Stochastic Geometry-Based Modeling and Analysis of Beam Management in 5G

Sanket S. Kalamkar,[1] Fuad M. Abinader Jr.,[2] François Baccelli,[1] Andrea S. Marcano Fani,[2] and Luis G. Uzeda Garcia[2]
Affiliation: [1]INRIA-ENS, Paris, France, [2]Nokia Bell Labs, Paris, France

***Abstract*—Beam management is central in the operation of dense 5G cellular networks. Focusing the energy radiated to mobile terminals (MTs) by increasing the number of beams per cell increases signal power and decreases interference, and has hence the potential to bring major improvements on area spectral efficiency (ASE). This benefit, however, comes with unavoidable overheads that increase with the number of beams and the MT speed. This paper proposes a first system-level stochastic geometry model encompassing major aspects of the beam management problem: frequencies, antennas, and propagation; physical layer, wireless links, and coding; network geometry, interference, and resource sharing; sensing, signaling, and mobility management. This model leads to a simple analytical expression for the effective ASE that the typical user gets in this context. This in turn allows one to find, for a wide variety of 5G network scenarios including millimeter wave (mmWave) and sub-6 GHz, the number of beams per cell that offers the best global trade-off between these benefits and costs. We finally provide numerical results that discuss the effects of different systemic trade-offs and performances of mmWave and sub-6 GHz 5G deployments.**

I. Introduction

A. Motivation

The ever-increasing demand in capacity for mobile communications makes it necessary to consider new implementation approaches/techniques that can significantly boost data rates and the area spectral efficiency (ASE) of mobile networks. One key enabler considered in 5G [1] to face this demand is the use of the spectrum beyond the sub-6 GHz frequencies, known as millimeter wave (mmWave). More specifically, 5G relies on the spectrum above 20 GHz, where bandwidth of up to 400 MHz can be used to offer very high data rates (above 10 Gbps peak rates) and increase the network capacity; nevertheless, the sub-6 GHz bands, with up to 100 MHz of bandwidth, are still needed to ensure wide area coverage and data rates up to a few Gbps.

One of the key difficulties that mmWave frequencies face is their challenging propagation characteristics: they are subject to high path loss, penetration loss, and diffraction loss due to their millimetric wavelength, thus reaching short distances typically within a few hundred meters. But what was once considered a limitation makes nowadays mmWave a suitable candidate for small cells deployments, which can be used for network densification and capacity boosting.

To overcome the propagation challenges at mmWave frequencies, steerable arrays with a large number of antenna elements are used to create highly directional beams that concentrate the transmitted energy to achieve high gains and make the signal more robust to increase coverage. Moreover, mmWave communications must be designed to operate under mobility conditions, covering users in LOS (Line-of-Sight) and NLOS (Non-LOS) at pedestrian and vehicular speeds. This can become quite challenging since mmWave frequencies are highly sensitive to changes in the environment. Thus, any mmWave-based system relies on beam management techniques to select the best beam during a base station (inter-cell) handover and quickly switch/reselect a new beam during intra-cell mobility to avoid beam misalignment and performance losses. Although more important for mmWave frequencies, beamforming techniques with highly directional narrow beams can also be used in sub-6 GHz frequencies to enhance the network performance.

Nevertheless, beam-based communications come with implementation challenges, beam management procedures such as beam refinement and beam failure detection and recovery that introduce different signaling and delay overheads. Also, an efficient beam management relies on capturing systemic trade-offs that depend on the mobile terminal (MT) mobility, cell sizes, and the number/width of the beams. For instance, a large number of narrow beams improves the signal-to-interference-plus-noise ratio (SINR), but it also leads to more frequent service interruption due to beam switching and beam misalignment, which degrade the network performance. Given such a complex nature of beam management, the system-level simulations, although crucial, are often time-consuming and expensive. Hence, to complement system-level simulations, in this paper, we aim to provide a tractable mathematical framework that permits a system-level analysis of beam management in 5G.

B. Contributions

1) We provide a first mathematical optimization framework for the beam management problem in 5G. Using the well-established tools from stochastic geometry [2], [3], our proposed optimization yields the number of beams that maximizes the area spectral efficiency.
2) The optimization permits a system-level analysis in 5G. In particular, it takes into account the average speed of the MTs, mobility-induced beam-misalignment error, beam selections during base station handovers and beam reselections within a cell, and time overheads associated

E-mail: sanket.kalamkar@inria.fr, fuad.abinader@nokia-bell-labs.com

with these beam (re)selections. For instance, given an average speed of the mix of MTs (pedestrians, bikes, and cars), the mix of geometries (some cells are bigger, other smaller, with the MTs either close or far from the serving base station), the mix for fading and blockages present in the network, our optimization gives the number of beams that provides the best "sum rate" for MTs of all types in a large ball.

3) While the proposed framework is generic and covers a variety of related problems, e.g., other beam-centric systems, we give numerical results to evaluate a 5G New Radio (NR)-compliant radio access network operating in a dense urban macro/picocell scenario, for both sub-6 GHz and mmWave frequencies. The results reveal the key inter-dependencies between network parameters and provide insights into the associated trade-offs.

C. Related Work

The study of user handovers in cellular networks and their effect on various performance metrics such as throughput is a rich area (see e.g., [4]–[6] and the references therein). Most of the works have focused on only base station handovers (or, equivalently, cell handovers) in cellular networks [7]–[12]. The work in [13] is probably the closest one to our work. This work studies both initial beam selections during base station handovers and beam reselections within a cell and their associated overheads in a beam-centric mmWave cellular network. In particular, authors in [13] have obtained analytical expressions of inter-cell and inter-beam handover rates, based on which, the average spectral efficiency is calculated subject to overheads due to handovers. But, the work in [13] considers only noise-limited scenario and ignores interference from other base stations, while our work considers both noise and interference in the analysis. Second, our work also studies the effect of handovers on optimizing the number of beams, while [13] assumes a fixed number of beams, i.e., no beam management. Third, unlike [13], we consider blockages and beam misalignment due to mobility. Overall, there is a lack of understanding on how the MT mobility affects the beam management in the presence of interference, blockages, and beam misalignment error.

II. System Model

A. Network Model

As shown in Fig. 1, we consider a downlink cellular network, where the base station (BS) locations are modeled as a homogeneous Poisson point process (PPP) $\Phi \subset \mathbb{R}^2$ with intensity λ. We assume that the omnidirectional MT moves on a randomly oriented straight line with speed v. Without loss of generality, thanks to the isotropy and the stationarity of the PPP [2], this line of MT motion can be considered to be along the X-axis and passing through the origin. We assume that each BS always has an MT to serve. Also, a BS serves one MT at a time per resource block. The MT associates itself with the nearest BS. Such an association results in BS cells forming a Poisson Voronoi tessellation as shown in Fig. 1. Let

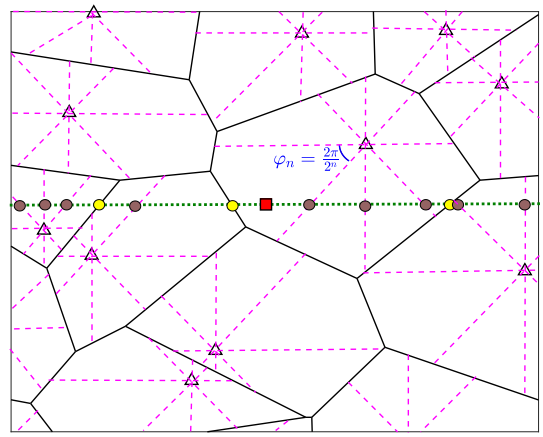

Fig. 1. A snapshot of a Poisson cellular network with directional beamforming to an MT. In this case, each BS has $2^3 = 8$ beams, i.e., $n = 3$. $\triangle$: Base station (BS), red square : MT location, brown filled circle : Beam reselection location, yellow filled circle : BS handover location, solid black lines : Cell boundaries, dashed lines : Beam boundaries, and dotted line : Path of the MT.

$X_0 \in \Phi$ denote the location of the serving BS of the MT at a given time. Without loss of generality, we can focus on the MT that is located at the origin at that time. After averaging over the PPP, this MT becomes the *typical* MT.

B. Beamforming Model

A BS at location $X \in \Phi$ uses directional beamforming to communicate with the typical MT. As shown in Fig. 1, we approximate the actual antenna pattern by a sectorized one, where each sector corresponds to one beam of the BS. For simplicity, we assume that each BS has 2^n beams with $n \in \mathbb{N}$, which corresponds to 2^{n-1} beam boundaries.[1] Hence, the beamwidth is

$$\varphi_n = \frac{2\pi}{2^n} = \frac{\pi}{2^{n-1}}. \tag{1}$$

We focus on a simple antenna pattern model where the main lobe is restricted to the beamwidth. Both the main and side lobe gains depend on the number 2^n of beams. In particular, the antenna gain G_n is expressed as

$$G_n(\psi) = \begin{cases} G_{\mathrm{m},n} & \text{if } |\psi| \leq \varphi_n/2 \\ G_{\mathrm{s},n} & \text{otherwise,} \end{cases} \tag{2}$$

where $G_{\mathrm{m},n}$ and $G_{\mathrm{s},n}$ denote the antenna gains within the main lobe and the side lobe, respectively, as a function of the number 2^n of beams, and ψ is the angle off the boresight direction.

[1]Here, a beam boundary is a line segment that connects two points on the cell boundary and passes through the location of the BS, as shown in Fig. 1.

C. Mobility-induced Beam Misalignment

In 5G NR, the beam (re)selection occurs during a synchronization signal burst (SSB) [14], which is done periodically with period τ. During an SSB, the beam that has the MT within its beamwidth is selected for communication. We call this beam the *reference beam*. Due to the mobility of the MT, it is possible that the MT has moved out of the beamwidth of the reference beam without selecting the new beam, and is still connected to the reference beam until the next SSB when the new reference beam is selected. Such a mobility-induced beam misalignment reduces the signal strength at the MT as it then lies within the side lobe of the reference beam chosen during a previous SSB. At a given time, the beam misalignment probability $p_{\rm bm}$ depends on the speed v of the MT, the duration τ between two SSBs, and the average distance between two beam reselections. Here, we propose a simple yet effective formula that captures the effects of these parameters on the probability of beam misalignment due to MT mobility in a snapshot of the network. We evaluate the probability $p_{\rm bm}$ that the MT is outside the beamwidth of the reference beam by the:

$$p_{\rm bm} = 1 - \exp\left(-\frac{v\tau}{1/\mu_{\rm s,b}}\right), \tag{3}$$

where $1/\mu_{\rm s,b}$ is the average distance between two consecutive beam reselections, with $\mu_{\rm s,b}$ being the linear intensity of beam reselection given in Theorem 2. We can interpret (3) as

$$1 - p_{\rm bm} = \mathbb{P}(T > \tau),$$

where T is an exponential random variable with mean $\frac{1}{v\mu_{\rm s,b}}$, with T interpreted as the time between two consecutive SSBs. Note that $\frac{1}{v\mu_{\rm s,b}}$ is the average time between two consecutive beam reselections. Hence, the average time between two SSBs is equal to the average time between two consecutive beam reselections.

Based on the probability of the beam misalignment, the antenna gain G_0 of the serving BS is given as

$$G_0 = \begin{cases} G_{{\rm m},n} & \text{w.p. } 1 - p_{\rm bm} \\ G_{{\rm s},n} & \text{w.p. } p_{\rm bm}. \end{cases} \tag{4}$$

D. Blockage and Channel Model

The presence of obstacles leads to LOS and NLOS propagation between the typical MT and a BS. We adopt a LOS ball model [15] to capture the effect of blockages. In particular, the propagation between a BS and the typical MT separated by distance d is LOS if $d < R_{\rm c}$ where $R_{\rm c}$ is the maximum distance for LOS propagation. The LOS and NLOS channel conditions induced by the blockage effect are characterized by different path loss exponents, denoted by $\alpha_{\rm L}$ and $\alpha_{\rm N}$, respectively. As shown in [16], typical values of these path loss exponents are $\alpha_{\rm L} \in [1.8, 2.5]$ and $\alpha_{\rm N} \in [2.5, 4.7]$.

The channel follows Rayleigh fading with unit mean power gain. Let h_X denote the channel power gain from the BS at location $X \in \Phi$ to the MT.[2] Note that, the random variables h_X are i.i.d. exponential with unit mean, i.e., $h_X \sim \exp(1)$. Let $|X|$ denote the distance between a BS at $X \in \Phi$ and the typical MT located at the origin. We consider the standard power-law path loss model with path loss function as

$$l(X) = \begin{cases} K|X|^{-\alpha_{\rm L}} & \text{if } |X| < R_{\rm c} \\ K|X|^{-\alpha_{\rm N}} & \text{if } |X| \geq R_{\rm c}, \end{cases} \tag{5}$$

where $K = \left(\frac{c}{4\pi f_{\rm c}}\right)^2$ is a frequency-dependent constant with c being the speed of light and $f_{\rm c}$ the carrier frequency. We assume that the typical MT can receive from its serving BS (also the nearest BS) in both LOS and NLOS conditions.[3]

E. Signal-to-Interference-Noise Ratio (SINR)

When the typical MT is associated with the BS located at $X_0 \in \Phi$ with antenna gain G_0 as in (4), the SINR at the typical MT is given by

$$\mathsf{SINR}_n = \frac{PG_0 h_{X_0} l(|X_0|)}{\sigma^2 + I_n}, \tag{6}$$

where P is the transmit power of a BS. Also, $\sigma^2 = WN_0$ is the noise power, where W and N_0 are the bandwidth and the noise spectral density, respectively. In the denominator of (6), I_n is the interference power at the typical MT given by

$$I_n = \sum_{X \in \Phi \setminus \{X_0\}} PG_n h_X l(|X|). \tag{7}$$

Since the beams of all BSs are oriented towards their respective MTs, the direction of arrivals between interfering BSs and the typical MT is distributed uniformly in $[-\pi, \pi]$. Thus, the gain G_n of an interfering BS is equal to $G_{{\rm m},n}$ with probability $\varphi_n/2\pi$ and $G_{{\rm s},n}$ with probability $1 - \varphi_n/2\pi$, where φ_n is the beamwidth given by (1).

III. Ergodic Shannon Rate

We are interested in calculating the downlink ergodic Shannon rate at the typical MT, which is given as

$$\mathcal{R}_n = W\mathbb{E}\left[\log\left(1 + \mathsf{SINR}_n\right)\right], \tag{8}$$

where $\mathbb{E}(\cdot)$ denotes expectation.

Definition 1 (Success Probability): The success probability $p_{\rm s}$ of the typical MT is the probability that the SINR at the typical MT exceeds a predefined threshold. Mathematically,

$$p_{\rm s}(n, \beta) \triangleq \mathbb{P}(\mathsf{SINR}_n > \beta), \tag{9}$$

where $\beta > 0$ is the predefined SINR threshold, which also parametrizes the transmission rate.

Lemma 1: Let

$$F(\alpha_{\rm S}, \alpha_{\rm I}, w) \triangleq 2\pi\lambda\left(1 - \frac{p_{I,n}}{1 + \frac{\beta r^{\alpha_{\rm S}} G_{{\rm m},n}}{G_0 w^{\alpha_{\rm I}}}} - \frac{1 - p_{I,n}}{1 + \frac{\beta r^{\alpha_{\rm S}} G_{{\rm s},n}}{G_0 w^{\alpha_{\rm I}}}}\right) w,$$

[2] Although the LOS channels are better modeled by Nakagami-m fading, Rayleigh fading allows us much better analytical tractability.

[3] The assumption of the nearest-BS association irrespective of LOS/NLOS conditions of the nearest BS is due to analytical tractability.

where $p_{I,n} = \varphi_n/2\pi$. The success probability $p_{\rm s}(n,\beta)$ is

$$p_{\rm s}(n,\beta) = (1-p_{\rm bm})q_{\rm s}(n,\beta,G_{{\rm m},n}) + p_{\rm bm}q_{\rm s}(n,\beta,G_{{\rm s},n}), \tag{10}$$

where $p_{\rm bm}$ is given by (3) and

$$q_{\rm s}(n,\beta,G_0) = \int_0^{R_{\rm c}} f_R(r)\exp\left(-\frac{\beta r^{\alpha_{\rm L}}\sigma^2}{PKG_0}\right) \times \exp\left(-\int_r^{R_{\rm c}} F(\alpha_{\rm L},\alpha_{\rm L},w){\rm d}w - \int_{R_{\rm c}}^{\infty} F(\alpha_{\rm L},\alpha_{\rm N},w){\rm d}w\right){\rm d}r + \int_{R_{\rm c}}^{\infty}\exp\left(-\frac{\beta r^{\alpha_{\rm N}}\sigma^2}{PKG_0} - \int_r^{\infty} F(\alpha_{\rm N},\alpha_{\rm N},w){\rm d}w\right) f_R(r){\rm d}r \tag{11}$$

with $f_R(r) = 2\pi\lambda r e^{-\lambda\pi r^2}$ where R is the distance to the nearest BS.

Proof: The proof is given in Appendix A. ■

Theorem 1: Imposing the limit $Q_{\max}$ on the maximum achievable SINR stemming from RF imperfections and modulation schemes, the ergodic Shannon rate per unit time is

$$\mathcal{R}_n = W\int_0^{Q_{\max}} \frac{p_{\rm s}(n,z)}{z+1}\,{\rm d}z, \tag{12}$$

where $p_{\rm s}(n,z)$ is given by (11).

Proof: The proof follows directly from

$$\mathcal{R}_n = W\mathbb{E}[\log(1+\min(\mathsf{SINR}_n, Q_{\max}))].$$

■

IV. Beam (Re)Selection

A. Beam Selection during BS handovers

When the MT performs a BS handover, it selects a beam with the new BS. We know from [17] that, for the Poisson-Voronoi tessellation, the linear intensity of *cell* boundary crossings, i.e., BS handovers, is $\mu_{\rm s,c} = \frac{4\sqrt{\lambda}}{\pi}$. Hence, the time intensity of BS handover (or, equivalently, beam selection) is

$$\mu_{\rm c} = \frac{4\sqrt{\lambda}}{\pi}v. \tag{13}$$

B. Beam Reselection

The beam reselection occurs at a beam boundary within the Voronoi cell of a BS. In Fig. 1, the locations of beam reselections are denoted by brown filled circles. We are interested in calculating the average number of beam reselections the typical MT performs per unit length. We also calculate the time intensity of beam reselection, i.e., the average rate of beam reselections.

Theorem 2 (Intensity of beam reselection): For 2^n beams and PPP of intensity λ, the linear intensity $\mu_{\rm s,b}$ of beam reselection for the typical MT moving on a straight line with speed v is $\frac{2^n\sqrt{\lambda}}{\pi}$. The time intensity $\mu_{\rm t,b}$ of beam reselection is $\mu_{\rm t,b}(n) = \frac{2^n\sqrt{\lambda}}{\pi}v$.

Proof: The proof is given in Appendix B. ■

Remark 1: When taken into account the average number of beam reselections that are skipped between two consecutive SSBs, the effective time intensity of beam reselection becomes

$$\mu_{\rm b}(n) = \frac{1}{\max\left(\tau, \frac{1}{\mu_{\rm t,b}(n)}\right)}, \tag{14}$$

TABLE I
Values of network parameters [20]

Parameter	FR1	FR2
Carrier frequency ($f_{\rm c}$)	3.5 GHz	28 GHz
Bandwidth (W)	100 MHz	400 MHz
Noise density (N_0)	-174 dBm/Hz	-174 dBm/Hz
Transmit power (P)	43 dBm	36 dBm
Beam reselection overhead ($T_{\rm b}$)	23 ms	23 ms
Cell handover overhead ($T_{\rm c}$)	43 ms	43 ms
SSB periodicity (τ)	20 ms	20 ms
MT speed (v)	$[3, 30, 120]$ km/h	$[3, 30]$ km/h
Inter-site distance (ISD)	$[250, 500, 1000]$ m	$[75, 125, 250]$ m
Maximum SINR ($Q_{\max}$)	30 dB	30 dB
Path loss exponent (α)	3.5	$\alpha_{\rm L} = 1.9$, $\alpha_{\rm N} = 3.5$
Blockage model	Implicit (NLOS)	LOS ball
LOS ball radius ($R_{\rm c}$)	–	75 m

where τ is the SSB periodicity.

V. Time Overhead due to Beam (Re)Selection

BS handovers and beam reselections may result in significant overheads in terms of time as a consequence of the time spent in beam sweeping and alignment, respectively. Such an overhead reduces the time available for data transmissions, in turn reducing the ergodic Shannon rate.

The typical MT moves on a straight line across different beams within the Voronoi cell of a BS as well as across Voronoi cells of different BSs. Hence, the two components that contribute to the time overhead are:

1) The time $T_{\rm c}$ for beam sweeping after each BS handover, i.e., *cell* boundary crossing, which includes the periodic SSB measurement, receiver processing time for the SSBs, and the handover interruption time due to cell switch and radio resource control (RRC) reconfiguration.
2) The time $T_{\rm b}$ for beam alignment after each beam reselection within the Voronoi cell of a BS, which includes the periodic SSB measurement and the receiver processing time for the SSBs.

Thus, the total average overhead per unit time is

$$T_{\rm o}(n) = \mu_{\rm b}(n)T_{\rm b} + \mu_{\rm c}T_{\rm c}. \tag{15}$$

The effective area spectral efficiency (ASE) per unit time is

$$\mathcal{R}_{\rm eff}(n) = \lambda\mathcal{R}_n(1-T_{\rm o}(n))^+, \tag{16}$$

where $\mathcal{R}_n$ is given by (12) and $(A)^+ = \max(0, A)$.

Our objective is to find the integer n that maximizes the effective ASE per unit time, i.e.,

$$n_* = \arg\max_{n\in\mathbb{N}}\ \mathcal{R}_{\rm eff}(n). \tag{17}$$

The value of n_* can easily be found by a linear search.

VI. Numerical Results and Discussions

To validate our proposed model, we consider a 5G NR-compliant radio access network operating in a dense urban macro/pico cell scenario. A summary of the model parameters for FR1 (sub-6 GHz) and FR2 (above 6 GHz) network deployments is provided in Table I.

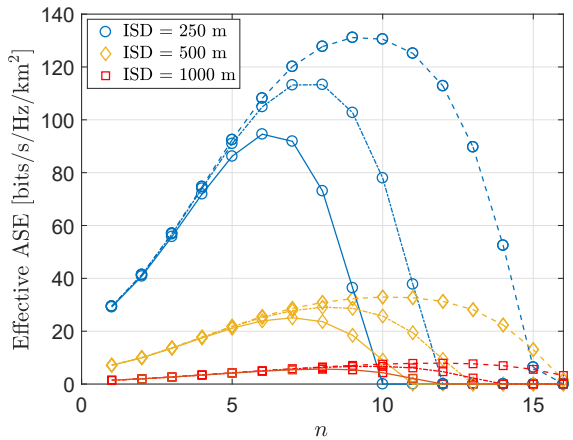

Fig. 2. **FR1**: Effect of average inter-site distance (ISD) on the effective ASE for different n. Dashed line: $v = 3$ km/h, Dashed-dotted line: $v = 30$ km/h, and Solid line: $v = 120$ km/h.

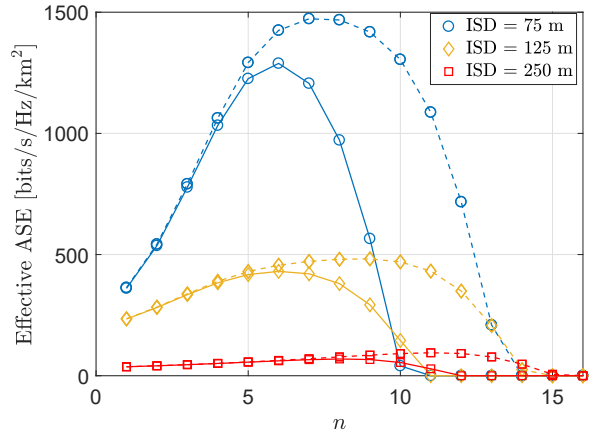

Fig. 3. **FR2**: Effect of average inter-site distance (ISD) on the effective ASE for different n. Dashed line: $v = 3$ km/h and Solid line: $v = 30$ km/h.

To explore the best potential of 5G NR networks in operational bands FR1 and FR2, we use the maximum bandwidth allowed as per 5G NR Release 15 [18], [19], namely 100 MHz for FR1 and 400 MHz for FR2. As the inter-site (or, equivalently inter-base station) distances (ISDs) for FR1 are expected to be larger due to lower frequency-dependent path loss, we choose a transmission power P = 43 dBm. For FR2, we can expect smaller ISDs due to (a) higher attenuation loss at higher carrier frequencies, and (b) because massive MIMO and high beamforming gains imposes challenges in terms of RF exposure and EMF limitations. Thus we decrease P to 36 dBm.

One important network planning decision is related to the number of beams per BS. In our proposed model, this is related to the choice of the values for n. The main and side lobe gains depend on the number 2^n of beams. In particular, with the increase in the number of beams, the main lobe gain increases while the side lobe gain decreases. Thus, without loss of generality, we assume the following antenna gains. The main lobe gain is $G_{\mathrm{m},n} = 2^n$, while the side lobe gain is $G_{\mathrm{s},n} = \frac{1}{2^n}$.[4]

Another important parameter is the intensity λ of BSs, which is directly related to the average cell size and the ISD. For an intensity λ of BSs, the average cell size is $1/\lambda$ [2]. The average cell radius in a network is defined as the radius r_{cell} of a ball having the same average area as the cell: $1/\lambda$. Then, the average ISD is $2r_{\mathrm{cell}} = 2/(\sqrt{\pi\lambda})$. Hence we simply adjust the value of the intensity λ of BSs to represent different ISD scenarios. For FR1 bands, we analyze the following ISDs: 1000 m, 500 m, and 250 m. As for FR2 bands, the effect of propagation attenuation requires us to decrease the ISD. Hence we consider the following ISDs: 250 m, 125 m, and 75 m.

[4]As seen from previous sections, our model is amenable to different modeling of antenna gain patterns without loss of insights obtained in this paper.

In Figs. 2, 3, and 4, we evaluate the performances of FR1 and FR2 deployments, for different MT speeds v and ISDs. The following discussions are valid for Figs. 2, 3, and 4.

In both FR1 and FR2 deployments, for a given ISD, as n (and hence the number 2^n of beams) increases, the beamforming gain increases and the interference from interfering BSs decreases due to narrower beams. But, at the same time, for a given v, the typical MT performs more frequent beam reselections due to a smaller beamwidth, increasing the time overhead. The negative impact of the increased time overhead gradually dominates the beamforming gain. Also, with an increase in n, the probability of beam misalignment (given by (3)) increases as the MT is more likely to move into the beamwidth of another beam between two consecutive SSB measurements. As a result of this tussle, we observe that there is a value n_* of n that maximizes the effective ASE.

Similarly, with an increase in v, the MT crosses beam and cell boundaries more frequently resulting in a higher number of beam reselections and BS handovers, respectively. Also, the probability of beam misalignment increases. Hence, with an increase in v, the optimal value n_* decreases, and for a fixed n, the effective ASE decreases.

Note that a smaller ISD means a denser cellular network. As the ISD increases, the average cell size increases. Thus, BSs need to increase the number 2^n of beams for a higher beamforming gain (equivalently, a smaller beamwidth) and smaller interference. Hence, the value of n_* increases with ISD. For a smaller n, the time overhead associated with beam reselections is small. Hence, a smaller ISD results in a higher effective ASE due to a smaller distance between the typical MT and its serving BS. This dominates the negative effects of increased interference power due to a denser deployment of BSs and increased time overhead due to more frequent BS handovers. But, for a large n, both beam reselections and BS handovers happen more frequently for a smaller ISD, and the

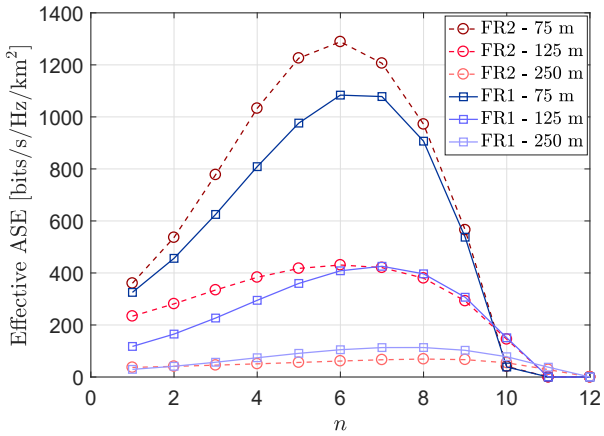

Fig. 4. Comparison between FR1 and FR2 deployments for different ISDs with $v = 30$ km/h.

network with a larger ISD achieves a higher effective ASE.

When comparing beamwidths for FR1 with those for FR2 given the same n value, we can expect that, even though the angular beamwidth for a given n is the same, the linear width of the beams will be larger in FR1 than that in FR2 due to a *usually* larger cell radius in the former.[5] Thus, it is expected that MTs operating in FR1 are less penalized due to overheads associated with beam reselections and BS handovers. Also, due to a larger ISD, the *distance-based* path loss when operating in FR1 is higher than when operating in FR2. Hence for the same value of n, we can expect that the interference in FR1 is smaller than that in FR2 despite a bit higher transmit power in FR1. This indicates that the additional received interference from a higher number of beams will not be as intense in FR2 as in FR1. As shown in Fig. 4, this behavior is captured by our model, where the optimal value n_* for the ISD $= 250$ m in FR1 ($n_* = 8$) is higher than that in FR2 ($n_* = 6$) with ISD $= 75$ m or in FR2 with ISD $= 125$ m ($n_* = 7$).

Also, as Fig. 4 shows, when we compare the performances of FR1 and FR2 deployments for the same ISD, we observe an interesting trade-off. Smaller ISDs (e.g., 75 m) benefit FR2 irrespective of the value of n as, for a given critical LOS distance R_{c}, it is more likely that the serving BS has LOS propagation to the MT (so smaller path loss) boosting signal power. Recall that, in FR1, even the serving BS always has NLOS propagation to the MT. As a result, the impact of increased time overheads with n on the effective ASE is less in FR2 than that in FR1. As the ISD increases, the probability of serving BS lying outside the LOS ball increases, in turn, reducing signal power significantly in FR2. Thus, when combined with higher path loss at higher frequencies in FR2, for higher values of n, the impact of time overheads is relatively higher in FR2 than FR1. Note that, as a result of the tussle between competing effects, the values of n_* in FR1 and FR2 are close to each other for the same ISD.

Overall, the numerical results discussed in this section validate the claim that the proposed model fits as a tool for system-level evaluation of network planning decisions in beam-based access networks such as 5G NR.

[5] Although nothing prohibits small ISDs for FR1 (e.g., Wi-Fi operation in 2.4 GHz and 5 GHz), pico/femto cell deployment models are the best fit for mmWave deployments.

VII. Conclusions and Future Directions

We presented a system-level stochastic geometry model for a 5G NR radio access network that allows one to capture essential scenario characteristics (e.g., operational frequency, blockage characteristics), technological features (e.g., beamforming configuration, delay overheads), and network deployment choices (e.g., ISD, carrier bandwidth). We analyzed the application of the model for a dense urban macro/pico cell scenario for FR1 and FR2. We demonstrated that the model accurately captures the existing trade-offs between ISDs, path loss, interference, and signaling overhead due to beam management. This also shows that the stochastic geometry model can be effective for conducting system-level analysis of beam-based radio access networks such as 5G NR.

In the future, we will extend our model to capture system-level impact of multiple MT panels and/or multiple MT beams. We also aim to improve the model to incorporate sectorized BSs with antenna panels and their effects on beam shaping/gains, MT blockage models, and others. Finally, we intend to extend our model to capture other trade-offs involved in designing a beam-based radio access network. For instance, with an increase in the number of beams, the average beam time-of-stay (ToS), i.e., the time a user remains with a given beam before switching to another beam, decreases, which impacts the available time for performing channel estimation, link adaptation, and power control. Also, increasing the periodicity of monitored resources may improve the effectiveness and response time of beam refinement, but it results in reduced spectral efficiency due to the control overhead.

References

[1] H. Holma, A. Toskala, and T. Nakamura, 5G Technology: 3GPP New Radio. John Wiley & Sons Ltd., 2020.
[2] F. Baccelli and B. Błaszczyszyn, Stochastic Geometry and Wireless Networks, NoW Publishers, 2009.
[3] M. Haenggi, Stochastic Geometry for Wireless Networks. Cambridge, U.K.: Cambridge University Press, 2012.
[4] M. Polese, M. Giordani, M. Mezzavilla, S. Rangan and M. Zorzi, "Improved handover through dual connectivity in 5G mmWave mobile networks," in IEEE Journal on Selected Areas in Communications, vol. 35, no. 9, pp. 2069-2084, Sept. 2017.
[5] M. Tayyab, X. Gelabert and R. Jäntti, "A survey on handover management: From LTE to NR," in IEEE Access, vol. 7, 2019.
[6] H. Tabassum, M. Salehi and E. Hossain, "Fundamentals of mobility-aware performance characterization of cellular networks: A tutorial," in IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2288-2308, third quarter 2019.
[7] K. Tokuyama and N. Miyoshi, "Data rate and handoff rate analysis for user mobility in cellular networks," IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, 2018, pp. 1-6.
[8] A. Chattopadhyay, B. Błaszczyszyn and E. Altman, "Two-tier cellular networks for throughput maximization of static and mobile users," in IEEE Transactions on Wireless Communications, vol. 18, no. 2, pp. 997-1010, Feb. 2019.

[9] E. Demarchou, C. Psomas and I. Krikidis, "Mobility management in ultra-dense networks: Handover skipping techniques," in IEEE Access, vol. 6, pp. 11921-11930, 2018.
[10] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri and M. Alouini, "Handover management in dense cellular networks: A stochastic geometry approach," IEEE International Conference on Communications (ICC), Kuala Lumpur, 2016, pp. 1-7.
[11] R. Arshad, H. ElSawy, S. Sorour, T. Y. Al-Naffouri and M. Alouini, "Velocity-aware handover management in two-tier cellular networks," in IEEE Transactions on Wireless Communications, vol. 16, no. 3, pp. 1851-1867, Mar. 2017.
[12] Y. Teng, A. Liu and V. K. N. Lau, "Stochastic geometry based handover probability analysis in dense cellular networks," International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, 2018, pp. 1-7.
[13] Z. Li and W. Wang, "Handover performance in dense mmWave cellular networks," International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, 2018, pp. 1-7.
[14] M. Giordani, M. Polese, A. Roy, D. Castor and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," in IEEE Communications Surveys and Tutorials, vol. 21, no. 1, pp. 173-196, First quarter 2019.
[15] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," in IEEE Transactions on Wireless Communications, vol. 14, no. 2, pp. 1100-1114, Feb. 2015.
[16] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," IEEE Transactions on Communications, vol. 63, no. 9, pp. 3029-3056, Sep. 2015.
[17] F. Baccelli and S. Zuyev, "Stochastic geometry models of mobile communication networks," Frontiers in Queueing, CRC Press, pp. 227-243, 1997.
[18] 3GPP, "NR; User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone," 3rd Generation Partnership Project (3GPP), TS 38.101-1, Version 15.9.0, April 2020.
[19] 3GPP, "NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone," 3rd Generation Partnership Project (3GPP), TS 38.101-2, Version 15.9.1, April 2020.
[20] https://portal.3gpp.org/desktopmodules/Specifications /SpecificationDetails.aspx?specificationId=319.

Appendix A
Proof of Lemma 1

The success probability is

$$p_{\rm s}(n,\beta) \triangleq \mathbb{P}(\mathsf{SINR}_n > \beta). \tag{18}$$

Depending on whether the serving BS lies within the LOS ball of radius $R_{\rm c}$ or not, we can write (18) as

$$\begin{aligned} p_{\rm s}(n,\beta) = &\int_0^{R_{\rm c}} \mathbb{P}(\mathsf{SINR}_n > \beta \mid r) f_R(r) {\rm d}r \\ &+ \int_{R_{\rm c}}^{\infty} \mathbb{P}(\mathsf{SINR}_n > \beta \mid r) f_R(r) {\rm d}r, \end{aligned} \tag{19}$$

where $f_R(r) = 2\pi\lambda r \exp(-\lambda\pi r^2)$ is the probability density function of the distance R of the typical MT to the nearest BS. When $0 < r < R_{\rm c}$, we have

$$\begin{aligned} \mathbb{P}(\mathsf{SINR}_n > \beta \mid r) &= \mathbb{P}\left(\frac{PKG_0 h_{X_0} r^{-\alpha_{\rm L}}}{N_0 + I_n} > \beta \mid r\right) \\ &= \mathbb{P}\left(h_{X_0} > \frac{\beta r^{\alpha_{\rm L}}(N_0 + I_n)}{PKG_0} \mid r\right) \\ &= e^{-\frac{\beta r^{\alpha_{\rm L}} N_0}{PKG_0}} \mathbb{E}\left[\exp\left(-\frac{\beta r^{\alpha_{\rm L}} I_n}{PKG_0}\right)\right]. \end{aligned} \tag{20}$$

We have

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\frac{\beta r^{\alpha_{\rm L}} I_n}{PKG_0}\right)\right] \\ &= \mathbb{E}\left[\exp\left(-\frac{\beta r^{\alpha_{\rm L}} \sum_{X\in\Phi\setminus\{X_0\}} PG_n h_X l(|X|)}{PKG_0}\right)\right] \\ &\overset{\rm (a)}{=} \mathbb{E}\left[\prod_{X\in\Phi\setminus\{X_0\}} \mathbb{E}\left(\exp\left(-\frac{\beta r^{\alpha_{\rm L}} G_n h_X l(|X|)}{KG_0}\right)\right)\right] \\ &\overset{\rm (b)}{=} \mathbb{E}\left[\prod_{X\in\Phi\setminus\{X_0\}} \mathbb{E}\left(\frac{1}{1+\frac{\beta r^{\alpha_{\rm L}} l(|X|) G_n}{KG_0}}\right)\right] \\ &\overset{\rm (c)}{=} \mathbb{E}\left[\prod_{X\in\Phi\setminus\{X_0\}} \left(\frac{p_{I,n}}{1+\frac{\beta r^{\alpha_{\rm L}} l(|X|) G_{\rm m,n}}{KG_0}} + \frac{1-p_{I,n}}{1+\frac{\beta r^{\alpha_{\rm L}} l(|X|) G_{\rm s,n}}{KG_0}}\right)\right], \end{aligned}$$

where (a) follows from the i.i.d. nature of fading random variables, (b) follows from averaging over the channel power gains on interfering channels, and (c) follows from the fact that the interference from an interfering BS at $X \in \Phi$ comes from its main lobe with probability $p_{I,n}$ and from its side lobe with probability $1 - p_{I,n}$ with $p_{I,n} = \varphi_n/2\pi$ the probability that the typical MT lies within the beamwidth of the main lobe of an interfering BS.

From the probability generating functional (PGFL) of PPP, it follows that

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\frac{\beta r^{\alpha_{\rm L}} I_n}{PKG_0}\right)\right] \\ &= \exp\left(-2\pi\lambda\int_r^{\infty}\left(1 - \frac{p_{I,n}}{1+\frac{\beta r^{\alpha_{\rm L}} l(w) G_{\rm m,n}}{KG_0}} - \frac{1-p_{I,n}}{1+\frac{\beta r^{\alpha_{\rm L}} l(w) G_{\rm s,n}}{KG_0}}\right) w {\rm d}w\right) \\ &= \exp\left(-\int_r^{R_{\rm c}} F(\alpha_{\rm L},\alpha_{\rm L},w){\rm d}w - \int_{R_{\rm c}}^{\infty} F(\alpha_{\rm L},\alpha_{\rm N},w){\rm d}w\right), \end{aligned} \tag{21}$$

where

$$F(\alpha_{\rm S},\alpha_{\rm I},w) \triangleq 2\pi\lambda\left(1 - \frac{p_{I,n}}{1+\frac{\beta r^{\alpha_{\rm S}} G_{\rm m,n}}{G_0 w^{\alpha_{\rm I}}}} - \frac{1-p_{I,n}}{1+\frac{\beta r^{\alpha_{\rm S}} G_{\rm s,n}}{G_0 w^{\alpha_{\rm I}}}}\right) w.$$

Similarly, we can obtain an expression for $\mathbb{P}(\mathsf{SINR}_n > \beta \mid r)$ when $R_{\rm c} \le r < \infty$.

Appendix B
Proof of Theorem 2

The beam reselections occur at beam boundaries of a BS. In Fig. 1, the locations of beam reselections are denoted by brown filled circles. As shown in Fig. 5, let θ denote the angle of a beam boundary with respect to the direction of the motion of the MT. This angle θ is distributed uniformly at random in $[0,\pi]$. In Fig. 5, a 'brown filled circle' denotes a point along the X-axis where the typical MT reselects the beam. Let $\Psi \subset \mathbb{R}$ denote the point process of beam reselection events, where the points of Ψ are indicated by 'brown filled circles.' We are interested in calculating the average number of beam reselections the typical MT performs per unit length, which is equivalent to calculating the intensity of Ψ.

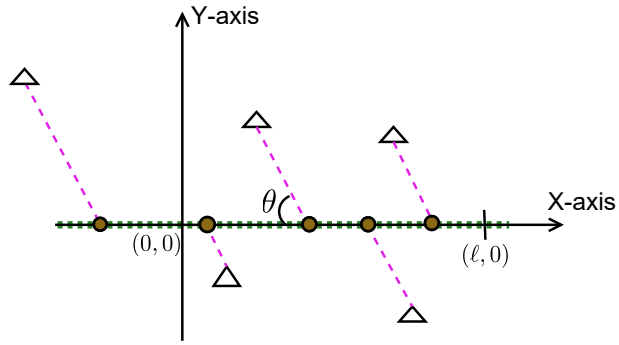

Fig. 5. At most one beam reselection for a BS. △ : Base station (BS), brown filled circle : Beam reselection location, dashed lines : Beam boundaries, and dotted line : Path of MT along the X-axis.

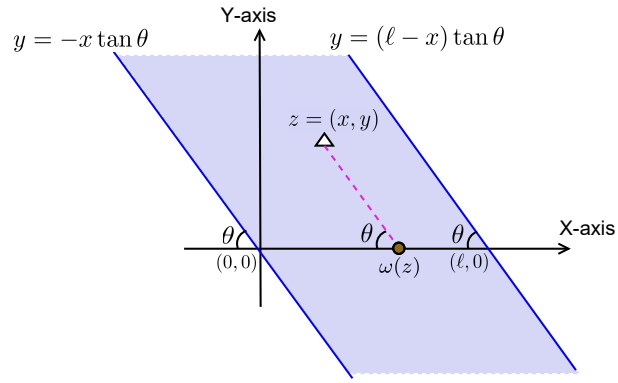

Fig. 6. Strip $\mathcal{S}$ (shaded area) between two blue solid lines with angle θ with the X-axis. △ : BS, brown filled circle : Beam reselection location, and dashed line : Beam boundary.

Without loss of generality, let us consider the motion of the typical MT in the interval $[0, \ell]$, which corresponds to the motion of the typical MT from $(0, 0)$ to $(\ell, 0)$ along the X-axis. First, we calculate the average number of beam reselections when there is at most one beam reselection corresponding to a BS. This happens when there are two beams of the same size for a BS. Using the result for two beams, we can obtain the intensity of beam reselections corresponding 2^n beams of a BS.

During the interval $[0, \ell]$, the typical MT may move through the Voronoi cells of multiple BSs. As shown in Fig. 6, let $\omega(X)$ denote the point of the beam reselection corresponding to the BS located at $X \in \Phi$. The event of beam reselection corresponding to a BS at $X \in \Phi$ occurs when the following two events occur simultaneously:

1) the point of the beam reselection lies in the Voronoi cell of the BS at $X \in \Phi$, i.e., $\omega(X) \in V_\Phi(X)$ where $V_\Phi(X)$ is the Voronoi cell of the BS located at $X \in \Phi$.
2) the point of the beam reselection lies on the line connecting $(0, 0)$ and $(\ell, 0)$, i.e, $\omega(X) \in [0, \ell]$.

Consequently, conditioning on θ, the average of number of beam reselections in $[0, \ell]$ is

$$\mathbb{E}(\Psi[0,\ell] \mid \theta) = \mathbb{E}\left(\sum_{X\in\Phi} \mathbf{1}_{\omega(X)\in V_\Phi(X)} \mathbf{1}_{\omega(X)\in[0,\ell]}\right),$$

where $\mathbf{1}$ is the indicator function. Conditioning on the fact that there is a BS at location $z \in \mathbb{R}^2$ and using the Campbell's theorem [2], it follows that

$$\mathbb{E}(\Psi[0,\ell] \mid \theta) = \lambda \int_{\mathbb{R}^2} \mathbb{P}_z(\omega(z) \in V_\Phi(z))\ \mathbf{1}_{\omega(z)\in[0,\ell]}\ \mathrm{d}z, \tag{22}$$

where $\mathbb{P}_z(\cdot)$ denotes the Palm probability.

Event $\boldsymbol{\omega(z) \in V_\Phi(z)}$: This event occurs when there is no BS closer to the typical MT than the one at location z. This is equivalent to the event that there is no BS in the ball of radius $\|z - \omega(z)\|$ centered at $\omega(z)$. Since the BS point process is a PPP with intensity λ, we have

$$\begin{aligned}\mathbb{P}_z(\omega(z) \in V_\Phi(z)) &= \exp\left(-\lambda\pi\|z-\omega(z)\|^2\right)\\ &= \exp\left(-\frac{\lambda\pi y^2}{\sin^2\theta}\right),\end{aligned} \tag{23}$$

where $\|z - \omega(z)\| = \frac{y}{\sin\theta}$ since $z = (x, y)$.

Event $\boldsymbol{\omega(z) \in [0, \ell]}$: This event occurs if the BS at z lies within the strip $\mathcal{S}$ between two lines passing through the origin $(0, 0)$ and $(\ell, 0)$ at angle θ as shown in Fig. 6. Equivalently, by representing z in Euclidean coordinates as $z = (x, y)$, it follows that

$$\mathbf{1}_{\omega(z)\in[0,\ell]}\ \mathrm{d}z = \mathbf{1}_{z\in\mathcal{S}}\ \mathrm{d}z = \mathbf{1}_{(x,y)\in\mathcal{S}}\ \mathrm{d}x\mathrm{d}y. \tag{24}$$

The left line passing through the origin at an angle θ can be given as $y = -x\tan\theta$, while the right line passing through $(\ell, 0)$ at an angle θ can be given as $y = (\ell - x)\tan\theta$. Thus, the BS at z lies in the strip $\mathcal{S}$ if

$$-\infty < y < \infty \quad \text{and} \quad -\frac{y}{\tan\theta} \le x \le \ell - \frac{y}{\tan\theta}. \tag{25}$$

From (23), (24), and (25), we can express (22) as

$$\begin{aligned}\mathbb{E}(\Psi[0,\ell] \mid \theta) &= \lambda \int_{y=-\infty}^{\infty} \mathrm{d}y \int_{x=-\frac{y}{\tan\theta}}^{\ell-\frac{y}{\tan\theta}} \exp\left(-\frac{\lambda\pi y^2}{\sin^2\theta}\right)\mathrm{d}x\\ &= \ell\left(\sqrt{\lambda}|\sin\theta|\right).\end{aligned} \tag{26}$$

Since θ is uniformly distributed in $[0, \pi]$, averaging over θ yields

$$\mathbb{E}(\Psi[0,\ell]) = \left(\frac{2\sqrt{\lambda}}{\pi}\right)\ell. \tag{27}$$

Hence, for the case of two beams, the linear intensity of beam reselections is $\frac{2\sqrt{\lambda}}{\pi}$.

For 2^n beams with $n \in \mathbb{N}$, there are 2^{n-1} possibilities of beam reselections corresponding to the serving BS. In this case, we can obtain the average number of beam reselections by superimposing the beam reselection events corresponding to each beam boundary crossing considered earlier in this

proof. Hence, the linear intensity of beam reselections for 2^n beams is

$$\mu_{\mathrm{s,b}}(n) = 2^{n-1}\frac{2\sqrt{\lambda}}{\pi} = \frac{2^n\sqrt{\lambda}}{\pi}. \tag{28}$$

By considering the speed v of the typical MT, we get the time intensity $\mu_{\mathrm{t,b}}$ of beam reselections as

$$\mu_{\mathrm{t,b}}(n) = \frac{2^n\sqrt{\lambda}}{\pi}v. \tag{29}$$