

Emerging Small Cell Wireless Technologies for 5G: Architectures and Applications

Lead Guest Editor: Mostafa Zaman Chowdhury

Guest Editors: Takeo Fujii, Gabriel-Miro Muntean, Ji-Woong Choi, and Giuseppe Araniti



Emerging Small Cell Wireless Technologies for 5G: Architectures and Applications

Wireless Communications and Mobile Computing

Emerging Small Cell Wireless Technologies for 5G: Architectures and Applications

Lead Guest Editor: Mostafa Zaman Chowdhury

Guest Editors: Takeo Fujii, Gabriel-Miro Muntean, Ji-Woong Choi,
and Giuseppe Araniti



Copyright © 2018 Hindawi. All rights reserved.

This is a special issue published in "Wireless Communications and Mobile Computing." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Javier Aguiar, Spain	Manuel Fernandez-Veiga, Spain	Antonella Molinaro, Italy
Wessam Ajib, Canada	Gianluigi Ferrari, Italy	Simone Morosi, Italy
Muhammad M. Alam, Portugal	Ilario Filippini, Italy	Kumudu S. Munasinghe, Australia
Eva Antonino-Daviu, Spain	Jesus Fontecha, Spain	Enrico Natalizio, France
Shlomi Arnon, Israel	Luca Foschini, Italy	Wing Kwan Ng, Australia
Leyre Azpilicueta, Mexico	Sabrina Gaito, Italy	Petros Nicopolitidis, Greece
Paolo Barsocchi, Italy	Óscar García, Spain	Giovanni Pau, Italy
Alessandro Bazzi, Italy	Manuel García Sánchez, Spain	Rafael Pérez-Jiménez, Spain
Zdenek Becvar, Czech Republic	L. J. García Villalba, Spain	Matteo Petracca, Italy
Francesco Benedetto, Italy	José A. García-Naya, Spain	Marco Picone, Italy
Olivier Berder, France	A.-J. García-Sánchez, Spain	Daniele Pinchera, Italy
Ana M. Bernardos, Spain	Vincent Gauthier, France	Giuseppe Piro, Italy
Mauro Biagi, Italy	Carlo Giannelli, Italy	Javier Prieto, Spain
Dario Bruneo, Italy	Carles Gomez, Spain	Rüdiger C. Pryss, Germany
Zhipeng Cai, USA	Juan A. Gomez-Pulido, Spain	Junaid Qadir, Pakistan
Claudia Campolo, Italy	Ke Guan, China	Rajib Rana, Australia
Gerardo Canfora, Italy	Daojing He, China	Luca Reggiani, Italy
Rolando Carrasco, UK	Paul Honeine, France	Daniel G. Reina, Spain
Vicente Casares-Giner, Spain	Sergio Ilarri, Spain	Jose Santa, Spain
Dajana Cassioli, Italy	Antonio Jara, Switzerland	Stefano Savazzi, Italy
Luis Castedo, Spain	Minho Jo, Republic of Korea	Hans Schotten, Germany
Lin Chen, France	Shigeru Kashihara, Japan	Patrick Seeling, USA
Hui Cheng, UK	Mario Kolberg, UK	Mohammad Shojafar, Italy
Luca Chiaravaglio, Italy	Nikos Komninos, UK	Giovanni Stea, Italy
Ernestina Cianca, Italy	Juan A. L. Riquelme, Spain	Enrique Stevens-Navarro, Mexico
Riccardo Colella, Italy	Pavlos I. Lazaridis, UK	Zhou Su, Japan
Mario Collotta, Italy	Tuan Anh Le, UK	Luis Suarez, Russia
Bernard Cousin, France	Xianfu Lei, China	Ville Syrjälä, Finland
Igor Curcio, Finland	Martín López-Nores, Spain	Pierre-Martin Tardif, Canada
Donatella Darsena, Italy	Javier D. S. Lorente, Spain	Mauro Tortonesi, Italy
Antonio De Domenico, France	Maode Ma, Singapore	Juan F. Valenzuela-Valdés, Spain
Antonio de la Oliva, Spain	Pietro Manzoni, Spain	Aline C. Viana, France
Gianluca De Marco, Italy	Álvaro Marco, Spain	Enrico M. Vitucci, Italy
Luca De Nardis, Italy	Gustavo Marfia, Italy	Honggang Wang, USA
Alessandra De Paola, Italy	Francisco J. Martinez, Spain	Jie Yang, USA
Oscar Esparza, Spain	Michael McGuire, Canada	Lian Zhao, Canada
Maria Fazio, Italy	Nathalie Mitton, France	
Mauro Femminella, Italy	Klaus Moessner, UK	

Contents

Emerging Small Cell Wireless Technologies for 5G: Architectures and Applications

Mostafa Zaman Chowdhury , Takeo Fujii , Gabriel-Miro Muntean , Ji-Woong Choi , and Giuseppe Araniti 

Volume 2018, Article ID 8969264, 2 pages

A Novel Indoor Mobile Localization System Based on Optical Camera Communication

Md. Tanvir Hossan , Mostafa Zaman Chowdhury, Amirul Islam , and Yeong Min Jang 

Volume 2018, Article ID 9353428, 17 pages

Downlink Linear Precoders Based on Statistical CSI for Multicell MIMO-OFDM

Ebrahim Baktash, Chi-Heng Lin, Xiaodong Wang, and Mahmood Karimi

Volume 2017, Article ID 5981659, 18 pages

Diversity-Multiplexing-Nulling Trade-Off Analysis of Multiuser MIMO System for Intercell Interference Coordination

Jinwoo Kim and Chung G. Kang

Volume 2017, Article ID 5709367, 7 pages

Interactive Smart Fashion Using User-Oriented Visible Light Communication: *The Case of Modular Strapped Cuffs and Zipper Slider Types*

Jai-Eun Kim, Youn-Hee Kim, Jung-Hun Oh, and Ki-Doo Kim

Volume 2017, Article ID 5203053, 13 pages

Relay Positioning Strategy for Traffic Data Collection of Multiple Unmanned Aerial Vehicles Using Hybrid Optimization Systems: A FANET-Based Case Study

José Jailton, Tassio Carvalho, Jasmine Araújo, and Renato Francês

Volume 2017, Article ID 2865482, 11 pages

Reliability-Centric Analysis of Offloaded Computation in Cooperative Wearable Applications

Aleksandr Ometov, Dmitry Kozyrev, Vladimir Rykov, Sergey Andreev, Yuliya Gaidamaka,

and Yevgeni Koucheryavy

Volume 2017, Article ID 9625687, 15 pages

Editorial

Emerging Small Cell Wireless Technologies for 5G: Architectures and Applications

Mostafa Zaman Chowdhury ,¹ Takeo Fujii ,² Gabriel-Miro Muntean ,³ Ji-Woong Choi ,⁴ and Giuseppe Araniti ⁵

¹Kookmin University, Seoul, Republic of Korea

²University of Electro-Communications, Tokyo, Japan

³Dublin City University (DCU), Dublin, Ireland

⁴Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, Republic of Korea

⁵DIIES Department, University Mediterranea of Reggio Calabria, Reggio Calabria, Italy

Correspondence should be addressed to Mostafa Zaman Chowdhury; mzaman@kookmin.ac.kr

Received 18 February 2018; Accepted 18 February 2018; Published 5 April 2018

Copyright © 2018 Mostafa Zaman Chowdhury et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

5th generation (5G) is the next step of mobile telecommunication standards. Researchers all over the world are focusing on the development of 5G communication systems, to be fully available for the users by 2020. This era of wireless communication will bring new exclusive network technologies and service capabilities. Small cell network technologies will be the key driver for the 5G communication. It is expected that 5G communication will consist of ultradense heterogeneous small cell networks. The use of the traditional radio frequency- (RF-) based wireless communication is being saturated to meet the demand of 5G. Hence, researchers are searching for new spectrum for wireless communication. The optical spectrum (infrared, visible light, and ultraviolet) is considered as a promising solution for the development of future high density and high capacity optical wireless communication (OWC) networks. The OWC offers unique advantages such as huge unregulated bandwidth and inherent security. However, the limitation of OWC is the vulnerability to obstacles. Therefore, the coexistence of RF-based omnidirectional and OWC-based directional small cells will be the excellent approach for the targeted multitier ultradense heterogeneous networks in 5G.

The motivation behind this special issue has been to solicit cutting-edge research relevant to architectures and applications of emerging small cell wireless technologies for 5G communication. This special issue invited papers that

address such issues. Following a rigorous review process (including a second review round), 6 outstanding papers have been finally selected for inclusion in the special issue. The accepted papers cover a wide range of research subjects in the broader area of small cell wireless technologies to meet the demand of 5G communications systems.

The paper “Reliability-Centric Analysis of Offloaded Computation in Cooperative Wearable Applications” by A. Ometov et al. brings into perspective the challenges related to heterogeneous communications and offloaded computation operating in cases of fault-tolerant computation, computing, and caching. The authors focus on the emerging augmented reality applications that require reliable delegation of the computing and caching functionality to proximate resource-rich devices.

The paper by E. Baktash et al. entitled “Downlink Linear Precoders Based on Statistical CSI for Multicell MIMO-OFDM” investigates the linear precoder design for downlink multicell multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) systems, where base stations (BSs) coordinate to reduce the interference across space and frequency. The authors apply the random matrix theory to approximate the ergodic weighted sum rate of the system with a closed form expression.

The paper “Diversity-Multiplexing-Nulling Trade-Off Analysis of Multiuser MIMO System for Intercell Interference Coordination” by J. Kim and C. G. Kang explores

the performance trade-off of multicell multiuser MIMO systems for achieving intercell and intracell interference-free conditions. The authors analyze the three-dimensional diversity multiplexing-nulling trade-off among the diversity order, multiplexing order, and nulling order. They also design a precoding matrix to mitigate both intercell and intracell interference for a linear precoding-based MU-MIMO system.

The paper by J.-E. Kim et al. entitled “Interactive Smart Fashion Using User-Oriented Visible Light Communication: The Case of Modular Strapped Cuffs and Zipper Slider Types” presents the design guideline for regulating the technical implementation, design strategies, and directions of interactive LED devices. The technology and design concepts are demonstrated through a case study of an existing LED light-based wearable interaction. They also develop a smart fashion of modular strapped cuffs and zipper slider types that can be attached and detached according to the user’s preferences.

The paper “Relay Positioning Strategy for Traffic Data Collection of Multiple Unmanned Aerial Vehicles Using Hybrid Optimization Systems: A FANET-Based Case Study” by J. Jailton et al. proposes a flight path planning model, which involves a metaheuristic optimization-based approach. The proposal relied on the artificial neural networks to optimize the positioning of the relay device, so that the throughput between the other devices could be increased.

The paper by M. T. Hossan et al. entitled “A Novel Indoor Mobile Localization System Based on Optical Camera Communication” proposes a localization technique for smartphones in indoor environments. This technique can calculate the coordinates of a smartphone using existing illumination infrastructure with LEDs. The system can locate smartphones without further modification of the existing LED light infrastructure.

Acknowledgments

We would like to thank all the authors who submitted their excellent research articles to this special issue and all the reviewers for providing their valuable and timely feedback through the review process, which helped to improve the quality of this special issue. The work of M. Z. Chowdhury was supported by the Korea Research Fellowship Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2016H1D3A1938180).

*Mostafa Zaman Chowdhury
Takeo Fujii
Gabriel-Miro Muntean
Ji-Woong Choi
Giuseppe Araniti*

Research Article

A Novel Indoor Mobile Localization System Based on Optical Camera Communication

Md. Tanvir Hossan , Mostafa Zaman Chowdhury, Amirul Islam , and Yeong Min Jang 

Department of Electronics Engineering, Kookmin University, Seoul, Republic of Korea

Correspondence should be addressed to Yeong Min Jang; yjang@kookmin.ac.kr

Received 23 September 2017; Accepted 10 January 2018; Published 25 February 2018

Academic Editor: Alessandra De Paola

Copyright © 2018 Md. Tanvir Hossan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Localizing smartphones in indoor environments offers excellent opportunities for e-commerce. In this paper, we propose a localization technique for smartphones in indoor environments. This technique can calculate the coordinates of a smartphone using existing illumination infrastructure with light-emitting diodes (LEDs). The system can locate smartphones without further modification of the existing LED light infrastructure. Smartphones do not have fixed position and they may move frequently anywhere in an environment. Our algorithm uses multiple (i.e., more than two) LED lights simultaneously. The smartphone gets the LED-IDs from the LED lights that are within the field of view (FOV) of the smartphone's camera. These LED-IDs contain the coordinate information (e.g., x - and y -coordinate) of the LED lights. Concurrently, the pixel area on the image sensor (IS) of projected image changes with the relative motion between the smartphone and each LED light which allows the algorithm to calculate the distance from the smartphone to that LED. At the end of this paper, we present simulated results for predicting the next possible location of the smartphone using Kalman filter to minimize the time delay for coordinate calculation. These simulated results demonstrate that the position resolution can be maintained within 10 cm.

1. Introduction

Nowadays, the number of mobile devices (i.e., smartphone) is dramatically increasing. These devices have an immense scope for commercial applications, for example, services, e-commerce, and e-banking. To move forward with consumer-facing commercial applications, location based services (LBS) for smartphones need to be improved. As the location of the mobile device is unpredictable; a reliable, dynamic, accurate, and situation-adaptive localization technique is required for LBS [1]. Moreover, this technique should be secure and interruption free. This LBS approaches will not increase consumer activity without a reliable localization system. The localization scheme design should be considered for indoor and outdoor environments. Indoor localization is most promising for e-commerce applications, as most of them are indoor based. The access density for smartphones is the highest in shopping malls, super markets, and transit stations (i.e., railway, bus, and subway). Indoor environments are hubs for almost all web based business applications.

Outdoor localization schemes are also promising and may be improved with insights from indoor localization solution.

Both industry and academic research institutes have recently shown interest in the issue of indoor smartphone localization, and various schemes have been proposed [2, 3]. The most common and widely used indoor localization system relies on the global positioning system (GPS) [4]. Moreover, this system has three particular limitations (e.g., poor GPS signal reception, loss of GPS signal, and limited localization accuracy) [5] especially in indoor environments. This system is not suitable for underground or indoor localization. Since the signal from the satellites to a receiver should be line-of-sight (LOS), and building, soil, water, trees, or even poor weather conditions inhibit this signal. Time of flight (TOF) cameras are another possible candidate for solving the localization scheme [6]. Besides its advantages, this system is too expensive and sometimes requires complex scenarios for implementation which makes it inappropriate for unique approach. TOF camera also has some other drawbacks; it is only useful for detection and ranging purposes and

does not facilitate the communication purpose [7]. Received signal strength indication (RSSI), time of arrival (TOA), time difference of arrival (TDOA), and angle of arrival (AOA) are physical parameters of radio signal that can be used for localization with special distributed monitors [8, 9]. Other approaches used for indoor localization include computer vision (CV) and artificial intelligence (AI) [10]. None of these approaches are ideal for indoor localization. For example, RSSI depends on environmental conditions. It is affected by shadowing, path loss, signal fading, and interference from neighboring cells [11]. Therefore, errors can be included in the calculated value. On the contrary, narrowband signals, reduced data transmission rate, and lower location precision can be inhibited in TDOA parameter [12]. Other approaches like combining RSSI and AI cannot mitigate the challenges in indoor localization because subjective and objective data from AI have an impact on new input data [13]. On the basis of system performance improvement, it may provide feedback to the input of the system. Concurrently, it will cause a huge impact on new data if input variables are varying fast. Another approach, photogrammetry, measures object location from the photographs. This is a very simple way to generate a map from sequentially taken photographs [14].

To deal with the existing challenges for indoor localization system, optical frequency band in electromagnetic spectrum is a novel candidate that avoids the problems typical with radio frequency (RF). The communication technology where the optical frequency is used is known as optical wireless communication (OWC). A promising subsystem of OWC is optical camera communication (OCC), where multiple LEDs are used as transmitters and a camera or image sensor (IS) is used as receiver [15]. In an indoor environment, the communication channel for OCC is uninterrupted, signal-to-noise (SNR) is high, security is ensured by LOS communication, simple signal processing is confirmed, and high-speed communication is possible [16]. OCC has multiple-input-multiple-output (MIMO) functionality [17] which is ideal for analyzing many objects at the same time. Therefore, OCC is a good technique for indoor localization systems. Moreover, OCC is useful not only for localization but also for communication. Since 2011, the IEEE has formed a new working group to finalize the standardization specification (i.e., IEEE 802.15.7) [18]. The development of the standardization specification for OCC will be finalized by mid-2018.

In the proposed scheme, smartphones are located through signal from LED light using OCC and photogrammetry. Using the OCC techniques, smartphone cameras receive the LED-ID signals from each LED light (or fixture). This LED-ID resembles coordinate (x, y) of the LED lights. The size of the image of the LED fixture on the image sensor of the camera changes with the relative distance between the smartphone and LED fixture. This distance is calculated using photogrammetry which helps to find the remaining coordinate of the smartphone. To optimize the resolution, Kalman filter is applied to accurately predict the next possible locations of the smartphone.

The rest of paper is organized as follows. In Section 2, we survey several literatures on indoor localization system. Our proposed scheme is illuminated in Section 3. In Section 4,

the channel modeling and communication are explained. The processes for distance calculation and localization are stated in Section 5. In Section 6, the performance of our proposed system using Kalman filter is evaluated. Finally, we summarize our work and future research direction in Section 7.

2. Related Works

In this section, several indoor localization schemes are summarized. OWC based indoor localization can be classified in three ways: triangulation, fingerprinting, and proximity-based approaches. A geometric position is required in triangulation. Fingerprinting is scene analysis approach and proximity-based approach is grid method. In [19], vision-based positioning and navigation are introduced with the help of a camera and newly defined 3D map in an indoor environment. The authors tried the possible ways to improve the accuracy and system reliability through several experiments. They mainly focused on different applications, for example, mobile robot navigation, transportation, visitor guiding, security, and emergency services. In the same manner, the authors in [20] propose a vision-based system using a camera as the main sensor for object detection, tracking, and localization for distributed sensing, communication, and parallel computing. They used camera nodes as inputs of sensor fusion techniques to reduce computer system encumbrances such as Extended Kalman Filter (EKF) and Maximum Likelihood (ML). They also test investigation for moving objects. The drawback of their proposed system is the lacking of automatic determination of sensor location. For an indoor environment, a vision-based navigation system using augmented reality was proposed in [21]. They recognize a location automatically using image sequences which is taken in the indoor environment and then realized augmented reality by flawlessly superimposing the user's view with detailed location information. They use wearable mobile PC with a camera for taking image sequences and it transmits the images to remote PCs for future processing. Using their proposed system, the average location recognition success rate was found around 89%. Moreover, its performance will deteriorate with harsh environmental scenarios. In [22], neural network and OCC based indoor positioning system is proposed. They estimate camera position for trained and untrained environments. The error for estimated camera position is less than 10 mm, which can increase up to 200 mm. However, the error depends upon the camera position. In [23], the authors demonstrated a system for positioning and orientation of overlying the location information on camera phones in an indoor environment. They find a location from images with the help of the existing standard hardware. They process less data within very short time to generate accurate data. The limitations of their proposed system in handling the whole environment is that they must improve the feature detection and match while maintaining low latency. Their proposed system is complex and very costly for navigation on a standard camera phone. In [24], the authors propose radiometry and camera geometry for OCC camera model. They consider a camera model for the indoor environment at 50 cm separation for LEDs and the distance between LEDs is 200 cm.

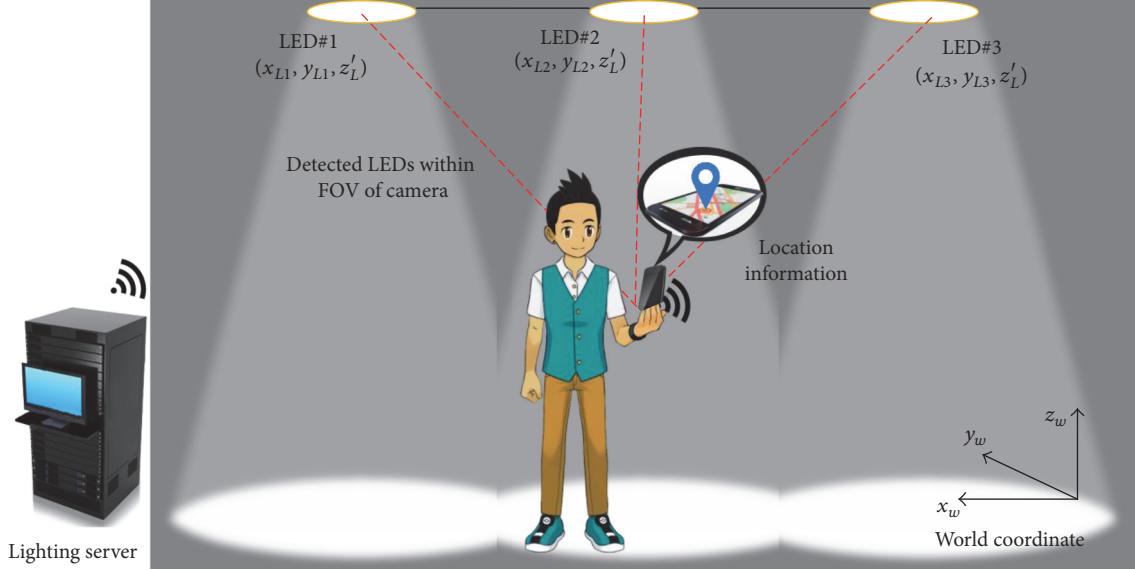


FIGURE 1: Overall architecture of the proposed indoor localization scheme.

3. Overall Architecture

We propose a localization scheme in which smartphones can be located in an indoor environment with the help of LED lights and the smartphone's camera. In our proposed scheme, we consider several factors to identify smartphone locations in indoor environments. As in Figure 1, all LED lights fixtures are attached to the ceiling, the distance between ceiling and floor is constant for particular indoor environment, the camera of smartphone must be under the illumination of the LED light fixture, and there must be at least three LED lights in the field of view (FOV) of the camera, while the smartphone maintains continuous communication with the lighting server. A lighting server provides API-based access to reproducible, web-based visualizations. The FOV of a camera is the solid angle through which the IS can sense electromagnetic radiation such as visible light. The system performance improves if the number of LED lights within the camera FOV is increased and our algorithm requires at least three light fixtures to deliver accurate location measurements. For our system, the LED lights fixtures coordinates are parallel to the world coordinate system. The vertical distance between ceiling and floor (i.e., z -coordinate) is equal for all LED lights where the z -coordinate of light fixture is in the inverse direction to the world z -coordinates. However, the camera coordinates of the smartphone change frequently with respect to the LED light fixture coordinates.

As shown in Figure 2, each LED light broadcasts its own coordinate information (i.e., x - and y -coordinate) as a modulated LED-ID signal. The z -coordinate is the same for each and every LED light in a certain indoor environment. Therefore, the term of z -coordinate in the LED-ID is ignored to reduce the complexity and data packet size. These LED-IDs are transmitted as a modulated optical signal. After receiving the signal from the LED lights, smartphones process the data in two different ways. They identify LED-IDs from received

signals and measure the distance of the LED light with the corresponding LED-ID using photogrammetry. This distance is measured by calculating the size of the light fixture and number of pixels of the corresponding LED light on the IS. The geometric image size of the LED on the IS varies with the distance between the light source and camera. If the LED is located far away from the camera, the size of the image is smaller on the IS and comparatively a large image will generate for the LED light which is near to the camera.

LED-IDs with the corresponding distance calculation are sent to the lighting server by the smartphone via a wireless fidelity (Wi-Fi) access point (AP). The coordinate information for each LED lights is stored in the lighting server. After receiving a signal from the smartphone, the lighting server matches the LED-ID signal with its stored location information. Then the algorithm can map the location of the smartphone.

Meanwhile, the location of the smartphone may change during this processing time. Therefore, the lighting server uses Kalman filter tracking algorithm which predicts the next possible location from the current location of the smartphone. This location information is sent to the smartphone. Additionally, the LED light's projected image on the IS changes as the smartphone moves. Therefore, the placement interval of the LED lights should be constant distance to ensure that at least three images of LED lights are available from any location of the room.

4. Channel Modeling and Communication

4.1. Propagation Model of Light from LEDs. The radiation pattern from LED lights is affected by the roughness of the chip faces and geometry of the encapsulating lens. Several models are used to describe the direction of light strength from a light source (i.e., LED). A popular approach is Monte

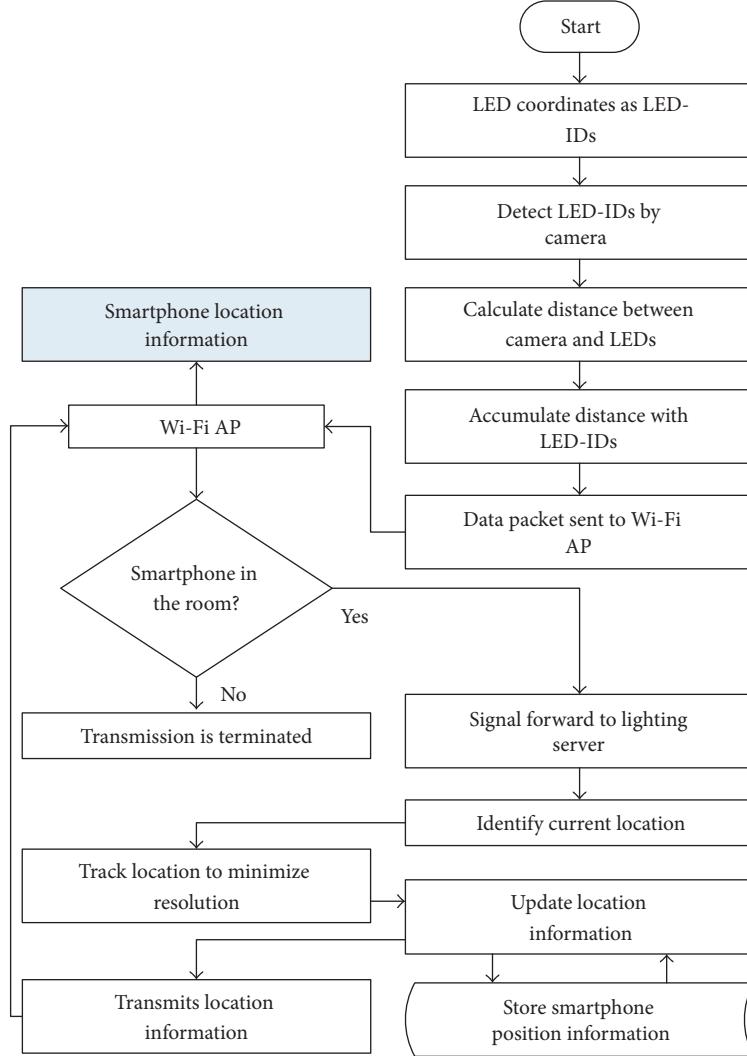


FIGURE 2: Flowchart for the proposed scheme.

Carlo ray tracing [25]. In a Gaussian (or cosine) power distribution [26], the ray of light is diffusely reflected or refracted. The final radiation pattern of light should appear linear super which are angularly shifted in function of the angle of incidence of every traced ray due to diffusely reflection or refraction characteristic shown in Figure 3.

The energy flux per solid angle known as luminous intensity and transmitted optical power are the two basic properties of a light source such as LED light. The luminous intensity is given as

$$I(\phi) = I(0) \cos^m(\phi), \quad (1)$$

where ϕ is the luminous flux and $I(0)$ is the center luminous intensity of an LED. This luminous flux can be defined as integration between minimum wavelength, λ_1 , and maximum wavelength, λ_2 , of working optical spectrum

$$\phi = K_m \int_{\lambda_1}^{\lambda_2} V(\lambda) \phi_e(\lambda) d\lambda, \quad (2)$$

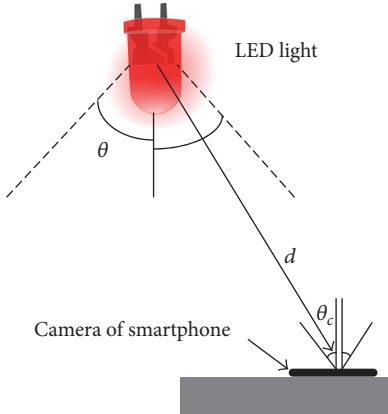
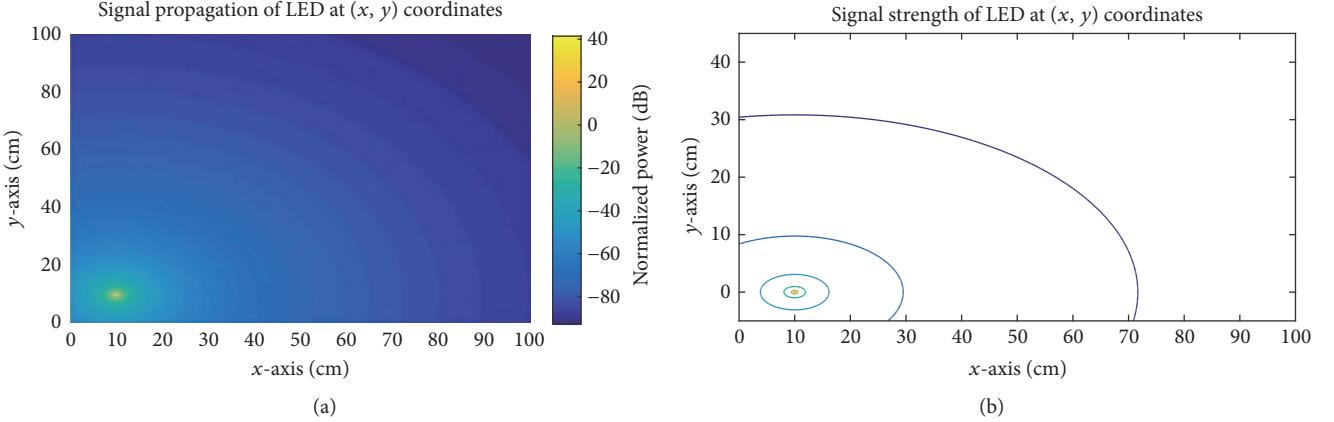


FIGURE 3: Geometric model for LOS communication.

where $V(\lambda)$ is the standard luminosity curve and K_m is the maximum spectral efficacy for vision. The integral of the

FIGURE 4: LED light (a) propagation and (b) average strength at (x, y) coordinates.

energy flux ϕ_e in all directions is the transmitted optical power P_t , given as

$$P_t = \int_0^{2\pi} \phi_e d\theta d\lambda. \quad (3)$$

Figure 4(a) shows the LED light propagation direction and Figure 4(b) shows the strength of the radiation from a single light source. The power level is maximum 40 dBm at the center of the light source which is represented as yellow in Figure 4(a) and deteriorates to -80 dBm represented as violet, from the center to the edge of the sphere. Joining average power strengths of light at certain x - and y -coordinate is forming ellipse shape in Figure 4(b). Therefore, weak signal from neighboring light sources causes signal interference. This problem can be mitigated by removing the background noise with a full control on camera shutter speed.

4.2. LED-ID for Optical Camera Communication. Each LED light has a fixed location and the coordinate information for each single LED light is different from the other LED lights in the same room. The coordinates of the LED light are parallel to the world coordinates. These coordinates are analogues to the LED-ID. Every LED light transmits its own ID to the camera which can be declared as a digital tag. For this purpose, the LED acts as a transmitter and camera as a receiver.

The data from the LED light is sent as a modulating signal by varying the intensity of the light using LED driver integrated circuit (IC). This driver controls the light intensity by dimming the LED through a variety of methods. Data from the LED light can be encoded in the phase of LED light signal and phase can be changed by turning the LED light on and off. However, turning the LED light on is not always possible after turning the lights off fully. Therefore, it is recommended to dim the light at minimum intensity. Here, this modulation is known as IM/DD (Intensity Modulated/Direct Detection) modulation [27]. The available modulation technique to transmit signal from LED light sources can be classified as follows:

- (i) On-off keying (OOK) [28]: the two logic signals in a digital transmission “1” and “0” are represented as

high and zero voltage at the transmitter end. This is achieved by applying flickering illumination of the LED light to represent the on-off state of the transmitter.

- (ii) Pulse width modulation (PWM) [29]: the modulated signal from the LED light is transmitted in the form of a square wave. The desire level of the pulse is obtained by adjusting the LED light dimming.
- (iii) Pulse-position modulation (PPM) [30]: light from the LED encoded message is transmitted through a single pulse in one-time shifts.
- (iv) Orthogonal frequency division multiplexing (OFDM) [31]: data is sent as parallel substreams of modulated data using multiple orthogonal subcarriers in a channel.
- (v) Frequency-shift keying (FSK): a modulated digital signal can be carried by the instantaneous frequency shifting with a constant amplitude.
- (vi) Phase-shift keying (PSK) [32]: a digital signal can be carried by the instantaneous phase shifting of the baseband signal.

For high speed data transmission channels, OFDM is used where the possibility of multipath fading and intersymbol interference is high [33]. Despite the enormous advantage afforded by OFDM, we do not implement it, because high speed data transmission is not required for most indoor localization applications. In these cases, OOK is a better choice for data transmission from LEDs. Figure 5 shows how, after encoding and modulating the light, the LED light fixture transmits data to the camera of the smartphone. Our system decodes LED coordinates, after demodulating and decoding the received signal from the image processor.

4.3. Channel Modeling for OCC. The pixel E_b/N_0 of the model of IS [34] can be calculated as

$$\text{Pixel } \frac{E_b}{N_0} = \frac{E[\rho^2]}{E[n^2]} \approx \frac{s^2 \Delta}{\alpha s \Delta + \beta}, \quad (4)$$

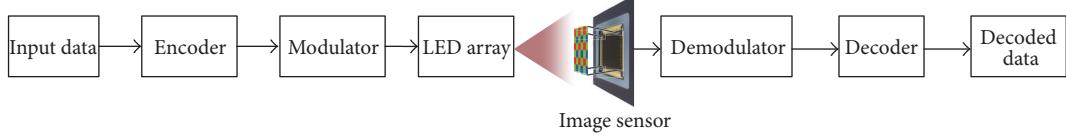


FIGURE 5: Basic architecture of optical camera communication.

where E_b represents the energy-per-bit, N_0 represents the spectral-noise-density, ρ is the unit pixel value, s is the amplitude of the signal, Δ is the camera exposure duration as a ratio of the signal cycle, n is a noise term, and α, β are model fitting parameters for system noise.

Considering the distortion in the channel, the signal-to-noise-plus-interference ratio (SNIR) can be determined as

$$\text{SNIR} = \frac{\sigma P_{\text{avg}}^2}{(1 - \sigma) P_{\text{avg}}^2 + \sigma_n^2}, \quad (5)$$

where P_{avg} is the average pixel intensity transmitted from the LED light, σ is the average distortion factor whose value lies between 1 and 0: that is, $0 \leq \sigma \leq 1$. $\sigma = 1$ indicates minimum signal lost where the LED light is focusing directly on the camera lens and $\sigma = 0$ indicates that no image pixel is generated on the image sensor of the camera.

With the additive white Gaussian noise (AWGN) characteristic on the camera channel [35], the channel capacity of the space time modulation can be expressed by Shannon capacity formula as

$$C = F_{fps} W_s \log_2 (1 + \text{SNIR}), \quad (6)$$

where F_{fps} is the frame rate of the smartphone camera and W_s is the spatial bandwidth, which represents how much information is carried by the pixels in each image frame. The spatial bandwidth is equivalent to the number of orthogonal or parallel channels in a MIMO system.

The bit error rate (BER) which depends on the SNIR and the modulation scheme measures the impact of the channel. The noise sources that affect the transmission of the light signal from the LEDs are intersymbol interference, background and transmitter LED shot noise, and thermal noise.

Contemporary smartphones use complementary metal oxide semiconductor (CMOS) based IS where shutter type is rolling shutter. With this shutter technique, light intensities on the IS are captured row by row and the whole image is composed of different pixel array. Therefore, the exposed time delay between pixel array lines records the changing state of illumination of the LED light as a group of pixels in one image. The optical channel DC gain $H(0)$ models the channel characteristic from LED lights to the camera which can be determined [36] as

$$H(0)$$

$$= \begin{cases} \frac{(m+1) A_{\text{cam}}}{2\pi d^2} \cos^m(\phi) T_s(\theta) \cos(\theta), & 0 \leq \theta \leq \theta_c \\ 0, & \theta \geq \theta_c, \end{cases} \quad (7)$$

where m is the order of Lambertian emission, A_{cam} is the area on IS, d is the distance between an LED and IS, $T_s(\theta)$ is the optical filter coefficient for signal transmission, θ is the angle of incidence, and θ_c is camera FOV semiangle. Here, m can be defined as

$$m = -\frac{\ln 2}{\ln (\cos_{\theta_c/2})}. \quad (8)$$

It is required to consider the channel noise N (which is independent of signal characteristic) and the number of LED-ID signal n to calculate the channel output u as follows:

$$v = R \sum_{i=1}^n h_i x_i + N, \quad (9)$$

where $u = [u_1 \ u_2 \ \dots \ u_n]^T$ and R is the camera, responsivity.

The average optical power of the IS of the camera can be calculated as

$$P_r = \frac{I(0) \cos^m(\phi) \cos(\psi)}{d^2}. \quad (10)$$

5. Distance Calculation and Localization

5.1. Distance Calculation between LED Light and Camera. The fundamental operation of a camera is diagrammed in Figure 6 where an image of a target LED light is projected on the IS of a camera. Light from the target LED passes through the camera lens and is concentrated on the IS plate. The projected image on the IS plate is an inverted image of LED light fixture.

Consider that f is the focal length of the camera, d is the distance from the camera lens to the target LED light, and e is the distance from the focal length to the projected image on the IS. Therefore, we can write

$$\frac{e}{d} = \frac{f}{d-f}. \quad (11)$$

The magnification of the lens is the ratio of the projected image size to the geometric size of LED light. If the camera projects a square image on the IS where the height and width of the target LED and the projected image are (a, b) and (a_i, b_i) , respectively, then the lens magnification can be expressed as

$$M = \frac{a_i}{a} = \frac{b_i}{b} = \frac{e}{d} = \frac{f}{d-f}. \quad (12)$$

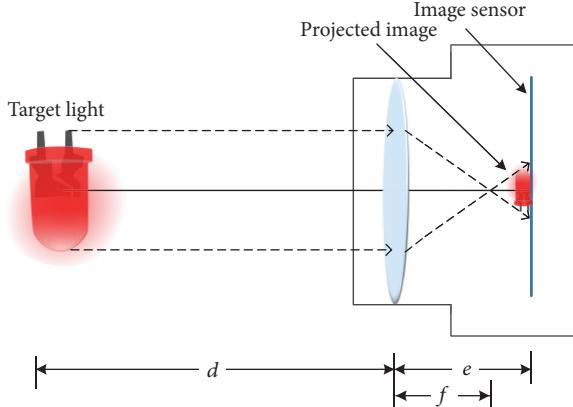


FIGURE 6: Illustration of projected image on the image sensor.

If we ignore the loss in the optical channel, then $d \gg f$. More precisely, $d - f \approx d$. By combining (11) with (12), we get

$$a_i b_i = M^2 ab. \quad (13)$$

The number of pixels of image is the ratio of the projected image size on the IS to the unit pixel area of the same sensor. If the number of pixels is η_i on the IS, ρ^2 is the unit pixel area of IS, and S is the area of the target LED light source, then we obtain

$$\eta_i = \frac{f^2 S}{\rho^2 d^2}. \quad (14)$$

Different shapes of light fixture are available in the market. However, rectangular/square and circle are typical light fixture shapes. For circular shape LED light fixture, if a is the radius of the circle, then the area of the light source will become $S = \pi a^2$. On the contrary, if a and b are the width and length of rectangular light sources, respectively, then the area of the light source will be $S = ab$.

If we only consider the LED lights within the smartphone FOV, the distances from the LED light to the smartphone camera are different for each and every light. The projected image from LED light facing the camera straight-on is larger than the image of an LED light located at an angle to the same camera. These distances can be determined by (14) which must be modified due to the relative motion between the camera and LED light fixtures. For each camera, the focal length is f , and the unite pixel area of IS is fixed. On the contrary, if we know the real-physical area of the lighting fixture, then we can write (14) as

$$d = \tau \sqrt{\frac{1}{\eta_i}}, \quad (15)$$

where $\tau = f(\sqrt{S}/\rho)$ is constant for each camera and LED light.

Figure 7 shows how the image areas are varying with the distance of the LED lights from the camera. From (15), the distance from IS to the LED light fixture is inversely proportional to the square root of the image area on the IS.

5.2. Scenario of Image Area on the IS for Dynamic of Camera.

The smartphone is the only moving device with respect to the LED lights in indoor environments. The FOV of the camera shifts with changes in the smartphone's location. Therefore, the images of LED lights on the IS of camera also change. Additionally, the lighting infrastructure must be designed in a way that keeps at least three LED lights within the FOV of each camera. Figure 8 describes a scenario in which the LED lights detected by the camera vary from 4 to 5 due to the movement of the smartphone from location 1 to location 5.

Square blocks, black dots, and circles represent LED lights, camera, and the FOV of the camera, respectively. For location 1 of the smartphone, five LED lights images are projected on IS. The number of projected images varies from 5 to 4 when the location of the smartphone changes from location 1 to location 5. If more than three LED lights images are projected on the IS, location of the smartphone can be calculated more accurately.

Next, we consider another scenario in which the smartphone moves from location 1 to location 2, as shown in Figure 9(b). Due to the change of smartphone location, the image of the LED lights on the IS also changes. We consider three possible location of LED lights (e.g., red, green, and blue), all attached to the ceiling. For location 1 of the smartphone in Figure 9(c), the blue LED light is relatively close to the camera. This means that the angle is smaller from camera to the blue LED light. This angle is large for the red LED light compared with the blue LED light. On the contrary, this angel is nearly zero for the green LED light. These distances are calculated using (15).

We describe the straight-line distances of red, green, and blue LED lights from lens of the camera which are d_{r1} , d_{g1} , and d_{b1} , respectively. Furthermore, η_{ir1} , η_{ig1} , and η_{ib1} are the number of pixels on the IS for red, green, and blue light fixtures, respectively, at location 1. Therefore, (15) can be written for each LED light as

$$d_{r1} \propto \frac{1}{\sqrt{\eta_{ir1}}}, \quad (16)$$

$$d_{g1} \propto \frac{1}{\sqrt{\eta_{ig1}}}, \quad (17)$$

$$d_{b1} \propto \frac{1}{\sqrt{\eta_{ib1}}}. \quad (18)$$

From Figure 9(c), the size of image on the IS is larger for the green LED light. This size gradually decreases from the blue to red LED lights. Therefore, comparing one particular image with the other image on IS can be mathematically represented as $\eta_{ig1} > \eta_{ib1} > \eta_{ir1}$. As we know from (15), the number of pixels projected on the IS for a certain object depends only on the distance between the camera and the object when other factors remain constant. The mathematical expression for each and every image on the IS and (16)–(18) gives a conclusion: $d_{r1} > d_{b1} > d_{g1}$.

At location 2 in Figure 9(a), the angular distance between the camera and the blue LED light fixture increases compared with the red LED light. Therefore, the image size increases accordingly. In contrast, the image size of the green LED light is almost identical due to the small shift in angle.

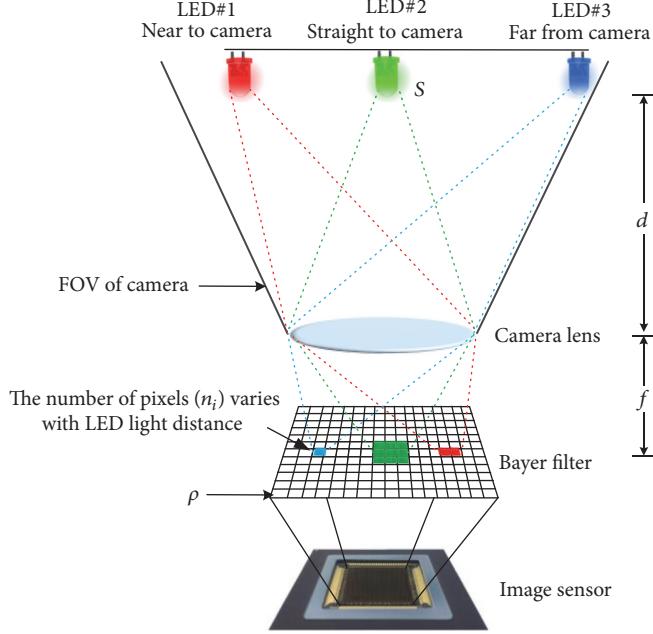


FIGURE 7: The projected image of three LEDs on the image sensor has been deleted.

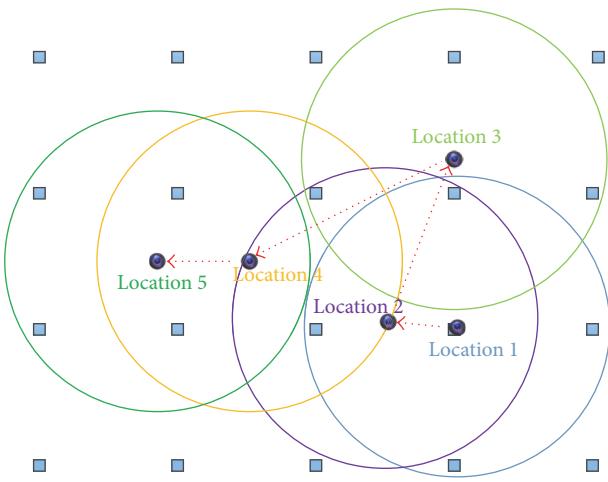


FIGURE 8: Change of considered LED lights within the FOV of camera with the change of smartphone.

For location 2, the distance from LED lights to the camera lens are denoted as d_{r2} , d_{g2} , and d_{b2} for red, green, and blue LEDs, respectively. Moreover, the numbers of pixels on the IS are labeled as η_{ir2} , η_{ig2} , and η_{ib2} . Therefore, if we apply (15) for three LED light images, we will get

$$d_{r2} \propto \frac{1}{\sqrt{\eta_{ir2}}}, \quad (19)$$

$$d_{g2} \propto \frac{1}{\sqrt{\eta_{ig2}}}, \quad (20)$$

$$d_{b2} \propto \frac{1}{\sqrt{\eta_{ib2}}}. \quad (21)$$

From Figure 9(a), the general expression for the number of image pixels is $\eta_{ig2} > \eta_{ir2} > \eta_{ib2}$ on the IS. From (19)–(21), the distance can be explained as $d_{b1} > d_{r1} > d_{g1}$.

5.3. Uploading Information to the Lighting Server. The smartphone begins to send distance information with the corresponding LED-ID to the lighting server via Wi-Fi AP. This information is sent as a packet with two slots where the first slot has the coordinate information of the LED light and second slot has its distance information. The information of LED light coordinates is already stored in the lighting server. After receiving a signal from the smartphone, the lighting server generates a virtual map of LED lights from extracting information from the packet. With the mathematical model of trilateration (or multilateration for more than three LED lights), the lighting server calculates the location of the smartphone.

5.4. Computing Smartphone Location. The beam of LED light propagates from ceiling to floor and illumination spreads 360° from the center of a light source. At a specific distance from the LED light, the intensity of light illumination is equal. If the smartphone is located at any of these locations, the projected image area on the IS will be found the same with the other projected images around that circle. Therefore, the chance of false location identification is higher for a single LED light source. Figure 10(a) shows the original location of the camera with the probable false location of the camera (in a faded image). This false location information creates errors during location mapping in the lighting server.

Consequently, to mitigate location-estimation errors when only a single LED light is visible, we consider another LED light as a reference for the first LED light. From Figure 10(b), we narrow down the location to the torus formed by the intersection of the two spheres. Though more

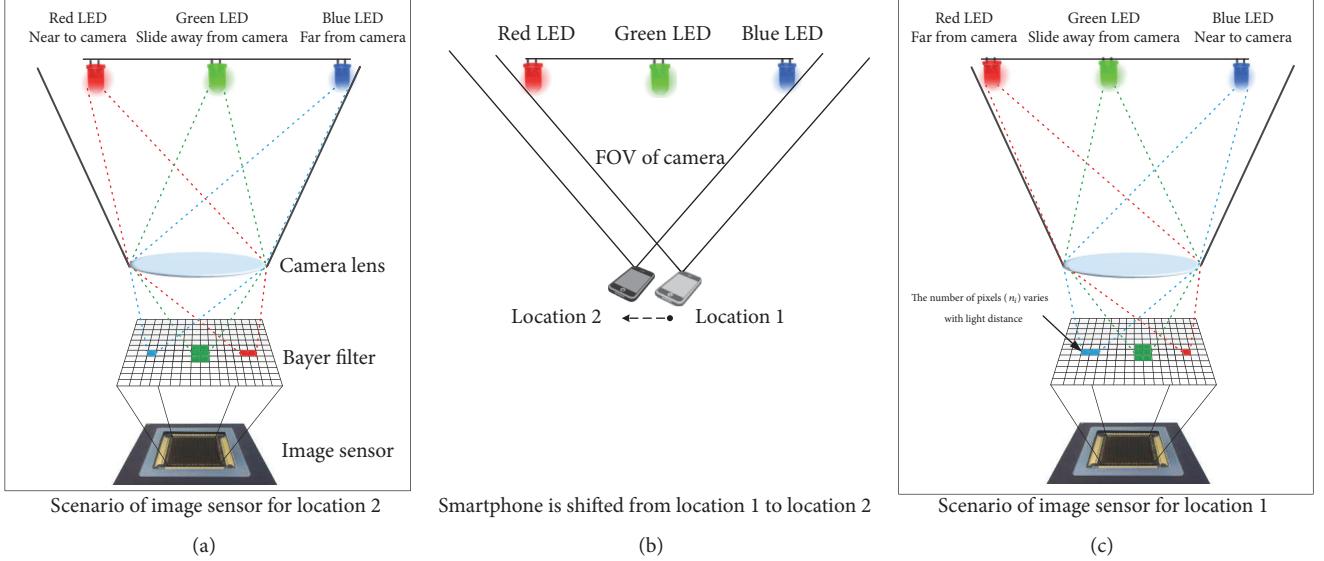


FIGURE 9: Scenarios of pixel size on IS for the movement of smartphone from location 1 to location 2; (a) position 2; (b) shift of FOV of the camera; and (c) location 1.

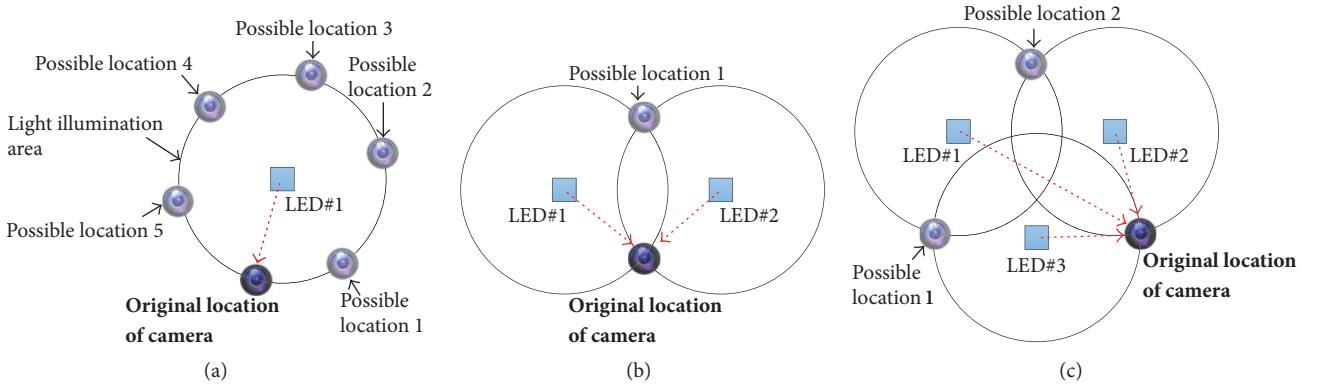


FIGURE 10: Narrowing down localization estimation error using (a) single LED light, (b) two LED lights, and (c) three LED lights.

LED lights are available in this scenario, the duplication still occurs. Hence, a third LED light is necessary to deliver accurate location measurements.

With a third LED light, we can narrow the location of the smartphone to one possible location. In Figure 10(c), three circles centered on each of the landmarks overlap at three different locations where the radius of each circle is equidistant from each landmark. Therefore, two other locations of the smartphone along with the original are still possible even with three landmarks. Moreover, other location information may not arise any confusion for smartphone position estimation. It is possible to estimate the smartphone's location accurately by comparing information from two LED lights with information from a third light.

The method to determine the location of smartphone using three fixed reference points (or LED lights) is known as

trilateration [37] or more than three points (which is known as multilateration). For trilateration, the measuring platform is simultaneous with three relevant nonlinear equations. The reference LED lights can be situated either in a triangle or in a straight line from each other.

If $P_i = (x_{Lj}, y_{Lj}, z'_{Lj})$ is the coordinates of any LED light under the ceiling, where $j = 1, 2, 3$, and the coordinate information of the smartphone camera is $P_c = (x, y, z)$, then the distance from the camera to the LED light can be represented with the following equation:

$$d_j^2 = (x - x_{Lj})^2 + (y - y_{Lj})^2 + (z - z'_{Lj})^2. \quad (22)$$

A matrix can be generated with the above equations for $i = 1, 2, 3$ as

$$\begin{bmatrix} 1 & -2x_{L1} & -2y_{L1} & -2z'_{L1} \\ 1 & -2x_{L2} & -2y_{L2} & -2z'_{L2} \\ 1 & -2x_{L3} & -2y_{L3} & -2z'_{L3} \end{bmatrix} \begin{bmatrix} x^2 + y^2 + z^2 \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1^2 & -x_{L1}^2 & -y_{L1}^2 & -z'_{L1}^2 \\ d_2^2 & -x_{L2}^2 & -y_{L2}^2 & -z'_{L2}^2 \\ d_3^2 & -x_{L3}^2 & -y_{L3}^2 & -z'_{L3}^2 \end{bmatrix}. \quad (23)$$

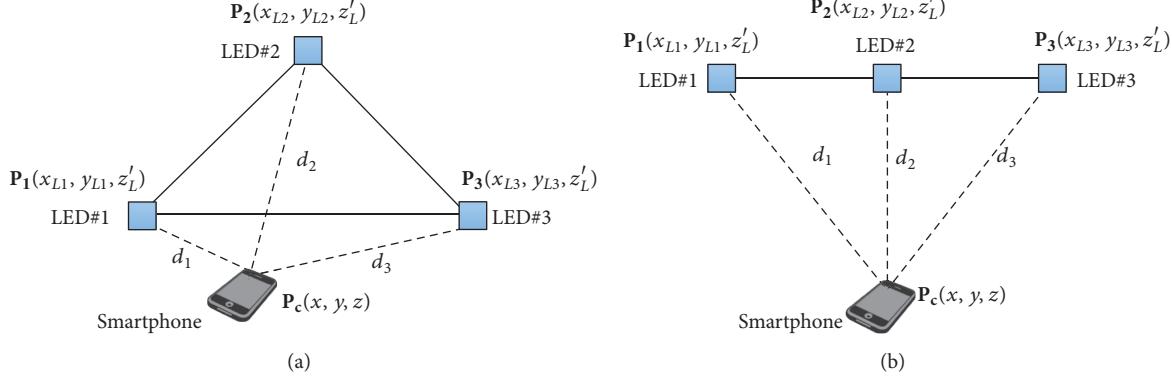


FIGURE 11: Determination of camera location from reference LED lights located at (a) the vertexes of triangle and (b) different points on a straight line.

The matrix equation can be replaced as

$$Z_0x = q_0. \quad (24)$$

Two different cases can occur when solving the trilateration problem for locating the smartphone. The LED lights can be distributed randomly as in Figure 11(a) or aligned in a straight line as shown in Figure 11(b).

To identify the location of a smartphone from the reference LED lights located at the vertexes of a triangle, the general solution of (24) can be expressed as

$$x = x_p + tx_h, \quad (25)$$

where x_p is denoted as the particular solution and t is the real parameter. If $Z_0x = 0$ is a homogeneous system, then x_h is its solution.

The matrix Z_0 is written as pseudoinverse matrix format to determine the solution for x_p . On the other hand, the value of t can be evaluated using the expression of $x_p = [x_{p0}, x_{p1}, x_{p2}, x_{p3}]^T$, $x_h = [x_{h0}, x_{h1}, x_{h2}, x_{h3}]^T$, and $x = [x_0, x_1, x_2, x_3]^T$.

The following solution can be generated after solving (25) as

$$x_1 = x_p + t_1 x_h, \quad (26)$$

$$x_2 = x_p + t_2 x_h. \quad (27)$$

To identify the location of a smartphone from reference LED lights located in a straight line, the general solution of (24) is expressed as

$$x = x_p + tx_{h1} + \kappa x_{h2}, \quad (28)$$

where homogeneous system $Z_0x = 0$; x_{h1} and x_{h2} are two solutions with real parameters κ .

The mathematical expression of (24) is different for the case with more than three LED lights. The solution can be found by solving multilateration problem. The relevant equation can be expressed as follows:

$$Zx = q. \quad (29)$$

On the base of the least squares methods, the solution of (29) can be found as

$$\hat{x} = (Z^T Z)^{-1} Z^T q. \quad (30)$$

In Figure 12, three LED lights are located at three points (i.e., P_1 , P_2 , and P_3) in a two-dimensional plane. Their illumination spheres intersect at two points (i.e., P_{C1} and P_{C2}) that are possible locations for the smartphone camera. The lighting server chooses between the multiple possible location of the smartphone using trilateration.

Figure 13 shows that systematic estimation error is generated when the lighting server estimates the position of the smartphone. The error is minimum for the horizontal bias (in Figure 13(a)) of the indoor environment and is much higher for the vertical bias (in Figure 13(b)). Therefore, system performance is much degraded when measuring vertical position.

Around a cluster of three LED lights, the possible locations of the smartphone are shown in Figure 15. Dotted lines represent the optical links between the camera and the LED lights and solid lines represent the fixed distances between the LED lights. In almost all cases, the distance from the smartphone to each LED light is different. In some cases (Figures 14(b), 14(c), 14(d), and 14(g)), the distance to two LED lights is equal compared with the distance from the other LED light. Additionally, there are few cases (Figures 14(e), 14(f), and 14(h)) where all three distances are different from each other. Concurrently, there is only one case (Figure 14(a)) where camera is equidistant from all LED lights. The algorithm can locate smartphones at these locations without error.

Figure 15 shows an example of the final stage of a server-side process for estimating the smartphone coordinates. Three LED lights are imaged within the FOV of the camera. The distance between each LED light is equal and in our tests this value is 150 cm. The LED light coordinates are $P_1(200, 0)$, $P_2(200, 150)$, and $P_3(200, 300)$, all in cm. Here, the x -coordinates are the same for these three LED lights but the y -coordinates are all different. We chose these coordinates to

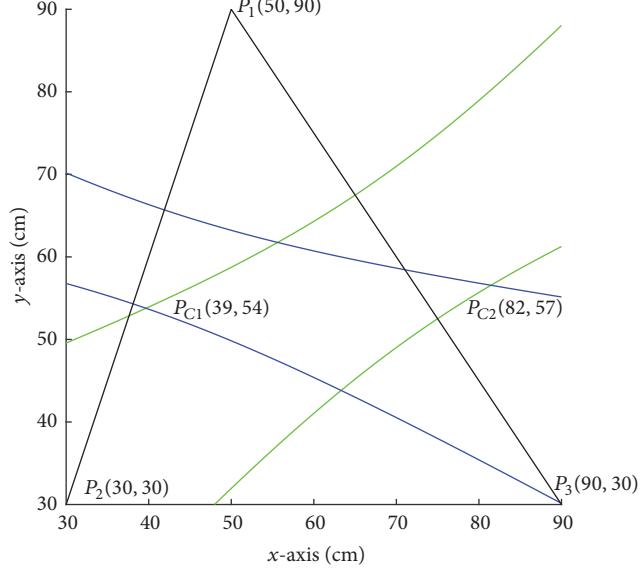


FIGURE 12: Trilateration using three LEDs in 2D plane.

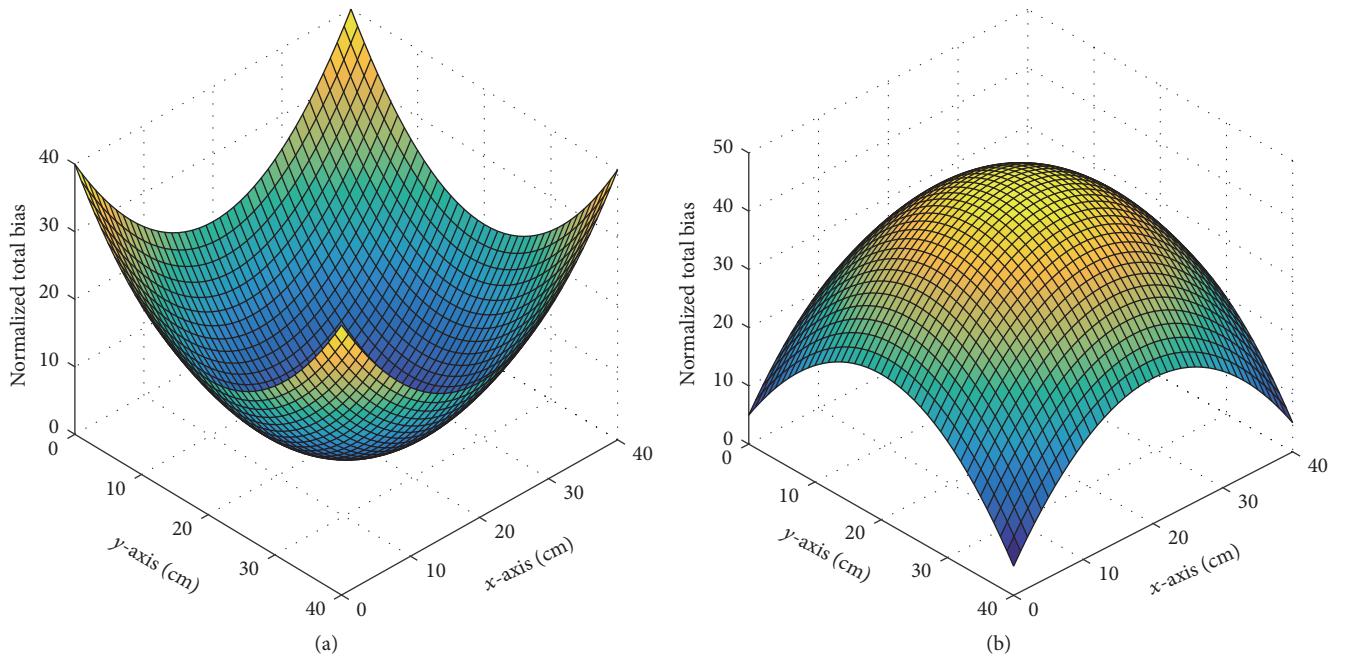


FIGURE 13: Systematic errors in (a) horizontal and (b) vertical direction.

simplify this example. The z -coordinate is equal for all the LED lights, so we ignore it in our calculations.

Let us consider a smartphone placed between $P_1(200, 0)$ and $P_2(200, 150)$ and far away from $P_3(200, 300)$. More precisely, this smartphone is closer to $P_1(200, 0)$ than $P_2(200, 150)$. The distances from the camera to $P_1(200, 0)$, $P_2(200, 150)$, and $P_3(200, 300)$ are 320 cm, 336.05 cm, 410.37 cm, respectively, which are measured by calculating the image sizes on the image sensor. The relative distances from these three LED lights show that the smartphone y -coordinate is 40 cm away from $P_1(200, 0)$ and is 110 cm

away from $P_2(200, 150)$. In this example, we consider that $P_1(200, 0)$ is located at the origin and the y -coordinate of the smartphone is calculated with respect to it. Here, 40 cm is the y -coordinate of the smartphone. The z -coordinate of the smartphone camera can be measured with the Pythagorean theorem. Therefore, the z -coordinate of smartphone can be calculated as 317.5. Finally, the estimated coordinate is $P_C(200, 40, 318)$.

5.5. Estimate the Next Location of the Smartphone. After finalizing the smartphone coordinates, the lighting server

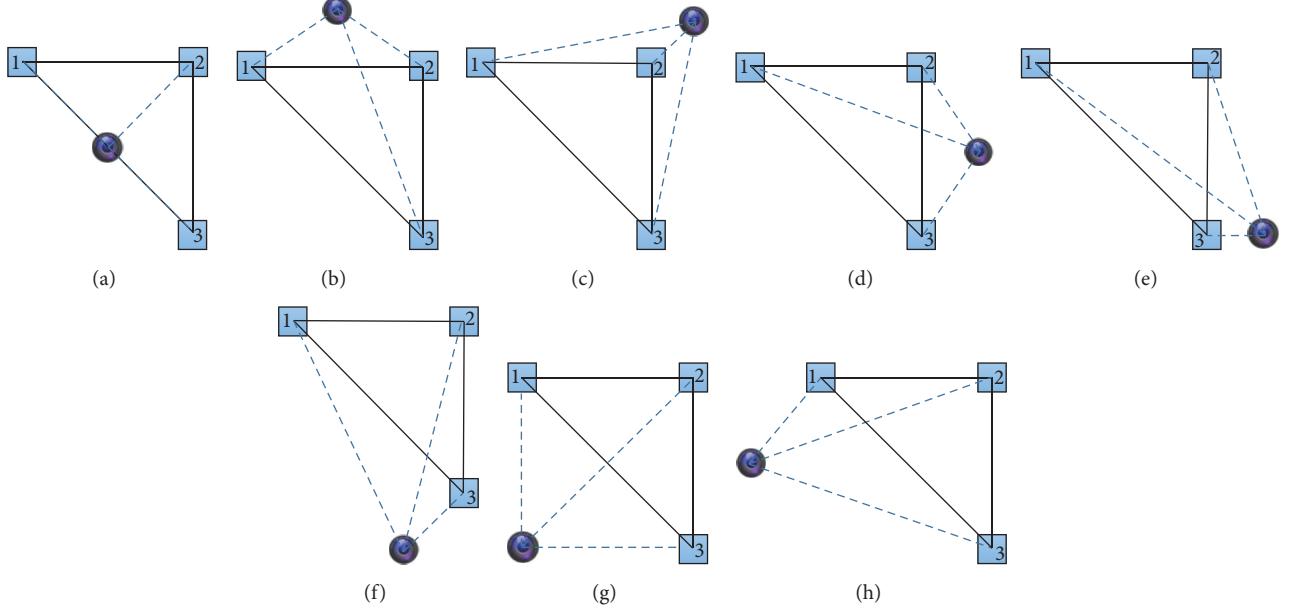


FIGURE 14: Relative distance from LED lights to camera of smartphone; ((b), (c), (d), (g)) two equal distances; ((e), (f), (h)) different distance; and (a) equal distance.

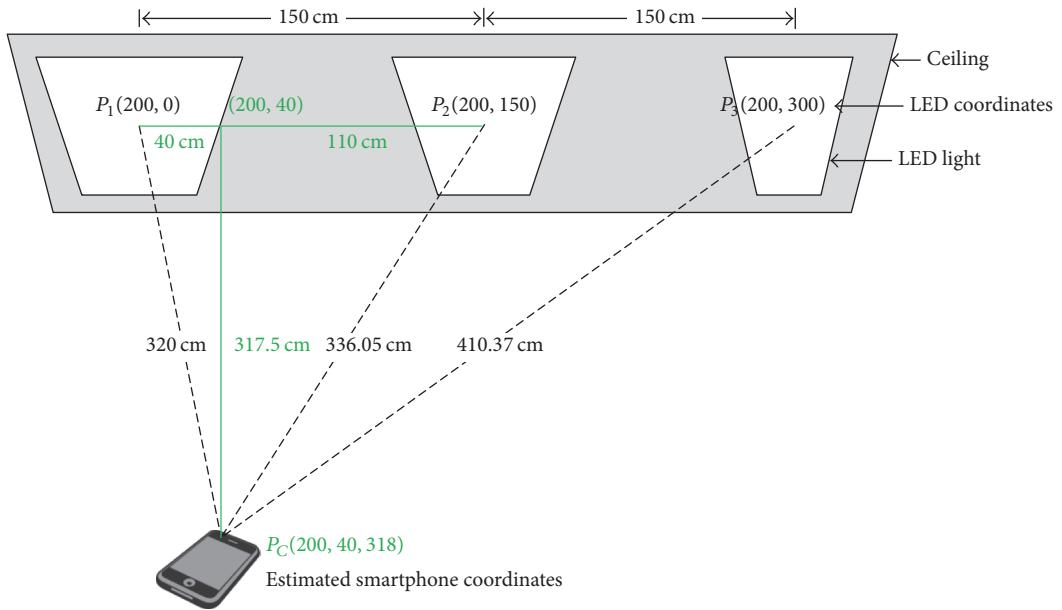


FIGURE 15: Lighting server estimates the coordinates of the smartphone.

sends coordinate information to the smartphone, updates this information, and stores the information for future use. The location of the smartphone is always changing. While the server estimates the smartphone position, the smartphone may have moved. It is required to run another server-side algorithm in parallel to estimate the velocity, acceleration, and next possible position of the smartphone. We use a Kalman filter to track the next position of the smartphone. This filter depends on the present input measurement instead

of previous information (e.g., velocity and acceleration) from the smartphone [38].

Kalman filter is a recursive estimator and linear filter mostly used to approximate errors in navigation applications with minimum variance estimate in a least squares sense under noise processes. Kalman filter gain, current estimation, and new error in the estimation are three important calculations in Figure 16. The Kalman filter gain places the special importance on the error in the estimate and the

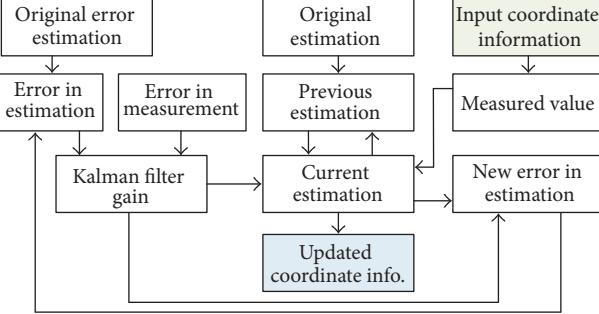


FIGURE 16: Flow diagram of Kalman filter.

error in the measurement. On the other hand, the current estimation depends on the previous estimation and the present measured value. The relative importance between previous estimate and present measured value is also fixed by the Kalman filter gain. Furthermore, Kalman filter gain and current estimation are needed to know the new error in the estimate which is passed onto error in future estimation. The preliminary estimated location of the smartphone can be described as

$$X_{k_p} = BX_{k-1} + w_k, \quad (31)$$

where X_{k-1} is the initial location of the smartphone, B is the state (or adaptation) matrix, and w_k is noise added to the initial location.

The measurements and state vectors are weighted by their respective processes' covariance matrices. The process covariance matrix (or error in the position estimation) can be represented as

$$P_{k_p} = BP_{k-1}B^T + Q_k, \quad (32)$$

where the initial process covariance matrix is P_{k-1} and Q_k is noise.

The filter deweights the measured value during large variance and low gain in comparison to the state estimate. This situation leads the filter to prioritize the prediction state rather than measurements. In different circumstances, the measured value is weighted more over the predicted value due to the small variance and high gain. The gain of the Kalman filter is known as the Kalman gain which depends on the error in the estimate and error in the measurement. Kalman gain, K_g , is the ratio of the error in estimate to the total error in both the estimate and measurement,

$$K_g = \frac{P_{k_p}H^T}{HP_{k_p}H^T + R}, \quad (33)$$

where R is the observation or measured error and H is transformation matrix, which converts the covariance matrix into Kalman filter gain matrix. The value of Kalman gain lies between 0 and 1 (i.e., $0 \leq K_g \leq 1$). If K_g is near to 1, it means error in the measurement is nearly 0. In this estimation,

TABLE 1: Physical parameter for the simulated scenarios.

<i>Camera parameters</i>	
FOV	120°
Focal length	5 mm
Pixel size	1 μm
Image size	640 × 320 pixels
Pixel edge length	7.1e - 3 mm
Frame rate	30 fps
Lens aperture	4
<i>Parameters for transmitter</i>	
LED diameter	170 mm
LED area	22,700 mm²
Half power emission	1500 mW
Radiation semiangle	20°
Size of LED panel	10 × 10 cm²
Modulation method	OOK

the estimates are unstable (large error in the estimate) and measurements are accurate.

The error in location estimate will decrease when the value of K_g is close to 0. Therefore, the difference between estimation and actual is narrowed down. The expression for current estimation can be written as

$$X_k = X_{k_p} + K_g [Y_k - HX_{k_p}], \quad (34)$$

where X_k is the present estimate, X_{k_p} is the previous estimate, and Y_k is the measured smartphone coordinate.

Similarly, if the Kalman filter gain is large, then the present error in the estimate is small. The new predicted state can be defined as follows:

$$P_k = [I - HK_g] P_{k_p}. \quad (35)$$

5.6. Postponed Signal Propagation. The smartphone location measurement and signal propagation stops if the smartphone user leaves the room. To recognize this situation, the lighting server broadcasts a message several times and waits for a reply. In the event of no reply, the server stops sending data and stores the position information.

6. Simulation Result

To evaluate the performance of our proposed scheme, we used a smartphone in 1600 sqft. indoor environment. The test instrument specifications are provided in Table 1. The simulation result will vary with the variation of the camera and luminaire parameters.

Figure 17 shows a graphical representation between BER versus SNIR for theoretical and simulation results. Both curves are almost merged together because we ignore the effect of channel noise in our simulation. It explains that BER is increased with the decrease for SNIR of OOK signaling.

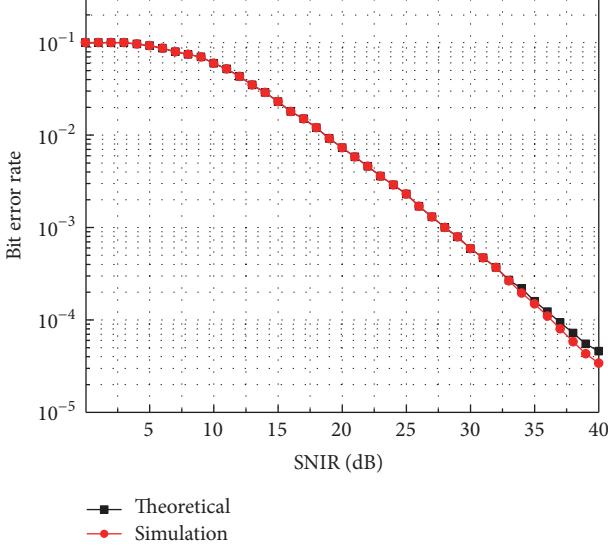


FIGURE 17: Simulation BER versus theoretical BER with SNIR for OOK signaling.

In Figure 18, the simulation result shows that initially the location of the smartphone is not identified and is mentioned as unassigned. Estimation accuracy was not good enough for the 1st estimation compared with the 4th estimation. This estimation process is improved sequentially after a few steps. Concurrently, the distance between each estimation is kept between 9 and 10 cm.

The possibility of changing coordinates in the z -direction is negligible. Therefore, we only have to calculate the x - and y -coordinates of the smartphone. In Figure 19, the green line shows the mean value of the smartphone location and red line is the estimated value. The location estimation using Kalman filter for x -axis is plotted in Figure 19(a) and y -axis of smartphone is plotted in Figure 19(b). Figure 19 states that there is a deviation of location estimation from the mean value of the location. We consider 1 Hz sampling rate and run time is 50 sec. Overall, 50 samples were considered for simulation.

Distance measurement using OCC depends on the size of the projected image of the LEDs on the IS. With increasing the distance, the projected image on the IS occupies less area rather than a shorter distance. Therefore, the possibility of smartphone localization is shrunk if the vertical distance between the smartphone and the LED lights at the ceiling is increased. In Figure 20, when the vertical distance from the camera to ceiling is remaining within 10 m, the image occupied area is greater than or equal to 4-pixel area. After 10 to 35 m, localization possibility is reduced due to decreasing pixel occupied area ($4 > \eta_i \geq 1$). Theoretically, it is required to occupy at least unit pixel area of an image sensor. However, it is difficult to ensure that the projected image merges with edge-by-edge of a pixel. Therefore, after 35 m, localization possibility is zero because pixel occupied area remained $\eta_i < 1$. We consider a fixed transmitter size and in that case, its image occupies less than unit pixel area after 35 m. If we

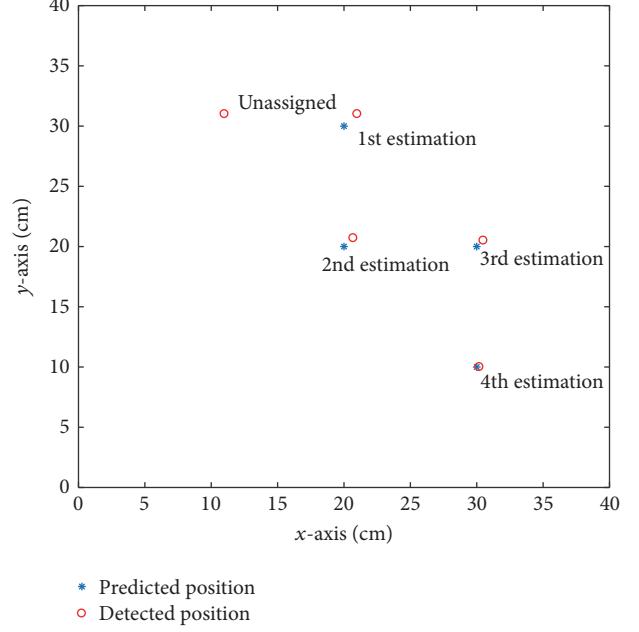


FIGURE 18: Predicting smartphone location using Kalman filter with the detected locations.

change the transmitter size, then the distance measurement performance will be changed.

Localization estimation error occurs due to change of smartphone position frequency. We test our algorithm both with Kalman filter and without Kalman filter. Then we generate a plot for comparing the significance between them in Figure 21. A significant deviation on performance was found from the figure. At the initial stage of measurement, both show the same percentage of errors in estimation, whereas in both cases, estimation errors are exponentially decreasing with simulation runtime. At 10 sec period, estimation error is near zero for Kalman filter based estimation. Meanwhile, at the same time, another estimation (i.e., without Kalman filter) shows 50% error in estimation.

7. Conclusion

In this paper, we proposed a smartphone localization system for an indoor environment. Using OCC for smartphone localization is a novel idea. We also use photogrammetry technique along with OCC. The localization resolution for the smartphone is kept within 10 cm. The proposed system relies upon a central processing lighting server for positioning calculations. Signaling from LED light fixtures and localization of smartphone are kept within certain indoor environment. Therefore, this localization scheme is more secure. Additionally, chance of error in the position estimation is more for the system where the implication of Kalman filter is ignored. We included Kalman filter to track the next possible location of the smartphone. Thus, the proposed scheme is more accurate than the existing localization scheme. These lighting fixtures are useful not only for localization but also for illumination for the user. In future work, we will test and evaluate

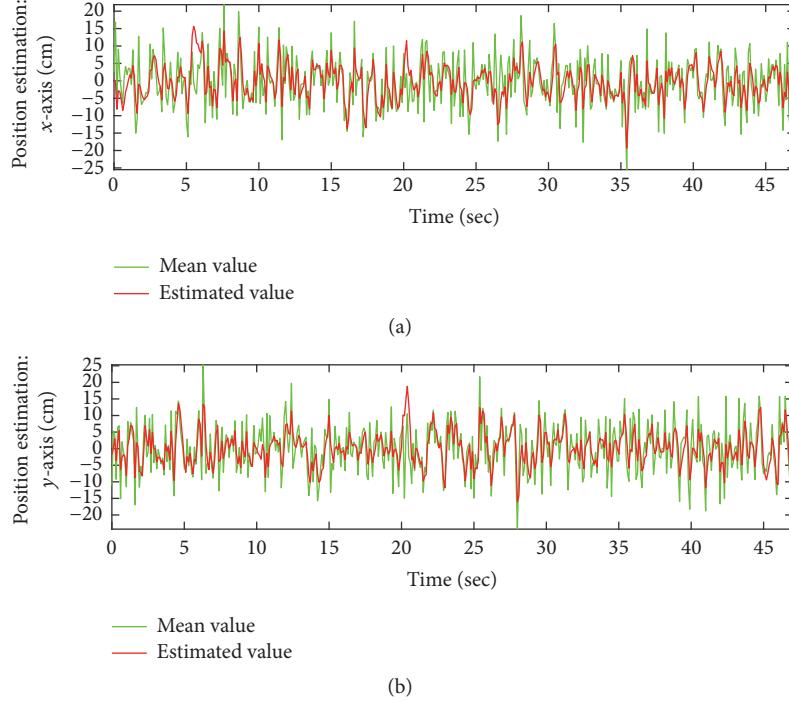


FIGURE 19: Tracking the mean value of the smartphone location using Kalman filter at the direction of (a) x -axis and (b) y -axis.

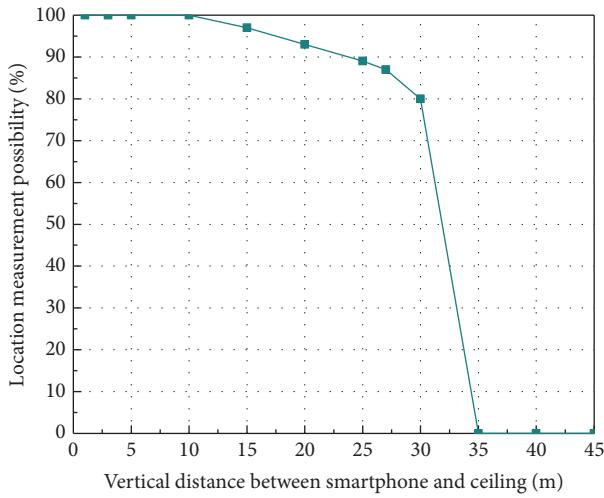


FIGURE 20: Comparison between indoor vertical distance between smartphone and ceiling versus localization measurement possibility.

the performance in different environmental scenarios (e.g., escalator and staircase). We will consider variation in height between smartphone and ceiling light fixtures. Meanwhile, we are also trying to optimize the position identification resolution without using Kalman filter to make the system simpler.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

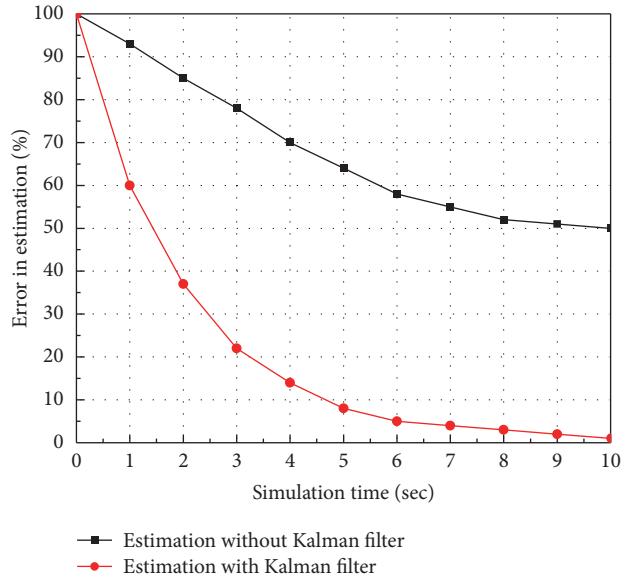


FIGURE 21: Simulation time versus error in estimation on the basis of using Kalman filter.

Acknowledgments

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Global IT Talent Support Program (IITP-2017-0-01806) supervised by the IITP (Institute for Information and Communication Technology Promotion).

References

- [1] A. Yassin, Y. Nasser, M. Awad et al., "Recent advances in indoor localization: a survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327–1346, 2017.
- [2] K. A. Nuaimi and H. Kamel, "A survey of indoor positioning systems and algorithms," in *Proceedings of the International Conference on Innovations in Information Technology (IIT '11)*, pp. 185–190, Abu Dhabi, UAE, June 2011.
- [3] G. Kul, T. Özyer, and B. Tavli, "IEEE 802.11 WLAN based real time indoor positioning: literature survey and experimental investigations," *Procedia Computer Science*, vol. 34, pp. 157–164, 2014.
- [4] B. H. Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice*, Springer, Wien, Austria, 5th edition, 2001.
- [5] A. Kleusberg and R. B. Langley, "The limitations of GPS," *GPS World*, vol. 1, no. 2, 1990.
- [6] P. Zanuttigh, G. Marin, C. Dal Mutto, F. Dominio, L. Minto, and G. M. Cortelazzo, *Time-of-Flight and Structured Light Depth Cameras: Technology and Applications*, Springer International Publishing, 1st edition, 2016.
- [7] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: a survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011.
- [8] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, 2007.
- [9] Z. Farid, R. Nordin, and M. Ismail, "Recent advances in wireless indoor localization techniques and system," *Journal of Computer Networks and Communications*, vol. 2013, Article ID 185138, 12 pages, 2013.
- [10] S.-H. Jung, B.-C. Moon, and D. Han, "Unsupervised learning for crowdsourced indoor localization in wireless networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2892–2906, 2016.
- [11] K. Heurtefeux and F. Valois, "Is RSSI a good choice for localization in wireless sensor network?" in *Proceedings of the 26th IEEE International Conference on Advanced Information Networking and Applications (AINA '12)*, pp. 732–739, Fukuoka, Japan, March 2012.
- [12] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 41–53, 2005.
- [13] Q. Liu, S. Levinson, Y. Wu, and T. Huang, "Interactive and incremental learning via a mixture of supervised and unsupervised learning strategies," in *Proceedings of the 5th Joint Conference on Information Sciences (JCIS '00)*, vol. 1, pp. 555–558, March 2000.
- [14] K. Kraus, *Photogrammetry: Geometry from Images and Laser Scans*, vol. 1, Walter de Gruyter, 2nd edition, 2007.
- [15] M. Uysal, C. Capsoni, Z. Ghassemlooy, A. Boucouvalas, and E. Udvaray, *Optical Wireless Communications: An Emerging Technology*, Springer International Publishing, Cham, Switzerland, 2016.
- [16] N. Saha, M. S. Iftekhar, N. T. Le, and Y. M. Jang, "Survey on optical camera communications: challenges and opportunities," *IET Optoelectronics*, vol. 9, no. 5, pp. 172–183, 2015.
- [17] P. Luo, M. Zhang, Z. Ghassemlooy et al., "Experimental demonstration of RGB LED-based optical camera communications," *IEEE Photonics Journal*, vol. 7, no. 5, pp. 1–12, 2015.
- [18] A. C. Boucouvalas, P. Chatzimisios, Z. Ghassemlooy, M. Uysal, and K. Yiannopoulos, "Standards for indoor optical wireless communications," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 24–31, 2015.
- [19] X. Li, J. Wang, A. Olesk, N. Knight, and W. Ding, "Indoor positioning within a single camera and 3D maps," in *Proceedings of the Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS '10)*, pp. 1–9, Kirkkonummi, Finland, October 2010.
- [20] J. M. S. Matamoros, J. R. M. Dios, and A. Ollero, "Cooperative localization and tracking with a camerabased WSN," in *Proceedings of the IEEE International Conference on Mechatronics (ICM '09)*, pp. 1–6, Malaga, Spain, April 2009.
- [21] J. Kim and H. Jun, "Vision-based location positioning using augmented reality for indoor navigation," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 3, pp. 954–962, 2008.
- [22] M. S. Iftekhar, N. Saha, and Y. M. Jang, "Neural network based indoor positioning technique in optical camera communication system," in *Proceedings of the 5th International Conference on Indoor Positioning and Indoor Navigation (IPIN '14)*, pp. 431–435, Busan, South Korea, October 2014.
- [23] H. Hile and G. Borriello, "Positioning and orientation in indoor environments using camera phones," *IEEE Computer Graphics and Applications*, vol. 28, no. 4, pp. 32–39, 2008.
- [24] M. S. Iftekhar, M. A. Hossain, C. H. Hong, and Y. M. Jang, "Radiometric and geometric camera model for optical camera communications," in *Proceedings of the 7th International Conference on Ubiquitous and Future Networks (ICUFN '15)*, pp. 53–57, Sapporo, Japan, July 2015.
- [25] G. J. Ward, F. M. Rubinstein, and R. D. Clear, "A ray tracing solution for diffuse interreflection," in *Proceedings of the 15th International Conference on Computer Graphics and Interactive Techniques*, pp. 85–92, August 1988.
- [26] I. Moreno and C.-C. Sun, "Modeling the radiation pattern of LEDs," *Optics Express*, vol. 16, no. 3, pp. 1808–1819, 2008.
- [27] S. Hranilovic and F. R. Kschischang, "Optical intensity-modulated direct detection channels: signal space and lattice codes," *IEEE Transactions on Information Theory*, vol. 49, no. 6, pp. 1385–1399, 2003.
- [28] S. Rajagopal, R. D. Roberts, and S.-K. Lim, "IEEE 802.15.7 visible light communication: modulation schemes and dimming support," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 72–82, 2012.
- [29] F. Vasca and L. Iannelli, *Dynamics and Control of Switched Electronic Systems*, Springer, London, UK, 1st edition, 2012.
- [30] E. D. J. Smith, R. J. Blaikie, and D. P. Taylor, "Performance enhancement of spectral-amplitude-coding optical CDMA using pulse-position modulation," *IEEE Transactions on Communications*, vol. 46, no. 9, pp. 1176–1185, 1998.
- [31] J. Armstrong, "OFDM for optical communications," *Journal of Lightwave Technology*, vol. 27, no. 3, pp. 189–204, 2009.
- [32] X. Song, F. Yang, and J. Cheng, "Subcarrier intensity modulated optical wireless communications in atmospheric turbulence with pointing errors," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 5, no. 4, pp. 349–358, 2013.
- [33] J. Armstrong, "Analysis of new and existing methods of reducing intercarrier interference due to carrier frequency offset in OFDM," *IEEE Transactions on Communications*, vol. 47, no. 3, pp. 365–369, 1999.

- [34] T. Nguyen, A. Islam, T. Hossan, and Y. M. Jang, "Current status and performance analysis of optical camera communication technologies for 5G networks," *IEEE Access*, vol. 5, pp. 4574–4594, 2017.
- [35] I. Djordjevic, W. Ryan, and B. Vasic, *Coding for Optical Channels*, Springer, New York, NY, USA, 2010.
- [36] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997.
- [37] F. Thomas and L. Ros, "Revisiting trilateration for robot localization," *IEEE Transactions on Robotics*, vol. 21, no. 1, pp. 93–101, 2005.
- [38] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, 2003.

Research Article

Downlink Linear Precoders Based on Statistical CSI for Multicell MIMO-OFDM

Ebrahim Baktash,¹ Chi-Heng Lin,² Xiaodong Wang,³ and Mahmood Karimi¹

¹School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

²Department of Statistics, Columbia University, New York, NY 10027, USA

³Department of Electrical Engineering, Columbia University, New York, NY 10027, USA

Correspondence should be addressed to Xiaodong Wang; wangx@ee.columbia.edu

Received 21 September 2017; Accepted 6 December 2017; Published 31 December 2017

Academic Editor: Mostafa Zaman Chowdhury

Copyright © 2017 Ebrahim Baktash et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With 5G communication systems on the horizon, efficient interference management in heterogeneous multicell networks is more vital than ever. This paper investigates the linear precoder design for downlink multicell multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) systems, where base stations (BSs) coordinate to reduce the interference across space and frequency. In order to minimize the overall feedback overhead in next-generation systems, we consider precoding schemes that require statistical channel state information (CSI) only. We apply the random matrix theory to approximate the ergodic weighted sum rate of the system with a closed form expression. After formulating the approximation for general channels, we reduce the results to a more compact form using the Kronecker channel model for which several multicarrier concepts such as frequency selectivity, channel tap correlations, and intercarrier interference (ICI) are rigorously represented. We find the local optimal solution for the maximization of the approximate rate using a gradient method that requires only the covariance structure of the MIMO-OFDM channels. Within this covariance structure are the channel tap correlations and ICI information, both of which are taken into consideration in the precoder design. Simulation results show that the rate approximation is very accurate even for very small MIMO-OFDM systems and the proposed method converges rapidly to a near-optimal solution that competes with networked MIMO and precoders based on instantaneous full CSI.

1. Introduction

Multicell multiple-input multiple-output orthogonal frequency-division multiplexing (MIMO-OFDM) is a promising technology for next-generation telecommunication networks. Both MIMO and OFDM are known to facilitate the wireless networks with excellent capabilities; when MIMO and OFDM are combined together, MIMO-OFDM achieves tremendous capacities that will be enjoyed by 5G systems in the near future. While MIMO-OFDM systems have superb performance, several challenges invoked by MIMO and OFDM characteristics are involved in the system design [1–6]. Other competing technologies in 5G networks are filtered-OFDM (F-OFDM) and universal filtered multicarrier (UFMC) that are receiving attention in the literature [7, 8].

MIMO wireless cellular systems can achieve excellent data-rates if proper coordination among base stations (BSs) is

employed to suppress the intercell and intracell interference. The highest system capacity is achieved by networked MIMO where both the channel state information (CSI) and the data streams of all users in different cells are made available at all BSs [9, 10]. A more practical and less complex form of coordination is through joint linear precoding with only the CSI of the users shared among BSs [11–13]. In this case, each BS obtains the CSI of its intracell users and shares it with its adjacent BSs. Such precoding approach is very efficient for small number of BS antennas and it can alleviate the interference to a significant degree. As the number of BS antennas increases, the amount of signaling overhead for passing the CSI among BSs becomes taxing [14]. This problem is aggravated in frequency-selective channels employing OFDM where the CSI is different on each subcarrier. To overcome this difficulty, the precoders can be designed based on statistical CSI at the transmitter rather than instantaneous CSI [3, 15–20]. Since

the statistical CSI changes much slower than instantaneous CSI, the BSs need to be updated less frequently, hence, much less signaling overhead. While there is a performance loss due to the lack of exact CSI, precoder design based on statistical CSI for large-scale MIMO systems is an efficient approach because the performance loss is negligible [18, 20].

In order to design MIMO precoders with statistical CSI, the ergodic rate, which is a function of the covariance matrices of the MIMO channels, must be maximized over the precoders [16]. To facilitate the maximization, the ergodic rate can be approximated by a compact deterministic expression when the MIMO system is large-scale. The optimization can then be carried out on this deterministic approximation which is a function of the statistics of the channels [20–23]; thus, the solution is a set of precoders that are functions of the statistical CSI. Deterministic approximation to the Shannon rate can be achieved based on techniques such as Bai and Silverstein [22, 24, 25], Gaussian method [26, 27], Lindeberg principle [20, 26, 28, 29], or a combination of them [22, Theorem 6.9], [26]. All of these methods approximate the Shannon transform of large dimensional random matrices, but they are based on slightly different assumptions.

The deterministic equivalents of the rates for large-scale MIMO systems that are derived in recent works [20–23, 25, 26, 29] are functions of the channel statistics, so they can be used as objective functions for precoder design with statistical CSI. This has been done in [21–23] where closed form optimal precoders are derived for simple single-user scenarios. In [20], an uplink multiuser multicell MIMO system is considered and suboptimal linear precoders are found based on the Lindeberg approximation to the ergodic rate. We shall extend this methodology to the problem of downlink multiuser multicell MIMO-OFDM linear precoder design with statistical CSI. Considering the downlink leads to a somewhat different power constraint compared to the uplink, and the OFDM assumption gives rise to concepts such as frequency selectivity [17, 30, 31], channel tap correlations [17, 31–34], and intercarrier interference (ICI) [1, 35–38], all of which must be taken into account in the precoder design. Our derivation of statistical precoding for the downlink multicell MIMO-OFDM system is a unification of previous works [20, 29], and while several details are different from these two works, this derivation is not the main contribution. The main point of this work is applying the already established results of random matrix theory to the MIMO-OFDM scenario in order to study the impact of frequency selectivity, tap correlations, and ICI on the statistical precoder design and system performance.

We shall study the downlink multicell MIMO-OFDM linear precoder design with statistical CSI using the deterministic approximation of the Shannon rate, a problem studied before only under instantaneous CSI [12]. In particular, with instantaneous CSI, the channel statistics are not used in precoder design and one may try to optimize the instantaneous rate, as is done in [12] where it has been shown that such an optimization problem is nonconvex and a suboptimal solution has been proposed using the Karush-Kuhn-Tucker (KKT) conditions. In this work, however, we assume that there is only statistical CSI available at the BSs and we design

linear precoders that maximize the ergodic weighted sum rate. As opposed to the instantaneous case, for analysis of the ergodic rate, we need a statistical model for the MIMO channel which is usually determined by a covariance structure [23, 39]. We first consider the general correlated channel model with arbitrary probability distribution described in [23] that includes several statistical models such as the independent and identically distributed (i.i.d.) Rayleigh fading channel and the Kronecker channel [39]. Then we reduce the results to a simpler form for the Kronecker channel model.

To maximize the ergodic rate, inspired by [20, 23, 29], we find a deterministic approximation to the Shannon rate of the MIMO-OFDM system using the random matrix theory [21, 40]. The methodology will be applied to the MIMO-OFDM system with tap correlations and intercarrier interference. Then, assuming that the BSs have the statistical CSI of all users in the form of covariance matrices, we form a weighted sum rate maximization problem and propose a suboptimal solution using the KKT conditions along with the gradient descent method. Our simulations show that the approximation is quite accurate even for small size MIMO-OFDM systems and the proposed algorithm converges rapidly to a maximum which has a substantial improvement over isotropic precoding. The results only slightly underperform those obtained with perfect instantaneous CSI [12]. We then extend to the case where the frequency-selective channel suffers from correlation among channel taps and we show that the precoders become frequency dependent under tap correlations. We study the effects of tap correlation on the precoder design and system performance. Next, we allow ICI among OFDM subcarriers which is caused by carrier frequency offset due to synchronization errors and Doppler shifts [1]. The ICI introduces a new source of interference in addition to the intercell and intracell interference. We find the deterministic approximation to the rate under ICI and then study its impact on the precoder design and system performance. Our simulations show that while spatial correlations, tap correlations, and ICI decrease the system sum rate, our method alleviates this performance loss by incorporating the correlation information and ICI intensity information into the precoder design. It must be noted that our statistical CSI based method is applicable to networked MIMO with full cooperation where the BSs share the channel statistics and the transmit data.

The remainder of the paper is organized as follows. In Section 5, we provide the system model and formulate the optimization problem. In Section 3, we obtain the deterministic approximation to the ergodic rate function. In Section 4, we give a gradient-descent-based algorithm to obtain the suboptimal precoders. In Section 5, we discuss the channel model and how the formulas simplify for the separable channel models. In Section 6, we extend to the case where there is ICI. In Section 7, we present simulation results, and finally, conclusions are given in Section 8.

We denote matrices, vectors, and scalars by upper-case bold letters as in \mathbf{X} , lower-case bold letters as in \mathbf{x} , and non-bold letters as in X and x , respectively. Moreover, \mathbf{X}^* , \mathbf{X}^T , and \mathbf{X}^H denote conjugate, transpose, and conjugate-transpose, respectively. The element on the i th row and j th column of

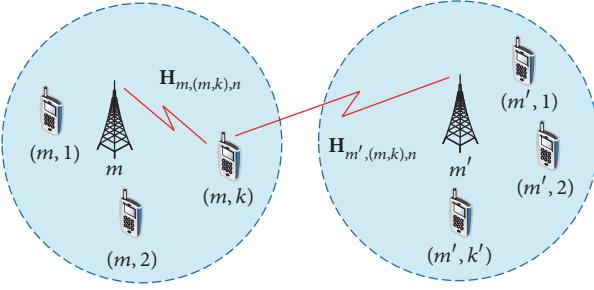


FIGURE 1: Multicell MIMO-OFDM system model.

a matrix is denoted by $[\mathbf{X}]_{i,j}$, and the i th element of a vector is denoted by $[\mathbf{x}]_i$. Vectorization, trace, and expected value operators are denoted by $\text{vec}\{\cdot\}$, $\text{tr}\{\cdot\}$, and $\mathbb{E}[\cdot]$, respectively. The all-one and all-zero vectors of size N and identity matrix of size $N \times N$ are denoted by $\mathbf{1}_N$, $\mathbf{0}_N$, and \mathbf{I}_N , respectively. The operators $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ represent real part and imaginary part, respectively.

2. System Model and Problem Formulation

2.1. System Model. We consider a downlink multicell MIMO-OFDM wireless network shown in Figure 1, with M BSs each serving K users over N subcarriers. All BSs have N_B transmitting antennas while each user has N_U receiving antennas with $N_B \geq N_U$. The k th user under BS m is denoted by (m, k) . The data vector to be precoded and transmitted by BS m to its k th user, that is, the data vector for user (m, k) , over subcarrier n is denoted by $\mathbf{x}_{(m,k),n} \in \mathbb{C}^{N_U}$, and the corresponding linear precoding matrix is denoted by $\mathbf{F}_{(m,k),n} \in \mathbb{C}^{N_B \times N_U}$. The downlink channel matrix from BS m' to user (m, k) over subcarrier n is denoted by $\mathbf{H}_{m',(m,k),n} \in \mathbb{C}^{N_U \times N_B}$. The thermal noise at user (m, k) on subcarrier n is denoted by $\mathbf{z}_{(m,k),n} \in \mathbb{C}^{N_U}$. Based on the above notations, the received signal at user (m, k) on subcarrier n is given by

$$\begin{aligned} \mathbf{y}_{(m,k),n} = & \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k),n} \mathbf{x}_{(m,k),n} \\ & + \sum_{k' \neq k}^K \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k'),n} \mathbf{x}_{(m,k'),n} \\ & + \sum_{m' \neq m}^M \sum_{k'=1}^K \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{x}_{(m',k'),n} \\ & + \mathbf{z}_{(m,k),n}, \end{aligned} \quad (1)$$

where the second and third terms represent the intracell and intercell interference, respectively. We assume that $\mathbf{x}_{(m,k),n}$ and $\mathbf{z}_{(m,k),n}$ are i.i.d. Gaussian with $\mathbb{E}[\mathbf{x}_{(m,k),n} \mathbf{x}_{(m,k),n}^H] = \mathbf{I}$ and $\mathbb{E}[\mathbf{z}_{(m,k),n} \mathbf{z}_{(m,k),n}^H] = \mathbf{I}$. We also assume that MIMO channels are independent across users and BSs.

Each user is assumed to know its own instantaneous CSI for detecting the data [16, 20, 22, 23, 29], which leads to the

instantaneous data-rate for user (m, k) on subcarrier n given by

$$r_{(m,k),n} = \log |\mathbf{I} + \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H \mathbf{H}_{m,(m,k),n}^H \Theta_{(m,k),n}^{-1}|, \quad (2)$$

where $\Theta_{(m,k),n}$ denotes the total interference plus noise on subcarrier n given by

$$\begin{aligned} \Theta_{(m,k),n} = & \mathbf{I} + \sum_{k' \neq k}^K \mathbf{H}_{m,(m,k),n} \mathbf{F}_{(m,k'),n} \mathbf{F}_{(m,k'),n}^H \mathbf{H}_{m,(m,k'),n}^H \\ & + \sum_{m' \neq m}^M \sum_{k'=1}^K \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H, \end{aligned} \quad (3)$$

where the first, second, and third terms represent the noise power, intracell interference power, and intercell interference power, respectively.

2.2. Problem Formulation. Our goal is to design precoder matrices based only on the second-order statistics of the MIMO channels. While the second-order statistics fully describe the statistical CSI for zero mean Gaussian channels, it is only partial statistical CSI for non-Gaussian channels. Defining $\mathcal{F} = \{\mathbf{F}_{(m,k),n}\}$ as the set of precoder matrices, we formulate the weighted sum rate optimization problem as

$$\begin{aligned} \max_{\mathcal{F}} \quad & \mathbb{E}_{\mathbf{H}} \left[\sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N r_{(m,k),n} \right] \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \text{tr}(\mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H) = P_m, \quad \forall m, \end{aligned} \quad (\mathbf{P}1)$$

where $\mathbb{E}_{\mathbf{H}}[\cdot]$ denotes expectation over the channel matrices $\{\mathbf{H}_{m',(m,k),n}\}$ and $w_{(m,k)}$ is the maximization weight associated with user (m, k) . The constraint is due to the fact that BS m is subject to the transmit power limit of P_m . Note that the power constraint introduces coupling across subcarriers, while the rate function given by (2) is independent for each subcarrier and coupled across users and BSs only, a feature of OFDM systems.

Define the covariance matrix of $\mathbf{H}_{m',(m,k),n}$ as

$$\Omega_{m',(m,k),n} = \mathbb{E}_{\mathbf{H}} [\text{vec} \{\mathbf{H}_{m',(m,k),n}\} \text{vec}^H \{\mathbf{H}_{m',(m,k),n}\}]. \quad (4)$$

Unlike designing precoders based instantaneous CSI as in [12], the precoders in \mathcal{F} are restricted to be functions of $\{\Omega_{m',(m,k),n}\}$. The above approach to precoder design has several advantages. It is very practical since acquiring full CSI for all $m', (m, k), n$ induces huge communication overhead due to CSI exchanges. Also, since the higher-order statistics can be costly to obtain, the second-order statistics often serve as the minimal statistics for various kinds of estimations.

In order to solve the nonconvex problem (P1), we need to find the expectation $\mathbb{E}_{\mathbf{H}}[r_{(m,k),n}]$. However, this expectation

makes the problem hard to analyse even for the single-user case [16]. Therefore, it is useful to replace the objective function by the analytically tractable approximation $\tilde{r}_{(m,k),n} \approx \mathbb{E}_H[r_{(m,k),n}]$. Thanks to the random matrix theory, as will be shown in the next section, we can find the deterministic approximation $\tilde{r}_{(m,k),n}$ that allows analytical expressions for the ergodic rate [20–23, 29]. The approximation is appropriate for our problem as $\tilde{r}_{(m,k),n}$ is a function of second-order statistics. While we have asymptotic convergence $\tilde{r}_{(m,k),n} \rightarrow \mathbb{E}_H[r_{(m,k),n}]$ when the number of BS antennas is large, our simulations show that the approximation is also accurate for MIMO-OFDM systems with very small number of antennas.

We now replace the ergodic rate $\mathbb{E}_H[r_{(m,k),n}]$ with its deterministic approximation $\tilde{r}_{(m,k),n}$ and form the analytically tractable problem

$$\begin{aligned} \max_{\mathcal{F}} \quad & \sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \tilde{r}_{(m,k),n} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \text{tr}(\mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H) = P_m, \quad \forall m. \end{aligned} \quad (\text{P2})$$

While analytically tractable, problem (P2) is still nonconvex. Since the global optimum to a nonconvex optimization problem is generally hard to obtain, we will develop an algorithm to find the local optimal solution to (P2). In the next sections, we will first derive the asymptotic approximation $\tilde{r}_{(m,k),n}$ and then propose an algorithm to obtain locally optimum precoder matrices.

3. Asymptotic Approximation of the Rate

In this section we derive the asymptotic approximation to the ergodic rate function. From (2) and (3), it follows that the rate function can be rewritten as

$$\begin{aligned} r_{(m,k),n} = & +\log \left| \mathbf{I} + \sum_{m'} \sum_{k'} \mathbf{H}_{m',(m,k),n} \right. \\ & \cdot \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. \right| - \log \left| \mathbf{I} \right. \\ & + \sum_{(m',k') \neq (m,k)} \sum_{m'} \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. \right|. \end{aligned} \quad (5)$$

To express the rate more compactly, define

$$\begin{aligned} \mathbf{A}_{(m,k),n} &= [\mathbf{H}_{1,(m,k),n} \bar{\mathbf{F}}_{1,n}, \dots, \mathbf{H}_{m,(m,k),n} \bar{\mathbf{F}}_{m,n}, \dots, \\ &\quad \mathbf{H}_{M,(m,k),n} \bar{\mathbf{F}}_{M,n}], \\ \mathbf{B}_{(m,k),n} &= [\mathbf{H}_{1,(m,k),n} \bar{\mathbf{F}}_{1,n}, \dots, \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n}, \dots, \\ &\quad \mathbf{H}_{M,(m,k),n} \bar{\mathbf{F}}_{M,n}], \end{aligned} \quad (6)$$

in which

$$\begin{aligned} \bar{\mathbf{F}}_{m',n} &= [\mathbf{F}_{(m',1),n}, \dots, \mathbf{F}_{(m',k),n}, \dots, \mathbf{F}_{(m',K),n}], \\ \check{\mathbf{F}}_{m,k,n} &= [\mathbf{F}_{(m,1),n}, \dots, \mathbf{F}_{(m,k-1),n}, \mathbf{F}_{(m,k+1),n}, \dots, \mathbf{F}_{(m,K),n}], \end{aligned} \quad (7)$$

so that the rate $r_{(m,k),n}$ is formed by the difference between two Shannon transforms as

$$\begin{aligned} r_{(m,k),n} = & \log \left| \mathbf{I} + \mathbf{A}_{(m,k),n} \mathbf{A}_{(m,k),n}^H \right| \\ & - \log \left| \mathbf{I} + \mathbf{B}_{(m,k),n} \mathbf{B}_{(m,k),n}^H \right|. \end{aligned} \quad (8)$$

The $N_U \times N_U KM$ matrix $\mathbf{A}_{(m,k),n}$ is associated with the total received signal while the $N_U \times N_U (KM - 1)$ matrix $\mathbf{B}_{(m,k),n}$ is associated with interference plus noise.

In order to approximate $\mathbb{E}_H[r_{(m,k),n}]$, we need to find the asymptotic approximation to $\mathbb{E}_X[\log |\mathbf{I} + \mathbf{XX}^H|]$. According to the random matrix theory [21, 22, 40], the ergodic Shannon transform of \mathbf{XX}^H can be asymptotically approximated by a deterministic function that only depends on the second-order statistics of \mathbf{X} . Therefore, we only need the second-order statistics of $\mathbf{A}_{(m,k),n}$ and $\mathbf{B}_{(m,k),n}$ to characterize $\tilde{r}_{(m,k),n} \approx \mathbb{E}_H[r_{(m,k),n}]$.

3.1. Second-Order Statistics of $\mathbf{A}_{(m,k),n}$ and $\mathbf{B}_{(m,k),n}$. We shall now calculate the $N_U^2 KM \times N_U^2 KM$ covariance matrix of $\mathbf{A}_{(m,k),n}$ defined as

$$\Sigma_{(m,k),n}^A = \mathbb{E}_H [\text{vec} \{\mathbf{A}_{(m,k),n}\} \text{vec}^H \{\mathbf{A}_{(m,k),n}\}], \quad (9)$$

which is associated with the total received signal, and the $N_U^2 (KM - 1) \times N_U^2 (KM - 1)$ covariance matrix of $\mathbf{B}_{(m,k),n}$ defined as

$$\Sigma_{(m,k),n}^B = \mathbb{E}_H [\text{vec} \{\mathbf{B}_{(m,k),n}\} \text{vec}^H \{\mathbf{B}_{(m,k),n}\}], \quad (10)$$

which is associated with interference plus noise. From the definition of $\mathbf{A}_{(m,k),n}$ and statistical independence of channel matrices across BSs, it follows that

$$\Sigma_{(m,k),n}^A = \begin{bmatrix} \Sigma_{1,(m,k),n} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \Sigma_{m,(m,k),n} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \Sigma_{M,(m,k),n} \end{bmatrix}, \quad (11)$$

where $\Sigma_{m',(m,k),n}$ is the $N_U^2 K \times N_U^2 K$ subcovariance matrix corresponding to BS m' and given by

$$\begin{aligned} \Sigma_{m',(m,k),n} &= \mathbb{E}_H [\text{vec} \{\mathbf{H}_{m',(m,k),n} \bar{\mathbf{F}}_{m',n}\} \text{vec}^H \{\mathbf{H}_{m',(m,k),n} \bar{\mathbf{F}}_{m',n}\}], \\ &= \mathbb{E}_H [\text{vec} \{\mathbf{H}_{m',(m,k),n} \bar{\mathbf{F}}_{m',n}\} \text{vec}^H \{\mathbf{H}_{m',(m,k),n} \bar{\mathbf{F}}_{m',n}\}]. \end{aligned} \quad (12)$$

Similarly, from the definition of $\mathbf{B}_{(m,k),n}$ and statistical independence of channel matrices across BSs, it follows that

$$\Sigma_{(m,k),n}^{\mathbf{B}} = \begin{bmatrix} \Sigma_{1,(m,k),n} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \cdots & \cdots & \mathbf{0} \\ \vdots & \vdots & \check{\Sigma}_{m,(m,k),n} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \cdots & \Sigma_{M,(m,k),n} \end{bmatrix}, \quad (13)$$

where $\check{\Sigma}_{m,(m,k),n}$ is the $N_U^2(K-1) \times N_U^2(K-1)$ subcovariance matrix corresponding to BS m' and given by

$$\check{\Sigma}_{m,(m,k),n} = \mathbb{E}_{\mathbf{H}} [\text{vec} \{ \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n} \} \text{vec}^H \{ \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n} \}]. \quad (14)$$

Now, we only need to calculate the subcovariance matrices $\Sigma_{m',(m,k),n}$ and $\check{\Sigma}_{m,(m,k),n}$ to completely determine the covariance matrices $\Sigma_{(m,k),n}^{\mathbf{A}}$ and $\Sigma_{(m,k),n}^{\mathbf{B}}$. It is straightforward to show that

$$\begin{aligned} \Sigma_{m',(m,k),n} &= (\check{\mathbf{F}}_{m',n}^T \otimes \mathbf{I}) \Omega_{m',(m,k),n} (\check{\mathbf{F}}_{m',n}^* \otimes \mathbf{I}), \\ \check{\Sigma}_{m,(m,k),n} &= (\check{\mathbf{F}}_{m,k,n}^T \otimes \mathbf{I}) \Omega_{m,(m,k),n} (\check{\mathbf{F}}_{m,k,n}^* \otimes \mathbf{I}), \end{aligned} \quad (15)$$

where $\Omega_{m',(m,k),n}$ is defined in (4).

We remark here that, in contrast to the uplink transmission considered in [20] where each subcovariance matrix is $N_U N_B \times N_U N_B$, in the downlink transmission each subcovariance matrix is $N_U^2 K \times N_U^2 K$, so each dimension is approximately K times larger. The increased size of the covariance matrix is due to the fact that the downlink transmitter (BS) has K intended receivers (users), while the uplink transmitter (user) has only one intended receiver (BS). In addition to this, there are N covariance matrix groups in the MIMO-OFDM system, one for each subcarrier.

The deterministic approximation will depend on the eigenvalues of $\Sigma_{(m,k),n}^{\mathbf{A}}$ and $\Sigma_{(m,k),n}^{\mathbf{B}}$, so we shall define two eigenvalue matrices. Define the $N_U \times N_U KM$ eigenvalue matrix $\Lambda_{(m,k),n}^{\mathbf{A}}$ and $N_U \times N_U (KM - 1)$ eigenvalue matrix $\Lambda_{(m,k),n}^{\mathbf{B}}$, which are formed by stacking the eigenvalues of $\Sigma_{(m,k),n}^{\mathbf{A}}$ and $\Sigma_{(m,k),n}^{\mathbf{B}}$ in $N_U \times 1$ sized columns, respectively. Such a definition is required to apply the theorems from random matrix theory. Note that the first dimension of eigenvalue matrices is the receiver antenna number N_U . It is easy to see that

$$\begin{aligned} \Lambda_{(m,k),n}^{\mathbf{A}} &= [\Lambda_{1,(m,k),n}, \dots, \Lambda_{m,(m,k),n}, \dots, \Lambda_{M,(m,k),n}], \\ \Lambda_{(m,k),n}^{\mathbf{B}} &= [\Lambda_{1,(m,k),n}, \dots, \check{\Lambda}_{m,(m,k),n}, \dots, \Lambda_{M,(m,k),n}], \end{aligned} \quad (16)$$

where $N_U \times N_U K$ eigenvalue matrix $\Lambda_{m',(m,k),n}$ and $N_U \times N_U (K-1)$ eigenvalue matrix $\check{\Lambda}_{m,(m,k),n}$ are formed by stacking the eigenvalues of $\Sigma_{m',(m,k),n}$ and $\check{\Sigma}_{m,(m,k),n}$ in $N_U \times 1$ sized columns, respectively.

3.2. Approximation of the Rate Function. In this subsection, we will derive the approximation to the rate based on the random matrix theory. We start by reviewing some concepts from the random matrix theory. For an $I \times J$ random matrix \mathbf{X} , the correlation function is defined as $R_{\mathbf{X}}(i, j; i', j') \triangleq \mathbb{E}[[\mathbf{X}]_{i,j} [\mathbf{X}]_{i',j'}]$. Define the kernel of the correlation function by a set of orthonormal eigenfunctions $\psi_{k,l}(i, j) \in \mathbb{C}$ satisfying $\sum_{i',j'} R_{\mathbf{X}}(i, j; i', j') \psi_{k,l}(i', j') = \lambda_{k,l} \psi_{k,l}(i, j)$ where $\lambda_{k,l}$ is the kernel eigenvalue [21]. We say the kernel is factorable if $\psi_{k,l}(i, j) = u_k(i) v_l(j)$ for some functions $u_k(i)$ and $v_l(j)$. Also, define $\Lambda^{\mathbf{X}} \in \mathbb{R}^{I \times J}$ as the eigenvalue matrix which is formed by stacking the eigenvalues of $\Sigma^{\mathbf{X}} = \mathbb{E}[\text{vec}\{\mathbf{X}\} \text{vec}^H \{\mathbf{X}\}]$ in $I \times 1$ sized columns. Matrices \mathbf{X} , $\Sigma^{\mathbf{X}}$, and $\Lambda^{\mathbf{X}}$ can represent $\mathbf{A}_{(m,k),n}$, $\Sigma_{(m,k),n}^{\mathbf{A}}$, and $\Lambda_{(m,k),n}^{\mathbf{A}}$ (or those of \mathbf{B}), respectively.

Proposition 1. Suppose the following technical conditions hold for the correlation function of $\mathbf{X} \in \mathbb{C}^{I \times J}$: (a) its kernel is factorable; (b) its eigenvalues multiplied by I are uniformly bounded. Then, based on [29, Theorem 2] and [23, Theorem 3], the random variable $(1/I) \log |\mathbf{I} + \mathbf{X} \mathbf{X}^H|$ will converge almost surely to its deterministic approximation $\tilde{f}_{\mathbf{X}}$ for J/I fixed as $I, J \rightarrow \infty$. We have

$$\begin{aligned} \tilde{f}_{\mathbf{X}} &\triangleq \frac{1}{I} \sum_{j=1}^J \log \left(1 + I \left[\Lambda^{\mathbf{X}} \mathbf{u} \right]_j \right) \\ &\quad + \frac{1}{I} \sum_{i=1}^I \log \left(1 + J \left[\Lambda^{\mathbf{X}} \mathbf{v} \right]_i \right) - J \mathbf{u}^T \Lambda^{\mathbf{X}} \mathbf{v} \log e, \end{aligned} \quad (17)$$

where $\Lambda^{\mathbf{X}}$ denotes the eigenvalue matrix and $\mathbf{u} \in \mathbb{R}^I$, $\mathbf{v} \in \mathbb{R}^J$ are solutions to the following fixed point equations:

$$\begin{aligned} [\mathbf{u}]_i^{-1} &= I \left(1 + J \left[\Lambda^{\mathbf{X}} \mathbf{v} \right]_i \right), \quad i = 1, \dots, I, \\ [\mathbf{v}]_j^{-1} &= J \left(1 + I \left[\Lambda^{\mathbf{X}} \mathbf{u} \right]_j \right), \quad j = 1, \dots, J. \end{aligned} \quad (18)$$

Now based on this approximation technique and the second-order statistics of $\mathbf{A}_{m',(m,k),n}$ and $\mathbf{B}_{m',(m,k),n}$ derived previously, we can obtain the asymptotic approximation to the ergodic rate in the following theorem.

Theorem 2. Suppose the following conditions are satisfied for the correlation functions of $\mathbf{A}_{(m,k),n}$ and $\mathbf{B}_{(m,k),n}$: (a) they have factorable kernels; (b) their eigenvalues multiplied by N_U are uniformly bounded. Then for $N_B \geq N_U \rightarrow \infty$ we have

$$\frac{r_{(m,k),n} - \tilde{r}_{(m,k),n}}{N_U} \rightarrow 0, \quad (19)$$

in which $\tilde{r}_{(m,k),n}$ is a deterministic function given by

$$\begin{aligned}
\tilde{r}_{(m,k),n} = & - \sum_{j=1}^{N_U KM} \log [\mathbf{v}_{(m,k),n}^A]_j \\
& + \sum_{j=1}^{N_U(KM-1)} \log [\mathbf{v}_{(m,k),n}^B]_j + \sum_{i=1}^{N_U} \left[\log \frac{[\mathbf{u}_{(m,k),n}^B]_i}{[\mathbf{u}_{(m,k),n}^A]_i} \right. \\
& \left. + N_U \log e ([\mathbf{u}_{(m,k),n}^A]_i - [\mathbf{u}_{(m,k),n}^B]_i) \right] - N_U \\
& \cdot \log (N_U (KM - 1)) + N_U KM \log \left(1 - \frac{1}{KM} \right),
\end{aligned} \tag{20}$$

where $\mathbf{u}_{(m,k),n}^A \in \mathbb{R}^{N_U}$, $\mathbf{v}_{(m,k),n}^A \in \mathbb{R}^{KMN_U}$, $\mathbf{u}_{(m,k),n}^B \in \mathbb{R}^{N_U}$, and $\mathbf{v}_{(m,k),n}^B \in \mathbb{R}^{(KM-1)N_U}$ are the solutions to the following fixed point equations:

$$[\mathbf{u}_{(m,k),n}^A]_i^{-1} = N_U (1 + KMN_U [\Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A]_i), \tag{21}$$

$$[\mathbf{v}_{(m,k),n}^A]_j^{-1} = N_U KM \left(1 + N_U [\Lambda_{(m,k),n}^A]^T \mathbf{u}_{(m,k),n}^A \right)_j, \tag{22}$$

$$\begin{aligned}
& [\mathbf{u}_{(m,k),n}^B]_i^{-1} \\
& = N_U (1 + N_U (KM - 1) [\Lambda_{(m,k),n}^B \mathbf{v}_{(m,k),n}^B]_i), \\
& [\mathbf{v}_{(m,k),n}^B]_j^{-1} \\
& = N_U (KM - 1) \left(1 + N_U [\Lambda_{(m,k),n}^B]^T \mathbf{u}_{(m,k),n}^B \right)_j,
\end{aligned} \tag{23}$$

where i and j range according to the vector sizes.

Proof. We will apply Proposition 1 to each term in (8). Consider the first term $\log |\mathbf{I} + \mathbf{A}_{(m,k),n} \mathbf{A}_{(m,k),n}^H|$ in which $\mathbf{A}_{(m,k),n} \in \mathbb{C}^{N_U \times N_U KM}$ must satisfy conditions (a) and (b) of Theorem 2 in accordance with the same conditions in Proposition 1. Since $N_U KM / N_U = KM$, the ratio of dimensions is always fixed as $N_B \geq N_U \rightarrow \infty$. Condition $N_B \geq N_U$ guarantees that the precoders have full column rank so (2) remains valid. Now, from Proposition 1, the first term in (8) divided by N_U is approximated as

$$\begin{aligned}
& \tilde{f}_{\mathbf{A}_{(m,k),n}} \\
& \triangleq \frac{1}{N_U} \sum_{j=1}^{N_U KM} \log \left(1 + N_U [\Lambda_{(m,k),n}^A]^T \mathbf{u}_{(m,k),n}^A \right)_j \\
& + \frac{1}{N_U} \sum_{i=1}^{N_U} \log \left(1 + KMN_U [\Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A]_i \right) \\
& - N_U KM \mathbf{u}_{(m,k),n}^A]^T \Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A \log e.
\end{aligned} \tag{24}$$

In the first term of approximation $\tilde{f}_{\mathbf{A}_{(m,k),n}}$, we can express $[\Lambda_{(m,k),n}^A]^T \mathbf{u}_{(m,k),n}^A$ in terms of $[\mathbf{v}_{(m,k),n}^A]_j$ using (22). In the second term of approximation $\tilde{f}_{\mathbf{A}_{(m,k),n}}$, we can express $[\Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A]_i$ in terms of $[\mathbf{u}_{(m,k),n}^A]_i$ using (21). For the third term of approximation $\tilde{f}_{\mathbf{A}_{(m,k),n}}$, we can write

$$\begin{aligned}
& \mathbf{u}_{(m,k),n}^A]^T \Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A \\
& = \sum_i \sum_j [\mathbf{u}_{(m,k),n}^A]_i [\Lambda_{(m,k),n}^A]_{i,j} [\mathbf{v}_{(m,k),n}^A]_j \\
& = \sum_i [\mathbf{u}_{(m,k),n}^A]_i \sum_j [\Lambda_{(m,k),n}^A]_{i,j} [\mathbf{v}_{(m,k),n}^A]_j \\
& = \sum_i [\mathbf{u}_{(m,k),n}^A]_i [\Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A]_i
\end{aligned} \tag{25}$$

and express $[\Lambda_{(m,k),n}^A \mathbf{v}_{(m,k),n}^A]_i$ in terms of $[\mathbf{u}_{(m,k),n}^A]_i$ using (21). So all the terms in $\tilde{f}_{\mathbf{A}_{(m,k),n}}$ are expressed in terms of $[\mathbf{u}_{(m,k),n}^A]_i$, $[\mathbf{v}_{(m,k),n}^A]_j$. In a similar manner, $\tilde{f}_{\mathbf{B}_{(m,k),n}}$ can be expressed in terms of $[\mathbf{u}_{(m,k),n}^B]_i$, $[\mathbf{v}_{(m,k),n}^B]_j$. Forming $(1/N_U) \tilde{r}_{(m,k),n} = \tilde{f}_{\mathbf{A}_{(m,k),n}} - \tilde{f}_{\mathbf{B}_{(m,k),n}}$ along with some mathematical manipulations completes the proof. \square

Some remarks are in order. It is shown in [23] that many channel models including the unitary-independent-unitary (UIU) model, Kronecker model, and independent nonidentically distributed (IND) model satisfy conditions (a) and (b) of Theorem 2. While the approximation is guaranteed to converge as $N_B \geq N_U \rightarrow \infty$, it is accurate even for very small number of antennas (e.g., $N_U = 3$, $N_B = 3$) as will be shown in the simulation section. The existence of positive solutions to the fixed point equations in Theorem 2 is proved in [21]. Computationally, the equations can be solved numerically by iteratively substituting the value on the right-hand side into the left-hand side. The convergence result suggests that the ergodic rate $\mathbb{E}_H[r_{(m,k),n}]$ can be approximated by $\tilde{r}_{(m,k),n}$.

4. Weighted Sum Rate Maximization

In this section, we develop an algorithm to obtain a sub-optimal solution to problem (P2). We will often need to differentiate one matrix with respect to another matrix; therefore, to facilitate calculations, the following definition is presented [41].

Definition 3. Let $\mathbf{f}(\mathbf{x})$ be a function mapping $\mathbb{R}^n \rightarrow \mathbb{R}^m$, define its derivative $\mathcal{D}_{\mathbf{x}} \mathbf{f}$ as the $m \times n$ matrix with elements

$$[\mathcal{D}_{\mathbf{x}} \mathbf{f}]_{i,j} = \frac{\partial \mathbf{f}_i(\mathbf{x})}{\partial \mathbf{x}_j}. \tag{26}$$

Furthermore, let $\mathbf{F}(\mathbf{X})$ be a function mapping $\mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{p \times q}$; define its derivative $\mathcal{D}_{\mathbf{X}}\mathbf{F}(\mathbf{X})$ as the $pq \times mn$ matrix

$$\mathcal{D}_{\mathbf{X}}\mathbf{F}(\mathbf{X}) = \mathcal{D}_{\text{vec}[\mathbf{X}]}\text{vec}\{\mathbf{F}(\mathbf{X})\}. \quad (27)$$

The most important property of this definition is that the chain rule holds, which enables the differentiation of complicated functions. Adopting this definition, we can now seek to solve the optimization problem.

Problem (P2) is nonconvex, so a global maximum is generally hard to obtain. Fortunately, the KKT conditions still serve as the necessary conditions for local optima, so we seek a suboptimal solution that satisfies the KKT conditions. To this end, we define

$$\mathbf{f}_{(m,k),n} = \begin{bmatrix} \text{vec}\{\text{Re}(\mathbf{F}_{(m,k),n})\} \\ \text{vec}\{\text{Im}(\mathbf{F}_{(m,k),n})\} \end{bmatrix} \in \mathbb{R}^{2N_U N_B \times 1}, \quad (28)$$

$$[\mathbf{f}_{(m,k),n}]_i = \begin{cases} \sqrt{P_m} \cos [\phi_{(m,k),n}]_i \prod_{(k',n',i') < (k,n,i)} \sin [\phi_{(m,k'),n'}]_{i'}, & (k,n,i) \neq (K,N,2N_B N_U) \\ \sqrt{P_m} \prod_{(k',n',i') < (K,N,2N_B N_U)} \sin [\phi_{(m,k'),n'}]_{i'}, & (k,n,i) = (K,N,2N_B N_U), \end{cases} \quad (30)$$

where $<$ denotes the lexicographical order of vectors. Under this parametrization, the power constraint is automatically satisfied since the variables are on the surface of a hypersphere with radius $\sqrt{P_m}$.

Since the local optima occur at the interior of the domain of angle parameters, the KKT condition for problem (P2) is simplified to

$$\sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n} = 0, \quad \forall m', \quad (31)$$

where $\theta_{m'} \in \mathbb{R}^{(2N_U N_B K N - 1) \times 1}$ holds all the angle parameters corresponding to BS m' . The above expression simply states that the gradient vanishes at any optimum point. So finding a local optimum solution is equivalent to pursuing a set of points $\theta_{m'}^{\text{opt}}$, $\forall m'$ at which the gradient vanishes. In practice, the equation for setting the gradient to zero is highly nonlinear and it is impossible to solve it directly. However, starting from any initial point for $\theta_{m'}$, we can use a simple gradient search method to increase the objective function after each iteration and gradually approach a point where the gradient is zero. In order to do so, we need to derive the $1 \times (2N_U N_B K N - 1)$ vector $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$, which is the transpose of the gradient vector. The remaining materials of this section will thus be devoted to the derivation of $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$.

4.1. Derivation of $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$. It immediately follows by the chain rule that

which is the complex-real isomorphism of $\mathbf{F}_{(m,k),n}$. Problem (P2) then becomes

$$\begin{aligned} \max_{\widetilde{\mathcal{F}}} \quad & \sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \tilde{r}_{(m,k),n} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \mathbf{f}_{(m,k),n}^T \mathbf{f}_{(m,k),n} = P_m, \quad \forall m, \end{aligned} \quad (29)$$

where $\widetilde{\mathcal{F}}$ is a set that has complex-real isomorphism with respect to \mathcal{F} .

To further simplify the problem, we can reformulate it as an unconstrained optimization. To do so, similar to [42], we make a change of variables from $\mathbf{f}_{(m,k),n}$ to the spherical coordinates $\phi_{(m,k),n}$ that belongs to $\mathbb{R}^{2N_B N_U}$ if $(k,n) \neq (K,N)$ and to $\mathbb{R}^{2N_B N_U - 1}$ if $(k,n) = (K,N)$. The transformation is described by

$$\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n} = \mathcal{D}_{\bar{\mathbf{f}}_{m',n}} \tilde{r}_{(m,k),n} \mathcal{D}_{\theta_{m'}} \bar{\mathbf{f}}_{m',n}, \quad (32)$$

$$\bar{\mathbf{f}}_{m',n} = [\mathbf{f}_{(m',1),n}^T, \mathbf{f}_{(m',2),n}^T, \dots, \mathbf{f}_{(m',K),n}^T]^T. \quad (33)$$

Now we need to find $\mathcal{D}_{\bar{\mathbf{f}}_{m',n}} \tilde{r}_{(m,k),n}$ and $\mathcal{D}_{\theta_{m'}} \bar{\mathbf{f}}_{m',n}$ in order to form the chain rule and obtain (33). We shall find derivatives with respect to the subvectors of $\bar{\mathbf{f}}_{m',n}$, that is, with respect to vectors $\mathbf{f}_{(m',k'),n}$. The components of $\mathcal{D}_{\theta_{m'}} \mathbf{f}_{(m',k'),n}$ can be obtained by differentiating (30) which yields the following:

$$\begin{aligned} \mathcal{D}_{\theta_{m'}} [\mathbf{f}_{(m',k'),n}]_i &= [\mathbf{f}_{(m',k'),n}]_i \\ &\times \begin{cases} \left[\begin{array}{c} \cot [\phi_{(m',1),1}]_1 \\ \vdots \\ \cot [\phi_{(m',k'),n}]_{i-1} \\ -\tan [\phi_{(m',k'),n}]_i \\ \mathbf{0} \\ \cot [\phi_{(m',1),1}]_1 \\ \vdots \\ \cot [\phi_{(m',k'),n}]_{i-2} \\ \cot [\phi_{(m',k'),n}]_{i-1} \end{array} \right]^T, & (k',n,i) \neq (K,N,2N_B^2) \\ \left[\begin{array}{c} \cot [\phi_{(m',1),1}]_1 \\ \vdots \\ \cot [\phi_{(m',k'),n}]_{i-2} \\ \cot [\phi_{(m',k'),n}]_{i-1} \end{array} \right]^T, & \text{o.w.} \end{cases} \end{aligned} \quad (34)$$

The components of $\mathcal{D}_{\mathbf{f}_{(m',k'),n}} \tilde{r}_{(m,k),n}$ can be found through the complex-real isomorphism as

$$\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{r}_{(m,k),n} = 2 \begin{bmatrix} \operatorname{Re}^T \left\{ \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{r}_{(m,k),n} \right\} \\ -\operatorname{Im}^T \left\{ \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{r}_{(m,k),n} \right\} \end{bmatrix}^T. \quad (35)$$

Substituting (34) and (35) in (33) completes the derivation provided that $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{r}_{(m,k),n}$ is available.

Calculation of $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{r}_{(m,k),n}$ is more complicated since, according to Theorem 2, $\tilde{r}_{(m,k),n}$ is a function of $\mathbf{u}_{(m,k),n}^A$, $\mathbf{v}_{(m,k),n}^A$, $\mathbf{u}_{(m,k),n}^B$, and $\mathbf{v}_{(m,k),n}^B$, which are in turn functions of $\Lambda_{(m,k),n}^A$ and $\Lambda_{(m,k),n}^B$. The chain rule gives

$$\begin{aligned} & \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \tilde{r}_{(m,k),n} \\ &= \mathcal{D}_{\mathbf{u}_{(m,k),n}^A} \tilde{r}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^A} \mathbf{u}_{(m,k),n}^A \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^A \\ &+ \mathcal{D}_{\mathbf{v}_{(m,k),n}^A} \tilde{r}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^A} \mathbf{v}_{(m,k),n}^A \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^A \quad (36) \\ &+ \mathcal{D}_{\mathbf{u}_{(m,k),n}^B} \tilde{r}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^B} \mathbf{u}_{(m,k),n}^B \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^B \\ &+ \mathcal{D}_{\mathbf{v}_{(m,k),n}^B} \tilde{r}_{(m,k),n} \mathcal{D}_{\Lambda_{(m,k),n}^B} \mathbf{v}_{(m,k),n}^B \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^B. \end{aligned}$$

There are three derivatives involved in each term of (36). In what follows, we shall calculate each of them.

4.1.1. First Chain of (36). The first chains can be obtained by differentiating $\tilde{r}_{(m,k),n}$, as given in Theorem 2, which yields

$$\begin{aligned} \mathcal{D}_{[\mathbf{u}_{(m,k),n}^A]_i} \tilde{r}_{(m,k),n} &= - \left([\mathbf{u}_{(m,k),n}^A]_i^{-1} - N_U \right) \log e, \\ \mathcal{D}_{[\mathbf{u}_{(m,k),n}^B]_i} \tilde{r}_{(m,k),n} &= \left([\mathbf{u}_{(m,k),n}^B]_i^{-1} - N_U \right) \log e, \quad (37) \\ \mathcal{D}_{[\mathbf{v}_{(m,k),n}^A]_i} \tilde{r}_{(m,k),n} &= - [\mathbf{v}_{(m,k),n}^A]_i^{-1} \log e, \\ \mathcal{D}_{[\mathbf{v}_{(m,k),n}^B]_i} \tilde{r}_{(m,k),n} &= [\mathbf{v}_{(m,k),n}^B]_i^{-1} \log e. \end{aligned}$$

4.1.2. Second Chain of (36). The second chains can be obtained by differentiating the fixed point equations given in Theorem 2. It can be shown that

$$\begin{aligned} & \begin{bmatrix} \mathcal{D}_{[\Lambda_{(m,k),n}^A]_{h,l}} \mathbf{u}_{(m,k),n}^A \\ \mathcal{D}_{[\Lambda_{(m,k),n}^A]_{h,l}} \mathbf{v}_{(m,k),n}^A \end{bmatrix} = -KN_U M \\ & \times \begin{bmatrix} \mathbf{U}_{(m,k),n}^A & KN_U^2 M \Lambda_{(m,k),n}^A \\ KN_U^2 M \Lambda_{(m,k),n}^A & \mathbf{V}_{(m,k),n}^A \end{bmatrix}^{-1} \\ & \cdot \begin{bmatrix} [\mathbf{v}_{(m,k),n}^A]_l \mathbf{e}_h^{N_U} \\ [\mathbf{u}_{(m,k),n}^A]_h \mathbf{e}_l^{KN_B M} \end{bmatrix}, \\ & \begin{bmatrix} \mathcal{D}_{[\Lambda_{(m,k),n}^B]_{h,l}} \mathbf{u}_{(m,k),n}^B \\ \mathcal{D}_{[\Lambda_{(m,k),n}^B]_{h,l}} \mathbf{v}_{(m,k),n}^B \end{bmatrix} = -(KM - 1) N_U \end{aligned}$$

$$\begin{aligned} & \times \begin{bmatrix} \mathbf{U}_{(m,k),n}^B & (KM - 1) N_U^2 \Lambda_{(m,k),n}^B \\ (KM - 1) N_U^2 \Lambda_{(m,k),n}^B & \mathbf{V}_{(m,k),n}^B \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} [\mathbf{v}_{(m,k),n}^B]_l \mathbf{e}_h^{N_U} \\ [\mathbf{u}_{(m,k),n}^B]_h \mathbf{e}_l^{(KM-1)N_B} \end{bmatrix}, \end{aligned} \quad (38)$$

where \mathbf{e}_l^P denotes the l th standard basis vector in \mathbb{R}^P and

$$\begin{aligned} \mathbf{U}_{(m,k),n}^A &= \operatorname{diag} \left\{ \left[\mathbf{u}_{(m,k),n}^A \right]_1^{-2}, \dots, \left[\mathbf{u}_{(m,k),n}^A \right]_{N_U}^{-2} \right\}, \\ \mathbf{V}_{(m,k),n}^A &= \operatorname{diag} \left\{ \left[\mathbf{v}_{(m,k),n}^A \right]_1^{-2}, \dots, \left[\mathbf{v}_{(m,k),n}^A \right]_{KN_U M}^{-2} \right\}, \\ \mathbf{U}_{(m,k),n}^B &= \operatorname{diag} \left\{ \left[\mathbf{u}_{(m,k),n}^B \right]_1^{-2}, \dots, \left[\mathbf{u}_{(m,k),n}^B \right]_{N_U}^{-2} \right\}, \\ \mathbf{V}_{(m,k),n}^B &= \operatorname{diag} \left\{ \left[\mathbf{v}_{(m,k),n}^B \right]_1^{-2}, \dots, \left[\mathbf{v}_{(m,k),n}^B \right]_{(KM-1)N_U}^{-2} \right\}. \end{aligned} \quad (39)$$

4.1.3. Third Chain of (36). Using the chain rule together with (16), we have

$$\begin{aligned} & \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^A \\ &= \begin{bmatrix} \mathbf{0} \\ \mathcal{D}_{\Sigma_{m',(m,k),n}} \Lambda_{m',(m,k),n} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Sigma_{m',(m,k),n} \\ \mathbf{0} \end{bmatrix}, \\ & \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^B \\ &= \begin{cases} \begin{bmatrix} \mathbf{0} \\ \mathcal{D}_{\Sigma_{m',(m,k),n}} \Lambda_{m',(m,k),n} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Sigma_{m',(m,k),n} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, & m' \neq m, \\ \begin{bmatrix} \mathcal{D}_{\check{\Sigma}_{m,(m,k),n}} \check{\Lambda}_{m,(m,k),n} \mathcal{D}_{\mathbf{F}_{(m',k'),n}} \check{\Sigma}_{m,(m,k),n} \\ \mathbf{0} \end{bmatrix}, & m' = m. \end{cases} \end{aligned} \quad (40)$$

We now need to calculate the terms in the above formulas. Note that the nonzero terms in the above matrices are equal to $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{m',(m,k),n}$ and $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \check{\Lambda}_{m,(m,k),n}$ expressed through the chain rule.

Based on the results in [41], the derivative of the eigenvalues with respect to the matrix can be explicitly written as a function of the eigenvectors, so we obtain

$$\begin{aligned} & \mathcal{D}_{\Sigma_{m',(m,k),n}} \Lambda_{m',(m,k),n} \\ &= \begin{bmatrix} \operatorname{vec}^T \left\{ \boldsymbol{\mu}_{m',(m,k),n}^1 \boldsymbol{\mu}_{m',(m,k),n}^1 \right\} \\ \vdots \\ \operatorname{vec}^T \left\{ \boldsymbol{\mu}_{m',(m,k),n}^{KN_U^2} \boldsymbol{\mu}_{m',(m,k),n}^{KN_U^2} \right\} \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} & \mathcal{D}_{\check{\Sigma}_{m,(m,k),n}} \check{\Lambda}_{m,(m,k),n} \\ &= \begin{bmatrix} \text{vec}^T \left\{ \check{\mu}_{m,(m,k),n}^1 * \check{\mu}_{m,(m,k),n}^{1^T} \right\} \\ \vdots \\ \text{vec}^T \left\{ \check{\mu}_{m,(m,k),n}^{(K-1)N_U^2} * \check{\mu}_{m,(m,k),n}^{(K-1)N_U^2 T} \right\} \end{bmatrix}, \end{aligned} \quad (41)$$

where $\check{\mu}_{m',(m,k),n}^i$ and $\check{\mu}_{m,(m,k),n}^i$ denote the i th eigenvector of $\Sigma_{m',(m,k),n}$ and $\check{\Sigma}_{m,(m,k),n}$, respectively.

Finally, differentiating (15) and using the results in [41], we obtain the formulas for $\mathcal{D}_{F_{(m',k'),n}} \Sigma_{m',(m,k),n}$ and $\mathcal{D}_{F_{(m,k'),n}} \check{\Sigma}_{m,(m,k),n}$ given by

$$\begin{aligned} \mathcal{D}_{F_{(m',k'),n}} \Sigma_{m',(m,k),n} &= \left\{ \left[\left(\bar{F}_{m',n}^H \otimes I_{N_U} \right) \left(\Omega_{m',(m,k),n}^{1/2} \right)^* \right] \right. \\ &\quad \times K_{N_U N_B, K N_U^2} \left\{ e_k^K \otimes \left[\left(\Omega_{m',(m,k),n}^{T/2} \otimes I_{N_U^2} \right) \right. \right. \\ &\quad \cdot \left(I_{N_B} \otimes K_{N_U, N_U} \otimes I_{N_U} \right) \left(I_{N_U N_B} \otimes \text{vec} \{ I_{N_U} \} \right) \\ &\quad \cdot \left. \left. K_{N_B, N_U} \right] \right\}, \end{aligned} \quad (42)$$

$$\begin{aligned} \mathcal{D}_{F_{(m,k'),n}} \check{\Sigma}_{m,(m,k),n} &= \left\{ \left[\left(\check{F}_{m,k',n}^H \otimes I_{N_U} \right) \left(\Omega_{m,(m,k),n}^{1/2} \right)^* \right] \right. \\ &\quad \times I_{(K-1)N_U^2} \left\{ K_{N_U N_B, (K-1)N_U^2} \left\{ u_{k,k'} \right. \right. \\ &\quad \otimes \left[\left(\Omega_{m,(m,k),n}^{T/2} \otimes I_{N_U^2} \right) \left(I_{N_B} \otimes K_{N_U, N_U} \otimes I_{N_U} \right) \right. \\ &\quad \cdot \left. \left. \left(I_{N_U N_B} \otimes \text{vec} \{ I_{N_U} \} \right) K_{N_B, N_U} \right] \right\}, \end{aligned} \quad (42)$$

in which $K_{k,l}$ is the commutation matrix defined in [41] satisfying $K_{k,l} \text{vec}\{X\} = \text{vec}\{X^T\}$ for every matrix $X \in \mathbb{C}^{k \times l}$, and

$$u_{k,k'} = \begin{cases} e_k^{K-1} & k' < k, \\ e_{k'-1}^{K-1} & k' > k, \\ 0_{K-1} & k' = k. \end{cases} \quad (43)$$

With all the chains derived, (36) is completely characterized which in turn enables an explicit expression for $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$ given by (33).

4.2. Local Optimum Solution. With the gradient $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$ computed, a gradient search method can be applied. We now propose Algorithm 1 to find a local optimum solution for problem (P2).

Algorithm 1 is not limited to multicell networks with partial cooperation, so with small modifications it can be used for precoder design based on statistical CSI for networked MIMO systems. In networked MIMO, in addition to the channel statistics, it is assumed that each BS has all the transmit data, so the whole system can be viewed as a MIMO super-cell with statistical CSI at the transmitter.

5. MIMO-OFDM Kronecker Channel Model

While the results obtained so far are valid for general correlation channel models, in this section, we will reduce the results to a more compact form with less computational complexity by considering the Kronecker channel model. The Kronecker model arises in practice when the immediate surrounding dominates the spatial correlation and the intermediate scattering clusters exist in a narrow angular range seen from the antennas [34, 39]. We shall start by describing the MIMO-OFDM channel and expressing it with the Kronecker structure.

5.1. Statistical Representation. A wideband MIMO channel is characterized by L channel taps $\tilde{H}_0, \dots, \tilde{H}_{L-1} \in \mathbb{C}^{I \times J}$. The channel matrix on the n th subcarrier is then given by $H_n = \sum_{l=0}^{L-1} \tilde{H}_l e^{-j2\pi nl/N} \in \mathbb{C}^{I \times J}$. Now the correlation among channel taps is given by the tap correlation matrix $\Upsilon \in \mathbb{C}^{L \times L}$ that is defined as $[\Upsilon]_{l,k} \triangleq \mathbb{E}\{[\tilde{H}_l]_{i,j} [\tilde{H}_k]_{i',j'}^*\}, \forall i, j, i', j'$. On the other hand, the Kronecker model assumes that the correlation of transmitter side and the receiver side is separable, so for each channel tap \tilde{H}_l , we have $\mathbb{E}\{[\tilde{H}_l]_{i,j} [\tilde{H}_l]_{i',j'}^*\} = [\mathbf{R}]_{i,i'} [\mathbf{T}]_{j,j'}$ where $\mathbf{R} \in \mathbb{C}^{I \times I}$ and $\mathbf{T} \in \mathbb{C}^{J \times J}$ are receiver and transmitter correlation matrices, respectively. Based on the above definitions, it is easy to show that the channel correlation matrix on subcarrier n is given by

$$\begin{aligned} \Omega_n &= \mathbb{E} \left\{ \text{vec} \{ H_n \} \text{vec}^H \{ H_n \} \right\} = \omega_n^\Upsilon (\mathbf{T} \otimes \mathbf{R}) \in \mathbb{C}^{IJ \times IJ}, \\ \omega_n^\Upsilon &\triangleq \mathbf{w}_n^H \Upsilon \mathbf{w}_n, \end{aligned} \quad (44)$$

$$\mathbf{w}_n = [e^{-j2\pi n(0)/N}, \dots, e^{-j2\pi n(L-1)/N}]^T.$$

Therefore, the channel correlation matrix Ω_n is characterized by the Kronecker product of transmitter and receiver correlation matrices \mathbf{T} and \mathbf{R} , multiplied by the quadratic form ω_n^Υ which depends on the tap correlation matrix and the n th Fourier vector. Finally, the following statistical representation can be considered for the channel:

$$H_n = \mathbf{R}^{1/2} \mathbf{H}_w \left[(\Upsilon^{1/2} \mathbf{w}_n) \otimes \mathbf{T}^{1/2} \right], \quad (45)$$

where $\mathbf{H}_w \in \mathbb{C}^{I \times JL}$ is a white random matrix whose elements are uncorrelated with zero mean and unit variance. Through the properties of the Kronecker product, the above statistical representation yields $\Omega_n = \omega_n^\Upsilon (\mathbf{T} \otimes \mathbf{R})$. We shall use (45) to describe the MIMO-OFDM Kronecker channel model.

If the channel taps are uncorrelated, Υ is diagonal and ω_n^Υ becomes independent of n ; hence, the channel correlation function Ω_n is the same over all subcarriers. However, whenever the channel taps are correlated, Υ is not diagonal and channel statistics are different on each subcarrier. Since the precoders depend on the channel statistics, when there is tap correlation, the MIMO-OFDM precoding matrices are frequency dependent which limits the system performance [17, 31–34]. But when there is no tap correlation, precoders are the same across all frequencies. We shall study the effect of tap correlation on the system sum rate in the simulation results section.

```

(1) Initialize: initialize  $\theta_{m'}(0)$ ,  $\forall m'$  and  $t \leftarrow 0$ ;
(2) while stopping criterion not met do
    (3) compute  $\mathbf{F}_{(m',k'),n}$ ,  $\forall m', k', n$  using (28), (30);
    (4) solve for  $\mathbf{u}_{(m,k),n}^A, \mathbf{v}_{(m,k),n}^A, \mathbf{u}_{(m,k),n}^B, \mathbf{v}_{(m,k),n}^B, \forall m, k, n$ , using Theorem 2;
    (5) compute  $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}, \forall m', m, k, n$  using (33);
    (6)  $\theta_{m'}(t+1) \leftarrow \theta_{m'}(t) + \alpha(t) \sum_{m,k,n} w_{(m,k)} \mathcal{D}_{\theta_{m'}}^T \tilde{r}_{(m,k),n}, \forall m' \text{ and } t \leftarrow t+1$ ;
(7) end while
(8) return  $\mathbf{F}_{(m',k'),n}$ ,  $\forall m'$ ;

```

ALGORITHM 1: Gradient search solution to problem (P2).

5.2. Eigendecomposition under Kronecker Model. Based on (45), the downlink channel matrix between BS m' and user (m, k) can be expressed as

$$\mathbf{H}_{m',(m,k),n} = \mathbf{R}_{m',(m,k),n}^{1/2} \mathbf{H}_w \left[(\mathbf{Y}_{m',(m,k)}^{1/2} \mathbf{w}_n) \otimes \mathbf{T}_{m',(m,k)}^{1/2} \right], \quad (46)$$

where $\mathbf{R}_{m',(m,k)} \in \mathbb{C}^{N_U \times N_U}$, $\mathbf{T}_{m',(m,k)} \in \mathbb{C}^{N_B \times N_B}$, and $\mathbf{Y}_{m',(m,k)} \in \mathbb{C}^{L \times L}$ are the corresponding receiver, transmitter, and tap correlation matrices, respectively, and $\mathbf{H}_w \in \mathbb{C}^{N_U \times N_B L}$ is a white random matrix whose elements are uncorrelated with zero mean and unit variance. Obviously, we have $\Omega_{m',(m,k),n} = \omega_n^{\mathbf{Y}_{m',(m,k)}} (\mathbf{T}_{m',(m,k)} \otimes \mathbf{R}_{m',(m,k)})$ that along with (15) yields

$$\begin{aligned} \Sigma_{m',(m,k),n} &= \omega_n^{\mathbf{Y}_{m',(m,k)}} \left(\bar{\mathbf{F}}_{m',n}^T \mathbf{T}_{m',(m,k)} \bar{\mathbf{F}}_{m',n}^* \right) \otimes \mathbf{R}_{m',(m,k)}, \\ \check{\Sigma}_{m,(m,k),n} &= \omega_n^{\mathbf{Y}_{m,(m,k)}} \left(\check{\mathbf{F}}_{m,k,n}^T \mathbf{T}_{m,(m,k)} \check{\mathbf{F}}_{m,k,n}^* \right) \otimes \mathbf{R}_{m,(m,k)}. \end{aligned} \quad (47)$$

Observe that the above subcovariance matrices decompose into a separated form, so their eigenvalue matrices admit the following decomposition:

$$\begin{aligned} \Lambda_{m',(m,k),n} &= \omega_n^{\mathbf{Y}_{m',(m,k)}} \lambda_{m',(m,k)}^R \lambda_{m',(m,k),n}^{\bar{T}}, \\ \check{\Lambda}_{m,(m,k),n} &= \omega_n^{\mathbf{Y}_{m,(m,k)}} \lambda_{m,(m,k)}^R \lambda_{m,(m,k),n}^{\check{T}}, \end{aligned} \quad (48)$$

in which the vector $\lambda_{m',(m,k)}^R \in \mathbb{C}^{N_U \times 1}$ holds the eigenvalues of $\mathbf{R}_{m',(m,k)}$, and the vectors $\lambda_{m',(m,k),n}^{\bar{T}} \in \mathbb{C}^{KN_U \times 1}$ and $\lambda_{m,(m,k),n}^{\check{T}} \in \mathbb{C}^{(K-1)N_U \times 1}$ hold the eigenvalues of $\bar{\mathbf{F}}_{m',n}^T \mathbf{T}_{m',(m,k)} \bar{\mathbf{F}}_{m',n}^*$ and $\check{\mathbf{F}}_{m,k,n}^T \mathbf{T}_{m,(m,k)} \check{\mathbf{F}}_{m,k,n}^*$, respectively. Now the eigenvalue matrices $\Lambda_{m',(m,k)}^A$ and $\Lambda_{m,(m,k)}^B$ can then be found by inserting (48) into (16).

Computationally, we no longer need to compute the eigenvalues of very large matrices $\Sigma_{m',(m,k),n} \in \mathbb{C}^{KN_U^2 \times KN_U^2}$ and $\check{\Sigma}_{m,(m,k),n} \in \mathbb{C}^{(K-1)N_U^2 \times (K-1)N_U^2}$, but instead, we need only compute the eigenvalues of the lower dimensional matrices $\mathbf{R}_{m',(m,k)} \in \mathbb{C}^{N_U \times N_U}$, $\bar{\mathbf{F}}_{m',n}^T \mathbf{T}_{m',(m,k)} \bar{\mathbf{F}}_{m',n}^* \in \mathbb{C}^{KN_U \times KN_U}$, and $\check{\mathbf{F}}_{m,k,n}^T \mathbf{T}_{m,(m,k)} \check{\mathbf{F}}_{m,k,n}^* \in \mathbb{C}^{(K-1)N_U \times (K-1)N_U}$. Using the above eigendecomposition, the computational complexity is reduced from $\mathcal{O}(K^3 N_U^6)$ to $\mathcal{O}(K^3 N_U^3)$.

5.3. Derivation of $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}$ under Kronecker Model. The procedure given in Section 4.1 can be applied here with great simplifications for $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^A$ and $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{(m,k),n}^B$ which constitute the third chain of (36) computed in Section 4.1.3. Note that the third chain was the most complicated component of the derivative. Considering (16), it suffices to find $\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{m',(m,k),n}$ and $\mathcal{D}_{\mathbf{F}_{(m,k'),n}} \check{\Lambda}_{m,(m,k),n}$. It can be shown from (48) that

$$\begin{aligned} &\mathcal{D}_{\mathbf{F}_{(m',k'),n}} \Lambda_{m',(m,k),n} \\ &= \omega_n^{\mathbf{Y}_{m',(m,k)}} (\mathbf{I}_{KN_U} \otimes \lambda_{m',(m,k)}^R) \\ &\quad \times \begin{bmatrix} \text{vec}^T \left\{ \bar{\mu}_{m',(m,k),n}^1 \bar{\mu}_{m',(m,k),n}^1 \right\} \\ \vdots \\ \text{vec}^T \left\{ \bar{\mu}_{m',(m,k),n}^{KN_U} \bar{\mu}_{m',(m,k),n}^{KN_U} \right\} \end{bmatrix} \\ &\quad \times \left(\bar{\mathbf{F}}_{m',n}^H \mathbf{T}_{m',(m,k)}^{1/2} \otimes \mathbf{I}_{KN_U} \right) \\ &\quad \times \mathbf{K}_{N_B, KN_U} \left(\mathbf{e}_{k'}^K \otimes \mathbf{I}_{N_U} \otimes \mathbf{T}_{m',(m,k)}^{1/2} \right), \end{aligned} \quad (49)$$

$$\begin{aligned} &\mathcal{D}_{\mathbf{F}_{(m,k'),n}} \check{\Lambda}_{m,(m,k),n} \\ &= \omega_n^{\mathbf{Y}_{m,(m,k)}} (\mathbf{I}_{KN_U} \otimes \lambda_{m,(m,k)}^R) \\ &\quad \times \begin{bmatrix} \text{vec}^T \left\{ \check{\mu}_{m,(m,k),n}^1 \check{\mu}_{m,(m,k),n}^1 \right\} \\ \vdots \\ \text{vec}^T \left\{ \check{\mu}_{m,(m,k),n}^{(K-1)N_U} \check{\mu}_{m,(m,k),n}^{(K-1)N_U} \right\} \end{bmatrix} \\ &\quad \times \left(\check{\mathbf{F}}_{m,n}^H \mathbf{T}_{m,(m,k)}^{1/2} \otimes \mathbf{I}_{KN_U} \right) \\ &\quad \times \mathbf{K}_{N_B, (K-1)N_U} \left(\mathbf{u}_{k,k'} \otimes \mathbf{I}_{N_U} \otimes \mathbf{T}_{m,(m,k)}^{1/2} \right), \end{aligned}$$

where $\mu_{m',(m,k),n}^i$ and $\check{\mu}_{m,(m,k),n}^i$ denote the i th eigenvector of $\bar{\mathbf{F}}_{m',n}^T \mathbf{T}_{m',(m,k)} \bar{\mathbf{F}}_{m',n}^*$ and $\check{\mathbf{F}}_{m,k,n}^T \mathbf{T}_{m,(m,k)} \check{\mathbf{F}}_{m,k,n}^*$, respectively, and $\mathbf{u}_{k,k'}$ was defined in (43). Using the above equations, the computational complexity is reduced from $\mathcal{O}(KN_U^3 N_B)$ to $\mathcal{O}(KN_U N_B)$.

6. Intercarrier Interference

In this section, we allow ICI among OFDM subcarriers. The case without ICI is then a special case of this scenario. The ICI occurs when there is carrier frequency offset due to synchronization errors and Doppler shifts [1, 38, 43]. This leads to loss of orthogonality among subcarriers which introduces more interference to the system. To accommodate our method to this situation, we will approximate the rate under ICI and then extend the proposed algorithm to facilitate the precoder design with statistical CSI.

6.1. System Model and Problem Formulation under ICI. When there is ICI, the received signal at user (m, k) on subcarrier n can be modelled as

$$\begin{aligned} \mathbf{y}_{(m,k),n} = & \sum_{m',k',n'} \sqrt{\rho_{n',n}} \mathbf{H}_{m',(m,k),n'} \mathbf{F}_{(m',k'),n'} \mathbf{x}_{(m',k'),n'} \\ & + \mathbf{z}_{(m,k),n}, \end{aligned} \quad (50)$$

where $|\rho_{n',n}| < 1$ models the power leaked from subcarrier n' to subcarrier n due to the ICI. Following [1], we model the ICI through the normalized frequency offset denoted by $0 < \epsilon < 1$, which is the ratio of the actual frequency offset to the intercarrier spacing. It is shown in [1] that $\rho_{n',n}$ relates to the ϵ by

$$\begin{aligned} \rho_{n',n} = & \left(\frac{\sin(\pi(n' - n + \epsilon))}{N \sin((\pi/N)(n' - n + \epsilon))} \right) \\ & \cdot \exp \left[j\pi \left(1 - \frac{1}{N} \right) \epsilon \right]^2. \end{aligned} \quad (51)$$

Here at the receiver, in addition to the summations over m', k' that model intercell and intracell interference, the received signal over subcarrier n depends also on all other subcarriers through ICI and thus the summation over n' . Note that the formulation reduces to that of the non-ICI scenario when $\epsilon = 0$, or equivalently $\rho_{n',n} = \delta_{n',n}$. The achievable instantaneous data-rate under ICI for user (m, k) on subcarrier n is then given by

$$\begin{aligned} r_{(m,k),n}^{\text{ICI}} = & + \log \left| \mathbf{I} + \sum_{m',k',n'} |\rho_{n',n}| \right. \\ & \cdot \mathbf{H}_{m',(m,k),n} \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. - \log \left| \mathbf{I} \right. \right. \\ & + \sum_{(m',k',n') \neq (m,k,n)} |\rho_{n',n}| \mathbf{H}_{m',(m,k),n} \\ & \cdot \mathbf{F}_{(m',k'),n} \mathbf{F}_{(m',k'),n}^H \mathbf{H}_{m',(m,k),n}^H \left. \left. \right| \right|. \end{aligned} \quad (52)$$

If we define

$$\begin{aligned} \tilde{\mathbf{A}}_{(m,k),n} \triangleq & \left[\sqrt{\rho_{1,n}} \mathbf{H}_{1,(m,k),1} \bar{\mathbf{F}}_{1,1}, \dots, \sqrt{\rho_{N,n}} \mathbf{H}_{1,(m,k),N} \bar{\mathbf{F}}_{1,N}, \right. \\ & \dots, \sqrt{\rho_{n,n}} \mathbf{H}_{m,(m,k),n} \bar{\mathbf{F}}_{m,n}, \dots, \sqrt{\rho_{1,n}} \mathbf{H}_{M,(m,k),1} \bar{\mathbf{F}}_{M,1}, \dots, \end{aligned}$$

$$\begin{aligned} & \left. \sqrt{\rho_{N,n}} \mathbf{H}_{M,(m,k),N} \bar{\mathbf{F}}_{M,N} \right], \\ \tilde{\mathbf{B}}_{(m,k),n} \triangleq & \left[\sqrt{\rho_{1,n}} \mathbf{H}_{1,(m,k),1} \bar{\mathbf{F}}_{1,1}, \dots, \sqrt{\rho_{N,n}} \mathbf{H}_{1,(m,k),N} \bar{\mathbf{F}}_{1,N}, \right. \\ & \dots, \sqrt{\rho_{n,n}} \mathbf{H}_{m,(m,k),n} \check{\mathbf{F}}_{m,k,n}, \dots, \sqrt{\rho_{1,n}} \mathbf{H}_{M,(m,k),1} \bar{\mathbf{F}}_{M,1}, \dots, \\ & \left. \sqrt{\rho_{N,n}} \mathbf{H}_{M,(m,k),N} \bar{\mathbf{F}}_{M,N} \right], \end{aligned} \quad (53)$$

then we can write

$$\begin{aligned} r_{(m,k),n}^{\text{ICI}} = & \log \left| \mathbf{I} + \tilde{\mathbf{A}}_{(m,k),n} \tilde{\mathbf{A}}_{(m,k),n}^H \right| \\ & - \log \left| \mathbf{I} + \tilde{\mathbf{B}}_{(m,k),n} \tilde{\mathbf{B}}_{(m,k),n}^H \right|. \end{aligned} \quad (54)$$

Based on this result, we can utilize the same random matrix method employed for the non-ICI scenario to derive the approximated rate $\tilde{r}_{(m,k),n}^{\text{ICI}} \approx \mathbb{E}_{\mathbf{H}}[r_{(m,k),n}^{\text{ICI}}]$.

The optimization problem we consider is similar to **(P1)** but using the rate under ICI, that is, $r_{(m,k),n}^{\text{ICI}}$. As before, we replace the rate $r_{(m,k),n}^{\text{ICI}}$ by its approximation $\tilde{r}_{(m,k),n}^{\text{ICI}}$ for tractability. Therefore, the following optimization problem is formed:

$$\begin{aligned} \max_{\mathcal{F}} \quad & \sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \tilde{r}_{(m,k),n}^{\text{ICI}} \\ \text{s.t.} \quad & \sum_{k=1}^K \sum_{n=1}^N \text{tr}(\mathbf{F}_{(m,k),n} \mathbf{F}_{(m,k),n}^H) = P_m, \quad \forall m. \end{aligned} \quad (\text{P3})$$

6.2. Second-Order Statistics under ICI. Due to the ICI, in addition to per-carrier statistics, the BSs now share the additional information of cross-carrier correlation. Specifically, the cross-carrier covariance matrices,

$$\begin{aligned} \Omega_{m',(m,k),n_1,n_2} \triangleq & \mathbb{E}_{\mathbf{H}} \left[\text{vec} \left\{ \mathbf{H}_{m',(m,k),n_1} \right\} \text{vec}^H \left\{ \mathbf{H}_{m',(m,k),n_2} \right\} \right], \end{aligned} \quad (55)$$

are assumed to be available at all BSs. This additional information is important for the BSs to suppress the ICI. For Kronecker model, the transmitter, receiver, and tap correlation matrices are sufficient because we have

$$\Omega_{m',(m,k),n_1,n_2} = \omega_{n_1,n_2}^{\Upsilon_{m',(m,k)}} (\mathbf{T}_{m',(m,k)} \otimes \mathbf{R}_{m',(m,k)}) \quad (56)$$

with $\omega_{n_1,n_2}^{\Upsilon_{m',(m,k)}} = \mathbf{w}_{n_1}^H \Upsilon_{m',(m,k)} \mathbf{w}_{n_2}$.

We will now find the second-order statistics of $\tilde{\mathbf{A}}_{(m,k),n}$ and $\tilde{\mathbf{B}}_{(m,k),n}$ which depend on $\Omega_{m',(m,k),n_1,n_2}$. It can be shown that

$$\begin{aligned} \Sigma_{(m,k),n}^{\tilde{\mathbf{A}}} &= \text{diag} \left\{ \Sigma_{1,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{m,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{M,(m,k),n}^{\text{ICI}} \right\} \\ \Sigma_{(m,k),n}^{\tilde{\mathbf{B}}} &= \text{diag} \left\{ \Sigma_{1,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{m,(m,k),n}^{\text{ICI}}, \dots, \Sigma_{M,(m,k),n}^{\text{ICI}} \right\} \end{aligned} \quad (57)$$

where $\Sigma_{m',(m,k),n}^{\text{ICI}}$ and $\check{\Sigma}_{m,(m,k),n}^{\text{ICI}}$ are block matrices holding N^2 blocks with block (n_1, n_2) given by

$$\begin{aligned} \Sigma_{m',(m,k),n}^{\text{ICI}(n_1, n_2)} &= \sqrt{\rho_{n_1,n}\rho_{n_2,n}} \\ &\times \left(\bar{\mathbf{F}}_{m',n_1}^T \otimes \mathbf{I}_{N_U} \right) \Omega_{m',(m,k),n_1, n_2} \left(\bar{\mathbf{F}}_{m',n_1}^T \otimes \mathbf{I}_{N_U} \right)^H, \end{aligned} \quad (58)$$

$$\check{\Sigma}_{m,(m,k),n}^{\text{ICI}(n_1, n_2)} = \sqrt{\rho_{n_1,n}\rho_{n_2,n}} \mathbf{U}_{m,k,n_1} \Omega_{m,(m,k),n_1, n_2} \mathbf{U}_{m,k,n_2}^H,$$

with $\mathbf{U}_{m,k,\tilde{n}} = \bar{\mathbf{F}}_{m,\tilde{n}}^T \otimes \mathbf{I}_{N_U}$, $\forall \tilde{n} \neq n$, and $\mathbf{U}_{m,k,\tilde{n}} = \check{\mathbf{F}}_{m,k,n}^T \otimes \mathbf{I}_{N_U}$, $\forall \tilde{n} \neq n$.

Finally, we define the eigenvalue matrices as

$$\begin{aligned} \Lambda_{(m,k),n}^{\tilde{\mathbf{A}}} &= \left[\Lambda_{1,(m,k),n}^{\text{ICI}}, \dots, \Lambda_{m,(m,k),n}^{\text{ICI}}, \dots, \Lambda_{M,(m,k),n}^{\text{ICI}} \right], \\ \Lambda_{(m,k),n}^{\tilde{\mathbf{B}}} &= \left[\Lambda_{1,(m,k),n}^{\text{ICI}}, \dots, \check{\Lambda}_{m,(m,k),n}^{\text{ICI}}, \dots, \Lambda_{M,(m,k),n}^{\text{ICI}} \right], \end{aligned} \quad (59)$$

where $N_U \times N_U KN$ eigenvalue matrix $\Lambda_{m',(m,k),n}^{\text{ICI}}$ and $N_U \times N_U (K - 1)N$ eigenvalue matrix $\check{\Lambda}_{m,(m,k),n}^{\text{ICI}}$ are formed by stacking the eigenvalues of $\Sigma_{m',(m,k),n}^{\text{ICI}}$ and $\check{\Sigma}_{m,(m,k),n}^{\text{ICI}}$ in $N_U \times 1$ sized columns, respectively.

6.3. Rate Approximation under ICI. Using the second-order statistics derived above, and by extending Theorem 2, we present the rate approximation under ICI in the following corollary.

Corollary 4. For $N_B \geq N_U \rightarrow \infty$ we have

$$\frac{r_{(m,k),n}^{\text{ICI}} - \tilde{r}_{(m,k),n}^{\text{ICI}}}{N_U} \longrightarrow 0, \quad (60)$$

in which $\tilde{r}_{(m,k),n}^{\text{ICI}}$ is a deterministic function given by

$$\begin{aligned} \tilde{r}_{(m,k),n}^{\text{ICI}} &= - \sum_{j=1}^{N_U KNM} \log \left[\mathbf{v}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_j \\ &+ \sum_{j=1}^{N_U (KNM-1)} \log \left[\mathbf{v}_{(m,k),n}^{\tilde{\mathbf{B}}} \right]_j + \sum_{i=1}^{N_U} \left[\log \frac{\left[\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{B}}} \right]_i}{\left[\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_i} \right. \\ &\quad \left. + N_U \log e \left(\left[\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_i - \left[\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{B}}} \right]_i \right) \right] - N_U \\ &\quad \cdot \log (N_U (KNM - 1)) + N_U K NM \log \left(1 - \frac{1}{KNM} \right), \end{aligned} \quad (61)$$

where $\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{A}}} \in \mathbb{R}^{N_U}$, $\mathbf{v}_{(m,k),n}^{\tilde{\mathbf{A}}} \in \mathbb{R}^{KN_U MN}$, $\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{B}}} \in \mathbb{R}^{N_U}$, and $\mathbf{v}_{(m,k),n}^{\tilde{\mathbf{B}}} \in \mathbb{R}^{(KNM-1)N_U}$ are the solutions to the following fixed point equations:

$$\begin{aligned} \left[\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_i^{-1} &= N_U \left(1 + KNM N_U \left[\Lambda_{(m,k),n}^{\tilde{\mathbf{A}}} \mathbf{v}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_i \right), \\ \left[\mathbf{v}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_j^{-1} &= N_U K NM \left(1 + N_U \left[\Lambda_{(m,k),n}^{\tilde{\mathbf{A}}} \right]^T \mathbf{u}_{(m,k),n}^{\tilde{\mathbf{A}}} \right]_j, \\ \left[\mathbf{u}_{(m,k),n}^{\tilde{\mathbf{B}}} \right]_i^{-1} &= N_U \left(1 + N_U (KNM - 1) \left[\Lambda_{(m,k),n}^{\tilde{\mathbf{B}}} \mathbf{v}_{(m,k),n}^{\tilde{\mathbf{B}}} \right]_i \right), \\ \left[\mathbf{v}_{(m,k),n}^{\tilde{\mathbf{B}}} \right]_j^{-1} &= N_U (KNM - 1) \left(1 + N_U \left[\Lambda_{(m,k),n}^{\tilde{\mathbf{B}}} \right]^T \mathbf{u}_{(m,k),n}^{\tilde{\mathbf{B}}} \right)_j, \end{aligned} \quad (62)$$

where i and j range according to the vector sizes.

6.4. Weighted Sum Rate Optimization under ICI. After the change of variables to spherical coordinates in a similar manner to Section 4, the KKT conditions for (P3) are $\sum_{m=1}^M \sum_{k=1}^K w_{(m,k)} \sum_{n=1}^N \mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}^{\text{ICI}} = 0$, $\forall m'$. Under ICI, the rate is a function of precoders over all subcarriers and not just n , so the gradient is modified as

$$\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}^{\text{ICI}} = \sum_{n'=1}^N \mathcal{D}_{\bar{\mathbf{f}}_{m',n'}} \tilde{r}_{(m,k),n}^{\text{ICI}} \mathcal{D}_{\theta_{m'}} \bar{\mathbf{f}}_{m',n'}. \quad (63)$$

Here, due to the ICI, we need to sum over all subcarriers to calculate the gradient function. The above derivative can be obtained with the chain rule in a similar manner to Section 4; however, we need the derivatives with respect to all subcarriers and not just n , as is seen from (63). Another difficulty involved under ICI is the calculation of $\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \Lambda_{m',(m,k),n}^{\text{ICI}}$ and $\mathcal{D}_{\mathbf{F}_{(m,k'),n'}} \check{\Lambda}_{m,(m,k),n}^{\text{ICI}}$ which are rather different from the case without ICI. We shall only derive the above two terms under ICI because other terms are straightforward to find.

For general channels, it can be shown that

$$\begin{aligned} &\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \Lambda_{m',(m,k),n}^{\text{ICI}} \\ &= \begin{bmatrix} \text{vec}^T \left\{ \mathbf{y}_{m',(m,k),n}^1 \mathbf{y}_{m',(m,k),n}^1 \right\}^T \\ \vdots \\ \text{vec}^T \left\{ \mathbf{y}_{m',(m,k),n}^{KN_N^2} \mathbf{y}_{m',(m,k),n}^{KN_N^2} \right\}^T \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{n',1} \\ \vdots \\ \mathbf{Y}_{n',N} \end{bmatrix}, \\ &\mathcal{D}_{\mathbf{F}_{(m,k'),n'}} \check{\Lambda}_{m,(m,k),n}^{\text{ICI}} \\ &= \begin{bmatrix} \text{vec}^T \left\{ \check{\mathbf{y}}_{m,(m,k),n}^1 \check{\mathbf{y}}_{m,(m,k),n}^1 \right\}^T \\ \vdots \\ \text{vec}^T \left\{ \check{\mathbf{y}}_{m,(m,k),n}^{(KN-1)N_U^2} \check{\mathbf{y}}_{m,(m,k),n}^{(KN-1)N_U^2} \right\}^T \end{bmatrix} \begin{bmatrix} \check{\mathbf{Y}}_{n',1} \\ \vdots \\ \check{\mathbf{Y}}_{n',N} \end{bmatrix}, \end{aligned} \quad (64)$$

where $\mathbf{y}_{m',(m,k),n}^i$ and $\check{\mathbf{y}}_{m,(m,k),n}^i$ are the i th eigenvectors of $\Sigma_{m',(m,k),n}^{\text{ICI}}$ and $\check{\Sigma}_{m,(m,k),n}^{\text{ICI}}$, respectively, and $\check{\mathbf{Y}}_{n',\tilde{n}}$ are given by

$$\begin{aligned}
\mathbf{Y}_{n',\tilde{n}} &= \sqrt{\rho_{n',n}\rho_{\tilde{n},n}} \mathbf{K}_{KN_U^2,KN_N_U^2} \left[\mathbf{0} \left(\mathbf{F}_{m',\tilde{n}}^T \otimes \mathbf{I}_{N_U} \right)^* \boldsymbol{\Omega}_{m',(m,k),n',\tilde{n}}^T \otimes \mathbf{I}_{KN_U^2} \mathbf{0} \right]^T \left(\mathbf{I}_{N_B} \otimes \mathbf{K}_{N_U,KN_N_U} \right) \\
&\times \left(\mathbf{I}_{KN_U N_B} \otimes \text{vec} \left\{ \mathbf{I}_{N_U} \right\} \right) \mathbf{K}_{N_B,KN_N_U} \left(\mathbf{e}_{k'}^K \otimes \mathbf{I}_{N_U N_B} \right), \\
\check{\mathbf{Y}}_{n',\tilde{n}} &= \sqrt{\rho_{n',n}\rho_{\tilde{n},n}} \times \begin{cases} \mathbf{K}_{(K-1)N_U^2,(KN-1)N_U^2}, & \tilde{n} = n' \\ \mathbf{K}_{KN_U^2,(KN-1)N_U^2}, & \tilde{n} \neq n' \end{cases} \times \left[\mathbf{0} \mathbf{U}_{\tilde{n}}^T \otimes \mathbf{I}_{(K-1)N_U^2} \mathbf{0} \right]^T \left(\mathbf{I}_{N_B} \otimes \mathbf{K}_{N_U,(K-1)N_U} \right) \\
&\times \left(\mathbf{I}_{(K-1)N_U N_B} \otimes \text{vec} \left\{ \mathbf{I}_{N_U} \right\} \right) \mathbf{K}_{N_B,(K-1)N_U} \left[\mathbf{0} \mathbf{I}_{N_U N_B} \mathbf{0} \right]^T.
\end{aligned} \tag{65}$$

For Kronecker channels, it can be shown that $\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \Lambda_{m',(m,k),n}^{\text{ICI}}$ and $\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \check{\Lambda}_{m',(m,k),n}^{\text{ICI}}$ are given by

$$\begin{aligned}
\mathcal{D}_{\mathbf{F}_{(m',k'),n'}} \Lambda_{m',(m,k),n}^{\text{ICI}} &= \left(\mathbf{I}_{KN_U N} \otimes \boldsymbol{\lambda}_{m',(m',k')}^R \right) \\
&\cdot \begin{bmatrix} \text{vec}^T \left\{ \boldsymbol{\beta}_{m',(m,k),n}^1 \right. & \left. \boldsymbol{\beta}_{m',(m,k),n}^1 \right\}^T \\ \vdots \\ \text{vec}^T \left\{ \boldsymbol{\beta}_{m',(m,k),n}^{KN_U N} \right. & \left. \boldsymbol{\beta}_{m',(m,k),n}^{KN_U N} \right\}^T \end{bmatrix} \left(\mathbf{W}_{m',(m,k),n} \right. \\
&\otimes \mathbf{I}_{KN_U N}) \times \mathbf{K}_{LN_B,KN_U N} \mathbf{e}_{(n'-1)K+k'}^{NK} \\
&\otimes \left[\sqrt{\rho_{n',n}} \left(\mathbf{K}_{N_U,L} \otimes \mathbf{I}_{N_B} \right) \right. \\
&\cdot \left(\text{vec} \left\{ \boldsymbol{\Upsilon}_{m',(m,k)}^{1/2} \right\} \mathbf{w}_n \otimes \mathbf{I}_{N_B N_U} \right) \\
&\cdot \left. \left(\mathbf{I}_{N_U} \otimes \mathbf{T}_{m',(m,k)}^{1/2} \right) \right], \\
\mathcal{D}_{\mathbf{F}_{(m,k'),n'}} \check{\Lambda}_{m',(m,k),n}^{\text{ICI}} &= \left(\mathbf{I}_{N_U(KN-1)} \otimes \boldsymbol{\lambda}_{m',(m,k)}^R \right) \\
&\cdot \begin{bmatrix} \text{vec}^T \left\{ \check{\boldsymbol{\beta}}_{m,(m,k),n}^1 \right. & \left. \check{\boldsymbol{\beta}}_{m,(m,k),n}^1 \right\}^T \\ \vdots \\ \text{vec}^T \left\{ \check{\boldsymbol{\beta}}_{m,(m,k),n}^{N_U(KN-1)} \right. & \left. \check{\boldsymbol{\beta}}_{m,(m,k),n}^{N_U(KN-1)} \right\}^T \end{bmatrix} \left(\check{\mathbf{W}}_{m,(m,k),n} \right. \\
&\otimes \mathbf{I}_{N_U(KN-1)}) \times \mathbf{K}_{LN_B,N_U(KN-1)} \mathbf{v}_{k',n'} \\
&\otimes \left[\sqrt{\rho_{n',n}} \left(\mathbf{K}_{N_U,L} \otimes \mathbf{I}_{N_B} \right) \right. \\
&\cdot \left. \left(\text{vec} \left\{ \boldsymbol{\Upsilon}_{m',(m,k)}^{1/2} \right\} \mathbf{w}_n \otimes \mathbf{I}_{N_B N_U} \right) \left(\mathbf{I}_{N_U} \otimes \mathbf{T}_{m',(m,k)}^{1/2} \right) \right]
\end{aligned} \tag{66}$$

in which

$$\begin{aligned}
\mathbf{W}_{m',(m,k),n} &= \begin{bmatrix} \sqrt{\rho_{1,n}} \mathbf{w}_1^T \boldsymbol{\Upsilon}_{m',(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m',1} \mathbf{T}_{m',(m,k)}^{1/2} \right)^T \\ \vdots \\ \sqrt{\rho_{N,n}} \mathbf{w}_N^T \boldsymbol{\Upsilon}_{m',(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m',N} \mathbf{T}_{m',(m,k)}^{1/2} \right)^T \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\check{\mathbf{W}}_{m,(m,k),n} &= \begin{bmatrix} \sqrt{\rho_{1,n}} \mathbf{w}_1^T \boldsymbol{\Upsilon}_{m,(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m,1} \mathbf{T}_{m,(m,k)}^{1/2} \right)^T \\ \vdots \\ \sqrt{\rho_{N,n}} \mathbf{w}_N^T \boldsymbol{\Upsilon}_{m,(m,k)}^{1/2} \otimes \left(\bar{\mathbf{F}}_{m,N} \mathbf{T}_{m,(m,k)}^{1/2} \right)^T \end{bmatrix},
\end{aligned} \tag{67}$$

and $\boldsymbol{\beta}_{m',(m,k),n}^i$ and $\check{\boldsymbol{\beta}}_{m,(m,k),n}^i$ are the i th eigenvectors of $\mathbf{W}_{m',(m,k),n}$ and $\check{\mathbf{W}}_{m,(m,k),n}$, respectively, and

$$\begin{aligned}
\mathbf{v}_{k',n'} &= \begin{cases} \mathbf{e}_{(n'-1)K+k'-1}^{NK-1}, & (n-1)K+k < (n'-1)K+k', \\ \mathbf{e}_{(n'-1)K+k'}^{NK-1}, & (n-1)K+k > (n'-1)K+k', \\ \mathbf{0}_{NK-1}, & (n-1)K+k = (n'-1)K+k'. \end{cases} \tag{68}
\end{aligned}$$

The computational complexity is greatly reduced for the Kronecker channel model from $\mathcal{O}(K^2 N_U^4 N)$ to $\mathcal{O}(LN_B KN_U N)$.

We can now propose Algorithm 2 to find a suboptimal solution to (P3). Note that this algorithm is computationally more cumbersome compared to Algorithm 1, since the number of components in $\mathcal{D}_{\theta_m} \tilde{r}_{(m,k),n}^{\text{ICI}}$ is N times larger due to the ICI.

7. Simulation Results

In this section, we demonstrate the approximation accuracy and evaluate the performance of our algorithms. We consider a downlink multicell MIMO-OFDM system, where the number of cells and users in each cell is set to be 4, and the number of subcarriers is set to 8, that is, $M = 4$, $K = 4$, and $N = 8$. The number of antennas for users and BSs is set to 3 unless stated otherwise; that is, $N_B = 3$ and $N_U = 3$. The users are uniformly distributed in cells with 0.5 Km radius. Without loss of generality, the weighting for the sum rate maximization is uniform in the simulations, that is, $w_{(m,k)} = 1/NM$, $\forall m, k$. We assume the power constraints are the same

```

(1) Initialize: initialize  $\theta_{m'}(0)$ ,  $\forall m'$  and  $t \leftarrow 0$ ;
(2) while stopping criterion not met do
(3)    compute  $\mathbf{F}_{(m',k'),n}$ ,  $\forall m', k', n$  using (28), (30);
(4)    solve for  $\mathbf{u}_{(m,k),n}^A, \mathbf{v}_{(m,k),n}^A, \mathbf{u}_{(m,k),n}^B, \mathbf{v}_{(m,k),n}^B$ ,  $\forall m, k, n$ , using Corollary 4;
(5)    compute  $\mathcal{D}_{\theta_{m'}} \tilde{r}_{(m,k),n}^{\text{ICI}}$ ,  $\forall m', m, k, n$  using (63);
(6)     $\theta_{m'}(t+1) \leftarrow \theta_{m'}(t) + \alpha(t) \sum_{m,k,n} w_{(m,k)} \mathcal{D}_{\theta_{m'}}^T \tilde{r}_{(m,k),n}^{\text{ICI}}$ ,  $\forall m'$  and  $t \leftarrow t+1$ ;
(7) end while
(8) return  $\mathbf{F}_{(m',k'),n}$ ,  $\forall m'$ ;

```

ALGORITHM 2: Gradient search solution to problem (P3).

for all BSs, that is, $P_m = P$, $\forall m$. Since the noise power was normalized to unity, the transmitter signal-to-noise ratio is $\text{SNR} = 10 \log_{10} P$.

We consider the Kronecker channel model introduced in Section 5 with correlation matrices

$$\begin{aligned} [\mathbf{T}_{m',(m,k)}]_{i,j} &= \left(\frac{\alpha}{d_{m',(m,k)}} \right)^2 \rho_T^{|i-j|}, \\ [\mathbf{R}_{m',(m,k)}]_{i,j} &= \left(\frac{\alpha}{d_{m',(m,k)}} \right)^2 \rho_R^{|i-j|}, \end{aligned} \quad (69)$$

in which $d_{m',(m,k)}$ denotes the distance between BS m' and user (m, k) , and ρ_T, ρ_R are the spatial correlation factors as defined in [44]. The constant α is chosen so that $\mathbb{E}[(\alpha/d_{m',(m,k)})^2] = 1$. While not explicitly specified, we also choose ρ_T, ρ_R to be random and uniformly distributed in the interval $[0.5, 1]$ for each user. The tap correlation matrix is chosen as

$$[\mathbf{Y}_{m',(m,k)}]_{i,j} = \sqrt{\gamma_i \gamma_j} \rho_Y^{|i-j|}, \quad (70)$$

where ρ_Y is the tap correlation factor and $\sum_{i=1}^L \gamma_i = 1$ with $\gamma_i/\gamma_{i+1} = 3$ dB as defined in [34]. The number of channel taps is assumed to be 3 throughout our simulations; that is, $L = 3$. The ICI is modelled according to (51).

7.1. Approximation Accuracy. We compare the exact ergodic sum rate obtained by simulations and its approximation given by Theorem 2 or Corollary 4. The results are depicted in Figure 2, where it is seen that the approximations are extremely accurate. Note that although the approximations are asymptotic, exceptional accuracy is observed for our small size MIMO system ($N_B = 3, N_U = 3$); this fact justifies this approximation approach for practical antenna sizes. It is also seen that the approximation is accurate across a wide range of SNRs.

7.2. System Performance. Now we demonstrate the performance of our proposed method given by Algorithm 1, which is based only on statistical CSI at the BSs. For comparison, we shall also depict the results of two other methods. First, the results of the algorithm are based on perfect instantaneous CSI proposed in [12], which assumes the same system model

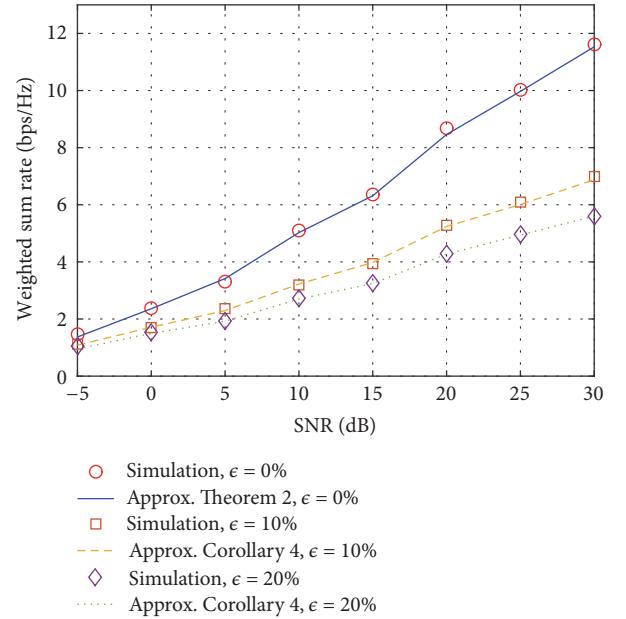


FIGURE 2: Approximation accuracy for the ergodic sum rate.

as this work but with full CSI at the BSs. Second, the results of networked MIMO are based on statistical CSI, where, in addition to the channel statistics, it is assumed that each BS has all the transmit data, so the whole system can be viewed as a MIMO super-cell with statistical CSI at the transmitter. The above two systems obviously require much heavier BS coordination, but as we shall see, our method competes with them. As mentioned before, our proposed algorithm is not limited to multicell networks with partial cooperation, so with small modifications it can be used for precoder design based on statistical CSI for networked MIMO systems.

Figure 3 depicts the results where it is observed that although our method uses statistical CSI with limited BS cooperation, the achievable weighted sum rate is comparable to that with full CSI and networked MIMO for a wide range of SNRs; note that statistical CSI incurs a much lower signaling overhead and requires less frequent update compared to full CSI. Here, we also see the results for different tap correlation factors. While tap correlations degrade system performance, using our method to incorporate the correlation information

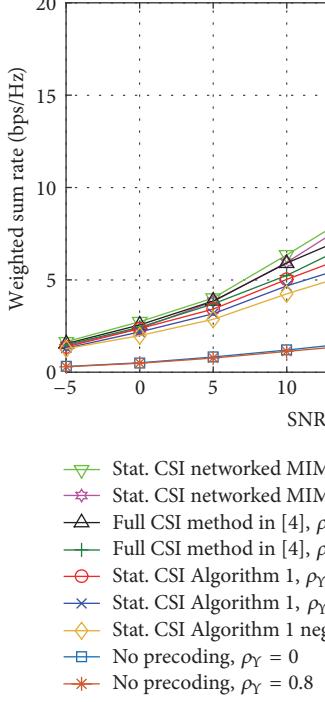


FIGURE 3: System performance under different conditions.

into precoder design alleviates the performance loss as seen in Figure 3. Moreover, it is observed that lack of knowledge about the tap correlations or neglecting them (assuming uncorrelated channel taps) leads to performance loss. When no precoders are employed (identity matrix precoding), there is a huge disadvantage, so precoder design is crucial in the MIMO-OFDM system.

7.2.1. Correlation Factors. In Figures 4 and 5, we investigate the effect of tap correlation and spatial correlation factors on the system performance. We see that both tap correlation and spatial correlation decrease the weighted sum rate. However, the spatial correlation shows more prominent effect on the sum rate. It is seen that as long as the correlation information is incorporated into precoder design via our proposed algorithm, the performance loss is not significant for a wide range of correlation degrees. It is noteworthy to mention that the performance loss becomes slightly larger at lower SNRs.

7.2.2. Number of Antennas. Figure 6 shows the sum rate as the number of transmitter antennas N_B increases; here the number of receiver antennas is fixed to $N_U = 2$. The weighted sum rate is seen to increase with N_B . As it is seen, the slope of performance increase for Algorithm 1 and that of the method in [12] based on full CSI is similar as N_B increases. This suggests that our proposed method is reliable for arbitrary large MIMO-OFDM systems.

7.3. System Performance under ICI. Now we present the system performance when there is ICI. Due to the cross

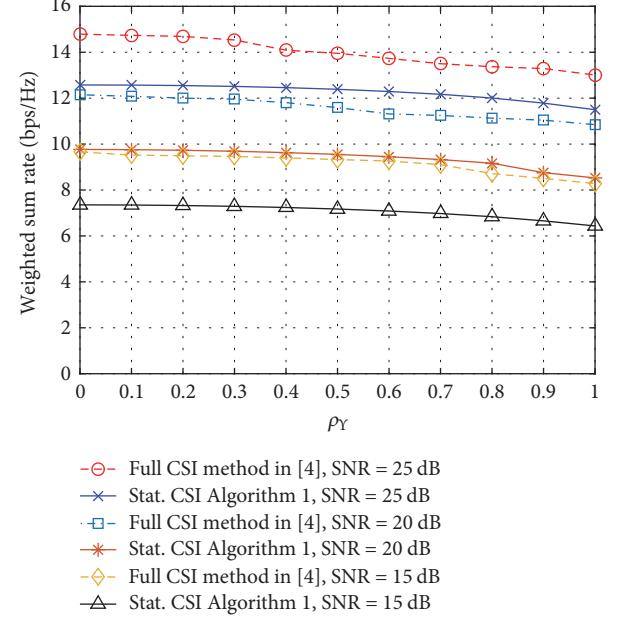


FIGURE 4: Impact of tap correlations.

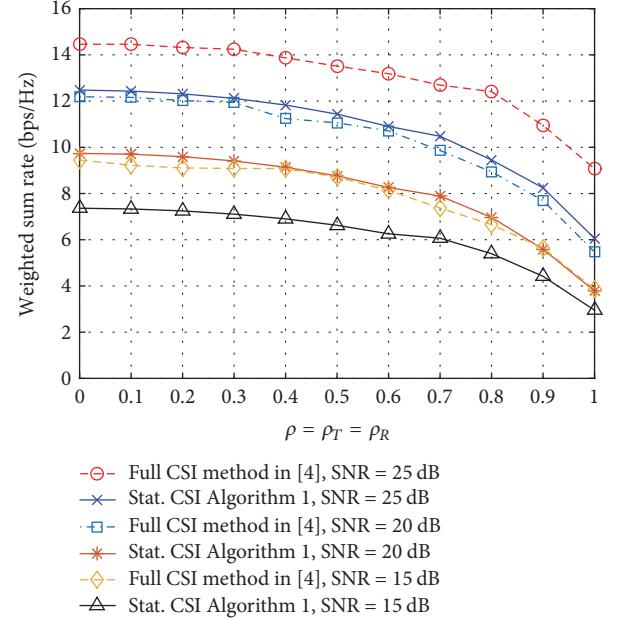


FIGURE 5: Impact of spatial correlations.

interference from other subcarriers, the sum rate decreases as is seen in Figure 7. It is seen that, for nonnegligible FO, there can be a serious decrease in the sum rate. However, smart precoder design implemented by Algorithm 2 can suppress ICI and achieve a reasonable sum rate. It is observed that the performance gain of our method over the nonprecoding scheme increases with SNR. Moreover, the advantages is more pronounced for smaller frequency offsets. However, we note that the gap between ICI-free system and system with ICI also increases with SNR.

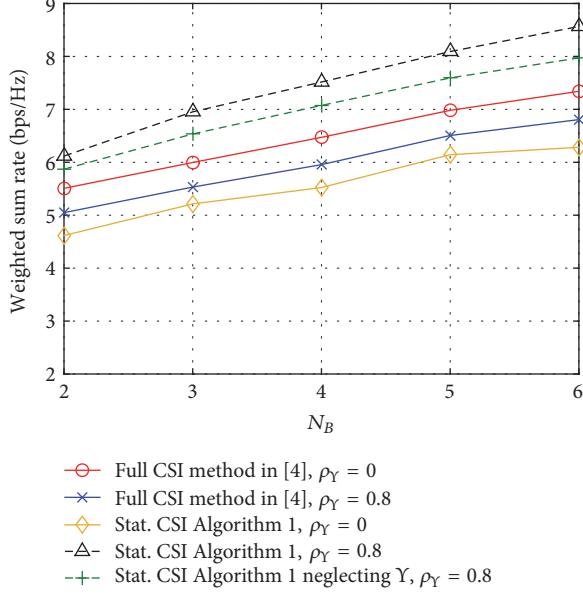


FIGURE 6: Impact of number of antennas.

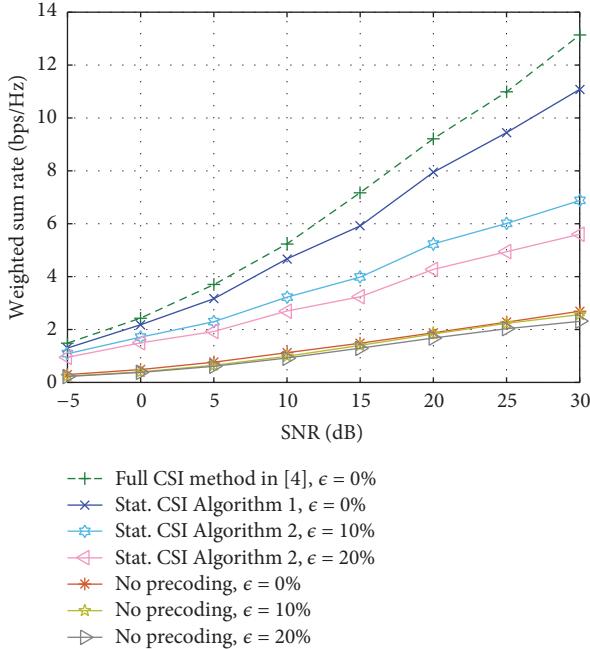


FIGURE 7: System performance under ICI.

In Figure 8, we plot the weighted sum rate versus the ICI intensity factor for $\text{SNR} = 15 \text{ dB}$. We see that the system sum rate is highly sensitive to ICI for small ϵ .

7.4. Convergence Rate. Now we compare the convergence rates for the Algorithms 1 and 2 for $\text{SNR} = 15 \text{ dB}$ and $\epsilon = 0\%, 1\%, 10\%, \text{ and } 20\%$. The initial precoders are chosen to be the identity matrix. Under various degrees of ICI, we see from Figure 9 that Algorithm 1 always converges faster

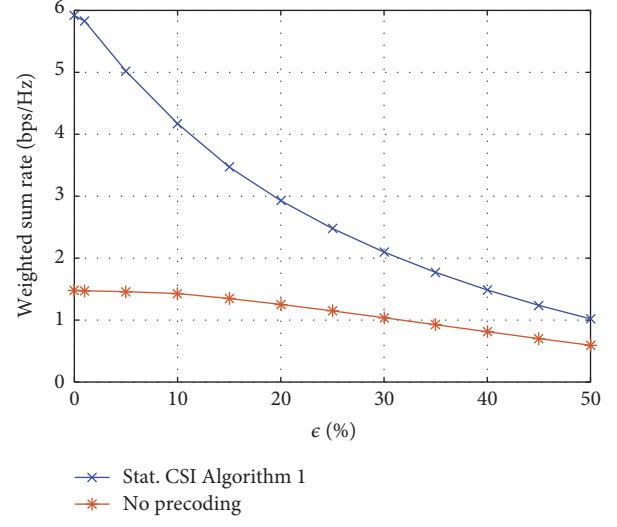


FIGURE 8: Impact of ICI intensity.

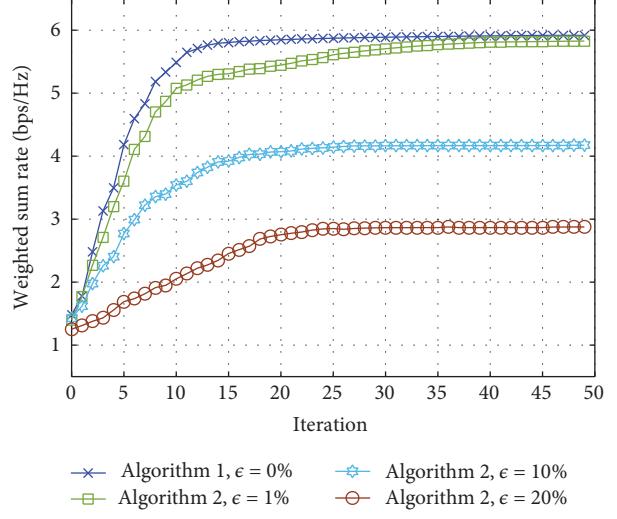


FIGURE 9: Convergence rate of Algorithms 1 and 2.

than Algorithm 2. When there is ICI, the convergence rate of Algorithm 2 is similar for various degrees of ICI.

8. Conclusions

We investigated linear precoding for downlink multicell MIMO-OFDM systems based on statistical CSI. The main contribution of this work was applying the already established results of random matrix theory to the MIMO-OFDM scenario in order to study the impact of frequency selectivity, tap correlations, and ICI on the statistical precoder design and system performance. The asymptotic approximations to the ergodic rates in ICI and ICI-free scenarios were derived, based on which, we formulated two nonconvex sum rate maximization problems and proposed locally optimum gradient based solutions to them. Simulation results showed that while spatial correlations, tap correlations, and ICI decrease

the system sum rate, our method alleviates this performance loss by incorporating the correlation information and ICI intensity information into the precoder design.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Leading Talents Program of Guangdong Province under Grant 00201510, the Basic Research Program of Shenzhen under Grant JCYJ20151117161854942, and the Shenzhen Peacock Program under Grant KQTD2015071715073798.

References

- [1] D. N. Dao and C. Tellambura, "Intercarrier interference self-cancellation space-frequency codes for MIMO-OFDM," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1729–1738, 2005.
- [2] J. M. García-Loygorri, L. de Haro, J. M. Riera, L. Cuéllar, and C. Rodríguez, "Keyholes in MIMO-OFDM: Train-to-Wayside Communications in Railway Tunnels," *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–7, 2017.
- [3] Y. Jin and X.-G. Xia, "A robust precoder design based on channel statistics for MIMO-OFDM systems with insufficient cyclic prefix," *IEEE Transactions on Communications*, vol. 62, no. 4, pp. 1249–1257, 2014.
- [4] C. Xu, J. Zhang, M. Liu, and C. Yin, "Pilot design for sparse channel estimation in large-scale MIMO-OFDM system," *International Journal of Antennas and Propagation*, vol. 2016, Article ID 6142574, 8 pages, 2016.
- [5] P. Xia, S. Zhou, and G. B. Giannakis, "Adaptive MIMO-OFDM based on partial channel state information," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 202–213, 2004.
- [6] R. El Chall, F. Nouvel, M. Hélard, and M. Liu, "Performance and complexity evaluation of iterative receiver for coded MIMO-OFDM systems," *Mobile Information Systems*, vol. 2016, Article ID 7642590, 22 pages, 2016.
- [7] J. Abdoli, M. Jia, and J. Ma, "Filtered OFDM: A new waveform for future wireless systems," in *Proceedings of the 16th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2015*, pp. 66–70, Sweden, July 2015.
- [8] P. N. Rani and C. S. Rani, "UFMC: The 5G modulation technique," in *Proceedings of the 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–3, Chennai, December 2016.
- [9] M. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications Magazine*, vol. 13, no. 4, pp. 56–61, 2006.
- [10] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. J. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2983–2989, 2009.
- [11] W. Wu, K. Wang, W. Zeng, Z. Ding, and C. Xiao, "Cooperative multi-cell MIMO downlink precoding with finite-alphabet inputs," *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 766–779, 2015.
- [12] K. Wang, X. Wang, W. Xu, and X. Zhang, "Coordinated linear precoding in downlink multicell MIMO-OFDMA networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4264–4277, 2012.
- [13] L. Sanguinetti, R. Couillet, and M. Debbah, "Large system analysis of base station cooperation for power minimization," *IEEE Transactions on Wireless Communications*, vol. 99, p. 1, 2016.
- [14] D. J. Love, R. W. Heath Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, 2008.
- [15] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4298–4310, 2010.
- [16] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.
- [17] E. Yoon, J. Hansen, and A. Paulraj, "Space-frequency preceding for an OFDM based system exploiting spatial and path correlation," in *Proceedings of the GLOBECOM'04 - IEEE Global Telecommunications Conference*, pp. 436–440, December 2004.
- [18] C. Artigue and P. Loubaton, "On the precoder design of flat fading MIMO systems equipped with MMSE receivers: a large-system approach," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 57, no. 7, pp. 4138–4155, 2011.
- [19] S. A. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1165–1175, 2004.
- [20] H. Wu, X. Gao, X. Wang, and X. You, "Sum-rate-optimal pre-coding for multi-cell large-scale MIMO uplink based on statistical CSI," *IEEE Transactions on Communications*, vol. 63, no. 8, pp. 2924–2935, 2015.
- [21] A. M. Tulino and S. Verdú, "Random Matrix Theory and Wireless Communications," *Foundations and Trends™ in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [22] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge University Press, Cambridge, UK, 2011.
- [23] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 51, no. 7, pp. 2491–2509, 2005.
- [24] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*, Springer Series in Statistics, Springer, New York, Second edition, 2010.
- [25] R. Couillet, M. Debbah, and J. . Silverstein, "A deterministic equivalent for the analysis of correlated MIMO multiple access channels," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 57, no. 6, pp. 3493–3514, 2011.
- [26] J. Zhang, C.-K. Wen, S. Jin, X. Gao, and K. Wong, "On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 133–148, 2013.
- [27] F. Dupuy and P. Loubaton, "On the capacity achieving covariance matrix for frequency selective MIMO channels using the asymptotic approach," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 57, no. 9, pp. 5737–5753, 2011.

- [28] S. Chatterjee, "A generalization of the Lindeberg principle," *Annals of Probability*, vol. 34, no. 6, pp. 2061–2076, 2006.
- [29] H. Wu, L. Wang, X. Wang, and X. You, "Asymptotic and non-asymptotic analysis of uplink sum rate for relay-assisted MIMO cellular systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 6, pp. 1348–1360, 2014.
- [30] Z. Liu, Y. Xin, and G. B. Giannakis, "Space-time-frequency coded OFDM over frequency-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2465–2476, 2002.
- [31] C. Xiao and Y. R. Zheng, "On the ergodic capacity of MIMO triply selective rayleigh fading channels," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2272–2279, 2008.
- [32] V. K. V. Gottumukkala and H. Minn, "Ergodic capacity analysis of MISO/SIMO-OFDM with arbitrary antenna and channel tap correlation," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 7, pp. 3062–3068, 2013.
- [33] J. Pang, J. Li, and L. Zhao, "Ergodic capacity formula of MIMO-OFDM systems under correlated frequency selective Rayleigh fading channels," *Journal of Electronics (China)*, vol. 24, no. 5, pp. 588–592, 2007.
- [34] E. Yoon, J. Hansen, and A. Paulraj, "Space-frequency precoding with space-tap correlation information at the transmitter," *IEEE Transactions on Communications*, vol. 55, no. 9, pp. 1702–1711, 2007.
- [35] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 191–193, 1995.
- [36] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2451–2464, 2002.
- [37] S. Diggavi, N. Al-Dhahir, and A. Stamoulis, "Intercarrier interference in MIMO OFDM," in *Proceedings of the 2002 International Conference on Communications (ICC 2002)*, pp. 485–489, usa, May 2002.
- [38] P. H. Moose, "Technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, no. 10, pp. 2908–2914, 1994.
- [39] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002.
- [40] V. L. Girko, "Random determinants," in *Probability theory. Mathematical statistics. Theoretical cybernetics, Vol. 24 (Russian)*, Itogi Nauki i Tekhniki, pp. 3–51, i, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1986.
- [41] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, part 1, pp. 2740–2746, 2007.
- [42] X. Wang, V. Krishnamurthy, and J. Wang, "Stochastic gradient algorithms for design of minimum error-rate linear dispersion codes in MIMO wireless systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1242–1255, 2006.
- [43] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, 2008.
- [44] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-time Wireless Communication*, Cambridge University Press, 2003.

Research Article

Diversity-Multiplexing-Nulling Trade-Off Analysis of Multiuser MIMO System for Intercell Interference Coordination

Jinwoo Kim¹ and Chung G. Kang²

¹IT and Mobile Communication Division, Samsung Electronics, Soowon, Republic of Korea

²School of Electrical Engineering, Korea University, Seoul, Republic of Korea

Correspondence should be addressed to Chung G. Kang; ccgkang@korea.ac.kr

Received 16 September 2017; Accepted 6 November 2017; Published 10 December 2017

Academic Editor: Mostafa Zaman Chowdhury

Copyright © 2017 Jinwoo Kim and Chung G. Kang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A fundamental performance trade-off of multicell multiuser multiple-input multiple-output (MU-MIMO) systems is explored for achieving intercell and intracell interference-free conditions. In particular, we analyze the three-dimensional diversity-multiplexing-nulling trade-off (DMNT) among the diversity order (i.e., the slope of the error performance curve), multiplexing order (i.e., the number of users that are simultaneously served by MU-MIMO), and nulling order (i.e., the number of users with zero interference in a victim cell). This trade-off quantifies the performance of MU-MIMO in terms of its diversity and multiplexing order, while nulling the intercell interference toward the victim cell in the neighbor. First, we design a precoding matrix to mitigate both intercell and intracell interference for a linear precoding-based MU-MIMO system. Then, the trade-off relationship is obtained by analyzing the distribution of the signal-to-noise ratio (SNR) at the user terminals. Furthermore, we demonstrate how DMNT can be applied to estimate the long-term throughput for each mobile station, which allows for determining the optimal number of multiplexing order and throughput loss due to the interference nulling.

1. Introduction

Multiuser MIMO (MU-MIMO) scheme, allowing a base station (BS) to communicate with multiple users simultaneously, provides an opportunity to boost the sum capacity through precoding, even when each user has only one antenna. For example, zero-forcing transmit beamforming (ZFBF) is one of the practical multiuser transmission strategies for MU-MIMO systems [1]. By designing one user's beamforming vector to be orthogonal to other selected users' channel vectors, ZFBF can completely eliminate the multiuser interference corresponding to intracell interference in cellular systems. Furthermore, using more transmit antenna can increase the number of users simultaneously served by MU-MIMO or enhance the error performance of each link between the BS and user.

Despite the theoretical attractiveness, the capacity gain promised by MIMO techniques has been shown to degrade severely in a multicell environment. To suppress the intercell interference, the authors in [2–5] investigated a coordinated

beamforming scheme using multiple antennas at the BS. The achievable rate region of the MISO interference channel, in the case where the full channel information is shared among BSs, was derived in [2, 3], with instantaneous and statistical CSI, respectively. Distributed beamforming with a virtual SINR framework was proposed in [4]. The theoretical results in [2–4], however, are limited to only one user in the victim cell. The authors in [5] assumed that the interference experienced by multiple users in the victim cells is suppressed. Some studies on the interference mitigation in the cooperative beamforming for multiuser systems have been studied from the perspective of scheduling issues [6–9]. A low-complexity random beamforming method, which only requires sharing of user indices, has been suggested with analytic throughput expressions [6]. In [7], the authors provided a transmission beamforming scheme for interference nulling with user selection. Also, the reduced complexity algorithms for joint user selection in adaptive coordination scheme were designed in [8]. However, the unrealistic special homogeneous case, in which all users have the same average SNR, is assumed in [8].

In [9], we have considered a generalized intercell interference coordination problem and proposed a two-step coordination procedure to choose a cell-edge user and decide the coordination. However, the research in [6–9] was not extended to MU-MIMO, that is, only dealing with the multiple users in the serving cell.

In this paper, we analyze the three-dimensional DMNT among the diversity order D (i.e., slope of error performance curve), multiplexing order L (i.e., the number of users simultaneously served by MU-MIMO scheme), and nulling order N (i.e., the number of other cell users subject to zero interference in a victim cell), while providing the victim cell with intercell interference nulling. We consider an interference-free environment in which the BS in each cell employs a precoding matrix with N_T antennas, so as to null the intracell interference while mitigating the intercell interference. It is assumed that all users are equipped with a single antenna. Our contribution is to reveal the fundamental property of performance trade-off, given by $N_T = D + L + N - 1$, in multi-cell MU-MIMO subject to intercell and intracell interference-free conditions. To the best of our knowledge, there has never been any rigorous justification for this particular property in previous works. Note that the current diversity-multiplexing-nulling trade-off (DMNT) is quite different from the well-known diversity-multiplexing trade-off (DMT) that deals with the multiple antenna gain to be achieved simultaneously by any coding scheme (e.g., space-time coding) in the point-to-point MIMO system [10]. Meanwhile, we demonstrate that our DMNT can be applicable to estimating the long-term user throughput, which allows for determining the optimal multiplexing order and throughput loss due to interference nulling.

The rest of this paper is organized as follows. We present some preliminaries for our analysis in Section 2. Section 3 presents a system model and the precoding matrix design under consideration. Our analysis results for the diversity-multiplexing-nulling trade-off are given in Section 4. Section 5 demonstrates how DMNT can be applied to estimate the long-term throughput for each MS. Finally, concluding remarks are given in Section 6.

2. Preliminaries

We first introduce some results from previous works, which are useful for our analysis.

Definition 1. Let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ be independent complex Gaussian random vectors with zero mean vector (i.e., $E\{\mathbf{x}_i\} = \mathbf{0}_p$) and identity covariance matrix (i.e., $E\{\mathbf{x}_i \mathbf{x}_i^H\} = \mathbf{I}_p$). If $\mathbf{A} = \mathbf{X} \mathbf{X}^H$, where $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$ is the $p \times n$ matrix, then \mathbf{A} is said to have a *complex Wishart distribution* with n degrees of freedom [11], that is, $\mathbf{A} \sim W_p(n)$.

By using n vectors, we may define the matrix \mathbf{A} as

$$\mathbf{A} = \mathbf{X} \mathbf{X}^H = \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^H. \quad (1)$$

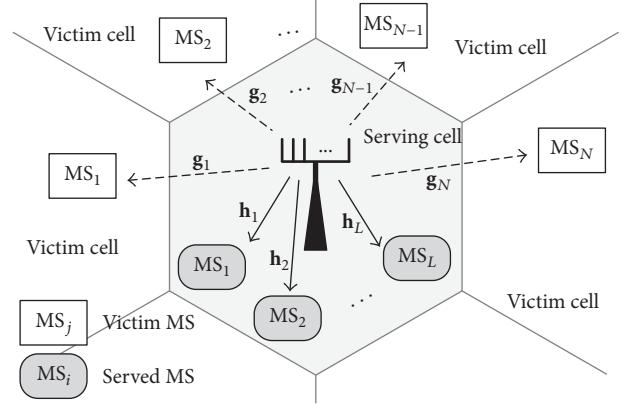


FIGURE 1: System model: serving cell and victim cells.

The probability density function of \mathbf{A} for $n \geq p$ is given as

$$p_W(\mathbf{A}) = \frac{\exp(-\text{tr}(\mathbf{A})) (\det \mathbf{A})^{n-p}}{\pi^{p(p-1)/2} \prod_{j=1}^p \Gamma(n-j+1)}, \quad (2)$$

where $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$. If $p = 1$, that is, \mathbf{A} is defined as a random variable, then it has a Chi-squared distribution with $2n$ degrees of freedom, that is, $\mathbf{A} \sim \chi^2(2n)$. This result immediately follows by substituting $p = 1$ into the probability density function (pdf) of the Wishart distribution.

Lemma 2. If $\mathbf{A} \sim W_p(n)$ and \mathbf{A} is partitioned as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad (3)$$

where \mathbf{A}_{11} is $k \times k$ and the Schur complement of block \mathbf{A}_{22} is also a Wishart matrix with a distribution of $W_k(n-p+k)$. In [12], the Schur complement of block \mathbf{A}_{22} is given as

$$\tilde{\mathbf{A}}_{11} = \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21}. \quad (4)$$

Proof. See proof of Theorem 3.2.10 in [13]. \square

3. Signal Model and Precoding Matrix Design

We consider MU-MIMO downlink systems in which the BS serves a set of selected mobile stations (MSs) simultaneously in a serving cell, while imparting interference to the MSs in victim cells, as illustrated in Figure 1. We assume that there are K and N MSs in the serving cell and victim cells, respectively. Let \mathcal{S} be a subset of indices for users that are intended for transmission by the BS ($\mathcal{S} \subset \{1, 2, \dots, K\}$, $|\mathcal{S}| = L \leq K$). The user set \mathcal{S} is dynamically selected by a scheduler in the BS. At the serving cell, we design a wireless link equipped with N_T transmit antennas at the BS and a single receive antenna at each MS. The MSs in the victim cells also employ a single receive antenna and do not perform any type of interference mitigation. Let us denote \mathbf{h}_k and \mathbf{w}_k as $N_T \times 1$ complex Gaussian channel vector and beamforming vector for the k th MS, respectively. For a subset \mathcal{S} , we define $\mathbf{H}(\mathcal{S}) =$

$[\mathbf{h}_1 \dots \mathbf{h}_L]$ and $\mathbf{W}(\mathcal{S}) = [\mathbf{w}_1 \dots \mathbf{w}_L]$. The received signal y_l at the l th MS in the subset \mathcal{S} is represented by

$$y_l = \mathbf{h}_l^H \mathbf{w}_l s_l + \sum_{j=1, j \neq l}^L \mathbf{h}_l^H \mathbf{w}_j s_j + n_l, \quad l = 1, 2, \dots, L \in \mathcal{S}, \quad (5)$$

where s_l and n_l are the data symbol and the Additive White Gaussian Noise (AWGN) with variance of N_0 , respectively. We impose average power constraints, $\|\mathbf{w}_l\|^2 = 1$ and $E\{|s_l|^2\} = E_s$. The received signals in (5) are rewritten by the aggregated received signal vector \mathbf{y} as

$$\mathbf{y} = \mathbf{H}^H \mathbf{W} \mathbf{s} + \mathbf{n}, \quad (6)$$

where $\mathbf{s} = [s_1 \dots s_L]^T$ and $\mathbf{n} = [n_1 \dots n_L]^T$. Let \mathbf{g}_j denote an $N_T \times 1$ channel vector from the serving BS to the j th MS in the victim cell. Note that N MSs in the victim cells are those subject to intercell interference. In our current system model in Figure 1, N MSs in the virtual victim cell can be considered as those multiplexed with L MSs in the serving cell while satisfying the intercell interference-free condition. For the MSs in the victim cell, the aggregated received signal vector $\tilde{\mathbf{y}}$ is given by

$$\tilde{\mathbf{y}} = \mathbf{G}^H \mathbf{W} \mathbf{s}, \quad (7)$$

where $\mathbf{G} = [\mathbf{g}_1 \dots \mathbf{g}_N]$, which can be known to the serving BS by a sounding signal [14]. Note that a desired signal of the victim MS is not represented in (7); that is, $\tilde{\mathbf{y}}$ is just an intercell interference vector for the victim MSs, which would be controlled by the BS in the serving cell.

Our proposed precoding matrix design focuses on achieving the interference-free communication for both interferences from the other cell and other user signal from the serving BS. First, the intercell interference-free condition is given as

$$\tilde{\mathbf{y}} = \mathbf{G}^H \mathbf{W} \mathbf{s} = \mathbf{0}_N, \quad (8)$$

where $\mathbf{0}_N$ denotes an $N \times 1$ column vector with all-zero elements. Our objective is to rigorously show how much the spatial degrees of freedom are lost in this situation. Note that the intercell interference-free condition (8) leads to the following proposition:

Proposition 3. *To satisfy the intercell interference-free condition, $\mathbf{W} \mathbf{s}$ should lie in the null space of \mathbf{G} , that is, the orthogonal complement of the subspace \mathcal{G} spanned by column vectors $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N$ of \mathbf{G} .*

From Proposition 3, the precoding matrix \mathbf{W} can be a cascade of matrices, $\mathbf{W}^{\text{inter}}$ and $\mathbf{W}^{\text{intra}}$, that is, $\mathbf{W} = \mathbf{W}^{\text{inter}} \mathbf{W}^{\text{intra}}$, where $\mathbf{W}^{\text{inter}}$ eliminates the intercell interference by adopting a projection matrix onto the orthogonal complement (PMOC) of \mathcal{G} . If $\mathbf{W}^{\text{inter}}$ is a PMOC of \mathcal{G} , $\mathbf{W} \mathbf{s}$ lies in the null space of \mathbf{G} , regardless of $\mathbf{W}^{\text{intra}}$. Therefore, we may design two precoding matrices, $\mathbf{W}^{\text{inter}}$ and $\mathbf{W}^{\text{intra}}$, independently, so as to meet each design constraint.

To produce a PMOC of \mathcal{G} , consider the following QR decomposition of \mathbf{G} :

$$\mathbf{G} = \mathbf{Q} \mathbf{R}, \quad (9)$$

where \mathbf{Q} is an $N_T \times N_T$ unitary matrix and \mathbf{R} is an $N_T \times N$ upper triangular matrix. As the bottom $(N_T - N)$ rows of \mathbf{R} consist of entire zeroes, it is often useful to partition \mathbf{Q} and \mathbf{R} as follows:

$$\mathbf{G} = \mathbf{Q} \mathbf{R} = \mathbf{Q} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} = [\mathbf{Q}_1 \ \mathbf{Q}_2] \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}_1 \mathbf{R}_1, \quad (10)$$

where \mathbf{R}_1 is an $N \times N$ upper triangular matrix, \mathbf{Q}_1 is $N_T \times N$ orthogonal matrix, \mathbf{Q}_2 is $N_T \times (N_T - N)$, and both \mathbf{Q}_1 and \mathbf{Q}_2 have orthogonal columns. Let \mathbf{P}_G^\perp denote a PMOC of \mathcal{G} . Then, we can obtain \mathbf{P}_G^\perp by using \mathbf{Q}_1 or \mathbf{Q}_2 as follows:

$$\mathbf{P}_G^\perp = \mathbf{I}_{N_T} - \mathbf{Q}_1 \mathbf{Q}_1^H = \mathbf{Q}_2 \mathbf{Q}_2^H. \quad (11)$$

Using $\mathbf{Q}_1 = \mathbf{G} \mathbf{R}_1^{-1}$ and $\mathbf{R}_1 = \mathbf{Q}_1^H \mathbf{G}$, alternative form can be represented as

$$\mathbf{P}_G^\perp = \mathbf{I}_{N_T} - \mathbf{Q}_1 \mathbf{Q}_1^H = \mathbf{I}_{N_T} - \mathbf{G} (\mathbf{G}^H \mathbf{G})^{-1} \mathbf{G}^H. \quad (12)$$

Let us now design $\mathbf{W}^{\text{intra}} = [\mathbf{w}_1^{\text{intra}} \dots \mathbf{w}_L^{\text{intra}}]$, whose purpose is to avoid interuser interference. Denoting $\tilde{\mathbf{H}}^H = \mathbf{H}^H \mathbf{P}_G^\perp = [\tilde{\mathbf{h}}_1 \ \tilde{\mathbf{h}}_2 \ \dots \ \tilde{\mathbf{h}}_L]^H$ for ZFBF, beamforming vectors are selected such that they satisfy the zero-interference condition $\tilde{\mathbf{h}}_k^H \mathbf{w}_l^{\text{intra}} = 0$ for $k \neq l$; that is, the beamforming vector for user l lies in the null space spanned by $\{\tilde{\mathbf{h}}_k, \forall k \neq l\}$. One easy choice of the precoding matrix that gives zero interuser interference is the pseudoinverse; that is,

$$\tilde{\mathbf{W}} = \tilde{\mathbf{H}} (\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1} = [\tilde{\mathbf{w}}_1 \ \tilde{\mathbf{w}}_2 \ \dots \ \tilde{\mathbf{w}}_L]. \quad (13)$$

The precoding matrix $\mathbf{W}^{\text{intra}}$ is formed by the unit-normalized columns of $\tilde{\mathbf{W}}$; that is, $\mathbf{w}_l^{\text{intra}} = \tilde{\mathbf{w}}_l / \|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_l\|$. In matrix form, $\mathbf{W}^{\text{intra}}$ is given by

$$\mathbf{W}^{\text{intra}} = \left[\frac{\tilde{\mathbf{w}}_1}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_1\|} \ \frac{\tilde{\mathbf{w}}_2}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_2\|} \ \dots \ \frac{\tilde{\mathbf{w}}_L}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_L\|} \right]. \quad (14)$$

Finally, the aggregated received signal vector \mathbf{y} is given by

$$\begin{aligned} \mathbf{y} &= \tilde{\mathbf{H}}^H \mathbf{W}^{\text{intra}} \mathbf{s} + \mathbf{n} \\ &= \tilde{\mathbf{H}}^H \left[\frac{\tilde{\mathbf{w}}_1}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_1\|} \ \dots \ \frac{\tilde{\mathbf{w}}_L}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_L\|} \right] \mathbf{s} + \mathbf{n} \\ &= \tilde{\mathbf{H}}^H \tilde{\mathbf{W}} \cdot \text{diag} \left(\frac{1}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_1\|}, \dots, \frac{1}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_L\|} \right) \mathbf{s} + \mathbf{n}. \end{aligned} \quad (15)$$

Since $\tilde{\mathbf{W}} = \tilde{\mathbf{H}} (\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1}$, the received signal y_l is given as

$$y_l = \frac{s_l}{\|\mathbf{P}_G^\perp \tilde{\mathbf{w}}_l\|} + n_l. \quad (16)$$

4. Analysis of Diversity-Multiplexing-Nulling Trade-Off (DMNT)

The effective SNR at the l th MS is given by $\gamma_l = \bar{\gamma}_l / \| \mathbf{P}_G^\perp \tilde{\mathbf{w}}_l \|^2$, where $\bar{\gamma}_l = E_s/N_0$ denotes the average SNR. As $\mathbf{P}_G^\perp \tilde{\mathbf{W}} = [\mathbf{P}_G^\perp \tilde{\mathbf{w}}_1 \ \mathbf{P}_G^\perp \tilde{\mathbf{w}}_2 \ \dots \ \mathbf{P}_G^\perp \tilde{\mathbf{w}}_L]$, we have

$$\| \mathbf{P}_G^\perp \tilde{\mathbf{w}}_l \|^2 = \left[(\mathbf{P}_G^\perp \tilde{\mathbf{W}})^H (\mathbf{P}_G^\perp \tilde{\mathbf{W}}) \right]_{ll}, \quad (17)$$

where $[\mathbf{X}]_{ll}$ denotes an (l, l) -diagonal element of \mathbf{X} . Using (17) and $\tilde{\mathbf{W}} = \tilde{\mathbf{H}}(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1}$, we have

$$\gamma_l = \frac{\bar{\gamma}_l}{\left[(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1} \tilde{\mathbf{H}}^H \mathbf{P}_G^{\perp H} \mathbf{P}_G^\perp \tilde{\mathbf{H}} (\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1} \right]_{ll}}. \quad (18)$$

Now, defining $\bar{\mathbf{H}} := \tilde{\mathbf{H}}^H$, we have

$$\gamma_l = \frac{\bar{\gamma}_l}{\left[(\bar{\mathbf{H}} \bar{\mathbf{H}}^H)^{-1} \bar{\mathbf{H}} \mathbf{P}_G^{\perp H} \mathbf{P}_G^\perp \bar{\mathbf{H}}^H (\bar{\mathbf{H}} \bar{\mathbf{H}}^H)^{-1} \right]_{ll}}. \quad (19)$$

Meanwhile, $\bar{\mathbf{H}} = \mathbf{H}^H \mathbf{P}_G^\perp$ from $\tilde{\mathbf{H}}^H = \mathbf{H}^H \mathbf{P}_G^\perp$ and $\bar{\mathbf{H}} = \tilde{\mathbf{H}}^H$, which gives

$$\begin{aligned} \bar{\mathbf{H}} \mathbf{P}_G^{\perp H} \mathbf{P}_G^\perp \bar{\mathbf{H}}^H &= \mathbf{H}^H \mathbf{P}_G^{\perp H} \mathbf{P}_G^{\perp H} \mathbf{P}_G^\perp \mathbf{P}_G^{\perp H} \mathbf{H} = \mathbf{H}^H \mathbf{P}_G^{\perp H} \mathbf{P}_G^{\perp H} \mathbf{H} \\ &= \bar{\mathbf{H}} \bar{\mathbf{H}}^H. \end{aligned} \quad (20)$$

The last step follows from the properties of the projection matrix; that is, $\mathbf{P}_G^{\perp H} = \mathbf{P}_G^\perp$ and $\mathbf{P}_G^{\perp 2} = \mathbf{P}_G^\perp$. Finally, we have

$$\gamma_l = \frac{\bar{\gamma}_l}{\left[(\bar{\mathbf{H}} \bar{\mathbf{H}}^H)^{-1} \right]_{ll}} = \frac{\bar{\gamma}_l}{\left[(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1} \right]_{ll}} \quad \left(\because \tilde{\mathbf{H}} = \bar{\mathbf{H}}^H \right). \quad (21)$$

Let $\mathbf{Z} = \tilde{\mathbf{H}}^H \tilde{\mathbf{H}}$. Without loss of generality, we obtain the SNR for $l = 1$ as follows:

$$\gamma_1 = \frac{\bar{\gamma}_1}{[\mathbf{Z}^{-1}]_{11}} = \frac{\bar{\gamma}_1 \det(\mathbf{Z})}{\text{adj}(\mathbf{Z})_{11}}. \quad (22)$$

The last equality in (22) follows from Cramer's rule [14] and the (i, j) -element of $\text{adj}(\mathbf{Z})$ is given by

$$\text{adj}(\mathbf{Z})_{ij} = (-1)^{i+j} \det(\mathbf{Z}_{ji}), \quad (23)$$

where \mathbf{Z}_{ji} is the $(L - 1) \times (L - 1)$ matrix formed by deleting the j th row and i th column of \mathbf{Z} . When \mathbf{Z} is partitioned as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_{11} & \mathbf{z}_{12} \\ \mathbf{z}_{21} & \mathbf{z}_{22} \end{bmatrix}, \quad (24)$$

$\text{adj}(\mathbf{Z})_{11} = \det(\mathbf{Z}_{22})$ from (23). Thus, (22) can be expressed as

$$\gamma_1 = \frac{\bar{\gamma}_1 \det(\mathbf{Z})}{\det(\mathbf{Z}_{22})}. \quad (25)$$

To proceed, we use the following property:

$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det(\mathbf{D}) \det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}). \quad (26)$$

This in turn yields

$$\gamma_1 = \bar{\gamma}_1 \det(z_{11} - \mathbf{z}_{12} \mathbf{Z}_{22}^{-1} \mathbf{z}_{21}) = \bar{\gamma}_1 \tilde{z}_{11}, \quad (27)$$

where $\tilde{z}_{11} = z_{11} - \mathbf{z}_{12} \mathbf{Z}_{22}^{-1} \mathbf{z}_{21}$; that is, \tilde{z}_{11} is the Schur complement of block \mathbf{Z}_{22} . We can show that the distribution of γ_1 is given by the following lemma, which will be useful for analyzing the DMNT.

Lemma 4. γ_1 has a chi-squared distribution with $2(N_T - N - L + 1)$ degrees of freedom.

Proof. See Appendix. \square

It is well-known that the diversity order (corresponding to a slope of an error performance curve) is n when the SNR is distributed as a chi-squared distribution with $2n$ degrees of freedom [15]. By Lemma 4, therefore, the diversity order of the interference-free MU-MIMO system under consideration is given by $N_T - N - L + 1$, which leads to the following result for the DMNT.

Theorem 5. Let N , L , and D denote the nulling order, multiplexing order, and diversity order for the downlink BS in the multicell MU-MIMO system. Assuming a BS equipped with N_T antennas and MSs equipped with a single antenna, the trade-off among the nulling order, multiplexing order, and diversity order for the interference-free condition is given by

$$N_T = D + L + N - 1. \quad (28)$$

5. Application of DMNT: Asymptotic Throughput Analysis

In this section, DMNT in Theorem 5 can be applied to estimate the long-term throughput for each MS when an opportunistic ZFBF scheduling is applied. It deals with an asymptotic throughput analysis, from which the optimal number of multiplexing order and the throughput loss due to the interference nulling can be achieved. In fact, it will be a useful analytical framework that can predict how DMN must be traded off so as to maximize the cell throughput or to maintain a required level of cell throughput.

Let $\gamma_l(\mathcal{S})$ represent the effective SNR for the l th user when \mathcal{S} is a set of the selected users that are transmitted at the same time, as defined in Section 3. The objective of the opportunistic scheduling is to determine a subset of users, \mathcal{S} , such that the sum rate is maximized; that is,

$$\mathcal{S}^* = \arg \max_{\mathcal{S}} \sum_{l=1}^L R_l(\mathcal{S}), \quad (29)$$

where $R_l(\mathcal{S}) = \log_2(1 + \gamma_l(\mathcal{S}))$. Assuming $L = 2$ (scalable to $L > 2$ with the same principle), the long-term throughput \bar{R}_k is given by

$$\bar{R}_k = \Pr(k \in \mathcal{S}^*) E\{R_k(\mathcal{S}^*) | k \in \mathcal{S}^*\}. \quad (30)$$

Since the channel vectors have independent and identical distribution, $\Pr(\mathcal{S}^* = \{k, 1\}) = (1/(K-1)) \cdot \Pr(k \in \mathcal{S}^*)$ and thus (30) can be expressed as

$$\begin{aligned} \bar{R}_k &= (K-1) \Pr(\mathcal{S}^* = \{k, 1\}) E\{R_k(\mathcal{S}^*) | \mathcal{S}^* = \{k, 1\}\}. \end{aligned} \quad (31)$$

Furthermore, since $\Pr(\mathcal{S}^* = \{k, 1\}) = 1/K(K-1)$, we have

$$\begin{aligned} \bar{R}_k &= \frac{1}{K} E\{R_k(\mathcal{S}^*) | \mathcal{S}^* = \{k, 1\}\} = \frac{1}{K} \\ &\cdot E\{R_k(\mathcal{S}^*) | R_k(\{k, 1\}) + R_1(\{k, 1\}) \geq R_i(\{i, 1\}) \\ &+ R_1(\{i, 1\}), \forall i \neq k\}. \end{aligned} \quad (32)$$

As $K \rightarrow \infty$, there will be multiple MSs (i.e., k th MS and others) who have near orthogonal channel vector with respect to MS 1. Thus, we can assume that $R_1(\{k, 1\}) \approx R_1(\{i, 1\})$ for sufficiently large K , reducing (32) into

$$\begin{aligned} \bar{R}_k &\approx \frac{1}{K} E\{R_k(\mathcal{S}^*) | R_k(\{k, 1\}) \geq R_i(\{i, 1\}), \forall i \neq k\} \\ &= \frac{1}{K} E\{R_k(\mathcal{S}^*) | \gamma_k \geq \gamma_i, \forall i \neq k\}. \end{aligned} \quad (33)$$

Equation (33) implies that, for large K , the long-term throughput can be obtained by a single user scheduling problem in which each MS has a diversity order of $(N_T - N - L + 1)$, as dictated by Theorem 5 and therefore, it follows the chi-squared distribution with $2(N_T - N - L + 1)$ degrees of freedom. We note that the long-term throughput for the single user scheduling can be obtained by using the extreme value theory [16]. Based on the results in [16], the analytic throughput result is obtained as

$$\begin{aligned} \bar{R}_k &= \frac{1}{K} \log_2 \left(1 + F_{\gamma_k}^{-1} \left(\frac{K-1}{K} \right) \right) + \frac{1}{K} \\ &\times E_0 \left(\log_2 \left(1 + F_{\gamma_k}^{-1} \left(\frac{K \cdot e - 1}{K \cdot e} \right) \right) \right. \\ &\left. - \log_2 \left(1 + F_{\gamma_k}^{-1} \left(\frac{K-1}{K} \right) \right) \right), \end{aligned} \quad (34)$$

where $F_{\gamma_k}(\gamma)$ is the CDF of γ_k in Lemma 4 and $E_0 = 0.5772$, which is the Euler constant [17].

In order to validate the accuracy of our asymptotic throughput analysis, we simulate the case with $N_T = 8$, $N = 2$, and $L = 2$, and uniformly distributed 30 MSs; that is, $K = 30$. Since all users have the different average SNR in the simulation, the different average throughput is observed for individual user as shown in Figure 2. We find that throughput measurement for each MS is acceptably close to the analytical result. Furthermore, the average cell throughput $\sum_k \bar{R}_k$ can be determined by the asymptotic throughput analysis for $N_T = 16$ and $K = 100$, as shown in Figures 3 and 4. Figure 3 shows that there exists a multiplexing order to maximize the cell throughput for the given nulling order, demonstrating the optimal number of multiplexing order and the throughput loss due to the interference nulling. For example,

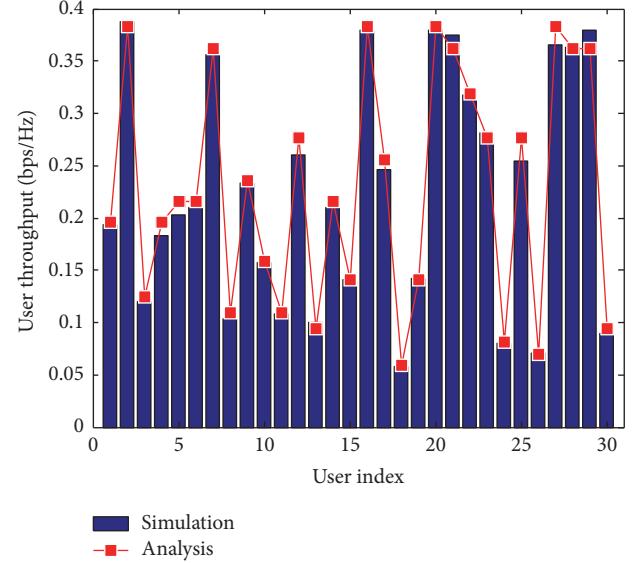


FIGURE 2: Throughput estimation for each MS: analysis versus simulation.

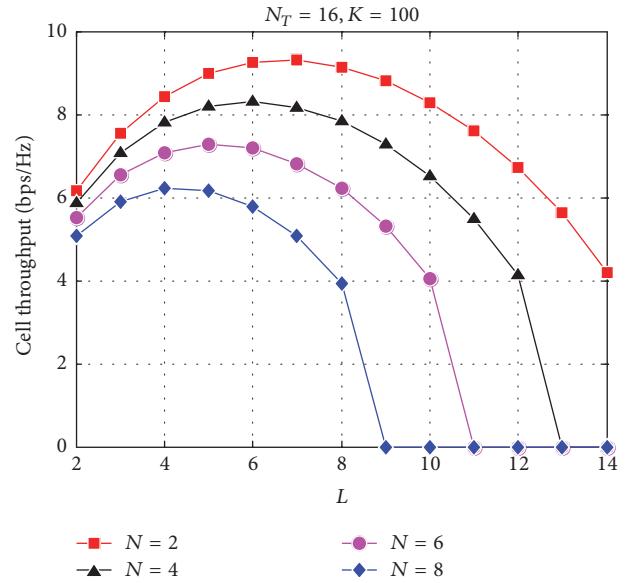


FIGURE 3: Cell throughput: diversity order (D) versus multiplexing order (L).

when two other cell users are nulled (i.e., $N = 2$), the optimal number of multiplexing gains is 7; that is, 7 active users must be served in the reference cell ($L = 7$). Furthermore, it is obvious from Figure 3 that the average cell throughput is reduced as increasing the nulling order. The current analysis implies that the multiplexing order should be adjusted adaptively for the different nulling order. Furthermore, it is observed that the cell throughput gained by reducing the interference-nulling effect subject to the same diversity gain becomes more conspicuous as the diversity order increases; that is, the multiplexing order decreases. Figure 4 shows the same observation for varying the nulling order for

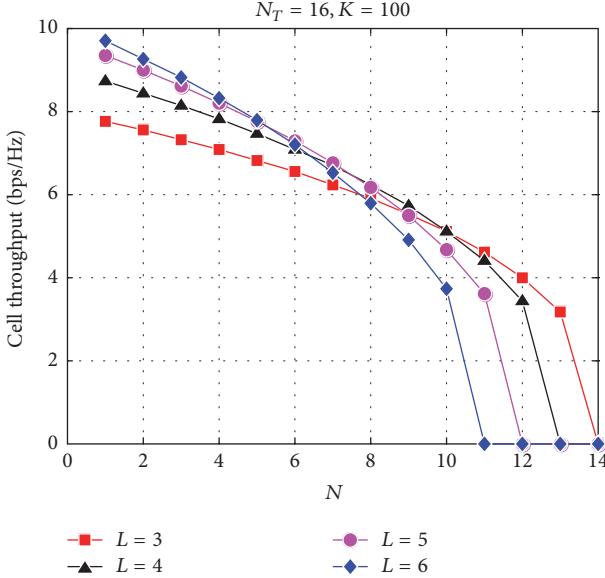


FIGURE 4: Cell throughput: diversity order (D) versus nulling order (N).

the given multiplexing order. Unlike Figure 3, there is no optimal operating point between N and D . Furthermore, it is observed that the multiplexing gain contributes to the cell throughput differently, depending on the nulling order. For example, spatial multiplexing is more important than diversity in improving the cell throughput if the nulling order is small. Otherwise, diversity gain is more contributing to the cell throughput than multiplexing gain. Meanwhile, the maximum nulling order N that is required to maintain the minimum target cell throughput can be determined by the curves in Figure 4.

6. Conclusion

In this paper, we have investigated a fundamental property of performance trade-off in the multicell multiuser multiple-input multiple-output (MU-MIMO) system when the intercell and intracell interference-free conditions must be satisfied. Assuming a BS equipped with N_T antennas and MSs equipped with a single antenna, the trade-off among the nulling order (N), multiplexing order (L), and diversity order (D) for the interfere-free condition is given by $N_T = D + L + N - 1$. By characterizing the three-dimensional diversity-multiplexing-nulling trade-off (DMNT), our analysis provides a quantitative framework for dealing with the intercell interference coordination in a multicell MU-MIMO system. Finally, we have demonstrated that our DMNT result can be used to perform an asymptotic throughput analysis, which predicts how DMN must be traded off so as to maximize the cell throughput or to maintain a required level of cell throughput in the system. The overall system performance can be optimized by selecting a set of users and the precoders at the same time. As the joint optimization involves enormous complexity, joint optimization seems to be unrealistic in practice. It will be worth investigating the suboptimal

approach that can be implemented in practice. However, user selection and codebook design subject to the limited feedback is beyond our scope in this paper. In other words, even if the overall system performance is governed by user selection and codebook design, along with channel estimation error, we aimed at demonstrating Diversity-Multiplexing-Nulling Trade-off (DMN) as an ideal performance characteristic, which would play a fundamental design guideline.

The interference-free environment (realized through the interference-nulling effect) may not be straightforward to achieve in practice. Since we assume that a full CSI is available at BS, an interesting venue for future work is to design a limited feedback system and, furthermore, to analyze the performance subject to channel estimation error. Whereas the current asymptotic analysis illustrates throughput of the reference cell only, we may need a multicell coordination framework in which all neighboring cells are coordinated to determine the optimal DMNT for the overall throughput maximization in practice.

Appendix

Proof of Lemma 4

Since $\tilde{\mathbf{H}}^H = \mathbf{H}^H(\mathbf{I}_{N_T} - \mathbf{Q}\mathbf{Q}^H)$, \mathbf{Z} can be expressed as

$$\mathbf{Z} = \mathbf{H}^H(\mathbf{I}_{N_T} - \mathbf{Q}_1\mathbf{Q}_1^H)\mathbf{H}. \quad (\text{A.1})$$

In (A.1), $\mathbf{Q}_1\mathbf{Q}_1^H$ can be eigen-decomposed as $\mathbf{Q}_1\mathbf{Q}_1^H = \mathbf{X}\Sigma\mathbf{X}^H$ where \mathbf{X} is a unitary matrix and Σ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues. Using $\mathbf{Q}\mathbf{Q}^H = \mathbf{X}\Sigma\mathbf{X}^H$, (A.1) can be expressed as

$$\begin{aligned} \mathbf{Z} &= \widehat{\mathbf{H}}^H(\mathbf{I}_{N_T} - \Sigma)\widehat{\mathbf{H}} \\ &= [\widehat{\mathbf{h}}_1 \dots \widehat{\mathbf{h}}_{N_T}] (\mathbf{I}_{N_T} - \Sigma) [\widehat{\mathbf{h}}_1 \dots \widehat{\mathbf{h}}_{N_T}]^H. \end{aligned} \quad (\text{A.2})$$

Given the orthogonality of \mathbf{X} , $\mathbf{X}^H\mathbf{H}$ is also an i.i.d. complex Gaussian channel matrix. Defining $\widehat{\mathbf{H}} = \mathbf{X}^H\mathbf{H} = [\widehat{\mathbf{h}}_1 \ \widehat{\mathbf{h}}_2 \ \dots \ \widehat{\mathbf{h}}_{N_T}]^H$, we have

$$\mathbf{Z} = \widehat{\mathbf{H}}^H(\mathbf{I}_{N_T} - \Sigma)\widehat{\mathbf{H}}. \quad (\text{A.3})$$

Meanwhile, \mathbf{Q}_1 consists of N orthonormal vectors. In other words, N eigenvalues of $\mathbf{Q}_1\mathbf{Q}_1^H$ have a value of 1, while the others are zero. Without loss of generality, we assume that the diagonal entries are sorted in increasing order. In matrix form, Σ is given by

$$\Sigma = \text{diag}\left(0, \dots, 0, \underbrace{1, \dots, 1}_{\text{length } N}\right). \quad (\text{A.4})$$

Applying this to (A.3), we have

$$\begin{aligned} \mathbf{Z} &= [\widehat{\mathbf{h}}_1 \dots \widehat{\mathbf{h}}_{N_T-N} \ \mathbf{0} \ \dots \ \mathbf{0}] [\widehat{\mathbf{h}}_1 \dots \widehat{\mathbf{h}}_{N_T}]^H \\ &= \sum_{i=1}^{N_T-N} \widehat{\mathbf{h}}_i \widehat{\mathbf{h}}_i^H. \end{aligned} \quad (\text{A.5})$$

From the definition of the Wishart distribution in (1), we conclude that \mathbf{Z} has a Wishart distribution with $(N_T - N)$ degrees of freedom. Furthermore, we also find by Lemma 2 that \tilde{z}_{11} in (27) has a Wishart distribution. Therefore, it is clear that $\tilde{z}_{11} \sim W_k(N_T - N - L + 1)$. By Definition 1, therefore, \tilde{z}_{11} is a chi-squared random variable with $2(N_T - N - L + 1)$ degrees of freedom.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (no. 2014-0-00282, Development of 5G Mobile Communication Technologies for Hyperconnected smart services) and in part by the Brain Korea 21 Plus Project in 2017.

References

- [1] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, 2006.
- [2] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, part 2, pp. 5292–5296, 2008.
- [3] J. Lindblom, E. Karipidis, and E. G. Larsson, "Selfishness and altruism on the MISO interference channel: the case of partial transmitter CSI," *IEEE Communications Letters*, vol. 13, no. 9, pp. 667–669, 2009.
- [4] R. Zakhour and D. Gesbert, "Coordination on the MISO interference channel using the virtual SINR framework," in *Proceedings of the International ITG Workshop on Smart Antennas*, Berlin, Germany, February 2009.
- [5] J. Zhang and J. G. Andrews, "Adaptive spatial intercell interference cancellation in multicell wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1455–1468, 2010.
- [6] J. Zhu and H.-C. Yang, "Low-complexity coordinated beamforming transmission for multiuser MISO systems and its performance analysis," in *Proceedings of the 53rd IEEE Global Communications Conference (GLOBECOM '10)*, Miami, Fla, USA, December 2010.
- [7] U. Jang, H. Son, J. Park, and S. Lee, "CoMP-CSB for ICI nulling with user selection," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 2982–2993, 2011.
- [8] S.-H. Moon, C. Lee, S.-R. Lee, and I. Lee, "A joint adaptive beamforming and user scheduling algorithm for downlink network MIMO systems," in *Proceedings of the IEEE International Conference on Communications (ICC '13)*, pp. 5392–5397, Budapest, Hungary, June 2013.
- [9] J. Kim, I. S. Hwang, and C. G. Kang, "Inter-cell coordinated beamforming with opportunistic scheduling," in *Proceedings of the 2013 IEEE International Conference on Communications (ICC '13)*, pp. 5699–5703, Budapest, Hungary, June 2013.
- [10] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, 2003.
- [11] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution. (an introduction)," *Annals of Mathematical Statistics*, vol. 34, pp. 152–177, 1963.
- [12] E. V. Haynsworth, "On the Schur complement," in *Basel Mathematical Notes*, University of Basel, June 1968.
- [13] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, New York, NY, USA, 1982.
- [14] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*, Wiley, 2009.
- [15] T. Muir, *The Theory of Determinants in the Historical Order of Development*, Dover Publications, New York, NY, USA, 1960.
- [16] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, UK, 2005.
- [17] G. Song and Y. Li, "Asymptotic throughput analysis for channel-aware scheduling," *IEEE Transactions on Communications*, vol. 54, no. 10, pp. 1827–1834, 2006.

Research Article

Interactive Smart Fashion Using User-Oriented Visible Light Communication: *The Case of Modular Strapped Cuffs and Zipper Slider Types*

Jai-Eun Kim,¹ Youn-Hee Kim,² Jung-Hun Oh,¹ and Ki-Doo Kim¹

¹Department of Electronics Engineering, Kookmin University, Seoul 02707, Republic of Korea

²Department of Fashion Design, Kookmin University, Seoul 02707, Republic of Korea

Correspondence should be addressed to Ki-Doo Kim; kdk@kookmin.ac.kr

Received 11 August 2017; Accepted 9 November 2017; Published 10 December 2017

Academic Editor: Ji-Woong Choi

Copyright © 2017 Jai-Eun Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Because LEDs offer flexible expressions such as brightness, color control, and various patterns, they are popularly used in multidevice interactions. Moreover, LEDs have excellent physical characteristics. However, existing LED light-based wearable interactions are designed for interest and attention. So, LED can be used in fashion as it can give new look to our style and at the same time also as an interaction device. Therefore, in this paper, we present the design guideline for regulating the technical implementation, design strategies, and directions of interactive LED devices. The technology and design concepts are demonstrated through a case study (analysis) of an existing LED light-based wearable interaction. We also design a scenario-based iterative collaborative design process model. Finally, we develop a smart fashion of modular strapped cuffs and zipper slider types that can be attached and detached according to the user's preference as the interactive smart fashion using user-oriented visible light communication, ultimately pursuing a visual-MIMO (Multiple-Input Multiple-Output) product through stepwise strategy.

1. Introduction

Wearable technology applied products are now contributing to high value-added businesses that cater to the emotional needs of users seeking creative newness as well as changing consumer lifestyles by fusing technology and fashion. The word “fashion,” in this paper, indicates new look of our dress combined with technology. The technology has two purposes: (i) giving our dress a distinctive style and (ii) sharing information about what we want to share. Because of this communication factor, we mention it as “smart fashion” or “interactive smart fashion” or “wearable interactions” or “wearable technology applied products.” A good example is smart fashion, which combines clothing with luminous bodies such as optical fibers, LEDs, and electroluminescent films. Smart fashion is being developed for various applications such as entertainment, communication, and safety protection.

The physical characteristics of LEDs are superior to those of other luminous bodies. Specifically, LEDs have long

lifetimes and low energy consumptions; moreover, they are more compact, thinner, lighter, and more environmentally friendly than other electronic devices. They react quickly and offer limitless interaction possibilities when combined with various sensors. In addition, the brightness, color, and pattern of LED-based interactions can be freely adjusted. Such flexible expression is ideal for realizing simultaneity in wearable interactions.

Research on light-applied smart fashion is still in the early stage and focuses on attracting the interest and attention of users rather than exploring commercialization possibilities. Current implementations are nonefficient, functionally limited, and cannot establish a social infrastructure. As such, they have failed to gain the understanding and sympathy of people. User-centric smart fashion that is easily accessible and human-friendly demands further research.

To acquire a solid position in the smart fashion market, designers must combine technology with consumer perceptions in the popular market rather than adopt a simple

technology-based approach [1]. To this end, we established an interdisciplinary research team of engineers and fashion designers and proposed smart fashion that detects the emotions of consumers and incorporates functionality and aesthetics into the products.

Our interdisciplinary research team has proposed visible light communication (VLC) for wearable interactions (smart fashion) through the following three-stage research strategy: (1) identifying the specialized factors of LED-based visual communication (information visualization) by surveying users; the users' needs then become the evaluation criteria for information visualization; (2) based on the user-needs survey results and the analysis results of interactive smart fashion, proposing design guidelines (technology and design concepts) for technical implementations and design strategies and directions; and (3) designing a process model for the scenario-based interactive collaborative design and developing a smart fashion of modular strapped cuffs and zipper slider types, which can be attached or detached by users. This VLC-based interactive smart fashion is developed via a stepwise strategy.

2. Related Works

VLC is not the latest topic but it has vast application. Visual-MIMO in smart fashion technology is a potential branch of future research. Researchers have tried to implement visual-MIMO in different applications. Corbellini et al. [2] described different user application like home networking, LED-to-LED communication for toys, interactive fashion and fabrics, toys without radio emission, and toy communication with smartphone. They gave an idea about fashion with VLC technology but no proper demonstration. Swartz et al. [3] designed wearable communication system that can control body attached device by personal area networks (PANs) and it has also access to the wide area networks (WANs). Hertleer et al. [4] introduced an intelligent fabric system that contains an antenna operating properly in the 2.4–2.4835-GHz industrial-scientific-medical bandwidth. The similar kind of work was done by Ito et al. [5]. They proposed a wearable jacket to employ endoscope system. Pyattaev et al. [6] showed current and emerging connectivity solutions for high-density wearable deployments, their relative performance, and open communication challenges. Pathak et al. [7] made survey paper on VLC system. They described VLC system components, communication medium for VLC, and the application of the VLC like indoor localization, gesture recognition, screen-camera communication, and vehicular networking.

Except [2], the other studies are about either VLC system or wearable devices for communication. None of them directly worked with both VLC and wearable technology. Reference [2] only mentioned VLC technology for interactive fashion design. We focused on the fusion of “wearable” technology and “LED communication.” And we developed the fashionable visual-MIMO prototype using user-oriented visible light communication. In our paper, we show that interactive fashion with VLC is good candidate for future fashion and technology industry.

3. Survey on LED Light for Usability

In interactive smart fashion, the diverse needs of users and the important elements in a usability assessment must be considered. To this end, our interdisciplinary research team conducted a preliminary survey on users' needs, which revealed the important factors reflecting the specialized characteristics of LED-based visual communication and used them as an evaluation scale for information visualization. Thus, users' opinions were actively collected and applied in the initial design stage, and their needs were analyzed through user-centric design approaches. The improvement factors were also identified.

The user-needs survey and analysis revealed three types of needs; infotainment (information + entertainment) communication, nonverbal communication for people with hearing impairments, and a wearable interface with detachable modules. The analysis results are outlined as follows:

- (1) Content Survey for Infotainment: a questionnaire based on a 5-point Likert scale was administered to 20 women in their twenties. The survey focused on the content elements needed for interactions. Eighty-five percent of the respondents reported that LED-based visual communication facilitates the sharing of sympathy and enjoyment. That is, most of the respondents desired both functionality and decorativeness of the LED element. Moreover, in the integration of communication devices into interactive fashion, the visual sensory element (the LED) was preferred over the tactile sensory element (vibration) (50% versus 45% of respondents). Forty-five and 30% of respondents chose the neckline and shoulders-and-arms (upper and lower arms), respectively, as the attachment site for an LED-based communication. Respondents also considered the around-the-arm interface format as the most recognizable in terms of usability.
- (2) Content Survey for Nonverbal Communication for People with Hearing Impairments: the questionnaire was carried out by categorizing the methods of LED-based visual communication and the preferred sensory parts of people with hearing impairments. The respondents were 30 members of the Korea Association of the Deaf in their twenties, thirties, or forties. All respondents expressed a need for smart fashion and had used a communication device or auxiliary equipment. Their scores were based on a 5-point Likert scale and in-depth interviews. The needs analysis revealed that smart fashion helps to resolve the communication disorders between people with and without hearing impairments. It also assists the hearing impaired individuals with their nonverbal communication through mutual feedback and allows them to partially communicate with individuals lacking knowledge of sign language. Finally, smart fashion promises to expand human functions and converts the information needed by people with hearing impairments into visual signals. These people communicate by cellular phone (46.6%), hearing aid

(26.6%), or both (26.8%). More than 80% of people with hearing impairments reported difficulty with communication exchanges. Respondents also desired to receive communications such as tactile (73.3%) and visual (26.7%) sensation through an intuitive interface. Visually perceptible wearable interface format, such as an LED, was rated satisfactory and average by 73.3% and 26.7% of respondents, respectively. These findings highlight the need for interactive smart fashion with an inbuilt visual communication feature, enabling intuitive interactions for people with hearing impairments.

- (3) Survey on Detachable Modules and Wearable Interface Designs: the subjects (32 men and women aged 20–28 years participated in a survey and they had an interest and experiences with smart fashion.) reported their responses on a questionnaire based on a 5-point Likert scale, and in interviews. Eighty-one percent of the respondents reacted positively to wearable interface designs with human-friendly fasteners (buttons, sleeve bands, zipper sliders, buckles, stoppers, and Velcro). The favored fasteners were buttons (38.5% of respondents), strapped cuffs (30.8%), and zipper sliders (28%). In addition, 83% of respondents preferred to attach and detach their favorite functions and designs according to time, place, and occasion (TPO). Reasons for this preference included detergency, ease of management, usability, economy, compatibility, and sociality. These findings highlight the need for further research.

Considering these findings, our research team set a direction of promoting a LED-based smart fashion market by developing a user-centric content with VLC that fuses engineering and fashion design. The resulting design will satisfy the formattiveness, wearability, purposefulness, economy, usability, satisfaction, and safety demands of the smart fashion market.

4. Method: Research through Explorative Design

4.1. LED as Design Material. LED, as a small device, can freely express the formative elements of line, surface, and space designs. They also create spatiotemporal patterns through light. Consistent with recent trends in emotion centric consumption, fashion design has increasingly focused on images and aesthetic sensitivity. With visualizing technologies of various colors, LED light-based wearable communication has the potential of creating a new market with original and creative high value-added fashion products that satisfy the fashion trends toward emotions and the smart fashion era. Recently, multidisciplinary researches have promoted these goals by exploiting the LED as design material. To highlight the use of LEDs as future design materials, this section introduces a typical case study of LED.

To improve the wearability of LEDs, researchers have applied textile-based methods or non-textile-based methods. First, the textile-based LED, based on a research on the

conductive textile, has been studied as one axis of the electronic components, such as transistors and sensors [8]. Among the textile-based methods, the polymer light emitting electrochemical cell (PLEC) method achieves flexible, lightweight LED-based materials that can be woven into textiles like conventional clothes and which provide the same brightness from any direction [9]. Furthermore, PLEC materials can be manufactured by a high-speed, low-cost, roll-to-roll process that applies the dip-coating techniques widely used in fiber, thread, and fabric dyeing. Figure 1 illustrates a color-adjustable textile fabricated by the PLEC process [9].

Rather than illuminating the textile itself, non-textile-based methods aim to improve the physical characteristics of the LEDs incorporated into wearable devices. Colloidal quantum dot LEDs (QLEDs) are non-textile-based devices with unique optoelectronic properties such as color tunability, narrow emission spectra, high quantum yield, and photo/air stability. Additional advantages include printability on various substrates, ultrathin active layers, and high luminescence at low operating voltages [10]. Recently, ultrathin and wearable RGB QLEDs have been fabricated by a high-resolution intaglio transfer printing technique. Such QLEDs can be laminated onto various soft and curvilinear surfaces without diminishing the high EL efficiency. High-definition, full-color deformable QLEDs are expected to be realized in the near future [10]. Figure 2 presents electronic tattoos based on ultrathin wearable RGB QLEDs constructed by the high-resolution intaglio transfer printing technique [10].

We expect that previous research will enable free design variations of wearable devices and also will provide more solutions to the designer of wearable devices by promoting the physical properties of LEDs in a wearable-friendly way.

4.2. Conventional Wearable Interactions Based on Led Light. With the growth of the wearable device market, now, it is not difficult to experience LED-based wearable interaction in the vicinity. This section analyzes existing LED-based interactive wearable devices, discusses their technical issues and the problems of conventional wearable interactions, and proposes new directions.

In the following representative cases, the functionality and decorativeness of LED light-based wearable technology are demonstrated in entertainment, communication, and safety functions.

Figure 3 presents Lumalive fabric with an entertainment function, which is used in the Lumalive textile garments of Philips [11]. This fabric creates dynamic advertising, graphics, and constantly changing color expressions through a luminous and flexible LED array inserted in the textile.

Figure 4 presents an interactive sound shirt with a communication function for people with hearing impairments [12]. By wearing this shirt, users can “listen” to an orchestra by feeling the vibrations of the instruments, which activate different parts of the shirt. An orchestral concert for people with hearing impairments was jointly organized and held by an orchestra team “Junge Symphoniker Hamburg” for people with hearing impairments and a company “CuteCircuit” which specializes in wearable technologies. The “Sound

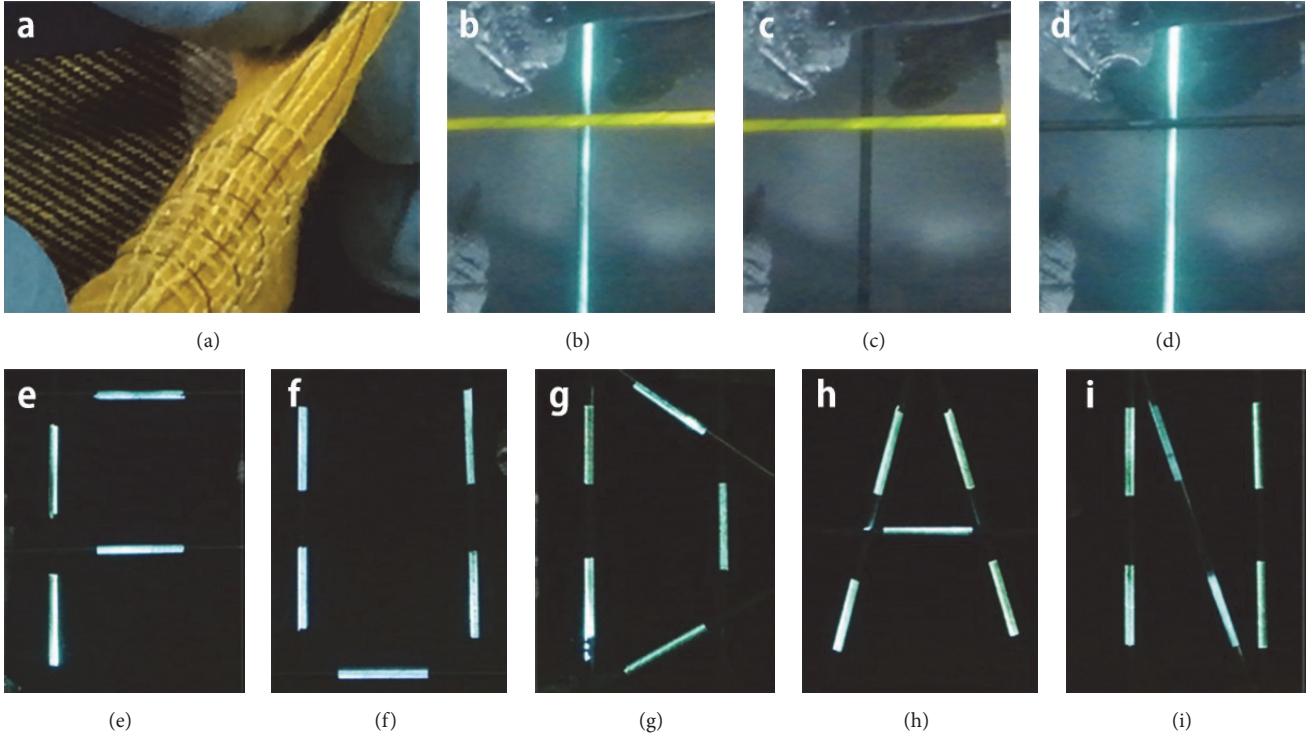


FIGURE 1: Integrated PLEC fibers and textiles: (a) textile subjected to bending and twisting, (b)–(d) two fiber-shaped PLECs selectively illuminated in different colors (biased at 10 Volts), and (e)–(i) weaving of fiber-shaped PLECs into a “FUDAN” pattern (biased at 9 Volts).

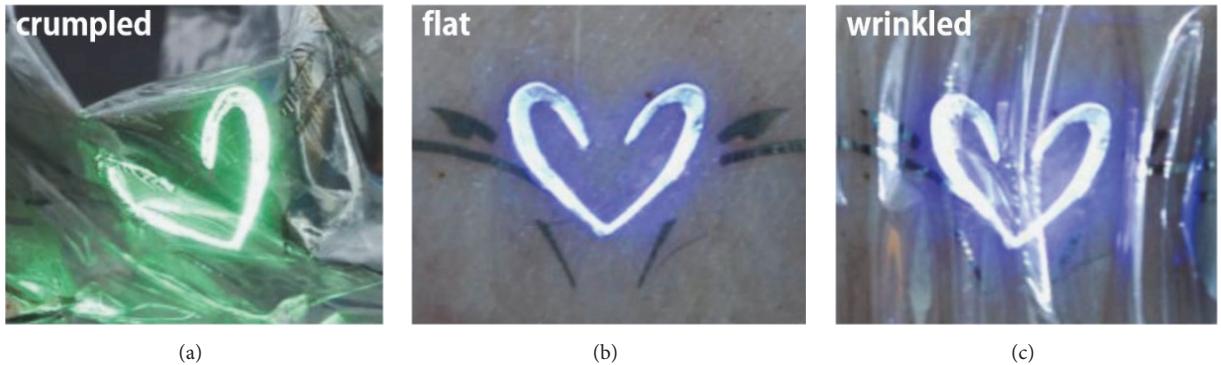


FIGURE 2: Electronic tattoo demonstrations based on ultrathin wearable QLEDs: (a) optical image of ultrathin green QLEDs laminated on crumpled Al foil and (b, c) photographs of the electronic tattoo (blue QLEDs) laminated on human skin (b). The wearable QLEDs maintain their optoelectronic performances even under skin deformations (c).



FIGURE 3: Lumalive textile garments with display functionality.

“Shirt” software includes sensors that detect the important sounds of the bass, cello, and percussion instruments and translate them into vibrations, enabling users to experience the orchestral performance.

Figure 5 presents a garment with a safety protection function [13]. LED-based turn signals are installed in the front and back of the garment, allowing the wearer to view approaching people from all directions and thereby avoid rear accidents.

Inspired by the preceding examples, our interdisciplinary research team analyzed the content and applicability of LED light-based wearable interactions. From a technological



FIGURE 4: A Sound Shirt for hearing impaired individuals.



FIGURE 5: LUMENUS smart cloth with a safety functionality.

viewpoint, conventional LED-based wearable interactions can be divided into three types (see Figure 6). The first is a simple lighting for aesthetic effects, and the second is a light emission by responding sensors embedded in the wearable device. Finally, the third is the light emission indicating the operating status of a separate device such as a smartphone. As shown in Figure 6, the conventional LEDs play a comparatively passive role that depends on the types or applications of the prescribed wearable devices.

Along with the Lumalive textile garments in Figure 3, which possess an infotainment functionality, the typical example of type 1 interaction in Figure 6 is CuteCircuit's Galaxy Dress, which is pictured in Figure 7 [14]. Galaxy Dress is composed of 24,000 full-color pixels and a LED of $2 \times 2 [\text{mm}^2]$ size is flat like paper. Also, the Galaxy Dress, as an integration of the electronic device technology, does not overheat and consumes very little electricity. The heaviest part of Galaxy Dress is not the technology, but the 40-layer pleated silk organza crinoline that widens the skirt [14]. In recognition of the merits of fusing technology and fashion, the Galaxy Dress is the centerpiece of the "Fast Forward: Inventing the Future" exhibition at the Museum of Science and Industry in Chicago [14]. However, the conceptual design is limited, as it does not consider the users' situations or purposes when wearable devices are gradually evolving into life-friendly products.

Next, a representative example of type 2 interaction shown in Figure 6 is the "Sound Shirt," shown in Figure 4, demonstrating the visual communication function. The luminescence pattern and color of the LED dynamically respond to the loudness sensor. However, the application range and usability of this interaction are limited by sensor type and the surrounding environment. Nevertheless, this sensor-based interaction caters to the situations and purposes of hearing impaired persons and hence expands human functionality.

The third interaction type in Figure 6 visually informs the status of an external device such as a smartphone. A representative example is the smart ring of RINGLY, which indirectly receives notifications of five smartphone applications selected by the user (see Figure 8). The user-selected applications are distinguished by five emission colors of the LED attached to the accessory and four vibration patterns [15]. Xiaomi's smart band Mi Band also reports the status of the user's smartphone and checks the user's heart rate, distance traveled (number of steps), and sleeping patterns. To remove the functional ambiguity of the three LED indicators embedded in the initial version [16], Xiaomi recently developed Mi Band 2, which replaces the LED indicators with small organic LED displays [17].

Considering the three kinds of generic wearable interactions, present LEDs are potentially useful in smart fashion design despite their limited applicability in smart materials. Given that most wearable devices include a sensor and a battery, LED light can visually and intuitively indicate the operating state of a device, which is a very attractive feature. Nevertheless, as LED lacks many of the functional factors demanded for smart fashion materials, it is generally avoided in costume design. Specifically, LEDs are limited to visual effects (light emission patterns and colors), which restricts their applicability in wearable device of everyday life. Therefore, to utilize LEDs in the smart fashion market, which is gradually becoming specialized and systematized, we require aesthetically pleasing, functional wearable interactions that adapt to the situations and purposes of users.

5. Technical Implementation: Color-Independent Visual-MIMO System

To incorporate visible-light communication in wearable devices and smart fashion, our interdisciplinary research team applied the color-space based generalized color modulation (GCM) technique proposed by ourselves [18]. The research adopted the color-independent visual-MIMO (Multiple-Input Multiple-Output) system. Here, the modifier "*color-independent*" indicates the independence of the variations in the light color and light intensity and *visual-MIMO* means the visible light communication (VLC) between the light emitting array (LEA) and the camera.

5.1. Visible Light Communication (VLC). Generally, VLC refers to the technology for sending information by blinking the LED at a speed that cannot be perceived by human eyes. Here, the LED is used for transmitting the information, while a photo detector (PD) or a camera image sensor can be utilized as the receiver (depending on the application). In

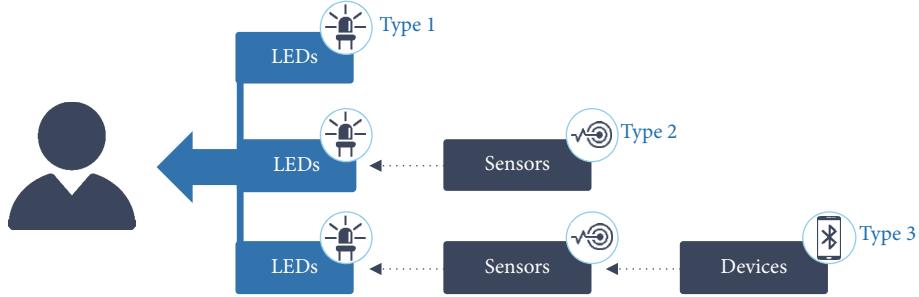


FIGURE 6: Three types of conventional wearable interactions based on LED light.



FIGURE 7: Galaxy Dress.



FIGURE 8: Smart ring of RINGLY.

the simplest case, digital data is sent through the LED by assigning the “on” and “off” of the LED light to the digital bits 1 and 0, respectively. The receiver decides the 1 or 0 status of the signal by detecting the changes in light intensity from the sensor. Recently, many studies about a camera image sensor as the transmission receiver have been conducted because most electronic devices (including smartphones) already contain a high-performance camera [2].

The low cost, energy efficiency, and environmental friendliness of LED elements and their extensive use of the existing infrastructure are advantageous for applying VLC in LED-based fashion. By adding a communication module and an appropriate smartphone application that controls the LED built-in smart fashion, any designer can incorporate LEDs into new interactions that are easily experienced by users. As in usual short-range wireless communication, the user can actively select and control the information to be delivered. Figure 9 demonstrates how VLC enables new wearable interactions in LED-based smart fashion. While maintaining its conventional functions, the LED can communicate with external devices through the VLC. The potential for new interactions will ultimately increase the value of LED-based design materials. Three types of VLC communication are discussed here.

Type 1: LEDs are attached to cloths and send some fixed information to Device 1 to Device N.

Type 2: LEDs can be attached to the locket. Sensors collect data and send the data to the device via LEDs.

Type 3: Almost similar to type 2 but here sensors receive data from another device via Bluetooth and LEDs are attached to the wearable device like watch.

In case of every type, multidevice interaction is possible, that is, one of the main advantages of the VLC communication.

5.2. Generalized Color Modulation (GCM). As a practical medium for transmitting data by VLC, the LED light must be carefully considered at the fashion design side because of its visible attributes. Most of the existing VLC systems have been developed for white lighting, which is unsuitable for fashionable clothing. The RGB LEDs in typical wearable devices (or smart fashions) are controlled by two parameters: color and intensity (brightness). In addition, as both parameters largely depend on the design requirements, the communication must perform under varying light color and brightness of the LEDs. This section introduces a modulation method that is optimized for combining the VLC technology with fashion design.

The first color-space-based modulation scheme, termed color shift keying (CSK), was proposed by the IEEE 802.15.7 task group [19, 20]. However, CSK is not suitable for communications when target color varies with time because

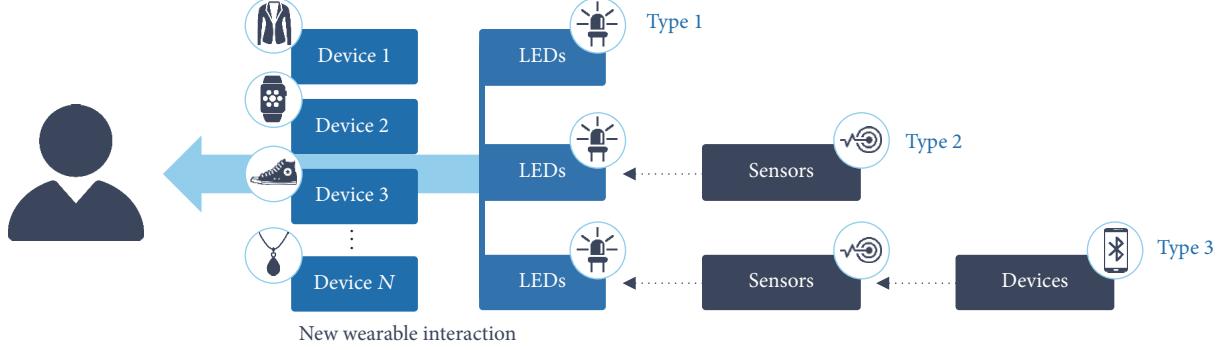


FIGURE 9: Block diagram of new wearable interactions through VLC.

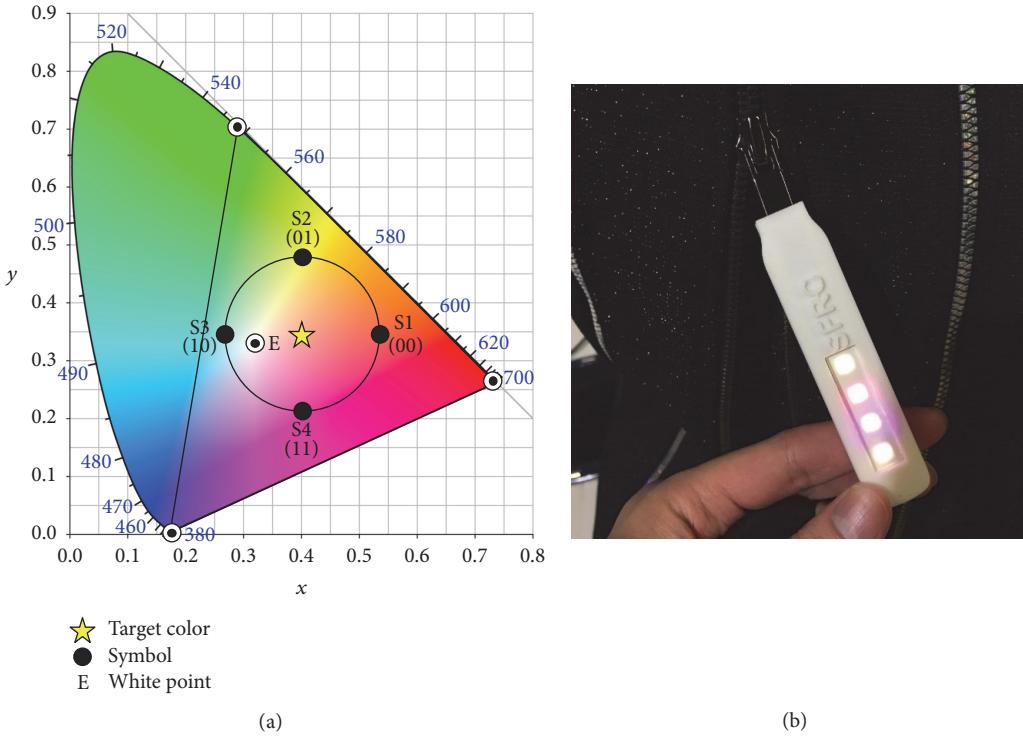


FIGURE 10: Example of GCM constellation and a prototype of color-independent VLC system: (a) example of GCM constellation and (b) prototype of color-independent VLC system.

it does not involve a systematic manner of coping with changing target color conditions. Here, the *target color* means the desired color of the LED lighting. In order to overcome this limitation, another color-space-based modulation scheme, termed as generalized color modulation (GCM), was proposed for color-independent VLC systems [18, 21]. It should be noted that the most distinctive feature and advantage of GCM over the other modulation schemes is color independency. GCM is able to generate any color within a gamut by combining some of the wavelengths or colors. Therefore, through GCM, we may achieve a VLC scheme that can maintain the original color and brightness while performing seamless communication. Figure 10 shows the example of color-space-based constellation and a prototype of the color-independent VLC transceiver. GCM constructs

a constellation diagram in a light color space to represent data symbols. Each constellation point in the color space represents a corresponding color. Symbol color indicates the illuminating color of the LED depending on the input data, and the *target color* means a color perceived by the human eye when the symbol color changes rapidly at random [22]. Our proposed GCM method selects the number of transmitting symbols and the corresponding colors of the symbols in CIE1931 color space (see Figure 10(a)) and assigns digital bits to each color. Typically, when many symbols are illuminating randomly and rapidly through the LEDs, the human eye perceives the color of the LED as the average of multiple symbol colors.

The LED-color independency of GCM may be its most indispensable feature in fashion design. Designers of smart

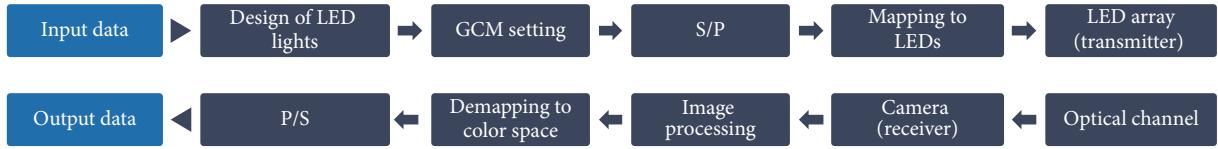


FIGURE 11: GCM-based visual-MIMO transceiving procedure with design of LED lights.

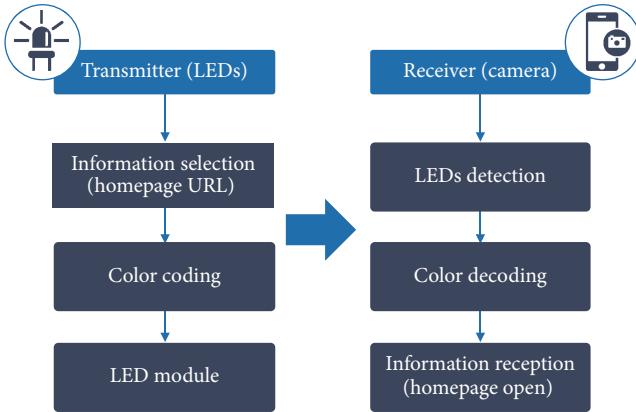


FIGURE 12: Block diagram of the transmitter (LEDs) and receiver (camera) in the “website-link” prototype.

wearable fashion must determine the color and brightness of the LEDs embedded in their products. The GCM method enables the engineer to maintain the color and brightness selected by the designer while performing seamless communication. Figure 11 shows a block diagram of the GCM-based visual-MIMO transceiving procedure, including the design of LED lights [23].

6. System Description: Transmitter (LEDs) and Receiver (Camera)

In the visual-MIMO system, the multiple transmitting elements of a light emitting array (LEA) are used as transmitters to communicate to the pixels of the camera (image sensor), which act as multiple receiving elements. This section explains how the transmitter and receiver of the color-independent visual-MIMO system interact through the “website links” prototype developed by our interdisciplinary research team. Figure 12 shows the overall operation of the “website link” prototype based on color-independent visual-MIMO.

The operation of each block in Figure 12 is outlined in the following.

6.1. Transmitter

- (1) Information selection: first, the user selects the data to transmit over the Bluetooth platform of a smartphone. In the prototype, the color independence of the visual-MIMO system is demonstrated on the relatively simple URL address of an internet homepage.

In actual use, the data type can be altered to suit the content of smart fashion.

- (2) Color coding: the selected data (in this case, the URL address) are converted into color information by the GCM method, which digitizes the data to fit the communication protocol of the color-independent visual-MIMO system.
- (3) LED module: the converted color information is sequentially mapped to multiple LEDs for fast LED emission.

6.2. Receiver

- (1) LED array detection: referring to the image captured by the camera of a smartphone application, the LED array is detected by an image processing algorithm [24].
- (2) Color decoding: each LED is located by analyzing the size and geometry of the LED array in the detected image. The symbol information is then obtained by determining and analyzing the color of each LED.
- (3) Information reception: the data are recovered by converting the color (or symbol) information into digital bits through our GCM method. Finally, the smartphone application opens the homepage using the recovered data.

7. Design Strategies and Results

In the design industry of the 21st century, LEDs, backed by technological advancements, can realize multiple-sensory emotions that express rich human emotions. The expressive effects are functionality, aesthetics, interactions, and amusement.

More specifically, the various colors and fluid graphics patterns of LED work as aesthetic elements. Along with the unique functionality of light, these elements create brand images and attract consumers to the design industry. The long life and environmentally friendly nature of LEDs enable efficient, low-cost designs. Combined with digital technologies, the new media features of LEDs can entertain users with designs involving visual and emotional interactions.

Our research team fused engineering and fashion design to actively capture the emotions of consumers in our design products. To achieve a smart fashion prototype based on VLC, our team developed modular strapped cuffs and zipper sliders with the physical properties of clothing, which can be detached as desired by users. The technological aspects of the design manifest in the changing functionality of

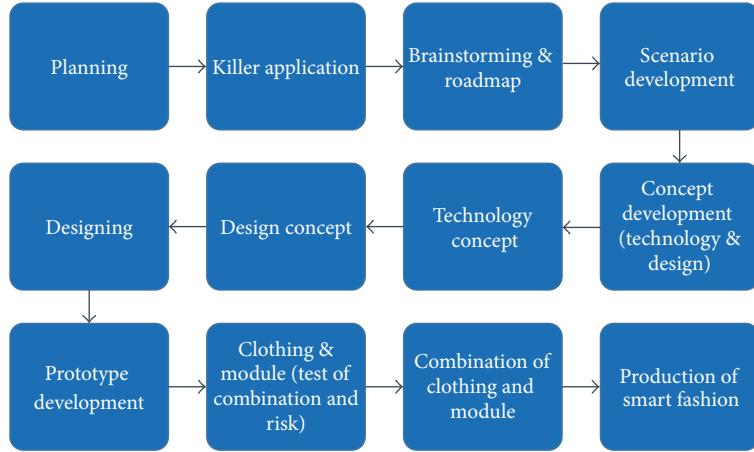


FIGURE 13: A process model of VLC-based smart fashion design.

the visual-MIMO communication (URL information in our scenario), which is implemented on the Bluetooth platform. In this design, users can adopt their own styles and individualities through the detachable modular approach.

7.1. Iterative Collaborative Design Process. The process model of the VLC-based smart fashion design fuses the processes with the desired characteristics developed in other fields. In developing the process model, we analyzed the methodologies of digital content, product design, and interaction design and incorporated them into the fashion process.

In the following we summarize the design process model by which we developed our stepwise strategy for the VLC-based smart fashion prototypes (see Figure 13):

- (1) In the Planning and Killer Application Development stage, the content is tailored to users by exploiting the VLC and the correlations among men, clothes, and technology.
 - (2) The Brainstorming and Roadmap stage analyzes the design needs of the user and identifies the specialized VLC factors through the technology. Finally, it designs a killer application.
 - (3) The Scenario Development stage determines the scope of the VLC-based scenarios and VLC applied services. It writes a composite story by predicting the actual situations, including the circumstances and goals of users, and facilitates solid judgments rather than a two-dimensional, fragmentary understanding of problems.
 - (4) The Concept Development (technology and design) stage identifies the hardware, software, and formative elements of the design. The design concept is based on *who, when, what, how, and why*, context and content, and user interface and user experience. During this critical stage, the elements stemmed from the problems of technology and design-centric thinking are improved by a practical and specific implementation plan, which is applied in actual research. This stage must develop the user-centric content and investigate

the fusion of engineering and design to meet the demands of users (namely, formativeness, wearability, purposefulness, economy, usability, satisfaction, and safety).

- (5) The Designing stage applies and materializes the formative design elements. This stage involves styling, sketch design, the design of choices for sample production, drawing, modeling, and fitting. In the present design, the strapped cuffs and zipper sliders were rendered detachable from the VLC system if desired by the user. Users can also reorganize various platforms containing the clothing components of the smart fashion module. The outcome was a playful design by which users can enjoy the technology and its operation while satisfying their emotional needs.

(6) In the Prototype Development stage, our team developed a prototype of the strapped cuffs and zipper slider types with inbuilt information (URL information in the current scenario) conveyed through a Bluetooth platform. The modules were combined with clothing at this stage, and the prototype was demonstrated.

7.2. Multipurpose Smart Fashion by User Interfaces. If clothes-module combination has superior functionality but is fixed and nondetachable, its smart fashion will be of limited utility and will fail to provide a user-centric design. To resolve this issue, the components of our modular system can be detached as desired by users and are compatible with the details (strapped cuffs and zipper sliders) and fastening systems of conventional clothes. The modular system vastly improves the multipurposefulness, economy, usability, and detergency of the earlier smart fashion lacked. The approach began with the modularization of simple form to expand the diversity without requiring software and hardware changes. Individual user experience-based designs can be implemented by building diverse platforms that are modularized for each user. By simulating the fastening systems of clothes (such as snaps, zippers, Velcro, and clips), we overcome the reluctance of users to wear computerized technologies. These fastenings are

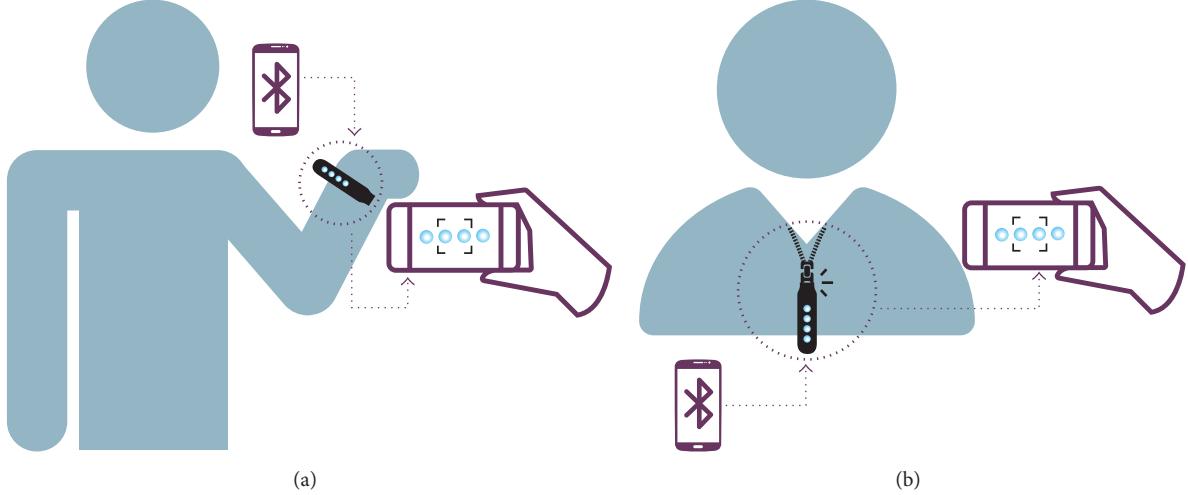


FIGURE 14: Communication between the transmitter (LEDs) and receiver (camera): (a) strapped cuffs type and (b) zipper slider type.

also familiar to users and are both functional and decorative. The developed fastening system is intuitively operated and adaptable to users' individual tastes.

The strapped cuffs and zipper sliders, proposed in our study, represent a multipurpose, adaptable smart fashion. Through the Bluetooth platform, users can alter the functions and designs of the smart fashion and continuously communicate with other users. The detachable modular system incorporates the functional module into clothing components that can be separated by fasteners (see Figure 14).

7.3. Interactive Smart Fashion for VLC. In the proposed scenario, two strapped cuffs and zipper slider types were applied to an interactive smart fashion, which nonverbally communicates information through VLC. This communication mode visually conveys the users' individuality.

Interaction (i) Scenario

- (1) A user takes a walk at night.
- (2) The walker meets an acquaintance by chance.
- (3) The acquaintance requests the walker's blog address.
- (4) With the walker's permission, the acquaintance (using the camera of a smartphone) photographs the walker's strapped cuffs (or zipper slider) embedding the visual-MIMO system.
- (5) The blog address is automatically sent to the smartphone.

Interaction (ii) Scenario

- (1) A user shops for clothes at the department store.
- (2) The shopper desires to know the size, material, color, and other details of a clothing item.
- (3) The shopper opens his 3D body avatar using the smartphone app.

- (4) The shopper photographs the tag shaped visual-MIMO with his smartphone camera.
- (5) The visual-MIMO system simulates the clothing item. The shopper receives the feedback without wearing the item.

(i) Strapped Cuffs Type. The cuffs band-type smart fashion module can be attached and detached through the Bluetooth platform developed by the research team. This fashionably smart module serves multiple purposes and functionalities by attaching and detaching the VLC-based strapped cuffs with a snap fastener, which is compatible with other clothes, or by attaching it to a zipper slider.

(ii) Zipper Slider Type. This module inserts a zipper slider with a clip fastener and can be detached according to users' tastes. The functionality and LED light source of this prototype are visually decorative, and the module alters the details of basic clothing items.

The research combined the elements of a smart fashion interface with shaped fastening systems that can be attached to clothes, alter the functions of clothes, and allow adjustment to different clothes (see Figure 15).

Design guidelines were incorporated into the technology and design concept, and the application planning of the elements was based on the analysis results of the user-needs survey. By this approach, the research team developed an interactive smart fashion prototype for the visual-MIMO communication. The prototype consisted of strapped cuffs and zipper slider type modules. To ensure a user-friendly product, the design guidelines were based on the four essential elements of a technology concept (usability, detachability, purposefulness, and economy) and of a design concept (usability, detachability, formative ness, and wearability).

Table 1 categorizes the technology and design concepts in the design guidelines of the interactive smart fashion for nonverbal communication through the visual-MIMO system.

TABLE 1: The design guidelines of interactive smart fashion for nonverbal communication through visual-MIMO communication.

Filed	Element	User needs	General plan
Technology concept	Usability	(i) Free of two hands, intuitional interface form needs. (ii) Communication: sight (LED), sense of touch (vibration sensor) (iii) Interactivity: using and care for simple, easy communication	(i) Transmitter (LEDs) and receiver (Camera) communication (Bluetooth Android Platform)
	Disadhesion	(i) Demand on disadhesion module in clothes (ii) Convenience through disadhesion (iii) Multipurpose and compatibility	(i) Disadhesion of module
	Purposefulness	(i) Taking into account situation and purpose to user	(i) Design disadhesion on user taste
	Economic feasibility	(i) Compatibility with other smart clothing. Take into account Economic feasibility	(i) Design of disadhesion a module
	Usability	(i) A decorative painting of technology (ii) Feeling in use of not computer instrument	(i) A source of light LED decorative painting of technology
	Disadhesion	(i) Fashionable needs to disadhesion and compatibility of module	(i) Disadhesion realization through fastener (ii) Strapped cuffs and zipper slider types
	Formativeness	(i) Everyday wearing positive, feeling in use of not computer instrument	(i) Details change on basic item
	Wearability	(i) Structural design of check body transformation	(i) Body move a little design on arm

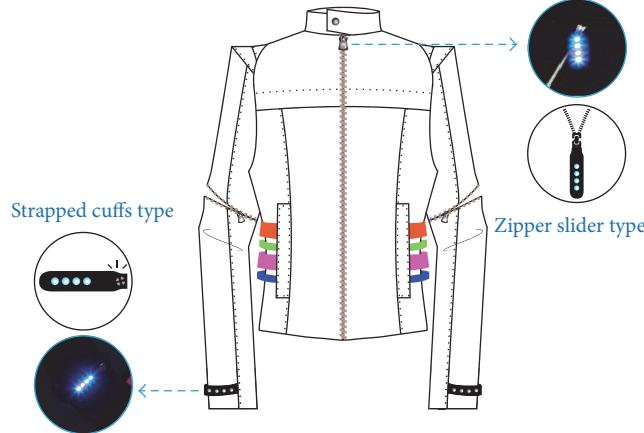


FIGURE 15: Dis-adhesion of smart module by a fastener: (left) strapped cuffs type and (right) zipper slider type.

Figure 16 presents the strapped cuffs and zipper slider prototypes of the interactive smart fashion for visual-MIMO communication.

Reference [25] shows the implementation of our proposed idea. LED can be used as wearable device as well as pattern for communication. In Figure 17, to verify the performance of the prototype, symbol error rate (SER) was analyzed for 100,000 random symbol data while varying the distance between the transmitter and the receiver; the distances used were 0.5, 1, 1.5, and 2 m, respectively. In the 4×4 LEA, there was no error for distances up to 1.5 m, but 12200 errors occurred at 2 m. For the 4×1 LEA, the number of errors was 700, 520, 2100, and 64400, respectively. In the case of 4×1 LEA, since the standard deviation (SD) value of the synchronization data increases, the difference from the SD value of the information data decreases, so the synchronization error occurs more frequently and the performance degrades [23, 25].

8. Conclusion

Interactive relationships in VLC-based smart fashion are formed between the individuals wearing LED-applied clothes and their observers, or among a group wearing LED-applied clothes. VLC-based smart fashion is organic, promoting interactive information flow through the wearable interactions and expanding the expressional forms of the wearable technology.

The proposed modular system can be detached as desired by the users. The system captures the details of real clothes (strapped cuffs and zipper sliders) and fastening systems by fusing technology with clothing design and is compatible with other clothes. The visual-MIMO system embedded in the clothing components is designed for versatile user tastes and allows users to operate from various platforms. The outcome was a playful design that satisfies the emotional needs of users through an enjoyable technology.



FIGURE 16: Prototypes of (left) strapped cuffs type and (right) zipper slider type.

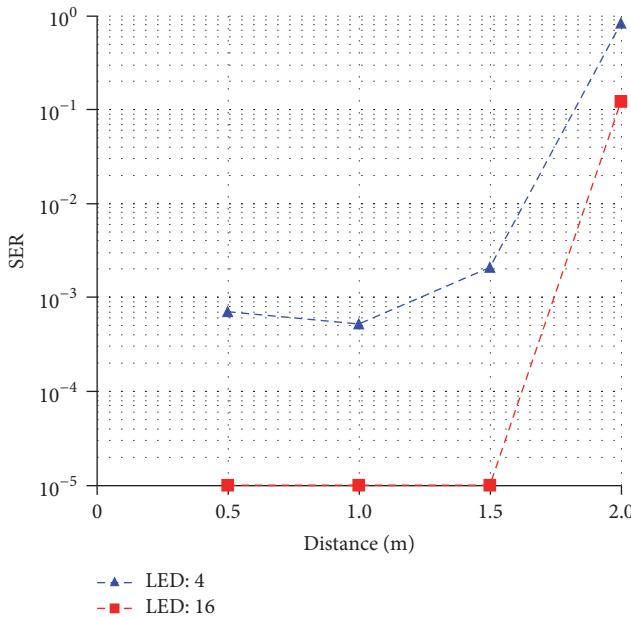


FIGURE 17: SER performance versus distance.

The system preserves the aesthetics of old wearable devices and smart fashion by preventing the flickering of LED color and brightness. In addition, an image processing algorithm at the receiver end was optimized for wearable devices and smart fashion, and an interactive conversation system that delivers information nonverbally via LED light was developed.

The proposed color-independent visual-MIMO system communicates only through LEDs and camera image sensors. As no separate communication network is required, the system is potentially compatible with location-based service (LBS) and internet of things (IoT) services with the old LED

infrastructure and smartphone cameras. Future technology is expected to further satisfy the goals of users through customized content development. The design characteristics of LED light (color and brightness) are visual and emotional elements. VLC-based technology will ultimately realize creative and original high value-added design products, giving the smart fashion industry a competitive edge in global markets.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Education [2015R1D1A1A01061396] and was also supported by the National Research Foundation of Korea Grant funded by the Ministry of Science, ICT, Future Planning [2015R1A5A7037615].

References

- [1] K. Nordby, "Conceptual designing and technology: Short-range RFID as design material," *International Journal of Design*, vol. 4, no. 1, pp. 29–44, 2010.
- [2] G. Corbellini, K. Akşit, S. Schmid, S. Mangold, and T. R. Gross, "Connecting networks of toys and smartphones with visible light communication," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 72–78, 2014.
- [3] Wearable communication system. US 6853293 B2, 2005.
- [4] C. Hertleer, H. Rogier, L. Vallozzi, and L. Van Langenhove, "A textile antenna for off-body communication integrated

- into protective clothing for firefighters," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 4, pp. 919–925, 2009.
- [5] E. Ito, M. Matsumoto, K. Tsuda, and M. Honjo, "Wearable jacket having communication function, and endoscope system employing wearable jacket," US 7109933 B2, 2006.
- [6] A. Pyattaev, K. Johnsson, S. Andreev, and Y. Koucheryavy, "Communication challenges in high-density deployments of wearable wireless devices," *IEEE Wireless Communications Magazine*, vol. 22, no. 1, pp. 12–18, 2015.
- [7] P. H. Pathak, X. Feng, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: a survey, potential and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2047–2077, 2015.
- [8] M. Stoppa and A. Chiolero, "Wearable electronics and smart textiles: a critical review," *Sensors*, vol. 14, no. 7, pp. 11957–11992, 2014.
- [9] Z. Zhang, K. Guo, Y. Li et al., "A colour-tunable, weavable fibre-shaped polymer light-emitting electrochemical cell," *Nature Photonics*, vol. 9, no. 4, pp. 233–238, 2015.
- [10] M. K. Choi, J. Yang, K. Kang et al., "Wearable red-green-blue quantum dot light-emitting diode array using high-resolution intaglio transfer printing," *Nature Communications*, vol. 6, article no. 7149, 2015.
- [11] PHILIPS. (December 10, 2016). <http://www.lighting.philips.com/main/products/luminous-textile/>.
- [12] JIMDO. (December 10, 2016). <http://soundshirt.jimdo.com/english-1/>.
- [13] LUMENUS. (December 10, 2016). <http://www.lumenus.com/>.
- [14] CUTECIRCUIT. (November 3, 2016). <https://cutecircuit.com/thegalaxy-dress/>.
- [15] RINGLY. (November 3, 2016). <https://ringly.com/>.
- [16] Xiaomi, (n.d.). (November 3, 2016) URL <http://www.mi.com/en/miband/>.
- [17] Xiaomi, (n.d.). (November 3, 2016). <http://www.mi.com/en/miband2/>.
- [18] P. Das, B.-Y. Kim, Y. Park, and K.-D. Kim, "Color-independent VLC based on a color space without sending target color information," *Optics Communications*, vol. 286, no. 1, pp. 69–73, 2013.
- [19] IEEE 802.15.7-2011, "Short-Range Wireless Optical Communication Using Visible Light", 2011.
- [20] A. Yokoi, "Color multiplex coding for VLC," *15 Working Group for Wireless Personal Area Networks*, 2008.
- [21] P. Das, B.-Y. Kim, Y. Park, and K.-D. Kim, "A new color space based constellation diagram and modulation scheme for color independent VLC," *Advances in Electrical and Computer Engineering*, vol. 12, no. 4, pp. 11–18, 2012.
- [22] T. H. Kwon, J. E. Kim, and K. D. Kim, "Synchronization method for color-independent visual-MIMO communication," in *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE)*, pp. 458-459, Las Vegas, NV, USA, January 2017.
- [23] J. E. Kim, J. W. Kim, Y. Park, and K. D. Kim, "Color-space-based visual-mimo for v2x communication," *Sensors*, vol. 16, no. 4, p. 591, 2016.
- [24] J.-E. Kim, J.-W. Kim, and K.-D. Kim, "LEA detection and tracking method for color-independent visual-mimo," *Sensors*, vol. 16, no. 7, p. 12, 2016.
- [25] T. H. Kwon, J. E. Kim, J. H. Oh, and K. D. Kim, "Implementation of color-space-based wearable visual-MIMO system," in *Proceedings of the Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 776–779, Milan, Italy, July 2017.

Research Article

Relay Positioning Strategy for Traffic Data Collection of Multiple Unmanned Aerial Vehicles Using Hybrid Optimization Systems: A FANET-Based Case Study

José Jailton, Tassio Carvalho, Jasmine Araújo, and Renato Francês

Institute of Technology, Federal University of Pará (UFPA), Belém, PA, Brazil

Correspondence should be addressed to José Jailton; josejailtonjunior@yahoo.com.br

Received 11 August 2017; Revised 23 October 2017; Accepted 2 November 2017; Published 7 December 2017

Academic Editor: Gabriel-Miro Muntean

Copyright © 2017 José Jailton et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the new context of Next Generation Networks, Aerial Ad Hoc Networks, also known as FANET (Flying Ad Hoc Network), are being used to monitor areas of difficult access. Owing to the dynamism and autoconfiguration of this type of network, a strategy is needed to position its devices (drones) to ensure it can achieve good performance. In light of this, this paper proposes a flight path planning model, which involves a metaheuristic optimization-based approach. The proposal relied on the artificial neural networks to optimize the positioning of the relay device, so that the throughput between the other devices could be increased; the benefits of the proposal were demonstrated through simulations.

1. Introduction

In recent years, as a result of technological advances in areas such as robotics, telecommunications, and computer networks, Unmanned Aerial Vehicles (UAVs) have emerged as an alternative means of providing diverse applications in both civilian and military areas. These UAVs aim to improve or create a network infrastructure in places that are difficult to access, such as areas of natural disasters or enemy territory. In this area, an Air Ad Hoc Network has emerged as an acceptable solution since it enables information to be collected in a more flexible, fast, and reliable way [1].

In [2], the authors state that the main challenge of FANET is to perform cooperative detection using multiple UAVs to cover an area that cannot be covered by a single UAV. In this way, it can create an overhead network in which its devices (drones) communicate by transmitting information to each other. Thus, it is necessary to have reliable and stable communication between the devices to maintain good levels of QoS (Quality of Service) and QoE (Quality of Experience).

In most FANET applications, the objective is to collect data from the environment and relay them to a ground base station. The topology of a FANET network consists

of several UAVs (Unmanned Aerial Vehicles) and a UAV relay, which is responsible for receiving information from the other UAVs and transmitting it to a base station on the ground. Establishing the position of the UAV relay in relation to the other UAVs is of paramount importance to ensure the satisfactory performance of the network since it must be accessible directly or indirectly (i.e., communication via intermediary devices) to all the other elements of the network. This has led to an increase in the number of studies on UAV relay positioning in the academic community.

This paper proposes a model that can establish the best real time positioning of a UAV relay among multiple UAVs used for flight missions in scenarios such as those shown in Figure 1. The objective is to fly the drone in real time to its optimal placement based on a knowledge of the position required for flight missions or the location of the optimal placement. Thus, the proposal aims to conduct the relay in real time to an “ideal” location (where a good level of communication can be maintained with all the other devices). This positioning will be determined by the location of the other multiple UAVs of the network and the proposal outlined in this article will calculate/optimize the new UAV relay location.

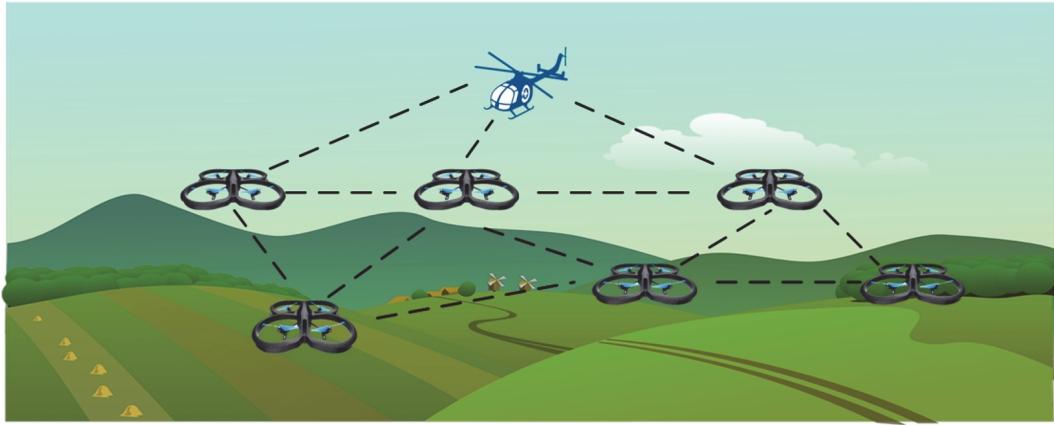


FIGURE 1: FANET scenario.

The proposal is divided into two key phases, which involve activating the mobility of the UAVs through the Network Simulator (NS-2) and using the mobility data of the UAVs in an Artificial Neural Network (AAN) in MatLab (the position and traffic of the UAVs with the relay were used for the training of the neural network), which will indicate the position of the UAV relay.

This article is structured as follows: Section 2 discusses reviews from related works; Section 3 describes the proposed optimizer in detail; the results obtained from the simulation are shown in Section 4; and Section 5 summarizes the conclusions and makes suggestions for future work.

2. Related Works

This section discusses the related work and examines the Aerial Ad Hoc Networks (FANETs). In addition, it calculates the positioning of the UAV relay with the aim of improving network performance.

In [4], UAVs are used for locating forest fires. Their approach leads to an improvement in the localization mission through a decision-making strategy, resulting from a temperature-based probabilistic model that is designed to estimate the distance to the forest fire. The UAV adjusts its trajectory according to the state of the knowledge of the fire, using a map to represent its knowledge, which is updated at each stage of its exploration.

In [5], an accurate analytical model is set out to reduce signal interference between the UAVs and the base station. The algorithm performs uniform frequency hopping within the bandwidth channel itself and has achieved service quality gains in overload or “interrupt cells.” However, in the article, there is no applied artificial intelligence technique that can make the task of optimization more efficient and faster, and in most situations, only a new relay UAV positioning would be able to achieve gains.

Reference [6] suggests forming a set of Aerial Ad Hoc Networks, including several backbones of FANETs that communicate with each other and also communicate with a base station. This composition of several backbones increases the coverage area of the monitored region, but there has been no

study on the proper way to position the devices to ensure that the network can perform in the best possible way.

The article of [7] includes a control scheme for the movement of the UAVs within a particular coverage area. The motion control scheme is based on the distance between the UAVs and their signal strength and allows them to have good communication, thus having good transmission rates. However, controlling the movement is not the most viable solution because certain areas of the coverage area are not sufficiently monitored.

The authors in [8] aim to minimize the number of MBSs needed to provide wireless coverage for a group of distributed ground terminals (GTs), ensuring that each GT is within the communication range of at least one MBS. A polynomial time algorithm with successive MBS placement is used, where the MBSs are placed sequentially starting on the area perimeter of the uncovered GTs along a spiral path toward the center, until all GTs are covered.

An efficient algorithm is necessary to maneuver each UAV so that the whole system can produce complex, adaptable, and flexible team behavior. The task-planning problem for UAV networks with connectivity constraint involves a number of parameters and interaction of dynamic variables. An algorithm has been proposed for such a situation [9]. There is also an algorithm for distributed intelligent agent systems in which agents autonomously coordinate, cooperate, negotiate, make decisions, and take actions to meet the objectives of a particular task. The connectivity constrained problem is NP-hard and a polynomial time heuristic has been proposed in the literature [10]. Reference [10] is the closest related work by also having proposed a heuristic algorithm, but without any evaluation of Quality of Service and Quality of Experience.

The articles cited in this section have led to several improvements being made within the context of FANET applications. However, no techniques were employed that aimed at maintaining high performance in the data transfer, which is achieved through the positioning of the UAVs. None of the papers cited used hybrid networks to manage the decision-making required for the best positioning between the mission-based UAVs and UAV relay. In light of this,

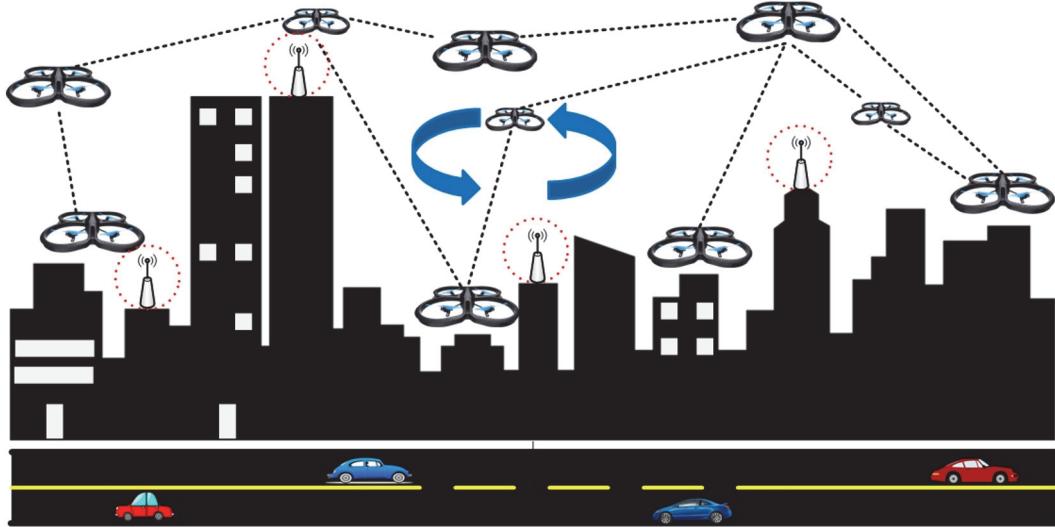


FIGURE 2: Urban FANET application.

this work employs techniques that allow the use of these networks both in ad hoc applications and in other wireless technologies.

3. FANET Application Features

The main feature of the Ad Hoc Networks, referred to by the IETF (Internet Engineering Task Force) as MANET (Mobile Ad Hoc Network), is that they do not have an infrastructure, and as a result, all network's functions must be carried out by the nodes themselves. This means that the nodes that make up an Ad Hoc Network must be able to communicate with each other as routers.

Ad hoc Networks are generally used in scenarios where there is a need to quickly deploy a network and usually where there is no proper infrastructure. Nodes can move arbitrarily within them and change the network topology in an unpredictable way, which requires the constant adaptation and reconfiguration of routes so that the nodes can still communicate with each other.

In the new context of Next Generation Networks (NGNs), a derivation of Ad Hoc Networks called FANET (Flying Ad Hoc Network) arises. FANETs are Ad Hoc Networks consisting of remotely controlled flying elements (UAVs: Unmanned Aerial Vehicles) that communicate with each other [11], and because FANETs have flexibility, versatility, and even ease of operability, they are used for both military and civilian applications, such as border control in risk zones, deforestation monitoring, and traffic flow control (Figure 2).

In FANET networks, the mobility index is much higher than in a traditional Ad Hoc Network; thus changes in topology are more frequent and even the coverage area is larger (depending on the type of hardware and technology used). That is why FANET networks have to be self-configuring and self-organizing; a FANET network must be prepared for sudden changes in its topology, in organization, and even in its communications systems.

The mobility of mobile units and their spatial arrangement are also very important in determining the communication routes. As a result of the movement, these routes are usually replanned to allow continuous interconnection of the nodes. For this reason, routing must be carried out dynamically and the activities involved in the routing protocol must be effective and simple, so as to give greater autonomy to the nodes and reduce the delay in the delivery of data from a sender node to a destination node.

The communication between the UAVs largely depends on their location, since all the information collected from the environment is concentrated in a UAV relay node (which is responsible for relaying the data to a control center). Hence, its positioning in relation to the other nodes must be strategic enough to ensure that the network maintains a good performance. It is not an ideal situation if the UAV relay has an excellent communication with some UAVs while at the same time having poor communications with the other UAVs in the network [12].

The UAVs are responsible for flying over the environment; they have sensors to collect information and a higher energy load than traditional sensor networks that means they can communicate with each other more easily by finding fewer obstacles in their line of sight. This should reduce the number of nodes needed to cover a given area, although climatic conditions can impair communication (wind, rain, etc.) [13].

It should be noted that the computational power of FANETs networks means that they have a greater capacity to transmit information, since in many cases they are responsible for transmitting information in real time (with videos from the monitored environment). There have been no specific routing protocols for FANETs so far, and the traditional protocols used in Ad Hoc Networks (e.g., AODV, OLSR) are also used in FANET.

Owing to the high mobility rate of UAVs in a FANET, updating the location of all the nodes in the network is a critical feature. The devices in the network must know the

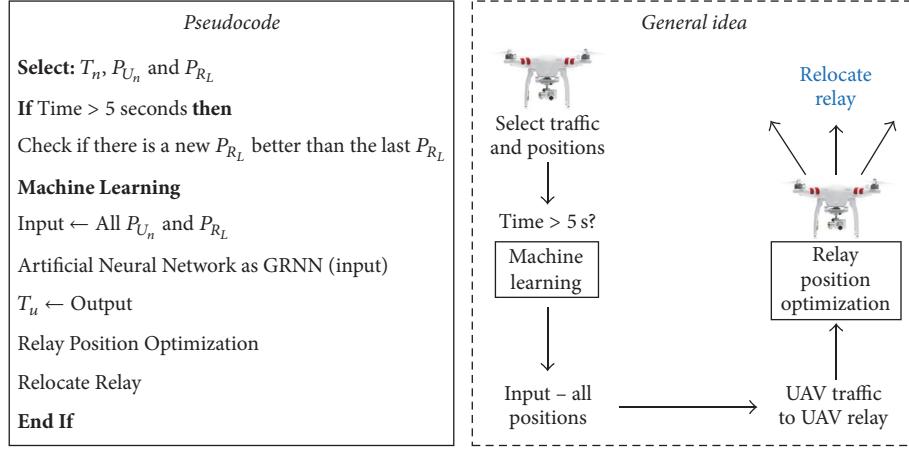


FIGURE 3: FANET model.

location of the other elements in real time, so in addition to the use of GPS (which, on average, sends information about the location every second), the UAVs have an Inertial Measurement Unit (IMU) capable of sending their location at a shorter interval than GPS at any time [14].

Therefore, the arrangement of the devices within a coverage area has a direct impact on the performance of the network and can either improve or worsen in accordance with their mobility. This is one of the main challenges to be solved in this type of network.

4. FANET Model for Optimization Position

As previously mentioned in the case of FANET, the position of the relay should be regarded as a problem to solve or as a means of improving the network performance. This work proposes a novel method for the establishment of communication between one UAV relay and other UAVs in mission flights. The proposal is largely based on the information that can be found in most of the FANET applications, and the aim is to collect data from the environment and to transmit this data to a ground base.

The objective is to drive the UAV relay in real time to the optimal placement with the knowledge of the UAVs positions and maximize the data traffic collection of the UAVs. For this optimization, an Artificial Neural Network (ANN) was used, the training phase estimates the best transmission rate considering the positions of the devices, and the bat algorithm based on the output of the ANN determines the new position for relay UAV through the objective function.

In the creation of the model of the proposal, this paper considers the following information that must be collected (by simulations) for the training of the Artificial Neural Network: (1) position of the UAV relay; (2) position of the UAVs in flight mission; (3) the throughput of each UAV in flight mission with the UAV relay; and (4) the Received Signal Strength Indication (RSSI) of each UAV in flight mission with the UAV relay.

In a multi-UAV system that comprises all the drones arranged in the FANET, U has been established as a set of

UAVs, the number of UAVs in U is denoted by $|U|$, and the UAV relay is denoted by R_L . Each U_n , $1 \leq n \leq |U|$, and R_L in a two-dimensional plane P has coordinates (x_n, y_n) that define its positions as $P_{U_n} = \{x_n, y_n\}$ and $P_{R_L} = \{x_n, y_n\}$. In addition, T_n is defined as the traffic between U_n and R_L , and also RSSI_n is defined as the RSSI between U_n and R_L .

In [10], it was established that, at any moment, the distance between $|U|$ is no closer to a certain threshold and, in this way, the UAVs can avoid collisions. Using the coordinates (x_n, y_n) and the T_n traffic as the database for this work, it was found that each U_n moved in a plane of 200×200 m in a random manner to generate traffic with R_L .

The heuristic algorithm for the traffic problem was used to solve the optimization problem (see Figure 3). This algorithm starts by selecting the positions of the UAVs and relay (and the traffic between them); after this, we look for a new relay positioning that is determined by the Artificial Neural Network. The output of the network will be the traffic that corresponds to positions P_{U_n} and P_{R_L} , and with these results, it is possible to reallocate the position of the relay to a position where the traffic is better than in the previous position.

5. Proposed Heuristics

In light of the scenario described in the previous section, this section focuses on machine learning, in particular on how the neural network works and the bat algorithm (bat) that was used as an optimization algorithm. In the case of the artificial neural networks, we used the Generalized Regression Neural Network (GRNN), the architecture of which is shown in Figure 4, and the data of P_{U_n} , P_{R_L} , and T_u (UAV position, UAV relay position, and UAV traffic to UAV relay, respectively).

The GRNNs have four layers: the input layer (in this work, P_U and P_{R_L} are used as inputs for the model), the intermediate layer, sum layer, and output layer (which is T_n for the model). The number of units in the input layer depends on the total number of variables. The neurons in the input layer transmit the information to the intermediate layer, and each neuron in the intermediate layer corresponds to a training pair. Thus, the number of neurons in the intermediate layer is equal to

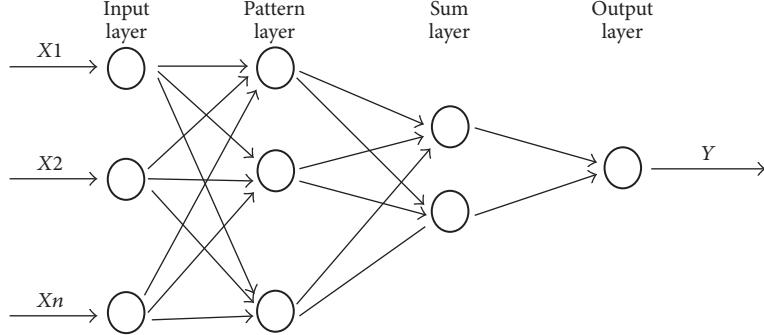


FIGURE 4: The typical architecture of a Generalized Regression Neural Network (GRNN) [3].

the number of training pairs. The third layer is the summation layer; there are only two neurons in this layer: the so-called D -sum and the second neuron, commonly called neuron S -sum. For an input vector X , the output Y of the GRNN is calculated by the following equation:

$$Y(X) = \frac{D}{S} = \frac{\sum_{i=1}^n Y^i \exp[-D_i^2/2\sigma^2]}{\sum_{i=1}^n \exp[-D_i^2/2\sigma^2]}. \quad (1)$$

The bat algorithm simulates the prey detection instinct of micro bats and was created by Yang et al. [15]. The basic rules of the bat algorithm are as follows:

- (i) All bats use the echo location to sense the distance and also “know” the difference between food/prey and the obstacles surrounding them.
- (ii) Bats randomly fly with velocity v_i at position x_i with a fixed frequency f and varied wavelength λ and intensity A_0 to search for prey. Additionally, they automatically adjust the wavelength (or frequency) of their emitted pulses and adjust the pulse emission rate $r \in [0, 1]$, depending on the proximity of their target.
- (iii) Although the loudness may vary in many ways, it has been established that the degree of intensity ranges from a large A_0 (positive) to a constant minimum value A_{\min} .

The bat algorithm yields a combination of UAV and relay positions and evaluates the positioning using the Artificial Neural Network already described. Afterwards, it continuously repeats these steps in order to find a solution that maximizes the throughput.

The optimization loop is in the focus of this work. It is assumed that each of the N neighboring UAVs in mission is in the range of the UAV relay. An initial population of relay positions is randomly created and is the basis for further optimization; hence, the genome $x = (x_1, \dots, x_d)^T$ with $x_1 \dots x_d$ = relay position. The algorithm makes use of the model described in Section 4 and illustrated in Figure 3. When the optimization starts, an initial population is generated and evaluated using the fitness function (GRNN) described along with the selection process. Each iteration of the bat algorithm works in the pseudocode in Algorithm 1.

Postprocess results and visualization: $\beta \in [0, 1]$ is a random vector defined from a uniform distribution. In general, $f_{q_{\min}}$ and $f_{q_{\max}}$ are defined according to the domain size of the problem of interest; at this work it was used for $f_{q_{\min}} = 1$ and for $f_{q_{\max}} = 2$. Pulse rate was defined initially as 0.7 and loudness as 0.5.

The loudness and pulse emission rates are regulated by the following equations:

$$\begin{aligned} A_i^{t+1} &= \alpha * A_i^t, \\ r_i^{t+1} &= r_i^0 * (1 - \exp(-\gamma * t)), \end{aligned} \quad (2)$$

where $0 < \alpha < 1$, $\gamma > 0$

being constants.

In bat algorithm, frequency tuning essentially acts as mutation, while selection pressure is relative constant via the use of current best solution x , found so far. There is no explicit crossover; however, mutation varies based on the variations of loudness and pulse emission. In addition, the variations of loudness and pulse emission rates also provide an auto-zooming ability so that exploitation becomes intensive as the search is approaching global optimality [16].

After the training phase in ANN, a new positioning of the UAVs was found, which formed a new scenario in which it would be necessary to adjust the position of the relay to achieve the best performance of the network. The bat algorithm finds the optimum position, between the UAVs and relay. By minimizing the objective function, called f_{\min} , it will be possible to adjust the position of the relay to obtain a better estimate of the traffic. This optimization heuristic can be seen in Algorithm 2.

Figure 5 shows the flowchart that illustrates the operation of the initial proposal where the bat optimizer algorithm is used to adjust the positioning of the relay in real time in search of the best traffic.

6. Performance Evaluation of FANET Simulations

This section analyzes the results of the computational experiments to prove the efficiency of the optimization algorithm. The proposed method is perturbation-free, does not require time separation between plant dynamics and the optimal

```

Initialize the bat population  $x$  and the velocities of the bats ( $v_i$ ).
Define pulse frequency  $f_i$  at  $x$ .
Initialize pulse rates  $r_i$  and the loudness  $A_i$ 
While not reach the number of maximum iterations:
    Generate more solutions by adjusting frequency and updating velocities and locations/solutions:
        
$$f_i = f_{\min} + (f_{\max} - f_{\min}) * \beta$$

        
$$v_i^t = v_i^{(t-1)} + (x_i^t - x^*) * f_i$$

        
$$x_i^t = x_i^{t-1} + v_i^t$$

        if (rand >  $r_i$ )
            Select a solution among the best solutions
            Generate a local solutions around the selected best solution
        End if
        Generate a new solution by flying randomly
         $f(x)$  is the use of the neural network, GRNN, to generate de traffics to evaluate  $x_i$ .
        If (rand <  $A_i$  &  $f(x)$  <  $f(x^*)$ )
            Accept the new solutions
            Increase  $r_i$  and reduce  $A_i$ 
        End if
        Rank the bats and find the current best  $x^*$ 
End while

```

ALGORITHM 1: Bat algorithm.

```

Bats population “ $b$ ”  $\leftarrow$  new  $P_{R_L}$ 
Input  $\leftarrow$  All new  $P_{U_n}$  and  $b$ 
Neural Network Training as TGRNN
 $T_u \leftarrow$  TGRNN Output
Objective Function  $f_{\min}$ 
 $f_{\min} \leftarrow 54 M * b - \text{sum (output)}$ 
 $f_{\min}$  It is the result that should be minimized

```

ALGORITHM 2: Objective function with GRNN.

placement-seeking loop, and does not require (for a while) knowledge of the propagation model.

The simulations were performed in the NS-2. The NS-2 simulator, like the other simulators in the network, does not provide support to three-dimensional simulations, and for this reason, the simulations carried out in this paper assume that the drones fly over at a similar height, similar as [8] that assumes the UAVs are flying at fixed altitude with LOS (line-of-sight) links. In the simulations, all the UAVs (multiple and relay) had random mobility within an area of 200×200 m. Multiple UAVs performed 500 kbps transmissions for the UAV relay. The simulation parameters used in the simulations are described in Table 1.

The data obtained from the simulations were drawn for the training of the neural network and also for the configuration of the bat algorithm. The parameters contained in Table 2 correspond to the best results for the configuration of the neural network topologies and the bat algorithm.

The evaluation was divided into three parts: (1) Computational Intelligence Results; (2) Quality of Service Results; and (3) Quality of Experience Results.

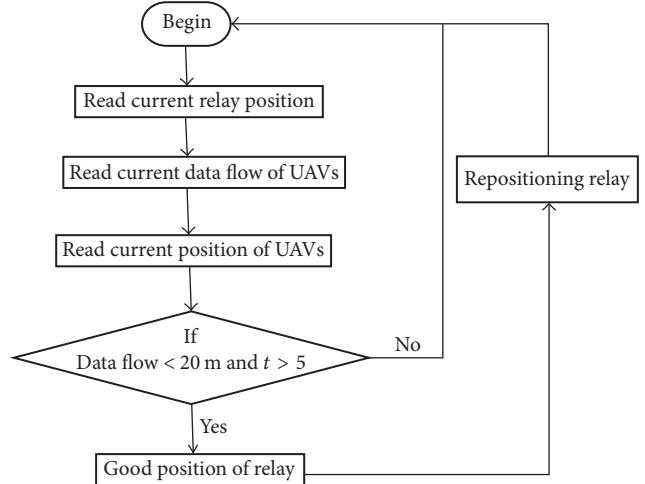


FIGURE 5: Flowchart of the initial proposal.

6.1. Computational Intelligence Results. The first stage entails training the neural network to learn how to relate the positioning of the UAVs and the relay to their respective network traffic. Figures 6(a) and 6(b) show the performance and training regression, respectively. In Figure 6(a), it can be observed that, through the use of cross-validation, a Mean Squared Error of 0.00072 was obtained.

After the neural network training, the bat algorithm was used by following the pseudocode shown in Figure 4, to obtain the minimization objective function. Thus, the smaller the result of f_{\min} , the greater the T_U . Figure 7 shows the evolutionary pattern of f_{\min} in 1000 generations of bats. In this way, each generation seeks to maintain or improve the

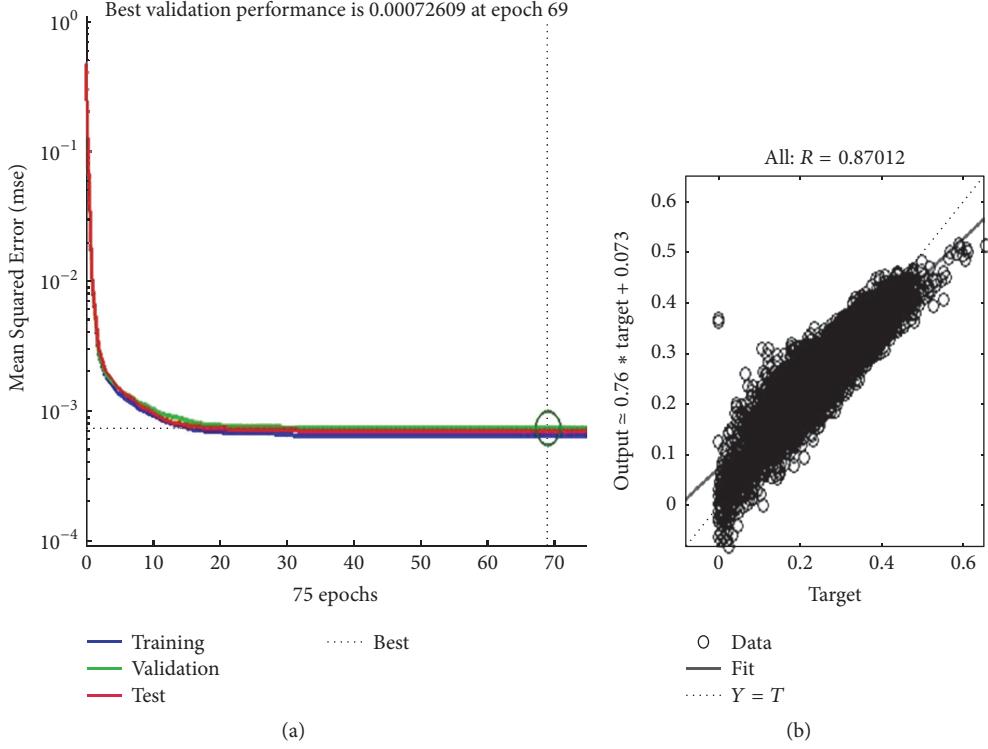


FIGURE 6: Performance and training regression.

TABLE 1: NS-2 configuration parameters.

Drones	4
Access technology	IEEE 802.11g
Propagation model	Shadowing
Mobility type	Random waypoint
Query	Droptail
Number of simulations	100
Confidence interval	95%
Transmission rate	500 Kbps
Frequency	2,4 Ghz
Area	$200 \text{ m} \times 200 \text{ m}$

existing population in order to find the best positioning between U_n and R_L .

The proposed model of this paper turned the complexity problem of relay positioning in a linear problem as showed in Figure 8. The number of iterations (10) versus increasing the number of UAVs was tested.

6.2. Quality of Service Results. This section makes an assessment of the performance of the network before the optimization and then with the optimization. The first evaluated result was the average throughput. As discussed earlier, the positioning of the relay in relation to the other UAVs has a direct impact on the application. It can be seen in Figure 9 that the average throughput with the optimization was higher than the average throughput without the optimization and that there was an improvement of 135% in an area of $200 \times 200 \text{ m}$.

TABLE 2: Main topology features.

GRNN	
Spread	0.3
Number of inputs	8
Number of outputs	3
Number of examples (training)	12402
Number of samples (validation)	50
Number of UAVs	3
Number of relays	1
Bat algorithm	
Population size	40
Number of generations	1000
Sound	0.5
Pulse rate r	0.7
Frequency	[0–2]

In the case of optimization, whenever the multiple UAVs are moved, a new position is found for the UAV relay and usually in a centralized position in relation to the other relays. In this way the optimizer prevents the UAV relay from being far away from the other UAVs during the movement. This optimization is possible because aside from knowing its own position, the UAV relay is given the location of the others and there is an estimation of the new position to which it must move so that it can maintain a good connection with all the multiple UAVs of the network.

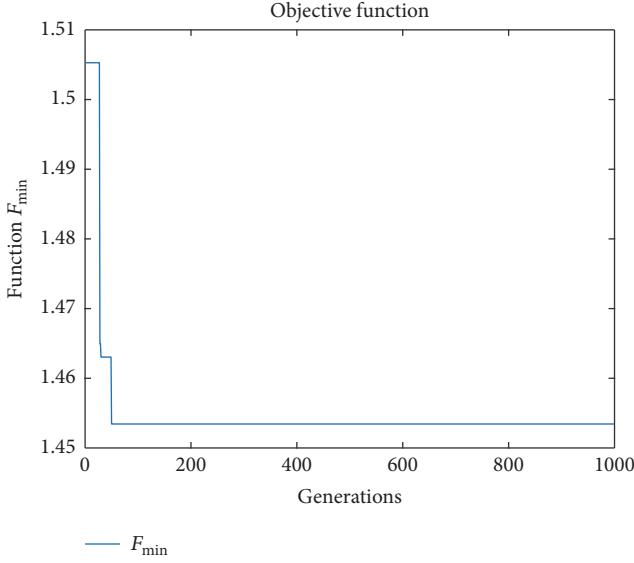


FIGURE 7: Evolutionary pattern.

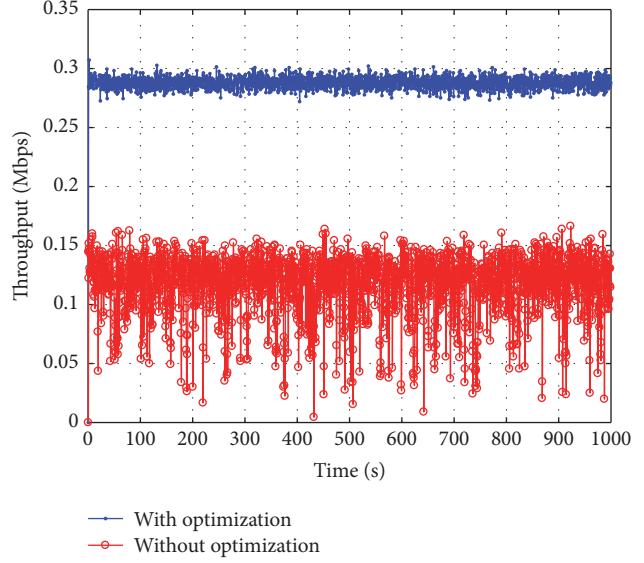


FIGURE 9: Throughput performance.

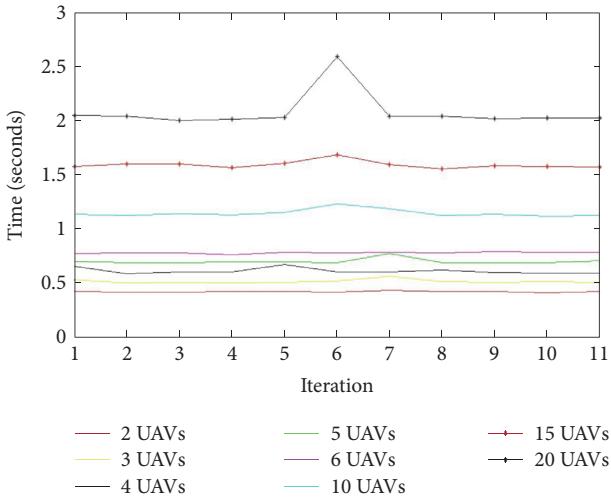


FIGURE 8: Linear problem.

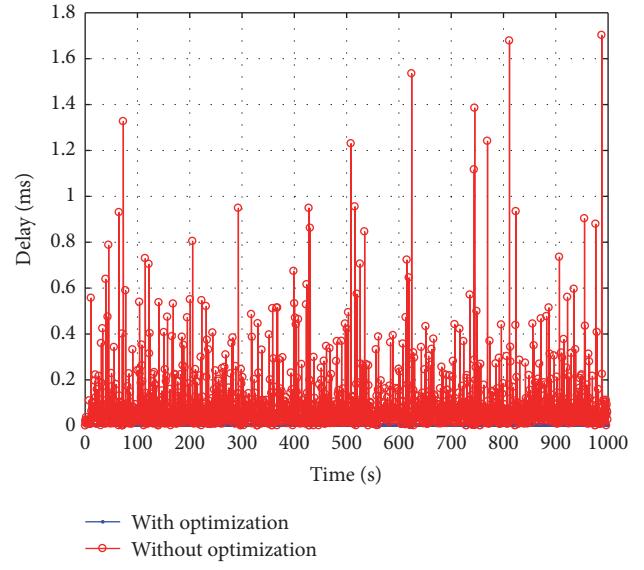


FIGURE 10: Delay performance.

Figure 10 shows the behavior of delays without optimization and after optimization. It can be concluded that without the optimization algorithm, the traffic between the UAVs and relay is subject to higher and more unstable delays (an average delay of 0.0512 ms). Owing to the mobility of the devices, there are occasions when the relay can be more distant from other devices, thus degrading the quality of communication. Therefore, a better result can be found through the optimizer, since there is not much delay variation (an average delay of 0.0034 ms), which makes it more reliable because the relay node is closer and more centralized in relation to the other nodes.

Based on the results obtained so far, Figure 11 shows the scenario for one of the samples used in the experiment in which the positioning of the relay can be seen before the optimization and its respective displacement to an optimal position. In view of the random nature of the mobility, it

is not possible to determine how far away the relay will be from the other relays, and this highlights the importance of the optimizer, which is able to estimate a new position that will allow suitable communication. Thus, after going through all the stages of the optimizer, the relay was moved to a new position so that there could be an improvement of the performance of the traffic between the UAVs.

6.3. Quality of Experience Results. This paper also analyzes the simulations with video applications. The video results are evaluated with the aid of the following objective QoE metrics: Peak Signal to Noise Ratio (PSNR), Structural Similarity Metric (SSIM), and Video Quality Metric (VQM). They are collected by means of the MSU Video Quality Measurement Tool (VQMT).

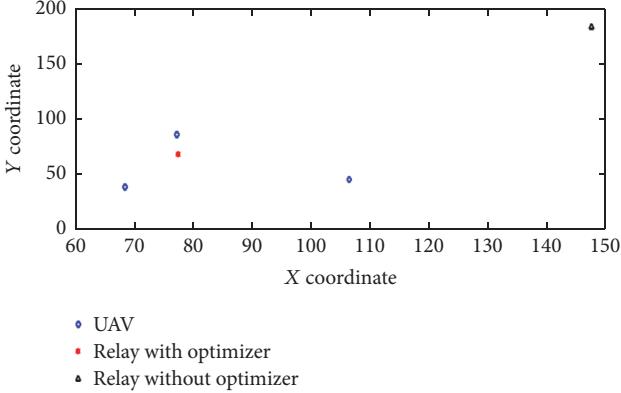


FIGURE 11: Positioning of the relay in multiple UAVs.

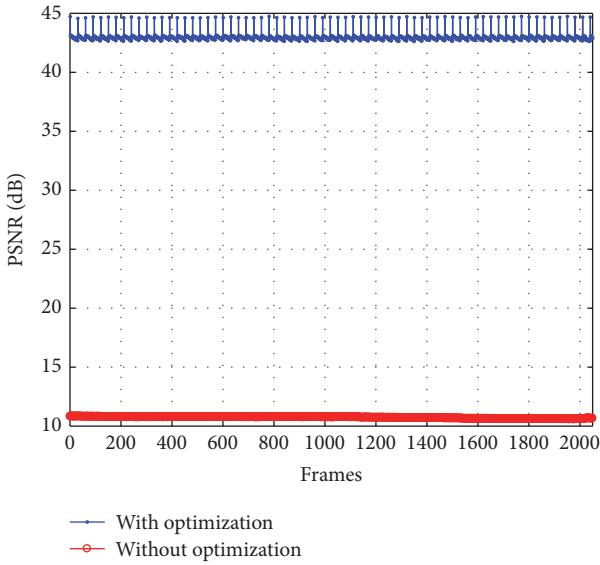


FIGURE 12: PSNR performance.

The PSNR is the most traditional objective metric and compares the quality of video received by the user with that of the original video. The PSNR is derived from another metric called the Mean Squared Error (MSE). The MSE calculates the average variation error of an actual attribute with which this same attribute was estimated. In other words, it compares the quality of the frame received by the user with regard to the original frame.

Figure 12 compares the PSNR of video transmitted both without and with the optimizer. The video streamed without the optimizer had a PSNR average of 10.7 dB; thus, it was rated as a poor quality video, while the video streamed with the optimizer had a PSNR average of 42.9 dB and was rated as an excellent quality video.

Unlike the PSNR that only analyzes the error rate of the video received in relation to the original video, the SSIM evaluates the received video with regard to features such as color, brightness, and structure. The SSIM index is a decimal value between 0 and 1, where 0 means zero correlation with the original image and 1 means exactly the same image.

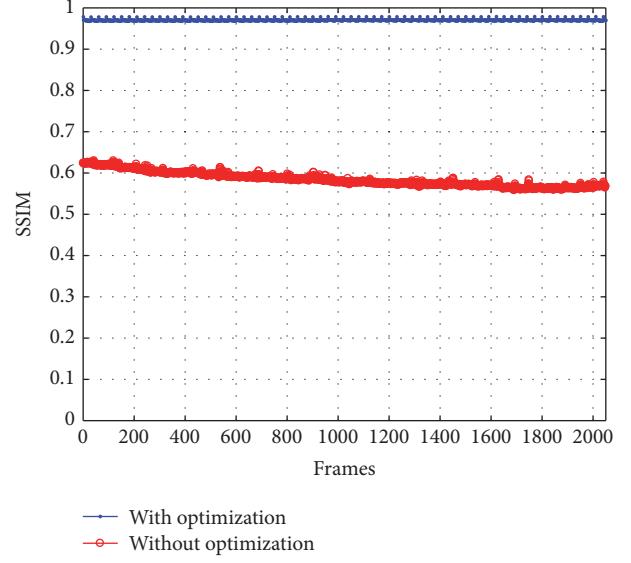


FIGURE 13: SSIM performance.

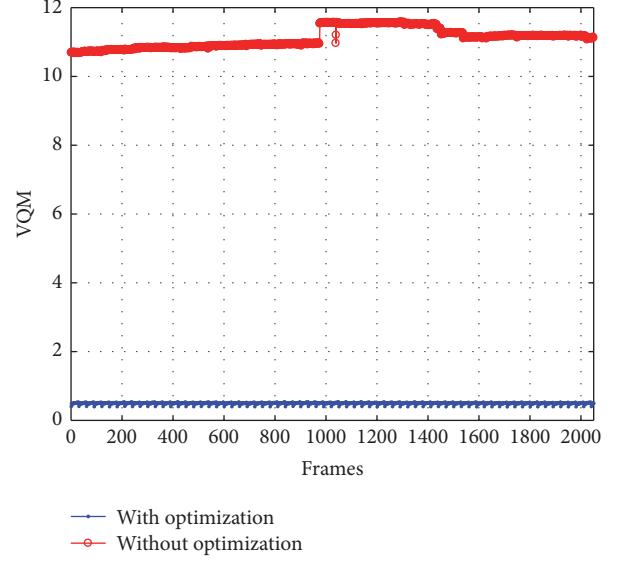


FIGURE 14: VQM performance.

The video streamed without the optimizer obtained an average SSIM of 0.57, while the video transmitted with the optimizer obtained an average SSIM of 0.96. Figure 13 shows the SSIM performance during the transmission.

The VQM is more complete than PSNR and SSIM because it evaluates the color, pixel, and noise distortion if the video is “erased.” The VQM determines the level of multimedia quality based on human eye perception and subjective factors, including blurring, global noise, block distortion, and color distortion. In the evaluation of the VQM results, zero is the best possible score. In the simulations, the video stream without optimization had a mean VQM of 11.1 while the video transmitted with the optimizer had a mean VQM of 0.4 (Figure 14).

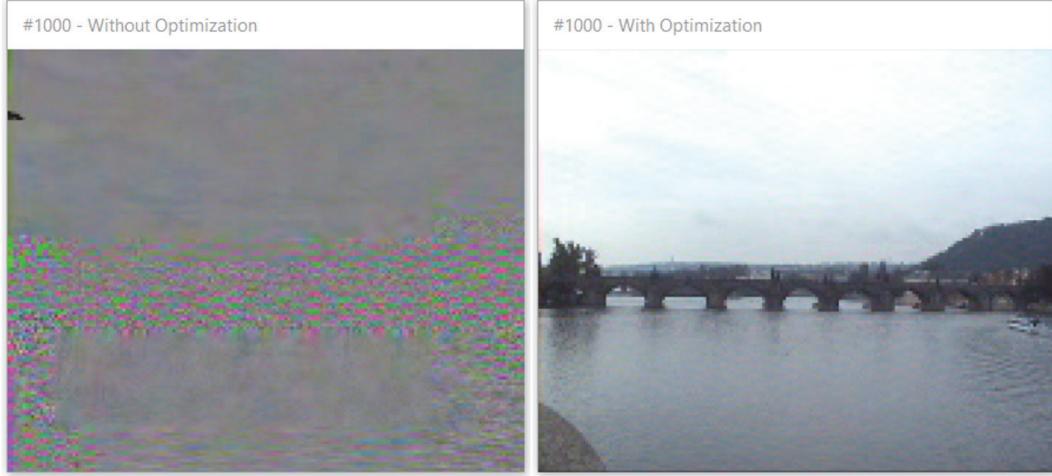


FIGURE 15: Comparison between frames without optimization and frames with optimization.

The efficiency of the optimizer can be proved by comparing a frame of the video transmitted without an optimizer with a frame of the video transmitted with the optimizer. In Figure 15, the superior quality of the video transmitted with the optimizer is evident, because it has moved the relay to a suitable position in relation to the other UAVs and it is able to improve the quality of the communication between them.

7. Final Considerations

With the creation of Ad Hoc Networks, a new challenge has emerged of improving their performance. One of the main problems to be solved is the positioning of UAV relay in relation to the other relays to improve communication in a FANET scenario. The results obtained by simulations underline the importance of combining areas such as telecommunications and computational intelligence, by means of a metaheuristic model, thus obtaining a better result and a better performance of the network as a whole. The use of the proposed model can meet the future needs that may arise in a FANET scenario by improving the communication system, both in civil and in military applications, and obtain satisfactory and more precise results in relation to the above-mentioned related works.

Other metaheuristic optimizers need to be tested to validate the method and confirm the best execution time for performance such as a genetic algorithm, Cuckoo search, and particle swarm. In addition, other scenarios can be tested such as an example of a performance evaluation of traffic caused by large crowds such as football matches, shows, and processions. The bat algorithm is part of a set of metaheuristic optimization algorithms with different applications that has demonstrated its applicability in computer networks as well as in this work.

In future work, the proposal could include a larger number of samples for optimization. This will involve expanding the FANET scenarios with communication to a 4G backbone network and carrying out simulations with other resolutions

(High Definition) video traffic to evaluate network performance in terms of Quality of Experience and the impact of different propagation mechanisms on the type of signal used.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] O. K. Sahingoz, "Networking models in flying Ad-hoc networks (FANETs): Concepts and challenges," *Journal of Intelligent & Robotic Systems*, vol. 74, no. 1-2, pp. 513–527, 2014.
- [2] E. A. Marconato, J. A. Maxa, D. F. Pigatto, A. S. R. Pinto, N. Larrieu, and K. R. L. J. C. Branco, "IEEE 802.11n vs. IEEE 802.15.4: A Study on Communication QoS to Provide Safe FANETs," in *Proceedings of the 46th IEEE/IFIP International Conference on Dependable Systems and Networks, DSN-W 2016*, pp. 184–191, July 2016.
- [3] S. Heddam, H. Lamda, and S. Filali, "Predicting Effluent Biochemical Oxygen Demand in a Wastewater Treatment Plant Using Generalized Regression Neural Network Based Approach: A Comparative Study," *Environmental Processes*, vol. 3, no. 1, pp. 153–165, 2016.
- [4] A. Belbachir, J. Escareno, E. Rubio, and H. Sossa, "Preliminary results on UAV-based forest fire localization based on decisional navigation," in *Proceedings of the Workshop on Research, Education and Development of Unmanned Aerial Systems, RED-UAS 2015*, pp. 377–382, November 2015.
- [5] S. Rohde, M. Putzke, and C. Wietfeld, "Ad hoc self-healing of OFDMA networks using UAV-based relays," *Ad Hoc Networks*, vol. 11, no. 7, pp. 1893–1906, 2013.
- [6] J. Li, Y. Zhou, and L. Lamont, "Communication architectures and protocols for networking unmanned aerial vehicles," in *Proceedings of the IEEE Globecom Workshops*, pp. 1415–1420, December 2013.
- [7] Z. Zhao and T. Braun, "Topology Control and Mobility Strategy for UAV Ad-hoc Networks: A Survey," in *Proceedings of the ERCIM eMobility and MobiSense Workshop*, 2013.

- [8] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement Optimization of UAV-Mounted Mobile Base Stations," *IEEE Communications Letters*, vol. 21, no. 3, pp. 604–607, 2017.
- [9] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," in *Nature Inspired Cooperative Strategies for Optimization (NISCO 2010)*, J. R. Gonzalez, D. A. Pelta, C. Cruz, G. Terrazas, and N. Krasnogor, Eds., vol. 284 of *Studies in Computational Intelligence*, pp. 65–74, Springer, Berlin, Germany, 2010.
- [10] I. Bekmezci, M. Ermis, and S. Kaplan, "Connected multi UAV task planning for Flying Ad Hoc Networks," in *Proceedings of the IEEE International Black Sea Conference on Communications and Networking, BlackSeaCom 2014*, pp. 28–32, May 2014.
- [11] I. Bekmezci, O. K. Sahingoz, and S. Temel, "Flying Ad-Hoc Networks (FANETs): a survey," *Ad Hoc Networks*, vol. 11, no. 3, pp. 1254–1270, 2013.
- [12] J. Gago, C. Douthé, R. E. Coopman et al., "UAVs challenge to assess water stress for sustainable agriculture," *Agricultural Water Management*, vol. 153, pp. 9–19, 2015.
- [13] D. Orfanus, E. P. De Freitas, and F. Eliassen, "Self-Organization as a Supporting Paradigm for Military UAV Relay Networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 804–807, 2016.
- [14] F. Wu, G. Ramchurn, and X. Chen, Coordinating human-avteams in disaster response, 2016.
- [15] Y. Yang, M. M. Polycarpou, and A. A. Minai, "Multi-UAV cooperative search using an opportunistic learning method," *Transactions of the ASME—Journal of Dynamic Systems, Measurement and Control*, vol. 129, no. 5, pp. 716–728, 2007.
- [16] X.-S. Yang, "Swarm intelligence based algorithms: A critical analysis," *Evolutionary Intelligence*, vol. 7, no. 1, pp. 17–28, 2014.

Research Article

Reliability-Centric Analysis of Offloaded Computation in Cooperative Wearable Applications

Aleksandr Ometov,^{1,2} Dmitry Kozyrev,² Vladimir Rykov,² Sergey Andreev,¹ Yuliya Gaidamaka,^{2,3} and Yevgeni Koucheryavy¹

¹Tampere University of Technology, Tampere, Finland

²RUDN University, Moscow, Russia

³Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow, Russia

Correspondence should be addressed to Aleksandr Ometov; aleksandr.ometov@tut.fi

Received 11 August 2017; Accepted 7 November 2017; Published 7 December 2017

Academic Editor: Mostafa Zaman Chowdhury

Copyright © 2017 Aleksandr Ometov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Motivated by the unprecedented penetration of mobile communications technology, this work carefully brings into perspective the challenges related to heterogeneous communications and offloaded computation operating in cases of fault-tolerant computation, computing, and caching. We specifically focus on the emerging augmented reality applications that require reliable delegation of the computing and caching functionality to proximate resource-rich devices. The corresponding mathematical model proposed in this work becomes of value to assess system-level reliability in cases where one or more nearby collaborating nodes become temporarily unavailable. Our produced analytical and simulation results corroborate the asymptotic insensitivity of the stationary reliability of the system in question (under the “fast” recovery of its elements) to the type of the “repair” time distribution, thus supporting the fault-tolerant system operation.

1. Introduction and Motivation

With the rapid proliferation of wireless communications technology, pushed by the avalanche of demanding services, mobile users are preparing to meet much higher levels of service experience than ever before. This poses an unprecedented burden on the existing as well as emerging wireless infrastructures [1]. Indeed, the numbers of connected handheld devices have increased dramatically over the past decade [2], and this growth is only expected to continue further with the introduction of wearables.

Today’s transition toward next-generation (5G) mobile technology brings along an astounding variety of networked devices, like low-end wearables, augmented and virtual reality eye-wear, and self-driving cars, among many others [3, 4]. This paradigm shift unlocks a large diversity of completely new applications and scenarios, which not only require conventional resources (e.g., bandwidth) but also impose other important constraints, primarily along the lines of latency and reliability [5]. The latter is due to an ongoing

evolution of pre-5G networking ecosystem, as it departs from serving primarily human-to-human (H2H) applications to embracing diverse machine-type communications (MTC) services [6, 7].

However, despite significant progress in circuit design, supported by cheaper production processes [8] and increasingly high pace of technological revolution [9], many application-level tasks are still infeasible to be executed in the resource-limited end-user devices [10]. This is because of different factors, such as reduced computation capabilities, high-speed mobility, and stringent security. The above is especially true for the more constrained Internet-of-Things (IoT) objects, including wearables, which are still designed so that their efficiency is maintained only when completing a certain set of predefined tasks, which may hamper their applicability.

To make matters worse, our present-day devices still considerably depend on energy efficiency of communication [11]. People are therefore forced to charge their carriable and wearable equipment on a daily basis, which is a dramatic

departure from the time when mobile phone was able to last for a week on one charging cycle. The energy harvesting and transfer techniques, including wireless charging or generating energy from heat, are still in their infancy, and their mass adoption may not take off any time soon [12], thus leaving the devices with extremely limited battery lifetimes.

Fortunately, in many cases a viable option for “on-the-fly” execution of a processing-hungry task is by offloading (delegating) it to more resource-rich and power-independent node(s). Then, after the input information has been processed remotely (possibly, in a distributed manner), it would be returned to the original constrained device hence saving its resources, as opposed to executing the task locally on the device [13]. Allocating more computing resources in the network and supplying user devices with appropriate communications technology to improve the energy efficiency via offloaded computation pose new challenges. These include the need for reliable wireless connectivity and failure-free remote processing of information, which could be resolved with the advent of communication and computing cooperation (3C) techniques [14, 15].

The development of 3C solutions has a long history behind, as it comes from conventional cloud computing [16], through fog structures [17], all the way to the network edge [18, 19]. This transition, in turn, underpins a new concept of mobile edge computing (MEC) as an emerging computing and caching paradigm [20]. MEC was initially employed in 2013 to facilitate the execution of services at the edge of a network, when IBM and NSN introduced a platform that could run applications on a mobile base station. These early efforts only had local impact [21] and did not consider other important aspects, like application migration and interoperability, among others.

MEC grew to its current importance a year after, when ETSI launched their Industry Specification Group (ISG) for mobile edge computing in 2014. With this specification, MEC aimed to “provide an IT service environment and cloud computing capabilities at the edge of the mobile network.” This group also pursues creation of an open ecosystem, where service providers can deploy their applications across multi-vendor MEC platforms, and the target services are expected to include augmented reality, intelligent video acceleration, connected cars, and IoT gateways, to name a few [22]. Once standardization is complete, telecommunication companies will become capable of deploying this service environment in their infrastructures.

In light of the above, the main contributions and the structure of this paper are as follows. We address an important question of whether MEC as part of the 3C system can operate reliably in advanced 5G-grade scenarios, such as “connected law enforcement officer” in urban settings. The subsequent Section 2 offers a brief review of computation offloading and communications techniques for smart mobile devices, mainly focusing on network-edge operation. Further, the use of specialized wearable equipment utilized in the proposed system design is discussed in Section 3. The paper continues with our developed analytical model in Section 4. In Section 5, we deliver the corresponding numerical results. Our study is concluded in the last section.

2. Communications and Computing Cooperation Strategies

2.1. Review of Emerging 3C Solutions. One of the most prominent solutions under the umbrella of 5G ecosystem from the 3C perspective is computation offloading [23]. This paradigm has been coined for resource-constrained and energy-hungry devices, which have difficulty in handling their heavy computation tasks and thus need to delegate those to more resourceful nodes. This allows to make a decisive step forward in the direction of content-oriented networking, in contrast to connection-oriented technology, thus targeting an increase in cooperative network capacity [24].

To this end, the stringent requirements of new applications and services are pushing the network operators to develop diverse communications capabilities in the form of heterogeneous networks (HetNets) [25–27]. For example, cellular macrocells can deliver coverage, while smaller cells may supply capacity, hence providing the users with better quality of experience (QoE) and lowering harmful interference [28]. Therefore, HetNets emerge as advanced architectures that are able to serve demanding user scenarios, by utilizing increasingly diverse wireless technologies across the same geographical area, but posing a complex challenge of backhaul efficiency [29].

Making backhaul more effective than in already deployed HetNet solutions may be costly for the network operators. On the other hand, radio access network (RAN) capacity can potentially be increased by deploying higher numbers of small cells. As one of the approaches to overcome this challenge, various caching techniques have been proposed and utilized [30, 31]. These provide the end users with easy access to content of high demand, which is stored locally instead of being transferred repeatedly over “expensive” communications links, especially during peak hours.

The range of network architectures to be utilized is very wide. Accounting for the fact that any of the proposed techniques can be executed not only on the infrastructure side but also by offloading tasks onto other user devices in proximity, device-to-device (D2D) technologies have recently seen a renewed research interest [32, 33]. Further facilitated by the latest progress in affordable memory capacity, transparent caching in key locations allows boosting the content delivery as well as improving the network resource utilization, even when the same content is not requested by the users simultaneously.

2.2. Emerging User Applications. As follows from the above, capable mobile devices (smartphones, high-end wearables, cars, etc.) are changing the lives of people. Already in 2015, the time spent by the average user online has been up to 5 hours daily [34] and it is commonly anticipated to grow further, since people turn more to their mobile devices as opposed to legacy stationary equipment. This, in turn, creates higher loads due to repetitive content downloads that could however be cached either on the network or on the user side.

A particularly challenging use case in the above context, enabled by the decisive advancements in wearable and computing technologies, is augmented reality (AR), which opens

the door to genuinely interactive user experience [35]. Different from virtual reality (VR), AR aims to supplement the real world instead of creating an entirely artificial environment. To this end, various objects in the person's surroundings become the target items for computer-generated annotations, which require complex and often real-time computation. This, in turn, substantiates the need for content delivery transformation, thus producing a diversity of computation-heavy and delay-sensitive AR applications [36].

While modern AR devices are already capable of advanced hardware acceleration [37], the opportunities for software-based data analysis remain rather limited. Today, the data is typically sent to the dedicated processing device by utilizing a cable connection. Hence, leading vendors are actively pursuing the implementation of mobile and truly wireless AR/VR technology [38]. However, the task of pattern matching, especially in the video processing applications, is still very challenging for the standalone smart glasses, and hence distributed/offloaded computation methods rapidly become one of the key areas of interest [39].

2.3. Computation Offloading Techniques. The notion of “cloud” first appeared in the late 70s, thus shaping the concept of distributed storage and computation. Since then, the evolution of cloud-based techniques counted a long history and in this paper we sketch the state of the art in this area only briefly.

(1) *Cloud and Fog Computing.* Generally, the concept of migrating the computation and data from portable and desktop devices to larger data centers is known as cloud computing [40]. The applications of interest are therefore delivered as services to the end users over the Internet, while the hardware for actually running those is “rented” from the service providers. Cloud-service clients help receive more capacity during the peak demand hours, reduce the device costs, and increase the network utilization dramatically, primarily at the network edge.

Another development in cloud computing is named fog computing [41]. The original definition was, “extension of the cloud computing paradigm (that) provides computation, storage, and networking services between end devices and traditional cloud servers” [42]. Technically, fog computing is a supplement for the conventional cloud computing that allows analyzing localized information “on-the-fly” by taking advantage of the locally deployed servers instead of forwarding all of the data to the remote server.

(2) *Mobile Edge Computing.* The core idea behind the concept of MEC is to execute the demanding applications even closer to the end user, that is, on the most proximate base stations [43]. The main benefits delivered to both the network operators and the end clients are as follows: (i) lower latency; (ii) higher bandwidth; and (iii) improved location awareness.

(3) *Mobile Cloud Computing.* Another noteworthy direction in the development of computation offloading from MEC techniques is mobile cloud computing (MCC) [44, 45]. Based on the concept of mobile delegation, MCC leads to

decentralized storage and facilitates distribution of execution between various mobile nodes that may operate at the network edge [46].

Solutions for MCC are many, but in cases where wireless connection to the base station (or, more generally, to the cloud) is not reliable, the devices can group together and solve the task collaboratively [47]. Existing multiplatform frameworks already allow migrating a part of the device's code to the neighboring “helper” nodes [48]. When such delegation is complete, the memory, CPU state, and other application-related resources are utilized independently. After the task at hand is finalized, the result is delivered back to the requester device.

In the following subsection, we briefly elaborate on the caching techniques that correspond to the reviewed offloading mechanisms.

2.4. Caching Strategies. Similarly to the case of classifying the computing strategies, two main groups of caching techniques could be identified. One involves using the network infrastructure, while the other is based on utilizing distributed user devices in proximity.

(1) *Conventional Caching Strategies.* First, infrastructure-based caching strategies are to be considered. In the Long Term Evolution (LTE) cellular networks by 3GPP, two types of locations appear attractive for deploying 5G-grade caches [49]: (i) in the Evolved Packet Core (reflecting the cloud computing paradigm) or (ii) in the RAN on the base station side (representing the MEC concept). For example, caching the content library at the femtocell stations (so-called femtocaching [50]) holds a promise to mitigate the backhaul requirements in next-generation 5G-grade HetNets [51].

(2) *Distributed Caching Strategies.* On the other hand, mobile user equipment already has a significant amount of data storage at its disposal. Since most of the users tend to consume similar content in their daily life [52] (e.g., news, social network data, blogs, and popular location-based files), the possibilities of utilizing direct D2D links for delivering popular content arise as an attractive solution [53]. Moreover, the costs of such technology integration appear to be more affordable as compared to the infrastructure-based caching.

The discussed development of 3C toward MEC and MCC solutions opens numerous opportunities for emerging 5G applications. The following section outlines our representative scenario of interest, where the goal is to select the most appropriate approach for supporting an advanced wearable application.

3. Utilization of High-End Wearables

In the rest of this paper, we focus on the *distributed* caching and computing strategies within the scope of MEC and MCC paradigms. Assuming that modern user devices are already capable of storing large amounts of information, this data needs to be delivered to more computationally powerful nodes for further processing [54]; see our considered example in Figure 1.



FIGURE 1: Connected law enforcement officer scenario.

One of the feasible approaches here is to enable dynamic data prefetching, where mobile but more computationally powerful devices (or infrastructure nodes) acquire and deliver the relevant information to less computationally powerful equipment [55]. The latter could become the basis for the following three main applications:

- (i) *Caching of popular content*: news, videos, maps, and more could be stored in the devices based on their proximity and then delivered directly if requested by other nodes in the network [56].
- (ii) *Incentive mechanisms*: users in proximity may assist the target device based on a certain incentive mechanism, be it operator bonuses, seller discounts, or access to another service.
- (iii) *Special cases*: users may be engaged in assisting the dedicated devices of special forces, that is, police, firefighters, and ambulance, among others, thus providing them with the required computing and caching resources upon request.

In all of the above listed cases, there may be situations when the device/network has the knowledge that it will be approached by another device in need later on. Then, the content can be precached in the “helper” device proactively by taking advantage of a reliable connection to the infrastructure, while refraining from utilizing this expensive

connectivity for the target requester device [57]. The security, privacy, and anonymity remain as important challenges to facilitate such operation but are left out of scope of this work. The authors assume that all data is encrypted and, for example, pairwise key cryptography is utilized.

3.1. Smart Devices of Interest. Today, various types of special forces are equipped with advanced devices [58], including smart cameras, health monitors, and communications systems. While health monitoring may not require real-time and heavy data analytics [59, 60], the assessment of the surrounding environment can have tremendous demands in terms of communications and computation.

Present-day technology revolution is primarily pushed by the industry giants. For example, Connected Law Enforcement Officer program by Motorola facilitates real-time collaboration experience between different specialized devices to improve the response times within the ecosystem of mission-critical services. This solution makes it possible for teams of people, in the field or in the office, to effectively communicate and cooperate by exchanging voice and data, securely, reliably, and in real time, regardless of network, carrier, or device, whether deployed inside the client’s network infrastructure or in the cloud as a hosted/managed solution.

In order to improve the operating efficiency and accelerate decision making, the devices need to collaborate

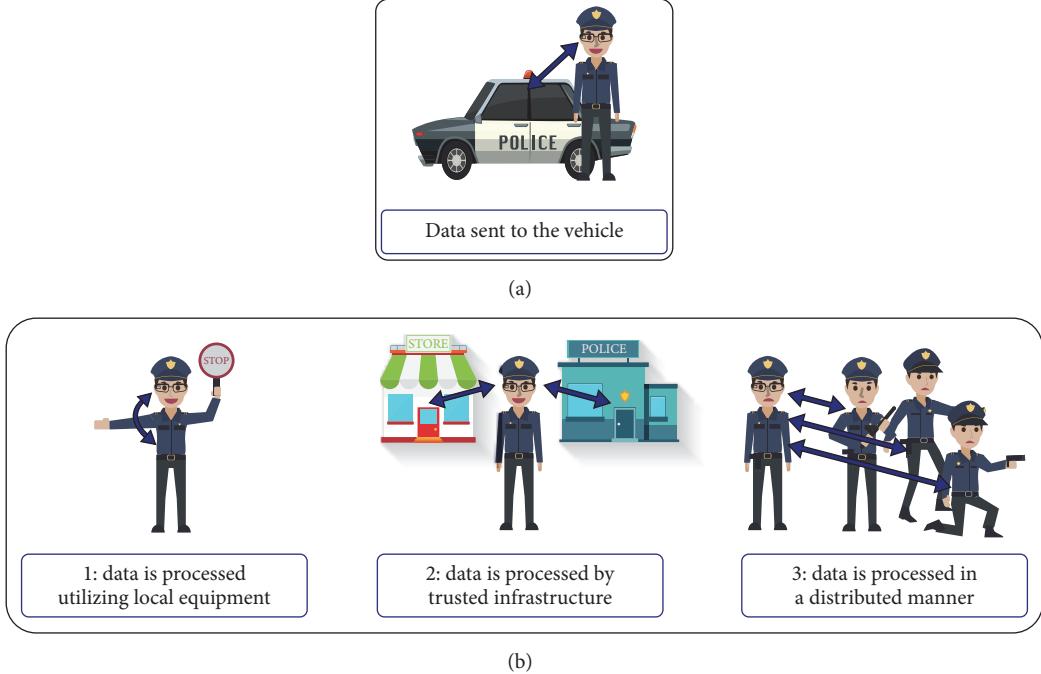


FIGURE 2: Our considered scenarios for real-time data offloading.

seamlessly and share information “on-the-fly.” The considered set of mobile devices typically includes the following:

- (i) *Video speaker microphone*: it combines a body-mounted camera with a high-quality radio speaker microphone and a touchscreen interface, for intuitive capture, storage, and management of video, audio, and image evidence.
- (ii) *Portable radio*: it provides reliable “push-to-talk” communications as well as superior audio quality for essential day-to-day and emergency two-way radio communications.
- (iii) *Handheld smartphone*: it offers reliable broadband connectivity and enables users to enjoy intelligent support of the entire team. With its rugged form factor, powerful data applications, and a quick-response user interface, it delivers “always-on” mission-critical user experience to improve situational awareness for police, firefighter, rescue, and emergency medical services as well as for organizations that support critical infrastructure.
- (iv) *Connected car*: it is equipped with, for example, TETRA radio, mobile computing features, and mission-critical 3GPP LTE broadband connectivity; it supports intuitive data applications as well as automatic number plate recognition (ANPR), video surveillance, lights, and sirens functionality, which can be accessed and managed effortlessly. This vehicle is particularly of interest for our case study due to its superior computation and caching capabilities, as well as more stable connectivity to the network infrastructure.

- (v) *AR glasses* are expected to join this set soon: they assist the user in a range of tasks and provide real-time assessment of the surrounding environment for faster response [61].

The considered application ecosystem includes real-time video streaming, computer-aided dispatch, unified “push-to-talk,” electronic citation, report writing, and data capture. In what follows, we primarily focus on the representative use case where a group of high-tech equipped officers on duty collaborate with each other. However, our findings generally hold valid for a range of similar mission-critical scenarios as discussed below.

3.2. Scenario of Interest. In this work, we concentrate on a study case that involves, for example, daily police officer operation, as demonstrated in Figure 2(a). The considered conventional duties include travel of the officer using the law enforcement vehicle, that is, a car, a motorcycle, or any other means of transport [62]. In this case, the resource-limited device is the AR glasses, which for the most of the time has the possibility of connecting to a resource-rich processing unit. The car can feature the required datasets and hardware to simultaneously process a high number of pattern-matching tasks from the AR glasses thus acting as the MEC/MCC unit. By this means, the police officer’s gear can effectively enjoy higher processing power.

Another study case reflects the scenarios, where the AR-user leaves the proximity with the powerful processing unit, as shown in Figure 2(b). This may happen for a long list of reasons, like an emergency call, due to operational causes, and so on. This scenario comprises the following modes:

- (1) *Patrolling (b1)*: if the processing unit is far away, the AR glasses may keep analyzing the most crucial

patterns in their vicinity according to the user preferences. Here, the battery resource becomes the main limiting factor, since running complex computation on this single device may significantly degrade energy efficiency.

- (2) *Visiting connected buildings (b2)*: if after losing direct connectivity to the processing node (discontinuing pure MCC operation), the AR-equipped user is moving toward, for example, a police station, shopping mall, or hospital, the respective pattern-matching sets can be securely transferred to the trusted infrastructure as in the MEC case. Hence, the AR glasses may be assisted by relying on the distributed pattern-matching functionality and therefore save their limited resources.
- (3) *Emergency call (b3)*: if an emergency situation occurs, the officer may need to arrive at the scene as part of the dispatched unit. In this case, the control center has the knowledge on who participates in the mission and can identify the vehicles and, potentially, other available (mobile) stations in proximity. Therefore, the required datasets can be migrated accordingly onto these nearby devices for the AR glasses to take advantage of the distributed collaborative processing.

On the one hand, utilizing “helper” devices with more abundant computation and energy resources becomes the desired option. However, in case a group of police officers turns out to be on a mission without direct access to such MEC/MCC unit (or the connection is not sufficiently reliable), the data in question may be distributed among the group of officers so that the latter could run collaborative computation in a distributed manner; that is, the number of simultaneously processing/caching “servers” becomes more than one.

Since the above cases *a* and *b1* do not pose much of a research challenge, we concentrate below on modeling a scenario built around *several* content-caching and processing devices that serve the tagged unit. We are specifically interested in assessing the reliability perspective of such operation, which is discussed in detail in the following section.

4. Proposed Reliability-Centric Modeling

This section introduces a mathematical framework suitable for analyzing the fault-tolerant collaborative operation of tagged and helper devices. In this work, we assume that the “active” (available) and the “repair” (unavailable) operation times for cooperating system elements are independent. The notation employed further on is summarized in Main System Model Notations.

4.1. General Considerations. Consider a random process $v(t)$, which corresponds to the number of failed collaborating elements at time t , where the set of states in the system is $E = 0, 1, \dots, k$. To characterize the behavior of the system in question by relying on a Markov process, we introduce an additional variable $x(t) \in R_+$, which is the overall duration spent at time t for the failed element recovery. We obtain

a two-dimensional process $(v(t), x(t))$ with an extended steady-state space $\mathcal{E} = \{(0), (1, x), (2, x), \dots, (k, x)\}$.

We further denote $p_0(t)$ as the probability that at time t the system is in the state $i = 0$, and $p_i(t; x)$ is the probability density function (in continuous component) that captures the probability that at time t the system is in state i ($i = \overline{1, k}$). The time taken to “repair” the failed element is in the range $(x, x + dx)$. The probabilities in question are calculated as

$$\begin{aligned} p_0(t) &= P\{\nu(t) = 0\}, \\ p_1(t, x) dx &= P\{\nu(t) = 1, x < x(t) < x + dx\}, \\ p_2(t, x) dx &= P\{\nu(t) = 2, x < x(t) < x + dx\}, \\ &\vdots \\ p_k(t, x) dx &= P\{\nu(t) = k, x < x(t) < x + dx\}. \end{aligned} \quad (1)$$

Consider the probabilities of two events that are important for our further discussion.

First Event. A system element has failed during time Δ under the condition that it has already worked for x units of time before its failure.

$$\begin{aligned} P\{x \leq A < x + \Delta \mid A \geq x\} &= \frac{P\{x \leq A < x + \Delta\}}{P\{A \geq x\}} \\ &= \frac{A(x + \Delta) - A(x)}{1 - P\{A < x\}} = \frac{a(x)\Delta}{1 - A(x)} = \frac{\alpha e^{-\alpha x}\Delta}{e^{-\alpha x}} \\ &= \alpha\Delta + o(\Delta). \end{aligned} \quad (2)$$

Second Event. Starting from the considered time instant, an element will be restored during the time Δ , provided that it has already been in “repair” for x time units.

$$\begin{aligned} P\{x \leq B < x + \Delta \mid B \geq x\} &= \frac{P\{x \leq B < x + \Delta\}}{P\{B \geq x\}} \\ &= \frac{B(x + \Delta) - B(x)}{1 - P\{B < x\}} = \frac{b(x)\Delta}{1 - B(x)} = \delta(x)\Delta. \end{aligned} \quad (3)$$

Steady-state transition diagram corresponding to the operation of the system in question is illustrated in Figure 3.

By applying the law of total probability, we transition to Kolmogorov’s system of differential equations, which allows obtaining the steady-state probabilities of the system at hand. Consider the probability $p_0(t + \Delta)$ that corresponds to the case where there are 0 failed elements at time instant $t + \Delta$. This situation may appear (i) if at time instant t either there were 0 failed elements or the active one has not failed during time instant Δ (with probability $1 - \alpha\Delta$) or (ii) if there was 1 failed element and it has been restored during time Δ (with probability $\delta(x)\Delta$), while no other active element has failed during this time (with probability $(1 - \alpha\Delta)$).

$$\begin{aligned} p_0(t + \Delta) &= P\{\nu(t + \Delta) = 0\} \\ &= p_0(t)(1 - \alpha\Delta) \end{aligned}$$

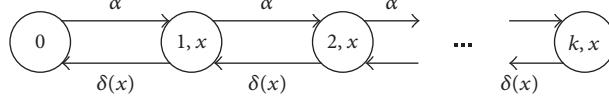


FIGURE 3: Steady-state transitions in the considered system.

$$\begin{aligned}
& + \int_0^t p_1(t, x) \cdot \delta(x) \cdot \Delta \cdot (1 - \alpha\Delta) dx \\
= & p_0(t) (1 - \alpha\Delta) \\
& + \int_0^t p_1(t, x) \cdot \delta(x) \cdot \Delta dx. \tag{4}
\end{aligned}$$

Hence, we establish the state-related equations as

$$p_0(t + \Delta) = p_0(t) \cdot (1 - \alpha\Delta) + \int_0^t p_1(t, x) \delta(x) \Delta dx, \tag{5}$$

and, similarly,

$$\begin{aligned}
& p_1(t + \Delta, x + \Delta) dx \\
= & P\{\nu(t + \Delta) = 1, x + \Delta < x(t + \Delta) < x + \Delta + dx\} \tag{6} \\
= & p_1(t, x) dx \cdot (1 - \alpha\Delta) \cdot (1 - \delta(x) \cdot \Delta).
\end{aligned}$$

After the necessary transformations, we obtain

$$p_1(t + \Delta, x + \Delta) = p_1(t, x) \cdot (1 - \alpha\Delta) \cdot (1 - \delta(x) \cdot \Delta). \tag{7}$$

At the next step, we have

$$\begin{aligned}
& p_2(t + \Delta, x + \Delta) dx \\
= & P\{\nu(t + \Delta) = 2, x + \Delta < x(t + \Delta) < x + \Delta + dx\} \tag{8} \\
= & p_2(t, x) dx \cdot (1 - \alpha\Delta) \cdot (1 - \delta(x) \cdot \Delta) \\
& + p_1(t, x) dx \cdot \alpha\Delta.
\end{aligned}$$

Further, we arrive at the equation

$$\begin{aligned}
p_2(t + \Delta, x + \Delta) = & p_2(t, x) \cdot (1 - \alpha\Delta) \cdot (1 - \delta(x) \cdot \Delta) \tag{9} \\
& + p_1(t, x) \cdot \alpha\Delta.
\end{aligned}$$

Similarly, we produce the equations for $p_i(t + \Delta, x + \Delta)$, $i = \overline{3, k-1}$ as

$$\begin{aligned}
p_i(t + \Delta, x + \Delta) = & p_i(t, x) \cdot (1 - \alpha\Delta) \cdot (1 - \delta(x) \Delta) \tag{10} \\
& + p_{i-1}(t, x) \alpha\Delta, \quad i = \overline{3, k-1}.
\end{aligned}$$

At the next step, we establish

$$\begin{aligned}
& p_k(t + \Delta, x + \Delta) dx \\
= & P\{\nu(t + \Delta) = k, x + \Delta < x(t + \Delta) < x + \Delta + dx\} \\
= & p_k(t, x) dx \cdot (1 - \delta(x) \cdot \Delta) + p_{k-1}(t, x) dx \cdot \alpha\Delta \\
& \cdot (1 - \delta(x) \cdot \Delta) \tag{11} \\
= & p_k(t, x) dx \cdot (1 - \delta(x) \cdot \Delta) + p_{k-1}(t, x) dx \cdot \alpha\Delta \\
& - p_{k-1}(t, x) dx \cdot \alpha \cdot \delta(x) \cdot \Delta^2 \\
= & p_k(t, x) dx \cdot (1 - \delta(x) \cdot \Delta) + p_{k-1}(t, x) dx \cdot \alpha\Delta,
\end{aligned}$$

where $\Delta^2 \ll \Delta \Rightarrow o(\Delta^2) \rightarrow 0$.

Then, we obtain for the general case the equation

$$\begin{aligned}
p_k(t + \Delta, x + \Delta) = & p_k(t, x) \cdot (1 - \delta(x) \cdot \Delta) \tag{12} \\
& + p_{k-1}(t, x) \cdot \alpha\Delta.
\end{aligned}$$

The boundary conditions in our system are

$$\begin{aligned}
p_1(t + \Delta, 0) = & p_0(t) \cdot \alpha + \int_0^t p_2(t, x) \delta(x) dx, \\
p_2(t + \Delta, 0) = & \int_0^t p_3(t, x) \delta(x) dx, \tag{13} \\
& \vdots \\
p_{k-1}(t + \Delta, 0) = & \int_0^t p_k(t, x) \delta(x) dx.
\end{aligned}$$

After the transformations, and based on (5), we establish

$$\begin{aligned}
& p_0(t + \Delta) \\
= & p_0(t) \cdot (1 - \alpha\Delta) + \int_0^t p_1(t, x) \cdot \delta(x) \cdot \Delta dx, \tag{14} \\
& p_0(t + \Delta) - p_0(t) \\
= & -p_0(t) \cdot \alpha\Delta + \int_0^t p_1(t, x) \cdot \delta(x) \cdot \Delta dx, \\
& \frac{1}{\Delta} \cdot (p_0(t + \Delta) - p_0(t)) \\
= & -p_0(t) \cdot \alpha + \int_0^t p_1(t, x) \cdot \delta(x) dx.
\end{aligned}$$

Passing to the limit $\Delta \rightarrow 0$, we obtain

$$\frac{dp_0(t)}{dt} = -\alpha \cdot p_0(t) + \int_0^t p_1(t, x) \cdot \delta(x) dx. \quad (15)$$

It follows from (7) that

$$\begin{aligned} p_1(t + \Delta, x + \Delta) &= p_1(t, x) \cdot (1 - \alpha\Delta) \cdot (1 - \delta(x) \cdot \Delta) \\ &= p_1(t, x) - p_1(t, x) \cdot \alpha\Delta - p_1(t, x) \cdot \delta(x) \cdot \Delta, \end{aligned} \quad (16)$$

$$\begin{aligned} p_1(t + \Delta, x + \Delta) &- p_1(t, x + \Delta) + p_1(t, x + \Delta) \\ &- p_1(t, x) = -p_1(t, x) \alpha\Delta - p_1(t, x) \delta(x) \Delta, \end{aligned}$$

$$\begin{aligned} \frac{p_1(t + \Delta, x + \Delta) - p_1(t, x + \Delta)}{\Delta} &+ \frac{p_1(t, x + \Delta) - p_1(t, x)}{\Delta} = -p_1(t, x) \alpha \\ &- p_1(t, x) \delta(x). \end{aligned} \quad (17)$$

Passing to the limit $\Delta \rightarrow 0$, we calculate

$$\frac{\partial p_1(t, x + \Delta)}{\partial t} + \frac{\partial p_1(t, x)}{\partial x} = -(\alpha + \delta(x)) \cdot p_1(t, x). \quad (18)$$

Similarly, based on (9), we have

$$\begin{aligned} \frac{p_2(t + \Delta, x + \Delta) - p_2(t, x + \Delta)}{\Delta} &+ \frac{p_2(t, x + \Delta) - p_2(t, x)}{\Delta} \\ &= -p_2(t, x) (\alpha + \delta(x)) + p_1(t, x) \alpha. \end{aligned} \quad (19)$$

Further, passing to the limit $\Delta \rightarrow 0$, we obtain

$$\begin{aligned} \frac{\partial p_2(t, x + \Delta)}{\partial t} + \frac{\partial p_2(t, x)}{\partial x} &= -(\alpha + \delta(x)) \cdot p_2(t, x) + \alpha p_1(t, x). \end{aligned} \quad (20)$$

Finally, it follows from (10) that

$$\begin{aligned} p_k(t + \Delta, x + \Delta) &= p_k(t, x) \cdot (1 - \delta(x) \cdot \Delta) \\ &+ p_{k-1}(t, x) \cdot \alpha\Delta, \\ p_k(t + \Delta, x + \Delta) &= p_k(t, x) - p_k(t, x) \cdot \delta(x) \cdot \Delta \\ &+ p_{k-1}(t, x) \cdot \alpha\Delta, \\ p_k(t + \Delta, x + \Delta) &- p_k(t, x + \Delta) + p_k(t, x + \Delta) \\ &- p_k(t, x) = -p_k(t, x) \cdot \delta(x) + p_{k-1}(t, x) \cdot \alpha, \\ \frac{p_k(t + \Delta, x + \Delta) - p_k(t, x + \Delta)}{\Delta} &+ \frac{p_k(t, x + \Delta) - p_k(t, x)}{\Delta} = -\delta(x) \cdot p_k(t, x) + \alpha \\ &\cdot p_{k-1}(t, x). \end{aligned} \quad (21)$$

Passing to the limit $\Delta \rightarrow 0$, we find

$$\begin{aligned} \frac{\partial p_k(t, x + \Delta)}{\partial t} + \frac{\partial p_k(t, x)}{\partial x} &= -\delta(x) \cdot p_k(t, x) + \alpha \cdot p_{k-1}(t, x). \end{aligned} \quad (22)$$

The boundary conditions with respect to the passage to the limit $\Delta \rightarrow 0$ are

$$\begin{aligned} p_1(t, 0) &= \alpha \cdot p_0(t) + \int_0^t p_2(t, x) \delta(x) dx, \\ p_2(t, 0) &= \int_0^t p_3(t, x) \delta(x) dx, \\ &\vdots \\ p_{k-1}(t, 0) &= \int_0^t p_k(t, x) \delta(x) dx. \end{aligned} \quad (23)$$

Since the process at hand is positively recurrent (due to finite state space as well as positive and bounded transition rates), there exist limit state probabilities under $t \rightarrow \infty$, which coincide with the stationary probabilities as

$$\begin{aligned} p_0 &= \lim_{t \rightarrow \infty} p_0(t), \\ p_i(x) &= \lim_{t \rightarrow \infty} p_i(t; x) \quad (i \in \{1, \dots, k\}). \end{aligned} \quad (24)$$

Hence, we arrive at the following system of equations for the steady-state mode (i.e., the system of equilibrium equations) in the form

$$\begin{aligned} \alpha p_0 &= \int_0^\infty p_1(x) \cdot \delta(x) dx, \\ \frac{dp_1(x)}{dx} &= -(\alpha + \delta(x)) \cdot p_1(x), \\ \frac{dp_2(x)}{dx} &= -(\alpha + \delta(x)) \cdot p_2(x) + \alpha \cdot p_1(x), \\ &\vdots \\ \frac{dp_{k-1}(x)}{dx} &= -(\alpha + \delta(x)) \cdot p_{k-1}(x) + \alpha \cdot p_{k-2}(x), \\ \frac{dp_k(x)}{dx} &= -\delta(x) \cdot p_k(x) + \alpha \cdot p_{k-1}(x), \\ p_1(0) &= \alpha \cdot p_0 + \int_0^\infty p_2(x) \cdot \delta(x) dx, \\ p_2(0) &= \int_0^\infty p_3(x) \cdot \delta(x) dx, \\ &\vdots \\ p_{k-1}(0) &= \int_0^\infty p_k(x) \cdot \delta(x) dx. \end{aligned} \quad (25)$$

As discussed previously, the established set of equations has an analytical closed-form solution, which we present in the following subsection for the special case of $k = 3$ redundant elements. This formulation, after the integration over the variable x , produces a result for the macrostate probabilities in the system. A macrostate probability p_i in the steady-state mode represents the limiting time-average fraction of time spent by the system in state i . Note that the sum $\sum_{i=0}^{k-1} p_i$ corresponds to the steady-state probability that at least one of k reserve (collaborating) elements is operational; that is, the system is functioning properly. Hence, in the general case we define the stationary reliability of the system as the steady-state probability of its failure-free operation, that is, $1 - p_k$.

Note that the established set of equations has an analytical closed-form solution for an arbitrary number of system elements. By increasing k , calculation of the undetermined coefficients becomes rather “bulky” while solving the corresponding inhomogeneous linear ordinary differential equations. In the following subsection, we consider an algorithm for solving the target system of differential equations in the special case.

4.2. Considerations for Target Scenario. To demonstrate the applicability of our general system modeling in the considered scenario, we further focus on the characteristic use case where a group of four police officers (one of them equipped with the AR glasses) are entering the building, while the squad leader wears a resource-hungry device that requires assistance from the devices of fellow officers. The squad moves along as a group but the connection to the assisting equipment may become unreliable (unavailable) due to unconstrained group mobility [63].

We begin with

$$\begin{aligned} \alpha p_0 &= \int_0^\infty p_1(x) \cdot \delta(x) dx, \\ \frac{dp_1(x)}{dx} &= -(\alpha + \delta(x)) \cdot p_1(x), \\ \frac{dp_2(x)}{dx} &= -(\alpha + \delta(x)) \cdot p_2(x) + \alpha \cdot p_1(x), \\ \frac{dp_3(x)}{dx} &= -\delta(x) \cdot p_3(x) + \alpha \cdot p_2(x), \\ p_1(0) &= \alpha \cdot p_0 + \int_0^\infty p_2(x) \cdot \delta(x) dx, \\ p_2(0) &= \int_0^\infty p_3(x) \cdot \delta(x) dx. \end{aligned} \quad (26)$$

Further, we focus on the solution obtained with Kolmogorov's system of differential equations. Let us initially consider the second equation

$$\begin{aligned} \frac{dp_1(x)}{dx} &= -(\alpha + \delta(x)) \cdot p_1(x), \\ \frac{dp_1(x)}{p_1(x)} &= -(\alpha + \delta(x)) dx, \end{aligned}$$

$$\begin{aligned} \frac{dp_1(x)}{p_1(x)} &= -\delta(x) dx - \alpha dx, \\ \frac{dp_1(x)}{p_1(x)} &= -\frac{b(x)}{1 - B(x)} dx - \alpha dx, \\ d \ln p_1(x) &= d \ln(1 - B(x)) + d \ln e^{-\alpha x}, \\ d \ln p_1(x) &= d \ln(1 - B(x)) \cdot e^{-\alpha x}, \\ \ln p_1(x) &= \ln(1 - B(x) \cdot e^{-\alpha x}) + \ln C_1, \\ \ln p_1(x) &= \ln(1 - B(x) \cdot e^{-\alpha x} \cdot C_1), \\ p_1(x) &= C_1 \cdot e^{-\alpha x} \cdot (1 - B(x)). \end{aligned} \quad (27)$$

We solve the differential equations for $p_2(x)$ and $p_3(x)$ by applying the parameters variation method.

$$\begin{aligned} p_2(x) &= (\alpha C_1 x + C_2) \cdot e^{-\alpha x} \cdot (1 - B(x)), \\ p_3(x) &= C_3 (1 - B(x)) - (C_2 + C_1 (\alpha x + 1)) \cdot e^{-\alpha x} \\ &\quad \cdot (1 - B(x)). \end{aligned} \quad (28)$$

To express the coefficients C_2 and C_3 in terms of C_1 , we utilize the boundary conditions

$$\begin{aligned} p_0 &= C_1 \cdot \alpha^{-1} \cdot \tilde{b}(\alpha), \\ p_1(x) &= C_1 \cdot e^{-\alpha x} \cdot (1 - B(x)), \\ p_2(x) &= C_1 \cdot e^{-\alpha x} \cdot (1 - B(x)) \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)}, \\ p_3(x) &= C_1 \cdot (1 - e^{-\alpha x}) \cdot (1 - B(x)) \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)}. \end{aligned} \quad (29)$$

Further, we evaluate the stationary probabilities for the macrostate p_1 .

$$\begin{aligned} p_1 &= \int_0^\infty p_1(x) dx = \int_0^\infty C_1 \cdot e^{-\alpha x} \cdot (1 - B(x)) dx \\ &= C_1 \cdot \left(\int_0^\infty e^{-\alpha x} dx - \int_0^\infty e^{-\alpha x} \cdot B(x) dx \right) \\ &= C_1 \cdot \left(\frac{1}{\alpha} - \frac{1}{\alpha} \tilde{b}(\alpha) \right) = C_1 \cdot \left(\frac{1 - \tilde{b}(\alpha)}{\alpha} \right), \end{aligned} \quad (30)$$

where $\tilde{b}(\alpha) = \alpha \tilde{B}(\alpha)$ is an LT for the PDF $b(x)$ at point $s = \alpha$. Next, for the probability p_2 we have

$$\begin{aligned} p_2 &= \int_0^\infty p_2(x) dx \\ &= \int_0^\infty C_1 \cdot e^{-\alpha x} \cdot (1 - B(x)) \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} dx \end{aligned}$$

$$\begin{aligned}
&= C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \int_0^\infty e^{-\alpha x} \cdot (1 - B(x)) dx \\
&= C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \left(\frac{1 - \tilde{b}(\alpha)}{\alpha} \right) \\
&= C_1 \cdot \frac{(1 - \tilde{b}(\alpha))^2}{\alpha \cdot \tilde{b}(\alpha)}. \tag{31}
\end{aligned}$$

Further, for p_3 we obtain

$$\begin{aligned}
p_3 &= \int_0^\infty p_3(x) dx = \int_0^\infty C_1 \cdot (1 - e^{-\alpha x}) \cdot (1 - B(x)) \\
&\quad \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} dx = C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \int_0^\infty (1 - e^{-\alpha x}) \\
&\quad \cdot (1 - B(x)) dx = C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \\
&\quad \times \left(\int_0^\infty (1 - B(x)) dx - \int_0^\infty e^{-\alpha x} dx \right. \\
&\quad \left. + \int_0^\infty e^{-\alpha x} \cdot B(x) dx \right) = C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \\
&\quad \cdot \left(- \int_0^\infty x \cdot (-b(x)) dx - \frac{1}{\alpha} + \frac{1}{\alpha} \tilde{b}(\alpha) \right) = C_1 \\
&\quad \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \left(\int_0^\infty x \cdot b(x) dx - \frac{1}{\alpha} + \frac{1}{\alpha} \tilde{b}(\alpha) \right) \\
&= C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \left(EB - \frac{1}{\alpha} + \frac{1}{\alpha} \tilde{b}(\alpha) \right) = C_1 \\
&\quad \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \left(b - \frac{1}{\alpha} + \frac{1}{\alpha} \tilde{b}(\alpha) \right) = C_1 \\
&\quad \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \left(b - \frac{1 - \tilde{b}(\alpha)}{\alpha} \right), \tag{32}
\end{aligned}$$

where b is the parameter of random “repair” time (mathematical expectation for the failed element).

Therefore, the probabilities of the system macrostates are given as

$$\begin{aligned}
p_0 &= C_1 \cdot \frac{\tilde{b}(\alpha)}{\alpha}, \\
p_1 &= C_1 \cdot \left(\frac{1 - \tilde{b}(\alpha)}{\alpha} \right), \\
p_2 &= C_1 \cdot \frac{(1 - \tilde{b}(\alpha))^2}{\alpha \cdot \tilde{b}(\alpha)},
\end{aligned}$$

$$p_3 = C_1 \cdot \frac{(1 - \tilde{b}(\alpha))}{\tilde{b}(\alpha)} \cdot \left(b - \frac{1 - \tilde{b}(\alpha)}{\alpha} \right). \tag{33}$$

By applying the normalization condition $p_0 + p_1 + p_2 + p_3 = 1$, we establish the constant C_1 as

$$C_1 = \frac{\alpha \cdot \tilde{b}(\alpha)}{\rho^{-1} (1 - \tilde{b}(\alpha)) + \tilde{b}(\alpha)}, \tag{34}$$

where $\rho^{-1} = EB/EA = b \cdot \alpha$.

Finally, we calculate the following steady-state probabilities for the considered “cold-standby repairable system” $\langle M_3/GI/1 \rangle$ as

$$\begin{aligned}
p_0 &= \frac{\tilde{b}^2(\alpha)}{\rho^{-1} (1 - \tilde{b}(\alpha)) + \tilde{b}(\alpha)}, \\
p_1 &= \frac{\tilde{b}(\alpha) (1 - \tilde{b}(\alpha))}{\rho^{-1} (1 - \tilde{b}(\alpha)) + \tilde{b}(\alpha)}, \\
p_2 &= \frac{(1 - \tilde{b}(\alpha))^2}{\rho^{-1} (1 - \tilde{b}(\alpha)) + \tilde{b}(\alpha)}, \\
p_3 &= \frac{(1 - \tilde{b}(\alpha)) (\rho^{-1} - 1 + \tilde{b}(\alpha))}{\rho^{-1} (1 - \tilde{b}(\alpha)) + \tilde{b}(\alpha)}, \tag{35}
\end{aligned}$$

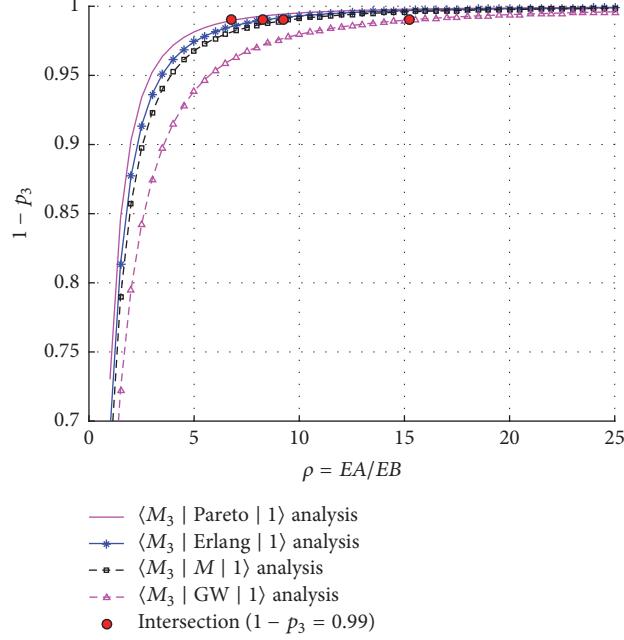
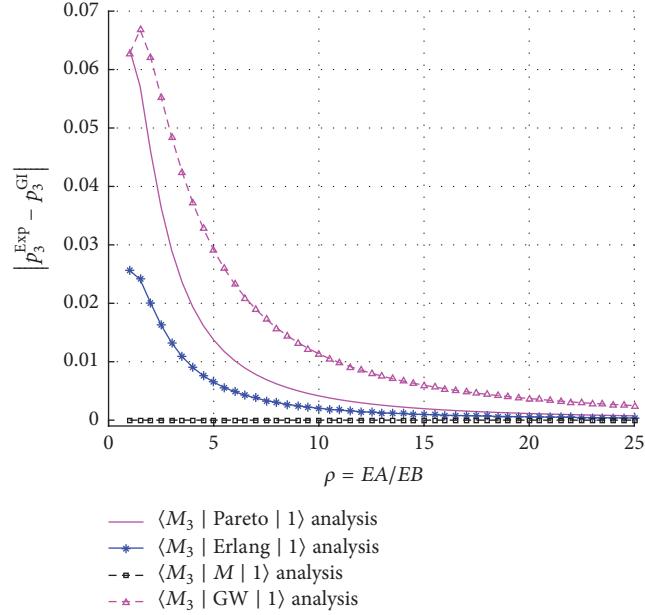
where $\rho^{-1} = b \cdot \alpha$.

As can be concluded from the produced expressions, there is a clear dependency of the system steady-state probabilities on the type of “repair” time distribution. However, this dependency becomes vanishingly small under the “fast repair” of failed elements. The numerical results are reported in the following section and they support this hypothesis.

5. Representative Numerical Results

This section summarizes selected numerical results obtained after applying our proposed analytical model to the scenario of interest described above. To this end, Figure 4 demonstrates the stationary system reliability $1 - p_3$ versus the relative recovery rate ρ for four different special cases of the $\langle M_3 | GI | 1 \rangle$ model, where GW stands for Gnedenko-Weibull distribution [64].

As it can be observed in the figure, the stationary reliability $1 - p_3$ becomes close to 1 as ρ tends to infinity irrespective of the shapes of the activity and repair time distributions. However, it reaches the high value of $1 - p_3 = 0.99$ at different values of ρ . This characterizes the sensitivity of the system stationary reliability with respect to the shape of the activity and repair time distributions (note different rates here). We can assess the rate of convergence of $1 - p_3$ with the means of quantiles for a given reliability level. The corresponding simulations have been conducted with ρ from 1 to 25 for 10,000 time slots over 200 executions, thus making it possible to reach the accuracy of 99.99%.

FIGURE 4: Analytical results for stationary system reliability $1 - p_3$ versus ρ for various “repair” time distributions.FIGURE 5: Absolute difference of values for system stationary reliability $1 - p_3$ versus ρ for various “repair” time distributions.

Further, the similarity measure (absolute difference) between the plots in Figure 4 is presented in Figure 5. The results support our above statement about high asymptotic insensitivity of the system stationary reliability, which is clearly observed due to the proximity of the corresponding curves. For instance, starting from $\rho = 20$, all of the curves are almost indistinguishable.

We also note that explicit analytical expressions for the steady-state distribution of the system under consideration are not always easy to obtain. For this case, we construct a simulation model that approximates the analytical results

for the system in question. We conduct our simulations by following the standard discrete event modeling approach. Previously discussed analytical values of the stationary system reliability $1 - p_3$ versus ρ (see Figure 4) are supported with simulated data, and the results are displayed in Figure 6. The differences between the produced curves become marginal very quickly. Even for the relatively small values of ρ , the stationary reliability of the system $1 - p_3$ is already very close to 1 for all the four cases.

Comparing the results obtained with both analytical and simulation approaches, we confirm that they are in close

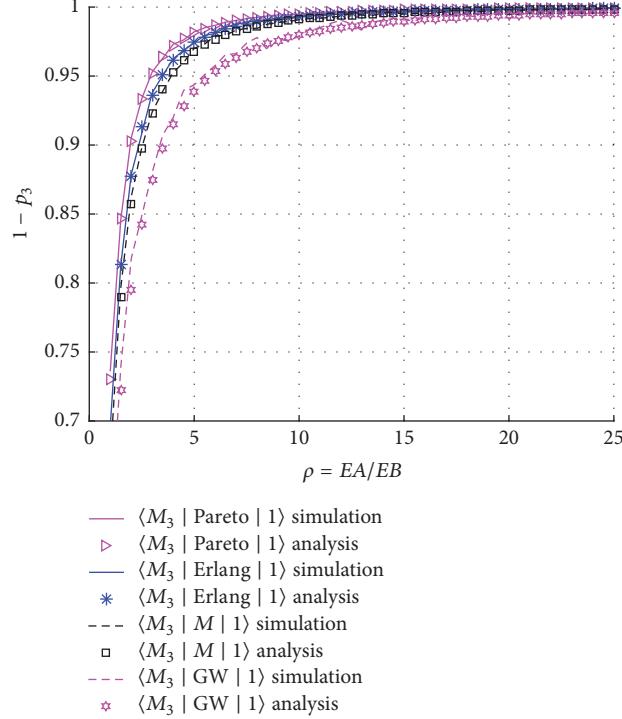


FIGURE 6: Analytical and simulation results of stationary system reliability $1 - p_3$ versus ρ (averaged values based on 200 replications) for various “repair” time distributions.

agreement. Hence, simulations can be employed in the cases where explicit analytical solution cannot be achieved due to its complexity or as part of a more complex simulation-based methodology that targets in-depth assessment of the fault-tolerant operation of MEC/MCC.

6. Main Conclusions

Today, mobile communications technology and resource-hungry applications are all over the market. As users expect to receive increasingly high levels of service experience, this introduces a burden on the existing wireless communications mechanisms. This work explores the challenge of delegating computing and caching functionality as the users move away from the network “center” to its edge, therefore supporting the emerging MEC paradigm.

With regard to the main contribution of this work, we propose a powerful mathematical methodology that allows assessing system-level reliability in cases where the “helper” nodes fail to provide reliable performance, thus allowing analysis of fault-tolerant MEC/MCC operation. The demonstrated modeling results reveal asymptotic insensitivity of the stationary reliability of the system under the “fast” recovery of its elements to the type of the “repair” time distribution.

In addition, the obtained steady-state stationary system distribution enables us to assess other operational variables, primarily with respect to the “helper” node performance, since the corresponding weighted coefficients may be added to the analysis readily. This permits characterizing the average stationary system performance where the assisting devices may belong to different classes.

Main System Model Notations

k :	Number of cooperating elements
$v(t)$:	Number of failed elements at time t
Δ :	Short time interval
A :	Time to failure (random variable)
α :	Failure rate of elements
$A(x) = 1 - e^{-\alpha x}, x \geq 0$:	Cumulative distribution function (CDF) of A
$a(x) = \alpha e^{-\alpha x}$:	Probability density function (PDF) of the random variable A
B :	Time to repair (random variable)
$B(x)$:	CDF of the random variable B
$b(x)$:	PDF of the random variable B
$\tilde{b}(s) = \int_0^{\infty} e^{-sx} b(x) dx$:	Laplace transform (LT) of the PDF $b(x)$
$EA = 1/\alpha$:	Mean time to failure of the main element
$b = EB$:	Mean time to repair the failed element
$\rho = EA/EB$:	Repair rate, relative recovery rate
$\delta(x) = b(x)/(1-B(x))$:	Conditional PDF of the residual repair time of the element, which has already spent time t in the repair state.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The publication was supported by the Ministry of Education and Science of the Russian Federation (Project no. 2.882.2017/4.6).

References

- [1] Cisco, *Global mobile data traffic forecast 2016–2021*, White Paper, 2017.
- [2] Ericsson, “Ericsson mobility report: On the pulse of the Networked Society,” <http://www.ericsson.com/mobility-report>, November 2017.
- [3] J. G. Andrews, S. Buzzi, and W. Choi, “What will 5G be?” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [4] J. Orlosky, K. Kiyokawa, and H. Takemura, “Virtual and augmented reality on the 5G highway,” *Journal of Information Processing*, vol. 25, pp. 133–141, 2017.
- [5] A. Orsino, A. Ometov, G. Fodor et al., “Effects of Heterogeneous Mobility on D2D- and Drone-Assisted Mission-Critical MTC in 5G,” *IEEE Communications Magazine*, vol. 55, no. 2, pp. 79–87, 2017.
- [6] S. Hamdoun, A. Rachedi, and Y. Ghamri-Doudane, “A flexible M2M radio resource sharing scheme in LTE networks within an H2H/M2M coexistence scenario,” in *Proceedings of the IEEE International Conference on Communications, ICC 2016*, Malaysia, May 2016.
- [7] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, “Radio resource allocation in LTE-advanced cellular networks with M2M communications,” *IEEE Communications Magazine*, vol. 50, no. 7, pp. 184–192, 2012.
- [8] “Smaller, faster, cheaper, over: the future of computer chips,” The New York Times, 2015: <https://www.nytimes.com/2015/09/27/technology/smaller-faster-cheaper-over-the-future-of-computer-chips.html?20r=0>.
- [9] “Why the wearable tech revolution has only just begun,” The Telegraph, 2016: <http://www.telegraph.co.uk/sponsored/technology/technology-trends/12044628/wearable-tech.html>.
- [10] M. Paterson and M. R. Glass, “The world through Glass: developing novel methods with wearable computing for urban videographic research,” *Journal of Geography in Higher Education*, vol. 39, no. 2, pp. 275–287, 2015.
- [11] R. Mahapatra, Y. Nijsure, G. Kaddoum, N. U. Hassan, and C. Yuen, “Energy efficiency tradeoff mechanism towards wireless green communication: a survey,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 686–705, 2016.
- [12] O. Galinina, H. Tabassum, K. Mikhaylov, S. Andreev, E. Hossein, and Y. Koucheryavy, “On feasibility of 5G-grade dedicated RF charging technology for wireless-powered wearables,” *IEEE Wireless Communications Magazine*, vol. 23, no. 2, pp. 28–37, 2016.
- [13] O. Galinina, A. Pyattaev, K. Johnsson, A. Turlikov, S. Andreev, and Y. Koucheryavy, “Assessing System-Level Energy Efficiency of mmWave-Based Wearable Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 923–937, 2016.
- [14] K. S. Deepak and A. V. Babu, “Improving energy efficiency of incremental relay based cooperative communications in wireless body area networks,” *International Journal of Communication Systems*, vol. 28, no. 1, pp. 91–111, 2013.
- [15] S. Ko, K. Huang, S. Kim, and H. Chae, “Live prefetching for mobile computation offloading,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3057–3071, 2017.
- [16] M. Peng, Y. Li, Z. Zhao, and C. Wang, “System architecture and key technologies for 5G heterogeneous cloud radio access networks,” *IEEE Network*, vol. 29, no. 2, pp. 6–14, 2015.
- [17] M. Peng, S. Yan, K. Zhang, and C. Wang, “Fog-computing-based radio access networks: Issues and challenges,” *IEEE Network*, vol. 30, no. 4, pp. 46–53, 2016.
- [18] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, “A survey on mobile edge computing: the communication perspective,” <https://arxiv.org/abs/1701.01090>.
- [19] C. You, K. Huang, H. Chae, and B.-H. Kim, “Energy-Efficient Resource Allocation for Mobile-Edge Computation Offloading,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1397–1411, 2017.
- [20] J. O. Fajardo, I. Taboada, and F. Liberal, “Radio-aware service-level scheduling to minimize downlink traffic delay through mobile edge computing,” *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNCS*, vol. 158, pp. 121–134, 2015.
- [21] M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, A. Neal et al., “Mobile-edge computing introductory technical white paper,” White Paper, Mobile-edge Computing (MEC) industry initiative, 2014.
- [22] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, “Mobile edge computing: A key technology towards 5G,” *ETSI White Paper*, vol. 11, 2015.
- [23] L. Person and P. Gotsur, “World LTE Market – Opportunities and Forecasts, 2012–2020,” Allied Market Research IC 14144, April 2014, p. 107.
- [24] M. Zhang, H. Luo, and H. Zhang, “A survey of caching mechanisms in information-centric networking,” *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1473–1499, 2015.
- [25] D. López-Pérez, I. Güvenç, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, “Enhanced inter-cell interference coordination challenges in heterogeneous networks,” *IEEE Wireless Communications Magazine*, vol. 18, no. 3, pp. 22–30, 2011.
- [26] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, “A survey of software-defined networking: past, present, and future of programmable networks,” *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1617–1634, 2014.
- [27] M. Kountouris and N. Pappas, “HetNets and massive MIMO: Modeling, potential gains, and performance analysis,” in *Proceedings of the 3rd IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications, IEEE APWC 2013*, pp. 1319–1322, Italy, September 2013.
- [28] H. Zhang, X. Chu, W. Guo, and S. Wang, “Coexistence of Wi-Fi and heterogeneous small cell networks sharing unlicensed spectrum,” *IEEE Communications Magazine*, vol. 53, no. 3, pp. 158–164, 2015.
- [29] S. Tombaz, P. Monti, F. Farias, M. Fiorani, L. Wosinska, and J. Zander, “Is backhaul becoming a bottleneck for green wireless access networks?” in *Proceedings of the 1st IEEE International Conference on Communications, ICC 2014*, pp. 4029–4035, Australia, June 2014.
- [30] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. C. M. Leung, “Cache in the air: exploiting content caching and delivery techniques for 5G systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 131–139, 2014.

- [31] X. Wang, X. Li, V. C. M. Leung, and P. Nasiopoulos, "A framework of cooperative cell caching for the future mobile networks," in *Proceedings of the 48th Annual Hawaii International Conference on System Sciences, HICSS 2015*, pp. 5404–5413, USA, January 2015.
- [32] Z. Chen and M. Kountouris, "D2D caching vs. small cell caching: Where to cache content in a wireless network?" in *Proceedings of the 17th IEEE International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2016*, UK, July 2016.
- [33] R. Wang, J. Zhang, S. H. Song, and K. B. Letaief, "Mobility-Aware Caching in D2D Networks," *IEEE Transactions on Wireless Communications*, no. 8, pp. 5001–5015, 2017.
- [34] "Americans spend an alarming amount of time checking social media on their phones," Digital Trends, 2015: <http://www.digitaltrends.com/mobile/informate-report-social-media-smartphone-use/>.
- [35] R. Fedorov, D. Frajberg, and P. Fraternali, "A framework for outdoor mobile augmented reality and its application to mountain peak detection," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 9768, pp. 281–301, 2016.
- [36] M. Mirahsan, R. Schoenen, and H. Yanikomeroglu, "HetHET-Nets: Heterogeneous Traffic Distribution in Heterogeneous Wireless Cellular Networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2252–2265, 2015.
- [37] "Augmented reality hardware is still way ahead of its software," Engadget, 2016: <https://www.engadget.com/2016/06/02/augmented-reality-hardware-and-software-still-needs-work/>.
- [38] "Rivvr brings wireless VR to the oculus rift and HTC vive," TechCrunch, 2016: <https://techcrunch.com/2016/12/15/rivvr-brings-wireless-vr-to-the-oculus-rift-and-htc-vive/>.
- [39] A. Y. C. Nee, S. K. Ong, G. Chryssolouris, and D. Mourtzis, "Augmented reality applications in design and manufacturing," *CIRP Annals - Manufacturing Technology*, vol. 61, no. 2, pp. 657–679, 2012.
- [40] M. D. Dikaiakos, D. Katsaros, P. Mehra, G. Pallis, and A. Vakali, "Cloud computing: distributed Internet computing for IT and scientific research," *IEEE Internet Computing*, vol. 13, no. 5, pp. 10–11, 2009.
- [41] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges," *Future Generation Computer Systems*, 2016.
- [42] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the Internet of Things," in *Proceedings of the 1st ACM Mobile Cloud Computing Workshop, MCC 2012*, pp. 13–15, Finland, August 2012.
- [43] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: Architecture, applications, and approaches," *Wireless Communications and Mobile Computing*, vol. 13, no. 18, pp. 1587–1611, 2013.
- [44] K. Wang, K. Yang, and C. Magurawalage, "Joint Energy minimization and resource allocation in C-RAN with mobile cloud," *IEEE Transactions on Cloud Computing*, 2016.
- [45] K. Wang, K. Yang, X. Wang, and C. S. Magurawalage, "Cost-effective resource allocation in C-RAN with mobile cloud," in *Proceedings of the IEEE International Conference on Communications, ICC 2016*, Malaysia, May 2016.
- [46] P. Bahl, R. Y. Han, L. E. Li, and M. Satyanarayanan, "Advancing the state of mobile cloud computing," in *Proceedings of the 3rd ACM Workshop on Mobile Cloud Computing and Services (MCS)*, pp. 21–28, Lake District, UK, June 2012.
- [47] S. Abolfazli, Z. Sanaei, E. Ahmed, A. Gani, and R. Buyya, "Cloud-based augmentation for mobile devices: motivation, taxonomies, and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 337–368, 2014.
- [48] K. Gai, M. Qiu, and H. Zhao, "Cost-aware multimedia data allocation for heterogeneous memory using genetic algorithm in cloud computing," *IEEE Transactions on Cloud Computing*, 2016.
- [49] S. Andreev, O. Galinina, A. Pyattaev et al., "Exploring synergy between communications, caching, and computing in 5G-grade deployments," *IEEE Communications Magazine*, vol. 54, no. 8, pp. 60–69, 2016.
- [50] N. Golrezaei, A. Molisch, A. G. Dimakis, and G. Caire, "Femto-caching and device-to-device collaboration: A new architecture for wireless video distribution," *IEEE Communications Magazine*, vol. 51, no. 4, pp. 142–149, 2013.
- [51] B. Zhou, Y. Cui, and M. Tao, "Stochastic content-centric multicast scheduling for cache-enabled heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6284–6297, 2016.
- [52] Y. Wang, A. Vasilakos, Q. Jin, and J. Ma, "Survey on mobile social networking in proximity (MSNP): Approaches, challenges and architecture," *Wireless Networks*, vol. 20, no. 6, pp. 1295–1311, 2014.
- [53] J. Hosek, M. Ries, P. Vajsar et al., "User's happiness in numbers: Understanding mobile YouTube quality expectations," in *Proceedings of the 38th International Conference on Telecommunications and Signal Processing, TSP 2015*, pp. 607–611, Czech Republic, July 2015.
- [54] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: the role of proactive caching in 5G wireless networks," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 82–89, 2014.
- [55] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5G networks: how location information can improve scalability, latency, and robustness of 5G," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112, 2014.
- [56] Y. Wang, L. Wei, A. V. Vasilakos, and Q. Jin, "Device-to-Device based mobile social networking in proximity (MSNP) on smartphones: Framework, challenges and prototype," *Future Generation Computer Systems*, 2015.
- [57] N. Golrezaei, P. Mansourifard, A. F. Molisch, and A. G. Dimakis, "Base-station assisted device-to-device communications for high-throughput wireless video networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3665–3676, 2014.
- [58] J. Kuula, P. Kettunen, V. Auvinen, S. Viitanen, O. Kauppinen, and T. Korhonen, "Smartphones as an alerting, command and control system for the preparedness groups and civilians: Results of preliminary tests with the Finnish police," in *Proceedings of the 10th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2013*, pp. 42–51, May 2013.
- [59] H. Wang, D. Peng, W. Wang, H. Sharif, H.-H. Chen, and A. Khoynezhad, "Resource-aware secure ECG healthcare monitoring through body sensor networks," *IEEE Wireless Communications Magazine*, vol. 17, no. 1, pp. 12–19, 2010.
- [60] T. Ma, P. L. Shrestha, M. Hempel, D. Peng, H. Sharif, and H.-H. Chen, "Assurance of energy efficiency and data security for ECG transmission in BASNs," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 1041–1048, 2012.

- [61] "Tampere traffic wardens work quicker thanks to Google Glass app," YLE UUTISET, 2016: http://yle.fi/uutiset/osasto/news/tampere_traffic_wardens_work_quicker_thanks_to_google_glass_app/8766025.
- [62] "2017 police vehicles article," Government Fleet-Feature, 2016: <http://www.government-fleet.com/article/story/2016/11/2017-police-vehicles.aspx>.
- [63] A. Orsino, D. Moltchanov, M. Gapeyenko et al., "Direct Connection on the Move: Characterization of User Mobility in Cellular-Assisted D2D Systems," *IEEE Vehicular Technology Magazine*, vol. 11, no. 3, pp. 38–48, 2016.
- [64] B. V. Gnedenko, Y. K. Belyayev, and A. D. Solov'yev, *Mathematical Methods of Reliability Theory*, Academic Press, 2014.