

# Equilibrium of an arbitrary bunch train with cavity resonators and short range wake: Enhanced iterative solution with Anderson acceleration

Robert Warnock<sup>\*</sup>

*SLAC National Accelerator Laboratory, Stanford University, Menlo Park, California 94025, USA  
and Department of Mathematics and Statistics, University of New Mexico,  
Albuquerque, New Mexico 87131, USA*



(Received 17 August 2021; accepted 4 October 2021; published 21 October 2021)

This paper continues the work of two previous treatments of bunch lengthening by a passive harmonic cavity in an electron storage ring. Such cavities, intended to reduce the effect of Touschek scattering, are a feature of fourth generation synchrotron light sources. The charge densities in the equilibrium state are given by solutions of coupled Haïssinski equations, which are nonlinear integral equations. If the only wake fields are from cavity resonators, the unknowns can be the Fourier transforms of bunch densities at the resonator frequencies. The solution scheme based on this choice of unknowns proved to be deficient at the design current when multiple resonators were included. Here we return to the conventional formulation of Haïssinski equations in coordinate space, the unknowns being charge densities at mesh points on a fine grid. This system would be awkward to solve by the Newton method used previously, because the Jacobian matrix is very large. Here a new solution is described, which is both Jacobian-free and much simpler. It is based on an elementary fixed point iteration, accelerated by Anderson's method. The scheme is notably fast and robust, accommodating even the case of extreme overstretching at current far beyond the design value. The Anderson method is promising for many problems in accelerator theory and beyond, since it is quite simple and can be used to attack all kinds of nonlinear and linear integral and differential equations. Results are presented for ALS-U with updated design parameters. The model includes harmonic and main rf cavities, compensation of beam loading of the main cavity by adjustment of the generator voltage, and a realistic short range wake field (rather than the broadband resonator wake invoked previously).

DOI: [10.1103/PhysRevAccelBeams.24.104402](https://doi.org/10.1103/PhysRevAccelBeams.24.104402)

## I. INTRODUCTION

The problem is to determine the longitudinal charge distributions of an arbitrary bunch train in an electron storage ring, in the state of equilibrium. This equilibrium could exist theoretically but be unstable in practice. Any coupling to transverse degrees of freedom is ignored. The train is arbitrary in the sense that there can be any distribution of gaps (unfilled buckets), and any distribution of bunch charge along the train.

Assuming that the collective motion is governed by coupled Vlasov-Fokker-Planck equations, the equilibrium state is determined by coupled Haïssinski equations [1]. The coupling arises from the long-range wake fields of high-Q cavity resonators. Every bunch in the train contributes to the excitation of these wakes, and thereby

influences all other bunches and even itself (by a very small amount). There is also the short range wake field (SR) from geometric aberrations in the vacuum chamber, affecting only the bunch that excites it. An achievement of the present work is to include this effect accurately, which has not been done before in a multibunch framework.

In Ref. [1], the problem was solved for the simplest model, in which the only wake field comes from a single passive higher harmonic cavity (HHC). In Ref. [2], the model was extended to include the wake field (beam loading) of the main accelerating cavity (MC), and also the SR, roughly approximated by a broadband resonator model. The rf generator voltage was adjusted by a least-squares algorithm so that the sum of the generator voltage and the induced voltage of the main cavity closely approximated the desired accelerating voltage, in amplitude and phase. Contrary to the supposition of Ref. [1], the main cavity compensated in this way played a substantial role, spoiling to some extent the desired effect of HHC. Also, the consequence of SR was not negligible.

Both Refs. [1] and [2] were based on a formulation of the Haïssinski equations in which the unknowns are the Fourier transforms of the bunch densities at the frequencies of the

<sup>\*</sup>warnock@slac.stanford.edu

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

cavity resonators. Of course, this is only possible if the entire wake field is due to resonators, narrow- or broad-band. Then, the number of real unknowns is  $n_u = 2n_b n_r$ , where  $n_b$  is the number of bunches and  $n_r$  is the number of resonators. For ALS-U, we have  $n_b = 284$  and  $n_r = 3$  for the model with HHC + MC + SR, thus  $n_u = 1704$ . Newton's method is readily feasible for solution of a system of this size or even much larger. In fact, the method worked beautifully in Ref. [1], where  $n_u = 656$ , but failed to converge for the full range of parameters desired in Ref. [2], with  $n_u = 1704$  or larger.

A possible way to avoid the divergence might be to return to the conventional formulation of the Haïssinski equations in coordinate space ( $z$ -space). The single-bunch equation in  $z$ -space, discretized on a mesh, is solved very robustly by Newton's method, even at currents far beyond realistic values [3]. Hoping for a similar success in the multibunch case, one encounters the problem of a very large Jacobian matrix. With a mesh of 100 cells and 284 bunches, the dimension of the matrix is  $28684 \times 28684$ , which is uncomfortable if not impossible on a standard personal computer. Moreover, other light source designs have more than 1000 bunches. Instead of a full Newton method, one could consider more economical quasi-Newton procedures such as Broyden's method [4,5].

Fortunately, a very simple and effective  $z$ -space solution turned up in the guise of a relaxed fixed point iteration, suggested by He, Li, Bai, and Wang [6]. This is Jacobian-free and involves little calculation beyond repeated evaluations of the potential function that appears in the exponent of the Haïssinski operator. This was successfully applied with parameters for ALS-U and other rings in Ref. [6], and I have verified the success for ALS-U. Many of the problems posed in Ref. [2] were solved in a simpler way by this method, but there were still some failures of convergence in cases of interest. Convergence of the method is slow, the number of iterations required being of order 100, but the total computation time is nevertheless modest.

This development turned my attention away from Newton-type methods and toward Jacobian-free fixed point iterations. There is a long history of efforts to accelerate iterative sequences [7]. One of particular interest is Anderson's proposal of 1965 [8–10]. It has the potential both to cure divergence and to promote fast convergence. Remarkably, it does both in our problem, providing a very fast and robust solution throughout the parameter domain of interest.

Section II describes the relaxed fixed point iteration. The discussion leads naturally to the continuation method, which is a more standard approach to nonlinear equations and a technique that can be related to Anderson acceleration. Section III introduces Anderson acceleration. Section IV presents results for ALS-U, with parameters from the latest design report, somewhat different from those

of our previous papers. Section V treats the relation of Anderson acceleration to Broyden's quasi-Newton method. Section VI presents conclusions and the outlook for future work.

## II. RELAXED FIXED POINT ITERATION AND THE CONTINUATION METHOD

We wish to solve  $n$  equations in  $n$  unknowns, written compactly as

$$x = g(x), \quad g: R^n \rightarrow R^n, \quad (1)$$

where  $g$  may be linear but is nonlinear in general. To relate to later discussions, we suppose that  $g$  has continuous first derivatives, as is true in our examples, although no derivatives appear in the numerical work. The function  $g$  will be called the basic map. The elementary fixed point iteration, or method of successive substitutions, tries to construct a solution by starting with some guess  $x_0$  and forming a sequence  $\{x_k\}$  as

$$x_{k+1} = g(x_k), \quad k = 0, 1, \dots, \quad (2)$$

hoping that the sequence will converge to a solution  $x$ . If  $g$  maps a ball  $B_r = \{x \mid \|x\| < r\}$  into itself, and reduces the distance between any two points in the ball,

$$\|g(x_1) - g(x_2)\| < \beta \|x_1 - x_2\|, \quad \text{all } x_1, x_2 \in B_r, \quad 0 < \beta < 1, \quad (3)$$

then the contraction mapping theorem ensures that the sequence converges to a solution, the only solution in  $B_r$ , for any  $x_0 \in B_r$ . Here  $\|\cdot\|$  can be any norm but for analytic estimates of  $\beta$ , a convenient choice is the maximum absolute value of components of  $x$ :  $\|x\| = \max_i |x^i|$ .

The discretized Haïssinski system has the contractive property for sufficiently small beam current, so we have the assurance of a unique solution at low current. Numerical calculations show that the sequence diverges for the larger currents of interest, so we seek a better algorithm.

We look for a better map  $h(x)$ , which ought to generate a sequence that will converge, at least for some appropriate  $x_0$ . The relaxed or damped fixed point iteration makes  $h(x)$  from  $g(x)$  in the simplest imaginable way,

$$h(x) = \alpha g(x) + (1 - \alpha)x, \quad 0 < \alpha < 1. \quad (4)$$

That is, if  $g$  produces too much change in  $x$ , reduce its contribution and use the current  $x$  itself for the rest of the iterate. It could be said that damping rather than relaxation is more descriptive of the process.

He *et al.* in Ref. [6] adopted this procedure to solve the coupled Haïssinski equations in the  $z$ -space formulation and found it to be remarkably effective. They called it a

“relaxation iteration.” The damping parameter  $\alpha$  was chosen by experiment. At high values of current a relatively small value is required for convergence, say  $\alpha = 0.1$ .

It looks as though He *et al.* generalized freely from the linear case, since their bibliography on the source of the method refers only to that case. When  $g$  is linear, relaxation is widely used in the SOR algorithm, successive over-relaxation, a variant of the Gauss-Seidel method intended to accelerate convergence. I have not seen much notice of the nonlinear application in the literature of numerical analysis, although one can find it in the context of particular problems. See for instance Sec. III of Ref. [11] where it is called “simple mixing”.

A bit more insight into Eq. (4) accrues if we invoke a differential equation. Define  $f(x) = g(x) - x$  and consider the system of ordinary differential equations,

$$\frac{dx}{dt} = f(x), \quad x(0) = x_0. \quad (5)$$

One might hope to find a solution of  $f(x) = 0$  as the limit of an asymptotically constant trajectory  $x(t)$  as  $t \rightarrow \infty$ . Euler’s method applied to Eq. (5) gives a sequence  $\{x_k\}$  defined by

$$\frac{x_{k+1} - x_k}{\Delta t} = f(x_k), \quad x_{k+1} = \Delta t g(x_k) + (1 - \Delta t)x_k. \quad (6)$$

Thus an attempt to find a constant asymptote by Euler’s method is the same as trying to find a solution by the relaxed fixed point iteration, with  $\alpha = \Delta t$ .

There is no proof that a solution of  $f(x) = 0$  is really to be found as the constant asymptote of a solution of Eq. (5). A more certain relation to a differential equation, explored in the literature, is obtained by considering a homotopy connecting an equation with known solution  $x_0$  to the equation of interest [12], § 7.5. Suppose that  $H: R^n \times R^1 \rightarrow R^n$  is a smooth function of both variables such that

$$H(x_0, t_0) = 0, \quad H(x, t_1) = f(x). \quad (7)$$

Then consider the trajectory  $x(t)$  defined by  $H(x(t), t) = 0$ . Differentiating we find

$$H_x(x, t) \frac{dx}{dt} + H_t(x, t) = 0. \quad (8)$$

As long as the inverse of the Jacobian  $H_x$  exists, we have the differential equation in explicit form,

$$\frac{dx}{dt} = -H_x^{-1}(x, t)H_t(x, t), \quad x(t_0) = x_0. \quad (9)$$

If this equation has a solution extending from  $t_0$  to  $t_1$ , we have achieved a solution of  $f(x) = 0$  as  $x = x(t_1)$ , according to Eq. (7). A procedure along these lines is called a continuation method.

There are of course myriad ways to choose  $H$ . An especially useful choice is

$$H(x, t) = f(x) - e^{-t}f(x_0), \quad t_0 = 0, \quad t_1 = \infty, \\ H_x(x, t) = f_x(x), \quad \text{and} \quad H_t(x, t) = e^{-t}f(x_0). \quad (10)$$

Since on the trajectory  $e^{-t}f(x_0) = f(x)$ , the Eq. (9) becomes

$$\frac{dx}{dt} = -f_x^{-1}(x)f(x). \quad (11)$$

When Euler’s method is applied to Eq. (11), we get

$$x_{k+1} = x_k - \Delta t f_x^{-1}(x_k)f(x_k), \quad (12)$$

which is the damped Newton method with damping factor  $\Delta t$ . It becomes the full Newton method for  $\Delta t = 1$ . If  $F(x) = -f_x(x)^{-1}f(x)$  obeys a Lipschitz condition, the differential Eq. (11) is subject to standard existence theorems. Boggs [13] has explored the use of more sophisticated integrators of Eq. (11), with the goal of approaching the asymptote more quickly.

Broyden’s method gives a way to approximate  $f_x^{-1}(x_k)$ , starting with a value for  $f_x^{-1}(x_0)$  [5]. The update from step  $k$  to step  $k + 1$  is obtained by adding a rank-1 matrix. With the definition  $G_k \approx f_x^{-1}(x_k)$ , the update takes the form

$$G_{k+1} = G_k + (\Delta x_k - G_k \Delta f_k) \frac{\Delta f_k^T}{\Delta f_k^T \Delta f_k}, \\ \Delta v_k = v_{k+1} - v_k, \quad (13)$$

where the row vector  $v^T$  is the transpose of a column vector  $v$ . This is called Broyden’s second method. His first method approximates the Jacobian itself in a similar way.

Sometimes it is adequate to take  $f_x(x_0) = -I$  where  $I$  is the unit matrix, which is equivalent to using the relaxed iteration of Eq. (6) for the first step. With this reasonable choice and Eq. (13), one can carry out an approximate version of the Newton iteration (12), often to good effect. The undamped iteration would be preferred but damping could be needed for convergence.

### III. ANDERSON ACCELERATION

Again we wish to solve  $x = g(x)$ . Step  $k$  of Anderson’s iteration makes use of the current and previous evaluations of the map,  $g(x_j)$ ,  $j = k, k - 1, \dots$ . These evaluations contain valuable information. The update  $x_{k+1}$  is formed from a favorable linear combination of  $g(x_j)$ .

With a given start  $x_0$ , we employ the following notations for  $k = 0, 1, \dots$ :

$$g_k = g(x_k), \quad f_k = g_k - x_k, \\ \|u\|^2 = \sum_{i=1}^n (u^i)^2, \quad u = (u^1, \dots, u^n). \quad (14)$$

Also choose an integer  $m \geq 1$  and define

$$m_k = \min(k, m), \quad (15)$$

which will be the number of previous map evaluations used at the  $k$ -th step, not more than  $m$ .

To find a good linear combination of the  $g_j$ , Anderson finds the coefficients in a minimal linear combination of the  $f_j$ . That is, he solves the constrained linear least-squares problem

$$(\alpha_0^k, \alpha_1^k, \dots, \alpha_{m_k}^k) = \arg \min \left\| \sum_{j=0}^{m_k} \alpha_j^k f_{k-m_k+j} \right\|^2, \\ \sum_{j=0}^{m_k} \alpha_j^k = 1. \quad (16)$$

Then, the next iterate is taken to be

$$x_{k+1} = \sum_{j=0}^{m_k} \alpha_j^k g_{k-m_k+j}. \quad (17)$$

For  $k = 0$ , the constraint alone determines the minimum, so that  $\alpha_0^0 = 1$  and  $x_1 = g(x_0)$ .

Anderson allowed extra flexibility by introducing a relaxation parameter  $\beta_k$ , with a corresponding update

$$x_{k+1} = \beta_k \sum_{j=0}^{m_k} \alpha_j^k g_{k-m_k+j} + (1 - \beta_k) \sum_{j=0}^{m_k} \alpha_j^k x_{k-m_k+j}. \quad (18)$$

In this scheme  $x_1 = \beta_0 g(x_0) + (1 - \beta_0)x_0$ , which is to say that the iteration starts with simple mixing. In view of the partial success of simple mixing, this would seem to be a good choice, at least for the first step. At later iterations one might put  $\beta_k = 1$ .

It is convenient, both for the calculation and for some steps in analysis, to recast the minimization problem without constraints. That is accomplished merely by a linear change of variables; see Ref. [9], Eq. (3.1) and following remarks. Define new constants  $\gamma_i^k$  such that

$$\alpha_0^k = \gamma_0^k, \quad \alpha_j^k = \gamma_j^k - \gamma_{j-1}^k, \\ 1 \leq j \leq m_k - 1, \quad \alpha_{m_k}^k = 1 - \gamma_{m_k-1}^k. \quad (19)$$

Now the sum of the  $\alpha_j^k$  is 1 for any choice of the  $\gamma_j^k$ , and the unconstrained minimization takes the form

$$(\gamma_0^k, \gamma_1^k, \dots, \gamma_{m_k}^k) = \arg \min \|f_k \\ + \sum_{j=0}^{m_k-1} \gamma_j^k (f_{k-m_k+j} - f_{k-m_k+j+1})\|^2. \quad (20)$$

Correspondingly, the next iterate is

$$x_{k+1} = g_k + \sum_{j=0}^{m_k-1} \gamma_j^k (g_{k-m_k+j} - g_{k-m_k+j+1}). \quad (21)$$

Of course, with relaxation this becomes

$$x_{k+1} = \beta_k \left( g_k + \sum_{j=0}^{m_k-1} \gamma_j^k (g_{k-m_k+j} - g_{k-m_k+j+1}) \right) \\ + (1 - \beta_k) \left( x_k + \sum_{j=0}^{m_k-1} \gamma_j^k (x_{k-m_k+j} - x_{k-m_k+j+1}) \right). \quad (22)$$

In the following work  $\beta_k = 1$  is a satisfactory choice, whereas in more difficult cases under investigation a value as small as 0.025 is required for convergence.

The relation of Anderson's method to Broyden's algorithm is discussed in Sec. V.

#### IV. RESULTS WITH ALS-U PARAMETERS

Parameters considered in the preliminary design report for the Advanced Light Source Upgrade (ALS-U) [14] of

TABLE I. Parameters from preliminary design report for ALS-U.

Ring circumference	$C$	196.5 m
Beam energy	$E_0$	2 GeV
Average bunch current	$I_{\text{avg}}$	500 mA
Momentum compaction	$\alpha$	$2.025 \times 10^{-4}$
Natural energy spread	$\sigma_\delta$	$1.02 \times 10^{-3}$
Natural rms bunch length	$\sigma_{z0}$	3.9 mm
Energy loss per turn (with insertion devices)	$U_0$	315–330 keV
Harmonic number	$h$	328
Main cavity frequency	$f_1$	500.390 MHz
Main cavity voltage	$V_1$	600 kV
Harmonic cavity harmonic number	3	
Harmonic cavity shunt impedance	$R_s$ (high $R/Q$ )	1.9 M $\Omega$
Harmonic cavity quality factor	$Q$ (high $R/Q$ )	$2.4 \times 10^4$
Harmonic cavity detuning	$f_r - 3f_1$ (high $R/Q$ )	317.80 kHz
Harmonic cavity shunt impedance	$R_s$ (low $R/Q$ )	1.4 M $\Omega$
Harmonic cavity quality factor	$Q$ (low $R/Q$ )	$3.4 \times 10^4$
Harmonic cavity detuning	$f_r - 3f_1$ (low $R/Q$ )	164.74 kHz
Main cavity shunt impedance (sum of two)	$R_s$ (unloaded)	9.8 M $\Omega$
Main cavity quality factor	$Q$ (unloaded)	$3.6 \times 10^4$
Main cavity detuning	$f_r - f_1$	-94.729 kHz
Main cavity coupling parameter	$\beta$ (optimum)	9.983
Main cavity coupling parameter	$\beta$ (ALS heritage)	3.1



October 2020 are listed in Table I. These parameters differ considerably from those adopted in Refs. [1,2], and [6], so part of the motivation for this report is to bring the study up to date. Two choices for the harmonic cavity parameters are contemplated, called the high- $R/Q$  and low- $R/Q$  options, which are alleged to have different implications for stability issues. The stability is of course important but does not concern us here.

At last notice the coupling coefficient  $\beta$  for the main rf cavity was still an undetermined feature of the design. The present ALS cavities might be used if their coupling could be increased enough to control the dc Robinson instability. For consistency with the calculations of Ref. [14], we take the “optimum” value of Table 1,  $\beta = 9.983$ . Appropriate values of impedance and quality factor for the calculation are the loaded values,  $R_{sL} = R_s/(1 + \beta)$  and

$Q_L = Q/(1 + \beta)$ . We use the main cavity detuning from the table, which realizes the “compensated condition” given by Eqs. (3.79) and (3.80) in Ref. [14].

The table in Ref. [14] gives  $U_0 = 330$  keV with insertion devices, but the reported calculations to be compared to ours have  $U_0 = 315$  keV, so we choose the latter.

The object is to solve the discretized coupled H $\ddot{a}$ issinski system, written compactly as

$$f(\rho) = 0, \quad f: R^n \rightarrow R^n. \quad (23)$$

Supposing that the mesh for discretization of each bunch density has  $n_m$  points, the vector  $\rho$  with  $n = n_m n_b$  components consists of  $n_b$  densities evaluated at the mesh points:

$$\rho = [\rho_1(z_1), \rho_1(z_2), \dots, \rho_1(z_{n_m}), \dots, \rho_{n_b}(z_1), \rho_{n_b}(z_2), \dots, \rho_{n_b}(z_{n_m})]. \quad (24)$$

Similarly,

$$f = [f_1(z_1), f_1(z_2), \dots, f_1(z_{n_m}), \dots, f_{n_b}(z_1), f_{n_b}(z_2), \dots, f_{n_b}(z_{n_m})], \quad (25)$$

where

$$f_i(z_j) = \frac{1}{A_i} \exp(-\mu U_i(z_j, \rho)) - \rho_i(z_j). \quad (26)$$

The denominator  $A_i$  is a normalization factor, the discretized integral of the numerator. In the exponent,  $U_i(z_j, \rho)$  is the potential well seen by the  $i$ th bunch and  $\mu$  is the constant of Eq. (49) in Ref. [1]. In the following,  $U_i$  consists of the expression defined in Eqs. (51), (57), and (60) of Ref. [1], plus the integral of the short range wake potential convolved with  $\rho_i$ , as follows:

$$U_i^{sr}(z) = \frac{1}{f_1} \int_{-\Sigma}^{\Sigma} S(z - \zeta) \rho_i(\zeta) d\zeta, \quad S(z) = \int_{\zeta_0}^z W(\zeta) d\zeta. \quad (27)$$

The wake potential  $W(z)$ , from detailed modeling by Dan Wang [15], is plotted in Fig. 13 of Ref. [2]. It is zero for  $z < \zeta_0$ , where  $\zeta_0$  is a small fraction of the bunch length, arising from the nonzero length of the drive bunch in the wake field simulation.

The solution vector  $x$  of the previous section is identified with  $\rho$  and the map vector  $g$  is from the first term in Eq. (26); that is  $g = f + \rho$ .

The least-squares step in the Anderson algorithm is done with the code “dgers” from the Intel Math Kernel Library. This solves the normal equation using the QR decomposition which is recommended in Ref. [9]. We take  $m_k = \min(k, \infty) = k$  so that at the  $k$ th iterate, the current

evaluation and all previous evaluations of the map  $g$  are employed. About the same results are obtained with a sufficiently large limit on the number used, say with  $m_k = \min(k, 8)$ , but this is bothersome to verify and gives no appreciable saving in computation time. The time for the least-squares step is negligible.

All results and CPU times are for a serial code in Fortran, running on a laptop. The code is arranged so that the result of any run can be taken as an initial guess for the next run. A result for a complete fill can then be used to initiate a run with a partial fill, or with a smaller detuning, or with a new wake component included, or with smaller error tolerances, etc.

In contrast to the algorithms used in Refs. [1] and [2], no continuation in current from small initial values is needed to achieve convergence. In spite of strong nonlinearities convergence is found immediately at the design current and even at much higher values.

### A. The case of high $R/Q$

The first step is to consider the complete fill, with all 328 buckets filled with the same charge and the entire wake coming from the HHC. Then every bunch comes out to have the same profile, even though that is not put in as a constraint. In Fig. 1, we show that profile for a decreasing sequence of detunings. The legend gives the detuning  $\delta f$  in kHz and  $\sigma/\sigma_0$ , the ratio of the rms bunch length to the natural bunch length of 3.9 mm. This plot agrees with Fig. (3.255) in Ref. [14]. With 201 mesh points per bunch, the CPU time is 6 s for each curve.

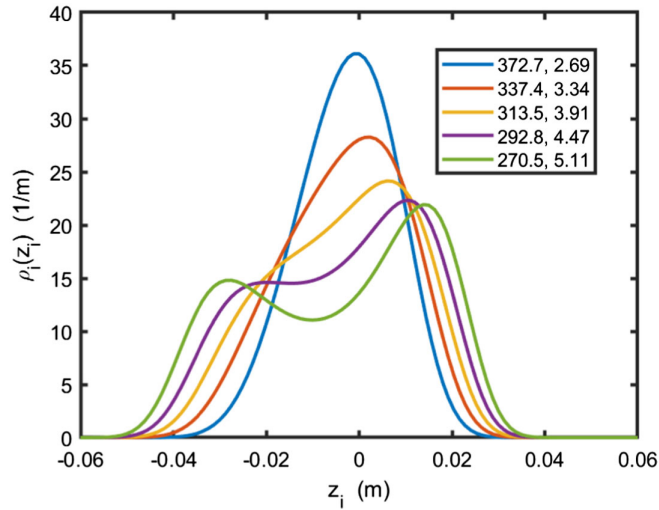


FIG. 1. Complete fill, HHC only, high  $R/Q$ . The legend gives the detuning  $\delta f = f_r - 3f_1$  in kHz, and the ratio  $\sigma/\sigma_0$  of rms bunch length to the natural bunch length.

The next step, again for a complete fill, is to see the effect of the main cavity beam loading, which is to be compensated by adjustment of the generator voltage. As expected, the compensation is essentially perfect and the bunch profile is the same to graphical accuracy. It is given by the blue curve in Fig. 2 for the nominal detuning of 317.8 kHz from Table I. Here, the increase in bunch length is  $\sigma/\sigma_0 = 3.79$ .

The compensation algorithm of Ref. [2] did not converge with the desired energy loss of  $U_0 = 315$  keV. Noticing that, it did converge in the work of Ref. [1] which had a smaller value of  $U_0$ , we reduced the value and then increased it in steps:  $U_0 = 260, 280, 300$ , and 315 keV. This procedure took 5 min. All subsequent calculations

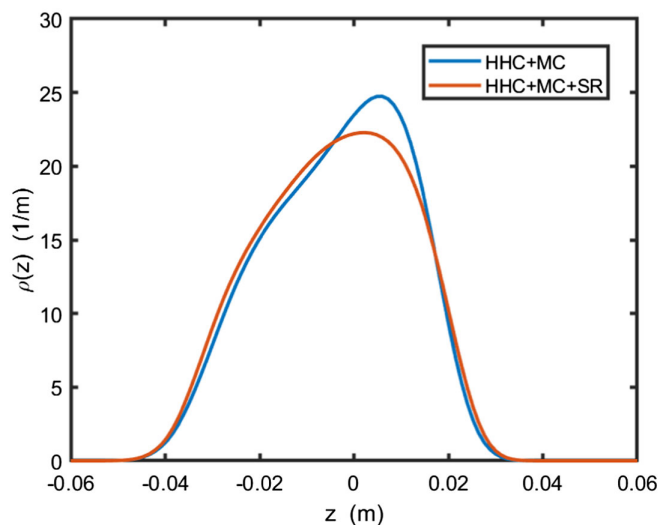


FIG. 2. Complete fill, high  $R/Q$ ,  $\delta f = 317.8$  kHz,  $\sigma/\sigma_0 = 3.80$  (blue), 3.92 (red).

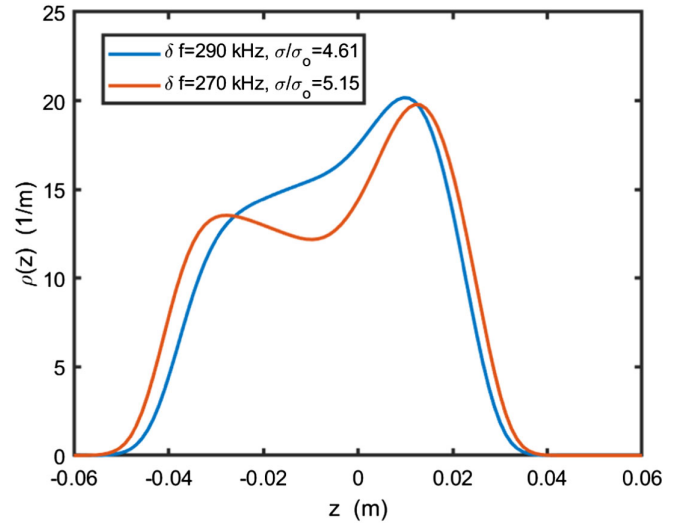


FIG. 3. Complete fill, high  $R/Q$ , with HHC + MC + SR. Two lower detunings.

were started with the result of a previous calculation and required less than one minute of additional time each.

The compensation algorithm is the standard Gauss-Newton method for nonlinear least squares, although it was not recognized as such in Ref. [2]. There are other algorithms, such as the Levenberg-Marquardt method, which can be more robust concerning the starting guess. Perhaps such a method could give a least squares solution directly for the desired  $U_0$  but not necessarily in a shorter time.

Turning on the short range wake we get the red curve in Fig. 2. The short range force reduces the asymmetry of the bunch and increases its rms length by 3%. This is different from the effect of the short range wake in the broadband resonator model [2] and perhaps more reasonable.

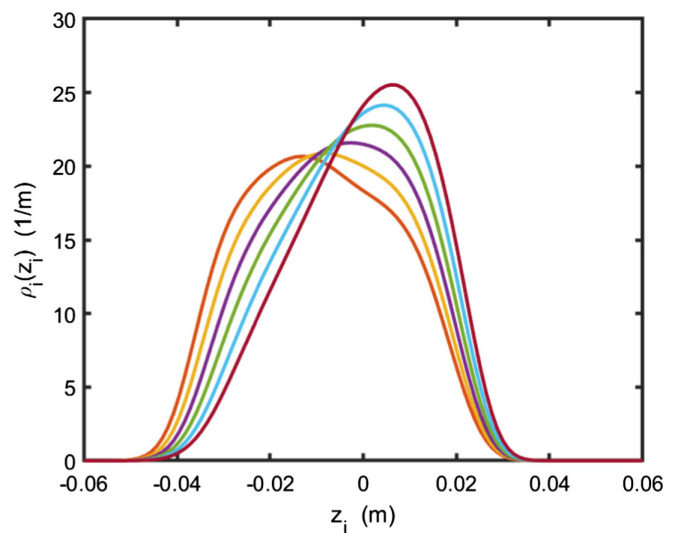


FIG. 4. Fill with distributed gaps, high  $R/Q$ , HHC + MC + SR,  $\delta f = 317.8$  kHz.

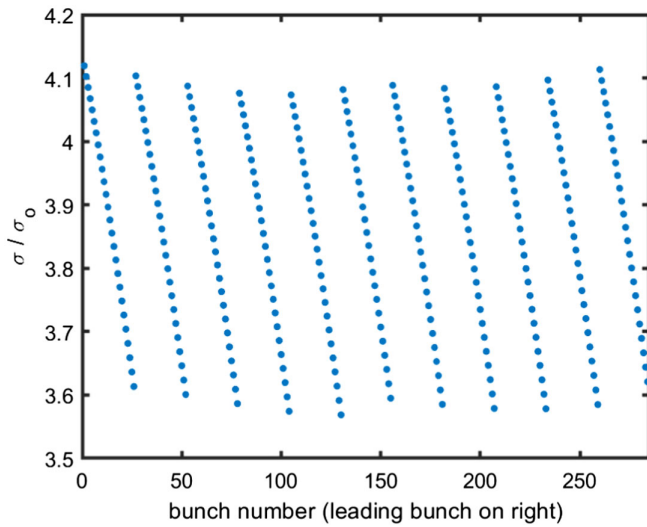


FIG. 5. Case of Fig. 4, ratio of bunch length to natural length, vs bunch number.

Additional bunch lengthening through a decrease in detuning is a possibility for the machine, discussed in Ref. [14]. The results of two smaller values are shown in Fig. 3. The transition to overstretching, when two peaks appear, occurs between the two.

The partial fill anticipated for the machine, which has been called Fill C2 in Ref. [16], has distributed gaps of four buckets each, with a total of 284 bunches. There are 11 subtrains, 9 with 27 bunches and 2 with 26, the latter two on opposite sides of the ring. All bunches have the same charge chosen to give the desired average current of 500 mA. Taking this case with the nominal detuning  $\delta f = 317.8$  kHz of Table 1, and including HHC, MC, and SR, we get the densities shown in Fig. 4. There are six bunches in the plot out of a typical subtrain of 27 bunches. The one

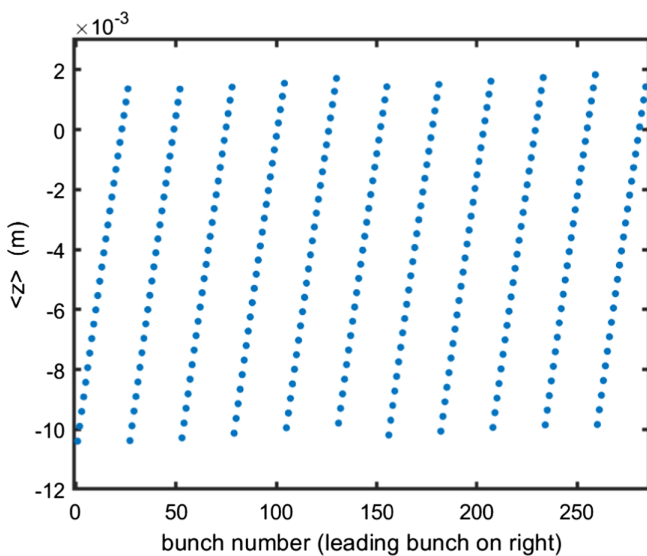


FIG. 6. Case of Fig. 4, bunch centroid vs bunch number.

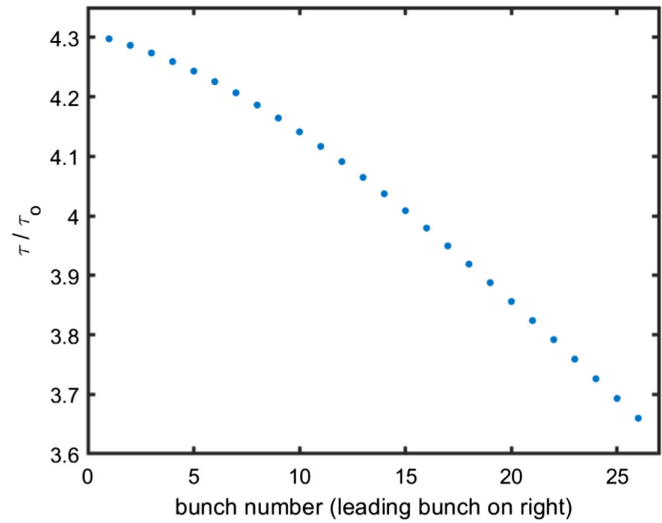


FIG. 7. Case of Fig. 4, increase of Touschek lifetime.

with maximum farthest to the right is nearest the front of the subtrain.

As was discovered in Ref. [2], the main cavity has a large influence when there are gaps in the train. The bunches near the front of the subtrain resemble that of the complete fill, whereas those at the middle and back are broader and flatter. The distributions of bunch length increase and centroid displacement along the full train are shown in Figs. 5 and 6. The bunch length increase and the centroid displacement are both largest at the back of a subtrain.

The most interesting figure of merit is the increase in the Touschek lifetime over the case without a harmonic cavity. This is plotted for a typical subtrain in Fig. 7. The factor of increase,  $\tau/\tau_0$ , is not far from the length increase  $\sigma/\sigma_0$ . The strong variation along the subtrain should be an issue

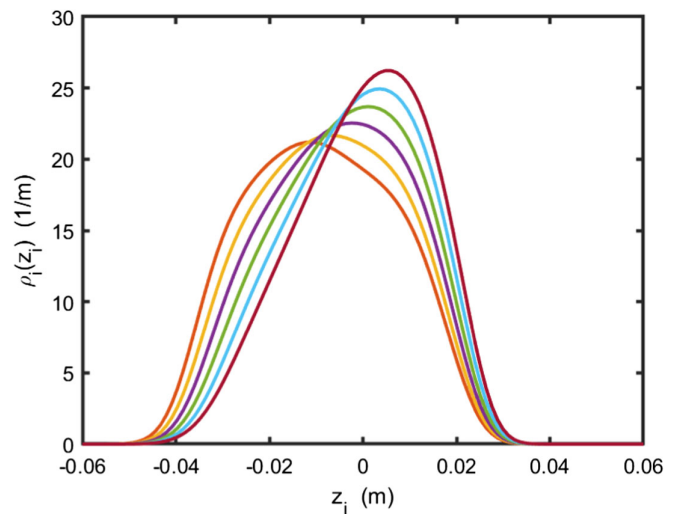


FIG. 8. Fill with distributed gaps, HHC + MC + SR,  $\delta f = 330$  kHz. This result resembles Fig. (3.256) in Ref. [14].

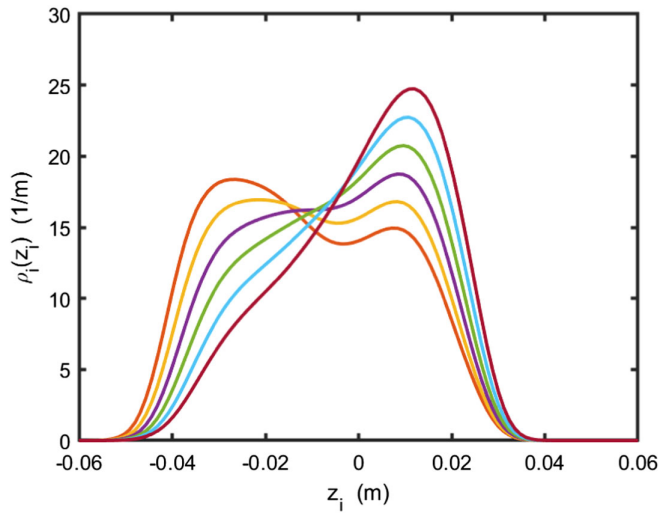


FIG. 9. Fill with distributed gaps, HHC + MC + SR,  $\delta f = 290$  kHz.

in determining the average beam lifetime, but that matter is beyond the scope of this work.

Fig. 4 is to be compared with Fig. (3.256) in Ref. [14] generated from a macroparticle simulation with rather severe noise. This plot is for a case different from ours in that it does not include the short range wake and most probably has a different account of the main cavity beam loading, which is mentioned but not described. Also, the detuning is not specified exactly but is said to be “right below the onset of the overstretching instability”. Our Fig. 4 has more broadening especially at the back of the train. For whatever reason, to get a result with a fair resemblance to Fig. (3.256), we have to increase the detuning to 330 kHz, getting the result in Fig. 8. Here, the results for bunch lengthening, centroid position, and lifetime increase are consistent with

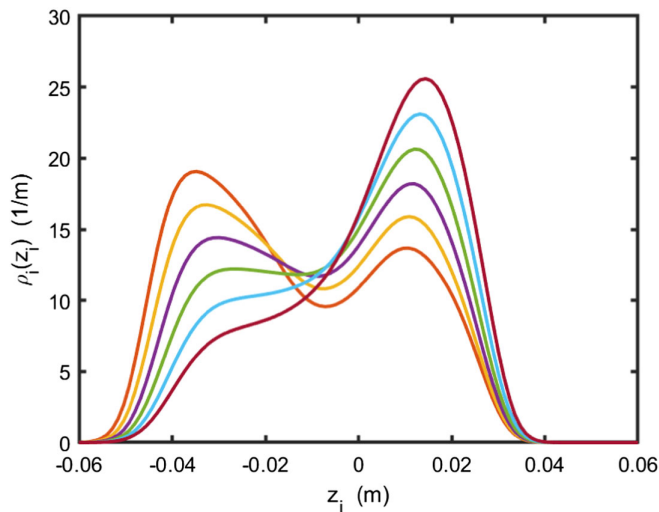


FIG. 10. Fill with distributed gaps, HHC + MC + SR,  $\delta f = 270$  kHz.

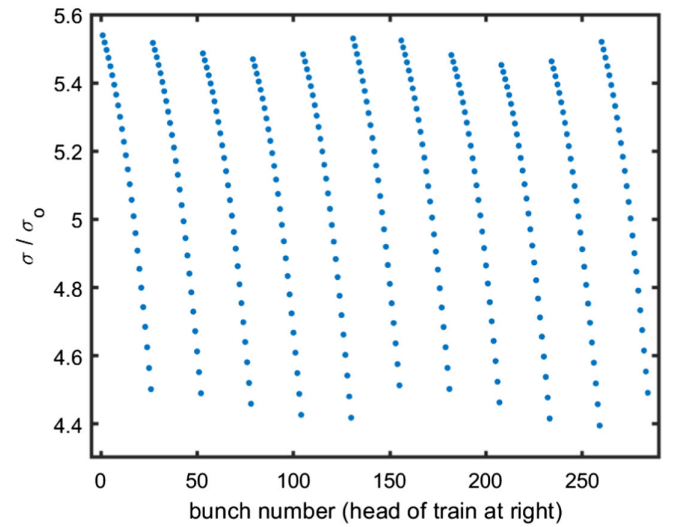


FIG. 11. Fill with distributed gaps, HHC + MC + SR,  $\delta f = 290$  kHz.

the results in Fig. (3.256) of Ref. [14] but a bit more favorable.

### B. High $R/Q$ with overstretching

What is the effect of overstretching with the partial fill, as compared to the result of Fig. 4 for the complete fill? It turns out that the threshold for overstretching is at larger  $\delta f$  than Fig. 3 would indicate for bunches at the back of a subtrain, but similar to Fig. 3 for bunches at the front. Figs. 9 and 10 show the densities for the same detunings as in Fig. 4. Figs. 11 and 12 show the corresponding bunch length increases.

This shows that there is no profit in overstretching beyond a certain point. Fig. 10 displays more undesirable

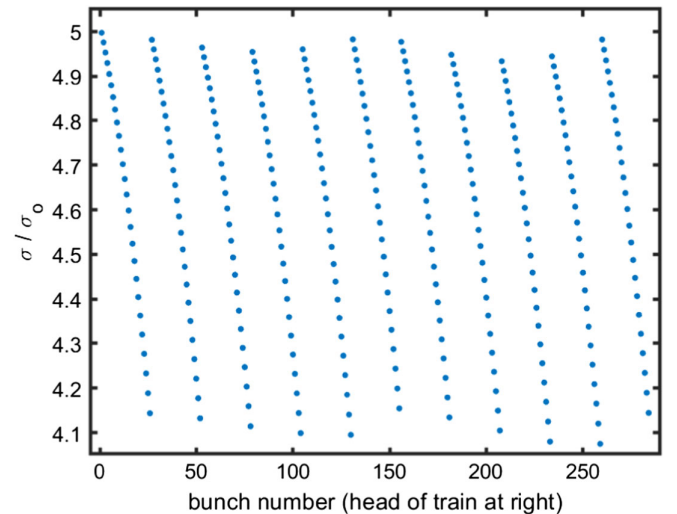


FIG. 12. Fill with distributed gaps, HHC + MC + SR,  $\delta f = 270$  kHz.



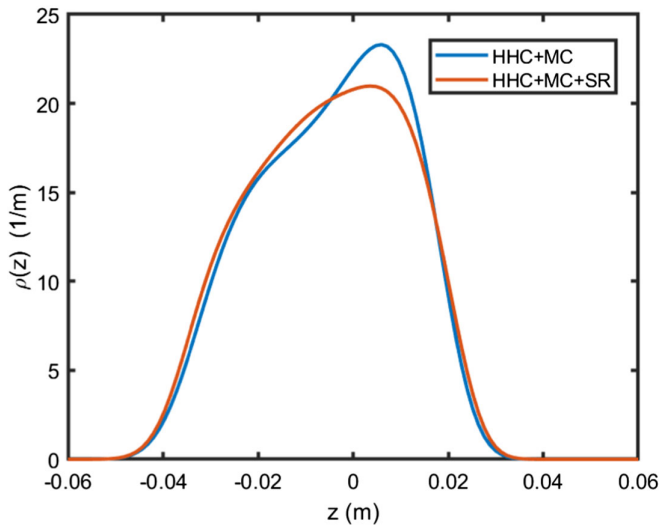


FIG. 13. Complete fill, low  $R/Q$ , HHC + MC in the blue curve, HHC + MC + SR in the red curve. Detuning 164.7 kHz from Table 1.  $\sigma/\sigma_0 = 3.96$  (blue), 4.07 (red).

bunch distortion than Fig. 9 without much increase in the average bunch length.

It took only 40 s to produce Fig. 9 starting with the solution of Fig. 4, and another 40 s to make Fig. 10 starting with Fig. 9. Through Anderson acceleration, we have gained a remarkable advance in technique compared to Ref. [2] in which these solutions could not be produced at all.

### C. The case of low $R/Q$

For low  $R/Q$ , the results for a complete fill are given in Fig. 13. Passing from this solution to the case of a partial fill, we find the pattern of Fig. 14.

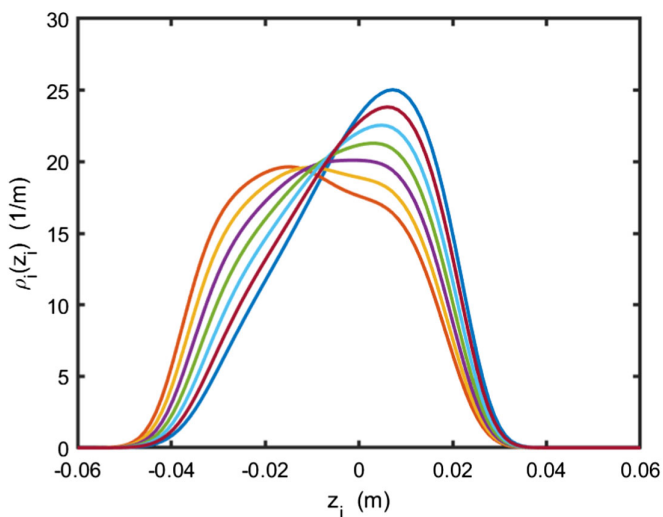


FIG. 14. Fill with distributed gaps, low  $R/Q$ , HHC + MC + SR,  $\delta f = 164.7$  kHz.

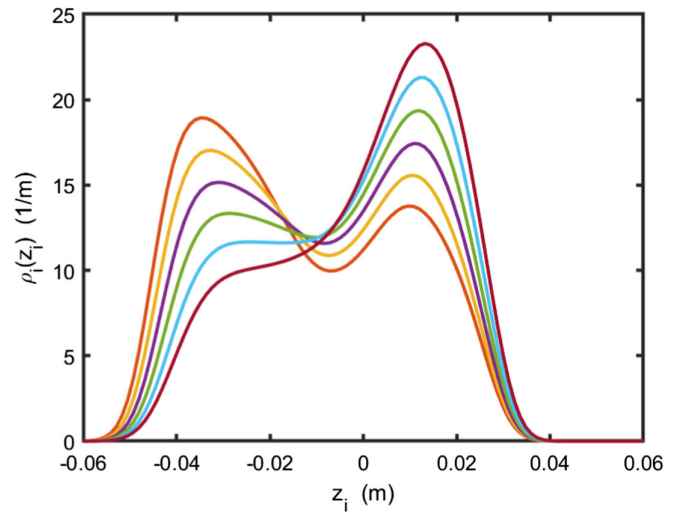


FIG. 15. Fill with distributed gaps, low  $R/Q$ , HHC + MC + SR,  $\delta f = 140$  kHz.

The preliminary design report [14] expresses an interest in running this case with 20% overstretching which is illustrated with a macroparticle simulation in Fig. (3.257). We obtain the closely similar result of Fig. 15 with a detuning of 140 kHz, reduced from 164.7 kHz. The corresponding outcomes for bunch lengthening, centroid distribution, and Touschek lifetime increase are plotted in Figs. 16, 17, and 18.

### V. THE RELATION OF ANDERSON ACCELERATION TO BROYDEN'S METHOD

By an argument of Eyert [17], revisited by Fang and Saad [10], Anderson's method is equivalent to a generalized form of Broyden's second method for updating an approximation to the inverse Jacobian as in Eq. (13). The usual Broyden method imposes the secant condition,

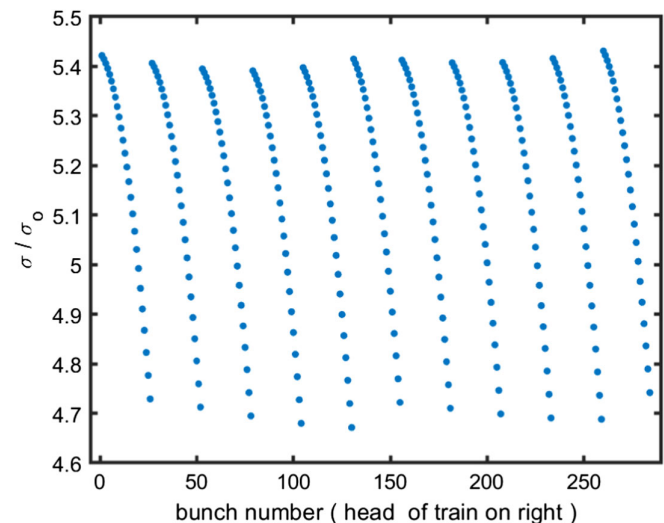


FIG. 16. Case of Fig. 14, ratio of bunch length to natural length vs bunch number.

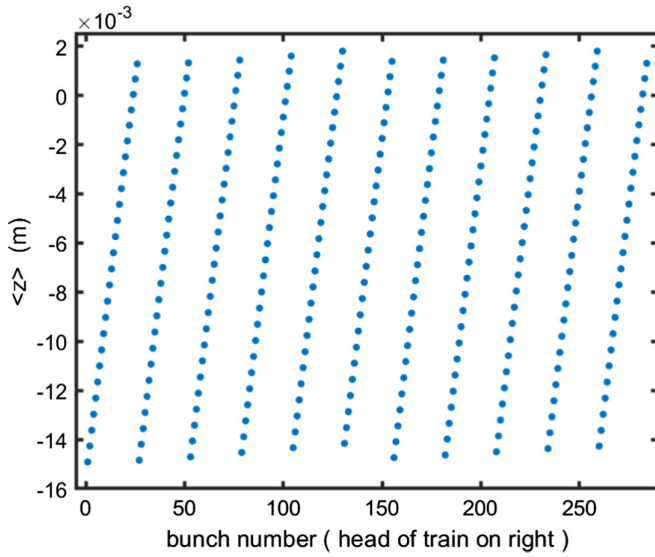


FIG. 17. Case of Fig. 14, bunch centroid vs bunch number.

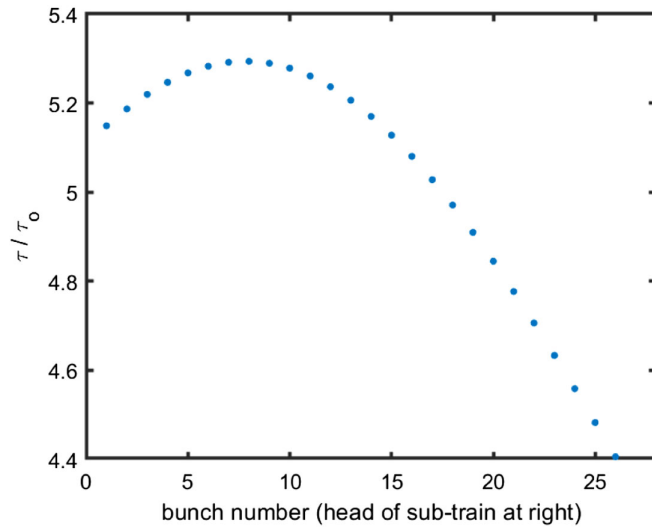


FIG. 18. Case of Fig. 14, increase of Touschek lifetime.

$$G_{k+1} \Delta f_k = \Delta x_k, \quad (28)$$

which is motivated by the linear approximation to  $f$ . A second condition requires that the Frobenius squared norm  $\|G_{k+1} - G_k\|_F^2$  be minimum with respect to  $G_{k+1}$ , subject to condition (28). These two conditions lead uniquely to Eq. (13).

A generalized Broyden method imposes multiseccant conditions,

$$G_k \Delta f_p = \Delta x_p, \quad p = k - m_k, \dots, k - 1, \quad (29)$$

thus taking account of previous iterates in the spirit of Anderson. Again imposing a minimum change of  $G_k$ , and

taking  $G_0 = -I$ , the updates  $x_{k+1} = x_k - G_k f_k$  are found to coincide with Anderson's. The matrix  $\mathcal{F}_k$  having the vectors  $\Delta f_p$  as columns is assumed to have full rank.

## VI. CONCLUSIONS AND OUTLOOK

I have described a new method to compute the equilibrium charge densities of an arbitrary bunch train with gaps, under the influence of an HHC, MC, and a realistic SR. It succeeds under more difficult conditions than previous methods, in fact for all conditions that arise in examples studied to date. Realized by a serial code on a laptop computer, the method takes only a few minutes for a thorough survey of the parameter space.

As an example, the parameter set for ALS-U in the preliminary design report was adopted. Results similar to those obtained by macroparticle simulations in the report could be obtained with a reasonable choice of detuning parameters. The report does not specify detuning exactly, and the physical model is different in not including the short range wake, and may have a different treatment of the MC beam loading.

This paper introduces the Anderson iterative method that is probably new to accelerator physics, and which seems very promising for further applications in the field. It is especially interesting for problems falling under rubrics such as “self consistency” or “phase space matching”, often formulated in terms of nonlinear integral or differential equations.

There are large-scale applications of Anderson's method and related ideas in the literature of *ab initio* quantum mechanical calculations of material and molecular properties [10,11,17–19]. These are based on the Kohn-Sham density functional formalism [20], which is similar in spirit if not in specifics to systems arising from the nonlinear Vlasov equation. It may be profitable to keep an eye on this work to see if there are any lessons to be learned for accelerator physics. Also, ongoing efforts by numerical analysts are interesting, specifically for Anderson acceleration [21].

## ACKNOWLEDGMENTS

I learned of Anderson acceleration through an inspiring colloquium for the University of New Mexico by Professor Hans De Sterck, University of Waterloo. I thank Dr. Dan Wang for a copy of her wake potential for ALS-U. This work was supported in part by the U.S. Department of Energy, Contract No. DE-AC03-76SF00515.

- [1] R. Warnock and M. Venturini, Equilibrium of an arbitrary bunch train in presence of a passive harmonic cavity: Solution through coupled Haïssinski equations, *Phys. Rev. Accel. Beams* **23**, 064403 (2020).

- [2] R. Warnock, Equilibrium of an arbitrary bunch train in the presence of multiple resonator wake fields, *Phys. Rev. Accel. Beams* **24**, 024401 (2021).
- [3] R. Warnock and K. Bane, Numerical solution of the Haïssinski equation for the equilibrium state of a stored electron beam, *Phys. Rev. Accel. Beams* **21**, 124401 (2018).
- [4] C. G. Broyden, A class of methods for solving nonlinear simultaneous equations, *Math. Comput.* **19**, 577 (1965).
- [5] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations* (SIAM, Philadelphia, 1995).
- [6] T. He, W. Li, Z. Bai, and L. Wang, Numerical solution of the coupled Haïssinski equations for the equilibrium state of an arbitrary bunch train in an electron storage ring, *Nucl. Instrum. Methods Phys. Res., Sect. A* **1006**, 165434 (2021); Longitudinal equilibrium density distribution of arbitrary filled bunches in presence of a passive harmonic cavity and the short range wakefield, *Phys. Rev. Accel. Beams* **24**, 044401 (2021).
- [7] C. Brezinski, Convergence acceleration during the 20th century, *J. Comput. Appl. Math.* **122**, 1 (2000).
- [8] D. G. Anderson, Iterative procedures for nonlinear integral equations, *J. Assoc. Comput. Mach.* **12**, 547 (1965).
- [9] H. F. Walker and P. Ni, Anderson acceleration for fixed-point iteration, *SIAM J. Numer. Anal.* **40**, 1715 (2011).
- [10] H. Fang and Y. Saad, Two classes of multisecond methods for nonlinear acceleration, *Numer. Linear Algebra Appl.* **16**, 197 (2009).
- [11] P. H. Dederichs and R. Zeller, Self-consistency iterations in electronic-structure calculations, *Phys. Rev. B* **28**, 5462 (1983).
- [12] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).
- [13] P. T. Boggs, The solution of nonlinear systems of equations by A-stable integration techniques, *SIAM J. Numer. Anal.* **8**, 767 (1971).
- [14] Preliminary Design Report for ALS-U, available as .pdf file at [https://drive.google.com/file/d/1B\\_jMmOJkZYFLZ3PnCxYOietSPYAtwV03/view](https://drive.google.com/file/d/1B_jMmOJkZYFLZ3PnCxYOietSPYAtwV03/view); (*Note the underscore following B*); see pp. 221–224 and 256–286.
- [15] Dan Wang, Lawrence Berkeley National Laboratory (private communication).
- [16] Z. Pan, S. De Santis, T. Hellert, C. Steier, C. Sun, C. Tang, and M. Venturini, Beam-loading transients and bunch shape in the operation of passive harmonic cavities in the ALS-U, paper THPAK037 in *Proceedings of IPAC18, Vancouver BC, Canada* (2018).
- [17] V. Eyert, A comparative study on methods for convergence acceleration of iterative vector sequences, *J. Comput. Phys.* **124**, 271 (1996).
- [18] G. Kresse and J. Furthmüller, Effective iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [19] P. Pulay, Convergence acceleration of iterative sequences. The case of SCF iteration, *Chem. Phys. Lett.* **73**, 393 (1980).
- [20] W. Kohn and L. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [21] H. De Sterck and Y. He, On the asymptotic linear convergence speed of Anderson acceleration, Nestorov acceleration, and nonlinear GEMRES, *SIAM J. Sci. Comput.* **0**, S21 (2021).