SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2017, Vol. 60 102301:1–102301:15
doi: 10.1007/s11432-016-9011-8

Fast degree-distribution optimization for BATS codes

Huakai ZHAO[1], Shenghao YANG[2*] & Guinian FENG[3]

[1]*Beijing Satellite Navigation Center, Beijing* 100094*, China;*
[2]*The Chinese University of Hong Kong, Shenzhen* 518172*, China;*
[3]*Beijing Institute of Tracking and Communication Technology, Beijing* 100094*, China*

Received September 26, 2016; accepted November 30, 2016; published online July 4, 2017

Abstract Batched sparse (BATS) codes have been proposed for communication through networks with packet loss. BATS codes include a matrix generalization of fountain codes as the outer code and random linear network coding at the intermediate network nodes as the inner code. BATS codes, however, do not possess a universal degree distribution that achieves an optimal rate for any distribution of the transfer matrix ranks. Therefore, it is important to have a fast degree-distribution optimization approach for finite-length BATS codes. In this paper, we propose the concept of batch release probability (BRP), and demonstrate some characteristics of BRPs from the degree distributions achieving nearly optimal performance. Based on these BRP characteristics, we propose a novel degree-distribution optimization approach that achieves the similar decoding performance with a much shorter optimization time, compared with the previous approach. Moreover, the universality of BRPs observed in this paper can further simplify the degree-distribution optimization of BATS codes.

Keywords network coding, batched sparse (BATS) code, finite-length analysis, degree distribution, batch release probability, optimization

Citation Zhao H K, Yang S H, Feng G N. Fast degree-distribution optimization for BATS codes. Sci China Inf Sci, 2017, 60(10): 102301, doi: 10.1007/s11432-016-9011-8

1 Introduction

For communication through networks with packet loss, a batched sparse (BATS) code consists of an outer code and an inner code [1, 2]. The outer code, a matrix generalization of fountain codes, can potentially generate an unlimited number of batches, each of which consists of M coded symbols. When $M = 1$, the outer code turns into a Luby transform (LT) code [3] (or raptor code [4], if precoding is applied). The inner code is random linear network coding [5, 6] at the intermediate network nodes, which is allowed only for the symbols belonging to the same batch, so that the network coding does not change the batch degree. The end-to-end transmission of a batch from the source node to a destination node can be modeled by a (batch) transfer matrix, which reflects the network coding at the intermediate network nodes and the packet loss at the network links.

BATS codes preserve salient features of fountain codes especially their rateless property and low encoding/decoding complexity. Compared with the ordinary random linear network coding schemes, BATS codes have a lower encoding/decoding complexity, as well as a smaller coefficient vector overhead and

* Corresponding author (email: shyang@cuhk.edu.cn)

intermediate node caching requirement. Compared with other low-complexity random linear network coding schemes like expander chunked (EC) codes [7], Gamma codes [8] and L-chunked codes [9], BATS codes generally achieve higher rates and have the extra feature that an unlimited number of batches can be generated. Application of BATS codes in various network communication scenarios have been studied in [10–12]. BATS code based network transmission protocols (BATS protocols) is proposed in [13].

Different from LT codes, however, BATS codes do not possess a universal degree distribution that achieves the optimal rate for any distribution of the transfer matrix ranks (also called the rank distribution), which captures the effect of the network characteristics (e.g., packet losses, network topology and packet scheduling) on the batches. To maximize the achievable rate, the degree distribution of a BATS code can be optimized in terms of the rank distribution. Therefore, how to fast optimize a degree distribution for various rank distributions is an important BATS code design issue, especially for applications where the rank distribution is difficult to predict and can only be measured online.

Two approaches exist for optimizing the degree distribution of a BATS code. Based on the asymptotic performance evaluation of BATS codes with belief propagation (BP) decoding, the degree distribution can be obtained by solving a linear optimization for a given rank distribution [2]. The obtained degree distribution achieves a nearly optimal rate when the number of input symbols is large (e.g., hundreds of thousands); however, the performance of such a degree distribution for a relatively small number of input symbols is poor.

The performance of BATS codes with a relatively small number of input symbols has practical importance. Based on the finite-length analysis of BATS codes proposed in [14], a greedy approach of finite-length degree-distribution optimization is proposed. The greedy approach evaluates the finite-length performance of a set of degree distributions and outputs the best one. Though the greedy approach has demonstrated significant performance gain compared with the asymptotic approach, the long runtime prevents it from being used in applications that require online degree-distribution optimization. The greedy approach does not use any prior knowledge about good (or nearly optimal) degree distributions, so that the search space is large.

In this paper, we propose a novel BRP-based degree-distribution optimization approach that can benefit from certain prior knowledge on good degree distribution and achieves the similar decoding performance with a much shorter optimization time, compared with the greedy approach. We define the concept of batch release probability (BRP) for capturing the characteristics of the good degree distributions. Based on the analysis of the robust soliton distribution for LT codes and the evaluation of the degree distributions obtained using the greedy approach, we demonstrate some characteristics of BRPs that can be simply described using a pair of real numbers (v_1, v_2) with $v_1 + v_2 < 1$ and $v_1, v_2 > 0$.

Moreover, we also observe certain universal property of BRPs in our experiments. Particularly, let (v_1^*, v_2^*) be the pair (v_1, v_2) optimized for certain rank distribution. For a randomly generated rank distribution $\boldsymbol{h}$, we obtain a degree distribution $\boldsymbol{\Psi}$ such that the BRP of $\boldsymbol{\Psi}$ and $\boldsymbol{h}$ is approximately characterized by (v_1^*, v_2^*). We find for a large fraction of the rank distributions, this degree distribution demonstrates nearly optimal performance. This universal property of BRPs can further reduce the computation cost of our BRP-based degree-distribution optimization approach.

2 BATS codes

2.1 Encoding and transmission

A finite field $\mathbb{F}_q$ is fixed with size q, called the base field. Let $K, n, M > 0$ be integers. A BATS code with K input symbols of the base field includes a sequence of n batches, $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$, formed by

$$\boldsymbol{X}_i = \boldsymbol{B}_i \boldsymbol{G}_i, \tag{1}$$

where $\boldsymbol{B}_i$ is a row vector consisting of d_i input symbols, and $\boldsymbol{G}_i$ is a $d_i \times M$ totally random matrix called the generator matrix. We call d_i the (batch) degree and M the (batch) size. The degree $d_i, i = 1, \ldots, n$

are i.i.d. random variables with a given distribution $\boldsymbol{\Psi} = (\Psi_1, \ldots, \Psi_K)$, i.e., $\Pr\{d_i = k\} = \Psi_k$. The distribution $\boldsymbol{\Psi}$ is called degree distribution and is the parameter of BATS codes to be studied in this paper.

The batches are transmitted through a network where intermediate nodes perform linear network coding only among the symbols belonging to the same batch. So the received symbols of i-th batch can be represented by a row vector,

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{H}_i = \boldsymbol{B}_i\boldsymbol{G}_i\boldsymbol{H}_i, \tag{2}$$

where $\boldsymbol{H}_i$ is an M-row random matrix called the (batch) transfer matrix. The number of columns of $\boldsymbol{H}_i$ corresponds to the number of symbols received for i-th batch, which may vary for different batches and is finite. Let $r_i = \mathrm{rk}(\boldsymbol{H}_i)$. We know that $r_i \leqslant M$.

2.2 GE-BP decoding

We describe the Gaussian-elimination belief-propagation (GE-BP) decoding algorithm of BATS codes provided in [2]. The input of the GE-BP decoding is the sequence $(\boldsymbol{Y}_i, \boldsymbol{G}_i\boldsymbol{H}_i)$, $i = 1, \ldots, n$. In other words, each batch is associated with a linear system of equations given in (2), where the input symbols in $\boldsymbol{B}_i$ are the variables to solve. The decoder knows the indices of the input symbols involved in each batch. A batch with degree d, generator matrix $\boldsymbol{G}$ and transfer matrix $\boldsymbol{H}$ is said to be decodable if $\mathrm{rk}(\boldsymbol{GH}) = d$, i.e., the associated linear equation of the batch is uniquely solvable, and an input symbol is decodable if it is involved in a decodable batch.

The GE-BP decoding has multiple steps. Suppose the step index starts at 0. For each step, a decodable input symbol b is selected, substituted into the undecodable batches that it is involved in, and marked as recovered. Consider an undecodable batch that involves b with degree d, generator matrix $\boldsymbol{G}$ and transfer matrix $\boldsymbol{H}$. The substitution will remove one row of $\boldsymbol{G}$ and reduce d by one, which may make the batch decodable and hence generate new decodable input symbols. The decoding stops when there are no decodable input symbols. The GE-BP decoding algorithm uses a given number n of batches, and is denoted by $\mathrm{BP}(n)$.

2.3 Solvability of a batch

The performance of $\mathrm{BP}(n)$ has been analyzed when n (and K) is a finite number (finite-length analysis) in [14]. Let us introduce some basic steps of these analysis that are useful in this paper. Assume that $\boldsymbol{H}_1, \boldsymbol{H}_2, \ldots, \boldsymbol{H}_n$ are independent and follow the same distribution of a random matrix $\boldsymbol{H}$.

Let us check the probability that a batch is decodable when its degree has a specific value. According to the decoding algorithm of $\mathrm{BP}(n)$, if a batch is decodable at step t, it is decodable at all steps $t' > t$ until the associated linear system has no variable left. We say a batch is decodable for the first time at step t if it is decodable at step t, but is not decodable at step $t-1$.

For $s = 0, 1, \ldots, M$, let $\boldsymbol{G}^{(s)}$ be an $s \times M$ totally random matrix over the base field $\mathbb{F}_q$. Define

$$\hbar_s \triangleq \Pr\left\{\mathrm{rk}\left(\begin{bmatrix}\boldsymbol{G}^{(1)}\\ \boldsymbol{G}^{(s)}\end{bmatrix}\boldsymbol{H}\right) = \boldsymbol{rk}(\boldsymbol{G}^{(s)}\boldsymbol{H}) = s\right\}, \tag{3}$$

$$\hbar'_s \triangleq \Pr\{\mathrm{rk}(\boldsymbol{G}^{(s)}\boldsymbol{H}) = s\}, \tag{4}$$

where $\boldsymbol{G}^{(1)}$ and $\boldsymbol{G}^{(s)}$ are statistically independent. We see that $\hbar_s$ is the probability that a batch is decodable for the first time when it involves s input symbols. Once a batch becomes decodable, it remains to be decodable until all input symbols it involves in are decoded. Note that $\hbar'_s = \sum_{k\geqslant s}\hbar_k$ for $0 \leqslant s \leqslant M$ and $\hbar_s = 0$ for $s > M$. As it was characterized in [2],

$$\hbar_s = \sum_{k=s}^{M}\frac{\zeta_s^k}{q^{k-s}}h_k \quad\text{and}\quad \hbar'_s = \sum_{k=s}^{M}\zeta_s^k h_k, \tag{5}$$

where

$$\zeta_s^k \triangleq \begin{cases} (1-q^{-k})(1-q^{-k+1})\cdots(1-q^{-k+s-1}), & s>0,\\ 1, & s=0. \end{cases} \tag{6}$$

Let $h_k \triangleq \Pr\{\mathrm{rk}(\boldsymbol{H})=k\}$, we call $\boldsymbol{h}=(h_0,\ldots,h_M)$ the rank distribution of $\boldsymbol{H}$.

For BP(n), we are interested in the step when the decoding stops, which is equal to the number of input symbols that are decoded. If BP(n) stops at step K, all the input symbols are decoded. Define $P_{\mathrm{err}}(n)$ as the probability that at least one input symbol is not decoded when decoding stops. In [14], a recursive formula is provided to calculate $P_{\mathrm{err}}(n)$. In this paper, we study how to optimize the degree distribution using $P_{\mathrm{err}}(n)$ as the performance measure.

2.4 Greedy approach for degree-distribution optimization

Let us first introduce the greedy approach of finite-length degree-distribution optimization in [14], which has the following two steps with an initial degree distribution $\boldsymbol{\Psi}^{(0)}$ (which can be trivial).

(i) Find one or multiple new degree distribution which may be potentially better than $\boldsymbol{\Psi}^{(0)}$.

(ii) Evaluate the performance of these new degree distributions in terms of an objective function (e.g., $P_{\mathrm{err}}(n)$), and select the degree distribution that outperforms $\boldsymbol{\Psi}^{(0)}$ the most.

These two steps can be applied repeatedly.

For the degree-distribution optimization of finite-length BATS codes in [14], the degree distribution obtained from the asymptotic analysis of BATS codes is used as the initial degree distribution $\boldsymbol{\Psi}^{(0)}$. In the first step, a new degree distribution $\boldsymbol{\Psi}^{(1)}$ is obtained by perturbing the degree distribution $\boldsymbol{\Psi}^{(0)}$ at certain degree d so that

$$\boldsymbol{\Psi}^{(1)} = (\boldsymbol{\Psi}^{(0)} + \delta \boldsymbol{e}_{d-1})/(1+\delta),$$

where δ is a real number and $\boldsymbol{e}_{d-1}$ is the all-zero vector except that the $(d-1)$-th component is 1. In the second step, the performances of $\boldsymbol{\Psi}^{(1)}$ and $\boldsymbol{\Psi}^{(0)}$ are compared based on the finite-length results of BATS codes and the better degree distribution is selected as $\boldsymbol{\Psi}^{(0)}$. These two steps are repeated for a number of iterations and output $\boldsymbol{\Psi}^{(0)}$.

Technically, the greedy approach can find a nearly optimal degree distribution if we use a sufficiently small δ and repeat the algorithm for a large number of iterations. In our experience, the greedy approach converges when the number of iterations is in the range 2000–5000 for $K=256$ and $M=16$. The greedy approach does not use any prior characteristics of the good degree distributions and therefore the number of degree distributions evaluation is large, which results in a long runtime.

3 Batch release probability

3.1 Definition of BRP

We say that a batch is released at step t $(1 \leqslant t \leqslant K-1)$ if it is decodable for the first time at step t and involves at least one decodable input symbol. At step 0, a decodable batch is also released. We use $r(t,d)$ to denote the probability that a batch with degree d is released at step t, which is called the degree release probability.

Proposition 1. (Degree release probability formula)

- For $t=0$ and $d \leqslant M$, $r(0,d)=\hbar'_d$.
- For $t=1,\ldots,K-1$ and $d=2,\ldots,\min\{K,M+t\}$,

$$r(t,d) = \sum_{s=\max\{1,d-t\}}^{\min\{M,K-t,d-1\}} \hbar_s \frac{\binom{t-1}{d-s-1}\binom{K-t}{s}}{\binom{K}{d}}. \tag{7}$$

- For all other t and d, $r(t,d)=0$.

Proof. At step 0, only a batch with degree $d \leqslant M$ can be decodable as well as released, and $r(0,d) = \Pr\{\mathrm{rk}(\boldsymbol{G}^{(d)}\boldsymbol{H}) = d\} = \hbar'_d$ (see Eq. (4)).

For $t \geqslant 1$, only a batch with degree $d \geqslant 2$ and $d \leqslant t + M$ can be released. For a batch that is released at step t, the number of input symbols involved in the batch is between 1 and M. Fix a batch with degree d, the degree released probability at step t can be decomposed into

$$\begin{aligned} r(t,d) &= \sum_{s=1}^{M} \Pr(\text{the batch is released at step } t \mid \text{the batch involves } s \text{ input symbols at step } t) \\ &\quad \times \Pr(\text{the batch involves } s \text{ input symbols at step } t) \\ &= \sum_{s=1}^{M} \Pr(\text{the batch is decodable for the first time when it involves } s \text{ input symbol at step } t) \\ &\quad \times \Pr(\text{the batch involves } s \text{ input symbols at step } t) \\ &= \sum_{s=\max\{1,d-t\}}^{\min\{M,K-t\}} \hbar_s \times \frac{\binom{t-1}{d-s-1}\binom{K-t}{s}}{\binom{K}{d}}, \end{aligned}$$

where the event that the batch involves s input symbols at step t has a hypergeometric distribution because the input symbols are randomly selected when decoding. Hence this event can be expressed as: $d - s - 1$ input symbols involved in the batch are recovered before step t, one input symbol involved in the batch is recovered at step t and the remaining s input symbols involved in the batch are among the $K - t$ unrecovered input symbols, which imply $d - t \leqslant s \leqslant \min\{K - t, d - 1\}$.

Lemma 1. For all degree d, $\sum_{t=0}^{K-1} r(t,d) = 1 - \hbar_0$.

Proof. For $d = 1$, $\sum_{t=0}^{K-1} r(t,d) = r(0,1) = \hbar'_1 = 1 - \hbar_0$. For $d > 1$,

$$\begin{aligned} \sum_{t=0}^{K-1} r(t,d) &= r(0,d) + \sum_{t=1}^{K-1} \sum_{s=\max\{1,d-t\}}^{\min\{M,K-t,d-1\}} \hbar_s \frac{\binom{t-1}{d-s-1}\binom{K-t}{s}}{\binom{K}{d}} \\ &= r(0,d) + \sum_{s=1}^{\min\{M,d-1\}} \hbar_s \sum_{t=\max\{1,d-s\}}^{\min\{K-1,K-s\}} \frac{\binom{t-1}{d-s-1}\binom{K-t}{s}}{\binom{K}{d}} = r(0,d) + \sum_{s=1}^{\min\{M,d-1\}} \hbar_s, \end{aligned} \tag{8}$$

if $1 < d \leqslant M$, Eq. (8) equals $\hbar'_d + \sum_{s=1}^{d-1} \hbar_s = \hbar'_1 = 1 - \hbar_0$. And if $d > M$, Eq. (8) equals $\sum_{s=1}^{M} \hbar_s = \hbar'_1 = 1 - \hbar_0$. Hence for all degree d, $\sum_{t=0}^{K-1} r(t,d) = 1 - \hbar_0$.

Definition 1 (Batch release probability, BRP)**.** Let $r(t,\boldsymbol{\Psi}) = \sum_{d\geqslant 1} \Psi_d r(t,d)$, called the batch release probability (BRP) at step t. Let $R(\boldsymbol{\Psi}) = (r(0,\boldsymbol{\Psi}), \ldots, r(K-1,\boldsymbol{\Psi}))$, which represent the BRP at each step and is called the BRP distribution. To specify the dependence of the BPR on the rank distribution $\boldsymbol{h}$, we also write $R(\boldsymbol{\Psi})$ as $R(\boldsymbol{\Psi},\boldsymbol{h})$.

Note that $\sum_{t=0}^{K-1} r(t,\boldsymbol{\Psi})$ is $1 - \hbar_0$ instead of 1, where the missing $\hbar_0$ is the probability that a batch is decodable for the first time when all its input symbols are previously decodable. In general, for a vector $\boldsymbol{L}$ of l entries, we denote by $\boldsymbol{L}[k]$ its kth entry and the starting index is 0. We say a length-l vector $\boldsymbol{L}$ a distribution vector if $\boldsymbol{L}[i] \geqslant 0$ and $\sum_{i=0}^{l-1} \boldsymbol{L}[i] > 0$. Note that we do not require that $\sum_{i=0}^{l-1} \boldsymbol{L}[i] = 1$ since, if desired, we can simply normalize the vector. Henceforth, when we say $\boldsymbol{\Psi} = (\Psi_1, \ldots, \Psi_K)$ is a degree distribution, it is possible that $\sum_d \Psi_d \neq 1$. The definition of BRP in Definition 1 can be extended to this case.

3.2 Linearity of BRP

As we will see in this paper, for the degree distributions that demonstrate good GE-BP decoding performance, the corresponding BRP distributions have a similar form. This observation will be used in our degree-distribution optimization approach, where one of the important steps is to obtain a degree distribution $\boldsymbol{\Psi}$ such that $R(\boldsymbol{\Psi})$ is (approximately) a given distribution vector. Note that it is possible that for a distribution vector $\boldsymbol{r}$, e.g., $\boldsymbol{r} = \boldsymbol{e}_{10}$, there exists no degree distribution $\boldsymbol{\Psi}$ such that $R(\boldsymbol{\Psi}) = \boldsymbol{r}$.

Algorithm 1 Transformation from a distribution vector to a degree distribution

Input:
- $K, N, M, \boldsymbol{h}$ are the number of input symbols, number of batches, batch size, and rank distribution, respectively.
- $\boldsymbol{L}$ is a length-K distribution vector.
- E_1, E_2 are two threshold parameters.

Output:

1: **for** $d = K$ to 1 **do**
2:     $\Psi_d \Leftarrow 1$
3:     **for** $t = 0$ to $K-1$ **do**
4:         **if** $r(t,d) \geqslant E_1$ and $\boldsymbol{L}[t] \geqslant E_1$ **then**
5:             $\Psi_d \Leftarrow \min\{\Psi_d, \boldsymbol{L}[t]/r(t,d)\}$
6:         **end if**
7:     **end for**
8:     **if** $\max_t(\Psi_d r(t,d) - \boldsymbol{L}[t]) \geqslant E_2$ **then**
9:         $\Psi_d \Leftarrow 0$
10:    **else**
11:         $\boldsymbol{L} \Leftarrow \boldsymbol{L} - \Psi_d \boldsymbol{V}_d$
12:    **end if**
13: **end for**
14: **return** $\boldsymbol{\Psi} = (\Psi_1, \ldots, \Psi_k)$

Before giving the transformation from a BRP distribution to a degree distribution, let us discuss a linearity property of BRP distributions. We can rewrite the BRP distribution $R(\boldsymbol{\Psi})$ as follows:

$$\begin{aligned} R(\boldsymbol{\Psi}) &= (r(0,\boldsymbol{\Psi}), \ldots, r(K-1,\boldsymbol{\Psi})) = \left(\sum_{d=1}^{K} \Psi_d r(0,d), \ldots, \sum_{d=1}^{K} \Psi_d r(K-1,d)\right) \\ &= \sum_{d=1}^{K} \Psi_d \left(r(0,d), \ldots, r(K-1,d)\right) = \sum_{d=1}^{K} \Psi_d \boldsymbol{V}_d, \end{aligned} \tag{9}$$

where $\boldsymbol{V}_d \triangleq (r(0,d), \ldots, r(K-1,d))$. Eq. (9) is a linear decomposition of BRP distribution and the coefficients of $\boldsymbol{V}_d$ $(d = 1, \ldots, K)$ constitute the corresponding degree distribution. Define a $K \times K$ matrix $\boldsymbol{V}$ as $(\boldsymbol{V}_1^{\mathrm{T}}, \ldots, \boldsymbol{V}_K^{\mathrm{T}})$, we have

$$R(\boldsymbol{\Psi}) = \boldsymbol{\Psi} \boldsymbol{V}^{\mathrm{T}}. \tag{10}$$

The properties of $\boldsymbol{V}_d$ are shown in Figure 1. From the sub-figures of the $\boldsymbol{V}_d$ with different values of d, we can see that (i) the peak value of $\boldsymbol{V}_d$ (i.e., $\max_t r(t,d)$) increases with degree d; (ii) the peak position (i.e., $\mathrm{argmax}_t\, r(t,d)$) increases with degree d.

3.3 Approximating a distribution vector using BRP

In this subsection, we introduce an approach to obtain a degree distribution $\boldsymbol{\Psi}$ such that $R(\boldsymbol{\Psi})$ approximates a length-K distribution vector $\boldsymbol{L}$. According to (10), the degree distribution can be obtained if matrix $\boldsymbol{V}$ is non-singular,

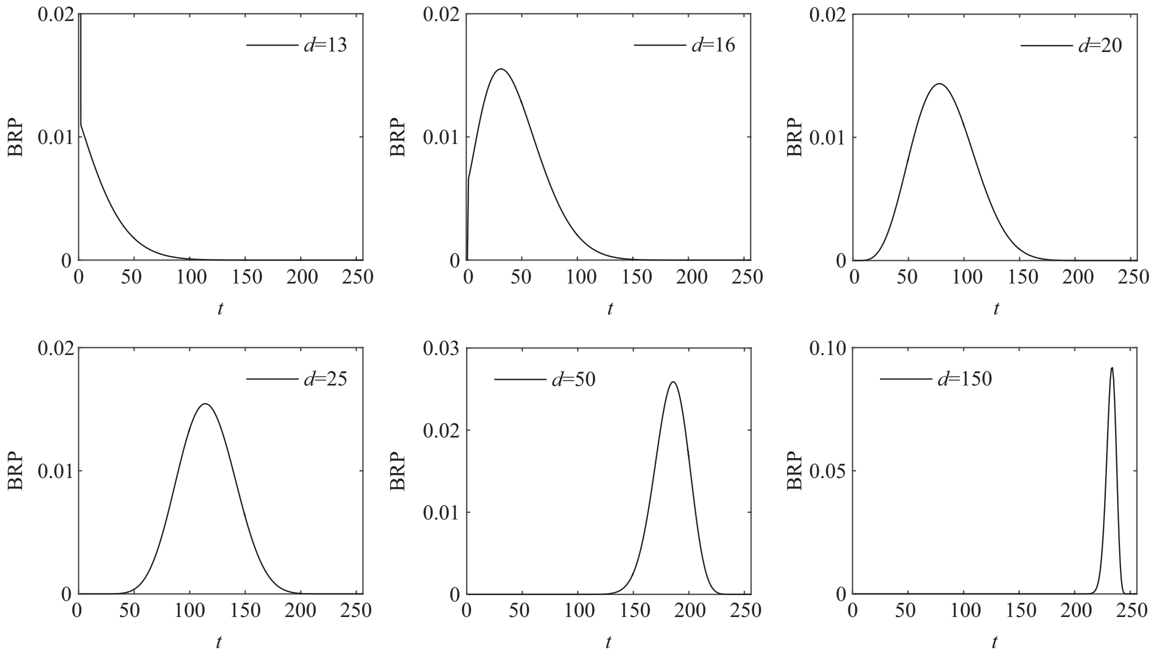
$$\boldsymbol{\Psi} = \boldsymbol{L} \boldsymbol{V}^{-\mathrm{T}}. \tag{11}$$

However, matrix $\boldsymbol{V}$ usually has a large or infinite condition number, which makes the numerical inversion infeasible. For our purpose, it is not necessary to obtain a degree distribution $\boldsymbol{\Psi}$ such that $R(\boldsymbol{\Psi}) = \boldsymbol{L}$, which may not exist; it is sufficient that $R(\boldsymbol{\Psi})$ is close enough to $\boldsymbol{L}$.

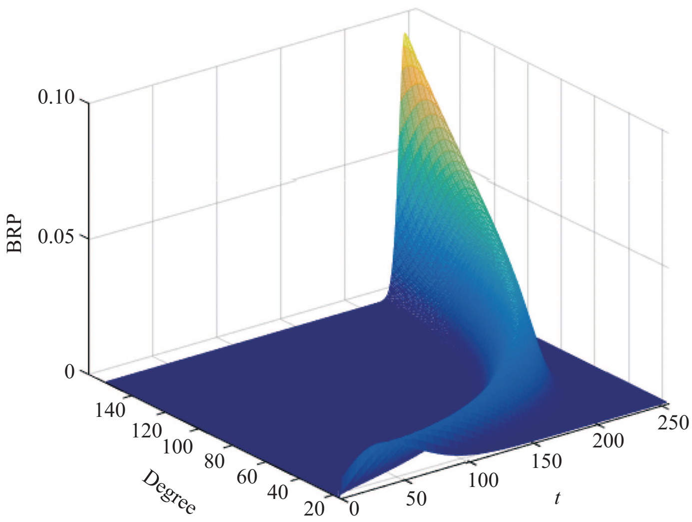
In view of the properties BRP distributions discussed in the previous subsection, our idea is to find a linear combination of $\boldsymbol{V}_d$ that approximates the given $\boldsymbol{L}$, where the combination coefficients Ψ_d of $\boldsymbol{V}_d$ constitute the desired $\boldsymbol{\Psi}$. Figure 2 shows an example of linear combination of $\boldsymbol{V}_d$ in ascending order of degree d, where the upper bound of the curves in Figure 2(d) is the given vector $\boldsymbol{L}$.

Particularly, we use Algorithm 1 to obtain such a degree distribution, where the searching of the linear combination coefficients Ψ_d of $\boldsymbol{V}_d$ is in the descending order of d. The pseudocodes from the 2nd line to the 7th line search a proper coefficient Ψ_d for the current $\boldsymbol{V}_d$. And the pseudocodes from the 8th line to the 12th line check if the coefficient Ψ_d is suitable by measuring the approximation error. The threshold E_1 is used to alleviate the effect of relatively small $r(t,d)$ or $\boldsymbol{L}(t)$ values, and the threshold E_2 is to control the approximation error. This algorithm is suitable for any length-K distribution vector $\boldsymbol{L}$. The algorithm guarantees that $\Psi_d \geqslant 0$.

Numerical results in Figure 3 illustrate the performance of Algorithm 1. In Figure 3(a), all the entries of the distribution vector $\boldsymbol{L}$ have almost the same value. There is no degree distribution $\boldsymbol{\Psi}$ such that $R(\boldsymbol{\Psi}) = \boldsymbol{L}$. The $\boldsymbol{L}$ in Figure 3(b) is the BRP distribution of a randomly generated degree distribution. We see that our algorithm can approximate this BRP distribution well.



(a)



(b)

Figure 1 (Color online) $\boldsymbol{V}_d$ with different degrees. (a) $d = 13, 16, 20, 25, 50, 150$, and $K = 256$; (b) $d = 20, \ldots, 150$, $K = 256$.

4 BRP-based approach for degree-distribution optimization

4.1 Heuristics from robust soliton distribution of LT codes

As a special case of BATS codes with $M = 1$, LT codes achieve very good BP decoding performance with the robust soliton distribution, denoted by $\boldsymbol{\Psi}^{\mathrm{RS}}$. For the general case with $M > 1$, however, we do not have a single degree distribution that can achieve good GE-BP decoding performance for all (or even a subset of) rank distributions. So it is not surprising that the robust soliton distribution cannot be used in BATS codes. But the study of the robust soliton distribution for LT codes would shed light the design of a good degree distribution for BATS codes.

A robust soliton distribution can be specified by two parameters (see [3] for the definition). $R(\boldsymbol{\Psi}^{\mathrm{RS}})$ is

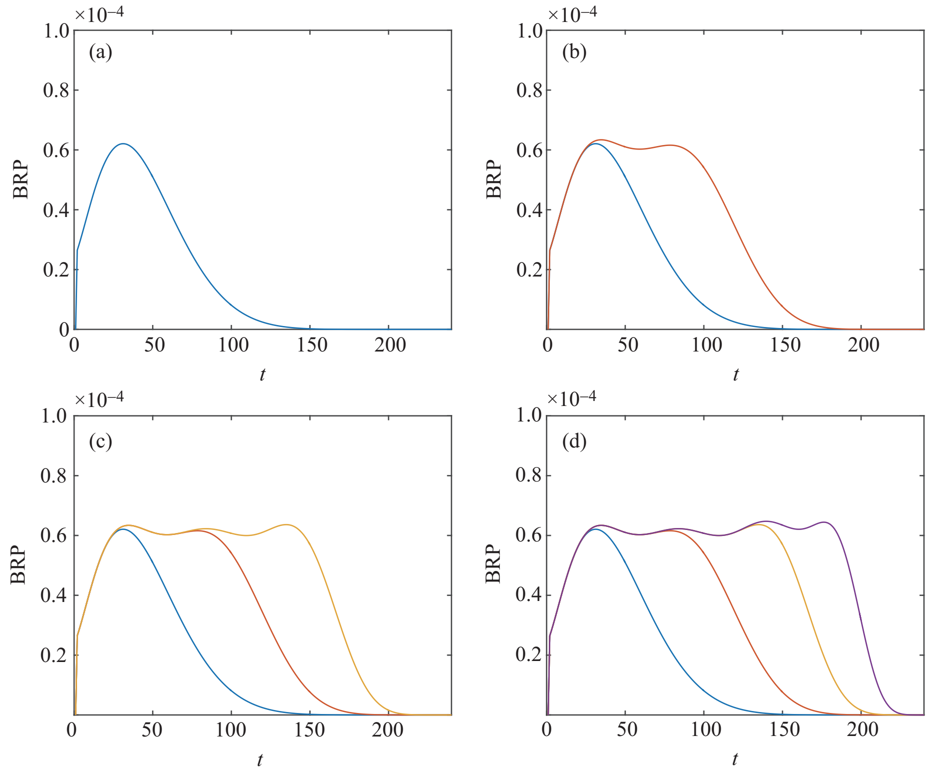


Figure 2 (Color online) Example of the linear combination of $\boldsymbol{V}_d$, (a) $d = 17$; (b) $d = 17, 23$; (c) $d = 17, 23, 33$; (d) $d = 17, 23, 33, 50$.

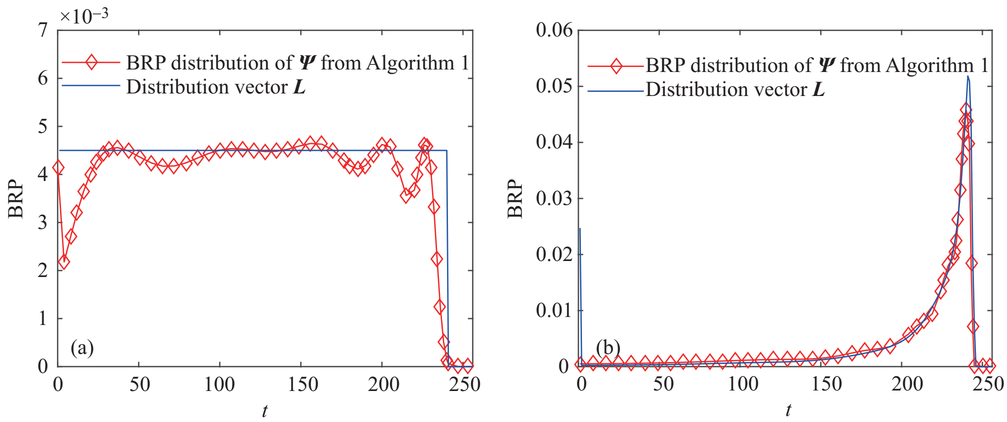


Figure 3 (Color online) Results of Algorithm 1 for two given $\boldsymbol{L}$. The curves without rhombus express the given $\boldsymbol{L}$. The rhombus curves are the BRP distributions $\boldsymbol{\Psi}$ of the degree distributions obtained from the Algorithm 1. (a) A horizontal $\boldsymbol{L}$; (b) BRP distribution of a random degree distribution.

the BRP distribution of the robust soliton distribution for LT codes [1)]. As shown in Figure 4, $R(\boldsymbol{\Psi}^{\mathrm{RS}})$ has the following characteristics: the BRP is high at step 0 and at the steps close to K, and is smooth and relatively low at other steps. When we tune the parameters of the robust soliton distribution in a feasible range, though BRP changes, these characteristics of BRP are preserved. We explain in the following paragraph how these characteristics of BRP affect the BP decoding performance of LT codes.

The BRP determines the number of released batches (though for LT codes, each batch has only one coded symbol) at each step, which may generate new decodable input symbols. Denote by R_t the number of the decodable input symbols at step t. Note that the BP decoding of LT codes stops when there exist

1) Our previous discussion can be applied to LT codes by letting $M = 1$, $\hbar_0 = 0$ and $\hbar_1 = 1$ (see [14]).

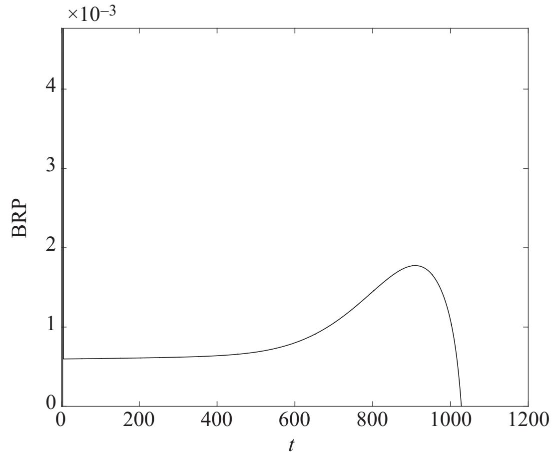


Figure 4 BRP distribution of the robust soliton distribution for the LT code with $K = 1024$.

no decodable input symbols. At step 0, the BRP is high to ensure a proper size of R_0. During each following decoding step, one input symbol is recovered from the set of decodable input symbols. The smooth and relatively low BRP can ensure that on average one previously undecodable input symbol become decodable at each step t until t reaches a value close to K, which can keep the number of decodable symbols roughly unchanged. Therefore, the number of decodable symbols R_t, $t = 0, 1, \ldots$, is approximately a random walk with the initial position R_0. We would like R_0 is large enough so that R_t is not zero with high probability before t is close to K, which accounts for the high BRP at step 0. When the decoding step is close to K, the input symbols involved in the released batches at step t are previously decodable with a high probability of $R_t/(K-t)$. Thus, a high BRP is needed to generate enough released symbols that all the undecodable symbols can become decodable with high probability.

Though there is not a single degree distribution with these characteristics for all rank distributions, these characteristics motivate us to use BRP for designing the degree distribution of BATS codes.

4.2 The approximate BRP vector and the corresponding degree distribution

Now let us check the BRP characteristics of the general BATS codes. Figure 5(c) and (d) are two BRP vectors of the degree distributions obtained from the greedy approach, which demonstrate similar characteristics of LT codes. Particularly, the BRP vector demonstrates two peaks, the first peak is at the step 0, the second peak is in a sequence of consecutive steps from $K-M$ to $K-1$. At other steps, the BRP vector is smooth and roughly flat. If we ignore the minor variations, the BRP vectors can be approximated by a bimodal vector $\boldsymbol{L}$ of '⊔' shape, illustrated in Figure 5(e).

For the vector $\boldsymbol{L}^*$ shown in Figure 5(e), our purpose is to obtain a degree distribution $\boldsymbol{\Psi}$ such that $R(\boldsymbol{\Psi})$ approximates $\boldsymbol{L}^*$. A simple approach is that first fix $\boldsymbol{L}^*$ and then obtain degree distribution $\boldsymbol{\Psi}$ by Algorithm 1. However, this approach needs $M+2$ parameters to formulate $\boldsymbol{L}^*$: one parameter formulates the BRP at step 0, one parameter formulates the BRP values from step 1 to step $K-M-1$, and M parameters formulate the BRP values from step $K-M$ to step $K-1$.

Here we give a simpler approach for obtaining a degree distribution $\boldsymbol{\Psi}$ that can approximate $\boldsymbol{L}^*$. First we separate $\boldsymbol{L}^*$ into two length-K vectors $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$, where

$$\boldsymbol{L}_1[t] = \begin{cases} \boldsymbol{L}^*[t], & 0 \leqslant t < K-M, \\ 0, & K-M \leqslant t, \end{cases} \qquad \boldsymbol{L}_2[t] = \begin{cases} 0, & 0 \leqslant t < K-M, \\ \boldsymbol{L}^*[t], & K-M \leqslant t. \end{cases}$$

$\boldsymbol{L}_1$ can be formulated by parameters v_1 and v_2,

$$\boldsymbol{L}_1[t] = \begin{cases} v_1, & t = 0, \\ \frac{1-v_1-v_2}{K-M-1}, & 1 \leqslant t < K-M, \\ 0, & t > K-M. \end{cases} \tag{12}$$

h_{12}	h_{13}	h_{14}	h_{15}	h_{16}
0.1	0.2	0.4	0.2	0.1

(a)

h_{10}	h_{11}	h_{12}	h_{13}	h_{14}	h_{15}	h_{16}
0.1	0.1	0.2	0.2	0.2	0.1	0.1

(b)

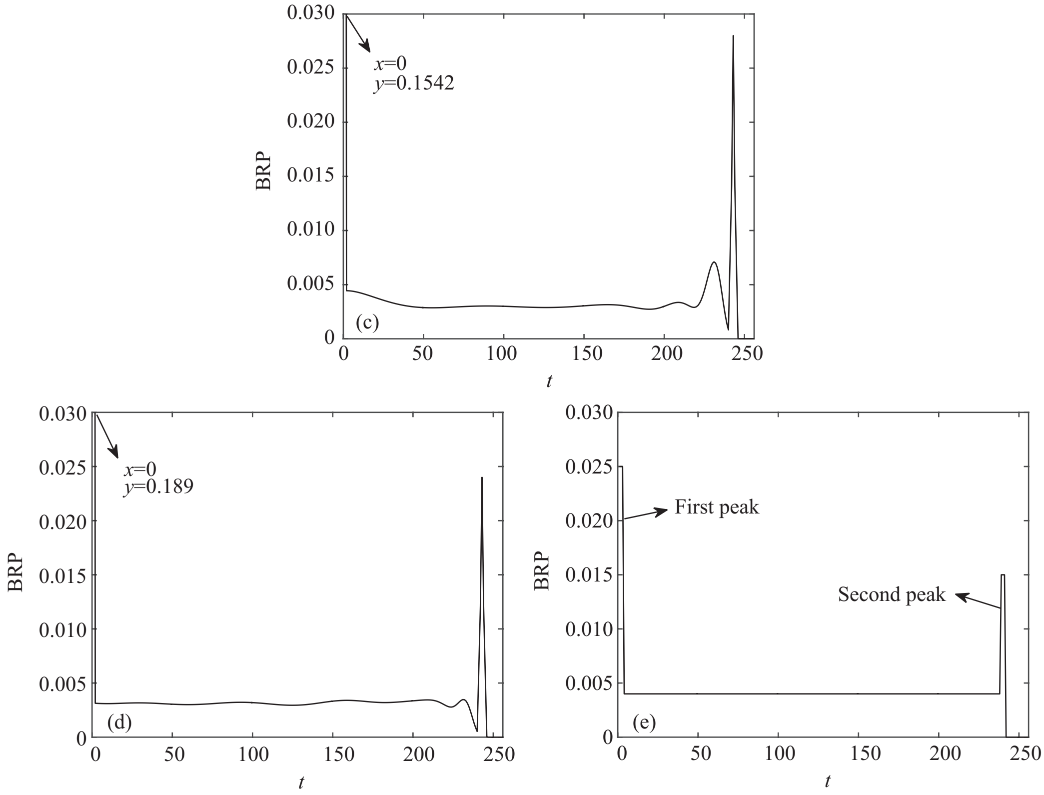


Figure 5 (a) Rank distribution $\boldsymbol{h}_1'$; (b) rank distribution $\boldsymbol{h}_2'$; (c) BRP distribution of the degree distribution obtained by the greedy approach for rank distribution $\boldsymbol{h}_1'$, $K = 256$, $M = 16$; (d) BRP distribution of the degree distribution obtained by the greedy approach for rank distribution $\boldsymbol{h}_2'$, $K = 256$, $M = 16$; (e) the approximate BRP vector of the degree distribution from the greedy approach.

Algorithm 2 Obtaining a degree distribution from v_1 and v_2

Input:
- $K, N, M, \boldsymbol{h}$ are the number of input symbols, number of batches, batch size, and rank distribution, respectively.
- v_1, v_2 are two positive real numbers that $v_1 + v_2 < 1$.

Output:

Initialize $\boldsymbol{L}_1$ using Equation (12)
Compute $\boldsymbol{\Psi}^1$ using $\boldsymbol{L}_1$ by Algorithm 1
Initialize $\boldsymbol{\Psi}^2$ using Equation (13)
$\boldsymbol{\Psi} \Leftarrow \boldsymbol{\Psi}^1 + \boldsymbol{\Psi}^2$
Normalize $\boldsymbol{\Psi}$
return $\boldsymbol{\Psi}$

Applying Algorithm 1 on $\boldsymbol{L}_1$, a degree distribution $\boldsymbol{\Psi}^1$ can be obtained.

Define $\boldsymbol{\Psi}^2$ as

$$\Psi_d^2 = \begin{cases} 0, & 1 \leqslant d < K, \\ v_2, & d = K. \end{cases} \tag{13}$$

Since a batch with degree K can not be released until step $K - M$, we have $r(t, K) = 0$ for $t = 0, 1, \ldots, K - M - 1$. Recalling the shape of $\boldsymbol{V}_K = (r(0, K), r(1, K), \ldots, r(K - 1, K))$ illustrated in Subsection 3.2, we can use $R(\boldsymbol{\Psi}^2) = \Psi_K^2 \boldsymbol{V}_K$ to approximate $\boldsymbol{L}_2$.

Let $\boldsymbol{\Psi} = \boldsymbol{\Psi}^1 + \boldsymbol{\Psi}^2$. $\boldsymbol{L}^*$ can be approximated by $R(\boldsymbol{\Psi})$. Here we summarize this process as an algorithm outlined in Algorithm 2 for obtaining a degree distribution using parameters v_1 and v_2.

Table 1 The rank distribution $\boldsymbol{h}_0$ for evaluation examples. $\boldsymbol{h}_0$ is the one of the length-2 homogeneous line network with link erasure probability 0.2 (see [2]). Here the BATS code has $q = 256$ and $M = 16$

h_5	h_6	h_7	h_8	h_9	h_{10}	h_{11}	h_{12}	h_{13}	h_{14}	h_{15}	h_{16}
0.0001	0.0004	0.0025	0.0110	0.0387	0.1040	0.2062	0.2797	0.2339	0.1038	0.0190	0.0008

4.3 BRP-based approach for degree-distribution optimization

We are now ready to present our fast degree-distribution optimization approach based on the BRP properties. The BRP-based approach uses Algorithm 2 and certain finite-length performance measure of BATS codes. The approach in general includes multiple iterations of the following four steps with an initial list $\mathcal{V}$ of (v_1, v_2) pairs:

(1) Obtain a degree distribution $\boldsymbol{\Psi}^{(v_1,v_2)}$ using Algorithm 2 for each (v_1, v_2) pair in the list $\mathcal{V}$.

(2) Evaluate the BATS code decoding performance of $\boldsymbol{\Psi}^{(v_1,v_2)}$ in terms of certain objective function for each (v_1, v_2) pair in the list $\mathcal{V}$.

(3) If a better degree distribution is found, mark the degree distribution as $\boldsymbol{\Psi}^*$ and the corresponding (v_1, v_2) pair as (v_1^*, v_2^*).

(4) Update list $\mathcal{V}$.

In this section, we use the error probability $P_{\rm err}(n)$ of GE-BP decoding as the objective function in Step (2), which can be evaluated using algorithms in [14]. Another objective function will be discussed in the next section. In Step (4), the list $\mathcal{V}$ can be updated using various approaches. For example, we can use $\{(v_1^* \pm S, v_2^* \pm S)\}$ as $\mathcal{V}$, excluding the pairs that have been previous evaluated, where S takes a predefined value, e.g., 0.02. If no better degree distribution can be found in Step (3), we decrease the value of S by half. The above iteration stops when S is small enough (e.g., 0.001) or the number of iterations exceeds a certain value (e.g., 20). The output of the algorithm is $\boldsymbol{\Psi}^*$ and (v_1^*, v_2^*).

4.4 Performance comparison with the greedy approach for GE-BP decoding

The computation cost of the BRP-based approach is mainly contributed by two parts. The first part is the calculation of $P_{\rm err}(n)$ for a degree distribution, which has a computational complexity of $\mathcal{O}(K^2n^2M)$ (see [14]). The second part is the evaluation of Algorithm 2, which has a computational complexity of $\mathcal{O}(K^2M)$. The first part actually dominates the computation cost of the BRP-based approach. The time complexity of the greedy approach and the BRP-based approach can be compared using the number of times of evaluating $P_{\rm err}(n)$.

We use an example to demonstrate the performance of the BRP-based approach for GE-BP decoding. Consider a BATS code with $K = 256$, $M = 16$, $q = 256$ and the rank distribution $\boldsymbol{h}_0$ shown in Table 1. Three degree distributions $\boldsymbol{\Psi}^{\rm asy}$, $\boldsymbol{\Psi}^{\rm g\text{-}B}$ and $\boldsymbol{\Psi}^{\rm brp\_B}$ (given in Table A1 in Appendix A) are used in our evaluation, where $\boldsymbol{\Psi}^{\rm asy}$ is obtained by the asymptotic analysis approach; $\boldsymbol{\Psi}^{\rm g\text{-}B}$ is obtained by the greedy approach of GE-BP decoding; $\boldsymbol{\Psi}^{\rm brp\_B}$ is obtained by our BRP-based approach for GE-BP decoding. We evaluate the error probability $P_{\rm err}(n)$, $n = 1, \ldots, 150$, for these three degree distributions. See Figure 6 for the evaluation results.

We observe that $\boldsymbol{\Psi}^{\rm g\text{-}B}$ and $\boldsymbol{\Psi}^{\rm brp\_B}$ have almost the same error probability for different batch numbers, both of which demonstrate much better decoding performance than $\boldsymbol{\Psi}^{\rm asy}$. To achieve this performance, the greedy approach calculates $P_{\rm err}(n)$ for about 2500 times, which takes about 30 min using for computer with Intel i5-4590 CPU, while the BRP-based approach calculates $P_{\rm err}(n)$ for less than 50 times, which takes about 1 min for the same computer. In other words, the BRP-based approach is roughly 50 times faster than the greedy approach in this example.

5 BRP-based approach for inactivation decoding

5.1 Inactivation decoding

GE-BP decoding stops with high probability before the desired fraction of input symbols are decoded

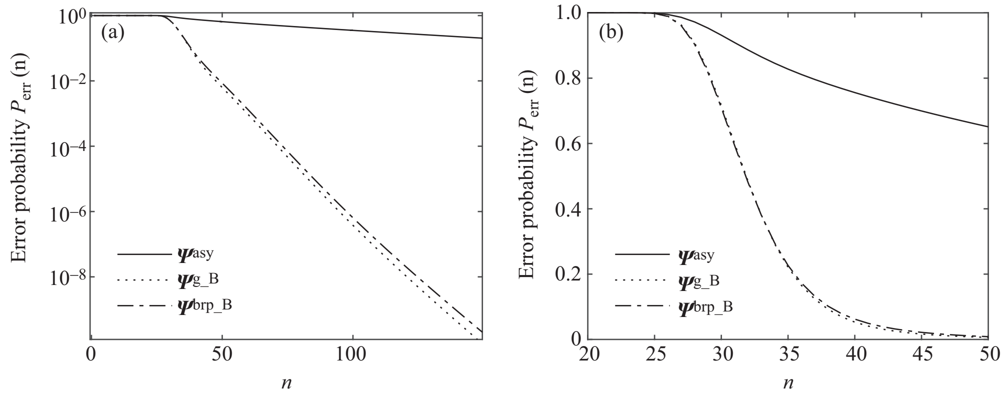


Figure 6 Error probability $P_{\rm err}(n)$ of GE-BP decoding for three degree distributions $\boldsymbol{\Psi}^{\rm asy}$, $\boldsymbol{\Psi}^{\rm g\text{-}B}$ and $\boldsymbol{\Psi}^{\rm brp\text{-}B}$. (a) n from 1 to 150; (b) zoom-in with n from 20 to 50.

when the number of input symbols is small. Though GE-BP decoding stops, Gaussian elimination can still be used to decode the remaining input symbols. But the decoding complexity of Gaussian elimination is much higher than that of GE-BP decoding. A better way to continue the decoding process is to use inactivation. We introduce the inactivation decoding provided in [14], which is a practical decoding method for BATS codes.

The inactivation decoding process for a given number n of batches is denoted by INAC(n). The decoding of INAC(n) is the same as BP(n) until there are no decodable symbols. Instead of stopping the decoding in BP(n), INAC(n) tries to resume the GE-BP decoding process by "inactivating" certain undecoded input symbols. Specifically, suppose that there are no decodable input symbols at step t, the decoder randomly picks an undecoded symbol b and marks it as inactive. The decoder substitutes the inactive b into the batches like a decoded symbol, except that b is an indeterminate and increases the step by one. Since the step is increased by one for each input symbol decoded or inactivated, the decoding process of INAC(n) is repeated until step K when all the input symbols are either decoded or inactive. The inactive input symbols can be recovered by solving a linear system of equations using Gaussian elimination. In a nutshell, inactivation decoding trades computation cost (decoding inactive input symbols using Gaussian elimination) with coding overhead.

For INAC(n), we are interested in the number of inactive symbols when the decoding stops. In [14], a recursive formula is provided to calculate the expected number of inactive symbols for inactivation decoding. The BRP-based degree-distribution optimization approach discussed in Subsection 4.3 can be used to optimize the degree distribution for inactivation decoding as well. We only need to use the expected number of the inactive symbols as the objective function in Step 2.

5.2 Performance comparison with the greedy approach for inactivation decoding

Let us demonstrate the performance of BRP-based approach for inactivation decoding by an example. Consider a BATS code with $K = 256$, $M = 16$, $q = 256$ and the rank distribution $\boldsymbol{h}_0$ in Table 1. Three degree distributions $\boldsymbol{\Psi}^{\rm asy}$, $\boldsymbol{\Psi}^{\rm g\text{-}I}$ and $\boldsymbol{\Psi}^{\rm brp\text{-}I}$ (given in Table A1 in Appendix A) are used in our evaluation, where $\boldsymbol{\Psi}^{\rm asy}$ is obtained by the asymptotic analysis approach of BATS codes in [2]; $\boldsymbol{\Psi}^{\rm g\text{-}I}$ is obtained by the greedy approach for inactivation decoding; $\boldsymbol{\Psi}^{\rm brp\text{-}I}$ is obtained by the BRP-based approach for inactivation decoding. We evaluate the expected number of inactive symbols of INAC(n), $n = 1, \ldots, 200$, for the three degree distributions. See Figure 7 for the evaluation results.

We observe that $\boldsymbol{\Psi}^{\rm g\text{-}I}$ and $\boldsymbol{\Psi}^{\rm brp\text{-}I}$ have almost the same expected number of inactive symbols for different batch numbers, which is far less than that of $\boldsymbol{\Psi}^{\rm asy}$. To achieve this performance, the BRP-based approach calculates the expected number of inactive symbols for about 40 times, while the greedy approach evaluates the expected number of inactive symbols for about 2000 times.

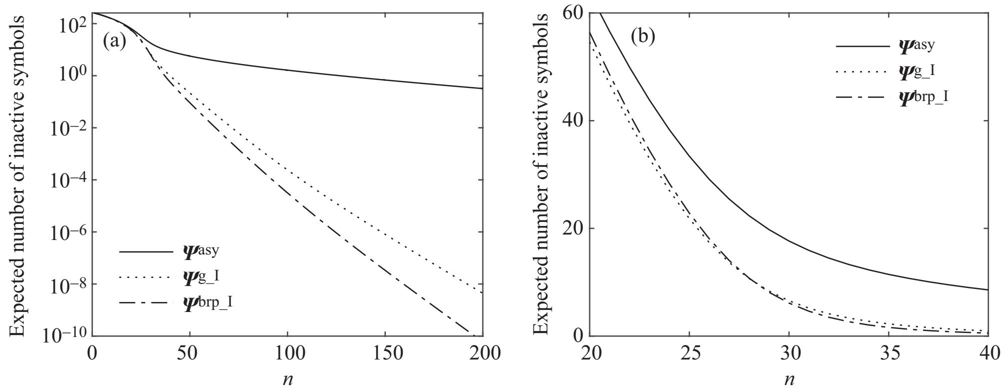


Figure 7 Expected number of inactive symbols for different degree distributions $\boldsymbol{\Psi}^{\rm asy}$, $\boldsymbol{\Psi}^{\rm g\text{-}I}$ and $\boldsymbol{\Psi}^{\rm brp\text{-}I}$. (a) n from 1 to 150; (b) zoom-in with n from 20 to 40.

Table 2 The rank distribution $\boldsymbol{h}_1$ for evaluation examples. $\boldsymbol{h}_1$ is the one of the length-3 line network with link erasure probability 0.2, 0.3 and 0.3, respectively

h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	h_{11}	h_{12}	h_{13}	h_{14}	h_{15}
0.0001	0.0005	0.0026	0.0113	0.0376	0.0971	0.1888	0.2622	0.2382	0.1254	0.0327	0.0034	0.0001

5.3 The universality of BRP

We know that a universal degree distribution does not exist for BATS codes, i.e., a degree distribution optimized for one rank distribution may not have a good performance for another rank distribution. As an example, recall the degree distribution $\boldsymbol{\Psi}^{\rm g\text{-}I}$ optimized for rank distribution $\boldsymbol{h}_0$ used in the last subsection, and consider another rank distribution $\boldsymbol{h}_1$ given in Table 2. The inactivation decoding performance of $\boldsymbol{\Psi}^{\rm g\text{-}I}$ for $\boldsymbol{h}_1$ is plotted in Figure 8(a), and is compared with the greedy-approach optimized degree distribution $\boldsymbol{\Psi}^{\rm g\text{-}I^*}$ for $\boldsymbol{h}_1$. We see that $\boldsymbol{\Psi}^{\rm g\text{-}I^*}$ is much better than $\boldsymbol{\Psi}^{\rm g\text{-}I}$ for $\boldsymbol{h}_1$.

We, however, find that the BRPs demonstrate certain universal property. Note that when using our BRP-based approaching for rank distribution $\boldsymbol{h}_0$, the algorithm outputs a pair (v_1^*, v_2^*) together with the degree distribution $\boldsymbol{\Psi}^{\rm brp\text{-}I}$. We can obtain another degree distribution $\boldsymbol{\Psi}^{\rm u\text{-}I}$ by Algorithm 2 w.r.t. (v_1^*, v_2^*) and rank distribution $\boldsymbol{h}_1$. Roughly speaking, $\boldsymbol{\Psi}^{\rm u\text{-}I}$ is a degree distribution such that BRP $R(\boldsymbol{\Psi}^{\rm u\text{-}I}, \boldsymbol{h}_1)$ can be approximately characterized by (v_1^*, v_2^*). As in Figure 8(a), $\boldsymbol{\Psi}^{\rm u\text{-}I}$ demonstrates a very similar performance as $\boldsymbol{\Psi}^{\rm g\text{-}I^*}$ for $\boldsymbol{h}_1$, where the latter is optimized by the greedy-approach for $\boldsymbol{h}_1$. This is surprising since (v_1^*, v_2^*) is optimized (by our BRP-based approach) for $\boldsymbol{h}_0$, but not for $\boldsymbol{h}_1$.

To verify whether the above observation is accidental, we use 2000 rank distributions with $M = 16$ generated randomly [15] to test the performance of (v_1^*, v_2^*), particularly $v_1^* = 0.015$, $v_2^* = 0.02$. For each rank distribution $\boldsymbol{h}$, two degree distributions $\boldsymbol{\Psi}^{\rm init}$ and $\boldsymbol{\Psi}^{\rm opt}$ are compared: $\boldsymbol{\Psi}^{\rm init}$ is obtained by Algorithm 2 w.r.t. (v_1^*, v_2^*) and $\boldsymbol{h}$, and $\boldsymbol{\Psi}^{\rm opt}$ is obtained using the BRP-based approach using $\{(v_1^*, v_2^*)\}$ as the initial list. The ratio α of the expected numbers of the inactive symbols for $\boldsymbol{\Psi}^{\rm init}$ and $\boldsymbol{\Psi}^{\rm opt}$ is calculated. Note that α is not larger than 1, and $\boldsymbol{\Psi}^{\rm init}$ is better if α is closer to 1.

The empirical distributions of the values α for these 2000 rank distributions are plotted in Figure 8(b). We can see that for more than 93.7% of the rank distributions, α is larger than 0.9; for all the rank distributions the smallest α is 0.747. Therefore, the (v_1^*, v_2^*) can achieve high performance for most of the rank distributions in our experiments.

This observation hints that, if the initial list is properly chosen, our BRP-based degree-distribution optimization only needs a small number of iterations to generate nearly optimal degree distributions. This conclusion can be explained as follows. The BRP vector of degree distribution obtained from greedy approach demonstrates two peaks, which are depicted by v_1 and v_2 (see Subsection 4.2). v_1 and v_2 for different BRP vectors are similar.

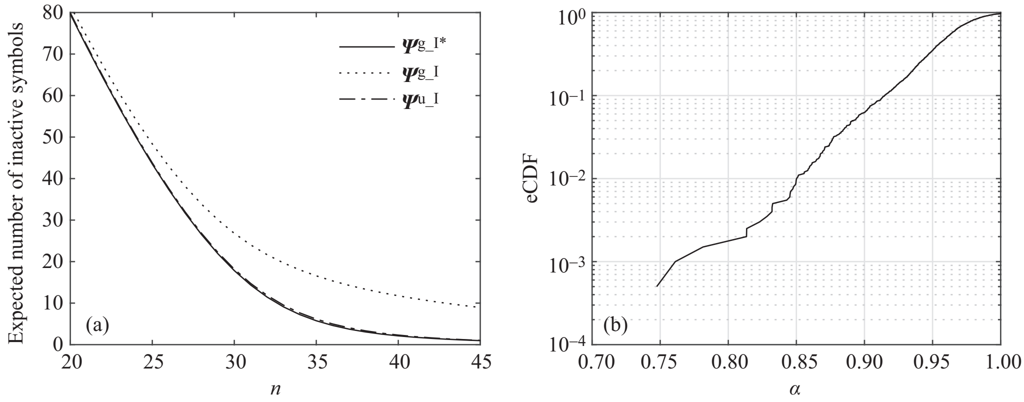


Figure 8 Some numerical results to demonstrate the universality of the (v_1, v_2). (a) The expected number of inactive symbols for three degree distributions $\boldsymbol{\Psi}^{\text{g\_I}^*}$, $\boldsymbol{\Psi}^{\text{g\_I}}$ and $\boldsymbol{\Psi}^{\text{u\_I}}$; (b) the empirical distributions of performance ratio α for 2000 different rank distributions for $(0.015, 0.02)$.

6 Conclusion

In this paper, we propose the concept of batch release probability (BRP) for capturing the characteristics of the good degree distributions of BATS codes. Based on these BRP characteristics, we propose a novel degree-distribution optimization approach. This approach can be significantly faster than the greedy approach proposed previously, and achieves nearly the same performance as the latter. The universal property observed in this paper deserves some further investigations towards better understanding of the decoding performance, as well as designing better BATS codes.

Acknowledgements This work was supported by Delay/Disruption Tolerant Communication Technology of the Chinese Academy of Sciences (Grant No. 2015AA703405A), and National Natural Science Foundation of China (Grant No. 61471215).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Yang S H, Yeung R W. Coding for a network coded fountain. In: Proceedings of IEEE International Symposium on Information Theory (ISIT), Saint-Petersburg, 2011. 2647–2651
2 Yang S H, Yeung R W. Batched sparse codes. IEEE Trans Inf Theory, 2014, 60: 5322–5346
3 Luby M. LT codes. In: Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, Los Alamitos, 2002. 271–280
4 Shokrollahi A. Raptor codes. IEEE/ACM Trans Netw, 2006, 14: 2551–2567
5 Li S-Y R, Yeung R W, Cai N. Linear network coding. IEEE Trans Inf Theory, 2003, 49: 371–381
6 Ho T, Medard M, Koetter R, et al. A random linear network coding approach to multicast. IEEE Trans Inf Theory, 2006, 52: 4413–4430
7 Tang B, Yang S, Ye B, et al. Expander chunked codes. Eurasip J Adv Signal Process, 2015, 2015: 1–13
8 Mahdaviani K, Ardakani M, Bagheri H, et al. Gamma codes: a low-overhead linear-complexity network coding solution. In: Proceedings of International Symposium on Network Coding, Cambridge, 2012. 125–130
9 Yang S, Tang B. From LDPC to chunked network codes. In: Proceedings of IEEE Information Theory Workshop, Hobart, 2014. 406–410
10 Huang Q, Sun K, Li X, et al. Just FUN: a joint fountain coding and network coding approach to loss-tolerant information spreading. In: Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing, Philadelphia, 2014. 83–92
11 Xu X, Praveen Kumar M S G, Guan Y L, et al. Two-phase cooperative broadcasting based on batched network code. IEEE Trans Commun, 2016, 64: 706–714
12 Xu X, Guan Y L. Reliable broadcast to a user group with limited source transmissions. In: Proceedings of IEEE International Conference on Communications, London, 2015
13 Yang S, Yeung R W, Cheung J H F, et al. BATS: network coding in action. In: Proceedings of the 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2015. 1204–1211
14 Yang S, Ng T C, Yeung R W. Finite-length analysis of BATS codes. In: Proceedings of International Symposium on Network Coding (NetCod), Calgary, 2013

15 Smith N A, Tromble R W. Sampling uniformly from the unit simplex. Johns Hopkins University, Tech. Rep, 2004. http://www.cs.cmu.edu/~nasmith/papers/smith+tromble.tr04.pdf

Appendix A Table of degree distribution

Several degree distributions used for the rank distribution $\boldsymbol{h}_0$ and $\boldsymbol{h}_1$ are listed in Tables A1 and A2, respectively.

Table A1 Several degree distributions for the rank distribution $\boldsymbol{h}_0$ in Table 1

(a) $\boldsymbol{\Psi}^{\rm asy}$: the degree distribution obtained using the asymptotic analysis in [14]

$\Psi^{\rm asy}_{14}$	$\Psi^{\rm asy}_{15}$	$\Psi^{\rm asy}_{20}$	$\Psi^{\rm asy}_{21}$	$\Psi^{\rm asy}_{27}$	$\Psi^{\rm asy}_{28}$	$\Psi^{\rm asy}_{37}$
0.0467	0.2502	0.1079	0.0781	0.0350	0.0968	0.0728
$\Psi^{\rm asy}_{38}$	$\Psi^{\rm asy}_{50}$	$\Psi^{\rm asy}_{51}$	$\Psi^{\rm asy}_{72}$	$\Psi^{\rm asy}_{116}$	$\Psi^{\rm asy}_{117}$	$\Psi^{\rm asy}_{256}$
0.0199	0.0676	0.0087	0.0697	0.0277	0.0312	0.0896

(b) $\boldsymbol{\Psi}^{\rm g\_B}$: the degree distribution obtained by the greedy approach for GE-BP decoding

$\Psi^{\rm g\_B}_{11}$	$\Psi^{\rm g\_B}_{12}$	$\Psi^{\rm g\_B}_{13}$	$\Psi^{\rm g\_B}_{14}$	$\Psi^{\rm g\_B}_{15}$	$\Psi^{\rm g\_B}_{20}$	$\Psi^{\rm g\_B}_{21}$	$\Psi^{\rm g\_B}_{27}$	$\Psi^{\rm g\_B}_{28}$
0.0826	0.0734	0.0550	0.0429	0.1745	0.0348	0.0809	0.0321	0.0888
$\Psi^{\rm g\_B}_{37}$	$\Psi^{\rm g\_B}_{38}$	$\Psi^{\rm g\_B}_{50}$	$\Psi^{\rm g\_B}_{51}$	$\Psi^{\rm g\_B}_{72}$	$\Psi^{\rm g\_I}_{116}$	$\Psi^{\rm g\_B}_{117}$	$\Psi^{\rm g\_B}_{256}$	
0.0484	0.0183	0.0620	0.0080	0.0623	0.0254	0.0286	0.0822	

(c) $\boldsymbol{\Psi}^{\rm g\_I}$: the degree distribution obtained by the greedy approach for inactivation decoding

$\Psi^{\rm g\_I}_{12}$	$\Psi^{\rm g\_I}_{13}$	$\Psi^{\rm g\_I}_{14}$	$\Psi^{\rm g\_I}_{15}$	$\Psi^{\rm g\_I}_{20}$	$\Psi^{\rm g\_I}_{21}$	$\Psi^{\rm g\_I}_{26}$	$\Psi^{\rm g\_I}_{27}$	$\Psi^{\rm g\_I}_{28}$
0.0796	0.0973	0.0414	0.2126	0.0955	0.0692	0.0088	0.0309	0.0857
$\Psi^{\rm g\_I}_{37}$	$\Psi^{\rm g\_I}_{38}$	$\Psi^{\rm g\_I}_{50}$	$\Psi^{\rm g\_I}_{51}$	$\Psi^{\rm g\_I}_{72}$	$\Psi^{\rm g\_I}_{116}$	$\Psi^{\rm g\_I}_{117}$	$\Psi^{\rm g\_I}_{256}$	
0.0644	0.0176	0.0598	0.0077	0.0512	0.0245	0.0276	0.0262	

(d) $\boldsymbol{\Psi}^{\rm brp\_B}$: the degree distribution obtained by BRP-based approach for GE-BP decoding

$\Psi^{\rm brp\_B}_{11}$	$\Psi^{\rm brp\_B}_{13}$	$\Psi^{\rm brp\_B}_{17}$	$\Psi^{\rm brp\_B}_{24}$	$\Psi^{\rm brp\_B}_{36}$	$\Psi^{\rm brp\_B}_{55}$	$\Psi^{\rm brp\_B}_{84}$	$\Psi^{\rm brp\_B}_{126}$	$\Psi^{\rm brp\_B}_{183}$	$\Psi^{\rm brp\_B}_{256}$
0.0705	0.2259	0.1773	0.1621	0.1180	0.0782	0.0496	0.0306	0.0184	0.0693

(e) $\boldsymbol{\Psi}^{\rm brp\_I}$: the degree distribution obtained by BRP-based approach for inactivation decoding

$\Psi^{\rm brp\_I}_{12}$	$\Psi^{\rm brp\_I}_{13}$	$\Psi^{\rm brp\_I}_{17}$	$\Psi^{\rm brp\_I}_{23}$	$\Psi^{\rm brp\_I}_{34}$	$\Psi^{\rm brp\_I}_{51}$	$\Psi^{\rm brp\_I}_{78}$	$\Psi^{\rm brp\_I}_{116}$	$\Psi^{\rm brp\_I}_{165}$	$\Psi^{\rm brp\_I}_{223}$	$\Psi^{\rm brp\_I}_{256}$
0.0435	0.2421	0.1874	0.1592	0.1335	0.0854	0.0579	0.0346	0.0207	0.0121	0.0235

Table A2 Several degree distributions for the rank distribution $\boldsymbol{h}_1$ in Table 2

(a) $\boldsymbol{\Psi}^{\rm g\_I}$: the degree distribution obtained by the greedy approach for inactivation decoding

$\Psi^{\rm g\_I}_{11}$	$\Psi^{\rm g\_I}_{12}$	$\Psi^{\rm g\_I}_{13}$	$\Psi^{\rm g\_I}_{16}$	$\Psi^{\rm g\_I}_{17}$	$\Psi^{\rm g\_I}_{18}$	$\Psi^{\rm g\_I}_{19}$	$\Psi^{\rm g\_I}_{20}$	$\Psi^{\rm g\_I}_{21}$	$\Psi^{\rm g\_I}_{22}$
0.0843	0.2734	0.0749	0.0391	0.0126	0.0190	0.0360	0.0588	0.0141	0.0461
$\Psi^{\rm g\_I}_{23}$	$\Psi^{\rm g\_I}_{25}$	$\Psi^{\rm g\_I}_{29}$	$\Psi^{\rm g\_I}_{34}$	$\Psi^{\rm g\_I}_{45}$	$\Psi^{\rm g\_I}_{55}$	$\Psi^{\rm g\_I}_{62}$	$\Psi^{\rm g\_I}_{63}$	$\Psi^{\rm g\_I}_{119}$	$\Psi^{\rm g\_I}_{256}$
0.0152	0.0278	0.0388	0.0616	0.0545	0.0283	0.0278	0.0124	0.0553	0.0202

(b) Degree distributions $\boldsymbol{\Psi}^{\rm u\_I}$ by BRP-based approach for inactivation decoding

$\Psi^{\rm u\_I}_{11}$	$\Psi^{\rm u\_I}_{15}$	$\Psi^{\rm u\_I}_{22}$	$\Psi^{\rm u\_I}_{34}$	$\Psi^{\rm u\_I}_{53}$	$\Psi^{\rm u\_I}_{95}$	$\Psi^{\rm u\_I}_{256}$
0.3032	0.2217	0.1853	0.1271	0.0817	0.0561	0.0249