Abstract—The available geometry-based stochastic channel models (GSCMs) at millimetre-wave (mmWave) frequencies do not necessarily retain spatial consistency for simulated channels, which is essential for small cells with ultra-dense users. In this paper, we work on cluster parameterization for the COST 2100 channel model using mobile channel simulations at 61 GHz in Helsinki Airport. The paper considers a ray-tracer which has been optimized to match measurements, to obtain double-directional channels at mmWave frequencies. A joint clustering-and-tracking framework is used to determine cluster parameters for the COST 2100 channel model. The KPowerMeans algorithm and the Kalman filter are employed to identify the cluster positions and to predict and track cluster positions respectively. The results confirm that the joint clustering-and-tracking is a suitable tool for cluster identification and tracking for our ray-tracer results. The movement of cluster centroids, cluster lifetime and number of clusters per snapshot are investigated for this set of ray-tracer results. Simulation results show that the multipath components (MPCs) are grouped into clusters at mmWave frequencies.

Index terms—Cluster identification, Kalman filter, KPowerMeans, millimetre wave, multi path components.

I. INTRODUCTION

Over the past few years, an abundance of techniques have been proposed as a means to efficiently scale the wireless capacity. It remains unclear which technology or set of technologies can meet the demand. One promising set of technologies for the 5th Generation (5G) cellular network is reviewed in [1]: the combination of large antenna arrays and short wavelength carrier waves. This combination allows for a greater bandwidth availability and extremely high spectral efficiency by utilizing a large number of antennas, whilst occupying a relatively small area. This technology is known as Massive multiple-input multiple-output (MIMO) in the millimeter-wavelength (mmWave) spectrum [2].

Most standardized MIMO channel models such as IEEE 802.11 [3] and the most recent 3GPP channel model [4] rely on clustering [3]. The same applies to the recent COST channel models, e.g., the COST 2100 model [5]– [8]. These models are geometry-based stochastic channel models (GSCMs) that are mathematically tractable, though to a limited extent, to investigate the performance of MIMO systems [9]. The concept of clustering is an essential basis of GSCMs to characterize scatterers in the cell environments. In [10]–[14], the authors use clusters to characterize measured multipath channels for a GSCM in mmWave bands. The available GSCMs at mm-waves do not necessarily retain the spatial consistency of simulated channels due to lack of cluster dynamics, which is essential for small cells with ultra-dense users. In this paper, we work on cluster parameterization to investigate the spatial consistency, using a ray-tracer which is adjusted to produce results consistent with measurements.

Unlike previously available clustering algorithms, in this paper the coordinates are exploited for which the multipath components (MPCs) interact with surrounding objects for a fixed position of mobile station (MS) and base station (BS). To the best of our knowledge, previously clustering has been performed in a double-directional setting, i.e., considering both angle of arrival (AoA) and angle of departure (AoD). A consistent scheme to identify and track clusters based on the spatial coordinates of the MPCs (the \([x, y, z]\)-coordinates of the MPCs) is presented. To investigate the performance of the proposed clustering scheme we exploit a set of ray-tracer results in Helsinki’s airport described in [15], which is very accurate to present the propagation properties such as specular reflections, diffraction, diffuse scattering [16]. The contributions of the paper are summarized as follows:

1. We study whether clusters exist or not.
2. For the first time, we perform clustering of dynamic multipath channels.
3. \([x, y, z]\) coordinate-based clustering.

A. Outline

The rest of the paper is organized as follows. Section II describes the ray-tracer and simulation area, and Section III provides the MPC clustering-and-tracking framework. The simulation results and discussion are
presented in Section IV while Section V concludes the paper.

B. Notation

The following notations are adopted in the rest of the paper. Uppercase and lowercase boldface letters are used for matrices and vectors, respectively. The notation $|x|$, $|X|_{\text{det}}$ and $|X|_{\text{size}}$ stand for the absolute value of $x$, determinant of matrix $X$, and the size of vector $x$, respectively. $X^{-1}$ and $X^T$ denote the inverse and transpose of matrix $X$, respectively. Moreover, $\mathbf{I}_n$ introduces identity matrix with size $n \times n$. The Kronecker product of $X$ and $Y$ is presented by $X \otimes Y$.

II. THE RAY-TRACER AND SIMULATION AREA

The in-house ray-tracer simulates multipath channels for a large number of links between BS and MS [15]. Note that our ray-tracer works with accurate descriptions of the environment in the form of point clouds, obtained by laser scanning, and has the ability of simulating relevant propagation properties such as specular reflections, diffraction, diffuse scattering and shadowing [16]. For more details on our ray-tracer refer to [15], [16]. A check-in hall of Helsinki airport as a representative small-cell scenario is considered as shown in Fig. 1. Exploiting the ray-tracer parameters in Fig. 1 we obtain the MPCs for links defined by BS and MS locations as in Fig. 1. The BS is located 1 m from a wall at a height of 5.7 m whereas the MS is placed at a height of 1.5 m at every 5 cm over a route. In total, 2639 links including 1816 line-of-sight (LOS) and 823 obstructed LOS (OLOS) are simulated.

In Fig. 1, we obtain the MPCs for links defined by BS and MS locations as in Fig. 1. The BS is located 1 m from a wall at a height of 5.7 m whereas the MS is placed at a height of 1.5 m at every 5 cm over a route. In total, 2639 links including 1816 line-of-sight (LOS) and 823 obstructed LOS (OLOS) are simulated.

As the ray-tracer calculates interactions of MPCs with physical objects in the environments, we save the first and last MPC interacting coordinates $[x, y, z]$ instead of the angle of departure and arrival of each MPC. We assume downlink where BS transmits and MS receives radio signals. The first and last interacting coordinates are the same for a single-bounce path, and are different for a multiple-bounce path. The ray-tracer also derives a complex gain for each MPC.

III. CLUSTERING-AND-TRACKING FRAMEWORK

Similar to standard clustering algorithms [17], [18], we independently perform clustering at each snapshot and thereafter the clusters are tracked. Consider $n = 1, \cdots, N$ data windows, where at each data window we have $L^{(n)}$ MPCs. Next, we define for each MPC $\mathbf{v}^{(n)}_{1,l} = [x^{(n)}_{M,S,l}, y^{(n)}_{M,S,l}, z^{(n)}_{M,S,l}]$ (the position of MPCs from MS side) and $\mathbf{v}^{(n)}_{2,l} = [x^{(n)}_{B,S,l}, y^{(n)}_{B,S,l}, z^{(n)}_{B,S,l}]$ (the position of MPCs from BS side), and finally we have

$$\chi_l^{(n)} = [\mathbf{v}^{(n)}_{1,l}, \mathbf{v}^{(n)}_{2,l}, \chi_l^{(n)}_{M,S,l}, \chi_l^{(n)}_{B,S,l}].$$

The same equality hold for the BS-side components. This enables us after visualising clusters to plot clusters separately for $\mathbf{v}^{(n)}_{1,l}$ and $\mathbf{v}^{(n)}_{2,l}$ in physical three-dimensional space as well as defining the matrix $\chi^{(n)} = [\chi_1^{(n)}, \cdots, \chi_L^{(n)}]$. Moreover, the $l$th MPC in window $n$ has a power represented by $p_l^{(n)}$ which enables us to define the power vector $\mathbf{p}^{(n)} = [p_1^{(n)}, \cdots, p_L^{(n)}]$.

A. Cluster Parameters

In next step, we define the following parameters for each cluster:

1. Cluster ID $c$.
2. Cluster power at time $n$: $\gamma_c^{(n)} = \sum_{l \in I_c^{(n)}} p_l^{(n)}$, where $I_c^{(n)}$ denotes the set of MPCs belonging to cluster $c$ at time $n$.
3. Total number of MPCs in cluster $c$ at time $n$: $L_c^{(n)} = |I_c^{(n)}|_{\text{size}}$.
4. Cluster centroid position:

$$\mu_c^{(n)} = \left[\sum_{l \in I_c^{(n)}} p_l^{(n)} x_{M,S,l}^{(n)}, \sum_{l \in I_c^{(n)}} p_l^{(n)} y_{M,S,l}^{(n)}, \sum_{l \in I_c^{(n)}} p_l^{(n)} z_{M,S,l}^{(n)}\right]^T$$

5. Combined cluster centroid position and speed:

$$\theta_c^{(n)} = \left[\mu_c^{(n)}, \Delta x_{M,S,c}, \Delta y_{M,S,c}, \Delta z_{M,S,c}\right]^T$$

6. Cluster spread matrix:

$$\mathbf{C}_c^{(n)} = \sum_{l \in I_c^{(n)}} \frac{1}{\gamma_c^{(n)}} \left(\chi_l^{(n)} - \mu_c^{(n)}\right)\left(\chi_l^{(n)} - \mu_c^{(n)}\right)^T$$

Next, similar to terminology in [17], a Kalman filter [19] is used to both track and predict the cluster positions over time. Moreover, an initial-guess process introduces an appropriate initial guess for cluster centroids, and finally the clustering algorithm determines
the clusters in the ray-tracer results exploiting the initial guess.

B. Kalman Filter to Track and Predict Cluster Positions

We exploit the cluster centroid positions and cluster centroid speeds for the Kalman tracking [19]. The following state equations are used:

\[
\begin{align*}
\theta_c^{(n)} &= A\theta_c^{(n-1)} + B^{(n)}, \\
\Phi &= I_3 \otimes \begin{bmatrix} 1 & 0 \\
0 & 1 \end{bmatrix}, \\
\mu_c^{(n)} &= D\theta_c^{(n)} + E^{(n)}, \\
D &= I_3 \otimes \begin{bmatrix} 1 & 0 \end{bmatrix},
\end{align*}
\]

where \(B^{(n)}\) and \(E^{(n)}\) refer to the state-noise with covariance matrix \(Q\) and the observation-noise with covariance matrix \(R\), respectively. Note that \(\mu_c^{(n)}\) introduces the observed cluster centroid position. The prediction and update equations are given by

Prediction:
\[
\begin{align*}
\theta_c^{(n|n-1)} &= A\theta_c^{(n-1|n-1)}, \\
\Phi^{(n|n-1)} &= A\Phi^{(n-1|n-1)} + Q
\end{align*}
\]

and update:
\[
\begin{align*}
K^{(n|n)} &= \Phi^{(n|n-1)}D^T \left( DM^{(n-1)}D^T + R \right)^{-1}, \\
\theta_c^{(n|n)} &= \theta_c^{(n|n-1)} + K^{(n|n)} \left( \mu_c - D\theta_c^{(n|n-1)} \right), \\
M^{(n|n)} &= \left( I - K^{(n|n)}D \right) M^{(n|n-1)}
\end{align*}
\]

C. Association of Clusters

Association of predicted targets to identified targets is a substantial challenge in any multi-target tracking [17]. Based on [17], the distance between a cluster with parameters \((\mu_c, C_c)\) and a cluster with centroid \(\tilde{\mu}\) is called the closeness function and is given by

\[
d_c(\mu|\mu_c, C_c) = \frac{1}{(2\pi)^{\frac{N}{2}} |C_c|^{\frac{1}{2}}} \exp \left( -\frac{1}{2} \left( \mu - \mu_c \right)^T C_c^{-1} \left( \mu - \mu_c \right) \right).
\]

First, the closeness function between the old clusters (with the old covariance matrix) and new centroids and the closeness function between the new clusters (with the old covariance matrix) and old centroids are calculated. Next, for each new cluster the closest old cluster and for each old cluster the closest new cluster is determined. Note that the closest cluster is determined by finding the maximum value of the closeness function. If the closeness function from both directions are exactly the same, these two clusters are associated and assumed to be one cluster. The clusters which are not assumed are assumed to be new ones.

D. Initial Guess for Clusters

The initial guess of the cluster centroids is a challenging task in clustering algorithms. In [17], the authors propose a novel initial guess to maximize the distances between the cluster centroids. If there is no cluster prediction available, the path having the strongest power is selected as the first centroid \(\mu_1\), whereas for the case of available cluster prediction, the initial-guess centroid from the prediction is to be as the current initial guess. Note that the multipath component distance (MCD) in this paper is different from the one used in [17], [20]. The distance measure between MPCs \(i\) and \(j\) is given by

\[
\text{MCD}_{ij} = \sqrt{||\text{MCD}_{x_{MS,i,j}}||^2 + ||\text{MCD}_{y_{MS,i,j}}||^2 + ||\text{MCD}_{z_{MS,i,j}}||^2}.
\]

Note that in (10) we have

\[
\text{MCD}_{x_{MS,i,j}} = \frac{|x_{MS,i} - x_{MS,j}|}{\Delta x_{MS,max}},
\]

where \(\Delta x_{MS,max} = \max \{ |x_{MS,i} - x_{MS,j}| \}\), and the other terms in (10) are evaluated in a similar way to (10). Next, the weighted distance matrix \(\Upsilon \in C^{l \times c}\) between all paths and all initial-guess centroids is evaluated as follows:

\[
\Upsilon (\chi^n_l - \hat{\mu}_c) = \log_{10} \left( p_{l}^{(n)} \right) \text{MCD} (\chi^n_l - \hat{\mu}_c)
\]

Following the terminology in [17], we select the path with the maximum minimum distance to any centroid as follows:

\[
l_{\text{sel}} = \max_l \left\{ \min_c \{ \Upsilon \} \right\}.
\]

We then assign all MPCs to their closest centroid and cluster power is evaluated. If we do not achieve the maximum number of clusters, and centroid powers are larger than 0.01% of the total snapshot power, we repeat the calculation of the weighted distance matrix \(\Upsilon \in C^{l \times c}\) in (11). Otherwise, the last centroid is ignored and the algorithm is stopped.

E. Clustering Algorithm

The KPowerMeans clustering algorithm is investigated in [21], and it performs as follows: the initial-guess algorithm is applied, and the KPowerMeans clustering algorithm is run only once as the initial guess as are constant. For more details on the KPowerMeans clustering algorithm refer to [21]. Note that if any cluster occupies less than 1% of total cluster power, we re-start the clustering algorithm with the initial guess, with the number of clusters is reduced by one. Therefore, it is possible that the algorithm ends with a single cluster.
IV. RESULTS AND DISCUSSION

The joint clustering-and-tracking algorithm is applied to the ray-tracer results at Helsinki airport, explained in Section II where we have 2639 links. Figs. 2 and 6 present the exemplary plots for different snapshots. The MPCs are shown by dots, where their power is shown by light blue (weak power) and violet (strong power). The clusters are shown by ellipsoids and always 99% of the total power is carried by the MPCs within clusters. We use different colors for ellipsoids just to make the cluster recognition easier. Each cluster is identified by a cluster ID which is written on each cluster. As these exemplary figures show for snapshots 2, 3 and 4, cluster 2 is always tracked while the other clusters are determined as new clusters.

Next, the lifetime of clusters for the available sets of ray-tracer results is investigated, for Tx-side clusters and Rx-side clusters separately. Figs. 7 and 8 show the histograms of cluster lifetimes for Rx-side (BS-side) and Tx-side (MS-side) scenarios, respectively. The figures show that in most cases clusters are active only for a few snapshots for this set of ray-tracer results. This requires more investigation. Moreover, the number of clusters per snapshot is presented in Figs. 9 and 10 for Rx-side and Tx-side clusters, respectively.

The other interesting phenomenon is the movement of the tracked cluster centroids, which is shown in Fig. 11. Based on these figures the cluster centroids moves rapidly in the $x$ or $y$ direction while its speed is very low in other direction. Moreover, the figure show for these clusters that the centroid’s speed is very low in the $z$ direction. Finally, Figs. 12 and 13 investigate the...
distribution of the percentage of power in Tx-side and Rx-side clusters.

V. CONCLUSIONS

In this paper, we have worked on parameterization for the COST 2100 channel model at 60 GHz band. We have worked on a ray-tracer, which has been optimized to match measurements, to get double-directional channels at mmWaves. We have combined clustering and tracking to improve the performance of consistent clustering. The results showed that the joint clustering-and-tracking allows for cluster identification and tracking for the ray-tracer results. Cluster lifetime and number of clusters per snapshot have been investigated.
REFERENCES

[1] E. Björnson, E. A. Jorswieck, M. Debbah, and B. Ottersten, “Multijobjective signal processing optimization: the way to balance conflicting metrics in 5G systems,” IEEE Signal Process. Mag., vol. 31, no. 6, pp. 14–23, Oct. 2014.

[2] A. Adhikary, E. A. Safadi, M. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, “Joint spatial division and multiplexing for mm-wave channels,” IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1259–1255, Jun. 2014.

[3] A. F. Molisch and F. Tufvesson, “Propagation channel models for next-generation wireless communications systems,” IEEE Trans. Veh. Technol., vol. E97-B, no. 10, pp. 2022–2034, Oct. 2014.

[4] LTE. Study on channel model for frequency spectrum above 6 GHz. 3GPP Specification 38.900, May 2016.

[5] L. Liu, J. Poutanen, F. Quitin, K. Haneda, F. Tufvesson, P. D. Doncker, P. Vainikainen, and C. Oestges, “The COST 2100 MIMO channel model,” IEEE Wireless Commun., vol. 19, no. 6, pp. 92–99, Dec. 2012.

[6] M. Bashar, A. Burr, K. Haneda, and K. Cumanan, “Robust user scheduling with COST 2100 channel model for Massive MIMO networks,” IET Microwaves, Antennas Propag., vol. 12, no. 11, pp. 1751–8725, Aug. 2018.

[7] M. Bashar, A. G. Burr, D. Maryopli, K. Haneda, and K. Cumanan, “Robust geometry-based user scheduling for large MIMO systems under realistic channel conditions,” in Proc. IEEE EW, May 2018, pp. 1–6.

[8] M. Bashar, A. Burr, K. Haneda, and K. Cumanan, “Evaluation of low complexity Massive MIMO techniques under realistic channel conditions,” IEEE Trans. Veh. Technol., Submitted.

[9] M. K. Samimi and T. S. Rappaport, “Statistical channel model with multi frequency and arbitrary antenna beamwidth for millimetre-wave outdoor communications,” in Proc. IEEE Globe com, Dec. 2015.

[10] A. Karttunen, J. Jarvelainen, A. Khattun, and K. Haneda, “Radio propagation measurements and WINNER II parametrization for a shopping mall at 61-65 GHz,” in Proc. IEEE VTC, May 2015, pp. 1–6.

[11] K. Haneda., L. Tian, H. Asplund, J. Li, Y. Wnag, D. Steer, C. Li, T. Balercia, S. Lee, Y.-S. Kim, A. Ghosh, T. Tomas, N. Nakamura, Y. Kakishima, T. Imai, H. Papadopoulis, T. S. Rappaport, G.-R. McCartney, M. K. Samimi, S. Sun, O. Koyimen, S. Hur, J. Park, J. Zhang, E. Mellios, A. F. Molisch, S. S. Ghassamzadeh, and A. Ghosh, “Indoor 5G 3GPP-like channel models for office and shopping mall environments,” in Proc. IEEE ICC Workshop, May 2016, pp. 694–699.

[12] C. Schneider, J. Geddschold, M. Kaske, R. S. Thoma, and G. D. Galdo, “Estimation and characterization of multipath clusters in urban scenarios,” in Proc. IEEE EuCAP, Apr. 2018, pp. 1–5.

[13] N. Iqbal, D. Dupleich, C. Schneider, J. Luo, R. Muller, S. Hafner, G. D. Galdo, and R. S. Thoma, “Time modeling of intra-cluster multipaths for 60 GHz fading channels,” in Proc. IEEE EuCAP, Apr. 2018, pp. 1–5.

[14] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, “On mm-wave multipath clustering and channel modeling,” IEEE Trans. Ant. Prop., vol. 62, no. 3, pp. 1445–1455, Mar. 2014.

[15] K. Haneda, J. Jarvelainen, and A. Karttunen, “[online]. available: https://arxiv.org/pdf/1802.08591.pdf,” in Proc. IEEE VTC, 2018, pp. 1–6.

[16] J. Jarvelainen, K. Haneda, and A. Karttunen, “Indoor propagation channel simulations at 60 ghz using point cloud data,” IEEE Trans. Ant. Prop., vol. 64, no. 8, pp. 4457–4467, Aug. 2016.

[17] N. Czink, K. Haneda, J. Zhang, E. Mellios, A. F. Molisch, S. Hafner, G. D. Galdo, and R. S. Thoma, “Time modeling of intra-cluster multipaths for 60 GHz fading channels,” in Proc. IEEE EuCAP, Apr. 2018, pp. 1–5.

[18] N. Czink, E. Bonek, L. Hentila, J. P. Nuutinen, and J. Yli-talo, “Cluster-based MIMO channel model parameters extracted from indoor time-variant measurements,” in Proc. IEEE Globe com, Nov. 2006, pp. 1–5.

[19] S. Kay, Fundamentals of statistical signal processing: Estimation theory. Englewood Cliffs, NJ, USA, Prentice-Hall, 1993.
[20] N. Czink, P. Cera, J. Salo, E. Bonek, J. P. Nuutinen, and J. Ylitalo, “Improving clustering performance using multipath component distance,” *Electronics Letters*, vol. 42, no. 1, pp. 1–2, Jan. 2006.

[21] N. Czink, P. Cera, J. Salo, E. Bonek, J. P. Nuutinen, and J. Ylitalo, “A framework for automatic clustering of parametric MIMO channel data including path powers,” in *Proc. IEEE VTC*, Sep. 2006, pp. 1–5.