly out of the training band, as will be analytically studied in Section IV. An intuitive

---

[1]This assumption is not very restrictive here given that, for typical values of B, the array can be considered frequency independent inside the training band B.

---

way to see this is to simply notice that the autocorrelation function $C_{h,m}(\Delta f)$ decays in $1/(\Delta f\tau_{\max})$. This implies that the extrapolation performance can only be satisfactory for extrapolation range f spaced about $1/\tau_{\max}$ away from the training band, i.e., $f \in [-\frac{B}{2} - \frac{1}{\tau_{\max}}, \frac{B}{2} + \frac{1}{\tau_{\max}}]$. For a typical delay spread of $\tau_{\max} = 2.5 \ \mu s$, we have $\frac{1}{\tau_{\max}} = 400$ kHz, which is much too low to deploy a typical FDD massive MIMO system.

Note that, to compute the expectation $\mathcal{E}[.]$, we used the long term statistics of the path parameters. Following the same idea, the LMMSE estimator presented here did not combine the LS samples from different antennas and hence did not leverage spatial correlation to improve the performance. This is justified for the case that the angular spread is sufficiently large (for a given antenna spacing) such that the correlation between antennas is close to zero, i.e., $\mathcal{E}\left(h_m(f)h_{m'}(f)^*\right) \approx 0$ if $m \neq m'$; e.g., if the antenna spacing is half a wavelength and the angular distribution of the scatterers is uniform [29]. This implies that combining the LS estimates from different antennas would not provide any significant gain.

On the other hand, an improved LMMSE estimator implementation would require to estimate the second-order statistics of the channel or "instantaneous" distribution of the path parameters. Under typical propagation conditions, paths delays and angles are clustered and thus far from being uniformly distributed. This implies that the frequency-space correlation function can be first estimated and then leveraged to significantly improve the LMMSE performance [4], [5]. However, this gain is properly taking into account by the high-resolution estimator proposed in the following.

B. High-Resolution Estimation

The poor extrapolation performance of the LMMSE estimator can be intuitively explained by the fact that the extrapolated channel does not exhibit a linear dependence on the in-band channel frequency response. On the other hand, high-resolution channel estimation allows to alleviate these limitations. As depicted in Fig. 4, instead of estimating the composite channel function $h_m(f)$, the HRPE approach directly estimates the parameters of each path. Taking advantage of the underlying (non-linear) physical dependence of the channel on its parameters can prove very useful to improve the extrapolation performance.

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, the high-resolution (HR) estimate of the extrapolated channel reads as

$$\hat{h}_{\mathrm{HR},m}(f) = \sum_{l=1}^{L}\hat{\alpha}_l a_m(\hat{\phi}_l, \hat{\theta}_l, f)e^{-j2\pi f\hat{\tau}_l}. \qquad (6)$$

Of course, intuitive reasoning tells us that the extrapolated channel will suffer from the estimation errors on the path parameters. Indeed, the finite bandwidth and aperture of the array directly induce a finite resolution in delay and angle, which leads to inaccuracy in the estimation of the delay and angle parameters. Moreover, the error on $\hat{h}_{\mathrm{HR},m}(f)$ becomes especially large as the extrapolation range becomes large. Indeed, one can note that the delay estimates $\hat{\tau}_l$ are multiplied
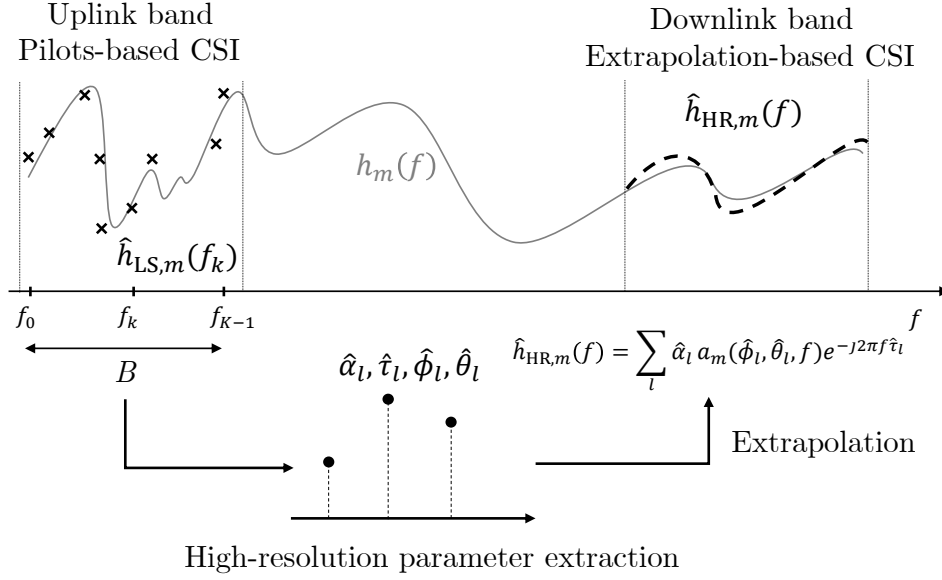
Fig. 4. Extrapolation-based FDD system: channel in the downlink band is extrapolated from pilots received in the uplink band, relying on high-resolution parameter extraction. CSI: channel state information.

by f so that the error increases as f increases. The impact of inaccurate parameter estimation will be carefully studied in Section IV-B. We also assume here that the parameters of the MPCs are independent of frequency. In other words, the channel is assumed to remain stationary on a frequency band including both uplink and downlink FDD disjoint bands. This is well fulfilled in most practical situations [30], since these parameters remain constant over a bandwidth corresponding to about 10% of the carrier frequency. For sub-6 GHz systems, this assumption is generally satisfied so that channel extrapolation could be used. Indeed, the LTE duplex spacing between uplink and downlink bands is: 10 to 70 MHz in the 700-900 MHz band and 45 to 190 MHz in the 1400-1700 MHz band. An exception are the Advanced Wireless Services (AWS) bands of LTE, which have up to 400 MHz spacing at 1700-2100 MHz carrier frequency.

To extract the path parameters $\hat{\tau}_l$, $\hat{\phi}_l$, $\hat{\theta}_l$ and $\hat{\alpha}_l$, we propose to use the SAGE algorithm originally described in [31] and widely adopted in the channel propagation community. We extended the algorithm to extract elevation angles and take into account the frequency dependence of the array. One should note that the main contribution of this paper is not to propose a novel efficient algorithm for high-resolution channel estimation but rather to show that the general performance bounds derived in the next section can be approached by conventional algorithms with reasonable complexity, such as the SAGE algorithm. Other HRPE algorithms could also be applied and may provide different results.

In brief, the algorithm aims at maximizing the likelihood of the received samples at pilot subcarriers as a function of the path parameters. Let us define $\boldsymbol{\psi} = (\boldsymbol{\psi}_1^T, \ldots, \boldsymbol{\psi}_L^T)^T \in \mathbb{R}^{5L \times 1}$ 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. The vectors $\hat{\boldsymbol{\psi}}$ and $\hat{\boldsymbol{\psi}}_l$ respectively de-

note their estimate. The optimization problem of maximizing the likelihood can be reformulated as

$$\min_{\hat{\boldsymbol{\psi}}} \sum_{m=1}^{M} \sum_{k=0}^{K-1} \left| r_m(f_k) - \sum_{l=1}^{L} \hat{\alpha}_l a_m(\hat{\phi}_l, \hat{\theta}_l, f) e^{-j2\pi f \hat{\tau}_l} s(f_k) \right|^2.$$

The problem is not easy to solve due to its high-dimensionality ($5L$) and the highly non-linear dependence on the path parameters. The SAGE algorithm provides an efficient suboptimal solution to the problem relying on an iterative approach. At each iteration, only the parameters corresponding to one path, $e.g.$, $\boldsymbol{\psi}_l$, are optimized while other path parameters keep their past value. This reduces the search dimensions from $5L$ dimensions to 5 dimensions at each iteration. Furthermore, inside each iteration, the 5-dimensional search is simplified into five one-dimensional searches optimizing each parameter one at a time using a line search. The algorithm iterates until convergence or if a maximal number of iterations is achieved. The initial estimates of each path are obtained by successive ordered cancellation.

IV. Performance Analysis

In this section, we start by studying the performance of the previously detailed algorithms in terms of the MSE of the estimated channel frequency response and the related error correlation matrix. In the last part of this section, we study the relationship between the MSE and the downlink user performance. More specifically, we will derive the expression of the SNR at the user side, taking into account the beamforming power loss induced by incorrect channel estimates.

We define the MSE of an estimate $\hat{h}_m(f)$ of $h_m(f)$ as

$$\text{MSE}_m(f) \triangleq \mathbb{E}\left[|\hat{h}_m(f) - h_m(f)|^2 \right], \tag{7}$$

where the expectation is taken over the noise realizations for a fixed channel realization $h_m(f)$ and thus the underlying

parameters $\boldsymbol{\psi}$. Similarly, the error correlation matrix of the channel vector at frequency f is defined as

$$\mathbf{E}(f) \triangleq \mathbb{E}\left[\left(\mathbf{h}(f) - \hat{\mathbf{h}}(f)\right)\left(\mathbf{h}(f) - \hat{\mathbf{h}}(f)\right)^H\right]$$

$$[\mathbf{E}(f)]_{m,m'} = \mathbb{E}\left[\left(\hat{h}_m(f) - h_m(f)\right)\left(\hat{h}_{m'}(f) - h_{m'}(f)\right)\right],$$

with $\mathrm{MSE}_m(f) = [\mathbf{E}(f)]_{m,m}$, $\mathbf{h}(f) = (h_1(f), ..., h_M(f))^T$ and $\hat{\mathbf{h}}(f) = (\hat{h}_1(f), ..., \hat{h}_M(f))^T$.

A. Conventional Low-Resolution Estimation

To simplify the following expressions of the performance of the LS and LMMSE estimators, we will assume equipowered pilot symbol, *i.e.*, $|s(f_k)|^2 = E_s$. This assumption is consistent with, *e.g.*, the Zhadoff-Chu training sequences in LTE and NR. Extension to the general case is straightforward. The total training power is then $E_T = \sum_{k=0}^{K-1} |s(f_k)|^2 = K E_s$. Based on the expression of the LS estimate in (3), the MSE expression at pilot subcarrier f_k is obtained as

$$\mathrm{MSE}_{\mathrm{LS},m}(f_k) = \mathbb{E}\left[|\hat{h}_{\mathrm{LS},m}(f_k) - h_m(f_k)|^2\right] = \frac{\sigma_w^2}{E_s}. \quad (8)$$

Moreover, the channel estimation errors at different antennas are uncorrelated given that noise samples are uncorrelated implying that $\mathbf{E}_{\mathrm{LS}}(f_k) = \frac{\sigma_w^2}{E_s}\mathbf{I}_M$.

Defining $\mathbf{h}_m \triangleq (h_m(f_0), \ldots, h_m(f_{K-1}))^T$ and using the expression of the LMMSE estimate in (4), the LMMSE performance for any frequency f in-band (interpolation) or out-of-band (extrapolation) is

$$\mathrm{MSE}_{\mathrm{LMMSE},m}(f) = |h_m(f)|^2 - 2\Re\left(h_m^*(f)\mathbf{p}_m^H(f)\mathbf{h}_m\right)$$
$$+ \mathbf{p}_m^H(f)\left(\mathbf{h}_m\mathbf{h}_m^H + \frac{\sigma_w^2}{E_s}\mathbf{I}_K\right)\mathbf{p}_m(f)$$

$$[\mathbf{E}_{\mathrm{LMMSE}}(f)]_{m,m'} = h_m(f)h_{m'}^*(f)$$
$$+ \mathbf{p}_m^H(f)\left(\mathbf{h}_m\mathbf{h}_{m'}^H + \frac{\sigma_w^2}{E_s}\mathbf{I}_K\delta_{m-m'}\right)\mathbf{p}_{m'}(f)$$
$$- h_m(f)\mathbf{h}_{m'}^H\mathbf{p}_{m'}(f) - \mathbf{p}_m^H(f)\mathbf{h}_m h_{m'}^*(f).$$

Note that we took the expectation only over the noise statistics and not the channel statistics ($\mathbb{E}(.)$ instead of $\mathcal{E}(.)$). Under these statistics, the LMMSE estimator is biased, *i.e.*, $\mathbb{E}\left[\hat{h}_{\mathrm{LMMSE},m}(f)\right] = \mathbf{p}_m^H(f)\mathbf{h}_m \neq h_m(f)$.

B. High-Resolution Estimation

The MSE performance of high-resolution estimation heavily depends on the choice of the algorithm and the result is typically not in closed-form. To circumvent this limitation and make our result more general and tractable, we will compute the CRLB of the channel estimate, which is by definition a theoretical bound and is independent of the choice of the algorithm. In other words, the goal of this paper is not to derive specific channel extrapolation algorithms and to study their computational complexity, but we propose general theoretical performance bounds. In the simulation section, we will show that SAGE performs close to the CRLB, implying that the bound can be approached by conventional algorithms and hence is useful.

To derive this bound, we can first notice that $\mathbf{h}(f)$ is a non-linear function of path parameters $\boldsymbol{\psi}$, as explicitly detailed in (2). Using this fact, we can apply the CRLB formula for non-linear transformation of parameters [32]. The bound tells us that for any unbiased estimator $\hat{\mathbf{h}}(f)$ of $\mathbf{h}(f)$, we have

$$\mathbf{E}(f) \succcurlyeq \mathbf{C}(f) \triangleq (\mathbf{G}(f))^H \mathbf{I}_\psi^{-1} \mathbf{G}(f), \quad (9)$$

where matrices $\mathbf{G}(f)$ and $\mathbf{I}_\psi$ are the Jacobian and Fisher information matrices respectively, whose forms are given in following subsections. The relationship $\mathbf{E}(f) \succcurlyeq \mathbf{C}(f)$ implies that the matrix $\mathbf{E}(f) - \mathbf{C}(f)$ is positive semidefinite, which directly implies that the MSE at antenna m and frequency f can be bounded by the corresponding diagonal element

$$\mathrm{MSE}_m(f) \geq [\mathbf{C}(f)]_{m,m} = \mathbf{g}_{m,f}^H \mathbf{I}_\psi^{-1} \mathbf{g}_{m,f}, \quad (10)$$

where $\mathbf{g}_{m,f}$ is the m-th column of $\mathbf{G}(f)$.

1) Fisher information matrix: Matrix $\mathbf{I}_\psi \in \mathbb{R}^{5L \times 5L}$ is the Fisher information matrix of the path parameters. Since the received samples $r_m(f_k)$ at each antenna and pilot subcarrier follow a circularly symmetric complex normal distribution with variance σ_w^2 and mean

$$\mu_{m,k} \triangleq \sum_{l=1}^{L} \alpha_l a_m(\phi_l, \theta_l, f_k) e^{-\jmath 2\pi f_k \tau_l} s(f_k),$$

we can directly use the CRLB formula for the general Gaussian case [32] to compute each element of the Fisher information matrix

$$[\mathbf{I}_\psi]_{u,v} = \frac{2}{\sigma_w^2} \sum_{k=0}^{K-1} \sum_{m=1}^{M} \Re\left(\frac{\partial \mu_{m,k}^*}{\partial \psi_u} \frac{\partial \mu_{m,k}}{\partial \psi_v}\right). \quad (11)$$

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

$$\mathbf{I}_\psi = \frac{2}{\sigma_w^2}\begin{pmatrix} \mathbf{I}_{\psi_1, \psi_1} & \cdots & \mathbf{I}_{\psi_1, \psi_L} \\ \vdots & \ddots & \vdots \\ \mathbf{I}_{\psi_L, \psi_1} & \cdots & \mathbf{I}_{\psi_L, \psi_L} \end{pmatrix}, \quad (12)$$

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

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 (11) as

$$\frac{d\mu_{m,k}}{d\tau_l} = \alpha_l a_m(\phi_l, \theta_l, f_k)(-\jmath 2\pi f_k)s(f_k)e^{-\jmath 2\pi f_k \tau_l}$$
$$\frac{d\mu_{m,k}}{d\phi_l} = \alpha_l \dot{a}_{m,\phi}(\phi, \theta, f_k)s(f_k)e^{-\jmath 2\pi f_k \tau_l}$$
$$\frac{d\mu_{m,k}}{d\theta_l} = \alpha_l \dot{a}_{m,\theta}(\phi, \theta, f_k)s(f_k)e^{-\jmath 2\pi f_k \tau_l}$$
$$\frac{d\mu_{m,k}}{d\alpha_l^R} = a_m(\phi_l, \theta_l, f_k)s(f_k)e^{-\jmath 2\pi f_k \tau_l}$$
$$\frac{d\mu_{m,k}}{d\alpha_l^I} = \jmath a_m(\phi_l, \theta_l, f_k)s(f_k)e^{-\jmath 2\pi f_k \tau_l}.$$

Inserting these partial derivatives in (11) and for a specific array pattern $a_m(\phi, \theta, f)$, the Fisher information matrix in (12) can be easily constructed. In the following, we will make the assumption.

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

In practice, a rank deficiency of $\mathbf{I}_\psi$ could arise if several paths become close in delay and angle, which would cause the determinant of $\mathbf{I}_\psi$ to go to zero. To be more accurate, the definition of "close distance in delay and angle" should always be measured relatively to the system Fourier resolution. For instance, if the system occupies a 10 MHz bandwidth ($B = 10$ MHz), inducing a resolution in delay of $1/B = 100$ ns, two rays are said to be close in delay if their spacing is much smaller than 100 ns. Depending on the underlying physical phenomenon inducing the presence of dense multipath components, different solutions may be possible to address a rank deficiency of $\mathbf{I}_\psi$.

If two rays, or more, are closely spaced in angle and delay and their delay separation is not only smaller than $1/B$ but also much smaller than $1/f$ where f is the targeted extrapolation range, then they can be replaced by a single ray whose complex gain is given by the sum of the complex amplitudes of the correlated rays. As an example, if the targeted extrapolation performance is $f = 100$ MHz, two rays with a spacing smaller than $1/f = 10$ ns can be combined without affecting the extrapolation performance significantly.

If the source of dense multipath components is related to wavefront curvature, a more advanced channel model can be taken into account to address them [33], [34]. Indeed, our channel model in (2) relies on a plane wave assumption. The presence of spherical waves would result in a large number of plane waves in (2), inducing a potential ill-conditioning of $\mathbf{I}_\psi$.

If, finally, the dense multipath components are present due to a truly rich scattering environment, instead of trying to estimate a large number of possibly unreliable paths, a better solution may be to consider them as random components. Rather than estimating their instantaneous values, their statistics can be estimated and included in the likelihood formulation to improve the conditioning of $\mathbf{I}_\psi$. This approach is used by, *e.g.*, the RIMAX algorithm [35].

2) Jacobian matrix: Matrix $\mathbf{G}(f) \in \mathbb{C}^{5L \times M}$ is the Jacobian matrix of the transformation defined as

$$\mathbf{G}(f) \triangleq \frac{\partial \mathbf{h}^T(f)}{\partial \psi}, \ [\mathbf{G}(f)]_{v,m} = \frac{\partial h_m(f)}{\partial \psi_v}.$$

It can be partitioned into columns corresponding to each antenna element as $\mathbf{G}(f) = (\mathbf{g}_{1,f}, ..., \mathbf{g}_{M,f})$. Furthermore, each vector $\mathbf{g}_{m,f}$ can be partitioned into different paths and

path parameters as

$$\mathbf{g}_{m,f} = (\mathbf{g}^T_{m,f,\boldsymbol{\psi}_1}, \ldots, \mathbf{g}^T_{m,f,\boldsymbol{\psi}_L})^T$$

$$\mathbf{g}_{m,f,\boldsymbol{\psi}_l} = (g_{m,f,\tau_l}, g_{m,f,\phi_l}, g_{m,f,\theta_l}, g_{m,f,\alpha_l^I}, g_{m,f,\alpha_l^R})^T$$

$$g_{m,f,\tau_l} = (-\jmath 2\pi f)\alpha_l a_m(\phi_l, \theta_l, f)e^{-\jmath 2\pi f \tau_l}$$

$$g_{m,f,\phi_l} = \alpha_l \dot{a}_{m,\phi}(\phi_l, \theta_l, f)e^{-\jmath 2\pi f \tau_l}$$

$$g_{m,f,\theta_l} = \alpha_l \dot{a}_{m,\theta}(\phi_l, \theta_l, f)e^{-\jmath 2\pi f \tau_l}$$

$$g_{m,f,\alpha_l^R} = a_m(\phi_l, \theta_l, f)e^{-\jmath 2\pi f \tau_l}$$

$$g_{m,f,\alpha_l^I} = \jmath a_m(\phi_l, \theta_l, f)e^{-\jmath 2\pi f \tau_l}.$$

3) Separated Rays: The CRLB of (10) 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 extrapolation range that can be expected. 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 does not generally depend on the channel but rather on the system parameters such as type of BS antennas, extrapolation range and carrier frequency. The assumption particularly 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**)

$$\mathbf{s}_l \triangleq \left(s(f_0)e^{-\jmath 2\pi f_0 \tau_l} \quad \ldots \quad s(f_{K-1})e^{-\jmath 2\pi f_{K-1}\tau_l}\right)^T$$

$$\dot{\mathbf{s}}_l \triangleq -\jmath 2\pi(f_0 s(f_0)e^{-\jmath 2\pi f_0 \tau_l}, ..., f_{K-1}s(f_{K-1})e^{-\jmath 2\pi f_{K-1}\tau_l})^T$$

$$\mathbf{a}_l \triangleq \left(a_1(\phi_l, \theta_l) \quad \ldots \quad a_M(\phi_l, \theta_l)\right)^T \in \mathbb{C}^{M \times 1}$$

$$\dot{\mathbf{a}}_{l,\phi} \triangleq \left(\dot{a}_{1,\phi}(\phi_l, \theta_l) \quad \ldots \quad \dot{a}_{M,\phi}(\phi_l, \theta_l)\right)^T \in \mathbb{C}^{M \times 1}$$

$$\dot{\mathbf{a}}_{l,\theta} \triangleq \left(\dot{a}_{1,\theta}(\phi_l, \theta_l) \ldots \quad \dot{a}_{M,\theta}(\phi_l, \theta_l)\right)^T \in \mathbb{C}^{M \times 1}.$$

(**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:

$$\mathbf{s}_l^H \mathbf{s}_{l'} = \dot{\mathbf{s}}_l^H \dot{\mathbf{s}}_{l'} = \dot{\mathbf{s}}_l^H \mathbf{s}_{l'} = 0. \quad (13)$$

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

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

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

The assumption (**As3**) is a strong assumption, whose accuracy will typically depend on different parameters. The specular paths will generally become relatively more separated in delay as the bandwidth of $s(f)$ increases, inducing higher resolution in delay. Similarly, the resolution and hence the separation in azimuth and elevation will be improved as the number of antenna elements M is increased. More generally,

for a given channel, 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. Moreover, the validity of (**As3**) will be assessed in Section V using practical channel models.

(**As4**): the transmitted pilots $s(f_k)$ have a symmetric energy distribution implying that $|s(f)|^2 = |s(-f)|^2$ and

$$\dot{\mathbf{s}}_l^H \mathbf{s}_l = (\jmath 2\pi) \sum_{k=0}^{K-1} f_k |s(f_k)|^2 = 0, \quad l = 1, \dots, L. \quad (14)$$

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

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

The symmetric condition on the pilot energy is satisfied in conventional systems such as LTE or NR since pilots have uniform energy while the condition on the array pattern is generally satisfied for symmetric arrays. For instance, it is easy to check that the condition is fulfilled for a rectangular array if each antenna element has an isotropic pattern according to (22) later studied in Section V. The following bound gives a particularization of the CRLB of (10) under additional assumptions (**As2**) − (**As4**) and for the MSE averaged over the receive antennas, *i.e.*,

$$\text{MSE}(f) \triangleq \frac{1}{M} \text{tr}\left[\mathbf{E}(f)\right] = \frac{1}{M} \sum_{m=1}^{M} \text{MSE}_m(f).$$

Proposition 1. *Under* (**As2**) − (**As4**)*, the expression of the CRLB of (10) averaged over the receive antennas simplifies to*

$$\text{MSE}(f) \geq \frac{\sigma_w^2}{E_T} \underbrace{\frac{L}{M}}_{\text{SNR gain}} \left(\underbrace{2}_{\text{Loss factor}} + \underbrace{\frac{1}{2}\left(\frac{f}{\sigma_F}\right)^2}_{\text{Extrapolation penalty}} \right),$$

$$(15)$$

where E_T is the total training power and σ_F^2 is the mean squared bandwidth of the transmit signal

$$E_T = \|\mathbf{s}_l\|^2 = \sum_{k=0}^{K-1} |s(f_k)|^2$$

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

Proof. The proof is given in the Appendix. □

By adding some assumptions, the CRLB can be greatly simplified and provides much insight into the physical meaning of the different terms of the bound. We can clearly identify the two main advantages of high-resolution channel estimation. As compared to the LS estimation performance that we derived in (8) where the total pilot power is $E_T = KE_s$, a gain of a factor $\frac{MK}{L}$ can be observed. This gain comes from two contributions: the array gain M and the estimation of only L channel coefficients instead of K as in the LS case. However, a loss factor of 2 appears, coming from the penalty of estimating the azimuth and elevations angles of each path. Moreover, the channel can be extrapolated in frequency at the cost of a MSE penalty that quadratically scales with the ratio f/σ_F,

which physically makes sense. Indeed, as the extrapolation range f increases, the estimate quality worsens. On the other hand, as the uplink training bandwidth increases, the delays of each path are better estimated, which leads to an improved extrapolation performance. Note that the denominator σ_F indicates that the extrapolation range can be quantified in multiples of the uplink training band B.

It is interesting to see that the simplified CRLB does not depend on the path parameters ψ for well separated paths. 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 CRLB, we can find a closed-form expression of the extrapolation range: we define the γ extrapolation range, denoted by $f_{\text{Extrapol}-\gamma}$, as the frequency f beyond which the extrapolation performance falls γ times below that of the conventional LS estimator given in (8). Using the expressions of (8) and (15), we easily find

$$f_{\text{Extrapol}-\gamma} = 2\sigma_F \sqrt{\frac{MK\gamma}{2L} - 1}. \quad (16)$$

Note that this definition is independent of the ratio E_s/σ_w^2.

C. Relationship between channel MSE and user SNR

The goal of this section is to find how the user performance is affected by imperfect CSI at the BS. Let us assume that the BS communicates in the downlink by beamforming in the direction of the user using a beamforming vector $\mathbf{g}(f) \in \mathbb{C}^{M \times 1}$. The frequency f denotes the pilot subcarrier frequency at which the symbol $d(f)$ is transmitted. The BS uses maximum ratio combining and normalizes the beamforming vector to have unit power so that

$$\mathbf{g}(f) = \frac{\left(\hat{\mathbf{h}}(f)\right)^*}{\|\hat{\mathbf{h}}(f)\|}.$$

The demodulated symbol at the user side is

$$r(f) = (\mathbf{h}(f))^T \mathbf{g}(f) d(f) + w(f),$$

and the related SNR, averaged over the statistics of the transmit symbols, the noise on uplink pilots and the noise on downlink symbols,

$$\text{SNR}_{\text{DL}}(f) = \frac{E_{d,f}}{\sigma_w^2} \mathbb{E}\left[|(\mathbf{h}(f))^T \mathbf{g}(f)|^2\right] \quad (17)$$

$$= \frac{E_{d,f} \|\mathbf{h}(f)\|^2}{\sigma_w^2} \underbrace{\mathbb{E}\left[\frac{|(\hat{\mathbf{h}}(f))^H \mathbf{h}(f)|^2}{\|\hat{\mathbf{h}}(f)\|^2 \|\mathbf{h}(f)\|^2}\right]}_{\eta(f)},$$

where the channel $\mathbf{h}(f)$ is considered as deterministic (not random) and $E_{d,f}$ is the energy of downlink symbol $d(f)$. The term $\eta(f)$ is the so-called beamforming efficiency bounded as $0 \leq \eta(f) \leq 1$. It represents the beamforming power loss due to imperfect CSI. If it close to 1, almost no loss is induced. On the other hand, if it is close to 0, the efficiency

is strongly affected. We assume that coherent demodulation can be practically achieved at the user side by including pilots in the downlink transmission frame. This allows the users to estimate and compensate their equivalent channel after beamforming by the BS. Note that these pilots are beamformed to each user. They are thus orthogonal and can be transmitted at the same time-frequency resource to different users, implying that their overhead is negligible. Moreover, they do not need to be fed back to the BS. Note that the use of such user-specific reference signals for channel estimation is a common feature foreseen for 5G NR deployment [36], [37]. In the following, we neglect their overhead in the spectral efficiency computation. This is motivated by the fact that they are also present in TDD systems, implying that the comparison remains fair.

The analytical expression of $\eta(f)$ is complicated to compute as it involves the expectation of a ratio of random variables. Therefore, we propose to approximate the expectation of the ratio by the ratio of expectations, which corresponds to the first-order Taylor expansion of the ratio around the mean of its numerator and denominator

$$\eta(f) \approx \hat{\eta}(f) = \frac{\mathbb{E}\left[|(\hat{\mathbf{h}}(f))^H \mathbf{h}(f)|^2\right]}{\mathbb{E}\left[\|\hat{\mathbf{h}}(f)\|^2 \|\mathbf{h}(f)\|^2\right]} \tag{18}$$

$$= \frac{\mathbb{E}\left[|(\mathbf{h}(f) + \mathbf{e}(f))^H \mathbf{h}(f)|^2\right]}{\mathbb{E}\left[\|\mathbf{h}(f) + \mathbf{e}(f)\|^2\right] \|\mathbf{h}(f)\|^2},$$

where we defined the estimation error $\mathbf{e}(f) \triangleq \hat{\mathbf{h}}(f) - \mathbf{h}(f)$. We expect the approximation $\hat{\eta}(f)$ to asymptotically converge to $\eta(f)$ as the number of antennas grows large. Indeed, one can note that both the numerator and the denominator of $\eta(f)$ involve a sum of M elements. Thus, as M grows large, the estimation error gets better averaged and we expect the numerator and denominator to converge to their expected value. Based on the statistics of the estimation error $\mathbf{e}(f)$ derived in previous sections, $\hat{\eta}(f)$ can be easily computed. Furthermore, for an unbiased estimator, we have $\mathbb{E}(\mathbf{e}(f)) = \mathbf{0}$ and $\hat{\eta}(f)$ simplifies to

$$\hat{\eta}(f) = \frac{\|\mathbf{h}(f)\|^2 + \frac{(\mathbf{h}(f))^H \mathbf{E}(f)\mathbf{h}(f)}{\|\mathbf{h}(f)\|^2}}{\|\mathbf{h}(f)\|^2 + \text{tr}\left[\mathbf{E}(f)\right]}. \tag{19}$$

The second term of the numerator has the form of a Rayleigh quotient, which is upper and lower bounded by the maximal and minimal eigenvalues of matrix $\mathbf{E}(f)$ respectively. This implies that the numerator is always larger than the denominator, ensuring that $\hat{\eta}(f) \leq 1$ as expected.

In the end, we have found an analytical expression that relates the channel estimation error correlation matrix to the loss in beamforming power at the user side. The corresponding loss in terms of capacity, spectral efficiency and bit error rate can be directly inferred from the beamforming efficiency $\hat{\eta}(f)$. Indeed, the spectral efficiency at extrapolation range f can be inferred as

$$C(f) = \log_2\left(1 + \text{SNR}_{\text{DL}}(f)\right) \tag{20}$$

$$\approx \log_2\left(1 + \frac{E_{d,f} \|\mathbf{h}(f)\|^2}{\sigma_w^2} \hat{\eta}(f)\right) \quad \text{[bits/symbols]}.$$

Similarly, the uncoded symbol error rate for M-QAM symbols can be inferred as [29]

$$\text{SER}(f) = 2\frac{\sqrt{M} - 1}{\sqrt{M}} \text{erfc}\left(\sqrt{\frac{3\text{SNR}_{\text{DL}}(f)}{2(M - 1)}}\right) \tag{21}$$

$$\approx 2\frac{\sqrt{M} - 1}{\sqrt{M}} \text{erfc}\left(\sqrt{\frac{3E_{d,f}\|\mathbf{h}(f)\|^2 \hat{\eta}(f)}{2(M - 1)\sigma_w^2}}\right),$$

where $\text{erfc}(.)$ is the complementary error function.

Note that we only considered the performance of a single-user. The same methodology could be extended to study the performance of multiple users communicating at the same time and frequency. This type of study can be conducted relying on the formulas derived in this work for the statistics of the channel estimation errors and using mathematical tools from the random matrix theory literature [38], [39]. Note that multi-user beamforming is more sensitive to channel estimation error as it can lead to inter-user interference.

V. Simulation Results

This section evaluates the performance of channel extrapolation for FDD massive MIMO system. The performance of the conventional LMMSE estimator will be compared to the high-resolution estimation based on the SAGE algorithm. The theoretical CRLB of the channel MSE derived in Section IV-B is also included as a benchmark. Furthermore, the beamforming efficiency studied in Section IV-C will be used to relate the MSE of the channel estimates to the user link performance in terms of SNR and spectral efficiency. Finally, graphs will include the performance of a corresponding TDD system. We emphasize that the loss of performance of the FDD versus TDD comes from the less accurate channel state information due to extrapolation. The TDD performance can be simply inferred from the previous analytical results as the performance related to the in-band channel estimation, which amounts to channel interpolation based on pilot estimates rather than extrapolation in the FDD mode.

In the simulations, we assumed that the pilots $s(f_k)$ have uniform energy distribution over the K frequency points f_k, i.e., $|s(f_k)|^2 = E_s$ for $k = 0, \ldots, K - 1$. The pilots are uniformly spaced across the uplink bandwidth B with spacing $1/\tau_{\max}$, i.e., $f_k = \left(k - \frac{K-1}{2}\right)/\tau_{\max}$ for $k = 0, \ldots, K - 1$ and $K = B\tau_{\max} + 1$. This assumption is consistent with, e.g., the Zhadoff-Chu training sequences in LTE and NR. We set $\tau_{\max} = 2.5\mu s$. The center frequency of the uplink training band is set to $f_c = 3.5$ GHz. Note that our system model is in baseband frequency. Hence, the carrier frequency f_c corresponds to a zero baseband frequency ($f = 0$). We consider a synthetic rectangular planar array at the BS 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

$$a_m(\phi, \theta, f) = e^{-\jmath 2\pi \frac{f_c + f}{c} \mathbf{r}_m \cdot \hat{\mathbf{e}}(\phi, \theta)}, \tag{22}$$

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 a reference
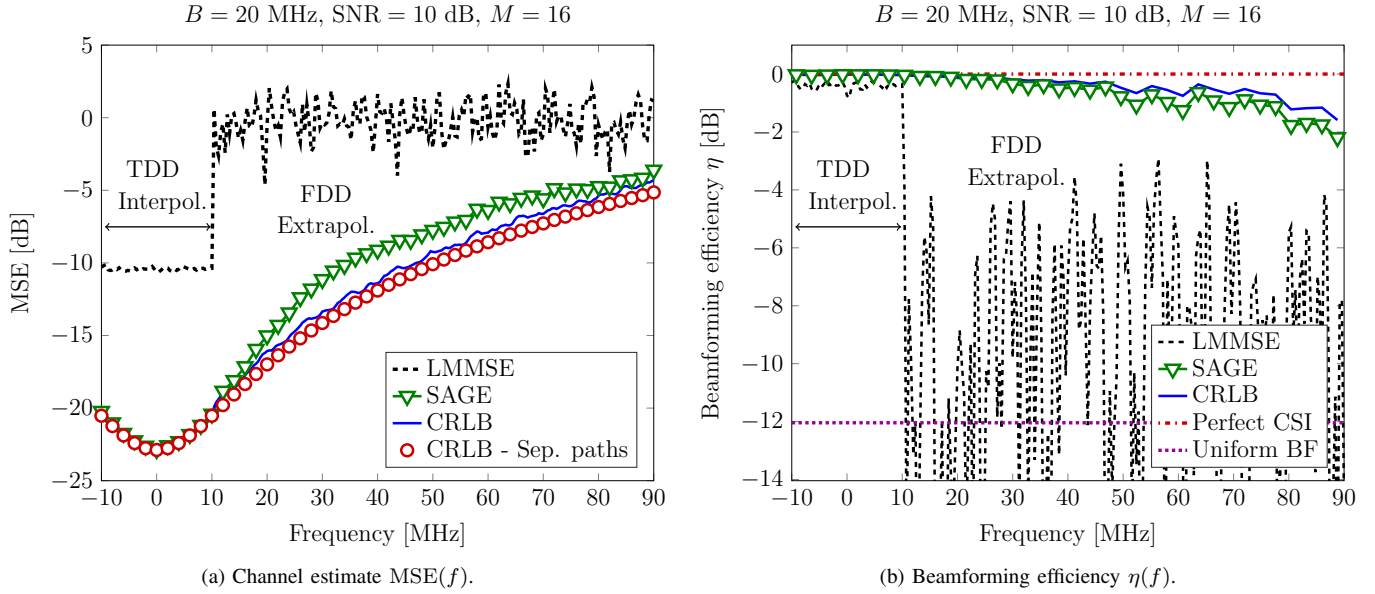
Fig. 5. The LMMSE estimator performs poorly in terms of extrapolation while the high resolution SAGE estimator performs much better and approaches the CRLB. The simplified CRLB, assuming well separated paths, performs close to the CRLB.

point. The reference point is chosen to ensure that $\sum_m \mathbf{r}_m = \mathbf{0}$. Note that (22) is frequency dependent because of the beam squint effect. Three planar array geometries are considered: $M = 4$ (2 Horiz. $\times$ 2 Vert.), $M = 16$ (4 Horiz. $\times$ 4 Vert.) and $M = 64$ (8 Horiz. $\times$ 8 Vert.).

The channel frequency response and received samples are generated according to (2) and (1) respectively. The path parameters ψ are generated by the QuaDRiGa toolbox [40] according to the 3D-UMa NLOS model defined by 3GPP TR 36.873 v12.5.0 specifications [41]. 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 average channel power is normalized to one and the per-pilot SNR, defined as SNR $\triangleq E_s/\sigma_w^2$, is set to 10 dB.

A. High-Resolution versus LMMSE Estimation

We start by analyzing the performance for a single realization of channel parameters. Fig. 5 compares the performance of the LMMSE and the high-resolution SAGE estimators and the system parameters $M = 16$, SNR $= 10$ dB and $B = 20$ MHz. This implies that the uplink pilots belong to the support $[-10, 10]$ MHz, which corresponds to the operating band of a corresponding TDD system since uplink and downlink bands are shared. The performance of the algorithms was averaged over 1000 noise realizations. A delay step size of $\frac{1}{50B} = 1$ ns and an angular grid size of 1 degree are used as parameters of the SAGE grid search.

Fig. 5. (a) depicts the performance in terms of the MSE of the channel estimates. The SAGE-based channel extrapolation approaches the MSE performance of the theoretical CRLB. This implies that the CRLB gives a good indication of the achievable MSE. Furthermore, we can see in the figure that the CRLB performs close to the simplified one of Proposition 1,

obtained under the assumption of well separated paths. On the other hand, the LMMSE estimator performs worse than high-resolution, especially out of the training band where the error appears to abruptly jump to higher values. Indeed, its transition zone is of the order of a few hundreds of kHz, which is negligible compared to the considered extrapolation range of about 100 MHz. In-band, it also performs worse than SAGE as it does not exploit the joint spatial-frequency structure of the channel but performs independent per-antenna estimation.

Fig. 5. (b) shows the beamforming efficiency $\eta(f)$ related to the different estimators. As a reminder, the beamforming efficiency, studied in Section IV-C, corresponds to the loss of received signal power due to the impact of incorrect channel estimates on the beamformer. In the case of perfect CSI, $\eta(f) = 1 = 0$ dB and the efficiency is maximized. Note that both estimators perform similarly in-band or in TDD mode. The reduction of $\eta(f)$ as the extrapolation frequency f increases can be seen as the loss induced in the FDD system as compared to a TDD system. Here again, the LMMSE performance degrades very quickly out of the band. Note that its performance gets even worse at some points than a simple uniform beamforming strategy in all directions, i.e., $\mathbf{g}(f) = \mathbf{1}/\sqrt{M}$ implying a beamforming efficiency $\eta(f) = 1/M \approx -12$ dB. On the other hand, the SAGE extrapolation approaches the efficiency related to the CRLB and only suffers from a beamforming power loss of about 2 dB at an extrapolation frequency of 90 MHz.

B. Impact of Number of Antennas

For the same set of channel parameters, Fig. 6 (a) depicts the extrapolation performance for different numbers of antenna elements. As the number of antennas increases, the resolution in the angle domain increases and the BS can better resolve the paths. This implies that (**As3**) becomes valid and the CRLB
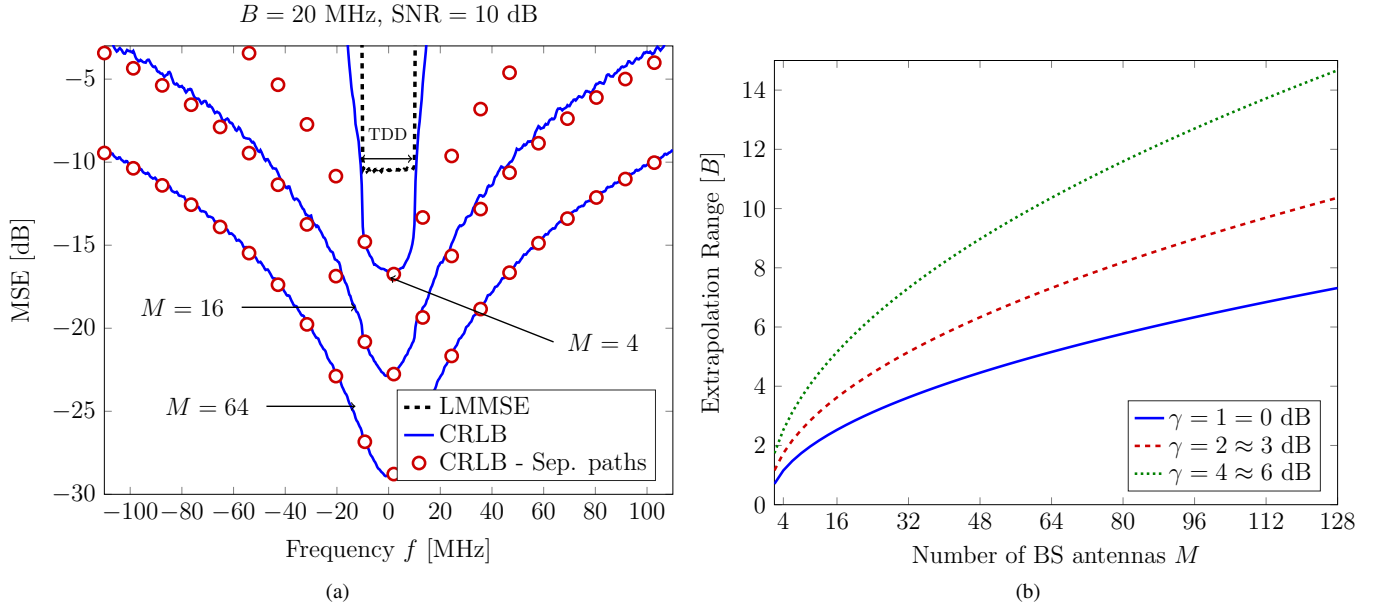
Fig. 6. (a) As number of antennas increases, the CRLB reaches the CRLB formula for separated paths while achieving an additional SNR gain. (b) Extrapolation range $f_{\text{Extrapol}-\gamma}$, quantified in multiples of training band B, is defined as the extrapolation range beyond which the CRLB falls below γ times the one of the LS estimator in-band. It is based on the well separated CRLB formula.

converges to the simplified CRLB. Note that, in the $M = 4$ case, extrapolation performance in terms of MSE deteriorates quickly as we move away from the uplink band. The simplified CRLB of Proposition 1 is very close to the full CRLB as soon as the array has 16 antennas. Moreover, as the number of antennas M increases, a corresponding in-band array gain is achieved implying that the curves are shifted 6 dB down as the number of antennas is multiplied by 4.

In Fig. 6 (b), the extrapolation range $f_{\text{Extrapol}-\gamma}$, expressed in (16), is plotted as a function of the number of antennas. As a reminder, the formula corresponds to the extrapolation range f beyond which the CRLB performance is γ times worse than the one of the conventional LS estimator in-band. The formula is independent of the SNR and assumes that the paths are well separated, *i.e.*, (**As3**) is valid. As an example, if the BS has 128 antennas and the uplink training band is $B = 20$ MHz, the FDD separation between uplink and downlink bands should be at most $7B = 140$ MHz to guarantee that the extrapolation-based FDD downlink MSE performance is equivalent to the corresponding one of a TDD system using conventional LS estimation.

C. Spectral Efficiency

To evaluate the loss in spectral efficiency of an extrapolation-based FDD system versus a corresponding TDD system, we use the QuaDRIiGa toolbox to generate the channel parameters related to 200 users locations randomly distributed in a radius of 200 meters around the BS; the BS height is 20m above ground. The channel model used is still the 3GPP 3D-UMa NLOS model. For each user location, we compute the spectral efficiency at frequency f according to (20) where $\hat{\eta}(f)$ is computed based on (19) and $\mathbf{E}(f) = \mathbf{C}(f)$ given in (9). The SNR related to uplink pilots is set to SNR $= 10$ dB and the downlink SNR on the received symbol

is $E_{d,f}/\sigma_w^2 = 10$ dB as well. The uplink training bandwidth is left to 20 MHz.

Fig. 7 (a) plots the cumulative density function (CDF) of the spectral efficiency for two antenna settings $M = 16$ and $M = 64$ and for 4 values of the extrapolation frequency f: $f = 0$ corresponds to the uplink carrier frequency, which is in-band and also corresponds to the downlink performance of a TDD system, $f \in \{80, 160, 240\}$ MHz corresponds to the FDD downlink performance of systems working at different extrapolation frequencies. The performance of a perfect CSI system is also plotted but can hardly be distinguished from the $f = 0$ TDD performance. As the extrapolation frequency f increases, the inaccurate CSI induces a beamforming power loss and a related loss of spectral efficiency for FDD systems as compared to TDD systems. Note that this loss is much more pronounced for the $M = 16$ case than the $M = 64$ case. This can be explained by the fact that the channel estimation and related extrapolation is more accurate in the $M = 64$ case as improved spatial resolution is available.

Fig. 7 (b) plots the uncoded symbol error rate performance, computed based on (21) averaged over the different user locations, for the same parameters in the $M = 64$ case. A 256-QAM constellation is considered. Here again, TDD performance ($f = 0$) and perfect CSI curves fit. As the extrapolation frequency f increases, the performance of the corresponding FDD system is degraded but still remains in a 2 dB range from the TDD performance.

VI. CONCLUSIONS

This paper investigated the performance of extrapolation-based FDD massive MIMO systems, relying on high-resolution parameter estimation. We demonstrated that, under a good calibration of the BS and favorable propagation conditions, channel extrapolation is a viable solution to deploy
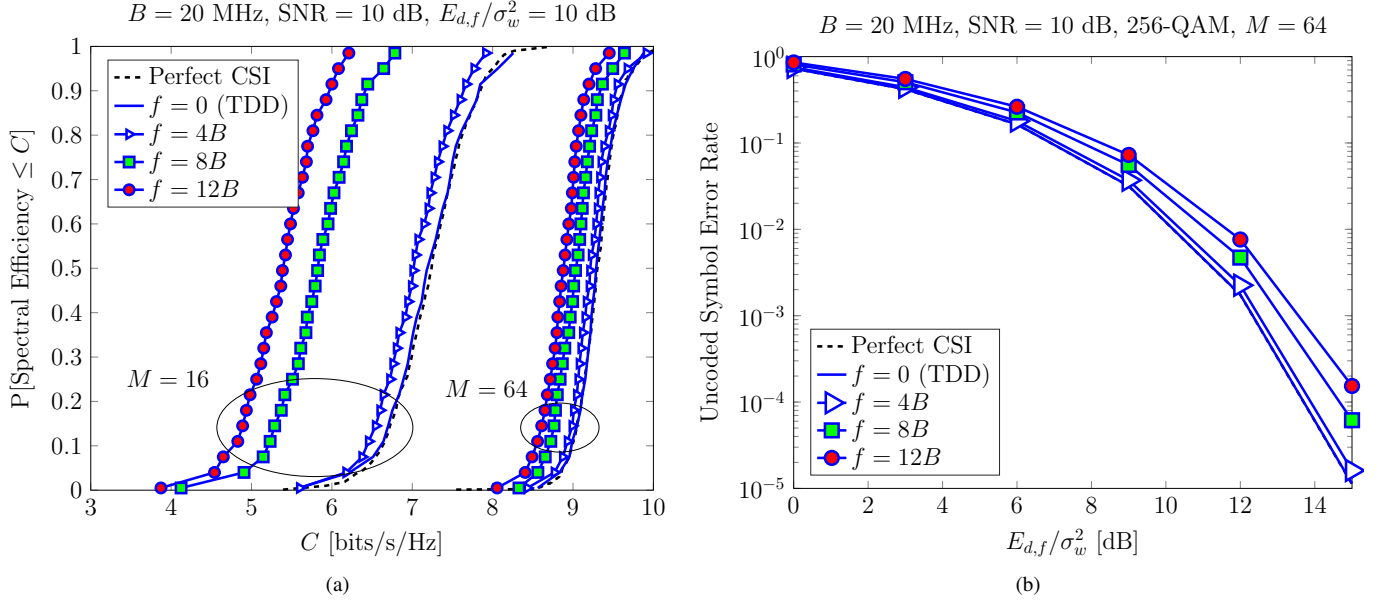
Fig. 7. (a) Cumulative density function of the spectral efficiency of extrapolation-based FDD systems as compared to a TDD system, for different number of BS antennas M and extrapolation frequency f (separation between uplink carrier frequency and downlink subcarrier frequency). (b) Corresponding uncoded symbol error rate performance for $M = 64$ antenna case.

FDD massive MIMO systems. It has the great advantage to drastically reduce the DL pilot overhead and completely remove the need for a feedback from the users. The price to pay is a reduction in the quality of the channel estimates, which results in a performance loss in the user downlink transmission. Theoretical CRLB for the MSE of the extrapolated channel and the related user SNR performance were derived and validated through numerical simulations. Our simulation results show that extrapolation-based FDD systems relying on high-resolution channel estimation are a feasible and attractive solution, even as compared to a corresponding TDD system. In particular, we showed that the FDD performance only suffers from a 1 to 3 dB reduction in beamforming power for extrapolation range as large as 200 MHz for a BS equipped with 64 antennas.

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.*, dense multipath components, will require further investigation. Another interesting perspective is to take into account the impact of multiple antennas at the user side. Intuitively, this should help the extrapolation as we saw that the performance highly depends on the path separability in at least one domain, which helps their estimation as being free from the interference of other paths. In SIMO, paths can only be separated in the delay and angle of arrival domain while in MIMO they can be additionally separated in the angle of departure domain. Similarly, accounting for the time variations of the channel could be beneficial too as paths could be separated in the Doppler domain.

VII. APPENDIX

Using (11), we can compute the different elements of the Fisher information matrix given in (12). 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 on 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

$$I_{\tau_l \tau_{l'}} = \Re\left(\alpha_l^* \alpha_{l'} \mathbf{a}_l^H \mathbf{a}_{l'} \dot{\mathbf{s}}_l^H \dot{\mathbf{s}}_{l'}\right) = |\alpha_l|^2 \|\mathbf{a}_l\|^2 \|\dot{\mathbf{s}}\|^2 \delta_{l-l'}$$

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

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

$$I_{\alpha_l^R \alpha_{l'}^R} = I_{\alpha_l^I \alpha_{l'}^I} = \Re\left(\mathbf{a}_l^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}\right) = \|\mathbf{a}_l\|^2 \|\mathbf{s}\|^2 \delta_{l-l'}.$$

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

$$I_{\tau_l \phi_{l'}} = -\Re\left(\alpha_l^* \alpha_{l'} \mathbf{a}_l^H \dot{\mathbf{a}}_{l',\phi} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\tau_l \theta_{l'}} = -\Re\left(\alpha_l^* \alpha_{l'} \mathbf{a}_l^H \dot{\mathbf{a}}_{l',\theta} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\tau_l \alpha_{l'}^R} = -\Re\left(\alpha_l^* \mathbf{a}_l^H \mathbf{a}_{l'} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\tau_l \alpha_{l'}^I} = \Im\left(\alpha_l^* \mathbf{a}_l^H \mathbf{a}_{l'} \dot{\mathbf{s}}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\phi_l \alpha^R} = \Re\left(\alpha_l^* \dot{\mathbf{a}}_{l,\phi}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\phi_l \alpha^I} = -\Im\left(\alpha_l^* \dot{\mathbf{a}}_{l,\phi}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\theta_l \alpha_{l'}^R} = \Re\left(\alpha_l^* \dot{\mathbf{a}}_{l,\theta}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\theta_l \alpha_{l'}^I} = -\Im\left(\alpha_l^* \dot{\mathbf{a}}_{l,\theta}^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}\right) = 0$$

$$I_{\alpha_l^R \alpha_{l'}^I} = -\Im\left(\mathbf{a}_l^H \mathbf{a}_{l'} \mathbf{s}_l^H \mathbf{s}_{l'}\right) = 0.$$

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.*,

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

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

$$\mathbf{I}_{\boldsymbol{\psi}} = \frac{2}{\sigma_w^2} \begin{pmatrix} \mathbf{I}_{\boldsymbol{\psi}_1,\boldsymbol{\psi}_1} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \mathbf{I}_{\boldsymbol{\psi}_L,\boldsymbol{\psi}_L} \end{pmatrix},$$

$$\mathbf{I}_{\boldsymbol{\psi}_l,\boldsymbol{\psi}_l} = \begin{pmatrix} I_{\tau_l\tau_l} & & & & \\ & I_{\phi_l\phi_l} & I_{\phi_l\theta_l} & & \\ & I_{\phi_l\theta_l} & I_{\theta\theta_l} & & \\ & & & I_{\alpha_l^R\alpha_l^R} & \\ & & & & I_{\alpha_l^I\alpha_l^I} \end{pmatrix}.$$

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 CRLB of (10) averaged over the receive antennas becomes

$$\text{MSE}(f) \geq \frac{1}{M} \sum_{m=1}^{M} \mathbf{g}_{m,f,\boldsymbol{\psi}}^H \mathbf{I}_{\boldsymbol{\psi}}^{-1} \mathbf{g}_{m,f,\boldsymbol{\psi}}$$

$$= \frac{\sigma_w^2}{2M} \sum_{m=1}^{M} \sum_{l=1}^{L} \mathbf{g}_{m,f,\boldsymbol{\psi}_l}^H \mathbf{I}_{\boldsymbol{\psi}_l,\boldsymbol{\psi}_l}^{-1} \mathbf{g}_{m,f,\boldsymbol{\psi}_l} \quad (23)$$

$$= \frac{\sigma_w^2}{2M} \sum_{m=1}^{M} \sum_{l=1}^{L} \left[\frac{|g_{m,f,\tau_l}|^2}{I_{\tau_l\tau_l}} + \frac{|g_{m,f,\alpha_l^R}|^2}{I_{\alpha_l^R\alpha_l^R}} + \frac{|g_{m,f,\alpha_l^I}|^2}{I_{\alpha_l^I\alpha_l^I}} \right.$$

$$\left. + \begin{pmatrix} g_{m,f,\phi_l}^* & g_{m,f,\theta_l}^* \end{pmatrix} \begin{pmatrix} I_{\phi_l\phi_l} & I_{\phi_l\theta_l} \\ I_{\phi_l\theta_l} & I_{\theta_l\theta_l} \end{pmatrix}^{-1} \begin{pmatrix} g_{m,f,\phi_l} \\ g_{m,f,\theta_l} \end{pmatrix} \right].$$

After some computations, we find that

$$\sum_{m=1}^{M} \sum_{l=1}^{L} \frac{|g_{m,f,\alpha_l^R}|^2}{I_{\alpha_l^R\alpha_l^R}} + \frac{|g_{m,f,\alpha_l^I}|^2}{I_{\alpha_l^I\alpha_l^I}} = \frac{2L}{\|\mathbf{s}\|^2}$$

$$\sum_{m,l} \begin{pmatrix} g_{m,f,\phi_l}^* & g_{m,f,\theta_l}^* \end{pmatrix} \begin{pmatrix} I_{\phi_l\phi_l} & I_{\phi_l\theta_l} \\ I_{\phi_l\theta_l} & I_{\theta_l\theta_l} \end{pmatrix}^{-1} \begin{pmatrix} g_{m,f,\phi_l} \\ g_{m,f,\theta_l} \end{pmatrix} = \frac{2L}{\|\mathbf{s}\|^2}$$

$$\sum_{m=1}^{M} \sum_{l=1}^{L} \frac{|g_{m,f,\tau_l}|^2}{I_{\tau_l\tau_l}} = \frac{L(2\pi f)^2}{\|\dot{\mathbf{s}}\|^2} = \frac{L}{\|\mathbf{s}\|^2} \frac{f^2}{\sigma_F^2},$$

where $\sigma_F^2 = \frac{\|\dot{\mathbf{s}}\|^2}{(2\pi)^2\|\mathbf{s}\|^2}$. Inserting the result of these last equations into (23) and using the definition $E_T = \|\mathbf{s}\|^2$, we find the result of Proposition 1.

REFERENCES

[1] F. Rottenberg, R. Wang, J. Zhang, and A. F. Molisch, "Channel Extrapolation in FDD Massive MIMO: Theoretical Analysis and Numerical Validation," in *accepted for presentation at 2019 IEEE Global Communications Conference*, Waikoloa, HI, USA, Dec 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] Y. Liao, H. Yao, Y. Hua, and C. Li, "CSI Feedback Based on Deep Learning for Massive MIMO Systems," *IEEE Access*, vol. 7, pp. 86 810–86 820, 2019.

[9] X. Li, S. Jin, H. A. Suraweera, J. Hou, and X. Gao, "Statistical 3-D Beamforming for Large-Scale MIMO Downlink Systems Over Rician Fading Channels," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1529–1543, April 2016.

[10] 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.

[11] M. Pun, A. F. Molisch, P. Orlik, and A. Okazaki, "Super-Resolution Blind Channel Modeling," in *2011 IEEE International Conference on Communications (ICC)*, Kyoto, Japan, June 2011, pp. 1–5.

[12] N. Jalden, H. Asplund, and J. Medbo, "Channel extrapolation based on wideband MIMO measurements," in *2012 6th European Conference on Antennas and Propagation (EUCAP)*, Prague, Czech Republic, March 2012, pp. 442–446.

[13] 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.

[14] 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)*, Kalamata, Greece, June 2018, pp. 1–5.

[15] 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.

[16] M. Haardt, M. D. Zoltowski, C. P. Mathews, and J. Nossek, "2D unitary ESPRIT for efficient 2D parameter estimation," in *1995 International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, Detroit, MI, USA, USA, May 1995, pp. 2096–2099 vol.3.

[17] W. Shen, L. Dai, B. Shim, Z. Wang, and R. W. Heath, "Channel Feedback Based on AoD-Adaptive Subspace Codebook in FDD Massive MIMO Systems," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5235–5248, Nov 2018.

[18] S. Qiu, D. Gesbert, D. Chen, and T. Jiang, "A Covariance-Based Hybrid Channel Feedback in FDD Massive MIMO Systems," *arXiv preprint arXiv:1902.05184*, 2019.

[19] 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.

[20] 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.

[21] T. Choi, F. Rottenberg, J. Gomez-Ponze, A. Ramesh, P. Luo, J. Zhang, and A. F. Molisch, "Channel extrapolation for FDD massive MIMO: procedure and experimental results," in *accepted for presentation at 2019 IEEE 90th Vehicular Technology Conference (VTC Fall)*, Honolulu, Hawaii, USA, Sep. 2019.

[22] 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.

[23] 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.

[24] 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.

[25] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *2010 Future Network Mobile Summit*, Florence, Italy, June 2010, pp. 1–10.

[26] J. Vieira, F. Rusek, and F. Tufvesson, "Reciprocity calibration methods for massive MIMO based on antenna coupling," in *2014 IEEE Global Communications Conference*, Austin, TX, USA, Dec 2014, pp. 3708–3712.

[27] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity Calibration for Massive MIMO: Proposal, Modeling, and

Validation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3042–3056, May 2017.

[28] 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.

[29] A. F. Molisch, *Wireless communications*. John Wiley & Sons, 2012, vol. 34.

[30] A. F. Molisch, "Ultra-Wide-Band Propagation Channels," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 353–371, Feb 2009.

[31] 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.

[32] S. M. Kay, "Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1)," *PTR Prentice-Hall, Englewood Cliffs*, 1993.

[33] X. Yin, S. Wang, N. Zhang, and B. Ai, "Scatterer Localization Using Large-Scale Antenna Arrays Based on a Spherical Wave-Front Parametric Model," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6543–6556, Oct 2017.

[34] A. Le Calvez, L. Le Magoarou, and S. Paquelet, "Massive MIMO channel estimation taking into account spherical waves," *arXiv preprint arXiv:1811.05669*, 2018.

[35] A. Richter, "Estimation of radio channel parameters: Models and algorithms." ISLE, 2005.

[36] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The New 5G Radio Access Technology," *IEEE Communications Standards Magazine*, vol. 1, no. 4, pp. 24–30, Dec 2017.

[37] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2018.

[38] L. Sanguinetti, A. Kammoun, and M. Debbah, "Theoretical Performance Limits of Massive MIMO With Uncorrelated Rician Fading Channels," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 1939–1955, March 2019.

[39] Ö. Özdogan, E. Björnson, and E. G. Larsson, "Massive MIMO With Spatially Correlated Rician Fading Channels," *IEEE Transactions on Communications*, vol. 67, no. 5, pp. 3234–3250, May 2019.

[40] 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.

[41] "3GPP TR 36.873 v12.5.0," Tech. Rep., 2017.

Peng Luo received the B.S. degree in electrical engineering from the Rose-Hulman Institute of Technology in 2018. He is studying his M.S. degree in electrical engineering from the University of Southern California and works as a directed research student in the WiDeS lab. He worked on large antenna array pattern measurement and helped to build outdoor massive (large-scale) MIMO channel sounder. His research interests are channel modeling and evaluation using modern and future wireless communication system.

Charlie Jianzhong Zhang (S'00M'03SM'09F'16) received the Ph.D. degree from the University of Wisconsin-Madison, Madison, WI, USA. He is currently a Vice President and the Head of the Standards and Mobility Innovation Lab, Samsung Research America, Richardson, TX, USA, where he leads research, prototyping, and standards for 5G and future multimedia networks. From 2009 to 2013, he was the Vice Chairman of the 3GPP RAN1 working group and led development of LTE and LTE-Advanced technologies such as 3D channel modeling, UL-MIMO, CoMP, carrier aggregation for TD-LTE.

François Rottenberg (S'15-M'18) received the M.Sc. in electrical engineering from the Université catholique de Louvain (UCLouvain), Louvain-la-Neuve, in 2014, and the Ph.D. degree from UCLouvain and Université libre de Bruxelles (ULB), Brussels, in 2018. From September 2018 to August 2019, he was a postdoctoral researcher with the University of Southern California (USC), Los Angeles, U.S.A, leading the massive MIMO research team, in collaboration with Samsung Research America. He is now a postdoctoral researcher/"Chargé de recherche" with UCLouvain and ULB, funded by the Belgian National Science Foundation (FNRS). He participated to various national, European and international projects. Since 2015, he has been a regular visitor of the Centre Tecnolgic Telecomunicacions Catalunya (CTTC), Castelldefels, Spain and National Institute of Information and Communications Technology (NICT), Tokyo, Japan.
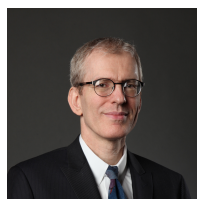
Thomas Choi is a Ph.D. student studying electrical and computer engineering at Wireless Devices and Systems (WiDeS) group in University of Southern California (USC). He received a B.S. degree in electrical engineering at USC in 2015 followed by a M.S. degree in aerospace engineering at Georgia Tech in 2017. His current research interests are evaluating and modeling real world wireless channels accurately for modern and future wireless systems through channel sounder developments and channel measurement campaigns.

Andreas F. Molisch (S'89-M'95-SM'00-F'05) received the Dipl. Ing., Ph.D., and habilitation degrees from the Technical University of Vienna, Vienna, Austria, in 1990, 1994, and 1999, respectively. He subsequently was with FTW (Austria), AT&T (Bell) Laboratories Research (USA); Lund University, Sweden, and Mitsubishi Electric Research Labs (USA). He is now a Professor and the Solomon Golomb Andrew and Erna Viterbi Chair at the University of Southern California, Los Angeles, CA, USA.

His current research interests are the measurement and modeling of mobile radio channels, multi-antenna systems, wireless video distribution, ultra-wideband communications and localization, and novel modulation formats. He has authored, coauthored, or edited four books (among them the textbook Wireless Communications, Wiley-IEEE Press), 19 book chapters, more than 250 journal papers, more than 340 conference papers, as well as more than 80 patents and 70 standards contributions.

Dr. Molisch has been an Editor of a number of journals and special issues, General Chair, Technical Program Committee Chair, or Symposium Chair of multiple international conferences, as well as Chairman of various international standardization groups. He is a Fellow of the National Academy of Inventors, Fellow of the AAAS, Fellow of the IET, an IEEE Distinguished Lecturer, and a Member of the Austrian Academy of Sciences. He has received numerous awards, among them the Donald Fink Prize of the IEEE, the IET Achievement Medal, the Armstrong Achievement Award of the IEEE Communications Society, and the Eric Sumner Award of the IEEE.