

Simultaneous optimization of beam emittance and dynamic aperture for electron storage ring using genetic algorithm

Weiwei Gao, Lin Wang, and Weimin Li

National Synchrotron Radiation Laboratory, University of Science and Technology of China, Hefei, Anhui, China
(Received 25 May 2011; published 14 September 2011)

Finding a high quality of lattice that simultaneously meets low beam emittance performance and acceptable dynamic aperture is a challenging task for the storage ring-based light source, especially for the next generation storage ring which is characterized with ultralow beam emittance. This paper presents an alternative method, based on the concept of genetic algorithm, to simultaneously optimize the beam emittance and dynamic aperture for low emittance lattice. Instead of analyzing the nonlinear indicators extracted from the high order nonlinear map, the algorithm can globally optimize the nonlinear performance by the direct dynamic aperture tracking result. So this method is more straightforward and efficient than analyzing the nonlinear driving terms. In order to illustrate this method, the quadrupole and sextupole strengths of a five-bend-achromatic lattice are simultaneously optimized by nondominated sorting genetic algorithm II (NSGA-II). Finally, the optimal linear optics for ultralow emittance lattices with better dynamic aperture are obtained. The result shows that the algorithm is particularly useful for the low emittance lattice design, where the beam emittance and the dynamic aperture always conflict with each other.

DOI: 10.1103/PhysRevSTAB.14.094001

PACS numbers: 29.20.db, 29.27.Bd

I. INTRODUCTION

The third generation synchrotron radiation light sources, with the order of nm rad beam emittance, have achieved unprecedented success in the fields of synchrotron radiation application, such as ESRF [1] and ALS [2,3]. To enhance the radiation brightness and transverse coherence, the next generation electron storage ring or ultimate storage ring is aiming at several tens of pm rad emittance. The equilibrium beam emittance in the storage ring is determined by radiation damping and quantum excitation processes. Generally, it can be expressed as a function of beam energy and optical functions as well as the bending angle of each dipole magnet. For a specific storage ring, the energy range of the beam is determined by scientific requirement, because it has a close relationship with the undulator radiation spectrum. Similarly, there are finite choices on the dipole magnet parameters while considering affordable storage ring circumference and construction capital. The simple but more efficient and economic way to minimize emittance is by adjusting the quadrupole magnet strength parameters to change the optical functions, such as betatron function and dispersion function, to fit the optimum beam emittance condition. Nevertheless, the design task becomes more and more demanding on our ability to achieve ultralow beam emittance and to control nonlinear aberrations simultaneously. The strong focusing

lattice, in general, is characterized with large natural chromaticity and small dispersion function. A deleterious result of this kind of lattice is that the dynamic aperture would be shrieked and become the limitation of storage ring performance, such as shorter Touschek lifetime, poorer injection efficiency, etc. In order to get high brightness and long beam lifetime, the low emittance lattice with adequate dynamic aperture becomes an important issue on the storage ring design.

In the past ten years, the optimization of dynamic aperture based on the strength of nonlinear driving terms has gained a comfortable result and been applied to many storage ring designs. Based on this method, several nonlinear optimization approaches have been developed to manipulate the dynamic aperture with appropriate choice of sextupole strengths and positions. For example, the sextupole scheme for the Swiss Light Source (SLS) is successfully optimized by OPA code [4]. Alternatively, the interleaved or noninterleaved sextupole scheme has been applied to optimize dynamic aperture for low emittance lattice. The subtle arrangement of sextupoles is essential to cancel the nonlinear effects introduced by each other [5–7]. But few such cases exist in realistic storage ring-based light sources. Furthermore, the octupoles as well as other multipoles [8] can be inserted into the lattice to cancel some aberration terms.

Normally, the above-mentioned methods optimize the linear optics and dynamic aperture of the storage ring in two separate processes. First, the quadrupole strengths are optimized for linear optics, beam emittance, tunes, Twiss functions, etc. For a fixed linear lattice whose tune and Twiss functions are determined by the given linear

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

elements, the sextupole strengths should meet the linear chromaticity compensation condition and good nonlinear performance. Then, the chromatic terms, geometric terms, chromaticities, and the tune shifts with amplitude are calculated based on an analytic method, for example Lie algebra [9] or normal form theory [10,11]. The best sextupole parameters are obtained by minimizing the above nonlinear parameters under the chromaticity compensation condition. Afterward, numerical tracking of dynamic aperture and other numerical methods, such as frequency map analysis [12], should be implemented to check nonlinear performance. The tune space survey, where a few quadrupole families are changed to adjust the horizontal and vertical tunes, is carried out to get the confidence of the best choice for linear optics design. In this way, reasonable betatron phase advances are supposed to improve nonlinear performance and the quadrupole strengths should be changed accordingly to minimize the nonlinear driving terms. In most cases, the linear optics design and nonlinear dynamic optimization process will be repeated several times to obtain large enough dynamic aperture with a tune far from resonance [13]. The problem with this methodology is that the iterative process will consume the designer's lots of time to get a reasonable lattice. In addition, it is hard to find the best nonlinear indicators as the dynamic aperture optimization goal.

With the proposal of an ultimate storage ring, provided with ultralow beam emittance, the issue of how to achieve the best linear and nonlinear lattice parameters turns into the first concern for light source design study. Unfortunately, the effectiveness of traditional nonlinear optimization may be limited, because there is no distinct one-to-one correspondence between dynamic aperture and nonlinear driving terms.

In this paper we exploit an optimization code based on direct dynamic aperture tracking and an artificial intelligence algorithm, multiobjective evolutionary algorithm [14,15], to simultaneously optimize the linear optics and dynamic aperture in a straightforward way. Actually speaking, the nondominated sorting-based multiobjective evolutionary algorithm (called nondominated sorting genetic algorithm II or NSGA-II), proposed by Srinivas and Deb, is used. The Pareto optimal front obtained by NSGA-II converges to the real optimal front for some test problems, and the details of this algorithm can be found in Ref. [16]. This algorithm is extremely suitable for solving the questions which are nonlinear, discontinuous, with conflicting objectives, many local optimal, and several decision variables. For the lattice design problem, we can optimize some parameters at one time, such as emittance, betatron tunes, momentum compaction factor, dispersion function, dynamic aperture, etc. In this paper, the quadrupole and sextupole strength values are varied to optimize the lattice with fixed magnet positions. The result of the code gives the global optimal solutions which compromised among all

of the objectives. The quantitative criterion of dynamic aperture tracking as an optimization goal shows powerful optimization capability, even though we do not study the physical origin restricting nonlinear performance in detail.

The paper is organized as follows. First of all, the physics computation concepts used in this optimization code are described briefly. Then we recapitulate the optimization process with an ultralow emittance lattice, whose emittance and dynamic aperture are optimized simultaneously by NSGA-II. At the end of the paper, evolution of some nonlinear parameters during optimization is given and the insight behind optimization process is discussed.

II. PHYSICS MODEL OF THE OPTIMIZATION CODE

We now briefly introduce the theories used to describe the particle motion in the optimization code.

The description of linear beam motions in the 6D phase spaces are based on the concept of the transfer matrix theory. The linear elements include drift space, quadrupole, bending magnets, and fringe field effect [17]. The linear optical parameters, such as Twiss parameters, tunes, radiation integrals [18], damping partition number, momentum compaction factor, beam emittance, etc., are calculated according to their standard definition [19–21], which is obtained by the multiplied final transfer matrix or the integration of the Twiss function in the dipole magnets.

Dynamic aperture is the particle survival space in physical space, and multiturn numerical tracking is the only reliable way to estimate dynamic aperture of the storage ring. For simplicity, the nonlinear component (sextupole) is treated by a symplectic thin lens transfer map for dynamic aperture tracking. The following exact sextupole Hamiltonian in the Frenet-Serret coordinate, as shown in Eq. (1), is used:

$$H(x, p_x, y, p_y, \sigma, p_\sigma; s) = \frac{\lambda}{6}(x^3 - 3xy^2) + \left\{ p_\sigma - [1 + f(p_\sigma)] \sqrt{1 - \frac{p_x^2 + p_y^2}{[1 + f(p_\sigma)]^2}} \right\}, \quad (1)$$

where $p_\sigma = \frac{1}{\beta_0^2} \frac{\Delta E}{E}$, p_x and p_y are the horizontal and vertical momentum, respectively, and

$$f(p_\sigma) = \frac{\Delta p}{p_0} = \frac{1}{\beta_0} \sqrt{(p_\sigma \beta_0^2 + 1)^2 - \frac{(m_0 c^2)^2}{E_0}} - 1.$$

The generating function method is used to obtain the final phase space coordinate of a particle passing through a sextupole. A more detailed derivation may be found in Refs. [22,23]. Based on these theories, the particle motion through an element can be described by a simple operator map \mathcal{M} :

$$X_f = \mathcal{M} \cdot X_i, \quad (2)$$

where $X = X(x, p_x, y, p_y, \sigma, p_\sigma)^T$, with T denoting the transposition of matrix. In Eq. (2), \mathcal{M} can be either a linear transfer map or a nonlinear operating map. X_i and X_f are the initial and final phase space coordinate of an element.

To accelerate dynamic aperture tracking speed, a simple trick is used in this code. Knowing the magnet arrangement of a given lattice, the sectionalized transfer matrix, multiplied by the linear transfer matrices between two adjacent sextupoles, needs to be calculated only once. Namely, the assembly of linear elements between the adjacent sextupoles can be considered as one matrix when tracking dynamic aperture. Hence, the whole ring is composed of \mathcal{N} nonlinear elements (sextupoles) and \mathcal{N} linear sections. Furthermore, for getting a fast tracking speed, the sextupole length is set to be very small and the tracking process through a sextupole can be accomplished in one step. So, the tracking process is simplified. Because of the huge computation task, fast tracking speed is an important issue for the practical dynamic aperture optimization.

III. SIMULTANEOUS OPTIMIZATION BY NSGA-II

A. Feasibility and superiority

The perturbation theory or normal form analysis indicated that the nonlinear Hamiltonian can be divided into phase-independent terms, such as chromaticities and tune shifts with amplitude, and phase-dependent terms, such as various order of nonlinear driving terms, which drive nonlinear resonances and may be the main reasons of nonlinear performance limitation. The discovery reveals that the beam instability condition is not only determined by the strength of nonlinear elements itself, possibly closely related to the betatron amplitude at nonlinear elements and, what is more, the betatron phase advances between nonlinear elements. This means the strengths of nonlinear driving terms are correlated with the linear optical function. So, the simultaneous optimization of linear optics and nonlinear beam dynamics becomes a preferable method for storage ring design.

Recently, with the increasing speed of a computer, optimization based on direct dynamic aperture calculation, instead of nonlinear resonance theories, becomes a powerful tool for lattice designers. The advantage of the above method is ignoring the indistinct relation between dynamic aperture and nonlinear driving terms; for example, BETA code used in SOLEIL [24] and ALBA [25] light source design. At the same time, various kinds of artificial intelligence algorithms showed a surprising capability in the parameter optimization problem, such as simulated annealing, particle swarm optimization, and genetic algorithm. Thus, motivation of the simultaneous optimization of linear optics and dynamic aperture for ultralow emittance lattice becomes reality.

The numerical optimization concept has been used in the field of accelerators and gained expected achievements in storage ring [26] lattice design, damping ring [27] lattice design, and other accelerator system design [28–30]. The simultaneous optimization concept leads us to globally search for the solutions which are compromised between all of the objectives. A genetic algorithm may be the first choice to carry out the optimization process because it can simultaneously optimize several objectives without giving the weight of every objective. What is more, it is especially suitable for solving the nonlinear problems without a detailed understanding of the intrinsic physical principle. Thus, it provides an easy method for lattice design. First, we can suppose a lattice in our mind and set limits to variable parameters according to available technology. The parameters we want to maximize or minimize are considered as the objectives. At the end of the optimization, the algorithm will put out the results compromising among all of the objectives and satisfying our constraint conditions. Furthermore, the result gives a global view of the proposed lattice, whether need for additional magnets or some of the magnet positions should be adjusted.

However, for large rings, which have many magnets in each superperiod, it is hard to carry out the optimization process where all of the strength parameters of the quadrupoles and sextupoles are varied in large variable ranges. In this case, high dimension space and large variable range will need huge computing capability, which means large population size and generation number are necessary. But, that is not an obstruction; we can let some of the magnet strengths to be fixed or find out some variable ranges that satisfy our design goal first. Observing the linear optics property of a lattice by the GLASS [31] method indicated that the variable parameter space is divided into several isolated regions which exhibit different optical properties. What is more, the stable solutions with low emittance property are concentrated in one or several of the regions. It is logical that, for a given lattice design goal, we can first determine the range of quadrupole strength regions which correspond to the low emittance lattice condition and provide a good behavior of linear optical functions. This process will effectively reduce the computation complexity for the next step without violating from the low emittance condition. Thus, the final global optimal solutions would not be affected by this process. Then, several families of sextupoles are used to compensate the natural chromaticity. In this step, the quadrupole and sextupole strength parameters are simultaneously optimized by the NSGA-II algorithm whose objective is dynamic aperture and beam emittance.

Note that finding the quadrupole strength regions is an auxiliary step used to reduce the computation complexity, not obligatory for all the simultaneous optimization problems. If only few magnet strengths have to be varied (approximately no more than 10), the optimization process can be directly carried on by NSGA-II. For example, if we

want to enlarge the dynamic aperture for a given linear lattice whose sextupole strengths are the only variables, the sextupole strengths may be directly optimized by NSGA-II with the dynamic aperture and momentum aperture or other nonlinear indications as objectives.

B. Optimize an ultralow emittance lattice

The dynamic aperture for a specific low emittance lattice can be enlarged by optimizing the sextupole strengths with the constraint that the natural chromaticities are corrected to zero. In fact, the code is provided with this function. But under this condition, the influence of the linear optics, such as tune, on the nonlinear beam instability is omitted. In reality, however, it is better for us to consider the influence of linear optics on the dynamic aperture size. Therefore, it is wise for us to simultaneously optimize the linear optics and dynamic aperture with the quadrupole and sextupole strengths as variables. However, the computational time comes to the significant issue.

Luckily, observing the beam dynamics in the storage ring indicated that the linear optical functions and beam emittance are mainly determined by the linear elements only, and the nonlinear effects can be ignored when choosing the linear lattice properties. So, in order to simplify the computational complexity, we can remove the linear infeasible solutions first. This means the solutions which are unstable or with poor linear optical functions or perhaps large emittance value are discarded. This is accomplished by only varying the quadrupole strengths to obtain the feasible regions which satisfy the linear optics design goal. This process is faster, because the linear optics calculation is much quicker than dynamic aperture tracking. When the feasible quadrupole strength regions are obtained, we should use a relatively small population size to carry out the dynamic aperture and low emittance lattice optimization process. This treatment can save lots of time, and make the simultaneous optimization easy to be implemented.

To give a detailed insight of how to optimize the linear and nonlinear properties of a low emittance lattice by the presented scheme, we optimize a very compact five-bend-achromatic lattice in this section to show how this method works. This lattice is designed for Hefei Advanced Light Source (HALS), which is a proposal of advanced ring-based vacuum ultraviolet and soft x-ray source.

TABLE I. The main design properties of the ultralow emittance lattice.

Circumference	420 m
Energy	1.5 GeV
Number of superperiod	20
Focusing type	Five bend achromatic
Nature emittance (design goal)	<0.075 nm rad
Straight section for insertion devices	20×5.634 m

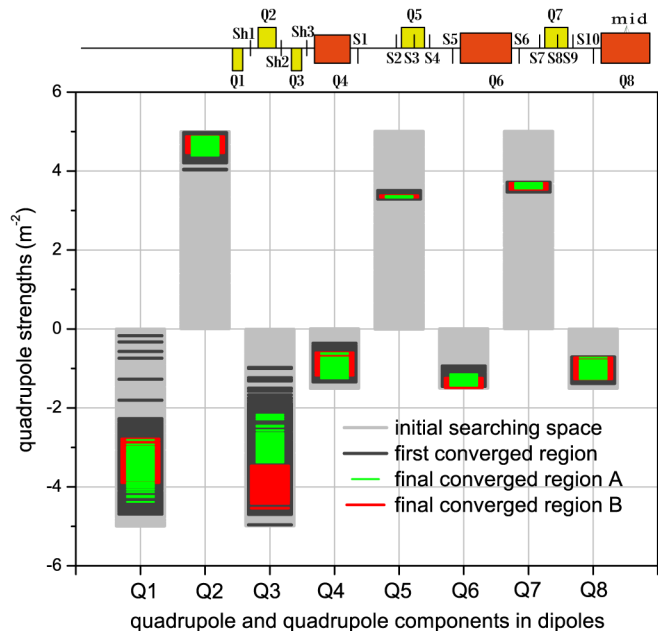


FIG. 1. The quadrupole strengths and the quadrupole components of the dipoles. Q1, Q2, and Q3 are the quadrupoles in the dispersion-free section, Q5 and Q7 are the quadrupoles in the dispersion section, while Q4, Q6, Q8 are the quadrupole components in the dipoles. The light grey dots are the initial searching spaces of each quadrupole magnet, the dark grey dots are the first time converged region, and the multicolor dots are the second time convergence regions. All of these converged regions satisfy the design emittance goal: $\epsilon_x < 0.075$ nm rad.

The main properties of this five-bend-achromatic lattice are summed up in Table I. The magnet arrangements are predetermined considering space for insertion devices, acceptable ring circumference, and other accelerator components. In order to enhance damping wiggler effects, the bending magnet length of the outer dipoles and inner dipoles are chosen to satisfy the low emittance condition for the achromatic condition [32]. What we should do is just to vary the quadrupole and sextupole magnet strengths, as given in Fig. 1, to obtain an ultralow emittance lattice with better dynamic aperture. This may be the most important procedure for the ultralow emittance lattice design.

1. Find the quadrupole strength regions satisfying the low emittance condition by single objective genetic algorithm

The goal of this step is to find out one or more convergence region of quadrupole strengths that meet the low emittance condition, so as to minimize the searching space for the next step without violating from the low emittance design goal.

The linear lattice optimization work has been carried out previously. The systematic optimization process by MOEA was reported in Ref. [33]. Now we use the single objective genetic algorithm instead of MOEA to find quadrupole

strengths, because there is only one objective, emittance, to be minimized. If other properties have to be optimized the MOEA may be more appropriate.

The betatron function is a critical parameter for linear lattice optics, and also the dispersion function. A storage ring's acceptance is determined by the maximum betatron function and vacuum chamber aperture (or dynamic aperture). The higher maximum betatron function is, the smaller acceptance will be. Also, many indication parameters are related to the betatron function, such as linear orbit error amplification factors, nonlinear driving terms, stop band widths, etc. [20,34]. So we hope to restrict the maximum betatron function at both the horizontal and vertical planes to some values. Experience shows that the maximum value is less than 40 m for existing storage rings, and we chose this value first. For a large ring, the tune can be easily adjusted by slightly changing one or some group of quadrupole strengths, so the tune values are not within our consideration in this step. Besides, for different sextupole arrangement it is hard to say which tunes are better. Hence, the tunes together with the nonlinear properties should be optimized in the following step. For preserving the good solutions, we set only four optimization constraints first. Namely, they are

$$\begin{aligned} \beta_x(\text{maximum}) < 40 \text{ m} & \quad \beta_y(\text{maximum}) < 40 \text{ m} \\ \eta_x(\text{maximum}) < 2 \text{ m} & \quad |\eta_x(\text{ID section})| < 0.01 \text{ m}. \end{aligned}$$

It is worthwhile noticing that the achievement of ultra-low emittance generally leads to a small dispersion function due to small dipole angle and strong focusing quadrupole. Consequently, the third constraints can be eliminated for this lattice while the final results are not affected. For the common situation, however, the maximum dispersion function may need to be limited. The fourth constraint is an approximate achromatic condition in order to enhance damping wiggler effects on the beam emittance reduction.

As it is clear from Fig. 1, there are eight families of quadrupoles (including the quadrupole components in the dipoles) in each superperiod. The corresponding quadrupole magnet names and their positions are plotted on the top of Fig. 1 (half cell of mirror symmetry). In this figure, the yellow rectangular represents quadrupole, the red color rectangular indicates dipole magnet, while the black bar demonstrates sextupole. The low emittance lattices that satisfy the above constraints can be acquired by varying the strengths of the eight families of quadrupoles. In reality, however, the magnet gradient is restricted by available magnet technology. In addition, the maximum strengths of each quadrupole and the quadrupole components of dipole magnets are different. Assuming room-temperature magnets are used, the initial searching regions are given in the bottom of Fig. 1 with a light grey strip symbol.

After the determination of variable ranges, constraints, and objective, we can proceed the optimization immediately.

In this step, a population size with 3000 individuals is used and the converged optimal solutions are obtained at generation 50. The result shows that the lowest emittance value under an achromatic condition is about 0.057 nm rad, and the natural chromaticities at the horizontal and vertical planes are $\xi_x = -158$, $\xi_y = -185$, respectively. The linear lattice optical functions of this low emittance lattice are plotted in Fig. 2. However, at about 10th generation, the population finds the low emittance regions meeting our design goal and satisfying all the constraint conditions. The variable range of this converged low emittance lattice area, at the 10th generation, is plotted in variable space, as shown in Fig. 1, with dark grey color.

It is interesting to note that the quadrupole strengths of the five magnets in the arc section are converged to a small region, particularly Q5 and Q7. The strengths of these two quadrupole families are about 3.3–3.4 m⁻² and 3.6–3.7 m⁻², respectively. The reason is obvious, that only one quadrupole located between the dipoles controls dispersion functions, which has a major contribution to the emittance value. The available Q3 strengths are located in large variable ranges, which means for a given Q3 strength we can change the combination of the other magnet strengths in the dark grey color range to get a solution satisfying the above constraints. Also, we observed that the optical function corresponds to the lowest emittance value, as shown in Fig. 2, and is not as prominent as we expected. Especially, the horizontal betatron function is too small in the straight section and the natural chromaticities are too large. So that, for further minimizing the quadrupole strength ranges and getting more appropriate linear optics, we can set the constraints more strictly to remove the useless solutions by repeating the above optimization process. The variable ranges are constrained to the dark grey color range. The constraints are as listed in the following:

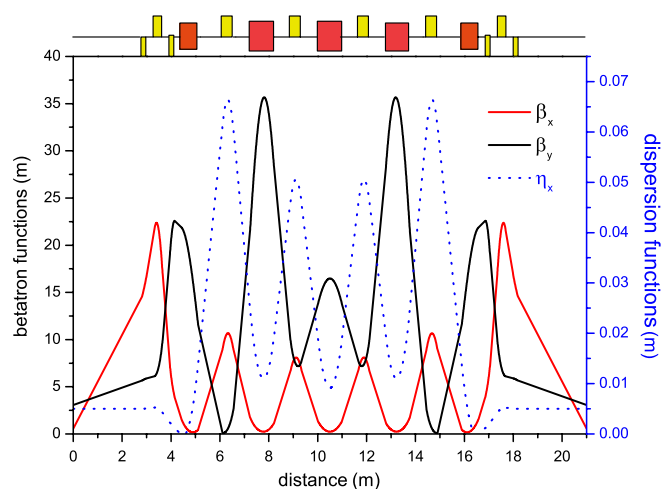


FIG. 2. The betatron function and dispersion function of the minimum emittance lattice. The emittance value is 57 pm rad.

$$\begin{aligned} \beta_x(\text{maximum}) &< 25 \text{ m} \\ \beta_y(\text{maximum}) &< 25 \text{ m} \\ 5 \text{ m} < \beta_x(\text{ID section}) &< 15 \text{ m} \\ \beta_y(\text{ID section}) &< \beta_x(\text{ID section}) \\ \eta_x(\text{maximum}) &< 2 \text{ m} \\ |\eta_x(\text{ID section})| &< 0.005 \text{ m.} \end{aligned}$$

In this repeating process, for obtaining smaller natural chromaticity the constraints of the maximum betatron functions are restricted to a small value: 25 m. The constraint of the horizontal betatron function in the insertion device section is limited within 5–15 m, which is selected for local bump injection. The values for betatron function constraints would not be exclusive, and we believed that these values should be different for different design purpose and different designers. In this repeating process, however, the variable ranges are relatively minimized, defined in the already obtained dark grey color range, and can be accomplished in a few hours. After the repeated process the feasible solutions satisfying the design objective with emittance value smaller than 0.075 nm rad are shown in Fig. 1 with green dots and red dots.

Although, the above two processes can be accomplished in a single process by getting the constraints as strict as possible. But, for large variable searching problems, it is wise for us to use the repeating optimization process to remove the useless regions step by step. This process should save lots of computation time and space, since it would need smaller population size and smaller generation number at each step. What is more, for every repeating process the variables are constrained to the more reasonable ranges, and the useless regions (for example, the unstable regions) are not calculated. This is beneficial to getting the best solutions.

Note that for this lattice the constraint of the absolute value of dispersion function at the straight section can be neglected. For this situation minimizing the emittance and the absolute value of dispersion function at the straight section are two objectives; as a result it turns to be multiobjective problems. The obtained minimum emittance values of achromatic lattice are not different, regardless of multiobjective or single objective evolutionary algorithm. However, if other objectives have to be minimized or maximized, the multiobjective will show its superiority. Hence, the multiobjective optimization might also be useful for this step.

2. Simultaneously optimize the dynamic aperture and emittance by NSGA-II

In this step we realize a rather relaxed low emittance lattice with moderate dynamic aperture. Unlike the optimization based on nonlinear perturbation terms, the quantitative criterion of dynamic aperture is used to directly

judge the nonlinear performance. Using the dynamic aperture as the optimizing objective is more efficient than using the nonlinear driving terms, since there is no need for us to optimize the weight of each driving terms.

To begin, we should first explain the quantitative criteria of dynamic aperture. Having this numerical standard, the code can automatically evaluate which lattice parameters are better in nonlinear performance. Our quantitative standard of the dynamic aperture in the optimization process is the average radius and the minimum radius of the dynamic aperture. It is based on a circle as shown in Fig. 3.

In general, the requirements of dynamic aperture size in the horizontal and vertical directions are different. Usually, larger horizontal direction size is preferred. If the minimum requirement of dynamic aperture size in the horizontal and the vertical direction are determined, we can make both directions to be equal by an appropriate scale. Hence, the desired dynamic aperture is converted to square from rectangular and the above quantitative standard of dynamic aperture is usable. In other words, for simplicity the real shape dynamic aperture (almost ellipse) is transformed to a circle in this numerical standard. The initial tracking particles are uniformly distributed in the square. The distance between the boundary lost particles' position and the origin point is viewed as one radius of the dynamic aperture. The average radius R_{avg} can be obtained by averaging every boundary lost particle's radius. The smallest one of all is known as the minimum radius R_{min} . These two radii are clearly plotted in Fig. 3. A more superior solution is the dynamic aperture with not only large average radius but also large minimum radius to remove the cases that stable islands exist in dynamic aperture. If both of these two radii are maximized the shape of the dynamic aperture is close to a circle (close to ellipse for realistic), and then the irregular bump shape of dynamic aperture can be avoided.

Having this numerical standard, we now evolve the initial population to minimize the beam emittance and maximize the dynamic aperture by NSGA-II at the same time. As it is clear from Fig. 1, the feasible regions of low emittance lattice are located in two regions (named as

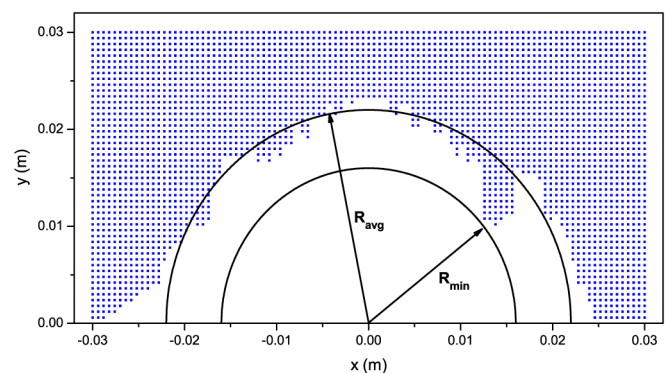


FIG. 3. The quantitative standard of dynamic aperture. R_{min} is the minimum radius, and R_{avg} is the average radius.

region A and region B), where there is no clear boundary except for Q3. In this step, we have two choices to carry out the simultaneous optimization process: the first one is vary the quadrupole strengths continuously in this two regions, which means the two regions can be viewed as one continuous region. The second strategy is optimizing the two regions separately, and the idea is to reduce the computational resources with better solutions.

We chose convergence region A as shown in Fig. 1 with green color for further optimization because, except for Q3, the other quadrupole strengths are much closer for region A and region B. Hence, for an economic way it reasonable for us to first consider region A, with smaller Q3 strength, for further optimization. In this step the quadrupole strength not only contributes to emittance, but also dynamic aperture. This is because the variation of betatron phase advances due to quadrupole strength should contribute to dynamic aperture with different nonlinear resonance excitation. Those eight families of quadrupoles together with seven families of chromatic sextupoles (S1, S2, S3, S5, S7, S8, S10) and three families of harmonic sextupoles (Sh1, Sh2, Sh3) are varied to obtain the largest dynamic aperture solutions under the constraint that the natural chromaticity is corrected to near zero. Among those sextupoles, S3 and S8 are located in the middle of the quadrupoles and the quadrupole-sextupole combined magnets are used. Those sextupole names and the corresponding positions can be seen from Fig. 1. However, from Fig. 1 we observed that for each super-period one can implement ten families of sextupoles. But now, there are only seven families of sextupoles to be variables. This is accomplished by letting some of the sextupole strength be equal due to their similar contribution to chromaticity compensation. They are $S2 = S4$, $S5 = S6$, and $S7 = S9$. The above artificial action has no special meaning in nonlinear optimization and the only merit is minimizing the variable spaces, so as to minimize the computational complexity.

Except for the constraints used in the first optimization step, the constraints of chromaticities are added. It means that the natural chromaticities are corrected to nearly zero. Summarizing, the detailed constraints are as listed in the following:

$$\begin{aligned} \xi_{x,y} &< 5 \\ \beta_x(\text{maximum}) &< 25 \text{ m} \\ \beta_y(\text{maximum}) &< 25 \text{ m} \\ |\eta_x(\text{ID section})| &< 0.005 \text{ m} \\ 5 \text{ m} &< \beta_x(\text{ID section}) < 15 \text{ m} \\ \beta_y(\text{ID section}) &< \beta_x(\text{ID section}). \end{aligned}$$

Additionally, we can also choose the tune, momentum compaction factor, and natural chromaticity as constraints, based on our own problem. Of course, for different lattice

or different design goal, the constraints number can be varied easily, and does not hamper the utility of the optimization method presented in this paper.

Our ultimate goal is globally finding the solutions with minimum emittance and better dynamic aperture. Namely, maximizing the dynamic aperture size and minimizing the emittance value are two objectives in this optimization process. The quality of the dynamic aperture is composed of the following two parts:

$$DA = kR_{\min} + R_{\text{avg}}.$$

R_{avg} and R_{\min} are the radius as shown in Fig. 3, while k is a ratio factor based on the weight of those two radii. For this optimization process we chose $k = 2.5$, which means the R_{\min} is more important than the R_{avg} . Alternatively, both radii together with the emittance can be viewed as three objectives.

Having these reasonable variable ranges, constraints, and objectives, the optimization process is carried out by a population with 20000 individuals and the maximum generation number is 30. In this process, the most time consuming work is tracking dynamic aperture, so that the infeasible solutions do not necessarily have to be tracked. We should artificially evaluate the dynamic aperture for those infeasible solutions. For example, let the value equal zero, which is no better than all of the feasible solutions. For those infeasible solutions, the smaller constraints violation, the better the rank will be. This means the infeasible solutions will converge to the feasible regions generation by generation, while the feasible solutions will generate better solutions gradually. After the convergence of the population, we get the Pareto optimal front in which the solutions are compromised between dynamic aperture and the emittance. The distribution of these solutions in the objective space is plotted in Fig. 4 with black dots.

For this lattice, the three sextupoles S2, S3, and S4 locate closely and there are not enough phases for optimization. Another group of sextupoles (S7, S8, and S9) is in the same situation. But those sextupoles can reduce the strength of each sextupole. For comparison we also carried out an optimization while eliminating the combined quadrupole-sextupoles: S3 and S8. The other constraints and objectives are the same as the above. The optimal front is also given in Fig. 4 with blue dots. Result shows that the lattice with more sextupoles has better dynamic aperture than the lattice with reduced number of sextupoles.

Three solutions (named as lattice A, lattice B, and lattice C) from the database (black dots) are chosen, and the behavior of optical functions as well as the dynamic aperture of the three solutions are plotted in Figs. 5–7. In order to clearly explain the superiority of the proposed method, the dynamic aperture obtained by MAD (HARMON command) [35] is also shown in these figures. For each lattice, the two dynamic apertures are tracked by the same turns and the same particle number.

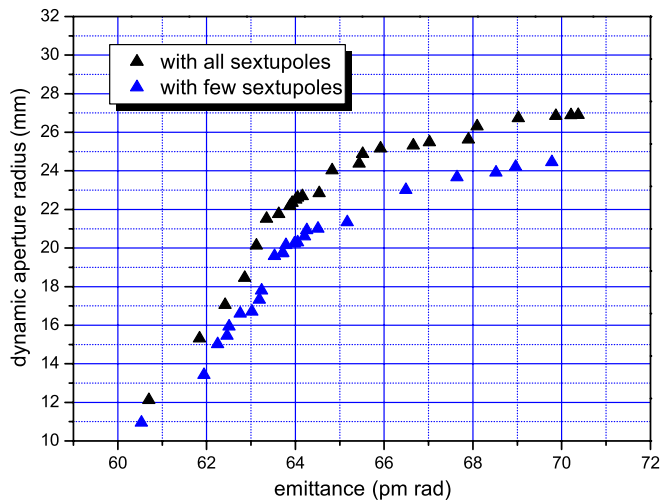


FIG. 4. The Pareto optimal front of the last generation. The horizontal axis is one of the objectives: emittance, and the vertical axis is the radius indicating dynamic aperture quantity: $DA = 2.5 \times R_{\min} + R_{\text{avg}}$. The black dots are the optimal front with all of the sextupoles on, and the blue dots are the optimal front where two families of combined sextupoles are eliminated.

Obviously, the dynamic aperture obtained by NSGA-II is much larger than the value obtained by MAD (HARMON command). On the other side, according to Fig. 4, the optimization process gives us a clear indication on the relationship between achieved beam emittance and largest dynamic aperture. Also, from Fig. 4 we will see that the dynamic aperture has a sharp increase at the emittance value about 63 pm rad. This means the dynamic aperture is difficult to be enlarged at the emittance value smaller than 63 pm rad, while the dynamic aperture size increased gradually at the emittance value between 64 to 70 pm rad. This method gives a straightforward way to obtain a relatively large dynamic aperture for low emittance lattice. The operation is very easy; what we need to do is just write the lattice input file, give the reasonable constraints as well as objectives. The remaining work will be completed by computer. At the end of the generation, the code will give us a clear database with low emittance and better dynamic aperture. The optimization results can give us another indication: there exists the lowest beam emittance with acceptable dynamic aperture, or the need of rearranging the quadrupole and sextupole positions to

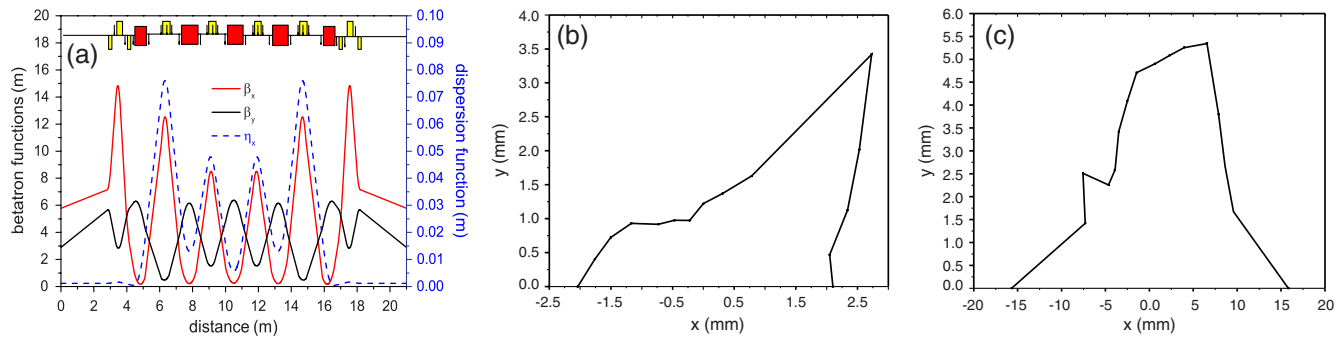


FIG. 5. Lattice A. The betatron function and dispersion function [plot (a)], the dynamic aperture obtained by the MAD HARMON command [plot (b)], and the dynamic aperture obtained by NSGA-II [plot (c)]. The emittance value is 67.9 pm rad. The tune in the horizontal and vertical plane is $\mu_x = 46.405\,505$ and $\mu_y = 26.484\,499$. The natural chromaticities of both planes are $\xi_x = -140.440\,135$ and $\xi_y = -44.136\,350$, respectively. The corrected chromaticities are $\xi_x = 0.122\,208$ and $\xi_y = 0.053\,921$.

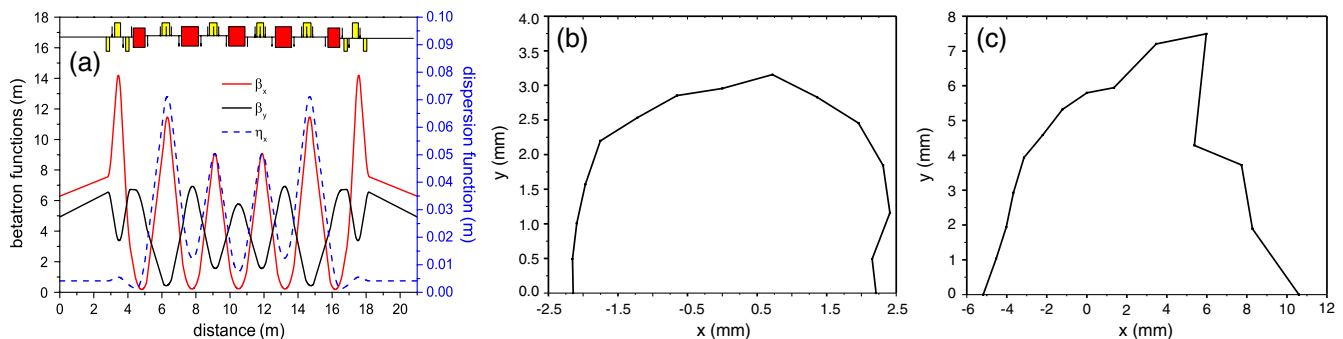


FIG. 6. Lattice B. The betatron function and dispersion function [plot (a)], the dynamic aperture obtained by the MAD HARMON command [plot (b)], and the dynamic aperture obtained by NSGA-II [plot (c)]. The emittance value is 64.1 pm rad. The tunes in the horizontal and vertical planes are $\mu_x = 46.112\,960$ and $\mu_y = 24.959\,340$. The natural chromaticities of both planes are $\xi_x = -134.567\,099$ and $\xi_y = -45.589\,299$, respectively. The corrected chromaticities are $\xi_x = 1.369\,799$ and $\xi_y = 4.075\,952$.

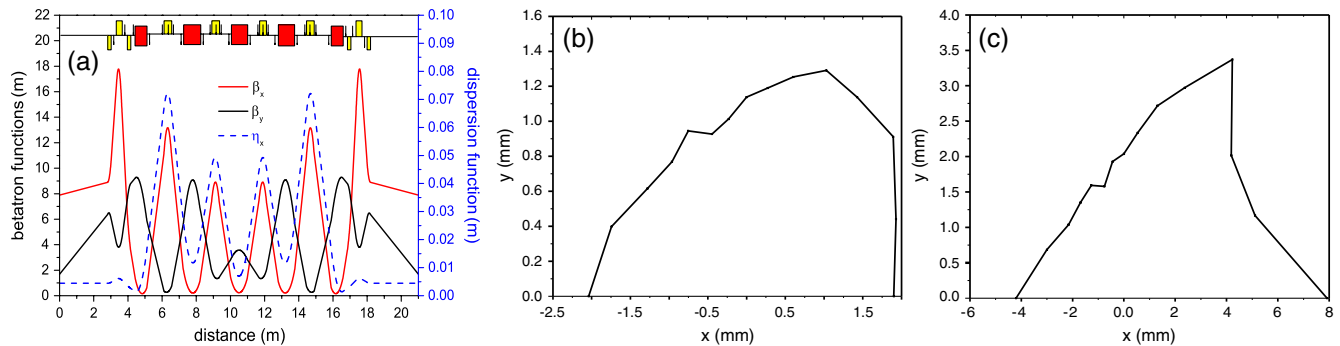


FIG. 7. Lattice C. The betatron function and dispersion function [plot (a)], the dynamic aperture obtained by the MAD HARMON command [plot (b)], and the dynamic aperture obtained by NSGA-II [plot (c)]. The emittance value is 60.7 pmrad. The tunes in the horizontal and vertical planes are $\mu_x = 45.877366$ and $\mu_y = 31.031640$. The natural chromaticities of both planes are $\xi_x = -152.995732$ and $\xi_y = -65.451753$, respectively. The corrected chromaticities are $\xi_x = 1.165934$ and $\xi_y = 0.322485$.

achieve a desirable low beam emittance and large dynamic aperture.

IV. DISCUSSION

The central ideas of the above method is the application of NSGA-II algorithm and optimization based direct dynamic aperture tracking results. The powerful optimizing capability of NSGA-II would give us confidence that the global optimum solution is found. The advantage of using direct dynamic aperture tracking results is more effective in nonlinear optimization, although the quantitative criterion of dynamic aperture is not exclusive or optimum. On the other side, the lack of this method is that we have achieved a lattice with ultralow beam emittance and acceptable dynamic aperture without a detailed understanding of the physical resources limiting dynamic aperture. The typical nonlinear indicators, such as higher order chromaticities, tune shifts with amplitude, first and second order geometric, and chromatic terms [36], are calculated for the three lattices (Fig. 5–7). As shown in Table II, we cannot find a clear decreasing nonlinear indicator. It is difficult for us to say which decreased term will bring out good dynamic aperture because one decreased term will always accompany with the other terms increased.

In order to further find the variation trend of the nonlinear driving terms with dynamic aperture size, we also plotted the tune shift with amplitudes with the emittance value (see Figs. 8 and 9). The solutions are chosen from the optimal front of some evenly spaced generations. Since the emittance value and dynamic aperture are two conflicting objectives, the emittance value increase means the dynamic aperture enlarged in these pictures.

From the two figures we will see that there is no regular variation trend of the nonlinear indicators with the emittance increased (the dynamic aperture size increased). Also, with the generation number increasing the driving terms are not decreased clearly. But the picture tells us that it is hard to find a good dynamic aperture if one of the

driving terms is too large. For example, the horizontal tune shift with amplitude should be smaller than 1000 m^{-1} etc. Because of limited calculation and study, we can roughly conclude that it is possible to get a big dynamic aperture as long as all of the nonlinear driving terms are smaller than a specific value. Unfortunately, we have made a manual change of sextupole parameters to minimize the above nonlinear indicators further and the dynamic aperture of the resulting lattice is smaller. It means we cannot say the smaller nonlinear driving terms will bring out larger dynamic aperture. We think that the above nonlinear indicators are insufficient to describe nonlinear motion for the ultralow emittance storage ring, and minimization of these

TABLE II. The nonlinear driving terms for some of the Pareto optimal solutions.

Driving terms	Lattice A	Lattice B	Lattice C
H21000	15.17	2.71	4.83
H3000	87.03	67.53	48.47
H10110	95.73	33.19	62.76
H10020	3.95	0.34	17.58
H10200	22.49	0.27	14.45
H20001	2.08	2.42	2.98
H00201	0.19	0.23	0.42
H31000	65.00	130.31	172.49
H40000	2.04	19.71	26.22
H20110	102.76	127.21	105.67
H11200	14.44	0.11	125.25
H20020	824.10	60.08	107.25
H20200	40.75	24.89	79.59
H00310	4.72	0.75	11.36
H00400	0.04	44.74	7.45
Second order ξ_x	-0.10	-0.24	-0.49
Second order ξ_y	0.26	0.20	0.12
$d\mu_x/dx$	-240.45	-338.12	-554.32
du_y/dy	-160.95	-152.52	-432.78
$d\mu_{xy}/dy, x$	-68.43	-50.22	-21.70

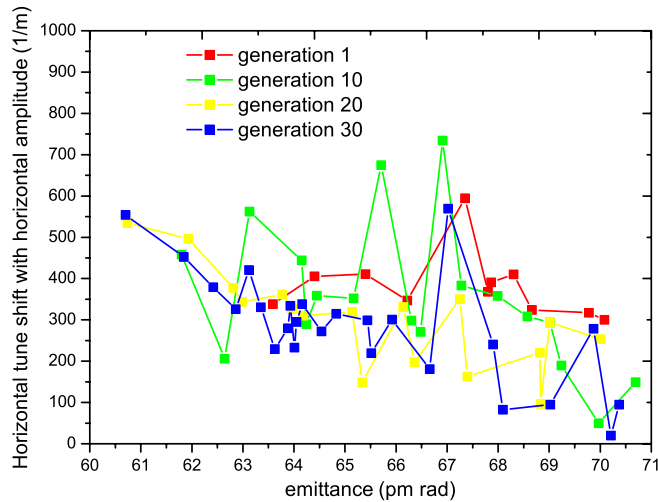


FIG. 8. The horizontal tune shift with horizontal amplitude. The different color means the optimal solutions obtained at different generation.

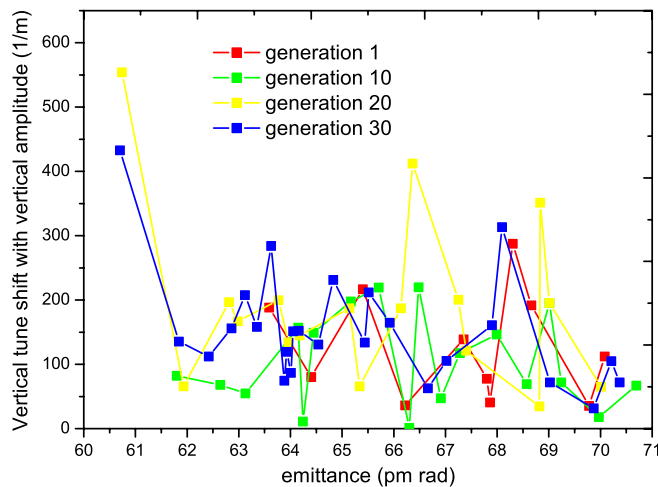


FIG. 9. The vertical tune shift with vertical amplitude. The different color means the optimal solutions obtained at different generation.

indicators cannot ensure large dynamic aperture. But we believe that adequate analytical nonlinear indicators should simplify optimization of lattice parameters, and we should make a detailed study of the relationship between dynamic aperture and nonlinear indicators in the near future.

V. CONCLUSION

An alternative way of lattice design by genetic algorithm is demonstrated. A storage ring lattice with ultralow emittance is designed to illustrate the effectiveness of the optimization method. Both the linear optics and nonlinear beam dynamics is optimized simultaneously by NSGA-II. The result shows that the approach will be applicable to the

problem of ultralow emittance lattice which has serious nonlinear beam dynamics. This method is very easy to implement. We need only propose a possible lattice, and let some optics be constraints. The code will automatically search for the low emittance lattice with better dynamic aperture. If there exists feasible solutions, the algorithm will return the solutions with the best objectives, and the global optimal solution can be obtained. If not, it will return the solutions with the lowest constraint violation. In this case, the result will guide us to relax the constraint conditions or even change the lattice structure.

Several works should be added to enhance the optimization capability of the code in the future, such as including transverse momentum aperture as objective, automatic optimization of quadrupole, and sextupole positions based on some practical limits and considerations. Of course, if the insertion device model and error effects are considered in the optimization process, the final results would be closer to the realistic storage ring.

- [1] A. Ropert, P. Elleaume, L. Farvacque, and M. Hasselweiler, in *Proceedings of the European Particle Accelerator Conference, Vienna, 2000* (EPS, Geneva, 2000).
- [2] D. Robin and B. Feinberg, in *Proceedings of the Particle Accelerator Conference, Chicago, IL, 2001* (IEEE, New York, 2001).
- [3] A. L. Robinson and A. S. Schlachter, in *Proceedings of the 1991 Particle Accelerator Conference, San Francisco, CA, 1991* (IEEE, New York, 1991).
- [4] Andreas Streun, OPA version α -170 (2009).
- [5] Maxim Korostelev and Frank Zimmermann, in *Proceedings of the 20th Particle Accelerator Conference, Portland, OR, 2003* (IEEE, New York, 2003).
- [6] Louis Emery, in *Proceedings of the 1989 Particle Accelerator Conference, Chicago, IL* (IEEE, New York, 1989).
- [7] Katsunobu Oide and Haruyo Koiso, *Phys. Rev. E* **47**, 2010 (1993).
- [8] L. Emery, in *Proceedings of the 1991 Particle Accelerator Conference, San Francisco, CA, 1991* (Ref. [3]).
- [9] Alex J. Dragt, Filippo Neri, and Govindan Rangarajan, *Annu. Rev. Nucl. Part. Sci.* **38**, 455 (1988).
- [10] M. Berz, in *Proceedings of the International Workshop on Nonlinear Problems in Accelerator Physics, Berlin* (Taylor & Francis, UK, 1993).
- [11] E. Nissen and B. Erdelyi, *Phys. Rev. ST Accel. Beams* **13**, 074001 (2010).
- [12] H. S. Dumas and J. Laskar, *Phys. Rev. Lett.* **70**, 2975 (1993).
- [13] James Ablett *et al.*, National Synchrotron Light Source II (conceptual Design Report) (2006).
- [14] John H. Holland, *Sci. Am.* **267**, 66 (1992).
- [15] Kalyanmoy Deb, *Multi-Objective Optimization using Evolutionary Algorithms* (Wiley, New York, 2001).
- [16] Kalyanmoy Deb, *IEEE Trans. Evol. Comput.* **6**, 182 (2002).

- [17] W. Hardt, J. Jäger, and D. Möhl, Report No. CERN-PS-LEA-Note-82-5, 1982.
- [18] R.H. Helm, M.J. Lee, and P.L. Morton, Report No. SLAC-PUB-1193, 1973.
- [19] Alexander Wu. Chao, *Handbook of Accelerator Physics and Engineering* (World Scientific Publishing, Singapore, 2002).
- [20] Helmut Wiedemann, *Particle Accelerator Physics* (Springer, New York, 2007).
- [21] S. Y. Lee, *Accelerator Physics* (World Scientific, Singapore, 2004).
- [22] D.P. Barber, K. Heinemann, and G. Ripken, Report No. DESY-96-156, 1996.
- [23] G. Ripken and F. Schmidt, Report No. CERN/SL/95-12 (AP), 1995.
- [24] L. Nadolski *et al.*, Proceedings of the European Particle Accelerator Conference, Vienna, 2000 (Ref. [1]).
- [25] D. Einfeld, in *Proceedings of the 11th European Particle Accelerator Conference, Genoa, 2008* (EPS-AG, Genoa, Italy, 2008).
- [26] Lingyun Yang *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **609**, 50 (2009).
- [27] L. Emery, in *Proceedings of the 21st Particle Accelerator Conference, Knoxville, 2005* (IEEE, Piscataway, NJ, 2005).
- [28] Ivan V. Bazarov and Charles K. Sinclair, *Phys. Rev. ST Accel. Beams* **8**, 034202 (2005).
- [29] Alexey A. Poklonskiy *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A* **558**, 135 (2006).
- [30] A. A. Poklonskiy and D. Neuffer, *Int. J. Mod. Phys. A* **24**, 959 (2009).
- [31] David S. Robin, Weishi Wan, and Fernando Sannibale, and Victor P. Suller, *Phys. Rev. ST Accel. Beams* **11**, 024002 (2008).
- [32] S. Y. Lee, *Phys. Rev. E* **54**, 1940 (1996).
- [33] W.W. Gao, L. Wang, and W.M. Li, in Proceedings of the First International Particle Accelerator Conference (IPAC 10), Kyoto, Japan (2010), pp. 4515–4517.
- [34] R. Tomàs, M. Giovannozzi, and R. de Maria, *Phys. Rev. ST Accel. Beams* **12**, 011002 (2009).
- [35] Hans Grote and F. Christoph Iselin, The MAD Program User's Reference Manual, Report No. CERN/SL/90-13 (AP), 1995.
- [36] Johan Bengtsson, SLS Report No. 9/97, 1997.