Optimal Sum-DoF of the K-user MISO BC with Current and Delayed Feedback

Paul de Kerret, David Gesbert, Jingjing Zhang, and Petros Elia

Abstract

This work identifies the optimal sum degrees-of-freedom (DoF) of the K-user MISO Broadcast Channel (BC), in the presence of delayed channel-state information at the transmitter (CSIT) and additional current CSIT of imperfect precision. In a setting where the current channel estimation error scales — in the high SNR regime — as $\text{SNR}^{-\alpha}$ for $\alpha \in [0,1]$, the optimal sum-DoF takes the simple form $(1-\alpha)K/(\sum_{k=1}^{K}\frac{1}{k})+\alpha K$, and is the result of a novel scheme which deviates from existing efforts as it i) digitally combines interference, ii) uses MAT-type alignment where all symbols are decoded irrespective of order, and iii) utilizes a hierarchical quantizer whose output is distributed across rounds in a way that minimizes unwanted interference. These jointly deliver, for the first time, the elusive simultaneous scaling of both MAT-type delayed-CSI gains as well as ZF-type imperfect-CSI gains.

I. Introduction

In the wireless BC, feedback accuracy and timeliness can crucially affect performance, but are also notoriously difficult to obtain. In terms of accuracy, it is well known that increasing feedback quality can elevate performance, from that of TDMA (sum DoF of 1), to the maximum possible interference-free performance of K sum DoF. As the recent result in [1] tells us, having imperfect instantaneous CSIT with an estimation error that scales (in the high-power P setting) as $P^{-\alpha}$ ($\alpha \in [0,1]$), can allow, using basic ZF precoding techniques, for an optimal sum-DoF of $1+(K-1)\alpha$.

On the other hand, when perfect-accuracy CSIT is obtained in a delayed manner — in the sense that the CSI is fed back with a delay that can exceed the channel coherence period — then, using more involved, retrospective, *MAT*-type space-time alignment [2], one can surprisingly get substantial DoF gains, reaching a sum-DoF of $\frac{K}{H_K}$, which nicely scales with K, approximately as $\frac{K}{\ln(K)}$.

This interplay between performance and feedback timeliness-and-quality, has sparked a plethora of works that considered a variety of feedback mechanisms with delayed and imperfect CSIT. Such works can be found in [3]–[5], and in [6] which — for the two-user MISO BC setting — studied the case where the CSIT can alternate between perfect, delayed (completely outdated), and non-existent (see also [7]).

An interesting approach came with the work in [8] which considered a feedback mechanism that offered a combination of imperfect-quality current (instantaneously available) CSIT, together with additional (perfect-accuracy) delayed CSIT. In this same setting — which reflected different scenarios, including that of using predictions to get an estimate of the current state of a time-correlated channel — the channel estimation error of the current channel state was assumed to scale in power as $P^{-\alpha}$, for some CSIT quality exponent $\alpha \geq 0$. Additional work — within the context of the BC — came in [9], [10] which established the maximal DoF in a two-user MISO BC scenario, as well as in [11], [12] which considered the case of imperfect-quality delayed CSIT. Additionally work can be found in [13], [14] which considered the broad setting of any-time any-quality feedback, and in [15], [16] which studied the two-user MIMO BC (and IC); all for the two-user case. This general challenge of dealing with imperfect feedback has also sparked very recent interest, with different publications that include [17]–[22].

This paper was submitted in part to ITW 2016.

The authors are with the Mobile Communications Department, EURECOM, Sophia Antipolis, France (email: {Paul.DeKerret,David.Gesbert,Jingjing.Zhang,Petros.Elia}@eurecom.fr).

D. Gesbert and P. de Kerret are supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (Agreement no. 670896). P. Elia is supported by the ANR project ECOLOGICAL-BITS-AND-FLOPS.

A. Simultaneous scaling of MAT and ZF gains

For the more general case of the K-user BC, again with joint delayed and imperfect-current CSIT, very little is known. For the particular case considered here, a general outer bound was provided in [23], and efforts to reach this bound can be found in [24]. The main goal has remained to secure simultaneous scaling of MAT-type gains (that exploit delayed CSIT), and ZF gains (that exploit imperfect-quality current CSI). To date, this has remained an elusive open problem, and any instance of providing such simultaneous gains was either limited to the 2-user case, or — as in the case of the scheme in [23] — resulted in MAT-type DoF gains that saturated at 2. This elusive open problem is resolved here, by inventing a new scheme, referred to as the Q-MAT scheme, that combines different new ingredients that jointly allow for MAT and ZF components to optimally coexist. Combined with the outer bound in [23], the achieved DoF establishes the optimal sum-DoF, which is here shown to be equal to $\alpha K + (1-\alpha)K/(\sum_{k=1}^{K}\frac{1}{k})$.

B. Notation

For $\mathcal{C}(P)$ denoting the sum capacity [25] of the MISO BC considered, we will place emphasis on the high-SNR degree of freedom approximation

$$\mathrm{DoF}^{\star} \triangleq \lim_{P\to\infty} \frac{\mathcal{C}(P)}{\log_2(P)}. \tag{1}$$

We will use the notation $H_K \triangleq \sum_{i=1}^{K}\frac{1}{i}$ to represent the K-th harmonic number. $\mathbb{Z}$ will represent the integers, $\mathbb{Z}^+$ the positive integers, $\mathbb{R}^+$ the positive real numbers, $\mathbb{C}$ the complex numbers, $\binom{n}{k}$ the n-choose-k operator, and $\oplus$ the bitwise XOR operation. We will use $\mathcal{K} \triangleq \{1, 2, \cdots, K\}$. If $\mathcal{S} \subset \mathcal{K}$ is a set, then $\bar{\mathcal{S}}$ will denote $\mathcal{K}\backslash\mathcal{S}$, and $|\mathcal{S}|$ will denote its cardinality. Complex vectors will be denoted by lower-case bold font. For any vector $\mathbf{x}$, we will use $\{\mathbf{x}\}_i$, $\|\mathbf{x}\|^2$ and $\mathbf{x}^{\mathrm{H}}$ to respectively denote the ith element of the vector, its magnitude-squared, and its conjugate transpose. We will also use $\doteq$ to denote *exponential equality*, i.e., we write $f(P) \doteq P^B$ to denote $\lim_{P\to\infty} \dfrac{\log(f(P))}{\log(P)} = B$. We write $\mathcal{N}_{\mathbb{C}}(0,\sigma^2)$ to denote the complex Gaussian distribution of zero mean and variance σ^2.

C. System Model

1) K-User MISO Broadcast Channel: This work considers the K-User MISO BC with fading, where the transmitter (TX) — which is equipped with M antennas ($M \geq K$) — serves K single-antenna receivers (RX). At any time t, the signal received at RX $k \in \mathcal{K}$, can be written as

$$y_k[t] = \boldsymbol{h}_k^{\mathrm{H}}[t]\mathbf{x}[t] + n_k[t] \tag{2}$$

where $\boldsymbol{h}_k^{\mathrm{H}}[t] \in \mathbb{C}^{1\times M}$ represents the channel to user k at time t, where $\mathbf{x}[t] \in \mathbb{C}^M$ represents the transmitted signal, and where $n_k[t] \in \mathbb{C}$ represents the additive noise at RX k, where this noise is distributed as $\mathcal{N}_{\mathbb{C}}(0,1)$, independently of the channel and of the transmitted signal. Furthermore, the transmitted signal $\mathbf{x}[t]$ fulfills the average asymptotic power constraint $\mathbb{E}[\|\mathbf{x}[t]\|^2] \doteq P$. The channel is assumed to be drawn from a continuous ergodic distribution such that all the channel matrices and all their sub-matrices are almost surely full rank.

2) Perfect delayed CSIT and imperfect current CSIT: Our CSIT model builds on the delayed CSIT model introduced in [2] and generalized in [8] to account for the availability of an imperfect estimate of the current channel state. For ease of exposition, we will here adopt the fast-fading channel model, and will assume that at any time t, the TX has access to the delayed CSI (with perfect accuracy) of all previous channel realizations up to time $t-1$, as well as has imperfect knowledge of the current channel state. Each current estimate $\hat{\boldsymbol{h}}_k^{\mathrm{H}}[t]$ for each channel $\boldsymbol{h}_k^{\mathrm{H}}[t]$, comes with an estimation error

$$\tilde{\boldsymbol{h}}_k^{\mathrm{H}}[t] = \boldsymbol{h}_k^{\mathrm{H}}[t] - \hat{\boldsymbol{h}}_k^{\mathrm{H}}[t] \tag{3}$$

whose entries are i.i.d. $\mathcal{N}_{\mathbb{C}}(0, P^{-\alpha})$ with power $P^{-\alpha}$ for some parameter $\alpha \in [0,1]$, which we refer to as the *CSIT quality exponent*, and which is used to parameterize the accuracy of the current CSIT[1]. We also adhere to

[1]Note that from a DoF perspective, we can restrict ourselves to $\alpha \in [0,1]$, since an estimation/quantization error with power scaling as P^{-1} ($\alpha = 1$), is essentially perfect. Similarly an estimation error with power scaling as P^0 ($\alpha = 0$), offers no DoF gains over the case of having no CSIT (cf. [1], see also [23]).

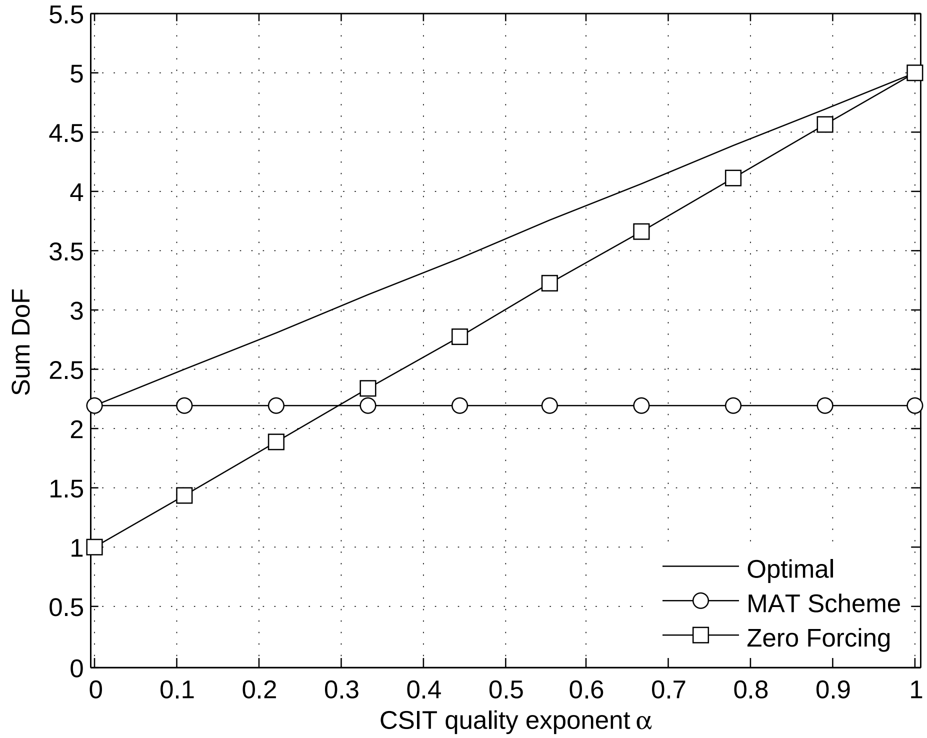

Fig. 1: Sum DoF achieved as a function of the CSIT scaling exponent α for $K = 5$ users.

the common convention (see [2], [8], [10]) of assuming perfect and global knowledge of channel state information at the receivers (perfect global CSIR), where the receivers know all channel states and all estimates. Finally we make the common assumption that the channel $\boldsymbol{h}_k^{\mathrm{H}}[t]$ is independent of all previous channel estimates and channel estimation errors, when conditioned on $\hat{\boldsymbol{h}}_k^{\mathrm{H}}[t]$.

II. Main Results

We proceed directly with the main result.

Theorem 1. *In the K-user MISO BC ($M \geq K$) with perfect delayed CSIT and α-quality current CSIT, the optimal sum DoF is*

$$\mathrm{DoF}^{\star}(\alpha) = (1-\alpha)\frac{K}{\sum_{k=1}^{K}\frac{1}{k}} + \alpha K. \tag{4}$$

Proof. The scheme that achieves this performance, is described in the next sections. The optimality follows from the fact that this scheme's performance matches the DoF outer bound in [23]. □

Figure 1 shows the optimal DoF for the $K = 5$ case, and compares this performance to the MAT-only and ZF-only DoF performance. We proceed to describe the scheme, following the exposition in [2], describing first the simpler case of $K = 2$, and then proceeding with the general K case.

III. Scheme for the $K = 2$ user case

As stated above, we begin with the scheme description for the two-user case, which offers intuition on all aspects of the construction[2].

[2]While a bit premature at this stage, we hasten to note that this exposition of the two-user case, nicely accentuates the difference between our scheme and previous efforts [9], [10] which — albeit optimal for the two user case — had to deviate from the MAT-type canonical structure in order to accommodate for the ZF component, thus making it difficult to extend to higher dimensions.

A. Encoding and Transmission

The transmission follows a structure that incorporates the MAT multi-destination multi-layer scheme [2] and is also divided into two phases (for the 2-user case), where phase 1 corresponds to the transmission of order-1 data symbols (meant for one user at a time) and spans 2 time slots (TS), while phase 2 corresponds to the transmission of order-2 data symbols (meant for both users simultaneously) and spans 1 TS.

A first deviation from existing schemes can be found in the fact that the scheme requires several *rounds*. Each round consists of two phases. The decoding of each round requires input from the next round, and the large number of rounds guarantees that the termination step does not have a notable effect on the overall DoF performance.

We describe the transmission for an arbitrary round N, and we will use the shorthand notation $\bullet^{(R_N)}$ to denote the fact that the index of the round is equal to N. The particularities of the first and last rounds will be clarified later on.

1) Phase 1: Phase 1 spans two time slots, $t_1^{(R_N)}$ and $t_2^{(R_N)}$. For $t = t_1^{(R_N)}$, the transmitted signal is given by

$$\mathbf{x}[t]=\mathbf{V}[t]\boldsymbol{m}[t]+\boldsymbol{v}_2^{\mathrm{ZF}}[t]a_2[t]+\sum_{k=1}^{2}\boldsymbol{v}_k^{\mathrm{ZF}}[t]s_k[t] \tag{5}$$

where

- $\boldsymbol{m}[t]\in\mathbb{C}^2$ is a vector containing two Q-MAT *data symbols* meant for user 1, each carrying $(1-\alpha)\log_2(P)$ bits, where the first symbol is allocated full power $\mathbb{E}\left[|\{\boldsymbol{m}[t]\}_1|^2\right]\doteq P$, while the second symbol is allocated lesser power $\mathbb{E}\left[|\{\boldsymbol{m}[t]\}_2|^2\right]\doteq P^{1-\alpha}$. Furthermore $\mathbf{V}[t]\in\mathbb{C}^{2\times 2}$ is defined as
$$\mathbf{V}[t]\triangleq\begin{bmatrix}\boldsymbol{v}_1^{\mathrm{ZF}}[t] & \boldsymbol{u}_1\end{bmatrix} \tag{6}$$
where $\boldsymbol{v}_1^{\mathrm{ZF}}[t]\in\mathbb{C}^2$ is the unit-norm ZF beamformer aimed at user 1 (i.e., which is orthogonal to the current estimate of the channel to user 2), while $\boldsymbol{u}_1\in\mathbb{C}^2$ is a unit-norm vector that is randomly drawn and isotropically distributed.
- $a_2[t]\in\mathbb{C}$ is a so-called *auxiliary symbol* meant for user 2, carrying $\min(1-\alpha,\alpha)\log_2(P)$ bits (generally from previous interfering terms), and allocated full power $\mathbb{E}\left[|a_2[t]|^2\right]\doteq P$.
- $s_k[t]\in\mathbb{C},\ k\in\{1,2\}$ are *ZF data symbols* meant for user k, each carrying $\alpha\log_2(P)$ bits and each having power $\mathbb{E}\left[|s_k[t]|^2\right]\doteq P^{\alpha}$.

Upon omitting the noise realizations for simplicity, the received symbols during the same $t = t_1^{(R_N)}$, can be written as

$$\begin{aligned} y_1[t] &= \underbrace{\boldsymbol{h}_1^{\mathrm{H}}[t]\mathbf{V}[t]\boldsymbol{m}[t]}_{\doteq P}+\underbrace{\boldsymbol{h}_1^{\mathrm{H}}[t]\boldsymbol{v}_2^{\mathrm{ZF}}[t]a_2[t]}_{\doteq P^{1-\alpha}}+\underbrace{z_1[t]}_{\doteq P^{\alpha}}\\ y_2[t] &= \underbrace{\boldsymbol{h}_2^{\mathrm{H}}[t]\boldsymbol{v}_2^{\mathrm{ZF}}[t]a_2[t]}_{\doteq P}+\underbrace{i_2[t]}_{\doteq P^{1-\alpha}}+\underbrace{z_2[t]}_{\doteq P^{\alpha}} \end{aligned} \tag{7}$$

where

$$i_2[t]\triangleq\underbrace{\boldsymbol{h}_2^{\mathrm{H}}[t]\mathbf{V}[t]\boldsymbol{m}[t]}_{\doteq P^{1-\alpha}} \tag{8}$$

and where for any $t\in\mathbb{Z}^+$,

$$z_k[t]\triangleq\underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t]\boldsymbol{v}_k^{\mathrm{ZF}}[t]s_k[t]}_{\doteq P^{\alpha}}+\underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t]\boldsymbol{v}_{\bar{k}}^{\mathrm{ZF}}[t]s_{\bar{k}}[t]}_{\doteq P^{0}},\ k\in\{1,2\}. \tag{9}$$

In the above equations, underneath each summand, we describe the asymptotic approximation of the power of the corresponding term.

For $t = t_2^{(R_N)}$, the transmission is described by the above equations, after exchanging the indices of the two users ($1\leftrightarrow 2$), and thus the received signal during $t = t_2^{(R_N)}$, takes the form

$$\begin{aligned} y_1[t] &= \underbrace{\boldsymbol{h}_1^{\mathrm{H}}[t]\boldsymbol{v}_1^{\mathrm{ZF}}[t]a_1[t]}_{\doteq P^{1-\alpha}}+\underbrace{i_1[t]}_{\doteq P^{1-\alpha}}+\underbrace{z_1[t]}_{\doteq P^{\alpha}}\\ y_2[t] &= \underbrace{\boldsymbol{h}_2^{\mathrm{H}}[t]\mathbf{V}[t]\boldsymbol{m}[t]}_{\doteq P}+\underbrace{\boldsymbol{h}_2^{\mathrm{H}}[t]\boldsymbol{v}_1^{\mathrm{ZF}}[t]a_1[t]}_{\doteq P^{1-\alpha}}+\underbrace{z_2[t]}_{\doteq P^{\alpha}} \end{aligned} \tag{10}$$

where now

$$i_1[t_2^{(R_N)}] \triangleq \underbrace{\boldsymbol{h}_1^{\mathrm{H}}[t_2^{(R_N)}]\mathbf{V}[t_2^{(R_N)}]\boldsymbol{m}[t_2^{(R_N)}]}_{\doteq P^{1-\alpha}}. \tag{11}$$

a) Interference quantization: At the end of phase 1, the TX can use its delayed CSIT to compute $i_2[t_1^{(R_N)}]$ and $i_1[t_2^{(R_N)}]$, on the basis of which it generates an order-2 data symbol carrying $(1-\alpha)\log_2(P)$ bits. Let us first consider the more involved case of $\alpha \leq \frac{1}{2}$, and let us consider a specifically designed quantization process which relies on the following lemma.

Lemma 1. *Let y be a random variable with a density, and with variance P^{β_1}, $\beta_1 \geq 0$, and let $Q_{\beta_1,\beta_2}(y)$ be the quantized output of a properly-scaled integer quantizer of rate $(\beta_1 - \beta_2)\log_2(P)$ bits for any positive $\beta_2 \leq \beta_1$.*

Then, for any random variable n (with density) and a variance that does not scale above P^{β_2}, there exists such a quantizer such that

$$\lim_{P\to\infty} \Pr\{Q_{\beta_1,\beta_2}(y+n) = Q_{\beta_1,\beta_2}(y)\} = 1. \tag{12}$$

Proof. The detailed proof is given in the appendix. □

Our aim is to quantize the interference term $i_2[t_1^{(R_N)}]$ (which has power that scales as $P^{1-\alpha}$). Considering the above quantizer Q_{β_1,β_2}, and setting $\beta_1 = 1-\alpha, \beta_2 = \alpha$, offers us a quantized version $Q_{1-\alpha,\alpha}(i_2[t_1^{(R_N)}])$ that carries $(1-2\alpha)\log_2(P)$ bits, and which guarantees (by construction; see the Appendix) that the resulting quantization noise $n_2[t_1^{(R_N)}]$ has power that scales as P^{α}. The idea is to then follow this with a second quantization step, where now the aforementioned quantization noise $n_2[t_1^{(R_N)}]$ is itself re-quantized using any standard optimal quantizer with $\alpha\log_2(P)$ bits, which is known [25] to guarantee quantization noise (from the second quantization) that scales as P^0. For $\hat{n}_2[t_1^{(R_N)}]$ denoting the quantized version of $n_2[t_1^{(R_N)}]$, we get the final combined estimate in the form

$$\hat{i}_2[t_1^{(R_N)}] \triangleq Q_{1-\alpha,\alpha}(i_2[t_1^{(R_N)}]) + \hat{n}_2[t_1^{(R_N)}] \tag{13}$$

where in the above the addition is over the complex numbers, and where we ommited signals at the level of the noise floor.

For the easier case where $\alpha \geq \frac{1}{2}$, the interference term is quantized using $(1-\alpha)\log_2(P)$ bits using a standard quantizer, which gives again $\hat{i}_2[t_1^{(R_N)}]$, which is again guaranteed (cf. [25]) to have quantization noise that scales in P^0.

The same quantization process is applied to $i_1[t_2^{(R_N)}]$ to obtain $\hat{i}_1[t_2^{(R_N)}]$ with the same rate and quantization-noise properties. All these quantized bits will be placed in the auxiliary data symbols of the next round, as we now describe.

b) Generation of auxiliary data symbols for phase 1 of round $N+1$: The auxiliary data symbol $a_2[t_1^{(R_{N+1})}]$ will carry

$$\begin{cases} a_2[t_1^{(R_{N+1})}] \longleftarrow \hat{n}_2[t_1^{(R_N)}] & ,\ \alpha \leq \frac{1}{2} \\ a_2[t_1^{(R_{N+1})}] \longleftarrow \hat{i}_2[t_1^{(R_N)}] & ,\alpha \geq \frac{1}{2} \end{cases} \tag{14}$$

and similarly $a_1[t_2^{(R_{N+1})}]$ will carry the following quantization bits

$$\begin{cases} a_1[t_2^{(R_{N+1})}] \longleftarrow \hat{n}_1[t_2^{(R_N)}] & ,\alpha \leq \frac{1}{2} \\ a_1[t_2^{(R_{N+1})}] \longleftarrow \hat{i}_1[t_2^{(R_N)}] & ,\alpha \geq \frac{1}{2}. \end{cases} \tag{15}$$

During the first round, all these auxiliary symbols are initialized to zero.

c) Generation of Q-MAT *data symbols for phase 2 of round N:* During the second phase of round N, using delayed CSIT, the TX generates the following order-2 Q-MAT data symbol

$$m[t_{12}^{(R_N)}] \longleftarrow (\hat{i}_2[t_1^{(R_N)}] \oplus \hat{i}_1[t_2^{(R_N)}]) \tag{16}$$

that carries a XOR combination $\hat{i}_2[t_1^{(R_N)}] \oplus \hat{i}_1[t_2^{(R_N)}]$ of the interference experienced at the two users.

2) Phase 2: Phase 2 consists of one TS $t = t_{12}^{(R_N)}$, during which

$$\mathbf{x}[t] = \boldsymbol{v}[t]m[t] + \sum_{k=1}^{2} \boldsymbol{v}_k^{\mathrm{ZF}}[t]s_k[t] \tag{17}$$

where

- $m[t] \in \mathbb{C}$ is an order-2 Q-MAT data symbol which carries $(1-\alpha)\log_2(P)$ bits that originate from the previous phase (of the same round), and which is meant for both users. The symbol is allocated full power $\mathbb{E}\left[|m[t]|^2\right] \doteq P$.
- $s_k[t],\ k \in \{1,2\}$ is a ZF data symbol meant for user k, carrying $\alpha \log_2(P)$ bits, and having power $\mathbb{E}\left[|s_k[t]|^2\right] \doteq P^{\alpha}$.

Upon omitting the noise realizations, the received symbols during TS $t = t_{12}^{(R_N)}$ take the form

$$\begin{aligned} y_1[t] &= \underbrace{\boldsymbol{h}_1^{\mathrm{H}}[t]\boldsymbol{v}[t]m[t]}_{\doteq P} + \underbrace{z_1[t]}_{\doteq P^{\alpha}} \\ y_2[t] &= \underbrace{\boldsymbol{h}_2^{\mathrm{H}}[t]\boldsymbol{v}[t]m[t]}_{\doteq P} + \underbrace{z_2[t]}_{\doteq P^{\alpha}}. \end{aligned} \tag{18}$$

B. Decoding (achievability proof by induction)

We now turn to the decoding part, which here — for purposes of clarify of exposition — will be assumed to commence after the end of transmission in all rounds and all phases. We will show that each user can decode all its desired data symbols. The proof has to be done by induction due to the fact that the auxiliary data symbols contain information coming from the previous round.

Let us consider without loss of generality the decoding at user 1. Our induction statement is that if $a_2[t_1^{(R_{N+1})}]$ is decoded at user 1, user 1 can decode all its corresponding Q-MAT data symbols ($\boldsymbol{m}[t_1^{(R_N)}]$ and $m[t_{12}^{(R_N)}]$) and ZF data symbols ($s_1[t_1^{(R_N)}], s_1[t_2^{(R_N)}], s_1[t_{12}^{(R_N)}]$) of round N, as well as $a_2[t_1^{(R_N)}]$.

After initializing (by setting to zero) all auxiliary data symbols of the first round, we directly proceed to consider an arbitrary round N. As part of the induction and the initialization, we consider $a_2[t_1^{(R_N)}]$ to be already decoded at user 1.

a) Decoding of phase 2: Before decoding phase 1 of round N, we directly apply successive decoding on (19), to retrieve the Q-MAT and ZF data symbols of phase 2 of round N. This is indeed achieved because, as we see from (19), the SINR of the Q-MAT data symbol is in the order of $P^{1-\alpha}$, which matches the scaling of the symbol's rate. We then can also follow with the decoding of the ZF data symbols.

b) Decoding the interference in phase 1 (round N): As a first step, receiver 1 uses the signals received during round $N+1$ to decode $a_1[t_2^{(R_{N+1})}]$. This is possible because, as we see from (10), the scaling of the SINR matches the scaling of the rate. The contents of $a_1[t_2^{(R_{N+1})}]$ depend on whether $\alpha \geq \frac{1}{2}$ or $\alpha \leq \frac{1}{2}$.

- If $\alpha \geq \frac{1}{2}$, user 1 has obtained $\hat{i}_1[t_2^{(R_N)}]$ from $a_1[t_2^{(R_{N+1})}]$.
- If $\alpha \leq \frac{1}{2}$, user 1 has obtained $\hat{n}_1[t_2^{(R_N)}]$ from $a_1[t_2^{(R_{N+1})}]$. To recover the quantized interference $\hat{i}_1[t_2^{(R_N)}]$, it is necessary for user 1 to also decode $Q_{1-\alpha,\alpha}\left(i_1[t_2^{(R_N)}]\right)$.

  This quantized interference can be obtained by quantization of the received signal $y_1[t_2^{(R_N)}]$ using the quantizer $Q_{1-\alpha,\alpha}$ that we used for encoding. Indeed, following Lemma 1, we see that

  $$Q_{1-\alpha,\alpha}\left(y_1[t_2^{(R_N)}]\right) = Q_{1-\alpha,\alpha}\left(i_1[t_2^{(R_N)}]\right) \tag{19}$$

  will be satisfied with probability that approaches 1. Thus, user 1 can decode $Q_{1-\alpha,\alpha}\left(y_1[t_2^{(R_N)}]\right)$, and then use $\hat{n}_1[t_2^{(R_N)}]$ to obtain $\hat{i}_1[t_2^{(R_N)}]$.

Thus in both cases, user 1 can obtain $\hat{i}_1[t_2^{(R_N)}]$.

c) Decoding of the Q-MAT *data symbols of phase* 1*:* We see that from phase 2, using (16), user 1 has decoded

$$m[t_{12}^{(R_N)}] = (\hat{i}_1[t_2^{(R_N)}] \oplus \hat{i}_2[t_1^{(R_N)}]). \tag{20}$$

Using $\hat{i}_1[t_2^{(R_N)}]$, user 1 obtains $\hat{i}_2[t_1^{(R_N)}]$, and thus now user 1 has knowledge (up to the noise level) of the following two components

$$\begin{aligned}&\underbrace{\boldsymbol{h}_1^{\mathrm{H}}[t_1^{(R_N)}]\mathbf{V}[t_1^{(R_N)}]\boldsymbol{m}[t_1^{(R_N)}] + \boldsymbol{h}_1^{\mathrm{H}}[t_1^{(R_N)}]\boldsymbol{v}_2^{\mathrm{ZF}}[t_1^{(R_N)}]a_2[t_1^{(R_N)}]}_{\doteq P} + \underbrace{z_1[t_1^{(R_N)}]}_{\doteq P^{\alpha}}\\&\underbrace{\boldsymbol{h}_2^{\mathrm{H}}[t_1^{(R_N)}]\mathbf{V}[t_1^{(R_N)}]\boldsymbol{m}[t_1^{(R_N)}]}_{\doteq P^{1-\alpha}}.\end{aligned} \tag{21}$$

By induction, $a_2[t_1^{(R_N)}]$ is assumed to be already decoded at user 1, such that its contribution to the received signal in (21) can be removed. Consequently, user 1 has obtained two signals with a SINR scaling in $P^{1-\alpha}$, and can decode its two corresponding data symbols in $\boldsymbol{m}[t_1^{(R_N)}]$.

d) Decoding of the ZF data symbols at user 1*:* We have shown that, during $t = t_1^{(R_N)}$ and $t = t_{12}^{(R_N)}$, user 1 could decode its corresponding Q-MAT data symbols, and thus that it is possible for user 1 to use successive decoding to decode the ZF data symbols. However, the Q-MAT data symbols sent during $t_2^{(R_N)}$ are only decoded at user 2, and thus successive decoding can not be used directly here. For this we use the fact that user 1 has reconstructed $i_1[t_2^{(R_N)}]$ (up to bounded noise). Therefore user 1 can remove the interference generated by the Q-MAT data symbols that were meant for user 2, and can thus decode its own ZF data symbols (eq. (10)).

e) Decoding of the auxiliary data symbols of round $N+1$*:* In order to conclude the inductive step, it remains to prove that it is possible for user 1 to decode $a_2[t_1^{(R_{N+1})}]$. This result follows directly from the definition of the auxiliary data symbol in (14). Indeed, the auxiliary data symbol of $t = t_1^{(R_{N+1})}$ depends only on the Q-MAT data symbols in $\boldsymbol{m}[t_1^{(R_N)}]$. These data symbols have already been decoded at user 1, who can thus decode $a_2[t_1^{(R_{N+1})}]$. This concludes the inductive step.

C. Calculation of the DoF

In this particular $K = 2$ setting, one round spans 3 TS during which 2 Q-MAT data symbols of rate $(1-\alpha)\log_2(P)$ bits are transmitted to each user. Furthermore, one ZF data symbol of rate $\alpha\log_2(P)$ bits is transmitted to each user during each TS. In the last round, termination is achieved by sending only auxiliary symbols, which induces a small loss in DoF performance, which though becomes negligible as the number of rounds increases. The resulting DoF is thus

$$\mathrm{DoF}^{\text{Q-MAT}} = \frac{4(1-\alpha)+6\alpha}{3} \tag{22}$$

as in Theorem 1.

IV. Scheme for the K-user case: Encoding and Transmission

We now describe the Q-MAT scheme for an arbitrary number of users. The transmission follows the multi-phase structure in MAT [2], with K phases (per round), where phase j aims to communicate order-j data symbols, and where each such symbol is meant for j users. For clarity, our description will trace the description in Section III.C of [2]. In contrast to the original MAT scheme, the order-j data symbols are now expected to be decoded at all the j corresponding receivers, the nature of the transmitted signals is different, the decoding process is different, and the scheme spreads over many rounds. We proceed to describe the transmission during round N, and assume that the transmissions up to round $N-1$ have already been concluded.

A. Phase $j = 1, 2, \dots, K-1$

Phase j transmits $(K-j+1)\binom{K}{j}$ order-j data symbols (each carrying $(1-\alpha)\log_2(P)$ bits), and in the process it generates $j\binom{K}{j+1}$ order-$(j+1)$ data symbols of rate $(1-\alpha)\log_2(P)$ bits, as well as generates auxiliary symbols to be sent during phase j of round $N+1$.

This phase j, during any round N, has $\binom{K}{j}$ time-slots $t_{\mathcal{S}}^{(R_N)}$, each dedicated to a subset $\mathcal{S}$ of users, for all $\mathcal{S} \subset \mathcal{K}$ of size $|\mathcal{S}| = j$.

1) Transmission at phase j: During $t = t_{\mathcal{S}}^{(R_N)}$, the transmit signal is given by

$$\mathbf{x}[t_{\mathcal{S}}^{(R_N)}] = \mathbf{V}[t_{\mathcal{S}}^{(R_N)}]\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}] + \sum_{\ell \in \bar{\mathcal{S}}} \boldsymbol{v}_\ell^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}]a_\ell[t_{\mathcal{S}}^{(R_N)}] + \sum_{k=1}^{K} \boldsymbol{v}_k^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}]s_k[t_{\mathcal{S}}^{(R_N)}] \tag{23}$$

where

- $\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}] \in \mathbb{C}^{K-j+1}$ is a vector containing $(K-j+1)$ Q-MAT order-j data symbols meant for the users in $\mathcal{S}$. Each symbol carries $(1-\alpha)\log_2(P)$ bits. The first symbol has full power $\mathbb{E}\left[|\{\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}]\}_1|^2\right] \doteq P$, while the rest have power $\mathbb{E}\left[|\{\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}]\}_i|^2\right] \doteq P^{1-\alpha}, \forall i \in \{2, \dots, K-j+1\}$. Furthermore, $\mathbf{V}[t_{\mathcal{S}}^{(R_N)}] \in \mathbb{C}^{K \times (K-j+1)}$ is defined as
$$\mathbf{V}[t_{\mathcal{S}}^{(R_N)}] \triangleq \begin{bmatrix} \boldsymbol{v}_{\mathcal{S}}^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}] & \mathbf{U}_j \end{bmatrix} \tag{24}$$
where $\boldsymbol{v}_{\mathcal{S}}^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}] \in \mathbb{C}^K$ is a unit-norm ZF precoder that is orthogonal to all current-CSI estimates of the channels of the users in $\bar{\mathcal{S}}$, and where $\mathbf{U}_j \in \mathbb{C}^{K \times (K-j)}$ is a randomly chosen, isotropically distributed unitary matrix.
- $a_k[t_{\mathcal{S}}^{(R_N)}] \in \mathbb{C},\ k \in \bar{\mathcal{S}}$ is an auxiliary data symbol meant for user k, having rate $\min(1-\alpha, \alpha)\log_2(P)$ bits and power $\mathbb{E}\left[|a_k[t_{\mathcal{S}}^{(R_N)}]|^2\right] \doteq P$.
- $s_k[t_{\mathcal{S}}^{(R_N)}],\ k \in \mathcal{K}$ is a ZF data symbol meant for user k, having rate $\alpha\log_2(P)$ bits and power $\mathbb{E}\left[|s_k[t_{\mathcal{S}}^{(R_N)}]|^2\right] \doteq P^\alpha$.

Again for $t = t_{\mathcal{S}}^{(R_N)}$, for user $k \in \mathcal{S}$, the received signal takes the form

$$y_k[t_{\mathcal{S}}^{(R_N)}] = \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t_{\mathcal{S}}^{(R_N)}]\mathbf{V}[t_{\mathcal{S}}^{(R_N)}]\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P} + \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t_{\mathcal{S}}^{(R_N)}]\sum_{\ell \in \bar{\mathcal{S}}} \boldsymbol{v}_\ell^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}]a_\ell[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P^{1-\alpha}} + \underbrace{z_k[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P^\alpha} \tag{25}$$

where for $k \in \mathcal{K}$, we have

$$z_k[t] \triangleq \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t]\boldsymbol{v}_k^{\mathrm{ZF}}[t]s_k[t]}_{\doteq P^\alpha} + \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t]\sum_{\ell=1, \ell \neq k}^{K} \boldsymbol{v}_\ell^{\mathrm{ZF}}[t]s_\ell[t]}_{\doteq P^0}. \tag{26}$$

For user $k \in \bar{\mathcal{S}}$, the received signal is written as

$$y_k[t_{\mathcal{S}}^{(R_N)}] = \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t_{\mathcal{S}}^{(R_N)}]\boldsymbol{v}_k^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}]a_k[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P} + \underbrace{i_k[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P^{1-\alpha}} + \underbrace{z_k[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P^\alpha} \tag{27}$$

where we have introduced the short-hand notation $i_k[t_{\mathcal{S}}^{(R_N)}]$ for $k \in \bar{\mathcal{S}}$ as

$$i_k[t_{\mathcal{S}}^{(R_N)}] \triangleq \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t_{\mathcal{S}}^{(R_N)}]\mathbf{V}[t_{\mathcal{S}}^{(R_N)}]\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P^{1-\alpha}} + \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t_{\mathcal{S}}^{(R_N)}]\sum_{\ell \in \bar{\mathcal{S}}, \ell \neq k} \boldsymbol{v}_\ell^{\mathrm{ZF}}[t_{\mathcal{S}}^{(R_N)}]a_\ell[t_{\mathcal{S}}^{(R_N)}]}_{\doteq P^{1-\alpha}}. \tag{28}$$

For the above, it is easy to see (cf. [26]) that ZF beamforming which uses α-quality current CSIT, reduces the scaling of the interference power by a multiplicative factor of $P^{-\alpha}$.

2) Generation of new data symbols carrying information on the interference: The generation of the data symbols that are transmitted in phase $j+1$ of round N and in phase j of round $N+1$ from the interference generated during phase j of round N is one of the key ingredients of our scheme. We now consider that the transmissions of phase j have ended for every possible set $\mathcal{S} \subset \mathcal{K}$ for which $|\mathcal{S}| = j$.

a) Interference quantization: At the end of phase j, using delayed CSIT, the transmitter reconstructs $i_k[t_{\mathcal{S}}^{(R_N)}],\ k \in \bar{\mathcal{S}}$, which have power that scales as $P^{1-\alpha}$. The next step depends on the value of α.

- If $\alpha \leq \frac{1}{2}$, the transmitter uses the quantizer of Lemma 1 to obtain
$$\hat{i}_k[t_{\mathcal{S}}^{(R_N)}] = Q_{1-\alpha,\alpha}(i_k[t_{\mathcal{S}}^{(R_N)}]) \tag{29}$$
that comes with a residual quantization noise $n_k[t_{\mathcal{S}}^{(R_N)}]$, which — by design, and directly from the proof of Lemma 1 — has power scaling in P^{α}. Then the transmitter quantizes the quantization noise $n_k[t_{\mathcal{S}}^{(R_N)}]$, to get $\hat{n}_k[t_{\mathcal{S}}^{(R_N)}]$, with $\alpha \log_2(P)$ bits, leaving a residual quantization noise that only scales in P^0 (cf. [25]). Finally the transmitter combines (over the complex numbers) the two quantized estimates, to get a total estimate
$$\hat{i}_k[t_{\mathcal{S}}^{(R_N)}] \triangleq Q_{1-\alpha,\alpha}(i_k[t_{\mathcal{S}}^{(R_N)}]) + \hat{n}_k[t_{\mathcal{S}}^{(R_N)}]. \tag{30}$$
- If $\alpha \geq \frac{1}{2}$, the interference term $i_k[t_{\mathcal{S}}^{(R_N)}],\ k \in \bar{\mathcal{S}}$ is quantized using any $(1-\alpha)\log_2(P)$ bit quantizer, to directly give $\hat{i}_k[t_{\mathcal{S}}^{(R_N)}]$ with quantization noise that scales in P^0 (cf. [25]).

b) Generation of auxiliary data symbols for phase j of round $N+1$: The above quantized estimates of the interference will be placed in *auxiliary* data symbols to be transmitted during round $N+1$. Depending on the value of α, the auxiliary symbol $a_k[t_{\mathcal{S}}^{(R_{N+1})}]$ is loaded as follows
$$\begin{cases} a_k[t_{\mathcal{S}}^{(R_{N+1})}] \longleftarrow \hat{n}_k[t_{\mathcal{S}}^{(R_N)}] & ,\alpha \leq \frac{1}{2} \\ a_k[t_{\mathcal{S}}^{(R_{N+1})}] \longleftarrow \hat{i}_k[t_{\mathcal{S}}^{(R_N)}] & ,\alpha \geq \frac{1}{2} \end{cases} \tag{31}$$
where the differentiation into two cases reflects that the rate of the auxiliary data symbol $a_k[t_{\mathcal{S}}^{(R_N)}]$ will be set equal to $\min(1-\alpha,\alpha)\log_2(P)$ bits. We note that for the first round, all auxiliary data symbols are set to zero.

c) Generation of Q-MAT *data symbols for phase $j+1$ of round N:* Let us consider an arbitrary set $\mathcal{P} \subset \mathcal{K}$ with $|\mathcal{P}| = j+1$, and let us denote its elements as
$$\mathcal{P} \triangleq \{p_1, \dots, p_{j+1}\}. \tag{32}$$
The j order-$(j+1)$ data symbols[3] $m_\ell[t_{\mathcal{P}}^{(R_N)}], \ell \in \{1,\dots,j\}$ *for phase $j+1$ of round N* are then defined as
$$m_\ell[t_{\mathcal{P}}^{(R_N)}] \triangleq \left(\hat{i}_{p_\ell}[t_{\mathcal{P}\setminus p_\ell}^{(R_N)}] \oplus \hat{i}_{p_{\ell+1}}[t_{\mathcal{P}\setminus p_{\ell+1}}^{(R_N)}]\right), \qquad \forall \ell \in \{1,\dots,j\}. \tag{33}$$
Each of the above $2j$ components $\{\hat{i}_{p_\ell}[t_{\mathcal{P}\setminus p_\ell}^{(R_N)}],\ \hat{i}_{p_{\ell+1}}[t_{\mathcal{P}\setminus p_{\ell+1}}^{(R_N)}]\}_{\ell\in\{1,\dots,j\}}$ has already been received at the different receivers, as some form of interference. In the next phase, the transmitter will recreate the j different linear combinations $m_\ell,\ \ell \in \{1,\dots,j\}$ and transmit them.

B. Phase K

Phase K is a broadcasting phase where a fully common message (meant for all users) is transmitted. During the single time-slot $t = t_{\mathcal{K}}^{(R_N)}$ of this phase, the transmitted signal takes the form
$$\mathbf{x}[t] = \boldsymbol{v}[t]m[t] + \sum_{k=1}^{K} \boldsymbol{v}_k^{\mathrm{ZF}}[t]s_k[t] \tag{34}$$
where

- $m[t] \in \mathbb{C}$ is a Q-MAT order-K data symbol, having rate $(1-\alpha)\log_2(P)$ and power $\mathbb{E}\left[|m[t]|^2\right] \doteq P$, and where $\boldsymbol{v}[t] \in \mathbb{C}^K$ is a random unit-norm vector.
- $s_k[t],\ k \in \mathcal{K}$ is a ZF data symbol meant for user k, having rate $\alpha\log_2(P)$ and power $\mathbb{E}\left[|s_k[t]|^2\right] \doteq P^{\alpha}$.

For the same $t = t_{\mathcal{K}}^{(R_N)}$, each user $k \in \mathcal{K}$ receives
$$y_k[t] = \underbrace{\boldsymbol{h}_k^{\mathrm{H}}[t]\boldsymbol{v}[t]m[t]}_{\doteq P} + \underbrace{z_k[t]}_{\doteq P^{\alpha}}. \tag{35}$$

In this phase, there is no auxiliary data symbol.

[3] See further down for a clarification on having to repeat phases to accumulate enough symbols. This is done exactly as in MAT, and it is transparent to the scheme and the performance here.

C. Remark: repetition of the phases

It is important to note that vector $\boldsymbol{m}[t_{\mathcal{S}}^{(R_N)}] \in \mathbb{C}^{K-j}$ ($|\mathcal{S}| = j+1$) contains $K-j$ data symbols. To cover the gap from the fact that we have only generated j order-$(j+1)$ data symbols $m_\ell[t_{\mathcal{P}}^{(R_N)}], \ell = 1, \ldots, j$ ($|\mathcal{P}| = j$), one must simply repeat phase j $\frac{K!}{j}$ times, exactly as in the original MAT scheme [2]. This is automatically accounted for in the DoF calculation, as we will note later.

V. Scheme for the K-user case: Decoding and DoF

A. Decoding: Proof by Induction

We now consider decoding, and for simplicity assume that all transmissions of all phases and all rounds, have been completed[4]. The proof is done by induction. Before making the inductive statement, we introduce the following set of decoded symbols

$$\mathcal{A}_{j,k}^{(R_N)} \triangleq \left\{ a_\ell[t_{\mathcal{S}}^{(R_N)}] \ : \ \forall \mathcal{S} \subset \mathcal{K}, \forall \ell \in \bar{\mathcal{S}}, |\mathcal{S}| = j, k \in \mathcal{S} \right\}. \tag{36}$$

For any round N and phase $j < K$, our induction statement is as follows:
If

- (i) Phase $j+1$ up to phase K of round N have been successfully decoded
- (ii) Any user k has decoded the set $\mathcal{A}_{j,k}^{(R_N)}$

then

- (i) Each user can decode all the Q-MAT order-j data symbols and its corresponding ZF data symbols transmitted during phase j of round N
- (ii) User k can decode the set $\mathcal{A}_{j,k}^{(R_{N+1})}$.

After initializing the auxiliary data symbols of the first round to zero, each user k can decode $\mathcal{A}_{j,k}^{(R_1)}$ (in (32)). Thus we proceed with the inductive step for an arbitrary round N and phase $j < K$, where by induction, it holds that all Q-MAT order-$(j+1)$ data symbols have been successfully decoded at the corresponding users, which means that user k has received the data symbols $m_\ell[t_{\mathcal{T}}^{(R_N)}], \ell = 1, \ldots, j$, for any set $\mathcal{T}$ for which $|\mathcal{T}| = j+1$ and for which $k \in \mathcal{T}$.

In addition, still by induction, each user k has decoded the set $\mathcal{A}_{j,k}^{(R_N)}$.

d) Decoding of the desired auxiliary data symbols: As a first step, user k uses the signal received during round $N+1$ to decode $a_k[t_{\mathcal{W}}^{(R_{N+1})}], \forall \mathcal{W} \subset \mathcal{K}, |\mathcal{W}| = j, k \in \bar{\mathcal{W}}$. Indeed, it can be seen from (27) that the scaling of the SINR matches the scaling of the rate. Again we differentiate between the cases $\alpha \geq \frac{1}{2}$ and $\alpha \leq \frac{1}{2}$, in terms of what we place inside the auxiliary data symbols.

- If $\alpha \leq \frac{1}{2}$, user k has decoded $\hat{n}_k[t_{\mathcal{W}}^{(R_N)}]$ for every set $\mathcal{W} \subset \mathcal{K}$ for which $k \in \bar{\mathcal{W}}$ and for which $|\mathcal{W}| = j$. To recover the quantized interference $\hat{i}_k[t_{\mathcal{W}}^{(R_N)}]$, it is necessary for user k to obtain $Q_{1-\alpha,\alpha}(i_k[t_{\mathcal{W}}^{(R_N)}])$ (see (30)), and this is achieved by quantizing the received signal $y_k[t_{\mathcal{W}}^{(R_N)}]$ using the quantizer $Q_{1-\alpha,\alpha}$. The idea is, as we see from Lemma 1, that this quantizer is invariant to addition of an independent noise scaling in P^α, and thus, in our case — with probability that approaches 1 — we get

$$Q_{1-\alpha,\alpha}\left(y_k[t_{\mathcal{W}}^{(R_N)}]\right) = Q_{1-\alpha,\alpha}\left(i_k[t_{\mathcal{W}}^{(R_N)}]\right). \tag{37}$$

- If $\alpha \geq \frac{1}{2}$, user k has decoded directly from the auxiliary data symbols the quantized interference $\hat{i}_k[t_{\mathcal{W}}^{(R_N)}]$ for every set $\mathcal{W}$ with $|\mathcal{A}| = j$ such that $k \in \bar{\mathcal{W}}$.

In both cases, user k obtains $\hat{\hat{i}}_k[t_{\mathcal{W}}^{(R_N)}], \forall \mathcal{W} \subset \mathcal{K}, |\mathcal{W}| = j, k \in \bar{\mathcal{W}}$.

[4] It will become clear that the data symbols of round N can be decoded as soon as round $N+1$ has ended. Thus, the delay in the decoding of one data symbol does not increase with the number of rounds.

e) Information at user k at the end of phase $j+1$: We now describe what are the data symbols available at user k after successfully decoding phase $j+1$.

Note that for any set $\mathcal{T}, |\mathcal{T}| = j+1$ for which $k \in \mathcal{T}$, we can write $\mathcal{T} = \{k, \mathcal{W}_{\mathcal{T},k}\}$ with $|\mathcal{W}_{\mathcal{T},k}| = j$ and $k \in \bar{\mathcal{W}}_{\mathcal{T},k}$. Thus from the previous paragraph, user k then knows $\hat{i}_k[t^{(R_N)}_{\mathcal{W}_{\mathcal{T},k}}]$. Using the order-$(j+1)$ data symbols $m_\ell[t^{(R_N)}_{\mathcal{T}}], \ell = 1, \dots, j$, user k is able to obtain all the quantized interference terms forming the order-$(j+1)$ data symbols, i.e., $\hat{i}_\ell[t^{(R_N)}_{\mathcal{W}_{\mathcal{T},\ell}}], \forall \ell \in \mathcal{T}$ with $\ell \neq k$ and $\mathcal{T} = \{\ell, W_{\mathcal{T},\ell}\}$, where again we note that $k \in \mathcal{W}_{\mathcal{T},\ell}$.

Considering all the decoded order-$(j+1)$ data symbols, user k has then decoded the set $\mathcal{O}^j_k$ defined as

$$\mathcal{O}^j_k \triangleq \left\{ \left\{ \hat{i}_\ell[t^{(R_N)}_{\mathcal{W}\backslash\ell}] \right\}_{\ell \in \mathcal{W}} \ : \ \forall \mathcal{W} \subset \mathcal{K}, |\mathcal{W}| = j+1, k \in \mathcal{W} \right\}. \tag{38}$$

f) Decoding of the Q-MAT *data symbols at user k:* Let us now consider an arbitrary user k and an arbitrary set $\mathcal{S} \subset \mathcal{K}, |\mathcal{S}| = j, k \in \mathcal{S}$. We will show that user k is able to decode its corresponding Q-MAT order-j data symbols in $\boldsymbol{m}[t^{(R_N)}_{\mathcal{S}}]$. For that purpose, the user needs $K - j + 1$ observations with a SINR scaling in $P^{1-\alpha}$. These observations will be:

$$y_k[t^{(R_N)}_{\mathcal{S}}] \tag{39}$$

$$i_\ell[t^{(R_N)}_{\mathcal{S}}] \qquad , \forall \ell \in \bar{\mathcal{S}}. \tag{40}$$

Indeed, we can rewrite the above quantized interference terms as $\hat{i}_\ell[t^{(R_N)}_{\mathcal{T}\backslash\ell}]$ with $\mathcal{T} = \{\mathcal{S}, \ell\}$ and $k \in \mathcal{W}$. Thus, all the quantized interference terms $\hat{i}_\ell[t^{(R_N)}_{\mathcal{S}}], \forall \ell \in \bar{\mathcal{S}}$ are part of $\mathcal{O}^j_k$. Hence, user k is able to obtain all the observations in (40), up to a quantization noise that has power that scales as P^0. However, the Q-MAT data symbols are interfered by the auxiliary data symbols $a_\ell[t^{(R_N)}_{\mathcal{S}}]$ (see (25) and (28)), but by induction, these data symbols are known at user k, and thus their interference can be removed. Consequently, user k is able to decode its corresponding Q-MAT data symbols in $\boldsymbol{m}[t^{(R_N)}_{\mathcal{S}}]$.

g) Decoding of the ZF data symbols at user k: With the Q-MAT order-j data symbols decoded, it is possible for user k to decode the ZF data symbols $s_k[t^{(R_N)}_{\mathcal{S}}]$ for all the time slots $t^{(R_N)}_{\mathcal{S}}$ for which $k \in \mathcal{S}$ ($|\mathcal{S}| = j$). When $k \in \bar{\mathcal{S}}$, user k cannot decode the Q-MAT data symbols transmitted, as it does not have sufficiently many observations. However, the user has decoded $\hat{i}_\ell[t^{(R_N)}_{\mathcal{S}}]$ for all sets $\mathcal{S} \subset \mathcal{K}$ with $|\mathcal{S}| = j$ and $k \in \bar{\mathcal{S}}$. Therefore it can remove the interference created by the Q-MAT data symbols up to the noise floor also for these TS, and decode its corresponding ZF data symbol $s_k[t^{(R_N)}_{\mathcal{S}}]$.

h) Decoding of the auxiliary data symbols of round $N+1$: In order to conclude the inductive step, it remains to prove that it is possible for user k to decode the auxiliary data symbols $a_\ell[t^{(R_{N+1})}_{\mathcal{S}}], \forall \ell \in \bar{\mathcal{S}}, |\mathcal{S}| = j, k \in \mathcal{S}$. This follows directly from the definition of the auxiliary data symbols in (31). Indeed, these auxiliary data symbols at round $N+1$ are a function of the auxiliary data symbols at round N ($a_\ell[t^{(R_N)}_{\mathcal{S}}]$ for $\ell \in \bar{\mathcal{S}}$, which are assumed to be known by induction), and the Q-MAT data symbols in $\boldsymbol{m}[t^{(R_N)}_{\mathcal{S}}]$ which have been decoded at user k. Therefore, user k is able to decode these auxiliary data symbols. This concludes the inductive step.

B. Calculation of the DoF

We first note that the last round is dedicated to transmitting only auxiliary data symbols, and that — if the number of rounds is sufficiently large — this DoF loss is negligible. Thus focusing on just one round, we note that we have followed closely the structure in different phases of the MAT scheme, so to calculate our DoF we first scale the MAT DoF by a factor of $(1-\alpha)$ to account for our reduced rate of the Q-MAT data symbols, and then note that during each TS, we additionally send $\alpha \log_2(P)$ bits to each user. Adding these two parts together, provides immediately the sum-DoF expression in Theorem 1, and concludes the proof.

VI. Conclusion

The work has provided the first ever communication scheme which, for the general K-user MISO BC, manages to simultaneously and optimally exploit both delayed and imperfect-quality current CSIT. This is achieved by providing a new way of jointly incorporating MAT-type alignment (based on delayed CSIT), and ZF-type separation (using imperfect-quality current CSIT). In addition to resolving a theoretical open problem, the Q-MAT scheme is designed to adapt to CSI considerations that span both timeliness and precision.

Appendix

Recall that y is a random variable with a density, and with variance P^{β_1}, $\beta_1 \geq 0$, and that n is a random variable with density and with variance that does not scale above P^{β_2}. To design the properly-scaled integer quantizer $Q_{\beta_1,\beta_2}(y)$ which will have rate $(\beta_1-\beta_2)\log_2(P)$ bits for any positive $\beta_2 \leq \beta_1$, let $\eta \triangleq \sqrt{P^{1-\alpha}}$ and let $Z_n[\imath] \triangleq \{a+bi : a,b \in [-\eta,\eta] \cap \mathbb{Z}\}$, with cardinality $|Z_n[i]| = \lfloor P^{1-\alpha} \rfloor$. Now let $\theta \triangleq \sqrt{L_1 P^{\alpha}}$, for some $L_1 \in \mathbb{R}^+$, where $\lim_{P\to\infty} \frac{L_1}{P^{\epsilon}} = 0, \forall \epsilon > 0$. θ is chosen to be arbitrarily large such that

$$\lim_{P\to\infty} \frac{\Pr(|q+n|^2 \geq L_1 P)}{P^{\epsilon}} = 0, \quad \forall q \in \theta \cdot Z_n[i]$$

and it is used to regulate the spacing (and the minimum distance) of the grid $\theta \cdot Z_n[i]$

Let $\mathcal{A} \triangleq \{x \in \mathbb{C} \ : \ \min_{q\in\theta\cdot Z_n[\imath]} |x-q| \geq \frac{(\sqrt{L_1}-\sqrt{L_2})P^{\frac{\alpha}{2}}}{2}\}$ for some $L_2 \gg 1, L_2 \ll L_1$, where L_2 is sufficiently large such that

$$\lim_{P\to\infty} \frac{\Pr(|n| > \sqrt{L_2 P^{\alpha}})}{P^{\epsilon}} = 0, \forall \epsilon > 0.$$

In the above, $\sqrt{L_2 P^{\alpha}}$ represents the width of region $\mathcal{A}$, which represents the y region that stands to be substantially affected by the addition of n (in terms of the quantization outcome).

Let $\hat{y} = Q_{1,\alpha}(y)$ and $\hat{y}_n = Q_{1,\alpha}(y+n)$, where $\hat{y}, \hat{y}_n \in \theta \cdot Z_n[\imath]$. Then we have

$$\begin{aligned}
\Pr(\hat{y} \neq \hat{y}_n) &\leq \Pr(\text{err}|y \in \mathcal{A}) \cdot \Pr(y \in \mathcal{A}) + \Pr(\text{err}|y \notin \mathcal{A}) \cdot \Pr(y \notin \mathcal{A}) + \Pr(|y+n|^2 \geq L_1 P) \\
&\leq \Pr(y \in \mathcal{A}) + \Pr(|n|^2 > L_2 P^{\alpha} + \Pr(|y+n|^2 > L_1 P) && (41)\\
&\leq \Pr(y \in \mathcal{A}) + \epsilon_2 && (42)\\
&\leq \epsilon_3 && (43)
\end{aligned}$$

where $\epsilon_2, \epsilon_3 (\epsilon_2 \leq \epsilon_3)$ can be made arbitrarily small. The last step is due to the fact that the volume of $\mathcal{A}$ can be made vanishingly small compared to P, and from the assumption of having well-behaved random variables y and n, that for the general case, guarantees that the measure concentration is not infinitely higher in some regions than in others. This concludes the proof of the lemma.

References

[1] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT," 2014. [Online]. Available: http://arxiv.org/abs/1403.1541

[2] M. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.

[3] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of two-user and certain three-user MIMO broadcast channel with delayed CSI," Dec. 2011, submitted to *IEEE Trans. Inf. Theory*, available on arXiv:1101.0306.

[4] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of three-user MIMO broadcast channel with delayed CSIT," Jul. 2011.

[5] N. Lee and R. W. Heath, "Not too delayed CSIT achieves the optimal degrees of freedom," in *Proc. Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012.

[6] R. Tandon, S. A. Jafar, S. Shamai (Shitz), and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO BC," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106–4128, 2013.

[7] B. Rassouli, C. Hao, and B. Clerckx, "DoF analysis of the K-user MISO broadcast channel with hybrid CSIT," in *Proc. IEEE International Conference on Communications (ICC)*, 2015.

[8] M. Kobayashi, S. Yang, D. Gesbert, and X. Yi, "On the degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2012.

[9] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.

[10] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1084–1087, Jul. 2012.

[11] J. Chen and P. Elia, "Can imperfect delayed CSIT be as useful as perfect delayed CSIT? DoF analysis and constructions for the BC," in *Proc. Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012.

[12] J. Chen, S. Yang, and P. Elia, "On the fundamental feedback-vs-performance tradeoff over the MISO-BC with imperfect and delayed CSIT," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2013.

[13] J. Chen and P. Elia, "MISO broadcast channel with delayed and evolving CSIT," Jul. 2013.

[14] ——, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8336–8356, Dec. 2013.

[15] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degrees of freedom region of temporally correlated MIMO networks with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 494–514, Jan. 2014.

[16] J. Chen and P. Elia, "Symmetric two-user MIMO BC with evolving feedback," in *Proc. Information Theory and Applications Workshop (ITA)*, 2014.

[17] M. Torrellas, A. Agustin, and J. Vidal, "Retrospective interference alignment for the MIMO interference broadcast channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, June 2015.

[18] B. Rassouli, C. Hao, and B. Clerckx, "DoF analysis of the MIMO broadcast channel with alternating/hybrid CSIT," *IEEE Trans. Inf. Theory.*, vol. 62, no. 3, March 2016.

[19] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *CoRR*, 2016. [Online]. Available: http://arxiv.org/abs/1602.09028

[20] B. C. B. Hao, Chenxi; Rassouli, "Achievable DoF regions of MIMO networks with imperfect CSIT," 2016. [Online]. Available: http://arxiv.org/abs/1603.07513

[21] A. Bracher and M. A. Wigger, "Feedback and partial message side-information on the semideterministic broadcast channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, June 2015.

[22] S. Lashgari, R. Tandon, and S. Avestimehr, "Three-user MISO broadcast channel: How much can CSIT heterogeneity help?" in *Proc. IEEE International Conference on Communications (ICC)*, June 2015, pp. 4187–4192.

[23] P. de Kerret, X. Yi, and D. Gesbert, "On the degrees of freedom of the K-user time correlated broadcast channel with delayed CSIT," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2013.

[24] Y. Luo, T. Ratnarajah, and A. K. Papazafeiropoulos, "Degrees-of-freedom regions for the K-user MISO time-correlated broadcast channel," 2014. [Online]. Available: http://arxiv.org/abs/1412.1023

[25] T. Cover and A. Thomas, *Elements of information theory*. Wiley-Interscience, Jul. 2006.

[26] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.