Performance Evaluation in Single or Multi-Cluster C-RAN Supporting Quasi-Random Traffic

I-A. Chousainov, I. Moscholios, A. Kaloxylos, Senior Member, IEEE, and M. Logothetis, Senior Member, IEEE

Abstract—In this paper, a cloud radio access network (C-RAN) is considered where the remote radio heads (RRHs) are separated from the baseband units (BBUs). The RRHs in the C-RAN are grouped in different clusters according to their capacity while the BBUs form a centralized pool of computational resource units. Each RRH services a finite number of mobile users, i.e., the call arrival process is the quasi-random process. A new call of a single service-class requires a radio and a computational resource unit in order to be accepted in the C-RAN for a generally distributed service time. If these resource units are unavailable, then the call is blocked and lost. To analyze the multi-cluster C-RAN, we model it as a single-rate loss system, show that a product form solution exists for the steady state probabilities and propose a convolution algorithm for the accurate determination of congestion probabilities. The accuracy of this algorithm is verified via simulation. The proposed model generalizes our recent model where the RRHs in the C-RAN are grouped in a single cluster and each RRH accommodates quasi-random traffic.

Index Terms—cloud, radio access, call blocking, product form, quasi-random.

I. INTRODUCTION

The cloud radio access network (C-RAN) is considered as a promising fifth generation (5G) network architecture which is anticipated to address the increasing demand for bandwidth-hungry applications, decreased latency (in the order of a few ms), improved data rate (up to 1 Gbps) and consequently the rapid growth of wireless traffic which is expected to reach 110 ExaBytes/month by the end of 2023 [1], [2].

The C-RAN architecture includes a large number of base stations where the remote radio heads (RRHs) are separated from the baseband units (BBUs) [3]. The BBUs form a centralized pool of computational resource units (RUs) which is connected to the RRHs via the common public radio interface (CPRI) with a high-capacity fronthaul [4].

To benefit from network function virtualization, we consider virtualized BBU resources (V-BBU) [3] where the BBU functionality and services have been virtualized in the form of virtual network functions (VNFs) [5]. We focus on the radio resource management, a BBU function that can be virtualized as a VNF, which is responsible not only for the radio resource allocation but also for the call admission control (CAC) of mobile users (MUs).

A CAC is a significant quality of service (QoS) mechanism that gives access to the RUs required by MUs and provides fairness among different mobile services/applications. Considering call-level traffic in the C-RAN, such a QoS mechanism is a bandwidth (resource) sharing policy since it affects call blocking probabilities (CBP). Generally speaking, it is desirable to have efficient recursive formulas or convolution algorithms for the determination of CBP. Such formulas or algorithms can be helpful in network planning and can be obtained via teletraffic loss/queueing models [6], [7]. In this paper, we focus on convolution algorithms which are adopted in the literature not only in the case of loss/queueing models whose steady state probabilities have a product form solution (PFS) but also in the case of non-PFS models (e.g., [8]–[16]). Such algorithms are advantageous not only due to their low computational complexity but also because they can incorporate various resource sharing policies such as the complete sharing policy, the bandwidth reservation policy and threshold-based policies (e.g., [17]–[24]).

On the one hand, various aspects of the C-RAN architecture have been investigated and analyzed the last few years, such as the capacity demands and possible functional splits on the fronthaul network [25], [26], energy and cost saving issues [27], [28], security challenges [29], resource allocation issues related to RRH selection, spectrum management and throughput maximization [30], as well as the dimensioning problem of the necessary number of V-BBU required to handle a specific number of RRHs [31], [32]. The latter focus only on the V-BBU and model them as a queuing system in which the arrival process of jobs follows a batched Poisson process and the service time is exponentially distributed. On the other hand, there are very few papers that consider CAC in the C-RAN and provide analytical formulas or algorithms for the CBP determination ([33]–[35], [37]–[39]).

In [33], all RRHs form a single cluster in the C-RAN. The latter accommodates calls of a single service-class that arrive in the system according to a Poisson process. A new call

Original scientific paper
requires a computational RU from the V-BBU and a radio
RU from the serving RRH. If both RUs are available, then
the new call is accepted in the serving RRH and remains
in the system for a generally distributed service time. Oth-
erwise, call blocking occurs. The analytical model of [33],
named single-class-single-cluster (SC-SC) model, is based on
a multidimensional reversible Markov chain which leads to a
PFS for the steady-state probabilities. The PFS leads to an
accurate CBP calculation either via the complex procedure of
enumeration and processing of the system’s state space or via
recursive formulas. In [34], [35], the SC-SC model is extended
to include the case of overlapping calls via the notion of a
direct routing network [36]. In [37], a convolution algorithm
is proposed for the CBP determination of the SC-SC model. In
[38], the SC-SC model is extended to include the case where
the RRHs can be grouped in multi-clusters (SC-MC model).
In that case, the RRHs that belong to a cluster have the same
amount of radio RUs. Recently, the authors have extended the
SC-SC model to include the case of a quasi-random call arrival
process [39]. This process appears when the RRHs serve calls
generated by a finite number of MUs. In that sense, the quasirandom process is smoother than the Poisson process (where
calls are generated by an infinite number of users) [7]. We
name the model of [39], finite SC-SC (f-SC-SC) model.

In this paper, we focus on the models of [38] and [39].
More precisely, we extend the SC-MC model of [38] by incor-
porating the quasi-random call arrival process. The proposed
model, named finite SC-MC (f-SC-MC) model, generalizes the
f-SC-SC model of [39], since it covers the more complicated
multi-cluster case. Our contribution is three-fold: 1) we review
the SC-MC model, provide some corrections for the CBP
formulas (presented in [38]) which are necessary for their
implementation and propose a convolution algorithm for the
CBP calculation, 2) we propose the f-SC-MC model, analyze it
via a multidimensional reversible Markov chain and show that
a PFS exists for the steady-state probability distribution and
3) we propose a convolution algorithm (which is based on the
proposed PFS) for the exact determination of the congestion
probabilities in the f-SC-MC model and compare the results
with those obtained via the SC-SC model of [33], the SC-MC
model of [38] and the f-SC-SC model of [39]. Simulation
results verify the accuracy of all models.

The remainder of this paper is as follows: In Section II, we
review the SC-MC model and propose a convolution algorithm
for the calculation of CBP. In Section III, we present the
proposed f-SC-MC model and a 3-step convolution algorithm
for the determination of the various performance measures.
In Section IV, we present analytical and simulation results
for the congestion probabilities of the f-SC-SC and the f-
SC-MC models and compare them with the corresponding
analytical results of [33] and [38], respectively. We conclude
in Section V.

II. REVIEW OF THE SC-MC MODEL

A. The Analytical Model

Consider the C-RAN of Fig. 1 where the RRHs are sepa-
rated from the V-BBU. There exist $Z$ different classes of
RRHs. Each class $z$ ($z = 1, \ldots, Z$) forms a cluster of $M_z$
RRHs. A RRH that belongs to class $z$ has $C_z$ radio RUs which
are allocated to the serving calls of the MUs. The V-BBU has
a capacity of $T$ computational RUs.

Similar to the SC-SC model of [33], new calls arrive in
the C-RAN according to a Poisson process. Let $\lambda_z$ be the
arrival rate of new calls in a serving RRH of class $z$. The
call admission mechanism is as follows: a new call requires
a computational RU from the V-BBU and a radio RU from
the serving RRH of class $z$. If both RUs are available, then
the new call is accepted in the serving RRH for a generally
distributed service time with mean $\mu^{-1}$. Otherwise, the call
is blocked and lost.

![Fig. 1. The SC-MC model.](image-url)

Let $n_{z,m}$ be the number of in-service calls in the $m$-th RRH
that belongs to class $z$ ($z = 1, \ldots, Z$ and $m = 1, \ldots, M_z$).
Then, the number of calls serviced in all RRHs of all Z
classes is expressed by the steady-state vector $n = (n_{1,1}, \ldots,
n_{1, M_1}, \ldots, n_{Z,1}, \ldots, n_{Z, M_z})$ and the cor-
responding steady-state probability distribution, $P(n)$ has the
following PFS [38]:

$$P(n) = G^{-1} \left( \prod_{z=1}^{Z} \prod_{m=1}^{M_z} \frac{\lambda_z n_{z,m}}{n_{z,m}!} \right)$$

where: $\alpha_z = \lambda_z / \mu$ is the offered traffic-load (in erl)
in every RRH that belongs to class $z$, $G(\Omega) = \sum_{n \in \Omega} \prod_{z=1}^{Z} \prod_{m=1}^{M_z} \frac{\alpha_z n_{z,m}}{n_{z,m}!}$ and $\Omega$ is the system’s state
space given by $\Omega = \{n : 0 \leq n_{z,1}, \ldots, n_{z,M_z} \leq C_z, 0 \leq \sum_{z=1}^{Z} \sum_{m=1}^{M_z} n_{z,m} \leq T\}$.

Based on $P(n)$, we can determine the total CBP, $B_{z,int}$, of
calls serviced in a RRH of class $z$ via:

$$B_{z,int} = B_{z,r} + B_{z,c}$$

where: $B_{z,r}, B_{z,c}$ refer to the blocking caused due to insufficient
radio and computational RUs, respectively.

The values of $B_{z,int}$ can be computed either via a brute-force
evaluation method (see Section II-B) or via a recursive method
(see Section II-C) proposed in [38] or via a convolution
algorithm (see Section II-D) proposed herein.
B. The Brute Force Evaluation Method

The values of $B_{z,r}$ can be computed via (1) as follows:

$$B_{z,r} = \frac{1}{G} \left[ \frac{a_{C_z}}{C_z!} \sum_{n \in \Omega_{z,C_z}} \left( \prod_{i=1}^{Z} \prod_{m=1}^{M} \frac{a_{w,m}}{n_{w,m}!} \right) \left( \prod_{m=2}^{M} \frac{a_{z,m}}{n_{z,m}!} \right) \right]$$

where: $\Omega_{z,T} = \{ n : n_{z,1} = C_z \}$, $\Omega_{z,T} = \{ n : \sum_{z=1}^{Z} n_{z,m} < T \}$.

Equation (3) gives the values of $B_{z,r}$ for the 1st RRH of class $z$. However, since we have assumed that the RRHs of class $z$ are identical and have the same offered traffic-load $\alpha_z$ and the same capacities $C_z$, then (3) refers to the blocking caused due to insufficient radio RUs in any RRH of class $z$.

Similarly, by denoting as $\Omega_{z,T} = \{ n : \sum_{z=1}^{Z} \sum_{m=1}^{M} n_{z,m} = T \}$, the values of $B_c$ are given by:

$$B_c = \sum_{n \in \Omega_{z,T}} P(n)$$

C. The Recursive Evaluation Method

The determination of $B_{z,r}$ and $B_c$ according to the recursive method of [38], requires the following functions:

$$C_R(T, M) = \sum_{n \in \Omega_{z,T} \times 1}^{Z} \prod_{m=1}^{M} \frac{a_{z,m}}{n_{z,m}!}$$

$$R_R(T, M) = \sum_{n \in \Omega_{z,T} \times 1}^{Z} \prod_{m=2}^{M} \frac{a_{z,m}}{n_{z,m}!}$$

where: $M = (M_1, \ldots, M_z, \ldots, M_Z)^T$ is a column vector of length $Z$.

The recursive calculation of $C_R(T, M)$ and $R_R(T, M)$ is based on (7) and (8), including some necessary corrections (for their software implementation) compared to the formulas that appear in [38]:

$$C_R(T, M) = \begin{cases} \frac{a_{X_{z,1}}^{X_{z,1}}}{X_{z,1}!}, & \text{for } M = e_z \\ \sum_{i=1}^{X_{z,1}} C_R(T-i, M-e_z), & \text{for } M > e_z \\ \end{cases}$$

$$R_R(T, M) = \begin{cases} \frac{a_{X_{z,1}}^{X_{z,1}}}{X_{z,1}!}, & \text{for } T = 1 \\ \frac{R_R(T+1,M) - C_R(T,M)}{M_i}, & \text{for } T \leq \sum_{z=1}^{Z} M_z C_z \\ \prod_{s=1}^{T} \left( \sum_{i=0}^{s} \frac{a_{z,s}^{s}}{s!} \right)^{M_i}, & \text{for } T = \sum_{z=1}^{Z} M_z C_z + 1 \\ \end{cases}$$

Based on (7), (8), the values of $B_{z,r}$ and $B_c$ can be computed via (9) and (10) respectively:

$$B_{z,r} = \frac{C_z C_R(T - C_z, M - e_z)}{C_z! \cdot R_R(T + 1, M)}$$

$$B_c = \frac{C_R(T, M)}{R_R(T + 1, M)}$$

Regarding the computational complexity of the previous recursive method, it is proved that it is upper bounded by the term $(\max C_z)^2 + \max C_z) M^2$. The latter is based on the assumption that $T \leq \sum_{z=1}^{Z} M_z C_z$. In the case of $T > \sum_{z=1}^{Z} M_z C_z$, which results in an over-provision of the computational RUs, a large pool approximation is proposed for the determination of blocking probabilities.

The previous method leads to an efficient way for the computation of $B_{z,r}$ and $B_c$ and consequently $B_{z,\text{tot}}$ (via (2)). However, it is not clear how: i) the calculation of the occupancy distribution of the computational RUs will be achieved and ii) this method can be extended when non-Poisson arrival processes should be studied. Finally, note that in the particular case of $T > \sum_{z=1}^{Z} M_z C_z$, the model behaves as $M$ individual Erlang B loss systems of capacity $C_z$ ($z = 1, \ldots, Z$) radio RUs and offered traffic-load equal to $\alpha_z$. In that case, $B_0 = 0$ and $B_{z,r} = E_{C_z}(\alpha_z)$ where $E_{C_z}(\alpha_z)$ is the Erlang B formula. The recursive evaluation method of (8) and (9) is not defined when $T > \sum_{z=1}^{Z} M_z C_z$. The previous drawbacks can be dealt with the proposed convolution algorithm that follows.

D. The Proposed Convolution Algorithm

In order to efficiently determine $B_{z,\text{tot}}$ as well as the occupancy distribution of the computational RUs, we propose the following convolution algorithm:

Step 1

For each class $z$ of RRHs ($z = 1, \ldots, Z$) determine the occupancy distribution of each of the $M_z$ RRHs, $q_{z,m}(j)$, where $m = 1, \ldots, M_z$ and $j = 1, \ldots, C_z$, via:

$$q_{z,m}(j) = \left\{ \begin{array}{ll} \frac{a_{z,m}^{j}}{j! q_{z,m}^{(0)}}, & \text{for } j = 1, \ldots, C_z \\ 0, & \text{for } j = C_z + 1, \ldots, T \end{array} \right.$$

The values of $q_{z,m}(j)$ can be normalized via the constant $G_{z,m} = \sum_{j=1}^{C_z} q_{z,m}(j)$ and are denoted via $q'_{z,m}(j) = q_{z,m}(j)/G_{z,m}$.

Step 2

Following the successive convolution of all RRHs (of all classes $Z$), apart from the $(z,1)$ RRH, determine the aggregated occupancy distribution:

$$Q_{z}(z,1) = q'_{z,1} \ast \cdots \ast q'_{z,1} \ast \cdots \ast q'_{z,1} \ast q'_{z,1} \ast \cdots \ast q'_{z,1}$$

The convolution operation between two occupancy distributions $q'_{z,\text{u}}$ and $q'_{z,\text{w}}$ is given by:

$$q'_{z,\text{u}} \ast q'_{z,\text{w}} = \left\{ \begin{array}{ll} q'_{z,\text{u}}^{(0)} \cdot q'_{z,\text{w}}^{(0)}, & \text{for } x = 0 \\ \sum_{x=0}^{T} q'_{z,\text{u}}^{(x)} \cdot q'_{z,\text{w}}^{(T-x)}, & \text{for } x = 1, \ldots, T \end{array} \right.$$
Step 3
Compute \( B_{z,n,t} \) based on the normalized values obtained from the convolution operation of step 2, as follows:

\[
B_{z,n,t} = B_{z,r} + B_c = \frac{1}{G} \left( q_{z,1}^{(C)} - z - C + 1 \right) \sum_{r=0}^{T-C-1} Q_{\lfloor z-r \rfloor}(r) + q(T) \tag{14}
\]

where: \( q_{z,1}^{(C)} \) refers to the case of unavailable radio RUs in the \((z,1)\) RRH (already determined in step 1), \( Q_{\lfloor z-r \rfloor}(r) \) are the normalized values of \( Q_{\lfloor z-r \rfloor}(r) \), while \( q(T) \) refers to the un-normalized probability of unavailable computational RUs, given by \( q(T) = \sum_{x=0}^{T} Q_{\lfloor x-1 \rfloor}(x) \cdot q_{z,1}^{(C)}(T-r) \) and \( G \) is the normalization constant of the final convolution operation \( Q_{\lfloor z-r \rfloor} \cdot q_{z,1}^{(C)} \) determined via (13).

The notation of (14) leads to the values of \( B_{z,n,t} \) for the 1st RRH of class \( z \). Since we have already assumed that all RRHs of the same class \( z \) have the same capacity \( C_z \) and the same offered traffic-load \( \alpha_z \), it is obvious that (14) refers to the \( B_{z,n,t} \) of any RRH of the same class \( z \).

Based on the convolution algorithm, we can obtain the occupancy distribution of the computational RUs, as follows:

\[
q^{(0)}(j) = \frac{Q_{\lfloor z-r \rfloor}(0) \cdot q_{z,1}^{(C)}}{G}, \quad j = 0 \tag{15}
\]

\[
q^{(j)}(j) = \sum_{r=0}^{T-C-1} Q_{\lfloor z-r \rfloor}(r) \cdot q_{z,1}^{(C)}(j-r) / G, \quad j = 1, \ldots, T \tag{15}
\]

where: \( G \) is the normalization constant of the final convolution operation \( Q_{\lfloor z-r \rfloor} \cdot q_{z,1}^{(C)} \).

Regarding the computational complexity of (13) for the SC-MC model, it is in the order of \( O(T^2) \) while the corresponding evaluation order in order to obtain the values of (15) is in the order of \( O(MT^2) \).

Note that when \( Z = 1 \), then all RRHs form a single cluster in the C-RAN and we have the SC-SC model. In that case, the interested reader may resort to [33] and [39] for the analysis of the brute force evaluation and the recursive evaluation methods of this model. Herein, we only present the corresponding 3-step convolution algorithm for the CBP calculation [39]:

Step 1
For each of the \( M \) RRHs of the C-RAN, determine the occupancy distribution \( q_m(j) \), where \( m = 1, \ldots, M \) and \( j = 1, \ldots, C \), via:

\[
q_m(j) = \begin{cases} \frac{q^{(0)}(j)}{G_m} & \text{for } j = 1, \ldots, C \\ 0 & \text{for } j = C + 1, \ldots, T \end{cases} \tag{16}
\]

The normalized values of \( q_m(j) \) are expressed as \( q_m^{(j)} = q_m(j) / G_m \) where \( G_m = \sum_{j=0}^{C} q_m(j) \).

Step 2
Following the convolution of all RRHs, apart from the 1st RRH, determine the aggregated occupancy distribution:

\[
Q_{\lfloor -1 \rfloor} = q_1^{(C)} \ast \ldots \ast q_m^{(C)} \ast \ldots \ast q_M^{(C)} \tag{17}
\]

where the convolution of \( q_1^{(C)} \) and \( q_1^{(C)} \) is given by (13).

Step 3
Calculate the values of the total CBP, \( B_{tot} \), according to the normalized values obtained in step 2, as follows:

\[
B_{tot} = B_{tot} + B_c = \frac{1}{G} \left( q_{1}^{(C)}(C) - z - C + 1 \right) \sum_{r=0}^{T-C-1} Q_{\lfloor -1 \rfloor}(r) + q(T) \tag{18}
\]

where: \( q_1^{(C)}(C) \) expresses the case where there are no radio RUs available in the 1st RRH while \( q(T) \) expresses the (un-normalized) probability that all computational RUs are unavailable, given by \( q(T) = \sum_{x=0}^{T} Q_{\lfloor x-1 \rfloor}(x) \cdot q_{z,1}(T-x) \) and \( G \) is the normalization constant of the final convolution operation \( Q_{\lfloor -1 \rfloor} \cdot q_{z,1}^{(C)} \) given via (13).

To determine the occupancy distribution of the computational RUs, we can use the formulas:

\[
q^{(0)}(j) = \frac{Q_{\lfloor -1 \rfloor}(0) \cdot q_{1}^{(C)}}{G}, \quad j = 0 \tag{19}
\]

\[
q^{(j)}(j) = \sum_{z=0}^{T-C-1} Q_{\lfloor -1 \rfloor}(z) \cdot q_{1}^{(C)}(j-z) / G, \quad j = 1, \ldots, T \tag{19}
\]

III. THE PROPOSED FINITE SC-MC MODEL

A. The Analytical Model

In the proposed f-SC-MC model, calls arrive in the \( m \)-th RRH that belongs to class \( z = 1, \ldots, Z \) and \( m = 1, \ldots, M_z \), according to a quasi-random process with mean arrival rate \( \lambda_{z,m} = (N_{z,m} - n_{z,m}) v_{z,m} \), where \( N_{z,m} \) is the finite population of MUs that can be served by the \( m \)-th RRH of class \( z \), \( n_{z,m} \) is the number of in-service calls in the \( m \)-th RRH of class \( z \) and \( v_{z,m} \) is the mean call arrival rate per idle MU in the \( m \)-th RRH of class \( z \). The offered traffic-load per idle MU in the \( m \)-th RRH of class \( z \) is \( \alpha_{z,m} = v_{z,m} / \mu \). Note that if \( N_{z,m} \to \infty \), for \( z = 1, \ldots, Z \) and \( m = 1, \ldots, M_z \) and the total offered traffic-load remains constant, then the call arrival process becomes Poisson. In that case, the SC-MC model results. Similar to the SC-MC model an arriving call requires a radio RU from the serving RRH and a computational RU from the V-BBU. If these RUs are available then the new call is accepted in the serving RRH and remains for a generally distributed service time with mean \( \mu^{-1} \). Otherwise, call blocking occurs.

Let the number of in-service calls in all RRHs be described by the steady-state vector \( n = (n_1, \ldots, n_{1,M_1}, n_2, \ldots, n_{1,M_1}, \ldots, n_{Z,1}, n_{Z,1}, n_{Z,2}, \ldots, n_{Z,2}, \ldots, n_{Z,M_Z}) \). Also denote the steady-state vectors \( n_{z,m} = (n_{1,1}, \ldots, n_{1,M_1}, n_{z,1}, \ldots, n_{z,m}, 1, \ldots, n_{z,m}, n_{z,m+1}, \ldots, n_{z,m+n_{z,m}+1}, n_{z,m+1}, \ldots, n_{z,m+n_{z,m}+1}, \ldots) \) and let \( P_{n}(n) \) be the probability distributions of states \( n \), \( n_{z,m} \) and \( n_{z,m}^{+} \), respectively. Assuming that states \( n_{z,m} \) and \( n_{z,m}^{+} \) belong to the state space \( \Omega \), we present in Fig. 2 the state transition diagram for the \( m \)-th RRH of the class \( z \) of the proposed model.

Based on Fig. 2 and since the corresponding Markov chain is reversible, we can write the following Lévy balance equations for states: a) \( n_{z,m} \) and \( n \) (see (20)) and b) \( n \) and \( n_{z,m} \) (see (21)) where \( z = 1, \ldots, Z \) and \( m = 1, \ldots, M_z \):

\[
(N_{z,m} - n_{z,m}) v_{z,m} P_{n}(n_{z,m}) = n_{z,m} \mu P_{n}(n) \tag{20}
\]

\[
(N_{z,m} - n_{z,m}) v_{z,m} P_{n}(n_{z,m}) = (n_{z,m} + 1) \mu P_{n}(n_{z,m}) \tag{21}
\]
The system of local balance equations is satisfied by the following PFS for $z = 1, \ldots, Z$, $m = 1, \ldots, M_z$, and $n \in \Omega$:

$$P_{\text{fin}}(n) = G^{-1} \left( \frac{Z}{n} \prod_{z=1}^{Z} \prod_{m=1}^{M_z} (n_{z,m})^{\alpha_{z,m,\text{idle}}} \right)$$

where: $G \equiv G(\Omega) = \sum_{n \in \Omega} \prod_{z=1}^{Z} \prod_{m=1}^{M_z} (n_{z,m})^{\alpha_{z,m,\text{idle}}}$ and $\Omega$ is the set of all possible states given by $\Omega = \{ n : 0 \leq n_{z,1}, \ldots, n_{z,M_z} \leq C_z, 0 \leq \sum_{z=1}^{Z} \sum_{m=1}^{M_z} n_{z,m} \leq T \}$.

In a system that services quasi-random traffic, we distinguish CBP in call congestion (CC) and time congestion (TC) probabilities. The latter refer to the proportion of time the system is congested. On the other hand, CC probabilities coincide with CBP. CC probabilities are only slightly lower than TC probabilities in quasi-random loss models, especially if the number of traffic sources is high [7]. Assuming Poisson arrivals (i.e., the call arrival process adopted in the SC-MC model), TC and CC probabilities coincide (PASTA property, [7]).

**B. The Proposed Convolution Algorithm**

For the efficient calculation of the TC probabilities $B_{z,r}^{TC}$, $B_{z,r}^{TC}$ (and consequently the total TC probability $B_{z,r}^{TC} = B_{z,r}^{TC} + B_{z,r}^{TC}$) which refer to the proportion of time that there exist insufficient radio and computational RUs, respectively, we exploit the fact that the SC-MC model has a PFS and propose the following 3-step convolution algorithm:

**Step 1**
For each class $z$ of RRHs ($z = 1, \ldots, Z$) determine the occupancy distribution of each of the $M_z$ RRHs, $q_{\text{fin},z,m}(j)$, where $m = 1, \ldots, M_z$ and $j = 1, \ldots, C_z$, via:

$$q_{\text{fin},z,m}(j) = \begin{cases} \left( \frac{N_z}{n} \right)^{\alpha_{z,m,\text{idle}} q_{\text{fin},z,m}(0)} & \text{for } j = 1, \ldots, C_z \\ 0 & \text{for } j = C_z + 1, \ldots, T \end{cases}$$

The values of $q_{\text{fin},z,m}(j)$ can be normalized by the normalization constant $G_{\text{fin},z,m} = \sum_{j=0}^{C_z} q_{\text{fin},z,m}(j)$ and are denoted via $q_{\text{fin},z,m}(j) = q_{\text{fin},z,m}(j)/G_{\text{fin},z,m}$. Note that the first step of the proposed convolution algorithm treats each RRH individually from the rest RRHs. This treatment does not lead to exact TC probabilities unless the model has a PFS for the steady-state probabilities.

**Step 2**
Following the successive convolution of all RRHs (of all classes $Z$), apart from the $(z, 1)$ RRH, determine the aggregated occupancy distribution:

$$Q_{\text{fin},z,-1}(j) = q_{\text{fin},z,1}(j) \ast q_{\text{fin},z,2}(j) \ast \cdots \ast q_{\text{fin},z,M_z}(j)$$

The convolution operation between two occupancy distributions $q_{\text{fin},z,u}$ and $q_{\text{fin},z,u}$ is given by (13).

**Step 3**
Calculate the values of $B_{z,r}^{TC}$ based on the normalized values obtained from the convolution operation of step 2, as follows:

$$B_{z,r}^{TC} = G_{\text{fin},z}^{-1} \left( q_{\text{fin},z,1}(C_z) \sum_{r=0}^{T-C_z-1} Q_{\text{fin},z,-1}(r) + q_{\text{fin},z-1}(T) \right)$$

where: $q_{\text{fin},z,1}(C_z)$ refers to the case of unavailable radio RUs in the $(z, 1)$ RRH (already determined in step 1), $Q_{\text{fin},z,-1}(r)$ are the normalized values of $Q_{\text{fin},z,-1}(r)$, while $q_{\text{fin},z-1}(T)$ refers to the un-normalized probability of unavailable computational RUs, given by $q_{\text{fin},z-1}(T) = \Sigma_{r=0}^{T} Q_{\text{fin},z,-1}(r) \cdot q_{\text{fin},z-1}(T-r)$ and $G$ is the normalization constant of the final convolution operation $Q_{\text{fin},z,-1}(z) \ast q_{\text{fin},z-1}$ determined via (13).

As far as the CC probabilities for a call in the $m$-th RRH of class $z$ are concerned, they are determined for a system with $N_{z,m} = 1$ traffic sources.

The occupancy distribution of the computational RUs can be obtained as follows:

$$q_{\text{fin},z-1}(j) = \sum_{r=0}^{T} Q_{\text{fin},z,-1}(r) \cdot q_{\text{fin},z-1}(j-r)/G_{\text{fin},z-1}, \quad j = 1, \ldots, T$$

where: $G_{\text{fin}}$ is the normalization constant of the final convolution operation $Q_{\text{fin},z,-1}(z) \ast q_{\text{fin},z-1}$.

Note that when $Z = 1$, then all RRHs form a single cluster in the C-RAN and we have the f-SC-SC model. In that case, the interested reader may resort to [39] for the analysis of the model. Herein, we only present the corresponding 3-step convolution algorithm for the calculation of the TC probabilities [39]:

**Step 1**
For each of the $M$ RRHs, calculate the occupancy distribution $q_{\text{fin},m}(j)$, where $m = 1, \ldots, M$ and $j = 1, \ldots, C$, via:

$$q_{\text{fin},m}(j) = \begin{cases} \left( \frac{N_z}{n} \right)^{\alpha_{z,m,\text{idle}} q_{\text{fin},m}(0)} & \text{for } j = 1, \ldots, C \\ 0 & \text{for } j = C + 1, \ldots, T \end{cases}$$

The normalized values of $q_{\text{fin},m}(j)$ are expressed as $q_{\text{fin},m}(j) = q_{\text{fin},m}(j)/G_{\text{fin},m}$, where $G_{\text{fin},m} = \Sigma_{j=0}^{C} q_{\text{fin},m}(j)$.

**Step 2**
Following the successive convolution of all RRHs, apart from the 1st RRH, determine the aggregated occupancy distribution:

$$Q_{\text{fin},-1} = q_{\text{fin},2} \ast \cdots \ast q_{\text{fin},M}$$

Taking into account two occupancy distributions $q_{\text{fin},u}$ and $q_{\text{fin},v}$, their convolution operation is given by (13).
Step 3
Calculate the values of the total TC probability, \( B^{TC}_{\text{tot}} \), based on the normalized values obtained in step 2, as follows:

\[
B^{TC}_{\text{tot}} = B^{TC}_{\text{c}} + B^{TC}_{\text{a}} = G_{\text{fin}}^{-1} \left( \sum_{r=0}^{T-C-1} Q_{\text{in},(-1)}(r) + q_{\text{fin}}(T) \right)
\]

where:

\[
q_{\text{fin}}(C) = \frac{Q_{\text{in}, (-1)}(C)}{G_{\text{fin}}}
\]

\[
q_{\text{fin}}(j) = \sum_{r=0}^{j} Q_{\text{in}, (-1)}(r) \cdot q_{\text{fin}}(j-r) / G_{\text{fin}}, \quad j = 1, \ldots, T
\]

where: \( G_{\text{fin}} \) is the normalization constant of the final convolution operation \( Q_{\text{in},(-1)} \ast q_{\text{fin}} \).

Figure 3 summarizes the proposed methodology which is required in order to determine TC probabilities in the f-SC-MC model. Initially, it is important to have a PFS for the steady-state probabilities of the proposed model. Due to the fact that the model can be analyzed via a reversible Markov chain, we can obtain the PFS of \( (22) \). Based on the PFS, we can propose a convolution algorithm for the exact determination of TC probabilities.

IV. NUMERICAL RESULTS

In this section, we present two application examples. In the first example, we provide analytical and simulation TC probabilities results of the f-SC-MC model of [39] and the corresponding analytical results of the SC-SC model of [33]. In the second example, we provide analytical and simulation TC probabilities results of the proposed f-SC-MC model and the corresponding analytical results of the SC-MC model of [38]. Simulation results are based on the Simscript III simulation language [40] and are mean values of 7 runs. In each run, two hundred million calls are generated while the first 5% of these generated calls are not considered in the TC probabilities results so as to account for a warm-up period. In addition, we do not present reliability ranges in the figures of this section, in order to increase their readability. These reliability ranges are less than two order of magnitude.

In the first example, consider a C-RAN of \( M = 6 \) RRHs in the form of a single cluster. All RRHs have the same capacity \( C = 5 \) radio RUs. On the other hand, the V-BBU consist of \( T \) computational RUs. We consider two values of \( T \): 1) \( T = M \cdot C = 30 \) and 2) \( T = 24 \). Regarding the number of finite sources (or MUs) per RRH, we consider three values: a) \( N = 10 \), b) \( N = 50 \) and c) \( N = 100 \) sources [39]. In each of the \( M \) RRHs, let the offered traffic-load per idle MU be \( \alpha_{\text{m,idle}} = \alpha_{m}/N \) where \( \alpha_{m} \) expresses the value of the Poisson traffic. Contrary to the initial values of \( \alpha_{m} \) chosen in [39] \( \alpha_{m} = 2.0 \) erl for \( m = 1, \ldots, 6 \), herein we choose \( \alpha_{m} = 3.0 \) erl for \( m = 1, \ldots, 6 \).

In the x-axis of Figs. 4-7, the value of \( \alpha_{m} \) increases in steps of 0.2 erl. So, point 1 is: \((\alpha_{1}, \alpha_{2}, \alpha_{3}, \alpha_{4}, \alpha_{5}, \alpha_{6}) = (3.0, 3.0, 3.0, 3.0, 3.0, 3.0, 3.0)\) while point 11 is: \((\alpha_{1}, \alpha_{2}, \alpha_{3}, \alpha_{4}, \alpha_{5}, \alpha_{6}) = (5.0, 5.0, 5.0, 5.0, 5.0, 5.0)\). Contrary to the results presented in [39], the values of \( \alpha_{m} \) chosen herein help us investigate an offered traffic-load region that leads to higher TC probabilities results (compare Figs. 3-6 of [39] with the corresponding figures of this example).

Figure 4 shows the simulation and analytical TC probabilities of \( B_{r} \) for all values of \( N \).
Figure 6 presents the simulation and analytical TC probabilities of $B_r$ for all values of $N$ and both values of $T$. We observe that a decrease in $T$ results in a decrease of $B_r$ especially for high values of the offered traffic-load. This behavior can be explained as follows: a decrease in $T$ leads many new calls to blocking due to the unavailability of computational RUs, and therefore the number of available radio RUs increases.

Figure 7 presents the simulation and analytical TC probabilities of $B_c$ for all values of $N$ and $T = 24$. We observe that the decrease of $T$ to 24 units, results in a remarkable increase of $B_c$ (compare Figs. 5 and 7) which reveals that particular attention is required when choosing the value of $T$ especially if it is much lower than $M \cdot C$.

In the second example, consider a C-RAN of $Z = 3$ classes of RRHs. The first class includes $M_1 = 10$ RRHs of capacity $C_1 = 8$ radio RUs. The second class includes $M_2 = 5$ RRHs of capacity $C_2 = 12$ radio RUs. Finally, the third class includes $M_3 = 2$ RRHs of capacity $C_3 = 15$ radio RUs. The V-BBU consists of $T = 90$ computational RUs. Regarding the number
of finite sources for the $n$-th RRH (of each class $z = 1, 2, 3$), we consider the values: a) $N_{z,m} = 20$, b) $N_{z,m} = 50$ and c) $N_{z,m} = 100$ sources. The offered traffic-load per idle MU in each of the $M$ RRHs is $\alpha_{z,m,\text{idle}} = \alpha_z/N_{z,m}$ where $\alpha_z$ refers to the value of the Poisson traffic. Initially, we assume that $\alpha_1 = 2.0$ erl for each RRH of the first class ($m = 1, \ldots, 10$). Similarly, let $\alpha_2 = 4.0$ erl for each RRH of the second class ($m = 1, \ldots, 5$) and $\alpha_3 = 6.0$ erl for each RRH of the third class ($m = 1, 2$).

In the x-axis of Figs. 8-10, the value of $\alpha_z$ increases in steps of 0.2 erl. So, point 1 is: $(\alpha_1, \alpha_3, \alpha_3) = (2.0, 4.0, 6.0)$ while point 11 is: $(\alpha_1, \alpha_3, \alpha_3) = (4.0, 6.0, 8.0)$.

Figures 8-10 present the analytical and simulation TC probabilities of $B_{z,r}$ (depicted for presentation purposes as $B_{z,r}$) for each class $z$ and the three values of $N_{z,m}$ (depicted for presentation purposes as $N$). For comparison, we also include the corresponding TC probabilities for the SC-MC of [38]. In all figures, we observe that: a) the existing SC-MC model provides much higher TC probabilities compared to those obtained via the f-SC-MC model, especially when the values of $N$ are low (e.g., when $N = 20$ or 50 sources), b) increasing $N$ results in TC probabilities which are closer to those obtained via the SC-MC model (where by definition $N$ is infinite) and c) simulation results are very close to the analytical results.

V. CONCLUSION

In this paper we propose a new loss model, named f-SC-MC model, for the analysis of a C-RAN that accommodates quasi-random traffic which is smoother than the classical Poisson process considered in [33], [38]. Two different cases are considered regarding the grouping of RRHs. The single cluster case, proposed in [39] and the multi-cluster case proposed herein. The latter generalizes the model of [39] since it assumes that RRHs are grouped in different clusters according to their capacities in radio RUs. Our first major contribution is to show that the steady-state probabilities of the proposed model in the multi-cluster case (f-SC-MC model) have a PFS. This is an essential step in order to proceed with our second major contribution which is a convolution algorithm for the exact determination of TC probabilities in the f-SC-MC model. Indeed, the first step of the proposed convolution algorithm treats each RRH individually from the rest RRHs. This treatment does not lead to exact TC probabilities unless the model has a PFS for the steady-state probabilities. The accuracy of the proposed algorithm is verified via simulation and found to be very satisfactory. As a future work we intend to study call arrival processes that are burstier than the Poisson process such as the batched Poisson process [41]–[44].

REFERENCES

[1] M. Habibi, M. Nasimi, B. Han, and H. Schotten, “A comprehensive survey of RAN architectures toward 5G mobile communication system”, IEEE Access, pp. 70371–70421, May 2019. doi: 10.1109/ACCESS.2019.2919657

[2] M. Hossain, A. Mahin, T. Debnath, F. Mosharraf, and K. Ziaul, “Recent research in cloud radio access network (C-RAN) for 5G cellular systems A survey”, Journal Netw. Comp. Appl., vol. 139, pp. 31–48, Aug. 2019. doi: 10.1016/j.jnca.2019.04.019

[3] A. Checko et al., “Cloud RAN for mobile networks — A technology overview”, IEEE Commun. Surveys Tuts., vol. 17, no. 1, pp. 405–426, 1st Quart., 2015. doi: 10.1109/COMST.2014.235525

[4] CPRI Consortium, “CPRI Specification V7.0 Common Public Radio Interface (CPRI). Interface Specification”, 2015.

[5] ETSI GS NFV-MAN 001 (V1.1.1), “Network Function Virtualisation (NFV): Management and Orchestration”, Dec. 2014.
Iskanter-Alexandros Chousainov received the B.S. degree from the Dept. of Mathematics of the University of the Aegean in 2002, the M.Sc degree in Information Systems from Hellenic Open University in 2017. He is currently a Ph.D. student at the Dept. of Informatics & Telecommunications, University of Peloponnese, Tripolis, Greece. His research interests include teletraffic engineering and performance evaluation of 5G networks.

Ioannis D. Moscholios was born in Athens, Greece, in 1976. He received the Dipl.-Eng. degree in Electrical & Computer Engineering from the University of Patras, Patras, Greece, in 1999, the M.Sc. degree in Spacecraft Technology & Satellite Communications from the University College London, UK, in 2000 and the Ph.D. degree in Electrical & Computer Engineering from the University of Patras, in 2005. From 2005 to 2009 he was a Research Associate at the Wire Communications Laboratory, Dept. of Electrical & Computer Engineering, University of Patras. From 2009 to 2013 he was a Lecturer in the Dept. of Telecommunications Science and Technology, University of Peloponnese, Tripolis, Greece. Currently, he is an Associate Professor in the Dept. of Informatics & Telecommunications, University of Peloponnese, Tripolis, Greece. His research interests include teletraffic engineering, simulation and performance analysis of communication networks. He has published over 130 papers in international journals, conferences and book chapters and has filed several patents for topics related to 5G networks.

Alexandros Kaloxyllos (B.Sc., M.Phil., Ph.D.) received his Ph.D. degree in Informatics and Telecommunications from the University of Athens in 1999. Since 1994, he has participated, as a researcher, in numerous projects realized in the context of EU Programmes as well as national initiatives. In 2002 he joined the faculty of the University of Peloponnese, where he is currently an Associate Professor in the Department of Informatics and Telecommunications. From 2014 until 2017, he worked as a Principal Researcher in Huawei’s European Research Center in Munich, where he led the Radio Access Network team in the design of 5G networks. During 2016, he was the vice-chair of the 5G PPP Architecture WG. Since October 2019, he is the Technology Board Chairman of 5G-PPP. He has published over 170 papers in international journals, conferences and book chapters and is a co-author of the book: Efficient Multirate Teletraffic Loss Models Beyond Erlang (IEEE Press, Wiley, April 2019).

Michael D. Logothetis (SM’08) received the Dipl.Eng. and Ph.D. degrees in electrical engineering from the University of Patras, Patras, Greece, in 1981 and 1990, respectively. From 1982 to 1990, he was a Teaching and Research Assistant with the Laboratory of Wire Communications, University of Patras. From 1991 to 1992, he was Research Associate with NTT’s Telecommunication Networks Laboratories, Tokyo, Japan. He was a Lecturer with the Electrical and Computer Engineering Department, University of Patras, and in 2009 was elected a (Full) Professor in the same department. He has authored over 220 conference/journal papers and is a co-author of the book: Efficient Multirate Teletraffic Loss Models Beyond Erlang (IEEE Press, Wiley, April 2019). His current research interests include teletraffic theory and engineering, traffic/network control, simulation and performance optimization of telecommunications networks. He organized (was General Chair) of the IEEE/IET CSNDSP 2006 and IEICE ICTF 2016. He has served as a TPC member of several international conferences while he has chaired many technical sessions. He was a Guest Editor of five journals: Mediterranean Journal of Electronics and Communications, Mediterranean Journal of Computers and Networks, IET Circuits, Devices & Systems, IET Networks, and Ubiquitous Computing and Communication Journal. He is a senior member of IEE, a member of the IARIA (Fellow), the IEICE, the FITCE, and the Technical Chamber of Greece.