COLUMN-BASED MATRIX APPROXIMATION WITH QUASI-POLYNOMIAL STRUCTURE

Jeongmin Chae[1]*, Praneeth Narayanamurthy*[1]*, Selin Bac*[2]*, Shaama Mallikarjun Sharada*[2] *and Urbashi Mitra*[1]

[1] Department of Electrical and Computer Engineering, University of Southern California
[2] Department of Chemical Engineering, University of Southern California

ABSTRACT

A novel matrix completion problem is considered herein: observations based on fully sampled columns and quasi-polynomial side information is exploited. The framework is motivated by quantum chemistry problems wherein full matrix computation is expensive, but partial computations only lead to column information. The proposed algorithm successfully estimates the row-space of a true matrix given *a priori* knowledge of the true matrix. A theoretical error bound is provided, which captures the possible inaccuracies of the side information. This work designs the first provable matrix approximation algorithm using just column samples. The proposed algorithm is validated via simulations which enable the characterization of the amount of information provided by the quasi-polynomial side information.

***Index Terms*—** Matrix approximation; high-rank matrix completion; quasi-polynomial; row space information.

1. INTRODUCTION

Low rank matrix approximation has been successfully applied to a variety of applications including recommendation systems, quantum state tomography, matrix regression and gene-expression [1–5]. Classical low-rank matrix approximation has assumed access to randomly collected entries of the true matrix. Recent work in computer vision, bioinformatics and quantum computing suggest that alternative sampling strategies may be more relevant [6–11]. For example, samples in [10] are concentrated around the main diagonal of a matrix and [11] allows for the sampling of complete columns and rows as well as some additional random samples. To accommodate these sampling approaches, new matrix reconstruction methods have been developed.

The current work considers column samples coupled with side information on the rowspace of the true matrix versus the typical random sampling. The sampling strategy is motivated by problems in quantum chemistry [12, 13], where full matrix computation is expensive and only full columns can be computed. While [11] shares some features of our proposed framework (column samples), it also necessitates fully sampled rows as well as additional random samples – thus, [11] cannot be applied to our quantum chemistry problem. To the best of our knowledge, matrix approximation with column samples and rowspace information as captured by a quasi-polynomial structure has not been previously investigated.

Contributions. We propose a new framework for matrix approximation given only a few columns and quasi-polynomial side infromation. The proposed algorithm, dubbed quasi-polynomial matrix approximation (QPMA), successfully integrates a known quasi-polynomial structure of a true matrix. A theoretical additive error bound is provided and shown to be comparable to that achieved by conventional matrix approximation methods. Simulation studies validate that the estimated rowspace information provided by QPMA provides useful side information and enables the characterization of how much information is provided via the quasi-polynomial structure.

Motivating application. For further context, we provide additional details on the $n \times m$ dimensional ground truth matrix $\mathbf{M}$ motivated by applications in quantum chemistry [12, 13]. Each column of the ground truth matrix $\mathbf{M}_{[:,i]}, i \in [m]$ consists of n eigenvalues of a Hessian calculated at one point on the reaction coordinate, $\mathbf{s}$. These eigenvalues are necessary for calculating chemical reaction rate coefficients with sophisticated rate theories. However, obtaining the Hessian matrix at each point is computationally expensive. In our prior work [13, 14], we showed that eigenvalue information in $\mathbf{M}$ has a quasi-polynomial expression along the rows. We further showed that low rank matrix completion can be employed to approximate $\mathbf{M}$. Our goal herein is to approximate the matrix $\mathbf{M}$ given only a few columns (i.e., eigenvalues calculated at a few points along the reaction coordinate) with the quasi-polynomial information and reduce the overall computational cost.

2. PROBLEM FORMULATION AND ALGORITHM

2.1. Problem setting

Let $\mathbf{M} \in \mathbb{R}^{n \times m}$ be the true matrix of rank k. Herein, we consider the problem of obtaining a rank-r approximation of $\mathbf{M}$ from d randomly sampled columns, where $r \leq d \leq k$, motivated by the the quantum chemistry application, wherein the actual matrices are high rank. Motivated by the chemical reaction rate processes in [12], we model quasi-polynomial structure, of degree l, of the true matrix $\mathbf{M}$ as follows,

$$\mathbf{M} = \mathbf{Q}\mathbf{S} + \mathbf{E}, \tag{1}$$

where $\mathbf{Q} \in \mathbb{R}^{n \times l}$ is an unknown polynomial coefficient matrix; $\mathbf{S} \in \mathbb{R}^{l \times m}$ encodes the known polynomial information

matrix; and $\mathbf{E}$ is the perturbation/noise matrix. Specifically, given a vector of reaction coordinate values, $\mathbf{s} = [s_1, \ldots, s_m]$ and the polynomial degree l, the polynomial side information is assumed to be captured by each i-th row of $\mathbf{S}$ possessing $i-1$ degree of polynomial with respect to $\mathbf{s}$, where $i \in [1, l+1]$. For example, i-th row of $\mathbf{S}$ has the form of $\mathbf{S}_{\{i,:\}} = [s_1^{i-1}, \ldots, s_m^{i-1}]$. Notice from (1) that the true matrix $\mathbf{M}$ is not purely polynomial.

This work considers the setting where a few columns of $\mathbf{M}$ are sampled, which we formally define next. Let $\mathcal{C} = \{c_1, \ldots, c_d\} \subset [m]$ denote the set of sampled column indices . We assume that each column is sampled independently from all other columns. Then the column sampling operator $\mathbf{\Psi} \in \{0,1\}^{m \times d}$ is given by $\mathbf{\Psi} \stackrel{\text{def}}{=} \mathbf{I}_{\mathcal{C}}$, where $\mathbf{I}$ is the identity matrix of dimension m and $\mathbf{I}_{\mathcal{C}}$ denotes the sub-matrix of $\mathbf{I}$ formed by columns indexed in the set $\mathcal{C}$. Thus, the observed matrix, $\mathbf{A} \in \mathbb{R}^{n \times d}$ can be equivalently expressed as $\mathbf{A} = \mathbf{M}\mathbf{\Psi}$.

Lastly, we define the SVD of $\mathbf{M}$ and $\mathbf{QS}$ as

$$\mathbf{M} \stackrel{\text{SVD}}{=} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T} = \underbrace{\mathbf{U}_M\mathbf{\Sigma}_M\mathbf{V}_M^\mathsf{T}}_{\text{rank-}r\text{-approximation}} + \underbrace{\mathbf{U}_{M,\perp}\mathbf{\Sigma}_{M,\perp}\mathbf{V}_{M,\perp}^\mathsf{T}}_{\text{remainder}}.$$

$$\text{and} \quad \mathbf{QS} \stackrel{\text{SVD}}{=} \underbrace{\mathbf{U}_{QS}\mathbf{\Sigma}_{QS}\mathbf{V}_{QS}^\mathsf{T}}_{\text{rank-}r\text{-approximation}} + \underbrace{\mathbf{U}_{QS,\perp}\mathbf{\Sigma}_{QS,\perp}\mathbf{V}_{QS,\perp}^\mathsf{T}}_{\text{remainder}}.$$

2.2. Quasi-Polynomial Matrix Approximation Algorithm

We next present the optimization strategy for Quasi Polynomial Matrix Approximation (QPMA). A natural optimization problem that takes into account the structural information provided by $\mathbf{QS}$ as well as the desired rank r approximation of $\mathbf{M}$ is given by

$$\min_{\mathbf{Q},\mathbf{X}} \ \|\mathbf{A} - \mathbf{X}\mathbf{\Psi}\|_F^2 \ \text{ subject to} \tag{2}$$
$$\|\mathbf{A} - \mathbf{QS}\mathbf{\Psi}\|_F^2 \leq \|\mathbf{E}\|_F^2 \|\mathbf{\Psi}\|_F^2 \text{ and } \text{rank}(\mathbf{X}) \leq r.$$

Notice that while sufficient information about the column-space of $\mathbf{M}$ is provided through $\mathbf{A}$, it is almost impossible to estimate the row-space without the prior quasi-polynomial structure, represented by $\mathbf{QS}$.

The above optimization problem is NP-hard, in general, due to the rank constraint [15]. Although there is a plethora of literature that attempts to circumvent the computational hardness – either through convex relaxation [16–18] or through non-convex approaches [19], these approaches are still either computationally or statistically sub-optimal respectively. Furthermore, it is not clear if these approaches can be easily modified to leverage the quasi-polynomial side information. Finally, we note that the above optimization problem does not explicitly leverage the row- and column-space information of $\mathbf{M}$. We observe that if the row- and column-space information of $\mathbf{M}$, i.e., $\mathbf{U}_M$ and $\mathbf{V}_M$ were known, a natural factorized approach [20] that relaxes (2) is given by

$$\min_{\mathbf{Z},\mathbf{Q}} \ \left\|\mathbf{A} - \mathbf{U}_M\mathbf{Z}\mathbf{V}_M^\mathsf{T}\mathbf{\Psi}\right\|_F^2 \tag{3}$$
$$\text{subject to} \ \ \|\mathbf{A} - \mathbf{QS}\mathbf{\Psi}\|_F^2 \leq \|\mathbf{E}\|_F^2 \|\mathbf{\Psi}\|_F^2.$$

However, since we do not know $\mathbf{U}_M$ and $\mathbf{V}_M$, they need to be estimated. QPMA is comprised of three stages: (i) estimating the column space of $\mathbf{M}$; (ii) followed by estimating the unknown polynomial coefficient matrix, $\mathbf{Q}$, and subsequently estimating the row-space of $\mathbf{M}$ by leveraging the quasi polynomial structure; and (iii) finally preforming matrix approximation with the constrained row- and column-space approximations obtained previously. The complete algorithm is summarized in Algorithm 1.

Column space estimation of $\mathbf{M}$. Notice that we are only given a subset of the columns of $\mathbf{M}$. We argue that as long as d is "large enough", and the columns are sampled independently, the top-r left singular vectors of the observed matrix $\mathbf{A}$ provide a good estimate of the (r-dimensional) column-space of $\mathbf{M}$. The estimate is given by the rank-r SVD as $\mathbf{A} \stackrel{\text{r-SVD}}{=} \mathbf{U}_A\mathbf{\Sigma}_A\mathbf{V}_A^\mathsf{T}$.

Row space estimation of $\mathbf{M}$. Next, we leverage the side information to estimate the row-space of $\mathbf{M}$. Recall $\mathbf{A}$ does not provide sufficient information for row-space of $\mathbf{M}$. Hence, we estimate the unknown polynomial coefficient matrix, $\mathbf{Q}$ as follows

$$\hat{\mathbf{Q}} = \underset{\bar{\mathbf{Q}}}{\text{argmin}} \ \left\|\mathbf{A} - \bar{\mathbf{Q}}\mathbf{S}\mathbf{\Psi}\right\|_F^2. \tag{4}$$

This is a standard regression problem, which can be efficiently solved by gradient descent [21]. With $\hat{\mathbf{Q}}$ obtained

Algorithm 1 Proposed QPMA

Input: $\mathbf{A} \in \mathbb{R}^{n \times d}$, $\mathbf{S} \in \mathbb{R}^{l \times m}$
Parameters: A target rank r, Degree of polynomial l, Step size η, # of iteration T_1 and T_2 for (4) and (5)
Initialization: Randomly generate $\hat{\mathbf{Q}}_1$ from $\mathcal{N}(0,1)$
Algorithm:
Column-space estimation
  Do rank-r SVD of $\mathbf{A}$ as $\mathbf{A} = \mathbf{U}_A\mathbf{\Lambda}_A\mathbf{V}_A^\mathsf{T}$
Row-space estimation
  For $t \in [T_1]$, do
    $\hat{\mathbf{Q}}_{t+1} = \hat{\mathbf{Q}}_t - \eta\left(\mathbf{A} - \hat{\mathbf{Q}}_t\mathbf{S}\mathbf{\Psi}\right)(\mathbf{S}\mathbf{\Psi})^\mathsf{T}$
  With $\hat{\mathbf{Q}} \equiv \hat{\mathbf{Q}}_T$, do rank-$r$ SVD of $\hat{\mathbf{Q}}\mathbf{S}$ as $\hat{\mathbf{Q}}\mathbf{S} = \hat{\mathbf{U}}_{QS}\hat{\mathbf{\Lambda}}_{QS}\hat{\mathbf{V}}_{QS}^\mathsf{T}$
Matrix approximation
  Using $\mathbf{U}_A$ and $\hat{\mathbf{V}}_{QS}$, for $t \in [T]$, do
    $\hat{\mathbf{Z}}_{t+1} = \hat{\mathbf{Z}}_t - \eta\mathbf{U}_A^\mathsf{T}\left(\mathbf{A} - \mathbf{U}_A\hat{\mathbf{Z}}_t\hat{\mathbf{V}}_{QS}^\mathsf{T}\mathbf{\Psi}\right)\left(\hat{\mathbf{V}}_{QS}^\mathsf{T}\mathbf{\Psi}\right)^\mathsf{T}$
  Obtain $\hat{\mathbf{Z}} = \hat{\mathbf{Z}}_T$ and complete $\hat{\mathbf{M}} = \mathbf{U}_A\hat{\mathbf{Z}}\hat{\mathbf{V}}_{QS}^\mathsf{T}$
Output: $\hat{\mathbf{M}} = \mathbf{U}_A\hat{\mathbf{Z}}\hat{\mathbf{V}}_{QS}^\mathsf{T}$.

by solving (4), we can estimate the row-space of $\mathbf{M}$, $\hat{\mathbf{V}}_{QS}$, through a rank-r SVD of $\hat{\mathbf{QS}}$ as $\hat{\mathbf{QS}} = \hat{\mathbf{U}}_{QS}\hat{\mathbf{\Sigma}}_{QS}\hat{\mathbf{V}}_{QS}^{\mathsf{T}}$.

Rank-r matrix approximation. Finally, with $\mathbf{U}_A$ and $\hat{\mathbf{V}}_{QS}$ obtained from previous steps, we can perform the low-rank approximation as follows,

$$\hat{\mathbf{Z}} = \arg\min_{\bar{\mathbf{Z}}} \left\| \mathbf{A} - \mathbf{U}_A \bar{\mathbf{Z}} \hat{\mathbf{V}}_{QS}^{\mathsf{T}} \mathbf{\Psi} \right\|_F^2. \tag{5}$$

This final optimization is also a regression problem, which can also be solved by gradient descent. This concludes the algorithm.

3. MAIN RESULT AND PROOF SKETCH

To provide the spectral error bound for QPMA, we require the following preliminaries:

Definition 1 (Incoherence). *[15] Let $\mathbf{X}$ be a $n \times m$ matrix of rank r and $\mathbf{X} \overset{r\text{-}SVD}{=} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{T}}$. Let $\mathbf{u}_i$ be the i-th row of $\mathbf{U}$ and $\mathbf{v}_j$ be the j-th row of $\mathbf{V}$. Then, the incoherence of $\mathbf{X}$ is given by $\bar{\mu}(\mathbf{X}) = \max\left(\max_{i\in[n]} \frac{n}{r}\|\mathbf{u}_i\|^2, \max_{j\in[m]} \frac{m}{r}\|\mathbf{v}_j\|^2\right)$.*

For clarity, we define $\mu \doteq \bar{\mu}(\mathbf{M})$ and $\hat{\mu} \doteq \bar{\mu}(\hat{\mathbf{M}})$ for the ground truth and its estimate, respectively. We denote $\sigma_i(\mathbf{QS})$ and $\sigma_j(\mathbf{M})$, $1 \le i, j \le m$, the i-th and j-th largest singular values of $\mathbf{QS}$ and $\mathbf{M}$ respectively which enable the following assumption.

Assumption 1 : There exists a $\delta > 0$ such that $\delta = \min\{\min_{i,j} |\sigma_i(\mathbf{QS}) - \sigma_j(\mathbf{M})|, \min_i \sigma_i(\mathbf{QS})\}$ for $1 \le i \le r, r \le j \le m$.

Theorem 1. *Assume there exists a $\delta > 0$ that satisfies the Assumption 1 and that d columns are sampled uniformly at random from the underlying ground truth, $\mathbf{M}$. Then, if $d \ge c_1 max\left\{\mu r \ln r, \frac{\hat{\mu}^2 r^2 \ln r}{n}\right\}$, with probability at least $1 - c_2 r^{-10}$ we have*

$$\frac{\|\mathbf{M} - \hat{\mathbf{M}}\|_2^2}{\|\mathbf{M}\|_2^2} \le \frac{4m}{d}\left[\frac{27(n+2m)}{d}\frac{\sigma_{r+1}^2(\mathbf{M})}{\sigma_1^2(\mathbf{M})} + \frac{2\sqrt{2}\|\mathbf{E}\|_F}{\delta} + \frac{72\|\mathbf{E}\|_F^2}{\delta^2}\right]$$

where c_1 and c_2 are positive constants.

Proof Sketch: The proofs and additional lemmas are provided in the extended version of this work [22]. The proof follows from applying large-deviation style results from random matrix theory [23] to ensure that the loss-function in (5) is well-behaved as long as we sample a sufficient number of columns, followed by a careful application of Wedin's theorem [24]. ⊠

Discussion. The first term in Theorem 1 represents the unrecoverable error due to the low rank approximation of a higher rank matrix. As $\mathbf{M}$ is only quasi-polynomial ($\mathbf{E} \neq \mathbf{0}$), the remaining terms capture the incomplete nature of the side information. Since $\text{rank}(\mathbf{M}) = k \gg r$ and $\text{rank}(\mathbf{QS}) = l \ge r$, we have $\delta = \sigma_r(\mathbf{M} - \mathbf{E}) - \sigma_{r+1}(\mathbf{M})$. As δ measures effective eigengap of $\mathbf{M}$, we interpret $\frac{\delta}{\|\mathbf{E}\|_F}$ as a signal-to-noise ratio (SNR). Thus, we can describe our approximation as scaling with $C\max(1/SNR, (1/SNR)^2)$. We also observe that QPMA suffers a multiplicative factor of $O(m/d)$ coupled with the best rank-r approximation error, that is $\|\mathbf{M} - \mathbf{M}_r\| = \sigma_{r+1}(\mathbf{M})$. We note that CUR+ [11] induces a multiplicative factor of $O(m\sqrt{m/d})$, and thus our bound is competitive with respect to scaling. Note that the error bound is quantified by the informativeness of side information, i.e., the bound gets tighter when $\|\mathbf{E}\|_F^2$ is small. We provide a more detailed discussion and other lemmas in the longer version of this paper [22].

4. SIMULATIONS

All numerical results are averaged over 100 independent runs. We generate the entries of the polynomial coefficients matrix $\mathbf{Q} \in \mathbb{R}^{n\times l}$ independently from $\mathcal{N}(0,1)$. The polynomial basis matrix $\mathbf{S} \in \mathbb{R}^{l\times m}$ is constructed using arbitrarily sampled reaction coordinate values $\mathbf{s}$. As in [12], the values of $\mathbf{s}$ are arranged in ascending order after uniformly sampling $\mathbf{s} = [1 + 0.01 * [m]]$. To control the rank of the underlying ground truth matrix, we generate the perturbation matrix $\mathbf{E}$ as $\mathbf{U}_{QS}\mathbf{R}_1\mathbf{V}_{QS}^{\mathsf{T}} + \mathbf{U}_{QS,\perp,[:,1:k-l]}\mathbf{R}_2\mathbf{V}_{QS,\perp,[:,1:k-l]}^{\mathsf{T}}$, where $\mathbf{QS} \overset{SVD}{=} \mathbf{U}_{QS}\mathbf{\Sigma}_{QS}\mathbf{V}_{QS}^{\mathsf{T}} + \mathbf{U}_{QS,\perp}\mathbf{\Sigma}_{QS,\perp}\mathbf{V}_{QS,\perp}^{\mathsf{T}}$. The entries of $\mathbf{R}_1 \in \mathbb{R}^{l\times l}$ and $\mathbf{R}2 \in \mathbb{R}^{k-l\times k-l}$ are generated i.i.d. from $\mathcal{N}(0,\sigma)$ with $\sigma = 0.0001$. With this, we obtain the rank-k ground truth matrix, $\mathbf{M} = \mathbf{QS} + \mathbf{E}$. We use the normalized mean squared error (NMSE), defined as $\text{NMSE}(\hat{\mathbf{M}}, \mathbf{M}) = \frac{\|\mathbf{M}-\hat{\mathbf{M}}\|_F}{\|\mathbf{M}\|_F}$ to measure the performance.

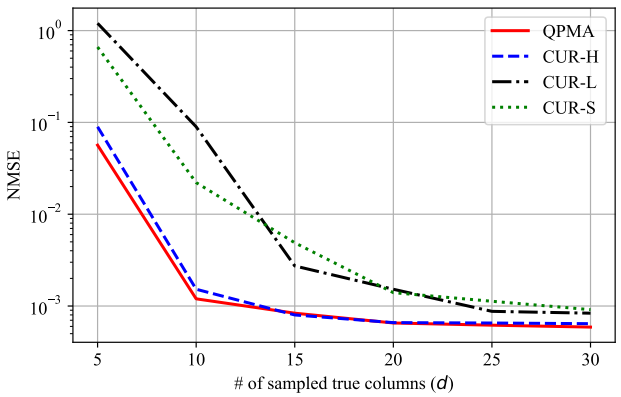

Fig. 1. The comparison of NMSE between QPMA and various types of CUR+ versus # of true sampled columns (d) when $k = 30$ and $l = 5$.

We compare the performance of QPMA with CUR+ [11] although we emphasize that CUR+ requires access to a sub-

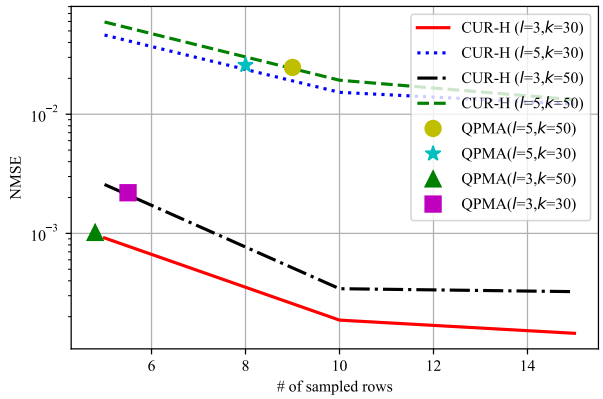

Fig. 2. The performance of QPMA corresponding to the same NMSE of CUR-H with fixed $d = 5$ and varying rows.

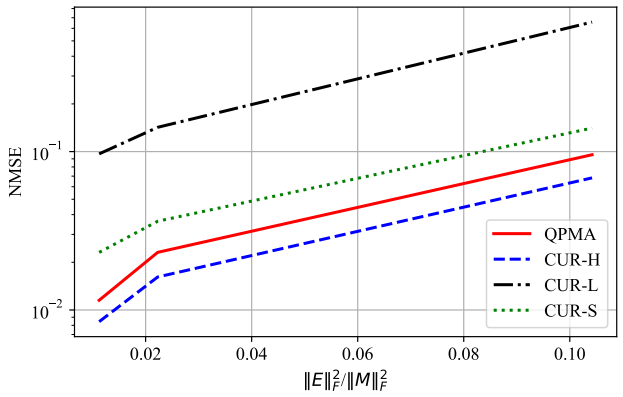

Fig. 3. The sensitivity to noise for QPMA and CUR+ algorithms. The parameters are set by $d = 5$, $l = 5$ and $k = 100$.

set of rows, columns and a few additional entries to work. To provide a fair comparison applicable to our setting, we implement three types of CUR+, named CUR-L (CUR with low # samples), CUR-S (CUR with same # samples) and CUR-H (CUR with high # samples, which is the default setting). QPMA observes d columns, CUR-L observes $d/2$ rows and columns, CUR-S observes $d/2$ rows and columns and an additional $d^2/4$ entries, and CUR-H observes d columns and rows in total. Thus, compared to QPMA, CUR-L observes $d^2/4$ fewer *distinct* entries, CUR-S observes the same number of entries, and CUR-H observes $nd - d^2/4$ more distinct entries.

We first investigate the effect of increasing the number of observed columns; the true rank is $k = 30$, and $\sigma = 0.0001$. The results are provided in Fig. 1. As expected, the NMSE for all algorithms reduces as we increase the number of observed samples. Furthermore, even though CUR-H observes many more entries than QPMA, the performance is comparable, and QPMA significantly outperforms CUR-S and CUR-L. Additionally, in the low-sample regime, QPMA has the lowest reconstruction error; thus the side information is being well leveraged by QPMA. We next attempt to answer the following question: *how* much (row-space) information is being captured by the quasi-polynomial side information[1]. To this end, for both QPMA and CUR-H, we fix the number of observed columns to $d = 5$. We vary the observed number of *rows* for CUR-H that attains the same numerical error as QPMA. The results are provided in Fig. 2. Notice that when the the degree of the quasi-polynomial side information is $l = 3$, for both values of rank, k, CUR-H requires access to at least $5 - 6$ rows to match the performance of QPMA. Additionally, when $l = 5$, CUR-H requires at least $8 - 9$ rows to match the performance of QPMA. This trend confirms our intuition that as the underlying degree (and hence the "goodness" of the side information) increases, the performance of existing vanilla algorithms degrades. Finally, we illustrate the sensitivity to noise, $\mathbf{E}$ for all algorithms in Fig. 3. Notice that the performance of all algorithms gracefully degrade with increasing noise.

5. CONCLUSIONS

In this work, we provided a novel framework for matrix approximation under the assumption that one can only collect fully sampled columns of the underlying matrix. Furthermore, one has access to side information with regards to the rowspace that is quasi-polynomial. This framework is motivated by a problem in quantum chemistry; our goal is to significantly reduce computational complexity of chemical reaction rate computation. We have further proposed the quasi-polynomial matrix approximation (QPMA) algorithm which successfully integrates a known quasi-polynomial structure of a true matrix with a few fully sampled columns. A bound on the approximation error is derived which clearly shows the trade off between the side information and how close the true matrix is to being fully polynomial. We showed that QPMA showed a competitive multiplicative factor of $O(m/d)$ coupled with the best low rank approximation error. QPMA is compared to modern row and column sampled approaches [11]; numerical simulations show that QPMA can capture more than about 2 times information of true rows compared to state-of-art method.

Acknowledgements

This work is funded in part by one or more of the following grants: NSF CCF-1817200, ARO W911NF1910269, DOE DE-SC0021417, Swedish Research Council 2018-04359, NSF CCF-2008927, NSF CCF-2200221, ONR 503400-78050, ONR N00014-15-1-2550 and USC+Amazon Center on Secure and Trusted Machine Learning.

[1]To keep the discussion simple, we only compare with CUR-H even though QPMA is technically at a disadvantage in this comparison.

6. REFERENCES

[1] Ivan Markovsky, "Structured low-rank approximation and its applications," *Automatica*, vol. 44, no. 4, pp. 891–909, 2008.

[2] Nathan Srebro and Tommi Jaakkola, "Weighted low-rank approximations," in *Proceedings of the 20th international conference on machine learning (ICML-03)*, 2003, pp. 720–727.

[3] Vladimir Koltchinskii, "Low rank matrix recovery: nuclear norm penalization," in *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*, pp. 191–234. Springer, 2011.

[4] Michael W Mahoney and Petros Drineas, "Cur matrix decompositions for improved data analysis," *Proceedings of the National Academy of Sciences*, vol. 106, no. 3, pp. 697–702, 2009.

[5] Christos Boutsidis, Michael W Mahoney, and Petros Drineas, "An improved approximation algorithm for the column subset selection problem," in *Proceedings of the twentieth annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 2009, pp. 968–977.

[6] Tianxi Cai, T Tony Cai, and Anru Zhang, "Structured matrix completion with applications to genomic data integration," *Journal of the American Statistical Association*, vol. 111, no. 514, pp. 621–633, 2016.

[7] Simon Foucart, Deanna Needell, Reese Pathak, Yaniv Plan, and Mary Wootters, "Weighted matrix completion from non-random, non-uniform sampling patterns," *IEEE Transactions on Information Theory*, vol. 67, no. 2, pp. 1264–1290, 2020.

[8] Suriya Gunasekar, Arindam Banerjee, and Joydeep Ghosh, "Unified view of matrix completion under general structural constraints," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[9] Yue Hu, Xiaohan Liu, and Mathews Jacob, "A generalized structured low-rank matrix completion algorithm for mr image recovery," *IEEE transactions on medical imaging*, vol. 38, no. 8, pp. 1841–1851, 2018.

[10] Guangcan Liu, Qingshan Liu, and Xiaotong Yuan, "A new theory for matrix completion," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[11] Miao Xu, Rong Jin, and Zhi-Hua Zhou, "Cur algorithm for partially observed matrices," in *International Conference on Machine Learning*. PMLR, 2015, pp. 1412–1421.

[12] Stephen Jon Quiton, Jeongmin Chae, Selin Bac, Kareesa Kron, Urbashi Mitra, and Shaama Mallikarjun Sharada, "PVMC," Mar. 2022, (accessed 2022-03-07).

[13] Selin Bac, Stephen Jon Du Quiton, Kareesa J Kron, Jeongmin Chae, Urbashi Mitra, and Shaama Mallikarjun Sharada, "A matrix completion algorithm for efficient calculation of quantum and variational effects in chemical reactions," *The Journal of Chemical Physics*, Apr. 2022, Publisher: American Institute of Physics.

[14] Stephen Jon Quiton, Urbashi Mitra, and Shaama Mallikarjun Sharada, "A matrix completion algorithm to recover modes orthogonal to the minimum energy path in chemical reactions," *The Journal of Chemical Physics*, vol. 153, no. 5, pp. 054122, Aug. 2020.

[15] Emmanuel J. Candès and Benjamin Recht, "Exact Matrix Completion via Convex Optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, Dec. 2009.

[16] E.J. Recht Candès, "Exact matrix completion via convex optimization," in *Foundations of Computational Mathematics*, 2009, vol. 9, p. 717–772.

[17] Emmanuel J Candes and Yaniv Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.

[18] Emmanuel J Candès and Terence Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.

[19] Ruoyu Sun and Zhi-Quan Luo, "Guaranteed matrix completion via non-convex factorization," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6535–6579, 2016.

[20] Samuel Burer and Renato DC Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Mathematical Programming*, vol. 95, no. 2, pp. 329–357, 2003.

[21] Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[22] Jeongmin Chae, Praneeth Narayanamurthy, Selin Bac, Shaama Mallikarjun Sharada, and Urbashi Mitra, "Matrix approximation with side information: When column sampling is enough," *arXiv preprint arXiv:2212.05469*, 2022.

[23] M.L. Mehta, "Random matrices," 2004.

[24] PÅ Wedin, "Perturbation bounds in connection with singular value decomposition," *BIT*, vol. 12, pp. 99–111, 1972.