

On the Robustness of Cross-Concentrated Sampling for Matrix Completion

HanQin Cai

Department of Statistics and Data Science
Department of Computer Science
University of Central Florida
Orlando, FL 32816, USA
hqcai@ucf.edu

Chandra Kundu

Department of Statistics and Data Science
University of Central Florida
Orlando, FL 32816, USA
chandra.kundu@ucf.edu

Longxiu Huang

Department of Computational Mathematics, Science and Engineering
Department of Mathematics
Michigan State University
East Lansing, MI 48824, USA
huangl3@msu.edu

Bowen Su

Department of Mathematics
Michigan State University
East Lansing, MI 48824, USA
subowen@msu.edu

Abstract—Matrix completion is one of the crucial tools in modern data science research. Recently, a novel sampling model for matrix completion coined cross-concentrated sampling (CCS) has caught much attention. However, the robustness of the CCS model against sparse outliers remains unclear in the existing studies. In this paper, we aim to answer this question by exploring a novel Robust CCS Completion problem. A highly efficient non-convex iterative algorithm, dubbed Robust CUR Completion (RCURC), is proposed. The empirical performance of the proposed algorithm, in terms of both efficiency and robustness, is verified in synthetic and real datasets.

Index Terms—Robust matrix completion, cross-concentrated sampling, CUR decomposition, outlier detection.

I. INTRODUCTION

MATRIX completion problem [1, 2] aims to recover the underlying low-rank matrix \mathbf{X} from some entry-wise partial observation. It has received tremendous attention in the past decades and has been widely applied in real-world applications such as collaborative filtering [3, 4], image processing [5, 6], and signal processing [7, 8]. While the vanilla matrix completion studies are often based on entry-wise uniform or Bernoulli sampling models, the recent development of a novel sampling model called *Cross-Concentrated Sampling* (CCS) [9] has caught the attention of the matrix completion community. The CCS model is summarized in Procedure 1.

As illustrated in Figure 1, the CCS model bridges two popular sampling models for matrix completion: *Uniform sampling* and *CUR sampling* which recovers the underlying low-rank matrix from observed rows and columns. The CCS model provides extra flexibility, with a theoretical guideline, for samples concentrated on certain rows and columns. This model has the potential for more cost- and time-efficient sampling strategies in applications such as recommendation systems.

The authors contributed equally. This work was partially supported by NSF DMS 2304489. (Corresponding authors: HanQin Cai and Longxiu Huang.)

Procedure 1 Cross-Concentrated Sampling (CCS) [9]

- 1: **Input:** \mathbf{Y} : access to the data matrix.
 - 2: Uniformly choose row and column indices \mathcal{I}, \mathcal{J} .
 - 3: Set $\mathbf{R} = [\mathbf{Y}]_{\mathcal{I}, :}$ and $\mathbf{C} = [\mathbf{Y}]_{:, \mathcal{J}}$.
 - 4: Uniformly sample entries in \mathbf{R} and \mathbf{C} , then record the sampled locations as $\Omega_{\mathbf{R}}$ and $\Omega_{\mathbf{C}}$, respectively.
 - 5: **Output:** $[\mathbf{Y}]_{\Omega_{\mathbf{R}} \cup \Omega_{\mathbf{C}}}$, $\Omega_{\mathbf{R}}$, $\Omega_{\mathbf{C}}$, \mathcal{I} , \mathcal{J} .
-

However, data in real-world applications are often corrupted by sparse outliers, and uniformly sampled data must be processed by some robust completion algorithms for reliable recovery results. That being said, a crucial question has to be answered before the CCS model can be widely used: “*Is CCS-based matrix completion robust to sparse outliers under some robust algorithms, like the uniform sampling model?*”

Mathematically, provided the partial observation $\mathcal{P}_{\Omega}(\mathbf{Y})$ of a corrupted data matrix $\mathbf{Y} = \mathbf{X} + \mathbf{S}$, we study the Robust CCS Completion problem:

$$\begin{aligned} & \underset{\mathbf{X}, \mathbf{S}}{\text{minimize}} \quad \frac{1}{2} \langle \mathcal{P}_{\Omega}(\mathbf{X} + \mathbf{S} - \mathbf{Y}), \mathbf{X} + \mathbf{S} - \mathbf{Y} \rangle \\ & \text{subject to} \quad \text{rank}(\mathbf{X}) = r \quad \text{and} \quad \mathbf{S} \text{ is } \alpha\text{-sparse}, \end{aligned} \quad (1)$$

where \mathcal{P} is the sampling operator defined as:

$$\mathcal{P}_{\Omega}(\mathbf{Y}) = \sum_{(i,j) \in \Omega} [\mathbf{Y}]_{i,j} \mathbf{e}_i \mathbf{e}_j^T$$

and the observation set Ω is generated according to the CCS model. Similar to the standard robust matrix completion, a sparse variable \mathbf{S} is introduced for better outlier tolerance, thus one can recover the underlying low-rank matrix \mathbf{X} accurately.

A. Related Work

The problem of low-rank matrix recovery in the presence of sparse outliers has been well-studied under the settings

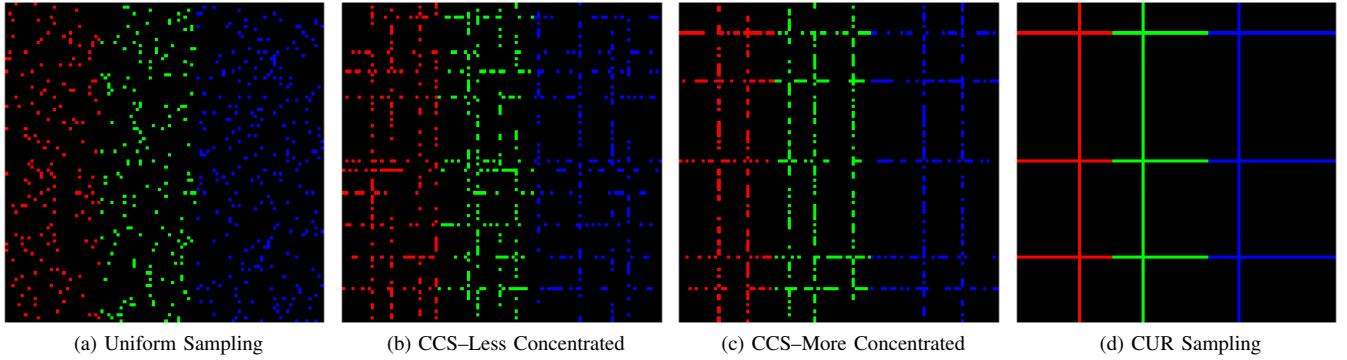


Fig. 1: [9] Visual illustrations of different sampling schemes. From left to right, sampling methods change from the uniform sampling style to the CUR sampling style with the same total observation rate. Colored pixels indicate observed entries, while black pixels mean missing entries.

of uniform sampling and Bernoulli sampling. This problem is known as *robust principal component analysis* (RPCA) when the corrupted data matrix is fully observed, and it is called *robust matrix completion* (RMC) when data is partially observed. The seminal work [10] considers both RPCA and RMC problems via convex relaxed formulations and provides recovery guarantees. In particular, under the μ -incoherence assumption for the low-rank \mathbf{X} , [10] requires the positions of outliers to be placed uniformly at random, and at least $0.1n^2$ entries are observed uniformly at random. Later, a series of non-convex algorithms [11–16] tackle RPCA and/or RMC problems with an improved, non-random α -sparsity assumptions for the outlier matrix \mathbf{S} . The typical recovery guarantee shows a linear convergence of a non-convex algorithm, provided $\alpha \leq \mathcal{O}(1/\text{poly}(\mu r))$; moreover, $\mathcal{O}(\text{poly}(\mu r)\text{polylog}(n)n)$ random samples are typically required for the RMC cases. Another line of work [7, 8, 17–19] focuses on the robust recovery of structured low-rank matrices, e.g., Hankel matrices, and they typically require merely $\mathcal{O}(\text{poly}(\mu r)\text{polylog}(n))$ samples by utilizing the structure, even in the presence of structured outliers. More recently, [20–22] study the robust CUR decomposition problem, that is, recovering the low-rank matrix from row- and column-wise observations with entry-wise corruptions. The studies show that robust CUR decomposition is as robust as RPCA, with better computational efficiency and data interoperability.

On the other hand, [9] shows that CCS-based matrix completion requires $\mathcal{O}(\mu^2 r^2 n \log^2 n)$ samples which is only a factor of $\log n$ worse than the state-of-the-art result; however, its outlier tolerance has not been studied.

B. Notation

For a matrix \mathbf{M} , $[\mathbf{M}]_{i,j}$, $[\mathbf{M}]_{\mathcal{I},:}$, $[\mathbf{M}]_{:,j}$, and $[\mathbf{M}]_{\mathcal{I},\mathcal{J}}$ denote its (i, j) -th entry, its row submatrix with row indices \mathcal{I} , its column submatrix with column indices \mathcal{J} , and its submatrix with row indices \mathcal{I} and column indices \mathcal{J} , respectively. $\|\mathbf{M}\|_{\text{F}} := (\sum_{i,j} [\mathbf{M}]_{i,j}^2)^{1/2}$ denotes the Frobenius norm, $\|\mathbf{M}\|_{2,\infty} := \max_i (\sum_j [\mathbf{M}]_{i,j}^2)^{1/2}$ denotes the largest row-wise

ℓ_2 -norm, $\|\mathbf{M}\|_{\infty} = \max_{i,j} |[\mathbf{M}]_{i,j}|$ denote the largest entry-wise magnitude, \mathbf{M}^\dagger represents the Moore–Penrose inverse of \mathbf{M} , and \mathbf{M}^\top is the transpose of \mathbf{M} . $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product. We denote e_i as the i -th standard basis of the real vector space. The symbol $[n]$ denotes the set $\{1, 2, \dots, n\}$ for all $n \in \mathbb{Z}^+$. Throughout the paper, uniform sampling is referred to as uniform sampling with replacement.

II. PRELIMINARIES

A. Assumptions

As discussed in the related work, μ -incoherence and α -sparsity are commonly used assumptions in RMC problems. We will use the same assumptions in this study and give the formal definition here:

Assumption 1 (μ -incoherence). *Let $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$ be a rank- r matrix. It is said μ -incoherent if*

$$\|\mathbf{U}\|_{2,\infty} \leq \sqrt{\frac{\mu r}{n_1}} \quad \text{and} \quad \|\mathbf{V}\|_{2,\infty} \leq \sqrt{\frac{\mu r}{n_2}}$$

for some positive μ , where $\mathbf{U}\Sigma\mathbf{V}^\top$ is the compact SVD of \mathbf{X} .

Assumption 2 (α -sparsity). *$\mathbf{S} \in \mathbb{R}^{n_1 \times n_2}$ is said α -sparse if there are at most α fraction of non-zero entries in each of its row and column. That is, for all $i \in [n_1]$ and $j \in [n_2]$, it holds*

$$\|[S]_{i,:}\|_0 \leq \alpha n_2 \quad \text{and} \quad \|[S]_{:,j}\|_0 \leq \alpha n_1.$$

Note that no randomness is required for the sparsity pattern.

B. CUR Approximation

CUR approximation, also known as skeleton decomposition, is the backbone of our algorithm design. We shall go over some background of CUR approximation for the reader's interest. CUR approximation is a self-expressive technique that fills the interpretability gap in matrix dimensionality reduction. Given a rank- r matrix $\mathbf{X} \in \mathbb{R}^{n \times n}$, if we select rows and columns from \mathbf{X} that span its row and column spaces respectively, then \mathbf{X} can be reconstructed from these submatrices. This has been affirmed in previous studies:

Theorem 3 ([23]). Consider row and column index sets $\mathcal{I}, \mathcal{J} \subseteq [n]$. Denote submatrices $\mathbf{C} = [\mathbf{X}]_{:, \mathcal{J}}$, $\mathbf{U} = [\mathbf{X}]_{\mathcal{I}, :}$ and $\mathbf{R} = [\mathbf{X}]_{\mathcal{I}, \mathcal{J}}$. If $\text{rank}(\mathbf{U}) = \text{rank}(\mathbf{X})$, then $\mathbf{X} = \mathbf{C}\mathbf{U}^\dagger\mathbf{R}$.

For further details of CUR approximation, including the historical context and formal proof, one can refer to [23] and reference therein. Note that through various row- and column-wise sampling methodologies, meeting this rank equivalence condition in Theorem 3 is highly probable, especially when a sufficient number of rows and columns are sampled. An example of this is detailed in Theorem 4.

Theorem 4 ([24, Theorem 1.1]). Let \mathbf{X} satisfy Assumption 1, and suppose we sample $|\mathcal{I}| = \mathcal{O}(r \log n)$ rows and $|\mathcal{J}| = \mathcal{O}(r \log n)$ columns uniformly at random. Then $\text{rank}(\mathbf{U}) = \text{rank}(\mathbf{X})$ with probability at least $1 - \mathcal{O}(rn^{-2})$.

There also exist many other sampling methodologies that ensure the recoverability of CUR approximation; for instance, the Discrete Empirical Interpolation Method (DEIM) [25, 26] and leverage scores sampling method [27]. The interested reader can find more details in [28, Table 1]. Nonetheless, they all sample completed rows and columns, thus are substantially different than the CCS model.

III. PROPOSED ALGORITHM

In this section, we will present a novel non-convex algorithm for solving the proposed Robust CCS Completion problem (1). This iterative algorithm is based on the idea of projected gradient descent where CUR approximation is used for fast low-rank approximation in each iteration. The algorithm is dubbed Robust CUR Completion (RCURC) and summarized in Algorithm 2.

Algorithm 2 Robust CUR Completion (RCURC)

```

1: Input:  $[\mathbf{Y} = \mathbf{X} + \mathbf{S}]_{\Omega_R \cup \Omega_C}$ : observed data;  $\Omega_R, \Omega_C$ :  
observation locations;  $\mathcal{I}, \mathcal{J}$ : row and column indices that  
define  $\mathbf{R}$  and  $\mathbf{C}$  respectively;  $\eta_R, \eta_C$ : step sizes;  $r$ : target  
rank;  $\varepsilon$ : target precision level;  $\zeta_0$ : initial thresholding value;  
 $\gamma$ : thresholding decay parameter.  
2:  $\mathbf{X}_0 = \mathbf{0}$ ;  $k = 0$   
3: while  $e_k > \varepsilon$  do ▷  $e_k$  is defined in (3)  
4:   // Updating sparse component  
5:    $\zeta_{k+1} = \gamma^k \zeta_0$   
6:    $[\mathbf{S}_{k+1}]_{\mathcal{I}, :} = \mathcal{T}_{\zeta_{k+1}} [\mathbf{Y} - \mathbf{X}_k]_{\mathcal{I}, :}$   
7:    $[\mathbf{S}_{k+1}]_{:, \mathcal{J}} = \mathcal{T}_{\zeta_{k+1}} [\mathbf{Y} - \mathbf{X}_k]_{:, \mathcal{J}}$   
8:   // Updating low-rank component  
9:    $\mathbf{R}_{k+1} = [\mathbf{X}_k]_{\mathcal{I}, :} + \eta_R [\mathcal{P}_{\Omega_R} (\mathbf{Y} - \mathbf{X}_k - \mathbf{S}_{k+1})]_{\mathcal{I}, :}$   
10:   $\mathbf{C}_{k+1} = [\mathbf{X}_k]_{:, \mathcal{J}} + \eta_C [\mathcal{P}_{\Omega_C} (\mathbf{Y} - \mathbf{X}_k - \mathbf{S}_{k+1})]_{:, \mathcal{J}}$   
11:   $\mathbf{U}_{k+1} = \mathcal{D}_r([\mathbf{R}_{k+1}]_{:, \mathcal{J}} \uplus [\mathbf{C}_{k+1}]_{\mathcal{I}, :})$   
12:   $[\mathbf{R}_{k+1}]_{:, \mathcal{J}} = \mathbf{U}_{k+1}$  and  $[\mathbf{C}_{k+1}]_{\mathcal{I}, :} = \mathbf{U}_{k+1}$   
13:   $\mathbf{X}_{k+1} = \mathbf{C}_{k+1} \mathbf{U}_{k+1}^\dagger \mathbf{R}_{k+1}$  ▷ Do not compute  
14:   $k = k + 1$   
15: Output:  $\mathbf{C}_k, \mathbf{U}_k, \mathbf{R}_k$ : The CUR components of the estimated low-rank matrix.

```

We will go over the algorithm step by step in the following paragraphs. For the ease of the presentation, we start with the low-rank component.

Updating low-rank component. To utilize the structure of cross-concentrated samples, it is efficient to enforce the low-rank constraint of \mathbf{X} with the CUR approximation technique. Let $\mathbf{R} = [\mathbf{X}]_{\mathcal{I}, :}$, $\mathbf{C} = [\mathbf{X}]_{:, \mathcal{J}}$, and $\mathbf{U} = [\mathbf{X}]_{\mathcal{I}, \mathcal{J}}$. Applying gradient descent directly on \mathbf{R} and \mathbf{C} gives:

$$\begin{aligned} \mathbf{R}_{k+1} &= [\mathbf{X}]_{\mathcal{I}, :} + \eta_R [\mathcal{P}_{\Omega_R} (\mathbf{Y} - \mathbf{X}_k - \mathbf{S}_{k+1})]_{\mathcal{I}, :} \\ \mathbf{C}_{k+1} &= [\mathbf{X}]_{:, \mathcal{J}} + \eta_C [\mathcal{P}_{\Omega_C} (\mathbf{Y} - \mathbf{X}_k - \mathbf{S}_{k+1})]_{:, \mathcal{J}}, \end{aligned}$$

where η_R and η_C are the stepsizes. However, When it comes to the intersection submatrix \mathbf{U} , it is more complicated as Ω_R and Ω_C can have overlaps. We abuse the notation \uplus and define an operator called *union sum* here:

$$[\mathbf{R}_{k+1}]_{:, \mathcal{J}} \uplus [\mathbf{C}_{k+1}]_{\mathcal{I}, :} = \begin{cases} [\mathbf{R}_{k+1}]_{i, j} & \text{if } (i, j) \in \Omega_R \setminus \Omega_C; \\ [\mathbf{C}_{k+1}]_{i, j} & \text{if } (i, j) \in \Omega_C \setminus \Omega_R; \\ \frac{\eta_R \eta_C}{\eta_R + \eta_C} \left(\frac{[\mathbf{R}_{k+1}]_{i, j}}{\eta_R} + \frac{[\mathbf{C}_{k+1}]_{i, j}}{\eta_C} \right) & \text{if } (i, j) \in \Omega_C \cap \Omega_R; \\ 0 & \text{otherwise.} \end{cases}$$

Basically, we take whatever value we have for the non-overlapped entries and take a weighted average for the overlaps where the weights are determined by the stepsizes used in the updates of \mathbf{R}_{k+1} and \mathbf{C}_{k+1} . To ensure the rank- r constraint, at least one of \mathbf{C}_{k+1} , \mathbf{U}_{k+1} or \mathbf{R}_{k+1} should be rank- r . For computational efficiency, we choose to put it on the smallest one. Thus,

$$\mathbf{U}_{k+1} = \mathcal{D}_r([\mathbf{R}_{k+1}]_{:, \mathcal{J}} \uplus [\mathbf{C}_{k+1}]_{\mathcal{I}, :}),$$

where \mathcal{D}_r is the best rank- r approximation operator via truncated SVD. After replacing the intersection part \mathbf{U}_{k+1} in the previously updated \mathbf{R}_{k+1} and \mathbf{C}_{k+1} , we have the new estimation of low-rank component:

$$\mathbf{X}_{k+1} = \mathbf{C}_{k+1} \mathbf{U}_{k+1}^\dagger \mathbf{R}_{k+1}. \quad (2)$$

However, (2) is just a conceptual step and one should never compute it. In fact, the full matrix \mathbf{X} is never needed and should not be formed in the algorithm as updating the corresponding CUR components is sufficient.

Updating sparse component. We detect the outliers and put them into the sparse matrix \mathbf{S} via hard-thresholding operator:

$$[\mathcal{T}_\zeta(\mathbf{M})]_{i, j} = \begin{cases} 0 & \text{if } |[\mathbf{M}]_{i, j}| < \zeta; \\ [\mathbf{M}]_{i, j} & \text{otherwise.} \end{cases}$$

The hard-thresholding on residue $\mathbf{Y} - \mathbf{X}_k$, paired with iterative decayed thresholding values:

$$\zeta_{k+1} = \gamma^k \zeta_0 \quad \text{with some } \gamma \in (0, 1),$$

has shown promising performance in outlier detection in prior art [14, 16, 29, 30]. Notice that we only need to remove outliers located on the selected rows and columns, i.e., \mathbf{R} and \mathbf{C} , since

they are the only components needed to update the low-rank component later. Therefore, for computational efficiency, we should only compute \mathbf{X}_k on the selected rows and columns to update \mathbf{S}_{k+1} correspondingly—as said, one should never compute the full \mathbf{X}_k in this algorithm. In particular,

$$[\mathbf{X}_k]_{\mathcal{I}, :} = [\mathbf{C}_k]_{\mathcal{I}, :} \mathbf{U}_k^\dagger \mathbf{R}_k \quad \text{and} \quad [\mathbf{X}_k]_{:, \mathcal{J}} = \mathbf{C}_k \mathbf{U}_k^\dagger [\mathbf{R}_k]_{:, \mathcal{J}}.$$

Stopping criteria and stepsizes. We finish the algorithm with the last few pieces. The stopping criteria is set to be $e_k \leq \varepsilon$ where ε is the targeted accuracy and the computational error is

$$e_k = \frac{\langle \mathcal{P}_{\Omega_R \cup \Omega_C}(\mathbf{S}_k + \mathbf{X}_k - \mathbf{Y}), \mathbf{S}_k + \mathbf{X}_k - \mathbf{Y} \rangle}{\langle \mathcal{P}_{\Omega_R \cup \Omega_C} \mathbf{Y}, \mathbf{Y} \rangle}. \quad (3)$$

The recommended stepsizes are $\eta_R = \frac{1}{p_R}$ and $\eta_C = \frac{1}{p_C}$ where p_R and p_C are the observation rates of Ω_R and Ω_C respectively. Smaller stepsizes should be used with larger α , i.e., more outliers.

IV. NUMERICAL EXPERIMENTS

In this section, we will verify the empirical performance of RCURC with both synthetic and real datasets. All experiments are implemented on Matlab R2022b and executed on a laptop equipped with Intel i7-11800H CPU (2.3GHz @ 8 cores) and 16GB DDR4 RAM.

A. Synthetic Datasets

In this simulation, we assess the computational efficiency of our algorithm, RCURC, in addressing the robust CCS completion problem. We construct $\mathbf{Y} = \mathbf{X} + \mathbf{S}$, a $d \times d$ matrix with $d = 3000$, where $\mathbf{X} = \mathbf{WV}^\top$ is a randomly generated rank- r matrix. To create the sparse outlier tensor \mathbf{S} , we randomly select α percent entries to form the support of \mathbf{S} . The values of the non-zero entries are then uniformly sampled from the range $[-c\mathbb{E}(|S_{ij}|), c\mathbb{E}(|S_{ij}|)]$. To generate the robust CCS completion problems, we set $\alpha = 0.2$, $\frac{|\mathcal{I}|}{d} = \frac{|\mathcal{J}|}{d} = 30\%$, and $\frac{|\Omega_R|}{|\mathcal{I}|d} = \frac{|\Omega_C|}{|\mathcal{J}|d} = 25\%$. The results are obtained by averaging over 50 runs and reported in Figure 2. Both figures in Figure 2 depict the relationship between the relative error e_k and computational time for our RCURC method with varying rank r and outlier amplification factor c . It is noteworthy that RCURC consistently achieves nearly linear convergence rates across different scenarios.

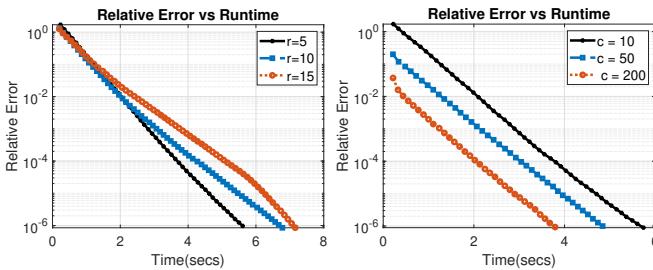


Fig. 2: Empirical convergence of RCURC. **Left:** $c = 10$ and varying r . **Right:** $r = 5$ and varying c .

TABLE I: Video size, runtime and PSNR. Herein \mathcal{S} represents *shoppingmall* and \mathcal{R} represents *restaurant*.

	FRAME SIZE	FRAME NUMBER	RUNTIME	PSNR
\mathcal{S}	256×320	1000	33.12s	41.87
\mathcal{R}	120×160	3055	31.16s	39.84

B. Video Background Subtraction

We have applied our robust CCS model and RCURC to the problem of background separation. Two popular videos, *shoppingmall* and *restaurant*, are used as benchmarks. The dimensions of the datasets are documented in Table I.

To evaluate the performance of our method, we apply IRCUR [31] to the entire video dataset. We use the background derived from IRCUR as the benchmark ground truth. Subsequently, we randomly select 40% of the rows and 40% of the columns to create subrow and subcolumn matrices. At the next stage, we generate 30% observations on these submatrices. The reconstructed backgrounds obtained using RCURC are illustrated in Figure 3. It is evident that the background recovered by RCURC closely resembles that obtained via IRCUR. Furthermore, we utilize the Peak Signal-to-Noise Ratio (PSNR) as a quantitative measure to evaluate the reconstruction quality of RCURC in comparison to the results achieved with IRCUR. These findings are detailed in Table I.

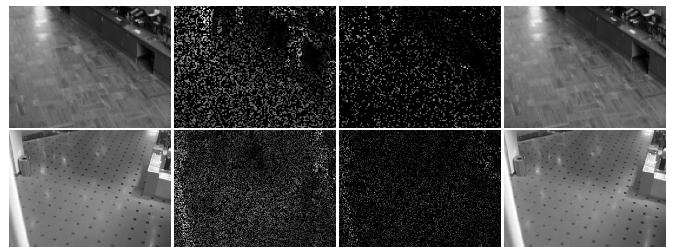


Fig. 3: Video background subtraction results: The first row shows *shoppingmall*, the second *restaurant*. Column one displays the ground truth background via IRCUR from full datasets. Columns two and three show two observed frames from our CCS model. The last column presents the background from RCURC using partial videos.

V. CONCLUSION REMARKS

This paper introduces a novel mathematical model for robust matrix completion problems with cross-concentrated samples. A highly efficient non-convex algorithm, dubbed RCURC, has been developed for the proposed model. The key techniques are projected gradient descent and CUR approximation. The numerical experiments, with both synthetic and real datasets, show high potential. In particular, we consistently observe linear convergence on RCURC.

As for future work, we will study the statistical properties of the proposed robust CCS completion model, such as theoretical sample complexities and outlier tolerance. The

recovery guarantee with a linear convergence rate will also be established for RCURC. We will also explore other real-world applications that suit the proposed model.

REFERENCES

- [1] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
- [2] B. Recht, “A simpler approach to matrix completion.” *J. Mach. Learn. Res.*, vol. 12, no. 12, 2011.
- [3] J. Bennett, C. Elkan, B. Liu, P. Smyth, and D. Tikk, “KDD cup and workshop 2007,” *SIGKDD Explor. Newsl.*, vol. 9, no. 2, p. 51–52, 2007.
- [4] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, “Using collaborative filtering to weave an information tapestry,” *Commun. ACM*, vol. 35, no. 12, pp. 61–70, 1992.
- [5] P. Chen and D. Suter, “Recovering the missing components in a large noisy low-rank matrix: Application to sfm,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1051–1063, 2004.
- [6] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, “Fast and accurate matrix completion via truncated nuclear norm regularization,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2117–2130, 2012.
- [7] J.-F. Cai, T. Wang, and K. Wei, “Fast and provable algorithms for spectrally sparse signal reconstruction via low-rank Hankel matrix completion,” *Appl. Comput. Harmon. Anal.*, vol. 46, no. 1, pp. 94–121, 2019.
- [8] H. Cai, J.-F. Cai, and J. You, “Structured gradient descent for fast robust low-rank Hankel matrix completion,” *SIAM J. Sci. Comput.*, vol. 45, no. 3, pp. A1172–A1198, 2023.
- [9] H. Cai, L. Huang, P. Li, and D. Needell, “Matrix completion with cross-concentrated sampling: Bridging uniform sampling and CUR sampling,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10100–10113, 2023.
- [10] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” *Journal of the ACM*, vol. 58, no. 3, pp. 1–37, 2011.
- [11] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain, “Non-convex robust PCA,” in *Adv. Neural Inf. Process Syst.*, 2014, pp. 1107–1115.
- [12] X. Yi, D. Park, Y. Chen, and C. Caramanis, “Fast algorithms for robust PCA via gradient descent,” in *Adv. Neural Inf. Process Syst.*, 2016, pp. 4152–4160.
- [13] T. Zhang and Y. Yang, “Robust pca by manifold optimization,” *J. Mach. Learn. Res.*, vol. 19, no. 1, pp. 3101–3139, 2018.
- [14] H. Cai, J.-F. Cai, and K. Wei, “Accelerated alternating projections for robust principal component analysis,” *J. Mach. Learn. Res.*, vol. 20, no. 1, pp. 685–717, 2019.
- [15] T. Tong, C. Ma, and Y. Chi, “Accelerating ill-conditioned low-rank matrix estimation via scaled gradient descent,” *J. Mach. Learn. Res.*, vol. 22, no. 150, pp. 1–63, 2021.
- [16] H. Cai, J. Liu, and W. Yin, “Learned robust PCA: A scalable deep unfolding approach for high-dimensional outlier detection,” in *Adv. Neural Inf. Process Syst.*, vol. 34, 2021, pp. 16977–16989.
- [17] Y. Chen and Y. Chi, “Robust spectral compressed sensing via structured matrix completion,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6576–6601, 2014.
- [18] S. Zhang and M. Wang, “Correction of corrupted columns through fast robust hankel matrix completion,” *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2580–2594, 2019.
- [19] H. Cai, J.-F. Cai, T. Wang, and G. Yin, “Accelerated structured alternating projections for robust spectrally sparse signal recovery,” *IEEE Trans. Signal Process.*, vol. 69, pp. 809–821, 2021.
- [20] H. Cai, K. Hamm, L. Huang, J. Li, and T. Wang, “Rapid robust principal component analysis: CUR accelerated inexact low rank estimation,” *IEEE Signal Process. Lett.*, vol. 28, pp. 116–120, 2021.
- [21] H. Cai, K. Hamm, L. Huang, and D. Needell, “Robust CUR decomposition: Theory and imaging applications,” *SIAM J. Imaging Sci.*, vol. 14, no. 4, pp. 1472–1503, 2021.
- [22] K. Hamm, M. Meskini, and H. Cai, “Riemannian CUR decompositions for robust principal component analysis,” in *Topological, Algebraic and Geometric Learning Workshops*, 2022, pp. 152–160.
- [23] K. Hamm and L. Huang, “Perspectives on CUR decompositions,” *Appl. Comput. Harmon. Anal.*, vol. 48, no. 3, pp. 1088–1099, 2020.
- [24] J. Chiu and L. Demanet, “Sublinear randomized algorithms for skeleton decompositions,” *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 1361–1383, 2013.
- [25] S. Chaturantabut and D. C. Sorensen, “Nonlinear model reduction via discrete empirical interpolation,” *SIAM J. Sci. Comput.*, vol. 32, no. 5, pp. 2737–2764, 2010.
- [26] D. C. Sorensen and M. Embree, “A DEIM induced CUR factorization,” *SIAM J. Sci. Comput.*, vol. 38, no. 3, pp. A1454–A1482, 2016.
- [27] P. Drineas, M. W. Mahoney, and S. Muthukrishnan, “Relative-error CUR matrix decompositions,” *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 2, pp. 844–881, 2008.
- [28] K. Hamm and L. Huang, “Stability of sampling for CUR decompositions,” *Found. Data Sci.*, vol. 2, no. 2, pp. 83–99, 2020.
- [29] H. Cai, Z. Chao, L. Huang, and D. Needell, “Fast robust tensor principal component analysis via fiber CUR decomposition,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2021, pp. 189–197.
- [30] ———, “Robust tensor CUR decompositions: Rapid low-tucker-rank tensor recovery with sparse corruption,” *SIAM J. Imaging Sci.*, 2024.
- [31] H. Cai, K. Hamm, L. Huang, J. Li, and T. Wang, “Rapid robust principal component analysis: CUR accelerated inexact low rank estimation,” *IEEE Signal Process. Lett.*, vol. 28, pp. 116–120, 2020.