Space–Time Power Schedule for Distributed MIMO Links Without Instantaneous Channel State Information at the Transmitting Nodes

Yue Rong, *Member, IEEE*, Yingbo Hua, *Fellow, IEEE*, Ananthram Swami, *Fellow, IEEE*, and A. Lee Swindlehurst, *Fellow, IEEE*

Abstract—A space–time optimal power schedule for multiple distributed multiple-input multiple-output (MIMO) links without the knowledge of the instantaneous channel state information (CSI) at the transmitting nodes is proposed. A readily computable expression for the ergodic sum capacity of the MIMO links is derived. Based on this expression, which is a non-convex function of power allocation vectors, a projected gradient algorithm is developed to optimize the power allocation. For a symmetric set of MIMO links with independent identically distributed channels, it is observed that the space–time optimal power schedule reduces to a uniform isotropic power schedule when nominal interference is low, or to an orthogonal isotropic power schedule when nominal interference is high. Furthermore, the transition region between the latter two schedules is seen to be very sharp in terms of nominal interference-to-noise ratio (INR). For MIMO links with correlated channels, the corresponding space–time optimal power schedule is developed based on the knowledge of the channel correlation matrices. It is shown that the channel correlation has a great impact on the ergodic capacity and the optimality of different power scheduling approaches.

Index Terms—Multiple-input multiple-output (MIMO) systems, space–time power schedule, wireless mesh networks.

I. INTRODUCTION

IN a large wireless mesh network of many multiple-input multiple-output (MIMO) nodes, multiple MIMO links must share a common frequency band concurrently to ensure a high spectral efficiency of the whole network [1]–[3]. Developing optimal power schedules for a set of co-channel, concurrent and neighboring MIMO links is therefore important.

Power scheduling for multiple MIMO links has been studied in [4]–[7]. In [4], a space-only (i.e., time-invariant) power schedule is presented, and an iterative algorithm leading to the Nash equilibrium [8] is developed. In [5], the same space-only criterion is used, but a projected gradient algorithm is developed that yields a better result. In [6], the space-only approach is considered without channel state information (CSI) at the transmitting nodes. In [7], a space–time power schedule is proposed that generalizes the approaches used in [4] and [5].

In this paper, we focus on networks with symmetric topology, and assume that the network topology is known to a scheduler, and each receiving node knows the instantaneous CSI of the link-of-interest. For a given set of transmitting–receiving pairs, we present a space–time optimal power schedule without instantaneous CSI at the transmitting nodes. This work goes beyond that in [7], which assumes instantaneous CSI at the transmitting nodes and also beyond [6] which assumes a time-invariant transmitting covariance matrix at each link.

In the absence of instantaneous CSI at the transmitting nodes, the knowledge of some statistical properties of the CSI is necessary for designing optimal power schedules. We will aim to optimize an ergodic network capacity with respect to the statistical distributions of the CSI. In other words, we consider a "fast fading" scenario where the CSI changes rapidly with respect to the delay requirement, or equivalently, the effective data rate that we are interested in is for a time interval during which the CSI changes many times.

We will consider two cases of the statistical properties of the CSI. In the first, the channel matrix between every two nodes consists of independent and identically distributed (i.i.d.) complex Gaussian entries. In the second, each channel matrix consists of correlated entries. The first case will be referred to as the case of i.i.d. channels. The second case will be referred to as the case of correlated channels.

An important result in this paper is a "closed form" expression for the ergodic sum capacity of multiple MIMO links. This result, following from the quadratic forms of Gaussian random matrices [9] and [10], is expressed in terms of finite sums and a simple one-dimensional integral. It is readily computable. Another important result in this paper is the development of a projected gradient algorithm that allows one to maximize the ergodic sum capacity and hence to compute the corresponding optimal power schedule.

To illustrate the potential of the space–time optimal power schedule, we will compare it with the following schemes.

- *Uniform isotropic power schedule*: All links conduct transmissions at the same time and the same frequency, and the source covariance matrix used by all links is proportional to the identity matrix.
- *Orthogonal isotropic power schedule*: All links conduct transmissions in orthogonal channels (such as different times and/or different frequencies), and the source covariance matrix used by all links is proportional to the identity matrix.
- *Low-rank power schedule*: Each link uses a low rank source covariance matrix.

For a network of symmetric MIMO links with i.i.d. channels, we will show that the space–time optimal power schedule consistently outperforms the low rank power schedule. Interestingly, the space–time optimal power schedule reduces to a uniform isotropic power schedule when nominal interference level is low, or to an orthogonal isotropic power schedule when nominal interference level is high. Furthermore, the transition between the latter two schedules is very sharp along a nominal interference-to-noise ratio (INR) axis. All space-only power schedules, including those in [6], are shown to be suboptimal compared to ours.

It is important to note that by INR, we refer to a *nominal* INR unless specified otherwise. The *nominal* INR can be kept constant while the *actual* INR changes, which will be further explained.

We will also show that the channel correlation reduces the network capacity when nominal interference level is low, but increases the network capacity when nominal interference level is high. When the channel correlation is known, we will show that the eigenspace of the optimal source covariance matrices of each transmitting node is equal to that of the corresponding channel covariance matrix. In other words, the channel should be "whitened" by the source covariance matrices.

The rest of this paper is organized as follows. In Section II, the system model of distributed MIMO links is introduced, and the problem formulation is given. A readily computable form of the ergodic sum capacity of the MIMO links with i.i.d. channels is derived in Section III. In Section IV, we present the space–time optimal power scheduling scheme. In Section V, we investigate the impact of the channel correlation on the network capacity and study the corresponding space–time optimal power schedule. Numerical examples are shown in Section VI. Conclusions are drawn in Section VII.

II. PROBLEM FORMULATION

We consider data transmissions in a wireless mesh network of MIMO nodes. We focus on a time window where a set of distinct pairs of transmitting and receiving nodes in a given neighborhood need to accomplish their data transmissions. Typically, a pair of transmitting and receiving nodes consists of two adjacent nodes for power and bandwidth efficiency. We also consider the spectral region where the channel frequency response is "flat" (i.e., frequency-invariant). The key problem that we will address in this paper is how to design the power allocation and the source covariance matrix for each transmitting node so that the network throughput during this time window is maximized.

There are two simple solutions. One is to force all transmitting nodes to transmit at the same time and at all time, and the source covariance matrices are optimized under this constraint. We call this solution the space-only power schedule. The other solution is to slice the time into multiple slots, and during each time slot only one transmitting node is actually transmitting. This is known as time-division multiplexing access (TDMA). With TDMA, the power allocation for each link is either zero or a full power during each time slot.

There is a more general solution where during an interval of multiple time slots, the power allocations as well as the source covariance matrices for all links are treated as functions of the time slots. These functions can be optimized to maximize the total network throughput. This is the approach that we will follow in this paper. We call this approach space–time optimal power schedule. Note that the time slots here can be replaced by frequency bins or a combination of time slots and frequency bins. Both time and frequency are temporal characteristics with respect to space. But for a simple exposure of the concept of our scheme, we will only refer to time slots.

For convenience, we will call the transmitting node and the receiving node of the ith link as the ith transmitting node and the ith receiving node, respectively. Assuming that there are L links, it is sufficient to design the optimal power schedule over L time slots [7]. We will assume that the channel matrices remain constant over L consecutive time slots but change randomly over an interval of many multiples of L time slots. We will design the power schedule to maximize an ergodic network capacity which is averaged over the statistical distribution of the channel matrices. This capacity is achievable (approximately) over the interval of many multiples of L time slots. Therefore, our theory is valid for "fast fading" channels, i.e., the time delay due to encoding and decoding over many multiples of L time slots is tolerable.

We assume that each node has N antennas. The vector of the received signal $\mathbf{y}_i$ at the ith receiving node and in the tth time slot can be written as

$$\mathbf{y}_i(t) = \sqrt{\beta_{i,i}}\mathbf{H}_{i,i}\mathbf{x}_i(t) + \sum_{j=1, j \neq i}^{L} \sqrt{\beta_{i,j}}\mathbf{H}_{i,j}\mathbf{x}_j(t) + \mathbf{v}_i(t).$$

(1)

Here, $\beta_{i,i}$ denotes the average gain of the ith transmitting node and the ith receiving node, which is for a link-of-interest. $\beta_{i,j}$, $j \neq i$, is the average gain between the ith transmitting node and the ith receiving node, which is for an interfering link. $\mathbf{H}_{i,j}$ is the instantaneous $N \times N$ channel matrix (normalized by the averaged channel gain) between the jth transmitting node and the jth receiving node. $\mathbf{x}_i(t)$, $i = 1, \ldots, L$, denotes the $N \times 1$ vector of the transmitted signal from the ith transmitting node within the slot t. Within time slot t, $\mathbf{x}_i(t)$ is a random function of the continuous time τ which is omitted. $\mathbf{v}_i(t)$, $i = 1, \ldots, L$, is the $N \times 1$ vector of the i.i.d. additive (slot-invariant) white Gaussian noise (AWGN) with zero mean and unit covariance matrix $\mathbf{C}_{\mathbf{v}_i} = \mathbf{I}_N$. And $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. Fig. 1 shows an example of a circular network with three concurrently communicating MIMO links. A list summarizing the notations used for the system model is given in Appendix A.
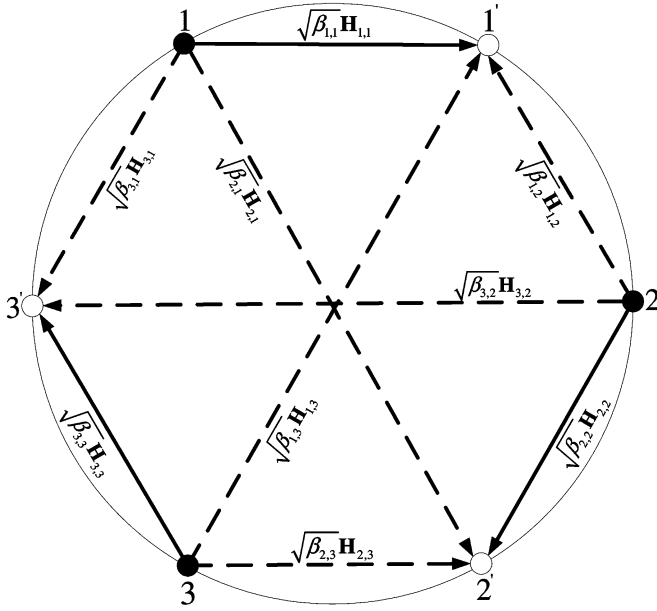
Fig. 1. Circular wireless network. Solid circles: source nodes; hollow circles: destination nodes. Solid lines: signal streams; dashed lines: interference streams.

The first term in (1) is the signal of interest at the ith receiving node, and the second term is the sum of interfering signals from all other $L-1$ transmitting nodes. We assume that all the transmitted signals are Gaussian distributed with zero mean and their covariance matrix is periodically time varying with the period equal to L time slots. We will write $\mathbf{P}_i(t) \triangleq E\{\mathbf{x}_i(t)\mathbf{x}_i(t)^H\}$, $t = 1, \ldots, L$, where $E\{\cdot\}$ stands for statistical expectation, and $(\cdot)^H$ denotes matrix Hermitian transpose. Moreover, we assume that $(1/L)\sum_{t=1}^{L}\mathrm{tr}\{\mathbf{P}_i(t)\} = P_i$, $i = 1, \ldots, L$, where P_i is the averaged transmit power from the ith transmitting node, and $\mathrm{tr}\{\cdot\}$ stands for the trace of a matrix. We will need further assumptions as shown below.

- There is no coding cooperation between interfering transmitting nodes. Namely, the source information is encoded independently at different transmitting nodes. Cooperative coding among interfering MIMO links would significantly increase the networking complexity, which is undesirable.
- The entries of $\mathbf{H}_{i,j}$, $i, j = 1, \ldots, L$, are complex Gaussian distributed with zero mean and unit variance. (As mentioned before, $\mathbf{H}_{i,j}$ remains constant over L time slots but varies randomly during an interval of many multiples of L time slots. For a stationary network, such random fading can be induced by randomly changing the phase of each transmitting antenna.) Thus, the averaged signal-to-noise ratio (SNR) at each antenna of the ith receiving node is given by $A_i \triangleq \beta_{i,i}P_i$, while the nominal INR from the jth transmitting node to each antenna of the ith receiving node is defined as $B_{i,j} \triangleq \beta_{i,j}P_j$. We treat $\beta_{i,j}$, $i = 1, \ldots, L$ as deterministic parameters both in the analysis and the simulations.
- There is no instantaneous CSI knowledge at any transmitting node, and the ith receiving node only knows the CSI of the link-of-interest, i.e., $\mathbf{H}_{i,i}$.

With a given set of transmit covariance matrices $\mathbf{P}_i(t)$, $i, t = 1, \ldots, L$, the averaged ergodic sum capacity of L links over L time slots can be written as

$$I_a(\bar{\mathbf{P}}) = \frac{1}{L}E_{\mathbf{H}}\left\{\sum_{t=1}^{L}\sum_{i=1}^{L}\log_2\left|\mathbf{I}_N + \beta_{i,i}\mathbf{H}_{i,i}\mathbf{P}_i(t)\mathbf{H}_{i,i}^H\mathbf{R}_i^{-1}(t)\right|\right\} \quad (2)$$

where $|\cdot|$ denotes the determinant of a matrix or the absolute value of a scalar, $E_{\mathbf{H}}\{\cdot\}$ stands for the statistical expectation with respect to all channel matrices $\mathbf{H} \triangleq \left[\mathbf{H}_{1,1}^T, \ldots, \mathbf{H}_{L,L}^T\right]^T$, and $(\cdot)^T$ denotes the matrix transpose. In (2), $\bar{\mathbf{P}}$ is a matrix stacking the source covariance matrices of all links:

$$\bar{\mathbf{P}} \triangleq \left[\bar{\mathbf{P}}_1^T, \ldots, \bar{\mathbf{P}}_L^T\right]^T$$
$$\bar{\mathbf{P}}_i \triangleq \left[\mathbf{P}_i^T(1), \ldots, \mathbf{P}_i^T(L)\right]^T \quad i = 1\ldots, L$$

and $\mathbf{R}_i(t)$ is the (actual) interference-plus-noise covariance matrix at the ith receiving node in the tth time slot

$$\mathbf{R}_i(t) = \sum_{j=1, j\neq i}^{L}\beta_{i,j}\mathbf{H}_{i,j}\mathbf{P}_j(t)\mathbf{H}_{i,j}^H + \mathbf{I}_N.$$

In this paper, we aim at maximizing the ergodic capacity (2) by choosing the transmit covariance matrices $\mathbf{P}_i(t)$, $i, t = 1, \ldots, L$. This problem can be compactly formulated as

$$\max_{\bar{\mathbf{P}}} \quad I_a(\bar{\mathbf{P}}) \quad (3)$$

$$\text{s.t.} \quad \frac{1}{L}\sum_{t=1}^{L}\mathrm{tr}\{\mathbf{P}_i(t)\} = P_i \quad (4)$$

$$\mathbf{P}_i(t) \geq 0 \quad i, t = 1, \ldots, L \quad (5)$$

where (4) represents the average transmit power constraint at each transmitting node. Note that in this paper, we focus on networks with a symmetric topology. Therefore, no link suffers a fairness problem under the ergodic sum capacity. When the network is asymmetric, fairness issues arise. Then, the cost functions must be changed. One commonly used criterion is known as "proportional fairness," which can be formulated as sum of logarithms of link capacities. The gradient-based method can be applied similarly. But the details would go beyond the scope of this paper.

III. ERGODIC SUM CAPACITY FOR MIMO LINKS WITH I.I.D. CHANNELS

In this section, we derive a closed-form expression for the ergodic sum capacity (2) for the case of i.i.d. channels. The correlated channel case will be addressed in Section V. For simplicity, we first derive a closed-form expression for the following ergodic sum capacity at one time slot:

$$I(\mathbf{P}_1(t), \ldots, \mathbf{P}_L(t)) = E_{\mathbf{H}}\left\{\sum_{i=1}^{L}\log_2\left|\mathbf{I}_N + \beta_{i,i}\mathbf{H}_{i,i}\mathbf{P}_i(t)\mathbf{H}_{i,i}^H\mathbf{R}_i^{-1}(t)\right|\right\}. \quad (6)$$

We will later generalize the result to obtain a closed-form expression for the ergodic sum capacity averaged over all time slots as shown in (2).

Let us denote

$$\mathbf{P}_i(t) = \mathbf{U}_i(t)\mathbf{D}_i(t)\mathbf{U}_i^H(t) \qquad (7)$$

as the eigendecomposition of $\mathbf{P}_i(t)$, where $\mathbf{U}_i(t)$ is the unitary matrix of eigenvectors, and

$$\mathbf{D}_i(t) = \text{diag}\{d_{i1}(t), d_{i2}(t), \ldots, d_{iN}(t)\}$$

is an $N \times N$ diagonal matrix of all eigenvalues. For convenience, we will use the $N \times 1$ column vectors $\mathbf{d}_i(t)$, $i = 1, \ldots, L$, defined as

$$\mathbf{d}_i(t) \triangleq [d_{i1}(t), d_{i2}(t), \ldots, d_{iN}(t)]^T.$$

Substituting (7) into (6), we can rewrite (6) as

$$I(\mathbf{d}_1(t), \ldots, \mathbf{d}_L(t)) = E_{\tilde{\mathbf{H}}(t)}\left\{\sum_{i=1}^{L}\log_2\left|\mathbf{I}_N \right.\right.$$
$$\left.\left. + \beta_{i,i}\tilde{\mathbf{H}}_{i,i}(t)\mathbf{D}_i(t)\tilde{\mathbf{H}}_{i,i}^H(t)\tilde{\mathbf{R}}_i^{-1}(t)\right|\right\} \qquad (8)$$

where

$$\tilde{\mathbf{H}}(t) \triangleq \left[\tilde{\mathbf{H}}_{1,1}^T(t), \ldots, \tilde{\mathbf{H}}_{L,L}^T(t)\right]^T$$
$$\tilde{\mathbf{H}}_{i,i}(t) = \mathbf{H}_{i,i}\mathbf{U}_i(t)$$
$$\tilde{\mathbf{H}}_{i,j}(t) = \mathbf{H}_{i,j}\mathbf{U}_j(t)$$
$$\tilde{\mathbf{R}}_i(t) = \sum_{j=1,j\neq i}^{L}\beta_{i,j}\tilde{\mathbf{H}}_{i,j}(t)\mathbf{D}_j(t)\tilde{\mathbf{H}}_{i,j}^H(t) + \mathbf{I}_N.$$

Since $\mathbf{H}_{i,j}$, $i, j = 1, \ldots, L$, has i.i.d. Gaussian entries and $\mathbf{U}_i(t)$ is unitary, the statistics of $\tilde{\mathbf{H}}_{i,j}(t)$ are identical to those of $\mathbf{H}_{i,j}$ [11]. Hereafter, for notational simplicity, we drop the tildes on the $\mathbf{H}$ matrices and the dependence of $\mathbf{H}$ matrices on t, and write the ergodic sum capacity expression (8) as

$$I(\mathbf{d}_1(t), \ldots, \mathbf{d}_L(t))$$
$$= E_{\mathbf{H}}\left\{\sum_{i=1}^{L}\log_2\left|\mathbf{I}_N + \beta_{i,i}\mathbf{H}_{i,i}\mathbf{D}_i(t)\mathbf{H}_{i,i}^H\mathbf{R}_i^{-1}(t)\right|\right\}$$
$$= \sum_{i=1}^{L}\left(E_{\mathbf{H}}\left\{\log_2\left|\mathbf{R}_i(t) + \beta_{i,i}\mathbf{H}_{i,i}\mathbf{D}_i(t)\mathbf{H}_{i,i}^H\right|\right\}\right.$$
$$\left. - E_{\mathbf{H}}\{\log_2|\mathbf{R}_i(t)|\}\right) \qquad (9)$$
$$= \sum_{i=1}^{L}\left(E_{\mathbf{H}_i}\left\{\log_2\left|\mathbf{I}_N + \mathbf{H}_i\mathbf{\Lambda}_i(t)\mathbf{H}_i^H\right|\right\}\right.$$
$$\left. - E_{\bar{\mathbf{H}}_i}\left\{\log_2\left|\mathbf{I}_N + \bar{\mathbf{H}}_i\bar{\mathbf{\Lambda}}_i(t)\bar{\mathbf{H}}_i^H\right|\right\}\right) \qquad (10)$$

where

$$\mathbf{R}_i(t) = \sum_{j=1,j\neq i}^{L}\beta_{i,j}\mathbf{H}_{i,j}\mathbf{D}_j(t)\mathbf{H}_{i,j}^H + \mathbf{I}_N$$
$$\mathbf{H}_i = [\mathbf{H}_{i,1}, \ldots, \mathbf{H}_{i,L}]$$
$$\bar{\mathbf{H}}_i = [\mathbf{H}_{i,1}, \ldots, \mathbf{H}_{i,i-1}, \mathbf{H}_{i,i+1}, \mathbf{H}_{i,L}]$$
$$\mathbf{\Lambda}_i(t) = \text{diag}\left\{\left[\boldsymbol{\lambda}_{i,1}^T(t), \ldots, \boldsymbol{\lambda}_{i,L}^T(t)\right]\right\}$$
$$\bar{\mathbf{\Lambda}}_i(t) = \text{diag}\left\{\left[\boldsymbol{\lambda}_{i,1}^T(t), \ldots, \boldsymbol{\lambda}_{i,i-1}^T(t),\right.\right.$$
$$\left.\left. \boldsymbol{\lambda}_{i,i+1}^T(t), \ldots, \boldsymbol{\lambda}_{i,L}^T(t)\right]\right\}$$
$$\boldsymbol{\lambda}_{i,j}(t) = \beta_{i,j}\mathbf{d}_j(t).$$

It can be seen from (10) that the ergodic sum capacity expression is a summation of $2L$ logarithmic terms all having a similar structure. For convenience, we can now focus on one of the terms

$$I_{i1}(\mathbf{\Lambda}_i(t)) = E_{\mathbf{H}_i}\left\{\log_2\left|\mathbf{I}_N + \mathbf{H}_i\mathbf{\Lambda}_i(t)\mathbf{H}_i^H\right|\right\}. \qquad (11)$$

It can be seen that (11) is equivalent to the ergodic capacity of a point-to-point MIMO system under AWGN with NL transmit antennas and N receive antennas. A closed-form expression for the ergodic capacity of such a system has been shown in [9], where a determinant representation for the distribution of quadratic forms of a complex Gaussian matrix [10] has been used. Using the results in [9], (11) can be written as

$$I_{i1}(\mathbf{\Lambda}_i(t)) = \log_2(e)\sum_{n=0}^{N-1}\sum_{k=1}^{NL}c_{tikn}Q(n, \lambda_{t,i,k})$$

where [see (12)–(14), shown at the bottom of the next page]. (Note that there are typos in (16) of [9], where the terms in the first summation should be $(i!(-1)^{(i-r)}/(i-r)!)\gamma^{(r+1)}e^{1/\gamma}S_1(1/\gamma)$, using the same notations as in [9] except for S_1.) Here, $\lambda_{t,i,k} \triangleq [\mathbf{\Lambda}_i(t)]_{k,k}$, $k = 1, \ldots, NL$, denotes the (k, k)th element of matrix $\mathbf{\Lambda}_i(t)$, and $S_1(x) \triangleq \int_x^{\infty}e^{-\tau}/\tau d\tau$ is the exponential integral function of order 1 [12, eq. 5.1.1]. It is worth noting that (12)–(14) are derived under the assumption that all $\lambda_{t,i,k}$, $k = 1, \ldots, NL$, have distinct values. When some of them are identical, we can obtain the closed-form ergodic sum capacity expression by deriving the limit of (11) with respect to those common values of $\lambda_{t,i,k}$ using L'Hospital's rule. For numerical evaluation, it is sufficient to slightly and randomly perturb these identical values of $\lambda_{t,i,k}$ because all functions are continuous and $\lambda_{t,i,k}$ are deterministic.

Similar to (11), we can write each of the $2L$ summands in (10) in a closed form. Thus, the ergodic sum capacity can be represented as

$$I(\mathbf{d}_1(t), \ldots, \mathbf{d}_L(t))$$
$$= \log_2(e)\sum_{i=1}^{L}\left[\sum_{n=0}^{N-1}\sum_{k=1}^{NL}c_{tikn}Q(n, \lambda_{t,i,k})\right.$$

$$-\sum_{n=0}^{N-1}\sum_{k=1}^{N(L-1)} u_{tikn} Q\left(n, \bar{\lambda}_{t,i,k}\right)\Bigg] \quad (15)$$

where $\bar{\lambda}_{t,i,k} \triangleq \left[\bar{\mathbf{\Lambda}}_i(t)\right]_{k,k}$, $k = 1, \ldots, N(L-1)$, c_{tikn} is given by (12) and (13), and u_{tikn} is given by (16) and (17), shown at the bottom of the page.

The result (15) can be readily used to write the averaged ergodic sum capacity (2) in closed form. First, we rewrite (2) as (18) and (19), shown at the bottom of the page. Then, by following the steps given above, we can write (19) in closed form as

$$I_a\left(\bar{\mathbf{d}}\right) = \frac{\log_2(e)}{L} \sum_{t=1}^{L}\sum_{i=1}^{L}\Bigg[\sum_{n=0}^{N-1}\sum_{k=1}^{NL} c_{tikn} Q(n, \lambda_{t,i,k})$$
$$-\sum_{n=0}^{N-1}\sum_{k=1}^{N(L-1)} u_{tikn} Q(n, \bar{\lambda}_{t,i,k})\Bigg] \quad (20)$$

where $\bar{\mathbf{d}}$ is a vector stacking the power scheduling parameters of all links at all L time slots

$$\bar{\mathbf{d}} \triangleq \left[\bar{\mathbf{d}}_1{}^T, \ldots, \bar{\mathbf{d}}_L{}^T\right]^T$$
$$\bar{\mathbf{d}}_i \triangleq \left[\mathbf{d}_i^T(1), \ldots, \mathbf{d}_i^T(L)\right]^T \quad i = 1 \ldots, L$$

and $\bar{\mathbf{d}}_i$ is a vector stacking the power scheduling parameters of the ith link. In (20), c_{tikn} and u_{tikn} can be written using (12) and (13), and (16) and (17), respectively.

As shown in (20), the ergodic sum capacity is now expressed as a finite summation involving rational functions and exponential integration functions of the power scheduling vectors $\mathbf{d}_i(t)$, $i, t = 1, \ldots, L$, of all transmitting nodes. The exponential integration function is available in many software such as MATLAB and Mathematica. Thus, (20) is easy to compute. For a better understanding, the function $Q(n, \lambda)$ is plotted in Fig. 2 versus λ for $n = 0, 1, 2, 3$. The closed-form expression (20) enables us to numerically optimize the averaged ergodic sum capacity and hence the power scheduling.

The accuracy of the ergodic sum capacity expression (20) is verified by Fig. 3. This figure shows both the analytical and the simulation results of the ergodic capacity versus the nominal INR (i.e., B) of a symmetric network ($B_{i,j} = B$), with $L = 2$ active links, where each node has $N = 2$ antennas. Here, we set $P_i = 20$ dB, $\beta_{i,i} = 1$, $i = 1, \ldots, L$, and $\beta_{i,j} = \beta$, $i, j = 1, \ldots, L, j \neq i$. The ergodic capacity shown is a per-link ergodic capacity which is the ergodic sum capacity divided by the number of active links L. For the simulation results, 1000 channel realizations were used. A randomly selected power allocation vector was used for each link. It can be seen from Fig. 3 that the analytical result is consistent with the simulation result.

$$c_{tikn} = \frac{(-1)^{N-n-1}\lambda_{t,i,k}^{N(L-1)-1}}{n!}\left(\prod_{h\neq k}(\lambda_{t,i,k} - \lambda_{t,i,h})\right)^{-1} b_{tikn} \quad (12)$$

$$b_{tikn} = \begin{cases} \displaystyle\sum_{1\leq j_1 < \cdots < j_{N-n-1} \leq NL}^{j_r \neq k} \lambda_{t,i,j_1}\ldots\lambda_{t,i,j_{N-n-1}}, & n = 0, \ldots, N-2 \\ 1, & n = N-1. \end{cases} \quad (13)$$

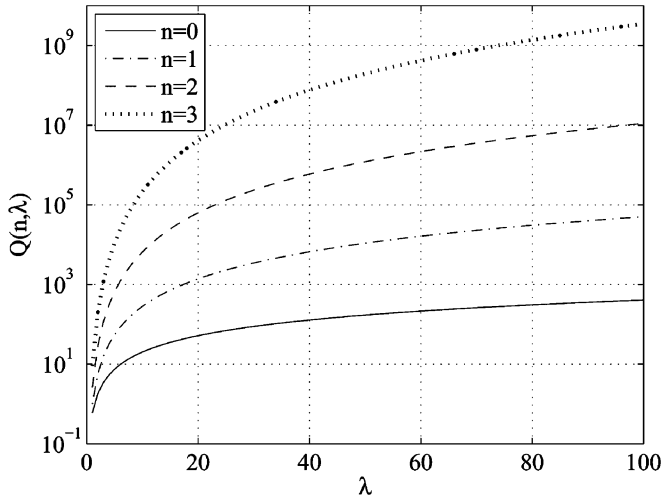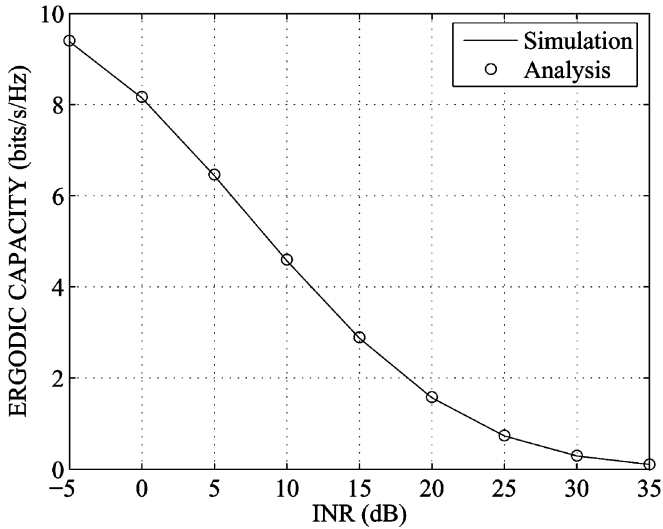$$Q(n, \lambda_{t,i,k}) = \int_0^\infty \ln(1+x)x^n e^{-(x/\lambda_{t,i,k})}dx$$
$$= \sum_{r=0}^{n} \frac{n!(-1)^{(n-r)}}{(n-r)!}\lambda_{t,i,k}^{r+1} e^{1/\lambda_{t,i,k}} S_1\left(\frac{1}{\lambda_{t,i,k}}\right) + \sum_{r=1}^{n}\sum_{s=0}^{r-1}\sum_{h=0}^{r-s-1}\frac{n!(-1)^{(n-r)}\lambda_{t,i,k}^{h+s+2}}{(n-r)!(r-s-1-h)!(r-s)}. \quad (14)$$

$$u_{tikn} = \frac{(-1)^{N-n-1}\bar{\lambda}_{t,i,k}^{N(L-2)-1}}{n!}\left(\prod_{h\neq k}\left(\bar{\lambda}_{t,i,k} - \bar{\lambda}_{t,i,h}\right)\right)^{-1} \bar{b}_{tikn} \quad (16)$$

$$\bar{b}_{tikn} = \begin{cases} \displaystyle\sum_{1\leq j_1 < \cdots < j_{N-n-1} \leq N(L-1)}^{j_r \neq k} \bar{\lambda}_{t,i,j_1}\ldots\bar{\lambda}_{t,i,j_{N-n-1}}, & n = 0, \ldots, N-2 \\ 1, & n = N-1. \end{cases} \quad (17)$$

$$I_a\left(\bar{\mathbf{P}}\right) = \frac{1}{L}\sum_{t=1}^{L}\sum_{i=1}^{L}\left(E_{\mathbf{H}}\left\{\log_2\left|\mathbf{R}_i(t) + \beta_{i,i}\mathbf{H}_{i,i}\mathbf{D}_i(t)\mathbf{H}_{i,i}^H\right|\right\} - E_{\mathbf{H}}\{\log_2|\mathbf{R}_i(t)|\}\right) \quad (18)$$

$$= \frac{1}{L}\sum_{t=1}^{L}\sum_{i=1}^{L}\left(E_{\mathbf{H}_i}\left\{\log_2\left|\mathbf{I}_N + \mathbf{H}_i\mathbf{\Lambda}_i(t)\mathbf{H}_i^H\right|\right\} - E_{\bar{\mathbf{H}}_i}\left\{\log_2\left|\mathbf{I}_N + \bar{\mathbf{H}}_i\bar{\mathbf{\Lambda}}_i(t)\bar{\mathbf{H}}_i^H\right|\right\}\right). \quad (19)$$

Fig. 2. Function $Q(n, \lambda)$ for $n = 0, 1, 2, 3$.

Fig. 3. Comparison of the analytical and simulation results of the per-link ergodic capacity versus nominal INR for the case of i.i.d. channels. Here, $N = 2$, $L = 2$.

IV. SPACE–TIME POWER SCHEDULE

It has been shown in [7] that by applying a space–time power schedule where the source covariance matrices are allowed to be functions of time, a larger (averaged) ergodic sum capacity can be achieved. However, the power scheduling scheme in [7] requires the instantaneous CSI knowledge at the transmitting nodes, which is valid for slow fading channels. In this section, we optimize the space–time power schedule for fast fading channels by maximizing the ergodic sum capacity expression derived in Section III for which the instantaneous CSI is not required at the transmitting nodes.

Taking the power constraint of each active link into account, our space–time power scheduling approach becomes the following optimization problem:

$$\max_{\bar{\mathbf{d}}} \quad I_a\left(\bar{\mathbf{d}}\right) \tag{21}$$

$$\text{s.t.} \quad \left\|\bar{\mathbf{d}}_i\right\|_1 = LP_i, \quad \bar{\mathbf{d}}_i \geq 0, \quad i = 1, \ldots, L \tag{22}$$

where (22) is the set of the transmit power constraints at all transmitting nodes, and $\|\cdot\|_1$ denotes the sum norm (or l_1 norm) of a vector. For a vector $\mathbf{x}$, $\mathbf{x} \geq 0$ means that each entry of $\mathbf{x}$ is nonnegative.

The results in [6] show that when the INR is sufficiently low, the ergodic sum capacity (15) is a concave function of the power allocation vectors $\mathbf{d}_1(t), \ldots, \mathbf{d}_L(t)$, but when the INR is sufficiently high, (15) becomes a convex function of the power allocation vectors. (Since the power allocation in [6] is independent of time, the INR in [6] is an instantaneous INR, an averaged INR and also the nominal INR.) However, in general, it can be seen from (10) that due to the mutual interference among different links, the ergodic sum capacity is neither a convex function, nor a concave function, of the power allocation vectors $\mathbf{d}_1(t), \ldots, \mathbf{d}_L(t)$. Similarly, (20) is neither a convex nor a concave function of the power scheduling vector $\bar{\mathbf{d}}$ in the general INR region. Thus, in general, (21)–(22) is a nonconvex optimization problem.

The objective function (21) is quite complicated and does not have a clear structure. The gradient-based methods are effective for such a problem [13, ch. 1]. These methods guarantee the convergence to a stationary point, provided that proper step sizes are chosen. Moreover, since the constraints (22) are simple linear constraints, the projected gradient technique [13, sec. 2.3] can be applied to solve the problem (21)–(22). We first give a brief introduction to the projected gradient algorithm.

Let us consider the following constrained optimization problem

$$\max_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t. } \mathbf{x} \in \mathcal{X}$$

where $f(\cdot)$ is a continuously differentiable scalar function, and $\mathcal{X}$ is a nonempty, closed, and convex set. The projected gradient algorithm is a feasible direction method. The algorithm starts at an initial point $\mathbf{x}^{(0)}$. At the kth iteration, $\mathbf{x}^{(k)}$ is updated as

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta_k \left(\tilde{\mathbf{x}}^{(k)} - \mathbf{x}^{(k)} \right) \tag{23}$$

where $\delta_k \in (0, 1]$ is a step size, and

$$\tilde{\mathbf{x}}^{(k)} = \text{proj}\left[\mathbf{x}^{(k)} + s_k \nabla f\left(\mathbf{x}^{(k)} \right) \right]. \tag{24}$$

Here, $\text{proj}[\cdot]$ denotes the projection onto the feasible set $\mathcal{X}$, and s_k is a positive scalar. From (23) and (24), we can see that to obtain the vector $\tilde{\mathbf{x}}^{(k)}$, we first take a step $s_k \nabla f\left(\mathbf{x}^{(k)} \right)$ along the gradient, similar to the steepest descent algorithm. Then we project the result $\mathbf{x}^{(k)} + s_k \nabla f\left(\mathbf{x}^{(k)} \right)$ onto the feasible set $\mathcal{X}$. It has been shown that after a sufficient number of iterations, the sequence $\mathbf{x}^{(k)}$ converges to a local optimal point $\mathbf{x}_{\text{opt}}$ [13, sec. 2.3].

Two major issues of a projected gradient algorithm are: 1) computing the gradient of the objective function, and 2) projecting the gradient-modified vector into the feasible set.

The derivation of the gradient of (20) is tedious but straightforward. The complete expression for the gradient of (20) with respect to $d_{lm}(t), l = 1, \ldots, L, m = 1, \ldots, N, t = 1, \ldots, L$ is listed in Appendix B.

In the sequel, we show how to carry out the projection operation. Let us take the ith link as an example, and denote the $NL \times 1$ column vector of gradient as

$$\mathbf{g}_i \triangleq \left[\frac{\partial I_a(\bar{\mathbf{d}})}{\partial d_{i1}(1)}, \ldots, \frac{\partial I_a(\bar{\mathbf{d}})}{\partial d_{iN}(L)} \right]^T.$$

Then taking a step along the positive gradient (24), the $NL \times 1$ power vector is updated as

$$\hat{\mathbf{d}}_i = \bar{\mathbf{d}}_i + s\mathbf{g}_i.$$

The next step of the projected gradient algorithm is to project $\hat{\mathbf{d}}_i$ onto the feasible region of power vectors defined by (22). In fact, the projection operation can be seen as searching for a point $\tilde{\mathbf{d}}_i$ in the region of (22), which has a minimal Euclidean distance to the point $\hat{\mathbf{d}}_i$. Thus for the ith link, the projection operation can be formulated as the following constrained optimization problem:

$$\min_{\tilde{\mathbf{d}}_i} \quad \left\| \tilde{\mathbf{d}}_i - \hat{\mathbf{d}}_i \right\|^2 \tag{25}$$

$$\text{s.t.} \quad \left\| \tilde{\mathbf{d}}_i \right\|_1 = LP_i, \quad \tilde{\mathbf{d}}_i \geq 0, \ i = 1, \ldots, L \tag{26}$$

where $\| \cdot \|$ denotes the Euclidean norm of a vector.

The problem (25)–(26) can be efficiently solved by using the Lagrange multiplier method [14, sec. 5.1]. Each element of $\tilde{\mathbf{d}}_i$ can be written as

$$\tilde{d}_{in} = \left[\hat{d}_{in} - \frac{\mu}{2} \right]^+, \quad n = 1, \ldots, NL \tag{27}$$

where μ is the Lagrange multiplier, and for a real scalar x, $[x]^+ \triangleq \max\{x, 0\}$.

The Lagrange multiplier μ can be obtained by substituting (27) back into (26). We have the following equation:

$$\sum_{n=1}^{NL} \max \left\{ \hat{d}_{in} - \frac{\mu}{2}, 0 \right\} = LP_i. \tag{28}$$

The left-hand side of (28) is a piecewise linear function and monotonically decreasing with respect to μ. Thus, (28) can be easily solved by, for example, the bisection method [14, pp. 145].

Note that under the space–time power scheduling scheme, the iterative computation of the power allocation vectors for all links can be done concurrently or sequentially. The concurrent iterations must be done at a single place (centralized). But the sequential iterations can be done in a distributed manner, i.e., each link updates its own power allocation vector at each iteration. At the end of each sequential iteration, all links must exchange their updated power allocation vectors before the next sequential iteration takes place. Due to space limitation, we only present the centralized technique in detail in this paper.

At the kth iteration of the centralized technique, the power scheduling vector is updated as

$$\bar{\mathbf{d}}_i^{(k+1)} = \bar{\mathbf{d}}_i^{(k)} + \delta_k \left(\tilde{\mathbf{d}}_i^{(k)} - \bar{\mathbf{d}}_i^{(k)} \right) \tag{29}$$

TABLE I
PROCEDURE OF THE PROJECTED GRADIENT POWER
ALLOCATION APPROACH

1) Initialize power allocation vectors $\bar{\mathbf{d}}_i$. Choose step sizes. Set $k = 0$.
2) Set $k := k + 1$. Calculate the gradient of (20) $\mathbf{g}_i^{(k)}$ using (53)-(59). Let $\hat{\mathbf{d}}_i^{(k)} = \bar{\mathbf{d}}_i^{(k)} + s_k \mathbf{g}_i^{(k)}$. Project $\hat{\mathbf{d}}_i^{(k)}$ to obtain $\tilde{\mathbf{d}}_i^{(k)}$ using (27). Update $\bar{\mathbf{d}}_i^{(k)}$ using (29)-(30).
3) If convergent, end. Else go to step 2.

$$\tilde{\mathbf{d}}_i^{(k)} = \text{proj} \left[\bar{\mathbf{d}}_i^{(k)} + s_k \mathbf{g}_i^{(k)} \right], \quad i = 1, \ldots, L \tag{30}$$

where δ_k and s_k are scalars of step size. They can be chosen, for example, according to the Armijo rule along the feasible direction [13, pp. 225–226]. In this rule, $s_k = s$ is a constant throughout the iterations, and $\delta_k = \theta^{m_k}$, where m_k is the minimal nonnegative integer that satisfies the following inequality:

$$I_a \left(\bar{\mathbf{d}}^{(k+1)} \right) - I_a \left(\bar{\mathbf{d}}^{(k)} \right)$$
$$\geq \sigma \theta^{m_k} \sum_{i=1}^{L} \left(\mathbf{g}_i^{(k)} \right)^T \left(\tilde{\mathbf{d}}_i^{(k)} - \bar{\mathbf{d}}_i^{(k)} \right).$$

Here, σ and θ are constants. According to [13, pp. 225–226], usually σ is chosen close to 0, and a proper choice of θ is from 0.1 to 0.5.

The steps of (29) and (30) are carried out for all L links and continue till the sequence $\bar{\mathbf{d}}^{(k)}$ converges. Such convergence is guaranteed according to the principle of the projected gradient algorithm using the Armijo rule along the feasible direction [13, Proposition 2.3.1]. The convergence criterion is taken as

$$\max \text{abs} \left\{ \bar{\mathbf{d}}^{(k+1)} - \bar{\mathbf{d}}^{(k)} \right\} \leq \varepsilon$$

where $\max \text{abs}\{\cdot\}$ denotes the maximal absolute value among all elements of a vector, and ε is a positive constant close to 0. Since the projected gradient algorithm only converges to a local optimal solution, multiple initializations can be used to improve the performance. In Section VI, we will study the achievable ergodic capacities with different number of initializations.

The procedure of applying the projected gradient technique to solve the problem (21)–(22) is summarized in Table I. Note that there may be peak transmit power constraint in some practical systems such that $d_i(t) \leq d_{\max}$, where $d_{\max}$ denotes the peak transmit power. The proposed projected gradient approach can be easily extended to solve the problem (21)–(22) with the additional peak transmit power constraint.

The proposed space–time power schedule requires only the knowledge of averaged SNR and nominal INR of each link at the transmitting nodes. This knowledge can be easily obtained by exploiting the topology of the wireless networks and the averaged transmit power of each transmitting node, which in turn, can be obtained by having each node report its current location

and power budget to the scheduler periodically. In practice, the optimization procedure can be run offline for different combinations of averaged SNR and nominal INR. The resulting optimal parameters can be tabulated. Then, in real-time applications, we only need to look up this table to select the optimal power parameters. It is worth noting that although the instantaneous SNR and the instantaneous INR may vary from time slot to time slot, the averaged SNR and the nominal INR are static for many multiples of L time slots according to the assumption in Section II.

Before finishing this section, we want to discuss a special case of the proposed space–time power schedule. When only one time slot is considered for power scheduling, we have the space-only power schedule, which can be written as the following constrained optimization problem:

$$\max_{\mathbf{d}_1,\ldots,\mathbf{d}_L} \quad I(\mathbf{d}_1,\ldots,\mathbf{d}_L) \tag{31}$$
$$\text{s.t.} \quad \|\mathbf{d}_i\|_1 = P_i, \quad \mathbf{d}_i \geq 0, \quad i = 1,\ldots,L. \tag{32}$$

It has been shown in [6] that at a sufficiently low interference level, the space-only optimal power schedule is a uniform isotropic power schedule where all links use the same source covariance matrix and the source covariance is the identity matrix. While at a sufficiently high interference level, it is shown in [6] that the space-only optimal power schedule becomes a low rank power schedule where each link uses a low rank source covariance matrix. However, the work [6] does not provide a good answer for the intermediate region of interference. Our optimization based on a single time slot yields the space-only optimal power schedule for any given interference level, which will be shown in Section VI.

Also note that for the space-only power schedule, the nominal interference referred to in this paper is also the actual interference.

V. CHANNEL CORRELATION

In Sections III and IV, we assumed that the channel matrix has i.i.d. Gaussian elements. Yet, in some practical applications, the elements of a channel matrix can be correlated due to reasons such as the short spacing between antenna elements, and/or lack of scatters around the antennas [15], [16]. In this section, we investigate the impact of channel correlation on the ergodic sum capacity, and also study the corresponding space–time optimal power schedule. As we aim to study the optimal power schedule, the channel correlation at the transmitter side is of great interest. A widely accepted model for such a channel correlation is given by [15], [16]

$$\mathbf{H}_{i,j} = \mathbf{H}_{i,j}^w \mathbf{\Omega}_j^{1/2} \tag{33}$$

where $\mathbf{H}_{i,j}^w$ has i.i.d. complex Gaussian distributed entries with zero mean and unit variance, and $\mathbf{\Omega}_i \geq 0$ characterizes the channel correlation at the transmit side. Although the MIMO channel realizations $\mathbf{H}_{i,j}$ can be time varying from slot to slot, we assume that the channel statistics $\mathbf{\Omega}_j$ is static for many multiples of L time slots.

Let us denote

$$\mathbf{\Omega}_j \triangleq \mathbf{V}_j \mathbf{\Gamma}_j \mathbf{V}_j^H \tag{34}$$

as the eigendecomposition of the channel covariance matrix $\mathbf{\Omega}_j$, where $\mathbf{V}_j$ is a unitary matrix of eigenvectors, and $\mathbf{\Gamma}_j$ is a diagonal matrix containing all the eigenvalues. Substituting (34) back into (33), we have

$$\mathbf{H}_{i,j} = \mathbf{H}_{i,j}^w \mathbf{V}_j \mathbf{\Gamma}_j^{1/2} \mathbf{V}_j^H.$$

Let us introduce the following new matrices

$$\mathbf{S}_{i,j} \triangleq \mathbf{H}_{i,j} \mathbf{V}_j \mathbf{\Gamma}_j^{-1/2} = \mathbf{H}_{i,j}^w \mathbf{V}_j \tag{35}$$
$$\mathbf{Q}_j(t) \triangleq \mathbf{\Gamma}_j^{1/2} \mathbf{V}_j^H \mathbf{P}_j(t) \mathbf{V}_j \mathbf{\Gamma}_j^{1/2}. \tag{36}$$

It can be seen from (35) that since $\mathbf{V}_j$ is a unitary matrix, $\mathbf{S}_{i,j}$ has the same distribution as $\mathbf{H}_{i,j}^w$[11]. In other words, the entries of $\mathbf{S}_{i,j}$ are i.i.d. Gaussian distributed with zero mean and unit variance. Substituting (35) and (36) back into (6), we have

$$I(\mathbf{Q}_1(t),\ldots,\mathbf{Q}_L(t)) = E_{\mathbf{S}}\Bigg\{\sum_{i=1}^{L}\log_2\Bigg|\mathbf{I}_N$$
$$+\beta_{i,i}\mathbf{S}_{i,i}\mathbf{Q}_i(t)\mathbf{S}_{i,i}^H\mathbf{T}_i^{-1}(t)\Bigg|\Bigg\} \tag{37}$$

where $\mathbf{S} \triangleq \left[\mathbf{S}_{1,1}^T,\ldots,\mathbf{S}_{L,L}^T\right]^T$, and

$$\mathbf{T}_i(t) = \sum_{j=1,j\neq i}^{L}\beta_{i,j}\mathbf{S}_{i,j}\mathbf{Q}_j(t)\mathbf{S}_{i,j}^H + \mathbf{I}_N.$$

When the channel covariance matrix $\mathbf{\Omega}_i$ is unknown at the transmitting nodes, a straightforward (suboptimal) space-only power schedule is $\mathbf{P}_i(t) = P_i/N\mathbf{I}_N$. By substituting (36) into (37), we have

$$I\left(\frac{P_1}{N}\mathbf{\Gamma}_N,\ldots,\frac{P_L}{N}\mathbf{\Gamma}_N\right) = E_{\mathbf{S}}\Bigg\{\sum_{i=1}^{L}\log_2\Bigg|\mathbf{I}_N$$
$$+\frac{\beta_{i,i}P_i}{N}\mathbf{S}_{i,i}\mathbf{\Gamma}_i\mathbf{S}_{i,i}^H\mathbf{F}_i^{-1}\Bigg|\Bigg\} \tag{38}$$

where

$$\mathbf{F}_i = \sum_{j=1,j\neq i}^{L}\frac{\beta_{i,j}P_j}{N}\mathbf{S}_{i,j}\mathbf{\Gamma}_j\mathbf{S}_{i,j}^H + \mathbf{I}_N. \tag{39}$$

From (38) and (39), we can see that in this case, the ergodic sum capacity is equivalent to that of the multiple MIMO links with i.i.d. fading channel and the source covariance matrices $P_i/N\mathbf{\Gamma}_i$, $i = 1,\ldots,L$. Thus, the impact of the channel correlation on the ergodic sum capacity for the space-only case can be analyzed as follows.
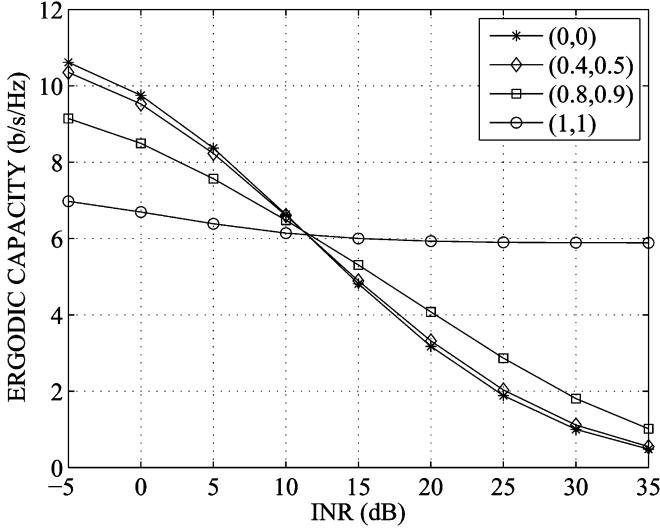
Fig. 4. Comparison of ergodic capacities of MIMO networks with channel correlation $\omega = (0,0)$, $\omega = (0.4, 0.5)$, $\omega = (0.8, 0.9)$, or $\omega = (1, 1)$. Here, $N = 2$, $L = 2$.

When the interference level is low such that it can be neglected, we have L parallel point-to-point MIMO links. It is well known [11] that for a single link with unknown CSI at the transmitter, the optimal source covariance matrix is a scaled identity matrix. If the channel is correlated, $\Gamma_i \neq c\mathbf{I}_N$, where c is a scalar, then the ergodic capacity is reduced by using Γ_i as the source covariance matrix. For the case of highly correlated channels, the values of the diagonal elements of Γ_i are widely spread. This may lead to a severe performance degradation of the ergodic sum capacity.

When the interference level is high, it has been shown in [6] that the optimal space-only power schedule is equivalent to a low rank power schedule. When the channel is highly correlated, Γ_i has the tendency to be rank-deficient. In this case, the channel correlation "assists" a low rank power schedule and hence could improve the ergodic sum capacity.

Fig. 4 shows a numerical example of the impact of channel correlation on the per-link ergodic capacity of distributed MIMO links with $L = 2$ active and symmetric links, where each node has $N = 2$ antennas and the space-only optimal power schedule is applied. Here, we set $P_i = 20$ dB, $\beta_{i,i} = 1$, $i = 1, \ldots, L$, and $\beta_{i,j} = \beta$, $i, j = 1, \ldots, L, j \neq i$. The exponential correlation model has been adopted to simulate the channel correlation, i.e., the (k, l)th element of matrix Ω_i is given by

$$[\Omega_i]_{k,l} = \omega_i^{|k-l|}, \quad 0 \leq \omega_i \leq 1,$$
$$i = 1, \ldots, L, \quad k, l = 1, \ldots, N. \quad (40)$$

We will use $\omega \triangleq (\omega_1, \ldots, \omega_L)$ to capture the channel correlation of all L links. Fig. 4 supports our analysis, i.e., the ergodic capacity decreases as the channel correlation increases when the interference is low, but the ergodic capacity increases as the channel correlation increases when the interference is high. In particular, the ergodic capacity becomes very robust against INR when the channels are fully correlated, i.e., when $\omega = (1, 1)$.

In the sequel of this section, we develop the space–time optimal power schedule for MIMO links with correlated channels. We assume that all channel covariance matrices of all links are known at each transmitting node.

The overall network ergodic sum capacity averaged over L time slots can be written as

$$I_a(\bar{\mathbf{Q}}) = \frac{1}{L} E_{\mathbf{S}} \left\{ \sum_{t=1}^{L} \sum_{i=1}^{L} \log_2 \left| \mathbf{I}_N + \beta_{i,i} \mathbf{S}_{i,i} \mathbf{Q}_i(t) \mathbf{S}_{i,i}^H \mathbf{T}_i^{-1}(t) \right| \right\}$$

where $\bar{\mathbf{Q}}$ is a matrix stacking the $\mathbf{Q}$ matrices of all links:

$$\bar{\mathbf{Q}} \triangleq \left[\bar{\mathbf{Q}}_1^T, \ldots, \bar{\mathbf{Q}}_L^T \right]^T$$
$$\bar{\mathbf{Q}}_i \triangleq \left[\mathbf{Q}_i^T(1), \ldots, \mathbf{Q}_i^T(L) \right]^T \quad i = 1 \ldots, L.$$

The space–time optimal power schedule can be formulated as the following constrained optimization problem:

$$\max_{\bar{\mathbf{Q}}} \quad I_a(\bar{\mathbf{Q}}) \quad (41)$$

$$\text{s.t.} \quad \frac{1}{L} \sum_{t=1}^{L} \text{tr} \left\{ \Gamma_i^{-1/2} \mathbf{Q}_i(t) \Gamma_i^{-1/2} \right\} = P_i \quad (42)$$

$$\mathbf{Q}_i(t) \geq 0, \quad i, t = 1, \ldots, L. \quad (43)$$

Note that the equality constraints in (42) follow from

$$\frac{1}{L} \sum_{t=1}^{L} \text{tr} \left\{ \Gamma_i^{-1/2} \mathbf{Q}_i(t) \Gamma_i^{-1/2} \right\} = \frac{1}{L} \sum_{t=1}^{L} \text{tr} \left\{ \mathbf{V}_i^H \mathbf{P}_i(t) \mathbf{V}_i \right\}$$
$$= \frac{1}{L} \sum_{t=1}^{L} \text{tr}\{\mathbf{P}_i(t)\} = P_i.$$

Let us denote $\mathbf{Q}_i(t) \triangleq \mathbf{U}_{\mathbf{Q}_i}(t) \mathbf{Z}_i(t) \mathbf{U}_{\mathbf{Q}_i}^H(t)$ as the eigendecomposition of $\mathbf{Q}_i(t)$, $i, t = 1, \ldots, L$, where $\mathbf{U}_{\mathbf{Q}_i}(t)$ is a unitary matrix of eigenvectors, and $\mathbf{Z}_i(t) = \text{diag}\{z_{i1}(t), z_{i2}(t), \ldots, z_{iN}(t)\}$ is an $N \times N$ diagonal matrix containing all the eigenvalues. Then, the following theorem is in order.

Theorem 1: The optimal solution $\mathbf{Q}_1^*(1), \ldots, \mathbf{Q}_L^*(L)$ of the problem (41)–(43) are diagonal matrices and given by $\mathbf{Q}_i^*(t) = \mathbf{Z}_i^*(t)$, $i, t = 1, \ldots, L$.

Proof: First, we note that since $\mathbf{S}_{i,j}$, $i, j = 1, \ldots, L$ have i.i.d. Gaussian distributed entries with zero mean and unit variance, the distribution of $\mathbf{S}_{i,j}$ is invariant under a unitary transform $\mathbf{S}_{i,j} \mathbf{U}_{\mathbf{Q}_j}(t)$. Thus, we have

$$I_a(\bar{\mathbf{Q}}) = I_a(\bar{\mathbf{z}}) \quad (44)$$

where

$$\bar{\mathbf{z}} \triangleq \left[\bar{\mathbf{z}}_1^T, \ldots, \bar{\mathbf{z}}_L^T \right]^T$$
$$\bar{\mathbf{z}}_i \triangleq \left[\mathbf{z}_i^T(1), \ldots, \mathbf{z}_i^T(L) \right]^T$$
$$\mathbf{z}_i(t) \triangleq [z_{i1}(t), z_{i2}(t), \ldots, z_{iN}(t)]^T \quad i, t = 1 \ldots, L.$$

Second, it has been shown in [17] using the theory of majorization [18] that if $\mathbf{Q}_i(t)$, $i, t = 1, \ldots, L$ satisfy the constraint (42), then $\mathbf{Z}_i(t)$ also satisfy the constraint

$$\frac{1}{L} \sum_{t=1}^{L} \text{tr} \left\{ \Gamma_i^{-1/2} \mathbf{Z}_i(t) \Gamma_i^{-1/2} \right\} = P_i,$$
$$\mathbf{Z}_i(t) \geq 0, \quad i, t = 1, \ldots, L. \quad (45)$$

From (44) and (45), it can be seen that $\mathbf{Q}_i^*(t) = \mathbf{Z}_i^*(t)$, $i, t = 1, \ldots, L$. $\qquad\qquad\square$

Using Theorem 1, we can see that since $\mathbf{Q}_i^*(t) = \mathbf{U}_{\mathbf{Q}_i}^*(t)\mathbf{Z}_i^*(t)(\mathbf{U}_{\mathbf{Q}_i}^*(t))^H = \mathbf{Z}_i^*(t)$ for any diagonal matrix $\mathbf{Z}_i^*(t)$, obviously, we have $\mathbf{U}_{\mathbf{Q}_i}^*(t) = \mathbf{I}_N$.

Substituting $\mathbf{Q}_i^*(t) = \mathbf{Z}_i^*(t)$ back into (36), we obtain the optimal $\mathbf{P}_i(t)$ as

$$\mathbf{P}_i^*(t) = \mathbf{V}_i\mathbf{\Gamma}_i^{-1/2}\mathbf{Z}_i^*(t)\mathbf{\Gamma}_i^{-1/2}\mathbf{V}_i^H, \quad i, t = 1, \ldots, L. \quad (46)$$

We can see from (46) that the eigenspace of the optimal source covariance matrix of each transmitting node is equal to that of the corresponding channel covariance matrix. In other words, the channel should be "whitened" by the source covariance matrices. This property greatly reduces the number of optimization variables from N^2L^2 to NL^2. Now the remaining task is to find the optimal $\bar{\mathbf{z}}$. It can be formulated as the following constrained optimization problem:

$$\max_{\bar{\mathbf{z}}} \quad I_a(\bar{\mathbf{z}}) \quad (47)$$

$$\text{s.t.} \quad \frac{1}{L}\sum_{t=1}^{L}\text{tr}\left\{\mathbf{\Gamma}_i^{-1/2}\mathbf{Z}_i(t)\mathbf{\Gamma}_i^{-1/2}\right\} = P_i \quad (48)$$

$$\mathbf{Z}_i(t) \geq 0, \quad i, t = 1, \ldots, L. \quad (49)$$

The objective function (47) can be further written as

$$I_a(\bar{\mathbf{z}}) = \frac{1}{L}\sum_{t=1}^{L}\sum_{i=1}^{L}\left(E_{\mathbf{S}}\left\{\log_2\left|\mathbf{T}_i(t) + \beta_{i,i}\mathbf{S}_{i,i}\mathbf{Z}_i(t)\mathbf{S}_{i,i}^H\right|\right\}\right.$$
$$\left. - E_{\mathbf{S}}\{\log_2|\mathbf{T}_i(t)|\}\right). \quad (50)$$

Comparing (50) with (18), we find that they have the same structure. Thus (50) can be written in the same closed form as (20) by following the computation steps listed in Section III. Since the constraints (48)–(49) are simple linear constraints, the problem (47)–(49) can be solved by the projected gradient technique developed in Section IV, and hence more details are omitted.

VI. NUMERICAL EXAMPLES

We now illustrate the performance of the space–time power scheduling scheme presented earlier. For comparison in the case of i.i.d. channel, we will consider the following five schemes.

- Scheme A1: Space–time optimal power schedule, which is based on (21) and (22).
- Scheme A2: Orthogonal isotropic power schedule, where all links conduct transmissions in orthogonal channels (such as different times and/or different frequencies), and the source covariance matrix used by all links is proportional to the identity matrix.
- Scheme A3: Uniform isotropic power schedule, where all links conduct transmissions at the same time and the same frequency, and the source covariance matrix used by all links is proportional to the identity matrix.
- Scheme A4: Space-only optimal power schedule, which is based on (31) and (32).
- Scheme A5: Low rank power schedule [6], where each link uses a low rank source covariance matrix where the corresponding ranks for L links are denoted by the string of integers $(r_1, r_2, \ldots, r_L)$. The ith link with the rank r_i uses
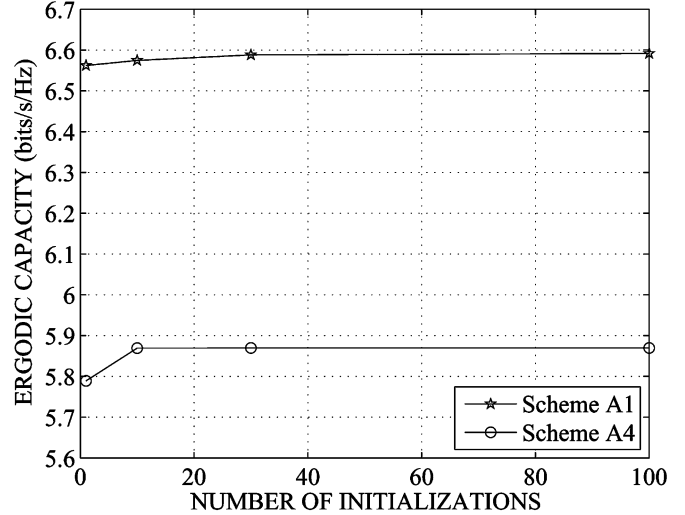

Fig. 5. Ergodic capacities of schemes A1 and A4 versus the number of initializations. Here, $N = 2$, $L = 2$, SNR = 20 dB, INR = 30 dB.

a power vector $\mathbf{d}_i$ of r_i nonzero equal entries and $N - r_i$ zero entries.

For the correlated channel case, we compare the following four schemes.

- Scheme B1: Space–time optimal power schedule, which is based on (46)–(49).
- Schemes B2–B4: These approaches are similar to Schemes A2, A3, and A5, respectively. However, for a fair comparison, in schemes B2, B3, and B4, the design of the source covariance matrices takes into account the channel covariance matrices.

Schemes A1, A2, B1, and B2 are space–time based, and all other schemes are space-only based. Schemes A1 and B1 are space–time optimal, while Schemes A2 and B2 are not.

For Schemes A1, A4, and B1, the following parameters were applied: $s = 1$, $\sigma = 0.1$, $\theta = 0.5$, and $\varepsilon = 10^{-4}$.

For all examples, we consider a symmetric and circular network where all nodes are uniformly located on a circle. Fig. 1 shows an example of such a network with $L = 3$ active links. For each case to be considered, we set the same transmit power for all transmitting nodes $P_i = P$, $i = 1, \ldots, L$, the same gain for all desired links $\beta_{i,i} = \rho$, $i = 1, \ldots, L$, and the same gain for all interfering links $\beta_{i,j} = \beta$, $i, j = 1, \ldots, L, j \neq i$. Thus, we have the same averaged SNR for all desired links $A_i = P\rho$, $i = 1, \ldots, L$, and the same nominal INR for all interfering links $B_{i,j} = P\beta$, $i, j = 1, \ldots, L, j \neq i$. For all simulations of ergodic capacities versus INR, we set $P = 20$ dB and $\rho = 1$. The INR is changed by varying β. Similarly, for all simulations of ergodic capacities versus SNR, we take $P = 10$ dB and $\beta = 1$. Then we change ρ to vary SNR. Since the nominal INR and the averaged SNR are quasi static, we assume that their values can be obtained with relatively high precision. The ergodic capacity shown in all figures is a per-link ergodic capacity.

For Schemes A1, A4, and B1, the power allocation vectors were initialized randomly. For each simulation point, a certain number of initializations were tried and the best result was chosen. Fig. 5 shows the ergodic capacities of Schemes A1 and A4 for the i.i.d. channel with $N = 2$, $L = 2$, SNR = 20 dB,

TABLE II
AVERAGE NUMBER OF ITERATIONS; $N = 2, L = 2$, SNR = 20 dB

INR (dB)	-5	0	5	10	15	20	25	30	35
Scheme A1	7	11	20	187	10	10	11	14	15
Scheme A4	6	8	14	87	14	15	17	18	20

Fig. 6. Comparison of ergodic capacities of schemes A1–A5; i.i.d. channel. Scheme A5 uses $(r_1, r_2) = (1, 1)$. Here, $N = 2, L = 2$.

Fig. 7. Comparison of ergodic capacities of schemes A1–A5; i.i.d. channel. Scheme A5 uses $(r_1, r_2) = (1, 1)$. Here, $N = 2, L = 2$.
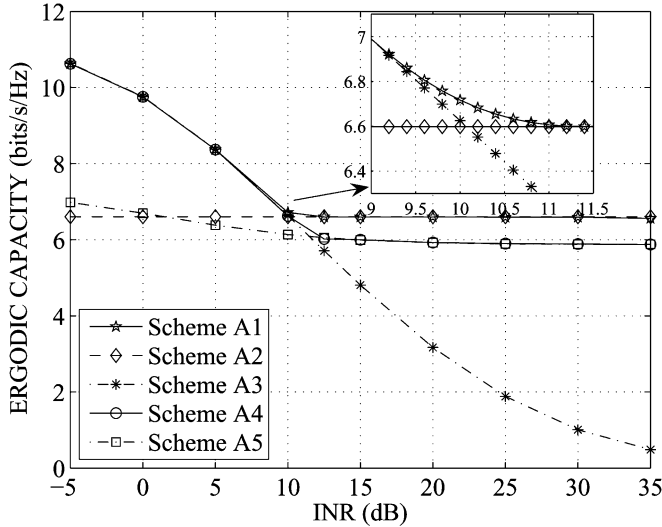

Fig. 8. Comparison of ergodic capacities of schemes A1–A5; i.i.d. channel. Scheme A5 uses $(r_1, r_2) = (1, 1)$. Here, $N = 3, L = 2$.

INR = 30 dB, and different number of initial points. From this figure, one can see that the difference caused by the number of initializations is not very large. In fact, after 30 initializations, the difference is almost zero. Thus, in the sequel, we use 30 initializations for Schemes A1, A4, and B1.

The number of iterations required by Schemes A1, A4 and B1 for convergence depends on many factors such as N, L, SNR, INR, and the choice of the initial point. As an example, in Table II we list the average number of iterations for convergence of Schemes A1 and A4 over 30 random initializations for different INR levels at $N = 2$, $L = 2$, and SNR = 20 dB. It is observed that the number of iterations required at INR = 10 dB is relatively large. It must be due to the insensitivity of the objective function around the optimal covariance matrices at this particular INR value. This INR value also coincides with the sharp transition point of the sum capacity of Scheme A1 (see Fig. 6).

In general, the number of iterations for convergence increases with the number of optimization variables (i.e., NL^2). However, an exact relationship of the number of iterations with respect to the number of links is difficult to obtain.

The performance comparison of Schemes A1–A5 for the i.i.d. channel environment are shown in Figs. 6–10, while the performance of Schemes B1–B4 for the correlated channel scenario are compared in Figs. 11–15. Fig. 6 compares the ergodic capacities of schemes A1–A5 with $L = 2$ and $N = 2$. From Fig. 6, we can see that Scheme A2 performs the same as Scheme A1 at high INR, and Scheme A3 has the same performance as Scheme A1 at low INR. More interestingly, the transition of the optimality from Scheme A2 to Scheme A3 along the INR axis is very sharp (within about 1.5 dB of INR). This property of
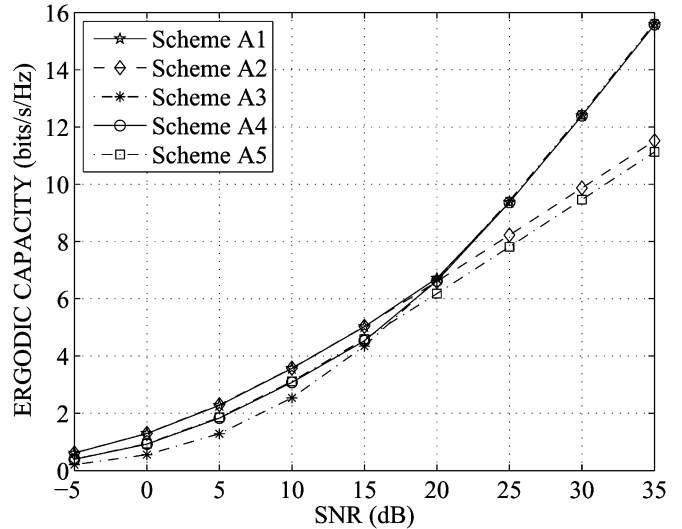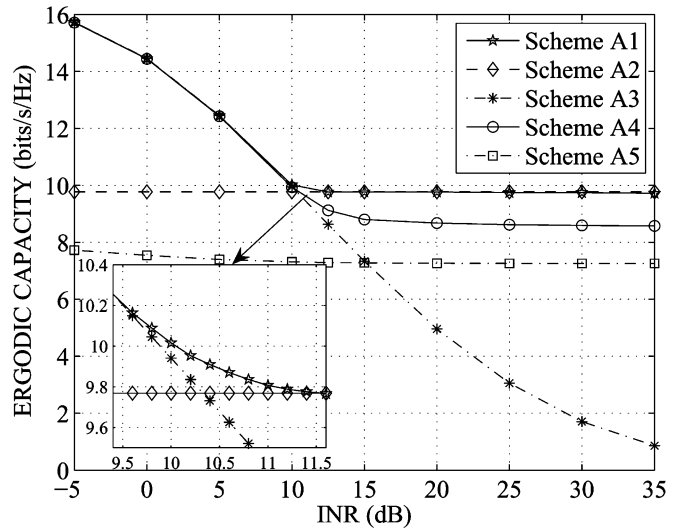
Schemes A2 and A3 is also observed in Figs. 8 and 9. The ergodic capacities versus SNR for the same simulation setups at INR = 10 dB is shown in Fig. 7. Similarly, we observe the optimality of Scheme A1 in the whole SNR region.

Fig. 8 compares the ergodic capacities of Schemes A1–A5 with $L = 2$ and $N = 3$. Here, Scheme A5 with $(r_1, r_2) = (1, 1)$ remains strongly suboptimal even compared with Scheme A4 over the whole range of INR. This is because under $N = 3$, there are effectively three independent streams. But Scheme A5 with $(r_1, r_2) = (1, 1)$ uses only two.

Fig. 9 compares the ergodic capacities of Schemes A1–A5 with $L = 3$ and $N = 2$. Here, the capacity of Scheme A5 with $(r_1, r_2, r_3) = (1, 1, 1)$ decreases as INR increases. This is because it over uses the total available (two) streams at high INR.

Fig. 10 compares Schemes A3, A4, and A5, which are all space-only schemes. Here, $N = 3$ and $L = 2$. Scheme A5
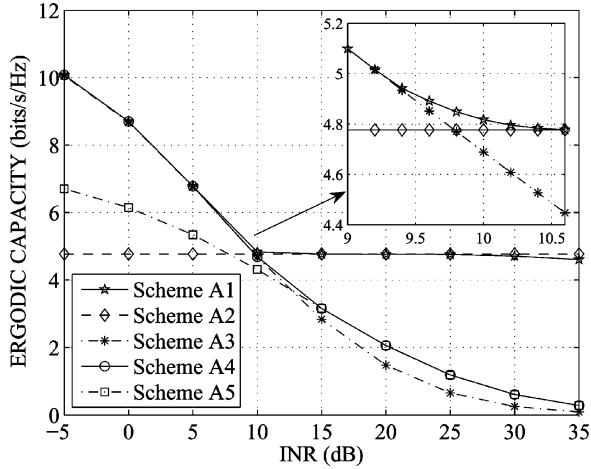
Fig. 9. Comparison of ergodic capacities of schemes A1–A5; i.i.d. channel. Scheme A5 uses $(r_1, r_2, r_3) = (1, 1, 1)$. Here, $N = 2, L = 3$.
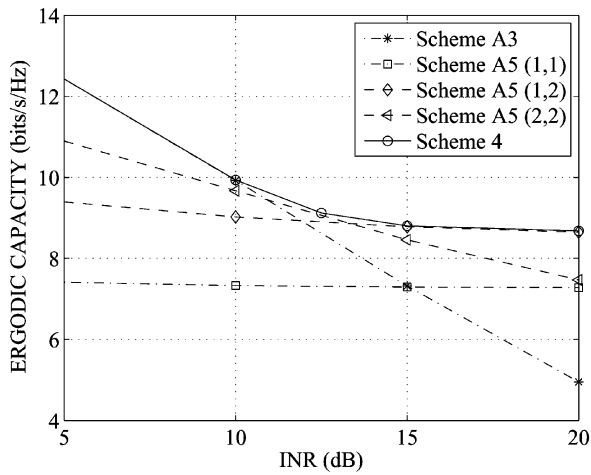

Fig. 10. Comparison of ergodic capacities of Schemes A3, A4, and A5; i.i.d. channel. Scheme A5 uses $(r_1, r_2) = (1, 1)$, $(r_1, r_2) = (1, 2)$, or $(r_1, r_2) = (2, 2)$. Here, $N = 3, L = 2$.

Fig. 11. Comparison of ergodic capacity of schemes B1–B4; correlated channel with $\boldsymbol{\omega} = (0.8, 0.9)$. Scheme B4 uses $(r_1, r_2) = (1, 1)$. Here, $N = 2, L = 2$.

Fig. 12. Comparison of ergodic capacity of schemes B1–B4; correlated channel with $\boldsymbol{\omega} = (0.8, 0.9)$. Scheme B4 uses $(r_1, r_2) = (1, 1)$. Here, $N = 2, L = 2$.

with different rank conditions is illustrated. Given $N = 3$ and $L = 2$, Scheme A5 with $(r_1, r_2) = (1, 2)$ performs the same as Scheme A4 at high INR. When $\sum_i^L r_i < N$, Scheme A5 remains strongly suboptimal compared with Scheme A4 at all INR.

Recall the optimality property of Schemes A2 and A3 that Scheme A2 performs the same as Scheme A1 when INR is larger than a threshold and Scheme A3 has the same performance as Scheme A1 when INR is less than the threshold. We can determine the threshold INR value B^* by solving the following nonlinear equation:

$$J_1\left(\frac{B^*}{N}, \frac{A}{N}, N, L\right) = \frac{J_2\left(\frac{AL}{N}, N, N\right)}{L} \qquad (51)$$

which is explained next.

In (51), $A \triangleq P\rho$, $J_2(\nu, p, q)$ denotes the ergodic sum capacity of a $p \times q$ MIMO link with SNR ν. Hence, $J_2(AL/N, N, N)/L$
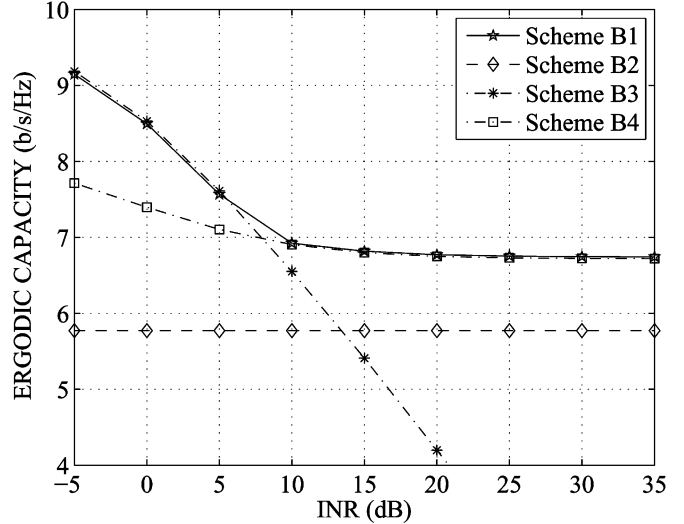
is the ergodic sum capacity using Scheme A2. Using the result in [19], $J_2(\nu, p, q)$ can be written as

$$J_2(\nu, p, q)$$
$$= \log_2(e) e^{1/\nu} \sum_{i=0}^{p-1} \sum_{j=0}^{i} \sum_{l=0}^{2j} \left\{ \frac{(-1)^l (2j)! (q - p + l)!}{2^{2i-l} j! l! (q - p + j)!} \right.$$
$$\left. \binom{2i - 2j}{i - j} \binom{2j + 2q - 2p}{2j - l} \sum_{k=0}^{q-p+l} S_{k+1}\left(\frac{1}{\nu}\right) \right\}.$$

Here, $S_n(x) \triangleq \int_1^{\infty} e^{-x\tau} \tau^{-n} d\tau$, $n = 0, 1, 2, \ldots$, is the exponential integral function of order n [12, eq. 5.1.4] and can be calculated using its recurrence relation

$$S_n(x) = \frac{e^{-x} - x S_{n-1}(x)}{n - 1}.$$
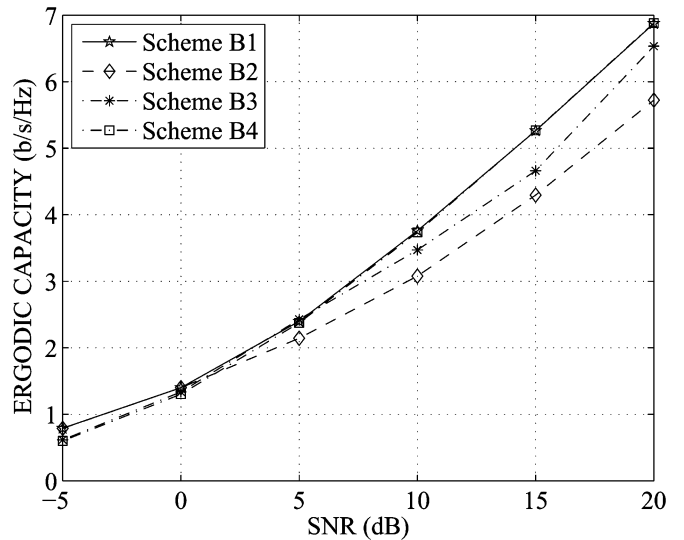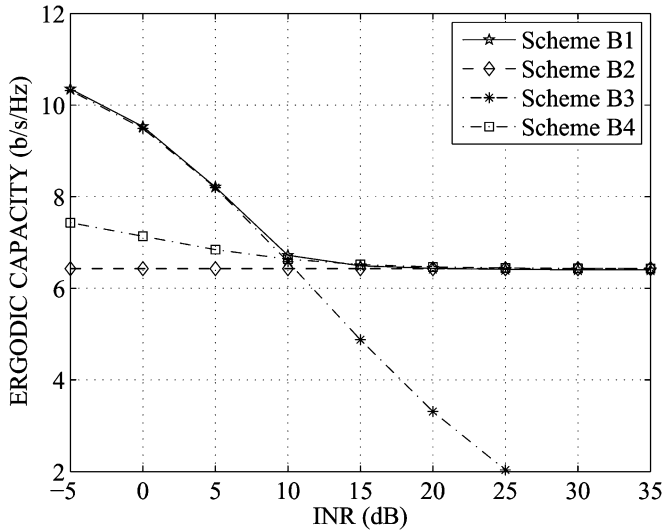
Fig. 13. Comparison of ergodic capacity of schemes B1–B4; correlated channel with $\boldsymbol{\omega} = (0.4, 0.5)$. Scheme B4 uses $(r_1, r_2) = (1, 1)$. Here, $N = 2, L = 2$.

Fig. 14. Comparison of ergodic capacity of schemes B1–B4; correlated channel with $\boldsymbol{\omega} = (0.7, 0.8, 0.9)$. Scheme B4 uses $(r_1, r_2) = (1, 1, 1)$. Here, $N = 2, L = 3$.

Also, $p \triangleq \min\{N_T, N_R\}$ and $q \triangleq \max\{N_T, N_R\}$, where N_T and N_R are the number of transmit and receive antennas, respectively. Thus, $p \leq q$.

In (51), $J_1(B^*/N, A/N, N, L)$ is the ergodic sum capacity of Scheme A3. A complete form of $J_1(\kappa, \nu, N, L)$ is as follows:

$$J_1(\kappa, \nu, N, L) = \log_2(e) |\boldsymbol{\Delta}|^{-1} \sum_{i=0}^{N-1} \frac{1}{i!} \left| \boldsymbol{\Delta}_{1:N(L-1)+i} \right.$$
$$\left. \times \boldsymbol{\xi}_i \boldsymbol{\Delta}_{N(L-1)+i+2:NL} \right| - J_2(\nu, N, NL - N)$$

where $\boldsymbol{\Delta}$ is a matrix whose (p, q)th element is given by

$$\Delta_{p,q} = \begin{cases} \left(\nu^{q-1}\right)^{[p-1]} & p = 1, \ldots, N \\ \left(\kappa^{q-1}\right)^{[p-N-1]} & p = N+1, \ldots, NL. \end{cases}$$

$\boldsymbol{\xi}_i$ is a column vector whose pth entry is given by (52), shown at the bottom of the page, and $\boldsymbol{\Delta}_{i:j}$ stands for the ith to jth columns of matrix $\boldsymbol{\Delta}$. Here $(\cdot)^{[i]}$ denotes the derivative of order i.

Since $g(B) = J_1(B/N, A/N, N, L) - J_2(AL/N, N, N)/L$ is a monotonically decreasing function of B, thus (51) can be easily solved, for instance, by the bisection method [14, pp. 145]. As an example, we applied (51) to the case of Fig. 9, and obtained $B^* = 9.78$ dB, which agrees well with the simulation result in Fig. 9.

In practice, the threshold INR B^* can be tabulated for different network parameters such as the number of links, the number of antennas of each node, and the SNR. Once this table is available, it can be looked up in real time to determine
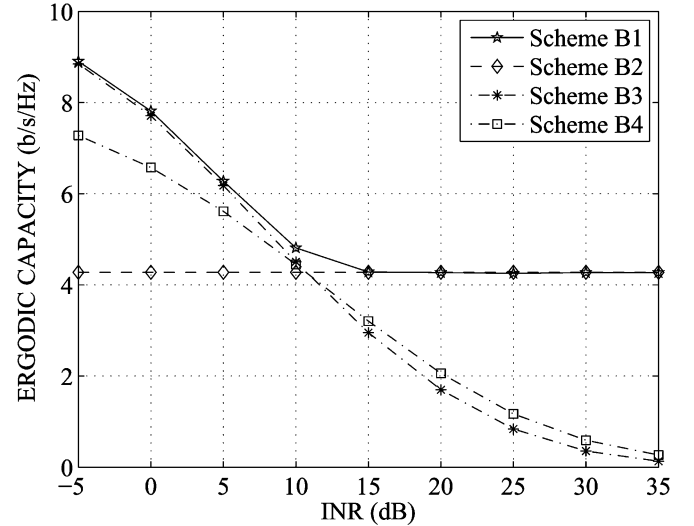
whether each node should be scheduled under Scheme A2 or Scheme A3.

In the following, we simulate the environment with correlated MIMO channels using the exponential correlation model (40). Fig. 11 compares the ergodic capacities of Schemes B1–B4 with $L = 2$ and $N = 2$. Here, the channel is set to be highly correlated with $\boldsymbol{\omega} = (0.8, 0.9)$. For Scheme B4, we have $(r_1, r_2) = (1, 1)$. It can be seen from Fig. 11 that Scheme B4 outperforms Scheme B2. Note that Scheme B2 is equivalent to TDMA. We now see that for correlated channels, TDMA is no longer optimal at high INR. That is, the channel correlation can change the optimality of a particular power schedule scheme (exact for the space–time optimal schemes A1 and B1). It can also be observed that Scheme B1 consistently provides the optimal ergodic capacity throughout the whole INR region. The optimality of Scheme B1 is also observed in Figs. 13–15.

The ergodic capacities versus SNR for the same simulation setups at INR = 10 dB is shown in Fig. 12. Similarly, we observe the optimality of Scheme B1 in the whole SNR region.

Fig. 13 illustrates the performance of Schemes B1–B4 with $L = 2$, $N = 2$, and low channel correlation $\boldsymbol{\omega} = (0.4, 0.5)$. Comparing Fig. 13 with Fig. 11, it can be seen that the ergodic capacity of Scheme B2 increases when the channel is less correlated, while the performance of Scheme B4 with $(r_1, r_2) = (1, 1)$ degrades. The reason is that when the channel is less correlated, $\boldsymbol{\Gamma}_i$ tends to be an identity matrix. Thus, Scheme B2 is close to be optimal in high interference

$$(\xi_i)_p = \begin{cases} \left(\nu^{N(L-1)-1} \int_0^\infty \ln(1+x) x^i e^{-x/\nu}\right)^{[p-1]} & p = 1, \ldots, N \\ \left(\kappa^{N(L-1)-1} \int_0^\infty \ln(1+x) x^i e^{-x/\kappa}\right)^{[p-N-1]} & p = N+1, \ldots, NL \end{cases} \qquad (52)$$

region. And Scheme B4 becomes suboptimal when the channel correlation is low. Interestingly, for $\boldsymbol{\omega} = (0.4, 0.5)$, Schemes B2 and B4 happen to yield a similar ergodic capacity in the high interference region.

Fig. 14 compares the ergodic capacities of Schemes B1–B4 with $L = 3$, $N = 2$, and $\boldsymbol{\omega} = (0.7, 0.8, 0.9)$. We find that in the presence of strong interference, Scheme B2 performs the same as Scheme B1. The capacity of Scheme B4 with $(r_1, r_2, r_3) = (1, 1, 1)$ decreases as INR increases, because it over uses the total available streams at high INR.

Finally, Fig. 15 shows the ergodic capacities of Schemes B1–B4 with $L = 2$, $N = 3$, and $\boldsymbol{\omega} = (0.8, 0.9)$. Interestingly, we can observe that Scheme B4 consistently outperforms Scheme B2. Here, the optimality result is different from that of the i.i.d. channel environment, where the orthogonal isotropic power schedule outperforms the low rank power schedule (see Fig. 8). In contrast to Fig. 11, Scheme B4 is suboptimal even in the high interference region. This fact demonstrates the importance of Scheme B1. The optimal ergodic capacity can not be achieved by a "prefixed" power

$$\frac{\partial I_a(\bar{\mathbf{d}})}{\partial d_{lm}(t)} = \frac{\log_2(e)}{L} \sum_{n=0}^{N-1} \sum_{k=1}^{NL} \left(c'_{tlkn}Q(n, \lambda_{t,l,k}) + c_{tlkn}Q'(n, \lambda_{t,l,k})\right) + \frac{\log_2(e)}{L} \sum_{i=1, i\neq l}^{L}$$
$$\left[\sum_{n=0}^{N-1} \sum_{k=1}^{NL} \left(c'_{tikn}Q(n, \lambda_{t,i,k}) + c_{tikn}Q'(n, \lambda_{t,i,k})\right) - \sum_{n=0}^{N-1} \sum_{k=1}^{N(L-1)} \left(u'_{tikn}Q(n, \bar{\lambda}_{t,i,k}) + u_{tikn}Q'(n, \bar{\lambda}_{t,i,k})\right)\right]$$

$$c'_{tikn} = \begin{cases} \frac{(-1)^{N-n-1}((L-1)N-1)}{n!} \lambda_{t,i,k}^{(L-1)N-2} \beta_{i,l} \left(\prod_{h\neq k}(\lambda_{t,i,k} - \lambda_{t,i,h})\right)^{-1} b_{tikn} \\ -\frac{(-1)^{N-n-1}}{n!} \lambda_{t,i,k}^{(L-1)N-1} \beta_{i,l} \sum_{j=1, j\neq k}^{NL} \left[\left(\prod_{h\neq k,j}(\lambda_{t,i,k} - \lambda_{t,i,h})\right)^{-1} (\lambda_{t,i,k} - \lambda_{t,i,j})^{-2}\right] b_{tikn}, \quad k = q \quad (53) \\ \frac{(-1)^{N-n-1}}{n!} \lambda_{t,i,k}^{(L-1)N-1} \beta_{i,l} \left(\prod_{h\neq k}(\lambda_{t,i,k} - \lambda_{t,i,h})\right)^{-1} \left[(\lambda_{t,i,k} - \lambda_{t,i,q})^{-1} b_{tikn} + a_{tikq}\right], \quad k \neq q \end{cases}$$

$$u'_{tikn} = \begin{cases} \frac{(-1)^{N-n-1}((L-2)N-1)}{n!} \bar{\lambda}_{t,i,k}^{(L-2)N-2} \beta_{i,l} \left(\prod_{h\neq k}\left(\bar{\lambda}_{t,i,k} - \bar{\lambda}_{t,i,h}\right)\right)^{-1} \bar{b}_{tikn} - \frac{(-1)^{N-n-1}}{n!} \bar{\lambda}_{t,i,k}^{(L-2)N-1} \\ \times \beta_{i,l} \sum_{j=1, j\neq k}^{N(L-1)} \left[\left(\prod_{h\neq k,j}\left(\bar{\lambda}_{t,i,k} - \bar{\lambda}_{t,i,h}\right)\right)^{-1} \left(\bar{\lambda}_{t,i,k} - \bar{\lambda}_{t,i,j}\right)^{-2}\right] \bar{b}_{tikn}, \quad k = \bar{q} \quad (54) \\ \frac{(-1)^{N-n-1}}{n!} \bar{\lambda}_{t,i,k}^{(L-2)N-1} \beta_{i,l} \left(\prod_{h\neq k}\left(\bar{\lambda}_{t,i,k} - \bar{\lambda}_{t,i,h}\right)\right)^{-1} \left[\left(\bar{\lambda}_{t,i,k} - \bar{\lambda}_{t,i,q}\right)^{-1} \bar{b}_{tikn} + \bar{a}_{tik\bar{q}}\right], \quad k \neq \bar{q} \end{cases}$$

$$Q'(n, \lambda_{t,i,k}) = \begin{cases} \sum_{r=0}^{n} \frac{n!(-1)^{(n-r)}}{(n-r)!} \lambda_{t,i,k}^{r} \beta_{i,l} \left[(r+1)e^{1/\lambda_{t,i,k}} S_1\left(\frac{1}{\lambda_{t,i,k}}\right) - \frac{1}{\lambda_{t,i,k}} e^{1/\lambda_{t,i,k}} S_1\left(\frac{1}{\lambda_{t,i,k}}\right) + 1\right] \\ + \sum_{r=1}^{n} \sum_{s=0}^{r-1} \sum_{h=0}^{r-s-1} \frac{n!(-1)^{(n-r)}(h+s+2)\lambda_{t,i,k}^{h+s+1}\beta_{i,l}}{(n-r)!(r-s-1-h)!(r-s)}, \qquad k = q \\ 0, \qquad k \neq q \end{cases} \quad (55)$$

$$Q'(n, \bar{\lambda}_{t,i,k}) = \begin{cases} \sum_{r=0}^{n} \frac{n!(-1)^{(n-r)}}{(n-r)!} \bar{\lambda}_{t,i,k}^{r} \beta_{i,l} \left[(r+1)e^{1/\bar{\lambda}_{t,i,k}} S_1\left(\frac{1}{\bar{\lambda}_{t,i,k}}\right) - \frac{1}{\bar{\lambda}_{t,i,k}} e^{1/\bar{\lambda}_{t,i,k}} S_1\left(\frac{1}{\bar{\lambda}_{t,i,k}}\right) + 1\right] \\ + \sum_{r=1}^{n} \sum_{s=0}^{r-1} \sum_{h=0}^{r-s-1} \frac{n!(-1)^{(n-r)}(h+s+2)\bar{\lambda}_{t,i,k}^{h+s+1}\beta_{i,l}}{(n-r)!(r-s-1-h)!(r-s)}, \qquad k = \bar{q} \\ 0, \qquad k \neq \bar{q}. \end{cases} \quad (56)$$

$$\bar{q} = \begin{cases} (l-2)N + m, & i < l \\ (l-1)N + m, & i > l. \end{cases}$$

$$a_{tikq} = \begin{cases} \sum_{1 \leq j_1 < \cdots < j_{N-n-2} \leq NL}^{j_r \neq k, q} \lambda_{t,i,j_1} \ldots \lambda_{t,i,j_{N-n-2}}, & n = 0, \ldots, N-3 \\ 1, & n = N-2 \\ 0, & n = N-1. \end{cases}$$

$$\bar{a}_{tik\bar{q}} = \begin{cases} \sum_{1 \leq j_1 < \cdots < j_{N-n-2} \leq N(L-1)}^{j_r \neq k, \bar{q}} \bar{\lambda}_{t,i,j_1} \ldots \bar{\lambda}_{t,i,j_{N-n-2}}, & n = 0, \ldots, N-3 \\ 1, & n = N-2 \\ 0, & n = N-1. \end{cases}$$
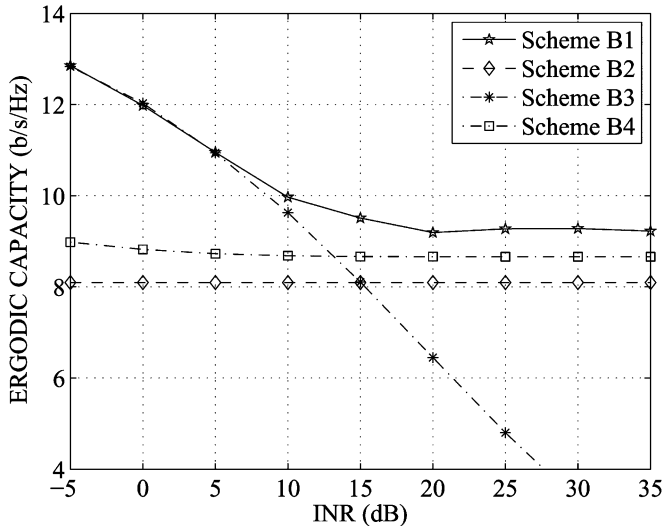
Fig. 15. Comparison of ergodic capacity of schemes B1–B4; correlated channel with $\boldsymbol{\omega} = (0.8, 0.9)$. Scheme B4 uses $(r_1, r_2) = (1, 1)$. Here, $N = 3, L = 2$.

scheduling schemes. Our schemes A1 and B1 are adaptive to the network conditions.

VII. CONCLUSION

We have proposed a space–time optimal power scheduling approach for multiple distributed MIMO links assuming no instantaneous CSI at the transmitting nodes. For the i.i.d. and correlated channel environments, this approach leads to Scheme A1 and Scheme B1, respectively. In particular, for a symmetric network of i.i.d. channels, using Scheme A1 as the optimal benchmark, we have observed that Scheme A2, an orthogonal isotropic power schedule (such as TDMA), is optimal when the INR is larger than a threshold, and Scheme A3, a uniform isotropic power schedule, is optimal when the INR is less than the same threshold. The threshold INR value can be computed based on the network topology and the transmit power level. For the correlated channel environment, we have shown that the channel correlation has a great impact on the ergodic capacity and the optimality of different power scheduling approaches. Future research should address issues arising from asymmetric networks.

APPENDIX A
SUMMARY OF NOTATIONS FOR SYSTEM MODEL

$\mathbf{H}_{i,j}$ MIMO channel between the ith transmitting node and the jth receiving node.

$\beta_{i,j}$ $i = j$, channel gain of a link-of-interest; $i \neq j$, channel gain of an interfering link.

$\mathbf{P}_i(t)$ Transmit covariance matrix of the ith transmitting node in the tth time slot.

$\mathbf{R}_i(t)$ Interference-plus-noise covariance matrix at the i receiving node in the tth time slot.

P_i Averaged transmit power from the ith transmitting node.

$\mathbf{x}_i(t)$ Signal vector transmitted from the ith transmitting node within time slot t.

$\mathbf{y}_i(t)$ Signal vector received by the ith destination node within time slot t.

I_a Averaged ergodic sum capacity.

I Ergodic sum capacity at one time slot.

$\boldsymbol{\Omega}_i$ Channel correlation matrix of the ith transmitting node.

APPENDIX B
THE GRADIENT OF (20)

See the first equation shown at the bottom of the previous page, where $(\cdot)'$ denotes the first order derivative, c'_{tikn}, u'_{tikn}, $Q'(n, \lambda_{t,i,k})$, and $Q'(n, \bar{\lambda}_{t,i,k})$ are given by (53)–(56), respectively, shown at the bottom of the previous page. Here, $q = (l - 1)N + m$.

REFERENCES

[1] Y. Hua, Y. Huang, and J. J. Garcia-Luna-Aceves, "Maximizing the throughput of large ad hoc wireless networks," *IEEE Signal. Process. Mag.*, vol. 23, no. 5, pp. 84–94, Sep. 2006.

[2] K. Hong and Y. Hua, "Throughput analysis of large wireless networks with regular topologies," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, 2007, Article ID 26760.

[3] B. Zhao and Y. Hua, "A distributed MAC scheme for large network of wireless routers," in *Proc. 25th Army Sci. Conf.*, Orlando, FL, Nov. 27–30, 2006.

[4] M. F. Demirkol and M. A. Ingram, "Power-controlled capacity for interfering MIMO links," in *Proc. IEEE Vehicular Technol. Conf. (VTC)*, Atlantic City, NJ, Oct. 2001, vol. 1, pp. 187–191.

[5] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[6] R. S. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun 2003.

[7] Y. Rong and Y. Hua, "Optimal power schedule for distributed MIMO links," *IEEE Trans. Wireless Commun.*, to be published.

[8] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.

[9] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.

[10] H. Gao and P. J. Smith, "A determinant representation for the distribution of quadratic forms in complex normal vectors," *J. Multivariate Anal.*, vol. 73, no. 2, pp. 155–165, May 2000.

[11] İ. E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[12] I. Abramowitz and M. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Washington, DC: U.S. Government Printing Office, 1972.

[13] D. P. Bertsekas, *Nonliner Programming*, 2nd ed. Bellmont, MA: Athena Scientific, 1995.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.

[16] D. Gesbert, H. Bölcskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[17] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receiver antennas with covariance feedback," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, Apr. 2001, pp. 2266–2270.

[18] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic, 1979.

[19] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.
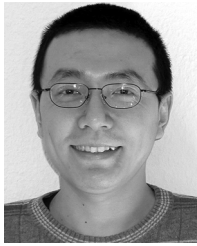
Yue Rong (S'03–M'06) was born in Jiangsu, China, in 1976. He received the B.E. degree from Shanghai Jiao Tong University, China, the M.Sc. degree from the University of Duisburg-Essen, Duisburg, Germany, and the Ph.D. degree (*summa cum laude*) from Darmstadt University of Technology, Darmstadt, Germany, all in electrical engineering, in 1999, 2002, and 2005, respectively.

From April 2001 to October 2001, he was a Research Assistant at the Fraunhofer Institute of Microelectronic Circuits and Systems, Duisburg, Germany. From October 2001 to March 2002, he was with Nokia, Ltd., Bochum, Germany. From November 2002 to March 2005, he was a Research Associate at the Department of Communication Systems in the University of Duisburg-Essen. From April 2005 to January 2006, he was with the Institute of Telecommunications at Darmstadt University of Technology, as a Research Associate. From February 2006 to November 2007, he was a Postdoctoral Researcher with the Department of Electrical Engineering, University of California, Riverside. Since December 2007, he has been with the Department of Electrical and Computer Engineering, Curtin University of Technology, Perth, Australia, as a Lecturer. His research interests include signal processing for communications, wireless communications, wireless networks, applications of linear algebra and optimization methods, and statistical and array signal processing.

Dr. Rong received the 2001–2002 Graduate Sponsoring Asia scholarship of DAAD/ABB (Germany) and the 2004 Chinese Government Award for Outstanding Self-Financed Students Abroad (China).
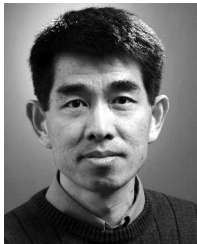
Yingbo Hua (S'86–M'88–SM'92–F'02) received the B.S. degree from the Nanjing Institute of Technology (Southeast University), Nanjing, China, in 1982 and the Ph.D. degree from Syracuse University, Syracuse, NY, in 1988.

He was a Professor with the University of Melbourne, Australia, from 1990 to 2001, and took sabbatical leave with Hong Kong University of Science and Technology during 1999–2000, and Microsoft Research, WA, in summer 2000. In February 2001, he joined the University of California, Riverside. He has guided 20 Ph.D. students, seven Master's-by-research students, and 11 postdoctoral fellows. He has published over 260 research articles in the fields of signal processing, system identification, wireless communications, and ad hoc networks. He is a co-editor of the books *Signal Processing Advances in Wireless and Mobile Communications* (Prentice-Hall, 2001) and *High-Resolution and Robust Signal Processing* (Marcel Dekker, 2003).

Prof. Hua has served on the Editorial Boards of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1994–1997), the IEEE SIGNAL PROCESSING LETTERS (1998–2002), the *EURASIP Journal on Signal Processing* (2005 to present), and the *IEEE Signal Processing Magazine* (2007 to present). He has been a member on several IEEE Signal Processing Society's Technical Committees and many international conference committees. He served as Chair of Executive Committee of Communications and Networking of University of California Industry and University Cooperative Research Program.

Ananthram Swami (S'79–M'79–SM'96–F'08) received the B.Tech. degree from Indian Institute of Technology (IIT), Bombay, India; the M.S. degree from Rice University, Houston, TX; and the Ph.D. degree from the University of Southern California (USC), all in electrical engineering.

He has held positions with Unocal Corporation, USC, CS-3, and Malgudi Systems. He was a Statistical Consultant to the California Lottery, developed a Matlab-based toolbox for non-Gaussian signal processing, and has held visiting faculty positions at INP, Toulouse, France. He is currently with the U.S. Army Research Laboratory, Adelphi, MD, where his work is in the broad area of signal processing, wireless communications and networking, including both sensor networks and MANETs.

Dr. Swami has been an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2002. He has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS (2002–2003), the IEEE TRANSACTIONS ON CIRCUITS & SYSTEMS—PART II (2002–2003), the *IEEE Signal Processing Magazine* (2002–2005), and the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1994–1996 and 2004–2007). He was Member (1998–2005), Vice-Chair (2002–2004), and Chair (2005–2006) of the IEEE Signal Processing Society (SPS) Technical Committee on Signal Processing for Communications and a Member of the IEEE SPS Statistical Signal and Array Processing (SSAP) Technical Committee from 1993 to 1998. He was Co-Organizer and Co-Chair of the 1993 IEEE SPS Higher Order Statistics (HOS) Workshop, the 1996 IEEE SPS SSAP Workshop, and the 1999 ASA-IMA Workshop on Heavy-Tailed Phenomena. He has served as Member of the Technical Committee for numerous workshops and conferences and is actively involved in the IEEE Military Communications Conference (MILCOM). He was Co-Guest Editor of the *IEEE Signal Processing Magazine* 2004 Special Issue on Signal Processing for Networking, the *IEEE Signal Processing Magazine* 2006 Special Issue on Distributed Signal Processing in Sensor Networks, the *EURASIP Journal on Advances in Signal Processing* 2006 Special Issue on Reliable Communications over Rapidly Time-Varying Channels, the *EURASIP Journal on Wireless Communications and Networking* 2007 Special Issue on Wireless Mobile Ad Hoc Networks, and the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING 2008 Special Issue on Signal Processing and Networking for Dynamic Spectrum Access.

A. Lee Swindlehurst (S'83–M'84–SM'99–F'04) received the B.S. (*summa cum laude*) and M.S. degrees from Brigham Young University, Provo, UT, in 1985 and 1986, respectively, and the Ph.D. degree from Stanford University, Stanford, CA, in 1991, all in electrical engineering.

From 1986 to 1990, he was employed at ESL, Inc., Sunnyvale, CA, where he was involved in the design of algorithms and architectures for several radar and sonar signal processing systems. He joined the faculty of the Department of Electrical and Computer Engineering at Brigham Young University in 1990, where he holds the position of Full Professor. During 1996–1997, he held a joint appointment as a visiting scholar at both Uppsala University, Uppsala, Sweden, and at the Royal Institute of Technology, Stockholm, Sweden. Currently, he is Vice President of Research for ArrayComm LLC, San Jose, CA. His research interests include sensor array signal processing for radar and wireless communications, detection and estimation theory, and system identification, and he has over 150 publications in these areas.

Dr. Swindlehurst is a past Secretary of the IEEE Signal Processing Society. He is currently serving as Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, as a member of the Editorial Boards for the *EURASIP Journal on Wireless Communications and Networking*, and the *IEEE Signal Processing Magazine*, and is a past Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is a recipient of several paper awards: the 2000 IEEE W. R. G. Baker Prize Paper Award, the 2006 IEEE Signal Processing Society's Best Paper Award, and the 2006 IEEE Communications Society Stephen O. Rice Prize in the Field of Communication Theory, and he is coauthor of a paper that received the IEEE Signal Processing Society Young Author Best Paper Award in 2001.