Robust Decoding for Convolutionally Coded Systems Impaired by Memoryless Impulsive Noise

Der-Feng Tseng, *Member, IEEE*, Yunghsiang S. Han, *Fellow, IEEE*, Wai Ho Mow, *Senior Member, IEEE*, Po-Ning Chen, *Senior Member, IEEE*, Jing Deng, *Senior Member, IEEE*, and A. J. Han Vinck, *Fellow, IEEE*

***Abstract*—It is well known that communication systems are susceptible to strong impulsive noises. To combat this, convolutional coding has long served as a cost-efficient tool against moderately frequent memoryless impulses with given statistics. Nevertheless, impulsive noise statistics are difficult to model accurately and are typically not time-invariant, making the system design challenging. In this paper, because of the lack of knowledge regarding the probability density function of impulsive noises, an efficient decoding scheme was devised for single-carrier narrowband communication systems; a design parameter was incorporated into recently introduced joint erasure marking and Viterbi decoding algorithm, dubbed the metric erasure Viterbi algorithm (MEVA). The proposed scheme involves incorporating a well-designed clipping operation into a Viterbi algorithm, in which the clipping threshold must be appropriately set. In contrast to previous publications that have resorted to extensive simulations, in the proposed scheme, the bit error probability performance associated with the clipping threshold was characterized by deriving its Chernoff bound. The results indicated that when the clipping threshold was judiciously selected, the MEVA can be on par with its optimal maximum-likelihood decoding counterpart under fairly general circumstances.**

***Index Terms*—Impulsive noise, Bernoulli-Gaussian channel, Middleton Class-A model, metric erasure Viterbi Algorithm (MEVA), power line communications.**

I. Introduction

THE communication community has expressed renewed interest in coded transmissions in communication systems that are impaired by non-Gaussian noises. Some of the primary impetuses are that the capacity-achieving coding is already relatively well understood in the context of the additive white Gaussian noise (AWGN) model, and nuisances, such as man-made electromagnetic interference and atmospheric noises, are the limiting factors in numerous realistic situations such as power line communications (PLC) [1], digital subscriber line (DSL) loops [2], and wireless communication systems [3]. Thus, communication system designs and their performance in non-Gaussian noise models must be further investigated. The subsequent man-made noise models, which have been widely adopted for performance analysis, can be found in [4] and [5]. Both [6] and [7] studied how binary signaling and quadrature amplitude modulation were affected by the impulsive noise model of [8], provided that the probability density function (PDF) of the impulsive noise was available at the receiver.

Previous studies have employed channel coding [1], [9], verifying that it withstands the obstruction caused by additive white impulsive noise. In addition to using channel coding to compensate for impulsive-noise-related losses, the orthogonal frequency division multiplexing (OFDM) technique, a multi-carrier modulation scheme, has been shown to resist impulses in certain circumstances [5]. Notably, [5] showed that, an OFDM system is outperformed by its single-carrier counterpart in terms of uncoded symbol error probability when substantial impulsive noise energy was combined with a moderately or severely high probability that impulses would occur. Another study [10] indicated that the information rate of OFDM systems was lower compared with their single carrier counterparts except for systems that operated at an extremely high spectral efficiency. These facts warranted the investigation of coded single carrier narrowband communication systems that are subject to frequent impulses, of which the average power is much greater than is that of the background noise.

The study focuses on single carrier narrowband communication systems such as those used in power line communications for Smart Grid applications. For example, so-called Low Data Rate (LDR)[1] Narrowband Power Line Communications (NB-PLC) are typically single-carrier based and transmit data at rates of a few kilobits per second (kbps). The NB-PLC includes devices that conform to the following standards: ISO/IEC 14908-3 (LonWorks), ISO/IEC, 14543-3-5 (KNX), CEA-600.31 (CEBus) and IEC 61334 (FSK and Spread-FSK). Several additional non-SDO-based examples include the Insteon, X10, HomePlug C&C, SITRED, Ariane, Controls, and BacNet. The listed transceivers (primarily the LonWorks and IEC 61334) are the most deployed transceivers in the world, and perhaps hundreds of millions have been employed. Compared with the LDR NB-PLC standards, the multi-carrier

Manuscript received February 12, 2013; revised August 10, 2013. The editor coordinating the review of this paper and approving it for publication was S. Galli.

D.-F. Tseng and Y. S. Han are with the Dept. of Electrical Engineering, National Taiwan Univ. of Sci. and Tech., Taipei, Taiwan. (e-mail: dtseng@mail.ntust.edu.tw).

W. H. Mow is with the Dept. of Electronic and Computer Engineering, Hong Kong Univ. of Sci. and Tech., Hong Kong.

P.-N. Chen is with the Dept. of Electrical and Computer Engineering, National Chiao Tung Univ., Hsinchu, Taiwan.

J. Deng is with the Dept. of Computer Science, Univ. of North Carolina, Greensboro, North Carolina.

A. J. Han Vinck is with the Institute for Experimental Mathematics, University of Duisburg-Essen, Essen, Germany.

This work was supported by the National Science Council of Taiwan under NSC-101-2221-E-011-072-MY2, NSC 101-2221-E-011-069-MY3, and NSC 101-2221-E-009-008-MY3. The work of the third author was supported by the Hong Kong RGC under GRF Project No. 616512. Parts of this work were presented at The Fifth International Workshop on Signal Design and its Applications in Communications (IWSDA'11).

Digital Object Identifier 10.1109/TCOMM.2013.101813.130122

[1]The taxonomy on LDR/HDR (together with the technologies listed) was first introduced in [11] (see Sect. I.B in [11])

(OFDM) based standards, which are referred to as High Data Rate (HDR) NB-PLC, allow higher data rates (typically up to a few hundred kbps). Examples of the HDR NB-PLC include the G3-PLC (ITU-T Recommendation G.9903) and PRIME (ITU-T Recommendation G.9904), which were both ratified in 2012, and IEEE P12901.2, which is expected to be ratified in 2014.

A. Related work

As previously mentioned, channel coding has been widely employed in communication systems that operate in impulsive noise environments. For instance, convolutional coding was applied in [12], turbo coding was applied in [13], and low density parity-check (LDPC) coding was applied in [14]. For more details, refer to Chapter 5 in [15] and the references therein. The performance limits for communication systems corrupted by impulsive noise were meticulously derived in [16] and [10], which provide insight into the degree to which the capacity loss is incurred when comparing them with the impulse-free systems. For a Discrete Multitone (DMT) system in an impulsive noise channel model, which is conceived as the concatenation of an AWGN channel and an erasure channel, the channel capacity is derived in [17], in which a capacity-approaching LDPC code is also pursued. Operating on the assumption that the PDF of the impulsive noise was known at the receiver, and based on the coding scheme of [1], [1] and [9] investigated the Chernoff bounds of the pairwise error probability (PEP) for the optimal and suboptimal receivers in real and complex Middleton Class-A channels. Subsequently, [18] presented general expression of the exact PEP in certain impulsive noise channels. The Chernoff bound on the bit error probability (BEP) for ideally interleaved coded OFDM systems was calculated in [19], which assumed that the Central Limit Theorem could be applied to the Fourier transform outputs of time-domain noise samples, in which the PDF is also assumed known.

In contrast to the aforementioned publications that acquired coding gain by accounting for the exact PDF of impulsive noise at the receiver, numerous studies [20]–[23] have investigated using a certain clipping-based decoding metric to reduce the performance degradation induced by strong impulsive noise and forgo the computationally demanding estimation of impulsive noise statistics. In situations where code concatenation is considered, [24] shows that the latency that results from interleaving can be greatly reduced using an outer code. Notably, in [24], the decoding metric that accommodates the impulsive noise was not provided for the inner code decoder, of which the implicated functionality is to precisely erase the corrupted bytes prior to outer code decoding. The results in [24] showed an enhanced data rate for a Discrete Multitone Very-high-bit-rate Digital Subscriber Loop (DMT-VDSL) system. Notably, in the aforementioned references, the clipping thresholds were primarily devised based on extensive computer simulation, which can be time-consuming or computationally prohibitive, especially in regard to low-bit-error-rate applications. Relatively recently, the joint erasure marking and Viterbi decoding algorithm (JEVA) introduced in [18] was shown to outperform a conventional scheme, in which the erasure marking to received signals and the erasure Viterbi decoding operate sequentially. The simulations in [18] showed that in certain circumstances, the JEVA could achieve a level of performance similar to that of the BEP performance of the optimal maximum likelihood decoding (MLD).

B. Main contributions of this paper

- Despite a lack of statistical knowledge regarding impulses, a robust decoding scheme is proposed, which incorporates a design parameter, p into the new approximate decoding metric derived from the MAP decoding rule. Because of the memoryless nature of the impulsive noise channels, the proposed decoding scheme is analogous to performing clipping based on the decoder metric [16], [22], [23], [21], in which the clipping threshold is induced by p. Note that a similar metric-clipping based decoding algorithm was reported in [16], [22], [23],[2] in which a fixed robust clipping threshold of $\Delta = 10^{-3}$ was chosen after extensive simulations. Although these studies used a similar clipping concept, their definition of the clipping threshold of Δ differed from that of the current study.
- After invoking the Chernoff bound on the BEP, an analytical method is proposed for choosing the value of the clipping threshold (or equivalently, the design parameter, p). Our uniquely derived method can be applied to scenarios in which performing extensive computer simulation becomes unaffordable (e.g., for low-BEP applications), shedding light on how the design parameters should be adjusted as the system parameters vary. When system settings similar to those in [16] and [22] were used, the results indicated that a fixed robust choice of p justified the choice of $\Delta = 10^{-3}$ in [16], [22], [23] (using the derived Chernoff bound in the study), potentially resulting in a level of system performance close to that of a maximum likelihood decoder, which requires extensive statistical knowledge of impulsive noises. The simulation results indicated that at a high probability of impulse occurrence and at various levels of Impulse-Gaussian power Ratio (IGR), both of which are assumed unavailable to the receiver, the clipping threshold suggested by the derived Chernoff bound on the BEP induces a substantial performance gain compared with the schemes of [16] and [22].
- To explore the feasibility of the proposed decoder, the effect of the analog-to-digital converter (ADC) on the level of BEP performance was also investigated using simulations. Similar to the conventional impulse-free case, it was shown that when the dynamic range of the ADC is property adjusted, the performance loss that results from quantization error is approximated to be a small constant in terms of the signal-to-noise ratio in decibel (dB), similar to that occurs in the conventional impulse-free case.

The paper is organized as follows. The primary impulsive noise models and their system frameworks are reviewed in

[2]The proposed robust decoding scheme was inspired by the joint erasure marking and Viterbi decoding algorithm [18]. We became aware of the relevant work reported in [16], [22], [23] after completing an early version of this paper. We acknowledge the anonymous reviewers, who brought [16], [22], [23] to our attention, suggesting the necessary comparisons.

Section II, and the details of the proposed efficient decoding algorithm are introduced in Section III. A performance evaluation for the proposed decoding scheme, which is based on the Chernoff bound on the BEP, is detailed in Section IV. The simulation results are presented in Section V, and a conclusion is offered in Section VI.

II. System Model

Let $\boldsymbol{C}$ represent a $(n,\ k,\ m)$ convolutional code (CC) comprising an information sequence of length kL bits, where n code bits are produced blockwise in response to each block input of k information bits, and m is the memory length of the convolutional encoder. Because m blocks of k-bit zeros are appended to the end of the information sequence to clear the contents of the shift registers, the length of each codeword is $N = n(L+m)$. During its transmission, the convolutionally coded sequence is subjected to memoryless impulsive noises that are characterized, but not restricted by, the Bernoulli-Gaussian model [5] or the Middleton Class-A model [8], in which each noise sample, n_j at any time j, is the sum of an AWGN noise g_j and an impulsive noise i_j (i.e., $n_j = g_j + i_j$). As a convention, the AWGN noise $\{g_j\}_{j=0}^{N-1}$ has a flat single-sided power spectral density (PSD) of height N_0.

In the Bernoulli-Gaussian model [5], the impulsive noise, i_j, can be further represented as a product of two independent variables; in other words, $i_j = b_j\omega_j$, where $b_j \in \{0,1\}$ is a Bernoulli random variable that expresses the probability of occurrence of impulses $\Pr(b_j = 1) = p_b$, and $\{\omega_j\}_{j=0}^{N-1}$ is an independent and identically distributed (*i.i.d.*) Gaussian random process with a mean of zero and a variance of $\frac{N_0}{2}\Gamma$. Throughout this paper, $\Pr(\cdot)$ denotes the probability of the event inside the parentheses. Notably, the constant Γ is the mean power ratio of impulsive noise, ω_j relative to AWGN noise g_j; thus, it is exactly the IGR. The PDF of n_j in the Bernoulli-Gaussian model is thus given by

$$\begin{aligned} f_{\mathrm{BG}}(x) &= (1-p_b)\cdot\frac{1}{\sqrt{N_0/2}}\varphi\left(\frac{x}{\sqrt{N_0/2}}\right) \\ &\quad + p_b\cdot\frac{1}{\sqrt{(N_0/2)(1+\Gamma)}}\varphi\left(\frac{x}{\sqrt{(N_0/2)(1+\Gamma)}}\right)\ , \end{aligned} \tag{1}$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}}\exp\{-x^2/2\}$ is the Gaussian PDF with zero mean and unit variance.

The Middleton Class-A model [8] has also been popularly adopted for characterizing channel noises by using impulsive characteristics, for which the PDF of *i.i.d.* additive noise samples $\{n_j = g_j + b_j\omega_j\}_{j=0}^{N-1}$ is described as [15]:

$$f_{\mathrm{M}}(x) = \sum_{\ell=0}^{\infty}\alpha_\ell\cdot\frac{1}{\sigma_\ell}\varphi\left(\frac{x}{\sigma_\ell}\right)\ , \tag{2}$$

where $\alpha_\ell = e^{-A}\frac{A^\ell}{\ell!}$ and $\sigma_\ell^2 = \frac{N_0}{2}(1+\frac{\ell}{\Lambda A})$. Notably, when specifying α_ℓ and σ_ℓ^2, the parameter, A is a constant typically referred to as the *impulsive index*, and Λ can be regarded as the mean power ratio of the AWGN component with respect to the impulsive noise component [15] because $\{\omega_j\}_{j=0}^{N-1}$ is an *i.i.d.* Gaussian random process with a mean of zero and a variance of $\left(\frac{N_0}{2}\right)/(\Lambda A)$. It is worth noting that b_j is a Poisson variable with mean A such that $\Pr(b_j = \ell) = \alpha_\ell$ for $\ell \geq 0$; hence, f_{M} exhibits the same form as does f_{BG} if $\alpha_0 = 1-p_b$, $\alpha_1 = p_b$, and $\alpha_\ell = 0$ for $\ell \geq 2$.

III. Metric Erasure Viterbi Algorithm (MEVA)

As previously mentioned, it may be difficult to obtain the precise PDF of impulsive noises. To address the adverse effect of decoding in the context of unknown impulse statistics, clipping was attempted on the decoding metrics; the clipping threshold is further investigated in Section IV. The evolution of erasure marking [18] to clipping is briefly described as follows.

By assuming the modulation format is Binary Phase-Shift Keying (BPSK),[3] the received symbol sequence $\boldsymbol{r} = (r_0, r_1, \ldots, r_{N-1})$ can be represented using the following equation:

$$\boldsymbol{r} = (-1)^{\boldsymbol{v}}\sqrt{E} + \boldsymbol{n}, \tag{3}$$

where $\boldsymbol{v} = (v_0, v_1, \ldots, v_{N-1}) \in \{0,1\}^N$ is the transmitted codeword and E is the energy of the modulated symbol. From the perspective of the decoder, the impulsive noise sequence $\boldsymbol{n} = (n_0, n_1, \ldots, n_{N-1})$ is only known to be *i.i.d.*, where n_j is an impulse plus AWGN with probability p, and only an AWGN with probability $(1-p)$. It should be stressed that the parameter p is likely different than the true probability of impulse occurrence p_b in the Bernoulli-Gaussian model [or $(1-\alpha_0)$ in the Middleton Class-A model)] because the impulsive noise statistics are unknown or extremely difficult to estimate at the receiver.

Furthermore, an *indicator*, e_j was defined for each received symbol to indicate whether this symbol was corrupted by an impulsive noise, where $e_j = 1$ is the arrival of an impulse, and $e_j = 0$ means that an impulse is absent. Referencing the notation e_j, the joint-maximum-likelihood decision of $(\hat{\boldsymbol{v}}, \hat{\boldsymbol{e}})$ was denoted for all the possible pairs $(\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{e}}) \in \boldsymbol{E}_N$, where $\boldsymbol{E}_N$ is the collection of all $(\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{e}})$ pairs, each of which corresponds to a codeword, $\tilde{\boldsymbol{v}}$, which is paired with an indicator sequence $\tilde{\boldsymbol{e}} = (\tilde{e}_0, \tilde{e}_1 \ldots, \tilde{e}_{N-1})$. This contrasts with the traditional Viterbi algorithm (VA), which finds the maximum likelihood codeword simply over the code book $\boldsymbol{C} = \{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_{2^{kL}-1}\}$. To clarify the notations in the rest of this paper, v_j is reserved to denote the code bit that was truly transmitted, $\tilde{v}_j$ denotes the code bit associated with the branch of the trellis currently decoded over, and $\hat{v}_j$ denotes the decoding decision. The same convention is applied for the notations of the indicator sequences.

Assuming an equal *a priori* probability of information sequences, the maximum *a posteriori* probability (MAP) decoding rule suggests that the optimal decision $(\hat{\boldsymbol{v}}, \hat{\boldsymbol{e}})$ should be given by

$$\begin{aligned} (\hat{\boldsymbol{v}}, \hat{\boldsymbol{e}}) &= \mathop{\arg\max}_{(\tilde{\boldsymbol{v}},\tilde{\boldsymbol{e}})\in\boldsymbol{E}_N} \ln f(\boldsymbol{r}|\tilde{\boldsymbol{v}},\tilde{\boldsymbol{e}})\Pr(\tilde{\boldsymbol{e}}) \\ &= \mathop{\arg\max}_{(\tilde{\boldsymbol{v}},\tilde{\boldsymbol{e}})\in\boldsymbol{E}_N} \sum_{j=0}^{N-1}\Big((1-\tilde{e}_j)\big[\ln f(r_j|\tilde{v}_j,\tilde{e}_j=0) + \ln\Pr(\tilde{e}_j=0)\big] \\ &\quad + \tilde{e}_j\big[\ln f(r_j|\tilde{v}_j,\tilde{e}_j=1) + \ln\Pr(\tilde{e}_j=1)\big]\Big), \end{aligned} \tag{4}$$

[3] The subsequent derivation applies directly to Quadrature Phase-Shift Keying (QPSK) with Gray mapping, in which the in-phase and quadrature components are equivalent to two BPSK symbols. In addition, it is simple to adapt this derivation for more general M-ary modulation schemes.

where f denotes a PDF. It is obvious that $f(r_j|\tilde{v}_j, \tilde{e}_j = 0) = \frac{1}{\sqrt{N_0/2}}\varphi\left(\frac{r_j-(-1)^{\tilde{v}_j}\sqrt{E}}{\sqrt{N_0/2}}\right)$ when the received sample, r_j is impulse-free. If $\tilde{e}_j = 1$, the contribution of r_j to the quantity $\ln f(\boldsymbol{r}|\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{e}}) \Pr(\tilde{\boldsymbol{e}})$ in (4) is $\ln f(r_j|\tilde{v}_j, \tilde{e}_j = 1) + \ln(p)$; without statistical data for the impulsive noise, using the same measure as that of the impulse-free case is detrimental because the decoding decision may be only dictated by the impulse-corrupted samples. From the perspective of a robust decoder design, a simple way to avert this problem is to neutralize the untrustworthy contribution by erasing the metric $\ln f(r_j|\tilde{v}_j, \tilde{e}_j = 1)$ in (4); thus, (4) is refined to (5).

The right side of (5) is proposed as the path metric of $(\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{e}})$, which is denoted by $M(\boldsymbol{r}|\tilde{\boldsymbol{v}}, \tilde{\boldsymbol{e}})$. Accordingly, the bit metric of $(\tilde{v}_j, \tilde{e}_j)$ is defined

$$M(r_j|\tilde{v}_j, \tilde{e}_j) = (1-\tilde{e}_j)\left(\frac{(r_j-(-1)^{\tilde{v}_j}\sqrt{E})^2}{N_0} - \ln\frac{(1-p)}{\sqrt{\pi N_0}}\right) - \tilde{e}_j\ln(p). \tag{6}$$

Note that the decoding scheme that uses (6) as its decoding metric was first named the *metric erasure Viterbi algorithm* (MEVA) in [25]. Depending whether the received sample, r_j is corrupted by an impulse, two probable bit metrics result from (6), and the smaller metric is retained for updating the path metric. Specifically, at time j, if the following inequality holds:

$$\underbrace{\frac{(r_j-(-1)^{\tilde{v}_j}\sqrt{E})^2}{N_0} - \ln\frac{(1-p)}{\sqrt{\pi N_0}}}_{M(r_j|\tilde{v}_j,\tilde{e}_j=0)} > \underbrace{-\ln(p)}_{M(r_j|\tilde{v}_j,\tilde{e}_j=1)}, \tag{7}$$

the MEVA deems the associated sample as impulse-corrupted. In this regard, the decoding behavior of the MEVA is equivalent to performing clipping on the Euclidean metric $(r_j-(-1)^{\tilde{v}_j}\sqrt{E})^2$, where the clipping threshold is measured using

$$T = N_0\ln\left(\frac{1-p}{p\sqrt{\pi N_0}}\right), \tag{8}$$

and (7) can be written using:

$$(r_j-(-1)^{\tilde{v}_j}\sqrt{E})^2 > T. \tag{9}$$

Based on (9), the path metric can be transformed into:

$$\begin{aligned}&\sum_{j=0}^{N-1} M(r_j|\tilde{v}_j, \tilde{e}_j)\\ &= \sum_{j=0}^{N-1}\min\{M(r_j|\tilde{v}_j, \tilde{e}_j=0), M(r_j|\tilde{v}_j, \tilde{e}_j=1)\}\\ &= \frac{1}{N_0}\sum_{j=0}^{N-1}\min\left\{(r_j-(-1)^{\tilde{v}_j}\sqrt{E})^2, T\right\} - N\ln\frac{(1-p)}{\sqrt{\pi N_0}}.\end{aligned} \tag{10}$$

By removing the last inconsequential term for decision-making from (10) and multiplying the resulting quantity by N_0, the implementation of the MEVA is equivalent to performing the Viterbi algorithm on the original (i.e., non-erasure-expanded) trellis of the CC with the clipped Euclidean bit metric

$$\tilde{M}(r_j|\tilde{v}_j) \triangleq \min\left\{(r_j-(-1)^{\tilde{v}_j}\sqrt{E})^2, T\right\}. \tag{11}$$

By abusing the notation, the corresponding path metric can be further denoted using:

$$\tilde{M}(\boldsymbol{r}|\tilde{\boldsymbol{v}}) = \sum_{j=0}^{N-1}\tilde{M}(r_j|\tilde{v}_j). \tag{12}$$

It should be mentioned that when p is larger than $\frac{1}{1+\sqrt{\pi N_0}}$, a negative clipping threshold $(T < 0)$ results. In this case, all the code bits are erased and the MEVA can only output a random guess on the transmitted codeword; the probability of a correct decoding equals to 2^{-kL}.

By contrast to applying the VA on the erasure-expanded trellis in [25], which increases the decoding complexity by 2^n times relative to the VA on the original trellis, the clipping mechanism, (12) applied to implement the MEVA, only induces an additional clipping operation (11) for each brach metric computation; hence the decoding complexity is increased by a constant multiplicative factor (< 2).

It must be emphasized that $f(r_j|\tilde{v}_j, \tilde{e}_j = 1)p$ is a good approximation to p when p is fairly small. Moreover, when impulsive noise statistics are unavailable, the parameter, p assumed at the decoder likely differs from the true probability of impulse occurrence, p_b. In particular, after erasing the untrustworthy metric $\ln f(r_j|\tilde{v}_j, \tilde{e}_j = 1)$ in (4), this parameter becomes an offset provider $-\tilde{e}_j\ln(p)$ in the bit metric in (6). Hence, it transforms into a *design parameter* rather than an estimate for the probability of impulse occurrence, p_b. The suggestion for selecting a suitable p for the MEVA is presented in Subsection V-A.

IV. Performance Bounds for Given Statistics of Impulsive Noises

The concept of clipping was inherently executed in the MEVA in Section III; now the PEP bound for the MEVA can be derived by applying the Chernoff bound technique for the two statistical models of impulsive noises in Section II.

For ease of understanding, only the analysis of the Bernoulli-Gaussian noise model is given. A simple analogous bound for the Middleton Class-A model is provided, and no detailed derivation is presented. These bounds are then used to select a suitable value for p when it is treated as a design parameter. Subsection V-A presents how to select p according to these bounds.

In this analysis, only the (n, k, m) CCs with $k = 1$ were focused. The derivation can be similarly extended to CCs where $k > 1$. For length-L information sequences, an $(n, 1, m)$ CC can be regarded as an (N, L) linear block code of length $N = n(L+m)$. Without losing generality, because of the linearity of the code and the memoryless property of the additive noise, the all-zero codeword, $\boldsymbol{v}_0$ can be assumed to be transmitted when the decoding error is calculated.

Based on the system setting in the previous paragraph, a well-known BEP bound P_B can be calculated using the following equation:

$$P_B \le \frac{1}{k}\sum_{w=d_{\min}}^{d_{\max}} B_w\cdot\Pr\left(\tilde{M}(\boldsymbol{r}|\boldsymbol{v}^{(w)}) \le \tilde{M}(\boldsymbol{r}|\boldsymbol{v}_0)\Big|\,\mathcal{A}_{\boldsymbol{0}}\right), \tag{13}$$

where $d_{\min}$ and $d_{\max}$ are the minimum and maximum pairwise Hamming distances of the (N, L) block code, $\boldsymbol{v}^{(w)}$ is any

$$\begin{aligned}
&(\hat{\boldsymbol{v}}, \hat{\boldsymbol{e}}) \\
=\ & \underset{(\tilde{\boldsymbol{v}},\tilde{\boldsymbol{e}})\in \boldsymbol{E}_N}{\arg\max} \sum_{j=0}^{N-1} \Big((1-\tilde{e}_j) \ln f(r_j|\tilde{v}_j, \tilde{e}_j = 0) + (1-\tilde{e}_j)\ln(1-p) + \tilde{e}_j \ln(p) \Big) \\
=\ & \underset{(\tilde{\boldsymbol{v}},\tilde{\boldsymbol{e}})\in \boldsymbol{E}_N}{\arg\min} \sum_{j=0}^{N-1} \left[(1-\tilde{e}_j) \left(\frac{(r_j - (-1)^{\tilde{v}_j}\sqrt{E})^2}{N_0} - \ln \frac{(1-p)}{\sqrt{\pi N_0}} \right) - \tilde{e}_j \ln(p) \right]. \qquad (5)
\end{aligned}$$

codeword that possesses Hamming weight w, B_w is the total number of nonzero information bits in all Hamming-weight-w codewords, and $\mathcal{A}_\mathbf{0}$ is the event for which the all-zero codeword is chosen for transmission. The remaining task is to derive $\Pr\left(\tilde{M}(\boldsymbol{r}|\hat{\boldsymbol{v}}) \leq \tilde{M}(\boldsymbol{r}|\boldsymbol{v}_0) \middle| \mathcal{A}_\mathbf{0} \right)$ as a function of the Hamming weight of $\hat{\boldsymbol{v}}$.

Continuing the derivation of the PEP in (13):

$$\begin{aligned}
&\Pr\left(\tilde{M}(\boldsymbol{r}|\hat{\boldsymbol{v}}) \leq \tilde{M}(\boldsymbol{r}|\boldsymbol{v}_0) \middle| \mathcal{A}_\mathbf{0} \right) \\
=\ & \Pr\left(\sum_{j\,:\,\hat{v}_j=1} \phi_j \leq 0 \middle| \mathcal{A}_\mathbf{0} \right) \\
\leq\ & \min_{s\,:\,s>0} E\left[e^{-s\sum_{j\,:\,\hat{v}_j=1}\phi_j} \middle| \mathcal{A}_0 \right], \qquad (14)
\end{aligned}$$

where for those j's in which $\hat{v}_j = 1$, the following formula is used:

$$\begin{aligned}
\phi_j &= \min\left((r_j+\sqrt{E})^2, T \right) - \min\left((r_j-\sqrt{E})^2, T \right) \\
&= \begin{cases} 4\sqrt{E}r_j, & (r_j+\sqrt{E})^2 < T \wedge (r_j-\sqrt{E})^2 < T; \\ T-(r_j-\sqrt{E})^2, & (r_j+\sqrt{E})^2 \geq T \wedge (r_j-\sqrt{E})^2 < T; \\ (r_j+\sqrt{E})^2 - T, & (r_j+\sqrt{E})^2 < T \wedge (r_j-\sqrt{E})^2 \geq T; \\ 0, & (r_j+\sqrt{E})^2 \geq T \wedge (r_j-\sqrt{E})^2 \geq T. \end{cases}
\end{aligned}$$

Depending on the gap between the clipping threshold, T and the symbol energy, E, two cases can be distinguished: $E \geq T$ and $E < T$, in which it is assumed that $T > 0$ and $t = \sqrt{T}$.

If $E \geq T$,

$$\phi_j = \begin{cases} T-(r_j-\sqrt{E})^2, & \sqrt{E}-t < r_j < \sqrt{E}+t \\ (r_j+\sqrt{E})^2 - T, & -\sqrt{E}-t < r_j < -\sqrt{E}+t \\ 0, & \text{otherwise.} \end{cases}$$

Else if $E < T$, then

$$\phi_j = \begin{cases} 4\sqrt{E}r_j, & \sqrt{E}-t < r_j < -\sqrt{E}+t \\ T-(r_j-\sqrt{E})^2, & -\sqrt{E}+t \leq r_j < \sqrt{E}+t \\ (r_j+\sqrt{E})^2 - T, & -\sqrt{E}-t \leq r_j \leq \sqrt{E}-t \\ 0, & \text{otherwise.} \end{cases}$$

Note that $r_j = \sqrt{E} + n_j$ when $\boldsymbol{v}_\mathbf{0}$ is transmitted; hence, the preceding equations can be written as: If $E \geq T$,

$$\phi_j = \begin{cases} T-n_j^2, & -t < n_j < t \\ (2\sqrt{E}+n_j)^2 - T, & -t-2\sqrt{E} < n_j < t-2\sqrt{E} \\ 0, & \text{otherwise,} \end{cases}$$

else,

$$\phi_j = \begin{cases} 4\sqrt{E}(\sqrt{E}+n_j), & -t < n_j < t-2\sqrt{E} \\ T-n_j^2, & t-2\sqrt{E} \leq n_j < t \\ (2\sqrt{E}+n_j)^2 - T, & -t-2\sqrt{E} \leq n_j \leq -t \\ 0, & \text{otherwise,} \end{cases}$$

where by following (1), the PDF of n_j is equal to:

$$\begin{aligned}
f_{n_j}(x) &= (1-p_b)\frac{1}{\sqrt{N_0/2}}\varphi\left(\frac{x}{\sqrt{N_0/2}} \right) \\
&+ p_b \frac{1}{\sqrt{N_0/2(1+\Gamma)}}\varphi\left(\frac{x}{\sqrt{(N_0/2)(1+\Gamma)}} \right).
\end{aligned}$$

Thus, when $E \geq T$, $\sigma_0 = \sqrt{N_0/2}$, $\sigma_1 = \sqrt{(N_0/2)(1+\Gamma)}$, and $\mu = \sqrt{E}$, the following equation is yielded:

$$\begin{aligned}
&E\left[e^{-s\phi_j} \middle| \mathcal{A}_\mathbf{0} \right] \\
&= \int_{-t}^{t} e^{-s(T-x^2)} f_{n_j}(x)dx + \int_{-t-2\mu}^{t-2\mu} e^{-s((2\mu+x)^2-T)} f_{n_j}(x)dx \\
&+ \left(1 - \int_{-t}^{t} f_{n_j}(x)dx - \int_{-t-2\mu}^{t-2\mu} f_{n_j}(x)dx \right) \\
&= e^{-sT}\left[(1-p_b)A_1(s,t,\sigma_0) + p_b A_1(s,t,\sigma_1) \right] \\
&+ e^{sT}\left[(1-p_b)B_1(\mu,s,t,\sigma_0) + p_b B_1(\mu,s,t,\sigma_1) \right] \\
&+ \left[(1-p_b)C_1(t,\sigma_0,\mu) + p_b C_1(t,\sigma_1,\mu) \right], \qquad (15)
\end{aligned}$$

where $Q(t) = \int_t^\infty \varphi(x)dx$ is the tail probability of the standard normal distribution, and the detailed derivations and the definitions of functions A_1, B_1, and C_1 are placed in the Appendix for improved readability. When $T > E$, the following equation is yielded:

$$\begin{aligned}
&E\left[e^{-s\phi_j} \middle| \mathcal{A}_\mathbf{0} \right] \\
&= \int_{-t}^{t-2\mu} e^{-4s\mu(\mu+x)} f_{n_j}(x)dx + \int_{t-2\mu}^{t} e^{-s(T-x^2)} f_{n_j}(x)dx \\
&+ \int_{-t-2\mu}^{-t} e^{-s((2\mu+x)^2-T)} f_{n_j}(x)dx + \left(1 - \int_{-t-2\mu}^{t} f_{n_j}(x)dx \right) \\
&= e^{-4s\mu^2}\left[(1-p_b)D(s,t,\sigma_0,\mu) + p_b D(s,t,\sigma_1,\mu) \right] \\
&+ e^{-sT}\left[(1-p_b)A_2(s,t,\sigma_0,\mu) + p_b A_2(s,t,\sigma_1,\mu) \right] \\
&+ e^{sT}\left[(1-p_b)B_2(s,t,\sigma_0,\mu) + p_b B_2(s,t,\sigma_1,\mu) \right] \\
&+ \left[(1-p_b)C_2(t,\sigma_0,\mu) + p_b C_2(t,\sigma_1,\mu) \right], \qquad (16)
\end{aligned}$$

where the detailed derivations and the definitions of functions A_2, B_2, C_2, and D are located in the Appendix.

Similarly, this derivation can be extended to the Middleton Class-A model, in which (15) and (16), respectively, become

$$\begin{aligned}
&E\left[e^{-s\phi_j} \middle| \mathcal{A}_\mathbf{0} \right] \\
=\ & e^{-sT}\sum_{\ell=0}^{\infty}\alpha_\ell A_1(s,t,\sigma_\ell) + e^{sT}\sum_{\ell=0}^{\infty}\alpha_\ell B_1(\mu,s,t,\sigma_\ell) \\
&+ \sum_{\ell=0}^{\infty}\alpha_\ell C_1(t,\sigma_\ell,\mu)
\end{aligned}$$

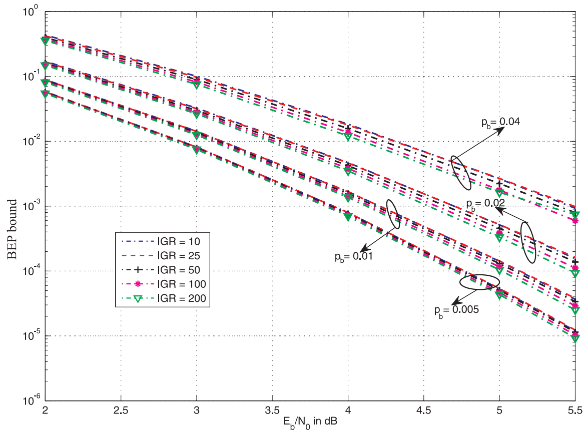


Fig. 1. Bit error probability bounds for the MEVA over Bernoulli-Gaussian channels.

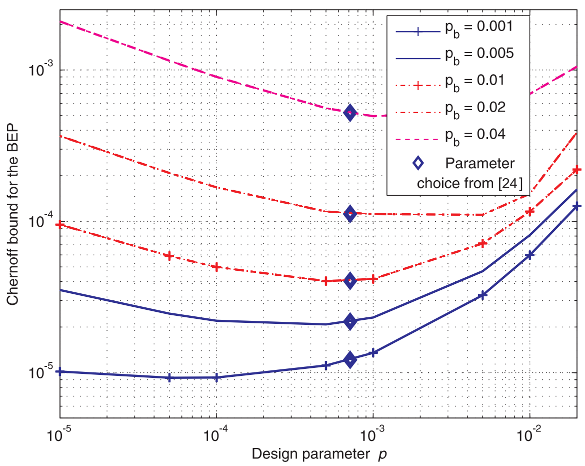


Fig. 2. Bit error probability bounds versus the design parameter p for the MEVA over Bernoulli-Gaussian channels with $\Gamma = 100$ at $E_b/N_0 = 5$ dB. Here the diamond markers correspond to the choice of clipping threshold suggested in [23].

and

$$
\begin{aligned}
&E\left[e^{-s\phi_j}\middle|\mathcal{A}_0\right]\\
&= e^{-4s\mu^2}\sum_{\ell=0}^{\infty}\alpha_\ell D(s,t,\sigma_\ell,\mu)+e^{-sT}\sum_{\ell=0}^{\infty}\alpha_\ell A_2(s,t,\sigma_\ell,\mu)\\
&\quad +e^{sT}\sum_{\ell=0}^{\infty}\alpha_\ell B_2(s,t,\sigma_\ell,\mu)+\sum_{\ell=0}^{\infty}\alpha_\ell C_2(t,\sigma_\ell,\mu).
\end{aligned}
$$

After the expression of $E\left[e^{-s\phi_j}\middle|\mathcal{A}_0\right]$ is obtained, the desired PEP bound follows directly after substituting it into (14).

V. Simulation Results

The proposed MEVA was experimented by conducting computer simulations to configure the transmission of the $(2,1,6)$ CC with generators 147,135 (octal) over Bernoulli-Gaussian and Middleton Class-A noise channel models. The number of information bits per data frame for the CC was fixed at $L = 500$. To simplify the exposition, the BPSK modulation was exclusively used in the following simulations.

A. The Chernoff bounds on the BEP of the MEVA

To attest our derived Chernoff bound on the BEP for the MEVA, it was first measured for a wide range of probability of

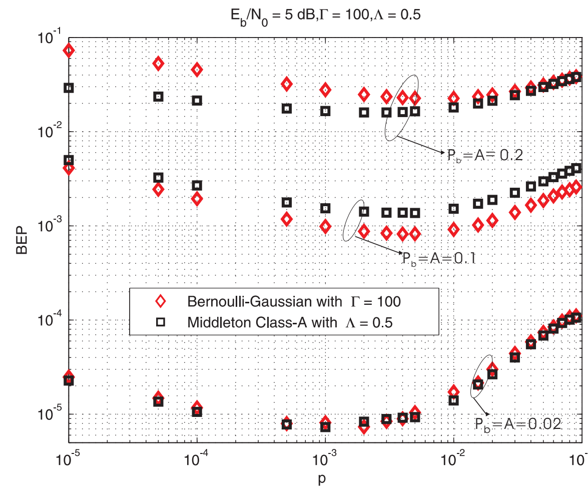
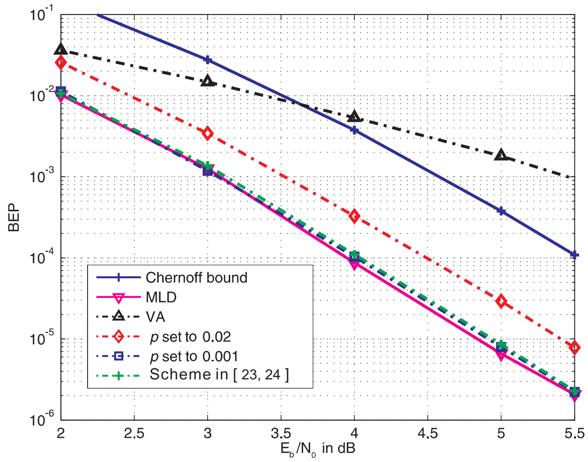


Fig. 3. Bit error probabilities versus the design parameter p with respect to various p_b and A at $E_b/N_0 = 5$ dB.

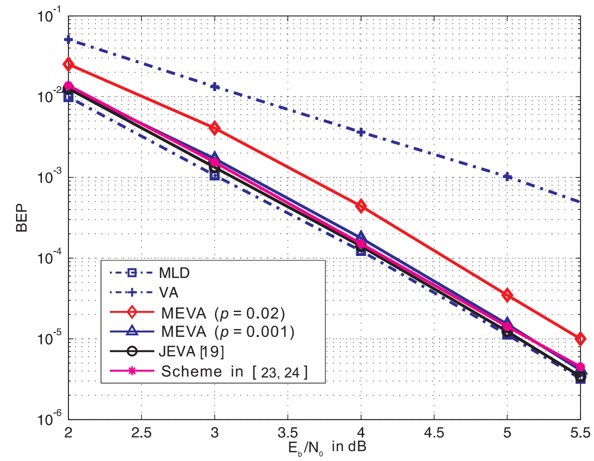
impulse occurrence p_b, and the results were plotted for various IGRs (e.g. parameter Γ) in Fig. 1. The design parameter p was set to equal p_b in this figure; although the value of its induced BEP bound is unfavorable, a favored p choice is later shown in Fig. 2. Fig. 1 shows that the curves of the BEP bounds are fairly close for each illustrated probability, p_b, indicating that the MEVA performs robustly against impulses that are unknown to the receiver. Notably, the value of the BEP performance bound is poor when the probability, p_b is large, indirectly reflecting that the effectiveness of the MEVA in couple with the CC is likely weakened when impulses occur extremely frequently.

Subsequently, to measure how the choice of p affects the performance of the MEVA, the derived Chernoff bound is employed to expedite the assessment prior to generating laborious computer simulations in the forthcoming subsections. Here, the values of p_b relevant to the target application scenario are assumed to vary between 0.005 and 0.04. Specifically, the Chernoff bound was measured for the BEP of the MEVA at $E_b/N_0 = 5$ dB; the system was set to that design parameter p, which is used to determine the clipping threshold, is likely deviated from the unknown p_b. The entailing result is shown in Fig. 2 in which the IGR was fixed at $\Gamma = 100$. Remarkably, regardless of the probability of impulse occurrence, p_b, the incurred BEP bounds were all U-shaped; thus, certain p ranges, such as $p \geq 10^{-2}$, were avoided. By contrast, it is possible to set a fixed clipping threshold T [see (8)] derived from a properly chosen parameter, p, which is independent of the probability, p_b. This allows the proposed decoding scheme to be used without requiring a sophisticated estimation algorithm for p_b and, at the same time, the resultant coding gain is likely to be least compromised. In this view, parameter $p = 0.001$ was selected for the following simulations, unless otherwise specified.

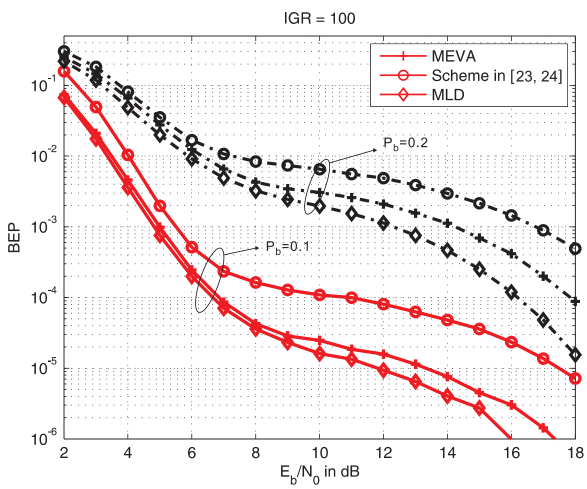
As previously mentioned, the clipping threshold in [23] was empirically determined and converted to conform with the clipping threshold formula in (8), which is marked in Fig. 2 (see the diamond markers in each curve). Clearly, the value associated with the diamonds is fairly close to that of the chosen $p = 0.001$. Thus, the simulation results should be similar to those of [23].



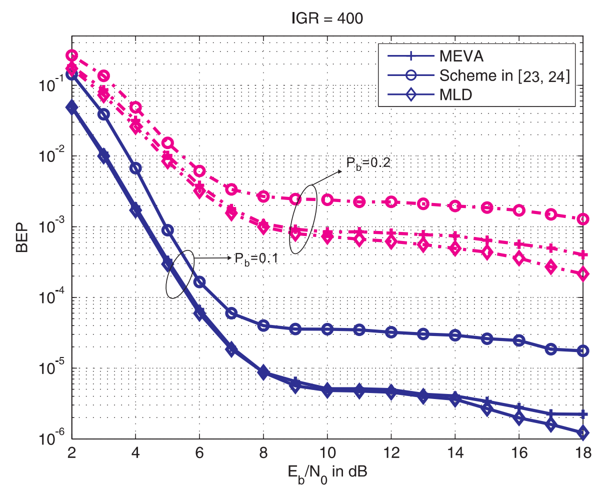
(a) Middleton Class-A channel with $A = 0.02$ and $\Lambda = 1/2$



(b) Bernoulli-Gaussian channel with $p_b = 0.02$ and $\Gamma = 15$



(c) Bernoulli-Gaussian channel with $\Gamma = 100$



(d) Bernoulli-Gaussian channel with $\Gamma = 400$

Fig. 4. Comparisons of bit error probabilities for various schemes.

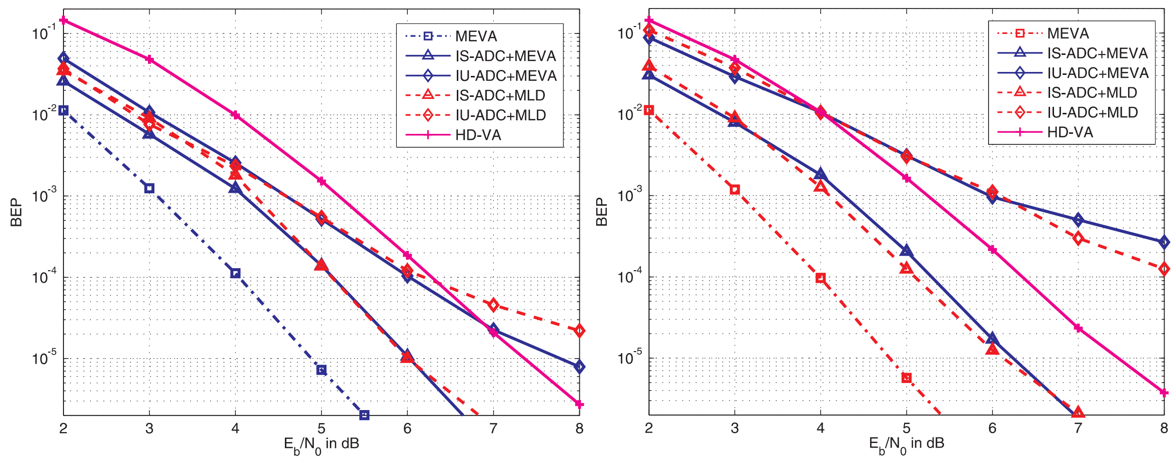
B. The impact of the design parameter p on the simulated BEPs

To demonstrate the effectiveness of the derived analytical bound in interpreting the behavioral performance of the MEVA, the BEPs were investigated at $E_b/N_0 = 5$ dB at a wide range of p in cases of $p_b = A = 0.02,\ 0.1,\ 0.2$ for Bernoulli-Gaussian and Middleton Class-A models. Fig. 3 shows that the BEP increases in the regime of exaggerated p (such as $p \geq p_b$). A similar conclusion was reached in [22]. As the value of p declines from $p = p_b$, the BEP first improves but later increases when an excessively small p is reached. This U-shape trend was demonstrated in the analytical results (Fig. 2). Particularly for lower p_b values, such as 0.02, a more substantial performance change is demonstrated in the presented range of design parameter p; thus, it is useful to derive the Chernoff bound for the BEP of a clipping-featured decoder to arrive at a suitable clipping threshold, averting underperformance regardless of the impulsive noise model.

C. Comparing the MEVA and the works of [22], [23], and [18]

To compare the MEVA with the existing works such as [22], [23], and [18], computer simulations were conducted, using the Bernoulli-Gaussian and Middleton Class-A noise models; the results are summarized in Fig. 4. The BEP curve that corresponds to the ideal situation, in which the receiver possesses perfect statistical knowledge of the noise model, served as the benchmark for the level of performance in these subfigures, and was labeled the MLD. The curve labeled VA shows the performance of using the conventional (non-clipped) Euclidean bit metric $(r_j - (-1)^{\tilde{v}_j}\sqrt{E})^2$; thus, the impulsive effect is completely neglected at the receiver.

Notably, in Fig. 4(a), the BEP curve derived from the MEVA, in which the clipping threshold is dictated by a fixed value of $p = 0.001$ (see the dash-dot line marked with a "□") almost overlaps the curve induced by the threshold in [22], [23] (see the dash-dot line marked with a "+"), and that of the MLD (see the solid line marked with a "▽"), indicating that both clipping methods demonstrate a robust performance against impulses and are as effective as the benchmark MLD. Moreover, for the MEVA, a 0.5 dB gain occurs by setting $p = 0.001$ when compared with p set to A, which is 0.02 (see the dash-dot line marked with a "◇"). Notably, the Chernoff bound for the BEP of the MEVA (see the solid line marked with a "+") exhibits similar trends to the simulated BEPs, indicating the efficacy of the derived bound.



(a) IGR $= 100$ (b) IGR $= 400$

Fig. 5. Bit error probabilities of the MEVA decoders with a preceding impulse-saturated (IS) and impulse-unsaturated (IU) ADC with $p = 0.001$ over the Bernoulli-Gaussian channel with $p_b = 0.02$ and two different IGRs.

Observations similar to those in Fig. 4(a) can be drawn from Fig. 4(b), in which the MEVA was compared with the iterative JEVA in [18]. Note that in Fig. 4(b), where $p_b = 0.02$ and $\Gamma = 15$ in the Bernoulli-Gaussian noise channel, the JEVA performs comparably to the MEVA and the scheme in [22] and [23], whereas a multi-fold complexity increase results from iterations, preventing the JEVA from being applying in practical systems.

When the probability of impulse occurrence p_b is extremely high, such as 0.1 or 0.2 in Figs. 4(c) and 4(d), the MEVA and the MLD perform similarly when the E_b/N_0 is low. However, the gap between the BEP curve induced by the MEVA and that of the MLD becomes visible at a high E_b/N_0. These subfigures show that this gap is further enlarged when the p_b increases from 0.1 to 0.2, especially when the IGR (i.e., Γ) is small. This is primary because in the context of a large p_b, a high E_b/N_0 value is inevitable for the BEP curve to decline; however, the model mismatch between the assumed $f(r_j|v_j, e_j = 1) \approx 1$ (cf., the paragraph immediately after Eq. (5)) and the conditional PDF of the impulsive noise $f_{\text{BG}}(r_j - (-1)^{v_j}\sqrt{E}|b_j = 1)$ exacerbates when the IGR is small. Despite this, the MEVA based on the induced clipping threshold can outperform the schemes proposed in [22] and [23] when the E_b/N_0 values are high.

D. Examining the impact of the ADC at the front end

The saturation effect of impulsive noises on the MEVA preceded by an ADC over the Bernoulli-Gaussian noise model was investigated. In this experiment, it was assumed that the ADC that preceded the MEVA decoder used 8-level quantization and uniform quantization spacing [26]. An impulse-unsaturated (IU) ADC, namely, an IU-ADC + MEVA, was employed in which the dynamic range was determined by the magnitude of the signal plus the impulsive background noise. The case corresponding to the typical 8-level uniform quantization [26] was referred to as "IS-ADC + MEVA," in which the dynamic range of the impulse-saturated (IS) ADC was the signal level plus the background noise level, excluding the impulsive noise level. Fig. 5 shows the results, based on which, four observations are made.

First, without a preceding ADC, the dash-dot lines marked with squares in Figs. 5(a) and (b) are nearly overlapped, reinforcing the immunity of the MEVA against strong impulses, where the IGR $= 100$ and 400 in Figs. 5(a) and (b), respectively. Second, regardless of the value of the IGR, adding an IS-ADC to the MEVA front end induces roughly 1 dB performance loss at BEP $= 10^{-5}$ because of the unavoidable quantization noise; nevertheless, the IS-ADC + MEVA decoder approaches the benchmark performance of the IS-ADC + MLD decoder. Third, because of the excess quantization noise that scales with the IGR, a much more pronounced performance loss occurs when an IU-ADC is used. For example, the IU-ADC + MEVA BEP curves that correspond to IGR $= 400$ (see the solid line marked with diamonds) do not decrease to 10^{-3} until $E_b/N_0 = 6$ dB; similarly, the excess quantization noise causes a serious performance degradation for the MLD (see the curves that correspond to IU-ADC + MLD). Fourth, the coding gain realized by the IS-ADC + MEVA decoder when compared with the hard-decision Viterbi algorithm, namely the HD-VA (Fig. 5) is between 1 dB and 1.5 dB, again confirming the superiority of the MEVA, even with a preceding ADC.

VI. Conclusion

A robust decoding scheme is indispensable for a convolutionally coded data stream that is affected by impulsive noises because achieving the maximum likelihood decoding performance relies heavily on an accurate estimate of the impulsive noise statistics; the time-varying nature taxes the complexity of the receiver. By incorporating design parameter, p, which emulates the unknown probability of impulse occurrence in the joint erasure marking and Viterbi decoding algorithm, the *metric erasure Viterbi algorithm* (MEVA) was developed; it can be applied despite a lack of detailed statistical knowledge of impulsive noises. Capitalizing on the memoryless attribute of impulsive noises (e.g., resulting from ideal interleaving), the MEVA can be implemented using the conventional VA and a well-designed clipping operation for the Euclidean metric, in which the clipping threshold is set according to the chosen design parameter, p.

To determine a feasible clipping threshold, the Chernoff bound was used to calculate the bit error probability (BEP) of the MEVA over the Bernoulli-Gaussian and Middleton Class-A impulsive noise models. Remarkably, by examining the analytical error bounds for a practical range of noise model parameters, choosing a fixed, robust value for design parameter, p is likely for a satisfactory BEP performance. Specifically, when the mean power ratio between impulses and background noises is large and the probability of impulse occurrence is moderate, the simulation results show that the MEVA that employs a fixed, robust choice of p is not only commensurate with the traditional VA in terms of the order of complexity, but equivalent to the ideal MLD in terms of the level of BEP performance. Because of its efficiency in decoding and robust performance, the MEVA provides an attractive solution for practical applications in impulsive noise channels such as power line communications. Notably, the BEP simulation results demonstrate the efficacy of the fixed robust choice of p, which is induced using the Chernoff bound. This justifies using the Chernoff bound on the BEP as an expeditious tool for analyzing the performance of the MEVA.

Finally, based on extensive simulations (not shown), more powerful channel codes compared with the convolutional codes should be employed if the probability of impulse occurrence is fairly high. For instance, the LDPC code [23] has been proposed when this probability is 0.1. Therefore, future studies could adapt the proposed MEVA approach to sophisticated channel codes and verify its effectiveness in harsh scenarios.

References

[1] J. Haring and A. J. H. Vinck, "Coding for impulsive noise channels," in *Proc. 2001 IEEE ISPLC*, pp. 103–108.

[2] J. Modestino, D. H. Sargard, and R. E. Bollen, "Use of coding to combat impulse noise on digital subscriber loops," *IEEE Trans. Commun.*, pp. 529–537, May 1988.

[3] K. L. Blackard and T. S. Rappaport, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Sel. Areas Commun.*, pp. 991–1001, Sept. 1993.

[4] D. Middleton, "Statistical-physical models of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, pp. 106–127, Aug. 1977.

[5] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Trans. Commun.*, pp. 145–147, Feb 1996.

[6] A. D. Spaulding and D. Middleton, "Optimum reception in the impulsive interference environment—part I: coherent detection," *IEEE Trans. Commun.*, pp. 910–923, Sept. 1977.

[7] M. K. S. Miyamoto and N. Morinaga, "Performance analysis of QAM systems under class A impulse noise environment," *IEEE Trans. Electromagn. Compat.*, pp. 260–267, May 1995.

[8] D. Middleton, "Canonical and quasi-canonical probability models of class A interference," *IEEE Trans. Electromagn. Compat.*, pp. 76–106, May 1983.

[9] J. Haring and A. J. H. Vinck, "Performance bounds for optimum and suboptimum reception under class A impulsive noise," *IEEE Trans. Commun.*, pp. 1130–1136, July 2002.

[10] R. Pighi, M. Franceschini, G. Ferrari, and R. Raheli, "Fundamental performance limits of communications systems impaired by impulse noise," *IEEE Trans. Commun.*, pp. 171–182, Jan 2009.

[11] S. Galli, A. Scaglione, and Z. Wang, "For the grid and through the grid: the role of power line communications in the smart grid," *Proc. IEEE*, pp. 1–26, June 2011.

[12] D. H. Sargrad and J. W. Modestino, "Error-and-erasures coding to combat impulse noise on digital subscriber loops," *IEEE Trans. Commun.*, pp. 1145–1155, Aug. 1990.

[13] D. Umehara, H. Yamaguchi, and Y. Morihiro, "Turbo decoding in impulsive noise environment," in *Proc. 2004 IEEE Global Telecom. Conf.*, pp. 194–198.

[14] H. Nakagawa, D. Umehara, S. Denno, and Y. Morihiro, "A decoding for low density parity check codes over impulsive noise channels," in *Proc. 2005 IEEE ISPLC*, pp. 85–89.

[15] H. C. Ferreira, L. Lampe, J. Newbury, and T. G. Swart, *Power Line Communications: Theory and Applications for Narrowband and Broadband Communications over Power Lines*. John Wiley and Sons, 2010.

[16] D. Fertonani and G. Colavolpe, "Theoretical limits and practical detection schemes for channels affected by Class-A impulse noise," in *Proc. 2007 IEEE Global Telecom. Conf.*, pp. 146–150.

[17] M. Ardakani, F. R. Kschischang, and W. Yu, "Near-capacity coding for discrete multitone systems with impulse noise," *EURASIP J. Applied Signal Process.*, pp. 1–10, vol. 2006, article ID 98738.

[18] T. Li, W. H. Mow, and M. Siu, "Joint erasure marking and Viterbi decoding algorithm for unknown impulsive noise channels," *IEEE Trans. Wireless Commun.*, pp. 3407–3416, Sept. 2008.

[19] P. Amirshahi, S. M. Navidpour, and M. Kavehrad, "Performance analysis of uncoded and coded OFDM broadband transmission over low voltage power-line channels with impulsive noise," *IEEE Trans. Power Delivery*, pp. 1927–1934, Oct 2006.

[20] T. C. Chuah, "Robust iterative decoding of turbo codes in heavy-tailed noise," *IEE-Proc.*, pp. 20–38, Feb. 2005.

[21] J. Mitra and L. Lampe, "Robust decoding for channels with impulse noise," in *Proc. 2006 IEEE Global Telecom. Conf.*, pp. 1–6.

[22] D. Fertonani and G. Colavolpe, "A simplified metric for soft-output detection in the presence of impulse noise," in *Proc. 2007 IEEE ISPLC*, pp. 121–126.

[23] ——, "A robust metric for soft-output detection in the presence of class-A noise," *IEEE Trans. Commun.*, pp. 36–40, Jan 2009.

[24] D. Toumpakaris, W. Yu, J. Cioffi, D. Gardan, and M. Ouzzif, "A byte-erasure method for improved impulse immunity in DSL systems using soft information from an inner code," in *Proc. 2003 IEEE Int. Conf. on Commun.*, pp. 2431–2435.

[25] D.-F. Tseng, Y. S. Han, W. H. Mow, and J. Deng, "Efficient decoding over power-line channels," in *Proc. 2011 Int'l Workshop on Signal Design and its Applications in Communications*, pp. 138–141.

[26] J. A. Heller and I. M. Jacobs, "Viterbi decoding for satellite and space communication," *IEEE Trans. Commun. Technol.*, pp. 835–848, Oct. 1971.

Der-Feng Tseng received his B.S.E.E. degree from National Chiao Tung University, and M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, Indiana. He was with TRW-ESL, Sunnyvale, CA, and Motorola Mobility, Pistcataway, NJ, as Senior Systems Engineer engaged in embedded software development for mobile devices before joining National Taiwan University of Science and Technology, where he has been with since February 2003. He was a visiting research scholar in the Department of Electrical Engineering at the University of Houston from June to September 2013. He was the co-recipient of "The Fred Ellersick MILCOM Award for Best Paper in the Unclassified Technical Program" at the IEEE MILCOM'98, Boston, MA. His research interests lie in the areas of communication systems, statistical signal processing, and channel coding techniques.

Yunghsiang S. Han (S'90–M'93–SM'08–F'11) was born in Taipei, Taiwan, 1962. He received B.Sc. and M.Sc. degrees in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1984 and 1986, respectively, and a Ph.D. degree from the School of Computer and Information Science, Syracuse University, Syracuse, NY, in 1993. He was from 1986 to 1988 a lecturer at Ming-Hsin Engineering College, Hsinchu, Taiwan. He was a teaching assistant from 1989 to 1992, and a research associate in the School of Computer and Information Science, Syracuse University from 1992 to 1993. He was, from 1993 to 1997, an Associate Professor in the Department of Electronic Engineering at Hua Fan College of Humanities and Technology, Taipei Hsien, Taiwan. He was with the Department of Computer Science and Information Engineering at National Chi Nan University, Nantou, Taiwan from 1997 to 2004. He was promoted to Professor in 1998. He was a visiting scholar in the Department of Electrical Engineering at University of Hawaii at Manoa, HI from June to October 2001, the SUPRIA visiting research scholar in the Department of Electrical Engineering and Computer Science and CASE center at Syracuse University, NY from September 2002 to January 2004 and July 2012 to June 2013, and the visiting scholar in the Department of Electrical and Computer Engineering at University of Texas at Austin, TX from August 2008 to June 2009.

He was with the Graduate Institute of Communication Engineering at National Taipei University, Taipei, Taiwan from August 2004 to July 2010. From August 2010, he is with the Department of Electrical Engineering at National Taiwan University of Science and Technology as Chair professor. His research interests are in error-control coding, wireless networks, and security. Dr. Han was a winner of the 1994 Syracuse University Doctoral Prize and a Fellow of IEEE.

Wai Ho Mow (S'89-M'93-SM'99) received his M.Phil. and Ph.D. degrees in information engineering from the Chinese University of Hong Kong in 1991 and 1993, respectively. From 1997 to 1999, he was with the Nanyang Technological University, Singapore. He has been with the Hong Kong University of Science and Technology (HKUST) since March 2000. He was the recipient of seven research/exchange fellowships from ve countries, including the Humboldt Research Fellowship. His research interests are in the areas of communications, coding, and information theory. He pioneered the lattice approach to signal detection problems (such as sphere decoding and complex lattice reduction-aided detection) and unied all known constructions of perfect roots-of-unity (or CAZAC) sequences (widely used as preambles and sounding sequences).

He has published one book, and has coauthored over 30 filed patent applications and over 150 technical publications, among which he is the sole author of over 40. He coauthored two papers that received the ISITA2002 Paper Award for Young Researchers and the APCC2013 Best Paper Award. Since 2002, he has been the principal investigator of over 16 funded research projects. In 2005, he chaired the Hong Kong Chapter of the IEEE Information Theory Society. He was the Technical Program Co-Chair of five conferences, and served the technical program committees of numerous conferences, such as ICC, Globecom, ITW, ISITA, VTC and APCC. He was a Guest Associate Editor for numerous special issues of the IEICE Transactions on Fundamentals. He was an industrial consultant for Huawei, ZTE, and Magnotech Ltd. He was a member of the Radio Spectrum Advisory Committee, Ofce of the Telecommunications Authority, Hong Kong S.A.R. Government from 2003 to 2008.

Po-Ning Chen (S'93-M'95-SM'01) was born in Taipei, R.O.C. in 1963. He received the B.S. and M.S. degrees in electrical engineering from National Tsing-Hua University, Taiwan, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from University of Maryland, College Park, in 1994.

From 1985 to 1987, he was with Image Processing Laboratory in National Tsing-Hua University, where he worked on the recognition of Chinese characters. During 1989, he was with Star Tech. Inc., where he focused on the development of finger-print recognition systems. After the reception of Ph.D. degree in 1994, he jointed Wan Ta Technology Inc. as a vice general manager, conducting several projects on Point-of-Sale systems. In 1995, he became a research staff in Advanced Technology Center, Computer and Communication Laboratory, Industrial Technology Research Institute in Taiwan, where he led a project on Java-based Network Managements. Since 1996, he has been an Associate Professor in Department of Communications Engineering at National Chiao-Tung University, Taiwan, and was promoted to a full professor since 2001. He was elected to be the Chair of IEEE Communications Society Taipei Chapter in 2006 and 2007, during which IEEE ComSoc Taipei Chapter won the 2007 IEEE ComSoc Chapter Achievement Awards (CAA) and 2007 IEEE ComSoc Chapter of the Year (CoY). He has served as the chairman of Department of Communications Engineering, National Chiao-Tung University, during 2007–2009. Since 2012, he becomes the associate chief director of Microelectronics and Information Systems Research Center, National Chiao Tung University, Taiwan, R.O.C.

Dr. Chen received the annual Research Awards from National Science Council, Taiwan, R.O.C., five years in a row since 1996. He then received the 2000 Young Scholar Paper Award from Academia Sinica, Taiwan. His Experimental Handouts for the course of Communication Networks Laboratory have been awarded as the Annual Best Teaching Materials for Communications Education by Ministry of Education, Taiwan, R.O.C., in 1998. He has been selected as the Outstanding Tutor Teacher of National Chiao-Tung University in 2002. He was also the recipient of Distinguished Teaching Award from College of Electrical and Computer Engineering, National Chiao-Tung University, Taiwan, in 2003. His research interests generally lie in information and coding theory, large deviation theory, distributed detection and sensor networks.

Dr. Jing Deng is an associate professor in the Department of Computer Science at the University of North Carolina at Greensboro (UNCG), Greensboro, North Carolina, USA. Dr. Deng visited the Department of Electrical Engineering at Princeton University and the Department of Electrical and Computer Engineering, WINLAB at Rutgers University in Fall of 2005. He was with the Department of Computer Science at the University of New Orleans from 2004 to 2008. He served as a Research Assistant Professor in the Department of Electrical Engineering and Computer Science at Syracuse University from 2002 to 2004. He received his Ph.D. degree from School of Electrical and Computer Engineering at Cornell University, Ithaca, New York, USA in January, 2002. He received his M.E. and B.E. degrees in Electronic Engineering at Tsinghua University in 1997 and 1994, respectively.

Dr. Deng is an editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. Dr. Deng's research interests include wireless network and security, information assurance, mobile ad hoc networks, and wireless sensor networks.

A. J. Han Vinck (F'06) is a full professor in Digital Communications at the University of Essen, Essen, Germany, since 1990. He studied electrical engineering at the University of Eindhoven, the Netherlands, where he obtained his Ph.D. in 1980. He accepted the position of visiting professor (2010-2012) at the University of Johannesburg, South Africa. He was invited (2011) to be consultant professor at the Harbin Institute of Technology, Harbin, China. In 2003 he was an adjoint professor at the Sun Yat-Sen University in Kaohsiung, Taiwan. In 1986 he was a visiting scientist at the German space Agency in Oberpfaffenhofen, Germany. His interest is in Information and Communication theory, Coding and Network aspects in digital communications. He is the author of the book *Coding Concepts and Reed-Solomon Codes*.

From 1991-1993, 1998-2000, 2006-2008, he was the director of the Institute for Experimental Mathematics in Essen. Professor Vinck was the director (1997-1999) of the Post-Graduate School on Networking, "CINEMA." From 2000-2004 he was the chairman for the communication division of the Institute for Critical Infrastructures, CRIS.

Professor Vinck served on the Board of Governors of the IEEE Information Theory Society from1997 until 2006. In 2003 he was elected President of the IEEE Information Theory Society. He served as Member-at-Large 2001-2002 in the Meetings and Services Committee for the IEEE. In 1999 he was the Program Chairman for the IEEE IT workshop in Kruger Park, South Africa (175 participants). In 1997 he acted as Co-chairman for the 1997 IEEE Information Theory symposium in Ulm, Germany (704 participants). Professor Vinck was founding Chairman (1995-1998) of the IEEE German Information Theory chapter. In 1990 he organized the IEEE Information Theory workshop in Veldhoven, the Netherlands (125 participants). IEEE elected him in 2006 as a fellow for his "Contributions to Coding Techniques." He is a distinguished lecturer for the Information theory as well as for the Communications society of the IEEE.

Professor Vinck is the initiator and organizer of the Japan-Benelux workshops on Information theory (now Asia-Europe workshop on "Concepts in Information Theory") and the International Winterschool on Coding, Cryptography and Information theory in Europe. He started (Essen, 1997) and still supports the organization of the series of conferences on Power Line Communications and its Applications, ISPLC (the 16th ISPLC2012 will be in Beijing). In 2006 he received the IEEE ISPLC2006 Achievement award in Orlando (FL, USA) for his contributions to Power Line Communications and for facilitating the transition of ISPLC to a fully financially and technically sponsored IEEE Communications Society conference.

The SA-IEE annual award was presented to him for the best paper published in the *SA-IEE Africa Research Journal* in the year 2008. He also received the best paper award at Chinacom 2013. He is the president of the Leibniz foundation. This foundation supports research and education in the field of Information theory, Neurosciences and biology.

APPENDIX

A. Detailed derivation of (15)

For $s > 0$, $E\left[e^{-s\phi_j}\,\middle|\,\mathcal{A}_{\mathbf{0}}\right] =$

$$\begin{aligned}
& \int_{-t}^{t} e^{-s(T-x^2)} f_{n_j}(x)dx + \int_{-t}^{t} e^{-s(x^2-T)} f_{n_j}(x-2\mu)dx + \left(1 - \int_{-t}^{t} f_{n_j}(x)dx - \int_{-t}^{t} f_{n_j}(x-2\mu)dx\right) \\
= \;& e^{-sT}\int_{-t}^{t} e^{sx^2}\left[(1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x}{\sigma_0}\right) + p_b\frac{1}{\sigma_1}\varphi\left(\frac{x}{\sigma_1}\right)\right]dx \\
& + e^{sT}\int_{-t}^{t} e^{-sx^2}\left[(1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x-2\mu}{\sigma_0}\right) + p_b\frac{1}{\sigma_1}\varphi\left(\frac{x-2\mu}{\sigma_1}\right)\right]dx \\
& + 1 - \int_{-t}^{t}\left[(1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x}{\sigma_0}\right) + p_b\frac{1}{\sigma_1}\varphi\left(\frac{x}{\sigma_1}\right) + (1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x-2\mu}{\sigma_0}\right) + p_b\frac{1}{\sigma_1}\varphi\left(\frac{x-2\mu}{\sigma_1}\right)\right]dx \\
= \;& e^{-sT}\left[(1-p_b)A_1(s,t,\sigma_0) + p_b A_1(s,t,\sigma_1)\right] + e^{sT}\left[(1-p_b)B_1(\mu,s,t,\sigma_0) + p_b B_1(\mu,s,t,\sigma_1)\right] \\
& + \left[(1-p_b)C_1(\mu,t,\sigma_0) + p_b C_1(\mu,t,\sigma_1)\right],
\end{aligned}$$

where by letting

$$a = \begin{cases} \frac{1}{\sqrt{1-2s\sigma^2}}, & 2s\sigma^2 < 1 \\ \infty, & 2s\sigma^2 = 1 \\ \frac{1}{\sqrt{2s\sigma^2-1}}, & 2s\sigma^2 > 1, \end{cases}$$

$b = \frac{1}{\sqrt{1+2s\sigma^2}}$, $\tilde{\varphi}(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}x^2}$, and $\tilde{Q}(y) = \int_0^y \tilde{\varphi}(x)dx$, we derive that

$$\begin{aligned}
A_1(s,t,\sigma) &= \int_{-t}^{t} e^{sx^2}\frac{1}{\sigma}\varphi\left(\frac{x}{\sigma}\right)dx = \int_{-t/\sigma}^{t/\sigma} e^{s\sigma^2 y^2}\varphi(y)\,dy = \int_{-t/\sigma}^{t/\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(1-2s\sigma^2)y^2}dy \\
&= \begin{cases} \int_{-t/(\sigma a)}^{t/(\sigma a)} a\cdot\varphi(x)\,dx, & 2s\sigma^2 < 1 \\ \frac{2t}{\sqrt{2\pi}\sigma}, & 2s\sigma^2 = 1 \\ \int_{-t/(\sigma a)}^{t/(\sigma a)} a\cdot\tilde{\varphi}(x)\,dx, & 2s\sigma^2 > 1 \end{cases} = \begin{cases} a\left(1 - 2Q\left(\frac{t}{\sigma a}\right)\right), & 2s\sigma^2 < 1 \\ \frac{2t}{\sqrt{2\pi}\sigma}, & 2s\sigma^2 = 1 \\ 2a\tilde{Q}\left(\frac{t}{\sigma a}\right), & 2s\sigma^2 > 1 \end{cases} \\
&= \begin{cases} \frac{1}{\sqrt{1-2s\sigma^2}}\left(1 - 2Q\left(\frac{\sqrt{1-2s\sigma^2}}{\sigma}t\right)\right), & 2s\sigma^2 < 1 \\ \frac{2t}{\sqrt{2\pi}\sigma}, & 2s\sigma^2 = 1 \\ \frac{2}{\sqrt{2s\sigma^2-1}}\tilde{Q}\left(\frac{\sqrt{2s\sigma^2-1}}{\sigma}t\right), & 2s\sigma^2 > 1 \end{cases}
\end{aligned}$$

and

$$\begin{aligned}
B_1(\mu,s,t,\sigma) &= \int_{-t}^{t} e^{-sx^2}\frac{1}{\sigma}\varphi\left(\frac{x-2\mu}{\sigma}\right)dx \\
&= e^{-2\mu^2(1-b^2)/\sigma^2}\int_{-t}^{t}\frac{1}{\sigma}\varphi\left(\frac{x-2\mu b^2}{\sigma b}\right)dx \\
&= e^{-2\mu^2(1-b^2)/\sigma^2}\int_{(-t-2\mu b^2)/(\sigma b)}^{(t-2\mu b^2)/(\sigma b)} b\varphi(y)\,dy \\
&= be^{-2\mu^2(1-b^2)/\sigma^2}\left(Q\left(\frac{2\mu b^2 - t}{\sigma b}\right) - Q\left(\frac{2\mu b^2 + t}{\sigma b}\right)\right) \\
&= \frac{1}{\sqrt{1+2s\sigma^2}}e^{-\frac{4s\mu^2}{1+2s\sigma^2}}\left(Q\left(\frac{2\mu - t(1+2s\sigma^2)}{\sigma\sqrt{1+2s\sigma^2}}\right) - Q\left(\frac{2\mu + t(1+2s\sigma^2)}{\sigma\sqrt{1+2s\sigma^2}}\right)\right),
\end{aligned}$$

and

$$C_1(t,\sigma,\mu) = 2Q\left(\frac{t}{\sigma}\right) - Q\left(\frac{2\mu - t}{\sigma}\right) + Q\left(\frac{2\mu + t}{\sigma}\right).$$

Notably, since $t \le \sqrt{E} = \mu$, it is always valid that for $2s\sigma^2 < 1$,

$$2\mu b^2 - t \ge \frac{2\mu}{1+2s\sigma^2} - \mu = \mu\frac{(1-2s\sigma^2)}{(1+2s\sigma^2)} > 0.$$

B. Detailed derivation of (16)

For $s > 0$,

$$\begin{aligned}
E\left[e^{-s\phi_j}\,\middle|\,\mathcal{A}_{\mathbf{0}}\right] &= e^{-4s\mu^2}\int_{-t}^{t-2\mu} e^{-4s\mu x} f_{n_j}(x)dx + e^{-sT}\int_{t-2\mu}^{t} e^{sx^2} f_{n_j}(x)dx \\
&\quad + e^{sT}\int_{-t}^{-t+2\mu} e^{-sx^2} f_{n_j}(x-2\mu)dx + \left(1 - \int_{-t-2\mu}^{t} f_{n_j}(x)dx\right)
\end{aligned}$$

$$\begin{aligned}
&= e^{-4s\mu^2}\int_{-t}^{t-2\mu} e^{-4s\mu x}\left[(1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x}{\sigma_0}\right)+p_b\frac{1}{\sigma_1}\varphi\left(\frac{x}{\sigma_1}\right)\right]dx\\
&\quad +e^{-sT}\int_{t-2\mu}^{t} e^{sx^2}\left[(1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x}{\sigma_0}\right)+p_b\frac{1}{\sigma_1}\varphi\left(\frac{x}{\sigma_1}\right)\right]dx\\
&\quad +e^{sT}\int_{-t}^{-t+2\mu} e^{-sx^2}\left[(1-p_b)b\frac{1}{\sigma_0}\varphi\left(\frac{x-2\mu}{\sigma_0}\right)+p_b\frac{1}{\sigma_1}\varphi\left(\frac{x-2\mu}{\sigma_1}\right)\right]dx\\
&\quad +\left(1-\int_{-t-2\mu}^{t}\left[(1-p_b)\frac{1}{\sigma_0}\varphi\left(\frac{x}{\sigma_0}\right)+p_b\frac{1}{\sigma_1}\varphi\left(\frac{x}{\sigma_1}\right)\right]dx\right)\\
&= e^{-4s\mu^2}\left[(1-p_b)D(s,t,\sigma_0,\mu)+p_bD(s,t,\sigma_1,\mu)\right]\\
&\quad +e^{-sT}\left[(1-p_b)A_2(s,t,\sigma_0,\mu)+p_bA_2(s,t,\sigma_1,\mu)\right]\\
&\quad +e^{sT}\left[(1-p_b)B_2(s,t,\sigma_0,\mu)+p_bB_2(s,t,\sigma_1,\mu)\right]\\
&\quad +\left[(1-p_b)C_2(\mu,t,\sigma_0)+p_bC_2(\mu,t,\sigma_1)\right]
\end{aligned}$$

where by letting

$$a=\begin{cases}\frac{1}{\sqrt{1-2s\sigma^2}}, & 2s\sigma^2<1\\ \infty, & 2s\sigma^2=1\\ \frac{1}{\sqrt{2s\sigma^2-1}}, & 2s\sigma^2>1\end{cases}$$

$b=\frac{1}{\sqrt{1+2s\sigma^2}}$, $\tilde{\varphi}(x)=\frac{1}{\sqrt{2\pi}}e^{\frac{1}{2}x^2}$, and $\tilde{Q}(y)=\int_0^y\tilde{\varphi}(x)dx$, we derive that

$$D(s,t,\sigma,\mu) = \int_{-t}^{t-2\mu} e^{-4s\mu x}\frac{1}{\sigma}\varphi\left(\frac{x}{\sigma}\right)dx=\int_{-t/\sigma}^{(t-2\mu)/\sigma} e^{-4s\mu\sigma y}\varphi(y)\,dy$$

$$\begin{aligned}
&= \int_{-t/\sigma}^{(t-2\mu)/\sigma} e^{8s^2\mu^2\sigma^2}\varphi(y+4s\mu\sigma)\,dy\\
&= e^{8s^2\mu^2\sigma^2}\int_{-t/\sigma+4s\mu\sigma}^{(t-2\mu)/\sigma+4s\mu\sigma}\varphi(x)\,dx\\
&= e^{8s^2\mu^2\sigma^2}\int_{(\mu/\sigma)(-t/\mu+4s\sigma^2)}^{(\mu/\sigma)(t/\mu-2+4s\sigma^2)}\varphi(x)\,dx\\
&= \begin{cases}
e^{8s^2\mu^2\sigma^2}\left(Q\left(\frac{2\mu-t}{\sigma}-4s\mu\sigma\right)-Q\left(\frac{t}{\sigma}-4s\mu\sigma\right)\right), & 4s\sigma^2<2-t/\mu\\
e^{8s^2\mu^2\sigma^2}\left(1-Q\left(\frac{t-2\mu}{\sigma}+4s\mu\sigma\right)-Q\left(\frac{t}{\sigma}-4s\mu\sigma\right)\right), & 2-t/\mu\le 4s\sigma^2<t/\mu\\
e^{8s^2\mu^2\sigma^2}\left(Q\left(-\frac{t}{\sigma}+4s\mu\sigma\right)-Q\left(\frac{t-2\mu}{\sigma}+4s\mu\sigma\right)\right), & t/\mu\le 4s\sigma^2
\end{cases},
\end{aligned}$$

$$\begin{aligned}
A_2(s,t,\sigma,\mu) &= \int_{t-2\mu}^{t} e^{sx^2}\frac{1}{\sigma}\varphi\left(\frac{x}{\sigma}\right)dx=\int_{(t-2\mu)/\sigma}^{t/\sigma} e^{s\sigma^2y^2}\varphi(y)\,dy=\int_{(t-2\mu)/\sigma}^{t/\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(1-2s\sigma^2)y^2}dy\\
&= \begin{cases}
\int_{(t-2\mu)/(\sigma a)}^{t/(\sigma a)} a\cdot\varphi(x)\,dx, & 2s\sigma^2<1\\
\frac{2\mu}{\sqrt{2\pi}\sigma}, & 2s\sigma^2=1\\
\int_{(t-2\mu)/(\sigma a)}^{t/(\sigma a)} a\cdot\tilde{\varphi}(x)\,dx, & 2s\sigma^2>1
\end{cases}\\
&= \begin{cases}
a\left(Q\left(\frac{t-2\mu}{\sigma a}\right)-Q\left(\frac{t}{\sigma a}\right)\right), & 2s\sigma^2<1 \text{ and } t>2\mu\\
a\left(1-Q\left(\frac{t}{\sigma a}\right)-Q\left(\frac{2\mu-t}{\sigma a}\right)\right), & 2s\sigma^2<1 \text{ and } \mu<t\le 2\mu\\
\frac{2\mu}{\sqrt{2\pi}\sigma}, & 2s\sigma^2=1\\
a\left(\tilde{Q}\left(\frac{t}{\sigma a}\right)-\tilde{Q}\left(\frac{t-2\mu}{\sigma a}\right)\right), & 2s\sigma^2>1 \text{ and } t>2\mu\\
a\left(\tilde{Q}\left(\frac{t}{\sigma a}\right)+\tilde{Q}\left(\frac{2\mu-t}{\sigma a}\right)\right), & 2s\sigma^2>1 \text{ and } \mu<t\le 2\mu
\end{cases}\\
&= \begin{cases}
\frac{1}{\sqrt{1-2s\sigma^2}}\left(Q\left(\frac{\sqrt{1-2s\sigma^2}}{\sigma}(t-2\mu)\right)-Q\left(\frac{\sqrt{1-2s\sigma^2}}{\sigma}t\right)\right), & t>2\mu\\
\frac{1}{\sqrt{1-2s\sigma^2}}\left(1-Q\left(\frac{\sqrt{1-2s\sigma^2}}{\sigma}t\right)-Q\left(\frac{\sqrt{1-2s\sigma^2}}{\sigma}(2\mu-t)\right)\right), & \mu<t\le 2\mu\\
\frac{2\mu}{\sqrt{2\pi}\sigma}, & 2s\sigma^2=1\\
\frac{1}{\sqrt{2s\sigma^2-1}}\left(\tilde{Q}\left(\frac{t}{\sigma a}\right)-\tilde{Q}\left(\frac{t-2\mu}{\sigma a}\right)\right), & 2s\sigma^2>1 \text{ and } t>2\mu\\
\frac{1}{\sqrt{2s\sigma^2-1}}\left(\tilde{Q}\left(\frac{t}{\sigma a}\right)+\tilde{Q}\left(\frac{2\mu-t}{\sigma a}\right)\right), & 2s\sigma^2>1 \text{ and } \mu<t\le 2\mu
\end{cases}
\end{aligned}$$

and

$$\begin{aligned}
B_2(s,t,\sigma,\mu) &= \int_{-t}^{-t+2\mu} e^{-sx^2}\frac{1}{\sigma}\varphi\left(\frac{x-2\mu}{\sigma}\right)dx = e^{-2\mu^2(1-b^2)/\sigma^2}\int_{-t}^{-t+2\mu}\frac{1}{\sigma}\varphi\left(\frac{x-2\mu b^2}{\sigma b}\right)dx \\
&= e^{-2\mu^2(1-b^2)/\sigma^2}\int_{(-t-2\mu b^2)/(\sigma b)}^{(-t+2\mu-2\mu b^2)/(\sigma b)} b\varphi\left(y\right)dy \\
&= be^{-2\mu^2(1-b^2)/\sigma^2}\left(Q\left(\frac{2\mu b^2+t-2\mu}{\sigma b}\right)-Q\left(\frac{2\mu b^2+t}{\sigma b}\right)\right), \\
&= \frac{1}{\sqrt{1+2s\sigma^2}}e^{-\frac{4s\mu^2}{1+2s\sigma^2}}\left(Q\left(\frac{t+2ts\sigma^2-4s\mu\sigma^2}{\sigma\sqrt{1+2s\sigma^2}}\right)-Q\left(\frac{t+2st\sigma^2+2\mu}{\sigma\sqrt{1+2s\sigma^2}}\right)\right),
\end{aligned}$$

and

$$C_2(t,\sigma,\mu) = Q\left(\frac{t}{\sigma}\right)+Q\left(\frac{t+2\mu}{\sigma}\right).$$

Notably, as $t > \mu$, it is always valid that

$$2\mu b^2+t-2\mu = t-\frac{4\mu s\sigma^2}{1+2s\sigma^2} > \mu-\frac{4\mu s\sigma^2}{1+2s\sigma^2} = \mu\frac{(1-2s\sigma^2)}{(1+2s\sigma^2)} > 0.$$