

Linear optics and coupling correction with closed orbit modulation

Xiaobiao Huang^{✉*}

Argonne National Laboratory, Lemont, Illinois 60439, USA



(Received 7 May 2021; accepted 8 July 2021; published 23 July 2021)

We propose a new method to measure and correct linear optics and coupling for storage rings with beam position monitor (BPM) data. Two correctors in each transverse plane are used to modulate the closed orbit in an appropriate pattern to sample the linear optics. The orbit modulation data are fitted to the lattice model in a similar manner as the usual method of fitting the orbit response matrix. The closed orbit modulation measurements can be done much faster, and the number of sample points is not limited by the number of correctors. The method also has a number of advantages over methods that rely on turn-by-turn BPMs. A simulation is done to demonstrate the working principles of the method.

DOI: [10.1103/PhysRevAccelBeams.24.072805](https://doi.org/10.1103/PhysRevAccelBeams.24.072805)

I. INTRODUCTION

Linear optics correction is critical for storage rings to achieve the design performance. A commonly adopted method for linear optics correction is linear optics from closed orbit (LOCO), which fits the measured orbit response matrix to a lattice model to determine the quadrupole errors in the machine [1]. The orbit response matrix consists of orbit deviations at beam position monitor (BPM) locations when the orbit correctors are changed, one at a time. Each column of the orbit response matrix corresponds to the response of one corrector. The LOCO method uses only orbit correctors and BPMs, both of which are available on a typical storage ring. Combined with the constrained fitting technique to address the degeneracy problem due to similarities in the optics perturbations by neighboring quadrupole magnets [2], the LOCO method can be applied to essentially any storage ring. When the cross-plane response data are included and the skew quadrupoles are fitted, linear coupling errors can also be determined and corrected.

One disadvantage of the LOCO method is that measuring the orbit response is a slow process, especially for machines with slow correctors. The time needed to measure one set of orbit response matrix data ranges from ~ 10 min to a few hours. This is not ideal, especially for cases where many datasets are needed, for example, during a lattice development shift, or when measuring the effects of an insertion device operated at various modes and gaps.

Fast measurement has been demonstrated by driving multiple correctors with rf modulation of different frequencies [3–5]. This method substantially reduces the data-taking time for LOCO. The method requires the correctors to have a fast response (e.g., 8 Hz in [3]). However, at many facilities, not all correctors are fast ones. Exclusion of the slow correctors reduces the number of optics sample points. In addition, driving multiple correctors in parallel increases measurement noise in the LOCO data [4].

Turn-by-turn (TBT) orbits taken with a beam undergoing coherent betatron oscillations provides a much faster way to sample the linear optics. The TBT BPM data can be processed with a number of methods to extract the information for optics correction. Some methods derive the linear optics functions (beta functions and betatron phase advances) from the oscillations [6–8], which can, in turn, be used to fit the lattice model [8,9]. Linear coupling can be determined and corrected using decoupled betatron normal modes by the independent component analysis method [10]. The measured phase advances from the TBT BPM data are also used to measure and correct the global optics section by section [11,12].

The TBT orbit data can also be fitted to the lattice model directly by comparing the measured orbits and the orbits obtained through particle-tracking simulation [13]. This method has been extended to one-pass systems such as the Linac Coherent Light Source linac and transport lines [14]. In the latter case, two correctors in each transverse plane were used to scan the trajectory on a grid as a way to sample the linear optics [15].

TBT orbits and closed orbit responses are essentially the same in terms of sampling the linear optics. Both orbits represent deviations from the phase space origin and are governed by the focusing forces along the beam path. Because the betatron tunes are typically away from lower-order resonances, the TBT orbits of many turns of

*huangxb@anl.gov

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

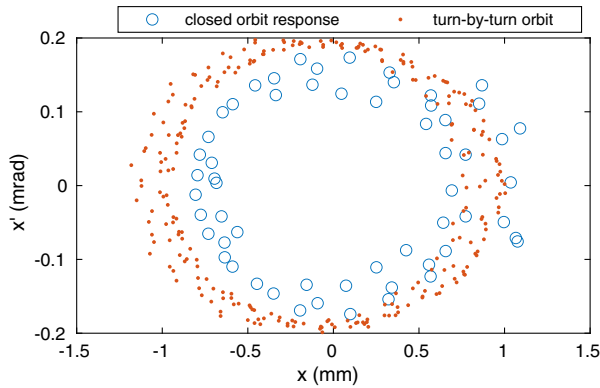


FIG. 1. Comparison of closed orbit responses with 58 correctors and TBT orbits for 256 turns in phase space.

an excited beam usually fills up the circumference of the phase space ellipse. The closed orbit responses of all correctors observed at one location also spread out in the phase space as the correctors are distributed along the ring. This is illustrated in Fig. 1, which shows a comparison of horizontal phase space for 256 turns of TBT orbits with betatron oscillations and closed orbit responses by the 58 correctors for the Stanford Positron Electron Asymmetric Ring (SPEAR3) storage ring.

The two ways to sample the linear optics both have their advantages and disadvantages. The TBT orbit data can be taken within seconds. However, it requires BPMs with TBT capability and a vertical pinger to excite betatron motion in the vertical plane. The data precision for TBT orbit is usually much lower than closed orbit measurements. In addition, decoherence of a kicked bunch can limit the number of turns of usable data to tens of turns. The closed orbit measurements have high precision. However, data taking is slow, and the number of data samples is limited by the number of corrector magnets.

In this study, we propose a new approach to sample the linear optics with BPMs, which can be used to determine the linear optics and coupling errors in the same fashion as the LOCO method or the method of fitting TBT orbit data but without the disadvantages of either method. In the new method, we use a pair of correctors in each transverse plane to modulate the closed orbit such that the orbits are distributed on the phase space ellipse.

The new method is similar to LOCO in that it also measures closed orbits. It is substantially different from LOCO in that it needs only two correctors in each plane and as many sample points can be obtained as necessary. The measurements can be done quickly. The two correctors can be simultaneously driven by sinusoidal signals of the same frequency with a proper phase difference as they are changed to scan the phase space ellipse. In this mode, the data-taking process can complete within a second. Even if the two correctors are discretely sampled, it would still be much faster than orbit response matrix measurement, as

faster correctors can be used and scanning the same correctors in small steps takes less time than switching to different correctors.

The new method may be referred to as linear optics from closed orbit modulation (LOCOM). In the next section, we describe the method in more detail. In Sec. III, simulation results are shown as a demonstration of the working principle. The conclusion is given in Sec. IV.

II. THE LOCOM METHOD

A. Orbit modulation

When one orbit corrector is changed to give an angular kick to the beam, the closed orbit changes in a certain pattern according to the closed orbit response of the corrector. This corresponds to one point away from the origin in the phase space (observed at a given location). Changing the strength of the kick further will generate another point. However, in terms of probing the linear optics, the second point is degenerate with the previous point, as it corresponds to the same pattern, only linearly scaled. New information is brought in only if the closed orbit is shifted to a different angle coordinate in the phase space.

With two orbit correctors in the same plane separated by a betatron phase advance, closed orbits corresponding to any angle coordinate shift can be generated. By varying the kicks of the two correctors, the closed orbit can thoroughly scan the phase space. For the best effect of sampling the linear optics, we prefer that the scanned points populate the circumference of the phase space ellipse of an appropriate size [14].

Suppose the two correctors are located at P_1 and P_2 and apply kicks θ_1 and θ_2 , respectively, as illustrated in Fig. 2. The closed orbit with kicks from the two correctors is the superposition of the closed orbits with each kick individually. Suppose the transfer matrix from P_2 to P_1 is \mathbf{m} ; then the phase space coordinate of an initially unperturbed

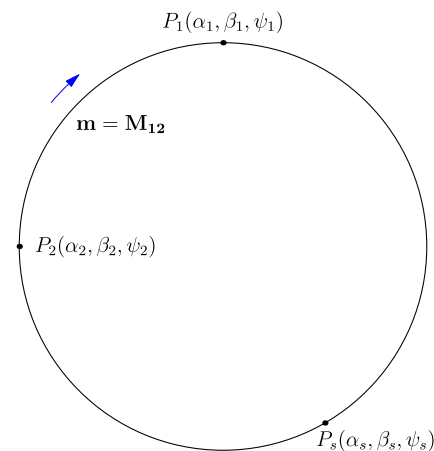


FIG. 2. Illustration of modulating the orbit with two correctors located at P_1 and P_2 .

particle (before P_2) at just downstream of P_1 will be

$$\begin{aligned} \mathbf{m} \begin{pmatrix} 0 \\ \theta_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \theta_1 \end{pmatrix} \\ = \begin{pmatrix} \sqrt{\beta_1 \beta_2} \theta_2 \sin \psi_{12} \\ \theta_1 + \sqrt{\frac{\beta_2}{\beta_1}} \theta_2 (\cos \psi_{12} - \alpha_1 \sin \psi_{12}) \end{pmatrix}, \quad (1) \end{aligned}$$

where $\alpha_{1,2}$ and $\beta_{1,2}$ are Courant-Snyder parameters with the subscript indicating locations P_1 and P_2 , respectively, and ψ_{12} is the phase advance from P_2 to P_1 . The normalized phase space coordinates with the two kicks downstream of P_1 is

$$\begin{pmatrix} x \\ \alpha_1 x + \beta_1 x' \end{pmatrix} = \sqrt{\beta_1} \begin{pmatrix} \sqrt{\beta_2} \theta_2 \sin \psi_{12} \\ \sqrt{\beta_1} \theta_1 + \sqrt{\beta_2} \theta_2 \cos \psi_{12} \end{pmatrix}. \quad (2)$$

Although this is not the closed orbit, we can expect that scanning the kicked coordinate over the local Courant-Snyder ellipse will have a similar effect for sampling the linear optics. It would be ideal if $\psi_{12} = k\pi + \frac{\pi}{2}$ (with integer k), in which case we can simply specify $\theta_1 = \theta_{\text{amp}} \cos \phi$ and $\theta_2 = \sqrt{\frac{\beta_1}{\beta_2}} \theta_{\text{amp}} \sin \phi$, with ϕ increasing with uniform steps.

It is preferred that ψ_{12} is close to the ideal value. In the general case, we could solve for the required kick angles θ_1 and θ_2 needed to uniformly scan the ellipse. However, a uniform scan is not strictly necessary. Since it is desirable to drive the two correctors with sinusoidal signals, we look for the form of

$$\theta_1 = \theta_{\text{amp}} \sin \phi, \quad \theta_2 = \sqrt{\frac{\beta_1}{\beta_2}} \theta_{\text{amp}} \cos(\phi + \chi), \quad (3)$$

with a phase shift χ between the waveforms of the two correctors. We would like to make the x and $\alpha_1 x + \beta_1 x'$ components orthogonal within a period of the drive signal. It is easy to show that this is achieved when

$$\chi = \frac{\pi}{2} - \psi_{12}. \quad (4)$$

Simulation with the SPEAR3 lattice is used to illustrate the concept of closed orbit modulations with two correctors. A pair of horizontal correctors, with phase advance difference $\psi_{x,12} = 1.60\pi$ and a pair of vertical correctors, with $\psi_{y,12} = 0.52\pi$ are used. Figure 3 shows the phase space plots in comparison to the Courant-Snyder ellipses at the same location for the two planes. For both planes, the choices of $\chi = -\psi_{12}$ and $\chi = \frac{\pi}{2} - \psi_{12}$ are shown for comparison. Clearly, the latter case paints out a trace similar to the ideal Courant-Snyder ellipses and is, thus, better suited for linear optics sampling. For the vertical

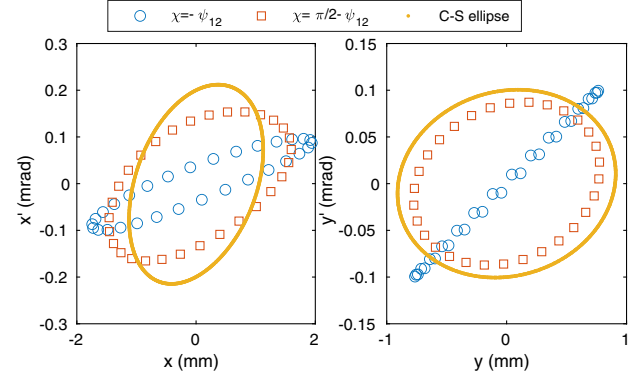


FIG. 3. Phase space plots for simulated closed orbit modulation with the SPEAR3 lattice. Left, horizontal; right, vertical.

plane, as the phase advance between the two correctors is closer to the ideal value of $\frac{\pi}{2}$, the scanned trace is closer to the ideal ellipse.

Figure 4 shows the horizontal orbit samples for the two values of the phase shift χ . The top plot shows the case with $\chi = -\psi_{x,12}$ and the bottom plot for $\chi = \frac{\pi}{2} - \psi_{x,12}$. The orbits shown in the top plot are nearly degenerate, as they are in a similar (but scaled) pattern. The case in the bottom plot is better, as there is a shift in phase between the orbits. Figure 5 shows the orbit samples for the vertical plane, where the degeneracy in the top plot is even more pronounced. The bottom plot in Fig. 5 illustrates the ideal case of orbit modulation, where the amplitude or orbit deviation remains nearly the same, but the phase of the orbit is shifted.

B. Fitting the lattice model

Similar to LOCO, the measured closed orbit patterns in the LOCOM method can be compared to its counterpart in simulation. The differences between measurements and simulation represent the optics errors of the machine, as well as imperfections in the measurements, such as

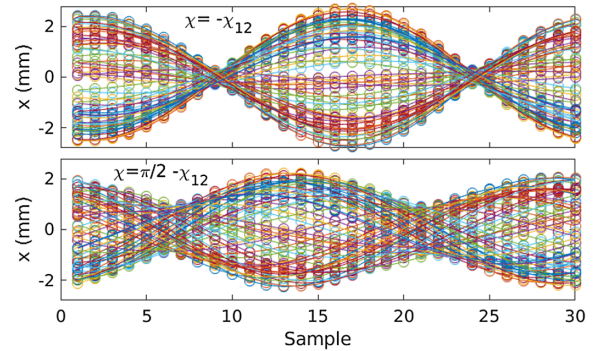


FIG. 4. Comparison of horizontal orbit samples phase shifts between the two correctors set to $\chi = -\psi_{x,12}$ (top) or $\chi = \frac{\pi}{2} - \psi_{x,12}$ (bottom) for the SPEAR3 example. Each sample is an orbit with 57 BPMs.

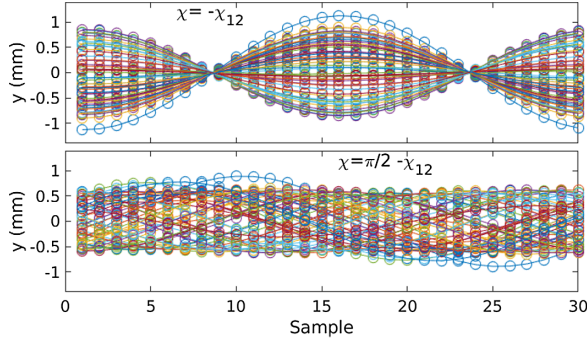


FIG. 5. Comparison of vertical orbit samples for the SPEAR3 example with $\chi = -\psi_{y,12}$ (top) or $\chi = \frac{\pi}{2} - \psi_{y,12}$ (bottom).

corrector and BPM calibration errors. These errors can be fitted to minimize the difference between the measurements and the model predictions. The cross-plane orbit changes (i.e., vertical orbit shifts due to horizontal correctors) represent linear coupling, which can also be fitted, with skew quadrupoles in the lattice model and rolls of correctors and BPMs as fitting parameters. The least-square fitting scheme is the same as LOCO [1] or fitting the TBT orbits [13].

The BPM readings (x and y) are related to the actual beam orbits (\tilde{x} and \tilde{y}) through [16]

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} g_x & c_x \\ c_y & g_y \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix}, \quad (5)$$

where $g_{x,y}$ are the horizontal and vertical gains and $c_{x,y}$ are the coupling coefficients. Similarly, the two correctors also have gain and roll errors. The actual kicks on the beam, $\tilde{\theta}_x$ and $\tilde{\theta}_y$, are related to the apparent kicks (as determined from the power supply) θ_x and θ_y via

$$\begin{pmatrix} \tilde{\theta}_x \\ \tilde{\theta}_y \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix} \begin{pmatrix} k_x \theta_x \\ k_y \theta_y \end{pmatrix}, \quad (6)$$

with corrector gain $k_{x,y}$ and roll ϕ . One advantage of the LOCOM method over LOCO is that only the gain and roll errors for the four correctors need to be fitted, while LOCO needs to include the corrector parameters for each orbit. Generally speaking, the reduced number of fitting parameters help avoid cross talk between the fitting parameters and increase the capability of the fitting scheme to resolve the actual errors.

Dispersion functions can be included as fitting data, as is done for the LOCO method. Because the closed orbit modulation data are similar to orbit response matrix data in nature, the robustness of the LOCOM method is expected to be similar to the LOCO method.

TABLE I. Corrector gains and rolls: Target vs fitted values.

H corrector	k_x target	k_x fitted	ϕ_x target	ϕ_x fitted
1	1.050	1.038	0.010	0.0106
2	0.970	0.965	0.005	0.0051
V corrector	k_y target	k_y fitted	ϕ_y target	ϕ_y fitted
1	0.950	0.941	0.005	0.0050
2	1.040	1.028	-0.005	-0.0066

III. SIMULATION

A simulation is done to test the scheme of fitting LOCOM data to discover lattice errors, using the SPEAR3 lattice model. Quadrupole gradient errors are added artificially to the model to generate optics distortion. Error is also added to a skew quadrupole to introduce coupling. Random gain and coupling errors are introduced to the 57 BPMs. The c_y and c_x coupling coefficients are chosen to be equal in magnitude but with opposite signs such that they represent rolls. Gain and roll errors are also added to the two corrector magnets.

The LOCOM data are generated by modulating the same pairs of correctors in each plane as in the previous section. The modulation waveforms differ in phase by $\frac{\pi}{2} - \psi_{12}$. A total of 30 orbits were generated with the maximum kick angle of 0.1 mrad for each transverse plane. The in-plane orbits are as shown in Figs. 4 and 5, bottom plots. Dispersion functions are also measured with the simulated lattice errors. The dispersion terms are given a weight factor of 25 in the least-square χ^2 definition. Random BPM noise with rms value $1 \mu\text{m}$ is added to the simulated BPM readings. There are a total of 6954 data points.

The fitting parameters are 78 quadrupole gradients, 13 skew quadrupole gradients, BPM gains and coupling coefficients (57×4), and corrector gains and rolls (2×4). There are a total of 327 fitting parameters. The initial χ^2 of the least-square problem is 1.61×10^5 (normalized by BPM noise sigma and the number of data points). This is reduced to 1.464×10^4 , 3070, and 550, respectively,

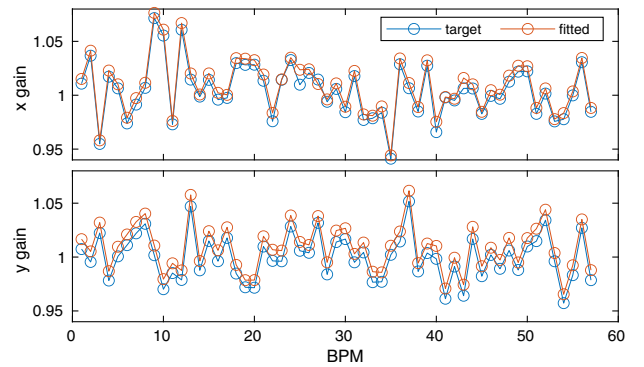


FIG. 6. Comparison of fitted BPM gains and the target values.

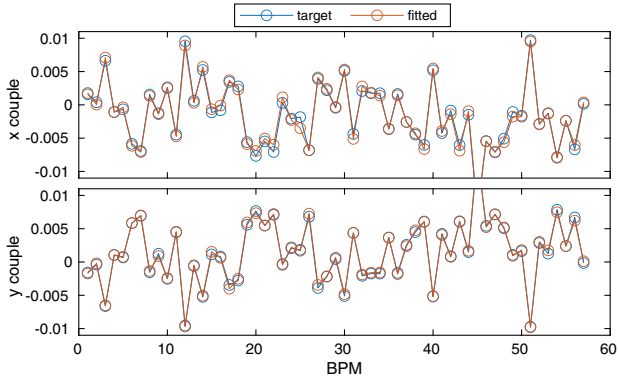


FIG. 7. Comparison of fitted BPM coupling coefficients and the target values. Note that the target values are deliberately made with $c_y = -c_x$.

after one, two, and three iterations. The final value is $\chi^2 = 3.2$ after six iterations.

The corrector gain and roll values are listed in Table I, with target values compared to the fitted values. The fitted BPM gains are shown in Fig. 6, while the fitted BPM coupling coefficients are in Fig. 7. The corrector rolls and BPM coupling coefficients are in good agreement with the expected values. The corrector gains are off by a small amount. However, a large fraction of it comes from the degeneracy due to simultaneously varying all corrector gains and BPM gains. On average, the fitted horizontal

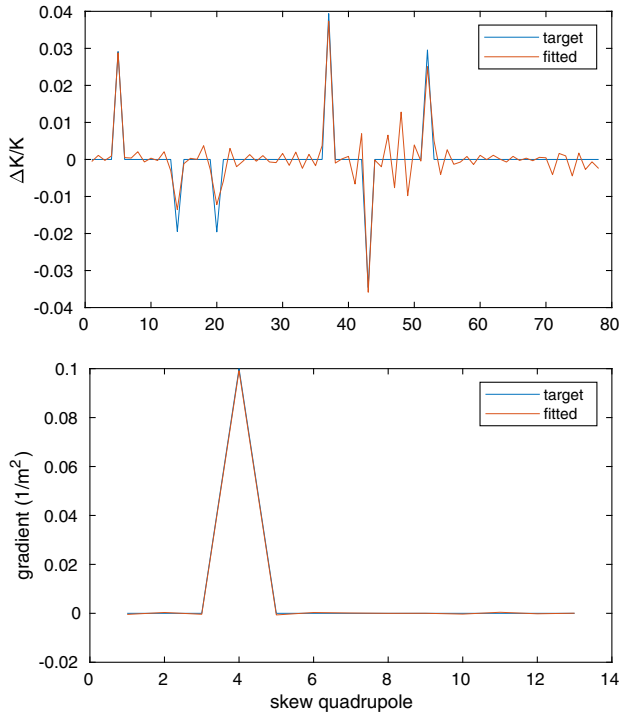


FIG. 8. Comparison of fitted fractional quadrupole gradient changes, $\Delta K/K$ (top), and skew quadrupole gradients (bottom) with respect to the expected values.

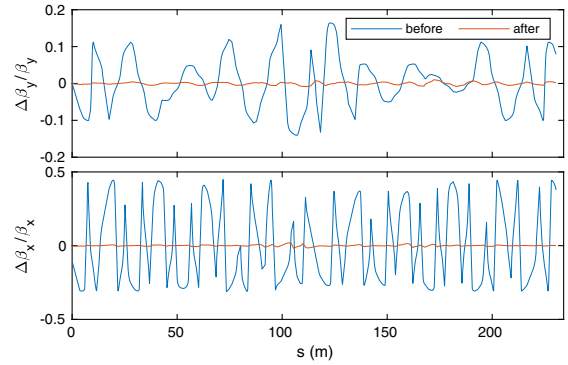


FIG. 9. Beta beating between the reference lattice (with errors) and the ideal lattice (“before”) is compared to beta beating between the reference lattice and the fitted lattice (“after”). Top, vertical; bottom, horizontal.

BPM gains are higher by 0.45%, while the fitted vertical BPM gains are higher by 0.91%.

The fitted quadrupole gradient changes and the skew quadrupole gradients are shown in Fig. 8. The quadrupole errors planted in the lattice model are mostly recovered, with some leakage to the nearby quadrupoles, which is common to lattice calibration by fitting [2]. The rms beta beating with respect to the reference lattice is reduced from 25% to 0.5% for the horizontal plane and from 7% to 0.4% for the vertical plane. The skew quadrupole error is found by fitting with high accuracy.

The reference lattice with artificial errors has rms beta beatings of 25% and 7% for the horizontal and vertical planes, respectively, while the beta beating of the fitted lattice with respect to the reference lattice is only 0.5% and 0.4%, respectively. The beta beating of reference lattice is compared to the beta beating between the reference lattice and the fitted lattice in Fig. 9. The dispersion functions for the reference lattice and the fitted lattice are compared in Fig. 10. The rms values of the dispersion function differences between the reference and fitted lattices are reduced to 8.6 and 1.2 mm for the horizontal and vertical planes, respectively.

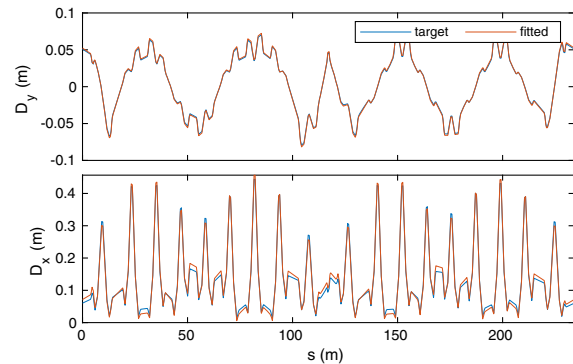


FIG. 10. The dispersion functions of the reference lattice are compared to the fitted lattice. Top, vertical; bottom, horizontal.

IV. CONCLUSION

We propose a new method to measure the linear optics and coupling of a storage ring with BPM data. This method, referred to as LOCOM, uses a pair of correctors in each transverse plane to modulate the closed orbit in a pattern that scans the angle coordinate along the Courant-Snyder ellipse in the phase space. The orbit modulation data are fitted with a lattice model, from which the lattice errors can be determined and used for correction. BPM gains and coupling coefficients as well as the gains and rolls of the four correctors are included in the fitting.

A simulation was done with the SPEAR3 lattice model to test the method. It was shown that the BPM, corrector, quadrupole, and skew quadrupole errors can be determined by fitting the closed orbit modulation data.

Compared to the commonly adopted method of LOCO [1], the LOCOM method has a number of advantages. It uses only two correctors per plane and can generate as many data points as needed. These correctors can be fast magnets, and the modulation scan can be done with small steps. Therefore, data taking is faster. The correctors can be driven with sinusoidal signals of the same frequency but with a proper phase shift, which could reduce the data-taking time to within a second. Compared to the fast orbit response matrix measurement method with sinusoidal corrector modulation [3–5], the present method should be faster as only two correctors are driven, data reduction is simpler as no demixing is needed, and the number of samples is flexible. Furthermore, only eight corrector parameters (gain and roll) need to be fitted, which would help increase the ability of the fitting scheme to resolve lattice errors.

The LOCOM method also has advantages over the methods that use TBT BPM data. It is applicable to practically any storage ring, while TBT BPMs are typically available only at the newer machines. Closed orbit measurements have much higher accuracy than TBT orbits. TBT orbits may have a limited number of usable turns due to decoherence, especially for high chromaticity or high oscillation amplitude. Many rings do not have a pinger to excite beam motion in the vertical plane, which is needed for the TBT approach.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357 (ANL).

- [1] J. Safranek, *Nucl. Instrum. Methods Phys. Res., Sect. A* **388**, 27 (1997).
- [2] X. Huang, J. Safranek, and G. Portmann, *ICFA Newsletter* **44**, 60 (2007), <https://inspirehep.net/files/873ea4d6d6e082f5adac55f41422b58f>.
- [3] I. Martin, M. Abbott, M. Furseman, G. Rehm, and R. Bartolini, in *Proceedings of the Fifth International Particle Accelerator Conference (IPAC 2014), Dresden, Germany* (JACoW, Geneva, 2014), pp. 1763–1765.
- [4] Z. Martí, G. Benedetti, M. Carlà, J. Fraxanet, U. Iriso, J. Moldes, A. Olmos, and R. Petrocelli, in *Proceedings of the Eighth International Particle Accelerator Conference (IPAC 2017), Copenhagen, Denmark* (JACoW, Geneva, 2017), pp. 365–367.
- [5] X. Yang, V. Smaluk, L. H. Yu, Y. Tian, and K. Ha, *Phys. Rev. Accel. Beams* **20**, 054001 (2017).
- [6] P. Castro, J. Borer, A. Burns, G. Morpurgo, and R. Schmidt, in *Proceedings of the 15th Particle Accelerator Conference, PAC-1993, Washington, DC, 1993* (IEEE, New York, 1993), pp. 2103–2105.
- [7] C.-X. Wang, V. Sajaev, and C.-Y. Yao, *Phys. Rev. Accel. Beams* **6**, 104001 (2003).
- [8] X. Huang, S. Y. Lee, E. Prebys, and R. Tomlin, *Phys. Rev. Accel. Beams* **8**, 064001 (2005).
- [9] X. Shen, S. Y. Lee, M. Bai, S. White, G. Robert-Demolaize, Y. Luo, A. Marusic, and R. Tomás, *Phys. Rev. Accel. Beams* **16**, 111001 (2013).
- [10] X. Yang and X. Huang, *Nucl. Instrum. Methods Phys. Res., Sect. A* **828**, 97 (2016).
- [11] M. Aiba, S. Fartoukh, A. Franchi, M. Giovannozzi, V. Kain, M. Lamont, R. Tomás, G. Vanbavinckhove, J. Wenninger, F. Zimmermann, R. Calaga, and A. Morita, *Phys. Rev. Accel. Beams* **12**, 081002 (2009).
- [12] R. Tomás, O. Brüning, M. Giovannozzi, P. Hagen, M. Lamont, F. Schmidt, G. Vanbavinckhove, M. Aiba, R. Calaga, and R. Miyamoto, *Phys. Rev. Accel. Beams* **13**, 121004 (2010).
- [13] X. Huang, J. Sebek, and D. Martin, *Phys. Rev. Accel. Beams* **13**, 114002 (2010).
- [14] T. Zhang, X. Huang, and T. Maxwell, *Phys. Rev. Accel. Beams* **21**, 092801 (2018).
- [15] P. Emma and W. Spencer, in *Proceedings of the 1991 Particle Accelerator Conference, San Francisco, CA, 1991* (IEEE, New York, 1991), pp. 1549–1551.
- [16] J. Safranek, G. Portmann, and A. Terebilo, in *Proceedings of the 8th European Particle Accelerator Conference, Paris, 2002* (EPS-IGA and CERN, Geneva, 2002), pp. 1184–1186.