The Tunnelling Salesman: Truncated Variational Approximations for Quantum Mechanical Global Optimization

Oliver B. Downs
Hagai Attias
Chris J.C. Burges
and Robert Rounthwaite

Abstract

In this work we present a novel method for global optimization, exploiting the mathematics of quantum mechanics, and in particular the tunnelling phenomenon. We consider a quantum mechanical system with a single particle in a potential specified by the cost function of our optimization problem. We assume the ground-state of the system is localised to the global minimum of the potential function, a condition which can be assured by choosing the particle mass sufficiently large. We then approximate this groundstate using a variational approach to optimise an expansion in terms of a truncated basis set of harmonic oscillator wavefunctions. We show how to encode integer factoring and travelling salesman problems in terms of finding the global minimum of a quartic polynomial cost function. We demonstrate our quantum algorithm on one- and two-dimensional toy problems, and apply it to factoring biprimes with 10 bits.

Introduction

The state of the art in global optimization for general problems is simulated annealing and its derivatives. The algorithm can be viewed in the physical context of a classical thermal system. This inspires the novel approach presented here, in which we draw upon intuition from non-classical physical systems which obey quantum mechanics (QM), and in particular exhibit the phenomenon of quantum tunnelling.

In QM, a potential function $\hat{V}(x)$ describes the energy terrain experienced by a particle, which may contain multiple minima separated by barriers. Combined with a term describing the kinetic energy of the particle, it defines the Hamiltonian operator:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + \hat{V} \qquad (1)$$

QM captures all available information on the state of the particle in a (complex) wave function $\psi(x)$. For instance, $|\psi(x)|^2$ is the probability distribution over the position x of the particle. To obtain $\psi(x)$ one solves the eigenvalue problem

$$\hat{H}\psi = E\psi \qquad (2)$$

known as the Schroedinger equation, where E is the energy. The solution is a set of wave functions termed eigenstates, and a matching set of energies. The eigenstate with the lowest energy is the ground state, ψ^0. By choosing the particle's mass sufficiently large, the probability density associated with the ground state wave function becomes localized to the global minimum of the potential.

Hence, a given optimization problem involving the minimization of a cost function can be restated as the problem of finding the ground state of the QM system with a potential specified by that cost function. In particular, computationally 'hard' integer problems, including factoring [1] and the travelling salesman problem, can be cast as QM problems involving a quartic polynomial potential, [3]. However, solving the Schroedinger equation exactly for these potentials is generally an analytically intractable problem.

In our approach we approximate the groundstate of such a system using a method based on perturbation theory.

2

A truncated variational approach for approximate ground-states

Assume that a local minimum of the potential (over an N-dimensional space) has been found. We make an anisotropic harmonic oscillator (HO) approximation to the potential function at the local minimum, by computing its Hessian. We change coordinates to the diagonal basis of the Hessian at the local minimum, thus factorizing the harmonic oscillator wavefunctions, which we use as a basis set for our groundstate approximation.

We then compute the matrix elements of the local harmonic oscillator wavefunctions $\{\phi^n\}$ under the full Hamiltonian of the system

$$M_{mn} = \langle \phi^m | \hat{H} | \phi^n \rangle \tag{3}$$

For the optimization problems treated here, which involve a quartic cost function, the full potential is simply a quartic perturbation to the HO approximation. We can therefore evaluate the matrix elements efficiently by representing the full Hamiltonian in terms of the N 1-dimensional "ladder" operators $\hat{A}_i^\dagger$, $\hat{A}_i$ of the HO approximation (see Appendix), since we can write polynomials in y_i in terms of these operators in the ith dimension, for example:

$$y_i^2 = \frac{1}{2\hbar} \left(\hat{A}_i + \hat{A}_i^\dagger \right)^2 \tag{4}$$

These matrix elements can therefore be computed analytically in arbitrarily high dimension N, with a computational cost $O(N^4)$; $O(N^3)$ for determining the diagonal basis of the Hessian at the local minimum, and an additional linear term in the dimensionality for computing the integrals as a product on the N variables.

Eigen-decomposition of this matrix yields an approximation for the energy levels in the full potential, the eigenvalues $\{E_{\mathrm{full},k}\}$, and the combinations of the HO basis functions making approximate corresponding eigenstate wavefunctions for the full potential $\psi_{\mathrm{full}}^{(k)}$, the eigenvectors $\{\vec{e}_k\}$. We *truncate* the basis set at K HO basis functions (the methodology for selecting these is described below)

$$\psi^{(k)}(\vec{x})_{\mathrm{full}} \approx \sum_{l=1}^{K} (\vec{e_k})_l \phi^l(\vec{x}) \tag{5}$$

We use this approximation for $\psi^{(0)}(\vec{x})_{\mathrm{full}}$ and compute the mean and covariance matrix of the resulting probability distribution over the location of the particle, $p^{(0)}(x) = |\psi^{(0)}(x)|^2$, which makes a localised single-peaked approximation for the location of the global minimum of the potential. A gradient method, initialized from the mean of the probability distribution, can then be used to exactly determine the location of the global minimum.
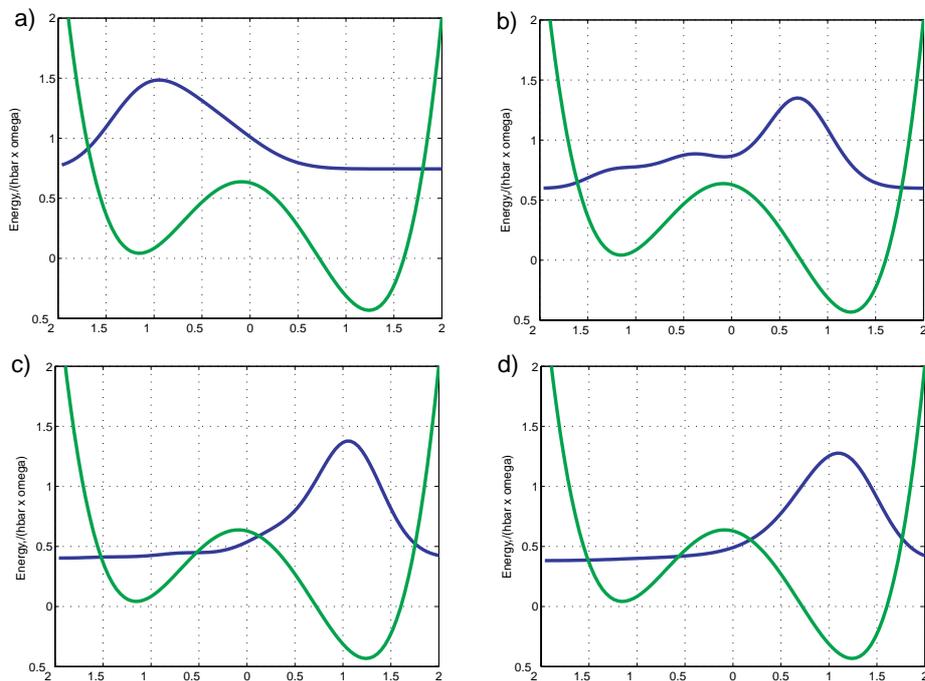
Figure 1: Truncated variational groundstate approximations using a) 3, b) 7, c) 10, d) 20 Harmonic oscillator eigenfunctions of the left-hand minimum of the potential, with $\epsilon = 0.5$, plotted at a vertical offset given by the energy of the state. (Potential is in grey, approximate groundstate is in black).

Toy Problems

Quartic potentials in 1 and 2 dimensions

We first consider approximations to the groundstate of the one-dimensional potential

$$\hat{V}(x) = (1 - x)^2 (1 + x)^2 - \epsilon x \qquad (6)$$

In Fig. 1 the approximate groundstate is shown when the harmonic oscillator basis of the left hand well is truncated at 3, 7, 10 and 20 basis functions. It can be seen that the approximate groundstate becomes localized to the global minimum similar to the true groundstate of the system.

This illustrates the quantum mechanical tunneling phenomenon. The energy of the resulting approximate state is found to be well below the energy of the interceding local maximum of the potential.

In two dimensions we consider the (separable) quartic potential function

$$\hat{V}(x, y) = x^4 + y^4 - 2x^2 - 2y^2 - x - \tfrac{1}{2}y + 2 \qquad (7)$$

which has four unequal minima, as shown in Fig. 2a).

4
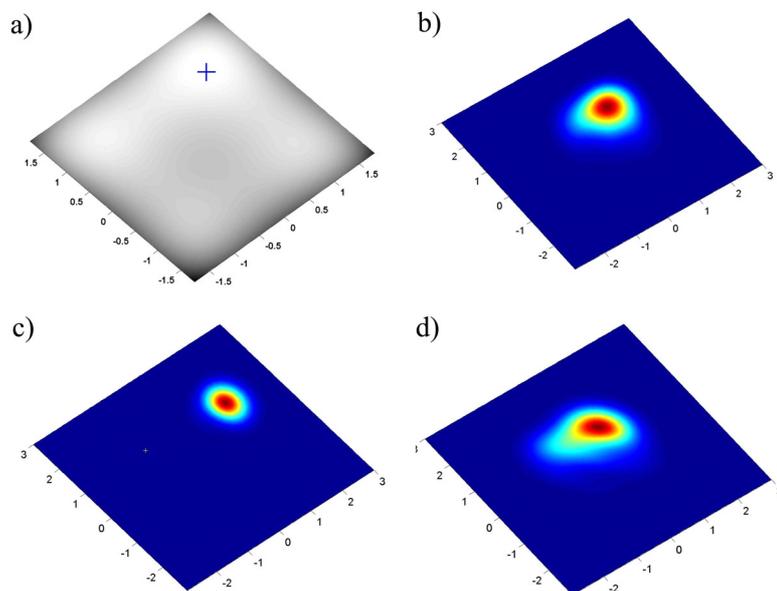
Figure 2: a) 4-minimum toy 2-dimensional quartic potential function, global min. indicated by '+'. Truncated variational groundstate density using an HO basis centred on the local minimum at: b) [-0.9301, 1.1072], c) [1.0575, -0.8376] and d) [1.0575 1.1072].

Using 50 (HO) basis functions centred on each of the three local minima, we obtain groundstate approximations for the system which are localised to the global minimum of the problem, Fig. 2b)-d).

Quartic Potential Functions for 'hard' problems

It is possible to express several computationally hard optimization problems in the context of finding the global minimum of a low-order polynomial potential function, explicitly here we present the factoring, and travelling salesman problems. Such "programming" by potential function is a long-established concept, [2].

Factoring Integers

We consider the problem of factoring the (large) integer h, assuming that h has as single pair of (prime) factors f and g. We represent the problem in a binary space in which $\{f_i\}, \{g_j\}$ are the sets of N bits representing f and g respectively $f = \sum_{i=1}^{N} 2^i f_i$ and $g = \sum_{j=1}^{N} 2^j g_j$.

We can then write the potential term which forces the factorisation as

$$V_f = \left(\sum_{i,j} 2^{i+j} f_i g_j - h \right)^2 \tag{8}$$

Since $\{f_i, g_j\}$ are jointly the variables of the problem, we concatenate them into a single set of binary variables $\{x_k\}$ with:

$$x_k = \begin{cases} f_k \text{ for } k = 1 \text{ to } N, \\ g_{k-N} \text{ for } k = N+1 \text{ to } 2N. \end{cases} \tag{9}$$

We add a binarising term

$$V_b = \sum_{i=1}^{2N} x_i^2 \left(1 - x_i \right)^2 \tag{10}$$

which enforces minima in the potential to be integers. We then need to add a term to enforce an ordering on the two prime factors to make for a single unique global minimum, by biasing the second factor to be larger than the first

$$V_o = \sum_{i=1}^{N} \left(2^i x_i - 2^i x_{N+i} \right) \tag{11}$$

The full potential function is then a quartic polynomial in the variables $\{x_i\}$ having a unique global minimum at the solution of the factorisation

$$\hat{V} = \left(\sum_{i=1}^{N} \sum_{j=1}^{N} 2^i 2^j x_i x_{N+j} - h \right)^2 + \sum_{i=1}^{2N} x_i^2 \left(1 - x_i \right)^2 + \sum_{i=1}^{N} \left(2^i x_i - 2^i x_{N+i} \right) \tag{12}$$

A less simplistic approach to encoding this problem is presented by Burges, [1], yielding a polynomial with all coefficients of $O(1)$. However it codes in $O(2N)$ variables.

Travelling Salesman

We consider a set of N cities, with pairwise distances between them described by the symmetric matrix D. We describe a journey by the N by N matrix X, if there is an edge of the journey between the mth and nth cities then $X_{mn} = 1$, and 0 otherwise. Stringing the rows of X end on end we obtain the N^2 by 1 vector x, then in the same space, D becomes the diagonal N^2 by N^2 matrix D'.

The total distance covered in a given journey x is then given by

$$C_{\text{distance}} = x^T D' x \tag{13}$$

We then have to add terms which enforcing the constraints on the journey - each city is visited only once, and that all cities are visited. This is the same as forcing only one entry of value 1 in each row and column of the matrix.

$$V_{\text{all cities once}} = \sum_m \left(\sum_n x_{mN+n} - 1 \right)^2 + \sum_n \left(\sum_m x_{mN+n} - 1 \right)^2 \tag{14}$$

We then need to enforce the binary nature of the variables, as in the factoring problem

$$C_b = \sum_k x_k{}^2 \left(1 - x_k \right)^2 \tag{15}$$

Thus we have programmed the N-city Travelling Salesman problem as that of globally minimising a quartic polynomial energy function in N^2 variables.

Implementation issues

An indication of how many functions are required from the ladder of HO basis functions can be obtained by considering the overlap of the nth HO basis function with a Gaussian of similar width located in one of the adjacent minima, as illustrated in Fig. 3a). Since the HO basis functions have a spatial span which increases with their energy, then for problems restricted to the binary hypercube the number of functions required is upper-bounded.

The most distant local minimum from the global will be at the diagonally opposite corner of the binary hypercube, assuming this case yields a discrete Poisson distribution over the overlap of the nth function

$$p(n) = e^{-\alpha^2} \frac{\alpha^{2n}}{n!} \tag{16}$$

with mean and variance, $\alpha^2 = \frac{N}{2} \frac{m\omega}{\hbar}$, where ω is the square-root of the curvature of the local minimum along the diagonal of the hypercube.

This still suggests a worst-case exponential algorithm, since the maximum number of basis functions required in each of the N dimensions is proportional to N. However empirically it is found that very few of the the full set of basis functions contribute strongly to the groundstate approximation, and so we need a criterion by which to
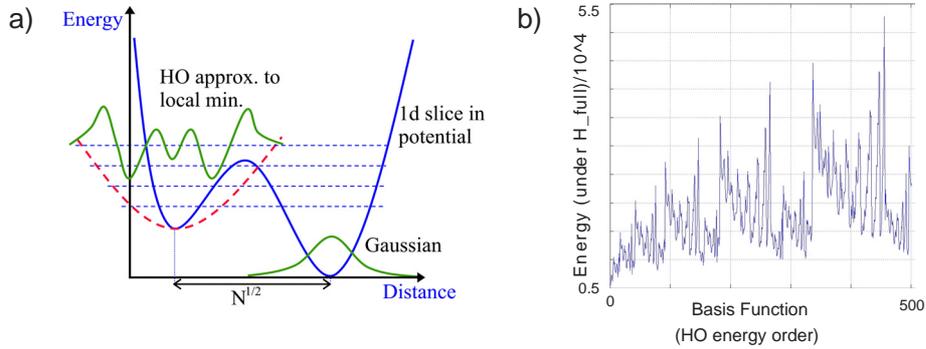
Figure 3: a) Prototype problem for determining the number of HO basis functions we expect to need in the approximation. We consider the overlap integrals of the spectrum of eigenfunctions $\{\phi_i\}$ of the HO approximation to the left hand local minimum with a Gaussian centred in the global minimum, g, along the diagonal line through the centre of the binary cube, with harmonic oscillator angular freq, ω. b) Energies of the first 500 HO basis functions under the full potential in 6 dimensions for factoring the number 35.

restrict the set of basis functions employed in the algorithm. We consider evaluating the diagonal elements of the matrix, M_{kl}. These elements reflect the average energy of the HO basis functions as states under the *full* Hamiltonian. The intuition is that the lower energy overlap states will contribute strongly to the groundstate approximation, and it is these which should be selected as the set of basis functions for the approximation, Fig. 3b). This is an area of ongoing work.

Discussion

In this work we have presented a deterministic approach to global optimization exploiting intuition and mathematics from quantum physics. The simplicity of the approach lies in the tractability of the harmonic oscillator basis set, which can be exploited for approximating the eigenstates of any polynomial potential. Here we have graphically demonstrated the concept of the work in 1 and 2 dimensional problems, and are able to report preliminary results on factoring the small biprimes 6, 35 and 899. Significant further work lies in better understanding the computational complexity of the algorithm, as embodied in determining the typical size of the required basis set. Further experimental work will focus on empirically exploring the scalability of the algorithm, and its reliability in identifying the global solution to different classes of optimization problem.

Acknowledgements

References

[1] Burges, CJC (2002). Factoring as Optimization. *Microsoft Research Technical Report*, MSR-TR-2002-83, (August 2002).

[2] Hopfield, JJ and Tank, DW (1985). 'Neural' computation of decisions in optimization problems. *Biological Cybernetics* 52. pp. 141–152.

[3] Bellare, M, Goldwasser S, Lund, C and Russell, A (1993). Efficient probabilistically checkable proofs and applications to approximation. *Proceedings 25th Annual Symposium on the Theory of Computing, ACM*.

Appendix

The 1-d HO "ladder" operators are defined as follows:

$$\hat{A} = \sqrt{\frac{m\omega}{2}}x + \frac{\hbar}{\sqrt{2m\omega}}\frac{\partial}{\partial x}, \ \hat{A}^\dagger = \sqrt{\frac{m\omega}{2}}x - \frac{\hbar}{\sqrt{2m\omega}}\frac{\partial}{\partial x} \tag{17}$$

and hence, using the standardised variables

$$y = \sqrt{\frac{m\omega}{\hbar}}x = \frac{1}{\sqrt{2\hbar}}\left(\hat{A} + \hat{A}^\dagger\right) \tag{18}$$

It is then simple to compute matrix-elements of the HO eigenfunctions $\{\phi_i^n\}$ for the ith-dimension

$$M_{mn} = \langle \phi_i^m | \hat{H} | \phi_i^n \rangle \tag{19}$$

since

$$\left(\hat{A_i}^\dagger\right)^m \phi_i^{n_i} = \left(\sqrt{\hbar}\right)^m \sqrt{\frac{(n_i+m)!}{n_i!}}\phi_i^{n_i+m}$$
$$\left(\hat{A_i}\right)^m \phi_i^{n_i} = \left(\sqrt{\hbar}\right)^m \sqrt{\frac{n_i!}{(n_i-m)!}}\phi_i^{n_i-m} \tag{20}$$

and where c is a constant,

$$\langle \phi_i^n | c\phi_i^m \rangle = c\delta_{nm} \tag{21}$$