To appear in Physics Letters A

Bayesian Approach to Inverse Time–Dependent Quantum Mechanics

J. C. Lemm*
Institut für Theoretische Physik, Universität Münster
48149 Münster, Germany

Abstract

Using a new Bayesian method for solving inverse quantum problems, potentials of quantum systems are reconstructed from time series obtained by coordinate measurements in non–stationary states. The approach is based on two basic inputs: 1. a likelihood model, providing the probabilistic description of the measurement process as given by the axioms of quantum mechanics, and 2. additional *a priori* information implemented in form of stochastic processes over potentials.

The first step to be done when applying quantum mechanics to a real world system is the reconstruction of its Hamiltonian from observational data. Such a reconstruction, also known as inverse problem, constitutes a typical example of empirical learning [1, 2, 3]. Whereas the determination of potentials from spectral and from scattering data has been studied in much detail in inverse spectral and inverse scattering theory [4, 5, 6], this Letter describes the reconstruction of potentials by measuring particle positions in coordinate space for finite quantum systems in *time–dependent* states. The presented method can easily be generalized to other forms of observational data.

In the last years much effort has been devoted to practical empirical learning in the form of regression, classification, or density estimation problems. Examples include, just to name a few, prediction of financial time series,

*E-mail: lemm@uni-muenster.de, WWW: http://pauli.uni-muenster.de/~lemm

medical diagnosis, and image or speech recognition. This lead to a variety of new learning algorithms, many of those being in principle also applicable to inverse quantum problems. In particular, we refer in the following to the *Bayesian* framework [7, 8, 9]. In the context of quantum mechanics the Bayesian framework has been used for the reconstruction of density operators [10, 11, 12] and the reconstruction of potentials for stationary quantum systems at finite temperature [13] and stationary many–body systems at zero temperature [14]. In this Letter a Bayesian approach is developed to solve problems of inverse time–dependent quantum mechanics (ITDQ).

To be specific, we study the following type of time series data generated by a quantum system: Preparing a particle in an eigenstate of the position operator with coordinates x_0 at time t_0, we let this state evolve in time according to the rules of quantum mechanics and measure its new position at time t_1, finding a value x_1. Continuing from this measured position x_1, we measure the particle position again at time t_2, and repeat this procedure until n data points x_i at times t_i have been collected. We thus end up with observational data of the form $D = \{(x_i, \Delta_i, x_{i-1}) | 1 \leq i \leq n\}$, where x_i is the result of the i-th coordinate measurement, $\Delta_i = t_i - t_{i-1}$ the time interval between two subsequent measurements and x_{i-1} the coordinates of the previous observation (or preparation) at time t_{i-1}. In particular, we consider systems with a time–independent Hamiltonian of the form $H = T + V$, consisting of a standard kinetic energy term T and a local potential $V(x, x') = \delta(x - x')v(x)$. In that case, our aim is the reconstruction of the function $v(x)$ from observational data D. (The restriction to local potentials simplifies the numerical calculations. Nonlocal Hamiltonians can be reconstructed similarly.)

Being interested in predictions of future events empirical learning requires *a priori* information to generalize from a finite number of observational data, i.e., to define the link between available and future data. At least implicitly, *a priori* information is therefore built–in in every learning algorithm. It is a special advantage of Bayesian approaches to deal with *a priori* information explicitly. This is especially important if the number of data is small relative to the complexity of the learning task. Technically, *a priori* information encompasses all data which are not considered part of the actual observational data, including information on the stochastic process which generates the function $v(x)$ from which data are sampled. Empirical measurement of *a priori* information can be related to *a posteriori* control of the generalization behavior at the time of testing the theory [15]. In practice, *a priori* informa-

tion typically represents qualitative knowledge, for example smoothness or approximate periodicity of the potential to be reconstructed. This practical interpretation of *a priori* information will be adopted in the following.

Setting up a Bayesian model requires the definition of two probabilities (or, for continuous variables, probability densities, respectively): 1. the probability $p(D|v)$ to measure data D given potential v, for D considered fixed known as the *likelihood* of v, and 2. a prior probability $p(v)$ implementing *a priori* information available for the potential to be reconstructed. Referring to a *maximum a posteriori approximation* (MAP) we will understand those potentials v to be solutions of the reconstruction problem which maximize $p(v|D)$, i.e., the *posterior* probability of v given all available data D [15]. The basic relation to be employed is then Bayes' theorem, according to which

$$p(v|D) \propto p(D|v)p(v). \tag{1}$$

Choosing a parametric ansatz for the potential v, a (nonuniform) prior term $p(v)$ is usually not necessary, so the MAP becomes a maximum likelihood approximation. In the following, however, we concentrate on nonparametric approaches which are less restrictive than their parametric counterparts. Their large flexibility makes it then essential to include (nonuniform) priors. Such priors can be formulated explicitly in terms of the function $v(x)$ and are conveniently chosen as a Gaussian process [16]. Gaussian process priors, which we use for ITDQ in the following, have also been used for nonparametric regression [17], classification [18], general density estimation [15], and stationary inverse quantum problems [13, 14]. (A Bayesian MAP approach is often technically equivalent to a classical regularization approach [16]. The interpretation of regularization and prior terms however differs slightly [19, 20].)

It is the specific likelihood model discussed next which defines ITDQ. According to the axioms of quantum mechanics the probability density that a particle is found at position x_i at time t_i, provided the particle has been at x_{i-1} at time t_{i-1}, is given by [24, 25]

$$p_i = p(x_i|\Delta_i, x_{i-1}, v) = |\phi_i(x_i)|^2, \tag{2}$$

where

$$\phi_i(x_i) \;=\; \langle x_i \,|\, \phi_i \rangle \;=\; \langle x_i \,|\, U_i \, x_{i-1} \rangle, \tag{3}$$

are matrix elements of the time evolution operator

$$U_i = \exp\left(-i\Delta_i H\right), \tag{4}$$

setting $\hbar = 1$. The transition amplitudes (3) can be calculated by inserting orthonormalized eigenstates ψ_α of H, with energies E_α,

$$\phi_i(x_i) = \sum_\alpha \exp\left(-i\Delta_i E_\alpha\right) \psi_\alpha(x_i)\psi_\alpha^*(x_{i-1}), \tag{5}$$

assuming H has a completely discrete spectrum, which can always be achieved by imposing zero or periodic boundary conditions.

Clearly, it is straightforward to modify (2) for measuring observables different from the particle position. It is also interesting to note that the transition probabilities (2) define a Markoff process with $W_i(x \to x') = p(x'|\Delta_i, x, v)$. For real Hamiltonians with real boundary conditions (in coordinate representation) which have real eigenfunctions $\psi_\alpha(x)$ [21], they obey the relation $W_i(x \to x') = W_i(x' \to x)$. It follows that the detailed balance condition, $p_{\rm stat}(x)W_i(x \to x') = p_{\rm stat}(x')W_i(x' \to x)$, is fulfilled for a uniform $p_{\rm stat}(x)$ which therefore represents the stationary state of the Markoff process of repeated position measurements.

Having defined the likelihood model of ITDQ (2), in the next step a prior for v has to be chosen. A convenient nonparametric prior $p(v)$ is a Gaussian

$$p_G(v) = \left(\det \frac{\mathbf{K}_0}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2}\langle v - v_0 \,|\, \mathbf{K}_0 \,|\, v - v_0\rangle\right), \tag{6}$$

with (real symmetric, positive semi–definite) inverse covariance $\mathbf{K}_0$, acting in the space of potentials, and mean $v_0(x)$ which can be considered as a reference potential for v. Typical examples are smoothness constraints on v which correspond to choosing differential operators for $\mathbf{K}_0$. Reference potentials can be made more flexible by allowing parameterized families $v_0(x;\theta)$. Within the context of Bayesian statistics such additional parameters θ are known as hyperparameters [7, 9, 22, 23]. In MAP approximation the optimal hyperparameters are determined by maximizing the posterior (1) simultaneously with respect to θ and $v(x)$ [26]. A simplified procedure consists in using a parametric approximation $v(\theta)$ which maximizes the likelihood $\prod_i p_i(\theta)$ as reference potential v_0 for the nonparametric reconstruction $v(x)$ [15].

If available, it is useful to include some information about the ground state energy $E_0(v)$ which helps to determine the depth of the potential. This can, for example, be a noisy measurement of the ground state energy which, assuming Gaussian noise, is implemented by

$$p_E \propto \exp\left(-\frac{\mu}{2}\left(E_0(v) - \kappa\right)^2\right). \tag{7}$$

Combining (6) and (7) with (2) for n repeated coordinate measurements starting from initial position x_0 we obtain for the posterior (1),

$$p(v|D) \propto p_G(v)\, p_E(v) \prod_{i=1}^{n} p_i. \tag{8}$$

To calculate the MAP solution v^* =argmax$_v p(v|D)$ we set the functional derivative (Fréchet derivative) of the posterior (8), or technically more convenient of its logarithm, with respect to v to zero,

$$0 = \delta_v \ln p(v|D) = \delta_v \ln p_G(v) + \delta_v \ln p_E(v) + \sum_i \delta_v \ln p_i, \tag{9}$$

where

$$\begin{aligned} \delta_v \ln p_G(v) &= -\mathbf{K}_0\,(v - v_0), & (10)\\ \delta_v \ln p_E(v) &= -\mu(E_0(v) - \kappa)\,\delta_v E_0(v), & (11)\\ \delta_v \ln p_i &= 2\mathrm{Re}[\phi_i^{-1}(x_i)\,\delta_v \phi_i(x_i)]. & (12) \end{aligned}$$

The functional derivative $\delta_v \phi_i$ can according to Eq. (5) be obtained from $\delta_v \psi_\alpha$. The still required $\delta_v \psi_\alpha$ and $\delta_v E_\alpha$ can then be found by calculating the functional derivative of the eigenvalue equation $H\psi_\alpha = E_\alpha \psi_\alpha$, i.e.,

$$(E_\alpha - H)|\,\delta_v \psi_\alpha\rangle = (\delta_v H - \delta_v E_\alpha)|\,\psi_\alpha\rangle. \tag{13}$$

Multiplying from the left with $\langle\psi_\alpha\,|$, using the conjugate eigenvalue equation and $\delta_{v(x)} v(x') = \delta(x - x')$ (where $\delta_{v(x)} = \langle x\,|\,\delta_v\,\cdot\,\rangle$ denotes the x component of the functional derivative δ_v) we get for normalized eigenfunctions,

$$\delta_{v(x)} E_\alpha = \langle\psi_\alpha\,|\,\delta_{v(x)} H\,|\,\psi_\alpha\rangle = |\psi_\alpha(x)|^2. \tag{14}$$

To solve Eq. (13) for $\delta_v \psi_\alpha$ we have to note that all eigenfunctions ψ_γ with $E_\gamma = E_\alpha$ are zero modes of the operator $E_\alpha - H$. For the sake of simplicity we assume nondegenerate eigenvalues, the generalization to the degenerate case being straightforward [27]. First, we notice that the problem is solvable because Eq. (14) ensures that the r.h.s. of (13) has no components in the null space of $E_\alpha - H$. Second, it is easy to verify that the Moore–Penrose inverse of $E_\alpha - H$ is $G_\alpha = \sum_{\gamma\neq\alpha} |\,\psi_\gamma\rangle\langle\psi_\gamma\,|/(E_\alpha - E_\gamma)$, i.e., $G_\alpha(E_\alpha - H)G_\alpha = G_\alpha$. Third, we choose the arbitrary contributions from the null space of $E_\alpha - H$

to be zero, i.e., $\langle \delta\psi_\alpha \,|\, \psi_\alpha \rangle = 0$. This can be achieved for real $\psi_\alpha(x)$ by fixing the norm, i.e., by requiring $\delta_v \langle \psi_\alpha \,|\, \psi_\alpha \rangle = 0$. For complex $\psi_\alpha(x)$ also phases must be controled. With $G_\alpha \psi_\alpha = 0$ we finally obtain $\delta_v \psi_\alpha = G_\alpha(\delta_v H)\psi_\alpha$, or, written in components

$$\delta_{v(x)}\psi_\alpha(x') = \sum_{\gamma\neq\alpha} \frac{1}{E_\alpha - E_\gamma}\, \psi_\gamma(x')\psi_\gamma^*(x)\psi_\alpha(x). \tag{15}$$

Collecting the results it follows from Eq. (5),

$$\begin{aligned}
\delta_{v(x)}\phi_i(x_i) \;=\; & \delta_{v(x)}\langle x_i \,|\, U_i\, x_{i-1}\rangle \;= \\
& \sum_\alpha \exp\left(-i\Delta_i E_\alpha\right)\Big[\left(-i\Delta_i |\psi_\alpha(x)|^2\right)\right) \psi_\alpha(x_i)\psi_\alpha^*(x_{i-1}) \\
& + \sum_{\gamma\neq\alpha} \frac{1}{E_\alpha - E_\gamma} \psi_\gamma(x_i)\psi_\gamma^*(x)\psi_\alpha(x)\psi_\alpha^*(x_{i-1}) \\
& + \sum_{\gamma\neq\alpha} \frac{1}{E_\alpha - E_\gamma} \psi_\gamma^*(x_{i-1})\psi_\gamma(x)\psi_\alpha^*(x)\psi_\alpha(x_i)\Big].
\end{aligned} \tag{16}$$

Inserting Eq. (14) for $\alpha = 0$ in Eq. (11) and Eq. (16) in Eq. (12) a MAP solution for the potential v can be found for example by iterating the stationarity equation (9) numerically on a lattice [15, 27, 28]. Clearly, such a straightforward discretization can only be expected to work for a low–dimensional x variable. Higher dimensional systems usually require additional approximations [14].

As the next step, we want to check the numerical feasibility of a non-parametric reconstruction of the potential v for a one–dimensional quantum system. For that purpose, we choose a system with the true potential

$$v_{\text{true}}(x) = \frac{c_1}{\sqrt{2\pi\sigma}} \exp \frac{(x - c_2)^2}{2\sigma^2}, \tag{17}$$

where $c_1 = -10$, $c_2 = -2$, and $\sigma = 2$. An example of the time evolution of an unobserved particle in the potential v_{true} is shown in Fig. 1. As input for the reconstruction algorithm 50 data points x_i are sampled from the corresponding true likelihoods $p(x|\Delta_i, x_{i-1}, v_{\text{true}})$. A corresponding path of an observed particle is shown in Fig. 2.

Besides a noisy energy measurement of the form (7) we include a truncated Radial Basis Function prior [29], i.e., a Gaussian prior (6) with smooth-

ness related inverse covariance

$$\mathbf{K}_0(x,x') = \delta(x-x')\lambda \sum_{k=0}^{3} (-1)^k \frac{\sigma_0^{2m}}{k!2^k} \left(\frac{\partial^2}{\partial x^2}\right)^k . \tag{18}$$

To obtain an adapted reference potential v_0 for the Gaussian prior, a parameterized potential of the form

$$v_0(a,b,c) = \min[0, a(x-b)^2 + c], \tag{19}$$

is optimized with respect to a, b, c by maximizing the "extended likelihood" $\sum_i \ln p_i(v_0) + \ln p_E(v_0)$. Using a parametric reference potential corresponds to the prior expectation that the potential to be reconstructed has a form similar (where similarity is defined by the inverse covariance) to (19). This exemplifies how the results of parametric reconstructions can be used as input for a more flexible nonparametric approach. Finally, the stationarity equation (9) is solved by iterating according to

$$\begin{aligned} v^{(r+1)} &= v^{(r)} + \eta \Big[v_0 - v^{(r)} + \\ &\quad \mathbf{K}_0^{-1} \{ 2 \sum_i^n \mathrm{Re}[\delta_v \ln \phi_i(x_i)] + \delta_v (\ln p_G + \ln p_E) \} \Big], \end{aligned} \tag{20}$$

optimizing the step width η by a line search algorithm. The resulting nonparametric ITDQ solution v_{ITDQ} (see Fig. 3) is a reasonable reconstruction of v_{true} and clearly better than the best parametric approximation $v_0(a,b,c)$. It is only the flat area near the right border where due to missing and unrepresentative data the reconstruction differs significantly from the true potential.

Fig. 4 compares the sum over empirical transition probabilities $\widetilde{p}_{\mathrm{emp}}(x) = \frac{1}{n}\sum_{i=1}^n \delta(x-x_i)$ as derived from the observational data D with the corresponding true $p_{\mathrm{true}} = \frac{1}{n}\sum_{i=1}^n p(x|\Delta_i, x_{i-1}, v_{\mathrm{true}})$ and reconstructed $p_{\mathrm{ITDQ}} = \frac{1}{n}\sum_{i=1}^n p(x|\Delta_i, x_{i-1}, v_{\mathrm{ITDQ}})$. Due to the summation over data points with different x_{i-1}, the quantities shown in Fig. 4 do not present the complete information which is available to the algorithm. Hence, Fig. 5 depicts the corresponding quantities for a fixed x_{i-1}. In particular, Fig. 5 compares the reconstructed transition probability (2) with the corresponding empirical and true transition probabilities for a particle having been at time t_{i-1} at position x_{i-1} =1. The ITDQ algorithm returns an approximation for all such transition probabilities.

Figs. 4 and 5 show that the reconstructed v_{ITDQ} tends to produce a better approximation of the empirical probabilities than the true potential v_{true}. Indeed, the error on the data or negative log–likelihood $\epsilon_D(v) = -\sum_i \ln p_i(v)$, being a canonical error measure in density estimation, is smaller for v_{ITDQ} than for v_{true}. A smaller λ, i.e., a lower influence of the prior, produces a still smaller error $\epsilon_D(v_{\mathrm{ITDQ}})$. At the same time, however, the reconstructed potential becomes more wiggly for smaller λ, being the symptom of the well known effect of "overfitting". On the other hand, the (true) generalization error $\epsilon_g(v) = -\int dx\, dx' p(x) p(x'|x, v_{\mathrm{true}}) \ln p(x'|x, v)$ [with uniform $p(x)$] can never be smaller for the reconstructed v_{ITDQ} than for v_{true}. As it is typical for most empirical learning problems, the generalization error $\epsilon_g(v_{\mathrm{ITDQ}})$ shows a minimum as function of λ. It is this minimum which gives the optimal value for λ. Knowledge of the true model allows in our case to calculate the generalization error exactly. If, as usual, the true model is not known, classical cross–validation [16] and bootstrap [30] techniques can be used to approximate the generalization error as function of λ empirically. Alternatively to optimizing λ or other hyperparameters one can integrate over them [26]. Similarly, studying the feasibility of a Bayesian Monte Carlo approach, contrasting the MAP approach of this paper, would certainly be interesting.

In summary, in this Letter a Bayesian approach has been developed to solve the inverse problem for time–dependent quantum systems. The approach is able to handle quite general types of observational data and allows a flexible and explicit implementation of *a priori* information, the latter representing for instance data collected in the past or qualitative knowledge about the potential to be reconstructed. Classical examples of *a priori* information are smoothness and other approximate symmetries. Numerical results have been presented for a one dimensional model.

References

[1] A. N. Tikhonov, Soviet Math. Dokl. 4 (1963) 1035–1038

[2] A. Tarantola, Inverse Problem Theory — Methods for Data Fitting and Model Parameter Estimation, Elsevier, Amsterdam, 1987.

[3] V. N. Vapnik, Statistical Learning Theory Wiley, New York, 1998.

[4] K. Chadan and P. C. Sabatier, Inverse Problems in Quantum Scattering Theory, Springer Verlag, New York, 1989.

[5] R. Snieder and A. Tarantola, Phys. Rev. A 39 (1989) 3303–3309

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

[7] J. O. Berger, Statistical Decision Theory and Bayesian Analysis (2nd ed.), Springer Verlag, New York, 1985.

[8] J. M. Bernado and A. F. Smith, Bayesian Theory, Wiley, New York, 1994.

[9] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, Bayesian Data Analysis, Chapman & Hall, New York, 1995.

[10] M. Tan, J. Mod. Opt. 44 (1997) 2233

[11] V. Bužek, G. Drobný, R. Derka, G. Adam, and H. Wiedemann, quant-ph/9805020.

[12] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A, 61 (1999)

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

[14] J. C. Lemm and J. Uhlig, Phys. Rev. Lett. 84 (2000) 4517

[15] J. C. Lemm, Bayesian Field Theory, Münster University, Technical Report No. MS-TP1-99-1, physics/9912005, 1999.

[16] G. Wahba, Spline Models for Observational Data, SIAM, Philadelphia, 1990.

[17] C. K. I. Williams and C. E. Rasmussen, in NIPS 8, edited by D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, The MIT Press, Cambridge, MA, 1996, pp. 514–520.

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

[19] J. C. Lemm, Prior Information and Generalized Questions, Massachusetts Institute of Technology, Artificial Intelligence Laboratory Memo No.1598, 1996.

[20] W. P. Gouveia and J. A. Scales, Inverse Problems 13 (1997) 323–349

[21] A. Messiah, Quantum Mechanics, Vol. II, North–Holland, Amsterdam, 1961.

[22] B. P. Carlin, T. A. Louis, Bayes and Empirical Bayes Methods for Data Analysis, Chapman & Hall/CRC, Boca Raton, 1996.

[23] J. C. Lemm, J. Uhlig, and A. Weiguny, Bayesian Reconstruction of Approximately Periodic Potentials for Quantum Systems at Finite Temperature, Münster University, Technical Report No. MS-TP1-00-4, quant-ph/0005122, 2000.

[24] C.W. Helstrom, Quantum Detection and Estimation Theory. Academic Press, New York, 1976.

[25] A.S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory. North–Holland, Amsterdam, 1982.

[26] D. J. C. MacKay, Hyperparameters: optimise or integrate out? in Maximum Entropy and Bayesian Methods, Santa Barbara 1993, edited by G. Heidbreder, Kluwer, Dordrecht, 1994.

[27] J. C. Lemm and J. Uhlig, Münster University,, Technical Report No. MS-TP1-99-15, quant-ph/0006027 (to appear in Few Body Physics) 1999.

[28] D. P. Bertsekas, Nonlinear Programming, Athena Scientific, Belmont, Mass. 1995.

[29] F. Girosi, M. Jones, and T. Poggio, Neural Computation 7 (1995) 219

[30] B. Efron and R. J. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.

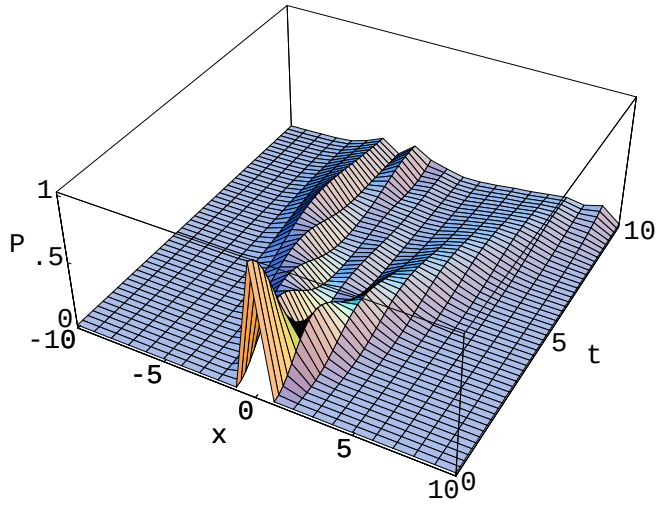

Figure 1: Time evolution of an unobserved particle with mass $m = 1$ started at time $t_0 = 0$ from $x_0 = 0$ in the potential (17). Shown is the transition probability $p(x|\Delta t = t, x_0 = 1, v_{\text{true}})$.

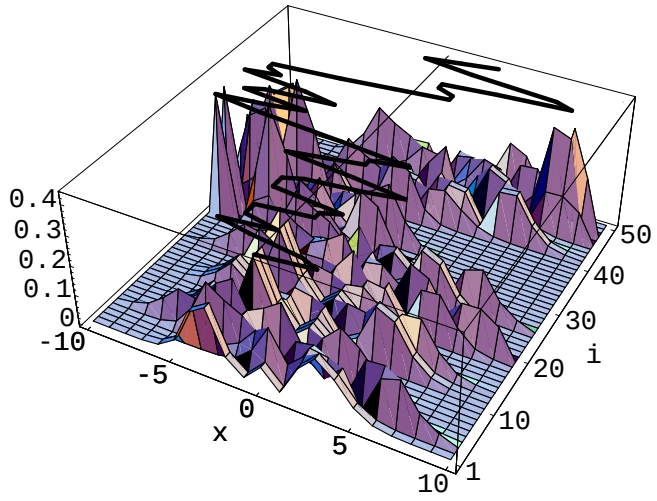

Figure 2: Time evolution of an observed particle with mass $m = 1$ in the potential (17). The figure shows for each data point i the probability $p(x|\Delta_i = 5, x_{i-1}, v_{\text{true}})$, starting from $x_0 = 0$. (Hence, the probability at i =1 corresponds to that shown in Fig. 1 at $t = 5$.) The actual data points x_i have been sampled from that probabilities and form the observed path shown on top as a thick line.

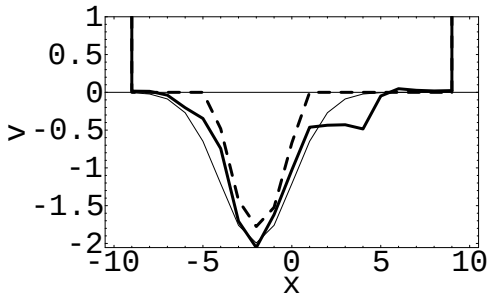

Figure 3: Numerical reconstruction of a potential from 50 coordinate measurements (see Fig. 2). Shown are the true potential $v_{\rm true}$ (thin line), the best parametric approximation used as reference potential v_0 (dashed line), and the reconstructed potential $v_{\rm ITDQ}$ (thick line). Parameters: $m = 1$, $\Delta_i = 5$, $v_{\rm true}$ of Eq. (17), v_0 of the form (19), Gaussian prior with $\mathbf{K}_0$ as in (18) with $\lambda = 0.1$ and $\sigma_0 = 3$, p_E (7) with $\mu = 10$ and $\kappa = E_0(v_{\rm true})$, periodic boundary conditions for ψ_α, fixed boundary values $v(-10) = v(10) = 10^5$ for $v_{\rm ITDQ}$, v_0 and $v_{\rm true}$, calculated on a lattice with 21 points. Errors: $\epsilon_D(v_{\rm ITDQ}) = 99.1$, $\epsilon_D(v_{\rm true}) = 104.4$, $\epsilon_g(v_{\rm ITDQ}) = 1.891$, $\epsilon_g(v_{\rm true}) = 1.818$.

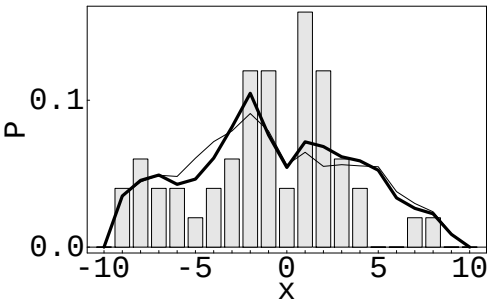

Figure 4: Sum of empirical transition probabilities $\widetilde{p}_{\rm emp}$ (bars), the corresponding true $p_{\rm true}$ (thick line) for $v_{\rm true}$, and the reconstructed $p_{\rm ITDQ}$ (thin line).

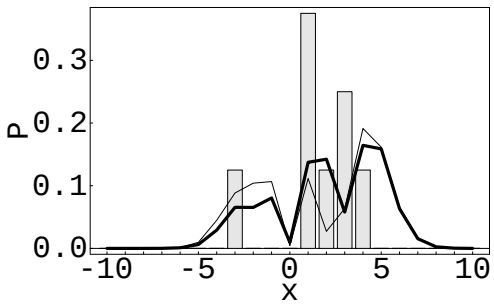

Figure 5: Same functions as in Fig. 4, but restricted to measurements of a particle which has been at position $x_{i-1} = 1$ at the time of the previous measurement.