Delay Optimization for Industrial Wireless Control Systems based on Channel Characterization

***Abstract*—Wireless communication is gaining popularity in industry for its simple deployment, mobility, and low cost. Ultra low latency and high reliability requirements of mission critical industrial applications are highly demanding for wireless communication, and the indoor industrial environment is hostile to wireless communication due to the richness of reflection and obstacles. Assessing the effect of the industrial environment on the reliability and latency of wireless communication is a crucial task, yet it is challenging to accurately model the wireless channel in various industrial sites. Based on the comprehensive channel measurement results from the National Institute of Standards and Technology at 2.245 GHz and 5.4 GHz, we quantify the reliability degradation of wireless communication in multi-path fading channels. A delay optimization based on the channel characterization is then proposed to minimize packet transmission times of a cyclic prefix orthogonal frequency division multiplexing system under a reliability constraint at the physical layer. When the transmission bandwidth is abundant and the payload is short, the minimum transmission time is found to be restricted by the optimal cyclic prefix duration, which is correlated with the communication distance. Results further reveal that using relays may in some cases reduce end-to-end latency in industrial sites, as achievable minimum transmission time significantly decreases at short communication ranges.**

***Index Terms*—channel measurement analysis, industrial applications, physical layer transmission time optimization, end-to-end latency reduction of relay.**

I. Introduction

Compared to mobile communication applications, industrial communication applications have a lower tolerance to faults and require real-time communication [1]. Different examples of industrial applications can be mentioned that pose strict latency and reliability constraints on the industrial communication network. By supervising the production activities, process automation aims at more efficient and safe operation of (for example) paper, mining, and cement industries [2]. Factory automation includes timing-constrained operational applications, such as those used for motion control and work-cell operation [3]. Power system automation refers to automatic monitor, control, and protection of power system components via instrumentation and control devices [4]. In this paper we generally refer to the high performance networks required in this scenario as high-performance wireless (WirelessHP).

Distinct from home and office environments, industrial sites are filled with reflective material such as sheet metal which cause rich resonance and fading effects. The multi-path delay spread will cause inter-symbol interference in wireless communication which can degrade the reliability performance and increase latency. Inter-symbol interference becomes more severe with longer delay spreads, which leads to more complex equalization algorithms and consequently increases processing delay. When equalization delays occur in delay sensitive applications, there is a trade-off between reliability and latency, both of which are demanded by many industrial applications.

Few papers characterize the delay in industrial wireless communication. The delay profiles for an empty hall and one filled with metallic materials at 433 MHz were presented in [5], and the root mean square (RMS) delay spread was found shorter in the empty hall. [6] presented the power delay profile for line-of-sight (LOS) and non-line-of-sight (NLOS) scenarios with different antenna polarization combinations in an industrial hall, whose dominant building material is concrete. RMS delay spread was reported to increase with distance in the industrial indoor scenarios [7], [8]. The above papers all chose RMS delay spread to characterize the multi-path fading. However, this metric fails to quantify reliability degradation accurately, especially when the targeted packet error rate (PER) is very low, $10^{-6} - 10^{-9}$. In [9], the authors proposed to characterize the ratio of the energy of the paths arriving within a duration range and the energy arriving outside this range, and delay window was defined as the minimum duration to satisfy the given ratio. A similar parameter at 5.85 GHz with short distance is also characterized in [10].

In this paper, we investigate delay window statistics in industrial sites looking at the comprehensive measurement campaign results from National Institute of Standards and Technology (NIST) [11], which include three representative industrial sites at 2.245 GHz and 5.4 GHz frequency bands. With delay window results, we propose a delay optimization based on channel characterization (CCDO) to design the optimal parameters in a cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) system with the goal of minimizing the transmission time under a given reliability requirement. CCDO is a modification to the packet transmission optimization in [12], where the impact of cyclic prefix (CP) length on reliability is not taken into consideration.

The remainder of the paper is organized as follows: In Section II, we introduce the background of WirelessHP [12], the measurement campaign conducted by NIST, and formulate the CCDO problem. In Section III, the CCDO algorithm is introduced, and the data analysis is given in Section IV. Section V concludes the paper and discusses future work.

II. Problem Formulation

A. Background of WirelessHP

The demanded performance of wireless communication varies among different industrial scenarios, in which different physical processes are monitored and/or controlled. The demands of the industrial scenarios can be transformed into quantified requirements including sampling rate, number of the nodes, and PER. The latency level of different scenarios can be characterized by the scheduling unit (SU), which is a link-level requirement indicating the minimum time in which a packet

TABLE I
CRITICAL INDUSTRIAL USE CASES AND ACCORDING REQUIREMENTS [13](PA: PROCESS AUTOMATION, FA: FACTORY AUTOMATION, PSA: POWER SYSTEMS AUTOMATION)

Use Case	Sampling rate (Hz)	# of Nodes	SU	PER
PA	$10^0 - 10^1$	$10^2 - 10^3$	$\approx$1 ms	$\approx 10^{-6}$
FA	$10^2 - 10^3$	$10^2 - 10^3$	$\approx$10 us	$\approx 10^{-6}$
PSA	$10^3 - 10^4$	$10^1 - 10^2$	$\approx$ x us	$\approx 10^{-9}$

can be exchanged between a controller and a sensor/actuator node [13]. With the rough assumptions of no multiplexing at the space and frequency domain, and no redundancy in the time domain, SU can be calculated as the cycle time (inverse of sampling rate) divided by the number of nodes. The SU can be modified by a multiplying factor when multiplexing or redundancy is used.

The requirements of the three industrial uses cases discussed in the previous section are listed in Table I [13]. From Table I, besides the high reliability requirement, the scenarios such as factory automation and power systems automation also set stringent delay requirements with SUs of several microseconds. Reviewing the existing wireless technologies, there are large gaps between these technologies and the aforementioned required SUs. WirelessHART [14] and ISA 100.11.a [15], which are built on the IEEE 802.15.4 Wireless Personal Area Network (WPAN) physical (PHY) layer and rely on timeslotted channel hopping, achieve a 10 ms slot time (SU). Wireless Interface for Sensors and Actuators (WISA) [16], built on IEEE 802.15.1 and exploiting timedivision multiple access (TDMA) and frequency hopping, can schedule 128 μs and 64 μs for uplink and downlink SU's, respectively. Finally, Wireless Network for Industrial Automation - Factory Automation (WIA-FA) [17], built on IEEE 802.11 and relying on TDMA, achieves a 100 μs SU.

The recently proposed WirelessHP network [12] aims to provide a solution to support multi-Gbps aggregate data rate, very high reliability ranging from 10^{-6} to 10^{-9} PER, and SU lower than 1 μs.

B. Problem Formulation for CCDO

At the physical (PHY) layer, minimizing the packet transmission time in the CP-OFDM system is an effective way to reduce the SU. The WirelessHP network proposes a PHY OFDM packet structure with flexible FFT size, we use this structure for the optimization in this paper. Each PHY packet consists of preamble symbols (the main purpose of which are packet detection, channel estimation, etc.) and data symbols (that are used to carry the data information). Preamble symbols use the same fast Fourier transform (FFT) size as the data symbol for easy implementation of channel estimation, thus the symbol duration T_{sym} is the same for the two kinds of symbols. T_{sym} is a function of the CP duration T_{CP} and the FFT size N_{FFT}:

$$T_{\text{sym}}(T_{\text{CP}}, N_{\text{FFT}}) = T_{\text{s}} \cdot (N_{\text{CP}} + N_{\text{FFT}}), \tag{1}$$

where T_{s} is the sample duration (which usually equals the inverse of the bandwidth B, $T_{\text{s}} = 1/B$), and N_{CP} is the number of samples for the CP ($N_{\text{CP}} = \lceil T_{\text{CP}}/T_{\text{s}} \rceil$).

In OFDM systems, the whole bandwidth is divided into N_{FFT} sub-carriers, which consists of N_{gsc} guard sub-carriers, N_{psc} pilot sub-carriers, N_{dcsc} 0 Hz null sub-carrier, and N_{dsc} data sub-carriers to allocate the data to be transmitted:

$$N_{\text{dsc}}(N_{\text{FFT}}) = N_{\text{FFT}} - N_{\text{gsc}} - N_{\text{dcsc}} - N_{\text{psc}}. \tag{2}$$

The number of OFDM data symbols $N_{\text{sym}}^{\text{data}}$ is determined by N_{dsc}, the raw payload size L in bit, coding rate R_c, modulation order M that transform $\log_2 M$ coded bits into one modulated symbol, and N_{ss} the number of spatial streams,

$$N_{\text{sym}}^{\text{data}}(N_{\text{FFT}}) = \lceil \frac{L}{N_{\text{dsc}} \cdot \log_2 M \cdot R_c \cdot N_{\text{ss}}} \rceil. \tag{3}$$

The packet transmission time T_{pkt} can be calculated as

$$\begin{aligned} T_{\text{pkt}}(T_{\text{CP}}, N_{\text{FFT}}) &= T_{\text{sym}} \cdot (N_{\text{sym}}^{\text{pre}} + N_{\text{sym}}^{\text{data}}) \\ &= T_{\text{s}} \cdot (N_{\text{CP}} + N_{\text{FFT}}) \cdot (N_{\text{sym}}^{\text{pre}} + N_{\text{sym}}^{\text{data}}) \\ &\geq (T_{\text{CP}} + T_{\text{s}} \cdot N_{\text{FFT}}) \cdot (N_{\text{sym}}^{\text{pre}} + N_{\text{sym}}^{\text{data}}). \end{aligned} \tag{4}$$

where number of preamble symbols $N_{\text{sym}}^{\text{pre}}$ is fixed, and does not change with the payload size. Thus given a payload size L, minimizing T_{pkt} equals to finding the best tradeoff between N_{FFT} and $N_{\text{sym}}^{\text{data}}$, and to determine the minimum required T_{CP}.

We are now ready to give the expression of the CCDO problem, which aims to minimize T_{pkt} while satisfying the reliability requirement:

$$\begin{aligned} \min_{N_{\text{FFT}}, T_{\text{CP}}} \quad & T_{\text{pkt}}(T_{\text{CP}}, N_{\text{FFT}}) \\ \text{s.t.} \quad & \frac{B}{N_{\text{FFT}}} \leq \frac{1}{T_{\text{CP}}}, && \text{(5a)} \\ & \frac{B}{N_{\text{FFT}}} \geq F_{\text{center}} \cdot 4\,\text{ppm}, && \text{(5b)} \\ & N_{\text{gsc}} \geq N_{\text{FFT}} \cdot \text{GBR}_{\text{min}}, && \text{(5c)} \\ & N_{\text{dsc}}(N_{\text{FFT}}) > 0, && \text{(5d)} \\ & N_{\text{FFT}} = 2^m, m \in N, && \text{(5e)} \\ & \text{PER}(T_{\text{CP}}) \leq \text{PER}_{\text{req}}. && \text{(5f)} \end{aligned}$$

We explain the meaning of the constraints in the following paragraphs. Constraint (5a) requires the sub-carrier bandwidth being smaller than the coherence bandwidth, which can be calculated as the inverse of T_{CP}. This constraint guarantees each sub-carrier to experience flat fading, and it also sets a lower limit to N_{FFT}. Combining (4) and (5a), we find the lower bound of T_{pkt},

$$T_{\text{pkt}} \geq 2T_{\text{CP}} \cdot (N_{\text{sym}}^{\text{pre}} + N_{\text{sym}}^{\text{data}}). \tag{6}$$

Constraint (5b) requires that the sub-carrier bandwidth be larger than the carrier frequency offset (CFO) for easy implementation of the sub-carrier equalization. This constraint sets an upper limit to N_{FFT}. CFO consists of Doppler frequency offset and that caused by the mismatch between the oscillators of the transmitter and receiver, and frequency offset caused by Doppler shift can be ignored compared to the oscillator mismatch. The oscillator precision tolerance is specified to be less than 20 ppm in a standard-compliant communication system, and with a frequency offset estimation to reduce the

TABLE II
INDUSTRIAL SITES IN THE NIST MEASUREMENT CAMPAIGN

Site	Dimension	Contents
Automotive Assembly Plant	400 m×400 m×12 m	machining areas, inspection machines, assembly work cells, stacked storage areas
Boiler	50 m×80 m×7.6 m	large machinery, overhead obstructions
Machine shop	12 m×15 m×7.6 m	machines

offset by 10 times, 4 ppm of the central frequency is used to represent the CFO.

Constraint (5c) sets the minimum required number of guard sub-carriers to avoid interference from the neighboring band by setting a guard band ratio $\mathrm{GBR}_{\mathrm{min}}$.

Constraint (5d) demands N_{dsc} to be positive. According to the IEEE 802.11g/n standard, N_{psc} is set to 4, and N_{dcsc} is set to 1 when calculating N_{dsc} in (2).

Constraint (5e) demands N_{FFT} to be a power of 2.

Constraint (5f) explicitly sets the PER requirement $\mathrm{PER}_{\mathrm{req}}$, which is an implicit function of T_{CP} as will be explained in Section III.

Proposition 1. *The optimal solution of T_{CP} in Problem (5) is given by the minimum solution of constraint (5f).*

Our CCDO problem formulation in (5) is substantially different from the optimization problem formulated in [12]: constraint (5f) plays a fundamental role in terms of reliability. If we remove constraint (5f) and take T_{CP} as an independent value, we would achieve the optimization problem in [12], which has two main drawbacks compared to CCDO. First, the optimization problem in [12] neither explicitly includes the reliability performance nor can it guarantee the reliability requirement. Moreover, in (5) we include as decision variable T_{CP}, which was not present in the problem in [12]. How to determine such a variable demands substantially new arguments compared to [12], as we show in the next sections.

In CCDO, the minimum required T_{CP} under the PER requirement is quantitatively analyzed and reported in Section III based on the industrial channel measurement carried out at NIST, which is described in the following subsection

C. Industrial Channel Measurement Campaign

The measurement campaign included three representative sites with different size and contents, which are listed in Table II. The layout figures and description of the three sites can be found in [11]. During the measurements, the transmitter was fixed, while the receiver was mounted on a cart, moving at constant speed along the chosen routes, which are presented with black solid lines in the figure.

2.245 GHz and 5.4 GHz were chosen as the central frequencies as they are licensed U.S. government bands, and are close to the unlicensed industrial, scientific, and medical radio bands at 2.4 and 5 GHz, respectively. The measurements were performed by the NIST Channel Sounder [18]. The channel impulse responses are stored with associated Cartesian coordinates and time information [19]. In this paper, we derive the delay window of the three industrial sites from the channel impulse responses.

TABLE III
$\mathrm{DW}_{\mathrm{min}}$ PARAMETERS FOR THE MEASUREMENTS AT AUTOMOTIVE ASSEMBLY PLANT.

	2.245 GHz		5.4 GHz	
	10 dB	20 dB	10 dB	20 dB
10 m	280 ns	450 ns	340 ns	550 ns
50 m	400 ns	600 ns	500 ns	800 ns
100 m	750 ns	1100 ns	950 ns	1400 ns

TABLE IV
$\mathrm{DW}_{\mathrm{min}}$ PARAMETERS FOR THE MEASUREMENTS AT BOILER.

	2.245 GHz		5.4 GHz	
	10 dB	20 dB	10 dB	20 dB
10 m	220 ns	350 ns	220 ns	380 ns
20 m	270 ns	410 ns	280 ns	430 ns
30 m	310 ns	450 ns	330 ns	440 ns

III. CCDO ALGORITHM

In this section, the two novel features of CCDO algorithm are explained. First, channel delay characterization is analyzed with delay window (DW) and self-signal-to-interference-ratio (self-SIR) based on the channel measurement results conducted at NIST. Then, by explicitly considering the reliability requirement, the minimum required DW is derived, which can be set for the T_{CP} in (5a) for transmission time optimization while satisfying the given reliability requirements.

A. Channel Delay Characterization

The CP-OFDM system can achieve full equalization if all the delayed paths arrive within the CP range. However, it is inefficient to choose the CP length equal to the delay spread (DS) in the industrial sites, as the delay spread is generally very long due to the rich reflective environment. Following the definition of delay window in [9], we denote with t_1 the determined start of the OFDM symbol and with DW a certain duration, so that the paths that arrive at [t_1, t_1+DW] are usable signal, while the others arriving outside this range will act as interference. DW is shorter than the overall DS.

We then define self-SIR to quantify the ratio of the usable signal power and the interference power, and "self" characterizes the interference is caused inherently by the desired signal rather than by other transmitters,

$$\text{self-SIR} = \frac{\int_{t_1}^{t_1+\mathrm{DW}} |h(t)|^2 dt}{\int_0^{\mathrm{DS}} |h(t)|^2 dt - \int_{t_1}^{t_1+\mathrm{DW}} |h(t)|^2 dt}, \qquad (7)$$

where $h(t)$ is the complex channel impulse response. For each DW value, we assume that t_1 to maximize the self-SIR can always be found. Thus for each self-SIR, the corresponding minimum required DW value $\mathrm{DW}_{\mathrm{min}}$ can be determined, which also equals the the minimum required CP length in the OFDM system.

We compute the $\mathrm{DW}_{\mathrm{min}}$ with 10 dB and 20 dB self-SIR, different distances in the three sites, and list the numerical results in Table III. IV, and V. It can be observed that, as

TABLE V
$\mathrm{DW_{min}}$ PARAMETERS FOR THE MEASUREMENTS AT MACHINE SHOP.

	2.245 GHz		5.4 GHz	
	10 dB	20 dB	10 dB	20 dB
5 m	300 ns	460 ns	330 ns	470 ns
10 m	480 ns	490 ns	350 ns	500 ns
20 m	450 ns	650 ns	450 ns	600 ns

expected, $\mathrm{DW_{min}}$ for 20 dB self-SIR threshold is longer than that with 10 dB self-SIR threshold, which is determined by the definition of self-SIR in (7). A more important insight is that in the industrial sites with rich reflections, the $\mathrm{DW_{min}}$ increases with communication distance. Considering the effect of the carrier frequency, the $\mathrm{DW_{min}}$ at 5.4 GHz is generally longer than that at 2.245 GHz at the same distance and with the same self-SIR threshold.

B. Delay Window under Reliability Requirement

Besides low latency, high reliability is the other strict requirement in industrial communications, which is usually quantified by PER. PER is related with modulation, coding rate, and SNR value. SNR is in turn affected by the transmit power, bandwidth, and path loss of the channel, and the self-SIR studied in the previous section. As PER acts as a constraint (5f) in CCDO, and is affected by the above mentioned parameters, it is best to determine the available range of the parameters to satisfy the PER constraint first. In the following, we take convolutional code as an example to show a theoretical way to determine the required self-SIR value and $\mathrm{DW_{min}}$, which can be easily extended to other codes. The convolutional code is suitable for short packets and is capable to start decoding before all the bits are collected, which avoids long decoding latency. Note that the derived result should be updated periodically due to the dynamic change of wireless propagation environment with moving reflectors and obstacles. This procedure can also be done by simulations given the large amount of channel impulse responses.

The PER of convolutional code is bounded by [20]

$$\mathrm{PER} < 1 - (1 - P_E)^{L_c}, \tag{8}$$

where L_c is the coded packet length in bits $L_c = \lceil L/R_c \rceil$, while P_E is the first-event error probability, which can be calculate given the modulation order, received SNR value, and the decoding algorithm.

In order to provide an example, we consider the convolutional code used by IEEE 802.11 a/g, which has a 6-stage shift register with generator polynomials q0 = 1011011, and q1 = 1111001 with rate 1/2. Punctured convolutional codes with higher rates (2/3, 3/4, 5/6) can be constructed by periodic puncturing of specific code symbol from the output of the 1/2 code [21]. Four modulation orders BPSK, QPSK, 8PSK and 16QAM are used, and hard decision demodulation is used. Given a PER threshold, the minimum required SNR for each modulation and coding rate can be calculate. The results are plotted in Fig. 1 for PER thresholds of 10^{-6} and 10^{-9}, according to the PER requirement in Table I.

In the SNR computation, we must take into account the impact of multi-path fading. Indeed, when setting the CP length equal to the $\mathrm{DW_{min}}$ with a given self-SIR threshold, the multi-path components that arrive outside the CP duration will cause interference. We hence define signal to echo and noise ratio (SENR) as

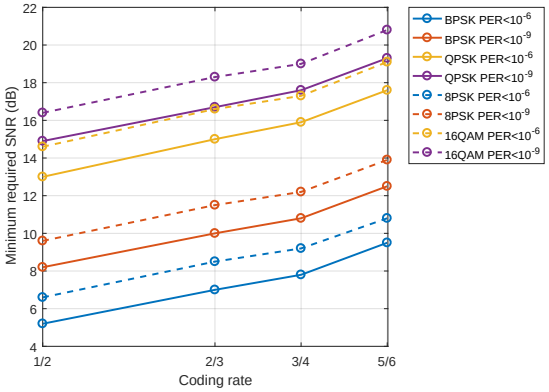

Fig. 1. Minimum required SNR of different modulation and coding rates for PER$< 10^{-6}$ and PER$< 10^{-9}$ with L = 100 bits packets.

$$\mathrm{SENR} = \frac{P_t \cdot \mathrm{PL} \cdot \frac{\text{self-SIR}}{1+\text{self-SIR}}}{P_t \cdot \mathrm{PL} \cdot \frac{1}{1+\text{self-SIR}} + \mathrm{AWGN}}, \tag{9}$$

where $P_t \cdot \mathrm{PL}$ represents the received power, self-SIR$/(1 +$ self-SIR$)$ and $1/(1 +$ self-SIR$)$ split the total received power into usable signal power and self-interference power according to the definition of self-SIR. SENR value is affected by path loss gain PL, self-SIR, additive white Gaussian noise (AWGN) and transmission power P_t. PL is mainly dependent on the communication distance and can be computed according to theoretical models such as free-space path loss equation. It can also be determined by site-specific measurements [11]. Finally, thermal noise AWGN is determined by the receiver noise figure and the bandwidth. We should note that SENR does not equal to SNR naturally, but can be considered as an accurate approximation. For a quantitative estimation of the effect of self-interference, it is assumed that its effect is simply additive to the AWGN. Hence setting SENR equal to the minimum required SNR to achieve a certain PER for a given combination of coding and modulation scheme from Fig. 1, the minimum required self-SIR can be calculated as

$$\text{self-SIR} = \frac{(\mathrm{AWGN} + P_t\,\mathrm{PL}) \cdot \mathrm{SNR}}{P_t\,\mathrm{PL} - \mathrm{AWGN} \cdot \mathrm{SNR}}. \tag{10}$$

Note that self-SIR a positive number. This requires the denominator of the righthand side of (10) to be positive, which sets a constraint to the choice of P_t, PL (related with distance), AWGN (determined by the bandwidth), and SNR (which varies among different modulation and coding rate).

We provide an example to show how to calculate the minimum required self-SIR value for different coding and modulation schemes. We set P_t to be 20 dBm, which is the maximum effective isotropic radiated power (EIRP) allowed at 5.4 GHz. Two PL values, -85 dB and -100 dB are chosen, corresponding to distances of 50 m and 100 m respectively, from the reported results in [11]. The AWGN power changes according to the bandwidth, which takes values from 5 MHz

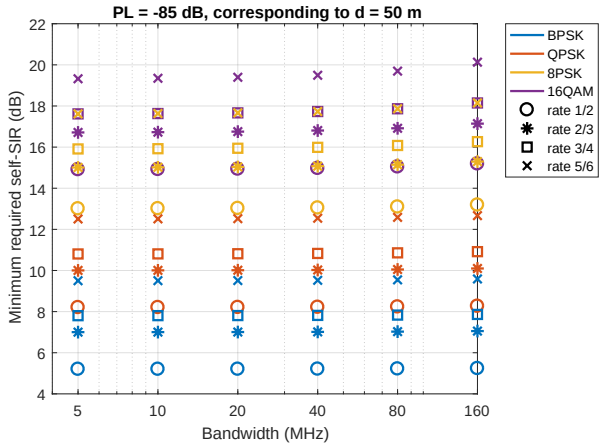

Fig. 2. Minimum required self-SIR of different modulation and coding rates for PER$< 10^{-6}$ with L = 100 bits packets, 20 dBm transmission power, and -85 dB path loss corresponding to 50 m communication distance.

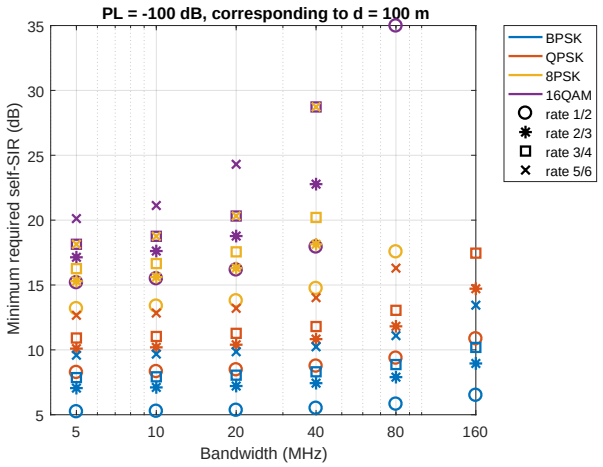

Fig. 3. Minimum required self-SIR of different modulation and coding rates for PER$< 10^{-6}$ with L = 100 bits packets, 20 dBm transmission power, and -100 dB path loss corresponding to 100 m communication distance.

to 160 MHz. PER requirement is set to be smaller than 10^{-6}. The results are shown in Fig. 2 and Fig. 3. From the figures, the larger the coding rate and modulation order, the larger the minimum required self-SIR, which is consistent with (10), since the minimum required SNR is larger with larger coding rate and modulation order. In Fig. 2, the minimum required self-SIR increases only marginally with the bandwidth. This is because when PL=-85 dB, the term $P_t\,\mathrm{PL}/\mathrm{SNR}$ is much larger than AWGN, thus the change of AWGN has small effect. Hence these results are valid for the interference-limited regime, rather than the noise-limited regime. With PL=-100 dB in Fig. 3, instead, the above two terms are comparable, thus the minimum required self-SIR increases clearly with bandwidth. The increasing slope is larger with higher modulation and coding rate, as the corresponding SNR is larger. Note that for quadrature phase shift keying (QPSK), eight phase shift keying (8PSK) and 16-quadrature amplitude modulation (16QAM) in Fig. 3, there is a limit to the maximum bandwidth with different coding rate with the given value of P_t and PL, which is due to positive constraint on (10).

C. Solution of the CCDO

The key idea of CCDO is to quantify the effect of DW duration on SNR and PER, and determine the $\mathrm{DW}_{\min}$ under given PER requirement. From the study of the previous two subsections, given a PER requirement and payload length, the minimum required self-SIR together with the corresponding $\mathrm{DW}_{\min}$ varies with coding, modulation, central frequency, distance and different sites. In the OFDM systems, the optimal T_{CP} can be set to $\mathrm{DW}_{\min}$ to derive the optimal packet transmission time while guaranteeing the reliability requirement.

In the optimization problem (5), constraint (5f) explicitly sets the PER requirement $\mathrm{PER}_{\mathrm{req}}$, which is an implicit function of T_{CP} as explained in the previous subsections. After the minimum required $T_{\mathrm{CP}} = \mathrm{DW}_{\min}$ is determined by (5f), it will be used to update T_{CP} in (5a) and $T_{\mathrm{pkt}}(T_{\mathrm{CP}}, N_{\mathrm{FFT}})$. Then for each combination set of L, B, M, R_c, and N_{ss}, the corresponding minimized T_{pkt} can be calculated by numerically finding the optimal combination of N_{FFT} and $N_{\mathrm{sym}}^{\mathrm{data}}$. The solution can be easily obtained off-line by standard numerical methods.

The possible applications of CCDO are manifold. First, it can act as a basic reference methodology for architectural design of new industrial wireless techniques or standards that want to take into account the multi-path effect in the industrial sites. Second, it can be used for pre-deployment assessment, to roughly estimate the involved parameter values and further design the scheduling parameters. Third, it can be used for run-time optimization, assuming that periodic monitoring is employed to measure the dynamic change of the wireless channel.

IV. Performance Analysis of CCDO

In this section, the advantage of using CCDO to optimize the packet transmission time is shown through numerical results. The results are based on the measurements carried out at AAplant as an example. A similar optimization can be applied to the other scenarios as well as to any site for which channel measurements are available. Moreover, in real operation processes, the measurement should be done periodically, and the CP length in CCDO should be tuned according to the updated results.

A. Optimal packet transmission time by CCDO

With 100 bit raw payload size, reliability requirement of PER$< 10^{-6}$, and communication link distance below 50 m and 100 m, the minimum required self-SIR corresponding to different modulation and coding schemes with bandwidth from 5 MHz to 160 MHz can be found in Fig. 2 and Fig. 3. For the sake of simplicity, self-SIR is set to be 10 dB if the minimum required self-SIR is less than 10 dB, and to be 20 dB if the minimum required self-SIR is larger than 10 dB but smaller than 20 dB. The corresponding $\mathrm{DW}_{\min}$ values for different frequency bands and communication distances (50 and 100 m) are listed in Table III. Then, T_{CP} is set equal to $\mathrm{DW}_{\min}$ and the optimization problem (5) is run. It is worth stressing that now

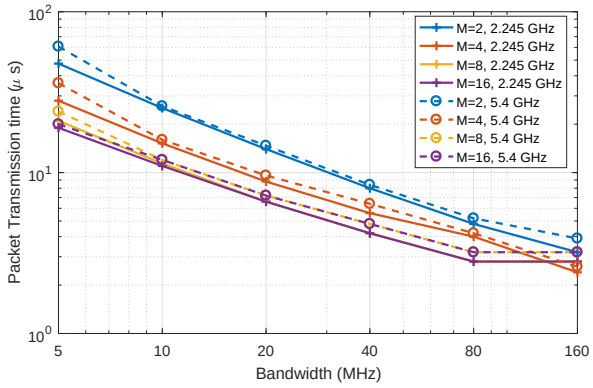

Fig. 4. Packet transmission time with CCDO in the 2.245/5.4 GHz bands for different modulation orders and bandwidth with L = 100 bits packets, PER$< 10^{-6}$.

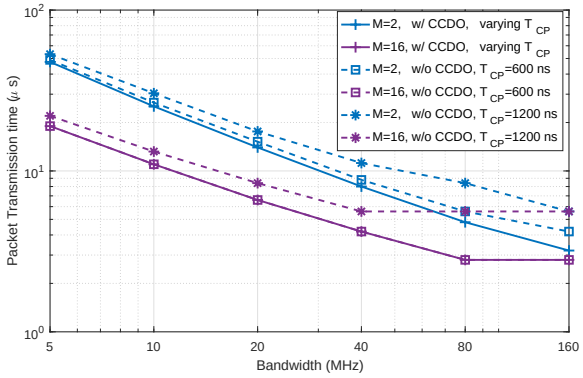

Fig. 5. Packet transmission time before and after CCDO in the 2.245 GHz band for BPSK and 16QAM and different bandwidth with L = 100 bits packets, PER$< 10^{-6}$ and 50 m communication distance.

T_{CP} changes with the modulation, coding, central frequency, and distance to ensure that PER$< 10^{-6}$ is satisfied.

When considering a communication distance of 50 m, the packet transmission times with varying bandwidth using different modulation and carrier frequencies are plotted in Fig. 4. For each modulation, the transmission times with coding rates 1/2, 2/3, 3/4 and 5/6 are all calculated and the minimum value is selected for the modulation. As T_{s} is the inverse of bandwidth, the packet transmission time T_{pkt} decreases when increasing the bandwidth by (4). T_{pkt} at 2.245 GHz is smaller than that at 5.4 GHz with the same modulation and bandwidth, as T_{CP} is smaller at 2.245 GHz.

When the bandwidth is small, T_{pkt} decreases significantly with bandwidth for each modulation, and higher modulations result in lower T_{pkt}. This is because $N_{\text{sym}}^{\text{data}}$ is reciprocally proportional to both M and the N_{dsc} (related with N_{FFT}). As N_{FFT} is upper bounded by (5b), when B increases, N_{dsc} increases together with N_{FFT}, and $N_{\text{sym}}^{\text{data}}$ decreases. When the bandwidth increases from 80 MHz to 160 MHz, instead, the decrease of T_{pkt} is limited for low order modulations and even null for high order modulations. This is due to the fact that with very short payload size ($L = 100$) and abundant bandwidth (80 MHz, 160 MHz), $N_{\text{sym}}^{\text{data}}$ = 1 for the high order modulations (8PSK, 16QAM), and $T_{\text{CP}} = \text{DW}_{\min}$ does not change from 80 MHz to 160 MHz. In this case, hence, the variables affecting T_{pkt} are only N_{FFT} and T_{s} from (4). By (5a), the lower bound of N_{FFT} is proportional to $B \cdot \text{DW}_{\min}$, thus $T_s \cdot N_{\text{FFT}}$, and T_{pkt} stay constant with abundant bandwidth and short payload size. In this case T_{CP} becomes the limiting factor to minimize T_{pkt} for the high order modulations.

B. Comparison before and after CCDO

In this subsection, we compare the T_{pkt} achieved by CCDO (5) and the optimization in [12] without reliability constraint or optimal T_{CP}. Unified T_{CP} = 600 ns and 1200 ns are selected for the optimization in [12], while T_{CP} for CCDO is selected in the same way as in Fig. 4, considering 50 m distance at AAplant. We compare the results at 2.245 GHz to which the results are similar at 5.4 GHz. Only the curves of M=2 (BPSK) and M=16 (16QAM) are plotted in Fig. 5 for better visibility. For BPSK, the optimal T_{CP} value with CCDO is 400 ns from 5 MHz to 160 MHz, thus the unified T_{CP} values of 600 ns and 1200 ns result in longer packet transmission time. For 16QAM, the optimal T_{CP} value with CCDO is 600 ns from 5 MHz to 160 MHz, thus unified T_{CP} 600 ns will not increase the transmission time, while 1200 ns T_{CP} will increase it. Though the increase using 1200 ns T_{CP} accounts for 15% at 5 MHz bandwidth compared to the optimal T_{CP} value in CCDO, the increase reaches 100% when the bandwidth is 160 MHz, which is detrimental for time critical applications.

Without CCDO, which includes T_{CP} as a variable to satisfy PER requirement, the randomly selected T_{CP} might either cause longer T_{pkt} with a T_{CP} longer than $\text{DW}_{\min}$, or violate the PER requirement by using a shorter T_{CP}.

C. Comparison with other protocols

In this subsection, we compare the physical layer parameters and the minimum achieved T_{pkt} for 100 bit payload among different wireless protocols and WirelessHP with CCDO in Table VI. IEEE 802.15.1 is the base for WISA [16], and it has fixed slot duration of 625 μs [22]. Though WISA reduces the slot duration, it will still take more than 100 μs since the data rate is 1 MHz. IEEE 802.15.4 is the base for ISA 100.11.a [15] and wirelessHART [14]. As direct sequence spread spectrum (DSSS) is used, the maximum data rate is only 250 kbps with 5 MHz bandwidth, and it takes 400 μs to transmit the payload and 256 μs for preamble and header [23]. As the base for WIA-FA [17], IEEE 802.11g can support multiple bandwidths up to 20 MHz, but FFT size is fixed at 64 [21]. With 20 MHz bandwidth, the maximum data rate is 54 MHz with 64-quadrature amplitude modulation (64QAM) and convolutional coding with 3/4 coding rate. The corresponding T_{pkt} for 100 bit payload is 24 μs including 4 μs for payload and 20 μs for preamble. WirelessHP also uses OFDM, but the FFT size is flexible and can adjust by the payload size. Moreover, WirelessHP greatly reduces the preamble to just one symbol, and the preamble symbol has the same FFT size as the data symbol. Taking the same modulation, bandwidth, and CP duration with IEEE 802.11g, the minimized T_{pkt} of WirelessHP is 6.4 μs with the optimal FFT size being 16,

TABLE VI
PHY PERFORMANCE BETWEEN WIRELESSHP CCDO AND OTHER PROTOCOLS.

	IEEE 802.15.1	IEEE 802.15.4	IEEE 802.11.g	WirelessHP	
Bandwidth	1 MHz	5 MHz	20 MHz	5 MHz	20 MHz
Data rate	1 Mbps	250 kbps	54 Mbps	5.83 Mbps	22.5 Mbps
Modulation	GFSK	DSSS OQPSK	OFDM 64QAM 3/4 coding rate	OFDM QPSK	OFDM 64QAM 3/4 coding rate
T_{pkt} (μs)	625	400 + 256	4 + 20	21.6 + 7.2	4.8 + 1.6

which is a tradeoff between preamble symbol duration and the data symbol duration. The data rate of WirelessHP follows the way of IEEE 802.11g, which is calculated as the payload bits in the data sub-carriers divided by the symbol duration.

The above data rate calculation does not consider the preamble duration, which is alright when payload size is much longer than the preamble size. However, with a very short payload size like 100 bits, the preamble duration (20 μs in IEEE 802.11g) can be much longer than the data symbol duration (4 μs). In this case, the data rate should be modified by payload size (100 bits) divided by packet transmission time, thus T_{pkt} is a more accurate index to assess the latency performance. Comparing the parameters of the protocols in Table VI, in order to achieve low T_{pkt}, high data rate is necessary but is not a guarantee. Flexible and efficient packet structure to consider the payload size and preamble duration is also necessary. In this way, WirelessHP with CCDO is superior than the other protocols to guarantee short T_{pkt} for short payload size.

D. Relay VS Direct transmission

Point-to-point transmission is intuitively considered to have lower latency compared to that with relays, which usually imply long queuing and forwarding delays. However, in typical industrial applications, the queuing delay at the relay nodes can be neglected by considering deterministic traffic patterns and optimal scheduling. Assuming the relay nodes work in full-duplex and that a decode-and-forward approach is used, the forwarding delay will be reduced to the time needed to collect signals before it can start to decode, and the processing time for decoding and re-encoding.

Some recent papers argue that a relay configuration has potential to achieve smaller end-to-end latency, because given a end-to-end reliability requirement, relays could provide a higher equivalent SNR value at the receiver, and further a higher transmission rate [24], [25]. However, the above analyses were based on AWGN channel and did not consider the effect caused by the delay spread to delay and reliability.

In order to get a deeper understanding of the impact of relays in realistic multi-path channels, we compare here the minimum required T_{pkt} by CCDO at 50 m and 100 m distances under the requirement of PER$< 10^{-6}$. The according self-SIR for each modulation and coding at the two distances can be found in Fig. 2 and Fig. 3, and the according T_{CP} can be chosen from Table III. The minimized T_{pkt} at 2.245 GHz with 50 m and 100 m distances is plotted in Fig. 6. The results for 5.8 GHz is similar with 2.245 GHz. From Fig. 6, it turns out that T_{pkt} at the 100 m distance is longer than that obtained at the 50 m distance, as in the latter case higher modulation and

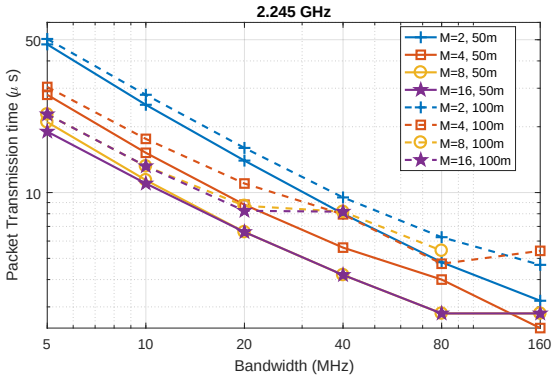

Fig. 6. Packet transmission time with CCDO for different modulation orders, bandwidth and communication distances (50 and 100 m), with L = 100 bits packets, PER$< 10^{-6}$ and 2.245 GHz center frequency.

coding can be used and T_{CP} is also shorter. There is a rise for T_{pkt} of QPSK at the 100 m distance from 80 MHz to 160 MHz, which is due to the fact that the maximum coding rate at 80 MHz is 5/6 while it is 3/4 at 160 MHz, causing an increase of $N_{\text{sym}}^{\text{data}}$ by (3). T_{pkt} of 8-PSK and 16-QAM at 160 MHz are not calculated, as the parameters cannot satisfy PER$< 10^{-6}$ with these two modulations at 160 MHz according to Fig. 3.

Finally, when the bandwidth is 40 MHz, T_{pkt} of 8-PSK and 16-QAM with 100 m communication distance is almost doubled with respect to the T_{pkt} with 50 m. This also holds for 8-PSK at 80 MHz, and QPSK at 160 MHz. Indeed, from (6), the packet transmission time is lower bounded by $T_{\text{pkt}} \geq 2T_{\text{CP}} \cdot (N_{\text{sym}}^{\text{pre}} + N_{\text{sym}}^{\text{data}})$, and the lower bound is approached with abundant bandwidth. As the upper bound and lower bound of N_{FFT} are both proportional to the bandwidth B from (5a) and (5b), when increasing B, N_{FFT} as well as N_{dsc} increase. With the increase of N_{dsc} along with B, $N_{\text{sym}}^{\text{data}}$ decreases down to 1 from (3), and T_{pkt} is proportional to T_{CP} under this case. As the T_{CP} with 10 dB and 20 dB self-SIR at 100 m is nearly double of the corresponding T_{CP} at 50 m as shown in Table III, the according T_{pkt} also doubles when the bandwidth is abundant for each modulation.

The observation that packet transmission time with 50 m distance is half the one with 100 m distance motivates the conclusion that using relay may be an effective way to reduce the end-to-end delay in the industrial sites, which have rich reflections. Indeed, besides the merit of using higher modulation order and code rate, short communication distance also results in short required T_{CP}, which is the limiting factor for reducing the transmission time when transmission bandwidth is abundant in the OFDM system. For other modulations that do not use the CP-based equalization, the DW length can be seen as a reflection about the distortion level of the signal. The

more severe the signal's distortion, the more advanced equalization algorithms should be used, which implies longer delay for the equalization processing. From the results obtained with CCDO, the minimum achievable latency is not constant at different distances, due to the fact that the optimal CP length varies with the distance. The shorter delay window at short communication ranges (lower than 50 m at AAplant) enables shorter transmission time compared to longer distances, thus it is helpful to add relays to make the distance of each hop within the range of short delay window length, especially when the direct path has a high probability to be blocked by moving obstacles or reflectors.

V. Conclusion and Future Work

An analysis of two crucial reliability parameters, namely self-SIR and delay window, was performed based on channel measurement at different sites, distances, and central frequencies. The results motivate us to propose a physical layer optimization method, CCDO, to minimize OFDM packet transmission time under reliability constraints. The key feature of CCDO is to determine the minimum delay window duration to guarantee the reliability requirement on the message transmission, and use this duration as cyclic prefix (CP) length for the optimization of the packet transmission time. Under a given reliability requirement, the minimum required delay window duration increases with communication distance in three representative industrial sites. For short packets with abundant transmission bandwidth (e.g., from 80 MHz for 100 bit payload), the CP duration is found to be the restricting factor to reduce packet transmission time. The results motivate the use of relay nodes in the industrial environment to limit the communication distance range and obtain a lower packet transmission time, the reduction of which is greatly due to that delay window duration is different at different distance ranges.

In this paper, the results are based on the channel measurement at 2.245 GHz and 5.4 GHz with single antennas at both transmitter and receiver side. It will be interesting to investigate whether the relation between the delay window and distance holds in the millimeter frequency band and with multiple antennas, which we will consider for future work. In such a case, to compare the delay for direct transmission and that with relay, a more complete performance comparison of direct and relay configurations will be needed.

References

[1] Z. Pang, M. Luvisotto, and D. Dzung, "High performance wireless communications for critical control applications," *IEEE Industrial Electronics Magazine*, 2017.

[2] Y. Liu, R. Candell, and N. Moayeri, "Effects of wireless packet loss in industrial process control systems," *ISA transactions*, vol. 68, pp. 412–424, 2017.

[3] D. Orfanus, R. Indergaard, G. Prytz, and T. Wien, "Ethercat-based platform for distributed control in high-performance industrial applications," in *Emerging Technologies & Factory Automation (ETFA), 2013 IEEE 18th Conference on*. IEEE, 2013, pp. 1–8.

[4] S. Feliciano, H. Sarmento, and J. Z. de Oliverira, "Field area network in a mv/lv substation: A technical and economical analysis," in *Intelligent Energy and Power Systems (IEPS), 2014 IEEE International Conference on*. IEEE, 2014, pp. 192–197.

[5] J. F. Coll, J. Chilo, and B. Slimane, "Radio-frequency electromagnetic characterization in factory infrastructures," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 3, pp. 708–711, 2012.

[6] D. P. Gaillot, E. Tanghe, W. Joseph, P. Laly, B. Hanssens, M. Liénard, and L. Martens, "Polarization properties of specular and dense multipath components in a large industrial hall," in *General Assembly and Scientific Symposium (URSI GASS), 2014 XXXIth URSI*. IEEE, 2014, pp. 1–4.

[7] J. Karedal, S. Wyne, P. Almers, F. Tufvesson, and A. F. Molisch, "A measurement-based statistical model for industrial ultra-wideband channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 3028–3037, 2007.

[8] Y. Ai, J. B. Andersen, and M. Cheffena, "Path-loss prediction for an industrial indoor environment based on room electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 7, pp. 3664–3674, 2017.

[9] J.-E. Berg, J. Ruprecht, J.-P. de Weck, and A. Mattsson, "Specular reflections from high-rise buildings in 900 mhz cellular systems," in *Vehicular Technology Conference, 1991. Gateway to the Future Technology in Motion., 41st IEEE*. IEEE, 1991, pp. 594–599.

[10] B. Holfeld, D. Wieruch, L. Raschkowski, T. Wirth, C. Pallasch, W. Herfs, and C. Brecher, "Radio channel characterization at 5.85 ghz for wireless m2m communication of industrial robots," in *Wireless Communications and Networking Conference (WCNC), 2016 IEEE*. IEEE, 2016, pp. 1–7.

[11] R. Candell, C. A. Remley, J. T. Quimby, D. R. Novotny, A. E. Curtin, P. B. Papazian, G. H. Koepke, J. E. Diener, and M. T. Hany, "Industrial wireless systems: Radio propagation measurements," *Technical Note (NIST TN)-1951*, 2017.

[12] M. Luvisotto, Z. Pang, D. Dzung, M. Zhan, and X. Jiang, "Physical layer design of high-performance wireless transmission for critical control applications," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 2844–2854, 2017.

[13] M. Luvisotto, Z. Pang, and D. Dzung, "Ultra High Performance Wireless Control for Critical Applications: Challenges and Directions," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1448–1459, 2017.

[14] *HART Field Communication Protocol Specification, Revision 7.0*, HART Communication Foundation Std., 2007. [Online]. Available: http://www.hartcomm.org/

[15] *ISA-100.11a Wireless Systems for Industrial Automation: Process Control and Related Applications*, International Society of Automation (ISA) Std., 2009.

[16] G. Scheible, D. Dzung, J. Endresen, and J. E. Frey, "Unplugged but connected [design and implementation of a truly wireless real-time sensor/actuator interface]," *IEEE Industrial Electronics Magazine*, vol. 1, no. 2, pp. 25–34, 2007.

[17] *PAS 62948: Industrial networks - Wireless communication network and communication profiles - WIA-FA*, International Electrotechnical Commission (IEC) Std., 2015.

[18] D. R. Novotny, A. E. Curtin, C. A. Remley, P. B. Papazian, J. T. Quimby, and R. Candell, "A tetherless, absolute-time channel sounder, processing, and results for a complex environment," *Proceedings of the Antenna Measurement Techniques Association, Austin, TX, United States*, 2016.

[19] R. Candell, "Radio frequency measurements for selected manufacturing and industrial environments," *Field Measurements of Wireless Systems for Industrial Environments*, 2016.

[20] M. Pursley and D. Taipale, "Error probabilities for spread-spectrum packet radio with convolutional codes and viterbi decoding," *IEEE Transactions on communications*, vol. 35, no. 1, pp. 1–12, 1987.

[21] IEEE 802.11 Working Group, "IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11*, 2016.

[22] IEEE 802.15 Working Group, "IEEE Standard for Information technology-Local and metropolitan area networks-Specific requirements-Part 15.1a: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications for Wireless Personal Area Networks (WPAN)," *IEEE Std 802.15.1-2005 (Revision of IEEE Std 802.15.1-2002)*, 2005.

[23] ——, "IEEE Standard for Information technology-Local and metropolitan area networks-Specific requirements-Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPANs)," *IEEE Std 802.15.4-2006 (Revision of IEEE Std 802.15.4-2003)*, 2006.

[24] I. Marić, "Low latency communications," in *Information Theory and Applications Workshop (ITA), 2013*. IEEE, 2013, pp. 1–6.

[25] A. Chaaban and A. Sezgin, "Multi-hop relaying: An end-to-end delay analysis," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2552–2561, 2016.