Modeling and Analysis of Differential CQI Feedback in 4G/5G OFDM Cellular Systems

Vineeth Kumar, *Student Member, IEEE*, and Neelesh B. Mehta, *Fellow, IEEE*

***Abstract*— Reduced feedback schemes are crucial in achieving the high data rates expected of orthogonal frequency-division multiplexing-based 4G and 5G cellular systems. They ensure that the feedback overhead required for acquiring the channel-state information to enable downlink scheduling and adaptive modulation and coding at the base station (BS) does not overwhelm the uplink. We present a novel modeling and analysis of the single-user and multi-user throughputs of the differential feedback scheme that is used in both 4G and 5G standards. In this feedback scheme, a user feeds back a 4-bit wideband channel quality indicator (CQI), which indicates the rate that the user can decode if the BS was to transmit to it over the entire system bandwidth, and a 2-bit differential CQI for each subband relative to it. Our analysis incorporates co-channel interference, different single-stream multi-antenna modes, and different schedulers, which cover a wide range of the tradeoff between the throughput and user fairness. It brings out several insights such as the increase in the throughput as the correlation between subbands increases and how differential feedback reduces the overhead significantly while only marginally reducing the throughput.**

***Index Terms*— Adaptive modulation and coding (AMC), differential feedback, exponential effective SNR mapping (EESM), orthogonal frequency division multiplexing (OFDM), scheduling.**

I. Introduction

CONTEMPORARY cellular communication standards such as 4G Long Term Evolution (LTE) use orthogonal frequency division multiplexing (OFDM) in the downlink [2]. OFDM has also recently been adopted in the 5G new radio standard, but with a more flexible numerology compared to LTE [3]. In it, the system bandwidth is divided into several narrow bandwidth orthogonal subcarriers, which are aggregated to form subchannels (SCs).

Frequency-domain scheduling, adaptive modulation and coding (AMC), and multiple-input-multiple-output (MIMO) are among the key techniques employed in such OFDM systems to enhance the spectral efficiency [2]. In scheduling, the base station (BS) determines which user to allocate to each SC. In AMC, the BS determines the modulation and coding scheme (MCS) to transmit for each SC. To enable these, the BS ideally needs each user to feed back the channel state information (CSI) for every SC. Such feedback is required not only in frequency-division duplexing systems, but also in time-division duplexing systems, due to imperfect calibration of the transmit and receive radio frequency chains and due to asymmetry in the uplink and downlink interferences.

Reduced feedback schemes are key to providing the BS with the requisite CSI and ensuring that the feedback overhead does not overwhelm the uplink. Several such schemes have been proposed in the literature and adopted in standards [3], [4]. We summarize them below. We do not delve into feedback techniques for MIMO as their design principles are different.

A. Reduced Feedback Schemes Used in OFDM Systems

Threshold-based, best-m, clustering-based, and differential feedback are among the popular reduced feedback schemes studied in OFDM systems [4]–[7]. In threshold-based feedback, each user feeds back a quantized value of the signal-to-noise ratio (SNR) of each SC [4], [5]. In best-m feedback, each user reports only the indices and the corresponding rates for m SCs that can support the highest rates [6], [7]. In clustering-based feedback, adjacent SCs are grouped into clusters and feedback is sent only for each cluster [8]. Differential schemes feed back, using fewer bits, only the difference between a reference value and the quantity to be fed back.

Practical 4G LTE and 5G systems employ a combination of the above schemes [2], [3]. In them, feedback is conveyed by the channel quality indicator (CQI), which indicates the MCS that can be decoded over a pre-specified transmission bandwidth. Three feedback schemes are specified in LTE. In *wideband feedback*, which is an example of clustering-based feedback, a user reports just a single wideband CQI. It is the highest-rate MCS that can be decoded if the entire system bandwidth were allocated to that user. To support frequency-domain scheduling, the following two schemes are instead used; they feed back extra information to supplement the wideband CQI. In *UE-selected subband feedback*, the indices and an average CQI for the best-m subbands, where each subband comprises several adjacent SCs, are sent. This is a variant of best-m feedback. In *higher layer-configured subband (HLCS) feedback*, for each subband, a 2-bit differential CQI that encodes the difference between the wideband CQI and the MCS that the user can decode on the subband is sent.

Manuscript received August 16, 2018; revised January 5, 2019; accepted February 21, 2019. Date of publication March 12, 2019; date of current version April 9, 2019. This work was supported in part by the DST -Swarnajayanti Fellowship Award under Grant DST/SJF/ETA-01/2014-15 and in part by the Qualcomm Innovation Fellowship (QInF) India 2018. This paper was presented in part at the IEEE International Conference on Communications, May 2018 [1]. The associate editor coordinating the review of this paper and approving it for publication was W. Chen. *(Corresponding author: Vineeth Kumar.)*

The authors are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: vineethkumar01@gmail.com; nbmehta@iisc.ac.in).

Digital Object Identifier 10.1109/TWC.2019.2903047

In 5G, wideband feedback and HLCS feedback have been specified thus far [3].

The throughput of wideband feedback is analyzed in [9] and [10]. The throughput of UE-selected subband feedback and a subband-level feedback scheme without differential feedback is analyzed in [11]. A medium access control (MAC)-level throughput analysis is presented in [12], but only for wideband feedback. The above papers also differ in their channel and cell layout models. While [9] and [10] consider the multi-cell scenario and model large-scale fading, [11] and [12] consider the single-cell scenario and do not model large-scale fading.

B. Focus and Contributions

In this paper, we present a novel and systematic modeling and analysis of the differential feedback-based HLCS feedback scheme. Such an analysis is not available in the literature, to the best of our knowledge. As we shall see, it requires the development of several new approximation methods. It gives valuable mathematical insights about the performance of a feedback scheme that is an integral component of both 4G and 5G standards. It enables a system designer to independently evaluate or verify the performance of the scheme over a wide range of system parameter settings without resorting to computationally intensive simulations.

We first derive an expression for the throughput of the HLCS feedback scheme for a single-user scenario. This incorporates the effect of both large-scale fading and small-scale fading. We then generalize it to a multi-cell, multi-user scenario for three different schedulers that provide vastly different trade-offs between cell throughput and user-fairness. We do so for the case when the subband fading gains conditioned on the large-scale fading are independent and identically distributed (i.i.d.), which is justified when the coherence bandwidth of the channel is comparable to the subband bandwidth. Our results apply to different single-stream multi-antenna diversity modes such as single-input-multiple-output (SIMO), multiple-input-single-output (MISO), and single-stream MIMO. We also extend the analysis to various multi-stream single user (SU)-MIMO, multi-user (MU)-MIMO, and massive MIMO modes for the single-cell scenario.

A key innovation in our approach is the modeling of wideband CQI using the exponential effective SNR mapping (EESM) [13], [14] and an accurate analysis of the joint statistics of EESM and differential CQI. EESM can be interpreted as the equivalent SNR that a codeword sees if it were transmitted over a flat-fading channel.[1]

We gain more insights by analyzing the following two extreme scenarios: (i) Users feed back complete CSI to the BS, for which we show that the throughput expression simplifies considerably and holds for any correlation between the subband fading gains, and (ii) Subband fading gains of the user are fully correlated, for which we show that differential feedback achieves the same throughput as full CSI feedback, but with significantly less overhead. This is a special property of differential feedback and is not true, in general, for other schemes.

Our analysis differs significantly from prior works on differential CQI feedback, which have been simulation-based and, thus, do not generalize easily [17], [18]. It differs from [19], [20], and the references therein, which study differential feedback in MIMO-OFDM systems but do not model CQI feedback and discrete rate adaptation, which are an integral part of the 4G and 5G standards. It is also very different from [11], which does not model differential feedback, considers only a single-cell scenario, and models the statistics of EESM and the SNR of single-stream MIMO differently.

C. Organization and Notation

This paper is organized as follows. Section II presents a brief overview of the 4G LTE and 5G standards, and the system model that they motivate. Section III analyzes the average throughput of the single-user scenario. Section IV extends the analysis to the multi-cell, multi-user scenario with frequency-domain scheduling. Numerical results are presented in Section V. Our conclusions follow in Section VI.

Notation: We denote the probability of an event A by $\Pr(A)$ and the joint probability of events A and B by $\Pr(A, B)$. The conditional probability of A given B is denoted by $\Pr(A \mid B)$. The probability density function (PDF) and cumulative distribution function (CDF) of a random variable (RV) X are denoted by $f_X(\cdot)$ and $F_X(\cdot)$, respectively. We denote the expectation with respect to an RV X by $\mathbb{E}_X[\cdot]$, and the absolute value by $|\cdot|$. We denote the cardinality of a set $\mathcal{V}$ by $\#(\mathcal{V})$ and its complement by $\mathcal{V}'$. We denote the complex conjugate of a complex number a by a^*. We denote vectors and matrices using boldface characters. For a matrix $\mathbf{H}$, we denote its $(j,j)^{\text{th}}$ element by $[\mathbf{H}]_{jj}$ and its Hermitian transpose by $\mathbf{H}^\dagger$.

II. Preliminaries and System Model

A. Overview of LTE Frame Structure and CQI Feedback

We first summarize the essential features of the LTE downlink, which motivates the system model that we describe next and sets up our terminology. We also comment on its commonalities with 5G. A downlink frame in LTE has a duration of 10 ms and comprises 10 subframes. Each subframe is of duration 1 ms and consists of two 0.5 ms slots. Each slot consists of seven OFDM symbols in the normal cyclic prefix (CP) mode. In the frequency domain, the system bandwidth is divided into several subcarriers, each of which has a bandwidth of 15 kHz. A physical resource block (PRB) consists of a group of 12 adjacent subcarriers and has a duration of one slot. 5G, in addition, supports a flexible numerology in which the subcarrier spacing can vary from 15 kHz to 240 kHz. A slot comprises 14 OFDM symbols. The frame duration is the same as in LTE, but the number of slots in a frame varies [21].

CQI Feedback: The CQI is a 4-bit value that indicates an estimate of the MCS that the user can reliably decode on

[1]Other approaches based on using the median of the subband CQI values [15] and mutual information-based effective SNR mapping (MIESM) [16] have been studied in the literature. However, the median-based approach does not ensure that a codeword generated using the MCS given by it can be decoded and the MIESM-based approach is analytically intractable.

the downlink. There are $2^4 = 16$ CQIs, whose MCSs and rates are tabulated in [2, Table 10.1]. The CQI is estimated by the user from the signal quality measurements of the reference signals it receives from the BS. The frequency of CQI feedback is controlled by the BS. The finest possible frequency resolution for CQI reporting is a subband, which consists of q adjacent PRBs, where $2 \leq q \leq 8$. This implies that a codeword can be transmitted over a frequency-selective channel even though only one MCS is reported for it.

Scheduling and AMC: Based on the coarser subband-level CQI received from all the users, the implementation-specific scheduler at the BS determines which user to allocate to a PRB. It makes these decisions at most once in every millisecond, and conveys it to the users on the downlink control channel.

In AMC, the BS determines the transmit rate to the scheduled user on each PRB by mapping the received CQI to an MCS, which consists of a modulation scheme and a code rate. The MCS chosen and the user scheduled, thus, depend on the channel conditions. LTE supports QPSK, 16QAM, and 64QAM and the code rate varies from 0.08 to 0.93. 5G also supports 256QAM. To reduce the control signaling overhead, the same MCS is used for transmission on all the PRBs allocated to a user in a subframe, which consists of two slots.

In this paper, we focus on single-stream transmission, which encompasses single-input-single-output (SISO), SIMO, MISO, and single-stream MIMO. This enables us to characterize the impact of CQI feedback on the throughput. The design and throughput analysis of codebook-based precoding for multi-stream transmission is different, and is an open problem.

B. System Model for a Single-User Scenario With Small-Scale and Large-Scale Fading

We first study a point-to-point transmission in the OFDM downlink between a BS that has N_t transmit antennas and a single user that has N_r receive antennas. This brings out the key ideas using notation that is considerably simpler than that for the multi-cell scenario, which we consider next in Section IV. Our model below captures several practically important – but not all – aspects of the system described above, and is tractable and insightful. The system bandwidth is divided into N_{PRB} orthogonal PRBs. Groups of q adjacent PRBs form subbands. Hence, the total number of subbands is $N_{\text{SB}} = \lceil N_{\text{PRB}}/q \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling function.

1) Channel Model: In the channel between the user and the BS, signals experience frequency-selective small-scale fading, which follows the Rayleigh distribution, and also undergo path-loss and shadowing, which have been referred to as large-scale fading in [22]. Shadowing follows the lognormal distribution. They are modeled as follows.

Small-Scale Fading: The complex downlink baseband fading gain $H_n(i,j)$ of subband n of the user between the i^{th} receive antenna of the user and the j^{th} transmit antenna of the BS is a circularly-symmetric complex Gaussian RV with zero mean and unit variance. For tractability, we assume that it remains constant over a subband. This is justified when the bandwidth of a subband, which ranges from 360 kHz to 1.4 MHz in LTE, is comparable to the coherence bandwidth of the channel. For instance, for the typical urban (TU) channel, the coherence bandwidth is 530 kHz [23, Ch. 3]. The fading gains between the different transmit antennas of the BS and the different receive antennas of the user are i.i.d. [23, Ch. 2]. The subband fading power gains $|H_n(i,j)|^2$, for $1 \leq n \leq N_{\text{SB}}$, are statistically identical, which follows from the uncorrelated scatterers assumption [23, Ch. 2].

Large-Scale Fading: ω denotes the lognormal RV that models the large-scale fading; it is the same for all subbands [10], [22]. In dB scale, it is a Gaussian RV with mean $\mu_\omega(x) = -\delta_{\text{dB}}(x_0) - 10\eta \log(x/x_0)$ and standard deviation σ_{shad}, where x is the distance between the user and the BS, $\delta_{\text{dB}}(x_0)$ is the path-loss at a reference distance x_0 from the BS, and η is the path-loss exponent [23, Ch. 1].

2) SNRs for Different Single-Stream Multi-Antenna Modes: The instantaneous SNR γ_n of subband n for SIMO ($N_r \geq 1$ and $N_t = 1$) with maximal ratio combining (MRC), MISO ($N_r = 1$ and $N_t \geq 1$) with maximal ratio transmission (MRT), and single-stream MIMO ($N_r \geq 2$ and $N_t \geq 2$) with singular value decomposition-based MIMO beamforming [24], is [10]

$$\gamma_n = \frac{P_T \omega \phi_n}{P_N} = \Omega \phi_n, \tag{1}$$

where P_T is the transmit power of the BS per subband, P_N is the additive white Gaussian noise (AWGN) power per subband, and ϕ_n depends on the multi-antenna mode. For SIMO $\phi_n = \sum_{j=1}^{N_r} |H_n(j,1)|^2$ is a gamma RV with PDF $f_{\phi_n}(v) = v^{d-1} e^{-\frac{v}{\lambda}} / (\Gamma(d)\lambda^d)$, for $v \geq 0$, where $\Gamma(\cdot)$ is the gamma function [25, (8.310)]. The PDF has a shape parameter $d = N_r$ and a scale parameter $\lambda = 1$ [24]. Similarly, for MISO, $\phi_n = \sum_{i=1}^{N_t} |H_n(1,i)|^2$ is a gamma RV with $d = N_t$ and $\lambda = 1$. For single-stream MIMO, ϕ_n can be accurately approximated as a gamma RV with $d = N_t N_r$ and $\lambda = ((N_t + N_r)/(N_t N_r + 1))^{\frac{2}{3}}$ [24].[2] In all cases, $\Omega = (P_T/P_N)\omega$ is a lognormal RV with dB scale parameters $\mu_\Omega = \mu_\omega(x) + \xi \log(P_T/P_N)$ and $\sigma_\Omega = \sigma_{\text{shad}}$, where $\xi = 10/\log(10)$.

3) MCS Set for AMC: The BS has available to it an *MCS set* $\mathcal{M} = \{1, 2, \ldots, M\}$ of MCSs to choose from. MCS m has a rate r_m and a decoding threshold T_m. The user can decode MCS m on subband n only if $\gamma_n \geq T_m$. Else, an outage occurs [9], [26]. Let $r_1 < r_2 < \cdots < r_M$ and $T_1 < T_2 < \cdots < T_M < T_{M+1}$, where $r_1 = 0$, $T_1 = 0$, and $T_{M+1} = \infty$.

4) Feedback: We now describe the HLCS feedback scheme and also contrast it with full CSI feedback, in which a user feeds back as CQI the highest-rate MCS that it can decode on each subband.

Full CSI Feedback: This is a threshold-based feedback scheme, in which the CQI fed back for subband n is m if $T_m \leq \gamma_n < T_{m+1}$. Given this CQI, the BS transmits with MCS $\tilde{S}_n = m$ on subband n [26]. This scheme requires $B = \log_2(M)$ bits per subband for feedback, which results

[2]For supporting MISO and single-stream MIMO, a precoding matrix indicator (PMI) and a rank indicator (RI) are fed back additionally in LTE and 5G [2, Ch. 11], [3]. To ensure tractability and to focus on the impact of differential CQI on the system throughput, we assume that this feedback resolution is sufficiently fine so that the actual SNR is close to the ideal SNR derived above. Our results, thus, provide an upper limit for MIMO systems with coarse PMI feedback.

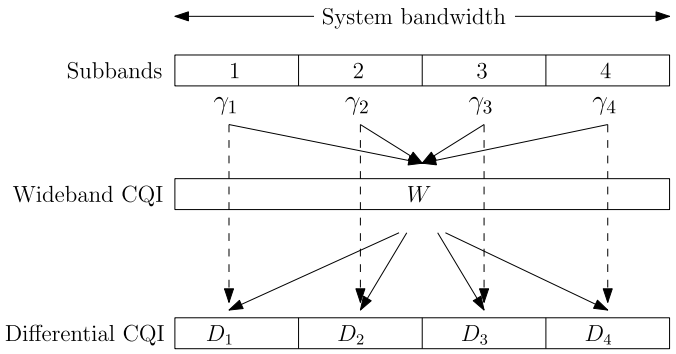

Fig. 1. Illustration of generation of wideband CQI and differential CQI offset values for the user for $N_{\text{SB}} = 4$ subbands.

in an overhead of $N_{\text{SB}}B$ bits per user. In LTE and 5G, M is set as 16 [2, Table 10.1], [3]. While one CQI table is specified in LTE, three CQI tables are specified in 5G; the BS selects one among them [3].

HLCS Feedback: In it, the user determines a wideband CQI W assuming that the entire system bandwidth is allocated to it. For each subband n, it also determines the subband CQI S_n, assuming that the subband is allocated to it. The user then generates a subband differential CQI offset value D_n using W and S_n. This is illustrated in Fig. 1. The details are as follows:

Step 1) Wideband CQI Generation: The MCS W corresponding to this wideband CQI must be such that a codeword transmitted using it over the entire system bandwidth can be decoded by the user. Since different subbands see different SNRs, the MCS to feed back is not obvious at first sight.

To do this systematically, we use EESM, whose accuracy has been extensively validated in [9], [13], and [14]. It maps a vector of N_{SB} subband SNRs $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{N_{\text{SB}}})$ seen by different parts of the codeword into a single effective SNR $\zeta^{(m)}$ for MCS m as follows [13]:

$$\zeta^{(m)} = -\beta_m \log\left(\frac{1}{N_{\text{SB}}}\sum_{n=1}^{N_{\text{SB}}} \exp\left(-\frac{\gamma_n}{\beta_m}\right)\right), \text{ for } m \geq 2, \quad (2)$$

where β_m is an MCS-dependent scaling constant [13]. Here, $\zeta^{(m)}$ is interpreted to be the equivalent SNR for MCS m in a frequency-flat AWGN channel that results in the same probability of error as that when transmitting over the frequency-selective channel. Consequently, the wideband MCS W is chosen as the highest-rate MCS for which $\zeta^{(m)} \geq T_m$. Thus,

$$W = m, \quad \text{if and only if } \zeta^{(m)} \geq T_m, \zeta^{(m+1)} < T_{m+1}, \ldots, \zeta^{(M)} < T_M, \text{ for } 2 \leq m \leq M. \quad (3)$$

Else, $W = 1$, which corresponds to the rate $r_1 = 0$ (no transmission).

Step 2) Differential CQI Generation: Let $S_n \in \mathcal{M}$ denote the subband MCS, which is the index of the highest-rate MCS that the user can decode on subband n. Clearly, $S_n = m$ if $T_m \leq \gamma_n < T_{m+1}$. In both LTE and 5G, it is encoded differentially with respect to W to generate the 2-bit subband differential CQI offset D_n as follows [2, Ch. 10], [3]:

$$D_n = \begin{cases} -1, & S_n - W \leq -1, \\ 0, & S_n - W = 0, \\ 1, & S_n - W = 1, \\ 2, & S_n - W \geq 2. \end{cases} \quad (4)$$

The vector $\mathbf{D} = (D_1, \ldots, D_{N_{\text{SB}}})$ of differential CQI offsets, which requires $2N_{\text{SB}}$ bits, and W, which requires 4 bits, are then fed back to the BS. Given D_n and W, the BS transmits with MCS $\tilde{S}_n$ on subband n, where

$$\tilde{S}_n = W + D_n. \quad (5)$$

We shall generalize this system model and notation to the multi-cell, multi-user scenario in Section IV. It can also be easily generalized to allow for less or more than 2 bits of differential feedback. We summarize the important variables and system parameters described above in Table I.

TABLE I
IMPORTANT VARIABLES AND SYSTEM PARAMETERS

Parameters	Definitions
N_{SB}	Number of subbands
N_t, N_r	Number of transmit and receive antennas
P_T	Transmit power of the BS per subband
P_N	AWGN power per subband
M	Number of MCSs
r_m, T_m	Rate and decoding threshold for MCS m
γ_n, $\zeta^{(m)}$	SNR of subband n, effective SNR for MCS m
Ω, ϕ_n	Large-scale fading and small-scale fading on subband n
d, λ	Shape and scale parameters of gamma RV
μ_Ω, σ_Ω	Mean and standard deviation of Ω
S_n, $\tilde{S}_n$	Subband MCS and transmit MCS for subband n
W, D_n	Wideband MCS and differential CQI offset for subband n
x, x_0	Distance between BS and user, reference distance
$\delta_{\text{dB}}(x_0)$, η	Path-loss at x_0, path-loss exponent

III. SINGLE-USER SCENARIO: AVERAGE THROUGHPUT ANALYSIS

For the single-user scenario, using the law of total expectation, the fading-averaged throughput $\bar{R}_n$ on subband n of the differential feedback scheme is given by

$$\bar{R}_n = \sum_{m=2}^{M} r_m \text{Pr}\left(\tilde{S}_n = m, \gamma_n \geq T_m\right), \quad (6)$$

since r_m bits/symbol is achieved only if the BS transmits with MCS m ($\tilde{S}_n = m$) and the codeword is decoded by the user ($\gamma_n \geq T_m$). The summation in (6) begins at $m = 2$ and not $m = 1$ since $r_1 = 0$. From (4) and (5), $\tilde{S}_n$ is m when one among the following four mutually exclusive events occurs: (i) $W = m + 1, S_n \leq m$; (ii) $W = m, S_n = m$; (iii) $W = m - 1, S_n = m$; or (iv) $W = m - 2, S_n \geq m$. Therefore,

$$\begin{aligned} \text{Pr}(\tilde{S}_n = m, \gamma_n \geq T_m) = {} & \text{Pr}(W = m+1, S_n \leq m, \gamma_n \geq T_m) \\ & + \text{Pr}(W = m, S_n = m, \gamma_n \geq T_m) \\ & + \text{Pr}(W = m-1, S_n = m, \gamma_n \geq T_m) \\ & + \text{Pr}(W = m-2, S_n \geq m, \gamma_n \geq T_m). \end{aligned} \quad (7)$$

For $1 \leq m \leq 2$, W can become zero or negative in one or more events in (7). For $m = M$, W can become $M+1$. We set the probabilities of such events to 0 in (7). The computation of each term in (7) involves the following three key steps, which are explained in detail below: (i) Reducing the dimension of the multi-variate PDFs involved using tight approximations; (ii) Developing a general integral form for computing the probabilities involving these reduced-dimension PDFs; and (iii) Using recent results about the statistics of EESM to evaluate this integral.

A. Reducing the Dimension of the Multi-Variate PDFs

We focus on the first term in (7) below for $3 \leq m \leq M-1$. From (3), it follows that

$$\begin{aligned}&\Pr\left(W = m+1, S_n \leq m, \gamma_n \geq T_m\right)\\ &\quad = \Pr(\zeta^{(m+1)} \geq T_{m+1}, \zeta^{(m+2)} < T_{m+2}, \ldots, \zeta^{(M)}\\ &\qquad < T_M, T_m \leq \gamma_n < T_{m+1}).\end{aligned} \tag{8}$$

The RVs $\zeta^{(1)}, \zeta^{(2)}, \ldots, \zeta^{(M)}$ are correlated since they are all functions of the same set of RVs $\gamma_1, \ldots, \gamma_{N_{\text{SB}}}$. Thus, computing (8) requires an $(M-m)$-dimensional joint PDF of the RVs $\zeta^{(m+1)}, \ldots, \zeta^{(M)}, \gamma_n$. However, no expression is known for it. An exact closed-form expression is not available even for the marginal PDF of $\zeta^{(m)}$ due to its involved form.

To solve this problem, we use the following intuition. If MCS $m+2$ cannot be decoded by the user, then it it is highly unlikely that a higher rate MCS can be decoded.[3] Therefore,

$$\begin{aligned}&\Pr\left(\zeta^{(m+1)} \geq T_{m+1}, \zeta^{(m+2)} < T_{m+2},\right.\\ &\qquad \left.\ldots, \zeta^{(M)} < T_M, T_m \leq \gamma_n < T_{m+1}\right)\\ &\quad \approx \Pr\left(\zeta^{(m+1)} \geq T_{m+1}, \zeta^{(m+2)} < T_{m+2},\right.\\ &\qquad \left. T_m \leq \gamma_n < T_{m+1}\right).\end{aligned} \tag{9}$$

Using the law of total probability, we then have

$$\begin{aligned}&\Pr\left(\zeta^{(m+1)} \geq T_{m+1}, \zeta^{(m+2)} < T_{m+2}, T_m \leq \gamma_n < T_{m+1}\right)\\ &\quad = \Pr\left(\zeta^{(m+1)} \geq T_{m+1}, T_m \leq \gamma_n < T_{m+1}\right)\\ &\qquad -\Pr\left(\zeta^{(m+1)} \geq T_{m+1}, \zeta^{(m+2)} \geq T_{m+2},\right.\\ &\qquad \left. T_m \leq \gamma_n < T_{m+1}\right).\end{aligned} \tag{10}$$

The above reasoning also implies that if MCS $m+2$ can be decoded, then it is highly likely that the lower rate MCS $m+1$ can also be decoded. Therefore, (10) simplifies to

$$\begin{aligned}&\Pr\left(\zeta^{(m+1)} \geq T_{m+1}, \zeta^{(m+2)} < T_{m+2}, T_m \leq \gamma_n < T_{m+1}\right)\\ &\quad \approx \Pr\left(\zeta^{(m+1)} \geq T_{m+1}, T_m \leq \gamma_n < T_{m+1}\right)\\ &\qquad -\Pr\left(\zeta^{(m+2)} \geq T_{m+2}, T_m \leq \gamma_n < T_{m+1}\right).\end{aligned} \tag{11}$$

[3]While this is trivially true for a flat-fading channel, this is not the case for a frequency-selective channel due to the non-linear form of EESM [10].

Similarly, the second, third, and fourth probability terms in (7) simplify as follows:

$$\begin{aligned}&\Pr(W = m, S_n = m, \gamma_n \geq T_m)\\ &\quad \approx \Pr\left(\zeta^{(m)} \geq T_m, T_m \leq \gamma_n < T_{m+1}\right)\\ &\qquad -\Pr\left(\zeta^{(m+1)} \geq T_{m+1}, T_m \leq \gamma_n < T_{m+1}\right),\end{aligned} \tag{12}$$

$$\begin{aligned}&\Pr(W = m-1, S_n = m, \gamma_n \geq T_m)\\ &\quad \approx \Pr\left(\zeta^{(m-1)} \geq T_{m-1}, T_m \leq \gamma_n < T_{m+1}\right)\\ &\qquad -\Pr\left(\zeta^{(m)} \geq T_m, T_m \leq \gamma_n < T_{m+1}\right),\end{aligned} \tag{13}$$

$$\begin{aligned}&\Pr(W = m-2, S_n \geq m, \gamma_n \geq T_m)\\ &\quad \approx \Pr\left(\zeta^{(m-2)} \geq T_{m-2}, \gamma_n \geq T_m\right)\\ &\qquad -\Pr\left(\zeta^{(m-1)} \geq T_{m-1}, \gamma_n \geq T_m\right).\end{aligned} \tag{14}$$

Combining (11)–(14) yields the following much simpler expression for (7):

$$\begin{aligned}\Pr\left(\tilde{S}_n = m, \gamma_n \geq T_m\right) \approx\ &\Pr\left(\zeta^{(m-2)} \geq T_{m-2}, \gamma_n \geq T_m\right)\\ &-\Pr\left(\zeta^{(m-1)} \geq T_{m-1}, \gamma_n \geq T_{m+1}\right)\\ &-\Pr\left(\zeta^{(m+2)} \geq T_{m+2}, \gamma_n \geq T_m\right)\\ &+\Pr\left(\zeta^{(m+2)} \geq T_{m+2}, \gamma_n \geq T_{m+1}\right).\end{aligned} \tag{15}$$

The corresponding expressions for $\Pr\left(\tilde{S}_n = m, \gamma_n \geq T_m\right)$ for the special cases of $m = 2$ and $m = M$ are

$$\begin{aligned}\Pr\left(\tilde{S}_n = 2, \gamma_n \geq T_2\right) \approx\ &\Pr(T_2 \leq \gamma_n \leq T_3)\\ &-\Pr\left(\zeta^{(4)} \geq T_4, \gamma_n \geq T_2\right)\\ &+\Pr\left(\zeta^{(4)} \geq T_4, \gamma_n \geq T_3\right),\end{aligned}$$

and

$$\Pr\left(\tilde{S}_n = M, \gamma_n \geq T_M\right) \approx \Pr\left(\zeta^{(M-2)} \geq T_{M-2}, \gamma_n \geq T_M\right).$$

B. General Integral Form for the Probabilities

Each term in (15) is of the general form $\Pr(\zeta^{(l)} \geq T_l, \gamma_n \geq t)$, where $l \in \{m-2, m-1, m+2\}$ and $t \in \{T_m, T_{m+1}\}$. It can be written in terms of the PDF $f_{\gamma_n}(\cdot)$ of γ_n as

$$\begin{aligned}&\Pr\left(\zeta^{(l)} \geq T_l, \gamma_n \geq t\right)\\ &\qquad = \int_t^{\infty} \Pr\left(\zeta^{(l)} \geq T_l | \gamma_n = v\right) f_{\gamma_n}(v)\, dv.\end{aligned} \tag{16}$$

As defined in (1), $\gamma_n = \Omega\phi_n$, where Ω is a lognormal RV and ϕ_n is a gamma RV whose parameters depend on the multi-antenna mode used. Its PDF is

$$\begin{aligned}f_{\gamma_n}(v) &= \mathbb{E}_{\Omega}[f_{\gamma_n}(v|\Omega = u)],\\ &= \int_0^{\infty} f_{\Omega}(u)\, \frac{1}{u} f_{\phi_n}\left(\frac{v}{u}\right)\, du.\end{aligned} \tag{17}$$

Substituting (17) and the lognormal PDF of Ω in (16), we get

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) &= \frac{\xi}{\sqrt{2\pi}\sigma_\Omega}\int_0^\infty \frac{1}{u^2} e^{-\left(\frac{\xi\log(u)-\mu_\Omega}{\sqrt{2}\sigma_\Omega}\right)^2} \\ &\times \int_t^\infty f_{\phi_n}\left(\frac{v}{u}\right)\Pr\Big(\zeta^{(l)} \geq T_l | \gamma_n = v, \Omega = u\Big)\, dv\, du.\end{aligned} \tag{18}$$

Using the variable substitution $y = (\xi\log(u) - \mu_\Omega)/\left(\sqrt{2}\sigma_\Omega\right)$ and expanding $\zeta^{(l)}$ as per (2) yields

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^\infty \frac{e^{-y^2}}{\Lambda_y}\int_t^\infty f_{\phi_n}\left(\frac{v}{\Lambda_y}\right) \\ \times\Pr\left(-\beta_l\log\left(\sum_{j=1}^{N_{\text{SB}}}\frac{e^{-\frac{\phi_j\Lambda_y}{\beta_l}}}{N_{\text{SB}}}\right)\right. \\ \left.\geq T_l \,\Big|\, \gamma_n = v, \Omega = u\Big)\, dv\, dy,\end{aligned} \tag{19}$$

where

$$\Lambda_y = e^{\left(\sqrt{2}\sigma_\Omega y + \mu_\Omega\right)/\xi}. \tag{20}$$

Notice that (19) is of the form $\frac{1}{\sqrt{\pi}}\int_{-\infty}^\infty e^{-y^2}h(y)\,dy$, where

$$\begin{aligned}h(y) = \frac{1}{\Lambda_y}\int_t^\infty f_{\phi_n}\left(\frac{v}{\Lambda_y}\right) \\ \times\Pr\left(-\beta_l\log\left(\sum_{j=1}^{N_{\text{SB}}}\frac{e^{-\frac{\phi_j\Lambda_y}{\beta_l}}}{N_{\text{SB}}}\right) \geq T_l \,\Big|\, \gamma_n = v, \Omega = u\right) dv.\end{aligned} \tag{21}$$

Applying Gauss-Hermite quadrature, (19) simplifies to

$$\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) \approx \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) h(y_i), \tag{22}$$

where $w_{\text{GH}}(i)$ and y_i, for $1 \leq i \leq N_{\text{GH}}$, are the weights and abscissas, respectively [27, (25.4.46)]. Substituting the expression for $h(y_i)$ from (21) in (22) yields

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) \approx \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N_{\text{GH}}}\frac{w_{\text{GH}}(i)}{\Lambda_{y_i}}\int_t^\infty f_{\phi_n}\left(\frac{v}{\Lambda_{y_i}}\right) \\ \times\Pr\left(-\beta_l\log\left(\sum_{j=1}^{N_{\text{SB}}}\frac{e^{-\frac{\phi_j\Lambda_{y_i}}{\beta_l}}}{N_{\text{SB}}}\right) \geq T_l \,\Big|\, \gamma_n = v, \Omega = u\right) dv.\end{aligned} \tag{23}$$

We now compute the expression for $\Pr\Big(-\beta_l\log\Big(\frac{1}{N_{\text{SB}}}\sum_{j=1}^{N_{\text{SB}}} e^{-\frac{\phi_j\Lambda_{y_i}}{\beta_l}}\Big) \geq T_l \,\Big|\, \gamma_n = v, \Omega = u\Big)$. Upon removing the log, rearranging terms, and using the fact that ϕ_j is independent of ϕ_n, for $j \neq n$, we get

$$\begin{aligned}\Pr\left(-\beta_l\log\left(\sum_{j=1}^{N_{\text{SB}}}\frac{e^{-\frac{\phi_j\Lambda_{y_i}}{\beta_l}}}{N_{\text{SB}}}\right) \geq T_l \,\Big|\, \gamma_n = v, \Omega = u\right) \\ = F_{X_l(i)}\left(x\left(v;l\right)\right),\end{aligned} \tag{24}$$

where $x\left(v;l\right) = (N_{\text{SB}}e^{\frac{-T_l}{\beta_l}} - e^{\frac{-v}{\beta_l}})/(N_{\text{SB}}-1)$ and

$$X_l(i) = \frac{1}{N_{\text{SB}}-1}\sum_{j=1, j\neq n}^{N_{\text{SB}}} e^{-\frac{\Lambda_{y_j}\phi_j}{\beta_l}}. \tag{25}$$

C. Exploiting EESM Statistics

From (2) and (25), we see that the RV $-\beta_l\log\left(X_l(i)\right)$ is related to the effective SNR of MCS l over $(N_{\text{SB}}-1)$ subbands. Recent results on EESM statistics in [9] show that $X_l(i)$ can also be well approximated as a Beta RV.[4] This is different from [11], which models $\zeta^{(l)}$ as a lognormal RV. Hence,

$$F_{X_l(i)}(x\left(v;l\right)) = \mathcal{B}(x\left(v;l\right); a_l(i), b_l(i)), \text{ for } x\left(v;l\right) \in [0,1], \tag{26}$$

where $\mathcal{B}(x\left(v;l\right); a_l(i), b_l(i)) = \frac{\int_0^{x(v;l)} z^{a_l(i)-1}(1-z)^{b_l(i)-1}\,dz}{\int_0^1 z^{a_l(i)-1}(1-z)^{b_l(i)-1}\,dz}$ is the regularized incomplete Beta function [25, (8.392)]. In terms of the mean $\mu_l(i)$ and variance $v_l(i)$ of $X_l(i)$, the parameters $a_l(i)$ and $b_l(i)$ obtained from [27, (26.1.33)] are

$$a_l(i) = \frac{\mu_l(i)\left[\mu_l(i) - (\mu_l(i))^2 - v_l(i)\right]}{v_l(i)}, \tag{27}$$

$$b_l(i) = \frac{(1-\mu_l(i))\left[\mu_l(i) - (\mu_l(i))^2 - v_l(i)\right]}{v_l(i)}. \tag{28}$$

Furthermore, from (25), it can be easily shown that $\mu_l(i)$ and $v_l(i)$ can be written as

$$\mu_l(i) = (1 + \Lambda_{y_i}\lambda\beta_l^{-1})^{-d}, \tag{29}$$

$$v_l(i) = \frac{1}{N_{\text{SB}}-1}\left((1 + 2\Lambda_{y_i}\lambda\beta_l^{-1})^{-d} - (\mu_l(i))^2\right). \tag{30}$$

Substituting (29) and (30) in (27) and (28) yields the expressions for $a_l(i)$ and $b_l(i)$. We shall test its accuracy in Section V.

Substituting the expression for the Beta CDF from (26) in (23) yields

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) \approx \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N_{\text{GH}}}\frac{w_{\text{GH}}(i)}{\Lambda_{y_i}}\int_t^\infty f_{\phi_n}\left(\frac{v}{\Lambda_{y_i}}\right) \\ \times\mathcal{B}(x(v;l); a_l(i), b_l(i))\, dv.\end{aligned} \tag{31}$$

As shown in Appendix A, this can be evaluated as a finite series as follows:

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) \\ \approx \frac{1}{\Gamma(d)\sqrt{\pi}}\sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) e^{\frac{-t}{\Lambda_{y_i}\lambda}} \\ \times\sum_{i_1=0}^{d-1}\binom{d-1}{i_1}\left(\frac{t}{\Lambda_{y_i}\lambda}\right)^{d-1-i_1}\sum_{i_2=1}^{N_{\text{GL}}} w_{\text{GL}}(i_1,i_2) \\ \times\mathcal{B}\left(x\left(\lambda\theta_{\text{GL}}(i_1,i_2)\Lambda_{y_i} + t; l\right); a_l(i), b_l(i)\right),\end{aligned} \tag{32}$$

[4]The Beta approximation for $X_l(i)$ is based on the fact that $e^{-\frac{\Lambda_{y_j}\phi_j}{\beta_l}}$ has a finite and positive support of $(0,1]$. When N such i.i.d. RVs with finite and positive support are added, the central limit approximation of Papoulis states that the sum is well approximated by a Beta RV even for small N [28].

where $w_{\text{GL}}(.,.)$ and $\theta_{\text{GL}}(.,.)$, for $1 \leq i_2 \leq N_{\text{GL}}$, are the weights and the corresponding abscissas, respectively, of generalized Gauss-Laguerre quadrature [29]. Recall that d and λ depend on the multi-antenna mode used. In (32), we have found that just $N_{\text{GH}} = 10$ and $N_{\text{GL}} = 8$ terms are sufficient for ensuring numerical accuracy for the entire range of parameters of interest.

Evaluating (31) for all values of l and t that arise in (15) and combining them yields $\Pr\left(\tilde{S}_n = m, \gamma_n \geq T_m\right)$. Thereafter, substituting it in (6) yields the final expression for $\bar{R}_n$. This is not shown due to its length and as it provides no additional insight.

D. Two Insightful Extreme Cases

To benchmark the performance of the differential feedback scheme and gain insights, we now analyze two extreme cases: (i) User feeds back full CSI to the BS and (ii) Differential feedback with fully correlated subbands, which occurs when the subband bandwidth is much less than the coherence bandwidth, as is the case, for example, in rural environments.

1) Full CSI Feedback With Arbitrary Subband Correlation:

Result *1*: With full CSI feedback, the fading-averaged throughput $\bar{R}_n$ on subband n is given by

$$\bar{R}_n = \frac{1}{\Gamma(d)\sqrt{\pi}} \sum_{m=2}^{M} r_m \left(\sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) \times \left[U\left(\frac{T_m}{\Lambda_{y_i}\lambda}, d\right) - U\left(\frac{T_{m+1}}{\Lambda_{y_i}\lambda}, d\right) \right] \right), \quad (33)$$

where $U(\cdot,\cdot)$ denotes the upper incomplete gamma function [27, (6.5.3)].

Proof: The proof is given in Appendix B. ∎

Equation (33) brings out the dependence of $\bar{R}_n$ on the MCS set through the parameters r_m and T_m, on the multi-antenna mode used through the parameters d and λ, and on large-scale fading through Λ_{y_i}.

2) Differential Feedback With Fully Correlated Subbands: In this case, for every channel realization, all subband SNRs of the user, which are statistically identical, take the same value with probability 1 [30], [31, Ch. 7]. Therefore, it follows from (2) that

$$\zeta^{(m)} = \gamma_1 = \cdots = \gamma_{N_{\text{SB}}}, \quad \text{for } 1 \leq m \leq M. \quad (34)$$

Substituting (34) in (5) yields $W = S_n$ and $D_n = 0$. Hence, $\tilde{S}_n = S_n$. Therefore, the expression for the fading-averaged throughput is the same as that in (33) for full CSI feedback. Note that (33) applies to any correlation between the subband SNRs. Also, full CSI feedback yields the highest throughput among all feedback schemes. Therefore, differential feedback achieves the highest throughput when the subband SNRs are fully correlated.

IV. Multi-Cell, Multi-User Scenario With Frequency-Domain Scheduling

In the multi-cell, multi-user scenario with co-channel interference, the BS serves K users in a cell. The frequency-domain scheduler at the BS schedules a user in each subband based on the differential and wideband CQIs fed back by the users. This is done once in every scheduling interval, which is of the order of one millisecond. We shall refer to cell 0 as the reference cell. It receives co-channel interference from C neighboring cells.

In order to track the user as well as the BS indices, we expand the notation as follows. Let $\mathbf{H}_{kn}^{(l)}$ denote an $N_r \times N_t$ matrix, whose $(i,j)^{\text{th}}$ element $H_{kn}^{(l)}(i,j)$ denotes the complex downlink baseband fading gain on subband n of user k between receive antenna i of the user and transmit antenna j of BS l, where $0 \leq l \leq C$. Each element is a circularly-symmetric complex Gaussian RV with unit variance. As before, $H_{kn}^{(l)}(i,j)$ for different i, j, k, n, and l are independent RVs since the users are spaced sufficiently far apart [10], [23]. The large-scale fading of the channel between user k and BS l is now denoted by ω_{lk}. In dB scale, ω_{lk} is a Gaussian RV with mean $\mu_{\omega_{lk}}(x_{lk}) = -\delta_{\text{dB}}(x_0) - 10\eta\log(x_{lk}/x_0)$ and standard deviation σ_{shad}, where x_{lk} is the distance between BS l and user k, and the parameters $\delta_{\text{dB}}(x_0)$ and η are defined in Section II-B.1. The presence of co-channel interference makes the modeling of the resultant signal-to-interference-plus-noise ratio (SINR) for single-stream MIMO intractable because of an additional inter-stream interference term in its denominator. Hence, we focus on SIMO and MISO in this section.

A. SINR for SIMO and MISO

For SIMO with MRC and for MISO with MRT, the SINR γ_{kn} on subband n of user k takes the following general form:

$$\gamma_{kn} = \frac{P_T}{P_N}\frac{\omega_{0k}}{\varepsilon_{kn}}\phi_{kn}, \quad (35)$$

where P_T and P_N are defined in Section II-B.2. Here, ϕ_{kn} tracks the small-scale fading and the denominator term $\varepsilon_{kn} = (P_T/P_N)\sum_{l=1}^{C}\omega_{lk}|I_{kn}(l)|^2 + 1$ tracks the interference and noise powers or their scaled versions thereof, as described below.

For SIMO with MRC, $\phi_{kn} = \sum_{j=1}^{N_r}|H_{kn}^{(0)}(j,1)|^2$ is a gamma RV with parameters $d_k = N_r$ and $\lambda_k = 1$, and $I_{kn}(l) = \frac{\sum_{j=1}^{N_r}\left(H_{kn}^{(0)}(j,1)\right)^* H_{kn}^{(l)}(j,1)}{\sqrt{\sum_{j=1}^{N_r}|H_{kn}^{(0)}(j,1)|^2}}$ is a scaled version of the interference caused by BS l to user k on subband n. It is a circularly-symmetric complex Gaussian RV with unit variance [32]. Hence, ε_{kn} is a sum of C lognormal RVs, which are modulated by exponential RVs that correspond to small-scale fading, and unity. It is well approximated by a lognormal RV with parameters $\mu_{\varepsilon_{kn}}$ and $\sigma_{\varepsilon_{kn}}$ [10]. These parameters are obtained by the moment generating function (MGF) matching method [33]. In it, the MGF $\Psi_{\varepsilon_{kn}}(s)$ of ε_{kn}, which is given below, is equated with that of the lognormal distribution, which is given in [33, (9)], at two different values of s. It can be shown that

$$\Psi_{\varepsilon_{kn}}(s) \approx \frac{e^{-s}}{\pi^{\frac{C}{2}}}\prod_{l=1}^{C}\left[\sum_{i=1}^{N_{\text{GH}}}\frac{w_{\text{GH}}(i)}{1+s\frac{P_T}{P_N}e^{\left(\mu_{\omega_{lk}}+\sqrt{2}y_i\sigma_{\text{shad}}\right)/\xi}}\right], \quad (36)$$

where $w_{\text{GH}}(i)$ and y_i are the weights and abscissas, respectively, of Gauss-Hermite quadrature.

For MISO with MRT, $\phi_{kn} = \left(\sum_{i=1}^{N_t} |H_{kn}^{(0)}(1,i)|^2\right)^2$ is the square of a gamma RV with parameters $d_k = N_t$ and $\lambda_k = 1$, and $I_{kn}(l) = \sum_{i=1}^{N_t} \left(H_{k'(l)n}^{(l)}(1,i)\right)^* H_{kn}^{(l)}(1,i)$, where $k'(l)$ is the user scheduled on subband n in cell l. This is because the beamforming weights used by BS l to serve the user $k'(l)$ are proportional to $H_{k'(l)n}^{(l)}(1,i)$, for $1 \leq i \leq N_t$. Here, ε_{kn} consists of a sum of C lognormal RVs, which are modulated by small-scale fading related RVs, and unity. As above, we approximate $(P_T/P_N)\omega_{lk}|I_{kn}(l)|^2$, for each l, by a lognormal RV whose parameters $\mu'_{\omega_{lk}}$ and $\sigma'_{\omega_{lk}}$ are obtained using the Fenton-Wilkinson method [23, Ch. 3]. We then approximate ε_{kn} by a lognormal RV, as above, using the MGF matching method. This yields

$$\Psi_{\varepsilon_{kn}}(s) \approx \frac{e^{-s}}{\pi^{\frac{C}{2}}} \prod_{l=1}^{C} \left[\sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) e^{-\left(s\exp\left(\mu'_{\omega_{lk}} + \sqrt{2}y_i\sigma'_{\omega_{lk}}\right)\right)}\right]. \tag{37}$$

For both SIMO and MISO, we match $\Psi_{\varepsilon_{kn}}(s)$ at $s = 0.2$ and $s = 1$, as prescribed in [33].

Since the ratio of two lognormal RVs is a lognormal RV, the SINR γ_{kn} in (35) for both SIMO and MISO can now be written as

$$\gamma_{kn} = \Omega_k \phi_{kn}, \tag{38}$$

where Ω_k is a lognormal RV with parameters $\mu_{\Omega_k} = \xi \log(P_T/P_N) + \mu_{\omega_{0k}}(x_{0k}) - \mu_{\varepsilon_{kn}}$ and $\sigma_{\Omega_k} = \sqrt{\sigma^2_{\omega_{0k}} + \sigma^2_{\varepsilon_{kn}}}$, and ϕ_{kn}, which tracks small-scale fading, is as defined above.

B. Cell Throughput Analysis With Scheduling

We now analyze the cell throughput per subband in the reference cell for the following three schedulers, which trade off between throughput and fairness differently:

1) Round-Robin (RR) Scheduler: It allocates subbands to users in a pre-determined, channel-agnostic manner to ensure fairness, but it does not exploit multi-user diversity. For example, in the k^{th} scheduling interval, all subbands are allocated to user k, for $1 \leq k \leq K$.

2) Greedy Scheduler: On each subband, it selects the user that fed back the highest MCS index. Hence, it exploits multi-user diversity but is not fair.

3) Modified Proportional Fair (MPF) Scheduler [34]: It selects the user with the highest ratio of reported rate to its fading-averaged value. Thus, it exploits multi-user diversity and also ensures fairness.

Let u denote the scheduled user on subband n. Let $\tilde{S}_{un}$ denote the MCS that the BS uses to transmit to user u on subband n. Since only one user is scheduled on subband n in each scheduling interval, using the law of total expectation, the average cell throughput $\bar{R}_{\text{cell}}(n)$ for subband n is

$$\bar{R}_{\text{cell}}(n) = \sum_{k=1}^{K}\sum_{m=2}^{M} r_m \text{Pr}\Big(u = k, \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big), \tag{39}$$

$$= \sum_{k=1}^{K}\sum_{m=2}^{M} r_m \text{Pr}\Big(\tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) \times \text{Pr}\Big(u = k | \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big). \tag{40}$$

Here, the expression for $\text{Pr}\Big(\tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big)$ is already derived in Section III, except for the change in notation. In the following, we present expressions for the conditional probability term $\text{Pr}\Big(u = k | \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big)$ in (40) for the three schedulers.

Result *2*: The conditional probability $\text{Pr}(u = k | \tilde{S}_{un} = m, \gamma_{un} \geq T_m)$ is given as follows:

1. *RR Scheduler:*

$$\text{Pr}\Big(u = k | \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) = \frac{1}{K}. \tag{41}$$

2. *Greedy Scheduler:*

$$\begin{aligned}\text{Pr}\Big(u = k | \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) &= \prod_{g=1, g\neq k}^{K} \text{Pr}\Big(\tilde{S}_{gn} < m\Big) + \sum_{q=1}^{K-1} \frac{1}{q+1} \\ &\times \sum_{\substack{\mathcal{V}\subset\{1,\ldots,K\}\setminus k \\ \#(V)=q}} \left[\prod_{\forall \nu \in \mathcal{V}} \text{Pr}\Big(\tilde{S}_{\nu n} = m\Big)\right] \\ &\times \left[\prod_{\forall j \in \mathcal{V}'} \text{Pr}\Big(\tilde{S}_{jn} < m\Big)\right],\end{aligned} \tag{42}$$

where $\text{Pr}\Big(\tilde{S}_{gn} < m\Big) = \sum_{p=1}^{m-1} \text{Pr}\Big(\tilde{S}_{gn} = p\Big)$. The term $1/(q+1)$ in (42) accounts for the probability of selecting user k when q other users belonging to the set $\mathcal{V}$ report the same highest MCS index $\tilde{S}_{un}$.

3. *MPF Scheduler:* For subband n of user g, let $R_{gn} \in \{r_1, \ldots, r_M\}$ denote the rate allocated by the BS. Let m_{gk} denote the index of the highest-rate MCS that is strictly less than $r_m \mathbb{E}[R_{gn}]/\mathbb{E}[R_{kn}]$, where $\mathbb{E}[R_{gn}] = \sum_{m=1}^{M} r_M \text{Pr}\Big(\tilde{S}_{gn} = m\Big)$. Then,

$$\text{Pr}\Big(u = k | \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) = \prod_{\substack{g=1, \\ g\neq k}}^{K} \left[\sum_{j=1}^{m_{gk}} \text{Pr}\Big(\tilde{S}_{gn} = j\Big)\right]. \tag{43}$$

Proof: The proof is relegated to Appendix C. ∎

We now evaluate probabilities of the form $\text{Pr}\Big(\tilde{S}_{gn} = p\Big)$ that arise in (42) and (43).

C. Evaluating Pr$\Big(\tilde{S}_{gn} = p\Big)$

From the definition of $\tilde{S}_{gn}$ in (5) and the law of total probability, we get[5]

$$\begin{aligned}\text{Pr}\Big(\tilde{S}_{gn} = p\Big) = \ &\text{Pr}(W_g = p+1, S_{gn} \leq p) \\ &+\text{Pr}(W_g = p, S_{gn} = p) \\ &+\text{Pr}(W_g = p-1, S_{gn} = p) \\ &+\text{Pr}(W_g = p-2, S_{gn} \geq p).\end{aligned} \tag{44}$$

[5]As in Section III-A, the probability terms in which $W_g \leq 0$ or $W_g > M$ are set to 0.

Writing (44) in terms of γ_{gn} and the effective SNR $\zeta_g^{(p)}$ corresponding to MCS p for user g in a manner similar to (8), and applying the simplifications described in Section III-A, we get

$$\begin{aligned}\Pr\left(\tilde{S}_{gn}=p\right) \approx\ & \Pr\left(\zeta_g^{(p-2)} \geq T_{p-2}, T_p \leq \gamma_{gn} < \infty\right) \\ & -\Pr\left(\zeta_g^{(p-1)} \geq T_{p-1}, T_{p+1} \leq \gamma_{gn} < \infty\right) \\ & +\Pr\left(\zeta_g^{(p+1)} \geq T_{p+1}, 0 \leq \gamma_{gn} \leq T_{p-1}\right) \\ & -\Pr\left(\zeta_g^{(p+2)} \geq T_{p+2}, 0 \leq \gamma_{gn} < T_p\right). \quad (45)\end{aligned}$$

The above expressions for $p = 2$ and M are similar to those in Section III-A. Each term in (45) is of the form $\Pr\left(\zeta_g^{(j)} \geq T_j, t_1 \leq \gamma_{gn} \leq t_2\right)$, where $j \in \{p-2, p-1, p+1, p+2\}$, $t_1 \in \{0, T_p, T_{p+1}\}$, and $t_2 \in \{T_{p-1}, T_p, \infty\}$. Writing in terms of the PDF $f_{\gamma_{gn}}(\cdot)$ of γ_{gn}, we get

$$\begin{aligned}&\Pr\left(\zeta_g^{(j)} \geq T_j, t_1 \leq \gamma_{gn} \leq t_2\right) \\ &\quad = \int_{t_1}^{\infty} f_{\gamma_{gn}}(v) \Pr\left(\zeta_g^{(j)} \geq T_j | \gamma_{gn} = v\right) dv \\ &\qquad - \int_{t_2}^{\infty} f_{\gamma_{gn}}(v) \Pr\left(\zeta_g^{(j)} \geq T_j | \gamma_{gn} = v\right) dv. \quad (46)\end{aligned}$$

Next, we evaluate $\int_t^{\infty} f_{\gamma_{gn}}(v) \Pr\left(\zeta_g^{(j)} \geq T_j | \gamma_{gn} = v\right) dv$, where $t \in \{t_1, t_2\}$, separately for SIMO and MISO because the PDF of ϕ_{gn} is different for them.

SIMO: From (38), the SINR γ_{gn} for user g on subband n is the product of the lognormal RV Ω_g and the gamma RV ϕ_{gn}. This is similar to the expression in (1) for the SNR except that it now tracks the user index g. Hence, the expressions derived in Section III-B directly apply to this scenario, except for the updated notation.

MISO: From (38) and Section IV-A, γ_{gn} for MISO is the product of a lognormal RV Ω_g and ϕ_{gn}, which is now the square of a Gamma RV whose PDF can be shown to be $f_{\phi_{gn}}(v) = (\sqrt{v})^{d_g-2} e^{-\frac{\sqrt{v}}{\lambda_g}} / (2\Gamma(d_g)\lambda^{d_g})$, for $v \geq 0$. The general steps (16)–(25) of Section III-B still apply in this case. The RV $X_j(i)$ in (25) is again approximated as a Beta RV as follows. From (25), the mean $\mu_j(i)$ of $X_j(i)$ is given by

$$\mu_j(i) = \mathbb{E}\left[X_j(i)\right] = \mathbb{E}\left[\exp\left(-\Delta_{y_i}^{(g)} \phi_{gn} / \beta_j\right)\right], \quad (47)$$

where $\Delta_{y_i}^{(g)} = e^{\left(\sqrt{2}\sigma_{\Omega_g} y_i + \mu_{\Omega_g}\right)/\xi}$ is the same as the constant Λ_{y_i} (cf. (20)) except that it also tracks the user index g. To compute the expectation in (47), we use the PDF of ϕ_{gn} from above and substitute $u = \sqrt{v}/\lambda_g$ in the resulting integral to get $\mu_j(i) = (1/\Gamma(d_g)) \int_0^{\infty} e^{-\left(\Delta_{y_i}^{(g)} (\lambda_g v)^2 / \beta_j\right)} u^{d_g - 1} e^{-u}\, du$. Applying generalized Gauss-Laguerre quadrature, as done in Appendix A, yields

$$\mu_j(i) \approx \frac{1}{\Gamma(d_g)} \sum_{i_2=1}^{N_{\text{GL}}} w_{\text{GL}}(d_g, i_2) e^{-\left(\frac{(\lambda_g \theta_{\text{GL}}(d_g, i_2))^2 \Delta_{y_i}^{(g)}}{\beta_j}\right)}, \quad (48)$$

where $w_{\text{GL}}(\cdot,\cdot)$ and $\theta_{\text{GL}}(\cdot,\cdot)$ are the weights and abscissas, respectively. Similarly, the variance $v_j(i) = \mathbb{E}\left[\left(X_j(i)\right)^2\right] - \left(\mu_j(i)\right)^2$ of $X_j(i)$ is given by

$$\begin{aligned}v_j(i) \approx & \frac{1}{N_{\text{SB}} - 1} \left(\frac{1}{\Gamma(d_g)} \sum_{i_2=1}^{N_{\text{GL}}} w_{\text{GL}}(d_g, i_2) \right. \\ & \left. \times \exp\left(-\frac{2\left(\lambda_g \theta_{\text{GL}}(d_g, i_2)\right)^2 \Delta_{y_i}^{(g)}}{\beta_j}\right) - \left(\mu_j(i)\right)^2 \right). \quad (49)\end{aligned}$$

Substituting (48) and (49) in (27) and (28) yields $a_j(i)$ and $b_j(i)$.

Upon substituting the Beta CDF with these parameters and the expression for $f_{\phi_{gn}}\left(v/\Delta_{y_i}^{(g)}\right)$ in (23), we get

$$\begin{aligned}&\Pr\left(\zeta_g^{(j)} \geq T_j, \gamma_{gn} \geq t\right) \\ &\approx \frac{1}{2\sqrt{\pi}\Gamma(d_g)(\lambda_g)^{d_g}} \sum_{i=1}^{N_{\text{GH}}} \frac{w_{\text{GH}}(i)}{\Delta_{y_i}^{(g)}} \\ &\quad \times \int_t^{\infty} \left(\frac{v}{\Delta_{y_i}^{(g)}}\right)^{\frac{d_g-2}{2}} e^{-\frac{1}{\lambda_g}\sqrt{\frac{v}{\Delta_{y_i}^{(g)}}}} \mathcal{B}(x(v;j); a_j(i), b_j(i))\, dv, \end{aligned} \quad (50)$$

where, similar to $x(v;l)$ in Section III-B, we have $x(v;j) = \left(N_{\text{SB}} e^{\frac{-T_j}{\beta_j}} - e^{\frac{-v}{\beta_j}}\right) / (N_{\text{SB}} - 1)$. Simplifying (50) along lines similar to (52)–(54) yields

$$\begin{aligned}&\Pr\left(\zeta_g^{(j)} \geq T_j, \gamma_{gn} \geq t\right) \\ &\approx \frac{1}{\Gamma(d_g)\sqrt{\pi}} \sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) e^{-\frac{t'}{\lambda_g}} \\ &\quad \times \sum_{i_1=0}^{d_g-1} \binom{d_g - 1}{i_1} \left(\frac{t'}{\lambda_g}\right)^{d_g - 1 - i_1} \sum_{i_2=1}^{N_{\text{GL}}} w_{\text{GL}}(i_1, i_2) \\ &\quad \times \mathcal{B}\left(x\left(\left(\lambda_g \theta_{\text{GL}}(i_1, i_2) + t'\right)^2 \Delta_{y_i}^{(g)}; j\right); a_j(i), b_j(i)\right), \end{aligned} \quad (51)$$

where $t' = \sqrt{t/\Delta_{y_i}^{(g)}}$. This completes our derivation of the cell throughput for the multi-cell scenario.

D. Extension to Multi-Stream SU-MIMO and MU-MIMO

The above analysis can be extended to multi-stream SU-MIMO, MU-MIMO, and massive MIMO modes for the single-cell scenario as follows. A tractable extension for the multi-cell scenario is an open problem. For ease of exposition, we drop the cell index l in the following discussion.

1) Open-Loop Spatial Multiplexing with a Zero-Forcing (ZF) Receiver with $N_r \geq N_t$: In this case, the BS transmits N_t parallel streams of data to user k on subband n. The user multiplies the received signal for the j^{th} stream by the j^{th} row of $\left(\mathbf{H}_{kn}{}^{\dagger}\mathbf{H}_{kn}\right)^{-1}\mathbf{H}_{kn}{}^{\dagger}$ to cancel the inter-stream interference. Hence, the SINR $\gamma_{kn}(j)$ of user k for subband n for the j^{th} stream is given by $\gamma_{kn}(j) = P_T(j)\omega_k\phi_{kn}(j)/P_N(j) = \Omega_k\phi_{kn}(j)$, where $P_T(j)$ is the transmit power of the BS

for the j^{th} stream, $P_N(j)$ is the AWGN power for the j^{th} stream, $\phi_{kn}(j) = 1 \Big/ \left[\left(\mathbf{H}_{kn}{}^{\dagger}\mathbf{H}_{kn}\right)^{-1}\right]_{jj}$ is a gamma RV with parameters $d_k = N_r - N_t + 1$ and $\lambda_k = 1$ [35], [36], and $\Omega_k = (P_T(j)/P_N(j))\omega_k$ is a lognormal RV with parameters $\mu_{\Omega_k} = \mu_{\omega_k}(x_k) + \xi \log\left(P_T(j)/P_N(j)\right)$ and $\sigma_{\Omega_k} = \sigma_{\text{shad}}$. This is now similar to (1). Therefore, the analysis subsequent to it applies.

2) Open-Loop Spatial Multiplexing with MRC at the Receiver with $N_r \geq N_t$: In this case, user k multiplies the received signal for the j^{th} stream by the $\mathbf{h}_{kn}^{\dagger}(j)$, where $\mathbf{h}_{kn}(j)$ is j^{th} column of $\mathbf{H}_{kn}$. Hence, $\gamma_{kn}(j) = P_T(j)\omega_k\phi_{kn}(j)/(P_N(j)\varepsilon_{jn})$, where $\phi_{kn}(j) = \| \mathbf{h}_{kn}(j) \|^2$ is a gamma RV with parameters $d_k = N_r$ and $\lambda_k = 1$ and $\varepsilon_{jn} = (P_T(j)/P_N(j)) \sum_{q=1,q\neq j}^{N_t} \omega_k I_{kn}(q,j) + 1$. Here, $I_{kn}(q,j) = \| \mathbf{h}_{kn}^{\dagger}(j)\mathbf{h}_{kn}(q) \|^2 / \| \mathbf{h}_{kn}(j) \|^2$ is a gamma RV with parameters $d_k = N_t - 1$ and $\lambda_k = 1$ [35]. Since $\gamma_{kn}(j)$ is similar to (35), its subsequent analysis applies.

3) MU-MIMO with ZF Precoding with $K \leq N_t$ and $N_r = 1$: In this case, the BS transmits one data stream to each user on subband n. Hence, the SINR γ_{kn} of user k on subband n is given by $\gamma_{kn} = P_T(k)\omega_k\phi_{kn}/P_N(k) = \Omega_k\phi_{kn}$, where $P_T(k)$ is the transmit power to user k, $P_N(k)$ is the AWGN power for user k, ϕ_{kn} is a gamma RV with parameters $d_k = N_t - K + 1$ and $\lambda_k = 1$ [36], and $\Omega_k = (P_T(k)/P_N(k))\omega_k$ is a lognormal RV with parameters $\mu_{\Omega_k} = \mu_{\omega_k}(x_k) + \xi \log\left(P_T(k)/P_N(k)\right)$ and $\sigma_{\Omega_k} = \sigma_{\text{shad}}$. Now γ_{kn} is similar to (1) and its subsequent analysis, therefore, applies.

4) Massive MIMO with Matched Filter (MF) Precoding with $N_t/K \to \infty$, and $N_r = 1$: In this case, due to channel hardening, it can be shown that [22] $\gamma_{kn} = P_T(k)\omega_k N_t/P_N(k)$ is a lognormal RV with parameters $\mu_{\gamma_{kn}} = \mu_{\omega_k}(x_k) + \xi \log\left(N_t P_T(k)/P_N(k)\right)$ and $\sigma_{\gamma_{kn}} = \sigma_{\text{shad}}$. Since γ_{kn} is the same for all n, this is similar to the fully correlated subbands scenario. Hence, its analysis applies.

V. Numerical Results

Simulation Setup and Parameters: The MCS set is as per LTE [2, Table 10.1]. It has $M = 16$ MCSs whose rates range from 0 to 5.55 bits/symbol. The associated decoding thresholds are generated as $T_m = (2^{R_m} - 1)/\kappa$, where κ is the coding loss factor that is set as 0.398 [6]. The values of β_l for different MCSs are given in [37], and $N_{\text{SB}} = 10$.

For the multi-cell, multi-user scenario, we consider a hexagonal cellular layout with cell radius $R = 300$ m, $C = 6$ co-channel interfering cells, and a frequency-reuse factor of one. The simulation results are averaged over 2000 realizations of small-scale and large-scale fading and 200 user drops. The user locations are distributed uniformly in the cell except for a circular area of radius 30 m around the BS [10]. The shadowing and path-loss parameters are $\sigma_{\text{shad}} = 6$ dB, $\eta = 3.5$, $x_0 = 30$ m, and $\delta_{\text{dB}}(x_0) = 10$ dB. The ratio P_T/P_N is set to achieve a pre-specified cell-corner SNR, which is varied in the figures. The analysis results are computed for each user drop using the final simplified form of (40) for each subband, and are then numerically averaged over the drops and subbands.

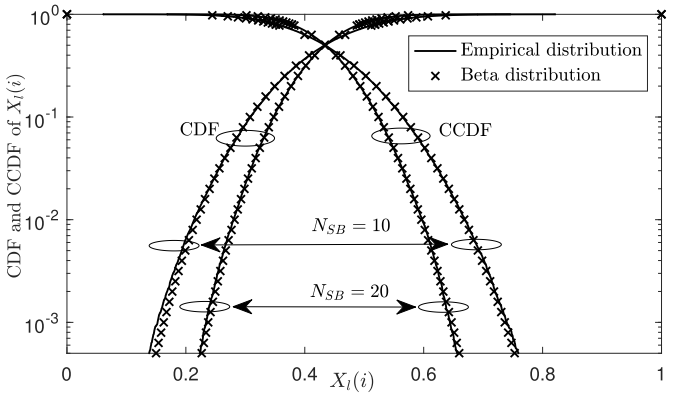

Fig. 2. CDF and CCDF of $X_l(i)$ for i.i.d. subband fading gains for SISO for $N_{\text{SB}} = 10$ and 20 ($l = 11$ and $\Lambda_{y_i} = 10$ dB).

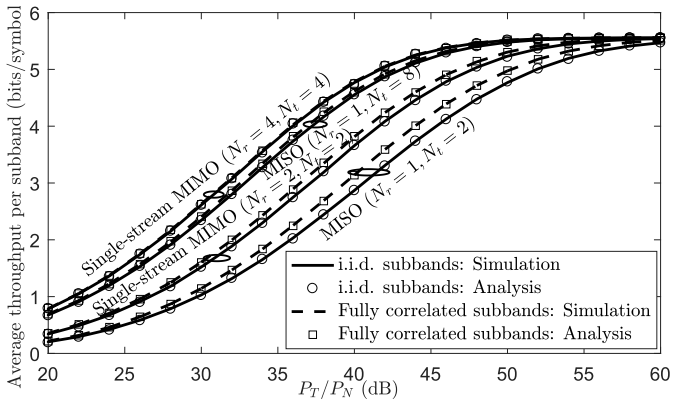

Fig. 3. Single-user scenario: Average throughput per subband of HLCS feedback for MISO and single-stream MIMO as a function of $\frac{P_T}{P_N}$ ($N_{\text{SB}} = 10$).

Statistics of $X_l(i)$ in EESM Expressions: We compare in Fig. 2 the CDF and the complementary CDF (CCDF) of the Beta RV with the empirical CDF and CCDF of $X_l(i)$ [6], [9]. The CDF captures the accuracy of the fit for small values of $X_l(i)$, but saturates to 1 for larger values. The CCDF helps visualize the accuracy of the fit better in the latter regime. This comparison differs from that in [6], which models $\zeta^{(l)}$ as a lognormal RV, and that in [9], which plots the CDF and CCDF of the RV $-\beta_l \log\left((1/N_{\text{SB}}) \sum_{n=1}^{N_{\text{SB}}} \exp\left(-\gamma_n/\beta_l\right)\right)$. This is done for MCS $l = 11$, which uses 64QAM and a 0.46 rate code, and SISO. We see that the Beta approximation is accurate over a three orders of magnitude range for $N_{\text{SB}} = 10$ and 20. The accuracy of the fit increases as N_{SB} increases, since the aforementioned central limit approximation of Papoulis [28] becomes more accurate.

Single-User Scenario: Fig. 3 plots for HLCS feedback the average throughput per subband of a user as a function of P_T/P_N for MISO with $N_t = 2$ and 8, and single-stream MIMO with $N_r = 2, N_t = 2$ and $N_r = 4, N_t = 4$, when the subband fading gains are i.i.d. The analysis curves are plotted using the final simplified form of (6) that is obtained in Section III. Also shown for comparison are the corresponding results when the subbands fading gains are fully correlated. As we saw in Section III-D.2, this has the same throughput as full CSI feedback for any correlation. Hence, the corresponding analysis curves are plotted using (33).

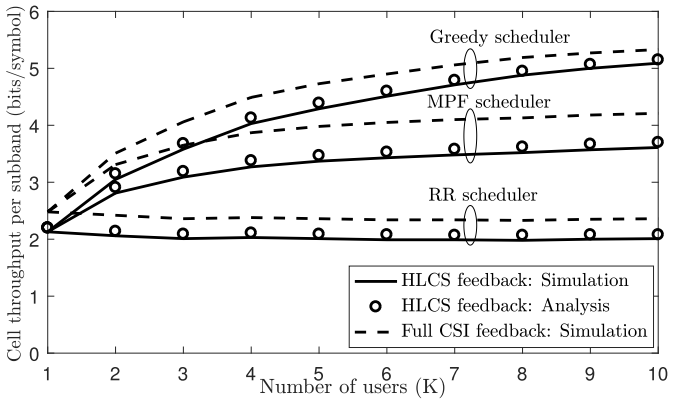

Fig. 4. Multi-cell, multi-user scenario: Cell throughput per subband of the three schedulers as a function of number of users K for MISO ($N_r = 1$, $N_t = 4$, $N_{\text{SB}} = 10$, and cell-corner SNR of 5 dB).

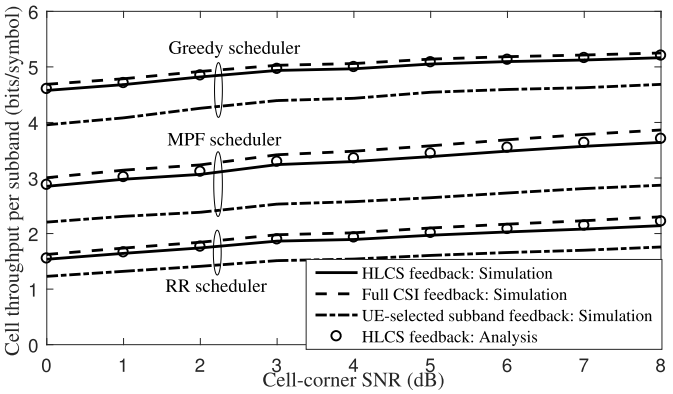

Fig. 5. Multi-cell, multi-user scenario: Cell throughput per subband of the three schedulers as a function of the cell-corner SNR for SIMO ($N_r = 4$, $N_t = 1$, $N_{\text{SB}} = 10$, and $K = 10$).

In all cases, we observe a good match between the analytical and simulation results over a large 40 dB range for P_T/P_N. We observe that the loss in throughput due to differential feedback is marginal as the number of antennas increases. In all cases, the throughput increases as P_T/P_N increases and saturates to 5.55 bits/symbol, which is the rate of the highest-rate MCS.

Multi-Cell, Multi-User Scenario with Scheduling: Fig. 4 plots the cell throughput per subband of HLCS feedback and full CSI feedback as a function of K for MISO ($N_t = 4$) for the three schedulers. We observe a good match between the analytical and simulation results. For any given K, we see that the RR scheduler has the lowest cell throughput and the greedy scheduler the highest. The cell throughputs of the greedy and MPF schedulers increase as K increases and saturate to 5.55 bits/symbol. A new observation is that compared to full CSI feedback, the reduction in throughput due to differential feedback is the most for the MPF scheduler. However, even here, it is at most 15.50% for all K.

Fig. 5 plots the cell throughput per subband of HLCS feedback as a function of the cell-corner SNR for SIMO ($N_r = 4$) for the three schedulers with $q = 2$ PRBs. It also compares the performance of HLCS feedback (24 bits overhead), full CSI feedback (40 bits overhead), and UE-selected subband feedback (17 bits overhead when the best three subbands are selected [2, Ch. 10]). For a given cell-corner SNR, the throughput of HLCS feedback is nearly the same as that of full CSI feedback for all three schedulers and exceeds that of UE-selected subband feedback. We note, however, that the feedback overheads of the three schemes are different.

VI. Conclusions

We presented an accurate analysis of the throughput of the differential feedback-based HLCS feedback scheme that is used in both 4G and 5G standards. We did so for different single-stream multi-antenna modes for the single-user scenario and the multi-cell, multi-user scenario with co-channel interference. EESM enabled an accurate and tractable modeling of the generation of the wideband and differential CQIs. We saw that when the subband fading gains were fully correlated, differential feedback was as good as full CSI feedback. Otherwise, at low SINRs, it incurred a marginal reduction in throughput. We also observed that the MPF scheduler was more sensitive to differential feedback. An interesting avenue for future work is to extend the proposed analysis to different multi-stream SU-MIMO modes with $N_r < N_t$, MU-MIMO, and massive MIMO for the multi-cell scenario.

Appendix

A. Expression for $Pr\big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\big)$ in (32)

In the integral in (31), substituting the gamma PDF $f_{\phi_n}(v/\Lambda_{y_i})$ from Section II-B.2 and thereafter using the variable transformation $z = (v/(\Lambda_{y_i}\lambda)) - (t/(\Lambda_{y_i}\lambda))$, we get

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) &= \frac{1}{\Gamma(d)\sqrt{\pi}} \sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) \\ &\quad \times \int_0^\infty \left(\frac{t}{\Lambda_{y_i}\lambda} + z\right)^{d-1} e^{-\left(z + \frac{t}{\Lambda_{y_i}\lambda}\right)} \\ &\quad \times \mathcal{B}\left(x\left(\lambda z \Lambda_{y_i} + t; l\right); a_l(i), b_l(i)\right) dz. \end{aligned} \tag{52}$$

Expanding the $(t/(\Lambda_{y_i}\lambda) + z)^{d-1}$ term as a binomial series,

$$\begin{aligned}\Pr\Big(\zeta^{(l)} \geq T_l, \gamma_n \geq t\Big) &= \frac{1}{\Gamma(d)\sqrt{\pi}} \sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) e^{\frac{-t}{\Lambda_{y_i}\lambda}} \\ &\quad \times \sum_{i_1=0}^{d-1} \binom{d-1}{i_1} \left(\frac{t}{\Lambda_{y_i}\lambda}\right)^{d-1-i_1} \int_0^\infty z^{i_1} e^{-z} \\ &\quad \times \mathcal{B}\left(x\left(\lambda z \Lambda_{y_i} + t; l\right); a_l(i), b_l(i)\right) dz. \end{aligned} \tag{53}$$

The integral $\int_0^\infty z^{i_1} e^{-z} \mathcal{B}\left(x\left(\lambda z \Lambda_{y_i} + t; l\right); a_l(i), b_l(i)\right) dz$ in (53) is of the form $\int_0^\infty z^{i_1} e^{-z} f(z) dz$, which can be accurately approximated by generalized Gauss-Laguerre quadrature as follows [29]:

$$\int_0^\infty z^{i_1} e^{-z} f(z) dz \approx \sum_{i_2=0}^{N_{\text{GL}}} w_{\text{GL}}(i_1, i_2)\, f\left(\theta_{\text{GL}}(i_1, i_2)\right), \tag{54}$$

where $w_{\text{GL}}(\cdot,\cdot)$ and $\theta_{\text{GL}}(\cdot,\cdot)$ are the weights and abscissas, respectively. The error in the approximation decreases as N_{GL} increases. Applying this in (53) yields (32).

B. Proof of Result 1

In full CSI feedback, the CQI fed back by the user for subband n is the index of the MCS that can be decoded by the user on that subband. Hence, $\Pr\Big(\tilde{S}_n = m, \gamma_n \geq T_m\Big) = \Pr(T_m \leq \gamma_n < T_{m+1})$. This holds even for correlated subband fading gains. Writing in terms of the PDFs of Ω and ϕ_n,

$$\Pr(T_m \leq \gamma_n < T_{m+1}) = \frac{\xi}{\sqrt{2\pi}\sigma_\Omega}\int_0^\infty \frac{1}{u^2} e^{-\left(\frac{\xi\log(u)-\mu_\Omega}{\sqrt{2}\sigma_\Omega}\right)^2} \times \int_{T_m}^{T_{m+1}} f_{\phi_n}\Big(\frac{v}{u}\Big)\, dv\, du. \quad (55)$$

Substituting $y = (\xi\log(u) - \mu_\Omega)/\big(\sqrt{2}\sigma_\Omega\big)$ and applying Gauss-Hermite quadrature yields

$$\Pr\Big(\tilde{S}_n = m, \gamma_n \geq T_m\Big) \approx \frac{1}{\sqrt{\pi}}\sum_{i=1}^{N_{\text{GH}}} \frac{w_{\text{GH}}(i)}{\Lambda_{y_i}} \times \int_{T_m}^{T_{m+1}} f_{\phi_n}\left(\frac{v}{\Lambda_{y_i}}\right) dv, \quad (56)$$

where $w_{\text{GH}}(i)$ and y_i, for $1 \leq i \leq N_{\text{GH}}$, are the weights and abscissas, respectively, and Λ_{y_i} is defined in (20). Substituting the expression for the gamma PDF $f_{\phi_n}(\cdot)$, the integral in (56) can be written in terms of the upper incomplete gamma function $U(\cdot,\cdot)$ as follows:

$$\Pr\Big(\tilde{S}_n = m, \gamma_n \geq T_m\Big) = \frac{1}{\Gamma(d)\sqrt{\pi}}\sum_{i=1}^{N_{\text{GH}}} w_{\text{GH}}(i) \times \left[U\left(\frac{T_m}{\Lambda_{y_i}\lambda}, d\right) - U\left(\frac{T_{m+1}}{\Lambda_{y_i}\lambda}, d\right)\right]. \quad (57)$$

Substituting (57) in (6) yields (33).

C. Deriving $Pr\Big(u = k|\tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big)$

1) RR Scheduler: Since it is feedback-agnostic, the probability of a user getting selected is the same for all K users and is, therefore, equal to $1/K$.

2) Greedy Scheduler: Let $\mathcal{V}$ be the set of users other than user k that report their highest MCS as $\tilde{S}_{un}$. Let $\#(\mathcal{V}) = q$. User k is selected for subband n in the following two cases:

(i) $q = 0$: In this case, the MCS $\tilde{S}_{un}$ reported by user k is higher than those reported by all the other users. Hence,

$$\Pr\Big(u = k|\tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) = \Pr\Big(\tilde{S}_{gn} < \tilde{S}_{un}, \forall g \neq k \,\Big|\, \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big), \quad (58)$$

$$= \prod_{g=1, g\neq k}^{K} \Pr\Big(\tilde{S}_{gn} < m\Big), \quad (59)$$

where (59) follows because the subband SINRs of different users are mutually independent.

(ii) $q \geq 1$: In this case, we break ties at random [10]. Therefore, user k is selected with probability $1/(q+1)$. Consequently, writing in terms of all possible realizations of $\mathcal{V}$, we get

$$\Pr\Big(u = k|\tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) = \sum_{q=1}^{K-1}\frac{1}{q+1}\sum_{\substack{\mathcal{V}\subset\{1,\ldots,K\}\setminus k \\ \#(V)=q}} \Pr\Big(\tilde{S}_{\nu n} = m, \tilde{S}_{jn} < m, \forall \nu \in \mathcal{V}, \forall j \in \mathcal{V}' \,\Big|\, \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big), \quad (60)$$

$$= \sum_{q=1}^{K-1}\frac{1}{q+1}\sum_{\substack{\mathcal{V}\subset\{1,\ldots,K\}\setminus k \\ \#(V)=q}} \left[\prod_{\forall \nu\in\mathcal{V}} \Pr\Big(\tilde{S}_{\nu n} = m\Big)\right] \times \left[\prod_{\forall j\in\mathcal{V}'} \Pr\Big(\tilde{S}_{jn} < m\Big)\right]. \quad (61)$$

Combining (59) and (61) yields (42).

3) MPF Scheduler: It schedules user k on subband n if it has the highest ratio of the rate reported to its average rate. Since this ratio is a real number, ties between users occur with zero probability and need not be considered. Furthermore, the event $R_{gn} = r_m$ is the same as the event $\tilde{S}_{gn} = m$. Thus,

$$\Pr\Big(u = k \,\Big|\, \tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) = \Pr\Big(\frac{R_{gn}}{\mathbb{E}[R_{gn}]} < \frac{R_{kn}}{\mathbb{E}[R_{kn}]}, \forall g \neq k \,\Big|\, R_{kn} = r_m, \gamma_{kn} \geq T_m\Big), \quad (62)$$

$$= \Pr\Big(R_{gn} < r_{m_{gk}}, \forall g \neq k \,\Big|\, R_{kn} = r_m, \gamma_{kn} \geq T_m\Big), \quad (63)$$

where m_{gk} is as defined in the result statement. Since the MCSs reported by different users are independent, we get

$$\Pr\Big(u = k|\tilde{S}_{un} = m, \gamma_{un} \geq T_m\Big) = \prod_{\substack{g=1, \\ g\neq k}}^{K} \Pr\big(R_{gn} < r_{m_{gk}}\big). \quad (64)$$

Furthermore, $\Pr\big(R_{gn} < r_{m_{gk}}\big) = \sum_{j=1}^{m_{gk}} \Pr(R_{gn} = r_j) = \sum_{j=1}^{m_{gk}} \Pr\Big(\tilde{S}_{gn} = j\Big)$. Substituting this in (64) yields (43).

REFERENCES

[1] V. Kumar and N. B. Mehta, "Modeling and performance analysis of differential CQI feedback in OFDM cellular systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[2] S. Sesia, I. Toufik, and M. Baker, *LTE—The UMTS long Term Evolution*, 2nd ed. Hoboken, NJ, USA: Wiley, 2009.

[3] *NR—Physical Layer Procedures for Data*, document 3GPP TS 38.214, v15.2.0, 2018.

[4] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[5] S. Sanayei and A. Nosratinia, "Opportunistic downlink transmission with limited feedback," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4363–4372, Nov. 2007.

[6] S. N. Donthi and N. B. Mehta, "Joint performance analysis of channel quality indicator feedback schemes and frequency-domain scheduling for LTE," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 3096–3109, Sep. 2011.

[7] J. Leinonen, J. Hämäläinen, and M. Juntti, "Performance analysis of downlink OFDMA resource allocation with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2927–2937, Jun. 2009.

[8] J. Chen, R. A. Berry, and M. L. Honig, "Limited feedback schemes for downlink OFDMA based on sub-channel groups," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1451–1461, Oct. 2008.
[9] J. Francis and N. B. Mehta, "EESM-based link adaptation in point-to-point and multi-cell OFDM systems: Modeling and analysis," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 407–417, Jan. 2014.
[10] J. Francis and N. B. Mehta, "Characterizing the impact of feedback delays on wideband rate adaptation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 960–971, Feb. 2015.
[11] S. N. Donthi and N. B. Mehta, "An accurate model for EESM and its application to analysis of CQI feedback schemes and scheduling in LTE," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3436–3448, Oct. 2011.
[12] A. Masaracchia, R. Bruno, A. Passarella, and S. Mangione, "Analysis of MAC-level throughput in LTE systems with link rate adaptation and HARQ protocols," in *Proc. IEEE 16th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2015, pp. 1–9.
[13] E. Westman, "Calibration and evaluation of the exponential effective SINR mapping (EESM) in 802.16," M.S. thesis, Dept. Elect. Eng., KTH Roy. Inst. Technol., Stockholm, Sweden, Sep. 2006.
[14] H. Song, R. Kwan, and J. Zhang, "General results on SNR statistics involving EESM-based frequency selective feedbacks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1790–1798, May 2010.
[15] G. Ku and J. M. Walsh, "Resource allocation and link adaptation in LTE and LTE advanced: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1605–1633, 3rd Quart., 2015.
[16] L. Wan, S. Tsai, and M. Almgren, "A fading-insensitive performance metric for a unified link quality model," in *Proc. IEEE Wireless Commun. Netw. conf. (WCNC)*, Apr. 2006, pp. 2110–2114.
[17] H. T. Nguyen, I. Z. Kovacs, Y. Wang, and K. I. Pedersen, "Feedback compression schemes for downlink carrier aggregation in LTE-Advanced," in *Proc. IEEE IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2011, pp. 1–5.
[18] J. Duplicy *et al.*, "MU-MIMO in LTE systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2011, no. 1, Mar. 2011, Art. no.496763. doi: 10.1155/2011/496763.
[19] Y.-P. Zhang, P. Wang, S. Feng, P. Zhang, and S. Tong, "On the efficient channel state information compression and feedback for downlink MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 63, no. 7, pp. 3263–3275, Sep. 2014.
[20] Y. Jeon, H.-M. Kim, Y.-S. Cho, and G.-H. Im, "Time-domain differential feedback for massive MISO-OFDM systems in correlated channels," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 630–642, Feb. 2016.
[21] *NR—Physical Channels and Modulation*, document 3GPP TS 38.211, v15.2.0, 2018.
[22] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
[23] G. L. Stüber, *Principles Mobile Communcation*, 4th ed. New York, NY, USA: Springer, 2017.
[24] Y. Li, L. Zhang, L. J. Cimini, and H. Zhang, "Statistical analysis of MIMO beamforming with co-channel unequal-power MIMO interferers under path-loss and Rayleigh fading," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3738–3748, Aug. 2011.
[25] L. S. Gradshteyn and L. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.
[26] M. López-Benítez, "Throughput performance models for adaptive modulation and coding under fading channels," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2016, pp. 1–6.
[27] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, 9th ed. New York, NY, USA: Dover, 1972.
[28] A. Papoulis, *The Fourier Integral and Its Applications*, 1st ed. New York, NY, USA: Mcgraw-Hill, 1962.
[29] P. Rabinowitz and G. Weiss, "Tables of abscissas and weights for numerical evaluation of integrals of the form $\int_0^\infty e^{-x} X^{nf}(X) dx$," *Math. Tables Aids Comput.*, vol. 13, no. 68, pp. 285–294, Oct. 1959.
[30] D. Gilat, "Some conditions under which two random variables are equal almost surely and a simple proof of a theorem of chung and fuchs," *Ann. Math. Statist.*, vol. 42, no. 5, pp. 1647–1655, Oct. 1971.
[31] J. L. Doob, *Stochastic Processes*, 1st ed. Hoboken, NJ, USA: Wiley, 1953.
[32] A. Shah and A. M. Haimovich, "Performance analysis of maximal ratio combining and comparison with optimum combining for mobile radio communications with cochannel interference," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1454–1463, Jul. 2000.
[33] N. B. Mehta, J. Wu, A. F. Molisch, and J. Zhang, "Approximating a sum of random variables with a lognormal," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2690–2699, Jul. 2007.
[34] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 766–778, Mar. 2007.
[35] R. H. Y. Louie, M. R. McKay, and I. B. Collings, "Open-loop spatial multiplexing and diversity communications in ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 317–344, Jan. 2011.
[36] X. Yu, C. Li, J. Zhang, M. Haenggi, and K.-B. Letaief, "A unified framework for the tractable analysis of multi-antenna wireless networks," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 7965–7980, Dec. 2018.
[37] J. Fan, Q. Yin, G. Y. Li, B. Peng, and X. Zhu, "MCS selection for throughput improvement in downlink LTE systems," in *Proc. 20th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2011, pp. 1–5.

Vineeth Kumar (S'16) received the B.Tech. degree in electronics and communications engineering from Amrita University, Kerala, in 2011, and the M.Tech. degree in signal processing from the National Institute of Technology Calicut in 2014. He is currently pursuing the Ph.D. degree with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru. His research interests include the design and performance analysis of limited feedback schemes for rate adaptation and resource allocation in orthogonal frequency-division multiplexing systems.

Neelesh B. Mehta (S'98–M'01–SM'06–F'19) received the B.Tech. degree in electronics and communications engineering from IIT Madras in 1996 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, CA, USA, in 1997 and 2001, respectively. He is currently a Professor with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru. He is a Fellow of the Indian National Academy of Engineering and The National Academy of Sciences, India. He was a recipient of the Shanti Swarup Bhatnagar Award 2017 and the Swarnajayanti Fellowship. He has served on the Board of Governors of the IEEE ComSoc from 2012 to 2015. He has served on the Executive Editorial Committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2014 to 2017, and as its Chair from 2017 to 2018. He is currently an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS.