

# On the Computation and Approximation of Outage Probability in Satellite Networks with Smart Gateway Diversity

Christos N. Efrem and Athanasios D. Panagopoulos, *Senior Member, IEEE*

**Abstract**—The utilization of extremely high frequency (EHF) bands can achieve very high throughput in satellite networks (SatNets). Nevertheless, the severe rain attenuation at EHF bands imposes strict limitations on the system availability. Smart gateway diversity (SGD) is considered indispensable in order to guarantee the required availability with reasonable cost. In this context, we examine a load-sharing SGD (LS-SGD) architecture, which has been recently proposed in the literature. For this diversity scheme, we define the system outage probability (SOP) using a rigorous probabilistic analysis based on the Poisson binomial distribution (PBD), and taking into consideration the traffic demand as well as the gateway (GW) capacity. Furthermore, we provide several methods for the exact and approximate calculation of SOP. As concerns the exact computation of SOP, a closed-form expression and an efficient algorithm based on a recursive formula are given, both with quadratic worst-case complexity in the number of GWs. Finally, the proposed approximation methods include well-known probability distributions (binomial, Poisson, normal) and a Chernoff bound. According to the numerical results, binomial and Poisson distributions are by far the most accurate approximation methods.

**Index Terms**—Satellite networks, EHF bands, feeder link, smart gateway diversity, outage probability, Poisson binomial distribution, recursive formula, approximation methods.

## I. INTRODUCTION

NEXT-generation broadband SatNets require very high data-rates (up to 1 Tbps) that can be accomplished by utilizing EHF bands (above 30 GHz) in the feeder links. Although the frequency shift from Ka (20/30 GHz) to Q/V (40/50 GHz) or W (75-110 GHz) bands provides more spectrum, the high levels of rain attenuation (tens of dB) cannot be tackled by the standard fade mitigation techniques (FMTs), such as uplink power control (ULPC), adaptive coding and modulation (ACM) and data rate adaptation (DRA). As a result, gateway diversity (GD) is necessary to achieve high system availability, since it is a more effective and powerful FMT (at the expense of installing additional GWs) [1]–[5]. Nevertheless, the conventional GD (where the same signal is transmitted by two or three GWs) is economically prohibitive for reaching the Tbps due to the large number of required GWs [6]. An alternative solution to achieve high availability with reasonable cost is the smart gateway diversity (SGD), where a user beam can be served by different GWs depending on the propagation conditions and the traffic load. In particular, if a

GW experiences deep fades then its traffic can be rerouted to other GWs with better propagation conditions.

### A. Related Work

In [6], two SGD techniques are examined, namely, the frequency multiplexing diversity and the  $N + P$  diversity. The performance analysis of these schemes is based on a simple probabilistic model, assuming the same outage probability for each GW (although unusual in practice) as well as independent propagation conditions over the GW locations. Moreover, the authors in [7] study the  $N$ -active diversity (with time or frequency multiplexing, taking into account the spatial correlation between the GWs) and the  $N + P$  diversity (where there are  $N$  active plus  $P$  redundant or idle GWs). In the former scheme, all the  $N$  GWs are active and each user beam is served by a group of GWs, whereas in the latter scheme each user beam is served by only one GW and switches to a redundant GW in case of outage.

A novel GW switching scheme for the  $N + P$  scenario is proposed in [8], using a dynamic rain attenuation model and considering two key performance indicators: the average outage probability and the average switching rate. Furthermore, a different SGD scheme, where there is no redundant GWs but each GW should have some spare capacity, is analyzed in [9]. Specifically, in nominal clear-sky conditions all GWs are active and operate using a maximum fraction of their full capacity, while if some GWs experience heavy rain attenuation then their traffic is served by the remaining GWs using their extra capacity. Finally, an extension of the well-known  $N$ -active and  $N + P$  diversity schemes to multiple-input-multiple-output (MIMO) architectures is presented in [10].

### B. Contribution

The main contributions of this work, in comparison with existing approaches, are as follows:

- In this paper, we analyze in detail a SGD architecture operating in *load-sharing mode*, where the GWs do not necessarily have equal outage probabilities. To the best of our knowledge, the concept of LS-SGD has been firstly introduced in [9], assuming that all GWs utilize the same fraction of their full capacity in clear-sky conditions; our analysis, however, does not make such an assumption.
- Unlike previous research, we present a *system-level approach* taking into account the *traffic demand* as well as the *GW capacity*. In particular, we are interested in the *system outage probability (SOP)*, defined as the probability of not satisfying the overall traffic demand, which is a *stricter performance metric* than the *user*

C. N. Efrem and A. D. Panagopoulos are with the School of Electrical and Computer Engineering, National Technical University of Athens, 15780 Athens, Greece (e-mails: chefr@central.ntua.gr, thpanag@ece.ntua.gr).

This article has been accepted for publication in *IEEE Transactions on Aerospace and Electronic Systems*, DOI: 10.1109/TAES.2020.3022437. Copyright © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See [http://www.ieee.org/publications\\_standards/publications/rights/index.html](http://www.ieee.org/publications_standards/publications/rights/index.html) for more information.

outage probability (UOP), i.e., the probability of not satisfying the traffic demand of a specific user.

- Furthermore, we study the performance improvement (in terms of SOP) that can be achieved by increasing the number of GWs in the LS-SGD scheme. For this purpose, we define two comparative metrics, namely, the *SOP-improvement factor* and the *generalized SOP-improvement factor*.
- In addition, exact methods for the computation of SOP are given, including a *closed-form expression* and an *efficient algorithm based on a recursive formula*. The worst-case complexity of both methods is quadratic in the number of GWs.
- Finally, we provide some approximation methods for the estimation of SOP. More specifically, the SOP can be approximated by various probability distributions (binomial, Poisson, normal) as well as a Chernoff bound. Ultimately, we conclude that *binomial* and *Poisson* distributions are the most appropriate approximation methods for SGD systems operating in EHF bands.

### C. Paper Organization

The remainder of this article is organized as follows. Firstly, Section II describes and analyzes in more detail the LS-SGD architecture. Moreover, Sections III and IV present exact and approximation methods for calculating the SOP, respectively. In addition, the performance of LS-SGD as well as the accuracy of approximation methods are examined in Section V. Finally, concluding remarks are given in Section VI.

### D. Mathematical Notation & Conventions

*Mathematical notation:*  $\mathbb{Z}^+ = \{1, 2, 3, \dots\}$ ,  $\mathbb{Z}_0^+ = \{0, 1, 2, \dots\}$ ,  $\mathcal{N} = \{1, 2, \dots, N\}$  and  $\mathcal{N}_0 = \{0, 1, \dots, N\}$ , where  $N \in \mathbb{Z}^+$ . Moreover,  $\mathbb{P}(\cdot)$  and  $\mathbb{E}(\cdot)$  denote probability and expectation, respectively.  $\lfloor \cdot \rfloor$  and  $\lceil \cdot \rceil$  are respectively the floor and ceiling functions. In addition,  $|x|$  represents the absolute value of a real number  $x$ , while  $|\mathcal{S}|$  stands for the cardinality of a set  $\mathcal{S}$ .  $\mathbf{0}_N$  and  $\mathbf{1}_N$  denote the  $N$ -dimensional all-zeros and all-ones vectors, respectively. Furthermore,  $\varphi(x) = (\sqrt{2\pi})^{-1} e^{-0.5x^2}$  is the probability density function (PDF),  $\Phi(x) = \int_{-\infty}^x \varphi(u) du$  is the cumulative distribution function (CDF), and  $Q(x) = 1 - \Phi(x)$  is the complementary CDF (CCDF) of the standard normal distribution. Finally, the *total variation distance* between two (discrete) random variables (RVs)  $X$  and  $Y$  on  $\mathbb{Z}_0^+$  is defined as follows:

$$d_{\text{TV}}(X, Y) = \sup_{\mathcal{A} \subseteq \mathbb{Z}_0^+} |\mathbb{P}(X \in \mathcal{A}) - \mathbb{P}(Y \in \mathcal{A})| = \frac{1}{2} \sum_{m \in \mathbb{Z}_0^+} |\mathbb{P}(X = m) - \mathbb{P}(Y = m)| \quad (1)$$

*Mathematical conventions:*  $\sum_{i \in \emptyset} a_i = 0$  and  $\prod_{i \in \emptyset} a_i = 1$ .

### E. Preliminaries on Discrete Probability Distributions

1) *Bernoulli Distribution:* A binary (0/1) RV follows a Bernoulli distribution with parameter  $p \in [0, 1]$ ,  $X \sim \text{Bern}(p)$ , if and only if (iff) its probability mass function (PMF) is given by:  $\mathbb{P}(X = 1) = p$  and  $\mathbb{P}(X = 0) = 1 - p$ .

2) *Binomial Distribution:* A discrete (integer-valued) RV  $X \sim \text{Bin}(N, p)$ , where  $N \in \mathbb{Z}^+$  and  $p \in [0, 1]$ , iff its PMF is:

$$\mathbb{P}(X = m) = \binom{N}{m} p^m (1 - p)^{N-m}, \quad \forall m \in \mathcal{N}_0 \quad (2)$$

The binomial distribution is a generalization of the Bernoulli distribution, because  $\text{Bin}(1, p) \equiv \text{Bern}(p)$ . Furthermore, if  $\{X_n\}_{n \in \mathcal{N}}$  is a set of *independent and identically distributed* (i.i.d.) Bernoulli RVs ( $X_n \sim \text{Bern}(p)$ ,  $\forall n \in \mathcal{N}$ ), then  $S = \sum_{n \in \mathcal{N}} X_n \sim \text{Bin}(N, p)$ .

3) *Poisson Binomial Distribution:* A discrete RV  $X \sim \text{PoisBin}(\mathbf{p})$ , where  $\mathbf{p} = [p_1, p_2, \dots, p_N] \in [0, 1]^N$  with  $N \in \mathbb{Z}^+$ , iff its PMF is given by:

$$\mathbb{P}(X = m) = \sum_{\mathcal{A} \in \mathcal{C}_m} \prod_{i \in \mathcal{A}} p_i \prod_{j \in \mathcal{N} \setminus \mathcal{A}} (1 - p_j), \quad \forall m \in \mathcal{N}_0 \quad (3)$$

where  $\mathcal{C}_m = \{\mathcal{A} \subseteq \mathcal{N} : |\mathcal{A}| = m\}$  (i.e., the set of all subsets of  $\mathcal{N}$  having  $m$  elements) with  $|\mathcal{C}_m| = \binom{N}{m} = \frac{N!}{m!(N-m)!}$ . The binomial distribution is a special case of the PBD, since  $\text{PoisBin}(p\mathbf{1}_N) \equiv \text{Bin}(N, p)$ . Moreover, if  $\{X_n\}_{n \in \mathcal{N}}$  is a set of *independent, but not necessarily identically distributed*, Bernoulli RVs ( $X_n \sim \text{Bern}(p_n)$ ,  $\forall n \in \mathcal{N}$ ), then  $S = \sum_{n \in \mathcal{N}} X_n \sim \text{PoisBin}(\mathbf{p})$ .

4) *Poisson Distribution:* A discrete RV  $X \sim \text{Pois}(\lambda)$ , where  $\lambda \geq 0$ , iff its PMF is expressed by:  $\mathbb{P}(X = m) = e^{-\lambda} \lambda^m (m!)^{-1}$ ,  $\forall m \in \mathbb{Z}_0^+$ .

## II. SMART GATEWAY DIVERSITY ARCHITECTURE

In this section, we describe and analyze a *load-sharing SGD* (LS-SGD) scheme, where the unused capacity of available (not in outage) GWs can be exploited to serve the users of the remaining GWs (which are in outage). To the best of our knowledge, this SGD architecture has been firstly proposed and analyzed in [9]. Nevertheless, our approach is somewhat different, since it explicitly takes into consideration the *traffic demand* as well as the *GW capacity*.

### A. System Model

Consider a SatNet consisting of a geostationary satellite and a ground network of  $N \in \mathbb{Z}^+$  (geographically distributed) GWs, which are denoted by the set  $\mathcal{N} = \{1, 2, \dots, N\}$ . All the GWs are connected to a *network control center* (NCC) through dedicated terrestrial links. The NCC performs, when necessary (in case of deep fading), the traffic switching/rerouting between the GWs.<sup>1</sup> Furthermore, the following analysis focuses on the *feeder links* (data transmission from the GWs to the satellite), considering ideal (without noise and interference) satellite-user links.<sup>2</sup>

In addition, the *distance between any two different GWs is large enough* (some hundreds of km), and thus the spatial correlation of the propagation impairments at the

<sup>1</sup>The details on the switching/handover procedure are beyond the scope of this paper; see [6], [8], [11] for more information on this important topic.

<sup>2</sup>As concerns the downlink of multibeam satellite systems, an energy-efficient power allocation in order to jointly minimize the unmet system capacity and the total radiated power is proposed in [12].

GW locations is extremely small [6], [13]. As a result, the rain attenuations/fades experienced by the GWs can be considered (*mutually*) *independent*. It is also assumed that there is no ACM, so each feeder link is either available at full capacity or completely unavailable.<sup>3</sup> Therefore, the feeder links can be mathematically modeled as a set  $\{X_n\}_{n \in \mathcal{N}}$  of *independent, but not necessarily identically distributed*, Bernoulli RVs ( $X_n \sim \text{Bern}(p_n)$ ,  $\forall n \in \mathcal{N}$ ), where  $p_n \in [0, 1]$  is the outage/exceedance probability of the  $n^{\text{th}}$  link/GW (i.e., the probability that the rain attenuation exceeds a specific threshold); some methods for calculating  $p_n$  are discussed in [9]. Moreover, we define the RV  $S_{\mathcal{N}} = \sum_{n \in \mathcal{N}} X_n \sim \text{PoisBin}(\mathbf{p}_{\mathcal{N}})$ , with  $\mathbf{p}_{\mathcal{N}} = [p_1, p_2, \dots, p_N]$ , which is the *total number of GWs that are in outage in the set  $\mathcal{N}$* .<sup>4</sup> The expectation, the standard deviation, and the 3<sup>rd</sup> central moment of  $S_{\mathcal{N}}$  are given respectively by:

$$\mu_{\mathcal{N}} = \mathbb{E}(S_{\mathcal{N}}) = \sum_{n \in \mathcal{N}} p_n \quad (4)$$

$$\sigma_{\mathcal{N}} = \sqrt{\mathbb{E}((S_{\mathcal{N}} - \mu_{\mathcal{N}})^2)} = \sqrt{\sum_{n \in \mathcal{N}} p_n(1 - p_n)} \quad (5)$$

$$\nu_{\mathcal{N}} = \mathbb{E}((S_{\mathcal{N}} - \mu_{\mathcal{N}})^3) = \sum_{n \in \mathcal{N}} p_n(1 - p_n)(1 - 2p_n) \quad (6)$$

Note that  $\mu_{\mathcal{N}} \geq \sigma_{\mathcal{N}}^2$ ,  $\mu_{\mathcal{N}} \in [0, N]$ ,  $\sigma_{\mathcal{N}}^2 \in [0, N/4]$ , and  $\nu_{\mathcal{N}} \in [-N/(6\sqrt{3}), N/(6\sqrt{3})]$ .

### B. System Outage Probability

In the sequel, suppose that the  $n^{\text{th}}$  GW can offer a *maximum data-rate (capacity)*  $R_n^{\max} > 0$ , and the *total requested data-rate (traffic demand)* is  $R_{\text{tot}}^{\text{req}} = \sum_{u \in \mathcal{U}} R_u^{\text{req}} > 0$ , where  $\mathcal{U} = \{1, 2, \dots, U\}$  is the set of users and  $R_u^{\text{req}} \geq 0$  is the requested data-rate of user  $u$ . Moreover, the operation of NCC ensures the following *load-sharing property*: all users receive their requested data-rate if and only if (iff) the overall capacity of the available (not in outage) GWs is greater than or equal to the traffic demand. Equivalently, there is at least one user that receives inadequate data-rate iff the overall capacity of the available GWs is less than the traffic demand.

**Definition 1** (General SOP expression). *The SOP is defined as follows:*

$$P_{\text{out}}^{\text{sys}} = \sum_{\mathcal{A} \in \mathcal{F}} \prod_{i \in \mathcal{A}} p_i \prod_{j \in \mathcal{N} \setminus \mathcal{A}} (1 - p_j) \quad (7)$$

<sup>3</sup>Classical FMTs, such as ULPC, ACM and DRA, can tackle impairments of a few dB (e.g., gaseous absorption and cloud attenuation). However, in EHF bands these techniques alone are no longer effective, because the rain attenuation can reach tens of dB. Hence, SGD has to be used in order to keep SOP at the required levels. In essence, due to the intense rain attenuation in EHF bands, *SGD is the primary FMT*, whereas ULPC, ACM and DRA are secondary/supplementary FMTs. As a result, the absence of ACM in the analysis of SGD is quite reasonable. In any case, our approach provides a *lower bound on the performance* of a more realistic system that utilizes SGD together with standard FMTs.

<sup>4</sup>According to Section I-E2, if  $p_n = p$ ,  $\forall n \in \mathcal{N}$  (i.i.d. Bernoulli RVs), then  $S_{\mathcal{N}} \sim \text{Bin}(N, p)$ . Note that this is rarely the case in practice.

where  $\mathcal{F} = \left\{ \mathcal{A} \subseteq \mathcal{N} : \sum_{j \in \mathcal{N} \setminus \mathcal{A}} R_j^{\max} < R_{\text{tot}}^{\text{req}} \right\}$ . In other words,  $\mathcal{F}$  contains all the subsets  $\mathcal{A}$  of the  $N$  GWs such that: if the GWs in  $\mathcal{A}$  are all in outage and the remaining GWs in  $\mathcal{N} \setminus \mathcal{A}$  are all available (not in outage), then the traffic demand cannot be satisfied by the latter group of GWs. In essence, the SOP expresses the probability of not satisfying the traffic demand of all users (or, equivalently, the probability that there is at least one user that receives inadequate data-rate). Similarly, we can define the system availability (SA) as the probability of the complementary event:  $P_{\text{avail}}^{\text{sys}} = 1 - P_{\text{out}}^{\text{sys}}$ .

For simplicity, we assume that all GWs have the same capacity,  $R_{\text{GW}}^{\max} > 0$ , in the rest of the paper; this is not such a strong assumption in practice, since the same frequency band is fully reused in each feeder link and the clear-sky link budget is almost identical for all GWs.

**Theorem 1** (Special SOP expression). *Suppose that all GWs have the same capacity, i.e.,  $R_n^{\max} = R_{\text{GW}}^{\max} > 0$ ,  $\forall n \in \mathcal{N}$ . Then, (7) reduces to the following expression:*

$$P_{\text{out}}^{\text{sys}} = P_{\text{out}}^{\text{sys}}(L, N) = \sum_{m=L}^N \sum_{\mathcal{A} \in \mathcal{C}_m} \prod_{i \in \mathcal{A}} p_i \prod_{j \in \mathcal{N} \setminus \mathcal{A}} (1 - p_j) \quad (8)$$

where  $\mathcal{C}_m = \{\mathcal{A} \subseteq \mathcal{N} : |\mathcal{A}| = m\}$  and  $L$  is given by:

$$L = N - \lceil r \rceil + 1 \quad (9)$$

where  $r > 0$  is the ratio of the traffic demand to the GW capacity, that is:

$$r = R_{\text{tot}}^{\text{req}} / R_{\text{GW}}^{\max} \quad (10)$$

*Proof.* Under the condition of equal GW capacities, we have that  $\mathcal{F} = \{\mathcal{A} \subseteq \mathcal{N} : (N - |\mathcal{A}|)R_{\text{GW}}^{\max} < R_{\text{tot}}^{\text{req}}\}$ . In addition,  $(N - |\mathcal{A}|)R_{\text{GW}}^{\max} < R_{\text{tot}}^{\text{req}} \Leftrightarrow N - |\mathcal{A}| < r \Leftrightarrow N - |\mathcal{A}| < \lceil r \rceil \Leftrightarrow N - |\mathcal{A}| \leq \lceil r \rceil - 1 \Leftrightarrow |\mathcal{A}| \geq N - \lceil r \rceil + 1$ . Consequently,  $\mathcal{F} = \{\mathcal{A} \subseteq \mathcal{N} : |\mathcal{A}| \geq L\} = \bigcup_{m=L}^N \mathcal{C}_m$  and then (8) follows immediately from (7).  $\square$

**Remark 1.** According to Section I-E3,  $P_{\text{out}}^{\text{sys}}(L, N) = \sum_{m=L}^N \mathbb{P}(S_{\mathcal{N}} = m) = \mathbb{P}(S_{\mathcal{N}} \geq L)$ , i.e., the SOP is the *probability of having at least  $L$  out of  $N$  GWs in outage*.<sup>5</sup>

Although in general  $L \in \mathcal{N}_0$ , for the diversity system under consideration  $L \in \mathcal{N}$  due to the fact that  $\lceil r \rceil \in \mathcal{N}$ , since a)  $r > 0 \Leftrightarrow \lceil r \rceil \geq 1$ , and b)  $NR_{\text{GW}}^{\max} \geq R_{\text{tot}}^{\text{req}} \Leftrightarrow N \geq r \Leftrightarrow N \geq \lceil r \rceil$  (note that  $N_{\min} = \lceil r \rceil$  is the *minimum required number of GWs*). Finally, we provide a result about the monotonicity of SOP.

**Proposition 1** (SOP monotonicity). *For a given set  $\mathcal{N}$  of GWs, the SOP is an increasing function of  $r$ .*

<sup>5</sup>Similar formula is also given in [9] and [14], however, without explicit dependence on the traffic demand and the GW capacity. Herein, this dependence is clearly expressed by (9) and (10). Note that this SOP definition is a generalization of the *classical SOP* (i.e., the probability of having all GWs in outage), which is obtained when  $\lceil r \rceil = 1 \Rightarrow L = N \Rightarrow P_{\text{out}}^{\text{sys}} = \prod_{n \in \mathcal{N}} p_n$ ; the classical SOP is used in [15] to select the (globally) minimum number of GWs satisfying SOP-requirements.

*Proof.* Let  $r_1 \geq r_2 \Rightarrow \lceil r_1 \rceil \geq \lceil r_2 \rceil \Rightarrow L_1 \leq L_2 \Rightarrow P_{\text{out}}^{\text{sys}}(L_1, N) \geq P_{\text{out}}^{\text{sys}}(L_2, N)$ .  $\square$

### C. SOP-Improvement Factor

Subsequently, we study the performance improvement (in terms of SOP) achieved by an  $N$ -GW diversity system in comparison with a single-GW system.

**Definition 2** (SOP-improvement factor). *Assuming the same  $\lceil r \rceil = 1$  and that  $P_{\text{out}}^{\text{sys}}(N, N) > 0$ , the SOP-improvement factor is defined as follows:*

$$I = \frac{P_{\text{out}}^{\text{sys}}(1, 1)}{P_{\text{out}}^{\text{sys}}(N, N)} = \frac{p_1}{\prod_{n \in \mathcal{N}} p_n} = \left( \prod_{n=2}^N p_n \right)^{-1} \quad (11)$$

Obviously, it holds that  $I \geq 1$ .

Next, consider a diversity system with  $N + K$  GWs ( $K \in \mathbb{Z}_0^+$ ) all of which have the same capacity  $R_{\text{GW}}^{\text{max}} > 0$ , and  $\lceil r \rceil \in \mathcal{N}$  (since  $1 \leq \lceil r \rceil \leq \min(N, N + K) = N$ ). Furthermore, let  $\mathcal{K} = \{N + 1, N + 2, \dots, N + K\}$  be the set of additional GWs, and  $\mathbf{p}_{\mathcal{N} \cup \mathcal{K}} = [\mathbf{p}_{\mathcal{N}}, \mathbf{p}_{\mathcal{K}}] = [p_1, p_2, \dots, p_{N+K}]$  be the vector of GW outage probabilities, where  $\mathbf{p}_{\mathcal{K}} = [p_{N+1}, p_{N+2}, \dots, p_{N+K}]$ . Suppose also that  $\{X_i\}_{i \in \mathcal{N} \cup \mathcal{K}}$  is a set of *independent, but not necessarily identically distributed*, Bernoulli RVs ( $X_i \sim \text{Bern}(p_i)$ ,  $\forall i \in \mathcal{N} \cup \mathcal{K}$ ). Besides  $S_{\mathcal{N}}$ , we define the RVs  $S_{\mathcal{K}} = \sum_{k \in \mathcal{K}} X_k \sim \text{PoisBin}(\mathbf{p}_{\mathcal{K}})$  and  $S_{\mathcal{N} \cup \mathcal{K}} = \sum_{i \in \mathcal{N} \cup \mathcal{K}} X_i = S_{\mathcal{N}} + S_{\mathcal{K}} \sim \text{PoisBin}(\mathbf{p}_{\mathcal{N} \cup \mathcal{K}})$  denoting the total number of GWs which are in outage in the sets  $\mathcal{K}$  and  $\mathcal{N} \cup \mathcal{K}$ , respectively. For this diversity system  $L' = N + K - \lceil r \rceil + 1 = L + K$ , with  $L' \in \{K + 1, K + 2, \dots, K + N\}$ .

**Proposition 2** (SOP reduction). *Let  $P_{\text{out}}^{\mathcal{N}} = P_{\text{out}}^{\text{sys}}(L, N) = \mathbb{P}(S_{\mathcal{N}} \geq L)$  and  $P_{\text{out}}^{\mathcal{N} \cup \mathcal{K}} = P_{\text{out}}^{\text{sys}}(L', N + K) = \mathbb{P}(S_{\mathcal{N} \cup \mathcal{K}} \geq L')$  stand for the SOP of the  $N$ -GW and  $(N + K)$ -GW diversity systems, respectively. Then, it holds that  $P_{\text{out}}^{\mathcal{N} \cup \mathcal{K}} \leq P_{\text{out}}^{\mathcal{N}}$ .*

*Proof.* See Appendix A.  $\square$

In view of this fact, we can generalize the definition of SOP-improvement factor.

**Definition 3** (Generalized SOP-improvement factor). *Assuming the same  $\lceil r \rceil \in \mathcal{N}$  and that  $P_{\text{out}}^{\mathcal{N} \cup \mathcal{K}} > 0$ , we define the generalized SOP-improvement factor of the  $(N + K)$ -GW over the  $N$ -GW diversity system as follows:<sup>6</sup>*

$$I_g = \frac{P_{\text{out}}^{\mathcal{N}}}{P_{\text{out}}^{\mathcal{N} \cup \mathcal{K}}} = \frac{P_{\text{out}}^{\text{sys}}(L, N)}{P_{\text{out}}^{\text{sys}}(L + K, N + K)} \Big|_{L=N-\lceil r \rceil+1} \quad (12)$$

According to Proposition 2, we have that  $I_g \geq 1$ .

Notice that by setting  $N = 1$  and  $K = N' - 1$  (thus  $\lceil r \rceil = 1$  and  $L = 1$ ), we obtain  $I_g = \frac{P_{\text{out}}^{\text{sys}}(1, 1)}{P_{\text{out}}^{\text{sys}}(N', N')} = I$ . Finally, we would like to emphasize that by increasing the number of GWs the SOP decreases, but higher GW connectivity is required; such connectivity issues are very important in the design and optimization of SatNets [16]. In other words, there is a trade-off between performance improvement and connectivity complexity.

<sup>6</sup>The generalized SOP-improvement factor  $I_g$  can be estimated using the approximation methods provided in Section IV.

TABLE I  
COMPLEXITY COMPARISON BETWEEN EXACT METHODS

Exact Method	Direct Computation	CFE	RF (Algorithm 1)	FFT-based Algorithm [20]
Time Complexity	$O(2^N N)$	$\Theta(N^2)$	$\Theta(L(N - L + 1)) = O(N^2)$	$O(N(\log N)^2)$

### III. EXACT METHODS FOR COMPUTING SOP

In the sequel, several techniques for the exact computation of SOP are presented. The time complexity of these methods is summarized in Table I.

#### A. Direct Computation

The direct computation of SOP is based on the analytic formula (8), which requires  $\sum_{m=L}^N |\mathcal{C}_m| N = N \sum_{m=L}^N \binom{N}{m} \leq N \sum_{m=0}^N \binom{N}{m} = 2^N N = O(2^N N)$  arithmetic operations. Because of its exponential worst-case complexity, this method is practicable only for very small  $N$ .

#### B. Closed-Form Expression

According to [17], the SOP can be calculated, using polynomial interpolation and discrete Fourier transform (DFT), by the following *closed-form expression (CFE)*:

$$P_{\text{out}}^{\text{sys}}(L, N) = 1 - \frac{1}{N+1} \left( L + \sum_{n \in \mathcal{N}} \frac{1-c^{-nL}}{1-c^{-n}} \prod_{m \in \mathcal{N}} (1 + (c^n - 1)p_m) \right) \quad (13)$$

where  $c = e^{j2\pi/(N+1)}$ , with  $j = \sqrt{-1}$  being the imaginary unit. It is interesting to note that the CFE comprises a sum of complex numbers, but the overall outcome is a real number in  $[0, 1]$ . The same formula is also derived in [18], using the characteristic function of the PBD as well as the DFT. Furthermore, the computational complexity of (13) is  $\Theta(N^2)$ .

#### C. Recursive Formula

In this part, we explore the power and beauty of recursion.

**Theorem 2** (SOP recursive formula). *The SOP is given by the following recursive formula (RF):*

$$P_{\text{out}}^{\text{sys}}(L, N) = (1 - p_N) P_{\text{out}}^{\text{sys}}(L, N - 1) + p_N P_{\text{out}}^{\text{sys}}(L - 1, N - 1) \quad (14)$$

with initial/boundary conditions: a)  $P_{\text{out}}^{\text{sys}}(0, N) = 1$  and b)  $P_{\text{out}}^{\text{sys}}(N + 1, N) = 0$ ,  $\forall N \in \mathbb{Z}^+$ .

*Proof.* See Appendix B.  $\square$

It can be verified, using mathematical induction, that (8) is the solution of (14). To the best of our knowledge, this RF is derived for the first time in [19], making use of symmetric switching functions. Our proof, however, is much simpler.

Algorithm 1 presents an efficient method to compute the SOP using the RF, which follows directly from the algorithm given in [19]. The *time complexity* of Algorithm 1 is  $\Theta(L(N - L + 1)) = O(N^2)$ , with best-case complexity  $\Theta(1)$  for  $L = 0$ , and worst-case complexity  $\Theta(N^2)$  for  $L = \lfloor N/2 \rfloor$ .

---

**Algorithm 1** Exact computation of SOP
 

---

**Input:**  $N \in \mathbb{Z}^+$ ,  $L \in \mathcal{N}_0$ , and  $\mathbf{p} = [p_1, p_2, \dots, p_N] \in [0, 1]^N$   
**Output:**  $P_{\text{out}}^{\text{sys}} = P_{\text{out}}^{\text{sys}}(L, N)$   
 1:  $D := N - L$ ,  $M := L + 1$ ,  $\alpha := \mathbf{0}_M$ ,  $\alpha_1 := 1$ ,  $\ell := 1$   
 2: **for**  $i := 1$  **to**  $N$  **step**  $+1$  **do**  
 3:    $h := i$   
 4:   **if**  $i > D + 1$  **then**  $\ell := i - D$  **end if**  
 5:   **if**  $i > L$  **then**  $h := L$  **end if**  
 6:   **for**  $j := h$  **to**  $\ell$  **step**  $-1$  **do**    $\triangleright h, \ell$ : high/low index  
 7:      $\alpha_{j+1} := (1 - p_i) \cdot \alpha_{j+1} + p_i \cdot \alpha_j$   
 8:   **end for**  
 9: **end for**  
 10:  $P_{\text{out}}^{\text{sys}} := \alpha_M$

---

and  $L = \lceil N/2 \rceil$ . Moreover, notice that the complexity is  $\Theta(N)$  for  $L = 1$  and  $L = N$ . As a result, Algorithm 1 has lower complexity in some cases than the CFE which requires  $\Theta(N^2)$  operations regardless of  $L$ . Finally, the *space complexity* of Algorithm 1 is  $\Theta(N + L) = \Theta(N)$ .

#### D. FFT-based Algorithm

An even more efficient and advanced algorithm for computing the SOP is provided in [20]. This method recursively applies the fast Fourier transform (FFT) to compute *generating function (GF)* products, thus achieving an overall complexity of  $O(N(\log N)^2)$ .

In particular, the PMF of  $S_{\mathcal{N}} \sim \text{PoisBin}(\mathbf{p}_{\mathcal{N}})$  can be written in the following form:

$$\begin{aligned} & [\mathbb{P}(S_{\mathcal{N}} = 0) \ \mathbb{P}(S_{\mathcal{N}} = 1) \ \dots \ \mathbb{P}(S_{\mathcal{N}} = N)] = \\ & = [q_1 \ p_1] * [q_2 \ p_2] * \dots * [q_N \ p_N] \end{aligned} \quad (15)$$

where  $*$  stands for the *convolution* operation and  $q_n = 1 - p_n$ ,  $\forall n \in \mathcal{N}$ . In addition, the GF of the Poisson-binomial PMF is given by:

$$\begin{aligned} g(z) &= \sum_{n \in \mathcal{N}_0} \mathbb{P}(S_{\mathcal{N}} = n) z^n = \prod_{n \in \mathcal{N}} (q_n + p_n z) = \\ &= g_{\pi} \prod_{n \in \mathcal{N}} (1 + a_n z) = g_{\pi} (1 + A(z)) \end{aligned} \quad (16)$$

where  $g_{\pi} = \prod_{n \in \mathcal{N}} q_n$  and  $a_n = p_n/q_n$ ,  $\forall n \in \mathcal{N}$ . Obviously, the SOP is simply the sum of the coefficients of  $z^m$  from  $m = L$  to  $N$  (see Remark 1). Since the product of two GF is equivalent to the convolution of two sequences formed from the GF coefficients, the FFT can be used to compute GF products more efficiently compared to the term-by-term calculation. The basic idea of the algorithm proposed in [20] is to apply the FFT to compute the GF  $A(z)$  using a *divide-and-conquer* approach. More details on the implementation of the algorithm can be found therein.

**Remark 2.** Despite the fact that the FFT-based algorithm is more sophisticated and has lower asymptotic complexity, CFE and Algorithm 1 are sufficient in the context of SGD, where the number of GWs  $N$  is relatively small.

#### IV. APPROXIMATION METHODS FOR ESTIMATING SOP

Afterwards, we introduce some useful methods to approximate the SOP, exploiting the fact that  $P_{\text{out}}^{\text{sys}}(L, N) = \mathbb{P}(S_{\mathcal{N}} \geq L) = 1 - \mathbb{P}(S_{\mathcal{N}} \leq L - 1)$ ,  $\forall L \in \mathcal{N}_0$ . These techniques consist of *probability distributions (binomial, Poisson, normal)* as well as a *Chernoff bound*. For convenience, a summary of approximation methods is given in Table II.

##### A. Binomial Approximation (BA)

The PBD can be approximated by the binomial distribution [24] in the following sense, defining  $\bar{p} = \frac{1}{N} \sum_{n \in \mathcal{N}} p_n$ ,  $\bar{q} = 1 - \bar{p}$ , and assuming  $\bar{p} \in (0, 1)$ :

a)  $d_{\text{TV}}(S_{\mathcal{N}}, Y) \leq (N/(N+1))(1 - \bar{p}^{N+1} - \bar{q}^{N+1})\delta_{\mathcal{N}}$ , where  $Y \sim \text{Bin}(N, \bar{p})$  and  $\delta_{\mathcal{N}} = 1 - (N\bar{p}\bar{q})^{-1}\sigma_{\mathcal{N}}^2$ , and b)  $d_{\text{TV}}(S_{\mathcal{N}}, Y) \rightarrow 0$  if and only if (iff)  $\delta_{\mathcal{N}} \rightarrow 0$  (or, equivalently,  $(N\bar{p}\bar{q})^{-1}\sigma_{\mathcal{N}}^2 \rightarrow 1$ ). It is interesting to note that when  $p_n = p$ ,  $\forall n \in \mathcal{N}$ , it holds that:  $\bar{p} = p$ ,  $\bar{q} = 1 - p$  and  $\sigma_{\mathcal{N}}^2 = N\bar{p}\bar{q} \Rightarrow \delta_{\mathcal{N}} = 0 \Rightarrow d_{\text{TV}}(S_{\mathcal{N}}, Y) = 0 \Rightarrow S_{\mathcal{N}} \sim \text{Bin}(N, p)$ , which is in agreement with Section I-E2. Hence, the BA is given by:

$$P_{\text{out}}^{\text{sys}}(L, N) \approx 1 - \mathbb{P}(Y \leq L - 1) = 1 - \sum_{m=0}^{L-1} \binom{N}{m} \bar{p}^m (1 - \bar{p})^{N-m} \quad (17)$$

##### B. Poisson Approximation (PA)

In 1960, Le Cam [25] established a remarkable inequality:  $d_{\text{TV}}(S_{\mathcal{N}}, Z) \leq \sum_{n \in \mathcal{N}} p_n^2$ , where  $Z \sim \text{Pois}(\mu_{\mathcal{N}})$ . It is obvious that if  $\sum_{n \in \mathcal{N}} p_n^2 \rightarrow 0$ , then  $d_{\text{TV}}(S_{\mathcal{N}}, Z) \rightarrow 0$ . As reported in [26], *Le Cam's theorem/finequality* admits various proofs using different techniques. Consequently, we have that:

$$P_{\text{out}}^{\text{sys}}(L, N) \approx 1 - \mathbb{P}(Z \leq L - 1) = 1 - e^{-\mu_{\mathcal{N}}} \sum_{m=0}^{L-1} \mu_{\mathcal{N}}^m (m!)^{-1} \quad (18)$$

##### C. Normal Approximation (NA)

According to [21], the *central limit theorem (CLT)* for the PBD states that:  $\lim_{N \rightarrow \infty} \Delta_{\mathcal{N}} = 0$  (asymptotic normality) of  $(S_{\mathcal{N}} - \mu_{\mathcal{N}})\sigma_{\mathcal{N}}^{-1}$  iff  $\lim_{N \rightarrow \infty} \sigma_{\mathcal{N}}^2 = \infty$ , where  $\Delta_{\mathcal{N}} = \sup_{s \in \mathbb{R}} |\mathbb{P}(S_{\mathcal{N}} \leq s) - \Phi((s - \mu_{\mathcal{N}})\sigma_{\mathcal{N}}^{-1})|$ . Therefore, by applying a *continuity correction*,<sup>7</sup> the SOP can be approximated by:

$$P_{\text{out}}^{\text{sys}}(L, N) \approx 1 - \Phi(\zeta) = Q(\zeta) \quad (19)$$

where  $\zeta = (L - \mu_{\mathcal{N}} - 0.5)\sigma_{\mathcal{N}}^{-1}$ .

<sup>7</sup>In probability theory, a *continuity correction* is an adjustment that is made when a discrete (probability) distribution is approximated by a continuous distribution. In particular, suppose that the continuous RV  $Y$  approximates the discrete RV  $X$ . Then,  $\mathbb{P}(X \leq m) = \mathbb{P}(X \leq m + 0.5) \approx \mathbb{P}(Y \leq m + 0.5)$ ,  $\forall m \in \mathbb{Z}$ .

TABLE II  
SUMMARY OF APPROXIMATION METHODS

Approximation Method	SOP Approximation Formula $\tilde{P}_{\text{out}}^{\text{sys}}(L, N)$	Parameters/Range of $L$	Condition for Higher Accuracy
Binomial Approximation (BA) <sup>a</sup>	$1 - \sum_{m=0}^{L-1} \binom{N}{m} \bar{p}^m \bar{q}^{N-m}$	$\bar{p} = \frac{1}{N} \sum_{n \in \mathcal{N}} p_n, \bar{q} = 1 - \bar{p}$	$(N\bar{p}\bar{q})^{-1} \sigma_{\mathcal{N}}^2 \rightarrow 1$
Poisson Approximation (PA) <sup>b</sup>	$1 - e^{-\mu_{\mathcal{N}}} \sum_{m=0}^{L-1} \frac{\mu_{\mathcal{N}}^m}{m!}$	—	$\sum_{n \in \mathcal{N}} p_n^2 \rightarrow 0$
Normal Approximation (NA)	$1 - \Phi(\zeta) = Q(\zeta)$	$\zeta = (L - \mu_{\mathcal{N}} - 0.5)\sigma_{\mathcal{N}}^{-1}$	$\sigma_{\mathcal{N}}^2 \rightarrow \infty$
Refined Normal Approximation (RNA)	$\min(\max(1 - G(\zeta), 0), 1)$		
Chernoff Bound (CB)	$(\mu_{\mathcal{N}}/L)^L e^{L - \mu_{\mathcal{N}}}$	$\forall L \in \{\lfloor \mu_{\mathcal{N}} \rfloor + 1, \lfloor \mu_{\mathcal{N}} \rfloor + 2, \dots, N\}$	—

<sup>a,b</sup> According to the numerical results (Section V), BA and PA are the most appropriate approximation methods for SGD systems operating in EHF bands.

#### D. Refined Normal Approximation (RNA)

Consider the following function:

$$G(x) = \Phi(x) + \nu_{\mathcal{N}}(6\sigma_{\mathcal{N}}^3)^{-1}(1 - x^2)\varphi(x) \quad (20)$$

According to [21]–[23], there exists a constant  $C < \infty$  such that  $\Delta'_{\mathcal{N}} = \sup_{s \in \mathbb{R}} |\mathbb{P}(S_{\mathcal{N}} \leq s) - G((s - \mu_{\mathcal{N}})\sigma_{\mathcal{N}}^{-1})| \leq C\sigma_{\mathcal{N}}^{-2} = O(\sigma_{\mathcal{N}}^{-2})$ . Observe that  $\lim_{N \rightarrow \infty} \Delta'_{\mathcal{N}} = 0$ , when  $\lim_{N \rightarrow \infty} \sigma_{\mathcal{N}}^2 = \infty$ . As a result, by applying the *continuity correction* once more, we obtain the following approximation:

$$P_{\text{out}}^{\text{sys}}(L, N) \approx \min\left(\max\left(\hat{P}_{\text{out}}^{\text{sys}}(L, N), 0\right), 1\right) \quad (21)$$

where  $\hat{P}_{\text{out}}^{\text{sys}}(L, N) = 1 - G(\zeta)$  and  $\zeta = (L - \mu_{\mathcal{N}} - 0.5)\sigma_{\mathcal{N}}^{-1}$ . Note that we make use of the above min-max formula in order to ensure that  $P_{\text{out}}^{\text{sys}}(L, N) \in [0, 1]$ , because  $\hat{P}_{\text{out}}^{\text{sys}}(L, N)$  may be outside the interval  $[0, 1]$  in some cases.

#### E. Chernoff Bound (CB)

A Chernoff (upper) bound can be constructed using a result given in [27] which states that:  $\mathbb{P}(S_{\mathcal{N}} \geq (1 + \delta)\mu_{\mathcal{N}}) \leq (e^{\delta}/(1 + \delta)^{1+\delta})^{\mu_{\mathcal{N}}}$ ,  $\forall \delta > 0$ . Specifically, by setting  $(1 + \delta)\mu_{\mathcal{N}} = L$  and assuming  $\mu_{\mathcal{N}} > 0$ , we obtain:

$$P_{\text{out}}^{\text{sys}}(L, N) \leq (\mu_{\mathcal{N}}/L)^L e^{L - \mu_{\mathcal{N}}} \quad (22)$$

which holds  $\forall L \in \{\lfloor \mu_{\mathcal{N}} \rfloor + 1, \lfloor \mu_{\mathcal{N}} \rfloor + 2, \dots, N\}$ , since  $\delta > 0 \Leftrightarrow L > \mu_{\mathcal{N}} \Leftrightarrow L > \lfloor \mu_{\mathcal{N}} \rfloor \Leftrightarrow L \geq \lfloor \mu_{\mathcal{N}} \rfloor + 1$ .

### V. NUMERICAL RESULTS AND DISCUSSION

In this section, all results present statistical averages derived from  $10^3$  independent system configurations, where the GW outage probabilities  $\{p_i\}_{i \in \mathcal{N} \cup \mathcal{K}}$  are uniformly distributed in  $(0, 0.02)$ , i.e., 98% to 100% link availability.

#### A. SOP Analysis

Firstly, we study the SOP as a function of the number of GWs,  $N$ , and the ratio of the traffic demand to the GW capacity,  $r$ . As shown in Fig. 1, the SOP increases with  $\lceil r \rceil$  for all values of  $N$ , which is in accordance with Proposition 1. Moreover, for any fixed  $\lceil r \rceil$ , we can observe that the SOP decreases with the increase of  $N$  (see Proposition 2).

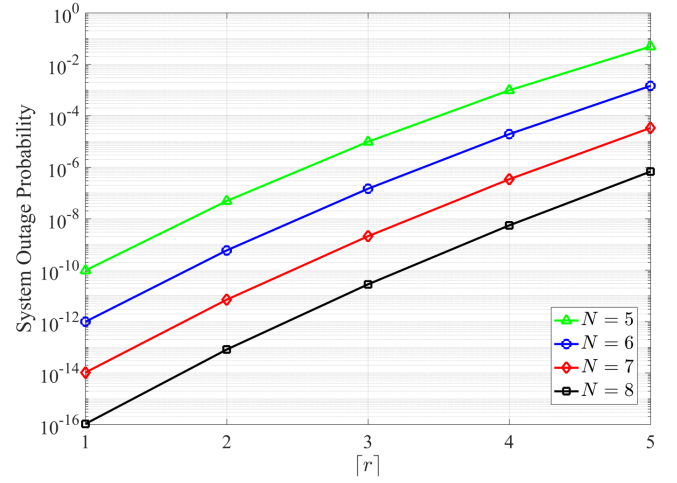


Fig. 1. System outage probability,  $P_{\text{out}}^{\text{sys}}$ , (calculated using Algorithm 1) versus the ceiling of  $r$  (the ratio of the traffic demand to the GW capacity).

Nevertheless, as mentioned at the end of Section II-C, this SOP improvement is achieved in exchange for higher connectivity complexity.

Secondly, we examine the performance enhancement achieved by a  $(5 + K)$ -GW compared to a 5-GW diversity system by means of the generalized SOP-improvement factor (where  $K \in \{1, 2, 3, 4\}$  is the number of additional GWs). Specifically, as illustrated in Fig. 2,  $I_g$  decreases with the increase of  $\lceil r \rceil$  for every value of  $K$ . Furthermore, for a given  $\lceil r \rceil$ , larger number of additional GWs results in higher performance improvement.

#### B. Performance of Approximation Methods

In order to evaluate the accuracy of a probability distribution and the tightness/sharpness of the Chernoff bound, we define the maximum absolute error (maxAE), the root-mean-square error (RMSE), and the mean absolute error (MAE) as follows:

$$\epsilon_{\text{max}}(N) = \max_{L \in \mathcal{S}} \left| P_{\text{out}}^{\text{sys}}(L, N) - \tilde{P}_{\text{out}}^{\text{sys}}(L, N) \right| \quad (23)$$

$$\epsilon_{\text{rms}}(N) = \sqrt{\frac{1}{|\mathcal{S}|} \sum_{L \in \mathcal{S}} \left( P_{\text{out}}^{\text{sys}}(L, N) - \tilde{P}_{\text{out}}^{\text{sys}}(L, N) \right)^2} \quad (24)$$

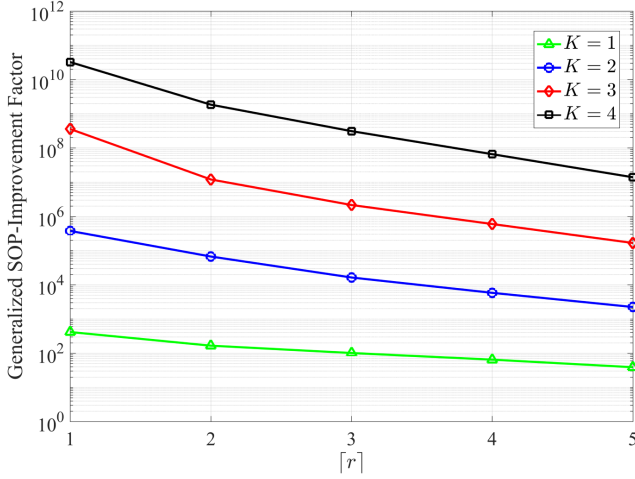


Fig. 2. Generalized SOP-improvement factor,  $I_g$ , (computed using Algorithm 1), in comparison with a diversity system consisting of  $N = 5$  GWs, versus the ceiling of  $r$  (the ratio of the traffic demand to the GW capacity).

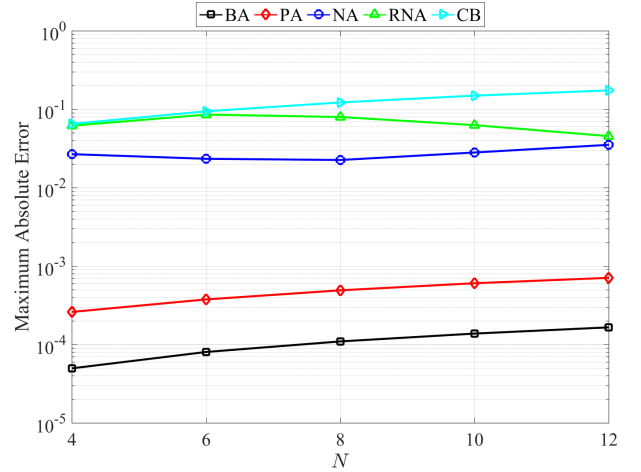
$$\epsilon_{\text{mean}}(N) = \frac{1}{|\mathcal{S}|} \sum_{L \in \mathcal{S}} \left| P_{\text{out}}^{\text{sys}}(L, N) - \tilde{P}_{\text{out}}^{\text{sys}}(L, N) \right| \quad (25)$$

where  $\tilde{P}_{\text{out}}^{\text{sys}}(L, N)$  is the approximate SOP. Moreover, for probability distributions  $\mathcal{S} = \mathcal{N}_0$  (with  $|\mathcal{S}| = N + 1$ ), while for CB  $\mathcal{S} = \{\lfloor \mu_N \rfloor + 1, \lfloor \mu_N \rfloor + 2, \dots, N\}$  (with  $|\mathcal{S}| = N - \lfloor \mu_N \rfloor \geq 1$ ). In general, it holds that  $\epsilon_{\text{max}}(N) \geq \epsilon_{\text{rms}}(N) \geq \epsilon_{\text{mean}}(N)$ .

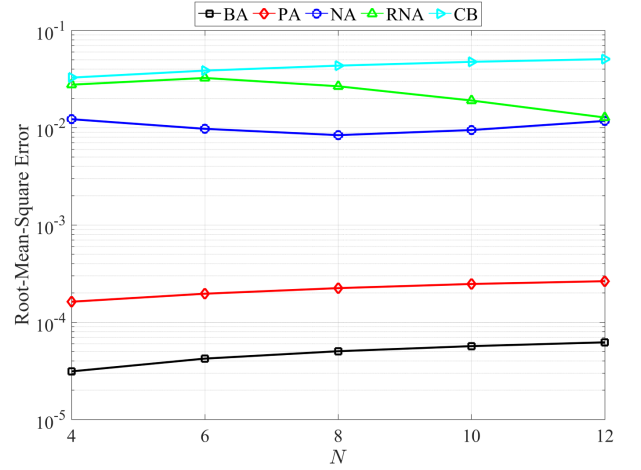
Fig. 3 presents the accuracy of approximation methods, in terms of maxAE, RMSE and MAE, versus the number of GWs. It can be observed that the approximation methods in descending-performance (or, equivalently, ascending-error) order are as follows: {BA, PA, NA, RNA, CB}. More specifically, BA and PA significantly outperform the other methods (the achieved errors are of the order of  $10^{-4}$  or  $10^{-5}$ ), while CB exhibits the lowest accuracy. At this point, we would like to give an explanation of the performance of BA, PA, NA and RNA. In practice, *the number of GWs is relatively small* ( $N \approx 4$  to  $7$ ) and *all the GW outage probabilities are very close to zero* (i.e.,  $p_n \approx 0$ ,  $\forall n \in \mathcal{N} \Rightarrow p_1 \approx p_2 \approx \dots \approx p_N$ ). As a result, the variance  $\sigma_N^2 = \sum_{n \in \mathcal{N}} p_n(1 - p_n)$  and the quantity  $\sum_{n \in \mathcal{N}} p_n^2$  are quite small, while  $\sigma_N^2 \approx N\bar{p}\bar{q}$  (see Section IV-A). Finally, according to Table II, it is clear that the condition for higher accuracy of BA/PA is well satisfied, whereas that of NA/RNA is not. In summary, BA and PA are the most suitable approximation methods for SGD systems.

## VI. CONCLUSION

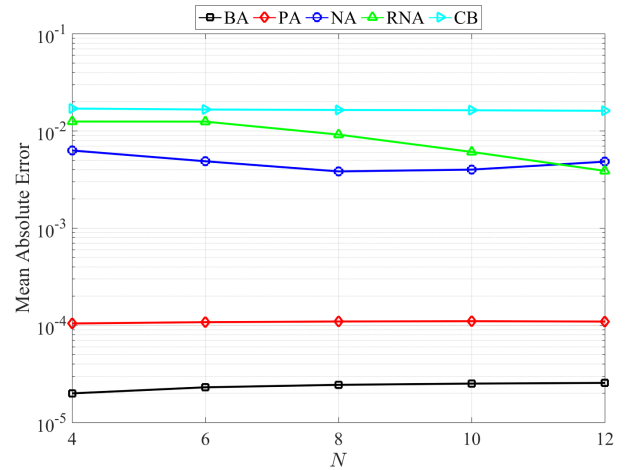
In this paper, we have studied in depth the LS-SGD scheme, which has been recently introduced in SatNets. Furthermore, a number of useful mathematical tools have been presented in order to compute and approximate the SOP. Finally, based on the numerical results, we conclude that the SOP can be well approximated by BA and PA, since these methods achieve remarkable accuracy. Such approximations may be useful for simplifying and solving hard optimization problems with SOP-constraints in SGD-based SatNets.



(a)



(b)



(c)

Fig. 3. Accuracy comparison of approximation methods: (a) maximum absolute error, (b) root-mean-square error, and (c) mean absolute error versus the number of GWs.



## APPENDIX

## A. Proof of Proposition 2

By virtue of the law/theorem of total probability, we obtain:

$$\begin{aligned}
 P_{\text{out}}^{\mathcal{N} \cup \mathcal{K}} &= \mathbb{P}(S_{\mathcal{N}} + S_{\mathcal{K}} \geq L + K) = \\
 &= \sum_{j=0}^K \mathbb{P}(S_{\mathcal{K}} = j) \mathbb{P}(S_{\mathcal{N}} + S_{\mathcal{K}} \geq L + K | S_{\mathcal{K}} = j) = \\
 &= \sum_{j=0}^K \mathbb{P}(S_{\mathcal{K}} = j) \mathbb{P}(S_{\mathcal{N}} \geq L + K - j) = \\
 &= \sum_{j=0}^K \mathbb{P}(S_{\mathcal{K}} = j) [\mathbb{P}(S_{\mathcal{N}} \geq L) - \mathbb{P}(L \leq S_{\mathcal{N}} \leq L + K - j - 1)] \leq \\
 &\leq \mathbb{P}(S_{\mathcal{N}} \geq L) \underbrace{\sum_{j=0}^K \mathbb{P}(S_{\mathcal{K}} = j)}_{=1} = \mathbb{P}(S_{\mathcal{N}} \geq L) = P_{\text{out}}^{\mathcal{N}}
 \end{aligned} \tag{26}$$

and the proposition follows.

## B. Proof of Theorem 2

Firstly, the initial conditions of the RF are trivially true. Secondly, from the law/theorem of total probability, the SOP  $P_{\text{out}}^{\text{sys}}(L, N) = \mathbb{P}(S_{\mathcal{N}} \geq L)$  can be written as follows:

$$\begin{aligned}
 P_{\text{out}}^{\text{sys}}(L, N) &= \sum_{j=0}^1 \mathbb{P}(X_N = j) \mathbb{P}(S_{\mathcal{N}} \geq L | X_N = j) = \\
 &= \sum_{j=0}^1 \mathbb{P}(X_N = j) \mathbb{P}(S_{\mathcal{N} \setminus \{N\}} \geq L - j) = \\
 &= \mathbb{P}(X_N = 0) \mathbb{P}(S_{\mathcal{N} \setminus \{N\}} \geq L) + \mathbb{P}(X_N = 1) \mathbb{P}(S_{\mathcal{N} \setminus \{N\}} \geq L - 1)
 \end{aligned} \tag{27}$$

where  $S_{\mathcal{N} \setminus \{N\}} = \sum_{n \in \mathcal{N} \setminus \{N\}} X_n = S_{\mathcal{N}} - X_N$ . Due to the fact that  $\mathbb{P}(X_N = 0) = 1 - p_N$  and  $\mathbb{P}(X_N = 1) = p_N$ , we get (14) and this completes the proof.

## REFERENCES

- [1] A. D. Panagopoulos *et al.*, "Satellite communications at Ku, Ka, and V bands: Propagation impairments and mitigation techniques," *IEEE Commun. Surv. & Tutor.*, vol. 6, no. 3, pp. 2-14, Third Quarter 2004.
- [2] A. D. Panagopoulos *et al.*, "Long-term rain attenuation probability and site diversity gain prediction formulas," *IEEE Trans. Antennas Propag.*, vol. 53, no. 7, pp. 2307-2313, July 2005.
- [3] C. I. Kourogiorgas *et al.*, "On the earth-space site diversity modeling: A novel physical-mathematical outage prediction model," *IEEE Trans. Antennas Propag.*, vol. 60, no. 9, pp. 4391-4397, Sept. 2012.
- [4] G. A. Karagiannis *et al.*, "Multidimensional rain attenuation stochastic dynamic modeling: Application to earth-space diversity systems," *IEEE Trans. Antennas Propag.*, vol. 60, no. 11, pp. 5400-5411, Nov. 2012.
- [5] A. D. Panagopoulos, "Propagation phenomena and modeling for fixed satellite systems: Evaluation of fade mitigation techniques," in *Radio wave propagation and channel modeling for earth-space systems*, Taylor & Francis Group, CRC Press, 2016.
- [6] N. Jeannin *et al.*, "Smart gateways for terabit/s satellite," *Int. J. Satell. Commun. Netw.*, vol. 32, no. 2, pp. 93-106, March 2014.
- [7] A. Kyrgiazos, B. G. Evans and P. Thompson, "On the gateway diversity for high throughput broadband satellite systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5411-5426, Oct. 2014.
- [8] A. Gharanjik *et al.*, "Multiple gateway transmit diversity in Q/V band feeder links," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 916-926, March 2015.
- [9] T. Rossi, M. De Sanctis and F. Maggio, "Evaluation of outage probability for satellite systems exploiting smart gateway configurations," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1541-1544, July 2017.
- [10] T. Delamotte and A. Knopp, "Smart diversity through MIMO satellite Q/V-band feeder links," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 285-300, Feb. 2020.
- [11] M. Muhammad, G. Giambene and T. de Cola, "QoS support in SGD-based high throughput satellite networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8477-8491, Dec. 2016.
- [12] C. N. Efreem and A. D. Panagopoulos, "Dynamic energy-efficient power allocation in multibeam satellite systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 228-231, Feb. 2020.
- [13] A. Z. Papafragkakis *et al.*, "ALPHASAT site diversity experiments in Greece and the UK at Ka band: Comparison of 2-years' results," *ITU Journal: ICT Discoveries*, vol. 2, no. 1, Nov. 2019.
- [14] T. Rossi *et al.*, "Smart gateway diversity optimization for EHF satellite networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 130-141, Feb. 2020.
- [15] C. N. Efreem and A. D. Panagopoulos, "Globally optimal selection of ground stations in satellite systems with site diversity," *IEEE Wireless Commun. Lett.*, to be published.
- [16] J. Pérez-Trufero *et al.*, "High throughput satellite system with Q/V-band gateways and its integration with terrestrial broadband communication networks," in *Proc. 32nd AIAA International Communications Satellite Systems Conference*, Aug. 2014, pp. 1-10.
- [17] M. Fernandez and S. Williams, "Closed-form expression for the Poisson-binomial probability density function," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 2, pp. 803-817, April 2010.
- [18] Y. Hong, "On computing the distribution function for the Poisson binomial distribution," *Computational Statistics & Data Analysis*, vol. 59, pp. 41-51, March 2013.
- [19] A. M. Rushdi, "Utilization of symmetric switching functions in the computation of k-out-of-n system reliability," *Microelectron. Reliab.*, vol. 26, no. 5, pp. 973-987, 1986.
- [20] L. A. Belfore, "An  $O(n(\log_2 n)^2)$  algorithm for computing the reliability of k-out-of-n:G and k-to-l-out-of-n:G systems," *IEEE Trans. Reliab.*, vol. 44, no. 1, pp. 132-136, March 1995.
- [21] P. Deheuvels, M. L. Puri, and S. S. Ralescu, "Asymptotic expansions for sums of nonidentically distributed Bernoulli random variables," *Journal of Multivariate Analysis*, vol. 28, no. 2, pp. 282-303, 1989.
- [22] V. G. Mikhailov, "On a refinement of the central limit theorem for sums of independent random indicators," *Theory Probab. Appl.*, vol. 38, no. 3, pp. 479-489, 1994.
- [23] A. Yu. Volkova, "A refinement of the central limit theorem for sums of independent random indicators," *Theory Probab. Appl.*, vol. 40, no. 4, pp. 791-794, 1996.
- [24] W. Ehm, "Binomial approximation to the Poisson binomial distribution," *Statist. & Probab. Lett.*, vol. 11, no. 1, pp. 7-16, Jan. 1991.
- [25] L. Le Cam, "An approximation theorem for the Poisson binomial distribution," *Pacific J. of Math.*, vol. 10, no. 4, pp. 1181-1197, 1960.
- [26] J. M. Steele, "Le Cam's inequality and Poisson approximations," *The American Mathematical Monthly*, vol. 101, no. 1, pp. 48-54, 1994.
- [27] T. Hagerup, and C. Rüb, "A guided tour of Chernoff bounds," *Information Processing Letters*, vol. 33, no. 6, pp. 305-308, Feb. 1990.