

Performance Analysis of Centralized and Partially Decentralized Co-Operative Networks

Rajitha Senanayake, *Member, IEEE*, Phee Lep Yeoh, *Member, IEEE*, and Jamie S. Evans, *Member, IEEE*

Abstract—We consider cellular networks with co-operative clusters of neighboring base stations detecting multiple *in-cluster* users subject to interference from *out-of-cluster* users. We assume that the base stations, equipped with multiple antennas, are connected to a central processor in each cluster. For such a network, we first consider centralized processing where all the *in-cluster* user signals are sent to the central processor for linear minimum mean squared error (LMMSE) estimation. Next, we consider partially decentralized processing where the *in-cluster* user signals are locally estimated at each base station, and the local estimates are combined at the central processor. For both processing architectures, we derive new expressions for the achievable rate of an *in-cluster* user when the channels between the users and base stations are subject to independent Rayleigh fading and distance-dependent path loss. The solutions are based on accurate approximations we derive for the characteristic function (CF) and the probability density function (PDF) of each user's signal-to-interference-plus-noise ratios (SINRs). Numerical examples highlight the accuracy of the analysis and compare the performance of centralized and partially decentralized processing under different cluster scenarios.

Index Terms—Base station co-operation, multiuser detection, LMMSE, Rayleigh fading.

I. INTRODUCTION

THE evolution of data traffic in recent years has witnessed an unprecedented growth in demand for high capacity wireless communications. In 2020, it is anticipated that fifth generation (5G) wireless networks will serve a hundred-fold increase in wireless devices and deliver a thousand-fold increase in data traffic [1], [2]. To meet this demand, 5G networks are seeking to overcome the major barrier of inter-cell interference through a more proactive interference-aware coordination of multiple cells, often termed as base station cooperation. This allows wireless networks to exploit interference by treating antennas of multiple cells as a virtual multiple antenna array [2]. In so doing, base station cooperation effectively treats inter-cell interference as useful information and significantly improves the reliability [3]–[7] and spectral efficiency [8]–[13].

Manuscript received May 6, 2015; revised September 22, 2015; accepted December 21, 2015. Date of publication December 29, 2015; date of current version February 12, 2016. The associate editor coordinating the review of this paper and approving it for publication was A. Ghayeb.

R. Senanayake and P. L. Yeoh are with the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, Vic. 3010, Australia (e-mail: r.senanayake@student.unimelb.edu.au; phee.yeoh@unimelb.edu.au).

J. S. Evans is with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Vic. 3800, Australia (e-mail: jamie.evans@monash.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2015.2512918

Due to limitations in the backhaul network, cooperation has to be limited to a finite number of geographically neighboring base stations which partitions the network into cooperating clusters. This suggests the formation of two classes of users—*in-cluster* users and *out-of-cluster* users. The importance of cluster-based cooperation is discussed in [14]–[16] where it was shown that the spectral efficiency of base station cooperation saturates above a certain transmit power threshold when subject to interference from *out-of-cluster* users. To characterize the performance of cluster-based cooperation, the fully cooperative linear Wyner model in [8] was limited to partial cooperation between neighbouring base stations in [17] where the asymptotic multiplexing gain per user was shown to grow monotonically with the number of cooperating transmitters and receivers. In [18], compressive sensing and spectral clustering were applied to select the set of cooperating base stations and decode the user signals using a sparse linear receive-filter. In [19], user power control with an energy-efficient opportunistic transmission strategy was proposed to reduce the inter-cluster interference for a one-dimensional cellular grid. In [20], [21], user scheduling and clustering based on the angular spread of the multipath channels was investigated to increase the spectral efficiency of two-dimensional multi-cell massive MIMO networks. More recently in [22], a distributed pricing-based optimization algorithm was proposed to maximize the weighted sum-rate with local maximal ratio-combining (MRC) in multi-cell cooperative networks.

In this paper, we consider the performance of cluster-based cooperation from the perspective of two practical signal processing architectures. First, we propose a centralized processing architecture where all the received signals at the base station antennas are sent to the central processor along with their individual channel state information (CSI) of fading gains and path gains [23]. In this architecture, no processing is done at the base stations. We employ a linear minimum mean squared error (LMMSE) estimator at the central processor to detect the *in-cluster* users while treating *out-of-cluster* users as interference. This processing architecture offers attractive performance gains when compared to a traditional non cooperative network. We also note that the processing complexity of centralized processing is less than fully cooperative networks, as the cooperation is limited to a set of geographically neighboring base stations. However, as the cluster size increases, we note that collecting all the received signals at one central location may not be feasible because of the wide geographical separation of base stations. Furthermore, the signal processing at the central processor can easily become computationally burdensome as the size of the cluster grows large.

Second, we propose a partially decentralized processing architecture where the received signals at each base station are locally estimated and the estimates of each user and an appropriate set of weights are sent to the central processor [24]. In the partially distributed processing architecture, the processing complexity at the central processor is reduced by distributing the signal processing among the base stations. The estimation of user signals is done in two steps: 1) The in-cluster user signals are estimated at each base station using LMMSE estimation, 2) The local estimates are optimally combined at the central processor such that the signal-to-interference-plus noise ratio (SINR) of a given user is maximized. This results in a lower signal processing load at the central processor and eliminates the need to feedback all CSI to the central processor.

In this paper, we examine the achievable rates of centralized and partially decentralized processing and derive new analytical expressions to compare the two architectures under different cluster-based network scenarios. Our novel contributions are detailed as follows:

- New closed-form expressions for the approximate characteristic function (CF) and probability density function (PDF) of the received SINR of an in-cluster user are derived for centralized and partially decentralized processing. Based on these, accurate approximations on the achievable rate of the in-cluster users are derived. While centralized processing outperforms partially decentralized processing, our results show that the performance gap between them is small.
- Our analytical expressions establish that the achievable rate saturates above a certain transmit power threshold, similar to the spectral efficiency in [14]. As such, we confirm that the achievable rate with clustered cooperation does not improve continuously with the SNR.
- We highlight that a fixed cluster arrangement results in a lower achievable rate due to out-of-cluster interference impacting on users at the cluster edge. We show that a dynamic cluster arrangement can improve the achievable rate by ensuring that all users are located in the center of their cooperating cluster, thus reducing the severity of out-of-cluster interference. Interestingly, we find that partially decentralized processing with dynamic clustering outperforms centralized processing with fixed clustering.
- We examine the effect of the path loss exponent on the achievable rate of cluster-based cooperation. We find that the achievable rate of both centralized and decentralized processing with out-of-cluster interference improves with increasing path loss exponent. This further illustrates the novel interference limited behavior of cluster-based cooperation.

The rest of the paper is organized as follows. In Section II we present the system model of the cluster-based cooperative network. The centralized processing architecture is analyzed in Section III and the partially decentralized processing architecture is analyzed in Section IV, where the CF, PDF and the achievable rate results are derived and discussed. Numerical examples are illustrated in Section V, followed by concluding remarks in Section VI.

II. SYSTEM DESCRIPTION

We consider the uplink of a cluster of R cooperating base stations, where each base station is equipped with K receive antennas. Due to the finite size of the cluster, the transmissions from N users within the cluster are necessarily subject to interference from \tilde{N} users outside the cluster. We assume that the base stations are linked to a central processor by high capacity delay-free links. As such, the received signal at antenna k of base station r can be written as

$$y_{rk} = \sum_{n=1}^N h_{rkn} s_n + \underbrace{\sum_{n=N+1}^{N+\tilde{N}} h_{rkn} s_n + z_{rk}}_{w_{rk}}, \quad (1)$$

where $h_{rkn} \sim \mathcal{CN}(0, g_{rkn}^2)$ is the Rayleigh distributed complex channel gain between user n and antenna k of base station r , g_{rkn} is the distance dependent path gain between user n and antenna k of base station r , s_n is the transmitted symbol from user n with $E\{|s_n|^2\} = 1$, and $z_{rk} \sim \mathcal{CN}(0, \sigma^2)$ is the complex Gaussian noise at antenna k of base station r . Note that the use of the Rayleigh distribution to model small scale fading is realistic in urban environments, where the channel is subject to a lot of reflections, refractions and scattering from the surrounding surfaces. Given that the second summation in (1) represents the out-of-cluster interference¹, we define the total noise plus out-of-cluster interference at antenna k of base station r as w_{rk} which has a variance of $\sigma_{rk}^2 = E\{w_{rk} w_{rk}^*\} = \sum_{n=N+1}^{N+\tilde{N}} g_{rkn}^2 + \sigma^2$.

A. Centralized Processing Architecture

Under the centralized processing architecture, illustrated in Fig. 1, all the $R \times K$ received signals at the base station antennas are sent to the central processor to jointly estimate the in-cluster users. Thus, the $\mathcal{C}^{RK \times 1}$ received vector at the central processor can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{w}, \quad (2)$$

where $\mathbf{s} = (s_1, s_2, \dots, s_N)^T$ is the $\mathcal{C}^{N \times 1}$ transmitted symbol vector from the N in-cluster users, $\mathbf{w} = (w_{11}, w_{12}, \dots, w_{RK})^T$ is the $\mathcal{C}^{RK \times 1}$ noise plus out-of-cluster interference vector, and \mathbf{H} is the $\mathcal{C}^{RK \times N}$ in-cluster channel matrix with independent entries, h_{rkn} .

We consider LMMSE estimation at the central processor, where a weighted matrix \mathbf{V} is applied to the received vector \mathbf{y} such that the mean squared error $E\{\|\mathbf{s} - \mathbf{V}^H \mathbf{y}\|^2\}$ is minimized. The notation $(A)^H$ denotes the Hermitian transpose of matrix A . The estimated symbol for an in-cluster user n is given by

$$\hat{s}_n = \mathbf{v}_n^H \mathbf{y} \quad (3)$$

where \mathbf{v}_n is the weight vector for user n defined as [25]

$$\mathbf{v}_n = \left(\mathbf{H}\mathbf{H}^H + \Sigma \right)^{-1} \mathbf{h}_n. \quad (4)$$

¹It is important to note that we do not need the out-of-cluster interference plus noise to be Gaussian, as the analysis presented in this paper only requires the second order statistics of the out-of-cluster interference plus noise term.

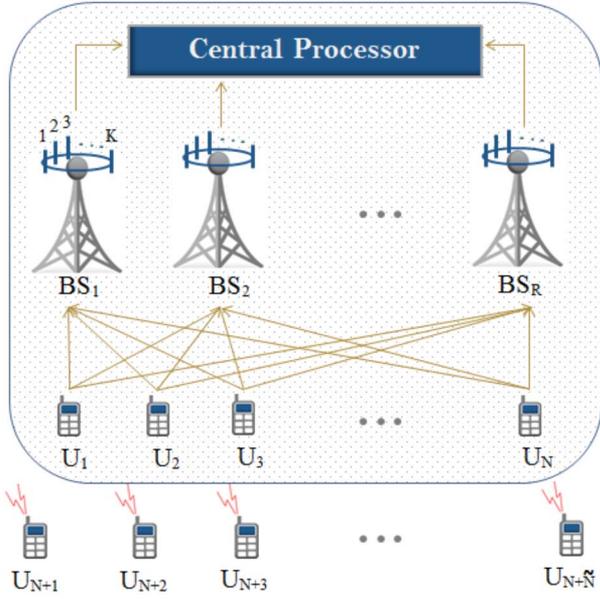


Fig. 1. A Centralized processing architecture with N in-cluster users and \tilde{N} out-of-cluster users.

In (4), $\mathbf{h}_n = [h_{11n}, h_{12n}, \dots, h_{RK n}]^T$ is the n -th column vector of \mathbf{H} containing the channel gains from user n to all the base station antennas, and

$$\Sigma = E\{\mathbf{w}\mathbf{w}^H\} = \text{diag}(\sigma_{11}^2, \sigma_{12}^2, \dots, \sigma_{RK}^2), \quad (5)$$

is the covariance matrix of the noise plus out-of-cluster interference which simplifies to a diagonal matrix due to the independence of terms in \mathbf{w} for each base station. In the signal estimation, we adopt the practical assumption of cluster-based cooperation where the base stations do not have knowledge of the individual channel gains of the out-of-cluster users and only utilize knowledge of the noise plus out-of-cluster interference variance, σ_{rk}^2 , in the weight vector.

Based on \hat{s}_n , we can write the instantaneous SINR of an in-cluster user n at the central processor as [25]

$$\gamma_n = \frac{\mathbf{v}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{v}_n}{\mathbf{v}_n^H \left[\sum_{k \neq n}^{N+\tilde{N}} \mathbf{h}_k \mathbf{h}_k^H + \sigma^2 \mathbf{I} \right] \mathbf{v}_n}, \quad (6)$$

where the denominator contains the instantaneous channel gains of both in-cluster and out-of-cluster users and \mathbf{I} is the identity matrix. We note that the instantaneous channel gains of the in-cluster users tend to be larger than that of the out-of-cluster users due to the distance dependence path loss. As such, in the following analysis we consider an approximation for the SINR of cluster-based cooperative networks given by

$$\gamma_n \approx \frac{\mathbf{v}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{v}_n}{\mathbf{v}_n^H [\tilde{\mathbf{H}}_n \tilde{\mathbf{H}}_n^H + \Sigma] \mathbf{v}_n} = \mathbf{h}_n^H \mathbf{R}_n^{-1} \mathbf{h}_n, \quad (7)$$

where we approximate the instantaneous channel gains of the out-of-cluster users with their expected values, and define the covariance matrix \mathbf{R}_n as

$$\mathbf{R}_n = \tilde{\mathbf{H}}_n \tilde{\mathbf{H}}_n^H + \Sigma, \quad (8)$$

with $\tilde{\mathbf{H}}_n \tilde{\mathbf{H}}_n^H = \sum_{k \neq n}^N \mathbf{h}_k \mathbf{h}_k^H$ containing the instantaneous channel gains of the other $N - 1$ in-cluster users and Σ containing the variances of the noise plus out-of-cluster interference. We will illustrate the accuracy of this approximation in Section V through numerical examples. It is important to point out that the SINR approximation in (7) is different from that in previous works such as [7] which analyzed base station cooperation without out-of-cluster interference.

B. Partially Decentralized Processing Architecture

Under the partially decentralized processing architecture the in-cluster user symbols are estimated according to a two-tier LMMSE estimation. In the first tier, we apply local LMMSE estimation at each base station to estimate the received signals from the N in-cluster users at the K antennas. As such, a local estimate of each in-cluster user $n \in N$ is made at each base station. In the second tier, we consider global LMMSE estimation at the central processor to combine the local estimates sent from the R base stations, such that the SINR of the in-cluster user n is maximized. These two tiers are further detailed next.

1) *Local Estimation:* The $\mathcal{C}^{K \times 1}$ received vector for local estimation at base station r can be written as

$$\mathbf{y}_r = \mathbf{H}_r \mathbf{s} + \mathbf{w}_r, \quad (9)$$

where \mathbf{H}_r is the $\mathcal{C}^{K \times N}$ channel matrix of independent channel coefficients from the in-cluster users to K antennas at base station r . The term $\mathbf{w}_r = (w_{r1}, w_{r2}, \dots, w_{rK})^T$ is the $\mathcal{C}^{K \times 1}$ noise plus out-of-cluster interference vector at base station r . We consider local LMMSE estimation at base station r , where a weighted vector \mathbf{v}_r is applied to the received vector \mathbf{y}_r such that the mean squared error $E\{\|\mathbf{s} - \mathbf{v}_r^H \mathbf{y}_r\|^2\}$ is minimized. This is essentially equivalent to applying the centralized processing architecture described in Section II-A at a single base station. As such, the estimated symbol for in-cluster user n at base station r can be written as

$$\hat{s}_{rn} = \mathbf{v}_{rn}^H \mathbf{y}_r, \quad (10)$$

where

$$\mathbf{v}_{rn} = \left(\mathbf{H}_r \mathbf{H}_r^H + \Sigma_r \right)^{-1} \mathbf{h}_{rn}, \quad (11)$$

with $\mathbf{h}_{rn} = [h_{r1n}, \dots, h_{rKn}]^T$ and $\Sigma_r = E\{\mathbf{w}_r \mathbf{w}_r^H\} = \text{diag}(\sigma_{r1}^2, \sigma_{r2}^2, \dots, \sigma_{rK}^2)$. We consider the same approximation as (7) and define the SINR of user n at base station r averaged over the out-of-cluster interference as

$$\gamma_{rn} \approx \mathbf{h}_{rn}^H \mathbf{R}_{rn}^{-1} \mathbf{h}_{rn}, \quad (12)$$

where the covariance matrix \mathbf{R}_{rn} is defined as

$$\mathbf{R}_{rn} = \tilde{\mathbf{H}}_{rn} \tilde{\mathbf{H}}_{rn}^H + \Sigma_r, \quad (13)$$

with $\tilde{\mathbf{H}}_{rn} \tilde{\mathbf{H}}_{rn}^H = \sum_{k \neq n}^N \mathbf{h}_{rk} \mathbf{h}_{rk}^H$ containing the channels of the other $N - 1$ in-cluster users to the r -th base station and Σ_r containing the variances of the noise plus out-of-cluster interference at the r -th base station.

2) *Central Combining*: After performing local estimation, the R base stations transmit their estimates, \hat{s}_{rn} , to the central processor together with a combining weight, ω_{rn} . These combining weights are calculated at each base station in such a way that the final SINR of user n is maximized. This is equivalent to implementing global LMMSE on the local estimates at the central processor [26]. Note that an equal weight could be used to combine the local estimates at the central processor when the backhaul capacity is severely limited. However, the proposed combining weights will significantly outperform equal weight combining for most users in a cooperative cellular network where the SINR from a given user to each base station is largely non-identical. To derive the weighted matrices for the central processor, we substitute (9) into (10) and reexpress the estimate of user n at base station r as

$$\hat{s}_{rn} = \underbrace{\mathbf{v}_{rn}^H \mathbf{h}_{rn} s_n}_{\text{signal}} + \underbrace{\mathbf{v}_{rn}^H \left(\sum_{i \neq n}^N \mathbf{h}_{ri} s_i + \mathbf{w}_r \right)}_{\text{noise+interference}}, \quad (14)$$

where we have separately identified the signal part and the noise plus interference part of \hat{s}_{rn} . To maximize the SINR at the central processor, the combining weight from base station r is defined as [27, eq. (13.23)]

$$\omega_{rn}^* = \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^*}{\xi_{rn}}, \quad (15)$$

where $(\cdot)^*$ denotes the complex conjugate and ξ_{rn} is the noise plus interference power of user n at base station r , which can be derived using (14) as

$$\xi_{rn} = E \left\{ \mathbf{v}_{rn}^H \left(\sum_{i \neq n}^N \mathbf{h}_{ri} s_i + \mathbf{w}_r \right) \left(\sum_{i \neq n}^N \mathbf{h}_{ri} s_i + \mathbf{w}_r \right)^H \mathbf{v}_{rn} \right\} \quad (16)$$

$$= \mathbf{v}_{rn}^H \left(\sum_{i \neq n}^N \mathbf{h}_{ri} \mathbf{h}_{ri}^H + \Sigma_r \right) \mathbf{v}_{rn} \quad (17)$$

where $E\{\cdot\}$ in (16) is the expectation over the data symbols and noise, and (17) follows from the independence between them.

Combining all R estimates according to the weights in (15) results in the final estimate of the symbol from user n

$$\hat{s}_n = \sum_{r=1}^R \omega_{rn}^* \hat{s}_{rn}. \quad (18)$$

Note that the only information the central processor requires from each base station are the N estimates \hat{s}_{rn} of the user symbols and the corresponding weights ω_{rn} . Using (18), we derive an approximation for the SINR of user n at the central processor as a sum of the SINRs from the R base stations given by

$$\hat{\gamma}_n \approx \sum_{r=1}^R \gamma_{rn}, \quad (19)$$

where a detailed derivation of $\hat{\gamma}_n$ is provided in Appendix A.

III. CENTRALIZED PROCESSING

In this section, we derive the achievable rate of a given in-cluster user based on the centralized processing architecture. It should be noted that the rate of each user is different since we consider non-identical path gains to each base station. Considering a Gaussian approximation for the distribution of the interference plus noise in γ_n , the achievable rate of user n can be written as

$$C_n = \int_0^\infty \log_2(1 + \gamma_n) f_{\gamma_n}(\gamma_n) d\gamma_n, \quad (20)$$

where $f_{\gamma_n}(\gamma_n)$ is the PDF of γ_n . We note that an exact analysis of $f_{\gamma_n}(\gamma_n)$ is extremely tedious even for the simple two user scenario with no out-of-cluster interference [28]. As such, to obtain more insightful results, we consider an accurate approximation for $f_{\gamma_n}(\gamma_n)$ using a Laplace-type approximation for the CF of γ_n .

A. Characteristic Function of γ_n

Based on (7), the CF of γ_n is defined as [29]

$$\phi_{\gamma_n}(t) = E \left\{ e^{jt\gamma_n} \right\} = E \left\{ e^{jt\mathbf{h}_n^H \mathbf{R}_n^{-1} \mathbf{h}_n} \right\}, \quad (21)$$

where the expectation is over the channel gains in \mathbf{h}_n and \mathbf{R}_n . First, we consider the CF conditioned on \mathbf{R}_n and derive the expectation in (21) over the distribution of \mathbf{h}_n . As the channels from user n to each antenna are independent, the joint PDF of the channel gains in $\mathbf{h}_n = [h_{1kn}, \dots, h_{RKn}]^T$ can be written as a product of $R \times K$ individual PDFs given by

$$\begin{aligned} f_{\mathbf{h}_n}(\mathbf{h}_n) &= \prod_{r=1}^R \prod_{k=1}^K \frac{1}{\pi g_{rkn}^2} e^{-\frac{|h_{rkn}|^2}{g_{rkn}^2}} \\ &= \frac{1}{\pi^{RK} |\mathbf{G}_n|} e^{-\mathbf{h}_n^H \mathbf{G}_n^{-1} \mathbf{h}_n}, \end{aligned} \quad (22)$$

where $\mathbf{G}_n = E \{ \mathbf{h}_n \mathbf{h}_n^H \} = \text{diag} (g_{11n}^2, g_{12n}^2, \dots, g_{RKn}^2)$, is the covariance matrix of \mathbf{h}_n and $|\cdot|$ denotes the determinant of a matrix. Substituting (22) into (21) we can write the CF conditioned on \mathbf{R}_n as given in (23), shown at the bottom of the page, where $dh_{rkn} = dh_{rknI} dh_{rknQ}$ with $h_{rknI} = \text{Re}(h_{rkn})$ and $h_{rknQ} = \text{Im}(h_{rkn})$. The multi-fold integral in (23) can be solved using the identity in [28, Lemma 02] which results in the CF conditioned on \mathbf{R}_n as

$$\phi_{\gamma_n}(t | \mathbf{R}_n) = \frac{1}{|\mathbf{I} - jt\mathbf{R}_n^{-1} \mathbf{G}_n|}. \quad (24)$$

Substituting $\mathbf{R}_n = \bar{\mathbf{H}}_n \bar{\mathbf{H}}_n^H + \Sigma$, the conditional CF can be reexpressed as

$$\phi_{\gamma_n}(t | \bar{\mathbf{H}}_n) = \frac{|\bar{\mathbf{H}}_n \bar{\mathbf{H}}_n^H + \Sigma|}{|\bar{\mathbf{H}}_n \bar{\mathbf{H}}_n^H + \Sigma - jt\mathbf{G}_n|}, \quad (25)$$

$$\phi_{\gamma_n}(t | \mathbf{R}_n) = \frac{1}{\pi^{RK} |\mathbf{G}_n|} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\mathbf{h}_n^H (\mathbf{G}_n^{-1} - jt\mathbf{R}_n^{-1}) \mathbf{h}_n} dh_{1kn} dh_{2kn} \dots dh_{RKn}, \quad (23)$$

which is now conditioned on the channel gains in $\bar{\mathbf{H}}_n$. The CF averaged over $\bar{\mathbf{H}}_n$ can be written as an expectation

$$\phi_{\gamma_n}(t) = E \left\{ \frac{|\mathbf{I} + \Sigma^{-1} \bar{\mathbf{H}}_n \bar{\mathbf{H}}_n^H|}{|\mathbf{I} - jt \Sigma^{-1} \mathbf{G}_n + \Sigma^{-1} \bar{\mathbf{H}}_n \bar{\mathbf{H}}_n^H|} \right\}, \quad (26)$$

where we have multiplied both numerator and denominator by Σ^{-1} . Using the determinant identity $|\mathbf{I} + \chi \chi^H| = |\mathbf{I} + \chi^H \chi|$ in the numerator and the denominator of (26) we get

$$\phi_{\gamma_n}(t) \approx \frac{1}{|\mathbf{X}|} E \left\{ \frac{|\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n|}{|\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \mathbf{X}^{-1} \bar{\mathbf{H}}_n|} \right\}, \quad (27)$$

where $\mathbf{X} = \mathbf{I} - jt \Sigma^{-1} \mathbf{G}_n$. From [28], we note that an exact analysis of (27) is extremely cumbersome even for a simple system model with 2 users where $\bar{\mathbf{H}}_n$ is a vector and no out-of-cluster interference. As the exact analysis of (27) is intractable, a more useful approach is to consider a Laplace-type approximation as in [30] which results in

$$\phi_{\gamma_n}(t) \approx \frac{E \{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n| \}}{|\mathbf{X}| E \{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \mathbf{X}^{-1} \bar{\mathbf{H}}_n| \}}. \quad (28)$$

This approximation, which is better known for scalar quadratic forms [30], has some motivation in the work of [7], [25], [31]². In Appendix B, we illustrate that the identity in both the numerator and the denominator of (27) can be expressed as a limit of a Wishart matrix, which results in a ratio of determinants of matrix quadratic forms. This ratio can then be decomposed to give a product of scalar quadratic forms as in [31], motivating the Laplace-type approximation given in (28). This approximation allows us to use a similar approach developed in [7] to derive the expectations for the determinants in (28). However, it is important to note that the structure of the numerator and the denominator of (28) are different from [7] as we consider the effect of out-of-cluster interference.

Let us first consider the numerator of (28). Let $\lambda_1, \lambda_2, \dots, \lambda_\vartheta$ be the set of ordered non zero eigenvalues of $\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n$ where $\vartheta = \min(RK, N-1)$. As such, the determinant in the numerator can be written as a product of the eigenvalues given by

$$E \left\{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n| \right\} = E \left\{ \prod_{i=1}^{\vartheta} (1 + \lambda_i) \right\}. \quad (29)$$

Using [32, eq. 1.2.9] and [32, eq. 1.2.12], we can reexpress the expectation as

$$E \left\{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n| \right\} = E \left\{ \sum_{i=0}^{\vartheta} \mathcal{E}_i(\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n) \right\}, \quad (30)$$

$$= \sum_{i=0}^{\vartheta} \sum_{\delta} E \left\{ \left| (\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n)_{\delta_{i,N-1}} \right| \right\}, \quad (31)$$

²Note that the extended Laplace-type approximation in [31] does not consider the effect of out-of-cluster interference. Hence the structure of the numerator and the denominator of [31, eq. 40] are different. However, we note that their argument directly applies to the case with out-of-cluster interference as both the numerator and the denominator in (28) contain the identity.

where $\mathcal{E}_i(A)$ in (30) is the i -th elementary symmetric function of a matrix A , which is further expressed in (31) as a sum of the determinants of all $i \times i$ principal minors of A [32], where $\delta_{i,N-1}$ is a length i subset of $\{1, 2, \dots, N-1\}$ and the sum in (31) is over all such subsets³.

Clearly, the term $E \left\{ \left| (\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n)_{\delta_{i,N-1}} \right| \right\}$ is now the building block of the overall expectation in (31). According to [7, Corollary 1], when the total number of measurements at central processor is larger than the number of interfering users, i.e., $RK \geq N-1$ we can write

$$E \left\{ \left| (\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n) \right| \right\} = \text{Perm}(\Sigma^{-1} \Lambda_n), \quad (32)$$

where $\Lambda_n = E \{ \bar{\mathbf{H}}_n \circ \bar{\mathbf{H}}_n^\dagger \}$ with \circ denoting the Hadamard product⁴, [33], \dagger denoting the element-wise conjugate, and $\text{Perm}(\cdot)$ denoting the permanent of a matrix [34]. When the number of measurements at each base station is smaller than the number of interfering users, i.e., $RK < N-1$, some eigenvalues of $\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n$ will be zero, thus $E \left\{ \left| (\bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n) \right| \right\}$ is zero when $i > RK$. As such, the expression in (31) can be simplified to,

$$E \left\{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n| \right\} = \sum_{i=0}^{\vartheta} \sum_{\delta} \text{Perm}((\Sigma^{-1} \Lambda_n)_{\delta_{i,N-1}}). \quad (33)$$

The expectation in (33) can be evaluated in closed-form due to the fact that Λ_n is an $RK \times N-1$ matrix with entries consisting of the path gains of the $N-1$ in-cluster users in $\bar{\mathbf{H}}_n$.

Similarly, we expand the denominator of (28) as

$$|\mathbf{X}| E \left\{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \mathbf{X}^{-1} \bar{\mathbf{H}}_n| \right\} = |\mathbf{X}| \sum_{i=0}^{\vartheta} \sum_{\delta} \text{Perm}((\Sigma^{-1} \mathbf{X}^{-1} \Lambda_n)_{\delta_{i,N-1}}) \quad (34)$$

$$= |\mathbf{X}| \sum_{i=0}^{\vartheta} \sum_{\delta} \frac{\text{Perm}((\Sigma^{-1} \Lambda_n)_{\delta_{i,RK}}^{N-1})}{|\mathbf{X}_{\delta_{i,RK}}|} \quad (35)$$

where the permanent in (34) is reexpressed as (35) using [7, Corollary 1]. Taking $|\mathbf{X}|$ inside the summation we can reexpress (35) as

$$|\mathbf{X}| E \left\{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \mathbf{X}^{-1} \bar{\mathbf{H}}_n| \right\} = \sum_{i=0}^{\vartheta} \sum_{\delta} |\mathbf{X}_{\bar{\delta}_{RK-i,RK}}| \text{Perm}((\Sigma^{-1} \Lambda_n)_{\delta_{i,RK}}^{N-1}), \quad (36)$$

where $\bar{\delta}_{RK-i,RK}$ is a length $RK-i$ subset of $\{1, 2, \dots, RK\}$ that is not in $\delta_{i,RK}$. Now, given $\mathbf{X} = \mathbf{I} - jt \Sigma^{-1} \mathbf{G}_n$, (36) can be written in terms of the elementary symmetric function as

$$|\mathbf{X}| E \left\{ |\mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \mathbf{X}^{-1} \bar{\mathbf{H}}_n| \right\} = \sum_{i=0}^{\vartheta} \sum_{\delta} \sum_{l=0}^{RK-i} (-jt)^l \mathcal{E}_l((\Sigma^{-1} \mathbf{G}_n)_{\bar{\delta}_{RK-i,RK}}) \times \text{Perm}((\Sigma^{-1} \Lambda_n)_{\delta_{i,RK}}^{N-1}) \quad (37)$$

³We use A_{ν}^{μ} to denote the submatrix formed by taking the rows and columns of A indexed by the sets ν and μ , respectively. If either ν or μ contain the complete set, the corresponding subscript/superscript is omitted.

⁴Hadamard product is an element-wise product of two matrices of the same dimension where each element (i, j) is the product of the (i, j) -th elements in the original two matrices.

$$= \sum_{l=0}^{rK} (-jt)^l \sum_{i=0}^{\vartheta} \sum_{\delta} \mathcal{E}_l((\mathbf{\Sigma}^{-1} \mathbf{G}_n)_{\delta_{RK-i, RK}}) \times \text{Perm}((\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)_{\delta_{i, RK}}^{N-1}), \quad (38)$$

where (38) follows from the fact that the l -th elementary symmetric function $\mathcal{E}_l((\mathbf{\Sigma}^{-1} \mathbf{G}_n)_{\delta_{RK-i, RK}}) = 0$ for $l > RK - i$.

Substituting (33) and (38) into (28), our approximation of the CF is given by

$$\phi_{\gamma_n}(t) \approx \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\sum_{l=0}^{RK} (-jt)^l \Upsilon_l}, \quad (39)$$

where

$$\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n) = \sum_{i=0}^{\vartheta} \sum_{\delta} \text{Perm}((\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)_{\delta_{i, N-1}}), \quad (40)$$

and

$$\Upsilon_l = \sum_{i=0}^{\vartheta} \sum_{\delta} \mathcal{E}_l((\mathbf{\Sigma}^{-1} \mathbf{G}_n)_{\delta_{RK-i, RK}}) \text{Perm}((\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)_{\delta_{i, RK}}^{N-1}). \quad (41)$$

Finally, noting that $\sum_{l=0}^{RK} (-jt)^l \Upsilon_l$ is a degree RK polynomial in t , we can further simplify (39) as

$$\phi_{\gamma_n}(t) \approx \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\Upsilon_{RK} \sum_{l=0}^{RK} (-jt)^l \frac{\Upsilon_l}{\Upsilon_{RK}}} \quad (42)$$

$$= \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\Upsilon_{RK} \prod_{l=1}^{RK} (\varepsilon_l - jt)}, \quad (43)$$

where ε_l are the roots of the denominator polynomial in (42), which can be computed using a standard root finding program. Using the CF expression in (43) we can derive important functions such as the PDF and the CDF of γ_n . However, we only derive the PDF expression in the following, as the focus of this chapter is on the derivation of the achievable rate.

B. Probability Density Function of γ_n

The PDF of γ_n can be derived from the CF as [29]

$$f_{\gamma_n}(\gamma_n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{\gamma_n}(t) e^{-jt\gamma_n} dt. \quad (44)$$

We proceed by reexpressing the CF in (43) to allow easy integration as

$$\phi_{\gamma_n}(t) \approx \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\Upsilon_{RK}} \sum_{l=1}^{RK} \frac{\rho_l}{(\varepsilon_l - jt)}, \quad (45)$$

where $\rho_l = \frac{1}{\prod_{j \neq l}^{RK} (\varepsilon_j - \varepsilon_l)}$. Substituting (45) into (44), results in an approximate expression for the PDF of γ_n given by

$$f_{\gamma_n}(\gamma_n) \approx \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{2\pi \Upsilon_{RK}} \int_{-\infty}^{\infty} \sum_{l=1}^{RK} \frac{\rho_l e^{-jt\gamma_n}}{(\varepsilon_l - jt)} dt. \quad (46)$$

The integral in (46) can be solved using the identity in [35, eq. 7.3.382] and the final expression for the approximate PDF can be derived as

$$f_{\gamma_n}(\gamma_n) \approx \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\Upsilon_{RK}} \sum_{l=1}^{RK} \rho_l e^{-\varepsilon_l \gamma_n}. \quad (47)$$

C. Achievable Rate Approximation

We derive an approximation on the achievable rate of user n with centralized processing by substituting (47) into (20) which results in

$$C_n \approx \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\Upsilon_{RK}} \sum_{l=1}^{RK} \rho_l \int_0^{\infty} \log_2(1 + \gamma_n) e^{-\varepsilon_l \gamma_n} d\gamma_n \quad (48)$$

$$= \frac{\Psi(\mathbf{\Sigma}^{-1} \mathbf{\Lambda}_n)}{\Upsilon_{RK} \ln 2} \sum_{l=1}^{RK} -\frac{\rho_l e^{\varepsilon_l} E_i(-\varepsilon_l)}{\varepsilon_l}, \quad (49)$$

where (49) is derived using the identity [35, eq. 2, 4.337] and E_i denotes the exponential integral function. Our new results in (49) provides an accurate approximation of the exact achievable rate of an arbitrary in-cluster user with centralized processing, which will be illustrated via numerical examples in Section V.

IV. PARTIALLY DECENTRALIZED PROCESSING

In this section, we derive an expression for the achievable rate of a given in-cluster user based on the partially decentralized processing architecture. Considering a Gaussian approximation for the distribution of the interference plus noise in $\hat{\gamma}_n$, the achievable rate of user n can be written as

$$C_n = \int_0^{\infty} \log_2(1 + \hat{\gamma}_n) f_{\hat{\gamma}_n}(\hat{\gamma}_n) d\hat{\gamma}_n, \quad (50)$$

where $f_{\hat{\gamma}_n}(\hat{\gamma}_n)$ is the PDF of $\hat{\gamma}_n$. Similar to Section III we consider an accurate approximation for $f_{\hat{\gamma}_n}(\hat{\gamma}_n)$ using a Laplace-type approximation for the CF of $\hat{\gamma}_n$.

A. Characteristic Function of $\hat{\gamma}_n$

As $\hat{\gamma}_n$ is the sum of R independent non-identically distributed SINRs from the base stations, the CF of $\hat{\gamma}_n$ can be defined as [29]

$$\phi_{\hat{\gamma}_n}(t) = E \left\{ e^{jt\hat{\gamma}_n} \right\} \approx \prod_{r=1}^R E \left\{ e^{jt\gamma_{rn}} \right\}, \quad (51)$$

where $E\{\cdot\}$ denotes the expectation taken over the channel gains. Thus the CF of $\hat{\gamma}_n$ is the product of R individual CFs, $\phi_{\gamma_{rn}}(t) \approx E \left\{ e^{jt\mathbf{h}_{rn}^H \mathbf{R}_{rn}^{-1} \mathbf{h}_{rn}} \right\}$.

Given that local estimation is performed at each base station with the use of only K received signals, an approximation of the CF at base station r can be derived by setting $R = 1$ in the centralized processing architecture. As such, using (28) we approximate $\phi_{\gamma_{rn}}(t)$ as

$$\phi_{\gamma_{rn}}(t) \approx \frac{E \left\{ \left| \mathbf{I} + \bar{\mathbf{H}}_{rn}^H \mathbf{\Sigma}_r^{-1} \bar{\mathbf{H}}_{rn} \right| \right\}}{|\mathbf{X}_{rn}| E \left\{ \left| \mathbf{I} + \bar{\mathbf{H}}_{rn}^H \mathbf{\Sigma}_r^{-1} \mathbf{X}_{rn}^{-1} \bar{\mathbf{H}}_{rn} \right| \right\}}. \quad (52)$$

where $\mathbf{X}_{rn} = \mathbf{I} - jt\boldsymbol{\Sigma}_r^{-1}\mathbf{G}_{rn}$. Following the same steps as with the centralized processing architecture, we derive expressions for expectations in the numerator and the denominator in (52). Let $\lambda_1, \lambda_2, \dots, \lambda_\varphi$ be the set of ordered non zero eigenvalues of $\bar{\mathbf{H}}_{rn}^H \boldsymbol{\Sigma}_r^{-1} \bar{\mathbf{H}}_{rn}$ where $\varphi = \min(K, N - 1)$. Similar to the derivation in (33) the expectation in the numerator of (52) can be expressed as

$$E \left\{ \left| \mathbf{I} + \bar{\mathbf{H}}_{rn}^H \boldsymbol{\Sigma}_r^{-1} \bar{\mathbf{H}}_{rn} \right| \right\} = \sum_{i=0}^{\varphi} \sum_{\delta} \text{Perm}((\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})^{\delta_{i,N-1}}). \quad (53)$$

Likewise based on (38), the denominator of (52) is derived as

$$\begin{aligned} |\mathbf{X}_{mn}| E \left\{ \left| \mathbf{I} + \bar{\mathbf{H}}_{rn}^H \boldsymbol{\Sigma}_r^{-1} \mathbf{X}_{rn}^{-1} \bar{\mathbf{H}}_{rn} \right| \right\} \\ = \sum_{l=0}^K (-jt)^l \sum_{i=0}^{\varphi} \sum_{\delta} \mathcal{E}_l((\boldsymbol{\Sigma}_r^{-1} \mathbf{G}_{rn})_{\delta_{K-i,K}}) \\ \times \text{Perm}((\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})_{\delta_{i,K}}^{N-1}). \end{aligned} \quad (54)$$

Based on (53) and (54), the approximate CF of γ_{rn} is given by

$$\phi_{\gamma_{rn}}(t) \approx \frac{\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{\sum_{l=0}^K (-jt)^l \Upsilon_{rl}}, \quad (55)$$

where

$$\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn}) = \sum_{i=0}^{\varphi} \sum_{\delta} \text{Perm}((\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})^{\delta_{i,N-1}}) \quad (56)$$

and

$$\Upsilon_{rl} = \sum_{i=0}^{\varphi} \sum_{\delta} \mathcal{E}_l((\boldsymbol{\Sigma}_r^{-1} \mathbf{G}_{rn})_{\delta_{K-i,K}}) \text{Perm}((\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})_{\delta_{i,K}}^{N-1}). \quad (57)$$

Substituting (55) into (51) we derive the approximate CF of $\hat{\gamma}_n$ as

$$\phi_{\hat{\gamma}_n}(t) \approx \prod_{r=1}^R \frac{\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn}) / \Upsilon_{rK}}{\sum_{l=0}^K (-jt)^l \Upsilon_{rl} / \Upsilon_{rK}}. \quad (58)$$

To further simplify the CF expression, we note that $\sum_{l=0}^K (-jt)^l \Upsilon_{rl} / \Upsilon_{rK}$ is a degree K polynomial in t which means that the denominator of (58) consists of a multiplication of R such polynomials. As such, we can re-arrange the denominator of (58) into a single polynomial of degree RK as

$$\phi_{\hat{\gamma}_n}(t) \approx \frac{\prod_{r=1}^R \Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{\Upsilon_{rK} \sum_{q=0}^{RK} \zeta_q (-jt)^q}, \quad (59)$$

where ζ_q corresponds to the sum of all polynomial terms that has degree q defined as

$$\zeta_q = \sum_{x_i \in \chi_q} \prod_{r=1}^R \frac{\Upsilon_{rx(r)}}{\Upsilon_{rK}}. \quad (60)$$

In (60), the first summation of χ_q is over the set of sets of R elements chosen from $\{0, 1, \dots, K\}$ such that $\sum_{r=1}^R x_i(r) = q$

where x_i is the i -th set of χ_q . Based on (59), the CF of $\phi_{\hat{\gamma}_n}$ can be further simplified as

$$\phi_{\hat{\gamma}_n}(t) \approx \frac{\prod_{r=1}^R \Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{\Upsilon_{rK} \prod_{q=1}^{RK} (\varepsilon_q - jt)}. \quad (61)$$

where ε_q are the roots of the denominator polynomial in (59), which can be computed using a standard root finding program.

B. Probability Density Function of $\hat{\gamma}_n$

Based on our CF in (61), we can directly derive an approximate expression for the PDF of $\hat{\gamma}_n$ as

$$\begin{aligned} f_{\hat{\gamma}_n}(\hat{\gamma}_n) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{\hat{\gamma}_n}(t) e^{-jt\hat{\gamma}_n} dt \\ &\approx \left[\prod_{r=1}^R \frac{\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{2\pi \Upsilon_{rK}} \right] \int_{-\infty}^{\infty} \sum_{q=1}^{RK} \frac{\rho_q e^{-jt\hat{\gamma}_n}}{(\varepsilon_q - jt)} dt, \end{aligned} \quad (62)$$

where we have reexpressed the CF for easy integration with

$$\rho_q = \frac{1}{\prod_{j \neq q}^{RK} (\varepsilon_j - \varepsilon_q)}. \quad (63)$$

Similar to Section III-B, we solve the integral in (62) using the identity in [35, eq. 7.3.382] which results in

$$f_{\hat{\gamma}_n}(\hat{\gamma}_n) \approx \left[\prod_{r=1}^R \frac{\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{\Upsilon_{rK}} \right] \sum_{q=1}^{RK} \rho_q e^{-\varepsilon_q \hat{\gamma}_n}. \quad (64)$$

C. Achievable Rate Approximation

Finally, we present our new approximation for the achievable rate of user n with partially decentralized processing as

$$C_n \approx \left[\prod_{r=1}^R \frac{\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{\Upsilon_{rK}} \right] \sum_{q=1}^{RK} \rho_q \int_0^{\infty} \log_2(1 + \hat{\gamma}_n) e^{-\varepsilon_q \hat{\gamma}_n} d\hat{\gamma}_n \quad (65)$$

$$= \left[\prod_{r=1}^R \frac{\Psi(\boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Lambda}_{rn})}{\Upsilon_{rK} \ln 2} \right] \sum_{q=1}^{RK} \frac{\rho_q e^{\varepsilon_q} \text{Ei}(-\varepsilon_q)}{\varepsilon_q}, \quad (66)$$

which is derived by substituting (64) into (50) and using the identity [35, eq. 2, 4.337] to solve the resulting integral. Our new results in (66) provides an accurate approximation of the exact achievable rate of an arbitrary in-cluster user with partially decentralized processing, which will be illustrated via numerical examples in Section V.

V. NUMERICAL EXAMPLES

In this section, we present numerical examples to highlight the achievable rate of centralized and partially decentralized processing with out-of-cluster interference under different network scenarios. We consider a network model with a cluster of three base stations (BS_1 - BS_3) where the base stations are equipped with three antennas each. We randomly place

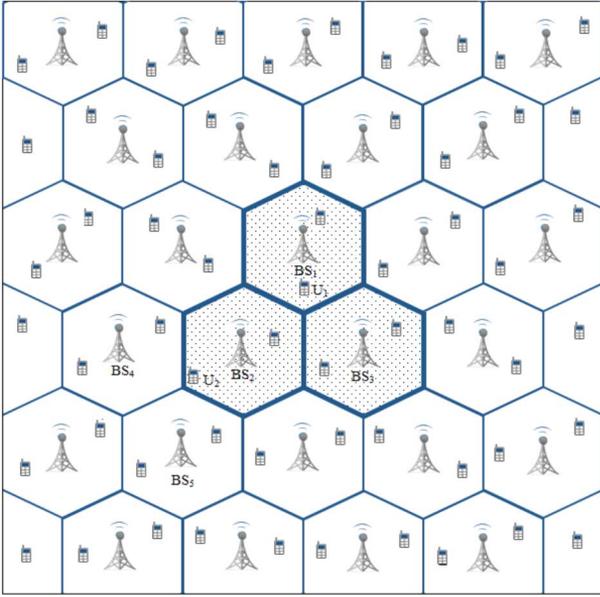


Fig. 2. Network model with a cluster of 3 cooperating base stations (BS1-BS3) equipped with 3 antennas each, receiving signals from users randomly located with 2 active users in each cell.

two active users in each cell as illustrated in Fig. 2. We define the path gain between user n and antenna k of base station r as

$$g_{rkn}^2 = \frac{1}{L(d_0)} \left(\frac{d_{rkn}}{d_0} \right)^{-\alpha}, \quad (67)$$

where d_{rkn} is the distance between user n and antenna k of base station r , $L(d_0)$ is the path loss measured at a reference distance d_0 , and α is the path loss exponent which is set to 4 unless otherwise stated. In the examples, we set the reference path loss as $L(d_0) = 100$ dB at $d_0 = 100$ m, and the cell radius is 120 m. As stated in Section II, the Rayleigh fading channel gains are complex Gaussian with zero mean and unit variance, and the additive noise is complex Gaussian with zero mean and variance σ^2 .

Fig. 3 plots the achievable rate of in-cluster user U_1 , shown in Fig. 2, versus the received SNR for centralized processing with $N = 6$ in-cluster users. The simulation points are obtained using Monte Carlo simulation with channel fading and noise components for each simulation trial drawn from an independent complex Gaussian distribution. The figure shows that our approximate achievable rate results, generated using (49), accurately predicts the exact simulation points, generated using (6), throughout the full range of SNRs. We compare the following three scenarios: 1) Achievable rate with out-of-cluster interference where $\tilde{N} = 18$ users in the first ring of 9 interfering cells outside the cooperating cluster, 2) Achievable rate without out-of-cluster interference where $\tilde{N} = 0$, and 3) Achievable rate with no cooperation where $\tilde{N} = 22$ and no backhaul processor is used. The plots with cooperation clearly shows that out-of-cluster interference gives rise to a saturation regime where further increase in SNR after a certain threshold value, does not noticeably improve the achievable rate. The saturation level of the achievable rate can be accurately approximated by setting

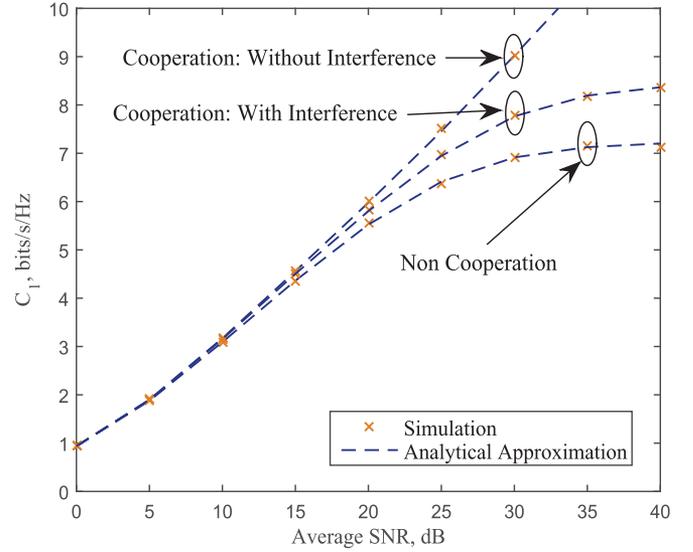


Fig. 3. Achievable rate of U_1 versus the received SNR with the centralized processing architecture.

$\sigma^2 = 0$ in (49) which results in 8.4 bits/s/Hz for this example scenario. The figure clearly illustrates that the achievable rate of a cooperating cluster within a large network is drastically impacted by out-of-cluster interference in the high SNR regime. The plot also highlights the performance gain obtained by cooperation with centralized processing when compared to a traditional non cooperative network. This gain can be increased further by increasing the cluster size. However, the increase in the performance gain comes at a cost of added backhaul complexity in real-life implementation.

Fig. 4 plots the achievable rate of the in-cluster user U_1 for partially decentralized processing with the same interference scenarios considered in Fig. 3. We see that our approximate achievable rate results, generated using (66), accurately predicts the exact simulation points throughout the full range of SNRs. In contrast to Fig. 3, we observe that the achievable rate of partially decentralized processing saturates at high SNRs in all three scenarios including the scenario without out-of-cluster interference. This is due to the fact that the number of antennas at each base station is less than the total number of in-cluster users. As local estimation in partially decentralized processing only utilizes the signals received at a certain base station the performance at high SNRs is limited by the in-cluster interference. Thus, the partially decentralized processing architecture is more suitable for smaller cluster sizes. The plot also highlights the performance gain obtained by cooperation with partially decentralized processing when compared to a traditional non cooperative network.

The plots in Fig. 3 and Fig. 4 were generated for a fixed set of user locations illustrated in Fig. 2, but in a cellular environment users can be scattered in different locations throughout the network. Therefore in Fig. 5, we compare the achievable rate of U_1 for centralized and partially decentralized processing averaged over different user locations. To do so, we randomly place the users throughout the network in Fig. 2 and numerically average the achievable rate of U_1 over the corresponding path gains

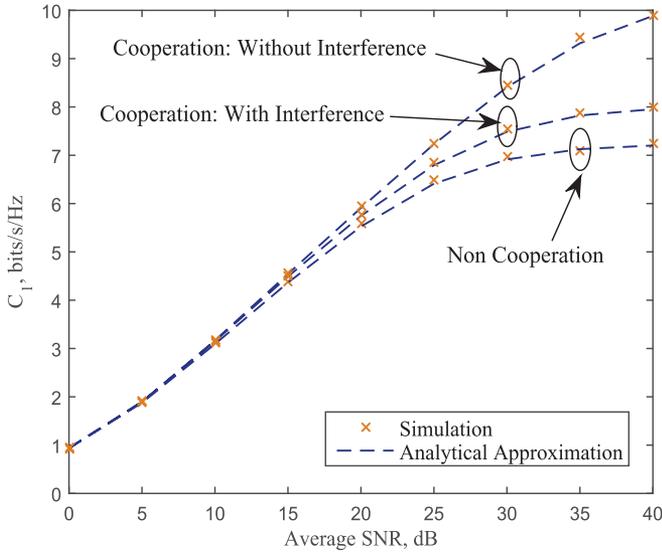


Fig. 4. Achievable rate of U_1 versus the received SNR with the partially decentralized processing architecture.

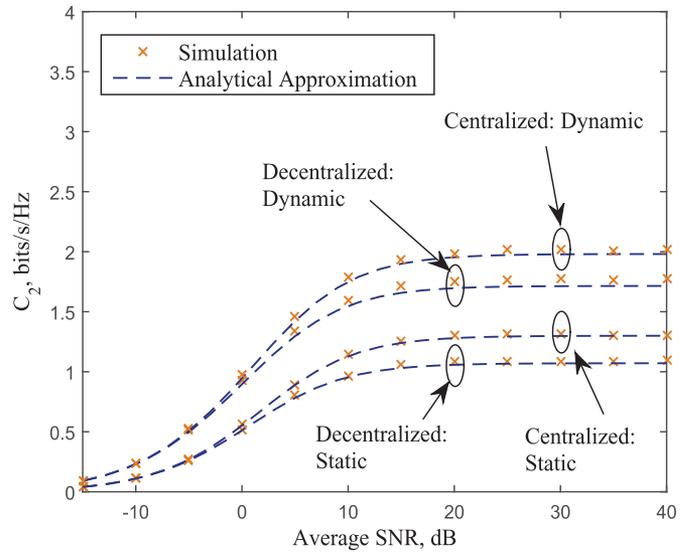


Fig. 6. Achievable rate of U_2 versus the received SNR with the static and dynamic cluster arrangements.

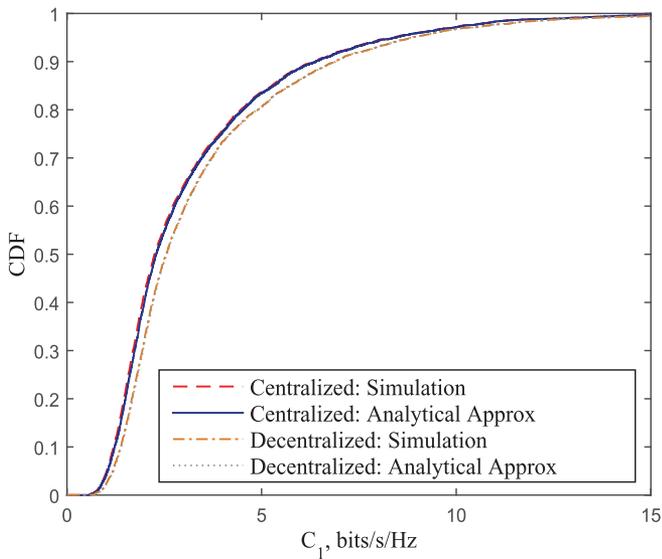


Fig. 5. CDF of the achievable rate of U_1 with $N = 6$ in-cluster users and $\tilde{N} = 9$ out-of-cluster interfering users.

with $N = 6$ in-cluster users and $\tilde{N} = 18$ out-of-cluster interfering users. We assume a uniform distribution of two users per cell and set the transmit power of all the users such that the average received SNR of U_1 is 30 dB. We observe in the figure that the performance of partially decentralized processing is very close to that of centralized processing when averaged over the user locations. We also observe that the accuracy of our analysis does not depend on specific user locations. We note that for a fixed cooperating cluster centralized processing performs better than partially decentralized processing. This is expected, as centralized processing uses $R \times K$ received signals at the central processor to achieve a better estimation whereas partially decentralized processing uses K received signals at each base station.

In Fig. 6, we consider different cluster configurations for user U_2 , located at the cluster edge as shown in Fig. 2. To do so, we compare static clustering with fixed base station clusters, with dynamic clustering where the base station clusters are selected based on the nearest user locations [2]. We note that the cooperating cluster of BS_1 - BS_3 (which we refer to as static clustering) results in user U_2 being located at the cluster edge. Applying dynamic clustering results in a cooperating cluster for U_2 comprising the three closest base stations BS_2 , BS_4 and BS_5 . As illustrated in Fig. 6, partially decentralized processing with static clustering has the worst performance. However, it has the simplest cluster implementation with low feedback overheads to the central processor and fixed base station clusters. As expected, centralized processing with dynamic clustering has the best performance but requires the highest amount of coordination and feedback overheads. Interestingly, we highlight that partially decentralized processing with dynamic clustering can outperform centralized processing with static clustering which indicates an attractive balance between the increased coordination complexity of dynamic clustering and the lower feedback overheads of partially decentralized processing.

Finally, in Figs. 7 and 8 we present the approximate achievable rate for different path loss exponents to further highlight the effect of interference in clustered cooperation.

Fig. 7 plots the achievable rate of U_1 with centralized processing versus the average received SNR for $\alpha = 2, 3$ and 4. As expected, the achievable rate without out-of-cluster interference decreases with increasing α . This is due to the fact that the received signals of the in-cluster users are weaker as α increases. In contrast, at high SNRs we find that the achievable rate with out-of-cluster interference improves with increasing path loss exponent even though the signals from both in-cluster and out-of-clusters users are weaker as α increases. This is due to the fact that increasing α weakens the out-of-cluster interference causing the saturation to occur at a higher SNR. This

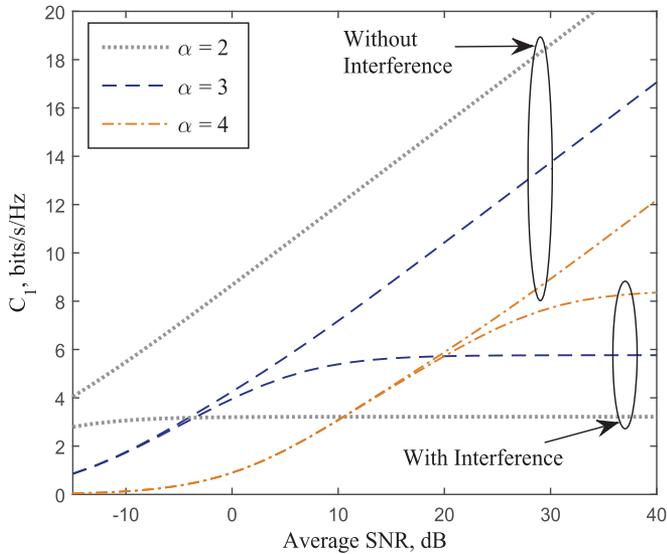


Fig. 7. Achievable rate of U_1 versus the received SNR with the centralized processing architecture for different path loss exponents.

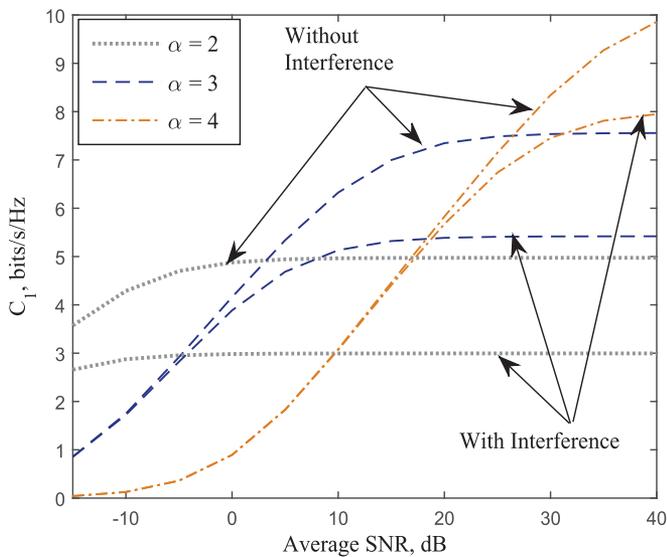


Fig. 8. Achievable rate of U_1 versus the received SNR with the partially decentralized processing architecture for different path loss exponents.

further illustrates the interference limited behavior of clustered cooperation where the out-of-cluster interference dominates the performance at high SNRs.

Fig. 8 plots the achievable rate of U_1 with partially decentralized processing versus the average received SNR for $\alpha = 2, 3$ and 4 . We observe that at high SNRs, the achievable rate improves with increasing path loss exponent for both cases of with and without out-of-cluster interference. This is due to the fact that the partially decentralized processing architecture is still limited by the in-cluster interference, as discussed in Fig. 4, even when there is no out-of-cluster interference. Thus, increasing α results in a higher achievable rate due to lower in-cluster interference.

VI. CONCLUDING REMARKS

We analyzed two different cooperative processing architectures, namely centralized processing and partially decentralized processing, for a cluster based cellular network. New approximations on the achievable rate were derived for a general multi-cell network with an arbitrary number of users transmitting to arbitrary number of cooperating base stations subject to independent Rayleigh fading channels and distance dependent path loss.

Our results provide an accurate approximation of the achievable rate for both centralized and partially decentralized processing architectures. In centralized processing, all the received signals at the $R \times K$ antennas are sent to the central processor to jointly estimate the user signals. Given that all the signal information is available at the central processor, centralized processing generally outperforms partially decentralized processing. However, the performance advantage of centralized processing comes at the cost of increased computational complexity relative to the partially decentralized processing architecture. Specifically, performing centralized LMMSE estimation involves the inversion of a large matrix of size $RK \times RK$, which is an order $(RK)^3$ operation [36]. Furthermore, the feedback overhead of obtaining all the received signals and channel gains at the central processor in large cooperative networks is daunting. In partially decentralized processing, the computational load is distributed amongst the R base stations where each base station processes K received signals. As such, the local LMMSE estimation performed at the base stations only requires the inversion of a smaller matrix of size $K \times K$, which is an order K^3 operation. Furthermore, the feedback overhead is significantly lower compared with centralized processing as individual channel gains of all the users are not required at the central processor. The only required data at the central processor are the local estimates and their corresponding weights for optimum combining.

We highlighted the dramatic impact of out-of-cluster interference on the achievable rate of both centralized and partially decentralized processing. We showed that the out-of-cluster interference gives rise to a saturation regime at high SNRs such that further increase in the SNR after a certain threshold does not noticeably improve the achievable rate. We noted that the performance difference between the two architectures is small with centralized processing performing slightly better than partially decentralized processing. We also revealed that dynamically selecting the cooperating cluster such that the desired user is in the cluster center allows partially decentralized processing to outperform centralized processing with static clustering. Furthermore, we found that increasing the path loss actually results in an improved achievable rate when the base station cluster is interference limited.

An interesting extension to this work is to consider cooperative clusters with limited feedback and propagation delays to the central processor. It would also be interesting to analyze the performance of these two processing architectures in other fading environments such as Rician. Determining how the achievable rate with centralized and partially decentralized processing behaves in heterogeneous networks with different cell

sizes is also desirable, as these are likely to predominate in 5G wireless networks and beyond.

APPENDIX A PROOF OF EQUATION (19)

Substituting (14) and (15) into (18) results in

$$\hat{s}_n = \underbrace{\sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^* \mathbf{v}_{rn}^H \mathbf{h}_{rn}}{\xi_{rn}} s_n}_{\text{signal}} + \underbrace{\sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^*}{\xi_{rn}} \mathbf{v}_{rn}^H \left(\sum_{i \neq n}^N \mathbf{h}_{ri} s_i + \mathbf{w}_r \right)}_{\text{noise+interference}}, \quad (68)$$

where we identify the signal part and the total noise plus interference part in the estimate \hat{s}_n with

$$\xi_{rn} = \mathbf{v}_{rn}^H \left(\sum_{i \neq n}^N \mathbf{h}_{ri} \mathbf{h}_{ri}^H + \boldsymbol{\Sigma}_r \right) \mathbf{v}_{rn} \quad (69)$$

$$\hat{\gamma}_n = \frac{\left(\sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^* (\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}} \right) \left(\sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^* (\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}} \right)^H}{\sum_{r=1}^R \left(\frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^*}{\xi_{rn}} \mathbf{v}_{rn}^H \left(\sum_{i \neq n}^{N+\tilde{N}} \mathbf{h}_{ri} \mathbf{h}_{ri}^H + \sigma^2 \mathbf{I} \right) \mathbf{v}_{rn} \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}} \right)}. \quad (70)$$

We proceed to derive the SINR of \hat{s}_n as the ratio of the signal power and total noise plus interference power as given in (70). Similar to (7), we approximate the received SINR of the partially decentralized processing architecture as

$$\begin{aligned} \hat{\gamma}_n &\approx \frac{\left(\sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^* (\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}} \right) \left(\sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^* (\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}} \right)^H}{\sum_{r=1}^R \left(\frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^*}{\xi_{rn}} \mathbf{v}_{rn}^H \left(\sum_{i \neq n}^N \mathbf{h}_{ri} \mathbf{h}_{ri}^H + \boldsymbol{\Sigma}_r \right) \mathbf{v}_{rn} \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}} \right)} \\ &= \sum_{r=1}^R \frac{(\mathbf{v}_{rn}^H \mathbf{h}_{rn})^* (\mathbf{v}_{rn}^H \mathbf{h}_{rn})}{\xi_{rn}}, \end{aligned} \quad (71)$$

where the instantaneous channel gains of the out-of-cluster users are replaced by their expected values. We will illustrate the accuracy of this approximation in Section V through numerical examples.

Substituting (11) and (17), we can reexpress (71) as

$$\hat{\gamma}_n \approx \sum_{r=1}^R \frac{(\mathbf{h}_{rn}^H (\mathbf{R}_r^{-1})^H \mathbf{h}_{rn}) (\mathbf{h}_{rn}^H \mathbf{R}_r^{-1} \mathbf{h}_{rn})}{\mathbf{h}_{rn}^H (\mathbf{R}_r^{-1})^H \mathbf{R}_r \mathbf{R}_r^{-1} \mathbf{h}_{rn}} \quad (72)$$

where $\mathbf{R}_r = (\mathbf{H}_r \mathbf{H}_r^H + \boldsymbol{\Sigma}_r)$. Finally, we apply the matrix inversion lemma to further simplify (72) and derive the SINR of user n at the central processor as

$$\begin{aligned} \hat{\gamma}_n &\approx \sum_{r=1}^R \frac{(\mathbf{h}_{rn}^H (\mathbf{R}_r^{-1})^H \mathbf{h}_{rn}) (\mathbf{h}_{rn}^H \mathbf{R}_r^{-1} \mathbf{h}_{rn})}{\mathbf{h}_{rn}^H (\mathbf{R}_r^{-1})^H \mathbf{h}_{rn}} \\ &= \sum_{r=1}^R \mathbf{h}_{rn}^H \mathbf{R}_r^{-1} \mathbf{h}_{rn} \approx \sum_{r=1}^R \gamma_{rn}. \end{aligned} \quad (73)$$

The above derivation shows that, when averaged over the out-of-cluster interference, the global LMMSE estimation at the central processor results in a sum of all R SINRs from applying local LMMSE at the base stations.

APPENDIX B PROOF OF EQUATION (28)

In (27), the CF of γ_n is given as

$$\phi_{\gamma_n}(t) = \frac{1}{|\mathbf{X}|} E \left\{ \frac{|\mathbf{I} + \bar{\mathbf{H}}_n^H \boldsymbol{\Sigma}^{-1} \bar{\mathbf{H}}_n|}{|\mathbf{I} + \bar{\mathbf{H}}_n^H \boldsymbol{\Xi} \bar{\mathbf{H}}_n|} \right\}, \quad (74)$$

where $\boldsymbol{\Xi} = \boldsymbol{\Sigma}^{-1} \mathbf{X}^{-1}$. Let $\mathbf{I} = E\{\mathbf{P}^H \mathbf{P}\}$ where \mathbf{P} is a $\varsigma \times N-1$ matrix with each element drawn from a complex Gaussian distribution $\mathcal{CN}(0, \frac{1}{\varsigma})$. In the limit of $\varsigma \rightarrow \infty$, we can reexpress the identity matrix as

$$\mathbf{I} = \lim_{\varsigma \rightarrow \infty} \{\mathbf{P}^H \mathbf{P}\}. \quad (75)$$

Substituting (75) into (74) we can write

$$\begin{aligned} \phi_{\gamma_n}(t) &= \frac{1}{|\mathbf{X}|} \lim_{\varsigma \rightarrow \infty} E \left\{ \frac{|\mathbf{P}^H \mathbf{P} + \bar{\mathbf{H}}_n^H \boldsymbol{\Sigma}^{-1} \bar{\mathbf{H}}_n|}{|\mathbf{P}^H \mathbf{P} + \bar{\mathbf{H}}_n^H \boldsymbol{\Xi} \bar{\mathbf{H}}_n|} \right\} \\ &= \frac{1}{|\mathbf{X}|} \lim_{\varsigma \rightarrow \infty} E \left\{ \frac{\left| \left(\mathbf{P}^H, \bar{\mathbf{H}}_n^H \boldsymbol{\Sigma}^{-\frac{1}{2}} \right) \left(\frac{\mathbf{P}}{\boldsymbol{\Sigma}^{-\frac{1}{2}} \bar{\mathbf{H}}_n} \right) \right|}{\left| \left(\mathbf{P}^H, \bar{\mathbf{H}}_n^H \boldsymbol{\Xi}^{\frac{1}{2}} \right) \left(\frac{\mathbf{P}}{\boldsymbol{\Xi}^{\frac{1}{2}} \bar{\mathbf{H}}_n} \right) \right|} \right\} \\ &= \frac{1}{|\mathbf{X}|} \lim_{\varsigma \rightarrow \infty} E \left\{ \frac{|\mathbf{Q}^H \bar{\boldsymbol{\Sigma}} \mathbf{Q}|}{|\mathbf{Q}^H \bar{\boldsymbol{\Xi}} \mathbf{Q}|} \right\}, \end{aligned} \quad (76)$$

where $\bar{\boldsymbol{\Sigma}} = \text{diag}(\mathbf{I}, \boldsymbol{\Sigma}^{-\frac{1}{2}})$, $\bar{\boldsymbol{\Xi}} = \text{diag}(\mathbf{I}, \boldsymbol{\Xi}^{-\frac{1}{2}})$, and $\mathbf{Q} = \begin{pmatrix} \mathbf{P} \\ \bar{\mathbf{H}}_n \end{pmatrix}$. Then using the fact that

$$|\mathbf{Q}^H \mathbf{Q}| = \prod_{i=1}^{N-1} \mathbf{q}_i^H \left(\mathbf{I} - \mathbf{Q}_i (\mathbf{Q}_i^H \mathbf{Q}_i)^{-1} \mathbf{Q}_i^H \right) \mathbf{q}_i, \quad (77)$$

where \mathbf{q}_i is the i -th column of \mathbf{Q} , \mathbf{Q}_i is \mathbf{Q} with columns $1, 2, \dots, i-1$, $|\mathbf{Q}_1^H \mathbf{Q}_1| = 1$ and $\mathbf{Q}_1 (\mathbf{Q}_1^H \mathbf{Q}_1)^{-1} \mathbf{Q}_1^H = \mathbf{0}$, we can re-express the denominator and the numerator of (76) for a large but finite value of ς as

$$\begin{aligned} \phi_{\gamma_n}(t) &\approx \frac{1}{|\mathbf{X}|} \prod_{i=1}^{N-1} E \left\{ \frac{\mathbf{q}_i^H \left(\bar{\boldsymbol{\Sigma}} - \bar{\boldsymbol{\Sigma}} \mathbf{Q}_i (\mathbf{Q}_i^H \bar{\boldsymbol{\Sigma}} \mathbf{Q}_i)^{-1} \mathbf{Q}_i^H \bar{\boldsymbol{\Sigma}} \right) \mathbf{q}_i}{\mathbf{q}_i^H \left(\bar{\boldsymbol{\Xi}} - \bar{\boldsymbol{\Xi}} \mathbf{Q}_i (\mathbf{Q}_i^H \bar{\boldsymbol{\Xi}} \mathbf{Q}_i)^{-1} \mathbf{Q}_i^H \bar{\boldsymbol{\Xi}} \right) \mathbf{q}_i} \right\}. \end{aligned} \quad (78)$$

Note that the above approximation assumes the product terms in (77) to be independent, which is true only when \mathbf{q}_i contains independent and identical elements. In the present setting of this paper with base station cooperation all elements in \mathbf{q}_i are not identical. However, this approximation is motivated by the fact that part of \mathbf{q}_i which is contributed by \mathbf{P} is identical. Next we

apply the standard Laplace type approximation [30] into (78) as follows.

$$\begin{aligned}
 \phi_{\gamma_n}(t) & \approx \frac{1}{|\mathbf{X}|} \prod_{i=1}^{N-1} \frac{E \left\{ \mathbf{q}_i^H \left(\bar{\Sigma} - \bar{\Sigma} \mathbf{Q}_i \left(\mathbf{Q}_i^H \bar{\Sigma} \mathbf{Q}_i \right)^{-1} \mathbf{Q}_i^H \bar{\Sigma} \right) \mathbf{q}_i \right\}}{E \left\{ \mathbf{q}_i^H \left(\bar{\Xi} - \bar{\Xi} \mathbf{Q}_i \left(\mathbf{Q}_i^H \bar{\Xi} \mathbf{Q}_i \right)^{-1} \mathbf{Q}_i^H \bar{\Xi} \right) \mathbf{q}_i \right\}}, \\
 & \approx \frac{1}{|\mathbf{X}|} \frac{E \left\{ \prod_{i=1}^{N-1} \mathbf{q}_i^H \left(\bar{\Sigma} - \bar{\Sigma} \mathbf{Q}_i \left(\mathbf{Q}_i^H \bar{\Sigma} \mathbf{Q}_i \right)^{-1} \mathbf{Q}_i^H \bar{\Sigma} \right) \mathbf{q}_i \right\}}{E \left\{ \prod_{i=1}^{N-1} \mathbf{q}_i^H \left(\bar{\Xi} - \bar{\Xi} \mathbf{Q}_i \left(\mathbf{Q}_i^H \bar{\Xi} \mathbf{Q}_i \right)^{-1} \mathbf{Q}_i^H \bar{\Xi} \right) \mathbf{q}_i \right\}}, \\
 & = \frac{1}{|\mathbf{X}|} \frac{E \left\{ \left| \mathbf{Q}_i^H \bar{\Sigma} \mathbf{Q}_i \right| \right\}}{E \left\{ \left| \mathbf{Q}_i^H \bar{\Xi} \mathbf{Q}_i \right| \right\}} = \frac{1}{|\mathbf{X}|} \frac{E \left\{ \left| \mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \bar{\mathbf{H}}_n \right| \right\}}{E \left\{ \left| \mathbf{I} + \bar{\mathbf{H}}_n^H \Sigma^{-1} \mathbf{X}^{-1} \bar{\mathbf{H}}_n \right| \right\}}. \tag{79}
 \end{aligned}$$

The accuracy of this approximation is illustrated in Section V using numerical examples.

REFERENCES

- [1] V. Jungnickel *et al.*, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, May 2014.
- [2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [3] E. Aktas, J. Evans, and S. Hanly, "Distributed decoding in a cellular multiple-access channel," *IEEE Commun. Mag.*, vol. 7, no. 1, pp. 241–250, Jan. 2008.
- [4] R. Senanayake, P. L. Yeoh, and J. S. Evans, "On the bit error probability of optimal multiuser detectors in cooperative cellular networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2472–2478, Jun. 2014.
- [5] R. Senanayake, P. L. Yeoh, and J. S. Evans, "Optimal multiuser detection in a cooperative two-cell network," in *Proc. Aust. Commun. Theory Workshop (AusCTW'13)*, Adelaide, Australia, Feb. 2013, pp. 1–6.
- [6] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "The effect of macrodiversity on the performance of MLD in flat Rayleigh/Rician fading," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1764–1767, Nov. 2012.
- [7] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "The effect of macrodiversity on the performance of MLD in flat Rayleigh/Rician fading," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2240–2251, May 2013.
- [8] S. Hanly and P. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommun. Syst.*, vol. 1, pp. 1–42, Jan. 1993.
- [9] M. N. Bacha, J. S. Evans, and S. V. Hanly, "On the capacity of MIMO cellular networks with macrodiversity," in *Proc. Aust. Commun. Theory Workshop (AusCTW'06)*, Perth, Australia, Feb. 2006, pp. 105–109.
- [10] S. Venkatesan and A. Lozano, "Network MIMO: Overcoming inter-cell interference in indoor wireless systems," in *Proc. Conf. Rec. 41st Asilomar Conf. Signals Syst. Comput. (ACSSC'07)*, Nov. 2007, pp. 83–87.
- [11] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.
- [12] P. Marsch, S. Khattak, and G. Fettweis, "A framework for determining realistic capacity bounds for distributed antenna systems," in *Proc. Inf. Theory Workshop (ITW'06)*, Punta del Este, Uruguay, Oct. 2006, pp. 571–575.
- [13] H. Huang, M. Trivellato, A. Hottinen, M. Shaf, P. J. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2983–2989, Jun. 2009.
- [14] A. Lozano, R. W. Heath, and J. G. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5213–5226, Sep. 2013.
- [15] A. Lozano, J. G. Andrews, and R. W. Heath, "Spectral efficiency limits in pilot-assisted cooperative communications," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT'12)*, Cambridge, MA, USA, Jul. 2012, pp. 1132–1136.
- [16] A. Lozano, J. G. Andrews, and R. W. Heath, "On the limitations of cooperation in wireless networks," in *Proc. Inf. Theory Appl. Workshop (ITA'12)*, San Diego, CA, USA, Feb. 2012, pp. 123–230.
- [17] A. Lapidoth, N. Levy, S. Shamai Shitz, and M. Wigger, "Cognitive Wyner networks with clustered decoding," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6342–6367, Oct. 2014.
- [18] S. J. Kim, S. Jain, and G. B. Giannakis, "Backhaul-constrained multi-cell cooperation using compressive sensing and spectral clustering," in *Proc. IEEE 13th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC'12)*, Jun. 2012, pp. 65–69.
- [19] E. Katranaras, M. A. Imran, M. Dianati, and R. Tafazolli, "Green inter-cluster interference management in uplink of multi-cell processing systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6580–6592, Dec. 2014.
- [20] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with interference in distributed large-scale MIMO systems: A statistical approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct. 2014.
- [21] E. Bjornson, E. G. Larsson, and M. Debbah, "Optimizing multi-cell massive MIMO for spectral efficiency: How many users should be scheduled?" in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP'14)*, London, U.K., Dec. 2014, pp. 612–616.
- [22] S. Nagaraj, M. L. Honig, and K. Zeineddine, "Distributed optimization of multi-cell uplink co-operation with Backhaul constraints," in *Proc. IEEE Int. Conf. Commun. (ICC'15)*, London, U.K., Jun. 2015, pp. 3987–3992.
- [23] R. Senanayake, P. L. Yeoh, and J. S. Evans, "On the sum capacity of cluster-based cooperative cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC'15)*, London, U.K., Jun. 2015, pp. 1613–1618.
- [24] R. Senanayake, P. L. Yeoh, and J. S. Evans, "Distributed LMMSE estimation in cooperative cellular networks," in *Proc. IEEE Int. Conf. Commun. (ICC'15)*, London, U.K., Jun. 2015, pp. 4083–4088.
- [25] H. Gao, P. J. Smith, and M. V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 666–672, May 1998.
- [26] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE J. Sel. Areas Commun.*, vol. AC-2, no. 4, pp. 528–539, Jul. 1984.
- [27] A. F. Molisch, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2011.
- [28] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "Performance analysis of dual-user macrodiversity MIMO systems with linear receivers in flat Rayleigh fading," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4394–4404, Dec. 2012.
- [29] M. K. Simon and M. S. Alouini, *Digital Communication Over Fading Channels*. New York, NY, USA: Wiley-Interscience, 2005.
- [30] O. Lieberman, "A Laplace approximation to the moments of a ratio of quadratic forms," *Biometrika*, vol. 81, p. 681–690, Dec. 1994.
- [31] D. A. Basnayaka, P. J. Smith, and P. A. Martin, "Ergodic sum capacity of macrodiversity MIMO systems in flat Rayleigh fading," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5257–5270, Sep. 2013.
- [32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [33] X. Gao, B. Jiang, X. Li, A. Gershman, and M. R. McKay, "Statistical eigenmode transmission over jointly correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3735–3750, Aug. 2009.
- [34] H. Minc, *Permanents*. Reading, MA, USA: Addison-Wesley, 1978.
- [35] I. Grashytn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.
- [36] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.



Rajitha Senanayake (S'12–M'15) received the B.E. degree in electrical and electronics engineering from the University of Peradeniya, Peradeniya, Sri Lanka, the B.I.T. degree in information technology from the University of Colombo, Colombo, Sri Lanka, and the Ph.D. degree in electrical and electronics engineering from the University of Melbourne, Melbourne, Vic., Australia, in 2009, 2010, and 2015, respectively. From 2009 to 2011, she was with the research and development team with Excel Technology, Colombo, Sri Lanka. She is currently a Research Fellow with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Australia. Her research interests include communications theory and wireless communications. She was the recipient of the Melbourne International Research Scholarship and the Melbourne International Fee Remission Scholarship awarded by the University of Melbourne.



Phee Lep Yeoh (S'08–M'12) received the B.E. and Ph.D. degrees (with University Medal) from the University of Sydney, Sydney, N.S.W., Australia, in 2004 and 2012, respectively. From 2008 to 2012, he was with the Telecommunications Laboratory, University of Sydney, and the Wireless and Networking Technologies Laboratory, Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia. In 2012, he joined the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne,

Vic., Australia. His research interests include heterogeneous wireless networks, large-scale MIMO networks, cognitive co-operative communications, and multiscale molecular communications. He was the recipient of the Australian Research Council Discovery Early Career Researcher Award, the University of Sydney Postgraduate Award, and the CSIRO Postgraduate Scholarship.



Jamie S. Evans (S'93–M'98) was born in Newcastle, Australia, in 1970. He received the B.S. degree in physics and the B.E. degree in computer engineering from the University of Newcastle, Callaghan, N.S.W., Australia, and the M.S. and Ph.D. degrees in electrical engineering from the University of Melbourne, Melbourne, Vic., Australia, in 1992, 1993, 1996, and 1998, respectively. From March 1998 to June 1999, he was a Visiting Researcher at the Department of Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA,

USA. He returned to Australia to take up a position as a Lecturer with the University of Sydney, Sydney, N.S.W., Australia, where he stayed until July 2001. From July 2001 to March 2012, he was with the Department of Electrical and Electronic Engineering, University of Melbourne. He is currently a Professor with the Department of Electrical and Computer Systems Engineering at Monash University, Melbourne, Vic., Australia. His research interests include communications theory, information theory, and statistical signal processing with a focus on wireless communications networks. He was the recipient of the University Medal upon graduation from the University of Newcastle, and the Chancellor's Prize for excellence for his Ph.D. thesis from University of Melbourne.