Power Minimization in Wireless Network Virtualization with Massive MIMO

Mohammadmoein Soltanizadeh†, Ben Liang†, Gary Boudreau*, and S. Hossein Seyedmehdi*
†Dept. of Electrical and Computer Engineering, University of Toronto, Canada
*Ericsson Canada

***Abstract*—This paper presents a new method of wireless network virtualization to share the downlink of a massive-MIMO base station among multiple Service Providers (SPs). The SPs are allowed to simultaneously utilize all antennas and channel resource of the Infrastructure Provider (InP). This can improve resource utilization but also requires the InP to control the interference between SPs that are oblivious of each other. We develop novel precoding schemes to minimize the InP's transmission power while satisfying certain prescribed maximum deviation between each SP's intended signal to its users and what the InP delivers. This problem is studied for both perfect and imperfect channel state information (CSI). Under perfect CSI, the proposed precoding is optimal and substantially outperforms a time-sharing alternative. Under imperfect CSI, we use a numerical lower bound to show that the proposed precoding is nearly optimal.**

I. Introduction

The capital and operational expenses of wide-area wireless networks discourage service providers from deploying modern technologies and also hinder new companies from entering the industry. In response, the concept of virtualization has been proposed to reduce the expenses of network deployment and operation by abstracting and sharing physical resources, and to make it easier to migrate to newer products and technologies by decoupling distinct parts of the network. Virtualization creates a set of logical entities from a given set of physical entities in a manner that is transparent to users [1]. By enabling abstraction and sharing of the physical resources, it maximizes utilization of the resources while providing the required quality of service to users and enforcing the isolation and security that users need.

A virtualized network is generally composed of Infrastructure Providers (InPs) that create and manage the infrastructure, and Service Providers (SPs) that utilize the resources to provide services to their subscribing users. An InP virtualizes the resources that it owns and splits them into slices. These slices consist of (virtualized) core networks and (virtualized) access networks corresponding to the wired slice and the wireless slice, respectively. The SPs lease these virtual resources, and operate them to provide end-to-end services to end-users without needing to know the underlying physical architecture of the InP.

Wireless network virtualization has been studied mainly under two categories in the literature. The first category focuses on resource allocation, spectrum partitioning, and enforcing fairness among users, e.g. [2], [3], while the second category studies how virtualization can be applied to different wireless network technologies e.g., [4], [5], [6], and [7]. Although these studies show the importance and potential of wireless network virtualization, there is scarce publication that relates wireless network virtualization to massive MIMO. The authors of [8] and [9] studied resource provisioning in the virtualization of massive MIMO networks. In [8], different SPs employ disjoint subsets of a base station's antennas, which can lead to substantial loss in data rate compared with complete sharing of all antennas. Furthermore, their system model does not concern the design details of virtualization such as precoding or scheduling. In [9], the InP uses a hierarchical auction to allocate separate sub-channels to different SPs. However, as we will show later, restricting the SPs to orthogonal channels can also lead to inefficient resource utilization.

In this work, we allow antenna and channel sharing among multiple SPs in a novel wireless network virtualization architecture. We show that, through proper precoding, the InP can enforce the requirements of isolation by leveraging massive MIMO beamforming, while improving the overall efficacy of the network. Although precoding has been well studied in wireless communication, new challenges arise when it comes to wireless network virtualization. Since each SP does not have access to the channel information of the users of the other SPs, handling the interference between the users of different SPs is challenging. For instance, if the SPs use typical schemes of precoding and have the InP use their precoding matrices without considering the other SPs, the system will likely incur a large amount of interference. Our goal is to design methods for the InPs to manage this interference, such that the users of each SP receive nearly the same signal that their own SP has designed for them. This problem has not been treated in traditional precoding.

The main contributions of this paper are summarized below:

- We first formulate the precoding problem in wireless network virtualization when perfect CSI is available, to minimize the InP's transmission power while satisfying certain prescribed maximum deviation between each SP's intended signal to its users and what the InP delivers. We observe that the problem is convex and derive a closed-form expression for its dual problem and the sub-gradient of the dual function.
- We then consider the precoding problem under imperfect

This work has been funded in part by Ericsson Canada and by the Natural Sciences and Engineering Research Council of Canada.

CSI using chance constraints, which is a non-convex optimization problem. Despite the general intractability of such a problem, we demonstrate that, under massive MIMO, there is a simple and effective solution. We further derive a lower bound to the optimal objective for performance benchmarking.

- We conduct simulation study on the performance of the proposed solutions. Under perfect CSI, we observe that the obtained optimal precoding substantially outperforms a time-sharing alternative. Under imperfect CSI, we assume a practical MMSE channel estimation scheme and show numerically that the proposed solution gives nearly optimal performance.

The rest of this paper is organized as follows. In Section II, we state the system model and problem formulation for the perfect and imperfect CSI cases. In Section III, we present an optimal solution for perfect CSI, and in Section IV we present the proposed solution for imperfect CSI. Performance evaluation and discussion are given in Section V. Finally, Section VI concludes the paper along with a possible extension to multi-cell networks.

Notation: We use $\|\cdot\|$ to denote either the two-norm of a vector or the Frobenius norm of a matrix. We use $(\cdot)^T$ and $(\cdot)^H$ to denote matrix transpose and Hermitian transpose, respectively.

II. System Model

For convenience and without loss of generality, we assume that all the components in a virtualized wireless network are encompassed into two entities: InP and SPs. The SPs are responsible for serving the subscribers and programming their services, and the InP owns the infrastructure and is responsible for executing virtualization. We assume that the other parts of the network, including the core network and computational resources, are already virtualized and can be utilized by the InP and the SPs.

A. Precoding Design by InP and SPs

We consider the downlink in each cell of a cellular network, comprised of an InP that owns a base station with N antennas, and M SPs with their own schedulers and precoders. The SPs share the same wireless spectrum provided by the InP. The SPs are able to design their desirable precoding matrix for their own users. Denote by K_m the number of users of SP m, $\mathbf{H}_m \in \mathbb{C}^{K_m \times N}$ the flat fading channel of the users of SP m from the base station, and $\mathbf{W}_m$ the desirable precoding matrix designed by SP m. For example, in Section V, for the purpose of illustration, we will assume that the Regularized Zero Forcing (RZF) scheme is used by SP m to compute $\mathbf{W}_m$ for $m \in \{1, 2, \dots, M\}$.

The InP is responsible for virtualizing the base station in a manner transparent to the users. For our purpose, this means that it designs a new precoding matrix, $\mathbf{V}$, based on the inputs from the SPs, such that the users of SP m receive signals with a defined quality of service that reflects the SP's choice of $\mathbf{W}_m$, despite SP m is ignorant of the interfering SPs. It is worth noting that the InP cannot simply transmit using a simple aggregation of the precoding matrices designed by the SPs, since that would lead to severe inter-SP interference.

In the basic scenario, each SP may utilize all N antennas at the InP's base station. However, in this work, we also allow a more general scenario where each SP m only has information about the channels from its users to a subset of antennas, which is of size $N_m \leq N$. For example, one reason for this may be to reduce the amount of channel information exchange. In such a case, we assume that N_m and the subset of antennas are predetermined, and denote by $\mathbf{G}_m$ the part of $\mathbf{H}_m$ that corresponds to only the subset of antennas.

An overview of the overall operation of the virtualized network is as follows. In the time domain, transmissions are divided into defined periods, which can be permanent or dynamic on a frame or subframe basis. In each transmission period, firstly, to each SP m, the InP communicates the corresponding channel information $\mathbf{G}_m$. Then, the SPs design the precoding matrices $\mathbf{W}_m$ for their users and communicate them to the InP. Finally, the InP collects all $\mathbf{W}_m$ and designs precoding matrix $\mathbf{V}$ to transmit all users' messages.

In addition to executing virtualization, and satisfying constraints that come from its requirement, the InP is concerned with additional internal constraints such as total transmit power. In this work, we aim to minimize the InP's transmission power, which can be represented by $\|\mathbf{V}\|^2$.

B. Constraint Formulation for the Perfect CSI Case

In this idealized case, the channel $\mathbf{H}_m$ is exactly known by the InP, and $\mathbf{G}_m$ is precisely communicated to SP m for all m over the transmission period under consideration. Let $\mathbf{x}_m \in \mathbb{C}^{K_m}$ represent the messages for the users of SP m, and define $\mathbf{x}$ as

$$\mathbf{x} = [\mathbf{x}_1^T, \dots, \mathbf{x}_M^T]^T. \tag{1}$$

Without loss of generality, assume that the messages are zero-mean, uncorrelated, and normalized to 1, i.e.,

$$\mathbb{E}\,\mathbf{x}_m = 0. \tag{2}$$

$$\mathbb{E}\mathbf{x}_m^H \mathbf{x}_n = \begin{cases} 0 & \text{if } m \neq n, \\ 1 & \text{if } m = n. \end{cases}$$

Let $\mathbf{V} = [\mathbf{V}_1, \dots, \mathbf{V}_M]$ be the precoding matrix designed by the InP, where $\mathbf{V}_m$ corresponds to the users of SP m. Then, the users of SP m receive

$$\mathbf{y}_m = \mathbf{H}_m \mathbf{V}_m \mathbf{x}_m + \sum_{i \neq m} \mathbf{H}_m \mathbf{V}_i \mathbf{x}_i. \tag{3}$$

Note that the second part of the above sum represents the interference to the users of SP m, from the signals intended for the users of the other SPs. The precoding matrix $\mathbf{V}$ should be designed in a way that the received signal in (3) does not deviate significantly from the signal that SP m expects its users to receive, i.e.,

$$\mathbf{y'}_m = \mathbf{G}_m \mathbf{W}_m \mathbf{x}_m. \tag{4}$$

Formally, for each SP m, the InP should satisfy the following inequality:

$$\mathbb{E}_{\mathbf{x}}\|\mathbf{y}_m - \mathbf{y'}_m\|^2 \leq \mathcal{I}_m^2. \quad (5)$$

where $\mathcal{I}_m$ is a predefined threshold, which is assumed to be known a priori as a part of the contractual agreement between the SP and the InP. By substituting (3) and (4), this can be re-written as

$$\mathbb{E}_{\mathbf{x}}\|\mathbf{H}_m\mathbf{V}\mathbf{x} - \mathbf{G}_m\mathbf{W}_m\mathbf{x}_m\|^2 \leq \mathcal{I}_m^2. \quad (6)$$

Furthermore, by taking the expected value, we have

$$\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m}\|\mathbf{H}_m\mathbf{V}_i\|^2 \leq \mathcal{I}_m^2. \quad (7)$$

C. Constraint Formulation for the Imperfect CSI Case

The assumption that the channel is perfectly known may be far from reality due to factors such as channel estimation error and pilot contamination. In order to take into account channel uncertainty, we adopt the model

$$\mathbf{H}_m = \hat{\mathbf{H}}_m + \mathbf{E}_m \quad (8)$$

where $\mathbf{H}_m$ is the unknown true channel of users of SP m, $\hat{\mathbf{H}}_m$ is the estimated channel, and $\mathbf{E}_m$ is the channel estimation error such that $\text{vec}(\mathbf{E}_m) = \mathcal{CN}(\mathbf{0}, \mathbf{C}_e)$, where $\mathbf{C}_e$ is a covariance matrix. We assume that $\hat{\mathbf{H}}_m$ is independent of $\mathbf{E}_m$, which for example is the case under the common MMSE channel estimation scheme as presented in Section V-B.

In this case, we redefine $\mathbf{G}_m$ according to $\hat{\mathbf{H}}_m$ and assume that it is precisely communicated to SP m. Thus, the left-hand side of (7) becomes a random variable. We thus re-express this constraint in probabilistic terms as follows:

$$\mathbb{P}\{\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m}\|\mathbf{H}_m\mathbf{V}_i\|^2 \leq \mathcal{I}_m^2\} \geq 1-\epsilon. \quad (9)$$

where ϵ is a small positive number that determines the uncertainty of meeting the deviation threshold $\mathcal{I}_m$.

III. Power Minimization for Perfect CSI

As explained in Section II-A, we focus on power minimization as the objective of the InP subject to the constraints given in (7), i.e., the InP needs to solve the optimization problem given by

$$\min_{\mathbf{V}} \quad \|\mathbf{V}\|^2 \quad (10)$$

$$\text{s.t.} \quad \|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m}\|\mathbf{H}_m\mathbf{V}_i\|^2 \leq \mathcal{I}_m^2 \quad \forall m. \quad (11)$$

The following lemma provides a sufficient condition for the feasibility of the problem and also describes a property of the optimal solution. We omitted its proof due to space limitation.

Lemma 1. *If $\sum_m K_m \leq N$, the problem (10)-(11) is always feasible, and all constraints are active at optimum.*

When perfect channel information is available, the optimization problem given in (10)-(11) is a convex program. Instead of directly using an existing numerical solver on this problem, in the following, we derive closed-form expressions for the dual function and its sub-gradient, to be used in a Lagrangian solution approach.

For any Lagrange multipliers $\boldsymbol{\lambda}$, The Lagrange function of the problem is

$$L(\mathbf{V},\boldsymbol{\lambda}) = \|\mathbf{V}\|^2 + \sum_m \lambda_m(\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m}\|\mathbf{H}_m\mathbf{V}_i\|^2 - \mathcal{I}_m^2). \quad (12)$$

By taking derivative with respect to $\mathbf{V}_m^*$ and setting it to zero we have

$$\frac{\partial L}{\partial \mathbf{V}_m^*} = \mathbf{V}_m + \sum_i \lambda_i \mathbf{H}_i^H\mathbf{H}_i\mathbf{V}_m - \mathbf{H}_m^H\mathbf{G}_m\mathbf{W}_m = 0. \quad (13)$$

Therefore,

$$\mathbf{V}_m^* = (\mathbf{I} + \sum_i \lambda_i\mathbf{H}_i^H\mathbf{H}_i)^{-1}\mathbf{H}_m^H\mathbf{G}_m\mathbf{W}_m. \quad (14)$$

Defining $\mathbf{V}^*(\boldsymbol{\lambda}) = [\mathbf{V}_1^*, \ldots, \mathbf{V}_M^*]$, the dual function is

$$g(\boldsymbol{\lambda}) = L(\boldsymbol{\lambda}, \mathbf{V}^*(\boldsymbol{\lambda})). \quad (15)$$

Then the subgradient method to maximize $g(\boldsymbol{\lambda})$ requires the following result, whose proof is omitted tue to space limitation.

Lemma 2. *The following choice of $\mathbf{d}$ is a sub-gradient for the dual function $g(\boldsymbol{\lambda})$:*

$$\mathbf{d} = (d_1, \ldots, d_M)^T \quad (16)$$

$$d_i = (\|\mathbf{H}_m\mathbf{V}_m^* - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m}\|\mathbf{H}_m\mathbf{V}_i^*\|^2) - \mathcal{I}_m^2 \quad (17)$$

The rest of the iterative updating method to solve (10)-(11) is standard and hence is omitted.

IV. Power Minimization for Imperfect CSI

Under imperfect channel information, the precoding problem is formulated by

$$\min_{\mathbf{V}} \quad \|\mathbf{V}\|^2 \quad (18)$$

$$\text{s.t.} \quad \mathbb{P}\Big\{\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m}\|\mathbf{H}_m\mathbf{V}_i\|^2 \leq \mathcal{I}_m^2\Big\} \geq 1-\epsilon \quad \forall m. \quad (19)$$

Optimization problems with chance constraints are generally hard to solve. However, in the following we present a solution to this problem that is computationally efficient, and as will be shown in Section V, has nearly optimal performance.

We first present a lemma which provides a formula to derive the probability in (19):

Lemma 3. *Given a Hermitian matrix* $\mathbf{A}$ *and complex vector* $\mathbf{b}$*, for circular complex Gaussian random vector* $\mathbf{e}$ *distributed as* $\mathcal{CN}(\mathbf{0}, \mathbf{C}_e)$*, the CDF of* $\|\mathbf{Ae}+\mathbf{b}\|^2$ *is given by*

$$\mathbf{P}(\|\mathbf{Ae}+\mathbf{b}\|^2 \le \tau) = \int_{-\infty}^{\infty} \frac{e^{\tau(\beta+j\omega)}}{(\beta+j\omega)} \frac{e^{-c_0(\omega)+c(\omega)}}{\prod_i (1+(\beta+j\omega)D_i^2)} d\omega \tag{20}$$

for any $\beta > 0$*, where*

$$c_0(\omega) = (\beta+j\omega)\mathbf{b}^H\mathbf{b}, \tag{21}$$

$$c(\omega) = (\beta+j\omega)^2 \sum_i \frac{|a_i|^2 D_i^2}{1+(\beta+j\omega)D_i^2}, \tag{22}$$

$$\mathbf{a} = \mathbf{U}^H\mathbf{b}, \tag{23}$$

$$\mathbf{UDQ} = \mathrm{svd}(\mathbf{AC}_e). \tag{24}$$

The proof of this lemma is inspired by [10], in which the authors provided a closed-form expression for the CDF of $\|\mathbf{e}-\mathbf{b}\|^2_{\mathbf{Q}}$ for a Hermitian matrix $\mathbf{Q}$. We follow a slightly different approach to obtain the CDF of $\|\mathbf{Ae}+\mathbf{b}\|^2$. We omit the detailed proof due to space limitation.

To invoke Lemma 3 for our purpose, we transform the left-hand side of (7) to vector form as follows:

$$\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m} \|\mathbf{H}_m\mathbf{V}_i\|^2 \tag{25}$$

$$=\|\mathbf{H}_m\mathbf{V} - \mathbf{G}_m[\mathbf{0},\ldots,\mathbf{W}_m,\ldots,\mathbf{0}]\|^2 \tag{26}$$

$$=\|(\hat{\mathbf{H}}_m+\mathbf{E}_m)\mathbf{V} - \mathbf{G}_m[\mathbf{0},\ldots,\mathbf{W}_m,\ldots,\mathbf{0}]\|^2 \tag{27}$$

$$=\|\mathbf{E}_m\mathbf{V} + \hat{\mathbf{H}}_m\mathbf{V} - \mathbf{G}_m[\mathbf{0},\ldots,\mathbf{W}_m,\ldots,\mathbf{0}]\|^2 \tag{28}$$

$$=\|\mathbf{Ae}+\mathbf{b}\|^2, \tag{29}$$

where

$$\mathbf{A} = \mathbf{V}^T \otimes \mathbf{I}_{K_m} \tag{30}$$

$$\mathbf{e}_m = \mathrm{vec}(\mathbf{E}_m) \tag{31}$$

$$\mathbf{b}_m = \mathrm{vec}(\hat{\mathbf{H}}_m\mathbf{V} - \mathbf{G}_m[\mathbf{0},\ldots,\mathbf{W}_m,\ldots,\mathbf{0}]). \tag{32}$$

Furthermore, we have the following lemma, whose proof is given in Appendix A.

Lemma 4. *The quantity expressed in (29) can be written as the sum of* KN *independent random variables, where* $K = \sum_m K_m$*.*

Hence, assuming that Lyapunov's Central Limit Theorem conditions are satisfied, since KN is a large number in the massive MIMO setting, (29) is approximately Gaussian.

Then, it is easy to show that, for $\epsilon \le 0.5$, the condition

$$\mathbb{P}\{\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m} \|\mathbf{H}_m\mathbf{V}_i\|^2 \le \mathcal{I}_m^2\} \ge 1-\epsilon \tag{33}$$

implies that

$$\mathbb{E}\{\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m} \|\mathbf{H}_m\mathbf{V}_i\|^2\} \le \mathcal{I}_m^2. \tag{34}$$

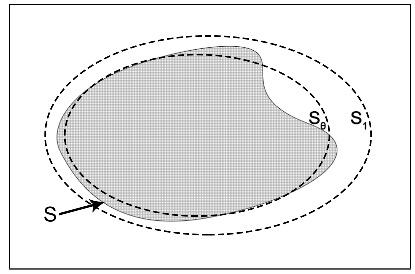

Fig. 1. The feasible sets of the the original and bounding problems.

Therefore, the following convex optimization problem provides a lower bound to problem (18)-(19):

$$\min_{\mathbf{V}} \quad \|\mathbf{V}\| \tag{35}$$

$$\text{s.t.} \quad \mathbb{E}\Big\{\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m} \|\mathbf{H}_m\mathbf{V}_i\|^2\Big\} \le \mathcal{I}_m^2 \quad \forall m \tag{36}$$

We may further simplify this problem by noting that the expected value in (36) can be re-written as

$$\|\mathbf{DV}\|^2 + \|\hat{\mathbf{H}}_m\mathbf{V} - \mathbf{G}_m[\mathbf{0},\ldots,\mathbf{W}_m,\ldots,\mathbf{0}]\|^2, \tag{37}$$

where $\mathbf{D}$ is a diagonal matrix, diag$(D_1,\ldots,D_N)$, with

$$D_i = \sqrt{\sum_i \sigma_{i,j}^2} \tag{38}$$

and $\sigma_{i,j}^2 = \sum_j \mathbb{E}|e_{ij}|^2$. The derivation of of the above result is omitted due to page limitation.

The proposed solution is based on problem (35)-(36). We describe the rationale of our solution by Figure 1. Let shaded area S be the feasibility set of problem (18)-(19), and let us define the following optimization problem for some positive θ:

$$P_\theta: \quad \min_{\mathbf{V}} \quad \|\mathbf{V}\| \tag{39}$$

$$\text{s.t.} \quad \mathbb{E}\Big\{\|\mathbf{H}_m\mathbf{V}_m - \mathbf{G}_m\mathbf{W}_m\|^2 + \sum_{i\neq m} \|\mathbf{H}_m\mathbf{V}_i\|^2\Big\} \le (\theta\mathcal{I}_m)^2 \quad \forall m. \tag{40}$$

Let areas S_θ be the feasibility set of problem P_θ. From the lower bound provided earlier, we have $S \subset S_1$. At each iteration, we shrink the feasibility set of P_θ by decreasing θ, and solve P_θ. The solution to P_θ is then checked by the formula given in (20) to verify whether the solution belongs to S. If so, then the algorithm ends; otherwise we shrink S_θ further to find another solution. This procedure continues until either a solution is found or P_θ becomes infeasible. It should be noted that P_θ is a convex programming problem and hence can be solved in polynomial time.

V. Performance Evaluation

We consider a hexagon cell with radius $R_c = 500m$ where the users of the SPs are uniformly distributed across the cell.

Our network consists of 4 SPs, each with 15 users, and an InP with a base station with $N = 100$ antennas placed at the center of the cell. We assume that all antennas are used by each SP, i.e., $N_m = 100$ for all m.

The baseband fading channel that links the base station to the kth user is modeled by [11]

$$\mathbf{h}_k = \beta_k^{1/2}\mathbf{g}_k \quad \forall k \tag{41}$$

$$\beta_k[\text{dB}] = -31.54 - 37.1\log_{10}(d_k) - 8\psi_k \quad \forall k \tag{42}$$

where $\mathbf{g}_k$, distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, models small-scale fading, β_k captures path loss and shadowing effects, d_k is the distance of user k from the BS, and ψ_k is a standard Gaussian random variable that accounts for large-scale shadowing.

As an example for illustration, we assume that RZF precoding is used by the SPs. This precoding is obtained by

$$\mathbf{W}_m = \alpha\mathbf{G}_m^H(\mathbf{G}_m\mathbf{G}_m^H + \frac{K_m}{\sigma_n^2}\mathbf{I})^{-1} \tag{43}$$

where σ_n is the noise power, and α is the normalizing factor and is set such that $\|\mathbf{W}_m\| = \frac{P}{\sqrt{M}}$, where P is the maximum amount of total power that the InP can dedicate to all the SPs.

We define deviation factors, ρ_m, and use them to set values of $\mathcal{I}_m$ by

$$\mathcal{I}_m = \rho_m\|\mathbf{H}_m\mathbf{W}_m\|. \tag{44}$$

The deviation factors can be interpreted as a received-signal proportional measure on how much the InP's precoding is allowed to deviate from the SP's intended precoding. To reduce the number of system variables in simulation, we assume $\rho_m = \rho$ for all m. Other default system parameters are P=8 dBm, and σ_n^2=-174 dBm/Hz, and we set the channel bandwidth to 10 KHz.

A. Performance under Perfect CSI

Figure 2 displays the norm of $\mathbf{V}$, which corresponds to the square root of transmission power, with respect to deviation factors ρ and N. It shows that InP can satisfy any value of ρ but at the expense of more transmission power. For the sake of comparison, we also plot the transmission power required for a time-sharing scheme where the InP serves the SPs one-by-one in equal time slots. We see that in most regions it is preferable to deploy the proposed scheme rather than simple sharing in time. This implies similar advantage of the proposed scheme over other orthogonal channel sharing schemes such as [9].

B. Performance under Imperfect CSI

For the purpose of illustration, we assume that the CSI is obtained through reverse channel estimation in TDD such that the users transmit L pilots with some given power P_{pilot} one by one, and the base station estimates the channel. The minimum mean square error (MMSE) channel estimate, $\hat{\mathbf{h}}_k$, of user k's channel, $\mathbf{h}_k$, is obtained by

$$\hat{\mathbf{h}}_k = \mathbf{C}_{\mathbf{h}_k}^T\mathbf{T}^T(\mathbf{T}\mathbf{C}_{\mathbf{h}_k}^T\mathbf{T}^T + \sigma_n^2\mathbf{I})^{-1}[\mathbf{y}_k^1; \ldots; \mathbf{y}_k^T] \tag{45}$$

where $\mathbf{C}_{\mathbf{h}_k} = \mathbb{E}\mathbf{h}_k\mathbf{h}_k^H$ and $\mathbf{T} = \sqrt{P_{\text{pilot}}}[\mathbf{I}, \ldots, \mathbf{I}]^T$. The error of this estimation can be modeled as $\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{e}_k$ where $\mathbf{e}_k$

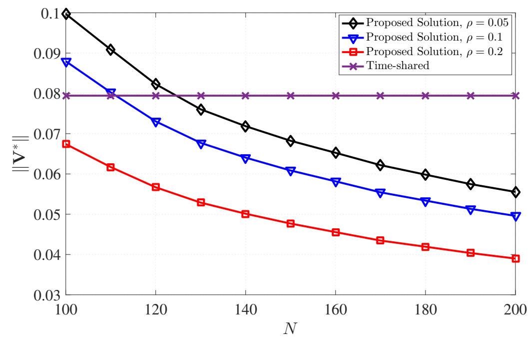

Fig. 2. Norm of $\mathbf{V}^*$ for perfect CSI.

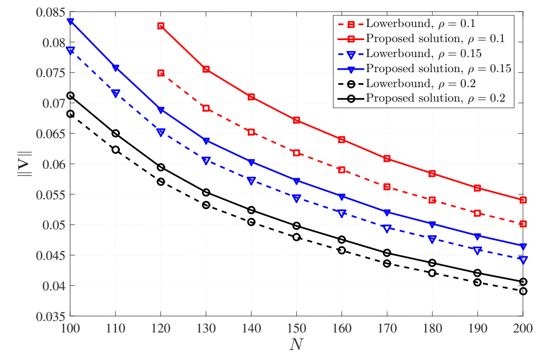

Fig. 3. Norm of $\mathbf{V}^*$ vs. N for imperfect CSI.

is independent of the estimated channel and is Gaussian with covariance given by

$$\mathbf{C}_e = \mathbf{C}_{\mathbf{h}_k} - \mathbf{C}_{\mathbf{h}_k}\mathbf{T}^T(\mathbf{T}\mathbf{C}_{\mathbf{h}_k}\mathbf{T} + \sigma_n^2\mathbf{I})^{-1}\mathbf{T}\mathbf{C}_{\mathbf{h}_k}. \tag{46}$$

We use the following default values: P_{pilot}=20dBm, ϵ=0.1, and L=5.

Figures 3 and 4 display the norm of $\mathbf{V}$ with respect to N and ρ, respectively. We also plot the lower bound as defined in Section IV for comparison. The most important observation is that the proposed solution is close to the lower bound, which suggests that the proposed solution is nearly optimal. These figures further suggests that using a large number of antennas can drastically reduce the InP's power consumption in the proposed virtualization scheme despite channel estimation error. Finally, we observe that as the value of ρ decreases, the norm of $\mathbf{V}$ increases until a point where the problem becomes infeasible. Indeed, having a larger number of antennas benefits the InP by giving a feasible solution for smaller values of ρ.

Figure 5 demonstrates the maximum channel error for which the problem is feasible for different values of ρ. Furthermore, it depicts an upper bound for the feasible solutions, obtained from the lower bound given in Section IV. This figure illustrates that our proposed algorithm almost always provides a solution when a feasible solution exists.

VI. Conclusion and Discussion

We have investigated virtualization in wireless networks with massive MIMO. To the best of our knowledge, this is

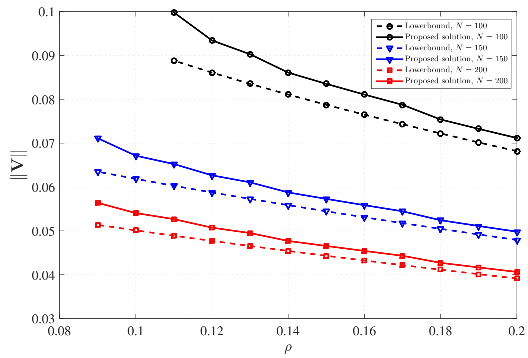

Fig. 4. Norm of $\mathbf{V}^*$ vs. ρ for imperfect CSI.

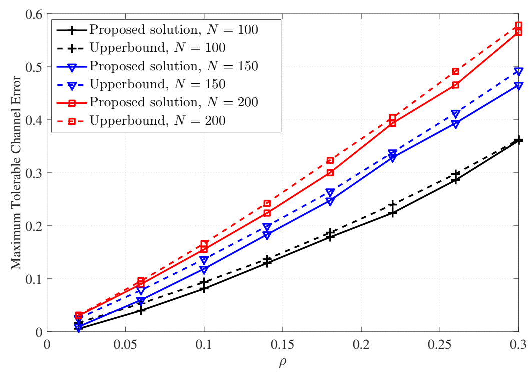

Fig. 5. Maximum tolerable channel error vs. ρ.

the first work to provide a formal formulation for virtualization of massive MIMO that considers the impact of precoding on interference control among SPs. We formulate the precoding problem under perfect and imperfect CSI scenarios, and propose simple but effective solutions for each. Simulation results demonstrate that the deviation of user received signal, between what is separately planned by each SP and what the InP delivers, can be small, while allowing the InP to reduce its energy consumption, especially when the InP possesses a large number of antennas.

Finally, we remark on how the proposed methods can be extended to a multi-cell network where each BS maintains its interference to the neighbouring cells under a certain threshold. Let $\mathbf{H}_0$ be the channel between the BS and the users of neighbouring cells. The interference to these users should be below a certain value, or formally,

$$\|\mathbf{H}_0\mathbf{V}\|_F \leq \mathcal{I}_0. \tag{47}$$

This constraint has the same form as the constraints for SP m but with $\mathbf{W}_m = 0$. Therefore, adding constraints to suppress the interference in the neighbouring cells leads to an optimization problem with the same form as (10)-(11) or (18)-(19), except for a larger number of constraints [12].

Appendix

A. Proof of Lemma 4

For each m, we can re-write the deviation into a vector as expressed in (29), where random vector $\mathbf{e}_m$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{C}_e)$. For notational simplicity, for the rest of this proof, we rescale both $\mathbf{A}$ as

$$\mathbf{A} = (\mathbf{V}^T \otimes \mathbf{I})\mathbf{C}_e^{1/2} \tag{48}$$

and $\mathbf{e}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{KN})$.

Let $[\mathbf{U}, \mathbf{D}, \mathbf{Q}] = \text{svd}(\mathbf{A})$ be the singular value decomposition of $\mathbf{A}$. Then from (29) we have

$$\begin{aligned}\|\mathbf{A}\mathbf{e}_m + \mathbf{b}\|^2 &= \|\mathbf{U}\mathbf{D}\mathbf{Q}\mathbf{e}_m + \mathbf{b}\|^2 && (49)\\ &= \|\mathbf{D}\mathbf{Q}\mathbf{e}_m + \mathbf{U}^H\mathbf{b}\|^2 && (50)\\ &= \|\mathbf{D}\mathbf{e}'_m + \mathbf{a}\|^2 && (51)\end{aligned}$$

where

$$\begin{aligned}\mathbf{e}'_m &= \mathbf{Q}\mathbf{e}_m && (52)\\ \mathbf{a} &= \mathbf{U}^H\mathbf{b}. && (53)\end{aligned}$$

We note that

$$\|\mathbf{D}\mathbf{e}'_m + \mathbf{a}\|^2 = \sum_{i=1}^{KN} |D_i e'^i_m + a_i|^2 \tag{54}$$

where D_i is the ith element on the diagonal of $\mathbf{D}$ and $\mathbf{e}'^i_m$ and a_i are the ith elements of vectors $\mathbf{e}'_m$ and $\mathbf{a}_m$, respectively. Furthermore, since $\mathbf{Q}$ is a unitary matrix, we have $\mathbf{e}'_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{KN})$, i.e., e'^i_m are independent over i. Furthermore, D_i and a_i are independent of the estimation error $\mathbf{E}_m$, since $\hat{\mathbf{H}}_m$ is independent of $\mathbf{E}_m$. Therefore, the summands in (54) are independent.

References

[1] X. Wang, P. Krishnamurthy, and D. Tipper, "Wireless network virtualization," in *Proc. International Conference on Computing, Networking and Communications (ICNC)*, Jan 2013.

[2] Y. Zhou, Y. Li, G. Sun, D. Jin, L. Su, and L. Zeng, "Game theory based bandwidth allocation scheme for network virtualization," in *Proc. IEEE Global Telecommunications Conf. (Globecom)*, Dec 2010.

[3] M. Yang, Y. Li, L. Zeng, D. Jin, and L. Su, "Karnaugh-map like online embedding algorithm of wireless virtualization," in *Proc. Intl. Symp. Wireless Personal Multimedia Communications*, Sep 2012.

[4] Y. Zaki, L. Zhao, C. Goerg, and A. Timm-Giel, "LTE mobile network virtualization," *Mobile Networks and Applications*, vol. 16, no. 4, pp. 424–432, Aug 2011.

[5] A. Banchs, P. Serrano, P. Patras, and M. Natkaniec, "Providing throughput and fairness guarantees in virtualized WLANs through control theory," *Mobile Networks and Applications*, vol. 17, no. 4, pp. 435–446, Aug 2012.

[6] G. Bhanage, D. Vete, I. Seskar, and D. Raychaudhuri, "SplitAP: leveraging wireless network virtualization for flexible sharing of WLANs," in *Proc. IEEE Global Telecommunications Conf. (Globecom)*, Dec 2010.

[7] D. Yun and Y. Yi, "Virtual network embedding in wireless multihop networks," in *Proc. Intl. Conf. Future Internet Technologies*, Jun 2011.

[8] V. Jumba, S. Parsaeefard, M. Derakhshani, and T. Le-Ngoc, "Resource provisioning in wireless virtualized networks via massive-MIMO," *IEEE Wireless Communications Letters*, vol. 4, no. 3, pp. 237–240, June 2015.

[9] K. Zhu and E. Hossain, "Virtualization of 5G cellular networks as a hierarchical combinatorial auction," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2640–2654, Oct 2016.

[10] T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite quadratic forms in Gaussian random variables," in *Proc. IEEE Intl. Symp. Info. Theory (ISIT)*, Jul 2009.

[11] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.

[12] M. Soltanizadeh, "Power Minimization in Wireless Network Virtualization with Massive MIMO," Master's thesis, University of Toronto, 2017.