An optimization-based approach to calculating neutrino flavor evolution

Eve Armstrong,[1,*] Amol V. Patwardhan,[2,3,†] Lucas Johns,[2,3,‡] Chad T. Kishimoto,[4,3,§] Henry D. I. Abarbanel,[2,5,‖] and George M. Fuller[2,3,¶]

[1]*BioCircuits Institute, University of California, San Diego, La Jolla, California 92093-0328, USA*
[2]*Department of Physics, University of California, San Diego, La Jolla, California 92093-0319, USA*
[3]*Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla, California 92093-0424, USA*
[4]*Department of Physics and Biophysics, University of San Diego, San Diego, California 92110, USA*
[5]*Marine Physical Laboratory (Scripps Institution of Oceanography), University of California, San Diego, La Jolla, California 92093-0210, USA*

(Received 14 December 2016; published 10 October 2017)

We assess the utility of an optimization-based data assimilation (D.A.) technique for treating the problem of nonlinear neutrino flavor transformation in core-collapse supernovae. D.A. uses measurements obtained from a physical system to estimate the state variable evolution and parameter values of the associated model. Formulated as an optimization procedure, D.A. can offer an integration-blind approach to predicting model evolution, which offers an advantage for models that thwart solution via traditional numerical integration techniques. Further, D.A. performs most optimally for models whose equations of motion are nonlinearly coupled. In this exploratory work, we consider a simple steady-state model with two monoenergetic neutrino beams coherently interacting with each other and a background medium. As this model can be solved via numerical integration, we have an independent consistency check for D.A. solutions. We find that the procedure can capture key features of flavor evolution over the entire trajectory, even given measurements of neutrino flavor only at the endpoint, and with an assumed known initial flavor distribution. Further, the procedure permits an examination of the sensitivity of flavor evolution to estimates of unknown model parameters, locates degeneracies in parameter space, and can identify the specific measurements required to break those degeneracies.

DOI: 10.1103/PhysRevD.96.083008

I. INTRODUCTION

We assess the efficacy of a data assimilation (D.A.) technique for constraining neutrino flavor evolution histories inside core-collapse supernovae. The specific technique of interest in this paper is an integration-blind procedure, which offers advantages for problems that thwart traditional numerical integration techniques. The procedure is crafted to efficiently find solutions for an extremized action from a sparse set of measurements [1]. Importantly, the technique of transporting information from the measured quantities through the complete model to the *unmeasured* quantities requires that the model's differential equations be coupled. For this reason, D.A. lends itself particularly well to the exploration of nonlinear systems. Notably, collective neutrino oscillation phenomena in astrophysical environments are essentially nonlinear. This feature begs the question of whether a D.A. approach to solving these problems is feasible. Here we investigate this issue of feasibility in the context of a simple toy model that captures nonlinearity.

Neutrino flavor evolution in compact object environments is a vexing and unsolved problem [2–52]. It is inherently nonlinear and fraught with difficulties. These difficulties are exacerbated by the limitations of our understanding of key supernova physics, for example: the equation of state and weak interaction properties of nuclei and nuclear matter in hot and dense conditions.

Even accounting for the inherent uncertainties in supernova microphysics, obtaining convincing numerical simulations of supernova neutrino flavor histories inside the supernova remains problematic. In just one example, inelastic neutrino backscattering could contribute to flavor evolution in some regimes [27,34,35], giving rise to the "neutrino halo" problem. The neutrino halo effect changes the flavor evolution problem from an initial value problem with neutrino fluxes and flavor content specified at the edge of the proto-neutron star (the "neutrino sphere"), to something more akin to a boundary value problem, where flavor phase information is propagating both outward and *inward*. The computational difficulties endemic to neutrino direction-changing scattering represent just one of the many challenges we face in transitioning from the usual coherent "index of refraction" treatment of neutrino flavor evolution to a full

*earmstrong@ucsd.edu
†apatward@ucsd.edu
‡ljohns@physics.ucsd.edu
§ckishimoto@sandiego.edu
‖habarbanel@ucsd.edu
¶gfuller@ucsd.edu

quantum kinetic approach [53–69]. Other outstanding problems in supernova neutrino flavor physics include fast flavor conversion [48–51] and spatial and temporal instabilities [26,36,37,40,52]. For background on neutrino physics in massive stars and supernovae, see Appendix D.

Against this backdrop of uncertainty in theoretical calculations, there is an effort to configure water or ice Čerenkov (for example, HyperK, IceCube); liquid argon (for example, DUNE); and liquid scintillator detectors (for example, JUNO), to be able to capture the neutrino signal from a Galactic core-collapse supernova [70–77]. The potential for such a detection to provide a probe of beyond-standard-model physics in the neutrino sector and insight into the nuclear equation of state, along with a host of astrophysical issues like neutrino heating and nucleosynthesis, is alluring. Consequently, the stakes are high when it comes to gaining confidence in modeling nonlinear neutrino flavor conversion. Given this context, it is worth exploring new techniques.

The extraction of information from measurements is a general procedure known in the geosciences as data assimilation [78–81]. D.A. has been used commonly in fluid dynamics [82,83] and more recently in neuroscience [84–87], where the available data from a physical system are sparse and the corresponding model consists of degrees of freedom coupled in a nonlinear manner. The aim of D.A. is to incorporate information contained in measurements directly into a model, to estimate unknown parameters and the dynamics of the model state variables—both measured and unmeasured. The test of a successful estimation is the ability of the completed model to predict the system state outside the times or locations at which the measurements were obtained.

D.A. has two key advantages over numerical integration. First, when cast within the framework of optimization, it can be written as an integration-blind procedure. An integration-blind approach may be amenable to boundary-value problems for which a solution via forward integration is unrealistic. Second, the procedure can systematically and efficiently identify the existence of degenerate solutions, and the specific measurements that are required to break degeneracy. For these reasons, we considered D.A. worth exploring as a possible alternative attack to the standard initial-value treatment of neutrino flavor evolution employed in, for example, the "bulb" model [2].

We examine the potential utility of D.A. in the context of collective neutrino oscillations by applying it to a simple model that can be solved via numerical integration—and thus where there exists an independent consistency check for D.A. solutions. The model describes the flavor evolution of two monoenergetic neutrino beams emanating from a supernova event. The measurements used are the flavor content of each neutrino beam at some final radius $r = R$ at which a detector might be placed, given an assumed known initial flavor distribution at the surface of the "neutrino sphere" ($r = 0$). We seek to determine whether the sparse measurements suffice to yield the flavor evolution history over the interim distance, and to estimate unknown model parameter values: namely, the strength of neutrino coupling to matter and the strength of neutrino-neutrino coupling.

In this first exploratory paper, three main results emerge. First, over repeated trials we obtain realistic overall flavor evolution history for both neutrinos, even given the extreme sparsity of measurements. We shall describe exceptions to this finding in Sec. V. Second, we find that generally the given measurements are insufficient to distinguish among multiple parameter solution sets. Third, and as a consequence of the first and second findings, we gain insight regarding the sensitivity of flavor evolution to these parameter values.

This paper proceeds as follows.

(i) Section II, Inverse Problems and Optimization, explains the general framework for data assimilation via optimization, and it states the specific objective function and method of evaluation used in this paper.
(ii) Section III, Model, describes the specific model used in this paper: a simplified version of the dynamics of neutrino flavor evolution that ensues from the surface of a core-collapse supernova event.
(iii) Section IV, The Experiments, explains the full procedure for simulated experiments given to an optimization algorithm, and our physical rationale for the experimental designs.
(iv) Section V, Results, describes the solutions.
(v) In Sec. VI, Discussion, we comment on the implications of the results with respect to more realistic problems in neutrino astrophysics, and we describe immediate future work.
(vi) Section VII, Summary, contains concluding remarks.
(vii) Appendix A gives the path-integral derivation of the objective function used in this paper.
(viii) Appendix B gives details of the optimization procedure.
(ix) Appendix C gives an overview of how to interpret our simplistic model in terms of a constant-entropy envelope model.
(x) Appendix D gives background on relevant neutrino astrophysics.

II. INVERSE PROBLEMS AND OPTIMIZATION

A. General framework

D.A. is a procedure whereby information in measurements is used to complete a model of the system from which the measurements were obtained (see Ref. [88] for an introduction to this "inverse problem" formulation). The model F is written as a set of D ordinary differential equations (ODE) that evolve in affine parameter r as

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

where the components x_a of the vector $\boldsymbol{x}$ are the model state variables. The affine parametrization r may be, for

example, time or distance. The unknown model parameters to be estimated are contained in $\boldsymbol{p}$; note that the model evolution depends on $\boldsymbol{p}$.

A subset L of the D state variables is associated with measured quantities. One seeks to estimate the p unknown parameters and the evolution of all state variables given the provided measurements, and to then use those estimates to predict the model evolution in regions where there exist no measurements. The prediction phase is the test of estimation quality.

In the simulated experiments described in this paper, we integrate forward the equations of motion from an initial known state at $r = 0$ (the surface of the neutrino sphere), to obtain a "measurement" of L variables at some detector location R. The question for D.A. is whether sufficient information can be propagated through the coupled equations—from measured to unmeasured variables—such that the parameters that had generated the measurements can be inferred. As noted, the test of a successful parameter estimation is its ability to predict the evolution of all D state variables in the interim $r \in (0, R)$ during which no measurements are obtained.

B. Specific optimization formulation used in this paper

One method commonly employed to solve an inverse problem is optimization. Optimization is the process of finding the extremum of a function, called the "cost (or objective or penalty) function." The cost function used in this paper is motivated from a path-integral-like formulation of D.A., and for this reason we nickname it an "action."

In constructing the action that will be used to yield parameter estimates, we consider three factors that will dictate those estimates: (1) measurements obtained from the physical system of interest, (2) the dynamics of the model describing that system, and (3) additional equality constraints that are specific to the model formulation. The means by which we incorporate information from measurements, and the manner in which each of the above three factors is considered, may be best understood via an examination of this cost function's specific formulation. We do this now, and then proceed to describe each term in turn.

The action A_0 used in this paper is written as

$$
\begin{aligned}
A_0 = {} & \frac{R_f}{(N-1)D} \sum_{n\in\{\text{odd}\}}^{N-2} \sum_{a=1}^{D} \left[\left\{ x_a(n+2) - x_a(n) - \frac{\delta r}{6} [F_a(\boldsymbol{x}(n),\boldsymbol{p}) + 4F_a(\boldsymbol{x}(n+1),\boldsymbol{p}) + F_a(\boldsymbol{x}(n+2),\boldsymbol{p})] \right\}^2 \right. \\
& \left. + \left\{ x_a(n+1) - \frac{1}{2}(x_a(n) + x_a(n+2)) - \frac{\delta r}{8} [F_a(\boldsymbol{x}(n),\boldsymbol{p}) - F_a(\boldsymbol{x}(n+2),\boldsymbol{p})] \right\}^2 \right] \\
& + \frac{R_m}{N_{\text{meas}}} \sum_j \sum_{l=1}^{L} (y_l(j) - x_l(j))^2 + k \sum_n^N |g_1(\boldsymbol{x}(n)) - 1|^2 + k \sum_n^N |g_2(\boldsymbol{x}(n)) - 1|^2. \qquad (1)
\end{aligned}
$$

We seek the path $\boldsymbol{X}^0 = \{\boldsymbol{x}(0), \ldots, \boldsymbol{x}(N), \boldsymbol{p}\}$ in state space on which A_0 attains a minimum value.

The two squared terms in the first double sum in Eq. (1) incorporate the model evolution of all D state variables x_a. Of these, the first term in curly braces represents error in the first derivative (with respect to r) of the state variables, whereas the second term corresponds to error in the second derivative. These terms can be derived from a consideration of Markov-chain transition probabilities. Here, the outer summation in n is taken over all odd-numbered grid points —discretized steps in r that parametrize the model equations of motion. The step size δr is defined as the distance between alternate grid points: $\delta r \equiv r_{n+2} - r_n$. The inner summation in a is taken over all D state variables.

The squared term in the second double sum governs the transfer of information from measurements y_l to states x_l. It derives from the concept of conditional mutual information of probability theory. The y_l are the measurements, and x_l are the model variables corresponding to the measurements. Here, summation on j runs over the set of all N_{meas} discretized grid points where the measurements are made, which may in general be some subset of all the model grid points. The summation in l is taken over the L measured quantities.

R_m and R_f are inverse covariance matrices for the measurement and model errors, respectively. In this paper we take the measurements to be mutually independent and the state variables to be independent, rendering these matrices diagonal. Additionally, we constrain R_f to be uniform across all state variables, and likewise R_m for all measurements. For our purposes, R_m and R_f are relative weighting terms; the utility of relative weighting will be described below in this section. For a short derivation of the first two terms, see Appendix A; for a full treatment, see Ref. [1].

The third and fourth terms (with coefficients k) are equality constraints, which were added to increase the efficiency of the search algorithm. These will be written out explicitly in Sec. III.

The optimization is performed at all locations along a path simultaneously, so as not to impart greater importance to a measurement at any particular location over another.

An integration-blind technique may lend itself well to problems that cannot be solved in a straightforward manner via forward integration. An example of such a problem—the "neutrino halo" —is discussed in Appendix D.

To minimize A_0 we employ a variational approach, via the open-source Interior-Point Optimizer (Ipopt) [89]; see Sec. IV.

C. Iterative reweighting of measurement and model contributions ("annealing")

The complete D.A. procedure involves an iteration that is aimed at identifying the parameter set corresponding to the global minimum of the action. Local minima will represent degenerate sets of parameter estimates that may fit the measurements well but which are poor predictors of state variable evolution outside of the locations at which the measurements were obtained. The remedy consists of recursively calculating A_0 as the ratio of the model and the measurement coefficients, R_f and R_m, respectively, is gradually increased. Specifically, we define $R_f = R_{f,0}\alpha^\beta$, where α is a small number greater than 1, and β is increased from 0 in uniform increments. For all simulations performed in this paper, R_f takes a uniform value over all state variables x_a. R_m is uniform over all measured variables x_l, and it is held fixed across iterations while R_f is incremented. For an explanation of why this iterative procedure—which we call "annealing"—aids in identifying the global minimum, see Appendix A.

III. MODEL

A. Toy model motivation and scheme

A complete description of neutrino flavor evolution in realistic astrophysical environments, such as supernovae, involves complications imposed by geometry, multiangle effects, realistic emission spectra, and nonforward scattering, among other effects. Turning D.A. into a useful tool to simulate the dynamics in such a model, while matching the fidelity and sophistication of current forward-integration codes like BULB, is a daunting task. We shall not attempt this here. Rather, our motivation in this paper is to take a first step toward assessing the efficacy of D.A. in treating the astrophysical neutrino flavor transformation problem. This first step requires the use of a vastly simplified model.

The model we craft possesses two key features. First, it is nonlinear—a key aspect of the physics that gives rise to collective neutrino flavor evolution in these environments. Notably, D.A. is particularly useful for estimating model evolution and parameter values in nonlinear models where only a subset of the state variables can be accessed experimentally. Such is the case for neutrino flavor evolution. A second key point is that the model is sufficiently simple to be solvable via traditional forward-integration techniques. This feature enables an independent consistency check for D.A. solutions.

For our model, we consider a scenario in which two monoenergetic neutrino beams with different energies interact with each other and with a background consisting of nuclei, free nucleons, and electrons. The densities of the background particles and of the neutrino beams themselves are taken to dilute as some functions of a position coordinate r, which could be interpreted as the distance from the neutrino sphere in a supernova. In what follows, we first discuss general two-flavor neutrino flavor oscillations, and we then adapt this formalism for our particular toy model.

B. Two-flavor neutrino flavor evolution

Since ν_μ and ν_τ neutrino flavors experience identical interactions in the supernova environment, the three-flavor problem can be reduced to a two-flavor mixing between ν_e and a state ν_x that is a particular superposition of ν_μ and ν_τ [90,91]. The flavor state of neutrinos of energy E, as a function of position, can then be expressed using a 2×2 density matrix, which in the flavor basis is given by

$$\rho_E(r) = \begin{pmatrix} \rho_{ee,E} & \rho_{ex,E} \\ \rho_{xe,E} & \rho_{xx,E} \end{pmatrix} = \begin{pmatrix} |a_{e,E}|^2 & a^*_{e,E}a_{x,E} \\ a_{e,E}a^*_{x,E} & |a_{x,E}|^2 \end{pmatrix}. \quad (2)$$

Here the last matrix representation of the density operator is for the special case where the neutrinos are in pure states, with $a_{e,E}$ and $a_{x,E}$ being the respective flavor amplitudes. The quantum kinetic equation (QKE) governing the evolution of the general density operator ρ_E has the form [16,53–69,92,93]

$$i\frac{d\rho_E(r)}{dr} = [H_E(r), \rho_E(r)] + i\mathcal{C}_E(r). \quad (3)$$

Here we are assuming that the neutrino density matrix elements and potentials carry no explicit time dependence; that is, they may vary only as functions of position along the neutrino trajectory—-a steady-state solution. $H_E(r)$ is the Hamiltonian driving coherent flavor evolution. The last term on the right side, $\mathcal{C}_E(r)$, captures the effects of collisions. Neglecting collisions, which *may* be justified in some supernova regions and epochs [35,55], results in the coherent limit in which neutrino flavor evolution is Schrödinger-like:

$$i\frac{\mathrm{d}\rho_E(r)}{\mathrm{d}r} = [H_E(r), \rho_E(r)]. \quad (4)$$

Here, the right side is trace-conserving, implying unitary evolution. Equation (4) can also be cast in the form of a standard path-integral extremization problem [24].

1. Spin basis and polarization vectors

Equation (4) can be recast in terms of Bloch vectors $\boldsymbol{P}_E$ and $\boldsymbol{H}_E$ by decomposing the density matrices ρ_E and Hamiltonians H_E, respectively, into a basis of Pauli spin matrices (for details see Ref. [65]):

$$\begin{aligned}\rho_E &= \frac{1}{2}(P_{E,0}\mathbb{I} + \boldsymbol{P}_E\cdot\boldsymbol{\sigma}) \\ &= \frac{1}{2}\begin{pmatrix} P_{E,0}+P_{E,z} & P_{E,x}-iP_{E,y} \\ P_{E,x}+iP_{E,y} & P_{E,0}-P_{E,z}\end{pmatrix},\end{aligned} \tag{5}$$

$$\begin{aligned}H_E &= \frac{1}{2}(H_{E,0}\mathbb{I} + \boldsymbol{H}_E\cdot\boldsymbol{\sigma}) \\ &= \frac{1}{2}\begin{pmatrix} H_{E,0}+H_{E,z} & H_{E,x}-iH_{E,y} \\ H_{E,x}+iH_{E,y} & H_{E,0}-H_{E,z}\end{pmatrix}.\end{aligned} \tag{6}$$

We refer to the quantities $\boldsymbol{P}_E$ as "polarization vectors," whereas the vectors $\boldsymbol{H}_E$ will inherit the name "Hamiltonians." Note that the subscripts x, y, z on the vector components above do not refer to spatial coordinates, but rather to directions in this $SU(2)$ "Bloch space." The advantage of Bloch-vector decomposition is twofold: (1) the dynamical variables of the system, that is, the components of $\boldsymbol{P}_E$, are now real numbers, unlike the complex amplitudes in the density matrices, and (2) the geometric representation makes for easier visualization of the often complex underlying dynamics. To illustrate the second point, we write the evolution equation in terms of these Bloch vectors:

$$\frac{\mathrm{d}P_{E,0}}{\mathrm{d}} = 0, \quad \text{and} \quad \frac{\mathrm{d}\boldsymbol{P_E}}{\mathrm{d}r} = \boldsymbol{H}_E(r)\times\boldsymbol{P}_E(r). \tag{7}$$

The first equation is simply a restatement of trace preservation, and in fact, for a normalized density matrix, $P_{E,0}=1$. The second equation, on the other hand, resembles Larmor precession of a magnetic moment, with the Hamiltonian in this case playing the role of a magnetic field. Note that $P_{E,z}$ represents the probability of a neutrino being detected as a ν_e over ν_x; that is, $P_{E,z} = |a_{e,E}|^2 - |a_{x,E}|^2$. Or, if one were to think in terms of a population of neutrinos,

$$P_{E,z}(r) = \frac{n_{\nu_e,E}(r) - n_{\nu_x,E}(r)}{n_{\nu,E}(r)}, \tag{8}$$

where $n_{\nu,E}(r)$ is the number density of neutrinos of energy E at a position r, and $n_{\nu_\alpha,E}(r) \equiv n_{\nu,E}(r)|a_{\alpha,E}(r)|^2$ is the "expected" number density of neutrinos in the flavor α at that energy and position.

2. The Hamiltonian and equations of motion for neutrino forward scattering

Having set up the evolution equations, let us now describe the specific Hamiltonians that drive flavor evolution in the coherent limit—first in matrix form and then in the Pauli spin representation. The Hamiltonian H_E consists of three contributing terms: $H_E(r) = H_{\mathrm{vac},E} + H_m(r) + H_{\nu\nu}(r)$. The first of these terms drives flavor oscillations in vacuum:

$$H_{\mathrm{vac},E} = \frac{\Delta}{2}\begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta\end{pmatrix}, \tag{9}$$

where θ is the mixing angle in vacuum, describing the unitary transformation between the weak interaction (flavor) eigenstates, and the energy (mass) eigenstates. Also, $\Delta \equiv \delta m^2/2E$, where δm^2 is the mass-squared splitting between the two energy eigenstates. The other two contributions arise from neutrino forward scattering on background matter particles (H_m, the "matter Hamiltonian") and other neutrinos ($H_{\nu\nu}$, the "neutrino-neutrino Hamiltonian"), respectively. In the scenario described above, they assume the forms

$$H_m = \frac{V(r)}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}, \tag{10}$$

$$H_{\nu\nu} = \sum_E \mu_E(r)\rho_E(r). \tag{11}$$

Here we have already subtracted the trace from the vacuum and matter Hamiltonian, since it has no bearing on the flavor evolution. In terms of the baryon number density $n_B(r)$, the electron fraction $Y_e(r)$, and the neutrino number density $n_{\nu,E}(r)$, the potentials are

$$\begin{aligned}V(r) &= \sqrt{2}G_F n_B(r) Y_e(r), \\ \mu_E(r) &= \sqrt{2}G_F \alpha(r) n_{\nu,E}(r),\end{aligned} \tag{12}$$

where G_F is the Fermi constant and $\alpha(r) \equiv 1-\cos\psi(r)$ is a factor that weights the neutrino-neutrino coupling according to the intersection angle $\psi(r)$ between the two neutrino streams. Our choice of particular functional forms for $V(r)$ and $\mu_E(r)$ is stated and explained later in this section (Sec. III B 4) and in Appendix C. The neutrino-neutrino Hamiltonian $H_{\nu\nu}$ ensures that the evolution equations are nonlinear, since the Hamiltonians driving the evolution of the density matrices depend on the density matrices themselves; moreover, the evolution histories of the two neutrino populations are now coupled to each other.

Gathering the above Hamiltonians and expressing them in the Pauli basis, one obtains the complete set of dynamical equations for the two neutrino beams:

$$\begin{aligned}
\frac{dP_{1,x}}{dr} &= (\Delta\cos 2\theta - V(r))P_{1,y} \\
&\quad + \mu(r)(P_{2,y}P_{1,z} - P_{2,z}P_{1,y}), \\
\frac{dP_{1,y}}{dr} &= -(\Delta\cos 2\theta - V(r))P_{1,x} - \Delta\sin 2\theta P_{1,z} \\
&\quad + \mu(r)(P_{2,z}P_{1,x} - P_{2,x}P_{z,1}), \\
\frac{dP_{1,z}}{dr} &= \Delta\sin 2\theta P_{1,y} \\
&\quad + \mu(r)(P_{2,x}P_{1,y} - P_{2,y}P_{1,x}), \qquad (13) \\
\frac{dP_{2,x}}{dr} &= (\Delta\cos 2\theta - V(r))P_{2,y} \\
&\quad + \mu(r)(P_{1,y}P_{2,z} - P_{1,z}P_{2,y}), \\
\frac{dP_{2,y}}{dr} &= -(\Delta\cos 2\theta - V(r))P_{2,x} - \Delta\sin 2\theta P_{2,z} \\
&\quad + \mu(r)(P_{1,z}P_{2,x} - P_{1,x}P_{z,2}), \\
\frac{dP_{2,z}}{dr} &= \Delta\sin 2\theta P_{2,y} + \mu(r)(P_{1,x}P_{2,y} - P_{1,y}P_{2,x}), \qquad (14)
\end{aligned}$$

where for simplicity we have assumed equal neutrino number densities at both energies $[n_{\nu,E_1}(r) = n_{\nu,E_2}(r)]$, so that $\mu_{E_1}(r) = \mu_{E_2}(r) = \mu(r)$. For brevity, we have used P_1 and P_2 in place of P_{E_1} and P_{E_2}.

3. Physics of the model: MSW resonance and collective effects

In principle, the various Hamiltonians driving neutrino flavor evolution can—in the adiabatic limit—be combined and expressed in the form of effective in-medium oscillation parameters:

$$\begin{aligned}
&H_{\rm vac} + H_m(r) + H_{\nu\nu}(r) \\
&\quad \equiv \frac{\Delta_m(r)}{2}\begin{pmatrix} -\cos 2\theta_m(r) & \sin 2\theta_m(r) \\ \sin 2\theta_m(r) & \cos 2\theta_m(r) \end{pmatrix}, \qquad (15)
\end{aligned}$$

where $\Delta_m(r) = \sqrt{(\Delta\cos 2\theta - V_{\rm eff}(r))^2 + \Delta^2\sin^2 2\theta}$ and $\sin^2 2\theta_m(r) = (\Delta^2\sin^2 2\theta)/\Delta_m^2(r)$ represent the effective in-medium mass-squared difference and mixing angle, respectively. Here, $V_{\rm eff}$ is taken to represent the effective matter + collective potential experienced by a neutrino. At $V_{\rm eff}(r) = \Delta\cos 2\theta$, the in-medium mixing angle θ_m achieves its maximal value of $\pi/4$ and flavor transformation becomes resonant. A system that passes adiabatically through this resonance is susceptible to highly efficient flavor conversion through the Mikheyev-Smirnov-Wolfenstein (MSW) mechanism [94,95], which is an essential feature of the scenarios we treat in this study. Near the neutrino sphere, H_m dominates, so that the heavier mass eigenstate essentially aligns with the ν_e flavor state. At large radii, however, $H_{\rm vac}$ takes over, and for small mixing angles, this means the heavier mass eigenstate aligning more closely with ν_x. If this transition is sufficiently adiabatic, a neutrino initially emitted as ν_e undergoes near-complete conversion into ν_x prior to its detection.

In the numerical calculations discussed below we choose neutrino energy ratios and matter potentials that can encompass highly adiabatic neutrino flavor transformation, so that neutrinos stay in instantaneous mass eigenstates. Knowing what the flavor states of neutrinos are at the beginning of our calculations, that is, at the neutrino sphere, we can then determine the flavor states at the end, without knowing the precise details of the intervening matter density profile. It is because of this reason that adiabatic neutrino flavor evolution presents a fundamental problem in interpreting a detected core-collapse neutrino signature: possible degeneracy of neutrino flavor histories. That is, any number of smooth matter density profiles, each transited by neutrinos adiabatically, will facilitate conversion of an initial ν_e into a ν_x, or vice versa. A key objective of this study is to ascertain whether optimization techniques can map out degeneracies. In Sec. VI, we suggest that introducing additional complexities in our model, including sharp features in the matter potential (such as shocks) that would engineer nonadiabaticity, can help break such degeneracies [96–99].

Introducing neutrino-neutrino coupling into this picture gives rise to an array of nonlinear flavor-transformation phenomena. Nonlinearity can manifest as various modes of collective neutrino flavor oscillation—see Ref. [12] for a review. In these collective modes significant fractions of the neutrinos in a range of energies and locations may undergo simultaneous, sometimes synchronized coherent flavor oscillations. In essence, neutrino-neutrino forward scattering serves to "inform" a neutrino about the flavor states of others, and the nonlinear nature of the interactions guides neutrino flavor states into lock-step coherence. Determining the locations in radius and energy of the transition in and out of such collective modes, or whether they even occur at all, will be an important objective for core-collapse supernova neutrino burst detection.

In a practical sense, collectivity engendered by nonlinear neutrino-neutrino forward-scattering potentials may add to the possible degeneracy in neutrino flavor histories, or it may tend to narrow the possibilities. To use optimization techniques to explore this question, we will now present simple functional forms for matter and neutrino-neutrino potentials.

4. Choice of the matter potential and the neutrino coupling term

The above formulation of 2×2 neutrino flavor evolution is general and has been used to calculate collective neutrino oscillation phenomena, often capturing the qualitative behavior of more sophisticated 3×3 multiangle simulations. Here, we seek to use D.A. to solve the two-flavor evolution embodied in an appropriately adapted version of Eqs. (13) and (14). This requires choices for the matter and neutrino background potentials.

The matter potential $V(r)$ is typically written in terms of the baryon density $n_B(r)$ and the electron fraction $Y_e(r)$. For simplicity we combine the two dependences and describe $V(r)$ using a single power law,[1]

$$V(r) = \frac{C}{r^3}, \tag{16}$$

where all constants, including the weak coupling G_F, as well as physical parameters such as the neutrino sphere radius, and n_B and Y_e at the neutrino sphere, have been absorbed into the dimensionful constant C, which we treat as a parameter to be determined by the data assimilation procedure.

We also choose a simplified structure for the neutrino-neutrino coupling term $\mu(r)$. Here, the dependences of the neutrino number density $n_{\nu,E}(r)$ and the effect of the intersection angle $\alpha(r) \equiv 1 - \cos\psi(r)$ are bundled together into a single power law,

$$\mu(r) = \frac{Q}{r^3}, \tag{17}$$

where, as with the matter potential, all constant parameters are absorbed into Q, which we will treat as a single parameter to be determined by the data assimilation procedure. These functional forms are adopted as coarse mock-ups of the matter and neutrino densities surrounding the neutrino sphere and are not meant to emulate realistic profiles to an accurate degree. Nevertheless, the physical motivation behind the choice of these functional forms is discussed in Appendix C.

The challenge posed to the data-assimilation machinery is to estimate the constants C and Q as well as the flavor-space trajectories $\boldsymbol{P}_1(r)$ and $\boldsymbol{P}_2(r)$ of the two neutrino beams as they propagate outward from the neutrino sphere (radius $r = 0$) towards some radius R. We imagine that a detector sits at R. For this exploratory D.A. study, we have chosen energies that allow us to examine how the procedure operates over different resonance locations relative to the detector location. Our motivation is to probe the utility of eventually adding constraints on physics within the envelope. The inputs to the D.A. machinery are (1) the model equations of motion and (2) the measurements P_z of each neutrino at $r = R$ (given a known initial state at $r = 0$), which are processed through the action-minimization procedure detailed in the previous section.

IV. THE EXPERIMENTS

A. The model-specific optimization procedure

Given (1) the model embodied in Eqs. (13) and (14), and with unknown parameters C (the weight of the matter potential) and Q (the weight of the coupling potential); (2) measurements of the model state variables $P_{1,z}$ and $P_{2,z}$ at $r = R$; and (3) an assumed initial known flavor state (P_z) of each neutrino at $r = 0$, we seek to identify the path $\boldsymbol{X}^0 = \{\boldsymbol{P}_1(0), \boldsymbol{P}_2(0), \ldots, \boldsymbol{P}_1(R), \boldsymbol{P}_2(R), C, Q\}$ in state space such that the cost function of Eq. (1) attains a minimum value. Within our model formulation, the measurements y_l are the P_z components of both neutrinos, and the components of $\boldsymbol{x}$ are the P_x, P_y, and P_z values of both neutrinos. (Obviously, the only potential "measurement" of supernova neutrinos would be in a terrestrial detector, and would correspond to an energy spectrum; see Sec. VI.)

The two equality constraints (with coefficients k) are designed to improve the efficiency of the search algorithm. The algorithm does not recognize relations among nonindependent state variables, but rather considers each independently. Because the model described by Eqs. (13) and (14) implicitly imposes $\boldsymbol{P}^2 =$ constant, the model is overdetermined in the Cartesian coordinate system. To minimize the computational expense, we added these equality constraints to strictly impose unitarity at the start of the annealing procedure ($R_m \gg R_f$), in which regime the model weight may not yet be sufficiently strong for its implicit requirement to be well respected. The functions g_1 and g_2 are

$$g_1(\boldsymbol{x}(n)) = P_{1,x}^2 + P_{1,y}^2 + P_{1,z}^2$$
$$g_2(\boldsymbol{x}(n)) = P_{2,x}^2 + P_{2,y}^2 + P_{2,z}^2;$$

the value of coefficient k in Eq. (1) was taken to be 1.

B. The experimental designs

We designed two sets of experiments for D.A., where the sets are distinguished by the ratio of the neutrino energies. In both cases the first neutrino (Neutrino 1) experiences the MSW resonance at a radius $r = 1.1$ in the absence of coupling ($Q = 0$). [We express both the location r and matter and neutrino potential coefficients C and Q as dimensionless quantities. We can provide dimensions, for example in cm and MeV cm^3, respectively, through Eqs. (18) and (C2); see Appendix C.] This requirement sets the value of the matter coefficient C. In the first set of experiments, the ratio of neutrino energies $E_{\nu_1}/E_{\nu_2} = 2.5$; in the second set of experiments, $E_{\nu_1}/E_{\nu_2} = 0.01$. For each of these two cases, we examined the model dynamics for four values of the coupling strength: $Q = 0$, 1, 100, and 1000. These choices were made to permit an examination of whether the quality of results is sensitive to neutrino coupling strength or energy ratio.

For all experiments, we assumed the neutrinos to be in pure ν_e flavor at $r = 0$, with a corresponding P_z value of $+1$. Here we make a note regarding our treatment of this bound constraint at $r = 0$. While the only actual measurement in this experiment occurs at $r = R$, for the purposes of the D.A. procedure we treated the known initial state (at $r = 0$) as a "measurement" as well. We then added a 1%

[1]In practice, we set the dependences as $1/(r + 0.1)^3$, to avoid infinities at $r = 0$, where 0.1 is in arbitrary units.

uncertainty to that "measurement." In this way, the initial known distribution was treated by the algorithm as a bound constraint of finite rigidity. To be reasonable, we also added 1% noise to the measurement of P_z at $r = R$.

Finally, for all eight experiments, a search was performed 20 times, each beginning from a randomly chosen location on the state space surface (for a description of the discretized space, see Appendix B, Sec. B 1).

The simulated data were generated by forward-integrating Eqs. (13) and (14) out to $r = R$, via a fourth-order adaptive Runge-Kutta scheme "odeINT," an open-access Python integrator. 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 integration step for the simulated data, and the discretization step, were each $\Delta r = 0.0001$. The resulting flavor evolution histories obtained in the interim ($r = 0$ to R) consist of 20,000 points. The annealing procedure took β from 0 to 30 in increments of 1, and $R_{f,0} = 0.01$.

The summations in Eq. (1), then, are constituted as follows. For the measurement error, the sum on location j has two terms, $r = 0$ and R, and the sum on the measured state variables l has two terms, $P_{1,z}$ and $P_{2,z}$. For the model error, the sum on location n has 20,000 terms, and the sum on all state variables a has six terms: $P_{1,x}$, $P_{1,y}$, $P_{1,z}$, $P_{2,x}$, $P_{2,y}$, and $P_{2,z}$. We performed the experiments for various values of R_m between 1 and 10,000.

A link to the Python codes that we used to interface with Ipopt is provided in Appendix B, Sec. B 2.

C. Rationale for experiments in light of astrophysical considerations

The specific experimental designs we adopt—namely, the two energy ratios E_{ν_1}/E_{ν_2} and various values of coupling strength Q—were crafted to be analogies to interesting physical scenarios. In Appendix C we illustrate these analogies by providing supernova envelope examples in which to "embed" the chosen neutrino energy ratios.

The examples with a neutrino energy ratio of 2.5 will correspond to situations where the location of our "detector," that is, our final location R, is well outside the supernova envelope, and well beyond any MSW resonances. With completely adiabatic flavor evolution this will correspond to the completely degenerate case, where the initial and final neutrino flavor states are essentially predetermined, and the flavor state history between these points is not uniquely determined. Consequently, we might expect the optimization algorithm to fail to converge consistently to a single C-value in the case where $Q = 0$. Nonadiabatic flavor evolution, for example, because of density ledges or shocks, would be expected to break this degeneracy.

In the examples with a neutrino energy ratio of 0.01, the final location R was held unchanged from the $E_{\nu_1}/E_{\nu_2} = 2.5$ case. Changing the energy ratio, however, results in changing the locations of the MSW resonances relative to $r = R$. In particular, in these examples, the resonance location of Neutrino 2 is shifted *beyond* the final location $r = R$, in the limit where $Q \ll C$. Physically, this corresponds to a scenario in which the final location R is *inside* the supernova envelope.

Our purpose behind this choice is twofold. First, we aimed to assess whether optimization techniques can capture with fidelity the neutrino flavor evolution, if information is specified about the flavor content at certain locations within the envelope. This information might be important for calculating nucleosynthesis or neutrino heating, since both of these issues hang mostly on the ν_e and $\bar{\nu}_e$ content of the local neutrino fluxes. For example, alpha-rich freeze-out in the postaccretion phase will lead to overproduction of nuclides such as ^{90}Zr, unless the electron fraction is larger than about 0.48 [100]. This process may place constraints on the electron neutrino/antineutrino energy spectra and fluxes in the postaccretion epoch. There also exist speculative theoretical models about the possible production of r-process nuclides in core-collapse supernova environments, for example, via neutrino-spallation-induced liberation of neutrons in the helium layer [101,102].

Our second motivation for choosing a value of R inside the supernova envelope was to facilitate a direct comparison between different energy ratios, in order to examine the sensitivity of the D.A. procedure for different final radii relative to the resonance locations. We include these results here because our aim throughout this paper is not to capture physical realism, but rather to examine the robustness of the D.A. machinery over various model regimes.

Finally, and before giving examples of specific supernova conditions (Appendix C), it will prove useful to note that the MSW resonance condition, $\delta m^2 \cos 2\theta = 2E_\nu V_{\rm eff}(r)$, where $V_{\rm eff}(r)$ is the flavor-diagonal potential from background matter and neutrinos at location r, lets us determine the resonance location for a given neutrino energy. For purely matter-driven ($Q = 0$) flavor evolution this means that the location of the resonance is

$$r_{\rm res} = \left(\frac{2E_\nu}{\delta m^2 \cos 2\theta}\right)^{1/3} C^{1/3}, \tag{18}$$

and consequently the ratio of the resonance locations for two different neutrino energies, E_{ν_1} and E_{ν_2}, is independent of C in this case:

$$\frac{r_{\rm res,1}}{r_{\rm res,2}} = \left(\frac{E_{\nu_1}}{E_{\nu_2}}\right)^{1/3}. \tag{19}$$

V. RESULTS

A. Key findings

Before presenting details of the results, we summarize key findings:

(i) For six out of the eight experiments (which were defined in Sec. IV), the measurements contain sufficient information to qualitatively capture overall flavor evolution through the MSW resonance, given the estimated parameter values. That is, the flavor evolution histories are consistent with the parameter estimates. These six experiments correspond to model regimes in which the equations of motion for the neutrinos are strongly coupled to each other, to the matter background, or both.

(ii) For these six experiments, the precise location of the MSW transition is estimated correctly in roughly 33% of trials.

(iii) For the other two (out of eight) experiments, the D.A. result failed to capture the model evolution. These two cases correspond to low coupling of the equations of motion of ν_2 to both matter and to ν_1.

(iv) There is insufficient information in the measurements to break the degeneracy in allowed sets of parameter values C and Q. This result is expected, as broad ranges of the parameters C and Q will leave flavor evolution completely adiabatic, thereby matching the neutrino flavor values imposed at the end points. The corresponding picture of the state space surface that emerges is not one riddled with clearly defined local minima of varying depth, but rather a wide, relatively flat basin.

(v) The sensitivity of model evolution to parameter values may depend on (1) energy ratio E_{ν_1}/E_{ν_2}, (2) the value of neutrino-neutrino coupling strength Q, (3) the location of the detector relative to the resonance location, or (4) any combination of the above. For example, when the detector sits outside the region of flavor transformation, the model is insensitive to the values of C and Q. When the measurement is made prior (in location) to complete flavor transformation, however, a correlation emerges between estimates of C and Q. This finding suggests that the addition of physical constraints *within* the supernova envelope (in a more complicated model) could prove useful for degeneracy breaking.

(vi) The high-frequency, low-amplitude oscillations that modulate the overall transformation are fit poorly, in comparison to the fit of the overall transformation. This is due to the high sensitivity of this aspect of the model to precise estimates of C and Q.

B. Predicted flavor evolution and parameter estimates

1. Predicted flavor evolution histories

Examples of flavor evolution for the eight experiments are depicted in Figs. 1 and 2, for neutrino energy ratios $E_{\nu_1}/E_{\nu_2} = 2.5$ and 0.01, respectively. For each experiment, polarization vector components $P_x(r)$, $P_y(r)$, and $P_z(r)$ are shown at left for ν_1 and at right for ν_2. The rows correspond to results for Q parameter values of 0, 1, 100, and 1000, respectively. Predicted flavor evolution by the D.A. procedure is shown in red, alongside evolution curves in blue corresponding to the correct evolution obtained by forward integration, for comparison. The best results, over all trials, were obtained by a choice of measurement weight R_m of 1 and annealing parameter β between 13 and 15.

For six out of the eight experiments, we found the model dynamics to be captured well. For two out of the eight experiments, flavor evolution was traced poorly over all trials. We will first discuss the six relatively successful experiments, and then separately the final two.

The six relative successes were all four experiments for the $E_{\nu_1}/E_{\nu_2} = 2.5$ case (for all four values of Q), and the two experiments for the $E_{\nu_1}/E_{\nu_2} = 0.01$ high coupling ($Q = 100$ and 1000). For all six of these experiments, the corresponding plots on Figs. 1 and 2 represent roughly one third of results over all trials. For the other roughly 66% of trials, the overall transformation history had the same qualitative appearance, but the precise location r_{res} of resonant flavor conversion was matched less precisely to the model evolution. By "less precisely," we mean roughly a discrepancy between model and prediction captured in the top left panel in Fig. 2: P_x, P_y, and P_z for ν_1 in the case for $Q = 0$.

As we will describe in Sec. VI, all of these results capture the expected behavior of resonant transformation modified by neutrino self-coupling; the offsets in r_{res} over the trials are due to different (degenerate) sets of parameter estimates.

Next we examine the two experiments for the $E_{\nu_1}/E_{\nu_2} = 0.01$ case, for low coupling: $Q = 0$ and 1. Unlike the results for the other six experiments, all of these D.A. solutions fail to match the features of the true evolution of ν_2, including the measurements at the end points. This result may be interpreted in terms of the efficiency of information flow among the state variables. As noted, if one has available as measurements only those corresponding to a subset of the model's total number of state variables, then in order to obtain information regarding the *unmeasured* states, the equations between measured and unmeasured states must be coupled to some significant degree. This interpretation is borne out by observations that generally, across physical models in other fields, D.A. tends to perform poorly when the equations of motion are not strongly coupled [103], as is the case here: ν_2 is not strongly coupled to ν_1, and it is also far from its resonance location.

2. Parameter estimates

The parameter estimates corresponding to the flavor evolution histories of Figs. 1 and 2 are listed in Table I. We found that for each experiment, the estimates varied across trials, with values of C and Q that spanned the permitted search ranges for each parameter (not shown); the search ranges are specified in the caption of Table I. Note that the

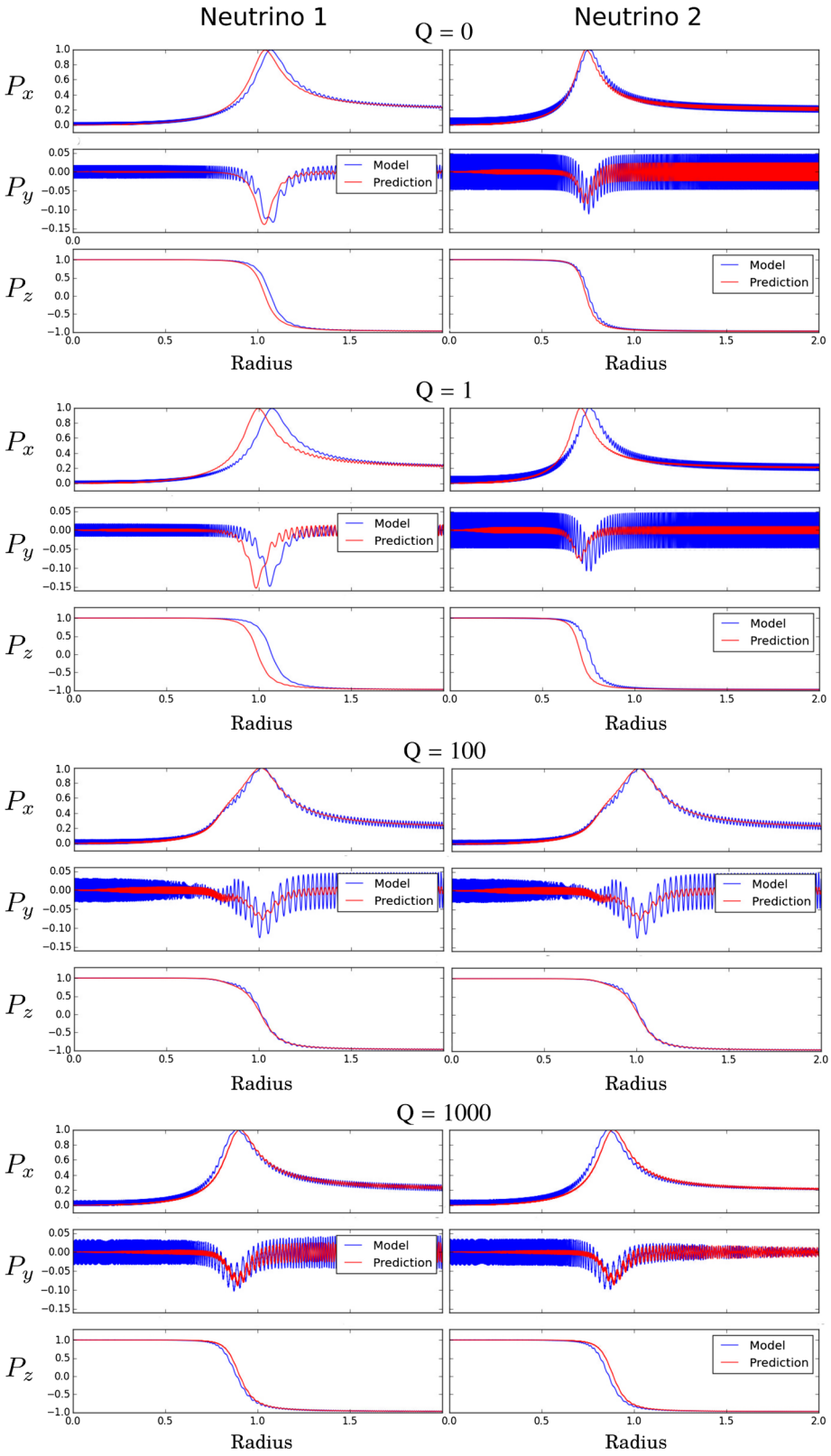

FIG. 1. Flavor evolution histories for the four experiments with neutrino energy ratio $E_{\nu_1}/E_{\nu_2} = 2.5$. For each experiment, polarization vector components $P_x(r)$, $P_y(r)$, and $P_z(r)$ are shown at left for ν_1 and at right for ν_2. Red curves represent neutrino flavor evolution histories predicted by Ipopt, given P_z measurements at the end points. The blue curves shown for comparison are the model solutions obtained by forward integration of the equations of motion. The rows correspond to results for Q parameter values of 0, 1, 100, and 1000, respectively. The true (model) value of C was 1304.5 for all four experiments.

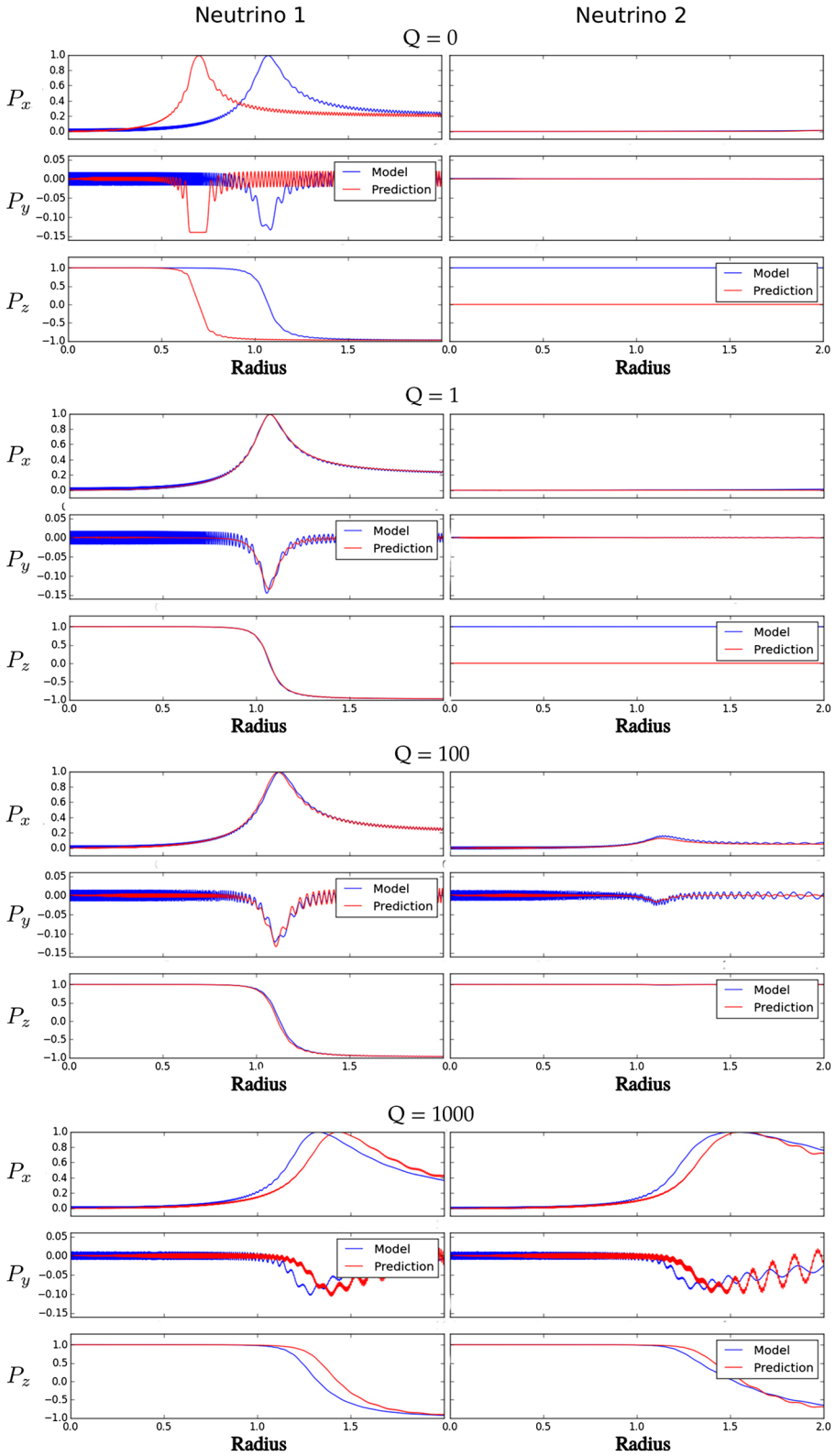

FIG. 2. Flavor evolution histories for the four experiments with neutrino energy ratio $E_{\nu_1}/E_{\nu_2} = 0.01$. For each experiment, polarization vector components $P_x(r)$, $P_y(r)$, and $P_z(r)$ are shown at left for ν_1 and at right for ν_2. Red curves represent neutrino flavor evolution histories predicted by Ipopt, given P_z measurements at the end points. The blue curves shown for comparison are the model solutions obtained by forward integration of the equations of motion. The rows correspond to results for Q parameter values of 0, 1, 100, and 1000, respectively. The true (model) value of C was 1304.5 for all four experiments. Note the poor match to the flavor evolution of ν_2 for the low-coupling cases ($Q = 0$ and 1). See text for comments.

TABLE I. Parameter estimates C and Q corresponding to the solutions whose state variable evolution is displayed in Figs. 1 and 2, for $E_{\nu_1}/E_{\nu_2} = 2.5$ (center columns) and 0.01 (right). The permitted search range for C is 500 to 1900, for all experiments. The permitted search range for $Q = 0$ is −10 to 10; for $Q = 1$, 0 to 10; for $Q = 100$, 0 to 200; for $Q = 1000$, 500 to 1900.

		$E_{\nu_1}/E_{\nu_2} = 2.5$		$E_{\nu_1}/E_{\nu_2} = 0.01$	
C (model)	Q (model)	C (DA)	Q (DA)	C (DA)	Q (DA)
1304.5	0	1457	0.3	563	10[a]
1304.5	1	1292	0.7	1586	9
1304.5	100	1621	169	1560	101
1304.5	1000	1681	1840	1641	1272

[a]Value is a permitted search range bound.

search ranges were different for each value of Q chosen; see Sec. VI. In addition, for comments on the lack of errors on these parameter estimates and on methods to quantify these errors, see Sec. VI.

To examine possible model sensitivity to parameter estimates in different model regimes, we explored $C - Q$ space for all trials over all eight experiments. For the four experiments with $E_{\nu_1}/E_{\nu_2} = 2.5$, no statistical trend emerged. Note that for this case, the detector sits at a location beyond the complete flavor conversion of both neutrinos. For $E_{\nu_1}/E_{\nu_2} = 0.01$, however, a trend emerged at $Q = 100$ and 1000. Figure 3 illustrates this trend. As Q is increased from 100 to 1000, it clearly emerges that a correlation between the C and Q values is required by the estimates. Note that for this case, the detector sits at a location at which the resonant flavor conversion of ν_2 is not complete. See Sec. VI for comments.

Finally, to test whether the low noise added to the measurements was in part responsible for the high degeneracy of these estimates, we repeated all experiments with zero noise in the measurements. Results were essentially unchanged.

C. Tests to ascertain success of prediction, given the parameter estimates

The degeneracy of parameter estimates described above gave rise to an important question: Could the deviation of the state prediction from precise MSW resonance be attributed to the particular estimates of C and Q? In other words, we sought to ascertain whether—over all trials for the six out of eight rather successful experiments—the evolution history was traced correctly given the respective parameter estimates.

To this end, we examined the value of A_0 over all values of the annealing parameter β. The top panel of Fig. 4 shows the value of the action A_0 over the annealing procedure corresponding to the first solution depicted in Fig. 1, where $E_{\nu_1}/E_{\nu_2} = 2.5$ and coupling strength $Q = 0$. This is a purely matter-driven neutrino evolution case. The y-axis is

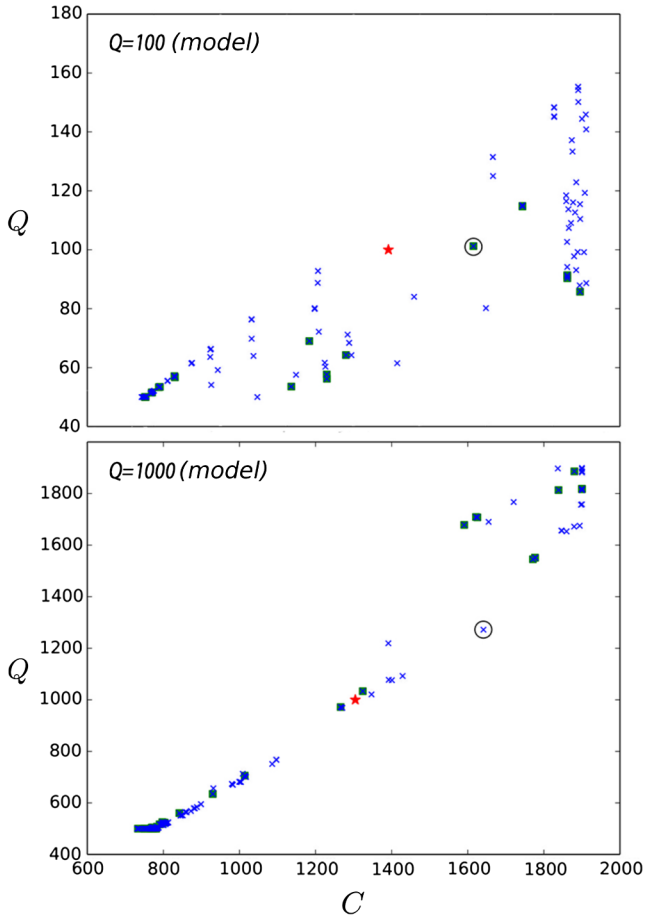

FIG. 3. Relation between Ipopt-predicted values of parameters C (matter coupling strength) and Q ($\nu - \nu$ coupling strength), for $E_{\nu_1}/E_{\nu_2} = 0.01$, for model $Q = 100$ (top) and model $Q = 1000$ (bottom). Symbol representations are as follows. Blue x's: estimates over all trials for all values of annealing parameter β; solid green squares: estimates over all trials for the values of β that consistently yielded the best fits to flavor evolution, $\beta = 13$ through 15; open black circle: the estimate corresponding to the plot of flavor evolution on Fig. 2; solid red star: the true model value. A weak positive correlation between C and Q appears for $Q = 100$, and gets stronger at $Q = 1000$. The same plot (of $C - Q$ space) for the experiments in which $E_{\nu_1}/E_{\nu_2} = 2.5$ revealed no significant correlation at any value of Q (not shown). See Sec. VI for comments.

the base-ten logarithm of A_0, and the x-axis is the parameter β defined by $R_f = R_{f,0}2^{\beta}$; $R_{f,0} = 0.01$ and $R_m = 1$.

This top panel of Fig. 4 is representative of the A_0-versus-β plots for all of the most-precisely-fit flavor evolution solutions presented in Figs. 1 and 2, that is, for roughly 33% of trials in six out of the eight experiments. Specifically, the logarithm of A_0 holds at a constant value of −4.0 (up to machine precision) for all values of β.

The bottom panel of Fig. 4 shows an overlaid plot of A_0-versus-β for all 20 trials corresponding to the first experiment (again, for $E_{\nu_1}/E_{\nu_2} = 2.5$ and $Q = 0$). An examination of each of the 20 results individually revealed

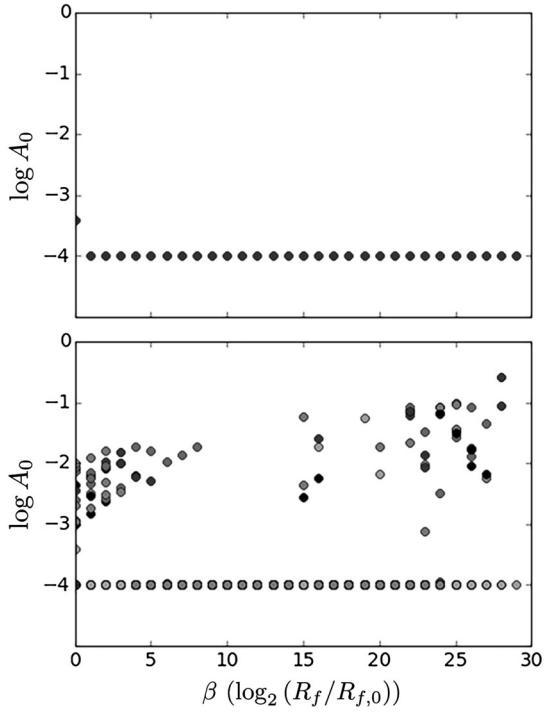

FIG. 4. The action A_0 over the annealing procedure. The y-axis is $\log_{10} A_0$, and the x-axis is the parameter β defined by $R_f = R_{f,0} 2^{\beta}$; $R_{f,0} = 0.01$ and $R_m = 1$. Top: A_0-versus-β for the first solution depicted in Fig. 1, where $E_{\nu_1}/E_{\nu_2} = 2.5$ and coupling strength $Q = 0$. Specifically, $\log_{10} A_0$ holds constant at -4.0 to machine precision. This panel is representative of the roughly 33% of solutions that fit flavor evolution well. Bottom: An overlaid plot of A_0-versus-β for all 20 solutions corresponding to the first experiment (for $E_{\nu_1}/E_{\nu_2} = 2.5$ and $Q = 0$). The degree of scatter at the extremes ($\beta \sim 0$ and 30) for each trial is individually correlated with the degree to which the corresponding flavor history solution matched the location of transformation.

that the value $\log_{10} A_0 = -4.0$ essentially held constant over the annealing procedure for each trial, but with varying degrees of scatter. The degree of scatter at the extremes ($R_m \gg R_f$ and $R_f \gg R_m$) on each plot was correlated with the degree to which the corresponding flavor history solution matched the precise location of transformation.

As there occurs no evolution of action with increasing R_f, we conclude that the dominant contribution to the action is the measurement error, and the model dynamics is obeyed well. Further, given that A_0 has the same minimum value for all parameter sets, we infer that all parameter sets—and corresponding evolution histories—capture the given measurements equally well.

Finally, we note two systematic features that we have identified for obtaining best-fit solutions to the flavor histories. Namely, all best fits, as exemplified in Figs. 1 and 2, correspond to values of β between 13 and 15 (out of the full span of zero to 30), and $R_m = 1$. The optimal values of these user-defined parameters vary across dynamical models, and for any given model it is important to identify a systematically reliable choice. Currently, this can only be done via trial and error; see Appendix B, Sec. 2 b.

VI. DISCUSSION

There are many issues in nonlinear compact object neutrino flavor evolution that are difficult to treat with standard initial-value-problem integrations of the neutrino flavor quantum kinetic equations as described in Sec. III. This concern constitutes our reason for considering numerical path integral-inspired approaches to this problem. To this end, here we have investigated the efficacy with which a particular D.A. algorithm (1) captures neutrino flavor evolution, and (2) identifies the sensitivity of that evolution to parameter values. It is significant that using one particular algorithm (Ipopt), we were able capture key features, like MSW resonance locations, of a simple supernova envelope neutrino flavor oscillation model. Furthermore, our results do indeed capture the degeneracy inherent in the highly adiabatic versions of our model, and they reveal that certain choices, for example, neutrino-neutrino coupling strength Q, neutrino energy ratio, and detector location, affect model sensitivity to the unknown parameters to be estimated. These findings are encouraging. In addition, and as discussed in the previous section, there exist particular realizations of parameter choices in our simple model that yield a MSW resonance prediction that is offset from the true location. It will be important to explore methods of breaking the parameter degeneracies that caused the offsets.

In this section we will (1) examine the significance of the results, (2) consider immediate enhancements to the model and to the D.A. experiment formulation, and (3) describe a tentative plan for ultimately investigating D.A. as an alternative method to solving the backscattering (halo) problem.

A. The physics captured by the results

Even in this simplistic two-neutrino model an array of flavor effects is evident. Perhaps the most prominent of these is the MSW conversion that occurs for one or both of the neutrinos in the scenarios depicted in Figs. 1 and 2. As the value of Q increases, the nonlinear coupling between ν_1 and ν_2 grows stronger and leads to modifications both in the locations of the resonances and in the flavor-space trajectories traversed by the neutrinos as they pass through resonance. The influence of nonlinearity is especially conspicuous in the scenario with $Q = 1000$ and an energy ratio of $E_{\nu_1}/E_{\nu_2} = 0.01$, wherein the resonance of ν_1, which would have occurred at $r \sim 1$ in the absence of neutrino-neutrino coupling, has synchronized with that of ν_2, which would have otherwise occurred at $r > 2$. The predicted solutions are able to capture the changing nature of resonant conversion across the entire range of Q values that we have considered.

In the most successful cases the solutions also exhibit more subtle features of flavor transformation. The small excursion

in P_y that is common to many of the results reflects the slight deviation from perfect adiabaticity: the polarization vector swings away from the xz-plane as it struggles to keep up with the Hamiltonian vector. Superimposed on top of these excursions are small-amplitude high-frequency oscillations, which correspond to the fast precession of $\boldsymbol{P}$ about $\boldsymbol{H}$, and although there are quantitative discrepancies between the model and prediction curves, the correct qualitative behavior is visible. Taken as a whole, Figs. 1 and 2 suggest that even with just two measurements per neutrino the data assimilation procedure is nevertheless sensitive to the fine features of flavor transformation.

B. Model sensitivity to parameter values

We now comment on our finding that the model's sensitivity to parameter values depends on detector location. As noted, we found that for experiments where the detector sat outside the range of resonant flavor transformation of both neutrinos, there existed high degeneracy in permitted values of parameters C and Q. This degeneracy is a result of the flavor conversion of neutrinos having completed prior to detection, thus rendering a measurement of P_z at that location relatively insensitive to both C and Q independently. In the cases where the measurement occurred prior (in location) to the complete transformation of ν_2, however, there emerged a correlation between the estimated values of C and Q. This result occurred for the energy ratio $E_{\nu_1}/E_{\nu_2} = 0.01$. The correlation between C and Q is strongly evident for $Q = 1000$ and more weakly for $Q = 100$.

Our interpretation of these findings is as follows. For the $Q = 1000$ case (Fig. 2, fourth row, right panel), ν_2 is still evolving through resonance at the location of measurement.[2] In this regime, the value of $P_{2,z}$ is highly sensitive to the value of Q at a particular location. Because both $V(r)$ and $\mu(r)$ are taken to scale as $1/r^3$ [see Eqs. (16) and (17)], the estimated value of Q essentially sets a corresponding value of C. For the case of $Q = 100$, the observed correlation between C and Q is weaker because (1) while some flavor oscillations of ν_2 are apparent, ν_2 is not yet evolving through resonance (Fig. 2, third row, right panel) and (2) since $Q \ll C$, the dependence of resonance location on Q is minimal. This discovery of detector-location-dependent model sensitivity to parameter values (Sec. IV C) suggests that it will be useful to add to a more realistic model additional known physics *within* the supernova envelope—as constraints, rather than as strict measurements (see immediately below).

[2]Note that in this regime the neutrino-neutrino potential is *negative*, since ν_1 has already gone through resonance and therefore has $P_{1,z} = -1$. This leads to the resonance of ν_2 being pulled farther in as compared to the case with no ν–ν coupling, an effect that was pointed out in Refs. [44,45] and is akin to the "matter-neutrino resonance" (MNR) [25,104–108].

C. Next steps

Ideally, we would next like to explore a model of similar simplicity, but with the addition of a backscattering term. Given, however, the degeneracy of parameter estimates in this simple model—and hence an unreliable predicted precision of the MSW transition, it is clear that we first must improve the current D.A. procedure. Improvements may involve amending the model, the D.A. formulation, or both. Here we outline a tentative stepwise plan for the ultimate employment of D.A. upon a realistically sized flavor evolution model that includes terms that are problematic for traditional numerical integration.

1. Tailor the D.A. procedure to the point at which we are consistently obtaining precise estimates of parameters and MSW transition location for a statistically significant fraction of trials—for example, over 99%. This step may require any of the following amendments, to be discussed in detail below: (1) addition to the model of more complicated potentials, (2) scaling of the model in terms of neutrino number, (3) imposing constraints on values of parameters and state variables within particular ranges of locations at particular epochs, and (4) modifying the particular D.A. formulation. Note that, in this stage, it is critical that the model continue to be solvable via numerical integration, so that we continue to have a consistency check for D.A. solutions. In addition, it is critical in this stage to add only the minimal amount of complexity to the model that is required to break degeneracy. A more realistic model—in which, for example, measurements correspond to energy spectra—would be vastly too large an initial step to take, given the difficulties already inherent with backscattering and with honing an appropriate D.A. methodology.
2. Once the procedure is consistently yielding precise estimates at, say, 99% confidence or higher, we add to the model a term that presents formidable difficulties to traditional numerical integration, namely, backscattering. No longer able to solve the model via standard forward integration, we obtain "blind" D.A. solutions. We are now permitting the D.A. to lead us in an understanding of the new physics that emerges.
3. Ultimately, having identified a D.A. protocol that works reliably for a toy model, we scale the model in neutrino number, consider multiangle calculations, and take the measurements to correspond to energy spectra.

D. Immediate next step: Seek methods for degeneracy breaking

1. Amending the model

Degeneracies in outcomes for different neutrino flavor evolution histories, for example, stemming from adiabatic

evolution through MSW resonances, are among the sought-after targets of a numerical calculation. In many ways, however, they offer a computational challenge. If we think physically about what could cause nonadiabatic neutrino flavor evolution in a supernova envelope we immediately think of physics that gives abrupt jumps or ledges in the neutrino forward-scattering potentials. Shocks or entropy jumps associated with supernova progenitor star fossil nuclear burning shells can produce these features in the matter potential. A goal might be to reveal such features by taking a detected Galactic supernova core-collapse neutrino burst signature, and then employing simulations to reverse engineer the neutrino flavor histories in the supernova envelope. Such jumps are tantamount to alterations in the neutrino forward-scattering potentials. Regarding our ongoing study of D.A. applied to this problem, then, one method to break parameter degeneracies may be to add to the model more complicated potentials that capture the physics just described. Another goal might be to ascertain whether evidence of nonlinear neutrino flavor evolution, for example, collective oscillations, appears in such a detected signature.

Another method of degeneracy breaking may be the addition of constraints within the supernova envelope, that is, between locations $r = 0$ and R. We have already stipulated that we know the neutrino fluxes, energy spectra, and flavor states at the neutrino sphere. What if, in addition, we stipulate that at a certain epoch (time postbounce) in the supernova, and in a range of locations (radius), the polarization vector components and energy spectra lie in given ranges? That is, we might want to examine how well we can fit detected neutrino data with assumed neutrino sphere flavor and energy distributions, while also demanding, for example, neutron-rich conditions in a certain region in the envelope. This neutron excess might arise from the $\nu_e + n \to p + e^-$, $\bar{\nu}_e + p \to n + e^+$ competition [42], or neutral-and-charged-current neutrino spallation of neutrons from nuclei. We can incorporate such considerations in the D.A. procedure as constraints on state variable values.

Finally, we may add neutrinos, as well as antineutrinos, to the model. As neutrinos in any given model are coupled all to all, the addition of neutrinos would tighten the coupling among the model ODE. In short, prior to adding back-scattering, we consider it important to examine embellishments to the current model for degeneracy breaking. First, we will add more complicated potentials (to represent shocks and changes in the electron fraction and density at particular locations r). A more complicated model will not only capture more faithfully the physics of interest, but it will furnish additional parameters to be estimated, thereby possibly providing more power to the measurements in distinguishing among solutions. Second, we will add more measurements, in terms of neutrino number, to examine the measurement dependence of degeneracy breaking. Third, we will examine the model dynamics at various epochs after the core bounce. Fourth, and separate from "measurements," we will add constraints to the values of variables within particular ranges of locations at various epochs.

2. Amending the D.A. procedure

In addition to model embellishments, various modifications to the D.A. procedure may improve parameter estimation.

First we note our choices for various user-defined search parametrizations: the relative weighting terms R_m and R_f, the coefficient k of the equality constraints, and the choice of discretized step size in radius r. There exists no universal optimal choice for any of these values. The optimal choices are model dependent and require extensive trial and error to identify. In this paper we experimented with various choices and took the set that yielded the most reliable results, but our experimentation was not exhaustive. As one example, the relative weighting terms R_f and R_m were taken to be equal at all points on any given path. We might examine whether loosening this constraint changes the results. (For comments on choosing an optimal ratio of R_m/R_f, see Appendix B, Sec. 2 b.) Finally, a value of $R_m = 1.0$ consistently produced the best results, and this was the value used for all the results presented in this paper.

Second, the user-defined search ranges for parameter values C and Q differed over all eight experiments, depending on the true model values of these quantities. Obviously, such bias has no place in a true D.A. experiment, that is, in an experiment where the measurements have been generated by a real physical process whose associated model parameter values are, of course, not known. A prerequisite to taking on real (experimental) measurements will be the elimination of this bias in simulated experiments.

More broadly, there exist many formulations for incorporating measurements and model evolution into an optimization framework. To discuss the myriad of existing techniques is beyond the scope of this paper. Here we note just two as examples. The first is variational synchronization (or "nudging"), where the model is imposed as a strong constraint upon the cost function rather than as a term within it. This formulation has demonstrated better performance for some problems in neuroscience [109].

As a second example, we cite Monte Carlo (MC) algorithms. This is a class entirely separate from variational approaches. MC methods have three notable advantages over the variational method. First, they are more efficient at searching a wider area, thereby offering a more global view of the action surface. Second, the map yields not only the depth of a minimum but also its width—thereby offering a quantification of errors on parameter estimates.[3] Third, MC

[3]On a related note, work is ongoing on a MC method to quantify errors on estimates obtained via the variational approach described in this paper [110].

methods are more readily parallelizable. This feature may be valuable when considering scalability.

E. Scalability

Ultimately, we will aim to build a D.A. formulation for a large-scale model. Because of the enormity of realistic astrophysical scenarios in terms of model degrees of freedom, the practicality of scalability is an important consideration. In numerical weather prediction, large models are being handled by GPUs. The model currently used at the European Centre for Medium Range Weather Forecasts, for example, contains 10^9 degrees of freedom, 10^7 of which are measured [111]. This size is comparable to the size of a neutrino flavor evolution model we would ultimately aim to build. See Ref. [112] for information on their computing facilities. Current state-of-the-art numerical radiation hydrodynamics simulations of core-collapse supernovae that incorporate detailed equation of state physics and include Boltzmann neutrino transport might utilize 10^{10} degrees of freedom. Simulations of this size may be augmented to treat the halo problem, but full quantum kinetic treatments of neutrino flavor and spin evolution may remain elusive, even at the exascale [113]. As noted, MC frameworks for D.A. may be better suited for large-scale simulations, as they are readily parallelizable. See Ref. [114] for a formulation of the MC algorithm using a cost function similar to that used in this paper; see Ref. [115] for an exploration of GPU processing capabilities for such MC structures. Finally, we note that other kinds of statistical approaches to gas dynamics and magnetohydrodynamics are being pursued, for example, Gaussian process modeling [116]. Whether any of these might be employed to tackle neutrino flavor evolution is an open question.

VII. SUMMARY

Our exploration of optimization-based data assimilation techniques for treating the neutrino flavor transformation problem in supernovae has yielded insights into the utility of these approaches. Advantages include (1) the ability of D.A. to search efficiently over ranges of input model parameters, and (2) the integration-blind feature offered by an optimization framework. Obviously, we have only scratched the surface of this problem. We plan immediate modifications to both the model and D.A. procedure, to obtain more precise and reliable results in advance of considering the direction-changing scattering problem in a toy model. Following that achievement, we envision scaling up the sophistication of the model, by adding more neutrinos, with many more neutrino energy bins, a more sophisticated geometry, realistic nonsmooth matter density profiles, and direction-changing neutrino scattering. We also envision treating other supernova epochs, for example, shock break-out and the associated neutronization burst.

ACKNOWLEDGMENTS

We thank Nirag Kadakia, Paul Rozdeba, Sasha Shirman, and Daniel Breen for discussions on data assimilation and experimental design. We also appreciate Jung-Tsung Li, James Tian, Sebastien Tawa, John Cherry, Brad Keister, Baha Balentekin, and Shashank Shalgar for discussions on the relevance of D.A. to neutrino astrophysics. Thanks also to Evan Grohs and an anonymous reviewer, for valuable feedback that has significantly improved the paper. This work was supported in part by NSF Grants No. PHY-1307372 and No. PHY-1614864. Finally, thanks to the village of Doylestown, Ohio—for everything.

APPENDIX A: A PATH-INTEGRAL-BASED FORMULATION OF STATISTICAL DATA ASSIMILATION

1. Summary of purpose and strategy

Here we lay out a derivation of the cost function A_0 used in this paper. We begin by seeking the probability of obtaining a path $\boldsymbol{X}$ in the model's state space given observations $\boldsymbol{Y}$, or $P(\boldsymbol{X}|\boldsymbol{Y})$. If we write

$$P(\boldsymbol{X}|\boldsymbol{Y}) = e^{-A_0(\boldsymbol{X},\boldsymbol{Y})},$$

the equation above will then mean "the path $\boldsymbol{X}$ for which the probability (given $\boldsymbol{Y}$) is greatest is the path that minimizes A_0." Now, if A_0 is sufficiently large (where "sufficiently" must be defined by the results of a particular D.A. experiment using a particular model), we can use Laplace's method to estimate the minimizing path on the surface of A_0.[4]

A formulation for A_0 will permit us to obtain the expectation value of any function $G(\boldsymbol{X})$ on a path $\boldsymbol{X}$; expectation values are the quantities of interest when the problem is statistical in nature. We can write the expectation value of $G(\boldsymbol{X})$ as

$$\langle G(\boldsymbol{X})\rangle = \frac{\int d\boldsymbol{X} G(\boldsymbol{X}) e^{-A_0(\boldsymbol{X},\boldsymbol{Y})}}{\int d\boldsymbol{X} e^{-A_0(\boldsymbol{X},\boldsymbol{Y})}}.$$

That is, the expectation value can be expressed as a weighted sum over all possible paths, where the weights are exponentially sensitive to A_0. The RMS variation, and higher moments of $G(\boldsymbol{X})$, can be calculated by taking the x_a to the appropriate higher exponents. If the quantity of interest is the path $\boldsymbol{X}$ itself, then we choose $G(\boldsymbol{X}) = \boldsymbol{X}$.

It remains, then, to write a functional form for A_0. This will take place in two steps. First we shall consider how

[4]Laplace's method was developed to approximate integrals of the form $\int e^{Mf(x)}dx$. For sufficiently high values of the coefficient M, significant contributions to the integral will come only from points in a neighborhood around the minimum, which can then be estimated.

measurements and model dynamics enter into the state and parameter estimation. This we will do via an examination of Bayesian probability theory and Markov chain transition probabilities, for the effect of measurements and model dynamics, respectively. Second, we shall make four simplifying assumptions: (1) the measurements taken at different times are independent, (2) both measurement and model errors have Gaussian distributions, (3) each measurement is taken to correspond directly to one model state variable, and (4) the minimizing path is independent of the guess—in state and parameter space—of the initial path.

In what follows, we shall describe this strategy in some detail (for a full treatment, see Ref. [1]). To remind the reader of the notation, the model consists of D ordinary differential equations, each of which represents the evolution of one of the model's D state variables. From the corresponding physical system, we are able to measure L quantities, each of which corresponds to one of the model's D state variables. Typically the measurements are sparse ($L \ll D$), and the sampling may be infrequent or irregular.

2. Considering model dynamics only (no measurements yet)

We shall first examine this formulation by considering the model's time evolution in the absence of measurements. We represent the model's path through state space as the set $\boldsymbol{X} = \{\boldsymbol{x}(t_0), \boldsymbol{x}(t_1), \ldots, \boldsymbol{x}(t_N), \boldsymbol{p}\}$, where t_N is the final "time point," the vector $\boldsymbol{x}(t)$ contains the values of the D total state variables, and $\boldsymbol{p}$ are the unknown parameters (here, the word "time" can also be taken to represent other grid parametrizations, for instance, location).

a. Assuming that a Markov process underlies the dynamics

If we assume that the dynamics are memoryless, or Markov, then $\boldsymbol{x}(t)$ is completely determined by $\boldsymbol{x}(t-\Delta t)$, where $t-\Delta t$ means "the time immediately preceding t" and an appropriate discretization of time Δt for our particular model has been chosen. A Markov process can be described in the continuous case by a differential equation, or as a set of differential equations:

$$\frac{\mathrm{d}x_a(t)}{\mathrm{d}t} = F_a(\boldsymbol{x}(t), \boldsymbol{p}); \qquad a = 1, 2, \ldots, D,$$

and we note that the model is an explicit function of the state variables $\boldsymbol{x}(t)$ and the unknown parameters $\boldsymbol{p}$. It is in this way that the unknown parameters are considered to be on equal footing with the variables; namely, they are variables with trivial dynamics.

In discrete time, that relation can be written in various forms. For our purposes, we use the trapezoidal rule,

$$x_a(n+1) = x_a(n) + \frac{\Delta t}{2}[F_a(\boldsymbol{x}(n+1)) + F_a(\boldsymbol{x}(n))],$$

where for simplicity we have taken n and $n+1$ to represent the values of t_n and t_{n+1}.

b. Permitting stochasticity in the model and recasting its evolution in terms of probabilities

We are interested in ascertaining the model evolution from time step to time step, where now we allow for some stochasticity in the model dynamics. In this scenario, the evolution can be formulated in terms of "transition probabilties," for example, $P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n))$, the probability of the system reaching a particular state at time $n+1$ given its state at time n. If the process were deterministic, then in our case $P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n))$ would simply reduce to $\delta^D(\boldsymbol{x}(n+1) - \boldsymbol{x}(n) - \frac{\Delta t}{2}[\boldsymbol{F}(\boldsymbol{x}(n+1)) + \boldsymbol{F}(\boldsymbol{x}(n))])$. We will revisit to this expression later in this appendix, under the section Approximating the Action.

For a Markov process, the transition probability from state $\boldsymbol{x}(n)$ to state $\boldsymbol{x}(n+1)$ represents the probability of reaching state $\boldsymbol{x}(n+1)$ given $\boldsymbol{x}(n)$ and $\boldsymbol{x}$ at *all* prior time steps, or

$$P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n)) = P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n), \boldsymbol{x}(n-1), \ldots, \boldsymbol{x}(0))$$

so that

$$\begin{aligned} P(\boldsymbol{X}) &\equiv P(\boldsymbol{x}(0), \boldsymbol{x}(1), \ldots, \boldsymbol{x}(N)) \\ &= \prod_{n=0}^{N-1} P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n)) P(\boldsymbol{x}(0)). \end{aligned}$$

We now write

$$P(\boldsymbol{X}) \equiv e^{-A_0(\boldsymbol{X})},$$

where A_0 is the action defined on the model's path $\boldsymbol{X}$ in state space (or the path that minimizes the action is the path most likely to occur).[5] Then the model term of the action, $A_{0,\mathrm{model}}$, can be written as

$$A_{0,\mathrm{model}} = -\sum \log[P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n))] - \log[P(\boldsymbol{x}(0))],$$

where the second term represents uncertainty in initial conditions.

3. Now with measurements

We now consider the effect of measurements. Let us define a complete set of measurements $\boldsymbol{Y}$ to be the set of all vectors $\boldsymbol{y}(n)$ at all times n—the analog of $\boldsymbol{X}$ for the complete set of state variable values. We shall examine

[5]The reader might find it of interest to note the quantum-mechanical analog of the transition probability, which involves the trivial addition of the term $\frac{i}{\hbar}$ in the exponent: $P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n)) = e^{\frac{i}{\hbar}A(t_{n+1}, t_n)}$, where A here is the classical action.

the effect of these measurements upon a model's dynamics by invoking the framework of "conditional mutual information" (CMI); for a useful definition of CMI, see Ref. [117].[6]

The expression $\mathrm{CMI}(\boldsymbol{x}(n),\boldsymbol{y}(n)|\boldsymbol{Y}(n-1))$ asks, "How much is learned about event $\boldsymbol{x}(n)$ upon observing event $\boldsymbol{y}(n)$, given that events $\boldsymbol{Y}(n-1)$ have been previously observed?" The CMI can be quantified as

$$\mathrm{CMI}(\boldsymbol{x}(n),\boldsymbol{y}(n)|\boldsymbol{Y}(n-1)) = \log\left[\frac{P(\boldsymbol{x}(n),\boldsymbol{y}(n)|\boldsymbol{Y}(n-1))}{P(\boldsymbol{x}(n)|\boldsymbol{Y}(n-1))P(\boldsymbol{y}(n)|\boldsymbol{Y}(n-1))}\right].$$

4. The complete action

With measurement considerations included, the action now becomes

$$A_0(\boldsymbol{X},\boldsymbol{Y}) = -\sum \log[P(\boldsymbol{x}(n+1)|\boldsymbol{x}(n))] - \log[P(\boldsymbol{x}(0))] - \sum \mathrm{CMI}(\boldsymbol{x}(n),\boldsymbol{y}(n)|\boldsymbol{Y}(n-1)),$$

where the first and second terms represent the model dynamics including initial conditions, and the third term represents the transfer of information from measurements. The summations are over time. As noted, this formulation positions us to calculate the expectation value of any function $G(\boldsymbol{X})$ on the path $\boldsymbol{X}$.

We now offer an interpretation of the measurement term. The measurement term can be considered to be a nudging (or synchronization) term. While nudging terms are often introduced rather artificially in the interest of model control, we have shown that the measurement term arises naturally through considering the effects of the information those measurements contain. For this reason, we prefer to regard the measurement term as a guiding potential. In the absence of the potential, we live in a state space restricted only by our model's degrees of freedom. The introduction of the measurements guides us to a solution within a *subspace* in which those particular measurements are possible.

5. Approximating the action

We now seek to simplify the action formulation for the purposes of calculation.

[6]The reader may find an intuitive understanding of our use of the CMI by the following consideration. The overall information, in bits, in a set A is defined as the Shannon entropy $H(A) = -\sum_A P(A)\log[P(A)]$. The CMI is a means to quantify the amount of information, in bits, that is transferred along a model trajectory within a particular temporal window. That information is equivalent to $-\sum_{n=0}^{N}\log[P(\boldsymbol{x}(n)|\boldsymbol{y}(n),\boldsymbol{Y}(n-1))]$.

a. The measurement term

Regarding the measurement term, we make four assumptions:

(i) The measurements taken at different times are independent of each other. This permits us to write the CMI simply as $P(\boldsymbol{x}(n)|\boldsymbol{y}(n))$, or

$$A_0(\boldsymbol{X},\boldsymbol{Y}) = -\log[P(\boldsymbol{X}|\boldsymbol{Y})].$$

(ii) There may be an additional relation between the measurements and the state variables to which those measurements correspond, which can be expressed with the use of some transfer function h_l: $h_l(\boldsymbol{x}(n)) = y_l(n)$.

(iii) For each of the L measured state variables, we allow for a noise term θ_l at each time point, for each measurement y_l that corresponds to a state variable x_l: $y_l(n) = h_l(\boldsymbol{x}(n)) + \theta_l(n)$. In this case, then, $P(\boldsymbol{x}(n)|\boldsymbol{y}(n))$ is simply some function of $h(\boldsymbol{x}(n)) - \boldsymbol{y}(n)$ at each time point.

(iv) The measurement noise has a Gaussian distribution.

Taking these assumptions, we arrive at

$$\mathrm{CMI}(\boldsymbol{x}(n),\boldsymbol{y}(n)|\boldsymbol{Y}(n-1)) = -\sum_{l,k=1}^{L}(h_l(\boldsymbol{x}(n)) - y_l(n)) \times \frac{[R_m(n)]_{lk}}{2}(h_k(\boldsymbol{x}(n)) - y_k(n)),$$

where R_m is the inverse covariance matrix of the measurements y_l.

b. The model term

We simplify the model term by assuming that the model may have errors, which will broaden the delta function in the expression noted earlier for the deterministic case. If we assume that the distribution of errors is Gaussian, then $\delta^D(z)$ becomes $\sqrt{\frac{\det R_f}{(2\pi)^D}}e^{[-z\frac{R_f}{2}z]}$, where R_f is the inverse covariance matrix for the model's state variables.

Taking both approximations together, assuming that the transfer function h_l is simply unity, and assuming that the minimizing path is independent of considerations of initial conditions, we obtain

$$A_0 = \sum_{n}^{N-1}\sum_{a}^{D}\frac{R_a^f}{2}\left(\frac{x_a(n+1)-x_a(n)}{t_{n+1}-t_n} - f_a(\boldsymbol{x}(n))\right)^2 + \sum_{j}\sum_{l}^{L}\frac{R_l^m}{2}(y_l(j)-x_l(j))^2,$$

where $f_a(\boldsymbol{x}(n)) \equiv \frac{1}{2}[F_a(\boldsymbol{x}(n)) + F_a(\boldsymbol{x}(n+1))]$. The first (model) term involves a summation over all D state variables, and the second (measurement) term involves a summation

over the L measured quantities. Note that here we write the model error term in a simpler, more general manner than the specific formulation used in this paper [Eq. (1)].

Finally, we allow in the cost function the addition of equality constraints, of the general form $kg(\boldsymbol{x}(n))$, where the coefficient k set the strength of the constraint function g. The specific equality constraints chosen in this paper are described in Sec. II.

APPENDIX B: DETAILS OF THE D.A. PROCEDURE

1. The discretized search space

The optimization procedure searches a $(D(N+1)+p)$-dimensional state space, where D is the number of state variables of a model, N is the number of discretized steps, and p is the number of unknown parameters. Note that each location point is considered a separate dimension. Thus the action, instead of being a functional of D functions, is a function of $(D(N+1)+p)$ variables.

Ipopt, the specific algorithm used in this paper, employs a Newton's, or descent-only, search. The spatial resolution is set by a user-defined step size. The user provides the objective function, model, the Jacobean and Hessian matrices of the model, permitted search ranges of variables and unknown parameters, and discretized step size. The algorithm iteratively searches for a path in the state space that minimizes the action subject to the requirements that the first derivative of the objective function at the minimizing path along any direction be zero and that its second derivative along any direction be positive definite. The resulting "path" is a set of state vectors, one at each discretized step, and specific values of the unknown parameters. Each path corresponds to a single point in the $(D(N+1)+p)$-dimensional space. In this way, the model parameters are considered on equal footing with the state variables; namely, the unknown parameters are state variables with trivial dynamics. Finally, to impose user-defined bounds placed upon the searches, Ipopt uses a barrier method. For details, see [89].

2. Specific choices governing D.A. experiments in this paper

a. Interface with Ipopt

Ipopt requires a user interface to discretize state space and calculate the model equations of motion, Jacobean, and Hessian matrices that are used in the minimization procedure. We used a suite of Python codes to generate this interface; it is available here: https://github.com/yejingxin/minAone.

b. Choosing R_f/R_m for best results

As noted in Sec. VI, there exists no universal rule for choosing an optimal ratio of model and measurement weights. An optimal value is model dependent and must be identified via trial and error. Generally, for many biophysical models of neurons, small neuronal networks, atmospheres, and chaotic Lorenz-63 and Lorenz-96 models, a value of β between 10 and 20 is found to be ideal [103]. The reader may compare this range to our identification of $\beta \in [13, 15]$, which we found yielded the best results.

Poor results at the extremes ($R_m \gg R_f$ and $R_f \gg R_m$) are expected for any model, for the following reasons. For low R_f, the model constraints are not yet sufficiently strict to require a converging solution. For high R_f, the failure of solutions has at least two potential causes. First, one encounters numerical problems with considering "infinite" model weight. The problem is ill conditioned when it involves a matrix whose elements are so large that the matrix is not invertible. The optimizing solution may thus become overly sensitive to changes in the state vector. Rounding errors may render these solutions invalid. A second possible cause is discretization error at high R_f. In taking a discretized derivative, one retains only the first term in a Taylor series. As the multiplicative factor grows, the higher-order terms—which are ignored—will become important.

APPENDIX C: EMBEDDING THE MODEL INTO A SIMPLIFIED ASTROPHYSICAL SYSTEM

1. Forms for matter and coupling potentials

The cubic radial dependence of the matter potential is actually close to the expected density run in the supernova envelope in some cases. For example, some seconds after a supernova explosion, perhaps 3 to 10s after core bounce, we can be left with a tenuous, near-hydrostatic envelope sitting in a gravitational potential well dominated by the hot, proto-neutron star. This envelope is being heated to high entropy by the intense neutrino radiation from the neutrino sphere, and driven off. This is the "neutrino-driven wind" epoch. It is a candidate site for r-process nucleosynthesis, but one fraught with challenges stemming from uncertain neutrino flavor transformation physics and the "alpha effect" : the interaction between charged current ν_e and $\bar{\nu}_e$ captures and aggressive alpha particle formation in the high entropy wind. In turn, the entropy of this wind is a complicated function of neutrino heating and flavor histories.

We can approximate the wind regime envelope as (1) being in hydrostatic equilibrium, with enthalpy per baryon equal to the local gravitational binding energy per baryon and (2) with the entropy of the material being carried entirely by relativistic particles, namely photons and electron-positron pairs. The latter assumption is tantamount to the entropy being high. We can combine (1) and (2) and find for a *constant* entropy envelope the baryon density dependence on radius r:

$$n_B(r) = \rho(r) N_{\rm A} = \left(\frac{2\pi^2}{45}\right) g \left[\frac{M_{\rm NS} m_{\rm p}}{m_{\rm pl}^2}\right]^3 \frac{1}{s^4}\frac{1}{r^3}, \qquad \text{(C1)}$$

where the baryon mass (energy) density is ρ; Avogadro's number is $N_{\rm A}$; and $M_{\rm NS}$, $m_{\rm p}$, and $m_{\rm pl}$ are, respectively, the neutron star mass, proton rest mass, and the Planck mass. Here s is the entropy per baryon in units of Boltzmann's constant $k_{\rm b}$, and g is the statistical weight in relativistic particles. In terms of this parametrization of the density run, our constant C in the expression for $V(r)$ is [118]

$$C = \sqrt{2} G_{\rm F} Y_e \left(\frac{11\pi^2}{45}\right)\left(\frac{g}{11/2}\right)\left[\frac{M_{\rm NS} m_{\rm p}}{m_{\rm pl}^2}\right]^3 \frac{1}{s^4}$$
$$\approx 2.9 \times 10^6 \ {\rm MeV\,cm^3} \left(\frac{g}{11/2}\right)\left[\frac{M_{\rm NS}}{1.4 M_\odot}\right]^3 \frac{Y_e}{s_{100}^4}, \qquad \text{(C2)}$$

where Y_e is the electron fraction, and the entropy per baryon in units of $100 k_{\rm b}$ is s_{100}.

In a spherical geometry, with neutrinos emitted from a *sharp* neutrino sphere, the radial dependence of the $\nu - \nu$ potential is $\mu(r) \sim 1/r^4$, as both the neutrino number flux $n_\nu(r)$ and the angle factor $\alpha(r)$ each dilute as $1/r^2$ in the far-field limit. In more complicated models, including those that incorporate backscattering, we expect a different radial dependence than $1/r^4$. Specifically, we expect the neutrino potential to drop less quickly with radius than in conventional bulb models. Here then, for simplicity and to enable a direct comparison to the matter potential, we choose $\mu(r) \sim 1/r^3$. That is, our motivation here was to introduce nonlinearity in a simple manner, while avoiding the use of different functional forms for $V(r)$ and $\mu(r)$.

2. Neutrino energy ratios set within an astrophysical context

For the energy ratio $E_{\nu_1}/E_{\nu_2} = 2.5$ we can give three plausible supernova envelope examples based on the constant-entropy, wind-like density profile given in Eqs. (C1) and (C2) and the atmospheric neutrino mass-squared splitting. If we take $s_{100} = 1$, $g = 11/2$, and $Y_e = 0.4$, all plausible conditions for a neutrino-driven wind that might form at > 3 s after core bounce, then the resonance locations for $E_{\nu_1} = 25$ MeV and $E_{\nu_2} = 10$ MeV are 289 km and 213 km, respectively, and the ratio in Eq. (19) is $\approx$1.4. Note that the corresponding resonance widths, $\delta r = |V/(dV/dr)|_{\rm res} \tan 2\theta \sim (r_{\rm res}/3)\sin 2\theta \sim 10$ km for $\theta = 0.1$, are small enough that the resonances are well separated for these neutrino energies. We can also consider the same neutrino energies, but now with a smaller entropy, $s_{100} = 0.1$, a slightly smaller electron fraction, $Y_e = 0.35$, and g-factor, $g = 2$. These choices will very crudely mock up an earlier accretion phase supernova envelope. In this case the resonance locations are 4254 km and 3135 km, respectively. If we consider the same envelope parameters but now take neutrino energies $E_{\nu_1} = 2.5$ MeV and $E_{\nu_2} = 1$ MeV, we obtain resonance locations at 1975 km and 1455 km, respectively. In all of these cases neutrino flavor evolution through these resonances will be adiabatic.

If we take the neutrino energy ratio of 0.01, with $E_{\nu_1} = 0.5$ MeV and $E_{\nu_2} = 50$ MeV, and the windlike higher entropy conditions described above, we obtain resonance locations at 79 km and 364 km, respectively, for the $Q = 0$ case. In this case, our experimental setup would put the final location R between these resonances, inside the supernova envelope. We study this scenario, for multiple values of coupling strength Q, in order to examine collective effects and explore the sensitivity of the D.A. procedure to flavor information deep in the supernova envelope.

APPENDIX D: EVOLUTION OF MASSIVE STARS, WEAK INTERACTIONS, AND NEUTRINO FLAVOR PHYSICS

The following is a pedagogical overview of neutrino physics in core-collapse supernovae [119–121].

1. Evolution of massive stars and the weak interaction

The weak interaction, the nuclear force responsible for changing neutrons to protons and vice versa, is the key to why stars shine, and why big stars collapse, explode, and synthesize the elements. The sun and stars like it burn hydrogen into helium, combining four protons into a helium nucleus, and thereby turning two of those protons into neutrons along the way. The fundamental weak reaction in the sun turns two protons into a deuterium nucleus with the emission of a positron and an accompanying electron-flavor neutrino, $p + p \to D + e^+ + \nu_e$.

Neutrinos experience only gravitation and the weak force, making them very "slippery," that is, able to escape from deep inside a dense star, and carry away energy. The weak interaction is aptly named, being some 20 orders of magnitude weaker than electromagnetic forces, at the relevant energy scales. Indeed, hydrogen burning in the sun is desperately slow. It will take 10^{10} years for the sun to burn through all of its hydrogen. In more massive stars, however, weak interactions, along with attendant neutrino emission, combined with gravitation, can nevertheless engineer their violent destruction.

Stars some ten or more times the mass of the sun ($M \geq 10 M_\odot$) evolve in millions of years through a series of nuclear burning epochs: hydrogen to helium, to carbon and oxygen, to magnesium, to silicon. Finally, silicon burns to "iron," forming a core with mass $\sim 1.4 M_\odot$ composed of relatively neutron-rich iron-peak nuclei (for example, ^{56}Fe, ^{48}Ca, etc.). From core carbon burning onward in these objects, the energy carried away by neutrinos *exceeds* that radiated by photons. Neutrinos carry away the heat

generated by nuclear reactions, forcing the star to contract and release more gravitational binding energy, accelerating nuclear burning, and so on. This jams the electrons in the star into a smaller and smaller volume, and the Pauli principle implies that they are consequently forced into higher and higher energy states—the electrons become relativistically degenerate. In turn, the energy dependent weak interactions, for example, electron capture on protons to make neutrons ($e^- + p \to n + \nu_e$) proceed faster. Though this iron core has a density more than ten orders of magnitude higher than that of water, it is essentially transparent to these neutrinos.

The end result is that neutrinos leave and refrigerate the core. Though the core has a temperature of nearly 10^{10} K ($\sim$1 MeV), it is desperately cold in a thermodynamic sense, highly ordered, with an entropy per baryon $\sim$1 unit of Boltzmann's constant, a factor of 10 or more lower than the entropy in the sun. This low entropy, or high order, sets up the core for instability. The pressure supporting the star against gravitation is coming mostly from the degenerate electrons, which are moving nearly at the speed of light. A consequence of the nonlinear nature of gravitation is that whenever the pressure support for a star comes from particles moving at the speed of light, that star is trembling on the verge of instability.

A variety of processes can shove the core over the edge, leading to dynamical collapse, with infall speeds in some cases approaching the free-fall rate. As the density rises, the electrons become even more energetic and electron capture proceeds even faster, making more neutrinos and "neutronizing" the collapsing core. When the density of the core reaches $\sim 10^{12}\ \mathrm{g\,cm^{-3}}$, roughly 1% of nuclear matter density, it becomes opaque to neutrinos. The neutrinos are trapped and quickly come into thermal and chemical equilibrium with the matter. As the collapse proceeds, the outer portions of the core are falling in supersonically. When the inner part of the core reaches nuclear density, the nucleons touch, and this region stops abruptly. The outer, supersonic part of the core slams into this "brick wall," generating a shock wave that propagates outward through the outer core.

In broad brush terms, the $\sim 1.4 M_\odot$ core collapses from a configuration with a radius like that of the Earth ($\sim 10^9$ cm) to one with a radius of roughly 45 km in about one second. Within another second or two it quasistatically shrinks down to a radius of 10 km. The upshot is a prodigious gravitational binding energy change, amounting to about 10% of the entire rest mass of the core. One percent of this energy largess resides in the bulk infall kinetic energy of the core (and consequently the initial energy in the shock wave), and the other 99%, some 10^{53} erg, is in the trapped seas of neutrinos of all kinds.

At the edge of the proto-neutron star, deemed the "neutrino sphere," the matter density and opacity to neutrinos drop off dramatically and neutrinos can more or less freely stream away, mostly unhindered by direction-changing or inelastic collisions with particles that carry weak charge, for example, neutrons, protons, electrons, and other neutrinos. The average energies of the neutrinos streaming out are of order $\sim$10 MeV. With a gravitational binding energy of 10^{53} erg ($\sim 10^{59}$ MeV), this amounts to some 10^{58} neutrinos carrying this energy away in a matter of a few seconds. These are titanic neutrino fluxes.

The shock wave that propagates through the supernova envelope is associated with an entropy jump across the shock front—the material that the shock plows into has an entropy per baryon $\sim 1 k_B$, whereas the material behind the shock has an entropy per baryon of $\sim 10 k_B$. As a result, the passing of the shock wave through the envelope results in the dissociation of nuclei into mostly free nucleons, a process that costs $\sim$8 MeV of energy per nucleon. This causes the shock wave to rapidly lose energy and stall at a radius of order a few hundred kilometers. Subsequently, within a second or so, charged-current captures of electron flavor neutrinos on neutrons and protons ($\nu_e + n \to p + e^-$ and $\bar{\nu}_e + p \to n + e^+$) may deposit enough energy in the matter behind the shock to reenergize it and get it moving again with an energy of 10^{51} erg, resulting in a supernova explosion. This process can be aided by hydrodynamic motion of the neutrino-heated material. In the end, about 1% of the total neutrino energy needs to be deposited in this material to get an explosion.

2. Collective neutrino flavor transformations in supernovae

It is known that neutrinos come in three "flavors," ν_e, ν_μ, and ν_τ, corresponding to each of the three charged leptons. These flavors denote weak-interaction eigenstates, essentially determining how these particles interact in matter. Each neutrino has an antiparticle, implying that there are six kinds of neutrinos: $\nu_e, \bar{\nu}_e, \nu_\mu, \bar{\nu}_\mu, \nu_\tau$, and $\bar{\nu}_\tau$. These particles are spin-1/2, electrically neutral, and have very small rest masses. We do not know what the masses are, but the differences of the squares of these masses are measured: the so-called solar mass-squared splitting $\delta m^2_\odot = m_2^2 - m_1^2 \approx 7.9 \times 10^{-5}\ \mathrm{eV}^2$, and the atmospheric mass-squared splitting $\delta m^2_{\mathrm{atm}} = m_3^2 - m_1^2 \approx 2.4 \times 10^{-3}\ \mathrm{eV}^2$, where m_1, m_2, and m_3 are the neutrino mass eigenvalues corresponding to the energy eigenstates (sometimes called "mass" states) of the neutrinos. Experiment shows that these neutrino mass states are not coincident with the flavor states and this can have consequences for the core-collapse supernova mechanism and for neutrino detection.

The fact that neutrino mass states are not coincident with flavor states means that neutrinos emitted initially in one flavor state can transform into another as they propagate, with consequences for the way these particles effect heating, nucleosynthesis, etc. Flavor transformations are modified in the presence of potentials arising from neutrino

forward scattering on particles that carry weak charge, such as leptons, nucleons, and other neutrinos. As the neutrinos stream away through the lower density material above the neutrino sphere, they acquire through forward scattering an "index of refraction," equivalent to an effective mass in medium. This is analogous to the way photons acquire an index of refraction and effective mass propagating through a transparent medium like glass. Unlike this optical case, however, the "medium" through which the supernova neutrinos pass consists, in part, of other neutrinos. This makes the neutrino flavor transformation problem fiercely nonlinear: the potentials that determine how neutrinos change their flavors depend on the flavor states of the neutrinos.

These nonlinear effects become important in environments where the neutrino fluxes are substantial, such as core-collapse supernovae, compact object mergers, and also the early Universe. A complete treatment of flavor-transformation physics in these environments is important, because the charged-current weak interactions are flavor dependent at typical temperatures (∼MeV)—the ν_e's participate, but ν_μ's and ν_τ's do not as there are no μ or τ leptons around to scatter on. As a result, the effective scattering cross sections for ν_e are larger than those for ν_μ and ν_τ, resulting in different energy deposition rates—relevant for the supernova explosion mechanism. Moreover, the charged-current weak processes $\nu_e + n \to p + e^-$ and $\bar{\nu}_e + p \to n + e^+$ determine the n/p ratio, and therefore knowing the flavor content is essential for evaluating the nucleosynthesis prospects in these environments.

Thermal processes during the core collapse manufacture neutrino-antineutrino pairs of all flavors and these thermalize with the electron capture-created ν_e's. The net result is a rough equipartition of energy among all six types of neutrinos. Neutrinos of different flavors, however, have correspondingly different interactions in the matter near the neutrino sphere. The result is that electron-flavor neutrinos, with the largest interactions, decouple furthest out, where it is coolest, and have lower average energies as a consequence. μ and τ flavor neutrinos and their antiparticles have no charged-current weak interactions, and so these neutrinos decouple deeper in, where it is hotter. Consequently, these are on average more energetic. Electron antineutrinos have energies in between those of the electron neutrinos and the μ or τ flavor neutrinos.

Neutrinos diffuse out of the hot proto-neutron star core with a typical random walk time of seconds. This rather long diffusion time also sets the time scale over which neutrino spectral parameters and fluxes change. The time scale for these changes can then be long compared to neutrino transit times across regions of interest. Numerical studies of supernova neutrino flavor evolution have traditionally sought to take advantage of this situation by seeking stationary, time-independent solutions to the evolution equations, wherein the neutrino fluxes/spectra depend only on position. These numerical studies, in which some $\sim 10^7$ nonlinearly coupled Schrödinger-like equations are solved on a supercomputer, have yielded unexpected and surprising results [2–47]. Nonlinearity in the neutrino flavor potentials can give rise to collective neutrino flavor oscillations, where significant populations of neutrinos in the supernova envelope can execute synchronized or other organized and simultaneous changes in flavor, across a range of neutrino energies and in a large region of space or time.

One of the limitations of current simulations of neutrino flavor evolution in supernovae is the failure to account for potentials arising from neutrino direction-changing scattering. This is the neutrino halo effect. Even though a relatively small fraction of neutrinos undergo direction-changing scattering, they could nevertheless contribute significantly to the forward-scattering potential felt by the outward-streaming neutrinos. This is a consequence of the peculiar intersection-angle dependence of the weak-interaction potential. In certain regions of the envelope, and for certain epochs, it has been shown that the potential term arising from the halo neutrinos could in fact be the dominant term [34,35]. A complete treatment of neutrino flavor evolution that includes the effects of both forward and direction-changing scattering necessitates the use of the so-called "quantum kinetic equations" (QKE) [16,53–69,92,93]. In high-density regions, where the scattering rates are large so that quantum mechanical phases do not have any time to build up, the QKEs reduce to a Boltzmann-like form. In the other limit, where the neutrinos essentially free-stream and only experience coherent forward scattering, the QKEs reduce to a Liouville–von Neumann (Schrödinger-like) equation.

If in the future we are lucky enough to detect the neutrino burst from a Galactic core-collapse event, we will want to know whether the detected signal indicates that the simple forward-scattering-based optical analogy is sufficient to explain the neutrino flavor data, or whether the halo must be invoked. A key objective will be to use this signal to potentially extract information regarding the conditions in the envelope and to ascertain whether collective oscillations and their signatures like spectral swaps/splits occurred. These issues prompt the exploration of alternative calculation techniques.

[1] H. Abarbanel, *Predicting the Future: Completing Models of Observed Complex Systems* (Springer Science & Business Media, New York, 2013).
[2] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Phys. Rev. D **74**, 105014 (2006).
[3] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Phys. Rev. Lett. **97**, 241101 (2006).
[4] H. Duan, G. M. Fuller, and Y.-Z. Qian, Phys. Rev. D **74**, 123004 (2006).
[5] H. Duan, G. M. Fuller, and Y.-Z. Qian, Phys. Rev. D **76**, 085013 (2007).
[6] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Phys. Rev. D **75**, 125005 (2007).
[7] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Phys. Rev. Lett. **99**, 241802 (2007).
[8] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Phys. Rev. Lett. **100**, 021101 (2008).
[9] H. Duan, G. M. Fuller, and Y.-Z. Qian, Phys. Rev. D **77**, 085016 (2008).
[10] H. Duan, G. M. Fuller, and J. Carlson, Comput. Sci. Discovery **1**, 015007 (2008).
[11] J. F. Cherry, G. M. Fuller, J. Carlson, H. Duan, and Y.-Z. Qian, Phys. Rev. D **82**, 085025 (2010).
[12] H. Duan, G. M. Fuller, and Y.-Z. Qian, Annu. Rev. Nucl. Part. Sci. **60**, 569 (2010).
[13] H. Duan and A. Friedland, Phys. Rev. Lett. **106**, 091101 (2011).
[14] J. F. Cherry, M.-R. Wu, J. Carlson, H. Duan, G. M. Fuller, and Y.-Z. Qian, Phys. Rev. D **84**, 105034 (2011).
[15] J. F. Cherry, M.-R. Wu, J. Carlson, H. Duan, G. M. Fuller, and Y.-Z. Qian, Phys. Rev. D **85**, 125010 (2012).
[16] Y. Zhang and A. Burrows, Phys. Rev. D **88**, 105009 (2013).
[17] A. B. Balantekin, in *American Institute of Physics Conference Series*, American Institute of Physics Conference Series, Vol. 1594, edited by S. Jeong, N. Imai, H. Miyatake, and T. Kajino (AIP, New York, 2014), pp. 313–318.
[18] C. Lunardini, in *American Institute of Physics Conference Series*, American Institute of Physics Conference Series, Vol. 1666 (2015), p. 070001.
[19] A. B. Balantekin, in *American Institute of Physics Conference Series*, American Institute of Physics Conference Series, Vol. 1743 (AIP, New York, 2016), p. 040001.
[20] E. Akhmedov and A. Mirizzi, Nucl. Phys. **B908**, 382 (2016).
[21] S. Chakraborty, R. S. Hansen, I. Izaguirre, and G. G. Raffelt, J. Cosmol. Astropart. Phys. 1 (2016) 028.
[22] R. Barbieri and A. Dolgov, Nucl. Phys. **B349**, 743 (1991).
[23] C. Volpe, Acta Phys. Pol. B Proc. Suppl. **9**, 769 (2016).
[24] A. B. Balantekin and Y. Pehlivan, J. Phys. G **34**, 47 (2007).
[25] L. Johns, M. Mina, V. Cirigliano, M. W. Paris, and G. M. Fuller, Phys. Rev. D **94**, 083505 (2016).
[26] G. Raffelt, S. Sarikas, and D. de Sousa Seixas, Phys. Rev. Lett. **111**, 091101 (2013).
[27] S. Sarikas, I. Tamborra, G. Raffelt, L. Hüdepohl, and H.-T. Janka, Phys. Rev. D **85**, 113007 (2012).
[28] G. G. Raffelt and A. Y. Smirnov, Phys. Rev. D **76**, 081301 (2007).
[29] S. Hannestad, G. G. Raffelt, G. Sigl, and Y. Y. Y. Wong, Phys. Rev. D **74**, 105010 (2006).
[30] B. Dasgupta, A. Dighe, G. G. Raffelt, and A. Y. Smirnov, Phys. Rev. Lett. **103**, 051105 (2009).
[31] D. Nötzold and G. Raffelt, Nucl. Phys. **B307**, 924 (1988).
[32] S. Pastor, G. Raffelt, and D. V. Semikoz, Phys. Rev. D **65**, 053011 (2002).
[33] B. Dasgupta, A. Dighe, A. Mirizzi, and G. Raffelt, Phys. Rev. D **78**, 033014 (2008).
[34] J. F. Cherry, J. Carlson, A. Friedland, G. M. Fuller, and A. Vlasenko, Phys. Rev. D **87**, 085037 (2013).
[35] J. F. Cherry, J. Carlson, A. Friedland, G. M. Fuller, and A. Vlasenko, Phys. Rev. Lett. **108**, 261104 (2012).
[36] B. Dasgupta and A. Mirizzi, Phys. Rev. D **92**, 125030 (2015).
[37] S. Abbar and H. Duan, Phys. Lett. B **751**, 43 (2015).
[38] H. Duan and S. Shalgar, Phys. Lett. B **747**, 139 (2015).
[39] H. Duan, Int. J. Mod. Phys. E **24**, 1541008 (2015).
[40] S. Abbar, H. Duan, and S. Shalgar, Phys. Rev. D **92**, 065019 (2015).
[41] H. Duan and S. Shalgar, J. Cosmol. Astropart. Phys. 10 (2014) 084.
[42] Y.-Z. Qian, G. M. Fuller, G. J. Mathews, R. W. Mayle, J. R. Wilson, and S. E. Woosley, Phys. Rev. Lett. **71**, 1965 (1993).
[43] Y.-Z. Qian and G. M. Fuller, Bull. Am. Astron. Soc. **24**, 1264 (1992).
[44] Y.-Z. Qian and G. M. Fuller, Phys. Rev. D **51**, 1479 (1995).
[45] Y.-Z. Qian and G. M. Fuller, Phys. Rev. D **52**, 656 (1995).
[46] A. Banerjee, A. Dighe, and G. Raffelt, Phys. Rev. D **84**, 053013 (2011).
[47] A. Friedland, Phys. Rev. Lett. **104**, 191102 (2010).
[48] R. F. Sawyer, Phys. Rev. D **72**, 045003 (2005).
[49] R. F. Sawyer, Phys. Rev. Lett. **116**, 081101 (2016).
[50] I. Izaguirre, G. Raffelt, and I. Tamborra, Phys. Rev. Lett. **118**, 021101 (2017).
[51] B. Dasgupta, A. Mirizzi, and M. Sen, J. Cosmol. Astropart. Phys. 02 (2017) 019.
[52] R. F. Sawyer, Phys. Rev. D **79**, 105003 (2009).
[53] A. Vlasenko, G. M. Fuller, and V. Cirigliano, arXiv: 1406.6724.
[54] C. Volpe, D. Väänänen, and C. Espinoza, Phys. Rev. D **87**, 113010 (2013).
[55] A. Vlasenko, G. M. Fuller, and V. Cirigliano, Phys. Rev. D **89**, 105004 (2014).
[56] J. Serreau and C. Volpe, Phys. Rev. D **90**, 125040 (2014).
[57] P. A. Andreev, Physica A (Amsterdam) **432A**, 108 (2015).
[58] C. Volpe, Int. J. Mod. Phys. E **24**, 1541009 (2015).
[59] V. Cirigliano, G. M. Fuller, and A. Vlasenko, Phys. Lett. B **747**, 27 (2015).
[60] A. Kartavtsev, G. Raffelt, and H. Vogel, Phys. Rev. D **91**, 125020 (2015).
[61] A. Dobrynina, A. Kartavtsev, and G. Raffelt, Phys. Rev. D **93**, 125030 (2016).
[62] A. Chatelain and C. Volpe, Phys. Rev. D **95**, 043005 (2017).
[63] C. Volpe, J. Phys. Conf. Ser. **718**, 062068 (2016).
[64] D. N. Blaschke and V. Cirigliano, Phys. Rev. D **94**, 033009 (2016).
[65] G. Sigl and G. Raffelt, Nucl. Phys. **B406**, 423 (1993).
[66] G. Raffelt, G. Sigl, and L. Stodolsky, Phys. Rev. Lett. **70**, 2363 (1993).

[67] G. Raffelt and G. Sigl, Astropart. Phys. **1**, 165 (1993).
[68] P. Strack and A. Burrows, Phys. Rev. D **71**, 093004 (2005).
[69] C. Y. Cardall, Phys. Rev. D **78**, 085017 (2008).
[70] K. Abe *et al.*, arXiv:1109.3262.
[71] K. Abe, H. Aihara, A. Aimi, R. Akutsu, C. Andreopoulos, I. Anghel, L. H. V. Anthony *et al.* (Hyper-Kamiokande Proto-Collaboration), arXiv:1611.06118.
[72] J. Ahrens *et al.*, Astropart. Phys. **16**, 345 (2002).
[73] R. Abbasi, Y. Abdou, T. Abu-Zayyad, M. Ackermann, J. Adams, J. A. Aguilar, M. Ahlers, M. M. Allen, D. Altmann *et al.* (IceCube Collaboration), Astron. Astrophys. **535**, A109 (2011); Astron. Astrophys. **563**, C1(E) (2014).
[74] R. Acciarri, M. A. Acero, M. Adamowski, C. Adams, P. Adamson, S. Adhikari, Z. Ahmad, C. H. Albright, T. Alion *et al.* (DUNE Collaboration), arXiv:1512.06148.
[75] R. Laha, J. F. Beacom, and S. K. Agarwalla, arXiv:1412.8425.
[76] F. An *et al.*, J. Phys. G **43**, 030401 (2016).
[77] J.-S. Lu, Y.-F. Li, and S. Zhou, Phys. Rev. D **94**, 023006 (2016).
[78] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming* (SIAM, 2010), Vol. 19.
[79] R. Kimura, J. Wind Eng. Ind. Aerodyn. **90**, 1403 (2002).
[80] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, Cambridge, England, 2003).
[81] G. Evensen, *Data Assimilation: The Ensemble Kalman Filter* (Springer Science & Business Media, New York, 2009).
[82] W. G. Whartenby, J. C. Quinn, and H. D. Abarbanel, Mon. Weather Rev. **141**, 2502 (2013).
[83] D. Rey, J. An, J. Ye, and H. D. Abarbanel (to be published).
[84] C. D. Meliza, M. Kostuk, H. Huang, A. Nogaret, D. Margoliash, and H. D. Abarbanel, Biol. Cybern. **108**, 495 (2014).
[85] N. Kadakia, E. Armstrong, D. Breen, U. Morone, A. Daou, D. Margoliash, and H. D. I. Abarbanel, arXiv:1608.04005.
[86] D. Breen, S. Shirman, E. Armstrong, N. Kadakia, and H. Abarbanel, arXiv:1608.04433.
[87] B. A. Toth, M. Kostuk, C. D. Meliza, D. Margoliash, and H. D. Abarbanel, Biol. Cybern. **105**, 217 (2011).
[88] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (SIAM, 2005).
[89] A. Wächter and L. T. Biegler, Math. Program. **106**, 25 (2006).
[90] A. B. Balantekin and G. M. Fuller, Phys. Lett. B **471**, 195 (1999).
[91] D. O. Caldwell, G. M. Fuller, and Y.-Z. Qian, Phys. Rev. D **61**, 123005 (2000).
[92] M. A. Rudzsky, Astrophys. Space Sci. **165**, 65 (1990).
[93] B. H. J. McKellar and M. J. Thomson, Phys. Rev. D **49**, 2710 (1994).
[94] L. Wolfenstein, Phys. Rev. D **17**, 2369 (1978).
[95] S. P. Mikheyev and A. Y. Smirnov, Yad. Fiz. **42** (1985) [Sov. J. Nucl. Phys. **42**, 913 (1985)].
[96] R. C. Schirato and G. M. Fuller, arXiv:astro-ph/0205390.
[97] R. Tomàs, M. Kachelrieß, G. Raffelt, A. Dighe, H.-T. Janka, and L. Scheck, J. Cosmol. Astropart. Phys. 9 (2004) 015.
[98] J. Xu, L.-J. Hu, R.-C. Li, X.-H. Guo, and B.-L. Young, arXiv:1412.7240.
[99] G. Fogli, E. Lisi, A. Mirizzi, and D. Montanino, J. Cosmol. Astropart. Phys. 06 (2006) 012.
[100] R. D. Hoffman, S. E. Woosley, G. M. Fuller, and B. S. Meyer, Astrophys. J. **460**, 478 (1996).
[101] R. I. Epstein, S. A. Colgate, and W. C. Haxton, Phys. Rev. Lett. **61**, 2038 (1988).
[102] P. Banerjee, Y.-Z. Qian, A. Heger, and W. Haxton, EPJ Web Conf. **109**, 06001 (2016).
[103] Abarbanel (private communication).
[104] A. Malkus, A. Friedland, and G. C. McLaughlin, arXiv:1403.5797.
[105] M.-R. Wu, H. Duan, and Y.-Z. Qian, Phys. Lett. B **752**, 89 (2016).
[106] D. Väänänen and G. C. McLaughlin, Phys. Rev. D **93**, 105044 (2016).
[107] A. Malkus, G. C. McLaughlin, and R. Surman, Phys. Rev. D **93**, 045021 (2016).
[108] Y.-L. Zhu, A. Perego, and G. C. McLaughlin, Phys. Rev. D **94**, 105006 (2016).
[109] D. Breen, Characterizing Real World Neural Systems Using Variational Methods of Data Assimilation (2017), http://escholarship.org/uc/item/8dz7c69k.
[110] S. Shirman (to be published).
[111] L. Isaksen, European Centre for Medium-Range Weather Forecasts, 2013.
[112] European Centre for Medium-Range Weather Forecasts, http://www.ecmwf.int/en/computing/our-facilities.
[113] DOE-ASCR Nuclear Physics computing review (2017), edited by J. Carlson and M. J. Savage (unpublished).
[114] J. C. Quinn and H. D. Abarbanel, Q. J. R. Meteorol. Soc. **136**, 1855 (2010).
[115] J. C. Quinn and H. D. Abarbanel, J. Comput. Phys. **230**, 8168 (2011).
[116] A. Reyes, D. Lee, C. Graziani, and P. Tzeferacos, arXiv:1611.08084.
[117] A. D. Wyner, Inf. Control **38**, 51 (1978).
[118] G. M. Fuller and Y.-Z. Qian, Phys. Rev. D **73**, 023004 (2006).
[119] H. A. Bethe, Rev. Mod. Phys. **62**, 801 (1990).
[120] A. Burrows, J. Hayes, and B. A. Fryxell, Astrophys. J. **450**, 830 (1995).
[121] H.-T. Janka, K. Langanke, A. Marek, G. Martínez-Pinedo, and B. Müller, Phys. Rep. **442**, 38 (2007).