

Received August 20, 2020, accepted September 15, 2020, date of publication September 18, 2020, date of current version October 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024978

# Pulsed Waveforms and Intermittently Nonlinear Filtering in Synthesis of Low-SNR and Covert Communications

ALEXEI V. NIKITIN<sup>1</sup>, (Member, IEEE), AND RUSLAN L. DAVIDCHACK<sup>2</sup>

<sup>1</sup>Nonlinear LLC, Wamego, KS 66457, USA

<sup>2</sup>School of Mathematics and Actuarial Science, University of Leicester, Leicester LE1 7RH, U.K.

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

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

**ABSTRACT** In traditional spread-spectrum techniques, a wideband transmit signal is obtained by modulating a wideband carrier by a narrowband signal containing a relatively low-rate message. In the receiver, the respective demodulation/despreading restores the information-carrying narrowband signal. In this paper, we introduce an alternative approach, where the low-rate information is encoded directly into a wideband waveform of a given bandwidth, without physical “spreading” of the carrier’s frequency. The main advantages of this approach lie in extended options for encoding the information, and in retaining a reversible control over the temporal and amplitude structures of the modulating wideband waveforms. Significant “excess bandwidth” (over that needed to carry the information) enables us to use allpass filters to manage statistical properties and time-domain appearances of these waveforms without changing their spectral composition. For example, a mixture of transmitted waveforms can be shaped as a low-crest-factor signal (e.g. to reduce the burden on the power amplifier), and/or made statistically indistinguishable from Gaussian noise (e.g. for covert transmissions and physical layer steganography), while the selected components of the received waveform can be transformed into high-crest-factor pulse trains suitable for multiplexing and/or low-SNR communications. Further, control over the temporal and amplitude structures of wideband waveforms carrying low-rate information enables effective use of nonlinear filtering techniques. Such techniques can be employed for robust real-time asynchronous extraction of the information, as well as for separation of wideband signal components with identical spectral content from each other. This can facilitate development of a large variety of low-SNR and covert communication configurations.

**INDEX TERMS** Aggregate spread pulse modulation, covert communications, hard-to-intercept communications, low-power communications, intermittently nonlinear filtering, physical layer, pileup effect, spread spectrum, steganography.

## I. INTRODUCTION

The additive white Gaussian noise (AWGN) capacity  $C$  of a channel operating in the power-limited regime (i.e. when the received signal-to-noise ratio (SNR) is small,  $\text{SNR} \ll 0\text{dB}$ ) can be expressed as  $C \approx \bar{P}/(N_0 \ln 2)$ , where  $\bar{P}$  is the average received power and  $N_0$  is the power spectral density (PSD) of the noise [1]. This capacity is linear in power and insensitive to bandwidth and, therefore, by spreading the average transmitted power of the information-carrying signal over a large frequency band, the average PSD of the signal could be made much smaller than the PSD of the

noise. This would “hide” the signal in the channel noise, making the transmission covert and insensitive to narrowband interference.

Various techniques to achieve such “spreading” are commonly referred to as *spread-spectrum*. These include such well known approaches as frequency-hopping spread spectrum (FHSS) and direct-sequence spread spectrum (DSSS) [2], [3], as well as chirp spread spectrum (CSS) [4], [5]. Within these techniques, a narrowband signal in the transmitter modulates a carrier that spans a wide frequency range. In the receiver, the respective demodulation, combined with despreading, is used to produce the information-carrying narrowband signal. Thus, even though the total SNR of the wideband transmitted signal can be low, obtaining the

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

information about the carrier enables us to detect the presence of a narrowband signal through spectral measurements. For example, the FHSS is widely used in legacy military equipment for low-probability-of-intercept (LPI) communications. However, using frequency hopping for covert communications is nearly obsolete today, since modern wideband software-defined radio (SDR) receivers can capture all of the hops and put them back together (J. E. Gilley, personal communication, Feb. 9, 2020).

In DSSS, a wideband unit-power pseudorandom signal, known as a spreading sequence, is modulated by a narrow-band information-carrying signal of a given power. After demodulation/despreading in the receiver, the original information-carrying signal is restored. However, such demodulation requires precise synchronization, which is perhaps the most difficult and expensive aspect of a DSSS receiver design. Also, while despreading cannot be performed without the knowledge of the spreading sequence by the receiver, the spreading code by itself may not be usable to secure the channel. For example, linear spreading codes are easily decipherable once a short sequential set of chips from the sequence is known. To improve security, it would be desirable to perform a “code hopping” in a manner akin to the frequency hopping. However, synchronization can be an extremely slow process for pseudorandom sequences, especially for large spreading waveforms (long codes), and thus such DSSS code hopping may be difficult to realize in practice.

In the power-limited regime, we would normally use binary coding and modulation (e.g. binary phase-shift keying (BPSK) or quadrature phase-shift keying (QPSK)) for the narrow-band information-carrying signal, and this signal will be significantly oversampled to enable wideband spreading. Thus an idealized narrow-band information-carrying signal that is to be “spread” can be viewed as a discrete-level signal that is a linear combination of analog Heaviside unit step functions [6] delayed by multiples of the bit duration. Such a signal would have a limited bandwidth and a finite power. Since the derivative of the Heaviside unit step function is the Dirac  $\delta$ -function [7], the derivative of this idealized signal will be a “pulse train” that is a linear combination of Dirac  $\delta$ -functions. This pulse train will contain all the information encoded in the discrete-level signal, but it will have infinitely wide bandwidth and infinitely large power. Both the bandwidth and the power can then be reduced to the desired levels by filtering the pulse train with a lowpass or bandpass filter. As discussed in Section II, if the time-bandwidth product (TBP) of the filter is sufficiently small so that the pulses in the filtered pulse train do not overlap, these pulses will still contain all the intended information. Such an approach is the basis for various “impulse,” “carrier-free,” and “baseband” communication and radar systems which are collectively referred to as “ultra wideband” (UWB) [8].

Relaxing the current technical (and mostly regulatory) limitations imposed on the term “UWB” (e.g., that the emitted signal bandwidth exceeds the lesser of 500 MHz or 20% of

the arithmetic center frequency [9]), the UWB concept can be extended to a pulsed signal in any provided frequency band. For instance, a baseband pulse train confined to a given physical band can be used for modulation of a single-frequency carrier. Then the relative bandwidth of the resulting transmit signal can be made arbitrary “wide” (e.g. larger than 20% of the carrier frequency) or “narrow” (e.g., <1%). If the pulse repetition rate in such a train is much smaller than its bandwidth, then such an approach can be considered a “spread-spectrum technique.”

On the one hand, transmitting low-rate information by a wideband pulsed waveform (pulse train) has an apparent appeal of no need for despreading: After demodulation, one can simply “capture” the pulses (e.g. their amplitudes, polarities, and/or interarrival times) to obtain the encoded information. On the other hand, at first glance such a pulse train is not suitable for use as a modulating signal in practical communication systems, especially for covert communications. Indeed, let us consider a pulse train with a given average pulse rate and power. The average PSD of this train can be made arbitrary small, since it is inversely proportional to the bandwidth. However, the peak-to-average power ratio (PAPR) of such a train would be proportional to the bandwidth, making the wideband signal extremely impulsive (super-Gaussian). This leads to several considerable challenges in adapting such a signal to covert transmissions. Firstly, a high crest factor of the pulse train can put a serious burden on the transmitter hardware, potentially making this burden prohibitive (e.g. for PAPR > 30 dB). Secondly, the high-PAPR structure of this waveform makes it easily detectable in the time domain by a large variety of techniques [8], even at very low signal-to-noise ratios (SNRs), seemingly making it unsuitable for covert communications. Thirdly, it may appear that sharing the wideband channel by multiple users would require explicit allocation of the transmit and/or pulse arrival times for each sub-channel, which would be impractical in most cases.

Favorably, the temporal and amplitude structure of a wideband pulse train is modifiable by linear filtering, and such filtering can convert a high-PAPR train into a low-PAPR signal, and *vice versa*. Therefore, as detailed in this paper, such PAPR-modifying filtering enables us to use pulsed waveforms for low-SNR covert communications. Fig. 1 provides a simplified illustration of such an approach. The designed digital pulse sequence is a “pulse train”  $\hat{x}[k]$  with only some of the samples having non-zero values:

$$\hat{x}[k] = \sum_j \llbracket k = k_j \rrbracket A_j, \quad (1)$$

where  $k_j$  is the sample index of the  $j$ -th pulse,  $A_j$  is its amplitude, and the double square brackets denote the *Iverson bracket* [10]

$$\llbracket P \rrbracket = \begin{cases} 1 & \text{if } P \text{ is true} \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$



## II. SCRAMBLING AND PAPR CONTROL UTILITY OF PILEUP EFFECT

A pulse train  $x(t)$  can be viewed as a sum of pulses with the same shape (impulse response)  $w(t)$ , same or different amplitudes  $A_j$ , and distinct arrival times  $t_j$ :  $x(t) = \sum_j A_j w(t - t_j)$ . When the width of the pulses in a train becomes greater than the distance between them, the pulses begin to overlap and interfere with each other. This is illustrated in Fig. 2: For the same interarrival times, the pulses in the sequence consisting of the narrow pulses  $w(t)$  remain separate, while the wider (more “spread out”) pulses  $g(t)$  are “piling up on top of each other.” In this example,  $g(t)$  is obtained by filtering  $w(t)$  with an allpass filter (consisting of 6 cascaded biquad sections), and thus the PSDs of the pulse sequences are identical. However, the “pileup effect” causes the temporal and amplitude structures of these sequences to be substantially different. For a random pulse train, when the ratio of the bandwidth and the pulse arrival rate becomes significantly smaller than the TBP of a pulse, the pileup effect causes the resulting signal to become effectively Gaussian [14, e.g.], making it impossible to distinguish between the individual pulses.

Indeed, let  $\hat{x}(t)$  be an “ideal” pulse train:  $\hat{x}(t) = \sum_j A_j \delta(t - t_j)$ , where  $\delta(x)$  is the Dirac  $\delta$ -function [7]. The *moving average* of this ideal train in a boxcar window of width  $2T$  can be represented by the convolution integral

$$\bar{x}(t) = \int_{-\infty}^{\infty} ds \frac{\theta(t+T) - \theta(t-T)}{2T} \hat{x}(t-s), \quad (5)$$

where  $\theta(x)$  is the Heaviside unit step function [6]. At any given time  $t_i$ , the value of  $\bar{x}(t_i)$  is proportional to the sum of  $A_j$  for the pulses that occur within the interval  $[t_i - T, t_i + T]$ . Then, if the amplitudes  $A_j$  and/or the interarrival times  $t_{j+1} - t_j$  are independent and identically distributed (i.i.d.) random variables with finite mean and variance, it follows from the Central Limit Theorem [15, e.g.] that the distribution of  $\bar{x}(t_i)$  approaches Gaussian for a sufficiently large interval  $[-T, T]$ .

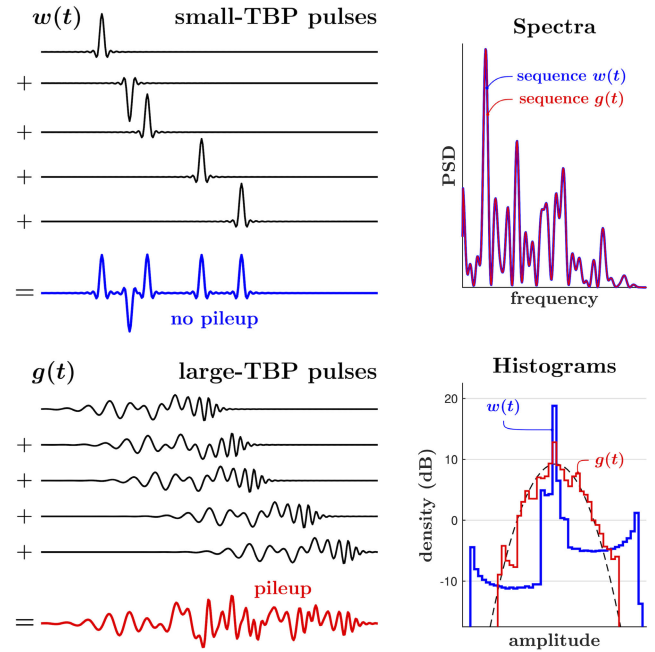
If we replace the boxcar weighting function in (5) with an arbitrary moving window  $w(t)$ , then (5) becomes a *weighted moving average*

$$x(t) = \int_{-\infty}^{\infty} ds w(t) \hat{x}(t-s) = (\hat{x} * w)(t) = \sum_j A_j w(t - t_j), \quad (6)$$

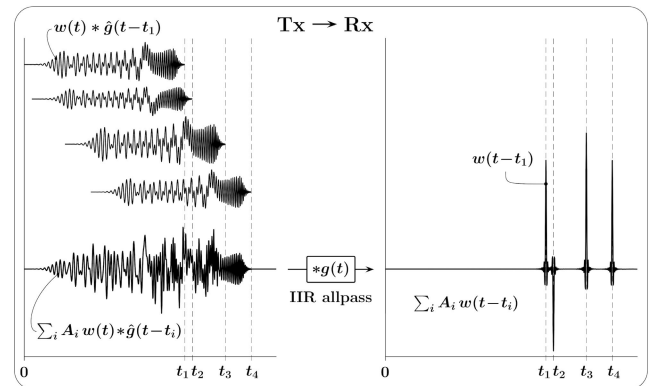
which is a “real” pulse train with the impulse response  $w(t)$ . If  $w(t)$  is normalized so that  $\int_{-\infty}^{\infty} ds w(s) = 1$ ,  $w(t)$  is an *averaging* (i.e. lowpass) filter. Then, if  $w(t)$  has both the bandwidth and the TBP similar to that of the boxcar pulse of width  $2T$ , the distribution of  $x(t_i)$  would be similar to that of  $\bar{x}(t_i)$  (e.g. Gaussian for a sufficiently large  $T$ ).

### A. PAPR CONTROL BY LARGE-TBP PULSE SHAPING

There are various ways to define the “time duration” and the “bandwidth” of a pulse [8, e.g.]. This can lead to a significant ambiguity in the definitions of the time-bandwidth products (TBPs), especially for waveforms with complicated temporal

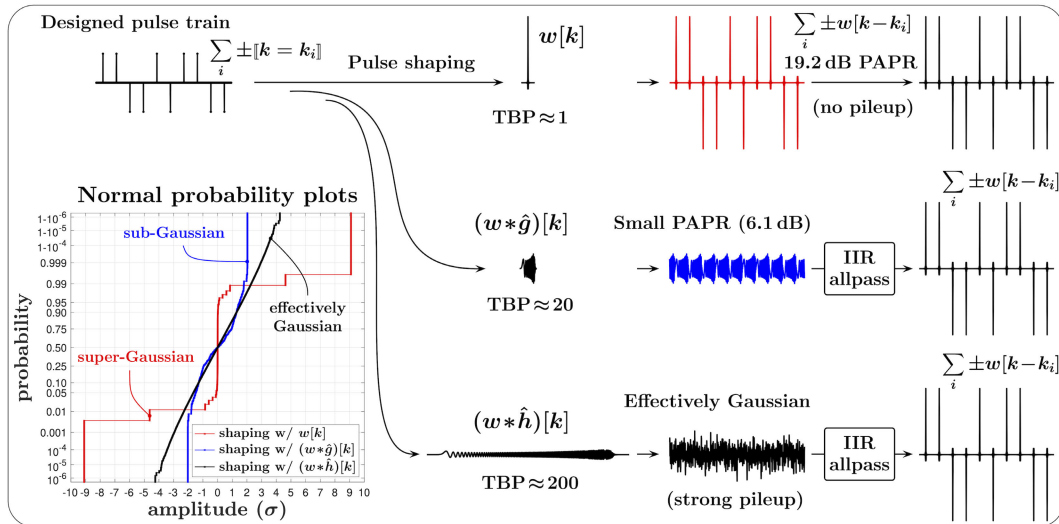


**FIGURE 2.** Illustration of pileup effect: When “width” of pulses becomes greater than distance between them, pulses begin to overlap and interfere with each other. For pulses with same spectral content, PSDs of pulse sequences are identical, yet their temporal and amplitude structures are substantially different.



**FIGURE 3.** Transmitter waveform is constructed as sum of scaled and time-shifted large-TBP pulses. In receiver, IIR allpass filter recovers small-TBP pulse train.

structures and/or frequency responses. For example, while compact support cannot be simultaneously achieved for the temporal and the spectral power densities of any pulse, the standard deviations,  $\sigma_t$  and  $\sigma_f$ , of these power densities can be used as measures of their width [16], [17]. Then, e.g., the TBP of a pulse can be defined as  $TBP = 4\pi\sigma_t\sigma_f \geq 1$ , with the equality (the smallest  $TBP = 1$ ) achieved for a Gaussian pulse. However, in the context of a PAPR control function of the pileup effect, our main concern is that the change in the TBP occurs only due to the change in the temporal structure of a filter, without the respective change in its spectral content. In this case, a change in the PAPR of a pulse is indicative of the change in its “sharpness” (or “resolution”) in the time domain, and the reciprocal of the PAPR can serve as a measure of the time duration of the pulse.



**FIGURE 4.** Using large-TBP filtering and pileup effect for PAPR control and obfuscation of temporal and amplitude structure of pulsed waveforms. In transmitter, pulse shaping with large-TBP filter reduces crest factor of pulse train, making it appear sub-Gaussian or effectively Gaussian. In receiver, signal’s distinct temporal and amplitude structure is restored.

For a single pulse  $w(t)$ , its PAPR can be expressed as

$$PAPR_w = \frac{\max(w^2(t))}{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} dt w^2(t)}, \quad (7)$$

where the interval  $[T_1, T_2]$  includes the effective time support of  $w(t)$ . Then for filters with the same spectral content but different impulse responses  $w(t)$  and  $g(t)$ , the ratio of their TBPs can be expressed as the reciprocal of the ratio of their PAPRs,

$$\frac{TBP_g}{TBP_w} = \frac{\max(w^2(t))}{\max(g^2(t))} = \frac{PAPR_w}{PAPR_g}, \quad (8)$$

where the PAPRs are calculated over a sufficiently long time interval  $[T_1, T_2]$  that includes the effective time support of both filters. Note that from (8) it follows that, among all possible pulses with the same spectral content, the one with the smallest TBP will contain a dominant large-magnitude peak. Hence any reasonable definition of a finite TBP for a particular filter with a given frequency response allows us to obtain comparable numerical values for the TBPs of all other filters with the same frequency response, regardless of their temporal structures.

There are multiple ways to construct pulses with identical spectral compositions (and thus identical autocorrelation functions) yet significantly different TBPs. For example, the autocorrelation function of an impulse response of any allpass filter is the Dirac  $\delta$ -function. Therefore, given a “seed” small-TBP pulse with finite (FIR) or infinite (IIR) impulse response  $w(t)$ , a large-TBP pulse with the same spectral content can be “grown” from  $w(t)$  by applying a sequence of IIR allpass filters that leave the PSD of the seed pulse unmodified [18, e.g.]. Then an FIR filter for pulse shaping in the transmitter can be obtained by (i) “spreading”  $w(t)$  with

an IIR allpass filter, (ii) truncating the pulse when it sufficiently decays to zero, and (iii) time-inverting the resulting waveform. Applying the same sequence of IIR allpass filters in the receiver to this waveform will produce the matched filter  $w(-t)$  to the original seed pulse.

In the example of Fig. 3, the transmitter waveform is a “piled-up” sum of thus constructed large-TBP pulses (obtained by “spreading”  $w(t)$  with the allpass filter  $\hat{g}(t)$ ), scaled and time-shifted. In the receiver, the IIR allpass filter  $g(t)$  (here consisting of 21 cascaded biquad allpass sections) recovers the underlying high-PAPR pulse train. Fig. 4 further illustrates how the pileup effect can be used to obscure (e.g. to mimic as Gaussian or sub-Gaussian) a large-PAPR (super-Gaussian) transmitted signal, while fully recovering its distinct temporal and amplitude structure in the receiver.

In Fig. 4, the first pulse shaping filter has a “narrow,” small-TBP impulse response  $w[k]$ . The impulse response  $(w * \hat{g})[k]$  of the second filter has about 20 times larger TBP, and it approximately “fills” one interpulse interval. The impulse response  $(w * \hat{h})[k]$  of the third pulse shaping filter spans several interpulse intervals. After shaping with  $w[k]$ , the pulse train remains highly impulsive, or super-Gaussian, and is suitable neither for modulation nor covert communications. After shaping with the second filter, the PAPR becomes about 20 times smaller, and the signal is sub-Gaussian. However, the individual pulse shapes are still mostly apparent in the time domain, and the signal is less “covert” than we might like. After the third filter, the waveform is effectively Gaussian, and it is impossible to distinguish between the individual pulses. While the PAPR of this aggregate waveform is larger than that of the sub-Gaussian signal (as can be seen from the normal probability plots shown in the lower left corner), it is not a serious practical limitation when using it for modulating a carrier. The relatively rare outliers in the modulating signal

can simply be “clipped,” without having much effect on the restored pulse train.

The seed  $w(t)$  used in Figs. 3 and 4 is an FIR root-raised-cosine (RRC) filter, and thus  $(w*w)(t)$  is a raised-cosine (RC) pulse [19, e.g.]. While TBPs of RC pulses are generally larger than those of a Gaussian or a Bessel pulse, compact frequency support of RC filters is appealing for communication applications. The TBPs of these filters remain relatively small for large roll-off factors (e.g.,  $TBP \lesssim 1.5$  for  $1/3 \lesssim \beta \leq 1$ ), and in the subsequent simulations and numerical examples we use FIR RC pulses with the roll-off factor  $\beta = 1/2$  ( $TBP \approx 1.27$ ).

### III. PULSED WAVEFORMS FOR LOW-SNR AND COVERT COMMUNICATIONS

Let us consider a pulse train consisting of pulses with a given TBP and bandwidth  $\Delta B$ , and with the average pulse arrival rate  $\mathcal{R}$ . When  $\mathcal{R}$  is sufficiently low, e.g.  $\mathcal{R} \ll \Delta B/TBP$ , pileup is negligible. As follows from the discussion in Section II, the PAPR of such a pulse train will be inversely proportional to the pulse rate,  $PAPR \propto \mathcal{R}^{-1}$ . Then, for a pulse train with a given bandwidth and average power, by reducing the pulse arrival rate the pulses can be made arbitrarily large and easily detectable even at low SNR. On the other hand, for a small PAPR and/or obfuscation of the temporal and amplitude structure of a pulsed waveform, pileup needs to be sufficiently strong. For that, for a given rate  $\mathcal{R}$  and bandwidth  $\Delta B$ , the TBP of the pulses needs to be sufficiently large, e.g.  $TBP \gtrsim \Delta B/\mathcal{R}$ .

For the designed pulse train  $\hat{x}(t) = \sum_j A_j \delta(t - t_j)$ , the waveforms  $x_w(t) = \sum_j A_j w(t - t_j)$  and  $x_g(t) = \sum_j A_j g(t - t_j)$  can be obtained by filtering  $\hat{x}(t)$  with filters having different impulse responses  $w(t)$  and  $g(t)$ , but the same frequency response and bandwidth  $\Delta B$ . Then  $x_w(t)$  would be a high-PAPR waveform suitable for messaging at a given  $SNR < 0$  dB, while  $x_g(t)$  would be a low-PAPR or effectively Gaussian “covert” waveform, when

$$TBP_w \frac{\mathcal{R}}{\Delta B} \ll SNR < 1 \lesssim TBP_g \frac{\mathcal{R}}{\Delta B}. \quad (9)$$

Thus using a link with a given bandwidth  $\Delta B$  for covert low-SNR communications would be mainly limited by the practically obtainable value of  $TBP_g$ . For example, for  $TBP_w$  of order unity and  $SNR = -20$  dB, the ratio  $\mathcal{R}/\Delta B$  would need to be smaller than approximately  $-30$  dB and, consequently,  $TBP_g \gtrsim 30$  dB.

For effective use of large-TBP pulse shaping for conversion of a high-PAPR pulse train with a distinct, super-Gaussian temporal and amplitude structure into an effectively Gaussian signal, the pulse train needs to be randomized. This can be accomplished by randomizing the amplitude of the pulses in the train, their arrival times, or both. The ways in which the pulse train is randomized affect the ways in which the information can be encoded and retrieved. For example, if the timing structure of the pulse train is known, synchronous pulse detection can be used. Otherwise, one may need to

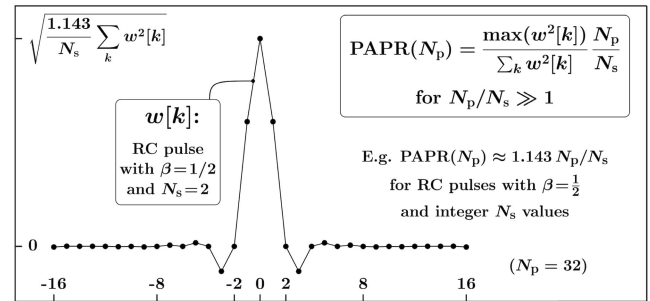


FIGURE 5. PAPR for train of equal-magnitude RC pulses separated by  $N_p \gg N_s$  is equal to PAPR of single pulse calculated on interval  $[-N_p/2, N_p/2]$ .

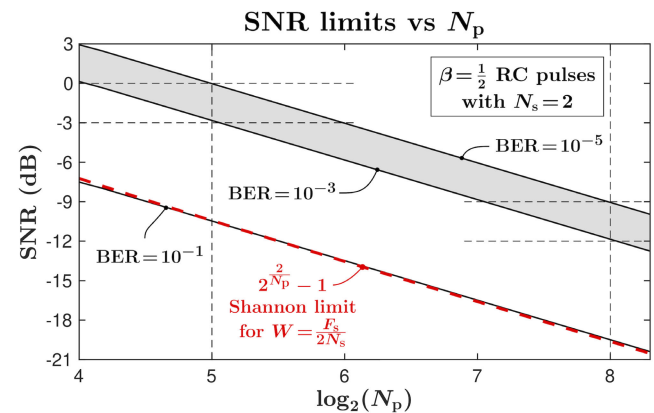


FIGURE 6. AWGN SNR limits for different BER as functions of samples between pulses for raised-cosine pulses with  $\beta = 1/2$  and  $N_s = 2$ .

employ asynchronous pulse detection (e.g. the pulse counting discussed in Section V).

#### A. SYNCHRONOUS PULSE DETECTION

Let us consider a pulse train consisting of pulses with the bandwidth  $\Delta B$  and a small TBP, so that a single large-magnitude peak in a pulse dominates, and assume that the arrival rate  $\mathcal{R}$  of the pulses is sufficiently small so that pileup is negligible (e.g.  $\mathcal{R} \ll \mathcal{R}_0 = \frac{1}{2} \Delta B/TBP$ ). When the arrival time of a pulse with the peak magnitude  $|A|$  is known, the probability of correctly detecting the polarity of this pulse in the presence of additive white Gaussian noise (AWGN) with zero mean and  $\sigma_n^2$  variance can be expressed, using the complementary error function, as  $\frac{1}{2} \text{erfc} \left( \frac{-|A|}{\sigma_n \sqrt{2}} \right)$ . Then the pulses with the magnitude  $|A| > \sigma_n \sqrt{2} \text{erfc}^{-1}(2\varepsilon)$  will have a pulse identification error rate smaller than  $\varepsilon$ . For example,  $\varepsilon \lesssim 1.3 \times 10^{-3}$  for  $|A| \gtrsim 3\sigma_n$ , and  $\varepsilon \lesssim 3.2 \times 10^{-5}$  for  $|A| \gtrsim 4\sigma_n$ .

The pulse rate in a digitally sampled train with regular (periodic) arrival times is  $\mathcal{R} = F_s/N_p$ , where  $F_s$  is the sampling frequency and  $N_p$  is the number of samples between two adjacent pulses in the train. For  $\mathcal{R}$  that is sufficiently smaller than  $\mathcal{R}_0$ , the PAPR of a train of equal-magnitude pulses with regular arrival times is an increasing function of the number of samples between two adjacent pulses  $N_p$ , and is proportional to  $N_p$ :

$$PAPR = PAPR(N_p) \propto N_p \quad \text{for large } N_p. \quad (10)$$

For example, for raised-cosine (RC) pulses  $\mathcal{R}_0 \approx (4T_s)^{-1}$ , where  $T_s$  is the symbol-period, and a “large  $N_p$ ” would mean  $N_p \gg T_s F_s = N_s$ , where  $N_s$  is the number of samples per symbol-period. As illustrated in Fig. 5,  $\text{PAPR}(N_p) \approx 1.143 N_p/N_s$  for  $N_p/N_s \gg 1$  for RC pulses with roll-off factor  $\beta = 1/2$  and integer values of  $N_s$ .

From the lower limit on the magnitude of a pulse for a given uncoded bit error rate (BER),

$$|A| = \sigma_n \sqrt{\text{SNR} \times \text{PAPR}} > \sigma_n \sqrt{2} \text{erfc}^{-1}(2 \times \text{BER}), \quad (11)$$

we can obtain the lower limit on the SNR for a given pulse rate:

$$\text{SNR}(N_p; \text{BER}) > \frac{2 [\text{erfc}^{-1}(2 \times \text{BER})]^2}{\text{PAPR}(N_p)} \propto N_p^{-1}, \quad (12)$$

or

$$\text{SNR}(N_p; \text{BER}) \gtrsim 1.75 \left[ \text{erfc}^{-1}(2 \times \text{BER}) \right]^2 \frac{N_s}{N_p} \quad (13)$$

for  $N_s/N_p \ll 1$  and RC pulses with  $\beta = 1/2$ . For example,  $\text{SNR}(N_p; 10^{-3}) \gtrsim 9.6/\text{PAPR}(N_p) \approx 8.4 N_s/N_p$ , and  $\text{SNR}(N_p; 10^{-5}) \gtrsim 18.2/\text{PAPR}(N_p) \approx 15.9 N_s/N_p$ .

Fig. 6 illustrates the SNR limits for different BER as functions of samples between pulses for RC pulses with  $\beta = 1/2$  and  $N_s = 2$ . For example, for the pulses separated by 128 symbol-periods,  $\text{BER} \lesssim 10^{-3}$  is achieved for  $\text{SNR} \gtrsim -12$  dB. For comparison, the AWGN Shannon capacity limit [20] for the bandwidth  $W = F_s/(2N_s)$ , which is the nominal bandwidth of the respective RRC filter, is also shown.

### B. ASYNCHRONOUS DETECTION (PULSE COUNTING)

When the time intervals between the pulses in a train do not compose a periodic structure, one must use asynchronous detection (pulse counting). In pulse counting, a pulse is detected when it crosses a certain non-zero threshold. A *false positive* (fp) detection occurs when such crossing is entirely due to noise, and a *false negative* (fn) detection happens when a pulse affected by the noise fails to cross the threshold. For a positive threshold  $\alpha^+ > 0$ , the false negative rate will be smaller than some tolerance rate  $\varepsilon_{fn}$  if the amplitude of a pulse is  $A > \alpha^+ + \sigma_n \sqrt{2} \text{erfc}^{-1}(2\varepsilon_{fn})$ .

As shown in [21], [22], for a filtered noise with zero mean and  $\sigma_n^2$  variance, its rate of up-crossing the threshold  $\alpha^+ > 0$  can be expressed as  $\mathcal{R}_{\max} \exp\left(-\frac{1}{2}(\alpha^+/\sigma_n)^2\right)$ , where the *saturation rate*  $\mathcal{R}_{\max}$  is determined entirely by the filter’s frequency response. Then, for the average pulse arrival rate  $\mathcal{R}$ , the threshold value needs to be  $\alpha^+ > \sigma_n \left[-2 \ln(\varepsilon_{fp} \mathcal{R}/\mathcal{R}_{\max})\right]^{\frac{1}{2}}$  in order to keep the false positive rate below  $\varepsilon_{fp}$ . For example, for  $\mathcal{R}/\mathcal{R}_{\max} = 1/10$ , the threshold value is  $\alpha^+ \gtrsim 4.3\sigma_n$  for  $\varepsilon_{fp} = 10^{-3}$ , and  $\alpha^+ \gtrsim 4.8\sigma_n$  for  $\varepsilon_{fp} = 10^{-4}$ . Note that, as shown in [21], for an ideal “brick wall” lowpass filter with the bandwidth  $\Delta B$  the saturation rate  $\mathcal{R}_{\max} = \Delta B/\sqrt{3}$ . Hence, for example, for a root-raised-cosine or a raised-cosine filter  $\mathcal{R}_{\max} \approx (2T_s\sqrt{3})^{-1}$ , where  $T_s$  is the reciprocal of the symbol-rate parameter of the filter.

For a pulse rate  $\mathcal{R}$  that is sufficiently smaller than  $\mathcal{R}_0 = \frac{1}{2} \Delta B/\text{TBP}$ , the PAPR of a train of equal-magnitude pulses is inversely proportional to  $\mathcal{R}$ . Then, for a given signal-to-noise ratio (SNR) of a pulse train affected by additive Gaussian noise, and for a given false negative rate constraint  $\varepsilon_{fn}$ , the pulse rate needs to be sufficiently small to ensure the pulse detection with the error rate below  $\varepsilon_{fn}$ .

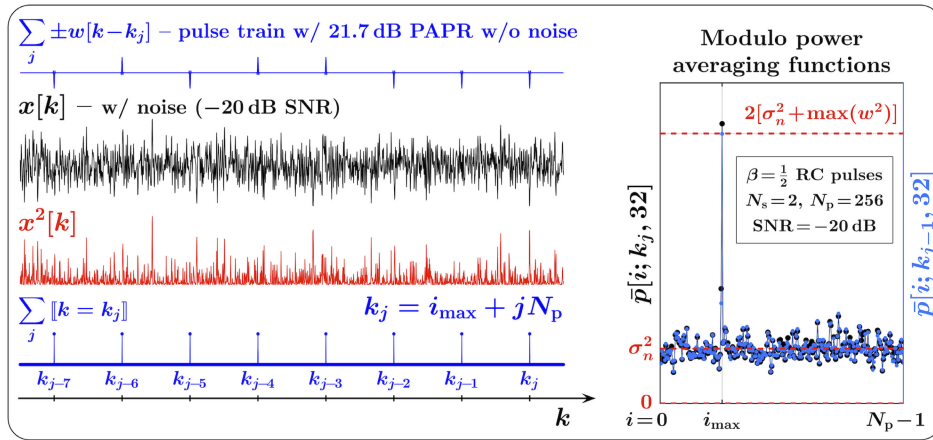
The asynchronous pulse detection (pulse counting) is discussed in more detail in Section V. While the rate limit for pulse counting is approximately an order of magnitude lower than for synchronous pulse detection with a similar BER, pulse counting does not rely on any *a priori* knowledge of pulse arrival times, and can be used as a backbone method for pulse detection. In addition, randomizing the pulse arrival times allows us to more effectively hide the temporal structure of the pulse train, prioritizing security over the data rates. Further, intermittently nonlinear filtering (INF) used in combination with synchronous and/or asynchronous pulse detection enables “layering” of pulse trains with significantly different powers, physical-layer steganography, and “friendly jamming” applications. However, since synchronous detection enables much higher data rates for the same SNR, the focus of the next section is on the technique that can be used for synchronous detection of pulses in a train with a periodic structure of interarrival times. In practice, both pulse counting and synchronous pulse detection can be used in combination. For example, given a constraint on the total power of the pulse train, counting of relatively rare, higher-magnitude pulses can be used to establish the timing patterns for synchronization, and synchronous detection of smaller, more frequent pulses can be used for a higher data rate.

### IV. SYNCHRONIZATION FOR REGULAR PULSE TRAINS

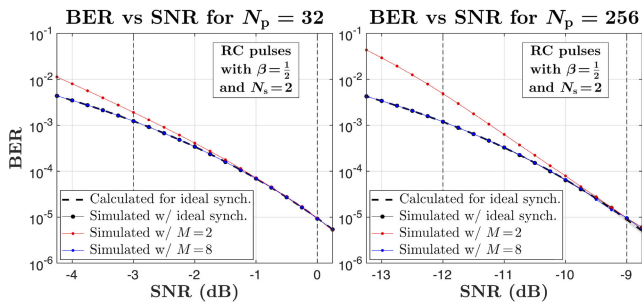
Let us consider the basic link shown in Fig. 1, where in the transmitter a binary sequence is encoded in a periodic pulse train according to (3), with  $N = N_p$ . To recover this sequence with minimal raw BER, we would need to sample the received pulse train  $x[k]$  at the instances where the pulses of the “ideal” (without noise) pulse train have maximum magnitudes. To determine the respective sampling indexes, the following *modulo power averaging* (MPA) function can be constructed as an exponentially decaying average of the instantaneous signal power  $x^2[k]$  in a window of size  $N_p + 1$ :

$$\begin{aligned} \bar{p}[i; k_j, M] &= \frac{M-1}{M} \bar{p}[i; k_{j-1}, M] \\ &+ \frac{1}{M} \sum_k \llbracket k \geq k_j - N_p \rrbracket \llbracket k \leq k_j \rrbracket \\ &\times \llbracket i = \text{mod}(k, N_p) \rrbracket x^2[k], \end{aligned} \quad (14)$$

where  $k_j$  is the sample index of the  $j$ -th pulse, and  $M > 1$ . Thus the window  $k_j - N_p \leq k \leq k_j$  includes two transmitted pulses,  $k_{j-1}$  and  $k_j$ , and the index  $i$  in  $\bar{p}[i; k_j, M]$  takes the values  $i = 0, \dots, N_p - 1$ . Note that using exponentially decaying average in (14) would roughly correspond to averaging  $N \approx 2M - 1$  such windows. The exponentially decaying average,



**FIGURE 7.** Illustration of synchronization procedure described by (14)–(16). AWGN SNR =  $-20$  dB is chosen to be low, and  $M = 32$  respectively high, to emphasize robustness even when  $\text{BER} \approx 1/3$ .



**FIGURE 8.** Calculated and simulated BERs as functions of AWGN SNRs for  $N_p = 32$  and  $N_p = 256$ . For shown SNR ranges, MPA function with  $M = 8$  provides reliable synchronization. (Compare with SNR limits in Fig. 6).

however, has the advantage of lower computational and memory burden, especially for large  $M$ , and faster adaptability to dynamically changing conditions.

For a sufficiently large  $M$ , the peak in  $\bar{p}[i; k_j, M]$  corresponding to the pulses of the ideal pulse train will dominate. Therefore, the index  $k_{j+1}$  for sampling of the  $(j + 1)$ -th pulse can be obtained as

$$k_{j+1} = i_{\max} + (j + 1)N_p, \quad (15)$$

where  $i_{\max}$  is given by

$$\bar{p}[i_{\max}; k_j, M] = \max_i (\bar{p}[i; k_j, M]). \quad (16)$$

Note that reliance on the global maximum of  $\bar{p}[i; k_j, M]$  is also likely to provide resilience to multipath interference, as this maximum will be mainly determined by the dominating (e.g. a line-of-sight) signal component.

Fig. 7 illustrates this synchronization procedure. The MPA function shown on the right-hand side of the figure is computed according to (14). To emphasize the robustness of this synchronization technique even when the bit error rates are very high, the SNR is chosen to be respectively low (SNR =  $-20$  dB,  $\text{BER} \approx 1/3$  in this example).

For the link shown in Fig. 1, and for the RC pulses with  $\beta = 1/2$  and  $N_s = 2$  used in Section III, Fig. 8 compares the calculated (dashed lines) and the simulated (dots connected

by solid lines) BERs, for the “ideal” synchronization (black dots), and for the synchronization with the MPA function described above. The AWGN noise is added at the receiver input, and the SNR is calculated at the output of the matched filter in the receiver. One can see that for  $M = 2$  (red dots) the errors in synchronization are relatively high, which increases the overall BER, but the MPA function with  $M = 8$  (blue dots) provides reliable yet still fast synchronization. The BERs and the respective SNRs in Fig. 8 are presented for the pulse repetition rates indicated by the vertical dashed lines in Fig. 6.

### A. MODULO MAGNITUDE AVERAGING

When a pulse train is used for communications rather than, say, radar applications, reliable synchronization may only need to be achievable for relatively low BER, e.g.  $\text{BER} \lesssim 1/10$ . Then the following *modulo magnitude averaging* (MMA) function can replace the MPA function in the synchronization procedure, in order to reduce the computational burden by avoiding squaring operations:

$$\begin{aligned} \bar{a}[i; k_j, M] &= \frac{M-1}{M} \bar{a}[i; k_{j-1}, M] \\ &+ \frac{1}{M} \sum_k \llbracket k > k_j - N_p \rrbracket \llbracket k \leq k_j \rrbracket \\ &\times \llbracket i = \text{mod}(k, N_p) \rrbracket |x[k]|. \end{aligned} \quad (17)$$

Note that the window  $k_j - N_p < k \leq k_j$  in (17) includes only the  $j$ -th transmitted pulse, instead of two pulses used in (14). The reason behind this is illustrated in Fig. 9, which compares (for AWGN) the MPA function  $\bar{p}[i; k_j, M]$  with the respective squared MMA function  $\bar{a}^2[i; k_j, M]$  computed for the window  $k_j - N_p \leq k \leq k_j$  that includes the “extra point” (the  $(j - 1)$ -th pulse). The relatively long averaging ( $M = 64$ ) is used to reduce the variations in the function values due to noise, and to make the comparison with the levels indicated by the dashed lines more apparent.

When a correct synchronization has already been obtained, and the maxima are “locked” at the correct  $i_{\max}$  values (black



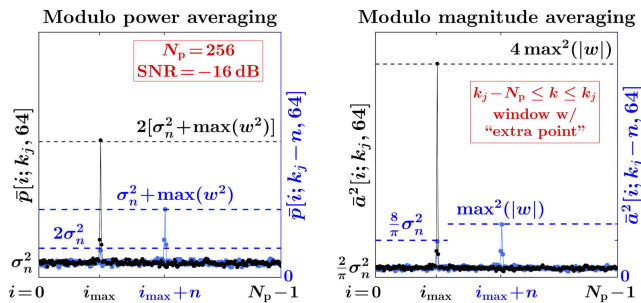


FIGURE 9. If used in modulo magnitude averaging, “extra point” significantly increases probability of synchronization failure.

dots connected by solid lines in Fig. 9), both the MPA and the MMA functions would adequately maintain the position of their maxima. However, an offset in the synchronization (e.g. by  $n$  points shown in the figure) significantly more unfavorably affects the margin between the extrema at  $i_{\max}$  and  $i_{\max} + n$  in the MMA function, compared with the MPA function (blue dots connected by solid lines). Thus the “extra point” may cause the “failure to synchronize” even at a relatively high SNR, and it should be removed from the calculation of the MMA function. Then, as illustrated in Fig. 10, for  $\text{BER} \lesssim 1/10$  synchronization with the MMA function  $\bar{a}[i; k_j, M]$  would be effectively equivalent to synchronization with the MPA function  $\bar{p}[i; k_j, M]$ . When reliable synchronization for larger BERs is desired (e.g. in timing and ranging applications), then the MPA given by (14) should be used.

V. INTERMITTENTLY NONLINEAR FILTERING (INF) FOR OUTLIER NOISE MITIGATION AND PULSE COUNTING

In addition to ever-present thermal noise (which can be appropriately modeled as AWGN), the received signal can contain significant amounts of non-Gaussian interference originating from a multitude of natural and technogenic (man-made) phenomena. Then the overall noise would be non-Gaussian and, depending on the noise coupling mechanisms and the system’s filtering properties and propagation conditions, it may contain distinct amplitude outliers when observed in the time domain. The presence of different types of such outlier noise is widely recognized in multiple applications under various general and application-specific names, most commonly as *impulsive*, *transient*, *burst*, or *crackling* noise. The outlier noise can be efficiently mitigated in real-time using intermittently nonlinear filters (INF) [12, e.g.], and, depending on the noise nature and composition, improvements in the quality of the signal of interest can vary from “no harm” to substantial.

Since in the power-limited regime the channel capacity is proportional to the SNR, even relatively small increase in the latter can be beneficial. For example, as can be seen in Fig. 6, for a given raw BER a 3 dB increase in the SNR enables doubling of the bit rate. Alternatively, for a given pulse rate, a 3 dB increase in the SNR can reduce the raw BER by several orders of magnitude. This is why, as illustrated in Fig. 1, it would be useful to deploy INF for mitigation of outlier noise in practical implementations of low-SNR links described in

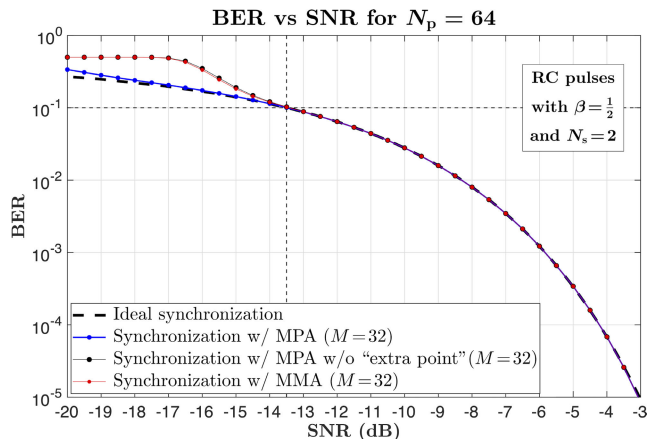


FIGURE 10. For BER smaller than about  $10^{-1}$ , less computationally expensive modulo magnitude averaging (e.g. given by (17)) can be used for synchronization. Modulo power averaging (with “extra point,” e.g. given by (14)) should be used when reliable synchronization for full BER range is desired.

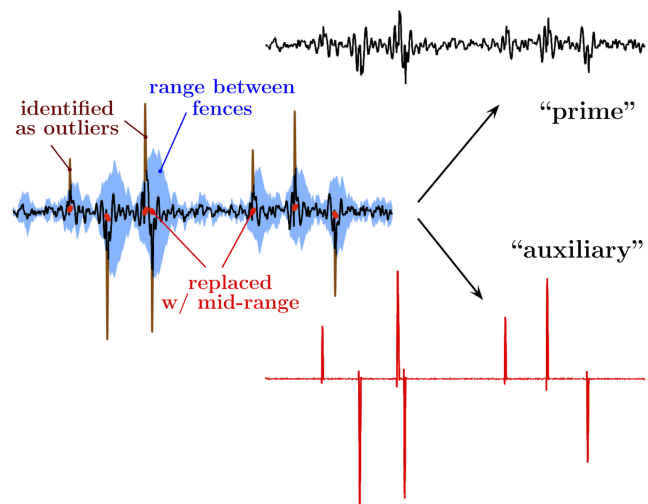


FIGURE 11. Intermittently Nonlinear Filtering (INF): Outliers are identified as protrusions outside of fenced range, and their values are replaced by those in mid-range. Otherwise, signal is not affected. “Auxiliary” output is difference between input and “prime” INF output.

this paper. Note that INF should be performed before the large-TBP filtering in the receiver, when the signal of interest is still Gaussian or sub-Gaussian, and the outlier structure of the noise is still apparent and not affected by the pileup.

In general, a nonlinear filter is capable of disproportionately affecting spectral densities of signals with distinct temporal and/or amplitude structures even when these signals have the same spectral content. In particular, the separation of a large-PAPR pulse train and a small-PAPR signal can be viewed as either (i) mitigation of impulsive noise affecting the small-PAPR signal, or (ii) extraction of impulsive signal from the small-PAPR background. In this paper, a specific type of Intermittently Nonlinear Filters (INF) is used to accomplish either or both tasks. While various INF configurations, their different uses, and the approaches to their analog and/or digital implementations are described elsewhere [12], [13], [23]–[25], Fig. 11 illustrates their basic concept. In an INF,

the upper and the lower fences establish a robust range that excludes high-amplitude pulses while effectively containing the small-PAPR component. The prime INF output simply contains the input signal in which the outliers (i.e. the pulses that protrude from the range) are replaced with mid-range values. This constitutes mitigation of impulsive noise affecting the small-PAPR signal. The auxiliary INF output is the difference between its input and the prime output. This is akin to extraction of impulsive signal from the small-PAPR background (or ‘‘pulse counting’’).

**A. ROBUST RANGE/FENCING IN INF**

For an INF to be effective in separation of small-PAPR and impulsive signals regardless of their relative powers, its range needs to be robust (insensitive) to the pulse train. Favorably, for a mixture of a small-PAPR signal with the bandwidth  $\Delta B$ , and a pulse train with the same bandwidth and the rate sufficiently below  $\mathcal{R}_0$ , when pileup effect is insignificant, the value of the interquartile range (IQR) of the mixture is insensitive to the power of the pulse train. This is illustrated in Fig. 12 for a pulse train affected by additive Gaussian noise. Thus robust upper ( $\alpha^+$ ) and lower ( $\alpha^-$ ) fences for INF can be constructed as linear combinations of the 1st ( $Q_{[1]}$ ) and the 3rd ( $Q_{[3]}$ ) quartiles of the signal (Tukey’s fences [26]) obtained in a moving time window:

$$[\alpha^-, \alpha^+] = [Q_{[1]} - \beta(Q_{[3]} - Q_{[1]}), Q_{[3]} + \beta(Q_{[3]} - Q_{[1]})], \tag{18}$$

where  $\alpha^+, \alpha^-, Q_{[1]}$ , and  $Q_{[3]}$  are time-varying quantities, and  $\beta$  is a scaling parameter of order unity. When an INF is used for pulse counting in the presence of AWGN, the particular value of  $\beta$  may be chosen based on the constraint on the relative rate  $\varepsilon_{fp}$  of false positive detections. Then, as follows from the discussion in Section III-B,

$$\beta \approx 1.05 \times \sqrt{\ln\left(\frac{\mathcal{R}_{max}}{\varepsilon_{fp}\mathcal{R}}\right)} - \frac{1}{2}. \tag{19}$$

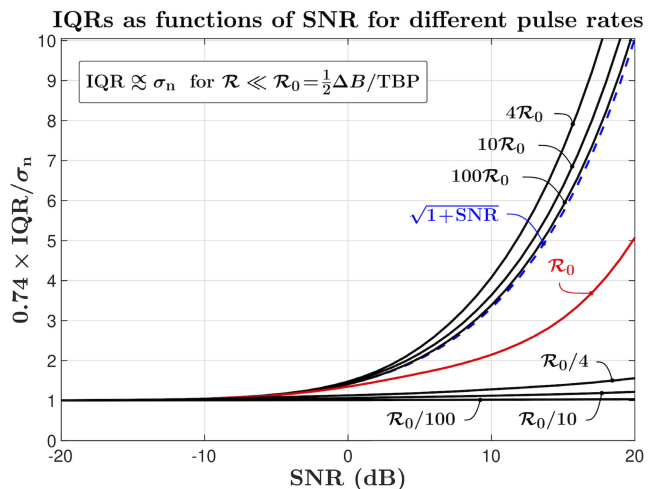
For example, for  $\mathcal{R}_{max}/\mathcal{R} = 10$ ,  $\beta \approx 2.7$  for  $\varepsilon_{fp} = 10^{-3}$ , and  $\beta \approx 3.1$  for  $\varepsilon_{fp} = 10^{-4}$ .

For Tukey’s fences with  $\beta$  given by (19), the probability of a false negative count of a pulse with the magnitude  $|A|$  can be expressed as

$$\varepsilon_{fn} \approx \frac{1}{2} \operatorname{erfc}\left(\frac{|A|}{\sigma_n \sqrt{2}} - \sqrt{\ln\left(\frac{\mathcal{R}_{max}}{\varepsilon_{fp}\mathcal{R}}\right)}\right). \tag{20}$$

Then, for example, for a train of equal-magnitude RC pulses with the roll-off factor 1/2, and with the average interarrival interval  $\bar{N}_p \gg N_s$  (with a possible constraint on the minimal interarrival distance to avoid pileup), the lower limit on the SNR for a given average pulse rate  $1/\bar{N}_p$  can be expressed as

$$\operatorname{SNR}(\bar{N}_p; \varepsilon_{fp}, \varepsilon_{fn}) \gtrsim 1.75 \left[ \operatorname{erfc}^{-1}(2\varepsilon_{fn}) + \sqrt{\ln\left(\frac{\bar{N}_p}{3.5 \varepsilon_{fp} N_s}\right)} \right]^2 \frac{N_s}{\bar{N}_p}. \tag{21}$$



**FIGURE 12.** For low pulse rates (e.g.  $\mathcal{R} \ll \frac{1}{2} \Delta B / \text{TBP}$ ), IQR provides reliable measure of additive Gaussian noise power,  $\sigma_n \propto \text{IQR}$ . Root-raised-cosine pulses (for which  $\mathcal{R}_0 \approx (47\varepsilon_s)^{-1}$ ) are used in this example. For completeness, IQRs for higher rates are also shown, but details of their change with SNR are not discussed.

Fig. 13 provides illustrative comparison of the AWGN SNR limits for synchronous and asynchronous detection. Note that while the rate limit for asynchronous detection is almost an order of magnitude lower than that for synchronous with a similar error rate, randomizing the pulse arrival times allows us to more effectively hide the temporal structure of the pulse train by a large-TBP pulse shaping, prioritizing security over the data rates.

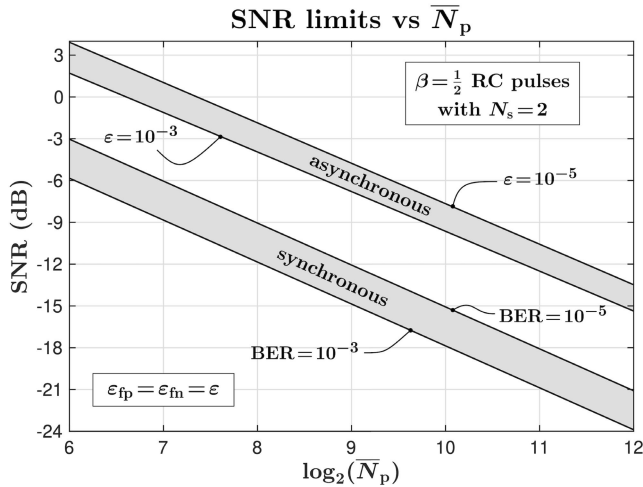
**B. QUANTILE TRACKING FILTERS FOR ROBUST FENCING**

As a practical matter, quantile tracking filters (QTFs) are an appealing choice for such robust fencing in INF, as QTFs are analog filters suitable for wideband real-time processing of continuous-time signals and are easily implemented in analog circuitry [12], [13], [23]–[25]. Further, their numerical computations are  $\mathcal{O}(1)$  per output value in both time and storage, which also enables their high-rate digital implementations in real time.

In brief, the signal  $Q_q(t)$  that is related to the given input  $x(t)$  by the equation

$$\frac{d}{dt} Q_q = \mu \left[ \lim_{\varepsilon \rightarrow 0} \mathcal{S}_\varepsilon(x - Q_q) + 2q - 1 \right], \tag{22}$$

where  $\mu$  is the rate parameter and  $0 < q < 1$  is the quantile parameter, can be used to approximate (‘‘track’’) the  $q$ -th quantile of  $x(t)$  for the purpose of establishing a robust range  $[\alpha^-, \alpha^+]$ . In (22), the comparator function  $\mathcal{S}_\varepsilon(x)$  can be any continuous function such that  $\mathcal{S}_\varepsilon(x) = \operatorname{sgn}(x)$  for  $|x| \gg \varepsilon$ , and  $\mathcal{S}_\varepsilon(x)$  changes monotonically from ‘‘−1’’ to ‘‘1’’ so that most of this change occurs over the range  $[-\varepsilon, \varepsilon]$ . As discussed in detail in [24], for a continuous stationary signal  $x(t)$  with a constant mean and a positive IQR, the outputs  $Q_{[1]}(t)$  and  $Q_{[3]}(t)$  of QTFs with a sufficiently small rate parameter  $\mu$  will approximate the 1st



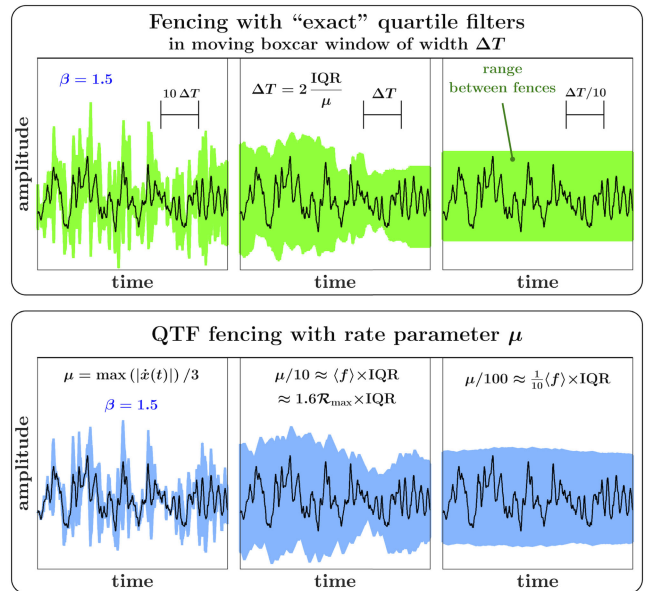
**FIGURE 13.** Illustrative comparison of SNR limits under AWGN for synchronous and asynchronous detection. Trains of equal-magnitude RC pulses with roll-off factor 1/2 and with average interarrival intervals  $\bar{N}_p \gg N_s$  are used.

and the 3rd quartiles, respectively, of the signal obtained in a moving boxcar time window with the width  $\Delta T$  of order  $2 \times \text{IQR}/\mu \gg \langle f \rangle^{-1}$ , where  $\langle f \rangle$  is the average crossing rate of  $x(t)$  with the 1st and the 3rd quartiles of  $x(t)$ . Consequently, as illustrated in Fig. 14, the overall behavior of the QTF fencing for a stationary constant-mean signal with a given IQR would be similar to the fencing with the “exact” quartile filters in a moving boxcar window  $[\theta(t) - \theta(t - \Delta T)]/\Delta T$ , where  $\Delta T = 2 \times \text{IQR}/\mu$  and  $\mu$  is the QTF rate parameter. However, for a sampling rate  $F_s$ , numerical computations of an “exact” quartile require  $\mathcal{O}(F_s \Delta T \log(F_s \Delta T))$  per output value in time, and  $\mathcal{O}(F_s \Delta T)$  in storage, becoming prohibitively expensive for high-rate real-time processing.

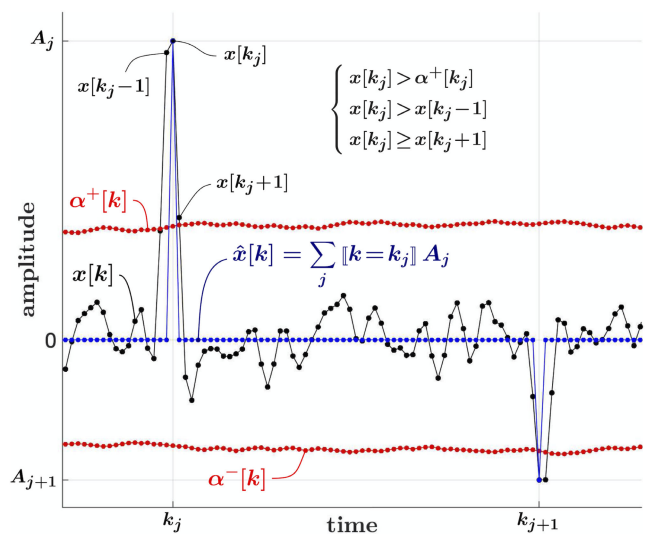
### C. ASYNCHRONOUS SAMPLING OF PULSE TRAINS

Let us assume that the designed pulse train  $\hat{x}[k]$  in the transmitter can be represented by (1), so that only  $k_j$ -th samples have non-zero values  $A_j$ , and that the information is encoded by the amplitudes  $A_j$  and/or the interarrival times  $k_j - k_{j-1}$ . Further, the impulse response  $w[k]$  of the small-TBP filter has a dominant peak, and the pulse arrival rate is small so that pileup is negligible. When a simple binary sequence is transmitted (say, “0” or “1” for negative or positive pulses), it can be recovered by measuring, for appropriately chosen thresholds  $\alpha^+ > 0$  and  $\alpha^- < 0$ , the upward and the downward crossings, respectively, of  $\alpha^+$  and  $\alpha^-$  by the pulse train  $x[k]$ .

The threshold crossings, however, do not characterize the amplitudes of the pulses. Also, the front edges of pulses with different magnitudes would have different slew rates, thus contributing to timing errors when the level crossings are used for timing. Instead, to extract the information about the amplitudes  $A_j$  in the train, and to reduce the error in obtaining the interarrival times  $k_j - k_{j-1}$ , the following pulse counting



**FIGURE 14.** Overall behavior of QTF fencing is similar to that with “exact” quartile filters in moving boxcar window of width  $\Delta T = 2 \times \text{IQR}/\mu$ .



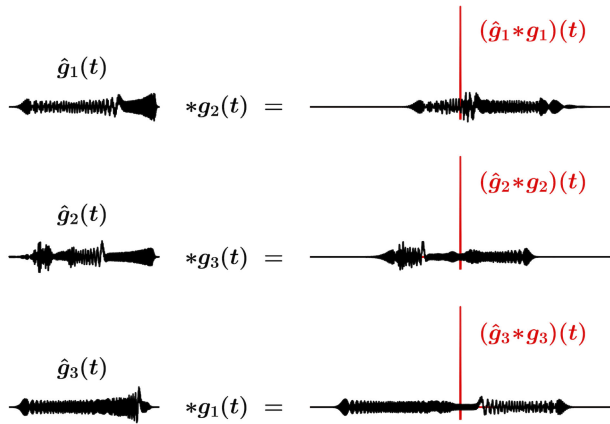
**FIGURE 15.** Sampling waveform  $x[k]$  at peaks of pulses protruding from range formed by Tukey’s fences  $[\alpha_k^-, \alpha_k^+]$ . Pulse train  $\hat{x}[k]$  is computed according to (24).

function  $C_{\alpha^\pm}^{\alpha^\pm}(x[k])$  can be used:

$$C_{\alpha^-}^{\alpha^-}(x_k) = \llbracket x_k > \alpha_k^+ \rrbracket \llbracket x_k > x_{k-1} \rrbracket \llbracket x_k \geq x_{k+1} \rrbracket + \llbracket x_k < \alpha_k^- \rrbracket \llbracket x_k < x_{k-1} \rrbracket \llbracket x_k \leq x_{k+1} \rrbracket. \quad (23)$$

This function takes unit values at the local extrema of  $x[k]$  that protrude from the range  $[\alpha_k^-, \alpha_k^+]$ , and is zero otherwise. (In (23), for better readability, we use  $x_k$  and  $\alpha_k^\pm$  in place of  $x[k]$  and  $\alpha^\pm[k]$ .) Then  $\hat{x}[k]$ , obtained as a product of the pulse train  $x[k]$  and the respective counting function,

$$\hat{x}[k] = C_{\alpha^\pm}^{\alpha^\pm}(x[k]) \cdot x[k] = \sum_j \llbracket k = k_j \rrbracket A_j, \quad (24)$$



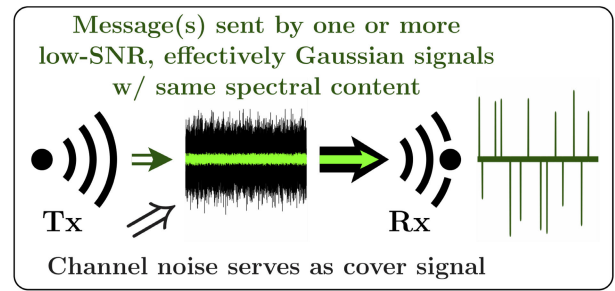
**FIGURE 16.** Example of autocorrelation (red) and cross-correlation responses for three filters  $\hat{g}_i(t)$  constructed by applying different allpass filters (consisting of 100 cascaded biquad sections) to RRC pulse.

will represent sampling of the train at the peaks of those pulses that protrude from the range formed by the fences  $\alpha^\pm[k]$ . This is illustrated in Fig. 15. When noise is negligible, and for appropriately chosen fences,  $\hat{x}[k]$  given by (24) will be equivalent to the designed pulse train in the transmitter.

### VI. PHYSICAL LAYER STEGANOGRAPHY

To meet the undetectability requirement, in a steganographic system the stego signals should be statistically indistinguishable from the cover signals [27]–[30]. For physical layer transmissions, this can perhaps be enhanced by requiring that the payload and the cover have the same bandwidth and spectral content, the same apparent temporal and amplitude structures, and that there are no explicit differences in the spectral and/or temporal allocations for the cover signals and the payload messages.

As discussed in Section II, by filtering a signal with a large-TBP filter one can “mimic” the signal as effectively Gaussian, without affecting its spectral composition. Given a small-TBP filter  $v(t)$  with a particular frequency response, one can construct a great variety of filters  $\hat{g}_i(t)$  and  $g_i(t) = \hat{g}_i(-t)$  with the same frequency response yet much larger TBPs (e.g., orders of magnitude larger). Thus the filters  $\hat{g}_i(t)$  can be used in the transmitter to transform both the cover and the payload signals into effectively Gaussian, making them statistically indistinguishable from each other, while having the same spectral content. These filters can be constructed in such a way that their combined matched (autocorrelation) responses are equal to each other,  $(\hat{g}_i * g_i)(t) = v(t) * v(-t) = w(t)$  for any  $i$ , and have a small TPB, yet the convolutions of any  $\hat{g}_i(t)$  with  $g_j(t)$  for  $i \neq j$  (cross-correlations) have large TBPs. Fig. 16 illustrates this for three filters constructed by applying different allpass filters (consisting of 100 cascaded biquad sections) to an RRC pulse. Then a selected  $i$ -th signal component can be “recovered” in the receiver by applying  $g_i(t)$ , while the rest of the signal would remain effectively Gaussian.



**FIGURE 17.** Using channel noise as cover signal.

For example, let us consider  $K \geq 2$  effectively Gaussian waveforms  $x_i = (\hat{x}_i * \hat{g}_i)(t) = \sum_m A_m \hat{g}_i(t - t_m)$ . Then any  $i$ -th component in the sum  $x = \sum_{i=1}^K x_i$  can be viewed as a “payload,” while the rest of the mixture can be considered a “cover.” To “extract” the payload, we apply  $g_i(t)$  to this mixture:

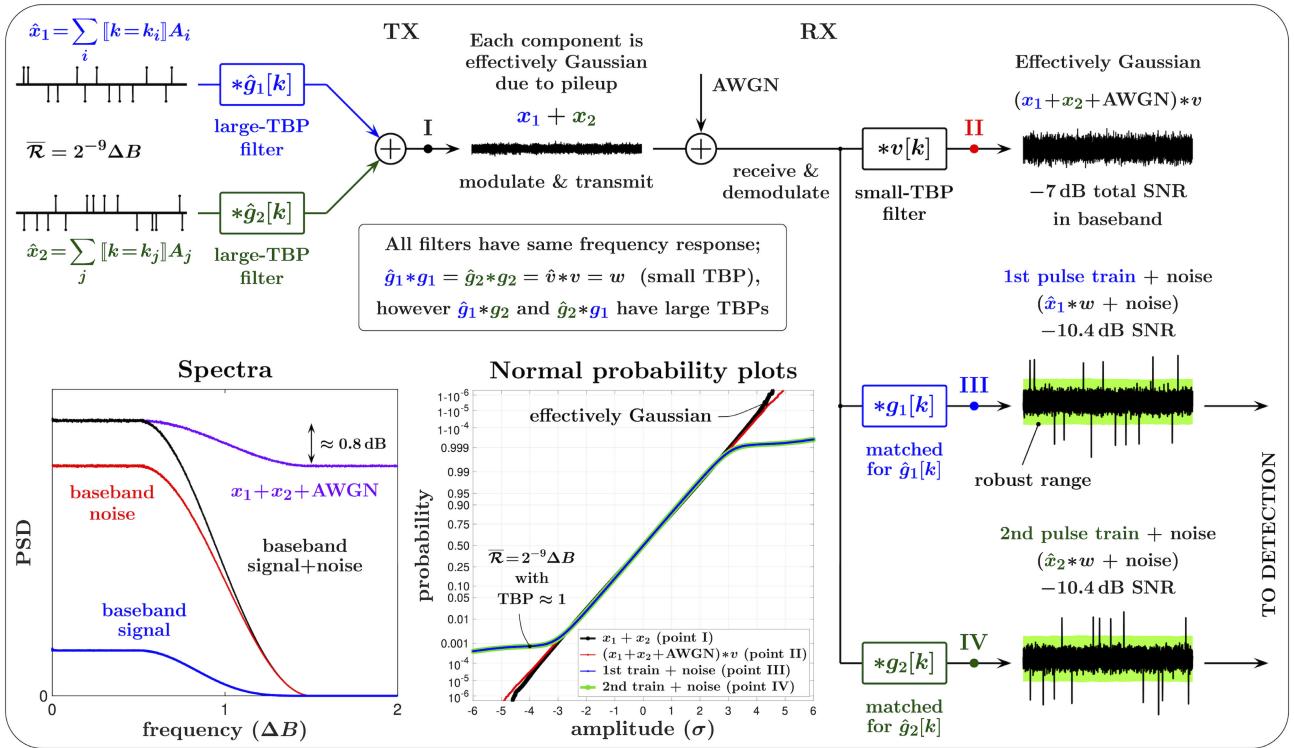
$$(x * g_i)(t) = \sum_m A_m w(t - t_m) + \sum_j \llbracket j \neq i \rrbracket (\hat{x}_j * g_{ji})(t), \quad (25)$$

where  $g_{ji}(t) = (\hat{g}_j * g_i)(t)$ . The right-hand side of (25) can be treated as a pulse train affected by Gaussian noise. When the external noise is negligible (i.e. when the total SNR is much larger than  $K$ ), for a mixture of equal-power waveforms the SNR for such a train will be  $\approx 1/(K - 1)$  (e.g.  $-7$  dB for a 6-component mixture).

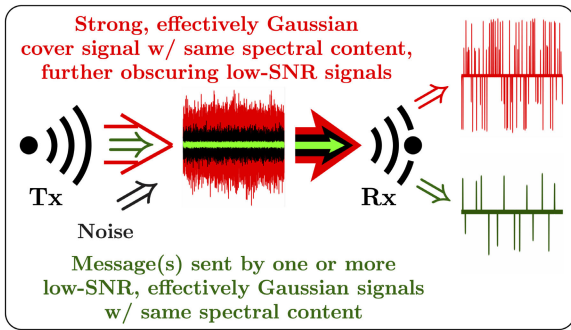
#### A. USING CHANNEL NOISE AS COVER SIGNAL

The very existence of a detectable carrier (cover signal) may be a dead giveaway for the stego payload. For example, the mere presence of a sheet of paper implies the possibility of a message written in invisible ink. Therefore, the best steganography should be “carrier-less,” when the payload is covertly embedded into something “ever-present.” In the physical layer, such “ideal” and unidentifiable cover signal is the channel noise. Such noise always includes the ever-present thermal noise as one of its components, and typically contains other (in general, non-Gaussian) natural and/or technogenic (man-made) components which make the noise non-white and characterized by time-variant parameters. Then, if the stego payload “pretends” to be Gaussian, and its PSD everywhere is small enough to be well within the “natural” variations in the PSD of the channel noise, any physically available band can be used to carry a virtually undetectable covert message. This approach is schematically illustrated in Fig. 17, where it is assumed that the messages are sent by pulse trains shaped with large-TBP filters.

In essence, all spread spectrum techniques for covert communications involve transmitting a signal that requires a limited bandwidth on a much wider bandwidth, thereby suppressing the PSD of the transmission below the noise floor [29], [30]. Even if the noise is stationary and accurate wideband measurements of its PSD  $\mathcal{N}$  are available, detection of the mere presence of a signal with the peak



**FIGURE 18.** Illustrative example of using channel noise as cover signal for two-component payload. Overall covert message is split into two individually covert components. In receiver, filter matched to selected pulse shaping filter in transmitter converts respective signal component into corresponding pulse train, while second component remains as effectively Gaussian contribution to noise.



**FIGURE 19.** Additional INF-removable cover for low-SNR payloads.

PSD  $\mathcal{S} \ll \mathcal{N}$  would require identifying an approximately  $4.3 \times \mathcal{S}/\mathcal{N} \lesssim 1$  dB small “bump” in the measured noise PSD during the transmission. (This task rapidly becomes more challenging for intermittent, short-duration transmissions, and/or for non-stationary noise.) Note, however, that even when the correct de-scrambling filter is applied to a pulse train shaped with a large-TBP filter, it does not change the average PSD of the transmitted signal, and so spectral measurements alone would still be insufficient for detection of the covert transmission. In contrast, for example, applying a known spreading sequence to a DSSS signal converts a wideband low-PSD signal into a narrowband high-PSD peak, and therefore enables its detection through spectral measurements.

As outlined in Fig. 17, the covert messages can be sent by several same-bandwidth pulse trains, each shaped with its own large-TBP filter, and Fig. 18 provides a detailed illustrative example for a two-component mixture. In Fig. 18, for each of the designed pulse trains  $\hat{x}_1[k]$  and  $\hat{x}_2[k]$  we use respective large-TBP pulse shaping filters,  $\hat{g}_1[k]$  and  $\hat{g}_2[k]$ , with the impulse responses extending over multiple interpulse intervals. By design,  $\hat{g}_1[k]$  and  $\hat{g}_2[k]$  have identical, small-TBP autocorrelation functions, but their cross-correlations have large TBPs. Then, after the demodulation in the receiver, a filter matched to the selected pulse shaping filter converts the respective signal component into a corresponding pulse train, while the second component remains as an effectively Gaussian contribution to the noise.

For  $K \geq 1$  equal-power components and dominant noise (e.g. when the noise power is much larger than the power of  $K - 1$  components), the SNR per component will be  $\approx \text{SNR}/K$ . Thus the available pulse rate for the total SNR can be equally shared among  $K$  low-SNR payloads, each requiring its own “key” to extract the information. This can be used, for example, for splitting the overall covert message into multiple, individually covert components.

**B. ADDING INF-REMOVABLE COVER FOR LOW-SNR PAYLOADS**

For a stego pulse train with a given rate, further increasing the power of the channel noise (say, by 10 dB) can make the pulse train undetectable. For example, when the pulse rate

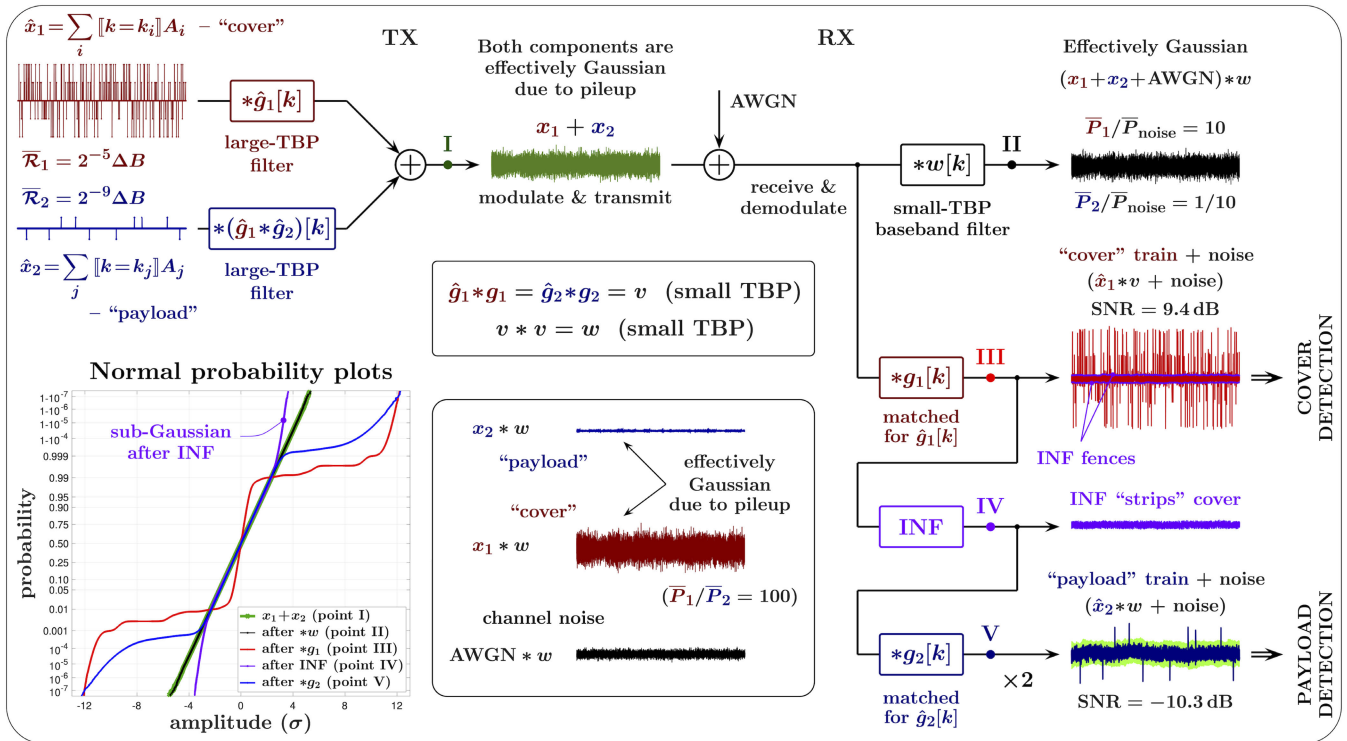


FIGURE 20. Illustrative example of adding INF-removable cover as outlined in Fig. 19. Power of cover exceeds power of payload by  $\times 100$  (20 dB), while both occupy effectively same spectral band. In receiver, removal of cover component by INF provides access to much weaker payload.

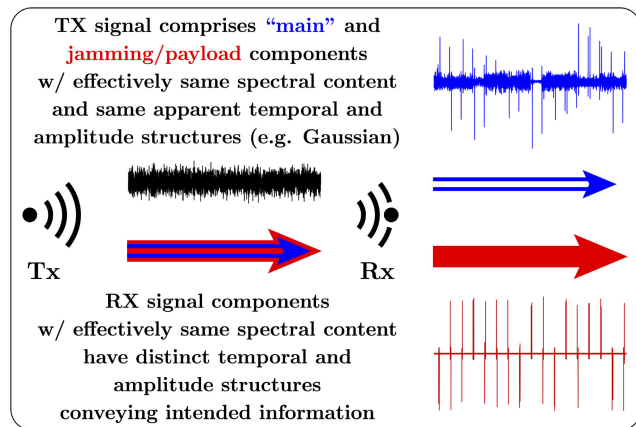


FIGURE 21. Basic concept of "friendly in-band jamming".

is higher than the Shannon limit for the given SNR, neither synchronous nor asynchronous detection would be possible. However, such increase in the channel noise power can be accomplished by an additional pulse train, simply disguised as Gaussian. Then an INF in the receiver, in combination with the respective "de-scrambling" filter, can effectively remove this additional noise, enabling the detection of the low-power payload. In addition, the higher-power pulse train can itself carry a lower-security (or decoy) message, and/or the timing information that enables synchronous pulse detection in the stego pulse train. Recovering this information from the "extra cover" signal would still require knowledge of the respective

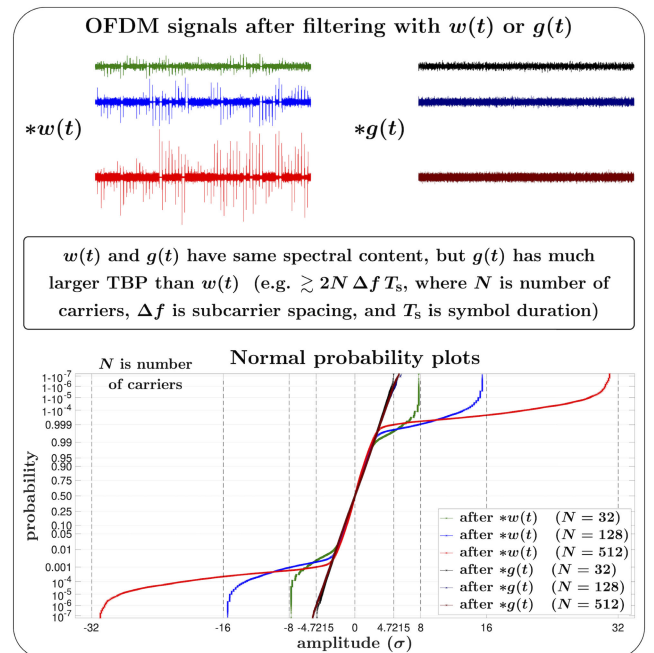
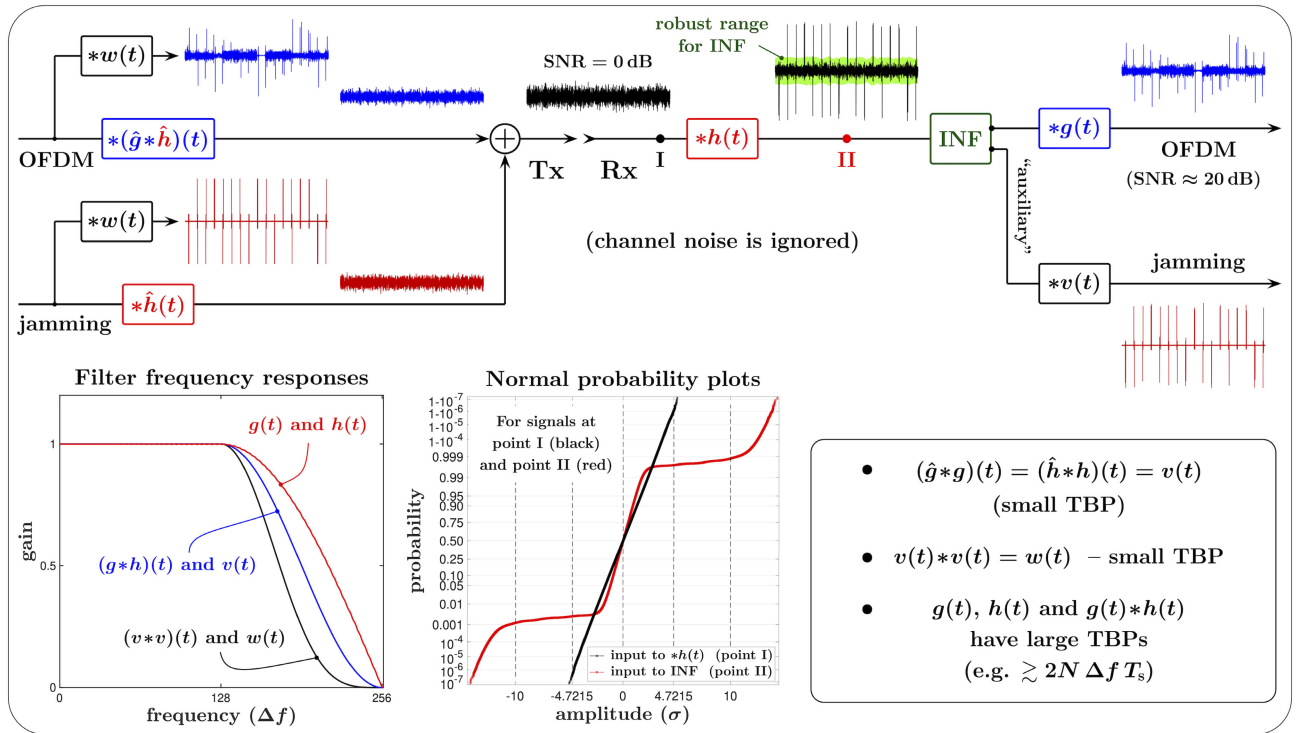


FIGURE 22. OFDM PAPR reduction by large-TBP filtering.

scrambling filter used by the transmitter. This concept is schematically illustrated in Fig. 19, and Fig. 20 provides its detailed illustrative example.

In Fig. 20, the large-TBP filters  $\hat{g}_1[k]$  and  $(\hat{g}_1 * \hat{g}_2)[k]$  are used for shaping the cover ( $\hat{x}_1[k]$ ) and the payload ( $\hat{x}_2[k]$ )



**FIGURE 23. Friendly in-band jamming of OFDM signal: Combination of linear and nonlinear filtering in receiver is used for effective separation of OFDM and “friendly jamming” signals, although both signals in received mixture have effectively same spectral characteristics and temporal and amplitude structures, and there are no explicit differences in their temporal allocations.**

pulse trains, and the power of the cover exceeds the power of the payload by a factor of 100 (by 20 dB). The TBP of  $\hat{g}_2[k]$  is also large, and the autocorrelation responses of  $\hat{g}_1[k]$  and  $\hat{g}_2[k]$  are identical and have a small TBP. In the receiver, the demodulated signal is first filtered with  $g_1[k]$ , which converts the cover component into an impulsive pulse train, while the rest of the signal (that contains the payload and the noise) remains a “diffused,” low-PAPR waveform. Thus INF can be used for effective separation of the cover pulse train from the rest of the signal. After removal of the cover signal by the INF, the filter  $g_2[k]$  is applied to the residual low-PAPR waveform. This converts the component corresponding to the weak payload pulse train  $\hat{x}_2[k]$  into a high-PAPR signal, enabling its reliable detection.

### C. FRIENDLY IN-BAND JAMMING

In our third example, the main message is transmitted using one of the existing communication protocols, but its temporal and amplitude structure is obscured by employing a large-TBP filter in the transmitter, e.g., made to be effectively Gaussian. This alone provides a certain level of security, since the intersymbol interference becomes excessively large and the signal cannot be recovered in the receiver without the knowledge of the scrambling filter. In addition, a jamming pulse train, disguised as Gaussian by another (and different) large-TBP filter, is added to the main signal. This jamming signal has effectively the same spectral content as the main

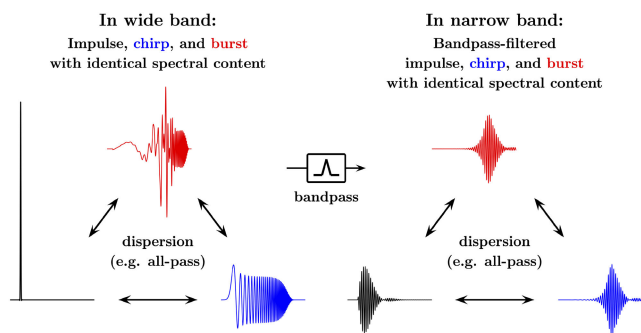
signal, and its power is sufficiently large so that the main signal is unrecoverable even if the first scrambling filter is known. In the receiver, the jamming pulse train is removed from the mixture (and recovered, if it itself contains information), enabling the subsequent recovery of the main message. This concept is schematically illustrated in Fig. 21.

#### 1) OFDM PAPR REDUCTION

In addition to improved security, applying a large-TBP filter to the main signal reduces PAPR of large-crest-factor signals such as those in orthogonal frequency-division multiplexing (OFDM), as illustrated in Fig. 22. Here, the simulated OFDM signals are generated without restrictions of the proportion of “ones” and “zeros” in a symbol, and thus they have the maximum achievable PAPRs (i.e.  $2N$ , where  $N$  is the number of carriers).

#### 2) ILLUSTRATIVE EXAMPLE

In Fig. 23, the main signal is a high-PAPR OFDM signal, and the jamming signal is a high-PAPR impulsive pulse train with the spectral content in effectively the same band (see the frequency responses of the filters in the lower left panel of Fig. 23). After the filtering with large-TBP filters  $\hat{g}*\hat{h}(t)$  and  $\hat{h}(t)$ , respectively, both the OFDM and the jamming signals become effectively Gaussian, and so does their mixture that is being transmitted and received (see the black line in the normal probability plots shown in the lower middle panel



**FIGURE 24.** Effect of filtering on temporal and amplitude structure of signal is more apparent at wider bandwidth.

of Fig. 23). (In this example, the channel noise is assumed to be relatively small and is not shown.) However, applying a filter  $h(t)$  matched for  $\hat{h}(t)$  in the receiver restores the high-PAPR structure of the jamming signal (see the red line in the normal probability plots), while the OFDM component remains Gaussian. Subsequently, the INF accomplishes both the mitigation of the jamming pulse train affecting the OFDM component and the extraction of the jamming signal. Applying the filter  $g(t)$  to the prime INF output effectively restores the original high-PAPR OFDM signal. If desired, the jamming pulse train is restored by applying the filter  $v(t)$  to the auxiliary INF output.

The main properties of the filters used in this example are listed in the lower right panel of Fig. 23, and their frequency responses are shown in the lower left panel of the figure.

### VII. CONCLUSION

For a finite pulse rate, the ideal pulse train  $\hat{x}(t) = \sum_j A_j \delta(t - t_j)$  has infinite bandwidth. Similarly, the designed digital pulse sequence  $\hat{x}[k] = \sum_j \llbracket k = k_j \rrbracket A_j$  occupies the full Nyquist range of frequencies. On the other hand, the information encoded in a low-rate pulse train can also be transmitted, with simple binary coding and modulation, utilizing much smaller bandwidth. This is why communication with pulse trains can be considered a spread-spectrum technique. In contrast to other such techniques, however, wideband pulse trains are constructed without actual physical “spreading” of a narrowband signal, and the “spreading factor” (i.e. the fraction of a given bandwidth that is utilized) can be managed by changes in the information rate and/or the bandwidth of the pulse shaping filter. The main advantage of this approach lies in retaining control over the temporal and amplitude structure of the pulsed waveform. Such control, especially when combined with nonlinear filtering techniques, enables the development of a large variety of low-SNR and covert communication configurations.

Somewhat simplistically, other spread-spectrum approaches can be viewed as modulating a wideband carrier by a narrowband signal in the transmitter, then recovering the narrowband signal in the receiver. For example, since multiplication is associative, the spreading sequence in DSSS

can be applied first to the carrier, forming a wideband carrier signal; then the wideband carrier is modulated by the narrowband signal. In FHSS and CSS, the frequency range of the carrier also spans a wide range. In the receiver, the respective demodulation (now combined with despreading) is used to produce the information-carrying narrowband signal. Because of the lack of “excess bandwidth” over that needed to carry the information (for the given modulation type), the statistical properties of such a narrowband signal can no longer be significantly changed by linear filtering that leaves the spectral composition of the signal intact. On the other hand, as we illustrate throughout the paper, the temporal and amplitude structures of wideband signals are easily managed, and distinct outliers in pulsed waveforms can be made appear, disappear, and reappear without affecting the signals’ spectral content.

In general, the effect of filtering on the temporal and/or the amplitude structure of a signal is more apparent at wider bandwidths, as a broader frequency range results in finer time resolution. This is illustrated in Fig. 24, where the impulse, the chirp, and the “burst” signals have the same spectral content, and only the phases in their Fourier representations are different. These three signals can be morphed into each other by all-pass filtering that leaves their power spectral densities unmodified. In a wide band, such filtering drastically changes the time-domain appearance of the signal and its amplitude density, while in a narrow band (after the bandpass filtering) these changes are much less apparent and all signals maintain similar temporal and amplitude structures. Thus wideband pulsed waveforms extend our options for encoding low-rate information. Since wider bandwidth offers increased time resolution, the message can be encoded not only in the amplitudes of the pulses (e.g. as a binary sequence given by (3)), but also in the pulse interarrival times, as represented by (1). This also adds to increased security, as various non-binary encoding protocols can be used for signals within the same physical band.

Further, control over the temporal and amplitude structures of wideband pulse trains carrying low-rate information provides for effective use of nonlinear filtering techniques. Such techniques can be employed for robust real-time asynchronous extraction of the information as well as for separation of signals with the same spectral content from each other. For example, the interquartile range can provide reliable measure for the average power of a Gaussian or sub-Gaussian signal, while being insensitive to changes in the power of such sparse signals as high-PAPR pulse trains. This enables the use of asynchronous detection (pulse counting) to extract both the amplitude and the timing information from randomized pulse trains. More generally, intermittently nonlinear filtering, when combined with waveform control by linear pulse shaping, allows effective separation of different signals with identical spectral profiles, facilitating development of versatile mixtures of same-band signals for covert and/or hard-to-intercept communications, and increasing robustness and quality of such communications in the presence of



non-Gaussian noise. Two particular illustrative examples of using such mixtures are detailed in the previous section.

## ACKNOWLEDGMENT

The authors would like to thank Kendall Castor-Perry (aka The Filter Wizard); James E. Gilley of BK Technologies; Jeff E. Smith of Sierra Nevada Corporation; Arlie Stonestreet II of Ultra Electronics ICE, and Kyle D. Tidball of Textron Aviation, for their valuable suggestions and critical comments. The authors would also like to extend their special gratitude to late Dr. Earl McCune (1956–2020), for encouraging them to address this topic.

## REFERENCES

- [1] S. Verdú, "Fifty years of Shannon theory," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2057–2078, Oct. 1998.
- [2] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*, Revised ed. New York, NY, USA: McGraw-Hill, 1994.
- [3] D. Torrieri, *Principles of Spread-Spectrum Communication Systems*, 4th ed. Cham, Switzerland: Springer, 2018.
- [4] S. D. Blunt and E. L. Mokole, "Overview of radar waveform diversity," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 31, no. 11, pp. 2–42, Nov. 2016.
- [5] N. Hosseini and D. W. Matolak, "Nonlinear quasi-synchronous multi user chirp spread spectrum signaling," 2019, *arXiv:1909.09887*. [Online]. Available: <http://arxiv.org/abs/1909.09887>
- [6] R. N. Bracewell, "Heaviside's unit step function,  $H(x)$ ," in *The Fourier Transform and Its Applications*, 3rd ed. New York, NY, USA: McGraw-Hill, 2000, pp. 61–65.
- [7] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. London, U.K.: Oxford Univ. Press, 1958.
- [8] T. W. Barrett, "History of ultra wideband communications and radar: Part I, UWB communications," *Microw. J.*, pp. 22–56, Jan. 2001.
- [9] *Characteristics Of Ultra-Wideband Technology*, Standard ITU-R SM.1755-0, 2006. [Online]. Available: <https://www.itu.int>
- [10] D. E. Knuth, "Two notes on notation," *Amer. Math. Monthly*, vol. 99, no. 5, pp. 403–422, May 1992.
- [11] B. Picinbono, "On instantaneous amplitude and phase of signals," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 552–560, Mar. 1997.
- [12] A. V. Nikitin and R. L. Davidchack, "Hidden outlier noise and its mitigation," *IEEE Access*, vol. 7, pp. 87873–87886, 2019.
- [13] A. V. Nikitin and R. L. Davidchack, "Complementary intermittently nonlinear filtering for mitigation of hidden outlier interference," in *Proc. MILCOM-IEEE Mil. Commun. Conf. (MILCOM)*, Norfolk, VA, USA, Nov. 2019, pp. 12–14.
- [14] A. V. Nikitin, "Pulse pileup effects in counting detectors," Ph.D. dissertation, Dept. Phys. Astron., Univ. Kansas, Lawrence, KS, USA, 1998.
- [15] A. D. Aleksandrov, A. N. Kolmogorov, and M. A. Lavrent'ev, Eds., *Mathematics: Its Content, Methods, and Meaning*. New York, NY, USA: Dover, 1999.
- [16] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, no. 3, pp. 429–457, 1946.
- [17] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Upper Saddle River, NJ, USA: Prentice-Hall, 1995.
- [18] P. A. Regalia, S. K. Mitra, and P. P. Vaidyanathan, "The digital all-pass filter: A versatile signal processing building block," *Proc. IEEE*, vol. 76, no. 1, pp. 19–37, Jan. 1988.
- [19] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 4th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2006.
- [20] C. E. Shannon, "Communication in the presence of noise," *Proc. Inst. Radio Eng.*, vol. 37, no. 1, pp. 10–21, Jan. 1949.
- [21] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst. Tech. J.*, vol. 23, no. 3, pp. 282–332, 1944.
- [22] A. V. Nikitin, R. L. Davidchack, and T. P. Armstrong, "The effect of pulse pile-up on threshold crossing rates in a system with a known impulse response," *Nucl. Instrum. Methods Phys. Res. A. Accel. Spectrom. Detect. Assoc. Equip.*, vol. 411, no. 1, pp. 159–171, Jul. 1998.
- [23] A. V. Nikitin, "Method and apparatus for nonlinear filtering and for mitigation of interference," U.S. Patent 10 637 490, Apr. 28, 2020.
- [24] A. V. Nikitin and R. L. Davidchack, "Quantile tracking filters for robust fencing in intermittently nonlinear filtering," 2019, *arXiv:1911.00736*. [Online]. Available: <http://arxiv.org/abs/1911.00736>
- [25] 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.
- [26] J. W. Tukey, *Exploratory Data Analysis*. Reading, MA, USA: Addison-Wesley, 1977.
- [27] J. Fridrich, M. Goljan, and D. Soukal, "Searching for the stego-key," *Proc. SPIE*, vol. 5306, pp. 70–82, Jun. 2004.
- [28] J. Fridrich, M. Goljan, P. Lisonek, and D. Soukal, "Writing on wet paper," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3923–3935, Oct. 2005.
- [29] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.
- [30] B. A. Bash, D. Goeckel, D. Towsley, and S. Guha, "Hiding information in noise: Fundamental limits of covert wireless communication," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 26–31, Dec. 2015.



**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 PhD in physics from the University of Kansas in 1998, Dr. Nikitin 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 US patents.



**RUSLAN L. DAVIDCHACK** is Professor of Mathematical Modelling and Computation in the School of Mathematics and Actuarial Science at the University of Leicester, United Kingdom. His research interests are in developing computational methods and tools for applications in molecular simulations, nonlinear dynamics and signal processing.

Dr. Davidchack received his undergraduate degree in theoretical condensed matter physics from Lviv University, Ukraine, and a PhD 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. Dr. Davidchack has an extensive track record of interdisciplinary collaborations which span academic institutions (e.g. Georgia Tech; UC Berkeley; Oxford, UK; EPFL, Switzerland; NTNU, Norway; KTH, Sweden), national laboratories in the USA (e.g. AFRL, Oak Ridge, Sandia), and UK (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, UK, France, Germany, South Korea, China, and Ukraine.

Dr. Davidchack is a co-inventor of the AVATAR methodology, which provided the initial foundation for the technology of intermittently nonlinear outlier noise mitigation.

• • •