Spatial Coverage Cross-Tier Correlation Analysis for Heterogeneous Cellular Networks

David M. Gutierrez-Estevez, *Student Member, IEEE*, Ian F. Akyildiz, *Fellow, IEEE*, and Etimad A. Fadel

***Abstract*—In the search for improved coverage and capacity, cellular networks are currently undergoing a major transformation. A thoroughly planned architecture comprised of macrocells served by large-coverage expensive base stations (BSs) is evolving toward a much more heterogeneous architecture where the macrocell network is underlaid by one or several tiers of small cells deployed in an irregular and unplanned fashion using universal frequency reuse. Major challenges of this new scenario are the problems of intercell interference (ICI) and cell association. In this paper, these two problems are tackled in a novel way by analyzing and exploiting the inherent spatial cross-tier coverage correlation due to the cochannel ICI for a two-tier network. A mathematical framework for the representation of the two-tier coverage maps and their correlation is developed based on the spatial statistical properties of the signal quality measurements reported by the users to the base station. Several semivariogram-based estimation models are applied and cross-validated. Furthermore, a closed-form expression for the cross-tier coverage correlation function depending only on the estimator's parameters is obtained. In addition, a practical application of this framework is proposed. Cross-tier correlation information is exploited in the design of a new cell association policy based on cell-specific biasing for small cells. Numerical results show that the mathematical framework can provide accurate representations of the coverage fields and their correlation. Moreover, the performance of our proposed correlation-aware cell association policy is shown to be promising enough to encourage further research in this direction.**

***Index Terms*—Cell association, cross correlation, heterogeneous networks (HetNets), interference, random field, semivariogram.**

I. Introduction

THE challenge of the ever-increasing demands for wireless data is shaping the design and development of fourth-generation and beyond cellular systems, such as the Third-Generation Partnership Project (3GPP)'s Long-Term Evolution (LTE) Advanced [1], [2]. However, historical data regarding capacity improvements in wireless networks show that the largest gains ever obtained come mainly from the use of smaller cells [3]. This is the major driver for the new network paradigm of heterogeneous networks (HetNets), where cells of different sizes and different radio access technologies must coexist. The main focus now for every operator across the world is the deployment of HetNets comprised of a macrocell network underlaid by one or several tiers of small cells [4]. Femtocells, metrocells, picocells, or microcells are representative technologies of this family. Tremendous activity on small cells has been conducted in recent years by academic institutions, companies, and standardization bodies, including 3GPP, 3GPP2, and the WiMAX forum [5], [6].

One of the major problems in this new network paradigm is intercell interference (ICI). This interference is particularly detrimental for cochannel scenarios where the operator does not carefully plan the small-cell deployment, unavoidably affecting the coverage maps of both tiers. The coverage effects of the interference, as well as the potential solutions for its proper management, have been extensively studied in the literature in the last few years. Analytical expressions for coverage on multitier networks were derived, for example, in [7], and tailored expressions for multiantenna settings can be found in [8]. However, most of the attention is focused on interference management for two-tier femtocell networks using techniques related to power control, spectrum management, or different access modes of operation [9]–[11]. Nevertheless, none of these previous studies investigate properly the fundamental correlation existing in the coverage maps of both tiers due to the inherent interference, which is a fundamental feature that can be exploited in the design of practical interference management approaches.

In addition, the problem of cell association is another challenge for HetNets. Deciding which users should be attached to which base station (BS) is a trivial problem in the case of homogeneous networks: The signal-to-interference-plus-noise ratio (SINR) is maximized, and the BS providing the highest SINR is always selected. However, this approach in the case of HetNets leads to high underutilization of the small cells since very few users would be attached to them. To solve this problem and increase the capacity gains of small cells, the currently agreed solution is to expand the range of the small cells by introducing a so-called bias factor [12]. However, this increases the interference on the user, thus coupling the ICI and cell association problems [13].

Manuscript received February 25, 2013; revised June 23, 2013 and October 8, 2013; accepted January 12, 2014. Date of publication February 13, 2014; date of current version October 14, 2014. This work was supported in part by the King Abdulaziz City for Science and Technology through the National Policy for Science, Technology, and Innovation Plan under Grant 12-INF2731-03, by the Unit for Science and Technology at King Abdulaziz University, by Fundación Caja Madrid, and by the Georgia Research Alliance. The review of this paper was coordinated by Prof. H.-H. Chen.

D. M. Gutierrez-Estevez is with the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30338 USA (e-mail: david.gutierrez@ece.gatech.edu).

I. F. Akyildiz is with the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30338 USA, and also with the Department of Computer Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: ian@ece.gatech.edu).

E. A. Fadel is with the Department of Computer Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: eafadel@kau.edu.sa).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2306260

In this paper, we propose a novel mathematical framework to characterize the spatial coverage cross-correlation between the macrocell and the small-cell tiers in the downlink. Furthermore, we apply this modeling to develop a solution of the cell association problem based on cell-specific bias. The observation that coverage in both tiers may have some level of correlation in both the spatial and temporal domains due to the existing interference is informally assumed in most previous studies. However, to the best of our knowledge, there exists no previous paper explicitly addressing the spatial coverage cross-tier correlation, and accordingly, there exists no analytical characterization of this correlation. Since we focus on the spatial dimension in this paper, we make use of spatial statistic techniques to derive analytical expressions for the two-tier coverage maps suitable for a subsequent cross-correlation analysis. In particular, we use random field estimation techniques to generate coverage fields, formalized in terms of downlink SINR fields, using the position and SINR values of the users in the network. A cross-validation is performed across different estimation models seeking for the best choice. To the best of our knowledge, only our previous interference management scheme [14] has applied this framework to the present two-tier cellular scenario. In addition, we derive an explicit formulation of the spatial coverage cross-tier correlation function depending on the previous model's parameters. Moreover, we present a novel approach for small-cell association where the range of the small cells is expanded following the cell biasing approach as proposed in 3GPP [15]. However, our bias value is specific to each small cell and is set by taking into account coverage correlation considerations that can be extracted from our mathematical framework.

The remainder of this paper is organized as follows. In Section II, we present the two-tier network model based on stochastic geometry utilized for this paper. In Section III, we develop the estimation of the two-tier SINR fields using random field theory. Section IV shows the derivation of the analytical cross-correlation function, whereas Section V presents the novel correlation-aware cell-specific bias solution. Both numerical and simulation results are shown in Section VI. Finally, the conclusions and future work plans are presented in Section VII.

II. Two-Tier Network System Model

We consider a two-tier cellular network consisting of a single macrocell BS and multiple small-cell BSs deployed within the coverage area of the macrocell, as shown in Fig. 1. The small cells share the properties of cochannel deployment, open subscriber group operation, and lack of coordination. This means that small cells share the same channels both with the macrocell and among themselves, any user assigned by the network can camp on any small cell, and no careful frequency planning has been carried out by the operator to minimize interference. Their coverage area will be determined by their transmit power ratio with respect to the macrocell BS, which is assumed the same for all small-cell BSs present in the network.

Fig. 1 also shows the different types of downlink interference present in this cochannel deployment, which are represented as dotted red lines. Cross-tier interference is shown in two ways:

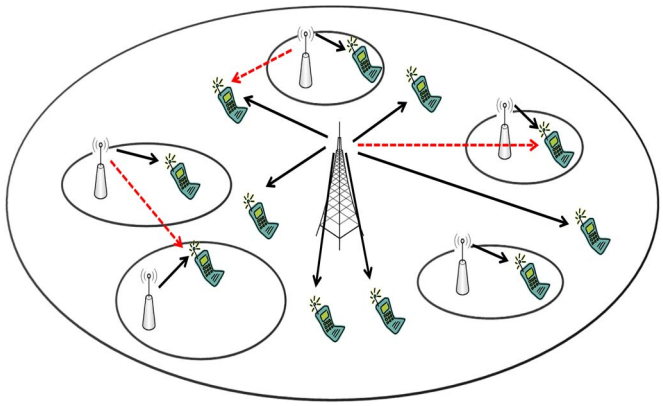

Fig. 1. Two-tier considered topology.

Macrocell users (MUs) receive interference from small-cell BSs, and small-cell users (SCUs) receive interference coming from the macrocell BS. In addition, cotier interference is also modeled as the interference caused by a small-cell BS to an SCU of an adjacent small cell.

Since there is no planning on the locations of the small-cell BSs, these are drawn from a stochastic point process [16]. The model assumes a Poisson point process (PPP) Φ_s for the spatial distribution of the BSs with density ν. Similarly, mobile users are also modeled by an independent PPP, i.e., Φ_u, with different density levels, namely β. The location of the users plays an important role in the performance of the estimation framework presented in Section III, which utilizes the users positions as inputs for the generation of the coverage maps. Further, the employed channel model between the BS and the user accounts for fading and path loss. The fading effect follows a Rayleigh distribution; hence, the channel coefficients are assumed drawn from independent and identically distributed exponential distributions. The path-loss effect is modeled using the standard path-loss function given by $\mathrm{PL} = \|d\|^{-\alpha}$, where $\alpha > 2$ corresponds to the path-loss exponent.

The approach that is followed for the attachment of a user to a BS is the maximum SINR criterion: For each user, a set of SINRs is calculated with respect to the BSs from which signal is being received, and the maximum is selected for attachment. We call $\mathcal{B}$ the set of BSs present in the network serving both macrocell and small-cell areas. Using the former criterion and the described model, the resulting SINR S of user u attached to maximum-SINR BS b can be expressed as follows:

$$S(u,\, b) = \frac{P^b h_u^b \left\| d(u,\, b) \right\|^{-\alpha}}{\sum_{a \in \mathcal{B} \setminus b} P^a h_u^a \left\| d(u,\, a) \right\|^{-\alpha} + \sigma^2} \tag{1}$$

where P^b is the transmitted power from BS b, $h_u^b \sim \exp(1)$ is the channel coefficient from BS b to user u, and σ^2 is the constant additive noise power. Fig. 2 shows one deployment example generated with this model and the attachment results. Users, represented as green dots, are always attached to a BS. In Fig. 2, the user association is explicitly shown if it takes place with a small-cell BS; if associated with the macrocell BS, the user is represented as a point with no attachment for clarity of the figure. Clearly, channel effects and interference contribute to the users not always attaching to the closest BS.

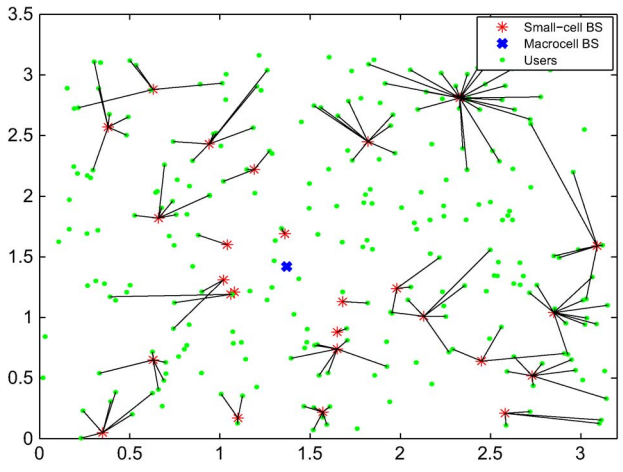

Fig. 2. Sample deployment scenario. Small-cell BSs and users are distributed as PPPs with densities $\nu = 2$ and $\beta = 12\nu$, respectively. Transmit power ratio of macrocell to small-cell BSs is given by $P_{\mathrm{mc}} = 100P_{\mathrm{sc}}$.

III. Two-Tier Coverage Map Estimation Using Random Fields

Spatial statistics constitute a powerful tool utilized in the solution of a broad spectrum of problems in different fields, ranging from sociology or medicine to engineering and technology. A comprehensive literature review of the field can be found in [17] and [18]. The particular field of wireless communications has also made use of different techniques drawn from this field [19], as Section II just showed for the case of point processes. Here, we focus on the particular subfield of random fields, i.e., a set of techniques previously utilized in the wireless network literature for problems, such as the characterization of spectrum in dynamic spectrum access systems [20] or the power spectral density estimation of signal and interference in traditional cellular systems [21]. In most cases, the focus of these studies is centered on modeling and learning about the state of a wireless network at runtime. Our objective is to make use of these techniques to derive suitable analytical SINR coverage expressions for cross-correlation analysis of the two tiers.

Originally developed in the field of geostatistics to describe the properties of soil [22], a random field $Z(\mathbf{x})$ is a stochastic process defined over some metric space or region of interest W. It is characterized by its mean and covariance function shown in the following:

$$\begin{aligned} \mu(\mathbf{x}) &= E\left[Z(\mathbf{x})\right] \\ C(\mathbf{x}, \mathbf{y}) &= E\left[Z(\mathbf{x})Z(\mathbf{y})\right] - \mu(\mathbf{x})\mu(\mathbf{y}). \end{aligned} \tag{2}$$

The application of this framework had the initial objective of reliably estimating soil properties at locations that cannot be sampled using correlation properties of the field. To characterize the spatial correlation of the collected signals in a given random field, the concept of *semivariogram* was introduced, which is calculated as follows:

$$\gamma(\mathbf{x}, \mathbf{y}) = 0.5 \times E\left[\left(Z(\mathbf{x}) - Z(\mathbf{y})\right)^2\right]. \tag{3}$$

The semivariogram will only depend on $\mathbf{h} = \mathbf{x} - \mathbf{y}$ if the random process Z is intrinsically stationary. Further, if Z is also isotropic, the semivariogram γ will only depend on the distance between the two points $\|\mathbf{h}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. Hence, our analysis is confined to planar deployments due to the irregular distribution of users in the vertical domain. Since the definition of semivariogram requires an infinite number of samples, we use the popular *empirical semivariogram* model to perform the estimation, which estimates a semivariogram value for each pair of samples. The following shows the expression, where $N(\mathbf{h})$ represents the pooled number of pairs separated by lag $\mathbf{h}$, which are grouped into lag bins:

$$\widehat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} \left[Z(\mathbf{x}_\alpha + \mathbf{h}) - Z(\mathbf{x}_\alpha)\right]^2. \tag{4}$$

The empirical semivariogram provides only a limited number of points. Hence, to provide reliable subsequent field estimation, we need semivariogram fitting models. Different parametric models exist, all of which are characterized in terms of three key empirical parameters: 1) *sill*, defined as the maximum spatial correlation level for any two points; 2) *range*, defined as the lag between two points at which the semivariogram reaches the sill value; and 3) *nugget*, defined as the semivariogram value for a lag value of zero. The three models considered in this paper are the exponential model $\gamma_{\exp}$, the Gaussian model γ_{gauss}, and the linear model γ_{lin}, given as follows:

$$\gamma_{\exp}(\mathbf{h}) = b + c \times \left(1 - \exp\left(-3\|\mathbf{h}\|/a\right)\right) \tag{5}$$

$$\gamma_{\mathrm{gauss}}(\mathbf{h}) = b + c \times \left(1 - \exp\left(-3\|\mathbf{h}\|^2/a\right)\right) \tag{6}$$

$$\gamma_{lin}(\mathbf{h}) = b + c + a \times \|\mathbf{h}\| \tag{7}$$

where a, b, and c represent the range, the nugget, and the sill.

In wireless systems, different continuous spatial phenomena can be modeled as random fields. Examples are signal strength, interference power, shadow fading, or the coverage metric SINR. In this paper, we shall treat two-tier downlink SINRs, i.e., $S^M(\mathbf{x})$ for the macrocell and $S^F(\mathbf{x})$ for the small-cell tier, as random fields and assume that a collection of sample measurements will be available provided by the scattered mobile users. To obtain the full coverage maps over the region of interest, we need to estimate the values of S^M and S^F at every position $\mathbf{x}_0$ for which a measurement is not directly available.

We select an ordinary Kriging estimator, which is a linear estimator of great simplicity extensively used in geostatistics. Kriging interpolates the values of observations of a random field at unobserved locations. For this purpose, a best linear unbiased estimator (BLUE) is used based on a stochastic model of the spatial dependence quantified by the semivariogram (see [17] for a complete discussion and derivation). To obtain the field value in a generic location given by the vector $\mathbf{x}$, the macrocell estimated field $\widehat{S}^M(\mathbf{x})$ can be calculated as follows:

$$\widehat{S}^M(\mathbf{x}) = \sum_{i=1}^{m} \lambda_i^M(\mathbf{x}) S^M(\mathbf{x}_i) \tag{8}$$

where m represents the total number of available measurements, $S^M(\mathbf{x}_i)$ represents the sample field measurements provided by the MUs, and the weights $\lambda_i^M(\mathbf{x})$ correspond to the Kriging coefficients of the estimator representing how much

amount of the measured field in a particular point will be contained in the estimation. The search for a solution to the BLUE optimization problem yields the following system of equations:

$$\begin{pmatrix} \boldsymbol{\Gamma}^M & \mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}^M \\ \mu^M \end{pmatrix} = \begin{pmatrix} \boldsymbol{\gamma}^M(\mathbf{x}) \\ 1 \end{pmatrix} \tag{9}$$

where $\boldsymbol{\lambda}^M$ is the desired vector of Kriging weights, and $\boldsymbol{\gamma}^M(\mathbf{x})$ is a vector of semivariogram values relating to the position at which one wishes to estimate $(\mathbf{x})$ to the MU measurement locations $(\mathbf{x}_i^M)$. Therefore, $\gamma_i^M(\mathbf{x}) = \gamma(\mathbf{x} - \mathbf{x}_i)$. The parameter μ is a Lagrange multiplier used in the minimization of the estimator variance whose value can be obtained from the given equations. Finally, Γ^M is the matrix of macrocell measurement semivariograms, which is given by the following:

$$\boldsymbol{\Gamma}^M = \begin{pmatrix} \gamma(\mathbf{0}) & \gamma\left(\mathbf{x}_1^M - \mathbf{x}_2^M\right) & \cdots & \gamma\left(\mathbf{x}_1^M - \mathbf{x}_m^M\right) \\ \gamma\left(\mathbf{x}_2^M - \mathbf{x}_1^M\right) & \gamma(\mathbf{0}) & \cdots & \gamma\left(\mathbf{x}_2^M - \mathbf{x}_m^M\right) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma\left(\mathbf{x}_m^M - \mathbf{x}_1^M\right) & \gamma\left(\mathbf{x}_m^M - \mathbf{x}_2^M\right) & \cdots & \gamma(\mathbf{0}) \end{pmatrix}. \tag{10}$$

In the given matrix, vectors $\mathbf{x}_i$ correspond to the positions of the m users providing measurements of the macrocell coverage field. A remark worth noting is the fact that, in general, the nugget value will be zero $(b = 0)$ since this parameter is used to account for the variability at distances smaller than the typical sample spacing, including the measurement error. Thus, for the selected semivariogram fitting models in this paper, the diagonal of Γ^M will consist of zeros.

Back to (9), a simplified expression containing the weight vector is the following one:

$$\boldsymbol{\Gamma}^M \cdot \boldsymbol{\lambda}^M + \mu^M = \boldsymbol{\gamma}^M. \tag{11}$$

Simple manipulation of the given equation yields

$$\boldsymbol{\lambda}^M = (\boldsymbol{\Gamma}^M)^{-1} \cdot \boldsymbol{\gamma}_\mu^M \tag{12}$$

where

$$\boldsymbol{\gamma}_\mu^M(\mathbf{x}) = \begin{pmatrix} \gamma\left(\mathbf{x} - \mathbf{x}_1^M\right) \\ \gamma\left(\mathbf{x} - \mathbf{x}_2^M\right) \\ \vdots \\ \gamma\left(\mathbf{x} - \mathbf{x}_m^M\right) \end{pmatrix} - \mu^M. \tag{13}$$

Equation (12) shows in a compact matrix form an expression for the Kriging weight vectors, which will be obtained according to the selected semivariogram model. Equation (13) contains the semivariogram values of the MUs minus the Lagrange multiplier of the Kriging equations.

Similarly, the derivation of the estimated small-cell tier field $\widehat{S}^F$ with measurements provided by f mobile users will be performed in an analogous manner. The small-cell coverage field $\widehat{S}^F$ is defined as the maximum SINR value that a mobile user could achieve at each point of the region of interest if a small cell were serving it.

We are now in place to give analytical expressions for the coverage fields of the two-tier network, which are given by as follows in compact matrix form:

$$\begin{aligned} \widehat{S}^M(\mathbf{x}) &= \boldsymbol{\lambda}^M(\mathbf{x}) \cdot \mathbf{s}^M \\ &= \left[(\boldsymbol{\Gamma}^M)^{-1} \cdot \boldsymbol{\gamma}_\mu^M(\mathbf{x})\right]^T \cdot \mathbf{s}^M \\ &= \boldsymbol{\gamma}_\mu^M(\mathbf{x})^T \cdot (\boldsymbol{\Gamma}^M)^{-1} \cdot \mathbf{s}^M \end{aligned} \tag{14}$$

$$\begin{aligned} \widehat{S}^F(\mathbf{x}) &= \boldsymbol{\lambda}^F(\mathbf{x}) \cdot \mathbf{s}^F \\ &= \left[(\boldsymbol{\Gamma}^F)^{-1} \cdot \boldsymbol{\gamma}_\mu^F(\mathbf{x})\right]^T \cdot \mathbf{s}^F \\ &= \boldsymbol{\gamma}_\mu^F(\mathbf{x})^T \cdot (\boldsymbol{\Gamma}^F)^{-1} \cdot \mathbf{s}^F. \end{aligned} \tag{15}$$

Here, T represents the transpose operator, and $\mathbf{s}^F$ and $\mathbf{s}^M$ represent the macrocell and small-cell tiers measurement vectors, respectively. The macrocell case is shown as example in

$$\mathbf{s}^M = \begin{pmatrix} S\left(\mathbf{x}_1^M\right) \\ S\left(\mathbf{x}_2^M\right) \\ \vdots \\ S\left(\mathbf{x}_m^M\right) \end{pmatrix}. \tag{16}$$

As a final remark, it is important to acknowledge the dependence of the model on the density of users per cell since the coverage maps are generated based on users' measurements. This paper considers a scenario with a realistic proportion of users to BSs according to the literature and 3GPP, but coverage estimation with this same model for underutilized cells would deserve further study.

IV. New Spatial Cross-Tier Correlation Function

Equations (14) and (15) provide *continuous* field estimations since they allow the calculation of the field values for any location $\mathbf{x}$ belonging to the region of interest W. Therefore, we can define the cross-tier correlation function as follows:

$$R_{\widehat{S}_M, \widehat{S}_F}(\mathbf{x}) = \int_{W_{u_x}} \int_{W_{u_y}} \widehat{S}_M(\mathbf{u}) \widehat{S}_F(\mathbf{u} + \mathbf{x}) d\mathbf{u} \tag{17}$$

where W_{u_x} and W_{u_y} represent the region of interest in the two spatial dimensions over which the network is deployed. By utilizing (14) and (15) and making some change of notation, we obtain the following:

$$\begin{aligned} &\int_{W_{u_x}} \int_{W_{u_y}} \widehat{S}_M(\mathbf{u}) \widehat{S}_F(\mathbf{u} + \mathbf{x}) d\mathbf{u} \\ &\quad = \int_{W_{u_x}} \int_{W_{u_y}} \boldsymbol{\gamma}_\mu^M(\mathbf{u}) \cdot \underbrace{(\boldsymbol{\Gamma}^M)^{-1} \cdot \mathbf{s}^M}_{\mathbf{b}^M} \\ &\qquad \cdot \boldsymbol{\gamma}_\mu^F(\mathbf{u} + \mathbf{x}) \cdot \underbrace{(\boldsymbol{\Gamma}^F)^{-1} \cdot \mathbf{s}^F}_{\mathbf{b}^F} d\mathbf{u} \\ &\quad = \int_{W_{u_x}} \int_{W_{u_y}} \boldsymbol{\gamma}_\mu^M(\mathbf{u}) \cdot \mathbf{b}^M \cdot \boldsymbol{\gamma}_\mu^F(\mathbf{u} + \mathbf{x}) \cdot \mathbf{b}^F d\mathbf{u} \end{aligned} \tag{18}$$

where vectors $\mathbf{b^M}$ and $\mathbf{b^F}$ capturing the information of the semivariogram matrix and measurements are introduced for ease of notation. Further mathematical manipulation yields

$$\begin{aligned}
&\int_{W_{u_x}}\int_{W_{u_y}} \boldsymbol{\gamma}_\mu^M(\mathbf{u})\cdot\mathbf{b}^M\cdot\boldsymbol{\gamma}_\mu^F(\mathbf{u}+\mathbf{x})\cdot\mathbf{b}^F d\mathbf{u}\\
&\quad=\int_{W_{u_x}}\int_{W_{u_y}} \left(\gamma_\mu\left(\mathbf{u}-\mathbf{x}_1^M\right)b_1^M+\cdots+\gamma_\mu\left(\mathbf{u}-\mathbf{x}_m^M\right)b_m^M\right)\\
&\qquad\cdot\left(\gamma_\mu\left(\mathbf{u}-\mathbf{x}_1^F\right)b_1^F+\cdots+\gamma_\mu\left(\mathbf{u}-\mathbf{x}_f^F\right)b_f^F\right)d\mathbf{u}\\
&\quad=\int_{W_{u_x}}\int_{W_{u_y}} \sum_{i=1}^{mf}\gamma_\mu\left(\mathbf{u}-\mathbf{x}_{f(i)}^M\right)b_{f(i)}^M\gamma_\mu\left(\mathbf{u}-\mathbf{x}_{g(i)}^F\right)b_{g(i)}^F d\mathbf{u}
\end{aligned} \tag{19}$$

where the subindices $f(i)$ and $g(i)$ have been introduced for determining either a specific measurement location or an element of a vector. Functions f and g are defined as follows:

$$\begin{aligned}
f(i)&=\left\lfloor\frac{i}{f}\right\rfloor\\
g(i)&=i \bmod f.
\end{aligned} \tag{20}$$

Switching the order of the sum and the integrals allows us to obtain the final cross-tier correlation expression given by

$$R_{\widehat{S}^M,\widehat{S}^F}(\mathbf{x})=\sum_{i=1}^{mf} b_{f(i)}^M b_{g(i)}^F K_i(\mathbf{x}) \tag{21}$$

where the correlation terms $K_i(\mathbf{x})$ are calculated according to the following:

$$K_i(\mathbf{x})=\int_{W_{u_x}}\int_{W_{u_y}} \gamma_\mu^M\left(\mathbf{u}-\mathbf{x}_{f(i)}^M\right)\gamma_\mu^F\left(\mathbf{u}+\mathbf{x}-\mathbf{x}_{g(i)}^F\right)d\mathbf{u}. \tag{22}$$

V. Correlation-Aware Cell-Specific Bias for Range Expansion

Here, we present one application of how two-tier coverage correlation can be used to solve the problem of cell association for HetNets. As explained in Section I, a bias factor is utilized to expand the range of small cells. This factor, which is applied at the receiver of the users, is a positive offset added to the actual received power from the small-cell BS that makes the user believe the received power is higher so that it performs the handover at an earlier point. As a consequence, it allows more users to be associated with the small cell. Several problems arise in this scenario. First, interference is incurred in the users as they will be handing over to the small cell when they are still receiving strong signal from the macrocell. Second, setting a universal bias factor for all the small cells in the network is a very rigid approach that may undermine performance since not all the coverage regions of the different small cells have the same features, even when their transmit power is the same. This will be shown in Section VI. For this reason, we introduce a correlation-aware cell-specific bias solution where the cross-tier correlation of the fields is exploited to provide a suitable cell bias value for each small cell.

TABLE I
Network Simulation Parameters

Parameter	Network 1	Network 2
Macrocell area (km)	5	10
Number of macrocell BSs	1	1
Small-cell BSs density (λ)	1	2
Mobile user density (β)	10	24
Power ratio (S/F)	100	100
Pathloss exponent	3	3
Channel exponential mean	-	1
BS attachment criterion	max-SINR	max-SINR

The approach can be summarized as follows: *The cell-specific bias will approximately keep its universal value when the respective increase and decrease in coverage in each of the two network tiers moving away from the small cell are correlated, i.e., when one is the cause of the other. Otherwise, its value will be modified in such a way that steep falls in the coverage of a small cell when moving away from its BS without a correlated increasing field of the surrounding macrocell will decrease the value of the bias factor. Similarly, steep falls in the macrocell coverage when approaching a small-cell BS that does not present a correlated small cell field will increase the bias value.* By doing this, the bias will adapt to the dual behavior of both fields: It will increase its value when there is gain in further extending small cell coverage, and it will decrease when the universal bias value is already causing losses. This adaptation can be also very useful to optimize the cell association when the two strongest power levels come from two small cells, but that problem is now left as future work.

To implement the given scheme, the coverage maps need to be generated to obtain relevant information regarding the fields. Then, the cell-specific bias value should be calculated. A simple implementation of the given principle for the derivation of the cell-specific bias is

$$B_i=B_{\text{un}}\cdot\frac{\delta}{\frac{|\widehat{m}_{F_i}|}{|\widehat{m}_M|}}=B_{\text{un}}\cdot\frac{\delta|\widehat{m}_M|}{|\widehat{m}_{F_i}|} \tag{23}$$

where B_i is the cell bias of small cell i, B_{un} is the universal bias value, and both m_{F_i} and $\widehat{m}_M$ are estimations of the steepness of the fields around small cell i. These values can be estimated in different ways. In Section VI-D, we show one particularly simple method using the coverage maps derived in Section III. Finally, δ is a system parameter that adjusts the ratio to make it close to one for the highest correlated case where the universal bias value is a wise design choice.

It is worthwhile to include a remark on the low implementation costs of this approach. Small-cell BSs just need to collect SINR values of their users, which they do already, whereas some signaling exchange between the two tiers would be carried out regarding the macrocell coverage map around the small-cell area. This communication could take place over a wired or over-the-air backhaul connection. However, other implementations of correlation-aware cell-specific bias can also be easily designed.

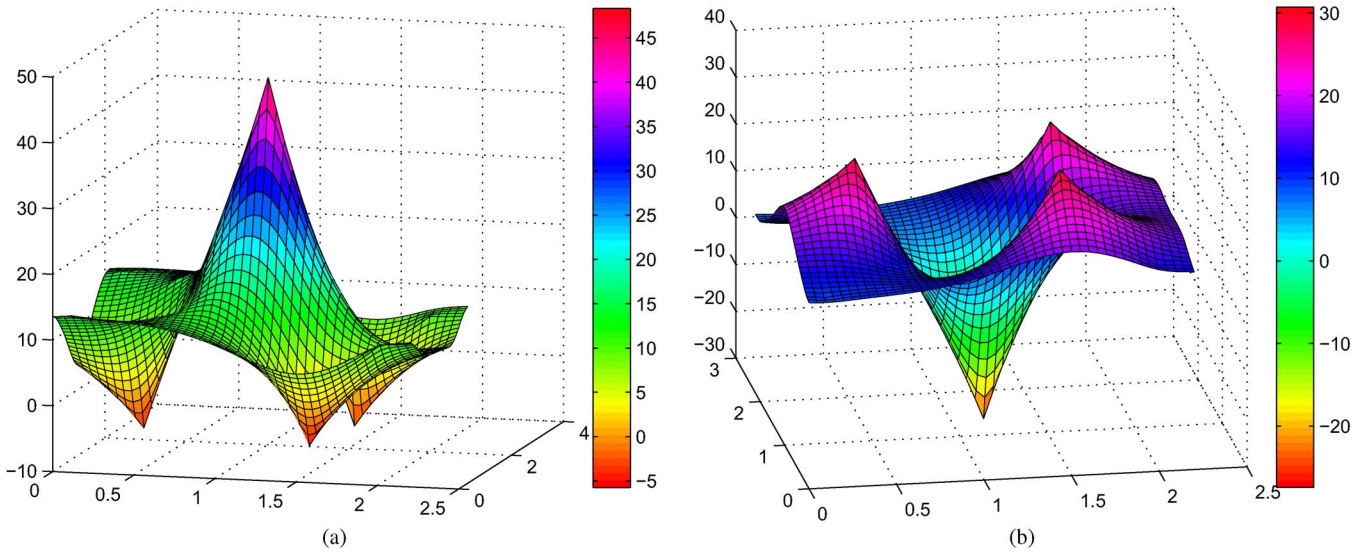

Fig. 3. Two-tier coverage field estimation with exponential semivariogram model for network 1. (a) Macrocell SINR field. (b) Small-cell SINR field.

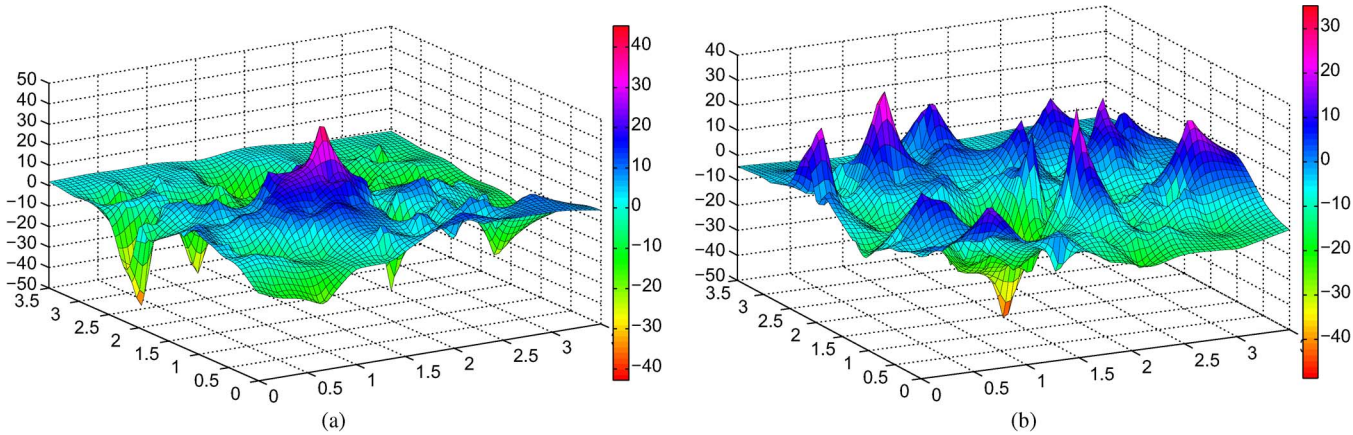

Fig. 4. Two-tier coverage field estimation with exponential semivariogram model for network 2. (a) Macrocell SINR field. (b) Small-cell SINR field.

VI. Performance Evaluation

Here, we show numerical results and performance evaluation with the target of illustrating the correlation analysis of Sections IV and V. The objective is to validate the analytical equations for a two-tier network consisting of a macrocell underlaid by a tier of open-access small cells that use the same transmit power. Furthermore, the performance of a cell-specific bias solution will be evaluated. The employed methodology can be summarized as follows. Following the system model described in Section II, we simulate two networks with the parameters given in Table I. Network 1 is a simplified case with a reduced number of small-cell BS where the channel effects are limited to the path loss due to distance. Network 2 is a more realistic case containing numerous BSs and the channel model presented in Section II. The simulated scenario is used to obtain the SINR values of the mobile users in the network and their BS attachment. With those values at hand, we show and analyze in Section VI-A the complete two-tier SINR field estimations obtained according to (14) and (15). In Section VI-B, we perform a cross-validation of the three semivariogram models presented in Section III to determine which of them performs the best estimation. Section VI-C shows the novel cross-tier correlation functions of networks 1 and 2 obtained by evaluating (21) and (22). Finally, Section VI-D evaluates the performance of our correlation-aware solution for cell association.

A. Coverage Field Generation

The estimations of the macrocell and small-cell coverage fields are shown in Figs. 3 and 4 for networks 1 and 2, respectively. Both of them have used the exponential semivariogram model for their generation. In the case of network 1, the structure of the SINR fields is very *clean* since the only present channel effect is the path loss; hence, no randomness affects the estimation. As expected, the macrocell SINR field exhibits a maximum at the location of the macrocell BS. The quality of the field goes down with the distance to the BS due to the path-loss effect and the interference coming from the three small cells present in the network. The minima of the SINR field are precisely met at the location of these small-cell BSs. Reciprocally, the small-cell tier exhibits also an expected structure, with maxima located at the small-cell BS locations and a minimum at the location of the macrocell BS. Thus, a clear cross correlation exists between the two fields.

The case of network 2 is shown in Fig. 4. Here, we observe fields of a less smooth structure. The randomness of the Rayleigh fading channel and the more complex nature of the

network are responsible for it. However, since our final objective is to determine and characterize a possible cross correlation between the fields, we do observe reciprocal behavior in the two-tier field despite the channel. Maxima and minima are still located at the positions of the BSs, whereas the type of the BS depends on whether it corresponds to the macrocell tier or to the small-cell tier. Thus, a characterization of the cross correlation is still meaningful, as will be shown in Section VI-C.

B. Semivariogram Model Cross-Validation

We now explore the different semivariogram fitting models introduced in Section III. The three cases that will be analyzed correspond to the exponential, Gaussian, and linear models employed for the best possible fitting of user measurements to an analytical semivariogram function. The method that we follow is a cross-validation of the estimation results for the case of network 2, where a larger pool of measurements is available. An estimation is performed at each location for which a measurement is available, and the error of the estimation is obtained by subtracting the estimated value from the real value. This method is repeated for the three models. Furthermore, two error metrics are shown: the complete cross-validation error distribution, where all the estimation errors for both fields are captured, and the mean square error (MSE) given by

$$E = \frac{\sum_i^N \left(\widehat{Z}(i) - Z(i)\right)^2}{N} \tag{24}$$

where N is the total number of measurements.

The results of the cross-validation experiment are shown in Fig. 5. In all cases, the distribution of the cross-validation error has a shape resembling a Gaussian distribution centered around an error of zero. Furthermore, all three models exhibit an acceptable estimation MSE, with the exponential model being the one that more accurately estimates the coverage values. Therefore, the exponential semivariogram model is the most suitable for reconstructing the coverage fields in a two-tier cellular network.

C. Cross-Correlation Function

This section shows the numerical results obtained by implementing the analytical expression for the cross-correlation function obtained in Section IV. Again, the cases of both networks 1 and 2 are illustrated. A 2-D grid is generated for the calculation of the cross-correlation function. Additionally, the raw results obtained with (21) for every grid position are informative only when compared with each other in a relative fashion. Therefore, it makes sense to further normalize them to obtain a metric similar to a cross-correlation coefficient with values between -1 and $+1$. However, as opposed to a cross-correlation coefficient, a value of $+1$ would not mean that the fields are identical but that their correlation would be only maximum. The normalizing operation is given by

$$\mathbf{R}_{\text{norm}} = 2 \times \frac{\mathbf{R} - R_{\min}}{R_{\max} - R_{\min}} - 1 \tag{25}$$

where $R_{\min}$ and $R_{R_{\max}}$ correspond to the minimum and maximum values of the analytical cross-correlation function.

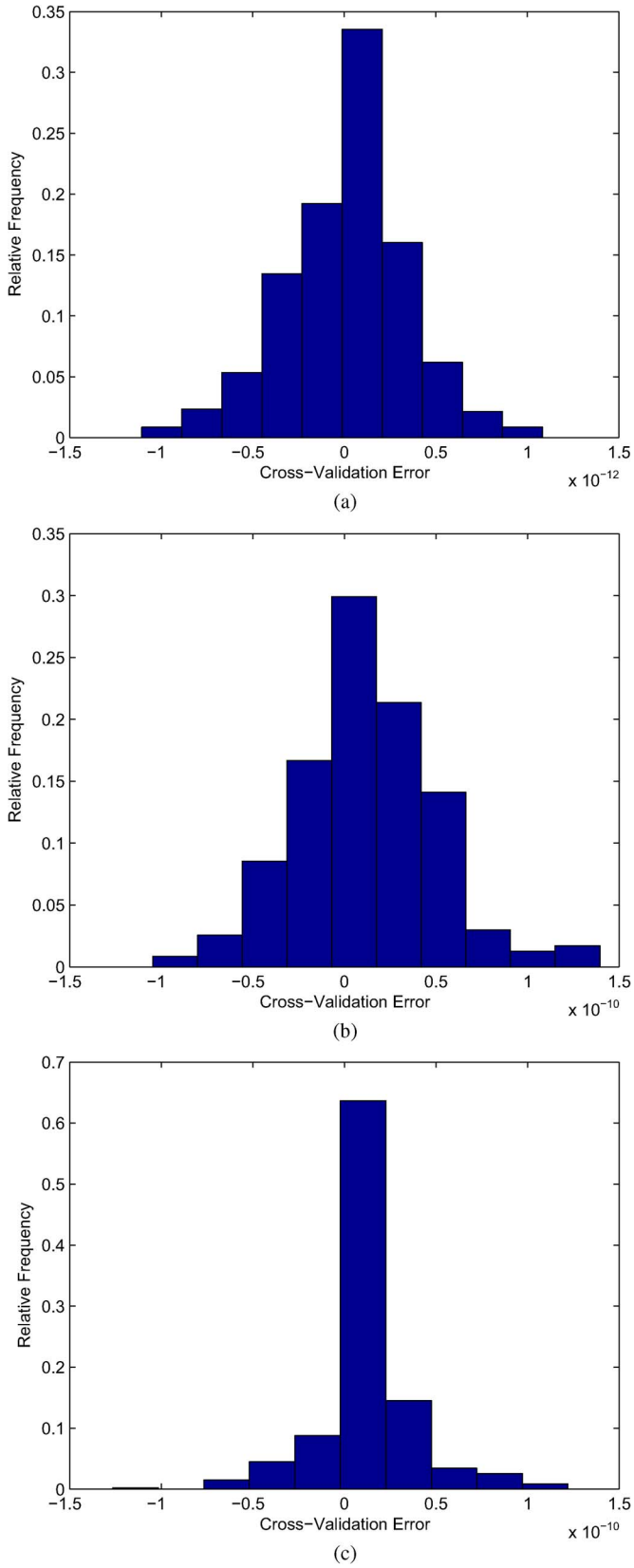

Fig. 5. Cross-validation error distribution of the two-tier field estimation. (a) Exponential model with $E = 1.23 \cdot 10^{-25}$. (b) Gaussian model with $E = 1.48 \cdot 10^{-21}$. (c) Linear model with $E = 6.49 \cdot 10^{-22}$.

Fig. 6 shows the cross-correlation results. The cross-tier correlation function of both networks exhibits a similar structure. A minimum is obtained exactly at the position (0, 0), which is a

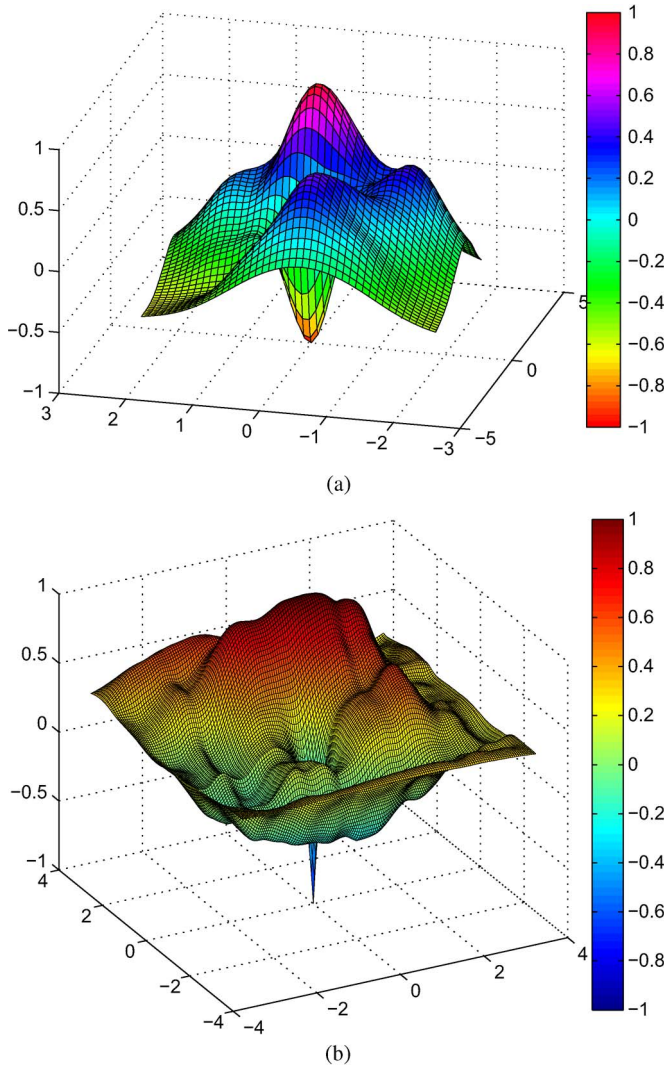

Fig. 6. Analytical cross-tier correlation results. (a) Network 1. (b) Network 2.

meaningful remark since it is precisely when the two unshifted fields overlap, causing minima and maxima to coincide in space and largely contribute to the cross-correlation value. Thus, it is in this location that the integral in (17) should yield a result of the highest absolute value. However, the two-tier fields can be considered *phase shifted* by 180° since maxima in one field are mapped to minima in the reciprocal field. Hence, the value is -1.

The cross-correlation function provides also some information regarding the location of the small cells within the network. No phase shift exists in this case between the two fields since the maximum corresponding to the macrocell BS in the macrocell field is matched with the maxima corresponding to the small-cell BS in the small-cell tier. In the case of network 1, the function is smoother due to the simplified architecture of the network, and three maxima can be detected, corresponding to the three small cells existing in the network. The complexity of network 2 does not allow differentiation of the locations of the small cells, but it does convey an idea of the density of small-cell BSs contained in a certain area.

Beyond the intuitive meaningfulness of the results, we further validate the cross-correlation model by obtaining a normalized spatial cross-correlation coefficient of the two fields with the built-in function of MATLAB *normxcorr2*. We compare these results with the analytical ones shown earlier. The graphs are omitted from this paper due to lack of space, but MSEs on the order of 10^{-8} and 10^{-6} for networks 1 and 2, respectively, suggest an analytical model accurate enough for the representation of the cross-tier correlation function.

These results validate the intuitive idea of coverage correlation between tiers due to interference. Hence, understanding the level of this correlation in the spatial dimension is very useful knowledge that can be applied in the design of practical solution for tiered networks, as it will be shown in Section VI-D.

D. Cell-Specific Bias

To evaluate the performance of the cell-specific bias approach, we simulate user rates and per-cell throughput values for a network without bias, with several universal bias values and with cell-specific bias. For the latter case, the steepness of the fields $\widehat{m}_M$ and $\widehat{m}_{F_i}$ is estimated by averaging the spatial slope of the coverage maps in the four main spatial directions (north, east, south, and west) when going away from the small-cell BS. For this experiment, the average is performed over the coverage radius of a bias-free small cell. Furthermore, the experiment dismisses users with too low rates that are considered in outage. The boundary is given by the 80% of the lowest rate in a bias-free scenario.

Fig. 7 shows the cumulative density functions of user rates and per-cell throughput for the case of network 2. The considered universal bias values are 10%, 15%, and 20% of the average received power across the network, and the reference B_{un} value for the cell-specific case is 15%. When the rates of all users are shown, the bias-free case shows the best results because the few users in the small cells will experience a high SINR. The larger the bias, the worse the geometry of the small cell. Cell-specific bias performs better than its reference value and all other larger bias values. However, when the cell throughput is considered, i.e., all the user rates per small cell are aggregated, the cell-specific bias solution performs best. These results show that, depending on the characteristics of the coverage map, a universal bias value may be too large for some small cells and too small for some others. In the former case, the geometry of the cell is hurt, i.e., users attached to it suffer from low SINR and subsequent low data rates. In the latter case, there is a loss of capacity since more users could be supported. Thus, both cases show a degradation of their throughputs when the bias values are not correctly chosen. On the other hand, a cell-specific bias is capable of providing each cell with an appropriate tradeoff value that allows it to increase the number of supported users and the overall throughput.

VII. Conclusion

In this paper, we have addressed the fundamental issue of the spatial coverage cross-tier correlation between the macrocell and the small-cell tiers in the presence of interference by introducing a new mathematical framework and a practical application. First, we derive analytical expressions for the

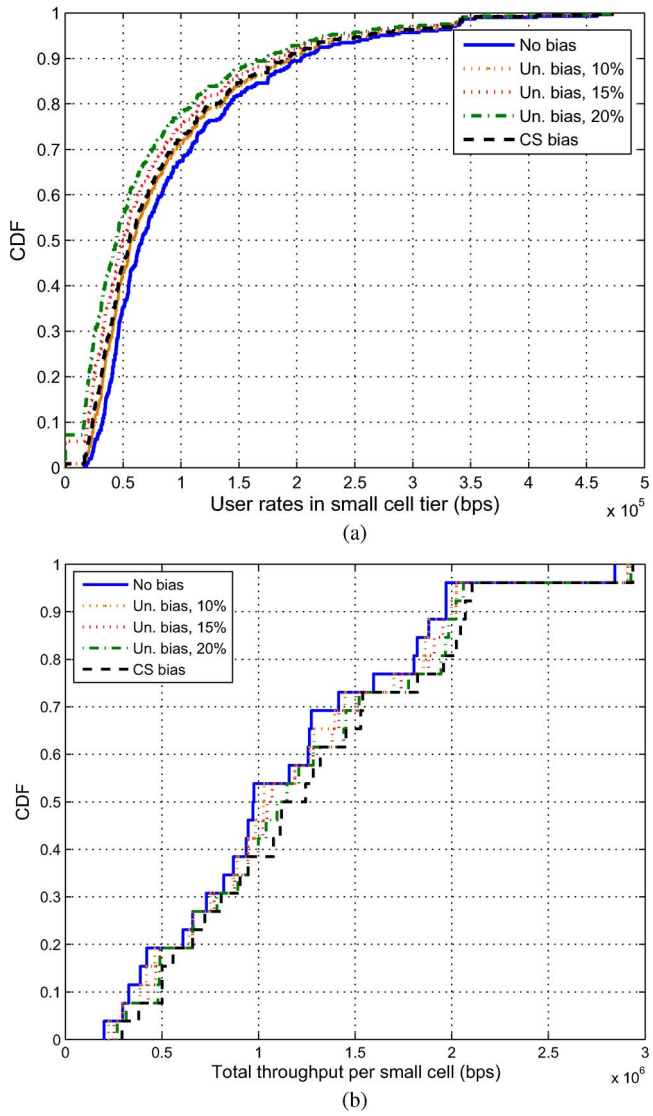

Fig. 7. Cell-specific bias results. (a) Rates. (b) Throughput.

estimated two-tier coverage fields using techniques from the random field theory. We utilize different estimation models and cross-validate them for an assessment of their suitability for the task at hand. Stemming from these expressions, we derive a novel closed-form expression for the cross-tier correlation function exclusively dependent on the estimator's parameters. Then, we propose a novel and simple cell-association scheme for small cells based on cell-specific bias. Results show the suitability of the model for the estimation of the fields and their correlation, as well as the potential of using cross-correlation to efficiently solve practical problems of HetNets. Moreover, this paper opens the path for further research on cross-tier coverage correlation and its applications.

REFERENCES

[1] I. F. Akyildiz, D. M. Gutierrez-Estevez, R. Balakrishnan, and E. Chavarria-Reyes, "LTE-advanced and the evolution to Beyond 4G(B4G) systems," *Phys. Commun.*, vol. 10, pp. 31–60, Mar. 2014.

[2] I. F. Akyildiz, D. M. Gutierrez-Estevez, and E. Chavarria-Reyes, "The evolution to 4G cellular systems: LTE-advanced," *Phys. Commun.*, vol. 3, no. 4, pp. 217–244, Dec. 2010.

[3] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[4] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. A. Thomas, J. G. Andrews, P. Xia, H. Jo, H. S. Dhillon, and T. D. Novlan, "Heterogeneous cellular networks: From theory to practice," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 54–64, Jun. 2012.

[5] "3GPP TS 25.467: UTRAN architecture for 3G Home Node B(HNB); Stage 2," Sophia-Antipolis, France, Tech. Rep., Dec. 2013.

[6] "WMF-T33-118-R016v02: Network architecture femtocell core specification," Clackamas, OR, USA, Tech. Rep., Nov. 2010.

[7] C. Mun, J. Moon, and J. G. Yook, "Self-optimized coverage coordination in femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 2977–2982, Oct. 2010.

[8] V. Chandrasekhar, M. Kountouris, and J. G. Andrews, "Coverage in multi-antenna two-tier networks," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 10, pp. 5314–5327, Oct. 2009.

[9] H.-S. Jo, C. Mun, J. Moon, and J.-G. Yook, "Interference mitigation using uplink power control for two-tier femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4906–4910, Oct. 2009.

[10] C. Young-June, K. C. Seung, and B. Saewoong, "Flexible design of frequency reuse factor in OFDMA cellular networks," in *Proc. IEEE ICC*, Jun. 2006, pp. 1784–1788.

[11] I. F. Akyildiz, E. Chavarria-Reyes, D. M. Gutierrez-Estevez, R. Balakrishnan, and J. R. Krier, "Enabling next generation small cells through femtorelays," *Phys. Commun.*, vol. 9, pp. 1–15, Dec. 2013.

[12] D. López-Pérez, X. Chu, and I. Guvenc, "On the expanded region of picocells in heterogeneous networks," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 3, pp. 281–294, Jun. 2012.

[13] H.-S. Jo, Y. J. Sang, P. Xia, and J. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[14] D. M. Gutierrez-Estevez, B. Canberk, and I. F. Akyildiz, "Spatio-temporal estimation for interference management in femtocell networks," in *Proc. IEEE Int. Symp. PIMRC*, Sep. 2012, pp. 1137–1142.

[15] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun. Mag.*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[16] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

[17] N. Cressie, *Statistics for Spatial Data*. Hoboken, NJ, USA: Wiley, 1993.

[18] B. D. Ripley, *Spatial Statistics*. Hoboken, NJ, USA: Wiley, 2004.

[19] J. Riihijarvi and P. Mahonen, "Spatial statistics for wireless networks research," *Proc. Environ. Sci.*, vol. 7, pp. 86–91, 2011.

[20] J. Riihijarvi, P. Mahonen, M. Wellens, and M. Gordziel, "Characterization and modelling of spectrum for dynamic spectrum access with spatial statistics and random fields," in *Proc. IEEE Int. Symp. Pers., Indoor, Mobile Radio Commun.*, Sep. 2008, pp. 1–6.

[21] J. Riihijarvi and P. Mahonen, "Estimating wireless network properties with spatial statistics and models," in *Proc. Int. Symp. Model. Optim. Mobile, Ad Hoc Wireless Netw. WiOpt*, May 2012, pp. 331–336.

[22] G. Bohling, Introduction to geostatistics and variogram analysis, 2005. [Online]. Available: http://gismyanmar.org/geofocus/wp-content/uploads/2013/01/Variograms.pdf

David M. Gutierrez-Estevez (S'13) received the Engineering degree in telecommunications from Universidad de Granada, Granada, Spain, in 2009 and the M.S. degree in electrical and computer engineering from Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2011, with a fellowship from Fundación la Caixa. He is currently working toward the Ph.D. degree with the Broadband Wireless Networking Laboratory, Georgia Tech, under the supervision of Prof. I. F. Akyildiz on the topic of heterogeneous networks for future cellular systems with a fellowship from Fundación Caja Madrid (2011–2013).

In 2007, he was an Intern with Fraunhofer Institute for Integrated Circuits, Erlangen, Germany. From April 2008 to June 2009, he was a Research Assistant with Fraunhofer Heinrich Hertz Institute, Berlin, Germany. In 2013, he was an Intern with Qualcomm, San Diego, CA, USA.

Ian F. Akyildiz (F'96) received the B.S., M.S., and Ph.D. degrees in computer engineering from the University of Erlangen-Nürnberg, Erlangen, Germany, in 1978, 1981, and 1984, respectively.

He is currently the Director of the Broadband Wireless Networking Laboratory, the Chair of the Telecommunication Group, and the Ken Byers Chair Professor in telecommunications with the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA. Since 2011, he has been a Consulting Chair Professor with the Department of Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. Since January 2013, he has been a Finland Distinguished Professor (supported by the Academy of Finland) with the Department of Communications Engineering, Tampere University of Technology, Tampere, Finland. He is also an Honorary Professor with the School of Electrical Engineering, Technical University of Catalonia (UPC), Barcelona, Spain, and with the Department of Electrical, Electronic, and Computer Engineering, University of Pretoria, Pretoria, South Africa. He is the Founder of the NaNoNetworking Center in Catalonia (N3Cat) and the Advanced Sensor Networks Laboratory. His research interests include Long-Term Evolution, advanced cellular systems, and nanoscale communication networks.

Dr. Akyildiz is a Fellow of the Association for Computing Machinery (ACM). He is currently the Editor-in-Chief of *Computer Networks* (Elsevier). He was the founding Editor-in-Chief of *Ad Hoc Networks* (Elsevier), *Physical Communication* (Elsevier), and *Nano Communication Networks* (Elsevier). He received numerous awards from the IEEE and the ACM.

Etimad A. Fadel received the B.S. degree in computer science from King Abdul Aziz University (KAU), Jeddah, Saudi Arabia, in 1994 and the M.phil/Ph.D. degree in computer science from De Montfort University, Leicester, U.K., in 2007.

She is currently an Assistant Professor with the Department of Computer Science, KAU. Her research interests include wireless networks, Internet of Things, Internet of Nano-Things, sensor networks, and Long-Term Evolution-Advanced cellular systems.