Large System Analysis of the Energy Consumption Distribution in Multi-User MIMO Systems with Mobility

Luca Sanguinetti, *Member, IEEE,* Aris L. Moustakas, *Senior Memebr, IEEE,*
Emil Björnson, *Member, IEEE,* and Merouane Debbah, *Fellow, IEEE*

Abstract—In this work, we consider the downlink of a single-cell multi-user MIMO system in which the base station (BS) makes use of N antennas to communicate with K single-antenna user equipments (UEs). The UEs move around in the cell according to a random walk mobility model. We aim at determining the energy consumption distribution when different linear precoding techniques are used at the BS to guarantee target rates within a finite time interval T. The analysis is conducted in the asymptotic regime where N and K grow large with fixed ratio under the assumption of perfect channel state information (CSI). Both recent and standard results from large system analysis are used to provide concise formulae for the asymptotic transmit powers and beamforming vectors for all considered schemes. These results are eventually used to provide a deterministic approximation of the energy consumption and to study its fluctuations around this value in the form of a central limit theorem. Closed-form expressions for the asymptotic means and variances are given. Numerical results are used to validate the accuracy of the theoretical analysis and to make comparisons. We show how the results can be used to approximate the probability that a battery-powered BS runs out of energy and also to design the cell radius for minimizing the energy consumption per unit area. The imperfect CSI case is also briefly considered.

Index Terms—Energy consumption, multi-user MIMO, downlink, linear pre-coding techniques, user mobility, random walk, Brownian motion, large system analysis, random matrix, central limit theorem.

I. INTRODUCTION

THE DATA traffic in cellular networks has been increasing exponentially for a long time and is expected to continue this trend, at least for the next five years [2]. Currently, one of the biggest challenges related to the traffic growth is the increasing energy consumption of the cellular infrastructure

equipments [3]. This means that the energy consumption must be taken into account from the outset when designing cellular networks of the future. This is particularly important when deploying base stations (BSs) in new rural regions of the world, where the electrical grid is unreliable or even non-existing. Off-grid deployments rely on combinations of diesel generators, batteries, and local energy harvesting (e.g., from solar panels) [3]. Since the supply of energy is either limited or fluctuates with the harvesting, it is of paramount importance to operate the BS such that it will not run out of energy (also known as power outage).

A lot of work has been done recently to better understand the energy consumption tradeoffs in cellular networks [4]. However, most of the analysis is carried out without taking into account the impact of user equipment (UE) mobility. This is a major issue since UE movements play a key role in determining the required energy in outdoor scenarios. Consider, for example, the case in which all UEs are positioned at the cell edge and assume that some target rate must be guaranteed at each UE. This would result into an increased power consumption, which might rapidly make the BS run out of energy if the UEs remain at the cell edge for an extended period of time (due to the lack of mobility). Therefore, it is of major importance to quantify not only how often such unlikely events happen, but also how long they last. All this depends on the UE mobility model and in particular on how their positions, velocities and accelerations change over time. A first attempt in this direction can be found in [5] in which the authors make use of stochastic geometry to model the energy consumption of a cellular network where each UE is connected to its closest single-antenna BS. In particular, a refined energy consumption model is proposed that includes the energy of broadcast messages, traffic activity, and user mobility.

The theory of random walks has been analyzed extensively in probability theory and it has been applied to a variety of fields. In a nutshell, it is a mathematical formalization of a path that consists of a succession of random steps [6]. If the time elapsed between two successive steps is relatively small (possibly zero), then the central limit theorem implies that the distribution of the corresponding path-valued process approaches that of a Brownian motion [7]. A more involved model (known as Lévy flight) corresponds to a random walk in which the step lengths have a probability distribution that is heavy tailed [6]. Another one that has been widely investigated is the so-called random waypoint process in which a mobile

user picks a random destination and travels with constant velocity to reach it. Although human mobility is rarely random, in [8], [9] the authors show that a Lévy flight contains some important statistical similarities with human mobility. This suggests that a random walk model might be considered as a good one to achieve a reasonable tradeoff between accuracy and analytical tractability [8], [9]. Despite the well-understood properties of the random mobility models above, not much progress has been made towards providing analytical results for the statistics of metrics in wireless communications.

Motivated by this lack, in this work we consider the downlink of a single-cell multi-user multiple-input multiple-output (MIMO) system. The BS is equipped with an array of N antennas and serves simultaneously K UEs by using linear precoding. The UEs are assumed to move around in the cell following (independent) random walks. Based on the random walk mobility pattern, we aim at determining the statistical distribution of the energy consumption required to guarantee a given set of UE data rates within a finite time interval of length T. The analysis is conducted in the asymptotic regime where N and K grow large with fixed ratio $c = K/N$ and it is applied to several linear precoding schemes, under the assumption of perfect channel state information (CSI) at the BS. In particular, we first focus on the asymptotic design of the precoder minimizing the transmit power at each time slot while satisfying the rate requirements [10], [11]. Differently from [12]–[14], the asymptotic transmit powers and beamforming vectors are computed using the approach adopted in [15], which provides us a much simpler means to overcome the technical difficulties arising with the application of standard random matrix theory tools (e.g. [13]). As already pointed out in [12]–[14], in the asymptotic regime the optimal values can be computed in closed-form through a nice and simple expression, which depends only on the average channel attenuations and rate requirements. The asymptotic analysis is then applied to design different heuristic linear precoding techniques, namely, maximum ratio transmission (MRT), zero-forcing (ZF), and regularized ZF (RZF) [16]–[19]. In particular, we concentrate on a RZF scheme (e.g. [17], [18]) that exploits knowledge of the average channel attenuations of UEs in designing the beamforming vectors and provide the closed-form expression for the optimal regularization parameter in the asymptotic regime.

The asymptotic analysis above is then used to provide a deterministic approximation of the energy consumption (which is asymptotically accurate in the large system limit) and to study its fluctuations around this value under the form of a central limit theorem. We compute closed-form expressions for the asymptotic means and variances of all considered schemes exploiting the connection between random walk and Brownian motion. It turns out that the fluctuations induced by the randomness in the UE positions make the variance of the energy decrease as $1/K$ for all considered schemes. This makes the variability induced by the small-scale fading less important for the analysis since it scales as $1/K^2$ [1], [20]. The analytical expressions are validated by means of extensive simulations and are shown to closely match the numerical results for different settings. Finally, we exemplify

how the new statistical characterization can be used to compute the probability that a battery-powered BS runs out of energy within a certain time period and also to design the cell radius for minimizing the energy consumption per unit area.

The remainder of this paper is organized as follows. [1] The next section introduces the system model and formulates the problem under investigation. Section III revisits the optimal linear precoder design in both finite and asymptotic regimes. Section IV deals with the asymptotic analysis and design of MRT, ZF and RZF. All this is then used in Section V to prove that in the asymptotic regime the energy consumption for all considered schemes converges to a Gaussian random variable whose mean and variance are derived analytically. In Section VI, numerical results are used to validate the theoretical analysis and to make comparisons among different processing schemes. Some applications of the statistical characterization of the energy consumption are also exemplified in Section VI. In addition, the imperfect CSI case is briefly addressed for the ZF and RZF precoding schemes. Finally, some conclusions and discussions are drawn in Section VII.

II. SYSTEM MODEL AND PROBLEM STATEMENT

A. System model

We consider the downlink of a single-cell multi-user MIMO system in which the BS makes use of N antennas to communicate with K single-antenna UEs over the whole channel bandwidth of W Hz. Linear processing is used for data precoding and perfect CSI is assumed to be available at the BS. As mentioned above, the imperfect CSI case will be briefly addressed in Section VI-D. The K active UEs are moving around within the coverage cell $\mathcal{C}$ of area $|\mathcal{C}|$. We assume that the number of UEs K can increase arbitrarily, while the area $|\mathcal{C}|$ is maintained fixed. To simplify the computations, we assume that the cell is circular with radius R such that $|\mathcal{C}| = \pi R^2$ and that the BS is located in the centre of the cell with its N transmit antennas adequately spaced apart (such that the channel components to any UE are uncorrelated). The location of UE k at time t is denoted by $\mathbf{x}_k(t) \in \mathbb{R}^2$ (in meters) and it is computed with respect to the BS. We call $\mathbf{X}_k = \{\mathbf{x}_k(t); t \geq 0\}$ and model $\mathbf{X}_k$ and $\mathbf{X}_j$ for $k \neq j$ as obtained from identical and independent random walks constrained into the region $\mathcal{C}$ [21]. We let ℓ be the (average) size of a step of the random walk and let ξ be the corresponding (average) time elapsed between two successive steps. The (average) velocity of each user between consecutive steps is thus given by $\sim \ell/\xi$. We

---

[1] The following notation is used throughout the paper. The notations $\mathsf{E}_{\mathbf{z}}\{\cdot\}$, $\mathsf{VAR}_{\mathbf{z}}\{\cdot\}$ and $\mathsf{COV}_{\mathbf{z}}\{\cdot\}$ indicate that the expectation, variance and covariance are computed with respect to $\mathbf{z}$. The notation $||\cdot||$ stand for the Euclidean norm whereas $|\mathcal{S}|$ is used to denote the cardinality of the enclosed set $\mathcal{S}$. We let $\mathbf{I}_K$ denote the $K \times K$ identity matrix, whereas $\mathbf{1}_K$ and $\mathbf{0}_K$ are the K-dimensional unit and null column vectors, respectively. We use $\mathcal{CN}(\cdot, \cdot)$ to denote a multivariate circularly-symmetric complex Gaussian distribution whereas $\mathcal{N}(\cdot, \cdot)$ stands for a real one. The notation $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution whereas $\xrightarrow{a.s.}$ stands for almost surely equivalent. We denote $J_n(\cdot)$ the n-order Bessel function and use $Q(z)$ to indicate the Gaussian tail function while $Q^{-1}(z)$ indicates its inverse. We use $\nabla f(\mathbf{z})$ to indicate the vector differential operator whereas $\nabla^2 f(\mathbf{z})$ stands for the Laplace operator. The notation $x \sim y$ with x, y real-valued numbers means that x is approximately equal to y.

assume that after each step the user performs a next step in an independent random direction.[2] Thus, $\mathbf{x}_k(t)$ can be expressed as a sum over $\lfloor t/\xi \rfloor$ independent and identically distributed zero-mean steps $\Delta\mathbf{x}_k(i)$, i.e., $\mathbf{x}_k(t) = \sum_{i=1}^{\lfloor t/\xi \rfloor} \Delta\mathbf{x}_k(i)$ with $|\Delta\mathbf{x}_k(i)| = \ell$. As it is known (see Appendix B for a short discussion on this), if $\ell \to 0$ and $\xi \to 0$ with ℓ^2/ξ kept fixed, then the random walk converges (due to the central limit theorem) to a Brownian motion with diffusion coefficient D such that $4D = \ell^2/\xi$.

We call $\mathbf{h}_k(t) = [h_{k,1}(t), h_{k,2}(t), \ldots, h_{k,N}(t)]^T \in \mathbb{C}^{N \times 1}$ the vector whose generic entry $h_{k,n}(t)$ represents the channel propagation coefficient between the nth antenna at the BS and the kth UE at time t. In particular, we assume that

$$\mathbf{h}_k(t) = \sqrt{l(\mathbf{x}_k(t))}\mathbf{w}_k(t) \tag{1}$$

where $\mathbf{w}_k(t) \sim \mathcal{CN}(0, \mathbf{I}_K)$ accounts for the small-scale fading channel and $l(\cdot) : \mathbb{R}^2 \to \mathbb{R}^+$ describes the large-scale channel fading at different user locations; that is, $l(\mathbf{x}_k(t))$ is the average channel attenuation due to pathloss and shadowing at distance $||\mathbf{x}_k(t)||$. As seen, $l(\mathbf{x}_k(t))$ is assumed to be independent from the transmit antenna index n. This is a reasonable assumption since the distances between UEs and BS are much larger than the distance between the antennas [22]. Since the forthcoming analysis does not depend on a particular choice of $l(\cdot)$ as long as it is a function of the user distance and is bounded from below, we keep it generic. Moreover, in all subsequent discussions we assume that $\{l(\mathbf{x}_k(t)); k = 1, 2, \ldots, K\}$ are known at the BS (see Section VI-D for a brief discussion on this).

Example 1. *In the simulations, we assume that the average channel attenuation at a generic position $\mathbf{x}$ is dominated by the pathloss and is evaluated as [23]*

$$l(\mathbf{x}) = 2L_{\bar{x}} \left(1 + \frac{||\mathbf{x}||^\beta}{\bar{x}^\beta}\right)^{-1} \tag{2}$$

where $\beta \geq 2$ is the pathloss exponent, $\bar{x} > 0$ is some cut-off parameter and $L_{\bar{x}}$ is a constant that regulates the attenuation at distance $\bar{x}$. The average inverse channel attenuation $\mathbb{E}_{\mathbf{x}}\{l^{-1}(\mathbf{x})\}$ plays a key role in all subsequent discussions. Simple integration shows that, if $l(\mathbf{x})$ is given by (2) and the initial positions of UEs are assumed to be uniformly distributed within the cell, then

$$\frac{1}{|\mathcal{C}|} \int_{\mathcal{C}} \frac{1}{l(\mathbf{x})} d\mathbf{x} = \frac{R^\beta}{2\bar{x}^\beta L_{\bar{x}}} \left(\frac{2}{2+\beta} + \frac{\bar{x}^\beta}{R^\beta}\right). \tag{3}$$

We assume that the temporal correlations of $\{\mathbf{w}_k(t); t \geq 0\}$ are such that the coherence time is $\Delta\tau \sim \lambda\xi/\ell$, where λ is the wavelength and ℓ/ξ is the velocity of each UE due to the random walk.

Example 2. *Consider a cell with radius $R = 500$ m and operating at the carrier frequency of $f_c = 2.4$ GHz so that $\lambda = 0.125$ m. Assume $\ell = 50$ m and $\xi = 30$ seconds. Therefore, the average velocity of each UE within two consecutive time steps is $\ell/\xi = 100$ m/minute and the diffusion*

---

[2]The model can be generalized to include correlated steps [7]. In this case, ℓ and ξ will correspond to the length and time over which the correlation in the direction of the walker is lost.

---

coefficient is $D = \ell^2/(4\xi) = 1250$ m²/minute. Under these assumptions, the coherence time $\Delta\tau$ of the small-scale fading is approximately $\Delta\tau \sim \lambda\xi/\ell = 0.075$ s. On the other hand, the coherence time of UE movements (also known as forgetting time) is $\sim R^2/D = 3.33$ hrs, which is much larger than $\Delta\tau$. All this will be used later on to give some intuitions about the different impact of small-scale fading and UE mobility on the energy fluctuations.

B. Problem statement

We call $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T \in \mathbb{C}^{K \times 1}$ the signal transmitted at time t and denote by $\mathbf{V}(t) = [\mathbf{v}_1(t), \mathbf{v}_2(t), \ldots, \mathbf{v}_K(t)] \in \mathbb{C}^{N \times K}$ its precoding matrix. We assume that $\mathbf{s}(t)$ originates from a Gaussian codebook with zero mean and covariance matrix $\mathbb{E}_{\mathbf{s}}[\mathbf{s}(t)\mathbf{s}^H(t)] = \mathbf{I}_K$. The sample $y_k(t)$ received at the kth UE at time t takes the form

$$y_k(t) = \mathbf{h}_k^H(t)\mathbf{V}(t)\mathbf{s}(t) + n_k(t) \tag{4}$$

where $n_k(t) \sim \mathcal{CN}(0, \sigma^2)$ is the additive noise. Under the assumption of perfect CSI at UEs, the SINR at the kth UE is easily written as [19]

$$\text{SINR}_k(t) = \frac{\left|\mathbf{h}_k^H(t)\mathbf{v}_k(t)\right|^2}{\sum_{i=1, i\neq k}^{K} \left|\mathbf{h}_k^H(t)\mathbf{v}_i(t)\right|^2 + \sigma^2} \tag{5}$$

whereas the achievable rate in bit/s/Hz is given by

$$R_k(t) = \log_2\left(1 + \text{SINR}_k(t)\right). \tag{6}$$

While conventional systems have large disparity between peak and average rates, we aim at designing the system so as to guarantee a fixed rate r_k at UE k at each time slot for $k = 1, 2, \ldots, K$. Imposing $R_k(t) = r_k$ into (6) yields

$$\text{SINR}_k(t) = \frac{\left|\mathbf{h}_k^H(t)\mathbf{v}_k(t)\right|^2}{\sum_{i=1, i\neq k}^{K} \left|\mathbf{h}_k^H(t)\mathbf{v}_i(t)\right|^2 + \sigma^2} = \gamma_k \tag{7}$$

where γ_k is the target SINR of UE k obtained as $\gamma_k = 2^{r_k} - 1$. The transmitted power $P(t) = \mathsf{E}_{\mathbf{s}}[||\mathbf{V}(t)\mathbf{s}(t)||^2]$ at time t is given by

$$P(t) = \text{tr}\left(\mathbf{V}(t)\mathbf{V}^H(t)\right) \tag{8}$$

while the energy consumption E_T within a given time interval $[0, T]$ is obtained as

$$E_T = \int_0^T P(t)dt = \int_0^T \text{tr}\left(\mathbf{V}(t)\mathbf{V}^H(t)\right) dt. \tag{9}$$

Since $\mathbf{V}(t)$ depends on the realizations of $\mathbf{W}(t) = [\mathbf{w}_1(t), \mathbf{w}_2(t), \ldots, \mathbf{w}_K(t)] \in \mathbb{C}^{N \times K}$ as well as on the user positions $\{\mathbf{x}_k(t); k = 1, 2, \ldots, K\}$, the energy E_T in (9) is clearly a random and time-dependent quantity. To characterize its statistics, we exploit the large system limit in which $K, N \to \infty$ with fixed ratio $c = K/N$. This will allow us to measure the energy consumption of the system when K is relatively large but also to capture the temporal correlations induced on the energy by UE mobility and its corresponding

asymptotic distribution. The asymptotic analysis will be conducted for most of the common linear precoding techniques. We begin with the precoder that minimizes the power $P(t)$ while satisfying the rate requirements. Then, we concentrate on the asymptotically design of MRT, ZF and RZF.

Remark 1. *A known problem with using the asymptotic analysis is that the target rates are not guaranteed to be achieved when N is finite and relatively small (see for example [14]). This is because the approximation errors are translated into fluctuations in the resulting SINR values. However, these errors vanish rapidly also in the finite regime when N is large enough, which is the regime envisioned for massive MIMO systems [24].*

III. OPTIMAL LINEAR PRECODING VIA LARGE SYSTEM ANALYSIS

We assume that at each time slot t the precoding matrix $\mathbf{V}(t)$ is designed as the solution of the following power minimization problem:

$$\min_{\mathbf{V}(t)} \quad P(t) = \text{tr}\left(\mathbf{V}(t)\mathbf{V}^H(t)\right) \tag{10}$$

$$\text{subject to} \quad \text{SINR}_k(t) \geq \gamma_k \quad k = 1, 2, \ldots, K. \tag{11}$$

As shown in [10], the above optimization problem is not convex but it can be put in a convex form by reformulating the SINR constraints as second-order cone constraints. The optimal linear precoder (OLP) $\mathbf{V}_{\text{OLP}}(t)$ is eventually found to be [10], [11]:

$$\mathbf{V}_{\text{OLP}}(t) = \left(\sum_{i=1}^{K} \lambda_i^\star(t)\mathbf{h}_i(t)\mathbf{h}_i^H(t) + N\mathbf{I}_N\right)^{-1} \mathbf{H}(t)\sqrt{\mathbf{P}^\star(t)} \tag{12}$$

where $\mathbf{H}(t) = [\mathbf{h}_1(t), \mathbf{h}_2(t), \ldots, \mathbf{h}_K(t)] \in \mathbb{C}^{N \times K}$ and $\boldsymbol{\lambda}^\star(t) = [\lambda_1^\star(t), \lambda_2^\star(t), \ldots, \lambda_K^\star(t)]^T$ is the positive unique fixed point of the following equations for $k = 1, 2, \ldots, K$ [10], [25]:

$$\left(1 + \frac{1}{\gamma_k}\right)\lambda_k^\star(t) =$$

$$\frac{1}{\mathbf{h}_k^H(t)\left(\sum_{i=1}^{K} \lambda_i^\star(t)\mathbf{h}_i(t)\mathbf{h}_i^H(t) + N\mathbf{I}_N\right)^{-1}\mathbf{h}_k(t)}. \tag{13}$$

Also, $\mathbf{P}^\star(t) = \text{diag}\{p_1^\star(t), p_2^\star(t), \ldots, p_K^\star(t)\}$ is a diagonal matrix whose entries are such that the SINR constraints are all satisfied with equality when $\mathbf{V}(t) = \mathbf{V}_{\text{OLP}}(t)$. Plugging (5) into (11), the optimal $\mathbf{p}^\star(t) = [p_1^\star(t), p_2^\star(t), \ldots, p_K^\star(t)]^T$ is obtained as

$$\mathbf{p}^\star(t) = \sigma^2 \mathbf{D}^{-1}(t)\mathbf{1}_K \tag{14}$$

where the (k, i)th element of $\mathbf{D}(t) \in \mathbb{C}^{K \times K}$ is [25]

$$[\mathbf{D}(t)]_{k,i} = \begin{cases} \frac{1}{\gamma_k}|\mathbf{h}_k^H(t)\mathbf{c}_k^\star(t)|^2 & \text{for } k = i \\ -|\mathbf{h}_k^H(t)\mathbf{c}_i^\star(t)|^2 & \text{for } k \neq i \end{cases} \tag{15}$$

with $\mathbf{c}_k^\star(t)$ being the kth column of

$$\mathbf{C}^\star(t) = \left(\sum_{i=1}^{K} \lambda_i^\star(t)\mathbf{h}_i(t)\mathbf{h}_i^H(t) + N\mathbf{I}_N\right)^{-1} \mathbf{H}(t). \tag{16}$$

As seen, $\mathbf{V}_{\text{OLP}}(t)$ in (12) is parameterized by $\boldsymbol{\lambda}^\star(t)$ and $\mathbf{p}^\star(t)$, where $\boldsymbol{\lambda}^\star(t)$ needs to be evaluated by an iterative procedure due to the fixed-point equations in (13). When $K, N \to \infty$ with $K/N = c \in (0, 1]$, some recent tools in large system analysis allow us to compute the so-called deterministic equivalents (see [26], [27] for more details on this subject) of $\boldsymbol{\lambda}^\star(t)$ and $\mathbf{p}^\star(t)$. For later convenience, we call

$$\eta = 1 - \frac{c}{K}\sum_{i=1}^{K} \frac{\gamma_i}{1 + \gamma_i} \tag{17}$$

and

$$A(t) = \frac{1}{K}\sum_{i=1}^{K} \frac{\gamma_i}{l(\mathbf{x}_i(t))}. \tag{18}$$

The following theorem provides the asymptotic value of the solution to (10).

Theorem 1. *If $K, N \to \infty$ with $K/N = c \in (0, 1]$, then*

$$\max_{k=1,2,\ldots,K} \left|\lambda_k^\star(t) - \overline{\lambda}_k(t)\right| \xrightarrow{a.s.} 0 \tag{19}$$

$$\max_{k=1,2,\ldots,K} |p_k^\star(t) - \overline{p}_k(t)| \xrightarrow{a.s.} 0 \tag{20}$$

where $\overline{\lambda}_k(t)$ and $\overline{p}_k(t)$ are the deterministic equivalents of $\lambda_k^\star(t)$ and $p_k^\star(t)$ at time slot t, respectively, and are given by

$$\overline{\lambda}_k(t) = \frac{\gamma_k}{l(\mathbf{x}_k(t))\eta} \tag{21}$$

$$\overline{p}_k(t) = \frac{\gamma_k}{l(\mathbf{x}_k(t))\eta^2}\left(\overline{P}_{\text{OLP}}(t) + \frac{\sigma^2}{l(\mathbf{x}_k(t))}\left(1 + \gamma_k\right)^2\right) \tag{22}$$

with

$$\overline{P}_{\text{OLP}}(t) = \frac{c\sigma^2}{\eta}A(t) \tag{23}$$

being the deterministic equivalent of the transmit power $P(t)$ in (8).

Proof: Similar results have previously been derived by applying standard random matrix theory tools to the right-hand-side of (13). However, the application of these tools to the problem at hand is not analytically correct since the Lagrange multipliers in (13) are a function of the channel vectors $\{\mathbf{h}_i(t)\}$. To overcome this issue, we make use of the same approach adopted in [15] whose main steps are sketched in Appendix A for completeness. On the other hand, (22) is proved using standard random matrix theory results and is omitted for space limitations. ∎

Observe that the Lagrange multiplier $\lambda_k(t)$ is known to act as a user priority parameter that implicitly determines how much interference the other UEs may cause to UE k [25]. Interestingly, its asymptotic value $\overline{\lambda}_k(t)$ in (21) turns out to be proportional to the target SINR γ_k and inversely proportional to $l(\mathbf{x}_k(t))$ such that users with weak channels have larger values. This means that in the asymptotic regime higher priority is given to users that require high performance

(i.e., $\gamma_k > 1$) and/or have weak average propagation conditions (i.e., $l(\mathbf{x}_k(t)) < 1$).

The following corollary can be easily obtained from Theorem 1 and will be useful later on.

Corollary 1 ([13])**.** *If the same target SINR is imposed for each user, i.e.,*

$$\boldsymbol{\gamma} = \gamma \mathbf{1}_k \tag{24}$$

then $\overline{\lambda}_k(t)$ *in (21) reduces to*

$$\overline{\lambda}_k(t) = \frac{\gamma}{l(\mathbf{x}_k(t))} \left(1 - c\frac{\gamma}{1+\gamma}\right)^{-1} \tag{25}$$

and $\overline{P}_{\mathrm{OLP}}(t)$ *becomes*

$$\overline{P}_{\mathrm{OLP}}(t) = c\sigma^2 \left(1 - c\frac{\gamma}{1+\gamma}\right)^{-1} A(t). \tag{26}$$

IV. HEURISTIC LINEAR PRECODING VIA LARGE SYSTEM ANALYSIS

Inspired by the optimal linear precoding in (12), we also consider suboptimal precoding techniques that build on heuristics [25]. To this end, we let $\mathbf{V}(t)$ take the generic form:

$$\mathbf{V}(t) = \left(\sum_{i=1}^{K} \alpha_i(t)\mathbf{h}_i(t)\mathbf{h}_i^H(t) + N\rho\mathbf{I}_N\right)^{-1} \mathbf{H}(t)\sqrt{\mathbf{P}(t)} \tag{27}$$

where $\boldsymbol{\alpha}(t) = [\alpha_1(t), \alpha_2(t), \ldots, \alpha_K(t)]^T$ is now a given vector with positive design parameters and $\rho > 0$ is another design parameter. Note that (27) is basically obtained from (12) by setting $\lambda_k(t) = \alpha_k(t)/\rho$ for all k. As before, the power allocation matrix $\mathbf{P}(t)$ is computed so as to satisfy all the SINR constraints with equality in the asymptotic regime.

If $\mathbf{V}(t)$ takes the generic heuristic form in (27), then for any given $\boldsymbol{\alpha}(t)$ and ρ the results of Theorem 1 in [16] lead to the following corollary.

Corollary 2 ([16])**.** *If* $\mathbf{V}(t)$ *is defined as in (27) and* $K, N \to \infty$ *with* $c \in (0, 1]$*, then*

$$P(t) - \overline{P}(t) \xrightarrow{a.s.} 0 \tag{28}$$

$$\mathrm{SINR}_k(t) - \overline{\mathrm{SINR}}_k(t) \xrightarrow{a.s.} 0 \tag{29}$$

where $\overline{P}(t)$ *and* $\overline{\mathrm{SINR}}_k(t)$ *are given by*

$$\overline{P}(t) = \frac{c\mu'}{K} \sum_{i=1}^{K} \frac{p_i(t)l(\mathbf{x}_i(t))}{(1 + \alpha_i(t)l(\mathbf{x}_i(t))\mu)^2} \tag{30}$$

$$\overline{\mathrm{SINR}}_k(t) = \frac{p_k(t)l(\mathbf{x}_k(t))\mu^2}{\overline{P}(t) + \frac{\sigma^2}{l(\mathbf{x}_k(t))} (1 + \alpha_k(t)l(\mathbf{x}_k(t))\mu)^2} \tag{31}$$

and μ *is the solution to the fixed point equation*

$$\mu = \left(\frac{1}{N} \sum_{i=1}^{K} \frac{\alpha_i(t)l(\mathbf{x}_i(t))}{1 + \alpha_i(t)l(\mathbf{x}_i(t))\mu} + \rho\right)^{-1} \tag{32}$$

with μ' *in (30) being its derivative with respect to* ρ*.*

The above corollary is now easily used to compute the asymptotic power $\overline{P}(t)$ required to satisfy $\overline{\mathrm{SINR}}_k(t) = \gamma_k$

for $k = 1, 2, \ldots, K$. Setting $\overline{\mathrm{SINR}}_k(t)$ in (31) equal to γ_k yields

$$\overline{p}_k(t) = \frac{\gamma_k}{l(\mathbf{x}_k(t))\mu^2} \left(\overline{P}(t) + \frac{\sigma^2}{l(\mathbf{x}_k(t))} (1 + \alpha_k(t)l(\mathbf{x}_k(t))\mu)^2\right). \tag{33}$$

Plugging (33) into (30) and solving with respect to $\overline{P}(t)$ leads to the following result.

Lemma 1. *If* $\overline{\mathrm{SINR}}_k(t) = \gamma_k$ *for* $k = 1, 2, \ldots, K$*, then* $\overline{p}_k(t)$ *takes the form in (33) and*

$$\overline{P}(t) = \frac{c\sigma^2 A(t)}{1 - \mu^2 F(t) - cB(t)} \tag{34}$$

with $A(t)$ *and* μ *being given by (18) and (32), respectively, and*

$$B(t) = \frac{1}{K} \sum_{i=1}^{K} \frac{\gamma_i}{(1 + \alpha_i(t)l(\mathbf{x}_i(t))\mu)^2} \tag{35}$$

$$F(t) = \frac{1}{N} \sum_{i=1}^{K} \frac{(\alpha_i(t)l(\mathbf{x}_i(t)))^2}{(1 + \alpha_i(t)l(\mathbf{x}_i(t))\mu)^2}. \tag{36}$$

The above results are next used to characterize the asymptotic structure of MRT, ZF and RZF.

A. Maximum Ratio Transmission

Setting $\boldsymbol{\alpha}(t) = \mathbf{0}_K$ and $\rho = 1$ into (27) leads to

$$\mathbf{V}_{\mathrm{MRT}}(t) = \frac{1}{N}\mathbf{H}(t)\sqrt{\mathbf{P}(t)} \tag{37}$$

which is the well known MRT precoder. The following corollary is obtained from Lemma 1.

Corollary 3. *If MRT is used, then* $\overline{p}_k(t)$ *reduces to* $\overline{p}_k(t) = \frac{\gamma_k}{l(\mathbf{x}_k(t))} \left(\overline{P}_{\mathrm{MRT}}(t) + \frac{\sigma^2}{l(\mathbf{x}_k(t))}\right)$ *where*

$$\overline{P}_{\mathrm{MRT}}(t) = \frac{c\sigma^2}{1 - c\underline{\gamma}} A(t) \tag{38}$$

and $\underline{\gamma}$ *is the average of target SINRs:*

$$\underline{\gamma} = \frac{1}{K} \sum_{i=1}^{K} \gamma_i. \tag{39}$$

Since $\overline{P}_{\mathrm{MRT}}(t)$ must be positive and finite, it follows that the asymptotic analysis can be applied to MRT only when the following condition is satisfied: $1 - c\underline{\gamma} > 0$. If $\boldsymbol{\gamma} = \gamma\mathbf{1}_k$, this implies $\gamma < 1/c$ or, equivalently, $r < \log_2\left(1 + \frac{1}{c}\right)$. If the above condition is met, then $\overline{P}_{\mathrm{MRT}}(t)$ is positive and finite. On the other hand, $\overline{P}_{\mathrm{MRT}}(t)$ diverges to infinity when $r = \log_2\left(1 + \frac{1}{c}\right)$.

B. Zero Forcing

Setting $\boldsymbol{\alpha}(t) = \mathbf{1}_K$ into (12) yields $\mathbf{V}(t) = \left(\mathbf{H}(t)\mathbf{H}^H(t) + N\rho\mathbf{I}_N\right)^{-1} \mathbf{H}(t)\sqrt{\mathbf{P}(t)}$ from which using the Woodbury matrix identity and imposing $\rho = 0$ the ZF precoder is obtained

$$\mathbf{V}_{\mathrm{ZF}}(t) = \mathbf{H}(t) \left(\mathbf{H}^H(t)\mathbf{H}(t)\right)^{-1} \sqrt{\mathbf{P}(t)}. \tag{40}$$

The following corollary can be obtained from Lemma 1.

Corollary 4. *If ZF is used, then $\overline{p}_k(t)$ reduces to $\overline{p}_k(t) = \gamma_k \sigma^2$ and*

$$\overline{P}_{\mathrm{ZF}}(t) = \frac{c\sigma^2}{1-c} A(t). \tag{41}$$

Proof. Setting $\boldsymbol{\alpha} = \mathbf{1}_K$ into (32) and using simple mathematical arguments yields

$$\rho\mu = 1 - c + \frac{1}{N} \sum_{i=1}^{K} \frac{1}{1 + l(\mathbf{x}_i(t))\mu}. \tag{42}$$

Letting $\varphi = \rho\mu$ one gets $\varphi = 1 - c + \frac{1}{N} \sum_{i=1}^{K} \frac{1}{1+l(\mathbf{x}_i(t))\varphi/\rho}$ from which it follows that $\varphi \to 1-c$ when $\rho \to 0$. This means that μ goes to infinity as $\frac{1-c}{\rho}$ when $\rho \to 0$. Therefore, from (33) one easily gets $\overline{p}_k(t) - \gamma_k \sigma^2 \to 0$ whereas (36) leads to $B(t) \to 0$ and $\mu^2 F(t) \to c$. Plugging these results into (34) yields (41). $\qquad\square$

C. Regularized Zero Forcing

Assume that $\alpha_i(t) = 1/l(\mathbf{x}_i(t))$ for any i, then the processing matrix $\mathbf{V}(t)$ in (12) reduces to (see also [17])

$$\mathbf{V}_{\mathrm{RZF}}(t) = \left(\sum_{i=1}^{K} \mathbf{w}_i(t)\mathbf{w}_i^H(t) + N\rho\mathbf{I}_N \right)^{-1} \mathbf{H}(t)\sqrt{\mathbf{P}(t)} \tag{43}$$

which is referred to as RZF precoder in the sequel. Differently from OLP that requires to compute the fixed point of a set of equations, the optimization of RZF requires only to look for the value of ρ. This can not generally be done in closed-form but requires a numerical optimization procedure [25]. If the asymptotic regime is analyzed, then the following result is obtained.

Lemma 2. *If a RZF precoder is used and $K, N \to \infty$ with $K/N = c \in (0,1]$, then the optimal regularization parameter is found to be*

$$\rho^\star = \frac{1}{\underline{\gamma}} - \frac{c}{1 + \underline{\gamma}} \tag{44}$$

with $\underline{\gamma}$ given by (39). The deterministic equivalent of the transmit power reduces to

$$\overline{P}_{\mathrm{RZF}}(t) = c\sigma^2 \left(1 - c\frac{\underline{\gamma}}{1+\underline{\gamma}} \right)^{-1} A(t) \tag{45}$$

whereas $\overline{p}_k(t) = \frac{\gamma_k}{l(\mathbf{x}_k(t))\underline{\gamma}^2} \left(\overline{P}_{\mathrm{RZF}}(t) + \frac{\sigma^2}{l(\mathbf{x}_k(t))} \left(1 + \underline{\gamma} \right)^2 \right).$

Proof: Setting $\alpha_i = 1/l(\mathbf{x}_i(t))$ for $i = 1, 2, \ldots, K$ into (32) and (34) – (36) yields

$$\overline{P}_{\mathrm{RZF}}(t) = c\sigma^2 \frac{(1+\mu)^2 A(t)}{\mu \left(c + \rho (1+\mu)^2 \right) - c\underline{\gamma}} \tag{46}$$

with

$$\mu = \left(\frac{c}{1+\mu} + \rho \right)^{-1}. \tag{47}$$

Taking the derivative of $\overline{P}_{\mathrm{RZF}}(t)$ in (46) with respect to ρ, one gets (the mathematical steps are omitted for space limitations)

$$\frac{\partial \overline{P}_{\mathrm{RZF}}(t)}{\partial \rho} = 2c^2\sigma^2 \frac{(\underline{\gamma} - \mu) A(t)}{\left(\mu \left(c + \rho (1+\mu)^2 \right) - c\underline{\gamma} \right)^2}. \tag{48}$$

From the above equation, it turns out that the minimum power is achieved when μ is such that $\mu = \underline{\gamma}$. Plugging this result into (47) yields (44). $\qquad\blacksquare$

To the best of our knowledge, this is the first time that the value of ρ that minimizes the power consumption in the asymptotic regime is given in the explicit form of (44). Most of works have only looked for the value of ρ that maximizes the sum rate of the network (see for example [16]). Interestingly, the results of Lemma 2 can be used to prove the following corollary, which states that RZF is asymptotically equivalent to the optimal linear precoder when the same rate constraint is imposed for all UEs.

Corollary 5. *If the same target SINR is imposed for each user, i.e., $\boldsymbol{\gamma} = \gamma\mathbf{1}_K$, then RZF becomes optimal in the asymptotic regime.*

Proof: If $\boldsymbol{\gamma} = \gamma\mathbf{1}_k$, then (44) reduces to $\rho^\star = \frac{1}{\gamma} - \frac{c}{1+\gamma}$ and $\mathbf{V}_{\mathrm{RZF}}(t)$ in (43) becomes equivalent to (12) after replacing $\lambda_k^\star$ with $\overline{\lambda}_k$ given by (25). $\qquad\blacksquare$

Observe that if $\boldsymbol{\alpha}(t)$ is set to $\mathbf{1}_K$, then $\mathbf{V}(t)$ in (12) reduces to [16]:

$$\mathbf{V}_{\mathrm{RZF}}(t) = \left(\sum_{i=1}^{K} \mathbf{h}_i(t)\mathbf{h}_i^H(t) + N\rho\mathbf{I}_N \right)^{-1} \mathbf{H}\sqrt{\overline{\mathbf{P}}}(t). \tag{49}$$

As shown in [18], if $K, N \to \infty$ with $c \in (0,1]$ and $\mathbf{V}_{\mathrm{RZF}}(t)$ is given as above, then the value of ρ minimizing $P(t) = \mathrm{tr}\left(\mathbf{V}(t)\mathbf{V}^H(t) \right)$ while satisfying the rate constraints is obtained as

$$\rho^\star = \frac{1}{\mu^\star} - \frac{1}{N} \sum_{i=1}^{K} \frac{l(\mathbf{x}_i(t))}{1 + l(\mathbf{x}_i(t))\mu^\star} \tag{50}$$

with $\mu^\star$ being solution of the following fixed point equation:

$$\mu^\star = \left(\sum_{i=1}^{K} \frac{l(\mathbf{x}_i(t))\gamma_i}{(1 + l(\mathbf{x}_i(t))\mu^\star)^3} \right) \left(\sum_{i=1}^{K} \frac{(l(\mathbf{x}_i(t)))^2}{(1 + l(\mathbf{x}_i(t))\mu^\star)^3} \right)^{-1}. \tag{51}$$

In addition, it turns out that [18]

$$\overline{P}_{\mathrm{RZF}}(t) = \frac{c\sigma^2 A(t)}{1 - (\mu^\star)^2 F(t) - cB(t)} \tag{52}$$

where $A(t)$ is given by (18) while $B(t)$ and $F(t)$ are obtained from (36) setting $\alpha_i(t) = 1 \ \forall i$ and replacing μ with $\mu^\star$. The above results are used in [18] to prove that the conventional RZF processing given by (49) becomes optimal in the asymptotic regime only if the ratio between γ_k and $l(\mathbf{x}_k(t))$ is the same for any k. Although possible, this assumption is not very realistic in practical systems as it would imply a strong dependence between rate requirements and UE positions. For this reason, we focus our analysis on the RZF given by (43) whose optimality holds true under a more reasonable assumption of identical target UE rates.

V. ENERGY FLUCTUATIONS

We are now left with studying the fluctuations of E_T in the large system limit. Towards this end, observe that the large system analysis above shows that, for all considered schemes, the transmit power $P(t) = \mathrm{tr}\left(\mathbf{V}(t)\mathbf{V}^H(t)\right)$ hardens to a deterministic quantity, which takes the following general form:

$$\overline{P}(t) = \frac{c\sigma^2}{\eta} A(t) = \frac{c\sigma^2}{\eta} \frac{1}{K} \sum_{i=1}^{K} \frac{\gamma_i}{l(\mathbf{x}_i(t))} \qquad (53)$$

with η being specifically given by

$$\eta = \begin{cases} 1 - \frac{c}{K} \sum_{i=1}^{K} \frac{\gamma_i}{1+\gamma_i} & \text{for OLP} \\[2mm] 1 - c\underline{\gamma} = 1 - \frac{c}{K} \sum_{i=1}^{K} \gamma_i & \text{for MRT} \\[2mm] 1 - c = 1 - \frac{K}{N} & \text{for ZF} \\[2mm] 1 - c\frac{\underline{\gamma}}{1+\underline{\gamma}} = 1 - c\frac{\sum_{i=1}^{K} \gamma_i}{K + \sum_{i=1}^{K} \gamma_i} & \text{for RZF.} \end{cases} \qquad (54)$$

As seen, in the large system limit the power expenditure of all considered schemes is a function of time due to the fluctuations induced on the average channel attenuation by $\{\mathbf{x}_i(t); i = 1, 2, \ldots, K\}$. The statistics of the large-scale UE movements can be exploited to calculate the temporal fluctuations of each term in the right-hand-side of (53). These results are stated in the following two lemmas, which are proved in Appendix C.

Lemma 3. *The average value of $l^{-1}(\mathbf{x}_k(t))$ does not depend on the UE movements and hardens to a deterministic quantity independent of time t, which is the same for all UEs and given by*

$$\mathsf{E}_{\mathbf{X}_k}\left[l^{-1}(\mathbf{x}_k(t))\right] = \frac{1}{|\mathcal{C}|} \int_{\mathcal{C}} \frac{1}{l(\mathbf{x})} d\mathbf{x} \qquad (55)$$

with $|\mathcal{C}| = \pi R^2$.

Lemma 4. *The covariance of $l^{-1}(\mathbf{x}_k(t))$ depends on the temporal correlation induced by UE movements and is such that*

$$\iint_0^T \mathsf{COV}_{\mathbf{X}_k}\left[l^{-1}(\mathbf{x}_k(\tau)), l^{-1}(\mathbf{x}_k(s))\right] d\tau ds = \frac{TR^2}{D}\Theta \qquad (56)$$

where

$$\Theta = \sum_{i=1}^{\infty} \frac{2\phi_i^2}{\kappa_i^2 J_0^2(\kappa_i)} \int_0^1 \left(1 - e^{-\frac{\kappa_i^2 DTt}{R^2}}\right)^2 dt \qquad (57)$$

and

$$\phi_i = 2 \int_0^1 \frac{1}{l(R\mathbf{z})} J_0(\kappa_i z) z dz \qquad (58)$$

with κ_i being the ith zero of $J_1(x)$.

The values of $\{\kappa_i\}$ in (57) and (58) can be found in [28, pp. 390] whereas $\{\phi_i\}$ can be calculated explicitly using formulae in [29, pp. 684]. Observe also that the sum in (57) requires only to compute a few terms as it converges fast to its effective value.

The following theorem summarizes one of the major results of this work.

Theorem 2. *If linear precoding is used at the BS, then the following convergence holds true in the large system limit:*

$$\sqrt{K}\left(\frac{E_T - \epsilon}{\sqrt{\Sigma}}\right) \xrightarrow[K,N\to\infty]{\mathcal{D}} \mathcal{N}(0,1) \qquad (59)$$

where

$$\epsilon = T\frac{c\sigma^2}{\eta}\left(\frac{1}{K}\sum_{i=1}^{K}\gamma_i\right)\frac{1}{|\mathcal{C}|}\int_{\mathcal{C}}\frac{1}{l(\mathbf{x})}d\mathbf{x} \qquad (60)$$

$$\Sigma = \left(\frac{c\sigma^2}{\eta}\right)^2\left(\frac{1}{K}\sum_{i=1}^{K}\gamma_i^2\right)\frac{TR^2}{D}\Theta \qquad (61)$$

with Θ being computed as in (57).

Proof: The proof is given in Appendix D. ∎

Theorem 2 says that the fluctuation of E_T around ϵ is approximately Gaussian with variance Σ/K. Some intuition on this result is as follows. As mentioned earlier, the temporal variability of the energy E_T basically depends through $\mathbf{V}(t)$ on the variations of the small-scale fading matrix $\{\mathbf{W}(t)\}$ and the UE positions $\{\mathbf{x}_i(t); i = 1, 2, \ldots, K\}$ in the cell. However, the statistics of these two contributions and their impact on the fluctuations of E_T are relatively different. To see how this comes about, recall that the coherence time $\Delta\tau$ of $\mathbf{W}(t)$ is $\sim \lambda\xi/\ell$ where λ is the wavelength and ℓ/ξ is the UE velocity between two successive steps. In addition, the variance of the power required to only compensate $\mathbf{W}(t)$ (which corresponds to the case of motionless UEs) is known to scale as $1/K^2$ due to the K^2 degrees of freedom of $\mathbf{W}(t)$ [20]. This means that the energy fluctuations induced by the compensation of the small-scale fading only behaves roughly as $\sim \Delta\tau/K^2$. Differently from the small-scale fading (where the coherence distance is $\sim \lambda\xi$), the coherence distance of UE movements (defined as the distance after which the correlation between the initial and final position becomes small) is related to the cell radius R and the corresponding coherence time can be roughly quantified as $\sim R^2/D$ with $4D = \ell^2/\xi$. The latter is much larger than $\Delta\tau \sim \lambda\xi/\ell$ for typical parameter settings (see Example 2). In addition, the variance of the power due to the randomness of $\{\mathbf{x}_i(t)\}$ scales as $1/K$ rather than as $1/K^2$ (see Appendix D). Putting these two facts together, it follows that the energy fluctuations induced by UE mobility are on the order of $\sim R^2/(KD)$ and thus largely dominate the variability of the small-scale fading, which is on the order of $\sim \Delta\tau/K^2$.

Example 3. *Assume that the pathloss function $l(\mathbf{x})$ is modelled as in (2). In these circumstances,*

$$\epsilon = T\frac{c\sigma^2}{\eta}\left(\frac{1}{K}\sum_{i=1}^{K}\gamma_i\right)\frac{R^\beta}{2\bar{x}^\beta L_{\bar{x}}}\left(\frac{2}{2+\beta} + \frac{\bar{x}^\beta}{R^\beta}\right) \qquad (62)$$

and the coefficients $\{\phi_i\}$ in (61) take the form:[3]

$$\phi_i = \frac{R^\beta}{\bar{x}^\beta L_{\bar{x}}}\int_0^1 x^\beta J_0(\kappa_i x) x dx. \qquad (63)$$

---

[3]Observe that $\int_0^1 J_0(\kappa_i x) x dx = 0$.

TABLE I
PARAMETER SETTING

Parameter	Value	Parameter	Value
Bandwidth	$W = 20$ MHz	Diffusion coefficient	$D = 1250$ m^2/minute
Cell radius	$R = 500$ m	Pathloss coefficient	$\beta = 4$
Cut-off parameter	$\bar{x} = 25$ m	Average pathloss attenuation at $\bar{x}$	$L_{\bar{x}} = -93$ dB
Carrier frequency	$f_c = 2.4$ GHz	Noise power	$\sigma^2 = -97.8$ dBm
Average step size length	$\ell = 50$ m	Time interval	$T = 3, 6, 12$ or 24 hrs

TABLE II
NUMERICAL VALUES OF $\mathsf{E}[E_T]/\epsilon$ AND $K\mathsf{VAR}[E_T]/\Sigma$ FOR THE PARAMETER SETTING IN FIGS. $1-3$. WHEN $T = 3, 6, 12$ AND 24 [HRS].

	Fig. 1	Fig. 1	Fig. 2	Fig. 2	Fig. 3	Fig. 3
	$\mathsf{E}[E_T]/\epsilon$	$K\mathsf{VAR}[E_T]/\Sigma$	$\mathsf{E}[E_T]/\epsilon$	$K\mathsf{VAR}[E_T]/\Sigma$	$\mathsf{E}[E_T]/\epsilon$	$K\mathsf{VAR}[E_T]/\Sigma$
$T = 3$ [hrs]	1.030	1.1887	1.023	1.1207	1.019	1.0675
$T = 6$ [hrs]	1.011	1.1287	1.008	1.0901	1.006	1.048
$T = 12$ [hrs]	1.005	1.0659	1.004	1.0428	1.003	1.021
$T = 24$ [hrs]	1.001	1.0459	1.000	1.0211	1.000	1.018

If $\beta = 4$, using the formulae in [29, pp. 684] one gets

$$\int_0^1 x^\beta J_0(\kappa_i x) x dx = 4 J_0(\kappa_i) \frac{\kappa_i^2 - 8}{\kappa_i^4}. \tag{64}$$

Plugging the above results into (57) yields $\Theta = \Omega R^{2\beta}$ with

$$\Omega = \frac{1}{\bar{x}^{2\beta} L_{\bar{x}}^2} \sum_{i=1}^{\infty} \frac{32(\kappa_i^2 - 8)^2}{\kappa_i^{10}} \int_0^1 \left(1 - e^{-\frac{\kappa_i^2 D T t}{R^2}}\right)^2 dt. \tag{65}$$

The variance Σ in (61) is eventually obtained as

$$\Sigma = \left(\frac{c\sigma^2}{\eta}\right)^2 \left(\frac{1}{K} \sum_{i=1}^{K} \gamma_i^2\right) \Omega \frac{TR^2}{D} R^{2\beta}. \tag{66}$$

From the closed-form expressions in (62) and (66), it follows that the mean and variance of E_T depend heavily on the values of the path-loss exponent β and the cell radius R.

Remark 2. *From the results of Theorem 2, it follows that the mean and variance of E_T are both proportional to the time interval T. This means that the variability of E_T/T will be less important as T becomes large. Observe also that if T is such that $DT/R^2 \gg 1$ then the integral in (57) converges to unity and the ratio Σ/T becomes independent from T. This large time limiting behavior is complementary to the large K and N analysis performed in this work. We will elaborate further on this result later on in Section VII as it can be exploited to consider an alternative regime in which K and N are finite while T grows large.*

Remark 3. *It is worth observing that the results of Theorem 2 can in principle be extended to other linear pre-coding techniques in which $\overline{P}(t)$ has a more involved structure. In particular, it can be extended to the classical RZF precoder whose asymptotic power $\overline{P}_{RZF}(t)$ is given in (52). Although being in a more complicated form than (53), the mean and variance of E_T are still computable as they basically require to evaluate the fluctuations induced by UE movements on the different terms $\mu^\star$, $A(t)$, $B(t)$ and $F(t)$.*

VI. NUMERICAL VALIDATION AND APPLICATIONS

A. Numerical validation

The accuracy of the above asymptotic statistical characterization is now validated numerically by Monte-Carlo simulations. The results are obtained for 1000 different initial positions $\mathbf{x}_k(0)$ for $k = 1, 2, \ldots, K$ within the coverage area. The simulations were performed using Matlab and the code is available for download[4], which enables reproducibility as well as simple testing of other parameter values. The parameter setting is given in Table I. The pathloss function $l(\mathbf{x})$ is modelled as in (2) with $\beta = 4$ and $L_{\bar{x}} = -93$ dB. The latter is such that for $f_c = 2.4$ GHz the attenuation at $\bar{x}$ is the same as that in the cellular model analyzed in [22].

Figs. $1-3$ show the cumulative distribution function (CDF) of E_T/T for different values of K, N and T. The ratio $c = K/N$ is kept fixed and equal to $c = 0.5$ whereas the user rates r_k are all set to 1.5 bit/s/Hz. As seen, the simulation results match pretty well with the theoretical ones for both OLP/RZF and ZF in all investigated scenarios.[5] This validates the theoretical analysis of this work and shows a substantial energy reduction in using OLP. The CDF of E_T/T for MRT has the same behaviour of CDFs of OLP and RZF and it is not reported for illustration purposes as the average power required by MRT is larger than ZF and OLP (or RZF) and thus it would have comprised the legibility of the results. In particular, it turns out that if MRT is used then the mean value of E_T/T is increased by a factor of 7.92 with respect to OLP. This is in agreement to what can be easily obtained computing the ratio between the two different values of (54) for OLP and MRT. A close inspection of Figs. $1-3$ reveals that for a given T there is a progressive agreement between numerical and theoretical results as K increases (see also the numerical values reported in Table II). This is due to the fact that the finite size effect of T becomes less relevant as K grows large. A similar behaviour is observed if one compares

[4]https://github.com/lucasanguinetti/energy-consumption-in-MU-MIMO-with-mobility.

[5]Recall that RZF becomes equivalent to OLP when the user rates are all the same (see Corollary 5).
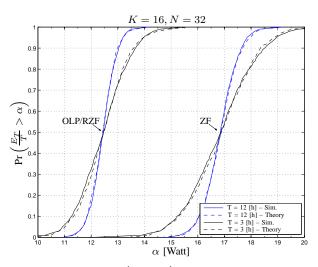
Fig. 1. Outage probability $\Pr\left(\frac{E_T}{T} > \alpha\right)$ of OLP and ZF when $K = 16$ and $T = 3$ or 12 hrs.
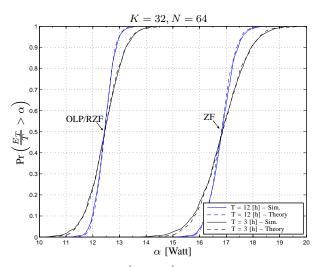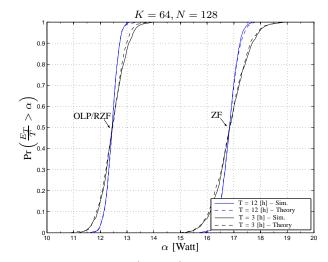

Fig. 3. Outage probability $\Pr\left(\frac{E_T}{T} > \alpha\right)$ of OLP and ZF when $K = 64$ and $T = 3$ or 12 hrs.

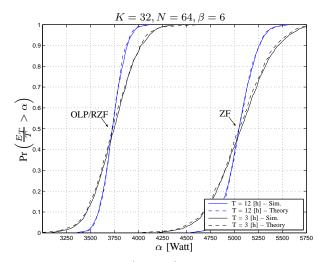Fig. 2. Outage probability $\Pr\left(\frac{E_T}{T} > \alpha\right)$ of OLP and ZF when $K = 32$ and $T = 3$ or 12 hrs.

Fig. 4. Outage probability $\Pr\left(\frac{E_T}{T} > \alpha\right)$ of OLP and ZF when $K = 32$, $\beta = 6$ and $T = 3$ or 12 hrs.

the results of Figs. 1–3 fixing K while letting T increase. This is because increasing T allows each UE to cover larger areas of the cell and, as a consequence, the energy consumption becomes less random and closer to the mean. This in turn reduces the corresponding fluctuations.

To evaluate the impact of the pathloss coefficient on the energy consumption statistics, we set $\beta = 6$ (as an extreme case). In these circumstances, the integral on the right-hand-side of (63) can be computed as [29]:

$$\int_0^1 x^\beta J_0(\kappa_i x) x \, dx = 6 J_0(\kappa_i) \frac{\kappa_i^4 - 24\kappa_i^2 + 192}{\kappa_i^6}. \qquad (67)$$

Fig. 4 show the CDF of E_T/T for $\beta = 6$ when $K = 32, N = 64$ and $r_k = 1,5$ bit/s/Hz. Comparing the results of Fig. 4 with those in Fig. 2, it is seen that both the mean and variance values of E_T/T largely increase as β becomes larger.

B. Dimensioning of cell battery

A possible application of the results of Theorem 2 is as follows. Assume that the energy level $\mathcal{E}$ of a battery-powered BS has to be designed such that a certain rate is guaranteed at each UE and the probability of running out of energy (before replacement or reloading) is smaller than some given threshold χ. Mathematically, this amounts to saying that $\Pr(E_T > \mathcal{E}) \leq \chi$. From the results of Theorem 2, one gets

$$\Pr(E_T > \mathcal{E}) = Q\left(\sqrt{K} \frac{\mathcal{E} - \epsilon}{\sqrt{\Sigma}}\right) \qquad (68)$$

from which it follows that

$$\mathcal{E} \geq \frac{\sqrt{\Sigma}}{\sqrt{K}} Q^{-1}(\chi) + \epsilon. \qquad (69)$$

Fig. 5 illustrates the battery level $\mathcal{E}$ as a function of rate for different values of K when $N = 128, \beta = 4$ and the replacing (or recharging) time T is 12 hrs. Marks indicate simulation results while solid lines are obtained theoretically.
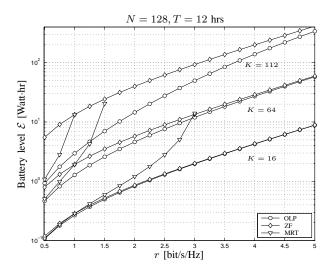
Fig. 5. Battery level that is required by OLP, ZF and MRT to meet $\Pr(E_T > \mathcal{E}) \leq 1\%$ when $K = 16, 64$ or 112.
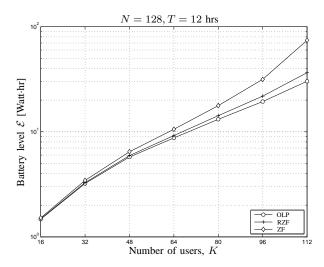

Fig. 7. Cell radius obtained with OLP, ZF and MRT to minimize the energy consumption per unit area when $K = 16, 64$ or 112.

Fig. 6. Battery level that is required by OLP, RZF and ZF to meet $\Pr(E_T > \mathcal{E}) \leq 1\%$ when $r_k \in [1, 4]$ bit/s/Hz.

Fig. 8. Cell radius obtained with OLP, RZF and ZF to minimize the energy consumption per unit area when $r_k \in [1, 4]$ bit/s/Hz.

Fig. 6 reports the values of $\mathcal{E}$ as a function of K in the same operating conditions of Fig. 5 with the only difference that now the UE rate requirements take values within the interval $[1, 4]$ bit/s/Hz. From the results of Figs. 5 and 6, it follows that OLP and RZF provide a substantial energy saving with respect to ZF when K increases. As expected, the saving is more relevant for moderate values of rate requirements r since ZF is known to be suboptimal in that regime. From Fig. 5, it follows that for low data rates (in the range of $r = 0.5$ bit/s/Hz) MRT requires the same battery level of OLP while a substantial increase is observed as r grows up to $\log_2(1 + 1/c)$.

Note that the condition $\chi = 1\%$ makes the necessary battery level $\mathcal{E}$ be substantially higher than $\mathsf{E}[E_T]$. It is also worth observing that the above value accounts only for the energy required to transmit the signal $\mathbf{s}(t)$ within the time interval T. An accurate design of the battery level should also take into account the power amplifier efficiency as well as the energy required for digital signal processing, channel coding and decoding, channel estimation and precoding, and so forth
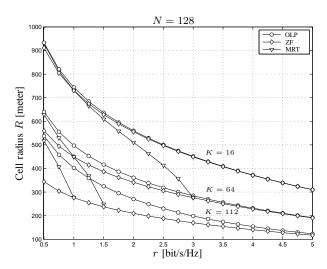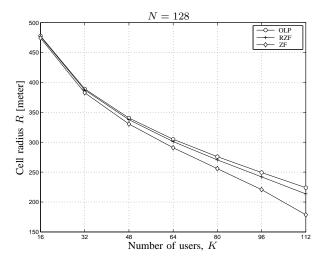
(see [30] for more details). However, all these quantities can be somehow quantified off-line and easily added to $\mathcal{E}$ for a correct design.

C. Optimization of the cell radius

Assume that a system designer must optimize the cell radius for minimizing the average energy consumption per unit area defined as:

$$\mathcal{F} = \frac{\epsilon + \vartheta T}{R^2} \tag{70}$$

where ϑ accounts for the power consumption that is needed for running a BS. Taking the derivative of (70) with respect to R yields

$$R^\star = \bar{x} \sqrt[\beta]{\left(1 + 2L_{\bar{x}}\frac{\eta}{c\gamma}\frac{\vartheta}{\sigma^2}\right)\frac{\beta + 2}{\beta - 2}}. \tag{71}$$

Figs. 7 and 8 show the optimal cell radius in the same setting of Figs. 5 and 6, respectively. Marks indicate simulation results

Fig. 9. Outage probability $\Pr\left(\frac{E_T}{T} > \alpha\right)$ of RZF and ZF with imperfect CSI, $K = 32$, $T = 3, 12$ hrs and $\tau_i^2 = \tau^2 = 0.05\ \forall i$.

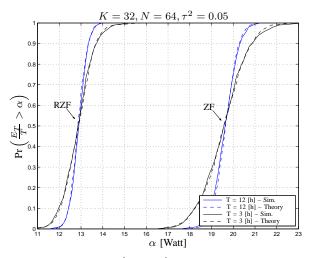Fig. 10. Outage probability $\Pr\left(\frac{E_T}{T} > \alpha\right)$ of RZF and ZF with imperfect CSI, $K = 32$, $T = 3, 12$ hrs and $\tau_i^2 = \tau^2 = 0.15\ \forall i$.

while solid lines are obtained theoretically. The parameter ϑ is set to 18 Watt and corresponds to the fixed power consumption required for control signals, backhaul, and so forth (this value is taken from [31]). From the results of Figs. 7 and 8, it follows that OLP and RZF allow a substantial increase of the coverage area compared to ZF especially when K is large and low rate requirements are imposed.

D. Imperfect CSI

So far, the analysis has been carried out under the assumption of perfect knowledge of $l(\mathbf{x}_k(t))$ and $\mathbf{h}_k(t)$ for $k = 1, 2, \ldots, K$. Observe that $l(\mathbf{x}_k(t))$ corresponds to the average channel attenuation, which changes in time (roughly) three orders of magnitudes slower that the fast fading channels $\mathbf{h}_k(t)$ [32]. More specifically, the characteristic time of change of $\mathbf{h}_k(t)$ is the time it takes for the UE k to move by a wavelength, i.e. $\sim \lambda\xi/\ell$ while that of $l(\mathbf{x}_k(t))$ is $\sim R\xi/\ell$. Therefore, $l(\mathbf{x}_k(t))$ changes far slower than $\mathbf{h}_k(t)$ by a factor $\sim R/\lambda$. In practice, this means that $l(\mathbf{x}_k(t))$ maintains constant for a sufficiently large number of reception phases to be accurately estimated at the BS (for example, through measurements of the received signal strength indicator). This makes it reasonable to assume perfect knowledge of $l(\mathbf{x}_k(t))$. On the other hand, the fast variations of $\mathbf{h}_k(t)$ might result into large estimation errors or outdated estimates (especially in high mobility environments) and thus should be taken into account to avoid a severe degradation of the system performance. This is why the impact of imperfect knowledge of $\mathbf{h}_k(t)$ has been extensively investigated in multi-user MIMO systems either in the finite number of antennas and users regime (e.g., [33]–[35]) or in the asymptotic regime [16], [36], [37]. While the power minimization problem with imperfect CSI is still much open for OLP in both regimes, it can be handled for both ZF and RZF. In the asymptotic regime, one might for example use the results illustrated in [37].

Although applied to a completely different context, the analytical results in [37] can be adapted to the system un-
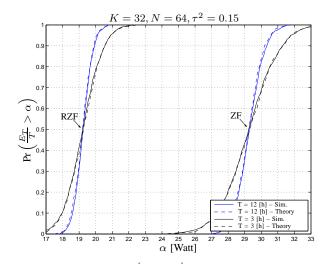
der investigation using standard random matrix theory tools. Omitting the mathematical details, from the results of Section IV in [37] the asymptotic power consumption with imperfect CSI takes form:

$$\overline{P}(t) = \frac{c\sigma^2}{\eta'}\frac{1}{K}\sum_{i=1}^{K}\frac{\gamma_i'}{l(\mathbf{x}_i(t))} \tag{72}$$

where $\gamma_i' = \frac{\gamma_i}{1-\tau_i^2}$ and

$$\eta' = \begin{cases} 1 - c - c\sum_{i=1}^{K}\gamma_i'\tau_i^2 & \text{for ZF} \\ 1 - c\frac{\sum_{i=1}^{K}\gamma_i}{K+\sum_{i=1}^{K}\gamma_i} - c\sum_{i=1}^{K}\gamma_i'\tau_i^2 & \text{for RZF} \end{cases} \tag{73}$$

with τ_i accounting for the accuracy or quality of the estimate of $\mathbf{h}_i(t)$ [37], i.e., $\tau_i = 0$ corresponds to perfect CSI, whereas for $\tau_i = 1$ the CSI is completely uncorrelated to the true channel. As seen, (72) has the same general form of (53) when imperfect CSI is available. Similar conclusions can be drawn for MRT. To validate these analytical results, Figs. 9 – 10 show the CDF of E_T/T in the same operating conditions of Fig. 3 when $\forall i\ \tau_i^2 = \tau^2 = 0.1$ and $\tau_i^2 = \tau^2 = 0.2$, respectively. The computation of ϵ and Σ is performed through (60) and (61) simply replacing η and $\{\gamma_i; i = 1, 2, \ldots, K\}$ with η' and $\{\gamma_i'; i = 1, 2, \ldots, K\}$. As seen, in both cases the simulation results match with the theoretical ones obtained from Theorem 2 using the asymptotic power given by (72). Comparing the results of Figs. 9 – 10 with those of Fig. 3, it is seen that imperfect CSI leads to an increase of the average energy consumption.

VII. CONCLUSIONS AND DISCUSSIONS

In this work, we have studied the energy consumption dynamics in the downlink of a single-cell MIMO network with N antennas at the BS in which K UEs move around according to a random walk model and linear precoding is used to guarantee target rates. The analysis has been conducted

for a finite time interval of length T when $K, N \rightarrow \infty$ with a fixed ratio when perfect CSI is available at the BS. Under these assumptions, we have shown that the energy consumption converges in distribution to a Gaussian random variable for most of the common precoding techniques and we have computed its mean and variance analytically. We have shown that user mobility plays the dominant role in determining the energy fluctuations. In particular, it turns out that the variance of energy consumption scales as $1/K$ rather than as $1/K^2$ as it happens for motionless UEs. Numerical results have been used to show that the analytical expressions yield accurate approximations when K is of the order of tens (see for example Table II). As an application of these results, we have dimensioned a battery-powered BS to satisfy a certain probability of running out of energy and we have also computed the cell radius to minimize the energy per unit area. The imperfect CSI has been also briefly addressed.

The analysis has also pointed out that in addition to K there is another parameter that plays a key role in describing the energy fluctuations. This parameter is given by DT/R^2, where R is the cell radius and D is the so-called diffusion constant of random walks (or Brownian motions). The quantity DT/R^2 roughly describes to what extent the UEs have been around in the cell area. When it is relatively small (as it has been assumed in this work), each UE has not had the chance to visit most of the cell and thus the energy consumed to serve all of them is a random quantity whose fluctuations are basically controlled by K. However, when DT/R^2 is relatively large, then each UE has likely had the time to move around all the coverage area, and thus the energy consumed to serve each of them becomes approximately deterministic. This result suggests an alternative regime for the asymptotic analysis of energy consumption: the finite K and N but large T (or more properly large DT/R^2) regime. This case has been analyzed in some detail in [38] but it is still pretty much open. In addition to this, another interesting direction for future work is to consider other mobility models such as the Lévy flight, which seems to contain some important statistical similarities with human mobility [8], [9]. The extension of the analysis to a multi-cell network with imperfect CSI and pilot contamination is also an interesting topic for future research.

APPENDIX A
PROOF OF THEOREM 1

We start assuming that the SINR constraints are such that the following assumption is satisfied:

$$\lim_{K \rightarrow \infty} \sup \frac{1}{N} \sum_{i=1}^{K} \frac{\gamma_i}{1 + \gamma_i} < 1. \tag{74}$$

Under these circumstances, η in (17) is a positive quantity. Observe that $c \in (0,1)$ is sufficient for the above condition to be satisfied. To proceed further, we rewrite (13) as follows (using simple calculus)

$$\frac{\gamma_k}{\lambda_k^\star} = \frac{1}{N} \mathbf{h}_k^H \left(\frac{1}{N} \sum_{i=1, i \neq k}^{K} \lambda_i^\star \mathbf{h}_i \mathbf{h}_i^H + N \mathbf{I}_N \right)^{-1} \mathbf{h}_k \tag{75}$$

from which letting

$$d_k = \frac{\gamma_k}{\lambda_k^\star l(\mathbf{x}_k)} \tag{76}$$

and using (1) we eventually obtain

$$d_k = \frac{1}{N} \mathbf{w}_k^H \left(\frac{1}{N} \sum_{i=1, i \neq k}^{K} \frac{\gamma_i}{d_i} \mathbf{w}_i \mathbf{w}_i^H + \mathbf{I}_N \right)^{-1} \mathbf{w}_k. \tag{77}$$

Assume that the quantities $\{d_k\}$ are well defined, positive and such that $0 \leq d_1 \leq d_2 \leq \ldots \leq d_K$. Then, using monotonicity arguments, from (77) it follows that

$$d_K \leq \frac{1}{N} \mathbf{w}_K^H \left(\frac{1}{N} \sum_{i=1}^{K-1} \frac{\gamma_i}{d_K} \mathbf{w}_i \mathbf{w}_i^H + \mathbf{I}_N \right)^{-1} \mathbf{w}_K =$$

$$= d_K \frac{1}{N} \mathbf{w}_K^H \left(\frac{1}{N} \sum_{i=1}^{K-1} \gamma_i \mathbf{w}_i \mathbf{w}_i^H + d_K \mathbf{I}_N \right)^{-1} \mathbf{w}_K \tag{78}$$

or, equivalently,

$$1 \leq \frac{1}{N} \mathbf{w}_K^H \left(\frac{1}{N} \sum_{i=1}^{K-1} \gamma_i \mathbf{w}_i \mathbf{w}_i^H + d_K \mathbf{I}_N \right)^{-1} \mathbf{w}_K. \tag{79}$$

Assume now that d_K is infinitely often larger than $\eta + \ell$ with η given by (17) and $\ell > 0$ some positive value [15]. Let us restrict ourselves to such a subsequence. From (79), using monotonicity arguments we obtain

$$1 \leq \frac{1}{N} \mathbf{w}_K^H \left(\frac{1}{N} \sum_{i=1}^{K-1} \gamma_i \mathbf{w}_i \mathbf{w}_i^H + (\eta + \ell) \mathbf{I}_N \right)^{-1} \mathbf{w}_K. \tag{80}$$

Applying standard results in random matrix theory along with the union bound and Markov inequality one gets [15]

$$\frac{1}{N} \mathbf{w}_K^H \left(\frac{1}{N} \sum_{i=1}^{K-1} \gamma_i \mathbf{w}_i \mathbf{w}_i^H + (\eta + \ell) \mathbf{I}_N \right)^{-1} \mathbf{w}_K - e(\ell) \rightarrow 0 \tag{81}$$

with $e(\ell)$ being the unique positive solution to

$$e(\ell) = \left(\frac{1}{N} \sum_{i=1}^{K} \frac{\gamma_i}{1 + \gamma_i e(\ell)} + \eta + \ell \right)^{-1}. \tag{82}$$

From (81), recalling (80) yields

$$\lim_{K \rightarrow \infty} \inf e(\ell) \geq 1. \tag{83}$$

Using the fact that $e(0) = 1$ (recall that η is defined as in (17)) and that $e(\ell)$ is a decreasing function of ℓ, it can be proved that for any $\ell > 0$ [15]

$$\lim_{K \rightarrow \infty} \sup e(\ell) < 1. \tag{84}$$

This however goes against (83) and creates a contradiction on the initial hypothesis that $d_K > \eta + \ell$ infinitely often. Therefore, we must admit that $d_K \leq \eta + \ell$ for all large values of K. Reverting all inequalities and using similar arguments yields $d_1 \geq \eta - \ell$ for all large values of K. Putting all these results together yields $\eta - \ell \leq d_1 \leq d_2 \leq \cdots \leq d_K \leq \eta + \ell$ from which we may write $\max_{k=1,2,\ldots,K} |d_k - \eta| \leq \ell$ for all large values of K [15]. Taking a countable sequence of ℓ going to

zero, we eventually obtain $\max_{k=1,2,\ldots,K} |d_k - \eta| \to 0$ from which using (76) and assuming that $\lim_{K\to\infty} \sup \frac{\gamma_k}{l(\mathbf{x}_k)} < \infty$ the result in (19) follows.

APPENDIX B
RANDOM WALK AND BROWNIAN MOTION

The convergence of a random walk towards the Brownian motion model is basically due to the central limit theorem. More precisely, observe that $\mathbf{x}_k(t) = \sum_{i=1}^{\lfloor t/\xi \rfloor} \Delta \mathbf{x}_k(i)$ has zero-mean and variance $\mathsf{E}\left[\|\mathbf{x}_k(t)\|^2\right] = t\ell^2/\xi$. Assume now that t is fixed and $\xi \to 0$, in the sense that $t/\xi \to \infty$. Thus, the variance of $\mathbf{x}_k(t)$ can be kept finite only if the ratio ℓ^2/ξ is kept fixed and finite. In this limit, $\mathbf{x}_k(t)$ becomes a Brownian motion with diffusion coefficient D given by $D = \ell^2/(4\xi)$. In practical systems, however, neither the step size ℓ nor the corresponding time ξ vanishes. Nevertheless, the equivalence still holds true since we are basically interested in the long term (corresponding to large values of t) and large distance statistics of a random walk with finite ℓ and ξ wherein each walker changes position according to a transition rule that only depends on its current position. In these circumstances, the random walk boils down to a Markov process, whose statistical properties in the large time and distance regime are still well captured by simply treating it as a Brownian motion.

A. Transition probability

The transition probability $\mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t')$ of a Brownian motion in a given area $\mathcal{C}$ is defined as the probability of being at a location $\mathbf{x}'$ at time t given that the walker was at $\mathbf{x}$ at time $t' < t$. Mathematically, $\mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t')$ is obtained solving the following diffusion equation [39]

$$\frac{\partial \mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t')}{\partial t} = \frac{D}{2} \nabla^2 \mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t') \tag{85}$$

under appropriate boundary conditions (specifying the behaviour of the user when reaching the boundary $\partial \mathcal{C}$ of $\mathcal{C}$) and subject to:

$$\int_{\mathcal{C}} \mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t') d\mathbf{x} = 1 \tag{86}$$

$$\lim_{t \to t'} \mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t') = \delta^2(\mathbf{x} - \mathbf{x}'). \tag{87}$$

Note also that $\mathbb{P}(\mathbf{x}, \mathbf{x}'; t-t') = \mathbb{P}(\mathbf{x}', \mathbf{x}; t-t')$ since the Laplacian operator in (85) is both real and Hermitian. Although a large variety of boundary conditions can be imposed [7], the most typical ones are the following two: *i)* the walker exits $\mathcal{C}$ when it reaches the boundary $\partial \mathcal{C}$; *ii)* the walker bounces back into $\mathcal{C}$ when it hits $\partial \mathcal{C}$. The former choice requires to impose $\mathbb{P}(\mathbf{x} \in \partial \mathcal{C}, \mathbf{x}'; t - t') = 0$ while the latter needs $\mathbf{r}^T \nabla \mathbb{P}(\mathbf{x} \in \partial \mathcal{C}, \mathbf{x}'; t - t') = 0$ with $\mathbf{r}$ being the unit vector along the radial direction.[6]

---

[6] Observe that if the walker bounces back, then $\mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t')$ is always properly normalized in $\mathcal{C}$ (since the probability of exiting $\mathcal{C}$ is zero). This condition is known in the literature as Neuman boundary condition.

B. Computation of the transition probability

The transition probability $\mathbb{P}(\mathbf{x}, \mathbf{x}'; t-t')$ solving (85) over a bounded domain $\mathcal{C}$ can be expressed in the following general form [40]:

$$\mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t') = \sum_n \sum_m g_{n,m}^*(\mathbf{x}) g_{n,m}(\mathbf{x}') e^{-\epsilon_{n,m}(t-t')} \tag{88}$$

where $g_{n,m}(\mathbf{x})$ is the (n, m)th (normalized) eigenfunction of the Laplacian operator ∇^2 in $\mathcal{C}$ satisfying the boundary conditions and $\epsilon_{n,m}$ is the corresponding eigenvalue. The above result basically follows from the fact that $\{g_{n,m}(\mathbf{x})\}$ form a basis in the Hilbert space $L^2(\mathcal{C})$ and thus any function can be written in terms of $\{g_{n,m}(\mathbf{x})\}$ (we refer the interested reader to [40] for a more detailed discussion on this). Observe that $\mathbb{P}(\mathbf{x}, \mathbf{x}'; t - t')$ in the form of (88) meets (86) and (87) since the eigenfunctions are such that [40]:

$$\int_{\mathcal{C}} g_{n_1,m_1}^*(\mathbf{x}) g_{n_2,m_2}(\mathbf{x}) d\mathbf{x} = \delta_{n_1,n_2} \delta_{m_1,m_2} \tag{89}$$

$$\sum_n \sum_m g_{n,m}^*(\mathbf{x}) g_{n,m}(\mathbf{x}') = \delta^2(\mathbf{x} - \mathbf{x}') \tag{90}$$

where $\delta^2(\mathbf{x})$ is the two-dimensional Dirac delta function. The conditions above can also be used to show that if $t'' < t' < t$ then

$$\int_{\mathcal{C}} \mathbb{P}(\mathbf{x}, \mathbf{x}'; t-t') \mathbb{P}(\mathbf{x}', \mathbf{x}''; t'-t'') d\mathbf{x}' = \mathbb{P}(\mathbf{x}, \mathbf{x}''; t-t'') \tag{91}$$

which is a consequence of the Markov property of Brownian motions.

Computing the eigenfunctions in closed-form is not always possible for any region $\mathcal{C}$. However, if $\mathcal{C}$ has a circular symmetry with $|\mathcal{C}| = \pi R^2$ and the reflecting boundary condition is imposed, then the eigenfunctions can be expressed in polar coordinates (r, ϕ) as [40]:

$$g_{n,m}(r, \phi) = (A_{n,m})^{1/2} J_m\left(\frac{\kappa_{n,m} r}{R}\right) e^{\mathrm{i}m\phi} \tag{92}$$

where $\kappa_{n,m}$ is the nth non-trivial zero of the first derivative of the $m-$Bessel function $J_m(\cdot)$ and $A_{n,m}^{-1} = \pi R^2 \left(J_m^2(\kappa_{n,m}) + J_{m-1}^2(\kappa_{n,m}) \right)$ and the eigenvalue is $\epsilon_{n,m} = \kappa_{n,m}^2$.

APPENDIX C
PROOFS OF LEMMAS 3 AND 4

In this appendix, we rely on the Brownian motion model (see Appendix B) to compute the mean value and covariance of $l^{-1}(\mathbf{x}_k(t))$. The former requires to evaluate

$$\mathsf{E}_{\mathbf{X}_k}\left[l^{-1}(\mathbf{x}_k(t))\right] = \mathsf{E}_{\mathbf{x}_k(0)}\left[\mathsf{E}_{\mathbf{X}_k|\mathbf{x}_k(0)}\left[l^{-1}(\mathbf{x}_k(t))|\mathbf{x}_k(0)\right]\right] \tag{93}$$

where $\mathsf{E}_{\mathbf{x}_k(0)}\left[z|\mathbf{x}_k(0)\right]$ denotes the conditional expectation of z with respect to the initial position $\mathbf{x}_k(0) \in \mathcal{C}$. On the other hand, the covariance of $l^{-1}(\mathbf{x}_k(t))$ takes the form in (94). To simplify the notation, in the following derivations we drop the UE index k and relabel $\mathbf{x}_k(t)$, $\mathbf{x}_k(t'')$ and $\mathbf{x}_k(0)$ as follows $\mathbf{x}_k(t) \to \mathbf{x}$, $\mathbf{x}_k(t'') \to \mathbf{x}''$ and $\mathbf{x}_k(0) \to \mathbf{x}'$. In addition, we call $F(\mathbf{x}, t|\mathbf{x}', 0)$ the probability that the UE reaches $\mathbf{x}$ at a generic time t if its position at time 0 is $\mathbf{x}'$. From Appendix B,

$$\mathrm{COV}_{\mathbf{X}_k}\left[l^{-1}(\mathbf{x}_k(t)), l^{-1}(\mathbf{x}_k(t''))\right] = \mathsf{E}_{\mathbf{x}_k(0)}\left[\mathsf{E}_{\mathbf{X}_k|\mathbf{x}_k(0)}\left[l^{-1}(\mathbf{x}_k(t))l^{-1}(\mathbf{x}_k(t''))|\mathbf{x}_k(0)\right]\right] -$$
$$- \mathsf{E}_{\mathbf{x}_k(0)}\left[\mathsf{E}_{\mathbf{X}_k|\mathbf{x}_k(0)}\left[l^{-1}(\mathbf{x}_k(t))|\mathbf{x}_k(0)\right]\mathsf{E}_{\mathbf{X}_k|\mathbf{x}_k(0)}\left[l^{-1}(\mathbf{x}_k(t''))|\mathbf{x}_k(0)\right]\right]. \quad (94)$$

we have that $F(\mathbf{x},t|\mathbf{x}',0) = \mathbb{P}(\mathbf{x},\mathbf{x}';t)$ with $\mathbb{P}(\mathbf{x},\mathbf{x}';t)$ being the solution of the diffusion equation in (85).

We start computing the mean of $l^{-1}(\mathbf{x}(t))$. Under the above assumptions, we may write

$$\mathsf{E}_{\mathbf{X}}\left[l^{-1}(\mathbf{x}(t))\right] = \frac{1}{|\mathcal{C}|}\iint_{\mathcal{C}}l^{-1}(\mathbf{x})\,\mathbb{P}(\mathbf{x},\mathbf{x}';t)d\mathbf{x}d\mathbf{x}' \quad (95)$$

where we have taken into account that $\mathbf{x}'$ is uniformly distributed within the circular cell of area $|\mathcal{C}| = \pi R^2$. Recalling[7] that $\mathbb{P}(\mathbf{x},\mathbf{x}';t) = \mathbb{P}(\mathbf{x}',\mathbf{x};t)$ and using $\int_{\mathcal{C}}\mathbb{P}(\mathbf{x}',\mathbf{x};t)d\mathbf{x}' = 1$ (see (86) in Appendix B) one gets

$$\mathsf{E}_{\mathbf{X}}\left[l^{-1}(\mathbf{x}(t))\right] = \frac{1}{|\mathcal{C}|}\int_{\mathcal{C}}l^{-1}(\mathbf{x})d\mathbf{x}. \quad (96)$$

The covariance requires to evaluate the two terms in the right-hand-side of (94). Let us start with the first one, which is given by

$$\mathsf{E}_{\mathbf{x}'}\left[\mathsf{E}_{\mathbf{X}|\mathbf{x}'}\left[l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')|\mathbf{x}'\right]\right] = \quad (97)$$
$$\frac{1}{|\mathcal{C}|}\iiint_{\mathcal{C}}l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')F(\mathbf{x},t,\mathbf{x}'',t''|\mathbf{x}',0)d\mathbf{x}d\mathbf{x}''d\mathbf{x}'. \quad (98)$$

Thanks to the Markov property of Brownian motions, one gets

$$F(\mathbf{x},t,\mathbf{x}'',t''|\mathbf{x}',0) = F(\mathbf{x},t|\mathbf{x}'',t'',\mathbf{x}',0)F(\mathbf{x}'',t''|\mathbf{x}',0)$$
$$= F(\mathbf{x},t|\mathbf{x}'',t'')F(\mathbf{x}'',s|\mathbf{x}',0) \quad (99)$$

or, equivalently, $F(\mathbf{x},t,\mathbf{x}'',t''|\mathbf{x}',0) = \mathbb{P}(\mathbf{x},\mathbf{x}'';t-t'')\mathbb{P}(\mathbf{x}'',\mathbf{x}';t'')$. Plugging this result into (97) and using $\mathbb{P}(\mathbf{x}'',\mathbf{x}';t'') = \mathbb{P}(\mathbf{x}',\mathbf{x}'';t'')$ with $\int_{\mathcal{C}}\mathbb{P}(\mathbf{x}',\mathbf{x}'';t'')d\mathbf{x}' = 1$ one gets

$$\mathsf{E}_{\mathbf{x}'}\left[\mathsf{E}_{\mathbf{X}|\mathbf{x}'}\left[l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')|\mathbf{x}'\right]\right] = \quad (100)$$
$$\frac{1}{|\mathcal{C}|}\iint_{\mathcal{C}}l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')\,\mathbb{P}(\mathbf{x},\mathbf{x}'';t-t'')d\mathbf{x}d\mathbf{x}''. \quad (101)$$

We are now left with the computation of the second term in (94), which is explicitly given by

$$\mathsf{E}_{\mathbf{x}'}\left[\mathsf{E}_{\mathbf{X}|\mathbf{x}'}\left[l^{-1}(\mathbf{x})|\mathbf{x}'\right]\mathsf{E}_{\mathbf{X}|\mathbf{x}'}\left[l^{-1}(\mathbf{x}'')|\mathbf{x}'\right]\right] = \quad (102)$$
$$\frac{1}{|\mathcal{C}|}\iiint_{\mathcal{C}}l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')\,\mathbb{P}(\mathbf{x},\mathbf{x}';t)\,\mathbb{P}(\mathbf{x}'',\mathbf{x}';t'')d\mathbf{x}d\mathbf{x}''d\mathbf{x}'. \quad (103)$$

Using $\mathbb{P}(\mathbf{x}'',\mathbf{x}';t'') = \mathbb{P}(\mathbf{x}',\mathbf{x}'';t'')$ and (91) lead to

$$\mathsf{E}_{\mathbf{x}'}\left[\mathsf{E}_{\mathbf{X}|\mathbf{x}'}\left[l^{-1}(\mathbf{x})|\mathbf{x}'\right]\mathsf{E}_{\mathbf{X}|\mathbf{x}'}\left[l^{-1}(\mathbf{x}'')|\mathbf{x}'\right]\right] = \quad (104)$$
$$\frac{1}{|\mathcal{C}|}\iint_{\mathcal{C}}l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')\,\mathbb{P}(\mathbf{x},\mathbf{x}'';t+t'')d\mathbf{x}d\mathbf{x}'' \quad (105)$$

[7]This comes from the fact that the probability of reaching (within the same time interval) the point $\mathbf{x}$ starting from $\mathbf{x}'$ is the same as that of reaching $\mathbf{x}'$ starting from $\mathbf{x}$ (see Appendix B).

Plugging the above results together into (94), we eventually obtain

$$\mathrm{COV}_{\mathbf{X}}\left[l^{-1}(\mathbf{x}), l^{-1}(\mathbf{x}'')\right] = \quad (106)$$
$$\frac{1}{|\mathcal{C}|}\iint_{\mathcal{C}}l^{-1}(\mathbf{x})l^{-1}(\mathbf{x}'')Q(\mathbf{x},\mathbf{x}'';t,t'')d\mathbf{x}d\mathbf{x}''. \quad (107)$$

where $Q(\mathbf{x},\mathbf{x}'';t,t'')$ is defined as

$$Q(\mathbf{x},\mathbf{x}'';t,t'') = \mathbb{P}(\mathbf{x},\mathbf{x}'';t-t'') - \mathbb{P}(\mathbf{x},\mathbf{x}'';t+t''). \quad (108)$$

We are now left with substituting into (106) the closed-form expressions of $\mathbb{P}(\mathbf{x},\mathbf{x}'';t+t'')$ and $\mathbb{P}(\mathbf{x},\mathbf{x}'';t-t'')$ as obtained through (88) using the eigenfunctions given by (92). After standard but lengthy computations (not shown for space limitations), we eventually get the result in (56) of Lemma 4.

APPENDIX D
PROOF OF THEOREM 2

In this appendix, we outline the proof for the central limit theorem of the energy consumption E_T. The first step is to observe that since the underlying Brownian motions of UEs are continuous, we may think of the integral in (9) as a limit of a finite sum, i.e.,

$$E_T = \int_0^T P(t)dt = \lim_{L\to\infty}\frac{T}{L}\sum_{n=1}^{L}P(\mathbf{x}(n),\mathbf{W}(n)) \quad (109)$$

where $\mathbf{x}(n)$ and $\mathbf{W}(n)$ corresponds to $\mathbf{x}(t)$ and $\mathbf{W}(t)$ evaluated at $t = nT/L = n\xi$, respectively. In writing the above equation, we have explicitly specified the dependence of $P(t)$ on the UE positions $\mathbf{x}(n) = \{\mathbf{x}_k(n); k=1,2,\ldots,K\}$ and on the fading coefficients $\mathbf{W}(n) = \{\mathbf{W}_k(n); k=1,2,\ldots,K\}$. We assume that for any k the channel fading coefficients in $\mathbf{W}_k(n)$ are independent over n since the fast fading channel $\mathbf{W}_k(t)$ decorrelates approximately as $\sim \lambda\xi/\ell$, which is much smaller than the sampling time ξ.

To proceed further, we let $\Delta E_T = E_T - \mathsf{E}[E_T]$ be the difference between E_T and its average value $\mathsf{E}[E_T]$ and rewrite ΔE_T as the sum of two terms, i.e., $\Delta E_T = \Delta_1 + \Delta_2$ with

$$\Delta_1 = E_T - \int_0^T \overline{P}(t)dt \quad (110)$$

$$\Delta_2 = \int_0^T \overline{P}(t)dt - \mathsf{E}[E_T] \quad (111)$$

Observe that the variance of Δ_1 is just T times the variance of $P(t) - \overline{P}(t)$ averaged over user locations, i.e.,

$$\mathsf{VAR}[\Delta_1] = T\mathsf{E}_X\left[\mathsf{VAR}_{\mathbf{W}}\left[P(t) - \overline{P}(t)\right]\right]. \quad (112)$$

In [20], the authors show that $K\mathsf{VAR}\left[P(t) - \overline{P}(t)\right]$ is finite in the limit $K \to \infty$. Since the pathloss function is bounded

from below, it follows that $K\mathsf{E}_X\left[\mathsf{VAR}\left[P(t)-\overline{P}(t)\right]\right]$ is also bounded, i.e.,

$$\lim_{K\to\infty}\sup K\mathsf{E}_X\left[\mathsf{VAR}\left[P(t)-\overline{P}(t)\right]\right] < \infty. \quad (113)$$

This means that the fluctuations of Δ_1 are of order $O(K^{-1})$.

To analyze the second term Δ_2, we use (53) and observe that the integral in (110) can be written as a sum over independent Brownian paths of the K users in the system. Using the results of Appendix C, the term corresponding to kth UE has mean and variance given by

$$\frac{1}{K}\frac{c\sigma^2}{\eta}\int_0^T \mathsf{E}_{\mathbf{x}_k}\left[\frac{\gamma_k}{\ell(\mathbf{x}_k(t))}\right] = \frac{\gamma_k}{|\mathcal{C}|}\frac{T}{K}\frac{c\sigma^2}{\eta}\int_{\mathcal{C}}\frac{1}{l(\mathbf{x}_k(\tau))}d\mathbf{x}_k(\tau) \quad (114)$$

and

$$\frac{1}{K^2}\left(\frac{c\sigma^2}{\eta}\right)^2\iint_0^T \mathsf{COV}_{\mathbf{x}_k}\left[\frac{\gamma_k}{\ell(\mathbf{x}_k(\tau))},\frac{\gamma_k}{\ell(\mathbf{x}_k(s))}\right] = $$
$$= \gamma_k^2\frac{TR^2}{DK^2}\left(\frac{c\sigma^2}{\eta}\right)^2\Theta \quad (115)$$

where Θ is obtained as in (57). Hence, Δ_2 has finite mean and fluctuations of order $K^{-1/2}$ and its variance can be obtained summing (115) over k. This leads to

$$\mathsf{VAR}\left[\Delta_2\right] = \left(\frac{c\sigma^2}{\eta}\right)^2\left(\frac{1}{K}\sum_{i=1}^K\gamma_i^2\right)\frac{TR^2}{KD}\Theta. \quad (116)$$

As a result of the above, we eventually have that

$$\sqrt{K}\left(\frac{E_T-\epsilon}{\sqrt{\Sigma}}\right)\xrightarrow[K,N\to\infty]{\mathcal{D}}\mathcal{N}(0,1) \quad (117)$$

as stated in Theorem 2. As seen, the proof is basically articulated in two steps. First, we exploit the results in [20] to point out that that the power variance due to the fast fading scales as $O(K^{-2})$ for all considered schemes. Then, we take advantage of (53) and use the results in Lemmas 3 and 4 to deal with the UE movements and to prove the results.

ACKNOWLEDGMENT

REFERENCES

[1] L. Sanguinetti, A. Moustakas, E. Bjornson, and M. Debbah, "Energy consumption in multi-user MIMO systems: Impact of user mobility," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, Florence, Italy, May 2014.

[2] "Global mobile data traffic forecast update, 2010-2015 (Cisco visual networking index)," *White Paper*, Feb. 2011.

[3] A. Fehske, G. Fettweis, J. Malmodin, and G. Biczok, "The global footprint of mobile communications: the ecological and economic perspective," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 55–62, Aug. 2011.

[4] Y. Chen, S. Zhang, S. Xu, and G. Li, "Fundamental trade-offs on green wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 30–37, June 2011.

[5] L. Decreusefond, P. Martins, and T.-T. Vu, "Modeling energy consumption in cellular networks," in *in Proceedings of the 25th International Teletraffic Conference, France*, 2013, pp. 1–5.

[6] W. Paul and J. Baschnagel, *Stochastic Processes: From Physics to Finance*. Springer Verlag, 2013, vol. 2nd Edition.

[7] P. Mörters and Y. Peres, *Brownian motion*. Cambridge University Press, 2010.

[8] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, "On the Levy-walk nature of human mobility," *IEEE/ACM Transactions on Networking*, vol. 19, no. 3, pp. 630 – 643, June 2011.

[9] N. Scafetta, "Understanding the complexity of the Levy-walk nature of human mobility with a multi-scale cost/benefit model," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 21, no. 4, 2011.

[10] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[11] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2508–2523, May 2012.

[12] Y. Huang, C. W. Tan, and B. Rao, "Large system analysis of power minimization in multiuser MISO downlink with transmit-side channel correlation," in *Int. Symp. Inf. Theory and its Applications (ISITA)*, Oct. 2012, pp. 240–244.

[13] R. Zakhour and S. Hanly, "Base station cooperation on the downlink: Large system analysis," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2079–2106, Apr. 2012.

[14] A. T. H. Asgharimoghaddam and N. Rajatheva, "Decentralizing the optimal multi-cell beamforming via large system analysis," in *Proc. of IEEE Int. Conf. Commun.*, Sydney, Australia, June 2014.

[15] R. Couillet and M. McKay, "Large dimensional analysis and optimization of robust shrinkage covariance matrix estimators," *Journal of Multivariate Analysis*, vol. 131, no. 0, pp. 99 – 120, 2014.

[16] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.

[17] R. Muharar, R. Zakhour, and J. Evans, "Optimal power allocation and user loading for multiuser MISO channels with regularized channel inversion," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5030–5041, Dec. 2013.

[18] L. Sanguinetti, E. Björnson, M. Debbah, and A. Moustakas, "Optimal linear pre-coding in multi-user MIMO systems: A large system analysis," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, Austin, Texas, USA, 2014.

[19] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013.

[20] Z. D. Bai and J. W. Silverstein, "CLT for linear spectral statistics of large dimensional sample covariance matrices," *The Annals of Probability*, vol. 32, pp. 553–605, 2004.

[21] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. & Mobile Comp.: Special Issue on Mobile Ad-Hoc Networking: Research, Trends and Applications*, vol. 2, 2002.

[22] G. Calcev, D. Chizhik, B. Goransson, S. Howard, H. Huang, A. Kogiantis, A. Molisch, A. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, March 2007.

[23] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing – The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441 – 6463, Oct 2013.

[24] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[25] E. Bjornson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142 – 148, July 2014.

[26] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *The Annals of Applied Probability*, vol. 17, no. 3, pp. 875–930, 2007.

[27] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge, New York: Cambridge University Press, 2011.

[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, Inc, 1972.

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*. New York: Academic Press, 1995.

[30] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO

the answer?" *submitted to IEEE Trans. Wireless Comm.*, April 2014. [Online]. Available: http://arxiv.org/abs/1403.6150

[31] G. Auer, O. Blume, V. Giannini, I. Godor, M. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda, *D2.3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown*. INFSO-ICT-247733 EARTH, ver. 2.0, 2012. [Online]. Available: http://www.ict-earth.eu/

[32] I. Viering, H. Hofstetter, and W. Utschick, "Spatial long-term variations in urban, rural and indoor environments," in *the 5th Meeting of COST273*, Lisbon, Portugal, 2002, pp. 103 – 108.

[33] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576–3587, Sept 2009.

[34] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June 2010.

[35] J. Gonzalez-Coma, M. Joham, P. Castro, and L. Castedo, "Power minimization in the multiuser downlink under user rate constraints and imperfect transmitter CSI," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, May 2013.

[36] A. Muller, E. Bjornson, R. Couillet, and M. Debbah, "Analysis and management of heterogeneous user mobility in large-scale downlink systems," in *Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 773 – 777.

[37] L. Sanguinetti, A. L. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD HetNets: A large system analysis," *submitted to J. Sel. Areas Commun.*, July 2014. [Online]. Available: http://arxiv.org/abs/1407.6481

[38] A. L. Moustakas, L. Sanguinetti, and M. Debbah, "Effects of mobility on user energy consumption and total throughput in a massive MIMO system," in *Information Theory Workshop (ITW), 2014 IEEE*, Hobart, Tasmania, Nov 2014, pp. 292–296.

[39] C. Itzykson and J. M. Drouffe, *Statistical Field Theory*. Cambridge: Cambridge Monographs on Mathematical Physics, 1991, vol. 1.

[40] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. New York: J. Wiley & Son, Inc., 1998.