Joint User Node Positioning and Clock Offset Estimation in 5G Ultra-Dense Networks

Janis Werner*, Mário Costa†, Aki Hakkarainen*, Kari Leppänen† and Mikko Valkama*
* Department of Electronics and Communications Engineering, Tampere University of Technology
† Huawei Technologies Oy (Finland) Co., Ltd, Finland R&D Center
Emails: {janis.werner, aki.hakkarainen, mikko.e.valkama}@tut.fi
Emails: {mariocosta, kari.leppanen}@huawei.com

Abstract—It is commonly expected that network densification will play an important role in achieving the capacity demands of 5G communication networks. While densification is introduced to improve the spectral efficiency and area-capacity, it also results in an infrastructure that is perfectly suitable for user node (UN) positioning. However, so far this compelling opportunity has not been clearly recognized in the literature. In this paper, we therefore propose to make “always on” positioning an integral part of 5G networks such that highly accurate UN position estimates are available at any given moment but without draining the UN batteries. We furthermore propose an extended Kalman filter (EKF) that tracks the UN position based on the fusion of direction of arrival (DoA) and time of arrival (ToA) estimates obtained at the access nodes (ANs) of the 5G network. Since ToA estimates are typically not useful for positioning unless the UN is synchronized with the network, we include a realistic clock model within the DoA/ToA EKF. This addition makes it possible to estimate the offset of the imperfect UN clock, along with the UN position. In an extensive analysis that is based on specific 5G simulation models, we then quantify the enormous potential of high accuracy positioning in 5G networks, in general, and the proposed DoA/ToA EKF, in particular. Moreover, we demonstrate that the proposed DoA/ToA EKF substantially outperforms the classical DoA-only EKF and is furthermore also able to handle practically extremely relevant situations where the DoA-only EKF fails to position the UN.

Index Terms—Angle-of-arrival, direction-of-arrival, localization, location-awareness, time-of-arrival, tracking, ultra-dense networks, 5G networks

I. INTRODUCTION

In order to achieve the demands expected for 5G wireless communication networks, it is expected that access nodes (ANs) with high spatial density are deployed (see, e.g., [1]). Consequently, user nodes (UNs) in 5G will operate within the range of several ANs simultaneously, such that devices are expected to be in the line of sight (LoS) of a few ANs for most of the time. In addition to meeting the increased communication demand, this also creates an opportunity for accurate device positioning based on time of arrival (ToA) estimates, obtained at the LoS ANs. Since it is expected that the ANs are also equipped with antenna arrays, ANs can furthermore estimate the direction of arrivals (DoAs), which can, in turn, be used to improve the device positioning accuracy. Overall, this results in great potential to develop and provide highly accurate device positioning within 5G networks that has significant advantages compared to already existing approaches.

On the technical side, the envisioned wide waveform bandwidths in 5G systems make it possible to obtain highly accurate ToA estimates that, in combination with the DoA estimates, can be fused into position estimates with extremely high accuracy. Therefore, 5G device positioning has the potential to substantially outperform existing techniques such as global navigation satellite systems (GNSSs) (≈ 5 m [2]), LTE observed time difference of arrival (OTDoA) (≈ 25 m [3]) or WLAN fingerprinting (3 – 4 m [4]). Second, since 5G positioning can potentially be carried out entirely on the network side, the power consumption within the user devices is exceptionally low. In fact, in the network-centric scenario, the only contribution to 5G positioning on the UN side is the transmission of uplink signals that are used to estimate the ToA and DoA within the ANs. However, these signals do not necessarily need to be dedicated positioning signals, but can be basically the same reference/pilot signals that are anyways exchanged between the devices and ANs for multiple-input multiple-output (MIMO) channel estimation [5]. Thus, the power consumption within the UNs, due to positioning capabilities, could be even a 100 times less compared to that of global positioning system (GPS) (≈ 100 – 150 mW [6]). As a consequence, and in contrast to GPS and other GNSSs, 5G positioning can continuously run in the background, providing highly accurate position estimates at any given moment. Third, the 5G positioning concepts will also work indoors, opening business opportunities to an important market. In contrast, classical GNSS-based solutions are not able to provide accurate positioning indoors as they require direct visibility to the satellites. Finally, 5G positioning would also allow for the determination of altitude, which is of particular importance in indoor positioning as well as for unmanned aerial vehicles (UAVs).

With all of these advantages, 5G positioning is expected to meet the high demands of future location based services and applications. Moreover, an accurate positioning infrastructure will be needed also for the navigation and mutual coordination of, e.g., robots that are expected to be commonplace by 2020-2030. Similarly, vehicles would benefit from fast, reliable and accurate positioning that enables advanced collision-avoidance as well as automatic driving gains. Ultimately, the UN position can also be capitalized in the 5G radio network itself in order to

This work was supported by the Doctoral Program of the President of Tampere University of Technology and the Finnish Funding Agency for Technology and Innovation (Tekes, under the project “5G Networks and Device Positioning”).

Online video material available at www.tut.fi/5G/GLOBECOM15.
predict mobility and overall to carry out enhanced location-based radio resource management (RRM), for example.

In summary, the deployment of 5G networks with a high density of ANs provides an unprecedented opportunity to create an advanced positioning system that will be able to meet the demands of future location-based services and applications, and greatly enhance the RRM of the 5G radio network. However, so far this opportunity has not been clearly recognized in the academy and industry. In this paper, we propose and demonstrate the capabilities of a positioning system that is specifically designed for high accuracy user device position estimation and tracking in 5G ultra-dense networks. More specifically, we propose an extended Kalman filter (EKF) that fuses the ToA and DoA estimates from a single or multiple LoS ANs into a position estimate. In the past, EKFs have been proposed for tracking the position of a target using DoA estimates only [7] as well as using a combination of DoA and ToA estimates [8]. However, the DoA/ToA-based method proposed in [8] assumes perfect timing synchronization among the sensors of the positioning system as well as with the device to be positioned. While synchronization among the ANs of ultra-dense networks is a reasonable assumption, it is unrealistic to assume synchronization with the UN. Obtaining the latter form of synchronization is particularly challenging due to the cheap oscillators that are commonly installed in user devices. These imperfect oscillators result in large and time-varying clock offsets, which, if not corrected, render the respective ToA estimates useless for positioning. As a consequence, the method proposed in [8] is not suitable for network-centric UN positioning.

In this paper, we address the problem of UN clock synchronization by including a realistic clock offset model from [9] within the EKF. The resulting joint DoA/ToA EKF is then able to estimate both the UN position as well as the clock offset of the user device. To test the proposed positioning system, we first demonstrate the achievable DoA/ToA estimation performance using the respective Cramer-Rao bound (CRB) in conjunction with the 5G simulation models from the “Mobile and wireless communications enablers for the twenty-twenty information society (METIS)” project [10], [11]. Thereafter, we use these results to show that the proposed joint DoA/ToA EKF provides position estimates with significantly higher accuracy than the classical DoA-only EKF [7], while, at the same time, obtaining also highly accurate clock offset estimates. Moreover, we show that, in contrast to DoA-only based solutions, the proposed joint DoA/ToA EKF is capable of tracking the user position even if only a single LoS AN exists.

In summary, the contributions of this paper are as follows:

- We propose and discuss “always on” UN positioning in 5G.
- We simulate and analyze the achievable DoA/ToA estimation performance in 5G networks using the respective CRB in combination with the METIS channel model.
- We propose and test a novel joint DoA/ToA-EKF for joint UN positioning and clock offset estimation using realistic simulation models.

The paper is organized as follows. In Section II we discuss the proposed 5G positioning engine as well as the applied clock offset models. The CRB that we use to evaluate the achievable DoA and ToA estimation performance is shortly summarized in Section III.
on azimuth and elevation DoA estimation is possible and follows straightforwardly from the work presented in this paper.

B. Clock Offset

In the literature, it is generally agreed that the clock offset \( \rho \) is a time-varying quantity due to imperfections of the device oscillators, see e.g., [9], [14]. For a measurement period \( T \), the clock offset can therefore be written in a recursive form as [9]
\[
\rho[n] = \rho[n-1] + \alpha[n]T
\]
(1)

where \( \alpha[n] \) is known as the clock skew. Some authors, e.g., those of [14] assume the clock skew to be constant, while some recent research based on measurements suggests that the clock skew is in fact time-dependent, at least over the large observation period (1.5 months) considered in [9]. However, taking the research and measurement results in [15], [16] into account, where devices are identified remotely based on an estimate of the average clock skew, one could assume that the average clock skew is indeed constant. This also matches with the measurement results in [9], where the clock skew seems to be fluctuating around a mean value. Nevertheless, the measurements in [9], [15], [16] were obtained indoors, i.e., in a temperature controlled environment. In practice, it would be surprising if large changes in the ambient temperature due to, e.g., a transition from indoors to outdoors would not affect the clock skew at all. Therefore, we adopt the more general model [9] of a time-varying clock skew, which also encompasses the constant clock skew model.

The clock skew in [9] is modeled as an auto-regressive (AR) process of order \( P \). While the measurement results in [9] reveal that modeling the clock offset as an AR process results in large performance gains compared to a constant clock skew model, an increase of the order beyond \( P = 1 \) does not seem to increase the accuracy of clock offset tracking significantly. In here, we consequently model the clock skew as an AR model of first order according to
\[
\alpha[n] = \beta \alpha[n-1] + \eta[n]
\]
(2)

where \( \beta \) is a (constant) parameter and \( \eta[n] \sim \mathcal{N}(0, \sigma^2) \) is additive white Gaussian noise (AWGN). Note that the joint DoA/ToA EKF proposed in Section IV could easily be extended to AR processes of higher orders.

III. CRB on DoA and ToA Estimation

The CRB is a lower bound on the estimation variance of any unbiased estimator. This section states the CRB on 2D DoA and delay estimation of the LoS path on a dense multipath channel. In this paper, the DoA and ToA estimates used by the EKF proposed in Section IV follow a distribution that is given by the CRB. We choose such an approach in order to identify the ultimate performance limits of DoA/ToA estimation in ultra-dense networks, independent of a particular estimator. In practice, the DoA and ToA estimates used by the EKF may be found from the RIMAX algorithm, for example [17]. Interestingly, the RIMAX algorithm is known to be statistically efficient in the asymptotic regime, and very close to the CRB in practice [17]. Thus, this approach does not limit the generality of the proposed EKF and is moreover also a good indication of the performance that can be achieved with practical estimators.

We model each uplink (UL)-single input multiple output (SIMO) channel as a superposition of a single specular component and diffuse components. As motivated in Section II, we consider the channels corresponding to the ANs that are in LoS with the UN, only. In particular, let \( z_k \in C^{M_f \times 1} \) denote the UL-SIMO channel of the \( k \)th AN for all of the \( M_f \) subcarriers, i.e. \( z_k = [\hat{z}_k(1), \ldots, \hat{z}_k(M_f)]^T \). The UL-SIMO channel may be estimated from UL pilot signals, and it is also needed in multiantenna wireless communications for spatial-division multiple-access and beamforming.

In this paper, the UL-SIMO channel is parameterized as follows:
\[
z_k(\xi) = \left( [\hat{a}_k^H(\theta, \varphi), \hat{a}_k^V(\theta, \varphi)] \otimes \hat{a}_k(\tau) \right) \gamma + n, \quad (3)
\]
where \( \otimes \) denotes the Kronecker product and the LoS-parameter vector \( \xi \in \mathbb{R}^{3 \times 1} \) is given by \( \xi = [\tau, \theta, \varphi]^T \). In particular, \( \tau \in \mathbb{R}_+^M \), \( \theta \in [0, \pi] \), \( \varphi \in [0, 2\pi] \) denote the ToA of the LoS path as well as the elevation and azimuth DoAs, respectively. Moreover, vectors \( a_k^H(\theta, \varphi) \in C^{N_R \times 1} \) and \( a_k^V(\theta, \varphi) \in C^{N_R \times 1} \) in (3) denote the array steering vectors of the \( k \)th AN for a horizontal and vertical polarization, respectively. Furthermore, vectors \( a_k(\tau) \in C^{M_f \times 1} \) denotes the frequency-response of the UL-SIMO channel, including the Tx/Rx RF-chains, and \( M_f \in \mathbb{N} \) denotes the number of subcarriers. For the sake of simplicity, we assume that the Rx RF-chains are identical. Finally, vector \( \gamma \in C^{2 \times 1} \) denotes the complex-valued weights of the LoS path in both horizontal and vertical polarizations while \( n \in C^{M_f \times N_R} \) models both the diffuse components of the channel and the measurement noise. Vector \( n \) is zero-mean complex–circular multivariate Gaussian distributed with a covariance matrix given by
\[
R_n = C^{M_f \times N_R \times M_f \times N_R}.
\]

The CRB for the LoS-parameters is given by \( \text{CRB}(\theta) = I^{-1}(\theta) \) where \( I(\theta) \in \mathbb{R}^{7 \times 7} \) denotes the Fisher information matrix (FIM) and \( \theta = [\xi^T, \Im\{\gamma\}^T, \Re\{\gamma\}^T]^T \in \mathbb{R}^{7 \times 1} \) the parameter vector. The FIM can be shown to equal [18, Sec. 4.1]
\[
I(\theta) = 2RI(D_{f}^H\gamma R_{f}^{-1}D_{f} \otimes D_{g}^H\gamma R_{g}^{-1}D_{g}), \quad (4)
\]
where \( \otimes \) denotes the Hadamard-Schur product and \( R_{f} \) and \( R_{g} \) denote the covariance matrices of the diffuse components in the delay and angular domain, respectively. Moreover, matrices \( D_{f} \in C^{M_f \times 7} \), \( D_{g} \in C^{1 \times 7} \), and \( D_{g} \in C^{N_R \times 7} \) are given by:
\[
D_{f} = \left[ \frac{\partial}{\partial \tau} a(\tau), 1^T \otimes a(\tau) \right], \quad D_{g} = [\hat{a}(\theta, \varphi), \frac{\partial}{\partial \theta} \hat{a}(\theta, \varphi), \frac{\partial}{\partial \varphi} \hat{a}(\theta, \varphi), \hat{a}(\theta, \varphi) \otimes 1^T], \quad (5a)
\]
\[
D_{g} = [\tilde{a}(\theta, \varphi), \frac{\partial}{\partial \theta} \tilde{a}(\theta, \varphi), \frac{\partial}{\partial \varphi} \tilde{a}(\theta, \varphi), \tilde{a}(\theta, \varphi) \otimes 1^T]. \quad (5b)
\]
where \( \tilde{a} \) denotes a vector of \( 1 \)'s with appropriate dimensions.

Note that the FIM expression in (4) assumes a noise covariance matrix given by \( R_n = R_f \otimes R_g \). This is a common assumption in channel modeling [17]. It is also important to note that \( R_n \) is unknown in practice and needs to be estimated, in addition to the LoS-parameters. However, the corresponding CRB block related to the LoS-parameters is still given by (4).

IV. Proposed EKF for Joint Tracking of User Position and Clock Offset

The EKF is a non-linear extension of the popular Kalman filter that iteratively estimates the state of a dynamic system. In this
section, we first propose the joint DoA/ToF EKF and then shortly discuss how to properly initialize the EKF for good tracking performance.

A. EKF Iterations

Within the joint DoA/ToF EKF, we estimate the UN position along with the clock offset and clock skew. Therefore, we obtain a state vector \( s[n] = [x[n], y[n], \nu_x[n], \nu_y[n], \rho[n], \alpha[n]]^T \) that evolves according to the state transition

\[
F = \begin{bmatrix}
I_2 & T \cdot I_2 & 0_2 \\
0_2 & I_2 & \rho_2 \\
0_2 & 0_2 & \beta
\end{bmatrix}, \quad F_c = \begin{bmatrix} 1 & 0 \\
0 & \beta
\end{bmatrix}
\]

with \( w[n] \sim N(0, Q) \), \( Q = \text{diag}(0, 0, \sigma_v^2, \sigma_v^2, 0, \sigma_\alpha^2) \). From (7), we can see that (6) consists of two decoupled parts. On the one hand, we have the part that corresponds to the movement of the user, i.e., position and velocity. This part originates from a conventional movement model as described in, e.g., [19, p. 459]. On the other hand, we have the part that describes the evolution of the clock offset and skew according to (1) and (2), respectively. Note that both (1) and (2) have been shown to be suitable for clock tracking in [9] using practical measurements. Unfortunately, the authors of [9] do not state values for \( \beta \) as determined in their experiments. However, they argue that the clock skew is quasi-stationary, even for very long measurement periods (1.5 months) such that we can calculate \( \beta = \sqrt{(\sigma_v^2 - \sigma_\alpha^2)/\sigma_\alpha^2} \), which becomes \( \beta \approx 0.98 \) for the values of \( \sigma_v^2 \) and \( \sigma_\alpha^2 \) given in [9]. In general, and although the authors of [9] propose some techniques to estimate \( \beta \), it is rather difficult for a network to obtain those estimates. Therefore, we set \( \beta = 1 \) within the EKF. However, according to our observations, the joint DoA/ToF EKF is not very sensitive to mismatches between \( \beta \) and the actual \( \beta \).

For every time step \( n \), denote \( L[n] \) as the number of ANs with a LoS to the UN and \( \ell = \{l_1, l_2, \ldots, l_{L[n]} \} \) as the indices of those ANs. At an individual AN \( k \in \ell \), the measurement equation then consists of two parts, i.e., the DoA estimate \( \hat{\varphi}_k[n] = \varphi_k[n] + \delta \hat{\varphi}_k[n] \) as well as the ToA estimate \( \hat{\tau}_k[n] = \tau_k[n] + \delta \hat{\tau}_k[n] \) with estimation errors \( \delta \hat{\varphi}_k[n] \) and \( \delta \hat{\tau}_k[n] \). We can hence write in short \( y_k = [\hat{\varphi}_k[n], \hat{\tau}_k[n]]^T = h_k(s[n]) + u[n] \), where \( u_k = [\delta \hat{\varphi}_k, \delta \hat{\tau}_k]^T \) is the estimation error with covariance \( R_k = E[u_k u_k^T] \) and \( h_k(s[n]) = [h_{k,1}(s[n]), h_{k,2}(s[n])]^T \). The vector function \( h_k : \mathbb{R}^6 \to \mathbb{R}^2 \) relates the measurement vector \( y_k \) to the UN state through the nonlinear equations

\[
h_{k,1}(s[n]) = \arctan\left(\frac{\Delta y_k[n]}{\Delta x_k[n]}\right)
\]

\[
h_{k,2}(s[n]) = \frac{d_k[n]}{c} + \rho[n]
\]

with \( \Delta x_k[n] = x[n] - x_k, \Delta y_k[n] = y[n] - y_k, \) \( d_k[n] = \sqrt{\Delta x_k[n]^2 + \Delta y_k[n]^2} \) and the speed of light \( c \). Finally, we write the complete measurement equation at time-step \( n \) by combining the \( y_k \) from all \( L[n] \) LoS-ANs into the \( L[n] \times 1 \) vector

\[
y[n] = h(s[n]) + u[n]
\]

where \( y = [y_T^1, y_T^2, \ldots, y_T^{L[n]}]^T \), \( h = [h_T^1, h_T^2, \ldots, h_T^{L[n]}]^T \) and \( u \sim N(0, R) \) with an \( L[n] \times L[n] \) block diagonal covariance matrix \( R = \text{blkdiag}(R_{l_1}, R_{l_2}, \ldots, R_{l_{L[n]}}) \).

We are now ready to state the well known EKF equations. In this paper, we use the same notation as in [20] where \( s^{-}[n] \) denotes an \textit{a priori} state estimate, i.e., an estimate obtained from the measurements up to but not including \( y[n] \), while \( s^{+}[n] \) denotes an \textit{a posteriori} estimate, i.e., an estimate obtained from measurements up to and including \( y[n] \). With this notation we can write the \textit{a priori} estimates of the state and its covariance at time-step \( n \) as

\[
s^{-}[n] = Fs^{-}[n-1] + w[n],
\]

\[
P^{-}[n] = FP^{-}[n-1]F^T + Q[n]
\]

while the \textit{a posteriori} estimates can be written as

\[
K[n] = P^{-}[n](H[n]P^{-}[n]H^T[n] + R[n])^{-1},
\]

\[
s^{+}[n] = s^{-}[n] + K[n](y[n] - h(s^{-}[n])),
\]

\[
P^{+}[n] = (I - K[n]H[n])P^{-}[n],
\]

In (13-15) we use the Jacobian matrix \( H = \frac{\partial h_k}{\partial s[n]} \) evaluated at \( s^{-}[n] \). For the joint DoA/ToF EKF, the elements of \( H \) become

\[
H_{2k-1 \cdot 1}[n] = [h_{k,1}, y_{k,1}]_{y[n] - h(s^{-}[n])},
\]

\[
H_{2k \cdot 1}[n] = [h_{k,2}, y_{k,2}]_{y[n] - h(s^{-}[n])},
\]

\[
H_{2k+1 \cdot 1}[n] = [h_{k,2}, y_{k,2}]_{y[n] - h(s^{-}[n])},
\]

It is then straightforward to show that

\[
h_{k,1}(s[n]) = \frac{\Delta y_k[n]}{d_k[n]} \quad \Rightarrow \quad [h_{k,1}, y_{k,1}]_{y[n] - h(s^{-}[n])} = \frac{\Delta x_k[n]}{d_k[n]} \quad \Rightarrow \quad [h_{k,2}, y_{k,2}]_{y[n] - h(s^{-}[n])} = \frac{\Delta y_k[n]}{d_k[n]}
\]

The UN position estimate at time step \( n \) is finally obtained as \( (s^{+}[n], \hat{\beta}_{s^{+}[n]}[n]) \) with an estimated covariance found as the upper-left-most \( 2 \times 2 \) submatrix of \( P^{+}[n] \). An estimate of the clock offset is given through \( \hat{\beta}_{s^{+}[n]}[n] \).

B. EKF Initialization

In general, the initialization, i.e., the choice of \( s^{+}[0] \) and \( P^{+}[0] \) is very important for the performance of an EKF. In the worst case, bad choices of \( s^{+}[0] \) might lead to divergence in the EKF. Fortunately, in ultra dense networks divergence of the UN position estimates is relatively easily noticed by checking whether the obtained estimates match with the locations of the ANs that can hear the UN. Nevertheless, we would obviously like to avoid divergence in the first place. Therefore, in the following we will briefly discuss possible initialization methods.

For the initialization of the UN position estimates, we can start to one of the many RSS/DoA/time difference of arrival (TDDoA)-based positioning techniques that are available in the literature. If available, the UN could even communicate a position estimate that it obtained itself using, e.g., GNSSs. It is important though that the chosen technique provides not only position estimates but also an estimate of the position covariance such that the respective elements in both \( s^{+}[0] \) and \( P^{+}[0] \) can be initialized. Unless the chosen positioning technique also provides

\[
\begin{align*}
\alpha[n] & = \frac{\Delta x_k[n]}{c} + \rho[n], \\
\beta[n] & = \frac{\Delta y_k[n]}{d_k[n]}, \\
\end{align*}
\]
a velocity estimate, the EKF can be initialized with a very coarse estimate of the average velocity that is easily obtained for a given location considering, e.g., speed limits and available means of transportation.

The UN clock offset can be limited to fairly low values by simply communicating the time from one of the LoS-ANs. Upon arrival at the UN, the communicated time can be used to set the time within the UN. Thereafter, the UN clock offset is determined by the transmission and reception delays occurring within the involved AN and UN, respectively, as well as the signal propagation time.

Manufactures typically report the clock skew of their oscillators in parts per million (ppm). An oscillator with a specification of 20 ppm, as an example, can be expected to result in a maximum clock offset of \( \pm 20 \mu s \) per 1 s of runtime. However, based on the results in [9], [15], [16] positive clock skews seem to be much more common than negative ones. In [15], the clock skews of 69 desktop computers have been estimated using transmission control protocol (TCP) timestamps over a period of 38 days. Out of the 69 machines, only two had slightly negative clock skews (\( > 6 \) ppm), while the average clock skew was about 21 ppm with a standard deviation (STD) of about 12 ppm. Similar results can be found in [16], where the clock skews of five smartphones have been estimated using internet control message protocol (ICMP) timestamps. Again, only a single smartphone had a slightly negative clock skew (\( \approx 3 \) ppm), while the average clock skew and STD were about 29 ppm and 34 ppm, respectively. Finally, in [9] the average clock skews of a low-powered microcontroller in combination with a 16 MHz and a 32.768 kHz clock were determined to be approximately 78 ppm and 39 ppm, respectively. Therefore, as a very general rule, we could initialize the clock skew to \( \hat{\alpha}[0] = 25 \) ppm with an STD of a few 10 ppm.

V. Numerical Evaluations and Analysis

In our evaluation, we first use the geometry-based stochastic METIS channel model [11] to determine the CRB on DoA/ToA estimation. The obtained CRB then serves as the basis to simulate the performance of the proposed DoA/ToA EKF. This approach greatly reduces the overall simulation run-time and moreover also simplifies the repeatability of our results.

A. DoA and Delay Estimation

The METIS geometry-based stochastic channel model (GSCM) stems from the WINNER+ channel model [21]. Our focus has been on the 3D urban micro (Umi) propagation scenario. In particular, the maximum number of clusters in the 3D UMi scenario is 12 for LoS and 19 for a NLoS condition. The clusters that are at least 25 dB weaker than the cluster with largest power are discarded. Each cluster comprises 20 propagation paths. Each propagation path is characterized by a power, delay, elevation angle, azimuth angle, complex-path weights and a cross-polarization ratio. Details regarding the calculation of the aforementioned parameters of the propagation paths may be found in [11], in addition to the corresponding probability distributions. The LoS propagation path is also included in case of a LoS condition. The Ricean K-factor is used to weight the LoS propagation path as well as the remaining clusters accordingly.

The parameters of the LoS propagation path, namely the elevation angle, azimuth angle, delay and path-weights, follow from the distance between the AN and UN as well as their relative locations.

For the beacons transmitted by the UNs we assume an OFDM waveform with a bandwidth of \( B = 200 \) MHz and a total number of \( N_{s,m} = 640 \) subcarriers. However, we exploit only \( N_i = 64 \) subcarriers for positioning purposes (1 every 10). The receiver noise in the ANs is modeled as iid circular symmetric complex Gaussian thermal noise with a variance of \( kT_B = -119 \) dBm per subcarrier, where \( T = 295 \) K, \( k \) is the Boltzmann’s constant, and \( B_s = B/N_{s,m} \). The radio frequency interference (RFI) at an AN is assumed to stem from UNs outside of the AN’s coordination area which we model as a circle of radius 250 m around the AN. In a separate simulation study (not included here due to lack of space), we have determined the RFI by placing the interfering UNs on a ring around the AN with outer and inner radius of 250 m and 500 m, respectively. The placement of the interfering UNs was according to a non-homogeneous Poisson process leading to an overall density of about 1000 UNs per 1 km². This value is according to the car density recommended by METIS [10]. Next, we assumed that the interfering UNs transmit an OFDM-based signal with 640 active subcarriers and a transmit power of 23 dBm. According to our simulation results, the overall interference at an AN can then be well approximated as normally distributed with a variance of \( \approx 86 \) dBm per subcarrier. Finally, the RFI is modeled as a spatially white process. Such an assumption significantly reduces the complexity of the simulations and it is in agreement with the recommendations in [22]. Finally, we have calculated the transmit power of the UNs to be positioned such that a LoS-AN 60 m away from the UN receives the positioning beacons with a signal-to-interference-plus-noise ratio (SINR) of 15 dB. This results in a UN transmit power of about 3 dBm.

Using the aforementioned setup, we have determined the CRB on the STD of DoA/ToA estimation as depicted in Fig. 2. The depicted results were obtained by numerically averaging over the CRB for \( 3 \cdot 10^5 \) channel realizations drawn randomly according to the METIS channel model [11]. The CRB for the METIS channel model has been obtained by fitting the multipath-components of the channel to the covariance matrix in (4). We have observed that in rare cases channel realizations occur which lead to a CRB on DoA estimates much larger than \((180°)^2\). This is a limitation of the CRB in use, which was derived neglecting the fact that DoAs are circular values defined on an interval of 360°. In order to avoid that the average CRB is dominated by these rare but very large and faulty values, we have removed the corresponding channel realizations. In a practical system we would also recommend to remove DoA/ToA estimates that do not match our earlier observations in any way. Thus, we believe that the removal of such channel realizations is in fact a very practical approach to solve the shortcomings of the CRB in use.

As expected, the CRB on ToA estimation increases monotonically with the UN-AN distance. This is caused by the propagation path loss that increases with the UN-AN distance. Obviously, an increasing path loss also results in a reduction of the UN SINR at the AN. This, in turn, leads to the observed worsening of the ToA
estimates. However, for the CRB on DoA estimation we observe a slightly different behavior. In fact, the CRB on DoA estimation decreases up to around 10 m and only thereafter, we observe the same monotonic increase with the UN-AN distance. While this initial decrease of the CRB might seem counter-intuitive at first, it is explained by the geometry of our estimation problem. According to our assumptions, the AN’s planar antenna array is mounted in a horizontal plane 8.5 m above the UN. Now if a UN is located directly underneath the AN azimuth DoA estimation is in fact unidentifiable with the assumed array. The conditions for azimuth DoA estimation only improve when the UN is moving away from the AN. In theory, the best performance is achieved when both the AN and UN are approximately co-planar. However, this also means that the UN is very far from the AN, which as discussed above significantly reduces the SINR. Due to these counteracting mechanisms we thus observe the lowest value for the CRB on DoA estimation at around 10 m. Finally, the CRB also reveals that DoA and ToA estimation errors are statistically independent since the off-diagonal elements of the CRB are practically zero (even in comparison to the small values of the CRB on ToA estimation).

B. User Position and Clock Offset Tracking

In this section, the performance of the proposed joint DoA/ToA-EKF is evaluated by tracking a UN moving through an urban environment similar to METIS’ Madrid model [10]. In accordance with the Madrid model, each block has dimensions of 120 m × 120 m and the street width is set to 12 m. The density of ANs is assumed to be 60 m, which is well in line with the assumptions in [5]. Overall, our street model consists of 4 × 4 blocks, resulting in the map depicted in Fig. 3a.

![Fig. 3. Maps used to test the tracking performance of the proposed joint DoA/ToA-EKF.](image)

A new UN is randomly and uniformly placed at the start of the N, E, S or W ends of either of the streets. When the UN approaches an intersection, its route continues in a randomly and uniformly chosen direction excluding the direction the UN is coming from. The route ends when the UN moves out of the map or when it has crossed a maximum of 6 intersections. For the sake of simplicity, it is assumed that the UN is moving in the middle of the street and with a constant velocity \( v = \sqrt{v_x^2 + v_y^2} = 15 \text{ km/h} \). Whenever a new UN is placed on the map, its clock offset and clock skew are randomly initialized according to \( \rho[0] \sim \mathcal{N}(0, \sigma_{\rho,0}^2), \sigma_{\rho,0} = 100 \mu \text{s} \) and \( \alpha[0] \sim \mathcal{N}(\mu_{\alpha,0}, \sigma_{\alpha,0}^2), \mu_{\alpha,0} = 25 \text{ ppm}, \sigma_{\alpha,0} = 30 \text{ ppm} \), respectively, as motivated in Section IV-B. Based on the measurement results in [9], we set the STD of the clock skew driving noise to \( \sigma_\eta = 6.3 \cdot 10^{-9} \).

We assume that an initial estimate \( \tilde{p}[0] = (\tilde{x}[0], \tilde{y}[0])^T \) of the UN position \( p[0] \) is available through, e.g., GNSS where \( \tilde{p}[0] \sim \mathcal{N}(p[0], \sigma_{p,0}^2 I_2) \), \( \sigma_{p,0} = 5 \text{ m} \). For each new UN, we then initialize the joint DoA/ToA-EKF as well as the classical DoA-only EKF [7] that serves as our positioning benchmark. The state vector and covariance of the joint DoA/ToA-EKF are initialized as \( \tilde{s}^+[0] = (\tilde{p}^T[0], 0, 0, 0, \mu_{\alpha,0})^T \) and \( P^+[0] = \text{diag}(\sigma_{p,0}^2, \sigma_{\rho,0}^2, \sigma_{\alpha,0}^2, \sigma_{\rho,0}^2, \sigma_{\alpha,0}^2, \sigma_{\alpha,0}^2, \sigma_{\alpha,0}^2) \), \( \sigma_{\alpha,0} = 5 \text{ m} \), respectively. The same values are then also used for the initialization of the DoA-only EKF. Thereafter, the EKFs are updated in an interval of \( T = N_T T_i \), where \( T_i = 167.3 \mu \text{s} \) is the duration of the 5G radio frame proposed in [5] and \( N_T \) is a positive integer. At every time-step \( n \), we assume that the two closest ANs not obscured by building blocks are in LoS with the UN. These LoS-ANs then produce a DoA and ToA estimate according to (10) using the CRBs obtained in the previous section. Thereby, the block matrix \( R_l \), of covariance \( \mathbf{R} \) becomes the CRB that was obtained for the distance closest to the actual distance between AN \( l \) and the UN. Within the EKFs, we have tuned the STD of the driving noises to \( \sigma_v = 0.1 \text{ m/s} \) and \( \sigma_\eta = 10^{-4} \). The latter value is much larger than the actual value, but leads to a much improved overall performance. Based on our observation, very small \( \sigma_\eta \) lead to a divergence in the joint DoA/ToA-EKF whenever the clock offset and clock skew estimates are very inaccurate as, e.g., the case in the initial tracking phase.

Table I summarizes the tracking performance of the proposed joint DoA/ToA-EKF in comparison to the classical DoA-only EKF. These results were obtained by averaging over 103 different UN routes, each with individual realizations of DoA/ToA estimation errors. In order to avoid that the tracking root-mean-squared errors (RMSEs) are dominated by the initial state estimates, we have excluded the first 20 EKF iterations in the RMSE calculation. Considering the clock offset, for example, we notice that the tracking RMSE in Table I is significantly smaller than the STD of the initial estimate (100 μs). However, these initial errors
are not determined by the EKFs and should therefore also not be included in the tracking RMSE calculation.

As expected, we can improve the estimation accuracy by decreasing the tracking period. Overall, the results also show that the proposed joint DoA/ToA-EKF greatly improves the positioning accuracy of the classical DoA-only EKF. Depending on the tracking period, the joint DoA/ToA-EKF achieves a positioning RMSE of 0.4 – 1.0 m, which is an improvement by one order of magnitude compared to the RMSE of the DoA-only EKF. Moreover and in contrast to the DoA-only EKF, the joint DoA/ToA-EKF also achieves highly accurate UN-network synchronization with an RMSE of only 4 ns.

The performance of the DoA-only EKF suffers in particular when the geometry of the two LoS-ANs and the UN resembles a line. In such cases information about the AN-UN angles suffices only to determine that the UN is located somewhere along that line, but not exactly where. This is a known problem of DoA-only based positioning in general and it has been shown that this problem can be solved by including the RSS, i.e., a measurement of the AN-UN distance into the positioning process along with the DoA [23]. Since the ToA is also a measurement of the AN-UN distance, the proposed joint DoA/ToA-EKF therefore does not suffer from such disadvantageous geometries either. In a 5G network, this is a very important property as the network is not primarily designed for positioning such that situations with geometries disadvantageous for positioning are inevitable. In fact, we cannot even guarantee that UNs are always in LoS with multiple ANs. Therefore, we have tested the tracking capabilities of the proposed joint DoA/ToA-EKF also for a second map with extremely challenging conditions as depicted in Fig. 3b. In this map the UN is in LoS with only a single AN for most of the time. Therefore, the classical DoA-only EKF is not able to track the UN at all. The proposed joint DoA/ToA-EKF on the other hand estimates the UN position and clock offset with a RMSE of about 3.0 m and 10.3 ns, respectively for \( N_T = 100 \). By increasing the number of antennas, we should be able to further increase the positioning accuracy. This makes the proposed joint DoA/ToA-EKF also interesting for positioning with massive MIMO systems, where the number of antennas is significantly larger compared to what we have assumed in this paper. This forms one of the aspects in our future research.

The behavior of both the DoA-only EKF and the joint DoA/ToA-EKF in tracking is also illustrated in the videos that we have uploaded to www.tut.fi/5G/GLOBECOM15.

VI. CONCLUSION

In this paper we showed that 5G ultra dense networks are able to provide UN positioning with an accuracy that may even be in the sub-meter range. More specifically, we have proposed a joint DoA/ToA-EKF that tracks both the clock offset and position of the UNs within a 5G network. Using the METIS channel models that were developed specifically for 5G networks, we then demonstrated the achievable DoA/ToA estimation performance by means of the respective CRB. Finally, we used the obtained results to simulate the performance of the proposed joint DoA/ToA-EKF and showed that the joint DoA/ToA-EKF significantly outperforms existing DoA-only solutions, while at the same time obtaining highly accurate UN clock offset estimates as a by-product. Moreover, we also demonstrated that the joint DoA/ToA-EKF is able to handle situations where DoA-only solutions are not able to estimate the UN position at all.

REFERENCES