

RESEARCH

Open Access



Rateless coding transmission over multi-state dying erasure channel for SATCOM

Shushi Gu¹, Jian Jiao^{1*} , Qinyu Zhang¹ and Xuemai Gu²

Abstract

Satellite communication (SATCOM) systems have attracted great attention from academic and industrial communities in recent years, and huge amount of data delivery over satellite downlinks is considered as a promising service in emerging 5G networks, such as multimedia broadcasting. Nevertheless, due to intermittent connections from LEO or MEO satellite to earth station, and high dynamic channel conditions over downlinks, satellites may not be able to transmit the large data files to the ground station on time. In this paper, we propose a new rateless coding transmission for multi-state dying erasure channels (MDEC) with random channel life span and time-varying packet error rates, to improve the transmitting capability over SATCOM downlinks. Firstly, a heuristic approach for suboptimal degree distributions based on AND-OR tree technique is presented to achieve higher intermediate performance and lower symbol error rate of our proposed rateless codes. Furthermore, the appropriate code length of the connective window is derived and analyzed for enhanced average throughput on MDEC that is also optimized by maximum problem solving. Simulations have been conducted to evaluate the effectiveness of our rateless coding transmission for large file delivery on dynamic channel conditions. The results demonstrate that our proposed transmission scheme outperforms existing conventional rateless codes with significantly better intermediate performance and throughput performance over unreliable SATCOM downlinks, under time-varying packet error rates and unpredictable occurrences of exhausted energy or cosmic ray attacks.

Keywords: Satellite communication, Multi-state dying erasure channel, Rateless codes, Degree distribution, Optimal code length

1 Introduction

Satellite communication (SATCOM) systems have attracted great attention from academic and industrial communities in recent years, which have been widely used for many military and civil services, e.g., weather forecast, environment monitoring, multimedia service, positioning system, and emergency rescue [1–4]. In comparison to terrestrial communication, SATCOM systems have the advantages of larger bandwidth and wider coverage, for providing the huge amount of data services in the emerging fifth generation (5G) networks, such as hybrid satellite-terrestrial communication systems [5, 6].

Nevertheless, satellite systems have significantly different link characteristics than terrestrial links [7]. Firstly,

since low earth orbit (LEO) or medium earth orbit (MEO) satellites (SA) fly along their own orbits and have very limited contact time with earth stations (ES), there is no sufficient contact time between SA and ES to download all data information [8]. Besides, the satellite downlink channel conditions are indeed time-varying due to atmospheric precipitation impacting high-frequency bands, so the time invariance assumption no longer holds [9]. Furthermore, the telecommunication systems on satellite may be subjected to serious effects, including the lack of battery energy and the attack of solar winds or cosmic rays, so that the communication links between SAs and ESs would randomly break down unpredictably. The above situations make great difficulties for the large amount of data transmitted back to the ground on the downlink channels. Therefore, it is necessary to give a theoretical channel model to describe the extremely dynamic characteristics on SATCOM channel, as well as to design an efficient

*Correspondence: jiaojian@hit.edu.cn

¹Communication Engineering Research Center, Harbin Institute of Technology-Shenzhen, The Xili University Town, Nanshan District, Shenzhen 518055, China

Full list of author information is available at the end of the article

and reliable transmission technique for large satellite data service downloading.

In [10], the authors firstly investigate a special type of channel with a finite and random channel length, termed *dying channel*. This type of channel may suddenly terminate due to communication links subjected to random fatal impacts, e.g., the sensor node may run out of power or be destroyed by fire attacks of military equipments in hostile environment, and the communication systems embedded in biological cells that may disappear unpredictably, due to excretion and digestion. It is critical to quantify how fast and reliably the information can be collected over this attacked channels with finite channel life span. Dying channels are modeled as the finite-state semi-Markov channels in [11]. It proves that lower Shannon limit which is very close to zero is almost reached on discrete memoryless channel that dies, so that arbitrarily small probability of error is not achievable. The authors optimize the sequences of finite block-length channel coding, to maximize the transmission volume communicated at fixed maximum message error probabilities by dynamic programming. Furthermore, the notion of the channel death is similar to the outage over fading channels. The outage capacity and the outage probability over dying channels are both given and derived in [12]. It also gives the optimization over the frame length and the power allocation over the constituting data blocks to minimize the outage probability. In consequence, the literatures above have motivated us to model the time-varying downlink between SA and ES subjected to unknown random blockages as a special form of dying channel.

Meanwhile, rateless codes, e.g., LT codes [13] and Raptor codes [14], as well-known erasure coding approaches, have attracted considerable interest in the areas of computing and communications. Unlike other error-correcting codes, i.e., RS and LDPC codes, rateless codes do not need a predetermined code-rate and have on-the-fly encoding and decoding property. It means that the encoder could continuously generate encoded symbols, while the decoder could gradually recover the original data files in real time. Rateless codes also have the following properties. (1) They can totally recover the original message from any subset of the encoded symbols of the length slightly larger than the original message that also can be seen as the near-MDS codes. (2) They allow very efficient encoding and decoding complexities, even as the logarithmic or linear number of the input symbols. In recent years, rateless codes have been widely used in vehicular networks for data dissemination by their high speed broadcasting ability [15]. Besides, based on unequal error protection (UEP) property, rateless codes are also employed for multimedia streaming service in terrestrial networks, and the adaptive recovery of the layered multimedia progressively improves the quality of experience

(QoE) of the receivers [16]. Therefore, it can be confirmed that the flexible adaptation and capacity-achieving make rateless codes natively appropriate for the large data bulk transmission on the satellite downlinks. There have been some studies about rateless codes used in SATCOM systems for burst error resistance and DVB-S file delivery [17, 18].

For rateless codes over the dying channel, the topic of good *intermediate performance* of rateless codes would be widely studied. The author of [19] firstly investigated the intermediate performance of LT codes through deriving the asymptotic error probability and provided a tight outer bound on the proportion of recoverable symbols. Building on the basis of [19], the authors in [20] presented two new methods (degree distribution optimization and encoded symbol sorting) to maximize the intermediate symbol recovery rate of rateless coding. However, in their work, the different degree distributions were needed to match the different numbers of received encoded symbols, which may cause the transmitting system more complex. The investigation on rateless codes over dying channel was firstly presented in [21]. The authors optimized the degree distribution of LT codes for both partial and full recovery situations, aiming to match the stochastic channel life span. But their model failed to consider a constituting procedure of large file transmission over multiple connections or multiple channel states over the dying channels. The authors of [22] proposed a new rateless code, named the growth codes, to improve intermediate performance and to maximize video transmission quality over error resilient channels. The transmitter continuously adapts the degree of the encoded symbols to guarantee that each received encoded packet has the highest instantaneous recoverable probability. The online fountain codes [23] is another class of rateless codes designed for higher intermediate performance. The lower redundancy overhead of online fountain codes is the most important improvement over the growth codes, and the feedback cost is roughly half that of growth codes. But the common drawback of growth codes and online codes is that feedback signal is required to inform the alterations of the decoding states in transmission process, in order to change an optimal encoding strategy at the transmitter, which may not be easily employed in some practical SATCOM networks.

In this paper, we consider a multi-state dying erasure channel (MDEC) model different from [21], which is window-connected with dynamic variation of the packet error rates and channel life spans. Different channel death distributions of MDEC are considered to describe the unpredictable effect factors on SATCOM links, such as device exhaustion or random attacks of cosmic radiation. Three channel states (good, bad and dead) on downlinks between SA and ES are modeled due to the influences

by atmospheric precipitation for high-frequency carrier waves, with a semi-Markov chain following time changes. The objective of this work is to improve the throughput of the large bulk files delivering over this multi-state dying erasure channel, through proposing a rateless coding transmission with appropriate degree distributions and code lengths for lower symbol error rate.

Our major contributions are summarized as three aspects: (i) we give the throughput expression of this multi-state dying erasure channel to describe intermittent and time-varying characteristic on SATCOM links, and also indicate the essential methods for rateless transmission designing over this channel model; (ii) we propose the sub-optimal degree distributions by a heuristic approach based on And-OR tree technique to minimize the symbol error rate over MDEC; and (iii) we analyze and derive the appropriate code lengths to improve the intermediate performance and throughput by a maximizing problem solving.

The remainder of this paper is organized as follows. Section 2 introduces the multi-state dying erasure channel modeling, presents the rateless transmission, and gives the throughput analysis. Relying on this model and problem statement, Section 3 proposes a certain degree distribution design method and optimizes the code length by solving a maximization problem. The simulation experiments and results are presented in Section 4, and Section 5 concludes the paper.

2 System model and performance analysis

2.1 Multi-state dying erasure channel

We present a new channel model termed in Fig. 1 for SATCOMs, named multi-state dying erasure channel (MDEC), which targets data packet transmission from SA to ES over the intermittent and time-varying binary erasure channel (BEC). Since LEO or MEO satellites move along their orbits periodically, MDEC channels have N ($N = 1, 2, 3, \dots$) connective windows, which are used to opportunistic convey the data packets from SA to ES. The unit time of each packet is τ , and thus, the connection

time for one window is fixed to $K\tau$. In one connective window, satellite link would be blockaded randomly, attributed to random events (energy exhausted or radiation damage) at unknown times. At this moment, we defined it as channel *dead state*, the connective window is unavailable for transmitting. The satellite must wait for recharging battery cells or restarting device, so the data transmission will break off unexpectedly until the next connection is established. We define the channel life span of MDEC as L , which follows an exponential distribution with mean $\frac{1}{\lambda}$ or a non-negative Gaussian distribution with mean μ and variance σ^2 , i.e., $L \sim E(\lambda)$ or $L \sim \mathcal{N}(\mu, \sigma^2)$. These two random distributions are able to accurately model the channel death caused by device exhaustion or random attacks of cosmic radiation, respectively.

Beside channel *dead state*, MDEC has another two channel states, defined as *good state* and *bad state*. Because the atmospheric environment conditions (weather and solar wind) is changing with time in every window-connection, the erasure probability of MDEC is also time-varying due to different noise temperatures of the receiver. Especially, when the transmitter of the emerging LEO satellite uses Ka-band (26 and 32 GHz), enlarging the capabilities of downlinks, the higher frequency produces serious atmospheric fading due to rain scintillation, cloud and gaseous attenuations. Denote p as the erasure probability in one window, which is modeled by a two-state semi-Markov chain, i.e., $p = p_G$ or p_B . The states evolve with two transition probabilities, i.e., λ_G and λ_B . The transition matrix \mathbf{P} of the semi-Markov process can be written follows:

$$\mathbf{P} = \begin{bmatrix} P(G|G) & P(B|G) \\ P(G|B) & P(B|B) \end{bmatrix} = \begin{bmatrix} 1 - \lambda_G & \lambda_G \\ \lambda_B & 1 - \lambda_B \end{bmatrix} \quad (1)$$

where λ_G indicates the probability that the channel transits from bad state to good state, while λ_B indicates the probability that the channel transits from good state to bad state.

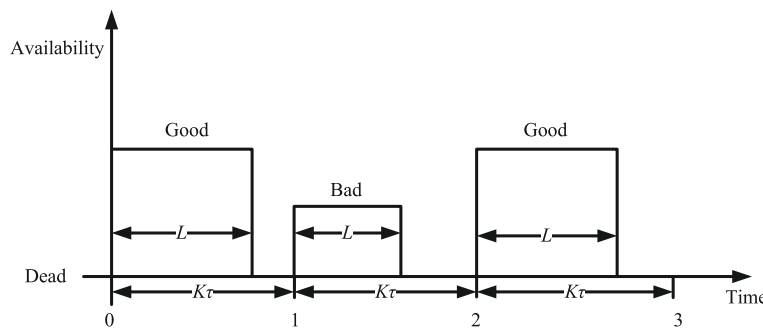


Fig. 1 Multi-state dying erasure channel model. MDEC model has three channel states, good, bad, and dead. The good and bad state transitions follow a semi-Markov process, and the duration of channel life span is random in one connective window

In summary, MDEC channel in Fig. 1 has three channel states, good, bad, and dead. Good and bad states (available) mean the channel is available for data transmission, but the transmitter cannot know the accurate state is good or bad of every window and just knows the statistical erasure probabilities of the two states and Markov transition matrix. The occurrence of dead state (unavailable) is also unknown by transmitter, and the duration of channel lifespan is random. MDEC model exhibits extremely quick dynamics, so SATCOM system cannot provide availability of reliable communications. Since limited contact time with ESs, satellite may not have sufficient contact time to download all its data to the ground.

2.2 Rateless coding transmission over MDEC

In order to resist the packets discarded and damaged, an appropriate transmission method is desired to improve the data reliability and efficiency over the MDEC with finite and random channel lifespan L and time-varying erasure probability p . It is important to note that automatic repeat-request (ARQ) is not suitable for MDEC. Since the connection time is limited, frequent retransmissions of lost or erroneous packets waste many opportunities of large data volume delivery from SA to ES. In addition, the feedback information is hard to inform channel state information (CSI) to the transmitter due to the dynamic channel conditions and the random channel death, which makes the advantages of ARQ mechanism degrade significantly.

Being able to adapt time-varying channel conditions, rateless codes have attracted a lot of interest as a promising forward error correction (FEC) technique. We use rateless codes to generate original data into encoded packets over MDEC, as illustrated in Fig. 2, in order to guarantee the reliability and improve the efficiency. Generally speaking, packets after rateless encoded are referred to output symbols, while source packets are regard as input symbols. Rateless codes generate output symbols with a specific degree distribution $\Omega(x) = \sum_{d=1}^D \Omega_d x^d$, where D is the maximal degree and Ω_d is the probability that d input symbols are randomly selected to be XORed to generate an output symbol.

Firstly, denote by k as the code length, which is the number of input symbols to be sent in one connective window. Secondly, the number of output symbols actually transmitted by SA is denoted as l . Apparently, l is a random variable decided by the channel lifespan L and length of window K , i.e., $l = \min(K, \lfloor \frac{L}{\tau} \rfloor)$. Once the channel switches to the dead state unpredictably, the transmitting flow of output symbols would break off at this moment, as $l = \lfloor \frac{L}{\tau} \rfloor$. If the channel state sustains available in one connective window, the transmitting flow of output symbols can be continuous until the end of connection, as

$l = K$. Lastly, the number of output symbols successfully received by the ES (l after erased with probability p) in one window is denoted as a random variable n , so the decoding overhead at the decoder is given by $\gamma = \frac{n}{k}$. If the channel dies or connective window finishes, the receiver will start to decode the output symbols and recover the original data. The decoder may partially or fully recover the original data in one window, which means that intermediate performance of rateless codes is necessarily considered.

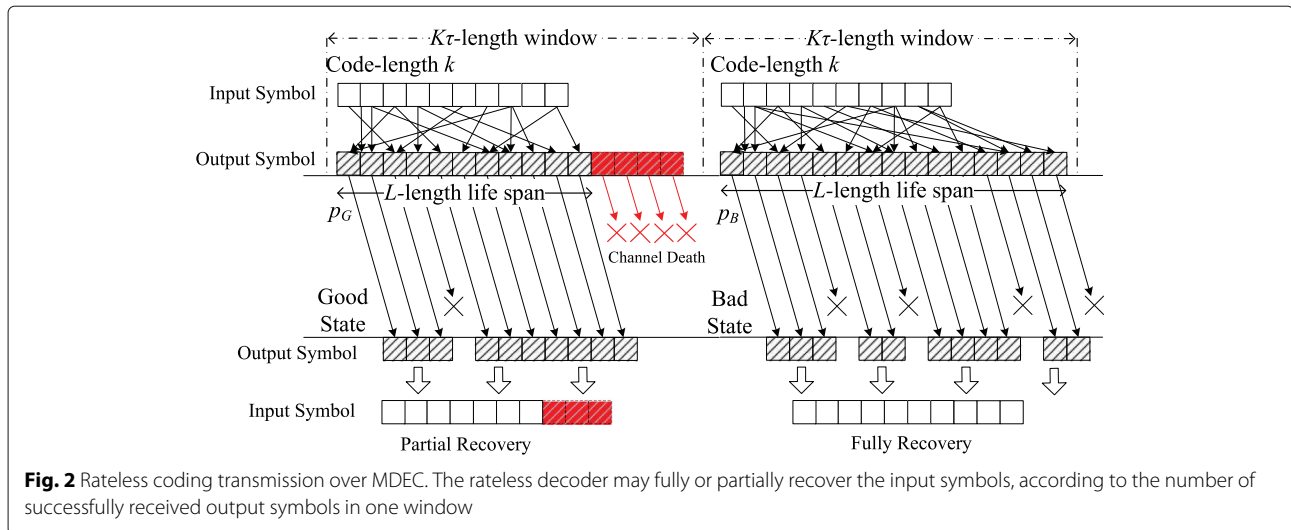
2.3 Throughput performance analysis

As we know, the rateless decoding performance depends largely on the decoding overhead γ as well as the degree distribution Ω . Due to channel erased or death, the receiver may not be able to collect enough output symbols (i.e., $n \leq k$ and $\gamma < 1$). We define $z(\gamma, \Omega)$ as the intermediate performance of our coding scheme, conditioned on the given γ and Ω , which represents the ratio of the recovered input symbols at the receiver to all the input symbols k . In MDEC model, the number of output symbols successfully received at the ES in every window, i.e., n , is randomly changing, due to the time-varying channel states transitions. Then, we define the average intermediate performance z_{avg} after infinite number of windows can be expressed as

$$z_{avg}(k, \Omega) = \sum_{n=1}^K z\left(\frac{n}{k}, \Omega\right) \Pr(n), \quad (2)$$

where $\Pr(n)$ is the probability of n output symbols successfully received by the receiver in one window, and as we know that $\sum_{n=1}^K \Pr(n) = 1$. Furthermore, we can also express the probability that output symbols erased on MDEC, named as the output symbol error rate SER_{out} . The detailed derivation about the relationship between $\Pr(n)$, SER_{out} and channel parameters will be given in Section 3 of this paper.

In Eq. (2), z_{avg} provides a reasonable measurement for the average input symbol recovery ratio over MDEC, conditioned upon two crucial parameters about the code length k and the encoding degree distribution Ω . We define the input symbol decoding error rate at the receiver as SER , which can be easily expressed as $SER = 1 - z_{avg}(k, \Omega)$. Therefore, larger k means less redundancy, and $\gamma = \frac{n}{k}$ decreases, resulting in reduced z_{avg} and raise SER . Besides, the intermediate performance z_{avg} and input symbol error rate SER are affected by the degree distribution Ω selections. When the average degree is higher (e.g., robust soliton distribution of LT codes), the rateless codes are suitable for the case that $\gamma > 1$, as full recovery. When the average degree is lower (i.e., only degree-one and degree-two encoded packets), rateless coding is suitable for the case that $\gamma < 1$, as partial recovery. Consequently,



one of the objectives of our work is to find the most appropriate k and $\Omega(x)$ under certain channel conditions, in an effort to maximize the intermediate performance z_{avg} and minimize input symbol error rate SER at the receiver.

Traditionally, if we let A_i denote the amount of data that SA has actually conveyed to the ES in the i th connective window, the average data throughput over N window-connective downlinks is defined as the data amount per unit time and can be computed as follows:

$$\text{throughput} \triangleq \frac{1}{NK\tau} \sum_{i=1}^N A_i. \quad (3)$$

In our works, for simplicity, we make the assumption that the data transmission between the SA and the ES uses the same packet size and same unit time. That is, the data size and transmission time are interchangeable. The amount of data received by the ES, i.e., A_i in Eq. (3), can be represented as the total number of input symbols recovered. Since each window in the MDEC is i.i.d, the average and normalized throughput of our rateless coding transmission on MDEC system is defined as the ratio of the average number of input symbols recovered in one window to the window length K , basing on (2), which can be expressed as follows:

$$\text{throughput} = \frac{kz_{avg}}{K} = \frac{k}{K}(1 - SER) = \frac{k}{K} \sum_{n=1}^K z\left(\frac{n}{k}, \Omega\right) \Pr(n). \quad (4)$$

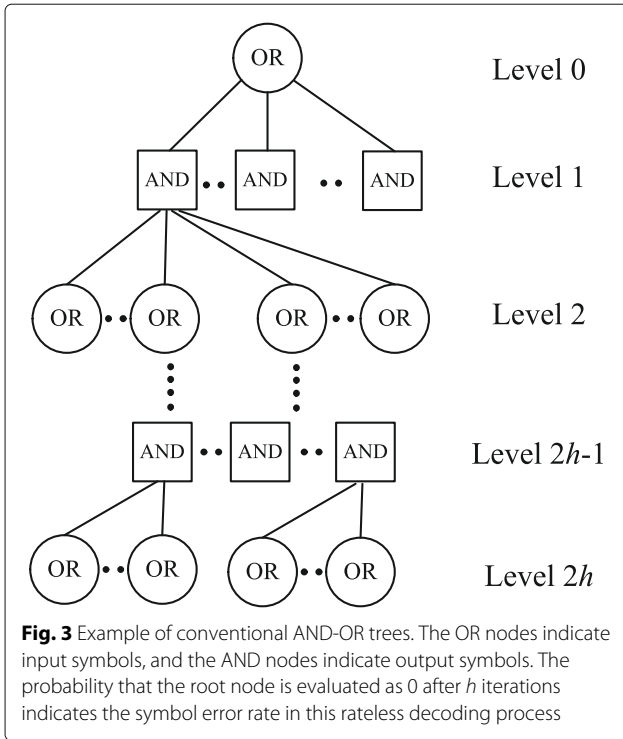
As can be easily seen that, if K is fixed, *throughput* in Eq. (4) improves by increasing k , which implies more input symbols can be transmitted. But as the above description, larger k makes lower intermediate performance z_{avg}

and higher input symbol error rate SER , so the *throughput* reduces especially when the lifespan L is shorter or the channel erasure probability p is higher. In addition, z_{avg} and SER are both closely decided by Ω in the decoding process. Therefore, for our final objective in this work, we need to maximize the *throughput* in Eq. (4), maximize the intermediate performance z_{avg} in Eq. (2), and minimize the input symbol error rate SER using the optimization method to tradeoff appropriate code lengths k and appropriate degree distributions $\Omega(x)$, based on the time-varying and dying channel conditions of MDEC model.

3 Degree distribution design and code length optimization

To analyze the potential performance of rateless codes, the AND-OR tree analysis technique [24] has been adopted in many literatures [20, 21, 25, 26]. Since we modified the AND-OR tree structure to analyze rateless coding over MDEC, a brief introduction of conventional AND-OR trees is as follows. An AND-OR tree, T_h , is a generated tree of depth $2h$ as shown in Fig. 3. The root of this tree is at depth 0. Its children are at depth 1, and their children are at depth 2, and so on. All nodes at depth 0, 2, 4, ..., $2h-2$ are labeled as OR nodes, and all nodes at depth 1, 3, 5, ..., $2h-1$ are labeled as AND nodes. An OR node with no children is assumed to have a value of zero, and an AND node with no children is assumed to have a value of one. It is shown that iterative BP decoding of rateless codes can be mapped to the logical operations in the AND-OR tree. An OR node represents an input symbol in rateless codes, and an AND node stands for an output symbol. More details can be found in [24].

In general, the asymptotic decoding performance predicted by AND-OR tree analysis is almost the same as that



of the finite-length rateless codes if the codes are sparse enough. Briefly, the asymptotic performance is given based on the standard approximation of the Binomial distribution $(1/k, \Omega'(1)k)$ by a Poisson input degree distribution $\delta(x) = e^{-\Omega'(1)\gamma(x-1)}$. Then, we define y_h as the asymptotic probability that one input symbol cannot be recovered after h iterations, equal to the input symbol error rate SER , is given by the iterative formula as follow

$$SER = y_h = \delta(1 - \beta(1 - y_{h-1})) \tag{5}$$

where $\beta(x) = \frac{\Omega'(x)}{\Omega'(1)}$ and $y_0 = 1$. Then, the input symbol error rate can be also derived based on decoding overhead γ and degree distribution $\Omega(x)$ as

$$SER(\gamma, \Omega) = \exp(-\gamma\Omega'(1 - y_{h-1})). \tag{6}$$

The sequence y_h converges with respect to the number of decoding iterations h . Consequently, as mentioned in Section 2, the average intermediate performance z_{avg} is the proportion of input symbols recovered in one window, which can be expressed as $z_{avg} = 1 - SER$. We will give the SER performances of our codes after degree distribution and code-length optimized, compared to those of LT and Raptor codes in the following sections.

3.1 Linearly combined degree distribution Ω_{LC}

Based upon the above discussions, we attempt a simple heuristic approach to design a sub-optimal degree distribution for rateless coding scheme in MDEC. We choose three classical degree distributions similarly to

the method in [19] and combine them to form a new degree distribution to improve the intermediate performance z with an assumption of finite code length k . That is, $\Omega^{(1)}(x) = x$, $\Omega^{(2)}(x) = x^2$, and the weak robust soliton distribution [14]:

$$\begin{aligned} \Omega^{(3)}(x) = & 0.008x + 0.494x^2 + 0.166x^3 + 0.072x^4 \\ & + 0.083x^5 + 0.056x^8 + 0.037x^9 + 0.056x^{19} \\ & + 0.025x^{64} + 0.003x^{66}. \end{aligned} \tag{7}$$

The first two degree distributions are chosen in consideration of the dead and bad states of MDEC, while the weak robust soliton distribution accounts for relatively general cases as well as the good state.

It is easy to see from [19] that $\ln 2$ and $\frac{3}{4} \ln 3$ are two dividing points for γ in the sense of the upper bound of the intermediate performances of the up bound. This observation motivates us to linearly combine $\Omega^{(1)}$, $\Omega^{(2)}$, and $\Omega^{(3)}$ to form a weighted degree distribution $\Omega_{LC}(x)$ shown below:

$$\Omega_{LC}(x) = \sum_{j=1}^3 \omega_j \Omega^{(j)}(x), \tag{8}$$

where $\omega = [\omega_1, \omega_2, \omega_3]$ is the weighting vector. The weighting coefficients are taken as the probabilities corresponding to the different decoding overhead intervals, i.e., $\omega_1 = \Pr\{0 \leq \gamma \leq \ln 2\}$, $\omega_2 = \Pr\{\ln 2 \leq \gamma \leq \frac{3}{4} \ln 3\}$, $\omega_3 = \Pr\{\frac{3}{4} \ln 3 \leq \gamma\}$. Similar to Raptor codes, the encoding and decoding complexities with Ω_{LC} is also linear, since the average degree of the output symbols is a constant. We will give the simulated results of average degree in the following section. This characteristic is important for rateless codes used in satellite systems with lower computation complexity. Using the And-OR tree technique with $y_{LC,0} = 1$, the asymptotic intermediate performance z_{LC} using Ω_{LC} can be readily shown as

$$z_{LC}(\gamma, \Omega_{LC}) = 1 - \exp(-\gamma\Omega'_{LC}(1 - y_{LC,h-1})). \tag{9}$$

The above heuristic degree distribution exhibits two notable advantages. Firstly, this linear combination approach only needs three weighting coefficients to generate a practical degree distribution with sub-optimum intermediate performance z_{LC} , which is able to save enormous computational power in hardware equipments of SA. Secondly, the three weighting coefficients are all related to the probability intervals of the decoding overhead γ . If the three weighting coefficients are set properly in accordance with the random distribution of lifespan L and the Markov transition probabilities λ_G and λ_B , the proposed rateless codes can dynamically adapt the degree distribution $\Omega_{LC}(x)$ to the random channel conditions.

3.2 Optimization of the code length k

Our next step is to optimize the code length k with the linear combined degree distribution $\Omega_{LC}(x)$, so as to maximize the *throughput* in Eq. (4). The average intermediate performance of the linear combined degree distribution, i.e., $z_{LC}(k, \omega)$, can be expressed as a function of k and $\omega = [\omega_1, \omega_2, \omega_3]$ similarly to Eqs. (8) and (9). Since K is fixed, maximizing throughput in Eq. (4) is equal to maximizing $kz_{LC}(k, \omega)$, which can be transformed into the following maximum problem:

$$\begin{aligned} \max_{k, \omega} \quad & kz_{LC}(k, \omega) & (10) \\ \text{s.t.} \quad & \omega_1 + \omega_2 + \omega_3 = 1 \\ & 0 < k \leq K. \end{aligned}$$

It is not difficult to see that k and ω in $\Omega_{LC}(x)$ are not independent. Because the weighting vector ω is defined as the probability distribution of $\gamma = \frac{n}{k}$ in different decoding overhead intervals. That is, the choice of k affects the value of weighting vector ω . As a result, ω_1, ω_2 , and ω_3 can be transformed to functions of k . Therefore, we attempt to transform the above maximum problem with multiple variables into a single variable optimization problem. We derive the relationship between ω and k over the MDEC with an exponentially distributed channel life span, while the derivation with a non-negative Gaussian distributed channel life span is similar.

Firstly, in the simplest BEC channel, the probability of n output symbols successfully received at the decoder follows a binomial distribution, i.e.,

$$\Pr_1(n) = \binom{l}{n} (1-p)^n p^{l-n}, \quad (11)$$

where l is the number of output symbols transmitted. Then, assuming the channel state follows the two-state Markov chain with the transition matrix \mathbf{P} in Eq. (1), the probability $\Pr_2(n)$ that n output symbols successfully received over Markov channels should be transformed based on $\Pr_1(n)$, i.e.,

$$\Pr_2(n) = q_B \binom{l}{n} (1-p_B)^n p_B^{l-n} + q_G \binom{l}{n} (1-p_G)^n p_G^{l-n}, \quad (12)$$

where the coefficients q_B and q_G can be expressed as $q_B = \frac{\lambda_G}{\lambda_G + \lambda_B}$ and $q_G = \frac{\lambda_B}{\lambda_G + \lambda_B}$.

Secondly, the probability of l output symbols in one window having been transmitted before death of the channel, i.e., $\Pr(l)$, can be derived as follows. If we assume that l symbols can be transmitted before the time instant of $(l+1)\tau$ in one window, $\Pr(l)$ can be shown as

$$\Pr(l) = \begin{cases} \int_{l\tau}^{(l+1)\tau} f(t) dt & 0 \leq l < K \\ \int_{K\tau}^{\infty} f(t) dt & l = K \end{cases}, \quad (13)$$

where $f(t) = \lambda e^{-\lambda t}$ is the probability density function of the exponential distribution of the channel life span $L(t)$. So, the probability of n output symbols successfully received by the ES, denoted as $\Pr(n)$, the same probability $\Pr(n)$ in Eq. (2), can be expressed as follows

$$\begin{aligned} \Pr(n) = \sum_{l=0}^K \Pr(l) & \left(q_B \binom{l}{n} (1-p_B)^n p_B^{l-n} \right. \\ & \left. + q_G \binom{l}{n} (1-p_G)^n p_G^{l-n} \right). \end{aligned} \quad (14)$$

Besides, similarly as above derivations, the probability of n output symbols erased over MDEC in one connective window, defined as $\Pr_{out}(n)$, can be shown as

$$\begin{aligned} \Pr_{out}(n) = \sum_{l=0}^K \Pr(l) & \left(q_B \binom{l}{n} (1-p_B)^{l-n} p_B^n \right. \\ & \left. + q_G \binom{l}{n} (1-p_G)^{l-n} p_G^n \right). \end{aligned} \quad (15)$$

Therefore, the output symbol error rate SE_{out} over MDEC in one K -length window can be expressed based on Eqs. (13) and (15) as

$$SE_{out} = \frac{1}{K} \sum_{n=0}^K n \Pr_{out}(n). \quad (16)$$

It can be easily seen that the output symbol error rate SE_{out} is only relevant to the channel conditions of MDEC, including erasure probabilities p_G and p_B , transition probabilities λ_G and λ_B , and channel lifespan $L(t)$. Hence, SE_{out} describes the dynamic switching of the MDEC states.

Thirdly, since $\gamma = \frac{n}{k}$ and k is fixed, the probability $\Pr(n)$ is equal to the probability of γ for all cases. If we put $\Pr(n)$ in Eq. (14) into the probability intervals of $\omega_1, \omega_2, \omega_3$ in Eq. (8), we can turn $\omega_1, \omega_2, \omega_3$ into the functions of k as follows:

$$\begin{aligned} \omega_1(k) &= \sum_{n=0}^{\lfloor \ln 2 k \rfloor} \Pr(n), \\ \omega_2(k) &= \sum_{n=\lceil \frac{3}{4} \ln 3 k \rceil}^{\lfloor \frac{3}{4} \ln 3 k \rfloor} \Pr(n), \\ \omega_3(k) &= \sum_{n=\lceil \frac{3}{4} \ln 3 k \rceil}^K \Pr(n). \end{aligned} \quad (17)$$

Lastly, the problem (10) can be converted to a single variable optimization problem (18) on k as shown below. The first constraint condition is transformed from Eq. (9) for the relationship between z_{LC} and k .

$$\begin{aligned}
 & \max_k \quad k \sum_{n=1}^K z_{LC}(k) \Pr(n) \tag{18} \\
 & \text{s.t.} \quad z_{LC} = 1 - \exp\left(-\frac{n}{k} (\omega_1(k) + 2\omega_2(k)(1 - y_{LC,h-1}) \right. \\
 & \quad \left. + \omega_3(k)\Omega^{(3)'}(1 - y_{LC,h-1}))\right) \\
 & \quad \omega_1(k) + \omega_2(k) + \omega_3(k) = 1 \\
 & \quad 0 < k \leq K.
 \end{aligned}$$

The initial value $y_{LC,0} = 1$. The intermediate performance $z_{LC} = 1 - y_{LC,h}$ converges roughly with respect to a constant number of decoding iterations h (according to the following simulations, $h = 40$ is adequate for convergence). Because of the on-the-fly encoding and decoding property, the iteration time delay will not severely impact the throughput performance at the receiver. We can ignore the decoding delay in our analysis.

Furthermore, it is easy to see that the complexity of problem (18) is $O(K^2)$ with exhaustive search algorithm used, since the feasible set of the optimum solution k_{opt} is finite and integer. Due to the computing capability of SA and ES, the complexity of $O(K^2)$ is reasonable for practical SATCOM scenarios. Once the value of k_{opt} is attained, the linear combined degree distribution $\Omega_{LC}(x)$ can be obtained. The optimization results of the code length and degree distribution for various channel conditions will be stored in several databases. As similarly as above, this optimization method of k_{opt} can also be applied in the situation where the channel lifespan follows a Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$. Although this approach is heuristic, the numerical results in Section 4 show that it could achieve a significant performance improvement over the conventional rateless codes.

4 Simulation and discussion

4.1 Symbol error rate and intermediate performance of proposed degree distribution

We will evaluate the symbol error rate SER and the intermediate performance z_{LC} of the linear combined degree distribution $\Omega_{LC}(x)$ obtained via the heuristic approach.

Firstly, without considering the erasure channel conditions, we give the simulation curves about the symbol error rate with respect to the decoding overhead at the receiver, as shown in Fig. 4. The code length is fixed to 2000, and the decoding overhead $\gamma = \frac{n}{k}$ is in the range of 0.7 to 1.3, so it includes full recovery and partial recovery situations. We use AND-OR tree analysis to describe the asymptotic decoding performance based on Eqs. (5) and (3), and use BP algorithm with Monte Carlo simulation time more than 5000 to describe the practical symbol error rates of the decoder. It can be found that the proposed degree distribution Ω_{LC} has much lower SER than

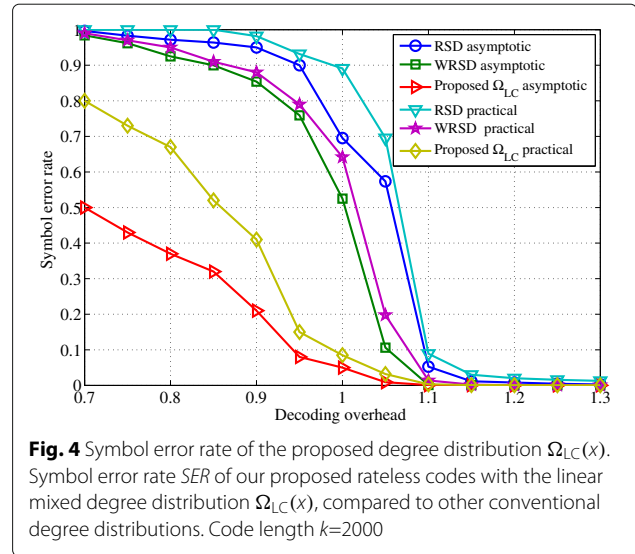
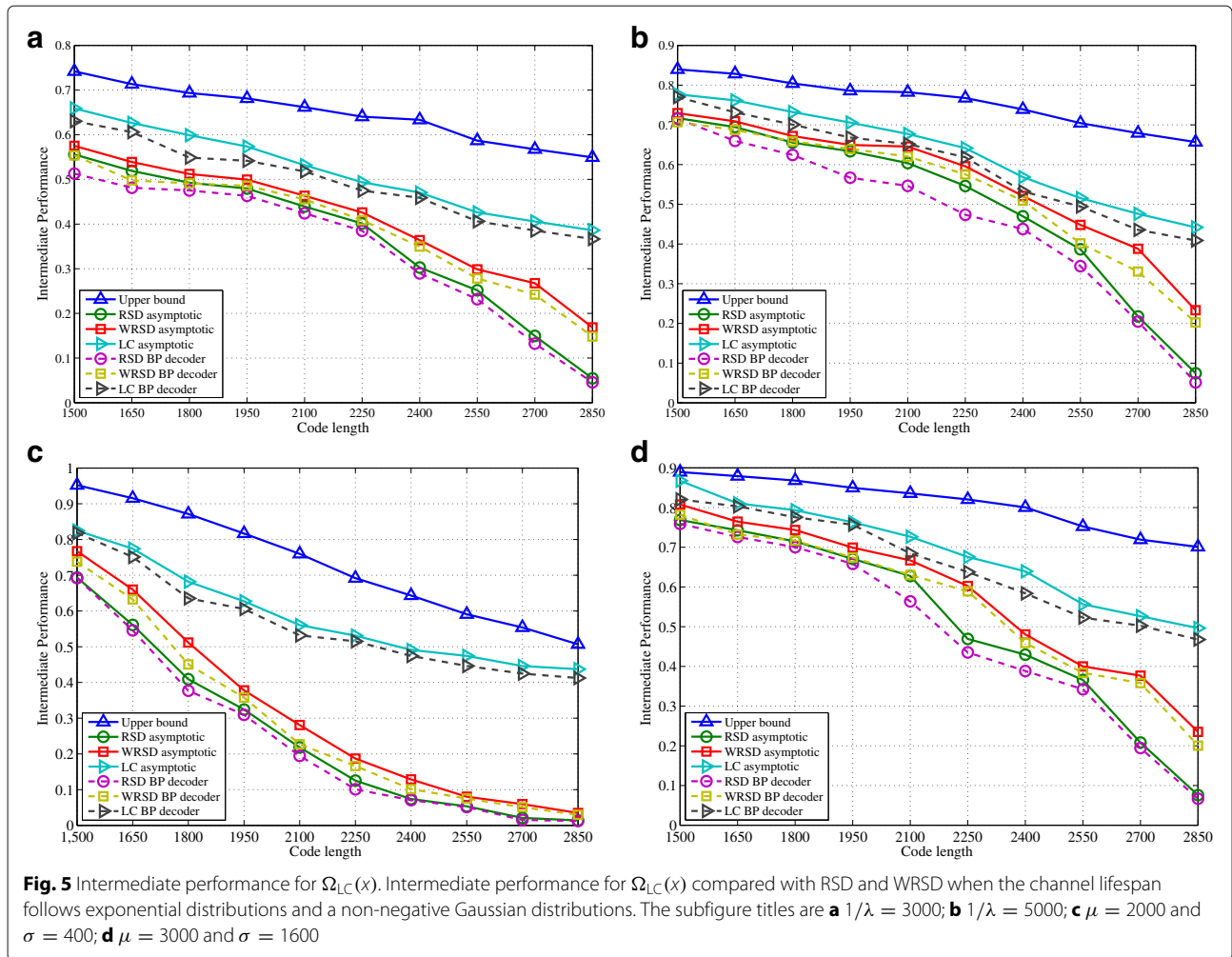


Fig. 4 Symbol error rate of the proposed degree distribution $\Omega_{LC}(x)$. Symbol error rate SER of our proposed rateless codes with the linear mixed degree distribution $\Omega_{LC}(x)$, compared to other conventional degree distributions. Code length $k=2000$

those of RSD and WRSD in the situation that $\gamma < 1$ and the same SER as that of WRSD in the situation that $\gamma > 1$. The reason is that the weighting coefficients of Ω_{LC} are properly designed corresponding to the different decoding overhead probability intervals. It proof that the SER of Ω_{LC} does not have the waterfall region at the point of decoding overhead slightly greater 1 as RSD and WRSD. This adapting property of Ω_{LC} is much suitable for the dynamic channel conditions and random channel lifespans.

Furthermore, in order to evaluate the intermediate performance z_{LC} of our designed degree distribution Ω_{LC} on SATCOM channel model, the parameters of the simulated MDEC are set as follows. The connective window length $K = 3000$. The channel lifespan L follows exponential distributions with means $1/\lambda = 3000$ and 5000 , or non-negative Gaussian distributions with means $\mu = 2000$ and 3000 , $\sigma = 400$ and 1600 . The two transition probabilities of the Markov channel are set $\lambda_G = 0.1$ and $\lambda_B = 0.2$, and the erasure probabilities of good and bad states are $p_G = 0.01$ and $p_B = 0.2$, respectively. The asymptotic intermediate performance in [19] is taken as the upper bound, which is calculated under the assumption that the decoding overhead γ in every window can be accurately estimated by the SA. We give the fixed range of the code length k from 1500 to 2850, so the $\Omega_{LC}(x)$ (named LC for short) with $\omega_1, \omega_2, \omega_3$ can be calculated by Eqs. (8) and (17).

As shown in Fig. 5, it can be observed that the BP decoding results by Monte Carlo simulation approximately match the asymptotic results obtained by the AND-OR tree analysis, which validates our previous inferences. The lower means of the channel lifespan shown in Fig. 5a, c indicate that the channel dies more frequently than the those with higher means in Fig. 5b, d, implying worse recovery ability for the rateless codes. As the code length k



increases, the intermediate performance gains of the three degree distributions all diminish obviously. The reason is that larger k results in smaller decoding overhead γ at the decoder, due to the fixed window length K . Particularly, the intermediate performance of $\Omega_{LC}(x)$ is significantly better than RSD and WRSD, approximately higher 17 to 20%, but still has a gap with the upper bound, attributed to its sub-optimality and the unknown decoding overheads at the decoder for each connective window of satellite downlink.

4.2 Average throughput of proposed transmission scheme

Next, we compare the throughput of the proposed coding scheme with that of the conventional rateless codes. MDEC lifespans with both exponential and non-negative Gaussian distributions are used to model the scenarios of energy exhaustion of the device and random attacks by radiation, respectively. We give three and four cases for the above two channel lifespan distributions with different means and variances.

As shown in Tables 1 and 2, the optimal code length k_{opt} and degree distribution Ω_{LC} are attained through solving the maximum problem (18). It shows that with an increasing mean, the optimal code length also increases, which explain the code redundancy decreases thanks to improved channel quality. Ω_{LC} under $\frac{1}{\lambda} = 3000$ has the highest proportion of degree-one nodes (nearly 50%). The channel may randomly die when its length is close to the window length K . As for the case of $\frac{1}{\lambda} = 10000$, the channel lifespan is mostly greater than the window length K , indicating that channel death rarely occurs. As a result, Ω_{LC} has only a slight higher proportion of degree-one nodes than that of WRSD to recover as many

Table 1 Optimal code length k_{opt} and weighting vector ω over the MDEC with an exponentially distributed channel lifespan

Case	$1/\lambda$	k_{opt}	$\omega = [\omega_1, \omega_2, \omega_3]$	d_{avg}
1	3000	2050	0.4006, 0.0555, 0.5439	5.58
2	5000	2280	0.2894, 0.0456, 0.665	6.59
3	10000	2550	0.1725, 0.0303, 0.7972	7.68

Table 2 Optimal code length k_{opt} and weighting vector ω over the MDEC with a non-negative Gaussian distributed channel lifespan

Case	μ	σ	k_{opt}	$\omega = [\omega_1, \omega_2, \omega_3]$	d_{avg}
1	2000	400	2850	0.6241, 0.2584, 0.1176	2.24
2	3000	400	2150	0.0005, 0.0062, 0.9933	9.29
3	3000	1600	2100	0.1905, 0.0515, 0.758	7.37
4	5000	1600	2250	0.0207, 0.0114, 0.9679	9.08

input symbols as possible. Similar results can be acquired for the channel lifespan with a Gaussian distributions. In the case of $\mu = 2000$ and $\sigma = 400$, most channel lengths are less than K , so lower degrees nodes (degree-one or degree-two) in Ω_{LC} dominate greater than

90%. The code length in this case is also larger, indicating that the transmitter sends the majority of input symbols un-coded to resist the severe channel conditions of MDEC. On the contrary, in the cases of $\mu = 3000, \sigma = 400$ and $\mu = 5000, \sigma = 1600$, the optimized results of Ω_{LC} are similar to those of the weak robust soliton distribution. This implies that a small variance or a large mean of the channel lifespan makes the channel condition steady and that Ω_{LC} adaptively evolves into the classical degree distribution with the optimal code length k_{opt} . Moreover, we also give the average degree d_{avg} of our optimized degree distributions in Tables 1 and 2 to show the linear encoding and decoding complexities. The d_{avg} can be calculated by the equation that $d_{avg} = \sum_{d=1}^D d\Omega_d$ of Ω_{LC} in Eq. (8)

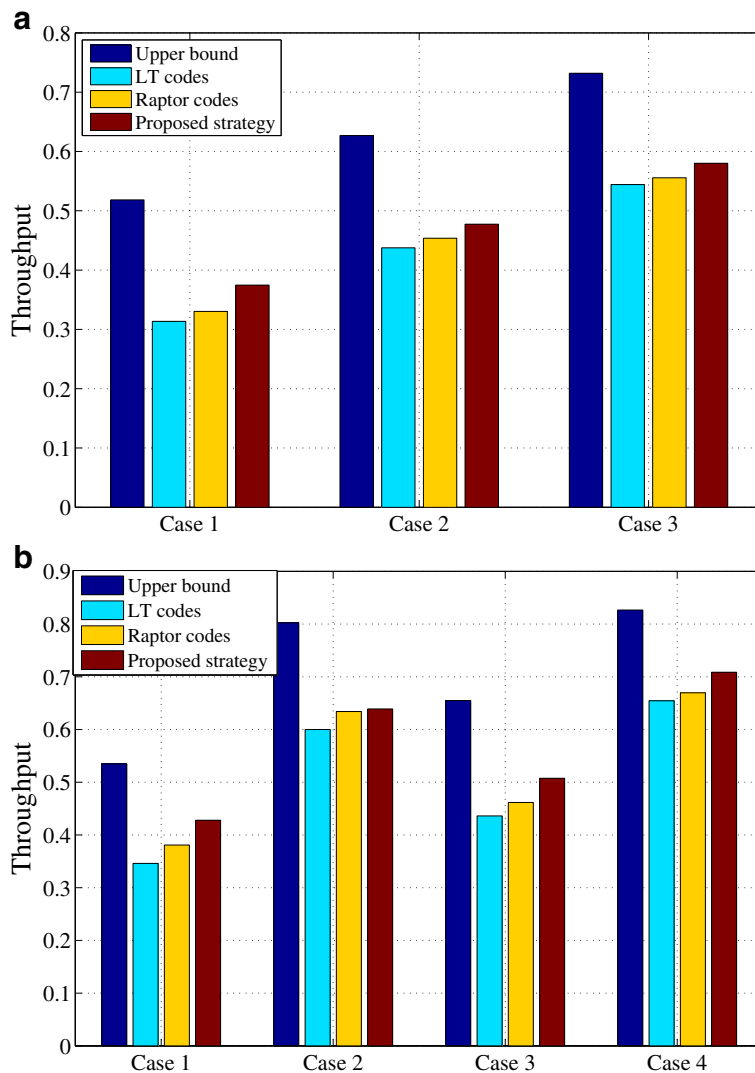


Fig. 6 Throughput performance of our proposed coding scheme compared with conventional LT and Raptor codes. Throughput over MDEC with exponential $L \sim E(\lambda)$ and non-negative Gaussian distributed $L \sim |\mathcal{N}(\mu, \sigma^2)|$ channel life spans. The subfigure titles are **a** Exponential distribution and **b** Non-negative Gaussian distribution

for different channel condition cases. It demonstrates that our proposed coding transmission is a practical technique for real-time delivery of large bulk data files in SATCOMs.

We give a SATCOM scenario that on a point-to-point MDEC between SA and ES with $N = 2000$ connective windows, and an infinite amount of data files to be transmitted. Figure 6a, b uses the bar charts to illustrate normalized throughput of the proposed rateless coding transmission under different cases of MDEC conditions in Tables 1 and 2, compared with the schemes of conventional LT and Raptor codes. For fair comparison, the code lengths of the LT and Raptor codes used in our simulations are also optimized by the results of the same maximum problem (10). Simulation results are based on the Eqs. (2) and (4), which demonstrate that our proposed coding transmission achieve better throughput performances compared to LT codes [13] and Raptor codes [14] with the channel randomly terminated by unpredictable attacks. When the channel conditions are getting worse, the proposed scheme has more significant outperformance. It is noted that there exists a performance gap between the upper bound and the proposed transmission schemes. Because the throughput of the upper bound is reachable only under a unrealistic assumption that the transmitter can accurately estimate the channel life span of every connective window.

5 Conclusions

In this paper, we considered the problem that large data file transmission over the connective-window and dynamic downlinks in SATCOM systems. Due to atmospheric precipitation and other emergency situations, such as energy exhausted and cosmic ray attacks, the channel state dramatically changes and unpredictably breaks off. A type of multi-state dying erasure channel model, named MDEC, is presented and characterized by a time-varying erasure probabilities and random channel lifespan. Moreover, a rateless coding transmission method with proper degree distribution and optimal code length was proposed to improve the average throughput and reduce the symbol error rate over MDEC. We firstly used a heuristic approach to design a weighted degree distribution based upon the AND-OR tree analysis technique to increase the intermediate performance. In addition, we have given the optimal code lengths by numerical analyses and calculations, to solve a maximize problem of throughput under random channel death probability distributions. We evaluated our transmission schemes by simulation based on different SATCOM channel conditions. The results demonstrated that our proposed rateless coding transmission attained bet-

ter intermediate performance and throughput than LT and Raptor codes with different channel state transitions and random distributions of unpredictable attacks.

Acknowledgements

This work has been supported in part by the National Natural Sciences Foundation of China (NSFC) under Grants 61525103, 61701136, 61771158, and 61371102, and Shenzhen Basic Research Program under Grants JCYJ 20160328163327348 and JCYJ 20150930150304185.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Communication Engineering Research Center, Harbin Institute of Technology-Shenzhen, The Xili University Town, Nanshan District, Shenzhen 518055, China. ²School of Electronics and Information Engineering, Harbin Institute of Technology, 92 West Dazhi Street, Nan Gang District, Harbin 150001, China.

Received: 4 July 2017 Accepted: 15 October 2017

Published online: 30 October 2017

References

1. C Niephaus, M Kretschmer, G Ghinea, Qos provisioning in converged satellite and terrestrial networks: a survey of the state-of-the-art. *IEEE Commun. Surv. Tutorials*. **18**(4), 2415–2441 (2016)
2. K Yu, C Rizo, D Burrage, AG Dempster, K Zhang, M Markgraf, An overview of GNSS remote sensing. *EURASIP J. Adv. Signal Process.* **2014**(1), 134 (2014)
3. SK Chronopoulos, C Koliopoulos, CT Angelis, in *ACM International Conference on Mobile Multimedia Communications (MobiMedia 07)*. Satellite multibeam signaling for multimedia services, (Nafpaktos, 2007), pp. 1–4
4. SK Chronopoulos, C Koliopoulos, A Pappa, CT Angelis, *Simulation of a Feasible Galileo System Operating in L1 and E5 Bands*. (Springer, Berlin Heidelberg, 2010), pp. 35–43
5. M Jia, X Gu, Q Guo, W Xiang, N Zhang, Broadband hybrid satellite-terrestrial communication systems based on cognitive radio toward 5g. *IEEE Wireless Commun.* **23**(6), 96–106 (2016)
6. G Araniti, I Bisio, MD Sanctis, A Orsino, J Cosmas, Multimedia content delivery for emerging 5g-satellite networks. *IEEE Trans. Broadcast.* **62**(1), 10–23 (2016)
7. M Jia, X Liu, X Gu, Q Guo, Joint cooperative spectrum sensing and channel selection optimization for satellite communication systems based on cognitive radio. *Int. J. Satell. Commun. Netw.* **35**(2), 139–150 (2017)
8. X Jia, T Lv, F He, H Huang, Collaborative data downloading by using inter-satellite links in leo satellite networks. *IEEE Trans. Wirel. Commun.* **16**(3), 1523–1532 (2017)
9. D Zhou, M Sheng, X Wang, C Xu, R Liu, J Li, Mission aware contact plan design in resource-limited small satellite networks. *IEEE Trans. Commun.* **65**(6), 2451–2466 (2017)
10. M Zeng, R Zhang, S Cui, in *IEEE Global Telecommunications Conference (GLOBECOM)*. On the outage capacity of a dying channel, (New Orleans, 2008), pp. 1–5
11. LR Varshney, SK Mitter, VK Goyal, An information-theoretic characterization of channels that die. *IEEE Trans. Inf. Theory*. **58**(9), 5711–5724 (2012)
12. M Zeng, R Zhang, S Cui, Outage capacity and optimal transmission for dying channels. *IEEE Trans. Commun.* **61**(1), 357–367 (2013)
13. M Luby, in *IEEE Symposium on Foundations of Computer Science*. LT codes, (Vancouver, 2002), pp. 271–280
14. A Shokrollahi, Raptor codes. *IEEE Trans. Inf. Theory*. **52**(6), 2551–2567 (2006)
15. MA Salkuyeh, F Hendessi, TA Gulliver, Data dissemination with rateless coding in a grid vehicular topology. *EURASIP J. Wireless Commun. Netw.* **2013**(1), 106 (2013)

16. Z Chen, L Yin, M Xu, J Lu, Layered multimedia broadcast using rateless codes with progressive recovery over cooperative mimo. *EURASIP J. Wireless Commun. Netw.* **2012**(1), 327 (2012)
17. MA Blanco, NV Burkhardt, C Chen, in *IEEE Military Communications Conference (MILCOM)*. Coding strategies for robust link blockage mitigation in satcom, (San Diego, 2013), pp. 611–616
18. P Cataldi, M Gerla, F Zampognaro, in *IEEE International Conference on Advances in Satellite and Space Communications (SPACOMM)*. Rateless codes for file transfer over dvb-s, (Colmar, 2009), pp. 7–12
19. S Sanghavi, in *IEEE Information Theory Workshop (ITW)*. Intermediate performance of rateless codes, (Tahoe City, 2007), pp. 478–482
20. A Talari, N Rahnavard, On the intermediate symbol recovery rate of rateless codes. *IEEE Trans. Commun.* **60**(5), 1237–1242 (2012)
21. M Zeng, R Calderbank, S Cui, On design of rateless codes over dying binary erasure channel. *IEEE Trans. Commun.* **60**(4), 889–894 (2012)
22. N Thomos, R Pulikkoonattu, P Frossard, Growth codes: intermediate performance analysis and application to video. *IEEE Trans. Commun.* **61**(11), 4710–4721 (2013)
23. Y Cassuto, A Shokrollahi, Online fountain codes with low overhead. *IEEE Trans. Inf. Theory.* **61**(6), 3137–3149 (2015)
24. MG Luby, M Mitzenmacher, MA Shokrollahi, in *ACM Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. Analysis of random processes via and-or tree evaluation, (San Francisco, 1998), pp. 364–373
25. SF Wang, Analytical multi-soliton solutions of a (2+1)-dimensional breaking soliton equation. *Springerplus.* **5**(1), 891 (2016)
26. G Ebadi, EV Krishnan, S Johnson, A Biswas, Cnoidal wave, snoidal wave, and soliton solutions of the d(m,n) equation. *Arab. J. Math.* **2**(1), 19–31 (2013)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
