Power Series Based Iterative Algorithms for Distributed Sparse Recovery

Shashank S, Ketan Atul Bapat, Mrityunjoy Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology Kharagpur, India
Email: ketan9318@gmail.com, shashanks@kgpian.iitkgp.ac.in, mrityun@ece.iitkgp.ac.in

***Abstract*—This paper proposes distributed versions of the recently proposed, power series based, (i) iterative hard thresholding (PSIHT), and, (ii) hard thresholding pursuit (PSHTP) algorithms for sparse signal recovery over diffusion networks. In the proposed framework, termed DiPSIHT and DiPSHTP, each node performs a local power-series-based gradient update using its local sensing matrix and measurements, and combines the updated estimates from its neighbors via diffusion weights before applying hard thresholding. An appropriate choice of the power series subsumes several existing distributed hard thresholding algorithms as special cases and also enables development of new distributed recovery algorithms. A rigorous convergence analysis of the proposed algorithms is presented via a vector inequality framework, yielding sufficient convergence conditions in terms of the restricted isometry constants (RICs) of the local sensing matrices and the spectral properties of the diffusion combination matrix. Simulation results show how by appropriate choice of the polynomials, significant performance improvement over existing methods can be achieved.**

***Index Terms*—Compressed sensing, diffusion networks, hard thresholding, power series, restricted isometry constant (RIC).**

I. Introduction

In distributed compressed sensing, a network of L nodes aims to collaboratively reconstruct a common K-sparse signal $\mathbf{x} \in \mathbb{C}^N$ from local (noisy) linear measurements

$$\mathbf{y}_v = \mathbf{A}_v\mathbf{x} + \mathbf{e}_v, \quad v = 1, \ldots, L, \tag{1}$$

where $\mathbf{A}_v \in \mathbb{C}^{m_v \times N}$, $m_v < N$, is the local sensing matrix, $\mathbf{y}_v$ is the measurement vector and $\mathbf{e}_v$ is additive noise, at the v-th node. This setup arises naturally in wireless sensor networks and multi-agent signal processing, where individual nodes have limited measurement capability but can significantly improve recovery performance through collaboration.

The centralized iterative hard thresholding (IHT) [1] and hard thresholding pursuit (HTP) [2] algorithms are widely studied, with well-established recovery guarantees under the restricted isometry property (RIP). Their distributed counterparts include the consensus-based distributed IHT (CB-DIHT) [3], the diffusion-based DiFIGHT and MoDiFIGHT [4], the distributed HTP (DHTP) [5], and others [6]. All these methods use the standard gradient $\mathbf{A}_v^H(\mathbf{y}_v - \mathbf{A}_v\mathbf{x}_v^n)$ for local updates.

Recently, Bapat and Chakraborty [7] proposed a unified power series framework for centralized hard thresholding, replacing the standard gradient update with $\mathbf{A}_v^H q(\mathbf{A}_v\mathbf{A}_v^H)(\mathbf{y}_v - \mathbf{A}_v\mathbf{x})$ for a suitable scalar power series $q(t)$. This framework subsumes IHT, HTP, pseudo-inverse based HT (PHT) [8], and Newton-step based IHT (NSIHT) [9] as special cases, and also enables new algorithms. Motivated by the well-known superiority of diffusion over consensus strategies in distributed estimation [10], this paper extends the power series framework of [7] to the distributed diffusion network setting. The main contributions made in this regard are as follows :

(i) We propose DiPSIHT and DiPSHTP, diffusion-based distributed extensions of the centralized PSIHT and PSHTP of [7].
(ii) We show that DiFIGHT [4] and related methods arise as special cases of the proposed algorithms.
(iii) We provide a vector-inequality convergence analysis establishing RIC-based sufficient conditions; the per-node condition is identical to the centralized case, supplemented by a mild network stability requirement.

II. Problem Setup and Proposed Algorithms

A. Notations

For a positive integer L, we use $[L]$ to denote the set $\{1, \ldots, L\}$. For a set $S \subseteq [N]$, $\overline{S} := [N] \backslash S$ and $|S|$ denote the complement and the cardinality of S respectively. The support of a vector $\mathbf{u}$, denoted by $\mathrm{supp}(\mathbf{u})$, is the index set of the nonzero entries of $\mathbf{u}$. For a set $\Omega \subseteq [N]$, $(\mathbf{u})_\Omega \in \mathbb{C}^N$ is a vector having entries $[(\mathbf{u})_\Omega]_i = u_i$ if $i \in \Omega$, and other entries set to zeros, and $\mathbf{u}_\Omega \in \mathbb{C}^{|\Omega|}$ is restriction of $\mathbf{u}$ to set Ω. For a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, we denote its spectral norm ($= +\sqrt{\mathbf{A}^H\mathbf{A}}$) by $\|\mathbf{A}\|_{2\to 2}$ and its induced ℓ_∞-norm ($= max\{\sum_{j=1}^n |A_{ij}| \ i = 1, 2, \cdots, m\}$, i.e., largest row sum of the magnitude of the entries of $\mathbf{A}$) by $\|\mathbf{A}\|_\infty$. Throughout the paper we assume that the target signal $\mathbf{x}$ is indeed K sparse, and an upperbound on K is known. Lastly, by $PS(\mathbf{y}, \mathbf{A}, T)$, we denote the vector $\mathbf{u} = \underset{\mathbf{z}\in\mathbb{C}^N}{\arg\min} \|\mathbf{y} - \mathbf{A}\mathbf{z}\|_2^2$ subject to $\mathrm{supp}(\mathbf{z}) = T$, i.e., $\mathbf{A}\mathbf{u}$ is the orthogonal projection of $\mathbf{y}$ on the column space of $\mathbf{A}_T$.

B. Network and Measurement Models

We consider a connected network of L nodes. The combination matrix $\mathbf{C} \in \mathbb{R}^{L\times L}$ specifies the connectivity between the different nodes in the network. The $(v, u)^{\text{th}}$ entry of $\mathbf{C}$, denoted by $c_{vu} \in [0, 1]$, is the weight assigned by node v to the estimate of node u. The graph $\mathcal{G}$ formed by the nodes and the matrix $\mathbf{C}$ are assumed to be undirected. Furthermore, $\mathbf{C}$ is assumed to be left-stochastic, i.e., $\mathbf{C}^t\mathbf{1} = \mathbf{1}$, meaning

TABLE I
DIPSIHT AND DIPSHTP ALGORITHMS

Initialization: $\mathbf{x}_v^0 = \mathbf{0},\ \forall\, v = 1, \ldots, L$

DiPSIHT — Iteration $n+1$ at node v:

S1 (Local PS update): for each $u \in \mathcal{N}_v$,

$$\boldsymbol{\psi}_u^{n+1} = \mathbf{x}_u^n + \mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)(\mathbf{y}_u - \mathbf{A}_u\mathbf{x}_u^n)$$

S2 (Diffuse-then-threshold):

$$\mathbf{x}_v^{n+1} = \mathcal{H}_K\left(\sum_{u\in\mathcal{N}_v} c_{vu}\,\boldsymbol{\psi}_u^{n+1}\right)$$

DiPSHTP — Iteration $n+1$ at node v:

S1 (Local PS hard threshold): for each $u \in \mathcal{N}_v$,

$$\boldsymbol{\psi}_u^{n+1} = \mathcal{H}_K\Big[\mathbf{x}_u^n + \mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)(\mathbf{y}_u - \mathbf{A}_u\mathbf{x}_u^n)\Big]$$

S2 (Diffuse and threshold):

$$\boldsymbol{\phi}_v^{n+1} = \mathcal{H}_K\left(\sum_{u\in\mathcal{N}_v} c_{vu}\,\boldsymbol{\psi}_u^{n+1}\right)$$

S3 (Local pursuit):

$$\mathbf{x}_v^{n+1} = \mathrm{PS}\big(\mathbf{y}_v, \mathbf{A}_v, \mathrm{supp}(\boldsymbol{\phi}_v^{n+1})\big)$$

$\sum_{v=1}^L c_{vu} = 1$ for all v. For any node v, the *neighborhood* of v is denoted by $\mathcal{N}_v$, defined as $\mathcal{N}_v = \{u \in \mathcal{V} : c_{vu} > 0\}$. We assume that for each node v, $c_{vv} > 0$, so that each node is a neighbor of itself.

Following [7], we consider a generalized cost function minimization at each v-th node, given by, $\min_{\mathbf{u}\in\mathbb{C}^N} \|\mathbf{Z}_v(\mathbf{y}_v - \mathbf{A}_v\mathbf{u}_v)\|_2^2$, subject to $\|\mathbf{u}_v\|_0 \le K$, where $\mathbf{Z}_v^H\mathbf{Z}_v = q_v(\mathbf{A}_v\mathbf{A}_v^H)$, with $q_v(t)$ being a power series in t for the v-th node with coefficients $\alpha_i^{(v)} \in \mathbb{R}$, given by $q_v(t) = \sum_{i=c_v}^{\infty} \alpha_i^{(v)} t^i$ where c_v is an integer. Let $\lambda_1^{(v)}, \ldots, \lambda_{m_v}^{(v)} \ge 0$ be the eigenvalues of the positive semidefinite matrix $\mathbf{A}_v\mathbf{A}_v^H$. Throughout the paper, we assume that for each v-th node, $q_v(t)$ is a convergent power series for $t = \lambda_i^{(v)}, i = 1, \cdots, m_v$ and $q_v(\lambda_i^{(v)}) > 0$. Since an eigenvalue $\lambda_i^{(v)}$ of $\mathbf{A}_v\mathbf{A}_v^H$ gives rise to an eigenvalue $q_v(\lambda_i^{(v)})$ for the matrix $q_v(\mathbf{A}_v\mathbf{A}_v^H)$, we have $q_v(\mathbf{A}_v\mathbf{A}_v^H)$ positive definite. We further define

$$\beta_{\min}^{(v)} := \min_i\, q_v\Big(\lambda_i^{(v)}\Big), \qquad \beta_{\max}^{(v)} := \max_i\, q_v\Big(\lambda_i^{(v)}\Big). \quad (2)$$

Since $q_v(\mathbf{A}_v\mathbf{A}_v^H)$ is positive definite as seen earlier, we have $\beta_{\min}^{(v)} > 0$ and $\beta_{\max}^{(v)} > 0$.

C. Proposed DiPSIHT and DiPSHTP Algorithms

Following the adapt-then-combine (ATC) diffusion paradigm [10], we propose the two algorithms in Table I. Here, $\mathbf{x}_v^n$ is the estimate of $\mathbf{x}$ at the v-th node at the n-th step of iteration and $\mathcal{H}_K[\cdot]$ is a hard thresholding operator retaining the K largest-magnitude entries.

Special Cases

(i) $q_v(t) = \mu_v$ (constant, $\mu_v > 0$): DiPSIHT reduces to the diffusion-based IHT, a special case of DiFIGHT [4].

(ii) $q_v(t) = \mu_v(\epsilon_v + t)^{-1}$ ($\mu_v > 0$, $\epsilon_v > 0$): yields a distributed Newton-step IHT/HTP, extending NSIHT/NSHTP [9].

(iii) $q_v(t) = \mu_v t^{-1}$ ($\mu_v > 0$): yields a distributed pseudo-inverse-based HT, extending PHT [8].

(iv) $q_v(t) = \mu_v(1 - t/\varepsilon_v)$ ($\mu_v > 0$, $\varepsilon_v > \sigma_{\max}^2(\mathbf{A}_v)$): yields a new distributed polynomial hard thresholding algorithm (DiPoly). The condition $\varepsilon_v > \sigma_{\max}^2(\mathbf{A}_v)$ ensures $q_v(\lambda_i^{(v)}) > 0$ for all eigenvalues.

(v) $q_v(t) = \mu_v e^{-\lambda_v t}$ ($\mu_v > 0$, $\lambda_v > 0$): yields a new distributed exponential hard thresholding algorithm (Di-Expo). The exponential series converges for all $t \ge 0$, and $q_v(\lambda_i^{(v)}) > 0$ holds automatically.

III. PRELIMINARIES

We collect here the definitions and lemmas used in the convergence analysis. Throughout, $\delta_s^{(v)}$ denotes the s-th order RIC of $\mathbf{A}_v$ (Definition 1), and $\eta = (\sqrt{5}+1)/2$.

Definition 1 (Restricted Isometry Constant [11])**.** *The s-th order restricted isometry constant (RIC) δ_s of a matrix $\mathbf{A} \in \mathbb{C}^{m\times N}$ is the smallest $\delta \ge 0$ such that*

$$(1-\delta)\|\mathbf{x}\|_2^2 \le \|\mathbf{A}\mathbf{x}\|_2^2 \le (1+\delta)\|\mathbf{x}\|_2^2 \quad (3)$$

holds for all s-sparse $\mathbf{x} \in \mathbb{C}^N$. Equivalently, the eigenvalues of $\mathbf{A}_S^H\mathbf{A}_S$ lie in $[1-\delta_s, 1+\delta_s]$ for every index set $S \subseteq [N]$ with $|S| \le s$.

Lemma 1 ([12], [13])**.** *Let $\mathbf{z} \in \mathbb{C}^N$ and let $\mathbf{x} \in \mathbb{C}^N$ be any K-sparse vector. Define $S := \mathrm{supp}(\mathbf{x})$ and $T := \mathrm{supp}(\mathcal{H}_K(\mathbf{z}))$. Then*

$$\|\mathbf{x} - \mathcal{H}_K(\mathbf{z})\|_2 \le \eta\,\big\|(\mathbf{x}-\mathbf{z})_{S\cup T}\big\|_2 \le \eta\,\|\mathbf{x}-\mathbf{z}\|_2, \quad (4)$$

where $\eta = (\sqrt{5}+1)/2 \approx 1.618$.

Lemma 2 ([7, Lemma 3.3])**.** *Let $\mathbf{A}_v \in \mathbb{C}^{m_v\times N}$, $\mathbf{e}_v \in \mathbb{C}^{m_v}$, and let $S \subseteq [N]$ with $|S| \le s$. Then*

$$\big\|\big(\mathbf{A}_v^H q_v(\mathbf{A}_v\mathbf{A}_v^H)\mathbf{e}_v\big)_S\big\|_2 \le \beta_{\max}^{(v)}\sqrt{1+\delta_s^{(v)}}\,\|\mathbf{e}_v\|_2. \quad (5)$$

Lemma 3 ([7, Lemma 3.4])**.** *Let $\mathbf{B}_v := \mathbf{A}_v^H q_v(\mathbf{A}_v\mathbf{A}_v^H)\mathbf{A}_v$, $\mathbf{u} \in \mathbb{C}^N$, and $T \subseteq [N]$ with $|T \cup \mathrm{supp}(\mathbf{u})| \le s$. Then*

$$\big\|\big[(\mathbf{I}-\mathbf{B}_v)\mathbf{u}\big]_T\big\|_2 \le \rho_s^{(v)}\,\|\mathbf{u}\|_2, \quad (6)$$

where, for any positive integer k,

$$\rho_k^{(v)} := \max\Big\{1-(1-\delta_k^{(v)})\beta_{\min}^{(v)},\ (1+\delta_k^{(v)})\beta_{\max}^{(v)} - 1\Big\}. \quad (7)$$

Lemma 4 ([7, Lemma 3.5])**.** *Let $\mathbf{y}_v = \mathbf{A}_v\mathbf{x} + \mathbf{e}_v$ be noisy measurements of a K-sparse signal $\mathbf{x}$, and let $T \subseteq [N]$ with $|T| = t$ and $\delta_{K+t}^{(v)} < 1$. Then $\mathbf{z}^* := \mathrm{PS}(\mathbf{y}_v, \mathbf{A}_v, T)$ satisfies*

$$\|\mathbf{x}-\mathbf{z}^*\|_2 \le \frac{\|(\mathbf{x})_{T^c}\|_2}{\sqrt{1-(\delta_{K+t}^{(v)})^2}} + \frac{\|(\mathbf{A}_v^H\mathbf{e}_v)_T\|_2}{1-\delta_{K+t}^{(v)}}. \quad (8)$$

Lemma 5 ([4, Lemma 4.3]). *Let $\mathbf{B} \in \mathbb{R}^{L\times L}$ be a non-negative matrix and $\mathbf{b} \in \mathbb{R}^L$ be a non-negative vector. Let $\{\mathbf{u}^k\}_{k\geq 0}$ be a sequence of non-negative vectors in $\mathbb{R}^L$ satisfying*

$$\mathbf{u}^{k+1} \preceq \mathbf{B}\,\mathbf{u}^k + \mathbf{b}, \quad k \geq 0, \tag{9}$$

where $\preceq$ denotes componentwise inequality. If $\mathbf{B}$ is stable (i.e., $\|\mathbf{B}\|_{2\to 2} < 1$) and $\mathbf{u}$ is any limit point of the sequence $\{\mathbf{u}^k\}_{k\geq 0}$, then

$$\mathbf{u} \preceq (\mathbf{I}-\mathbf{B})^{-1}\mathbf{b}. \tag{10}$$

IV. Convergence Analysis

A. Notation

For each node v, define the per-node error and the error vector:

$$h_v^n := \|\mathbf{x}_v^n - \mathbf{x}\|_2, \qquad \mathbf{h}^n := [h_1^n, \ldots, h_L^n]^T. \tag{11}$$

For DiPSIHT.

$$\omega_v := \rho_{3K}^{(v)}, \quad \mathbf{\Omega} := \operatorname{diag}(\omega_1, \ldots, \omega_L), \tag{12}$$

$$g_v := \eta\,\beta_{\max}^{(v)}\sqrt{1+\delta_{2K}^{(v)}}\,\|\mathbf{e}_v\|_2, \quad \mathbf{g} := [g_1, \ldots, g_L]^T. \tag{13}$$

For DiPSHTP.

$$d_v := \frac{\eta^2}{\sqrt{1-(\delta_{2K}^{(v)})^2}}, \quad \widetilde{\mathbf{D}} := \operatorname{diag}(d_1, \ldots, d_L), \tag{14}$$

$$\mathbf{\Omega}_0 := \operatorname{diag}\big(\rho_{3K}^{(1)}, \ldots, \rho_{3K}^{(L)}\big), \tag{15}$$

$$[\mathbf{g}_0]_u := \beta_{\max}^{(u)}\sqrt{1+\delta_{2K}^{(u)}}\,\|\mathbf{e}_u\|_2, \tag{16}$$

$$[\tilde{\mathbf{g}}_2]_v := \frac{\sqrt{1+\delta_K^{(v)}}}{1-\delta_{2K}^{(v)}}\,\|\mathbf{e}_v\|_2. \tag{17}$$

The (v,u) entry of the DiPSHTP iteration matrix $\mathbf{M} := \widetilde{\mathbf{D}}\mathbf{C}^T\mathbf{\Omega}_0$ is $[\mathbf{M}]_{vu} = d_v c_{vu}\rho_{3K}^{(u)}$, where the row-index factor d_v arises from the local pursuit at node v and the column-index factor $\rho_{3K}^{(u)}$ from the contraction at neighbor u.

B. DiPSIHT Convergence

Theorem 1 (DiPSIHT Convergence). *If for each $v = 1, \ldots, L$, the sensing matrix $\mathbf{A}_v$ satisfies $\delta_{3K}^{(v)} < 1/\eta$ and $q_v(t)$ is chosen so that*

$$\frac{1-\frac{1}{\eta}}{1-\delta_{3K}^{(v)}} < \beta_{\min}^{(v)} \leq \beta_{\max}^{(v)} < \frac{1+\frac{1}{\eta}}{1+\delta_{3K}^{(v)}}, \tag{18}$$

then the DiPSIHT iterates $\{\mathbf{x}_v^n\}_{n\geq 1}$ satisfy

$$\mathbf{h}^{n+1} \preceq \eta\,\mathbf{C}^T\mathbf{\Omega}\,\mathbf{h}^n + \mathbf{C}^T\mathbf{g}. \tag{19}$$

Condition (18) *implies $\omega_v = \rho_{3K}^{(v)} < 1/\eta$ for each v [7]. If additionally*

$$\max_v \sum_{u=1}^{L} c_{vu}\,\omega_u < \frac{1}{\eta}, \tag{20}$$

then $\|\eta\mathbf{C}^T\mathbf{\Omega}\|_\infty < 1$, implying $\|\eta\mathbf{C}^T\mathbf{\Omega}\|_{2\to 2} < 1$ (i.e., $\eta\mathbf{C}^T\mathbf{\Omega}$ is stable) in which case any limit point $\mathbf{h}$ of the sequence $\{\mathbf{h}^n\}_{n\geq 0}$ satisfies

$$\mathbf{h} \preceq \big(\mathbf{I} - \eta\mathbf{C}^T\mathbf{\Omega}\big)^{-1}\mathbf{C}^T\mathbf{g}. \tag{21}$$

Proof. Fix any node v. Let $S := \operatorname{supp}(\mathbf{x})$ ($|S| \leq K$), $S_v^{n+1} := \operatorname{supp}(\mathbf{x}_v^{n+1})$ ($|S_v^{n+1}| \leq K$), and $T := S \cup S_v^{n+1}$ ($|T| \leq 2K$).
Step 1. Since $\mathbf{x}_v^{n+1} = \mathcal{H}_K\big(\sum_u c_{vu}\boldsymbol{\psi}_u^{n+1}\big)$, Lemma 1 with K-sparse reference $\mathbf{x}$ gives

$$h_v^{n+1} \leq \eta\,\Big\|\Big(\sum_{u=1}^{L} c_{vu}\boldsymbol{\psi}_u^{n+1} - \mathbf{x}\Big)_T\Big\|_2. \tag{22}$$

Step 2. Since $\sum_u c_{vu} = 1$, we have $\sum_u c_{vu}\boldsymbol{\psi}_u^{n+1} - \mathbf{x} = \sum_u c_{vu}(\boldsymbol{\psi}_u^{n+1} - \mathbf{x})$. Substituting the Step S1 update and using $\mathbf{y}_u = \mathbf{A}_u\mathbf{x} + \mathbf{e}_u$:

$$\boldsymbol{\psi}_u^{n+1} - \mathbf{x} = (\mathbf{I} - \mathbf{B}_u)(\mathbf{x}_u^n - \mathbf{x}) + \mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)\mathbf{e}_u, \tag{23}$$

where $\mathbf{B}_u := \mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)\mathbf{A}_u$.
Step 3. Since $|\operatorname{supp}(\mathbf{x}_u^n)| \leq K$ and $|T| \leq 2K$, we have $|T \cup \operatorname{supp}(\mathbf{x}_u^n - \mathbf{x})| \leq 3K$. Applying Lemma 3 ($s = 3K$) to the first term of (23):

$$\big\|\big[(\mathbf{I} - \mathbf{B}_u)(\mathbf{x}_u^n - \mathbf{x})\big]_T\big\|_2 \leq \rho_{3K}^{(u)}\,h_u^n. \tag{24}$$

Applying Lemma 2 ($s = 2K$, since $|T| \leq 2K$) to the second term:

$$\big\|\big[\mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)\mathbf{e}_u\big]_T\big\|_2 \leq \beta_{\max}^{(u)}\sqrt{1+\delta_{2K}^{(u)}}\,\|\mathbf{e}_u\|_2. \tag{25}$$

Step 4. Applying the triangle inequality to (22), using (23) and bounds (24)–(25):

$$h_v^{n+1} \leq \eta\sum_{u=1}^{L} c_{vu}\Big[\rho_{3K}^{(u)}h_u^n + \beta_{\max}^{(u)}\sqrt{1+\delta_{2K}^{(u)}}\,\|\mathbf{e}_u\|_2\Big]. \tag{26}$$

Recalling $\omega_u = \rho_{3K}^{(u)}$ and $g_u = \eta\beta_{\max}^{(u)}\sqrt{1+\delta_{2K}^{(u)}}\|\mathbf{e}_u\|_2$ from (12)–(13), equation (26) reads

$$h_v^{n+1} \leq \eta\sum_u c_{vu}\omega_u h_u^n + \sum_u c_{vu}g_u = \big[\eta\mathbf{C}^T\mathbf{\Omega}\,\mathbf{h}^n + \mathbf{C}^T\mathbf{g}\big]_v,$$

which is (19) componentwise.
Step 5. Condition (20) states that the maximum row sum of $\eta\mathbf{C}^T\mathbf{\Omega}$ is strictly less than one, i.e., $\|\eta\mathbf{C}^T\mathbf{\Omega}\|_\infty < 1$. Since $\|\eta\mathbf{C}^T\mathbf{\Omega}\|_{2\to 2} \leq \|\eta\mathbf{C}^T\mathbf{\Omega}\|_\infty < 1$, $\eta\mathbf{C}^T\mathbf{\Omega}$ is stable. Also, the sequence $\{\mathbf{h}^n\}_{n\geq 0}$ is non-negative by construction. Applying Lemma 5 with $\mathbf{B} = \eta\mathbf{C}^T\mathbf{\Omega}$ and $\mathbf{b} = \mathbf{C}^T\mathbf{g}$, any limit point $\mathbf{h}$ of $\{\mathbf{h}^n\}$ satisfies (21). □

In the noiseless case, $\mathbf{g} = \mathbf{0}$, so $\mathbf{h} \preceq \mathbf{0}$, meaning every limit point is zero and $\mathbf{h}^n \to \mathbf{0}$ as $n \to \infty$. Since $\mathbf{C}$ is left-stochastic, $\sum_u c_{vu}\omega_u \leq \max_u \omega_u$, so condition (20) is satisfied whenever $\max_v \omega_v < 1/\eta$ holds independently of the network topology.

C. DiPSHTP Convergence

Theorem 2 (DiPSHTP Convergence). *Suppose the per-node conditions of Theorem 1 hold for each v, and additionally $\delta_{2K}^{(v)} < 1$ for each v. Then the DiPSHTP iterates $\{\mathbf{x}_v^n\}_{n\geq 1}$ satisfy*

$$\mathbf{h}^{n+1} \preceq \widetilde{\mathbf{D}}\,\mathbf{C}^T\mathbf{\Omega}_0\,\mathbf{h}^n + \widetilde{\mathbf{D}}\,\mathbf{C}^T\mathbf{g}_0 + \tilde{\mathbf{g}}_2, \tag{27}$$

where $\widetilde{\mathbf{D}}$, $\mathbf{\Omega}_0$, $\mathbf{g}_0$, $\tilde{\mathbf{g}}_2$ *are as in* (14)–(17) *and* $\mathbf{M} := \widetilde{\mathbf{D}}\mathbf{C}^T\mathbf{\Omega}_0$. *If additionally*

$$\max_v d_v \sum_{u=1}^{L} c_{vu}\, \rho_{3K}^{(u)} \;<\; 1, \tag{28}$$

then $\|\mathbf{M}\|_\infty < 1$, *meaning,* $\|\mathbf{M}\|_{2\to 2} < 1$ *and* $\mathbf{M}$ *is stable. By Lemma 5, any limit point* $\mathbf{h}$ *of the sequence* $\{\mathbf{h}^n\}_{n\geq 0}$ *satisfies*

$$\mathbf{h} \preceq (\mathbf{I} - \mathbf{M})^{-1}\big(\widetilde{\mathbf{D}}\,\mathbf{C}^T\mathbf{g}_0 + \tilde{\mathbf{g}}_2\big). \tag{29}$$

In the noiseless case, $\mathbf{g}_0 = \tilde{\mathbf{g}}_2 = \mathbf{0}$, *so every limit point satisfies* $\mathbf{h} \preceq \mathbf{0}$, *hence* $\mathbf{h}^n \to \mathbf{0}$ *as* $n \to \infty$. *Since* $\mathbf{C}$ *is left-stochastic,* $\sum_u c_{vu}\rho_{3K}^{(u)} \leq \max_u \rho_{3K}^{(u)}$, *so condition* (28) *holds whenever* $\max_v d_v \cdot \max_u \rho_{3K}^{(u)} < 1$, *a node-level condition independent of* $\mathbf{C}$.

Proof. Fix any node v. Let $\Lambda := \mathrm{supp}(\mathbf{x})$, $|\Lambda| \leq K$.

Stage A (Step S1: per-neighbor local PS hard threshold). Fix a neighbor $u \in \mathcal{N}_v$. Define the pre-threshold vector

$$\mathbf{z}_u^{n+1} := \mathbf{x}_u^n + \mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)(\mathbf{y}_u - \mathbf{A}_u\mathbf{x}_u^n), \tag{30}$$

so $\boldsymbol{\psi}_u^{n+1} = \mathcal{H}_K[\mathbf{z}_u^{n+1}]$ with $|\mathrm{supp}(\boldsymbol{\psi}_u^{n+1})| \leq K$. Let $\hat{S}_u := \mathrm{supp}(\boldsymbol{\psi}_u^{n+1})$. Using $\mathbf{y}_u = \mathbf{A}_u\mathbf{x} + \mathbf{e}_u$:

$$\mathbf{z}_u^{n+1} - \mathbf{x} = (\mathbf{I} - \mathbf{B}_u)(\mathbf{x}_u^n - \mathbf{x}) + \mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)\mathbf{e}_u. \tag{31}$$

Applying Lemma 1 to $\boldsymbol{\psi}_u^{n+1} = \mathcal{H}_K[\mathbf{z}_u^{n+1}]$ with K-sparse reference $\mathbf{x}$, $S_{\mathbf{x}} := \Lambda$, $T_{\boldsymbol{\psi}} := \hat{S}_u$:

$$\|\boldsymbol{\psi}_u^{n+1} - \mathbf{x}\|_2 \leq \eta\, \big\|(\mathbf{z}_u^{n+1} - \mathbf{x})_{\Lambda\cup\hat{S}_u}\big\|_2. \tag{32}$$

The cardinality bounds $|\Lambda| \leq K$, $|\hat{S}_u| \leq K$, and $|\mathrm{supp}(\mathbf{x}_u^n)| \leq K$ give $|(\Lambda\cup\hat{S}_u)\cup\mathrm{supp}(\mathbf{x}_u^n - \mathbf{x})| \leq 3K$. Applying the triangle inequality to (31), then Lemma 3 ($s = 3K$) and Lemma 2 ($s = 2K$, since $|\Lambda \cup \hat{S}_u| \leq 2K$):

$$\big\|\big[(\mathbf{I} - \mathbf{B}_u)(\mathbf{x}_u^n - \mathbf{x})\big]_{\Lambda\cup\hat{S}_u}\big\|_2 \leq \rho_{3K}^{(u)}\, h_u^n, \tag{33}$$

$$\big\|\big[\mathbf{A}_u^H q_u(\mathbf{A}_u\mathbf{A}_u^H)\mathbf{e}_u\big]_{\Lambda\cup\hat{S}_u}\big\|_2 \leq \beta_{\max}^{(u)}\sqrt{1+\delta_{2K}^{(u)}}\,\|\mathbf{e}_u\|_2. \tag{34}$$

Substituting (33)–(34) into (32):

$$\|\boldsymbol{\psi}_u^{n+1} - \mathbf{x}\|_2 \leq \eta\Big[\rho_{3K}^{(u)}\, h_u^n + \beta_{\max}^{(u)}\sqrt{1+\delta_{2K}^{(u)}}\,\|\mathbf{e}_u\|_2\Big] =: \eta\,\mu_u^n, \tag{35}$$

where $\mu_u^n := \rho_{3K}^{(u)} h_u^n + [\mathbf{g}_0]_u$.

Stage B (Step S2: diffuse and threshold). Let $\mathbf{w}_v := \sum_u c_{vu}\boldsymbol{\psi}_u^{n+1}$, $\boldsymbol{\phi}_v^{n+1} := \mathcal{H}_K(\mathbf{w}_v)$, $T_v := \mathrm{supp}(\boldsymbol{\phi}_v^{n+1})$. Define $S_v := \Lambda \cup T_v$; since $|\Lambda| \leq K$ and $|T_v| \leq K$, we have $|S_v| \leq 2K$. Since $\boldsymbol{\phi}_v^{n+1}$ vanishes on T_v^c (and hence on $S_v^c \subseteq T_v^c$):

$$\|(\mathbf{x})_{T_v^c}\|_2 = \|(\mathbf{x} - \boldsymbol{\phi}_v^{n+1})_{T_v^c}\|_2 \leq \|\mathbf{x} - \boldsymbol{\phi}_v^{n+1}\|_2. \tag{36}$$

Applying Lemma 1 to $\boldsymbol{\phi}_v^{n+1} = \mathcal{H}_K(\mathbf{w}_v)$ with K-sparse reference $\mathbf{x}$, support set Λ, and thresholded support T_v:

$$\|\boldsymbol{\phi}_v^{n+1} - \mathbf{x}\|_2 \leq \eta\,\big\|(\mathbf{w}_v - \mathbf{x})_{S_v}\big\|_2 \leq \eta\,\|\mathbf{w}_v - \mathbf{x}\|_2, \tag{37}$$

where the second inequality drops the restriction to S_v. Since $\sum_u c_{vu} = 1$, the triangle inequality gives

$$\|\mathbf{w}_v - \mathbf{x}\|_2 \leq \sum_{u=1}^{L} c_{vu}\|\boldsymbol{\psi}_u^{n+1} - \mathbf{x}\|_2. \tag{38}$$

Combining (36)–(38) with (35):

$$\|(\mathbf{x})_{T_v^c}\|_2 \leq \eta^2 \sum_{u=1}^{L} c_{vu}\Big[\rho_{3K}^{(u)} h_u^n + [\mathbf{g}_0]_u\Big]. \tag{39}$$

The factor η^2 arises from two sequential applications of Lemma 1: once in Stage A (the per-neighbor $\mathcal{H}_K$ in Step S1) and once in Stage B (the $\mathcal{H}_K$ at node v in Step S2).

Stage C (Step S3: local pursuit). Since $|T_v| \leq K$, we have $K + |T_v| \leq 2K$, so $\delta_{K+|T_v|}^{(v)} \leq \delta_{2K}^{(v)} < 1$. Applying Lemma 4 with $T = T_v$:

$$h_v^{n+1} \leq \frac{\|(\mathbf{x})_{T_v^c}\|_2}{\sqrt{1-(\delta_{2K}^{(v)})^2}} + \frac{\|(\mathbf{A}_v^H\mathbf{e}_v)_{T_v}\|_2}{1-\delta_{2K}^{(v)}}. \tag{40}$$

For the noise term, since $\mathbf{A}_{v,T_v} := (\mathbf{A}_v)_{T_v}$ is the column submatrix and $|T_v| \leq K$, the RIP (Definition 1) gives $\rho(\mathbf{A}_{v,T_v}) \leq \sqrt{1+\delta_K^{(v)}}$, hence

$$\|(\mathbf{A}_v^H\mathbf{e}_v)_{T_v}\|_2 = \|\mathbf{A}_{v,T_v}^H\mathbf{e}_v\|_2 \leq \sqrt{1+\delta_K^{(v)}}\,\|\mathbf{e}_v\|_2. \tag{41}$$

Substituting (39)–(41) into (40) and recalling d_v from (14):

$$h_v^{n+1} \leq d_v \sum_{u=1}^{L} c_{vu}\,\rho_{3K}^{(u)}\, h_u^n + d_v \sum_{u=1}^{L} c_{vu}\,[\mathbf{g}_0]_u + [\tilde{\mathbf{g}}_2]_v. \tag{42}$$

Writing (42) for all $v = 1, \ldots, L$ and noting $[\widetilde{\mathbf{D}}\mathbf{C}^T\mathbf{\Omega}_0]_{vu} = d_v c_{vu}\rho_{3K}^{(u)}$ and $[\widetilde{\mathbf{D}}\mathbf{C}^T\mathbf{g}_0]_v = d_v\sum_u c_{vu}[\mathbf{g}_0]_u$, yields (27) in vector form.

The matrix $\mathbf{M} = \widetilde{\mathbf{D}}\mathbf{C}^T\mathbf{\Omega}_0$ is non-negative with $\|\mathbf{M}\|_\infty = \max_v d_v\sum_u c_{vu}\rho_{3K}^{(u)}$. Condition (28) implies, $\|\mathbf{M}\|_\infty < 1$, meaning, $\|\mathbf{M}\|_{2\to 2} \leq \|\mathbf{M}\|_\infty < 1$ and $\mathbf{M}$ is stable. Also, the sequence $\{\mathbf{h}^n\}_{n\geq 0}$ is non-negative by construction. Applying Lemma 5 with $\mathbf{B} = \mathbf{M}$ and $\mathbf{b} = \widetilde{\mathbf{D}}\mathbf{C}^T\mathbf{g}_0 + \tilde{\mathbf{g}}_2$, it then follows that any limit point $\mathbf{h}$ of $\{\mathbf{h}^n\}$ satisfies (29). □

V. Simulation Results

We consider a noiseless setting, where the target K-sparse signal $\mathbf{x}$ is of dimension $N = 1000$, with support chosen uniformly at random and the non-zero entries drawn i.i.d. from either a normal or a Bernoulli distribution (± 1, equi-probable). We consider a network of $L = 10$ nodes connected via an Erdős–Rényi graph with an edge connection probability of $\log(L)/L$; self-loops are included to ensure diffusion stability, and the adjacency matrix is normalized to be left-stochastic. The local sensing matrices $\mathbf{A}_v$ are populated with i.i.d. Gaussian entries having zero mean and variance $1/m$.

For the proposed DiPSIHT and DiPSHTP algorithms, we evaluate two choices for the local power series $q_v(t)$: the

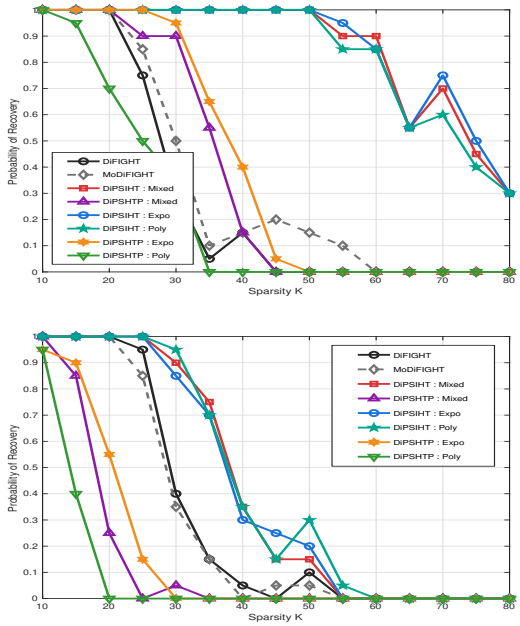

Fig. 1. Recovery rate vs. sparsity K, fixed $m = 50$. Top: Gaussian. Bottom: Bernoulli.

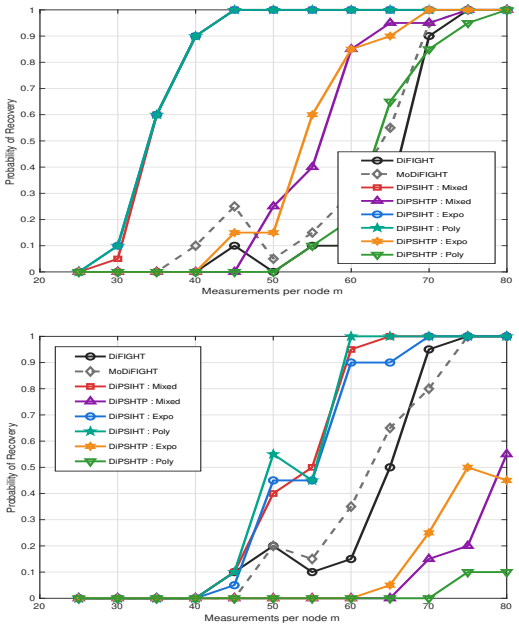

Fig. 2. Recovery rate vs. measurements m, fixed $K = 40$. Top: Gaussian. Bottom: Bernoulli.

exponential function $q_v(t) = \mu_v \exp(-\lambda_v t)$ and the first-order polynomial $q_v(t) = \mu_v(1 - t/\epsilon_v)$. The power series parameters $(\mu_v, \lambda_v, \epsilon_v)$ for DiPSIHT and DiPSHTP are selected for each specific algorithm and underlying signal distribution to optimize recovery performance. Specifically, for the Bernoulli distribution, the chosen parameters are: DiPSIHT (Expo) $\mu = 0.24$, $\lambda = 0.0125$; DiPSIHT (Poly) $\mu = 0.3$, $\epsilon = 60$; DiPSHTP (Expo) $\mu = 0.1375$, $\lambda = 0.0075$; and DiPSHTP (Poly) $\mu = 0.0875$, $\epsilon = 38$. For the Gaussian distribution, the parameters are: DiPSIHT (Expo) $\mu = 0.4$, $\lambda = 0.025$; DiPSIHT (Poly) $\mu = 0.3$, $\epsilon = 80$; DiPSHTP (Expo) $\mu = 0.2$, $\lambda = 0.0075$; and DiPSHTP (Poly) $\mu = 0.0875$, $\epsilon = 50$.

To demonstrate performance, we test three network configurations: homogeneous exponential (DiPSIHT/DiPSHTP-Expo), in which all nodes use the exponential $q_v(t)$; homogeneous polynomial (DiPSIHT/DiPSHTP-Poly), in which all nodes use the polynomial $q_v(t)$; and heterogeneous mixed (DiPSIHT/DiPSHTP-Mixed), in which each node independently uses either the exponential or polynomial $q_v(t)$, chosen uniformly at random. These are compared against DiFIGHT and MoDiFIGHT [4] ($\mu = 0.5$). Recovery is successful if $\mathrm{NMSD}(n) := \|\mathbf{X}^n - \mathbf{x}^\star\|_F^2/(L\|\mathbf{x}^\star\|_2^2)$ falls below 10^{-4} within 100 iterations, with results averaged over 20 independent trials. Figs. 1 and 2 show the probability of successful recovery versus sparsity K (fixed $m = 50$) and versus measurements m (fixed $K = 40$), for both signal distributions. The proposed algorithms consistently and substantially outperform the baselines across all configurations.

REFERENCES

[1] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[2] S. Foucart, "Hard thresholding pursuit: An algorithm for compressive sensing," *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, 2011.

[3] S. Patterson, Y. C. Eldar, and I. Keidar, "Distributed compressed sensing for static and time-varying networks," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4931–4946, Oct. 2014.

[4] S. Mukhopadhyay and M. Chakraborty, "Deterministic and randomized diffusion based iterative generalized hard thresholding (DiFIGHT) for distributed recovery of sparse signals," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 8, pp. 25–36, 2022.

[5] A. Zaki, P. P. Mitra, L. K. Rasmussen, and S. Chatterjee, "Estimate exchange over network is good for distributed hard thresholding pursuit," *Signal Process.*, vol. 156, pp. 1–11, 2019.

[6] S. Chouvardas, G. Mileounis, N. Kalouptsidis, and S. Theodoridis, "Greedy sparsity-promoting algorithms for distributed learning," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1419–1432, Mar. 2015.

[7] K. A. Bapat and M. Chakraborty, "Power series based hard thresholding algorithms for sparse signal recovery," *IEEE Signal Process. Lett.*, vol. 32, pp. 2474–2478, Jun. 2025.

[8] J. Wen, H. He, Z. He, and F. Zhu, "A pseudo-inverse-based hard thresholding algorithm for sparse signal recovery," *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 7, pp. 7621–7630, Jul. 2023.

[9] N. Meng and Y.-B. Zhao, "Newton-step-based hard thresholding algorithms for sparse signal recovery," *IEEE Trans. Signal Process.*, vol. 68, pp. 6594–6606, 2020.

[10] S.-Y. Tu and A. H. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6217–6234, Dec. 2012.

[11] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. New York, NY, USA: Birkhäuser, 2013.

[12] J. Shen and P. Li, "A tight bound of hard thresholding," *J. Mach. Learn. Res.*, vol. 18, no. 208, pp. 1–42, 2018.

[13] Y. Zhao and Z. Luo, "Improved RIP-based bounds for guaranteed performance of two compressed sensing algorithms," *Sci. China Math.*, vol. 66, no. 5, pp. 1123–1140, 2023.