it is unfair for our protocol since the real-time capability of the compared system is reduced. As shown in Fig. 2(b), the proposed scheme with precoding selection can achieve a much higher diversity order given the same multiplexing gain.

VI. Conclusion

In this paper, a new network-coding-based transmission protocol for generalized MIMO Y channels has been proposed and studied. By applying SSA-NC and interference nulling, the proposed protocol is able to avoid cochannel interference among multiple destinations. A simple mapping method is proposed to ensure low-complexity demodulation at the relay. We have developed analytical results, such as SER and diversity gain of the protocol. Precoding selection has also been presented to further improve the reception reliability. Monte Carlo simulation is provided to demonstrate the performance of the proposed scheme.

References

[1] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proc. ACM SIGCOMM*, 2007, pp. 397–408.

[2] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," in *Proc. ACM SIGCOMM*, 2006, pp. 243–254.

[3] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[4] R. Zhang, Y. Liang, C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, Jun. 2009.

[5] D. Gündüz, A. Goldsmith, and H. Poor, "MIMO two-way relay channel: Diversity-multiplexing tradeoff analysis," in *Proc. 42nd Asilomar Conf. Signals, Syst., Comput.*, 2008, pp. 1474–1478.

[6] Z. Zhang, S. Liew, and P. Lam, Physical Layer Network Coding, 2007, Arxiv preprint arXiv:0704.2475.

[7] T. Koike-Akino, P. Popovski, and V. Tarokh, "Optimized constellations for two-way wireless relaying with physical network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 773–787, Jun. 2009.

[8] D. Gündüz, A. Yener, A. Goldsmith, and H. Poor, "The multi-way relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 339–343.

[9] N. Lee and J. Lim, "A novel signaling for communication on MIMO Y channel: Signal space alignment for network coding," in *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 2892–2896.

[10] S. Jafer and S. Shamai, "Degrees of freedom region of the MIMO X channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[11] N. Lee, J. Lim, and J. Chun, "Degrees of freedom of the MIMO Y channel: Signal space alignment for network coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3332–3342, Jul. 2010.

[12] A. Wiesel and Y. Eldar, "Zero-forcing precoding and generalized inverses," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[13] M. Joham, W. Utschick, and J. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, pt. 1, no. 8, pp. 2700–2712, Aug. 2005.

[14] M. Rupp, C. Mecklenbrauker, and G. Gritsch, "High diversity with simple space time block-codes and linear receivers," in *Proc. IEEE GLOBECOM*, 2003, vol. 1, pp. 302–306.

[15] Z. Ding, Z. Ma, and K. Leung, "Impact of network coding on system delay for multisource–multidestination scenarios," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 831–841, Feb. 2010.

[16] Z. Ding, T. Wang, M. Peng, W. Wang, and K. Leung, "On the design of network coding for multiple two-way relaying channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1820–1832, Jun. 2011.

[17] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*. Amsterdam, Netherlands: Elsevier, 2007.

[18] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

Distributed Probabilistic-Data-Association-Based Soft Reception Employing Base Station Cooperation in MIMO-Aided Multiuser Multicell Systems

Shaoshi Yang, *Student Member, IEEE*, Tiejun Lv, *Member, IEEE*, Robert G. Maunder, *Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—Intercell cochannel interference (CCI) mitigation is investigated in the context of cellular systems relying on dense frequency reuse (FR). A distributed base-station (BS)-cooperation-aided soft reception scheme using the probabilistic data association (PDA) algorithm and soft combining (SC) is proposed for the uplink of multiuser multicell MIMO systems. The realistic 19-cell hexagonal cellular model relying on unity FR is considered, where both the BSs and the mobile stations (MSs) are equipped with multiple antennas. Local-cooperation-based message passing is used, instead of a global message passing chain for the sake of reducing the backhaul traffic. The PDA algorithm is employed as a low-complexity solution for producing soft information, which facilitates the employment of SC at the individual BSs to generate the final soft decision metric. Our simulations and analysis demonstrate that, despite its low additional complexity and backhaul traffic, the proposed distributed PDA-aided SC (DPDA-SC) reception scheme significantly outperforms the conventional noncooperative benchmarkers. Furthermore, since only the index of the possible discrete value of the quantized converged soft information has to be exchanged for SC in practice, the proposed DPDA-SC scheme is relatively robust to the quantization errors of the soft information exchanged. As a beneficial result, the backhaul traffic is dramatically reduced at negligible performance degradation.**

***Index Terms*—Base station (BS) cooperation, cochannel interference (CCI), distributed multiple-input–multiple-output (MIMO), message passing, multicell processing, probabilistic data association (PDA), soft combining (SC).**

I. Introduction

Spectrally efficient techniques, such as multiple-input–multiple-output (MIMO) antennas and near-unity frequency reuse (FR) are expected to be employed in the next-generation cellular networks. In this context, the achievable performance gain of multiuser multicell MIMO systems is predominantly limited by the effect of intercell cochannel interference (CCI) [1]. Recently, advanced receiver techniques using base station (BS) cooperation for exploiting the potential capacity of cellular systems were investigated [2]–[4]. The simplest conceptual approach to BS cooperation is to assume that there is a controller, or central processing unit (CPU), that coordinates the operation of all BSs [2]. However, the CPU constitutes a single point of potential failure; thus, the entire network is vulnerable. Additionally, since the

Manuscript received June 20, 2010; revised November 4, 2010 and April 18, 2011; accepted June 8, 2011. Date of publication June 16, 2011; date of current version September 19, 2011. This work was supported in part by the National Natural Science Foundation of China under Grant NSFC-60972075, by the China Scholarship Council, by the EPSRC, U.K. under the auspices of the China-U.K. Science Bridge, by the RC-U.K. under the India-U.K. Advanced Technology Centre (IU-ATC), and by the EU under the auspices of the Concerto project. The review of this paper was coordinated by Prof. E. Bonek.

S. Yang is with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K., and also with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: sy7g09@ecs.soton.ac.uk).

T. Lv is with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: lvtiejun@bupt.edu.cn).

R. G. Maunder and L. Hanzo are with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K. (e-mail: rm@ecs.soton.ac.uk; lh@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TVT.2011.2159822

complexity of multiuser detection (MUD) is dominated by the number of users, having a CPU imposes a potentially excessive computational burden and huge backhaul traffic, and thus may become less attractive.

A distributed implementation of the iterative interference cancellation framework used in [2] was then developed for single-antenna-aided multiuser systems in [3]. Although this scheme was shown to strike an attractive compromise, it has an exponentially increased computational complexity imposed by the computation of the soft information using the max-log maximum *a posteriori* (MAP) algorithm. Furthermore, the multiple rounds of iterative message exchange operations between the cooperative BSs may still impose potentially excessive backhaul traffic. The belief propagation (BP) algorithm was also applied to the problem of distributed detection in the single-antenna-aided 2-D Wyner model [4], which performs a chainlike message passing between all the BSs and provides a globally near-optimum solution. However, it relies on network-wide optimum information exchange, which unfortunately results in potentially excessive backhaul traffic and latency, especially for a starlike architecture routinely used for interconnecting the BSs [3].

The probabilistic data association (PDA) method, which was originally proposed for target tracking [5], may also be developed into a reduced-complexity design alternative for the MAP algorithm [6]–[8]. The PDA technique may be regarded as a promising detection technique, owing to its attractive properties. First, it may achieve a near-optimal MUD performance, particularly in the context of code-division multiple-access (CDMA) systems [6]. Second, it has polynomial complexity, increasing no faster than $O(L^3)$, where L is the number of transmit antennas in MIMO systems [8] or the number of users in CDMA [6]. Furthermore, the higher the number of transmit antennas or users, the better the attainable performance, provided that the channel is not rank-deficient [10].

In this paper, we propose a simple but effective soft-combining (SC) technique, and we develop the PDA algorithm into a distributed multiuser multicell soft-reception scheme to mitigate the prohibitive computational complexity and the huge amount of backhaul traffic faced by multicell processing employing BS cooperation. A realistic 19-cell hexagonal cellular MIMO-aided network model relying on either perfect or imperfect channel estimation is considered. In this model, the entire channel consists of multiple matrix subchannels, rather than of scalar/vector subchannels, as in [2]–[4]. SC is used at each BS to generate the final soft-decision information, which indicates that the *fundamental philosophy* of the proposed method is not that of "*interference cancellation*" but "*knowledge sharing and data fusion*." Additionally, we investigate the impact of quantization on both the backhaul traffic and the performance of the proposed scheme. Since, in practice, only the *index* of the possible discrete value of the converged soft information has to be exchanged for SC operation, the proposed scheme is relatively robust to quantization errors of the soft information exchanged, which dramatically reduces the backhaul traffic at negligible performance degradation. We also considered the challenging rank-deficient scenario, where the number of transmitters is higher than that of the receivers. Despite its significant performance gain over the conventional noncooperative MUD schemes, the proposed distributed PDA-aided SC (DPDA-SC) approach imposes a modest *cubically* increasing complexity as a function of the number of streams processed while maintaining low backhaul traffic. Low complexity is achieved as a benefit of the PDA's rapid convergence since only converged soft information is exchanged among the BSs of the specific cooperative BS cluster, requiring a single action.

II. Hexagonal Cellular Network Model

Consider a hexagonal cellular network model, where both the BSs and MSs are equipped with multiple antennas. Therefore, instead of having a conventional point-to-point channel impulse response (CIR) between each BS and MS [3], [4], we have a CIR matrix, where the interference imposed on each MS stems from not only other MSs but also their own multiple antennas. A unity FR is employed for all the cells, and an orthogonal multiple access technique may be applied. Therefore, the intracell interference is assumed to be negligible, whereas the CCI imposed by MSs of the interfering cells is dominant.

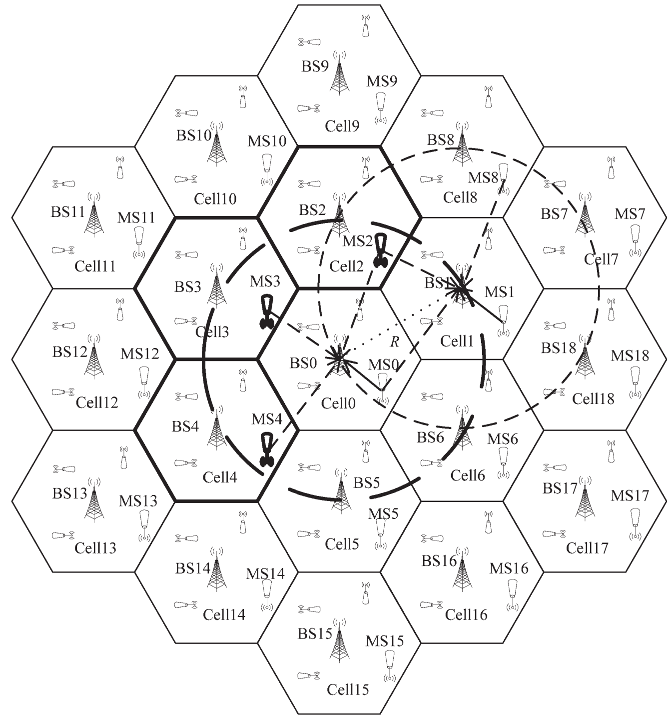

Fig. 1. Example setup showing a 19-cell hexagonal cellular model with CCI and unity FR.

Let us consider the topology shown in Fig. 1 as an example, which constitutes a *snapshot* of the dynamic network at a specific scheduling interval. Assume that there are N_b BSs and N_u MSs in this network (for the sake of comparability with classic cellular networks, N_b is set to 19 in Fig. 1), and K_i MSs in each cell, $i = 1, 2, \ldots, N_b$. For simplicity of analysis, we assume that each of them is equipped with M_b and M_u antennas, respectively. The position of each network entity is represented by its *dynamically updated polar coordinates* with respect to the specific *serving* BS. This dynamic coordinate system naturally lends itself to distributed processing. For example, if the central cell (Cell0) of Fig. 1 is considered, the point at BS0 may be defined as the origin and the BS in the upper right adjacent cell of Cell0 may be described as $\text{BS1} = (R, \pi/6)$, where R is the distance between two immediate neighbor BSs.

The available frequency band may be dynamically allocated to the MSs according to their uplink channel quality information evaluated by each BS. For simplicity of graphical illustration, the same frequency slot is tentatively assigned to the MSs situated at the same relative position in their corresponding home cell, as shown in Fig. 1. For example, MS0, MS1, MS2, ..., MS18 in the lower right corner of each cell are all the cochannel users. *Not all* the cochannel users but only those located within the *detectable range (DR)* of a specific BS are considered to be interfering with the desired user. Hence, the number of *effective* interferers with respect to each user may be different[1]. For instance, when the signal of MS0 is expected to be detected by its home

[1]The strong interference that is decodable is explicitly considered here, whereas the weak interference is implicitly treated as noise.

BS, i.e., by BS0, only MS2, MS3, and MS4, which are emphasized as bold, are the effective interferers. By contrast, the other cochannel users such as MS1, MS5, and MS6 are not deemed to be effective interferers since they are outside the disk centered at BS0 and having a radius of R. More explicitly, they are outside the DR of BS0. For MS7, MS10, and MS11, however, the number of interferers is two, one, and zero, respectively, although they are all at the boundary of the network. We emphasize that the interfering MSs contaminating the reception of each MS may be different in the next scheduling interval, owing to the dynamic nature of the network. On the other hand, the signal of MS0 may also be adequately received at BS1, BS5, and BS6, which therefore have the potential to act as the serving BSs of MS0. Similarly, the number of adjacent BSs supporting each MS may be different as well. We assume in general that, for each served MS, there are C_u such effective cochannel MSs and C_b adjacent serving BSs, respectively. Then, the four-tuple (M_b, M_u, C_b, C_u) may be used to represent the cooperating BS cluster, which is dynamically changing for the different served MSs.

III. Cooperative Distributed Soft Reception

A. Signal Model

Based on the hexagonal topology of Fig. 1, we consider an idealized synchronous uplink where the signal received at BSk is modeled as

$$\mathbf{y}^k = \mathbf{H}_k^k \mathbf{x}^k + \sum_{\alpha_i \neq k} \mathbf{H}_{\alpha_i}^k \mathbf{x}^{\alpha_i} + \mathbf{n}^k = \mathbf{H}_k^k \mathbf{x}^k + \mathbf{N}^k + \mathbf{n}^k \quad (1)$$

where $\mathbf{x}^{\alpha_i}$ is the length-M_u vector of symbols transmitted from MSα_i in Cellα_i, and each symbol is from the modulation constellation $\mathcal{A} = \{a_1, a_2, \ldots, a_M\}$ with cardinality M. Still referring to (1), $\mathbf{H}_{\alpha_i}^k$ is the $(M_b \times M_u)$-element channel matrix between MSα_i and BSk, $i = 1, 2, \ldots, C_u$, $k = 1, 2, \ldots, N_b$, whereas $\mathbf{n}^k$ is the length-M_b complex-valued circular symmetric Gaussian noise vector[2] with zero mean and covariance matrix $N_0 \mathbf{I}_{M_b}$ at BSk, where $\mathbf{I}_{M_b}$ is an $(M_b \times M_b)$-element identity matrix.

Let us now define the interference intensity as the channel gain ratio of the interfering users over that of the local desired user, i.e., as $\rho_{\alpha_i}^k = \|\mathbf{H}_{\alpha_i}^k\|_F / \|\mathbf{H}_k^k\|_F$, $0 \le \rho_{\alpha_i}^k \le 1$, where $\|\cdot\|_F$ represents the Frobenius norm.

As opposed to conventional noncooperative detection, the distributed detection of $\mathbf{x}^k$ carried out with the aid of BS cooperation detects not only the desired user's signal in the local cell but also all the cochannel users' signals overheard from the neighboring cells. More explicitly, the cochannel users' signals are no longer considered as detrimental interference; we rather consider these cochannel users' soft decision information as a useful source of further information to be exploited by cooperative processing via message passing among the BSs. To this end, the received signal model of (1) may be reformulated as a virtual or distributed MIMO model, where the cooperating BSs may be viewed as MIMO elements, yielding

$$\mathbf{y}^k = \mathbf{G}^k \mathbf{s}^k + \mathbf{n}^k \quad (2)$$

where we have $\mathbf{G}^k = [\mathbf{H}_k^k, \mathbf{H}_{\alpha_1}^k, \ldots, \mathbf{H}_{\alpha_{C_u}}^k]$, $\mathbf{s}^k = [(\mathbf{x}^k)^T, (\mathbf{x}^{\alpha_1})^T, \ldots, (\mathbf{x}^{\alpha_{C_u}})^T]^T$, and the elements of $\mathbf{s}^k$ are denoted as s_t^k, $t = 1, 2, \ldots, M_u\ (C_u + 1)$. For the sake of generality, $M_u\ (C_u + 1) \le M_b$ is not assumed here.

[2]Note that the noise considered here consists of both ordinary channel noise and weak interference that is not decodable. Treating the weak interference as noise has been recently proved optimal in the weak interference regime [11].

In the case of *imperfect* channel knowledge, the estimated channel matrices $\hat{\mathbf{H}}_k^k$ and $\hat{\mathbf{H}}_{\alpha_i}^k$ are respectively associated with the channel-estimation error matrices $\mathbf{E}_k$ and $\mathbf{E}_{\alpha_i}$, which may be deemed to obey the standard Gaussian distribution of $\mathcal{CN}(0, 1)$. They can be written as

$$\hat{\mathbf{H}}_k^k = \beta_k \mathbf{H}_k^k + \sqrt{1 - \beta_k^2}\mathbf{E}_k \quad (3)$$

$$\hat{\mathbf{H}}_{\alpha_i}^k = \beta_{\alpha_i} \mathbf{H}_{\alpha_i}^k + \sqrt{1 - \beta_{\alpha_i}^2}\mathbf{E}_{\alpha_i} \quad (4)$$

respectively [12], where β_k and β_{α_i} indicate the channel estimation quality and may be assumed to be close to 1 but not higher than 1. Thus, the received signal models of (1) and (2) may be rewritten as

$$\mathbf{y}^k = \hat{\mathbf{H}}_k^k \mathbf{x}^k + \sum_{\alpha_i \neq k} \hat{\mathbf{H}}_{\alpha_i}^k \mathbf{x}^{\alpha_i} + \mathbf{n}^k \quad (5)$$

$$\mathbf{y}^k = \hat{\mathbf{G}}^k \mathbf{s}^k + \mathbf{n}^k \quad (6)$$

respectively, where we have $\hat{\mathbf{G}}^k = [\hat{\mathbf{H}}_k^k, \hat{\mathbf{H}}_{\alpha_1}^k, \ldots, \hat{\mathbf{H}}_{\alpha_{C_u}}^k]$, or $\hat{\mathbf{G}}^k = \bar{\beta}_k \mathbf{G}^k + \sqrt{1 - \bar{\beta}_k^2}\bar{\mathbf{E}}_k$, with $\bar{\beta}_k$ and $\bar{\mathbf{E}}_k$ being the composite-channel estimation error indicators. Note that, when we have $\beta_k = \beta_{\alpha_i} = 1$, the signal model under imperfect CSI transforms into that of perfect CSI. Without loss of generality, here, we will continue by considering perfect channel estimation while presenting the proposed DPDA-SC scheme. The case of imperfect channel knowledge may be readily considered by the substitution of the corresponding perfect channels with the estimated channels.

B. Parallel Detection Using the PDA Algorithm

The first action of the DPDA-SC scheme is that the BSs perform parallel detection, employing the PDA algorithm as a low-complexity solution to estimate the *a posteriori* probability (APP) of each transmitted symbol without carrying out an exhaustive search in the space of all possible symbol combinations. Each BS jointly detects the signals of multiple users, including both the local user and other cells' users roaming close to this BS, which would be termed as interfering users in conventional noncooperative systems. For ease of exposition, we consider detection at BSk as an example and omit the BS index k in our forthcoming exposition.

Case I: When we have $M_u\ (C_u + 1) \le M_b$, for the sake of computational efficiency, the decorrelated signal model of [7] is adopted; hence, (2) may be further formulated as

$$\tilde{\mathbf{y}} = \mathbf{s} + \tilde{\mathbf{n}} = s_t \mathbf{e}_t + \sum_{l \neq t} s_l \mathbf{e}_l + \tilde{\mathbf{n}} \triangleq s_t \mathbf{e}_t + \mathbf{v}_t \quad (7)$$

where $\tilde{\mathbf{y}} = (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H \mathbf{y}$, $\tilde{\mathbf{n}}$ is a colored Gaussian noise with zero mean and covariance of $N_0\ (\mathbf{G}^H \mathbf{G})^{-1}$, $\mathbf{e}_l$ is a column vector with 1 in the lth position and 0 elsewhere, and $\mathbf{v}_t$ denotes the interference plus noise term for symbol s_t, for $t, l = 1, 2, \ldots, M_u\ (C_u + 1)$.

For each symbol s_t, we have a probability vector $\mathbf{P}(t)$ whose mth element $P_m(s_t|\mathbf{y})$ is the current estimate of the APP of having $s_t = a_m$, where $m = 1, 2, \ldots, M$, with a_m being the mth element of the modulation constellation $\mathcal{A}$. The key philosophy of the PDA algorithm is to approximate $\mathbf{v}_t$ obeying the multimodal Gaussian mixture distribution as a single multivariate colored Gaussian distributed random vector [6] with an updated mean of $\mathrm{E}(\mathbf{v}_t) = \sum_{l \neq t} \bar{s}_l \mathbf{e}_l$, covariance of $\mathrm{V}(\mathbf{v}_t) = \sum_{l \neq t} \mathrm{V}\{s_l\} \mathbf{e}_l \mathbf{e}_l^T + N_0 (\mathbf{G}^H \mathbf{G})^{-1}$, and pseudocovariance [9] of $\mathrm{U}(\mathbf{v}_t) = \sum_{l \neq t} \mathrm{U}\{s_l\} \mathbf{e}_l \mathbf{e}_l^T$, where

$$\bar{s}_l = \sum_{m=1}^{M} a_m P_m(s_l|\mathbf{y}) \quad (8)$$

$$\mathrm{V}\{s_l\} = \sum_{m=1}^{M} (a_m - \bar{s}_l)(a_m - \bar{s}_l)^* P_m(s_l|\mathbf{y}) \tag{9}$$

$$\mathrm{U}\{s_l\} = \sum_{m=1}^{M} (a_m - \bar{s}_l)(a_m - \bar{s}_l)^T P_m(s_l|\mathbf{y}). \tag{10}$$

Here, $P_m(s_l|\mathbf{y})$ is initialized as a uniform distribution and will be replaced with an updated probability at each iteration of the PDA algorithm.

Let

$$\mathbf{w}_m^{(t)} = \tilde{\mathbf{y}} - a_m^{(t)}\mathbf{e}_t - \sum_{l\neq t} \bar{s}_l \mathbf{e}_l \tag{11}$$

$$\varphi_m(s_t) \triangleq \exp\left(-\begin{pmatrix} \Re\left(\mathbf{w}_m^{(t)}\right) \\ \Im\left(\mathbf{w}_m^{(t)}\right)\end{pmatrix}^T \mathbf{\Lambda}_t^{-1} \begin{pmatrix} \Re\left(\mathbf{w}_m^{(t)}\right) \\ \Im\left(\mathbf{w}_m^{(t)}\right)\end{pmatrix}\right) \tag{12}$$

where we have

$$\mathbf{\Lambda}_t \triangleq \begin{pmatrix} \Re\left(\mathrm{V}(\mathbf{v}_t) + \mathrm{U}(\mathbf{v}_t)\right) & -\Im\left(\mathrm{V}(\mathbf{v}_t) - \mathrm{U}(\mathbf{v}_t)\right) \\ \Im\left(\mathrm{V}(\mathbf{v}_t) + \mathrm{U}(\mathbf{v}_t)\right) & \Re\left(\mathrm{V}(\mathbf{v}_t) - \mathrm{U}(\mathbf{v}_t)\right)\end{pmatrix} \tag{13}$$

and $a_m^{(t)}$ indicates that a_m is assigned to s_t, and $\Re(\cdot)$ and $\Im(\cdot)$ represent the real and imaginary parts of a complex variable, respectively.

Since it is assumed that the transmitted symbols have equal *a priori* probabilities, the APP of s_t is given as

$$P_m(s_t|\mathbf{y}) = \frac{p_m(\mathbf{y}|s_t)P(s_t = a_m)}{\sum\limits_{m=1}^{M} p_m(\mathbf{y}|s_t)P(s_t = a_m)} \approx \frac{\varphi_m(s_t)}{\sum\limits_{m=1}^{M} \varphi_m(s_t)}. \tag{14}$$

In summary, the PDA algorithm proceeds as follows:

Step 1) Initialization: set the initial values of the symbol probabilities $P_m(s_t|\mathbf{y})$ using a uniform distribution for $\forall t = 1, 2, \ldots, M_u \ (C_u + 1)$, $\forall m = 1, 2, \ldots, M$, i.e., $P_m(s_t|\mathbf{y}) = 1/M$; set the iteration counter to $z = 1$.

Step 2) Set the symbol index to $t = 1$.

Step 3) Based on the current values of $\{\mathbf{P}(l)\}_{l\neq t}$, compute $P_m(s_t|\mathbf{y})$ via (8) –(14), which will replace the corresponding elements of $\mathbf{P}(t)$.

Step 4) If $t < M_u \ (C_u + 1)$, let $t = t + 1$, and go to step 3. Otherwise, go to step 5.

Step 5) If $\mathbf{P}(t)$ has converged for $\forall t$ or the iteration index has reached its maximum, terminate the iteration. Otherwise, let $z = z + 1$, and return to step 2.

Case 2: When we have $M_u \ (C_u + 1) > M_b$, the appropriately modified version of the PDA method [13] may be applied to the current problem. Alternatively, the nondecorrelated signal model of [10] may be applied, which yields an equivalent performance to that of the decorrelated-signal-model-based PDA [14] when $M_u \ (C_u + 1) \leq M_b$. In the case of the nondecorrelated model, (2) may be expanded as

$$\mathbf{y} = \mathbf{g}_t s_t + \sum_{l\neq t} \mathbf{g}_l s_l + \mathbf{n} \triangleq \mathbf{g}_t s_t + \mathbf{u}_t \tag{15}$$

where $\mathbf{g}_l$ is the lth column of $\mathbf{G}^k$. Then, the PDA algorithm is obtained using a similar derivation to that of its counterpart in Case 1, as outlined throughout (8)–(14).

C. Parallel Message Exchange via UCS Mode

The effective neighboring BSs incorporated in the same cooperative BS cluster will then exchange their soft decision information produced by the PDA algorithm in parallel, assuming the presence of an idealized optical fiber backbone. It is emphasized that each BS plays the role of both client and server. In other words, each BS operates in a unified-client-server (UCS) mode. As a server, it helps detect the signals of all cochannel users at all the cooperating cells, and then, the soft decision information is sent to each user's home BS. This message passing action substantially benefits the signal detection process in neighboring cells. As a client, each BS receives multiple copies of soft decision information for its own desired user's signal. The exchange of soft information is carried out with the aid of BS cooperation. For example, BS0 estimates the APP of its own user MS0 and additionally forwards the APPs of MS2, MS3, and MS4 to the corresponding sites of BS2, BS3, and BS4, respectively. On the other hand, to aid the detection of MS0, the surrounding BS0, BS1, BS5, and BS6 output $P_m(s_t^0|\mathbf{y}^0)$, $P_m(s_t^0|\mathbf{y}^1)$, $P_m(s_t^0|\mathbf{y}^5)$, and $P_m(s_t^0|\mathbf{y}^6)$, respectively, and all these probabilities will be collected at BS0, i.e., the home BS of MS0. Therefore, BS0, BS1, BS5, and BS6 assist in the detection of MS0.

D. SC and Final Decision

Based on the aggregated soft decision information, each BS individually performs the SC of all the copies of its own desired user's soft information according to

$$P_m(s_t|\mathbf{y}_{\mathrm{coop}}) = P_m(s_t|\mathbf{y}^k) \prod_{j=1}^{C_b} P_m(s_t|\mathbf{y}^{\beta_j}) \tag{16}$$

where $\mathbf{y}_{\mathrm{coop}}$ stands for the received signal used for BS cooperation, i.e., $\mathbf{y}^k$ and $\mathbf{y}^{\beta_j}$, $j = 1, \ldots, C_b$. $P_m(s_t|\mathbf{y}_{\mathrm{coop}})$ represents the composite soft decision information[3]. Again, let us consider the detection of MS0's signal as an example, where the composite soft decision information is $P_m(s_t^0|\mathbf{y}_{\mathrm{coop}}) = P_m(s_t^0|\mathbf{y}^0) \times P_m(s_t^0|\mathbf{y}^1)\ P_m(s_t^0|\mathbf{y}^5)P_m(s_t^0|\mathbf{y}^6)$. Note that, for the sake of numerical stability, the soft information should be further normalized as

$$P_m(s_t|\mathbf{y}_{\mathrm{coop}})_{\mathrm{norm}} = \frac{P_m(s_t|\mathbf{y}_{\mathrm{coop}})}{\sum\limits_m P_m(s_t|\mathbf{y}_{\mathrm{coop}})}. \tag{17}$$

Finally, make a decision for each transmitted symbol s_t, yielding $\hat{s}_t = a_{m'}$ at each corresponding BS, where

$$m' = \arg \max_{d=1,2,\ldots,M} \left\{P_d(s_t|\mathbf{y}_{\mathrm{coop}})_{\mathrm{norm}}\right\}. \tag{18}$$

E. Complexity Analysis

The proposed DPDA-SC scheme has worst-case complexity at each BS per iteration, which is on the order of $O[(M_u\ (C_u + 1))^3]$, provided that the Sherman–Morrison–Woodbury formula is applied for the computation of $\mathbf{\Lambda}_t^{-1}$ [6]. No exhaustive network-wide information exchange is applied since this would impose excessive complexity. As a reduced-complexity alternative, the converged APPs are exchanged among the adjacent BSs in the cooperative BS cluster only once, i.e., after the PDA detection was completed at each of the participating BSs. Furthermore, SC requires only a few simple arithmetic operations, as seen in (16). Hence, both the complexity and the backhaul traffic imposed by the associated message exchange and SC

[3]Equation (16) may also be interpreted as the sum of bit log-likelihood rates, where "multiplication" is converted to "addition" in the logarithmic domain.

remain modest, as will be demonstrated by the numerical complexity comparison results of Fig. 7 in Section IV. With respect to the backhaul traffic, in the entire reception process of a symbol vector, only $C_u M$ messages are passed from each cooperating BS to the others, and it can be further reduced by transferring the index of the quantized soft information, instead of the soft information itself. As will be shown in Fig. 4 of Section IV, the uniform quantization using even just a single bit imposes only a negligible performance loss.

IV. Simulation Results and Discussions

In this section, we characterize the achievable performance of the proposed DPDA-SC approach using Monte Carlo simulations in the hexagonal cellular network of Fig. 1 consisting of 19 cells. We use the decorrelated-model-based DPDA-SC for the scenario of $M_u\ (C_u+1) \leq M_b$ and the nondecorrelated-model-aided DPDA-SC for $M_u\ (C_u+1) > M_b$. Quadrature phase-shift keying (QPSK) modulation is used, and the knowledge of the average equivalent SNR per receive antenna formulated as $\text{SNR} \triangleq 10\log_{10}(E\{\|\mathbf{Gs}\|^2/M_b\}/N_0)$ is exploited at each BS. Flat Rayleigh-fading channels are considered, i.e., the entries of each subchannel matrix between an MS and a BS are chosen as independent and identically distributed (i.i.d.) zero-mean unit-variance complex-valued Gaussian random variables. A new realization of each channel matrix is drawn for each data burst consisting of 1000 transmitted symbol vectors, and each element of the noise vector $\mathbf{n}^k$ is i.i.d. $\mathcal{CN}(0, N_0)$. We set $(M_u = 2, M_b = 8)$ and $(M_u = 4, M_b = 8)$ for the scenarios of $M_u\ (C_u+1) \leq M_b$ corresponding to Section IV-A–C, E and F, and $M_u\ (C_u+1) > M_b$ corresponding to Section IV-D, respectively. For simplicity, we consider MS0 in the following investigation, where $C_u = 3$, without loss of generality. Since the PDA algorithm typically converges within three to five 5 iterations [6], we set the maximum number of iterations to $I = 5$.

A. Perfect CSI

In the case of *perfect CSI*, Fig. 2 compares the bit error ratio (BER) performance of nine different setups, including the PDA and the maximum-likelihood (ML) single-user bounds recorded at BS0 for MS0, where ρ represents the interference intensity $\rho^k_{\alpha_i}$ defined in Section III. To be specific, each $\rho^k_{\alpha_i}$ value may be different, but for simplicity, an identical interference intensity was assumed for the interfering MSs. This is justified because all the MSs imposing interference on each of the desired MSs are situated in the neighboring cells and have similar distances from the desired MS's home BS.

The "SCSU LZF" scheme refers to the linear zero-forcing (LZF)-based single-cell single-user detection (SUD) invoked at each BS, where the cochannel users' signals arriving from the other cells are simply treated as background noise. Naturally, this low-complexity SUD leads to poor performance. The "MMSE-OSIC JD," the "PDA JD," and the "ML JD" refer to the joint detection (JD) of multiple cochannel MSs at each BS using the minimum-mean-square-error-based successive interference cancellation with optimal ordering (MMSE-OSIC), the PDA, and the ML approaches, respectively, where again, no SC is invoked. The ML detector is implemented with the aid of a reduced-complexity sphere decoder [15], where the sphere radius is adaptively adjusted according to the prevalent SNR-level to avoid a search failure. All the JD setups require knowledge of all the channel matrices between the cochannel users and the local BS for mitigating the CCI, but they do not share soft decision information with other cells since no message exchange and no SC is used. Nonetheless, a substantial BER improvement is shown in comparison to the SCSU LZF, particularly when ρ is small.

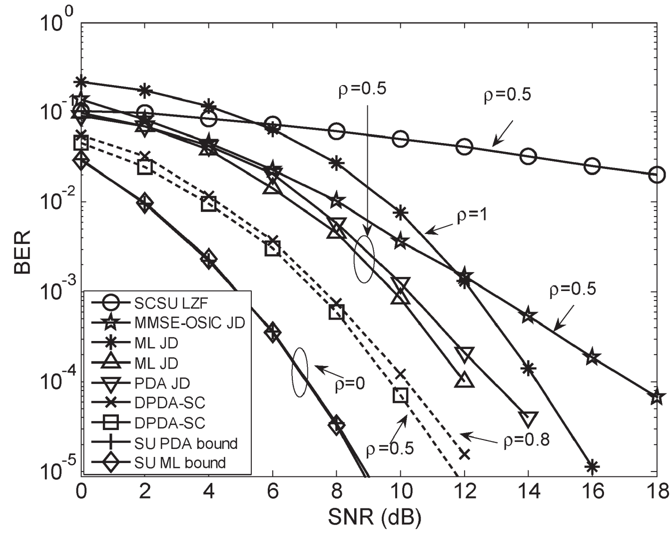

Fig. 2. Performance comparison of DPDA-SC, ML JD, SCSU LZF, SU ML bound, and SU PDA bound under perfect CSI and different interference intensities ρ, using QPSK modulation.

The dashed curves represent the proposed DPDA-SC scheme operating under $\rho = 0.5$ and $\rho = 0.8$. Observe in Fig. 2 that a significant further BER improvement is achieved, which is attributed to the spatial diversity gain provided by joint cooperative BS processing. The PDA and the ML single-user bounds, i.e., the "SU PDA bound" and "SU ML bound," are obtained by setting $\rho = 0$, which implies that the CCI vanishes. This scenario is equivalent to a single-user (2×8)-element spatial multiplexing MIMO system. It is observed in Fig. 2 that the PDA bound is extremely close to the ML bound. The results recorded in Fig. 2 for different ρ values characterize the impact of ρ on the attainable reception performance. It may be concluded from Fig. 2 that the interference intensity ρ is the key factor limiting the achievable performance of cellular MIMO networks.

B. Imperfect CSI

When considering the more practical *imperfect CSI* scenario, Fig. 3 compares the performance of the proposed DPDA-SC scheme to the MMSE-OSIC JD, the ML JD, and the PDA JD schemes at different levels of channel estimation quality of $\beta = 0.98$ and $\beta = 0.99$, and a given interference intensity of $\rho = 0.8$. It is observed in Fig. 3 that the achievable performance of all the schemes considered is dramatically degraded by the channel estimation error. Basically, different error floors are observed for these schemes because the fixed level of relatively strong interference plays a dominant role, when the SNR is high. More specifically, in the imperfect CSI scenario of Fig. 3, the PDA JD scheme only marginally outperforms the MMSE-OSIC JD scheme, although it enjoys a notable advantage in the perfect CSI scenario. In other words, the PDA JD scheme is more sensitive to the channel estimation error than the MMSE-OSIC JD scheme. This is because the MMSE-OSIC JD is a hard-decision method, whereas the PDA JD vitally relies on an iterative soft information update process. In general, the accuracy of the soft information is a key factor in determining the success of soft-information-based algorithms. Therefore, we can further observe in Fig. 3 that the DPDA-SC scheme has a substantial performance advantage over that of the PDA JD because the spatial diversity originating from BS cooperation using SC considerably improves the accuracy of the soft information. Compared to the ML JD, the DPDA-SC is remarkably superior in the moderate-SNR region of practical interest, although its advantage erodes in the high-SNR region of Fig. 3. This phenomenon is a consequence of

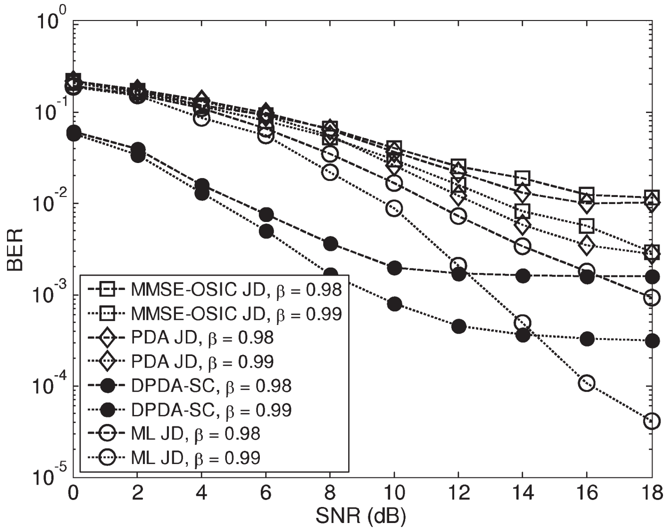

Fig. 3. Performance comparison of different reception schemes under different levels of channel estimation quality $\beta = 0.98$ and $\beta = 0.99$ and a given interference intensity $\rho = 0.8$, using QPSK modulation.

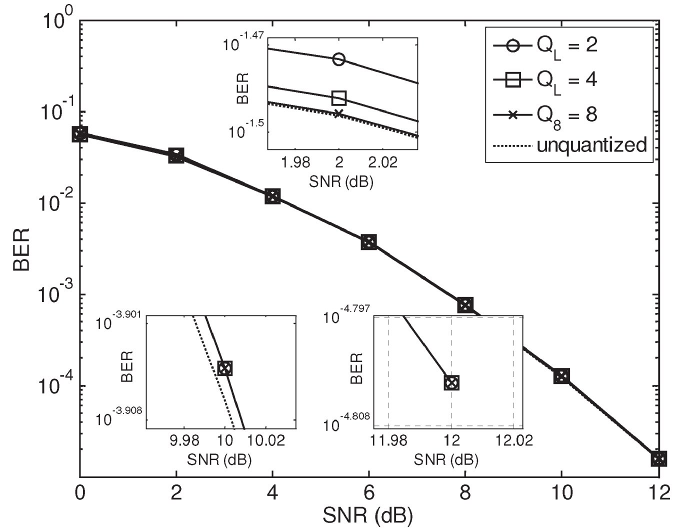

Fig. 4. Uniform quantization impact on the performance of DPDA-SC under perfect CSI and a given interference intensity $\rho = 0.8$, using QPSK modulation.

the different sensitivities of soft- and hard-information-based methods. Therefore, we conclude that the BS-cooperation-aided DPDA-SC scheme is capable of mitigating the effects of the error floor imposed by imperfect CSI by exploiting that the strong interference becomes a useful source of increased signal energy as a benefit of the more sophisticated distributed processing. Note that. although the error floor is not very obvious for the ML JD in the SNR range considered in Fig. 3, it is expected to become more evident at higher SNRs.

C. Impact of Quantization on the Backhaul Traffic and Performance

It is of practical interest to investigate the cost of backhaul traffic in the parallel message exchange stage using the UCS mode in the proposed DPDA-SC scheme. Instead of transferring the quantized probability value itself, the index of each probability value is transferred between the cooperative BSs, where a quantized probability lookup table is prestored. When considering the perfect CSI scenario, where the interference intensity is $\rho = 0.8$, Fig. 4 shows the performance of the DPDA-SC employing uniform quantization, which is performed on the converged probabilities before they are transferred to the cooperative BSs. It is shown that the performance loss due to the different quantization levels in uniform quantization is marginal in the proposed DPDA-SC scheme. This performance loss diminishes as the SNR increases and completely vanishes at high SNRs. This is because the converged soft information has been individually obtained using the PDA algorithm at each BS before the parallel message exchange stage, and the SC-aided final decision relies on which probability is the highest, rather than on the exact values of the probabilities themselves. By comparison, the soft information is obtained via multiple iterations between BSs in [3], where the performance is more sensitive to the quantization loss. Based on our results, if the $Q_L = 2$ uniform quantization scheme is used, only $M \log_2 Q_L = 4$ bits will be transferred from each BS to one of the cooperating BSs, when QPSK is used.

D. Rank-Deficient Scenario

Fig. 5 shows the performance of the proposed DPDA-SC scheme in the scenario of $M_u(C_u + 1) > M_b$, where $M_u = 4$, $M_b = 8$, $C_u = 3$. The nondecorrelated signal model is applied in both the PDA JD and the DPDA-SC schemes. We observed in Fig. 5 that the DPDA-SC is still superior to the PDA-JD, although both of them suffered significant performance loss due to the much stronger interlayer interference compared to the scenario of $M_u(C_u + 1) \leq M_b$.

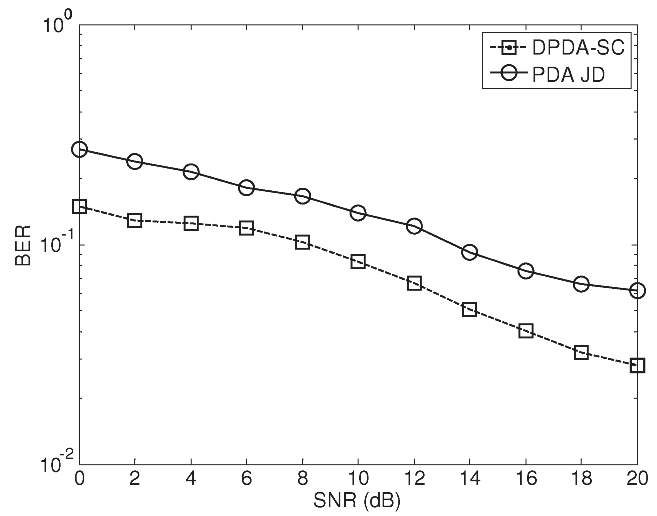

Fig. 5. Performance of DPDA-SC in rank-deficient scenario, $M_b = 8$, $M_u = 4$, $C_u = 3$, $C_b = 3$, with perfect CSI and a given interference intensity $\rho = 0.8$, using QPSK modulation.

E. Convergence Property

In Fig. 6, we characterize the convergence performance of the proposed DPDA-SC scheme under both perfect and imperfect CSI conditions. It may be observed in Fig. 6 that the DPDA-SC converges within a few iterations, which is a contributing factor of the low complexity of the DPDA-SC, as shown in Fig. 7.

F. Complexity Comparison

Finally, the complexity comparison between the DPDA-SC, the PDA JD, the MMSE-OSIC JD, and the ML JD is provided in Fig. 7. The complexity is quantified in terms of the number of equivalent additions required for decoding a single bit. We may observe that, as the number of concurrent transmissions increases, the complexity

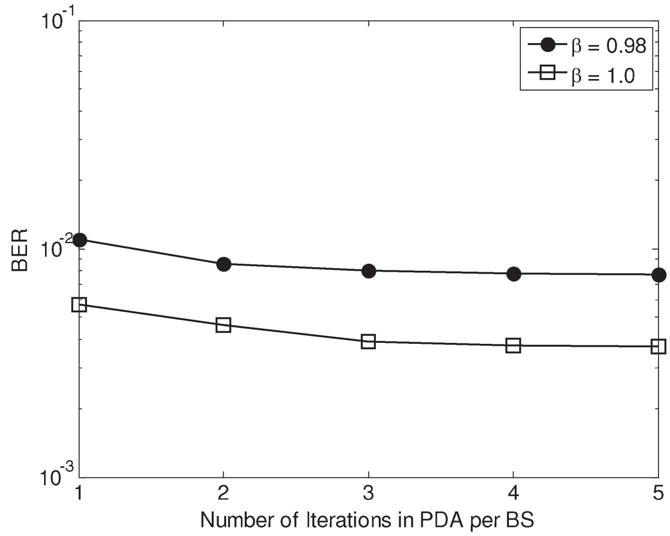

Fig. 6. Convergence property of DPDA-SC, SNR $= 6$ dB, QPSK, and $\rho = 0.8$.

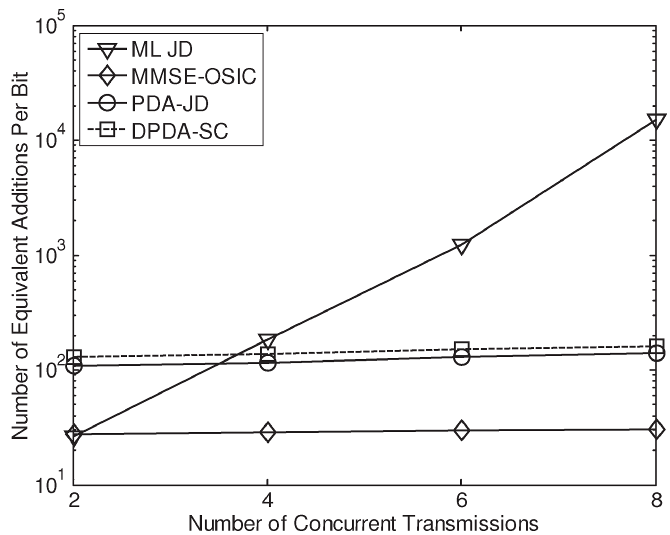

Fig. 7. Complexity comparison of different reception schemes, measured in terms of the number of equivalent additions per bit, SNR $= 6$ dB. QPSK, $\beta = 0.98$, and $\rho = 0.8$.

of ML JD rapidly increases, whereas the other three maintain a near-constant normalized complexity. Furthermore, the complexity of the proposed DPDA-SC is marginally higher (contributed by SC) than that of the conventional PDA JD but significantly lower than that of ML JD (about 1% when the number of concurrent transmissions is 8). Although the relatively coarse complexity analysis using $O(\cdot)$ function shows that the DPDA-SC approach enjoys the same $O(L^3)$ order of complexity as the SIC-based algorithm, the numerical results show that the DPDA-SC is about five times more complex than the MMSE-OSIC JD scheme.

V. Conclusion

We have proposed a DPDA-SC scheme for BS cooperation in the uplink of multiuser multicell MIMO systems. The realistic hexagonal cellular model relying on unity FR has been considered. The DPDA-SC scheme has been shown to converge in few iterations; hence, it constitutes a low-complexity solution for jointly estimating the initial soft-decision information at each BS. Each BS shares the MSs' soft information with the aid of their message exchange and generates the final soft decision information with the aid of SC. The simulation results, as well as our complexity analysis, demonstrate that the proposed scheme significantly outperforms the conventional noncooperative schemes while imposing a modest additional complexity and backhaul traffic. We have also investigated the performance of the proposed DPDA-SC scheme in the more practical imperfect CSI scenario and demonstrated that the DPDA-SC scheme succeeds in mitigating the system's error floor. The impact of quantization on both the backhaul traffic and the achievable performance has also been investigated, and it has been shown that, even when using single-bit uniform quantization, the performance loss is trivial while leading to low backhaul traffic.

References

[1] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO system with joint detection," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 442–453, Mar. 2004.
[2] T. Mayer, H. Jenkac, and J. Hagenauer, "Turbo base-station cooperation for intercell interference cancellation," in *Proc. IEEE ICC*, Jun. 2006, pp. 4977–4982.
[3] S. Khattak, W. Rave, and G. Fettweis (2008, Jan.). Distributed iterative multiuser detection through base station cooperation. *EURASIP J. Wireless Commun. Netw.* [Online]. *2008*, pp. 1–15. Available: http://dx.doi.org/10.1155/2008/390489
[4] E. Aktas, J. Evans, and S. Hanly, "Distributed decoding in a cellular multiple-access channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 241–250, Jan. 2008.
[5] Y. Bar-Shalon and X. R. Li, *Estimation and Tracking: Principles, Techniques and Software.*. Norwood, MA: Artech House, 1993.
[6] J. Luo, K. R. Pattipati, P. K. Willett, and F. Hasegawa, "Near optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 361–363, Sep. 2001.
[7] Y. Jia, C. M. Vithanage, C. Andrieu, and R. J. Piechocki, "Probabilistic data association for symbol detection in MIMO systems," *Electron. Lett.*, vol. 42, no. 1, pp. 38–40, Jan. 2006.
[8] S. Yang, T. Lv, R. G. Maunder, and L. Hanzo, "Unified bit-based probabilistic data association aided MIMO detection for high-order QAM constellations," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 981–991, Mar. 2011.
[9] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.
[10] J. Fricke, M. Sandell, J. Mietzner, and P. Hoeher, "Impact of the Gaussian approximation on the performance of the probabilistic data association MIMO decoder," *EURASIP J. Wireless Commun. Netw.*, vol. 5, no. 5, pp. 796–800, Oct. 2005.
[11] V. Annapureeddy and V. Veeravalli, "Gaussian interference networks: Sum capacity in the low interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.
[12] T. Weber, A. Sklavos, and M. Meurer, "Imperfect channel-state information in MIMO transmission," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 543–552, Mar. 2006.
[13] J. Luo, "Improved multiuser detection in code-division multiple access systems," Ph.D. dissertation, Univ. Connecticut, Storrs, CT, May, 2002.
[14] F. Cao, J. Li, and J. Yang, "On the relation between PDA and MMSE-ISDIC," *IEEE Signal Process. Lett.*, vol. 14, no. 9, pp. 597–600, Sep. 2007.
[15] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.