Accurate and efficient spin integration for particle accelerators

Dan T. Abell[*] and Dominic Meiser

Tech-X Corporation, 5621 Arapahoe Avenue, Boulder, Colorado 80303, USA

Vahid H. Ranjbar

Brookhaven National Laboratory, Upton, New York 11973, USA

Desmond P. Barber[†]

Deutsches Elektronen-Synchroton (DESY), 22603 Hamburg, Germany

Accurate spin tracking is a valuable tool for understanding spin dynamics in particle accelerators and can help improve the performance of an accelerator. In this paper, we present a detailed discussion of the integrators in the spin tracking code GPUSPINTRACK. We have implemented orbital integrators based on drift-kick, bend-kick, and matrix-kick splits. On top of the orbital integrators, we have implemented various integrators for the spin motion. These integrators use quaternions and Romberg quadratures to accelerate both the computation and the convergence of spin rotations. We evaluate their performance and accuracy in quantitative detail for individual elements as well as for the entire RHIC lattice. We exploit the inherently data-parallel nature of spin tracking to accelerate our algorithms on graphics processing units.

I. INTRODUCTION

The origin of nucleon spin remains an enduring puzzle in nuclear physics [1], and elucidating this puzzle is the principal focus of polarized beam experiments at the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory [2–4]. Because statistical uncertainties scale inversely with the square of the polarization [5], optimizing the polarization is essential for the efficient use of experimental resources.

Computer simulations serve an important role in understanding and improving beam polarization. For example, the invariant spin field (ISF) places an important upper bound on the maximum attainable polarization of a stored beam [6,7], and hence a knowledge of the ISF and how it varies with the machine optics is essential to optimizing the beam polarization.

Another motivation for spin-orbit tracking simulations—especially of high accuracy—derives from proposals to use storage rings in searches for a permanent electric dipole moment (EDM) in protons and deuterons [8]. Assessing the sensitivity of such experiments will require long-term spin-orbit simulations of unprecedented accuracy [9]. Other projects that can benefit from similar studies include MEIC [10,11], the LHeC [12,13], and the muon $g - 2$ experiment [14,15]. In addition, accurate spin-orbit tracking enables one to perform careful tests of mathematical concepts related to spin dynamics [16].

In spite of the manifest need for accurate simulations of spin dynamics, relatively few efforts have been made to develop codes that include spin tracking. In the context of hadron machines, which is the context of this paper, spin-orbit tracking codes for large storage rings and accelerators first emerged in the 1990s in response to the needs of specific projects. Méot added spin tracking to ZGOUBI in 1992 [17,18]. Hoffstätter and Berz did the same for COSY-INFINITY in 1993 [19]. Luccio developed SPINK in the mid-1990s to support the goal of accelerating polarized protons in RHIC [20,21]. Hoffstätter and Vogt created SPRINT to study the feasibility of attaining proton polarization at very high energy in HERA [7,22]. The code FORGET-ME-NOT, by Golubeva and Balandin, was developed for this purpose also [23]. In the mid-2000s spin motion was added to PTC by Forest [24–26] and to BMAD by Sagan [27]. Mane has recently added the code ELIMS to this list [28].

Some of these codes also include effects particular to electrons—especially synchrotron radiation. In addition, there are other more specialized codes that handle electrons: Of signal importance is the development of SLIM by Chao [29] in the early 1980s.

The original version of SPINK used orbit transport matrices generated by MAD8—later by MAD-X—to compute the orbital motion. That version was used for extensive studies of the beam polarization in RHIC [20].

---

[*]Corresponding author.
dabell@txcorp.com
[†]Also at Cockcroft Institute, Sci-Tech Daresbury, UK, and at the University of Liverpool, UK.

SPINK was later incorporated into the UAL framework and modified to use TEAPOT's orbital integrators. The current work arose out of an effort to port UAL-SPINK to GPU platforms. In the course of that work, however, it was discovered that even when using TEAPOT's orbital integrators, the code had difficulties with spin convergence [30,31], especially in the neighborhood of a strong spin resonance. Addressing this issue meant slowing down what were already numerically demanding spin tracking simulations.

Here we present a very accurate and efficient spin-orbit tracking code, GPUSPINTRACK, which we have incorporated into the UAL framework [32]. Because we have found that the accuracy of the orbital data has a significant impact on the accuracy of the spin tracking, our code relies on first performing very accurate symplectic integration for the orbital motion. With orbital data in hand, we then use the Thomas-Bargman-Michel-Telegdi (Thomas-BMT, or T-BMT) equation [33,34] to integrate the spin motion. Our integration methods include the effects of acceleration on both orbit and spin degrees of freedom. We note that because GPUSPINTRACK tracks the full nonlinear orbital motion and the full 3D spin motion, it is sensitive to the full range of spin-orbit resonances. This is particularly important in the context of acceleration across spin-orbit resonances. The most significant aspect of this paper is that we have found a means of accelerating the convergence of the spin integration.

To obtain reliable results from numerical studies, one must have a quantitative understanding of the errors inherent in a given simulation. With a quantitative model for the errors, one can perform simulations of the desired accuracy without wasting computational effort motivated by overly pessimistic error estimates. Another significant contribution of this paper is a detailed analysis of the accuracy of our integrators. We study how the orbital and spin motion converge after one turn, and also the long-time stability of our algorithms. For the spin dynamics, we also examine the errors across individual elements.

Numerical efficiency is another important consideration for spin-orbit tracking codes. To increase the speed of our simulations, we have implemented all our integrators on a graphics processing unit (GPU). The embarrassingly parallel nature of spin-orbit tracking (in the absence of space charge) makes this type of computation an ideal fit to the highly parallel architecture of GPUs.

In the next three sections, we describe the dynamical model used for our simulations (Sec. II), the orbital integrators we implement (Sec. III), and our approach to spin integration (Sec. IV). The latter section describes what we refer to as *Romberg quadratures for spin*, our method for accelerating the convergence of spin integration. In Sec. V, we describe the performance of both orbit and spin integrators. We give a brief description of the GPU implementation in Sec. VI. Finally, in Sec. VII, we conclude by discussing our results.

II. THE DYNAMICAL MODEL

In this study, we ignore the effects of synchrotron radiation and space-charge forces. We therefore model the orbital dynamics in an accelerator using a single-particle Hamiltonian appropriate to the externally applied magnetic and electric fields of a particle accelerator. For the spin dynamics, we treat a particle's spin expectation value as a unit three vector that obeys the Thomas-BMT equation [33–35].

A. Model for orbital dynamics

To describe the orbital motion of particles in a beam line element, we use longitudinal distance s along the element as our independent variable, and we use the same canonical phase-space coordinates as MAD [36]:

$$\vec{z} = (X, P_x, Y, P_y, T, P_t). \tag{1}$$

Here X and Y denote the horizontal and vertical coordinates in the local reference frame of our beam line element; P_x and P_y denote the corresponding conjugate momenta, divided by a fixed *scale momentum* $p_o = m\gamma_o\beta_o c$; $T = -c\Delta t$ measures the flight time (times the speed of light, c) relative to a reference particle with momentum p_o; and P_t denotes the energy deviation scaled by $p_o c$, so that

$$P_t = \frac{m\gamma c^2 - m\gamma_o c^2}{p_o c} = \frac{m\gamma c^2}{p_o c} - \frac{1}{\beta_o}. \tag{2}$$

Note that the minus sign in the definition of T means that a *positive* value for T implies that our particle arrives *earlier* than the reference particle.

For magnetostatic elements with strictly transverse fields that do not vary along the length, our Hamiltonian for a particle of charge q has the general form

$$H = -(1 + hX)\sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2 - P_x^2 - P_y^2}$$

$$- \frac{q}{p_o}(1 + hX)A_s(X, Y) + \frac{1}{\beta_o}P_t, \tag{3}$$

where $h = 1/\rho_c$ denotes the curvature of the local coordinate frame—h will vanish except in sector bends—and $A_s(X, Y)$ denotes the longitudinal component of the element's vector potential. (We choose a gauge in which the transverse components of the vector potential vanish.) Note that this Hamiltonian assumes any bending occurs only in the horizontal plane; for bends in any other plane, we simply apply rotations before and after a horizontal bend. The last term, P_t/β_o, accounts for the flight time of a particle with the reference momentum traversing the orbit $X = Y = 0$.

When integrating through a dipole, one must exercise some care concerning the relation between the fixed geometry, defined by a magnet's curvature and placement, and the variable physics, defined by a magnet's field

strength. For sector bends, the fixed geometry means the curvature h of the local frame is fixed. In the usual practice, one then sets the scale momentum p_o according to this curvature, so that $p_o = qB_1\rho_c$, where B_1 denotes the magnet's design field strength. Then $qB_1/p_o = h$, and the vector potential term in (3) simplifies to $\frac{1}{2}(1 + hX)^2$, but *only* if the magnet is correctly powered. To include the possibility of mispowering errors, one must not thus confuse the geometry with the physics; one must instead retain the dependence of the Hamiltonian on the actual magnetic field strength.

For rectangular bends, which have curvature $h = 0$, a more fundamental difficulty arises from the fact that the orbital motion is most simply integrated using Cartesian coordinates, whereas the curved design orbit does not follow the straight magnet axis. This means the horizontal phase-space variables (X, P_x), which describe deviations from the magnet axis, cannot describe deviations from any choice of reference orbit. Moreover, for a given fixed bending angle, the path length—hence also the flight time— will depend on the entrance angle. As a consequence, the flight time along even a reference orbit will depend on how the rectangular bend is oriented relative to adjacent beam line elements. If we wish the phase-space variables of (1) at both entrance and exit of this magnet to represent deviations, then we must view the Hamiltonian (3) as written in a set of variables *internal* to this magnet (and we must drop the term P_t/β_o). We then perform the appropriate conversions at entrance and exit.

Of course real magnets, because they obey the dictates of Maxwell, do *not* have "strictly transverse fields that do not vary along the length." That fiction—to which no sensible spectroscopist subscribes—proves useful in much of accelerator physics because (a) long magnets have fields for which that fiction is close to the truth, (b) short magnets yield orbital motion that is dominated by the integrated strength, and (c) fringe-field effects in multipole magnets, which arise primarily from the longitudinal field component, have contributions from both entrance and exit that tend to cancel, particularly in short magnets [37]. Still the field variation across a magnet fringe *can* have a significant impact on the beam dynamics. Cases where this is especially true include solenoids, whose longitudinal fields naturally have extended fringe regions; and rectangular bends, for which non-normal entrance and exit angles lead to vertical focusing [38]. In addition, nonlinear contributions in, for example, quadrupole fringe fields can cause large-emittance beams to occupy a large footprint in tune space [39].

In this paper we shall, for the most part, avoid detailed discussions of fringe fields. We make this choice for a couple of reasons. (i) The simulations shown in this paper were all done in the context of the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Lab, for which fringe field effects are relatively small. (ii) The subject of

transfer maps for magnetic fringe fields is complicated, and very good discussions are available elsewhere. In addition to the work cited above, the reader may consult Forest and colleagues [40–43] for work done mostly in the context of a hard-edge limit, or Dragt and colleagues [44–47] for work on transfer maps across realistic magnets that include fringe fields.

B. Model for spin dynamics

We describe a particle's spin expectation value by a unit three vector $\vec{S}$. The precession of this spin in a magnetic field $\vec{B}$ is governed by the Thomas-BMT equation [33,34], which says that in the rest frame of the magnet—our laboratory frame—the spin precesses according to the rule

$$\frac{d\vec{S}}{dt} = \vec{S} \times \frac{q}{m\gamma}[(1 + G\gamma)\vec{B} - G(\gamma - 1)(\hat{u} \cdot \vec{B})\hat{u}]. \quad (4)$$

An additional term, not shown here, must be included in the presence of electric fields. In this equation, G denotes the gyromagnetic anomaly, $(g - 2)/2$, with g the particle's gyromagnetic ratio; and $\hat{u}$ denotes the unit velocity vector, obtained by normalizing the momentum vector $(P_x, P_y, P_s)^T$. Here P_s denotes the longitudinal component of the scaled momentum, given by

$$P_s = \sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2 - P_x^2 - P_y^2}. \quad (5)$$

As in the case of orbital motion, it proves convenient to transform the equation of spin motion (4) to one with s as the independent variable. To do this, we multiply both sides of (4) by $dt/ds = 1/\dot{s}$. It is important to note that in a curved Frenet-Serret coordinate system, $\dot{s} \neq v_s$. Rather, simple geometry tells us that

$$v_s = (\rho_c + X)\dot{\theta} = (1 + hX)\rho_c\dot{\theta} = (1 + hX)\dot{s}.$$

Using this result, we now obtain the modified T-BMT equation in the form

$$\frac{d\vec{S}}{ds} = \vec{\Omega} \times \vec{S}, \quad (6a)$$

where

$$\vec{\Omega} = -\frac{1 + hX}{(B\rho)_o P_s}[(1 + G\gamma)\vec{B} - G(\gamma - 1)(\hat{u} \cdot \vec{B})\hat{u}] + h\hat{y}, \quad (6b)$$

and $(B\rho)_o = p_o/q$ denotes the (signed) reference value of the magnetic rigidity. The last term, $h\hat{y}$, accounts for the local frame rotation (assuming the rotation is about the axis $\hat{y}$).

Our model does not include spin-dependent contributions to the orbital motion, so-called Stern-Gerlach forces, because these are completely negligible for essentially all accelerators [48].

III. ORBIT INTEGRATION

To integrate a particle's orbital motion through a given beam line element, we use the standard technique of writing the relevant Hamiltonian as a sum of one or several integrable pieces. One then obtains an approximate solution for the total Hamiltonian by an appropriate concatenation of exact solutions for the separate pieces. An important advantage of this approach of splitting the Hamiltonian is that it produces a symplectic integrator—i.e. an integrator that preserves the fundamental structure that underlies any Hamiltonian system. A second advantage is that by composing the partial solutions in a symmetric fashion, one can achieve second-order convergence to the exact solution. Indeed, more sophisticated symmetric compositions allow one to achieve even higher order convergence [37,49]. For an extensive and valuable introduction to the literature on such *splitting methods*, see McLachlan and Quispel [50].

There are several ways one might split the Hamiltonian (3). One of the oldest—still widely used—is the drift-kick split that is the basis of TEAPOT [37,51]. One virtue of the drift-kick split is that it applies to a wide variety of beam line elements. If, however, we focus on particular elements, we can tailor the split accordingly. For ideal bending magnets—sector bends and rectangular bends—one can derive exact solutions, see Secs. III A and III B. For quadrupoles, one can split off the linear transverse motion, leaving a much smaller nonlinear kick as a correction, see Sec. III C. For higher-order multipoles, we return to the drift-kick split of TEAPOT.

A. Sector bend

For a pure sector bend, one may write the vector potential term in (3) as

$$-\frac{q}{p_\mathrm{o}}A_s(X,Y) = \frac{b_1}{2h(B\rho)_\mathrm{o}}(1+hX)^2, \qquad (7)$$

where b_1 denotes the actual dipole field strength, and h denotes the (fixed) physical curvature $1/\rho_c$ of the magnet. One can solve the corresponding equations of motion analytically [41]. First define the scaled momentum in the horizontal plane,

$$P_\alpha = \sqrt{1+\frac{2}{\beta_\mathrm{o}}P_t + P_t^2 - P_y^2}. \qquad (8)$$

This quantity is a constant of the motion. Also define the dimensionless parameter

$$\eta = \frac{b_1}{h(B\rho)_\mathrm{o}} = \frac{b_1/h}{p_\mathrm{o}/q}, \qquad (9)$$

which measures the magnetic field strength relative to its design value. Then, setting $s=0$ at the magnet entrance, we obtain the sector bend trajectory in the form

$$X(s) = \frac{1}{h}\left[\frac{1}{\eta}\left(\sqrt{P_\alpha^2 - P_x^2(s)} - \frac{1}{h}P_x'(s)\right) - 1\right], \qquad (10a)$$

$$P_x(s) = P_x^i\cos(hs) + \left[\sqrt{P_\alpha^2 - (P_x^i)^2} - \eta(1+hX^i)\right]\sin(hs), \qquad (10b)$$

$$Y(s) = Y^i + \frac{1}{\eta}P_y s + \frac{1}{\eta h}P_y \\ \times\left[\sin^{-1}\left(\frac{P_x^i}{P_\alpha}\right) - \sin^{-1}\left(\frac{P_x(s)}{P_\alpha}\right)\right], \qquad (10c)$$

$$P_y(s) = P_y^i, \qquad (10d)$$

$$T(s) = T^i - \frac{1}{\eta}P_t s - \frac{1}{\eta h}\left(\frac{1}{\beta_0}+P_t\right) \\ \times\left[\sin^{-1}\left(\frac{P_x^i}{P_\alpha}\right) - \sin^{-1}\left(\frac{P_x(s)}{P_\alpha}\right)\right] + \left(1-\frac{1}{\eta}\right)\frac{s}{\beta_\mathrm{o}}, \qquad (10e)$$

$$P_t(s) = P_t^i, \qquad (10f)$$

where the superscript i denotes values at the entrance.

Most simulations set the scale momentum p_o such that $p_\mathrm{o}/q = B_1/h$, where B_1 denotes the design value of the magnetic field strength. With this choice, the parameter η becomes

$$\eta = \frac{b_1/h}{B_1/h} = \frac{B_1 + \Delta B}{B_1} = 1 + \frac{\Delta B}{B_1} \equiv 1 + \delta_B,$$

where δ_B denotes the relative magnet mispowering error. For this choice of scaling, and for a correctly powered magnet, i.e., $\delta_B = 0$ and $\eta = 1$, an on-energy, on-axis particle will enter the magnet with $\vec{z} = 0$, and will also exit the magnet with $\vec{z} = 0$. In other words, this choice of scaling means that the map (10) with $\eta = 1$ preserves the origin.

B. Rectangular bend

For an ideal rectangular bend, one may write the vector potential term in (3) as

$$-\frac{q}{p_{\rm o}}A_s(X,Y) = \frac{b_1}{(B\rho)_{\rm o}}X, \qquad (11)$$

where b_1 and $(B\rho)_{\rm o}$ have the same meanings as for the sector bend. One can again solve the corresponding equations of motion analytically [41]. Here we define the length parameter

$$\rho_b = \frac{(B\rho)_{\rm o}}{b_1}, \qquad (12)$$

and we retain the definition of P_α as in (8). Then, again setting $s = 0$ at the magnet entrance,

$$X(s) = X^i + \rho_b\left[\sqrt{P_\alpha^2 - P_x^2(s)} - \sqrt{P_\alpha^2 - (P_x^i)^2}\right], \qquad (13a)$$

$$P_x(s) = P_x^i - \frac{s}{\rho_b}, \qquad (13b)$$

$$Y(s) = Y^i + P_y\rho_b\left[\sin^{-1}\left(\frac{P_x^i}{P_\alpha}\right) - \sin^{-1}\left(\frac{P_x(s)}{P_\alpha}\right)\right], \qquad (13c)$$

$$P_y(s) = P_y^i, \qquad (13d)$$

$$T(s) = T^i - \left(\frac{1}{\beta_0} + P_t\right)$$
$$\times \rho_b\left[\sin^{-1}\left(\frac{P_x^i}{P_\alpha}\right) - \sin^{-1}\left(\frac{P_x(s)}{P_\alpha}\right)\right], \qquad (13e)$$

$$P_t(s) = P_t^i. \qquad (13f)$$

Remember that, per our discussion near the end of Sec. II A, the phase-space coordinates here are *internal* to the rectangular bend. We require transformations that connect these coordinates to those used outside this magnet. See (16) and its solution (17) below.

Since a beam usually enters and exits a rectangular bend at some nonzero angle ϕ with respect to the magnet face, one must (a) perform dynamic rotations that transform the beam into and out of the Cartesian frame of the magnet, and (b) account for the vertical focusing effect. We describe the latter first.

At the entrance, or exit, of a rectangular bend, a typical beam has a significant horizontal momentum P_x with respect to a Cartesian frame aligned with the corresponding magnet face—approximately the sine of half the total bend angle of that magnet. Because the magnet fringe field includes a small horizontal component proportional (at lowest order) to the vertical displacement, a particle in that beam experiences a focusing kick [see (15d) below]

$$\Delta P_y \propto -(\pm\mathfrak{b})P_x Y, \qquad$$

where $\mathfrak{b} = eb_1/p_{\rm o} = b_1/(B\rho)_{\rm o} = 1/\rho_b$ denotes the scaled dipole field strength, and the upper (lower) sign refers to

the magnet entrance (exit). This kick, however, cannot represent the whole story, for this by itself cannot constitute a canonical transformation.

Treating the magnetic field of a rectangular bend as midplane symmetric and as having no dependence on the horizontal coordinate X, one may write the corresponding vector potential in the form

$$\vec{A} = \left[\sum_{k=1}^{\infty}(-1)^k B_{\rm o}^{(2k-1)}(z)\frac{y^{2k}}{(2k)!}, \; 0, \; -xB_{\rm o}(z)\right], \qquad (14)$$

where $B_{\rm o}(z)$ denotes the midplane magnetic field. In the thin-fringe limit, $B_{\rm o}(z)$ becomes a step function, and the $B_{\rm o}^{(j)}(z)$ become a delta function and multiple derivatives thereof. From here there remains a long journey to reach the goal of a symplectic transfer map that describes orbital motion across a thin fringe. We simply quote the result. In the absence of a finite-gap correction, one may obtain [41]

$$X^f = X^i + \frac{\pm\mathfrak{b}}{2P_s}\left\{\frac{1+(x')^2}{1+(y')^2} - \frac{2(x'y')^2}{[1+(y')^2]^2}\right\}(Y^f)^2, \qquad (15a)$$

$$P_x^f = P_x^i, \qquad (15b)$$

$$Y^f = 2Y^i\bigg/\left[1 + \sqrt{1 \pm \mathfrak{b}\frac{x'y'}{1+(y')^2}2Y^i}\right], \qquad (15c)$$

$$P_y^f = P_y^i - (\pm\mathfrak{b})\frac{x'}{1+(y')^2}Y^f, \qquad (15d)$$

$$T^f = T^i - (\pm\mathfrak{b})\left(\frac{1}{\beta_{\rm o}} + P_t\right)\frac{x'}{2P_s^2}\frac{1-(y')^2}{[1+(y')^2]^2}(Y^f)^2, \qquad (15e)$$

$$P_t^f = P_t^i. \qquad (15f)$$

Here, x' and y' respectively denote P_x/P_s and P_y/P_s, and P_s remains as given in (5). For the result including a finite-gap correction, consult [42,43].

The dynamic rotations that transform the beam into and out of the Cartesian frame of the magnet amount to drifts in cylindrical coordinates. Because these transformations rotate the spin, they are important; on the other hand, they play no role in the *accuracy* of our spin integration. Nevertheless, for the sake of completeness, we include them here: The relevant Hamiltonian,

$$H_{\rm prot} = -X\sqrt{1 + \frac{2}{\beta_{\rm o}}P_t + P_t^2 - P_x^2 - P_y^2} = -XP_s, \qquad (16)$$

yields the solution [37,41]

$$X(\phi) = \frac{X^i\sec\phi}{1 - (P_x^i/P_s^i)\tan\phi}, \qquad (17a)$$

$$P_x(\phi) = P_x^i \cos\phi + P_s^i \sin\phi, \tag{17b}$$

$$Y(\phi) = Y^i + \frac{P_y^i}{P_s^i} \cdot \frac{X^i \tan\phi}{1 - (P_x^i/P_s^i)\tan\phi}, \tag{17c}$$

$$P_y(\phi) = P_y^i, \tag{17d}$$

$$T(\phi) = T^i + \frac{1/\beta_o + P_t}{P_s^i} \cdot \frac{X^i \tan\phi}{1 - (P_x^i/P_s^i)\tan\phi}, \tag{17e}$$

$$P_t(\phi) = P_t^i. \tag{17f}$$

One needs to apply this transformation at the entrance with angle ϕ^{entr}, and at the exit with angle $-\phi^{\text{exit}}$. In addition, if the design orbit is not symmetric across the magnet, i.e., if $\phi^{\text{exit}} \neq -\phi^{\text{entr}}$, one must correct for the horizontal offset

$$X^{\text{offset}} = -L \frac{\cos\phi^{\text{entr}} - \cos\phi^{\text{exit}}}{\sin\phi^{\text{entr}} - \sin\phi^{\text{exit}}}. \tag{18}$$

Last, if the temporal variable represents a deviation, then one must subtract from T the amount

$$T_o = -\frac{L_{\text{arc}}}{\beta_o} = -\frac{L}{\beta_o} \cdot \frac{\phi^{\text{entr}} - \phi^{\text{exit}}}{\sin\phi^{\text{entr}} - \sin\phi^{\text{exit}}} \tag{19}$$

corresponding to the time for the reference particle to cross this magnet.

C. Quadrupole

For an ideal quadrupole, we write the vector potential term in (3) as

$$-\frac{q}{p_o} A_s(X, Y) = \frac{b_2}{2(B\rho)_o}(X^2 - Y^2), \tag{20}$$

where b_2 denotes the quadrupole gradient. The most accurate known split of the resulting Hamiltonian separates out the linear transverse motion, retaining the exact dependence on the energy deviation [41]; thus,

$$H_q = H_{q\text{L}} + H_{q\text{NL}}, \tag{21a}$$

where

$$H_{q\text{L}} = \frac{1}{2} \frac{P_x^2 + P_y^2}{\sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2}} + \frac{b_2}{2(B\rho)_o}(X^2 - Y^2), \tag{21b}$$

and

$$H_{q\text{NL}} = -\sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2 - P_x^2 - P_y^2}$$
$$-\frac{1}{2} \frac{P_x^2 + P_y^2}{\sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2}} + \frac{P_t}{\beta_o}. \tag{21c}$$

This is the splitting we use for the quadrupole Hamiltonian.

An alternative to the above split [52] does not include the correction for the energy deviation—the square root in (21b). While that split fails to produce the correct tune for off-energy particles, it does have the virtue of speed, because the same transfer matrix applies to all particles.

To solve the Hamiltonians in (21), we first define the scaled total momentum,

$$P = \sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2}, \tag{22}$$

which is a constant of the motion in the quadrupole, and also the (energy-dependent) focusing strength,

$$\kappa = \sqrt{\frac{b_2}{(B\rho)_o P}}. \tag{23}$$

Then, setting $s = 0$ at the magnet entrance, one obtains the solution for the linear transverse motion, exact in the energy deviation, as

$$X^{(\text{L})}(s) = X^i \cos(\kappa s) + s \frac{P_x^i}{P} \frac{\sin(\kappa s)}{\kappa s}, \tag{24a}$$

$$P_x^{(\text{L})}(s) = P_x^i \cos(\kappa s) - \kappa P X^i \sin(\kappa s), \tag{24b}$$

$$Y^{(\text{L})}(s) = Y^i \cosh(\kappa s) + s \frac{P_y^i}{P} \frac{\sinh(\kappa s)}{\kappa s}, \tag{24c}$$

$$P_y^{(\text{L})}(s) = P_y^i \cosh(\kappa s) + \kappa P Y^i \sinh(\kappa s), \tag{24d}$$

$$\begin{aligned}
T^{(\text{L})}(s) = T^i - \frac{1}{P}\left(\frac{1}{\beta_0} + P_t\right) \\
\times \left\{ \frac{s}{4}\left[\frac{(P_x^i)^2 + (P_y^i)^2}{P^2} + \kappa^2((X^i)^2 - (Y^i)^2)\right] \right. \\
+ \frac{s}{4}\left[\left(\frac{P_x^i}{P}\right)^2 - (\kappa X^i)^2\right]\frac{\sin(2\kappa s)}{2\kappa s} \\
+ \frac{s}{4}\left[\left(\frac{P_y^i}{P}\right)^2 + (\kappa Y^i)^2\right]\frac{\sinh(2\kappa s)}{2\kappa s} \\
\left. - \frac{\kappa s}{2}\left[X^i \frac{P_x^i}{P}\frac{\sin^2(\kappa s)}{\kappa s} - Y^i \frac{P_y^i}{P}\frac{\sinh^2(\kappa s)}{\kappa s}\right]\right\}, \tag{24e}
\end{aligned}$$

$$P_t^{(\text{L})}(s) = P_t^i. \tag{24f}$$

This solution assumes a positive value for the quadrupole gradient b_2, so that this magnet focuses in the X-P_x plane and defocuses in the Y-P_y plane. We can always convert a skew quadrupole, or a more general field orientation, to this case by applying rotations in the transverse plane—i.e. about the longitudinal axis—before and after the element.

The nonlinear part of the quadrupole Hamiltonian, $H_{q\mathrm{NL}}$, represents a "kick" of the position coordinates. A straightforward integration yields the solution

$$X^{(\mathrm{NL})}(s) = X^i + sP_x^i\left(\frac{1}{P_s^i} - \frac{1}{P}\right), \tag{25a}$$

$$P_x^{(\mathrm{NL})}(s) = P_x^i, \tag{25b}$$

$$Y^{(\mathrm{NL})}(s) = Y^i + sP_y^i\left(\frac{1}{P_s^i} - \frac{1}{P}\right), \tag{25c}$$

$$P_y^{(\mathrm{NL})}(s) = P_y^i, \tag{25d}$$

$$\begin{aligned}
T^{(\mathrm{NL})}(s) = T^i - s&\left(\frac{1}{\beta_\mathrm{o}} + P_t\right) \\
&\times \left(\frac{1}{P_s^i} - \frac{1}{P}\frac{(P_x^i)^2 + (P_y^i)^2}{2P^2}\right) + \frac{s}{\beta_\mathrm{o}},
\end{aligned} \tag{25e}$$

$$P_t^{(\mathrm{NL})}(s) = P_t^i, \tag{25f}$$

where

$$P_s^i = \sqrt{1 + 2P_t/\beta_\mathrm{o} + P_t^2 - (P_x^i)^2 - (P_y^i)^2}. \tag{26}$$

In addition to producing the correct tunes for off-energy particles, the Hamiltonian split shown in (21) has the additional advantage that the term $H_{q\mathrm{NL}}$ is small in the paraxial approximation. In particular, it is fourth order in the transverse momenta: $H_{q\mathrm{NL}} \sim \mathcal{O}(P_x^m P_y^n)$ with $m + n = 4$.

We employ the linear and nonlinear solutions shown here by concatenating the corresponding maps in a time-reversal-symmetric manner. Using the simplest symmetric approximation,

$$e^{s:H_q:} \approx e^{s/2:H_{q\mathrm{L}}:}e^{s:H_{q\mathrm{NL}}:}e^{s/2:H_{q\mathrm{L}}:}, \tag{27}$$

we obtain results that are accurate through to second order in the step size s. (Here we have used the colon notation introduced by Dragt [53,54].) If desired, one may use higher-order versions of (27) [37,49,50].

For information about transfer maps for quadrupole fringe fields, consult [40].

D. Higher-order multipole

For an ideal straight multipole magnet—sextupole, octupole, etc.—one may write the vector potential term in (3) as

$$-\frac{q}{p_\mathrm{o}}A_s^{(m)}(X, Y) = \mathrm{Re}\frac{b_m + \mathrm{i}a_m}{m(B\rho)_\mathrm{o}}(X + \mathrm{i}Y)^m, \tag{28}$$

where the b_m and a_m respectively denote the normal and skew multipole coefficient for a magnet with $2m$ poles. We split the resulting Hamiltonian into two pieces:

$$H_{\mathrm{mult}} = H_{\mathrm{drift}} + H_m, \tag{29a}$$

where

$$\begin{aligned}
H_{\mathrm{drift}} &= -\sqrt{1 + \frac{2}{\beta_\mathrm{o}}P_t + P_t^2 - P_x^2 - P_y^2} + \frac{1}{\beta_\mathrm{o}}P_t \\
&= -P_s + \frac{1}{\beta_\mathrm{o}}P_t,
\end{aligned} \tag{29b}$$

and H_m denotes the multipole term in (28), or possibly a superposition of such terms. This constitutes the well-known *drift-kick* split. Because H_{drift} depends only on momenta, and H_m only on coordinates, they each generate very simple motion. We then solve the total Hamiltonian, H_{mult}, in a manner exactly analogous to our solution in (27) of the quadrupole Hamiltonian (21).

For information about transfer maps for multipole fringe fields, consult [40].

E. Thin lens

A thin lens is a representation of a (short) multipole magnet wherein we imagine the length shrinks to zero and the strengths b_m and a_m grow to infinity in such a manner that their product remains constant. Except for taking this limit, we treat the thin lens just as we do the multipole of Sec. III D. Since the drift vanishes, one may think of this as "a drift-kick split without the drift."

F. TEAPOT

The orbital integration performed in TEAPOT [51] is essentially identical in spirit to the integration of multipoles described above in Sec. III D: One splits the general Hamiltonian (3) into a drift (in Frenet-Serret coordinates for a sector bend) and a momentum kick due to the magnetic field. Thus

$$H_T = H_D + H_K, \tag{30a}$$

where, in general,

$$H_D = -(1 + hX)\sqrt{1 + \frac{2}{\beta_\mathrm{o}}P_t + P_t^2 - P_x^2 - P_y^2}, \tag{30b}$$

and

$$H_K = -\frac{q}{p_0}(1 + hX)A_s(X, Y). \qquad (30c)$$

When h vanishes, one obtains exactly the multipole drift-kick split of Sec. III D. In the case of nonzero curvature, i.e., $h \neq 0$, the vector potential A_s becomes quite complicated. Nevertheless, because H_K depends only on coordinates, one may easily compute the motion generated.

The drift Hamiltonian H_D differs only slightly from that in (16), and we obtain the solution from (17) by replacing X with $1/h + X$, and ϕ with hs:

$$X(s) = \frac{(1/h + X^i)\sec(hs)}{1 - (P_x^i/P_s^i)\tan(hs)} - \frac{1}{h}, \qquad (31a)$$

$$P_x(s) = P_x^i\cos(hs) + P_s^i\sin(hs), \qquad (31b)$$

$$Y(s) = Y^i + \frac{P_y^i}{P_s^i} \cdot \frac{(1/h + X^i)\tan(hs)}{1 - (P_x^i/P_s^i)\tan(hs)}, \qquad (31c)$$

$$P_y(s) = P_y^i, \qquad (31d)$$

$$T(s) = T^i - \frac{1/\beta_0 + P_t}{P_s^i} \cdot \frac{(1/h + X^i)\tan(hs)}{1 - (P_x^i/P_s^i)\tan(hs)}, \qquad (31e)$$

$$P_t(\phi) = P_t^i. \qquad (31f)$$

As in (27), we then combine the two solutions—the one above for H_D, and the one for H_K—to approximate the motion generated by the full Hamiltonian H_T.

A disadvantage of this integration scheme is that the two pieces of the Hamiltonian H_T in (30a) may be of similar order. While the resulting integrator is still second-order accurate in s, it has a larger constant factor in the error term. This yields, overall, a less accurate integration for a given number of slices. This proves to be a significant issue when we go on to integrate the spin motion.

On the other hand, drift-kick integration has the advantage that it applies to any magnetic field described by a longitudinal vector potential $A_s(X, Y)$. Moreover, it does not suffer from the apparent discontinuity in the map that occurs for the quadrupole when using the matrix-kick integrator of Sec. III C. Avoiding the use of control-flow statements [if (focusingQuad) {···} else {···}] can be an important consideration when using an integrator for map computation rather than particle tracking [37].

G. Solenoid

For the solenoid, we use a model which has a uniform longitudinal field in the body and the flux return confined to a thin-pancake fringe region. The longitudinal body field necessitates a transverse vector potential $(b_0/2)\rho\hat{\phi}$, which we write in the Cartesian form

$$A_x = -\frac{1}{2}b_0 Y, \qquad A_y = \frac{1}{2}b_0 X, \qquad (32)$$

where b_0 denotes the magnetic field strength.

For our thin-pancake fringe model, we multiply the components (32) by a profile function $S_\varepsilon(z)$ that rises from zero to one across an entrance fringe of length ε, remains constant across the body, and then returns to zero across an exit fringe of length ε. We eventually take the limit $\varepsilon \to 0$. This model means our pancake fringe has a longitudinal field that rises, or falls, linearly across the fringe, and a radial field with flux $b_0\pi\rho^2$ crossing an area $2\pi\rho\varepsilon$ at radius ρ.

Because we have a transverse vector potential, we must modify the general Hamiltonian (3): Since we use a Cartesian reference frame, we set $h = 0$; of course we set $A_s = 0$; and the canonical momentum now equals kinetic momentum plus vector potential. We therefore write the Hamiltonian for a magnetic solenoid in the form

$$H_{\text{sol}} = -\left[1 + \frac{2}{\beta_0}P_t + P_t^2 - \left(P_x - \frac{e}{p_0}A_x\right)^2\right.$$
$$\left. - \left(P_y - \frac{e}{p_0}A_y\right)^2\right]^{1/2} + \frac{1}{\beta_0}P_t$$
$$= -P_s + \frac{1}{\beta_0}P_t. \qquad (33)$$

On crossing one of the solenoid fringes, a particle experiences a transverse kick from the radial component of the magnetic field. The transverse vector potential also changes across the same fringe. It turns out that for our thin-pancake fringe model, those changes—to the kinetic momentum and the vector potential—cancel in a manner that leaves the transverse canonical momentum unchanged across the fringe. Moreover, in the small-ε limit, the transverse coordinates do not change across a fringe. Finally, we ignore the effect on the longitudinal momentum, because it is second order in the transverse dynamical variables. The net result is that for our thin-pancake fringe, the Hamiltonian (33) with vector potential (32) describes the entire solenoid.

Given the above model, one can compute an exact analytic result for the motion generated by the Hamiltonian (33). The solution comprises the product of a pair of commuting matrices that act on the transverse variables (X, P_x, Y, P_y), plus a shift of the temporal variable T:

$$R_1 = \begin{bmatrix} \cos(\kappa s) & \frac{1}{\eta}\sin(\kappa s) & 0 & 0 \\ -\eta\sin(\kappa s) & \cos(\kappa s) & 0 & 0 \\ 0 & 0 & \cos(\kappa s) & \frac{1}{\eta}\sin(\kappa s) \\ 0 & 0 & -\eta\sin(\kappa s) & \cos(\kappa s) \end{bmatrix},$$
$$(34a)$$

$$R_2 = \begin{bmatrix} \cos(\kappa s) & 0 & \sin(\kappa s) & 0 \\ 0 & \cos(\kappa s) & 0 & \sin(\kappa s) \\ -\sin(\kappa s) & 0 & \cos(\kappa s) & 0 \\ 0 & -\sin(\kappa s) & 0 & \cos(\kappa s) \end{bmatrix}, \quad (34b)$$

$$T(s) = T^i - \left(\frac{1}{\beta_o} + P_t \right) \frac{s}{P_s^i} + \frac{s}{\beta_o}. \quad (34c)$$

In the above expressions, the parameter

$$\eta = \frac{eb_o}{2p_o} = \frac{b_o}{2(B\rho)_o}, \quad (35)$$

P_s^i denotes the initial value of the square root in the Hamiltonian (33), and $\kappa = \eta/P_s^i$.

The reader should note that the presence of P_s^i in (34) means that—despite superficial appearances—the transformation represented by (34) is actually nonlinear.

Later, when computing the spin precession vector $\vec{\Omega}$, we shall need to compute the direction vector $\hat{u}$. All magnets discussed previously can be described using a longitudinal vector potential. Then, as mentioned in Sec. II B, we can obtain $\hat{u}$ simply by normalizing the momentum vector $(P_x, P_y, P_s)^T$. For the solenoid, however, we must be sure to subtract the transverse components of the vector potential from the transverse canonical momenta to obtain the (scaled) kinetic momentum vector,

$$m\gamma \vec{V} = \left(P_x - \frac{e}{p_o} A_x, \ P_y - \frac{e}{p_o} A_y, \ P_s \right)^T. \quad (36)$$

From this we compute $\hat{u} = m\gamma \vec{V}/\|m\gamma \vec{V}\|$.

IV. SPIN INTEGRATION

High-quality orbital integration is a prerequisite for accurate spin integration, because it yields accurate values for the precession vector $\vec{\Omega}$. However, discretizing the spin motion leads to a separate source of error that can cause inaccurate spin precession, even in the context of perfect orbital integration. To address this difficulty, one might, for example, choose to adjust the number of spin slices independently of the number of orbital slices. In this section, we discuss this issue along with a description of our methods for tracking spin.

A. Piece-wise constant (PWC) spin precession

A commonly used integration strategy for spin treats the magnetic fields and velocity vectors in (6b) as constant throughout a slice of length Δs. One then integrates (6), the Thomas-BMT equation, to the form

$$\vec{S}(s + \Delta s) = \mathbf{R}(\vec{\omega}) \cdot \vec{S}(s), \quad (37a)$$

where $\mathbf{R}(\vec{\omega})$ denotes the 3×3 matrix that describes rotation about axis $\vec{\omega}$ by angle $|\vec{\omega}|$. We approximate this rotation vector as $\vec{\omega} = \Delta s \, \vec{\Omega}$, with $\vec{\Omega}$ given in (6b). Then to compute the spin rotation across a whole element, one simply multiplies the contributions from each slice. For four slices, say, one thus transports an initial spin $\vec{S}^i$ to a final spin $\vec{S}^f$ according to

$$\vec{S}^f = \mathbf{R}(\vec{\omega}_4) \cdot \mathbf{R}(\vec{\omega}_3) \cdot \mathbf{R}(\vec{\omega}_2) \cdot \mathbf{R}(\vec{\omega}_1) \cdot \vec{S}^i. \quad (37b)$$

There are two sources of errors in this approach to spin integration. First, errors in the orbital integration feed into the spin integration via errors in the fields and velocities needed in the right-hand side of the Thomas-BMT equation. Second, treating the rotation axis as a piece-wise constant introduces errors that arise from the noncommutativity of spin rotations around nonparallel axes.

When using drift-kick integrators for the orbital motion, the first source of error usually dominates. Then increasing the number of orbital slices to improve the orbital accuracy automatically diminishes the magnitude of the second type of error. We thus find that with drift-kick integrators, increasing the number of orbital slices is essential for accurate spin integration.

The situation changes when we use the more accurate bend-kick and matrix-kick integrators described above for dipoles and quadrupoles. These integrators allow us to take such large steps through both dipoles and quadrupoles that the lack of commutativity between consecutive spin rotations can now become an issue.

To speed the accumulation of spin rotations across a set of orbital slices, we use quaternions to represent rotations [55]. Compared to matrix multiplication, this saves a factor of about 2 in the required arithmetic operations. In addition, the spin rotation angles are very often sufficiently small that one may approximate the needed trigonometric functions by low-order polynomials. Every so often, we explicitly normalize the resulting quaternions to ensure that they remain unit quaternions [56].

B. Bézier interpolation

To treat the second source of error—arising from the fact that nonparallel rotations do not commute—we could do the same as for the first source of error: increase the number of orbital slices. That approach, however, seems unnecessarily expensive when we use very accurate, or exact, orbital integrators.

It turns out that the rotation vector $\vec{\Omega}$ varies across a slice in a sufficiently simple manner that we may interpolate it across a slice using piece-wise cubic polynomials, constructed as Bézier curves [57]. To use this technique, we compute, at the edge of each slice during the orbital integration, not only $\vec{\Omega}$, but also the derivatives of its three components [58]. With that information, we construct a
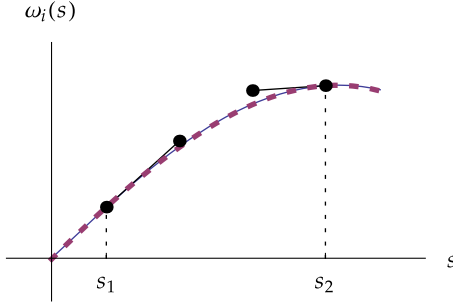
FIG. 1. Schematic of interpolating a function using a cubic Bézier curve. The function to be interpolated is shown as a thin solid curve, and the constructed interpolating Bézier curve is shown by the heavy dashed curve. The large dots indicate the control points derived from the value and slope of the function at s_1 and s_2.

cubic interpolating polynomial with the correct slope at the end points; see Fig. 1 for a schematic illustration. For the $\vec{\Omega}$ variations that occur in our simulations, this approach yields very accurate interpolating functions. Moreover, the low cost of evaluating this interpolating polynomial makes it possible to take many spin steps per orbital slice.

The PWC computation of the spin precession in (37b) will converge with a sufficiently large number of steps. If convergence requires an excessive number of steps, the use of cubic Bézier curves can speed up the computation. What the Bézier approach does *not* do is increase the rate at which the computation converges. For that we must look elsewhere, and that is the topic of the next section.

C. Romberg quadratures for spin precession

As we have seen, the spin precession across a beam line element is not, in reality, piece-wise constant; rather, within an element it varies smoothly with s. More significantly, the integrators based on (37) exhibit second-order convergence (see Sec. V B). This quadratic convergence suggests the use of some accelerating technique to cancel the errors. In particular, we have applied a Romberg approach [59,60] to spin integration, and, as we show later in this paper, the improvement is dramatic.

Instead of computing the rotation vectors at the middle of each slice, we now, for this new approach, compute them at the edges of each slice. We then accumulate the spin precession as in (37). For the first and last edges, however, we use a half-step; i.e., we replace Δs by $\Delta s/2$. This is akin to using the trapezoidal rule for integration [59,61]. We do this using N slices, with N a multiple of some power of 2.

During the orbital integration, we record the value of $\vec{\Omega}$ at the edge of each slice. Then by keeping every other $\vec{\Omega}$, or every fourth, or every eighth, etc., we can approximate the net spin precession using a range of additional step sizes that are related to the original by powers of 2. In other words, we compute for these different step sizes a net quaternion that represents the spin precession of a given

particle across an element. We obtain thus the sequence of quaternions,

$$Q(h), Q(h/2), Q(h/2^2), Q(h/2^3), \ldots, \quad (38)$$

where h denotes the coarsest step size.

For example, when crossing an element using eight slices, we will compute a total of nine $\vec{\Omega}$s: $\vec{\Omega}_0$ at the beginning of the first slice, $\vec{\Omega}_1$ at the end of the first slice, on through $\vec{\Omega}_8$ at the end of the last slice. Now, assuming each slice has length Δs, we multiply the first and last of these by $\Delta s/2$, and all the others by Δs, to obtain the sequence

$$\frac{1}{2}\Delta s\,\vec{\Omega}_0, \Delta s\,\vec{\Omega}_1, \Delta s\,\vec{\Omega}_2, \ldots, \frac{1}{2}\Delta s\,\vec{\Omega}_8.$$

After constructing the quaternions that correspond to each of these nine precession vectors, we multiply them together to obtain a net quaternion that represents one approximation to the exact spin precession across this element. This is one of the entries in the sequence (38) above. To obtain the entry to its immediate left, we drop every other $\vec{\Omega}$ and compensate by doubling the step size; the sequence

$$\Delta s\,\vec{\Omega}_0, 2\Delta s\,\vec{\Omega}_2, \ldots, \Delta s\,\vec{\Omega}_8$$

thus yields another—the next coarser—approximation to the spin precession across this element, and so on.

We then compute the *Romberg limit*: First define by

$$Q_{0k} = Q(h/2^k) \quad (39a)$$

the sequence of approximations in (38). Then use the rule

$$Q_{j+1,k} = \frac{4^{j+1}Q_{j,k} - Q_{j,k-1}}{4^{j+1} - 1} \quad (39b)$$

to construct the Romberg table:

$$
\begin{array}{llll}
Q_{00} & & & \\
Q_{01} & Q_{11} & & \\
Q_{02} & Q_{12} & Q_{22} & \\
Q_{03} & Q_{13} & Q_{23} & Q_{33}.
\end{array} \quad (39c)
$$

This table may have more or fewer rows than indicated here, but the number at the bottom right is the Romberg limit of the initial data given in the first column. We normalize the resulting quaternion at the end of each element.

When integrating a well-behaved function over a finite interval, the trapezoidal rule together with the Romberg limit performs remarkably well with modest computational effort. Its efficiency, however, derives from the structure of the error term seen in the Euler-Maclaurin summation formula [60,61], and the manner in which the Romberg algorithm

cancels those errors. Here, on the other hand, we have a *product* of quaternions, and hence no *a priori* reason to suspect that the above will actually work. We tried it for a lark.

D. Solenoid fringe

In the thin-pancake model discussed earlier for the solenoid fringe, there exists a well-defined limit of the fringe length ε times the precession vector $\vec{\Omega}$, as $\varepsilon \to 0$. In that limit, a spin crossing a solenoid fringe experiences a net rotation given by the vector

$$\vec{\omega} = \pm \frac{b_0/2}{(B\rho)_{\mathrm{o}} P_s} \left[(1 + G\gamma)(X, Y, 0) \right.$$
$$\left. - G(\gamma - 1) \frac{XP_x + YP_y}{P^2} m\gamma \vec{V} \right], \qquad (40)$$

where we choose the leading plus or minus for entrance or exit, respectively. Here the kinetic momentum vector $m\gamma\vec{V}$ is as given in (36), *except* that we must insert a factor of $\frac{1}{2}$ in front of A_x and A_y. In a similar manner, P_s denotes the square root in (33), but with the same factor of $\frac{1}{2}$ inserted before A_x and A_y. These factors of $\frac{1}{2}$ enter because in going to the limit of a zero-length fringe, we evaluate the field strength in the middle of that fringe region.

E. Dipole and multipole fringes

A particle's spin also experiences a kick when crossing the fringe of a multipole magnet. This spin kick arises predominantly from the longitudinal field component in the fringe. In the limit of a thin fringe, and to lowest order in the dynamical variables, the spin kick caused by the fringe of a normal (i.e., not skew) $2m$-pole magnet of strength b_m is given by the rotation vector

$$\vec{\omega}_m = -\frac{\pm b_m}{(B\rho)_{\mathrm{o}}} (1 + G) \frac{r^m}{m} \sin(m\phi)\hat{z}, \qquad (41)$$

where we convert $r^m \sin(m\phi)$ to a function of X and Y using $(X + iY)^m = r^m e^{im\phi}$. For a skew multipole, in (41) replace sin by cos, and b_m by a_m.

The above result (41) is appropriate for beams that cross a fringe roughly normal to the entrance or exit face of the magnet. When entering, or exiting, a rectangular bend, however, a particle typically sees the longitudinal field B_z as having a significant component transverse to its velocity. As long as P_x is not too large, one may account for this effect by simply adding an extra term to (41) to obtain the rotation vector

$$\vec{\omega}_{\mathrm{rb}} = \vec{\omega}_1 + \frac{\pm b_1}{(B\rho)_{\mathrm{o}}} G(\gamma - 1) P_x \hat{x}, \qquad (42)$$

which describes the spin kick experienced by a particle crossing the fringe of a rectangular bend. Note the factor of

$G\gamma$ in the second term: it is one good reason for not using rectangular bends in machines designed for polarized beams, particularly at high energy.

If one requires a higher order computation of the spin kick across a magnet fringe, it is possible to apply a procedure analogous to that used for orbital kicks across fringe fields in [41].

V. PERFORMANCE OF INTEGRATORS

In this section, we examine the performance of the orbital and spin integrators described in the previous two sections. Our principal focus is on the integrators for spin. We examine the orbital integrators—much more briefly—with the goal of understanding their impact on spin integration.

For the orbital integrators, we examine their performance in a particular context: that of Brookhaven's Relativistic Heavy Ion Collider (RHIC) [2] operating at approximately 250 GeV, with the optics settings used before the beta squeeze. For the spin integrators, we examine their single-element performance, as well as their performance in the context of spin tracking for RHIC. In this context, we note that RHIC has two full Siberian snakes on opposite sides of the ring to flip the spins, with the snake angles set to $\pm 45°$. These settings mean that for a perfectly aligned RHIC, the design orbit will have a spin tune of exactly $\frac{1}{2}$ [3,4]. Sextupoles are adjusted to set the chromaticities to $(\xi_x, \xi_y) = (0.70, 0.74)$, and the spin rotators near the interaction points are turned off. The latter are needed only after the beta squeeze, when the counterrotating RHIC beams are brought to collision, and the experiments require longitudinal polarization at the interaction point. For our simulations, then, the equilibrium polarization is roughly vertical throughout the ring.

A. Orbit integrators

For the simulations discussed here, we read a RHIC lattice description from a file in SXF format [62]. Which integrator to use for which elements is set in a separate ".apdf" file (for *Accelerator Propagator Description Format*). When referring to the integrators, we use the acronym BK/MK [63], or sometimes just "bend-kick," to mean the use of (10) for sector bends, together with (24) and (25) for quadrupoles. The Siberian snakes were modeled as thin elements: they are transparent to the orbital motion, and they act on the spin according to snake angles defined in an input file. In the simulations described here, we use a base number n_b of slices for all dipoles and most quadrupoles. Then for the strongest quadrupoles, those in the interaction region, we use $4n_b$ slices.

1. Poincaré sections

To give a sense of how well the drift-kick and BK/MK integrators reproduce gross features of the orbital motion,

we show Poincaré sections of the X-P_x plane; see Fig. 2. The range of amplitudes shown is representative of beams at this azimuth in RHIC—the interaction point, before the beta squeeze is applied to the optics.

Figure 2 indicates that the drift-kick and BK/MK integrators produce the same gross features of the orbital motion. There are, however, differences in the details. In particular, the drift-kick traces with $n_b = 4$ seem more elliptical, while the remaining traces seem more ovoid. The trace widths also seem more regular for drift-kick integration with $n_b = 4$, as compared to all the others. Differences between the two BK/MK integrators, with $n_b = 1$ and $n_b = 16$, are not visible with the bare eye. Differences between (b), (c), and (d) are much smaller than those between (a), on the one hand, and (c) and (d), on the other.

2. Orbital spectra and tune fitting

In Fig. 3 we show the spectrum of X obtained by taking a discrete Fourier transform of the X coordinates obtained over the course of $T = 1000$ turns,

$$S_X(\nu) = \left| \frac{1}{T} \sum_{t=0}^{T} e^{-i2\pi t\nu} X(t) \right|^2. \tag{43}$$

The spectra were averaged over a bunch containing $N = 1024$ particles with amplitudes typical for RHIC. The tunes marked in the figure (see top edge) are the design values for RHIC at this energy (215.735 GeV kinetic): $\nu_x = 28.6835$, $\nu_y = 29.6742$. Note that we see peaks not only at ν_x and ν_y, but also at $2\nu_x$ and $2\nu_y$—evidence of significant non-linearity in the orbital motion.
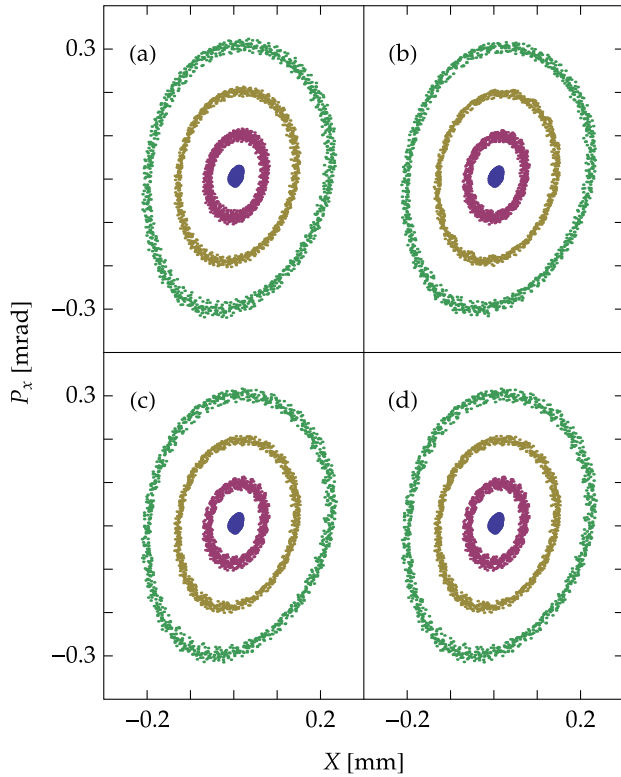
As shown in Fig. 3, one can obtain a close match to the design orbital tunes using either drift-kick or BK/MK integration. However, in the case of drift-kick integration, one must adjust the strengths of the main quadrupoles by an amount that depends on the number of orbital slices used. No such adjustment is required when using BK/MK integration.

In the context of purely orbital tracking, adjusting the quadrupole strengths to fit the desired tunes is a sensible means of ensuring the correct linear behavior in accelerator simulations. In the context of spin tracking, however, such adjustments might be problematic, because changing a quadrupole's strength changes the amount of spin precession a particle experiences as it crosses that element. In other words, fitting the orbital tune can perturb the integrated spin precession angle of a given particle—particularly at high energy, i.e., at large values of the naive spin tune, $G\gamma$.

We can estimate the importance of this effect as follows. On crossing a quadrupole, a spin experiences a net precession with magnitude given approximately by $\omega = -(1 + G\gamma)Ab_2L/(B\rho)_o$, where A denotes the particle amplitude, and L the quadrupole length. For the RHIC lattices we studied, the tune fitting required a relative change in b_2 that varied greatly with the step size used for crossing the quadrupoles. For finely sliced lattices, the relative change in b_2 (from the value used for BK/MK integration) could be as small as a few times 10^{-5}. At RHIC energies, this adjustment yields a sub-μrad change in the spin precession, and perhaps a few μrads over a full turn. As we shall see later, when one uses drift-kick orbital integration, this variation is negligible compared to other

FIG. 2. Poincaré sections for four sample particles using drift-kick integration with (a) $n_b = 4$ and (b) $n_b = 16$, and using BK/MK integration with (c) $n_b = 1$ and (d) $n_b = 16$.
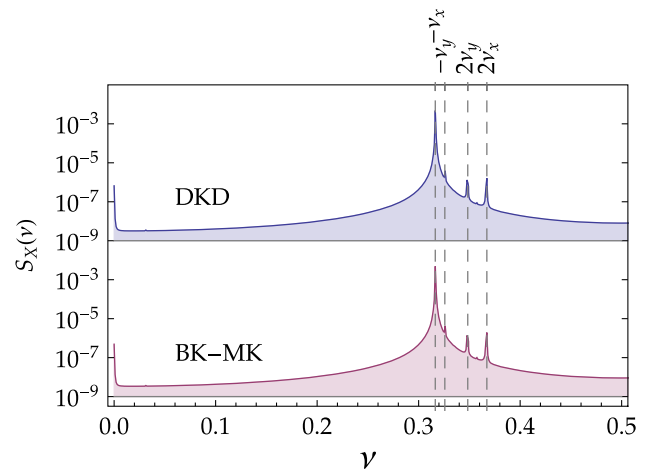

FIG. 3. Spectrum of X obtained using (a) drift-kick integration with $n_b = 8$, or (b) BK/MK integration with $n_b = 1$. The indicated tunes are all taken modulo 1.
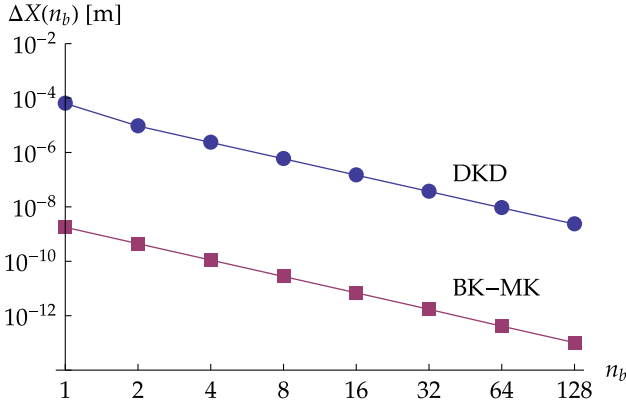
FIG. 4.   One-turn errors in X versus n_b for drift-kick integration (upper curve), and BK/MK integration (lower curve). The errors were computed by comparing with a data set produced using BK/MK integration and $n_b = 256$.

sources of error. On the other hand, for very coarse slicing, the relative change in b_2 can exceed 0.1%, which leads to a small, though noticeable, change in the simulated spin precession angle.

3. Single-turn errors

For accurate spin tracking, it does not suffice to integrate the orbital motion in a manner that merely reproduces global properties of the phase-space distribution. Errors in particle orbits introduce errors in the spin rotations, possibly changing the character of the simulated spin dynamics. As a consequence, one must compute accurately the trajectories of individual particles.

To gain an understanding of the orbital errors on a per-particle basis, we look at the rms deviation between computed particle orbits and a reference solution for each particle. For the X coordinate after t turns, we thus compute

$$\Delta X(t; n_b) = \sqrt{\frac{1}{N} \sum_{j=1}^{N} [X_j(t; n_b) - X_j^{(\mathrm{ref})}(t)]^2}. \qquad (44)$$

The sum is over the particles in the bunch. The reference solution is computed with $n_b = 128$ using the BK/MK integrators, i.e., $X_j^{(\mathrm{ref})}(t) = X_j(t; 128)$.

Figure 4 shows the absolute error $\Delta X(1; n_b)$ after one turn for a RHIC beam at approximately 216 GeV with a 95% horizontal emittance ε_x of 15π mm mrad. The slope -2 on the log-log scale shows that both drift-kick and BK/MK integrators exhibit second-order convergence in s. The drift-kick integrator starts out with a relative error of order one (cf. Fig. 2), indicating that with this very large step size, the phases of the orbital motion are completely off: the errors are as large as they can be with a symplectic integrator. At $n_b = 64$, the relative errors have decreased to about a part in 10^4. On the other hand, the BK/MK integrators exhibit relative errors which start out below a

part in 10^6 and decrease until they are below about a part in 10^9.

Figure 5 also shows one-turn errors in X, but here as a function of particle emittance. These results, which are consistent with those of Fig. 4, tell us that the accuracies of the different integrators have roughly the same amplitude dependence. To reduce the noise level in this graphic, we have averaged the results within nonoverlapping bins that each contained 20 particles in a narrow slice of emittance. The bands indicate one standard deviation above and below the computed average.

4. Evolution of orbital errors over many turns

Figure 6 shows the errors in the X coordinate as a function of the number of turns for drift-kick integration and BK/MK integration. If we do not adjust the quadrupole strengths so as to fit the desired tunes, then the drift-kick integration errors oscillate periodically because of the tune errors. We see that behavior in the uppermost (tan) curve,
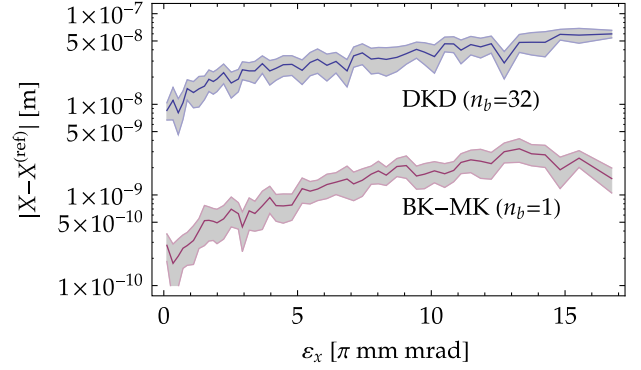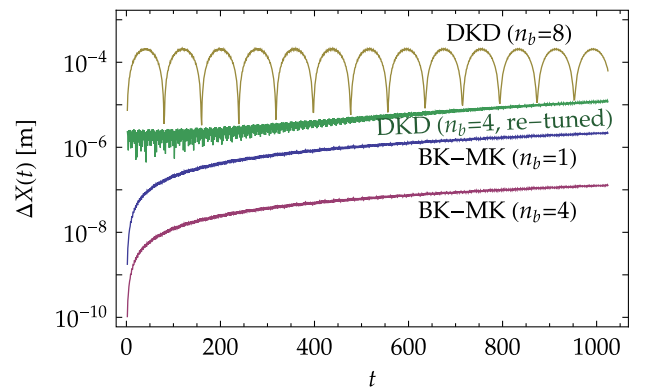

FIG. 5.   One-turn errors in X versus amplitude for several different integrators: drift-kick with $n_b = 32$ (upper curve), and BK/MK with $n_b = 1$ (lower curve).

FIG. 6.   Evolution of the orbital error in X as a function of turn number t. The four curves correspond to different methods of integration: from top to bottom they are drift-kick with $n_b = 8$ and no tune fit (tan), drift-kick after tune fitting with $n_b = 4$ (green), and BK/MK with $n_b = 1$ (purple), and 4 (maroon). For the reference solution, we used BK/MK with $n_b = 16$.

for drift-kick integration with $n_b = 8$. After fitting the tunes, we note that drift-kick integration errors with $n_b = 4$ now track the minima of the tan oscillatory errors.

While the errors observed are much smaller for BK/MK integration—even with $n_b = 1$—they always exhibit the same qualitative behavior: After a few hundred turns, even when using BK/MK integration, the errors grow approximately linearly because of unavoidable tiny differences in the tunes. We believe this is not a significant issue.

B. Spin integrators

Here we describe the accuracy of our spin integrators. First, we discuss the errors seen when tracking across various individual beam line elements, where we focus on the impact of both step-size and Romberg iterations. We then discuss, more briefly, the spin integration errors seen over the course of a single turn and many turns.

1. Single-element errors

The graphics in Fig. 7 show the spin integration errors we computed when tracking a particular particle [64] across several different beam line elements. From top to bottom, those elements are a quadrupole, a sector bend, a solenoid, and another solenoid (but with a very low energy proton, such that the total spin rotation is comparable to 90°). In each of those graphics, the solid curve labeled PWC corresponds to piece-wise constant spin integration, where for the orbital motion we used the integrators given in Secs. III A, III C, and III G. In all cases the slope −2 reveals second-order convergence for PWC spin integration. The remaining curves correspond to k Romberg iterations (i.e., Q_{kk}) applied to the PWC data obtained for the given number of slices. Since we do not know the result of exact spin integration, we have here estimated the spin integration error as the absolute difference between the result $\vec{\mathcal{S}}^N$ obtained using N slices and what we considered our "best" result, $\vec{\mathcal{S}}_{\text{best}}$.

In Fig. 7, we have used for $\vec{\mathcal{S}}_{\text{best}}$ the results we obtained using 256 slices and three Romberg iterations. If we instead use for $\vec{\mathcal{S}}_{\text{best}}$ our most finely sliced PWC result, then small details in these graphics change, but the overall implication holds—that one or more iterations of the Romberg procedure can dramatically reduce the errors made by PWC spin integration. In the case of the quadrupole integrated using four slices, we see that two Romberg iterations yield a four-decade reduction in the error—to a level that requires some four-hundred slices using just PWC spin integration. For the solenoid, a similar statement holds in the case of a low-energy proton; and in the high-energy case, one Romberg iteration suffices to reduce the spin integration error to the level of round-off. For the sector bend, we see a less dramatic absolute reduction in the error; but even there, between 8 and 64 slices, the slope −4 on the $k = 1$ curve tells us that one Romberg iteration converts second-order PWC integration to fourth order.
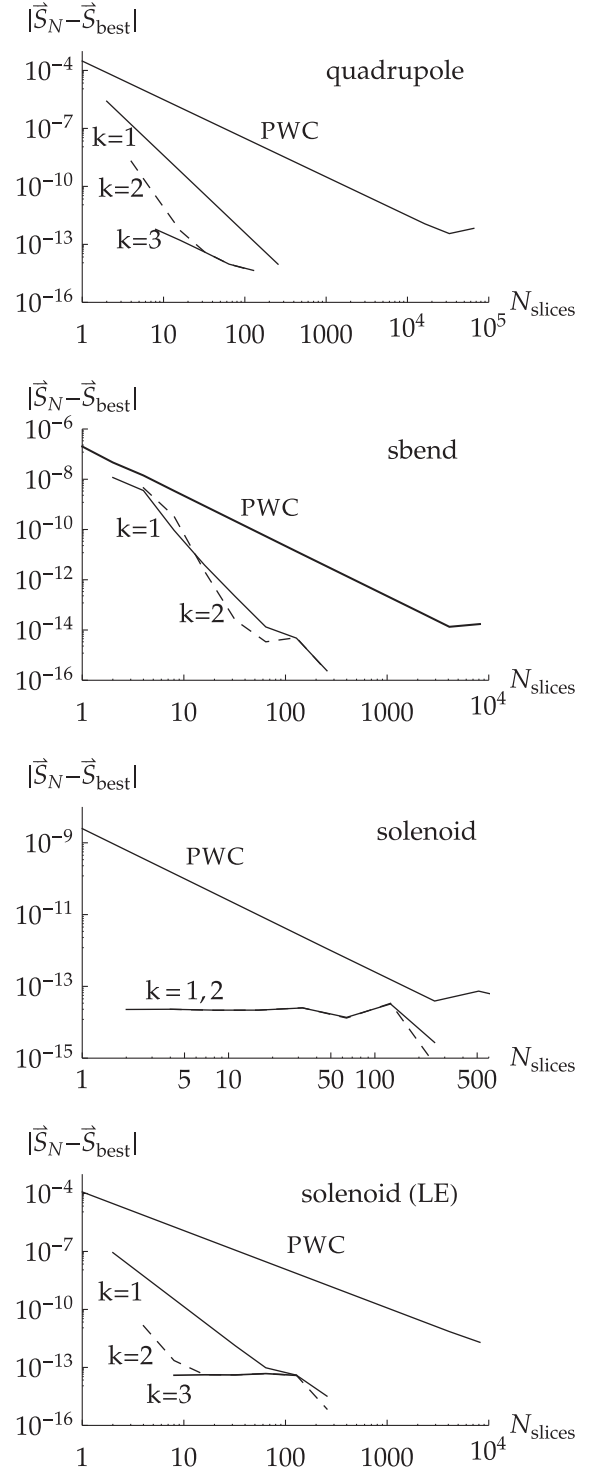

FIG. 7. Absolute error in the computed spin versus the number of slices for a proton crossing several different beam line elements. The upper three graphics show results for a 200 GeV p^+ crossing a RHIC IR quad (top), a RHIC arc dipole (second), and a 2.1 m, 1.3 T solenoid (third). The bottom graphic shows the result for a 25 GeV p^+ traversing the same solenoid. Each curve corresponds to a different method of integrating the spin: PWC for piece-wise constant integration; the k values indicate the number of Romberg iterations.

The graphics in Fig. 7 illustrate the scaling of spin integration errors for one particular particle, and the impact of Romberg iterations on PWC spin integration. Were we just lucky, or do we see similar behavior across the relevant phase space? The next several figures address this question, illustrating the effect of Romberg quadratures on spin integration for particles covering a range of initial conditions in phase space.

For the sector bend, we examine spin integration errors across a RHIC dipole: a 9.45 m long magnet set to bend 200 GeV protons 2.225°, so with magnetic field strength $B \approx 2.75$ T. As a proxy for exact tracking, we settled on, after considerable testing, one Romberg iteration applied to 128 slices of PWC spin integration. Furthermore, because a magnetostatic dipole is translationally invariant in both Y and T, we examine how the spin integration errors vary with respect to just X, P_x, P_y, and P_t.

In Fig. 8, we show the absolute error in the computed spin using PWC spin integration with either four or eight slices (upper and lower rows, respectively). Figure 9 shows the corresponding results obtained after we apply one Romberg iteration. To help the reader make a proper comparison, we note further details in the captions of those two figures.

A comparison of Figs. 8 and 9 suggests that applying Romberg quadratures to spin integration across a sector

bend does indeed reduce the error throughout phase space. But we are also struck by some of the features seen in Fig. 8: the pronounced valley that runs diagonally across the four graphics on the left; and the much flatter and somewhat larger errors in the presence of an energy deviation, seen in the two graphics on the right.

To gain some understanding of these features, we examined a much larger range of phase space; a range that is, in fact unphysically large for an arc dipole in RHIC. Figures 10 and 11 show the results. Each of the 100 small graphics in those two figures covers the same range of horizontal phase space: ± 5 cm in X, and ± 6 mrad in P_x. They are laid out in four matrices of 25 graphics each, where the columns correspond to five different values of P_t^i, evenly spaced in the range $\pm 2 \times 10^{-4}$; and the rows correspond to five different values of P_y^i, evenly spaced in the range ± 0.2 mrad. (See the labels at the top and to the right of each matrix of graphics.) To obtain the graphics in Fig. 10, we used four slices of PWC spin integration for the upper set of 25 graphics, and four slices plus one Romberg iteration for the lower set. We obtained the data shown in Fig. 11 using eight slices. In all the graphics, the gray-scale covers the same (logarithmic) range of errors: from black at 10^{-15} to white at $10^{-7.5}$.

The graphics shown in Figs. 10 and 11 reiterate the message that applying a Romberg step to the result of
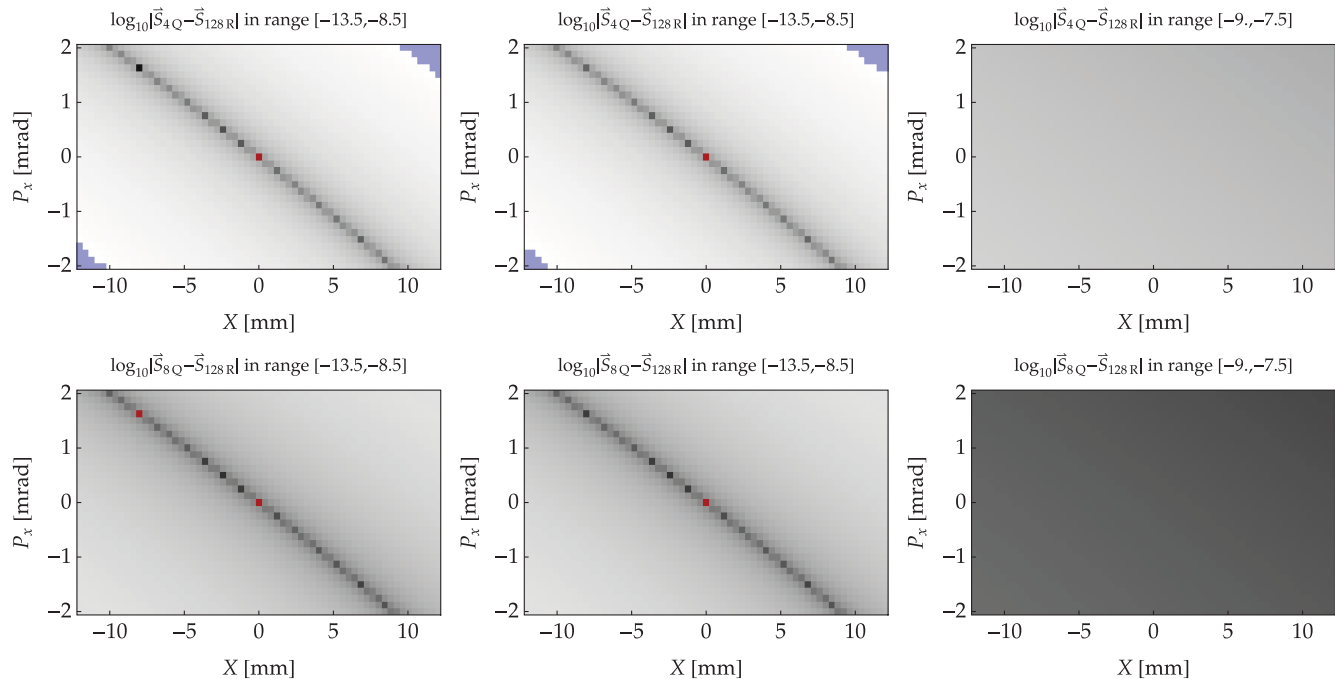

FIG. 8. Absolute error in the computed spin across a RHIC dipole for a range of initial conditions in orbital phase space. Darker (lighter) grays represent smaller (larger) errors. Points outside the range indicated above each graphic are colored red if below, and blue if above. The error range covers the same five decades in the four graphics to the left. In the right-hand two graphics, the error range is again the same, but covers just a decade and a half that is mostly above the range in the other graphics. The upper and lower rows respectively show results obtained using PWC spin integration with four and eight slices. The three columns correspond to different values for P_y^i and P_t^i: $P_y^i = 0$, $P_t^i = 0$ (left); $P_y^i = 20$ μrad, $P_t^i = 0$ (middle); $P_y^i = 20$ μrad, $P_t^i \approx 0.35 \times 10^{-3}$ (right).
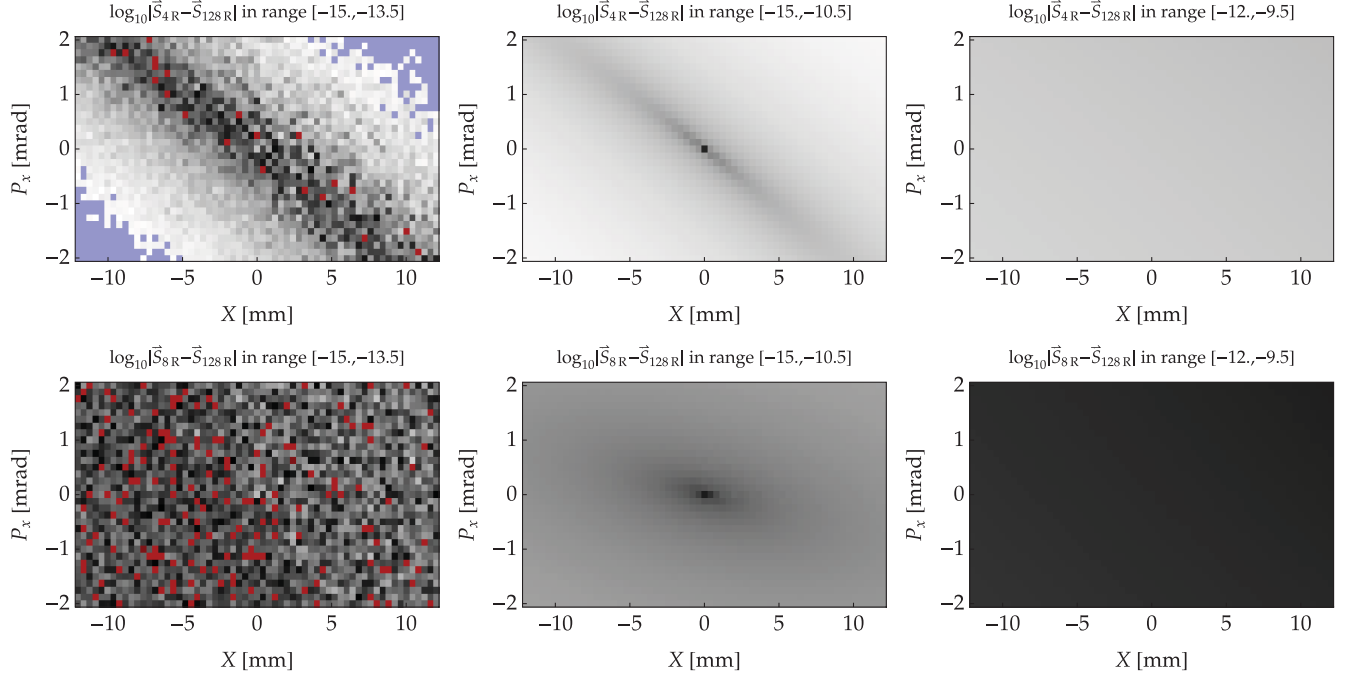
FIG. 9. Absolute error in the computed spin across a RHIC dipole for a range of initial conditions in orbital phase space. Each of the graphics in this figure shows results for the same orbital initial conditions and the same number of slices as the corresponding graphic in Fig. 8. The differences are that the integration here includes one Romberg step, and the error ranges are significantly lower.

PWC spin integration can reduce the error throughout the phase space of orbital initial conditions. In addition, however, we draw the reader's attention to the yellow curve, overlaid on each graphic, which clearly follows the deep valley seen in those graphics for which no Romberg step was applied. In each graphic, this curve is the locus of points (X^i, P_x^i) for which the value of P_x at the magnet exit equals the value P_x^i at the magnet entrance. This value is given by

$$P_x^{\text{sym}}(X^i) = \cos\left(\frac{hL}{2}\right)\left[\sqrt{1 + \frac{2}{\beta_o}P_t + P_t^2 - P_y^2 - (1 + hX^i)^2\cos^2\left(\frac{hL}{2}\right)} - (1 + hX^i)\sin\left(\frac{hL}{2}\right)\right]. \tag{45}$$

Note that the dependence of P_x^{sym} on P_y and P_t occurs only under the square root, and that there it is linear in P_t and quadratic in P_y. This difference (together with the fact that P_y and P_t denote small quantities) explains why the yellow curves in Figs. 10 and 11 show little dependence on P_y, but very sensitive dependence on P_t.

Earlier, when examining Fig. 8, we noted that the two right-hand graphics have much flatter and somewhat larger error profiles than appear in the left-hand four graphics. Now, on comparing with the graphics in Fig. 10, we see what happened: the valley simply moved "off stage" as we increased the energy deviation.

For the quadrupole, we examine spin integration errors across a RHIC interaction region (IR) quadrupole: in this case a 1.83 m long magnet set to focus 200 GeV protons with a magnetic quadrupole gradient of $b_2 = 70.0$ Tm^{-1}. As a proxy for exact tracking, we settled on, again after much testing, three Romberg iterations applied to 128 slices of PWC spin integration.

A coarse examination over the range of initial conditions ± 20 mm in X and Y, ± 20 μrad in P_x and P_y, and $\pm 0.35 \times 10^{-3}$ in P_t indicate that (at least within this phase-space domain) spin integration errors change little with respect to variations in the transverse momenta P_x and P_y, or the relative energy deviation P_t. As a consequence, here we show graphics of the spin integration errors for this element only in the (X, Y) plane, with $P_x^i = P_y^i = P_t^i = 0$.

In the case of the sector bend, the spin integration error does not appear to depend on the initial spin orientation. For the quadrupole, however, the accuracy of spin integration very definitely depends on the initial spin, so we have included that dependence in the results shown here.

In Fig. 12, we show the absolute error in the computed spin across a RHIC IR quadrupole. The first two rows show results obtained with PWC spin integration using either four slices (first row), or eight slices (second). The results shown in the lower two rows correspond to applying our Romberg procedure to the first two rows: four slices plus
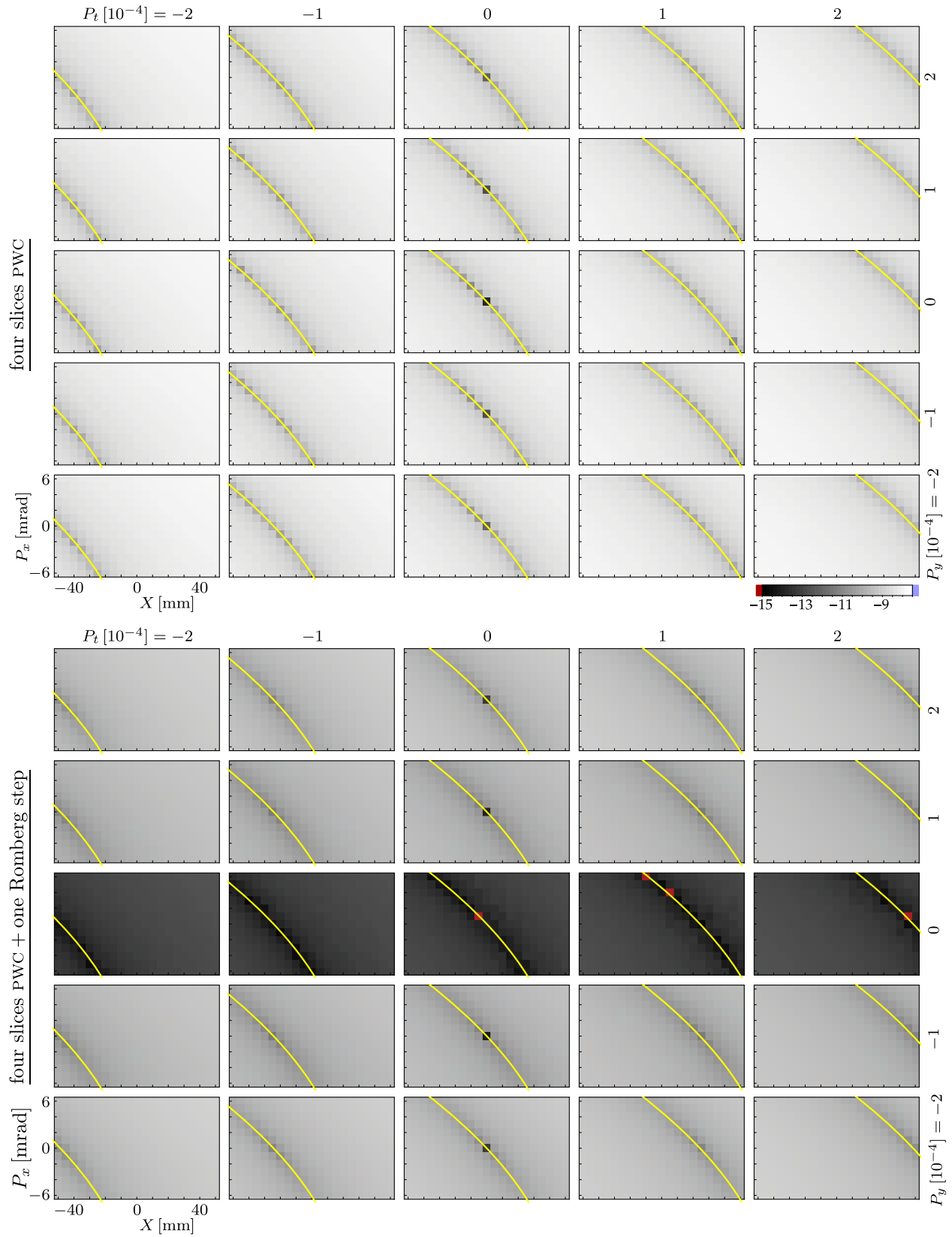
FIG. 10. Absolute error in the computed spin using four slices across a RHIC dipole. Each small graphic covers a range of initial conditions in horizontal phase space (see labels at lower left). The rows and columns correspond respectively to different initial values for P_y and P_t (see labels along right and top edges). For the lower set of graphics, we added one Romberg iteration.
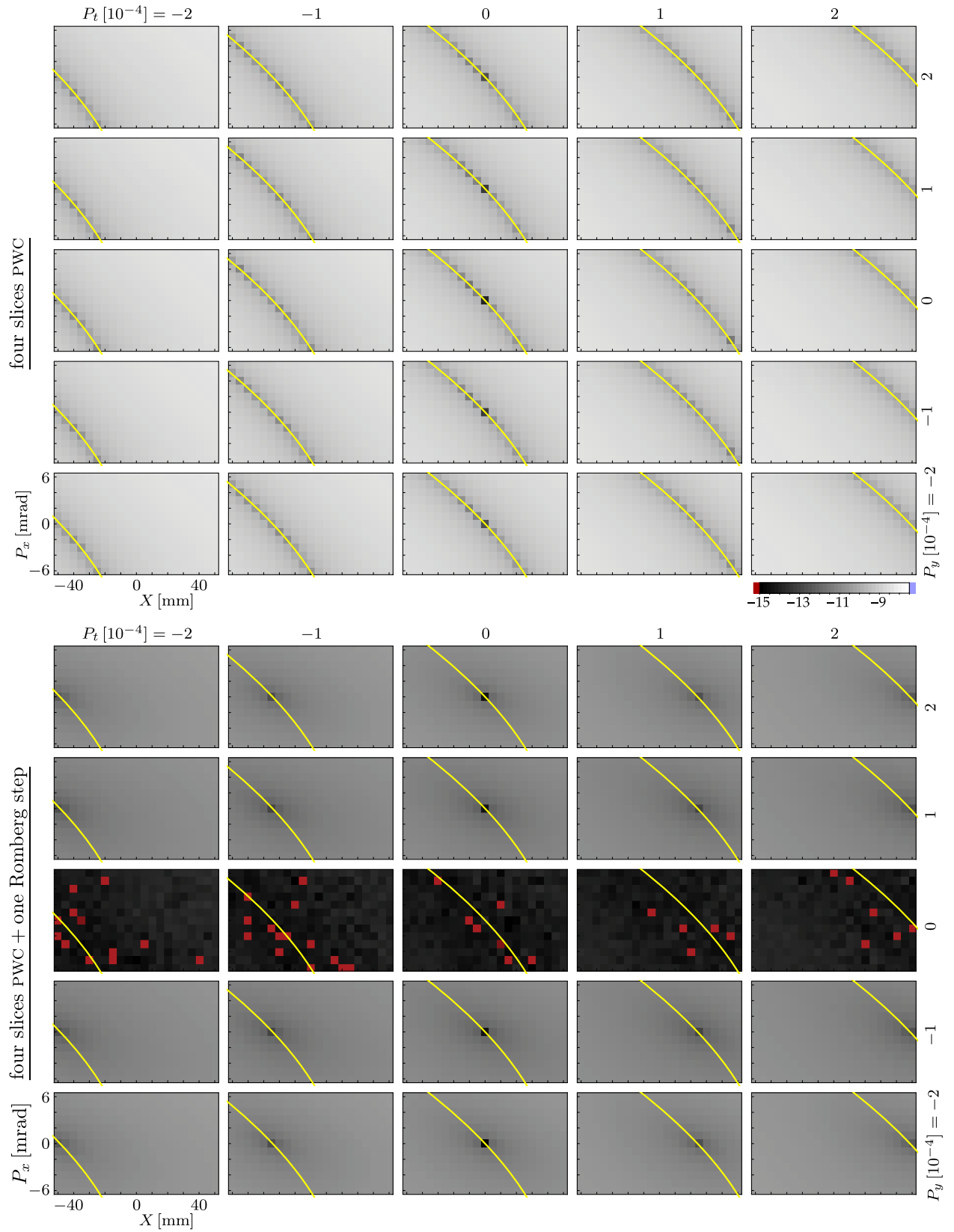
FIG. 11.   Absolute error in the computed spin using eight slices across a RHIC dipole for a range of initial conditions. See also the caption of Fig. 10. The two sets of 25 graphics correspond to PWC spin integration (upper set) and PWC plus one Romberg iteration (lower set).
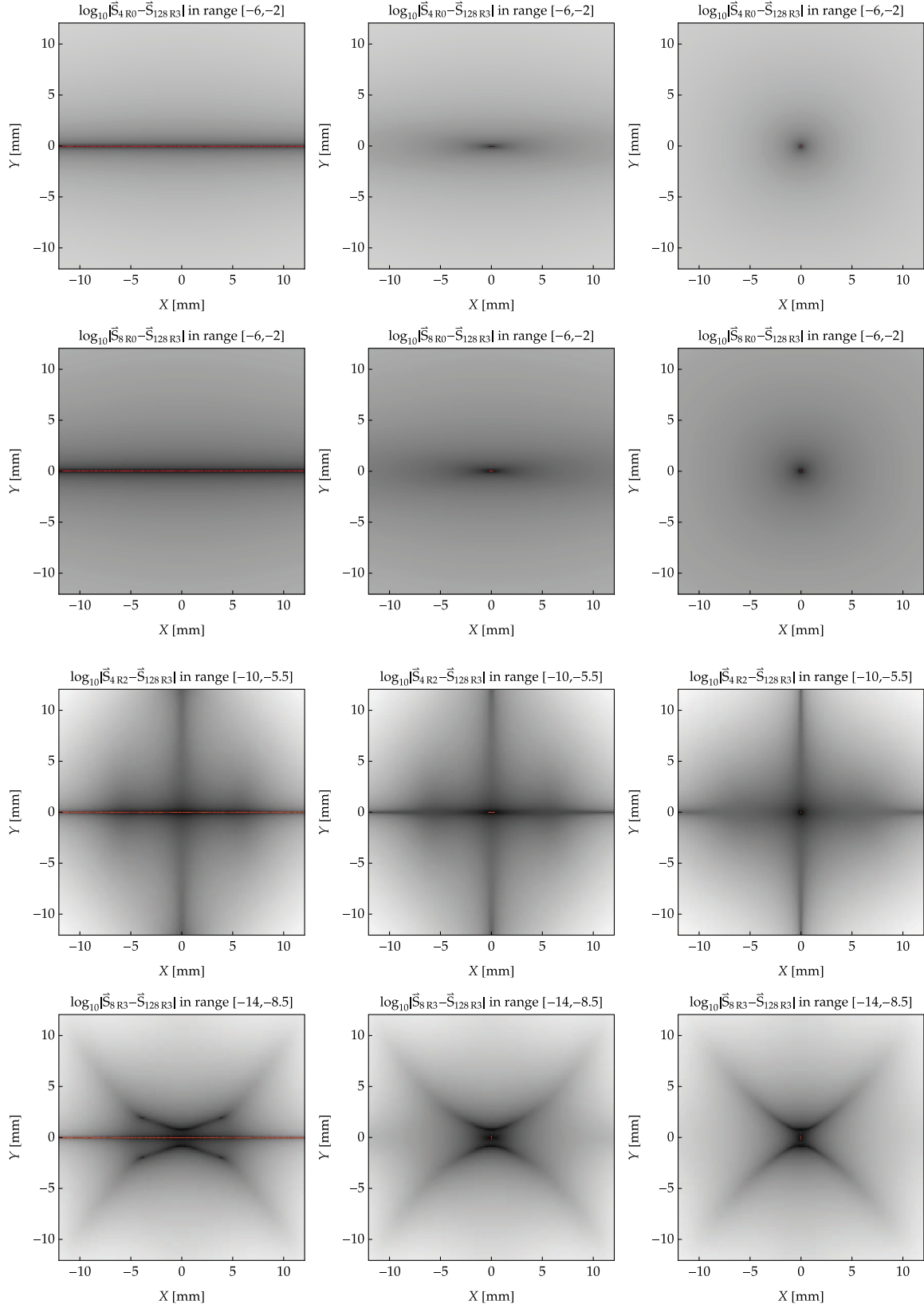
FIG. 12.   Absolute error in the computed spin across a RHIC IR quadrupole. The rows, from top to bottom, show results obtained using four slices, eight slices, four slices plus two Romberg steps, and eight slices plus three Romberg steps. The columns, from left to right, show results averaged over spins covering an opening angle about the vertical of 0°, 20°, and 180°.

two Romberg steps (third row), or eight slices plus three Romberg steps (bottom row). As indicated above each graphic, the gray-scale in the first two rows covers the same (logarithmic) range of errors: from black at 10^{-6} to white at 10^{-2}. In the third row, the range goes from 10^{-10} to $10^{-5.5}$, which means most of the errors in the third row (all except small triangles in the corners) lie below the errors in the upper two rows. In the last row, the range of 10^{-14} to $10^{-8.5}$ means the errors here lie entirely below those of the upper two rows; and they lie below most of the errors in the third row.

The columns in Fig. 12 correspond to averaging over various sets of initial spins. We generated a quasirandom distribution of spins (i.e., points on the unit sphere [65]) with opening angle α about the vertical: $\alpha = 0°$ (left column), 20° (middle), and 180° (right).

In the case of a vertical spin ($\alpha = 0°$), the errors vanish for orbital motion in the midplane because the vertical spin remains parallel to the vertical field everywhere along those orbits and therefore does not precess. What about other particles, those launched still with a vertical initial spin, but away from the midplane? For particles traversing a RHIC IR quad, the $\vec{\Omega}$ of (6b) lies predominantly along the quadrupole magnetic field, $(b_2 Y, b_2 X, 0)$. Since the component of $\vec{\Omega}$ orthogonal to the spin is the component that rotates the spin, it follows that an initially vertical spin should be most sensitive to variations in Y, and rather insensitive to variations in X. This suggests why the two graphics in the upper left of Fig. 12—those for $\alpha = 0°$ and no Romberg iterations—show essentially no variation with respect to X^i. Moreover, the fact that the orbital variation in Y is (to first order) proportional to Y^i suggests why the errors in those same two graphics increase away from the horizontal midplane.

As we allow the distribution of initial spins to open up ($\alpha > 0°$), the distribution of errors (before the application of any Romberg iterations) quickly becomes more symmetric about the quadrupole axis. See the four graphics in the upper right of Fig. 12. After we apply one or more Romberg steps, the quadrupole spin-integration errors become smaller and develop additional structure across the (X, Y) plane. See the bottom two rows of Fig. 12. We do not yet understand the origin of this structure.

For the solenoid, we examine spin integration errors across a 1.3 T, 2.1 m long magnet. For the 200 GeV protons we used in previous elements, we find that four slices together with one or two levels of Romberg quadratures suffice to reduce errors to the level of round-off. In this case, however, the particle experiences very little spin rotation—of order 10 mrad. We therefore chose to look at the spin integration errors for a *much* lower energy proton, 1.2 GeV, traversing this same magnet. In this case the spin rotation is of order 0.6 rad. As a proxy for exact tracking, we again settled on three Romberg iterations applied to 128 slices of PWC spin integration.

Given the homogeneous nature of a solenoid's body magnetic field in our thin-fringe approximation, we expect spin integration errors across this element to exhibit very little dependence on X^i, Y^i, or P_t^i. A coarse examination over a range of initial conditions in phase space indicates that this is indeed the case. As a consequence, here we show graphics of the spin integration errors for this element only in the (P_x, P_y) plane.

In Fig. 13, we show the absolute error in the computed spin for a 1.2 GeV proton traversing the solenoid described above. The top graphic shows the result obtained using eight slices of PWC spin integration. The two graphics below show results obtained using four slices plus two Romberg steps (middle), and eight slices plus three Romberg steps (bottom). As indicated by the ranges noted above each graphic, the error ranges in these three graphics do not overlap. Indeed, the error range drops by several decades with each step down the page.

2. Single-turn errors

In this section, we examine spin integration errors after a full turn in the same RHIC lattice used for the orbital studies in Sec. V A. See Figs. 14 and 15. As our measure of the error, we use the mean difference between the spin computed using a given set of numerical algorithms, and a reference solution that we use as a proxy for the exact result. The mean is taken over an ensemble of spins in a beam. We compute thus, at turn t,

$$\Delta \vec{\mathcal{S}}(t) = \frac{1}{N} \sum_{j=1}^{N} |\vec{\mathcal{S}}_j(t) - \vec{\mathcal{S}}_j^{(\mathrm{ref})}(t)|. \tag{46}$$

For the results shown here, we computed the reference solution using BK/MK integration with $n_b = 256$ and three Romberg iterations.

Figure 14 shows the absolute spin error as a function of the number of spin slices, here chosen identical to the number of orbital slices. No Romberg steps were applied, except when computing the reference solution. As a consequence, differences seen here between drift-kick results (upper two curves) and BK/MK results (lower curve) derive solely from differences in the orbital data passed to the Thomas-BMT equation.

At small numbers of orbital slices, spin integration based on drift-kick orbital integration reproduces only the one-turn spin-precession axis (close to the vertical). The precession angle about that axis, however, is completely off: compared to the reference solution, the phase of the spin precession is distributed uniformly over a 2π interval. Increasing the number of orbital slices improves the accuracy of spin integration, and at $n_b = 64$ the single-turn spin-phase errors fall below a mrad.

When performing spin integration based on BK/MK orbital integration, we obtain errors that are consistently
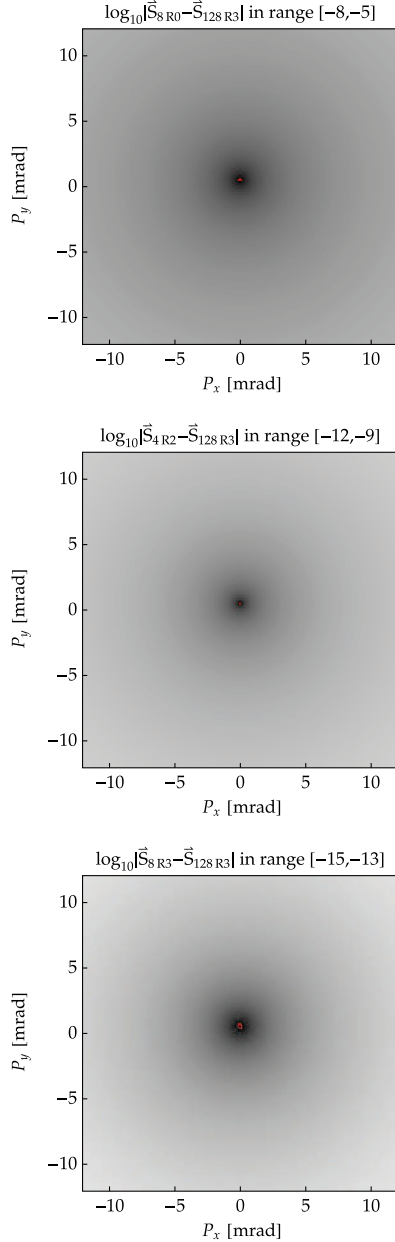
FIG. 13. Absolute error in the computed spin across a solenoid. We computed the data shown in these three graphics using, from top to bottom, eight slices PWC spin integration, four slices PWC plus two Romberg iterations, and eight slices PWC plus three Romberg iterations. Note the different error ranges indicated above each graphic: with each step down the page, the errors fall by several decades.

two-and-a-half decades below those based on drift-kick integration (see the lowest curve in Fig. 14). This result, we emphasize, obtains in the *absence* of Romberg quadrature.

Figure 15 also shows the absolute spin error, but now as a function of particle emittance, after a single turn in the RHIC lattice. As in Fig. 14, no Romberg steps were applied— except in computing the reference solution—so that here also the differences seen derive solely from differences in the
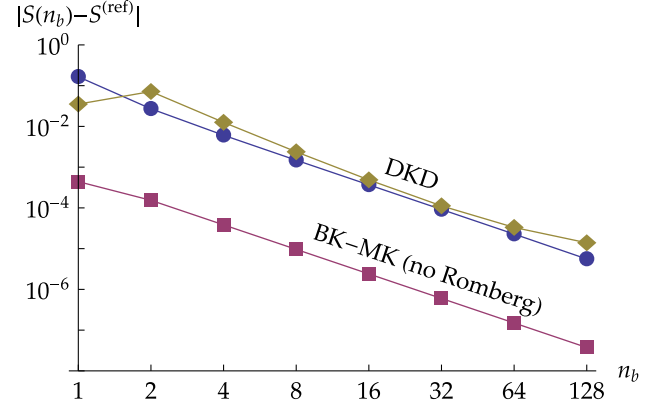

FIG. 14. Single-turn spin errors as a function of the number of slices for several different integrators: drift-kick (blue circles, tan diamonds), and BK/MK (purple squares). The blue circles represent DKSKD: traverse each slice using half-drift, half-kick, spin rotation, half-kick, half-drift. The purple squares represent DKDSDKD: traverse each slice using half-DKD, spin rotation, half-DKD. The reference solution used BK/MK with $n_b = 256$ together with three Romberg iterations.

accuracy of the orbital integration. In addition, we have reduced the noise level in this graphic by, as in Fig. 5, averaging over narrow slices of emittance.

3. Evolution of spin errors

We also examined how spin errors evolve over the course of many simulated turns in RHIC. In Fig. 16 we use the accumulated spin error to compare different methods of integration. The green curve shows results obtained using the previous standard: a TEAPOT split with four slices for most elements, but 16 slices for the strong elements around the interaction regions. The lowest two curves we obtained using the new integrators, also with $n_b = 4$. For the blue curve, we used just the "trapezoidal" rule (PWC of Fig. 7). For the lowest curve, we computed a Romberg limit using a maximum k of 3.
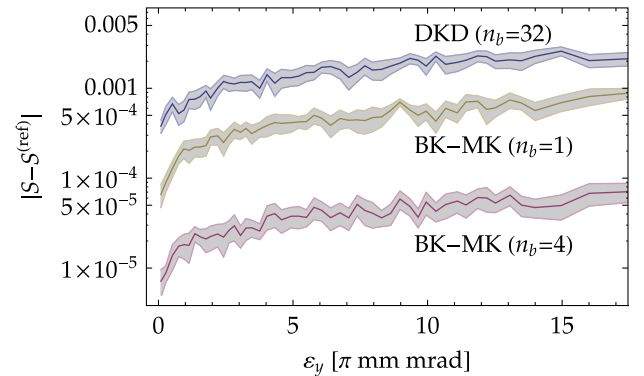

FIG. 15. Amplitude dependence of the single-turn spin errors for three different integrators: drift-kick with $n_b = 32$ (upper), and BK/MK with $n_b = 1$ (middle), and $n_b = 4$ (lower).
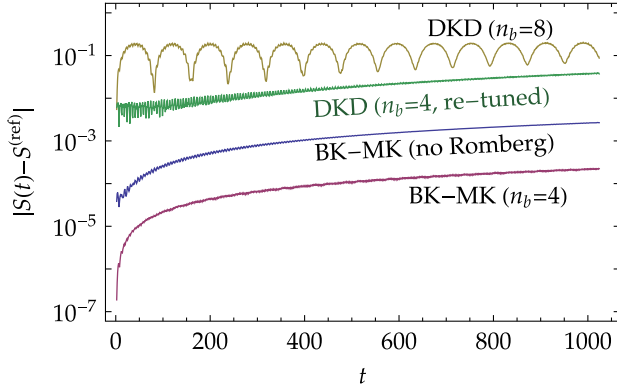
FIG. 16.   Accumulated spin error as a function of turn number, for multiturn tracking in a RHIC lattice, using different methods of integration. From top to bottom: drift-kick integration with $n_b = 8$ and no tune-fit (tan); the previous standard, a TEAPOT split with $n_b = 4$ (green); BK/MK with $n_b = 4$ and no Romberg (blue); and BK/MK with $n_b = 4$ (purple). The simulated beam had a 95% emittance of 15π mm mrad.

The topmost curve, obtained using drift-kick integration with $n_b = 8$, but without retuning the lattice, exhibits pronounced oscillations. The tune errors in the orbital integration translate into these periodic oscillations of the spin errors.

VI. GPU-ACCELERATED IMPLEMENTATION

In recent years, graphics processing units (GPUs) have emerged as an effective means to accelerate compute-intensive workloads. Because they are designed for massively data-parallel workloads with very regular control flow and memory access patterns, GPUs can dedicate a much larger fraction of their transistors to processing elements (registers and arithmetic and logic units) than can CPUs. As a consequence, they can achieve significantly larger computational throughput, performance per watt, and performance per dollar for workloads that are well suited to their architecture. Particle tracking in the absence of space-charge effects is an embarrassingly parallel problem, and is therefore a natural fit for GPUs.

To map the spin-orbit tracking problem onto GPU hardware, we assign a GPU thread to each particle. For the actual tracking, we first transfer the state of a bunch to global memory on the GPU. This state comprises six-dimensional orbit data, three-dimensional spin data, plus the quaternions required for accumulating spin rotations. Particle tracking through individual beam line elements is broken up into several GPU kernels. We first compute the orbital motion through a particular beam line element. This requires loading orbit data at the element entrance from global memory into registers, computing the orbital motion using the integrators discussed in detail above, and then, at the element exit, writing the updated state back to global memory. During the orbital computation, we record in global memory the spin
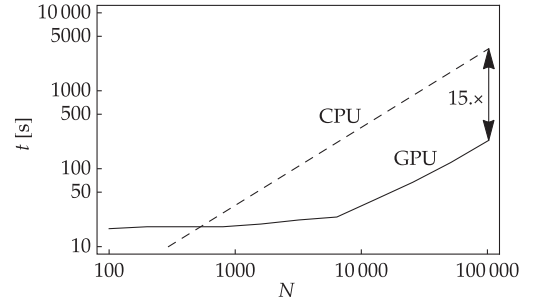

FIG. 17.   Time required to track N particles through 100 turns of the RHIC lattice with GPU-accelerated integrators (solid). The CPU time (dashed) is an estimate; see text for details.

precession vector $\vec{\Omega}$ at the edge of each orbital slice. After integrating the orbital data through the given element, we then use the recorded $\vec{\Omega}$ data to update the quaternions using PWC spin integration plus one or more Romberg steps. We then update the spin.

The particle data resides on the GPU throughout spin-orbit tracking. By maintaining the data on the GPU, we avoid unnecessary data transfers between GPU and CPU, which can significantly slow down simulations. Particle data is occasionally copied back to the CPU in order to perform analysis; e.g., to compute (or update) the invariant spin field (ISF) at a particular location in the accelerator.

Figure 17 shows a comparison of the performance of spin-orbit tracking on a GPU and a CPU. The GPU simulations were carried out on the Dirac cluster at the National Energy Research Scientific Computing center (NERSC). The CPU performance is an estimate for an eight-core CPU, based on the assumption that the simulation scales perfectly to eight CPU cores and that each CPU core is a factor 4 faster than a GPU core. For up to a few hundred particles the CPU is faster than the GPU. For larger numbers of particles, the GPU accelerated simulations are faster. For a few ten-thousands of particles, the GPU accelerated simulations are approximately 15 times faster than the corresponding CPU implementation.

VII. CONCLUSION

In Sec. V A, we showed that the BK/MK integrators can be much more efficient than the drift-kick integrators. This, of course, is well known (but see Forest's discussion [37] for why drift-kick integration remains useful). Our purpose in including such a comparison here has to do with the influence of accurate orbit integration on the accuracy of spin integration.

Figure 14 tells us that better orbit data yields more accurate spin integration. Using BK/MK integration yields a two-decade reduction of the spin errors as compared to drift-kick integration. We see that the improvement with the number of orbital slices is, in large part, due to the improvement in the orbital accuracy. In other words,

with drift-kick integration the accuracy of the spin motion will improve just by increasing the number of orbital slices while holding the number of spin integration steps fixed.

On the other hand, for BK/MK integration, the discretization of the spin motion appears to be a significant source of error: A comparison of Figs. 4 and 5 with Figs. 14 and 15 show that the relative error in the orbital motion is much smaller than that in the spin rotation angle.

In Sec. IV B we discussed interpolating the spin precession vectors, using cubic Bézier polynomials, as a means of rapidly computing multiple spin steps per orbital slice. To compute a Bézier curve for $\vec{\Omega}$, one must first determine the associated control points (see Fig. 1), and doing so requires computing the longitudinal derivative of $\vec{\Omega}$ at the beginning and end of each slice. For most elements, this is a nontrivial, hence time-consuming, computation. In other words, if we wish to enjoy the fast computation of many $\vec{\Omega}$s along an interpolating Bézier curve, we must first invest some time computing $\vec{\Omega}'$. As a trade-off, this means we benefit only if we plan to interpolate $\vec{\Omega}$ at some (possibly large) number of intermediate locations.

In the bargain, however, we do not improve the *rate* at which spin integration converges: The accuracy of this approach cannot improve upon what we obtain by taking smaller steps (i.e., using more orbital slices). In other words, the best we might hope for from interpolating the precession vector $\vec{\Omega}$, is a faster computation of PWC spin integration, which exhibits just second-order convergence.

As shown in Fig. 7, as well as in Figs. 8 through 13, the application of Romberg quadratures to PWC spin integration *does* increase the rate of convergence. Moreover, this approach requires only a modest number of orbital slices. Applying the Romberg algorithm to a sequence of quaternion products constitutes an easy and fast means of accelerating the convergence of spin integration. As a consequence, we have not pursued further the use of Bézier curves to speed up the computation of $\vec{\Omega}$.

Note that our *initial* motivation for using quaternions was the factor of 2 savings in arithmetic they convey. After discovering that Romberg applied to quaternions accelerates the convergence of the spin motion, we have another—much more significant—reason to use quaternions when integrating the Thomas-BMT equation for spin motion.

Concerning the convergence of spin motion, we here comment on what happens when simulating a system near a spin-orbit resonance [16]. First, note that in such a system a particle's spin orientation becomes very sensitive to even very tiny perturbations. As a consequence, if we follow a single spin and ask for its motion to converge as we slice the lattice ever more finely, we may never see that happen. As with the orbital motion, we do not—cannot—ask for the "exact result," because the machine we build is not the one we designed (though we hope it is close). We instead ask

for accurate qualitative dynamics, as revealed by phase-space portraits.

Second, note that the integrators are particular to given elements. In other words, they care only about the mapping of phase-space coordinates across a single element: z^i at entrance $\rightarrow z^f$ at exit. On the other hand, a resonance condition can exist only in a ring: it is *defined* by the ring. One might therefore expect that the choice of integrator is, in some sense, independent of the proximity and strength of a resonance. But this expectation can hold *only if the integrators are independent of, or have no effect on, whatever properties define the resonance.*

When using drift-kick integrators, we know the tune seen in tracking data depends on how finely we slice our elements. What happens, then, as we approach a resonance? If we increase the slicing so as to ensure convergence, the effective tune changes. Hence the distance to resonance changes, and the spin dynamics can appear quite different—especially because of its sensitivity in the neighborhood of a resonance. If, on increasing the slicing we also retune the lattice, then the distance to resonance, as defined by the orbital tunes, remains the same. But the quadrupole strengths do change, and this implies a change in the spin tune for a given simulated particle [16]. And so again the distance to resonance changes, and again the spin dynamics can appear quite different because of its sensitivity in the neighborhood of a resonance.

Because BK/MK integrators yield the same linear orbital tunes independently of slicing, they are not subject to the same difficulties.

In addition to the improved accuracy and accelerated convergence of spin tracking, our code benefits from the significant speed-up it derives just from its implementation on GPUs. As shown in Fig. 17, a reasonable estimate for that benefit yields a speed-up factor of about 15. On a more practical level, we note that on one of NVIDIA's GeForce TITAN GPU nodes, computing an ISF for RHIC at 10^4 phase-space locations using 10^3 turns takes less than fifteen minutes.

Finally, we are currently implementing integrators for other types of optical elements. These include electrostatic lenses, which are important for simulating EDM experiments, as well as better models for Siberian snakes.

ABELL *et al.*

Phys. Rev. ST Accel. Beams **18,** 024001 (2015)

[1] C. A. Aidala, S. D. Bass, D. Hasch, and G. K. Mallot, Rev. Mod. Phys. **85**, 655 (2013).

[2] M. Harrison, S. G. Peggs, and T. Roser, Annu. Rev. Nucl. Part. Sci. **52**, 425 (2002).

[3] M. Bai, Eur. Phys. J. Spec. Top. **162**, 181 (2008).

[4] H. Huang, L. Ahrens, I. G. Alekseev, E. Aschnauer *et al.*, in *Proceedings of the 2nd International Particle Accelerator Conference, San Sebastian, Spain, 2011*, edited by C. Petit-Jean-Genaz *et al.* (2011), pp. 1888–1890.

[5] W. Fischer and A. Bazilevsky, Phys. Rev. ST Accel. Beams **15**, 041001 (2012).

[6] K. A. Heinemann and G. H. Hoffstätter, Phys. Rev. E **54**, 4240 (1996).

[7] G. H. Hoffstaetter, *High-Energy Polarized Proton Beams: A Modern View*, Springer Tracts in Modern Physics (Springer, New York, 2006), Vol. 218.

[8] C. J. G. Onderwater, J. Phys. Conf. Ser. **295**, 012008 (2011).

[9] A. Lehrach, in *Proceedings of the 11th International Computational Accelerator Physics Conference, Rostock-Warnemünde, Germany, 2012*, edited by D. Hecht (2012), pp. 7–11.

[10] A. Accardi, V. Guzey, A. Prokudin, and C. Weiss, arXiv:1110.1031.

[11] S. Abeyratne, A. Accardi, S. Ahmed *et al.*, arXiv:1209.0757.

[12] J. L. Abelleira Fernandez, C. Adolphsen, A. N. Akay *et al.*, arXiv:1206.2913.

[13] O. Brüning and M. Klein, arXiv:1305.2090.

[14] J. P. Miller, E. de Rafael, and B. L. Roberts, Rep. Prog. Phys. **70**, 795 (2007).

[15] B. L. Roberts, arXiv:1001.2898v2.

[16] D. P. Barber, J. A. Ellison, and K. A. Heinemann, Phys. Rev. ST Accel. Beams **7**, 124002 (2004), an extended version is available at http://arxiv.org/abs/physics/0412157.

[17] F. Méot, Nucl. Instrum. Methods Phys. Res., Sect. A **313**, 492 (1992).

[18] F. Méot, ZGOUBI, http://sourceforge.net/projects/zgoubi/.

[19] M. Berz, COSY-INFINITY, http://cosyinfinity.org.

[20] A. U. Luccio, Brookhaven National Laboratory Technical Report No. BNL-52481, 1995.

[21] A. U. Luccio, A. Lehrach, J. Niederer, T. Roser, M. J. Syphers, and N. Tsoupas, in *Proceedings of the 1999 Particle Accelerator Conference, New York, NY, 1999*, edited by A. U. Luccio and W. W. MacKay (IEEE, Piscataway, NJ, 1999).

[22] M. Vogt, Ph.D. dissertation, Universität Hamburg, 2000.

[23] V. V. Balandin and N. I. Golubeva, in *Computational Accelerator Physics*, edited by J. J. Bisognano and A. A. Mondelli, AIP Conference Proceedings (AIP Press, Woodbury, NY, 1997), Vol. 391, see also http://arxiv.org/abs/physics/9903032.

[24] É. Forest, F. Schmidt, and E. McIntosh, KEK Report No. 2002-3, 2002.

[25] É. Forest, Y. Nogiwa, and F. Schmidt, in *Proceedings of the 9th International Computational Accelerator Physics Conference, Chamonix, France, 2006* (2006), pp. 17–21.

[26] É. Forest, KEK Report No. 2010-39, 2010.

[27] D. Sagan, BMAD, http://www.lns.cornell.edu/~dcs/bmad/.

[28] S. R. Mane, Nucl. Instrum. Methods Phys. Res., Sect. A **680**, 35 (2012).

[29] A. W. Chao, Nucl. Instrum. Methods **180**, 29 (1981).

[30] V. H. Ranjbar, F. Méot, and M. Bai, in *Proceedings of the 2011 Particle Accelerator Conference, New York, NY, 2011*, edited by T. Satogata (IEEE, Piscataway, NJ, 2011), pp. 1624–1626.

[31] D. T. Abell, D. Meiser, D. P. Barber, M. Bai, and V. H. Ranjbar, in *Proceedings of the 4th International Particle Accelerator Conference, Shanghai, China, 2013* edited by C. Petit-Jean-Genaz *et al.* (2013), pp. 1037–1039.

[32] N. Malitsky, R. Talman, M. Blaskiewicz, R. Calaga, R. Fliller III, A. Luccio, T. Satogata, and J. Wei, in *Proceedings of the 2003 Particle Accelerator Conference, Portland, OR, 2003*, edited by J. T. Chew, P. W. Lucas, and S. Webber (IEEE, Piscataway, NJ, 2003), p. 272.

[33] L. H. Thomas, Philos. Mag. S7 **3**, 1 (1927).

[34] V. Bargmann, L. Michel, and V. L. Telegdi, Phys. Rev. Lett. **2**, 435 (1959).

[35] Note that for particles with integer spin, there can be a state with vanishing spin expectation value. But this state is usually of little interest for experiments with polarized beams.

[36] *MAD—Methodical Accelerator Design* (European Organization for Nuclear Research, Geneva, 2002).

[37] É. Forest, J. Phys. A **39**, 5321 (2006).

[38] K. L. Brown, SLAC Report No. 75, 1982.

[39] Y. Papaphilipou and D. T. Abell, in *Proceedings of the 7th European Particle Accelerator Conference, Vienna, Austria, 2000*, edited by A. Gallo (European Physical Society, Geneva, 2000), pp. 1453–1455.

[40] É. Forest and J. Milutinović, Nucl. Instrum. Methods Phys. Res., Sect. A **269**, 474 (1988).

[41] É. Forest, *Beam Dynamics: A New Attitude and Framework*, The Physics and Technology of Particle and Photon Beams, Vol. 8 (Harwood Academic Publishers, Amsterdam, 1998).

[42] É. Forest, S. C. Leemann, and F. Schmidt, KEK Report No. 2005-109, 2006.

[43] É. Forest, S. C. Leemann, and F. Schmidt, KEK Report No. 2005-110, 2006.

[44] M. Venturini and A. J. Dragt, Nucl. Instrum. Methods Phys. Res., Sect. A **427**, 387 (1999).

[45] M. Venturini and A. J. Dragt, in *Proceedings of the 1999 Particle Accelerator Conference, New York, NY, 1999*, edited by A. U. Luccio and W. W. MacKay (IEEE, Piscataway, NJ, 1999), pp. 1593–1595.

[46] C. E. Mitchell and A. J. Dragt, Phys. Rev. ST Accel. Beams **13**, 064001 (2010).

[47] A. J. Dragt, *Lie Methods for Nonlinear Dynamics with Applications to Accelerator Physics*, latest version available at http://www.physics.umd.edu/dsat/.

[48] K. A. Heinemann, Technical Report No. DESY-96-229, 1996, available at http://arXiv.org/abs/physics/9611001.

[49] H. Yoshida, Phys. Lett. A **150**, 262 (1990).

[50] R. I. McLachlan and G. R. W. Quispel, Acta Numer. **11**, 341 (2002).

[51] L. Schachinger and R. M. Talman, SSC Central Design Group Technical Report No. SSC-52, 1985, http://lss.fnal.gov/archive/other/ssc/ssc-52.pdf.

[52] É. Forest, M. F. Reusch, D. L. Bruhwiler, and A. Amiry, *Part. Accel.* **45**, 65 (1994).

[53] A. J. Dragt, in *Physics of High Energy Particle Accelerators: Fermilab Summer School, Batavia, IL, 1981*, AIP Conf. Proc. No. 87, edited by R. A. Carrigan, F. R. Huson, and M. Month (AIP Press, Woodbury, NY, 1982), pp. 147–313.

[54] A. J. Dragt and É. Forest, J. Math. Phys. (N.Y.) **24**, 2734 (1983).

[55] L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics: Theory and Applications*, Encyclopedia of Mathematics and Its Applications, Vol. 8 (Addison-Wesley Publishing Co., Reading, MA, 1981).

[56] To simplify the bookkeeping, we do this at the end of each element, but once per turn would likely suffice.

[57] B. Casselman, *Mathematical Illustrations: A Manual of Geometry and PostScript* (Cambridge University Press, Cambridge, UK, 2005).

[58] Computing the longitudinal derivative of $\vec{\Omega}$ is tedious, but straightforward: Express the magnetic field in terms of particle location—so that $\vec{\Omega}$ depends solely on the orbital phase-space variables. Then use the chain rule to compute

$$\frac{\mathrm{d}\vec{\Omega}}{\mathrm{d}s} = \frac{\partial\vec{\Omega}}{\partial X}\frac{\partial X}{\partial s} + \frac{\partial\vec{\Omega}}{\partial P_x}\frac{\partial P_x}{\partial s} + \cdots.$$

The quantities $\partial_s X, \partial_s P_x, \ldots$ can be obtained from the orbital equations of motion.

[59] F. S. Acton, *Real Computing Made Real: Preventing Errors in Scientific and Engineering Calculations* (Princeton University Press, Princeton, NJ, 1996).

[60] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2002).

[61] F. B. Hildebrand, *Introduction to Numerical Analysis*, 2nd ed. (Dover Publications, Inc., Mineola, NY, 1987), unabridged, slightly corrected republication of the second edition (1974) of the work first published by McGraw-Hill, Inc., in 1956.

[62] H. Gröte, J. Holt, N. Malitsky, F. Pilat, R. Talman, and C. G. Trahern, Brookhaven National Laboratory Technical Report No. RHIC/AP/155, 1998.

[63] Bend-Kick/Matrix-Kick.

[64] While not essential to our discussion, we note that this particle had the initial phase-space values $X^i = 1.0$ mm, $P_x^i = 0.1$ mrad, $P_y^i = -0.2$ mrad, and P_t^i corresponding to a relative momentum deviation $\delta = 10^{-3}$. We set both Y^i and T^i to zero. In addition, we set this particle's initial spin to $\vec{S} \approx (0.1952, 0.9759, 0.0976)$.

[65] G. Marsaglia, Ann. Math. Stat. **43**, 645 (1972).