A Quantum Field Evolution Strategy
An Adaptive Surrogate Approach

Jörg Bremer[1] and Sebastian Lehnhoff[2]
[1]*Department of Computing Science, University of Oldenburg, Uhlhornsweg, Oldenburg, Germany*
[2]*R&D Division Energy, OFFIS – Institute for Information Technology, Escherweg, Oldenburg, Germany*

Keywords: Evolution Strategies, Global Optimization, Surrogate Optimization, Quantum Potential.

Abstract: Evolution strategies have been successfully applied to optimization problems with rugged, multi-modal fitness landscapes, to non linear problems, and to derivative free optimization. Usually evolution is performed by exploiting the structure of the objective function. In this paper, we present an approach that harnesses the adapting quantum potential field determined by the spatial distribution of elitist solutions as guidance for the next generation. The potential field evolves to a smoother surface leveling local optima but keeping the global structure what in turn allows for a faster convergence of the solution set. We demonstrate the applicability and the competitiveness of our approach compared with particle swarm optimization and the well established evolution strategy CMA-ES.

1 INTRODUCTION

Evolution Strategies have shown excellent performance in global optimization especially when it comes to complex multi-modal, high dimensional, real valued problems (Kramer, 2010; Ulmer et al., 2003). A major drawback of population based algorithms is the large number of objective function evaluations. Real world problems often face computational efforts for fitness evaluations; e. g. in Smart Grid load planning scenarios, fitness evaluation involves simulating a large number of energy resources and their behaviour (Bremer and Sonnenschein, 2014).

We propose a surrogate approach that harnesses a continuously updating quantum potential field determined by elitist solutions. On the one hand side, the quantum field exploits global information by aggregating over scattered fitness information similar to scale space approaches (Horn and Gottlieb, 2001; Leung et al., 2000), on the other side by continuously adapting to elitist solutions, the quantum field surface quickly flattens to a smooth surrogate for guiding further sampling directions. To achieve this, the surrogate is a result of Schrödinger's equation of which a probability function is derived that determines the potential function after a clustering approach (cf. (Horn and Gottlieb, 2001)). We associate minima of the potential field, created in denser regions of good solutions' positions with areas of interest for further investigation. Thus, offspring solutions are generated with a trend in descending the potential field. By harnessing the quantum potential as a surrogate, we achieve a faster convergence with less objective function calls compared with using the objective function alone. In lieu thereof the potential field has to be evaluated at selected point. Although a fine grained computation of the potential field would be a computationally hard task in higher dimensions (Horn and Gottlieb, 2001), we achieve a better overall performance because we need to calculate the field only at isolated data points.

The paper starts with a review of using quantum mechanics in computational intelligence and in particular in evolutionary algorithms; we briefly recap the quantum potential approach for clustering and present our adaption for integration into evolution strategies. We conclude with an evaluation of the approach with the help of several well-known benchmark test functions and demonstrate the competitiveness to two competitive algorithms: particle swarm optimization (PSO) and co-variance matrix adaption evolution strategy (CMA-ES).

2 RELATED WORK

Several evolutionary algorithms have been introduced to solve nonlinear complex optimization problems with multi-modal, rugged fitness landscapes. Each

Bremer, J. and Lehnhoff, S.
A Quantum Field Evolution Strategy - An Adaptive Surrogate Approach.
DOI: 10.5220/0006037000210029
In *Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016) - Volume 1: ECTA*, pages 21-29
ISBN: 978-989-758-201-1

of these method has its own characteristics, strengths and weaknesses. A common characteristics in all EAs is the generation of an offspring solution set in order to explore the characteristics of the objective function in the neighbourhood existing solutions. When the solution space is hard to explore or objective evaluations are costly, computational effort is a common drawback for all population-based schemes. Many efforts have been already spent to accelerate convergence of these methods. Example techniques are: improved population initialization (Rahnamayan et al., 2007), adaptive populations sizes (Ahrari and Shariat-Panahi, 2013) or exploiting sub-populations (Rigling and Moore, 1999).

Sometimes a surrogate model is used in case of computational expensive objective functions (Loshchilov et al., 2012) to substitute a share of objective function evaluations with cheap surrogate model evaluations. The surrogate model represents a learned model of the original objective function. Recent approaches use Radial Basis Functions, Polynomial Regression, Support Vector Regression, Artificial Neural Network or Kriging (Gano et al., 2006); each approach with individual advantages and drawbacks.

At the same time, quantum mechanics has inspired several fields of computational intelligence such as data mining, pattern recognition or optimization. In (Horn and Gottlieb, 2001) a quantum mechanics based method for clustering has been introduced. Quantum clustering extends the ideas of Scale Space algorithms and Support Vector Clustering (Ben-Hur et al., 2001; Bremer et al., 2010) by representing an operator in Hilbert space by a scaled Schrödinger equation that yields a probability function as result. The inherent potential function of the equation that can be analytically derived from the probability function is used to identify barycenters of data cluster by associating minima with centers. In (Weinstein and Horn, 2009a) this approach has been extended to a dynamic approach that uses the fully fledged time dependant variant of the Schrödinger equation to allow for a interactive visual data mining especially for large data sets (Weinstein et al., 2013).

We adapted and extended the quantum field part of the clustering approach to optimization and use the potential function to associate already found solutions from the objective domain with feature vectors in Hilbert space; but with keeping an emphasis of the total sum (cf. (Horn and Gottlieb, 2001)) and thus with keeping in mind all improvements of the ongoing optimum search.

(Rapp and Bremer, 2012) used a quantum potential approach derived from quantum clustering to detect abnormal events in multidimensional data streams. (Yu et al., 2010) used quantum clustering for weighing linear programming support vector regression. In this work, we derive a sampling method for a $(\mu+\lambda)$-ES from the quantum potential approach originally used for clustering by (Horn and Gottlieb, 2002).

A quantum mechanical extension to particle swarm optimization has been presented e.g. in (Sun et al., 2004; Feng and Xu, 2004). Here particles move according to quantum mechanical behavior in contrast to the classical mechanics ruled movement of particles in standard PSO. Usually a harmonic oscillator is used. In (Loo and Mastorakis, 2007) both methods quantum clustering and quantum PSO have been combined by deriving good particle starting positions from the clustering method first. For the simulated Annealing (SA) approach also a quantum extension has been developed (Suzuki and Nishimori, 2007). Whereas in classical SA local minima are escaped by leaping over the barrier with a thermal jump, quantum SA introduces the quantum mechanical tunneling effect for such escapes.

We integrated the quantum concept into evolution strategies; but by using a different approach: we harness the information in the quantum field about the so far gained success as a surrogate for generating the offspring generation. By using the potential field, information from all samples at the same time is condensed into a directed generation of the next generation.

3 THE SCHRÖDINGER POTENTIAL

We start with a brief recap of the Schrödinger potential and describe the concept following (Horn and Gottlieb, 2002; Horn and Gottlieb, 2001; Weinstein and Horn, 2009b). Let

$$H\psi \equiv (-\frac{\sigma_{\text{pot}}^2}{2}\nabla^2 + V(\boldsymbol{x}))\psi(\boldsymbol{x}) = E\psi(\boldsymbol{x}) \quad (1)$$

be the Schrödinger equation rescaled to a single free parameter σ_{pot} and eigenstate $\psi(\boldsymbol{x})$. H denotes the Hamiltonian operator corresponding to the total energy E of the system. ψ is the wave function of the quantum system and ∇^2 denotes the Laplacian differential operator. V corresponds to the potential energy in the system. In case of a single point at $\boldsymbol{x}_0$ Eq. (1) results in

$$V = \frac{1}{2}\sigma_{\text{pot}}^2(\boldsymbol{x}-\boldsymbol{x}_0)^2 \quad (2)$$

$$E = \frac{d}{2} \quad (3)$$

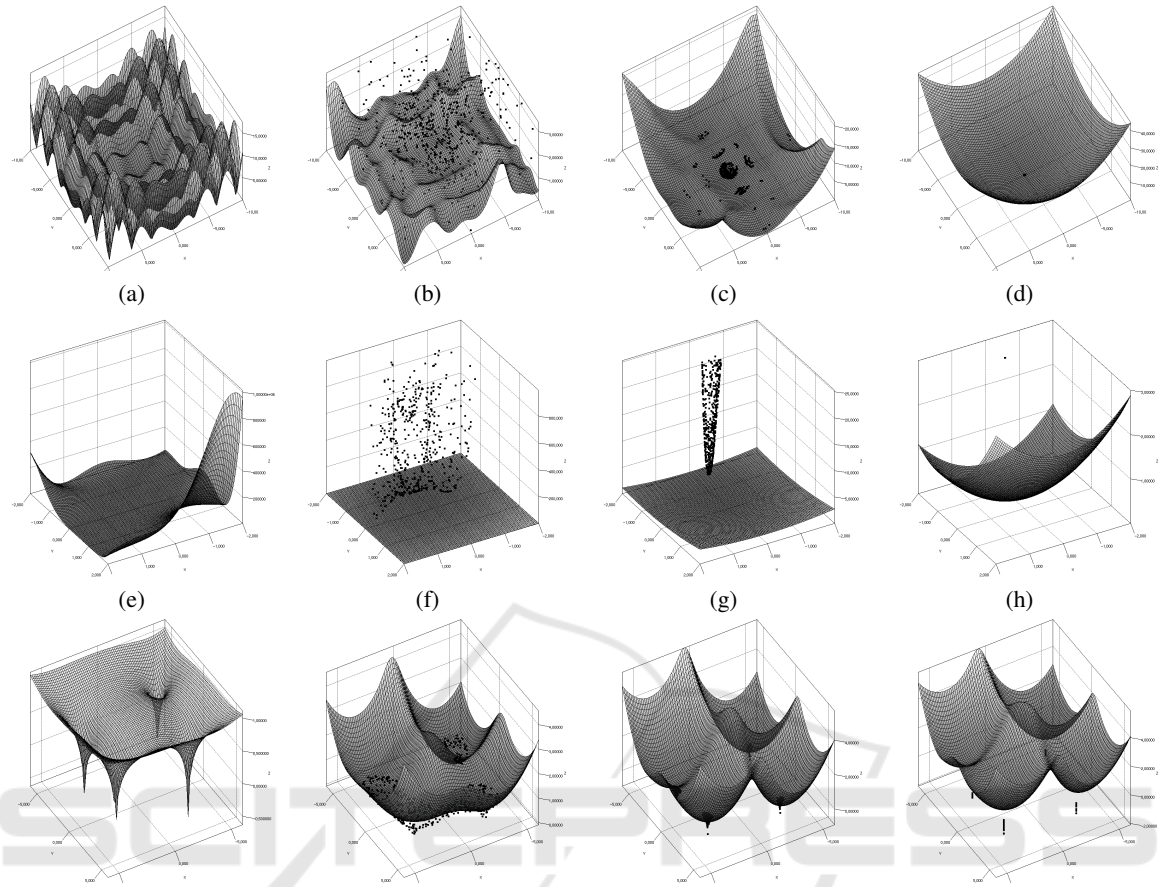

Figure 1: Function (left column) and exemplary evolution (after 1, 3 and 8 iterations) of the quantum potential that guides the search for minima for the test functions: Alpine, Goldstein-Price and Himmelblau (from top to bottom).

with d denoting the dimension of the field. In this case ψ is the ground state of a harmonic oscillator. Given an arbitrary set of points, the potential at a point $\boldsymbol{x}$ can be expressed by

$$\begin{aligned} V(\boldsymbol{x}) &= E + \frac{\frac{\sigma_{\text{pot}}^2}{2}\nabla^2\psi}{\psi} \\ &= E - \frac{d}{2} + \frac{1}{2\sigma_{\text{pot}}^2\psi}\sum_i (\boldsymbol{x}-\boldsymbol{x}_i)^2 e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_i)^2}{2\sigma_{\text{pot}}^2}}. \end{aligned} \tag{4}$$

In Eq. (4) the Gaussian wave function

$$\psi(\boldsymbol{x}) = \sum_i (\boldsymbol{x}-\boldsymbol{x}_i)^2 e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}_i)^2}{2\sigma_{\text{pot}}^2}} \tag{5}$$

is associated to each point and summed up. Please note that the bandwidth parameter (usually named σ, cf. (Horn and Gottlieb, 2002)) has been denoted σ_{pot} to discriminate the bandwidth of the wave function and the variance σ in the mutation used in the evolutions strategy later. In quantum mechanics, usually the potential $V(\boldsymbol{x})$ is given and solutions or eigenfunctions $\psi(\boldsymbol{x})$ are sought. In our approach we are already given $\psi(\boldsymbol{x})$ determined by a set of data points. The set of data points is given by elitist solutions. We then look for the potential $V(x)$ whose solution is $\psi(\boldsymbol{x})$.

The wave function part corresponds with the Parzen window estimator approach for data clustering (Parzen, 1962) or with scale-space clustering (Leung et al., 2000) that interprets this wave function as the density function that could have generated the underlying data set. The maxima of this density function correspond therefore with data centers.

In quantum clustering and by requiring ψ to be the ground state of the Hamiltonian H the potential field V establishes a surface that shows more pronounced minima (Weinstein and Horn, 2009b). V is unique up to a constant factor. By setting the minimum of V to zero it follows that

$$E = -min\frac{\frac{\sigma_{\text{pot}}^2}{2}\nabla^2\psi}{\psi}. \tag{6}$$

With this convention V is determined uniquely with $0 \leqslant E \leqslant \frac{d}{2}$. In this case, E is the lowest eigenvalue of the operator H and thus describes the ground state.

V is expected to exhibit one or more minima within some region around the data set and grow quadratically on the outside (Horn and Gottlieb, 2002). In quantum clustering, these minima are associated with cluster centers. We will interpret them as balance points or nuclei where the minimum of an associated function f lies if the set of data points that defines V is a selection of good points (in the sense of a good fitness according to f).

4 THE ALGORITHM

We start with a general description of the idea. In our approach we generate the quantum potential field of an elitist selection of samples. Out of a sample of λ solutions the best μ are selected according to the objective function. These μ solutions then define a quantum potential field that exhibits troughs at the barycenters of good solution taking into account all good solutions at the same time. In the next step this potential field is used to guide the sampling of the next generation of λ offspring solutions from which the next generation is selected that defines the new field. In this way, the potential field continuously adapts in each iteration to the so far found best solutions.

The advantage of using the potential field results from its good performance in identifying the barycenters of data points. Horn and Gottlieb (Horn and Gottlieb, 2001) demonstrated the superior performance compared with density based approaches like Scale Space or Parzen Window approaches (Roberts, 1997; Parzen, 1962). Transfered to optimization this means the quantum potential allows for a better identification of local optima. As they can be explored faster they can be neglected earlier which in turn leads to a faster convergence of the potential field towards the global optimum (cf. Figure 1).

Figure 1 gives an impression of the adaption process that transforms the quantum field into an easier searchable function. Each row shows the situation after 1, 3 and 8 iterations for different 2-dimensional objective functions. The left column displays the original objective function; from left to right the evolving potential field is displayed together with the respective offspring solutions that represent the so far best. The minimum of the potential field evolves towards the minimum of the objective function (or towards more than one optimum if applicable).

Figure 2 shows the approach formally. Starting from an initially generated sample $\mathcal{X}$ equally distributed across the whole search domain defined by a box constraint in each dimension. Next, the offspring is generated by sampling λ points normally distributed around the μ solutions from the old generation with an each time randomly chosen parent solution as expectation and with variance σ^2 that decreases with each generation. We use a rejection sampling approach with the metropolis criterion (Metropolis et al., 1953) for acceptance applied to the difference in the potential field between a new candidate solution and the old solution. The new sample is accepted with probability

$$p_a = \min(1, e^{\Delta V}). \tag{7}$$

$\Delta V = V(x_{old}) - V(x_{new})$ denotes the level difference in quantum field. A descent within the potential field is always accepted. A (temporary) degradation in quantum potential level is accepted with a probability P_a (eq. 7) determined by the level of degradation.

As long as there exists at least one pair $\{\boldsymbol{x}_1, \boldsymbol{x}_2\} \subset \mathcal{S}$ with $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, the potential field has a minimum at $\boldsymbol{x}'$ with $\boldsymbol{x}' \neq \boldsymbol{x}_1 \wedge \boldsymbol{x}' \neq \boldsymbol{x}_1$. Thus, the sampling will find new candidates. The sample variance σ^2 is decreased in each iteration by a rate ω. Finally, for the next iteration, the solution set $\mathcal{X}$ is updated by selecting the μ best from offspring $\mathcal{S}$.

The process is repeated until any stopping criterion is met; apart from having come near enough the minimum we regularly used an upper bound for the maximum number of iterations (or rather: number of fitness evaluations respectively).

$\mathcal{X} \coloneqq \{\boldsymbol{x}_i \sim \mathcal{U}(x_{lo}, x_{up})^d\},\ 1 \leqslant i \leqslant n$
repeat
  $\mathcal{S} \leftarrow \varnothing$
  repeat
    $\boldsymbol{x}_z \coloneqq \boldsymbol{x}_i \in \mathcal{X},\ i \sim \mathcal{U}(1, |\mathcal{X}|)$
    $\boldsymbol{s} \sim \mathcal{N}(\boldsymbol{x}_z, \sigma^2)$
    if $p \leqslant e^{V(\boldsymbol{x}_z) - V(\boldsymbol{s})},\ p \sim \mathcal{U}(0,1)$ **then**
      $\boldsymbol{S} \leftarrow \boldsymbol{S} \cup \boldsymbol{s}$
    end if
  until $|\mathcal{S}| == \lambda$
  $V \leftarrow V(\mathcal{S}, \sigma_{\text{pot}})$
  $\mathcal{X} \leftarrow \text{select}(\mathcal{S}, f, \mu)$
  $\sigma \leftarrow \sigma \cdot \omega$
until $\|f(x_{best}) - f(x^*)\| \leqslant \varepsilon$

Figure 2: Basic scheme for the Quantum Sampling ES Algorithm.

5 RESULTS

We evaluated our evolution strategy with a set of well known test functions developed for benchmark-

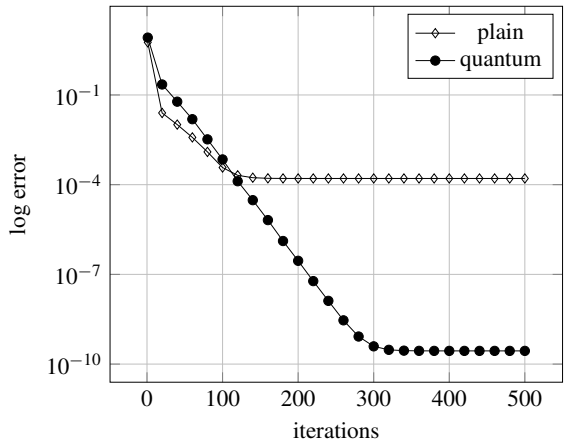

Figure 3: Comparing the convergence of using the quantum field as surrogate with the plain approach working on the objective function directly. For testing, the 2-dimensional Alpine function has been used.

ing optimization heuristics. We used the following functions: Alpine, Goldstein-Price, Himmelblau, Bohachevsky 1, generalized Rosenbrock, Griewank, Sphere, Booth, Chichinadze and Zakharov (Ulmer et al., 2003; Ahrari and Shariat-Panahi, 2013; Himmelblau, 1972; Yao et al., 1999); see also appendix A. These functions represent a mix of multi-modal, 2-dimensional and multi-dimensional functions, partly with a huge number of local minima and steep as well as shoal surroundings of the global optimum and broad variations in characteristics and domain sizes.

Figure 1 shows some of the used functions (left column) together with the respective evolution of the quantum potential field that guides the search towards the minimum at $(0,0)$ for the Alpine function and $(0,-1)$ for Goldstein-Price; the Himmelblau function has four global minima which are all found. The figure also shows the evolution of the solution population. In the next step, we tested the performance of our approach against competitive approaches.

First, we tested the effect of using the quantum field as adaptive surrogate compared with the same update strategy working directly on the fitness landscape of the objective function. Figure 3 shows the convergence of the error on the Alpine test function for both cases. Although the approach with surrogate converges slightly slower in the beginning it clearly outperforms the plain approach without quantum surrogate. Figure 4 shows the same effect for the 20-dimensional case. Both results show the convergence of the mean error for 100 runs each.

In a next step, we compared our approach with two well-known heuristics: particle swarm optimization (PSO) from (Kennedy and Eberhart, 1995) and the covariance matrix adaption

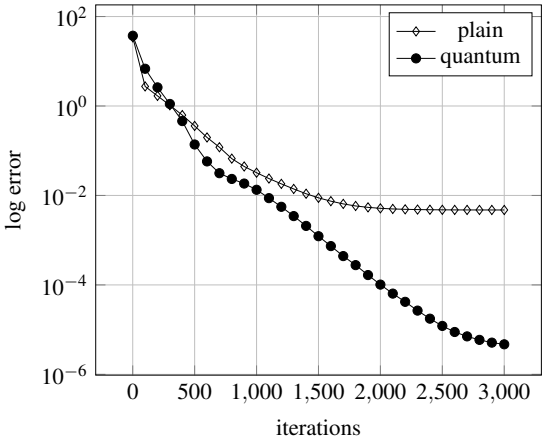

Figure 4: Convergence of using the quantum field as surrogate compared to the plain approach. Depicted are the means of 100 runs on the 20-dimensional Alpine function.

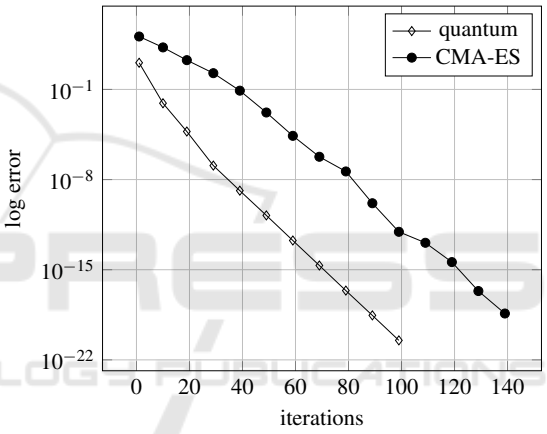

Figure 5: Comparing the convergence (means of 100 runs) of CMA-ES and the quantum approach on the 2-dimensional Booth function.

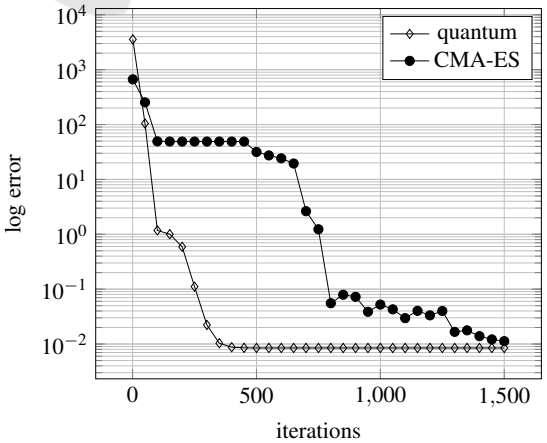

Figure 6: Comparing the convergence of CMA-ES and the quantum approach on the 20-dimensional Griewank function on the domain $[-2048, 2048]^{20}$. Both algorithms have been stopped after 1500 iterations.

evolution strategy (CMA-ES) by (Hansen, 2011). Both strategies are well-known, established, and have been applied to wide range of optimization problems. We used readily available and evaluated implementations from Jswarm-PSO (http://jswarm-pso.sourceforge.net) and commons math (http://commons.apache.org/proper/commons-math).

All algorithms have an individual, strategy specific set of parameters that usually can be tweaked to some degree for a problem specific adaption. Nevertheless, default values that are applicable for a wide range of functions are usually available. For our experiments, we used the following default settings. For the CMA-ES, the (external) strategy parameters are $\lambda, \mu, w_{i=1\ldots\mu}$, controlling selection and recombination; c_σ and d_σ for step size control and c_c and μ_{cov} controlling the covariance matrix adaption. We have chosen to set these values after (Hansen, 2011).

$$\lambda = 4 + \lfloor 3 \ln n \rfloor, \quad \mu = \left\lfloor \frac{\lambda}{2} \right\rfloor, \tag{8}$$

$$\boldsymbol{w}_i = \frac{\ln(\frac{\lambda}{2} + 0.5) - \ln i}{\sum_{\mu}^{j=1} \frac{\lambda}{2} + 0.5) - \ln i}, \; i = 1, \ldots, \mu \tag{9}$$

$$C_c = \frac{4}{n+4}, \; \mu_{\text{cov}} = \mu_{eff}, \tag{10}$$

$$C_{\text{cov}} = \frac{1}{\mu_{\text{cov}}} \frac{2}{(n+\sqrt{2})^2} + \left(1 - \frac{1}{\mu_{\text{cov}}}\right) \min\left(1, \frac{2\mu_{\text{cov}} - 1}{(n+2)^2 + \mu_{\text{cov}}}\right), \tag{11}$$

These settings are specific to the dimension N of the objective function. An in-depth discussion of these parameters is also given in (Hansen and Ostermeier, 2001).

For the PSO, we used values of 0.9 for the weights and 1 for the inertia parameter as default setting (Shi and Eberhart, 1998).

For the quantum field strategy, we empirically found the following values as a useful setting for a range of objective functions. The initial mutation variance has been set to $\sigma = d/10$ for an initial diameter d of the search space (domain of the objective function). The shrinking rate of the variance has been set to $\omega = 0.98$ and the bandwidth in the potential equation (4) has been set to $\sigma_{\text{pot}} = 0.4$. For the population size we chose $\mu = 10$ and $\lambda = 50$ if not otherwise stated.

First, we compared the convergence of the quantum strategy with the CMA-ES. Figure 5 shows a first result for the 2-dimensinal Booth function (with a minimum of zero). The quantum approach has been stopped at errors below 1×10^{-21} to avoid numerical instabilities. The used CMA-ES implementation has a similar condition integrated into its code. Comparing iterations, the quantum approach converges faster than CMA-ES. Figure 5 shows the result for the 20-dimensional Griewank function with a search domain of $[-2049, 2048]^{20}$. Here, the quantum approach achieves about the same result as the CMA-ES within less iterations. Comparing iterations does not yet shed light on performance.

As the performance is determined by the number of operations that have to be conducted in each iteration, the following experiments consider the number of function evaluation calls rather than iterations. Table 1 shows the results for a bunch of 2-dimensional test functions. For each test function and each algorithm the achieved mean (averaged over 100 runs) solution quality and the needed number of function evaluations is displayed. The solution quality is expressed as the error in terms of remaining difference to the known global optimum. As stopping criterion this time each algorithm has been equipped with two conditions: error below 5×10^{-17} and a given budget of at most 5×10^7 evaluations. For the quantum approach two counts of evaluation functions are given, because due to the nature of surrogate approaches a share of function evaluations is substituted by surrogate evaluations. Thus, total evaluations refers to the sum of function and surrogate evaluations.

The results show that CMA-ES is in general unbeatable in terms of function evaluations whereas the quantum approach in half of the cases gains the more accurate result. The PSO succeeds for the Griewank and the Chichinadze function. For the 20-dimensional cases in Table 2 the quantum approach gains the most accurate result in most of the cases. The CMA-ES winning margin of a low number of evaluations decreases compared with the quantum approach, but is still prominent. Nevertheless, the number of necessary function evaluations for the quantum approach can still be reduced when using a lower population size. But, such tuning is subject to the problem at hand. On the other hand, notwithstanding the low number of objective evaluations, the CMA-ES needs higher processing time for high-dimensional problems due to the fact that CMA-ES needs – among others – to conduct eigenvalue decompositions of its covariance matrix ($O(n^3)$) with number of dimensions n (Knight and Lunacek, 2007). Table 3 gives an expression for necessary computation (CPU-) times (Java 8, 2.7 GHz Quadcore) for the 100-dimensional Sphere function for CMA-ES and a quantum approach with reduced populations size ($\mu = 4$, $\lambda = 12$).

All in all, the quantum approach is competitive to the established algorithms and in some cases even superior.

Table 1: Results for comparing CMA-ES, PSO and the quantum approach with a set of 2-dimensional test functions. The error denotes the remaining difference to the known optimum; obj. evaluations and total evaluations refer to the number of conducted function evaluations and the sum of function and quantum surrogate evaluations respectively. The latter is only applicable to the quantum approach.

problem	algorithm	error	obj. evaluations	total evaluations
Alpine	CMA-ES	$3.954 \times 10^{-12} \pm 4.797 \times 10^{-12}$	746.38 ± 90.81	n/a
	PSO	$5.543 \times 10^{-9} \pm 3.971 \times 10^{-8}$	500000.00 ± 0.00	n/a
	quantum	$8.356 \times 10^{-16} \pm 4.135 \times 10^{-16}$	99363.00 ± 66100.02	198996.53 ± 132200.50
Griewank	CMA-ES	$2.821 \times 10^{2} \pm 2.254 \times 10^{2}$	174.22 ± 137.63	n/a
	PSO	$4.192 \times 10^{-4} \pm 1.683 \times 10^{-3}$	500000.00 ± 0.00	n/a
	quantum	$6.577 \times 10^{-3} \pm 5.175 \times 10^{-3}$	201361.50 ± 47207.02	472348.75 ± 94420.25
GoldsteinPrice	CMA-ES	$0.459 \times 10^{1} \pm 1.194 \times 10^{2}$	613.96 ± 203.03	n/a
	PSO	$1.698 \times 10^{0} \pm 1.196 \times 10^{1}$	500000.00 ± 0.00	n/a
	quantum	$1.130 \times 10^{0} \pm 1.255 \times 10^{1}$	250000.00 ± 0.00	500012.75 ± 3.48
Bohachevsky1	CMA-ES	$3.301 \times 10^{-2} \pm 1.094 \times 10^{-1}$	672.70 ± 59.91	n/a
	PSO	$1.626 \times 10^{-10} \pm 1.568 \times 10^{-9}$	500000.00 ± 0.00	n/a
	quantum	$4.707 \times 10^{-16} \pm 3.246 \times 10^{-16}$	39641.50 ± 1672.63	87548.03 ± 3377.00
Booth	CMA-ES	$4.826 \times 10^{-17} \pm 1.124 \times 10^{-16}$	605.68 ± 49.55	n/a
	PSO	$5.985 \times 10^{-14} \pm 3.743 \times 10^{-13}$	500000.00 ± 0.00	n/a
	quantum	$5.695 \times 10^{-16} \pm 2.957 \times 10^{-16}$	33696.00 ± 1357.07	68283.18 ± 2711.51
Chichinadze	CMA-ES	$0.953 \times 10^{1} \pm 4.125 \times 10^{1}$	698.44 ± 200.31	n/a
	PSO	$0.226 \times 10^{0} \pm 2.247 \times 10^{-1}$	500000.00 ± 0.00	n/a
	quantum	$1.327 \times 10^{1} \pm 8.376 \times 10^{0}$	50.00 ± 0.00	180.68 ± 12.63

Table 2: Results for comparing CMA-ES, PSO and the quantum approach with a set of 20-dimensional test functions with the same setting as in Table 1.

problem	algorithm	error	obj. evaluations	total evaluations
Rosenbrock	CMA-ES	$1.594 \times 10^{7} \pm 6.524 \times 10^{7}$	10678.60 ± 6514.47	n/a
	PSO	$1.629 \times 10^{10} \pm 1.759 \times 10^{10}$	50000000 ± 0.00	n/a
	quantum	$3.884 \times 10^{7} \pm 1.470 \times 10^{8}$	23291850.00 ± 3166.67	50001008.27 ± 557.53
Griewank	CMA-ES	$5.878 \times 10^{1} \pm 1.082 \times 10^{2}$	8734.48 ± 7556.50	n/a
	PSO	$1.429 \times 10^{2} \pm 3.539 \times 10^{2}$	50000000 ± 0.00	n/a
	quantum	$2.267 \times 10^{-3} \pm 3.973 \times 10^{-3}$	6929090.0 ± 974963.9	17269617.9 ± 1949920.8
Zakharov	CMA-ES	$9.184 \times 10^{-16} \pm 1.068 \times 10^{-15}$	8902.24 ± 845.05	n/a
	PSO	$7.711 \times 10^{1} \pm 5.961 \times 10^{1}$	50000000 ± 0.00	n/a
	quantum	$8.978 \times 10^{-17} \pm 9.645 \times 10^{-18}$	1021370.00 ± 4802.47	2962737.19 ± 10126.34
Spherical	CMA-ES	$1.200 \times 10^{-15} \pm 1.258 \times 10^{-15}$	8678.80 ± 912.13	n/a
	PSO	$2.637 \times 10^{0} \pm 7.517 \times 10^{0}$	50000000 ± 0.00	n/a
	quantum	$8.943 \times 10^{-17} \pm 1.036 \times 10^{-17}$	973750.00 ± 4243.53	2674716.70 ± 8844.04
Alpine	CMA-ES	$9.490 \times 10^{-12} \pm 3.331 \times 10^{-11}$	15196.48 ± 570.76	n/a
	PSO	$4.021 \times 10^{0} \pm 2.442 \times 10^{0}$	50000000 ± 0.0	n/a
	quantum	$9.272 \times 10^{-17} \pm 5.857 \times 10^{-18}$	1867830.0 ± 3795.41	4771580.5 ± 8390.4

Table 3: Comparing computational performace of CMA-ES and quantum approach with the 100-dimensional Sphere function.

algorithm	error	total evaluations	CPU time / nsec.
CMA-ES	$1.593 \times 10^{-16} \pm 1.379 \times 10^{-16}$	129204.4 ± 18336.1	$5.189 \times 10^{10} \pm 6.918 \times 10^{9}$
quantum	$9.861 \times 10^{-17} \pm 1.503 \times 10^{-18}$	161172.8 ± 626.8	$4.029 \times 10^{9} \pm 1.040 \times 10^{8}$

6 CONCLUSION

We introduced a novel evolution strategy for global optimization that uses the quantum potential field defined by elitist solutions for generating the offspring solution set.

By using the quantum potential, information about the fitness landscape of scattered points is condensed into a surrogate for guiding further sampling instead of looking at different single solutions; one at a time. In this way, the quantum surrogate tries to fit the search distribution to the shape of the objective function like CMA-ES (Hansen, 2006). The quantum surrogate adapts continuously as the optimization process zooms into areas of interest.

Compared with a population based solver and CMA-ES as established evolution strategy, we achieved a competitive and sometimes faster convergence with less objective function calls. We tested our method on ill-conditioned problems as well as on simple problems finding it performing equally good on both.

REFERENCES

Ahrari, A. and Shariat-Panahi, M. (2013). An improved evolution strategy with adaptive population size. *Optimization*, 64(12):1–20.

Ben-Hur, A., Siegelmann, H. T., Horn, D., and Vapnik, V. (2001). Support vector clustering. *Journal of Machine Learning Research*, 2:125–137.

Bremer, J., Rapp, B., and Sonnenschein, M. (2010). Support vector based encoding of distributed energy resources' feasible load spaces. In *IEEE PES Conference on Innovative Smart Grid Technologies Europe*, Chalmers Lindholmen, Gothenburg, Sweden.

Bremer, J. and Sonnenschein, M. (2014). Parallel tempering for constrained many criteria optimization in dynamic virtual power plants. In *Computational Intelligence Applications in Smart Grid (CIASG), 2014 IEEE Symposium on*, pages 1–8.

Feng, B. and Xu, W. (2004). Quantum oscillator model of particle swarm system. In *ICARCV*, pages 1454–1459. IEEE.

Gano, S. E., Kim, H., and Brown II, D. E. (2006). Comparison of three surrogate modeling techniques: Datascape, kriging, and second order regression. In *Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, AIAA-2006-7048*, Portsmouth, Virginia.

Hansen, N. (2006). The CMA evolution strategy: a comparing review. In Lozano, J., Larranaga, P., Inza, I., and Bengoetxea, E., editors, *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, pages 75–102. Springer.

Hansen, N. (2011). The CMA Evolution Strategy: A Tutorial. Technical report.

Hansen, N. and Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.*, 9(2):159–195.

Himmelblau, D. M. (1972). *Applied nonlinear programming [by] David M. Himmelblau*. McGraw-Hill New York.

Horn, D. and Gottlieb, A. (2001). The Method of Quantum Clustering. In *Neural Information Processing Systems*, pages 769–776.

Horn, D. and Gottlieb, A. (2002). Algorithm for data clustering in pattern recognition problems based on quantum mechanics. *Phys Rev Lett*, 88(1).

Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948 vol.4. IEEE.

Knight, J. N. and Lunacek, M. (2007). Reducing the space-time complexity of the CMA-ES. In *Genetic and Evolutionary Computation Conference*, pages 658–665.

Kramer, O. (2010). A review of constraint-handling techniques for evolution strategies. *Appl. Comp. Intell. Soft Comput.*, 2010:1–19.

Leung, Y., Zhang, J.-S., and Xu, Z.-B. (2000). Clustering by scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1396–1410.

Loo, C. K. and Mastorakis, N. E. (2007). Quantum potential swarm optimization of pd controller for cargo ship steering. In *Proceedings of the 11th WSEAS International Conference on APPLIED MATHEMATICS*, Dallas, USA.

Loshchilov, I., Schoenauer, M., and Sebag, M. (2012). Self-adaptive surrogate-assisted covariance matrix adaptation evolution strategy. *CoRR*, abs/1204.2356.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

Mishra, S. (2006). Some new test functions for global optimization and performance of repulsive particle swarm method. Technical report.

Parzen, E. (1962). On estimation of a probability den-

sity function and mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076.

Rahnamayan, S., Tizhoosh, H. R., and Salama, M. M. (2007). A novel population initialization method for accelerating evolutionary algorithms. *Computers & Mathematics with Applications*, 53(10):1605 – 1614.

Rapp, B. and Bremer, J. (2012). Design of an event engine for next generation cemis: A use case. In Hans-Knud Arndt, Gerlinde Knetsch, W. P. E., editor, *EnviroInfo 2012 – 26th International Conference on Informatics for Environmental Protection*, pages 753–760. Shaker Verlag. ISBN 978-3-8440-1248-4.

Rigling, B. D. and Moore, F. W. (1999). Exploitation of sub-populations in evolution strategies for improved numerical optimization. *Ann Arbor*, 1001:48105.

Roberts, S. (1997). Parametric and non-parametric unsupervised cluster analysis. *Pattern Recognition*, 30(2):261–272.

Shi, Y. and Eberhart, R. (1998). A modified particle swarm optimizer. In *International Conference on Evolutionary Computation*.

Sun, J., Feng, B., and Xu, W. (2004). Particle swarm optimization with particles having quantum behavior. In *Evolutionary Computation, 2004. CEC2004. Congress on*, volume 1, pages 325–331 Vol.1.

Suzuki, S. and Nishimori, H. (2007). Quantum annealing by transverse ferromagnetic interaction. In Pietronero, L., Loreto, V., and Zapperi, S., editors, *Abstract Book of the XXIII IUPAP International Conference on Statistical Physics*. Genova, Italy.

Ulmer, H., Streichert, F., and Zell, A. (2003). Evolution strategies assisted by gaussian processes with improved pre-selection criterion. In *in IEEE Congress on Evolutionary Computation,CEC 2003*, pages 692–699.

Weinstein, M. and Horn, D. (2009a). Dynamic quantum clustering: A method for visual exploration of structures in data. *Phys. Rev. E*, 80:066117.

Weinstein, M. and Horn, D. (2009b). Dynamic quantum clustering: a method for visual exploration of structures in data. *Computing Research Repository*, abs/0908.2.

Weinstein, M., Meirer, F., Hume, A., Sciau, P., Shaked, G., Hofstetter, R., Persi, E., Mehta, A., and Horn, D. (2013). Analyzing big data with dynamic quantum clustering. *CoRR*, abs/1310.2700.

Yao, X., Liu, Y., and Lin, G. (1999). Evolutionary programming made faster. *IEEE Trans. Evolutionary Computation*, 3(2):82–102.

Yu, Y., Qian, F., and Liu, H. (2010). Quantum clustering-based weighted linear programming support vector regression for multivariable nonlinear problem. *Soft Computing*, 14(9):921–929.

APPENDIX

Used test functions (Ulmer et al., 2003; Ahrari and Shariat-Panahi, 2013; Himmelblau, 1972; Yao et al., 1999; Mishra, 2006).

Alpine:

$$f_1(\boldsymbol{x}) = \sum_{i=1}^{n} |x_i \sin(x_i) + 0.1x_i|, \tag{12}$$

$-10 \leqslant x_i \leqslant 10$ with $\boldsymbol{x}^* = (0,\ldots,0)$ and $f_1(\boldsymbol{x}^*) = 0$.

Goldstein-Price:

$$\begin{aligned} f_2(\boldsymbol{x}) =& (1+(x_1+x_2+1)^2 \cdot \\ & (19-14x_1+2x_1^2-14x_2+6x_2x_2+3x_2^2)) \cdot \\ & (30+(2x_1-3x_2)^2 \cdot \\ & (18-32x_1+12x_1^2+48x_2-36x1x2+27x2^2)), \end{aligned} \tag{13}$$

$-2 \leqslant x_1,x_2 \leqslant 2$ with $\boldsymbol{x}^* = (0,-1)$ and $f_2(\boldsymbol{x}^*) = 3$.

Himmelblau:

$$f_3(\boldsymbol{x}) = (x_1^2+x_2-11)^2 + (x_1+x_2^2-7)^2 \tag{14}$$

$-10 \leqslant x_1,x_2 \leqslant 10$ with $f_3(\boldsymbol{x}^*) = 0$ at four identical local minima.

Bohachevsky 1:

$$f_4(\boldsymbol{x}) = x_1^2 + 2x_2^2 - 0.3\cos(2\pi x_1) - 0.4\cos(4\pi x_2) + 0.7, \tag{15}$$

$-100 \leqslant x_1,x_2 \leqslant 100$ with $\boldsymbol{x}^* = (0,0)$ and $f_4(\boldsymbol{x}^*) = 0$.

Generalized Rosenbrock:

$$f_5(\boldsymbol{x}) = \sum_{i=1}^{n-1} ((1-x_i)^2 + 100(x_{i+1}-x_i^2)^2), \tag{16}$$

$-2048 \leqslant x_i \leqslant 2048$ with $\boldsymbol{x}^* = (1,\ldots,1)$ $f_5(\boldsymbol{x}^*) = 0$.

Griewank:

$$f_6(\boldsymbol{x}) = 1 + \frac{1}{200}\sum_{i=1}^{n} x_i^2 - \prod_{i=0}^{n} \cos(\frac{x_i}{\sqrt{i}}), \tag{17}$$

$-100 \leqslant x_i \leqslant 100$ with $\boldsymbol{x}^* = (0,\ldots,0)$ $f_6(\boldsymbol{x}^*) = 0$.

Zakharov:

$$f_7(\boldsymbol{x}) = \sum_{i=1}^{n} x_i^2 + (\sum_{i=1}^{n} 0.5ix_i)^2 + (\sum_{i=1}^{n} 0.5ix_i)^4, \tag{18}$$

$-5 \leqslant x_i \leqslant 10$ with $\boldsymbol{x}^* = (0,\ldots,0)$ $f_7(\boldsymbol{x}^*) = 0$.

Sphere:

$$f_8(\boldsymbol{x}) = \sum_{i=1}^{n} x_i^2, \tag{19}$$

$-5 \leqslant x_i \leqslant 5$ with $\boldsymbol{x}^* = (0,\ldots,0)$ $f_8(\boldsymbol{x}^*) = 0$.

Chichinadze:

$$\begin{aligned} f_9(\boldsymbol{x}) =& x_1^2 - 12x_1 + 11 + 10\cos(\pi x_1/2) \\ & + 8\sin(5\pi x_1) - (1/5)^{0.5} e^{-0.5(x_2-0.5)^2}, \end{aligned} \tag{20}$$

$-30 \leqslant x_1,x_2 \leqslant 30$ with $\boldsymbol{x}^* = (5.90133, 0.5)$ $f_9(\boldsymbol{x}^*) = -43.3159$.

Booth:

$$f_10(\boldsymbol{x}) = (x_1+2x_2-7)^2(2x_1+x_2-5)^2, \tag{21}$$

$-20 \leqslant x_1,x_2 \leqslant 20$ with $\boldsymbol{x}^* = (1,3)$ $f_9(\boldsymbol{x}^*) = 0$.