Dihedral multi-reference alignment

Tamir Bendory, Dan Edidin, William Leeb, and Nir Sharon

July 14, 2021

Abstract

We study the dihedral multi-reference alignment problem of estimating the orbit of a signal from multiple noisy observations of the signal, translated by random elements of the dihedral group. We show that if the group elements are drawn from a generic distribution, the orbit of a generic signal is uniquely determined from the second moment of the observations. This implies that the optimal estimation rate in the high noise regime is proportional to the square of the variance of the noise. This is the first result of this type for multi-reference alignment over a non-abelian group with a non-uniform distribution of group elements. Based on tools from invariant theory and algebraic geometry, we also delineate conditions for unique orbit recovery for multi-reference alignment models over finite groups (namely, when the dihedral group is replaced by a general finite group) when the group elements are drawn from a generic distribution. Finally, we design and study numerically three computational frameworks for estimating the signal based on group synchronization, expectation-maximization, and the method of moments.

1 Introduction

We study the dihedral multi-reference alignment (MRA) model

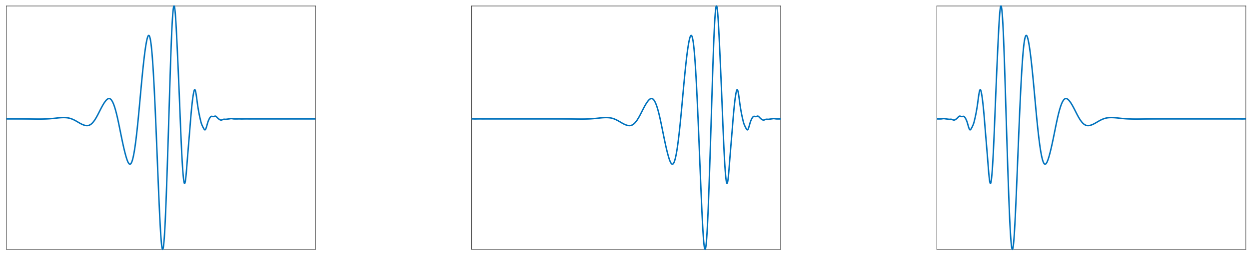
$$y = g \cdot x + \varepsilon, \quad g \sim \rho, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2 I), \tag{1.1}$$

where

- $x \in \mathbb{R}^L$ is a fixed (deterministic) signal to be estimated;
- ρ is an unknown distribution defined over the simplex Δ_{2L};
- g is a random element of the dihedral group D_{2L}, drawn i.i.d. from ρ, and acting on the signal by circular translation and reflection (see Figure 1);
- ε is a normal isotropic i.i.d. noise with zero mean and variance σ^2.

We wish to estimate the signal x from n realizations (observations) of y,

$$y_i = g_i \cdot x + \varepsilon_i, \quad i = 1, \dots, n, \tag{1.2}$$



(a) Signal (b) Shifted signal (c) Shifted and reflected signal

Figure 1: An example of the action of the dihedral group. The MRA problem (1.1) entails estimating a signal, up to a global circular shift and reflection, from multiple noisy copies of the signal acted upon by random elements of the dihedral group.

while the corresponding group elements $g_1, \ldots, g_n$ are unknown. We note, however, that the signal can be identified only up to the action of an arbitrary element of the dihedral group. Therefore, unless a prior information on the signal is available, the goal is estimating the orbit of signals $\{g \cdot x | g \in D_{2L}\}$. This type of problems is often dubbed an orbit recovery problem.

The model (1.1) is an instance of the more general MRA problem that was studied thoroughly in recent years [9, 14, 8, 3, 2, 18, 4, 27, 24, 11, 10, 31, 1, 21, 6, **?**, **?**, **?**, **?**, **?**]. In its generalized version, the MRA model is formulated as (1.1), but the signal x may lie in an arbitrary vector space (not necessarily $\mathbb{R}^L$), the dihedral group D_{2L} is replaced by an arbitrary group G, and $g \sim \rho$ is a distribution over G (in some cases, an additional fixed linear operator acting on the signal is also considered, e.g., [10, 8, 16, 12]). The goal is estimating the orbit of x, under the action of the group G.

Most of the previous studies on MRA considered a uniform (or Haar) distribution ρ of the group elements. In particular, it was shown that in many cases, such as $x \in \mathbb{R}^L$ and a uniform distribution over $\mathbb{Z}_L$, the third moment suffices to recover a generic signal uniquely [8, 23, 27, 14]. This result is of particular interest since it determines the sample complexity of MRA in the low signal-to-noise (SNR) regime. Specifically, in the low SNR regime $\sigma \to \infty$ (with a fixed dimension L), a necessary condition for signal identification is $n/\sigma^{2d} \to \infty$, where d is the lowest order moment that identifies the orbit of signals uniquely [8, 4, 11, 27] (see [31] for sample complexity analysis in high dimensions). For example, for $x \in \mathbb{R}^L$ with a uniform distribution over $\mathbb{Z}_L$, $d = 3$ and thus $n/\sigma^6 \to \infty$ is a necessary condition for accurate estimation of generic signal.

The effect of non-uniform distribution on the sample complexity was first studied in [2] for the abelian group $\mathbb{Z}_L$ and $x \in \mathbb{R}^L$. It was shown that in this case the second moment suffices to identify the orbit of generic signals uniquely for almost any non-uniform distribution (rather than the third moment if the distribution is uniform). In this work, we extend [2] for the non-abelian group D_{2L} and show that for a generic distribution and signal, the second moment identifies the orbit of solutions. This implies that a necessary condition for accurate orbit recovery under the model (1.1) for $\sigma \to \infty$ and fixed L is $n/\sigma^4 \to \infty$. This is the first result of this type for multi-reference alignment over a non-abelian group with a non-uniform distribution of group elements. The main theoretical results are summarized as follows.

Theorem 1.1 (informal statement of the main theorem)**.** *Consider the dihedral MRA prob-*

lem (1.1) *with a generic probability distribution ρ. Then, the first and second order moments of y are sufficient to recover almost all orbits.*

Corollary 1.2 (sample complexity)**.** *Consider the dihedral MRA problem* (1.1) *in the low SNR regime $\sigma \to \infty$. For a generic probability distribution and a generic signal, $n/\sigma^4 \to \infty$ is a necessary condition for accurate orbit estimation.*

Theorem 1.1 is formulated in technical terms in Theorem 2.2, which is proved in Section 2.2. The proof is based on algebraic geometry tools. Section 2.2 also discusses the precise meaning of the notion of generic signal and distribution. In Section 2.3, we use invariant theory to delineate general conditions for orbit recovery from the second moment in general MRA models over finite groups.

The MRA model is mainly motivated by the molecular structure reconstruction problem in single-particle cryo-electron microscopy (cryo-EM) [12]. The aim of a cryo-EM experiment is constituting a 3-D molecular structure from multiple observations. In each observation, the 3-D structure is acted upon by a random element of the non-abelian group of 3-D rotation SO(3). In addition, the distribution over SO(3) is usually non-uniform and unknown [38, 25, 7, 34]. Therefore, this paper is an important step towards understanding the statistical properties and sample complexity of the cryo-EM problem

Section 3 introduces three statistical estimation frameworks to recover the orbit of x. The first framework is based on estimating the missing group elements using the method of group synchronization [36, 10]. Once the group elements were accurately estimated, estimating the signal can be obtained by aligning the observations and averaging out the noise. However, reliable estimation of group elements is possible only if the noise level is low enough. To estimate the signal in high noise levels, we also suggest maximizing the marginalized maximum likelihood using expectation-maximization (EM). EM provides accurate estimations in a wide range of SNR regimes, however, its computational burden rapidly increases with the number of observations n and the noise level. As a third method, we propose an estimator based on the method of moments, which works quite well in all SNRs and its computational burden is roughly constant with the noise level and moderately increases with n. According to Theorem 1.1, we only use the first and second moments for the estimation.

2 Theory

2.1 The dihedral group

The dihedral group D_{2L} is a group of order $2L$, which is usually defined as the group of symmetries of a regular L-gon in $\mathbb{R}^2$. It is generated by a rotation r of order L corresponding to rotation by an angle $2\pi/L$ and a reflection s of order 2. Since rotation does not commute with reflection, the group D_{2L} is not abelian, but the relation $rs = sr^{-1}$ holds instead. Since r has order L, $r^{-1} = r^{L-1}$. The elements of D_{2L} can be enumerated as

$$\{1, r, \ldots, r^{L-1}, s, rs, \ldots, r^{L-1}s\},$$

where 1 is the identity element. Note that the subset $\{1, r, \ldots, r^{L-1}\}$ is a normal subgroup[1] isomorphic to the cyclic group $\mathbb{Z}_L$. MRA over the group $\mathbb{Z}_L$ was studied thoroughly, see for

[1] A subgroup $H < G$ is normal if it is invariant under conjugation by elements of G.

example [9, 14, 2].

There are two natural actions of the dihedral group ,D_{2L}, on $\mathbb{R}^L$ one in the time (or spatial) domain and one in the Fourier (frequency) domain. Explicitly, in the time domain, the action of the dihedral group on a signal $x \in \mathbb{R}^L$ is given by

$$\begin{aligned} (r \cdot x)[\ell] &= x[(\ell - 1) \bmod L], \\ (s \cdot x)[\ell] &= x[-\ell \bmod L]. \end{aligned} \tag{2.1}$$

Namely, r cyclically shifts a signal by one entry, and s reflects the signal. The action of the dihedral group is illustrated in Figure 1.

If we apply the discrete Fourier transform to $\mathbb{R}^L$, then we can identify $\mathbb{R}^L$ with the real subspace of $\mathbb{C}^L$ consisting of L-tuples $(\hat{x}[0], \dots, \hat{x}[L-1]) \in \mathbb{C}^L$ satisfying the condition $\overline{\hat{x}[\ell]} = \hat{x}[-\ell \bmod L]$, where $\overline{\hat{x}[\ell]}$ is the conjugate of $\hat{x}[\ell]$. In this case, the action of D_{2L} is given by:

$$\begin{aligned} (r \cdot \hat{x})[\ell] &= e^{2\pi \iota r \ell / L} \hat{x}[\ell], \\ (s \cdot \hat{x})[\ell] &= \overline{\hat{x}[\ell]} = \hat{x}[-\ell \bmod L]. \end{aligned} \tag{2.2}$$

2.2 Unique orbit recovery in dihedral MRA

We are now ready to present and prove the main result of this paper. Let $\rho \in \Delta_{2L}$ be a probability distribution on D_{2L}. We denote the probability of r^k by $p[k]$ and the probability of $r^k s$ by $q[k]$. Let p and q represent the vectors $(p[0], \dots, p[L-1])$ and $(q[0], \dots, q[L-1])$, respectively. Let $C_z \in \mathbb{R}^{L \times L}$ be a circulant matrix generated by $z \in \mathbb{R}^L$, and let $D_z \in \mathbb{R}^{L \times L}$ be a diagonal matrix whose entries are z. A direct calculation shows that the first two moments of the observations of (1.1) are given by the following expressions (compare with [2]).

Lemma 2.1. *Consider the dihedral MRA model* (1.1)*. The first moment of* y*,* $M^1 \in \mathbb{R}^L$*, is given by*

$$\mathbb{E}y := M^1(x, \rho) = C_x p + C_{sx} q = C_p x + C_q s x. \tag{2.3}$$

The second moment of y*,* $M^2 \in \mathbb{R}^{L \times L}$*, is given by*

$$\mathbb{E}yy^T := M^2(x, \rho) = C_x D_p C_x^T + C_{sx} D_q C_{sx}^T. \tag{2.4}$$

To present the main result of this paper, it will be convenient to consider the Fourier counterpart of the moments, defined by

$$\begin{aligned} \hat{M}^1 &= \mathbb{E}Fy = FM^1(x, \rho), \\ \hat{M}^2 &= \mathbb{E}Fy(Fy)^* = FM^2(x, \rho)F^*, \end{aligned}$$

where $F \in \mathbb{C}^{L \times L}$ is the discrete Fourier transform (DFT) matrix.

We say that a condition holds for generic signals (or distributions) if the set of signals (distributions) for which the condition does not hold is defined by polynomial conditions. The precise meaning of generic signals, in the context of this work, is discussed at the end of this section.

The main result of this paper is as follows.

Theorem 2.2 (Orbit recovery)**.** *For generic signal x and generic distribution ρ, the D_{2L} orbit of x is uniquely determined by $\hat{M}^1[0] = \hat{x}[0]$ and at most $\sim 2.5L$ entries of the matrix $\hat{M}^2$. More precisely, there exist non-zero polynomials $Q_1, \ldots, Q_r$ such that if $Q_1(x, \rho), \ldots, Q_r(x, \rho)$ are not all zero, then for any (z, ρ') with $M^1(z, \rho') = M^1(x, \rho)$ and $M^2(z, \rho') = M^2(x, \rho)$, z is in the same D_{2L} orbit as x.*

Remark 2.3. *Our method of proof necessarily requires that all of the entries of Fx are non-zero, where F is the discrete Fourier transform matrix; similar assumptions are often stated in the MRA literature, see for example [14, 3, 27, 11]. However, our proof also requires that additional, less explicit, polynomials in the entries of x, ρ be non-vanishing. This is discussed at the end of the proof.*

Proof. Let us define

$$\hat{p} = (\hat{p}[0], \ldots, \hat{p}[L-1]),$$
$$\hat{q} = (\hat{q}[0], \ldots, \hat{q}[L-1]),$$
$$\hat{x} = (\hat{x}[0], \ldots, \hat{x}[L-1]).$$

Note that the second moment in Fourier domain can be written as

$$\hat{M}^2 = \frac{1}{L}\left(D_{\hat{x}} C_{\hat{p}} D_{\overline{\hat{x}}} + D_{\overline{\hat{x}}} C_{\hat{q}} D_{\hat{x}}\right). \tag{2.5}$$

Moreover, since p, q, x are real, we have

$$\hat{x}[L-i] = \overline{x[i]},$$
$$\hat{p}[L-i] = \overline{\hat{p}[i]},$$
$$\hat{q}[L-i] = \overline{\hat{q}[i]}.$$

Define

$$M_{i,j} = \hat{p}[i+j]\hat{x}[i]\hat{x}[j] + \hat{q}[L-i-j]\hat{x}[L-i]\hat{x}[L-j], \tag{2.6}$$

so $M_{i,j}$ is $L\hat{M}^2[i, L-j]$. Our goal is to show that knowledge of $\hat{M}^1[0]$ and $O(L)$ of the entries $M_{i,j}$ determine the orbit of x.

Since ρ is a probability distribution, we note that

$$\hat{p}[0] + \hat{q}[0] = p[0] + \ldots + p[L-1] + q[0] + \ldots + q[L-1] = 1. \tag{2.7}$$

Thus, $M_{i,-i} = |\hat{x}[i]|^2$. It follows that knowledge of $M^2(x, \rho)$ determines the power spectrum of x. Replacing x by the vector whose Fourier transform has entries $\hat{x}[i]/|\hat{x}[i]|$, we may assume that each $\hat{x}[i]$ lies on the unit circle. Since $\hat{x}[0]$ is real, we take $\hat{x}[0] = 1$. With this assumption, the formula for $M_{i,j}$ can be written as

$$M_{i,j} = \hat{p}[i+j]\hat{x}[i]\hat{x}[j] + \hat{q}[L-i-j]/(\hat{x}[i]\hat{x}[j]). \tag{2.8}$$

Given a vector x and distribution ρ, consider the set I of vectors $z \in \mathbb{R}^L$ such that $M^1(z, \rho') = M^1(x, \rho)$ and $M^2(z, \rho') = M^2(x, \rho)$ for some probability distribution ρ' on the dihedral group D_{2L}. We will show that for generic (x, ρ) there are only $2L$ possible z's in this set. Note that the distribution is uniquely determined by the signal z, because the moments are linear functions of the distribution. Since the D_{2L} orbit of x is contained in the set I, we conclude that the orbit of x is determined by the moments of degree one and two.

Determining that the set I consists of at most $2L$ vectors is equivalent to showing that the following system of equations has at most $2L$ solutions:

$$\hat{p}'[i+j]\hat{z}[i]\hat{z}[j] + \hat{q}'[L-i-j]/(\hat{z}[i]\hat{z}[j]) = M_{i,j}, \tag{2.9}$$

where $|\hat{z}[i]| = 1$, $\hat{z}$ is the Fourier transform of a vector in $\mathbb{R}^L$, and $\hat{\rho}' = (\hat{p}', \hat{q}')$ is the Fourier transform of a probability distribution on D_{2L}. Consider the equations

$$\hat{p}'[1]\hat{z}[\ell]\hat{z}[1-\ell] + \hat{q}'[L-1]/(\hat{z}[\ell]\hat{z}[1-\ell]) = M_{\ell,1-\ell}, \quad \ell = 0,\ldots,L-1, \tag{2.10}$$

where the indices are taken modulo L. For each fixed ℓ, we can view equation (2.10) as a linear equation in $\hat{p}'[1], \hat{q}'[L-1]$. For the system to have a solution, it must be consistent. Taking the pair of equations when $\ell = 1$ and $\ell = m+1$ with $m \geq 1$, we obtain

$$\hat{q}'[L-1] = \frac{\left(\hat{z}[1]\hat{z}[m+1]^2 M_{1,0} - \hat{z}[1]^2\hat{z}[m]\hat{z}[m+1]M_{m+1,-m}\right)}{\hat{z}[m+1]^2 - \hat{z}[1]^2\hat{z}[m]^2}. \tag{2.11}$$

Equating equation (2.11) with $m = 1$ and $m = n+1$, we see that $\hat{z}[n+1]$ satisfies the following quadratic equation in terms of $\hat{z}[1], \hat{z}[2], \hat{z}[n]$:

$$\begin{aligned}(M_{2,-1}^2\hat{z}[1]\hat{z}[2] - M_{1,0}\hat{z}[1]^3)\hat{z}[n+1]^2 + M_{n+1,-n}^2(\hat{z}[1]^4\hat{z}[n] - \hat{z}[2]^2\hat{z}[n])\hat{z}[n+1] +& \\ M_{1,0}^2\hat{z}[1]\hat{z}^2[2]\hat{z}[n]^2 - M_{2,-1}^2\hat{z}[1]^3\hat{z}[2]\hat{z}^2[n] = 0.&\end{aligned} \tag{2.12}$$

Note that expressions of the form $M_{i,j}^2$ refer to exponents in this formula.

If $n > 2$, the three equations from (2.9) with $(i,j) = (n+1,0), (n,1), (n-1,2)$, respectively, yield three linear equations for $\hat{p}'[n+1], \hat{q}'[L-n-1]$ whose coefficients are rational expressions in $\hat{z}[1], \hat{z}[2], \hat{z}[n-1], \hat{z}[n], \hat{z}[n+1]$. The same analysis as above shows that $\hat{z}[n+1]$ satisfies an additional quadratic equation in $\hat{z}[1], \hat{z}[2], \hat{z}[n-1], \hat{z}[n]$:

$$\begin{aligned}&(M_{n,1}^2\hat{z}[1]\hat{z}[n] - M_{n-1,2}^2\hat{z}[2]\hat{z}[n-1])\hat{z}[n+1]^2 + M_{n+1,0}(\hat{z}[2]^2\hat{z}[n-1]^2 - \hat{z}[1]^2\hat{z}[n]^2)\hat{z}[n+1] \\ &+ (M_{n-1,2}^2\hat{z}[1]^2\hat{z}[2]\hat{z}[n-1]\hat{z}[n]^2 - M_{n,1}^2\hat{z}[1]\hat{z}[2]^2\hat{z}[n-1]^2\hat{z}[n]) = 0.\end{aligned} \tag{2.13}$$

Since $\hat{z}[n+1]$ satisfies the two non-equivalent quadratic equations (2.12) and (2.13), we can solve for $\hat{z}[n+1]$ in terms of $\hat{z}[1], \hat{z}[2]$ and $\hat{z}[n]$ and we obtain the following expression for $\hat{z}[n+1]$ as a rational function of $\hat{z}[1], \hat{z}[2], \hat{z}[n-1], \hat{z}[n]$:

$$\hat{z}[n+1] = \frac{a}{b}, \tag{2.14}$$

where

$$\begin{aligned}a = {}& M_{1,0}^2M_{n,1}^2(\hat{z}[1]^2\hat{z}[2]^2\hat{z}[n]^3 - \hat{z}[1]^4\hat{z}[2]^2\hat{z}[n-1]^2\hat{z}_n) + M_{2,-1}^2M_{n,1}^2(\hat{z}[1]^2\hat{z}[2]^3\hat{z}[n-1]^2\hat{z}[n] - \hat{z}[1]^4\hat{z}[2]\hat{z}[n]^3) \\ &+ M_{1,0}^2M_{n-1,2}^2(\hat{z}[1]^5\hat{z}[2]\hat{z}[n-1]\hat{z}[n]^2 - \hat{z}[1]\hat{z}[2]^3\hat{z}[n-1]\hat{z}[n]^2),\end{aligned} \tag{2.15}$$

and

$$b = M_{n-1,1}M_{n+1,-n}(\hat{z}[1]\hat{z}[2]^2\hat{z}[n]^2 - \hat{z}[1]^5\hat{z}[n]^2) + M_{n-1,2}^2M_{n+1,-n}^2(\hat{z}[1]^4\hat{z}[2]\hat{z}[n-1]\hat{z}[n] - \hat{z}[2]^3\hat{z}[n-1]\hat{z}[n]) + M_{n+1,0}^2M_{2,-1}^2(\hat{z}[1]\hat{z}[2]^3\hat{z}[n-1]^2 - \hat{z}[1]^3\hat{z}[2]\hat{z}[n]^2) + M_{n+1,0}^2M_{1,0}^2(\hat{z}[1]^5\hat{z}[n]^2 - \hat{z}[1]^3\hat{z}[2]^2\hat{z}[n-1]^2). \tag{2.16}$$

When $n = 2$, the equations in (2.9) corresponding to $(i, j) = (2, 1)$ and $(i, j) = (1, 2)$ are identical so we need another method to express $\hat{z}[3]$ as a rational function of $\hat{z}[1], \hat{z}[2]$. To get a second quadratic equation in this case, consider the equations of (2.9) corresponding to the pairs $(2, 0), (1, 1), (3, -1)$ to obtain the quadratic equation

$$(M_{1,1}^2\hat{z}[1]^2 - M_{2,0}^2\hat{z}[2])\hat{z}[3]^2 + M_{3,-1}^2\hat{z}[1](\hat{z}[2]^2 - \hat{z}[1]^4) + \hat{z}[1]^4\hat{z}[2](M_{2,0}^2\hat{z}[1]^2 - M_{1,1}^2\hat{z}[2]) = 0. \tag{2.17}$$

We then obtain the following expression for $\hat{z}[3]$ as a rational function of $\hat{z}[1]$ and $\hat{z}[2]$:

$$\frac{\hat{z}[1]\hat{z}[2](M_{1,0}^2M_{2,0}^2\hat{z}[1]^4 - M_{1,0}M_{1,1}^2\hat{z}[1]^2\hat{z}[2] - M_{2,0}M_{2,-1}\hat{z}[1]^2\hat{z}[2] + M_{1,0}M_{2,0}\hat{z}[2]^2)}{(M_{1,0}M_{3,-1}\hat{z}[1]^4 - M_{2,-1}M_{3,-1}\hat{z}[1]^2\hat{z}[2] - M_{1,1}M_{3,-2}\hat{z}[1]^2\hat{z}[2] + M_{2,0}M_{3,-2}\hat{z}[2]^2)}. \tag{2.18}$$

At this point we have shown that knowledge of $\hat{z}[1], \hat{z}[2]$ determine $\hat{z}[n+1]$ for $n \geq 2$, assuming that the rational expressions (2.14) and (2.18) are well defined (see the discussion at the end of the proof). We can also use the quadratic equations (2.12) (with $n = 3$) and (2.17) to obtain a second expression for $\hat{z}[3]^2$ as a rational function of $\hat{z}[1]$ and $\hat{z}[2]$. Equating this expression for $\hat{z}[3]^2$ with the square of the expression for $\hat{z}[3]$ given by (2.18), we obtain the following palindromic quartic equation for $\hat{z}[2]$ in terms of $\hat{z}[1]$:

$$A_0\hat{z}[1]^8 + A_1\hat{z}[1]^6\hat{z}[2] + A_2\hat{z}[1]^4\hat{z}[2]^2 + A_1\hat{z}[1]^2\hat{z}[2]^3 + A_0\hat{z}[1]^4 = 0, \tag{2.19}$$

where

$$\begin{aligned}
A_0 &= M_{1,0}^2M_{2,0}^2 - M_{1,0}M_{2,0}M_{3,-2}M_{3,-1} \\
A_1 &= -2M_{1,0}^2M_{1,1}M_{2,0} - 2M_{1,0}M_{2,-1}M_{2,0}^2 + M_{1,1}M_{2,0}M_{3,-2}^2 \\
&\quad +M_{1,0}M_{1,1}M_{3,-2}M_{3,-1} + M_{2,-1}M_{2,0}M_{3,-2}M_{3,-1} + M_{1,0}M_{2,-1}M_{3,-1}^2 \\
A_2 &= M_{1,0}^2M_{1,1}^2 + 2M_{1,0}M_{1,1}M_{2,-1}M_{2,0} + 2M_{1,0}^2M_{2,0}^2 + M_{2,-1}^2M_{2,0}^2 \\
&\quad -M_{1,1}^2M_{3,-2}^2 - M_{2,0}^2M_{3,-2}^2 - 2M_{1,1}M_{2,-1}M_{3,-2}M_{3,-1} - M_{1,0}^2M_{3,-1}^2 - M_{2,-1}^2M_{3,-1}^2.
\end{aligned}$$

Taking the complex conjugate of (2.19) and using the fact that $\hat{z}[i]$ lies on the unit circle so $\overline{\hat{z}[i]} = \hat{z}[i]^{-1}$, we obtain

$$\overline{A_0}\hat{z}[1]^{-8} + \overline{A_1}\hat{z}[1]^{-6}\hat{z}[2]^{-1} + \overline{A_2}\hat{z}[1]^{-4}\hat{z}[2]^{-2} + \overline{A_1}\hat{z}[1]^{-2}\hat{z}[2]^{-3} + \overline{A_0}\hat{z}[1]^{-4}. \tag{2.20}$$

Multiplying (2.20) by $\hat{z}[1]^8\hat{z}[2]^4$, we obtain a second quartic equation satisfied by $\hat{z}[2]$:

$$\overline{A_0}\hat{z}[1]^8 + \overline{A_1}\hat{z}[1]^6\hat{z}[2] + \overline{A_2}\hat{z}[1]^4\hat{z}[2]^2 + \overline{A_1}\hat{z}[1]^2\hat{z}[2]^3 + \overline{A_0}\hat{z}[2]^4. \tag{2.21}$$

Now take $\sqrt{-1}(A_0(2.21) - \overline{A_0}(2.19))$ and we obtain the following equation with real coefficients

$$\hat{z}[1]^2\hat{z}[2](B_1\hat{z}[1]^4 + B_2\hat{z}[1]^2\hat{z}[2] + B_1\hat{z}[2]^2) = 0, \tag{2.22}$$

where $B_1 = 2\Im(\overline{A_0}A_1)$ and $B_2 = 2\Im(\overline{A_0}A_2)$ ($\Im$ stands for the imaginary part of a complex number). Since $\hat{z}[1], \hat{z}[2] \neq 0$, we see that $\hat{z}[2]$ satisfies the real palindromic equation

$$B_1\hat{z}[1]^4 + B_2\hat{z}[1]^2\hat{z}[2] + B_1\hat{z}[2]^2 = 0. \tag{2.23}$$

Since the equation is palindromic, if $\hat{z}[2]$ is a root then $1/\hat{z}[2] = \overline{\hat{z}[2]}$ is also necessarily a root.

At this point we have shown that given $\hat{z}[1]$, there are (at most) two possible values for $\hat{z}[2]$ provided that B_1, B_2 are non-zero. Once we have $\hat{z}[1], \hat{z}[2]$, the values of $\hat{z}[3], \ldots, \hat{z}[L/2]$ are uniquely determined, assuming that the rational expressions (2.14) and (2.18) are well-defined. However, we have no constraints on $\hat{z}[1]$ other than it lies on the unit circle. Indeed all of the moment equations we considered are weighted homogeneous where the variable $\hat{z}[n]$ has weight n. In other words, if $(\hat{z}[1], \hat{z}[2], \hat{z}[3], \ldots, \hat{z}[L/2])$ is a solution, then $(\lambda\hat{z}[1], \lambda^2\hat{z}[2], \ldots, \lambda^{L/2}\hat{z}[L/2])$ will be a solution for any $\lambda \in S^1$. When L is even we obtain a constraint on $\hat{z}[1]$ by noting that $\hat{z}[L/2]^2 = 1$ since $z[L/2] = 1/z[L/2]$ because $L/2 = L - L/2$. Hence we must have $(\lambda^{L/2})^2 = 1$; i.e., $\lambda^L = 1$, so λ is an L-th root of unity. Hence our system can have at most $2L$ solutions. When L is odd, we observe that $\hat{z}[(L-1)/2] = \overline{\hat{z}[L-(L-1)/2]} = \overline{\hat{z}[(L+1)/2]} = \hat{z}[(L+1)/2]^{-1}$, and so $\hat{z}[(L-1)/2]\hat{z}[(L+1)/2] = 1$; and replacing $\hat{z}[k]$ with $\lambda^k\hat{z}[k]$, we find $\lambda^L\hat{z}[(L-1)/2]\hat{z}[(L+1)/2] = 1$, i.e., $\lambda^L = 1$. Hence, our system can only have at most $2L$ solutions in this case as well.

Generic Conditions. To complete the proof, we explain why for generic (x, ρ) with all $\hat{x}[i]$ non-zero, the quadratic equation (2.23) is non-zero and the rational expressions (2.14) and (2.18) are well-defined. To show that (2.23) is non-vanishing for generic (x, ρ) we must show that $\overline{A_0}A_1$ and $\overline{A_0}A_2$ are not pure real. This is a real polynomial condition on A_0, A_1, A_2, which are themselves polynomials in the entries of ρ and x. To prove that this condition holds generically, it suffices to prove that this is the case for a single choice of (x, ρ). Moreover, since the simplex is Zariski dense in the linear subspace $\sum p[i] + q[i] = 1$, it suffices to verify this when the vector ρ lies in this subspace without necessarily being a probability distribution. Applying the Fourier transform, it suffices to verify that the condition holds for a single pair $(\hat{x}, \hat{p})$ with $\hat{p}[0] + \hat{q}[0] = 1$. The expressions for A_0, A_1, A_2 are determined by the moment entries $M_{1,0}, M_{2,-1}, M_{3,-2}, M_{2,0}, M_{1,1}, M_{3,-2}$, which are in turn determined by the seven values $\hat{x}[1], \hat{x}[2], \hat{x}[3], \hat{p}[1], \hat{p}[2], \hat{q}[L-1], \hat{q}[L-2]$. In particular if we set $\{x[1], x[2], x[3]\} = \{1, 1, \sqrt{-1}\}$, $\{\hat{p}[1], \hat{p}[2]\} = \{1, 1\}$ and $\hat{\{}q[L-1], q[L-2]\} = \{1 + \sqrt{-1}, \sqrt{-1}\}$ then $B_1 = 16$ and $B_2 = -32$.

Since $(z, \rho') = (x, \rho)$ automatically satisfies the system of equations (2.9), it follows that $\hat{z}[1] = \lambda x[1]$, where λ is an L-th root of unity. Moreover, we know that that when $\hat{z}[1] = \hat{x}[1]$, the quadratic equation (2.23) has solutions $\hat{z}[2] = \{\hat{x}[2], 1/\hat{x}[2]\}$. Hence, if $\hat{z}[1] = \lambda\hat{x}[1]$ then (2.23) has solutions $\hat{z}[2] = \{\lambda^2\hat{x}[2], 1/(\lambda^2\hat{x}[2])\}$. It follows that the rational expression (2.18) is well-defined as long the polynomial expressions

$$(M_{1,0}^2 M_{2,0}^2 \hat{z}[1]^4 - M_{1,0}M_{1,1}^2\hat{z}[1]^2\hat{z}[2] - M_{2,0}M_{2,-1}\hat{z}[1]^2\hat{z}[2] + M_{1,0}M_{2,0}\hat{z}[2]^2),$$

and

$$(M_{1,0}M_{3,-1}\hat{z}[1]^4 - M_{2,-1}M_{3,-1}\hat{z}[1]^2\hat{z}[2] - M_{1,1}M_{3,-2}\hat{z}[1]^2\hat{z}[2] + M_{2,0}M_{3,-2}\hat{z}[2]^2),$$

are both non-zero when $\{\hat{z}[1], \hat{z}[2]\} = \{\lambda\hat{x}[1], \lambda^2\hat{x}[2]\}$ or $\{\hat{z}[1], \hat{z}[2]\} = \{\lambda\hat{x}[1], 1/(\lambda^2\hat{x}[2])\}$. If this is the case, then it follows that $\hat{z}[3] = \lambda^3\hat{x}[3]$ or $\hat{z}[3] = 1/(\lambda^3\hat{x}[3])$ because we know that $(\hat{x}[1], \lambda\hat{x}[2], \lambda^3\hat{x}[3])$ and $(\hat{x}[1], 1/(\lambda^2\hat{x}[2]), 1/(\lambda^2\hat{x}[3]))$ are the first three entries of a vector in the D_{2N} orbit of the vector $\hat{x}$. Using (2.15) and (2.16), we can now continue recursively to obtain sufficient genericity conditions on the pair (x, ρ).

$\square$

Remark 2.4. *As can be seen from the proof, we only use $\sim 5L/2$ of the entries of $M_{i,j}$ (out of L^2 entries overall) to determine the orbit of x. Precisely, we only use the $\sim 5L/2$ entries $M_{\ell,-\ell}, M_{\ell,1-\ell}, M_{\ell+1,0}, M_{\ell,1}$ and $M_{\ell-1,2}$ for $\ell = 0, \ldots, L/2$. A similar observation was made in [15].*

2.3 General theory for MRA with a general distribution over finite groups

The dihedral MRA model (1.1) we consider here is a special case of the general MRA problem considered in [8]. The general MRA problem involves recovering a signal $x \in V$ from measurements $g_i \cdot x + \epsilon_i$, where the group elements g_i are chosen from a compact group G and V is a vector space. In this section, we analyze a discrete version of the generalized MRA problem. In our model, G is a finite group and V is a finite-dimensional vector space. The group elements g_i are taken from a *generic distribution* on the finite group G, rather than a uniform distribution (with respect to the Haar measure) as in [8]. To make this notion precise, we observe that a probability distribution on a finite group is a function $\rho\colon G \to \mathbb{R}$ satisfying the condition $\rho(g) \geq 0$ for all $g \in G$ and $\sum_{g\in G} \rho(g) = 1$. The set of functions $G \to \mathbb{R}$ is a finite-dimensional representation of G called the *regular representation*, denoted by $R(G)$, and the set of probability distributions is a simplex in this vector space. The group G acts on $R(G)$ by the formula $(g \circ f)(h) = f(g^{-1}h)$.

In the case of a generic distribution ρ on a finite group G acting on a vector space V, the problem of recovery from n-th order moments can be viewed as a problem in invariant theory of the product $R(G) \times V$. By definition, the n-th order moment associated to a probability distribution ρ on G, $M^n := \sum_{g\in G} \rho(g)(gx)^{\otimes n}$ is a n-tensor of invariant polynomials of bidegree $(1, n)$ on $R(G) \times V$. Of particular interest in this paper is the second order moment $M^2(x, \rho) = \sum_{g\in G} = \rho(g)(gx)(gx)^T$, when G is the dihedral group. In this case, the second order moment gives a collection of invariant functions of total degree 3 on $R(G)\times V$. This is in contrast to the model considered in [8], where the distribution is assumed to be uniform and thus the n-th order moment is a collection of invariant polynomials of degree n on the vector space V.

The following result states that the orbits of a generic pair $(\rho, x) \in R(G)$ can be determined from degree 3 invariant polynomials.

Proposition 2.5. *The set of all degree 3 invariants on $R(G) \times V$ determines the G-orbit of a generic pair (ρ, x).*

Remark 2.6. *Note that there are invariant polynomials of total degree 3 on $R(G)\times V$ which cannot be expressed as linear combinations of the entries of the second moment. For example, invariant functions of degree $(3, 0)$ have nothing to do with moments in our sense and are just the degree three invariant of the regular representation $R(G)$.*

Proof. Since the representation $R(G)\times V$ contains a copy of the regular representation, [8, Theorem D2] proves that the generic orbit is linearly independent. It follows that the generic vector can be recovered from degree 3 invariants. $\square$

Following the terminology of [8, Section 1.4], we say that a signal x admits *list recovery* from a set of moment measurements if there are a finite number of orbits with same moments. Using the

Jacobian criterion, it is possible to test if a signal admits list recovery from a collection of moment measurements [8, Section 4.2.2]. Here we use the fact that for unknown probability distributions the moments are linear functions of the probabilities to obtain another criterion for list recovery as well as a criterion for orbit recovery.

To motivate our criterion, we recall the strategy used in the proof of Theorem 2.2. Given a generic probability distribution $\rho = \{p_g\}_{g \in G = D_{2L}}$ and a generic vector $x \in V = \mathbb{R}^L$ we proved that the following system of bilinear equations in the unknowns x', p'_g has at most $2L = |G|$ solutions

$$\sum_g (p_g (gx)^T gx - p'_g (gx')^T gx') = 0. \tag{2.24}$$

The next proposition shows that our verification was equivalent to proving a statement about an *incidence variety* associated to the group G and vector space $V = \mathbb{R}^L$. To formulate the result we first establish notation for the action of a finite group G on a vector space V. Let $I \subset (R(G) \times V)^2$ be the subvariety defined by the bilinear equations (2.24), where the $x, x', \{p_g\}, \{p'_g\}$ are all considered variables. In the language of algebraic geometry, I is called an *incidence variety.* The geometry of the incidence variety I characterizes when orbit and list recovery are possible.

Proposition 2.7. *Let G be a finite group acting on a vector space V, and let $I \subset (R(G) \times V)^2$ be an incidence defined in* (2.24)*.*

1. *If $\dim I = \dim V + |G|$, then for a generic signal x and probability distribution $\rho = \{p_g\}_{g \in G}$, list recovery is possible from the second order moment $M^2(x, \rho)$.*

2. *If $\dim I = \dim V + |G|$ and $\deg I = |G|$, then for a generic signal x and probability distribution $\rho = \{p_g\}_{g \in G}$, orbit recovery is possible from the second order moment $M^2(x, \rho)$.*

Proof. Consider the projection $\pi \colon I \to R(G) \times V$ defined by $(x, \rho, x', \rho') \mapsto (x, \rho)$. If $\dim I = \dim V + R(G)$, then the generic fiber of π must be 0-dimensional. Hence, for a generic vector $x \in V$ and probability distribution $\rho \in R(G)$, there can be at most a finite number of pairs $(x, \rho, x', \rho') \in I$. In other words, there are finite number of vectors x' such that there exist a distribution ρ' with the property that $M^2(x, \rho) = M^2(x', \rho')$ This proves part (i).

Note that for each $g \in G$, the set $X_g = \{(x, \rho, gx, g\rho) | x \in V, \rho \in R(G)\}$ is a $\dim V + |G|$-dimensional subvariety of I, which is isomorphic to $R(G) \times V$. In particular, if $\dim I = \dim V + |G|$ then it must necessarily be an irreducible component of the variety V in the sense of algebraic geometry. Hence, if $\dim I = \dim V + |G|$ then I has at least $|G|$ irreducible components. and therefore its degree must be at least $|G|$. Hence, if $\dim I = \dim |V| + |G|$ and $\deg I = |G|$, then I has exactly $|G|$ irreducible components and for generic x, ρ there will be exactly $|G|$ pairs $(x, \rho, x', \rho') \in I$. Hence each x' must necessarily equal gx for some $g \in G$. Therefore, the second order moments recover generic orbits x in this case. □

Corollary 2.8. *If $|G| > \frac{1}{2}\binom{L+1}{2}$, then list recovery is impossible from second order moments.*

Proof. Since the second order moment tensor is symmetric, the bilinear system (2.24) has $\binom{L+1}{2}$ distinct equations. It follows that if we fix x, ρ then there will be infinitely many solutions (x', ρ'). □

3 Algorithms

In this section, we introduce three algorithmic paradigms to estimate the signal x from dihedral MRA observations $y_1, \ldots, y_n$ as in (1.2). We first introduce the three methods, and then compare them numerically in Section 3.4.

3.1 Group synchronization

If the group elements $g_1, \ldots, g_n \in D_{2L}$ were known, estimating the signal can be done by aligning the observations and averaging out the noise:

$$x_{\text{est}} = \frac{1}{n}\sum_{i=1}^{n} g_i^{-1} y_i. \tag{3.1}$$

This motivates synchronization methods to estimate the unknown group elements from the observations. Synchronization starts by aligning all pairs of observations y_i, y_j, $i \neq j$, so that

$$y_i \approx g_{ij} y_j, \tag{3.2}$$

for some group element $g_{ij} \in D_{2L}$. A standard alignment procedure is based on cross-correlating the observations. In more general groups, other common features can be harnessed; see for example [20, 37]. The relation (3.2) is merely a proxy to $g_i \cdot x \approx g_{ij} g_j \cdot x$, which in turn means that $g_i g_j^{-1} \approx g_{ij}$. At this stage, one reduces the MRA problem to the problem of *group synchronization* [36], where we aim at estimating the unknown group elements $g_1, \ldots, g_n$ from a subset of their ratios $g_i g_j^{-1}$, often corrupted with noise.

Early synchronization studies addressed the problem over compact groups, such as, finite groups, phases, and rotations. The common property of all synchronization cases over compact groups is that we can reduce them all to synchronization over rotations, or a subgroup of rotations, by using a faithful orthogonal representation [17, 19, 28, 39]. Further generalizations extended synchronization methods to non-compact groups, and in particular to the Euclidean group, see e.g., [26, 32].

Specifically for the dihedral MRA problem (1.1), we start by computing the cross-correlation between any observation y_i and any other observation y_j and its reflection sy_j. The maximal value indicates the best alignment as in (3.2). The resulting ratios $g_{ij} \approx g_i g_j^{-1}$ serve as an input for a standard spectral algorithm [36], which uses a rounding procedure onto the dihedral group, resulting in estimates of the group elements $\tilde{g}_1, \ldots, \tilde{g}_n$. The orbit of the signal is then estimated by averaging over the synchronized observations

$$x_{\text{est}} = \frac{1}{n}\sum_{i=1}^{n} \tilde{g}_i^{-1} y_i. \tag{3.3}$$

Unfortunately, in low SNR environments the error of estimating the ratios $g_i g_j^{-1}$, and thus of estimating $g_1, \ldots, g_n$, grows rapidly [5, 30, 29, 13]. Thus, in such regimes we consider techniques which aim to recover the signal x directly, bypassing the estimation of the missing group elements $\{g_i\}_{i=1}^n$. Next, we present two such methods, based on expectation-maximization and the method of moments.

3.2 Maximum likelihood estimation using expectation-maximization

The log-likelihood function of (1.1) is given by

$$\ell(x,\rho) = \log p(y_1,\ldots,y_n;x,\rho) = \sum_{i=1}^{N} \log \sum_{j=1}^{2L} \rho[j] \frac{1}{(2\pi\sigma^2)^{L/2}} e^{-\frac{||y_i - g[j]\cdot x||^2}{2\sigma^2}}, \tag{3.4}$$

where $g[1],\ldots,g[2L]$ are the elements of D_{2L}. This is the standard likelihood function of a Gaussian mixture model, but all centers are connected through the orbit of D_{2L} acting on x. We wish to find the signal x and distribution ρ that maximize (3.4). In the sequel, we assume no prior information on the signal and the distribution. If such information is available, then it is useful to consider the log-posterior distribution $\log p(x,\rho|y_1,\ldots,y_n)$, which is equal to the log-likelihood plus the log of the prior terms.

To maximize the likelihood function, we devise an expectation-maximization (EM) algorithm [?]. The EM algorithm has been successfully applied to other MRA setups [14, 2, 24, 22] as well as for cryo-EM [33, 35, 12]. Although EM is not guaranteed to achieve the maximum of the non-convex likelihood function (3.4), it is guaranteed that each EM iteration does not reduce the likelihood.

EM is an iterative algorithm, and each step consists of two steps. In the first step, called the E-step, the expectation of the complete likelihood (namely, the joint likelihood of x, ρ and the group elements) is computed. The expectation is taken with respect to the group elements (i.e., the nuisance variables), given the current estimates of the signal x_t and the distribution ρ_t:

$$\begin{aligned} Q(x,\rho|x_t,\rho_t) &= \mathbb{E}\left\{\log p(y_1,\ldots,y_n,g_1,\ldots,g_n;x,\rho)\right\} \\ &= \sum_{i=1}^{n} \mathbb{E}\left\{-\frac{1}{2\sigma^2}\|y_i - g_i\cdot x\|^2 + \log\rho[g_i]\right\} + \text{constant} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{2L} w_{i,j}\left\{-\frac{1}{2\sigma^2}\|y_i - g[j]\cdot x\|^2 + \log\rho[j]\right\} + \text{constant}, \end{aligned} \tag{3.5}$$

where

$$w_{i,j} = \frac{\rho_t[j] e^{\frac{-1}{2\sigma^2}\|y_i - g[j]\cdot x_t\|^2}}{\sum_{j=1}^{2L}\rho_t[j] e^{\frac{-1}{2\sigma^2}\|y_i - g[j]\cdot x_t\|^2}}. \tag{3.6}$$

The second step, called M-step, maximizes Q with respect to x and ρ. In our case, the update step reads:

$$\begin{aligned} x_{t+1} &= \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{2L} w_{i,j} g^{-1}[j] y_i \\ \rho_{t+1}[j] &= \frac{\sum_{i=1}^{n} w_{i,j}}{\sum_{i=1}^{n}\sum_{j=1}^{2L} w_{i,j}}. \end{aligned} \tag{3.7}$$

If prior information is available (and thus the EM tries to maximize the posterior distribution rather than the likelihood), then it will act as a regularizer on the solution of the M-step. The EM algorithm iterates between computing the weights (3.6) and updating the parameters (3.7) until a stopping criterion is met.

3.3 The method of moments

The idea behind the method of moments is finding a pair (x, ρ) whose moments match the empirical moments of the observations. In particular, according to Theorem 1.1, only the first two moments are required to characterize uniquely the orbit of generic x and ρ. The empirical moments can be computed from the data simply by averaging:

$$\begin{aligned} M^1_{\text{est}} &= \frac{1}{n}\sum_{i=1}^{n} y_i, \\ M^2_{\text{est}} &= \frac{1}{n}\sum_{i=1}^{n} y_i y_i^T. \end{aligned} \tag{3.8}$$

By the law of large numbers, and using Lemma 2.1, for large n we have

$$\begin{aligned} M^1_{\text{est}} &\approx M^1(x,\rho) = C_{\mathbf{p}}x + C_{\mathbf{q}}sx, \\ M^2_{\text{est}} &\approx M^2(x,\rho) = C_x D_{\mathbf{p}} C_x^T + C_{sx} D_{\mathbf{q}} C_{sx}^T, \end{aligned} \tag{3.9}$$

where $C_z \in \mathbb{R}^{L\times L}$ is a circulant matrix generated by $z \in \mathbb{R}^L$, and $D_z \in \mathbb{R}^{L\times L}$ is a diagonal matrix whose entries are z. As $n \to \infty$, $M^1_{\text{est}} \to M^1(x,\rho)$ and $M^2_{\text{est}} \to M^2(x,\rho)$ almost surely.

A common practice is to estimate (x, ρ) from M^1_{est} and M^2_{est} by minimizing a non-convex least squares objective:

$$\min_{\tilde{x}\in\mathbb{R}^L, [p,q]\in\Delta^{2L}} \left\| M^2_{\text{est}} - C_{\tilde{x}} D_p C_{\tilde{x}}^T - C_{s\tilde{x}} D_q C_{s\tilde{x}}^T \right\|_{\text{F}}^2 + \lambda \left\| M^1_{\text{est}} - C_p\tilde{x} - C_q s\tilde{x} \right\|_2^2. \tag{3.10}$$

The solution of (3.10) is the method of moments estimator. While the objective function (3.10) is non-convex, it seems to provide accurate estimates in many cases.

3.4 Numerical experiments

This section compares numerically the algorithmic methods discussed above: synchronization, expectation-maximization, and the method of moments. We define signal-to-noise ratio (SNR) as $\|x\|^2/(L\sigma^2)$. To account for the group symmetry, we define relative error as

$$\text{relative error} = \min_{g\in D_{2L}} \frac{\|g\cdot x_{\text{est}} - x\|}{\|x\|}, \tag{3.11}$$

where x_{est} is the signal estimate.

We consider two regimes: (i) a relatively small number of observations ($n = 1000$) and moderate SNR levels, and (ii) large n and low SNR. The code to reproduce all experiments is publicly available at `https://github.com/nirsharon/DihedralMRA`. The results below represent the average over 50 trials. We initialized the EM algorithm from a single random point and halted it when the difference of the likelihood between two consecutive iterations dropped below 10^{-4}, or after a maximum of 400 iterations. For the method of moments, we minimized (3.10) using the trust-regions method; we initialized the optimization algorithm from 10 different random initial guesses and chose the one that yields the least value of the cost (3.10). The number of trust-regions iterations was limited to 200.

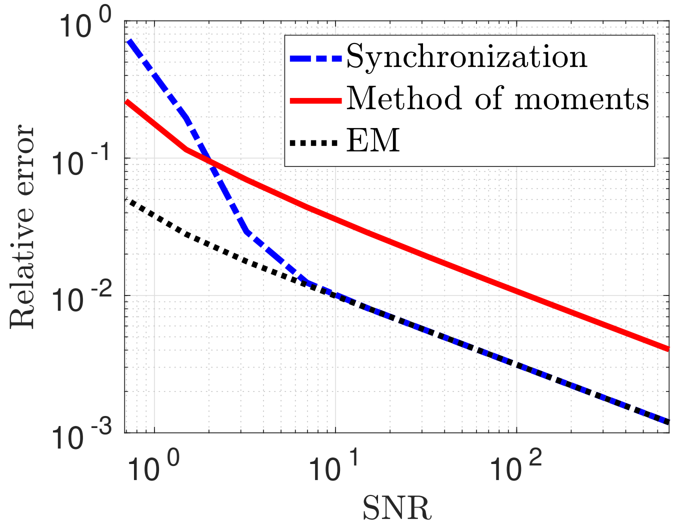


Figure 2: The relative error of the three methods under moderate SNR levels with $n = 1000$ observations. As the SNR deteriorates, the synchronization method fails to estimate the group elements, and thus the signal, accurately.

Moderate SNR regime. We begin with a noise regime where the synchronization approach presents a viable alternative to EM and the method of moments. Figure 2 shows the relative error of the three methods as a function of the SNR with $n = 1000$ observations. The method of moments shows inferior results compared to synchronization and EM since the empirical moments do not approximate the population moments accurately enough for such a small number of observations. For high SNR, the performance of synchronization and EM are comparable. However, as the SNR drops, the synchronization fails to estimate the group elements accurately, while both the method of moments and EM present consistent error rates. As the SNR approaches 1, when the signal and the noise are of the same order, the synchronization method introduces relative error close to 1, meaning it contributes no information about the solution.

Low SNR. We discard the synchronization algorithm in the low SNR regime as it cannot cope with high noise levels, as demonstrated in Figure 2. In addition, since the first step of the synchronization method involves pairwise alignment, the synchronization input consists of $\mathcal{O}(n^2)$ group elements, and so the computational complexity of this method makes it impractical for as many as $n = 10^5$ observations.

Figure 3a shows relative errors as a function of SNR. The EM outperforms the method of moments for SNR values above 1/10. For lower SNR levels, the method of moments shows similar estimation rates. In the high SNR regime, the error curves of both methods scale as $\mathrm{SNR}^{-1/2}$, namely as σ, which is the same estimation rate as if the group elements were known. In particular, the numerical slope of the EM method is -0.4999 and the method of moments presents a numerical slope of -0.5104. In the low SNR regime, however, the error curves scale as $\mathrm{SNR}^{-1} \propto \sigma^2$. While this slope is expected for the method of moments that directly uses the first two moments (and thus its standard deviation is proportional to σ^2), the moments do not appear explicitly in the EM iterations. Specifically, the numerical slopes for SNR values below $1/10$ were -1.0561 and -1.1058

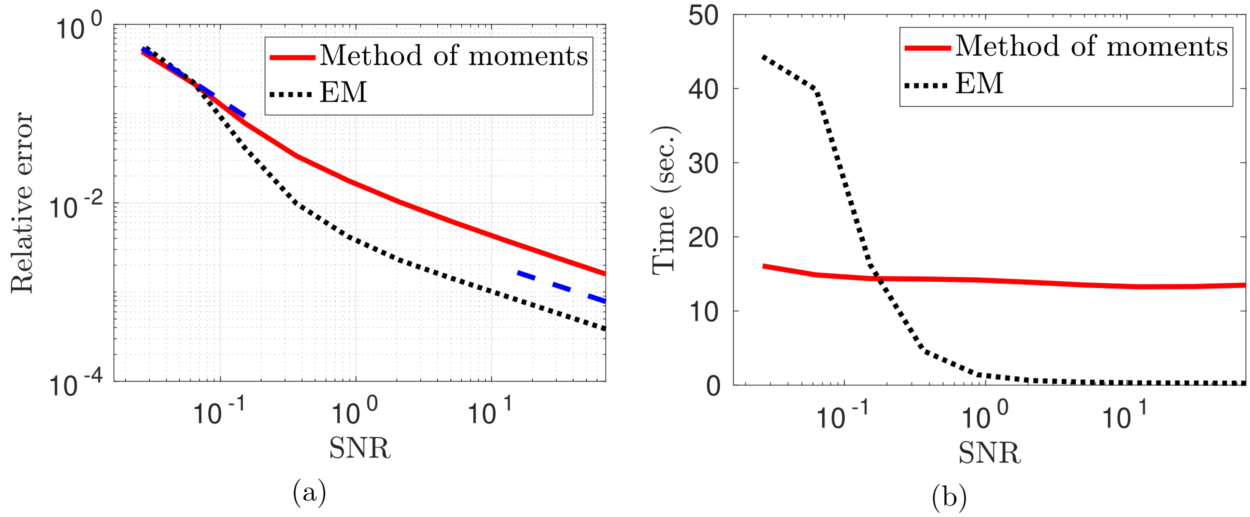


Figure 3: Relative error (left panel) and runtime (right panel) of the method of moments and EM as a function of the SNR for $n = 10^5$ observations. In the high SNR regime, the slope of the error curves (dashed blue line), for both methods, scales as $\mathrm{SNR}^{-1/2} \propto \sigma$, which is the same estimation rate as if the group elements were known. In the low SNR regime, however, the error curves scale as $\mathrm{SNR}^{-1} \propto \sigma^2$, corroborating our theoretical findings.

for the method of moments and EM, respectively. This rate implies that accurate estimation requires $n \gg \sigma^4$, corroborating our theoretical findings (Corollary 1.2) that no algorithm can achieve better estimation rates in the low SNR regime. A similar phenomenon was observed by previous MRA studies [35, 14, 2, 16]. For the connection between EM and the method of moments in the low SNR regime, see [**?**, **?**, **?**].

Figure 3b presents the corresponding average runtime. The runtime of EM increases as the SNR decreases, while the runtime of the method of moments remains roughly constant. The reason for the growth in runtime is revealed in Figure 4, where we display the average number of EM iterations as a function of SNR. The figure shows that the number of iterations is inversely proportional to the SNR.

Acknowledgment

W.L. and N.S. are partially supported by BSF grant no. 2018230. T.B. and N.S are partially supported by the NSF-BSF award 2019752. D.E. is supported by Simons Collaboration grant 708560. W.L. is partially supported by NSF award IIS-1837992.

References

[1] Asaf Abas, Tamir Bendory, and Nir Sharon. The generalized method of moments for multi-reference alignment. *arXiv preprint arXiv:2103.02215*, 2021.

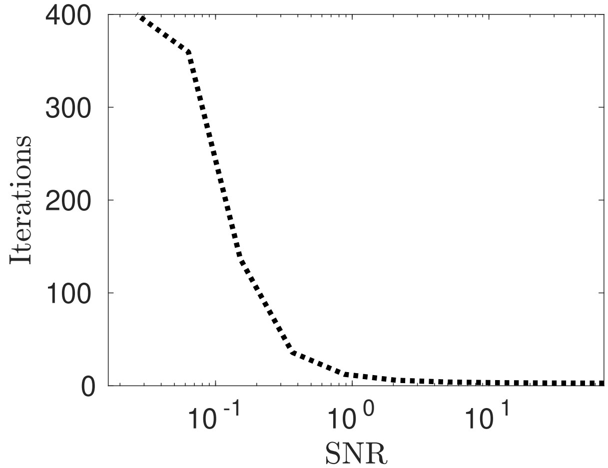


Figure 4: The average number of EM iterations as a function of the SNR. The maximum number of iterations is set to 400.

[2] Emmanuel Abbe, Tamir Bendory, William Leeb, João M Pereira, Nir Sharon, and Amit Singer. Multireference alignment is easier with an aperiodic translation distribution. *IEEE Transactions on Information Theory*, 65(6):3565–3584, 2018.

[3] Emmanuel Abbe, João M Pereira, and Amit Singer. Sample complexity of the boolean multireference alignment problem. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 1316–1320. IEEE, 2017.

[4] Emmanuel Abbe, João M Pereira, and Amit Singer. Estimation in the group action channel. In *2018 IEEE International Symposium on Information Theory (ISIT)*, pages 561–565. IEEE, 2018.

[5] Cecilia Aguerrebere, Mauricio Delbracio, Alberto Bartesaghi, and Guillermo Sapiro. Fundamental limits in multi-image alignment. *IEEE Transactions on Signal Processing*, 64(21):5707–5722, 2016.

[6] Yariv Aizenbud, Boris Landa, and Yoel Shkolnisky. Rank-one multi-reference factor analysis. *Statistics and Computing*, 31(1):1–31, 2021.

[7] Philip R Baldwin and Dmitry Lyumkis. Non-uniformity of projection distributions attenuates resolution in cryo-EM. *Progress in biophysics and molecular biology*, 150:160–183, 2020.

[8] Afonso S Bandeira, Ben Blum-Smith, Joe Kileel, Amelia Perry, Jonathan Weed, and Alexander S Wein. Estimation under group actions: recovering orbits from invariants. *arXiv preprint arXiv:1712.10163*, 2017.

[9] Afonso S Bandeira, Moses Charikar, Amit Singer, and Andy Zhu. Multireference alignment using semidefinite programming. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 459–470, 2014.

[10] Afonso S Bandeira, Yutong Chen, Roy R Lederman, and Amit Singer. Non-unique games over compact groups and orientation estimation in cryo-EM. *Inverse Problems*, 36(6):064002, 2020.

[11] Afonso S Bandeira, Jonathan Niles-Weed, and Philippe Rigollet. Optimal rates of estimation for multi-reference alignment. *Mathematical Statistics and Learning*, 2(1):25–75, 2020.

[12] Tamir Bendory, Alberto Bartesaghi, and Amit Singer. Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities. *IEEE signal processing magazine*, 37(2):58–76, 2020.

[13] Tamir Bendory, Nicolas Boumal, William Leeb, Eitan Levin, and Amit Singer. Multi-target detection with application to cryo-electron microscopy. *Inverse Problems*, 35(10):104003, 2019.

[14] Tamir Bendory, Nicolas Boumal, Chao Ma, Zhizhen Zhao, and Amit Singer. Bispectrum inversion with application to multireference alignment. *IEEE Transactions on signal processing*, 66(4):1037–1050, 2017.

[15] Tamir Bendory, Dan Edidin, and Shay Kreymer. Signal recovery from a few linear measurements of its high-order spectra. *arXiv preprint arXiv:2103.01551*, 2021.

[16] Tamir Bendory, Ariel Jaffe, William Leeb, Nir Sharon, and Amit Singer. Super-resolution multi-reference alignment. *arXiv preprint arXiv:2006.15354*, 2020.

[17] Nicolas Boumal. Nonconvex phase synchronization. *SIAM Journal on Optimization*, 26(4):2355–2377, 2016.

[18] Nicolas Boumal, Tamir Bendory, Roy R Lederman, and Amit Singer. Heterogeneous multireference alignment: A single pass approach. In *2018 52nd Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6. IEEE, 2018.

[19] Luca Carlone, Roberto Tron, Kostas Daniilidis, and Frank Dellaert. Initialization techniques for 3D SLAM: a survey on rotation estimation and its use in pose graph optimization. In *2015 IEEE international conference on robotics and automation (ICRA)*, pages 4597–4604. IEEE, 2015.

[20] Klaus Häming and Gabriele Peters. The structure-from-motion reconstruction pipeline–a survey with focus on short image sequences. *Kybernetika*, 46(5):926–937, 2010.

[21] Matthew Hirn and Anna Little. Wavelet invariants for statistically robust multi-reference alignment. *arXiv preprint arXiv:1909.11062*, 2019.

[22] Noam Janco and Tamir Bendory. An accelerated expectation-maximization for multi-reference alignment. *arXiv preprint arXiv:2105.07372*, 2021.

[23] Ramakrishna Kakarala. Completeness of bispectrum on compact groups. *arXiv preprint arXiv:0902.0196*, 1, 2009.

[24] Chao Ma, Tamir Bendory, Nicolas Boumal, Fred Sigworth, and Amit Singer. Heterogeneous multireference alignment for images with application to 2D classification in single particle reconstruction. *IEEE Transactions on Image Processing*, 29:1699–1710, 2019.

[25] Katerina Naydenova and Christopher J Russo. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nature Communications*, 8(1):1–5, 2017.

[26] Onur Ozyesil, Nir Sharon, and Amit Singer. Synchronization over Cartan motion groups via contraction. *SIAM Journal on Applied Algebra and Geometry*, 2(2):207–241, 2018.

[27] Amelia Perry, Jonathan Weed, Afonso S Bandeira, Philippe Rigollet, and Amit Singer. The sample complexity of multireference alignment. *SIAM Journal on Mathematics of Data Science*, 1(3):497–517, 2019.

[28] Amelia Perry, Alexander S Wein, Afonso S Bandeira, and Ankur Moitra. Message-passing algorithms for synchronization problems over compact groups. *Communications on Pure and Applied Mathematics*, 71(11):2275–2322, 2018.

[29] Dirk Robinson, Sina Farsiu, and Peyman Milanfar. Optimal registration of aliased images using variable projection with applications to super-resolution. *The Computer Journal*, 52(1):31–42, 2009.

[30] Dirk Robinson and Peyman Milanfar. Fundamental performance limits in image registration. *IEEE Transactions on Image Processing*, 13(9):1185–1199, 2004.

[31] Elad Romanov, Tamir Bendory, and Or Ordentlich. Multi-reference alignment in high dimensions: sample complexity and phase transition. *SIAM Journal on Mathematics of Data Science*, 3(2):494–523, 2021.

[32] David M Rosen, Luca Carlone, Afonso S Bandeira, and John J Leonard. SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group. *The International Journal of Robotics Research*, 38(2-3):95–125, 2019.

[33] Sjors HW Scheres. RELION: implementation of a bayesian approach to cryo-em structure determination. *Journal of structural biology*, 180(3):519–530, 2012.

[34] Nir Sharon, Joe Kileel, Yuehaw Khoo, Boris Landa, and Amit Singer. Method of moments for 3D single particle ab initio modeling with non-uniform distribution of viewing angles. *Inverse Problems*, 36(4):044003, 2020.

[35] Fred J Sigworth. A maximum-likelihood approach to single-particle image refinement. *Journal of structural biology*, 122(3):328–339, 1998.

[36] Amit Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and computational harmonic analysis*, 30(1):20–36, 2011.

[37] Amit Singer and Yoel Shkolnisky. Three-dimensional structure determination from common lines in cryo-EM by eigenvectors and semidefinite programming. *SIAM journal on imaging sciences*, 4(2):543–572, 2011.

[38] Yong Zi Tan, Philip R Baldwin, Joseph H Davis, James R Williamson, Clinton S Potter, Bridget Carragher, and Dmitry Lyumkis. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature methods*, 14(8):793–796, 2017.

[39] Roberto Tron, Xiaowei Zhou, and Kostas Daniilidis. A survey on rotation optimization in structure from motion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 77–85, 2016.