IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 17 September 2023, accepted 2 October 2023, date of publication 10 October 2023, date of current version 25 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3323022

RESEARCH ARTICLE

Pattern Based Mobility Management in 5G Networks With a Game Theoretic-Jump Markov Linear System Approach

MASOTO CHIPUTA [1], (Student Member, IEEE), MINGLONG ZHANG [2], (Member, IEEE), AND PETER HAN JOO CHONG [1], (Senior Member, IEEE)
[1]Department of Electrical and Electronic Engineering, Auckland University of Technology, Auckland 1010, New Zealand
[2]Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762, USA
Corresponding author: Masoto Chiputa (masoto.chiputa@aut.ac.nz)

ABSTRACT The fifth generation (5G) mobile communication adopted the usage of Millimeter Wave (mmWave) bands to ignite prospects of gigabit data rates in mobile networks. However, mmWave propagation is highly susceptible to competing factors of user and topographic dynamics: they formulate irregular cell patterns. The irregularities in mmWave cell patterns cause unreliable connectivity and can instigate unnecessary Handoffs (HOs). This behavior ultimately increases the risk of 5G link failures. To improve mmWave link connectivity hence guarantee continuous connectivity in 5G mobile communication, this paper proposes a HO scheme that predicts target link deterioration patterns to select the most reliable mmWave link for a mobile user. The scheme is based on Game Theory (GT) and Jump Markov Linear Systems (JMLS). JMLSs are known to account for abrupt/erratic changes in system dynamic predictions. We amalgamate GT with JMLS capability to predict target mmWave link pattern/behavior after the HO execution. Specifically, given channel gain and received power variation over distance, the GT-JMLS HO scheme predicts the sustainability of the signal-interference-noise ratio (SINR) pattern of a target link above threshold. This is paramount to reducing the selection of mmWave links that prematurely fail or require multiple HOs to sustain connectivity over a short distance or period. Our simulation results show that our proposed HO scheme offers target links with higher: throughput, energy efficiency, reliability, and longer dwell time between HOs than classical HO schemes.

INDEX TERMS Millimeter wave (mmWave), handover (HO), received power, game theory, jump Markov linear systems, 5G.

I. INTRODUCTION

The millimeter wave (mmWave) bands (30–300 GHz) hold great prospects of offering gigabit data rates in mobile networks, particularly in the fifth generation (5G) mobile networks and beyond. However, their sensitivity to concrete, water, street infrastructure, humans, among other things degrades their prospects of deployment in mobile networking. For instance, mmWave path loss is around 20–25 dB greater than that in microwave. Received power between line of sight (LOS) and non-line of sight (NLOS) is also known to hugely degrade in orders of 30 dB. Further, materials such as brick/mortar, attenuate mmWave signals by as much as 40–80 dB. Human bodies also cause losses of 20–35dB. Further, the doppler effects can be over 3 kHz (e.g., at 60 km/h for 60 GHz bands) with channels changing in orders of hundreds per microsecond faster than microwaves [1]. Thus, in mmWave communication, unlike microwave, several performance deteriorating factors co-exist in competition. This mmWave propagation behavior significantly degrades network reliability in 5G. Particularly at system level, mmWave sensitivity to user mobility and topographic dynamics induces too early, unnecessary or too late Handoffs [2]. To therefore enhance mmWave link connectivity hence reliability in 5G

The associate editor coordinating the review of this manuscript and approving it for publication was Stefan Schwarz.

mobile communication and, meet the desired data rate for mobile users, several mitigation measures for mobility management have been studied and proposed. These include the use of statistical, artificial intelligence (AI) [3] and game theory (GT) techniques in Handoff (HO) schemes [4]. Others include the use of dual connectivity (DC) model, with the interoperability of the LTE and mmWave protocol [5]. However, the challenges with most of the HO techniques is that they do not always predict nor consider the coexistence/competing effects of user and topographic network performance deterioration factors. This leads to inaccurate selection of the target cells, waste of transmission power, link failure and huge system overhead. Ultimately, ineffective handling of mmWave reduces in 5G can increase overhead, and cause too early, late, or incorrect HOs.

To thus be more viable and ensure 5G mobile network connectivity is continuous, we propose an augmented use of the Jump Markov Linear Systems (JMLS) modeling and Game Theory to predict the long-term deterioration pattern of target links. Our proposed HO scheme ultimately selects target links that sustain mmWave connectivity longer than the rest. JMLS modelling is known to account for both gradual and abrupt changes [6]. We use this JMLS capability in the 5G network to learn to predict the mmWave target link's abrupt and gradual deterioration pattern and vulnerability prior to executing a HO. Technically, given the deterioration power pattern by JMLS, we trade off channel gain and power effects on signal-interference-noise ratio (SINR) values in both LoS and NLoS to improve target link behavior predictions. Simulation results show that our proposed scheme can achieve higher throughput averaging over 20% than any other scheme. Additionally, the scheme can select more reliable target links with longer connectivity time between HOs.

A. CONTRIBUTIONS

- We propose a GT-JMLS based HO model with the link selection criteria considering not just the initial performance of a target link after HO but also the long-term behavior between HO. The scheme predicts the likely abrupt and gradual changes on the target link as the user moves. This helps assess the target link's ability to not just meet the desired QoS but sustain connectivity before another HO.
- We propose a HO scheme that learns to predict the deterioration performance pattern of a link prior to HO. In this regard, we amalgamate the learning capability of JMLS and GT.
- We explore the effects of relying on received power and channel gain when selecting target links in HOs. The concept explores the tradeoffs between the channel gain and the power requirements in meeting the desired SINR in both NLOS and LOS scenarios. This is vital for link connectivity survival in vulnerable situations, e.g., NLOS scenarios. It is also vital to minimize HOs where the LOS obstruction is temporal.
- We study different user types and their respective effects on network performance. We study the impact of each user type behavior on the network performance and target link deterioration pattern.

B. RELATED WORKS

Considering multiple coexisting and competing mmWave link deteriorating factors as described earlier, various HO schemes based on tradeoffs and machine learning techniques including GT and artificial Intelligence (AI) have been proposed for 5G mobile networks [6]. In [7], for instance, a tradeoff between high mmWave Base Station (mm-BS) spatial density is done against limited front and backhaul capacity. A decentralized user-centric game theoretic energy efficiency HO system is proposed in [8] to minimize backhaul load and increase mm-BS spatial density. Another classic game theory concept in [3] is used to make interactive decision-making policies for energy and bandwidth resource optimization. The challenge however in 5G mmWave networks is that they are a lot of competing factors to be traded off against. Unlike the classical game theory concepts in previous generation networks, different groups/players face dynamic performance deteriorating factors requiring different QoS analyses at different times/place [9]. For instance, topographic effects may have higher priority in one area than user mobility effects in determining the selection of target links. However, the HO trade off parameters may ignore the impact of topographic effects or vice-versa at the time when it is needed. To that effect, current classical games with prechosen tradeoff factors like in [4], [8], [9], [10], and [11] may not always yield desirable results and inevitably may cause too early/late or wasteful Hos leading to link failures in 5G.

On the other hand, multiple computing techniques have been amalgamated to get better results and overcome the shortfalls of learning concepts for 5G mobile connectivity. For instance, Deep Reinforcement Learning (DRL) techniques have been used to simultaneously learn co-existing effects of network load and user speed in 5G networks. One such example is the simultaneous deployment and tradeoff of Mobility Robust Optimization (MRO) [13] and Mobility Load Balancing (MLB) techniques in 5G mobility management in [12]. Authors in [14] use deep Q-learning for tradeoff offloading and selecting optimal edge links. The challenge, however, compared to game theory is that DRL requires thousands of samples to learn useful policies. Additionally, DRL is terribly unstable where there is a large source of local channel variances e.g., as exhibited in mmWave [2]. In a short term, most DRL-based HO policies pose huge 5G connectivity inconsistences. Therefore, they fail to warrant longer connectivity time. To warrant better useful policies within a short training time, studies in [15], [17], [18], [20], [21], and [22] incorporate RL with JMLS and GT techniques. In [21] the particle-filter-based RL is incorporated to predict a finite number of disturbances and states within a randomly chosen sample of trajectories. The author in [22] amalgamates JMLS formulation into game theory concepts to tackle distributed decision-making scenarios. All these concepts aim to improve GT and DRL performance. We thus expound on game theory with JMLS concepts to robustly expediate the analysis of deterioration predictions in 5G links and guarantee stability in HOs.

II. PROPOSED MODEL

To find reliable target links during HOs in 5G networks for a diverse user, we consider the mmWave target link deterioration behavior/pattern for three types of users including cars, pedestrians, and cyclists. For a given velocity of a user type, channel coherence time is linear with carrier frequency and velocity [1]. Thus, given the channel state for each user type, and to sustain SINR above a given threshold, we first learn to predict the mmWave received power deterioration pattern at the user after a HO. We use JMLS modelling to model the pattern. Due to limited and sometimes fast changing channel state information (CSI), the initial values of the deterioration pattern are inferred using Expected Maximization (EM) algorithm. However, EM is intractable in instances where CSI changes are drastic [6]. Thus,it is important to understand the rapid fluctuation and intermittent connectivity in mmWave networks [4]. Particularly to understand, anticipate and cope with the dynamics of rapid mmWave target link deterioration behavior, we augment EM inference with mean field approximation [2]. Here no single set of players' CSI is used to predict the global optimum deterioration path expected on target links. Instead, the aggregative behavior of all user types predicts the payoff(s); hence, the pattern of the target link deterioration is likely to take. We trade off the effects of channel gain against energy cost (received power) in LOS and NLOS to optimize the EM's initial estimations and find the likely target links that will consistently meet the desired SINR after HOs. Technically, we incorporate deterioration prediction samples initiated by EM into a game value function. This used as initial experience by mean field games to optimize the target link received power pattern prediction during HO selection. Understanding the behavior of the received power pattern and hence limits, for instance, in NLOS reveals a wealth of information about the underlying SINR distribution as users moves towards/away from serving cells. For instance, as a HO problem in our scheme, we are given n possible values of received powers, $x_1, \ldots, x_n$, over a given range of transmission states, $s_1, \ldots, s_n$, by JMLS-GM algorithm. The values from the JMLS-GM training are then used as set values required by the target link to, at least, meet the threshold of the desired SINR/data rate with a certain level of energy efficiency (reward r). During HO, the goal is to find a target link that can simultaneously satisfy, $x_1, \ldots, x_n$, and corresponding SINR over n states, i.e., $s_1, \ldots, s_n$. Particularly, at HO, the scheme looks at how each user's potential received power pattern for a specific/individual target link over a given transmission range given $s_1, \ldots, s_n$ will deviate from the optimal global optimal pattern $x_1, \ldots, x_n$. The level of deviation of a target link's expected deterioration pattern determines its reliability after HO for a given number of states.

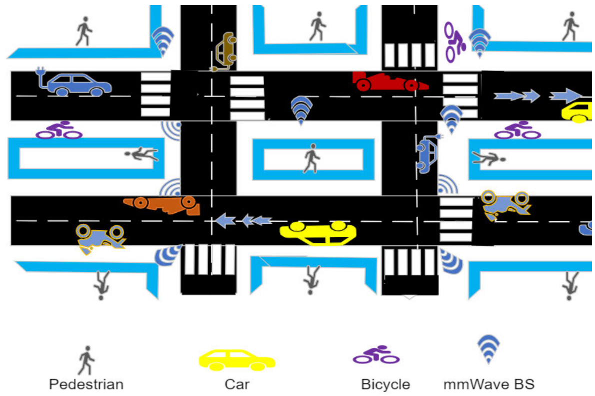

FIGURE 1. Multiuser-mobility model.

A. MANHATTAN GRID MOBILITY MODEL

A Manhattan grid model [12] is used to model the road network with streets and intersections in an urban scenario as shown in Fig. 1. The network area is 100m x 50m. We have three types of users: *pedestrians*, *cyclists*, and *cars* totaling 200. A third of the total number of users are pedestrians with speeds of 1.4m/s. Another third are cyclists with velocities between 7 and 8m/s. The other third are car users with average velocities of between 10 and 14m/s. Cars in the range of 3m or less to each other adjust their velocities by between 1 and 3m/s to avoid collision. Car velocities are updated every 3s to decrease/increase. Each street consists of right and left lanes for each user-type. Given user directions, i.e., $\mathfrak{y}$, the probability, $\mathbb{P}_{\mathfrak{y}}$, of delay or slowing at the crossing is [13] and [15]

$$\mathbb{P}_{\mathfrak{y}} = \frac{1}{k}\sum_{i=1}^{k} \frac{str\,(G_i) - 1}{link\,(G_i) + \left(P^i / T^i \delta_{int} v\right)} \tag{1}$$

where k is the number of all possible crossings at the intersection, $str\,(G_i)$ is the number of crossings available after the i^{th} street is blocked or not allowed due to traffic rules, $link\,(G_i)$ is the number of street available to cross to after the $\boldsymbol{i}^{\text{th}}$ number of streets is blocked. G_i represents the crossing after the $\boldsymbol{i}^{\text{th}}$ combination of streets. T^i and P^i are the average delay time at crossing and the probability that the user will be in LOS with a serving mm-BS, respectively. δ_{int} and v are user density and lane average velocity, respectively. Here, $\mathbb{P}_{\mathfrak{y}}$ lies in between (0, 1]. It holds true that higher $\mathbb{P}_{\mathfrak{y}}$ values imply, low robustness at crossing, i.e., longer delays with link degradation.

B. DATA RATE

The data rate, r^m, given a transmit power, P, is defined as [16]

$$r^m = blog_2\left(1 + \frac{P\left|h^H \boldsymbol{p}\right|^2}{(1+d^{\propto})} F_x\left(\left|\theta_k^l\right|\right)\right) \tag{2a}$$

$$h^H p = \sum_{\hbar=1}^{K} g_{\hbar}(t)\, e^{-2\pi i f_d \cos(\varphi_{\hbar}^{l}) t}, \tag{2b}$$

$$\varphi_k^l\left(^\circ\right) = \frac{1.4\, x10^4}{f_c\,(GHz)\,.v\,(km/h)}, \tag{2c}$$

where $\theta_k^l = \frac{2d \sin \varphi_k^l}{\lambda}$ is the normalized central angle of arrival of beam, p, with respect to a user beam given user velocity, v, under 50 Km/h and carrier frequency, f_c, as in (2b). f_d is the maximum Doppler shift given the central $\varphi_k^l\left(^\circ\right).\left|h^H \boldsymbol{p}\right|^2$ is the channel gain [6] coefficient. $F_x\left(\left|\theta_k^l\right|\right)$ denotes the Fejér kernel value variation such that as SINR approaches maximum, as user speed approaches zero, i.e., $F_x\left(\left|\theta_k^l\right|\right) \rightarrow$ ú1, particularly because the user beam aligns with the transmission beam. Otherwise, F_x approaches 0 as v increases [4] due to rapid variation. α is the path loss exponent [9]. The respective LOS and NLOS pathloss exponents are denoted by $\propto^{k_L}$ and $\propto^{kN_L}$, $\forall k \in \{1, \ldots, K\}$, in LOS and NLOS. $g_{\hbar}(t)$ is the time-varying gain of the channel over K clusters.

C. RESOURCE ALLOCATION PROBLEM

Assuming Θ is a set of optimization parameters from (2a) to (2c) with access policy, π. The outage probability, P_π, over observable signal set, Y_k, can be defined as [2], [11], and [23]

$$P_\pi\left(Y_k|\Theta\right) \triangleq P\left(\sum_l \sum_{s_k} b_l log_2\left(1 + \hat{\gamma}_t(x)\right) \geq \acute{Y} r_l^{m\zeta}\right), \tag{3a}$$

where $\hat{\gamma}_t$ is the measured SINR and $r_l^{m\zeta}$ is the targeted data rate given channel state, $s_t \in S$. b_l is the bandwidth given link channel l on signal $y \in Y_k$. We assume three sources of outage for a mmWave link, and all mm-BSs directionally transmit equal maximum power P. Firstly, all users have a receiver sensitivity of x_{kmin} and a threshold x_{k0}, thus must sustain $x_{k0} > x_{kmin}$ to avoid outage. Secondly, any user-mm-BS link that requires transmit power that exceeds maximum P to meet target data rate will not be established, i.e., will experience truncation outage [16] at a distance, d. That is for LOS and NLOS scenarios, users at distances greater than, $(P/x_{k0})^{\frac{1}{\propto \o^{kL}}}$ and $(P/x_{k0})^{\frac{1}{\propto \o^{kNL}}}$ [4], from target BSs are unable to communicate, respectively. Thirdly, outage may be due to not just meeting the desired rate, $r_l^{m\zeta}$, in (3). The minimum rate requirement problem of $\mathrm{R^m}$ under total power constraint and minimal outage requirement at t is defined as:

$$\mathbb{R}^m : \sum_t \sum_{S_t, l}\left[\left(1 - \beta_l\left(P_\pi^{m|x_t} P_{NL_k} + P_\pi^{m|u_t} P_{NL_k}\right)\right) r_l^{m\zeta}\,(y)\right], \tag{3b}$$

where $P_\pi^{m|x_t}$ and $P_\pi^{m|u_t}$ are LOS or NLOS conditional outage probabilities for a user in the m^{th} state of a user, respectively, given the mean outage probability, P_{NL_k}, at a particular transmission distance. $r_l^{m\zeta}$ is the optimal attainable rate for the l^{th} link at d. β_l is a binary factor. It is 1 or 0 for either the first or second component of **(3a)** depending on whether the expected outage condition at d is given access policy π is in LOS or NLOS. The corresponding energy cost, E_c, consists of two parts; transmission energy consumption in LOS and energy consumption in NLOS as defined as [7] and [8]

$$E_c = \beta\left\{P_i \frac{c\,(t-w)}{\mathbb{R}^m} + e_0 * \zeta c\,(t-w)\right\}, \tag{3c}$$

where P_i is the transmission power in link i, $c\,(t-w)$ is the total actual number of packets received at time t during window w at a distance $d_t \in s_t$ and e_0 is the unit energy per packet size in byte. $\zeta c\,(t-w)$ is the lost number of packets given the total expected packet over window w. β is the cost per unit energy. Thus, under the constraints of maximum transmission power, receiver sensitivity, gain and bandwidth constraints in NLOS and LOS scenarios [5], the problem is maximizing link utility with efficient and least power cost. To maximize long-term link utility of a target link, the long-term projected power, x_{t+1}, over for a minimum latency, $c\,(t-w)/\mathbb{R}^m$, is estimated as (4c):

$$x_{t+1} = \max \sum_{x_t, u_t}\left\{\frac{\gamma^{min}}{\hat{\gamma}} x_t - \frac{\alpha u_t^2}{\beta \hat{\gamma}^2}\right\}^+, \tag{3d}$$

where $\{.\}^+ = \{max, 0\}\,.x_t$ is current LOS received power; and u_t is likely discrepancy in NLOS received power over the same distance between user and mm-BS. γ^{min} is the minimum required SINR to satisfy the desire QoS and $\hat{\gamma}$ is the measured SINR. The second term parameters, α and β, are power and SINR scaling factors, respectively, account discrepancies between NLOS and LOS scenarios for the same distance.

We utilize JMLS to model feasible optimal received power for the long-term link utility of target links with optimal latency and minimal energy cost. JMLS is known for modelling abrupt and continuous behavior changes in failure prone systems [16]; thus, can incorporate abrupt effects of mmWave's NLOS and LOS switching dynamics.

III. JMLS-MFGT SYSTEM DEFINITION

We first reformulate the link optimization problems in **(3b)–(3d)** into JMLS learning problem.

A. THE JMLS REPRESENTATION

Let's the values of (3a)-(3c) be defined as JMLS equations [12], [16]

$$\begin{cases} x_{t+1} = A(s_t)x_t + B(s_t)\,u_t + g\,(s_t)\,w_t \\ y_t = \gamma^{min}\,(s_t)\,x_{t+1} + Q\,(s_t)\,v_t, \\ \mathcal{M} = (\Theta, P\,(S)\,, \pi) \end{cases} \tag{4a}$$

where $x_t \in X$ is the current received power in LOS given state, s_t, with the initial value, x_0, at an initial distance, d_o; $u_t \in \mathcal{U}$ is the associated power discrepancy due to blockage/ NLOS effects; $A(s_t)$ and $B(s_t)$ are dynamic SINR over power coefficient matrices with respect to (4c), $Q(s_t)$ and $g\,(s_t)$ are dynamic weighted noise given channel-gain measurement on SINR and EE(γ^{min}), respectively. s_t denotes the state governing parameter set $\Theta = \left\{A, B, R, \gamma^{min}, Q\right\}$ and belongs to Markov process at time t as defined as: $s_t = \{r_t, T_t, d_t, \eta_t, v, b\}$, where:

$\boldsymbol{v} = [\boldsymbol{v_1}, \ldots, \boldsymbol{v_T}]$:	is a vector of user velocity,
$\boldsymbol{r_t} = [\boldsymbol{r_1}, \ldots, \boldsymbol{r_t}]$:	is a vector representing user data rate,
$\boldsymbol{T_t} = [\boldsymbol{t_1^m}, \ldots, \boldsymbol{t_N^m}]$	is a vector of average service time,
$\boldsymbol{d} = [\boldsymbol{d_t^m}, \ldots, \boldsymbol{d_T^m}]$	is a vector of transmission distance points that have matching SINR and power values,
$\boldsymbol{\eta} = [\boldsymbol{\eta_1}, \ldots, \boldsymbol{\eta_N}]$	is a vector of user direction in n^{th} sample,
$\boldsymbol{b} = [\boldsymbol{b_1}, \ldots, \boldsymbol{b_T}]$	is a vector of mm BS.

$r^{min}(s_t)$ is the immediate reward following a HO to a link with received power, x_{t+1}, given observable signal $y_t \in Y$. $r^{min}(s_t)$ is defined as a function of energy efficiency at the receiver.

$$r^{min}(s_t) = \frac{r^m(s_t, P_{Los})}{P_{Los} + P_C}, \tag{4b}$$

where $r^m(s_t, P_{Los})$ is the maximum expected data rate, $P_{Los} \le P$ is the maximum transmission power at distance, $d \in s_t$, P_C is the circuit power consumption. The transition probability function between different states is given by:

$$P(S) \triangleq P\left(s_{t+1} = m_j | s_t = m_i\right), \tag{4c}$$

The normalized transmission energy cost index, $J(x_t)$, over N sampled distances/states from BSs is defined by:

$$\mathcal{J}(x_t) = \boldsymbol{E}\left\{\sum_{j=1}^{N} \|x_t\|^2_{Q(s_t)} + \sum_{j=0}^{N-1} \|u_t\|^2_{R(s_t)}\right\}, \tag{4d}$$

where $\|x_t\|^2_{Q(s_t)} > 0$ and $\|u_t\|^2_{R(s_t)} > 0$ represent weighted norm energy costs for actual received packets and lost packets (e.g., due to NLOS effects), respectively, given the expected number of packets over window, w. $J(x_t) \triangleq \varepsilon E_c$, where ε is normalization factor of the energy cost, E_c.

B. INITIAL TRAINING

Letting Y_T, X_T and S_T denote patterns of observed channels $\{y_1, \ldots, y_T\}$, the corresponding received power pattern $\{x_0, \ldots x_T\}$, and states $\{s_1, \ldots, s_T\}$ until time T, respectively. Given a finite set, Y_T, is observed over S_T, the JMLS model learning problem is to predict the best pattern, X_T, for parameter set Θ, over a finite distance $\{d\} \in S_T$, i.e., the desired QoS, particularly the SINR given latency and energy efficiency effects in mmWaves. Here, given $P(y_t, s_t | x_t, \Theta)$ is the conditional distribution of y_t, for, x_t estimates in state s_t, we use Expectation-Maximization (EM) algorithm in [7] where Bayesian inference [6] automatically infer initial unknown values of parameter set, Θ, over S_T and X_T, $Q\left(\Theta \mid \Theta^k\right)$, as defined by:

$$Q\left(\Theta \mid \Theta^k\right) = \mathbb{E}\left[\log P\left(X_{T-1}, S_T, Y_T | \Theta\right) | X_k, \Theta^k\right], \tag{5a}$$

$$\Theta^k = \arg\max_{A \in \Theta} Q\left(\Theta \mid \Theta^k\right), \tag{5b}$$

where Θ^k is the current known parameter estimate at the $\boldsymbol{k}^{\text{th}}$ iteration. The values are optimized based on the updates of the SINR changes corresponding energy cost in (4d). At every state, S_T, the received power pattern of a target over distance needs to be enough to meet the desired SINR to avert outage. In scenarios where the LOS is not clear, signals are deflected, diffused and reflected. We thus look at the extent of channel gain influence for a transmitted signal to meet SINR value per state. We note however, that the influence of channel gain is hidden from EM estimates. We thus use Mean Field Game Theory to simultaneously and strategically trade off the influence of received power against channel gain in predicting the SINR hence data rate and EE over distance for a target network. The MFGT considers three possible values and choose the target link with the least energy cost: hence best EE. The three energy cost factors to estimate on target link with reference to (4d) are:

1) the energy cost in LOS given a maximum received power is received for a maximum SINR value, $a \in A$

$$\mathcal{J}(x_t, a) = \max_{(x,a)} \boldsymbol{E}\left\{\sum_{0j=1}^{N} \|x_t\|^2_{Q(s_t)}\right\}; \tag{5c}$$

2) the energy cost in NLOS given a channel gain, h, compensates limits in received power for a max-SINR value, $a \in A$ (power limited)

$$\mathcal{J}(h_t) = \max_{(x,a)} \boldsymbol{E}\left\{\sum_{j=0}^{N-1} \|u_t\|^2_{R(s_t)}\right\}; \tag{5d}$$

3) the minimal energy cost based a balance in h and x_t is

$$\mathcal{J}(h_t, x_t) = \min \boldsymbol{E}\{\alpha J(x_t)\} + \min \boldsymbol{E}\{\beta J(h_t)\}. \tag{5e}$$

Considering $c(J(h_t), J(x_t, a), J(h_t, x_t))$ is the ultimate cost of the three, i.e., a convex function of $J(h_t)$. For the agent, in state s_t, the learning problem for the desired x_t becomes;

$$x_t(y_t) \triangleq \arg\min c\left(J(h_t), J(x_t), J(h_t, x_t)\right) + \arg\max r^{min}(s_i). \tag{5f}$$

We use mean field game theory [8], [9] to know the best energy cost over long-term mmWave connectivity. With mean field GT, we do not need the full analysis of all possible optimal power deterioration patterns. We just need to know the distribution of the similar energy cost experience, $J(h_t, x_t)$ given received power, x_t, and channel gain, h_t as factors influencing the SINR value in (5). The following section defines the mean field function value.

C. MEAN FIELD GAME VALUE FUNCTION

Let a be the SINR value under consideration and $\hat{a}$ denote the SINR estimates by EM for $x \in X$. We define the expected power value, x_t, as a variable of the immediate reward $r^{min}\left(\hat{a}, s_t, a\right)$ given the mean field value function, $\boldsymbol{V}(\boldsymbol{x_t})$:

$$\boldsymbol{V}(\boldsymbol{x_t}) = \max_{\pi} E\left\{\sum_{k=0}^{K} \delta^k r^{min}\left(\hat{a}, s_k, a_k\right) |_{\pi(x_t \mid \theta\pi)}\right\}, \tag{6a}$$

with policy, π, at each state, s defined as

$$\pi = \arg\max_{a_0,\ldots a_k}\left[\mathscr{r}^{min}\left(, s_k, a\right)+\sum_{s} P\left(s_{k+1}|s_k, a\right) V\left(s_{t+1}\right)\right], \tag{6b}$$

where $\delta \in (0, 1)$ is the discount factor and moves to closer to 1 as energy cost factor J, reduces.

The long-term value function, $V\left(\check{x}_k\right)$, is approximated as

$$V\left(\check{x}_k\right) = \sum_{a_0,\ldots a_k} P_\pi\left(x_k|y_t, \theta(a_k)\right)\left\{\mathscr{r}^{min}\left(-a, s_k, a_k\right)\right. \\ \left.+\delta^k\sum_{x} P\left(s_{k+1}|s_k, a, x\right) V\left(s_{t+1}\right)\right\} \tag{6c}$$

where $P_\pi\left(x_k|a_k\right)$ denotes the probability of obtaining a SINR value, a, given policy π. The mean field game looks for a target link whose SINR values satisfy a condition $a \geq \hat{a}$ over s_k when less than $\boldsymbol{x_k}$ is received power needed to satisfy $\hat{a}$. The discount factor, δ^k, is a tolerate factor over S_K states. Principally, to get the link with the desired SINR, using EM estimates, the mean field game theory selects target links with high channel gain to compensate for any received power loss e.g., in NLoS scenarios. And the following section explains in detail.

D. DETERIORATION PATH PROJECTION EHHANCEMENT

To find the optimal power pattern over k states, we let $g\left(x\right) : \mathbb{R}^d \rightarrow \mathbb{R}$ represent the mmWave link terminating cost factor and $r\left(x, a, -a\right) : \mathbb{R}^d x\mathbb{R}^m \rightarrow \mathbb{R}$ denote its running reward given the function value, $V\left(\check{x}_k\right)$, in each state over set, Y_K, **HJB** such that

$$V\left(\check{x}_k\right)\max_{a\in A} -E\left\{\int_0^K \delta^k r\left(x, a, -a\right) ds\left(t\right) + \delta^k g\left(x\right)|Y_k\right\} = 0, \tag{7a}$$

where the terminal energy cost factor, $g\left(x\right)$, corresponds to the outage costs, $c\left(*\right)$ satisfying the receiver sensitivity threshold condition, $x_{k0} > x_{kmin}$. Here for every given step in $Q\left(\Theta\,|\,\Theta^k\right)$ by EM, the pattern $X = \{x_0, \ldots x_k\}$ is calculated using the Hamilton–Jacobi–Bellman (HJB) equation in (7a):

$$x \triangleq V\left(s_t\right) + \sup_{hEU} \mathcal{H}\left(x, h^H p\right), \tag{7b}$$

where the Hamiltonian function, $\mathcal{H}\left(x, h^H p\right)$, is defined by a Hamilton principle [9]. The principle of HJB states that the true evolution of points in $X\left(t\right)$, i.e., between x_1 and x_2, is described by K states $\{V_1\left(s_t\right), \ldots, V_k\left(s_t\right)\}$ where $a \in A(t)$ is a stationary value (a SINR point of x where the variation of SINR is zero):

$$\boldsymbol{\mathcal{H}}\left(\boldsymbol{x}, \boldsymbol{h^H p}, \boldsymbol{a}\right) = \boldsymbol{0}. \tag{7c}$$

This is achieved by ensuring any received power value change between states, is compensated by a feasible channel gain value, $h^H p$, to satisfy desired SINR. Technically, HJB values for a selected value, x must satisfy the conditions in (7c) in each, k^{th} state for the serving cell to sustain connectivity.

For the mean Field principle [16], the best SINR link state s_j following user transition from s_i is determined by trading off the cost of receiving $x_j \in X_T$ with channel gain to obtain the same or better SINR value for a given set of valuables, Γ, in the game,

$$\Gamma < \left\{\mathcal{J}_{ij}\right\}, \left\{S_{ij}\right\}, H\left(\mathcal{u}, x_i, t\right) >, \tag{7d}$$

A Nash Equilibria here is thus a set of received power $\{\bar{x}_i, \ldots, \bar{x}_N\}$ over N states $\left(\bar{s}_1, \ldots, \bar{s}_N\right)$ such that it is "costly", i.e., in terms of energy, for a player/user to select a link that deviates from receiving power pattern $\{\bar{x}_i, \ldots, \bar{x}_N\}$ over $\left(\bar{s}_1, \ldots, \bar{s}_N\right)$ states to meet the threshold of the desired SINR over N states. Thus, any target link that meets the equilibria condition is worth the HO. Technically, the target link that meets the equilibria conditions $\{\bar{x}_i, \ldots, \bar{x}_N\}$ is likely to sustain connectivity hence SINR above threshold over, N states $\left(\bar{s}_1, \ldots, \bar{s}_N\right)$.

E. THE MIXED NASH EQUILIBRIUM

Pure Nash equilibria with $\{\bar{x}_i, \ldots, \bar{x}_N\}$ over $\left(\bar{s}_1, \ldots, \bar{s}_N\right)$ states does not however exist, Simultaneous effects of partial, and full NLOS, LOS and block over. Wave propagation makes it challenging to yield $\{\bar{x}_i, \ldots, \bar{x}_N\}$ values over. Thus, we introduce the notion of mixed values. A certain innovation, $F_{,i\rightarrow j}$, cost determined by the tradeoff between received power and channel gain variation:

$$F_{,i\rightarrow j} = \max_{a\in A}\left[V_{ij}x_i x_j + H\left(\mathcal{u}, x_i, t\right)\right], \tag{8a}$$

This means that each player anticipates a certain set of received power values from a given mm-BS connection with a certain probability, $\mathcal{P}\left(\mathcal{V}_i\right)$, where x_i, does not meet the desired SINR. Particularly, given $a^d \in \tilde{A}$ is the desired SINR value at a distance $d \in s$. The sufficient and necessary condition for a^d given corresponding EM value $\hat{a}^d$ needs is asymmetric Nash Equilibrium defined as:

$$\begin{aligned} &\mathscr{r}^{min}\left(\hat{a}^d, s_k, a^d\left(s\right)\right) \\ &\quad + \delta^k\sum_{x} P\left(s_{k+1}|s_k, \hat{a}^d, a^d(s), x\right) V\left(s_{k+1}\right) \\ &\geq \mathscr{r}^{min}\left(\hat{a}, s_k, a\left(s\right)\right) \\ &\quad + \delta^k\sum_{x} P\left(s_{k+1}|s_k, \hat{a}^d, a^d(s), x\right) V\left(s_{k+1}\right). \end{aligned} \tag{8b}$$

By solving the above conditions, the mean-field approximation as the function, F^*, of the innovation over all time $t = 1, \ldots, T$ can be defined by (8c) as [9]

$$F^* = \left\{\sum_{i,j} V_{i,j}\mathbb{E}_\mu\left[X_i X_j\right] + \sum_{i} H\left(\frac{x_{i+1}}{2}\right)\right\}, \tag{8c}$$

where μ is the probability that the user value x_i at distance d can be calculated based on the partial costs as

$$\mu_{d,i}(t) = \frac{1}{1 + \exp\left(\sum_{i \neq j} V_{d,x_i} - V_{d,x_j}\right)}. \tag{8d}$$

The HJB equation corresponding best pattern X_T is solved backwards in time, starting from a state with $t = T$ and ending at $t = 0$ [12] i.e., given S, for X.

At HO, the divergence of individual target link pattern is compared to global optimal pattern is measured generated from a collection of all data. The higher the divergence, the less reliable the target link and connection is. We use the Kullback-Leibler (KL) divergence to measure the level of unreliability of a target cell and is explained in detail in the following sections below.

F. DETERIORATION PATTERN ANALYSIS

We use the KL divergence to measure the divergence of target (local) pattern X, and function values, $V_{Los}(\check{s}_t)$ and $V_{NLOS}(\check{s}_t)$, in LOS and NLOS, respectively, from the global pattern values $V_i(\check{s}_t)$. KL measures how either $x_{Los}(\check{s}_{i+t})$ or$x_{NLOS}(\check{s}_t)$ with reference to their respective value functions, $V_{Los}(s_i)$ and $V_{NLOS}(s_i)$, will deviate from desired $x(\check{s}_t) \in V_i(\check{s}_t)\, \forall \bar{X}$. The KL equation and three conditions used for prediction of target link behavior are given by

$$KL\left(V(s_t)\,||V(s_{t+1})\right) = \mathbb{E}\left[\log\left(\frac{V(s_t)}{V(s_{t+1})}\right)\right]. \tag{9}$$

1) Using forward KL [12], the difference between $V(\check{x}_k)$ and $V_{Los}(s_{i+1})/V_{NLos}(s_{i+1})$ is weighted by $V(\check{x}_k)$, If, for instance, the reward is zero, i.e., $V(s_t) = 0$ then the current predicted pattern $\bar{X}$ needs to be updated to know which target cell would deteriorate optimally. In other words, $\bar{x}_t$ as minimal received power is not desirable and will not satisfy receiver sensitivity condition $x_t > x_{ko}$ to avert outage.
2) Conversely, if $V(s_t) > 0$, then the $\log\left(\frac{V(s_t)}{V(s_{t+1})}\right)$ term values will contribute to deterioration pattern $(\bar{x}_1, \ldots, \bar{x}_N)$ updates using pattern $(x_1, \ldots, x_N)$. If the divergence is high, this is not good because our objective is to minimize KL divergence with discrepancy predicted, $\bar{x}_t$, and the target cell local value x_t. We measure $V(s_i)$ with either $V_{Los}(s_{i+1})$ or $V_{NLOS}(s_{i+1})$ and whichever target link gives minimal KL divergence value will cause less abrupt changes in received power requirement. It also implies the target link can get the desired SINR over different states. Particularly by using the channel gain to balance up received power deficits in achieving the desired SINR within each transmission range.
3) We use Reverse KL to assess the target cells to meet the conditions in (1) and (2). The target cell whose KL difference between (1) and (2) is lower than that of other cells is more reliable. A higher reverse KL implies a wide divergence on how a target link is likely to perform after HO. This also indicates to how much the likely target link predicted desired pattern, $x_1, \ldots, x_N$, will likely not be meet given the optimal desirable deterioration pattern $(\bar{x}_1, \ldots, \bar{x}_N)$. This form of KL Divergence indicates how global $V(s_t)$ and local $V(s_t)$ estimate difference approach 0 on some areas where the link of a target link is more reliable and stable.

IV. ONLINE UPDATE OF THE PROPOSED MODEL

They are two ways in which the deterioration experience is updated. We consider two forms of link deterioration experience; 1) locally on individual mmWave BSs and 2) a global deterioration pattern, i.e., aggregative experience of distribute mmWave BSs. In this approach, the overall experience is independent of the number of mmWave BSs and may be scaled as desired at the cost of additional communication overhead. To find X_T over Y_T, the online optimization process is summarized in Algorithm I. It follows the classical EM algorithms [7] and alternates it with game theory at each step. Given the current iteration, the new path, X_t, uses the previous pattern, X_{t-1}, as a warm restart. The coefficients computed during the previous iterations for X_{t-1} are used as aggregate information for X_t. The information from past coefficients is thus carried forward in matrices as initial input to the later parameters. For instance, the SINR values, $\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_t$, are carried forward in matrices as in step (10):

$$A_t \leftarrow A_{t-1} + \gamma^{min}\hat{\gamma} \text{ and } B_t \leftarrow B_{t-1} + \hat{\gamma} u_t. \tag{10}$$

This allows the HO scheme to update received power patterns based on previous information without accessing the old data samples. The projected, X_{T+1}, is then optimized by using (10) and the previous global pattern, X_{T-1}, as input to X_T. This accelerates convergence rate and pattern prediction than classical batch algorithms. The learning scales up gracefully to use large data sets as more data samples build.

V. PERFORMANCE EVALUATION

In this Section, we describe the simulation model used and present the simulated results to study the performances of our proposed HO scheme.

A. NUMERICAL ANALYSIS MODEL

We take into account mmWave BS downlinks and concentrate on user experiences with SINR outages. With unique spatial densities, gains, receiver sensitivity, blockage parameters, and pathloss exponents, (see Table 1) BSs are dispersed randomly. A density λ_u, users are dispersed spatially within a voronoi circle denoting a mmWave cell using independent homogeneous Poisson point process (PPP) for user connection. BSs equally are dispersed in a homogenous PPP denoted by Φ and the BS density λ_{BS}. Each BS serves at least user selected at random from a Voronoi cell [11] using a round robin scheduler, hence the densities of users and BSs are assumed similar. Even after forming a single association with BS, we expect that the active users continue with a PPP [12].

The approximation does not take into account the connection of BS and served user point operations and assumes free space attenuation of a link. For analysis, the correlation between the reference BS and user link reliability is a derivation of SINR outage probability at various receiver sensitivity threshold with respect to KL divergence calculation $V\left(s_t, \lambda_{Bs}, \lambda_u\right)$, $V\left(s_{t+1}, \lambda_{Bs}, \lambda_u\right)$ in (9). Analytical results are depicted in Fig. 10.

B. DESCRIPTION OF THE SIMULATION MODEL

We assume that the state information, e.g., distance, speed, etc., can be known by any mm-BS in the network using a location-based service such as GLS [4] or HLS [5]. For our simulation, we use the DC LTE-mmWave model introduced by the NYU and the University of Padova [3] where LTE BSs (L-BS) manage mmWave BS (mm-BS). The model carefully considers an end-to-end mmWave cellular network performance. It uses ns-3 simulator featuring 3GPP channel model for frequencies above 6 GHz and a 3GPP-like cellular protocol stack [8]. NS-3 employs a Spatial Channel Model (SCM) in this evaluation of NR networks at mmWave frequencies. Here, the channel matrix, $\left|h^H \boldsymbol{p}\right|^2$ with an entry transmitter, t, and receiver, r, models the channel between the t-th and r-th antenna elements at the transmitter and receiver, respectively. Evert link entry $(t;\ r)$ is determined by the contribution of N clusters to depict NLOS reflections and the direct LOS path (if in existence). To replicate each cluster, different powers and durations are used, which is reliant on a number of rays spread around a common cluster angle of arrival and departure (see (1) - (4)]. Parameters here rely on the time varying Rician K factor determined by the instantaneous velocity of the measured vehicle [2], [5], [22]. A summary of simulation parameters is given in Table 1 while more details can be found in [8]. The initial stages of our model determine possible received power deterioration patterns for different users. We use the EM estimator on JMLS. The latter stages involve optimizing the initial pattern using game theory. At HO selection stage, i.e., to choose the best target cell, individual cell deterioration patterns with respect to user type are compared to the global JMLS-mean field GT pattern optimized. We use reverse KL divergence [12] to understand the margin of divergence of the target cell pattern from that of the global optimized pattern. The smaller the divergence the more reliable the target cell is during and the post HO process. The learning algorithm is online processed at the core network and is developed on an augmented ns-3-OpenAI Gym [10] toolkit as summarized in Algorithm I. Open AI Gym is integrable with ns-3 and supports the teaching agents for the variety of applications.

C. IMPLEMENTATION OF THE SIMULATION MODEL

Dynamic system level simulations are performed to evaluate the performances. Fig. 2 shows the flow chart of HO execution. The distribution of the received powers with respect to different states of the target links is predicted using EM estimations given a few initial values obtained from user connections and expert data values. Initial optimal received power patterns are then generated. The patterns are classified into two sets. One set is assumed to be influenced by high gain due to NLoS and while the other set is influenced by high received power values due to LoS to achieve the same SINR. In the meantime, the total reward, which is the energy efficiency (EE) of each pattern state, is also calculated based on the received power pattern estimations in the flowchart in Fig. 2 using the following steps.

TABLE 1. Simulation parameters [4].

Parameter	Value
mmWave	28GHz
mmWave bandwidth	1GHz
3GPP Channel Scenario	Urban Micro, Urban Macro
mmWave max outage	-5dB
mmWave transmission Power	46dBm
mmWave max PHY Rate	3.2Gbps
X2 link latency	1ms
S1 link latency	10ms
RLC buffer Size	5MB
S1 MME link latency	10ms
User speed	[1,50] m/s
UDP Source rate	200Mbits/sec
Receiver Sensitivity Threshold	-45 dBm
Directivity gain Maximum	7dBi BS, 5dBi
$\propto^{k_L}, \propto^{k_{N_L}}$	2,4
mmWave BS density	10 BS/km^2
Default number of users	200

Step 1: Using EM, the scheme estimates the received power in LOS and NLOS.

Step 2: For the received power values in LOS set, the objective is to find values that have least energy cost. A local optimization strategy set is obtained in this step by Hamilton–Jacobi–Bellman function conditions in (7)-(8).

Step 3: Received values from the partially NLOS user set and those from the LOS set are involved in a game. The users' received power values in the NLOS set(s) are assumed to be low but compensated by the high channel gain to get the same SINR value(s) obtained in LOS. Users with improved or degraded received power patterns given different states are moved to or from the NLOS set, while game values in the LOS sets are improved per state. Users failing to meet desired received power values are moved to NLOS set. The scheme skips to Step 4 if none of the parameters for users with regard to improving the network into the LOS value set. Otherwise, if there is a new value for a user state, we go back to Step 2.

Step 4: Repeat Step 3 until the shifting/HO condition for each state is met, i.e., attainable or empirical minimum received powers pattern. For the best target link, each user's received power about the mean at different states (***distances***) is not supposed to drop beyond user type pattern values to avoid link failure or outage.

Step 5: Calculate the total reward, EE using the final energy cost and received power pattern in (9). The power pattern

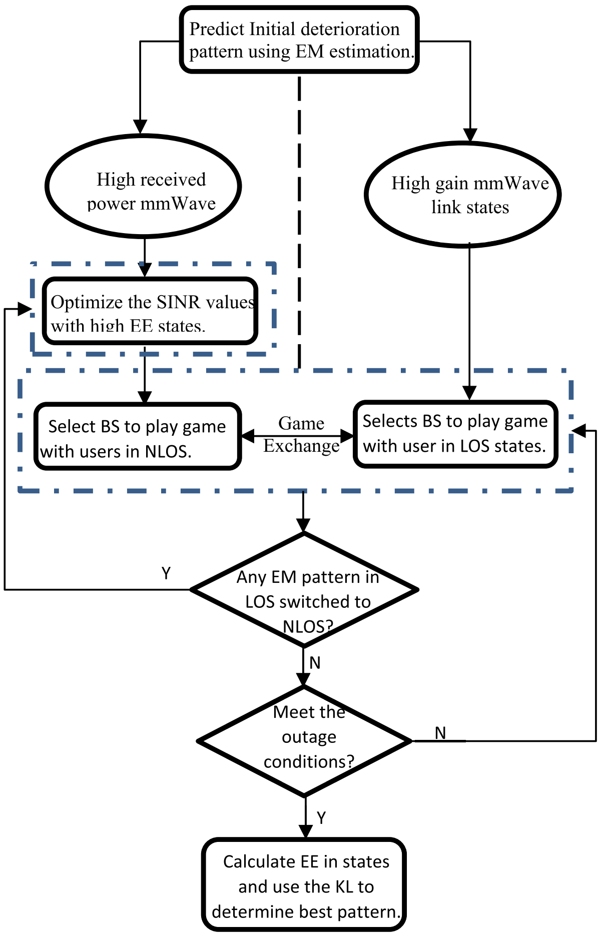

FIGURE 2. The flow chart of received power pattern estimations.

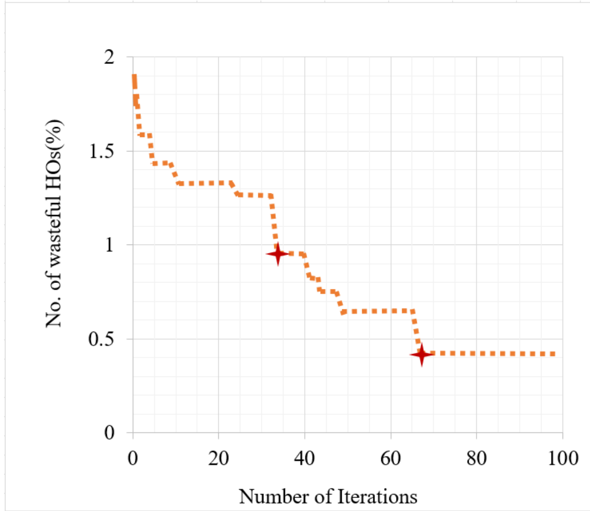

FIGURE 3. Wasteful/repeated HO vs. number of iterations.

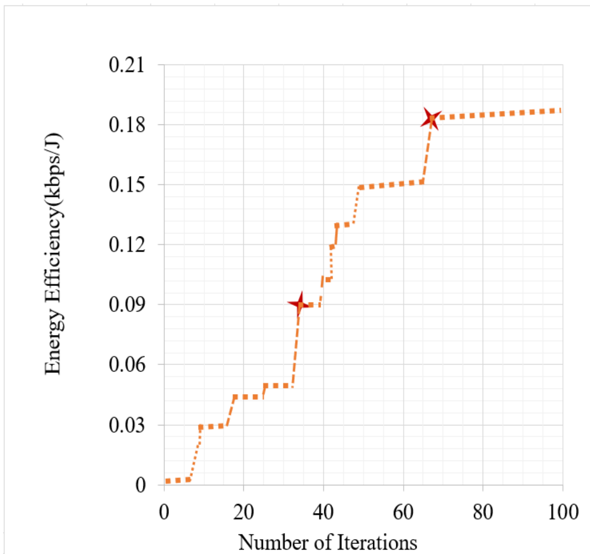

FIGURE 4. Energy efficiency vs number of iteration.

corresponds to the highest EE values defining the pattern **X**at each state.

The flow in Fig. 2 summarizes the above steps. The first dashed box is the LOS SINR value set and the second dashed box is the game's SINR parameters optimized using the Hamilton–Jacobi–Bellman solution in (7c).

D. RESULTS AND ANALYSIS

Wasteful HOs (or repeated HOs) refer to unnecessary handovers to the same serving BS. This is because the same serving BS triggers/refreshes a HO process instead of just maintaining the existing link or selecting another BS with a better link that would support longer connectivity. An increase in the number of wasteful HOs increase the HO overhead. Thus, we investigate the percentage of wasteful HOs over a period of time of training our proposed scheme as shown in Fig. 3. It is observed that the percentage of wasteful/repeated HOs using our proposed HO scheme decreases over time. In fact, after the 70^{th} iteration, the training converges because the HO link selection is optimized. It shows that any link selection/HO process initiated after the 70^{th} iteration has less than 0.5% chance of being wasteful/being repeated. This performance is attributed to the reliability of the proposed HO scheme to accurately predicted the behavioural pattern of mmWave links and ultimately select the most stable target links.

Fig. 4 shows the energy efficiency (EE) in terms of bits received over the amount of energy transmitted. The EE increases over time and thus there is an equivalent reduction in the energy cost. The increase of the EE also corresponds to the reduction in wasteful HOs over time as shown in Fig. 3. For instance, steep/swift increase in EE over time observed at around the 40^{th} episodes in Fig. 4 correspond to swift reduction in wasteful HO observed around the 40^{th} episode in Fig. 3. The mark, "$*$", in Fig. 3 corresponds to the sharp changes associated with learning pattern rate of our scheme i.e., drastic learning improvement once minimal but considerable training data has been collected online. Additionally sharp improvements further show how

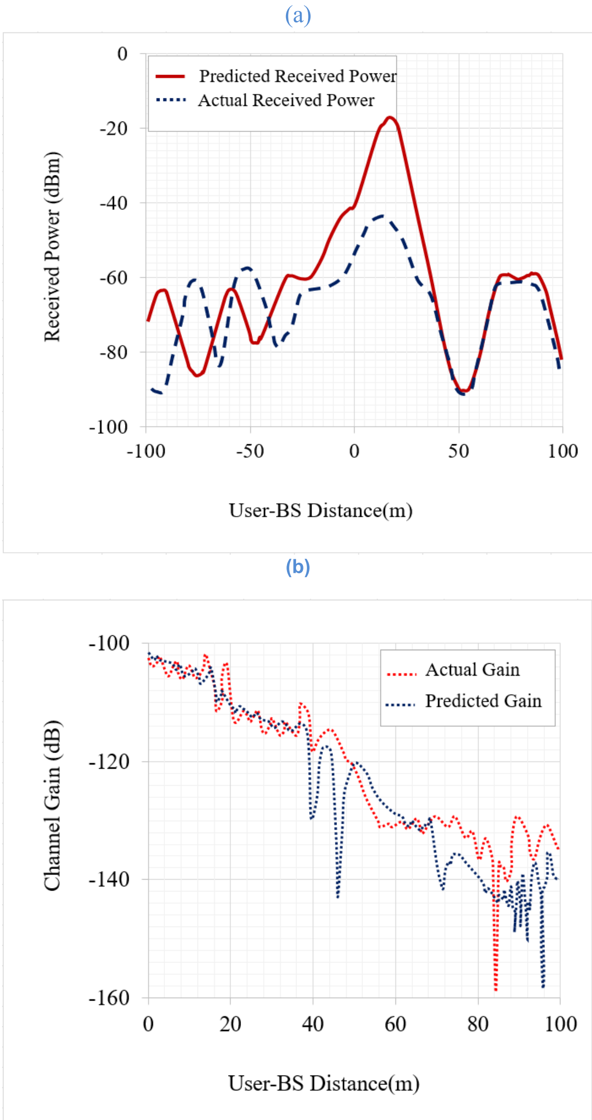

FIGURE 5. (a) Pedestrian user average received power vs distance from target cell. (b) Channel gain(dB) vs distance (m).

the proposed scheme is able to make sufficient improvements with minimal training data, e.g., within less than 40 training episodes. Particularly, with respect to user position and movements with respect to target mm-BS position. Our proposed JMLS-MFGT technique has the ability to learn fast and select more LOS target links (less wasteful HO) compared to NLOS links once a considerable but minimal amount of training data samples/iteration have been accumulated.

Fig. 5(a), shows the pedestrian user received power from the target cell vs. the user distance from the target cell in which a user is moving towards (−) and from (+) the mm-BS within a stretch of 100m. The dotted blue line shows the actual user received power within a coverage of 100m of a selected target mm-BS. The solid red line shows the predicted user received power by our proposed scheme from a selected target mm-BS that meets the desired SINR over

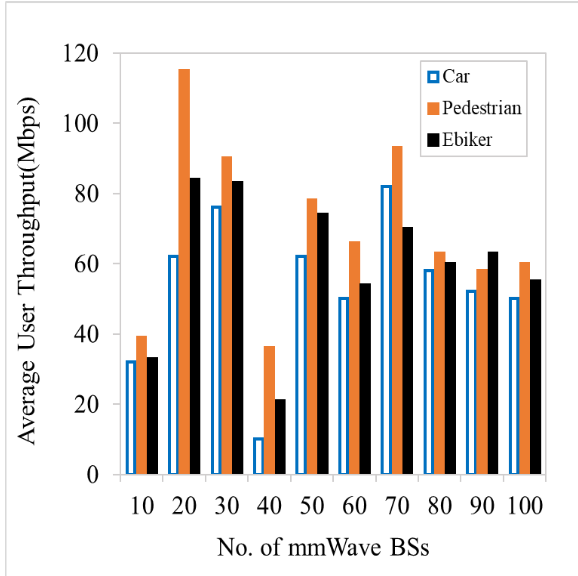

FIGURE 6. Average user throughput vs distance for different user types.

the transmission range. This closeness between the predicted and actual received power pattern proves that our proposed HO scheme can accurately predict the likely network performance behavior of the target link after a HO is executed. Fig. 5 further affirms the ability to predict accurately the received power pattern of target links prior to HO. Fig. 5(b), further shows the corresponding channel gain to distance behavior in respect to predicted and actual values in a multi array mmWave propagation condition.

Fig. 6 shows the average throughput for different user types at various user-mm-BS distances. The results suggest that pedestrian users have higher average throughput compared to any other user types. The is attributed to the ability of the proposed HO scheme to predict more accurately the behavior of the user's receive power pattern hence SINR of target links connecting to low speed users. Particularly, the channel state information (CSI) with respect to SINR performance parameters as denoted in equation no. 3 changes gradually for slower users. This gives ample time for the proposed HO scheme to predict more accurately the likely changes in the target link performance pattern from one state to the other as users move. This is unlike high-speed user training data where the change in CSI is more than often rapid, abrupt and inconsiderate of the usually slow learning rate. This ultimately negatively affects the incremental behavioral prediction of the pattern of a target link by our proposed scheme. Additionally, we closely analyzed how the corresponding SINR to the throughput changes over distance in Fig. 6. It was revealed that the throughput spikes are generally influenced by an increase in received power, e.g., due to improved LoS states.

Fig. 7 shows the HO failure rate against number of mm-BSs for different HO schemes. We compare the performance of our proposed HO scheme to other HO schemes based on mobility awareness [6], SINR-based [4] and DRL-JMLS [5].

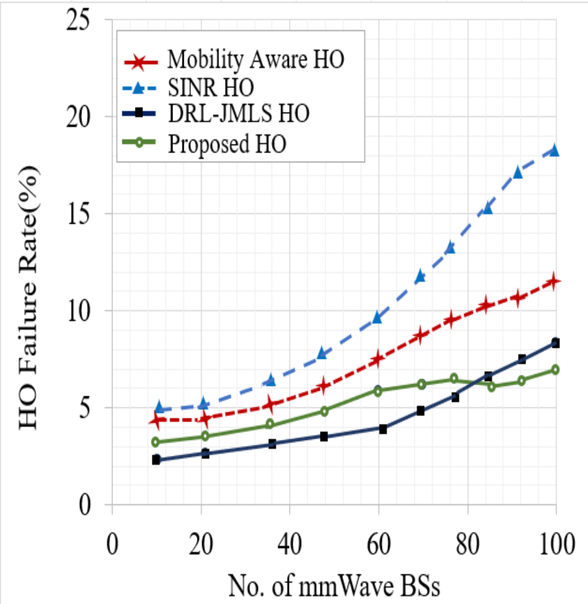

FIGURE 7. HO failure rate vs No. of mmWave BS for different HO schemes.

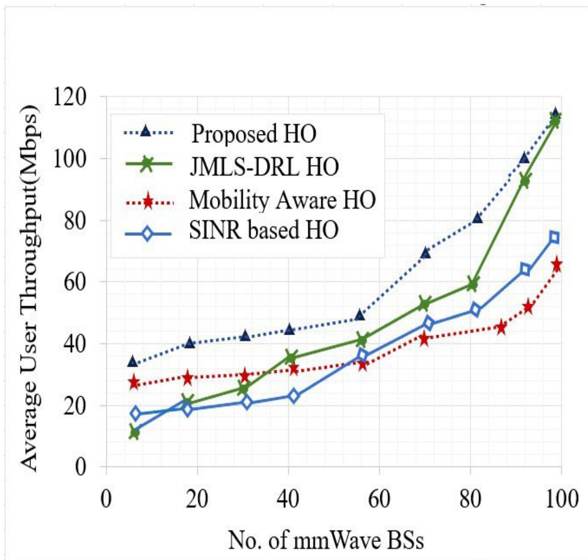

FIGURE 8. Average data rate vs no. of BSs for different HO schemes.

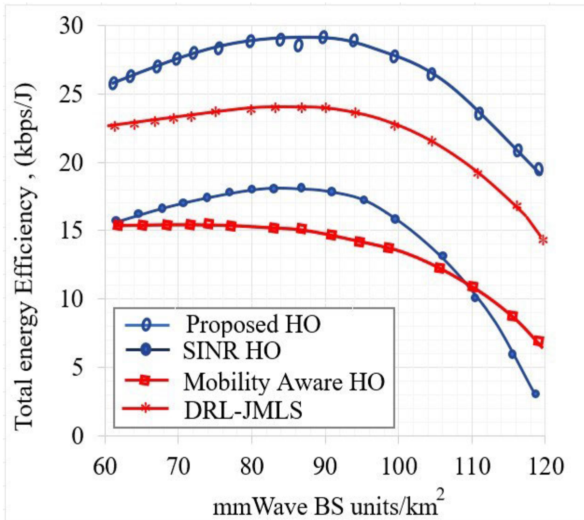

FIGURE 9. The Energy Efficiency vs mmWave BS Density variation.

The mobility awareness in [6] executes HO commands based on user speed. The SINR-based in [4] executes HO commands based on the next mm-BS with the highest SINR/data rate. The DRL-JMLS [5] is based on DRL and JMLS models to complete the HOs. As shown in Fig. 7, our proposed scheme experiences lesser HO failures with the number of mm-BS. This is because the scheme does not just select target BSs because of their high initial data rate/SINR/ received power. It also considers received power and channel gain variance post-HO using the KL divergence test (see section IVB). However, for a larger number of mm-BSs, the HO failure rate for DRL-JMLS and our proposed scheme is similar. This attributed to the fact that both have diverse data sufficient for training and making reliable HO policies within a short period of data collection from BSs. It must be emphasized that for a smaller number of BS our proposed HO scheme performs better than any other HO scheme in Fig. 7. This makes the proposed scheme more appealing for quick adjustment in HO failure decision and additionally less expensive to manage (requires less BSs to reduce HO failure rate). Further, it can make more reliable HO decisions with less resources (BS) and diversity of training data. Particularly, the lesser the number of BSs, the smaller the number of training samples the training process can collect in each training episode. This also reduces the diversity of training data collected per episode and ultimately affects the accuracy and the diversity of HO decisions.

Fig. 8 shows the average data rate of the network when different HO schemes are used. Similar to Fig.7, our proposed HO scheme gives the highest average data rate as compared to others. For a large number of mm-BSs, the average data rate for DRL-JMLS and our proposed scheme similar as both are able to scope sufficient and diverse training data within a short time3arn to select better target links, i.e., links with good LOS. Our proposed scheme however outperforms DRL-JMLS scheme when using fewer BS. This is because our proposed scheme even in limited training data (as explained in Fig. 4), requires lesser data for training and making more reliable HO decisions.

Fig. 9 compares the variation of energy efficiency for different HO schemes against the density of mm-BS. For the same value of mm-BS/km^2, the proposed HO scheme has higher EE than other 3 schemes. At peak, i.e., 90 BS/km^2, our proposed scheme is about 17% to 40% better than other 3 schemes. As the number of the mm-BSs increases, the EE rose to 90 BSs/km^2. It then slowly decreases when the number of mm-BS increases beyond 90. The decrease emanates from increasing distortion brought about by mmWave transmission sensitivity. We also observed that EE decreases with the

TABLE 2. Summary of pros and cons of different ho schemes.

	Pros	Cons
Proposed HO Model	- Better Energy Efficiency - High average data rate and lower HO failure rate. -Easy to scale for large networks. Agents (BS) HO decision based on their distribution in the network. - Mean field approximation relies on symmetry and homogeneity assumptions and enables to scale to an infinite number of agents.	-The impact of previously formulated HO policies vanishes with time elapses and user changing space, depicted by no difference in long-term average. - The HO scheme needs to constantly make HO decisions based on user space in terms of LoS and NLOS -Too complex for very small and simple mobile network environment
DRL-JMLS HO Model	- Can learn and build knowledge about a complex dynamic wireless communication environment. - Deep neural networks with suitable architectures are arguably mobile network choice, thanks to their ease of use and their generalization capabilities	- hard to scale in terms of number of agents, the HO scheme has to keep track of all agents to make HO decision. - centralized HO and power control system. -Too complex for small/simple mobile networks -require large training data set to make reliable HO policies
Mobility Aware HO Model	- Good HO decision for non-dynamic user mobility behavior -easy to scale up, ideal for small or simple networks	-Too reliant on user the consistent of user speed - centralized and poor HO decision
SINR HO	-easy to implement, does not consider user behavior -easy to scale up easy to scale up, ideal for small or simple networks	-lacks intelligence, highly reliant on only SINR variation -poor centralized HO, power control and decisions

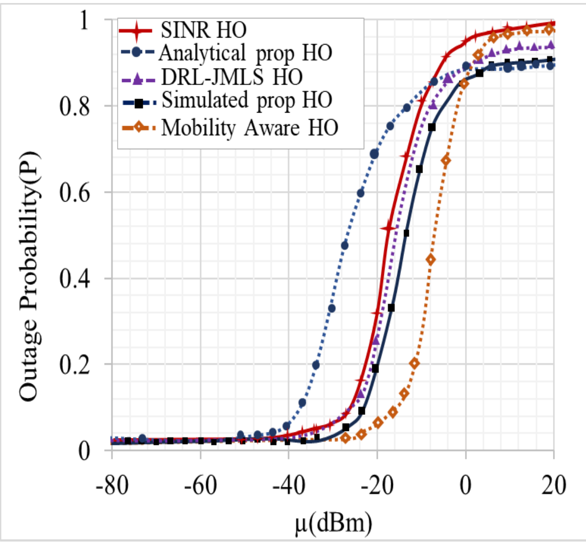

FIGURE 10. The truncation outage vs receiver sensitivity threshold.

increase in cellular number. This can be attributed to the increase in cellular power interference.

In Fig. 10, we assess the performance of different HO schemes in terms of outage probability of the mmWave link given the changes in receiver sensitivity threshold values. Particularly, any user experiences a truncation outage with respect to the serving link if the user transmitted power fails to meet the receiver sensitivity threshold $\boldsymbol{\mu}$. The proposed scheme outperformed other HO schemes especially at higher thresholds i.e., above −20dbm. This can be attributed to the scheme's ability to correctly predict the receive power pattern of the target link and select reliable (continuously above threshold) links.

Additionally, based on (3a), (7c), (8b) and (9), we theoretically analyzed outage in Fig. 10. The analytical results are in tandem with the simulated results for the most part of the threshold choices. In fact, given the variation of received power sensitivity with outage, further information was deduced via comparison of the analytical and simulated results. Particularly, it is observed in Fig. 10 that the theoretical outage at low received power sensitivity threshold values behave the same as the simulated results. This is despite assuming the received power behaviour to be the same as the free space attenuation in our theoretical analysis. This similarity is because for simulated results, the reception of the reflected rays (considered) and their contribution on the received power to be above the threshold is small due to the large reflection angles and small sender-receiver distance at HO. Therefore, the contribution to the total received power being above the receiver sensitivity threshold is also small. This makes the effects of both theoretical and simulated received power results on outage be similar at low threshold. However, as the threshold increases, given the dynamic changes in link distances for simulated environments, the reception of the reflected rays increases. The contribution to the overall received power to avert outage also increases as shown in Fig. 10. The increase in received power due to reflected rays negates outage probability in simulated environments. This causes a sharper rise in outage in our theoretical analysis because effects of reflected rays are ignored. Table 2 summarizes the pros and cons of different HO schemes used. It particularly highlights the computational, environmental and scalability advantage and disadvantages of different HO compared to the proposed scheme.

VI. CONCLUSION

In this work, we investigated the performance of using the deterioration pattern to determine reliable links in 5G mobile networks. An efficient JMLS-MFGT modelling of likely target link variation/pattern was presented to solve the challenges of selecting reliable mmWave links during HO processes in 5G mobile networks. Specifically, mean field-based game theory has been applied to optimize EM estimates in the HO selection matrix. Simulation results show that our proposed joint trained JMLS-MFGT scheme outperforms the existing HO algorithms with robust EE and longer stable connectivity. Thus, we conclude that the proposed JMLS-MFGT HO scheme can be employed as an alternative to classic HO schemes to optimize 5G networks.

REFERENCES

[1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014, doi: 10.1109/JPROC.2014.2299397.

[2] M. Rebato, M. Polese, and M. Zorzi, "Multi-sector and multi-panel performance in 5G mmWave cellular networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6, doi: 10.1109/GLOCOM.2018.8647528.

[3] X. Wang, X. Li, and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges," *IEEE Access*, vol. 3, pp. 1379–1391, 2015, doi: 10.1109/ACCESS.2015.2467174.

[4] H. Munir, H. Pervaiz, S. A. Hassan, L. Musavian, Q. Ni, M. A. Imran, and R. Tafazolli, "Computationally intelligent techniques for resource management in mmWave small cell networks," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 32–39, Aug. 2018.

[5] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan, and M. Zorzi, "Improved handover through dual connectivity in 5G mmWave mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2069–2084, Sep. 2017.

[6] M. Chiputa, M. Zhang, G. G. M. N. Ali, P. H. J. Chong, H. Sabit, A. Kumar, and H. Li, "Enhancing handover for 5G mmWave mobile networks using jump Markov linear system and deep reinforcement learning," *Sensors*, vol. 22, no. 3, p. 746, Jan. 2022, doi: 10.3390/s22030746.

[7] L. Ryzhik. (2018). *Lecture: Mean-Field Game*. [Online]. Available: https://math.stanford.edu/~ryzhik/STANFORD/MEAN-FIELD-GAMES/notes-mean-field.pdf

[8] V. Jain, F. Koehler, and E. Mossel, "The mean-field approximation: Information inequalities, algorithms, and complexity," in *Proc. Conf. Learn. Theory*, vol. 75, Jul. 2018, pp. 1326–1347. [Online]. Available: http://proceedings.mlr.press/v75/jain18b/jain18b.pdf

[9] O. Guéant, J. M. Lasry, and P. L. Lions, "Mean field games and applications," in *Paris-Princeton Lectures on Mathematical Finance 2010*. Berlin, Germany: Springer, 2010, pp. 205–266, doi: 10.1007/978-3-642-14660-2_3.

[10] J. Park, S. Y. Jung, S.-L. Kim, M. Bennis, and M. Debbah, "User-centric mobility management in ultra-dense cellular networks under spatio-temporal dynamics," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6, doi: 10.1109/GLOCOM.2016.7842367.

[11] B. Shubyn and T. Maksymyuk, "Intelligent handover management in 5G mobile networks based on recurrent neural networks," in *Proc. 3rd Int. Conf. Adv. Inf. Commun. Technol. (AICT)*, Jul. 2019, pp. 348–351, doi: 10.1109/AIACT.2019.8847734.

[12] M. Joud, M. García-Lozano, and S. Ruiz, "User specific cell clustering to improve mobility robustness in 5G ultra-dense cellular networks," in *Proc. 14th Annu. Conf. Wireless Demand Netw. Syst. Services (WONS)*, Feb. 2018, pp. 45–50.

[13] Y. Dai, D. Xu, S. Maharjan, and Y. Zhang, "Joint load balancing and offloading in vehicular edge computing and networks," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4377–4387, Jun. 2019.

[14] J. Moon and T. Başar, "Robust mean field games for coupled Markov jump linear systems," *Int. J. Control*, vol. 89, no. 7, pp. 1367–1381, Jul. 2016.

[15] Z. Zhou, J. Feng, Z. Chang, and X. Shen, "Energy-efficient edge computing service provisioning for vehicular networks: A consensus ADMM approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 5087–5099, May 2019.

[16] R. Xu and J. Shi, "ϵ-nash mean-field games for linear-quadratic systems with random jumps and applications," *Int. J. Control*, vol. 94, no. 5, pp. 1415–1425, May 2021, doi: 10.1080/00207179.2019.1651940.

[17] B. Anahtarci, C. D. Kariksiz, and N. Saldi, "Q-learning in regularized mean-field games," 2022, *arXiv:2003.12151v3*. [Online]. Available: https://arxiv.org/abs/2003.12151v3

[18] S. Shalev-Shwartz, "Online learning: Theory, algorithms, and applications," Ph.D. thesis, Dept. Comput. Sci. Eng., Hebrew Univ. Jerusalem, Jerusalem, Isreal, 2007.

[19] D. Bahdanau, P. Brakel, K. Xu, A. Goyal, R. Lowe, J. Pineau, A. Courville, and Y. Bengio, "An actor-critic algorithm for sequence prediction," 2016, *arXiv:1607.07086*.

[20] J. L. Carrera Villacrés, Z. Zhao, T. Braun, and Z. Li, "A particle filter-based reinforcement learning approach for reliable wireless indoor positioning," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2457–2473, Nov. 2019.

[21] Y. Chen, Y. Gao, C. Jiang, and K. J. R. Liu, "Game theoretic Markov decision processes for optimal decision making in social systems," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2014, pp. 268–272.

[22] J. Blumenstein, A. Prokes, J. Vychodil, M. Pospisil, and T. Mikulasek, "Time-varying K factor of the mm-wave vehicular channel: Velocity, vibrations and the road quality influence," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–5.

[23] M. Chiputa, P. H. J. Chong, S. U. Rehman, and A. Kumar, "Investigating mobility robustness in 5G networks using user-adaptive handoff strategies," in *Proc. Int. Conf. Ad Hoc Netw.*, vol. 306. Cham, Switzerland: Springer, 2019, pp. 99–112.

MASOTO CHIPUTA (Student Member, IEEE) received the B.Eng. degree in electrical and electronics engineering from The University of Zambia, Lusaka, Zambia, in 2011, the M.Eng. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2016, and the Ph.D. degree in electrical engineering from the Auckland University of Technology, Auckland, New Zealand, in 2023. He was a Research Assistant with the Auckland University of Technology. His research interests include machine learning and artificial intelligence with emphasis on wireless communication, heterogeneous networks, mmWave communication, and 5G and beyond cellular networks. He was a recipient of the AUT Ph.D. Students Fellowship Scholarship.

MINGLONG ZHANG (Member, IEEE) received the M.S. degree in electronics engineering and computer science from the Peking University, China, and the Ph.D. degree in electronical and electronic engineering from the Auckland University of Technology, New Zealand. From 2021 to 2022, he was a Lecturer and a Research Associate with the Auckland University of Technology. He is currently a Research Assistant Professor with Mississippi State University. His research interests include 5G/6G V2X communications, AI-enabled ORAN, and software-defined radios.

PETER HAN JOO CHONG (Senior Member, IEEE) received the Ph.D. degree from The University of British Columbia, Canada, in 2000. He was an Associate Professor (tenured) with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. From July 2000 and January 2001, he with the Advanced Networks Division, Agilent Technologies Canada Inc., Canada. From February 2001 to May 2002, he was a Research Engineer with the Nokia Research Center, Finland. From 2013 and 2016, he was the Director of the Infinitus, Centre for Infocomm Technology, NTU. He is currently a Full Professor and an Associate Head of the School (Research), School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, New Zealand. He is an Adjunct Professor with the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong, and an Honorary Professor with Amity University, India. He has published more than 300 research articles. His research interests include the areas of mobile communications systems, including MANETs/VANETs, C-V2X, the Internet of Things/Vehicles, AI for wireless networks, and 5G networks. He is a fellow of the Institution of Engineering and Technology (FIET).