

r -LOCAL UNLABELED SENSING: IMPROVED ALGORITHM AND APPLICATIONS

Ahmed Ali Abbasi[†] Abiy Tasissa^{*} Shuchin Aeron[†]

[†] Department of ECE, Tufts University, Medford, MA 02155

^{*} Department of Mathematics, Tufts University, Medford, MA 02155

ABSTRACT

The unlabeled sensing problem is to solve a noisy linear system of equations under unknown permutation of the measurements. We study a particular case of the problem where the permutations are restricted to be r -local, i.e. the permutation matrix is block diagonal with $r \times r$ blocks. Assuming a Gaussian measurement matrix, we argue that the r -local permutation model is more challenging compared to a recent sparse permutation model. We propose a proximal alternating minimization algorithm for the general unlabeled sensing problem that provably converges to a first order stationary point. Applied to the r -local model, we show that the resulting algorithm is efficient. We validate the algorithm on synthetic and real datasets. We also formulate the 1-d unassigned distance geometry problem as an unlabeled sensing problem with a structured measurement matrix.

Index Terms— Unlabeled sensing, unassigned distance geometry problem, proximal alternating minimization.

1. INTRODUCTION

The standard least squares problem is to recover signal $\mathbf{X}^* \in \mathbb{R}^{d \times m}$ given measurements $\mathbf{Y} \in \mathbb{R}^{n \times m}$ and the measurement matrix \mathbf{B} . The signal matrix \mathbf{X}^* is estimated by solving the least squares minimization problem $\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{B}\mathbf{X}\|_F^2$. This formulation assumes that there is one to one correspondence between measurements (the rows of \mathbf{Y}) and the measurement vectors (the rows of \mathbf{B}), i.e., we know which measurement corresponds to what measurement vector. However, in many problems of applied interest such as header free communication for mobile wireless networks [1], sampling in the presence of clock jitter [2], linear regression without correspondence [3] and point cloud registration [4], this mapping is not explicitly given. This motivates the study of the unlabeled sensing problem with the prototypical form given by

$$\mathbf{Y} = \mathbf{P}^* \mathbf{B} \mathbf{X}^* + \mathbf{W}, \quad (1)$$

where $\mathbf{P} \in \mathbb{R}^{n \times n}$ denotes the unknown permutation matrix and $\mathbf{W} \in \mathbb{R}^{n \times m}$ denotes an additive Gaussian noise with $\mathbf{W}_{ij} = \mathcal{N}(0, \sigma^2)$. The problem is to estimate \mathbf{X}^* and \mathbf{P}^* given the measurement matrix \mathbf{B} and measurements \mathbf{Y} .

Related work: The unlabeled sensing problem was first considered with a Gaussian measurement matrix \mathbf{B} in the single-view setup, where $m = 1$, in [5]. There, the authors showed that $n = 2d$ noiseless measurements are necessary and sufficient for recovery of \mathbf{x}^* . The same result was subsequently proven in [6] using an algebraic argument. For the case of Gaussian \mathbf{B} , the work in [7] showed that with $\text{SNR} \triangleq \|\mathbf{X}^*\|_F^2 / m\sigma^2$, $\text{SNR} = \Omega(n^2)$ is necessary for recovering \mathbf{P}^* exactly. For the multi-view setup, where $m > 1$, the result in [8] shows that the necessary SNR for recovery decreases as m increases, with necessary $\text{SNR} = O(1)$ for $m = \Omega(\log n)$.

Several algorithms have been proposed for the single view unlabeled sensing problem [1, 7, 9]. Algorithms for the multi-view problem setup have been proposed in works [8, 10, 11, 12, 13]. As computing the maximum likelihood estimate of \mathbf{P}^* may be NP-hard [11], the works in [8, 10, 12] impose a k -sparse structure on \mathbf{P}^* (see Figure 1a) where $n - k$ elements of the permutation are on the diagonal. In [13], the authors introduced the r -local model (see section 2) along with an algorithm that is based on graph matching. A permutation matrix \mathbf{P}^* of size $n \times n$ is r -local if it is composed of n/r blocks. Formally, $\mathbf{P}^* = \text{diag}(\mathbf{P}_1^*, \dots, \mathbf{P}_{n/r}^*)$, where $\mathbf{P}_i^* \in \Pi_r$ denotes an $r \times r$ permutation matrix (see Figure 1b).

Applications: In this paper, we consider two applications of the r -local unlabeled sensing model. The first application is the unassigned distance geometry problem (uDGP) [14] (see section 3). The second application is image unscrambling [15] (see section 5). Another application of the r -local permutation model is linear regression with blocking [16].

Contributions: (a) We compare the r -local model with the k -sparse model and argue that the r -local problem is more challenging under the most widely studied case of Gaussian measurement matrix. (b) We formulate the 1-d unassigned distance geometry problem (uDGP) [14] as an unlabeled sensing problem with a structured measurement matrix. (c) We propose a proximal alternating minimization algorithm for the r -local unlabeled sensing problem. Simulation results show that the proposed algorithm outperforms the algorithms in [8, 11, 13]. We also compare the proposed algorithm with [12] to show that algorithms for the k -sparse model are not applicable to the r -local model.

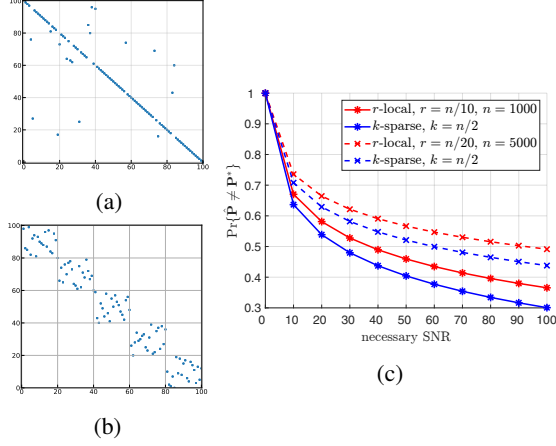


Fig. 1: (a). A sparse (or partially shuffled) permutation considered in [8, 10, 12, 17]. (b). The r -local permutation structure considered in this paper, with $r = 20$. (c). Figure plots the lower bound on the event $\Pr\{\hat{\mathbf{P}} \neq \mathbf{P}^*\}$ against SNR for the two models (see equation (2)).

2. r -LOCAL MODEL VS k -SPARSE MODEL

A permutation matrix \mathbf{P} is k -sparse if $n - k$ of its elements are on the diagonal, i.e., $\langle \mathbf{I}, \mathbf{P} \rangle = n - k$. Figures 1a and 1b show a k -sparse and an r -local permutation matrix, respectively. We compare the two models by the difficulty of recovering \mathbf{P}^* under each model when the matrix \mathbf{B} is assumed to be random Gaussian. Under this assumption, for any estimator $\hat{\mathbf{P}}$, the probability of error $\Pr\{\hat{\mathbf{P}} \neq \mathbf{P}^*\}$ is lower bounded by the result in [8]

$$\Pr\{\hat{\mathbf{P}} \neq \mathbf{P}^*\} \geq 1 - \frac{1 + \frac{nm}{2} \log_2(1 + \text{SNR})}{H(\mathbf{P}^*)}, \quad (2)$$

where \mathbf{P}^* is a random variable and $H(\mathbf{P}^*)$ denotes the entropy of \mathbf{P}^* . For uniformly distributed \mathbf{P}^* , the entropy is $H(\mathbf{P}^*) = \log_2 a$, where a is the number of possible permutations. Without any assumption on the permutation structure, the entropy of $\mathbf{P}^* \in \mathbb{R}^{n \times n}$ is $H(\mathbf{P}^*) = \log_2(n!)$. The entropy for r -local permutations is $H_r(\mathbf{P}^*) = \log_2 r!^{n/r}$. The entropy for k -sparse permutations is $H_k(\mathbf{P}^*) = \log_2(n!/(n-k)!)$. Performance guarantees for proposed algorithms are given for $k = n/8$ in [8] and $k = n/4$ in [12]. All simulation results for permutation recovery in [10, 17] are for $k \leq n/2$. Figure 1c shows that, for the same SNR, the lower bound in (2) is higher for the r -local model, $r = \{n/10, n/20\}$, than for the k -sparse model, with $k = n/2$. This suggests that permutation recovery under the r -local model is challenging in the large n regime.

3. 1-D UNASSIGNED DISTANCE GEOMETRY AS AN UNLABELED SENSING PROBLEM

The one-dimensional unassigned distance geometry problem (1-d uDGP) (see Figure 2) is to recover point coordinates

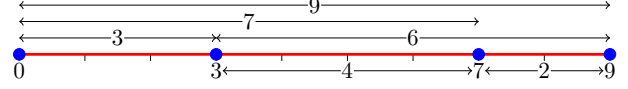


Fig. 2: The 1-d unassigned distance geometry problem (uDGP) is to recover the point coordinates $(0, 3, 7, 9)$ from their unlabeled pairwise distances $\{3, 2, 4, 6, 7, 9\}$. uDGP can be formulated as an unlabeled sensing problem with a deterministic measurement matrix \mathbf{B}_u , see equation (3).

from their unlabeled pairwise distances [14, 18, 19]. The distances are unlabeled as the pair (i, j) corresponding to the distance $|x_i - x_j|$ is not known. Let $\bar{\mathbf{x}} \in \mathbb{R}^d = [\bar{x}_1, \dots, \bar{x}_d]^T$ be the vector of unknown coordinates. Without loss of generality, we assume that the unknown point coordinates are sorted in decreasing order. It follows that each pairwise distance $\bar{x}_i - \bar{x}_j$ is a linear measurement of $\bar{\mathbf{x}}$. As the distances are translation invariant, we can also assume that the left most point is at the origin. The distance vector $\mathbf{y} \in \mathbb{R}^{d(d-1)/2}$ is the matrix-vector product $\mathbf{y} = \mathbf{B}_u \mathbf{x}^*$ where $\mathbf{x}^* \in \mathbb{R}^{d-1}$. It can be verified that \mathbf{B}_u has the following form.

$$\mathbf{B}_u = [\mathbf{I}_{(d-1) \times (d-1)}; \mathbf{B}_2; \dots; \mathbf{B}_{d-1}],$$

$$[\mathbf{B}_i \in \mathbb{R}^{(d-i) \times (d-1)}]_{pq} = \begin{cases} 1 & \text{if } q = i - 1 \\ -1 & \text{if } q = p + i - 1 \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where $[\mathbf{A}_1; \mathbf{A}_2]$ denotes vertical concatenation of the matrices $\mathbf{A}_1, \mathbf{A}_2$. The unlabeled sensing formulation for the example in Figure 2, with $d = 4$ points, is given below

$$\begin{pmatrix} 3 \\ 2 \\ 4 \\ 6 \\ 7 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 9 \\ 7 \\ 3 \end{pmatrix}.$$

$\mathbf{y} \quad \mathbf{B}_u \quad \mathbf{x}^*$

Why r -local model for uDGP? Imposing an r -local structure on \mathbf{P}^* has a practical interpretation for the uDGP problem. The problem where, in addition to the distances $|x_i - x_j|$, one of the two corresponding indices (i, j) is known, is modelled by an r -local $\mathbf{P}^* = \text{blkdiag}(d-1, \dots, 2)$. The r -local structure also models the problem where distance assignments are known up to a cluster of points. For example, for two clusters, each distance is known to be between a pair of points in cluster 1 or cluster 2 (intra-cluster) or a point in cluster 1 and a point in cluster 2 (inter-cluster) but the pair of points corresponding to each distance is still unknown.

4. PROPOSED APPROACH AND ALGORITHM

In this section, we present a new algorithm for the r -local unlabeled sensing problem. To motivate our proposed algorithm, we note that for i.i.d Gaussian noise \mathbf{W} in (1), the maximum likelihood estimate of \mathbf{P}^* is given by

$$\arg \min_{\mathbf{X}, \mathbf{P} \in \Pi_n} F(\mathbf{X}, \mathbf{P}) = \|\mathbf{Y} - \mathbf{P}\mathbf{B}\mathbf{X}\|_F^2, \quad (4)$$

where Π_n denotes the set of permutations of order n . The optimization problem in (4) is non-convex since the set of permutations is discrete. A natural optimization scheme for the problem in (4) is the alternating minimization algorithm which alternates between updating the signal matrix \mathbf{X} and the permutation \mathbf{P} . The works in [10, 11, 12, 17] propose one step estimators for \mathbf{P}^* , \mathbf{X}^* . We consider complete alternating minimization for (4) yielding the following updates

$$\mathbf{P}^{(t)} = \underset{\mathbf{P} \in \Pi_n}{\operatorname{argmin}} \langle -\mathbf{Y}\mathbf{X}^{(t)\top} \mathbf{B}^\top, \mathbf{P} \rangle, \quad (5)$$

$$\mathbf{X}^{(t+1)} = \underset{\mathbf{X}}{\operatorname{argmin}} F(\mathbf{X}, \mathbf{P}^{(t)}) = (\mathbf{P}^{(t)} \mathbf{B})^\dagger \mathbf{Y}. \quad (6)$$

The iterates $(\mathbf{P}^{(t)}, \mathbf{X}^{(t+1)})$ above are monotonically decreasing $F(\mathbf{P}^{(t+1)}, \mathbf{X}^{(t+1)}) \leq F(\mathbf{P}^{(t)}, \mathbf{X}^{(t+1)})$. However, this is not sufficient to establish convergence of the iterates. Specifically, we can not claim that $\lim_{t \rightarrow \infty} \{(\mathbf{P}^{(t)}, \mathbf{X}^{(t)})\}$ exists. We propose to use proximal alternating minimization (PAM) [20] instead which, as we show in section 4.1, converges to a first order stationary point of the objective in (4). PAM adds a regularization term that encourages the new iterate to be close to the current iterate. Formally, for the objective in (4), the PAM updates are given by

$$\mathbf{P}^{(t)} = \underset{\mathbf{P} \in \Pi_n}{\operatorname{argmin}} F(\mathbf{X}^{(t)}, \mathbf{P}) + \lambda^{(t)} \|\mathbf{P} - \mathbf{P}^{(t-1)}\|_F^2, \quad (7)$$

$$\mathbf{X}^{(t+1)} = \underset{\mathbf{X}}{\operatorname{argmin}} F(\mathbf{X}, \mathbf{P}^{(t)}) + \lambda^{(t)} \|\mathbf{X} - \mathbf{X}^{(t)}\|_F^2, \quad (8)$$

where $\lambda^{(t)} > 0$. Our algorithm follows from the updates in (7) and (8). Similar to (5) and (6), the updates are a linear program and a regularized least squares problem, respectively. Assuming an r -local structure on \mathbf{P}^* , the update in (7) is equivalent to the simpler update of n/r blocks of size r . A summary of our algorithm is given in Algorithm 1.

Algorithmic details: We use the *collapsed* initialization, as also done in [13], to initialize $\hat{\mathbf{Y}}^{(0)} = \mathbf{B}\hat{\mathbf{X}}^{(0)}$. The initialization estimates $\hat{\mathbf{X}}$ from n/r measurements given by summing r consecutive rows in \mathbf{B} and \mathbf{Y} . The convergence criteria is based on the relative change in the objective $(F^{(t-1)} - F^{(t)})/F^{(t-1)}$. The total cost of updating all the local permutations by linear programs is $O(nr^2)$. In line 8, \dagger denotes the Moore-Penrose pseudo-inverse. The regularized least squares problem can be solved in $O(2nd^2 - \frac{2}{3}d^3)$,

4.1. Convergence analysis

Proposition 1. *The sequence of iterates $\{(\mathbf{X}^{(t)}, \mathbf{P}^{(t)})\}$ generated by (7), (8) converges to a first order stationary point of the objective in (4).*

Proof. The result follows from Theorem 9 in [20]. To use Theorem 9, we need to verify that the Kurdyka-Lojasiewicz (KL) inequality holds for the objective $F(\mathbf{X}, \mathbf{P}) = \|\mathbf{Y} - \mathbf{P}\mathbf{B}\mathbf{X}\|_F^2 + \mathbf{I}_C(\mathbf{P})$ where $\mathbf{I}_C(\mathbf{P})$ is the indicator function of the set of permutations. First, note that the set of permutations of order n is semi-algebraic because each of the $n!$ permutations can be specified via n^2 linear equality constraints.

Algorithm 1 Proximal Alt-Min for r -local unlabeled sensing

Input: $\mathbf{B}, \mathbf{Y}, \hat{\mathbf{Y}}^{(0)}, \hat{\mathbf{X}}^{(0)}, r, \lambda, \epsilon$

Output: $\hat{\mathbf{P}}$

```

1:  $t \leftarrow 0$ 
2:  $\hat{\mathbf{P}}_k^{(t)} \leftarrow \mathbf{1}_{r \times r} / r$ 
3: while relative change  $> \epsilon$  do
4:   for  $k \in 1 \dots n/r$  do
5:      $\hat{\mathbf{P}}_k = \underset{\mathbf{P}_k \in \Pi_r}{\operatorname{argmin}} -\langle \mathbf{Y}\hat{\mathbf{Y}}^{(t)} + \lambda\hat{\mathbf{P}}_k^{(t)}, \mathbf{P}_k \rangle$ 
6:   end for
7:    $\hat{\mathbf{P}}^{(t+1)} \leftarrow \operatorname{diag}(\hat{\mathbf{P}}_1, \dots, \hat{\mathbf{P}}_{n/r})$ 
8:    $\hat{\mathbf{X}}^{(t+1)} \leftarrow \left[ \frac{\mathbf{B}_{n \times d}}{\sqrt{\lambda} \mathbf{I}_{d \times d}} \right]^\dagger \left[ \frac{\hat{\mathbf{P}}^{(t+1)\top} \mathbf{Y}}{\sqrt{\lambda} \hat{\mathbf{X}}^{(t)}} \right]$ 
9:    $\hat{\mathbf{Y}}^{(t+1)} \leftarrow \mathbf{B}\hat{\mathbf{X}}^{(t+1)}$ 
10:   $t \leftarrow t + 1$ 
11: end while
```

The indicator function of a semi-algebraic function is a semi-algebraic function [20]. The first term in the objective is polynomial, and therefore real-analytic. The sum of a real-analytic function and a semi-algebraic function is sub-analytic, and sub-analytic functions satisfy the KL inequality [21].

5. EXPERIMENTS

For all experiments using Algorithm 1, we set the tolerance $\epsilon = 0.01$. The proximal regularization is set to $\lambda^{(1)} = 100$ and scaled as $\lambda^{(t+1)} = \lambda^{(t)}/10$. MATLAB code for the experiments is available at the first author's github account: github.com/aabbas02/Proximal-Alt-Min-for-ULS-UDGP.

5.1. Synthetic simulations

Data generation: The entries of the signal matrix \mathbf{X}^* , the sensing matrix \mathbf{B} , and the noise \mathbf{W} are sampled i.i.d. from the normal distribution. The matrix \mathbf{W} is subsequently scaled by σ to set a specific SNR. The permutation \mathbf{P}^* is picked uniformly at random from the set of r -local permutations. All results are averages of 75 Monte Carlo (MC) runs.

Baselines: We compare our algorithm to four algorithms proposed in [8, 11, 12, 13]. First is a biconvex relaxation of (4) proposed in [8] and solved via the alternating direction method of multipliers (ADMM) algorithm. Second is the Levsort algorithm in [11]. For $m = d$ views and $\sigma^2 = 0$, the Levsort algorithm provably recovers \mathbf{P}^* exactly. The third algorithm in [12] approximates $\hat{\mathbf{X}} \simeq \mathbf{B}^\top \mathbf{Y}$ in (5) and estimates $\hat{\mathbf{P}} = \underset{\mathbf{P} \in \Pi}{\operatorname{argmin}} \langle \mathbf{Y}\mathbf{Y}^\top \mathbf{B}\mathbf{B}^\top, \mathbf{P} \rangle$. The approximation is justified when matrix \mathbf{B} is Gaussian i.i.d. and the rows of \mathbf{B} are partially shuffled. In order to ensure a fair comparison with the algorithms in [8, 11, 12], we specify an additional constraint to constrain the permutation solutions to be r -local. For each MC run, the ADMM penalty parameter ρ for [8] was tuned in the range 10^{-3} to 10^5 and the best results were retained. We also compare to the algorithm in [13] that solves a quadratic assignment problem to estimate \mathbf{P}^* . We refer to the

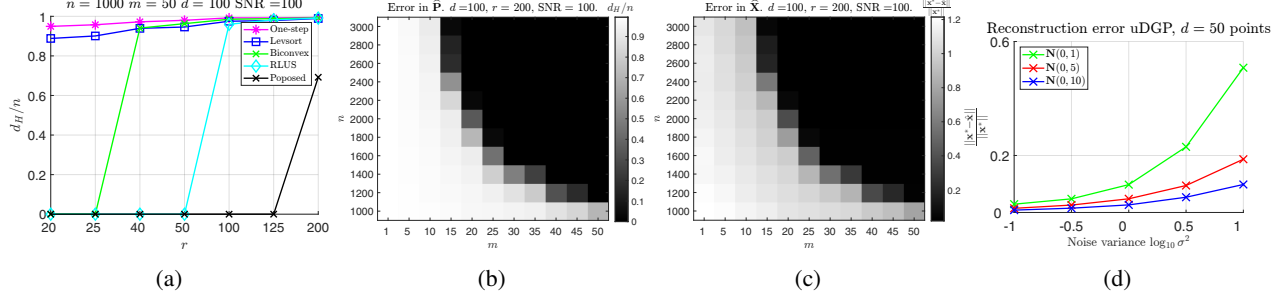


Fig. 3: Synthetic simulations for $\mathbf{Y} = \mathbf{P}^* \mathbf{B}_{n \times d} \mathbf{X}_{d \times m}^* + \mathbf{W}$. (a). Figure plots the fractional Hamming distortion d_H/n against block size r . The Hamming distortion is the number of mismatches in estimate $\hat{\mathbf{P}}$ of \mathbf{P}^* and is defined as $d_H = \sum_i \mathbb{1}(\hat{\mathbf{P}}(i) \neq \mathbf{P}^*(i))$, where $\mathbf{P}(i)$ denotes the column index of the 1 entry in the i^{th} row of the permutation matrix \mathbf{P} . (b,c). Figures plot d_H/n and the relative error $\|\hat{\mathbf{X}} - \mathbf{X}^*\|/\|\mathbf{X}^*\|$, respectively, against (m, n) for fixed (d, r) . (d). **uDGP**. Figure plots the relative error in the estimated point coordinates against increasing noise variance σ^2 . The $d = 50$ points are sampled i.i.d. from the normal distribution with variances $\{1, 5, 10\}$, and the permutation \mathbf{P}^* is r -local with blocks of sizes $(d - 1, \dots, 2)$.

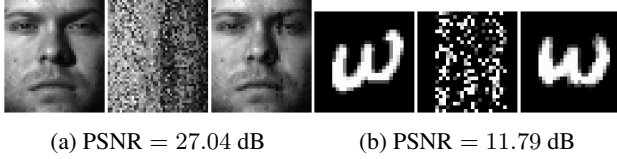


Fig. 4: (a). Left. Unscrambled image $\mathbf{y} \in \mathbb{R}^n$, $n = 48 \times 42$ from the YALE B dataset. Middle. Scrambled input image $\mathbf{P}^* \mathbf{y}$ via an r -local permutation, $r = 96$. Right. Reconstructed image $\hat{\mathbf{y}} = \hat{\mathbf{P}}^T \mathbf{P}^* \mathbf{y}$. (b). Left: Unscrambled image $\mathbf{y} \in \mathbb{R}^n$, $n = 28 \times 28$ from the MNIST dataset. Middle: Scrambled input image via an r -local permutation, $r = 28$. Right: Reconstructed image. The peak signal to noise ratio PSNR is defined in terms of the mean square reconstruction error $e = \frac{1}{n} \|\mathbf{y} - \hat{\mathbf{y}}\|_2^2$ as $\text{PSNR} = 10 \log_{10}(\frac{1}{e^2})$.

algorithms in [8, 11, 12, 13] as ‘Biconvex’, ‘Levsort’, ‘One-step’, and ‘RLUS’, respectively.

Results: In Figure 3a, we compare our proposed algorithm to the baselines. One-step [12] fails to recover \mathbf{P}^* for even small values of r , underscoring that algorithms to recover sparse permutations cannot be adapted to the r -local model. The Levsort algorithm [11] also fails to recover \mathbf{P}^* . The Biconvex algorithm [8] only recovers \mathbf{P}^* for $r = 20$. RLUS [13] gives comparable recovery of \mathbf{P}^* when $r \leq 50$, but is outperformed by the proposed algorithm at higher values of r . In Figures 3b and 3c, we set $(d, r) = (100, 200)$ and vary (n, m) . The errors in $\hat{\mathbf{X}}, \hat{\mathbf{P}}$ decrease as n increases. This is because the initialization to the algorithm (see section 4) improves with increasing n . The estimates also improve with increasing m . This observation agrees with the discussion in section 1 and the result in (2).

5.2. Scrambled image restoration

We consider a variation of the problem in [15]. Therein, given scrambled images $\mathbf{P}^* \mathbf{y}$ and unscrambled training data \mathbf{y} , a neural network is trained to unscramble $\mathbf{P}^* \mathbf{y}$. For our experiment, the images are drawn from the YALE B dataset [22]

	Proposed	RLUS [13]
MNIST	12.31 ± 2.57	12.03 ± 2.47
YALE	28.30 ± 2.87	26.43 ± 2.95

Table 1: Scrambled image restoration problem. For each dataset, the PSNR values (dB) are averaged over 10 classes.

and the MNIST dataset [23]. For each class in the MNIST (YALE B) dataset, the matrix $\mathbf{B} \in \mathbb{R}^{n \times d}$ contains the $d = 5$ (10) principal components of the unscrambled data. A class in the MNIST (YALE B) dataset comprises all images corresponding to a single digit (particular face). The problem is to recover the unknown permutation \mathbf{P}^* given the scrambled image $\mathbf{P}^* \mathbf{y}$ and the matrix \mathbf{B} . The setup is summarized in Figure 4. We compare our algorithm to RLUS [13] and the results are given in Table 1. The results show that the proposed algorithm outperforms RLUS [13].

5.3. 1-d uDGP

We consider the 1-dimensional uDGP problem (Section 3) where the distances are corrupted with i.i.d. Gaussian noise of variance σ^2 . The results in Figure 3d show that the relative error in the recovered points is low even for high noise variance. The results are noteworthy because the general 1-d uDGP problem with noisy distances is NP-hard [24].

6. CONCLUSION

We propose a proximal alternating minimization algorithm for the unlabeled sensing problem and apply it to the setting where the unknown permutation is r -local. The resulting algorithm is efficient and theoretically converges to a first order stationary point. Experiments on synthetic and real data show that the algorithm outperforms competing baselines. We formulate the 1-d unassigned distance geometry problem (uDGP) as an unlabeled sensing problem with a structured measurement matrix. Future work will explore information-theoretic inachievability results for the r -local permutation model and uDGP.

7. REFERENCES

- [1] L. Peng, X. Song, MC Tsakiris, H. Choi, L. Kneip, and Y. Shi, "Algebraically-initialized expectation maximization for header-free communication," in *IEEE Int. Conf. on Acous., Speech and Signal Process. (ICASSP)*. IEEE, 2019, pp. 5182–5186.
- [2] A. Balakrishnan, "On the problem of time jitter in sampling," *IRE Trans. on Inf. Theory*, vol. 8, no. 3, pp. 226–236, 1962.
- [3] Zhidong Bai and Tailen Hsing, "The broken sample problem," *Probability theory and related fields*, vol. 131, no. 4, pp. 528–552, 2005.
- [4] Manuel Marques, Marko Stošić, and João Costeira, "Subspace matching: Unique solution to point matching with geometric constraints," in *2009 IEEE 12th Int. Conf. on Computer Vision*, 2009, pp. 1288–1294.
- [5] Jayakrishnan Unnikrishnan, Saeid Haghighatshoar, and Martin Vetterli, "Unlabeled sensing with random linear measurements," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3237–3253, 2018.
- [6] I. Dokmanić, "Permutations unlabeled beyond sampling unknown," *IEEE Signal Processing Letters*, vol. 26, no. 6, pp. 823–827, 2019.
- [7] A. Pananjady, M. J. Wainwright, and T. A. Courtade, "Linear regression with shuffled data: Statistical and computational limits of permutation recovery," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3286–3300, 2018.
- [8] H. Zhang, M. Slawski, and P. Li, "Permutation recovery from multiple measurement vectors in unlabeled sensing," in *2019 IEEE Int. Symposium on Inf. Theory (ISIT)*, 2019, pp. 1857–1861.
- [9] L. Peng and M.C. Tsakiris, "Linear regression without correspondences via concave minimization," *IEEE Signal Proces. Letters*, vol. 27, pp. 1580–1584, 2020.
- [10] M. Slawski, E. Ben-David, and P. Li, "Two-stage approach to multivariate linear regression with sparsely mismatched data," *J. Mach. Learn. Res.*, vol. 21, no. 204, pp. 1–42, 2020.
- [11] A. Pananjady, M. J. Wainwright, and T. A. Courtade, "Denoising linear models with permuted data," in *2017 IEEE Int. Symposium on Inf. Theory (ISIT)*, 2017, pp. 446–450.
- [12] Hang Zhang and Ping Li, "Optimal estimator for unlabeled linear regression," in *Proceedings of the 37th Int. Conf. on Machine Learning (ICML)*, 2020.
- [13] Ahmed Ali Abbasi, Abiy Tasissa, and Shuchin Aeron, "R-local unlabeled sensing: A novel graph matching approach for multiview unlabeled sensing under local permutations," *IEEE Open Journal of Signal Process.*, vol. 2, pp. 309–317, 2021.
- [14] Shuai Huang and Ivan Dokmanic, "Reconstructing point sets from distance distributions," *IEEE Trans. Signal Process.*, vol. 69, pp. 1811–1827, 2021.
- [15] Gonzalo Mena, David Belanger, Scott Linderman, and Jasper Snoek, "Learning latent permutations with gumbel-sinkhorn networks," 2018.
- [16] P. Lahiri and M. D Larsen, "Regression analysis with linked data," *Jrnl of the American statistical association*, vol. 100, no. 469, pp. 222–230, 2005.
- [17] M. Slawski, M. Rahmani, and Ping Li, "A sparse representation-based approach to linear regression with partially shuffled labels," in *Uncertainty in Artificial Intelligence*. PMLR, 2020, pp. 38–48.
- [18] P. Duxbury, C. Lavor, L. Liberti, and L. de Salles-Neto, "Unassigned distance geometry and molecular conformation problems," *Journal of Global Optimization*, pp. 1–10, 2021.
- [19] Tamara Dakic, *On the turnpike problem*, Simon Fraser University BC, Canada, 2000.
- [20] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the kurdyka-Łojasiewicz inequality," *Mathematics of Operations Research*, vol. 35, no. 2, pp. 438–457, 2010.
- [21] Yangyang Xu and Wotao Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM Journal on imaging sciences*, vol. 6, no. 3, pp. 1758–1789, 2013.
- [22] A.S. Georgiades, P.N. Belhumeur, and D.J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.
- [23] Li Deng, "The mnist database of handwritten digit images for machine learning research," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.
- [24] Mark Cieliebak and Stephan Eidenbenz, "Measurement errors make the partial digest problem np-hard," in *Latin American Symposium on Theoretical Informatics*. Springer, 2004, pp. 379–390.