Energy-Reliability Aware Link Optimization for Battery-Powered IoT Devices with Non-Ideal Power Amplifiers

Aamir Mahmood, Member, IEEE, M. M. Aftab Hossain, Cicek Cavdar, and Mikael Gidlund, Senior Member, IEEE

Abstract—In this paper, we study cross-layer optimization of low-power wireless links for reliability-aware applications while considering both the constraints and the non-ideal characteristics of the hardware in IoT devices. Specifically, we define an energy consumption (EC) model that captures the energy cost—of transceiver circuitry, power amplifier, packet error statistics, packet overhead, etc.—in delivering a useful data bit. We derive the EC models for an ideal and two realistic non-linear power amplifier models. To incorporate packet error statistics, we develop a simple, in the form of elementary functions, and accurate closed-form packet error rate (PER) approximation in Rayleigh block-fading. Using the EC models, we derive energy-optimal yet reliability and hardware compliant conditions for limiting unconstrained optimal signal-to-noise ratio (SNR), and payload size. Together with these conditions, we develop a semi-analytic algorithm for resource-constrained IoT devices to jointly optimize parameters on physical (modulation size, SNR) and MAC (payload size and number of retransmissions) layers in relation to link distance. Our results show that despite reliability constraints, the common notion—higher M-ary modulations are energy optimal for short-range communication—prevails, and can provide up to 180% lifetime extension as compared to often employed OQPSK modulation in IoT devices. However, the reliability constraints reduce both their range and the energy efficiency, while non-ideal traditional PA reduces the range further by 50%, and diminishes the energy gains unless a better PA is employed.

Index Terms—Energy-efficiency, reliability, short-range communication, cross-layer design, IoT, non-linear power amplifiers.

I. INTRODUCTION

In rapidly growing Internet-of-things (IoT), a broad-range of applications—from time-critical services to connectivity of massive autonomous devices [2]—rely on energy- and hardware-constrained radio devices for local-area machine-to-machine type communication (MTC). In MTC, the devices exchange short packets at low duty cycle [3] with a controller or directly between two devices [4] with guaranteed reliability and latency. When the devices are battery-powered, the main design concern is to minimize energy consumption in wireless communication module in order to prolong the lifetime as much as possible. However, energy- and reliability-aware communication are two conflicting demands in MTC that often necessitate design trade-offs.

Reliable transfer of information has a direct impact on the needed energy consumption [5]. This is because the reliability depends on the bit error or packet error statistics of the wireless channel, which in turn are influenced by the choice of link parameters such as transmit power, modulation scheme, packet length, etc. If the packet error probability has to be reduced so as to successfully transfer a packet with a limited number of retransmissions, the link parameters must be optimized jointly while restraining the energy consumption. Note that transmitting the same packet to provide reliability not only translates into a higher transmission cost but also causes delay violations. As a result, it is becoming imperative to look not only into physical (PHY) and medium access (MAC) layers separately, but across layers with awareness to extreme energy and reliability constraints posed by MTC-IoT devices.

In any energy-aware design, the impact of constraints and the characteristics of components in the hardware layer cannot be ignored. In an RF device, a major chunk of energy is consumed at the transmitter for RF signal generation. This is because the signal generation at the desired output power requires amplification via a power amplifier (PA). However, due to the limited efficiency and linearity of traditional PAs, typically 50%–80% of energy is consumed by the PAs [6], [7]. Therefore, the low-power low cost IoT devices require a PA-centric joint design of PHY and MAC parameters.

In this paper, we study the selection of optimal—energy consumption minimizing and reliability aware—link (MAC and PHY) parameters in block-fading channels, however, under the often neglected hardware constraints and non-ideal PA characteristics of the IoT devices. To such end, we analyze the inter-play among optimal parameters under a widely used hypothetical PA model, and two realistic non-linear models: prevailing traditional PA (TPA), and envelop-tracking PA (ETPA) [7], [8], which to our best knowledge is the first such study.

A. Related Works and Contributions

The issue of cross-layer optimization of PHY and MAC parameters to minimize energy-consumption in low-power wireless sensor networks (WSNs) has been researched by many authors. How to select the modulation order, and transmission power to attain energy-efficient communication are studied in AWGN channel [9]–[11], and in fading channels [12] [13]. In general, these studies suggest using higher order
modulations at smaller distances, as opposed to the common notion followed in WSNs where low-order modulations are preferred for their low SNR requirement. For instance, highly-proliferated low-power transceivers CC1100 and CC2420 in WSNs employ BPSK and OQPSK modulations, respectively. In fading channels, it is shown in \cite{12, 13} that there exist an optimal SNR and packet length for each modulation scheme at which the required energy for successful transfer of an information bit is minimized. In \cite{12}, the optimal SNR is conditioned on the maximum transmit power, but this constraint is ignored in \cite{13}.

In all these studies, by assuming the system is delay-tolerant, no restriction on the number of retransmissions is imposed. As a result, the optimal link parameters are not bound to satisfy the time-critical MTC \cite{14}. In contrast, we limit the number of retransmissions of the same data packet based on the reliability and latency constraint. Moreover, constant PA efficiency, irrespective of the utilized transmit power, is assumed, which is far from the realistic PA characteristics. However, higher order QAM modulations, which carry information in both phase and amplitude, the realistic PAs suffer from poor power efficiency due to their non-constant envelope signals. Therefore, it is yet to be analyzed how these modulations behave at short distances under realistic PAs.

When optimizing the parameters of a wireless link, a valuable cross-layer metric capturing the cost-benefit trade-off, is packet error rate (PER) \cite{15}. However, a generic and exact analytical expression of PER in fading channels is not found in the literature. A generic upper bound on average PER in Rayleigh block-fading channels is \[ 1 - \exp(\omega_0/\text{SNR}) \], where \( \omega_0 \) — the waterfall threshold — is defined in terms of an integral of the PER function in the AWGN channel. Its approximation based on log-domain linear approximation of \( \omega_0 \) is developed for uncoded schemes in \cite{16}, and for (un)coded schemes in \cite{13}. However, the approximation in \cite{16} is tight in asymptotic regime while in \cite{13} the approximation parameters are simulation-aided. As a result, the integral is evaluated numerically, though not computationally intensive, but does not offer insights regarding what parameters determine the system performance.

A summary of our main contributions follows.

- We develop an energy consumption (EC) model giving energy cost per reliably transmitted information bit, and derive it for a hypothetical and two real-life non-linear power amplifier (PA) models.
- We show that the waterfall threshold of the PER upper bound in Rayleigh block-fading is tightly approximated by the expected value of an asymptotic PER distribution in AWGN, which leads to a PER approximation more accurate than \cite{13} and \cite{16} while maintaining explicit relation with the PHY and MAC layer parameters unlike \cite{13}.
- We find the closed-form expressions of optimal SNR and payload size, and find the necessary conditions for energy-reliability aware selection of SNR and payload size for a power-limited system.
- We develop a joint optimization algorithm to find optimal link parameters with respect to distance under the prescribed reliability constraints. We show that the right selection of parameters can increase the link’s lifetime up to 180 % compared to OQPSK modulation.
- Our analysis with non-linear PA models gives several new insights: under a realistic PA, the energy consumption to reliably operate a link increases; a traditional PA diminishes the gain in employing higher (PAPR) modulations at short distances; PAs with better linearity (e.g., envelop-tracking PA) can improve this situation, and must be studied for power-limited devices.

The rest of the paper is organized as follows. Section II presents the system model and energy-efficiency metrics. Section III develops the EC model for various PA models, and drives expression for the associated PER function. Section IV performs cross-layer optimization with reliability constraints and develops an algorithm for link’s parameters optimization. Section V gives the insightful results, and discusses the positive impact of the proposed optimization on the lifetime of the link. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODEL

A. Communication Model

We consider a master-slave communication model, typical for industrial automation applications \cite{17, 18}; a gateway (GW) acting as the master, and battery-operated wireless sensors as the slaves (see Fig. 1(a)). The sensors are placed at random locations within the communication range of the GW. The GW schedules a transmission by broadcasting a short control message, which indicates the sensor that shall transmit a packet exclusively in time and/or frequency resource. The control message also indicates the permissible maximum delay, \( D_{\text{max}} \), and the target packet error probability, \( P \), of the sanctioned transaction (see Fig. 1(c)). Note that the design of a scheduling policy that could satisfy the end-to-end latency including the delay in scheduling of a sensor is beyond the scope of this paper, although one may refer to \cite{4}.
B. Metrics, QoS Model and Objectives

In this work, we are interested in minimizing the energy consumption of the wireless sensor devices in transmitting the packet to the gateway. While optimizing the energy efficiency, we also want to ensure that the QoS requirements are fulfilled.

Energy Efficiency Metric. In general, to capture the energy consumption and reliability trade-off, the cost-benefit of the link is expressed as the ratio of the energy consumption to the corresponding data reliably delivered as

$$E = \left( \frac{1}{1-p} \right) E_0,$$

where the first term is the average number of transmissions, $\tau$, over an i.i.d channel realizations between transmissions, which is the mean of a geometric random variable with parameter $p$—the packet error rate (PER). The term $E_0$, to be defined precisely in Sec. III, accounts for the energy consumed by the radio circuitry to operate the link. The measure unit of (1) is [Joule/bit] as it represents the amount of energy required to transmit a given amount of data or, as stated otherwise, the energy cost per reliably transmitted information bit.

The measure of energy consumption in (1), with its relation to PER, depends on the physical parameters as SNR or transmit power, modulation order and symbol rate, and MAC layer parameters including packet size—information and overhead bits (see Fig. 1(b)—and the number of retransmissions.

Energy Efficiency Metric with Probabilistic QoS Model. Since delay $D_{\text{max}}$ is finite, the maximum number of retransmissions has to be bounded i.e., $\tau_r^{\text{max}} = \lfloor D_{\text{max}}/t_T \rfloor - 1$ where $t_T$ is the transmission time of the packet. Moreover, due to finite retransmissions, an error-free delivery cannot be ensured. Since a packet loses its value after $D_{\text{max}}$ in typical control applications, the packet is dropped after $\tau_r^{\text{max}}$ retransmissions. Therefore, the condition: the probability of packet error after $\tau_r^{\text{max}}$ retransmissions is not greater than $P_r$, defines a probabilistic QoS model

$$p\tau_r^{\text{max}} + 1 \leq P_r.$$  \hspace{1cm} (2)

Solving (2) for $p$ gives the reliability needed at the PHY layer under a limited number of transmissions at the MAC layer

$$p \leq P_r^{1/(\tau_r^{\text{max}} + 1)} : = P_{\text{req}}.$$  \hspace{1cm} (3)

If (3) is satisfied for each transmission attempt at the physical layer, the application QoS requirement in (1) with maximum $\tau_r^{\text{max}}$ retransmissions will be satisfied with the average number of transmissions per packet $\tilde{\tau}_{\text{trunc}} = 1 - p\tau_r^{\text{max}} + 1/(1 - p)$. In this case, the energy consumption follows from (1) as

$$E_{\text{trunc}} = \left( \frac{1 - p\tau_r^{\text{max}} + 1}{1 - p} \right) E_0.$$  \hspace{1cm} (4)

Objectives. From the above-defined probabilistic QoS model, our objective is to find the modulation scheme and its operational parameters: SNR or transmit power at the PHY layer, and packet size at the MAC layer such that the required PER in (3) is satisfied by implementing $\tau_r^{\text{max}}$ retransmissions at the MAC while energy efficiency in (1) for the sanctioned transaction is maximized.

III. ELEMENTS OF ENERGY CONSUMPTION MODEL

In this section, we define energy consumption model $E_0$, and derive it for various power amplifier models. Also, we find a PER expression that maintains an explicit connection with the modulation order and the associated parameters.

A. Average Energy Per Information Bit

The energy consumption of the signal path between the transmitter and receiver is comprised of baseband processing blocks and radio-frequency (RF) chain. The RF chain consists of a power amplifier (PA) and other electronic components such as analog-to-digital and digital-to-analog converters, low-noise amplifier, filters, mixers and frequency synthesizers. However, for an energy-constrained wireless system, the energy consumption of RF chain is the orders of magnitude higher than the consumption of baseband processing blocks. The power consumption of PA is considered to be proportional to the transmit power $P_t$ as $P_{\text{PA}} = \xi P_t/\eta(P_t)$, where $\xi$ is the modulation scheme dependent peak-to-average power ratio (PAPR) and $\eta(P_t)$ is the $P_t$-dependent drain efficiency of the PA (see Sec. III-B). If baseband power consumption is neglected and the power consumption of all the other components in RF chain excluding PA is denoted as $P_c$, a simple power consumption model is $P_{\text{on}} = \xi P_t/\eta(P_t) + P_c$. From [9], it leads to total energy consumption to transmit and receive a symbol $E_{\text{sym}}$

$$E_{\text{sym}} = \frac{\xi E_r}{\eta(P_t)} + \frac{P_c}{R_s},$$

where $E_r$ is the average transmission energy of a symbol and $R_s$ the symbol rate at PHY. For FSK, BPSK and QPSK modulations, PAPR $\xi = 1$, for OQPSK $\xi = 2.138$, and for a square MQAM modulations, it can be approximated as $\xi = 3(\sqrt{M} - 1/M + 1)$ [9].

Let $E_b = E_r/\log_2 M$ be the average received energy per uncoded bit where $E_r$ is the average received energy per symbol and $M$ is the constellation size. Then, the average SNR ($\bar{\gamma}$) at the receiver is

$$\bar{\gamma} = \frac{E_r}{N_0 \log_2 M}.$$  \hspace{1cm} (6)

Assuming a $\kappa$th-power path-loss model, the transmission energy at distance $d$ from (6) can be expressed as [9]

$$E_t \triangleq E_r G_d = \left( \frac{\bar{\gamma} N_0 \log_2 M}{G_d} \right) G_d,$$  \hspace{1cm} (7)

where $G_d \triangleq G_1 d^\kappa M_I$ is the path-loss gain, where $G_1$ is the gain factor at unit distance, which depends on the transmit and receive antenna gains and carrier frequency, and $M_I$ is the link margin.

In a packet based wireless system, the information bits are encapsulated into packets each carrying $n_p$ payload and $n_h$ overhead bits; thus the number of symbols in each packet are $n_s = (n_h + n_p)/\log_2 M$. Therefore, the average energy required to transmit and receive an information bit, from (3) and (7), is

$$E_0 = \left( \frac{n_p + n_h}{n_p \log_2 M} \right) E_{\text{sym}}.$$  \hspace{1cm} (8)
B. \( E_0 \) for Different Power Amplifier Models

As PA is the most power consuming component of a wireless link, its real-life characteristics must be considered for energy efficient communication. The power consumption of PA is high mainly due to its inefficiency in signal amplification, and non-linearity in signal amplification outside the limited dynamic range. As a result, if the PA input signal is higher than its linear region, the output signal is distorted; whereas, if PA input signal falls below from a saturation point—where the output power is maximum—the PA drain efficiency drops significantly. It is energy efficient to operate the PA at its saturation point [6], but due to the dynamic range of input signals (i.e., higher order modulations with high PAPR), it cannot operate at the saturation point and the PA must backoff at certain output power.

In earlier studies, a PA model with constant drain efficiency is assumed (e.g., see [9], [13]), which is far from reality. To capture the trade-off in energy efficiency and reliability, the link dynamics under realistic PA models must be studied. In the following, we compare a PA with constant efficiency (CPA) with two non-linear PAs, a) commonly used traditional PA (TPA) and b) envelop tracking PA (ETPA) that utilizes a linear PA along with a supply modulation circuitry in which the supply voltage tracks the input signal envelope.

**Constant PA.** In this case, the drain efficiency of a PA \( \eta \) at the output power \( P_t \) is assumed to be constant i.e., \( \eta(P_t) = \eta_0 \).

**Traditional PA.** Let \( \eta(P_t) \) be the power efficiency of PA at output power \( P_t \), then it is given by [19]

\[
\eta(P_t) = \eta_{\text{max}} \left( \frac{P_t}{P_{t,\text{max}}} \right)^{1/2},
\]

where \( P_{t,\text{max}} \) is the maximum designed output power and \( \eta_{\text{max}} \) is the maximum PA efficiency, which is achieved when \( P_t = P_{t,\text{max}} \).

At maximum output power e.g., \( P_t = P_0 \), a PA must be able to handle peak power, therefore \( P_0 \leq P_{t,\text{max}}/\xi \). It means that, for instance, if PAPR of a modulation scheme is 8 dB then \( P_0 \) must be 8 dB less than \( P_{t,\text{max}} \).

**Envelop tracking PA.** For ET-PA, the power efficiency \( \eta(P_t) \) is modeled as [7]

\[
\eta(P_t) = \eta_{\text{max}} \frac{P_t (1 + c)}{P_t + cP_{t,\text{max}}},
\]

where \( c = 0.0082 \) is a PA-dependent constant.

Fig. 2 depicts the efficiency response of these PA models. It can be clearly observed how optimistic CPA model is from real-life PA models, while ETPA model is expected to improve energy efficiency in comparison with TPA.

Using these power efficiency relations, we derive energy consumption per information bit \( (E_0, [9]) \). Table I shows \( E_0 \) and its associated parameters for the consider PA models, where \( R_b = W \log_2 M \) is the PHY layer bit rate in bandwidth \( W \).

| \( E_0 \) | CPA | TPA | ETPA |
|---|---|---|---|
| \( \frac{n_p + n_h}{n_p} A \gamma + B \) | \( \frac{n_p + n_h}{n_p} A \sqrt{\gamma} + B \) | Same as for CPA |
| \( \xi N_0 G_d / \eta_0 \) | \( \xi N_0 G_d / \eta_{\text{max}} \sqrt{N_0 G_d R_b} \) | \( \xi N_0 G_d / \eta_{\text{max}} (c + 1) \) |
| \( P_0 / R_b \) | \( P_{t,\text{max}} / R_b \) | \( 1 / R_b \sqrt{\eta_{\text{max}}(c + 1) + P_0} \) |

**C. Packet Error Rate**

To minimize energy consumption per information bit \( E \) in \([1]\), we require a generic packet error rate (PER) approximation that is accurate and also maintains an explicit connection with the parameters defining the system performance. An accurate approximation for uncoded schemes is proposed in [20], which however cannot be utilized for optimizing the function \( E \). Another, approximation is the upper bound on PER for (un)coded schemes in Rayleigh fading

\[
\tilde{P}_e(\gamma) \leq 1 - \exp \left( -\frac{\omega_0}{\gamma} \right),
\]

where \( \gamma \) is the SNR and \( \omega_0 \) is the waterfall threshold [21]. The threshold is defined as an integral of the PER function in the AWGN channel, \( f(\gamma) \)

\[
\omega_0 = \int_0^{\infty} f(\gamma) \, d\gamma.
\]

For \( N \)-bit uncoded packet with bit error rate (BER) function \( b_e(\gamma) \), \( f(\gamma) \) is defined as

\[
f(\gamma) = 1 - \left( 1 - b_e(\gamma) \right)^N.
\]

A log-domain linear approximation of \( \omega_0 \) is proposed for uncoded schemes in [16], and for (un)coded schemes in [13]. However, the approximation in [16] is tight for large packets only while in [13] the approximation parameters for a given modulation scheme are calculated by simulations. On the other hand, the following new proposition shows that the waterfall threshold in [12] is tightly approximated by the expected value of an asymptotic distribution of \( f(\gamma) \).

**Proposition 1:** For uncoded transmission of a packet with length \( N \), with a BER function described by \( e^{-\kappa m \gamma} \) or
\( c_m Q(\sqrt{k_m} \gamma) \) where \( 0 < c_m \leq 1 \) and \( k_m > 0 \), the threshold. \( \omega_0 \), in Rayleigh fading channel is approximated by the expected value of the Gumbel distribution for sample maximum,

\[
\omega_0 \approx E[\gamma] \triangleq a_N + \gamma_e b_N,
\]

where \( a_N \) and \( b_N \) are the parameters of the Gumbel distribution, and \( \gamma_e \) is the Euler constant.

**Proof:** See Appendix.

In [20], we have shown that the normalizing constants for BER functions of the form \( c_m e^{-k_m \gamma} \) and \( c_m Q(\sqrt{k_m} \gamma) \) are

\[
a_N = \log(Nc_m), \quad b_N = \frac{1}{k_m},
\]

where the optimal parameters for the BER function \( c_m Q(\sqrt{k_m} \gamma) \) are \( c_m \Rightarrow 0.2114c_m \) and \( k_m \Rightarrow 0.5598k_m \) [1].

Inserting (14) and (15) in (11) leads to a simple PER approximation which maintains explicit connection with the modulation specific parameters \( (c_m, k_m) \), packet size \( (N) \) and SNR \( (\gamma) \)

\[
P_e(\gamma) \approx 1 - (N)^{-\frac{1}{k_m \gamma}} \exp \left( -\frac{\log c_m + \gamma_e}{k_m \gamma} \right).
\]

The PER approximation (16) tightly matches to the upper bound, as compared to existing solutions, for any modulation scheme with BER functions in the form of exponential and the Gaussian-Q function. In order to show the accuracy of proposed approximation, using as an example of uncoded 16 QAM, Fig. 3 compares it with the numerical evaluation of (11), and also with the prior approximations in terms of relative error (RE) with respect to real PER. It can be observed that the RE of the proposed approximation is quite close to the upper bound (11) for small to large packet lengths. In comparison, the RE of approximations in [13, 16] is small at low SNR, however it increases rapidly especially for small packet lengths.

**PER for Coded Schemes.** Note that we can easily use the approximation (16) for PER evaluation of coded schemes by equating \( k_M = 1/k_m \), and \( b_M = (\log c_m + \gamma_e)/k_m \), where \( k_M \) and \( b_M \) are calculated by simulations for various coded schemes in [13] Table I. Therefore, one can perform the energy-reliability analysis presented hereafter in this paper for coded schemes also. However, as the low-cost and low-power IoT devices rely on retransmissions instead of error-correction to reduce the system complexity [5], we adhere to uncoded schemes to be compatible with low-cost IoT devices.

IV. LINK OPTIMIZATION WITH MINIMUM ENERGY CONSUMPTION

With two main components of our objective function \( E \) in place, now we find energy-optimal yet QoS-aware (i.e., maintaining the PER constraint in (3)) PHY and MAC parameters.

A. Optimal Average SNR / Transmit Power

To find an energy-optimal SNR, we fix the packet payload \( n_p \). It also represents a case where the sensors have to send fixed-size reports. Moreover, the optimal SNR must satisfy the reliability constraints set via PER and the constraints on the maximum transmit power, which could be either due to hardware limitations or frequency band regulations. Therefore, the optimization of objective function (1) with respect to SNR \( \gamma \) can be written as

\[
\text{minimize} \quad E(\gamma) \\
\text{subject to} \quad \bar{\gamma}_{\text{min}} \leq \gamma \leq \bar{\gamma}_{\text{max}},
\]

where \( \bar{\gamma}_{\text{min}} \) is the minimum average SNR requirement under PER bound in (3). From (16), \( \bar{\gamma}_{\text{min}} \) is

\[
\bar{\gamma}_{\text{min}} = -\frac{\gamma_e + \log \left( c_m (n_h + n_p) \right)}{k_m \log (1 - P_{\text{req}})}.
\]

While \( \bar{\gamma}_{\text{max}} \) is the maximum achievable SNR at transmit power limit \( P_0 \). Since the transmit power \( P_t \) cannot exceed \( P_0 \), i.e., \( P_t \leq P_0 \), it translates to \( \gamma \leq \bar{\gamma}_{\text{max}} \) with \( \bar{\gamma}_{\text{max}} \) from (7) as

\[
\bar{\gamma}_{\text{max}} = \frac{P_0}{W N_0 G_d}.
\]

From (18) and (19), the required SNR, denoted as \( \bar{\gamma}_{\text{req}}^* \) relates to an optimal SNR \( \bar{\gamma}^* \) without any constraints as

\[
\bar{\gamma}_{\text{req}}^* = \begin{cases} \bar{\gamma}_{\text{min}}, & \bar{\gamma}^* < \bar{\gamma}_{\text{min}} \\ \bar{\gamma}_{\text{max}}, & \bar{\gamma}_{\text{min}} < \bar{\gamma}^* < \bar{\gamma}_{\text{max}} \\ \bar{\gamma}^*, & \text{otherwise} \end{cases}
\]

which holds only if the condition \( \bar{\gamma}_{\text{min}} < \bar{\gamma}_{\text{max}} \) is satisfied. Otherwise, the reliability target cannot be satisfied for a given modulation scheme due to transmit power constraint.

The conditions in (20) can be visualized with Fig. 4 obtained using the parameters in Sec. IV Table. II under constant PA. Fig. 4 shows an example case of minimum required SNR \( \bar{\gamma}_{\text{min}} \), to satisfy certain QoS at various distances. It also depicts the unconstrained optimal SNR \( \bar{\gamma}^* \) and maximum achievable SNR \( \bar{\gamma}_{\text{max}} \). At \( d = 10 \), \( \bar{\gamma}_{\text{min}} \) is less than \( \bar{\gamma}^* \), therefore \( \bar{\gamma}^* \) is energy optimal and is preferred over \( \bar{\gamma}_{\text{min}} \). While at \( d = 30 \), \( \bar{\gamma}^* \) cannot satisfy the target and \( \bar{\gamma}_{\text{min}} \), though not energy optimal, is selected. At \( d = 70 \), the reliability target is not satisfied as \( \bar{\gamma}_{\text{min}} > \bar{\gamma}_{\text{max}} \).
We can find $\bar{\gamma}^*$, which is finally conditioned by (20), based on the following unconstrained optimization problem

$$
\text{minimize } E(\bar{\gamma}) \\
\text{subject to } \bar{\gamma} \in [0, \infty). \quad (21)
$$

Note that $E(\bar{\gamma})$ is a product of two functions: the number of retransmissions $\bar{\tau}(\bar{\gamma})$ with $\bar{\tau}(\bar{\gamma}) \leq 0$, and the average energy per transmission attempt $E_0(\bar{\gamma})$ such that $E_0(\bar{\gamma}) \geq 0$ where $\bar{\gamma}$ denotes the first derivative. Given that both $\bar{\tau}(\bar{\gamma})$ and $E_0(\bar{\gamma})$ are convex, the function $E(\bar{\gamma})$ is also convex [13] Lemma 1].

**CPA/ETPA models.** For CPA/ETPA models with $E_0$ in Table [1] $\bar{\gamma}^*$ can be obtained by solving $\frac{dE}{d\bar{\gamma}} = 0$ in (21), which yields a quadratic equation with a positive root as

$$
\bar{\gamma}^* = \frac{\omega_0}{2} + \sqrt{\omega_0^2 + \frac{\omega_0}{4} + \frac{B}{A} \frac{n_p}{n_h + n_p}}. \quad (22)
$$

**TPA model.** For $E_0$ with TPA model in Table [1] although $E$ is convex, the optimal $\bar{\gamma}^*$ is not found explicitly. However, it can be determined numerically from the following non-convex equation

$$
A K_m (\bar{\gamma})^{3/2} - 2 A \omega_0 (\bar{\gamma})^{1/2} - 2 \omega_0 B \left( \frac{n_p}{n_p + n_h} \right) = 0, \quad (23)
$$

Eq. (23) has a real positive root when the condition $27 B^2 k_m (n_p/(n_p + n_h))^2 > 8 A^2 \omega_0$ is satisfied

$$
\bar{\gamma}^* = \frac{1}{3} \left( \frac{4 \omega_0}{k_m} + \frac{4 \omega_0^2}{A^2} + \frac{A^2}{k_m^2} \right), \quad (24)
$$

where

$$
A = \frac{1}{A^2} \left[ 54 B^2 k_m \omega_0^2 \left( \frac{n_p}{n_p + n_h} \right)^2 - 8 A^2 k_m \omega_0^3 + 6 \sqrt{3} \times B^2 k_m \omega_0^2 \left( \frac{n_p}{n_p + n_h} \right)^2 \left( 27 B^2 k_m \left( \frac{n_p}{n_p + n_h} \right)^2 - 8 A^2 \omega_0 \right) \right].
$$

**B. Optimal Packet Payload**

Now we find the optimal $n_p$—the payload that minimizes the objective function (1)—by keeping SNR constant. The upper limit on the maximum payload size $n_p, \text{max}$ is determined by required $\gamma_\text{min}$ to satisfy a PER target. Therefore, our objective function is

$$
\begin{align*}
\text{minimize} & \quad E(n_p) \\
\text{subject to} & \quad 0 \leq n_p \leq n_p, \text{max}, \\
& \quad n_p, \text{max} \text{ from (3)}, \text{is}
\end{align*}
$$

$$
n_p = \frac{n_h \bar{\gamma} \left( k_m (1 - \bar{\gamma}) + \sqrt{k_m^2 \bar{\gamma}^2 + 2 k_m \bar{\gamma} + 4 B K_m + 1} \right)}{2 \bar{\gamma}^2 + B^2 \bar{\gamma}^2}. \quad (27)
$$

**TPA model.** For $E_0$ with TPA model in Table [1] although $E$ is convex, the optimal $\bar{\gamma}^*$ is not found explicitly. However, it can be determined numerically from the following non-convex equation

$$
A k_m (\bar{\gamma})^{3/2} - 2 A \omega_0 (\bar{\gamma})^{1/2} - 2 \omega_0 B \left( \frac{n_p}{n_p + n_h} \right) = 0, \quad (23)
$$

Eq. (23) has a real positive root when the condition $27 B^2 k_m (n_p/(n_p + n_h))^2 > 8 A^2 \omega_0$ is satisfied

$$
\bar{\gamma}^* = \frac{1}{3} \left( \frac{4 \omega_0}{k_m} + \frac{4 \omega_0^2}{A^2} + \frac{A^2}{k_m^2} \right), \quad (24)
$$

where

$$
A = \frac{1}{A^2} \left[ 54 B^2 k_m \omega_0^2 \left( \frac{n_p}{n_p + n_h} \right)^2 - 8 A^2 k_m \omega_0^3 + 6 \sqrt{3} \times B^2 k_m \omega_0^2 \left( \frac{n_p}{n_p + n_h} \right)^2 \left( 27 B^2 k_m \left( \frac{n_p}{n_p + n_h} \right)^2 - 8 A^2 \omega_0 \right) \right].
$$

**C. Joint Energy Optimal Parameters—$\bar{\gamma}, n_p, M, \tau^\text{max}$**

As the IoT devices will be used in diverse monitoring and control applications, it might be important in many to find the optimal SNR, payload size, modulation order and the number of retransmissions for energy efficient communication. For example, after deployment in harsh and inaccessible areas, the devices can optimize these parameters at the start of their operation, and then continue with the optimal link setting. The problem to jointly optimize these parameters can be formulated as

$$
\begin{align*}
\text{minimize} & \quad E(\bar{\gamma}, n_p, M, \tau^\text{max}) \\
\text{subject to} & \quad \bar{\gamma} \in \{\text{FSK, MPSK, MQAM}\}, \quad \tau^\text{max} = i, \quad i \geq 1,
\end{align*}
$$

where $M \in \{\text{FSK, MPSK, MQAM}\}$ and $\tau^\text{max} = i, \quad i \geq 1$. Note that the IoT devices will support only a few modulation schemes $M$. In addition, a small value of maximum retransmissions $\tau^\text{max}$ is feasible for operation with minimum energy consumption [13]. As a result, an exhaustive search over the combination of $M$ and $\tau^\text{max}$ will not be computationally demanding.

On the other hand, for each combination of $M$ and $\tau^\text{max}$, the joint optimum of $\bar{\gamma}$ and $n_p$ can be found from (22) and (27) either by solving a system of two non-linear equations or by iteratively invoking these equations. In either case, we must ensure that the reliability conditions in (20) and (25) are satisfied. However, the former approach requires numerical evaluation that might not be computationally feasible for hardwired-constrained IoT devices. Whereas, by iteratively invoking (22) and (27), $\bar{\gamma}$ and $n_p$ can efficiently converge to joint energy
Algorithm 1 Joint Energy-Reliability Aware Optimization of a Link’s Parameters

Input: \( P_{\text{req}}, \gamma_{\text{max}}, \delta \)
Output: \( \gamma^*, n_p^*, \tau_r^*, M^* \)

1: for \( M \in \{\text{FSK, MPSK, MQAM}\} \) do
2: for \( i = 1 \) to \( \gamma_{\text{max}} \) do
3: \( n_p \leftarrow 0 \)
4: while \( \Delta > \delta \) do
5: \( \bar{\gamma} \leftarrow \text{Evaluate } (22) \) or \( (23) \), \( \tau_{\text{req}} \leftarrow \text{Evaluate } (18) \)
6: \( \gamma_{\max} \leftarrow \text{Evaluate } (19), n_{p,\max} \leftarrow \text{Evaluate } (26) \)
7: if \( (\gamma_{\min} > \bar{\gamma}) \) then
8: \( \text{break; } \)
9: else
10: \( \gamma_{\text{req}} \leftarrow \text{Evaluate } (20) \)
11: end if
12: \( n_p \leftarrow \text{Evaluate } (27) \) or \( (28) \) with \( \bar{\gamma} = \gamma_{\text{req}} \)
13: if \( (n_p > n_{p,\max}) \) then
14: \( n_p \leftarrow n_{p,\max} \)
15: end if
16: \( E \leftarrow \text{Evaluate } (4) \), Print \( E, \gamma, n_p, \tau_r, M \)
17: \( \Delta \leftarrow \gamma_{\text{req}} - \bar{\gamma}, \bar{\gamma} = \gamma_{\text{req}} \)
18: end while
19: end for
20: return \( \gamma^*, n_p^*, \tau_r^*, M^* \) giving minimum \( E \)

optimum values while satisfying the reliability conditions. It is straightforward to develop the proof of convergence of the iterative approach [13] Corollary 3 under the probabilistic QoS model. We observed that by initializing \( n_p \) and \( \bar{\gamma} \) to any random values, this approach converges to their optimum values within a few iterations.

A pseudocode of the proposed joint optimization of (29) is given in Algorithm 1.

V. NUMERICAL RESULTS

In this section, we present the numerical results of link optimization. The simulation parameters are given in Table II where the PER target of \( P = 0.001 \) translates to 99.9% reliability.

### TABLE II

**Simulation Parameters**

| Parameter                  | Symbol | Value      |
|----------------------------|--------|------------|
| Max. transmit power        | \( P_0 \) | 10 mW      |
| Noise power density        | \( N_0/2 \) | -174 dBm/Hz |
| Path loss exponent         | \( \kappa \) | 3.5        |
| Path gain at unit distance | \( G_1 \) | 30 dB      |
| Link margin                | \( M_f \) | 40 dB      |
| Bandwidth                  | \( W \) | 10 kHz     |
| Circuit power - MQAM, MFSK| \( P_c \) | 310 mW, 265 mW |
| Max. PA efficiency         | \( \eta_0 \) | 80 %      |
| Packet overhead            | \( n_h \) | 48 bit    |
| Target PER                 | \( P \) | 0.001      |
| Max. retransmission        | \( \tau_{\text{max}} \) | 3          |

When operated at optimal SNR and payload size, the energy consumption (EC) of the selected modulation schemes with respect to distance with (solid lines) and without (dotted-marker lines) reliability constraints is shown in Fig. 5. These results are obtained based on constant PA. It is observed that there is an energy-optimal modulation scheme at each distance that also satisfies the reliability target: high-order modulations at lower distance and lower-order at higher distance. However, for a given transmit power limit, the distance until which the reliability target is satisfied decreases as the reliability requirements get tight. Also, how energy efficiency takes a toll under reliability constraints, compared to a link with unlimited allowed transmissions, can also be observed especially for high-order modulations. Note that the EC gap increases with reliability constraints becoming stringent. This is because higher than the energy-optimal is required to meet the reliability target. We observed the similar EC trend for other PA models however with exceptions that need in depth analysis.

Using Algorithm 1, Fig. 6 gives an holistic view of how energy-optimal parameters—modulation size (\( M \)), SNR (\( \gamma \)) or transmit power (\( P_r \)), payload (\( n_p \)) and the number of retransmissions (\( \tau_r \))—vary with the link distance. Mainly, it compares the impact of PA models on the EC while closely looking at the optimal system parameters. Fig. 6(a) shows that the EC under constant PA (CPA)-model is considerably optimistic compared to realistic PA models. Although the EC for envelope-tracking PA (ETPA) is higher and closely follows the EC for CPA, it is considerably higher for traditional PA (TPA). As a result, TPA model sees a link switching to a lower-order modulation at shorter distances. Thus, TPA reduces the gain in using high-order modulation with respect to a device utilizing only OQPSK modulation.

In Fig. 6(b)–Fig. 6(d) we can observe the trade-off or interplay among the link parameters in minimizing EC under ETPA and TPA, where the parameters under CPA are omitted for the clarity of figures. Nevertheless, their trend is similar to the parameters for ETPA. In general, the behavior of these parameters must be interpreted together with the PA-efficiency curves in Fig. 2. It can be noticed that, for the devices at short distances, high-order modulation and high-power transmissions are energy efficient under ETPA. Although ETPA has higher power consumption (\( P_{\text{req}} = \xi P_r/n(P_r) \)) than TPA at the same distances (see Fig. 6(b)(top)) for distance
Fig. 6. Impact of PA models on energy consumption and optimal PHY/MAC parameters. Figures (a) and (b) are essentially reciprocal results for $\bar{\gamma}$ and $P_t$, respectively, while (b) shows additional information on PA power consumption ($P_{pa}$).

1 m–11 m), thanks to better linearity of ETPA, high-power transmissions are capitalized by the devices to successfully transfer of (small-sized) information bits with a low number of retransmissions. On the other hand, power consumption of TPA is smaller but, with the increase in distance, it causes the increase in EC at a higher rate than the ETPA. As a result, TPA leads to almost third modulation degradation (i.e., to OQPSK) while 64 QAM is still energy optimal under ETPA. The reason behind low energy efficiency of TPA can be explained from its inefficiency at low output powers. In order to compensate for limited gain at transmitting at high-power at short distances, the devices opt for small transmit power but with a higher payload size and higher number of retransmissions.

In general, the optimal modulation becomes smaller as the distance increases. However with ETPA, at distances around 11 m–24 m, 16 QAM achieves better energy efficiency than TPA owing to its better trade-off in efficiency and the other link parameters (as discussed earlier). An interesting observation in this range is the rapid increase in payload size with the distance, which starts at around 17 m. However, note that, the corresponding SNR (in Fig. 6(c)) remains constant in this distance range. That is because increasing the transmit power with increasing packet size is optimal until next smaller modulation becomes energy optimal. If observed carefully, a small jump in payload size before the ETPA link downgrades its modulation from 64 QAM to 16 QAM can also be noticed.

When comparing the results where the same modulation scheme is employed under both the ETPA and TPA (i.e., at distances $\geq 24$ m), it can be observed that both transmit power and power consumption for TPA jumps higher than ETPA for the first time and their corresponding EC become almost identical. This is because the TPA efficiency increase exponentially in this range of output power, and also the small PAPR of OQPSK modulation reduces the amount of backoff from the saturation region. As consequence, this has an effect on the link parameters under TPA where the payload starts reducing with the distance and the number of retransmissions also reduce at some distances. However, as soon as the link operates at its maximum allowed transmit power, the link parameters under two PAs, including the modulation scheme (i.e., BPSK), become identical.

A. Lifetime Analysis of IoT Links/Devices

The optimal link parameters, as discussed earlier, allow analyzing the lifetime of IoT devices while considering their non-ideal hardware characteristics. For the analysis, we assume
that each device is battery-powered, and it reserves a charge capacity of 2 Ah exclusively for data communication. Also, each device, located at a certain distance from the gateway, assumed to be transmitting 5 kbit of sensory data on the average within a period of 5 min.

Fig. 7 shows the lifetime of a device with respect to link distance for the studied PA models, when operated with optimized link parameters. It also depicts the lifetime of a device employing only OQPSK modulation, while the selection of other link parameters at any distance is energy-reliability optimal. We observe that spectral efficient modulations can significantly prolong the lifetime of the devices located at short-range distances, i.e., within 1 m–24 m under CPA and ETPA. The expected lifetime increases as the distance decreases, and compared to a commonly employed low-order modulation (i.e., OQPSK) in IoT devices, the lifetime can be extended up to 180% in case of an ideal PA and 125% for ETPA. On the other hand, traditional PA (TPA) not only halves the range in which high-order modulation can help in extending the device/link life, it also brings down any gains of employing high-order modulations significantly.

VI. CONCLUSIONS

We studied cross-layer optimization of battery-operated wireless links with energy consumption (EC) minimization objective while considering: (a) reliability requirements of IoT applications, and (b) operational constraints and non-ideal characteristics of the IoT devices’ hardware. To this end, we derived EC models while capturing energy cost of ideal and realistic power amplifiers (PAs), and packet error statistics in particular. For packet error statistics, we developed an accurate PER approximation in Rayleigh block-fading, which is simple to exploit for cross-layer link design and optimization. Using the EC models, we derived energy-optimal, reliability-aware, and hardware-compliant conditions for SNR and payload size, which we exploited for developing a holistic algorithm to jointly optimize the link parameters. The algorithm can be utilized by resource-constrained devices for link adaptation based on the optimal selection of parameters.

Our path to link optimization allowed to make useful observations, especially when the target is to simultaneously minimize energy consumption and ensure certain reliability. We found that at each distance there is an optimal SNR and payload size, which leads to an energy optimal modulation scheme where the modulation order increases for short-range links. However a reliability-aware link compared to a delay tolerant system reduces this distance, and increases the energy consumption. Also, when non-realistic power amplifier (PA) characteristics are taken into account, we found that the traditional PA significantly diminishes the gain in the usage of higher modulations (causing both the higher energy consumption for higher order modulations, and switching to smaller modulations at shorter distances). While an envelop-tracking PA behaves closely to an ideal PA. Since the higher order modulations offer packing more information bits, they are energy-optimal at short distances, because circuit power exceeds PA power consumption. However, PA’s non-linearity makes these powers comparable, and leads to a higher energy consumption compared to an ideal PA.

Our lifetime analysis found that under ideal PA the lifetime can be extended up to 180% by selecting higher-order modulations in short-range networks compared to OQPSK. However, this gain reduces significantly under traditional PA, and the distance up to which a higher-order modulation can provide any gain reduces to half. Whereas, the envelop-tracking PA can still provide the lifetime extension of up to 125%. These results highlight the need to investigate efficient PAs for short-range IoT networks in order to gain from spectral efficient modulations.

APPENDIX

PROOF OF PROPOSITION 1

For packet length $N$, the PER function in [15] for BER functions described by $c_m e^{-k_m \gamma}$ and $c_m Q(\sqrt{k_m \gamma})$ is asymptotically approximated by the Gumbel distribution function for the sample minimum [20]

$$f(\gamma) \simeq 1 - \exp \left( -\exp \left( -\frac{\gamma - a_N}{b_N} \right) \right),$$

(30)

where $a_N$ and $b_N > 0$ are the normalizing constants.

Let $G(\gamma) = \exp(-\exp(-\frac{h_{\max}}{b_N}))$ be the cumulative distribution function (CDF) of the Gumbel distribution for the sample maximum, then from (30) and (12) we have

$$\omega_0 \approx \int_0^\infty \left( 1 - G(\gamma) \right) d\gamma,$$

(31)

which is the expected value of a continuous and non-negative random variable $\gamma$.

REFERENCES

[1] A. Mahmood, M. M. A. Hossain, and M. Gidlund, “Cross-layer optimization of wireless links under reliability and energy constraints,” in IEEE WCNC, April 2018, pp. 1–6.

[2] H. Shariatmadari, R. Ratasuk, S. Iraji, A. Laya, T. Taleb, R. Jianiti, and A. Ghosh, “Machine-type communications: current status and future perspectives toward 5G systems,” IEEE Commun. Mag., vol. 53, no. 9, pp. 10–17, September 2015.
[3] G. Durisi, T. Koch, and P. Popovski, “Toward massive, ultrareliable, and low-latency wireless communication with short packets,” Proc. of IEEE, vol. 104, no. 9, pp. 1711–1726, Sept 2016.

[4] Y. Chen, H. Zhang, N. Fisher, L. Y. Wang, and G. Yin, “Probabilistic per-packet real-time guarantees for wireless networked sensing and control,” IEEE Trans. Ind. Informat., vol. 14, no. 5, pp. 2133–2145, May 2018.

[5] T. Shafique, A. M. Abdelhady, O. Amin, and M. Alouini, “Energy efficiency, spectral efficiency and delay analysis for selective ARQ multichannel systems,” IEEE Trans. Green Commun. Netw., vol. 2, no. 3, pp. 612–622, Sept 2018.

[6] J. Joung, C. K. Ho, K. Adachi, and S. Sun, “A survey on power-amplifier-centric techniques for spectrum- and energy-efficient wireless communications,” IEEE Commun. Surveys Tuts., vol. 17, no. 1, pp. 315–333, Firstquarter 2015.

[7] M. A. Hossain and R. Jantti, “Impact of efficient power amplifiers in wireless access,” in IEEE Online Conf. on Green Communications (GreenCom), 2011, pp. 36–40.

[8] Q. Cui, Y. Zhang, W. Ni, M. Valkama, and R. Jantti, “Energy efficiency maximization of full-duplex two-way relay with non-ideal power amplifiers and non-negligible circuit power,” IEEE Trans. W. Commun., vol. 16, no. 9, pp. 6264–6278, Sept 2017.

[9] S. Cui, A. J. Goldsmith, and A. Bahai, “Energy-constrained modulation optimization,” IEEE Trans. W. Commun., vol. 4, no. 5, pp. 2349–2360, 2005.

[10] T. Wang, W. Heinzelman, and A. Seyedi, “Minimization of transceiver energy consumption in wireless sensor networks with AWGN channels,” in 40th Conf. on Commun., Control, and Computing, 2008, pp. 62–66.

[11] Y. Hou, M. Hamamura, and S. Zhang, “Performance tradeoff with adaptive frame length and modulation in wireless network,” in 5th Intl. Conf. on Computer and Inf. Tech., 2005, pp. 490–494.

[12] F. Rosas and C. Oberli, “Modulation and SNR optimization for achieving energy-efficient communications over short-range fading channels,” IEEE Trans. W. Commun., vol. 11, no. 12, pp. 4286–4295, 2012.

[13] J. Wu, G. Wang, and Y. R. Zheng, “Energy efficiency and spectral efficiency tradeoff in type-I ARQ systems,” IEEE J. Sel. A. in Commun., vol. 32, no. 2, pp. 356–366, 2014.

[14] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, “Industrial internet of things: Challenges, opportunities, and directions,” IEEE Trans. Indus. Informat., pp. 1–1, 2018, DOI:10.1109/TII.2018.2852491.

[15] A. Zappone, L. Sanguinetti, and M. Debbah, “Energy-delay efficient power control in wireless networks,” IEEE Trans. Commun., vol. 66, no. 1, pp. 418–431, Jan 2018.

[16] S. Liu, X. Wu, Y. Xi, and J. Wei, “On the throughput and optimal packet length of an uncoded ARQ system over slow Rayleigh fading channels,” IEEE Commun. Lett., vol. 16, no. 8, pp. 1173–1175, 2012.

[17] S. Vitturi, L. Seno, F. Tramarin, and M. Bertocco, “On the rate adaptation techniques of IEEE 802.11 networks for industrial applications,” IEEE Trans. Indus. Informat., vol. 9, no. 1, pp. 198–208, Feb 2013.

[18] Q. Wang and J. Jiang, “Comparative examination on architecture and protocol of industrial wireless sensor network standards,” IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 2197–2219, thirdquarter 2016.

[19] K. Han, Y. Choi, S. Choi, and Y. Kwon, “Power amplifier characteristic-aware energy-efficient transmission strategy,” Lecture Notes in Computer Science, vol. 4479, p. 37, 2007.

[20] A. Mahmood and R. Jantti, “Packet error rate analysis of uncoded schemes in block-fading channels using extreme value theory,” IEEE Commun. Lett., vol. 21, no. 1, pp. 208–211, Jan 2017.

[21] Y. Xi, A. Burr, J. Wei, and D. Grace, “A general upper bound to evaluate packet error rate over quasi-static fading channels,” IEEE Trans. W. Commun., vol. 10, no. 5, pp. 1373–1377, 2011.