IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 30, 2018, accepted June 5, 2018, date of publication June 11, 2018, date of current version June 26, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2845453

Resource Allocation With Minimum Outage Probability in Multicarrier Multicast Systems

DUC-PHUC VUONG[1], MINH-QUAN DAO[1], AND DANG-KHANH LE[2]
[1]Department of Electric and Electronics Engineering, Vietnam Maritime University, Haiphong 180000, Vietnam
[2]VMU College, Vietnam Maritime University, Haiphong 180000, Vietnam

Corresponding author: Dang-Khanh Le (khanh.ld@vimaru-vmc.edu.vn)

ABSTRACT Several innovative applications are emerging that require significant multimedia data transmissions, which unfortunately, current communication systems struggle to provide. In this paper, we examine the high outage tendencies of conventional multimedia multicasting (CVM) due to the bottleneck imposed by the user with minimum signal-to-noise ratio (SNR) and we present a technique based on the statistical approximations of the amortized weighted averaging (AWG) of users' SNR. For completeness, we propose a suboptimal multicast resource allocation algorithm using memoization with stochastic rounding. We compare the system performance of AWG scheme with the CVM scheme in multicarrier wireless network using our newly derived system outage probability and average throughput as performance metrics. Numerical results not only show that the performance gap of our suboptimal algorithm is reasonably within 10% from the optimal solution, but also show that the AWG scheme admits more users and is more energy efficient than the CVM especially when low power is available to the system.

INDEX TERMS Multicasting, outage probability, weighting factor, memoization, stochastic rounding, multimedia broadcast multicast systems (MBMS), MSR, CVM, minSNR, AWG.

I. INTRODUCTION

Radio spectrum is a scarce resource. The heavy demands for high data rates and the need to support large number of users with flexible quality of service (QoS) requirements implies a large number of future wireless devices must compete for the limited resources. Therefore, mobile phone operators worldwide have been hogging their data networks like a fat kid clutching a candy jar in an attempt to satisfy the ever-increasing needs of subscribers. The need is not expected to reduce anytime soon, in fact, it has been predicted that global mobile data traffic will explode very soon. Interestingly, video data from mobile-connected devices will be the most dominant traffic constituting almost 70 percent of the whole global multimedia data transmission [1].

Future wireless communication system is expected to support several disruptive applications which may require transmission to selected groups of users within close proximity, similar interest, or channel quality. One key enabling technology expected to facilitate the widespread adoption of these interesting group-based applications is multimedia multicasting [2]. The technology have been widely adopted as evolved Multimedia Broadcast Multicast Services (eMBMS) for future wireless cellular standards such as 3GPP Long Term Evolution-Advanced (LTE-A) to provide high rates for nomadic and mobile users [3]. The eMBMS is a point-to-multipoint interface specification for current and future cellular network which allows multiple users within the same or adjacent cell requesting similar multimedia contents to form groups and share allocated system resources. The technology will support geographic information updates such as traffic reports, local news, weather forecast, stock prices, and location-based adverts as well as multimedia entertainments such as IPTV, mobile TV, video-conferencing, and other related services [4]–[6]. The technology has been proved to both maximize spectral efficiency and resource utilizations [3], [7].

For decades, researchers have assumed that to avoid service outage in conventional MBMS (CVM), the group throughput and quality of services (QoS) should be defined in terms of the rates at which the user with the minimum signal-to-noise-ratio (minSNR) in the group or at the cell edge can decode successfully. While the minSNR approach effectively ensures users in the group achieve uniform success rate, it has been shown as an inefficient solution especially when minimum rate for guaranteed QoS and error-free decoding are required. In other words, the minSNR approach in CVM evidently allows all users in the group to receive transmitted data; but at the system level, it quickly leads to system saturation because it is too conservative, pessimistic, and cannot exploit multiuser channel diversities to improve network performance [8]–[10].

To address this problem, resource optimization techniques are often employed with a view of maximizing the system capacity. Solving such problem typically lends itself as a constrained optimization which may result in higher computation complexity and increased system design overhead. So far, there has been no easy way to utilize the minSNR approach in CVM to improve system performance without introducing optimization complexities [6], [8], [11].

Our focus in this paper is two-fold: First, we study a novel multicast transmission scheme where each user's SNR is weighted with a stochastic weighting coefficient. The main idea is to obtain a systematic *amortized*[1] *weighted averaging (AWG)* [10]. We examine the scheme and propose its closed-form approximation and then derive a new performance metric for evaluation. Secondly, we propose an inherently fast QoS-aware multicast resource allocation algorithm which allows us to evaluate the performance of AWG scheme in comparison with the existing minSNR scheme. To the best of our knowledge, this is the first comprehensive work in this domain, providing alternative single-rate multicasting scheme to improve system-level outage performance in multicarrier systems.

The remainder of this paper is structured as follows: Section II highlights existing works related to our research. Section III describes the system model with discussions on the conventional scheme and AWG-based SNR scheme. Section III-C introduces the statistical closed-form approximations for the AWG scheme. Section V analyzes the system performance of both AWG-based and CVM-based SNR in terms of the system average throughput and outage probability. Section VI describes the resource allocation problem and presents the proposed QoS-aware suboptimal allocation algorithm with memoization and stochastic rounding (MSR). Section VII compares the performance of AWG scheme with the CVM scheme and presents some performance evaluation results for both broadcast and multicast systems. Finally, section VIII provides a succinct conclusion of the paper and draws out some final remarks.

II. RELATED WORKS

Existing body of works in multimedia multicasting can be divided into *single-rate* and *multi-rate* transmission schemes. In single-rate, the BS transmits to all users in each multicast group at the same rate irrespective of their non-uniform capabilities, whereas in multi-rate, the BS transmits to each user in each multicast group at different rates based on what each user can decode [4]–[6], [8], [9], [11]–[13].

Across literatures, single-rate transmission has been quite popular and widely accepted due to its implementation simplicity. Single-rate multicast services must be transmitted at a rate low enough for user with minimum SNR to decode and high enough to maximally utilize system resources. Hence, the major problem becomes how to determine the most efficient rate to transmit to each group without being insensitive to users with bad channel quality or unfair to users with high throughput potentials.

One way to resolve this problem in cellular networks is to transmit based on user with the minimum SNR (minSNR) in each group. In fact, several existing works have used variations of minSNR as their underlining approach for multicasting. Shrestha *et al.* [14] proposed a MaxiMin-type optimization technique that attempts to transmit same data on multiple channel pair such that each group receives the *minimum rates* out of possible maximum. But channel pairing also reduces usable resources by almost 50%. Liu *et al.* [15] proposed a group partitioning with MaxiMin-type technique where users are partitioned into groups based on comparable pathloss instead of their data requirements. The *minimum SNRs* in each partition are selected as baselines. However, the approach also requires more signal processing overhead and multiple transmission of same data. Note that the aforementioned schemes use minSNR as baseline (cf. [8], [9], [14], eq.(10), [15], eq.(1)). Hence, we categorize their approaches as MaxiMin-based conventional MBMS (CVM) scheme.

A not very popular scheme is to transmit based on the the average or median SNR [16]–[18]. This centrality-based approach guarantees delivery to almost 50% of users and potentially provides higher capacity than minSNR scheme. However, the approach is also too optimistic as packet loss may be inevitable especially for weak users that are far from the BS. Other ideas have also been proposed for multi-rate multicast transmission; however, multi-rate techniques require difficult coding and synchronization complexities unlike single-rate schemes which are less complex and more practical to implement. Note that most of the the aforementioned efforts on wireless multicasting have focussed on group level capacity enhancement. Very few works have considered the system level impact on the network in terms of outage, coverage and reliability [10], [19]. *Our current work is distinct from [10] in particular as the authors have not presented a closed-form expression for AWG and they only considered outage in single carrier system, thus, do not require a resource allocation algorithm. In what follows, we propose an approximation for the AWG scheme and also propose a new multicast resource allocation algorithm to study the system-level impact of both CVM and AWG over a multicarrier wireless cellular network.*

III. SYSTEM MODEL

In this section, we first provide a detailed system description used in this work. Subsequently, we briefly explain the mathematical model for the existing minSNR-based CVM scheme and then provide a description of the new AWG scheme.

[1] monotonic reduction in value through systematic spreading.

A. SYSTEM DESCRIPTION

Consider an OFDMA-based cellular network with a base station (BS) and user $k \in \kappa_g$, where κ_g is the set of users in multicast group g receiving downlink traffic flow from the central base station. Number of users in each group is K_g and total users in the system is $T = \sum_{g=1}^{G} K_g$, where G is the total number of groups. Each group g has fixed or variable number of users who may be very closely located or sufficiently far-away located with different channel characteristics.

Similarly, there are N_s total number of available subcarriers and P_{Tot} is the system power available to the central BS. The transmit power on each subcarrier n is denoted as P_n. Each subcarrier has equal bandwidth spacing of $B_W = \frac{W}{N_0\ N_s}$, where W is the total bandwidth of the system and N_0 is the single-sided power spectral density of the white noise per unit of subcarrier. The SNR of each user is denoted as a random variable (r.v.) X. For each data transmitted to group g on subcarrier n, the channel coefficient is denoted as $c_{g,n}$. The wireless link is assumed to be an i.i.d. block Rayleigh fading channel. This means, $c_{g,n}$ is constant over time slot and changes independently across subcarriers according to a circular symmetric complex Gaussian distribution with zero mean and variance $\gamma_{g,n}^2$, i.e. $c_{g,n} \sim \mathcal{CN}(0, \gamma_{g,n}^2)$. It can therefore be verified that $X_{k,g,n}$ becomes an i.i.d. exponential r.v., $X_{k,g,n} \sim Exp(\lambda_{g,n})$ where $\lambda_{g,n} = \frac{P_n \gamma_{g,n}^2}{N_o B_W}$ is the average SNR of group g on subcarrier n - denoted (g, n).

For simplicity and notational conveniences, we assume equal power distribution on all subcarriers with unity N_o and B_W. Hence, we have $\lambda_{g,n} = P\gamma_{g,n}^2$, where $P = P_n = \frac{P_{Tot}}{N_s}$. In addition, when $G = 1$, we have a broadcast system, consequently, we can omit subscripts such that $X_{k,g,n} = X_k$ and $\lambda_{g,n}$ is simply λ. Moreover, for $G > 1$, for simplicity, we assume a frame-based system in which decisions on multicast throughput are made at the beginning of each time slot and a user who cannot correctly decode received data may have better channel gain in the next frame transmission [14].

B. PDF OF MINSNR-BASED CONVENTIONAL MULTICASTING (CVM)

In conventional multicasting (CVM), data must be transmitted at a rate low enough for users with worst channel gains to decode and high enough to efficiently utilize system resources. Therefore, the minSNR is often assumed for the group. Let $X_{min,n} = \min_{k \in \kappa_g} X_k$ be the r.v. representing the SNR of the group transmitting on subcarrier n. Using ordered statistics, it can be verified that the PDF $f_{X_{min}} = f_{X_{(1)}}$, where $X_{(1)} = X_{min} = Min\left[X_1, \ldots, X_k \ldots, X_{K_g}\right]$ is the first ordered statistics. In general, for set of i.i.d. r.v. X, the order of SNR of all users is denoted as: $X_{(1)}, \ldots, X_{(r)}, \ldots, X_{(K_g)}$, where $X_{(r)}$ is the r-th order of $X_{k \in \kappa_g}$. Thus, for any τ, $f_{X_{(\tau)}}$ is given as:

$$f_{X_{(\tau)}}(x) = \frac{n! f(x) F(x)^{\tau-1} (1-F(x))^{n-\tau}}{(\tau-1)!(n-\tau)!}. \quad (1)$$

SNR	X_1	...	X_r	...	X_{K_g-1}	X_{K_g}	Ascen.

$\otimes$

WMF	Y_{K_g}	...	$Y_{\tilde{r}=K_g-r+1}$	...	Y_2	Y_1	Descen.

FIGURE 1. Schematic of the weighting coefficient. r and $\tilde{r}$ are post-sort index.

Hence, using (1), $f_{X_{(r)}}(x)$ can be computed as:

$$f_{X_{(r)}}(x) = \frac{Q_r}{\lambda} Exp[-\frac{x(K_g+1-r)}{\lambda}] \left(1 - e^{-\frac{x}{\lambda}}\right)^{r-1}, \quad (2)$$

where $Q_r = \frac{K_g!}{\Gamma(r)\Gamma(K_g-r+1)}$. When $r = 1$ in (2), we obtain PDF $f_{X_{min}}$ for conventional multicasting. Coincidentally, $f_{X_{min}}$ is also exponential with parameter $\left(K_g \lambda_{g,n}\right)$. Note that the ordered statistics are dependent and non-identical.

C. CDF & PDF OF AWG-BASED SNR SCHEME (AWG)

In this section, we provide a step-wise overview of the AWG-based SNR scheme and highlight the procedure towards obtaining its PDF and CDF [10]. We also present some new key results on the statistics of the scheme.

1) THE AMORTIZED WEIGHTED AVERAGING (AWG)

Let $X_{(r)}$ be users' SNR arranged in ascending order, r being the post-sort index of X_k. Also, let $Y_{\tilde{r}}$ denotes a *positive-value stochastic amortized weighting coefficient* with PDF $f_{Y_{(r)}}$ sorted in descending order as shown in Fig. 1. A user's weighted SNR is then defined as:

$$Z_r = X_{(r)} \times Y_{\tilde{r}=K_g-(r)+1} \quad for \quad r = 1, 2, \ldots, K_g, \quad (3)$$

where $f_{Y_{(r)}}(y)$ can be computed from (1) as:

$$f_{Y_{(r)}}(y) = Q_r (1-y)^{r-1} y^{K_g-r}. \quad (4)$$

From eq. (3), we note that the SNR of each user in the group after ranking is effectively weighted by the corresponding $Y_{\tilde{r}}$ to offset the impact on other users in the group. For special cases where $Y_{\tilde{r}} = 1$, the minSNR $X_{(1)}$ is unaffected and when $Y_{\tilde{r}} = 0$, the highest SNR $X_{(K_g)}$ is treated as outlier which make sense in multicasting.

The AWG-SNR for each multicast group is then defined as:

$$H = \frac{1}{K_g} \sum_{r=1}^{K_g} Z_r. \quad (5)$$

To obtain PDF f_H, we should obtain f_{Z_r} for the r.v. Z_r which is the product of two i.i.d. r.v. $X_{(r)} \sim Exp(K_g \lambda_{g,n})$ and $Y_{\tilde{r}}$. Note that the derivation of density f_{Z_r} is not straightforward due to the orderedness of the random variates. Here, for simplicity, we model $Y_{\tilde{r}}$ as a continuous standard uniform r.v. with $Y_{\tilde{r}} \sim Unif[0, 1]$ as proposed in [10].

2) PDF $F_{Z_R}(z)$ and MGF M_{Z_r} of User Weighted SNR Z_r

Let $X_{(r)} \sim f_{X_{(r)}}$ and $Y_{(r)} \sim f_{Y_{(r)}}$ be two independent r.v. The PDF $f_{Z_r}(z)$, for $Z_r = X_{(r)} \cdot Y_{\tilde{r}}$ is given as:

$$f_{Z_r}(z) = \sum_{m=0}^{r-1}\sum_{n=0}^{r-1} A\left(\frac{\lambda}{wz}\right)^{-v}\Gamma\left(-v, \frac{wz}{\lambda}\right), \quad (6)$$

where $\underset{m,n,r}{A} = \frac{(K_g!)^2(-1)^{m+n}\binom{r-1}{m}\binom{r-1}{n}}{\lambda(\Gamma(r))^2(\Gamma(K_g-r+1))^2}$, $\underset{m,r}{v} = (K_g + m - r)$ and $\underset{n,r}{w} = (K_g + n - r + 1)$. $\Gamma(\cdot)$ is the incomplete Gamma function; m and n are the indices of the binomial expansions.

Similarly, let the MGF of $M_{Z_r}(-s) = E[e^{-sz}]$. Then, $M_H(-s)$ for r.v. H defined in (5) becomes:

$$M_H(-s) = \prod_{r=1}^{K_g} M_{Z_r}\left(-\frac{s}{K_g}\right) \quad (7)$$

$$= \prod_{r=1}^{K_g}\sum_{m=0}^{r-1}\sum_{n=0}^{r-1} P\,{}_2F_1(1, v+1; v+2; -u), \quad (8)$$

where $\underset{m,n,r}{P} = \frac{(K_g!)^2(-1)^{m+n}\binom{r-1}{m}\binom{r-1}{n}}{(v+1)w(\Gamma(r))^2(\Gamma(K_g-r+1))^2}$, $\underset{m,r}{v} = (K_g + m - r)$ and $\underset{n,r}{u} = \frac{\lambda s}{K_g w}$. Symbol ${}_2F_1(\cdot;\cdot;\cdot;\cdot)$ is the Gauss hypergeometric series [20]. For detailed proof, interested readers are invited to review [10]. As proof of concept, the plots of (6) for values of r and specific $r = 3$ are shown in Fig. 2 and Fig. 3.

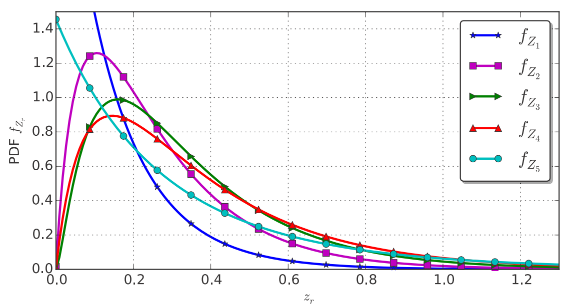

FIGURE 2. Derived $f_{Z_{(r)}}$ in (6) plotted for $r = \{1\ldots5\}$, $K_g = 5$ and $\lambda = 2$. Note that at $r = 1$ and $r = K_g$, $f_{Z_{(r)}}$ appear as exponential function.

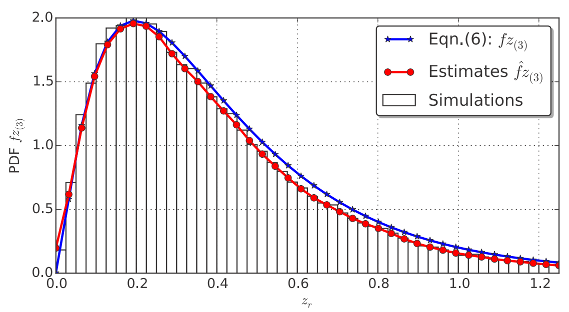

FIGURE 3. As a double check, the PDF $f_{Z_{(3)}}$ in (6) was validated with Monte Carlo simulation for $r = 3$, $K_g = 5$ and $\lambda = 1$.

Finally, the CDF of the AWG-based SNR of group (g, n) is derived using the differential property of Laplace transform on the moment generating function (MGF) of $H_{g,n}$ written as:

$$F_{H\,g,n}(h) = \mathcal{L}^{-1}\left[\frac{M_{H\,g,n}(-s)}{s}\right], \quad (9)$$

$$f_{H\,g,n}(h) = \mathcal{L}^{-1}[M_{H\,g,n}(-s)]. \quad (10)$$

where $\mathcal{L}^{-1}[\cdot]$ denotes the familiar inverse Laplace transform. Substituting (8) into (9) and solving with a multi-precision Laplace inversion method gives the CDF $F_{H\,g,n}(h)$ and PDF $f_{H\,g,n}(h)$ [21]. Due to the complexity of the MGF $M_{H\,g,n}(-s)$ in eq. (8), Afolabi *et al.* [10] could not directly obtain a closed-form expression for $F_{H\,g,n}(h)$ and $f_{H\,g,n}(h)$. In the next section, we propose a statistical approximation method which allows us to obtain an approximate closed-form.

IV. PROPOSED STATISTICAL APPROXIMATION FOR AWG

One solution we use in this work to obtain an approximate closed-form is to evaluate the inverse Laplace transform $f_H(h) = \mathcal{L}^{-1}[M_H(-s)]$ numerically and then develop a nonlinear model fitting equation from the interpolated data-points using the Levenberg-Marquardt nonlinear least square algorithm to obtain expressions $f_H^\star(h)$ and $F_H^\star(h)$ denoting the approximation of $f_H(h)$ and $F_H(h)$ in (9) and (10). Considering the shape of the resulting numerical evaluations of $f_H(h)$ and $F_H(h)$, we assume a Lognormal distribution for the model equation. We observe that since the numerical data do not contain random error, outliers or scatter, we can reliably interpolate the datapoints, capture the trends, and obtain a statistical approximate equations for $f_H^\star(h)$ and $F_H^\star(h)$. Our use of Lognormal distribution for the approximation is particularly motivated by recognizing the fact that a variable may follow the Lognormal distribution if it can be expressed as a product of r.v.s [22], [23]. More specifically, the logarithm of the product of r.v.s is a sum of r.v.s which tends to follow the Gaussian distribution as the number of r.v.s goes large. Therefore, the product of r.v.s tends to follow the Lognormal distribution. As can be seen in (7), H is such r.v. composed of products of other r.v. and thus directly lends itself towards a Lognormal model.

Suppose, $\Phi[\centerdot]$ is the transformation function, then the procedure for the statistical approximation can be succinctly explained for multicast groups in a multicarrier system as:

$$\Phi[\centerdot] : f_{H_{g,n}}(h) \rightarrow f_{H_{g,n}}(h, K_g, \lambda_{g,n}) \rightarrow H_{g,n} \sim \mathcal{LN}(\mu_{g,n}, \sigma_{g,n}) \quad (11)$$

where $\mu_{g,n}$ and $\sigma_{g,n}$, are respectively, the location and scale parameters which are obtained for every datapoint of the distribution. The approximate equations finally becomes:

$$F_H^\star(h) = \Phi\left[\mathcal{L}^{-1}\left[\frac{M_H(-s)}{s}\right]\right] = \frac{1}{2}\left(1 - \mathrm{Erf}\left[\frac{\mu_{g,n} - \log(h)}{\sqrt{2}\sigma_{g,n}}\right]\right) \quad (12)$$

$$f_H^\star(h) = \Phi\left[\mathcal{L}^{-1}[M_H(-s)]\right] = \frac{Exp\left[-\frac{(\mu_{g,n}-\log(h))^2}{2\sigma_{g,n}^2}\right]}{\sqrt{2\pi}h\sigma_{g,n}} \quad (13)$$

where Erf[·] is the error function. With $H_{g,n} \sim F^{\star}_{H_{g,n}}(h)$, we have the distributions of the instantaneous AWG-based SNR of each group (g, n). The plots of the numerical evaluation vs. the approximate form for the PDFs $f^{\star}_{H}(h), f_{H}(h)$ and CDF $F^{\star}_{H}(h), F_{H}(h)$ are presented in Fig. 4 and Fig. 5 respectively. It is observed that our approximations follow the numerical values excellently.

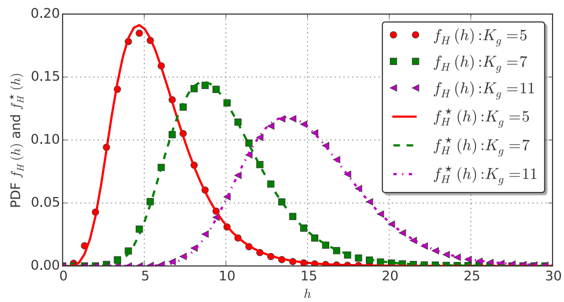

FIGURE 4. Numerical $f_H(h)$ vs Approx.$f^{\star}_H(h)$ for group g with different K_g.

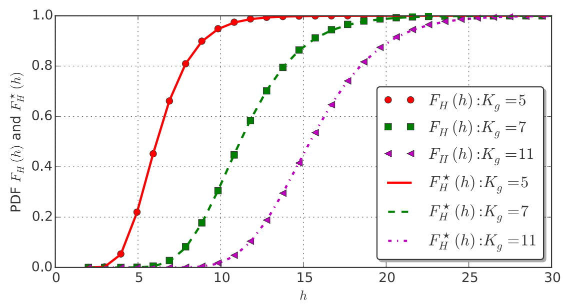

FIGURE 5. Numerical $F_H(h)$ vs Approx.$F^{\star}_H(h)$ for group g with different K_g.

V. PERFORMANCE ANALYSIS OF CVM & AWG

In this section, we analyze the performance of both CVM and AWG schemes using two metrics: the *average throughput* and *average outage probability*. Both metrics are complementary. The throughput shows the possible *system experience* while outage drills deeper and allows us to gain better insight into *why* certain behaviour occur.

A. AVERAGE THROUGHPUT ACROSS ALL ALLOCATED SUBCHANNELS

Let $^\circ$ denote placeholder V (for CVM) or W (for AWG), then aggregate throughput R of group (g, n) is given as:

$$R^{\circ}_{g,n} = Q(\tau_{g,n}) = K_g \text{Log}_2\left(1 + \tau_{g,n}\right), \tag{14}$$

where $\tau_{g,n}$ is the SNR for V or W on group (g, n). Using transformation of the distribution of an r.v., the PDF $f_{R^{\circ}_{g,n}}$ can be derived. Also, average throughput $E[R^{\circ}_{g,n}] = \int_0^{\infty} r^{\circ} f_{R^{\circ}_{g,n}}(r^{\circ})\, dr^{\circ}$ can directly be derived as [16]:

$$E[R^{\circ}_{g,n}] = \int_0^{\infty} Q(\tau_{g,n}) f_{\tau_{g,n}}(\tau_{g,n}) d\tau_{g,n}. \tag{15}$$

1) AWG AVERAGE THROUGHPUT

The average throughput of a group (g, n) with K_g users is $E\left[R^W_{g,n}\right]$. Recalling (14), we have $R^W_{g,n} = K_g \text{Log}_2\left(1 + H_{g,n}\right)$. Thus, $f_{R^W_{g,n}}(w)$ becomes:

$$f_{R^W_{g,n}}(w) = \frac{\text{Log}[2]\sqrt{2}\mathcal{S}}{\mathcal{S} K_g \sqrt{\pi}\sigma_{g,n}} Exp\left[-\frac{(\mu_{g,n} - \text{Log}[\mathcal{S}])^2}{2\sigma^2_{g,n}}\right], \tag{16}$$

where $\mathcal{S} = 2^{\frac{w}{K_g}} - 1$. Note that intergroup subcarrier sharing is disabled in this work since users in each group may have unique QoS and multimedia data requirement. Hence, to preserve channel orthogonality and ensure that not more than one group can be allocated to a single subcarrier, we define a subcarrier allocation index $\beta_{g,n}$ which indicates if a subchannel n has been allocated to multicast group g:

$$\beta_{g,n} = \begin{cases} 1, & \text{if subcarrier } n \text{ is allocated to group } g. \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

The *group-level, per channel average throughput* is given as:

$$E[R^W_{g,n}] = \frac{K_g \beta_{g,n}}{2\text{Log}[2]}\left(A + \sqrt{\frac{2}{\pi}}\sigma_{g,n} B + C\right), \tag{18}$$

where $A = \mu_{g,n}\text{Erf}\left[\frac{\mu_{g,n}}{\sqrt{2}\sigma_{g,n}}\right]$, $B = e^{-\frac{\mu^2_{g,n}}{2\sigma^2_{g,n}}}$ and $C = \sqrt{\frac{1}{\sigma^2_{g,n}}}\mu_{g,n}\sigma_{g,n}$. The system average throughput of G groups across all allocated N_s subcarriers is therefore:

$$C^W_M = \sum_{g=1}^{G}\sum_{n=1}^{N_s} E[R^W_{g,n}], \tag{19}$$

2) CVM AVERAGE THROUGHPUT

To obtain equivalent form for $f_{R^V_{g,n}}(v)$, $E[R^V_{g,n}]$ and C^V_M, we simply substitute $f_{X_{min,n}}$ for $f_{\tau_{g,n}}$ in (14),(15) and we can directly obtain:

$$f_{R^V_{g,n}}(v) = \frac{\mathcal{T}\text{Log}[2]}{P\gamma^2_{g,n}}\text{Exp}\left[\mathcal{V}(1 - \mathcal{T})\right] \tag{20}$$

$$E[R^V_{g,n}] = -\frac{K_g \beta_{g,n} e^{\mathcal{V}}}{\text{Log}[2]}\text{Ei}\left[-\mathcal{V}\right] \tag{21}$$

$$C^V_M = \sum_{g=1}^{G}\sum_{n=1}^{N_s} E[R^V_{g,n}] \tag{22}$$

where Ei[·] is the exponential integral function, $\mathcal{T} = 2^{\frac{v}{K_g}}$ and $\mathcal{V} = \frac{K_g}{P\gamma^2_{g,n}}$. Observe that average throughput C^W_M (19) for AWG is a function of $\mu_{g,n}$ and $\sigma_{g,n}$ computed from the approximation process in (11)-(13), while the average throughput C^V_M for CVM is directly derivable from $f_{X_{min}}(x)$ explained in (2).

B. SYSTEM AVERAGE OUTAGE PROBABILITY

Since AWG uses a modified averaging scheme, there is a possibility some users in the group may not be able to successfully decode received data. Hence, in this subsection, we examine the performance of AWG for broadcast and multicast systems in terms of outage probability.

1) OUTAGE PROBABILITY FOR BROADCAST SYSTEM

The broadcast outage measures the probability that *at least one* of the users in the multicast group cannot satisfy the minimum target data rate R_{qos} over all N_s subcarriers. We define R_{qos} as the minimum threshold required for guaranteed QoS. Users in group (g, n) can successfully decode received signal when $\left(R^{\circ}_{g,n} \geq \rho\right)$, where $\rho = \frac{R_{qos}}{N_s}$ is the average target rate per subcarrier. Probability that all user K_g correctly receive transmitted data in a broadcast system is then given as:

$$Pr\left(R^{\circ}_{g,n} \geq \rho\right)^{K_g} = \left[1 - F_{R^{\circ}_{g,n}}(\rho)\right]^{K_g}. \tag{23}$$

Broadcast outage where at least one user in group g cannot decode signals transmitted on the N_s subchannels can be derived:

$$\begin{aligned} G^{\circ}_{out} &= \frac{1}{N_s}\sum_{n=1}^{N_s} 1 - \left(1 - \beta_{g,n}\int_0^{\rho} f_{R^{\circ}_{g,n}}(r^{\circ})\, dr^{\circ}\right)^{K_g} \\ &= \frac{1}{N_s}\sum_{n=1}^{N_s} 1 - \left(1 - \beta_{g,n}Q^{\star}_{\rho}\right)^{K_g}, \end{aligned} \tag{24}$$

where $Q^{\star}_{\rho} = \int_0^{\rho} f_{R^{\circ}_{g,n}}(r^{\circ})\, dr^{\circ}$. Note that for broadcast system with $G = 1$, all N_s subchannel are used to facilitate transmissions of a single group.

2) OUTAGE PROBABILITY FOR MULTICAST SYSTEM

Similar to the above, multicast outage occur when certain multicast groups are locked out of resource allocation because the groups are deemed incapable of satisfying the average target rate on all allocated subchannels. Thus, outage probability for the multicast system is given as:

$$Pr\left(R^{\circ}_{g,n} \leq \rho\right) = F_{R^{\circ}_{g,n}}(\rho) = \int_0^{\rho} f_{R^{\circ}_{g,n}}(r^{\circ})\, dr^{\circ}. \tag{25}$$

Consequently, the system average outage probability S°_{out} of G multicast groups over all N_s subchannels becomes:

$$S^{\circ}_{out} = \frac{1}{GN_s}\sum_{g=1}^{G}\sum_{n=1}^{N_s} \beta_{g,n}Q^{\star}_{\rho}. \tag{26}$$

Substituting $f_{R^{V}_{g,n}}(v)$ and $f_{R^{W}_{g,n}}(w)$ for $f_{R^{\circ}_{g,n}}(r^{\circ})$ in (24) and (25), we can finally obtain both CVM S^{V}_{out} and AWG S^{W}_{out} for broadcast and multicast system average outage respectively. However, due to difficulty in derivation, we do not present a closed-form expression for S°_{out}, however, given all relevant parameters, a numerical evaluation is feasible.

VI. QOS-AWARE MULTICAST RESOURCE ALLOCATION

In this section, we first present details of the resource allocation problem and then proposed a low-complexity suboptimal algorithm using memoization and stochastic rounding (MSR). The resulting algorithm is used in evaluating the system performance.

A. MINIMIZING THE AVERAGE SYSTEM OUTAGE PROBABILITY

The system average outage probability minimization problem with G groups and N_s subchannels is formulated as:

$$\min_{\beta_{g,n}} \frac{1}{GN_s}\sum_{g=1}^{G}\sum_{n=1}^{N_s} 1 - \left(1 - \beta_{g,n}Q^{\star}_{\rho}\right)^{K_g} \tag{27a}$$

subject to:

$$\sum_{g=1}^{G} \beta_{g,n} = 1, \quad \forall n \tag{27b}$$

$$\beta_{g,n} \in [0, 1] \quad \forall g, \forall n \tag{27c}$$

$$\sum_{n=1}^{N_s} \beta_{g,n}Log_2\left(1 + P\gamma^2_{g,n}\right) \geq R_{qos} \quad \forall g, \tag{27d}$$

where (27c) is the integer constraint $\beta_{g,n}$ indicating which subchannel is allocated to each group, (27b) shows the exclusive allocation of subcarriers to unique groups to prevent co-channel interference. The fairness measure in (27d) guarantees that aggregate data rate of group g on all allocated subchannels satisfies the minimum target rate R_{qos} constraints. Observe also that equal power P in (27d) introduces suboptimality, but it enables us to find suitable AWG approximations in (11).

B. MAXIMIZING THE AVERAGE THROUGHPUT

The rate maximization problem with G groups and N_s subchannels is formulated similar to problem (27a)-(27d) by replacing (27a) with (28a) but under the same constraint equations. The goal here is to maximize the average throughput such that the QoS constraints and additional constraints on the system model are satisfied, i.e.

$$\max_{\beta_{g,n}} \sum_{g=1}^{G}\sum_{n=1}^{N_s} E[R^{\circ}_{g,n}]. \tag{28a}$$

subject to:

$$\sum_{g=1}^{G} \beta_{g,n} = 1, \quad \forall n \tag{28b}$$

$$\beta_{g,n} \in [0, 1] \quad \forall g, \forall n \tag{28c}$$

$$\sum_{n=1}^{N_s} \beta_{g,n}Log_2\left(1 + P\gamma^2_{g,n}\right) \geq R_{qos} \quad \forall g, \tag{28d}$$

C. OPTIMALITY, COMPLEXITY & SUBOPTIMAL ALGORITHM

The resulting optimization problems in subsections VI-A and VI-B become an integer linear programming (ILP) problem since the rank-2 matrix $\beta_{g,n}$ is linear in both cost function and constraints. However, the binary integer requirement makes the optimization an NP-Hard[2] problem therefore, it is difficult and time-consuming to solve in practise. For each subchannel in our minimization problem for

[2]problems for which solutions cannot be found in polynomial time

example, there are G^{N_s} possible combinations of subchannel allocation by brute force search. This approach becomes more prohibitive with increases in N_s and G. The exponential complexity makes optimal algorithm to the integer linear programming (ILP) problem unwieldy and impractical for real-time implementation.

One way to reduce the computation complexity is to transform the integer constraint $\beta_{g,n}$ in (27c) from discreet binary integer to continuous interval $C[0,1]$ having $\beta_{g,n} \geq 0$ and then solve using the *Lagrangian method with Karush-Kuhn-Tucker* (LKKT) conditions [24], [25]:

$$\left\{\beta_{g,n} \middle| \left\{\beta_{g,n} \in \{0,1\}\right\} \Longrightarrow \left\{\beta_{g,n} | 0 \leq \beta_{g,n} \leq 1\right\}\right\}. \quad (29)$$

The relaxation enlarges the solution space and transforms the NP-Hard IP problem to a tractable convex LP problem whose feasible solution provide a *lower bound* (in case of minimization problem) or *upper bound* (for maximization problem) on optimal value of the original optimization problem. To obtain an exact *global optimal* solution, the Lagrangian method can be combined with exhaustive search using adaptive sampling and partitioning techniques on the problem boundaries. This technique known as *branch and bound algorithm* ensures convergence to an optimal solution (OPT) but it is quite inefficient, slow and requires huge system resources despite the relaxation [25]. Although we present result for the global optimal solution (OPT) in this work, we omit the analytical derivation and focus more on an achievable suboptimal solution. In what follows, we propose a fast LKKT-based method using memoization and stochastic rounding (MSR) to speed-up computation of the feasible solutions.

D. PROPOSED MEMOIZATION WITH STOCHASTIC ROUNDING (MSR)

In this subsection, we briefly describe the memoization technique and the use the LKKT method to obtain the lower bound and upper bound for the average outage and throughput respectively. Thereafter, we apply the stochastic rounding algorithm on the intermediate results.

1) ENHANCING COMPUTATION USING MEMOIZATION

Finding the optimal solution requires executing several repetitive task of computing, searching, and comparing. We can speed-up computation by defining the cost function as a memoized function of the form $f[x] := f(x) = f(\dot{x})f(\ddot{x})$ and then solving for feasible point using LKKT. Memoization implicitly serve as a built-in lookup table that prevents re-computation of expensive function calls thus forcing the function logic to remember and return results of previous computations when same input arguments are received. Hence, we trade-off more memory for speed. For lack of better term, we define memoized form for the average outage probability as:

$$\mathcal{G}[\cdot] := \mathcal{G}\,[\cdot] = 1 - \left(1 - \beta_{g,n} Q_{\rho}^{\star}\right)^{K_g}, \quad (30)$$

$$\mathcal{C}_s = \frac{1}{GN_s} \sum_{g=1}^{G} \sum_{n=1}^{N_s} \beta_{g^*,n^*} \mathcal{G}\,[\cdot]. \quad (31)$$

Similarly, we can define memoized form for the average throughput as:

$$\mathcal{G}_{\mathcal{X}}[\cdot] := \mathcal{G}_{\mathcal{X}}\,[\cdot] = E[R^{\circ}_{g,n}], \quad (32)$$

$$\mathcal{C}_X = \sum_{g=1}^{G} \sum_{n=1}^{N_s} E[R^{\circ}_{g,n}] \quad (33)$$

where $\mathcal{G}[\cdot]$ and $\mathcal{G}_{\mathcal{X}}[\cdot]$ are functional forms and $[\cdot]$ is the input arguments G, N_s, K_g, ρ, P required by the memoized function. The memoized functions $\mathcal{G}[\cdot]$ and $\mathcal{G}_{\mathcal{X}}[\cdot]$ are used in computing the fractional intermediate results $\hat{\beta}_{g,n}$ before the stochastic rounding technique is applied [26]. Details of the proposed MSR resource utilization scheme is provided in Algorithm 1.

2) UPPER BOUND FOR QOS-AWARE RATE MAXIMIZATION

To construct the Lagrangian function, we define three variables $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{N_s})$, $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_G)$ and $\boldsymbol{\alpha} = (\alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{G,N_s})$ for the Lagrangian multipliers. By these definitions, the Lagrangian function for maximizing the average throughput using (18), (27a)-(27d) together with (28a) is defined as follows:

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\alpha}) = &\sum_{g=1}^{G} \sum_{n=1}^{N_s} \tfrac{K_g \beta_{g,n}}{log[4]} \left(A + \sqrt{\tfrac{2}{\pi}} \sigma_{g,n} B + C\right) \\ &+ \sum_{n=1}^{N_s} \lambda_n \left(1 - \sum_{g=1}^{G} \beta_{g,n}\right) + \sum_{g=1}^{G} \sum_{n=1}^{N_s} \alpha_{g,n} \beta_{g,n} \\ &+ \sum_{g=1}^{G} \eta_g \left(\sum_{n=1}^{N_s} \beta_{g,n} L_p - R_{qos}\right) \end{aligned} \quad (34)$$

where $L_p = Log_2(1 + P\gamma^2_{g,n})$. To have the optimal solution, the K.K.T necessary conditions should be satisfied [24]:

$$\nabla_{\hat{\beta}_{g,n}} L(\hat{\beta}_{g,n}, \hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\alpha}}) = 0 \quad (35)$$

$$\sum_{n=1}^{N_s} \hat{\lambda}_n \left(1 - \sum_{g=1}^{G} \hat{\beta}_{g,n}\right) = 0 \quad (36)$$

$$\sum_{g=1}^{G} \hat{\eta}_g \left(\sum_{n=1}^{N_s} \hat{\beta}_{g,n} L_p - R_{qos}\right) = 0 \quad (37)$$

where $\hat{\beta}_{g,n}, \hat{\eta}_g, \hat{\lambda}_n, \hat{\alpha}_{g,n}$ are points satisfying the above conditions. Using (34), expression in (35) can be rewritten as:

$$\begin{aligned} \nabla_{\beta_{g,n}} L(\cdot) = &\frac{\partial}{\partial \beta_{g,n}} \sum_{g=1}^{G} \sum_{n=1}^{N_s} \tfrac{K_g \beta_{g,n}}{2Log[2]} \left(A + \sqrt{\tfrac{2}{\pi}} \sigma_{g,n} B + C\right) \\ &+ \sum_{n=1}^{N_s} \lambda_n \left(1 - \sum_{g=1}^{G} \beta_{g,n}\right) + \sum_{g=1}^{G} \sum_{n=1}^{N_s} \alpha_{g,n} \beta_{g,n} \\ &+ \sum_{g=1}^{G} \eta_g \left(R_{qos} - \sum_{n=1}^{N_s} \beta_{g,n} L_p\right) = 0, \end{aligned} \quad (38)$$

Solving (38) and recognizing the KKT necessary conditions, we can obtain results for $\hat{\beta}_{g,n}$ under two different conditions. When $\hat{\beta}g, n = 0$, equality (36) is simplified, yielding: $\sum_{g=1}^{G}\sum_{n=1}^{N_s}\lambda_n = 0$. With some simplifications and substitution, the maximizer $\hat{\beta}g, n = 0$ should satisfy condition K_1:

$$K_1 : 2\text{Log}[2]\eta_g L_p = K_g(\mu_{g,n}A + \sqrt{\frac{2}{\pi}}\sigma_{g,n}B + C) \quad (39)$$

For the second condition where $\hat{\beta}_{g,n} > 0$, minimizer $\hat{\beta}_{g,n}$ should satisfy K_2:

$$K_2 : 2\text{Log}[2]\eta_g L_p < K_g(\mu_{g,n}A + \sqrt{\frac{2}{\pi}}\sigma_{g,n}B + C) \quad (40)$$

3) LOWER BOUND FOR QOS-AWARE OUTAGE MINIMIZATION

Using the Lagrangian function (35)-(37), minimization problem in (27a)-(27d), with constraint transformation in (29) gives:

$$\begin{aligned}\nabla_{\beta_{g,n}}L(\cdot) = \frac{\partial}{\partial\beta_{g,n}}\frac{1}{GN_s}\sum_{g=1}^{G}\sum_{n=1}^{N_s}1-\left(1-\beta_{g,n}Q_\rho^\star\right)^{K_g} \\ +\sum_{n=1}^{N_s}\lambda_n\left(1-\sum_{g=1}^{G}\beta_{g,n}\right)+\sum_{g=1}^{G}\sum_{n=1}^{N_s}\alpha_{g,n}\beta_{g,n} \\ +\sum_{g=1}^{G}\eta_g\left(R_{qos}-\sum_{g=1}^{G}\beta_{g,n}L_p\right)=0, \quad (41)\end{aligned}$$

where as before, $\eta_g, \lambda_n, \alpha_{g,n}$ are the Lagrangian multipliers for the inequality and equality constraints. By the KKT necessary conditions and some equation manipulation, the corresponding minimizer $\hat{\beta}_{g,n}$ should satisfy K_3 for $\hat{\beta}_{g,n} = 0$:

$$K_3 : \frac{K_g Q_\rho^\star}{GN_s}\left(1-Q_\rho^\star\hat{\beta}_{g,n}\right)^{K_g-1} = \eta_g L_p. \quad (42)$$

Also for $\hat{\beta}_{g,n} > 0$, minimizer $\hat{\beta}_{g,n}$ should satisfy K_4:

$$K_4 : \frac{K_g Q_\rho^\star}{GN_s}\left(1-Q_\rho^\star\hat{\beta}_{g,n}\right)^{K_g-1} < \eta_g L_p. \quad (43)$$

The resulting feasible points $\hat{\beta}_{g,n}$ are fractional optimizers. In what follows, we describe a stochastic scheme to obtain suboptimal integer points β_{g^*,n^*}.

4) THE STOCHASTIC ROUNDING PROCESS

First, the algorithm starts with zero allocation and proper initialization of required variables including the memoization function. We then compute the lower bound for system outage and upper bound for throughput using the LKKT method described in subsections VI-D.2 and VI-D.3. At this point, we should note that a group may be fractionally allocated to more than one subchannels. Lines 4-6 sorts the fractional parts for subchannels in decreasing order. This step directly implies throughput per channel decreasing sort for each group.

Algorithm 1 MSR for Outage Minimization: MSR_{OUT}

Require: $\Omega_g := \{\}, \beta_{g^*,n^*} := 0, \hat{\beta}_{g,n} := 0, \forall g, \forall n$

1: Define *outage cost function* in (27a) as memoized cost coefficient $\mathcal{G}[\cdot]$.
2: Using the LKKT method, compute $\hat{\beta}_{g,n}\,\forall g, \forall n$ on problem (27a)-(27d).
3: **repeat**
4: $\quad n \leftarrow n+1$
5: $\quad$ sort $\hat{\beta}_{g,n}, \quad g = \{1, 2, \ldots, G\}$ in decreasing order.
6: **until** $n == N_s$
7: **for** $n = 1$ to N_s **do**
8: $\quad$ **repeat**
9: $\quad\quad$ **for** $g = 1$ to G **do**
10: $\quad\quad\quad$ **if** $\hat{\beta}_{g,n} > 0$ **then**
11: $\quad\quad\quad\quad$ **if** $(Prob[\beta_{g^*,n^*} == 1] \geq \hat{\beta}_{g,n})\&\&$
12: $\quad\quad\quad\quad$ $(\beta_{g^*,n^*} \notin \Omega_g)$ **then**
13: $\quad\quad\quad\quad\quad$ set $\beta_{g^*,n^*} := 1$
14: $\quad\quad\quad\quad\quad$ update $\Omega_g := \beta_{g^*,n^*} \cup \Omega_g$
15: $\quad\quad\quad\quad\quad$ break
16: $\quad\quad\quad\quad$ **end if**
17: $\quad\quad\quad$ **end if**
18: $\quad\quad$ **end for**
19: $\quad$ **until** $\sum_{g^*=1}^{G}\beta_{g^*,n^*} == 1$
20: **end for**
21: Resulting β_{g^*,n^*} is the minimizer for the problem.
22: substituting β_{g^*,n^*} into $\mathcal{C}_s$ in (31) gives the rational optimal solution.

Lines 4 through 21 show the stochastic rounding process. In this technique, β_{g^*,n^*} is set to 1, if a random variate, using standard uniform distribution $\mathcal{U}(0, 1)$ is greater than the fractional solution $\hat{\beta}_{g,n}$ and a group has not been previously allocated to the subchannel n. Line 11 ensures that the allocation steps are only executed for $\hat{\beta}_{g,n}$ that have been previously allocated in line 3 as they are more likely to improve system performance. Following the allocation steps, line 15 updates $\beta_{g^*_n,n^*}$ to set of subchannels that have been allocated to groups.

The procedure continues until only one group is specifically allocated to a subchannel. The stochastic rounding can be viewed as a stepwise refinement process where number of groups on a subcarrier is transformed to strictly $= 1$ [26]. When the post-rounding β_{g^*,n^*} is substituted into (31), we can directly obtain the final solution which unsurprisingly is a suboptimal solution because the resulting 0-1 solution is rational optimum and not global optimum.

VII. NUMERICAL EVALUATION, RESULTS & DISCUSSIONS

We perform numerical evaluation to validate our proposals and analysis in section V and VI. Fig. 6 and 7 respectively discuss outage and rate trade-off for broadcast system, while Fig. 8, 10 discuss outage in multicast systems and Fig. 9, 11 explain the rate trade-off for multicast systems.

Algorithm 2 MSR for Rate Maximization: MSR_{TRX}

Require: $\Omega_g := \{\}, \beta_{g^*,n^*} := 0, \hat{\beta}_{g,n} := 0, \forall g, \forall n$

Define *throughput cost function* in (28a) as memoized cost coefficient $\mathcal{G}_{\mathcal{X}}[\cdot]$.
Compute $\hat{\beta}_{g,n}\ \forall g, \forall n$ on problem (28a)-(28d) using the LKKT method.
repeat
  $n \leftarrow n+1$
  sort $\hat{\beta}_{g,n}, \quad g = \{1, 2, \ldots, G\}$ in decreasing order.
until $n == N_s$
for $n = 1$ to N_s **do**
  repeat
    for $g = 1$ to G **do**
      if $\hat{\beta}_{g,n} > 0$ **then**
        if $(Prob[\beta_{g^*,n^*} == 1] \geq \hat{\beta}_{g,n})\&\&$
        $(\beta_{g^*,n^*} \notin \Omega_g)$ **then**
          set $\beta_{g^*,n^*} := 1$
          update $\Omega_g := \beta_{g^*,n^*} \cup \Omega_g$
          break
        end if
      end if
    end for
  until $\sum_{g^*=1}^{G} \beta_{g^*,n^*} == 1$
end for
the resulting β_{g^*,n^*} is the feasible set for the maximization problem.
substituting β_{g^*,n^*} into $\mathcal{C}_X$ in (33) gives the rational optimal solution.

In Fig. 6, we show the average system outage probability of both CVM and proposed AWG approximation for broadcast system using the optimal OPT and the proposed suboptimal MSR algorithms for evaluation. The results shown here are for $K_g = \{2, 3\}$, $N_s = 16$, $R_{qos} = 3$ and $P_{Tot} = 50dB$. Observe that for CVM, outage increases as number of users increases. This behaviour clearly shows that the system will become saturated at some point as K_g grows because the min-SNR limits system performance. For AWG however, notice that the outage begins to drop drastically with increase in K_g and P_{Tot} while still satisfying R_{qos} because AWG does not just use the minSNR but it amortizes the SNR of all users in the group. This results clearly confirms that *outage in a broadcast system with uniform data rate requirement depends on the number of users in the group* which interestingly is consistent with [10].

Fig. 7 shows the consequence of the minimum outage benefit in Fig. 6. For example, when $G = 1$ for both OPT and MSR, CVM obviously has higher transmission gain. However, the rate performance of AWG on both OPT and MSR is better at low P_{Tot}. This apparently means that for CVM to have any significant benefit, the BS needs to expend more transmit power. In essence, *rate gap between CVM and AWG imply that the benefit of AWG becomes obvious only when we transmit at low power in the downlink.* Although lower rate is

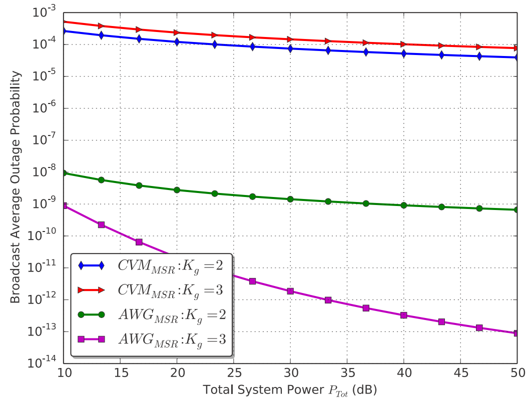

FIGURE 6. Broadcast outage probability vs. system transmit power with different number of users per group $K_g = \{2, 3\}$, $Ns = 16$, $R_{qos} = 3Mbps$.

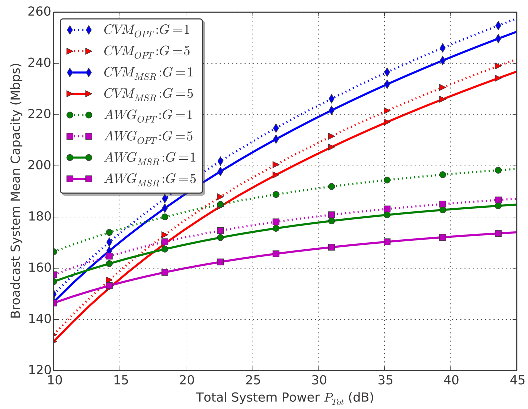

FIGURE 7. Throughput vs. system power for broadcast $G = 1$ and multicast $G = 5$ with random $K_g \leq 5$, $P_{Tot} = 45dB$, $Ns = 16$, $R_{qos} = 4$.

not desirable from network operators' perspective, but AWG will allow operators to provide energy efficient solutions with guaranteed rate requirements R_{qos} that satisfy users' multimedia needs. As a validation of our MSR allocation algorithm, notice that for both CVM and AWG, performance gap of MSR is *less than 12%* from the OPT.

Fig. 8 shows the average outage probability for multicast groups with random $K_g \leq 5$, $Rqos = \{2, 4\}Mbps$, $Ns = 16$, $G = 3$, $P_{Tot} = 50dB$ using suboptimal MSR algorithm. As expected, for both CVM and AWG, a higher system outage may occur when QoS rate requirement increases from $2Mbps$ to $4Mbps$ for a particular fixed transmit power P_{Tot}. More interesting point in this Fig. 8 is the low outage of AWG compared to CVM. This result further strengthens our observation in Fig. 6 and strongly establishes our position on the significance of AWG as an amortization scheme that effectively mitigates the detrimental impact of low and high SNRs in the multicast group. However, to put this results in accurate perspective, we need to examine its trade-off which is shown in Fig. 9.

Two key points immediately pop out from Fig. 9: First, at the low transmit power region ($P_{Tot} \leq 25dB$), the AWG scheme outperforms the CVM scheme just like in the

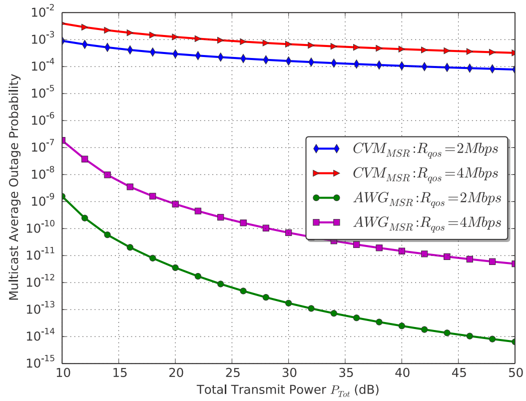

FIGURE 8. **Multicast outage probability vs. system power for different rate requirements $Rqos = \{2, 4\}Mbps$, $Ns = 16$, $G = 3$, $P_{Tot} = 50dB$.**

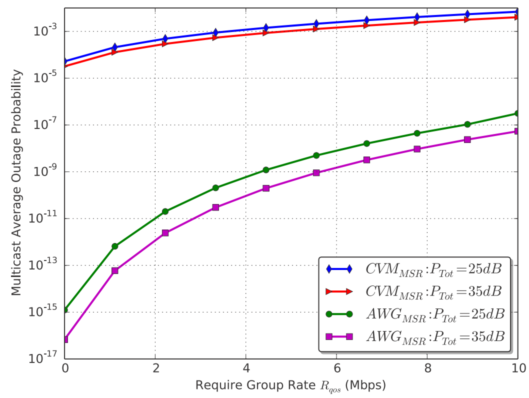

FIGURE 10. **Multicast outage probability vs minimum group rate requirement for different transmission power $P_{Tot} = \{25dB, 35dB\}$, $Ns = 16$, $G = 3$.**

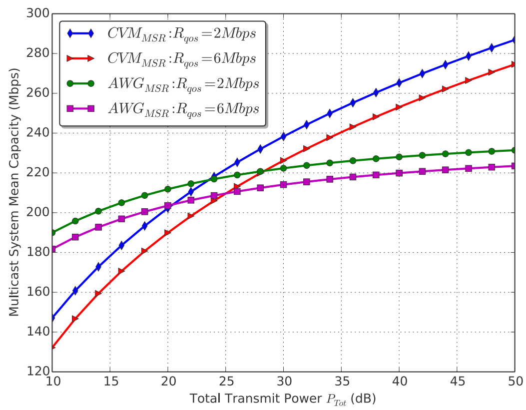

FIGURE 9. **Multicast average throughput vs. system transmit power for different group target rates $Rqos = \{2, 6\}$, $Ns = 16$, $G = 3$, $P_{Tot} = 50dB$.**

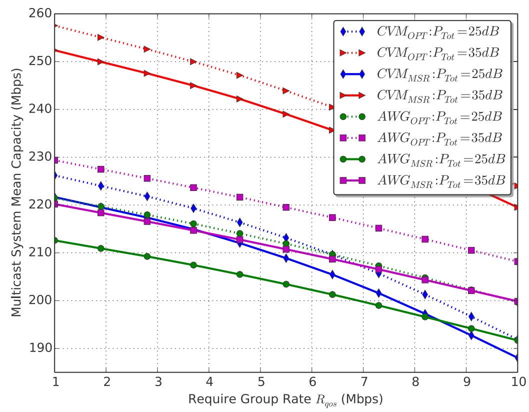

FIGURE 11. **Multicast average throughput vs minimum group rate requirement for different transmission power $P_{Tot} = \{25dB, 35dB\}$, $Ns = 16$, $G = 3$.**

broadcast system discussed in Fig. 7. This means there is a rate trade-off between CVM and AWG subject to the amount of transmit power available to the network. Secondly, when the group rate requirement R_{qos} increases from $2Mbps$ to $6Mbps$, there is a corresponding drop in throughput. This is because groups potentially need more subchannel resources to satisfy the $6Mbps$ requirements, therefore depleting the number of available resources that could be allocated. This means one or more groups that do not satisfy the $6Mbps$ requirements will not receive allocation and may have to wait another time frame. This result is consistent with Fig. 8 as low throughput could result in lower outage as in this case.

In Fig. 10, we show the average outage probability for multicast group with random $K_g \leq 5$, $P_{Tot} = \{25dB, 35dB\}$, $Ns = 16$, $G = 3$, and R_{qos} up to $10Mbps$ using the proposed suboptimal MSR algorithm. It is clear that rate of outage monotonously increases with change in required rate R_{qos} for different fixed transmit power. Observe that outage for CVM is significantly higher than for AWG even when transmit power increases. This is because when a group is limited by users at the cell edge (minSNR), the potential per subchannel reception rate becomes lower, therefore, for the group to satisfy the increasing R_{qos}, higher number of subchannel (bandwidth) should be allocated to the group. Such resource distribution undoubtedly depletes resources that could have been assigned to other groups to facilitate their transmissions. For AWG however, group per subchannel reception rate is potentially higher and lesser resources are required, so outage becomes lower. However, this implies some users in the group would not be able to successfully decode received data. In this case, a multicast retransmission or error correction scheme could be applied [27]. Such works are outside the scope of this paper.

Fig. 11 directly shows three important observations. First, when rate requirements increases without corresponding increase in subchannel and power resources, the system rate degrades. This behaviour is same for both CVM and AWG. Secondly, it highlights the consequences of the missed transmission of AWG earlier explained in Fig. 10. Observe that at the group level CVM (by definition) satisfy all users in group by requesting least throughput while AWG does not because some users miss out of successful reception. However, at the system (multigroup) level, AWG utilize less resources, make more resources available and satisfies more groups so outage is lower and by implication (of the missed transmission), average throughput is also lower than CVM. Lastly, we make a point about the efficiency of the suboptimal MSR algorithm that it also effectively performs up to 90% of the OPT algorithm which is well known to have high computation complexity.

VIII. CONCLUSION

In this paper, we have considered the outage probability for the emergent multimedia broadcast and multicast systems (MBMS). We analysed a unique amortized weighted averaging (AWG) and proposed its approximation as an alternative to conventional minSNR based assumption in multicasting. We presented the statistical distribution as well as its density function. We also analysed the system average outage probability and system average throughput for both broadcast and multicast systems. Through our analytical results, we show that contrary to popular opinion in conventional multicasting (CVM), *it is possible to systematically exploit the users' channel perception and prevent system saturation as number of users in a multicast group increases.* We also showed that AWG avoids outage, accommodates more users and is *generally more energy efficient than the CVM especially when low power is available to the system*. In practical systems, AWG can be implemented as a cross-layer radio resource management (RRM) submodule together with an error-correction code in the physical layer where the throughput can be determined before the system resources are allocated. Possible future research directions may include the integration of the emerging technologies such as non-orthogonal multiple access and index modulation into the multicast systems. In addition, considering physical layer security issues of the broadcast and multicast systems is also interesting [28], [29].

REFERENCES

[1] G. Araniti, M. Condoluci, P. Scopelliti, A. Molinaro, and A. Iera, "Multicasting over emerging 5G networks: Challenges and perspectives," *IEEE Netw.*, vol. 31, no. 2, pp. 80–89, Mar./Apr. 2017.

[2] J. Montalban *et al.*, "Multimedia multicast services in 5G networks: Subgrouping and non-orthogonal multiple access techniques," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 95–96, Mar. 2018.

[3] D. Lecompte and F. Gabin, "Evolved multimedia broadcast/multicast service (eMBMS) in LTE-advanced: Overview and Rel-11 enhancements," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 68–74, Nov. 2012.

[4] L. Zhang, Y. Wu, G. K. Walker, W. Li, K. Salehian, and A. Florea, "Improving LTE *e* MBMS with extended OFDM parameters and layered-division-multiplexing," *IEEE Trans. Broadcast.*, vol. 63, no. 1, pp. 32–47, Mar. 2016.

[5] C. P. Lau, A. Alabbasi, and B. Shihada, "An efficient live TV scheduling system for 4G LTE broadcast," *IEEE Syst. J.*, vol. 11, no. 4, pp. 2737–2748, Dec. 2017.

[6] K. Bakanoglu, W. Mingquan, L. Hang, and M. Saurabh, "Adaptive resource allocation in multicast OFDMA systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.

[7] A. M. C. Correia, J. C. M. Silva, N. M. B. Souto, L. A. C. Silva, A. B. Boal, and A. B. Soares, "Multi-resolution broadcast/multicast systems for MBMS," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 224–234, Mar. 2007.

[8] C. Suh and J. Mo, "Resource allocation for multicast services in multicarrier wireless communications," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 27–31, Jan. 2008.

[9] W. Xu, K. Niu, J. Lin, and Z. He, "Resource allocation in multicast OFDM systems: Lower/upper bounds and suboptimal algorithm," *IEEE Commun. Lett.*, vol. 15, no. 7, pp. 722–724, Jul. 2011.

[10] R. O. Afolabi, B. V. Nguyen, and K. Kim, "Amortized weighted averaging for multimedia broadcast and multicast systems," *IEEE Commun. Lett.*, vol. 18, no. 5, pp. 837–840, May 2014.

[11] J. Liu, W. Chen, Z. Cao, and K. B. Letaief, "Dynamic power and subcarrier allocation for OFDMA-based wireless multicast systems," in *Proc. IEEE Int. Conf. Commun.,* May 2008, pp. 2607–2611.

[12] Y. Sun and K. J. Ray Liu, "Transmit diversity techniques for multicasting over wireless networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2004, pp. 593–598.

[13] F. Hou, L. X. Cai, P. H. Ho, X. Shen, and J. Zhang, "A cooperative multicast scheduling scheme for multimedia services in IEEE 802.16 networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1508–1519, Mar. 2009.

[14] N. Shrestha, P. Saengudomlert, and Y. Ji, "Dynamic subcarrier allocation with transmit diversity for OFDMA-based wireless multicast transmissions," in *Proc. Int. Conf. Elect. Eng./Electron., Comput., Telecommun. Inf. Technol.*, May 2010, pp. 410–414.

[15] J. Liu, W. Chen, Y. J. Zhang, and Z. Cao, "A utility maximization framework for fair and efficient multicasting in multicarrier wireless cellular networks," *IEEE/ACM Trans. Netw.*, vol. 21, no. 1, pp. 110–120, Feb. 2013.

[16] P. K. Gopala and H. El Gamal, "On the throughput-delay tradeoff in cellular multicast," in *Proc. Int. Conf. Wireless Netw., Commun. Mobile Comput.*, Jun. 2005, pp. 1401–1406.

[17] C. H. Koh and Y. Y. Kim, "A proportional fair scheduling for multicast services in wireless cellular networks," in *Proc. Veh. Technol. Conf.*, Sep. 2006, pp. 1–5.

[18] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[19] T. Girici and G. D. Kurt, "Minimum-outage broadcast in wireless networks with fading channels," *IEEE Commun. Lett.*, vol. 14, no. 7, pp. 617–619, Jul. 2010.

[20] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series and Products*, 7th ed. New York, NY, USA: Academic, 2007.

[21] J. Abate and P. P. Valkó, "Multi-precision Laplace transform inversion," *Int. J. Numer. Methods Eng.*, vol. 60, no. 5, pp. 979–993, 2004.

[22] H. Mouri, "Log-normal distribution from a process that is not multiplicative but is additive," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 88, no. 4, p. 042124, 2013.

[23] M. Mitzenmacher, "A brief history of generative models for power law and lognormal distributions," *Internet Math.*, vol. 1, no. 2, pp. 226–251, 2004.

[24] M. L. Fisher, "Multi-precision Laplace transform inversion," *Manage. Sci.*, vol. 50, no. 12, pp. 1861–1871, Dec. 2004.

[25] S. Boyd and L. Vanderberghe, *Convex Optimization*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[26] P. Raghavan and C. D. Tompson, "Randomized rounding: A technique for provably good algorithms and algorithmic proofs," *Combinatorica*, vol. 7, no. 4, pp. 365–374, 1987.

[27] T. Mladenov, S. Nooshabadi, and K. Kim, "Efficient incremental raptor decoding over BEC for 3GPP MBMS and DVB IP-datacast services," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 313–318, Jun. 2011.

[28] B. V. Nguyen and K. Kim, "Secrecy outage probability of optimal relay selection for secure AnF cooperative networks," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2086–2089, Dec. 2015.

[29] B. Van Nguyen, H. Jung, and K. Kim, "Physical layer security schemes for full-duplex cooperative systems: State of the art and beyond," *IEEE Commun. Mag.,*, to be published, doi: 10.1109/MCOM.2017.1700588.

DUC-PHUC VUONG received the B.S. and M.S. degrees in marine electrical engineering and automation from Vietnam Maritime University in 2004 and 2007, respectively, and the Ph.D. degree in electric and control engineering from Mokpo National Maritime University, South Korea, in 2014. He is currently a Lecturer with the Department of Electric and Electronics Engineering, Vietnam Maritime University. His research interests include broadcast-multicast systems, industrial IoT applications, control of industrial machines, robot control, LabVIEW software, and recycle energy.

MINH-QUAN DAO received the B.S. and M.S. degrees in marine electrical engineering and automation from Vietnam Maritime University in 1999 and 2004, respectively, and the Ph.D. degree in engines and power plants from the Odessa National Maritime Academy in 2011. He is currently a Lecturer with the Department of Electric and Electronics Engineering, Vietnam Maritime University. His research interests include broadcast-multicast system, industrial IoT applications, control of industrial machines, robot control, recycle energy, ship power plants, ship control and automation, and maritime traffic.

DANG-KHANH LE received the B.S. degree from the Marine Engineering Department, Vietnam Maritime University, Haiphong, Vietnam, in 2006, and the M.S. and Ph.D. degrees from the Department of Marine Engineering, Mokpo National Maritime University, South Korea in 2012 and 2016, respectively. He held a post-doctoral position with iSL, Mokpo National University, in 2016. He is currently a Lecturer with the VMU College, Vietnam Maritime University. His research interests include instrumentation and automation control, UAV modeling and simulation, ultrasonic technologies and applications, iterative learning control, and intelligent transportation systems.