IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 12 August 2022, accepted 26 August 2022, date of publication 8 September 2022, date of current version 19 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3205009

RESEARCH ARTICLE

M-Ary Aggregate Spread Pulse Modulation for Robust and Scalable Low-Power Wireless Networks

ALEXEI V. NIKITIN[1], (Member, IEEE), AND RUSLAN L. DAVIDCHACK[2]
[1]Nonlinear LLC, Wamego, KS 66457, USA
[2]School of Computing and Mathematical Sciences, University of Leicester, Leicester LE1 7RH, U.K.

Corresponding author: Alexei V. Nikitin (avn@nonlineacorp.com)

This work was supported in part by Pizzi Inc., Denton, TX 76205, USA.

ABSTRACT In this paper we present an overview of the M-ary Aggregate Spread Pulse Modulation (M-ASPM), and provide an assessment of its suitability and advantages for use in low-power wide-area networks (LPWANs). Notably, M-ASPM combines high energy-per-bit efficiency, robustness, resistance to interference, and a number of other favorable technical characteristics, with the spread-spectrum ability to maintain the network capacity while extending its range. We quantify the impact of mutual interference of multiple M-ASPM transmitters and demonstrate how such capacity-preserving range extension can be achieved for numerous desired areal distributions of the uplink nodes. Throughout the paper, LoRa is used for benchmark comparison and quantification of various M-ASPM features. In particular, we show that, while sharing many essential properties with LoRa, M-ASPM provides far more effective network range extension. When used in the same manner as LoRa, M-ASPM can serve as an appealing LoRa alternative. In addition, while being different LPWAN solutions, M-ASPM and LoRa can be designed to concurrently operate in the same spectral band and geographical area, cooperatively complementing each other's coverage.

INDEX TERMS Aggregate spread pulse modulation (ASPM), intermittently nonlinear filtering (INF), Internet of Things (IoT), LoRa, low-power wide-area network (LPWAN), M-ary ASPM (M-ASPM), physical layer (PHY), spread spectrum, time-bandwidth product (TBP).

I. INTRODUCTION

By providing long range wireless access to the Internet of Things (IoT), Low-Power Wide Area Network (LPWAN) technologies have been one of the main drivers of the IoT expansion. The common feature of different LPWANs is a long range with relatively low throughput. However, an LPWAN is not a uniquely defined solution for a fixed set of requirements, but has a large space of characteristics that require tradeoffs and optimizations [1], [2].

For example, when extending the range of a wireless network, it may be desirable to trade the energy efficiency of a single link for the number of transmitting nodes that can coexist and concurrently operate in a given spectral band (say, with the total bandwidth B) in the extended range. Conceptually, perhaps the simplest way to achieve this tradeoff under a transmit power constraint is to allocate separate, narrower spectral sub-bands to different transmitters, with the bandwidths ΔB_i such that $\sum_i \Delta B_i = B$. In practice, it can be accomplished by various frequency-division multiple access (FDMA) schemes, for example, by the single-carrier FDMA (SC-FDMA) used in the Narrowband Internet of Things (NB-IoT) [3].

For a signal with a given power, reduction in its bandwidth from B to ΔB increases its power spectral density (PSD) as $\mathrm{PSD} \propto B/\Delta B$. Then, if the PSD of the noise is constant across the spectral band and the signal attenuation is frequency-independent, and for a given modulation and power-law path loss, the range d of a link can be related to its bandwidth ΔB as $d^\gamma \propto 1/\Delta B$, where γ is the path-loss exponent [4]. Thus the number of available sub-bands increases with range as d^γ. On the other hand, for the same-size payloads and

The associate editor coordinating the review of this manuscript and approving it for publication was Stefano Scanzio.

transmission rates, the number of non-interfering transmitters that can be deployed in a given sub-band is proportional to ΔB. Therefore, at best, when extending the distance between the transmitters and the receiver, one can only maintain (but not increase) the total number of the transmitting nodes that can be placed at the given range d, at the penalty of the energy consumption per node increasing with range as d^γ.

The effectiveness of the FDMA-based approaches to managing the network range and capacity largely depends on the validity of the relation $d_i^\gamma \propto 1/\Delta B_i$ for any i-th sub-band. However, signals in different sub-bands may be affected very differently by the propagation conditions, e.g. delay and Doppler spreads, and the assumption of a constant noise PSD would hardly hold for unlicensed spectral bands. Thus, alternatively, we may want to use the full spectral band B for all nodes as a common shared resource, and, instead of changing the bandwidth, achieve the desired range of a link by changing the spectral efficiency of a modulation with a given energy-per-bit efficiency. Say, we can use an increase in the processing gain $B/\Delta B$ provided by a spread spectrum (SS) technique [5] to extend the range of a link. For example, for the code-division multiple access (CDMA) with orthogonal codes, it will lead to the same relation between the range and the number of nodes as the FDMA-based approaches. That is, while increasing the distance between the transmitters and the receiver, we can maintain the total number of the transmitting nodes that can be placed at the given range, with the same penalty on the energy efficiency.

Short for "Long Range," LoRa [6], [7] is a modulation technique derived from chirp spread spectrum (CSS), that has recently become one of the prevailing technologies in LPWANs for the IoT applications. In LoRa, the main parameter of the modulation is the spreading factor (SF), which can range from SF $= 6$ through SF $= 12$, and is defined as the number of bits carried by the chirp waveforms. For a given spectral band, the SF channels can be considered (nearly) orthogonal. A single increment in the SF doubles the time duration of the chirp (thus reducing the spectral efficiency) while, by encoding more bits per chirp, also incrementing the energy-per-bit efficiency. This results in a longer range for the channels with larger SFs.

Thus, in LoRa the extension of the range is accomplished by a decrease in the spectral efficiency as well as an increment in the energy-per-bit efficiency. At first glance, it enables increase in both, the range (defined, for example, as the mean distance of the transmitters from the receiver, weighted by their payloads) and the capacity of the network. However, in LoRa the energy-per-bit efficiency increments are insignificant in comparison with the spectral efficiency reductions, and any decrease in the spectral efficiency is not accompanied by the respective increment in the number of available orthogonal channels operating with this spectral efficiency. In fact, the opposite is true: In LoRa, for a given spectral band, the number of the SF channels that can be deployed beyond a given range is a (stepwise) decreasing function of this range.

For example, in a single-gateway LoRa network all seven SF channels can be used in the range below that for SF $= 6$. However, the range between SF $= 10$ and SF $= 11$ can be served by only two channels, and only one SF channel (with SF $= 12$) can be employed in the range above SF $= 11$. At the same time, a single increment in the SF approximately doubles the time-on-air (ToA) of a given payload, proportionally increasing the energy consumption and reducing the number of nodes. As a result, the number of LoRa uplink nodes that can be placed at a given distance from the gateway is a rapidly decreasing (stepwise) function of the distance, with the overall decrease noticeably faster than the path attenuation (e.g., faster than $1/d^\gamma$ for power-law path loss).

Nevertheless, due to its high energy-per-bit efficiency, combined with robustness, resistance to interference, and a number of favorable technical characteristics (e.g., LoRa is a constant-envelope modulation), LoRa retains a strong appeal for various LPWANs. For example, by increasing the SF from 6 to 12, LoRa offers approximately sevenfold free-space range extension. While, for a single gateway, more than half on the total network capacity is confined to the range for SF $= 7$, and only less than 1.5% of the capacity is provided by the nodes with SF $= 12$, such a substantial range increase can benefit a large number of applications that do not require high areal densities of the end nodes at long ranges. At the same time, the contribution of the large-SF nodes to the total network capacity is insignificant. Therefore, when a long-distance areal coverage is desired, the effective range of LoRa coverage becomes limited to that of the relatively small SFs, e.g., 6 to 8, eroding the long-range benefits of large SFs and restricting LoRaWAN's scalability.

One approach to addressing this LoRa scalability limitation is to design a physical layer (PHY) modulation scheme that retains most, if not all, LoRa's advantages, including in the energy consumption and robustness, while providing the ability to better sustain the network capacity when extending its range. The M-ary Aggregate Spread Pulse Modulation (M-ASPM) [8] is an example of such modulation. In this paper, we present an overview of the M-ASPM, and provide an assessment of its suitability and advantages for use in LPWANs. Throughout, we use LoRa as a benchmark for comparison and quantification of various M-ASPM features.

In Section II, we describe the M-ASPM's physical layer, and briefly review its properties in an additive white Gaussian noise (AWGN) channel. Noticeably, for a single-sideband M-ASPM with constant magnitude pulses and $\log_2 M$ bits per waveform, the uncoded AWGN energy per bit efficiency is the same as of the LoRa modulation with the spreading factor SF $= \log_2 M$, for both coherent and noncoherent detection.

In Section III, we assess the overall suitability of using M-ASPM in LPWANs, and discuss the relationship between the M-ASPM's range and spectral efficiency. In particular, the maximum spectral efficiency of M-ASPM equals that of LoRa with SF $= \log_2 M$ for noncoherent detection, and exceeds that of LoRa by a factor of 4 for coherent detection. At the same time, the M-ASPM is a "true" spread spectrum

technique, and its spectral efficiency is inversely proportional to the average interpulse interval used in M-ASPM. Consequently, the M-ASPM's receiver sensitivity is proportional to this interval, and the desired physical range can be achieved for any value of M. In contrast, LoRa always operates at maximum spectral efficiency for a given SF, with the range determined by the specific SF value.

In Section IV, we discuss the use of multiple pulse shaping filters (PSFs) in M-ASPM, in a manner akin to asynchronous CDMA, and assess the impact of inter-PSF collisions. As the signal-to-interference ratio (SIR) margin of an M-ASPM signal is proportional to its average interpulse interval, and thus increases with the distance between the transmitter and the receiver (e.g., as d^{γ} for power-law path loss), multiple transmitters with different PSFs can be deployed, at sufficiently long ranges, with insignificant impact of the mutual interference. Notably, we provide a qualitative assessment of this impact.

Consequently, in Section V we address the task of sustaining M-ASPM's network capacity when extending its range. We show that the increase in the M-ASPM SIR margins with range allows us, by employing a larger number of the PSF channels, to maintain the number of equal-payload nodes at long ranges, in a manner similar to the FDMA and CDMA-based approaches. Further, this enables achieving various desired areal distributions of the end nodes without unduly sacrificing the total capacity of the M-ASPM network.

Thus, unlike LoRa, M-ASPM has the ability to sustain the network capacity when extending its range, and to accommodate various desired areal distributions of the end nodes with minimal impact on the total number of nodes. At the same time, LoRa may have an edge in a number of technical characteristics that are important in a variety of particular LPWAN use cases. Favorably, however, LoRa and M-ASPM can be designed to concurrently operate in the same spectral band and geographical area, cooperatively complementing each other's coverage. This is discussed and illustrated in Section VI, followed by the conclusion and the outline of further research directions in Section VII.

The description of M-ary ASPM was previously provided only in [8], where we evaluate the bit error probability for coherent and noncoherent M-ASPM links in an AWGN channel. Thus many of the promising features of this modulation have not yet been explored and/or quantified. While we briefly outline some of these features in Section VII, the primary focus of this paper is on the spread-spectrum properties of M-ASPM, as these directly affect the M-ASPM networks' scalability. In this context, two most important attributes of M-ASPM are: (1) the processing gain is decoupled from the value of M and, for a given M, is directly proportional to the average interpulse interval; and (2) particular "shapes" (e.g., the "lengths" and the temporal and amplitude structures) of PSF waveforms are not directly constrained by either the value of M or the processing gain. Then, for example, depending on the areal coverage requirements and the practical constraints on the ToA of specific payloads, tens of even hundreds of different PSF channels can be simultaneously deployed within a given spectral band. This alone provides extensive versatility in trading multiple M-ASPM parameters to reconcile often conflicting LPWAN technical constraints.

In what follows, we may interchangeably employ continuous-time (analog) and discrete (digital) representations for time-varying quantities. We use the analog representation of a signal $x(t)$ when there are no explicit constraints on its bandwidth. When a discrete (digital) representation $x[k]$ is used, it is assumed that $x(t)$ is band-limited, and it is appropriately sampled so that $x(t)$ is completely determined by $x[k]$. Further, while the average interpulse interval N_{p} in M-ASPM is an integer, its practical values are rather large (e.g., $N_{\mathrm{p}} \gg 10$), and we routinely treat N_{p} as a continuous variable, in particular, when relating it to other continuous quantities (e.g. the range).

II. AGGREGATE SPREAD PULSE MODULATION

In the Aggregate Spread Pulse Modulation (ASPM) [8], [9], [10], the information is encoded in the amplitudes A_j and/or the "arrival times" k_j of the pulses in a digital "pulse train" $\hat{x}[k]$ with only relatively small fraction of samples having non-zero values:

$$\hat{x}[k] = \sum_j [\![k = k_j]\!]\, A_j, \tag{1}$$

where k is the sample index, k_j is the sample index of the j-th pulse, A_j is the amplitude of the j-th pulse, and the double square brackets denote the *Iverson bracket* [11]

$$[\![P]\!] = \begin{cases} 1 & \text{if } P \text{ is true} \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where P is a statement that can be true or false. The average "pulse rate" f_{p} in such a train is $f_{\mathrm{p}} = F_{\mathrm{s}}/N_{\mathrm{p}}$, where F_{s} is the sample rate, and $N_{\mathrm{p}} = \langle k_j - k_{j-1} \rangle$ is the average interpulse interval. Note that for $N_{\mathrm{p}} \gg 1$ the pulse rate is much smaller than the Nyquist rate. Also note that for $N_{\mathrm{p}} \gg 1$ this train has a large peak-to-average power ratio (PAPR) even when $|A_j| = \text{const}$, and is generally unsuitable for use as a modulating signal. However, the designed pulse train $\hat{x}[k]$ given by (1) can be "re-shaped" by linear filtering:

$$x[k] = (\hat{x} * \hat{g})[k] = \sum_j A_j\, \hat{g}[k - k_j], \tag{3}$$

where $\hat{g}[k]$ is the impulse response of the filter and the asterisk denotes convolution. The filter $\hat{g}[k]$ can be, for example, a lowpass filter with a given bandwidth B. If the filter $\hat{g}[k]$ has a sufficiently large time-bandwidth product (TBP) [12], [13], most of the samples in the reshaped train $x[k]$ will have non-zero values, and $x[k]$ will have a much smaller PAPR than the designed sequence $\hat{x}[k]$. Such low-PAPR signal can then be used for modulating a carrier. If the combination of the amplitude A_j and the arrival time k_j of a pulse provides M distinct "states," each pulse can encode $\log_2 M$ bits, and the raw bit rate f_{b} in such a train is $f_{\mathrm{b}} = f_{\mathrm{p}} \log_2 M$. Such signaling

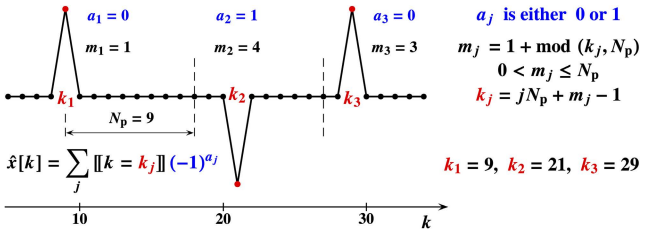

FIGURE 1. Example of designed pulse train with bipolar M-ary encoding.

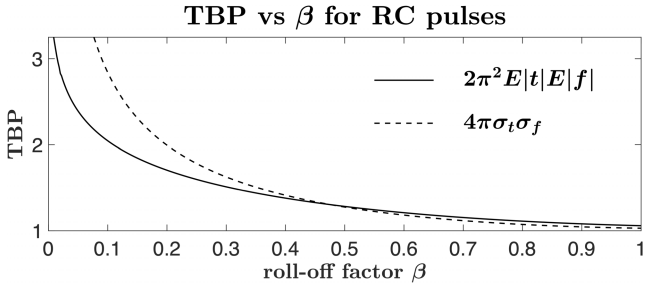

FIGURE 2. TBP of raised-cosine pulses as function of roll-off factor β.

is commonly referred to as "M-ary." When $B \gg f_\mathrm{b} = (F_\mathrm{s}/N_\mathrm{p})\log_2 M$, it results in a low-rate message encoded in a wideband waveform, and thus ASPM is a spread-spectrum technique.

To keep the energy per bit low, we may prefer not to use the magnitudes of the pulses for encoding, only their polarities (i.e., keep $|A_j| = \mathrm{const}$) and the arrival times, since those can be changed without changing the energy of the pulses. For example, for the arrival times in (1) one can use

$$k_j = jN_\mathrm{p} + \Delta k[m_j], \tag{4}$$

where $m_j \le M$ is a positive integer and $\Delta k[m]$ is an integer-valued invertible function, such that $0 \le \Delta k[m] < N_\mathrm{p}$ and $\Delta k[m] \ne \Delta k[l]$ for $m \ne l$. Then for $m_j \in \{1, 2, \ldots, M\}$ and $A_j = \mathrm{const}$ the pulse train given by (1) encodes $\log_2 M$ bits per pulse. We will refer to such M-ary encoding with $A_j = \mathrm{const}$ as "unipolar."

Another bit can be added by using $A_j = (-1)^{a_j}$, where a_j is either "0" or "1," and we will refer to such signaling as "bipolar." Then for bipolar M-ary signaling equation (1) can be rewritten as

$$\hat{x}[k] = \sum_j [\![k = jN_\mathrm{p} + \Delta k[m_j]]\!]\,(-1)^{a_j}, \tag{5}$$

where $m_j \in \{1, 2, \ldots, M/2\}$ and $a_j \in \{0, 1\}$. Fig. 1 illustrates a designed pulse train with such bipolar M-ary encoding according to (5), with $\Delta k[m_j] = m_j - 1$. In this example, $N_\mathrm{p} = 9$ and for $m_j \le N_\mathrm{p}$ each pulse can have $M = 2N_\mathrm{p} = 18$ distinct states.

For a given designed pulse sequence $\hat{x}[k]$, the spectral, temporal and amplitude structures of the reshaped train $x[k]$ will be determined by the choice of $\hat{g}[k]$. In particular, it may be desirable to select a filter $\hat{g}[k]$ that minimizes the PAPR of $x[k]$. Note that if the time duration of $\hat{g}[k]$ extends over multiple interpulse intervals, the instantaneous amplitudes and/or phases [14] of the resulting waveform are no longer representative of individual pulses. Instead, they are a "piled-up" aggregate of the contributions from multiple "stretched" pulses.

The key property of the large-TBP pulse shaping filter (PSF) $\hat{g}[k]$ is that its autocorrelation function (ACF), i.e., the convolution of $\hat{g}[k]$ with its matched filter $g[k] = \hat{g}[-k]$, has a much smaller TBP, in particular, sufficiently smaller than the ratio B/f_p. Then, after demodulation and analog-to-digital (A/D) conversion in the receiver, the encoded binary sequence can be recovered by filtering with $g[k]$ and sampling the resulting pulse train at $k = jN_\mathrm{p} + \Delta k[m]$, where $m \in \{1, 2, \ldots, M/2\}$ for bipolar, and $m \in \{1, 2, \ldots, M\}$ for unipolar encoding (i.e., using $g[k]$ as a decimation filter).

A good choice for the PSF would be a pulse that combines a small TBP of its ACF (e.g., close to that of a Gaussian pulse) with ACF's compact frequency support. An example of such ACF would be a raised-cosine (RC) pulse [15] with a sufficiently large roll-off factor β. While compact support cannot be simultaneously achieved for the temporal and the spectral power densities of any pulse, the standard deviations, σ_t and σ_f, of these power densities are typically used as measures of their width [12], [13]. Then, e.g., the TBP of a pulse can be defined as $\mathrm{TBP} = 4\pi\sigma_t\sigma_f \ge 1$, with the equality (the smallest $\mathrm{TBP} = 1$) achieved for a Gaussian pulse. However, for the temporal and spectral power densities of RC pulses, the absolute deviations around zero, $E|t|$ and $E|f|$, can be used instead. This makes the TBP measure less sensitive to the long "tails" of the RC pulses in the time domain. Then the TBP can be defined as $\mathrm{TBP} = 2\pi^2 E|t|E|f| \ge 1$. As can be seen in Fig. 2, with either definition the TBPs of RC pulses remain relatively small for large roll-off factors (e.g., $\mathrm{TBP} \lesssim 2$ for $1/5 \lesssim \beta \le 1$). Further, for a PSF with an RC ACF, the sample rate F_s can be chosen as $F_\mathrm{s} = 2N_\mathrm{s}B$, where $1 \le N_\mathrm{s} = 2/(1+\beta) < 2$ is the oversampling factor, and the spectral efficiency of the M-ASPM with such pulse shaping can be expressed as $\eta = f_\mathrm{b}/B = 4\log_2 M/N_\mathrm{p}/(1+\beta)$. For example, for $\beta = 1/3$ and $M = 16$, $\eta = 12/N_\mathrm{p}$.

Since for a given designed pulse sequence $\hat{x}[k]$ the temporal and amplitude structures of the reshaped train $x[k]$ are determined by the PSF $\hat{g}[k]$, these structures can be substantially different even for the PSFs with the same ACF. As discussed in [9] and [10], one can construct a great variety of large-TBP PSFs $\hat{g}_1[k]$, $\hat{g}_2[k]$, and so on, with the same small-TBP ACF $w[k]$, so that $(\hat{g}_i * g_i)[k] = w[k]$ for any i, while the convolutions of any $\hat{g}_i[k]$ with $g_j[k]$ for $i \ne j$ (cross-correlations) have large TBPs. Further, this property will also effectively hold for the PSFs $\hat{h}_i[k]$ such that $\hat{h}_i[k]$ is the discrete Hilbert transform of $\hat{g}_i[k]$, i.e., $\hat{h}_i[k] = H\{\hat{g}_i[k]\}$ [16], [17]. Therefore, using various combinations of PSFs $\hat{g}[k] \in \{\hat{g}_1[k], \hat{g}_2[k], \ldots\}$ and $\hat{h}[k] \in \{\hat{h}_1[k], \hat{h}_2[k], \ldots\}$, we can design different coherent and non-coherent modulation schemes with emphasis on particular spectral and/or temporal properties of the modulated signal.

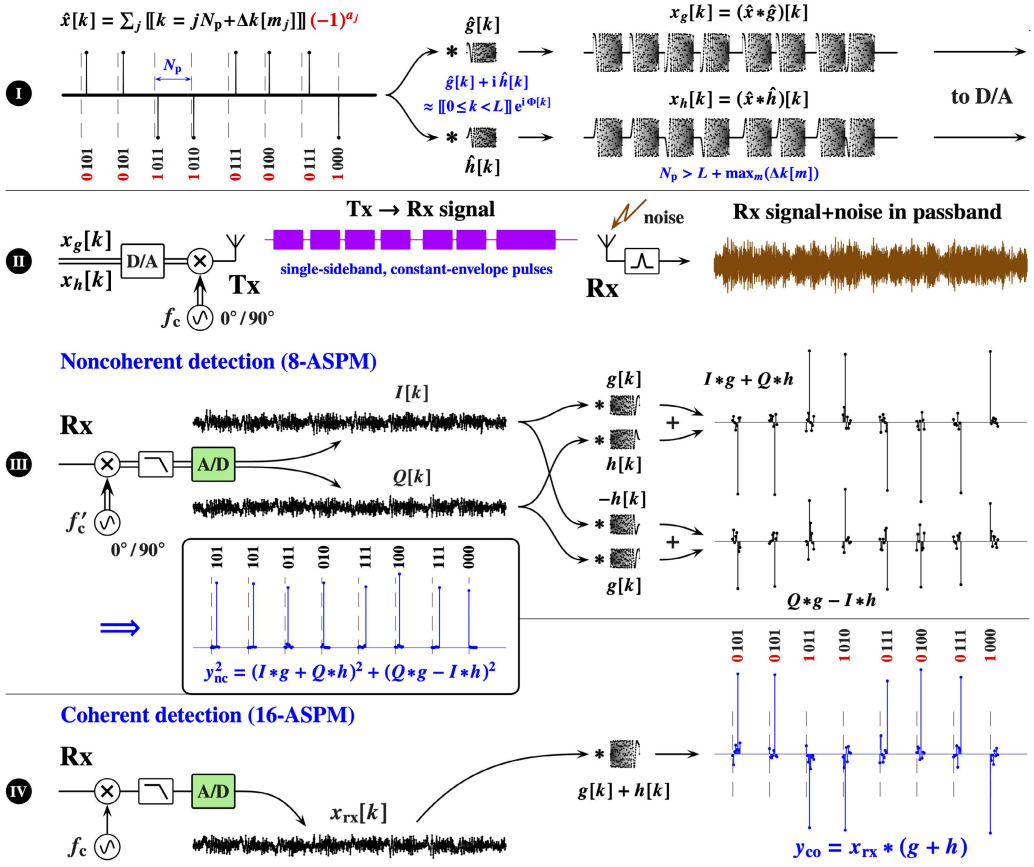

FIGURE 3. Illustration of single-sideband M-ary ASPM link with constant-envelope pulses and their noncoherent and coherent detection.

A. SINGLE-SIDEBAND M-ARY ASPM WITH CONSTANT-ENVELOPE PULSES

For example, Fig. 3 illustrates a single-sideband M-ary ASPM link which uses constant-envelope transmitted pulses and is suitable for both coherent ('co') and noncoherent ('nc') detection.

In the designed pulse train $\hat{x}[k]$ according to (5), we use 8 distinct possible pulse locations relative to jN_p (see Fig. 3(I)). This train is then filtered with $\hat{g}[k]$ and $\hat{h}[k]$ to form the shaped trains $x_g[k]$ and $x_h[k]$. After digital-to-analog (D/A) conversion, $x_g(t)$ and $x_h(t)$ are used for quadrature amplitude modulation of a carrier with frequency f_c, providing the transmitted waveform $x_g(t)\sin(2\pi f_\mathrm{c}t) + x_h(t)\cos(2\pi f_\mathrm{c}t)$ (Fig. 3(II)). If $\hat{g}[k]$ and $\hat{h}[k]$ are, say, the real and imaginary parts, respectively, of a nonlinear chirp with the desired ACF, e.g.

$$\hat{g}[k] + \mathrm{i}\,\hat{h}[k] = [\![0 \le k < L]\!]\,\exp(\mathrm{i}\,\Phi[k])\,, \tag{6}$$

where L is a positive integer (the "length" of the pulse in samples) and $\Phi[k]$ is the phase, then this waveform will occupy only a single sideband with the physical bandwidth B equal to the baseband bandwidth of the chirp. In addition, if the pulses do not overlap (e.g., $N_\mathrm{p} > L + \max_m(\Delta k[m])$), this waveform will consist of constant-envelope pulses.

In the receiver, we can use either noncoherent or coherent detection. While demodulation techniques may vary, Figs. 3(III) and 3(IV) provide particular examples.

For noncoherent detection (Fig. 3(III)), in the receiver's (Rx) quadrature demodulator the noisy passband signal is multiplied by the orthogonal sinusoidal signals from a local oscillator, lowpassed, and converted to the in-phase and quadrature digital signals $I[k]$ and $Q[k]$. We then use the matched filters $g[k]$ and $h[k]$, as shown in Fig. 3(III), to obtain 8 samples per pulse of the high-peakedness pulse train $y_\mathrm{nc}[k]$ corresponding to the designed pulse train. Out of each 8 samples, the position of the sample with the largest magnitude will correspond to the position of the respective pulse in the designed train.

Noncoherent detection does not require precise carrier synchronization, neither in phase nor frequency, but it does not recover the polarity of the pulses. Thus we do not obtain the most significant bit in the symbols encoded in the pulses of the designed train. For coherent detection, we would need to recover the phase of the carrier. But then we can also measure, for each pulse, the polarity of the sample with the largest magnitude, and thus obtain an extra bit per pulse.

For coherent detection (Fig. 3(IV)), after multiplication by $\sin(2\pi f_\mathrm{c}t + \pi/4)$, lowpass filtering, and A/D conversion

in the receiver, the resulting signal $x_{\rm rx}[k]$ is filtered with $g[k]+h[k]$ to form the bipolar baseband pulse train $y_{\rm co} = x_{\rm rx} * (g+h)$ corresponding to the designed train $\hat{x}[k]$.

Without loss of generality, the ACFs of $\hat{g}[k]$ and $\hat{h}[k]$ can be normalized to have the peak magnitudes equal to unity. Then, to avoid the interpulse interference, we can require that for coherent detection

$$w[\Delta k[m]-\Delta k[l]] = [\![m=l]\!], \tag{7}$$

where $w = \frac{1}{2}(\hat{g}*g + \hat{h}*h)$, and, for noncoherent detection,

$$v^2[\Delta k[m]-\Delta k[l]] = [\![m=l]\!], \tag{8}$$

where $v^2 = w^2 + \frac{1}{4}(\hat{h}*g - \hat{g}*h)^2$.

1) IMPROVING RESISTANCE to MULTIPATH DELAY AND DOPPLER SPREADS, and TO IMPULSIVE NOISE

For PSFs with a given ACF, the primary parameters of M-ASPM, affecting its spectral and energy-per-bit efficiencies, are the values of M and the average interpulse interval $N_{\rm p}$. As discussed in the subsequent sections, when M-ASPM with a given M is viewed as a spread spectrum technique, the value of $N_{\rm p}$ in particular is proportional to the processing gain. Thus, in the context of the overall M-ASPM's suitability for use in LPWANs, our primarily focus in the rest of this paper is the impact of the interpulse interval on the M-ASPM's properties.

However, we would like to mention in passing that other parameters of a particular M-ASPM implementation may have significant influence on its behavior under specific practical scenarios (e.g., moving network nodes, urban environments, etc.). For example, for the pulse-position encoding expressed by (4), the minimum time interval between the pulse positions corresponding to different symbols is given by the ratio $\min_{l\neq m}(|\Delta k[l]-\Delta k[m]|)/F_{\rm s}$. Thus performance of noncoherent M-ASPM in multipath propagation can generally be improved by increasing this interval, so it becomes sufficiently large with respect to the delay spread. Or, in the constant-envelope link described above, the time support of a PSF given by (6) is equal to $L/F_{\rm s}$. Therefore, this link would be insensitive to the relative velocity Δv between the transmitter and the receiver if $|\Delta v|/c \lesssim F_{\rm s}/(Lf_{\rm c})$, where c is the speed of light. This enables us to control M-ASPM's Doppler tolerance for a wide rage of spectral efficiencies (i.e., the values of $N_{\rm p}$).

Further, note that the A/D conversion in the ASPM receiver can be combined with intermittently nonlinear filtering (INF) described in [18], [19], and [20]. In INF, we establish a robust range that excludes noise outliers while including the signal of interest [21]. Then, we replace the outlier values with those in mid-range. Note that INF affects only a relatively small fraction of the samples in the signal+noise mixture, only those with exceedingly large magnitudes.

In Fig. 3, it is assumed that the A/D conversion can be performed in such a manner. As INF modifies the samples in the demodulated signal *before* the subsequent large-TBP filtering, its effect on the baseband signal of interest will be insignificant, while the contribution of the large-power noise outliers to the baseband noise will be reduced. This makes the link robust to outlier interferences, e.g., wideband impulsive noise commonly present in industrial environments [22], and increases the baseband signal-to-noise ratio (SNR) in the presence of such interferences. Since in the power-limited regime the channel capacity is proportional to the SNR, even relatively small increase in the latter will be beneficial.

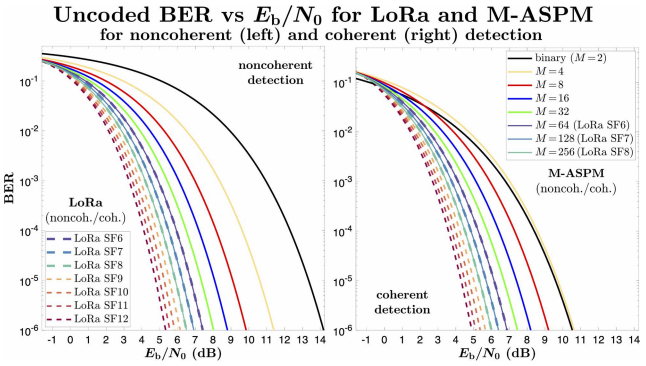

FIGURE 4. Uncoded BER vs $E_{\rm b}/N_0$ performances of LoRa (dashed lines in both subfigures) and single-sideband M-ASPM (solid lines in both subfigures) in AWGN channel.

B. UNCODED BER PERFORMANCE OF M-ASPM IN AWGN CHANNEL

Additive white Gaussian noise (AWGN) is only a "background" noise component in the congested spectrum of most IoT applications. Nevertheless, assessment of the M-ASPM properties in an AWGN channel provides a suitable benchmark for the subsequent evaluation of the M-ASPM performance under various more realistic propagation conditions and interference scenarios, and such assessment is given in [8]. For example, in an AWGN channel, the uncoded bit error rates (BER) performance of the coherent binary ASPM ($M = 2$) is identical to that of the binary phase-shift keying (BPSK) modulation. Noticeably, for large values of M, it is shown that the energy per bit efficiency of the M-ASPM (with constant magnitude pulses, $|A_j| = \text{const}$) is the same as of the LoRa modulation [6], [7] with the spreading factor equal to $\log_2 M$, for both coherent and noncoherent detection. This is illustrated in Fig. 4, where $E_{\rm b}$ is the energy per bit and N_0 is the (one-sided) PSD of the noise.

1) SIMULATED BER VS SNR PERFORMANCE OF 16-ASPM

Figure 5 compares the calculated (according to the expressions provided in [8], dashed lines) and the simulated (markers connected by solid lines) BERs for both coherent and noncoherent 16-ASPM links with different spreading factors $B/f_{\rm b}$. In this example, the ACF of the PSF $\hat{g}[k]$ is an RC pulse with the roll-off factor $\beta = 1/4$, and we use oversampling with $N_{\rm s} = 2/(1+\beta) = 8/5$. Then $B/f_{\rm b} = N_{\rm p}/(2N_{\rm s}\log_2 M) = 5N_{\rm p}/64$. In the figure, Γ denotes the SNR defined as $\Gamma = (E_{\rm b}/N_0)\times(f_{\rm b}/B)$.

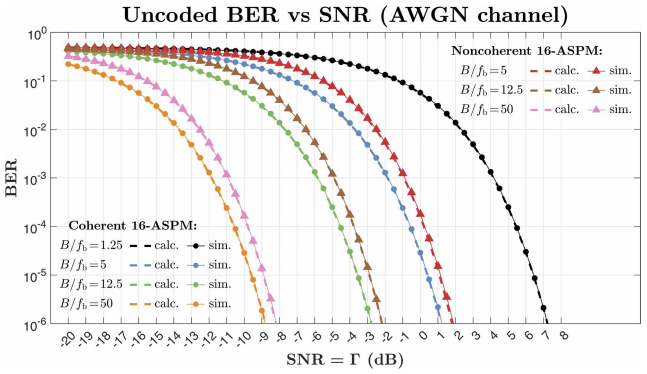

FIGURE 5. Calculated and simulated BERs as functions of AWGN SNRs for both coherent and noncoherent 16-ASPM with different values of B/f_b.

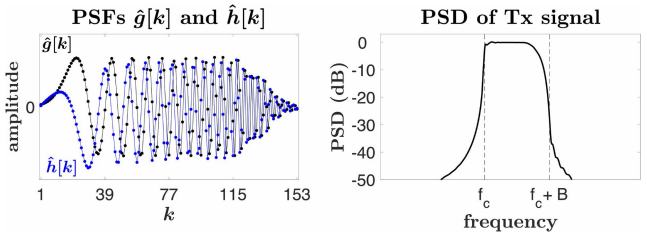

FIGURE 6. Example of PSFs $\hat{g}[k]$ and $\hat{h}[k]$, and PSD of modulated carrier in simulations shown in Fig. 5.

For the coherent 16-ASPM, the designed pulse train $\hat{x}[k]$ is given by

$$\hat{x}[k] = \sum_j [\![k = jN_\mathrm{p} + (4a_j + 2b_j + c_j)n]\!]\,(-1)^{d_j}, \qquad (9)$$

where $n = 2$, which encodes a 4-bit sequence $(a_1b_1c_1d_1\, a_2b_2c_2d_2 \ldots a_jb_jc_jd_j \ldots)$. For the noncoherent 16-ASPM, the designed pulse train is

$$\hat{x}[k] = \sum_j [\![k = jN_\mathrm{p} + (8a_j + 4b_j + 2c_j + d_j)n]\!], \qquad (10)$$

where $n = 4$. In the transmitter, filtering $\hat{x}[k]$ with the PSF $\hat{g}[k]$ forms the modulating component $x_\mathrm{I}[k]$, and filtering $\hat{x}[k]$ with the PSF $\hat{h}[k]$ forms the modulating component $x_\mathrm{Q}[k]$. The filter $\hat{h}[k]$ approximates the discrete Hilbert transform of $\hat{g}[k]$, i.e., $\hat{h}[k] \approx H\left\{\hat{g}[k]\right\}$ [16], [17], and thus $x_\mathrm{Q}[k]$ approximates the discrete Hilbert transform of $x_\mathrm{I}[k]$, i.e., $x_\mathrm{Q}[k] \approx H\{x_\mathrm{I}[k]\}$. Therefore, if after digital-to-analog conversion $x_\mathrm{I}(t)$ and $x_\mathrm{Q}(t)$ are used for quadrature amplitude modulation of a carrier with frequency f_c, the resulting modulated waveform $x_\mathrm{I}(t)\sin(2\pi f_\mathrm{c}t) + x_\mathrm{Q}(t)\cos(2\pi f_\mathrm{c}t)$ effectively occupies only a single sideband with the physical bandwidth B equal to the baseband bandwidth of $\hat{g}[k]$. Figure 6 illustrates both the PSFs $\hat{g}[k]$ and $\hat{h}[k]$ (left panel), and the PSD of the modulated carrier used in the simulations (right panel).

In the coherent receiver, the noisy passband signal is multiplied by the signal $\sin(2\pi f_\mathrm{c}t + \pi/4)$ from the local oscillator, lowpassed, and A/D converted to form the digital signal $x_\mathrm{rx}[k]$, which is then filtered with $g[k] + h[k]$ to form the baseband pulse train

$$y_\mathrm{co} = x_\mathrm{rx} * (g + h). \qquad (11)$$

For noncoherent detection, in the receiver's quadrature demodulator the noisy passband signal is multiplied by $\sin(2\pi f_\mathrm{c}t + \Delta\varphi)$ and $\cos(2\pi f_\mathrm{c}t + \Delta\varphi)$, lowpassed, and A/D converted to the in-phase (I) and quadrature (Q) digital signals $I[k]$ and $Q[k]$. Then the received unipolar pulse train is formed as

$$y^2_\mathrm{nc} = (I * g + Q * h)^2 + (Q * g - I * h)^2. \qquad (12)$$

In the simulations, the bit error rates are determined by comparing the bit sequences extracted from the "ideal" transmitted signals (without noise), and from the transmitted signals affected by AWGN with a given PSD N_0.

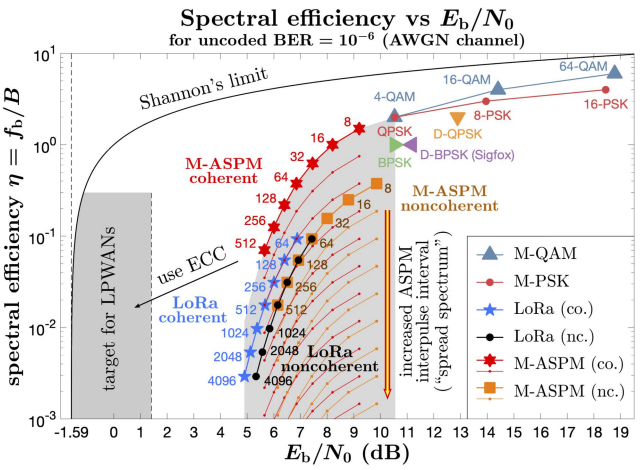

FIGURE 7. Assessing M-ASPM's suitability for use in LPWANs.

III. M-ASPM's SUITABILITY FOR USE IN LPWANs

Let us now briefly assess ASPM's overall suitability for use in LPWANs, for example, to provide long range wireless access in the IoT applications.

A. M-ASPM's SPECTRAL EFFICIENCY VS E_b/N_0

To begin, Fig. 7 compares the M-ASPM's spectral efficiency vs E_b/N_0 with those of several other modulations, including such commonly used in LPWANs as Sigfox and LoRa, for uncoded bit error rates BER $= 10^{-6}$ in an AWGN channel.

As can be seen in the figure, for a given M the maximum spectral efficiency of M-ASPM equals that of LoRa for noncoherent detection, and exceeds that of LoRa by a factor of 4 for coherent detection. (The M-ASPM's spectral efficiency can be maximized when the PSF's ACF is an RC pulse with zero roll-off factor, i.e., the sinc function). However, unlike the LoRa (which operates at maximum spectral efficiency for a given $M = 2^{\mathrm{SF}}$), the M-ASPM is a "true" spread spectrum technique, and its spectral efficiency is simply inversely proportional to the average interpulse interval N_p. This is shown

by the red and orange curves for the M-ASPM, with the factor of 2 difference in N_p for any two adjacent same-color curves. Consequently, the M-ASPM's receiver sensitivity is proportional to the average interpulse interval, and the desired BER can be achieved, for any SNR, by changing N_p.

For LPWANs, we ultimately want to be somewhere in the shaded region on the left in Fig. 7, where both the spectral efficiency and the energy per bit are low. In M-ASPM, by changing M and/or N_p, we "access" multiple values of spectral and energy per bit efficiencies within the shaded region in the middle. Note that this region fully contains the LoRa's values. Further, these values are for uncoded bit error rates. We can then use error correction coding (ECC) to improve the effective energy per bit performance, and to "move" the M-ASPM's values toward the target region for LPWANs. Thus M-ASPM offers to be a suitable modulation technique for LPWANs.

We will further focus on the noncoherent M-ASPM in particular, as it does not require precise carrier synchronization and is more resilient to various types of fading. Also, in quantitative examples we will use M-ASPM with $M = 16$, as a compromise between the energy-per-bit efficiency and the computational intensity of signal processing.

B. CONTROLLING RANGE BY INTERPULSE INTERVAL

For both noncoherent LoRa and noncoherent M-ASPM, the bit error probability P_b in AWGN channel can be expressed as [8]

$$P_\mathrm{b} = P_\mathrm{b}\left(\frac{\Gamma}{\eta}\right) = \frac{1}{2(M-1)}\sum_{k=2}^{M}(-1)^k\binom{M}{k}\exp\left(-\frac{k-1}{k}\,\frac{\Gamma}{\eta}\,\log_2 M\right), \tag{13}$$

where $\binom{n}{m} = \frac{n!}{(n-m)!\,m!}$ is the binomial coefficient, $\Gamma = (E_\mathrm{b}/N_0)\times(f_\mathrm{b}/B)$ is the SNR, and $\eta = f_\mathrm{b}/B$ is the spectral efficiency. (For LoRa, $M = 2^{\mathrm{SF}}$.)

From now on, to distinguish between the respective quantities for LoRa and M-ASPM, we mark those for LoRa by overhead tildes. Then the spectral efficiency of LoRa modulation is

$$\widetilde{\eta} = \widetilde{\eta}\left(\widetilde{M}\right) = \frac{\log_2\widetilde{M}}{\widetilde{M}}. \tag{14}$$

We will further use PSFs with RC ACFs for M-ASPM, and the sample rate $F_\mathrm{s} = 4B/(1+\beta)$, as discussed in Section II. Consequently, for M-ASPM

$$\eta = \eta\left(M, N_\mathrm{p}\right) = \frac{4\log_2 M}{(1+\beta)N_\mathrm{p}} = \widetilde{\eta}(M)\,\frac{4M}{(1+\beta)N_\mathrm{p}}, \tag{15}$$

and, for a given M, the spectral efficiency is inversely proportional to the average interpulse interval N_p. The minimum N_p value that can be used in noncoherent M-ASPM is $4M$, and thus the maximum spectral efficiency of M-ASPM is

$$\eta_\mathrm{max} = \eta\left(M, 4M\right) = \widetilde{\eta}(M)/(1+\beta). \tag{16}$$

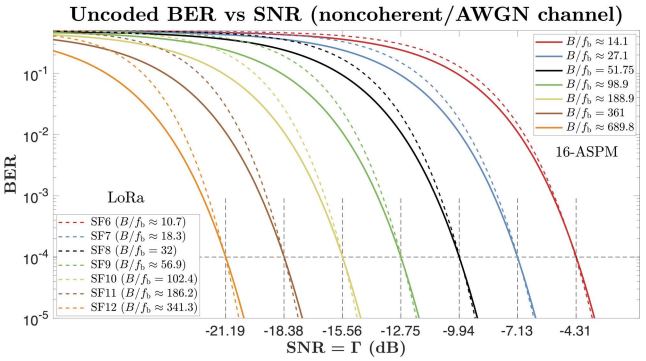

FIGURE 8. Uncoded BER vs SNR performances of LoRa (dashed lines) and single-sideband 16-ASPM (solid lines) for noncoherent detection in AWGN channel.

Note that, when $\beta = 0$ (sinc function ACF), it is equal to the spectral efficiency of noncoherent LoRa with $\widetilde{M} = M$.

Henceforth, whenever we compare M-ASPM and LoRa, we assume identical physical parameters of the links. For example, we assume the same physical frequency band, transmit power, antenna gains, and various system attenuations such as insertion and matching losses, etc.

If we desire to achieve the same BER performance at the same range (i.e., at the same SNR Γ) for LoRa (with a given $\widetilde{M}$) and M-ASPM (with a given M), the value of N_p can be obtained as a solution of the equalities

$$P_\mathrm{b}\left(\Gamma; M, N_\mathrm{p}\right) = \widetilde{P}_\mathrm{b}\left(\Gamma; \widetilde{M}\right) = \mathrm{BER}. \tag{17}$$

An example is given in Fig. 8, for $M = 16$ and $\mathrm{BER} = 10^{-4}$.

Normally, the received power decreases with the distance d between the transmitter and the receiver, and the SNR is a decreasing function of d. For example, for power-law path loss $\Gamma \propto d^{-\gamma}$, where γ is the path-loss exponent. For free-space path loss $\gamma = 2$, and it can be 2-3 times larger for harsh environments [23], [24], [25]. Then, from the condition $\Gamma/\eta = \text{const}$ it follows that, for the power-law path loss, the M-ASPM range $d \propto \eta^{-1/\gamma} \propto {N_\mathrm{p}}^{1/\gamma}$. While N_p is an integer, it is rather large ($N_\mathrm{p} \geq 4M$ for noncoherent M-ASPM) and, for a given M (e.g. $M = 16$), the M-ASPM's spectral efficiency can be treated as a continuous quantity. This is in contrast with LoRa, where LoRa's spectral efficiency is constant for a given spreading factor.

For example, Fig. 9 illustrates 16-ASPM's spectral efficiency vs. range under power-law path loss model, at AWGN $\mathrm{BER} = 10^{-4}$, as compared with LoRa. For 16-ASPM, η is the solution of the equality $P_\mathrm{b}\,(d; M = 16, \eta) = \mathrm{BER}$. For LoRa, the spectral efficiency is the maximum value of $\widetilde{\eta}$ satisfying the inequality $\widetilde{P}_\mathrm{b}\,(d; \widetilde{\eta}) \leq \mathrm{BER}$. As can be seen in the figure, for the same range, the spectral efficiency of 16-ASPM varies from more than double of LoRa at short ranges (e.g., at ranges sufficiently smaller than that for LoRa with $\mathrm{SF} = 6$), down to about a half of LoRa at the maximum LoRa range (for $\mathrm{SF} = 12$). Overall, for ranges above LoRa with $\mathrm{SF} = 7$, the spectral efficiency of 16-ASPM is somewhat lower than that of LoRa.

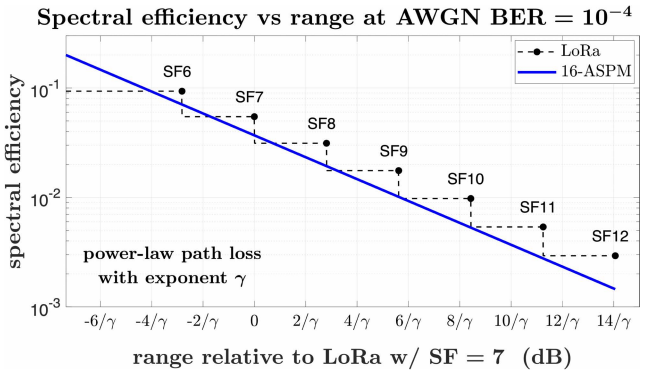

FIGURE 9. Spectral efficiency vs range under power-law path loss model (γ is path-loss exponent), for LoRa and 16-ASPM at AWGN BER = 10^{-4} (noncoherent detection).

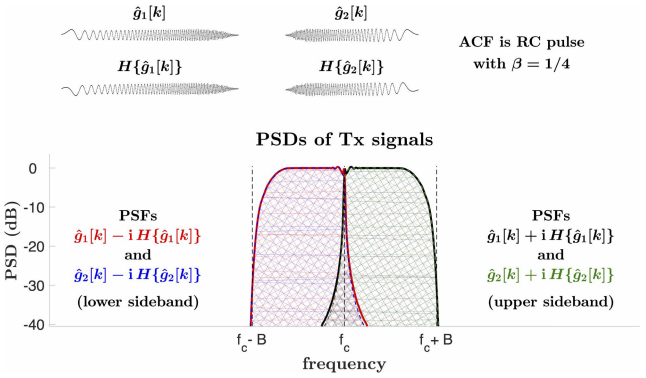

FIGURE 10. Simulated example of transmitted adjacent bands for PSFs with ACF as RC pulse with $\beta = 1/4$.

At first glance, this is an unfavorable comparison for 16-ASPM at long ranges, since lower spectral efficiency leads to longer ToA, exacerbating collisions from multiple transmitters. However, as discussed in the next section, the ToA limitation is only as severe as the impact of such collisions. If this impact is sufficiently small, then, for example, the 16-ASPM's ability to extend the range in smaller, more controllable increments enables us to place a significantly larger (say, by an order of magnitude) number of transmitters at longer ranges, thus more than compensating for a smaller spectral efficiency of a single transmitter.

IV. INTER-PSF COLLISIONS IN M-ASPM

When considering the impact of mutual interference of multiple M-ASPM transmitters, we shall recall that different M-ASPM transmitters can employ substantially different PSFs, in a manner similar to using different spreading sequences in asynchronous CDMA. Then the constraints on the actual PSF shape would be derived from the constraints on the physical length of the filters and the PAPR of the transmitted signal.

A. MUTUAL INTERFERENCE OF TWO M-ASPM TRANSMITTERS

Let us first examine mutual interference of two single-sideband M-ASPM transmitters, with the PSFs $\hat{g}_1[k]$ and $\hat{h}_1[k] = H\left\{\hat{g}_1[k]\right\}$ for the 1st transmitter, and $\hat{g}_2[k]$ and $\hat{h}_2[k] = H\left\{\hat{g}_2[k]\right\}$ for the 2nd transmitter.

First note that, for the matched filters in the receiver, $h_1[k] = -H\{g_1[k]\}$ and $h_2[k] = -H\{g_2[k]\}$. Further note that [26] $H(H(g))(t) = -g(t)$ and, for the Hilbert transform of the convolution of $g(t)$ and $h(t)$,

$$H(g*h) = H(g)*h = g*H(h). \tag{18}$$

With these equalities,

$$(\hat{g}_1 + \mathrm{i}\,\hat{h}_1)*(g_2 - \mathrm{i}\,h_2) = 2\left[(\hat{g}_1*g_2) + \mathrm{i}\,H(\hat{g}_1*g_2)\right]. \tag{19}$$

Now, without loss of generality, the in-phase and quadrature signals in the receiver can be represented as

$$I = A_1\hat{x}_1*\hat{g}_1 + A_2\hat{x}_2*\left(\hat{g}_2\cos\Delta\varphi + \hat{h}_2\sin\Delta\varphi\right) \tag{20}$$

and

$$Q = A_1\hat{x}_1*\hat{h}_1 + A_2\hat{x}_2*\left(\hat{h}_2\cos\Delta\varphi - \hat{g}_2\sin\Delta\varphi\right), \tag{21}$$

where $\hat{x}_1[k]$ and $\hat{x}_2[k]$ are the designed pulse trains. Then

$$\begin{aligned} I*g_1 + Q*h_1 = {} & 2A_1\hat{x}_1 * w + 2A_2\hat{x}_2 \\ & * \left((\hat{g}_2*g_1)\cos\Delta\varphi + H(\hat{g}_2*g_1)\sin\Delta\varphi\right), \end{aligned} \tag{22}$$

where $w[k] = (\hat{g}_1 * g_1)[k] = (\hat{g}_2 * g_2)[k]$ is the ACF, and

$$Q*g_1 - I*h_1 = -H\left(I*g_1 + Q*h_1\right). \tag{23}$$

Therefore, the pulse train obtained using the filters g_1 and h_1 in the receiver is the squared *complex envelope* [26] of the signal represented by (22). Note that the first term on the right-hand side of (22) is a high peakedness pulse train. If the convolution $\hat{g}_2 * g_1$ has a sufficiently large TBP, then the second term is a low-peakedness signal, and its impact on the resulting signal will be akin to the impact of a noise with relatively low PAPR. While such noise is non-Gaussian in general, its Gaussian approximation would be mostly adequate for the assessment of its effect on the BER, especially at low SNRs [18], [19], [20], [21].

It is worth mentioning at this point that changing the sign of one of the PSFs in one of the transmitters effectively eliminates the mutual interference. Indeed, e.g., using $\hat{h}_1[k] = -H\left\{\hat{g}_1[k]\right\}$ in the 1st transmitter "flips" the sideband of its modulated signal. Then the 1st and the 2nd transmitted signals will occupy different (adjacent) frequency bands and, predictably, will not interfere with each other. This is illustrated in Fig. 10, which provides a simulated (in a manner described in Section II-B1) example of such adjacent bands for 16-ASPM, when the PSF's ACF is an RC pulse with $\beta = 1/4$. Equivalently, the absence of such interference follows from the equality

$$(\hat{g}_1 - \mathrm{i}\,\hat{h}_1)*(g_2 - \mathrm{i}\,h_2) = 0. \tag{24}$$

Favorably, this change in the frequency band is accomplished without changing the frequency of the local oscillators in the transmitter and the receiver.

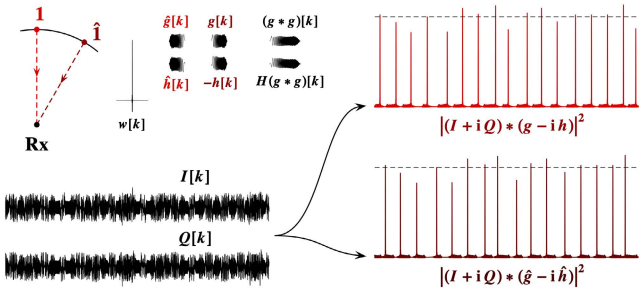

FIGURE 11. For large spreading factors ($B/f_b = 30$ in this example), mutual interference of two single-sideband M-ASPM transmitters with "flip" PSFs is insignificant for PSFs with sufficiently large TBP.

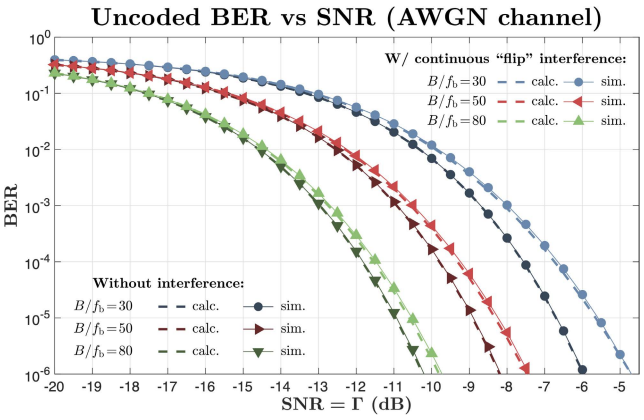

FIGURE 12. Calculated (dashed lines) and simulated (markers connected by solid lines) BERs for noncoherent 16-ASPM links with different spreading factors B/f_b (30, 50 and 80), and with and without continuous interference from "flip" transmitter.

1) "Flip" PULSE SHAPING FILTERS

Note that the TBP of the convolution of any PSF $\hat{g}$ with itself is always larger than the TBP of $\hat{g}$. In particular, for "flip" PSFs such that $\hat{g}_2 = g_1$ and $\hat{h}_2 = -h_1$, the TBP of $\hat{g}_2 * g_1 = g_1 * g_1$ is about twice as large as the (already large) TBP of g_1. Consequently, mutual interference of two single-sideband M-ASPM transmitters (operating in the same sideband, at the same power, similar distances from the receiver, and with similar average interpulse interval, i.e., similar spectral efficiencies) with "flip" PSFs is insignificant for PSFs with a sufficiently large TBP. This is illustrated in Fig. 11. (The horizontal dashed lines in the right-hand side indicate the magnitudes of the interference-free pulses in the received pulse trains.) In this example $B/f_b = 30$, which, in terms of the range, roughly corresponds to the range of LoRa with SF = 7. In fact, the interference from a "flip" transmitter can be treated as AWGN with the power equal to the signal power, i.e., with the SIR 0 dB. Thus its impact becomes smaller as the spectral efficiency decreases (for larger N_p). This is illustrated in Fig. 12, which compares the calculated (dashed lines) and the simulated (markers connected by solid lines) BERs for noncoherent 16-ASPM links with different spreading factors B/f_b (30, 50 and 80), and with and without interference from a "flip" transmitter.

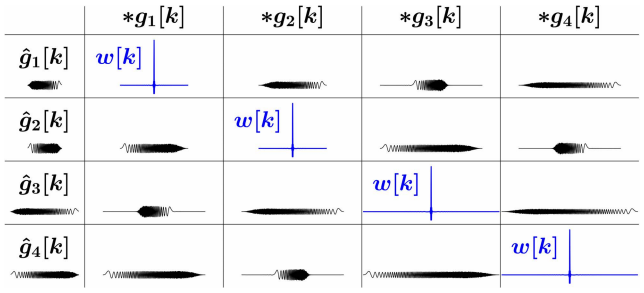

FIGURE 13. Example of PSFs with identical small-TBP ACFs and large-TBP cross-correlations.

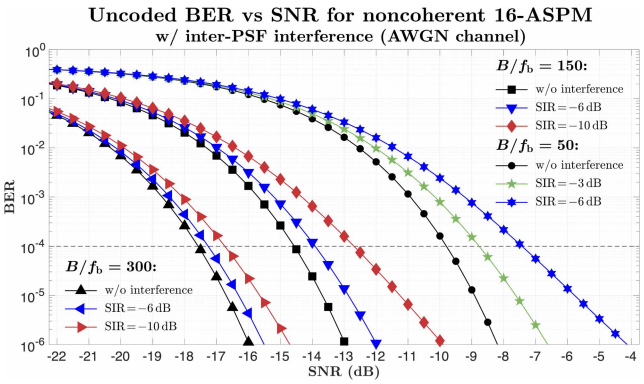

FIGURE 14. Impact of interference on M-ASPM decreases with average interpulse interval.

Note that the impact of the interference shown in Fig. 12 is for *continuous* transmission from the 2nd transmitter, when a "full" collision with the signal from the 1st transmitter is a certainty. This impact would be much smaller when the probability of such collisions is significantly less than unity. Therefore, by employing "flip" PSFs for transmitters operating at the same power, similar distances from the receiver, and with similar average interpulse interval (i.e., similar spectral efficiencies), we can effectively double the number of "collision-free" nodes in an M-ASPM network, especially for the "outer" nodes operating at longer range.

B. MULTIPLE INTER-PSF COLLISIONS

As discussed in Section II (also see [9], [10]), one can construct many large-TBP PSFs $\hat{g}_1[k]$, $\hat{g}_2[k]$, and so on, with the same small-TBP ACF $w[k]$, so that $(\hat{g}_i * g_i)[k] = w[k]$ for any i, while the convolutions of any $\hat{g}_i[(t)]k]$ with $g_j[(t)]k]$ for $i \neq j$ (cross-correlations) have large TBPs. An example of such PSFs is shown in Fig. 13.

Then the impact of the interference from transmitters with $\hat{g}_j[k] \in \{\hat{g}_2[k], \hat{g}_3[k], \dots\}$ (i.e., when $j \neq 1$) on the signal from the transmitter with $\hat{g}_1[k]$ would be akin to the impact of Gaussian noise with the power equal to the combined power of the interfering signals at the receiver. Since, as illustrated in Fig. 14, this impact decreases with average interpulse interval, becoming relatively insignificant for large values of B/f_b, this adds to the flexibility in achieving desired profiles of node densities in M-ASPM LPWANs.

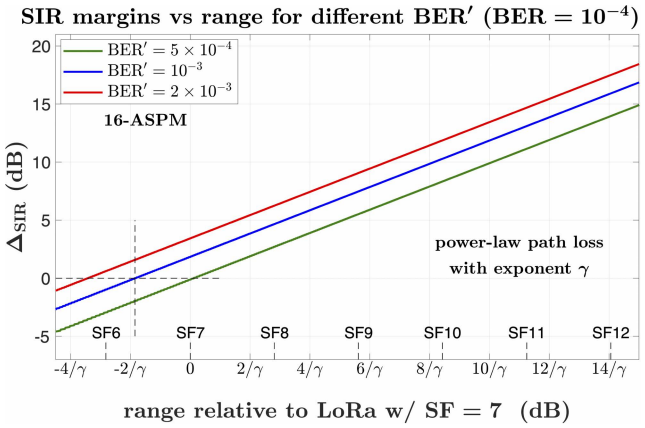

FIGURE 15. SIR margins for inter-PSF interference in M-ASPM increase with range ($\Delta_{\mathrm{SIR}} \propto N_{\mathrm{p}} \propto d^{\gamma}$ for sufficiently small BER).

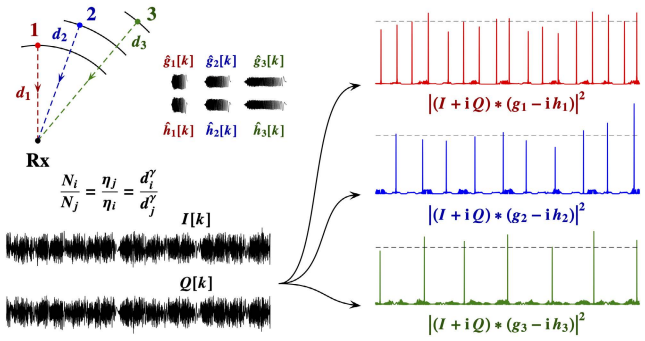

FIGURE 16. Illustration of impact of inter-PSF collisions for three 16-ASPM transmitters placed at particular distances from receiver. Respective spectral efficiencies are $\eta_1 = 1/30$, $\eta_2 = 1/50$, and $\eta_3 = 1/80$, and specific ranges shown are for free-space path loss. (These ranges roughly correspond to those from SF = 7 to mid-range between SF = 8 and SF = 9 for LoRa.)

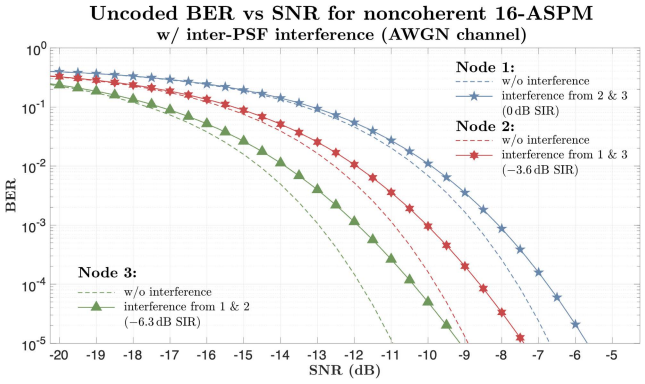

FIGURE 17. Uncoded BER vs. AWGN SNR for mutual interference example shown in Fig. 16.

The SIR can be related to the SNR Γ and the signal-to-interference-plus-noise ratio (SINR) Γ' as

$$\mathrm{SIR} = \frac{\Gamma\,\Gamma'}{\Gamma - \Gamma'}, \tag{25}$$

and the impact of the interference with a given SIR can be quantified by the deterioration (increase) in the BER when Γ becomes Γ'. If this interference increases the BER to $\mathrm{BER}' > \mathrm{BER}$, we may call the reciprocal of the respective SIR, i.e. $\Delta_{\mathrm{SIR}} = \mathrm{SIR}^{-1}$, an "SIR margin" for the given Γ, BER, and BER'. Then, as follows from the discussion in Section III-B, the SIR margins for M-ASPM can be determined from (25) and the condition

$$P_{\mathrm{b}}\left(\Gamma; M, N_{\mathrm{p}}\right) = \frac{\mathrm{BER}}{\mathrm{BER}'}\, P_{\mathrm{b}}\left(\Gamma'; M, N_{\mathrm{p}}\right) = \mathrm{BER}. \tag{26}$$

For example, for a power-law path loss $\Gamma \propto 1/d^{\gamma}$ and, for sufficiently small BER and BER', SIR margins for inter-PSF interference in M-ASPM increase with the range as $\Delta_{\mathrm{SIR}} \propto N_{\mathrm{p}} \propto d^{\gamma}$, becoming substantially high at long ranges (e.g., $\Delta_{\mathrm{SIR}} \gtrsim 10$ dB for the ranges beyond LoRa with SF = 10). This is quantified in Fig. 15.

Further, Figs. 16 through 19 provide two particular illustrations of mutual interference for multiple equal-power M-ASPM transmitters with different PSFs, placed at different distances from the receiver, such that the received BERs would be equal in the absence of the interference.

The average power of an i-th received pulse train with equal-magnitude pulses is inversely proportional to its

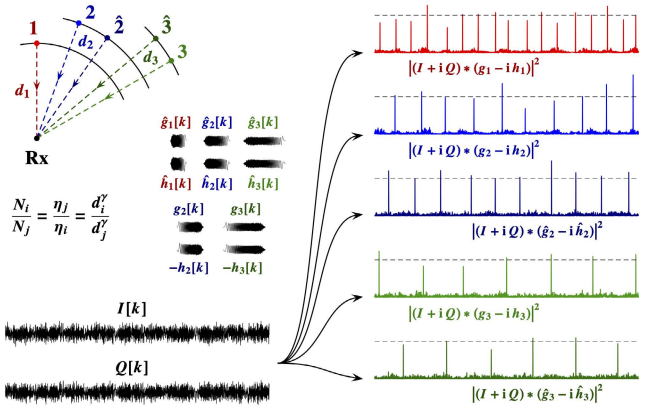

FIGURE 18. Illustration of impact of inter-PSF collisions for five 16-ASPM transmitters placed at particular distances from receiver. Respective spectral efficiencies are $\eta_1 = 1/30$, $\eta_2 = 1/50$, and $\eta_3 = 1/80$, and specific ranges shown are for free-space path loss. (These ranges roughly correspond to those from SF = 7 to mid-range between SF = 8 and SF = 9 for LoRa.)

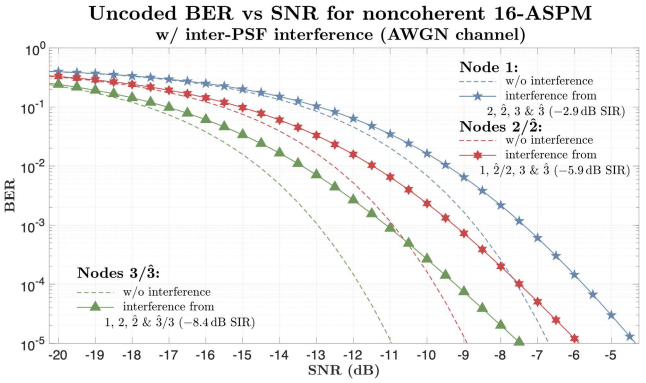

FIGURE 19. Uncoded BER vs. AWGN SNR for mutual interference example shown in Fig. 18.

average interpulse interval N_i. Then, for the power-law path loss, in the absence of interference the magnitudes of the pulses in the i-th and j-th received trains, and thus the received BERs, will be identical if $N_i/N_j = (d_i/d_j)^\gamma$. For the specific ranges shown in Figs. 16 and 18, the free-space path loss is assumed, and the horizontal dashed lines in the right-hand sides of these figures indicate the magnitudes of the interference-free pulses in the received pulse trains.

For larger interpulse intervals, we can use PSFs with proportionally larger TBPs, without increasing the PAPRs of the transmitted signals. This simplifies satisfying the large-TBP requirement for the PSF's cross-correlations. Also, for the nodes $\hat{2}$ and $\hat{3}$ in Fig. 18, we use the "flip" PSFs of the nodes 2 and 3, respectively.

Note that in Figs. 16 through 19 the shown impact of the interference is for continuous transmissions, when a full coincidence collision with the signals from all interfering transmitters is a certainty. This impact will be much less significant when the probability of such multiple coincidence collisions is rather small. Further, in these examples d_1 roughly corresponds to the range of LoRa with SF = 7, and d_3 corresponds to the mid-range between SF = 8 and SF = 9. For longer ranges, the SIR margins increase rather substantially (see Fig. 15), and the impact of the mutual interference will be much smaller even for continuous transmissions.

V. ACHIEVING DESIRED END NODE DISTRIBUTIONS IN M-ASPM LPWANs

In a design of a practical network, we may be given the coordinates of the end nodes placed at the desired locations, e.g., sensors co-located with traceable physical assets. In general, these coordinates can be time-variant, but we may initially assume that they vary sufficiently slowly and can be considered stationary during the ToA of any transmission. For a given placement of a gateway, the locations of the end nodes can be characterized by the distribution (density) function expressed in polar coordinates centered at the gateway, $\Phi(\varphi, r)$, where φ is the angular coordinate and $r > 0$ is the distance from the gateway.

A simple practical example of such a density function would be

$$\Phi(\varphi, r) = \frac{1}{4r\,\Delta s\,\Delta r\,N} \times \sum_{i=1}^{N} [\![\overline{|\varphi-\varphi_i|}\, r_i \le \Delta s]\!]\, [\![|r-r_i| \le \Delta r]\!]\, r_i, \quad (27)$$

where $\overline{|\varphi-\varphi_i|} = \min(|\varphi-\varphi_i|, 2\pi - |\varphi-\varphi_i|)$, (φ_i, r_i) are the coordinates of the i-th node, N is the total number of nodes, and the parameters Δr (distance increment) and Δs (arc length) represent the dimensions of an area element. Figure 20 shows an example of $\Phi(\varphi, r)$ obtained using (27) from the locations of the nodes for a given position of the gateway (at the intersection of the horizontal and vertical lines in the left-hand side of the figure).

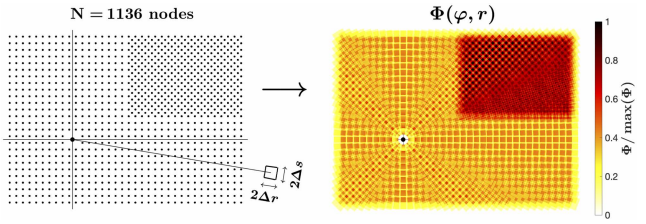

FIGURE 20. Example of representing discrete node locations as density function $\Phi(\varphi, r)$ of two continuous variables, φ and r.

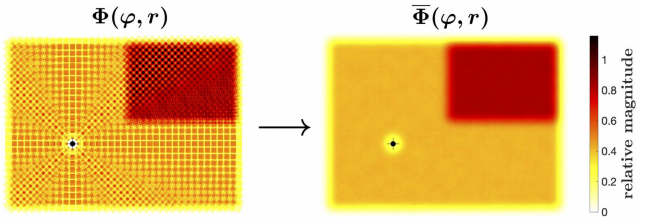

FIGURE 21. For "smooth" visual representation, $\Phi(\varphi, r)$ can be appropriately interpolated to obtain $\overline{\Phi}(\varphi, r)$. Both Φ and $\overline{\Phi}$ can be used interchangeably.

The main purpose for introducing such a density function is to characterize the aggregate of multiple node locations by a function of two *continuous* variables, φ and r, which simplifies optimization of the transmit and receive parameters of the nodes according to their spatial positions. If a "smooth" visual representation of the node distribution is also a goal, then the values of $\Phi(\varphi, r)$ can be appropriately interpolated (e.g., by 2D lowpass filtering) to obtain such a smooth representation $\overline{\Phi}(\varphi, r)$, as illustrated in Fig. 21. Both Φ and $\overline{\Phi}$ can be used interchangeably for obtaining the node parameters.

A. KEY IDEALIZATIONS

In order to focus on the most essential scaling properties of M-ASPM networks, we make a number of simplifying assumptions, which we will accept as valid unless explicitly stated otherwise. These simplifications include the recognition that various equalities and/or mathematical functions used below are approximations that represent quantities with only finite precision. Once a benchmark scaling model is established, it can be subsequently modified to account for various modifications, for example, for specific practical propagation channels and access protocols.

First, we will assume an uplink-focused network, with the main limitation on its capacity due to the uplink connections, when the end nodes are transmitters and the gateway is a receiver. All nodes transmit at the same power, and in the same frequency band. Next, we assume the "data equality" of the transmitting nodes, so that each node has the same data payload per unit time. In other words, the product of the data payload per transmission and the rate (average number per unit time) of transmissions for each node is a

constant. Note that the product of the transmission's ToA and the transmission rate is the *average duty cycle* of the node. Thus the equal-payload nodes with a smaller ToA per transmission will have a proportionally smaller average duty cycle.

To completely avoid co-PSF collisions for transmitting nodes with given PSFs, the product of the number of the nodes and their duty cycle must be smaller than unity. We shall denote this product as α, $0 < \alpha \leq 1$. When $\alpha = 1$, yet there are no collisions, the total combined power and data throughput for all same-PSF nodes will be equivalent to continuous transmission from a single node. For a random access protocol, the value of α that maximizes the throughput would be smaller than unity. For example, for pure ALOHA [27] the maximum throughput is achieved for $\alpha = 1/2$. For a given α, and given transmission rate and data payload per transmission, the available number of nodes with the same PSFs would be inversely proportional to the ToA of a transmission. When comparing the available number of the M-ASPM nodes with that of LoRa, we will assume the equality $\widetilde{\alpha} = \alpha$.

As before, whenever we compare M-ASPM and LoRa, we assume identical physical parameters of the links. For example, we assume the same physical frequency band, transmit power, antenna gains, and various system attenuations such as insertion and matching losses, etc.

Further, we will assume that the ToA per transmission is inversely proportional to the spectral efficiency of the node. For a constant-size data payload, this would typically hold for M-ASPM, i.e., ToA $\propto N_\mathrm{p}$, as the overhead (e.g. used for the header and synchronization) would normally utilize the same number of pulses for each transmission. For LoRa, if the number of overhead frames remains constant, then the number of overhead bits increases with the spreading factor, and the $\widetilde{\mathrm{ToA}}$ grows somewhat faster than the reciprocal of the spectral efficiency. However, for comparison with LoRa, we can assume that the data payload is large enough so that the number of LoRa overhead frames does not significantly affect the proportionality between the ToA and the reciprocal of the spectral efficiency.

Importantly, we will assume that a node can be placed only within its range, with the latter determined by the given constraint on the uncoded AWGN BER.

As discussed in Section III-B, for M-ASPM the range d is a monotonically increasing function of the average interpulse interval N_p. Inversely, for a given BER constraint, N_p is a nondecreasing function of the range. However, more generally, N_p would be a function of both φ and r, $N_\mathrm{p}(\varphi, r)$, as the path loss may depend on the angular coordinate φ, and N_p is a nondecreasing function of the range only for a given φ. For example, there may be obstacles in a certain direction from the receiver, or different multipath conditions. Then, e.g., for the average ToA of equal-payload nodes

$$\langle \mathrm{ToA} \rangle \propto \int_0^{2\pi} d\varphi \int_0^{\infty} dr\, r\, N_\mathrm{p}(\varphi, r)\, \Phi(\varphi, r). \tag{28}$$

If we assume that the path loss (and thus N_p) is independent of the direction φ, then (28) simplifies to

$$\langle \mathrm{ToA} \rangle \propto \int_0^{\infty} dr\, N_\mathrm{p}(r)\, \phi(r), \tag{29}$$

where $\phi(r)$ is the radial node density

$$\phi(r) = r \int_0^{2\pi} d\varphi\ \Phi(\varphi, r). \tag{30}$$

If $\Phi(\varphi, r)$ is given by (27), then $\phi(r)$ is simply

$$\phi(r) = \frac{1}{2N\Delta r} \sum_{i=1}^{N} [\![|r - r_i| \leq \Delta r]\!]. \tag{31}$$

B. NUMBER OF END NODES AT GIVEN DISTANCE FROM RECEIVER

If the inter-SF collisions in LoRa can be ignored, then the number of end nodes that can be placed at a given distance r from the receiver, $\widetilde{C}(r)$, would be limited only by the co-SF collisions, and can be expressed as

$$\widetilde{C}(r) = \sum_{i=1}^{n} [\![r \leq \widetilde{d}_i]\!]\, \widetilde{C}_i, \tag{32}$$

where $\widetilde{C}_i \propto \widetilde{\eta}(\widetilde{d}_i)$ is the number of end nodes that use the i-th spreading factor, $\widetilde{d}_i$ is the range for the i-th spreading factor, and n is the total number of the available spreading factors. That is, in LoRa the number of the end nodes that can be deployed at a given range is a (stepwise) decreasing function of this range.

In M-ASPM, the range d_i is an increasing function of the average interpulse interval N_p, which can be adjusted to match the given distance r from the receiver, $d_i = r$. If we employ m different PSF channels such that the inter-PSF interference is negligible, then

$$C(r) = \max_{\{d_1,\ldots,d_m\}} \sum_{i=1}^{m} [\![r \leq d_i]\!]\, C_1(d_i) = m\, C_1(r), \tag{33}$$

where $C_1(r)$ is the number of nodes for a single PSF channel.

In practice, the inter-PSF collisions can be ignored only for sufficiently small m, depending on the SIR margin for the given range. Then the number of the M-ASPM end nodes that can be placed at r can be expressed, for example, as

$$C(r) = \left\lceil \frac{\Delta_\mathrm{SIR}(r)}{\alpha} \right\rceil C_1(r), \tag{34}$$

where $\Delta_\mathrm{SIR}(r)$ is the SIR margin for the given target BER/BER$'$ (see Section IV-B and Fig. 15), and $\lceil x \rceil$ is the ceiling function.

Recall that for a sufficiently small target BER $\Delta_\mathrm{SIR}(r) \propto r^\gamma$, while $C_1(r) \propto \eta(r) \propto 1/r^\gamma$. Therefore, it follows from (34) that the increase in the M-ASPM SIR margins with range enables us to maintain the number of nodes at long ranges by employing a larger number of the PSF channels. As discussed in Section I, this is the desired target property of M-ASPM LPWANs, akin to the FDMA and CDMA-based

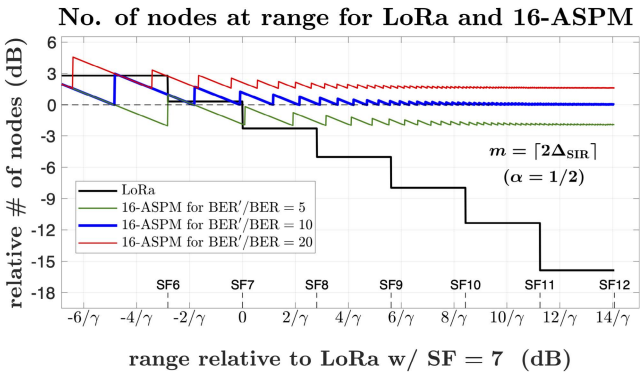

FIGURE 22. Number of end nodes that can be placed at given range in 16-ASPM and LoRa, under power-law path loss model (γ is path-loss exponent) and AWGN target BER = 10^{-4}. In LoRa, SF = 6 through SF = 12 are used, and inter-SF collisions are ignored. In 16-ASPM, number m of employed PSF channels varies according to SIR margins $\Delta_{\mathrm{SIR}}(r)$ for different BER′/BER.

approaches. Noticeably, Fig. 22 illustrates the contrast in the number of end nodes that can be placed at given range in 16-ASPM and LoRa. For example, for ranges beyond those of LoRa with SF = 11, 16-ASPM offers more than 40-fold larger network capacity, even when using rather conservative SIR margin constraints.

C. NETWORK RANGE EXTENSION FOR M-ASPM WITH MULTIPLE PULSE SHAPING FILTERS

For the end nodes placed over a wide area according to a given target distribution $\Phi(\varphi, r)$, we can obtain their average interpulse intervals in a manner that maximizes the network capacity.

Let us first consider the use of M-ASPM for extending the range of the network from d_0 to $d_0 + \Delta d$ when we can ignore the interference with the existing network (i.e., we can ignore collisions with the nodes at ranges below d_0).

If in the extended range we use the nodes with a given PSF $\hat{g}[k]$, then the extended capacity $\Delta\mathcal{C}_1$ (and hence the number of the nodes added to the network) will be constrained by co-PSF collisions. Since the ToA for a node is proportional to its average interpulse interval, if we use the nodes with $N_{\mathrm{p}} = N_{\mathrm{p}}(d_0 + \Delta d)$, then

$$\Delta\mathcal{C}_1 = \mathcal{C}_0 \frac{N_{\mathrm{p}}(d_0)}{N_{\mathrm{p}}(d_0 + \Delta d)}, \tag{35}$$

and for power-law path loss

$$\Delta\mathcal{C}_1 = \frac{\mathcal{C}_0}{\left(1 + \frac{\Delta d}{d_0}\right)^{\gamma}}. \tag{36}$$

In (35) and (36), $\mathcal{C}_0$ is the number of nodes that can be used with the interpulse interval $N_{\mathrm{p}}(d_0)$, limited by the co-PSF collisions. Unfavorably, the extended capacity $\Delta\mathcal{C}_1$ decreases with the increment in the range.

Instead, it would be beneficial to extend the range in several smaller increments, using different PSFs for each incremental extension.

As discussed in Section IV, with properly designed PSFs the impact of the inter-PSF collisions would be akin to the impact of Gaussian noise with the power equal to the combined power of the interfering signals at the receiver. The extent to which this impact can be considered insignificant (or easily compensable by relatively small adjustments in the interpulse intervals) would depend on the network geometry (including the desired space distribution of the nodes), the access protocol, channel conditions, and other factors. However, as follows from the discussion in Section IV-A, most practical range extensions would benefit from at least one intermediate extension step, which would more than double the number of the additional nodes. Further, since the SIR margins increase with the increase in the average interpulse interval, and thus with the increase in the range, a larger range extension can generally benefit from a larger number of the intermediate steps.

If the impact of inter-PSF is negligible, then the total number $\Delta\mathcal{C}_m$ of the nodes added through $m \geq 2$ incremental range extensions can be expressed as

$$m\,\Delta\mathcal{C}_1 < \Delta\mathcal{C}_m = \mathcal{C}_0 \sum_{i=1}^{m} \frac{N_{\mathrm{p}}(d_0)}{N_{\mathrm{p}}(d_i)}, \tag{37}$$

where m is the total number of range increments, $d_i = d_{i-1} + \Delta d_i$, and where $\sum_{i=1}^{m} \Delta d_i = \Delta d$ (i.e., $d_m = d_0 + \Delta d$). For power-law path loss (37) becomes

$$m\,\Delta\mathcal{C}_1 < \Delta\mathcal{C}_m = \mathcal{C}_0 \sum_{i=1}^{m} \left(\frac{d_0}{d_i}\right)^{\gamma}. \tag{38}$$

1) INCREMENTAL RANGE EXTENSION WITH DESIRED NODE DISTRIBUTION

For the nodes in the extended range (i.e., for the nodes with $r \geq d_0$ only), the density function can be normalized to unity as

$$\int_0^{2\pi} d\varphi \int_{d_0}^{d_0+\Delta d} dr\, r\, \Phi(\varphi, r) = 1. \tag{39}$$

Then the areal density ϱ_A of the nodes (i.e., the number of nodes per unit area, or the network capacity per unit area) is given by

$$\varrho_A(\varphi, r) = \Delta\mathcal{C}_m\, \Phi(\varphi, r), \tag{40}$$

where $\Delta\mathcal{C}_m$ is the total number of nodes in the extended range (see (37)).

Given the radial node density $\phi(r)$ and ignoring inter-PSF collisions, the desired rage increments can be obtained from the condition

$$\sum_{j=1}^{m} \frac{N_{\mathrm{p}}(d_i)}{N_{\mathrm{p}}(d_j)} \int_{d_{i-1}}^{d_i} dr\, \phi(r) = 1 \tag{41}$$

for any $i \in \{1, 2, \ldots, m\}$, which represents the equality of the total ToA for all nodes in each incremental range. Consequently, by solving the system of nonlinear equations represented by (41), we can obtain the incremental ranges d_i

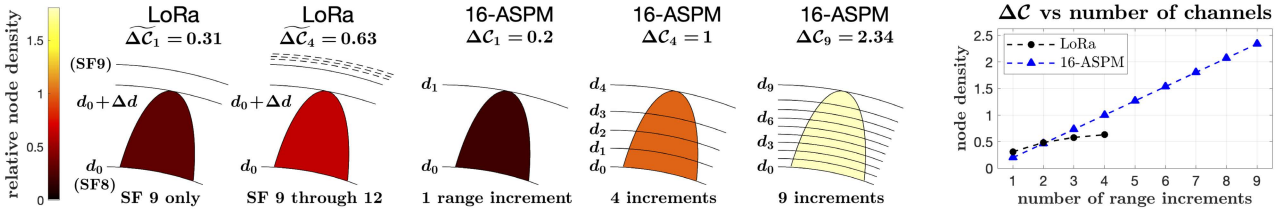

FIGURE 23. Increasing capacity of M-ASPM network through incremental range extension.

that allow us to preserve the desired node distribution $\Phi(\varphi, r)$. Note that for power-law path loss $N_{\rm p}(d_i) \propto d_i^{\gamma}$, and (41) becomes

$$\sum_{j=1}^{m} \left(\frac{d_i}{d_j}\right)^{\gamma} \int_{d_{i-1}}^{d_i} dr\, \phi(r) = 1. \tag{42}$$

2) NUMBER OF INCREMENTAL RANGES

For the power-law path loss, an estimate for the number of the incremental ranges m that can be used for efficient range extension can be obtained, based on the SIR margins, from the following constraint on the impact of inter-PSF collisions:

$$\langle \mathrm{SIR} \rangle_i^{-1} = \frac{\sum_{j=1}^{m} \langle r^{-\gamma} \rangle_j}{\langle r^{-\gamma} \rangle_i} - 1 \lesssim \frac{\Delta_{\mathrm{SIR}}\left(\langle r \rangle_i\right)}{\alpha}, \tag{43}$$

where

$$\langle f(r) \rangle_i = \frac{\int_{d_{i-1}}^{d_i} dr\, f(r)\, \phi(r)}{\int_{d_{i-1}}^{d_i} dr\, \phi(r)}. \tag{44}$$

In general, this number can be larger when a larger portion of the end nodes is placed at longer ranges, and vice versa.

When all nodes are placed at the same distance r from the receiver, then $\phi(r) = \delta(r)$ is the Dirac δ-function [28], and the condition (43) reduces to

$$m = \left\lceil \frac{\Delta_{\mathrm{SIR}}(r)}{\alpha} \right\rceil \leq 1 + \frac{\Delta_{\mathrm{SIR}}(r)}{\alpha}, \tag{45}$$

which is used to obtain (34) from (33).

For illustration of the effectiveness of the incremental range extension, in Fig. 23 we extend the range from d_0, corresponding to that of LoRa with SF $= 8$, to $d_0 + \Delta d$, which is within the range for SF $= 9$. The nodes are placed with uniform areal density within the boundary.

In LoRa, four SF channels (with SFs from 9 through 12) can be used for this range extension. However, using all four available LoRa SF channels only doubles the number of nodes, as compared with a single-channel extension with the highest spectral efficiency (SF $= 9$).

In contrast, the same number of channels (four), when used for incremental 16-ASPM range extension, increases the capacity by a factor of 5, as compared with a single-channel extension, and using 8 channels increases the capacity by more than an order of magnitude. In the example of Fig. 23, using 9 incremental ranges still satisfies the SIR margin constraints expressed by (43), for $\mathrm{BER}'/\mathrm{BER} = 10^{-3}/10^{-4}$, $\alpha = 1/2$, and full practical range of the values for the path-loss exponent γ.

D. SIMPLIFIED DESIRED AREAL COVERAGE

Although the incremental range extension minimizes the air-time and provides the maximum areal density of the nodes for a given distribution, it requires a reasonable degree of precision in the transmission settings according to the nodes' locations. Such precision may be difficult to obtain in practical deployments. In addition, the PSFs and the interpulse intervals of the nodes may need to be re-evaluated from time to time due to environmental changes, and/or due to relative motion of the nodes within the coverage area. Thus, in M-ASPM LPWANs that rely on the incremental range extension, a network managing protocol would need to be used to determine and dynamically adjust their settings. Instead, for simplicity of practical deployments, we may want to employ static and conservative transmission settings. An illustrative example of using such simplified settings for achieving uniform coverage within a hexagonal cell, centered at the gateway, is given in Fig. 24.

The transmitters within the inner circle of radius d_0 ("inner" nodes) will have the smallest interpulse interval and, therefore, the smallest SIR margin. Thus they may be highly susceptible to the interference from the nodes outside of this circle ("outer" nodes). At the same time, the nodes in close proximity to the gateway may produce high-power interference that can easily exceed the SIR margins of the outer nodes. However, as discussed in Section IV-A, if the i-th transmitter uses $\hat{h}_i[k] = H\left\{\hat{g}_i[k]\right\}$ while the j-th transmitter uses $\hat{h}_j[k] = -H\left\{\hat{g}_j[k]\right\}$, then the i-th and the j-th transmitted signals will occupy different (adjacent) frequency bands and will not interfere with each other. Then we can use one sideband for the inner nodes (within d_0), and the adjacent sideband for the outer nodes (between d_0 and $d_{\max}$).

In the example of Fig. 24, the inner nodes have the same PSFs and the interpulse intervals $N_{\rm p}(d_0)$. For the outer nodes, we use $m/2$ pairs of "flip" PSFs, and the interpulse intervals greater or equal to $N_{\rm p}(d_{\max})$. Further, we use the increment $\Delta N_{\rm p}$ in the interpulse interval between the j-th and $(j+1)$-th pairs. An increment in the interpulse interval allows us to raise the PSF's TBP without increasing the PAPR of the transmitted signal, and to ensure that all PSF's cross-correlations have sufficiently large TBPs.

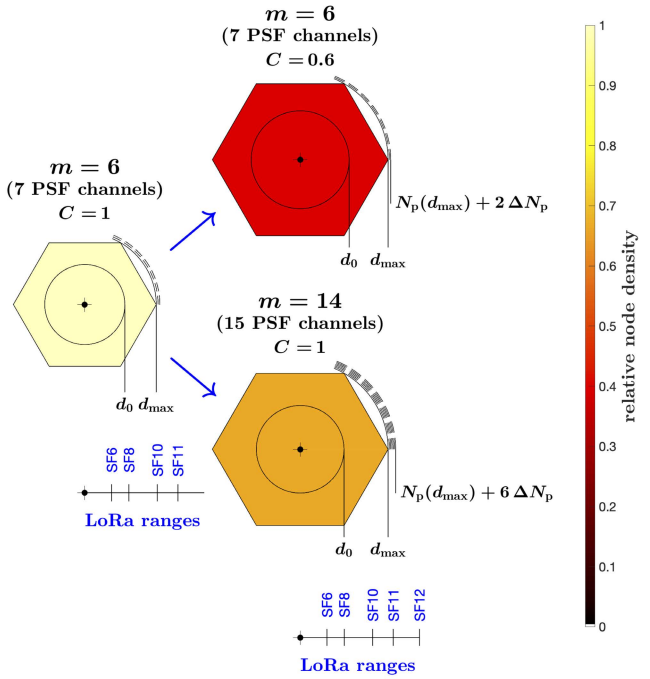

FIGURE 24. Example of using simplified conservative 16-ASPM transmission settings for achieving uniform coverage within hexagonal cell. "Proportional" range extension (with same number of PSF channels) results in decreased total capacity. Adding PSF channels (while remaining compliant with SIR margin constraints on m, as expressed by (47)) enables range extension without reducing total capacity (number of end nodes) within coverage area. SIR margin constraints are for BER′/BER = $10^{-3}/10^{-4}$ and $\alpha = 1/2$, and physical ranges shown are for path-loss exponent $\gamma = 2.6$.

The condition for achieving the desired areal density of the nodes can be expressed as

$$2\sum_{j=0}^{\frac{m}{2}-1}\frac{N_{\rm p}(d_0)}{N_{\rm p}(d_{\rm max})+j\,\Delta N_{\rm p}}=\frac{\int_{d_0}^{d_{\rm max}}dr\,\phi(r)}{\int_0^{d_0}dr\,\phi(r)}, \tag{46}$$

where m is constrained by the SIR margin at $d_{\rm max}$ as

$$m\lesssim 1+\frac{\Delta_{\rm SIR}(d_{\rm max})}{\alpha\left\langle\left(\frac{d_{\rm max}}{r}\right)^{\gamma}\right\rangle}, \tag{47}$$

and where

$$\left\langle\left(\frac{d_{\rm max}}{r}\right)^{\gamma}\right\rangle=\frac{\int_{d_0}^{d_{\rm max}}dr\,\left(\frac{d_{\rm max}}{r}\right)^{\gamma}\phi(r)}{\int_{d_0}^{d_{\rm max}}dr\,\phi(r)}. \tag{48}$$

From equations (46)–(48), we can then obtain the value of d_0. While such conservative settings do not necessarily maximize the capacity of the network, they may significantly simplify its deployment and management.

In the example of Fig. 24, the SIR margin constraints are for BER′/BER $= 10^{-3}/10^{-4}$ and $\alpha = 1/2$, and the physical ranges shown are for the path-loss exponent $\gamma = 2.6$. These values lead to $m = 6$ for the smallest-area hexagon (left), and to $m = 14$ for the larger hexagons (which have approximately 50% larger area).

As can be seen in Fig. 24, a "proportional" range extension (using the same number of the PSF channels) results in a decreased total capacity (down to about 60% in this example). Combined with the extended areal coverage, it leads to a fast deterioration in the node density (declining below 40% of the small-sell density). However, using additional PSF channels, while remaining compliant with the SIR margin constraints on m expressed by (47), maintains the total capacity (the number of end nodes) within the coverage area.

VI. COMPLEMENTARY USE OF LoRa AND M-ASPM

As has been discussed so far, the main advantage of M-ASPM over LoRa lies in its better scalability: As a "true" spread-spectrum technique, M-ASPM has the ability to better sustain the network capacity when extending its range, and to accommodate various desired areal distributions of the end nodes with minimal impact on the total number of nodes. At the same time, LoRa may have an edge in a number of technical characteristics that are important in a variety of particular LPWAN use cases.

For example, the average energy efficiency of the 16-ASPM nodes placed, in some distributed fashion, between the LoRa ranges for SF = 9 and SF = 12 would be in mid-range between 50% and 100% of that for LoRa, depending on the actual areal distribution of the nodes and the path-loss exponent. (This can be deduced from examining the spectral efficiency vs. range dependencies for LoRa and 16-ASPM shown in Fig. 9.) Combined with the "true" constant-envelope property of the LoRa modulation, it can result in significant energy savings when using LoRa, which is especially important for battery-powered end nodes.

However, for M-ASPM and LoRa operating in the same spectral band, the inter-SF, inter-PSF, and the "SF-PSF" collisions would have comparable impacts. Thus, favorably, LoRa and M-ASPM can be designed to concurrently operate in the same spectral band and geographical area, cooperatively complementing each other's coverage.

Say, our goal is a wide-area coverage with the desired end node density $\Phi_0(\varphi, r)$, which is relatively high beyond the range for SF = 8. However, for a LoRa network operating at full capacity, most of the nodes must be placed within this range. Thus, due to its limitations on the relative number of nodes at long ranges, LoRa may not be able to accommodate this goal without reducing the total capacity of the network. Instead, a full-capacity LoRa coverage can provide the density $\widetilde{\Phi}(\varphi, r)$, while M-ASPM, which is not a subject to these limitations, can complement this coverage to achieve the desired density $\Phi_0(\varphi, r)$.

The 16-ASPM node density $\Phi(\varphi, r)$ that complements $\widetilde{\Phi}(\varphi, r)$ to $\Phi_0(\varphi, r)$ can be expressed as

$$\Phi(\varphi,r)=\left(1+\frac{\widetilde{\mathcal{C}}}{\mathcal{C}}\right)\Phi_0(\varphi,r)-\frac{\widetilde{\mathcal{C}}}{\mathcal{C}}\,\widetilde{\Phi}(\varphi,r), \tag{49}$$

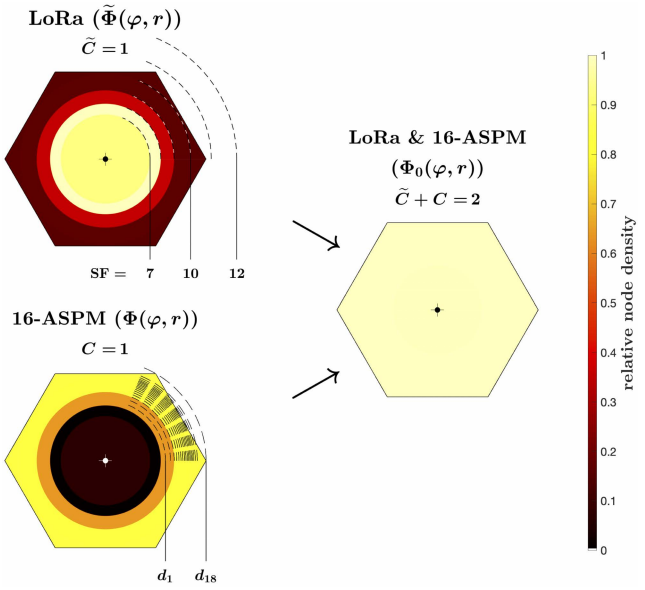

FIGURE 25. Illustrative example of complementary use of LoRa and M-ASPM for achieving uniform node density in large hexagonal cell centered at gateway. Physical ranges shown are for path-loss exponent $\gamma = 3$.

where $\widetilde{\mathcal{C}}$ and $\mathcal{C}$ are the total numbers of nodes for LoRa and 16-ASPM, respectively. Then we can use

$$\Phi'(\varphi, r) = [\![\Phi(\varphi, r) \geq 0]\!]\, \Phi(\varphi, r) \tag{50}$$

as the target 16-ASPM node density in an iterative procedure (subject to the constraints on the SIR margins) to obtain the 16-ASPM incremental ranges, both their number and their values, that enable us to achieve the required complementary 16-ASPM density $\Phi(\varphi, r)$.

An illustrative example of such complementary use of LoRa and M-ASPM is provided in Fig. 25. Here, the goal is to achieve uniform node density within a hexagonal cell centered at the gateway. (The specific physical ranges shown in the figure are for the path-loss exponent $\gamma = 3$.) For LoRa (the density function $\widetilde{\Phi}(\varphi, r)$), the area within the "effective" range (for SF = 9), containing over 87% of the total capacity, is only about half of the total area of the cell (and thus the coverage area). In contrast, the opposite is true for 16-ASPM (the density function $\Phi(\varphi, r)$): For essentially the same capacity, over 96% of the nodes lie outside the range for SF = 8. As a result, together LoRa and 16-ASPM ensure coverage with uniform density within the cell ($\Phi_0(\varphi, r)$), providing twice the capacity of the individual networks.

For simplicity, for both LoRa and 16-ASPM in Fig. 25, we ignore the impact of the high-power interference from the nodes in close proximity to the gateway with the low-power nodes placed at the ranges beyond SF = 7, and still represent as uniform the areal densities of the nodes within the range for SF = 7.

VII. CONCLUSION AND FUTURE DIRECTIONS

M-ary ASPM is a recently introduced modulation technique and many of its promising features have not yet been explored and/or quantified. Its description was previously provided only in [8], where we evaluate the bit error probability for coherent and noncoherent M-ASPM links in an AWGN channel. For a more comprehensive presentation, we have incorporated the main results of [8] into this paper, along with a brief discussion of improving such M-ASPM link properties as their resistance to multipath delay and Doppler spreads, and to impulsive noise commonly present in industrial environments. However, here our primary focus was on quantifying M-ASPM networks' scalability, that is, on the spread-spectrum properties of M-ASPM.

As discussed in Section III, the smallest average interpulse interval N_p, and thus the largest spectral efficiency η, that can be used in M-ASPM is constrained by the chosen value of M. For example, for the ACF as an RC pulse with the roll-off factor $\beta = 0$, and the sampling rate used in the paper, $N_\mathrm{p} \geq M$ for coherent M-ASPM and $N_\mathrm{p} \geq 4M$ – for noncoherent. For larger values of N_p, the M-ASPM's spectral efficiency is decoupled from the value of M and, for a given M, η becomes inversely proportional to N_p: $\eta \propto 1/N_\mathrm{p}$. Then, since the bit error probability is a monotonically decreasing function of the ratio of the SNR Γ and the spectral efficiency η, for a given SNR the M-ASPM's BER becomes a monotonically decreasing function of N_p. Consequently, as the M-ASPM's receiver sensitivity becomes proportional to the average interpulse interval, the desired BER can be achieved, for any M and SNR, by changing N_p.

In other words, when M-ASPM is used as a spread-spectrum technique (that is, when it operates at the spectral efficiencies below the maximum for a given M), its processing gain is proportional to the average interpulse interval N_p. As a result, N_p directly affects such link properties as its ToA, the SIR margin Δ_SIR and, for a given transmit power, the range d: ToA $\propto N_\mathrm{p}$, $\Delta_\mathrm{SIR} \propto N_\mathrm{p}$, and $d \propto {N_\mathrm{p}}^{1/\gamma}$. In particular, both the ToA and the SIR margin are proportional to N_p, which enables us to maintain the M-ASPM network's capacity while extending its range. Importantly, we demonstrate that such capacity-preserving range extension can be achieved for numerous desired areal distributions of the uplink nodes.

To establish the inverse proportionality between η and N_p in M-ASPM, we need to operate at the spectral efficiencies smaller than the maximum spectral efficiency for a given M, $\eta_\mathrm{max}(M)$ (see Fig. 7). That was one of the motivations for choosing $M = 16$ in most of the quantitative examples of this paper, as $\eta_\mathrm{max}(16)$ is 8/3 times higher than the spectral efficiency of LoRa with SF = 6 (for noncoherent detection), and thus the relation $d \propto \eta^{-1/\gamma} \propto {N_\mathrm{p}}^{1/\gamma}$ for 16-ASPM holds for the ranges above a fraction of the range for LoRa with SF = 6. This streamlined our presentation and simplified obtaining numerical solutions of the nonlinear equations. However, the price for this simplification is somewhat lower

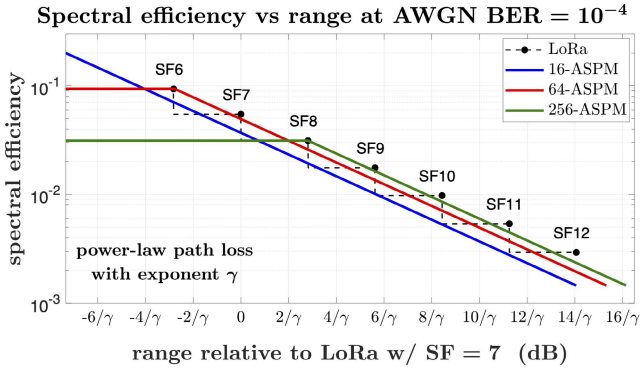

FIGURE 26. Spectral efficiency vs range for LoRa and M-ASPM (noncoherent detection).

16-ASPM overall spectral efficiency, and thus proportionally lower energy efficiency, as compared to LoRa for ranges above those with SF $=$ 7. Yet, despite such lower spectral efficiency and thus longer ToA, which exacerbates collisions, 16-ASPM offers significant increase in the network capacity for the ranges beyond LoRa's effective range.

Further, as illustrated in Fig. 26, larger values of M can be used in M-ASPM at longer ranges, to increase its spectral and energy efficiencies while maintaining the inverse proportionality between η and N_p. For example, for ranges above those for LoRa with SF $=$ 6, 64-ASPM offers approximately a factor of 3/4 reduction in the ToA, and thus in the energy consumption of the uplink nodes, over 16-ASPM. (Note that, while increasing M impacts the computational intensity of the signal processing in the receiver, with the scaling dependent on the implementation algorithm, for the pulse-position encoding there is no computational penalty on the transmitter's signal processing for large M.) As a result (and after the adjustment for a small decrease in the SIR margins), for ranges beyond those of LoRa with SF $=$ 11, 64-ASPM offers about 50-fold larger network capacity than LoRa, as compared with only 40-fold capacity increase provided by 16-ASPM (as discussed in Section V-B). Favorably, this is also accompanied by a slight reduction in the total energy consumption of the uplink nodes.

As can be seen in Fig. 26, for a given payload the ToA for 16-ASPM would be generally longer than for LoRa for ranges above SF $=$ 8, yet smaller than LoRa's for ranges below SF $=$ 8. However, for 256-ASPM the opposite is true. At the same time, for the ranges between SF $=$ 7 and SF $=$ 8 the spectral efficiencies of LoRa and 256-ASPM are identical. Such reduction in the nodes' ToA for larger M, combined with the increase in the total network capacity, also widens the appeal of complementary use of LoRa and M-ASPM outlined in Section VI, as it enables the extension of LoRa's coverage for longer ranges and/or different environments and desired areal node distributions.

To maintain the focus on the most essential scaling properties of M-ASPM networks, in this paper we have made a number of simplifying assumptions. For example, we assumed transmissions with a constant average power, without discussing the effect of different PSFs on the PAPR of the modulated signal, and thus the efficiency of the power amplifier (PA). For such a constant-power transmission, $d \propto N_\mathrm{p}{}^{1/\gamma}$ and the range is controlled, for a given M, by the data rate only.

However, for M-ASPM with constant-envelope pulses described in Section II-A, constraining the peak rather than the average power enables us to control the transmit power without sacrificing the PA efficiency. Then the average transmit power is proportional to the "pulse duty cycle" $D = L/N_\mathrm{p} < 1$, where L is the length of the pulse expressed by (6), and the range becomes a function of the product of D and N_p: $d \propto (D\,N_\mathrm{p})^{1/\gamma}$. The pulse duty cycle controls the average transmit power, while N_p controls the spectral efficiency (the data rate) and the SIR margins. This adds to the flexibility of achieving energy-efficient range extensions for numerous desired areal distributions of the network nodes, and is an example of many M-ASPM features yet to be considered and explored.

Overall, given the enormous diversity and the rapid growth of the IoT uses, it is extremely challenging to design a single LPWAN solution that would equally satisfy every need. However, by consolidating various benefits of different LPWAN approaches, M-ASPM allows extensive versatility in trading multiple PHY parameters to reconcile often conflicting LPWAN technical constraints. Therefore, the M-ASPM approach may offer a significant step towards the development of such a unified solution, and it merits further exploration.

APPENDIX A
ACRONYMS

ACF: autocorrelation function; A/D: Analog-to-Digital; ASPM: Aggregate Spread Pulse Modulation; AWGN: Additive White Gaussian Noise; BER: Bit Error Rate; CDMA: Code Division Multiple Access; CSS: Chirp Spread Spectrum; FDMA: Frequency-Division Multiple Access; INF: Intermittently Nonlinear Filtering; IoT: Internet of Things; LoRa: Long Range (CSS-based modulation technique for LPWANs); LoRaWAN: Long Range Wide Area Network; LPWAN: Low-Power Wide Area Network; M-ASPM: M-ary ASPM; NB-IoT: narrowband IoT; PA: Power Amplifier; PAPR: Peak-to-Average Power Ratio; PHY: physical layer; PSD: Power Spectral Density; PSF: Pulse Shaping Filter; RC: Raised-Cosine; SC-FDMA: Single-Carrier FDMA; SF: Spreading Factor (for LoRa); SIR: Signal-to-Interference Ratio; SINR: Signal-to-Interference-plus-Noise Ratio; SNR: Signal-to-Noise Ratio; SS: Spread Spectrum; TBP: Time-Bandwidth Product; ToA: Time-on-Air.

APPENDIX B
COMMENTS ON NOTATIONS

Whenever a particular notation is introduced in the paper, it is immediately defined. Some notations are used only once. The notations that are used consistently (and more that once) throughout the paper include:

α	product of number of nodes with given PSFs and their average transmission duty cycle
B	bandwidth
β	roll-off factor of RC pulse
C	network capacity or number of end nodes (with or without subscripts)
d	range or distance between transmitter and receiver (with or without subscripts)
Δ_{SIR}	SIR margin
E_{b}	energy per bit
η	spectral efficiency (with or without subscripts)
f_{b}	bit rate
f_{p}	pulse rate
F_{s}	sample rate
φ	angular coordinate (in density function; with or without subscripts)
$\phi(r)$	radial node density
$\Phi(\varphi, r)$	node density function expressed in polar coordinates centered at gateway
γ	path-loss exponent (in power-law path loss)
Γ	SNR
Γ'	SINR
k	sample index (in digital signal representations)
M	number of states in M-ary encoding
N_0	one-sided noise PSD
N_{p}	average interpulse interval
N_{s}	oversampling factor
P_{b}	bit error probability
r	distance from receiver/gateway (radial coordinate in density function)

Some notations may have different contextual meaning in different sections of the paper. This change normally affects the letters commonly representing integer numbers (e.g., "i" and "j"), such as subscripts and/or summation indices. For example, in Section II, the letter "j" is exclusively used as the (integer) number indicating a particular pulse in the designed pulse sequence $\hat{x}[k]$ (e.g., k_j is the sample index of the j-th pulse, A_j is the amplitude of the j-th pulse, and m_j is its "state"). However, in the subsequent sections j may relate to the j-th transmitter, or j-th range increment, or j-th pair of "flip" PSFs. Whenever such change occurs, the respective clarification is provided.

Throughout the paper, in the mathematical notations we reserve the letters "g" and "h" for pulse shaping filters. For example, we denote the finite impulse response of a PSF applied to a designed pulse train as $\hat{g}[k]$, where k is the sample index. As the focus in this paper is on the single-sideband M-ASPM, the PSFs g and h are related to each other through the Hilbert transform, e.g., $h(t) = \pm H(g)(t)$ (in analog domain) or $h[k] = \pm H\{g[k]\}$ (in digital representation).

Further, we find it convenient to use the "hat" operator for $\hat{g}[k]$ and $\hat{h}[k]$ to distinguish between these PSFs and their respective matched filters $g[k] = \hat{g}[-k]$ and $h[k] = \hat{h}[-k]$. We also use the hat symbol in Sections II and IV to denote the designed pulse trains $\hat{x}[k]$, $\hat{x}_1[k]$, and $\hat{x}_2[k]$, as opposed to the shaped trains obtained by applying PSFs to the designed pulse sequences.

While "N_{p}" is used for the average interpulse interval throughout the paper, in Section IV-B (including Figs. 16 and 18), the average interpulse intervals for the i-th and j-th transmitters are also denoted as N_i and N_j, respectively.

REFERENCES

[1] N. Naik, "LPWAN technologies for IoT systems: Choice between ultra narrow band and spread spectrum," in *Proc. IEEE Int. Syst. Eng. Symp. (ISSE)*, Oct. 2018, pp. 1–8.
[2] B. S. Chaudhari, M. Zennaro, and S. Borkar, "LPWAN technologies: Emerging application characteristics, requirements, and design considerations," *Future Internet*, vol. 12, no. 3, p. 46, Mar. 2020. [Online]. Available: https://www.mdpi.com/1999-5903/12/3/46
[3] C. B. Mwakwata, H. Malik, M. M. Alam, Y. Le Moullec, S. Parand, and S. Mumtaz, "Narrowband Internet of Things (NB-IoT): From physical (PHY) and media access control (MAC) layers perspectives," *Sensors*, vol. 19, no. 11, p. 2613, Jun. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/11/2613
[4] A. Kumar, D. Manjunath, and J. Kuri, *Wireless Networking*. San Mateo, CA, USA: Morgan Kaufmann, 2008.
[5] D. Torrieri, *Principles of Spread-Spectrum Communication Systems*, 4th ed. Boston, MA, USA: Springer, 2018.
[6] L. Vangelista, "Frequency shift chirp modulation: The LoRa modulation," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1818–1821, Dec. 2017.
[7] G. Baruffa, L. Rugini, L. Germani, and F. Frescura, "Error probability performance of chirp modulation in uncoded and coded LoRa systems," *Digit. Signal Process.*, vol. 106, Nov. 2020, Art. no. 102828.
[8] A. V. Nikitin and R. L. Davidchack, "M-ary aggregate spread pulse modulation in LPWANs for IoT applications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Madrid, Spain, Dec. 2021, pp. 1–7.
[9] A. V. Nikitin and R. L. Davidchack, "Pulsed waveforms and intermittently nonlinear filtering in synthesis of low-SNR and covert communications," *IEEE Access*, vol. 8, pp. 173250–173266, 2020.
[10] A. V. Nikitin, "Method and apparatus for nonlinear filtering and for secure communications," U.S. Patent 11 050 591, Jun. 29, 2021.
[11] D. E. Knuth, "Two notes on notation," *Amer. Math. Monthly*, vol. 99, no. 5, pp. 403–422, May 1992.
[12] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng. III, Radio Commun. Eng.*, vol. 93, no. 3, pp. 429–457, Nov. 1946.
[13] M. Vetterli and J. Kovačevic, *Wavelets and Subband Coding*. Upper Saddle River, NJ, USA: Prentice-Hall, 1995.
[14] B. Picinbono, "On instantaneous amplitude and phase of signals," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 552–560, Mar. 1997.
[15] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 4th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2006.
[16] R. N. Bracewell, *The Fourier Transform and its Applications*, 3rd ed. New York, NY, USA: McGraw-Hill, 2000.
[17] G. Todoran, R. Holonec, and C. Iakab, "Discrete Hilbert transform. Numeric algorithms," *Acta Electroteh.*, vol. 49, no. 4, pp. 485–490, 2008.
[18] A. V. Nikitin and R. L. Davidchack, "Analog-domain mitigation of outlier noise in the process of analog-to-digital conversion," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kansas City, MO, USA, May 2018, pp. 20–24.
[19] A. V. Nikitin and R. L. Davidchack, "Hidden outlier noise and its mitigation," *IEEE Access*, vol. 7, pp. 87873–87886, 2019.
[20] A. V. Nikitin and R. L. Davidchack, "Complementary intermittently nonlinear filtering for mitigation of hidden outlier interference," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Norfolk, VA, USA, Nov. 2019, pp. 12–14.
[21] A. V. Nikitin and R. L. Davidchack, "Quantile tracking filters for robust fencing in intermittently nonlinear filtering," 2019, *arXiv:1911.00736*.
[22] J. Courjault, B. Vrigneau, O. Berder, and M. R. Bhatnagar, "How robust is a LoRa communication against impulsive noise?" in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2020, pp. 1–6.

[23] A. E. Ferreira, F. M. Ortiz, L. H. M. K. Costa, B. Foubert, I. Amadou, and N. Mitton, "A study of the LoRa signal propagation in forest, urban, and suburban environments," *Ann. Telecommun.*, vol. 75, nos. 7–8, pp. 333–351, Aug. 2020, doi: 10.1007/s12243-020-00789-w.
[24] R. Ramesh, M. Arunachalam, H. K. Atluri, C. Kumar, S. V. R. Anand, P. Arumugam, and B. Amrutur, "LoRaWAN for smart cities: Experimental study in a campus deployment," in *LPWAN Technologies for IoT and M2M Applications*, B. S. Chaudhari and M. Zennaro, Eds. New York, NY, USA: Academic, 2020, pp. 327–345.
[25] M. Rademacher, H. Linka, T. Horstmann, and M. Henze, "Path loss in urban Lora networks: A large-scale measurement study," in *Proc. IEEE 94th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2021, pp. 1–6.
[26] D. Zwillinger, *Standard Mathematical Tables and Formulae*, 30th ed. Boca Raton, FL, USA: CRC Press, 1996, pp. 547–550.
[27] N. Abramson, "The ALOHA system: Another alternative for computer communications," in *Int. Workshop Manag. Requirements Knowl. (AFIPS)*, Houston, TX, USA, Nov. 1970, pp. 281–285.
[28] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. London, U.K.: Oxford Univ. Press, 1958.

ALEXEI V. NIKITIN (Member, IEEE) is a co-founder and Chief Science Officer of the Kansas-based Nonlinear LLC. He initiated his undergraduate and graduate studies in physics, chemistry, and engineering in the former USSR at Novosibirsk State University in Novosibirsk and Karpov Institute of Physical Chemistry in Moscow. After receiving a Ph.D. degree in physics from the University of Kansas in 1998, he led R&D work focused on methods and tools in nonlinear signal processing at several startup companies, some of which were subsequently acquired, specializing in applications in communications, power electronics, navigation, geophysical sciences, neurology, and biometrics. He is named as a first or sole inventor on over 30 issued U.S. patents.

RUSLAN L. DAVIDCHACK is a Professor of Mathematical Modeling and Computation in the School of Computing and Mathematical Sciences at the University of Leicester, U.K. His research interests are in developing computational methods and tools for applications in molecular simulations, nonlinear dynamics and signal processing.

He received his undergraduate degree in theoretical condensed matter physics from Lviv University, Ukraine, and the Ph.D. degree in computational statistical physics from the University of Kansas. He has developed several unique computational approaches and made significant contributions to a variety of fields in computational chemistry, signal processing, mathematics, statistical physics, and biochemistry. He has an extensive track record of interdisciplinary collaborations which span academic institutions (e.g., Georgia Tech; UC Berkeley; Oxford, U.K.; EPFL, Switzerland; NTNU, Norway; KTH, Sweden), national laboratories in the USA (e.g., AFRL, Oak Ridge, Sandia), and U.K. (Daresbury, Institute of Pharmaceutical Innovation), and industrial partners (Tata Steel, Rolls-Royce, GlaxoSmithKline, Procter & Gamble). He has given invited talks at many international conferences in the USA, U.K., France, Germany, South Korea, China, and Ukraine. He is a co-inventor of the AVATAR methodology, which provided the initial foundation for the technology of intermittently nonlinear outlier noise mitigation.