AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Space Physics

TECHNIQUE

10.1002/2014JA020127

Key Points:
- A fully 3-D radiation belt model in adiabatic invariant coordinates is developed
- Correct form of the zero second invariant boundary condition is derived
- Model reveals distinctive electron sources with and without local acceleration

Correspondence to:
L. Zheng,
zhengliheng@rice.edu

Citation:
Zheng, L., A. A. Chan, J. M. Albert, S. R. Elkington, J. Koller, R. B. Horne, S. A. Glauert, and N. P. Meredith (2014), Three-dimensional stochastic modeling of radiation belts in adiabatic invariant coordinates, *J. Geophys. Res. Space Physics*, *119*, 7615–7635, doi:10.1002/2014JA020127.

Received 24 APR 2014
Accepted 20 AUG 2014
Accepted article online 26 AUG 2014
Published online 29 SEP 2014

Three-dimensional stochastic modeling of radiation belts in adiabatic invariant coordinates

Liheng Zheng[1], Anthony A. Chan[1], Jay M. Albert[2], Scot R. Elkington[3], Josef Koller[4], Richard B. Horne[5], Sarah A. Glauert[5], and Nigel P. Meredith[5]

[1]Department of Physics and Astronomy, William Marsh Rice University, Houston, Texas, USA, [2]Air Force Research Laboratory, Space Vehicles Directorate, Albuquerque, New Mexico, USA, [3]Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, Colorado, USA, [4]Space Science and Application, Los Alamos National Laboratory, Los Alamos, New Mexico, USA, [5]British Antarctic Survey, Natural Environment Research Council, Cambridge, UK

Abstract A 3-D model for solving the radiation belt diffusion equation in adiabatic invariant coordinates has been developed and tested. The model, named Radbelt Electron Model, obtains a probabilistic solution by solving a set of Itô stochastic differential equations that are mathematically equivalent to the diffusion equation. This method is capable of solving diffusion equations with a full 3-D diffusion tensor, including the radial-local cross diffusion components. The correct form of the boundary condition at equatorial pitch angle $\alpha_0 = 90°$ is also derived. The model is applied to a simulation of the October 2002 storm event. At α_0 near 90°, our results are quantitatively consistent with GPS observations of phase space density (PSD) increases, suggesting dominance of radial diffusion; at smaller α_0, the observed PSD increases are overestimated by the model, possibly due to the α_0-independent radial diffusion coefficients, or to insufficient electron loss in the model, or both. Statistical analysis of the stochastic processes provides further insights into the diffusion processes, showing distinctive electron source distributions with and without local acceleration.

1. Introduction

The Earth's outer radiation belt is a torus-shaped region around the Earth consisting of relativistic electrons that pose a threat to human space assets [e.g., *Baker et al.*, 1997]. Fluxes of these energetic electrons in the outer radiation belt respond dynamically to magnetospheric perturbations and can vary by several orders of magnitude during storm times [e.g., *Reeves et al.*, 2003; *Baker and Kanekal*, 2008; *Tu et al.*, 2009]. The flux variations may be caused by two types of physical processes: adiabatic modulations in which the perturbation time scale is much greater than the electron drift period [e.g., *Kim and Chan*, 1997; *Tu and Li*, 2011; *Su et al.*, 2011] and nonadiabatic processes, caused by electron resonance with waves near the cyclotron and drift frequencies that violate one or more of the electron's adiabatic invariants [e.g., *Schulz and Lanzerotti*, 1974; *Elkington et al.*, 1999; *Thorne*, 2010].

Quasi-linear diffusion theory has been formulated to quantify the nonadiabatic changes of energetic electron fluxes, and diffusion models that are based on solving an electron Fokker-Planck equation have been developed [e.g., *Albert et al.*, 2009; *Subbotin et al.*, 2010; *Subbotin and Shprits*, 2012; *Fok et al.*, 2008; *Varotsou et al.*, 2008; *Tao et al.*, 2008, 2009; *Tu et al.*, 2009, 2013; *Camporeale et al.*, 2013a; *Glauert et al.*, 2014]. Typically, solving the Fokker-Planck equation involves using a numerical grid and finite-difference methods. However, unphysical negative solutions occur when off-diagonal diffusion tensor components exist [*Albert*, 2009, 2013; *Camporeale et al.*, 2013b, 2013c]. To overcome this in two dimensional (2-D) energy and pitch angle diffusion simulations, *Albert and Young* [2005] globally diagonalized the diffusion tensor by coordinate transformation and applied a standard finite-difference method thereafter. But in 3-D, where radial diffusion is included, such matrix diagonalization is not easily achieved. Instead, an operator-splitting technique is widely adopted, which uses two sets of grids, one for the radial diffusion and the other for the energy and pitch angle diffusion. To communicate between grids, frequent high-accuracy interpolations are required. As a result, a high grid resolution is required in these methods [*Subbotin et al.*, 2010]. Moreover, the two sets of grids imply exclusion of possible radial-local cross diffusion; in other words, the diffusion tensor is at most a 2-D energy and pitch angle block plus a 1-D radial diffusion coefficient, which is not fully 3-D.

Alternatively, it is useful to model radiation belt diffusion in the space of the adiabatic invariants. In adiabatic invariant coordinates, phase space density (PSD) is conserved under adiabatic processes, and the model results can be mapped to particle fluxes for a variety of magnetic field models. *Beutier and Boscher* [1995] wrote the Salammbô model using one 3-D grid in the three action integrals J_1, J_2, and J_3. However, a box computational domain in the action integral space transforms to a highly extended region of complicated shape in the L shell energy and pitch angle (hereafter denoted $\{L, E, \alpha_0\}$) coordinates, and vice versa. This can make setting boundary conditions difficult. Later, the Salammbô model was converted to the operator-splitting technique in the $\{L, E, \alpha_0\}$ coordinates [*Varotsou et al.*, 2008]. *Subbotin and Shprits* [2012] designed a new set of adiabatic invariant variables from mixtures of the action integrals, whose coordinate lines, by clever choice of a constant related to magnetic field, roughly follow those of the $\{L, E, \alpha_0\}$ coordinates. Currently, their model is also not fully 3-D.

From an entirely different approach, *Tao et al.* [2008] demonstrated successful application of the stochastic differential equation (SDE) method in solving the 2-D energy and pitch angle Fokker-Planck equation. *Selesnick* [2012] used this method in an investigation of electron decay rates in the inner radiation belt, where nondiffusive scattering becomes significant. The SDE method avoids global matrix diagonalization or reliance on a grid; this is advantageous compared to finite-difference methods in dealing with off-diagonal diffusion components and complex domain geometries. In the current study, we present the application of the SDE method to modeling the outer radiation belt in a fully 3-D adiabatic invariant coordinate system. Section 2 contains a brief introduction to SDE theory, derivations of the specific SDEs and boundary conditions of the radiation belt Fokker-Planck equation in adiabatic invariant variables, and a description of the numerical schemes to solve the system. In section 3, our SDE model, named REM (for Radbelt Electron Model), is applied to simulation of a storm event of October 2002. Insights into the diffusion process are provided via statistical analysis of the stochastic processes. Section 4 gives a summary and discussion of this work.

2. Solving the Radiation Belt Electron Diffusion Equation in Adiabatic Invariant Coordinates

From the perspective of SDEs, a diffusion process can be viewed as an ensemble of stochastic processes and the solution of the diffusion equation can be calculated from the expectation of path integrals of these stochastic processes. See Appendix A and the references therein for a formulation of the SDE theory and a derivation of the Feynman-Kac formula, which gives a solution to an initial value diffusion problem. In this section, we generalize the Feynman-Kac formula to diffusion problems with boundary conditions of the first kind (Dirichlet) and the second kind (Neumann) (section 2.1) and then apply it to the radiation belt electron diffusion equation. For this purpose, section 2.2 discusses the transformation of the Fokker-Planck equation into adiabatic invariant variables and then derives the corresponding SDEs, including a description of the diffusion coefficients in the new variables. Section 2.3 discusses the complex computational domain in the new coordinates, and the particular Neumann boundary condition that arises there. In section 2.4, we give a brief description of numerical implementation and validation of this method.

2.1. SDE Method for Solving Diffusion Problems With Boundary Conditions

As in Appendix A equation (A38), consider an n-D diffusion equation written in the Kolmogorov backward equation form (assuming summation convention)

$$\begin{aligned}\frac{\partial}{\partial t}u(t,\boldsymbol{\xi}) &= \widetilde{\mathcal{L}}u(t,\boldsymbol{\xi}) \\ &= \frac{1}{2}a_{ij}(t,\boldsymbol{\xi})\frac{\partial^2 u}{\partial\xi_i\partial\xi_j} + b_i(t,\boldsymbol{\xi})\frac{\partial u}{\partial\xi_i} + c(t,\boldsymbol{\xi})u,\end{aligned} \tag{1}$$

where a_{ij} are components of diffusion tensor $\mathbf{a}$, b_i relate to a_{ij} as in (A24) (in Cartesian coordinates), and c is a term characterizing source and loss. The diffusion process described by equation (1) is related to an n-D Itô stochastic process $\mathbf{X}_s^t$ whose equation of motion is given by the following SDE

$$\mathrm{d}\mathbf{X}_s^t = \mathbf{b}(t-s,\mathbf{X}_s^t)\mathrm{d}s + \boldsymbol{\sigma}(t-s,\mathbf{X}_s^t)\mathrm{d}\mathbf{W}_s, \quad 0 \le s \le t, \tag{2}$$

with initial position $\mathbf{X}_0^t = \boldsymbol{\xi}$. In equation (2) s is the running time, increasing from 0 to t, $\mathbf{W}_s$ is an n-D Wiener process, $\mathbf{b}$ is an n vector with components b_i, and $\boldsymbol{\sigma}$ is an $n \times n$ matrix obtained from the relation

$$\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathsf{T}} = \mathbf{a}. \tag{3}$$

Note that since the time argument of the coefficients in the SDE is $t - s$, the Itô process travels backward in time from t to 0.

The matrix $\boldsymbol{\sigma}$ is not uniquely determined by equation (3). However, according to Levi's theorem [*Freidlin*, 1985], different $\boldsymbol{\sigma}$ matrices satisfying (3) generate equivalent stochastic processes and hence yield the same solution of the diffusion equation. To see this, we observe that for any orthogonal matrix $\mathbf{U}$, the new $\boldsymbol{\sigma}$ obtained from $\tilde{\boldsymbol{\sigma}} \equiv \boldsymbol{\sigma}\mathbf{U}$ still satisfies equation (3), and, when inserted into the SDE (2), $\tilde{\boldsymbol{\sigma}}\mathrm{d}\mathbf{W}_s$ is equivalent to $\boldsymbol{\sigma}\mathrm{d}\widetilde{\mathbf{W}}_s$, where $\widetilde{\mathbf{W}}_s$ is merely the Wiener process $\mathbf{W}_s$ after a rotation and perhaps a reflection of coordinate axes, which is still a Wiener process [*Zhang*, 1999].

Given an initial condition

$$u(0, \boldsymbol{\xi}) = f(\boldsymbol{\xi}), \tag{4}$$

the diffusion equation (1) can be solved by defining the functional of the stochastic paths

$$F^{t,\boldsymbol{\xi}} = f(\mathbf{X}_t^t) \exp\left(\int_0^t c(t-s, \mathbf{X}_s^t)\mathrm{d}s\right), \tag{5}$$

where the superscripts in $F^{t,\boldsymbol{\xi}}$ indicate the starting time and position of $\mathbf{X}_s^t$, and $\mathbf{X}_t^t$ is a shorthand for $\mathbf{X}_{s=t}^t$. The Feynman-Kac formula (A41) asserts that equation (1), with initial condition (4) and boundaries at infinity, has the solution

$$u(t, \boldsymbol{\xi}) = \mathbb{E}[F^{t,\boldsymbol{\xi}}], \tag{6}$$

where the symbol $\mathbb{E}$ denotes expectation over an ensemble of stochastic processes. Observing that the Itô processes travel backward in time, the meaning of the Feynman-Kac formula is apparent: the solution at position $\boldsymbol{\xi}$ and time t is obtained by averaging the initial condition and the propagation information carried by the Itô processes along their paths.

To implement a Dirichlet boundary condition in a diffusion problem, we introduce the "stopping time," which is denoted by τ henceforth. Given a domain D, the stopping time is the time s until the Itô process first reaches the boundary and after which it remains at the same place. Such boundaries are denoted as $\partial_1 D$, where the subscript 1 stands for a boundary condition of "the first kind," also known as a Dirichlet boundary condition [*Øksendal*, 1998; *Freidlin*, 1985]. If $\partial_1 D$ can never be reached, τ is $+\infty$.

Let us consider the initial value and Dirichlet boundary value problem:

$$\frac{\partial}{\partial t}u(t, \boldsymbol{\xi}) = \widetilde{\mathcal{L}}u(t, \boldsymbol{\xi}), \tag{7}$$

$$u(0, \boldsymbol{\xi}) = f(\boldsymbol{\xi}), \quad \boldsymbol{\xi} \in D, \tag{8}$$

$$u(t, \boldsymbol{\xi}) = g(t, \boldsymbol{\xi}), \quad \boldsymbol{\xi} \in \partial_1 D, \tag{9}$$

in which $g(t, \boldsymbol{\xi})$ is the Dirichlet boundary condition and where $g(0, \boldsymbol{\xi}) = f(\boldsymbol{\xi})$ for $\boldsymbol{\xi} \in \partial_1 D$. To solve this problem, we generalize the functional $F^{t,\boldsymbol{\xi}}$ as

$$\begin{aligned} F^{t,\boldsymbol{\xi}} = {} & \mathbb{I}_{\tau \geq t} f(\mathbf{X}_t^t) \exp\left(\int_0^t c(t-s, \mathbf{X}_s^t)\mathrm{d}s\right) \\ & + \mathbb{I}_{\tau < t} g(t-\tau, \mathbf{X}_\tau^t) \exp\left(\int_0^\tau c(t-s, \mathbf{X}_s^t)\mathrm{d}s\right), \end{aligned} \tag{10}$$

where the indicator $\mathbb{I}_C$ has value 1 when condition C is true, and 0 otherwise [*Freidlin*, 1985]. The meaning of (10) is simple: if the Itô process remains in D when time s runs out, apply the initial condition; if it stops on

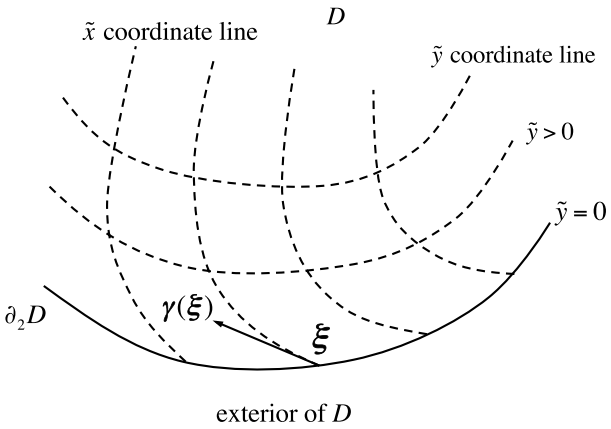

Figure 1. A schematic illustration of a 2-D domain D with a Neumann-type boundary $\partial_2 D$ and a local curvilinear coordinate system $\{\tilde{x}, \tilde{y}\}$ constructed near the boundary. The $\tilde{x}$ coordinate lines are tangent to the $\boldsymbol{\gamma}$ vectors on $\partial_2 D$, and the $\tilde{y} = 0$ coordinate line coincides with the boundary. The global coordinate is labeled $\boldsymbol{\xi}$.

$\partial_1 D$ before that, apply the Dirichlet boundary condition. The solution of this problem is then given by the expectation of the new functional (10) as in equation (6).

A Neumann-type boundary $\partial_2 D$ is different from a Dirichlet-type boundary $\partial_1 D$. Instead of stopping on the boundary, the Itô process is reflected on $\partial_2 D$ [*Freidlin*, 1985; *Øksendal*, 1998; *Costantini et al.*, 1998]. For a homogeneous Neumann boundary condition of the form

$$\boldsymbol{\gamma}(\boldsymbol{\xi}) \cdot \nabla u(t, \boldsymbol{\xi}) = 0, \quad \boldsymbol{\xi} \in \partial_2 D, \tag{11}$$

where $\boldsymbol{\gamma}(\boldsymbol{\xi})$ is a continuous unit vector field pointing toward the interior of D, the Itô process undergoes a reflection about the direction of $\boldsymbol{\gamma}$ at the place where it reaches $\partial_2 D$. In the SDE, this reflection is implemented by introducing an additional term involving $\boldsymbol{\gamma}$ such that

$$\begin{aligned} \mathrm{d}\mathbf{X}_s^t &= \mathbf{b}(t-s, \mathbf{X}_s^t)\mathrm{d}s + \boldsymbol{\sigma}(t-s, \mathbf{X}_s^t)\mathrm{d}\mathbf{W}_s + \boldsymbol{\gamma}(\mathbf{X}_s^t)\mathrm{d}k_s, \\ & 0 \le s \le t, \end{aligned} \tag{12}$$

where k_s is a nondecreasing 1-D stochastic variable, called local time, with initial value $k_0 = 0$. The increment $\mathrm{d}k_s$ is nonzero only when $\mathbf{X}_s^t$ reflects from $\partial_2 D$, and at other times it remains zero. In this sense, the vector $\boldsymbol{\gamma}$ can be looked upon as a "local velocity" of the stochastic motion on $\partial_2 D$. Consider a 2-D domain as sketched in Figure 1, for example. Because $\boldsymbol{\gamma}$ is not tangent to $\partial_2 D$, in the vicinity of a segment of the boundary we are able to construct a local curvilinear coordinate system, such that the $\tilde{x}$ coordinate lines are tangent to $\boldsymbol{\gamma}$ on $\partial_2 D$, the boundary coincides with the $\tilde{y} = 0$ coordinate line, and the interior of the domain has $\tilde{y} > 0$. When $\mathbf{X}_s^t$ reaches $\partial_2 D$, we can decompose it into the $\{\tilde{x}, \tilde{y}\}$ coordinates, and the local time k_s increases to ensure that the $\tilde{y}$ component of $\mathbf{X}_s^t$ is always greater than 0, while its $\tilde{x}$ component is left unaffected. When $\mathbf{X}_s^t$ is away from $\partial_2 D$, k_s then remains unchanged.

Finally, we consider the initial value and mixed boundary condition problem, which is the type of problem we must solve to solve the radiation belt diffusion equation in this paper:

$$\frac{\partial}{\partial t} u(t, \boldsymbol{\xi}) = \tilde{\mathcal{L}} u(t, \boldsymbol{\xi}), \tag{13}$$

$$u(0, \boldsymbol{\xi}) = f(\boldsymbol{\xi}), \quad \boldsymbol{\xi} \in D, \tag{14}$$

$$u(t, \boldsymbol{\xi}) = g(t, \boldsymbol{\xi}), \quad \boldsymbol{\xi} \in \partial_1 D, \tag{15}$$

$$\boldsymbol{\gamma}(\boldsymbol{\xi}) \cdot \nabla u(t, \boldsymbol{\xi}) = 0, \quad \boldsymbol{\xi} \in \partial_2 D. \tag{16}$$

The functional $F^{t,\boldsymbol{\xi}}$ for this problem has the same form as (10), and the solution of the problem is still given by (6), except now $\mathbf{X}_s^t$ must be expressed by (12).

2.2. Transformation of the Fokker-Planck Equation and Connection to SDEs

Diffusive evolution of the radiation belt electrons is described by a Fokker-Planck equation. In a general set of phase space coordinates $\{Q_i\}$, the Fokker-Planck equation assumes the form

$$\frac{\partial \bar{f}}{\partial t} = \frac{1}{G} \frac{\partial}{\partial Q_i} \left(G D_{Q_i Q_j} \frac{\partial \bar{f}}{\partial Q_j} \right), \tag{17}$$

where $\bar{f}(t, Q_i)$ is the phase-averaged PSD of the electrons, $G = \det(\partial J_i/\partial Q_j)$ is the Jacobian for the transformation from the three adiabatic invariant action integrals $\{J_i\}$ to the new variables $\{Q_j\}$, and $D_{Q_iQ_j}$ denotes components of the diffusion tensor in the $\{Q_j\}$ coordinates [e.g., *Haerendel*, 1968; *Schulz*, 1991]. $D_{Q_iQ_j}$ are related to the diffusion coefficients $D_{J_iJ_j}$ in the $\{J_i\}$ coordinates by the transformation

$$D_{Q_iQ_j} = \frac{\partial Q_i}{\partial J_k} D_{J_kJ_l} \frac{\partial Q_j}{\partial J_l}. \tag{18}$$

In the Fokker-Planck equation (17), we have neglected any source or loss term, as well as the frictional term caused by interparticle interactions, because these terms are relatively unimportant in the evolution of the outer electron radiation belt compared to the wave-particle interactions [*Schulz and Lanzerotti*, 1974]. However, from the formulation in the previous subsection, we note that the Fokker-Planck equation with these terms present is still solvable by the SDE method.

Instead of the action integrals, it is customary to regard $\bar{f}$ as a function of the more convenient adiabatic invariants M, K, and L (Roederer's L [*Roederer*, 1970]), which are related to J_i by physical constants, and whose expressions can be found in, e.g., *Schulz* [1991]. For computational purposes, $u \equiv \ln(M)$ is a better variable than M, for the following three reasons. First, in the typical energy range of electrons in the outer radiation belt, the value of M spans 4 orders of magnitude, from a few to over 10,000 MeV/G. Second, changing to $\ln(M)$ sends the lower bound of the variable from 0 to $-\infty$, which prevents unphysical scenarios of stochastic processes traveling into "negative M" regions during numerical implementation. Third, in the phase space with the $\ln(M)$ coordinate, the family of constant momentum surfaces have the same shape in a dipole magnetic field, which is not true using M; this feature makes the computational domain less irregular.

Following equation (17), we write the Fokker-Planck equation in the coordinates $\{u, K, L\}$. The Jacobian determinant for dipole magnetic field is

$$G = \frac{8\sqrt{2}\pi^2 \mu_0 m_0^{3/2}}{R_E} \frac{\exp(\frac{3}{2}u)}{L^2}, \tag{19}$$

where μ_0 is the Earth's magnetic moment, m_0 is electron mass, and R_E is Earth radius. The Fokker-Planck equation, transformed to the Kolmogorov backward equation form (1), becomes

$$\frac{\partial}{\partial t}\bar{f} = \frac{1}{2} a_{ij} \frac{\partial^2}{\partial Q_i \partial Q_j}\bar{f} + b_i \frac{\partial}{\partial Q_i}\bar{f}, \tag{20}$$

and from now on Q_i represents u, K, and L for $i = 1, 2, 3$, respectively. This can be accomplished by directly expanding the partial derivatives in (17) and collecting terms, due to the fact that $\bar{f}$ is the PSD with respect to canonical variables, though expressed in the $\{Q_i\}$ coordinates (see Appendix A2). The coefficients a_{ij} and b_i are given by

$$\mathbf{a} = 2\begin{pmatrix} D_{uu} & D_{uK} & D_{uL} \\ D_{uK} & D_{KK} & D_{KL} \\ D_{uL} & D_{KL} & D_{LL} \end{pmatrix}, \tag{21}$$

and

$$b_1 = \exp\left(-\tfrac{3}{2}u\right) \frac{\partial}{\partial u}\left[\exp\left(\tfrac{3}{2}u\right) D_{uu}\right] + \frac{\partial}{\partial K} D_{uK} + L^2 \frac{\partial}{\partial L}\left(\frac{D_{uL}}{L^2}\right), \tag{22}$$

$$b_2 = \exp\left(-\tfrac{3}{2}u\right) \frac{\partial}{\partial u}\left[\exp\left(\tfrac{3}{2}u\right) D_{uK}\right] + \frac{\partial}{\partial K} D_{KK} + L^2 \frac{\partial}{\partial L}\left(\frac{D_{KL}}{L^2}\right), \tag{23}$$

$$b_3 = \exp\left(-\tfrac{3}{2}u\right) \frac{\partial}{\partial u}\left[\exp\left(\tfrac{3}{2}u\right) D_{uL}\right] + \frac{\partial}{\partial K} D_{KL} + L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\right). \tag{24}$$

Following the results of Appendix A, the 3-D SDE associated with the Fokker-Planck equation (20), in the form shown by (12), has coefficients $\mathbf{b}$ given by expressions (22)–(24), and the matrix $\boldsymbol{\sigma}$ calculated from the decomposition of $\mathbf{a}$ in (21) according to (3). The reflection vector $\boldsymbol{\gamma}$ will be addressed in section 2.3, together with the boundary conditions.

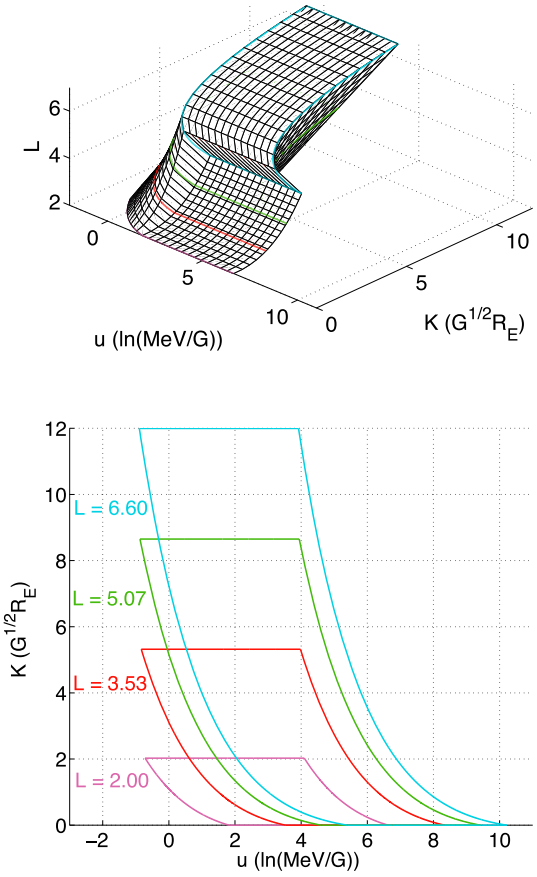

Figure 2. (top) An example computational domain in the *u-K-L* space in a dipole magnetic field, defined by *E* from 0.2 to 5.0 MeV, α_0 from bounce loss cone to 90°, and *L* from 2.0 to 6.6. (bottom) The projections in the *u-K* plane of the colored contours in Figure 2 (top), labeled by their *L* values.

The 3-D diffusion tensor $D_{Q_iQ_j}$ in the $\{u, K, L\}$ coordinates needs additional specification. Among its six independent components, the radial diffusion coefficient D_{LL} may be provided by empirical models [e.g., *Brautigam and Albert*, 2000; *Ozeke et al.*, 2012] or calculated using global MHD simulations [e.g., *Fei et al.*, 2006; *Tu et al.*, 2012]. The off-diagonal terms D_{uL} and D_{KL} are related to drift-shell splitting and Shabansky-type orbits [e.g., *Shabansky*, 1971; *Ukhorskiy et al.*, 2011; *O'Brien*, 2014]; and since they are largely unavailable to date, we use zero value for them in this paper. The remaining components D_{uu}, D_{uK}, and D_{KK} are related to energy and pitch angle scattering of the electrons, and they can be transformed from the more commonly computed diffusion coefficients in the equatorial pitch angle (α_0) and momentum (p) coordinates, by a transformation like (18). Written explicitly, they are

$$\begin{pmatrix} D_{uu} & D_{uK} \\ D_{uK} & D_{KK} \end{pmatrix} = \begin{pmatrix} \frac{\partial u}{\partial \alpha_0} & \frac{\partial u}{\partial p} \\ \frac{\partial K}{\partial \alpha_0} & \frac{\partial K}{\partial p} \end{pmatrix} \begin{pmatrix} D_{\alpha_0\alpha_0} & D_{\alpha_0 p} \\ D_{\alpha_0 p} & D_{pp} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial \alpha_0} & \frac{\partial K}{\partial \alpha_0} \\ \frac{\partial u}{\partial p} & \frac{\partial K}{\partial p} \end{pmatrix}. \tag{25}$$

The partial derivatives on the right-hand side of equation (25) should be regarded as functions of *u*, *K*, and *L* and are field geometry dependent. For a dipole magnetic field, these derivatives can be evaluated using published formulae [e.g., *Schulz*, 1991]. We note that $\partial K/\partial p$ is always zero, even in general magnetic fields, as long as the field is slowly varying (compared to a bounce period) and the particles experience no external force other than the magnetic force [*Roederer*, 1970].

2.3. Computational Domain and Boundary Conditions

The ability to deal with complicated boundary geometries gives the SDE method advantages in solving the electron Fokker-Planck equation in adiabatic invariant coordinates. In this paper, we demonstrate the SDE method by solving the Fokker-Planck equation (20) in a dipole magnetic field. The computational domain is determined by selecting a range of electron kinetic energy *E*, usually from a few hundred keV to several MeV, α_0 from 90° to the local bounce loss cone angle, and a range of *L*. Using dipole field geometry, these boundaries are transformed into the *u-K-L* space and are illustrated in Figure 2. Figure 2 (top) shows a 3-D view of the computational domain, a skewed hexahedron, with colored contours at four fixed *L* values. Visible on the left- and right-hand side of the hexahedron are the curved surfaces of constant minimum and maximum energies, which have the same shape due to the dipole field geometry. The front surface is the plane of the $K = 0$ boundary corresponding to $\alpha_0 = 90°$. The top surface is the upper *L* plane. The maximum *K* surface (corresponding to dipole bounce loss cone angles) and the lower *L* plane are not visible from this point of view. Figure 2 (bottom) shows the projection of the colored constant *L* contours in the *u-K* plane. As *L* increases, the maximum *K* value increases as a result of the diminishing loss cone, and the contour moves as a whole toward the larger *u* region.

Boundary conditions are assigned as follows. Dirichlet boundary conditions are specified on the surfaces of constant maximum energy (E_{max}), constant minimum energy (E_{min}), maximum *L* (L_{max}), minimum *L* (L_{min}), and loss cone *K* (K_{max}). In particular, we assume negligible electron flux at the maximum energy and the loss cone; therefore, the E_{max} and K_{max} boundary values are zero. In other words, Itô processes are lost upon

reaching the E_{max} or K_{max} boundaries. These boundary conditions are apparent from the nature of the diffusion problem, though the loss cone could be nonempty in the strong diffusion limit.

It is, however, not so apparent what form the boundary condition takes for the $K = 0$ boundary. To answer this question, we examine the Fokker-Planck equation in light of the variational theory of partial differential equations (PDEs). Starting from equation (17) and assuming that the spatial and temporal variables are separable, i.e., $\bar{f}(t, Q_i) = \Psi_n(t)\Phi_n(Q_i)$ with n indexing the nth eigenmode, equation (17) can be separated into (with no summation over repeated n in (26), (27), and (29))

$$\frac{d\Psi_n}{dt} = -\lambda_n \Psi_n, \tag{26}$$

and

$$\frac{\partial}{\partial Q_i}\left(GD_{Q_iQ_j}\frac{\partial \Phi_n}{\partial Q_j}\right) + \lambda_n G\Phi_n = 0, \tag{27}$$

where λ_n is the eigenvalue (physically, the temporal decay rate of the nth eigenmode). For each n, equation (27) is the Euler-Lagrange equation of the functional

$$\mathcal{I}[\Phi_n] = \int_D \mathcal{F}(\Phi_{n,i}, \Phi_n, Q_i) dQ_1 dQ_2 dQ_3, \tag{28}$$

in which $\Phi_{n,i} \equiv \partial\Phi_n/\partial Q_i$, and the integrand function $\mathcal{F}$ is

$$\mathcal{F}(\Phi_{n,i}, \Phi_n, Q_i) = G\left(D_{Q_iQ_j}\Phi_{n,i}\Phi_{n,j} - \frac{1}{2}\lambda_n \Phi_n^2\right). \tag{29}$$

Whereas $\bar{f}$, and hence Φ_n, are often held fixed on the domain boundaries, they are free to change on the $Q_2 = 0$ boundary. As a result, to make the functional $\mathcal{I}[\Phi_n]$ stationary against variations $\delta\Phi_n$, in addition to the Euler-Lagrange equation, we must also require the condition [e.g., *Mathews and Walker*, 1970]

$$\left.\frac{\partial \mathcal{F}}{\partial \Phi_{n,2}}\right|_{Q_2 = 0} = 0. \tag{30}$$

Applied to (29), equation (30) then gives a Neumann boundary condition, written explicitly in terms of $\bar{f}$ and $\{u, K, L\}$,

$$\left.\left(D_{uK}\frac{\partial \bar{f}}{\partial u} + D_{KK}\frac{\partial \bar{f}}{\partial K} + D_{KL}\frac{\partial \bar{f}}{\partial L}\right)\right|_{K = 0} = 0. \tag{31}$$

In the SDE language, this boundary condition indicates that, on the $K = 0$ boundary, the stochastic processes are reflected locally about the unit vector $\boldsymbol{\gamma}(\boldsymbol{\xi}) \propto (D_{uK}, D_{KK}, D_{KL})^{\intercal}$. From another point of view, the left-hand side of (31) is just the negative of the K component of the PSD diffusive flux (in a curvilinear coordinate system). Therefore, the physical meaning of (31) is that the PSD is not allowed to flow across the $K = 0$ surface, which is natural since the region of "negative K" is physically meaningless. This conclusion confirms the criteria of adjoint operators discussed in Appendix A2 (equations (A25) and (A26)).

Neumann-type boundaries require that the diffusion tensor must be nondegenerate in the direction of $\boldsymbol{\gamma}$ on $\partial_2 D$, which means the following quadratic form inequality must always hold [*Freidlin*, 1985]

$$\boldsymbol{\gamma}^{\intercal}(\boldsymbol{\xi})\mathbf{a}(t, \boldsymbol{\xi})\boldsymbol{\gamma}(\boldsymbol{\xi}) > 0, \quad \boldsymbol{\xi} \in \partial_2 D. \tag{32}$$

For the $K = 0$ boundary with boundary condition (31) to be nondegenerate, Appendix B gives proof that the sufficient and necessary condition is $D_{KK} > 0$ on the boundary. However, by Taylor expanding magnetic intensity along a field line about its minimum in terms of arc length, it can be proved that $\partial K/\partial\alpha_0$ vanishes at $K = 0$ even in realistic magnetic fields. Since $\partial K/\partial p$ is zero too, the transformation matrix in (25) is singular at $K = 0$, and the transformed D_{uK} and D_{KK} are both zero there. This is an artifact of calculating the u-K diffusion coefficients from the α_0-p coefficients. To circumvent this artifact, we invoke small but nonzero bounce-resonance effects in the equatorial plane (as argued physically by *Roberts* [1969]) to give a nonzero D_{KK}. Then, with zero D_{KL}, the boundary condition (31) reduces to

$$\left.\frac{\partial \bar{f}}{\partial K}\right|_{K = 0} = 0, \tag{33}$$

and the reflection vector $\boldsymbol{\gamma}$ is now simply $(0, 1, 0)^{\mathsf{T}}$, i.e., the normal to the $K = 0$ boundary. The Salammbô code [*Beutier and Boscher*, 1995] used (without explanation) the Neumann boundary condition $\partial\bar{f}/\partial J_2 = 0$ at $J_2 = 0$, which is equivalent to (33). Lastly, we emphasize that the variational method of determining the Neumann boundary condition and the discussion of the nondegeneracy condition are general; they apply to the Fokker-Planck equation in any other coordinate variables.

2.4. Numerical Implementation and Validation

Implementation of the SDE method starts from numerically integrating the SDE (12) in time. In the interior of the domain, we use the Euler-Maruyama scheme [*Kloeden and Platen*, 1992] as a time discrete approximation of the SDE:

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \mathbf{b}(t - s_n, \mathbf{X}_n)\Delta s + \boldsymbol{\sigma}(t - s_n, \mathbf{X}_n)\Delta \mathbf{W}_n, \tag{34}$$

where the equidistant time step $\Delta s = t/N = s_{n+1} - s_n$ with N the total number of steps and $\mathbf{X}_n$ is the approximation of $\mathbf{X}_s^t$ at time s_n. The matrix $\boldsymbol{\sigma}$ is calculated from (3) by Cholesky decomposition of $\mathbf{a}$, and the result is a lower triangular matrix. The increment of the Wiener process $\Delta\mathbf{W}_n$ is generated by

$$\Delta \mathbf{W}_n = \sqrt{s_{n+1} - s_n} \begin{pmatrix} \mathcal{N}_1 \\ \mathcal{N}_2 \\ \mathcal{N}_3 \end{pmatrix}, \tag{35}$$

where $\mathcal{N}_1$, $\mathcal{N}_2$, and $\mathcal{N}_3$ are three standard Gaussian random numbers each with zero mean and unit variance, which are transformed from three independent uniform deviations via the Box-Muller algorithm [*Press et al.*, 1992]. The uniform deviations are generated by the Dynamic Creator of Mersenne Twisters [*Matsumoto and Nishimura*, 1998, 2000], a fast multiple-stream pseudorandom number generator based on Mersenne primes, to ensure their mutual statistical independence. For evaluation of functional expectations like ours, where only the statistical distribution of random walks matters but not their individual paths, the convergence of the Euler-Maruyama scheme is of order 1.0, meaning that the mean error of the scheme is proportional to the first power of Δs [*Kloeden and Platen*, 1992].

Stochastic paths are nowhere differentiable. Therefore, integration of (34) only provides an approximation to the stochastic path at the time discretization steps s_n. An approximation of the entire path over the continuous time interval $0 \le s \le t$ is obtained by linear interpolations between $\mathbf{X}_n$ and $\mathbf{X}_{n+1}$, i.e., connecting them by straight line segments. On Dirichlet-type boundaries, the stopping position of an Itô process is then determined by the intersection between the boundary and this linear-interpolated path. The accuracy of this stopping strategy is of order 0.5 [*Gobet*, 2001]. Sophisticated order 1.0 schemes are also available [e.g., *Gobet*, 2001], but their implementations often require much more complicated computation near the boundary. On the Neumann-type boundary, we use the Symmetrized Euler Scheme (order 1.0) [*Bossy et al.*, 2004] to reflect the Itô processes.

The functional expectation $\mathbb{E}[F^{t,\xi}]$ is evaluated by the arithmetic mean of the functional values obtained from various stochastic paths. In practice, only a finite number of Itô processes can be simulated, which introduces statistical fluctuations to the result. Consequently, the error of the SDE method comes from two parts: one is the systematic error intrinsic to the numerical schemes, which can be directly reduced by using smaller step size (but above the round-off error limit); the other is the statistical error from the finite number of simulations, which is proportional to $1/\sqrt{\Gamma}$ with sufficiently large total number of simulations Γ, and has zero expectation. For the systematic error, schemes of higher order than Euler can be adopted to improve the convergence, but with a penalty of more intricate computations; and for the statistical error, variance reduction techniques, which use cleverer but more expensive ways to estimate expectation than the arithmetic mean, are possible mitigations [*Kloeden and Platen*, 1992]. Nonetheless, with the current schemes, our code is still able to give assessment of confidence intervals of the model solutions by statistical analysis of the simulations; or conversely, to make the model adaptive in the number of simulations to meet a prescribed confidence interval.

The REM has been tested against known solutions for 1-D radial diffusion (by setting *u-K* diffusion coefficients to zero) and for 2-D chorus wave diffusion (by setting the radial diffusion coefficient to zero). The number of steps and number of simulations vary in these tests to balance accuracy versus computational

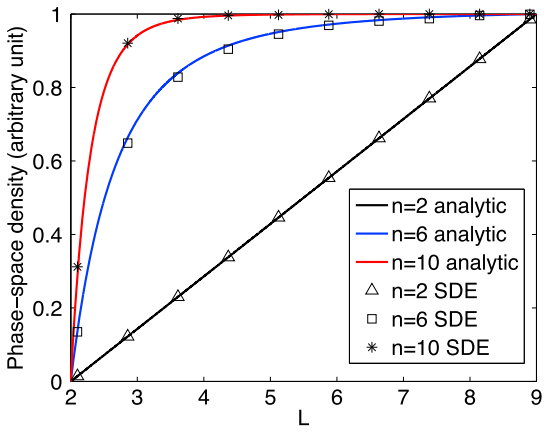

Figure 3. The SDE solutions (discrete marks) and the analytic solutions (solid curves) of equation (36) for $n = 2, 6, 10$ (see text). The solid black line coincides with the initial condition.

effort. In general, they are in the order of 10^3 and 10^4, respectively. In the radial diffusion test, we solve the steady state diffusion equation

$$0 = L^2 \frac{\partial}{\partial L}\left(\frac{D_{LL}}{L^2}\frac{\partial \bar{f}}{\partial L}\right) \tag{36}$$

with $D_{LL} \propto L^n$ for $n = 2, 6, 10$, a linear initial distribution of $\bar{f}$ from 0 at $L = 2$ to 1 at $L = 9$, and fixed-value Dirichlet boundary conditions. Figure 3 shows the comparison of the SDE results with analytic solutions of equation (36). For $n = 2$, equation (36) is degenerate, and the solution remains the same as the initial condition, as shown in Figure 3 by the black straight line. The SDE results are clearly in agreement with the analytic solutions.

In the chorus wave diffusion test, we compare the SDE solutions with those from *Albert and Young* [2005], using the same initial and boundary conditions and the same diffusion coefficients. *Albert and Young* [2005] solved the α_0-p electron Fokker-Planck equation at $L = 4.5$, with off-diagonal diffusion components, by a finite-difference method with matrix diagonalization. To compare, we reproduce the fluxes for $E = 0.5$ MeV and $E = 2.0$ MeV electrons after 0.1 day and 1.0 day, with α_0 ranging from 6° to 88° with 2° spacing (see Figure 4). Statistical variance is clearly observed in the 2.0 MeV and 1.0 day case, making the SDE solution fluctuate around its mean value. Within numerical errors associated with each of the methods, they yield solutions in very good agreement.

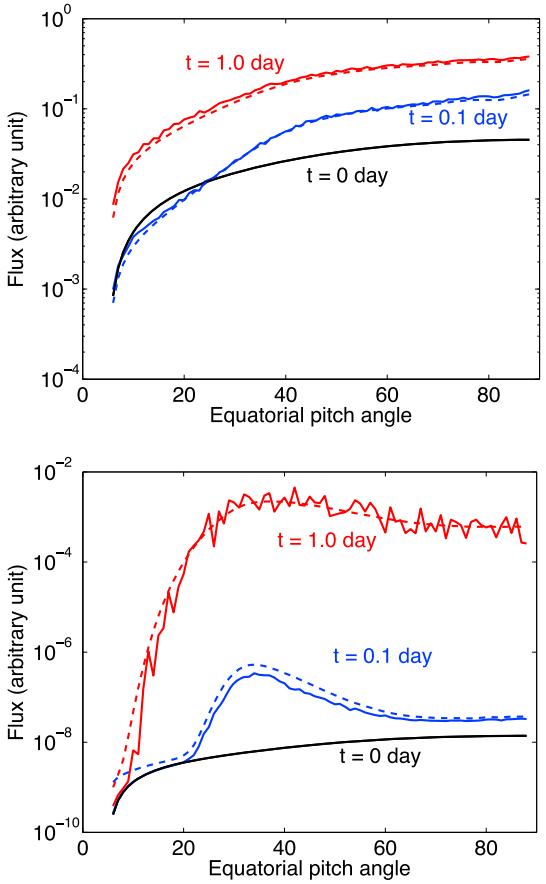

Figure 4. Comparisons between solutions from the SDE method (solid lines) and the *Albert and Young* [2005] method (dashed lines) for (top) 0.5 MeV and (bottom) 2.0 MeV electron fluxes after $t = 0.1$ day (blue) and $t = 1.0$ day (red) diffusion. Solid black lines show the initial condition.

3. Model Application to the October 2002 Storm

Motivated by observations of the PSD as a function of adiabatic invariants made using NASA's Polar satellite, the Los Alamos National Laboratory geostationary Earth orbit (GEO) satellites and the Global Positioning System (GPS) constellation [*Chen et al.*, 2007; *Koller et al.*, 2007], we apply the REM to a simplified simulation of the PSD increase at GPS orbit during a moderate high-speed stream (HSS) storm with minimum $Dst \sim -70$ nT from 15 October (Day 288) to 20 October (Day 293) 2002. *Chen et al.* [2007] gave a thorough description of the observational data during this storm. In that work, for $M = 2083$ MeV/G and $K = 0.03G^{1/2}R_E$, they concluded that during the recovery phase of this storm chorus wave acceleration played an essential role in replenishing the outer belt and generating the PSD peak observable at GEO.

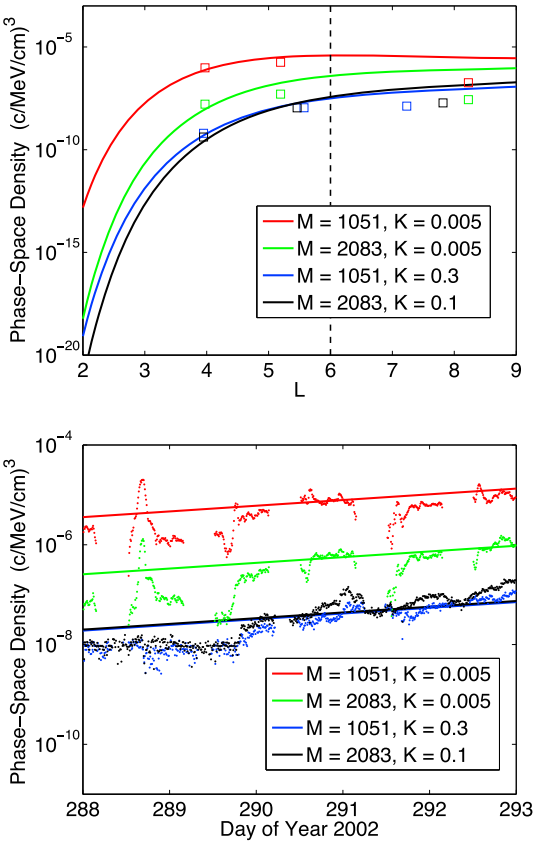

Figure 5. (top) Initial condition and (bottom) the outer L boundary condition used in the October 2002 HSS storm simulation. In Figure 5 (top), markers represent satellite data at the beginning of 15 October (Day 288) from GPS ($L \approx 4$), GEO ($L \approx 6$), and Polar ($L \approx 8$). Marker colors identify the corresponding M and K values to the fitted curves. The dashed line at $L = 6$ denotes the location of the outer L boundary. In Figure 5 (bottom), GEO observations are shown in colored dots. Simplified exponential increases are drawn as straight lines.

Figure 5 summarizes the initial condition and the outer L boundary condition in our simulation. To set up an initial 3-D PSD distribution, we fit the PSD data at a list of M values (167, 462, 1051, and 2083 MeV/G) and K values (0.005, 0.01, 0.03, 0.1, and 0.3 $G^{1/2}R_E$) in three L groups, corresponding to GPS, GEO, and Polar orbits, respectively. Figure 5 (top) illustrates the fitted initial PSD radial profile at the four pairs of M and K that we simulate. On the outer L boundary at geosynchronous distance ($L = 6$), we simplify the observed PSD temporal variations as exponential increases from the beginning to the end of the storm, as shown in Figure 5 (bottom). On the inner L boundary ($L = 2$), PSD is kept the same as its initial value over time. To make the boundary conditions consistent, the exponential increase is also carried onto the E_{min} boundary in such a way that the increase rate diminishes as L decreases from 6 to 2. Thus, on the E_{min} boundary, at $L = 6$, the increase is at the full rate; whereas at $L = 2$, the PSD is unchanged.

During the storm, the Kp index varied only a small amount about an average value of about 3; for simplicity, we fix $Kp = 3$ in assigning the diffusion coefficients. D_{LL} uses the *Brautigam and Albert* [2000] magnetic radial diffusion coefficient (henceforth denoted D_{LL}^{M}[B and A]), because in the outer belt it dominates over the electric field counterpart (D_{LL}^{E}[B and A]). Chorus wave diffusion coefficients in the $\{u, K, L\}$ coordinates are transformed from the British Antarctic Survey (BAS) drift- and bounce-averaged diffusion matrix, by the method described in section 2.2 and assuming dipole field geometry. The BAS diffusion matrix was calculated using the Pitch Angle and Energy Diffusion of Ions and Electrons code [*Glauert and Horne*, 2005] and plasma wave observations from the CRRES spacecraft [*Meredith et al.*, 2001, 2003] as described in *Varotsou et al.* [2005]. *Varotsou et al.* [2005] considered the effects of equatorial chorus, defined as chorus wave power within ±15° of the magnetic equator. In this paper we have included wave power up to ±30° to include the effects of strong midlatitude chorus that is observed on the dayside [*Tsurutani and Smith*, 1977; *Meredith et al.*, 2001, 2012]. Further, we have extrapolated for energies between 3 and 4 MeV. Plots of the transformed chorus wave diffusion rates are shown in Figure 6. Inward of geosynchronous orbit, the diffusion rates grow monotonically with L. In contrast to the $\{\alpha_0, p\}$ coordinates, where the off-diagonal term often changes sign, the transformed D_{uK} values are all negative throughout the entire L range of the BAS diffusion matrix for this Kp level.

Model simulations are performed to obtain PSD at four phase space positions at $L = 4$, with (M, K) values of $(1051, 0.005)$, $(2083, 0.005)$, $(1051, 0.3)$, and $(2083, 0.1)$. In dipole field, these phase space positions correspond respectively to energies (in MeV) and α_0 values of $(1.85, 81.1°)$, $(2.78, 81.1°)$, $(3.26, 37.6°)$, and $(3.47, 54.4°)$. At each position, three independent simulations are carried out: one with radial diffusion only, one with chorus wave diffusion only, and one with the two mechanisms combined. The results are given in Figure 7.

In Figures 7a and 7b, where α_0 values are near 90°, chorus waves have very little effect on the PSD variation, as seen from the almost flat "chorus only" curves. This could also be appreciated from the close alignment

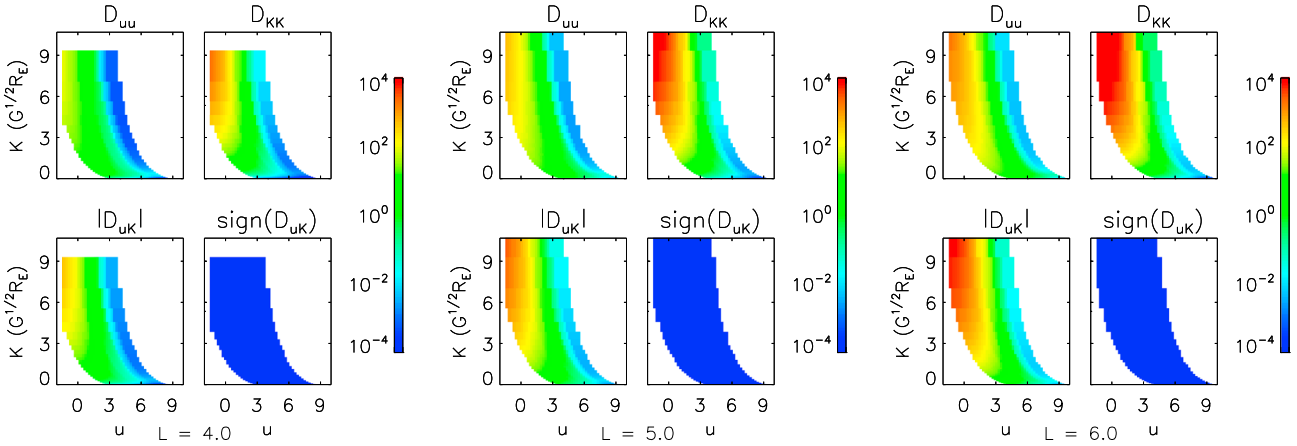

Figure 6. Normalized *u-K* diffusion rates (day^{-1}) transformed from the BAS drift- and bounce-averaged chorus wave diffusion matrix at three *L* values. D_{uu} is dimensionless since *u* is dimensionless. D_{uK} and D_{KK} are normalized against $K = 1\ G^{1/2}R_E$. The uniform blue color in each sign(D_{uK}) panel indicates negative regions of D_{uK}. Positive regions would have been red.

between the "radial only" and the "combined" curves. As argued in *Horne et al.* [2005] and *Shprits* [2009], this is because energy diffusion by chorus waves is less effective at large pitch angles (i.e., ~80°). In Figures 7a and 7b radial diffusion curves largely follow the observed data points. Thus, the equatorial PSD increases in this HSS storm are consistent with mainly radial diffusion acting on the particles.

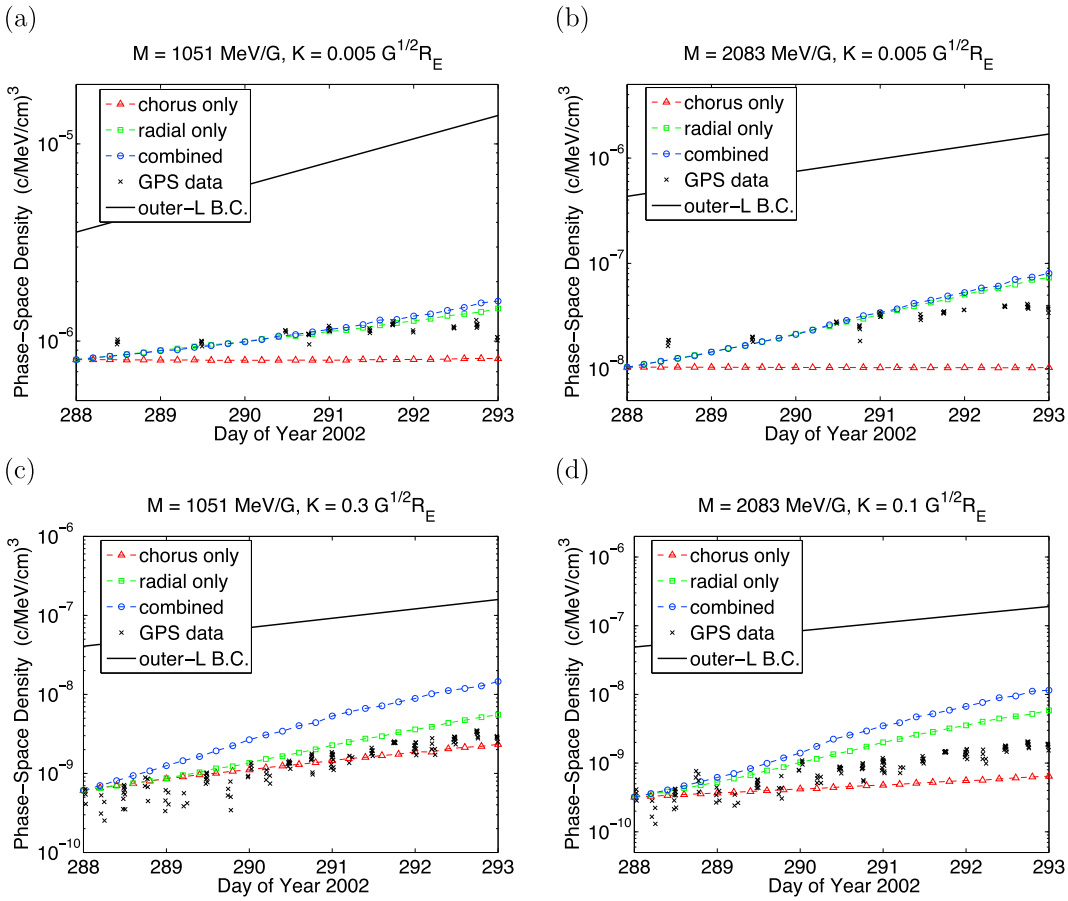

Figure 7. Comparison of the simulated PSD temporal changes with GPS observations at $L = 4$. In each of the panels, observed data are marked with black crosses. Black straight lines indicate the driving boundary conditions at $L = 6$ at each specified *M* and *K*. Note that the ordinate scales are different in these panels.

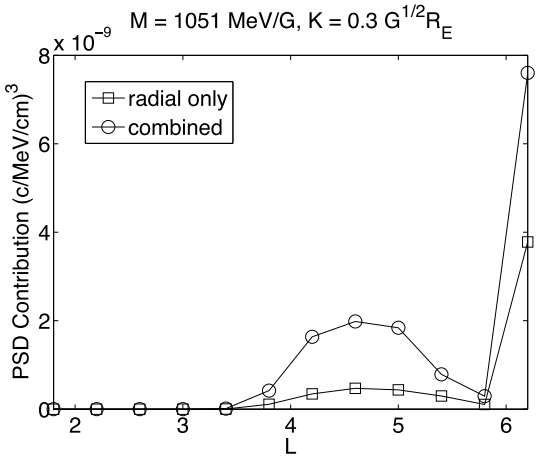

Figure 8. Radial distributions of the PSD contribution for the solution at $M = 1051$ MeV/G, $K = 0.3\ G^{1/2}R_E$, and $L = 4$ at $t = 5$ days, from the radial-only (square) and the combined (circle) simulations. The contributions are binned in L with bin size 0.4. The first and the last bins are cumulative; thus, the first marker represents all contributions from $L < 2.0$, and the last represents all contributions from $L \geq 6.0$. The interior markers reside at the center of each bin and give the value in that bin.

In Figures 7c and 7d, where α_0 values are in the range of peak chorus wave effectiveness (30°–60°, according to *Horne et al.* [2005]), chorus wave effects are more significant. Nevertheless, chorus wave diffusion alone is insufficient to explain the observed PSD increases. Radial diffusion alone, which reproduces the mild increase of PSD (less than 1 order of magnitude) in the near-equatorial cases, overestimates the observed increases here. As a result, the PSD variations of the combined diffusion are even more exaggerated. Apart from the simplified inputs to this simulation, we suggest two possible reasons for this overestimation. First, $D_{LL}^{M}[B \text{ and } A]$ is derived for equatorially bouncing electrons, whereas radial diffusion rates caused by ULF waves are α_0 dependent. From the estimate given in *Schulz and Lanzerotti* [1974], it follows that the radial diffusion rate in the α_0 range of 35°–55° is only about 0.2–0.4 times that at 90°. Therefore, using $D_{LL}^{M}[B \text{ and } A]$ will result in exaggerated radial diffusion at small α_0 values. Second, $L = 4$ is likely on the edge of the plasmasphere for this moderate storm, and electron loss from electromagnetic ion cyclotron (EMIC) waves and plasmaspheric hiss waves are not included in our simulation. At this L, effects of magnetopause shadowing or drift-shell splitting are expected to be negligible.

In addition to solving the diffusion equation, the SDE method also provides insight into the diffusion process that is not easily obtained in other methods. Take, for example, the solutions at $M = 1051$ MeV/G, $K = 0.3\ G^{1/2}R_E$, and $t = 5$ days (Day 293 in Figure 7c), where solutions of radial-only and combined diffusions are well separated. By recording the diffusion times of the stochastic processes, it suggests that the electrons take about 1.2 days to diffuse from GEO to GPS orbit in our simulation. On the other hand, by binning the stopping position distribution of the time backward stochastic processes, we are able to determine the PSD contribution from different phase space regions to the solution; in other words, to determine a phase space distribution of the energetic electron sources. In Figure 8, we illustrate these "source distributions" by plotting radial profiles of the source PSD contribution in the radial-only and the combined simulations. The radial profile in the chorus-only simulation is trivial, since it is simply a spike at $L = 4$.

As illustrated in the figure, at the end of the storm at GPS orbit, the majority of the PSD comes from the outer L boundary in both simulations. Radial diffusion thus plays an indispensable role in energizing all these electrons. However, the outer L boundary contributes about twice the amount of PSD in the combined simulation as in the radial-only simulation, demonstrating that with local acceleration turned on, the outer L boundary becomes more efficient in providing source electrons. In a synergistic manner, more low-energy electrons are accelerated by chorus waves while diffusing inward; these particles could not have been energized as much with radial diffusion only. A hump is clearly observed in the combined curve between $L = 3.8$ and 5.4, whereas the radial-only curve remains almost flat in this range. This difference indicates the regions where chorus wave acceleration is most important for the electron population at $L = 4$. We note that, from Figure 6, although chorus wave diffusions are stronger in larger L regions, their effect on the $L = 4$ solution is nevertheless radially localized. This is due to both the length of diffusion time and the shape of the initial PSD distribution.

4. Summary and Discussion

Stochastic differential equation (SDE) theory, which relates a diffusion problem to stochastic processes and thereby facilitates solutions to the diffusion equation from functional expectations of the stochastic processes, is presented in this work. Compared to finite-difference-based methods of solving diffusion

equations, the SDE method has several advantages when applied to radiation belt simulations. First and foremost, the SDE method does not need a numerical grid. This allows the method to deal with boundary geometries with great complexity, which facilitates the use of electron adiabatic invariant coordinates. Consequently, adiabatic variations are completely separated from nonadiabatic diffusion in the computations, and yet adiabatic losses due to the change of loss cone can still be easily modeled (as a time-varying boundary). The absence of a grid also makes the method more efficient for obtaining solutions at a limited number of phase space positions, which is often the case with spacecraft observations, since the method only samples the phase space regions that contribute to the solutions. Even in the opposite situation, where simulations are extended into larger regions, the method can be parallelized very efficiently to achieve the necessary coverage. Second, the SDE method can solve fully multidimensional diffusion problems. Hence, it provides a computational tool to assess diffusion effects of the electron Shabansky orbits and drift-shell splitting, which manifest themselves in the *M*-*L* and *K*-*L* components of the diffusion tensor. Last but not least, because solutions in the SDE method are obtained by summing over contributions from stochastic processes and then taking the average, the method is very robust. In particular, it can tolerate ranges of PSD of several orders of magnitude, and it never gives a negative PSD.

Based on SDE theory, we have formulated the SDE representation of the electron Fokker-Planck equation in the adiabatic invariant coordinates $\{u, K, L\}$, where $u \equiv \ln(M)$. We have also mathematically clarified the form of the Neumann boundary condition at $K = 0$ (or $\alpha_0 = 90°$), which has often been simply implemented as $\partial\bar{f}/\partial\alpha_0 = 0$ (this form is only correct when off-diagonal diffusion components are zero on the boundary). A fully 3-D numerical code in adiabatic invariant coordinates (named REM) has been constructed and successfully tested against known solutions of 1-D radial diffusion and 2-D chorus wave diffusion with off-diagonal components.

The REM was then applied to simulating with simplified conditions the PSD increase observed at GPS locations during the October 2002 high-speed stream storm. It has been argued in *Chen et al.* [2007] that chorus wave acceleration was active during this storm and was responsible for the PSD peak observed on geosynchronous orbit. For near-equatorial electrons (Figures 7a and 7b), the model gives reasonably accurate simulations of the observed GPS increase, consistent with radial diffusion from the outer *L* boundary and weak chorus wave diffusion at large equatorial pitch angles. At smaller equatorial pitch angles (Figures 7c and 7d) radial diffusion, and consequently the combined diffusion, overestimates the increase of PSD at GPS locations. Possible explanations could be the neglect of equatorial pitch angle dependence in the radial diffusion coefficients, or the neglect of electron loss from pitch angle scattering by hiss and EMIC waves in the model.

The SDE method allows us to extract information such as the electron diffusion time and PSD source distribution, by analyzing the statistics of the stochastic processes. Our model reveals the intrinsic differences between radial and combined diffusion from a new point of view, by plotting the radial distribution of the source PSD. In the case exemplified in this work, although both diffusion scenarios have the majority of the PSD coming from the outer boundary, with chorus waves present the efficiency of the outer boundary in providing seed electron populations is increased by a factor of 2. Chorus waves also generate a peak to the PSD source distribution well within the outer radiation belt, indicating the *L* ranges in which they are of particular interest.

In addition to its use as an independent model, the REM in the $\{u, K\}$ coordinates can also be coupled with MHD-particle transport simulations [e.g., *Elkington et al.*, 2002, 2004] to construct 4-D models that incorporate both diffusive and nondiffusive radiation belt dynamics, including both cyclotron frequency and drift frequency wave-particle interactions. With the unparalleled coverage and resolution from NASA's Van Allen Probes mission, the REM is a powerful and promising tool for investigating and simulating physical processes in the radiation belts.

Appendix A: Mathematical Formulation of the Stochastic Differential Equation Method

In this appendix we present enough of the theory of stochastic differential equations (SDEs) to enable us to apply the SDE method to solution of the radiation belt diffusion equations. For simplicity, these formulations

are initially developed in Cartesian coordinates, and generalization into curvilinear coordinates is outlined at the end of section A2.

A1. The Wiener Process and Itô Stochastic Calculus

Just as calculus deals with differentials, integrals, and differential equations, stochastic calculus deals with these objects, but with stochastic processes included. A stochastic process, X_t, is a time sequence that takes random values $\{x_0, x_1, x_2, \dots\}$ at time instances $0 \le t_0 < t_1 < t_2 <, \dots$. Throughout this article, we will assign capital letters to stochastic processes and use the corresponding lowercase letter to denote their values.

For simplicity, let us start from a 1-D time-homogeneous SDE of the form

$$\mathrm{d}X_t(\omega) = b\left(X_t(\omega)\right)\mathrm{d}t + \sigma\left(X_t(\omega)\right)\mathrm{d}W_t(\omega), \tag{A1}$$

in which $X_t(\omega)$ is the unknown stochastic process and ω is one sample in the sample space Ω. For example, if Ω is a set of pollen grains undergoing random motions, ω refers to one grain among them, and the stochastic process $X_t(\omega)$ is a time sequence of the spatial positions of that grain. The terminology "time-homogeneous" means the coefficients $b(X_t)$ and $\sigma(X_t)$ depend only on the value of the stochastic process, but not explicitly on t. $W_t(\omega)$ is a 1-D stochastic process called the Wiener process of ω. The Wiener process W_t is a Gaussian random process with initial value $W_0 = 0$ (with probability 1) and independent increments $\mathrm{d}W_t = W_{t+\mathrm{d}t} - W_t$, such that

$$\mathbb{E}[W_t] = 0, \qquad \mathbb{E}[W_t^2] = t, \tag{A2}$$

where the symbol $\mathbb{E}$ denotes expectation over Ω [*Øksendal*, 1998; *Kloeden and Platen*, 1992]. For different ω values, the corresponding Wiener processes $W_t(\omega)$ must be statistically independent.

The stochastic process $X_t(\omega)$ can be regarded as a mapping from the variable set $\{t, \omega\}$ to a real number x. For each fixed t, the mapping from ω to $X_t(\omega)$ defines a random variable; on the other hand, fixing ω we can consider $X_t(\omega)$ as a function of t, which is called a path of $X_t(\omega)$. It can be proved that such a function is continuous in t everywhere with probability 1, but nowhere differentiable [*Øksendal*, 1998; *Gardiner*, 2004]. With these concepts in mind, for brevity we will henceforth suppress the variable ω except when that results in ambiguity.

The solution of the 1-D SDE (A1) is formally represented by

$$X_t = X_0 + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s. \tag{A3}$$

On the right-hand side of equation (A3), the first integral is a Riemann integral over time, i.e., its value converges as we repeatedly divide the integration interval, no matter where we choose to evaluate the integrand within each subinterval. The second integral is a stochastic integral. In contrast to Riemann integral, its value does depend on the choices of s within subintervals because of the nondifferentiablity of W_s; this implies that $\mathrm{d}W_s$ has (with probability 1) infinite variation on every subinterval [*Øksendal*, 1998; *Freidlin*, 1985]. A widely used and well-studied choice is to evaluate the integrand at the beginning of every subinterval (so its value does not depend on the future), and this defines the so-called Itô integral

$$\int_0^t \sigma(X_s)\mathrm{d}W_s \equiv \sum_{n=0}^{N} \sigma(X_{s_n})(W_{s_{n+1}} - W_{s_n}), \qquad \max(s_{n+1} - s_n) \to 0. \tag{A4}$$

Other choices of s define different types of stochastic integral [e.g., *Kloeden and Platen*, 1992]. We use the Itô integral because of its close relation to diffusion processes [*Freidlin*, 1985].

Aside from linearity, the Itô integral has the particular properties that [*Øksendal*, 1998; *Kloeden and Platen*, 1992], for $0 \le S \le T$ and arbitrary (measurable) functions of time and stochastic processes $g(t, \omega)$ and $h(t, \omega)$,

$$\mathbb{E}\left[\int_S^T g\,\mathrm{d}W_t\right] = 0, \tag{A5}$$

$$\mathbb{E}\left[\int_S^T g\,\mathrm{d}W_t \int_S^T h\,\mathrm{d}W_t\right] = \int_S^T \mathbb{E}[gh]\mathrm{d}t. \tag{A6}$$

Property (A6) reveals that the differential $(dW_t)^2$ is of the same infinitesimal order as dt in the mean square sense. Combining (A5) and (A6), we may also write

$$\mathbb{E}\left[\int_S^T g\, dW_{ti} \int_S^T h\, dW_{tj}\right] = \delta_{ij} \int_S^T \mathbb{E}[gh] dt, \tag{A7}$$

in which δ_{ij} is the Kronecker delta. When $i \neq j$, W_{ti} and W_{tj} are independent Wiener processes so that the left-hand side of (A7) splits to the product of two expectations of each individual Itô integral, both equal to zero following (A5). In the mean square sense, we symbolically denote this $dW_{ti} dW_{tj} = \delta_{ij} dt$.

In multiple dimensions, the SDE (A1) becomes a vector equation

$$d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)dt + \boldsymbol{\sigma}(\mathbf{X}_t)d\mathbf{W}_t, \tag{A8}$$

where $\mathbf{X}_t$ is an n-D stochastic process, $\mathbf{b}$ is an n-D vector, $\boldsymbol{\sigma}$ is an $n \times m$ matrix, and $\mathbf{W}_t$ is an m-D Wiener process with each of its components an independent 1-D Wiener process.

Now let us consider a function $f(\mathbf{x})$, with continuous second-order partial derivatives, acting on $\mathbf{X}_t$. Property (A7) allows differentiation of $f(\mathbf{X}_t)$ in the mean square sense, with the aid of the SDE (A8), which yields

$$\begin{aligned} df(\mathbf{X}_t) &= \frac{\partial f}{\partial x_i} dX_{ti} + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j} dX_{ti} dX_{tj} \\ &= \frac{\partial f}{\partial x_i}(b_i dt + \sigma_{ij} dW_{tj}) + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j} \sigma_{ik} \sigma_{jk} dt, \end{aligned} \tag{A9}$$

where we have employed the summation convention. Collecting infinitesimals of the same order and defining the differential operator

$$\mathcal{L}f(\mathbf{x}) \equiv \frac{1}{2} a_{ij}(\mathbf{x}) \frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x}) + b_i(\mathbf{x}) \frac{\partial}{\partial x_i} f(\mathbf{x}), \tag{A10}$$

with

$$a_{ij} \equiv \sigma_{ik} \sigma_{jk} = (\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathsf{T}})_{ij}, \tag{A11}$$

equation (A9) gives rise to the following Itô formula [e.g., *Øksendal*, 1998], which describes the SDE of the new stochastic process $f(\mathbf{X}_t)$:

$$df(\mathbf{X}_t) = \mathcal{L}f(\mathbf{X}_t)dt + \sigma_{ij}(\mathbf{X}_t) \frac{\partial}{\partial x_i} f(\mathbf{X}_t) dW_{tj}. \tag{A12}$$

Note that a matrix $\mathbf{a}$ given by (A11) must be symmetric and semipositive definite (meaning all its eigenvalues are nonnegative and so is its determinant).

A2. Relation Between the Stochastic Differential Equation and the Fokker-Planck Equation

Having introduced stochastic calculus, we establish in this subsection the relation between the SDE (A8) and the corresponding Fokker-Planck equation through manipulations of the Itô formula (A12) and the differential operator $\mathcal{L}$. Subsequently, this relation gives a way to obtain solutions of the Fokker-Planck equation.

Integration of the Itô formula (A12) over time yields

$$f(\mathbf{X}_t) = f(\boldsymbol{\xi}) + \int_0^t \mathcal{L}f(\mathbf{X}_s) ds + \int_0^t \sigma_{ij}(\mathbf{X}_s) \frac{\partial f}{\partial x_i} dW_{sj}, \tag{A13}$$

in which $\boldsymbol{\xi} \equiv \mathbf{x}_0$ marks the initial position of the Itô process. Taking expectation values of both sides of (A13) and using property (A5) of the Itô integral gives

$$\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)] = f(\boldsymbol{\xi}) + \int_0^t \mathbb{E}^{\boldsymbol{\xi}}[\mathcal{L}f(\mathbf{X}_s)] ds, \tag{A14}$$

where the superscript of $\mathbb{E}^{\boldsymbol{\xi}}$ denotes the common initial position of the stochastic processes. If we let t approach 0 (from above), the integrand in equation (A14) can be approximated by its value at $s = 0$ and $\mathbf{X}_s$

takes the nonrandom initial value $\boldsymbol{\xi}$. In this situation, $\mathbb{E}$ can be removed from the integrand, and (A14) then gives another expression of $\mathcal{L}$ in terms of the limit

$$\mathcal{L}f(\boldsymbol{\xi}) = \lim_{t\downarrow 0} \frac{\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)] - f(\boldsymbol{\xi})}{t}. \tag{A15}$$

With this interpretation of $\mathcal{L}$ and regarding $\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_s)]$ as a function of $\boldsymbol{\xi}$, the integrand in equation (A14) can be transformed as follows:

$$\begin{aligned}\mathbb{E}^{\boldsymbol{\xi}}[\mathcal{L}f(\mathbf{X}_s)] &= \lim_{r\downarrow 0} \frac{\mathbb{E}^{\boldsymbol{\xi}}\left[\mathbb{E}^{\mathbf{X}_s}[f(\mathbf{X}_r)]\right] - \mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_s)]}{r} \\ &= \lim_{r\downarrow 0} \frac{\mathbb{E}^{\boldsymbol{\xi}}\left[\mathbb{E}^{\mathbf{X}_r}[f(\mathbf{X}_s)]\right] - \mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_s)]}{r} \\ &= \mathcal{L}\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_s)],\end{aligned} \tag{A16}$$

where the interchange of "$\mathbf{X}_s$" and "$\mathbf{X}_r$" is permissible due to the history-independent (Markovian) property of the stochastic process described by our SDE [*Øksendal*, 1998]. Inserting (A16) into (A14) and taking a time derivative, we have

$$\frac{\partial}{\partial t}\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)] = \mathcal{L}\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)]. \tag{A17}$$

Denoting $\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)]$ as $u(t, \boldsymbol{\xi})$, we have thus derived the PDE that the functional expectation of the stochastic processes satisfies. Together with the initial condition, they form an initial value problem:

$$\frac{\partial}{\partial t}u(t, \boldsymbol{\xi}) = \mathcal{L}u(t, \boldsymbol{\xi}), \tag{A18}$$

$$u(0, \boldsymbol{\xi}) = \mathbb{E}^{\boldsymbol{\xi}}[f(\boldsymbol{\xi})] = f(\boldsymbol{\xi}). \tag{A19}$$

Equation (A18) is the so-called Kolmogorov backward equation associated with the Itô process (A8).

From another point of view, consider the transition probability density $p_t(\mathbf{x}, \mathbf{y})$ of the Itô process $\mathbf{X}_t$, i.e., the probability density of a stochastic process traveling from $\mathbf{x}$ to $\mathbf{y}$ in time period t [*Øksendal*, 1998]. By the definition of p_t, the expectation $\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)]$ can be calculated from

$$\mathbb{E}^{\boldsymbol{\xi}}[f(\mathbf{X}_t)] = \int_D f(\boldsymbol{\eta})p_t(\boldsymbol{\xi}, \boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta}, \tag{A20}$$

where the integration is over the domain D in which f is defined. In light of this, the transition probability representation of relation (A16) reveals the useful property of $\mathcal{L}$

$$\int_D p_t(\boldsymbol{\xi}, \boldsymbol{\eta})\mathcal{L}_{\boldsymbol{\eta}}f(\boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta} = \mathcal{L}_{\boldsymbol{\xi}}\int_D p_t(\boldsymbol{\xi}, \boldsymbol{\eta})f(\boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta}, \tag{A21}$$

where the different subscripts of $\mathcal{L}$ indicate the different variables it is operating on.

Applying (A20) and (A21) to the Kolmogorov backward equation (A17) gives the integral equation

$$\int_D f(\boldsymbol{\eta})\frac{\partial}{\partial t}p_t(\boldsymbol{\xi}, \boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta} = \int_D p_t(\boldsymbol{\xi}, \boldsymbol{\eta})\mathcal{L}_{\boldsymbol{\eta}}f(\boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta}. \tag{A22}$$

To further transform, it is desirable to introduce the adjoint operator $\mathcal{L}^*$ of $\mathcal{L}$ [*Øksendal*, 1998; *Arfken and Weber*, 1995], defined as

$$\mathcal{L}^*\phi(\mathbf{x}) \equiv \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}\left[a_{ij}(\mathbf{x})\phi(\mathbf{x})\right] - \frac{\partial}{\partial x_i}\left[b_i(\mathbf{x})\phi(\mathbf{x})\right]. \tag{A23}$$

For diffusion processes, a_{ij} and b_i are not independent but are related by $\mathbf{b} = \frac{1}{2}\nabla \cdot \mathbf{a}$ as components of the rank two and one tensors, respectively [e.g., *Haerendel*, 1968]. Even in more general cases, we can always split b_i into a diffusive part and an advective part, such that

$$b_i = \frac{1}{2}\frac{\partial a_{ij}}{\partial x_j} + h_i. \tag{A24}$$

Using this relation, we have, for arbitrary functions u and v with continuous second-order partial derivatives, the Green's identity for these operators

$$v\mathcal{L}u - u\mathcal{L}^*v = \frac{\partial}{\partial x_i}\left(\frac{1}{2}va_{ij}\frac{\partial u}{\partial x_j} - \frac{1}{2}ua_{ij}\frac{\partial v}{\partial x_j} + uvh_i\right)$$
$$= \frac{\partial}{\partial x_i}(uF_i^v - vF_i^u + uvh_i), \tag{A25}$$

in which $F_i^u = -\frac{1}{2}a_{ij}\partial u/\partial x_j$ and $F_i^v = -\frac{1}{2}a_{ij}\partial v/\partial x_j$ are the i components of the diffusive fluxes of u and v, respectively. Then it follows that under natural boundary conditions, the condition of adjoint operators

$$\int_D v\mathcal{L}u\mathrm{d}\mathbf{x} = \int_D u\mathcal{L}^*v\mathrm{d}\mathbf{x} \tag{A26}$$

is satisfied because either the functions u and v, or their crossing-boundary fluxes, vanish on the boundary ∂D. The latter is indeed the case in the radiation belt diffusion equation in adiabatic invariant coordinates.

Under the assumption that f satisfies the aforementioned conditions of u and v, applying (A26) to the right-hand side of (A22) immediately gives

$$\int_D f(\boldsymbol{\eta})\frac{\partial}{\partial t}p_t(\boldsymbol{\xi},\boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta} = \int_D f(\boldsymbol{\eta})\mathcal{L}_{\boldsymbol{\eta}}^*p_t(\boldsymbol{\xi},\boldsymbol{\eta})\mathrm{d}\boldsymbol{\eta}. \tag{A27}$$

Since f is arbitrary, we conclude that the transition probability density $p_t(\boldsymbol{\xi},\boldsymbol{\eta})$, as a function of t and $\boldsymbol{\eta}$, is a solution of the PDE

$$\frac{\partial}{\partial t}p_t(\boldsymbol{\xi},\boldsymbol{\eta}) = \mathcal{L}_{\boldsymbol{\eta}}^*p_t(\boldsymbol{\xi},\boldsymbol{\eta}). \tag{A28}$$

This equation is called the Kolmogorov forward equation, or Fokker-Planck equation. Given knowledge of the initial distribution function $\phi(0,\boldsymbol{\xi})$ of some diffusion process, the distribution function at a later time t is calculated from

$$\phi(t,\boldsymbol{\eta}) = \int_D \phi(0,\boldsymbol{\xi})p_t(\boldsymbol{\xi},\boldsymbol{\eta})\mathrm{d}\boldsymbol{\xi}. \tag{A29}$$

Applying this integration to both sides of (A28) then yields the Fokker-Planck equation in a more familiar form

$$\frac{\partial}{\partial t}\phi(t,\boldsymbol{\eta}) = \mathcal{L}_{\boldsymbol{\eta}}^*\phi(t,\boldsymbol{\eta}). \tag{A30}$$

Next, we show that $\phi(t,\boldsymbol{\eta})$ also satisfies the Kolmogorov backward equation. Looking again at the integral equation (A27), taking a particular $f(\boldsymbol{\eta}) = \phi(0,\boldsymbol{\eta})$, and reversing the derivations from (A27) back to (A17) gives

$$\frac{\partial}{\partial t}\mathbb{E}^{\boldsymbol{\xi}}[\phi(0,\mathbf{X}_t)] = \mathcal{L}_{\boldsymbol{\xi}}\mathbb{E}^{\boldsymbol{\xi}}[\phi(0,\mathbf{X}_t)]. \tag{A31}$$

To see the meaning of $\mathbb{E}^{\boldsymbol{\xi}}[\phi(0,\mathbf{X}_t)]$, comparison between equations (A20) and (A29) suggests that we construct a time backward stochastic process $\mathbf{Z}_r$ (r decreasing from t to 0) with transition probability density $\pi_t(\boldsymbol{\eta},\boldsymbol{\xi}) = p_t(\boldsymbol{\xi},\boldsymbol{\eta})$. Then, following (A29), $\phi(t,\boldsymbol{\eta})$ is represented by

$$\phi(t,\boldsymbol{\eta}) = \int_D \phi(0,\boldsymbol{\xi})\pi_t(\boldsymbol{\eta},\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi} = \mathbb{E}^{t,\boldsymbol{\eta}}[\phi(0,\mathbf{Z}_0)], \tag{A32}$$

in which the time interval t is indicated in the superscript of $\mathbb{E}$. However, since t only appears as a time interval, it actually does not matter whether time is increasing or decreasing in the stochastic process. Therefore, going from $\boldsymbol{\eta}$ to $\boldsymbol{\xi}$ while time is decreasing is identical to going from $\boldsymbol{\eta}$ to $\boldsymbol{\xi}$ while time is increasing. This amounts to saying that $\pi_t(\boldsymbol{\eta},\boldsymbol{\xi}) = p_t(\boldsymbol{\eta},\boldsymbol{\xi})$, and $\mathbf{Z}_r = \mathbf{X}_s$. As a result,

$$\mathbb{E}^{\boldsymbol{\xi}}[\phi(0,\mathbf{X}_t)] = \mathbb{E}^{t,\boldsymbol{\xi}}[\phi(0,\mathbf{Z}_0)] = \phi(t,\boldsymbol{\xi}). \tag{A33}$$

Therefore, the distribution function $\phi(t,\boldsymbol{\eta})$ is also a solution of the Kolmogorov backward equation

$$\frac{\partial}{\partial t}\phi(t,\boldsymbol{\eta}) = \mathcal{L}_{\boldsymbol{\eta}}\phi(t,\boldsymbol{\eta}), \tag{A34}$$

as claimed. This ensures that given a Fokker-Planck equation of form (A30), we can always write it in the Kolmogorov backward equation form (A34), and from the latter, we can determine the corresponding SDE of the time backward stochastic process of form (A8) and solve for $\phi(t, \boldsymbol{\eta})$ with the expectation in (A33).

We now outline the results (without derivations) that allow generalization of this conclusion to curvilinear coordinates. Transforming from Cartesian coordinates $\boldsymbol{\eta}$ to curvilinear coordinates $\boldsymbol{\zeta}$, the distribution function transforms to $\psi(t, \boldsymbol{\zeta}) = \phi(t, \boldsymbol{\eta}(\boldsymbol{\zeta}))G(\boldsymbol{\eta}; \boldsymbol{\zeta})$, where G is the transformation Jacobian. Equations (A30) and (A34) transform to

$$\frac{\partial}{\partial t}\psi(t, \boldsymbol{\zeta}) = \mathcal{L}^*_{\boldsymbol{\zeta}}\psi(t, \boldsymbol{\zeta}), \tag{A30′}$$

$$\frac{\partial}{\partial t}\phi(t, \boldsymbol{\zeta}) = \mathcal{L}_{\boldsymbol{\zeta}}\phi(t, \boldsymbol{\zeta}), \tag{A34′}$$

with the coefficients in $\mathcal{L}^*_{\boldsymbol{\zeta}}$ and $\mathcal{L}_{\boldsymbol{\zeta}}$ changed accordingly, and the corresponding expectation representation is

$$\mathbb{E}^{\boldsymbol{\zeta}}[\phi(0, \widetilde{\mathbf{X}}_t)] = \phi(t, \boldsymbol{\zeta}), \tag{A33′}$$

where $\widetilde{\mathbf{X}}_t$ is the stochastic process in $\boldsymbol{\zeta}$ coordinates. Note that although the function in the Fokker-Planck equation (A30′) becomes the new distribution function ψ, the function in the Kolmogorov backward equation (A34′) and the expectation representation (A33′) remains ϕ, the distribution function with respect to Cartesian coordinates, except for a change of variables. In mechanics, the role of Cartesian coordinates in phase space is played by canonical variables, and transformations between canonical variables have unit Jacobian determinant [e.g., *Landau and Lifshitz*, 1976]. Hence, there is but one distribution function for all sets of canonical variables. Consequently, regardless of what distribution function we start with in (A30′), we always end up with the distribution function with respect to canonical variables when writing (A34′) and (A33′). Given that (A34) and (A33) have the same form as (A34′) and (A33′), we thus no longer distinguish between Cartesian and curvilinear coordinates in applying the SDE method.

A3. Solution of General Diffusion Equations: The Feynman-Kac Formula

In this subsection we consider two generalizations of the conclusions in the previous subsection. First, in time-inhomogeneous diffusion processes, the coefficients $a_{ij}(t, \mathbf{x})$, and $b_i(t, \mathbf{x})$ depend on both spatial coordinates and time t. To construct the operator $\mathcal{L}$ with explicit time dependence from the Itô formula, the strategy is to treat time as one extra "stochastic" variable and construct a new Itô process in the extended space $t \otimes \mathbb{R}^n$ as [*Freidlin*, 1985]

$$\mathbf{Y}_s \equiv \begin{pmatrix} t - s \\ \mathbf{X}_s \end{pmatrix}, \quad 0 \le s \le t, \tag{A35}$$

and the corresponding SDE is

$$\begin{aligned} \mathrm{d}\mathbf{Y}_s &= \begin{pmatrix} -1 \\ \mathbf{b}(t - s, \mathbf{X}_s) \end{pmatrix} \mathrm{d}s + \begin{pmatrix} 0 \cdots\cdots 0 \\ \boldsymbol{\sigma}\left(t - s, \mathbf{X}_s\right) \end{pmatrix} \mathrm{d}\mathbf{W}_s \\ &\equiv \tilde{\mathbf{b}}(\mathbf{Y}_s)\mathrm{d}s + \tilde{\boldsymbol{\sigma}}(\mathbf{Y}_s)\mathrm{d}\mathbf{W}_s. \end{aligned} \tag{A36}$$

Equation (A36) shows that $\mathbf{Y}_s$ is a time-homogeneous Itô process, and thus, all our previous conclusions apply. Repeating those procedures, it is straightforward to verify that the operators $\mathcal{L}$ and $\mathcal{L}^*$ associated with $\mathbf{Y}_s$ are indeed the time-dependent operators we wanted to construct. On the other hand, the projection of $\mathbf{Y}_s$ onto the $\mathbb{R}^n$ space is a time-inhomogeneous Itô process

$$\mathrm{d}\mathbf{X}^t_s = \mathbf{b}\left(t - s, \mathbf{X}^t_s\right)\mathrm{d}s + \boldsymbol{\sigma}\left(t - s, \mathbf{X}^t_s\right)\mathrm{d}\mathbf{W}_s, \quad 0 \le s \le t. \tag{A37}$$

Consequently, we can say that the time-dependent operators $\mathcal{L}$ and $\mathcal{L}^*$ are associated with this time-inhomogeneous Itô process, and the expectation representation from the previous subsection should have the SDE (A8) replaced by (A37).

Second, a general diffusion equation (in Kolmogorov backward equation form) might contain a source (loss) term characterized by $c(t, \boldsymbol{\xi})$, such as

$$\frac{\partial}{\partial t}u(t, \boldsymbol{\xi}) = \mathcal{L}u(t, \boldsymbol{\xi}) + c(t, \boldsymbol{\xi})u(t, \boldsymbol{\xi}). \tag{A38}$$

Defining the operator

$$\widetilde{\mathcal{L}} \equiv \mathcal{L} + c(t, \mathbf{x}), \tag{A39}$$

it can be verified that, to generate $\widetilde{\mathcal{L}}$ instead of $\mathcal{L}$ in the Itô formula (A12), we only need to replace the function f therein with a functional $F[c]$, regarded as a function of t and $\mathbf{x}$ for given c, such that [*Freidlin*, 1985; *Øksendal*, 1998]

$$F(t, \mathbf{x}) \equiv f(\mathbf{x}) \exp\left(\int_0^t c(t-s, \mathbf{x})\mathrm{d}s\right). \tag{A40}$$

Then, by the same derivations as for the Kolmogorov backward equation (A18), the solution of the general diffusion equation (A38) with initial condition (A19) is

$$u(t, \boldsymbol{\xi}) = \mathbb{E}^{\boldsymbol{\xi}}\left[f(\mathbf{X}_t^t) \exp\left(\int_0^t c(t-s, \mathbf{X}_s^t)\mathrm{d}s\right)\right], \tag{A41}$$

in which the symbol $\mathbf{X}_t^t$ is a shorthand of $\mathbf{X}_{s=t}^t$. This formula, known as the Feynman-Kac formula, forms the starting point for our SDE formulation of radiation belt diffusion in section 2.

Appendix B: Nondegeneracy Condition for Neumann Boundary Condition (31)

For convenience of discussion, we will use numeric indices $\{1, 2, 3\}$, instead of $\{u, K, L\}$, as the subscripts of components in this appendix. The unnormalized reflection vector

$$\widetilde{\boldsymbol{\gamma}} \equiv (D_{12}, D_{22}, D_{32})^{\mathsf{T}} \tag{B1}$$

coincides with the second column of the diffusion tensor $\mathbf{D}$. Because $\mathbf{D}$ is symmetric, the quadratic form of $\mathbf{D}$ with its own column vector is equal to a diagonal component of the matrix product $\mathbf{D}^3$. In this case, it is

$$\widetilde{\boldsymbol{\gamma}}^{\mathsf{T}}\mathbf{D}\widetilde{\boldsymbol{\gamma}} = (\mathbf{D}^3)_{22}. \tag{B2}$$

A real symmetric matrix can be diagonalized via orthogonal transformations, i.e., $\mathbf{D} = \mathbf{P}^{\mathsf{T}}\boldsymbol{\Lambda}\mathbf{P}$, where $\boldsymbol{\Lambda}$ is a real diagonal matrix of $\mathbf{D}$'s eigenvalues, and $\mathbf{P}$ is an orthogonal matrix composed of $\mathbf{D}$'s orthogonalized eigenvectors. Consequently,

$$\mathbf{D}^3 = \mathbf{P}^{\mathsf{T}}\boldsymbol{\Lambda}^3\mathbf{P}. \tag{B3}$$

Therefore, the diagonal components of $\mathbf{D}$ and $\mathbf{D}^3$ can be expressed respectively by

$$D_{ii} = \Lambda_k P_{ki}^2, \tag{B4}$$

and

$$(\mathbf{D}^3)_{ii} = \Lambda_k^3 P_{ki}^2, \tag{B5}$$

where Λ_k is the kth diagonal component of $\boldsymbol{\Lambda}$, i.e., the kth eigenvalue of $\mathbf{D}$. Diffusion processes require that the diffusion tensor must be at least semipositive definite (this can be seen from equation (A11) and the discussion thereafter), which implies Λ_k must be nonnegative. Hence, from equations (B4) and (B5), it follows that there is always $(\mathbf{D}^3)_{ii} \geq 0$; and the only situation for $(\mathbf{D}^3)_{ii} = 0$ is when at least either P_{ki}^2 or Λ_k is zero for each k, which is also the situation for $D_{ii} = 0$. In other words, $D_{ii} > 0$ is a sufficient and necessary condition for $(\mathbf{D}^3)_{ii} > 0$. Relating this conclusion to equation (B2), equation (21), and the nondegeneracy requirement of Neumann boundary conditions (32), we see that the sufficient and necessary condition for the boundary condition (31) to be nondegenerate is D_{22}, or D_{KK}, being greater than zero on the boundary.

Acknowledgments
This material is based upon work supported by the National Aeronautics and Space Administration under grants NNX11AJ38G and NNX10AL02G issued through the Science Mission Directorate. We gratefully acknowledge Los Alamos National Lab for providing phase space density data and Air Force Research Lab and British Antarctic Survey for supplying chorus wave diffusion coefficients. Liheng Zheng wishes to thank Xin Tao for valuable discussions about the SDE method.

Michael Liemohn thanks Xin Tao and another reviewer for their assistance in evaluating the paper.

References

Albert, J. M. (2009), The coupling of quasi-linear pitch angle and energy diffusion, *J. Atmos. Sol. Terr. Phys.*, *71*(16), 1664–1668, doi:10.1016/j.jastp.2008.11.014.

Albert, J. M. (2013), Comment on "On the numerical simulation of particle dynamics in the radiation belt. Part I: Implicit and semi-implicit schemes" and "On the numerical simulation of particle dynamics in the radiation belt. Part II: Procedure based on the diagonalization of the diffusion tensor" by E. Camporeale et al., *J. Geophys. Res. Space Physics*, *118*, 7762–7764, doi:10.1002/2013JA019126.

Albert, J. M., and S. L. Young (2005), Multidimensional quasi-linear diffusion of radiation belt electrons, *Geophys. Res. Lett.*, *32*, L14110, doi:10.1029/2005GL023191.

Albert, J. M., N. P. Meredith, and R. B. Horne (2009), Three-dimensional diffusion simulation of outer radiation belt electrons during the 9 October 1990 magnetic storm, *J. Geophys. Res.*, *114*, A09214, doi:10.1029/2009JA014336.

Arfken, G. B., and H. J. Weber (1995), *Mathematical Methods for Physicists,* 4th ed., Academic Press, San Diego, Calif.

Baker, D., and S. Kanekal (2008), Solar cycle changes, geomagnetic variations, and energetic particle properties in the inner magnetosphere, *J. Atmos. Sol. Terr. Phys.*, *70*(2–4), 195–206, doi:10.1016/j.jastp.2007.08.031.

Baker, D. N., et al. (1997), Recurrent geomagnetic storms and relativistic electron enhancements in the outer magnetosphere: ISTP coordinated measurements, *J. Geophys. Res.*, *102*, 14,141–14,148.

Beutier, T., and D. Boscher (1995), A three-dimensional analysis of the electron radiation belt by the Salammbô code, *J. Geophys. Res.*, *100*(A8), 14,853–14,861.

Bossy, M., E. Gobet, and D. Talay (2004), A symmetrized Euler scheme for an efficient approximation of reflected diffusions, *J. Appl. Prob.*, *41*, 877–889, doi:10.1239/jap/1091543431.

Brautigam, D. H., and J. M. Albert (2000), Radial diffusion analysis of outer radiation belt electrons during the October 9, 1990, magnetic storm, *J. Geophys. Res.*, *105*(A1), 291–309.

Camporeale, E., G. L. Delzanno, S. Zaharia, and J. Koller (2013a), On the numerical simulation of particle dynamics in the radiation belt: 1. Implicit and semi-implicit schemes, *J. Geophys. Res. Space Physics*, *118*, 3463–3475, doi:10.1002/jgra.50293.

Camporeale, E., G. L. Delzanno, S. Zaharia, and J. Koller (2013b), On the numerical simulation of particle dynamics in the radiation belt: 2. Procedure based on the diagonalization of the diffusion tensor, *J. Geophys. Res. Space Physics*, *118*, 3476–3484, doi:10.1002/jgra.50278.

Camporeale, E., G. L. Delzanno, S. Zaharia, and J. Koller (2013c), Reply to comment by J. M. Albert on "On the numerical simulation of particle dynamics in the radiation belt. Part I: Implicit and semi-implicit schemes" and "On the numerical simulation of particle dynamics in the radiation belt. Part II: Procedure based on the diagonalization of the diffusion tensor", *J. Geophys. Res. Space Physics*, *118*, 7765–7767, doi:10.1002/2013JA019389.

Chen, Y., G. D. Reeves, and R. H. W. Friedel (2007), The energization of relativistic electrons in the outer Van Allen radiation belt, *Nat. Phys.*, *3*, 614–617, doi:10.1038/nphys655.

Costantini, C., B. Pacchiarotti, and F. Sartoretto (1998), Numerical approximation for functionals of reflecting diffusion processes, *SIAM J. Appl. Math.*, *58*(1), 73–102.

Elkington, S. R., M. K. Hudson, and A. A. Chan (1999), Acceleration of relativistic electrons via drift-resonant interaction with toroidal-mode Pc-5 ULF oscillations, *Geophys. Res. Lett.*, *26*, 3273–3276.

Elkington, S. R., M. K. Hudson, M. J. Wiltberger, and J. G. Lyon (2002), MHD/particle simulations of radiation belt dynamics, *J. Atmos. Sol. Terr. Phys.*, *64*, 607–615.

Elkington, S. R., M. Wiltberger, A. A. Chan, and D. N. Baker (2004), Physical models of the geospace radiation environment, *J. Atmos. Sol. Terr. Phys.*, *66*(15–16), 1371–1387.

Fei, Y., A. A. Chan, S. R. Elkington, and M. J. Wiltberger (2006), Radial diffusion and MHD particle simulations of relativistic electron transport by ULF waves in the September 1998 storm, *J. Geophys. Res.*, *111*, A12209, doi:10.1029/2005JA011211.

Fok, M.-C., R. B. Horne, N. P. Meredith, and S. A. Glauert (2008), Radiation belt environment model: Application to space weather nowcasting, *J. Geophys. Res.*, *113*, A03S08, doi:10.1029/2007JA012558.

Freidlin, M. (1985), *Functional Integration and Partial Differential Equations, Annals of Mathematics Studies,* vol. 109, Princeton Univ. Press, Princeton, N. J.

Gardiner, C. W. (2004), *Handbook of Stochastic Methods,* 3rd ed., Springer, Berlin, Germany.

Glauert, S. A., and R. B. Horne (2005), Calculation of pitch angle and energy diffusion coefficients with the PADIE code, *J. Geophys. Res.*, *110*, A04206, doi:10.1029/2004JA010851.

Glauert, S. A., R. B. Horne, and N. P. Meredith (2014), Three-dimensional electron radiation belt simulations using the BAS radiation belt model with new diffusion models for chorus, plasmaspheric hiss, and lightning-generated whistlers, *J. Geophys. Res. Space Physics*, *119*, 268–289, doi:10.1002/2013JA019281.

Gobet, E. (2001), Euler schemes and half-space approximation for the simulation of diffusion in a domain, *ESAIM Probab. Stat.*, *5*, 261–297, doi:10.1051/ps:2001112.

Haerendel, G. (1968), Diffusion theory of trapped particles and the observed proton distribution, in *Earth's Particles and Fields; Proceedings of the NATO Advanced Study Institute Held at Freising, Germany, July 31–August 11, 1967*, edited by B. M. McCormac, p. 171, Reinhold, New York.

Horne, R. B., R. M. Thorne, S. A. Glauert, J. M. Albert, N. P. Meredith, and R. R. Anderson (2005), Timescale for radiation belt electron acceleration by whistler mode chorus waves, *J. Geophys. Res.*, *110*, A03225, doi:10.1029/2004JA010811.

Kim, H.-J., and A. A. Chan (1997), Fully-adiabatic changes in storm-time relativistic electron fluxes, *J. Geophys. Res.*, *102*, 22,107–22,116.

Kloeden, P. E., and E. Platen (1992), *Numerical Solution of Stochastic Differential Equations, Applications of Mathematics*, vol. 23, Springer, Berlin, Germany.

Koller, J., Y. Chen, G. D. Reeves, R. H. W. Friedel, T. E. Cayton, and J. A. Vrugt (2007), Identifying the radiation belt source region by data assimilation, *J. Geophys. Res.*, *112*, A06244, doi:10.1029/2006JA012196.

Landau, L. D., and E. M. Lifshitz (1976), *Mechanics,* 3rd ed., Pergamon Press, Oxford, U. K.

Mathews, J., and R. L. Walker (1970), *Mathematical Methods of Physics,* 2nd ed., Addison-Wesley, Boston, Mass.

Matsumoto, M., and T. Nishimura (1998), Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator, *ACM Trans. Model. Comput. Simul.*, *8*(1), 3–30, doi:10.1145/272991.272995.

Matsumoto, M., and T. Nishimura (2000), Dynamic creation of pseudorandom number generators, in *Monte Carlo and Quasi-Monte Carlo Methods 1998*, edited by H. Niederreite and J. Spanier, pp. 56–69, Springer, Berlin, Germany.

Meredith, N. P., R. B. Horne, and R. R. Anderson (2001), Substorm dependence of chorus amplitudes: Implications for the acceleration of electrons to relativistic energies, *J. Geophys. Res.*, *106*(A7), 13,165–13,178.

Meredith, N. P., R. B. Horne, R. M. Thorne, and R. R. Anderson (2003), Favored regions for chorus-driven electron acceleration to relativistic energies in the Earth's outer radiation belt, *Geophy. Res. Lett.*, *30*(16), 1871, doi:10.1029/2003GL017698.

Meredith, N. P., R. B. Horne, A. Sicard-Piet, D. Boscher, K. H. Yearby, W. Li, and R. M. Thorne (2012), Global model of lower band and upper band chorus from multiple satellite observations, *J. Geophys. Res.*, *117*, A10225, doi:10.1029/2012JA017978.

O'Brien, T. P. (2014), Breaking all the invariants: Anomalous electron radiation belt diffusion by pitch angle scattering in the presence of split magnetic drift shells, *Geophy. Res. Lett.*, *41*, 216–222, doi:10.1002/2013GL058712.

Øksendal, B. (1998), *Stochastic Differential Equations,* 5th ed., Springer, Berlin, Germany.

Ozeke, L. G., I. R. Mann, K. R. Murphy, I. J. Rae, D. K. Milling, S. R. Elkington, A. A. Chan, and H. J. Singer (2012), ULF wave derived radiation belt radial diffusion coefficients, *J. Geophys. Res.*, *117*, A04222, doi:10.1029/2011JA017463.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992), *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Cambridge Univ. Press, New York.
Reeves, G. D., K. L. McAdams, R. H. W. Friedel, and T. P. O'Brien (2003), Acceleration and loss of relativistic electrons during geomagnetic storms, *Geophys. Res. Lett.*, *30*(10), 1529, doi:10.1029/2002GL016513.
Roberts, C. S. (1969), Pitch-angle diffusion of electrons in the magnetosphere, *Rev. Geophys.*, *7*(1–2), 305–337, doi:10.1029/RG007i001p00305.
Roederer, J. G. (1970), *Dynamics of Geomagnetically Trapped Radiation, Physics and Chemistry in Space*, vol. 2, Springer, Berlin, Heidelberg, Germany.
Schulz, M. (1991), The magnetosphere, in *Geomagnetism*, vol. 4, edited by J. A. Jacobs, pp. 87–293, Academic Press, Waltham, Mass.
Schulz, M., and L. J. Lanzerotti (1974), *Particle Diffusion in the Radiation Belts*, Springer, New York.
Selesnick, R. S. (2012), Atmospheric scattering and decay of inner radiation belt electrons, *J. Geophys. Res.*, *117*, A08218, doi:10.1029/2012JA017793.
Shabansky, V. P. (1971), Some processes in the magnetosphere, *Space Sci. Rev.*, *12*, 299–418.
Shprits, Y. Y. (2009), Potential waves for pitch-angle scattering of near-equatorially mirroring energetic electrons due to the violation of the second adiabatic invariant, *Geophy. Res. Lett.*, *36*, L12106, doi:10.1029/2009GL038322.
Su, Z., F. Xiao, H. Zheng, and S. Wang (2011), Radiation belt electron dynamics driven by adiabatic transport, radial diffusion, and wave-particle interactions, *J. Geophys. Res.*, *116*, A04205, doi:10.1029/2010JA016228.
Subbotin, D., Y. Shprits, and B. Ni (2010), Three-dimensional VERB radiation belt simulations including mixed diffusion, *J. Geophys. Res.*, *115*, A03205, doi:10.1029/2009JA015070.
Subbotin, D. A., and Y. Y. Shprits (2012), Three-dimensional radiation belt simulations in terms of adiabatic invariants using a single numerical grid, *J. Geophys. Res.*, *117*, A05205, doi:10.1029/2011JA017467.
Tao, X., A. A. Chan, J. M. Albert, and J. A. Miller (2008), Stochastic modeling of multidimensional diffusion in the radiation belts, *J. Geophys. Res.*, *113*, A07212, doi:10.1029/2007JA012985.
Tao, X., J. M. Albert, and A. A. Chan (2009), Numerical modeling of multidimensional diffusion in the radiation belts using layer methods, *J. Geophys. Res.*, *114*, A02215, doi:10.1029/2008JA013826.
Thorne, R. M. (2010), Radiation belt dynamics: The importance of wave-particle interactions, *Geophy. Res. Lett.*, *37*, L22107, doi:10.1029/2010GL044990.
Tsurutani, B. T., and E. J. Smith (1977), Two types of magnetospheric ELF chorus and their substorm dependences, *J. Geophys. Res.*, *82*(32), 5112–5128, doi:10.1029/JA082i032p05112.
Tu, W., and X. Li (2011), Adiabatic effects on radiation belt electrons at low altitude, *J. Geophys. Res.*, *116*, A09201, doi:10.1029/2011JA016468.
Tu, W., X. Li, Y. Chen, G. D. Reeves, and M. Temerin (2009), Storm-dependent radiation belt electron dynamics, *J. Geophys. Res.*, *114*, A02217, doi:10.1029/2008JA013480.
Tu, W., S. R. Elkington, X. Li, W. Liu, and J. Bonnell (2012), Quantifying radial diffusion coefficients of radiation belt electrons based on global MHD simulation and spacecraft measurements, *J. Geophys. Res.*, *117*, A10210, doi:10.1029/2012JA017901.
Tu, W., G. S. Cunningham, Y. Chen, M. G. Henderson, E. Camporeale, and G. D. Reeves (2013), Modeling radiation belt electron dynamics during GEM challenge intervals with the DREAM3D diffusion model, *J. Geophys. Res. Space Physics*, *118*, 6197–6211, doi:10.1002/jgra.50560.
Ukhorskiy, A. Y., M. I. Sitnov, R. M. Millan, and B. T. Kress (2011), The role of drift orbit bifurcations in energization and loss of electrons in the outer radiation belt, *J. Geophys. Res.*, *116*, A09208, doi:10.1029/2011JA016623.
Varotsou, A., D. Boscher, S. Bourdarie, R. B. Horne, S. A. Glauert, and N. P. Meredith (2005), Simulation of the outer radiation belt electrons near geosynchronous orbit including both radial diffusion and resonant interaction with Whistler-mode chorus waves, *Geophys. Res. Lett.*, *32*, L19106, doi:10.1029/2005GL023282.
Varotsou, A., D. Boscher, S. Bourdarie, R. B. Horne, N. P. Meredith, S. A. Glauert, and R. H. Friedel (2008), Three-dimensional test simulations of the outer radiation belt electron dynamics including electron-chorus resonant interactions, *J. Geophys. Res.*, *113*, A12212, doi:10.1029/2007JA012862.
Zhang, M. (1999), A markov stochastic process theory of cosmic-ray modulation, *Astrophys. J.*, *513*(1), 409.