Published in *Phys. Rev. Lett.* **84** (2000), 4517

Hartree–Fock approximation for inverse many–body problems

J. C. Lemm and J. Uhlig
Institut für Theoretische Physik I, Universität Münster, 48149 Münster, Germany

A new method is presented to reconstruct the potential of a quantum mechanical many–body system from observational data, combining a nonparametric Bayesian approach with a Hartree–Fock approximation. *A priori* information is implemented as a stochastic process, defined on the space of potentials. The method is computationally feasible and provides a general framework to treat inverse problems for quantum mechanical many–body systems.

03.65.-w, 02.50-r, 21.60.Jz, 71.10.-w

The reconstruction of inter–particle forces from observational data is of key importance for any application of quantum mechanics to real world systems. Such inverse problems have been studied intensively in inverse scattering theory and in inverse spectral theory for one–body systems in one and, later, in three dimensions [1,2]. In this Letter we now outline a method, designed to deal with inverse problems for many–body systems.

Inverse problems are notoriously ill–posed [3]. It is well known that for ill–posed problems additional *a priori* information is required to obtain a unique and stable solution. In this Letter we refer to a nonparametric Bayesian framework [4,5] where *a priori* information is implemented explicitly, i.e., in form of stochastic processes over potentials [6].

Calculating an exact solution is typically not feasible for inverse many–body problems. As an example of a possible approximation scheme we will study in the following an 'Inverse Hartree–Fock Approximation' (IHFA). For situations where a Hartree–Fock (HF) ansatz is not sufficient, the inverse problem would have to be solved on top of other approximation schemes. A Random Phase Approximation or a full Time–Dependent Hartree–Fock Approximation [7–10], for example, would go beyond HF.

Bayesian methods can easily be adapted to different learning situations and have therefore been applied to a variety of empirical learning problems, including classification, regression, density estimation [11–13], and, recently, to quantum statistics [6]. In particular, using a Bayesian approach for quantum systems it is straightforward to deal with measurements of arbitrary quantum mechanical observables, to include classical noise processes, and to implement *a priori* information explicitly in terms of the potential.

To define the problem let us consider many–fermion systems with Hamiltonians, $H = T + V$, consisting of a one–body part T (e.g., in coordinate space representation $-(1/2m)\Delta$, with Laplacian Δ, mass m, $\hbar = 1$), and a two–body potential V. Introducing fermionic creation and annihilation operators $a^\dagger_\alpha$, a_α, corresponding to a complete single particle basis $|\varphi_\alpha>$, such Hamiltonians can be written,

$$H = \sum_{\alpha\beta} T_{\alpha\beta}\, a^\dagger_\alpha a_\beta + \frac{1}{4} \sum_{\alpha\beta\gamma\delta} V_{\alpha\beta\gamma\delta}\, a^\dagger_\alpha a^\dagger_\beta a_\delta a_\gamma , \qquad (1)$$

with $T_{\alpha\beta} = <\varphi_\alpha|T|\varphi_\beta>$ and antisymmetrized matrix elements $V_{\alpha\beta\gamma\delta} = <\varphi_\alpha\varphi_\beta|V|\varphi_\gamma\varphi_\delta>$. In particular, not to complicate the numerical calculations unnecessarily, we will study local two–body potentials depending only on the distance between two particles $z = |x_{(1)} - x_{(2)}|$, i.e., $V_{x_{(1)}x_{(2)}x'_{(1)}x'_{(2)}} = v(z)\left[\delta_{x_{(1)},x'_{(1)}}\delta_{x_{(2)},x'_{(2)}} - \delta_{x_{(1)},x'_{(2)}}\delta_{x_{(2)},x'_{(1)}}\right]$. (Being interested in numerical applications we will use discretized x variables.) Our aim is to reconstruct the function $v(z)$ from observational data.

In order to apply the Bayesian framework we need two model inputs: first, the probability $p(D|v)$ of measuring the observational data D given a potential v (for fixed D also known as *likelihood* of v), and, second, a *prior* probability $p(v)$ for v. The probability for v given data D, also called the *posterior* probability for v, is then obtained according to Bayes' rule,

$$p(v|D) \propto p(D|v)p(v). \qquad (2)$$

As solution of the reconstruction problem we will consider the potential with the maximal posterior probability (2). This approach, also known as *Maximum A Posteriori Approximation* (MAP), can be regarded as resulting from an approximation for the *predictive* probability $p(D_F|D) = \sum_v p(D_F|v)p(v|D)$ of future observations D_F given data D. The sum $\sum_v$, standing for a summation over all values $v(z)$ of the potential v (a functional integral for continuous z), can typically not be calculated exactly. Assuming that the main contribution to the sum comes from the potential with maximal posterior probability (MAP), i.e., $p(D_F|D) \approx p(D_F|v^*)$ where $v^* = \mathrm{argmax}_v p(v|D)$, we are left with maximizing (2). This can be done by setting the derivative of the posterior probability (or, technically often more convenient, its logarithm) with respect to $v(z)$ to zero.

As the first step, we have to identify the likelihood $p(D|v)$ of v for observational data D. In order to be able to obtain information about the potential, the system has to be prepared in a v–dependent state. Such a state can be a stationary statistical state, e.g. a canonical ensemble, or a time–dependent state evolving according

to the Hamiltonian of the system. In the following we will discuss many–body systems prepared in their ground state ψ_0. The (normalized) N–particle ground state wave function ψ_0 depends on v and is antisymmetrized for fermions. In particular, we will study two kinds of observational data $D = \{D_i | 1 \leq i \leq n\}$: (A) n simultaneous measurements of the coordinates of all N particles, (B) n measurements of the coordinates of a single particle. In case A, the ith measurement results in a vector $D_i = \vec{x}_i$, consisting of N components $x_{i(j)}$, each representing the coordinates of a single particle (which may also form a vector, e.g., a three dimensional one). According to the axioms of quantum mechanics, the probability of measuring the coordinate vector $\vec{x}_i$, given v, is,

$$p(\vec{x}_i|v) = |\psi_0(\vec{x}_i)|^2 = |\psi_0(x_{i(1)}, \cdots, x_{i(N)})|^2. \tag{3}$$

In case B, the probability for the single particle coordinates x to take the values $D_i = x_i$ is obtained by summing over the unobserved coordinates

$$p(x_i|v) = \sum_{x_{(2)}} \cdots \sum_{x_{(N)}} |\psi_0(x_i, x_{(2)}, \cdots, x_{(N)})|^2. \tag{4}$$

Similarly, it is for example straightforward to obtain the likelihoods for measuring inter–particle distances z, the particle number in a certain region, or particle momenta.

In contrast to an ideal measurement of a classical system, the state of a quantum system is typically changed by the measurement process. In particular, its state is projected in the space of eigenfunctions of the measured observable with eigenvalue equal to the measurement result. Hence, if we want to deal with independent, identically distributed data, the system must be prepared in the same state before each measurement. Under that condition the total likelihood factorizes

$$p(D|v) = \prod_{i=1}^{n} p(D_i|v). \tag{5}$$

As the second step, we have to choose a prior probability $p(v)$. A common and convenient choice are Gaussian prior probabilities (or, for functions, Gaussian processes), for which

$$p(v) = \left(\det \frac{\mathbf{K}_0}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{1}{2}<v-v_0|\mathbf{K}_0|v-v_0>}, \tag{6}$$

with positive (semi–)definite covariance ${\mathbf{K}_0}^{-1}$, and mean $v_0(z)$, playing the role of a reference potential. A typical choice for the inverse covariance is the (discretized) negative Laplacian multiplied with a 'regularization parameter' λ, i.e., $\mathbf{K}_0 = -\lambda\Delta$, favoring smooth potentials. Higher order differential operators are for example often included in $\mathbf{K}_0$ for regression problems to get differentiable regression functions Covariances can be constructed to implement general approximate symmetries, like for example approximate periodicity of v [6].

Having defined a likelihood $p(D|v)$ for many–body quantum systems and a prior probability $p(v)$ the next step is to solve the stationarity equation for the posterior probability (2)

$$0 = \partial_{v(z)} \ln p(v) + \sum_i \partial_{v(z)} \ln p(\vec{x}_i|v), \tag{7}$$

where we introduced the notation $\partial_{v(z)} = \partial/\partial v(z)$. The derivative of a Gaussian prior with respect to $v(z)$ is easily found as

$$\partial_v \ln p(v) = -\mathbf{K}_0(v - v_0), \tag{8}$$

∂_v denoting the gradient with respect to v, having components $\partial_{v(z)}$. To calculate the derivatives of the likelihoods (3) and (4) we need to know $\partial_{v(z)}\psi_0$. It is straightforward to show, by taking the derivative of $H\psi_0 = E_0\psi_0$ with respect to $v(z)$ that, for a nondegenerate ground state,

$$|\partial_{v(z)}\psi_0> = \sum_{\gamma\neq 0} \frac{1}{E_0 - E_\gamma} |\psi_\gamma><\psi_\gamma|\partial_{v(z)}H|\,\psi_0>, \tag{9}$$

with a complete basis of eigenstates ψ_γ, energies E_γ, and requiring $<\psi_0|\partial_{v(z)}\psi_0> = 0$ to fix norm and phase. Furthermore, from $\partial_{v(z)}v(z') = \delta_{z,|x'_{(1)} - x'_{(2)}|}$ directly follows

$$\partial_{v(z)}H = \frac{1}{2}\sum_x a_x^\dagger (a_{x-z}^\dagger a_{x-z} + a_{x+z}^\dagger a_{x+z}) a_x. \tag{10}$$

Typically, a direct solution of the many–body equation (9) is not feasible. To get a solvable problem we treat the many–body system in Hartree–Fock approximation [7,8,10] (For nonhermitian H see [14]). Thus, as the first step of an IHFA, we approximate the many–body Hamiltonian H by a one–body Hamiltonian h defined self–consistently by

$$h_{xx'} = T_{xx'} + \sum_{k=1}^{N} <x\varphi_k\,|\,V\,|\,x'\varphi_k>, \tag{11}$$

φ_k being the N–lowest orthonormalized eigenstates of h, i.e.,

$$h\varphi_k = \epsilon_k\varphi_k, \qquad \epsilon_1 \leq \epsilon_k \leq \epsilon_N. \tag{12}$$

The corresponding normalized N–particle Hartree–Fock ground state $|\Phi_0> = |\varphi_1, \cdots, \varphi_N>$ is the Slater determinant made of the N–lowest single particle orbitals φ_k. The Hartree–Fock likelihood replacing (3) becomes,

$$p_{\rm HF}(\vec{x}_i|v) = |\Phi_0(\vec{x}_i)|^2 = |\det B_i|^2, \tag{13}$$

defining the overlap matrix B_i with matrix elements $B_{kl;i} = \varphi_k(x_{i(l)})$. Analogously, the likelihood (4) becomes

$$p_{\rm HF}(x_i|v) = \sum_{k=1}^{N} |\varphi_k(x_i)|^2. \tag{14}$$

From (13) follows $\partial_{v(z)} p_{\rm HF}(\vec{x}_i|v) = \Phi_0(\vec{x}_i)\partial_{v(z)}\Phi_0^*(\vec{x}_i) + \Phi_0^*(\vec{x}_i)\partial_{v(z)}\Phi_0(\vec{x}_i)$, so that expanding the determinant $\det B_i = \sum_l^N M_{kl;i}B_{kl;i}$ in terms of the cofactors, $M_{kl,i} = (B_i^{-1})_{lk}\det B_i$, yields

$$\partial_{v(z)}\Phi_0(\vec{x}_i) = \sum_{kl}^{N} M_{kl;i}\,\partial_{v(z)}\varphi_k(x_{i(l)}). \qquad (15)$$

Hence, to obtain the derivatives of the HF likelihoods (13) or (14) with respect to $v(z)$, we have to find $\partial_{v(z)}\varphi_k$.

Again, we proceed by taking the derivative of Eq. (12) and obtain after standard manipulations (for nondegenerate ϵ_k and $<\varphi_k|\partial_{v(z)}\varphi_k> = 0$),

$$|\partial_{v(z)}\varphi_k> = \sum_{l\neq k}\frac{1}{\epsilon_k-\epsilon_l}|\varphi_l><\varphi_l|\partial_{v(z)}h|\varphi_k>. \qquad (16)$$

Furthermore, from Eq. (11) follows

$$\begin{aligned}\partial_{v(z)}h_{xx'} = \sum_{j=1}^{N}\Big(&<x\,\varphi_j\,|\,\partial_{v(z)}V\,|\,x'\,\varphi_j> \\ &+ <x\,\partial_{v(z)}\varphi_j\,|\,V\,|\,x'\,\varphi_j> \\ &+ <x\,\varphi_j\,|\,V\,|\,x'\,\partial_{v(z)}\varphi_j>\Big). \end{aligned} \qquad (17)$$

Therefore, the derivative of the orbitals is found by inserting Eq. (17) into Eq. (16)

$$\begin{aligned}\partial_{v(z)}\varphi_k(x) = \sum_{l\neq k}\frac{\varphi_l(x)}{\epsilon_k-\epsilon_l}\sum_{j=1}^{N}\Big(&<\varphi_l\varphi_j\,|\,\partial_{v(z)}V\,|\,\varphi_k\varphi_j> \\ &+ <\varphi_l\partial_{v(z)}\varphi_j\,|\,V\,|\,\varphi_k\varphi_j> \\ &+ <\varphi_l\varphi_j\,|\,V\,|\,\varphi_k\partial_{v(z)}\varphi_j>\Big). \end{aligned} \qquad (18)$$

That "inverse HF equation" (18) can quite effectively be solved by iteration, starting for example with initial guess $\partial_{v(z)}\varphi_j(x) = 0$. Because $\partial_{v(z)}\varphi_k(x)$ depends on z and x, Eq. (18), being the central equation of the IHFA, has the dimension of a two–body equation for the lowest N orbitals. Introducing, analogously to B_i, the matrix $\Delta_i(z)$ with $\Delta_{kl;i}(z) = \partial_{v(z)}\varphi_k(x_{i(l)})$ (for $z > 0$, $1 \le l \le N$, $1 \le k \le N$, $1 \le i \le n$), it is straightforward to check that the derivatives of the logarithm of the HF likelihoods (13) and (14) are given by

$$\partial_{v(z)}\ln p_{\rm HF}(\vec{x}_i|v) = 2{\rm Re}\left[{\rm Tr}(B_i^{-1}\Delta_i(z))\right], \qquad (19)$$

$$\partial_{v(z)}\ln p_{\rm HF}(x_i|v) = 2\frac{{\rm Re}\left[\sum_{k=1}^{N}\varphi_k^*(x_i)\partial_{v(z)}\varphi_k(x_i)\right]}{\sum_{l=1}^{N}|\varphi_l(x_i)|^2}. \qquad (20)$$

The stationarity equation (7) can now be solved by iteration, for example,

$$v^{(r+1)} = v^{(r)} + \eta\big[v_0 - v^{(r)} + \mathbf{K}_0^{-1}\sum_i \partial_v \ln p_{\rm HF}(\vec{x}_i|v^{(r)})\big], \qquad (21)$$

choosing a positive step width η and starting from an initial guess $v^{(0)}$.

In conclusion, reconstructing a potential from data by IHFA is based on the definition of a *prior probability* for v and requires the iterative solution of

1. the *stationarity equation for the potential* (7), needing as input for each iteration step (21)

2. the *derivatives of the likelihoods* (19) or (20), respectively, obtained by solving the (two–body–like) equation (18) for given

3. *single particle orbitals*, defined in (12) as solutions of the direct (one–body) Hartree–Fock equation.

We tested the numerical feasibility of the IHFA for a Hamiltonian $H = T + V_1 + V$ including a local two–body potential V to be reconstructed and a local one–body potential V_1, with diagonal elements $v_1(x) = 0$ for $-a \le x \le a$ and $v_1(x) = \infty$ elsewhere, to confine the particles and to break translational symmetry.

To check the validity of the IHFA we must be able to sample artificial data from the exact true many–body likelihood (3) for a given true potential $V_{\rm true}$. Because this requires to solve the corresponding many–body problem exactly, we have chosen a two–particle system with one–dimensional x (on a grid with 21 points, $a = 10$) for which the true ground state can be calculated by diagonalizing H numerically. We want to stress, however, that, in the case of real data, application of the IHFA to systems with $N > 2$ particles is straightforward and only requires to solve Eq. (18) for N instead for two orbitals.

We selected a local true two–body potential $V_{\rm true}$ with diagonal elements $v_{\rm true}(z) = b/(1 + e^{-2\gamma(z-L/2)/L})$ and parameter values ($b = 1$, $\gamma = 10$, $L = 2a + 1 = 21$ for mass $m = 0.1$) for which the iteration of the HF equation (12) converges. (For two–body systems the HF iteration leads easily to oscillations.) Calculating the true ground state ψ_0 for $V_{\rm true}$ and the corresponding true likelihoods (3) and (4) we then sampled: (A) $n = 100$ two–particle data from the true likelihood (3) and (B) $n = 100$ single particle data from the true likelihood (4). (See Figs. 1a and 2a.)

The calculations have been done for a Gaussian prior probability $p(v)$ as in (6) with $\lambda(\mathbf{I} - \Delta)/2$ (with identity $\mathbf{I}$, $\lambda = 5$) as inverse covariance $\mathbf{K}_0$, and a reference potential $v_0(z)$ of the form of $v_{\rm true}$, but with $\gamma = 1$ (so it becomes nearly linear in the shown interval.) Furthermore, we have set all potentials to zero at the origin and constant beyond the right boundary. The reconstructed potential $v_{\rm IHFA}$ has then been obtained by iterating according to Eq. (21) and solving Eqs. (12) and (18) within each iteration step.

The resulting IHFA likelihoods $p(\vec{x}_i|v_{\rm IHFA})$ (case A, Fig. 1a) or $p(x_i|v_{\rm IHFA})$ (case B, Fig. 2a) did indeed fit well the true likelihoods $p(\vec{x}_i|v_{\rm true})$ or $p(x_i|v_{\rm true})$, respectively. (Fig. 1a shows instead of the two–dimensional $p(\vec{x}_i|v)$ for vectors $\vec{x}_i$ the one–dimensional $p(z|v) = \sum_x [p(x, x-z|v) + p(x, x+z|v)]$ for the inter–particle

distance z.) In particular, in case A the reconstructed likelihood $p(z|v_{\rm IHFA})$ is over the whole range an improvement over the reference likelihood $p(z|v_0)$, while in case B the IHFA solution $p(x_i|v_{\rm IHFA})$ is nearly exactly the same as the true likelihood $p(x_i|v_{\rm true})$. That perfect result for case B is due to the fact that reconstructing the likelihood $p(x_i|v)$ for single particle data is a much simpler task than reconstructing the full $p(\vec{x}_i|v)$.

The situation is more complex for potentials (Figs. 1b and 2b). Firstly, one sees that the correlation information contained in the two–particle data of case A yields a better reconstruction for $v_{\rm IHFA}$ than the less informative single particle data of case B. In both cases, however, the true potential is only well approximated at medium inter–particle distances. For large and small distances, on the other hand, the IHFA solution is still dominated by the reference potential of the prior probability $p(v)$. This effect is a consequence of the lack of empirical data in those regions: The probability to find particles at large distances is small, because the true potential has its maximum at large distances. Also, measuring small distances is unlikely, because antisymmetry forbids two fermions to be at the same place. In such low data regions one must therefore rely on *a priori* information.

We are grateful to A. Weiguny for stimulating discussions.

[1] R. Newton, *Inverse Schrödinger Scattering in Three Dimensions* (Springer Verlag, Berlin, 1989).

[2] K. Chadan, D. Colton, L. Päivärinta, and W. Rundell, *An Introduction to Inverse Scattering and Inverse Spectral Problems* (SIAM, Philadelphia, 1997).

[3] A. Tikhonov and V. Arsenin, *Solution of Ill–posed Problems* (Wiley, New York, 1977).

[4] D. Sivia, *Data Analysis: A Bayesian Tutorial* (Oxford University Press, Oxford, 1996).

[5] Introductions to Bayesian statistics can be found at: http://bayes.wustl.edu/.

[6] J. Lemm, J. Uhlig, and A. Weiguny, Phys. Rev. Lett. **84**, 2068 (2000).

[7] J. Eisenberg and W. Greiner, *Microscopic Theory of the Nucleus* (North–Holland, Amsterdam, 1972).

[8] P. Ring and P. Schuck, *The Nuclear Many–Body Problem* (Springer Verlag, New York, 1980).

[9] K. Goeke, R. Cusson, F. Grümmer, P.-G. Reinhard, and H. Reinhardt, Prog. Theor. Phys. (Suppl.) **74 & 75**, 33 (1983).

[10] J.-P. Blaizot and G. Ripka, *Quantum Theory of Finite Systems* (The MIT Press, Cambridge, MA, 1986).

[11] C. K. I. Williams and C. E. Rasmussen, in *Advances in Neural Information Processing Systems 8*, edited by D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo (The MIT Press, Cambridge, MA, 1996), pp. 514–520.

[12] C. Williams and D. Barber, IEEE Trans. on Pattern Analysis and Machine Intelligence **20**, 1342 (1998).

[13] J. Lemm, Technical Report No. MS-TP1-99-1, Univ. of Münster, physics/9912005.

[14] J. C. Lemm, Annals of Physics **244**, 136 (1995).

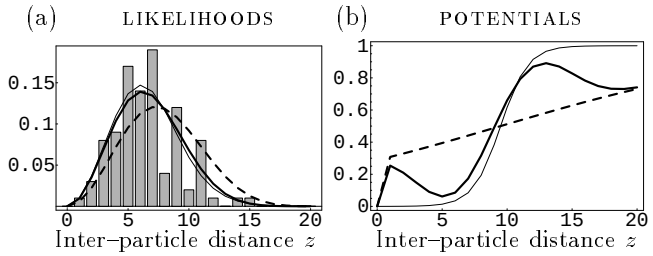

FIG. 1. Two–particle data (case A): (a) True likelihood $p(z|v_{\rm true})$ (thin line), empirical probability (relative frequency) $p_{\rm emp}(z) = (1/n)\sum_{i=1}^{n}\delta_{z,|x_{i(1)}-x_{i(2)}|}$ (bars) obtained from the $n = 100$ sampled two–particle data $\vec{x}_i$, reference likelihood $p(z|v_0)$ (dashed), and reconstructed likelihood $p(z|v_{\rm IHFA})$ (thick line) as function of the inter–particle distance $z = |x_{(1)} - x_{(2)}|$. (b) True potential $v_{\rm true}(z)$ (thin line), reference potential $v_0(z)$ (dashed), and reconstructed IHFA potential $v_{\rm IHFA}(z)$ (thick line) as function of the inter–particle distance z.

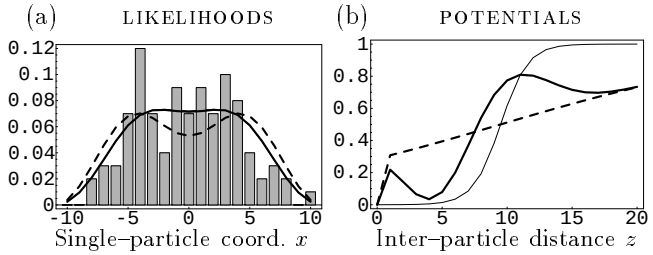

FIG. 2. Single particle data (case B): (a) True likelihood $p(x|v_{\rm true})$ (occluded by thick line), empirical probability $p_{\rm emp}(x) = (1/n)\sum_{i=1}^{n}\delta_{x,x_i}$ (bars) obtained from the $n = 100$ sampled single particle data x_i, reference likelihood $p(x|v_0)$ (dashed), and reconstructed likelihood $p(x|v_{\rm IHFA})$ (thick line) as function of single particle coordinates x. (b) True potential $v_{\rm true}(z)$ (thin line), reference potential $v_0(z)$ (dashed), and reconstructed IHFA potential $v_{\rm IHFA}(z)$ (thick line) as function of the inter–particle distance z.