High-Precision LoS Localization Using Composite Nakagami-m Log-Normal Model

Fangqing Xiao, Xiyao Zhou, Zunqi Li, Hongwei Hou, Dirk Slock
Communication Systems Department
Eurecom, France
Email: {fangqing.xiao, xiyao.zhou, zunqi.li, hongwei.hou, dirk.slock}@eurecom.fr

***Abstract*—This paper presents a high-precision localization method tailored for line-of-sight (LoS)-dominated wireless environments. Unlike conventional received signal strength (RSS) techniques that indiscriminately utilize multipath components, our approach leverages LoS path characteristics through a refined Composite Nakagami-m Log-Normal statistical model. We demonstrate that the LoS signal amplitude follows a Nakagami-m distribution (equivalent to Gamma-distributed energy), while its scale parameter exhibits log-normal variation due to free-space path loss. This dual-scale model accurately captures both small-scale fading and distance-dependent attenuation unique to LoS propagation. For practical implementation, we develop an expectation-maximization (EM) algorithm enhanced with Langevin Monte Carlo (LMC) sampling, enabling efficient maximum likelihood estimation. Experimental validation using Quadriga-generated LoS channel data confirms consistent high accuracy, outperforming standard RSS methods in comparable LoS conditions. Although primarily designed for LoS scenarios, the framework offers theoretical insights for future extensions to mixed propagation environments. The solution remains practical for existing infrastructure, providing immediate benefits for 5G/6G systems in LoS-predominant deployments such as urban canyons and millimeter-wave cells.**

***Index Terms*—Localization, Line-of-sight, Nakagami-m, Log-normal, Expectation maximization, Langevin Monte Carlo**

I. Introduction

Precise localization is essential for contemporary wireless systems, supporting applications such as autonomous navigation [1], smart infrastructure [2], and augmented reality [3]. While GPS functions effectively outdoors, alternative methods are required for indoor and urban environments. Cellular network-based localization, which utilizes existing base stations (BSs) and communication signals, offers a viable solution without necessitating additional hardware.

Current RSS-based localization methods are favored for their simplicity and compatibility with standard devices, as they circumvent the need for precise time synchronization or antenna arrays [4], [5], [6]. Nonetheless, these traditional RSS techniques are susceptible to significant errors due to multipath interference and oversimplified channel models [7]. To mitigate these issues, an enhancement in the channel estimation phase is proposed. This involves identifying and modeling the line-of-sight (LoS) path. By accurately characterizing the LoS attenuation and isolating it from the non-line-of-sight (NLoS) components, this method not only improves the accuracy of the model but also reduces interference, all without the need for additional equipment.

To enhance localization accuracy, we propose a novel method that accurately models both large-scale and small-scale fading effects using a Composite Nakagami-m Log-Normal Distribution, also known as Composite Gamma Log-Normal Distribution [8], [9], validated in various industry environments [10], [11]. Specifically, this model uses a Nakagami-m distribution to characterize signal amplitude, indicative of signal energy following a Gamma distribution, while assuming a uniform distribution for the phase. For large-scale attenuation, we model the scale parameter Ω of the Gamma distribution within a log-normal distribution framework, with its mean dictated by a free-space path-loss model. This dual approach effectively captures both the distance-dependent signal decay and rapid small-scale variations.

To determine the user equipment (UE) position from this model, we employ a maximum likelihood estimation (MLE) framework [12]. Given the presence of hidden variables, such as fading parameters, we implement an expectation-maximization (EM) algorithm [13], [14]. The E-step, involving the computation of posterior expectations, presents analytical challenges. To address this, we employ Langevin Monte Carlo (LMC) sampling [15], which facilitates efficient numerical integration and ensures robust convergence with maintained computational efficiency.

Finally, we validate our method using channel data generated by Quadriga [16], which simulates realistic urban environments. Our experimental results indicate significant improvements over traditional RSS-based methods, achieving sub-meter accuracy in typical LoS/NLoS mixed scenarios. The proposed model exhibits superior robustness against small-scale fading. Moreover, the EM-LMC estimator facilitates reliable and efficient position recovery, further enhancing the practicality and accuracy of our localization approach.

A. Notations

The notation $p(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the probability density function of a complex circularly symmetric Gaussian random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. For any matrix $\mathbf{H} \in \mathbb{C}^{\mathbf{M}\times\mathbf{N}}$, $\mathbf{h_i}$ represents its i-th column vector.

For any vector $\boldsymbol{a} \in \mathbb{C}^N$, a_i denotes its i-th element. $\mathbf{I_M} \in \mathbb{R}^{\mathbf{M}\times\mathbf{M}}$ stands for the $M \times M$ identity matrix.

II. System model

Consider a multi-base station (BS) localization system where the user equipment (UE) estimates its position using orthogonal frequency-division multiplexing (OFDM) pilot signals transmitted from K spatially distributed BSs. The system operates under the following assumptions:

A. Pilot Structure and Delay Modeling

Each BS transmits known pilot symbols $\mathbf{S}_k \in \mathbb{C}^{N\times N}$ (diagonal matrix, N subcarriers), and the channel exhibits sparse multipath propagation with *integer-delay* Line-of-Sight (LoS) and Non-LoS (NLoS) components. The delay matrix $\mathbf{T}_k \in \mathbb{C}^{N\times L}$ (where L is the maximum delay spread) is constructed as a Toeplitz matrix incorporating cyclic shifts corresponding to the LoS/NLoS path delays. These delays remain *time-invariant* over short observation intervals due to negligible geometric changes between the UE and BSs.

B. Time-Varying Channel Gains

The complex channel gain vector $\mathbf{a}_k[n] \in \mathbb{C}^{L\times 1}$, representing the amplitudes and phases of the multipath components, varies across time instances n due to small-scale fading. The received signal $\mathbf{y}_k[n] \in \mathbb{C}^{N\times 1}$ at the User Equipment (UE) from the k-th Base Station (BS) at time n is given by:

$$\mathbf{y}_k[n] = \mathbf{S}_k\mathbf{T}_k\mathbf{a}_k[n] + \mathbf{v}_k[n], \quad k = 1,\dots,K, \tag{1}$$

where $\mathbf{v}_k[n] \sim \mathcal{CN}(0,\sigma_v^2\mathbf{I}_N)$ represents the additive white Gaussian noise (AWGN). The matrices $\mathbf{S}_k$ (pilots) and $\mathbf{T}_k$ (delays) are deterministic and known, and the time-invariant delay matrix $\mathbf{T}_k$ encodes delay information between Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) components.

For collaborative localization, the aggregate received signal from all K BSs is stacked into a composite vector:

$$\tilde{\mathbf{y}}[n] = \begin{bmatrix}\mathbf{y}_1[n]\\ \vdots \\ \mathbf{y}_K[n]\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{S}_1\mathbf{T}_1 & \cdots & \mathbf{0}\\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{S}_K\mathbf{T}_K\end{bmatrix}}_{\text{Block-diagonal pilot-delay matrix } \boldsymbol{\Phi}} \begin{bmatrix}\mathbf{a}_1[n]\\ \vdots \\ \mathbf{a}_K[n]\end{bmatrix} + \tilde{\mathbf{v}}[n]. \tag{2}$$

The block-diagonal structure of $\boldsymbol{\Phi}$, resulting from orthogonal pilot assignments across BSs, ensures inter-cell interference-free channel estimation.

C. LoS Channel Matrix Acquisition

The input LoS channel fading matrix $\mathbf{X} \in \mathbb{C}^{M\times K}$, containing the complex gains of the line-of-sight (LoS) paths across M measurement intervals and K base stations, is obtained through the following channel estimation procedure:

$$\mathbf{X} = \begin{bmatrix}\hat{a}_1^{\text{LoS}}[1] & \cdots & \hat{a}_K^{\text{LoS}}[1]\\ \vdots & \ddots & \vdots \\ \hat{a}_1^{\text{LoS}}[M] & \cdots & \hat{a}_K^{\text{LoS}}[M]\end{bmatrix}, \tag{3}$$

where each element $\hat{a}_k^{\text{LoS}}[n]$ represents the estimated LoS channel gain for BS k at time interval n, obtained through the following steps:

1) Pilot-Based Channel Estimation: For each time interval n and BS k, the receiver first estimates the complete multipath channel vector:

$$\hat{\mathbf{a}}_k[n] = (\mathbf{S}_k\mathbf{T}_k)^\dagger\mathbf{y}_k[n]. \tag{4}$$

Moreover, the LoS component is extracted as:

$$\hat{a}_k^{\text{LoS}}[n] = \mathbf{e}_1^\top\hat{\mathbf{a}}_k[n] \tag{5}$$

where $\mathbf{e}_1 = [1,0,\dots,0]^\top$ selects the first component corresponding to the LoS path in the delay-ordered channel vector.

2) Statistical Properties: The estimation error for each LoS component follows:

$$\hat{a}_k^{\text{LoS}}[n] = a_k^{\text{LoS}}[n] + \epsilon_k[n], \quad \epsilon_k[n] \sim \mathcal{CN}(0,\sigma_{\epsilon,k}^2) \tag{6}$$

with estimation variance:

$$\sigma_{\epsilon,k}^2 = \sigma_v^2\|(\mathbf{S}_k\mathbf{T}_k)^\dagger\mathbf{e}_1\|^2 \tag{7}$$

3) Matrix Construction: The complete LoS channel matrix is constructed by collecting estimates across all time intervals and BSs:

$$\mathbf{X} = \mathbf{X}_{\text{true}} + \mathbf{E} \tag{8}$$

where $\mathbf{E}$ is the estimation error matrix with i.i.d. elements $\epsilon_k[n]$. The accuracy of $\mathbf{X}$ directly impacts the subsequent parameter estimation in the EM framework. The EM module operates under the default assumption that matrix $\mathbf{X}$ is accurate.

D. Statistical Modeling of LoS Channel Gain

The LoS channel fading for each base station is modeled as a compound Nakagami-m and log-normal distribution, capturing both small-scale and large-scale fading effects. For simplicity, we define the vector $\mathbf{x}_k \in \mathbb{C}^{M\times 1}$ to collect the M-group LoS channel fading of the k-th BS as:

$$\mathbf{x}_k = [x_{k1},\cdots,x_{kM}]^T = [a_k^{\text{LoS}}[1],\cdots,a_k^{\text{LoS}}[M]]^T.$$

For small-scale fading, each $a_k^{\text{LoS}}[n]$ follows a Nakagami-m distribution in magnitude and a uniform phase distribution over $[0,2\pi)$. Thus, the probability density function (PDF) of x_{kn} is:

$$p_x(x_{kn}|m,\Omega_{kn}) = \frac{m^m|x_{kn}|^{2m-2}}{\pi\Gamma(m)\Omega_{kn}^m}\exp\left[-\frac{m|x_{kn}|^2}{\Omega_{kn}}\right],$$

where Ω_{kn} represents the mean received signal power in the n-th interval between the k-th BS and the user equipment (UE). This is modeled as a hidden random variable with a log-normal distribution:

$$p_\Omega(\Omega_{kn}|\mu_k,\sigma_k^2) = \frac{\eta}{\Omega_{kn}\sqrt{2\pi\sigma_k^2}}\exp\left[-\frac{(\eta\ln\Omega_{kn}-\mu_k)^2}{2\sigma_k^2}\right],$$

where $\eta = 10/\ln 10$, and μ_k and σ_k^2 (in dB) are the scale and shape parameters, respectively:

$$\mu_k = \mathbb{E}\left\{\eta \ln \Omega_{kn}\right\}, \quad \sigma_k^2 = \mathbb{E}\left\{(\eta \ln \Omega_{kn} - \mu_k)^2\right\}.$$

Here, μ_k is defined as:

$$\mu_k = \eta \left(\ln G_0 + n_{\text{LoS}} \ln \frac{\lambda}{4\pi} - n_{\text{LoS}} d_k \right),$$

where G_0 combines transmit power and antenna gain, λ is the wavelength, n_{LoS} is the propagation fading factor and d_k is the LoS distance between the UE and the k-th BS.

Assuming the UE's position is $\mathbf{r}^{\text{UE}} = [r_x^{\text{UE}}, r_y^{\text{UE}}, r_z^{\text{UE}}]$ and the k-th BS's position is $\mathbf{r}^{\text{BS}_k} = [r_x^{\text{BS}_k}, r_y^{\text{BS}_k}, r_z^{\text{BS}_k}]$, the distance d_k is:

$$d_k = \|\mathbf{r}^{\text{UE}} - \mathbf{r}^{\text{BS}_k}\|_2,$$

with $\|\cdot\|_2$ denoting the Euclidean distance:

$$\|\mathbf{r}^{\text{UE}} - \mathbf{r}^{\text{BS}_k}\|_2 = \sqrt{\sum_{i \in \{x,y,z\}} (r_i^{\text{UE}} - r_i^{\text{BS}_k})^2}. \tag{9}$$

Since μ_k decreases linearly with d_k (due to the $-n_{\text{LoS}} d_k$ term), we can estimate the UE's position using multiple BS measurements.

E. Estimation Problem

Assuming we have perfectly accurate observations of $\mathbf{X} \in \mathbb{C}^{M \times K}$ given as $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_K]$, we define a hierarchical probabilistic model for the joint density function of:

- The measurements $\mathbf{X}$,
- The hidden signal $\mathbf{\Omega} \in \mathbb{R}^{M \times K} = [\mathbf{\Omega}_1, \cdots, \mathbf{\Omega}_K]$,
- The unknown parameters $\boldsymbol{\theta} = \left\{m, \{\mu_k^2\}_{k=1}^K, \{\sigma_k^2\}_{k=1}^K\right\}$.

This joint density is expressed as:

$$p(\mathbf{X}, \mathbf{\Omega}; \boldsymbol{\theta}) = \prod_{k=1}^{K} \prod_{n=1}^{M} p_x(x_{kn} | m, \Omega_{kn}) \times p_\Omega(\Omega_{kn} | \mu_k(\mathbf{r}^{\text{UE}}, n_{\text{LoS}}), \sigma_k^2). \tag{10}$$

Our goal is to estimate $\boldsymbol{\theta}$ directly from $\mathbf{X}$ using the maximum likelihood estimator (MLE):

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}; \mathbf{X}) = \arg\max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}; \mathbf{X}), \tag{11}$$

where $\mathcal{L}(\cdot)$ and $\ell(\cdot)$ denote the likelihood and log-likelihood functions, respectively. The likelihood function is given by:

$$\mathcal{L}(\boldsymbol{\theta}; \mathbf{X}) = p(\mathbf{X}|\boldsymbol{\theta}) = \int p(\mathbf{X}, \mathbf{\Omega}; \boldsymbol{\theta})\, d\mathbf{\Omega}. \tag{12}$$

However, directly solving this integral to obtain $p(\mathbf{X}|\boldsymbol{\theta})$ is intractable due to the lack of an analytical solution. Additionally, the latent variables $\mathbf{\Omega}$ are unobserved, and their distributions depend on the unknown $\boldsymbol{\theta}$. To address these challenges, we introduce the Expectation Maximization (EM)- Langevin Monte Carlo (LMC) algorithm in the following section.

III. Expectation Maximization (EM) - Langevin Monte Carlo (LMC) Algorithm

A. Review of Expectation Maximization

To solve the optimization problem in (11), we employ the expectation-maximization (EM) algorithm, which is particularly effective for estimation problems involving latent variables like $\mathbf{\Omega}$. Using the minorization-maximization (MM) framework, we iteratively maximize a lower bound of the log-likelihood function. At the t-th iteration, given the current estimate $\boldsymbol{\theta}^{(t)}$, we apply Jensen's inequality and the concavity of $\ln(\cdot)$ to construct a tractable lower bound for the log-likelihood. This allows us to approximate the optimal parameters through iterative refinement as below:

$$\ell(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta}^{(t)}) \geq \iint p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta}^{(t)}) \ln \frac{p(\mathbf{X}, \mathbf{\Omega}; \boldsymbol{\theta})}{p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta}^{(t)})} d\mathbf{\Omega} - \int p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta}^{(t)}) \ln p(\mathbf{X}|\boldsymbol{\theta}) d\mathbf{\Omega}. \tag{13}$$

where $p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta})$ is the posterior distribution of $\mathbf{\Omega}$ as which can be expressed by Bayes' rule as:

$$p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta}) = \frac{p(\mathbf{X}, \mathbf{\Omega}; \boldsymbol{\theta})}{\int p(\mathbf{X}, \mathbf{\Omega}; \boldsymbol{\theta}) d\mathbf{\Omega}}. \tag{14}$$

By observing (13), it is obvious that optimizing $\boldsymbol{\theta}$ by maximizing $\ell(\boldsymbol{\theta})$ also equals to maximizing the loss function as below:

$$\boldsymbol{\theta}^{(t+1)} = \arg\max_{\boldsymbol{\theta}} \mathbb{E}_{p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta}^{(t)})}[\ln p(\mathbf{X}, \mathbf{\Omega}; \boldsymbol{\theta})]. \tag{15}$$

Furthermore, it can be shown that at convergence, we obtain $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t+1)}$. At this point, Jensen's inequality becomes an equality, demonstrating that the EM algorithm converges to a (local) optimum. We now derive the parameter update rule for $\boldsymbol{\theta}^{(t+1)}$ based on the current estimate $\boldsymbol{\theta}^{(t)}$.

1) Optimization of Nakagami-m Parameter:: The estimation of the Nakagami-m parameter proceeds through the following rigorous derivation within the EM framework. First, we construct the complete-data log-likelihood function for parameter m given the observed measurements $\mathbf{X}$ and latent variables $\mathbf{\Omega}$:

$$\ell_c(m) = \sum_{k=1}^{K} \sum_{n=1}^{M} [m \ln m - m \ln \Omega_{kn} - \ln \Gamma(m) + (m-1) \ln |x_{kn}|^2 - \frac{m|x_{kn}|^2}{\Omega_{kn}} - \ln |x_{kn}|^2]. \tag{16}$$

Taking the conditional expectation with respect to $p(\mathbf{\Omega}|\mathbf{X}, \boldsymbol{\theta}^{(t)})$ yields the Q-function:

$$Q(m|\boldsymbol{\theta}^{(t)}) = m \ln m - m \ln[\Gamma(m)] + m\xi^{(t)}, \tag{17}$$

where

$$\xi^{(t)} = \frac{1}{KM}\sum_{k=1}^{K}\sum_{n=1}^{M}\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\ln\Omega_{kn}] + \frac{1}{KM}\sum_{m=1}^{N}\sum_{k=1}^{K}[|x_{kn}|^2]\mathbb{E}_{p(\Omega_{kn}|\mathbf{X},\boldsymbol{\theta}^{(t)})}[\Omega_{kn}^{-1}] - \frac{1}{KM}\sum_{n=1}^{M}\sum_{k=1}^{K}[\ln|x_k(l)|^2], \tag{18}$$

The maximization of $Q(m|\boldsymbol{\theta}^{(t)})$ leads to the following non-linear equation for $m^{(t+1)}$:

$$\ln m - \psi(m) + 1 = \xi^{(t)} \tag{19}$$

where $\psi(m) = \Gamma'(m)/\Gamma(m)$ is the digamma function. This transcendental equation requires numerical solution via the Newton-Raphson method. The iterative update rule is derived by computing the first and second derivatives of the Q-function:

$$\frac{\partial Q}{\partial m} = (\ln m - \psi(m)) + \xi^{(t)}, \quad \frac{\partial^2 Q}{\partial m^2} = \frac{1}{m} - \psi'(m), \tag{20}$$

where $\psi'(m)$ is the trigamma function. The Newton-Raphson iteration thus becomes:

$$m^{(i+1)} = m^{(i)} - \left[\frac{\partial^2 Q}{\partial m^2}\right]^{-1}\left[\frac{\partial Q}{\partial m}\right]. \tag{21}$$

For practical implementation, we initialize the Newton-Raphson method using the moment-based estimator [12]:

$$m_{\text{init}}^{(t+1)} = \frac{1+2\xi^{(t)}}{(2\xi^{(t)}-2)(2+\xi^{(t)})}. \tag{22}$$

The iteration continues until convergence is achieved, typically when $|m^{(i+1)} - m^{(i)}| < \epsilon$ for some small tolerance $\epsilon > 0$. The final value is taken as $m^{(t+1)} = m^{(\infty)}$. The expectations in these expressions are computed as:

$$\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\ln\Omega_{kn}] = \int_0^\infty \ln\Omega_{kn}\, p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})d\Omega_{kn}, \tag{23}$$

$$\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\Omega_{kn}^{-1}] = \int_0^\infty \Omega_{kn}^{-1}\, p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})d\Omega_{kn}, \tag{24}$$

where the conditional distribution $p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})$ is derived from the Nakagami-m and log-normal mixture model. The complete algorithm guarantees monotonic increase in the likelihood function and convergence to a local maximum under standard regularity conditions.

2) Optimization of Log-Normal Parameters: The closed-form updates for the log-normal distribution parameters μ_k and σ_k^2 are derived as follows:

$$\mu_k^{(t+1)} = \frac{\eta}{M}\sum_{n=1}^{M}\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}\left[\ln\Omega_{kn}\right], \tag{25}$$

$$(\sigma_k^2)^{(t+1)} = \frac{\eta^2}{M}\sum_{n=1}^{M}\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}\left[(\ln\Omega_{kn})^2\right] - (\mu_k^{(t+1)})^2, \tag{26}$$

where

$$\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})} = \int_0^\infty (\ln\Omega_{kn})^2 p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})d\Omega_{kn}. \tag{27}$$

B. Langevin Monte Carlo Algorithm for Expectaion Approximation

In the maximization step, we have to calculate the expectation $\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\ln\Omega_{kl}]$, $\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[(\ln\Omega_{kl})^2]$ and $\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\Omega_{kn}^{-1}]$. The posterior distribution $p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})$ is intractable in closed form, but its gradient $\nabla_{\Omega_{kn}}\ln p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})$ can be computed (up to a normalization constant) via:

$$\nabla_{\Omega_{kn}} p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)}) = \nabla_{\Omega_{kn}}\ln p_\Omega(\Omega;\mu_k^{(t)},(\sigma_k^2)^{(t)}) + \nabla_{\Omega_{kn}}\ln p(x_{kn}|\Omega_{kn},m^{(t)}). \tag{28}$$

By using the Langevin Monte Carlo (LMC) algorithm, the samples $\{\Omega_{kn}^{(i)}\}_{i=1}^P$ from $p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})$ by iterating the following dynamics:

$$\Omega_{kn}^{(i+1)} = \Omega_{kn}^{(i)} + \alpha\nabla_{\Omega_{kn}} p(\Omega_{kn}^{(i)}|x_{kn},\boldsymbol{\theta}^{(t)}) + \sqrt{2\alpha}\gamma^{(i)}, \tag{29}$$

where $\alpha > 0$ is the step size, $\gamma^{(i)}$ is the standard complex Gaussian noise with zero mean and one variance. Using T samples $\{\Omega_{kn}^{(t)}\}_{t=1}^T$ from LMC, the posterior moments are estimated empirically:

$$\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\ln\Omega_{kl}] \approx \frac{1}{T}\sum_{t=1}^{T}\ln(\Omega_{kl}^{(t)}), \tag{30}$$

$$\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[(\ln\Omega_{kl})^2] \approx \frac{1}{T}\sum_{t=1}^{T}(\ln(\Omega_{kl}^{(t)})^2, \tag{31}$$

$$\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\Omega_{kn}^{-1}] \approx \frac{1}{T}\sum_{t=1}^{T}\frac{1}{\Omega_{kl}^{(t)}}. \tag{32}$$

C. Optimization of UE Position and Path Loss Exponent:

If the EM-LMC converged, we have the $\{\hat{\mu}_k\}_{k=1}^K$ and $\{\hat{\sigma}_k^2\}_{k=1}^K$ the joint optimization of user equipment position $\mathbf{r}^{\text{UE}}$ and path loss exponent n_{LoS} is formulated as:

$$[(\hat{\mathbf{r}}^{\text{UE}}), \hat{n}_{\text{LoS}}] = \arg\min_{\mathbf{r}^{\text{UE}},n_{\text{LoS}}}\ \psi(\mathbf{r}^{\text{UE}},n_{\text{LoS}}) = \arg\min_{\mathbf{r}^{\text{UE}},n_{\text{LoS}}}\sum_{k=1}^{K}\frac{1}{(\hat{\sigma}_k^2)}\left\|\hat{\mu}_k - f(\mathbf{r}^{\text{UE}},\mathbf{r}^{\text{BS}_k},n_{\text{LoS}})\right\|^2, \tag{33}$$

where the channel propagation model $f(\cdot)$ is defined as:

$$f(\mathbf{r}^{\mathrm{UE}}, \mathbf{r}^{\mathrm{BS}_k}, n_{\mathrm{LoS}}) = \eta\Bigg(\ln G_0 + n_{\mathrm{LoS}} \ln \frac{\lambda}{4\pi} - n_{\mathrm{LoS}} \|\mathbf{r}^{\mathrm{UE}} - \mathbf{r}^{\mathrm{BS}_k}\|_2 \Bigg). \quad (34)$$

The joint optimization problem in (33) is solved through the following steps:

1) Initialization:
   - Set $(\mathbf{r}^{\mathrm{UE}})^{(0)}$ as the centroid of BS positions
   - Initialize $n_{\mathrm{LoS}}^{(0)} = 2$ (typical free-space value)
   - Compute initial $\zeta_k^{(0)} = f((\mathbf{r}^{\mathrm{UE}})^{(0)}, \mathbf{r}^{\mathrm{BS}_k}, n_{\mathrm{LoS}}^{(0)})$ using (34)
2) Gradient Computation: The gradient components are:

$$\frac{\partial \psi}{\partial \mathbf{r}^{\mathrm{UE}}} = 2 n_{\mathrm{LoS}} \eta \sum_{k=1}^{K} \frac{\Delta \psi_k}{\hat{\sigma}_k^2} \cdot \frac{\mathbf{r}^{\mathrm{UE}} - \mathbf{r}^{\mathrm{BS}_k}}{\|\mathbf{r}^{\mathrm{UE}} - \mathbf{r}^{\mathrm{BS}_k}\|_2}, \quad (35)$$

$$\frac{\partial \psi}{\partial n_{\mathrm{LoS}}} = 2\eta \sum_{k=1}^{K} \frac{\Delta \psi_k}{\hat{\sigma}_k^2} \left(\ln \frac{\lambda}{4\pi} - \|\mathbf{r}^{\mathrm{UE}} - \mathbf{r}^{\mathrm{BS}_k}\|_2 \right) \quad (36)$$

   where $\Delta \mu_k = \mu_k^{(t+1)} - f(\mathbf{r}^{\mathrm{UE}}, \mathbf{r}^{\mathrm{BS}_k}, n_{\mathrm{LoS}})$.
3) Iterative Optimization:
   - Apply limited-memory BFGS (L-BFGS) with Wolfe conditions [17].
   - Enforce $n_{\mathrm{LoS}} \in [1.6, 3.5]$ (empirical range).
   - Use Armijo line search with $\alpha_0 = 0.5$.
4) Termination Criteria:

$$\|\mathbf{r}_{(i+1)}^{\mathrm{UE}} - \mathbf{r}_{(i)}^{\mathrm{UE}}\|_2 < \epsilon_r \quad (\epsilon_r = 0.5\text{meter}), \quad (37)$$

$$|n_{\mathrm{LoS}}^{(i+1)} - n_{\mathrm{LoS}}^{(i)}| < \epsilon_n \quad (\epsilon_n = 0.01). \quad (38)$$

D. EM-LMC Algorithm Implementation

The complete Expectation-Maximization Langevin Monte Carlo (EM-LMC) algorithm is presented in Alg. 1. The implementation employs the following carefully tuned parameters:

- Convergence Threshold: $\epsilon = 0.001$ (relative change in log-likelihood).
- Learning Rate: $\alpha = 0.2$ (for Langevin dynamics step).
- Particle Count: $P = 200$ (parallel Markov chains).
- Thermalization Steps: $T = 1000$ (burn-in iterations per EM step).
- Maximum Iterations: $I_{\max} = 200$ (EM cycle limit).

IV. Simulation Result

A. Simulation Configuration

In this section, we present simulation results to evaluate the positioning performance of our proposed algorithm. We utilize the QuaDriGa channel simulator, which is consistent with the third Generation Partnership Program (3GPP) New Radio (NR) specifications [18] and has been validated in various field trials [16]. The simulator models the 3GPP urban macro (UMa) line-of-sight (LoS) scenarios, comprising ten clusters with ten subpaths each, sharing similar physical parameters. We configure four base stations (BSs) with coordinates set at $\mathbf{r}^{\mathrm{BS}_1} = [0, 0, 10]$, $\mathbf{r}^{\mathrm{BS}_2} = [50, 0, 10]$, $\mathbf{r}^{\mathrm{BS}_3} = [0, 50, 10]$, and $\mathbf{r}^{\mathrm{BS}_4} = [50, 50, 10]$ in meters. The user equipment (UE) position is $\mathbf{r}^{\mathrm{UE}} = [25, 25, 1]$ in meters, and the altitude of UE is supposed to be known. The carrier frequency is set to 6.7 GHz, a critical value for future communications in the upper mid-band. We use 64 pilot symbols with a subcarrier spacing of 120 kHz. The channel noise level is set at 20 dB SNR. Signal recovery of $\mathbf{X}$ is performed using the least squares estimate. Additionally, the antenna gain G_0 is specified as 0 dB. To underscore the effectiveness of our LoS CSI-based localization method, we apply the RSS-based localization method via the Log-normal

Algorithm 1 EM-LMC Based Localization

Input: LoS Channel Fading Matrix: $\mathbf{X}$ based on (3) and (5)
Output: UE's Positioning: $\hat{\mathbf{r}}^{\mathrm{UE}}$
Initialize: $\{m^{(0)}, \{(\mu_k^2)^{(0)}, (\sigma_k^2)^{(0)}\}_{k=1}^K\}$
while *stopping criterion not fulfilled* **do**
  // Expectation Step
  while *All expectation w.r.t.* $\mathbf{\Omega}$ *are obtained* **do**
    Generate LMC Samples $\{\Omega_{kn}^{(i)}\}_{i=1}^P$ based on (29)
    Update $\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\ln \Omega_{kl}]$ based on (30)
    Update $\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[(\ln \Omega_{kl})^2]$ based on (31)
    Update $\mathbb{E}_{p(\Omega_{kn}|x_{kn},\boldsymbol{\theta}^{(t)})}[\Omega_{kn}^{-1}]$ based on (32)
  // Maximization Step
  while *All parameters are optimized* **do**
    Update $m^{(t)}$ based on (21) and (22)
    Update $\{(\mu_k^2)^{(t)}, (\sigma_k^2)^{(t)}\}_{k=1}^K$ based on (25) and (26)
// UE Position Estimation
Optimize $\hat{\mathbf{r}}^{\mathrm{UE}}$ and $\hat{n}_{\mathrm{LoS}}$ based on (34) abd III-C

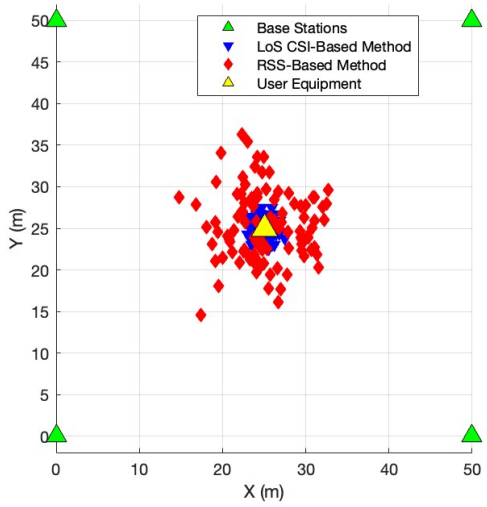

Fig. 1: Localization Results.

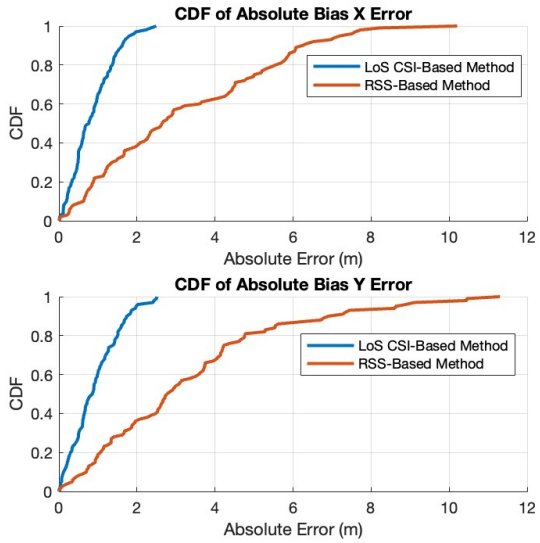

Fig. 2: CDF of Absolute Bias Error.

statistical model[19].

B. Simulation Results

In our experiments, we compare the performance of our proposed method, represented in blue, against a benchmark method, represented in red. Our approach estimates the n_{LoS} at approximately 2.27 and m at 1.1. Fig. 1 displays the outcomes of 100 estimation trials. It is evident that our results cluster more closely around the true values, showcasing the higher reliability of our method. In Fig. 2, which plots the results on x and y coordinates, we evaluate the performance from the cumulative distribution function (CDF) perspective. The 95% confidence interval for the x-axis error is 1.8 meters, and for the y-axis, it is 1.97 meters. These values significantly surpass those of the comparison trial. These findings substantiate the effectiveness of our method, highlighting its superior accuracy and reliability in practical scenarios.

V. Conclusions

This paper proposes a high-precision localization method for LoS-dominated environments by modeling LoS path characteristics via a Composite Nakagami-m Log-Normal distribution, which jointly captures small-scale fading (Nakagami-m) and large-scale path loss (log-normal). We develop an EM-LMC estimator to efficiently solve the associated MLE problem, addressing hidden variables through Langevin Monte Carlo sampling. Validated with 3GPP-compliant Quadriga simulations, our method achieves sub-meter accuracy (95% CI: 1.8 m x-axis, 1.97 m y-axis), outperforming traditional RSS techniques in LoS conditions. The framework's computational efficiency and compatibility with existing infrastructure make it practical for 5G/6G deployments in urban canyons or mmWave cells. While optimized for LoS scenarios, future work will extend the model to mixed LoS/NLoS environments via hybrid measurements and adaptive tuning, further enhancing its applicability for autonomous navigation, smart cities, etc..

Acknowledgements EURECOM's research is partially supported by its industrial members: ORANGE, BMW, SAP, iABG, Norton LifeLock, by the Franco-German project 5G-OPERA (BPI), the French project YACARI (PEPR-5G), the EU INFRA project CONVERGE, and by a Huawei France funded Chair towards Future Wireless Networks.

References

[1] M. Eskandari, A. V. Savkin, and W. Ni, "Consensus-based autonomous navigation of a team of RIS-equipped UAVs for LoS wireless communication with mobile nodes in high-density areas," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 2, pp. 923–935, 2022.

[2] A. L. Imoize, O. Adedeji, N. Tandiya, and S. Shetty, "6G enabled smart infrastructure for sustainable society: Opportunities, challenges, and research roadmap," *Sensors*, vol. 21, no. 5, p. 1709, 2021.

[3] Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "A survey on mobile augmented reality with 5G mobile edge computing: Architectures, applications, and technical aspects," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 2, pp. 1160–1192, 2021.

[4] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv. (CSUR)*, vol. 46, no. 2, pp. 1–32, 2013.

[5] S. Sadowski and P. Spachos, "Rssi-based indoor localization with the internet of things," *IEEE Access*, vol. 6, pp. 30 149–30 161, 2018.

[6] R.-H. Wu, Y.-H. Lee, H.-W. Tseng, Y.-G. Jan, and M.-H. Chuang, "Study of characteristics of RSSI signal," in *2008 IEEE Int. Conf. Industrial Technol.*, pp. 1–3.

[7] F. Xiao, Z. Zhao, and D. T. Slock, "Multipath Component Power Delay Profile Based Ranging," *IEEE J. Sel. Top. Signal Process.*, 2024.

[8] F. Hansen and F. I. Meno, "Mobile fading—Rayleigh and lognormal superimposed," *IEEE Trans. Veh. Technol.*, vol. 26, no. 4, pp. 332–335, 1977.

[9] A. C. Turlapaty, "Variational Bayesian estimation of statistical properties of composite gamma log-normal distribution," *IEEE Trans. Signal Process.*, vol. 68, pp. 6481–6492, 2020.

[10] T. Olofsson, A. Ahlen, and M. Gidlund, "Modeling of the fading statistics of wireless sensor network channels in industrial environments," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3021–3034, 2016.

[11] K. Cho, J. Lee, and C. G. Kang, "Stochastic geometry-based coverage and rate analysis under Nakagami & log-normal composite fading channel for downlink cellular networks," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1437–1440, 2017.

[12] A. Dogandzic and J. Jin, "Maximum likelihood estimation of statistical properties of composite gamma-lognormal fading channels," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2940–2945, 2004.

[13] S. Borman, "The expectation maximization algorithm-a short tutorial," *Submitted for publication*, vol. 41, 2004.

[14] F. Xiao and D. Slock, "Parameter estimation via expectation maximization-expectation consistent algorithm," in *ICASSP 2024*. IEEE, pp. 9506–9510.

[15] M. Girolami and B. Calderhead, "Riemann manifold langevin and hamiltonian monte carlo methods," *J. R. Stat. Soc. Ser. B Stat. Methodol.*, vol. 73, no. 2, pp. 123–214, 2011.

[16] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, 2014.

[17] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical programming*, vol. 45, no. 1, pp. 503–528, 1989.

[18] "Study on channel model for frequencies from 0.5 to 100 ghz," 3GPP, Technical Report TR 38.901, 2024, release 18.

[19] Y. Chu, W. Guo, K. You, L. Zhao, T. Peng, and W. Wang, "RSS-Based Multiple Co-Channel Sources Localization With Unknown Shadow Fading and Transmitted Power," *IEEE Trans. Veh. Technol.*, vol. 71, no. 12, pp. 13 344–13 355, 2022.