Linear Interference Alignment in Full-Duplex MIMO Networks with Imperfect CSI

Linear Interference Alignment in Full-Duplex MIMO Networks with Imperfect CSI

Paula Aquilina, *Student Member, IEEE* and Tharmalingam Ratnarajah, *Senior Member, IEEE*

***Abstract*—In this paper we consider a system where full-duplex (FD) base-stations (BSs) communicate with half-duplex (HD) downlink (DL) and uplink (UL) users in a multi-user multi-cell network, where all nodes are equipped with multiple antennas. The introduction of FD BSs offers potential to increase spectral efficiency, however it also causes a surge in the number of interference links compared to the HD network counterpart. Here, we apply linear interference alignment (IA) to manage interference in this network under imperfect channel state information (CSI). Firstly, we characterize the performance losses incurred with respect to the achievable sum rate and degrees of freedom (DoF). Results show that the general trend in performance loss is mainly determined by how the error scales with the signal-to-noise ratio (SNR). In particular, full UL and DL DoF can be achieved even under imperfect CSI when the channel error is at least inversely proportional to SNR. Moreover, in such cases the sum rate loss is always finite, and either goes to zero or is upper bounded by a derived value. Secondly, we design two linear IA algorithms applicable to the system under consideration. These are based on minimizing the mean square error (MMSE) and maximizing the signal-to-interference-plus-noise ratio (Max-SINR), and take into account statistical knowledge of the CSI error for added robustness. The proposed algorithms follow specific design principles that distribute the different interference components amongst the various beamformers and result in unitary receivers and precoders. Additionally, we show that under certain conditions both designs result in identical beamforming solutions, even though the MMSE algorithm has lower computational complexity. Thirdly, we also derive the proper condition for IA feasibility in the multi-cell system under consideration.**

***Index Terms*—Beamformer design, degrees of freedom, full-duplex, imperfect CSI, interference alignment, Max-SINR, MMSE, sum rate.**

I. Introduction

THE demand for wireless network resources is constantly on the rise, pushing for the design of new technologies that are able to handle unprecedented rates. One such technology is full-duplex (FD) communication. Whilst traditional half-duplex (HD) networks require separate time or frequency resources for downlink (DL) and uplink (UL) communication, FD caters for simultaneous DL and UL transmission, making it an attractive candidate solution to the ever growing spectrum demand problem.

FD operation was traditionally considered to be infeasible due to self-interference (SI) where power from DL transmission interferes with UL received signals at FD nodes. However, recent years have seen numerous breakthroughs in hardware design. By combining three types of SI cancellation techniques, namely, (a) propagation domain SI suppression, (b) analog circuit domain SI cancellation, and (c) digital circuit domain SI cancellation, it is now possible to suppress significant levels of SI. For example, [2] combines signal inversion and digital domain techniques to achieve 73 dB of SI suppression for a 10 MHz orthogonal frequency division multiplexed signal. In [3] all three classes of SI suppression methods are combined to achieve an average cancellation of 85 dB over a 20 MHz signal. Additionally, [4] proposes a single-antenna design that cancels up to 110 dB of SI over an 80 MHz bandwidth.

The promise of increased spectral efficiency has motivated a wide range of research into FD communication and its possible uses. Moreover, the new found ability to mitigate SI up to acceptable levels has brought to light new challenges that need to be addressed for the practical realization of FD networks. The most significant one is the issue of efficient interference management [5], [6], since the application of FD nodes brings along a huge surge in the amount of interference present in the network.

Consider for example the G-cell system in Fig. 1, where each cell has an FD BS serving one DL and one UL user per cell. In this network, for UL communication base-stations (BSs) have additional SI and BS-to-BS interference, while DL users have additional co-channel interference (CCI) from UL users both from the same cell and from other cells.

There have been a number of information-theoretic studies with the aim of understanding the fundamental capacity limits of FD systems, particularly the characterization of achievable degrees of freedom (DoF), as a first order characterization of the achievable rate. These DoF studies exploit interference alignment (IA) [7] in order to maximize capacity. Within this context, [8] studies the DoF region for single-cell systems with one multi-antenna FD BS and K single antenna FD users, and proposes an achievable scheme based on ergodic IA. The DoF region for an FD BS communicating with HD users [9], [10] and a point-to-point multiple-input multiple-output (MIMO) link [10] have also been studied, with the authors proposing achievable schemes based on asymptotic IA in each case. Additionally, in [11] the authors consider systems where an FD BS communicates with either FD or HD users; for each scenario the sum DoF are characterized and achievable schemes based on a combination of interference nulling and asymptotic IA are proposed.

While the ergodic and asymptotic IA techniques exploited in the literature mentioned so far are highly beneficial from a

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/L025299/1. Part of this work was presented at ICC 2017 [1].

P. Aquilina and T. Ratnarajah are with the Institute of Digital Communications, School of Engineering, The University of Edinburgh, Alexander Graham Bell Building, Kings Buildings, Mayfield Road, Edinburgh, EH9 3JL, United Kingdom. (email: {p.aquilina, t.ratnarajah}@ed.ac.uk).

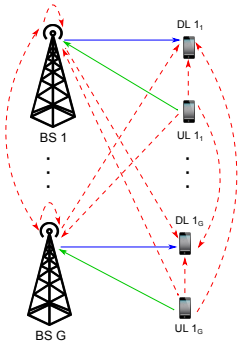

Fig. 1: G-cell network with an FD BS, and one DL and one UL user per cell. Solid lines represent desired links, dashed ones represent interference links.

theoretical standpoint, they are rather difficult to implement in practice. Ergodic IA requires the occurrence of complementary channels to cancel interference, whilst asymptotic IA requires an infinitely large amount of time extensions. Therefore, other works focus on different types of IA which can be realized more easily. For example, [12] considers the application of blind IA in a single-cell single-input single-output (SISO) system with an FD BS and HD users, and reconfigurable antennas that are able to generate a pre-determined set of radiation patterns. Additionally, [13] studies the feasibility of linear IA in single-cell MIMO systems with an FD BS and HD users. Linear IA creates a set of beamformers that simultaneously cancel all unwanted signals and leave the desired signals untouched, and is particularly attractive from a practical perspective because it provides a one-shot solution to the interference management problem and does not require the use of special equipment [7].

The merging of FD operation along with IA offers a highly promising dual approach solution to cope with the increasing spectrum demand problem, making it a very relevant research direction. However, the aforementioned literature assumes that the available channel state information (CSI) is perfect and considers only a single-cell context. In practice, CSI is likely to be imperfect, and real world implementation involves multiple cells. Motivated by these two challenges, and the appealing qualities of linear IA, here we consider a multi-cell system with an FD BS in each cell serving multiple legacy HD users, and apply linear IA to manage interference under imperfect CSI conditions. The combined FD BS / HD user system is chosen as opposed to a fully FD enabled one, because the shift from HD to FD is expected to involve significant expenditure, making it more practically relevant to initially consider scenarios where only the infrastructure elements (i.e. BSs) are upgraded to FD with user devices still operating in HD. The use of IA in such multi-cell FD systems has so far only been considered in [14], where the authors derive a scaling law for the multiplexing gain of FD over HD under a perfect CSI assumption, and with network MIMO capability between the BSs (i.e. assuming no BS-to-BS interference).

Our work aims to characterize the performance of IA in FD systems with imperfect CSI. Such an effect has already been considered for HD networks, see for example [15]–[18] and references therein, but is yet unexplored for FD ones. The CSI error model applied allows us to represent the CSI both as a function of the signal-to-noise ratio (SNR) or as completely independent of it and can represent different channels scenarios, including both reciprocal and non-reciprocal channels. Based on this error model, we characterize the mean asymptotic sum rate loss and the loss in achievable DoF. Results show that the general performance trend is highly dependent on how the error scales with SNR. Both losses go to zero under certain conditions, implying that the effect of the CSI error is negligible in such cases, and are otherwise quantified in terms of the system configuration and the CSI error parameters.

Next, we shift our focus to linear IA algorithms. No such methods are available for FD systems in literature so far, here we propose two novel techniques applicable to the FD network under consideration. Our algorithms are inspired by methods initially proposed for HD interference channels (ICs), namely, the minimum mean square error (MMSE) method from [19] and the maximum signal-to-interference-plus-noise ratio (Max-SINR) technique from [20]. These algorithms are not straightforward extensions of the original HD ones; (a) they separate the various interference components amongst the different available beamformers, (b) they exploit statistical knowlege of the CSI error resulting in a more robust design, and (c) they produce unitary beamformers. The algorithms are first derived for the single-cell case, since the feasibility of linear IA in such networks is already known [13], and later extended to the multi-cell case. For the multi-cell case, we also derive the proper condition for IA feasibility. This condition, along with the multi-cell version of the algorithms, can serve as tools to help future efforts into the determination of a full set of linear IA feasibility conditions for multi-cell FD enabled systems.

The remaining sections of the paper are organized as follows. Section II presents some preliminaries, namely, the system model, the CSI error model applied and the IA conditions. Section III presents two theorems that characterize the performance loss due to imperfect CSI. Next, in Section IV, we derive the MMSE and Max-SINR based algorithms for the single-cell context, and establish an equivalence between the two. Section V deals with multiple cell considerations, here we derive a proper condition for multi-cell systems and also present the multi-cell extensions for the IA algorithms. Section VI presents simulation results, and finally conclusions are provided in Section VII. Moreover, there are three appendices; the first two contain proofs for the derived theorems, while the last one provides some useful lemmas.

Notation: Scalars are represented using lower case standard font, vectors are represented using lower case bold font and matrices are represented using upper case bold font. $|\cdot|$, $\|\cdot\|$ and $\mathrm{Tr}(\cdot)$ denote the absolute value, the Euclidean norm and the trace respectively. $\mathbb{E}\{\mathbf{A}\}$ represents the expected value of $\mathbf{A}$. $\mathbb{QR}(\mathbf{A})$ takes the unitary part of the QR-decomposition of $\mathbf{A}$.

II. Preliminaries

A. System model

We consider a scenario having G cells, where each cell g has one FD BS, K_d DL users requiring b_d streams each and K_u UL users requiring b_u streams each. BSs are equipped

with M_B FD antennas, DL users are equipped with M_d HD antennas and UL users are equipped with M_u HD antennas. The received signal at the kth DL user in cell g and at BS g are given by (1) and (2) respectively.

$$\mathbf{y}_{k_g^d} = \sum_{j=1}^{G}\sum_{i=1}^{K_d} \kappa_{k_g^d,j}\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} + \sum_{j=1}^{G}\sum_{i=1}^{K_u} \kappa_{k_g^d,i_j^u}\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u} + \mathbf{n}_{k_g^d} \tag{1}$$

$$\mathbf{y}_g = \sum_{j=1}^{G}\sum_{i=1}^{K_u} \kappa_{g,i_j^u}\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u} + \sum_{j=1}^{G}\sum_{i=1}^{K_d} \kappa_{g,j}\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} + \mathbf{n}_g \tag{2}$$

Here, $\mathbf{H}_{k_g^d,j} \in \mathbb{C}^{M_d \times M_B}$ represents the channel going from BS j to DL user k_g^d, $\mathbf{H}_{k_g^d,i_j^u} \in \mathbb{C}^{M_d \times M_u}$ is the channel going from UL user i_j^u to DL user k_g^d, $\mathbf{H}_{g,j} \in \mathbb{C}^{M_B \times M_B}$ is the channel going from BS j to BS g and $\mathbf{H}_{g,i_j^u} \in \mathbb{C}^{M_B \times M_u}$ is the channel going from UL user i_j^u to BS g. All channel elements are distributed as $\mathcal{CN}(0,1)$. $\mathbf{V}_{i_j^d} \in \mathbb{C}^{M_B \times b_d}$ is the precoder for $\mathbf{s}_{i_j^d}$, with $\mathbf{s}_{i_j^d} \in \mathbb{C}^{b_d \times 1}$ being the data intended for the ith DL user in cell j, such that $\mathbb{E}\{\mathbf{s}_{i_j^d}\mathbf{s}_{i_j^d}^H\} = P\mathbf{I}$. $\mathbf{V}_{i_j^u} \in \mathbb{C}^{M_u \times b_u}$ is the precoder for $\mathbf{s}_{i_j^u} \in \mathbb{C}^{b_u \times 1}$, with $\mathbf{s}_{i_j^u}$ being the data transmitted by the ith UL user in cell j, such that $\mathbb{E}\{\mathbf{s}_{i_j^u}\mathbf{s}_{i_j^u}^H\} = P\mathbf{I}$. Moreover, $\mathbf{n}_{k_g^d}$ and $\mathbf{n}_g$ represent the noise with elements drawn from $\mathcal{CN}(0,\sigma^2)$. Additionally, $\kappa_{x,y}$ represents the pathloss. We adopt the generic pathloss model from [21] such that $\kappa_{x,y} = \psi r_{x,y}^{\mu}$, where ψ is the pathloss at unit distance, $r_{x,y}$ is the distance between nodes x and y, and μ is the pathloss exponent.

The estimated DL and UL received signals are given by

$$\begin{aligned}\widehat{\mathbf{s}}_{k_g^d} = &\sum_{j=1}^{G}\sum_{i=1}^{K_d} \kappa_{k_g^d,j}\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} + \mathbf{U}_{k_g^d}^H\mathbf{n}_{k_g^d} \\ &+ \sum_{j=1}^{G}\sum_{i=1}^{K_u} \kappa_{k_g^d,i_j^u}\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u}\end{aligned} \tag{3}$$

$$\begin{aligned}\widehat{\mathbf{s}}_{k_g^u} = &\sum_{j=1}^{G}\sum_{i=1}^{K_u} \kappa_{g,i_j^u}\mathbf{U}_{k_g^u}^H\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{s}_{i_j^u} + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_d} \kappa_{g,j}\mathbf{U}_{k_g^u}^H\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{s}_{i_j^d} \\ &+ \mathbf{U}_{k_g^u}^H\mathbf{n}_g + \underbrace{\Theta\sum_{i=1}^{K_d} \kappa_{g,g}\mathbf{U}_{k_g^u}^H\boldsymbol{\Upsilon}_{g,g}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d}}_{\substack{\text{residual SI for imperfect}\\ \text{SI cancellation scenarios}}}\end{aligned} \tag{4}$$

where $\mathbf{U}_{k_g^d} \in \mathbb{C}^{M_d \times b_d}$ is the receive beamformer applied at DL user k_g^d, $\mathbf{U}_{k_g^u} \in \mathbb{C}^{M_B \times b_u}$ is the receive beamformer applied at BS g to extract the data transmitted by UL user k_g^u and Υ represents the auxiliary error matrix which will be discussed in further detail later in Section II-B. Here (3) is obtained as $\mathbf{U}_{k_g^d}^H\mathbf{y}_{k_g^d}$. Additionally, for the perfect CSI case $\mathbf{H}_{g,g}\sum_{i=1}^{K_d}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d}$ is known at BS g, thus we obtain (4) with $\Theta = 0$ as $\mathbf{U}_{k_g^u}^H(\mathbf{y}_g - \mathbf{H}_{g,g}\sum_{i=1}^{K_d}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d})$. The parameter Θ is a binary term used to differentiate between perfect SI cancellation and imperfect SI cancellation. For perfect CSI, similar to other FD DoF studies [8]–[14], we assume that SI is always perfectly canceled, therefore $\Theta = 0$. For imperfect CSI, perfect SI cancellation is not guaranteed leading to a residual SI term, further details are provided in Section II-B.

Similar to prior literature dealing with interference management we require knowledge of the CSI. While going into the actual details of the CSI acquirement process is not within our scope, it is important to note that all the required channels can indeed be learned. Channels between users and BSs and between the BSs themselves can be estimated through standard 3GPP LTE channel estimation protocols applied in current HD systems. Additionally, channels between users can be learned via neighbour discovery methods applicable to device-to-device (D2D) communication, such as sounding reference signals (SRS) in 3GPP LTE. (See for example [22]–[24].)

B. Imperfect CSI considerations

The imperfect CSI is modeled as

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{E} \tag{5}$$

where $\widehat{\mathbf{H}}$ is the available CSI, $\mathbf{H}$ is the perfect channel, and $\mathbf{E}$ is the channel estimation error. Here, $\mathbf{E}$ is independent of $\mathbf{H}$ and modeled as $\mathbf{E} \sim \mathcal{CN}(0, \eta\mathbf{I})$, where $\eta = \beta\rho^{-\alpha}$ with $\rho = \frac{P}{\sigma^2}$ representing the nominal SNR.

The introduction of parameters α and β renders our CSI error model highly versatile, allowing us to represent the error as either dependent on or independent of SNR as required. Thus, for any $\alpha \geq 0$ and $\beta > 0$, η can be used to capture a variety of CSI acquisition scenarios. Of particular interest are the following instances.

- *Perfect CSI*: As $\alpha \to \infty$, perfect CSI is obtained for $\rho \geq 1$.
- *Reciprocal channels*: In reciprocal systems, channels are assumed to be identical regardless of the direction of communication. Thus, the CSI error is dependent on the ratio of the pilot power to the noise level at the pilot receiving node, i.e. it is inversely proportional to SNR. Therefore, the CSI error can be modeled by setting $\alpha = 1$.
- *Non-reciprocal channels*: In non-reciprocal systems, the channel information is different for different directions of communication. Thus, acquired CSI needs to be fedback over a dedicated feedback link. Transmissions over this link are subject to quantization errors, resulting in CSI error that is independent of SNR. This can therefore be modeled by setting $\alpha = 0$.

Further details on how to set α and β to represent different scenarios can be found in [18], [25] and references therein.

Alternatively, η as a whole can be interpreted as a single parameter that encapsulates the CSI quality. Its value may be assumed to be known *a priori*, and can be determined depending on the channel dynamics and the CSI estimation schemes applied. For additional details refer to [26] and reference within.

For our analysis we need the statistical properties of $\mathbf{H}$ conditioned $\widehat{\mathbf{H}}$. Knowing that $\mathbf{E}$ and $\mathbf{H}$ are independent, then $\widehat{\mathbf{H}}$ and $\mathbf{H}$ are jointly Gaussian. Thus we can express the perfect channel, $\mathbf{H}$, as a function of the available CSI, $\widehat{\mathbf{H}}$, and an auxiliary error matrix, $\boldsymbol{\Upsilon}$, as [27]

$$\mathbf{H} = \frac{1}{1+\eta}\widehat{\mathbf{H}} + \boldsymbol{\Upsilon} \tag{6}$$

$$\widehat{J}_{k_g^d} = \sum_{\substack{j=1 \\ (j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \kappa^2_{k_g^d,j} P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,j}\mathbf{v}_{i_j^d}^{m}|^2 + \sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} \kappa^2_{k_g^d,i_j^u} P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,i_j^u}\mathbf{v}_{i_j^u}^{m}|^2 \tag{8}$$

$$\widehat{J}_{k_g^u} = \sum_{\substack{j=1 \\ (j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} \kappa^2_{g,i_j^u} P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,i_j^u}\mathbf{v}_{i_j^u}^{m}|^2 + \sum_{\substack{j=1 \\ j\neq g}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \kappa^2_{g,j} P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,j}\mathbf{v}_{i_j^d}^{m}|^2 + \Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \kappa^2_{g,g} P|\mathbf{u}_{k_g^u}^{n\,H}\boldsymbol{\Upsilon}_{g,g}\mathbf{v}_{i_g^d}^{m}|^2 \tag{9}$$

$$\begin{aligned}\Delta R_{\text{ FD}} \leq & \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{1+\eta}\left(\sum_{\substack{j=1 \\ (j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{k_g^d,j}+\sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\kappa^2_{k_g^d,i_j^u}\right)\right) \\ & + \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{1+\eta}\left(\sum_{\substack{j=1 \\ (j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\kappa^2_{g,i_j^u}+\sum_{\substack{j=1 \\ j\neq g}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{g,j}+\Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{g,g}\right)\right)\end{aligned} \tag{11}$$

where $\text{vec}(\boldsymbol{\Upsilon}) \sim \mathcal{CN}\left(0, \frac{\eta}{1+\eta}\mathbf{I}\right)$ and is independent of $\widehat{\mathbf{H}}$.

Generally for theoretical studies perfect SI cancellation is assumed [8]–[14]. This can also be applied to our scenario, where one can assume that each BS has perfect knowledge of its SI channel and imperfect CSI for the remaining channels. However, one may also consider the case where the SI channel is also known imperfectly. For such situations $\widehat{\mathbf{H}}_{g,g}$ is available at BS g instead of $\mathbf{H}_{g,g}$, thus only $\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,g}\sum_{i=1}^{K_d}\mathbf{V}_{i_g^d}\mathbf{s}_{i_g^d}$ can be subtracted from (2). This results in the estimated UL data, $\widehat{\mathbf{s}}_{k_g^u}$, being given by (4) with $\Theta = 1$.

C. IA conditions

Provided that the system configuration (i.e. the combination of the number of cells, users, streams and antennas at the various nodes) is such that IA is feasible under perfect CSI, then the following IA conditions are observed

$$\begin{aligned}
|\mathbf{u}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}| &> 0 \quad \forall\, n,k,g \\
|\mathbf{u}_{k_g^u}^{n\,H}\widehat{\mathbf{H}}_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}| &> 0 \quad \forall\, n,k,g \\
\mathbf{u}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,j}\mathbf{v}_{i_j^d}^{m} &= 0 \quad \forall\, n,m,k,i,g,j,(n,k,g\neq m,i,j) \\
\mathbf{u}_{k_g^u}^{n\,H}\widehat{\mathbf{H}}_{g,j}\mathbf{v}_{i_j^d}^{m} &= 0 \quad \forall\, n,m,k,i,g,j,(g\neq j) \\
\mathbf{u}_{k_g^d}^{n\,H}\widehat{\mathbf{H}}_{k_g^d,i_j^u}\mathbf{v}_{i_j^u}^{m} &= 0 \quad \forall\, n,m,k,i,g,j \\
\mathbf{u}_{k_g^u}^{n\,H}\widehat{\mathbf{H}}_{g,i_j^u}\mathbf{v}_{i_j^u}^{m} &= 0 \quad \forall\, n,m,k,i,g,j,(n,k,g\neq m,i,j)
\end{aligned} \tag{7}$$

where $\mathbf{u}_{k_g^d}^{n}$, $\mathbf{u}_{k_g^u}^{n}$, $\mathbf{v}_{k_g^d}^{n}$ and $\mathbf{v}_{k_g^u}^{n}$ refer to the nth column of beamformers $\mathbf{U}_{k_g^d}$, $\mathbf{U}_{k_g^u}$, $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$ respectively.

Note that we use the term $\widehat{\mathbf{H}}$ in (7) since this represents the CSI available for beamformer calculation. For perfect CSI the IA conditions observed are (7) with $\widehat{\mathbf{H}}$ replaced by $\mathbf{H}$.

III. Performance losses due to imperfect CSI

With perfect CSI, $\eta = 0$ and $\widehat{\mathbf{H}} = \mathbf{H}$, thus satisfying the IA conditions in (7) results in perfect interference cancellation. However, when the available CSI is imperfect, not all interference is canceled leading to a significant amount of interference leakage, given by (8) for the DL and (9) for the UL. Residual leakage has an adverse effect on achievable sum rate and DoF, understanding its extent is a fundamental aspect towards obtaining a more realistic characterization of practical system performance. Here, we focus on the losses incurred in terms of achievable sum rate and DoF.

A. Sum rate loss

For i.i.d. Gaussian inputs, the network achievable rate under imperfect CSI is given by

$$\begin{aligned}
&\widehat{R}_{\text{ FD TOT}} \\
&= \widehat{R}_{\text{ FD DL}} + \widehat{R}_{\text{ FD UL}} \\
&= \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\kappa^2_{k_g^d,g}P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^{n}|^2}{\widehat{J}_{k_g^d}+\sigma^2}\right) \\
&\quad + \sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\kappa^2_{g,k_g^u}P|\mathbf{u}_{k_g^u}^{n\,H}\kappa^2_{g,k_g^u}\mathbf{v}_{k_g^u}^{n}|^2}{\widehat{J}_{k_g^u}+\sigma^2}\right).
\end{aligned} \tag{10}$$

For perfect CSI there is no interference leakage, thus the achievable rate is equal to $R_{\text{ FD TOT}} = R_{\text{ FD DL}} + R_{\text{ FD UL}}$, given by (10) with $\widehat{J}_{k_g^d} = \widehat{J}_{k_g^u} = 0$.

The sum rate loss due to imperfect CSI, $\Delta R_{\text{ FD}}$, is defined as the expected value of the difference between $R_{\text{ FD TOT}}$ and $\widehat{R}_{\text{ FD TOT}}$, and is characterized as shown in the following theorem.

Theorem 1. *Consider a G-cell system where each cell has one FD BS, K_d DL users requiring b_d streams each and K_u UL users requiring b_u streams each. For this system, under imperfect CSI with zero mean and error variance $\eta = \beta\rho^{-\alpha}$, the rate loss can be characterized as in (11).*

Under a homogeneous pathloss assumption, this can be further expressed as:

$$\lim_{SNR\to\infty} \Delta\breve{R}_{FD} \begin{cases} = 0 & \alpha > 1 \\ \leq \Omega & \alpha = 1 \\ \to \infty & 0 \leq \alpha < 1 \end{cases} \tag{12}$$

where

$$\Omega = GK_d b_d \Big(\log_2\big(1+\beta(GK_d b_d + GK_u b_u - 1)\big)\Big) + GK_u b_u \Big(\log_2\big(1+\beta(GK_u b_u - 1 + (G-1+\Theta)K_d b_d)\big)\Big).$$

Proof. See Appendix A. □

B. DoF loss

The DoF achievable under imperfect CSI are given by

$$\widehat{D}_{\text{FD TOT}} = \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}_{\text{FD DL}}\}\}}{\log_2 P} + \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{R}_{\text{FD UL}}\}\}}{\log_2 P}. \tag{13}$$

For perfect CSI there is no interference leakage (i.e. $\widehat{J}_{k_g^d} = \widehat{J}_{k_g^u} = 0$), thus the achievable DOF, $D_{\text{FD TOT}}$, is defined as (13) with $\widehat{R}_{\text{FD DL}}$ and $\widehat{R}_{\text{FD UL}}$ replaced by $R_{\text{FD DL}}$ and $R_{\text{FD UL}}$ respectively. Additionally, for feasible IA, $D_{\text{FD TOT}}$ can be further represented as

$$D_{\text{FD TOT}} = G(K_d b_d + K_u bu)\,. \tag{14}$$

The DoF loss, ΔD_{FD}, is the difference between $D_{\text{FD TOT}}$ and $\widehat{D}_{\text{FD TOT}}$, and is characterized as shown in the following theorem.

Theorem 2. *Consider a G-cell system where each cell has one FD BS, K_d DL users requiring b_d streams each and K_u UL users requiring b_u streams each. For this system, under imperfect CSI with zero mean and error variance $\eta = \beta\rho^{-\alpha}$:*

$$\Delta D_{FD} = \begin{cases} 0 & \alpha \geq 1 \\ (1-\alpha)\,G\,(K_d b_d + K_u b_u) & 0 \leq \alpha < 1\,. \end{cases} \tag{15}$$

Proof. See Appendix B. □

Remark 1. *The results of Theorems 1 and 2 show that the way the error scales with SNR, as reflected by the α parameter, is highly important in determining the general performance trend of IA, both from a rate and DoF perspective. The implications of the two theorems are intrinsically related. For example, in the range of $\alpha \geq 1$ Theorem 1 states that the sum rate loss is either zero or finite. This is also reflected in Theorem 2, where no DoF loss is expected within the same α range. The fact that the rate loss is zero only for $\alpha > 1$, while the DoF loss is zero for $\alpha \geq 1$, shows that while $\alpha = 1$ corresponds to perfect CSI from a DoF perspective, it does not correspond to perfect CSI from a more practical rate perspective. In the range of $0 \leq \alpha < 1$, Theorem 2 shows that a DoF loss is inevitable. This is also reflected in Theorem 1, where for the same α range the rate loss increases unboundedly with SNR.*

IV. Linear IA algorithms

While the bounds derived so far provide an understanding of the expected behavior of linear IA within the system model considered, it is also necessary to have algorithms that work within this context. Such algorithms are not yet available in literature for systems with FD BSs and HD users, therefore here we propose two different approaches: (a) an MMSE based solution, and (b) a Max-SINR based one.

These proposed algorithms are not straightforward extensions of the original HD ones from [19] and [20]; (a) they separate the various interference components amongst the different available beamformers rather than treating all interference equivalently, i.e. they are based on design principles that are specifically catered to the new system model, (b) they exploit statistical knowledge of the CSI error to provide added robustness, and (c) they result in unitary beamformers. The use of unitary beamformers has gained significant attention in recent years due to its role in codebook design for limited feedback scenarios. It has been selected for both single-user and multi-user mode operation for evolved universal terrestrial radio access [28], with advantages that include added simplicity of application and improved robustness to channel estimation errors [29]. Additionally, it has been shown to lower complexity for MMSE based algorithms [30] by avoiding the need for an extra linear search to enforce transmit power constraints when generating precoders, and also improve performance for Max-SINR based ones in multi-stream applications [31]. Note that linear IA aims to cancel out interference by ensuring that the desired signal subspace is separate from the subspace occupied by the interference and noise. Applying a QR decomposition to make the beamformers unitary has no effect on their subspace, thus the signal and interference-plus-noise are separable using both the unitary and non-unitary versions; implying that both versions observe the conditions in (7) and achieve IA.

The interference that needs to be handled by our IA solutions can be classified into four main categories:

(i) *Intra-DL interference* - interference caused by undesired DL data for other users in the same cell;
(ii) *Intra-UL interference* - interference caused by undesired UL data for other users in the same cell;
(iii) *CCI-OC* - co-channel interference caused by nodes located in other cells (includes both DL and UL data);
(iv) *R-SI* - residual self-interference at the BSs due to imperfect CSI knowledge.

While it is possible to create beamformers that handle all the interference jointly, prior results for HD systems [33], [34] indicate that this approach is not suited to interference scenarios that are more complex than the initially studied HD IC. Similar behavior has also been noted for our FD system, thus we base our IA algorithms on specific design principles (see Design Principle 1 and Design Principle 2 outlined later on).

Focus in this section will be on the derivation of the algorithms for a single-cell system, since the feasibility of such configurations has already been explored in current literature [13], and also due to the relevant compactness of the corresponding expressions in comparison to multi-cell ones.

$$F_{k^u} = \mathbb{E}\left\{\|\mathbf{f}_{k^u} - \mathbf{s}_{k^u}\|^2\right\} = \mathbb{E}\left\{\mathrm{Tr}\left(\left(\mathbf{U}^H \sum_{i=1}^{K_u} \kappa_{g,i^u}\mathbf{H}_{g,i^u}\mathbf{V}_{i^u}\mathbf{s}_{i^u} + \Theta\mathbf{U}_{k^u}^H \sum_{i=1}^{K_d} \kappa_{g,g}\boldsymbol{\Upsilon}_{g,g}\mathbf{V}_{i^d}\mathbf{s}_{i^d} + \mathbf{U}_{k^u}^H\mathbf{n}_g - \mathbf{s}_{k^u}\right)\right.\right.$$
$$\left.\left.\left(\mathbf{U}_{k^u}^H \sum_{i=1}^{K_u} \kappa_{g,i^u}\mathbf{H}_{g,i^u}\mathbf{V}_{i^u}\mathbf{s}_{i^u} + \Theta\mathbf{U}_{k^u}^H \sum_{i=1}^{K_d} \kappa_{g,g}\boldsymbol{\Upsilon}_{g,g}\mathbf{V}_{i^d}\mathbf{s}_{i^d} + \mathbf{U}_{k^u}^H\mathbf{n}_g - \mathbf{s}_{k^u}\right)^H\right)\right\}$$
$$\overset{(a)}{=} \mathrm{Tr}\Bigg(P\mathbf{U}_{k^u}^H \sum_{i=1}^{K_u} \kappa_{g,i^u}^2\mathbf{H}_{g,i^u}\mathbf{V}_{i^u}\mathbf{V}_{i^u}^H\mathbf{H}_{g,i^u}^H\mathbf{U}_{k^u} + \sigma^2\mathbf{U}_{k^u}^H\mathbf{U}_{k^u} - P\kappa_{g,k^u}\mathbf{U}_{k^u}^H\mathbf{H}_{g,k^u}\mathbf{V}_{k^u}$$
$$- P\kappa_{g,k^u}\mathbf{V}_{k^u}^H\mathbf{H}_{g,k^u}^H\mathbf{U}_{k^u} + \Theta P\mathbf{U}_{k^u}^H \sum_{i=1}^{K_d} \kappa_{g,g}^2\boldsymbol{\Upsilon}_{g,g}\mathbf{V}_{i^d}\mathbf{V}_{i^d}^H\boldsymbol{\Upsilon}_{g,g}^H\mathbf{U}_{k^u}\Bigg) + Pb_u \tag{16}$$

$$\frac{\partial F_{k^u}}{\partial \mathbf{U}_{k^u}} = P\mathbf{U}_{k^u}^H \sum_{i=1}^{K_u} \kappa_{g,i^u}^2 \left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,i^u} + \boldsymbol{\Upsilon}_{g,i^u}\right)\mathbf{V}_{i^u}\mathbf{V}_{i^u}^H\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,i^u} + \boldsymbol{\Upsilon}_{g,i^u}\right)^H$$
$$+ \sigma^2\mathbf{U}_{k^u}^H - P\kappa_{g,k^u}\mathbf{V}_{k^u}^H\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{g,k^u} + \boldsymbol{\Upsilon}_{g,k^u}\right)^H + \Theta P\mathbf{U}_{k^u}^H \sum_{i=1}^{K_d} \kappa_{g,g}^2\boldsymbol{\Upsilon}_{g,g}\mathbf{V}_{i^d}\mathbf{V}_{i^d}^H\boldsymbol{\Upsilon}_{g,g}^H\,. \tag{17}$$

The results for the multi-cell extensions are presented later in Section V-B. Note that when considering the single-cell case, where by definition $G = 1$, we drop the index g to indicate which cell a user belongs to, i.e. we use k^u to indicate the kth UL user in the cell and k^d to represent the kth DL user. However, the notation g is still used in channel and pathloss related indices to represent the BS.

The single-cell versions of the algorithms follow Design Principle 1 outlined below.

Design Principle 1. *Intra-UL interference is only handled by the receivers. Intra-DL interference is only handled by the precoders. R-SI is handled by both the transmit and receive beamformers at the BS.*

A. MMSE based design for single-cell systems

This algorithm focuses on minimizing the mean squared error, and designs beamformers which aim to find a balance between aligning the interference and ensuring that the signal level is suitably above noise. It was originally proposed for the IC with perfect CSI and a single-stream per user in [19], and later generalized to the multi-stream case in [32]. The designs in [19], [32] carry out a separate linear search (using techniques such as for example the bisection method) to enforce transmit power constraints for each of the precoders generated. The added computational costs incurred by such numerical searches can be avoided by ensuring that the beamformers produced are unitary [30]. Our MMSE design incorporates this lower complexity feature, and produces unitary beamformers via the inclusion of QR decomposition stages (see Steps 4 and 6 in Algorithm 1).

Starting with UL communication in the intended direction, with fixed $\mathbf{V}$ and in accordance to Design Principle 1, the optimization problem to find the BS receivers, $\mathbf{U}_{k^u}$, is given by

$$\min_{\mathbf{U}_{k^u}} \mathbb{E}\left\{\|\mathbf{f}_{k^u} - \mathbf{s}_{k^u}\|^2\right\} \;\forall\; k$$

where

$$\mathbf{f}_{k^u} = \mathbf{U}_{k^u}^H \sum_{i=1}^{K_u} \kappa_{g,i^u}\mathbf{H}_{g,i^u}\mathbf{V}_{i^u}\mathbf{s}_{i^u} + \Theta\mathbf{U}_{k^u}^H \sum_{i=1}^{K_d} \kappa_{g,g}\boldsymbol{\Upsilon}_{g,g}\mathbf{V}_{i^d}\mathbf{s}_{i^d} + \mathbf{U}_{k^u}^H\mathbf{n}_g.$$

The optimization function can be defined as in (16) where (a) follows since the transmitted data consists of i.i.d symbols, allowing us to use $\mathbb{E}\left\{\mathbf{s}_{k^u}\mathbf{s}_{i^u}^H\right\} = \mathbb{E}\left\{\mathbf{s}_{k^d}\mathbf{s}_{i^d}^H\right\} = \mathbf{0}\;\forall\, k, i, (k \neq i)$, $\mathbb{E}\left\{\mathbf{s}_{k^u}\mathbf{s}_{i^d}^H\right\} = \mathbb{E}\left\{\mathbf{s}_{k^d}\mathbf{s}_{i^u}^H\right\} = \mathbf{0}\;\forall\, k, i$ and $\mathbb{E}\left\{\mathbf{s}_{k^u}\mathbf{s}_{k^u}^H\right\} = \mathbb{E}\left\{\mathbf{s}_{k^d}\mathbf{s}_{k^d}^H\right\} = P\mathbf{I}\;\forall\; k$. Differentiating with respect to $\mathbf{U}_{k^u}$ and replacing $\mathbf{H}$ by (6) results in (17).

This can be made dependent on the imperfect CSI $\widehat{\mathbf{H}}$ only by using the statistical information we have on the error. Thus, taking expectations with respect to $\boldsymbol{\Upsilon}$, and using Lemmas 1 and 2 from Appendix C we obtain

$$\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\frac{\partial F_{k^u}}{\partial \mathbf{U}_{k^u}}\right\} = \sigma^2\mathbf{U}_{k^u}^H - \frac{P}{(1+\eta)}\kappa_{g,k^u}\mathbf{V}_{k^u}^H\widehat{\mathbf{H}}_{g,k^u}^H$$
$$+ \frac{P}{(1+\eta)^2}\mathbf{U}_{k^u}^H \sum_{i=1}^{K_u} \kappa_{g,i^u}^2\widehat{\mathbf{H}}_{g,i^u}\mathbf{V}_{i^u}\mathbf{V}i^{u\,H}\widehat{\mathbf{H}}_{g,i^u}^H$$
$$+ \frac{P\eta}{(1+\eta)}\mathbf{U}_{k^u}^H\left(b_u \sum_{i=1}^{K_u} \kappa_{g,i^u}^2 + \Theta b_d \sum_{i=1}^{K_d} \kappa_{g,g}^2\right). \tag{18}$$

The receiver which minimizes the UL mean square error is obtained by setting (18) equal to zero, resulting in

$$\mathbf{U}_{k^u} = \left(\sum_{i=1}^{K_u} \kappa_{g,i^u}^2\widehat{\mathbf{H}}_{g,i^u}\mathbf{V}_{i^u}\mathbf{V}_{i^u}^H\widehat{\mathbf{H}}_{g,i^u}^H + \gamma_{k^u}\mathbf{I}\right)^{-1}(1+\eta)\kappa_{g,k^u}\widehat{\mathbf{H}}_{g,k^u}\mathbf{V}_{k^u} \tag{19}$$

with

$$\gamma_{k^u} = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)\left(b_u \sum_{i=1}^{K_u} \kappa_{g,i^u}^2 + \Theta b_d \sum_{i=1}^{K_d} \kappa_{g,g}^2\right). \tag{20}$$

Using a similar process for DL communication in the intended direction, with fixed $\mathbf{V}$ and in accordance to Design

Principle 1, we solve

$$\min_{\mathbf{U}_{k^d}} \mathbb{E}\left\{\|\mathbf{f}_{k^d} - \mathbf{s}_{k^d}\|^2\right\} \ \forall\ k$$

where

$$\mathbf{f}_{k^d} = \kappa^2_{k^d,g}\mathbf{U}_{k^d}\mathbf{H}_{k^d,g}\mathbf{V}_{k^d}\mathbf{s}_{k^d} + \mathbf{U}_{k^d}\sum_{i=1}^{K_u}\kappa^2_{k^d,i^u}\mathbf{H}_{k^d,i^u}\mathbf{V}_{i^u}\mathbf{s}_{i^u} + \mathbf{U}_{k^d}\mathbf{n}_{k_d}.$$

This results in

$$\mathbf{U}_{k^d} = \left(\kappa^2_{k^d,g}\widehat{\mathbf{H}}_{k^d,g}\mathbf{V}_{k^d}\mathbf{V}^H_{k^d}\widehat{\mathbf{H}}^H_{k^d,g} + \sum_{i=1}^{K_u}\kappa^2_{k^d,i^u}\widehat{\mathbf{H}}_{k^d,i^u}\mathbf{V}_{i^u}\mathbf{V}^H_{i^u}\widehat{\mathbf{H}}^H_{k^d,i^u} + \gamma_{k^d}\mathbf{I}\right)^{-1}(1+\eta)\quad\kappa_{k^d,g}\quad\widehat{\mathbf{H}}_{k^d,g}\mathbf{V}_{k^d} \tag{21}$$

with

$$\gamma_{k^d} = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)\left(b_d\kappa^2_{k^d,g} + b_u\sum_{i=1}^{K_u}\kappa^2_{k^d,i^u}\right). \tag{22}$$

Next, considering the reciprocal network we can also apply a similar method to solve for $\mathbf{V}$ given fixed $\mathbf{U}$. In this network we assume that all directions of communication are reversed, i.e. UL users are receiving information from the BS, while DL users are transmitting information to the BS. Accordingly, $\mathbf{V}$ now act as receive beamformers and $\mathbf{U}$ act as precoders. We use $\overleftarrow{\mathbf{H}}_{a,b} = \mathbf{H}^H_{b,a}$ to represent the channel going from node b to node a in the reciprocal network. For communication by UL users in the reciprocal network, we solve

$$\min_{\mathbf{V}_{k^u}} \mathbb{E}\{\|\overleftarrow{\mathbf{f}}_{k^u} - \mathbf{s}_{k^u}\|^2\}\ \forall\ k$$

where

$$\overleftarrow{\mathbf{f}}_{k^u} = \kappa_{k^u,g}\mathbf{V}_{k^u}\overleftarrow{\mathbf{H}}_{k^u,g}\mathbf{U}_{k^u}\mathbf{s}_{k^u} + \mathbf{V}_{k^u}\sum_{i=1}^{K_d}\kappa_{k^d,i^u}\overleftarrow{\mathbf{H}}_{k^d,i^u}\mathbf{U}_{i^d}\mathbf{s}_{i^d} + \mathbf{V}_{k^u}\overleftarrow{\mathbf{n}}_{k_u}$$

to obtain

$$\mathbf{V}_{k^u} = \left(\kappa^2_{k^u,g}\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g}\mathbf{U}_{k^u}\mathbf{U}^H_{k^u}\overleftarrow{\widehat{\mathbf{H}}}^H_{k^u,g} + \sum_{i=1}^{K_d}\kappa^2_{k^u,i^d}\overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d}\mathbf{U}_{i^d}\mathbf{U}^H_{i^d}\overleftarrow{\widehat{\mathbf{H}}}^H_{k^u,i^d} + \overleftarrow{\gamma}_{k^u}\mathbf{I}\right)^{-1}(1+\eta)\quad\kappa_{k^u,g}\quad\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g}\mathbf{U}_{k^u} \tag{23}$$

with

$$\overleftarrow{\gamma}_{k^u} = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)\left(b_u\kappa^2_{k^u,g} + b_d\sum_{i=1}^{K_d}\kappa^2_{k^u,i^d}\right). \tag{24}$$

For communication by DL users in the reciprocal network, we solve

$$\min_{\mathbf{V}_{k^d}} \mathbb{E}\{\|\overleftarrow{\mathbf{f}}_{k^d} - \mathbf{s}_{k^d}\|^2\}\ \forall\ k$$

where

$$\overleftarrow{\mathbf{f}}_{k^d} = \sum_{i=1}^{K_d}\kappa_{g,i^d}\mathbf{V}_{k^d}\overleftarrow{\mathbf{H}}_{g,i^d}\mathbf{U}_{i^d}\mathbf{s}_{i^d} + \Theta\mathbf{V}_{k^d}\sum_{i=1}^{K_u}\kappa_{g,g}\overleftarrow{\boldsymbol{\Upsilon}}_{g,g}\mathbf{U}_{i^u}\mathbf{s}_{i^u} + \mathbf{V}_{k^d}\overleftarrow{\mathbf{n}}_g$$

to obtain

$$\mathbf{V}_{k^d} = \left(\sum_{i=1}^{K_d}\kappa^2_{g,i^d}\overleftarrow{\widehat{\mathbf{H}}}_{g,i^d}\mathbf{U}_{i^d}\mathbf{U}^H_{i^d}\overleftarrow{\widehat{\mathbf{H}}}^H_{g,i^d} + \overleftarrow{\gamma}_{k^d}\mathbf{I}\right)^{-1} \times (1+\eta)\kappa_{g,k^d}\overleftarrow{\widehat{\mathbf{H}}}_{g,k^d}\mathbf{U}_{k^d} \tag{25}$$

with

$$\overleftarrow{\gamma}_{k^d} = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)\left(b_d\sum_{i=1}^{K_d}\kappa^2_{g,i^d} + \Theta b_u\sum_{i=1}^{K_u}\kappa^2_{g,g}\right). \tag{26}$$

The resulting MMSE algorithm which exploits statistical knowledge of the CSI error (MMSE-SKCE) is as outlined in Algorithm 1.

Algorithm 1: MMSE-SKCE algorithm

1. Set γ_{k^u}, γ_{k^d}, $\overleftarrow{\gamma}_{k^u}$ and $\overleftarrow{\gamma}_{k^d}$ as (20), (22), (24) and (26) respectively.
2. Initialize $\mathbf{V}_{k^u}$ and $\mathbf{V}_{k^d}$ as random unitary matrices $\forall\ k$.
3. Obtain the receivers $\mathbf{U}_{k^u_g}$ and $\mathbf{U}_{k^d_g}$ using (19) and (21) $\forall\ k$.
4. Set $\mathbf{U}_{k^u} = \mathbb{QR}(\mathbf{U}_{k^u})$ and $\mathbf{U}_{k^d} = \mathbb{QR}(\mathbf{U}_{k^d})\ \forall\ k$.
5. Obtain the precoders $\mathbf{V}_{k^u_g}$ and $\mathbf{V}_{k^d_g}$ using (23) and (25) $\forall\ k$.
6. Set $\mathbf{V}_{k^u} = \mathbb{QR}(\mathbf{V}_{k^u})$ and $\mathbf{V}_{k^d} = \mathbb{QR}(\mathbf{V}_{k^d})\ \forall\ k$.
7. Repeat from Step 3 until convergence or for a fixed number of iterates.

B. Max-SINR based design for single-cell systems

This algorithm maximizes the signal-to-interference-plus-noise ratio on a per-stream basis. It was originally proposed as an IA algorithm in [20] for the MIMO IC under perfect CSI and later adapted for a variety of objectives in HD networks (see for example [18], [31], [33], [34] and references therein). In contrast to a naive approach, which would simply use the CSI provided assuming it is perfect, our algorithm exploits knowledge of η to compute more accurate beamformers. Moreover, apart from following Design Principle 1, our interference-plus-noise covariance matrices also take into account inter-stream interference for the data required at each node.

Starting with UL communication in the intended direction, the interference-plus-noise covariance matrix is given by (27), where (a) follows by replacing $\mathbf{H}$ with (6). Using knowledge of η, we can simplify (27) further by replacing all the elements that contain $\boldsymbol{\Upsilon}$ by their expected values. Applying Lemmas 1 and 2 from Appendix C, $\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\mathrm{A}\} = \mathbf{0}$ and $\mathbb{E}_{\boldsymbol{\Upsilon}}\{\mathrm{B}\} = \mathbb{E}_{\boldsymbol{\Upsilon}}\{\mathrm{C}\} = \eta/(1+\eta)\mathbf{I}$. Thus instead of (27), we can use

$$\widehat{\mathbf{Q}}^n_{k^u} = \sum_{\substack{i=1\\(i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u}\tau\kappa^2_{g,i^u}\widehat{\mathbf{H}}_{g,i^u}\mathbf{v}^m_{i^u}\mathbf{v}^{m\,H}_{i^u}\widehat{\mathbf{H}}^H_{g,i^u} + \xi_{k^u}\mathbf{I} \tag{28}$$

where

$$\tau = \frac{P}{(1+\eta)^2} \tag{29}$$

$$\mathbf{Q}_{k^u}^n = \sum_{\substack{i=1 \\ (i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u} P\kappa_{g,i^u}^2 \mathbf{H}_{g,i^u}\mathbf{v}_{i^u}^m \mathbf{v}_{i^u}^{m\,H}\mathbf{H}_{g,i^u}^H + \Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P\kappa_{g,g}^2 \boldsymbol{\Upsilon}_{g,g}\mathbf{v}_{i^d}^m \mathbf{v}_{i^d}^{m\,H}\boldsymbol{\Upsilon}_{g,g}^H + \sigma^2\mathbf{I}$$

$$\overset{(a)}{=} \sum_{\substack{i=1 \\ (i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u} P\kappa_{g,i^u}^2 \Bigg[\frac{1}{(1+\eta)^2}\widehat{\mathbf{H}}_{g,i^u}\mathbf{v}_{i^u}^m \mathbf{v}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{g,i^u}^H + \frac{1}{(1+\eta)}\underbrace{\left(\widehat{\mathbf{H}}_{g,i^u}\mathbf{v}_{i^u}^m \mathbf{v}_{i^u}^{m\,H}\boldsymbol{\Upsilon}_{g,i^u}^H + \boldsymbol{\Upsilon}_{g,i^u}^H\mathbf{v}_{i^u}^m \mathbf{v}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{g,i^u}\right)}_{\text{A}}$$
$$+ \sum_{\substack{i=1 \\ (i,m)\neq(k,n)}}^{K_u}\sum_{m=1}^{b_u} \underbrace{\boldsymbol{\Upsilon}_{g,i^u}\mathbf{v}_{i^u}^m \mathbf{v}_{i^u}^{m\,H}\boldsymbol{\Upsilon}_{g,i^u}^H}_{\text{B}}\Bigg] + \Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} P\kappa_{g,g}^2 \underbrace{\boldsymbol{\Upsilon}_{g,g}\mathbf{v}_{i^d}^m \mathbf{v}_{i^d}^{m\,H}\boldsymbol{\Upsilon}_{g,g}^H}_{\text{C}} + \sigma^2\mathbf{I} \tag{27}$$

and

$$\xi_{k^u} = \sigma^2 + \frac{P\eta}{(1+\eta)}\left(b_u\sum_{i=1}^{K_u}\kappa_{g,i^u}^2 - \kappa_{g,k^u}^2 + \Theta b_d\sum_{i=1}^{K_d}\kappa_{g,g}^2\right). \tag{30}$$

Similarly for DL communication in the intended direction we obtain

$$\widehat{\mathbf{Q}}_{k^d}^n = \sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\tau\kappa_{k^d,i^u}^2\widehat{\mathbf{H}}_{k^d,i^u}\mathbf{v}_{i^u}^m\mathbf{v}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{k^d,i^u}^H + \sum_{\substack{m=1 \\ m\neq n}}^{b_d}\tau\kappa_{k^d,g}^2\widehat{\mathbf{H}}_{k^d,g}\mathbf{v}_{k^d}^m\mathbf{v}_{k^d}^{m\,H}\widehat{\mathbf{H}}_{k^d,g}^H + \xi_d\mathbf{I} \tag{31}$$

where

$$\xi_d = \sigma^2 + \frac{P\eta}{(1+\eta)}\left(b_u\sum_{i=1}^{K_u}\kappa_{k^d,i^u}^2 + (b_d-1)\kappa_{k^d,g}^2\right). \tag{32}$$

Next, reversing the direction of communication, for UL users in the reciprocal network, we obtain

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k^u}^n = \sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\tau\kappa_{k^u,i^d}^2\overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d}\mathbf{u}_{i^d}^m\mathbf{u}_{i^d}^{m\,H}\overleftarrow{\widehat{\mathbf{H}}}_{k^u,i^d}^H + \sum_{\substack{m=1 \\ m\neq n}}^{b_u}\tau\kappa_{k^u,g}^2\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g}\mathbf{u}_{k^u}^m\mathbf{u}_{k^u}^{m\,H}\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g} + \overleftarrow{\xi}_{\,u}\mathbf{I} \tag{33}$$

where

$$\overleftarrow{\xi}_{\,u} = \sigma^2 + \frac{P\eta}{(1+\eta)}\left(b_d\sum_{i=1}^{K_d}\kappa_{k^u,i^d}^2 + (b_u-1)\kappa_{k^u,g}^2\right). \tag{34}$$

Additionally for DL users in the reciprocal network, we have

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k^d}^n = \sum_{\substack{i=1 \\ (i,m)\neq(k,n)}}^{K_d}\sum_{m=1}^{b_d}\tau\kappa_{g,i^d}^2\overleftarrow{\widehat{\mathbf{H}}}_{g,i^d}\mathbf{u}_{i^d}^m\mathbf{u}_{i^d}^{m\,H}\overleftarrow{\widehat{\mathbf{H}}}_{g,i^d}^H + \overleftarrow{\xi}_{\,k^d}\mathbf{I} \tag{35}$$

where

$$\overleftarrow{\xi}_{\,k^d} = \sigma^2 + \frac{P\eta}{(1+\eta)}(b_d\sum_{i=1}^{K_d}\kappa_{g,i^d}^2 - \kappa_{g,k^d}^2 + \Theta b_u\sum_{i=1}^{K_u}\kappa_{g,g}^2\Bigg). \tag{36}$$

The resulting Max-SINR algorithm which exploits statistical knowledge of the CSI error (Max-SINR-SKCE) is as outlined in Algorithm 2. Note that the original Max-SINR IA based algorithm from [20] does not contain a QR decomposition stage, but instead normalizes the per-stream beamformers. Having unitary beamformers was later shown to improve performance for multi-stream applications [31]. By including a QR decomposition stage in Steps 5 and 8 of Algorithm 2 we produce unitary beamformers, thereby ensuring we obtain the multi-stream advantages, and also eliminating the need for separate normalization steps since the resultant beamformers inherently consist of unit-norm vectors.

Algorithm 2: Max-SINR-SKCE algorithm

1. Set τ, ξ_{k^u}, ξ_d, $\overleftarrow{\xi}_{\,u}$ and $\overleftarrow{\xi}_{\,k^d}$ as (30), (29), (32), (34) and (36) respectively.
2. Calculate $\widehat{\mathbf{Q}}_{k^u}^n$ and $\widehat{\mathbf{Q}}_{k^d}^n$ using (28) and (31) $\forall\, n,k$.
3. Obtain the receive filters as $\mathbf{u}_{k^u}^n = \kappa_{g,k^u}(\widehat{\mathbf{Q}}_{k^u}^n)^{-1}\widehat{\mathbf{H}}_{g,k^u}\mathbf{v}_{k^u}^n$ and $\mathbf{u}_{k^d}^n = \kappa_{k^d,g}(\kappa_{k^d,g}\widehat{\mathbf{Q}}_{k^d}^n)^{-1}\widehat{\mathbf{H}}_{k^d,g}\mathbf{v}_{k^d}^n$ $\forall\, n,k$.
4. Set $\mathbf{U}_{k^u} = \mathbb{QR}(\mathbf{U}_{k^u})$ and $\mathbf{U}_{k^d} = \mathbb{QR}(\mathbf{U}_{k^d})$ $\forall\, k$.
5. Compute $\overleftarrow{\widehat{\mathbf{Q}}}_{k^u}^n$ and $\overleftarrow{\widehat{\mathbf{Q}}}_{k^d}^n$ using (33) and (35) $\forall\, n,k$.
6. Obtain the precoders as $\mathbf{v}_{k^u}^n = \kappa_{k^u,g}(\overleftarrow{\widehat{\mathbf{Q}}}_{k^u}^n)^{-1}\overleftarrow{\widehat{\mathbf{H}}}_{k^u,g}\mathbf{u}_{k^u}^n$ and $\mathbf{v}_{k^d}^n = \kappa_{g,k^d}(\overleftarrow{\widehat{\mathbf{Q}}}_{k^d}^n)^{-1}\overleftarrow{\widehat{\mathbf{H}}}_{g,k^d}\mathbf{u}_{k^d}^n$ $\forall\, n,k$.
7. Set $\mathbf{V}_{k^u} = \mathbb{QR}(\mathbf{V}_{k^u})$ and $\mathbf{V}_{k^d} = \mathbb{QR}(\mathbf{V}_{k^d})$ $\forall\, k$.
8. Repeat from Step 3 until convergence or for a fixed number of iterates.

Remark 2. *For perfect CSI or for imperfect CSI scenarios where statistical knowledge of the CSI error is unavailable, a naive version of Algorithms 1 and 2 can be implemented. For such situations we have $\eta = 0$ in the expressions for beamformer calculation. Thus, for MMSE-Naive we set $\gamma_u = \gamma_d = \overleftarrow{\gamma}_{\,u} = \overleftarrow{\gamma}_{\,d} = \frac{\sigma^2}{P}$ in Step 1. While, for Max-SINR-Naive we set $\tau = P$ and $\xi_u = \xi_d = \overleftarrow{\xi}_{\,u} = \overleftarrow{\xi}_{\,d} = \sigma^2$ in Step 1. Additionally, for the perfect CSI case $\mathbf{H}$ is used in place of $\widehat{\mathbf{H}}$ throughout. Note that the naive versions of the algorithms have the same computational complexity as those originally presented in Algorithms 1 and 2.*

C. Equivalence between MMSE and Max-SINR designs

Under certain conditions the beamformers obtained by the proposed MMSE and Max-SINR algorithms are equivalent,

implying that at each iteration both result in identical precoders and receivers.

Consider the definition for $\mathbf{u}_{k^u}^n$ from Step 4 of Algorithm 2. Defining $\widehat{\mathbf{Q}}_{k^u}^n \triangleq \mathbf{A}_{k^u}^n \frac{P}{(1+\eta)^2}$, we can express

$$\mathbf{u}_{k^u}^n = \kappa_{g,k^u} (\mathbf{A}_{k^u}^n)^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n \frac{(1+\eta)^2}{P} .$$

Additionally $\mathbf{A}_{k^u}^n$ may be represented as

$$\mathbf{A}_{k^u}^n = \mathbf{B}_{k^u} - \kappa_{g,k^u}^2 \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n \mathbf{v}_{k^u}^{n\,H} \widehat{\mathbf{H}}_{g,k^u}^H$$

where

$$\mathbf{B}_{k^u} = \sum_{i=1}^{K_u} \kappa_{g,i^u} \widehat{\mathbf{H}}_{g,i^u} \mathbf{V}_{i^u} \mathbf{V}_{i^u}^H \widehat{\mathbf{H}}_{g,i^u}^H + \xi_{k^u} \frac{(1+\eta)^2}{P} \mathbf{I} .$$

Applying Lemma 4 to $(\mathbf{A}_{k^u}^n)^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n$ we obtain

$$(\mathbf{A}_{k^u}^n)^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{g,k^u}^n = \frac{(\mathbf{B}_{k^u})^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n}{1 - \mathbf{v}_{k^u}^{n\,H} \widehat{\mathbf{H}}_{g,k^u}^H (\mathbf{B}_{k^u})^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n} .$$

Letting $\lambda_{k^u}^n = 1 - \mathbf{v}_{k^u}^{n\,H} \widehat{\mathbf{H}}_{g,k^u}^H (\mathbf{B}_{k^u})^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n$, we can represent the receiver as

$$\mathbf{u}_{k^u}^n = \kappa_{g,k^u} (\mathbf{B}_{k^u})^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{v}_{k^u}^n \frac{(1+\eta)^2}{P} \frac{1}{\lambda_{k^u}^n} .$$

Next, $\mathbf{u}_{k^u}^n \ \forall\, n = 1 \dots b_u$ can be horizontally concatenated to obtain the receiver across all streams as

$$\mathbf{U}_{k^u} = \kappa_{g,k^u} (\mathbf{B}_{k^u})^{-1} \widehat{\mathbf{H}}_{g,k^u} \mathbf{V}_{k^u} \mathbf{\Lambda}_{k^u} \tag{37}$$

where

$$\mathbf{\Lambda}_{k^u} = \frac{(1+\eta)^2}{P} \begin{bmatrix} \frac{1}{\lambda_{k^u}^1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_{k^u}^{b_u}} \end{bmatrix} .$$

Comparing (37) with the MMSE derived expression in (19), it can be noticed that they are very similar. For the naive and perfect CSI versions of the algorithms where $\eta = 0$, the term inside the inverse for (37) and (19) is equivalent. The only difference is an additional post-multiplication by $\mathbf{\Lambda}_{k^u}$ in (37); this matrix essentially multiplies each column vector with a scalar and thus has no effect on the resultant unitary part after the QR decomposition, therefore both algorithms obtain the same $\mathbf{U}_{k^u}$. A similar argument can be made for each of $\mathbf{U}_{k^d}$, $\mathbf{V}_{k^u}$ and $\mathbf{V}_{k^d}$. Thus in cases where η is actually 0, or unknown and assumed to be 0, (i.e. Max-SINR-Naive and MMSE-Naive) the two algorithms are equivalent.

Remark 3. *Note that even in cases where the Max-SINR/MMSE equivalence holds, the MMSE algorithm is less computationally complex than the Max-SINR algorithm, since the former operates on a per-user basis whilst the latter operates on a per-stream basis. Consider for example the number of matrix inverses involved; the MMSE algorithm requires a total of $2(K_u + K_d)$ inverses per iteration to compute the beamformers in Steps 3 and 5, while the Max-SINR algorithm requires a total of $2(K_u b_u + K_d b_d)$ inverses in Steps 3 and 5.*

D. Convergence of the proposed algorithms

Firstly, it is important to note that the convergence of Max-SINR based algorithms to achieve IA cannot be proven analytically, not even for the simplest case of the HD interference channel [20]. Considering that the Max-SINR algorithm proposed in this paper is based in principle on the original one from [20], but with increased complexity in the resultant expressions (due to the more complex system model), it follows by extension that the convergence of our algorithm cannot be analytically proven. However, the overall consensus in literature is that Max-SINR based algorithms for IA generally seem to converge to a constant value, as shown numerically in [35], and proven for sufficiently high SNR in [36]. Finally, it is also important to note that these convergence remarks also apply to our MMSE based algorithm, due to the equivalence established in Section IV-C. (For further information on the convergence behavior please refer to the simulations in Section VI-D.)

V. Multi-cell considerations

A significant body of literature related to linear IA focuses on the analytic derivation of feasibility conditions, for example, [37], [38] study this issue for the interference channel, [33], [39] consider interference broadcast channels and [13] derives feasibility conditions for linear IA in single-cell systems with an FD BS communicating with both UL and DL users. However, no feasibility conditions are available in literature so far for multi-cell multi-user systems with FB BSs and HD users. Here, we look into this issue by deriving the proper condition for this network type and also by extending the linear IA algorithms proposed in Section IV to the multi-cell case. The derivation of this condition and the algorithms can aid future work in this direction by serving as a starting point that provides insight into the theoretical feasibility of linear IA for different antenna configurations and DoF requirements.

A. Proper condition

The proper condition relates the feasibility of IA to the issue of determining the resolvability of a system represented by multivariate polynomial equations. A system of equations is classified as *proper* if the number of equations, N_e, does not exceed the number of variables, N_v, i.e. if $N_v \geq N_e$. Prior studies [37]–[39] show that for systems classified as *improper*, IA is surely infeasible. However, classifying a system as *proper* is not a sufficient condition to prove IA feasibility, i.e. systems that are proper but for which IA is infeasible may also exist.

We follow the method from [37] to derive expressions for N_v and N_e, and obtain the proper condition for the FD enabled multi-cell scenario considered in this work. Focusing on a symmetric system where $K_d = K_u = K$, $b_d = b_u = b$ and $M_d = M_u = N$ to simplify notation, we obtain

$$N_v = 2GKb(M_B + N - 2b)$$

$$\mathbf{U}_{k_g^d} = \left(\sum_{j=1}^{G} \sum_{i=1}^{K_u} \widehat{\mathbf{H}}_{k_g^d, i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \widehat{\mathbf{H}}_{k_g^d, i_j^u}^H + \sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{i=1}^{K_d} \widehat{\mathbf{H}}_{k_g^d, j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \widehat{\mathbf{H}}_{k_g^d, j}^H + \widehat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \mathbf{V}_{k_g^d}^H \widehat{\mathbf{H}}_{k_g^d, g}^H + \gamma_d \mathbf{I} \right)^{-1} (1+\eta) \widehat{\mathbf{H}}_{k_g^d, g} \mathbf{V}_{k_g^d} \quad (39)$$

$$\mathbf{U}_{k_g^u} = \left(\sum_{j=1}^{G} \sum_{i=1}^{K_u} \widehat{\mathbf{H}}_{g, i_j^u} \mathbf{V}_{i_j^u} \mathbf{V}_{i_j^u}^H \widehat{\mathbf{H}}_{g, i_j^u}^H + \sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{i=1}^{K_d} \widehat{\mathbf{H}}_{g, j} \mathbf{V}_{i_j^d} \mathbf{V}_{i_j^d}^H \widehat{\mathbf{H}}_{g, j}^H + \gamma_u \mathbf{I} \right)^{-1} (1+\eta) \quad \widehat{\mathbf{H}}_{g, k_g^u} \mathbf{V}_{k_g^u} \quad (40)$$

$$\mathbf{V}_{k_g^u} = \left(\sum_{j=1}^{G} \sum_{i=1}^{K_d} \overleftarrow{\widehat{\mathbf{H}}}_{i_j^d, k_g^u} \mathbf{U}_{i_j^d} \mathbf{U}_{i_j^d}^H \overleftarrow{\widehat{\mathbf{H}}}_{i_j^d, k_g^u}^H + \sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{i=1}^{K_u} \overleftarrow{\widehat{\mathbf{H}}}_{j, k_g^u} \mathbf{U}_{i_j^u} \mathbf{U}_{i_j^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{j, k_g^u}^H + \overleftarrow{\widehat{\mathbf{H}}}_{g, k_g^u} \mathbf{U}_{k_g^u} \mathbf{U}_{k_g^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{g, k_g^u}^H + \overleftarrow{\gamma}_u \mathbf{I} \right)^{-1} (1+\eta) \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u, g} \mathbf{U}_{k_g^u} \quad (41)$$

$$\mathbf{V}_{k_g^d} = \left(\sum_{j=1}^{G} \sum_{i=1}^{K_d} \overleftarrow{\widehat{\mathbf{H}}}_{g, i_j^d} \mathbf{U}_{i_j^d} \mathbf{U}_{i_j^d}^H \overleftarrow{\widehat{\mathbf{H}}}_{g, i_j^d}^H + \sum_{\substack{j=1 \\ j \neq g}}^{G} \sum_{i=1}^{K_u} \overleftarrow{\widehat{\mathbf{H}}}_{g, j} \mathbf{U}_{i_j^u} \mathbf{U}_{i_j^u}^H \overleftarrow{\widehat{\mathbf{H}}}_{g, j}^H + \overleftarrow{\gamma}_d \mathbf{I} \right)^{-1} (1+\eta) \quad \overleftarrow{\widehat{\mathbf{H}}}_{g, k_g^d} \mathbf{U}_{k_g^d} \quad (42)$$

and

$$N_e = (Kb)^2 (4G^2 - 2 - G) \, .$$

This allows us to express the proper condition as

$$\frac{2G(M_B + N)}{4G + K(4G^2 - 2 - G)} \geq b \, . \quad (38)$$

B. Multi-cell algorithm extension

Here we extend our IA algorithms to the multi-cell case. Since as outlined earlier in this section, the goal of the multi-cell version of the algorithms is to aid future work into understanding the theoretical feasibility of IA in multi-cell FD systems, the algorithms presented here are obtained by considering a homogeneous pathloss scenario, i.e. setting all κ terms to be equal to 1. This approach allows for equations that are clearer in terms of presentation, and is suited to IA feasibility studies since pathloss has no effect in the DoF domain where power $\to \infty$.

The design of the multi-cell algorithms is largely analogous to the single-cell ones, thus derivation details are omitted. The major difference in the derivation process is that instead of following Design Principle 1, we follow Design Principle 2 which includes additional considerations for CCI-OC which is now present.

Design Principle 2. *Intra-UL interference is only handled by the receivers. Intra-DL interference is only handled by the precoders. R-SI is handled by both the transmit and receive beamformers at the BSs. CCI-OC is handled by all beamformers.*

1) Multi-cell version of MMSE algorithm: The multi-cell version of MMSE-SKCE follows the general steps outlined for the single-cell version in Algorithm 1, with the following differences.

- In Step 1 set γ_d, γ_u, $\overleftarrow{\gamma}_d$ and $\overleftarrow{\gamma}_u$ as follows.

$$\gamma_d = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_u b_u + (G-1)K_d b_d + b_d - 1)$$
$$\gamma_u = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_u b_u - 1 + (G-1+\Theta)K_d b_d)$$
$$\overleftarrow{\gamma}_d = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_d b_d - 1 + (G-1+\Theta)K_u b_u)$$
$$\overleftarrow{\gamma}_u = \frac{\sigma^2(1+\eta)^2}{P} + \eta(1+\eta)(GK_d b_d + (G-1)K_u b_u + b_u - 1)$$

- In Step 3 calculate $\mathbf{U}_{k_g^d}$ and $\mathbf{U}_{k_g^u}$ using (39) and (40).
- In Step 5 find $\mathbf{V}_{k_g^d}$ and $\mathbf{V}_{k_g^u}$ using (41) and (42).

2) Multi-cell version of Max-SINR algorithm: The multi-cell version of Max-SINR-SKCE follows the general steps outlined for the single-cell version in Algorithm 2, with the following differences.

- In Step 1 set ξ_d, ξ_u, $\overleftarrow{\xi}_d$ and $\overleftarrow{\xi}_u$ as follows.

$$\xi_d = \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_u b_u + (G-1)K_d b_d + b_d)$$
$$\xi_u = \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_u b_u + (G-1+\Theta)K_d b_d)$$
$$\overleftarrow{\xi}_d = \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_d b_d + (G-1+\Theta)K_u b_u)$$
$$\overleftarrow{\xi}_u = \sigma^2 + \frac{P\eta}{(1+\eta)}(GK_d b_d + (G-1)K_u b_u + b_u)$$

- In Step 3 the forward interference-plus-noise covariances matrices $\widehat{\mathbf{Q}}_{k_g^d}^n$ and $\widehat{\mathbf{Q}}_{k_g^u}^n$ are calculated using (43) and (44).
- In Step 4 use $\mathbf{u}_{k_g^u}^n = (\widehat{\mathbf{Q}}_{k_g^u}^n)^{-1} \widehat{\mathbf{H}}_{g, k_g^u} \mathbf{v}_{k_g^u}^n$ and $\mathbf{u}_{k_g^d}^n = (\widehat{\mathbf{Q}}_{k_g^d}^n)^{-1} \widehat{\mathbf{H}}_{k_g^d, g} \mathbf{v}_{k_g^d}^n \; \forall \, n, k, g$.
- In Step 6 the backward interference-plus-noise covariance matrices $\overleftarrow{\mathbf{Q}}_{k_g^d}^n$ and $\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^u}^n$ are calculated using (45) and (46).
- In Step 7 use $\mathbf{v}_{k_g^u}^n = (\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^u}^n)^{-1} \overleftarrow{\widehat{\mathbf{H}}}_{k_g^u, g} \mathbf{u}_{k_g^u}^n$ and $\mathbf{v}_{k_g^d}^n = (\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^d}^n)^{-1} \overleftarrow{\widehat{\mathbf{H}}}_{g, k_g^d} \mathbf{u}_{k_g^d}^n \; \forall \, n, k, g$.

VI. Simulations

For the purpose of our simulations we set $\sigma^2 = 1$, making the SNR equivalent to the transmit power, and we consider a homogeneous pathloss scenario by setting $\kappa_{k_g^d, j} = \kappa_{k_g^d, i_j^u} = \kappa_{g, j} = \kappa_{g, i_j^u} = 1 \; \forall \; k, g, i, j$. Additionally, all results are averaged in a Monte-Carlo fashion over a number of dif-

$$\widehat{\mathbf{Q}}_{k_g^d}^n = \sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} \tau\widehat{\mathbf{H}}_{k_g^d,i_j^u}\mathbf{v}_{i_j^u}^m{}^H\widehat{\mathbf{H}}_{k_g^d,i_j^u}^H + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \tau\widehat{\mathbf{H}}_{k_g^d,j}\mathbf{v}_{i_j^d}^m\mathbf{v}_{i_j^d}^m{}^H\widehat{\mathbf{H}}_{k_g^d,j}^H + \sum_{\substack{m=1\\ m\neq n}}^{b_d} \tau\widehat{\mathbf{H}}_{k_g^d,g}\mathbf{v}_{k_g^d}^m\mathbf{v}_{k_g^d}^m{}^H\widehat{\mathbf{H}}_{k_g^d,g}^H + \xi_d\mathbf{I} \quad (43)$$

$$\widehat{\mathbf{Q}}_{k_g^u}^n = \sum_{\substack{j=1\\ (j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} \tau\widehat{\mathbf{H}}_{g,i_j^u}\mathbf{v}_{i_j^u}^m\mathbf{v}_{i_j^u}^m{}^H\widehat{\mathbf{H}}_{g,i_j^u}^H + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \tau\widehat{\mathbf{H}}_{g,j}\mathbf{v}_{i_j^d}^m\mathbf{v}_{i_j^d}^m{}^H\widehat{\mathbf{H}}_{g,j}^H + \xi_u\mathbf{I} \quad (44)$$

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^d}^n = \sum_{\substack{j=1\\ (j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \tau\overleftarrow{\widehat{\mathbf{H}}}_{g,i_j^d}\mathbf{u}_{i_j^d}^m\mathbf{u}_{i_j^d}^m{}^H\overleftarrow{\widehat{\mathbf{H}}}_{g,i_j^d}^H + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} \tau\overleftarrow{\widehat{\mathbf{H}}}_{g,j}\mathbf{u}_{i_j^d}^m\mathbf{u}_{i_j^d}^m{}^H\overleftarrow{\widehat{\mathbf{H}}}_{g,j}^H + \overleftarrow{\xi}_d\mathbf{I} \quad (45)$$

$$\overleftarrow{\widehat{\mathbf{Q}}}_{k_g^u}^n = \sum_{j=1}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d} \tau\overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,i_j^d}\mathbf{u}_{i_j^d}^m\mathbf{u}_{i_j^d}^m{}^H\overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,i_j^d}^H + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u} \tau\overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,j}\mathbf{u}_{i_j^u}^m\mathbf{u}_{i_j^u}^m{}^H\overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,j}^H + \sum_{\substack{m=1\\ m\neq n}}^{b_u} \tau\overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,g}\mathbf{u}_{k_g^u}^m\mathbf{u}_{k_g^u}^m{}^H\overleftarrow{\widehat{\mathbf{H}}}_{k_g^u,g}^H + \overleftarrow{\xi}_u\mathbf{I} \quad (46)$$

ferent channel realizations. Treating all interference as noise, throughout the simulations we calculate the sum rate as

$$R = \sum_{g=1}^{G}\sum_{k=1}^{K_d}\log_2\det\Big(\mathbf{I} + (\mathbf{X}_{k_g^d} + \sigma^2\mathbf{I})^{-1}\mathbf{S}_{k_g^d}\Big) + \sum_{g=1}^{G}\sum_{k=1}^{K_u}\log_2\det\Big(\mathbf{I} + (\mathbf{X}_{k_g^u} + \sigma^2\mathbf{I})^{-1}\mathbf{S}_{k_g^u}\Big)$$

where the signal covariance matrices are

$$\mathbf{S}_{k_g^d} = P\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,g}\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H\mathbf{H}_{k_g^d,g}^H\mathbf{U}_{k_g^d}$$

and

$$\mathbf{S}_{k_g^u} = P\mathbf{U}_{k_g^u}^H\mathbf{H}_{g,k_g^u}\mathbf{V}_{k_g^u}\mathbf{V}_{k_g^u}^H\mathbf{H}_{g,k_g^u}^H\mathbf{U}_{k_g^u}\ .$$

Additionally, $\mathbf{X}_{k_g^d}$ and $\mathbf{X}_{k_g^u}$ represent the DL and UL interference covariance matrices and are respectively given by

$$\mathbf{X}_{k_g^d} = \sum_{\substack{j=1\\ (j,i)\neq(g,k)}}^{G}\sum_{i=1}^{K_d} P\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{k_g^d,j}^H\mathbf{U}_{k_g^d} + \sum_{j=1}^{G}\sum_{i=1}^{K_u} P\mathbf{U}_{k_g^d}^H\mathbf{H}_{k_g^d,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{k_g^d,i_j^u}^H\mathbf{U}_{k_g^d}$$

and

$$\mathbf{X}_{k_g^u} = \sum_{\substack{j=1\\ (j,i)\neq(g,k)}}^{G}\sum_{i=1}^{K_u} P\mathbf{U}_{k_g^u}^H\mathbf{H}_{g,i_j^u}\mathbf{V}_{i_j^u}\mathbf{V}_{i_j^u}^H\mathbf{H}_{g,i_j^u}^H\mathbf{U}_{k_g^u} + \sum_{\substack{j=1\\ j\neq g}}^{G}\sum_{i=1}^{K_d} P\mathbf{U}_{k_g^u}^H\mathbf{H}_{g,j}\mathbf{V}_{i_j^d}\mathbf{V}_{i_j^d}^H\mathbf{H}_{g,j}^H\mathbf{U}_{k_g^u} + \Theta\sum_{i=1}^{K_d} P\mathbf{U}_{k_g^u}^H\mathbf{\Upsilon}_{g,g}\mathbf{V}_{k_g^d}\mathbf{V}_{k_g^d}^H\mathbf{\Upsilon}_{g,g}^H\mathbf{U}_{k_g^u}\ .$$

Note that for imperfect CSI scenarios, throughout this section we fix $\Theta = 1$ (i.e. we assume imperfect SI cancellation).

A. Results for theoretically derived bounds

We simulate a system having $G = 1$, $K_d = K_u = 4$, $b_d = b_u = 1$, $M_B = 4$ and $M_d = M_u = 3$ to obtain Fig. 2 and Fig. 3. For this system IA is known to be feasible [13] and the achievable DoF for perfect CSI are given by $G(K_d b_d + K_u b_u) = 8$. Its HD counterpart has $M_B = 4$, $N = M_d = M_u = 3$ antennas at the users, and serves all $K_{\text{TOT/cell}} = K_d + K_u = 8$ in the same direction simultaneously. The total achievable DoF using an HD BS are given by $\min\{M_B, NK_{\text{TOT/cell}}\} = 4$, which corresponds to exactly half those achieved when using an FD BS.

From Theorem 2 we know that full DoF are achievable for values of $\alpha \geq 1$. This can be verified by focusing on the results for $\alpha = 1.75$ and $\alpha = 1$ in Fig. 2, which have slopes that are equivalent to that of the perfect CSI case. However, while the $\alpha = 1.75$ result overlaps completely with the perfect CSI one at high SNR, the $\alpha = 1$ result demonstrates a constant gap for SNR > 30 dB. This behavior follows from Theorem 1, which predicts a finite loss upper bounded by Ω for $\alpha = 1$. For the system under consideration with $\beta = 10$, Ω is equates to 49.20 bits per channel use. Measuring the actual difference between the perfect CSI and the $\alpha = 1$ results from Fig. 2, we obtain a value of 47.9 bits per channel use, proving that the derived upper bound is not excessively loose.

Focusing on the range of $\alpha < 1$, Theorem 1 indicates that the sum rate loss is unbounded. This can be easily verified by considering the results for $\alpha = 0.75$ and $\alpha = 0$ in Fig. 2, all of which deviate from the perfect CSI result. From a DoF perspective, in the range of $\alpha < 1$ Theorem 2 indicates a loss equal to $(1 - \alpha)$ of the full Dof. For example for $\alpha = 0.75$, Theorem 2 predicts that only 75% of the full DoF are achievable. This can be confirmed by comparing the high SNR slopes for the perfect CSI curve, which achieves 8 DoF, and the one for $\alpha = 0.75$, which achieves 6 DoF. For $\alpha = 0$ the same theorem predicts 0 DoF achievable, and indeed both $\alpha = 0$ curves lie flat in the high SNR region. Additionally comparing the result for $\alpha = 0, \beta = 0.01$ and $\alpha = 0, \beta = 0.1$, it is clear that while the β value does not affect DoF behavior, it has a strong effect on the overall achievable rate. The curve for the smallest β settles at the higher value, which is expected since this indicates the smallest error. Note that for any β, $\alpha = 0$ represents the worst-case scenario with the CSI error variance being equal to β itself; this causes a significant amount of

interference leakage, making the system interference limited and eventually causing sum rate saturation.

For $\alpha < 1$, Theorem 2 also shows how the DoF loss is distributed amongst the DL and UL users. In Fig. 3 we plot the corresponding rates separately to verify this behavior. As can be seen for $\alpha = 1, \beta = 10$ total of 8 DoF are achieved; due to the symmetry of the simulated system where $K_u = K_d$ and $b_d = b_u$, this amounts to 4 DoF each for UL and DL. Focusing on $\alpha = 0.75, \beta = 10$ it can noticed that both DL and UL results have a high SNR slope that corresponds to 3 DoF, while for $\alpha = 0, \beta = 0.1$ the slopes corresponds to 0 DoF. In both cases the achieved DoF are equivalent to $\alpha G K_d b_d$ for the DL and $\alpha G K_u b_u$ for the UL, which confirms our expectations from Theorem 2.

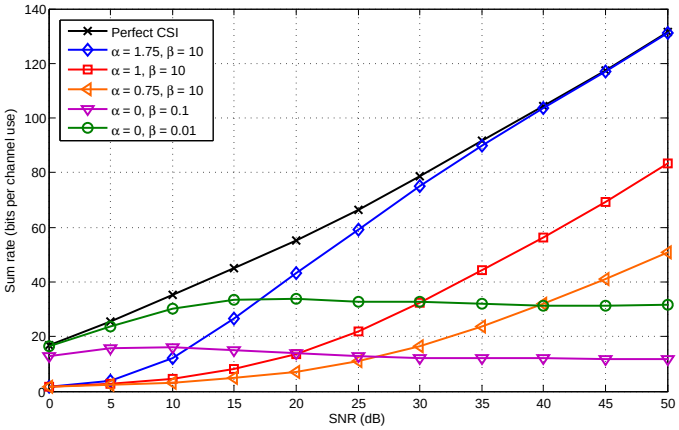

Fig. 2: Sum rate performance obtained using both Max-SINR-Naive and MMSE-Naive algorithms for scenario with $G = 1$, $K_d = K_u = 4$, $b_d = b_u = 1$, $M_B = 4$ and $M_d = M_u = 3$ under different imperfect CSI error conditions.

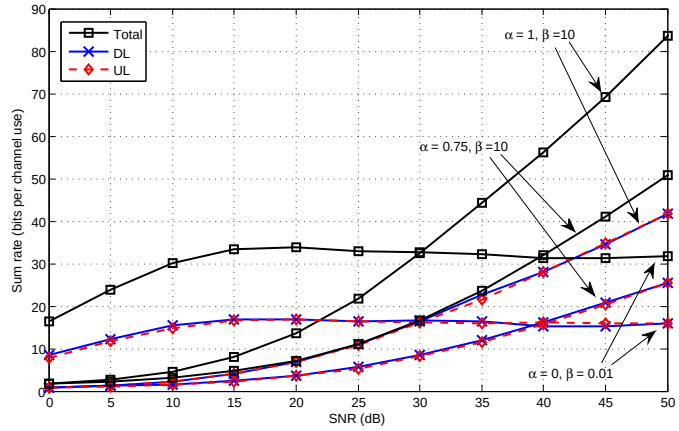

Fig. 3: Total, DL and UL rate performance obtained by using Max-SINR-Naive and MMSE-Naive algorithms for scenario with $G = 1$, $K_d = K_u = 4$, $b_d = b_u = 1$, $M_B = 4$ and $M_d = M_u = 3$ under different imperfect CSI error conditions.

Remark 4. *The imperfect CSI scenarios modeled in this section can be directly related to the CSI acquisition techniques outlined in Section II-B. For example, $\alpha = 0$ corresponds to the non-reciprocal channel scenario. Looking at the corresponding results in Fig. 2 it is clear that within this context IA works better in the lower SNR region. The overall performance depends on the quality of the CSI, which for the case of $\alpha = 0$ is a function of the amount of quantization. The lower the β, the smaller is the error due to quantization and the better IA performs. On the other hand for reciprocal channels, modeled by $\alpha = 1$, IA fares better in the higher SNR region. In this case the CSI error is inversely proportional to SNR, therefore its effect decreases with increasing SNR, leading to a better performance of the IA techniques.*

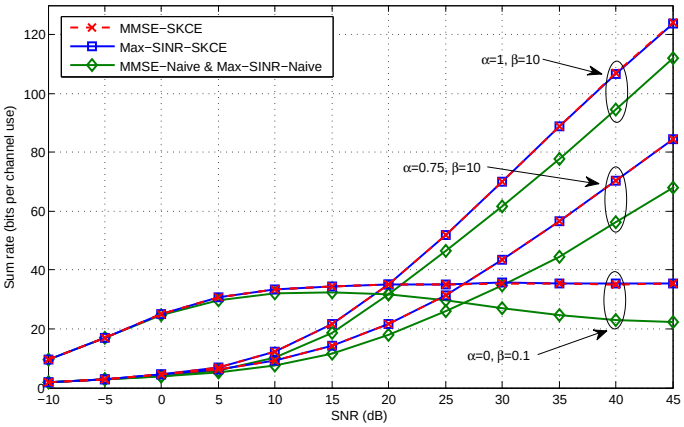

Fig. 4: Sum rate performance for scenario with $G = 1$, $K_d = K_u = 3$, $b_d = b_u = 2$ and $M_B = M_d = M_u = 6$ under different imperfect CSI error conditions.

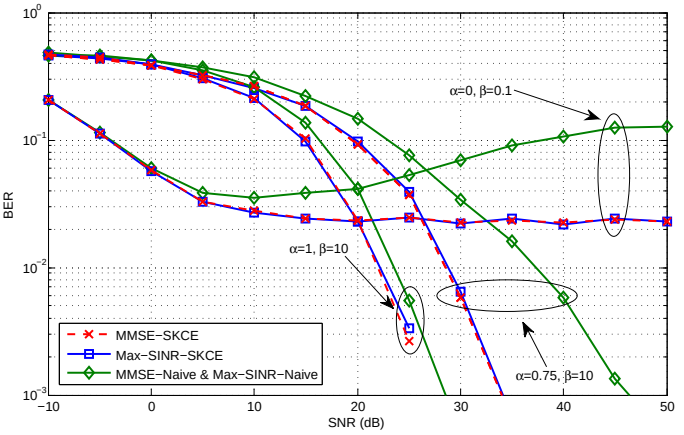

Fig. 5: BER performance for scenario with $G = 1$, $K_d = K_u = 3$, $b_d = b_u = 2$ and $M_B = M_d = M_u = 6$ under different imperfect CSI error conditions using QPSK modulation.

B. Results for SKCE algorithms

We use a system having $G = 1$, $K_d = K_u = 3$, $b_d = b_u = 2$ and $M_B = M_d = M_u = 6$, which is known to be feasible [13], to obtain Fig. 4 and Fig. 5. As can be seen from both figures while the SKCE versions of the algorithms produce results that are very close, the curves don't overlap completely in the manner that results for the naive versions do. Such behavior is expected since the Max-SINR and MMSE equivalence established in Section IV-C holds only for cases where η is set to 0 for beamformer calculation.

As seen from Fig. 4 and Fig. 5, the SKCE versions of the algorithms outperform the naive versions both in terms of sum rate and BER. For example for $\alpha = 1$, $\beta = 10$ at an SNR of 40 dB, MMSE-SKCE has a sum rate improvement of 12.3

bits per channel use, while Max-SINR-SKCE has a gain of 12.1 bits per channel. For the same α and β combination, MMMSE-SKCE achieves a BER of 1×10^{-2} at around 21.9 dB and Max-SINR-SKCE achieves it at 22.1 dB, while the naive version requires approximately 23.6 dB to obtain the same performance. Analogously, for $\alpha = 0.75$, $\beta = 10$ we have a rate gain of 14.2 bits per channel use for MMSE-SKCE and 14.1 bits per channel use for Max-SINR-SKCE. In term of BER for $\alpha = 0.75$, $\beta = 10$, MMSE-SKCE requires approximately 8.8 dB less than the naive version to reach a BER level of 1×10^{-2}, whilst Max-SINR-SKCE requires around 8.5 dB less than Max-SINR-Naive.

Considering the results for MMSE-Naive and Max-SINR-Naive with $\alpha = 0$, $\beta = 0.1$ in Fig. 4 and Fig. 5 it can be noticed that performance initially improves in the region of -10 dB up to around 10 dB, and then starts to degrade until it eventually settles to a steady state value for SNR ≥ 45 dB. For this specification of α and β, the CSI error is quite significant and independent of SNR. In the range of -10 db to 10 dB the power of the leakage is reasonably small, since the power levels we are dealing with are low; this allows for performance improvement across the region. However, once SNR increases beyond 10 dB the interference leakage starts to become more significant, resulting in an interference limited system; this leads to a degradation in performance that eventually settles to a steady state value. Such behavior is avoided by the SKCE version of the algorithms, which also improve the overall performance. In fact for $\alpha = 0$, $\beta = 0.1$ the SKCE algorithms settle at approximately 13.0 bits per channel use above their naive counterparts. Additionally, in terms of BER, MMSE-SKCE and Max-SINR-SKCE both settle at around 2.3×10^{-2}, while the naive versions settle at 1.3×10^{-1}.

C. Determining IA feasibility in multi-cell systems

Next we focus on how the proposed algorithms can be used to give an indication of IA feasibility for FD multi-cell systems with HD users. For example, consider a system having $G = 2$ and $K_d = K_u = 2$ with each user requiring 2 streams, i.e. $b_d = b_u = 2$. We want to determine the antennas required at the BS, M_B, and at the users, $N = M_d = M_u$, to ensure that full DoF equal to $G(K_d b_d + K_u b_u) = 16$ are achievable.

If all nodes have 16 antennas, i.e. $\{M_B = N = 16\}$, the desired number of streams can easily be delivered, however from an achievable DoF perspective this leads to an unnecessarily large number of antennas; with IA we should achieve the same DoF with less antennas. For a HD system, with M_B BS antennas and N user antennas, to deliver 2 streams each to $K_{\text{TOT/cell}} = 4$ across two cells (i.e achieve total DoF of 16), we need $M_B \geq 2(4+p)$ and $N \geq 2(5-p)$ where $p \in \{1, 2, 3, 4\}$ [39]. With $p = 1$ this evaluates to $M_B \geq 10$ and $N \geq 8$, implying that $\{M_B = 10, N = 8\}$ is the minimum number of antennas required to achieve 16 DoF in the HD system.

Moving on to our FD system, we use the proposed multi-cell algorithms from Section V with perfect CSI to obtain the results in Fig. 6. As can be seen results for $\{M_B = N = 16\}$, $\{M_B = 10, N = 8\}$, $\{M_B = 10, N = 7\}$ and $\{M_B = 9, N = 8\}$ have the same slope and achieve full DoF. However for

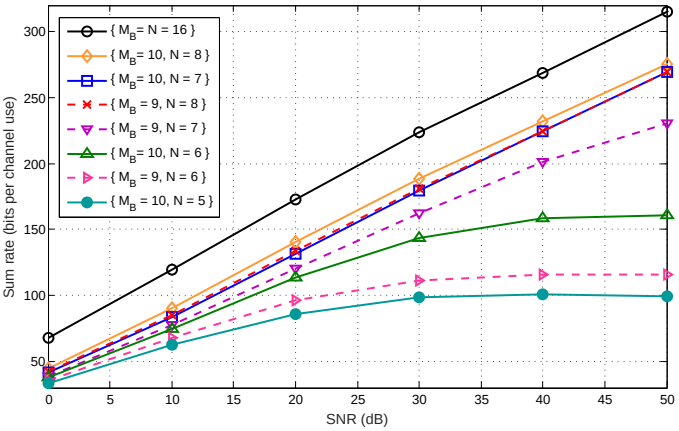

Fig. 6: Sum rate performance obtained by using both Max-SINR-Naive and MMSE-Naive algorithms under perfect CSI conditions for system with $G = 2$, $K_d = K_u = 2$, $b_d = b_u = 2$ and varying antenna numbers.

$\{M_B = 9, N = 7\}$, $\{M_B = 10, N = 6\}$, $\{M_B = 9, N = 6\}$, $\{M_B = 10, N = 5\}$, the sum rate flattens out as SNR increases, indicating that IA is infeasible. Table I relates the feasibility of the various system configurations simulated in Fig. 6 with the properness of the system according to (38). As can be seen systems marked as improper are always infeasible, however systems marked as proper are not necessarily feasible. In fact for $\{M_B = 9, N = 7\}$ and $\{M_B = 10, N = 6\}$, where the properness condition is met with equality, the resulting scenario is proper but infeasible.

TABLE I: Properness and IA feasibility for systems simulated in Fig. 6.

M_B	N	*Properness of system*	*IA Feasibility*
16	16	Proper with $N_v > N_e$	Feasible
10	8	Proper with $N_v > N_e$	Feasible
10	7	Proper with $N_v > N_e$	Feasible
9	8	Proper with $N_v > N_e$	Feasible
9	7	Proper with $N_v = N_e$	Infeasible
10	6	Proper with $N_v = N_e$	Infeasible
9	6	Improper with $N_v < N_e$	Infeasible
10	5	Improper with $N_v < N_e$	Infeasible

Moreover, it can be noticed that results for $\{M_B = 10, N = 8\}$, $\{M_B = 10, N = 7\}$ and $\{M_B = 9, N = 8\}$ obtain very similar rates with a marginal increase for an increasing number of antennas. The rate for $\{M_B = N = 16\}$ is the highest across the whole SNR range; however this rate advantage comes from having a significantly larger number of antennas compared to the other configurations where IA is also feasible.

D. Convergence results

Fig. 7 shows the convergence behavior of the designed IA algorithms. For each scenario plotted we consider an SNR of 10 dB and average the results over 200 channel realizations under perfect CSI. As can be seen for all scenarios the proposed algorithms do indeed converge to a constant value.

VII. Conclusion

The combination of FD technology and IA provides a promising solution to tackle the ever increasing resource demand problem in wireless networks. While the advantages

$$\Delta R_{FD} = \mathbb{E}_{\mathbf{H}}\{R_{\text{ FD TOT}}\} - \mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\{\widehat{R}_{\text{ FD TOT}}\}\}$$
$$= \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\kappa_{k_g^d,g}^2 P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^n|^2}{\sigma^2}\right)\right\} - \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\kappa_{k_g^d,g}^2 P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^n|^2}{\widehat{J}_{k_g^d}+\sigma^2}\right)\right\}\right\}$$
$$+\mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\kappa_{g,k_g^u}^2 P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^n|^2}{\sigma^2}\right)\right\} - \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\kappa_{g,k_g^u}^2 P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^n|^2}{\widehat{J}_{k_g^u}+\sigma^2}\right)\right\}\right\} \tag{47}$$

$$= \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\kappa_{k_g^d,g}^2 P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^n|^2}{\sigma^2}\right)\right\} + \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\kappa_{g,k_g^u}^2 P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^n|^2}{\sigma^2}\right)\right\}$$
$$- \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g^d}+\kappa_{k_g^d,g}^2 P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^n|^2}{\sigma^2}\right)\right\}\right\} + \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g^d}}{\sigma^2}\right)\right\}\right\}$$
$$- \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\widehat{J}_{k_g^u}+\kappa_{g,k_g^u}^2 P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^n|^2}{\sigma^2}\right)\right\}\right\} + \mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\widehat{J}_{k_g^u}}{\sigma^2}\right)\right\}\right\} \tag{48}$$

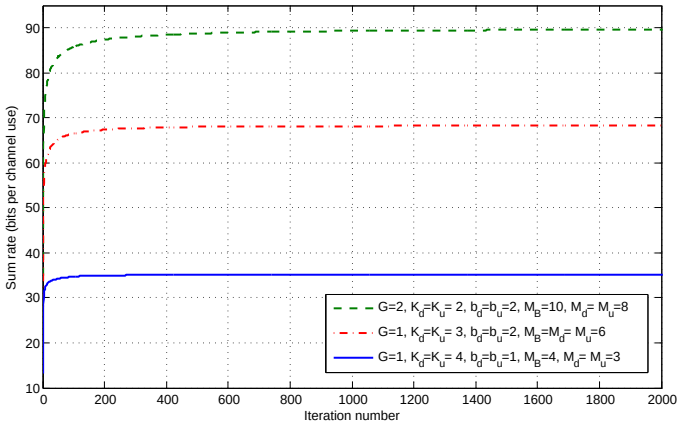

Fig. 7: Sum rate convergence trend averaged over 200 channel realizations for both Max-SINR and MMSE based algorithm designs at an SNR of 10 db, under perfect CSI.

are clear for perfect CSI, it is important to consider imperfect CSI scenarios to obtain a more practical characterization of the system's behavior. Here we apply linear IA in a multi-user multi-cell system with FD BSs and legacy HD users. We start by characterizing the performance losses incurred due to imperfect CSI with respect to the achievable sum rate and DoF. Results show that the way the CSI error behaves with SNR is highly important in determining the overall performance loss trend. When the two parameters are inversely proportional, the rate loss is upper bounded by a derived value and full DoF can be achieved. When the error variance scales with SNR to the power of a negative proper fraction, DoF loss proportional to α is experienced and rate loss is unbounded. We also propose two novel IA algorithms based on MMSE and Max-SINR that are applicable to an FD multi-cell system with HD users. Our designs exploit statistical knowledge of the CSI error and produce unitary beamformers. Moreover, they are shown to be equivalent for cases where η is set to 0, and provide significant performance improvements over the naive designs under imperfect CSI. Additionally, we also derive the proper condition for IA feasibility in the multi-cell system under consideration.

Appendix A
Proof of Theorem 1

Using the sum rate loss definition provided in Section III-A and taking expectations, the mean sum rate loss can be expressed as (47), where for the purpose of this proof we use notation $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$ to distinguish the imperfect CSI beamformers from the perfect CSI ones which we will continue to represent as $\mathbf{u}$ and $\mathbf{v}$. After a number of algebraic manipulations (47) can be further represented as (48). Additionally, since the interference leakage is by definition greater or equal to zero for both DL and UL communication, i.e. $\widehat{J}_{k_g^d} \geq 0$ and $\widehat{J}_{k_g^u} \geq 0$, we can establish the following inequalities.

$$\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\widehat{J}_{k_g^d}+\kappa_{k_g^d,g}^2 P|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^n|^2}{\sigma^2}\right)\right\}\right\}$$
$$\geq \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\kappa_{k_g^d,g}^2 P|\mathbf{u}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\mathbf{v}_{k_g^d}^n|^2}{\sigma^2}\right)\right\} \tag{49}$$

$$\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\widehat{J}_{k_g^u}+\kappa_{g,k_g^u}^2 P|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^n|^2}{\sigma^2}\right)\right\}\right\}$$
$$\geq \mathbb{E}_{\mathbf{H}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\kappa_{g,k_g^u}^2 P|\mathbf{u}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\mathbf{v}_{k_g^u}^n|^2}{\sigma^2}\right)\right\} \tag{50}$$

Using (49) and (50) into (48), and applying Jensen's inequality, results in

$$\Delta R_{\text{ FD}} \leq \sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}}{\sigma^2}\right)$$

$$+\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}}{\sigma^2}\right). \quad (51)$$

Therefore to quantify $\Delta R_{\text{ FD}}$ we need to find expressions for $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}$ and $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}$. Starting with $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}$, having already defined $\widehat{J}_{k_g^d}$ in (9), we can combine it with the channel model from (6) to obtain

$$\begin{aligned}
&\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}\\
&=\sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{k_g^d,j}P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}^{n\,H}_{k_g^d}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g^d,j}+\boldsymbol{\Upsilon}_{k_g^d,j}\right)\widehat{\mathbf{v}}^m_{i_j^d}\right|^2\right\}\\
&+\sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\kappa^2_{k_g^d,i_j^u}P\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}^{n\,H}_{k_g^d}\left(\frac{1}{1+\eta}\widehat{\mathbf{H}}_{k_g^d,i_j^u}+\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\right)\widehat{\mathbf{v}}^m_{i_j^u}\right|^2\right\}.
\end{aligned}$$

Using the IA conditions in (7), particularly $\left[\mathbf{u}^{n\,H}_{k_g^d}\widehat{\mathbf{H}}_{k_g^d,j}\mathbf{v}^m_{i_j^d}=0\ \forall\ n,m,k,i,g,j\ (n,k,g\neq m,i,j)\right]$ and $\left[\mathbf{u}^{n\,H}_{k_g^d}\widehat{\mathbf{H}}_{k_g^d,i_j^u}\mathbf{v}^m_{i_j^u}=0\ \forall\ n,m,k,i,g,j\ \right]$, this can be further simplified to

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}&=\sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{k_g^d,j}P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}^{n\,H}_{k_g^d}\boldsymbol{\Upsilon}_{k_g^d,j}\widehat{\mathbf{v}}^m_{i_j^d}\right|^2\right\}\\
&+\sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\kappa^2_{k_g^d,i_j^u}P\mathbb{E}_{\boldsymbol{\Upsilon}}\left\{\left|\widehat{\mathbf{u}}^{n\,H}_{k_g^d}\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\widehat{\mathbf{v}}^m_{i_j^u}\right|^2\right\}\\
&\overset{(a)}{=}P\frac{\eta}{1+\eta}\left(\sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{k_g^d,j}+\sum_{j=1}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\kappa^2_{k_g^d,i_j^u}\right).
\end{aligned} \quad (52)$$

where (a) follows by applying Lemma 3 from Appendix C.

Following a similar process, for the UL we obtain

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}&=P\frac{\eta}{1+\eta}\left(\sum_{\substack{j=1\\(j,i,m)\neq(g,k,n)}}^{G}\sum_{i=1}^{K_u}\sum_{m=1}^{b_u}\kappa^2_{g,i_j^u}\right.\\
&\left.+\sum_{\substack{j=1\\j\neq g}}^{G}\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{g,j}+\Theta\sum_{i=1}^{K_d}\sum_{m=1}^{b_d}\kappa^2_{g,g}\right).
\end{aligned} \quad (53)$$

Applying (52) and (53) into (51) results in (11) proving the first equation in Theorem 1.

Next, to obtain a more tractable equation for the rate loss we consider a homogeneous pathloss assumption where $\kappa=1$ $\forall k,g,i,j$. Under this assumption the interference leakage for DL and UL is given by (54) and (55) respectively.

$$\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\breve{J}_{k_g^d}\}\}=P\frac{\eta}{1+\eta}\left(GK_db_d-1+GK_ub_u\right). \quad (54)$$

$$\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\breve{J}_{k_g^u}\}\}=P\frac{\eta}{1+\eta}\left(GK_ub_u-1+(G-1+\Theta)K_db_d\right). \quad (55)$$

Applying the results from (54) and (55) into (51), results in

$$\begin{aligned}
\Delta\breve{R}_{\text{ FD}}&\leq\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{1+\eta}\left(GK_db_d-1+GK_ub_u\right)\right)\\
&+\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(1+\frac{P}{\sigma^2}\frac{\eta}{1+\eta}\left(GK_ub_u-1+(G-1+\Theta)K_db_d\right)\right)
\end{aligned}$$

which using $\eta=\beta\rho^{-\alpha}$, can be represented as

$$\begin{aligned}
&\Delta\breve{R}_{\text{ FD}}\\
&\leq GK_db_d\left[\log_2\left(1+(GK_db_d+GK_ub_u-1)\frac{\beta\rho^{1-\alpha}}{1+\beta\rho^{-\alpha}}\right)\right]\\
&+GK_ub_u\left[\log_2\left(1+GK_ub_u-1+(G-1+\Theta)K_db_d)\frac{\beta\rho^{1-\alpha}}{1+\beta\rho^{-\alpha}}\right)\right].
\end{aligned}$$

Finally, letting $P\to\infty$ the asymptotic sum rate loss under homogenous pathloss can be expressed as in (12), proving the second and final equation of Theorem 1.

Appendix B
Proof of Theroem 2

Starting with (13) and replacing $\widehat{R}_{\text{ FD DL}}$ and $\widehat{R}_{\text{ FD UL}}$ with the corresponding expressions from (10), after some further algebraic manipulations we obtain (56), where for the purpose of this proof we use notation $\widehat{\mathbf{u}}$ and $\widehat{\mathbf{v}}$ to distinguish the imperfect CSI beamformers from the perfect CSI ones which we will continue to represent as $\mathbf{u}$ and $\mathbf{v}$. Next, discarding the interference-plus-noise noise terms in A and B and applying Jensen's inequality to C and D results in (57). Additionally, (57) can be expressed as (58). This follows, (a) since for unitary beamformers, analogous to [40, Lemma 2] it can be shown that $|\mathbf{u}^{n\,H}_{k_g^d}\mathbf{H}_{k_g^d,g}\mathbf{v}^n_{k_g^d}|^2$ and $|\widehat{\mathbf{u}}^{n\,H}_{k_g^u}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}^n_{k_g^u}|^2$ are exponentially distributed with both mean and variance one, and (b) by replacing $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^d}\}\}$ and $\mathbb{E}_{\widehat{\mathbf{H}}}\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{\widehat{J}_{k_g^u}\}\}$ with (52) and (53) respectively, and considering the fact that since for the DoF metric $P\to\infty$, then pathloss effects have no implications in the DoF domain.

$$\begin{aligned}
\widehat{D}_{\text{ FD}}&=GK_db_d+GK_ub_u\\
&-\lim_{P\to\infty}\frac{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(P\frac{\eta}{1+\eta}\left(GK_db_d-1+GK_ub_u\right)+\sigma^2\right)}{\log_2 P}\\
&-\lim_{P\to\infty}\frac{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(P\frac{\eta}{1+\eta}(GK_ub_u-1+(G-1+\Theta)K_db_d)+\sigma^2\right)}{\log_2 P}
\end{aligned} \quad (58)$$

Substituting $\eta=\beta\rho^{-\alpha}=\beta P^{-\alpha}\sigma^{2\alpha}$ in (58) and letting $P\to\infty$, we obtain

$$\widehat{D}_{\text{ FD}}=\begin{cases}G(K_db_d+K_ub_u) & \alpha\geq 1\\ \alpha G(K_db_d+K_ub_u) & 0\leq\alpha<1.\end{cases} \quad (59)$$

Finally, using (59) and noting that $D_{\text{ FD}}=G(K_db_d+K_ub_u)$ from (14), we can compute the DoF loss as $\Delta D_{\text{ FD}}=D_{\text{ FD}}-\widehat{D}_{\text{ FD}}$ to obtain (15).

Appendix C
Useful Lemmas

Lemma 1. *[18] Consider $\mathbf{A}\in\mathbb{C}^{M\times N}$ to be a Gaussian matrix with entries that are i.i.d. with mean zero and variance*

$$\widehat{D}_{\text{FD}}$$

$$= \lim_{P\to\infty} \underbrace{\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\widehat{J}_{k_g^d}+\sigma^2+\kappa_{k_g^d,g}^2 P\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}\right|^2\right)\right\}\right\}}{\log_2 P}}_{\text{A}} - \lim_{P\to\infty} \underbrace{\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\widehat{J}_{k_g^d}+\sigma^2\right)\right\}\right\}}{\log_2 P}}_{\text{C}}$$

$$+ \lim_{P\to\infty} \underbrace{\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\widehat{J}_{k_g^u}+\sigma^2+\kappa_{g,k_g^u}^2 P\left|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}\right|^2\right)\right\}\right\}}{\log_2 P}}_{\text{B}} - \lim_{P\to\infty} \underbrace{\frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\widehat{J}_{k_g^u}+\sigma^2\right)\right\}\right\}}{\log_2 P}}_{\text{D}} \tag{56}$$

$$\geq \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\kappa_{k_g^d,g}^2 P\left|\widehat{\mathbf{u}}_{k_g^d}^{n\,H}\mathbf{H}_{k_g^d,g}\widehat{\mathbf{v}}_{k_g^d}^{n}\right|^2\right)\right\}\right\}}{\log_2 P} - \lim_{P\to\infty} \frac{\sum_{g=1}^{G}\sum_{k=1}^{K_d}\sum_{n=1}^{b_d}\log_2\left(\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\widehat{J}_{k_g^d}\right\}\right\}+\sigma^2\right)}{\log_2 P}$$

$$+ \lim_{P\to\infty} \frac{\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\kappa_{g,k_g^u}^2 P\left|\widehat{\mathbf{u}}_{k_g^u}^{n\,H}\mathbf{H}_{g,k_g^u}\widehat{\mathbf{v}}_{k_g^u}^{n}\right|^2\right)\right\}\right\}}{\log_2 P} - \lim_{P\to\infty} \frac{\sum_{g=1}^{G}\sum_{k=1}^{K_u}\sum_{n=1}^{b_u}\log_2\left(\mathbb{E}_{\widehat{\mathbf{H}}}\left\{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{\widehat{J}_{k_g^u}\right\}\right\}+\sigma^2\right)}{\log_2 P} \tag{57}$$

ω, and $\mathbf{b}\in\mathbb{C}^{N\times 1}$ to be a unit-norm vector independent of $\mathbf{A}$, then $\mathbb{E}_{\mathbf{A}}\{\mathbf{A}\mathbf{b}\mathbf{b}^H\mathbf{A}^H\}=\omega I$.

Lemma 2. $\mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\widehat{\mathbf{H}}_{g,i^u}\mathbf{v}_{i^u}^m\mathbf{v}_{i^u}^{m\,H}\boldsymbol{\Upsilon}_{g,i^u}^H\} = \mathbb{E}_{\widehat{\mathbf{H}},\boldsymbol{\Upsilon}}\{\boldsymbol{\Upsilon}_{g,i^u}\mathbf{v}_{i^u}^m\mathbf{v}_{i^u}^{m\,H}\widehat{\mathbf{H}}_{g,i^u}^H\} = 0\ \forall\ m,i.$

Proof. The precoders are computed based on $\widehat{\mathbf{H}}$, therefore from the CSI error definition in Section II-B, they are directly independent of $\boldsymbol{\Upsilon}$. □

Lemma 3. $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\mathbf{u}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\mathbf{v}_{i_j^d}^m|^2\} = \mathbb{E}_{\boldsymbol{\Upsilon}}\{|\mathbf{u}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\mathbf{v}_{i_j^u}^m|^2\} = \eta/(1+\eta)\ \ \forall\ \ k,g,n,i,\ \ j,m.$

Proof. Let us first focus on $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\mathbf{u}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\mathbf{v}_{i_j^d}^m|^2\}$. From the CSI error model definition in Section II-B, $\widehat{\mathbf{H}}$ and $\boldsymbol{\Upsilon}$ are independent. Moreover, $\mathbf{u}_{k_g^d}^n$ and $\mathbf{v}_{i_j^d}^m$ are computed based on $\widehat{\mathbf{H}}$, implying they are independent of $\boldsymbol{\Upsilon}$. Since $\boldsymbol{\Upsilon}$ is Gaussian and bi-unitarily invariant [41], $\mathbf{u}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\mathbf{v}_{i_j^d}^m\ \ \forall\ k,g,n,i,\ j,m$ is a Gaussian random variable with mean zero and variance $\eta/(1+\eta)$, thus $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\mathbf{u}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,j}\mathbf{v}_{i_j^d}^m|^2\} = \eta/(1+\eta)$. Applying a similar argument based on $\mathbf{u}_{k_g^d}^n$ and $\mathbf{v}_{i_j^u}^m$, $\mathbb{E}_{\boldsymbol{\Upsilon}}\{|\mathbf{u}_{k_g^d}^{n\,H}\boldsymbol{\Upsilon}_{k_g^d,i_j^u}\mathbf{v}_{i_j^u}^m|^2\} = \eta/(1+\eta)$. □

Lemma 4. *[42] For matrix $\mathbf{A}\in\mathbb{C}^{M\times M}$ and vector $\mathbf{b}\in\mathbb{C}^{M\times 1}$*

$$(\mathbf{A}-\mathbf{b}\mathbf{b}^H)^{-1}\mathbf{b} = \frac{\mathbf{A}^{-1}\mathbf{b}}{1-\mathbf{b}^H\mathbf{A}^{-1}\mathbf{b}}\ .$$

REFERENCES

[1] P. Aquilina and T. Ratnarajah, "Beamformer design for interference alignment in full-duplex cellular networks with imperfect CSI," in *Proc. IEEE ICC*, May 2017, pp. 1-6.

[2] M. Jain *et al.*, "Practical, real-time, full duplex wireless," in *Proc. MOBICOM*, Sep. 2011, pp. 301-312.

[3] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of FD wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296-4307, Dec. 2012.

[4] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. USENIX*, Aug. 2014, pp. 359-372.

[5] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637-1652, Sept. 2014.

[6] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2017-2046, Fourthquarter 2015.

[7] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, pp. 1-136, Jun. 2011.

[8] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," in *Proc. IEEE ITW*, Sep. 2013, pp. 1-5.

[9] J. Bai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of multi-user MIMO full-duplex network," in *Proc. IEEE ISIT*, Jun. 2015, pp. 864-868.

[10] S. H. Chae and S. H. Lim, "Degrees of freedom of cellular networks: Gain from full-duplex operation at a base station," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 4048-4053.

[11] S.-W. Jeon, S. H. Chae, and S. H. Lim, "Degrees of freedom of full-duplex multi-antenna cellular networks," in *Proc. IEEE ISIT,* Jun. 2015, pp. 869-873.

[12] M. Yang, S.-W. Jeon, and D. K. Kim, "On the degrees of freedom of full-duplex cellular networks," in *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2314-2326, Apr. 2017.

[13] K. Kim, S. W. Jeon, and D. K. Kim, "The feasibility of interference alignment for full-duplex MIMO cellular networks," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1500-1503, Sep. 2015.

[14] M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Farajzadeh-Tehrani, "Scaling wireless full-duplex in multi-cell networks," in *Proc. IEEE INFOCOM*, Apr. 2015, pp. 1751-1759.

[15] R. T. Krishnamachari and M. K. Varanasi, "Interference alignment under limited feedback for MIMO interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3908-3917, Aug. 2013.

[16] O. E. Ayach and R. W. Heath, "Interference alignment with analog channel state feedback," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 626-636, Feb. 2012.

[17] R. Tresch and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. IEEE ICC*, Jun. 2009.

[18] P. Aquilina and T. Ratnarajah, "Performance analysis of IA techniques in the MIMO IBC with imperfect CSI," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1259-1270, Apr. 2015.

[19] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *Proc. Asilomar Conf. Signal Syst. Comput.*, Nov. 2009, pp. 1106-1110.

[20] K. Gomadam, V. R. Cadambe and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309-3322, June 2011.

[21] R. K. Mungara and A. Lozano, "Interference surge in full-duplex wireless systems," in *Proc. Asilomar Conf.*, Nov. 2015.

[22] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3933-3938, Aug. 2009.

[23] S. Goyal, P. Liu, and S. Panwar, "User selection and power allocation in full duplex multi-cell networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2408-2422, Mar. 2017.

[24] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Multi-objective optimization for robust power efficient and secure full-duplex wireless communication systems," *IEEE Trans. on Wireless Commun.*, vol. 15, no. 8, pp. 5511-5526, Aug. 2016.

[25] J. Maurer, J. Jaldén, D. Seethaler, and G. Matz, "Vector perturbation precoding revisited," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 315-328, Jan. 2011.

[26] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203-2214, May 2006.

[27] S. M. Kay, *Fundamentals of statistical signal processing, volume I: Estimation theory*, Prentice-Hall, 1993.

[28] J. Zhu, J. Liu, X.She, and L. Chen, "Investigation on precoding techniques in E-UTRA and proposed adaptive precoding scheme for MIMO systems," in *Proc. APCC*, Oct. 2008, pp. 1-5.

[29] R. de Francisco and D. T. M. Slock, "An optimized unitary beamforming technique for MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 990-1000, Mar. 2010.

[30] S. M. Razavi and T. Ratnarajah, "Adaptive LS- and MMSE-Based beamformer design for multiuser MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 132-144, Jan. 2016.

[31] I. Santamaria, O. Gonzalez, R. W. Heath, and S. W. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1-6.

[32] S. W. Peters and R. W. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 206-218, Jan. 2011.

[33] B. Zhuang, R. A. Berry, and M. L. Honig, "Interference alignment in MIMO cellular networks," in *Proc. IEEE ICASSP*, May 2011, pp. 3356-3359.

[34] J. Schreck and G. Wunder, "Distributed Iinterference alignment in cellular systems: Analysis and algorithms," in *Proc. European Wireless Conference*, Apr. 2011, pp. 1-8.

[35] C. Wilson and V. Veeravalli, "A convergent version of the Max SINR algorithm for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2952-2961, June 2013.

[36] J. Park, Y. Sung, and H. Vincent Poor, "On beamformer design for multiuser MIMO interference channels." [Online]. Available: https://arxiv.org/pdf/1011.6121.pdf

[37] C. M. Yetis, T. Gou, S. A. Jafar and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks" *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771-4782, Sept. 2010.

[38] M. Razaviyayn, G. Lyubeznik and Z. Q. Luo, "On the degrees of freedom achievable through interference alignment in a MIMO Interference channel," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 812-821, Feb. 2012.

[39] T. Liu and C. Yang, "On the feasibility of linear interference alignment for MIMO interference broadcast channels with constant coefficients," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2178-2191, May 2013.

[40] S. M. Razavi and T. Ratnarajah, "Performance analysis of interference alignment under CSI mismatch," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4740-4748, Nov. 2014.

[41] A. M. Tulino and S. Verdú, "Random Matrix Theory and Wireless Communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1-182, Jun. 2004.

[42] H. V. Henderson and S. R. Searle, "On deriving the inverse of a sum of matrices", *SIAM Review.*, vol. 23, no. 1, pp. 53-60, Jan. 1981.

Paula Aquilina (S'14) received her B.Eng.(Hons.) degree in Electrical Engineering from the University of Malta in 2010, her M.Sc. in Signal Processing and Communications in 2013 and her PhD. degree in Engineering (Digital Communications) in 2017, both from the University of Edinburgh. Her main area of research is wireless communications, with particular focus on interference management and network information theory.

Tharmalingam Ratnarajah (A'96-M'05- SM'05) is currently with the Institute for Digital Communications, University of Edinburgh, Edinburgh, UK, as a Professor in Digital Communications and Signal Processing and the Head of Institute for Digital Communications. His research interests include signal processing and information theoretic aspects of 5G and beyond wireless networks, full-duplex radio, mmWave communications, random matrices theory, interference alignment, statistical and array signal processing and quantum information theory. He has published over 325 publications in these areas and holds four U.S. patents. He was the coordinator of the FP7 projects ADEL (3.7M€) in the area of licensed shared access for 5G wireless networks and HARP (3.2M€) in the area of highly distributed MIMO and FP7 Future and Emerging Technologies projects HIATUS (2.7M€) in the area of interference alignment and CROWN (2.3M€) in the area of cognitive radio networks. Dr Ratnarajah is a Fellow of Higher Education Academy (FHEA), U.K..