Inverse Quantum Statistics

Jörg Lemm

Joerg Uhlig

Achim Weiguny

Institut für Theoretische Physik I
Universität Münster

Email: lemm@uni-muenster.de
Web: http://pauli.uni-muenster.de/~lemm

Bayesian Approach

Variables:

a. v: Potential to be reconstructed (hidden variables, state of "Nature")

b. O: Measured observable

c. x: Measurement results

d. D: Data, consisting of training data $D_T = (x_T, O_T)$ and a priori information D_0.

Probability densities:

1. likelihood model: $p(x|O, v)$

2. prior model: $p(v|D_0)$

3. posterior: $p(v|D) \propto p(x_T|O_T, v)\, p(v|D_0)$

4. predictive: $p(x|O, D) = \int dv\, p(x|O, v)\, p(v|D)$

Likelihood and Prior Model

likelihood model	prior model
describes	
measurement process	generalization behavior
is determined by	
parameters	hyperparameters
Examples include	
density estimation regression classification inverse quantum theory	hard constraints Gaussian prior factors mixtures of Gaussians non–quadratic potentials

Likelihood Model of Quantum Mechanics

The probability of obtaining the value x when measuring the observable O for a quantum mechanical system in a state described by density operator ρ is given by

$$p(x|O,v) = \mathrm{Tr}(P_O\ \rho(v))$$

where $P_O(x) = \sum_\xi |x,\xi\rangle\langle x,\xi|$ is the projector on the space spanned by the orthonormalized eigenfunctions $|x,\xi>$ of O with eigenvalue x. (The label ξ distinguishes eigenfunctions which are degenerated with respect to O.)

For a *canonical ensemble* the density operator is given by

$$\rho = \frac{\exp(-\beta H)}{\mathrm{Tr}\ \exp(-\beta H)}$$

(with $H = T + V$ and setting $\hbar = 1$.)

Example: *Coordinate measurements*

$$p(x|\hat{x},v) = \sum_\alpha p_\alpha |\phi_\alpha(x)|^2 =< |\phi(x)|^2 >$$

with $H|\phi_\alpha> = E_\alpha|\phi_\alpha>$, $p_\alpha = e^{-\beta E_\alpha}/Z$, $Z = \sum_\alpha e^{-\beta E_\alpha}$.

Maximum a Posteriori Approximation

Calculate predictive density

$$p(x|O, D) = \int dv \, p(x|O, v) \, p(v|D)$$

in *Maximum a Posteriori Approximation* (MAP), i.e.,

$$p(x|O, D) \approx p(x|O, v^*)$$

with

$$\begin{aligned} v^* &= \mathrm{argmax}_v \, p(v|D) \\ &= \mathrm{argmax}_v \, p(v|D_0) \prod_{i=1}^{N} p(x_i|O_i, v) \\ &= \mathrm{argmax}_v \; \ln p(v|D_0) + \sum_{i=1}^{N} \ln p(x_i|O_i, v) \end{aligned}$$

This requires to solve (assuming a local potential, i.e., $V(x, x') = v(x)\delta(x - x')$)

$$0 = \delta_{v(x)} \ln p(v|D)$$

where

$$\delta_{v(x)} \ln p(v|D) = \frac{\delta \ln p(v|D)}{\delta v(x)}$$

Functional Derivative of the Likelihood

The functional derivative of a likelihood term with respect to $v(x)$ becomes

$$\delta_{v(x)} p(x_i|\hat{x}, v) =$$

$$\frac{\delta}{\delta v(x)} \left(\frac{e^{-\beta E_\alpha} |\phi_\alpha(x_i)|^2}{\sum_{\alpha'} e^{-\beta E_{\alpha'}}} \right) =$$
$$< \phi^*(x_i) \delta_{v(x)} \phi(x_i) > + < \delta_{v(x)} \phi^*(x_i) \phi(x_i) >$$
$$-\beta \left(< \delta_{v(x)} E |\phi(x_i)|^2 > - < \delta_{v(x)} E >< |\phi(x_i)|^2 > \right)$$

Problem: How can we find $\delta_{v(x)} \phi(x_i)$ and $\delta_{v(x)} E_\alpha$?

Functional Derivative of the Likelihood

Solution: Take functional derivative of the Schrödinger equation $H|\phi_\alpha> = E_\alpha|\phi_\alpha>$,

$$\left(\delta_{v(x)}H\right)|\phi_\alpha> + H\delta_{v(x)}|\phi_\alpha>$$

$$= \left(\delta_{v(x)}E_\alpha\right)|\phi_\alpha> + E_\alpha\delta_{v(x)}|\phi_\alpha>$$

One finds for orthonormalized, nondegenerated eigenfunctions,

$$\begin{aligned}
\delta_{v(x)}E_\alpha &= <\phi_\alpha|\delta_{v(x)}H|\phi_\alpha> = |\phi_\alpha(x)|^2 \\
\delta_{v(x)}\phi_\alpha(x') &= \sum_{\gamma\neq\alpha}\frac{1}{E_\alpha - E_\gamma}\phi_\gamma(x')\phi^*_\gamma(x)\phi_\alpha(x)
\end{aligned}$$

using

$$\delta_{v(x)}H(x',x'') = \delta_{v(x)}V(x',x'') = \delta(x-x')\delta(x'-x'')$$

Prior processes

Example: *Gaussian prior process*

$$p_0(v) = \left(\det \frac{\lambda \mathbf{K}_0}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{\lambda}{2} < v - v_0 | \mathbf{K}_0 | v - v_0 >}$$

with mean v_0 (reference potential) and covariance operator $\mathbf{K}_0$.

Example: *Mixture of Gaussian prior processes*

$$p_0(v) = \sum_{k=1}^{M} p_0(k)\, p_0(v|k)$$

with mean v_0 (reference potential) and covariance operator $\mathbf{K}_0$.

Functional derivative of prior with respect to $v(x)$

$$\delta_{v(x)} \ln p_0 = -\lambda \mathbf{K}_0 (v - v_0)$$

For mixture model $\mathbf{K}_0 = \sum_k p_0(k|v)\, \mathbf{K}_k$ and $v_0 = \mathbf{K}_0^{-1} \sum_k p_0(k|v)\, \mathbf{K}_k v_k$ where $p_0(k|v) = p_0(v|k) p_0(k)/p_0(v)$.

Covariances and Approximate Symmetries

For an approximately invariant potential

$$V \approx S^\dagger V S = \mathbf{S}V$$

a natural prior would be $p_0 \propto \exp(-E_S)$ with

$$E_S = < V - \mathbf{S}V | V - \mathbf{S}V >/2 = < V|\mathbf{K}_0|V >/2$$

where $\mathbf{K}_0 = (\mathbf{I}-\mathbf{S})^\dagger(\mathbf{I}-\mathbf{S})$, $\mathbf{I}$ = identity. Similarly, for a Lie group $\mathbf{S}(\theta) = \exp(\theta\mathbf{s})$ this yields $\mathbf{K}_0 = \mathbf{s}^\dagger\mathbf{s}$.

Example 1: *Smoothness = approximate invariance under infinitesimal translations*

$\mathbf{s} = d/dx \Rightarrow \mathbf{K}_0 = -\Delta$ (negative Laplacian).

Smoother potentials follow from Radial Basis Function prior $\mathbf{K}_0 = \exp\left(-\sigma^2_{\rm RBF}\Delta/2\right)$.

Example 2: *Approximate periodicity*

A. Periodic covariance:

$\mathbf{K}_0 = \Delta_\theta$ with $<v|\Delta_\theta|v> = \int dx |v(x) - v(x+\theta)|^2$.

B. Periodic mean:

$E_S = < V - V_0 | V - V_0 > /2$ for $V_0 = S^\dagger V_0 S$.

Energy measurements

Sometimes it is useful to include besides smoothness also a priori information or data which are related more directly to the depth of the potential. One such possibility is to include information about the average energy

$$U = < E > = \sum_\alpha p_\alpha E_\alpha$$

The average energy can then be controled by introducing a Lagrange multiplier term $\mu(U - \kappa)$, or, technically sometimes easier, by a term representing noisy energy data κ,

$$p_U \propto e^{-E_U}, \quad E_U = \frac{\mu}{2}(U - \kappa)^2$$

so that $U \to \kappa$ for $\mu \to \infty$. Using

$$\delta_{v(x)} U = < \delta_{v(x)} E > - \beta < E \, \delta_{v(x)} E > + \beta < E > < \delta_{v(x)} E >$$

one finds for the functional derivative of E_U

$$\delta_{v(x)} E_U = \mu \, (U - \kappa) < |\phi(x)|^2 \, (1 - \beta \, (E - U)) >$$

Functional Derivative of the Posterior

The functional derivative of the posterior with Gaussian prior process and energy penalty

$$0 = \delta_{v(x)} \ln p(v|D)$$

$$= \delta_{v(x)} \left(\sum_i \ln p(x_i|\hat{x}, v) - \frac{\lambda}{2} < v - v_0 | \mathbf{K}_0 | v - v_0 > - E_U \right)$$

can now be calculated by collecting:

1. the functional derivative of the likelihood

$$\delta_{v(x)} p(x_i|\hat{x}, v) =$$
$$< \left(\delta_{v(x)} \phi^*(x_i) \right) \phi(x_i) > + < \phi^*(x_i) \delta_{v(x)} \phi(x_i) >$$
$$-\beta \left(< |\phi(x_i)|^2 |\phi(x)|^2 > - < |\phi(x_i)|^2 > < |\phi(x)|^2 > \right)$$

2. the functional derivative of the prior

$$\delta_{v(x)} \ln p_0 = -\lambda \mathbf{K}_0 (v - v_0)$$

3. the functional derivative of the energy penalty term

$$\delta_{v(x)} E_U = \mu \left(U - \kappa \right) < |\phi(x)|^2 \left(1 - \beta \left(E \ - U \right) \right) >$$

Iteration Procedures = Learning Algorithms

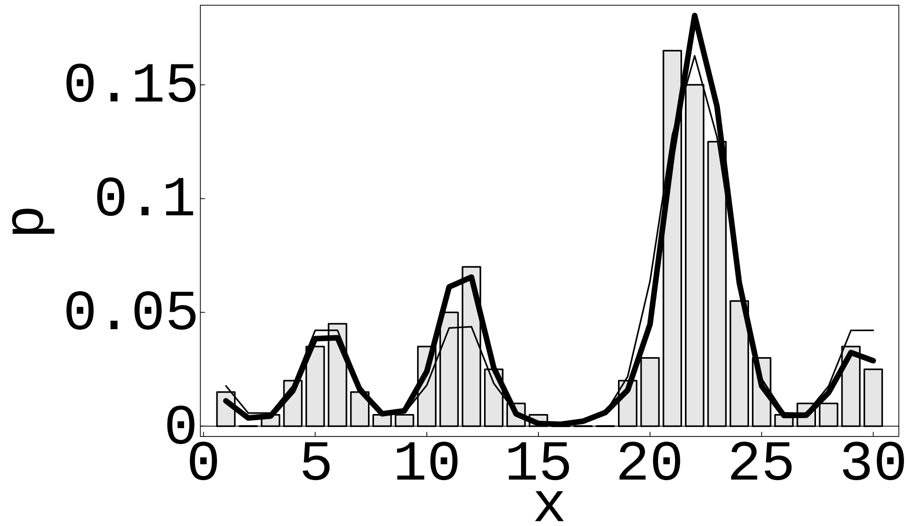
Starting with an initial guess $v^{(0)}$, choosing a step width η and a positive definite matrix $\mathbf{A}$ the stationarity equation can be solved by iterating according to

$$
\begin{aligned}
v^{(r+1)} &= v^{(r)} + \eta \mathbf{A}^{-1} (\mathbf{K}_0^{(r)} (v_0 - v^{(r)}) \\
&\quad + \sum_i \delta_{v(x)} \ln p(x_i | \hat{x}, v^{(r)}) - \delta_{v(x)} E_U^{(r)})
\end{aligned}
$$

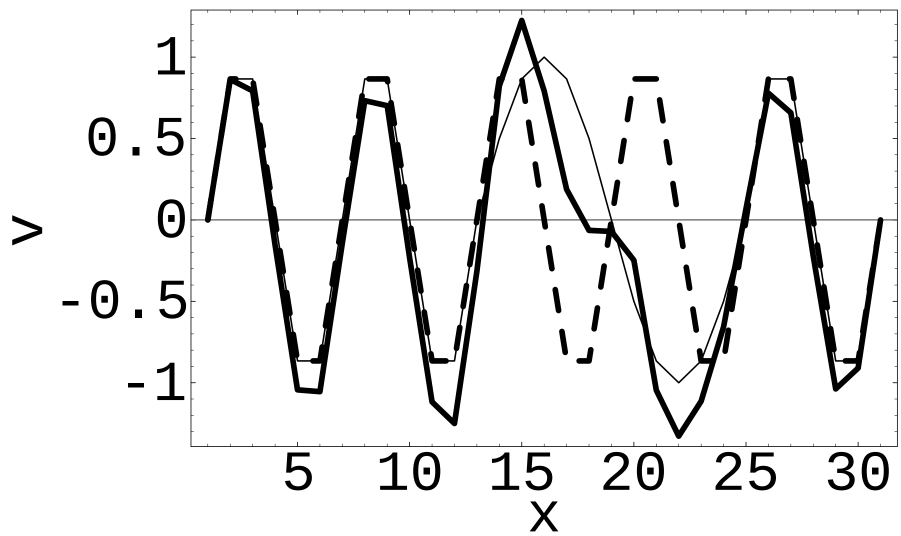
Typically, η and often also $\mathbf{A}$ are adapted during iteration. In the numerical examples we have studied, $\mathbf{A} = \lambda \mathbf{K}_0$ proved to be a good choice. Note that in general different initial guesses $v^{(0)}$ can yield different solutions.

Approximately Periodic Potentials

(a) Likelihoods



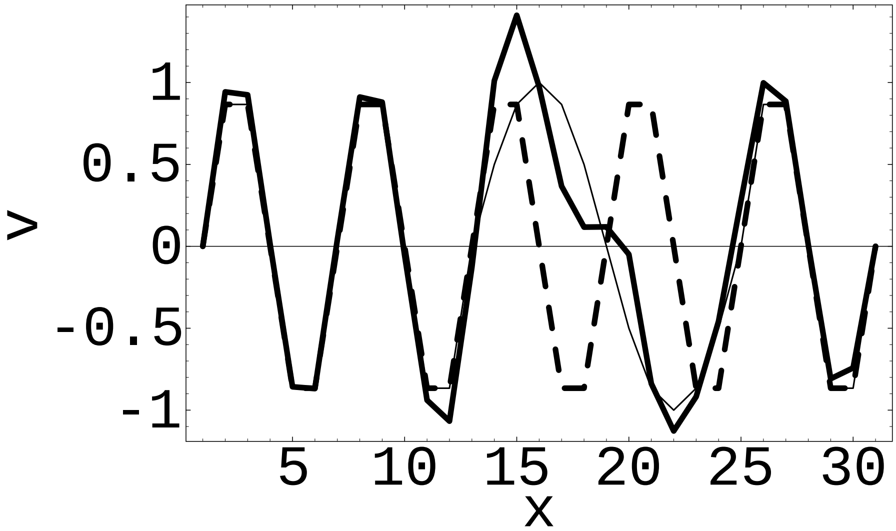
(b) Potentials



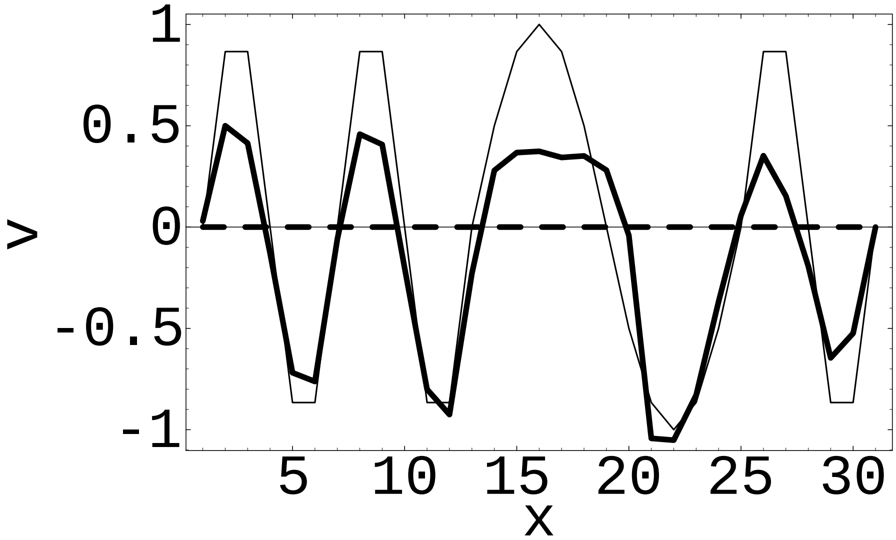
(a) Empirical density, i.e., relative frequencies of sampled data (bars), true (thin line) and reconstructed likelihood (thick line) (b) $v_{\rm true}$ (thin line) and reference potential $v_0 = \sin(\pi x/3)$ (dashed), reconstructed potential v (thick line). (200 data points, mesh with 30 points, m = 0.25 for $\hbar$ = 1, β = 4, $\mathbf{K}_0 = -\Delta$, λ = 0.2, μ = 0, $U(v_{\rm true}) = -0.354$, $U(v) = -0.552$.)

Reconstruction of Approximately Periodic Potentials with Energy penalty

(a) With periodic reference potential



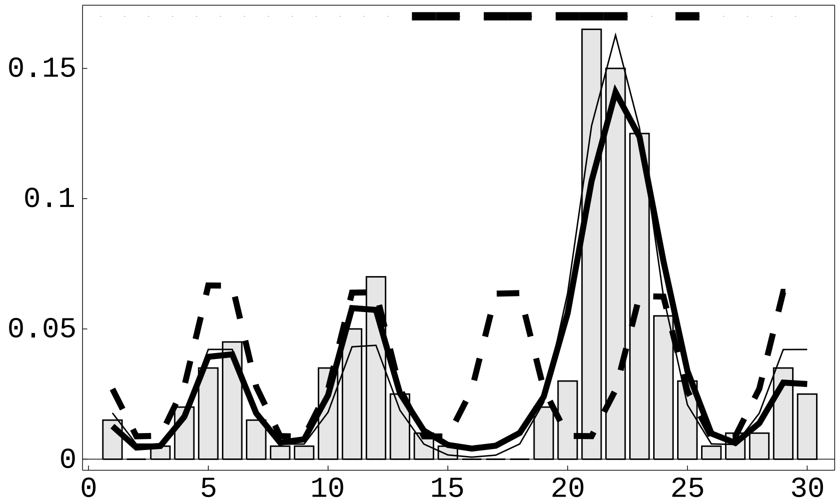
(b) With periodic prior covariance



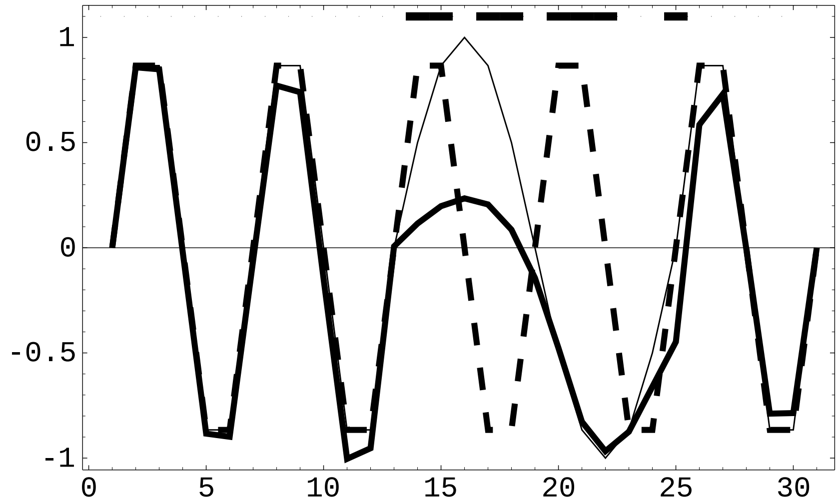
Reconstruction with energy penalty term E_U with $\kappa = U(v_{\rm true}) = -0.354$. (a) $\mathbf{K}_0 = -\Delta$, $v_0(x) = \sin(\pi x/3)$, $\mu = 1000$, $U(v) = -0.353$. (b) $\mathbf{K}_0 = -\Delta - \gamma\Delta_\theta$, $v_0 \equiv 0$, $\theta = 6$, $\gamma = 0.12$, $\lambda = 0.05$, $\mu = 10$, $U(v) = -0.351$.

Local Hyperparameters

(a) Likelihoods



(b) Potentials



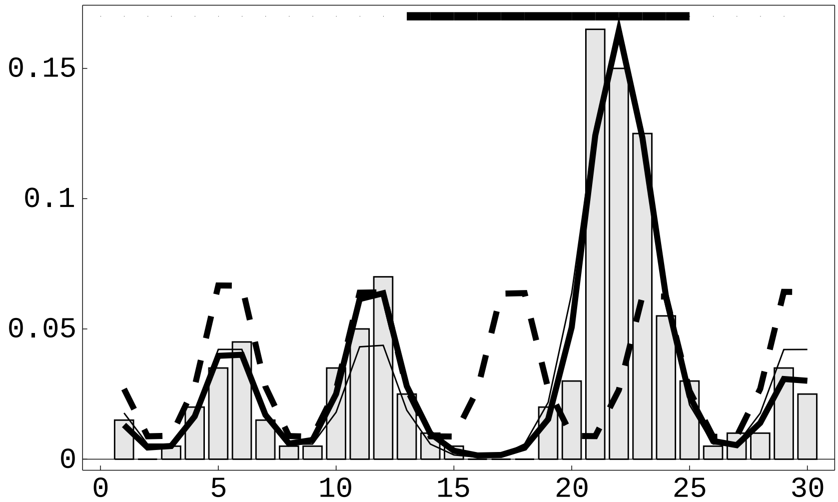
Reference potential is locally switched off if $(v(x) - v_0(x))^2 > \vartheta$, Laplace prior is kept,

$$E_0 = \frac{\lambda}{2} \int dx \, (v(x) - v_0(x))^2 \Theta \left(\vartheta - (v(x) - v_0(x))^2 \right) + \frac{\gamma}{2} < v|\Delta|v > + \frac{\mu}{2}(U - \kappa)^2$$

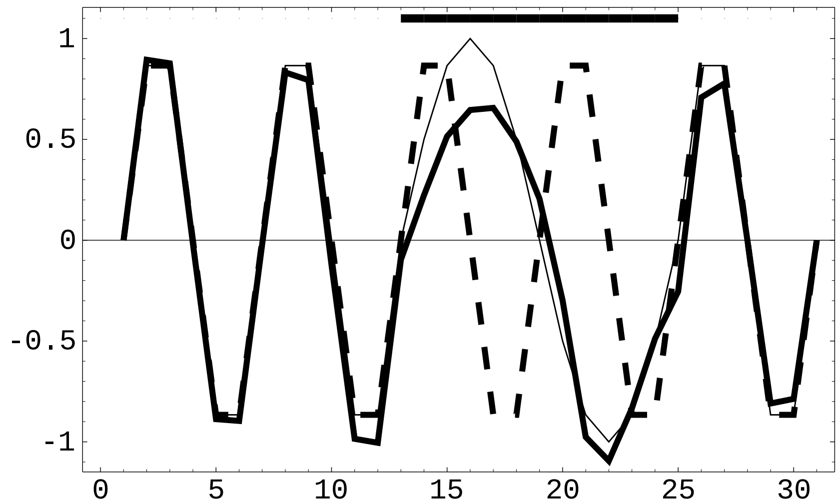
Empirical density $p_{\rm emp}$ (bars), true $p_{\rm true}$, $v_{\rm true}$ (thin), reference p_0, v_0 (dashed), solutions p, v (thick).

Local Hyperparameters

(a) Likelihoods



(b) Potentials



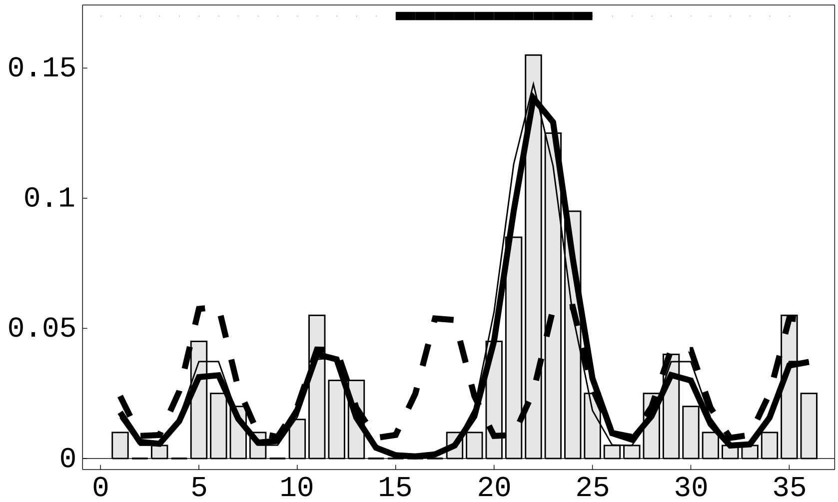
Switching between non–zero reference v_1 and $v_2 = 0$ including Laplacian prior.

$$E_0 = \frac{\lambda}{2}\int dx\,(\omega_1(x))^2(1-B(x)) + \int dx\,(\omega_2(x))^2 B(x) + \frac{\mu}{2}(U-\kappa)^2$$

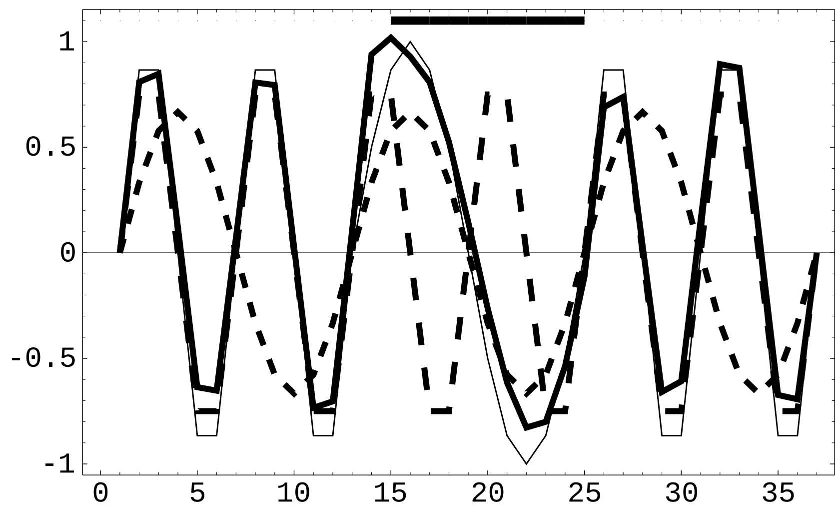
with $B(x) = \Theta(\omega_1(x) - \omega_2(x))$ and $\omega_1(x) = (\partial/\partial x)(v(x) - v_1(x))$, $\omega_2(x) = (\partial/\partial x)(v(x) - v_2(x))$.

Local Hyperparameters

(a) Likelihoods



(b) Potentials



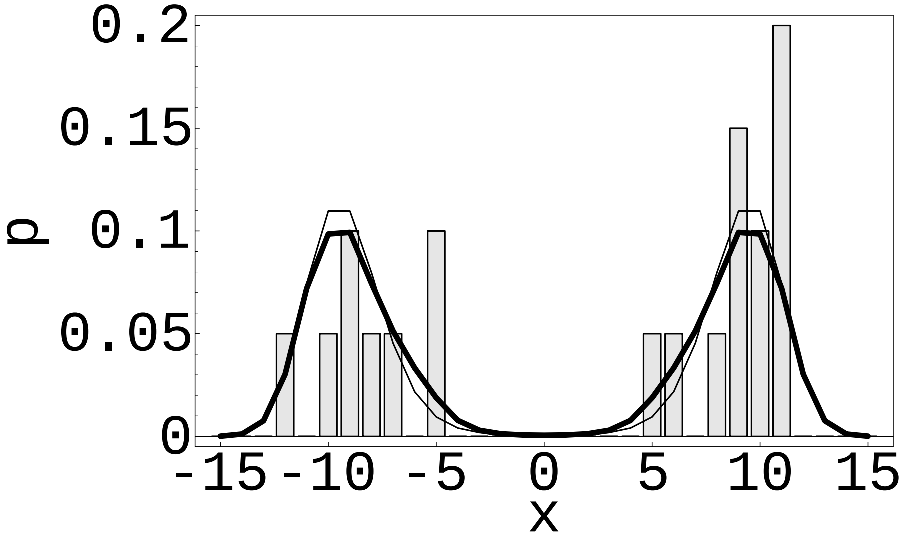
Switching between nonzero references v_1 and v_2 with additional penalty for discontinuities in $B(x) = \Theta\left(\omega_1(x) - \omega_2(x)\right)$,

$$E_0 = \frac{\lambda}{2}\int dx\,\left(\omega_1(x)\right)^2(1 - B(x)) + \int dx\,\left(\omega_2(x)\right)^2 B(x) + \frac{\mu}{2}(U - \kappa)^2 - \ln p(B)$$

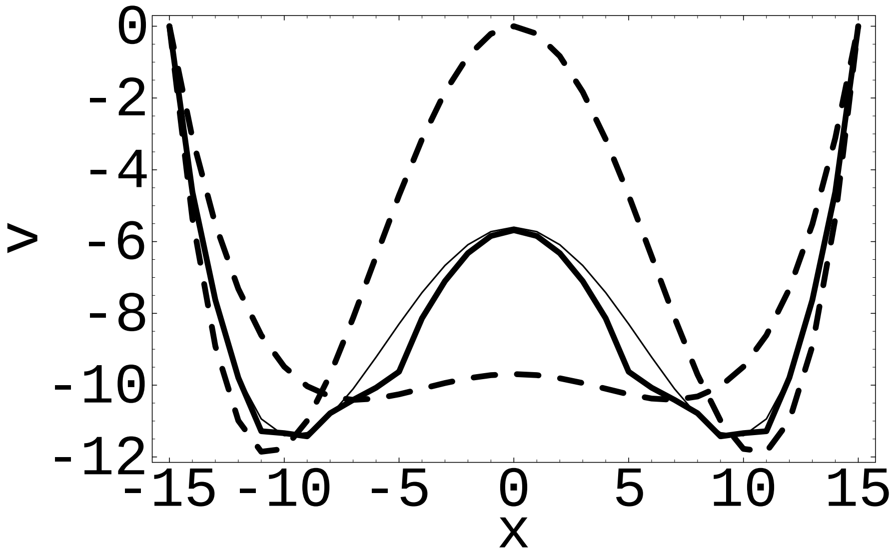
with hyperprior $\ln P(B)$ counting the discontinuities of B. Optimized using "simulated annealing".

Prior Mixture Models

(a) Likelihoods



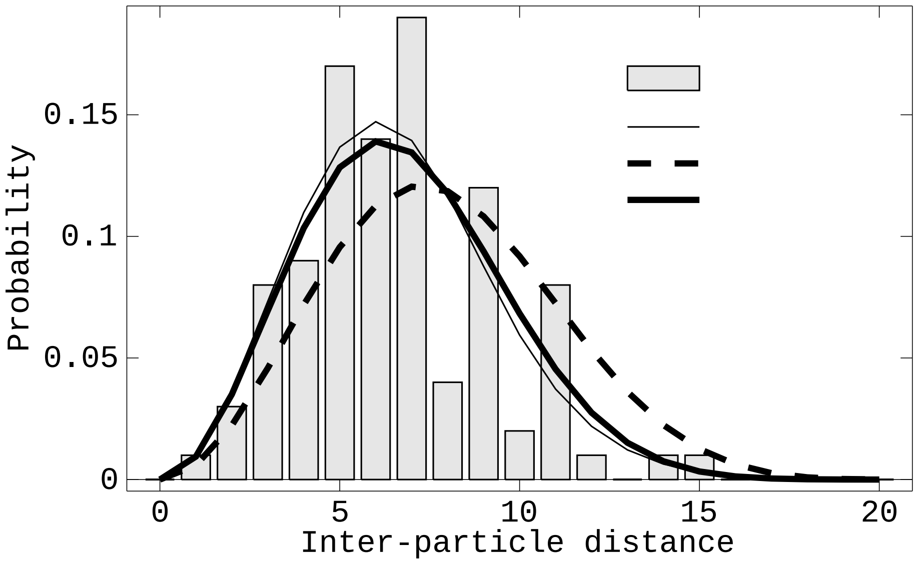
(b) Potentials



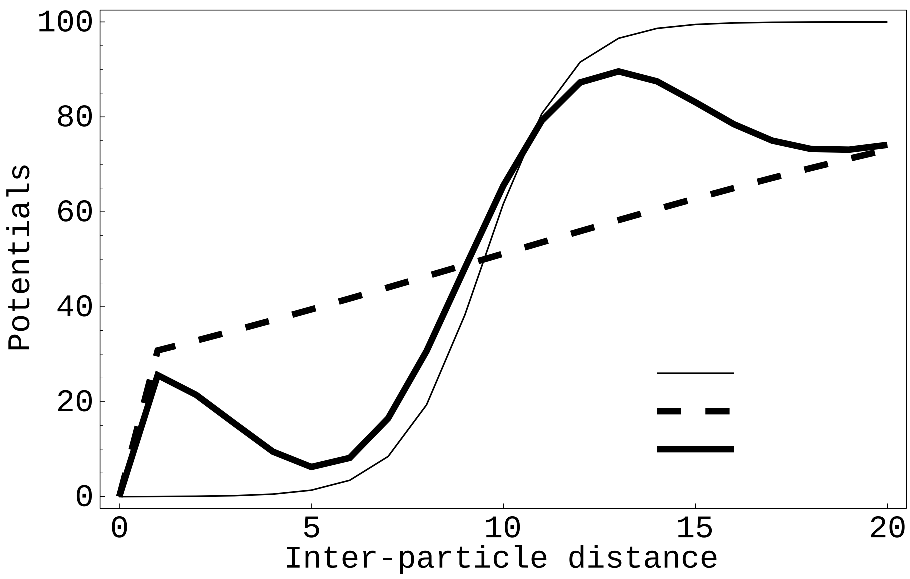
Final mixture coefficients $p_0(1|v) = 0.57$, $p_0(2|v) = 0.43$, 20 data points, mesh with 31 points, $v(x) = v(-x)$, boundary condition $v(\pm 15) = 0$, and $m = 0.1$, $\beta = 1$, $\mathbf{K}_0 = -\Delta$, $\lambda = 0.5$, $\mu = 10$, $\kappa = -10, 01 = U(v_{\mathrm{true}})$ resulting in $U(v) = -10.1$, 50 iterations, prior iteration.

Inverse Hartree–Fock Theory

(a) Likelihoods



(b) Potentials



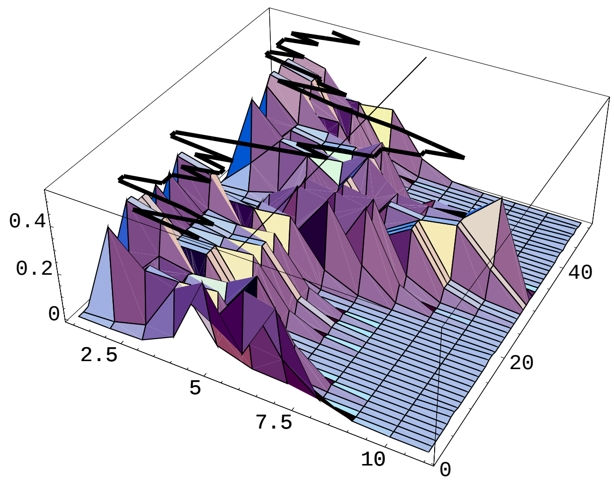
True potential v_{true}, reference potential v_0, and reconstructed IHFA potential as function of inter–particle distance.

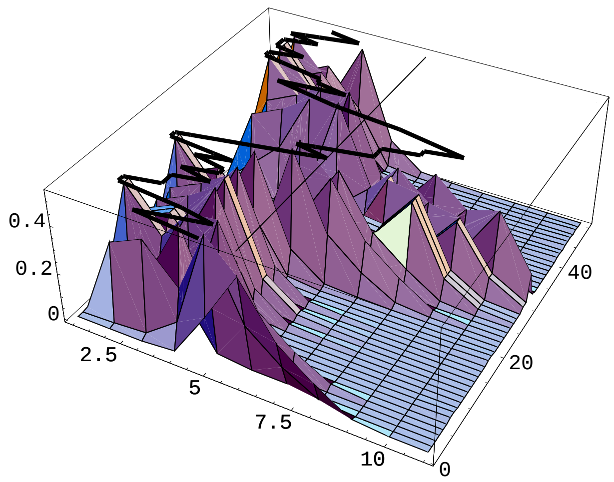
Time–Dependent Inverse Quantum Mechanics

Path of Repeatedly Measured Particle

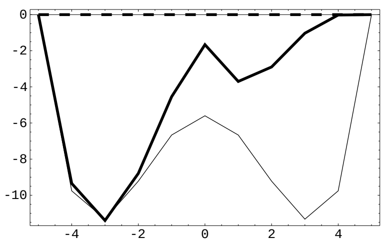
Original reconstructed

likelihood





Potential



Likelihood