BEAM TRANSPORT AND OPTIMIZATION TOOLS BASED ON EVOLUTIONARY STRATEGIES

L. Catani
Istituto Nazionale di Fisica Nucleare - Sezione di Roma II, ROMA ITALY

Abstract
This paper presents experimental results and off-line evaluations of a software tool for automatic beam transport optimization that uses an optimization algorithm based on Evolutionary Strategies, developed for the LISA 25 MeV linear superconducting accelerator of INFN-LNF.
The main advantage of this approach is that the definition of the optimization procedure doesn't require the analysis of the beamline. In this way the optimization of the beam shape, or its automatic steering, can be obtained even when the beamline's model is unavailable or insufficient diagnostic impede the definition of the response matrix.

1. INTRODUCTION

Genetic and Evolutionary Algorithms are well known techniques for numerical optimization. They allow a wide search for the optimum configuration of a given system by generating and selecting best sets of input parameters on the basis of some previously defined quality factor. The latter, to be defined in such a way as to best represent the desired characteristic and the behavior of the system to be optimized, is the only information needed by the algorithm. Experimental results and off-line evaluations of a software tool for automatic beam transport optimization that uses an optimization algorithm based on Evolutionary Strategies (ES), developed for the LISA 25 MeV linear superconducting accelerator of INFN-LNF [3] are presented in the following. The main advantage of the ES approach in our case is that the definition of the optimization procedure does not require a detailed understanding of the beamline behavior. Optimization of the beam shape, or automatic beam steering, can thus be obtained even when detailed modeling of the beamline is not available or when insufficient diagnostic impedes the definition of the response matrix. The described preliminary experiment evidenced other benefits of the technique such as its ease of development and implementation, independence from readout noise and set point drifts, and a high degree of flexibility that makes the tool particularly useful for setting up at commissioning and later optimization of the beamline. An off-line study of the algorithm has also shown that its main limitation, speed, may be improved with an appropriate choice of the generating operator and the quality factor. Other possible applications can be envisaged, such as the simultaneous optimization of beam parameters obtained from mixed type of diagnostics (e.g.BPMs + Toroids + visual beam spot position detectors +...) and the fine tuning of beamline models to be used by model-based optimization algorithms.

2. APPLYING EVOLUTION STRATEGY TO BEAM TRANSPORT OPTIMIZATION

The basic idea of the algorithm used for the transport optimization comes from the so called Evolutionary Strategy [1, 2] an attempt to translate into mathematical form the optimization method of biological systems.

Following the ES, the optimization of a set of N variables ("*genes*") x_i, i=0,..N-1, starts from an initial population of *M* possible configurations P_j, j=0,..M-1. These will be the *parents* of a new generation of *D* new configurations P'_k, k=0,..D-1 produced applying the ES operators to P_j. Using the Genetic Algorithms terminology those operators can be called as *selection*, *reproduction*, *crossover*, *mutation*. In general they select the parental points, generate new configurations (descendants) mixing their values x_i, add a random mutation and evaluate for every P'_k its fitness or quality factor, the latter being the only information given to the algorithm to perform optimization. Among the *D* new points, the *M* fittest ones will be chosen as parents for the next generation. The key feature of the algorithm is that the generation of the offspring is performed considering not only the genes of the parents but also the mutation amplitudes that generated them. One thus selects both the best values and the most effective mutation amplitude toward the optimized value. The (*M+D*)–*strategy*, the M best of all M+D individuals survive to became parents in the new generation, has been chosen for this application.

The descent to the deepest point of a surface in a N-dimensional space is the pictorial view of the process. If the surface is smooth and represented by a monotone function of all x_i (as in the minimization of N independent real numbers) then the procedure will easily find the optimized parameter's configuration. However, in the optimization, say, of a set of beam correctors in minimizing the RMS displacements • at some downstream Beam Position Monitors (BPM), the defined quality factor is obviously not a monotone function of the corrector values x_i. Moreover, the larger is the number of the correctors, the harder the optimization will be.

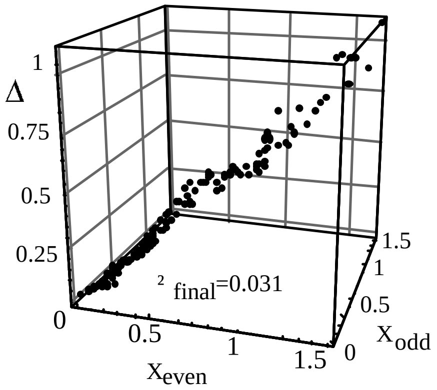

Fig. 1a

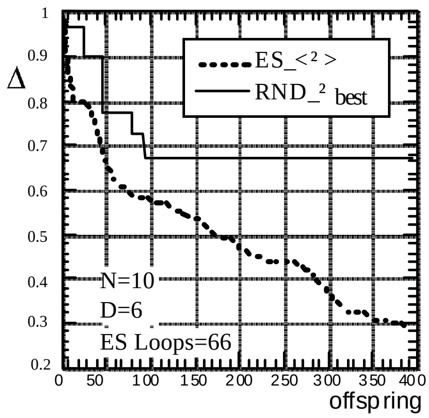

Fig. 1b

In spite of that, the algorithm has proven to be able to find the way to achieve a better than the initial correctors configuration. Fig.1a shows the 3D optimization path obtained from the simulation for a system of 8 correctors trough 200 generation loop. Each point in the graph is the average of all parent values for every generation; its coordinates are (X_{even}, X_{odd}, •) where $X_{even} = \sqrt{\sum_i \tilde{x}_{2i}}$, $X_{odd} = \sqrt{\sum_i \tilde{x}_{2i+1}}, i = 0,..\left[\frac{N}{2}\right]$; $\tilde{x}_i = \left|x_{f_i} - x_i\right|$, x_i and x_{f_i} are respectively the values of i-th corrector is its final value at the end of the procedure. One we can see from the graph that the procedure defines a search path to the optimized configuration, while the result of the application to the same problem of a simple random search would have been, in the same 3D graph, a cloud of scattered point. The effectiveness of ES with respect to a random search RND is again evident in Fig.1b. The RND curve represents the best • produced after a number of iterations, 396 in total. For the ES curve, the 396 new configurations are produced in 66 loops of 6 descendants. For every loop the parents average fitness is shown.

3. EXPERIMENTAL RESULTS

In the experimental set up N_s=6 was the number of controlled correctors and • the RMS displacement of the beam to be minimized on the N_{BPM} downstream beam position monitor. M=4 parents (at the beginning all identical and equal to original corrector's values) and D=12 offspring per generation were used.

The results of one of the tests are presented in Fig.2 showing the • vs. generation number curve and the initial and final beam position at the chosen BPMs.

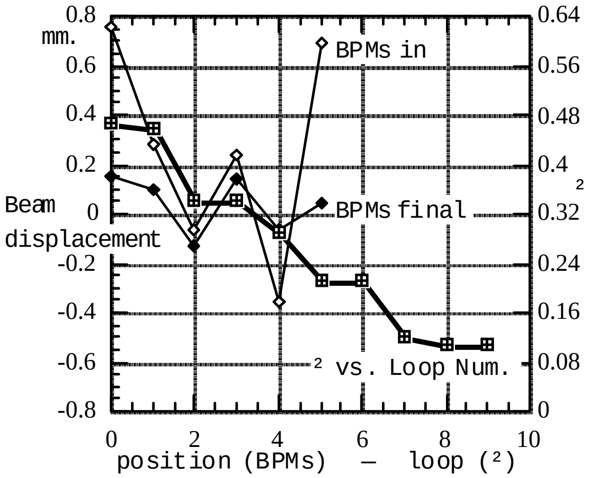

Fig. 2 Result of Transport Optimization Algorithm.

In this particular case the optimization started from a bad transport condition and after a few loops a better corrector's configuration was found. The • curve stopped decreasing when the values of the BPM signals were of the same order of the readout noise.

Instrumentation noise as well as the modifications of the beamline optics performed while optimizing have been evidenced as sources of disturbance for the process but did not prevent the evolution toward an optimized configuration, although they did in some case influence its convergence speed. Parasitic effects like electrostatic charging of the beam pipe or power supplies drift can be considered as undesired (usually weak) slow varying extra correctors; their effect will also be compensated by the procedure.

For on-line optimization applications of the method a reasonable speed of evolution towards the optimized configuration is also required. In this respect an unavoidable limitation of the method is that the evaluation of quality factors requires every new generated configuration to be applied to the beamline and the beam response to be measured.

In this preliminary application approximately 6 minutes were needed to run trough the 9 generations shown in Fig.2 when, with the injector repetition rate set at 1Hz, a 3 sec. delay was allowed between the setting of new values of correctors and the readout of BPMs.

4. SIMULATIONS AND ALGORITHM OPTIMIZATION

Evolutionary Strategies techniques have two main kinds of "tuning knobs" to adjust the algorithm to the particular application. The first are the operators applied at each generation to generate and select new offspring. The second is the definition of the quality factor that best fits the system and the desired result while giving the proper feedback to the optimization algorithm.

For instance, in the original version of the procedure every new offspring was generated selecting each of its *genes* x_i at random from anyone of the actual parents; then a random mutation was applied. One can argue that this works well when the x_i are independent which is not the case here because the value of a corrector clearly depends on those of preceding ones. Instead of this mixing of variables one could use something like a *clonation* or a mixing of two halves of correctors set. One should also recall that, while mixing of parent genes is a fundamental requirement for the ES to be able to search through the largest number of configurations, for this particular on-line application the main task is not to find the absolutely best configuration but only a better than the initial one and in a reasonable time. The described modification of the ES algorithm does actually improves its speed for particular applications, such as the trajectory optimization one we are considering while it may be less effective for other kinds of optimization. The selection of the parents to be duplicated should be guided by their quality factor and the latter could even be used to control the mutation amplitude, in order to have a fine search when the procedure is approaching the optimum. In defining the quality factor it is also advantageous to use a weighted RMS displacement at the BPMs rather than a pure RMS one. With appropriate weight functions assigned *a priori* to the various correctors, the procedure will start optimizing first the elements at the beginning of the beamline, as a human operator would.

To analyze these modifications four types of optimization algorithms have been defined and tested by simulations on a test beamline; the results of several runs have been statistically analyzed. For all runs the number of loops allowed n_L depended on the number of offspring to be produced and tested at every generation loop, D, in such a way as to keep $D•n_L=h$, with h=240. In evaluating the performance of the different configurations, execution time was not considered, on the assumption that, in an actual on-line application, it would have been the same for all.

The graph in Fig.3a shows distributions of • values, where $•=<•_{initial}/•_{final}>$, for all examined types of optimizing algorithm including the original one *t1*. The poorer performance of type *t1* is, we believe due to the random selection of corrector values from different parental configuration as in the original ES algorithm. Fig.3b shows the same distribution curves for type *t4* algorithm and the same six correctors but differently weighted for the purpose of computing the quality factor. The weight function used is $\exp\left(-{}^{i}\!/\!_{N\cdot\tau}\right)$ where *N* and *i* represent respectively the number of correctors and their order as seen by the beam.

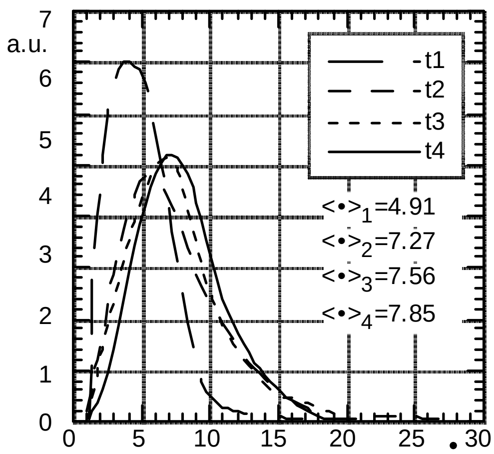

Fig. 3a

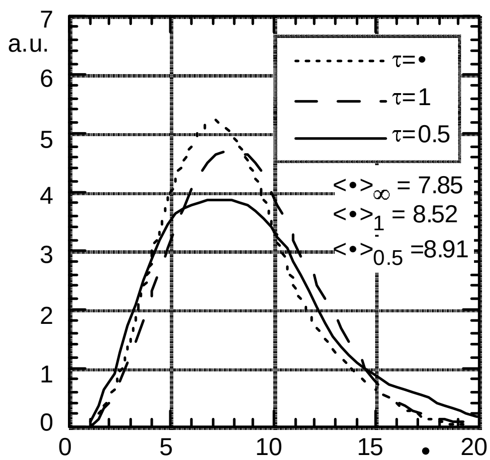

Fig. 3b

The parameter τ is the weight function parameter. The lower the value of τ, the higher will be the weight of the first correctors in the quality factor •. The graph shows that an appropriate choice of τ produces better performances; i.e. the algorithm finds better correctors configurations after the same number of iterations.

A weak dependence on the number of descendants D is also evidenced, particularly when few correctors are used; in such cases, for a given value of the product D•Loops, it is better to have larger number of generations. Random fluctuation of readout signals (noise) proved to be less dangerous than expected. The RMS noise level mainly determines the lowest achievable value of •. Finally, for a given algorithm the highest evolution speed should follow from the best compromise between the number D of descendants (i.e. the evaluations to be performed) per generation loop, and the reduced ability of the algorithm to evolve towards the optimized configuration resulting from a limited offspring (and parents) population.

5. ANOTHER EXAMPLE

With the beamline simulation program we have also tried to test the application of the algorithm to other optimization problems. The following example shows the optimization a beamline consisting of quadrupoles and drifts and equipped with a set of six dipole correctors. The beam is assummed to have a constant, given energy spread and the algorithm does search for the corrector configuration that gives both the best beam alignment along the channel and the lowest beam size at the end of it, taking into account the dispersion introduced by the correctors themselves. The following pictures show the evolution of trajectories (Fig.5a), beam envelope (Fig.5b) and corrector values (Fig.5c) before and after (thick lines) optimization and at intermediate optimization stages (i.e. after 5, 10, 100, 200 generations). Fig.5d compares the dispersion function, computed using *BeamOptics* [6], before and after (thick line) optimization. It can be seen that both the beam alignment and its size at the end of the beamline have been reduced. Furthermore, beam size is actually reduced all along the channel, mainly through reduction of the corrector RMS strength (almost halved) that minimizes and compensates the corrector generated dispersion. Similar results have been obtained even with non zero dispersion at the entrance of the channel, showing that the method can compensate an unknown initial dispersion. It is worth reminding once again that we cannot prove that the result obtained is the best possible achievable for the given beamline and, obviously, we cannot guarantee that an identical result is produced at each attempt. One has rather to stop the process once a good enough for the purpose result has been obtained.

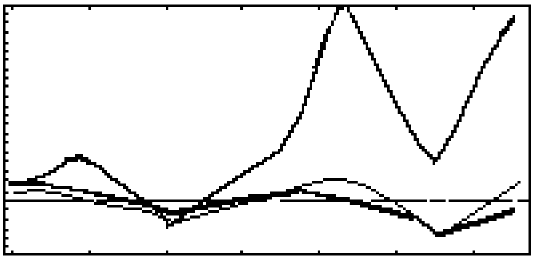

Fig.5a – Beam Trajectories

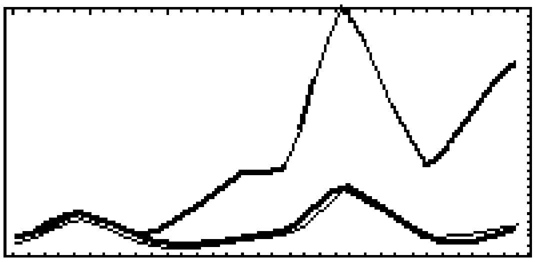

Fig.5b – Beam Envelope

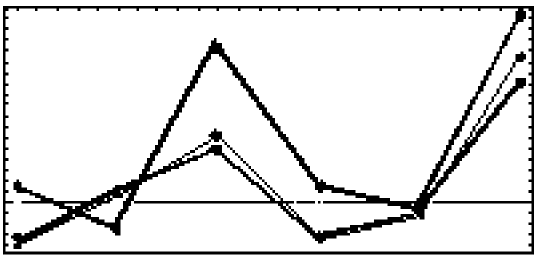

Fig.5c – Steerer Values

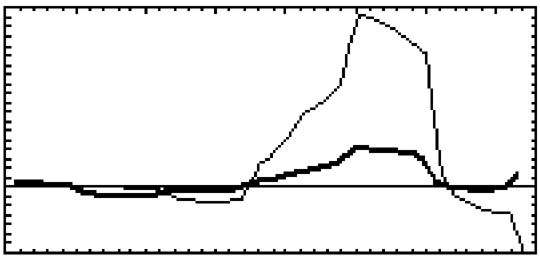

Fig.5d - Dispersion

6. CONCLUSION

While the described approach to automatic beam optimization may not be able to compete with model based or analytical methods, it can be extremely valuable when the previous ones cannot be used either because a model is not available or because insufficient diagnostic prevents one from defining a proper corrector matrix. Furthermore, it can also be easily configured to optimize a set of different parameters and beam characteristics, as seen in the previous example, or to fine tune the model to be used by other, model based optimization systems.

A rather long execution time of the optimization procedure is the main drawback of the proposed ES algorithm. A fast control system and a high beam repetition rate help to reduce the evaluation time of input configurations, and thus the overall execution time. Obviously because the settings produced during the process are not predictable, a low power operation mode must be available so that the beam can be driven through any trajectory without risking damage to the hardware. Further studies of the technique are foreseen, to investigate other possible applications and to improve mainly its speed and efficiency.

REFERENCES

[1] I.Rechenberg, "Evolution Strategy: Nature's way of Optimization", From: "Optimization: Methods and Applications, Possibilities and Limitation" — Lecture Notes in Engineering, Vol.47: Ed. H.W.Bergmann. Springer, Berlin 1989

[2] I.Rechenberg, "The Evolution Strategy. A Mathematical Model of Darwinian Evolution" — Springer Series in Synergetics, 22, Springer, Berlin-Heidelberg

[3] L.Catani "A Simple Tool for Beamline Commissioning and Transport Optimization" — CAP96, Williamsburg, Virginia 1996 – AIP Conference Proceedings 391 309-315

[4] M.Castellano et al. "Commissioning and Operating Experience with the LISA Superconducting Linac" — Nucl. Instr. Meth. in Phys. Res. A391 (1997) 385-397

[5] L.Catani et al., "Commissioning and Operation of the Control System of the LISA Superconducting Linac" — Nucl. Instr. Meth. in Phys. Res. A352 (1994) 71-74

[6] B.Autin et al., "*BeamOptics* A Program for Analytical Beam Optics" —CERN.Geneva, August 1998 - CERN 98-06