Inference solves a boundary-value collision problem with relevance to neutrino flavor transformation

Eve Armstrong[*]

Department of Physics, New York Institute of Technology, New York, New York 10023, USA
and Department of Astrophysics, American Museum of Natural History, New York, New York 10024, USA

Understanding neutrino flavor transformation in dense environments such as core-collapse supernovae (CCSN) is critical for inferring the physics of these events and interpreting a detected neutrino signal. The role of direction-changing collisions in shaping the neutrino flavor field in these environments is important and poorly understood; it has not been treated self-consistently. There has been progress, via numerical integration, to include the effects of collisions in the dynamics of the neutrino flavor field. While this has led to important insights, integration is limited by its requirement that full initial conditions must be assumed known. On the contrary, it has been shown in recent years that feedback from collisions to the field is a boundary value problem. Numerical integration techniques are poorly equipped to handle that formulation. This paper demonstrates that an inference formulation of the problem can solve a simple collisions-only model representing a CCSN core–without full knowledge of initial conditions. Specifically, the procedure solves a two-point boundary value problem with partial information at the bounds. The model is sufficiently simple that physical reasoning may be used as a confidence check on the inference-based solution, and the procedure recovers the expected model dynamics. This result demonstrates that inference can solve a problem that is artificially hidden from integration techniques—-a problem that is an important feature of flavor evolution in dense environments. Thus, it is worthwhile to explore means of augmenting the existing powerful integration tools with inference-based approaches.

I. INTRODUCTION

The physics of flavor evolution in the neutrino field in high-density environments such as core collapse supernovae (CCSN) can significantly affect the transport of energy, entropy, and lepton number, with implications regarding nucleosynthesis, the mechanism of explosion, and the mass assembly histories of galaxies [1–3]. Due to the fierce nonlinearity of various features of this problem, we are far from a complete understanding. One important and poorly studied problem is the role of direction-changing collisions in shaping the neutrino density profile in these environments.

It is well-established that coherent forward scattering of neutrinos with matter can lead to resonant effects [4–6] and that forward scattering of neutrinos with each other also affects flavor evolution [7–10]. These potentials, which govern neutrino flavor transformation, themselves depend on the flavor states of the neutrinos. This nonlinearity has led to the development of highly sophisticated numerical approaches, including the multiangle Neutrino BULB code [11,12] and the IsotropicSQA code [13].

The ability of direction-changing scattering, in addition, to significantly impact flavor evolution in the SN envelope was first shown by Ref. [14]. There have since been efforts to understand this "halo effect" on flavor transformation [15–18], although a self-consistent solution has not been obtained. More recent studies have shown that fast pairwise conversions can happen due to a crossing between the angular distributions of neutrinos and antineutrinos (See Ref. [19] and citations therein)—a phenomenon that is sensitive to the exact shapes of the angular distributions.

Most recently, direction-changing collisions (hereafter "collisions") have been shown to enhance fast flavor conversion in high-density environments [20–22]. This result is counter to the expectation that collisions should dampen flavor evolution via dynamical decoherence [23]. Moreover, even in small-scale CCSN models, accounting for collisions has yielded significant and counterintuitive changes in the angular distributions of the neutrino density profiles. For a more comprehensive summary of the role of collisions in flavor evolution, see Ref. [20] and citations therein.

To date, efforts to include collisions in the neutrino quantum kinetic equations have used numerical integration to solve an initial value problem (e.g., Refs. [16,20,24]). It is not necessarily desirable to place assumptions on initial

[*]evearmstrong.physics@gmail.com

conditions. On the contrary, when feedback between collisions and flavor evolution is included in the calculation, the problem becomes a two-point boundary-value problem with no guarantee of a unique solution [20].

Now, it is important to distinguish here between a particle- versus field-based framework for tackling the flavor evolution problem. In the particle framework, the evolution of each neutrino is an initial-value problem: The initial conditions at the source are known, and any collision simply follows deterministically from the evolution equations. This framework works well if all background targets with which a neutrino interacts are static in flavor space—that is, they do not themselves undergo flavor evolution.

To understand neutrino-neutrino interactions, however, the particle framework is intractable because the neutrino background is evolving in flavor as well.[1] To study flavor evolution then, we adopt the "field" approach (for examples, see Refs. [25–31]). In this framework, it is assumed that there exist continuous, steady-state flavor fields of back-propagating neutrinos and forward-propagating neutrinos that interact, where the instantaneous direction of a particular neutrino is not known.[2] Reference [15] offers perhaps the best explanation for why the initial-value formulation is insufficient for the backscattering problem within the context of flavor evolution. Instead, one tractable formulation is to use a boundary condition at the scattering surface, which must be determined self-consistently. One effort has been made to do this through iterative forward-backward integration [16]. Such a method is numerically unstable and time-consuming and hence, unfit to add to a large-scale simulation.

Within that context, in this paper, we present an inference-based method for solving the collision problem. We formulate it not as an initial-value problem, but rather as a two-point boundary value problem with partial information at the boundaries. We do this via an inference—or "inverse"—formulation [32]. Inference is a means to optimize a model with data and to predict model evolution at locations outside of those where the data have been provided. The inference formulation does not require that initial conditions be known. Rather, it requires that *some* condition(s) [or constraint(s)] be known *at one or more* locations that parametrize the model dynamics. These conditions may be placed at the end points of model evolution, or more generally at any point(s) along the model trajectory. Further, the formulation of inference employed in this paper offers a mean to determine whether a unique solution exists and—if it does not—the degree of degeneracy present.

In this paper, we use a specific formulation of inference that was built to handle the case of extremely sparse data. Within the geosciences, this technique is known as statistical data assimilation (SDA). SDA was invented for numerical weather prediction [33–38] and has gained considerable traction in neurobiology [39–45]. Within astrophysics, inference has been used mainly for pattern recognition (e.g., [46]), while its utility for model completion is gaining notice in the exoplanet community [47] and solar physics [48]. Recently, SDA has been applied to small-scale models of neutrino flavor evolution [49–51]. In those papers, the model contained coherent forward scattering only, and thus, numerical integration could be taken as a confidence check on SDA solutions.

By contrast to Refs. [49–51], this work tackles a problem formulation that numerical integration cannot access. For this reason, we shall examine a simple model whose solution can be intuited via physical reasoning. The model represents the high-density region of a supernova core, wherein neutrinos are effectively trapped and distributed isotropically. In this context, we consider *collisions alone*: absent neutrino-matter and neutrino-neutrino coupling potentials as a vital test that the procedure can reliably yield a physically intuitable solution. We must obtain that confidence prior to trusting the procedure to yield correct solutions once the interaction terms are added—as the solutions in that case will be too complicated to physically intuit. The model consists of four neutrino beams: two outgoing and two incoming and four angular bins. Absent flavor evolution, we will seek to examine the angular distributions of the neutrino number densities as functions of radius due to collisions alone.

Given partial information at the boundaries, the SDA procedure is able to recover the density profiles that follow from a simple physical argument, thereby illuminating physics that is artificially hidden from numerical integration techniques. Further, multiple independent trials converge to one solution.

This paper proceeds as follows. Section II describes the model and imposed boundary conditions and presents a simple physical argument regarding the expected model evolution. Section III describes the inference methodology and the simulated experiments performed. Section IV shows how the SDA procedure predicts the expected result. Finally, Sec. V comments on the implications for tackling the full quantum kinetic equations that includes nonlinear feedback between collisions and flavor evolution and advocates for exploring means to augment existing integration methodologies with inference-based techniques. The long-term ambition is to incorporate an

---

[1]For example, to predict the evolution of a forward-propagating "test" neutrino, one would need to know the evolution history of the "background" neutrinos that were emitted at an earlier time in the explosion and interacted with this test neutrino after they were back-scattered. Then, the evolution of these background neutrinos themselves would depend on their interactions with neutrinos that were emitted even earlier, and so on.

[2]To be clear, in the field framework, we cannot identify backward-propagating neutrinos to be forward-propagating neutrinos from an "earlier time" and trace their flavor evolution back to the source.

inference-based solution for the flavor field into existing large-scale hydrodynamics simulations.

II. STATIONARY BOX MODEL OF A SUPERNOVA CORE

The goal in crafting the model is twofold: (1) to create an opportunity to observe the effect on the neutrino angular distributions due to collisions alone and (2) to retain a simplicity such that the solution can be intuited via physical reasoning, so as to provide a confidence check on the SDA solution. To this end, we employ a two-dimensional flat "box" model and four neutrino beams: two incoming and two outgoing.

The model schematic of the supernova core is shown in Fig. 1. Neutrinos radiate from the center (the x-axis at $y = 0$) out to some final radius R (at $y = y_{\mathrm{max}}$). To examine the effects of collisions alone, we omit flavor evolution and adopt a single-flavor model with no antineutrinos. There is a sink at $r = R$ where the outgoing beams may escape. There is also a source and sink permitted at $y = 0$, as particles are permitted to travel in the negative $\hat{y}$ direction.

Neutrino flavor evolution can be expressed in terms of a density matrix ρ (and $\bar{\rho}$ for antineutrinos). For each neutrino momentum mode $\vec{p}$, we can write the following quantum-kinetic equation [52,53]:

$$\left(\frac{\partial}{\partial t} + \vec{v} \cdot \nabla\right)\rho(\vec{x}, \vec{p}, t) = -i[H(\vec{x}, \vec{p}, t), \rho(\vec{x}, \vec{p}, t)]$$
$$+ C\{\rho(\vec{x}, \vec{p}, t), \bar{\rho}(\vec{x}, \vec{p}, t)\}, \quad (1)$$

where $\vec{x}$ and t are spatial and temporal coordinates, respectively. On the right side, the Hamiltonian term H describes neutrino mixing in vacuum, neutrino interactions with the matter background, and neutrino-neutrino self interactions; it contains coherent scattering. The second term C is the "collision term," which accounts for direction-changing scattering.[3]
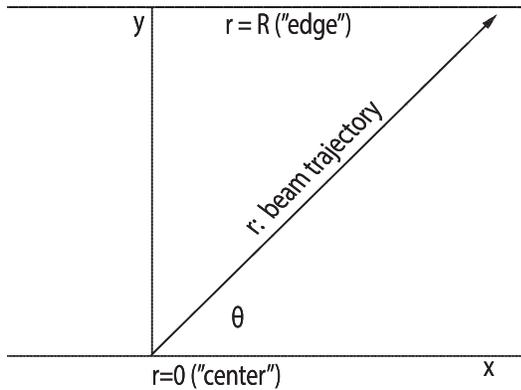
In this paper, we neglect the (coherent) Hamiltonian term and focus on the collision term C. We shall assume that a steady state can be defined on a timescale shorter than the leakage timescale, and so we eliminate time from Eq. (1). That is, we will solve Eq. (1) depending on scattering angle and not on momentum. In this context, the velocity in the second operator on the left side shall be considered a spatial—rather than a temporal—velocity. Then our equation of motion simplifies to

$$\cos\theta_i \frac{d\rho_i(r)}{dr} = \sum_j (-C^{\mathrm{loss}}\rho_i(r) + C^{\mathrm{gain}}\rho_j(r))\frac{\Delta\cos\theta_j}{2}. \quad (2)$$

Here, subscripts (i, j) denote indices of angular bins, $\Delta\cos\theta_j$ is the width of the jth angular bin, and the cosine term on the left side accounts for the dependence of path length on zenith angle. C^{loss} and C^{gain} are the coefficients for loss and gain for each bin; they represent the strength of interaction between neutrinos and nucleons.[4] Thus, for our box model, there are four Eqs. (2), each governing the evolution of the neutrino density in one of four angular bins.

To create a problem with a physically intuitable solution, we make further simplifications. First, the coefficients C^{loss} and C^{gain} are constant numbers throughout the medium. That is, the matter density is constant within the box and drops discontinuously to zero at $r = R$. Next, we set C^{loss} equal to C^{gain} to conserve particle number at each interaction.[5]

Finally, we choose a value of C to approximate nuclear density. Near nuclear density, neutrinos are effectively trapped, and so we expect an isotropic distribution of scatterers: At each collision, a neutrino has equal probability of scattering inward as outward. To choose C appropriately, we perform a dimensional analysis of Eq. (2). The left side is a derivative, with units of 1/length. The right side also has units of 1/length, as the only dimensionful quantities on the right side are the coefficients C, with units of inverse length. Physically, as C represents the strength of the interaction between neutrinos and matter, C can be considered the inverse of the neutrino mean free path, with units of m^{-1}. For

---

[3]Note that the separation of coherent from direction-changing scattering is an artificial one; see Sec. V.

[4]Physically, the loss and gain coefficients represent the product of the number of scatterers and the cross-section for each particular interaction channel averaged over some distribution of neutrino energy [24].

[5]Note that setting C^{loss} equal to C^{gain} renders Eq. (2) trivial, unless there are slight asymmetries in the values of the $\cos\theta$ terms on the left side of the equations. For this reason, we introduce slight asymmetries between the fluxes of outgoing versus incoming beams. The $\cos\theta$ values are 0.8 and 0.9 for the outgoing beams and -0.6 and -0.7 for those incoming.

FIG. 1. Two-dimensional "box" model of the supernova core. The core center is the x-axis, at $r = 0$. The positive $\hat{y}$ direction is outward toward the edge. Two beams are directed outward and two inward, each along some angle θ.

example, taking a mean free path on the order of meters to represent nuclear density, we choose C to be 0.2 m^{-1}.

A. The boundary conditions

We place on this model partially known boundary conditions. First, the values of the outgoing beams at the core center ($r = 0$) are known. At that location, we give them both a (dimensionless) number density of 0.5. Second, the values of incoming beams at the core edge ($r = R$) are zero. That is, by definition, at $r = R$, there occurs no further backscattering. This is the sole information furnished to the inference procedure. The procedure is tasked with taking this information together with the model equations of motion (Eq. (2)) to predict the complete density profile at all other spatial locations for all four angular bins—including the values at $r = 0$ of the incoming beams and the values at $r = R$ of the outgoing beams.

B. Expected density profile, to be recovered by the SDA procedure

Given a value of C within our box that corresponds to nuclear density, and the fact that a system represented as ordinary differential equations has no memory, we expect to observe isotropy for the vast majority of the trajectories of all four beams. That is, at each interaction, the particles possess no preferred direction. Thus, given the condition that the outgoing beams begin their trajectory at $r = 0$ with a value of 0.5, we expect the densities of all beams to remain approximately 0.5 throughout the majority of their their trajectories. Then, as we approach $r = R$, we should see a precipitous drop. Near R, the densities of the incoming beams should drop to zero, in accordance with the imposed boundary condition there, and the inference procedure should *predict* that the densities of the outgoing beams drop to some number below 0.5—that number to scale inversely with the value of C. Stated physically: As C weakens, more particles should escape.

Note again that the model possesses a leak at $r = R$ (and at $r = 0$), as scatterers may pass beyond zero—in the negative $\hat{y}$ of Fig. 1. We do not expect that a steady-state solution exists for longer than a duration corresponding to the leakage timescale. Rather, we expect that different choices for the coefficient C will yield different solutions, each representing one freeze frame of a steady state solution. Specifically, as C decreases, the mean free path increases—and equivalently, the total duration increases, as increasingly more neutrinos have had time to escape the core.

III. INFERENCE METHODOLOGY

SDA is an inference procedure in which a dynamical system is assumed to underlie any measured quantities. This model F can be written as a set of D ordinary differential equations that evolve in some parameterization r as

$$\frac{\mathrm{d}x_a(r)}{\mathrm{d}r} = F_a(\boldsymbol{x}(r), \boldsymbol{p}(r)); \quad a = 1, 2, ..., D,$$

where the components x_a of the vector $\mathbf{x}$ are the model state variables. Any unknown parameters to be estimated are contained in $\boldsymbol{p}$ and may themselves be variable. In this paper, model F is the set of four equations (Eq. (2) governing the neutrino density profile in each of four angular bins.

A subset L of the D state variables is associated with measured quantities. In this paper, the "measurements" are the four boundary conditions placed on the outgoing beams at $r = 0$ and incoming beams at $r = R$. One seeks to estimate the evolution of all state variables that is consistent with the measurements provided to predict model evolution at parametrized locations where measurements are not present.

We formulate the SDA procedure as an optimization wherein a cost function is extremized. An optimization formulation of inference does not require knowledge of initial conditions. Rather, it requires that *some* constraints be placed on the model at *some* location(s) on the coordinate axis that parametrizes the model equations of motion. These conditions may exist at the end points of the problem, as is the case in this paper. Importantly, however, the conditions need not be at the bounds. This flexibility is the key advantage of inference, compared to integration, that we aim to demonstrate.

The cost function is written in two terms.[6] One term represents measurement error: the difference between state prediction and any measurements made. The second term represents model error: the difference between state prediction and adherence to the model dynamics. It will be shown in Sec. IV that treating the model error as finite offers a means to determine whether a particular solution is consistent with both measurements and model dynamics, as well as a means to assess uniqueness.

We search the surface of the cost function via the variational method. The procedure in its entirety—that is, a variational approach to minimization coupled with an annealing method to identify a lowest minimum of the cost function (which will be described below in this section)—is referred to as variational annealing (VA). The procedure searches a $(D(N + 1) + p)$-dimensional state space, where N is the number of discretized steps. One seeks the path $X^0 = \boldsymbol{x}(0), ..., \boldsymbol{x}(N), \boldsymbol{p}(0), ...\boldsymbol{p}(N)$ in state space on which the cost function attains a minimum value. We refer to the cost function A_0 as the action because it can be derived from the concept of a classical action of a particle on a path in a state space [54]. Reference [50] demonstrated that the action formulation offers a simple litmus test for identifying correct solutions: Namely, they are solutions that correspond to the path of least action.

---

[6]Additional terms representing equality constraints may be added to the cost function, depending on the aim of a particular optimization procedure.

The cost function A_0 used in this paper is written as

$$A_0 = R_f A_{\text{model}} + R_m A_{\text{meas}}$$

$$A_{\text{model}} = \frac{1}{ND} \sum_{n \in \{\text{odd}\}}^{N-2} \sum_{a=1}^{D} \left[\left\{ x_a(n+2) - x_a(n) - \frac{\delta r}{6} [F_a(\boldsymbol{x}(n), \boldsymbol{p}) + 4F_a(\boldsymbol{x}(n+1), \boldsymbol{p}) + F_a(\boldsymbol{x}(n+2), \boldsymbol{p})] \right\}^2 \right.$$

$$\left. + \left\{ x_a(n+1) - \frac{1}{2}(x_a(n) + x_a(n+2)) - \frac{\delta r}{8} [F_a(\boldsymbol{x}(n), \boldsymbol{p}) - F_a(\boldsymbol{x}(n+2), \boldsymbol{p})] \right\}^2 \right]$$

$$A_{\text{meas}} = \frac{1}{N_{\text{meas}}} \sum_{j} \sum_{l=1}^{L} (y_l(j) - x_l(j))^2. \tag{3}$$

The second (simpler) squared term of Eq. (3) governs the transfer of information from measurements y_l to model states x_l.[7] Here, the summation on j runs over all discretized locations J at which measurements are made, which may be some subset of all discretized steps of the model. The summation on l is taken over all L measured quantities. In the simulations of this paper, the measured quantities are the number densities in the angular bins, the outgoing beams sampled at radius of $r = 0$ (at $j = 1$ only), and the incoming beams sampled only at $r = R$ (at the value of j that denotes the final radial location R).

The first squared term of Eq. (3) incorporates the model evolution of all D state variables x_a. Here, the outer sum on n is taken over all discretized time points of the model equations of motion. The sum on a is taken over all D state variables at all discretized locations.[8] In our model, these measured quantities are all four beams or $L = D = 4$. The first and second bracketed terms represent error in the first and second derivative of the model, respectively. For further details on the action formulation, see Appendix A of Ref. [49].

Importantly, note that the SDA procedure is tasked with inferring from the sparse boundary conditions the complete model evolution at *all* locations on r.

Finally, if the model is nonlinear, the cost function surface will be nonconvex: multiple minima may exist. For this reason, we employ an annealing procedure to identify the global minimum of the problem and to ascertain whether there exists a single global minimum—that is, a unique solution. Details of this annealing procedure will be presented in Sec. IV.

A. Details of the simulated experiments

We performed the simulated experiment for three distinct values of the coefficient $C^{\text{loss}} = C^{\text{gain}}$: 2.0, 0.2, and 0.02 m^{-1}, where the choice of 0.2 m^{-1} represents nuclear

density (a mean free path of 5 meters). At each discretized model location r, the procedure was permitted to search the full dynamical range of each state variable of [0:1]. For each of those three experiments, we took the size of the core to be $R = 50$ km, and examined the robustness of the solution across three distinct choices for step size dr: 1 m, 10 m, and 100 m.[9]

To examine the uniqueness of solutions, for each of the (nine) experiments described, 20 independent paths were searched, each beginning at a randomly generated set of initial guesses for the four state variables at each parametrized location r.

To perform the optimization, we used the open-source Interior Point Optimizer (Ipopt) [55]. Ipopt uses a Simpson's rule method of finite differences to discretize the state space, a Newton's method to search, and a barrier method to impose user-defined bounds that are placed upon the searches. The discretization of the state space, the calculations of the model Jacobean and Hessian matrices, and the annealing procedure were performed via an interface with Ipopt that was written in C and PYTHON [56]. All simulations were run on a 720-core, 1440-GB, 64-bit CPU cluster.

IV. RESULT

A. General findings

Results are fourfold.

(i) As expected for an isotropic distribution of scatterers, for a large value of the coefficient $C^{\text{loss}} = C^{\text{gain}}$, the number densities of the four angular bins remain constant at 0.5 for the majority of the trajectory, with a precipitous drop near the bound at $r = R$. As the value of C decreases, the rate of depletion of the number densities as a function of r slows, and the predicted values for the outgoing beams at $r = R$

---

[7]This term derives from the mutual information of probability theory [54].

[8]This term can be derived via consideration of Markov-chain transition probabilities [54].

[9]Within the dimensionless mathematical framework of the SDA procedure, those choices of 1, 10, and 100 m^{-1} corresponded to: a step size of 0.01 with 10101 steps, 0.1 with 1011 steps, and 1.0 with 101 steps, respectively—keeping to a total size of 101.
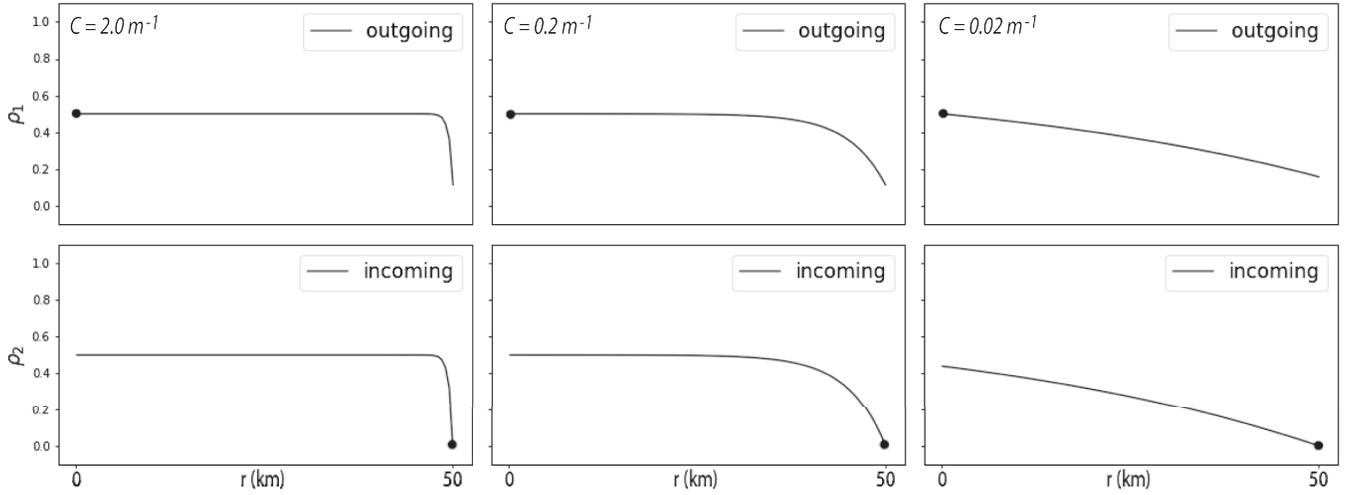
FIG. 2.   SDA result for the density profiles as functions of radius r of one outgoing (top) and one incoming beam (bottom), for three values of coefficients C. For left, middle, and right columns: $C = 2.0$, 0.2, and 0.02 m^{-1}, respectively. Circles denote locations of "measurements", or boundary conditions, provided to the procedure. The remaining trajectories are the SDA prediction. See the text for an interpretation in light of the expectation set forth in Sec. II.

show that the densities in those bins have been depleted to a lesser degree, as more particles escape the core. In the limit of no collisions ($C = 0$), the solution to Eq. (2) is trivial, and all particles escape.

(ii) The results are invariant, to one part in 10^6, across choices of discretized step size that span three orders of magnitude.

(iii) When the procedure receives one additional constraint such that the rates of depletion of the bin populations near R are captured, and meanwhile the value of C is withheld, the procedure is able to correctly *infer* the value of C that corresponds to those rates of depletion.

(iv) For all simulations performed, the plot of the cost function over the course of an annealing procedure (to be described below) shows that (i) all results converge; that is, solutions are consistent with both boundary conditions and model dynamics and (ii) each solution is unique; that is, all paths searched converge to one solution.

B. Effect of collision strength C

We examined the solution for three distinct values of collision strength $C^{\text{loss}} = C^{\text{gain}}$ to determine whether the optimization would recover our expectation that the rate of depletion of the number densities in the angular bins should slow as C weakens. Figure 2 shows the result. The left, middle, and right columns correspond to values of C of 2.0, $C = 0.2$, and 0.02 m^{-1}, respectively. The top and bottom rows show the density profile as a function of r for one outgoing beam ρ_1 and one incoming beam ρ_2, respectively. (The trajectories of the other two beams in the four-beam model are identical; not shown.) The circles denote the

locations of the "measurements" or boundary conditions, that were given as information to the SDA procedure. The remainder of the trajectories are the SDA prediction, given knowledge of the model dynamics.

From left to right in Fig. 2, note three effects. First, the number densities near $r = 0$ appear roughly the same, which follows from the argument regarding isotropy presented in Sec. II.

Second, as C weakens, the number densities fall off near $r = R$ increasingly gradually. Specifically, the predicted value for the outgoing beam ρ_1 at $r = R$ is 0.114 (left), 0.115 (middle), and 0.158 (right). This follows from the expectation that as the mean free path lengthens, more particles escape.

Third, as C weakens, fewer incoming particles are predicted to reach $r = 0$. Specifically, the value for the incoming beam ρ_2 at $r = 0$ is 5.0 (left), 4.9 (middle), and 4.4 (right). These numbers are summarized in Table I.

Recalling that the model contains a leak at $r = R$, we can consider the left, middle, and right panels of Fig. 2 to be a succession of solutions in time, each approximating a steady state over a timescale shorter than the leakage timescale. As C weakens from left to right, the time increases. In the limit where $C = 0$, the solution to

TABLE I.   Predicted values of outgoing and incoming beams at the endpoints, across values of C. As C dilutes from left to right, top: more outgoing neutrinos escape to $r = R$; bottom: fewer incoming neutrinos return to $r = 0$.

C (m^{-1})	2.0	0.2	0.02
ρ_1 at $r = R$ (outgoing)	0.114	0.115	0.158
ρ_2 at $r = 0$ (incoming)	0.5	0.499	0.44

Equation (2) is trivial, and all particles escape as time goes to infinity (not shown).

Finally, we make several notes regarding the robustness of the results. The procedure imposes the boundary conditions ($\rho_1 = 0.5$ at $r = 0$ and $\rho_2 = 0.0$ at $r = R$) to one part in 10^7. The results are invariant, to one part in 10^6, across the three choices of discretized step size dr corresponding to 1, 10, and 100 m.[10] The limit on that accuracy is likely due to a discretization error internal to Ipopt, the optimizer used in the procedure.

C. Correct inference of collision strength C as an unknown parameter

Having determined that the SDA procedure can predict the rate of depletion of the angular bins given a particular strength of coefficients C, we sought to determine whether it could perform the inverse: to infer C given a rate of depletion.

To this end, we repeated the experiment leaving C as an unknown parameter to be estimated. As before, we gave the SDA procedure the original four boundary conditions. In addition, from the original solutions shown in Fig. 2, we gave to the procedure the values of densities in each angular bin at five discretized locations prior to R, with the aim of capturing the rate of depletion of the bins. For each of the three cases, the procedure correctly estimated the unknown value of C to sixth-decimal precision.

D. Convergence and uniqueness

For a nonlinear model, the surface of the action, or cost function, will be nonconvex. The complete SDA procedure anneals in terms of the ratio of model and measurement error, R_f and R_m, respectively,[11] to gradually freeze out a lowest-minimum of the action [57]. This iteration works as follows.

We define the coefficient of measurement error R_m to be 1.0 and write the coefficient of model error R_f as $R_f = R_{f,0}\alpha^\beta$, where $R_{f,0} = 10^{-1}$, $\alpha = 1.5$, and β is initialized at zero. Parameter β is the annealing parameter. When $\beta = 0$, relatively free from model constraints, the action surface is smooth, and there exists one minimum of the variational problem that is consistent with the measurements. We estimate that minimum. Then we increase the weight of the model term slightly, via an integer increment in β, and recalculate the action. We do this recursively toward the deterministic limit of $R_f \gg R_m$. The aim is to

---

[10]For the reader interested in mapping these choices to a realistic physical scenario: a choice of $C = 0.2$ m^{-1}, which corresponds to a mean free path of 5 m, together with a step size dr of 1 m, would adequately sample a particle's path.

[11]More generally, R_m and R_f are inverse covariance matrices for the measurement and model errors, respectively. In this paper, the measurements are taken to be mutually independent, rendering these matrices diagonal.
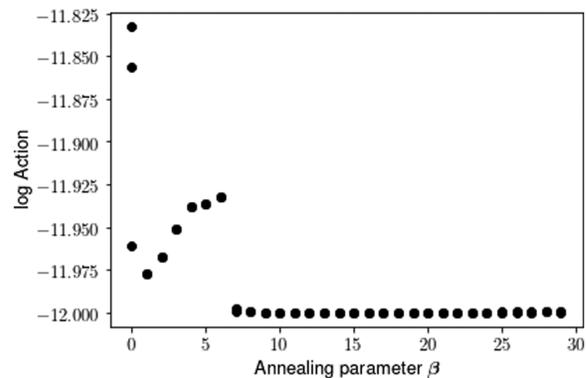
FIG. 3. Action, or cost function, as function of annealing parameter β, for a value of C of 0.2, representative of all simulations. As β passes 7, the action flatlines, indicating convergence to a solution that is consistent with both measurements (or boundary conditions) and model dynamics.

remain sufficiently near to the lowest minimum so as not to become trapped in a local minimum as the surface acquires the structure imposed by the model dynamics.

Reference [50] demonstrated that, for a particular path searched, the plot of action as a function of annealing parameter β indicates whether a particular solution is stable as the model dynamics are imposed increasingly rigidly. Figure 3 shows this result for the case of $C = 0.2$, representative of all experiments performed in this paper. Note that at a value of $\beta = 7$, the action flatlines. This behavior indicates that a solution has been found that is consistent with both imposed boundary conditions and with model dynamics. For a detailed explanation, see Ref. [50].

Finally, the 20 paths initialized for each experiment in this paper each converged to a single solution. This was demonstrated by identical state variable evolution (e.g., Fig. 2) and identical action(β) plots (e.g., Fig. 3) for each path sampled. Of course, to ascertain whether a solution is truly unique, one must sample an infinite number of paths. The specific number chosen for any particular model should depend on that model's complexity and dimensionality.

V. DISCUSSION

We have considered a simple two-point boundary value collision model with partially known boundary conditions and challenged an inference procedure to solve it. The model is sufficiently simple that a physical argument can intuit the expected result, thereby serving as a confidence check on the inference-based solution. Specifically, an optimization formulation of SDA recovers the expected angular distributions of the density profiles of particles in a collision-only flat two-dimensional model, where the particles represent neutrino-nucleon interactions near nuclear density. The behavior of the cost function over an iterative reweighting of the terms that impose the boundary

conditions and the equations of motion, respectively, demonstrates that the SDA solution is consistent with both the boundary conditions and the model dynamics. This was the case over a range of choices for the collision strength that spanned three orders of magnitude, and the results are robust to the step size used for discretization. Multiple randomly initiated searches of the state-and-parameter space converged to one solution.

This finding has important implications regarding the potential of inference to inform flavor transformation models governed by the full quantum kinetic equation (Eq. (1)), that is, taking the Hamiltonian together with direction-changing collisions. An inference formulation avoids reliance on initial conditions and rather seeks a solution that is compatible with constraints placed only at locations where we have high confidence in our understanding of the physics there. In the simulations performed in this paper, those constraints were placed at the end points, but they need not be. Further, within the optimization formulation, uniqueness can be investigated via the initialization of multiple independent searches. The ability of this SDA procedure to probe model degeneracy has been demonstrated in significantly more detail in Refs. [49,50].

It is relevant to note here another potential advantage of inference for augmenting numerical integration tools: efficiency. Including feedback from collisions self-consistently in the quantum kinetic calculations dramatically increases the computational complexity, as both the collisions and the flavor evolution will simultaneously shape the neutrino angular distributions [20]. The simple experiment presented in this paper did not offer the opportunity to showcase the high efficiency with which this SDA formulation performs state-and-parameter estimation; it has been demonstrated for a small-scale flavor evolution model in Ref. [50].

The question of *how* to implement inference within the existing large-scale numerical integration framework has yet to be explored. It is relevant here to note that the separation of coherent versus direction-changing scattering in Eq. (1)—the former into the Hamiltonian and the latter into the collision term C—is an artificial choice. It might be worthwhile to consider possible advantages of rewriting the quantum kinetic formulation to treat coherent and direction-changing scattering *collectively*. Of course, this is a daunting suggestion, as the formulation of Eq. (1) is the foundation of extremely large and powerful codes. Nevertheless, it is worth asking to what extent that formulation limits our ability to access the full dynamical range of the flavor field in compact object environments.

In parallel with exploring means to fold inference into the existing codes, it will be instructive to continue examining how inference performs on collision-only models of increasing complexity. For example, in the model used in this paper, neutrino number density was conserved at each collision; that is, collisions were completely elastic and did not absorb or emit neutrinos. Further, the matter density was held constant, with a discontinuous fall to zero at the core's edge. A more realistic matter profile will contain some radial dependence. In addition, a spherically symmetric geometry should be examined. Finally, because adding complexity will obscure a simple expectation regarding the result, it will be instructive to identify alternative comparisons to inference-based solutions, such as random-walk formulations.

[1] A. Mirizzi, I. Tamborra, H.-T. Janka, N. Saviano, K. Scholberg, R. Bollig, L. Hüdepohl, and S. Chakraborty, Riv. Nuovo Cimento **39**, 1 (2016).

[2] H. Duan, G. M. Fuller, and Y.-Z. Qian, Annu. Rev. Nucl. Part. Sci. **60**, 569 (2010).

[3] S. Chakraborty, R. Hansen, I. Izaguirre, and G. Raffelt, Nucl. Phys. **B908**, 366 (2016).

[4] S. Mikheev and A. Y. Smirnov, arXiv:0706.0454.

[5] S. Mikheev and A. Y. Smirnov, Sov. J. Nucl. Phys. **42**, 913 (1985).

[6] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).

[7] G. M. Fuller, R. W. Mayle, J. R. Wilson, and D. N. Schramm, Astrophys. J. **322**, 795 (1987).

[8] D. Nötzold and G. Raffelt, Nucl. Phys. **B307**, 924 (1988).

[9] J. Pantaleone, Phys. Lett. B **287**, 128 (1992).

[10] J. Pantaleone, Phys. Rev. D **46**, 510 (1992).

[11] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Phys. Rev. D **74**, 105014 (2006).

[12] H. Duan, G. M. Fuller, and J. Carlson, Comput. Sci. Discovery **1**, 015007 (2008).

[13] S. A. Richers, G. C. McLaughlin, J. P. Kneller, and A. Vlasenko, Phys. Rev. D **99**, 123014 (2019).

[14] J. F. Cherry, J. Carlson, A. Friedland, G. M. Fuller, and A. Vlasenko, Phys. Rev. Lett. **108,** 261104 (2012).
[15] J. F. Cherry, J. Carlson, A. Friedland, G. M. Fuller, and A. Vlasenko, Phys. Rev. D **87,** 085037 (2013).
[16] V. Cirigliano, M. Paris, and S. Shalgar, J. Cosmol. Astropart. Phys. 11 (2018) 019.
[17] M. Zaizen, J. F. Cherry, T. Takiwaki, S. Horiuchi, K. Kotake, H. Umeda, and T. Yoshida, J. Cosmol. Astropart. Phys. 06 (2020) 011.
[18] J. F. Cherry, G. M. Fuller, S. Horiuchi, K. Kotake, T. Takiwaki, and T. Fischer, Phys. Rev. D **102,** 023022 (2020).
[19] I. Tamborra and S. Shalgar, Annu. Rev. Nucl. Part. Sci. **71,** 165 (2020).
[20] S. Shalgar and I. Tamborra, Phys. Rev. D **103,** 063002 (2021).
[21] L. Johns, arXiv:2104.11369.
[22] H. Sasaki and T. Takiwaki, arXiv:2109.14011.
[23] J. D. Martin, J. Carlson, V. Cirigliano, and H. Duan, Phys. Rev. D **103,** 063001 (2021).
[24] S. Shalgar and I. Tamborra, Astrophys. J. **883,** 80 (2019).
[25] J. Kopp, J. High Energy Phys. 06 (2009) 049.
[26] E. K. Akhmedov, J. Kopp, and M. Lindner, J. High Energy Phys. 05 (2008) 005.
[27] A. Dolgov, O. Lychkovskiy, A. Mamonov, L. Okun, M. Rotaev, and M. Schepkin, Nucl. Phys. **B729,** 79 (2005).
[28] M. Beuthe, Phys. Rev. D **66,** 013003 (2002).
[29] A. Ioannisian and A. Pilaftsis, Phys. Rev. D **59,** 053003 (1999).
[30] C. Y. Cardall, Phys. Rev. D **61,** 073006 (2000).
[31] E. K. Akhmedov and J. Kopp, J. High Energy Phys. 04 (2010) 001.
[32] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM: Society for Industrial and Applied Mathematics, 2005).
[33] R. Kimura, J. Wind Eng. Ind. Aerodyn. **90,** 1403 (2002).
[34] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, Cambridge, England, 2003).
[35] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter* (Springer Science & Business Media, New York, 2009).
[36] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming* (SIAM: Society for Industrial and Applied Mathematics, 2010), Vol. 19.
[37] W. G. Whartenby, J. C. Quinn, and H. D. Abarbanel, Mon. Weather Rev. **141,** 2502 (2013).

[38] Z. An, D. Rey, J. Ye, and H. D. Abarbanel, Nonlinear Processes Geophys. **24** 9 (2017).
[39] S. J. Schiff, in *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, New York, 2009), pp. 3318–3321.
[40] B. A. Toth, M. Kostuk, C. D. Meliza, D. Margoliash, and H. D. Abarbanel, Biol. Cybern. **105,** 217 (2011).
[41] M. Kostuk, B. A. Toth, C. D. Meliza, D. Margoliash, and H. D. Abarbanel, Biol. Cybern. **106,** 155 (2012).
[42] F. Hamilton, T. Berry, N. Peixoto, and T. Sauer, Phys. Rev. E **88,** 052715 (2013).
[43] C. D. Meliza, M. Kostuk, H. Huang, A. Nogaret, D. Margoliash, and H. D. Abarbanel, Biol. Cybern. **108,** 495 (2014).
[44] A. Nogaret, C. D. Meliza, D. Margoliash, and H. D. Abarbanel, Sci. Rep. **6,** 1 (2016).
[45] E. Armstrong, Phys. Rev. E **101,** 012415 (2020).
[46] S. G. Djorgovski, C. Donalek, A. Mahabal, R. Williams, A. J. Drake, M. J. Graham, and E. Glikman, in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)* (2006), Vol. 1, pp. 856–863.
[47] N. Madhusudhan, Atmospheric retrieval of exoplanets, in *Handbook of Exoplanets*, edited by H. J. Deeg and J. A. Belmonte (Springer International Publishing, Cham, 2018), pp. 2153–2182.
[48] I. Kitiashvili and A. G. Kosovichev, Astrophys. J. **688,** L49 (2008).
[49] E. Armstrong, A. V. Patwardhan, L. Johns, C. T. Kishimoto, H. D. Abarbanel, and G. M. Fuller, Phys. Rev. D **96,** 083008 (2017).
[50] E. Armstrong, A. V. Patwardhan, E. Rrapaj, S. F. Ardizi, and G. M. Fuller, Phys. Rev. D **102,** 043013 (2020).
[51] E. Rrapaj, A. V. Patwardhan, E. Armstrong, and G. M. Fuller, Phys. Rev. D **103,** 043006 (2021).
[52] C. Volpe, Int. J. Mod. Phys. E **24,** 1541009 (2015).
[53] A. Vlasenko, G. M. Fuller, and V. Cirigliano, Phys. Rev. D **89,** 105004 (2014).
[54] H. D. Abarbanel, *Predicting the Future* (Springer, New York, 2013), Vol. 10.
[55] A. Wächter, in Dagstuhl Seminar Proceedings (Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2009).
[56] minAone interface with Interior-point Optimizer, https://github.com/yejingxin/minAone (accessed: 2021-10-13).
[57] J. Ye, D. Rey, N. Kadakia, M. Eldridge, U. I. Morone, P. Rozdeba, H. D. Abarbanel, and J. C. Quinn, Phys. Rev. E **92,** 052901 (2015).