Maximum Likelihood Estimation of Statistical Properties of Composite Gamma-Lognormal Fading Channels

Aleksandar Dogandžić and Jinghua Jin

Abstract—We propose maximum likelihood (ML) methods for estimating the parameters of composite gamma-lognormal fading channels. Newton–Raphson and expectation–maximization (EM) algorithms are developed to compute the ML estimates of the mean and variance of the shadowing component, and the Nakagami-m parameter of the fading component. We also derive Cramér–Rao bounds (CRBs) for the unknown parameters. Numerical simulations demonstrate the performance of the proposed method.

Index Terms—Composite gamma-lognormal fading channels, EM algorithm, Newton–Raphson iteration.

I. INTRODUCTION

Composite fading-shadowing models are used to describe the statistical properties of wireless communication channels in congested downtown areas [1]–[6], satellite communication systems [7]–[10], and distributed antenna systems [11], [12]. In this paper (see also [13]), we present maximum likelihood (ML) algorithms for estimating the parameters of the *composite gamma-lognormal* model in [1]–[4]. This model is fairly general and includes as special cases the Rayleigh-lognormal [6]–[8] and classical Nakagami-m fading and lognormal shadowing scenarios; see, e.g., [1] and [2]. Once obtained, the parameter estimates can be used to design and analyze the performance of wireless communication systems [1]–[14] and to compute minimum mean-square error (MMSE) estimates of mean-signal (shadow) powers; [1] see [15].

In Section II, we introduce the measurement model. The Newton–Raphson and expectation–maximization (EM) algorithms for ML estimation are presented in Sections II-A and B, the initialization of the proposed algorithms is discussed in Section II-C, and the Cramér–Rao bound (CRB) matrix for the unknown parameters is derived in Section II-D. In Section III, numerical examples are used to evaluate the accuracy of the proposed estimators. Concluding remarks are given in Section IV.

II. MEASUREMENT MODEL AND ML ESTIMATION

We present a composite gamma-lognormal fading model and ML methods for estimating the unknown fading and shadowing parameters. Assume that N *instantaneous signal powers* $y_k(t), t = 1, 2, \ldots, N$ have been collected in the kth observation interval, where $k = 1, 2, \ldots, K$, and define

$$\boldsymbol{y}_k = [y_k(1), y_k(2), \ldots, y_k(N)]^T \qquad (1)$$

where "T" denotes a transpose. [If the samples $y_k(t)$ are scaled by the noise power, they can be viewed as instantaneous signal-to-noise ratios

[1]Accurate estimation of the mean-signal powers is required to implement adaptive modulation techniques and algorithms for handoff, channel access, and power control; see, e.g., [1], [16], and [17].

(SNRs).] We model $y_k(t), t = 1, 2, \ldots, N$ as conditionally independent gamma random variables with the following probability density functions (pdfs):[2]

$$p_{y\,|\,u}(y_k(t)\,|\,u_k; m) = \frac{m^m y_k(t)^{m-1}}{u_k^m \Gamma(m)} \exp\left[-\frac{m y_k(t)}{u_k}\right] \qquad (2)$$

where u_k is the *mean-signal (shadow) power* in the kth interval, $\Gamma(\cdot)$ denotes the gamma function, and m is the Nakagami-m fading parameter. We then model the mean-signal powers as independent, identically distributed (i.i.d.) random variables with lognormal pdf

$$p_u(u_k; \mu, \sigma^2) = \frac{\xi}{u_k \sqrt{2\pi\sigma^2}} \exp\left[-\frac{(10\log_{10} u_k - \mu)^2}{2\sigma^2}\right]. \qquad (3)$$

In other words, u_k are assumed to be constant within an observation interval but vary randomly from one interval to another. [The assumption that the mean-signal powers u_k are independent is valid if the observation intervals are sufficiently separated in time.] Here, μ (in decibels) and σ (also in decibels) are the mean and standard deviation of $10\log_{10} u_k$, which are also known as the *area mean* and *shadow standard deviation*, respectively (see, e.g., [1]), and

$$\xi = \frac{10}{\ln 10}. \qquad (4)$$

Our goal is to find the ML estimates of m, μ, and σ^2 using the instantaneous power observations $y_k(t), t = 1, 2 \ldots, N, k = 1, 2, \ldots, K$. Define the unknown parameter vector $\boldsymbol{\theta} = [m, \mu, \sigma^2]^T$. The marginal distribution of $\boldsymbol{y}_k$ follows from (2) and (3):

$$p_{\boldsymbol{y}}(\boldsymbol{y}_k; \boldsymbol{\theta}) = \int_0^\infty \left[\prod_{t=1}^N p_{y\,|\,u}(y_k(t)\,|\,u; m)\right] \cdot p_u(u; \mu, \sigma^2)\, du \qquad (5)$$

for $k = 1, 2 \ldots, K$. The ML estimate of $\boldsymbol{\theta}$ is obtained by maximizing the log-likelihood function of all the measurements $\boldsymbol{y} = [\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \ldots, \boldsymbol{y}_K^T]^T$:

$$L(\boldsymbol{y}; \boldsymbol{\theta}) = \sum_{k=1}^K \ln p_{\boldsymbol{y}}(\boldsymbol{y}_k; \boldsymbol{\theta}). \qquad (6)$$

As observed in [5], the difficulty in estimating the parameters of the composite fading-shadowing models arises due to the integral form of the density function (5). In the following, we present Newton–Raphson and EM algorithms for finding the ML estimates of $\boldsymbol{\theta}$.

A. Newton–Raphson Method

We derive the Newton–Raphson algorithm for maximizing (6). [A quasi-Newton modification of the Newton–Raphson iteration is discussed in Section II-A1.] First, we apply the change-of-variable transformation

$$x = \frac{10\log_{10} u - \mu}{\sqrt{2\sigma^2}} \qquad (7)$$

to (6):

$$\begin{aligned} L(\boldsymbol{y}; \boldsymbol{\theta}) = &-\frac{K}{2}\ln\pi + (m-1)\sum_{k=1}^K \sum_{t=1}^N \ln y_k(t) \\ &+ KNm\ln m - KN\ln\Gamma(m) \\ &+ \sum_{k=1}^K \ln\left\{\int_{-\infty}^\infty q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2)\, dx\right\} \end{aligned} \qquad (8)$$

[2]Hence, $y_{k_1}(t_1)\,|\,u_{k_1}$ and $y_{k_2}(t_2)\,|\,u_{k_2}$ are independent for $k_1 \neq k_2$ or $t_1 \neq t_2$ or both, where $k_1, k_2 \in \{1, 2, \ldots, K\}$ and $t_1, t_2 \in \{1, 2, \ldots, N\}$.

where

$$q(x, \bar{y}_k, \boldsymbol{\theta}) = \exp\left(-mN \cdot \bar{y}_k \cdot 10^{-(\sqrt{2\sigma^2} \cdot x + \mu)/10}\right)$$
$$\cdot 10^{-mN \cdot (\sqrt{2\sigma^2} \cdot x + \mu)/10} \tag{9a}$$

and

$$\bar{y}_k = \frac{1}{N} \sum_{t=1}^{N} y_k(t) \tag{9b}$$

is the sample-mean estimate of the mean-signal power in the kth observation interval.

The gradient vector $\partial L(\boldsymbol{y}; \boldsymbol{\theta})/\partial \boldsymbol{\theta}$ and Hessian matrix $\partial^2 L(\boldsymbol{y}; \boldsymbol{\theta})/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ can be computed using (10a)–(11b), shown at the bottom of the page. The integral expressions in (8), (10), and (11) are efficiently and accurately evaluated using the Gauss–Hermite quadrature formula:

$$\int_{-\infty}^{\infty} f(x) \exp(-x^2) \, dx \approx \sum_{l=1}^{L} h_{x_l} f(x_l) \tag{12}$$

where $f(x)$ is an arbitrary real function, L is the quadrature order (determining approximation accuracy), $x_l, l = 1, \ldots, L$ are the zeroes of the Lth-order Hermite polynomial, and $h_{x_l}, l = 1, \ldots, L$ are the Gauss–Hermite quadrature weight factors tabulated in e.g., [18]. We have omitted the expressions for the derivatives $\partial q(x, \bar{y}_k, \boldsymbol{\theta})/\partial \theta_i$ and $\partial^2 q(x, \bar{y}_k, \boldsymbol{\theta})/\partial \theta_i \partial \theta_j, i, j \in \{1, 2, 3\}$, which are cumbersome but easy to compute. The (damped) Newton–Raphson algorithm updates the estimates of $\boldsymbol{\theta}$ as follows (see, e.g., [19, (13.25)], [20, Ch. 9.7], [21], [22], and [23, Ch. 9.5]):

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} - \lambda^{(i)} \cdot \left[\frac{\partial^2 L\left(\boldsymbol{y}; \boldsymbol{\theta}^{(i)}\right)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\right]^{-1} \frac{\partial L\left(\boldsymbol{y}; \boldsymbol{\theta}^{(i)}\right)}{\partial \boldsymbol{\theta}} \tag{13}$$

where the *damping factor* $0 < \lambda^{(i)} \le 1$ is chosen (at every step i) to ensure that the log-likelihood function (6) increases and that the parameter estimates remain in the allowable parameter space (i.e., $m, \sigma^2 >$

0). The negative inverse of the Hessian matrix evaluated at the ML estimate $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(\infty)}$

$$-\left[\frac{\partial^2 L(\boldsymbol{y}; \hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\right]^{-1} \tag{14}$$

can be used to estimate the covariance matrix of $\hat{\boldsymbol{\theta}}$ and to construct confidence regions for the unknown parameters; see, e.g., ([22, Ch. 4.1.3]. The Hessian matrix formulas (11) will be also utilized to compute the CRB matrix for the unknown parameters; see Section II-D.

1) BFGS Quasi-Newton Algorithm: To accelerate the Newton–Raphson algorithm, we propose the Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton method that approximates the Hessian matrices $\partial^2 L(\boldsymbol{y}; \boldsymbol{\theta}^{(i)})/\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ in (13) with the following estimates (see [21, (9.2.10)] and [22, (4.3.7)]):

$$H^{(i+1)} = H^{(i)} - \frac{H^{(i)} \boldsymbol{d}^{(i)} (\boldsymbol{d}^{(i)})^T H^{(i)}}{(\boldsymbol{d}^{(i)})^T H^{(i)} \boldsymbol{d}^{(i)}} + \frac{\boldsymbol{g}^{(i)} (\boldsymbol{g}^{(i)})^T}{(\boldsymbol{d}^{(i)})^T \boldsymbol{g}^{(i)}} \tag{15}$$

where

$$\boldsymbol{d}^{(i)} = \boldsymbol{\theta}^{(i+1)} - \boldsymbol{\theta}^{(i)} \tag{16a}$$

$$\boldsymbol{g}^{(i)} = \frac{\partial L\left(\boldsymbol{y}; \boldsymbol{\theta}^{(i+1)}\right)}{\partial \boldsymbol{\theta}} - \frac{\partial L\left(\boldsymbol{y}; \boldsymbol{\theta}^{(i)}\right)}{\partial \boldsymbol{\theta}} \tag{16b}$$

and the initial value $H^{(0)}$ can be obtained by computing the exact Hessian at the initial estimate of the unknown parameter vector $\boldsymbol{\theta}^{(0)}$:

$$H^{(0)} = \frac{\partial^2 L\left(\boldsymbol{y}; \boldsymbol{\theta}^{(0)}\right)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}. \tag{17}$$

The Hessian approximation (15) is also known as the *positive definite secant update*; see [21, Ch. 9.2]. Compared with the Newton–Raphson method, the BFGS quasi-Newton algorithm requires more iterations to converge, but each iteration has lower computational complexity; see also Section III.

$$\frac{\partial L(\boldsymbol{y}; \boldsymbol{\theta})}{\partial m} = KN \ln m + KN - KN \cdot \frac{\Gamma'(m)}{\Gamma(m)} + \sum_{k=1}^{K} \sum_{t=1}^{N} \ln y_k(t) + \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial q(x, \bar{y}_k, \boldsymbol{\theta})/\partial m \cdot \exp(-x^2) \, dx}{\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx} \tag{10a}$$

$$\frac{\partial L(\boldsymbol{y}; \boldsymbol{\theta})}{\partial \mu} = \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial q(x, \bar{y}_k, \boldsymbol{\theta}) / \partial \mu \cdot \exp(-x^2) \, dx}{\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx} \tag{10b}$$

$$\frac{\partial L(\boldsymbol{y}; \boldsymbol{\theta})}{\partial \sigma^2} = \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial q(x, \bar{y}_k, \boldsymbol{\theta})/\partial \sigma^2 \cdot \exp(-x^2) \, dx}{\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx}. \tag{10c}$$

and

$$\frac{\partial^2 L(\boldsymbol{y}; \boldsymbol{\theta})}{\partial m^2}$$

$$= KNm^{-1} - KN \cdot \frac{\Gamma(m)\Gamma''(m) - [\Gamma'(m)]^2}{[\Gamma(m)]^2}$$

$$+ \sum_{k=1}^{K} \frac{\int_{-\infty}^{\infty} \partial^2 q(x, \bar{y}_k, \boldsymbol{\theta})/\partial m^2 \cdot \exp(-x^2) \, dx}{\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx} - \left[\frac{\int_{-\infty}^{\infty} \partial q(x, \bar{y}_k, \boldsymbol{\theta})/\partial m \cdot \exp(-x^2) \, dx}{\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx}\right]^2 \tag{11a}$$

$$\left.\frac{\partial^2 L(\boldsymbol{y}; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right|_{(i,j) \ne (1,1)}$$

$$= \sum_{k=1}^{K} \left\{ \frac{\int_{-\infty}^{\infty} \partial^2 q(x, \bar{y}_k, \boldsymbol{\theta})/\partial \theta_i \partial \theta_j \cdot \exp(-x^2) \, dx}{\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx} \right.$$

$$\left. - \frac{\int_{-\infty}^{\infty} \partial q(x, \bar{y}_k, \boldsymbol{\theta})/\partial \theta_i \cdot \exp(-x^2) \, dx \cdot \int_{-\infty}^{\infty} \partial q(x, \bar{y}_k, \boldsymbol{\theta})/\partial \theta_j \cdot \exp(-x^2) \, dx}{\left[\int_{-\infty}^{\infty} q(x, \bar{y}_k, \boldsymbol{\theta}) \cdot \exp(-x^2) \, dx\right]^2} \right\}. \tag{11b}$$

B. EM Algorithm

We present an EM algorithm (see, e.g., [24], [25], and [26, Ch. 2.4.4]) to compute the ML estimates of $\boldsymbol{\theta}$. Define the vector of the mean-signal (unobserved) powers $\boldsymbol{u} = [u_1, u_2, \ldots, u_K]^T$. By treating $\boldsymbol{u}$ as the *unobserved* (or missing) data, we derive the following iteration between the expectation (E) and maximization (M) steps (see the Appendix):

E Step: Compute

$$\mathcal{T}_1\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) = \frac{1}{K}\sum_{k=1}^{K}\mathrm{E}_{u\,|\,\boldsymbol{y}}\left[\ln u_k\,\Big|\,\boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}\right] \tag{18a}$$

$$\mathcal{T}_2\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) = \frac{1}{K}\sum_{k=1}^{K}\mathrm{E}_{u\,|\,\boldsymbol{y}}\left[(\ln u_k)^2\,\Big|\,\boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}\right] \tag{18b}$$

$$\mathcal{T}_3\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) = \frac{1}{K}\sum_{k=1}^{K}\mathrm{E}_{u\,|\,\boldsymbol{y}}\left[u_k^{-1}\,\Big|\,\boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}\right]\cdot\bar{y}_k$$
$$-\frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t) \tag{18c}$$

where

$$\boldsymbol{\theta}^{(i)} = \left[m^{(i)}, \mu^{(i)}, (\sigma^2)^{(i)}\right]^T \tag{18d}$$

is an estimate of $\boldsymbol{\theta}$ in the ith iteration, and (18a)–(18c) are computed using (19a) and (19b), shown at the bottom of the page, where $t(u_k) = \ln u_k, (\ln u_k)^2$, and u_k^{-1}, for $k = 1, 2, \ldots, K$.

M Step: Compute

$$\mu^{(i+1)} = \xi \cdot \mathcal{T}_1\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) \tag{20a}$$

$$(\sigma^2)^{(i+1)} = \xi^2 \cdot \mathcal{T}_2\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) - \left(\mu^{(i+1)}\right)^2 \tag{20b}$$

and find $m^{(i+1)}$ that maximizes

$$m^{(i+1)} = \arg\max_m \left\{ m\ln m - \ln[\Gamma(m)] \right.$$
$$\left. - m\mathcal{T}_1\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) - m\mathcal{T}_3\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) \right\}. \tag{20c}$$

Upon convergence (i.e., as $i \to \infty$), the above algorithm also provides (estimated) MMSE estimates of the shadow powers in decibels [see (18a)]

$$\mathrm{E}_{u\,|\,\boldsymbol{y}}[10\log_{10} u_k\,|\,\boldsymbol{y}_k;\hat{\boldsymbol{\theta}}] = \xi \cdot \mathrm{E}_{u\,|\,\boldsymbol{y}}[\ln u_k\,|\,\boldsymbol{y}_k;\hat{\boldsymbol{\theta}}] \tag{21}$$

where the unknown parameter vector $\boldsymbol{\theta}$ is replaced with its ML estimate $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(\infty)}$. Note that estimates of the shadow powers in decibels are being utilized by most handoff algorithms, as well as for channel access and power control; see [17].

We now discuss computing the conditional expectation in (19) and maximizing (20c). The approximation (19b) was derived by applying the change-of-variable transformation (7) to the numerator and denominator in (19a) and using the Gauss–Hermite quadrature (12) to numerically evaluate the obtained integrals. Due to the cancellations of the

common terms in the numerator and denominator of (19a), (19b) is remarkably simple. In [15], (19b) was used to compute the MMSE estimates of $t(u_k) = u_k$.

The computation of $m^{(i+1)}$ requires maximizing (20c), which was performed using the Newton–Raphson method (embedded within the "outer" EM iteration) with the initial values chosen as (see [27, Ch. 8.3.6])

$$m_{\mathrm{init}}^{(i+1)} = \frac{3 + 2\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})}{2\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})\cdot\left[3 + \phi\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right)\right]} \tag{22a}$$

where

$$\phi\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) = \mathcal{T}_1\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) + \mathcal{T}_3\left(\boldsymbol{y};\boldsymbol{\theta}^{(i)}\right) - 1. \tag{22b}$$

The Newton–Raphson iteration for maximizing (20c) converges rapidly when initialized with the approximate ML estimate in (22a); see Section III. The derivatives needed to implement this iteration are shown in (A.5) in the Appendix, where $\varphi(\boldsymbol{y},\boldsymbol{u})$ should be replaced with $\phi(\boldsymbol{y};\boldsymbol{\theta}^{(i)})$.

C. Choosing the Initial Values

The proposed algorithms can be initialized by fitting the simple log-normal shadowing model, which leads to the following initial estimates of the shadowing parameters:

$$\mu^{(0)} = \frac{1}{K}\sum_{k=1}^{K}10\log_{10}\bar{y}_k \tag{23a}$$

$$(\sigma^2)^{(0)} = \frac{1}{K}\sum_{k=1}^{K}[(10\log_{10}\bar{y}_k)^2] - \left(\mu^{(0)}\right)^2. \tag{23b}$$

For $N > 1$, an initial estimate of m to start the EM iteration can be obtained using an approximate estimator similar to (22a):

$$m^{(0)} = \frac{3 + 2\varphi_0(\boldsymbol{y})}{2\varphi_0(\boldsymbol{y})\cdot[3 + \varphi_0(\boldsymbol{y})]} \tag{24a}$$

where

$$\varphi_0(\boldsymbol{y}) = \frac{1}{K}\cdot\sum_{k=1}^{K}(\ln\bar{y}_k) - \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t) \tag{24b}$$

is obtained by replacing the unobserved shadow powers u_k with their sample-mean estimates $\bar{y}_k$, $k = 1, 2, \ldots, K$ in the expression for $\varphi(\boldsymbol{y},\boldsymbol{u})$ in (A.6); see the Appendix.

D. Cramér–Rao Bounds

The CRB matrix for the unknown parameter vector $\boldsymbol{\theta}$ can be computed by inverting the expected negative Hessian matrix [see (11)], where the expectation is performed with respect to the distribution of $\boldsymbol{y}$ (see, e.g., [26, Ch. 3.4.2] and [28, Ch. 3.7]):

$$\mathrm{CRB}(\boldsymbol{\theta}) = -\left\{\mathrm{E}_{\boldsymbol{y}}\left[\frac{\partial^2 L(\boldsymbol{y};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right]\right\}^{-1}. \tag{25}$$

$$\mathrm{E}_{u\,|\,\boldsymbol{y}}\left[t(u_k)\,\Big|\,\boldsymbol{y}_k;\boldsymbol{\theta}^{(i)}\right]$$

$$= \frac{\int_0^\infty t(u)\cdot\left[\prod_{t=1}^{N}p_{y\,|\,u}\left(y_k(t)\,\Big|\,u;m^{(i)}\right)\right]\cdot p_u\left(u;\mu^{(i)},(\sigma^2)^{(i)}\right)du}{\int_0^\infty\left[\prod_{t=1}^{N}p_{y\,|\,u}\left(y_k(t)\,|\,u;m^{(i)}\right)\right]\cdot p_u\left(u;\mu^{(i)},(\sigma^2)^{(i)}\right)du} \tag{19a}$$

$$\approx \frac{\sum_{l=1}^{L}h_{x_l}\cdot t\left(10^{(\sqrt{2\sigma^2}\cdot x_l+\mu)/10}\right)\cdot\exp\left(-mN\cdot\bar{y}_k\cdot 10^{-(\sqrt{2\sigma^2}\cdot x_l+\mu)/10}\right)\cdot 10^{-mN\cdot\sqrt{2\sigma^2}\cdot x_l/10}}{\sum_{l=1}^{L}h_{x_l}\exp\left(-mN\cdot\bar{y}_k\cdot 10^{-(\sqrt{2\sigma^2}\cdot x_l+\mu)/10}\right)\cdot 10^{-mN\cdot\sqrt{2\sigma^2}\cdot x_l/10}} \tag{19b}$$
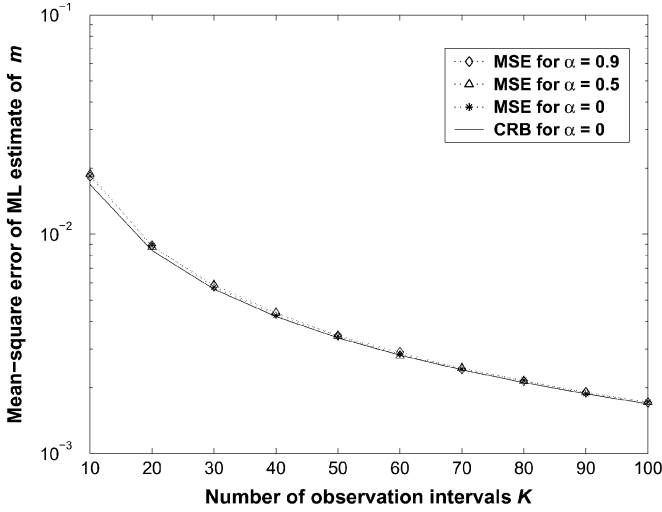
Fig. 1.   Mean-square error and Cramér–Rao bound for the proposed estimator of m as a function of K assuming uncorrelated shadow powers ($\alpha = 0$), correlated shadow powers with the AR coefficient $\alpha = 0.5$, and correlated shadow powers with $\alpha = 0.9$.
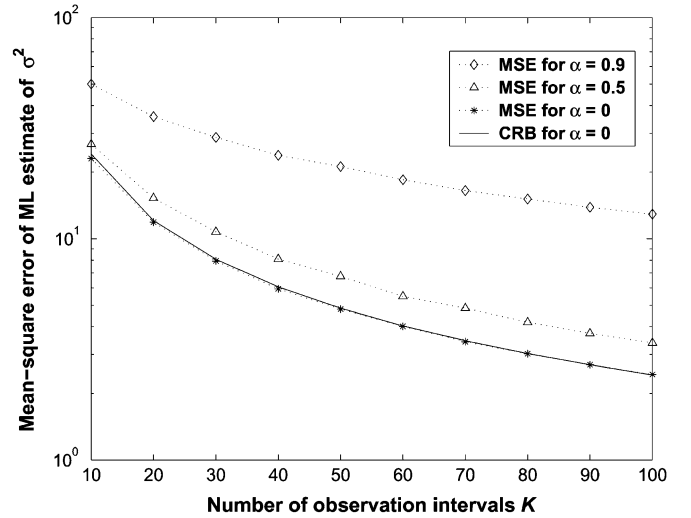

Fig. 2.   Mean-square error and Cramér–Rao bound for the proposed estimator of μ as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$, and correlated shadow powers with $\alpha = 0.9$.

Fig. 3.   Mean-square error and Cramér–Rao bound for the proposed estimator of σ^2 as a function of K assuming uncorrelated shadow powers, correlated shadow powers with $\alpha = 0.5$, and correlated shadow powers with $\alpha = 0.9$.
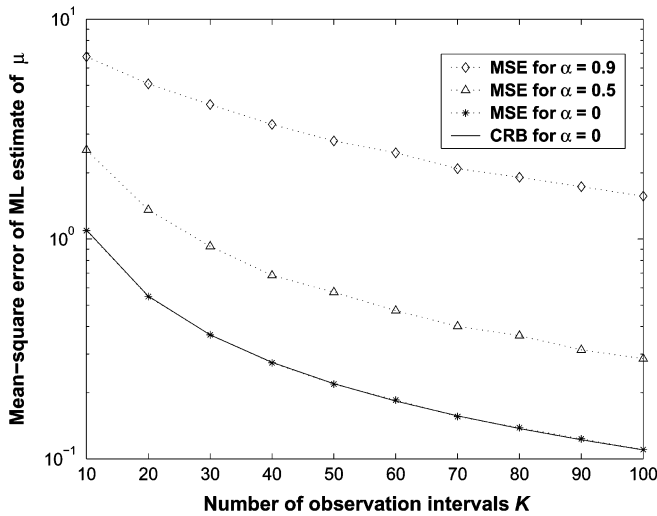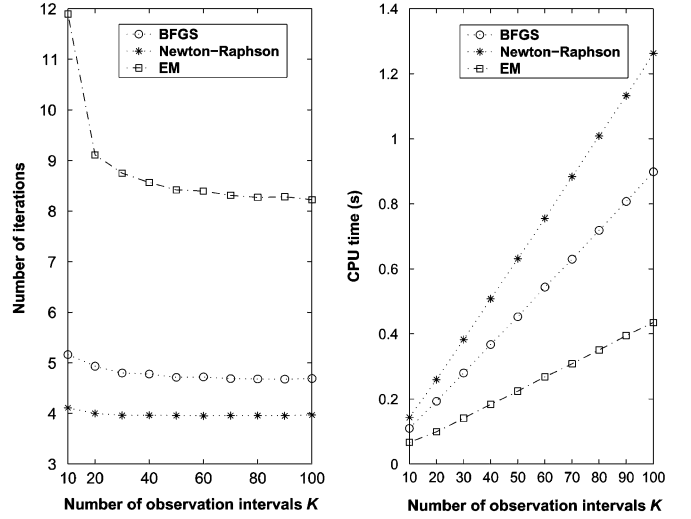

Fig. 4.   (Left) Average number of iterations and (right) CPU time of the EM, Newton–Raphson, and BFGS algorithms as a function of K.

The above expectation requires multidimensional integration, which can be performed using Monte Carlo integration, i.e., by averaging $\partial^2 L(\boldsymbol{y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ over many realizations of $\boldsymbol{y}$.

III. NUMERICAL EXAMPLES

The numerical examples presented here assess the estimation accuracy of the ML estimates of $\boldsymbol{\theta}$. Our performance metric is the mean-square error (MSE) of an estimator, calculated using 60 000 independent trials. [Note that the MSEs of the Newton–Raphson and EM algorithms coincide since the convergence points of both algorithms coincide and are equal to the ML estimate of $\boldsymbol{\theta}$.] The measurements $y_k(t), t = 1, 2, \ldots, N, k = 1, 2, \ldots, K$ were simulated from the composite gamma-lognormal distribution with $N = 10$ samples per observation interval, $10 \leq K \leq 100, m = 1$ (i.e., Rayleigh fading), $\mu = 5$ dB, and $\sigma = 3$ dB. The quadrature order of the Gauss–Hermite approximation in (12) [see also (19b)] was $L = 20$. In Figs. 1–3, we

show the MSEs (and corresponding CRBs[3]) for the ML estimates of m, μ, and σ^2, respectively, as functions of the number of observation intervals K. The ML estimators are "almost efficient" in this scenario, i.e., their MSEs are very close to the corresponding CRBs. In Figs. 1–3, we also show the performance of the proposed algorithms when the shadow powers u_k are correlated. We adopt the first-order autoregressive [AR(1)] correlation model for the shadow process in decibels (see, e.g., [17] and [29]):

$$10 \log_{10} u_k = \alpha \cdot 10 \log_{10} u_{k-1} + \omega_k \qquad (26)$$

where ω_k are i.i.d. Gaussian random variables with mean $(1 - \alpha) \cdot \mu = (1 - \alpha) \cdot 5$ dB and standard deviation $\sqrt{1 - \alpha^2} \cdot \sigma = \sqrt{1 - \alpha^2} \cdot 3$ dB. The MSEs of the proposed estimators are shown for $\alpha = 0.5$ and $\alpha = 0.9$. Interestingly, the MSE performance of the estimator of m is insensitive to the value of the correlation coefficient α; see

---

[3]The CRB matrix was computed using (25), where the expectation with respect to the distribution of $\boldsymbol{y}$ was performed using Monte Carlo integration with 60 000 trials.

Fig. 1. However, the estimation of the shadowing parameters μ and σ^2 is affected by α; see Figs. 2 and 3.

We now evaluate the computational efficiency of the proposed methods. In Fig. 4, we show the numbers of iterations and CPU times (averaged over 2000 trials) of the EM, Newton–Raphson, and BFGS algorithms (implemented using MATLAB), as functions of K. The EM algorithm converged within 12 iteration steps,[4] whereas the Newton–Raphson algorithm converged in four iterations (on average); see the left side of Fig. 4. In terms of CPU time, however, the EM algorithm was faster than the Newton–Raphson method [see the right side of Fig. 4], which can be explained by the fact that a single EM iteration is significantly faster than a Newton–Raphson iteration. In particular, the Newton–Raphson algorithm requires computing and inverting the Hessian matrix, which counterbalances its advantage in speed of convergence. This is a well-known drawback of the Newton–Raphson method (see, e.g., [22, Ch. 4.3.2] and [26, Ch. 2.4.3], which can be surmounted if the derivatives in (10) and (11) are computed in parallel. The BFGS quasi-Newton algorithm converged in five iterations (on average) and outperformed the Newton–Raphson method in terms of CPU time; however, it was slower than the EM algorithm; see the right side of Fig. 4.

IV. CONCLUSION

We derived maximum likelihood methods for estimating the parameters of composite gamma-lognormal fading channels. The ML estimates of the unknown fading and shadowing parameters were computed using Newton–Raphson and EM algorithms. We also applied the BFGS quasi-Newton algorithm, discussed initializing the proposed algorithms, and computed Cramér–Rao bounds for the unknown parameters. The proposed algorithms can be extended to other composite fading-shadowing scenarios, such as Rice-lognormal [9], [10] and multi-input multi-output (MIMO) fading scenarios. Further research will include

- combining quasi-Newton and EM methods (see [30] and references therein);
- developing ML methods for estimating parameters in combined shadowed/unshadowed fading channels for land-mobile satellite scenarios; see, e.g., [2, Ch. 2.2.4] and [31].

APPENDIX
EM ALGORITHM DERIVATION

We derive the EM algorithm presented in Section II-B. Observe that the complete-data log-likelihood can be written as

$$
\begin{aligned}
&L_c(\boldsymbol{y}, \boldsymbol{u}; \boldsymbol{\theta}) \\
&= \sum_{k=1}^{K} \ln p_u(u_k; \mu, \sigma^2) + \sum_{k=1}^{K}\sum_{t=1}^{N} \ln p_{y\,|\,u}(y_k(t)\,|\,u_k; m) \\
&= K \cdot \left\{ \ln\left(\frac{\xi}{\sqrt{2\pi}}\right) - \frac{1}{2}\ln\sigma^2 + Nm\ln m - N\ln[\Gamma(m)] \right. \\
&\quad + (m-1)N \cdot \left[\frac{1}{KN}\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t)\right] - mN \\
&\quad \cdot \left[\frac{1}{KN}\sum_{k=1}^{K}\sum_{t=1}^{N}\frac{y_k(t)}{u_k}\right] - \frac{\xi^2}{2\sigma^2}\cdot\frac{1}{K}\sum_{k=1}^{K}(\ln u_k)^2 \\
&\quad \left. + \left(\frac{\xi\mu}{\sigma^2} - mN - 1\right)\cdot\left(\frac{1}{K}\sum_{k=1}^{K}\ln u_k\right) - \frac{\mu^2}{2\sigma^2} \right\}. \quad \text{(A.1)}
\end{aligned}
$$

[4]The scalar Newton–Raphson iteration embedded within the "outer" EM iteration converged within three steps and has low computational complexity compared with the E step in (18).

Therefore, the complete-data sufficient statistics are

$$
T_1(\boldsymbol{u}) = \frac{1}{K}\cdot\sum_{k=1}^{K}\ln u_k \tag{A.2a}
$$

$$
T_2(\boldsymbol{u}) = \frac{1}{K}\cdot\sum_{k=1}^{K}(\ln u_k)^2 \tag{A.2b}
$$

$$
\begin{aligned}
T_3(\boldsymbol{y}, \boldsymbol{u}) &= \frac{1}{KN}\sum_{k=1}^{K}\sum_{t=1}^{N}\left[\frac{y_k(t)}{u_k}\right] - \frac{1}{KN}\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t) \\
&= \frac{1}{K}\cdot\sum_{k=1}^{K}\left(\frac{\bar{y}_k}{u_k}\right) - \frac{1}{KN}\cdot\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t) \quad \text{(A.2c)}
\end{aligned}
$$

where $\bar{y}_k$ was defined in (9b). The complete-data log-likelihood (A.1) is easily maximized with respect to μ and σ^2, yielding the following estimates:

$$
\hat{\mu} = \frac{1}{K}\sum_{k=1}^{K} 10\log_{10}(u_k) = \xi T_1(\boldsymbol{u}) \tag{A.3a}
$$

$$
\hat{\sigma}^2 = \frac{1}{K}\sum_{k=1}^{K}[10\log_{10}(u_k) - \hat{\mu}]^2 = \xi^2\,T_2(\boldsymbol{u}) - \hat{\mu}^2. \tag{A.3b}
$$

Then, to find the ML estimate of m based on the complete data, we need to maximize

$$
\Xi(\boldsymbol{y}, \boldsymbol{u}; m) = m\ln m - \ln[\Gamma(m)] - m\,T_1(\boldsymbol{u}) - m\,T_3(\boldsymbol{y}, \boldsymbol{u}) \tag{A.4}
$$

with respect to m. The above expression follows by dividing the concentrated complete-data log-likelihood function $L_c(\boldsymbol{y}, \boldsymbol{u}; [m, \hat{\mu}, \hat{\sigma}^2]^T)$ by KN and neglecting terms that are independent of m. It can be maximized using the Newton–Raphson iteration, which requires the first two derivatives of $\Xi(\boldsymbol{y}, \boldsymbol{u}; m)$ with respect to m:

$$
\frac{\partial\Xi(\boldsymbol{y}, \boldsymbol{u}; m)}{\partial m} = \ln m - \frac{\Gamma'(m)}{\Gamma(m)} - \varphi(\boldsymbol{y}, \boldsymbol{u}) \tag{A.5a}
$$

$$
\frac{\partial^2\Xi(\boldsymbol{y}, \boldsymbol{u}; m)}{\partial m^2} = \frac{1}{m} - \frac{\Gamma(m)\Gamma''(m) - [\Gamma'(m)]^2}{\Gamma(m)^2} \tag{A.5b}
$$

where

$$
\varphi(\boldsymbol{y}, \boldsymbol{u}) = T_1(\boldsymbol{u}) + T_3(\boldsymbol{y}, \boldsymbol{u}) - 1. \tag{A.6}
$$

The complete-data likelihood belongs to an exponential family of distributions, i.e., the log-likelihood (A.1) is linear in the *natural sufficient statistics* (A.2); see, e.g., [26, Ch. 1.6.2] for the definition of the multiparameter exponential family and natural sufficient statistics. In addition, the number of parameters is equal to the number of sufficient statistics. In this case, the EM algorithm is easily derived as follows (see, e.g., [25, Ch. 1.5.3] or [26, Th. 2.4.3, pp. 135–136]).

- The E step reduces to computing the conditional expectations of the complete-data natural sufficient statistics [in (A.2)] given the observed data $\boldsymbol{y}$; see (18). [Note that $\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t)/(KN)$ is constant with respect to this conditional expectation, and hence, $E_{u\,|\,\boldsymbol{y}}[\sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t)/(KN)\,|\,\boldsymbol{y}; \boldsymbol{\theta}] = \sum_{k=1}^{K}\sum_{t=1}^{N}\ln y_k(t)/(KN)$.]
- The M step is reduced to finding the expressions for (obtaining) the complete-data ML estimates of $\boldsymbol{\theta}$ [see (A.3)–(A.6)] and replacing the complete-data sufficient statistics (A.2) that occur in these expressions with their conditional expectations computed in the E step; see (20) and (22).

REFERENCES

[1] G. Stüber, *Principles of Mobile Communication*, Second ed. Norwell, MA: Kluwer, 2001.

[2] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*. New York: Wiley, 2000.

[3] M.-J. Ho and G. Stüber, "Co-channel interference of microcellular systems on shadowed Nakagami fading channels," in *Proc. 43rd Veh. Technol. Conf.*, Secaucus, NJ, May 1993, pp. 568–571.

[4] E. K. Al-Hussaini, A. M. Al-Bassiouni, H. M. Mourad, and H. Al-Shennawy, "Composite macroscopic and microscopic diversity of sectorized macrocellular and microcellular mobile radio systems employing RAKE receiver over Nakagami fading plus lognormal shadowing channel," *Wireless Pers. Commun.*, vol. 21, pp. 309–328, June 2002.

[5] H. Suzuki, "A statistical model for urban radio propagation," *IEEE Trans. Commun.*, vol. COM-25, pp. 673–680, July 1977.

[6] F. Hansen and F. I. Meno, "Mobile fading—Rayleigh and lognormal superimposed," *IEEE Trans. Veh. Technol.*, vol. VT-26, pp. 332–335, Nov. 1977.

[7] T. J. Moulsley and E. Vilar, "Experimental and theoretical statistics of microwave amplitude scintillations on satellite down-links," *IEEE Trans. Antennas Propag.*, vol. AP-30, pp. 1099–1106, Nov. 1982.

[8] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, and W. Papke, "The land mobile satellite communication channel—Recording, statistics, and channel model," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 375–386, May 1991.

[9] G. E. Corazza and F. Vatalaro, "A statistical model for land mobile satellite channels and its application to nongeostationary orbit systems," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 738–742, Aug. 1994.

[10] F. Vatalaro, F. Mazzenga, G. De Maio, and A. Forcella, "The generalized rice lognormal channel model—First and second order statistical characterization and simulation," *Int. J. Satell. Commun.*, vol. 20, pp. 29–45, Jan.–Feb. 2002.

[11] W. Roh and A. Paulraj, "Outage performance of the distributed antenna systems in a composite fading channel," in *Proc. 56th Veh. Technol. Conf.*, Vancouver, BC, Canada, Sept. 2002, pp. 1520–1524.

[12] ——, "MIMO channel capacity for the distributed antenna systems," in *Proc. 56th Veh. Technol. Conf.*, Vancouver, BC, Canada, Sept. 2002, pp. 706–709.

[13] A. Dogandžić and J. Jin, "Estimating statistical properties of composite gamma-lognormal fading channels," in *Proc. Globecom Conf.*, San Francisco, CA, Dec. 2003, pp. 2406–2410.

[14] C. van der Plas and J.-P. M. G. Linnartz, "Stability of mobile slotted ALOHA network with Rayleigh fading, shadowing, and near-far effect," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 359–366, Nov. 1990.

[15] Y.-C. Ko and M.-S. Alouini, "Estimation of Nakagami-m fading channel parameters with application to optimized transmitter diversity systems," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 250–259, Mar. 2003.

[16] A. Duel-Hallen, S. Q. Hu, and H. Hallen, "Long-range prediction of fading signals: Enabling adapting transmission for mobile radio channels," *IEEE Signal Processing Mag.*, vol. 17, pp. 62–75, May 2000.

[17] T. Jiang, N. D. Sidiropoulos, and G. B. Giannakis, "Kalman filtering for power estimation in mobile communications," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 151–161, Jan. 2003.

[18] *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 1972. Ninth printing. New York: Dover, M. Abramowitz and I. A. Stegun, Eds..

[19] G. A. F. Seber and C. J. Wild, *Nonlinear Regression*. New York: Wiley, 1989.

[20] W. H. Press, S. A. Teukolsky, W. T. Vettering, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Second ed, Cambridge, U.K.: Cambirdge Univ. Press, 1992.

[21] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[22] R. A. Thisted, *Elements of Statistical Computing: Numerical Computation*. New York: Chapman & Hall, 1988.

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Stanford, CA, [Online] Available: http://www.stanford.edu/~boyd/cvxbook.html: Stanford Univ.

[24] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc., Ser. B*, vol. 39, pp. 1–38, July 1977.

[25] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley, 1997.

[26] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Second ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

[27] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, Second ed, London, U.K.: Chapman & Hall, 1989.

[28] S. M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[29] A. J. Goldsmith, L. J. Greenstein, and G. J. Foschini, "Error statistics of real-time power measurements in cellular channels with multipath and shadowing," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 439–446, Aug. 1994.

[30] M. Jamshidian and R. I. Jennrich, "Acceleration of the EM algorithm by using quasi-Newton methods," *J. R. Stat. Soc., Ser. B*, vol. 59, no. 3, pp. 569–587, 1997.

[31] Y. Karasawa, K. Kimura, and K. Minamisono, "Analysis of availability improvement in LMSS by means of satellite diversity based on three-state propagation channel model," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 1047–1056, Nov. 1997.