Received: 10 April 2018 | Revised: 1 November 2018 | Accepted: 4 January 2019

DOI: 10.1002/dac.3917

RESEARCH ARTICLE

WILEY

# Time diversity scheme and adaptive signal clipping with blanking applied to G3 systems for narrowband power-line communications

Yves-François Rivard | Ling Cheng

School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

**Correspondence**
Ling Cheng, School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa.
Email: Ling.Cheng@wits.ac.za

**Funding information**
National Research Foundation, Grant/Award Number: SFH150710124964 and 112248

**Summary**

In this paper, an investigation into the behavior of modern narrowband power-line communication systems when concatenated with both a fountain code named Luby transform code and a combination of nonlinear preprocessing techniques called signal clipping with blanking is presented. The new systems can be used in applications such as the automatic meter reading component of smart grid technology. Testing is performed on a range of systems provided by the G3 standard and the results show that a bit error rate improvement can be achieved with the correct encoder/decoder settings. The recommended settings include an adaptive technique that uses Gaussian elimination for Luby transform code decoding and code rates that depend on the system specifications used. From there, it is shown that a further improvement is achieved when both the Luby transform code and adaptive signal clipping/blanking based on signal power are combined. It is found that these modifications present a complexity trade-off, which is higher on the receiver side. This added complexity can be well-tolerated in automatic meter-reading systems because of the star topology and asymmetric nature of the network, which has a centralized agent acting as a receiver that collects data from various transmitter nodes.

**KEYWORDS**

CENELEC, LT code, nonlinear preprocessor, power-line communications, time diversity

## 1 | INTRODUCTION

Power-line communications (PLC) is a technique that allows for the installation of a communication system over long or short distances by the reuse of existing power lines, acting as a communication channel. The advantage of this technique is that the costs of installation are reduced, the disadvantage, however, is that PLC systems operate over a very noisy channel. Other electrical systems sharing the network introduce various types of interference on the channel. As the demand for smart power management increases, techniques such as automatic meter reading that utilize the PLC channel for communication purposes have been developed and are being actively researched to provide an efficient backbone for smart grid applications.[1-5]

In an attempt to commercialize narrowband PLC networks and ensure compatibility as well as safety compliance, several organizations have established standards giving insight as to which specifications these systems should adhere to. These standards typically operate in low-frequency bands such as the CENELEC bands, ranging from 3 kHz to 148.5 kHz

  

as defined in the EN 50065-1 standard.[6] Two leading standards in this domain which are often researched[7,8] are the eRDF G3-PLC standard[9] and the Ibrerdola PRIME standard,[10] both specifying the physical layer as systems designed around baseband, real-valued orthogonal frequency-division multiplexing (OFDM) transmission signals with specific forward-error correction (FEC) chains.

Time diversity techniques have been proposed for overcoming the challenges inherent to using the PLC channel as a communications medium.[11,12] The proposed methods involve concatenating systems with low complexity FEC techniques such as a repetition code or higher complexity techniques belonging to a family of codes named fountain codes. This research focusses on Luby transform (LT) codes, which are a type of rateless fountain code where any number of packets $n$ can be generated from a set of $k$ input packets.[13] LT codes have successfully been applied in the context of indoor broadband PLC including applications such as the distribution of high-quality video over PLC.[14,15] Further research into LT codes resulted in the development of raptor codes, thus increasing the performance by precoding the input data with the use of a low-density parity-check (LDPC) code.[16] By reversing the order of the raptor code and LDPC code in the FEC block chain, ie, using the LT code as an outer code and LDPC code as an inner code, a further increase in performance can be achieved when applied over the PLC channel.[17] An improvement is obtained by using properties of the LDPC code to identify decoding failures and thus treat faulty packets as erasures. The use of the LDPC code detection technique thus approximates the behavior of LT codes over an erasure channel and prevents errors from propagating during the LT decoding process. Furthermore, time diversity can be combined with spatial diversity to further improve PLC systems.[18,19] The author of a previous study[18] suggests that a space-time coded multiple-input and single-output configuration can outperform traditional single-input single-output PLC systems. In the context of low-voltage networks, this type of space-time diversity technique has been achieved by transmitting the PLC signal over the three conductors of the three-phase electrical network (spatial diversity) to a single receiver and using time-division multiplexing (time diversity). Space-time diversity has also been applied in the form of distributed space-time block code in systems consisting of a PLC transceiver and a series of repeaters that are either set with an adaptive hybrid amplify-and-forward mode or decode-and-forward mode.[19]

Nonlinear preprocessing techniques that include clipping, nulling (blanking), hybrid clipping/blanking, and deep clipping are typically used at the transmitter of OFDM systems to reduce the peak-to-average power ratio but can also be used at the receiver of communication systems when the channel is affected by impulsive noise (IN) such as in the case of PLC.[20-23] The signal clipping method involves limiting the possible amplitude values of a waveform to a specified threshold value $T_c$. Clipping has been used successfully for combating IN in multicarrier systems operating over noisy environments such as terrestrial digital video broadcasting[24,25] as well as the PLC channel.[26-28] Compared with clipping, which limits the signal amplitudes to the value $T_c$, the blanking technique involves setting the waveform amplitude value to zero when the received signal amplitude is above a certain threshold $T_b$. In both the cases of clipping and blanking, formulas have been devised to calculate the optimal $T_c$ and $T_b$ threshold values based on the modeled characteristics of the IN. The challenge with the optimal threshold formulas is that the characteristics of the IN need to be known accurately by the receiver a priori,[21] which is generally not the case in real-world systems. As such, methods have been devised where the thresholds can be determined dynamically from SNR estimations or signal peak value calculations performed in real time by the receiver. As an example, significant improvement in bit error rate (BER) for systems using OFDM in conjunction with quadrature phase-shift keying and affected by IN can be achieved through the use of an adaptive clipping technique where the threshold $T_c$ is dependent on the signal-to-noise ratio. Furthermore, clipping is often combined with blanking where the threshold $T_b$ is higher than the threshold $T_c$.[28,29] For these hybrid clipping/blanking systems, the threshold $T_b$ is usually set to a scalar multiple of the clipping threshold, ie, $T_b = \alpha T_c$, where $\alpha$ is typically set to 1.4, but it has been shown that when it is made to be adaptive based on the received signal characteristics, a further increase in performance is achieved.[23]

These ideas can be taken a step further by combining and applying them to systems provided by current day standards such as G3 systems, which can be used to communicate over various networks for smart grid applications. The first advantage is that G3 smart meter systems have already been adopted in several areas,[30,31] making the task of performing the modification on live systems more practical should the new idea prove to be beneficial. The second advantage is that in G3 systems, inner codes are short especially when compared with LDPC codes, which require much larger codes for a significant increase in performance. Shorter inner codes also result in protection over smaller packets being used resulting in less corrupted data per packet when a burst of bits get damaged.

The paper is organized as follows: in Section 2 the research system models are fully specified: First, both the selected current up-to-date systems used to obtain benchmark results provided by the G3 standard and the channel model are given. Second, modifications to the G3 systems are performed by concatenating them with an LT code used as an outer

code to obtain new original systems. Third, the clipping and blanking techniques used to further modify the new systems in the research are described. The results obtained are presented and analyzed in Section 3 on both performance and complexity level. Possible future work stemming from this research is discussed in Section 4; and finally, in Section 5, a conclusion is drawn.

## 2 | SYSTEM MODELS

### 2.1 | Benchmark G3 system specifications

To establish baseline results for comparison and make the findings of this research applicable to the current industry, G3 systems provided by the current up-to-date standard are used as benchmark systems. Alternatively, systems provided by PRIME could have been used but have been shown to be inferior in performance albeit being less complex.[7] A G3 system can possess different characteristics based on the data rate and performance required by the application. The general overall system block diagram composed of a Reed-Solomon (RS) code, a convolutional code (CC), a repetition code, an interleaver, and an OFDM transceiver can be seen in Figure 1.

The input to the system is a randomly generated stream of binary data represented by $\mathbf{u}$. As an example, $\mathbf{u}$ could contain information about the readings of electrical meters that are being monitored over the PLC network. The FEC chain is a concatenated system that is comprised of an RS code for the outer layer, which converts the binary stream $\mathbf{u}$ to the encoded symbols $\mathbf{v_{RS}}$, followed by a CC, which further encodes the RS symbols into the binary vector $\mathbf{v_{CC}}$. Following this, the output of the CC is passed through a repetition code, which generates the repeated binary vector $\mathbf{v_{REP}}$. Finally, the interleaver produces the input to the OFDM modulator $\mathbf{X_k}$ by spreading out the data both in time across OFDM symbols and in frequency across OFDM subcarriers, which protects the data against burst errors. The outer code is an RS encoder over GF($2^m$) where $m$ is equal to 8 and which possesses either eight or 16 parity bits depending on the robustness required. Furthermore, shortening of the RS code is then performed either on a (255,239) or a (255,247) code depending on the frame size and required data rate. The convolutional encoder is a zero-terminated, half-rate encoder defined by the generator $\mathbf{g} = [171\ 133]$ in octal form, a constraint length of 7, and a free distance of 10 bits. The repetition code repeats data a total number of four times and is thus a (4, 1) code.

The decoders in the receiver associated with the FEC schemes are not defined in the standard as it is deemed to be outside the scope of the available documents but they have been implemented as follows for the purpose of this research: First, the de-interleaver component transforms the received symbols $\mathbf{Y_k}$ into $\overline{\mathbf{v_{REP}}}$ by performing the inverse binary mixing operation as that of the interleaver block. Majority logic decoding is used for the repetition code to convert the vector $\overline{\mathbf{v_{REP}}}$ to $\overline{\mathbf{v_{CC}}}$, where a bit is picked at random in the event of a tie. Decoding of the CC to obtain $\overline{\mathbf{v_{RS}}}$ is performed by the Viterbi algorithm using hard decision decoding where it is assumed that a traceback length of five times the constraint length is used for the optimal trade-off between performance and decoding time.[32] Finally, the decoder in the RS layer uses the Berlekamp-Massey algorithm, which results in $\overline{\mathbf{u}}$.

Signal transmission is performed by a baseband OFDM modulator and demodulator, which can operate with binary phase-shift keying, differential binary phase-shift keying (DBPSK), or differential quaternary phase-shift keying resulting in a varying modulation order $M$. The OFDM time signal can be represented mathematically by

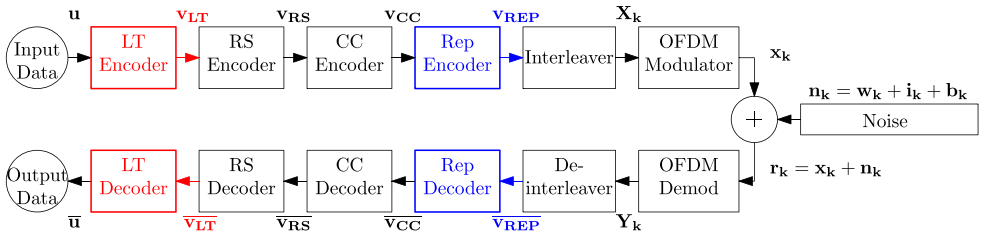


**FIGURE 1** System block diagram where the repetition code (represented by the blue blocks) is used in the baseline G3 power-line communication (PLC) system, and the Luby transform (LT) code (represented by the red blocks) is used to replace the repetition encoder in the LT-modified G3 system[9]

$$x(t) = \sum_{n=0}^{N_{SC}-1} X_n e^{j2\pi nt/T}, \quad 0 \le t < T, \tag{1}$$

where $N_{SC}$ is the number of subcarriers and $T$ is the OFDM symbol period. It should be noted that in the case of G3 systems, only real-valued signals are used. The signal is then sampled to produce $\mathbf{x_k}$. G3 systems operating with time diversity are set with the most robust settings available and thus make use of DBPSK only. The OFDM component uses a cyclic prefix of 30 samples, a raised cosine window spanning a range of 8 bits on both ends of each symbol and 256 FFT points. The frame size depends on the number of OFDM subcarriers $N_{SC}$, which is usually of 36 spread across the 35.9 kHz to 90.6 kHz frequency range as well as the number of OFDM symbols $N_{SYM}$ used per frame, which can vary between 40 and 252 symbols. The modulator settings result in a data rate, which varies between 2423 bit/s to 5592 bit/s.

## 2.2 | PLC channel model

The PLC channel is harsh and plagued with several types of non-Gaussian noises as it is shared with many other electrical systems. The noises can be categorized as colored background noise (BN), narrowband interference (NBI), and IN.[33-35]

The BN $w(k)$ is colored and therefore has a power spectral density that varies with frequency but is usually simulated as additive white Gaussian noise (AWGN). BN is generally more severe on the medium voltage network than on the low voltage network.[36]

NBI typically appears as a modulated sinusoidal signal and therefore occurs in specific frequency bands. This type of noise usually originates from sources such as television horizontal scanning frequencies or AM radio broadcasting stations. The NBI time signal can be represented mathematically from a modified equation[37] as follows:

$$b(t) = \sum_{i=1}^{N} A_i(t) sin(2\pi f_i + \theta_i) imp\left(\frac{t - t_{arr,i}}{t_{w,i}}\right), \tag{2}$$

where the $N$ NBI waveforms are defined by five parameters, namely the NBI amplitudes as a function of time $A_i(t)$, the NBI center frequencies $f_i$, the NBI phases $\theta_i$, the NBI arrival times $t_{arr,i}$, and the NBI durations $t_{w,i}$. It should also be noted that $imp(t)$ is a generalized impulse function that has an amplitude value of one and a width of one. In the frequency domain, NBI can be viewed as affecting a subset of the total OFDM subcarriers used in the available spectrum. As the G3 system possesses an adaptive tone-mapping feature that prevents the OFDM system from using bands that are deemed to be unusable after channel estimation, carriers affected by NBI can be avoided. The resulting assumption is therefore that the sampled NBI signal $b(k)$ can be ignored for the purpose of this research.

IN is the most destructive in the PLC channel as it can have a high power spectral density value resulting in it taking over the transmitted signal. The IN time signal is mathematically defined as[37]

$$i(t) = \sum_{i} A_i \; imp\left(\frac{t - t_{arr,i}}{t_{w,i}}\right), \tag{3}$$

where $A_i$ is the set of amplitudes of each impulse, $t_{w,i}$ is the width of the impulses, $t_{arr,i}$ is the time when each of the impulses starts, and $imp(t)$ is as previously defined. The IN amplitudes $A_i$ are randomly generated to provide values resulting in a maximum power spectral density reaching levels up to 50 dB greater than that of the BN.[34] The width $t_{w,i}$ of the impulses is uniformly distributed over the 0 to 1 ms range. When sampled, the IN signal is represented by $i(k)$. Finally, the arrival times are calculated from the inter-arrival times generated by inverse transform sampling of an exponential cumulative distribution function (CDF) of a Poisson process

$$F(t) = 1 - e^{\lambda t}, \tag{4}$$

where $\lambda$ is the rate parameter taking values between 1/0.015 and 1/0.003 and which determines how often impulses occur. Higher $\lambda$ values therefore result in worse channel conditions. Once the impulses and transmission signal samples have been generated over the duration of the current transmission window, they are both added to generate the received signal as follows:

$$r(k) = x(k) + w(k) + i(k) = x(k) + n(k), \tag{5}$$

where $r(k)$ is the summation of the transmission signal samples $x(k)$ and noise signal samples $n(k)$.

## 2.3 | LT-modified G3 system specifications

In the new proposed system, time diversity is implemented by concatenating G3 systems with an LT code as opposed to the repetition code. The LT code is thus used as an outer code as shown in Figure 1 by the red block. This particular setup is chosen such that the LT decoder may use information provided by the inner code, which attempts to correct errors and aid during the decoding process.[17] The first advantage of using the new configuration is that the LT code is concatenated with the code provided in the specification, which results in a new scheme, which is highly compatible with existing G3 systems. The second advantage is that this configuration has the effect of allowing the LT code to operate under conditions approximating a binary erasure symmetric channel by dropping corrupted LT packets in instances where errors are successfully detected. Finally, the third advantage is that as the LT code replaces the repetition code in the system, the data rate of the payload significantly increases.

In the encoding section of the transmitter, the outer LT code encoder produces the vector $\mathbf{v_{LT}}$ from the input data, which then gets further processed by the remainder of the FEC chain. On the receiver side, the LT code decoder generates $\overline{\mathbf{u}}$ from $\overline{\mathbf{v_{LT}}}$ by using a Gaussian elimination (GE) based decoder.[38] Further details about both the LT code encoder and decoder algorithms can be found in Appendix A.

## 2.4 | Signal clipping and blanking

As stated in Section 1, nonlinear preprocessing techniques such as clipping and blanking can be implemented based on the IN model characteristics. The problem with determining and utilizing the IN model characteristics in live systems is that they vary over time and errors in their estimation results in a non-negligible system degradation.[21] For the aforementioned reason, in this research, various clipping thresholds $T_c$ are tested based on the standard deviation $\sigma$ of the transmitted signal affected by AWGN and IN. The motivation behind using this technique can be seen by plotting the probability density function (PDF) of the transmitted signal affected by the different types of noises. First, when plotting the PDF of the transmitted OFDM signal time samples, a discrete distribution approximating a Gaussian distribution is obtained with standard deviation $\sigma_{OFDM}$ as seen in Figure 2. Second, plotting the PDF of the time samples originating from the OFDM signal affected by AWGN results in a discrete distribution where the variance $\sigma^2$ is equal to the addition of the clean signal variance $\sigma^2_{OFDM}$ and AWGN variance $\sigma^2_{AWGN}$ as seen in Figure 3. Finally, the PDF of the signal with AWGN and IN is seen in Figure 4.

As observed, most of the sampled values occur in the same region as in the PDF of Figure 3, characterised by a large spike centered around 0. Several smaller spikes caused by the IN then appear at magnitudes much larger than the range covered by the main spike and are therefore further away from it. It can thus be concluded that signal clipping should be able to remove these spikes occurring at larger amplitude values and help protect the system against IN. To further improve the performance of the modified systems as shown in Rabie and Alsusa's study,[23] blanking is also added with the

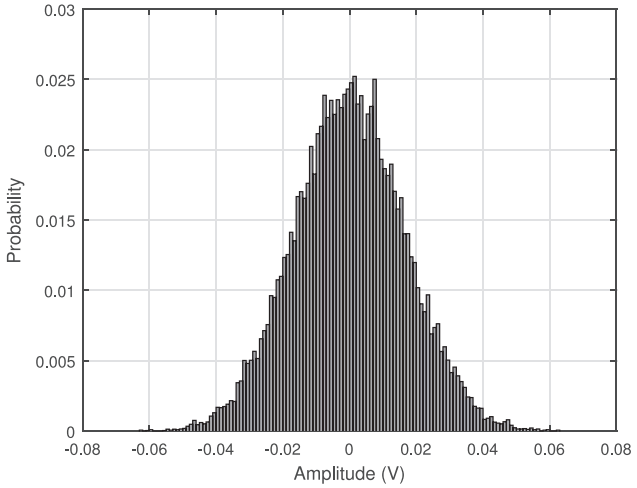


**FIGURE 2** Probability density function (PDF) of transmitted OFDM signal without noise and with a variance of $\sigma^2_{OFDM}$ with a value of $5.5 \cdot 10^{-4}$

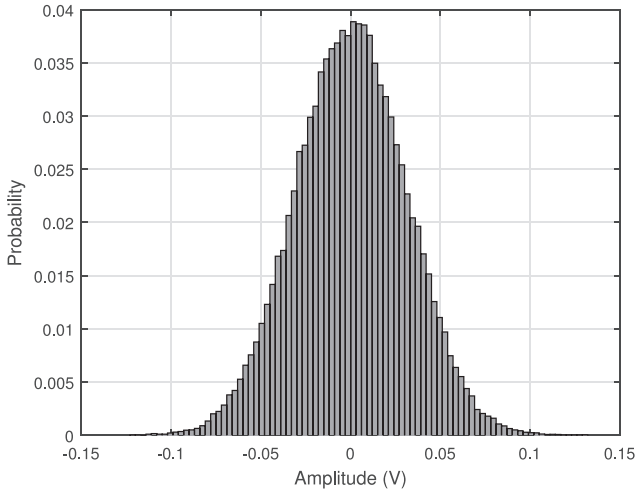


**FIGURE 3** Probability density function (PDF) of transmitted orthogonal frequency-division multiplexing (OFDM) signal affected by additive white Gaussian noise (AWGN) with variance $\sigma^2$ equal to $\sigma^2_{OFDM} + \sigma^2_{AWGN}$ where $\sigma_{OFDM}$ has the same value as in Figure 2 and $\sigma_{AWGN}$ has a value of $1.7 \cdot 10^{-4}$

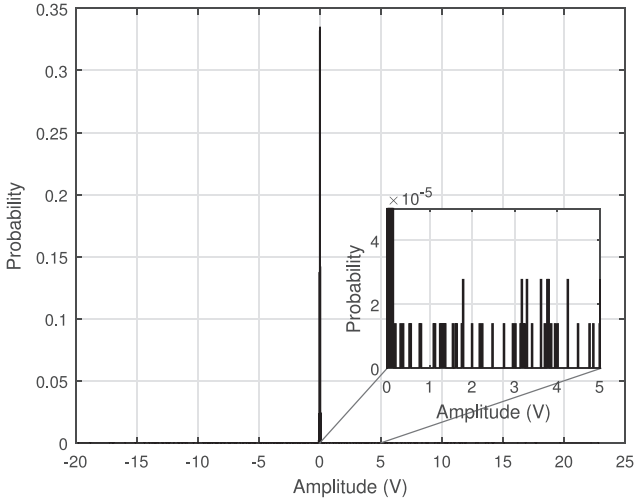


**FIGURE 4** PDF of transmitted orthogonal frequency-division multiplexing (OFDM) signal affected by additive white Gaussian noise (AWGN) and impulsive noise (IN)

blanking threshold $T_b$ set at the value of $1.4T_c$. The discrete signal $\bar{r}(k)$ is then obtained by applying the combined clipping and blanking procedure to the received signal $r(k)$ as follows[39]:

$$\bar{r}(k) = \begin{cases} r(k) & |r(k)| \leq I\sigma \\ I\sigma e^{j \arg(r(k))} & I\sigma < |r(k)| \leq 1.4I\sigma \\ 0 & |r(k)| > 1.4I\sigma \end{cases}, \tag{6}$$

where $k \in \{0, 1, \ldots, N-1\}$ and the scalar $I > 0$ is a constant. For the implementation of this technique, it is assumed that the value of the signal standard deviation $\sigma$ is available to the receiver from channel estimation. The PDF of a clipped signal with an $I$ value of 3 such that $T_c = 3\sigma$ can be seen in Figure 5, where the presence of two spikes at the amplitude values of $T_c$ and $-T_c$ can be observed.

The receivers of both the benchmark G3 and original LT-modified G3 systems are then modified further with the addition of an adaptive clipping with blanking component to test the validity of the new scheme. The performance provided by the adaptive clipping with blanking technique is then analyzed by the result of a second set of simulations performed with the system configurations shown in Figure 6.

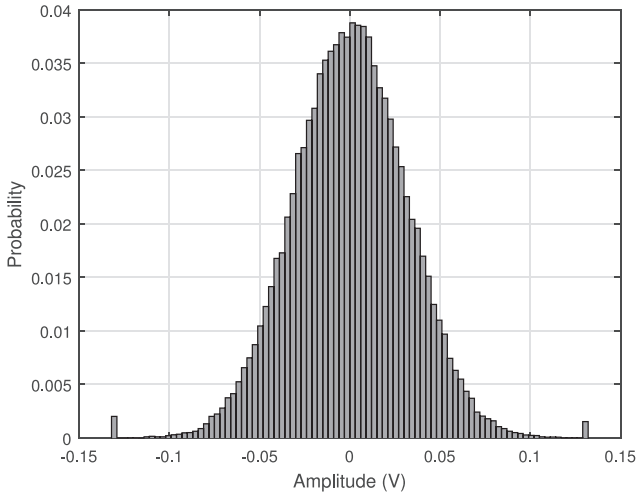


**FIGURE 5** PDF of transmitted clipped orthogonal frequency-division multiplexing (OFDM) signal affected by additive white Gaussian noise (AWGN) and impulsive noise (IN) with a $T_c$ value of $3\sigma$

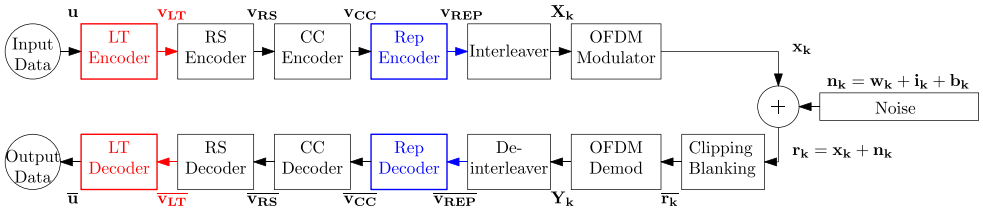


**FIGURE 6** System block diagram with the addition of nonlinear pre-processor signal clipping with blanking, where the repetition code (represented by the blue blocks) is used in the clipping + blanking G3 power-line communications (PLC) system, and the Luby transform (LT) code (represented by the red blocks) is used to replace the repetition encoder in the new clipping + blanking LT-modified G3 system[9]

## 3 | SIMULATION RESULTS AND ANALYSIS

In this section, the plots resulting from the simulations are first shown in Figure 7 to Figure 9 and subsequently analyzed. The simulation result plots, grouped based on IN channel conditions, contain the simulation results for all system configurations and are color-coded based on their $N_{SYM}$ values, which can be of 40, 56, or 252 OFDM symbols. In all cases, the metric utilized to determine when the simulation should end is the total number of bit errors. For statistical accuracy, every simulation therefore ends when at least 1000 bit errors have occured.[32]

### 3.1 | Current industry standard G3 performance analysis

To establish a baseline, simulations are first run with G3 systems provided by the standards. The baseline results are shown by the clean solid line in Figure 7 to Figure 9. Three different G3 systems are tested to cover the range of system specifications containing time diversity under the form of repetition coding. The tested systems are therefore based on implementations with an $N_{sym}$ value of 40, 56, and 252 OFDM symbols resulting in data rates of 2423 bit/s, 3257 bit/s, and 5592 bit/s, respectively. The G3 systems are then tested over three PLC channels, which have different impulse rate characteristics, ie, worst case scenario ($\lambda = 1/0.003$ in Figure 7), best case scenario ($\lambda = 1/0.015$ in Figure 9) and the midpoint ($\lambda = 1/0.009$ in Figure 8), resulting in a total of nine different test cases.

From the G3 simulation results, it can be concluded that the presence of IN in the PLC channel results in a performance behavior different from channels affected by AWGN only. Specifically, a noise floor is present following the initial waterfall region, which occurs at different $E_b/N_0$ and BER values depending on the system code specifications. As $\lambda$ increases, the occurrence of impulses increases, resulting in a worse performance. Furthermore, as the $N_{SYM}$ value decreases, the code rate of the RS codes get lowered while the error correcting capability remains constant, resulting in a smaller energy per

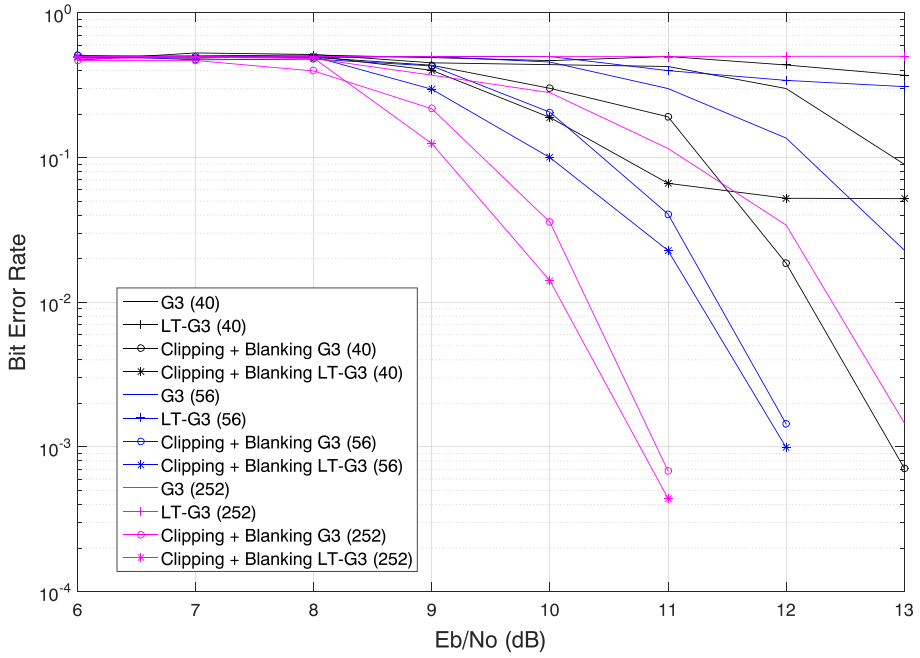


**FIGURE 7** Bit error rate (BER) vs $E_b/N_0$ for systems with 40 orthogonal frequency-division multiplexing (OFDM) symbols ((21, 13) RS), 56 OFDM symbols ((30, 22) RS) and 252 OFDM symbols ((141, 133) RS) over power-line communications (PLC) channel with a $\lambda$ value of 1/0.003

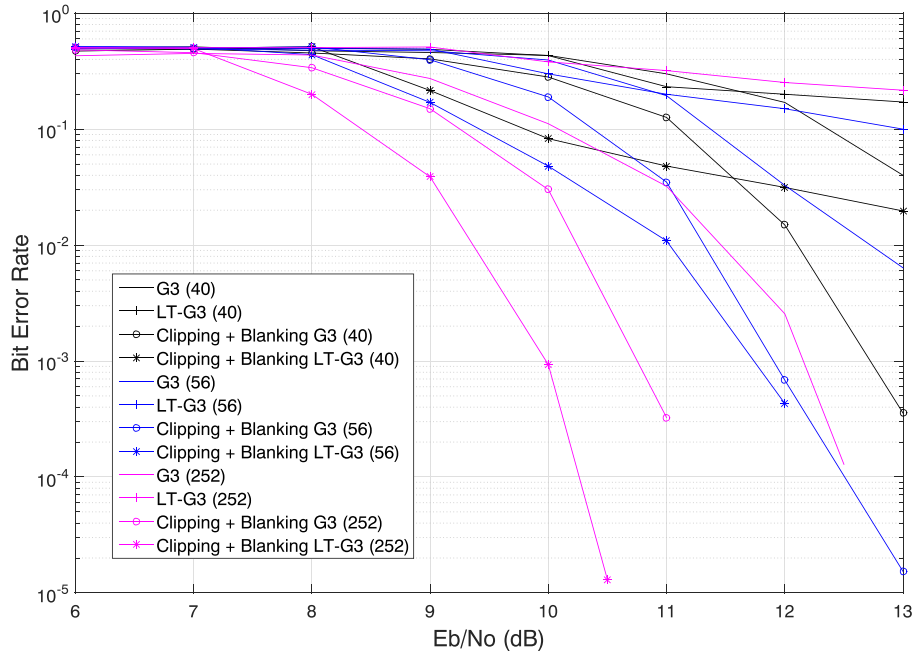


**FIGURE 8** Bit error rate (BER) vs $E_b/N_0$ for systems with 40 orthogonal frequency-division multiplexing (OFDM) symbols ((21, 13) RS), 56 OFDM symbols ((30, 22) RS) and 252 OFDM symbols ((141, 133) RS) over power-line communications (PLC) channel with a $\lambda$ value of 1/0.009

bit value $E_b$ and thus worsening the performance.[40] An example of worsening behavior from lowering the $N_{SYM}$ value can be seen when comparing the G3 plots in Figure 7. The set of plots in all simulation Figures (Figure 7 to Figure 9) can be ranked from best to worst in all three channel conditions based on the $N_{SYM}$ values: systems with 252 symbols outperform the ones with 56 OFDM symbols, which in turn outperform G3 set-ups with 40 OFDM symbols.

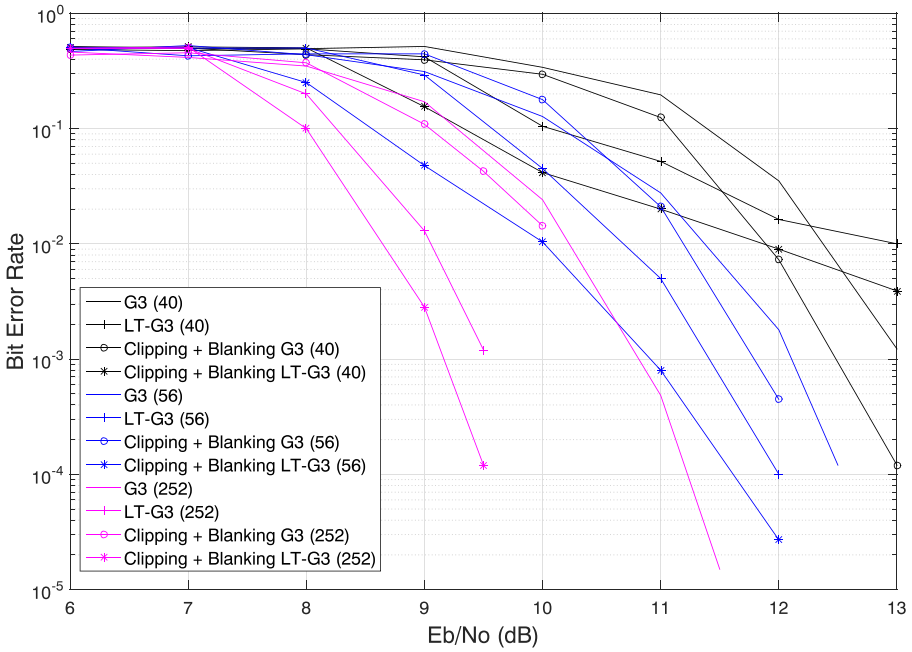


**FIGURE 9** Bit error rate (BER) vs $E_b/N_0$ for systems with 40 orthogonal frequency-division multiplexing (OFDM) symbols ((21, 13) RS), 56 OFDM symbols ((30, 22) RS) and 252 OFDM symbols ((141, 133) RS) over power-line communications (PLC) channel with a $\lambda$ value of 1/0.015

## 3.2 | LT-modified G3 system performance analysis

Simulations are performed to determine the behavior of the new LT-G3 scheme on the PLC channel under the same conditions as those used previously in the baseline systems, ie, $N_{sym}$ values of 40, 56, and 252 with $\lambda$ values of 1/0.015, 1/0.009, and 1/0.003, respectively, while keeping the RS code rates associated with systems of different OFDM symbol lengths constant. To implement the new LT-G3 systems, the number of LT source packets $k$ and the variables of the robust Soliton distribution must first be selected. Generally, it is proven that LT codes perform better as the number of source packets $k$ is increased.[17] In the case of G3 systems, this could be achieved by splitting data from the LT transmitter across multiple frames, but the drawback would be added non-negligible latency. The latency would be due to the receiver waiting on all or most of the frames to arrive before successful decoding could occur. To prevent this and adhere to the G3 specifications as much as possible, an enforced system constraint is that LT encoded data is restricted to a single frame. It should be kept in mind that because of this constraint, the number of total available bits per frame is dictated by the frame size, which is of $N_{sc} \cdot N_{sym} \cdot M$ bits. For the robust Soliton distribution, a $\delta$ value of 0.02 is typically chosen[17] with a spike value $Q$ dependent on the number of source packets chosen in each system. In these scenarios, the $Q$ values resulting in the best performance is empirically determined to occur at 10, 20, and 50 for the systems with an $N_{SYM}$ value of 40, 56, and 252, respectively. The smallest practical $l$ value in all system cases is 10. Following this, the optimal LT data rates must then be selected for all systems. As performance is analyzed based on the BER vs $E_b/N_0$ values, there exists a relationship between performance and LT code rate for the decoder implementation as shown in Figure 10.

As observed, for the LT modified G3 system with an $N_{SYM}$ value of 252 and using GE, the best code rate found that can be used is 1/1.05. The same process must be repeated for every system with a different $N_{SYM}$ value, as this varies in each case. To understand the behavior of the decoding method, tests are performed by collecting samples of how many packets are required at every iteration for successful decoding when packets are being continuously transmitted. From this data, the CDF of the decoding probability depending on the code rate is plotted. Example results obtained from systems with 252 OFDM symbols and 304 LT source packets are shown in Figure 11 for comparison.

The series of plots from Figure 11 represent three points along the $E_s/N_0$ range where the waterfall region occurs. $E_s/N_0$ values lower than those shown result in an unusable system with a BER of 0.5, while higher values result in a CDF with a faster rise time. It is reminded that in this example, because of the frame size limit constraint, and the selected source packet value $k$ of 304, that code rates greater than 1/1.4 cannot be used as this would result in the LT code data being sent through multiple frames, but the results have been shown here to aid in illustrating that this would further increase the decoding probability for a given $E_s/N_0$. In the curves, different identifiable steps occurring at regular intervals can be

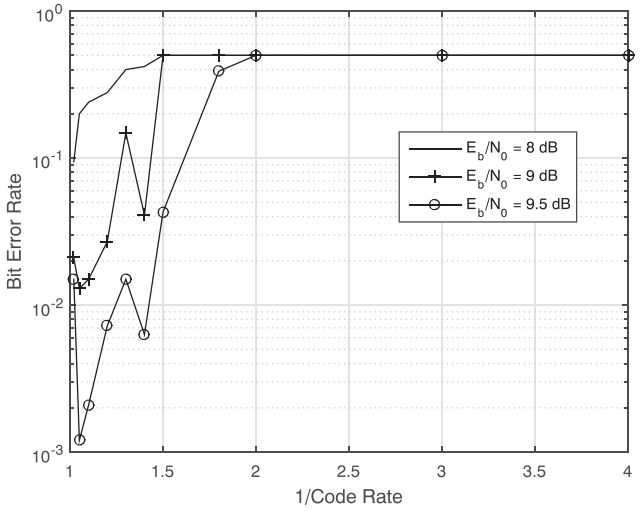


**FIGURE 10** Performance depending on code rate at various $E_b/N_0$ values for Luby transform (LT)-modified G3 system with 252 orthogonal frequency-division multiplexing (OFDM) symbols, (141, 133) Reed-Solomon (RS) code, $\lambda$ value of 1/0.015 and a $Q$ value of 50 using a Gaussian elimination (GE) decoder

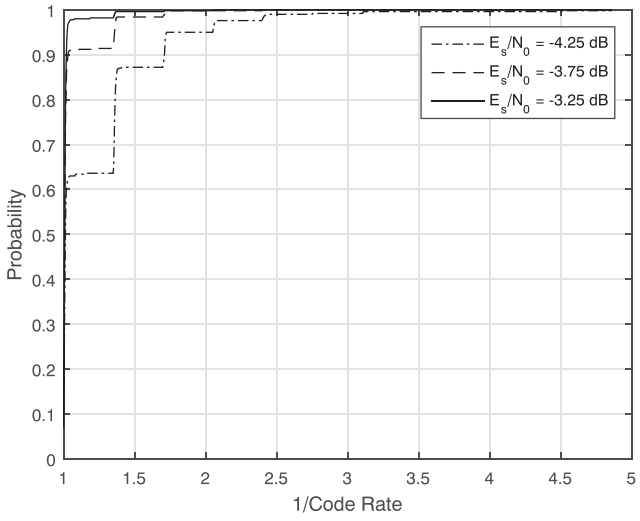


**FIGURE 11** Decoding probability based on Luby transform (LT) code rate for an LT-modified G3 system with 252 orthogonal frequency-division multiplexing (OFDM) symbols, (141, 133) RS code, $\lambda$ value of 1/0.015 and a $Q$ value of 50 using a Gaussian elimination (GE) decoder

noticed. This is explained by the fact that new LT packets arrive in bulk, which are the size of an RS decoded codeword. In this new system, each frame contains four RS codewords since the RS code is kept constant and the removed repetition code has a rate of 1/4. As seen, the GE method has a probability of decoding from a low packet number (even when the code rate is 1). The G3 and LT-G3 results are summarized in Table 1, wherein each row representing a different test condition, the best system is highlighted.

When the LT-G3 simulations are compared with the results of the G3 systems, a few observations can be noted. First, it is observed that for two of the nine cases, a non-negligible coding gain is observed. Both of the cases with positive results occur under the best case channel conditions with a $\lambda$ value of 1/0.015, which is shown in Figure 9. The first, labelled "LT-G3 (56)," is for the OFDM system with 56 symbols where a coding gain of up to about 0.5 dB can be observed followed by the OFDM system labelled "LT-G3 (252)" with 252 symbols where a coding gain of up to about 1 dB is obtained. For the rest of the systems (specifically "LT-G3 (40)" in Figure 9 and all systems of Figure 7 and Figure 8), noise floors can be examined, which occur earlier than with the unmodified G3 systems. It should also be noted that for lower values of $\lambda$, systems start performing worse as the number of OFDM symbols $N_{SYM}$ is decreased. This behavior is once again because

**TABLE 1** G3 and Luby transform (LT)-modified G3 performance comparison

| System Specifications | $\frac{E_b}{N_0}$ | G3 System | LT-G3 System |
|---|---|---|---|
| $N_{SYM} = 40$ | - | - | - |
| $\lambda = 1/0.003$ | 13 | $9 \cdot 10^{-2}$ | $3.7 \cdot 10^{-1}$ |
| $\lambda = 1/0.009$ | 13 | $4 \cdot 10^{-2}$ | $1.7 \cdot 10^{-1}$ |
| $\lambda = 1/0.015$ | 13 | $1.2 \cdot 10^{-3}$ | $1 \cdot 10^{-2}$ |
| $N_{SYM} = 56$ | | - | - |
| $\lambda = 1/0.003$ | 12 | $1.3 \cdot 10^{-1}$ | $3.4 \cdot 10^{-1}$ |
| $\lambda = 1/0.009$ | 12 | $3.2 \cdot 10^{-2}$ | $1.5 \cdot 10^{-1}$ |
| $\lambda = 1/0.015$ | 12 | $1.8 \cdot 10^{-3}$ | $1.4 \cdot 10^{-4}$ |
| $N_{SYM} = 252$ | - | - | - |
| $\lambda = 1/0.003$ | 11 | $1.1 \cdot 10^{-1}$ | $5 \cdot 10^{-1}$ |
| $\lambda = 1/0.009$ | 10.5 | $7.1 \cdot 10^{-2}$ | $3.4 \cdot 10^{-1}$ |
| $\lambda = 1/0.015$ | 9.5 | $9.8 \cdot 10^{-2}$ | $1.2 \cdot 10^{-3}$ |

of the worsening performance of shortened RS codes, as is the case with the benchmark systems. For cases with a higher value of $\lambda$, the situation changes and systems with a higher $N_{SYM}$ value can start performing worse. The reason behind this behavior is that as the $N_{SYM}$ value increases and thus RS codeword length increases, so does the probability of there being multiple impulses per codeword, which results in a threshold of packets being corrupted beyond the decoding capability of the decoder.

The cause of the problem resulting in a worse performance in most of the test cases is twofold: First, the inner code is not strong enough to correct the errors and allow for the LT code to operate successfully. Second, not enough information is provided by the inner code for the LT decoder to pinpoint which packets have been corrupted in a batch of packets the size of a RS codeword, resulting in too many packets being discarded. Nevertheless, should it be deemed to be a worthwhile upgrade, it is recommended that an adaptive system be implemented. In this new adaptive system, an LT code should be used when the channel conditions are deemed to be better than a certain threshold and the initially specified repetition code should be used when the channel conditions are below that same threshold, resulting in a system that has the best available performance according to the results presented. This topic can therefore be researched further to determine what the exact threshold conditions should be.

## 3.3 | Clipping with blanking performance analysis

As stated in Section 2.4, to determine whether the combination of LT code and nonlinear preprocessing techniques called signal clipping and blanking offers an improvement, tests are performed by adding the combined clipping and blanking method to both the benchmark G3 systems and LT-modified G3 systems. Once again, simulations are performed to cover the same range of $N_{SYM}$ and $\lambda$ values as the previous tests so that the results may be compared. As with the LT code, the first step involves finding the optimal settings, if any, resulting in the best achievable BER. In this case, the BER is calculated for various clipping thresholds $T_c$ and plotted so that the best value may be identified for multiple systems. It is reminded that the blanking threshold $T_b$ is set to $1.4T_c$. An example is given in Figure 12 for a clipping-modified G3 system.

As can be seen, for G3 systems with 252 OFDM symbols, the best adaptive $T_c$ value found is at $1.5\sigma$. The decrease in performance from lowering $T_c$ below $1.5\sigma$ is because of too much of the original signal getting clipped, while the decrease in performance for increasing $T_c$ above $1.5\sigma$ is caused by less of the impulses being clipped. It is found that the optimal adaptive value of $1.5\sigma$ is constant for both the G3 and LT-modified G3 systems with various combinations of $E_b/N_0$ and $\lambda$ values. Simulations are run to analyze the performance with the addition of a clipping module using $T_c = 1.5\sigma$ and a blanking module using $T_b = 1.4T_c$. The summary of the results that can be compared with the results of Table 1 are tabulated in Table 2, where once again the best overall result in each case is highlighted.

From the data, it is seen that in all cases, systems using adaptive clipping with blanking have a better BER than their nonclipped counterparts. It is also observed that for six of the nine tested settings, concatenating an LT code in combination with adaptive signal clipping with blanking provides an improvement in BER performance when compared with the clipped-only G3 systems. The systems that do not benefit from the combination of both an LT code and adaptive signal

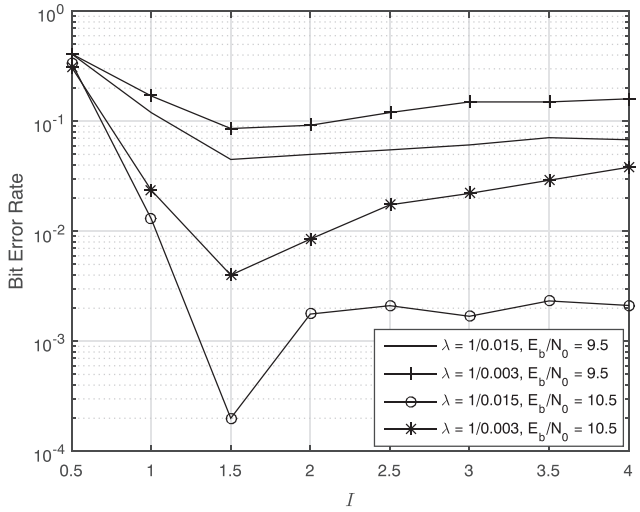


**FIGURE 12** Bit error rate (BER) vs $I$ values (refer to Equation 6) for G3 system with 252 orthogonal frequency-division multiplexing (OFDM) symbols, varying impulse rate parameter and varying $E_b/N_0$

**TABLE 2** Performance comparison of G3 and LT-modified G3 systems using adaptive signal clipping with blanking

| System Specifications | $\frac{E_b}{N_0}$ | Clipped | |
|---|---|---|---|
| | | G3 | LT-G3 |
| $N_{SYM} = 40$ | - | - | - |
| $\lambda = 1/0.003$ | 13 | $7.1 \cdot 10^{-4}$ | $5.2 \cdot 10^{-2}$ |
| $\lambda = 1/0.009$ | 13 | $3.6 \cdot 10^{-4}$ | $2 \cdot 10^{-2}$ |
| $\lambda = 1/0.015$ | 13 | $1.2 \cdot 10^{-4}$ | $3.9 \cdot 10^{-3}$ |
| $N_{SYM} = 56$ | - | - | - |
| $\lambda = 1/0.003$ | 12 | $1.4 \cdot 10^{-3}$ | $9.9 \cdot 10^{-4}$ |
| $\lambda = 1/0.009$ | 12 | $6.8 \cdot 10^{-4}$ | $4.3 \cdot 10^{-4}$ |
| $\lambda = 1/0.015$ | 12 | $4.4 \cdot 10^{-4}$ | $2.7 \cdot 10^{-5}$ |
| $N_{SYM} = 252$ | - | - | - |
| $\lambda = 1/0.003$ | 11 | $6.8 \cdot 10^{-4}$ | $4.4 \cdot 10^{-4}$ |
| $\lambda = 1/0.009$ | 10.5 | $3 \cdot 10^{-2}$ | $9.3 \cdot 10^{-4}$ |
| $\lambda = 1/0.015$ | 9.5 | $4.2 \cdot 10^{-2}$ | $1.2 \cdot 10^{-4}$ |

Abbreviation: LT, Luby transform.

clipping with blanking are the systems with the smallest $N_{SYM}$ value of 40, although it should be noted that all clipped G3 systems perform better than their baseline counterparts. It is thus recommended to use adaptive signal clipping with blanking in all cases when operating over the PLC channel but to only use a concatenated LT code when a high enough $N_{SYM}$ value is used, which depends on the data rate requirements of the application.

## 3.4 | System complexity comparison

Differences in complexity between the G3 and clipped LT-modified G3 systems originate from the components that are not shared between them. Specifically, the differences in complexity reside in the LT code and clipping and blanking components that have been concatenated, as the RS code, CC code, and interleaver components are kept constant amongst implementations that use the same number of OFDM symbols $N_{SYM}$.

The encoding of LT codes is performed by the XOR of $O\left(ln\left(\frac{k}{\delta}\right)\right)$ input packets on average as this is the average degree of an encoded packet.[13] Each symbol XOR operation is composed of $l$ individual XOR operations resulting in an encoding cost the order $O(\log k)$ after simplification.[41] The decoder implementation with GE has a complexity that depends on the

size of the decoding matrix where a matrix with $k$ rows is solved with costs the order of $O(k^3)$.[42] For signal clipping with blanking, the complexity is of the order $O(N)$ where $N$ is the number of sample points in the time signal.

As can be seen from the analysis, concatenating G3 systems with both an LT code and adaptive signal clipping with blanking increases performance under certain conditions but has the drawback of increasing complexity, especially on the receiver side. The transmitter modification is considered to be of low complexity as XOR operations are cheap to perform. Automatic meter reading systems are typically asymmetric as they are built as a network composed of several nodes (electrical/gas/water meters) communicating through the PLC channel with a centralized agent, which then processes the data.[43] This can be seen as a star network topology. Modification through the use of LT codes and signal clipping with blanking would thus be well-tolerated by the system as the receiver modifications that have higher complexity would be implemented in the centralized agent component, which possesses more available resources, while only the transmitter modifications with lower complexity would need to be installed on the individual meter nodes.

During the decoding process, another variable that needs to be considered is the delay brought on by the decoder. Specifically, decoding using GE can add considerable delay as decoding only starts once at least $k$ packets have been received, which are then all processed as a batch. Each time the process fails, a newly received packet is added to the matrices and the whole GE decoding process starts over. It has been shown that this delay can be reduced by modifying the GE decoding algorithm to process the received packet as they arrive and not completely restart the GE process when decoding is unsuccessful if more packets are still to be available.[38,44]

## 4 | FUTURE WORK

Several different paths can be explored stemming from this original research: First, as stated in Section 3.2, research could be performed to determine what the IN threshold value is to decide when exactly the adaptive LT-modified G3 system should switch between using the repetition code and the LT code. Second, further research could be made to determine the optimal adaptive threshold value $\alpha$ based on the IN characteristics. Finally, since the results presented in this research are obtained from software simulation, possible future work could involve obtaining physical measurements, which further validate the research. It is suggested that the physical measurements should be obtained from a system built using FPGA technology as a proof of concept.

## 5 | CONCLUSION

It has been shown that current day systems provided by up-to-date standards such as the G3 standard can be improved under certain conditions by concatenating them with both a type of fountain code named LT code and a nonlinear preprocessing techniques called hybrid signal clipping and blanking.

The addition of an LT code is demonstrated to successfully improve the performance of G3 systems by providing a non-negligible coding gain in two of the nine scenarios when using an LT decoder based on GE and with an optimal code rate. In the remaining seven cases that did not result in a performance improvement, the cause is found to be that the inner code is not powerful enough to provide the required quantity of clean packets for successful error-free decoding by the LT decoder, resulting in errors that propagate across packets. In addition, the inner code also does not provide enough information for the LT decoder to identify corrupted packets and therefore drop them. With the use of LT coding only, an adaptive scheme is thus recommended, which switches between LT coding and repetition coding based on the channel conditions (more specifically how severe the IN is).

When G3 systems are combined with both LT coding and adaptive signal clipping with blanking, six of the nine systems show an improvement when compared with the benchmark results. In G3 systems combining LT codes and clipping with blanking, coding gains are obtained provided that the number of transmitted symbols is high enough, resulting in a higher code rate, which in turn means that a larger number of clean packets are available to the LT decoder. It should be noted that the trade-off from increasing the number of transmitted symbols is a larger decoding delay at the receiver as the decoder then has to wait for more encoded packets to arrive before the decoding process can be completed. The overall recommendation is thus to use the LT-clipping-blanking combination with the clipping threshold $T_c$, based on the noise levels, set at $1.5\sigma$ and the blanking threshold set to $1.4T_c$ when G3 systems are used with OFDM symbols not set on the lowest setting (ie, G3 systems with higher code rates). It is also noted that adaptive clipping with blanking should

be used in all system configuration cases as from the simulation results, it always showed a non-negligible performance improvement when compared with the baseline systems.

The modifications result in a complexity trade-off, which is investigated. It is noted that the added complexity can be well-tolerated considering that it resides mainly in the receiver and that in the context of smart grid automatic meter reading applications, the receiver is the centralized system that possesses more available resources than the smaller units attached to the meters being monitored.

## ACKNOWLEDGMENTS

The financial assistance of the South African National Research Foundation (NRF) toward this research is hereby acknowledged (grant number SFH150710124964 and 112248). Opinions expressed and conclusions arrived at, are those of the author and not necessarily to be attributed to the NRF.

The author would like to thank Gareth Timm for the time spent proof reading this paper.

## CONFLICTS OF INTEREST

The authors declare no potential conflict of interests.

## FINANCIAL DISCLOSURE

None reported.

## ORCID

*Yves-François Rivard* https://orcid.org/0000-0002-3136-9375
*Ling Cheng* http://orcid.org/0000-0001-7873-8206

## REFERENCES

1. Park BS, Hyun DH. Implementation of AMR system using power line communication. In: IEEE Transmission and Distribution Conference and Exhibition; 18; Yokohama, Japan (6-10 October 2002).
2. Sivaneasan B, Gunawan E, So PL. Modeling and performance analysis of automatic meter-reading systems using PLC under impulsive noise interference. *IEEE Trans Power Delivery*. 2010;25(3):1465-1475.
3. Papadopoulos TA, Kaloudas CG, Chrysochos AI, Papagiannis GK. Application of narrowband power-line communication in medium-voltage smart distribution grids. *IEEE Trans Power Delivery*. 2013;28(2):981-988.
4. Kim Y, Bae JN, Kim JY. Performance of power line communication systems with noise reduction scheme for smart grid applications. *IEEE Trans Consum Electron*. 2011;57(1):46-52.
5. Son Y-S., Moon K-D, Kim C. Home energy management system based on power line communication. *IEEE Trans Consum Electronics*. 2010;56(3):1380-1386.
6. EN 50065-1. *Signalling on Low-voltage Electrical installations in the Frequency Range 3 kHz to 148.5 kHz - Part 1: General Requirements, Frequency Bands and Electromagnetic Disturbances*. Belgium: CENELEC; 2011.
7. Hochg M. Comparison of PLC G3 and PRIME. In: IEEE International Symposium on Power Line Communications and its Applications; 2011; Udine, Italy (3-6 April 2011):165-169.
8. Ogunyanda K, Familua AD, Swart TG, Ferreira HC, Cheng L. Evaluation of mixed permutation codes in PLC channels, using Hamming distance profile. *Springer Telecommun Syst*. 2016;65(1):1-11.
9. PLC G3 OFDM. *PLC G3 Physical Layer Specification*. France: eRDF; 2008.
10. Revision 1.4. *PRIME Specification for PoweRline Intelligent Metering Evolution*. Belgium: Iberdrola; 2014.
11. Cheng L, Swart TG, Ferreira C. Adaptive rateless permutation coding scheme for OFDM-based PLC. In: IEEE 17th International Symposium on Power Line Communications and Its Applications; 2013; Johannesburg, South Africa:242-246. (24-27 March 2013).
12. Cheng L, Ferreira HC. Time-diversity permutation coding scheme for narrow-band power-line channels. In: IEEE International Symposium on Power Line Communications and Its Applications; 2012; Beijing, China:120-125. (27-30 March 2012).
13. Luby M. LT codes. In: The 43rd Annual IEEE Symposium on Foundations of Computer Science; 2003; Vancouver, BC, Canada:271-280. (19-19 November 2002).
14. Amirshahi P, Navidpour SM, Kavehrad M. Fountain codes for impulsive noise correction in low-voltage indoor power-line broadband communications. In: 3rd IEEE Consumer Communications and Networking Conference; 2006; Las Vegas, NV, USA:473-477. (8-10 January 2006).
15. Luby M, Watson M, Gasiba T, Stockhammer T. High-quality video distribution using power line communication and application layer forward error correction. In: IEEE International Symposium on Power Line Communications and Its Applications; 2007; Pisa, Italy:431-436. (26-28 March 2007).

16. Shokrollahi A. Raptor codes. *IEEE Trans Inf Theory*. 2006;52(6):2551-2567.
17. Andreadou N, Tonello AM. On the mitigation of impulsive noise in power-line communications with LT codes. *IEEE Trans Power Delivery*. 2013;28(3):1483-1490.
18. Mahasneh H. In designing of ST-coded in-home PLC transmission scheme. *Int J Comput Appl*. 2016;141(2):1-5.
19. Tseng S-M, Lee T-L, Ho Y-C, Tseng D-F. Distributed space-time block codes with embedded adaptive AAF/DAF elements and opportunistic listening for multihop power line communication networks. *Int J Commun Syst*. 2015;30(1):2950.
20. Juwono FH, Guo Q, Huang D, Wong KP. Deep clipping for impulsive noise mitigation in OFDM-based power-line communications. *IEEE Trans Power Delivery*. 2014;49(3):1335-1343.
21. Alsusa E, Rabie KM. Dynamic peak-based threshold estimation method for mitigating impulsive noise in power-line communication systems. *IEEE Trans Power Delivery*. 2013;28(4):2201-2013.
22. Juwono FH, Guo Q, Chen Y, Xu L, Huang D, Wong KP. Linear combining of nonlinear preprocessor for OFDM-based power-line communications. *IEEE Trans Smart Grid*. 2016;7(1):253-260.
23. Rabie KM, Alsusa E. Performance analysis of adaptive hybrid nonlinear preprocessors for impulsive noise mitigation over power-line channels. In: IEEE ICC 2015 SAC - Communications for the Smart Grid; 2015; London, UK:728-733. (8-12 June 2015).
24. Zhidkov SV. Impulsive noise suppression in OFDM based communication systems. *IEEE Trans Consum Electron*. 2003;49(4):5-9.
25. Suraweera HA, Chai C, Shentu J, Armstrong J. Analysis of impulse noise mitigation techniques for digital television systems. In: 8th International OFDM Workshop; Hamburg, Germany:172-176. (24-25 September 2003).
26. Al-Mawali KS, Hussain ZM. Adaptive-threshold clipping for impulsive noise reduction in OFDM-based power line communications. In: IEEE 2009 International Conference on Advanced Technologies for Communications; 2009; Hai Phong, Vietnam:43-48. (12-14 October 2009).
27. Al-Mawali KS, Sadik AZ, Hussain ZM. Joint time-domain/frequency-domain impulsive noise reduction in OFDM-based power line communications. In: IEEE Australasian Telecommunication Networks and Applications Conference; 2008; Adelaide, SA, Australia:138-142. 7-10 December 2008.
28. Papilaya VN, Han Vinck AJ. Investigation on a new combined impulsive noise mitigation scheme for OFDM transmission. In: 17th IEEE International Symposium on Power Line Communications and its Applications; 2013; Johannesburg, South Africa:86-91. (24-27 March 2013).
29. Zhidkov SV. Analysis and comparison of several simple impulsive noise mitigation schemes for OFDM receivers. *IEEE Trans Commun*. 2008;53(1):5-9.
30. Cano C, Pittolo A, Malone D, Lampe L, Tonello AM, Dabak A. State-of-the-art in power line communications: from the applications to the medium. *IEEE J Sel Areas Commun*. 2016;34(7):1935-1952.
31. Zannoni T, Campagna F. LTE and PLC technologies for MV network supervision and automation. In: CIRED Workshop; Roma, Italy: 1-4. (11-12 June 2014).
32. Haykin S. *Communication Systems*. New York: John Wiley and Sons; 2000.
33. Hooijen OG. On the channel capacity of the residential power circuit used as a digital communications medium. *IEEE Commun Lett*. 1998;2(10):267-268.
34. Zimmermann M, Dostert K. Analysis and modeling of impulsive noise in broad-band powerline communication. *IEEE Trans Electromagn Compat*. 2002;44(1):249-258.
35. Shongwe T, Papilaya VN, Han Vinck AJ. Narrow-band interference model for OFDM systems for powerline communications. In: 17th IEEE International Symposium on Power Line Communications and Its Applications; March 24-27, 2013. Johannesburg, South Africa. 268-272.
36. Lee J-J, Choi S-J, Oh H-M, Lee W-T, Kim K-H, Lee D-Y. Measurements of the communications environment in medium voltage power distribution lines for wide-band power line communications. In: International Proceedings on Power Line Communications and its Applications; 2004; Zaragosa, Spain. 69-74.
37. Zimmermann M, Dostert K. An analysis of the broadband noise scenario in powerline networks. In: IEEE International Symposium on Powerline Communications; 2000; Limerick, Ireland. 131-138.
38. Boglio V, Grangetto M, Gaeta R, Sereno M. On the fly gaussian elimination for LT codes. *IEEE Commun Lett*. 2009;13(12):953-955.
39. Mowla M, Ali Y, Aoni RA. Performance comparison of two clipping based filtering methods for PAPR reduction of OFDM signal. *Int J Mobile Network Commun Telematics*. 2014;4(1):23-34.
40. Deutschi LJ. The effects of Reed-Solomon code shortening on the performance of coded telemetry systems. TDA Progress Report. NASA; 14-20; 1983.
41. Guede MA. Optimization of the belief propagation algorithm for Luby transform decoding over the binary erasure channel. *Master thesis*: Delft University of Technology; 2011.
42. Ryaben'kii VS, Tsynkov SV. *A Theoretical Introduction to Numerical Analysis*. New York: Chapman and Hall; 2006.
43. Wei C, Yang J. Implementation of automatic meter reading system using PLC and GPRS. *J Inf Comput Sci*. 2008;8:4343-4350.
44. Kim S, Ko K, Chung SY. Incremental Gaussian elimination decoding of Raptor codes over BEC. *IEEE Commun Lett*. 2008;12(4):307-309.
45. Tirronen T. Optimal degree distributions for LT codes in small cases. *Master thesis*: Helsinki University of Technology; 2005.
46. Rivard Y-F. Comparative study of a time diversity scheme applied to G3 systems for narrowband power-line communications. *Master thesis*: University of the Witwatersrand; 2016.

**How to cite this article:** Rivard Y-F, Cheng L. Time diversity scheme and adaptive signal clipping with blanking applied to G3 systems for narrowband power-line communications. *Int J Commun Syst*. 2019;e3917. https://doi.org/10.1002/dac.3917

## APPENDIX A: LUBY TRANSFORM CODE THEORY

This Appendix provides further insight about the inner workings of both the Luby transform (LT) code encoder and decoder used in this paper. The aim of this section is to provide theoretical knowledge about LT codes so that the reader may have a greater understanding of the results presented in this research.

### A.1 | LT code encoder

The LT encoder produces any number $n$ of encoded packets represented as $v_{LT}$ from $k$ input packets as follows: First, a degree $d$ is generated from a probabilistic distribution of choice. Second, $d$ packets are selected at random from the set of source packets by means of a uniform distribution. Finally, these packets are combined by the XOR operation, which results in an encoded packet that is ready for transmission. The first probabilistic distribution designed for the operation of LT codes is known as the Soliton distribution and is defined by the probability density function (PDF)[13]

$$p(1) = \frac{1}{k}, \tag{A1}$$

$$p(d) = \frac{1}{d(d-1)} \quad \text{for} \quad d = 2, 3...k. \tag{A2}$$

This distribution results in a PDF, which possesses a single spike at a low value resulting in many encoded packets of low degree $d$. Modifications to the ideal distribution exist to counteract the inefficiencies in the decoding process resulting in the creation of the robust Soliton distribution. Contrary to the ideal version, this PDF possesses an extra spike at a larger value $Q$, which is determined by the parameters of this distribution. The robust Soliton distribution PDF is[45]

$$\tau(d) = \frac{1}{iQ} \quad \text{for } i = 1, 2, \ldots, Q-1, \tag{A3}$$

$$\tau(d) = \frac{\ln\left(\frac{R}{\delta}\right)}{Q} \quad \text{for} \quad i = Q. \tag{A4}$$

$$\tau(d) = 0 \quad \text{for} \quad i = Q+1, \ldots, k, \tag{A5}$$

$$Z = \sum_{d=1}^{k} p(d) + \tau(d), \tag{A6}$$

$$\mu(d) = \frac{p(d) + \tau(d)}{Z}, \tag{A7}$$

where $R$ is equal to $k/Q$ and $Z$ is a normalization factor to ensure $\mu(d)$ satisfies the property of a PDF, namely that the total area under it is equal to 1. The choice of values for the parameters $\delta$ and $Q$ has an influence on the average number of combined packets per transmitted packet and the number of packets which are required at the decoder for the decoding operation to be successful.

The final variable that needs to be specified for the LT code component is the length $l$ of the LT packets. This variable should be picked such that it is as small as possible resulting in more source packets from a given amount of input data but has the constraint that it must be large enough such that the possible packet space, which is of size $2^l$, is larger than the number of source packets, ie, the minimum number of packets required in the decoding process. This condition must be met to avoid packets that are linearly dependent and which provide no new information during the decoding process, ie, resulting in a packet filled with bits of the value 0 after the XOR operation.

## A.2 | LT code decoder

Decoding of LT codes that generates $\overline{\mathbf{u}}$ from $\overline{\mathbf{v_{LT}}}$ can be performed by an iterative message passing algorithm such as belief propagation, or processes involving GE.[38] The choice of belief propagation or GE for the LT decoding procedure presents a trade-off between performance and complexity.[41] An LT decoder based on GE is implemented in this research as it has been shown that they outperform belief propagation decoders when used with G3 systems over the PLC channel.[46] For the decoding operation to begin, information about which source packets have been combined in each received encoded packet is required. It is possible to transmit the encoding information alongside the information data, but this method poses the risk of it getting corrupted and also results in a slower data rate. For this reason, it is assumed that both the transmitter and receiver have this prior information available to them in the form of matching lookup tables in memory or generated from a pseudorandom number generator with a common seed.

The GE decoding method can be accomplished once at least $k$ linearly independent encoded packets are received. A set of equations is first set up in the form $\mathbf{A}x = \mathbf{b}$. The elements of matrix $\mathbf{A}$ are 1 if packet $x_i$ where $i \in \{1, 2, \ldots, k\}$ is part of encoded packet $y_j$ where $j \in \{1, 2, \ldots, n\}$ and are 0 if it is not. The vector $x$ represents all the source packets $x_i$ and the vector $\mathbf{b}$ represents the encoded packets $y_j$. An example of the GE procedure is shown in Figure A1.

$$(a)\begin{bmatrix}1 & 0 & \ldots & 0\\ 1 & 1 & \ldots & 0\\ 1 & 1 & \ldots & 1\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 1 & \ldots & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_k\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ y_3\\ \vdots\\ y_n\end{bmatrix}\quad b)\begin{bmatrix}1 & 0 & \ldots & 0\\ 0 & 1 & \ldots & 0\\ 0 & 0 & \ldots & 1\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_k\end{bmatrix}=\begin{bmatrix}y'_1\\ y'_2\\ y'_3\\ \vdots\\ y'_n\end{bmatrix}$$

$$(c)\begin{bmatrix}1 & 0 & \ldots & 0\\ 0 & 1 & \ldots & 0\\ 0 & 0 & \ldots & 1\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_k\end{bmatrix}=\begin{bmatrix}x_1\\ x_2\\ x_3\\ \vdots\\ 0\end{bmatrix}$$

**FIGURE A1** Gaussian elimination (GE) decoding procedure overview. (A) Received packets inserted into matrix. (B) Matrix **A** is triangulated. (C) Back-substitution is performed to obtain the original source packets