A Cooperative Channel Estimation Approach for Coordinated Multipoint Transmission Networks

Qianrui Li*†, David Gesbert†, Nicolas Gresset*
*Mitsubishi Electric R&D Centre Europe (MERCE), Rennes, France
Email: {q.li,n.gresset}@fr.merce.mee.com
†Mobile Communication Department, Eurecom, Sophia Antipolis, France
Email: {liqi,gesbert}@eurecom.fr

***Abstract*—This paper considers the problem of coordinated multi-point transmission from transmitter devices that acquire noisy and partial channel state information (CSI) estimates from local feedback links. As cooperative multi-antenna transmission often relies on the availability of global channel estimates, we propose a novel decentralized algorithm that produces MMSE-optimal channel estimates on the basis of combining local feedback and inter-transmitter communications. To this end, we assume the devices are equipped with rate-limited bi-directional communication links over which they exchange a finite number of CSI-related bits. We propose a low-complexity cooperative channel estimation algorithm which exploits the local communications near-optimally and is robust to arbitrary feedback noise statistics. The proposed method has application in future decentralized CoMP scenarios, where we show clear advantages over the conventional channel state information exchange mechanisms.**

I. Introduction

Coordinated multi-point (CoMP) transmission methods such as network MIMO [1] [2] [3], coordinated beamforming and scheduling [4] [5] [6], and interference alignment [7] often require the acquisition of global channel state information (CSI) at all the devices that engage in the coordination [8]. In practice, in Frequency Division Duplex (FDD) scenarios, CSI acquisition at the transmitter exploits local feedback links that provide at best a noisy and partial estimate of the total (network-wide) CSI matrix at each base station. In the current LTE release for instance, only the CSI related to a subset of the users served by one base station are acquired by that base station. In turn the cooperating transmitters can rely on signaling interfaces (such as X2 [9] alike interface in LTE) that provide limited inter-device communication capabilities so as to enable global channel estimates, hence facilitate cooperation. Currently, such links are subject to a coarse design however, limited to communicating the CSI coefficients that are missing at one transmitter, with no regard the *statistical quality* of the local information already existing at that transmitter, and ignoring potential benefits of *correlated channel estimates* between two transmitters, i.e, the fact that two base station may both obtain estimates of the CSI related to a common user.

In this paper we recast this problem into a more general and systematic *decentralized channel estimation problem with side information*. We start off with transmitters having acquired an initial amount of CSI from local feedback links. It is essential to note that this initial local CSI can be of completely arbitrary nature in our framework, including scenarios such as limited feedback from an arbitrary subset of users, with arbitrary estimation error statistics. Our framework also includes so-called hierarchical CSI scenarios where some transmitters are called to play a role of *master* in a cluster of small-cells, collecting more information than surrounding *slave* transmitters.

Based on the given initial local CSI structure, we design a two-step decentralized channel estimation aiming at producing minimum mean squared error (MMSE) optimal estimates at all cooperating transmitters. The new approach involves the transmitters (1) exchanging with each other *shaped* quantization representations of their local channel data, under a rate constraint on the inter-transmitter communication. Secondly, (2) the transmitters construct a final channel estimate based on a suitable combination of their local CSI and the data received from surrounding devices. Interestingly, the steps (1) and (2) are intertwined but fortunately can be solved jointly. Finally, when the inter-device communication links are bidirectional and total amount of bits for all coordination links are constrained, an interesting question is how many bits of channel-related information will flow in each directions. In other words which transmitters shall teach more to other transmitters (resp. learn more from them) about the channel state.

It has to be noted that the problem of cooperative channel estimation is rooted in the information theoretic framework of network vector quantization [10]. Specifically, for two transmitters cooperation, it is the problem of lossy source coding with side information (a.k.a. Wyner-Ziv coding) [11]. When more transmitters are involved in the cooperation, it is a generalization of Wyner-Ziv coding with the decoder takes into account multiple instances of correlated and encoded information as well as the side information. The information theoretical bound and asymptotically bound-achieving quantizer design for Wyner-Ziv coding are well analyzed, for example, in [12], [13], [14], [15], [16] and [17].

In this paper, we propose a novel algorithm (referred to as *coordination shaping*). For the two transmitters cooperation scenario, the proposed algorithm outperforms the asymptotic optimal Wyner-Ziv coding algorithm [17] in the low rate region and the performance asymptotically achieves the Wyner-Ziv bound in the high rate region. The proposed algorithm works for multi-transmitter cooperation scenarios as well, which are generalizations of Wyner-Ziv coding, to jointly optimize all the involved quantizers. Finally, our optimization framework allows to find the optimal bits allocation to each coordination links in the case when the total amount of bits for the coordination links are constrained.

More specifically, our contributions are as follows:

- We present a framework for cooperative decentralized channel estimation which adapts to scenarios with arbitrary noise and feedbacks, under a finite bit communication link between cooperators.
- We derive a low complexity optimal algorithm which determines what relevant information about the channel state is best to exchange between the devices. The solution takes the form of a vector quantizer based on a weighted distortion measure, where the weight is adapted to the CSI error covariance.
- We exhibit how to combine the exchanged limited CSI together with the local CSI. We use a weighted linear combination where the weights can be obtained from a provably convergent algorithm.
- In the case the communication between the cooperating devices is over a bi-directional link with a total amount of bits constraint over all coordination links, we present an algorithm that determines the optimal bits allocation across the links so as to minimize an average mean squared error (MSE) metric.

The rest of the paper is arranged as follows. Section II introduces the system model and describes the problem. Section III provides the optimization of the reconstruction function and Section IV performs the optimization of the quantizer. Section V solves the coordination link bit allocation problem and numeric result for both decentralized channel estimation and the coordination link bit allocation are presented in Section VI. Section VII concludes the paper.

II. System model and problem description

We consider a communication system where K transmitters (TX) such as small cell base stations communicate with L receivers (RX) over a network MIMO channel. Each transmitter TX$i, i = 1, \dots, K$ is equipped with M transmit antennas while each receiver RX$j, j = 1, \dots, L$ is equipped with N receive antennas. The propagation channel between TXi and RXj is denoted as $\mathbf{H}_{ji} \in \mathbb{C}^{N \times M}$.

We now assume that each transmitter acquires an initial estimate of the global channel state from any pilot-based, digital or analog feedback mechanism. The CSI made initially available at TXi is modeled as:

$$\begin{aligned} \hat{\mathbf{H}}^{(i)} &= \mathbf{H} + \mathbf{E}^{(i)} \\ \mathbf{H} &= \begin{bmatrix} \mathbf{H}_{11} & \dots & \mathbf{H}_{1K} \\ \vdots & \dots & \vdots \\ \mathbf{H}_{L1} & \dots & \mathbf{H}_{LK} \end{bmatrix} \end{aligned} \tag{1}$$

where $\hat{\mathbf{H}}^{(i)} \in \mathbb{C}^{NL \times MK}$ is an arbitrary CSI estimate for $\mathbf{H}$ at TXi. $\mathbf{H} \in \mathbb{C}^{NL \times MK}$ is the true CSI of the network MIMO channel and $\mathbf{E}^{(i)} \in \mathbb{C}^{NL \times MK}$ is the estimation error at TXi. For ease of exposition, we assume that the true network MIMO channel satisfies $\mathbf{h} = vec(\mathbf{H}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Hence, the actual channels are uncorrelated antenna-wise and user-wise, while the *estimates* at various TXs are correlated through $\mathbf{H}$. The channel independent estimation error $\mathbf{E}^{(i)}$ satisfies $\mathbf{e}^{(i)} = vec(\mathbf{E}^{(i)}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_i)$. The estimation errors for different TXs are assumed independent, i.e, $\mathbb{E}\{\mathbf{e}^{(i)} \cdot \mathbf{e}^{(j)^H}\} = \mathbf{0}, \forall i \neq j$.

Note that this CSI feedback model is quite general and includes diverse scenarios from *local feedback* to *global noisy feedback*. Representing various CSI information structure can be done by varying the statistics (i.e. $\mathbf{Q}_i$) at the various transmitters, since larger coefficients on some elements of $\mathbf{Q}_i$ induces a reduced information for the corresponding channels at TXi, inducing partial CSI. Note that some feedback designs may lead to each transmitter having some *noisy* information about *all* the user channels. For instance in the broadcast feedback scenario, each RX acquires its downlink CSI from all transmitters (using pilots), quantizes it, then broadcasts it on the uplink feedback channel. Since the link strength between RX and each TX varies for each TX-RX pair, each TX then receives a different noisy version of the same global CSI.

In Fig. 1, the cooperation information exchange between two transmitters TXi and TXk are illustrated. TXk send to TXi a suitably quantized version of his local CSI estimate $\hat{\mathbf{h}}^{(k)}$ through the coordination link, denoted here by $\mathbf{z}_{ki}$. Similar operation is performed by TXi to provide TXk $\mathbf{z}_{ik}$. Based on the local CSI and exchanged CSI, both transmitters compute a final estimate.

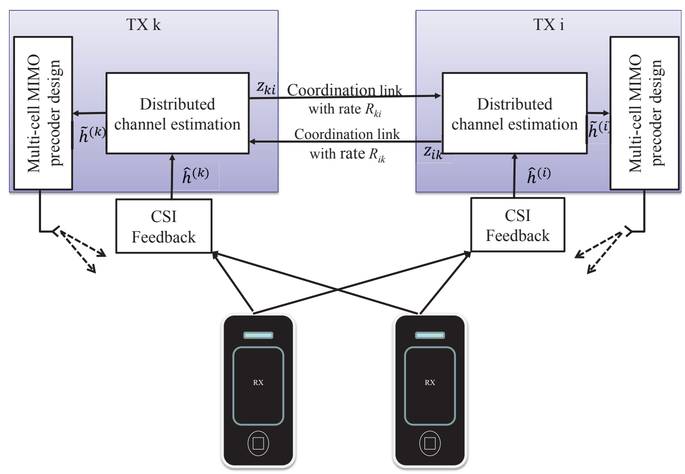

Fig. 1. Decentralized cooperative channel estimation across two base stations engaged in multi-cell MIMO downlink precoding.

The final estimate at TXi is denoted by $\tilde{\mathbf{h}}^{(i)}$, a reconstruction function $g_i(.)$ is used at TXi which combines local CSI $\hat{\mathbf{h}}^{(i)}$ and exchanged CSI $\mathbf{z}_{ki}$ to get $\tilde{\mathbf{h}}^{(i)}$. The quantization operation associated to the link from TXk to TXi is defined as $\mathcal{Q}_{ki} : \mathbb{C}^{NMKL \times 1} \mapsto \mathcal{C}_{ki}$, $\mathbf{z}_{ki} \in \mathcal{C}_{ki}$, $\mid \mathcal{C}_{ki} \mid = 2^{R_{ki}}$ where $\mathcal{C}_{ki}$ is the codebook for the quantizer $\mathcal{Q}_{ki}$. Finally R_{ki} is the finite number of bits on the coordination link between TXk and TXi. All cooperation information is exchanged in a single shot and simultaneously with other transmitters.

The goal of this paper is to find (i) the optimal reconstruction function $g_i(.)$ and (ii) the optimal fix rate quantizer $\mathcal{Q}_{ki}$ such that the expectation of the per dimensional MSE for the final channel estimate

$$D^{(i)} = \frac{1}{n}\mathbb{E}\{\| \mathbf{h} - \tilde{\mathbf{h}}^{(i)} \|^2\} \tag{2}$$

is minimized. Later on in this paper, we also optimize the values of R_{ki} under a global backhaul constraint.

Regarding the quantization on coordination link, the optimal vector quantization (VQ) is depicted in Fig. 2. $\mathbf{z}_{ki}$ and quantization error $\mathbf{e}_{\mathcal{Q}_{ki}}$ are shown to be uncorrelated, $\mathbf{e}_{\mathcal{Q}_{ki}}$ and $\hat{\mathbf{h}}^{(k)}$ are dependent [18]. The covariance matrices for $\hat{\mathbf{h}}^{(k)}$, $\mathbf{z}_{ki}$ and $\mathbf{e}_{\mathcal{Q}_{ki}}$ satisfy $\mathbf{Q}_{\hat{\mathbf{h}}^{(k)}} = \mathbf{Q}_{\mathbf{z}_{ki}} + \mathbf{Q}_{\mathcal{Q}_{ki}}$. Since the input of the quantizer $\hat{\mathbf{h}}^{(k)}$ is Gaussian, we obtain an upper bound of the impact of quantization by assuming that the quantization error $\mathbf{e}_{\mathcal{Q}_{ki}} = \hat{\mathbf{h}}^{(k)} - \mathbf{z}_{ki}$ is also Gaussian distributed as $\mathbf{e}_{\mathcal{Q}_{ki}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_{\mathcal{Q}_{ki}})$ [19]. Similar to [20] and based on the assumption for $\mathbf{e}_{\mathcal{Q}_{ki}}$ as Gaussian, we can approximate the VQ procedure by a gain-plus-additive-noise model (similar to the scalar quantizer case in [21]) shown in Fig.3.

Proposition 1. *Assume the quantization error $\mathbf{e}_{\mathcal{Q}_{ki}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_{\mathcal{Q}_{ki}})$, the optimal vector quantization procedure for $\hat{\mathbf{h}}^{(k)} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}+\mathbf{Q}_k)$, $\hat{\mathbf{h}}^{(k)} = \mathbf{e}_{\mathcal{Q}_{ki}} + \mathbf{z}_{ki}$ can be approximated by a gain-plus-additive-noise model with the quantization result $\mathbf{z}_{ki}$ satisfying:*

$$\mathbf{z}_{ki} = (\mathbf{I} + \mathbf{Q}_k - \mathbf{Q}_{\mathcal{Q}_{ki}})(\mathbf{I} + \mathbf{Q}_k)^{-1}\hat{\mathbf{h}}^{(k)} + \mathbf{q}_{ki} \tag{3}$$

where $\mathbf{q}_{ki}$ and $\hat{\mathbf{h}}^{(k)}$ are uncorrelated random vectors, $\mathbf{q}_{ki} \sim \mathcal{CN}(\mathbf{0}, (\mathbf{I} + \mathbf{Q}_k - \mathbf{Q}_{\mathcal{Q}_{ki}})(\mathbf{I} + \mathbf{Q}_k)^{-1}\mathbf{Q}_{\mathcal{Q}_{ki}})$.

Proof: This is just a trivial generalization of gain-plus-additive-noise model for optimal scalar quantizer in [21]. ■

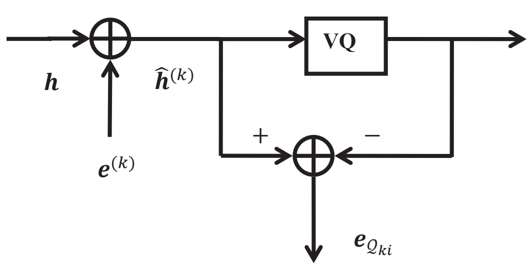

Fig. 2. Quantizer model for optimal vector quantization.

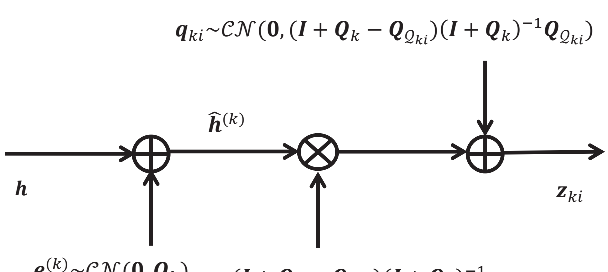

Fig. 3. Gain-plus-additive-noise model to approximate the optimal vector quantization procedure.

In the reminder of this work, both the design of reconstruction functions and optimal quantizers will be based on (3).

III. Reconstruction Function Design

This section addresses the aforementioned sub-problem of the optimal reconstruction function $g_i(.)$ design in general setting of multi-transmitter cooperation ($K \geq 2$). For the reason of simplicity, we only consider the reconstruction function as a weighted linear combination of estimates at TXi.

Therefore, the final estimate at TXi is:

$$\tilde{\mathbf{h}}^{(i)} = \sum_{k \in \mathcal{A}_i} \mathbf{W}_{ki}\mathbf{z}_{ki} + \mathbf{W}_{ii}\hat{\mathbf{h}}^{(i)} \tag{4}$$

where the set $\mathcal{A}_i$ contains the indices of TXs that are cooperating with TXi. Thus, according to (3), with equations (2) and (4) the optimal weight combining matrices can be derived.

Proposition 2. *The optimal weight combing matrices $\{\mathbf{W}^*_{ki}, \mathbf{W}^*_{ii}, k \in \mathcal{A}_i\}$ for the optimization problem:*

$$\begin{array}{cl} \min\limits_{\mathbf{W}_{ki}, \mathbf{W}_{ii}, k \in \mathcal{A}_i} & \frac{1}{n}\mathbb{E}\{\| \mathbf{h} - \tilde{\mathbf{h}}^{(i)} \|^2\} \\ s.t. & \tilde{\mathbf{h}}^{(i)} = \sum\limits_{k \in \mathcal{A}_i} \mathbf{W}_{ki}\mathbf{z}_{ki} + \mathbf{W}_{ii}\hat{\mathbf{h}}^{(i)} \end{array}$$

can be obtained from:

$$\left[\begin{array}{cccc} \mathbf{W}^*_{ii} & \mathbf{W}^*_{l_1 i} & \dots & \mathbf{W}^*_{l_{C_i} i} \end{array}\right] = \mathbf{\Upsilon}_i \mathbf{\Omega}_i^{-1}. \tag{5}$$

The optimum per dimensional MSE attained is:

$$D^{(i)*} = \frac{1}{n}tr\{\mathbf{I} + \mathbf{Q}_i^{-1} + \sum_{k \in \mathcal{A}_i} \left((\mathbf{I} + \mathbf{Q}_k)(\mathbf{I} + \mathbf{Q}_k - \mathbf{Q}_{\mathcal{Q}_{ki}})^{-1}(\mathbf{I} + \mathbf{Q}_k) - \mathbf{I}\right)^{-1}\}^{-1} \tag{6}$$

where $\mathbf{\Upsilon}_i, \mathbf{\Omega}_i$ are given below, the set $\mathcal{A}_i$ has cardinal $|\mathcal{A}_i| = C_i$ and is denoted by $\mathcal{A}_i = \{l_1, \dots, l_{C_i}\}$.

$$\begin{array}{rcl} \mathbf{\Upsilon}_i & = & \left[\begin{array}{cccc} \mathbf{I} & \mathbf{A}^H_{l_1 i} & \dots & \mathbf{A}^H_{l_{C_i} i} \end{array}\right] \\ \mathbf{\Omega}_i & = & \left[\begin{array}{cccc} \mathbf{I} + \mathbf{Q}_i & \mathbf{A}^H_{l_1 i} & \dots & \mathbf{A}^H_{l_{C_i} i} \\ \mathbf{A}_{l_1 i} & \mathbf{P}_{l_1 i} & \dots & \mathbf{A}_{l_1 i}\mathbf{A}^H_{l_{C_i} i} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{l_{C_i} i} & \mathbf{A}_{l_{C_i} i}\mathbf{A}^H_{l_1 i} & \dots & \mathbf{P}_{l_{C_i} i} \end{array}\right] \\ \mathbf{P}_{l_j i} & = & \mathbf{I} + \mathbf{Q}_{l_j} - \mathbf{Q}_{\mathcal{Q}_{l_j i}} \\ \mathbf{A}_{l_j i} & = & (\mathbf{I} + \mathbf{Q}_{l_j} - \mathbf{Q}_{\mathcal{Q}_{l_j i}})(\mathbf{I} + \mathbf{Q}_{l_j})^{-1}, j = 1 \dots C_i \end{array}$$

Proof: Due to lack of space, we only provide a sketch for the proof, the details are provided in the full size version of this paper [22]. According to (3), substitute $\mathbf{z}_{ki}$ into (4) and then (2), set the partial derivatives $\frac{\partial D^{(i)}}{\partial \mathbf{W}_{ii}} = 0, \frac{\partial D^{(i)}}{\partial \mathbf{W}_{ki}} = 0, k \in \mathcal{A}_i$ and solve this equation system, we can get (5). Insert (5) into (4) and then (2), after a few simplification we can get (6). ■

IV. Quantizer Design

Based on the weight expressions above, we can now proceed with the task of optimizing the quantizer used for communicating CSI-related bits from one transmitter to another. The main goal is to optimize $\mathbf{Q}_{\mathcal{Q}_{ki}}$ such that $D^{(i)*}$ can be minimized.

First, it's obvious that a conventional VQ will not provide the best $\mathbf{z}_{ki}$ as it typically minimizes the distortion in a way that ignores the local CSI available at TXi.

As $\mathbf{Q}_i, i = 1 \ldots K$ reflect the accuracy of local CSI and are assumed to be known to all TXs, the quantizer should allocate the quantization resource where most needed, i.e. in channel elements or directions that are known best by the TXk and least known by TXi. Thus, we propose to use the weighted square error distortion: $d_{\mathcal{Q}_{ki}}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^H \mathbf{B}_{ki}(\mathbf{x} - \mathbf{y})$ as the distortion measure of the quantizer $\mathcal{Q}_{ki}$, where the positive definite weight matrix $\mathbf{B}_{ki}$ is a variable to be optimized.

Importantly, we can calculate $\mathbf{Q}_{\mathcal{Q}_{ki}}$ in the asymptotic case when the given coordination link rate R_{ki} is sufficiently large, i.e. the large quantization rate regime.

Proposition 3. *For a quantizer with S levels and S is sufficiently large, the quantization error covariance $\mathbf{Q}_\mathbf{x}$ for a n dimensional complex random vector source $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Gamma})$ using optimal weighted mean square error distortion with weight matrix $\mathbf{B}$ satisfies:*

$$\mathbf{Q}_\mathbf{x} = 2S^{-\frac{1}{n}} M_n 2\pi (\frac{n+1}{n})^{n+1} \det(\mathbf{\Phi})^{\frac{1}{2n}} \mathbf{B}^{-1}$$

where M_n is a constant representing the minimal normalized second moment for all space-filling polytopes in $2n$ dimension [23], and $\mathbf{\Phi}$ is defined as:

$$\mathbf{\Phi} = \frac{1}{2} \begin{bmatrix} \Re(\mathbf{B}^{\frac{1}{2}} \mathbf{\Gamma} \mathbf{B}^{\frac{H}{2}}) & \Im(\mathbf{B}^{\frac{1}{2}} \mathbf{\Gamma} \mathbf{B}^{\frac{H}{2}}) \\ \Im(\mathbf{B}^{\frac{1}{2}} \mathbf{\Gamma} \mathbf{B}^{\frac{H}{2}}) & \Re(\mathbf{B}^{\frac{1}{2}} \mathbf{\Gamma} \mathbf{B}^{\frac{H}{2}}) \end{bmatrix}.$$

Proof: See Appendix. ∎

we now exploit Proposition 3 in order to derive $\mathbf{Q}_{\mathcal{Q}_{ki}}$:

$$\mathbf{Q}_{\mathcal{Q}_{ki}} = 2^{1-\frac{R_{ki}}{n}} M_n 2\pi \left(\frac{n+1}{n}\right)^{n+1} \det(\mathbf{\Phi}_{ki})^{\frac{1}{2n}} \mathbf{B}_{ki}^{-1} \quad (7)$$

where

$$\mathbf{\Phi}_{ki} = \frac{1}{2} \begin{bmatrix} \Re(\mathbf{B}_{ki}^{\frac{1}{2}}(\mathbf{I}+\mathbf{Q}_k)\mathbf{B}_{ki}^{\frac{H}{2}}) & \Im(\mathbf{B}_{ki}^{\frac{1}{2}}(\mathbf{I}+\mathbf{Q}_k)\mathbf{B}_{ki}^{\frac{H}{2}}) \\ \Im(\mathbf{B}_{ki}^{\frac{1}{2}}(\mathbf{I}+\mathbf{Q}_k)\mathbf{B}_{ki}^{\frac{H}{2}}) & \Re(\mathbf{B}_{ki}^{\frac{1}{2}}(\mathbf{I}+\mathbf{Q}_k)\mathbf{B}_{ki}^{\frac{H}{2}}) \end{bmatrix}$$

According to (6), (7), we can now optimize the quantization taking side information into account by solely optimizing the value of $\mathbf{B}_{ki}$. When R_{ki} is sufficiently large, the optimal $\mathbf{B}_{ki}$ can be solved as follows:

$$\begin{array}{ll} \min\limits_{\mathbf{B}_{ki}, k\in\mathcal{A}_i} & D^{(i)*} \\ s.t. & \det(\mathbf{B}_{ki}) = 1, \mathbf{B}_{ki} \succeq 0 \\ & D^{(i)*} \text{ defined in (6)} \end{array} \quad (8)$$

The constraints on $\mathbf{B}_{ki}$ matrices ensure that the optimization is performed among quantizers with the same quantization distortion. Note that a rigorous convexity analysis for optimization problem (8) is left for [22], only an approximation is provided here. Since we are considering the coordination link rate R_{ki} to be sufficiently large, use matrix inverse approximation on (6):

$$\begin{aligned} D^{(i)*} &= \frac{1}{n} tr\{\mathbf{I} + \mathbf{Q}_i^{-1} \\ &\quad + \sum_{k\in\mathcal{A}_i} \left((\mathbf{I}+\mathbf{Q}_k)(\mathbf{I}+\mathbf{Q}_k - \mathbf{Q}_{\mathcal{Q}_{ki}})^{-1}(\mathbf{I}+\mathbf{Q}_k) - \mathbf{I}\right)^{-1}\}^{-1} \\ &\simeq \frac{1}{n} tr\{\sum_{k\in\mathcal{A}_i} (\mathbf{Q}_k + \mathbf{Q}_{\mathcal{Q}_{ki}})^{-1} + \mathbf{Q}_i^{-1} + \mathbf{I}\}^{-1} \\ &\simeq \frac{1}{n} tr\{\sum_{k\in\mathcal{A}_i} \mathbf{Q}_k^{-1} + \mathbf{Q}_i^{-1} + \mathbf{I} - \sum_{k\in\mathcal{A}_i} \mathbf{Q}_k^{-1}\mathbf{Q}_{\mathcal{Q}_{ki}}\mathbf{Q}_k^{-1}\}^{-1} \end{aligned}$$

Therefore, the objective function is approximated as a trace of the inverse of positive semi-definite matrix, which is a convex optimization problem.

Note that once the optimal weight matrix $\mathbf{B}_{ki}^*$ is obtained, the codebook for optimal quantizer $\mathcal{Q}_{ki}^*$ can be calculated based on Lloyd algorithm and a training set. The optimal weight matrices for estimation combine can be calculated according to (5).

V. Coordination link bit allocation problem

An interesting consequence of the above analysis is the optimization of coordination where multiple transmitters can exchange simultaneously CSI-related information to each other under a global constraint on the coordination bits. The global optimization problem over all coordination links now becomes:

$$\begin{array}{ll} \min\limits_{\substack{\mathbf{B}_{ki}, R_{ki} \\ k\in\mathcal{A}_i, i=1,\ldots,K}} & \frac{1}{K}\sum\limits_{i=1}^{K} D^{(i)*} \\ s.t. & \det(\mathbf{B}_{ki}) = 1, \mathbf{B}_{ki} \succeq 0 \\ & \sum\limits_{i=1}^{K}\sum\limits_{k\in\mathcal{A}_i} R_{ki} = R_{tot}, R_{ki} \in \mathbb{N}^+ \\ & D^{(i)*} \text{ defined in (6).} \end{array}$$

Due to the integer constraints on R_{ki}, this problem becomes non-convex optimization. However, conventional alternating algorithms can be applied to perform the optimization in a two-step iterative approach: (i) initialize $\mathbf{B}_{ki}$, (ii) optimize the coordination bit R_{ki} with fixed $\mathbf{B}_{ki}$, (iii) optimize the weighting matrices $\mathbf{B}_{ki}$ with fixed R_{ki}, (iv) iterate between (ii) and (iii) until converges. The optimization in (ii) is an integer programming and the optimization in (iii) is a convex optimization, therefore, many conventional algorithms can be applied in both steps. Note that the optimum is not guaranteed by the alternating algorithm, it might converge to a local optimal point as well.

VI. Numerical performance analysis

In this section, the per dimensional MSE for decentralized channel estimation is evaluated for different settings using Monte-Carlo simulations over 10^5 channel realizations. In all simulations $K = 2$ and $L = 2$, $M = N = 1$. The channel $\mathbf{h} \in \mathbb{C}^{4\times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ and the rates on coordination link from TX2 to TX1 and from TX1 to TX2 are denoted R_{21} and R_{12}, respectively. The parameter M_n is set as $929/12960$ which is related to the $E8$ lattice since we are considering 8 real dimensions. In Fig. 4 and Fig. 5, the local CSI information structure is characterized by $\mathbf{Q}_1 = diag(0.1, 0.1, 0.9, 0.9)$ and

$\mathbf{Q}_2 = (0.9, 0.9, 0.1, 0.1)$ which corresponds to an example where TX1 has more accurate CSI about RX1 and less accurate CSI about RX2, while TX2 has more accurate CSI about RX2 and less accurate CSI about RX1. The $diag(.)$ operator represents a diagonal matrix with diagonal elements in the parenthesis.

Fig.4 shows the per dimensional MSE for the final estimation at TX1. The shaped coordination curve applies the proposed algorithm. The unshaped coordination implements the traditional optimal VQ and finds $\mathbf{W}_{21}, \mathbf{W}_{11}$ accordingly using (5). From the figure we can conclude that the shaped coordination algorithm outperforms the unshaped coordination algorithm, which not surprisingly shows the benefit of taking the priori statistic information into account. The asymptotically optimal Wyner-Ziv coding curve refers to the Wyner-Ziv quantizer for noisy source in [17]. It reveals that our algorithm outperforms the asymptotically optimal Wyner-Ziv coding algorithm in low coordination rate region. In fact, in two TXs cooperation case, it is trivial to prove that the proposed algorithm converges asymptotically to the Wyner-Ziv bound D_∞ [17] as well.

Fig. 5 exhibits the sum rate for a 2 TX cooperation system. The rate on the coordination links satisfies $R_{21} = R_{12}$. The zero forcing (ZF) precoder is constructed at each TX based on its final estimate of the network MIMO channel. We also provide the sum rate for the case when coordination links have infinite bandwidth. The figure shows that the proposed shaped coordination algorithm will improve the system sum rate beyond the unshaped coordination algorithm and asymptotically optimal Wyner-Ziv coding algorithm when a simple ZF precoder is implemented. As the rate on coordination link increases, the sum rate for all algorithms will converge to the infinite coordination rate case.

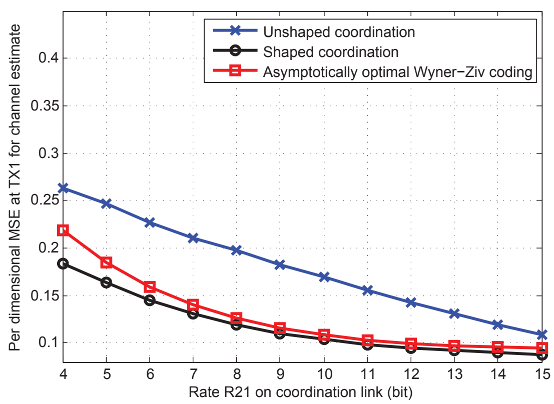

Fig. 4. Mean square error for the final channel estimation at TX1 as a function of coordination link rate R_{21}, local CSI error covariance matrices are $\mathbf{Q}_1 = diag(0.1, 0.1, 0.9, 0.9)$, $\mathbf{Q}_2 = diag(0.9, 0.9, 0.1, 0.1)$.

Fig.6 considers the coordination link bit allocation problem for a 2 TX cooperation system. The total amount of bits for coordination link is $R_{12} + R_{21} = 30$bits, the error covariance matrices $\mathbf{Q}_1 = diag(0.1, 0.1, 0.9, 0.9), \mathbf{Q}_2 = diag(0.5, 0.5, 0.5, 0.5)$. The figure reveals that the cooperation information exchange is not necessarily symmetric, the optimal coordination link bit allocation strategy is to let TX2 share the cooperation information to TX1 through a $R_{21} = 12$bits

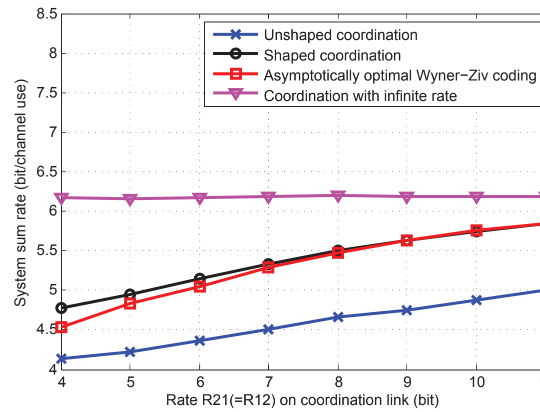

Fig. 5. Sum rate for the 2 TX cooperation system, each TX implements a ZF precoder based on its final channel estimation, SNR is 20dB per TX, local CSI error covariance matrices are $\mathbf{Q}_1 = diag(0.1, 0.1, 0.9, 0.9)$, $\mathbf{Q}_2 = diag(0.9, 0.9, 0.1, 0.1)$.

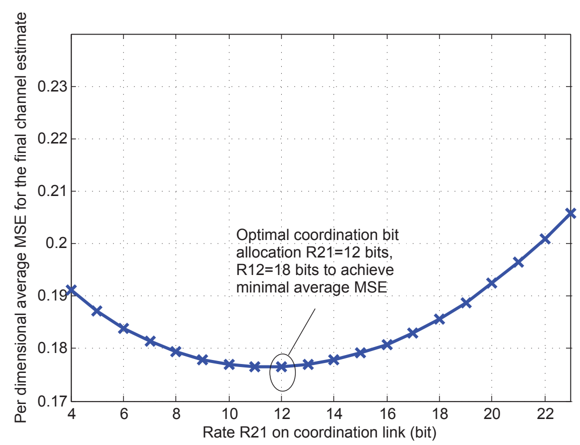

Fig. 6. Per dimensional average mean square error for the final channel estimation at TX1, TX2 when sum bit for coordination link $R_{tot} = R_{21} + R_{12} = 30$bits, $\mathbf{Q}_1 = diag(0.1, 0.1, 0.9, 0.9)$, $\mathbf{Q}_2 = diag(0.5, 0.5, 0.5, 0.5)$.

coordination link and vice versa through a $R_{12} = 18$bits coordination link. It reveals that if one TX has a better local CSI, it is more encouraged to share his information through a higher rate coordination link.

VII. Conclusion

We study the decentralized cooperative channel estimation for coordinated multipoint transmission networks. We derive a low-complexity algorithm which exploits local communications near-optimally and is robust to arbitrary feedback noise statistics. We exhibit clear advantages over CSI acquisition and exploitation methods used in conventional CoMP systems.

VIII. Appendix

Proof for Proposition 3: In order to prove this proposition, we need the following result: for a random vector $\mathbf{x}$, if an optimal S levels Euclidean distance distortion based quantizer applied on a random vector $\mathbf{y} = \mathbf{B}^{\frac{1}{2}}\mathbf{x}$ has the codebook $\{\mathbf{y}_1, \ldots, \mathbf{y}_S\}$ and associated partition $\{\mathcal{P}_1, \ldots, \mathcal{P}_S\}$, then the optimal S level weighted square error distortion based quantizer applied on the random vector $\mathbf{x}$ with weight matrix $\mathbf{B}$

will have the codebook $\{\mathbf{B}^{-\frac{1}{2}}\mathbf{y}_1, \ldots, \mathbf{B}^{-\frac{1}{2}}\mathbf{y}_S\}$ and associated partition $\{\mathbf{B}^{-\frac{1}{2}}[\mathcal{P}_1], \ldots, \mathbf{B}^{-\frac{1}{2}}[\mathcal{P}_S]\}$,where the $\mathbf{B}^{-\frac{1}{2}}[\mathcal{P}_i]$ is defined as $\mathbf{B}^{-\frac{1}{2}}[\mathcal{P}_i] = \{\mathbf{x} : \exists \mathbf{y} \in \mathcal{P}_i \text{ st. } \mathbf{x} = \mathbf{B}^{-\frac{1}{2}}\mathbf{y}\}$. This result can be easily obtained by change of variables in the integral expression of the average quantizer distortion. Thus,

$$\begin{aligned}
\mathbf{Q}_\mathbf{x} &= \mathbb{E}\{(\mathbf{x} - \mathcal{Q}(\mathbf{x}))(\mathbf{x} - \mathcal{Q}(\mathbf{x}))^H\} \\
&= \mathbf{B}^{-\frac{1}{2}}\mathbb{E}\{(\mathbf{y} - \mathcal{Q}(\mathbf{y}))(\mathbf{y} - \mathcal{Q}(\mathbf{y}))^H\}\mathbf{B}^{-\frac{H}{2}} \\
&= \mathbf{B}^{-\frac{1}{2}}\mathbf{Q}_\mathbf{y}\mathbf{B}^{-\frac{H}{2}},
\end{aligned}$$

where $\mathbf{y} = \mathbf{B}^{\frac{1}{2}}\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{B}^{\frac{1}{2}}\mathbf{\Gamma}\mathbf{B}^{\frac{H}{2}})$. Let $\mathbf{t} = [\Re(\mathbf{y})^T \Im(\mathbf{y})^T]^T$, then $\mathbf{t} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Phi})$ and

$$\mathbf{\Phi} = \frac{1}{2}\left[\begin{array}{cc} \Re(\mathbf{B}^{\frac{1}{2}}\mathbf{\Gamma}\mathbf{B}^{\frac{H}{2}}) & \Im(\mathbf{B}^{\frac{1}{2}}\mathbf{\Gamma}\mathbf{B}^{\frac{H}{2}}) \\ \Im(\mathbf{B}^{\frac{1}{2}}\mathbf{\Gamma}\mathbf{B}^{\frac{H}{2}}) & \Re(\mathbf{B}^{\frac{1}{2}}\mathbf{\Gamma}\mathbf{B}^{\frac{H}{2}}) \end{array}\right].$$

Furthermore, it is proven in [24] that

$$\mathbf{Q}_\mathbf{t} = \mathbb{E}\{(\mathbf{t} - \mathcal{Q}(\mathbf{t}))(\mathbf{t} - \mathcal{Q}(\mathbf{t}))^T\} = D_\mathbf{t}\mathbf{I}_n,$$

where the average distortion $D_\mathbf{t} = \frac{1}{n}tr(\mathbf{Q}_\mathbf{t})$ is obtained from [25] for large S as

$$\begin{aligned}
D_\mathbf{t} &= S^{-\frac{2}{n}} M_n \left(\int f_\mathbf{t}(\mathbf{t})^{\frac{n}{n+2}} d\mathbf{t}\right)^{\frac{n+2}{n}} \\
&= S^{-\frac{2}{n}} M_n 2\pi \left(\frac{n+2}{n}\right)^{\frac{n}{2}+1} \det(\mathbf{\Phi})^{\frac{1}{n}}
\end{aligned}$$

where $f_\mathbf{t}(.)$ is the probability density function (p.d.f) of $\mathbf{t}$. Finally, from the expression of $\mathbf{Q}_\mathbf{t}$ and by a real-to-complex conversion, we get

$$\mathbf{Q}_\mathbf{y} = 2S^{-\frac{1}{n}} M_n 2\pi (\frac{n+1}{n})^{n+1} \det(\mathbf{\Phi})^{\frac{1}{2n}} \mathbf{I}_n$$

which leads to

$$\mathbf{Q}_\mathbf{x} = 2S^{-\frac{1}{n}} M_n 2\pi (\frac{n+1}{n})^{n+1} \det(\mathbf{\Phi})^{\frac{1}{2n}} \mathbf{B}^{-1}.$$

REFERENCES

[1] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1380–1408, 2010.

[2] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on*, 2007, pp. 83–87.

[3] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *Wireless Communications, IEEE*, vol. 13, no. 4, pp. 56–61, 2006.

[4] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell multi-antenna wireless system," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 5, pp. 1748–1759, 2010.

[5] R. Zakhour and D. Gesbert, "Distributed multicell-MISO precoding using the layered virtual SINR framework," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 8, pp. 2444–2448, 2010.

[6] E. Bjornson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *Signal Processing, IEEE Transactions on*, vol. 60, no. 5, pp. 2508–2523, 2012.

[7] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the-user interference channel," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3425–3441, 2008.

[8] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: from theory to practice*. Cambridge University Press, 2011.

[9] S. Sesia, I. Toufik, and M. Baker, *LTE: the UMTS long term evolution*. Wiley Online Library, 2009.

[10] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," *Information Theory, IEEE Transactions on*, vol. 50, no. 8, pp. 1584–1604, 2004.

[11] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *Information Theory, IEEE Transactions on*, vol. 22, no. 1, pp. 1–10, 1976.

[12] H. Yamamoto, "Source coding theory for multiterminal communication systems with a remote source," *IEICE TRANSACTIONS (1976-1990)*, vol. 63, no. 10, pp. 700–706, 1980.

[13] S. C. Draper, "Successive structuring of source coding algorithms for data fusion, buffering, and distribution in networks," Ph.D. dissertation, Massachusetts Institute of Technology, 2002.

[14] D. Rebollo-Monedero and B. Girod, "Generalization of the rate-distortion function for wyner-ziv coding of noisy sources in the quadratic-gaussian case," in *Data Compression Conference, 2005. Proceedings. DCC 2005*. IEEE, 2005, pp. 23–32.

[15] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1250–1276, 2002.

[16] Z. Liu, S. Cheng, A. D. Liveris, and Z. Xiong, "Slepian-wolf coded nested lattice quantization for wyner-ziv coding: high-rate performance analysis and code design," *Information Theory, IEEE Transactions on*, vol. 52, no. 10, pp. 4358–4379, 2006.

[17] D. Rebollo-Monedero, S. Rane, and B. Girod, "Wyner-ziv quantization and transform coding of noisy sources at high rates," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, vol. 2. IEEE, 2004, pp. 2084–2088.

[18] I. Jee and R. Haddaad, "Optimum design of vector-quantized subband codecs," *Signal Processing, IEEE Transactions on*, vol. 46, no. 8, pp. 2239–2243, Aug 1998.

[19] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[20] P. Marsch and G. Fettweis, "On downlink network mimo under a constrained backhaul and imperfect channel knowledge," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, Nov 2009, pp. 1–6.

[21] P. H. Westerink, J. Biemond, and D. E. Boekee, "Scalar quantization error analysis for image subband coding using qmfs," *Signal Processing, IEEE Transactions on*, vol. 40, no. 2, pp. 421–428, 1992.

[22] Q. Li, D. Gesbert, and N. Gresset, "A cooperative channel estimation approach for coordinated multipoint transmission networks," *in preparation*.

[23] J. Conway and N. Sloane, "Voronoi regions of lattices, second moments of polytopes, and quantization," *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 211–226, Mar 1982.

[24] R. Zamir and M. Feder, "On lattice quantization noise," *Information Theory, IEEE Transactions on*, vol. 42, no. 4, pp. 1152–1159, Jul 1996.

[25] P. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 139–149, Mar 1982.