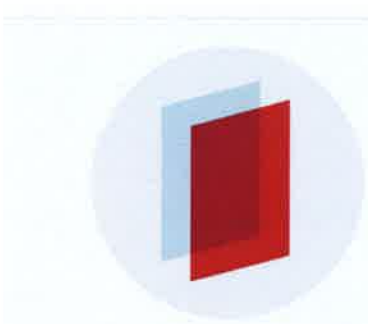
Journal of Statistical Mechanics:
Theory and Experiment

PAPER: DISORDERED SYSTEMS, CLASSICAL AND QUANTUM

Inversion problems for Fourier transforms of particle distributions



Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Disordered systems, classical and quantum

Inversion problems for Fourier transforms of particle distributions

Jaeuk Kim[1], Ge Zhang[2], Frank H Stillinger[3] and Salvatore Torquato[1,3,4,5]

[1] Department of Physics, Princeton University, Princeton, NJ 08544, United States of America
[2] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104-6396, United States of America
[3] Department of Chemistry, Princeton University, Princeton, NJ 08544, United States of America
[4] Princeton Institute for the Science and Technology of Materials, Princeton University, Princeton, NJ 08544, United States of America
[5] Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, United States of America
E-mail: torquato@electron.princeton.edu

Received 12 July 2018
Accepted for publication 4 October 2018
Published 30 November 2018

Online at stacks.iop.org/JSTAT/2018/113302
https://doi.org/10.1088/1742-5468/aae84c

Abstract. Collective coordinates in a many-particle system are complex Fourier components of the local particle density $n(\boldsymbol{x}) \equiv \sum_{j=1}^{N} \delta(\boldsymbol{x} - \boldsymbol{r}_j)$, and often provide useful physical insights. However, given collective coordinates, it is desirable to infer the particle coordinates via inverse transformations. In principle, a sufficiently large set of collective coordinates are equivalent to particle coordinates, but the nonlinear relation between collective and particle coordinates makes the inversion procedure highly nontrivial. Given a 'target' configuration in one-dimensional (1D) Euclidean space, we investigate the minimal set of its collective coordinates that can be uniquely inverted into particle coordinates. For this purpose, we treat a finite number M of the real and/or the imaginary parts of collective coordinates of the target configuration as constraints, and then reconstruct 'solution' configurations whose collective coordinates satisfy these constraints. Both theoretical and numerical investigations reveal that the number of numerically distinct solutions depends sensitively on the chosen collective-coordinate constraints and target configurations. From detailed analysis, we conclude that collective

coordinates at the $\lceil \frac{N}{2} \rceil$ smallest wavevectors is the minimal set of constraints for unique inversion, where $\lceil \cdot \rceil$ represents the ceiling function. This result provides useful groundwork to the inverse transform of collective coordinates in higher-dimensional systems.

Keywords: random/ordered microstructures, structural correlations

Contents

1. Introduction

For N identical point particles at positions of $\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N$ in a periodic fundamental cell Ω, the particle distribution can be described by the local particle density $n(\boldsymbol{x}) \equiv \sum_{j=1}^{N} \delta(\boldsymbol{x} - \boldsymbol{r}_j)$. Equivalently, this function can be represented by the (complex) Fourier components at wavevectors $\boldsymbol{k}$'s, associated with the geometry of Ω, i.e.

$$\tilde{n}(\boldsymbol{k}) \equiv \sum_{j=1}^{N} \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j}, \tag{1}$$

called *collective coordinates*. These quantities are often found to be a natural way to describe the distribution of particles, and thereby provide useful insights into many physical problems, e.g. excited states of liquid helium [1], conduction electrons in metals [2], general theory of simple liquids [3], and quantification of density fluctuations

[4, 5]. Furthermore, using functional Fourier transformation, governing equations of many-body systems, such as the Fokker–Planck equation, can be expressed in terms of collective coordinates [6].

It is often desirable to infer the particle coordinates from given collective coordinates via inverse transformations. Importantly, amplitudes of collective coordinates, or equivalently, the *structure factor*, $S(\mathbf{k})$, have long been used to probe the particle distributions, since $S(\boldsymbol{k})$ can be ascertained from scattering experiments [7]. However, unless the particle distribution is a perfect crystal, the structure factor alone cannot uniquely determine the particle distribution because it does not contain phase information. To solve this problem in x-ray crystallography, additional information is acquired from other physical properties, such as the interference pattern with known molecules (specific site labeling) [8], anomalous dispersion relations [9, 10], or sequential projections onto constrained hyperplanes [11]. Such inversion tasks are called the *phase-retrieval problems* [11–13] because the tasks are essentially equivalent to retrieving the 'phase' information contained in collective coordinates, the complete set of which are in principle invertible into particle coordinates. Even if the phase information is incorporated, however, this inversion task is still highly nontrivial, due to the nonlinear relation between collective and particle coordinates.

Given a *target* point configuration in 1D Euclidean space $\mathbb{R}$, our primary objective in this paper is to find the minimal set of its collective coordinates that uniquely determine particle coordinates aside from exchange of particle indices. This minimal set, therefore, uniquely determines collective coordinates at other wavevectors. To carry out this search, we treat the number M of the real and/or the imaginary parts of collective coordinates of a target configuration as *constraints*, and find all configurations, called *solutions*, whose collective coordinates satisfy these constraints. The number of constraints M is increased one-by-one until we have a unique solution that is, of course, identical to the target pattern.

Previous studies on this inversion task [5, 14–16] focused on some special types of constraints on the collective coordinates (defined by equation (1)) for a given set of wavevectors, such as the *stealthy* constraints, where $\tilde{n}(\boldsymbol{k}) = 0$, and amplitude-constraints for a prescribed radial function $f(r)$, i.e. $|\tilde{n}(\boldsymbol{k})| = f(|\boldsymbol{k}|)$. This inversion task is often carried out via the *collective-coordinate optimization technique* [15–19] that is designed to find ground-state configurations of the potential associated with those constraints. Here, it is useful to define a new parameter $\chi \equiv \mathcal{M}/(dN)$ [15, 17] that represents the relative fraction of the number of constrained collective coordinates $\mathcal{M}$ to the total number of degrees of freedom; see figure 1 for typical arrangements of the constraints in $d = 1, 2$. These studies analytically or numerically showed that when the stealthy constraints are imposed for $\chi < 1/2$, the associated ground states, called stealthy disordered hyperuniform systems [5, 15–17], are disordered, highly degenerate, and statistically isotropic. Importantly, it has been shown that systems derived from these special disordered point configurations by decorating the points with particles of certain shapes, are endowed with some novel photonic and transport properties [20–27]; see also [28] and references therein. Under the stealthy constraints with $\chi \geqslant 1/2$, on the other hand, virtually all configurations are crystalline in the first three spatial dimensions [5, 14, 17]. From the uniqueness of the solution at $\chi = 1/2$ in $d = 1$ [14] as well as the importance of phase information of collective coordinates, one can argue

that each constrained collective coordinate $\tilde{n}(\boldsymbol{k})$ removes *two* degrees of freedom in the accessible configurational space. Thus, it is natural to surmise that the minimum value of M for the unique inversion would be $M = d(N-1)$.

In the present work, we consider more general type of constraints, in which the real and/or the imaginary part of each collective coordinate are independently prescribed. For simplicity, we focus on 1D systems. For such systems, we show that the minimal set of collective-coordinate constraints consists of collective coordinates at the $\lceil N/2 \rceil$ smallest wavevectors, i.e. $M = 2\lceil N/2 \rceil$ rather than N. This result also implies that for a collective coordinate at a wavevector $\boldsymbol{k}$, both its real and imaginary parts must be specified. We analytically show this result for small systems of $N \leqslant 3$. However, this result is invalid if the target configurations are the integer lattice because one cannot determine its center of mass without a collective coordinate at the first Bragg peak. In our numerical studies for larger systems, we exclude the pathological case (i.e. the integer lattice), and consider two distinct ensembles of target configurations: perturbed lattices [29] via uniformly distributed displacements, and Poisson point distribution configurations. For each of these target configurations, we find solutions numerically via the collective-coordinate optimization technique. Our numerical results show that these two types of ensembles occupy qualitatively different energy landscapes: those in perturbed lattices are relatively simpler than those in Poisson ones.

In section 2, we present basic definitions and background. In section 3, we describe the numerical method that we employ to find solutions. In section 4, we theoretically and numerically determine the minimal sets of collective coordinates for small systems. Larger systems are numerically investigated in section 5. Finally, we provide concluding remarks in section 6.

2. Basic definitions and background

2.1. General properties of collective coordinates

For a N-particle point configuration within a periodic fundamental cell Ω, collective coordinates (1), which are also known as *collective density variables*, are complex-valued quantities that are defined at certain real-valued discrete wavevectors $\boldsymbol{k}$'s. Here, the available wavevectors correspond to the reciprocal lattice vectors of the cell Ω. For instance, if Ω is a $L_1 \times \cdots \times L_d$ rectangular box, then $\boldsymbol{k}$'s can be described as follows: $\boldsymbol{k} = 2\pi(\frac{m_1}{L_1}, \cdots, \frac{m_d}{L_d})$ for $(m_1, \cdots, m_d) \in \mathbb{Z}^d$. For simplicity, we focus on one-dimensional systems in the rest of this paper, and thus use the following short-hand notation:

$$k_m = 2\pi m/L. \tag{2}$$

At two different wavevectors, the collective coordinates are not always independent. For instance, the complex conjugate of a collective coordinate by definition is equal to its parity inversion, i.e. $\tilde{n}^*(\boldsymbol{k}) = \tilde{n}(-\boldsymbol{k})$. Thus, if we constrain such a pair of collective coordinates, only one of them is considered independent. For this reason, the relative fraction χ of constrained degrees of freedom is defined as not $2\mathcal{M}/(dN)$, but $\mathcal{M}/(dN)$; see figure 1.

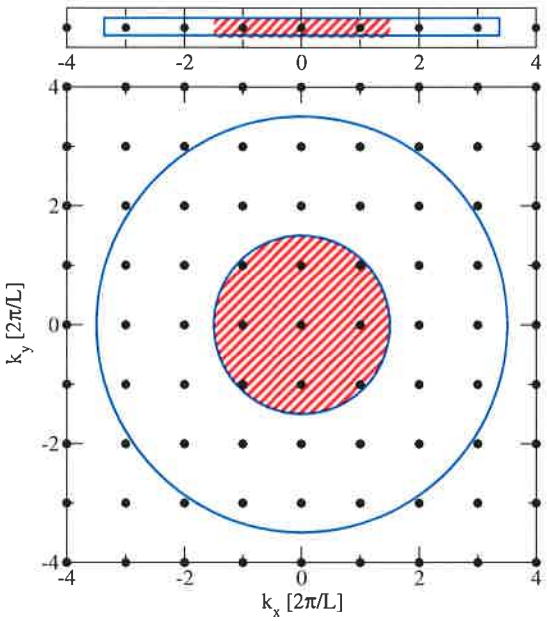


Figure 1. Schematics of typical arrangements of collective-coordinate constraints in Fourier space for a periodic d-dimensional square fundamental cell of side length L. Here, upper and lower panels represent cases for $d = 1$ and 2, respectively. Constraints are taken from $\tilde{n}(\boldsymbol{k})$'s at wavevectors between two concentric circles centered at the origin: there are $2\mathcal{M}$ wavevectors (black dots) within the blue circle, except for $2N_k + 1$ wavevectors inside the red-shaded region. In [5, 14–17], a spherical region with $N_k = 0$ was considered; see a list of available $\mathcal{M}$ values for two-dimensional cases in table II in [17]. For our present purposes, the number of constraints is denoted by $M = 2\mathcal{M}$ because the real and/or the imaginary parts of collective coordinates are considered independently.

Only certain sets of complex numbers can be collective coordinates of a 'realizable' point configuration. For example, there are some trivial necessary conditions of realizable collective coordinates, such as $|\tilde{n}(\boldsymbol{k})| \leqslant N$ for any wavevector $\boldsymbol{k}$, and $\tilde{n}(\boldsymbol{0}) = N$. However, it is highly nontrivial to find sufficient and necessary conditions of realizable collective coordinates. To avoid such realizability problems [30], we take constraints from the collective coordinates of a target configuration.

The value of a collective coordinate is independent of the choice of particle permutations: when we invert collective coordinates, the resulting particle coordinates also should be equivalent aside from exchange of particle indices.

2.2. Definitions

In the rest of this work, we clearly distinguish a target and a solution configurations by using separate notations $\boldsymbol{R}^N = \{R_1, R_2, \cdots, R_N\}$ and $\boldsymbol{r}^N = \{r_1, r_2, \cdots, r_N\}$, respectively. The corresponding collective coordinates are denoted by $\tilde{n}_T(k)$ and $\tilde{n}(k)$, respectively.

In numerical studies, two types of target configurations at unit number density are considered:

(i) Perturbed lattices [29, 31], generated from the integer lattice by independently displacing each particle via a uniform distribution in $[-\delta, \delta]$, and
(ii) Poisson point distribution configurations.

We note that the perturbed lattices become identical to the Poisson point distribution configurations if $\delta = N/2$ under the periodic boundary condition.

We denote M constraints, used in the inversion task, by $E_i = 0$ for $i = 1, 2, \cdots, M$. Starting from the origin in the Fourier space, we skip the first N_k wavenumbers and constrain the collective coordinates at the next $\lfloor M/2 \rfloor$ wavenumbers:

$$E_i \equiv \begin{cases} \mathrm{Re}\left[\tilde{n}_T(k_{N_k+m}) - \tilde{n}(k_{N_k+m})\right], & i = 2m-1 \ (i < M) \\ \mathrm{Im}\left[\tilde{n}_T(k_{N_k+m}) - \tilde{n}(k_{N_k+m})\right], & i = 2m, \quad (i \leqslant M) \end{cases} \tag{3}$$

where $\lfloor x \rfloor$ is the floor function, $m \in \mathbb{N}$, and $\mathrm{Re}\,[x]$ and $\mathrm{Im}\,[x]$ represent the real and the imaginary parts of a complex number x, respectively. Thus, if M is an even number, both the real and the imaginary parts of collective coordinates at $M/2$ consecutive wavenumbers are constrained. If M is an odd number, we prescribe the last term E_M via two conditions, each of which is concerning either the real or the imaginary parts of a target collective coordinate as follows:

$$E_M = \begin{cases} \mathrm{Re}\left[\tilde{n}_T\big(k_{N_k+\lceil M/2 \rceil}\big) - \tilde{n}\big(k_{N_k+\lceil M/2 \rceil}\big)\right], & (4) \\ \mathrm{Im}\left[\tilde{n}_T\big(k_{N_k+\lceil M/2 \rceil}\big) - \tilde{n}\big(k_{N_k+\lceil M/2 \rceil}\big)\right], & (5) \end{cases}$$

where $\lceil x \rceil$ is the ceiling function. Table 1 lists some examples of constraints.

3. Numerical method

Given a target configuration $\boldsymbol{R}^N$ of $N \geqslant 3$, we take M constraints from its collective coordinates, and numerically find solution configurations $\boldsymbol{r}^N$ via a modified 'collective-coordinate optimization technique' [15–19] that was initially designed to generate disordered classical point configurations, such as stealthy ground states [5, 15, 32], and the perfect-glass model [33]. The detailed procedure is described as follows:

(i) Starting from a random initial configuration $\{r_i^{(0)}\}_{i=1}^N$ of N particles, numerically search for an energy-minimizing configuration $\boldsymbol{r}^N \equiv \{\boldsymbol{r}_i\}_{i=1}^N$ for the following potential energy,

$$\begin{aligned} \Phi\big(\boldsymbol{r}^N; \boldsymbol{R}^N\big) &\equiv \sum_{l=1}^{M} \left|E_l\big(\boldsymbol{r}^N; \boldsymbol{R}^N\big)\right|^2 \\ &= \begin{cases} \sum_{l=N_k+1}^{M/2+N_k} \left|\tilde{n}_T(k_l) - \tilde{n}(k_l)\right|^2, & M \text{ is even} \\ \sum_{l=N_k+1}^{\lfloor M/2 \rfloor+N_k} \left|\tilde{n}_T(k_l) - \tilde{n}(k_l)\right|^2 + \left|E_M\big(\boldsymbol{r}^N; \boldsymbol{R}^N\big)\right|^2, & M \text{ is odd.} \end{cases} \end{aligned} \tag{6}$$

The jth component of its gradient is given by

Table 1. Examples of constraints E_{i} for corresponding shorthand notations. We note that when M is an even number, the real condition (4) and the imaginary condition (5) give the identical collective-coordinate constraints.

	E_1	E_2	E_3	E_4
$N_k = 0$ and $M = 4$	$\mathrm{Re}\,[\tilde{n}_T(k_1) - \tilde{n}(k_1)]$	$\mathrm{Im}\,[\tilde{n}_T(k_1) - \tilde{n}(k_1)]$	$\mathrm{Re}\,[\tilde{n}_T(k_2) - \tilde{n}(k_2)]$	$\mathrm{Im}\,[\tilde{n}_T(k_2) - \tilde{n}(k_2)]$
$N_k = 1$ and $M = 4$	$\mathrm{Re}\,[\tilde{n}_T(k_2) - \tilde{n}(k_2)]$	$\mathrm{Im}\,[\tilde{n}_T(k_2) - \tilde{n}(k_2)]$	$\mathrm{Re}\,[\tilde{n}_T(k_3) - \tilde{n}(k_3)]$	$\mathrm{Im}\,[\tilde{n}_T(k_3) - \tilde{n}(k_3)]$
$N_k = 0$, $M = 3$, and the real condition (4)	$\mathrm{Re}\,[\tilde{n}_T(k_1) - \tilde{n}(k_1)]$	$\mathrm{Im}\,[\tilde{n}_T(k_1) - \tilde{n}(k_1)]$	$\mathrm{Re}\,[\tilde{n}_T(k_2) - \tilde{n}(k_2)]$	·
$N_k = 0$, $M = 3$, and the imaginary condition (5)	$\mathrm{Re}\,[\tilde{n}_T(k_1) - \tilde{n}(k_1)]$	$\mathrm{Im}\,[\tilde{n}_T(k_1) - \tilde{n}(k_1)]$	$\mathrm{Im}\,[\tilde{n}_T(k_2) - \tilde{n}(k_2)]$	·

$$F_j(\boldsymbol{r}^N; \boldsymbol{R}^N) \equiv -\frac{\partial \Phi}{\partial r_j}(\boldsymbol{r}^N; \boldsymbol{R}^N) = \begin{cases} \sum_{l=N_k+1}^{M/2+N_k} 2k_l \mathrm{Im}\left[(\tilde{n}(k_l) - \tilde{n}_T(k_l))\, \mathrm{e}^{\mathrm{i}k_l\, r_j}\right], & M \text{ is even} \\ \sum_{l=N_k+1}^{\lfloor M/2 \rfloor + N_k} 2k_l \mathrm{Im}\left[(\tilde{n}(k_l) - \tilde{n}_T(k_l))\, \mathrm{e}^{\mathrm{i}k_l\, r_j}\right] - 2E_M \frac{\partial E_M}{\partial r_j}, & M \text{ is odd.} \end{cases} \tag{7}$$

where E_l is defined by (3), and for an odd number M, E_M is defined by one of two conditions (4) and (5). This configuration is called a 'solution' if $\Phi(\boldsymbol{r}^N; \boldsymbol{R}^N) < \epsilon_E$ for a specified small tolerance ϵ_E.

(ii) Test if this solution $\boldsymbol{r}^N$ agrees with the target configuration $\boldsymbol{R}^N$ or other solutions found previously within another small tolerance ϵ_X, i.e. $\max_{i=1}^N \{\min_{j=1}^N \{|r_{\mathrm{i}} - R_j|\}\} < \epsilon_X$. If they agree, then $\boldsymbol{r}^N$ is deemed to be identical to one of the previous solutions, and we increase the solution's count. Otherwise, we record $\boldsymbol{r}^N$ as a new solution.

(iii) Repeat the steps (i)–(ii) for N_{I} random initial configurations.

(iv) Repeat the steps (i)–(iii) for N_{T} different target configurations.

Roughly speaking, the potential (6) represents a 'deviation' or numerical error of a solution configuration from the target configuration in terms of given collective-coordinate constraints. In step (i), we mainly use two different optimization algorithms: the low-storage BFGS (L-BFGS) algorithm [34, 35] with the MINOP algorithm [15, 36], and the steepest descent algorithm [37]. We repeat this inversion task for many distinct initial configurations $\{r_{\mathrm{i}}^{(0)}\}_{i=1}^N$s and target configurations $\boldsymbol{R}^N$s. Unless stated otherwise, we use parameters as follows: $N_{\mathrm{I}} = 1000$, $N_{\mathrm{T}} = 1000$, and $\epsilon_X = 10^{-6}$.

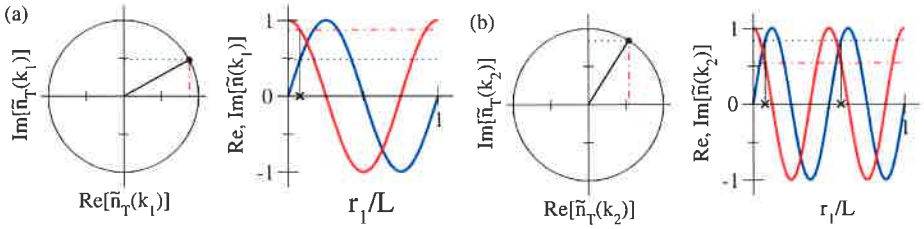


Figure 2. Illustrations for solutions of the inversion problem for a single-particle target configuration. (a) Cases with $N_k = 0$ and $M = 2$. When $\tilde{n}_T(k_1)$ is given as constraints (left), both its real and imaginary parts are required for a unique solution; see the cross (×) mark in the right panel. Red and blue lines represent the real and the imaginary parts of $\tilde{n}(k_1)$ of a solution, respectively. (b) Cases with $N_k = 1$ and $M = 2$. When $\tilde{n}_T(k_2)$ is given, we have two solutions.

For all numerically distinct solutions $\{\boldsymbol{r}^N\}$ of a target configuration $\boldsymbol{R}^N$, the *trivial solution* refers to the one that is identical to the target ($\boldsymbol{r}^N = \boldsymbol{R}^N$), while *nontrivial* solutions refer to the others ($\boldsymbol{r}^N \neq \boldsymbol{R}^N$).

4. Results for $N \leq 3$

Here, we theoretically and numerically investigate solutions for small target configurations.

4.1. $N = 1$

For a single-particle configuration, $\tilde{n}(k_1) = \mathrm{e}^{-\mathrm{i}2\pi r_1/L}$ is a one-to-one function from $r_1 \in [0, L)$ onto the unit circle on the complex plane, i.e. $\{z \in \mathbb{C} : |z| = 1\}$. Thus, it is straightforward to show that there is a unique solution, given constraints $\tilde{n}_T(k_1)$ that correspond to the cases of $N_k = 0$ and $M = 2$. Equivalently, collective coordinates at larger wavenumbers can be expressed in terms of $\tilde{n}_T(k_1)$, i.e. $\tilde{n}_T(k_m) = \tilde{n}_T(k_1)^m$. On the other hand, cases of $N_k = 0$ and $M = 1$, i.e. a single constraint of either $\mathrm{Re}\,[\tilde{n}_T(k_1)]$ or $\mathrm{Im}\,[\tilde{n}_T(k_1)]$, give two solutions; see figure 2(a). Thus, we need at least two constraints ($M = 2$) for the unique inversion of a single-particle configuration.

We note that $\tilde{n}_T(k_1)$ is the minimal set of constraints for single-particle systems. This is because when $m > 1$, $\tilde{n}_T(k_m)$ is no longer a one-to-one function from $r_1 \in [0, L)$ onto the unit circle on $\mathbb{C}$, and thus cases with $N_k = m$ and $M = 2$ for $m > 1$ give m distinct solutions; see figure 2(b).

4.2. $N = 2$

Using graphical solutions, one can straightforwardly show a single constraint ($N_k = 0$ and $M = 1$) gives infinitely many solutions; see one of the solid or dashed lines in figure 3. However, figure 3 also immediately shows that the following equation ($N_k = 0$ and $M = 2$)

$$\tilde{n}_T(k_1) = \mathrm{e}^{-\mathrm{i}2\pi r_1/L} + \mathrm{e}^{-\mathrm{i}2\pi r_2/L}, \tag{8}$$

yields a unique solution aside from exchange of particle indices, as follows:

$$e^{-i2\pi r_1/L} = \frac{\tilde{n}_T(k_1)}{2}\left(1 \pm i\sqrt{\frac{4}{|\tilde{n}_T(k_1)|^2} - 1}\right) \tag{9}$$

$$e^{-i2\pi r_2/L} = \frac{\tilde{n}_T(k_1)}{2}\left(1 \mp i\sqrt{\frac{4}{|\tilde{n}_T(k_1)|^2} - 1}\right), \tag{10}$$

if $\tilde{n}_T(k_1) \neq 0$, or equivalently, $|R_1 - R_2| \neq 0.5L$. Otherwise, the periodic image of the target configuration becomes the integer lattice, and all of translated lattices are solutions of (8), i.e. there are infinitely many solutions, as shown in figure 3(c).

If the target configuration is the integer lattice, in order to obtain a unique solution, the collective coordinate at the first Bragg peak (i.e. $\tilde{n}_T(k_2)$) should be additionally specified, which corresponds to the cases with $N_k = 0$ and $M = 4$. Then, the unique solution is

$$e^{-i2\pi r_1/L} = \frac{1}{2}\left(\tilde{n}_T(k_1) \pm \sqrt{2\tilde{n}_T(k_2)^2 - \tilde{n}_T(k_1)^2}\right), \tag{11}$$

$$e^{-i2\pi r_2/L} = \frac{1}{2}\left(\tilde{n}_T(k_1) \mp \sqrt{2\tilde{n}_T(k_2)^2 - \tilde{n}_T(k_1)^2}\right). \tag{12}$$

This is because the collective coordinate at the first Bragg peak provides the center of mass of this lattice configuration.

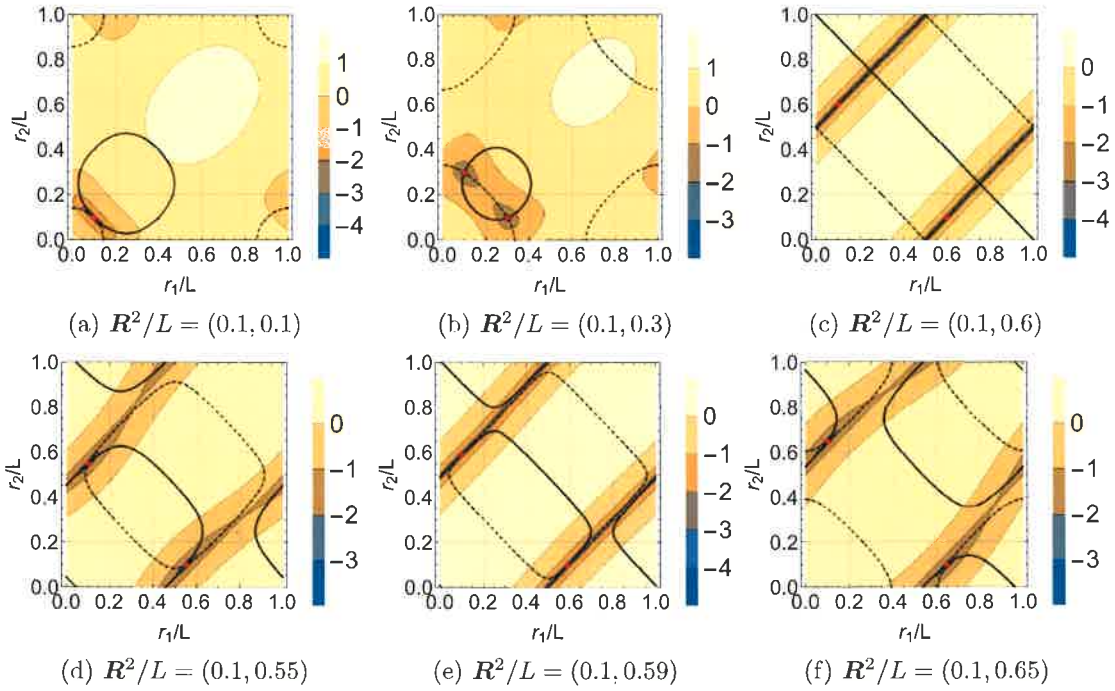
We note that the constraint $\tilde{n}_T(k_2)$ alone (i.e. $N_k = 1$ and $M = 2$) cannot be uniquely inverted into particle coordinates. It can be straightforwardly shown that there exist at least four distinct solutions, i.e. $(r_1, r_2) = \boldsymbol{a} + (R_1, R_2)$, where $\boldsymbol{a}/L = (0,0)$, $(0, 1/2)$, $(1/2, 0)$, and $(1/2, 1/2)$. By the same analysis, one can identify there are at least m^2 distinct solutions if only $\tilde{n}_T(2\pi m/L)$ is given. Therefore, we can conclude that for a two-particle configuration that is not the integer lattice, the minimal set of constraints for a unique solution is $\{\tilde{n}_T(k_1)\}$.

Remarks

(i) For a configuration of particle number $N > 1$, Fan *et al* [14] proved that $\tilde{n}(k_m) = 0$ for $m = 1, \cdots, \lfloor \frac{N}{2} \rfloor$ is a sufficient and necessary condition for the configuration to be the integer lattice or its translations. Thus, if one inverts collective coordinates at the $\lceil N/2 \rceil$ smallest wavenumbers of the integer lattice, its solutions are inevitably degenerated with a translational degree of freedom; see figure 3(c) for example.

4.3. $N = 3$

In the previous sections, we show that there is a unique solution in the inversion procedure with parameters $N_k = 0$ and $M = \lceil N/2 \rceil$, unless the target configuration is a



(a) $\boldsymbol{R}^2/L = (0.1, 0.1)$ (b) $\boldsymbol{R}^2/L = (0.1, 0.3)$ (c) $\boldsymbol{R}^2/L = (0.1, 0.6)$

(d) $\boldsymbol{R}^2/L = (0.1, 0.55)$ (e) $\boldsymbol{R}^2/L = (0.1, 0.59)$ (f) $\boldsymbol{R}^2/L = (0.1, 0.65)$

Figure 3. Graphical solutions of (8) for given respective target configurations. In each panel, black solid lines and dashed ones represent solutions of the 'real' and the 'imaginary' parts of (8), respectively. Contour plots depict potential energy landscape (i.e. $\log_{10}(\Phi(\boldsymbol{r}^2; \boldsymbol{R}^2))$) for each target configuration. Solutions (intersections of solid and dashed lines) are unique and identical to the target configuration (red dots), unless it is the integer lattice (i.e. $|R_1 - R_2| = L/2$) as shown in (c). Otherwise, there are infinitely many solutions, and one needs additional constraint $\tilde{n}_T(k_2)$ for unique solutions.

pathological case (i.e. either the integer lattice or its translations). Otherwise, there are infinitely many solutions. It implies that there would be a sudden transition in the number of distinct solutions varying with the type of target configurations. For this reason and simplicity in analysis, our target configurations are restricted here to perturbed lattices that can continuously interpolate between the integer lattice to Poisson configurations via the displacement parameter δ; see section 2.2.

For a perturbed lattice, its particle coordinates are described as $r_i = (i-1) + N\delta_i$ for $i = 1, \cdots, N$. Assuming weak perturbations (i.e. $|\delta_i| \ll 1$) for $N = 3$, collective-coordinate constraints can be approximated up to the second order of displacements;

$$\mathrm{Re}\left[\tilde{n}(k_m)\right] \approx \begin{cases} 3 - 2(m\pi)^2({\delta_1}^2 + {\delta_2}^2 + {\delta_3}^2), & m = 3i \\ \sqrt{3}m\pi(\delta_2 - \delta_3) + (m\pi)^2(-2{\delta_1}^2 + {\delta_2}^2 + {\delta_3}^2), & m = 3i+1 \\ -\sqrt{3}m\pi(\delta_2 - \delta_3) + (m\pi)^2(-2{\delta_1}^2 + {\delta_2}^2 + {\delta_3}^2), & m = 3i+2 \end{cases} \tag{13}$$

$$\mathrm{Im}\left[\tilde{n}(k_m)\right] \approx \begin{cases} 2m\pi\left(\delta_1 + \delta_2 + \delta_3\right), & m = 3i \\ m\pi\left(2\delta_1 - \delta_2 - \delta_3\right) + \sqrt{3}(m\pi)^2\left({\delta_2}^2 - {\delta_3}^2\right), & m = 3i+1 \\ m\pi\left(2\delta_1 - \delta_2 - \delta_3\right) - \sqrt{3}(m\pi)^2\left({\delta_2}^2 - {\delta_3}^2\right), & m = 3i+2 \end{cases} \tag{14}$$

where i represents non-negative integers.

For parameters $N_k = 0$ and $M = 3$ with the real condition (4) (or the imaginary one (5)), the quadratic approximations (13) and (14) yield at most two distinct solutions (A.1): the trivial solution ($\boldsymbol{r}^3 = \boldsymbol{R}^3$), and a nontrivial one ($\boldsymbol{r}^3 \neq \boldsymbol{R}^3$). This prediction is consistently observed in numerical results; see figure 4(a). Thus, the set of numerically distinct solutions abruptly changes from an uncountably many set into a finite one, as δ becomes nonzero. Figure 4(a) also shows that if δ increases, while the maximal number of numerically distinct solutions remains two, its occurrence decreases regardless of constraint conditions (4) and (5).

In numerical studies, it is important to know how results depend on the optimization algorithms and values of parameters, such as ϵ_E and ϵ_X. For this purpose, we investigate the energy distributions of numerical solutions obtained in the parameters of $N_k = 0$ and $M = 3$, and various conditions, as shown in figure 5. From figures 5(a) and (b), we see that given a target configuration, both trivial and nontrivial solutions have qualitatively similar energy profiles, regardless of the real (4) and the imaginary (5) conditions. Figure 5(c) demonstrates that the energy profiles of numerical solutions vary with optimization algorithms, but for a given algorithm both trivial and nontrivial solutions still have qualitatively similar energy profiles. Thus, a nontrivial solution cannot be eliminated by lowering the energy tolerance ϵ_E when $N = M = 3$. In the rest of this paper, we mainly use the BFGS and MINOP algorithms because the solutions obtained via these algorithms tend to have lower numerical errors than those via the steepest descent method.

For parameters $N_k = 0$ and $M = 4$, a unique solution can be obtained. This also can be deduced from the observation in the cases with $N_k = 0$ and $M = 3$ that given a target configuration, nontrivial solutions, respectively obtained by the real (4) and the imaginary (5) conditions, are numerically distinct; see figure 4(b). Thus, the common solution from two conditions (4) and (5) should be identical to the target. The unique solution also can be obtained from the quadratic approximations (13) and (14) as follows:

$$\delta_1 = \frac{1}{12\pi}\left[\frac{6\mathrm{Im}\left[2\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]} + \mathrm{Im}\left[4\,\tilde{n}_T(k_1) + \tilde{n}_T(k_2)\right]\right] \tag{15}$$

$$\delta_2 = \frac{1}{12\pi}\left[\frac{6\mathrm{Im}\left[2\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]} - \frac{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{\sqrt{3}}\right] \tag{16}$$

$$\delta_3 = \frac{1}{12\pi}\left[\frac{6\mathrm{Im}\left[2\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]} + \frac{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{\sqrt{3}}\right], \tag{17}$$

and thus the minimal set for three-particle systems is (both real and imaginary parts of) collective coordinates at the two smallest wavenumbers.

Remarks

(i) For parameters $N_k = 0$, $M = 3$, and the real condition (4), the quadratic approximations (13) and (14) give two exact solutions (A.1). While one of the solutions is the same as the target configuration up to some numerical errors, another solution cannot precisely predict the nontrivial solution partly because the nontrivial one is not a perturbed lattice with small displacements.

(ii) For parameters $N_k = 1$ and $M = 4$, a unique solution is obtained; see (A.3)–(A.5).

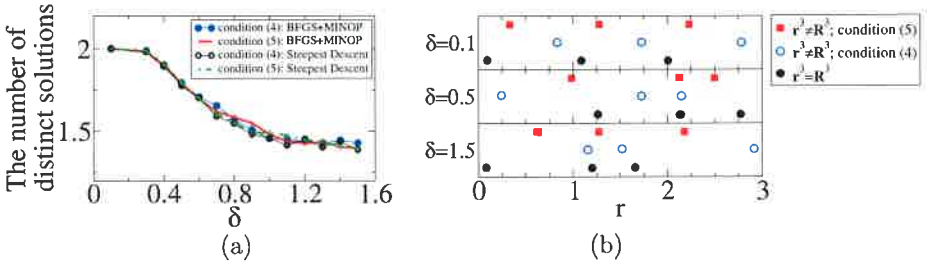


Figure 4. Numerical results of the inversion procedure for three-particle perturbed lattices in cases with $N_k = 0$ and $M = 3$. (a) The average number of distinct solutions per a target configuration. Two different optimization algorithms (BFGS + MINOP and the steepest descent) and two constraint conditions (the real (4) and the imaginary (5) ones) are used for comparison with the energy tolerance $\epsilon_E = 10^{-29}$. For any target configuration, the number of distinct solutions is at most two, but the average can vary with the target configurations. (b) Examples of nontrivial solutions for a given target perturbed lattice with various displacements δ. Nontrivial solutions by the real (4) or the imaginary (5) conditions, respectively, are different from each other, and are not translations of the target.

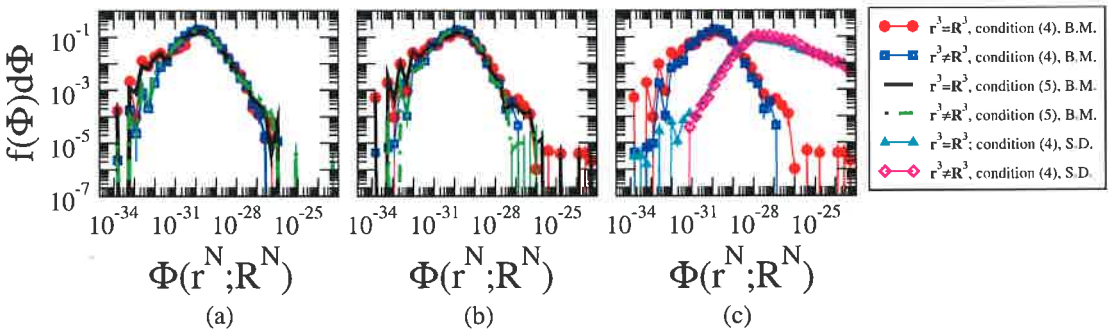


Figure 5. Log–log plots of histograms for energy distributions of numerically distinct solutions $\{\boldsymbol{r}^3\}$ of a three-particle target configuration $\boldsymbol{R}^3$ for parameters $N_k = 0$, $M = 3$, and $\epsilon_E = 10^{-20}$. Given a target configuration, there are at most two distinct solutions; a trivial solution and a nontrivial one. (a) and (b) Results from two constraint types (i.e. the real condition (4) and the imaginary condition (5)) are compared for two different types of target configurations: (a) perturbed lattices with $\delta = 0.1$ and (b) Poisson configurations. Here, BFGS+MINOP (B.M.) algorithms are used. (c) For Poissonian target configurations, we compare results from two different optimization algorithms: B.M., and steepest descent (S.D.). Here, the real condition (4) is considered.

5. Results for $N > 3$

Here, we numerically investigate the properties of the inversion procedure from collective coordinates, such as proper values of the tolerances ϵ_E and ϵ_X. For this purpose, we obtain distributions of energy $\Phi(\boldsymbol{r}^N; \boldsymbol{R}^N)$ for numerically distinct solutions, as we did in figure 5. Our results, shown in figures 6 and 7, demonstrate that the energy distributions sensitively depend on the number of skipped collective-coordinate constraints N_k as well as target configurations and the particle number N.

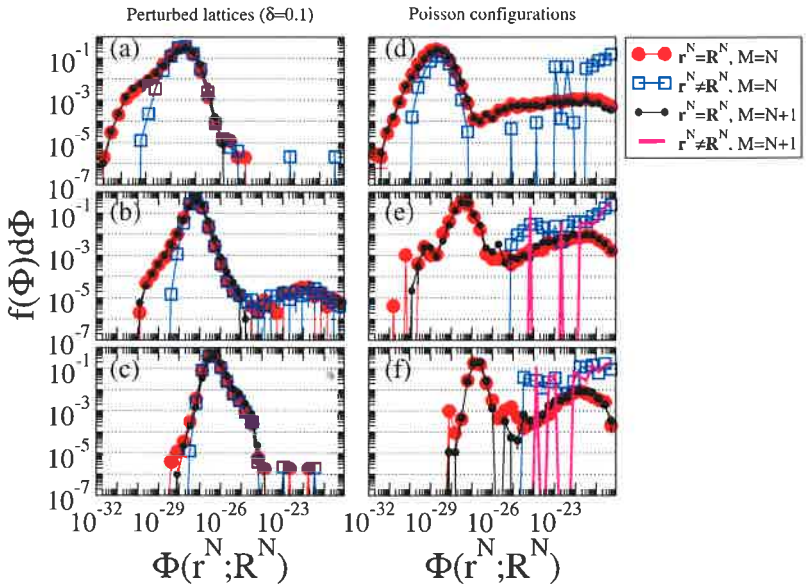


Figure 6. Log–log plots of histograms for energy distribution of numerically distinct solutions $\{\boldsymbol{r}^N\}$ for odd-number system sizes: $N = 9$ (a) and (d), 19 (b) and (e), and 29 ((c) and (f)). Using the real condition (4) condition and parameters $N_k = 0$ and $\epsilon_E = 10^{-20}$, two types of target configurations are considered: ((a)–(c)) perturbed lattices with $\delta = 0.1$ and (d)–(f) Poisson configurations. When $M = N$, while a target perturbed lattice has a single nontrivial solution ($\boldsymbol{r}^N \neq \boldsymbol{R}^N$), whose occurrence rate is similar to that of trivial ones, a Poissonian target mainly has the trivial solution but occasionally can have multiple nontrivial solutions. When $M = N + 1$ is an even number, while there is a unique solution for perturbed lattices, there can be more than one solution for a Poisson target configuration in relatively lower occurrence rates. Even in the latter case, however, the nontrivial solutions can be eliminated by lowering the tolerance ϵ_E around 10^{-25}.

At first, we consider the cases with $N_k = 0$ (figure 6). When there are even-number N of particles, $M \geqslant N$ constraints can give unique solutions for both types of target configurations: perturbed lattices and Poisson point distribution configurations. If N is an odd number, however, $M = N$ constraints no longer ensure unique solutions. When perturbed lattices are the target configurations (figures 6(a)–(c)) and $M = N$ constraints are considered, the energy $\Phi(\boldsymbol{r}^N; \boldsymbol{R}^N)$ always has two global minima, which correspond to the trivial solution ($\boldsymbol{r}^N = \boldsymbol{R}^N$) and a nontrivial one ($\boldsymbol{r}^N \neq \boldsymbol{R}^N$), respectively. On the other hand, the energy $\Phi(\boldsymbol{r}^N; \boldsymbol{R}^N)$ of a Poissonian target configuration (figures 6(d)–(f)) mostly has a single minimum that is identical to the target ($\boldsymbol{r}^N = \boldsymbol{R}^N$) but occasionally has more than two nontrivial solutions. Given parameters $N_k = 0$ and $M = N + 1$, while when the target is a perturbed lattice the inversion procedure gives a unique solution, when the target is a Poisson configuration this procedure may give multiple solutions. However, since the nontrivial solutions in the latter case have qualitatively different energy profiles from the trivial solution (see figures 6(d)–(f)), the nontrivial solutions can be eliminated by lowering the tolerance ϵ_E to a proper level. Thus, when N is an odd number, $M = N + 1$ constraints are required for the unique determination.

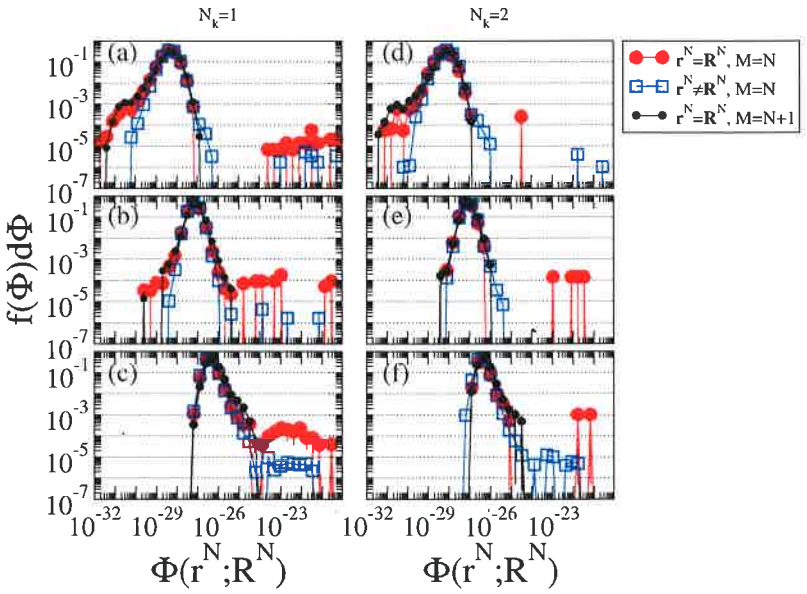


Figure 7. Log–log plots of histograms for energy distribution of numerically distinct solutions $\{\boldsymbol{r}^N\}$ for $N_k > 0$ and odd-number system sizes: $N = 9$ (a) and (d), 19 (b) and (e), and 29 (c) and (f). Considering perturbed lattices with $\delta = 0.1$ as the target configurations, we search solution configurations under the real condition (4) and the tolerance $\epsilon_E = 10^{-20}$, and via the BFGS + MINOP algorithms. We note that there is no nontrivial solution with $\Phi(\boldsymbol{r}^N; \boldsymbol{R}^N) < 10^{-20}$ if $N_k > 0$ and $M = N + 1$.

When first few collective coordinates are skipped ($N_k > 0$), there is no advantage of even-number particles, i.e. one cannot determine unique solutions with $M = N$ successive collective-coordinate constraints when N is an even number. Figure 7 shows the histograms for energies of numerical solutions obtained in the inversion procedure with an odd-number particles and $N_k > 0$. In figure 7, we note that for $M = N$ constraints there can be more than one nontrivial solutions whose energy profiles are similar to that of the trivial solutions. However, $M = N + 1$ constraints allow us to find the trivial solutions without any nontrivial one.

In general, as the system size N increases, both trivial and nontrivial solutions tend to have higher energies, i.e. larger numerical errors. Moreover, for parameters $N_k = 0$ and $M = N$, although for smaller systems the distribution of trivial and nontrivial solutions have tails in the low-energy regime (figures 6(a) and (d)), for larger systems the tails are shifted to the high-energy regime (figures 6(c) and (f)); see also figure 7 for cases with $N_k > 0$. This observation implies that it becomes less probable to obtain numerical solutions, whether they are trivial or not, as the particle number N increases, or the energy tolerance ϵ_E is lowered.

The average number of numerically distinct solutions, obtained in the inversion procedure, is shown in figure 8. This figure clearly demonstrates that for Poissonian targets (figures 8(d)–(f)) the two curves ($M = N$ and $N + 1$) collapse into a single line as N increases, and thus $\min M \to N$ as N increases. On the other hand, these two curves are clearly separated for perturbed lattices (figures 8(a)–(c)), and thus $\min M$ is determined by the cases where perturbed lattices are the target configurations. Figure 9

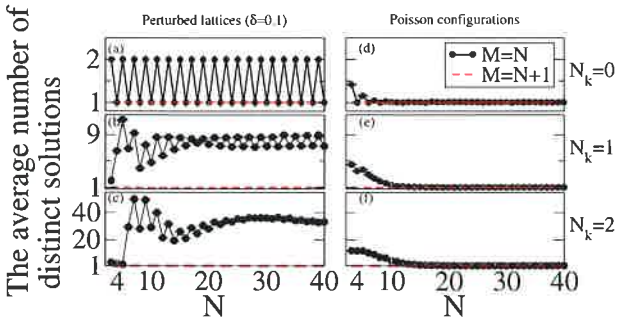


Figure 8. Numerical results for the average number of numerically distinct solutions per a target configuration of particle number N with various values of N_k. Using the real condition (4) and BFGS + MINOP algorithms, we consider two types of target configurations: (a)–(c) perturbed lattices with $\delta = 0.1$ and the tolerance $\epsilon_E = 10^{-20}$, and (d)–(f) Poisson configurations with $\epsilon_E = 10^{-25}$. When $N_k = 0$, both types of target configurations require $M = N$ constraints for an even-number N, and $M = N + 1$ is the minimal for an odd-number N: the minimal number of M is $2\lceil N/2 \rceil$. If $N_k > 0$, for both types of target configurations, the minimal number of constraints becomes $M = N + 1$.

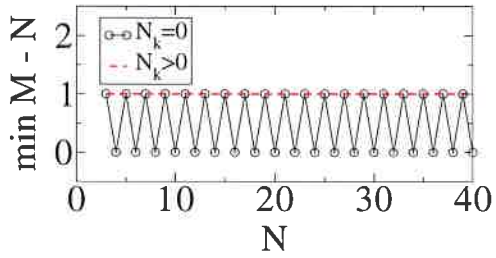


Figure 9. The minimal number of successive collective-coordinate constraints min M as a function of particle number N for various N_k.

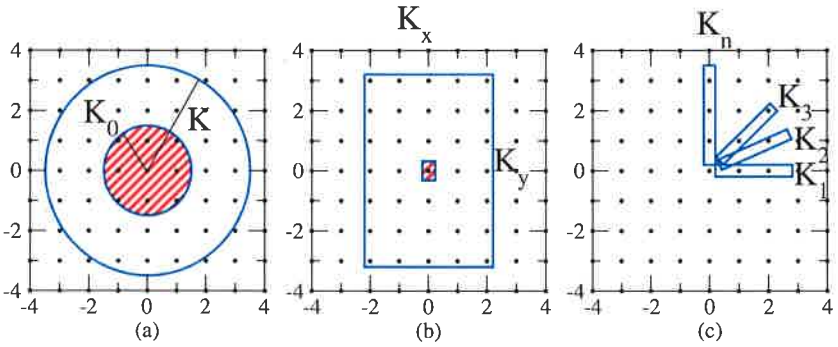


Figure 10. Schematics of some possible ways to select collective-coordinate constraints in the two-dimensional Fourier space. Collective coordinates are specified at wavevectors inside (a) an annular region of outer radius K and inner radius K_0 (see figure 1), (b) a rectangular region of width K_x and height K_y, and (c) n mutually non-parallel strips whose lengths are K_i, $i = 1, \cdots, n$. We note that the red-shaded region is excluded.

summarizes the results from analytic investigation into small systems (section 4) and numerical studies on larger systems (section 5). One can uniquely determine particle coordinates from collective coordinates at the $\lceil \frac{N}{2} \rceil$ smallest wavenumbers, i.e. parameters of $N_k = 0$ and $M = 2\lceil \frac{N}{2} \rceil$, by properly selecting ϵ_E. On the other hand, if $N_k > 0$, one requires $M = N + 1$ successive collective-coordinate constraints to uniquely determine particle coordinates. Therefore, when both cases are considered, the minimal set of collective-coordinate constraints are collective coordinates at the $\lceil \frac{N}{2} \rceil$ smallest wavenumbers.

6. Conclusions and discussions

In this work, we have investigated the minimal set of collective-coordinate constraints as a function of the particle number N to uniquely determine the progenitor particle coordinates in one dimension. We also considered how the minimal collective-coordinate constraints depend on constraint types (the real (4) and imaginary (5) conditions) and types of target configurations, i.e. perturbed lattices and Poisson point distribution configurations. As shown in figure 9, the minimal set of constraints are collective coordinates at the $\lceil \frac{N}{2} \rceil$ smallest wavenumbers: it corresponds to the parameters of $N_k = 0$ and $M = 2\lceil N/2 \rceil$. In other words, the removed number of degrees of freedom in the solution space will vary with each collective-coordinate constraint, and the real and the imaginary parts of a collective coordinate are not completely independent.

For this result to accommodate the pathological case, i.e. the integer lattice, one needs to regard all of its translations to be equivalent. As we noted in section 4.2, this is because translations of the integer lattice cannot be distinguished in terms of $\tilde{n}_T(k_m)$ for $m = 1, \cdots, \lceil N/2 \rceil$, since their collective coordinates are identically zero, except at the Bragg peaks, i.e. $k = 2\pi, 4\pi, \cdots$. An additional constraint $\tilde{n}_T(k_N) \equiv \tilde{n}_T(2\pi)$ at the first Bragg peak is necessary to remove the translational degree of freedom. However, we note that non-Bravais lattices are not pathological cases because their lattice constants are larger than one, and thus their first Bragg peaks should appear within the range of $|k| \leqslant \pi$.

It is useful to compare this conclusion with the result of Fan *et al* [14]. These authors proved that for a one-dimensional system one needs its collective coordinates at the $\lfloor \frac{N}{2} \rfloor$ smallest wavenumbers as well as the center of mass in order to determine all of its collective coordinates; see appendix B for the detailed summary. In the same context, our investigation shows that if the center of mass is unknown, one needs collective coordinates at the $\lceil \frac{N}{2} \rceil$ smallest wavenumbers. Moreover, when there are an even-number of particles, the knowledge of the center of mass does not reduce the necessary information.

While the present work focused on 1D systems for simplicity, it is useful to discuss implications of our results for the inversion problem in higher-dimensional systems. Unlike 1D systems, higher-dimensional systems allow many different ways to select collective-coordinate constraints; see figure 10. Consider here the case (c) where selected wavevectors form n nonparallel strips orienting toward the origin. Based on our present results, if the ith strip has a slope $s_i = n_i/m_i$, where n_i and m_i are integers and coprime,

and includes the smallest $\lceil N/2 \rceil$ wavevectors, then one can uniquely determine values of the coordinates on a line, i.e. $m_i x_j + n_i y_j$ for $j = 1, \cdots, N$. Thus, by using two perpendicular strips that include a total of $2\lceil N/2 \rceil$ collective-coordinate constraints, one can 'separately' determine the x and y coordinates of particle positions. In order to determine the pairing between the x and y coordinates, one needs collective-coordinate constraints along additional strips in the Fourier space, as shown in figure 10(c). Therefore, in this scheme at least $3\lceil N/2 \rceil$ collective-coordinate constraints are required.

It is interesting to compare collective coordinates with Fourier components in the discrete Fourier transform (DFT). While a Fourier component X_k in the DFT is a linear function of a complex sequence $\{x_n\}_{n=0}^{\mathcal{N}-1}$, a collective coordinate $\tilde{n}(k_m)$ is a nonlinear function of particle coordinates $\{R_j\}_{j=1}^{N}$. In both cases, wavenumbers are restricted to be equally spaced due to the periodic boundary conditions in direct spaces. On the contrary, the direct spaces are different in the two cases in that while the direct spaces in the DFT are digitized into $\mathcal{N}$ pixels, those in collective coordinates are continuous. If one discretizes the space of a point configuration with $\mathcal{N}$ pixels of width Δx, the configuration can be described by a real-valued sequence $\{x_n\}$, where x_n represents the number of particles in the nth pixel. Then, this conversion can be straightforwardly written as follows:

Particle coordinates:	$\{R_i\}_{i=1}^{N} \subset \mathbb{R}$	$\Rightarrow$	$\{x_n\}_{n=0}^{\mathcal{N}-1} \subset \mathbb{N} \cup \{0\}$
Collective coordinates:	$\tilde{n}(k_m) = \sum_{i=1}^{N} \exp(-\mathrm{i} k_m x_i)$	$\Rightarrow$	$X_m = \sum_{j=0}^{\mathcal{N}-1} x_j \exp\left[-\mathrm{i}\frac{2\pi m}{\mathcal{N}\Delta x}(j\Delta x)\right]$.

Thus, the mth collective coordinate $\tilde{n}(k_m)$ of a point configuration corresponds to the mth Fourier component X_m of its digitized version. From this relationship, one can surmise that the inverse DFT with the first $\mathcal{N}/2$ collective coordinates will give a discretized point configuration with a position precision Δx. In other words, one needs around 10^7 Fourier components to achieve $\Delta x \sim \mathcal{O}(10^{-7})$, which is a typical error in our solution configurations.

In the present work, we focused on the search for the minimal set of constraints, rather than computational costs. Our inversion procedure is intuitive and provides easy-to-estimate numerical errors in solutions (i.e. energy $\Phi(\boldsymbol{r}^N; \boldsymbol{R}^N)$), but this computational method is inefficient for large systems. For instance, as system size N increases, the computation cost grows at least in the order of N^2. Furthermore, since this method tends to have larger numerical errors in solution configurations as N increases (see figures 6 and 7), it becomes more likely to fail to find any solution with a given value of the energy tolerance ϵ_E. The failure rate becomes especially much higher when a target is more complicated. Therefore, for future studies, it would be important to develop more efficient computational procedures to invert collective coordinates into particle coordinates.

Acknowledgment

This work was supported partially by the National Science Foundation under Grant No. CBET-1701843.

Appendix A. Approximate solutions of equations (13) and (14)

For parameters $N_k = 0$ and $M = 3$, and the real condition (4), from (13) and (14), one can find two solutions as follows:

$$\begin{aligned}
\delta_1 &\approx \frac{(-18\mathrm{Im}\left[\tilde{n}_T(k_1)\right] \pm D)}{\pi(\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right])\left(\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right] - 12\right)} \\
\delta_2 &\approx -\frac{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{12\sqrt{3}\pi} + \frac{6}{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right] - 6}\left(\frac{\mathrm{Im}\left[\tilde{n}_T(k_1)\right]}{2\pi} - \delta_1\right) \\
\delta_3 &\approx \frac{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]}{12\sqrt{3}\pi} + \frac{6}{\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right] - 6}\left(\frac{\mathrm{Im}\left[\tilde{n}_T(k_1)\right]}{2\pi} - \delta_1\right),
\end{aligned} \tag{A.1}$$

where the discriminant D is written as

$$\begin{aligned}
D \equiv &\frac{1}{12\sqrt{3}}\left(\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right] - 6\right)\Big[\left(\left(\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right] - 6\right)^2 - 36\right) \\
&\times\left(\mathrm{Re}\left[4\,\tilde{n}_T(k_1) - \tilde{n}_T(k_2)\right]^2 - 36\mathrm{Re}\left[2\,\tilde{n}_T(k_1) + \tilde{n}_T(k_2)\right]\right) + 3888\mathrm{Im}\left[\tilde{n}_T(k_1)\right]^2\Big]^{1/2}.
\end{aligned} \tag{A.2}$$

Here, a trivial solution is obtained from (A.1) when a minus sign is taken in δ_1. Otherwise, (A.1) become a nontrivial solution.

For parameters $N_k = 1$ and $M = 4$, (13) and (14) give a single solution:

$$\delta_2 = -\frac{\delta_1}{2} + \frac{\mathrm{Im}\left[\tilde{n}_T(k_3)\right]}{12\pi} + \left[\sqrt{3}\pi{\delta_1}^2 + \frac{1}{4\sqrt{3}\pi}\left(\mathrm{Re}\left[\tilde{n}_T(k_2) + \frac{2}{9}\,\tilde{n}_T(k_3)\right] - \frac{2}{3}\right)\right] \tag{A.3}$$

$$\delta_3 = -\frac{\delta_1}{2} + \frac{\mathrm{Im}\left[\tilde{n}_T(k_3)\right]}{12\pi} - \left[\sqrt{3}\pi{\delta_1}^2 + \frac{1}{4\sqrt{3}\pi}\left(\mathrm{Re}\left[\tilde{n}_T(k_2) + \frac{2}{9}\,\tilde{n}_T(k_3)\right] - \frac{2}{3}\right)\right], \tag{A.4}$$

where δ_1 is determined by the following cubic equation:

$$\begin{aligned}
{\delta_1}^3 - \frac{\mathrm{Im}\left[\tilde{n}_T(k_3)\right]}{6\pi}{\delta_1}^2 - \frac{1}{12\pi^2}\left(\mathrm{Re}\left[\tilde{n}_T(k_2) + \frac{2}{9}\,\tilde{n}_T(k_3)\right] - \frac{11}{3}\right)\delta_1 \\
+ \frac{1}{72\pi^3}\left[\mathrm{Im}\left[\tilde{n}_T(k_3)\right]\left(\mathrm{Re}\left[\tilde{n}_T(k_2) + \frac{2}{9}\,\tilde{n}_T(k_3)\right] - \frac{5}{3}\right) - 3\mathrm{Im}\left[\tilde{n}_T(k_2)\right]\right] = 0,
\end{aligned} \tag{A.5}$$

which has a single real root for realizable and nonzero collective-coordinate constraints.

Appendix B. The uniqueness of solutions for the inversion problem

Using the generating function argument [14], one can prove that there is the unique configuration to satisfy N prescribed collective coordinates. Let us define a generating function as

$$f(z) \equiv \sum_{m=1}^{\infty} \frac{\tilde{n}(k_m)}{m} z^m, \tag{B.1}$$

which is well-defined for $|z| < 1$ because $|\tilde{n}(k_m)|$ is bounded. Using the definition (1) and power series expansion of the log function [$\ln(1-z) = \sum_{n=1}^{\infty} z^n/n$ for $|z| < 1$],

$$f(z) = \sum_{n=1}^{\infty}\left(\sum_{j=1}^{N} \mathrm{e}^{-\mathrm{i}nx_j}\right)\frac{z^n}{n} = \sum_{j=1}^{N}\sum_{n=1}^{\infty}\frac{\left(z\mathrm{e}^{-\mathrm{i}x_j}\right)^n}{n} = \sum_{j=1}^{N} -\ln\left(1 - z\mathrm{e}^{-\mathrm{i}x_j}\right)$$
$$= -\ln\left[\prod_{j=1}^{N}\left(1 - z\mathrm{e}^{-\mathrm{i}x_j}\right)\right]. \tag{B.2}$$

Since the term inside square brackets of logarithm is a polynomial of order N, $\exp[f(z)]$ also should be a polynomial of order N.

$$\prod_{j=1}^{N}\left(1 - z\mathrm{e}^{-\mathrm{i}x_j}\right) = \exp(-f(z)) = \mathbb{P}_N \exp(-f(z)) = \mathbb{P}_N \exp(-\mathbb{P}_N f(z))$$
$$= \mathbb{P}_N \exp\left(-\sum_{m=1}^{N}\frac{\tilde{n}(k_m)}{m}z^m\right), \tag{B.3}$$

where $\mathbb{P}_N$ represents a projection to a degree N polynomial of z.

By substituting (B.3) into (B.2) and doing further analysis, Fan *et al* [14] derived the following identity:

$$\sum_{m=1}^{N}\frac{\tilde{n}(k_m)}{m}z^m = -\ln\left[\mathbb{P}_{\lfloor\frac{N}{2}\rfloor}\exp\left(-\sum_{m=1}^{\lfloor N/2\rfloor}\frac{\tilde{n}(k_m)}{m}z^m\right) - \omega z^N \mathbb{P}_{-\lfloor\frac{N}{2}\rfloor}\exp\left(\sum_{m=1}^{\lfloor N/2\rfloor}\frac{\tilde{n}(-k_m)}{m}z^{-m}\right)\right], \tag{B.4}$$

where $\omega \equiv \exp\left(-\mathrm{i}2\pi\sum_{n=1}^{N}x_n\right)$, and $\lfloor x\rfloor$ is the floor function of x. Since $\tilde{n}(k_m) = \tilde{n}(-k_m)^*$, if collective coordinates at the $\lfloor\frac{N}{2}\rfloor$ smallest wavenumbers and the center of mass are known, in principle one can determine collective coordinates at other wavenumbers. In other words, there is a unique point configuration that satisfy these conditions.

References

[1] Feynman R 1954 *Phys. Rev.* **94** 262–77
[2] Pines D and Bohm D 1952 *Phys. Rev.* **85** 338–53
[3] Percus J and Yevick G 1958 *Phys. Rev.* **110** 1–13
[4] Torquato S and Stillinger F 2003 *Phys. Rev.* E **68** 041113
[5] Torquato S, Zhang G and Stillinger F 2015 *Phys. Rev.* X **5** 021020
[6] Edwards S and Schwartz M 2003 *J. Stat. Phys.* **110** 497–502
[7] Ashcroft N and Mermin N 1976 *Solid State Physics* (*Davis Drive* vol 10) (Belmont: Brooks/Cole, Cengage Learning)
[8] Kam Z 1977 *Macromolecules* **10** 927–34
[9] Pähler A, Smith J and Hendrickson W 1990 *Acta Crystallogr.* A **46** 537–40
[10] Hendrickson W and Ogata C 1997 *Phase Determination from Multiwavelength Anomalous Diffraction Measurements* vol 276 (New York: Academic) pp 494–523
[11] Elser V 2003 *J. Opt. Soc. Am.* A **20** 40–55
[12] Shechtman Y, Eldar Y, Cohen O, Chapman H, Miao J and Segev M 2015 *IEEE Signal Process. Mag.* **32** 87–109
[13] Harrison R 1993 *J. Opt. Soc. Am.* A **10** 1046–55
[14] Fan Y, Percus J, Stillinger D and Stillinger F 1991 *Phys. Rev.* A **44** 2394–402
[15] Zhang G, Stillinger F and Torquato S 2015 *Phys. Rev.* E **92** 022119
[16] Zhang G, Stillinger F and Torquato S 2015 *Phys. Rev.* E **92** 022120
[17] Uche O, Stillinger F and Torquato S 2004 *Phys. Rev.* E **70** 046122

[18] Uche O, Torquato S and Stillinger F 2006 *Phys. Rev.* E **74** 031104
[19] Zhang G, Stillinger F and Torquato S 2017 *Phys. Rev.* E **96** 042146
[20] Florescu M, Torquato S and Steinhardt P 2009 *Proc. Natl Acad. Sci. USA* **106** 20658–63
[21] Zhang G, Stillinger F and Torquato S 2016 *J. Chem. Phys.* **145** 244109
[22] Zhang G, Stillinger F and Torquato S 2017 *Soft Matter* **13** 6197–207
[23] Chen D and Torquato S 2018 *Acta Mater.* **142** 152–61
[24] Leseur O, Pierrat R and Carminati R 2016 *Optica* **3** 763–7
[25] Froufe-Pérez L, Engel M, Sáenz J and Scheffold F 2017 *Proc. Natl Acad. Sci. USA* **114** 9570–4
[26] Gkantzounis G, Amoah T and Florescu M 2017 *Phys. Rev.* B **95** 094120
[27] Wu B, Sheng X and Hao Y 2017 *PLoS One* **12** e0185921
[28] Torquato S 2018 *Phys. Rep.* **745** 1–95
[29] Gabrielli A 2004 *Phys. Rev.* E **70** 066131
[30] Kuna T, Lebowitz J and Speer E 2007 *J. Stat. Phys.* **129** 417–39
[31] Welberry T, Miller G and Carroll C 1980 *Acta Crystallogr.* A **36** 921–9
[32] Batten R, Stillinger F and Torquato S 2008 *J. Appl. Phys.* **104** 033504
[33] Zhang G, Stillinger F and Torquato S 2016 *Sci. Rep.* **6** 36963
[34] Nocedal J 1980 *Math. Comput.* **35** 773–82
[35] Johnson S 2014 NLopt nonlinear-optimization package, Version 2.4.2 (http://ab-initio.mit.edu/nlopt)
[36] Dennis J and Mei H 1979 *J. Optim. Theory Appl.* **28** 453–82
[37] Galassi M, Davies J, Theiler J, Gough B, Jungman G, Booth M and Rossi F 2009 *GNU Scientific Library Reference Manual* 3rd edn (United Kingdom: Network Theory Ltd.)