

Interval-Passing Algorithm for Non-negative Measurement Matrices: Performance and Reconstruction Analysis

Vida Ravanmehr, Ludovic Danjean,
Bane Vasić, *Fellow, IEEE*, and David Declercq, *Senior Member, IEEE*

Abstract—We consider the Interval-Passing Algorithm (IPA), an iterative reconstruction algorithm for reconstruction of non-negative sparse real-valued signals from noise-free measurements. We first generalize the IPA by relaxing the original constraint that the measurement matrix must be binary. The new algorithm operates on any non-negative sparse measurement matrix. We give a performance comparison of the generalized IPA with the reconstruction algorithms based on (i) linear programming and (ii) verification decoding. Then we identify signals not recoverable by the IPA on a given measurement matrix, and show that these signals are related to stopping sets responsible to failures of iterative decoding algorithms on the binary erasure channel (BEC). Contrary to the results of the iterative decoding on the BEC, the smallest stopping set of a measurement matrix is not the smallest configuration on which the IPA fails. We analyze the recovery of sparse signals on subsets of stopping sets via the IPA and provide sufficient conditions on the exact recovery of sparse signals. Reconstruction performance of the IPA using the IEEE 802.16e LDPC codes as measurement matrices are given to show the effect of stopping sets in the performance of the IPA.

Index Terms—Compressed Sensing, Interval-Passing Algorithm, Iterative Reconstruction Algorithm, Low-Density Parity-Check Codes, Message-Passing Algorithm, Stopping Sets.

I. INTRODUCTION

RECONSTRUCTING a sparse signal from a small set of measurements via compressed sensing [1] has attracted significant attention in the last few years. A k -sparse signal $\mathbf{x} \in \mathbb{R}^n$, i.e. a signal \mathbf{x} with at most k non-zero values, is observed indirectly through a shorter measurement vector $\mathbf{y} \in \mathbb{R}^m$ and obtained from the linear equations $\mathbf{y} = \mathbf{A}\mathbf{x}$ where \mathbf{A} is an $m \times n$ measurement matrix, with $m \ll n$. The task of compressed sensing is to recover \mathbf{x} from \mathbf{y} . The first approach to solve the compressed sensing problem is to find a signal \mathbf{x} with the smallest ℓ_0 -norm. The ℓ_0 -norm minimization of compressed sensing is NP-hard [2], [3]. Instead, the ℓ_1 -norm minimization solution based on linear programming (LP) was introduced to reconstruct \mathbf{x} . The LP technique [4] for the compressed sensing problem, called Basis Pursuit [5],

[6], has a remarkable performance, but its high complexity and running time makes it impractical in some applications which require fast reconstruction, or when the dimension of a measurement matrix is large. To tackle the issue of complexity, message-passing algorithms for compressed sensing have been proposed, originating from channel coding. Sarvothan *et al.* [7] were amongst the first who introduced a reconstruction algorithm based on belief propagation. Another application of belief propagation in compressed sensing was presented by Pham *et al.* in [8] where the authors provided two low-complexity algorithms, list decoding and multiple-based belief propagation. The iterative thresholding algorithm, also called approximate message passing, was proposed by Donoho *et al.* [9]. This algorithm can be viewed as the variant of the bit-flipping algorithm which is used for the decoding of low-density parity-check (LDPC) codes. A message passing algorithm, to which we refer in this paper as the Interval-Passing Algorithm (IPA), was introduced by Chandar *et al.* [10]. This is a simple iterative algorithm inspired by the parallel bit-flipping decoding algorithm to reconstruct a non-negative vector \mathbf{x} . In the binary measurement matrix case, the performance of this algorithm in both noise-free and noisy measurements was presented in [11], and a comparison between the IPA, the verification decoding algorithm [12], [13] and Basis Pursuit was given in [14].

Krishnan *et al.* [15] modified the IPA and showed that the IPA fails on the stopping sets which are the well-known configurations for the failure of the iterative decoding of LDPC codes over the binary erasure channel (BEC). Moreover, they showed that the IPA may fail even if the non-zero values of the signal do not contain a stopping set. In this paper, we analyze the recovery of the IPA in stopping sets and provide sufficient conditions for reconstructing a sparse signal with sparsity less than the size of a stopping set.

The rest of the paper is organized as follows. Section II provides preliminaries on compressed sensing, LDPC based measurement matrices, and a brief introduction to some iterative reconstruction algorithms. Section III provides a detailed explanation of the IPA. While the original version of the IPA uses binary measurement matrices, we modify this algorithm to work with non-negative sparse measurement matrices and present a comparison of the IPA with two other reconstruction techniques, namely the LP reconstruction, and the verification decoding. In Section IV, we give a theoretical analysis of the IPA recovery and provide sufficient conditions on the recovery

Manuscript received on February 29, 2012.

This work is funded by DARPA Knowledge Enhanced Compressive Measurements Project (KECoM) through contract #N66001-10-1-4079 and by NSF under grants CCF-0830245 and CCF-0963726.

V. Ravanmehr, L. Danjean, and B. Vasic, are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721, USA (e-mail: {vravanmehr, danjean, vasic}@email.arizona.edu).

L. Danjean and D. Declercq are with ETIS, ENSEA / University of Cergy-Pontoise / CNRS F-95000 Cergy-Pontoise, France (email: {danjean, declercq}@ensea.fr).

of signals whose non-zero values are a subset of a stopping set. In Section V, we provide simulation results exhibiting the IPA reconstruction performance using some LDPC matrices as the measurement matrices and link this performance with the stopping set distribution of these matrices. Finally, a conclusion and discussion are provided in Section VI.

II. PRELIMINARIES

In this section, we provide the notations used throughout the paper regarding compressed sensing and LDPC codes.

A. Compressed Sensing

Here we give definitions and main results related to compressed sensing. Let $\mathbf{x} \in \mathbb{R}^n$ be an n -dimensional k -sparse signal, and let $A_{m \times n}$ be a measurement matrix. Compressed sensing concerns the recovery of \mathbf{x} from measurements $\mathbf{y} = A\mathbf{x} \in \mathbb{R}^m$ where $m \ll n$ and $k \ll n$. The first approach to recover \mathbf{x} from the measurements \mathbf{y} , is to find a k -sparse signal from all the $\binom{n}{k}$ k -sparse possible signals which is known to be an NP-hard problem [2], [3]. This problem is called the ℓ_0 -minimization, and it is given by:

$$\hat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{x}\|_0 \text{ s.t. } \mathbf{y} = A\mathbf{x} \quad (1)$$

Throughout this paper, for a column vector $\mathbf{x} = [x_1, x_2, \dots, x_n]^t$, $\|\mathbf{x}\|_0 = |\operatorname{supp}(\mathbf{x})| = |\{k : x_k \neq 0\}|$, and $\|\mathbf{x}\|_1 = \sum_{k=1}^n x_k$.

The LP technique introduced by Chen *et al.* [5], called Basis Pursuit, provides a minimum ℓ_1 -norm for \mathbf{x} . This problem is called the ℓ_1 -minimization, as well as the recovery via convex optimization and it is given by:

$$\hat{\mathbf{x}} = \operatorname{argmin} \|\mathbf{x}\|_1 \text{ s.t. } \mathbf{y} = A\mathbf{x} \quad (2)$$

Candes *et al.* [6] showed that if the measurement matrix A satisfies a condition, namely the Restricted Isometry Property (RIP), Basis Pursuit can recover a sufficiently sparse signal. Roughly, a measurement matrix satisfies the RIP if it does not significantly distort the ℓ_2 -norm of any sufficiently sparse vector.

B. LDPC Codes

Low-density parity-check (LDPC) codes [16] are widely used in channel coding. Let \mathcal{C} denote an $(n, n - m)$ binary LDPC code. \mathcal{C} is defined as the null space of a parity check matrix H of size $m \times n$. H is known to be the adjacency matrix of its equivalent bipartite graph G , called the Tanner graph [17]. The Tanner graph G of an LDPC code is defined as $G = (V \cup C, E)$, where $V = \{v_1, \dots, v_n\}$ is the set of n variable nodes corresponding to the n columns of H , and $C = \{c_1, \dots, c_m\}$ is the set of m check nodes corresponding to the m rows of H . E is the set of edges in G , and an edge connects the variable node v_j to the check node c_i iff $h_{ij} = 1$ where h_{ij} is the element at the i^{th} row and j^{th} column in H . The code is said to be d_v -left-regular if all variable nodes in G have the same degree d_v (or equivalently if all the columns in H have weight d_v). The code is said to be d_c -right-regular if all check nodes in G have the same degree d_c (or equivalently

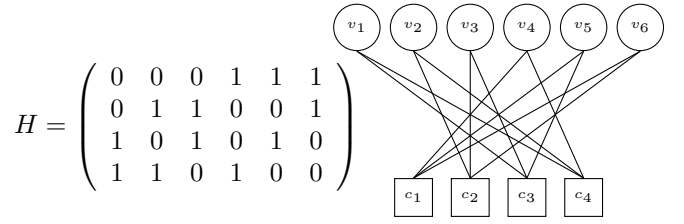


Fig. 1: Parity check matrix and Tanner graph of a $(2, 3)$ -regular LDPC code.

if all the rows in H have weight d_c). The code is said to be (d_v, d_c) -regular when it is both d_v -left-regular and d_c -right-regular. The girth g of a the Tanner graph G is the length of the shortest cycle in G . In the graphical representation of G , a white circle represents a variable node, and a white square represents a check node. The Fig. 1 shows a small example of a girth-6 $(2, 3)$ -regular LDPC matrix with its related Tanner graph. In this paper, we only consider (d_v, d_c) -regular LDPC codes with $d_v \geq 2$.

C. Message-Passing Algorithms

The relation between LDPC codes and compressed sensing was studied in [18] and [19] where the authors showed that good LDPC matrices are also good for compressed sensing using Basis Pursuit. Following the natural relation between measurement matrices for compressed sensing and LDPC matrices, several iterative algorithms have recently been proposed to recover sparse signals using message-passing. The first message-passing algorithms were introduced by Donoho *et al.* in [9] for noise free measurements. These algorithms, such as the second algorithm of [9], called the Approximate Message-Passing (AMP) algorithm, are iterative thresholding algorithms, i.e. at each iteration it ensures to provide the sparsest estimate of \mathbf{x} . These algorithms can be analyzed, as in [20], using the phase transition diagram. In this diagram the failure or success of the AMP is given as a function of different parameters of the system (k, m, n) .

Other message-passing algorithms exist in the literature, such as the list decoding and multiple-basis belief propagation from Pham *et al.* [8]. Chandar *et al.* [10] introduced a simple message passing algorithm that we investigate and modify in Section III.

We also present another kind of message-passing based algorithm used for comparison in our simulation results¹. Zhang and Pfister's verification algorithm [13] is an iterative algorithm to reconstruct strictly sparse signals. In this algorithm, summarized in Appendix A, the messages correspond to the vertices (variable and check nodes in the factor graph) and not to the edges. However, due to its extreme low complexity ($O(n)$) it is a good comparison for our simulation results.

III. INTERVAL-PASSING ALGORITHM

Chandar *et al.* [10] introduced a simple message passing algorithm for reconstructing non-negative signals using binary

¹This algorithm is presented in detail in [14].

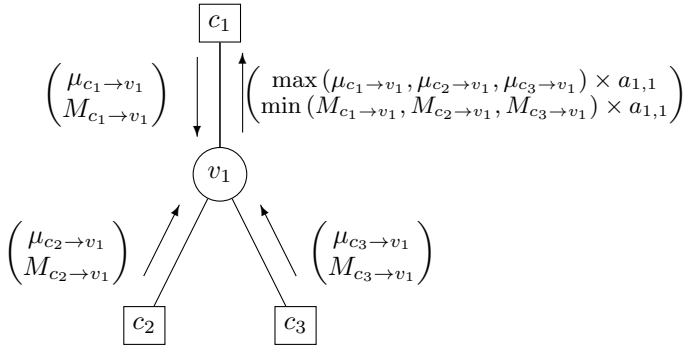


Fig. 2: IPA: Updating messages from the variable node v_1 to the measurement node c_1 .

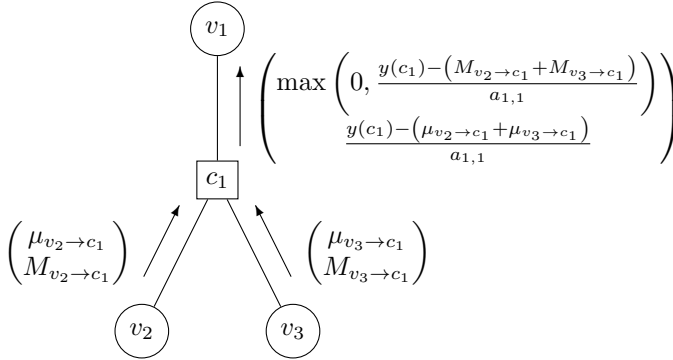


Fig. 3: IPA: Updating messages from the measurement node c_1 to the variable node v_1 .

measurement matrices. We refer to this algorithm as the IPA [11], [14], [15]. We modify this algorithm in order to deal with non-negative real-valued measurement matrices. From [10], the complexity of the algorithm is $\mathcal{O}(n(\log(\frac{n}{k}))^2 \log(k))$ which is a good trade-off between the polynomial complexity of the LP reconstruction, and the linear complexity of the simple verification decoding [13].

A. Description of the Algorithm

The IPA is an iterative algorithm, and thus messages are associated with the graphical representation of the measurement matrix to perform reconstruction. Let $V = \{v_1, v_2, \dots, v_n\}$ and $C = \{c_1, c_2, \dots, c_m\}$ be respectively the sets of variable nodes and measurement nodes² in the graphical representation of the measurement matrix $A = \{a_{j,i}\}$ for $1 \leq j \leq m$ and $1 \leq i \leq n$. The graphical representation of A is actually the Tanner graph of the binary image³ of A , whose edges are labeled by real values at the non-zero position in A . The graphical representation of A has the flavor of a non-binary LDPC code Tanner graph [21], [22].

In the IPA, the messages passing through edges are intervals $[\mu, M]$ corresponding to the lower and upper bounds of the

²Analogous to check nodes in LDPC codes.

³A matrix $H = \{h_{j,i}\}$ is said to be the binary image of a matrix $A = \{a_{j,i}\}$ if $h_{j,i} = 1$ if $a_{j,i} \neq 0$ and $h_{j,i} = 0$ if $a_{j,i} = 0$

Algorithm 1: Interval-Passing Algorithm

Input : \mathbf{y} and A such that $\mathbf{y} = A\mathbf{x}$, L .

Output : $\hat{\mathbf{x}}$ the estimate of \mathbf{x} .

Initialization: $\forall c_j \in C, \forall v_i \in \mathcal{N}(c_j), \mu_{c_j \rightarrow v_i}^{(0)} = 0$ and $M_{c_j \rightarrow v_i}^{(0)} = y(c_j)/a_{j,i}$;

for $l = 1$ **to** L **do**

```

1  foreach  $v_i \in V$  do
    foreach  $c_j \in \mathcal{N}(v_i)$  do
         $\mu_{v_i \rightarrow c_j}^{(l)} = \max_{c'_j \in \mathcal{N}(v_i)} (\mu_{c'_j \rightarrow v_i}^{(l-1)} \times a_{j,i};$ 
         $M_{v_i \rightarrow c_j}^{(l)} = \min_{c'_j \in \mathcal{N}(v_i)} (M_{c'_j \rightarrow v_i}^{(l-1)} \times a_{j,i};$ 
2  foreach  $c_j \in C$  do
    foreach  $v_i \in \mathcal{N}(c_j)$  do
         $\mu_{c_j \rightarrow v_i}^{(l)} = \frac{1}{a_{j,i}} \left( y(c_j) - \sum_{\substack{v'_i \in \mathcal{N}(c_j) \\ v'_i \neq v_i}} M_{v'_i \rightarrow c_j}^{(l)} \right);$ 
        if  $\mu_{c_j \rightarrow v_i}^{(l)} < 0$  then
             $\mu_{c_j \rightarrow v_i}^{(l)} = 0;$ 
         $M_{c_j \rightarrow v_i}^{(l)} = \frac{1}{a_{j,i}} \left( y(c_j) - \sum_{\substack{v'_i \in \mathcal{N}(c_j) \\ v'_i \neq v_i}} \mu_{v'_i \rightarrow c_j}^{(l)} \right);$ 
3  for  $v_i \in V$  do
    if  $(l > 1 \ \& \ \mu_{v_i \rightarrow \mathcal{N}(v_i)}^{(l)} = M_{\mathcal{N}(v_i) \rightarrow v_i}^{(l)}) \parallel l = L$ 
    then
         $\hat{x}(v_i) = \mu_{v_i \rightarrow \mathcal{N}(v_i)}^{(l)};$ 

```

estimation of the connected variable node. At each iteration l , the message update from the variable v_i to the measurement node c_j is given by:

$$\mu_{v_i \rightarrow c_j}^{(l)} = \max_{c'_j \in \mathcal{N}(v_i)} (\mu_{c'_j \rightarrow v_i}^{(l-1)} \times a_{j,i} \quad (3)$$

$$M_{v_i \rightarrow c_j}^{(l)} = \min_{c'_j \in \mathcal{N}(v_i)} (M_{c'_j \rightarrow v_i}^{(l-1)} \times a_{j,i} \quad (4)$$

and the messages from the measurement node c_j to the variable node v_i are updated as:

$$\mu_{c_j \rightarrow v_i}^{(l)} = \max \left\{ 0, \frac{y_j - \sum_{v'_i \in \mathcal{N}(c_j) \setminus \{v_i\}} M_{v'_i \rightarrow c_j}^{(l)}}{a_{j,i}} \right\} \quad (5)$$

$$M_{c_j \rightarrow v_i}^{(l)} = \frac{y_j - \sum_{v'_i \in \mathcal{N}(c_j) \setminus \{v_i\}} \mu_{v'_i \rightarrow c_j}^{(l)}}{a_{j,i}} \quad (6)$$

where $\mathcal{N}(v_i)$ (resp. $\mathcal{N}(c_j)$) is the set of measurement (resp. variable) nodes which are the neighbors of v_i (resp. c_j) in the Tanner graph of A . Updating messages from a variable (resp. measurement) node to a measurement (resp. variable) node is shown in Fig. 2 and 3, respectively.

The IPA is formally given in the Algorithm 1, where L represents the maximum number of reconstruction iterations. The reconstruction process stops when the maximum number

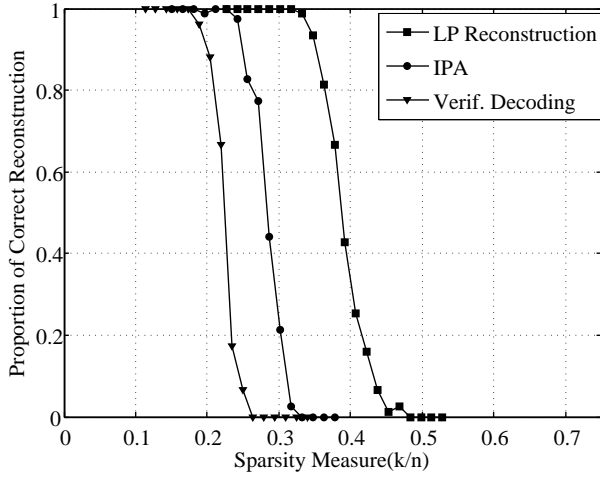


Fig. 4: Simulation results using the designed measurement matrix A .

of iterations is reached, or the lower bound and the upper bound of the interval from variable nodes to check nodes has converged to a common value for every variable node. This common value is set as the estimate of each connected variable node value.

B. Comparison Results

We now compare the IPA performance using the non-negative real valued measurement matrix with the complex LP reconstruction algorithm, and with simple verification decoding. We used the LDPC matrix design from an array of permutation matrices from [23]. We designed an $(2, 3)$ -regular LDPC matrix with $m = 159$ and $n = 265$ and then substituted the non zero elements in H by a random positive number to obtain the matrix A .

For each sparsity k , at least 75 random k -sparse signals \mathbf{x} are generated and 50 reconstruction iterations are performed. The proportion of correct reconstruction results are summarized in the plot of the Fig. 4 for the IPA, the verification decoding algorithm, and the complex LP. We can see then that the IPA is a good trade-off between performance and complexity.

IV. ANALYSIS

In this section, we study the recovery of the IPA on non-negative real-valued signals. We present the analysis on binary measurement matrices for the sake of clarity, but the extension of these results to non-negative real-valued measurement matrices is straightforward. First, we give a theorem given in [15] which proves the failure of the IPA on stopping sets. A stopping set is defined as follows.

Definition 1 ([24]). *A stopping set T is a subset of the set of variable nodes V such that all neighbors of T are connected to T at least twice.*

The cardinality of a stopping set is called the size of the stopping set.

Theorem 1 ([15]). *Let $A_{m \times n}$ be a binary measurement matrix. The IPA fails on the recovery of a signal \mathbf{x} if the non-zero entries contain a stopping set in A .*

Proof: We prove that if all variable nodes in T have non-zero values, the IPA cannot recover them. In other words, we prove that the bounds of the intervals passing through the edges of the graphical representation of A never converges, i.e. we show that $\forall v \in T$ such that $x(v) > 0$, then $\mu_{v \rightarrow c}^{(l)} < x(v) < M_{c \rightarrow v}^{(l)}$, $\forall l \geq 0, \forall c \in \mathcal{N}(v)$. Suppose $\forall v \in T, x(v) > 0$, from the definition of a stopping set $\forall c \in \mathcal{N}(v), y(c) > x(v)$. Then, at the initialization ($l = 0$) we have:

$$\mu_{c \rightarrow v}^{(0)} = 0 < x(v) < y(c) = M_{c \rightarrow v}^{(0)} \quad (7)$$

At the first iteration we have:

$$\mu_{v \rightarrow c}^{(1)} = \max_{c \in \mathcal{N}(v)} \left(\mu_{c \rightarrow v}^{(0)} \right) = 0 < x(v) \quad (8)$$

$$M_{v \rightarrow c}^{(1)} = \min_{c \in \mathcal{N}(v)} \left(M_{c \rightarrow v}^{(0)} \right) > x(v) \quad (9)$$

Now consider the update at the measurement node. For the lower bound of the interval we have:

$$\mu_{c \rightarrow v}^{(1)} = \max \left(0, y(c) - \sum_{v' \in \mathcal{N}(s) \setminus \{v\}} M_{v' \rightarrow c}^{(1)} \right) \quad (10)$$

$$< \left(y(c) - \sum_{v' \in \mathcal{N}(s) \setminus \{v\}} x(v') \right) \quad (11)$$

$$= x(v) \quad (12)$$

The last equation stands from the initialization $x(v) < y(c) = M_{v \rightarrow c}^{(1)}$. Similarly, for the upper bound at the measurement node we have:

$$M_{c \rightarrow v}^{(1)} = y(c) - \sum_{v' \in \mathcal{N}(s) \setminus \{v\}} \mu_{v' \rightarrow c}^{(1)} \quad (13)$$

$$> y(c) - \sum_{v' \in \mathcal{N}(s) \setminus \{v\}} x(v') \quad (14)$$

$$= x(v) \quad (15)$$

The last equation results from the initialization $x(v) > y(c) = \mu_{v \rightarrow c}^{(1)}$.

Thus we obtain $\mu_{c \rightarrow v}^{(1)} < x(v) < M_{c \rightarrow v}^{(1)}$. For $l > 1$ the messages from variable nodes to check nodes are simply the intersection of intervals from the check nodes at the previous iteration, and then we still have $\mu_{v \rightarrow c}^{(l)} < x(v) < M_{v \rightarrow c}^{(l)}$. The proof is completed by induction for every $l > 0$. ■

Theorem 1 also indicates that the IPA fails on reconstruction of a signal \mathbf{x} whose non-zero values form the smallest stopping set in the measurement matrix A . However, as we explain in the following example, the smallest stopping set is not the smallest configuration on which the IPA fails.

Example 1. Consider a stopping set of size four as given in Fig. 5. According to Theorem 1, the IPA cannot recover a signal \mathbf{x} with non-zero values on $\{v_1, v_2, v_3, v_4\}$. The algorithm also fails on a 2-sparse signal \mathbf{x} whose non-zero values are $\{v_1, v_3\}$ or $\{v_2, v_4\}$, which implies that the variable

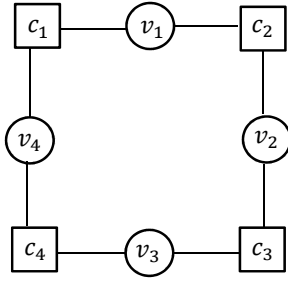


Fig. 5: Stopping set of size 4

nodes forming a smallest stopping set are not necessarily the smallest configuration on which the IPA fails. However, the IPA can recover a 2-sparse signal with non-zero values on $\{v_1, v_2\}$ or $\{v_1, v_4\}$.

Example 1 shows the main difference between the iterative decoding on the BEC and the signal recovery of the iterative IPA in compressed sensing. The iterative decoder over the BEC fails if and only if the erasures contain a stopping set, while the IPA fails even if the non-zero values do not involve a stopping set. The following results identify recoverable signals whose non-zero values are subsets of stopping sets. First, we show that every zero-value variable node is recoverable by the IPA. In this paper, we say that a node is zero if its value is equal to zero.

Lemma 1. *The IPA can recover all zero variable nodes.*

Proof: Suppose v is a variable node with value 0 and $\{c_1, c_2, \dots, c_{d_v}\}$ is the set of the d_v measurement neighboring nodes of v with measurement values $\{\alpha_1, \alpha_2, \dots, \alpha_{d_v}\}$. At each iteration of the IPA, the message which is sent from c_j ($j = 1, \dots, d_v$) to v is either $[0, 0]$ or $[0, \beta_j]$ where $0 < \beta_j \leq \alpha_j$. If v receives at least one $[0, 0]$ from one of its neighbors, the value of v is recovered as 0. If all messages from every c_j to v are $[0, \beta_j]$, the decision rule of the algorithm leads to recover the value of v to the maximum value of lower bounds of the intervals $[0, \beta_j]$, which is 0. ■

Since all zero variable nodes are recovered by the IPA, it is enough to study the recovery of non-zero variable nodes.

Definition 2. *A set of variable nodes S is called a minimal stopping set, if S forms a stopping set and it does not contain a smaller stopping set.*

It is clear that the smallest stopping set in a measurement matrix A is a minimal stopping set while a minimal stopping set is not necessarily the smallest stopping set. The size of the smallest stopping set is called the *stopping distance* [25] and plays a significant role in iterative decoding of LDPC codes over the BEC.

Theorem 2. *Let $A_{m \times n}$ be a binary measurement matrix and $V_S = \{v_1, v_2, \dots, v_k\}$ be a subset of variable nodes forming a minimal stopping set. Let $\mathbf{x} = [x_1, x_2, \dots, x_n]^t \in \mathbb{R}_{\geq 0}^n$ be a signal with at most $k - 2$ non-zero values, i.e. $\|\mathbf{x}\|_0 \leq k - 2$ such that the set of non-zero variables is a subset of V_S . Then, the IPA can recover \mathbf{x} if there exists at least one zero measurement node among the neighbors of V_S .*

Proof: Let $M = \{c_1, c_2, \dots, c_q\}$, $1 \leq q \leq m$ be the set of measurement nodes connected to V_S . Suppose that c_1 is the only zero measurement node. Since V_S forms a stopping set, there exist at least two zero variable nodes, say v_1, v_2 , connecting to c_1 and there exist non-zero value measurement nodes connected to v_1, v_2 . Moreover, V_S is a minimal stopping set, so there exists at least one measurement node, $c_2 \in M \setminus \{c_1\}$ with only one neighbor in $V_S \setminus \{v_1, v_2\}$, namely v_3 . Otherwise, $V_S \setminus \{v_1, v_2\}$ will be a smaller stopping set, which is a contradiction. Suppose c_2 has the value α . At the first iteration, c_1 sends $[0, 0]$ to v_1 and v_2 and v_1 and v_2 send $[0, 0]$ to their neighbors. Based on the above explanation, c_2 is a neighbor of v_1 or v_2 or both of them with only one neighbor in $V_S \setminus \{v_1, v_2\}$. At the second iteration, c_2 will send $[\alpha, \alpha]$ to the only variable node v_3 . So, the value of the variable node v_3 is recovered to α . Since the value of this variable node is recovered, we can consider this variable node to have value zero by subtracting α from the value of all measurement nodes which are the neighbors of v_3 . Now, we have three variable nodes v_1, v_2, v_3 whose values have been determined. Again, with the same discussion, there exists a measurement node $c_3 \in M \setminus \{c_1, c_2\}$ with only one neighbor in $V_S \setminus \{v_1, v_2, v_3\}$ which can be recovered in the next iteration. Continuing the same process will recover all variable nodes. ■

Example 2. *Fig. 6 illustrates the previous theorem considering the recovery of a signal with 4 non-zero variable nodes in a minimal stopping set of size 6 in which there exists one zero measurement node c_1 and two zero variable nodes v_1 and v_2 . Let $\alpha, \beta, \gamma, \delta, \eta$ be the non-zero values of c_2, c_3, c_4, c_5 and c_6 , respectively. Note that c_2 is one the measurement nodes with exactly one neighbor among $\{v_3, v_4, v_5, v_6\}$. At initialization, the zero measurement node c_1 sends $[0, 0]$ to v_1 and v_2 . And c_2 sends $[0, \alpha]$ to v_2 and v_3 . At the first iteration, v_2 sends $[0, 0]$ to c_2 which causes c_2 sends $[\alpha, \alpha]$ to v_3 . Thus, the value of v_3 is recovered as α . Another measurement node with only one neighbor in $\{v_4, v_5, v_6\}$ is c_3 . Again this measurement node sends $[\beta, \beta]$ to v_4 and so the value of v_4 is recovered as β . The same process results in the recovery of all variable nodes.*

In Theorem 2 we proved that the existence of at least one zero measurement node is enough to reconstruct a signal \mathbf{x} whose non-zero values are a subset of a minimal stopping set. As we will show in the following Lemma and Corollary, in regular measurement matrices we can give an upper bound on the number of variable nodes forming a subset of a minimal stopping set S such that there exists at least one measurement node among the neighbors of S with no connection to this subset of variable nodes. This result shows that in a minimal stopping set, if the number of non-zero values is bounded by a fixed number, the IPA can recover the signal \mathbf{x} .

Lemma 2. *Let $A_{m \times n}$ be a binary (d_v, d_c) -regular measurement matrix and let $g = 2r$ be the girth of the Tanner graph corresponding to A . Then, every subset N of variable nodes such that*

$$|N| < \frac{\sum_{i=0}^{r-1} (d_v - 1)^{\lceil \frac{i}{2} \rceil} (d_c - 1)^{\lfloor \frac{i}{2} \rfloor}}{d_v} \quad (16)$$

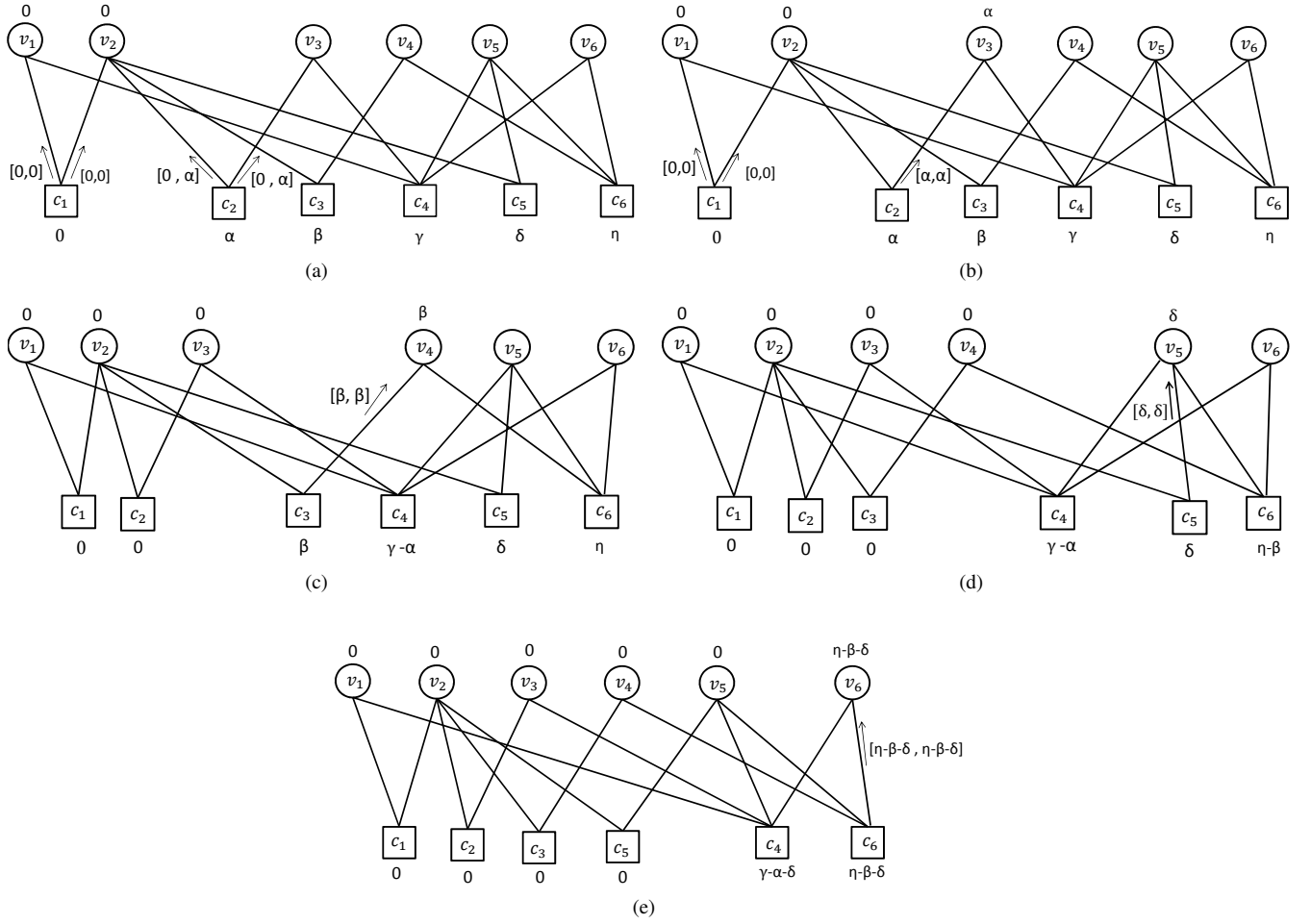


Fig. 6: Reconstruction of a signal with non-zero elements in a minimal stopping set of size 6 with the IPA when there exists one zero measurement node. (a) the first messages sending from measurement nodes to variable nodes, (b) messages from the measurement nodes to variable nodes in the first iteration (recovery of v_3), (c) the message to v_4 in the second iteration (recovery of v_4), (d) the message to v_5 in the third iteration (recovery of v_5), (e) the message to v_6 in the fourth iteration (recovery of v_6).

contains a measurement node which does not have any neighbor in N .

Proof: The proof is obtained by using lower bounds on the number of measurement nodes given in [26]. According to this bound,

$$m \geq \sum_{i=0}^{r-1} (d_v - 1)^{\lceil \frac{i}{2} \rceil} (d_c - 1)^{\lfloor \frac{i}{2} \rfloor}. \quad (17)$$

Now, suppose N is a subset of variable nodes. Since the matrix is regular, the maximum number of measurement nodes that can be connected to the variable nodes in N is at most $|N|d_v$. If $|N|d_v < \sum_{i=0}^{r-1} (d_v - 1)^{\lceil \frac{i}{2} \rceil} (d_c - 1)^{\lfloor \frac{i}{2} \rfloor}$ which results $|N| < \frac{\sum_{i=0}^{r-1} (d_v - 1)^{\lceil \frac{i}{2} \rceil} (d_c - 1)^{\lfloor \frac{i}{2} \rfloor}}{d_v}$, there exists at least one measurement node which does not have any connection to N . ■

Corollary 1. Suppose $A_{m \times n}$ is a binary (d_v, d_c) -regular measurement matrix with girth $g = 2r$. Let N be a subset

of k variable nodes that forms a minimal stopping set. If $|N| < \frac{\sum_{i=0}^{r-1} (d_v - 1)^{\lceil \frac{i}{2} \rceil} (d_c - 1)^{\lfloor \frac{i}{2} \rfloor}}{d_v}$, then the IPA can recover a signal \mathbf{x} with non-zero values in N . In the case that the girth is 6, $|N|$ is bounded by $|N| \leq \frac{(d_v - 1)d_c}{d_v} = \frac{(d_v - 1)k}{m}$. If the girth is 8, $|N|$ is bounded by $|N| \leq \frac{(d_v - 1)d_c + (d_v - 1)(d_c - 1)}{d_v}$.

The following theorem gives a sufficient condition on exact recovery of a signal whose support is a subset of a minimal stopping set and all neighboring measurement nodes are non-zero.

Theorem 3. Let $A_{m \times n}$ be a binary measurement matrix and $V_S = \{v_1, v_2, \dots, v_k\}$ be a subset of variable nodes forming a minimal stopping set. Let $\mathbf{x} = [x_1, x_2, \dots, x_n]^t \in \mathbb{R}_{\geq 0}^n$ be a signal with at most $k - 1$ non-zero values, i.e. $\|\mathbf{x}\|_0 \leq k - 1$ such that the set of non-zero variables is a subset of V_S . Suppose all measurement nodes have non-zero values. Then, the IPA can recover \mathbf{x} if

1. There exists at least one measurement node c_j such that

the variable nodes $\{v_1, v_2, \dots, v_p\}$ which are connected to c_j have non-zero values and do not share a measurement node other than c_j .

2. The measurement nodes $\{c_1, c_2, \dots, c_l\}$ connected to $\{v_1, v_2, \dots, v_p\}$ do not have non-zero neighboring variable nodes excluding $\{v_1, v_2, \dots, v_p\}$.

Proof: Suppose $\{v_1, v_2, \dots, v_p\}$ have non-zero values $\{\alpha_1, \alpha_2, \dots, \alpha_p\}$. Since $\{c_1, c_2, \dots, c_l\}$ are connected to zero variable nodes except for $\{v_1, v_2, \dots, v_p\}$ and $\{c_1, c_2, \dots, c_l\}$ are not shared by more than one variable node in $\{v_1, v_2, \dots, v_p\}$, the value of every measurement node in $\{c_1, c_2, \dots, c_l\}$ lies in $\{\alpha_1, \alpha_2, \dots, \alpha_p\}$. There exists a measurement node in $\{c_1, c_2, \dots, c_l\}$ that has only one neighbor in $V_S \setminus \{v_1, v_2, \dots, v_p\}$. Otherwise, $V_S \setminus \{v_1, v_2, \dots, v_p\}$ will be a smaller stopping set. Without loss of generality, suppose c_1 is a measurement node with this property which is connected to v_1 with the value α_1 . In the first iteration, c_j sends $\sum_{i=1}^p \alpha_i$ to its neighbors in $\{v_1, v_2, \dots, v_p\}$ and c_1 sends $[0, \alpha_1]$ to its neighbors. In the second iteration, c_j sends $[\alpha_1, \sum_{i=1}^p \alpha_i]$ to v_1 and c_1 and other neighbors of v_1 send intervals with the upper bound α_1 to v_1 which results that the message $[\alpha_1, \alpha_1]$ is sent from v_1 to its neighbors. Thus, the value of v_1 is recovered as α_1 and all measurement nodes which are the neighbors of v_1 are satisfied. So, c_1 is satisfied and its value can be considered as zero. The recovery of other variable nodes is followed from Theorem 2 which implies that existence of at least one zero measurement node is enough to recovery of all variable nodes in a configuration forms as a stopping set. ■

The following example shows how the IPA can recover a vector \mathbf{x} under the conditions of the Theorem 3.

Example 3. Fig. 7 depicts the recovery of the three non-zero variable nodes in $\{v_1, v_2, v_3\}$ a minimal stopping set of size 5 in which all measurement nodes are non-zero. First, note that the measurement node c_1 and the variable nodes v_1, v_2, v_3 satisfy the two conditions of Theorem 3. Thus, if the non-zero variable nodes v_1, v_2, v_3 have values α, β, γ , then c_1 has value $\lambda = \alpha + \beta + \gamma$ and other measurement nodes c_2, c_3, c_4, c_5 have the values α, β, γ and γ respectively. For simplicity, we just show how the value of the variable node v_3 is recovered. At initialization, c_1 sends $[0, \alpha + \beta + \gamma]$ to v_3 and c_5 sends $[0, \gamma]$ to v_3 . In the first iteration, c_1 receives $[0, \alpha]$ and $[0, \beta]$ from v_1 and v_2 , respectively and c_5 receives $[0, \min(\alpha, \gamma)]$ from v_5 . Then, c_1 sends $[\gamma, \alpha + \beta + \gamma]$ to v_3 and c_5 sends $[0, \gamma]$ or $[\gamma - \min(\alpha, \gamma), \gamma] = [\gamma - \alpha, \gamma]$ to v_3 . In the second iteration, v_3 sends $[\gamma, \gamma]$ to its neighbors which makes c_5 is satisfied and can be considered as a zero measurement node. Now, there exists a zero measurement node in this minimal stopping set. Theorem 2 results the recovery of other variable nodes.

Theorems 2 and 3 give sufficient conditions on the recovery of signals whose non-zero values form a subset of a minimal stopping set. To show how small stopping sets affect the performance of the IPA, we provide simulation results of the performance of the IPA in the next section.

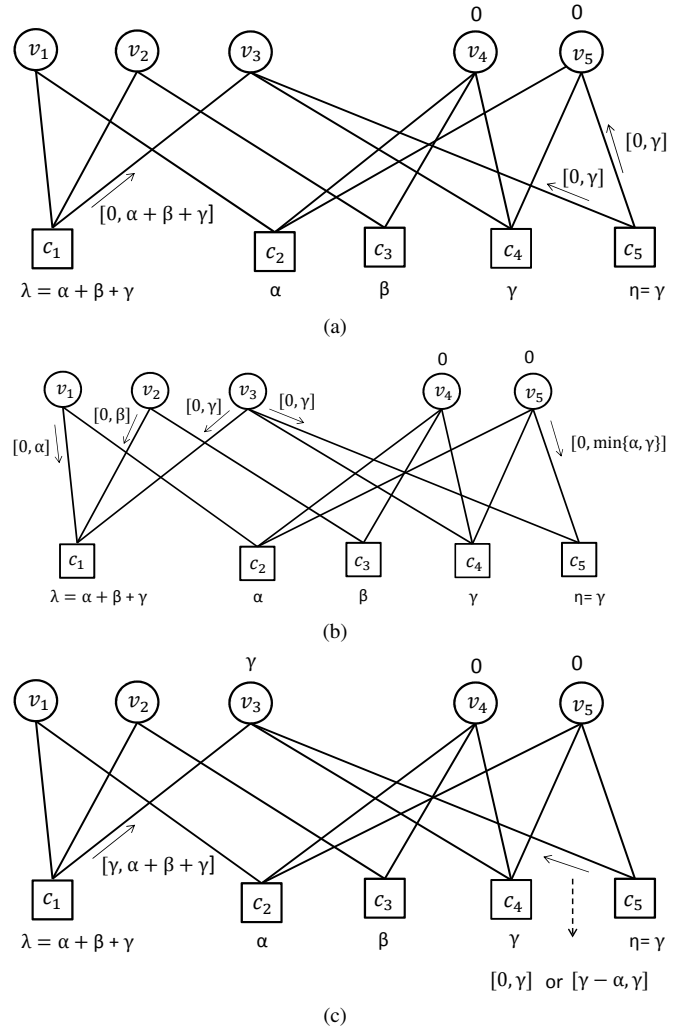


Fig. 7: Reconstruction of a signal with non-zero elements in a support of a stopping set of size 5 with IPA and with no zero measurement node. (a) the first messages which are sent from check nodes to variable nodes, (b) messages from variable nodes in the first iteration, (c) messages from check nodes in the second iteration.

V. SIMULATION RESULTS

In this section we provide simulation results of the reconstruction performance of the IPA by establishing the link with the analysis done previously related to stopping sets. From Theorem 1, we know that if the support of the signal \mathbf{x} is, or contains, a stopping set, the IPA cannot recover it. The stopping sets in the measurement matrix A are then responsible for most of the failures of the reconstruction via the IPA.

Orlitsky *et al.* derived a formula to obtain the average distribution $E(s)$ for any size s of stopping sets for a (d_v, d_c) -regular LDPC code of size n [27]:

$$E(s) = \binom{n}{s} \frac{\text{coef} \left(\left((1+x)^{d_c} - x d_c \right)^{\frac{d_v}{d_c} n}, x^{s d_v} \right)}{\binom{n d_c}{s d_v}} \quad (18)$$

TABLE I: Stopping distance s_{min} for the length $n = 768$ LDPC codes from the IEEE802.16 standard with different rates R . The number of stopping sets of weight s_{min} is denoted $N_{s_{min}}$ (from [28] and [30])

R	s_{min}	$N_{s_{min}}$	$N_{s_{min}+1}$	$N_{s_{min}+2}$	$N_{s_{min}+3}$	$N_{s_{min}+4}$
1/2	14	32	32	0	32	32
2/3 A	8	64	0	0	160	160
2/3 B	12	64	128	384	1120	3352
3/4 A	9	32	128	576	2192	9696
3/4 B	4	16	0	0	32	216
5/6	6	96	672	5376	36512	280128

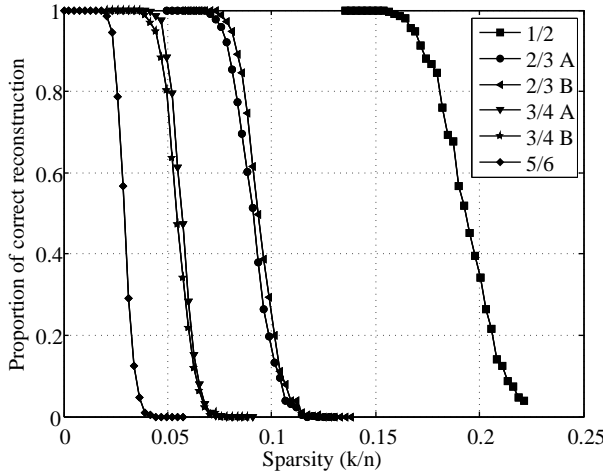


Fig. 8: IPA performance on the IEEE 802.16e LDPC codes of length $n = 768$.

where $\text{coef}(p(x), x^i)$ denotes the coefficient of x^i in the polynomial $p(x)$. The average distribution of stopping sets can theoretically be computed using the previous formula, however it has to be done by numerical methods due to its large complexity. To be able to practically find stopping sets, algorithms have been proposed as in [28] and [29] in the channel coding context. In their recent work, Rosnes *et al.* [28], [30] provide the stopping set repartition on various LDPC codes based on their algorithm to find small stopping sets. In [30] they focused more specifically on the LDPC codes from the IEEE 802.16e standard [31], referred as the WiMax codes. These codes are circulant-based LDPC codes, and the IEEE standard provides the design of codes for 19 different lengths. Also, one model matrix to design codes with rates 1/2, and 5/6 is provided, and two model matrices are provided for codes with rates 2/3 and 3/4 (denoted by A and B). We generated all the codes of length $n = 768$ according to the IEEE standard and we remind the stopping set distribution of these codes in Table I in which we adopted the notation of Rosnes *et al.* [30]. We used these 6 codes as measurement matrices and simulate the recovery performance via the IPA. The simulation results are shown in Fig. 8 where the proportion of correct reconstruction of sparse signals are plotted versus the sparsity measure k/n . For each sparsity k and for each matrix, 500 k -sparse signals are generated, and a maximum of 50 iterations of the IPA for the reconstruction are done.

These results emphasize the connection between the stop-

ping set distribution and the performance of the IPA from the theorems of the previous section. For instance, it is clear from Table I and the Fig. 8 that the stopping set spectra are responsible for most of the failures of the IPA. Indeed at a constant rate (e.g. 2/3), the measurement matrix with higher stopping distance has a slightly better performance. However, we can see that even if the matrix A with rate 3/4 has a better stopping distance than the rate 5/6 matrix, it performs better. This observation comes from the stopping set spectra as matrix A with rate 3/4 has only 16 stopping sets of weight 4, and the next ones have weight 7 (and there are a few of them) whereas for the rate 5/6, the number of stopping sets of weight 6 or 7 are very numerous. Then, although the stopping set spectra gives an insight of the performance of the IPA on a given measurement matrix, it is not obvious to foresee this performance because it depends on the stopping distance and on the number of stopping sets of each weight.

VI. CONCLUSION AND DISCUSSION

In this paper, we described in details the IPA that we modified in order to deal with the non-negative measurement matrices. As expected this algorithm presents better results than the simple verification decoding, and is less complex than the l_1 -minimization reconstruction via the LP. We also provided an analysis for the signal recovery of the IPA on stopping sets. Indeed the IPA fails to recover non-zero values when these values correspond with a stopping set of the measurement matrix. However we provided results that give sufficient conditions on which the IPA can recover a k' -sparse signal in a minimal stopping set of size k whose support is a subset of the stopping set and $k' < k$. Simulation results on LDPC codes whose stopping set distribution is known demonstrate the relation between the stopping sets and the performance of the IPA. However the exact influence of the stopping distance and the number of stopping sets present in the measurement matrix will be addressed in future publication.

Future work also includes to analyze the failure of the IPA in presence of noise and find a scheme to recover at least the support of the signal \mathbf{x} . Finding configurations on which the IPA fails when the measurements are noisy might help to construct better measurement matrices free of topological structures that leads to failures of reconstruction.

ACKNOWLEDGMENT

The authors thank Dung V. Nguyen for interesting and inspiring conversations.

APPENDIX A VERIFICATION DECODING

In the verification algorithm each variable node can have two states; one *unverified state* [*] (no value has yet been estimated at the corresponding variable node), and one *verified state* (the variable node has been estimated). Once a variable node has been estimated the assigned value cannot be modified.

The different steps of the algorithm are summarized in the

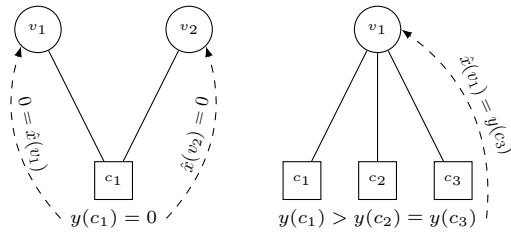


Fig. 9: Verification decoding algorithm: First Step (Left); Third Step (Right)

next four steps.

Step 1. Verify as 0 the variable nodes which are the neighbors of zero-value measurement nodes.

Step 2. Verify variable nodes connected to measurement nodes with degree one (only one edge connected) as the value of the measurement node.

Step 3. Verify a single variable node connected to two measurement nodes with the same measurement value as the common value of the measurement nodes.

Step 4. Subtract the values of the verified variable nodes from the neighboring measurement nodes and then remove all verified variable nodes and edges connected to them on the factor graph.

Steps 1 to 4 are repeated until the success of the reconstruction or until no more progress in the reconstruction is witnessed. The Fig. 9 sketch the Steps 1 and 3. The justification of Step 3 is based on the observation that over large alphabets the probability that two independent measurements are equal is quite small. Then, any two common measured values are likely to be generated by the same set of non-zero data. This observation holds for large alphabets and for real numbers too.

REFERENCES

- [1] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [2] —, "For most large underdetermined systems of linear equations, the minimal ℓ_1 norm solution is also the sparsest solution," *Comm. on Pure and Applied Math.*, vol. 59, no. 6, pp. 797–829, Jun. 2006.
- [3] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, pp. 227–234, 1995.
- [4] J. Feldman, M. Wainwright, and D. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 954–972, Mar. 2005.
- [5] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, Aug. 1998.
- [6] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [7] S. Sarvotham, D. Baron, and R. G. Baraniuk, "Compressed sensing reconstruction via belief propagation," Tech. Rep., 2006.
- [8] H. V. Pham, W. Dai, and O. Milenkovic, "Sublinear compressive sensing reconstruction via belief propagation decoding," in *Proc. IEEE Int. Symp. Inform. Theory*, Seoul, Korea, Jun. 28–Jul. 3 2009, pp. 674–678.
- [9] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing," *Proc. Nat. Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, Sep. 2009.
- [10] V. Chandar, D. Shah, and G. Wornell, "A simple message-passing algorithm for compressed sensing," in *Proc. IEEE Int. Symp. Inform. Theory*, Austin, TX, USA, Jun. 13–18 2010, pp. 1968–1972.
- [11] V. Ravanmehr, L. Danjean, D. Declercq, and B. Vasic, "On iterative compressed sensing reconstruction of sparse non-negative vectors," in *Proc. 4th Int. Symp. on Appl. Sci. in Biomedical and Commun. Technologies*, Barcelona, Spain, Oct. 26–29 2011.
- [12] F. Zhang and H. Pfister, "Compressed sensing and linear codes over real numbers," in *Inf. Theory and Applicat. Workshop*, Feb. 2008, pp. 558–561.
- [13] F. Zhang and H. D. Pfister, "On the iterative decoding of high-rate LDPC codes with applications in compressed sensing," *IEEE Trans. Inf. Theory*, (to appear) 2012. [Online]. Available: <http://arxiv.org/abs/0903.2232>
- [14] L. Danjean, V. Ravanmehr, D. Declercq, and B. Vasic, "Iterative reconstruction algorithms in compressed sensing," in *Proc. 19th Telecommun. Forum*, Belgrade, Serbia, Nov. 22–24 2011, pp. 537–541.
- [15] A. R. Krishnan, S. Sankaraman, and B. Vasic, "Graph-based iterative reconstruction of sparse signals for compressed sensing," in *Proc. 10th Int. Conf. on Telecommun. in Modern Satellite, Cable and Broadcasting Services*, Nis, Serbia, Oct. 5–8 2011, pp. 133–137.
- [16] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA, USA: M.I.T. Press, 1963.
- [17] R. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 533–547, Sep. 1981.
- [18] A. G. Dimakis and P. Vontobel, "LP decoding meets LP decoding: a connection between channel coding and compressed sensing," in *Proc. 47th Allerton Conf. on Commun., Control, and Computing*, Monticello, Illinois, USA, Sep. 30–Oct. 2 2009, pp. 8–15.
- [19] A. G. Dimakis, R. Smarandache, and P. O. Vontobel, "LDPC codes for compressed sensing," *submitted to IEEE Trans. Inf. Theory*, Dec. 2010. [Online]. Available: <http://arxiv.org/abs/1012.0602>
- [20] A. Maleki and A. Montanari, "Analysis of approximate message passing algorithm," in *Proc. 44th Annu. Conf. on Inform. Sci. and Syst.*, Princeton, NJ, USA, Mar. 17–19 2010, pp. 1–7.
- [21] M. C. Davey and D. J. C. MacKay, "Low-density parity check codes over $GF(q)$," *IEEE Commun. Lett.*, vol. 2, no. 6, pp. 165–167, Jun. 1998.
- [22] D. Declercq and M. Fossorier, "Decoding Algorithms for Nonbinary LDPC Codes Over $GF(q)$," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 633–643, Apr. 2007.
- [23] D. V. Nguyen, S. K. Chilappagari, B. Vasic, and M. W. Marcellin, "On the construction of structured LDPC codes free of small trapping sets," *IEEE Trans. Inf. Theory*, (to appear) 2012. [Online]. Available: <http://arxiv.org/abs/1008.4177>
- [24] C. Di, D. Proietti, I. Telatar, T. Richardson, and R. Urbanke, "Finite-length analysis of low-density parity-check codes on the binary erasure channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1570–1579, Jun. 2002.
- [25] M. Schwartz and A. Vardy, "On the stopping distance and the stopping redundancy of codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 922–932, Mar. 2006.
- [26] S. Hoory, "The size of bipartite graphs with a given girth," *J. Combinatorial Theory Series B*, vol. 86, no. 2, pp. 215–220, Nov. 2002.
- [27] A. Orlitsky, K. Viswanathan, and J. Zhang, "Stopping set distribution of LDPC code ensembles," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 929–953, Mar. 2005.
- [28] E. Rosnes and O. Ytrehus, "An efficient algorithm to find all small-size stopping sets of low-density parity-check matrices," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4167–4178, Sep. 2009.
- [29] G. Richter, "Finding small stopping sets in the tanner graphs of LDPC codes," Munich, Germany, Apr. 3–7 2006.
- [30] E. Rosnes, O. Ytrehus, M. Ambroze, and M. Tomlinson, "Addendum to: An efficient algorithm to find all small-size stopping sets of low-density parity-check matrices," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 164–171, Jan. 2012.
- [31] "IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands and corrigendum 1," *IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004)*, Feb 2006.