Signal Recovery Performance of the Interval-Passing Algorithm

Vida Ravanmehr†, Ludovic Danjean†‡, Bane Vasić†, David Declercq‡

† Dept. of Electrical and Computer Eng.
University of Arizona
Tucson, AZ 85721, USA
Email: {vravanmehr,danjean,vasic}@ece.arizona.edu

‡ ETIS
ENSEA / Univ. Cergy-Pontoise / CNRS
F-95000 Cergy-Pontoise, France
Email: declercq@ensea.fr

Abstract—This paper considers an iterative algorithm called the Interval-Passing Algorithm (IPA) which is used to reconstruct non-negative real signals using binary measurement matrices in compressed sensing (CS). The failures of the algorithm on stopping sets, also non-decodable configurations in iterative decoding of LDPC codes over the binary erasure channel (BEC), shows a connection between iterative reconstruction algorithm in CS and iterative decoding of LDPC codes over the BEC. In this paper, a stopping-set based approach is used to analyze the recovery of the IPA. We show that a smallest stopping set is not necessarily a smallest configuration on which the IPA fails and provide sufficient conditions under which the IPA recovers a sparse signal whose non-zero values lie on a subset of a stopping set. Reconstruction performance of the IPA using IEEE 802.16e LDPC measurement matrices are provided to show the effect of the stopping sets in the performance of the IPA.

I. INTRODUCTION

Compressed sensing (CS) is a relatively new field in signal processing, which is concerned with the recovery of large sparse signals from a small set of measurements. The straightforward approach to reconstruct sparse signals is based on ℓ₀-norm minimization and is known to be NP-hard [1], [2]. Relaxing the ℓ₀-norm minimization to ℓ₁-norm minimization leads to a linear programming technique to solve the CS problem called Basis Pursuit [3]. The relation between Low-Density Parity-Check (LDPC) codes and CS was studied by Dimakis et al. in [4] and [5] where the authors showed that good LDPC matrices for LP decoding are also good for CS using Basis Pursuit. Despite its remarkable performance, ℓ₁-norm minimization suffers from a problem related to its complexity. The high complexity of the algorithm and consequently its large running time make it impractical especially in the case of large measurement matrices. Instead, iterative reconstruction algorithms, inspired by the iterative decoding algorithms of LDPC codes, were developed for the CS problem. Sarvotham et al. [6] introduced an iterative algorithm based on belief propagation to reconstruct approximately sparse signals. In addition, they proposed the Sudocodes algorithm to reconstruct strictly sparse signals in [7]. Two low-complexity algorithms, list decoding and multiple-basis belief propagation were proposed by Pham et al. in [8]. These algorithms are another application of belief propagation in CS with good performance to reconstruct sparse signals. Zhang and Pfister [9] proposed two algorithms, LM1 and LM2 based on verification decoding to reconstruct strictly sparse signals. In [10], Donoho et al. proposed an iterative thresholding algorithm called the approximate message passing algorithm to reconstruct sparse signals from noise-free measurements using random Gaussian measurement matrices. This algorithm can be viewed as the variant of bit-flipping algorithm which is used for decoding of LDPC codes.

A message passing algorithm was introduced by Chandar et al. [11] to reconstruct non-negative real-valued signals using binary measurement matrices. This simple iterative algorithm with low-complexity and consequently low running time is suitable to reconstruct a long non-negative sparse signal. This algorithm, referred to as the Interval-Passing Algorithm (IPA), was modified by Krishnan et al. in [12], where the authors showed that the IPA fails on stopping sets. Stopping sets are the non-decodable configurations in the context of decoding of LDPC codes over the binary measure channel (BEC) [13]. The performance of the IPA in both noise-free and noisy measurements was studied in [14] and a comparison between the IPA, verification algorithm and Basis Pursuit was shown in [15]. A generalization of the IPA to use non-negative measurement matrices and reconstruction analysis of the algorithm were presented in [16].

In this paper, we analyze the recovery of the IPA on 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. In Section II, we describe the IPA and give some required definitions and concepts. In Section III, we provide an analysis of the recovery of the IPA where sufficient conditions for exact recovery of signals whose indices of non-zero values form a subset of stopping sets are given. In Section IV, we provide simulation results exhibiting the IPA reconstruction performance using some LDPC measurement matrices. Section V gives the summary of the results and provides possible directions for future work.

II. PRELIMINARIES AND PROBLEM SETTING

Iterative message passing algorithms, used for decoding of LDPC codes, were also developed in CS. Iterative algorithms can be easily described on the Tanner graph [17] of the
message matrix. The Tanner graph $T$ corresponding to the measurement matrix $A$ is a bipartite graph with two sets of nodes: variable nodes and measurement nodes. Each variable (resp. measurement) node represents a column (resp. row) of $A$. Let $A_{m \times n}$ be a binary measurement matrix, $V$ be the set of $n$ variable nodes and $C$ be the set of $m$ measurement nodes in the Tanner graph corresponding to $A$, respectively. There exists an edge between the variable node $v_i$ and the measurement node $c_j$ if and only if $A_{ij} = 1$. The set of measurement nodes connected to the variable node $v$ are called the neighbors of $v$ and is denoted by $N(v)$. Similarly, the neighbors of a measurement node $c$ are denoted by $N(c)$. For a subset $S$ of variable nodes $V$, $N(S)$ is the set of measurement nodes which are connected to the variable nodes in $S$.

Let $x \in \mathbb{R}^n$, be a signal which is observed indirectly through a shorter measurement vector $y \in \mathbb{R}^m$ such that $m \ll n$ and obtained from the linear equations $y = Ax$. In the IPA, messages passing through edges in the Tanner graph are closed intervals of non-negative real numbers as $[L, U]$. Updating messages is done as follows. First, every measurement node sends messages to its neighboring variable nodes. That is at $t = 0$, $L_i^{(0)} = 0$ and $U_i^{(0)} = y$ where $y$ is a component of $y$. For $t > 0$, the messages from the variable node $v$ to the measurement node $c$ are updated as:

$$L_{v\rightarrow c}^{(t+1)} = \max_{c' \in N(v)} \left(L_{c'\rightarrow v}^{(t)}\right),$$

$$U_{v\rightarrow c}^{(t+1)} = \min_{c' \in N(v)} \left(U_{c'\rightarrow v}^{(t)}\right),$$

and the messages from the measurement node $c$ to the variable node $v$ are updated to:

$$L_{c\rightarrow v}^{(t+1)} = \max\{0, y - \sum_{v' \in N(c) \setminus \{v\}} U_{v'\rightarrow c}^{(t+1)}\},$$

$$U_{c\rightarrow v}^{(t+1)} = y - \sum_{v' \in N(c) \setminus \{v\}} L_{v'\rightarrow c}^{(t+1)}.$$

The algorithm stops if the maximum number of iterations is reached or if for each variable node, the maximum of lower bounds ($L$) and the minimum of upper bounds ($U$) are equal. If one of the above conditions holds, the value of a variable node is determined as $L$.

In [12], Krishnan et al. showed that the IPA fails on recovery a signal whose non-zero entries form a stopping set. The notion of “stopping set” was first introduced in [13] in the context of analysis the failure of iterative decoding of LDPC codes over the BEC, where the authors proved that the set of erasures which will remain undecoded when the decoder stops equals to the largest stopping set of the erasures. A stopping set is defined as follows.

**Definition 1:** [13] A subset $S$ of variable nodes $V$ is called a stopping set if the neighbors of $S$ are connected to $S$ at least twice. The cardinality of a stopping set is called the size of the stopping set.

**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 a smallest stopping set in a measurement matrix $A$ is a minimal stopping set. However, a minimal stopping set is not necessarily a smallest stopping set.

The size of the smallest stopping set is called the stopping distance [18] and plays a significant role in iterative decoding of LDPC codes over the BEC.

### III. MAIN RESULTS

In this section, we study the recovery of the IPA on nonnegative real-valued signals. More details and examples are given in [16]. First, we give a theorem given in [12] which proves the failure of the IPA on stopping sets. In this paper, we consider binary measurement matrices whose columns have at least weight two. Also, a $k$-sparse signal is referred to a signal with at most $k$ non-zero values.

**Theorem 1:** [12] Let $A_{m \times n}$ be a binary measurement matrix. The IPA fails on the recovery of a signal $x$ if the non-zero values contain a stopping set in $A$.

**Proof:** The proof was given in [12]. However, we provide a slightly different proof with more details in Appendix A. ■

Theorem 1 also indicates the failure of the IPA on reconstruction a signal $x$ whose non-zero values form a smallest stopping set in the measurement matrix $A$. However, as we explain in the following example, a smallest stopping set is not a smallest configuration on which the IPA fails. In fact, the IPA may fail even if its non-zero values do not contain any stopping set.

**Example 1:** A stopping set of size four which is also a smallest stopping set is given in Fig. 1. Theorem 1 indicates the failure of the IPA on the 4-sparse signal $x = [x_1, x_2, x_3, x_4]^t$ (non-zero values on $\{v_1, v_2, v_3, v_4\}$). The algorithm also fails on 2-sparse signals $x = [x_1, 0, x_3, 0]^t$ and $x = [0, x_2, 0, x_4]^t$ (non-zero values on $\{v_1, v_3\}$ and $\{v_2, v_4\}$, respectively), which implies that the variable nodes forming a smallest stopping set are not necessarily a smallest configuration on which the IPA fails. However, the IPA can recover other 2-sparse signals including $x = [x_1, x_2, 0, 0]^t$ and $x = [x_1, 0, 0, x_4]^t$ (non-zero values on $\{v_1, v_2\}$ and $\{v_1, v_4\}$, respectively).

Analyzing the recovery of the IPA and finding configurations on which the IPA fails or succeeds might be helpful to construct better measurement matrices or improving the algorithm in a way to avoid the failure of the algorithm. The
following results identify some 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 zero variable node and $\{c_1, c_2, ..., c_k\}$ is the set of neighbors of $v$ with values $\{\alpha_1, \alpha_2, ..., \alpha_k\}$. At each iteration of the IPA, the message which is sent from $c_j$ ($j = 1, ..., k$) to $v$ is either $[0, 0]$ or $[0, \beta_j]$ where $0 \leq \beta_j \leq \alpha_j$. If $v$ receives at least one $[0, 0]$ from one of its neighbors, the value of $v$ is recovered as 0. If all messages to $v$ are $[0, \beta_j]$, the decision rule of the algorithm recovers the value of $v$ as 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.

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

**Proof:** The proof is given in [16]. ■

**Example 2:** Fig. 2 illustrates the recovery of a signal with 3 non-zero variable nodes in a minimal stopping set of size 5 in which $c_1$ is the only zero measurement node and there are two zero variable nodes $v_1$ and $v_2$. Suppose $c_2, c_3, c_4$ and $c_5$ have the values 3, 10, 8 and 10, respectively. Lemma 1 indicates the recovery of two zero variable nodes $v_1$ and $v_2$. Since $c_2$ is a non-zero measurement node with exactly one non-zero neighbor among $\{v_3, v_4, v_5\}$, at the first iteration $[3, 3]$ is sent to $v_3$ and hence the value of $v_3$ is recovered to 3. Then, the value of $v_3$ is subtracted from the values of its neighbors $c_2, c_3, c_4$ and $c_5$. Again, there exists a measurement node $c_4$ with only one non-zero neighbor $v_5$. So, in the second iteration, the value of $v_5$ is recovered to 5. And finally, the value of $v_4$ is set to 7.

The next 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 $V_S = \{v_1, v_2, ..., v_k\}$ be a subset of variable nodes forming a minimal stopping set. Let $x = [x_1, x_2, ..., x_n]^T \in \mathbb{R}_{\geq 0}^n$ be a signal vector with at most $k - 1$ non-zero values, i.e. $\|x\|_0 \leq k - 1$ 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 $x$ if

1. There exists at least one measurement node $c_j$ such that the variable nodes $\{v_1, v_2, ..., v_p\}$ which are connected to $c_j$ have non-zero values and do not share a measurement node other than $c_j$ and
2. The measurement nodes $\{c_1, c_2, ..., c_j\}$ connected to $v_1, v_2, ..., v_p$ do not have non-zero neighboring variable nodes excluding the variable nodes $\{v_1, v_2, ..., v_p\}$.

**Proof:** The proof follows from the properties of a minimal stopping set and Theorem 2 and is given in [16]. ■

**IV. SIMULATION RESULTS**

In this section we provide simulation results of the reconstruction performance of the IPA using different binary LDPC
codes as measurement matrices. Recently, Rosnes et al. [19], [20] have provided the distribution of stopping sets for various LDPC codes based on their algorithm to find small stopping sets. In [20], they focused more specifically on the LDPC codes from the IEEE 802.16e standard [21]. 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 the codes of length \( n = 2304 \), being the largest length, for the six different rate matrices according to the IEEE standard and we remind the stopping set distribution of these codes in Table I from [20].

We used these 6 codes as measurement matrices and simulate the recovery performance via the IPA. The simulation results are shown in Fig. 3 where the proportion of correct reconstruction of sparse vectors is plotted versus the sparsity. For each sparsity \( k \) and for each matrix, 500 \( k \)-sparse vectors are generated, and a maximum of 50 iterations of the IPA for the reconstruction are done. A random \( k \)-sparse vector is said to be correctly recovered if each of its \( n \) samples are correctly estimated as close as \( 10^{-6} \).

<table>
<thead>
<tr>
<th>( R )</th>
<th>( s_{min} )</th>
<th>( N_{min} )</th>
<th>( N_{min} + 1 )</th>
<th>( N_{min} + 2 )</th>
<th>( N_{min} + 3 )</th>
<th>( N_{min} + 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>28</td>
<td>96</td>
<td>288</td>
<td>288</td>
<td>624</td>
<td></td>
</tr>
<tr>
<td>2/3</td>
<td>15</td>
<td>96</td>
<td>0</td>
<td>96</td>
<td>480</td>
<td>768</td>
</tr>
<tr>
<td>2/3 A</td>
<td>15</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2/3 B</td>
<td>12</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3/4</td>
<td>32</td>
<td>192</td>
<td>288</td>
<td>1920</td>
<td>8616</td>
<td>43584</td>
</tr>
<tr>
<td>3/4 A</td>
<td>12</td>
<td>16</td>
<td>96</td>
<td>0</td>
<td>672</td>
<td>1824</td>
</tr>
<tr>
<td>3/4 B</td>
<td>12</td>
<td>16</td>
<td>96</td>
<td>0</td>
<td>672</td>
<td>1824</td>
</tr>
</tbody>
</table>

These results emphasize the connection between the stopping set distribution and the performance of the IPA from the theorems of the previous section. For instance, it is clear from Table I and Fig. 3 that the stopping set spectra are responsible for most of the failures of the IPA. Although the stopping distance is the same at a constant rate, we can see that the IPA performance is better for matrices with less stopping sets at given sizes. For example, the measurement matrix 2/3 A performs slightly worse than the measurement matrix 2/3 B because of the presence of stopping sets of size 17, 18 and 19 whereas the matrix 2/3 B does not contain any stopping set of these sizes. Also, we note that the matrix with the largest stopping distance performs the best (measurement matrix 1/2 with \( s_{min} = 28 \)), and reciprocally, the matrix with the smallest stopping distance performs the worst (measurement matrix 5/6 with \( s_{min} = 9 \)). Although the measurement matrix with the smallest rate gives more measurements, we observed in our simulations that the majority of the failures of the IPA for all the studied matrices have an induced sub-graph corresponding to a stopping set. The precise degree of connection between the failures of the IPA and the stopping set distribution of a measurement matrix will be addressed as further work, but from this first approach it seems pretty clear that the connection is very tight.

We also plot the averaged mean-square error (MSE) as a function of the number of iterations for a fixed sparsity \( k = 150 \) in Fig. 4 which corresponds to \( k/n = 0.065 \) in Fig. 3. The mean-square estimation is done by summing over the MSE of the interval bounds of all variable nodes after a given number of iterations. If the reconstruction is successful, the minimum and maximum bounds of the interval converge to a single value, and the MSE for this variable is zero. However, if the exact recovery fails, this MSE will be non-zero. The plot of Fig. 4 shows the averaged MSE for different number of iterations (the averaging being done over the number of sparse-vector generated and the length of the sparse-vector). The overall results shows that the MSE goes quickly to zero when 100% of the generated signals are correctly recovered, and tends to a fixed value when this is not the case.


V. DISCUSSION AND FUTURE WORK

In this paper, we provide an analysis for the signal recovery of the Interval-Passing algorithm on stopping sets. The results 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 $k'<k$. We also provided the results exhibiting the performance of the IPA using different LDPC matrices with different size of a smallest stopping set.

The future work includes to analyze the failure of the IPA in the presence of noise. Finding configurations on which the IPA fails when the measurements are noisy might help to construct better measurement matrices. Moreover, extending the IPA to reconstruct real-valued signals is another interesting problem in this context.

ACKNOWLEDGMENTS

This work is funded by DARPA under KeCoProgram through contract #N66001-10-1-4079 and by the NSF under grants CCF-0963726 and CCF-0830245. The authors would like to thank the anonymous reviewers for helpful suggestions, which improved the quality of the paper.

APPENDIX A

PROOF OF THEOREM 1

Let $V_S = \{v_1, v_2, ..., v_k\}$ be a stopping set in a measurement matrix $A_{m \times n}$ and let $x = [x_1, x_2, ..., x_n]^T$ be a signal whose indices of non-zero values contain the variable nodes in $V_S$. We prove that the IPA fails in recovery of $x$. From the definition of a stopping set, we have that for all $c \in N(V_S)$, $c$ is connected to at least two variable nodes in $V_S$. Thus, at initialization,

$$L_{c \rightarrow v}^{(0)} = 0 < x(v) < U_{c \rightarrow v}^{(0)},$$

where $x(v)$ is the estimated value of the signal $x$ corresponding to the variable node $v \in V_S$.

In the first iteration, the messages from variable nodes to check nodes are:

$$L_{v \leftarrow c}^{(1)} = 0 < x(v)$$

and the messages that variable nodes send to their neighboring variable nodes are:

$$L_{c \rightarrow v}^{(1)} = \max \{0, y - \sum_{v' \in N(c) \setminus \{v\}} U_{v' \rightarrow c}^{(1)}\} < y - \sum_{v' \in N(c) \setminus \{v\}} x(v') = x(v),$$

and

$$U_{c \rightarrow v}^{(1)} = y - \sum_{v' \in N(c) \setminus \{v\}} L_{v' \rightarrow c}^{(1)} > y - \sum_{v' \in N(c) \setminus \{v\}} x(v') = x(v).$$

Thus, at the first iteration, $L_{c \rightarrow v}^{(1)} = 0 < x(v) < U_{c \rightarrow v}^{(1)}$. The same process shows that for all iterations $t > 1$, $L_{c \rightarrow v}^{(t)} = 0 < x(v) < U_{c \rightarrow v}^{(t)}$. Hence, the value of $x(v)$ never converges to a fixed value and the IPA fails on recovery of $x$.

REFERENCES


