Interactive Reconciliation with Low-Density Parity-Check Codes

Jesús Martínez-Mateo, David Elkouss and Vicente Martín

Research group on Quantum Information and Computation
Universidad Politécnica de Madrid (UPM)
Campus de Montegancedo, 28660 Boadilla del Monte (Madrid), Spain
e-mail: {jmartinez, delkouss, vicente}@fi.upm.es

Abstract—**Efficient information reconciliation is crucial in several scenarios, being quantum key distribution a remarkable example. However, efficiency is not the only requirement for determining the quality of the information reconciliation process. In some of these scenarios we find other relevant parameters such as the interactivity or the adaptability to different channel statistics. We propose an interactive protocol for information reconciliation based on low-density parity-check codes. The coding rate is adapted in real time by using simultaneously puncturing and shortening strategies, allowing it to cover a predefined error rate range with just a single code. The efficiency of the information reconciliation process using the proposed protocol is considerably better than the efficiency of its non-interactive version.**

I. INTRODUCTION

Since the publication of the first quantum protocol, more than 25 years ago, quantum key distribution (QKD) [1] has evolved into a functional and commercial technology, and nowadays it is already possible to find commercial QKD systems by several manufacturers. A QKD system is used to create secret keys between two parties connected through a quantum channel, i.e. for instance an optic fibre. However, this technology is still far from reaching its real potential due to the lack of suitable developments in some of its fundamental processes, such as error correction. In a QKD protocol, error correction is included within a broader process known as secret key distillation [2]. In this process, error correction is a procedure used to reconcile discrepancies between two bit sequences, for this reason this procedure is known as information reconciliation. In order to accomplish it, the parties must exchange additional information over a public but authenticated channel: it can be read but not modified by an hypothetical eavesdropper. Since the information exchanged for reconciliation provides information about the key, the parties must agree on an additional procedure, called privacy amplification [3], used to reduce the information that may have been derived by any eavesdropper. An optimal reconciliation procedure provides the minimum information required for correcting the discrepancies between two sequences, minimising the key material that must be discarded during the privacy amplification, therefore maximising the final secret-key length.

One of the first methods proposed for correcting errors in a QKD system was *Cascade* [4]. Currently, it is probably the most widely used procedure for this purpose in QKD, due
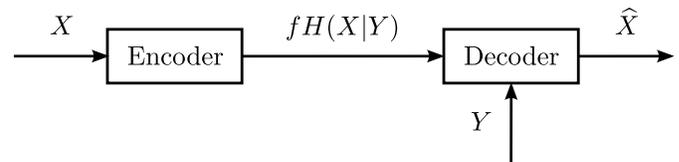

Fig. 1. Source coding with side information for one-way reconciliation.

to its simplicity and relatively good efficiency (see Fig. 3). However, *Cascade* is a highly interactive process that requires many communication rounds. The parties have to exchange a large number of messages where parities of different blocks and subblocks of a key are published.

A better alternative for error correction in QKD systems is provided by other strategies such as low-density parity-check (LDPC) codes. These codes were introduced by Gallager in the early 60s [5], and recently several proposals have emerged for using LDPC codes in the information reconciliation process [6], [7]. In this paper we propose a new protocol for error correction using rate adaptive LDPC codes. The protocol is able to correct errors within a known error rate range, iteratively transmitting more symbols in order to minimise the information transmitted for correction.

The paper is organised as follows: First, in section II, it is described the problem of information reconciliation in the secret-key agreement context. Then, in section III, a new protocol is proposed to improve the reconciliation process using interactive communication between the parties. Finally, in section IV, results with this new protocol are shown. These results are also compared with two different approaches: a similar proposal using rate adaptive codes but without interactive communication, and the simplest approach using LDPC codes without rate modulation.

II. INFORMATION RECONCILIATION

The problem of information reconciliation, when only one-way transmissions are allowed, can be modelled by the more general problem of source coding with side information. In this section we describe this more general approach, and are reviewed those techniques used to adapt LDPC codes in the information reconciliation context.

Fig. 2.   Examples of puncturing and shortening strategies applied to an LDPC code in order to modulate the coding rate.

A. Source Coding with Side Information

Let X and Y be two discrete random variables representing two correlated sources, and let X^n and Y^n be two correlated sequences obtained from both sources respectively. Assuming that these sources are separated into two legitimate parties: Alice and Bob. Information reconciliation allows Bob to recover X^n with the help of Y^n and sending M messages over a lossless channel. In the source coding with side information description, one of the parties encode the sequence X^n, and the other recovers $\hat{X}^n$ using the information provided by the encoded sequence and with help of side information Y^n, such that $X^n = \hat{X}^n$ with high probability. The minimum rate for encoding the source X in order to get $X = \hat{X}$ with the side information provided by Y was determined by Slepian-Wolf to be $H(X|Y)$ [8] (see Fig. 1). Both problems, information reconciliation and source coding with side information, are equivalent if only one-way transmissions are allowed, and $H(X|Y)$ is the minimum rate that could be used in a reconciliation protocol. However, even though the problem is formally different, an interactive reconciliation process shares the same lower bound [9]. Thus the reconciliation efficiency, f, for both one-way and interactive protocols can be defined by:

$$f = \frac{R_X}{H(X|Y)} \geq 1 \qquad (1)$$

In the quantum cryptography context information reconciliation arise after the basis reconciliation process, this is when both parties of a QKD system, Alice and Bob, share a raw key with discrepancies that should be removed by following a key distillation process. In most of QKD protocols, e.g. BB84 [10] or SARG [11], these discrepancies are uncorrelated and symmetric, such that they can be interpreted as errors in a communication made through a binary symmetric channel (BSC).

B. LDPC Codes and Syndrome Decoding

LDPC codes are known to achieve coding rates near the capacity of several channels under belief propagation decod-

ing [12]. It has been also shown that these codes can be used to encode near the theoretical limit for source coding with side information [13]. A modified decoder was proposed for syndrome decoding and applying the bin approach by Wyner [14]. The use of LDPC codes for encoding correlated sources was later formalised [15].

Following this exposition, information reconciliation can be solved for many QKD protocols by using good LDPC codes for the BSC. This problem has been already addressed, and good families of these codes have been found for different coding rates [6]. However, an LDPC code is constructed for a fixed coding rate. In consequence, in those scenarios of varying characteristics, such as QKD, if the parties do not share a suitable number of codes, the efficiency curve shows a saw behaviour (see Fig. 6). In order to solve this behaviour, in the next section we describe a new protocol able to adapt the coding rate of an LDPC code, minimising the information revealed for reconciliation.

III. PROTOCOL

A. Rateless Coding

Puncturing and shortening are two suitable strategies able to adapt the rate of a channel code as we have already shown in a previous work [7]. When p punctured symbols of a codeword are removed, a $[n, k]$ code is converted into a $[n - p, k]$ code (see Fig. 2a). Whereas, when shortening, s symbols are removed during the encoding process, and a $[n, k]$ code is converted into a $[n - s, k - s]$ code (see Fig. 2b).

Supposing that R_0 is the original coding rate of a family of LDPC codes defined by:

$$R_0 = 1 - \frac{\sum_i \lambda_i / i}{\sum_j \rho_j / j} \qquad (2)$$

where λ_i and ρ_i are the coefficients of their generating polynomials. This rate can be modulated applying puncturing
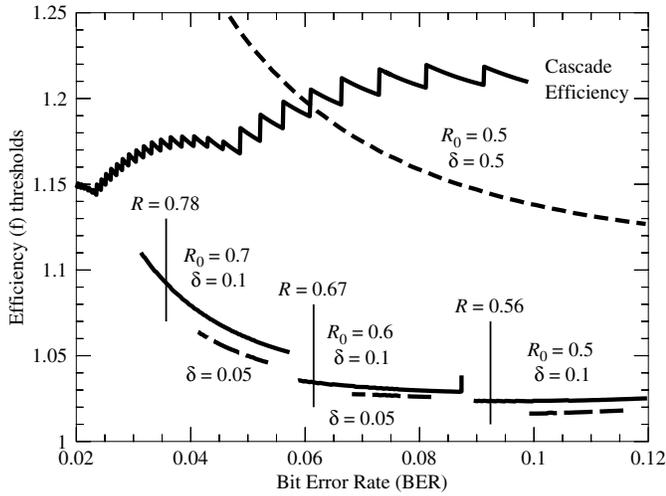
Fig. 3. Efficiency of *Cascade* [4] and efficiency thresholds of LDPC codes with rate modulation for an error rate from 2% to 10%. The range is representative of good to very bad quantum channels for the QKD case. Efficiencies have been calculated, using the expression defined in Eq. 1, for three different codes with rates R_0: 0.5, 0.6 and 0.7. Two δ values, 0.1 and 0.05, have been used for the rate modulation of these codes. An additional $\delta = 0.5$ has been used with $R_0 = 0.5$ covering the entire error rate range. The curves show how the efficiency of the LDPC code depends on the ratio of puncturing and shortening, δ, and the original coding rate, R_0. Higher δ values imply a bigger range of coding rates covered, unfortunately the efficiency drops for high values of δ [16]. The efficiency drop increases for simultaneous use of high δ values with high coding rates.

and shortening procedures as defined below. The modulated rate is then calculated as:

$$R = \frac{k - s}{n - p - s} = \frac{R_0 - \sigma}{1 - \pi - \sigma} \qquad (3)$$

where $\pi = p/n$ and $\sigma = s/n$ are the ratios of punctured symbols and shortened symbols respectively.

Both strategies, puncturing and shortening, may be applied in an isolated way in order to increase or decrease the coding rate respectively. However, we propose the use of both strategies simultaneously defining a new constant parameter, $\delta = \pi + \sigma$. This proposal is based primarily on two reasons:

1) Applying the same proportion of puncturing and shortening in every modulation. In consequence, regardless of the coding rate, the key length that can be corrected using the modulated code is known in advance.
2) As it is discussed in the next section, the use of puncturing and shortening simultaneously allow us to modify a previously modulated code in order to decrease the coding rate and to repeat an unsatisfactory correction process.

The efficiency, defined in Eq. 1, of the modulated code depends on the coding rate and the ratio δ of puncturing and shortening as shown in Fig. 3.

B. Interactive Reconciliation

The process of adapting the coding rate of an LDPC code is usually done with a previous estimate of the error rate to
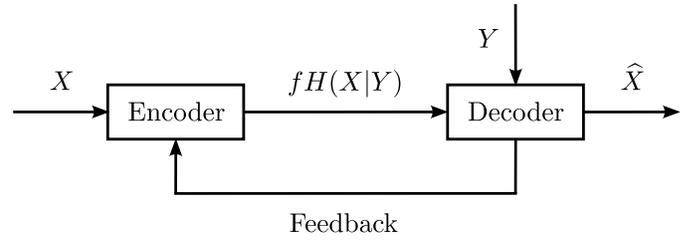
be corrected, this estimation is traditionally carried out by exchanging a sample of the sequence on the public channel. We propose here a new protocol for information reconciliation with LDPC codes that does not require to estimate the error rate. The protocol is based on syndrome coding, but adding a new functionality: feedback information about the success of the decoding process (see Fig. 4). With this feedback the original one-way approach becomes an interactive protocol, as described below, with more flexibility in order to correct optimally a range of error rates. The protocol is blind in the sense that it is able to adapt to different channel configurations without a prior estimate and, the reconciliation is successful as long as the channel's characteristics are within a pre-established range.

The protocol is described by the following three steps:

Step 0) Raw Key Exchange: Initially it is assumed that two sources, X and Y, generate two correlated symbol sequences, **x** and **y** belonging to Alice and Bob respectively. Moreover, it is also supposed that the two symbol sequences have discrepancies within a bounded error rate range, $[e_0, e_1]$. From this hypothesis, Alice and Bob can choose an LDPC code with an information rate aimed to correct an intermediate point in the interval. Depending on the range and the efficiency target, the parties agree on a δ value to cover the entire range of required coding rates.

$$R_{\min} = \frac{R_0 - \delta}{1 - \delta} \leq R \leq \frac{R_0}{1 - \delta} = R_{\max} \qquad (4)$$

such that $R_{\min} \leq 1 - h_2(e_1)$ and $R_{\max} \geq 1 - h_2(e_0)$, where h_2 is the binary Shannon entropy.

As initial coding rate is chosen the highest value, $R = R_{\max}$, such that all symbols used to modulate the rate correspond to a punctured symbol, i.e. $\delta = \pi$ and $\sigma = 0$. In this case the protocol provides the minimum amount of information.

Step 1) Encoding: Once it has been established a value for the coding rate, both parties compute the number of symbols to be punctured and shortened, p and s respectively:

$$s = \lceil (R_0 - R(1 - \delta)) \, n \rceil$$
$$p = \lfloor \delta n \rfloor - s \qquad (5)$$

The first time this step is run, Alice randomly chooses the symbols to be punctured —there are no shortened symbols in the first round—, and set them with random values. Once Alice knows which positions correspond with punctured symbols,

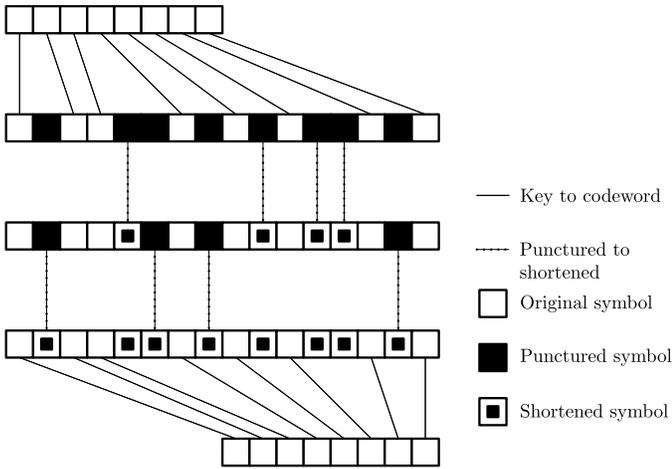Fig. 4. Source coding with side information and feedback.

Fig. 5. Example of a complete execution of the interactive protocol. The example shows how the symbols of an initial sequence are distributed in different positions of a codeword. The remaining positions of the codeword are initially marked as punctured symbols. In each round, the protocol replaces a proportion of punctured symbols with shortened symbols, thus reducing the coding rate.

and their values, she calculates the syndrome (compressed information), $\mathbf{z} = \mathbf{x}H^t$, and sends it to Bob along with their positions.

In subsequent runs of this step, Alice chooses randomly a preestablished proportion of punctured symbols that will be converted to shortened symbols and transmits to Bob their positions and values. The proportion of converted symbols in each round must be agreed by both parties at the beginning of the protocol, and it depends on the maximum number of rounds that is allowed and the desired efficiency.

Step 2) Decoding: Bob uses his correlated sequence, $\mathbf{y}$, and the information provided by punctured and shortened symbols as starting point to find a sequence with a syndrome that matches the syndrome received from Alice, $\mathbf{z}$. The protocol is successfully concluded in this step when Bob decodes a word matching the syndrome received. Otherwise, if the decoding process is stopped because the maximum number of pre-set iterations has been reached, then Bob agrees with Alice a decrease in the coding rate, if possible, and they return to the previous Step 1.

The protocol fails if the coding rate takes its minimum value and decoding is unsuccessful, i.e. $R = (R_0 - \delta)/(1 - \delta)$, which happens for $\delta = \sigma$ and $\pi = 0$.

A graphic description of this protocol is shown in Fig. 5. The figure illustrates an example showing three executing rounds.

Different executions of the proposed protocol may conclude with different ratios for puncturing and shortening, π and σ respectively, i.e. different protocol executions reconcile the original sequences with different efficiencies. Efficiency for a single protocol execution is defined in Eq. 1, the efficiency of this protocol can be measured by taking an average value. Puncturing and shortening ratios are then calculated by:

TABLE I
PROPORTION AND NUMBER OF PUNCTURED AND SHORTENED SYMBOLS, AND MODULATED RATE PER ROUND.

Round	δ	π^*	σ^*	p	s	R_0	R
0	0.1	1.00	0.00	20000	0	0.6	0.67
1	0.1	0.83	0.17	16666	3334	0.6	0.65
2	0.1	0.67	0.33	13332	6668	0.6	0.63
3	0.1	0.50	0.50	9999	10001	0.6	0.61
4	0.1	0.33	0.67	6666	13334	0.6	0.59
5	0.1	0.17	0.83	3333	16667	0.6	0.57
6	0.1	0.00	1.00	0	20000	0.6	0.56

$$\hat{\pi} = \frac{1}{M}\sum_{i=1}^{M}\pi^{(i)} \; ; \quad \hat{\sigma} = \frac{1}{M}\sum_{i=1}^{M}\sigma^{(i)} \qquad (6)$$

where M is the number of executions. The average efficiency is then calculated as:

$$\hat{f} = \frac{1}{h_2(e)} - \frac{R_0 - \hat{\sigma}}{(1 - \delta)h_2(e)} = \frac{1 - R_0 - \hat{\pi}}{(1 - \delta)h_2(e)} \qquad (7)$$

where e is the crossover probability that has been corrected.

Let Q be the maximum number of rounds, the efficiency value in an isolated execution is increased in each round by a constant factor, ϵ, that depends on the proportion of new shortened symbols, $q = \delta/Q$ such that $\pi_{j+1} = \pi_j - q$ and $\sigma_{j+1} = \sigma_j + q$. The efficiency of an execution that concludes in the round j can be also calculated by $f_j = f_0 + j\epsilon$, where:

$$f_0 = \frac{1 - R_0 - \delta}{(1 - \delta)h_2(e)} \; ; \quad \epsilon = \frac{q}{(1 - \delta)h_2(e)} \qquad (8)$$

IV. RESULTS

In order to produce a representative set of simulations that demonstrate the operation of the proposed protocol, we decided to build a single LDPC code of length $n = 2 \times 10^5$ and rate $R_0 = 0.6$ using a family of codes proposed by Elkouss [6] for the BSC. The code, with a puncturing and shortening proportion of $\delta = 0.1$, is able to modulate coding rates from $R_{\min} = 0.56$ to $R_{\max} = 0.67$, i.e. it is possible to construct codes for correcting error rates in an approximate range from 6% to 9%. Table I shows the proportion and number of punctured and shortened symbols in each round, assuming that only a maximum of seven rounds can be executed. For convenience, the proportions of punctured and shortened symbols has been normalised to 1, such that $\pi = \delta\pi^*$ and $\sigma = \delta\sigma^*$.

Table II shows the average number of rounds, $\hat{N}$, needed for correcting different error rates using a previously constructed LDPC code. The table also includes the average number of punctured symbols, $\hat{p}$, and the average number of shortened symbols, $\hat{s}$, that have been used in the last round of the correction process. From the average number of punctured and shortened symbols, the average proportion of punctured and shortened symbols, $\hat{\pi}$ and $\hat{\sigma}$ respectively, have been calculated together with the average efficiency, $\hat{f}$, as defined in Eq. 7.

Finally, Fig. 6 shows the reconciliation efficiency curves, calculated according to Eq. 1, but obtained with three different

TABLE II
AVERAGE NUMBER OF ROUNDS NEEDED TO CORRECT DIFFERENT ERROR
RATES WITH A SINGLE BUT MODULATED LDPC CODE.

BER	$\hat{N}$	$\hat{p}$	$\hat{s}$	$\hat{\pi}$	$\hat{\sigma}$	$\hat{f}$
0.055	0.03	19900	100	0.0995	0.0005	1.08664
0.060	1.12	16266	3734	0.0813	0.0187	1.08144
0.065	2.36	12132	7868	0.0607	0.0393	1.08651
0.070	3.12	9599	10400	0.0480	0.0520	1.06883
0.075	4.38	5399	14601	0.0270	0.0730	1.07841
0.080	5.00	3333	16667	0.0167	0.0833	1.05895
0.085	6.00	0	20000	–	–	–
0.090	6.00	0	20000	–	–	–
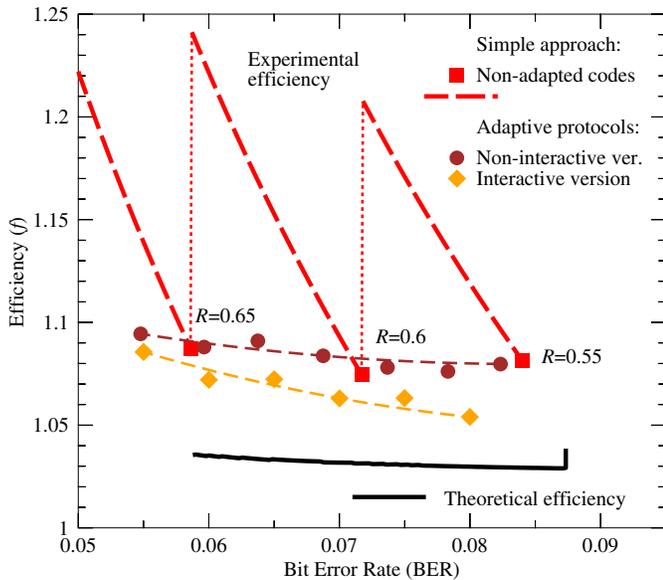
Fig. 6. Comparison between the efficiency obtained in the original rate adaptive protocol (non-interactive) using LDPC codes [7] and the interactive version proposed here. It is also included the theoretical efficiency computed for the same code, and the efficiency calculated for LDPC codes without rate modulation [6].

approaches: i) the protocol proposed here where the maximum number of allowed rounds has been increased to 20, ii) the original non-interactive protocol [7], and iii) an ensemble of LDPC codes for different rates.

V. CONCLUSIONS

In this paper we have studied an interactive information reconciliation protocol. The protocol has been analysed empirically and the different trade-offs in terms of decoding complexity, interactivity and efficiency have been described.

The protocol has several advantages. The interactive nature of the reconciliation improves the decoding process, whenever the decoding fails, a fraction of the punctured symbols is revealed thus allowing for virtually zero error rate after decoding. The adaptive characteristic of the protocol allows to skip measuring the error rate on the channel. Several applications can boost their performance if this step is skipped: an important example is secret key distillation in QKD protocols. In this context the error rate is measured by publicly showing a subset of the sequences and discarding the shown symbols,

the elimination of this step allows the parties to distill a significantly higher secret key rate.

The protocol presented on this paper can find a broad range of applications as it allows the parties to achieve reconciliation efficiencies as low as desired and, in several scenarios, to avoid the waste of a relevant part of the sequence for sampling purposes.

REFERENCES

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 145–195, Mar. 2002.
[2] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
[3] C. Bennett, G. Brassard, C. Crepeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.
[4] G. Brassard and L. Salvail, "Secret-Key Reconciliation by Public Discussion," in *Eurocrypt'93, Workshop on the theory and application of cryptographic techniques on Advances in cryptology*, ser. Lecture Notes in Computer Science, vol. 765. Springer-Verlag, 1994, pp. 410–423.
[5] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.
[6] D. Elkouss, A. Leverrier, R. Alleaume, and J. J. Boutros, "Efficient reconciliation protocol for discrete-variable quantum key distribution," in *IEEE International Symposium on Information Theory*, Jul. 2009, pp. 1879–1883.
[7] D. Elkouss, J. Martínez, D. Lancho, and V. Martín, "Rate Compatible Protocol for Information Reconciliation: An application to QKD," in *IEEE Information Theory Workshop*, Jan. 2010, pp. 145–149.
[8] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.
[9] G. Van Assche, *Quantum Cryptography and Secret-Key Distillation*. Cambridge University Press, 2006.
[10] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Int. Conference on Computers, Systems and Signal Processing*, 1984, pp. 175–179.
[11] V. Scarani, A. Acín, G. Ribordy, and N. Gisin, "Quantum Cryptography Protocols Robust against Photon Number Splitting Attacks for Weak Laser Pulse Implementations," *Phys. Rev. Lett.*, vol. 92, no. 5, p. 057901, Feb. 2004.
[12] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001.
[13] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.
[14] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1250–1276, 2002.
[15] J. Muramatsu, T. Uyematsu, and T. Wadayama, "Low-density parity-check matrices for coding of correlated sources," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3645–3654, Oct. 2005.
[16] H. Pishro-Nik and F. Fekri, "Results on Punctured Low-Density Parity-Check Codes and Improved Iterative Decoding Techniques," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 599–614, Feb. 2007.