On Optimizing Power Allocation For Reliable Communication over Fading Channels with Uninformed Transmitter

M. Majid Butt, *Senior Member, IEEE,* Eduard A. Jorswieck, *Senior Member, IEEE* and Nicola Marchetti, *Senior Member, IEEE*

***Abstract*—We investigate energy efficient packet scheduling and power allocation problem for the services which require reliable communication to guarantee a certain quality of experience (QoE). We establish links between average transmit power and reliability of data transfer, which depends on both average amount of data transfer and short term rate guarantees. We consider a slow-fading point-to-point channel without channel state information at the transmitter side (CSIT). In the absence of CSIT, the slow fading channel has an outage probability associated with every transmit power. As a function of data loss tolerance parameters, and minimum rate and peak power constraints, we formulate an optimization problem that adapts rate and power to minimize the average transmit power for the user equipment (UE). Then, a relaxed optimization problem is formulated where transmission rate is assumed to be fixed for each packet transmission. We use Markov chain to model constraints of the optimization problem. The corresponding problem is not convex for both of the formulated problems, therefore a stochastic optimization technique, namely the simulated annealing algorithm, is used to solve them. The numerical results quantify the effect of various system parameters on average transmit power and show significant energy savings when the service has less stringent requirements on timely and reliable communication.**

***Index Terms*—Energy efficiency, power control, packet scheduling, bursty packet loss, stochastic optimization, simulated annealing, URLLC.**

I. Introduction

Internet of things (IoT) is one of the use cases of 5G wireless communications to serve the heterogeneous services. Services like smart city, smart buildings and smart transportation systems depend heavily on efficient information processing and reliable communication techniques. The use of thousands of smart and tiny sensors to communicate regular measurements, e.g., temperature, traffic volume, etc., makes it extremely important to look at the energy efficiency aspect of the problem. Achieving ultra reliability and low latency communication (URLLC) at low energy for the emerging applications in 5th generation (5G) of wireless communication is considered very challenging [2]. Due to short packet size in IoT and machine type communication, finite block-length channel codes, novel diversity techniques, packet dropping mechanisms and control plane communication strategies are considered to enable URLLC [3]–[6].

In 5G networks, context aware scheduling is believed to play key role in smart use of resources [7] and the requirements on reliability and latency are dictated by the nature of the application. More specifically, IoT applications have extremely heterogenous requirements in terms of (average or deadline) latency, reliability and frequency of packet transmissions, and require quality of service (QoS) aware resource allocation mechanisms [8]. Depending on the application's context, it may not be necessary to receive every packet correctly at the receiver side to avoid experiencing a serious degradation in quality of experience (QoE). For instance, ITU recommendation ITU-T G.1080 (12/2008) specifies a set of requirements for picture/audio that define the quality impairments in addition to average packet loss rates [9]. If some packets are lost, the application may tolerate the loss without requiring retransmissions of the lost packets. The application loss tolerance without degrading quality can effectively be exploited to reduce average energy consumption of the devices.

We investigate energy efficient power allocation for the wireless systems with data loss constraints. The packet loss constraints are defined in terms of average packet loss and the maximum number of successively lost packets. The reliability aspect of the communication systems is conventionally handled at upper layers of communication using error correction codes and/or hybrid automatic repeat request (HARQ). Feedback based link adaptation applied in HARQ is dictated by the latency constraints of the application [10]. Our approach is different from the HARQ scheme because the simple device nodes do not possess a data buffer, which makes implementing HARQ systems impossible. Instead, we assume that the applications's QoE does not require every packet to be received successfully, i.e., loss of successive packets can be tolerated, but it must be bounded and parameterized. Video streaming, video conferencing, disaster management systems and interactive gaming are examples of such applications.

In literature, some earlier works have addressed similar problems in different settings and contexts (more at network

The material in this paper has been presented in part in ICC 2017, Paris, France [1].

M. Majid Butt and Nicola Marchetti are with CONNECT center for future networks, Trinity College Dublin, Dublin 2, Ireland. Email:{majid.butt, nicola.marchetti}@tcd.ie.

Eduard A. Jorswieck is with Department of Electrical Engineering and Information Technology, TU Dresden, Germany. Email: eduard.jorswieck@tu-dresden.de.

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund under Grant Number 13/RC/2077.

The work of E. Jorswieck was supported by the German Research Foundation, Deutsche Forschungsgemeinschaft, Collaborative Research Center 912, through the Highly Adaptive Energy- Efficient Computing.

level). In video streaming applications, it is important to select source coding parameters for various encoded representations of the same content in order to minimize the consumption power while maintaining a high quality of experience for the users [11]. Packet scheduling for media streaming with network coding has been studied in [12] with the objective to improve the perceived media quality. Wu *et al.* address the problem of bursty packet loss over internet in [13]. The authors propose a transmission scheme that trades delay to reduce the distortion in transmitted data. Similar works in [14], [15] address delay-quality tradeoff in video transmission over communication links. In [16], the authors evaluate the subjective and objective performance of video traffic for bursty loss patterns. Reference [17] considers real-time packet forwarding over wireless multi-hop networks with lossy and bursty links. The objective is to maximize the probability that individual packets reach their destination before a hard delay deadline. In a similar study, the authors in [18] investigate a scenario where multimedia packets are considered lost if they arrive after their associated deadlines. Lost packets degrade the perceived quality at the receiver, which is quantified in terms of the "distortion cost" associated with each packet. The goal of the work in [18] is to design a scheduler which minimizes the aggregate distortion cost over all receivers.

The energy efficiency aspect of the problem has been discussed in many works. Energy can be saved by relaxing various QoS constraints for data transmission. Delay and loss tolerance are two key dimensions to exploit for reducing transmit power. Various works in literature deal with exploiting delay tolerance to optimize transmit power in time varying wireless channels, e.g. [19]–[22]. If the latency requirements for the data permit, the transmission can be delayed and the effect of the random nature of fading wireless channels can be minimized by opportunistic scheduling schemes. The energy aspect of the problem has been addressed in [23] where the authors investigate intentional packet dropping mechanisms for delay limited systems to minimize energy cost over fading links.

Most of the works in literature characterize performance of the wireless network for average packet loss. In addition to average packet loss, bursty data loss is an important phenomenon which needs to be defined, characterized and analyzed. Some works analyze system for bursty traffic, e.g., the effect of access router buffer size on packet loss rate is studied in [24] when bursty traffic is present. However, assumption of bursty traffic is different from the notion of bursty data loss. An analytical framework to dimension the packet loss burstiness over generic wireless channels is considered in [25] and a new metric to characterize the packet loss burstiness is proposed. However, these works do not characterize the effect of average and bursty packet loss on the consumed energy at link level. Some recent studies in [1], [26], [27] characterize the effect of packet loss burstiness on average system energy for a multiuser wireless communication system where the transmit channel state information (CSIT) is fully available or erroneous.

In this work, no CSIT is assumed to be available, which poses new challenges for communication and scheduler design. When CSIT is not available for slow fading channels, channel state dependent power control cannot be applied and outage free communication cannot be guaranteed. For the no-CSIT case, we characterize the average power consumption of the point-to-point wireless channel for various average and bursty packet drop parameters, as well as the outage probability that the application can tolerate loss of a full sequence of packets (successively).

The main contributions of the work are summarized as follows:

- We model and formulate the power optimization problem for a point to point system using a Markov chain. The problem constraints involve various parameters that help characterizing QoE for a particular application, including average and successive packet loss bounds, as well as minimum packet size and long term average rate guarantees. We show that the formulated optimization problem is combinatorial and no closed form solution exists.
- We propose a solution of the optimization problem based on a low complexity stochastic optimization algorithm, namely Simulated Annealing (SA). The algorithm is based on randomization of input parameters. We numerically evaluate the performance of the proposed solution and verify that the algorithm produces results that are very close to the analytical solution for a special case of the problem.
- To reduce the complexity of the problem, we propose a fixed-rate adaptive power transmission scheme. The fixed rate transmission scheme inherits all the constraints of the original optimization problem, but the transmitted rate is the same for every transmission. This helps in reducing computational complexity for the problem.
- Simulation results show that our power allocation scheme exploits packet loss tolerance of the application to save considerable amount of energy; and thereby significantly improves the energy efficiency of the network as compared to lossless application case.

The rest of the paper is organized as follows. The system model for the work is introduced in Section II and state space description of the proposed scheme is discussed in Section III. We formulate the optimization problem to minimize average transmit power in Section IV. Then, we discuss a modified optimization problem in Section V where all the transmissions are of fixed rate. We discuss solution of both of the optimization problems using SA algorithm in Section VI. The performance of the proposed framework is numerically evaluated in Section VII and Section VIII summarizes the main results of the paper.

II. System Model

We consider a point-to-point system such that the transmitter user equipment (UE) has a single packet to transmit in each time slot. The packets are assumed to be variable in size, measured in bits/s/Hz. This is achieved by rate adaptation at physical layer using well known adaptive modulation and coding techniques. Time is slotted and the UE experiences quasi-static independent and identically distributed (i.i.d) block flat-fading such that the fading channel remains constant for

the duration of a block, but varies from block to block whereas duration of the block equals one time slot.

We assume no CSIT, but the transmitter is aware of the fading channel distribution. Depending on the scheduling state i (explained later in Section III), the UE transmits with a fixed power $P_i \leq P_m$ to transmit a packet with size R_i bits/s/Hz, and waits for the feedback. P_m is the peak transmit power constraint for the transmitter. For convenience, the distance between the transmitter and the receiver is assumed to be normalized.

For a transmit power P_i, and channel fading coefficient h, the outage probability for the failed transmission (channel outage) is denoted by ϵ_i such that,

$$\epsilon_i = \Pr\left[\log_2\left(1+\frac{P_i|h|^2}{N_0}\right) < R_i\right] \tag{1}$$

where N_0 is additive white Gaussian noise power.

If the transmitted packet is received at the receiver correctly, the receiver sends back a positive acknowledgement (ACK) message to the UE. If it is not decoded at the receiver, a negative acknowledgement (NAK) is fed-back to the UE. The feedback is assumed to be perfect without error. Note that a power and/or rate adaptation based on the feedback can be applied even without CSIT. Feedback based power allocation belongs to Restless Multi-armed Bandit Processes where the states of the UE in the system stochastically evolve based on the current state and the action taken. The UE receives a reward depending on its state and action. The next action depends on the reward received and the resulting new state. In this work, we investigate the effect of feedback based sequential decisions in terms of UE consumed average power.

A. Problem Statement

A single packet arrives at the transmit buffer of the UE in every time slot. The UE's data buffer has no capacity to store more than one packet.[1] This is a typical scenario for a wireless sensor network application where data measurements arrive constantly after regular fixed time intervals. The UE is battery powered, which needs to be replaced after regular intervals. It is therefore, important to save transmit energy as much as possible. Depending on the application, the UE has two constraints on the reliability of data packet transfer [26], [27]:

1) *Average packet drop/loss rate* γ is the parameter that constraints the average number of packets dropped/lost,

$$\gamma = \lim_{t\to\infty} \frac{\text{Packets dropped}}{\text{Packets transmitted}} \tag{2}$$

2) Maximum number of packets dropped successively. This is called *bursty packet drop constraint*. The parameter N denotes the *maximum number of packets allowed to be dropped successively without degrading QoE below a certain level*. Mathematically, the distance $r(q, q-1)$ between q^{th} and $q^{th}-1$ correctly received packets measured in terms of number of successively lost packets is constrained by parameter N, i.e.,

$$r(q, q-1) \leq N. \tag{3}$$

[1]Note that the buffer capacity is given by the largest rate R_{max} bits/s/Hz, that can be transmitted in any state.

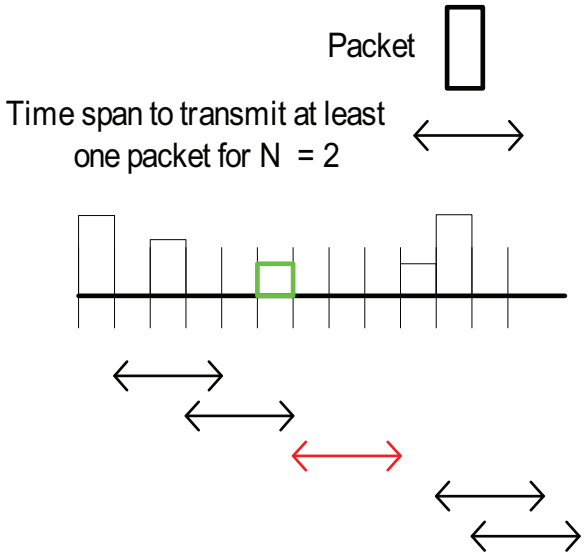


Fig. 1. Schematic diagram of the system with $N = 2$. The transmitted packets can be of variable size as shown in the diagram. The time span for the successive packet loss constraint has been drawn. The red period shows the instance when violation of successive packet loss occurs. The transmission of green packet in third time slot shows that constraint $N = 2$ was met.

Due to transmit power constraint, it is not possible to provide the guarantee in (3) with probability one. Given at least N packets have been lost successively by time instant $t-1$, we define a parameter ϵ_{out} at an instant t by the probability that the $N+1-th$ packet is lost, i.e.,

$$\epsilon_{out} = \Pr\left(r_t(q,q-1) = r_{t-1}(q,q-1)+1 | r_{t-1}(q,q-1) \geq N\right) \tag{4}$$

Fig. 1 shows the schematic diagram for the system. The successive packet drop parameter N equals 2 in the diagram. Whenever a packet is transmitted successfully, it is permitted to drop 2 packets in the next 2 time slots. If a packet is not transmitted successively in 3^{rd} successive time slot, it counts as an outage ϵ_{out}. The span of 3 successive packet drops have been depicted as red in the schematic diagram. Note that the transmitted packets have variable size in bits/s/Hz, which is constrained by minimum and maximum rate R_{min} and R_{max}, respectively. If a packet is transmitted immediately after the transmission of a packet in previous time slot, its size (rate) is more as compared to the packets transmitted when the packets have been lost already. We come back to rate adaptation and optimization later in the next section.

All of these parameters described in this section contribute to the QoE for the application. The average packet drop rate is commonly used to characterize a wireless network and bounds the QoE for the application. However, bursty packet loss in the applications like smart monitoring sensors can degrade the performance enormously due to absence of contiguous data measurements. On the other hand, the UE can exploit the

parameters γ and N to optimize average energy consumption if the application is more loss tolerant. If the application is loss tolerant and packet size is fixed, it is advantageous to transmit with a small power if a packet has just been received successfully in the last time slot because the impact of packet loss due to outage is not so severe on cumulative QoE. The consideration of bursty (successive) packet loss poses a new challenge in system modeling as the number of packets lost in previous time slots affect the power allocation decision at time slot t. Clearly, there is a trade-off between transmitting a packet at time t with small power based on the success of transmission in time slots $[t-1, t-2, \dots]$, and transmitting with large power to lower the risk of outage. This trade-off determines the power allocation policy. Let us illustrate the impact of ACKs and NAKs on the tightness of the constraints in the following:

If the permitted average packet loss rate γ is very high but N is small, i.e., it is not permitted to lose more than N packets successively without degrading QoE, the effective average packet drop rate becomes much lower than the permitted γ in this case. It may work to transmit with small power due to large γ, but parameter N does not allow it. Due to successive packet drop constraint N, transmission of a packet in a time slot t may not be as critical as in any other time slot with $t' \neq t$. If a packet was transmitted successfully in a time slot $t-1$, it implies that transmitting a packet with a lower power is not as risky in time slot t. However, when the number of successively lost packets approaches N, power allocation needs to be increased proportionally to avoid/minimise the event of missing N packets successively, which may cause loss of important information for wireless sensor networks.

A similar justification can be provided for rate adaptation for transmission in various states. If a transmission is made in the beginning, packet rate can be chosen a bit higher as risk involved due to dropping of a packet is not that great. When more successive packets are dropped, the rate must be decreased to increase the probability of success as depicted in Fig. 1. The rate is lower bounded by R_{min}, a system parameter, while the upper bound is obtained from the Shanon capacity with the peak power constraint P_m such that,

$$R_{max} = \log_2\left(1 + \frac{P_m|h|^2}{N_0}\right) \tag{5}$$

With every unsuccessful packet transmission, the response of the transmitter is to reduce the rate to increase the success probability, though it is not straight forward to see how this adaptation needs to be applied. The objective is to reduce the average transmit power, therefore rate and power adaptation with every unsuccessful transmission should be optimized in a way that QoE in terms of successful data transfer according to the parameters provided should be met and the transmit power is not wasted unnecessarily.

III. State Space Description

To model the problem, we take history of the packet transmissions in the last N time slots into account. If a NAK is received in time slot $t-1$, it needs to be determined whether transmission in time slot $t-2$ was an ACK or NAK. To capture the time evolution of the packet transmissions in successive time slots, we model the problem using a Markov chain where the next state only depends on the current state and is independent of the history. In a Markov chain, Markov state i is defined by the number of packets lost successively at the transmit time t. If a packet was transmitted successfully in time slot $t-1$, the current state $i = 0$. If two successive packets are lost in time slots $t-1$ and $t-2$, $i = 2$. The maximum number of Markov states is determined by parameter N, i.e., the bursty packet drop constraint. In the following, we explain how Markov chain process can effectively be used to model and formulate power allocation optimization problem for average and bursty packet drop constraints.

To explain the state transition mechanism, let us examine the power allocation policy first. At the beginning of the Markov chain process, a packet is transmitted with power P_0 and rate R_0 in a time slot t with initial Markov state $i = 0$. The channel has an outage probability of ϵ_i (defined in (1)). If the received feedback is ACK, the process moves back to state 0, otherwise moves to state 1. The lost packet is dropped permanently as UE has no buffer. In state $i = 1$, the new arriving packet is transmitted with power P_1 and rate $R_1 \geq R_{min}$. Thus, power allocation in state i is a function of outage probability ϵ_i and the rate R_i,

$$P_i = f(\epsilon_i, R_i). \tag{6}$$

If the packet is transmitted successfully, the next state is zero, 2 otherwise. Similarly, the Markov chain makes a transition to either state $i+1$ or state zero corresponding to the event of unsuccessful or successful transmission, respectively. When $i = N$ (termination state) and a packet is not transmitted successfully, this defines the outage event for successive packet loss. This is modeled by self state transition probability α_{NN} of staying in state S_N such that,

$$\alpha_{NN} = \epsilon_N = \Pr(S_{t+1} = N | S_t = N). \tag{7}$$

P_N is chosen such that $\alpha_{NN} \leq \epsilon_{out}$ where ϵ_{out} is a system parameter defined in (4). If a packet is lost in state N, we want Markov process to stay in state N for the next time slot to minimize further degradation in QoE as rate and power levels in state N are designed to maximize the possibility of successful transmission.

The state transitions from state i to j occur with a state transition probability α_{ij}. It is a function of parameters γ, N and channel distribution. For every transmit power P_i, there is an associated state transition probability α_{ij}.

Formally, the state transition probability α_{ij} from the current state $S_t = i$ to next state $S_{t+1} = j$ is defined by,

$$\alpha_{ij} = \Pr(S_{t+1} = j | S_t = i) \tag{8}$$

$$= \begin{cases} 1-\epsilon_i, & \text{if ACK Received}, \forall i, j = 0 \\ \epsilon_i, & \text{if NAK Received}, i \neq N, j = i+1, \\ & 0 \leq \epsilon_i \leq 1 \\ \epsilon_N, & \text{if NAK Received}, i = N, j = N \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

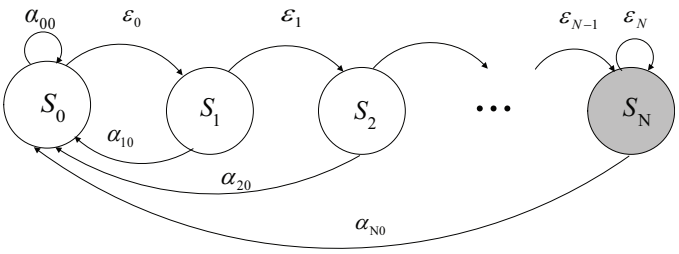


Fig. 2. State diagram for the Markov chain for the UE power allocation scheme.

where ϵ_i is given by (1). The resulting state diagram is shown in Fig. 2. The state transition probability matrix $\mathbf{A} = [\alpha_{ij}]_{i,j=0}^{N}$ takes the form

$$\mathbf{A} = \begin{pmatrix} 1-\epsilon_0 & \epsilon_0 & 0 & \dots & 0 \\ 1-\epsilon_1 & 0 & \epsilon_1 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 1-\epsilon_{N-1} & 0 & 0 & \dots & \epsilon_{N-1} \\ 1-\epsilon_N & 0 & 0 & \dots & \epsilon_N \end{pmatrix} \tag{10}$$

For a time homogeneous Markov chain, the steady state probability for state j, π_j is defined by

$$\pi_j = \sum_{i \in \mathcal{S}} \alpha_{ij} \pi_i \tag{11}$$

where $\mathcal{S}$ defines the state space for the UE states. Assuming $N_0 = 1$, for Rayleigh fading and state i, the outage probability is given by,

$$\epsilon_i \quad = \quad 1 - \exp\Big(\frac{-(2^{R_i}-1)}{P_i}\Big) \tag{12}$$

After some algebraic manipulation, the required transmit power P_i is calculated by,

$$P_i = \frac{1-2^{R_i}}{\log(1-\epsilon_i)} \tag{13}$$

From the transmit power for every state i, the average transmit power consumed is given by,

$$\bar{P} = \sum_{i=0}^{N} P_i \pi_i. \tag{14}$$

IV. Optimization Problem Formulation

The optimization problem is to jointly compute a vector of power values $\mathbf{P} = [P_0, P_1, \dots P_N]$ and $\mathbf{R} = [R_0, R_1, \dots R_N]$, which satisfies the constraints on packet dropping parameters and minimizes average system energy. The problem is mathematically formulated as,

$$\min_{\mathbf{P},\mathbf{R}} \bar{P} = \sum_{i=0}^{N} P_i \pi_i \tag{15}$$

$$s.t. \quad \begin{cases} \mathcal{C}_1 : \sum_{i=0}^{N} R_i \pi_i \geq R, \quad 0 \leq i \leq N \\ \mathcal{C}_2 : \sum_{i=0}^{N} \epsilon_i \pi_i \leq \gamma \\ \mathcal{C}_3 : \epsilon_N \leq \epsilon_{out} \\ \mathcal{C}_4 : R_{min} \leq R_i \leq R_{max} \end{cases} \tag{16}$$

The constraints are explained in the following:

- $\mathcal{C}_1$ is the average rate constraint, i.e., the average transmitted rate should be greater than R.
- $\mathcal{C}_2$ is the average packet loss constraint for the target average packet loss probability γ. The left hand side term is denoted by achieved average packet loss probability γ_r. From the state space model described in Section III, it is computed by the sum of the products of steady state and forward state transition probabilities.[2]
  The outage probability ϵ_i and the corresponding transmit power P_i for a UE in state i is computed such that the average packet dropping probability constraint $\mathcal{C}_2$ holds.
- $\mathcal{C}_3$ is the outage constraint. For $i = N$, $\epsilon_N \leq \epsilon_{out}$ where ϵ_{out} is defined in (4). ϵ_i cannot be determined directly and needs to be optimized for the system parameters.
  $$\epsilon_i = f(\gamma, N, \epsilon_{out}, h_X(x), R) \tag{17}$$
  where $h_X(x)$ is the fading channel distribution.
- In $\mathcal{C}_4$, rate R_i is constrained by $R_{min} \leq R_i \leq R_{max}$. R_{min} is the minimum rate that a packet is expected to provide; and depends on the application and the chosen modulation and coding schemes. If we take the example of IoT, we can define a minimum non-zero value of the rate R_{min} that carries the minimum information about the sensed data.

This solution of the problem provides successive as well as average packet loss guarantees. Similar to effective capacity, which gives the delay-limited capacity depending on the buffer decay rate [28], this solution also provides a minimum statistical rate guarantee R_{min} with outage ϵ_{out} over the span of N successive time slots and average rate guarantee R when $t \to \infty$. The packet size R_i is an optimization variable to be jointly computed with P_i for $\forall i$. The offline computed power allocation solution holds for online power allocation as long as the channel distribution remains the same.

The optimization problem is to jointly find rate $\mathbf{R}$ and power $\mathbf{P}$ vectors that result in minimum average power. If we choose P_i too high for small states (states with low number of successive outages), the packets will more likely be transmitted too early at the expense of larger power budget without exploiting loss tolerance of the application and provide good (but unnecessary) QoE. On the other side, if P_i is chosen too low in the beginning, the packets will be lost mostly and we have to transmit with much higher power to meet the *forced* condition that at least one packet has to be transmitted to avoid the sequence of N lost packets.

The formulated optimization problem covers the discussed practical aspects of the IoT applications. We take care of both packet reception frequency (average), packet reception order (successive packet) as well as min rate R_{min} transmitted within any span of N transmissions, and long term average rate R transported by these packets.

A. Complexity of the Programming Problem

There are two main difficulties associated with programming problem (15). The first challenge is to get a tractable

[2]State N is exception where self state transition represents packet loss.

expression for the steady state distribution π. It is obtained from the eigenvector of the state transition matrix $\mathbf{A}$ whose components α_{ij} depend on the outage probability vector $\vec{\epsilon}$ which depends on the transmission rate and power allocation (and thus on the optimization variables $\mathbf{P}$ and $\mathbf{R}$). Even though the state transition matrix $\mathbf{A}$ has a special structure as pointed out in (10), there is no closed form solution for the eigenvectors of this structured matrix. Therefore, the dependency of π on $\mathbf{P}, \mathbf{R}$ via $\mathbf{A}$ prohibits an analytical presentation.

The second challenge arises from the structure of the programming problem as such. In order to derive an efficient algorithm to solve (15), the problem should be jointly convex in $\mathbf{R}$ and $\mathbf{P}$. Let us only consider the constraint on the average dropping rate $\sum_{i=0}^{N} \epsilon_i \pi_i \leq \gamma$. The left side contains the expression $\epsilon_i = 1 - \exp^{-\frac{1-2^R_i}{P_i}} = \phi(R_i, P_i)$ which is a function of R_i and P_i. The constraint in (16) requires an upper bound on the average dropping rate. Since $\phi(R_i, P_i)$ is a concave function, this leads to a non-convex constraint. Therefore, even if a closed form solution for the steady state distribution could be derived, it will not lead to a convex programming problem. In order to gain more insights into the solution structure of the programming problem (15), we consider the special case $N = 1$ next.

B. Special Case: $N = 1$

In this case, the state transition probability matrix $\mathbf{A}$ reads,

$$\mathbf{A} = \begin{pmatrix} 1-\epsilon_0 & \epsilon_0 \\ 1-\epsilon_1 & \epsilon_1 \end{pmatrix} \tag{18}$$

Steady state transition probabilities for states 0 and 1 are calculated as,

$$\pi_0 = \frac{1-\epsilon_1}{1+\epsilon_0-\epsilon_1} \tag{19}$$

$$\pi_1 = \frac{\epsilon_0}{1+\epsilon_0-\epsilon_1}. \tag{20}$$

Computing γ_r for $\epsilon_1 = \epsilon_{out}$ and π_0 and π_1 calculated above

$$\gamma_r = \frac{\epsilon_0}{1+\epsilon_0-\epsilon_{out}}. \tag{21}$$

We can compute the value of ϵ_0 in closed form that satisfies constraints $\mathcal{C}_2$ and $\mathcal{C}_3$ with equality. Solving (21) and $\mathcal{C}_2$ in (16) with equality gives,

$$\epsilon_0 = (1-\epsilon_{out})\frac{\gamma}{1-\gamma}. \tag{22}$$

To compute power levels P_0 and P_1 for the computed ϵ_0 and ϵ_1, we require rates R_0 and R_1 that meet $\mathcal{C}_1$ and $\mathcal{C}_4$ and the resulting power levels minimize $\bar{P}$ from (14). There could exist many (R_0, R_1) pairs that meet $\mathcal{C}_1$ and $\mathcal{C}_4$ and computing the unique combination that minimizes $\bar{P}$ in closed form is not possible. However, under certain assumption on $\mathcal{C}_1$ and $\mathcal{C}_4$, it is possible to compute achievable rates R_0, R_1 in closed form (but not the optimal ones). As minimum acceptable transmit rate is R_{min} from $\mathcal{C}_4$ and we have N as the critical state, we assume that the rate transmitted in state N is R_{min} to maximize the chance of successful transmission.

For $N = 1$ case, using $R_1 = R_{min}$, and meeting $\mathcal{C}_1$ with equality, the rate R_0 turns out to be,

$$R_0 = \frac{R - R_{min}\pi_1}{\pi_0}. \tag{23}$$

Note that the solution is feasible only if $R_0 \leq R_{max}$. If $R_0 > R_{max}$, it implies R_{min} is not enough to meet $\mathcal{C}_1$ and needs to be increased.

The resulting power allocation P_0 and P_1 are given by

$$P_0 = \frac{1 - 2^{\left(\frac{R - R_{min}\pi_1}{\pi_0}\right)}}{\log(1 - \frac{\gamma(1-\epsilon_{out})}{1-\gamma})} \tag{24}$$

$$P_1 = \frac{1 - 2^{R_{min}}}{\log(1-\epsilon_{out})} \tag{25}$$

The resulting average power $\bar{P}$ in closed form is given by using (14),

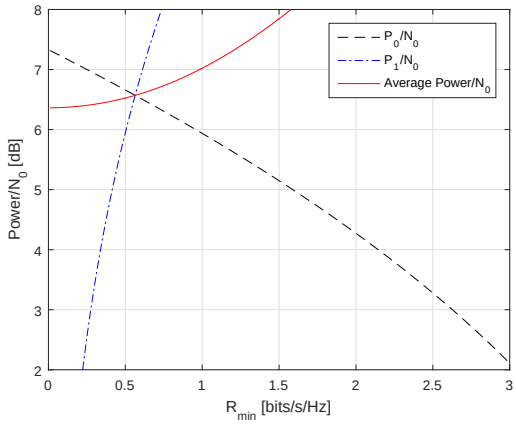
$$\bar{P} = \frac{1 - 2^{\left(\frac{R - R_{min}\pi_1}{\pi_0}\right)}}{\log\left(1 - \frac{\gamma(1-\epsilon_{out})}{1-\gamma}\right)}\pi_0 + \frac{1 - 2^{R_{min}}}{\log(1-\epsilon_{out})}\pi_1 \tag{26}$$

For this specific case, we evaluate the average power in Fig. 3. We fix $\epsilon_1 = \epsilon_{out}$ and compute ϵ_0 using (22). Then, for the (ϵ_0, ϵ_1) pair, we vary $R_1 = R_{min}$ and compute $\bar{P}$ for R_0 in (23). For a fixed rate R, γ, ϵ_0 and ϵ_1, various combinations of (R_0, R_1) provide various average power. We chose values of ϵ_1 such that it is less than γ in Fig. 3(a) and equal to γ in Fig. 3(b). For every ϵ_{out}, minimum $\bar{P}$ is achieved at a certain R_{min}. Please note that the intention is not to optimize $\bar{P}$ in Fig. 3; but to show how $\bar{P}$ varies as a function of R_0 and there is no mechanism to compute optimal $\bar{P}$ in closed form even for the simplest case of $N = 1$. The optimal power allocation can be computed only by jointly searching all possible combinations of (ϵ_0, ϵ_1) and (R_0, R_1). We observe that as ϵ_{out} increases, minimum power is achieved at large R_1. At $\epsilon_{out} = 0.1$, the optimal policy is to transmit with very small $R_1 = R_{min}$ while optimal power is achieved at larger $R_1 = R_{min}$ when $\epsilon_{out} = 0.2$.

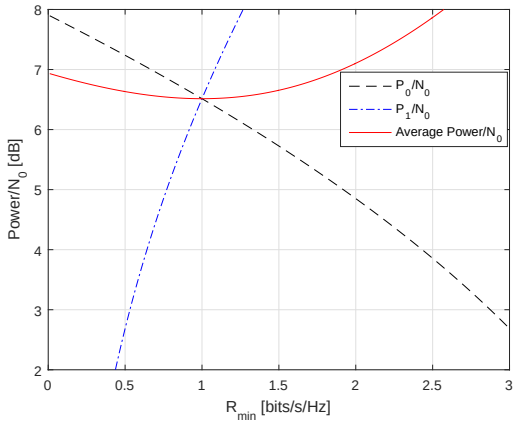
Fig. 3 provides us an interesting insight on optimal power allocation for the problem formulated in (15). When ϵ_{out} and/or target rate R is high, the difference between R_0 and R_1 is not very large for the optimal power allocation. We use this intuition to propose a relaxed optimization problem in next section and compare performance of solutions of both problems in Section VII.

V. Fixed Rate Transmission

In the optimization problem formulated in Section IV, the short term rate guarantee R_{min} over N successive time slots can be quite small as compared to average rate R. However, some applications require higher short term minimum rate guarantees such that $R_{min} \to R$. This leads us to a more restrictive optimization problem where packet size for each



(a) $\epsilon_{out} = 0.1$



(b) $\epsilon_{out} = 0.2$

Fig. 3. System parameters are $R = 1, \gamma = 0.2, N = 1, N_0 = 1$.

transmission is fixed to R bits/s/Hz. The resulting optimization problem is formulated as,

$$\min_{\mathbf{P}} \bar{P} \tag{27}$$

$$s.t. \quad \begin{cases} \mathcal{C}_1 : \gamma_r \leq \gamma, & 0 \leq \gamma \leq 1 \\ \mathcal{C}_2 : \epsilon_N \leq \epsilon_{out} & 0 \leq \epsilon_{out} \leq 1 \\ \mathcal{C}_3 : R_i = R & \forall i \\ \mathcal{C}_4 : P_i \leq P_m, & \forall i, j \end{cases} \tag{28}$$

where $\mathcal{C}_3$ represents the fixed rate constraint and $\mathcal{C}_4$ is the peak power constraint. This implies that largest transmit power at the UE cannot exceed P_m in any state i, regardless of the rate. Note that peak power constraint does not explicitly appear in (16) because R_{max} depends on P_m via (5) and appears in $\mathcal{C}_4$. For the modified optimization problem, the objective is to compute power vector $\mathbf{P}$ for fixed rate transmission. The constraints related to packet reception remain the same.

Lemma 1. *For the optimal power allocation in the fixed rate transmission, it holds $P_i \leq P_{i+1}$ for all $i \in [0, N]$.*

Proof. It is straight forward to prove by contradiction. If $P_i > P_{i+1}$ and the UE is allowed to enter state $i+1$, an optimal decision is not to transmit in state i at all and wait for a transmission in state $i+1$ which requires less power. This is a birth death process where after every $N-1$ time slots, one transmission is made in state N with power P_N. This clearly is suboptimal solution, and makes solving problem for most of the realistic γ and N values infeasible. □

Corollary 1. *The peak power constraint $P_i \leq P_m, \forall i, j$, reduces to $P_N \leq P_m$.*

Proof. From Lemma 1, $P_i \leq P_{i+1}, \forall i$. This implies, P_N is the largest transmit power for any state. Constraining $P_N \leq P_m$ is therefore, enough to apply peak power constraint to full system. □

From Corollary 1, P_N is constrained by P_m. However, P_N is also constrained by the power resulting from system parameter ϵ_{out} via $\mathcal{C}_2$. This implies that the problem is only feasible if the solution satisfies both outage probabilities resulting from the peak power constraint and the outage constraint ϵ_{out}. Denoting power consumption from ϵ_{out} by P_{out}, the solution is feasible if

$$P_{out} \leq P_N \leq P_m. \tag{29}$$

This problem is less flexible as compared to the general optimization problem as no rate adaptation is required at the transmit side. It is worth noting that in spite of reduction in complexity of the problem, the closed form solution of the problem is still not possible due to the challenges explained in Subsection IV-A. In the next section, we discuss special case to get some insight into the problem for the fixed rate transmission case.

A. $N = 1$ Case for Fixed Rate Transmission

Let us analyze a special case with $N = 1$ for the fixed transmission case. We compute ϵ_0, ϵ_1 for a given γ for the the fixed transmission rate case, i.e., $R_0 = R_1 = R$ (as in Section IV-B). It is possible to compute power levels P_0 and P_1 for the fixed transmission case.

The power levels P_0 and P_1 are computed from (13) and yield,

$$P_0 = \frac{1 - 2^R}{\log(1 - \frac{\gamma(1-\epsilon_{out})}{1-\gamma})} \tag{30}$$

$$P_1 = \frac{1 - 2^R}{\log(1 - \epsilon_{out})} \tag{31}$$

The resulting average power $\bar{P}$ in closed form is given by using (14),

$$\begin{aligned} \bar{P} &= \frac{1 - 2^R}{\log\left(1 - \frac{\gamma(1-\epsilon_{out})}{1-\gamma}\right)}\pi_0 + \frac{1 - 2^R}{\log(1 - \epsilon_{out})}\pi_1 \\ &= (1 - 2^R)\left(\frac{1}{\log\left(1 - \frac{\gamma(1-\epsilon_{out})}{1-\gamma}\right)}\pi_0 + \frac{1}{\log(1 - \epsilon_{out})}\pi_1\right) \end{aligned} \tag{32}$$

We can classify two distinct regions for the analysis of $\bar{P}$.

- $\epsilon_{out} \leq \gamma$: For this case, $\epsilon_N = \epsilon_{out}$ and the closed form expressions above hold. When $\epsilon_{out} \leq \gamma$, it implies that outage ϵ_N in state N should be less than the average outage probability γ. The optimal decision in the sense of power efficiency in this case is to transmit with power P_N that results in maximum permitted outage ϵ_{out}. This region is termed as bursty packet loss dominant region.
- $\epsilon_{out} > \gamma$: For this case, $\epsilon_N = \gamma$. Though, it is permitted to transmit with power P_N that results in $\epsilon_N < \gamma$, but this is not optimal in the sense of minimizing $\bar{P}$. If $\epsilon_{out} > \gamma$, the optimal decision is to transmit with power $P_i, \forall i$ that results in $\epsilon_i = \gamma$, i.e. the power allocation policy is independent of successive packet loss and only determined by average packet loss parameter γ.

From the above characterization, it is clear that $\epsilon_{out} \leq \gamma$ is the most critical region where average power consumption is determined by both the constraints on average packet loss and burst packet loss. We numerically verify in Section VII that the power levels computed in closed form for the boundary condition $\epsilon_N = \epsilon_{out}$ are not optimal for $\epsilon_{out} > \gamma$.

As with variable rate problem, the expressions for the power levels cannot be obtained in closed form for $N > 1$ when $\epsilon_{out} \leq \gamma$ in spite of fixing rate R for each transmission. The variables $\epsilon_0, \epsilon_1 \ldots \epsilon_N$ are unknown and it is not possible to compute a unique set of $\epsilon_i, \forall i$ in closed form that satisfies $\mathcal{C}_1$ in (28). The optimization problem in (27) is a combinatorial problem as it is hard to compute a unique solution in terms of $\epsilon_i, \forall i$ due to sum of product term in computation of γ_r. It is therefore, difficult to compute $\mathbf{P}$ that minimizes $\bar{P}$ using convex optimization techniques.

B. Characterization of Critical Regions

In the optimization problem in (27), average packet drop rate γ, successive packet loss constraint N and outage probability ϵ_{out} affect the average power consumption. It is worth noting that parameter γ is the only parameter that controls the 'quantity' of data loss. The parameters N and ϵ_{out} determine the qualitative effect for the average packet loss rate γ, i.e., for a fixed γ, different values of N and ϵ_{out} result in different QoE for the end user. If we relax N and ϵ_{out} constraints, we can save more energy at the expense of degradation in QoE without actually dropping more packets.

It is trivial that an increase in the acceptable average packet loss rate γ results in a monotonic decrease of average power consumption. However, it is not straightforward to understand the effect of parameters N and ϵ_{out} on the average power. In [26], it has been characterized that there exists a maximum N for a fixed γ that results in maximum energy efficiency for the system. Increasing N further, does not result in higher energy efficiency. We further characterize the energy efficiency as a function of qualitative parameters N, ϵ_{out} by the following lemma:

Lemma 2. *For a fixed γ and N, there exists a maximum $\epsilon_{out} = \gamma$ that results in minimum average power consumption. Increasing $\epsilon_{out} > \gamma$ does not help to reduce average power consumption.*

Lemma 2 can be proved following the proof of Lemma 1 in [26].

We numerically quantify the effect of these parameters on the average system energy consumption in Section VII. For a fixed γ, increasing N and/or ϵ_{out} helps in saving energy in the beginning. This implies that the system is in a region where avoiding successive packet loss has significant effect on average power consumption. An increase in N and/or ϵ_{out} helps system to drop a fraction γ of the packets with more degrees of freedom. When we increase N further, the system enters the region where the gap of N packet drops between two successful packet receptions almost never happens for a given γ. At this point, it does not matter if the system is allowed to drop more than N successive packets are not. Note that increasing ϵ_{out} has similar effect as increasing N; both permitting packets to be dropped successively within some margins. For small N, increasing ϵ_{out} has significant effect on average energy consumption as compared to large N. We provide numerical evidence of this characterization in Section VII.

VI. Stochastic Optimization

The combinatorial optimization problems in (15) and (27), which are not solvable with regular optimization techniques, can approximately be solved using stochastic optimization methods. There are a few heuristic techniques in literature to solve such problems like genetic algorithm, Q-learning, neural networks, etc. All of these techniques rely on randomized inputs to compute a solution at reduced computational complexity as compared to exhaustive search. Simulated Annealing (SA) is another similar stochastic optimization algorithm with the distinct feature that it helps avoid the solution to get stuck in local minima by introducing a probabilistic process called 'muting' as explained later in this section. The algorithm was originally introduced in statistical mechanics, and has been applied successfully to networking problems [26], [27]. Based on its ability to compute global minima with high probability, we use SA algorithm to solve optimization problem in (15) and (27).

In SA algorithm, a random configuration in terms of transition probability matrix $\mathbf{A}$ is generated in each iteration. Average power $\bar{P}$ is evaluated only if constraints in (16) are met. If the evaluated $\bar{P}$ is less than the previously computed best solution, the candidate set of outage probabilities ϵ_i, $\forall i$ are selected as the best available solution. However, the candidate set ϵ_i, $\forall i$ can be treated as the best solution with a certain temperature dependent probability even if the new solution is worse than the best known solution. This step is called *muting* and helps the system to avoid local minima. The muting occurs frequently at the start of the process as the selected temperature is very high and decrease as temperature is decreased gradually, where temperature denotes a numerical value that controls the muting process.

In literature, various cooling temperature schedules have been employed according to the problem requirements, such as Boltzmann annealing, fast annealing and adaptive cooling. The cooling schedule determines the convergence rate of the

Algorithm 1: Optimization by SA Algorithm for the General Case

Input: $(\mathbf{A}, T_m, \gamma, \epsilon_{out}, P_m)$;
T_m = lower bound on temperature;
$P_{a,0}$= Compute $\bar{P}$ as a function of initial $\mathbf{A}$;
$\bar{P}^* = P_{a,0}$; $\mathbf{A}^* = \mathbf{A}$;
$T_b = T_0$;
while $T_b \geq T_m$ **do**
  $T_b = \frac{T_0}{c_{\text{sa}} \cdot b+1}$;
  for *i=0* **to** *n* **do**
    Generate a random $\hat{\mathbf{A}}$;
    Compute γ_r for $\hat{\mathbf{A}}$;
    Evaluate $\mathcal{C}_2$ and $\mathcal{C}_3$;
    if $\mathcal{C}_2$ *and* $\mathcal{C}_3$ *satisfied* **then**
      for *j=0* **to** *r* **do**
        Generate $\mathbf{R}$ and evaluate $\mathcal{C}_1$ and $\mathcal{C}_4$;
        if $\mathcal{C}_1$ *and* $\mathcal{C}_4$ *are satisfied* **then**
          Solution feasible;
          Compute power vector $\hat{\mathbf{P}}$ as a function of $\hat{\mathbf{A}}$ using (13);
          Compute average power $\hat{P}_a$ in (14);
          s = A random number in range $[0, 1]$;
          if $s < \exp\left(\frac{-(\hat{\bar{P}}-\bar{P}_a)}{T}\right)$ **then**
            $\bar{P}_a = \hat{P}_a$;
            if *(*$\hat{\bar{P}} \leq \bar{P}^*$*)* **then**
              $\bar{P}^* = \hat{\bar{P}}$;
            end if
          end if
        end if
        else
          Solution Infeasible;
        end if
      end for
    end if
    else
      Solution Infeasible;
    end if
  end for
end while
Output: $(\bar{P}^*, \mathbf{A}^*)$;

Algorithm 2: Optimization by SA Algorithm For Fixed Rate Transmission

$T_b = \frac{T_0}{c_{\text{sa}} \cdot b+1}$;
for *i=0* **to** *n* **do**
  Generate a random $\hat{\mathbf{A}}$ and compute P_N;
  if $(\max(\mathbf{P}) \leq P_m)\mathbf{AND}(P_N \geq P_{out})$ **then**
    Solution feasible;
    Compute γ_r for $\hat{\mathbf{A}}$;
    if $\gamma_r < \gamma$ **then**
      Compute power vector $\hat{\boldsymbol{P}}$ as a function of $\hat{\mathbf{A}}$ using (13);
      Compute average power $\hat{P}_a$ in (14);
      s = A random number in range $[0, 1]$;
      if $s < \exp\left(\frac{-(\hat{\bar{P}}-\bar{P}_a)}{T}\right)$ **then**
        $\bar{P}_a = \hat{P}_a$;
        if *(*$\hat{\bar{P}} \leq \bar{P}^*$*)* **then**
          $\bar{P}^* = \hat{\bar{P}}$;
        end if
      end if
    end if
  end if
  else
    Solution not feasible;
  end if
end for
Output: $(\bar{P}^*, \boldsymbol{A^*})$;

solution. If temperature cools down at a fast rate, the optimal solution can be missed. On the other hand, if it cools down too slowly, optimization requires large amount of time. In this work, we employ fast annealing (FA) [29] because it provides us reasonably good results. In FA, it is sufficient to decrease the temperature linearly in each step b such that,

$$T_b = \frac{T_0}{c_{\text{sa}} \cdot b + 1} \tag{33}$$

where T_0 is a suitable starting temperature and c_{sa} is a constant, which depends on the requirements of the problem. After a fixed number of temperature iterations, when muting fully stops, the best solution is accepted as an approximation to the optimal solution. Note that the solution provided after a fixed number of temperature iterations is used to keep the computational complexity manageable. To show the convergence behaviour of the solution provided by SA, we compare the SA approximated results with the analytical results for the $N = 1$ case.[3]

To apply the SA algorithm and solve the optimization problem in (15), we use the following 2-step process:

1) First generate a random set of $\epsilon_i \forall i$ and evaluate if $\mathcal{C}_2$ and $\mathcal{C}_3$ are met. The candidate solutions which do not meet $\mathcal{C}_2$ and $\mathcal{C}_3$, are not feasible solutions and they are dropped.
2) For the candidate solutions that meet $\mathcal{C}_2$ and $\mathcal{C}_3$, we solve the following programming problem:
   Find $\mathbf{R}$ that meets $\mathcal{C}_1$ and $\mathcal{C}_4$. That constitutes a feasible solution. For all feasible solutions, we evaluate $\bar{P}$ and choose the $\mathbf{P}, \mathbf{R}$ vectors that minimize $\bar{P}$. Note that we need to generate randomized vector $\mathbf{R}$ between values R_{min} and R_{max} to generate a candidate solution.

Pseudocode for the optimization of problem using SA is presented in Algorithm 1. The complexity of the solution depends on parameter N. For large N, size of transition probability matrix grows and it becomes computationally expensive to calculate the optimal matrix.

The solution for the fixed rate transmission optimization problem in (27) is computed using the SA algorithm in a

[3]A lot of literature is available on providing more accurate measures of convergence for SA algorithm [30], but going in rigorous mathematical details on the convergence of the approximated solution is out of scope of this paper.

similar way, but it is relatively less complex. As rate is fixed for every transmission, a one step feasible solution comprising $\epsilon_i \forall i$ is selected from the randomly generated candidate solutions that meets $\mathcal{C}_1 - \mathcal{C}_4$. For every feasible solution, the objective function is evaluated and the solution that minimizes $\bar{P}$ is selected. To minimize the repetition, pseudocode for one temperature iteration is presented in Algorithm 2.

VII. Numerical Results

We perform a numerical evaluation of the proposed scheduling scheme in this section. We consider a Rayleigh fading channel with mean 1 for the point to point link. Peak power is set to 20 dBW for all numerical examples while the noise variance N_0 equals one. R_{min} is set to a small value of 0.001 bits/s/Hz to allow system to choose almost any $R_0 \leq R_m, R_1 \leq R_m$ combination that minimizes $\bar{P}$. The cooling schedule from (33) is applied in SA algorithm where number of iterations per temperature value is fixed.

We study the effect of packet loss parameters on the average power consumption for the special case $N = 1$ in Fig. 4, where the results are evaluated using both closed form expressions derived in Sections IV-B and V-A; and the SA framework developed in Section VI. Average transmit power is plotted for $N = 1$ and $\gamma = 0.2$ in Fig. 4. Note that $\epsilon_{out} = \epsilon_N$ in the closed form expression. For the fixed power case, the average power consumption is a convex function in ϵ_{out} for a fixed γ and N, and a unique optimal ϵ_{out} can be identified. Let us call it ϵ^*_{out}. If system parameter $\epsilon_{out} \leq \epsilon^*_{out}$, it results in high average power. However, if $\epsilon_{out} > \epsilon^*_{out}$, the system has more flexibility and it is optimal to set $\epsilon_N = \epsilon^*_{out}$ instead to save power. The optimized results with the SA method match closely with the analytically computed results for $\epsilon_{out} \leq \epsilon^*_{out}$ which validate the accuracy of the solution provided by the SA algorithm. For $\epsilon_{out} > \epsilon^*_{out}$, SA method provides the optimal solution in contrast to the suboptimal solution where $\epsilon_N = \epsilon_{out}$ is enforced. For the case of variable rates power allocation, the analytical results and the solutions from the SA method match closely. However, the difference is more in the case of fixed rate. Note that the results for the analytical case cannot be fully computed in closed form and the optimal rate is computed using a scalar search for the optimal R_{min} at fixed ϵ_0 values which introduces numerical inaccuracy for the 'semi-analytical' solution for $N = 1$.

Fig. 5 compares the average power for both schemes for $N = 1$ case. We use the SA method to compute the results for all the examples in rest of this section.[4] The results clearly show that variable rate scheme performs better than the fixed scheme at small ϵ_{out}. This is attributed to more flexibility in choosing rates for different states. Another important observation can be noted for $R = 3$ bits/s/Hz. Fixed rate scheme cannot provide any quality of service for $\epsilon_{out} < 0.07$ and $\gamma = 0.2$. The flexibility of variable rate scheme allows to achieve almost identical average power for all ϵ_{out} including the smaller ones. This leads us to the conclusion that rate

[4]It should be noticed that the results obtained from the SA algorithm always show some irregular points due to inherent randomness of the heuristic algorithm in computing the solution.

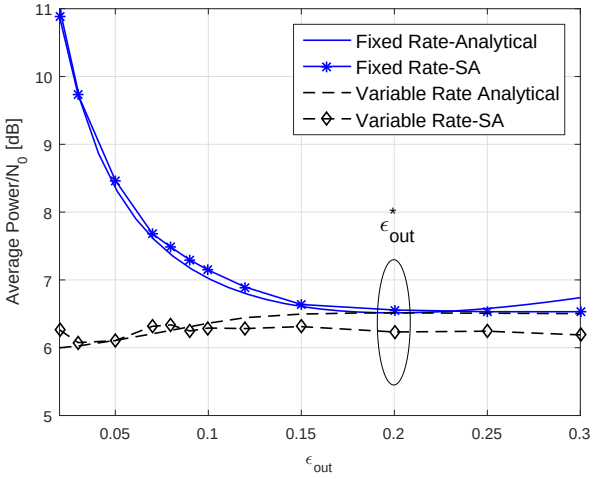


Fig. 4. Average power for both variable and fixed rate cases for the special case $N = 1$. Analytically computed solutions are compared with the solutions produced using the SA algorithm. The target rate R is set to 1 bits/s/Hz and γ is fixed to 0.2.

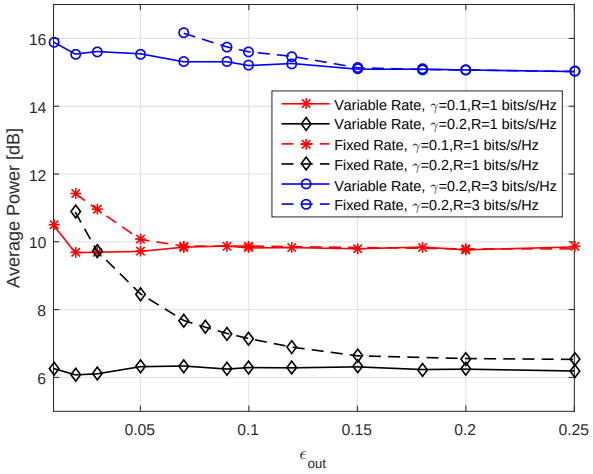
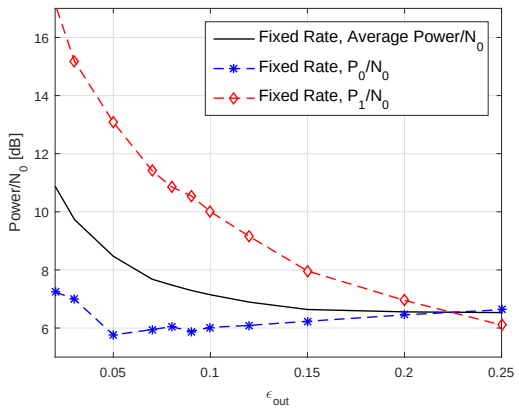


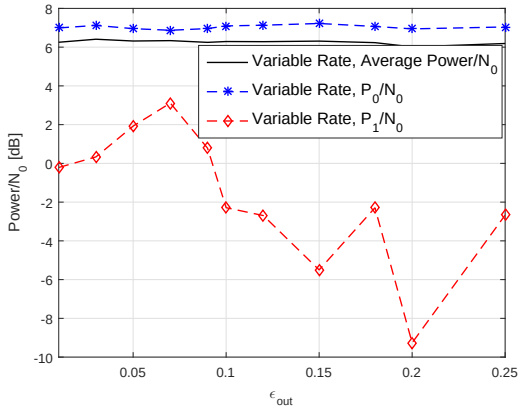
Fig. 5. Average power consumption for variable and fixed rate schemes for $N = 1$ and the SA algorithm.

adaptation is useful at small ϵ_{out}, while fixed rate transmission becomes almost as efficient at high ϵ_{out} at reduced complexity. This is clearly more evident at higher rates ($R = 3$ bits/s/Hz).

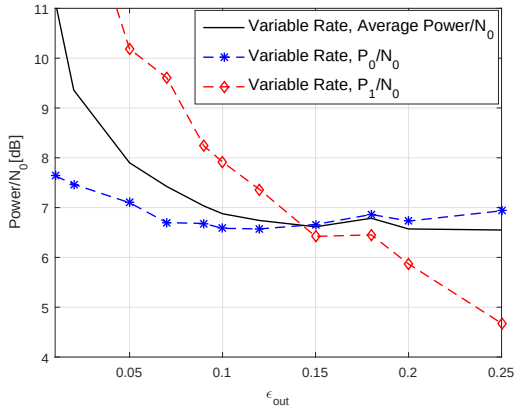
In Fig. 6, we investigate the power allocation for individual states for both schemes. Interestingly, power allocation is opposite in both schemes at small ϵ_{out}. For the fixed rate scheme, the optimal power allocation requires to transmit with small power in state '0' and large power in state '1'. As $R = 1$ in both states, it implies $\epsilon_0 \leq \epsilon_1 \leq \epsilon_{out}$ for the optimal power allocation. For the variable rate case when R_{min} is small (0.001 bits/s/Hz) as in Fig. 6(b) and Fig. 6(d), the optimal power allocation requires $P_0 \geq P_1$. The optimal policy is to transmit with higher rate and power $P_0 \leq P_m$ in state '0' and with substantially small power and rate in state 1. This results in optimal average power when rate adaptation is allowed. In Fig. 6(c), when R_{min} is increased to 0.5 bits/s/Hz for $R = 1$, the power allocation for variable rate scheme resembles



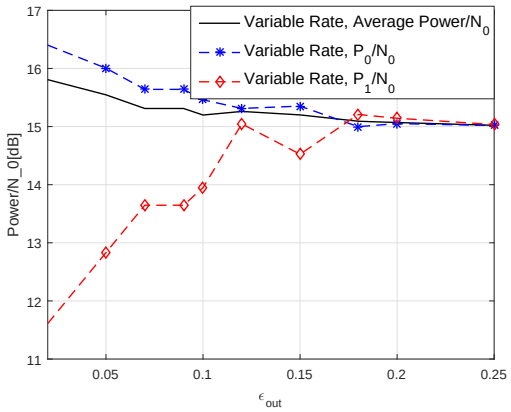
(a) Fixed Rate, R=1 bits/s/Hz



(b) Variable Rate, R=1 bits/s/Hz, $R_{min} = 0.001$ bits/s/Hz



(c) Variable Rate, R=1 bits/s/Hz, $R_{min} = 0.5$ bits/s/Hz



(d) Variable Rate, R=3 bits/s/Hz, $R_{min} = 0.001$ bits/s/Hz

Fig. 6. Power levels for the fixed and variable rate scheme for the individual states. The system parameters are $N = 1, \gamma = 0.2$.

more to fixed rate scheme due to decrease in flexibility in rate allocation. The fixed transmission policy suffers from the constraint $R_0 = R$ and any dropped packets have to be compensated in state 1 with more power. When $\epsilon_{out} \to \epsilon^*_{out}$ for variable rate power allocation, P_1 is relatively small as compared to P_0 if R is small (Fig. 6(b)) and peak power constraint is large. If R is large as in Fig. 6(d), P_1 is increased to meet the rate constraints. However, the intention of the system is to exploit state '0' by adapting power and rate as there is no constraint on ϵ_0.

Fig. 7 compares the average power consumption for the fixed rate scheme for the case $N = 1, 2, 3$ and $\gamma = 0.2$. The power levels are optimized using SA algorithm. It is evident that the resulting average power converges for all N to the same minimum value at ϵ^*_{out}. When $\epsilon_{out} \leq \epsilon^*_{out}$, an increase in N for a fixed γ helps to reduce average power consumption in general (specially at small ϵ_{out}). More flexibility in packet dropping parameters provides more degrees of freedom and results in energy savings. When $\epsilon_{out} > \epsilon^*_{out}$, the effect of large N vanishes and power saving depends solely on average packet dropping parameter.

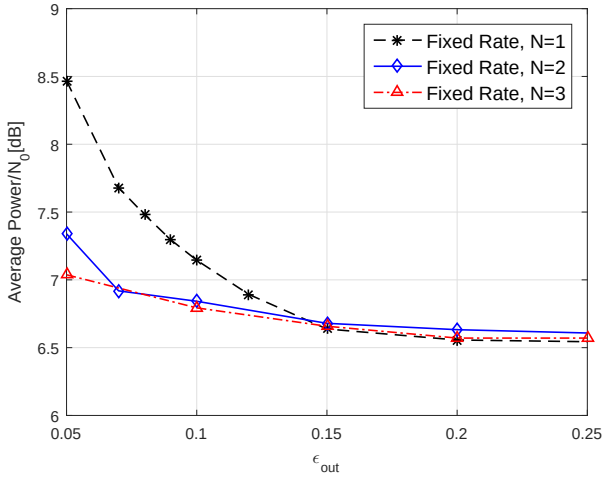


Fig. 7. Average power as a function of packet loss parameters for different N for the fixed rate scheme. γ is fixed to 0.2 and $R = 1$ bits/s/Hz.

VIII. Conclusion

We consider energy efficient scheduling and power allocation for the loss tolerant IoT applications. Data loss is characterized as a function of average and successive packet loss, and the probability that the successive packet loss constraint is not met. These parameters jointly define the QoE and context for an IoT application. In contrast to average

packet loss parameter, other loss parameters depend on the packet loss order without actually changing the number of lost packets. By considering bursty packet loss a form of contextual information, we provide another degree of freedom in the scheduling algorithm which can be exploited to reduce energy consumption. Without CSIT, we formulate the average power optimization problem as a function of data loss parameters. First, the generalized power optimization problem is discussed where the transmitted packets are adapted in size such that an average rate and minimum packet size guarantee is provided. Then, we relax the problem to the case where the packet size is fixed for all transmissions. Both of the optimization problems are combinatorial in nature and require a stochastic optimization technique to solve them. For both problems, we compute analytical expressions of average power as a function of system parameters for the special case $N = 1$ and compare it with the solution obtained from the proposed simulated annealing algorithm. Both of the analytical results match quite well and validate the solution provided by the heuristic simulated annealing algorithm. We numerically study performance of both schemes and show dependency of power consumption on the parameters that depend on the order of packet drop in addition to packet drop rate.

REFERENCES

[1] M. M. Butt, E. A. Jorswieck, and N. Marchetti, "Energy efficient scheduling for loss tolerant IoT applications with uninformed transmitter," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, Paris, France, May 2017.

[2] C. She, C. Yang, and T. Q. S. Quek, "Radio resource management for ultra-reliable and low-latency communications," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 72–78, 2017.

[3] K. F. Trillingsgaard and P. Popovski, "Downlink transmission of short packets: Framing and control information revisited," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2048–2061, May 2017.

[4] C. She, C. Yang, and T. Q. S. Quek, "Cross-layer optimization for ultra-reliable and low-latency radio access networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 127–141, Jan 2018.

[5] M. C. Gursoy, "Throughput analysis of buffer-constrained wireless systems in the finite blocklength regime," *EURASIP Journal on Wireless Communications and Networking*, vol. 2013, no. 1, p. 290, Dec 2013.

[6] S. Xu, T. H. Chang, S. C. Lin, C. Shen, and G. Zhu, "Energy-efficient packet scheduling with finite blocklength codes: Convexity analysis and efficient algorithms," *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5527–5540, Aug 2016.

[7] P. Pahalawatta, R. Berry, T. Pappas, and A. Katsaggelos, "Content-aware resource allocation and packet scheduling for video transmission over wireless networks," *IEEE Journal on selected areas in communication (JSAC)*, vol. 25, no. 4, pp. 749–759, May 2007.

[8] L. Li, S. Li, and S. Zhao, "QoS-aware scheduling of services-oriented internet of things," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1497–1505, May 2014.

[9] "G.1080: Quality of experience requirements for IPTV services," International Telecommunication Union, Standard, Dec. 2008.

[10] J. Choi and J. Ha, "On the energy efficiency of AMC and HARQ-IR with QoS constraints," *IEEE Trans. Vehicular Technology*, vol. 62, pp. 3261–3270, 2013.

[11] C. Li, L. Toni, J. Zou, H. Xiong, and P. Frossard, "Delay-power-rate-distortion optimization of video representations for dynamic adaptive streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2017.

[12] A. M. Sheikh, A. Fiandrotti, and E. Magli, "Distributed scheduling for low-delay and loss-resilient media streaming with network coding," *IEEE Transactions on Multimedia*, vol. 16, no. 8, pp. 2294–2306, Dec 2014.

[13] J. Wu, B. Cheng, C. Yuen, N. M. Cheung, and J. Chen, "Trading delay for distortion in one-way video communication over the internet," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 4, pp. 711–723, April 2016.

[14] J. Wu, C. Yuen, B. Cheng, M. Wang, and J. Chen, "Energy-minimized multipath video transport to mobile devices in heterogeneous wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, pp. 1160–1178, May 2016.

[15] J. Wu, B. Cheng, M. Wang, and J. Chen, "Energy-efficient bandwidth aggregation for delay-constrained video over heterogeneous wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 1, pp. 30–49, Jan 2017.

[16] M. M. Nasralla, C. T. E. R. Hewage, and M. G. Martini, "Subjective and objective evaluation and packet loss modeling for 3D video transmission over LTE networks," in *International Conference on Telecommunications and Multimedia*, Heraklion, Greece, July 2014.

[17] Z. Zou and M. Johansson, "Deadline-constrained maximum reliability packet forwarding with limited channel state information," in *IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, pp. 1721–1726.

[18] A. Dua, C. W. Chan, N. Bambos, and J. Apostolopoulos, "Channel, deadline, and distortion (CD2) aware scheduling for video streams over wireless," *IEEE Trans. Wireless. Communications*, vol. 9, no. 3, pp. 1001–1011, Mar. 2010.

[19] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 125–144, Jan 2004.

[20] J. Lee and N. Jindal, "Energy-efficient scheduling of delay constrained traffic over fading channels," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1866–1875, April 2009.

[21] ——, "Asymptotically optimal policies for hard-deadline scheduling over fading channels," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2482–2500, April 2013.

[22] R. A. Berry, "Optimal power-delay tradeoffs in fading channels-small delay asymptotics," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3939–3952, June 2013.

[23] M. J. Neely, "Intelligent packet dropping for optimal energy-delay tradeoffs in wireless downlinks," *IEEE Trans. on Automatic Control*, vol. 54, no. 3, pp. 565–579, March 2009.

[24] L. Sequeira, J. Fernandez-Navajas, L. Casadesus, J. Saldana, I. Quintana, and J. Ruiz-Mas, "The influence of the buffer size in packet loss for competing multimedia and bursty traffic," in *International Symposium on Performance Evaluation of Computer and Telecommunication Systems*, Toronto, Canada, July 2013.

[25] F. Liu, T. H. Luan, X. S. Shen, and C. Lin, "Dimensioning the packet loss burstiness over wireless channels: a novel metric, its analysis and application," *Wireless Communications and Mobile Computing*, 2012.

[26] M. M. Butt and E. A. Jorswieck, "Maximizing system energy efficiency by exploiting multiuser diversity and loss tolerance of the applications," *IEEE Trans. Wireless Communications*, vol. 12, no. 9, pp. 4392–4401, 2013.

[27] M. M. Butt, E. A. Jorswieck, and A. Mohamed, "Energy and bursty packet loss tradeoff over fading channels: A system-level model," *IEEE Systems Journal*, vol. 12, no. 1, pp. 527–538, March 2018.

[28] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 630–643, July 2003.

[29] H. Szu and R. Hartley, "Fast simulated annealing," *Physics Letters A*, vol. 122, no. 3-4, pp. 157–162, June 1987.

[30] V. Granville, M. Krivanek, and J. P. Rasson, "Simulated annealing: a proof of convergence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 652–656, Jun 1994.

M. Majid Butt (S'07 – M'10 – SM'15) received the MSc degree in Digital Communications from Christian Albrechts University, Kiel, Germany, in 2005, and the PhD degree in Telecommunications from the Norwegian University of Science and Technology, Trondheim, Norway, in 2011. He is an Assistant Professor at University of Glasgow as well as an adjunct Assistant Professor at Trinity College Dublin, Ireland. Before that, he has held senior researcher positions at Trinity College Dublin, Ireland and Qatar University. He is recipient of Marie Curie Alain Bensoussan postdoctoral fellowship from European Research Consortium for Informatics and Mathematics (ERCIM). He held ERCIM postdoc fellow positions at Fraunhofer Heinrich Hertz Institute, Germany, and University of Luxembourg. Dr. Majid's major areas of research interest include communication techniques for wireless networks with focus on radio resource allocation, scheduling algorithms, energy efficiency and cross layer design. He has authored more than 50 peer reviewed conference and journal publications in these areas. He has served as TPC chair for various communication workshops in conjunction with IEEE WCNC, ICUWB, CROWNCOM, IEEE Greencom and IEEE Globecom. He is a senior member of IEEE and serves as an associate editor for IEEE Access journal and IEEE Communication Magazine since 2016.

Eduard A. Jorswieck (S'01 – M'05 – SM'08) was born in 1975 in Berlin, Germany. He received his Ph.D. degree in electrical engineering and computer science from the Technische Universitt Berlin, Germany, in 2004. Since February 2008, he has been the head of the Chair of Communications Theory and Full Professor at Dresden University of Technology (TUD), Germany. Eduards main research interests are in the area of signal processing for communications and networks, applied information theory, and communications theory. He has published more than 100 journal papers, 11 book chapters, 3 monographs, and some 275 conference papers on these topics. Dr. Jorswieck is senior member of IEEE. Since 2017 he acts as Editor-in-Chief for the Springer EURASIP Journal on Wireless Communications and Networking. Since 2015, he is member of the IEEE SAM TC. Since 2016, he serves as Associate Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. He has served on the editorial boards of IEEE SIGNAL PROCESSING LETTERS, IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. In 2006, he received the IEEE Signal Processing Society Best Paper Award.

Nicola Marchetti is currently Assistant Professor in Wireless Communications at Trinity College Dublin, Ireland. He performs his research under the Trinity Information and Complexity Labs (TRICKLE) and the Irish Research Centre for Future Networks and Communications (CONNECT). He received the PhD in Wireless Communications from Aalborg University, Denmark in 2007, and the M.Sc. in Electronic Engineering from University of Ferrara, Italy in 2003. He also holds an M.Sc. in Mathematics which he received from Aalborg University in 2010. His collaborations include research projects in cooperation with Nokia Bell Labs and US Air Force Office of Scientific Research, among others. His research interests include Adaptive and Self-Organizing Networks, Complex Systems Science for Communication Networks, PHY Layer, Radio Resource Management. He has authored 110 journals and conference papers, 2 books and 7 book chapters, holds 2 patents, and received 4 best paper awards.