STEIN'S METHOD AND APPROXIMATING THE MULTIDIMENSIONAL QUANTUM HARMONIC OSCILLATOR

IAN W. MCKEAGUE,* *Columbia University*

YVIK SWAN (iD),** *Université libre de Bruxelles*

Abstract

Stein's method is used to study discrete representations of multidimensional distributions that arise as approximations of states of quantum harmonic oscillators. These representations model how quantum effects result from the interaction of finitely many classical 'worlds', with the role of sample size played by the number of worlds. Each approximation arises as the ground state of a Hamiltonian involving a particular interworld potential function. Our approach, framed in terms of spherical coordinates, provides the rate of convergence of the discrete approximation to the ground state in terms of Wasserstein distance. Applying a novel Stein's method technique to the radial component of the ground state solution, the fastest rate of convergence to the ground state is found to occur in three dimensions.

Keywords: Coupling; Hamiltonian; point configurations; Wasserstein distance

2020 Mathematics Subject Classification: Primary 60F05
Secondary 81Q65

1. Introduction

The many interacting worlds (MIW) approach to quantum mechanics [9] posits a Hamiltonian for a one-dimensional harmonic oscillator of the form $H_1(\mathbf{x}, \mathbf{p}) = E(\mathbf{p}) + V_1(\mathbf{x}) + U_1(\mathbf{x})$, where the locations of particles having identical mass $m > 0$ in N worlds are specified in dimensionless coordinates as $\mathbf{x} = (x_1, \ldots, x_N)$ with $x_1 > x_2 > \cdots > x_N$, and their momenta by $\mathbf{p} = (p_1, \ldots, p_N)$. Here, $E(\mathbf{p}) = \sum_{n=1}^{N} p_n^2/2m$ is the kinetic energy, $V_1(\mathbf{x}) = \sum_{n=1}^{N} x_n^2$ is the potential energy (for the parabolic trap), and

$$U_1(\mathbf{x}) = \sum_{n=1}^{N} \left(\frac{1}{x_{n+1} - x_n} - \frac{1}{x_n - x_{n-1}} \right)^2$$

is called the 'interworld' potential, where $x_0 = \infty$ and $x_{N+1} = -\infty$. The dimensionless coordinate x_n is related to the actual position x'_n of the particle by $x_n = \sqrt{2m\omega/\hbar}\, x'_n$, where $\hbar$ is the reduced Planck constant and ω is the angular frequency of the quantum harmonic oscillator. We restrict our attention to dimensionless coordinates throughout the paper.

In the MIW approach, the state of the system evolves deterministically according to Hamilton's equations of motion. The (stationary) ground state **x** is the (unique) particle configuration that minimizes the above Hamiltonian, with no particle movement involved (i.e. $p_n = 0$ for $n = 1, \ldots, N$). The resulting configuration x_n derived in [9] is symmetric about the origin and satisfies the nonlinear recursion equation

$$x_{n+1} = x_n - \frac{1}{x_1 + \cdots + x_n}. \tag{1}$$

It was shown in [13] that the empirical distribution of this ground-state solution, putting mass $1/N$ on each point in $\{x_n, n = 1, \ldots, N\}$, tends to the classical standard Gaussian quantum ground state as $N \to \infty$, and it was conjectured that the optimal rate of convergence in Wasserstein distance is of order $\sqrt{\log N}/N$; this conjecture was recently proved in [2].

An excited state of a quantum system is any state that has a higher energy than the ground state. The first excited state of the one-dimensional quantum harmonic oscillator can be represented in the MIW setting via the extension of the interworld potential U_1 introduced in [14], and leads to the two-sided Maxwell distribution as the limit (agreeing with quantum theory). Nonlocality was studied in [7] by introducing other extensions of the MIW interworld potential for the first excited state using higher-order smoothing methods. There is a close connection between classical Gibbs distributions and quantum energy states, as recently elucidated in [3].

In this article we study the question of how the MIW approach can be extended to general d-dimensional settings, for both ground states and excited states. The case $d = 2$ was examined in [10], which proposed using Delaunay triangulations and Voronoi tesselations of point configurations to extend the notion of the interworld potential and developed a numerical algorithm to estimate the resulting ground-state configuration, but the question of whether it is possible to establish an asymptotically valid approximation of such a solution to the classical quantum harmonic oscillator ground-state solution was not addressed.

Our contribution is twofold: (i) we introduce new interworld potentials that apply in general d-dimensional settings and that lead to tractable ground-state solutions (as well as some excited-state solutions) for the corresponding MIW Hamiltonian; and (ii) we provide a new version of the discrete density approach to Stein's method that furnishes upper bounds on the convergence rate of the radial components of these ground-state solutions. Evaluating these bounds numerically leads to optimal rates of convergence (in terms of Wasserstein distance) of the full MIW configurations that apply to general multidimensional settings. In contrast to [13, 14], which rely on the coupling version of Stein's method, the present approach leads to optimal rates. In the Supplementary Material we explore the types of upper bounds that can be obtained using the coupling approach in the multidimensional case.

The paper is organized as follows. Section 2 develops the proposed interworld potentials, first for the two- and three-dimensional cases, then for the general d-dimensional setting, including some that apply to excited states. The resulting MIW ground-state solutions are described in terms of their spherical coordinates, and plots of the solutions are provided for $d = 2$ and 3. An upper bound on the Wasserstein distance between two probability measures that have independent radial and directional components is provided at the end of Section 2. Section 3 restricts attention to the radial component and develops the discrete density approach to Stein's method mentioned above. In Section 4 we wrap up by providing optimal rates for full d-dimensional ground-state solutions. The Supplementary Material includes discussion of the coupling approach mentioned above, along with computer code.

2. Extending interworld potentials to higher dimensions

In this section we introduce interworld potentials that lead to satisfactory MIW approximations to the ground states and excited states of the d-dimensional isotropic quantum harmonic oscillator. First we consider the two-dimensional case.

2.1. Two-dimensional case

When $d = 2$, the idea is to couch the problem in terms of polar coordinates, which allows the interworld potential to adapt to the basic geometry of the problem, specifically via a separation of its angular (directional) and radial components. To that end, a configuration $\mathbf{x}$ of points in $\mathbb{R}^2$ that excludes the origin and is not confined to a single line passing through the origin is specified in terms of (signed) polar coordinates as $\{(r_{nj}, \theta_j) : n = 1, \ldots, N_j, j = 1, \ldots, M\}$, where there are $M = M(\mathbf{x}) \geq 2$ (distinct) angular coordinate values satisfying $0 \leq \theta_1 < \theta_2 < \cdots < \theta_M < \pi$. The radial coordinates r_{nj} in direction θ_j in this representation are signed (can be negative as well as positive) and satisfy $r_{1j} > r_{2j} > \cdots > r_{N_j j}$; a point (r, θ) with $r < 0$ is understood to correspond to the point with polar coordinates $(|r|, \theta + \pi)$. The total number of points is $N = N_1 + \cdots + N_M$.

We propose the following ansatz for the interworld potential:

$$U_2(\mathbf{x}) = 4 \sum_{j=1}^{M} \sum_{n=1}^{N_j} \left[\frac{1}{R_2(r_{n+1,j}) - R_2(r_{nj})} - \frac{1}{R_2(r_{nj}) - R_2(r_{n-1,j})} \right]^2 r_{nj}^2$$

$$+ \pi \sum_{j \sim k} \frac{1}{|\theta_j - \theta_k|_\pi} + \frac{N^2}{M^2} L \left(\sum_{j \sim k} |N_j - N_k| \right), \tag{2}$$

where the functions $R_d(r) = |r|^d \, \text{sign}(r)$, $r > 0$ (for a given $d \geq 2$), and $L(u) = \max(1, u/2)$, $u \geq 0$, will be used frequently in what follows. By convention, we set $R_d(r_{0j}) = \infty$ and $R_d(r_{N_j+1,j}) = -\infty$. The summations over all $j \sim k$ in the second and third terms refer to the $M \geq 2$ neighboring pairs $\{j, k\} \subset \{1, \ldots, M\}$ with $k - j = 1 \pmod{M}$. Also, $|\cdot|_\pi$ denotes absolute value mod π; hereafter we use similar notation with π replaced by other positive real numbers, depending on the context.

The intuition behind this ansatz comes from the well-known representation of the standard Gaussian density in two dimensions in terms of polar coordinates. That is, the angular coordinate (in the sense we defined above) can be taken as uniformly distributed on $[0, \pi]$, independent of the (signed) radial coordinate, which has the (two-sided) Rayleigh density $p(r) = b(r)\varphi(r)$, where φ is a standard normal density and $b(r) = \sqrt{\pi/2}|r|$, $r \in \mathbb{R}$.

The first term in $U_2(\mathbf{x})$ is based on a derivation given in [14], in which an interworld potential is proposed for densities of the general form $p(x) = b(x)\varphi(x)$, $x \in \mathbb{R}$. The focus in that paper was on the case of the two-sided Maxwell distribution, for which $b(x) = x^2$. The general ansatz for the interworld potential of a one-dimensional N-point configuration $x_1 > x_2 > \cdots > x_N$ was proposed to be

$$U_b(\mathbf{x}) = \sum_{n=1}^{N} \left[\frac{1}{B(x_{n+1}) - B(x_n)} - \frac{1}{B(x_n) - B(x_{n-1})} \right]^2 b(x_n)^2, \tag{3}$$

where $B(x) = \int_0^x b(t) \, dt$, $B(x_0) = \infty$, and $B(x_{N+1}) = -\infty$. When $b(x)$ is proportional to $|x|^k$ for some nonnegative integer k, as for the Rayleigh density ($k = 1$) or for the Gaussian density
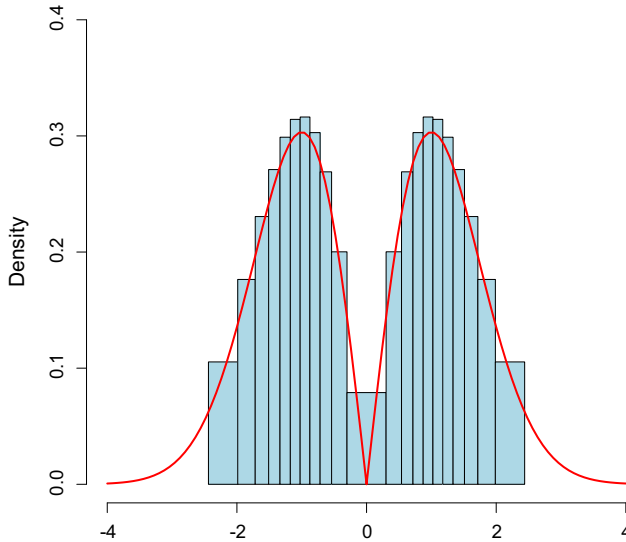
FIGURE 1. Histogram of the symmetric solution of the recursion (5) for $N = 22$ compared with the two-sided Rayleigh density; the breaks in the histogram are $x_1, \ldots, x_N$.

($k = 0$), it can be shown that the minimizer of $U_b(\mathbf{x}) + V_1(\mathbf{x})$ is a symmetric solution of the recursion

$$B(x_{n+1}) = B(x_n) - \left(\sum_{i=1}^{n} \frac{x_i}{b(x_i)} \right)^{-1}. \tag{4}$$

For the Rayleigh distribution, $b(x) = \sqrt{\pi/2}|x|$, so the interworld potential (3) becomes

$$U_b(\mathbf{x}) = 4 \sum_{n=1}^{N} \left[\frac{1}{R_2(x_{n+1}) - R_2(x_n)} - \frac{1}{R_2(x_n) - R_2(x_{n-1})} \right]^2 x_n^2.$$

From (4), the ground state is then a symmetric solution to the recursion equation

$$R_2(x_{n+1}) = R_2(x_n) - 2 \left(\sum_{i=1}^{n} \text{sign}(x_i) \right)^{-1}. \tag{5}$$

Further, it follows from analogous arguments in [14, Section 2] that the ground-state value of the Hamiltonian in this case is $4(N-1)$. Figure 1 compares the empirical distribution of the solution of the above recursion (for $N = 22$ points) with the two-sided Rayleigh density, showing that the agreement is remarkably accurate.

Returning now to the two-dimensional setting, after setting the kinetic energy to zero, the problem is to minimize

$$U_2(\mathbf{x}) + \sum_{j=1}^{M} \sum_{n=1}^{N_j} r_{nj}^2. \tag{6}$$

Note that the M components of the first term of the interworld potential $U_2(\mathbf{x})$ can be combined with the corresponding potential energy terms and separately minimized (using the recursion (5)), giving a combined contribution of $\sum_{j=1}^{M} 4(N_j - 1) = 4(N - M)$ to (6). It remains to minimize the sum of the second and third terms in $U_2(\mathbf{x})$ to furnish the complete ground-state solution. The strategy is to fix the value of $M \geq 2$ and then find the minimizing $\mathbf{x}$ (based on M radial directions). Note that, by Cauchy–Schwarz,

$$M = \sum_{j \sim k} \frac{|\theta_j - \theta_k|_\pi^{1/2}}{|\theta_j - \theta_k|_\pi^{1/2}} \leq \left(\sum_{j \sim k} |\theta_j - \theta_k|_\pi \right)^{1/2} \left(\sum_{j \sim k} \frac{1}{|\theta_j - \theta_k|_\pi} \right)^{1/2} = \left(\pi \sum_{j \sim k} \frac{1}{|\theta_j - \theta_k|_\pi} \right)^{1/2},$$

so the minimum of the second term is M^2, attained by setting $\theta_j = (j-1)\pi/M$, $j = 1, \ldots, M$, because that produces equality throughout the above display. With $M \leq N/2$ fixed, the third term in $U_2(\mathbf{x})$ is minimized by distributing the N points as evenly as possible among the M directions (with at least two in each direction) and then allotting the remaining $(N \bmod M)$ points to a sequence of neighboring directions. This results in $N_j = N_k$ for all neighbors j and k except possibly for two pairs of neighbors in which $|N_j - N_k| = 1$.

The sum of the last two terms of $U_2(\mathbf{x})$ therefore has the minimal value $M^2 + N^2/M^2$, and when combined with the contribution $4(N - M)$ arising from the first term in $U_2(\mathbf{x})$ along with the potential energy, as discussed earlier, the ground state minimizes $4(N - M) + M^2 + N^2/M^2$ as a function of M. The value of $M \leq N/2$ that attains this minimum (for a given value of N) is a monotone increasing function of N, and $M \sim \sqrt{N}$ as $N \to \infty$. It also follows that the empirical distribution of $\{N_j/N, j = 1, \ldots, M\}$ in the ground state converges weakly to uniform on $(0, 1)$ as $N \to \infty$.

Combined with the fact that the empirical distribution (conditional on M) of any set of minimizing angular coordinates $\{\theta_j, j = 1, \ldots, M\}$ converges weakly to uniform on $(0, \pi)$ as $M \to \infty$, we conclude that the empirical distribution of the angular coordinates of the N points in the ground state has the same limit as $N \to \infty$. Thus, provided we can show that the empirical distribution of the radial coordinates of the ground state converges to the Rayleigh distribution, it will follow that the full empirical distribution of the ground state (as illustrated by the left panel of Fig. 2) converges to the standard two-dimensional Gaussian distribution.

2.2. Interworld potentials for $d \geq 3$

In the three-dimensional case, a configuration $\mathbf{x}$ of points in $\mathbb{R}^3$ is specified in terms of (signed) spherical coordinates as

$$\{(r_{njk}, \theta_j, \phi_{jk}) : n = 1, \ldots, N_{jk}, \ j = 1, \ldots, M, \ k = 1, \ldots, K_j\}, \tag{7}$$

where there are (distinct) directions corresponding to points with polar angles $0 \leq \theta_1 < \theta_2 < \cdots < \theta_M \leq \pi/2$ and azimuthal angles $0 \leq \phi_{j1} < \phi_{j2} < \cdots < \phi_{jK_j} < 2\pi$, along with $N_{jk} \geq 2$ points in each direction (θ_j, ϕ_{jk}) determined by their (signed) radial distances $r_{1jk} < \cdots < r_{N_{jk}jk}$. There are a total of $N_{j\cdot} = \sum_{k=1}^{K_j} N_{jk}$ points corresponding to the K_j longitudinal directions in the jth parallel (or 'line of latitude') with polar angle θ_j. Over all M parallels, there are a total of $N = \sum_{j=1}^{M} N_{j\cdot}$ points arising from $K = \sum_{j=1}^{M} K_j$ different directions.

The spherical coordinates of a three-dimensional standard Gaussian random vector are independent. The signed radial component has the two-sided Maxwell distribution with density $r^2 \varphi(r)$, $r \in \mathbb{R}$, and the azimuthal angle is uniformly distributed on $[0, 2\pi)$. The polar
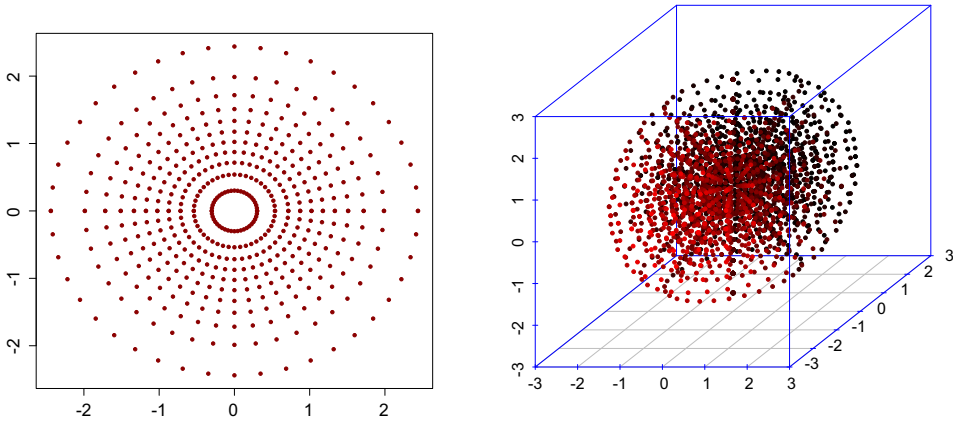
FIGURE 2. Ground states: $d = 2$ (left), $N = 22^2 = 484$ points, $M = 22$ directions, and $N_j = 22$ points in each direction; $d = 3$ (right), $N = 2744$ points, $M = 7$ parallels (polar angles in $[0, \pi/2]$), $K_j = 28$ azimuthal angles in each parallel, and $N_{jk} = 14$ points in each radial direction.

angle has cumulative distribution function (CDF) $1 - \cos(\theta)$, $\theta \in [0, \pi/2]$. Motivated by similar considerations to the two-dimensional case, the proposed interworld potential is taken to be

$$U_3(\mathbf{x}) = 9 \sum_{j=1}^{M} \sum_{k=1}^{K_j} \sum_{n=1}^{N_{jk}} \left(\frac{1}{r_{n+1,jk}^3 - r_{njk}^3} - \frac{1}{r_{njk}^3 - r_{n-1,jk}^3} \right)^2 r_{njk}^4$$

$$+ \sum_{i \sim j} \frac{1}{|\cos(\theta_i) - \cos(\theta_j)|} + \frac{\pi}{2} \sum_{j=1}^{M} \sum_{k \sim l} \frac{1}{|\phi_{jk} - \phi_{jl}|_{2\pi}}$$

$$+ \sum_{j=1}^{M} \frac{N_{j.}^2}{K_j^2} L\left(\sum_{k \sim l} |N_{jk} - N_{jl}| \right) + \frac{N}{4M} L\left(\sum_{i \sim j} |N_{i.} - N_{j.}| \right). \tag{8}$$

The last term above has a distinctive form that is not present in $U_2(\mathbf{x})$. The denominator in this term is designed to control the number of polar angles (M) in the second term, which in turn needs to be balanced with the K_j, since the polar angles range over an interval only one fourth the length of that for the azimuthal angles. Numerical experiments show that the resulting directional component in the ground state is close to uniformly distributed over the sphere, as illustrated in the right panel of Fig. 2.

We now seek to minimize the Hamiltonian $H_3(\mathbf{x}) = U_3(\mathbf{x}) + \sum_{j=1}^{M} \sum_{k=1}^{K_j} \sum_{n=1}^{N_{jk}} r_{njk}^2$. The components with distinct values of (j, k) in the first term of the interworld potential $U_3(\mathbf{x})$ can be combined with the corresponding potential energy terms and separately minimized (in terms of the recursion (4) with $b(x) = x^2$ that was studied in [14]), giving a combined contribution of $\sum_{j=1}^{M} \sum_{k=1}^{K_j} 6(N_{jk} - 1) = 6(N - K)$ to the Hamiltonian.

The next step is to minimize each of the last three terms in (8) for fixed $M, K_1, \ldots, K_M$. Using a similar Cauchy–Schwarz argument to what we used to minimize the second term in (2), the second term in (8) is minimized for

$$\theta_j = \cos^{-1}(1 - (j-1)/(M-1)), \qquad j = 1, \ldots, M, \tag{9}$$

and the minimum is $(M-1)^2$. Similarly, the minimum of the third term in (8) is attained by setting $\phi_{jk} = 2\pi(k-1)/K_j$, $k = 1, \ldots, K_j$, $j = 1, \ldots, M$, and the minimum is $\sum_{j=1}^{M} K_j^2/4$. The minimum of the fourth term given fixed values of the K_j and $N_{j\cdot}$ is found using the same argument as in the two-dimensional case for the third term in (2), resulting in the minimum $\sum_{j=1}^{M} N_{j\cdot}^2/K_j^2$ being attained when $N_{jk} \sim N_{j\cdot}/K_j$ for $k = 1, \ldots, K_j$. This reduces the minimal value of the Hamiltonian to

$$6(N-K) + (M-1)^2 + \sum_{j=1}^{M} [K_j^2/4 + N_{j\cdot}^2/K_j^2] + \frac{N}{4M} L\left(\sum_{i \sim j} |N_{i\cdot} - N_{j\cdot}| \right).$$

For fixed M, this expression is minimized when $K_j \sim \sqrt{2N_{j\cdot}}$ as $N_{j\cdot} \to \infty$, and it remains to minimize

$$7N + (M-1)^2 + \frac{N}{4M} L\left(\sum_{i \sim j} |N_{i\cdot} - N_{j\cdot}| \right).$$

This expression is minimized by $N_{j\cdot} \sim N/M$ and $M \sim N^{1/3}/2$ as $N \to \infty$. This implies

$$N_{jk} \sim N_{j\cdot}/K_j \sim \sqrt{N_{j\cdot}/2} \sim \sqrt{N/(2M)} \sim N^{1/3}, \qquad K_j \sim 2N^{1/3}.$$

We conclude that, as $N \to \infty$, the ground state consists of $K_j = 2N^{1/3}$ directions in each of $M = N^{1/3}/2$ parallels, with $N_{jk} = N^{1/3}$ points falling in each direction. An example with $N = 2744$ points is provided in the right panel of Fig. 2.

The extension to general $d \geq 3$ is straightforward. The radial component of the d-dimensional standard Gaussian distribution has density $p(r) = c_d |r|^{d-1} \varphi(r)$, $r \in \mathbb{R}$, where c_d is a normalizing constant, and is independent of the (signed) directional component, which is uniformly distributed (in the sense of Hausdorff measure) over the unit hemisphere $\{x \in \mathbb{R}^d : x_1 \geq 0, \ |x| = 1\}$. The configuration $\mathbf{x}$ is now specified using signed *hyperspherical* coordinates, precisely as in (7), except each polar angle θ_j is now required to be a *vector* $\theta_j = (\theta_{jl})$ of $d-2$ polar angles, with each component satisfying $0 \leq \theta_{1l} < \theta_{2l} < \cdots < \theta_{Ml} \leq \pi/2$, $l = 1, \ldots, d-2$. The notation for the azimuthal angles, the M parallels, and the K directions remains the same.

The $d-2$ polar angles are independent and identically distributed under the d-dimensional standard Gaussian, and their CDF can be expressed using a simple formula for the area of a hyperspherical cap [12]:

$$F_d(\theta) = B\left(\sin^2 \theta, \frac{d-1}{2}, \frac{1}{2} \right), \qquad \theta \in [0, \pi/2],$$

where $B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}\, dt$ is the incomplete Beta function. In particular, $F_4(\theta) = (2\theta - \sin(2\theta))/\pi$ and $F_5(\theta) = 1 - (3/2)\cos\theta + (1/2)\cos^3\theta$. This leads to the interworld potential

$$U_d(\mathbf{x}) = d^2 \sum_{j=1}^M \sum_{k=1}^{K_j} \sum_{n=1}^{N_{jk}} \left(\frac{1}{R_d(r_{n+1,jk}) - R_d(r_{njk})} - \frac{1}{R_d(r_{njk}) - R_d(r_{n-1,jk})} \right)^2 r_{njk}^{2(d-1)}$$

$$+ \frac{1}{(d-2)} \sum_{l=1}^{d-2} \sum_{i \sim j} \frac{1}{|F_d(\theta_{il}) - F_d(\theta_{jl})|} + \frac{\pi}{2} \sum_{j=1}^M \sum_{k \sim l} \frac{1}{|\phi_{jk} - \phi_{jl}|_{2\pi}}$$

$$+ \sum_{j=1}^M \frac{N_{j\cdot}^2}{K_j^2} L \left(\sum_{k \sim l} |N_{jk} - N_{jl}| \right) + \frac{N}{4(d-2)M^{d-2}} L \left(\sum_{i \sim j} |N_{i\cdot} - N_{j\cdot}| \right). \qquad (10)$$

It follows by a routine extension of the argument we used in the $d = 3$ case that the ground state asymptotically consists of $K_j = 2N^{1/2 - 1/(2d)}$ directions in each of $M = N^{1/d}/2$ parallels, with $N_{jk} = N^{1/2 - 1/(2d)}$ points falling in each direction.

A possible alternative approach is to specify the directions in the unit hemisphere by a minimal Riesz energy point configuration, without reference to polar and azimuthal components. As surveyed in [1], such configurations can provide asymptotically uniformly distributed point sets with respect to surface area (Hausdorff measure), with important applications in quasi-Monte Carlo, approximation theory, and material physics. However, there does not seem to be a way of linking the numbers of radial points with the directions obtained from minimizing Riesz energy that would lead to a Gaussian approximation. In contrast, the Hamiltonian based on the interworld potential (10) is readily minimized by following the same argument we used to minimize (8); the only difference in the solution is that $\cos^{-1}$ in (9) is replaced by F_d^{-1} in the expression for each polar angle θ_{jl}. Moreover, as explained in the following sections, our proposed approach leads to explicit rates of convergence (in terms of Wasserstein distance) of the empirical measure of the ground state to standard Gaussian.

2.3. Interworld potentials for excited states

The eigenstates of the classical d-dimensional isotropic quantum harmonic oscillator consist of products of d one-dimensional eigenfunctions, with separate Euclidean coordinates in each component. In this section we show how the approach in the previous sections extends naturally to the case when some of these one-dimensional components are in excited (higher-energy) states. The various eigenstates of the full system are described by vectors of quantum numbers indicating the energy level of each component (with 0 indicating the ground state). When expressed in spherical or polar coordinates, there is a separation of variables in the various eigenfunctions, which implies that the corresponding densities have independent contributions from the radial and angular components. This allows a separation of the radial and angular components in the interworld potential, as we now make explicit in examples of excited states for $d = 2$ and 3. The idea readily extends to any excited state of the MIW quantum harmonic oscillator.

For $d = 2$, the lowest three excited states have quantum numbers $(1, 0)$, $(0, 1)$, and $(1, 1)$. The density of the $(1, 0)$ state in signed polar coordinates is proportional to $\cos^2(\theta)|r|^3 \varphi(r)$, $\theta \in [0, \pi)$, $r \in \mathbb{R}$. The CDF of the angular component is $G_2(\theta) = (\theta + \sin(\theta)\cos(\theta))/\pi$, and the density of the signed radial component is proportional to $|r|^3 \varphi(r)$. This leads to the following interworld potential similar to (2):
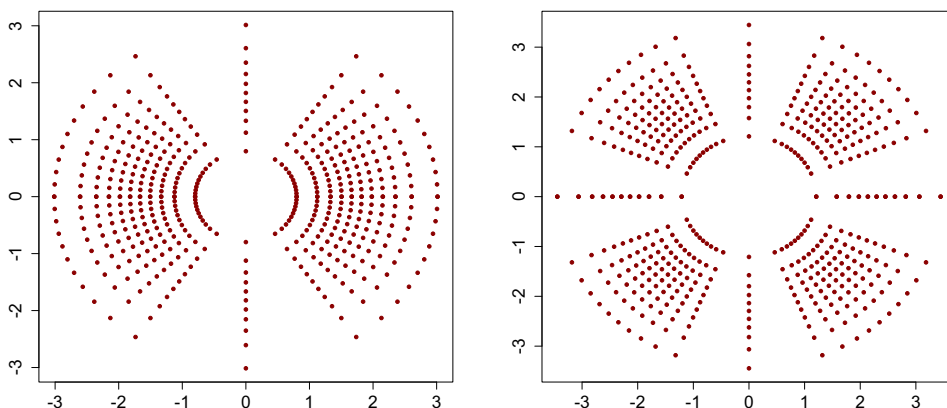
FIGURE 3. Excited states $(1, 0)$ (left) and $(1, 1)$ (right) for $d = 2$, $N = 22^2 = 484$ points, $M = 22$, $N_j = 22$; cf. the ground state in Fig. 2.

$$U_2(\mathbf{x}) = 16 \sum_{j=1}^{M} \sum_{n=1}^{N_j} \left[\frac{1}{R_4(r_{n+1,j}) - R_4(r_{nj})} - \frac{1}{R_4(r_{nj}) - R_4(r_{n-1,j})} \right]^2 r_{nj}^6$$

$$+ \sum_{j \sim k} \frac{1}{|G_2(\theta_j) - G_2(\theta_k)|_1} + \frac{N^2}{M^2} L \left(\sum_{j \sim k} |N_j - N_k| \right).$$

Following similar arguments to those used earlier, the Hamiltonian (6) based on this potential is minimized with the radial coordinates given by the symmetric solution to the recursion (4) with $b(x) = |x|^3$, and setting $\theta_j = G_2^{-1}((j-1)/M)$ with $M \sim \sqrt{N}$. Figure 3 shows the $(1, 0)$ and $(1, 1)$ excited states resulting from $N = 484$ points. The $(0, 1)$ state is the same as $(1, 0)$ except rotated by $90°$.

For $d = 3$, the density of the $(1, 0, 0)$ excited state in signed spherical coordinates is proportional to $\sin^3(\theta) \cos^2(\phi) r^4 \varphi(r)$, $\theta \in [0, \pi/2)$, $\phi \in [0, 2\pi)$, $r \in \mathbb{R}$. The CDF of the polar angle is $G_3(\theta) = (\cos(3\theta) - 9\cos(\theta))/8$, and the CDF of the azimuthal angle is $A(\phi) = (\phi - \sin(\phi)\cos(\phi))/(2\pi)$. The interworld potential is similar to (8):

$$U_3(\mathbf{x}) = 25 \sum_{j=1}^{M} \sum_{k=1}^{K_j} \sum_{n=1}^{N_{jk}} \left(\frac{1}{r_{n+1,jk}^5 - r_{njk}^5} - \frac{1}{r_{njk}^5 - r_{n-1,jk}^5} \right)^2 r_{njk}^8$$

$$+ \sum_{i \sim j} \frac{1}{|G_3(\theta_i) - G_3(\theta_j)|} + \frac{\pi}{2} \sum_{j=1}^{M} \sum_{k \sim l} \frac{1}{|A(\phi_{jk}) - A(\phi_{jl})|_{2\pi}}$$

$$+ \sum_{j=1}^{M} \frac{N_{j\cdot}^2}{K_j^2} L \left(\sum_{k \sim l} |N_{jk} - N_{jl}| \right) + \frac{N}{4M} L \left(\sum_{i \sim j} |N_{i\cdot} - N_{j\cdot}| \right).$$

The configurations of the $(1, 0, 0)$, $(1, 1, 0)$, and $(1, 1, 1)$ excited states resulting from $N = 2744$ points are displayed in Fig. 4.
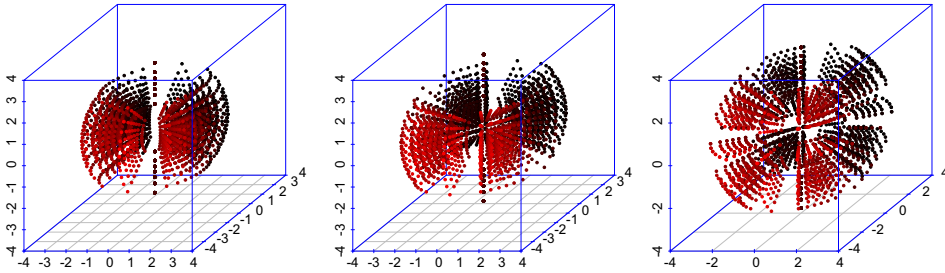
FIGURE 4. Excited states $(1, 0, 0)$ (left), $(1, 1, 0)$ (middle), and $(1, 1, 1)$ (right) for $d = 3$, $N = 2744$ points, $M = 7$, $K_j = 28$, $N_{jk} = 14$; cf. the ground state in the right panel of Fig. 2.

2.4. Wasserstein distance in spherical coordinates

The Wasserstein distance between probability measures μ and ν on a Polish metric space (S, ρ) can be defined equivalently as

$$d_{\mathrm{W}}(\mu, \nu) = \sup_{h \in \mathcal{H}} \left| \int_S h \, \mathrm{d}(\mu - \nu) \right| = \inf \mathbb{E} \rho(X, Y),$$

where $\mathcal{H}$ is the family of Lipschitz functions $h : S \to \mathbb{R}$ with $\mathrm{Lip}(h) \le 1$, and the (attained) infimum is over all coupled S-valued random elements $X \sim \mu$ and $Y \sim \nu$. A coupling that attains the above infimum exists by the Monge–Kantorovich theorem. When $S = \mathbb{R}$ and μ and ν have CDFs F and G, their Wasserstein distance coincides with the L_1-distance $\int |F(x) - G(x)| \, \mathrm{d}x$. The following inequality bounds the Wasserstein distance between two product measures $\mu = \prod_{i=1}^d \mu_i$ and $\nu = \prod_{i=1}^d \nu_i$ on $\mathbb{R}^d$ endowed with the Euclidean metric: $d_{\mathrm{W}}(\mu, \nu) \le \sum_{i=1}^d d_{\mathrm{W}}(\mu_i, \nu_i)$ (see [15, Lemma 3]).

We now present a variation of this result to enable the study of the convergence of the Wasserstein distance between the empirical distribution of the energy-minimizing configuration $\mathbf{x}$ and the distribution specified by quantum theory. Concentrating on the case $d = 3$ for simplicity, the following result provides a bound on the Wasserstein distance between two probability measures μ and ν on $\mathbb{R}^3$ (of the type that is relevant here) in terms of the Wasserstein distance between the distributions of their respective (signed) spherical coordinates (denoted μ_r, μ_θ, μ_ϕ, etc.).

Lemma 1. *Let $X \sim \mu$ and $Y \sim \nu$ be random vectors in $\mathbb{R}^3$. If the signed spherical coordinates of X are independent, and the same holds for Y, then*

$$d_{\mathrm{W}}(\mu, \nu) \le d_{\mathrm{W}}(\mu_r, \nu_r) + \sqrt{m_\mu m_\nu}[d_{\mathrm{W}}(\mu_\theta, \nu_\theta) + d_{\mathrm{W}}(\mu_\phi, \nu_\phi)],$$

where $m_\mu = \int |x| \, \mathrm{d}\mu_r(x)$ is the mean absolute deviation of the radial coordinate of X, and similarly for m_ν.

The result is the same in the case $d = 2$, except the azimuthal contribution $d_{\mathrm{W}}(\mu_\phi, \nu_\phi)$ is absent. The result naturally extends to general $d \ge 3$, with additional terms of the form $d_{\mathrm{W}}(\mu_\theta, \nu_\theta)$ representing each of the $(d - 2)$ polar angles. The directional contributions of the lowest-energy configuration will be shown to have a faster rate of convergence (as $N \to \infty$)

than the radial contribution, so it will suffice to restrict attention to the convergence of the latter, as we do in the next section.

Proof. For points $\underline{x}_1, \underline{x}_2 \in \mathbb{R}^3$, the squared Euclidean distance between them in terms of their signed spherical coordinates (r_i, θ_i, ϕ_i), $i = 1, 2$, is

$$\|\underline{x}_1 - \underline{x}_2\|^2 = r_1^2 + r_2^2 - 2r_1r_2[\cos(\theta_1 - \theta_2) + \sin(\theta_1)\sin(\theta_2)(\cos(\phi_1 - \phi_2) - 1)]$$

$$= (r_1 - r_2)^2 - 2r_1r_2[\cos(\theta_1 - \theta_2) - 1 + \sin(\theta_1)\sin(\theta_2)(\cos(\phi_1 - \phi_2) - 1)]$$

$$\leq (r_1 - r_2)^2 + r_1r_2[(\theta_1 - \theta_2)^2 + (\phi_1 - \phi_2)^2],$$

using the inequality $|\cos(x) - 1| \leq x^2/2$ for all $x \in \mathbb{R}$. Then, using the inequality $(|a| + |b| + |c|)^{1/2} \leq |a|^{1/2} + |b|^{1/2} + |c|^{1/2}$ for any $a, b, c \in \mathbb{R}$, we obtain

$$\|\underline{x}_1 - \underline{x}_2\| \leq |r_1 - r_2| + \sqrt{|r_1||r_2|}[|\theta_1 - \theta_2| + |\phi_1 - \phi_2|]. \tag{11}$$

By the Monge–Kantorovich theorem referred to above, we can choose coupled random vectors $X^\star \sim \mu$ and $Y^\star \sim \nu$ such that $d_W(\mu_r, \nu_r) = \mathbb{E}|X_r^\star - Y_r^\star|$, $d_W(\mu_\theta, \nu_\theta) = \mathbb{E}|X_\theta^\star - Y_\theta^\star|$, and $d_W(\mu_\phi, \nu_\phi) = \mathbb{E}|X_\phi^\star - Y_\phi^\star|$, with the random vectors $(X_r^\star, Y_r^\star)$, $(X_\theta^\star, Y_\theta^\star)$, and $(X_\phi^\star, Y_\phi^\star)$ being independent. Substituting $X^\star$ and $Y^\star$ for $\underline{x}_1$ and $\underline{x}_2$ in (11) and taking expectations of both sides leads to

$$d_W(\mu, \nu) \leq \mathbb{E}\|X^\star - Y^\star\| \leq \mathbb{E}|X_r^\star - Y_r^\star| + [\mathbb{E}\sqrt{|X_r^\star||Y_r^\star|}][\mathbb{E}|X_\theta^\star - Y_\theta^\star| + \mathbb{E}|X_\phi^\star - Y_\phi^\star|].$$

The proof is completed using Cauchy–Schwarz to obtain $\mathbb{E}\sqrt{|X_r^\star||Y_r^\star|} \leq \sqrt{m_\mu m_\nu}$. $\qquad\square$

3. Stein's method for approximating radial distributions

We begin this section with some generalities concerning the version of Stein's method developed in [5, 6]. We only give a summary here and refer the reader unfamiliar with Stein's method to the Supplementary Material for more information and background.

Consider a standardized (i.e mean 0, variance 1) random variable F_∞ with density p_∞ (in the following, p_∞ will take the form of a radial distribution), CDF $P_\infty = \int_{-\infty}^{\cdot} p_\infty(y)\,dy$, and survival function $\bar{P}_\infty = 1 - P_\infty$. With p_∞ we associate the 'density Stein operators'

$$\mathcal{T}_\infty f(x) = \frac{(f(x)p_\infty(x))'}{p_\infty(x)} = f'(x) + \frac{p'_\infty(x)}{p_\infty(x)}f(x),$$

$$\mathcal{L}_\infty h(x) = -\int_{-\infty}^{\infty} h'(y)\frac{P_\infty(y \wedge x)\bar{P}_\infty(y \vee x)}{p_\infty(x)}\,dy$$

acting on absolutely continuous functions f and h such that, moreover, f' and h are integrable with respect to p_∞. Readers familiar with Stein's method will recognize the operator $\mathcal{L}_\infty$ once it is noticed that it is equivalent, for differentiable test functions h, to $\mathcal{L}_\infty h(x) = p_\infty(x)^{-1}\int_{-\infty}^{x}(h(u) - \mathbb{E}h(F_\infty))\,du$; see the Supplementary Material. It can be shown that $\mathcal{T}_\infty\mathcal{L}_\infty h(x) = h(x) - \mathbb{E}h(F_\infty)$, where $\mathcal{L}_\infty\mathcal{T}_\infty f(x) = f(x)$ for all appropriate h and f. The choice $h(x) = -x$ gives the *Stein kernel*,

$$\tau_\infty(x) := \mathcal{L}_\infty h(x) = \int_{-\infty}^{\infty} \frac{P_\infty(y \wedge x)\bar{P}_\infty(y \vee x)}{p_\infty(x)}\,dy.$$

It is now well documented that Stein kernels, if available, provide valuable insight into the properties of the underlying densities. Following [4, 11], we propose to apply Stein's method based on the following Stein operator:

$$\mathcal{A}_\infty g(x) = \mathcal{T}_\infty(\tau_\infty(x)g(x)) = \tau_\infty(x)g'(x) - xg(x).$$

Note that, by design, $\mathbb{E}[\mathcal{A}_\infty g(F_\infty)] = 0$ for all g for which the moments exist. The Stein equation for p_∞ associated with operator $\mathcal{A}_\infty$ is then

$$\tau_\infty(x)g'(x) - xg(x) = h(x) - \mathbb{E}h(F_\infty) \tag{12}$$

for all appropriate h, and one can easily check that the function

$$g_h(x) = \frac{\mathcal{L}_\infty h(x)}{\tau_\infty(x)} \tag{13}$$

solves (12) at all x such that $\tau_\infty(x) \neq 0$, so that, for all real random variables F such that $P(\tau_\infty(F) = 0) = 0$,

$$\mathbb{E}h(F) - \mathbb{E}h(F_\infty) = \mathbb{E}[\tau_\infty(F)g_h'(F) - Fg_h(F)], \tag{14}$$

with g_h as given in (13). By the definition of the Wasserstein distance, we conclude that

$$d_W(\mathbb{F}, \mathbb{F}_\infty) \leq \sup_{h \in \mathcal{H}} \mathbb{E}[\tau_\infty(F)g_h'(F) - Fg_h(F)]|, \tag{15}$$

where $\mathbb{F}$ is the probability distribution of F, $\mathbb{F}_\infty$ is that of F_∞, and $\mathcal{H}$ is the family of Lipschitz functions with constant less than 1 (recall the first lines of Section 2.4).

Specializing to densities of the form $p_\infty(x) = b(x)\varphi(x)$, with $\varphi(x)$ the standard normal density, leads to the following nonuniform bounds for g; they are based on results from [6] and are proved in the Supplementary Material (see Lemma 2 therein).

Proposition 1. *Let $p_\infty(x) = b(x)\varphi(x)$ be a standardized density on the real line, with $b(x)$ a positive function such that $b(x) \neq 0$ for all $x \neq 0$. Define*

$$R_\infty(x) = \frac{1}{p_\infty(x)} \int_{-\infty}^{x} P_\infty(u)\, du \int_{x}^{\infty} \overline{P}_\infty(u)\, du.$$

Let $g(= g_h)$ be as in (13), with $h \in \mathcal{H}$. Then

$$|g(x)| \leq 1, \tag{16}$$

$$|g'(x)| \leq 2\frac{R_\infty(x)}{\tau_\infty(x)^2} =: \Psi_1(x), \tag{17}$$

$$|g''(x)| \leq \frac{2}{\tau_\infty(x)}\left(1 + \left|\frac{2x}{\tau_\infty(x)} - x + \frac{b'(x)}{b(x)}\right|\frac{R_\infty(x)}{\tau_\infty(x)}\right) =: 2\Psi_2(x). \tag{18}$$

An important aspect of the inequalities (16), (17), and (18) is that they hold uniformly over all $h \in \mathcal{H}$; our approach to Stein's method consists in using this information in combination with (15) in order to bound the Wasserstein distance.

We now further specialize to the choice $b(x) \propto |x|^k$, where $k = d - 1 \geq 0$ in the case of a d-dimensional radial distribution; throughout the remainder of this section we use the notation k
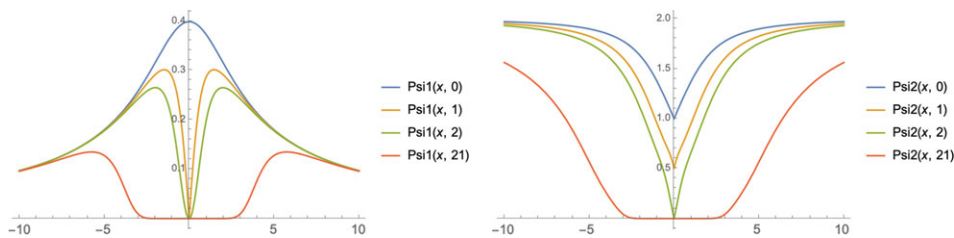
FIGURE 5. Functions Ψ_1 (left) and Ψ_2 (right) defined in (17) and (18) for $b(x) \propto |x|^k$ with $k = 0$ (blue), $k = 1$ (orange), $k = 2$ (green), and $k = 21$ (red).

rather than $d - 1$, but then in the final section we recast our results in terms of the dimension d. The following explicit expression for the Stein kernel in this case is obtained by straightforward integration of the expressions involved.

Lemma 2. *Let $p_\infty(x) \propto |x|^k \varphi(x)$ for given $k \geq 0$ with $\varphi(x)$ the standard Gaussian density. Then the Stein kernel is*

$$\tau_\infty(x; k) = 2^{k/2} e^{x^2/2} |x|^{-k} \Gamma(1 + k/2, x^2/2), \tag{19}$$

where $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t}\, dt$ is the (upper) incomplete gamma function.

Example 1. The following results are immediate from (19).

- If $k = 0$ (i.e. p_∞ is standard Gaussian density) then $\tau_\infty(x) = 1$.

- If $k = 1$ (i.e. p_∞ is two-sided Rayleigh density) then

$$\tau_\infty(x) = 1 + \sqrt{\pi/2} |x|^{-1} e^{x^2/2} \mathrm{Erfc}(x/\sqrt{2})$$

  (which behaves like $1/|x| + 1$).

- If $k = 2$ (i.e. p_∞ is two-sided Maxwell density) then $\tau_\infty(x) = 2/x^2 + 1$.

It is not hard to obtain expressions for the Stein kernel at any level $k \in \mathbb{N}$; they are provided in the Supplementary Material (see Lemma 5 therein).

The following properties of the functions defined in (17) and (18) (with τ_∞ as given in (19) and Ψ_1 and Ψ_2 positive) will be useful as well. (See Fig. 5 for an illustration. The Supplementary Material contains more information on these functions.)

- If $k = 0$ then $\Psi_1(x)$ is unimodal with maximum $\Psi_1(0) = \sqrt{2/\pi}$ at $x = 0$ and strictly decreasing towards 0 for $x \geq 0$; $\Psi_2(x)$ is unimodal with minimum $\Psi_2(0) = 1$ at $x = 0$ and strictly increasing towards 2 for $x \geq 0$.

- If $k = 1$ then $\Psi_1(x)$ is bimodal with minimum $\Psi_1(0) = 0$ at $x = 0$, maximum value less than 1, after which it is strictly decreasing towards 0 for $x \geq 0$; $\Psi_2(x)$ is unimodal with minimum $\Psi_2(0) = \frac{1}{2}$ at $x = 0$ and strictly increasing towards 2 for $x \geq 0$.

- If $k \geq 2$ then $\Psi_1(x)$ is bimodal with minimum $\Psi_1(0) = 0$ at $x = 0$, maximum value less than 1, after which it is strictly decreasing towards 0 for $x \geq 0$; $\Psi_2(x)$ is unimodal with minimum $\Psi_2(0) = 0$ at $x = 0$ and strictly increasing towards 2 for $x \geq 0$.

Having presented the key elements of Stein's method for radial distributions, we now give some properties of the sequences of discrete distributions whose proximity to these radial distributions we wish to assess. The following lemma, ensuring a unique symmetric and monotone solution to (4), is a routine extension of a result in [13, 14] to general k. In the following we assume N is even and define the median of an ordered set $\{x_1, \ldots, x_N\}$ to be $(x_m + x_{m+1})/2$, where $m = N/2$.

Lemma 3. *Every zero-median solution $\{x_1, \ldots, x_N\}$ of* (4) *when b(x) is proportional to $|x|^k$ for a given integer $k \geq 0$ satisfies the following properties.*

    (P1) *Zero mean:* $x_1 + \cdots + x_N = 0$

    (P2) *Variance bound:* $x_1^2 + \cdots + x_N^2 = (k+1)(N-1)$

    (P3) *Symmetry:* $x_n = -x_{N+1-n}$ *for* $n = 1, \ldots, N$

Further, there exists a unique solution $\{x_1, \ldots, x_N\}$ such that (P1) and

    (P4) *Strictly decreasing:* $x_1 > \ldots > x_N$

hold. This solution has the zero median property, and thus also satisfies (P2) and (P3).

Let $\mathbb{F}_N$ be the empirical distribution of the unique solution $\{x_1, \ldots, x_N\}$. For the case $k = 0$, the following result gives the optimal rate of convergence in Wasserstein distance of $\mathbb{F}_N$ to standard Gaussian. This result is not new, as the upper bound follows by adapting the coupling argument in [14], and the lower bound, which follows from a very delicate analysis of the x_i, was recently proved in [2]. We provide a new proof for the upper bound in Example 2.

Proposition 2. *The empirical distribution $\mathbb{F}_N$ on the unique symmetric and monotone solution to* (1) *(or* (4) *with $k = 0$) satisfies $d_W(\mathbb{F}_N, \mathbb{F}_\infty) \asymp \sqrt{\log N}/N$, where $\mathbb{F}_\infty$ is the standard Gaussian distribution.*

Upper bounds on the rate of convergence for values of $k \geq 1$ can also be obtained by adapting the arguments in [14] using the coupling approach (as developed in detail in the Supplementary Material). The following result for the case $k = 1$ is based on this approach (and is detailed in Section 2.3 of the Supplementary Material), but it is not expected to yield the optimal rate which, as argued below, we expect to be the same as the Gaussian case in the above proposition.

Proposition 3. *The empirical distribution $\mathbb{F}_N$ on the unique symmetric and monotone solution to* (5) *(or* (4) *with $k = 1$) satisfies $d_W(\mathbb{F}_N, \mathbb{F}_\infty) = O(\log N/N)$, where $\mathbb{F}_\infty$ is the two-sided Rayleigh distribution.*

The main result of this section is the following theorem based on a discrete density approach to Stein's method; cf. the approach of [8] in the case of points on an integer grid. This result is entirely new and may be of independent interest as the upper bound it provides does not rely on any specific structure of $x_1, \ldots, x_N$ apart from symmetry and monotonicity, and therefore may be much more widely applicable. It leads to the new proof of the tight upper bound in Proposition 2 for the case $k = 0$, as shown in Example 2. For $k \geq 1$, it becomes more difficult to analytically determine the order of the bound, but it nevertheless provides a tight bound that can be evaluated numerically and that agrees with the asymptotic order of $d_W(\mathbb{F}_N, \mathbb{F}_\infty)$ as $N \to \infty$, as discussed in Remark 3.
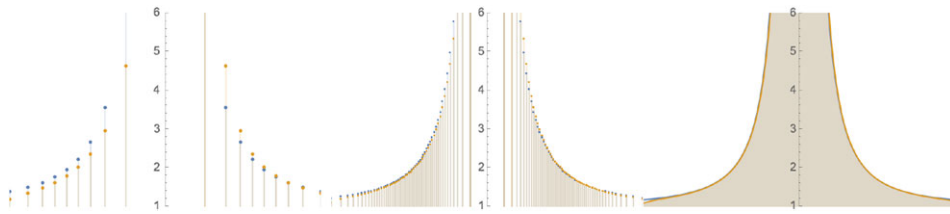
FIGURE 6. Superimposition of $\tau_N(x_i)$, $1 \leq i \leq N$, (orange bars) and $\tau_\infty(x_i)$, $1 \leq i \leq N$, (blue bars) when $\tau_\infty(x) = 1 + 2/x^2$ is the Stein kernel for the two-sided Maxwell distribution and $(x_i)_{1 \leq i \leq N}$ is as in Lemma 3 with $k = 2$ and $N = 20$ (left), $N = 110$ (middle), and $N = 300$ (right).
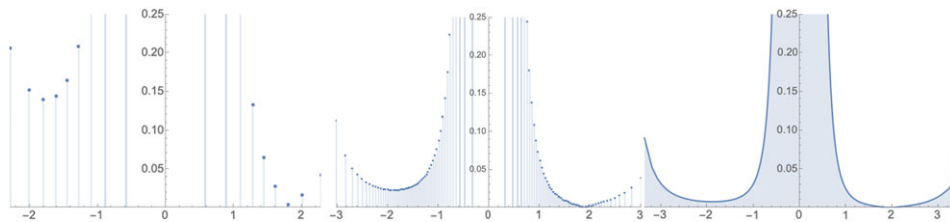

FIGURE 7. Plot of $|\tau_N(x_i) - \tau_\infty(x_i)|$, $1 \leq i \leq N$, when $\tau_\infty(x) = 1 + 2/x^2$ is the Stein kernel for the two-sided Maxwell distribution and $(x_i)_{1 \leq i \leq N}$ is as in Lemma 3 with $k = 2$ and $N = 20$ (left), $N = 110$ (middle), and $N = 300$ (right).

Theorem 1. *Let N be an even integer, and $x_1, \ldots, x_N$ be any symmetric and strictly monotone decreasing sequence. Set $\tau_N(x_i) = (x_i - x_{i+1}) \sum_{j=1}^{i} x_j$ for $1 \leq i \leq N - 1$, and $\tau_N(x_N) = 0$. Let $\mathbb{F}_N$ be the empirical distribution of these points, and let $\mathbb{F}_\infty$ be the probability measure with density p_∞ satisfying the assumptions of Proposition 1. Then*

$$d_W(\mathbb{F}_N, \mathbb{F}_\infty) \leq \frac{1}{N} \sum_{i=1}^{N} |\tau_\infty(x_i) - \tau_N(x_i)| \Psi_1(x_i) \tag{20}$$

$$+ \frac{1}{N} \sum_{i=1}^{N-1} |x_i - x_{i+1}| \tau_N(x_i) \max\{\Psi_2(x_i), \Psi_2(x_{i+1})\}, \tag{21}$$

where Ψ_1 and Ψ_2 are given in (17) and (18), respectively.

Remark 1. Equation (20) encourages us to consider the quantity $\tau_N(x_i) = (x_i - x_{i+1}) \sum_{j=1}^{i} x_j$ as a Stein kernel for the uniform distribution on the set $\{x_1, \ldots, x_N\}$. This function is nonnegative and, under the conditions that we are considering, we have $\tau_N(x) \approx \tau_\infty(x)$ throughout the support of $\mathbb{F}_N$ except close to the edges and the origin (see Figs. 6 and 7).

Remark 2. Since both Ψ_1 and Ψ_2 are uniformly bounded by 2, we immediately obtain the simpler upper bound

$$d_{\mathrm{W}}(\mathbb{F}_N, \mathbb{F}_\infty) \leq \frac{2}{N} \sum_{i=1}^{N} |\tau_\infty(x_i) - \tau_N(x_i)| + \frac{2}{N} \sum_{i=1}^{N-1} |x_i - x_{i+1}| \tau_N(x_i).$$

This simplicity comes at the cost of a loss of precision in the rate. For instance, in the case that we consider here, the presence of the functions Ψ_1 and Ψ_2 in the bounds controls the problems around the origin.

Example 2. (*Case $k = 0$.*) The sequence $\{x_n\}$ satisfies $x_n - x_{n+1} = (\sum_{i=1}^{n} x_i)^{-1}$ so that $\tau_N(x_i) = 1$ for all $i = 1, \ldots, N - 1$. Since $\tau_\infty(x_i) = 1$, (20) vanishes except for the term corresponding to $i = N$, leading to $\Psi_1(x_N)/N \leq 1/N$. Also, (21) reads

$$\frac{1}{N} \sum_{i=1}^{N-1} |x_{i+1} - x_i| \max\{\Psi_2(x_i), \Psi_2(x_{i+1})\}.$$

Using $\Psi_2 \leq 2$ along with the symmetry of the sequence, we then have

$$d_{\mathrm{W}}(\mathbb{F}_N, \mathbb{F}_\infty) \leq \frac{1}{N} + \frac{2}{N} \sum_{i=1}^{N-1} |x_{i+1} - x_i| = \frac{1 + 4x_1}{N} = O(\sqrt{\log N}/N),$$

where the last step follows by a version of the argument in [14, Lemma 4.8] showing that $x_1 = O(\sqrt{\log N})$. This proves the upper bound part of Proposition 2.

Remark 3. (*Cases $k \geq 1$.*) Evaluating the bound in Theorem 1 based on numerical solutions to (4) and comparing the results with values of $d_{\mathrm{W}}(\mathbb{F}_N, \mathbb{F}_\infty)$ calculated using numerical integration, over a range of values of N, suggests that the bound is of the same order as $d_{\mathrm{W}}(\mathbb{F}_N, \mathbb{F}_\infty)$, and moreover that

$$d_{\mathrm{W}}(\mathbb{F}_N, \mathbb{F}_\infty) \asymp \sqrt{\log N}/N \qquad \text{for } k = 1, 2,$$

$$d_{\mathrm{W}}(\mathbb{F}_N, \mathbb{F}_\infty) \asymp (\log N)^6/N^2 \qquad \text{for } k \geq 3.$$

To obtain these rates we used Mathematica to compute the bounds and R to compute the Wasserstein distance using its representation as the L_1 distance between the CDFs (the programs are provided in the Supplementary Material).

Remark 4. We emphasize that the bound in Theorem 1 holds irrespective of the sequence that is chosen; it may be of interest to optimize the approximation of p_∞ by a sample $x_1 > \cdots > x_N$ which minimizes the right-hand side. One simple way to do this is to require that the sequence satisfies the recurrence

$$x_{i+1} = x_i - \frac{\tau_\infty(x_i)}{\sum_{j=1}^{i} x_j},$$

thereby canceling out (20) and only leaving (21); note that in the Gaussian case $\tau_\infty = 1$ so we are back with the recurrence (1).

Proof of Theorem 1. The idea is to apply a version of Stein's density method to $F_N \sim \mathbb{F}_N$. Note that a discrete 'derivative' at x_i, $i = 1, \ldots, N - 1$, is given by $D_N f(x_i) = f(x_i) - f(x_{i+1})$.

Writing $\sigma(x_i) = \sum_{j=1}^{i} x_j$, straightforward summation, along with the identity $x_N = -\sum_{j=1}^{N-1} x_j$, shows that $\mathbb{E}[\sigma(F_N)D_Ng(F_N) - F_Ng(F_N)] = 0$ for all summable functions g. Subtracting this from (14) applied to $F = F_N$ gives

$$\mathbb{E}h(F_N) - \mathbb{E}h(F_\infty) = \mathbb{E}[\tau_\infty(F_N)g'(F_N) - F_Ng(F_N)]$$
$$= \mathbb{E}[\tau_\infty(F_N)g'(F_N) - \sigma(F_N)D_Ng(F_N)].$$

From (15) it then follows that

$$d_W(\mathbb{F}_N, \mathbb{F}_\infty) \leq \sup_{h \in \mathcal{H}} |\mathbb{E}[\tau_\infty(F_N)g'(F_N) - \sigma(F_N)D_Ng(F_N)]|,$$

with $g\ (= g_h)$ satisfying the bounds in Proposition 1. By Taylor's theorem,

$$D_Nf(x_i) = (x_i - x_{i+1})f'(x_i) - \tfrac{1}{2}(x_i - x_{i+1})^2f''(c_i) \qquad \text{for all } 1 \leq i \leq N-1,$$

where c_i is between x_{i+1} and x_i. Hence,

$$\mathbb{E}h(F_N) - \mathbb{E}h(F_\infty) = \frac{1}{N}\sum_{i=1}^{N-1}(\tau_\infty(x_i) - \sigma(x_i)(x_i - x_{i+1}))g'(x_i) - \frac{1}{2N}\sum_{i=1}^{N-1}\sigma(x_i)(x_i - x_{i+1})^2g''(c_i)$$

$$+ \tau_\infty(x_N)g'(x_N)$$

$$= \frac{1}{N}\sum_{i=1}^{N}(\tau_\infty(x_i) - \tau_N(x_i))g'(x_i) - \frac{1}{2N}\sum_{i=1}^{N-1}(x_{i+1} - x_i)\tau_N(x_i)g''(c_i)$$

(recall that $\tau_N(x_N) = 0$). Taking absolute values, using (17) and (18), along with the symmetry/unimodality of the function Ψ_2, we get the result. $\qquad\square$

4. Convergence of multidimensional ground states

In this section we apply the results of the previous section to the N-point empirical distributions $\mathbb{P}_N$ of the d-dimensional ground states discussed in Sections 2.1 and 2.2.

These states have (signed) radial components given by the unique zero-median solution $x_1 > \cdots > x_R$ to the recursion (4) with $b(x) = |x|^k$, where $k = d - 1 \geq 1$ and $R = N^{1/2-1/(2d)}$ plays the role of N in (4), corresponding to R points in each of $N^{1/2+1/(2d)}$ radial directions. By translating the observations in Remark 3 into the d-dimensional setting, the optimal convergence rates of the radial part of $\mathbb{P}_N$ to the radial part of the d-dimensional normal distribution $\mathcal{N}_d$ are seen to be given by

$$d = 2, 3: \qquad d_W\big(\mathbb{P}_N^{\text{radial}}, \mathcal{N}_d^{\text{radial}}\big) \asymp \sqrt{\log R}/R;$$

$$d \geq 4: \qquad d_W\big(\mathbb{P}_N^{\text{radial}}, \mathcal{N}_d^{\text{radial}}\big) \asymp (\log R)^6/R^2.$$

The number of distinct polar angles in the directional part of the ground state (each with $d - 2$ components taking values in $[0, \pi/2]$ when $d \geq 3$, and a single component in $[0, \pi)$ when $d = 2$) is of order $N^{1/d}$. For $d \geq 3$, a larger number (namely $N^{1/2-1/(2d)}$) of azimuthal angles corresponding to each polar angle are available, so the polar part has the slower convergence rate. In general, the Wasserstein distance between the directional part of $\mathbb{P}_N$ and that of $\mathcal{N}_d$

is thus of order $N^{-1/d}$ (this follows easily using the representation of Wasserstein distance as the L_1 distance between two CDFs). For $d \geq 4$, in view of the upper bound in Lemma 1, namely a sum involving separate contributions from the the radial part and the directional parts, we conclude that the *overall rate* at which $\mathbb{P}_N$ tends to $\mathcal{N}_d$ is of order $N^{-1/d}$. For $d = 3$, the radial component involving $N^{1/3}$ points dominates, however, so the overall rate is of order $\sqrt{\log N}/N^{1/3}$. Similarly, the radial component involving $N^{1/4}$ points dominates in the case $d = 2$, so the overall rate is of order $\sqrt{\log N}/N^{1/4}$.

In summary, we obtain

$$d = 2: \qquad d_{\mathrm{W}}(\mathbb{P}_N, \mathcal{N}_d) \asymp \sqrt{\log N}/N^{1/4};$$

$$d = 3: \qquad d_{\mathrm{W}}(\mathbb{P}_N, \mathcal{N}_d) \asymp \sqrt{\log N}/N^{1/3};$$

$$d \geq 4: \qquad d_{\mathrm{W}}(\mathbb{P}_N, \mathcal{N}_d) \asymp N^{-1/d}.$$

Interestingly, the fastest rate of convergence to the ground state occurs in three dimensions.

For the excited states discussed in Section 2.3, the rates of convergence are of order $N^{-1/d}$; in this case the *directional* components dominate because the radial component has the faster rate (as seen from Remark 3 with $k = d + 1$ for $d = 2, 3$).

Acknowledgements

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

Supplementary material

In the Supplementary Material we provide complementary proofs, along with alternative developments of some Wasserstein bounds via a coupling argument. We also give source code and links to files containing the sequences. The Supplementary Material for this article can be found at https://doi.org/10.1017/jpr.2022.125.

References

[1] BRAUCHART, J. S. AND GRABNER, P. J. (2015). Distributing many points on spheres: Minimal energy and designs. *J. Complexity* **31**, 293–326.

[2] CHEN, L. H. AND THÀNH, L. V. (2023). Optimal bounds in normal approximation for many interacting worlds. *Ann. Appl. Prob.* **33**, 625–642

[3] CLAEYS, P. W. AND POLKOVNIKOV, A. (2021). Quantum eigenstates from classical Gibbs distributions. *SciPost Phys.* **10**, 014.

[4] DÖBLER, C. (2015). Stein's method of exchangeable pairs for the Beta distribution and generalizations. *Electron. J. Prob.* **20**, 1–34.

[5] ERNST, M., REINERT, G. AND SWAN, Y. (2020). First-order covariance inequalities via Stein's method. *Bernoulli* **26**, 2051–2081.

[6] ERNST, M. AND SWAN, Y. (2020). Distances between distributions via Stein's method. *J. Theoret. Prob.* **35**, 949–987.

[7] GHADIMI, M., HALL, M. J. W. AND WISEMAN, H. M. (2018). Nonlocality in Bell's theorem, in Bohm's theory, and in many interacting worlds theorising. *Entropy* **20**, 567.

[8] GOLDSTEIN, L. AND REINERT, G. (2013). Stein's method for the Beta distribution and the Pólya–Eggenberger urn. *J. Appl. Prob.* **50**, 1187–1205.

[9] HALL, M. J. W., DECKERT, D.-A. AND WISEMAN, H. M. (2014). Quantum phenomena modeled by interactions between many classical worlds. *Phys, Rev. X* **4** 041013.

[10] HERRMANN, H., HALL, M. J. W., WISEMAN, H. M. AND DECKERT, D. A. (2018). Ground states in the many interacting worlds approach. Preprint, arXiv:1712.01918.

[11] LEY, C. REINERT, G. and SWAN, Y. (2017). Distances between nested densities and a measure of the impact of the prior in Bayesian statistics. *Ann. Appl. Prob.* **27**, 216–241.

[12] LI, S. (2011). Concise formulas for the area and volume of a hyperspherical cap. *Asian J. Math. Statist.* **4**, 66–70.

[13] MCKEAGUE, I. W. AND LEVIN, B. (2016). Convergence of empirical distributions in an interpretation of quantum mechanics. *Ann. Appl. Prob.* **26**, 2540–2555.

[14] MCKEAGUE, I. W., PEKÖZ, E. A. AND SWAN, Y. (2019). Stein's method and approximating the quantum harmonic oscillator. *Bernoulli* **25**, 89–111.

[15] MARIUCCI, E. AND REISS, M. (2018). Wasserstein and total variation distance between marginals of Lévy processes. *Electron. J. Statist.* **12**, 2482–2514.