# Decentralized-Precoding Aided Rateless Codes for Wireless Sensor Networks

Shinya Sugiura, *Member, IEEE*

***Abstract*—In this letter, we propose a novel decentralized-precoding aided rateless code, which is conceived for Wireless Sensor Networks (WSNs). More specifically, the proposed algorithm allows a collection of source nodes to precode their information symbols in an uncoordinated manner. The precoded symbols are then channel-encoded by a distributed rateless code at each sensor node. This architecture enables us to eliminate an error floor imposed on the conventional Luby Transform (LT)-code based WSNs, without requiring a central coordinator or any elaborate cooperations between the nodes. Our simulation results demonstrate that the proposed distributed rateless code is capable of attaining achievable BER performance comparable to that of a Raptor code coordinated by a central coordinator, while outperforming a distributed LT code.**

***Index Terms*—Distributed storage, fountain codes, graphical codes, network coding, rateless codes, sensor network, belief propagation algorithm.**

## I. Introduction

Wireless Sensor Networks (WSNs) [1], [2] consist of unreliable devices, sensors or vehicles, which can be used for instance as a distributed storage system [3]. Typically, the employment of an efficient Forward Error Correction (FEC) scheme is necessary for such a virtual storage system, in order to protect source information stored in distributed sensor nodes. Furthermore, it is desirable for WSNs to be able to complete the channel-encoding process in a decentralized and distributed manner, which is for the sake of avoiding any substantial overhead and energy consumption.

In comparison to the conventional fixed-rate FEC schemes, the recent class of rateless codes [4] is capable of producing a potentially infinite length of codeword, implying that its code rate as well as the amount of redundancy does not have to be fixed before the transmissions. The first practical rateless code, which is the so-called Luby Transform (LT) code [5], was invented for the Internet Binary Erasure Channels (BECs). While LT codes tend to exhibit a high error floor for noisy channels, such as Additive White Gaussian Noise (AWGN) and Rayleigh fading channels, Raptor codes [6] were conceived for the sake of combating this limitation, where a high-rate Low-Density Parity-Check (LDPC) channel encoder is inserted as a precoder in advance of the LT encoder. As the result, Raptor codes benefit from an error-free performance in noisy channel environments as well as from a practically low decoding complexity.



The author is with Toyota Central R&D Laboratories, Inc., Nagakute, Aichi, 480-1192, Japan (e-mail: sugiura@ieee.org).

Owing to their beneficial characteristics, rateless codes were applied to several WSN scenarios [7], [8]. For instance, in [7] an LT code was employed at source nodes to spread their channel-coded information symbols to WSNs without necessitating any negotiation between the nodes. Furthermore, the destination receiver can decode the information symbols by gathering any subset of the LT-coded symbols, which enables the band-efficient implementation of Incremental Redundancy (IR), which is especially useful for broadcast- or multicast-scenarios. However, since most of the previous studies related to distributed storage are based on LT codes, rather than Raptor codes, the above-mentioned error floor imposed by noisy channels is unavoidable. Unfortunately, it is impractical to configure a Raptor-coded WSN, because it requires the insertion of a precoder, which has to be aided with a central coordinator or with substantial information exchanges between the nodes, as mentioned in [7]. One exceptional approach is constituted by the exploitation of random walk [9], where the redundant channel-encoded symbols are generated during the additional transmissions in advance of the distributed LT coding. Note that the achievable performance of this scheme over noisy channels was not documented.

Against this background, the novel contributions of this letter are as follows. We first propose an efficient decentralized-precoding aided rateless code, where source nodes generate redundant symbols in an uncoordinated manner. More specifically, the redundant symbols are calculated so as to have appropriate correlations with the source symbols. Hence, this distributed-precoding architecture allows us to eliminate an error floor imposed on the distributed LT code, similarly to the centralized Raptor code. Importantly, owing to the fact that the proposed precoding scheme does not rely on either cooperative precoding or an additional random walk process, the cost imposed on encoding the source information is as low as that of the distributed LT code [7]. Furthermore, while most of the previous studies with respect to the distributed rateless codes considered a BEC, we provide simulation results of several distributed rateless codes in the context of binary-input AWGN channels.

The remainder of this letter is organized as follows. Section II provides the conventional rateless-code assisted WSN and the proposed decentralized-precoding aided rateless codes. The related numerical analysis is carried out in Section III. Finally, our conclusions are presented in Section IV.

## II. Proposed Distributed Rateless Codes

In this section, we firstly describe the network model adopted in this letter. Then, we highlight the conventional distributed LT-code [7] based WSNs, which is followed by the proposal of the novel distributed rateless-coding framework.

### A. Network Model

Let us consider a WSN, which consists of $N$ storage sensor nodes, $K$ source nodes as well as the communication links between them. We assume for simplicity that each source node has a single binary information symbol $d_k \in \{-1, 1\}$ $(k = 1, \cdots, K)$, which is saved at the $N$ sensor nodes. Moreover, a common destination node communicates with a subset of the sensor nodes in order to attain the stored source information. Hence, the following two-phase transmissions are considered in this scenario. In the broadcast phase, the $K$ source nodes transmit their symbols to the $N$ sensor nodes, where the received symbols are channel encoded at each sensor node. Then, in the collection phase the destination node receives the channel-coded symbols in order to decode the $K$ source symbols.[1]

### B. The Conventional Distributed LT Codes

The distributed LT code enables each of the $N$ sensor nodes to independently generate its own channel-coded symbol. This also allows the destination node to decode the symbols received from any subset of the $N$ channel coded symbols, which is achieved as the explicit benefit of the LT-code based rateless code. More specifically, in the broadcast phase, the $K$ source nodes independently transmit their symbols to the $N$ sensor nodes, noting that in this scheme any restriction may not be imposed on the type of the Medium Access Control (MAC) protocol employed. Then, letting the $n$th sensor node of the interest, $e_n$ number of source symbols randomly chosen, according to a predetermined probability generating function $\Omega(x)$ [5]. Finally, a symbol $c_n$ is determined by XORing the $e_n$ number of selected symbols, which is stored at the $n$th sensor node. Here, the associated edge information, represented by the indices of the selected source symbols, is also saved in the memory of the associated sensor node.

To be more specific, $\Omega(x)$ is represented by

$$\Omega(x) = \delta_1 x + \delta_2 x^2 + \cdots + \delta_k x^k + \cdots + \delta_K x^K, \quad (1)$$

where $\delta_k$ $(k = 1, \cdots, K)$ represents the particular fraction of check nodes having the degree $\delta_k$, noting that we have the relationship of $\sum_{k=1}^{K} \delta_k = 1$. Also, the average number of edges connected from the variable nodes to the check nodes is given by $\Omega'(1) = \sum_{k=1}^{K-1} k\delta_k$. Here, the check nodes are represented by XOR operations at the sensor nodes, while the variable nodes correspond to the source symbols of the source nodes [10].

In the collection phase, the destination node receives $(1 + \epsilon)K < N$ output symbols from arbitrary sensor nodes together with the associated edge information, where $\epsilon$ represents the portion of overhead symbols received at the destination node. Then, the Belief Propagation (BP) algorithm [10] is invoked for decoding the $K$ source symbols. In this process, the mutual information increases upon increasing the number of iterations between check- and variable-nodes.

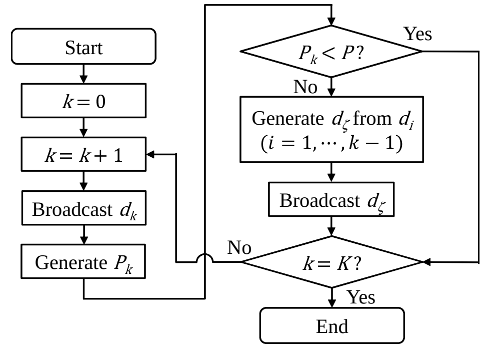


Fig. 1. Flowchart of the proposed decentralized-precoding assisted source transmissions.

### C. The Proposed Decentralized-Precoding Aided Rateless Codes

Fig. 1 depicts the broadcast phase of the proposed rateless code, where the $K$ source nodes transmit in total $M > K$ precoded symbols. This implies that $(M - K)$ redundant symbols are added in this precoding process. Here, a certain MAC protocol, such as Time-Division Multiple Access (TDMA) and Carrier Sense Multiple Access (CSMA), is assumed to be employed, where each of the $K$ source nodes carries out its transmissions from the first to the $K$th source nodes in order, without loss of any generality.[2]

More specifically, as portrayed in Fig. 1, the $k$th source node broadcasts its own symbol $d_k$ to the other source nodes as well as to the sensor nodes. Then, the $k$th source node generates a real-valued random variable $P_k$, which is uniformly distributed between 0 and 1. Here, if $P_k$ is lower than a preassigned threshold $P$, the $k$th source node's transmission is completed. Otherwise, the $k$th source node calculates another symbol $d_\zeta$ from the $(k-1)$ previously-transmitted source symbols $d_i$ $(i = 1, \cdots, k-1)$. More specifically, $\min(k-1, e)$ out of the $(k-1)$ source symbols are chosen so that the selected symbols have the least number of connections at the instant of the $k$th node's transmission [11].[3] The integer value $e$ obeys the probability generating function $\bar{\Omega}(x) = \sum_{d=1}^{K} \bar{\delta}_k x^k$, similarly to Eq. (1). Then, a redundant symbol $d_\zeta$ is generated by XORing the selected $\min(k-1, e)$ source symbols, which is transmitted to the $N$ sensor nodes.[4] Finally, the $N$ sensor nodes operate the distributed LT encoding of Section II-B based on the $M$ precoded symbols, which are received during the broadcast phase.

[1] In order to elaborate a little further, although we considered in this letter the scenario of $K$ source nodes, each having a single source bit, the proposed scheme can be readily applicable to a more general scenario of $K/B$ source nodes, each having $B \geq 1$ source bits.

[2] In order to expound a little further, our rateless code does not support MAC protocols, relying on purely simultaneous transmissions from the nodes, which are for example Code-Division Multiple Access (CDMA) and Orthogonal Frequency Division Multiple Access (OFDMA). We note nevertheless that the hybrid scheme, such as Time-Division (TD)/CDMA and TD/OFDMA schemes, may still be applicable to the proposed rateless code.

[3] This is possible because edge information corresponding to the first $(k-1)$ source nodes are known at the moment of the $k$th source node's transmission.

[4] To avoid potential errors for the packets saved at storage nodes, Cyclic Redundancy Checking (CRC) codes are typically employed at the source nodes, which enables error detection at the source- and storage-nodes.

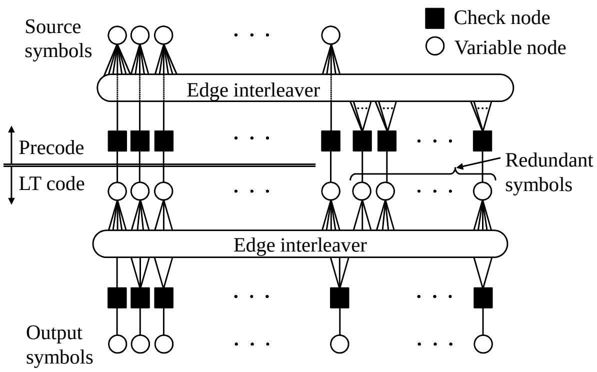


Fig. 2. Tanner graph of the proposed decentralized-precoding aided rateless code.

To provide further insights, in the proposed precoding scheme the edges connected from the check nodes to the variable nodes are directional, rather than symmetrically random. This indicates that the redundant check node of the $k$th source node, satisfying the relationship of $P_k \geq P$, potentially has the connections with the previously transmitted $(k-1)$ source symbols, rather than with the rest of the $(K-k)$ source symbols, unlike the conventional Raptor code. Although this restriction may induce some performance degradation, we will demonstrate later in Section III that the proposed rateless code is capable of attaining the performance comparable to that of the Raptor code.

We note that the above-mentioned value $P$, which is introduced as the threshold of the statistical redundant-symbol generations, corresponds to the average code-rate of the proposed decentralized precoder, hence the normalized transmission rate of the proposed rateless code may be defined by $R = P/(1+\epsilon)$. Furthermore, a higher $P$ value leads to the lower number of redundant symbols, hence resulting in the degradation of the error-correction capability. This implies that the optimization of $P$ imposes a design tradeoff between the achievable error-rate performance and the transmission rate.

Fig. 2 shows the Tanner graph of the proposed rateless code, which forms a two-stage serial-concatenated graph code, i.e. the precoder equivalent to the systematic Low-Density Generator Matrix (LDGM) code and the LT encoder. While there are several potential solutions of this graph code, in this letter we employ the tandem iterative detection [12], which is typical for the conventional Raptor code. To be more specific, the soft outputs of the inner LT decoder are calculated with the aid of the number of inner iterations $I_{\rm LT}$, which are then input to the outer iterative decoder, where we have the number of outer iterations $I_{\rm pre}$. Hence, the total number of iterations becomes $(I_{\rm LT} + I_{\rm pre})$. Also, we note that a higher number of edges $d_\zeta$ associated with the redundant symbols of Fig. 2 imposes a higher complexity at the distributed precoder and the destination node's receiver.

## III. Simulation Results

In this section the performance of our distributed rateless code is characterized by carrying out Monte Carlo simulations. We also considered two benchmark schemes, namely the distributed LT code of Section II-B as well as the Raptor code [6], which was assumed to be coordinated by a central node.

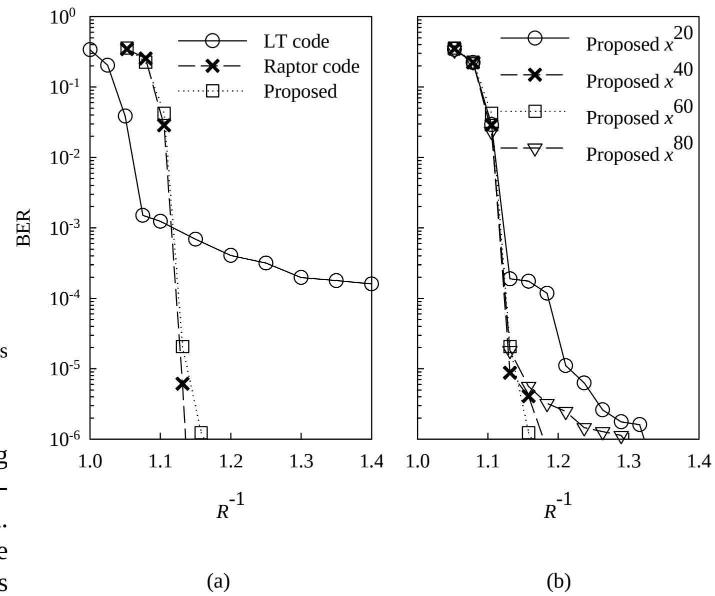


Fig. 3. Achievable BER performance of the distributed LT code, the coordinated Raptor code and the proposed decentralized-precoding aided rateless code, each employing BPSK modulation under the assumption of AWGN channel, exhibiting SNR = 5 dB. Furthermore, we employed (a) $\bar{\Omega}(x) = x^{60}$ and (b) $\bar{\Omega}(x) = x^{20}, x^{40}, x^{60}$ and $x^{80}$ as the precoder's random generating functions.

In our simulations, we set the LT-code's random distribution $\Omega(x)$ used at the $N$ sensor nodes as [13]

$$\begin{aligned}\Omega(x) &= 0.007969x + 0.493570x^2 + 0.166220x^3 \\ &+ 0.072646x^4 + 0.082558x^5 + 0.056058x^8 \\ &+ 0.037229x^9 + 0.055590x^{19} + 0.025023x^{65} \\ &+ 0.0003135x^{66}, \qquad (2)\end{aligned}$$

while that of the proposed precoder $\bar{\Omega}(x) = x^{D_c}$ was chosen as follows: $D_c = 20, 40, 60$ and $80$. The maximum number of inner and outer iterations was set to $(I_{\rm LT}, I_{\rm pre})$ = $(200, 200)$. Furthermore, the threshold $P$ was chosen as $P = 0.95$, which corresponds to the average outer code-rate. We also considered the Binary Phase-Shift Keying (BPSK) modulation and the AWGN channels for the links between the sensor nodes and the destination node.

Fig. 3(a) shows the achievable BER performance of the three rateless codes, namely the LT code, the Raptor code and the proposed rateless code at the Signal-to-Noise Ratio (SNR) of 5 dB. For the Raptor code and the proposed rateless code, we assumed for $K = 9\,500$ source bits and the $P = 0.95$ code-rate precoder. More specifically, the Raptor code employed the systematic LDPC precoder, where each source symbol has the degree four and the edges to the redundant check nodes selected uniformly, similarly to [11]. Moreover, 10 000 source bits as well as the probability generating function $\Omega(x)$ of Eq. (2) was used for the distributed LT code. Here,

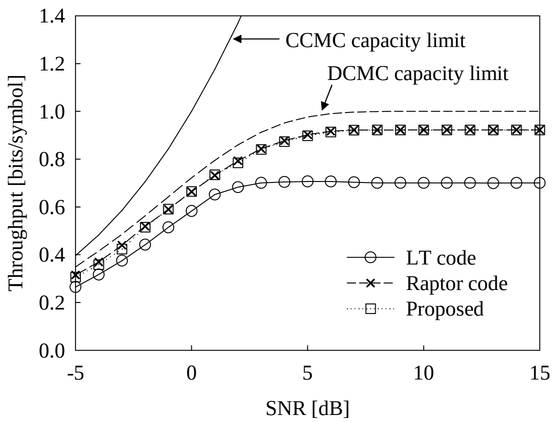


Fig. 4. Effective throughputs of the distributed LT code, the coordinated Raptor code and the proposed decentralized-precoding aided rateless code, each employing BPSK modulation under the assumption of AWGN channel Here, $\bar{\Omega}(x) = x^{60}$ was used at the proposed scheme's precoder.

we note that since the distributed LT code did not utilize a precoder, its normalized throughput may be simply expressed as $R = 1/(1+\epsilon)$. The random generating function $\bar{\Omega}(x)$ of our rateless code's precoder was given by $\bar{\Omega}(x) = x^{60}$. It was found from Fig. 3(a) that the proposed rateless code closely approached the performance of the coordinated Raptor code, while a marginal performance loss was seen for BER $= 10^{-6}$. As predicted, the LT-code based system exhibited a high error-floor for approximately BER = $10^{-4}$ in the simulated regime.[5]

Next, Fig. 3(b) investigates the effects of the random generating function $\bar{\Omega}(x) = x^{D_c}$ on the achievable BER performance of the proposed rateless code, where we varied $D_c$ as $D_c = 20, 40, 60$ and $80$. It can be seen in Fig. 3(b) that the $D_c = 60$ scenario attained a higher transmission rate for BER $= 10^{-6}$ than other $D_c$ scenarios. Furthermore, the scenarios of $D_c = 20$ and $80$ exhibited an error floor, similarly to the distributed LT code of Fig. 3(a). This ensures that the appropriate design of $\bar{\Omega}(x)$ is vital for our distributed rateless code. More specifically, although an increase in the value of $D_c$ tended to result in the performance improvement, an excessively high $D_c$ may degrade the performance owing to the so-called marginalization problem [10], which is caused by the associated graph codes with cycles.

Finally, in Fig. 4 we compared the effective throughputs of the three rateless codes and their bounds. In each frame transmission of $K$ source bits, the initial number of the received symbols was set to 10 000 and additional 100 bits were received until the iterative decoder output the correct $K$ source symbols. Here, we also plotted the two corresponding capacity-limit curves, which are represented by Continuous input Continuous output Memoryless Channel (CCMC) capacity [14] and Discrete input Continuous output Memoryless Channel (DCMC) capacity [15]. Observe in Fig. 4 that both the proposed rateless code and the Raptor code exhibited a similar performance, which approached the DCMC capacity limit, while outperforming the LT code. In the simulated scenario, the LT code's throughput was approximately 24% lower than other two codes at high SNRs.

## IV. Conclusions

In this letter, we proposed the novel rateless code conceived for WSNs, which can be operated in a fully decentralized manner, while combating the error-floor limitation of the conventional LT-code based counterpart. More specifically, each source node stochastically generates a redundant symbol, such that it has correlations with the previously transmitted source symbols. Hence, the resultant graph-code structure becomes similar to that of a centralized Raptor code. It was shown in our simulations that the proposed decentralized rateless code is capable of approaching the BER performance of the Raptor code, while outperforming the distributed LT code. The proposed rateless-code architecture may also be useful for improving multi-source cooperative communications based on the distributed LT code [16].

## References


[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[2] F. Wang and J. Liu, "Networked wireless sensor data collection: Issues, challenges, and approaches," *IEEE Communications Surveys & Tutorials*, 2011, doi:10.1109/SURV.2011.060710.00066, 12 pages in press.

[3] A. Dimakis, P. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4539–4551, 2010.

[4] N. Bonello, S. Chen, and L. Hanzo, "Low-density parity-check codes and their rateless relatives," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 1, pp. 3–26, Feb. 2011.

[5] M. Luby, "LT codes," in *the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'02)*, 2002, pp. 271–280.

[6] A. Shokrollahi, "Raptor codes," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.

[7] A. Dimakis, V. Prabhakaran, and K. Ramchandran, "Decentralized erasure codes for distributed networked storage," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2809–2816, 2006.

[8] K. Pang, Z. Lin, Y. Li, and B. Vucetic, "Design of distributed network-channel codes for wireless sensor networks," in *IEEE International Conference on Communications (ICC 2011)*, Kyoto, Japan, 5–9, June 2011, pp. 1–5.

[9] Z. Kong, S. Aly, and E. Soljanin, "Decentralized coding algorithms for distributed storage in wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 2, pp. 261–267, 2010.

[10] T. Richardson and R. Urbanke, *Modern coding theory*. Cambridge University Press, 2008.

[11] N. Bonello, R. Zhang, S. Chen, and L. Hanzo, "Reconfigurable rateless codes," *IEEE Transactions on Wireless Communications*, vol. 8, no. 11, pp. 5592–5600, 2009.

[12] A. Venkiah, C. Poulliat, and D. Declercq, "Jointly decoded Raptor codes: Analysis and design for the biawgn channel," *EURASIP Journal on Wireless Communications and Networking*, 2009, 11 pages, doi:10.1155/2009/657970.

[13] R. Palanki and J. Yedidia, "Rateless codes on noisy channels," in *International Symposium on Information Theory (ISIT 2004)*, Chicago, Illinois, USA, 27 June – 2 July 2004, pp. 1–12.

[14] J. Proakis, *Digital communications*. McGraw-Hill, New York, 2001.

[15] S. X. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 528–536, 2006.

[16] S. Puducheri, J. Kliewer, and T. Fuja, "The design and performance of distributed LT codes," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3740–3754, 2007.


[5] Although we employed the fixed number of edges at the proposed precoder's redundant check nodes, which is represented by $\bar{\Omega}(x) = x^{D_c}$, it is possible to reduce a marginal performance gap between the Raptor code and the proposed code of Fig. 3(a) by considering its distribution $\bar{\Omega}(x)$, similarly to Eq. (2). However, the detailed optimization and analysis will be left for our future study.