ABSTRACT

Core-collapse supernovae are among the most magnificent events in the observable universe. They produce many of the chemical elements necessary for life to exist and their remnants—neutron stars and black holes—are interesting astrophysical objects in their own right. However, despite millennia of observations and almost a century of astrophysical study, the explosion mechanism of core-collapse supernovae is not yet well understood.

Hyper-Kamiokande is a next-generation neutrino detector that will be able to observe the neutrino flux from the next galactic core-collapse supernova in unprecedented detail. We focus on the first 500 ms of the neutrino burst, corresponding to the accretion phase, and use a newly-developed, high-precision supernova event generator to simulate Hyper-Kamiokande’s response to five different supernova models. We show that Hyper-Kamiokande will be able to distinguish between these models with high accuracy for a supernova at a distance of up to 100 kpc.

Once the next galactic supernova happens, this ability will be a powerful tool for guiding simulations towards a precise reproduction of the explosion mechanism observed in nature.

1. INTRODUCTION

A star with a mass of at least 8 $M_{\odot}$ typically dies in a core-collapse supernova (ccSN). In the process, large amounts of intermediate-mass chemical elements are created and ejected into interstellar space, influencing the star formation rate and stellar evolution in their galactic neighbourhood. The compact remnant, meanwhile, is a neutron star or a black hole—important subjects of astrophysical research in their own right. Understanding the ccSN explosion mechanism is therefore one of the central goals of astrophysics.

The electromagnetic emission from a ccSN begins minutes to hours after the initial explosion, when the outgoing shock wave breaks through the surface of the star. It is therefore largely decoupled from the processes that occur during the explosion. The observation of neutrinos from SN1987A was consistent with basic features predicted by the delayed neutrino-driven explosion mechanism developed by Wilson and Bethe in the 1980s (Wilson 1985; Bethe & Wilson 1985); specifically, the presence of an accretion phase in the first $\sim 500$ ms (Loredo & Lamb 2002). However, with a total of two dozen events detected in the Kamiokande (Hirata et al. 1987, 1988), IMB (Bionta et al. 1987)
and Baksan (Alekseev et al. 1987) detectors, the available statistics were too low to determine details of the explosion mechanism.

In the decades since, progress in this area has largely relied on computer simulations. While these simulations have made considerable progress and increasingly sophisticated three-dimensional models have become available in recent years (see e.g. Hanke et al. (2013); O’Connor & Couch (2018a); Burrows et al. (2020)), they are still limited by the available computing power and exhibit significant quantitative and in many cases even qualitative differences.

Once the next galactic ccSN happens, current and next-generation neutrino detectors will make a high-statistics observation of the neutrino burst which will provide valuable input to simulations. Previous work has demonstrated that this would make it possible to identify whether the signal exhibits certain features like SASI (Lund et al. 2010; Tamborra et al. 2013) or LESA (Tamborra et al. 2014). For recent reviews of expected features of the neutrino signal, see Mirizzi et al. (2016); Horiuchi & Kneller (2018). However, while model discrimination based on these features may allow us to exclude some classes of models, no general method for distinguishing between any two different ccSN models based on their neutrino signal has yet been presented.

In this paper, we present a log-likelihood method that makes optimal use of the full time and energy information available from many neutrino detectors to identify which supernova model best matches a set of observed events. Using a newly-developed, high-precision supernova event generator and a realistic detector simulation and event reconstruction, we investigate Hyper-Kamiokande’s response to five supernova models simulated by different groups around the world. We show that this method requires just 100 (300) events within the first 500 ms of the supernova burst—corresponding to a supernova distance of at least 102 kpc (59 kpc) for normal mass ordering or 97 kpc (56 kpc) for inverted mass ordering—to distinguish between different supernova models with (high) accuracy.

This paper is organized as follows. We briefly introduce the Hyper-Kamiokande detector and summarise its sensitivity to supernova neutrinos in section 2. Section 3 describes our simulations, giving an overview over the supernova models employed (section 3.1), event generation (section 3.2), simulation and reconstruction (section 3.3) and data reduction (section 3.4). In section 4, we present our likelihood function and determine how accurately it can distinguish between supernova models, before concluding in section 5.

2. HYPER-KAMIOKANDE

Hyper-Kamiokande (Abe et al. 2018) is a next-generation water Cherenkov detector that will be built near the town Kamioka in Japan’s Gifu Prefecture, approximately 8 km south of the currently operating Super-Kamiokande detector (Fukuda et al. 2003). Located beneath the peak of Mount Nijugo, it will have an overburden of 650 m of rock (1750 m. w. e.). Construction has started in 2020, with data-taking scheduled to start in 2027. Its physics goals include precision measurements of neutrino oscillation parameters (by measuring atmospheric neutrinos, accelerator neutrinos from the upgraded J-PARC beamline (Abe et al. 2019) and solar neutrinos) as well as searches for proton decay and for astrophysical neutrinos from a wide range of sources. In this section, we first give a brief overview over the detector design and then discuss Hyper-Kamiokande’s sensitivity to supernova neutrinos.

2.1. Detector Design

The basic design of Hyper-Kamiokande (see figure 1) is similar to that of Super-Kamiokande. It is a large, cylindrical detector with a height of 71 m and diameter of 68 m, filled with 258 kton of ultra-pure water.1 It is optically separated into an outer detector with a width of 2 m at the top and bottom or 1 m at the sides, which acts as both shielding and active veto, and a 217 kton cylindrical inner detector. The structure dividing both detector regions has a diameter of 60 cm and contains an array of photosensors as well as front-end electronics to collect and digitize signals from these photosensors.

The exact photosensor configuration of Hyper-Kamiokande has not yet been determined. The baseline design uses a 40% photocoverage with a new model of 50 cm photomultiplier tube (PMT) that offers improved time and charge resolution compared to the PMT model used in Super-Kamiokande. Alternative designs are currently being finalized. Here, as a very conservative estimate, we assume a 20% photocoverage with the new 50 cm PMT model.

2.2. Supernova Observations

1 The simulations throughout this study used an earlier design with a height of 60 m and diameter of 74 m. The fiducial volume, and thus the results of this study, are not affected.
Figure 1. Drawing of the Hyper-Kamiokande detector.

Figure 2. Left: Expected number of events as a function of supernova distance. Right: True energy spectra of prompt events in the full inner detector for a supernova at 10 kpc; for reference, the energy threshold used in this analysis (see section 3.4) is indicated by a dashed grey line. Both panels assume the supernova model by Totani et al. (1998). Solid (dashed) lines correspond to normal (inverted) mass ordering, while different colours correspond to the interaction channels inverse beta decay (black), $\nu e$-scattering (red), $\nu_e + ^{16}O$ CC (purple) and $\bar{\nu}_e + ^{16}O$ CC (light blue).

For a galactic supernova at a fiducial distance of 10 kpc, Hyper-Kamiokande is expected to observe 54 000 to 90 000 events in a burst with a total duration of a few tens of seconds. For a nearby supernova (e.g. Betelgeuse at 0.2 kpc), the peak event rate could reach $10^8$ Hz. This rate was taken into account during the design of the DAQ system. As shown in the left panel of figure 2, the large volume also gives Hyper-Kamiokande an unprecedented ability to detect
neutrinos from supernovae beyond the Milky Way: For a supernova in the Large Magellanic Cloud at 50 kpc distance, it would still detect about 3000 events, while for a supernova in the Andromeda galaxy (M31) at 780 kpc distance, O(10) events are expected.

Hyper-Kamiokande can reconstruct the time and energy of each individual event, allowing it to reconstruct the neutrino spectrum. The right panel of figure 2 shows energy spectra for the interaction channels considered in this paper.

The main interaction channel, inverse beta decay ($\bar{\nu}_e + p \rightarrow n + e^+$), is responsible for about 90% of events, making Hyper-Kamiokande most sensitive to $\bar{\nu}_e$. Elastic neutrino-electron scattering ($\nu + e^- \rightarrow \nu + e^-$) is a subdominant interaction channel to which all neutrino flavours contribute. The angular distribution of elastically scattered electrons is strongly peaked into a forward direction, which can be used to determine the direction of a supernova at the fiducial distance of 10 kpc with an accuracy of about 1° (Abe et al. 2018). Charged-current interactions of $\nu_e$ and $\bar{\nu}_e$ on $^{16}$O nuclei are subdominant channels. Due to their high energy threshold and the steep energy dependence of their cross sections, both channels are a very sensitive probe of the high-energy tail of the supernova neutrino flux, making up anywhere from <1% to about 10% of observed events.

3. SIMULATIONS

3.1. Supernova Models

While computer simulations of core-collapse supernovae have made significant advances in recent decades, they are still limited by the available computing power. To overcome this problem, modelling groups employ a variety of different approximations and simplifying assumptions in their models, which lead to significant quantitative and in many cases even qualitative differences between different simulations.

Due to these uncertainties, and in order to demonstrate the broad applicability of the model discrimination method introduced in this work, we use a selection of five unrelated models here: a one-dimensional model that is primarily of historic interest (see section 3.1.1), two one-dimensional models from recent parametric studies (see sections 3.1.2 and 3.1.3) and two more complex multi-dimensional models (see sections 3.1.4 and 3.1.5). These simulations were performed by different groups using a variety of progenitors and simulation codes. They are intended to represent the much wider range of available models.

Figure 3 shows an overview over these models. In this section, we briefly describe these models.

3.1.1. Totani

This one-dimensional model (Totani et al. 1998), which is also referred to as the “Livermore model” or “Wilson model” in literature, is one of a small number of models that include the late-time evolution of the neutrino emission. While it is now dated and has been surpassed by more accurate models, it is still used as a baseline model in many recent publications (Acciarri et al. 2015; Abe et al. 2016a, 2018).

It uses a 20 M$_\odot$ progenitor, which was modeled to reproduce the light curve of SN1987A, and a simulation code developed by Wilson and Mayle (Wilson et al. 1986; Mayle et al. 1987). Neutrino transport is modelled by the flux-limited diffusion approximation with 20 logarithmically spaced energy groups up to 322.5 MeV. The simulation is one-dimensional and was performed from the start of collapse to 18 s after the core bounce.

3.1.2. Nakazato

This family of models (Nakazato et al. 2013) contains progenitors with different initial masses and metallicities. In this work, we focus on the 20 M$_\odot$ progenitor with solar metallicity ($Z = 0.02$). The one-dimensional simulation was performed from the start of collapse to 20 s after the core bounce in two stages. Here, we only use data from the first stage of the simulation, which contains the first 520 ms post bounce. It used the equation of state by Shen et al. (1998), 20 variably spaced energy groups up to 300 MeV and a general relativistic neutrino-radiation-hydrodynamics ($\nu$RHD) code that solves the differential equations for hydrodynamics and neutrino transport simultaneously (Sumiyoshi et al. 2005).

3.1.3. Couch

This family of models (Couch et al. 2019; Warren 2019) uses an approach for including effects of convection and turbulence in a one-dimensional simulation, which the authors call STIR (Supernova Turbulence In Reduced-dimensionality). In this approach, the effective strength of convection depends on one parameter, $a_\Lambda$, which can be tuned to reproduce
Figure 3. Event rate (top) and mean energy (bottom) of observed events in Hyper-Kamiokande, as predicted by the five supernova models used in this paper for normal (left) or inverted (right) mass ordering. All plots show the time interval from 20 ms to 520 ms after core bounce. The event rate is normalized to produce the same total number of events for each model, reflecting the assumption made in this paper that the distance of the supernova is unknown.

results from a three-dimensional simulation of the same progenitor (O’Connor & Couch 2018a). The model family contains 138 solar-metallicity progenitors with masses from 9 M\(_\odot\) to 120 M\(_\odot\). Here, we use results from the simulation of a 20 M\(_\odot\) progenitor (Sukhbold & Woosley 2014) with \(\alpha_A = 0.8\).\(^2\)

The simulation was implemented in the FLASH simulation framework (Fryxell et al. 2000; Dubey et al. 2009) using a newly-implemented hydrodynamics solver with a modified effective potential to approximate effects of general relativity (Marek et al. 2006; O’Connor & Couch 2018b) and the SFHo equation of state (Steiner et al. 2013).

Neutrino transport is simulated using a so-called “M1” transport scheme (O’Connor 2015; O’Connor & Couch 2018b) with 12 logarithmically spaced energy groups up to 250 MeV. Starting at 5 ms post-bounce, effects of neutrino-electron scattering with energy transfer are turned off to reduce the computational resources required. While this has little impact on the supernova dynamics, it does result in an increased mean energy for \(\nu_x\) (O’Connor & Couch 2018b).

3.1.4. Tamborra

This model (Hanke et al. 2013; Tamborra et al. 2014) is a pioneering three-dimensional supernova simulation with energy-dependent neutrino transport. We use results from the simulation of a 27 M\(_\odot\) progenitor (Woosley et al. 2002).

The simulation was performed using the PROMETHEUS-VERTEX code consisting of the hydrodynamics solver PROMETHEUS (Fryxell et al. 1991), which implements the piecewise-parabolic method (Colella & Woodward 1984), and the neutrino transport code VERTEX (Rampp & Janka 2002), which uses the “ray-by-ray-plus” approach for velocity-

\(^2\) This is an early version of the simulations, which differs slightly from the simulations described in the final version of Couch et al. (2019).
and energy-dependent neutrino transport (Buras et al. 2006). The simulation uses a recent set of neutrino interaction rates (Müller et al. 2012), the Lattimer and Swesty equation of state with compressibility $K = 220\text{MeV}$ (Lattimer & Swesty 1991) and an effective potential to account for general relativistic corrections to Newtonian gravity (Marek et al. 2006). In multi-dimensional simulations, the neutrino signal inherently depends on the direction of the observer relative to the progenitor. Here, we use the fluxes in the “violet” observer direction identified in Tamborra et al. (2014), which exhibits a large amplitude of the SASI oscillations in the luminosity and mean energy of neutrinos.

### 3.1.5. Vartanyan

This model is a recent two-dimensional simulation of a 9 $\text{M}_\odot$ progenitor with solar metallicity (Sukhbold et al. 2016). It is similar to the simulations presented in Radice et al. (2017); Seadrow et al. (2018) but used a different equation of state and grid resolution, which caused some physical and numerical differences. As a result, while the luminosity and mean energy are qualitatively very similar to those described in Seadrow et al. (2018), exact values may differ by several percent.

This simulation was performed using the neutrino-radiation-hydrodynamics code Fornax (Skinner et al. 2019), which combines a radiation hydrodynamics solver using a generalized variant of the piecewise-parabolic method (Colella & Woodward 1984) with neutrino transport using the “M1” scheme (Thorne 1981; Shibata et al. 2011; Murchikova et al. 2017). It used 20 logarithmically spaced energy groups with energies up to 300 MeV (100 MeV) for $\nu_e$ and $\overline{\nu}_e$, a detailed set of neutrino-matter interactions (Burrows et al. 2006), the SFHo equation of state (Steiner et al. 2013) and an effective potential to account for general relativistic corrections to Newtonian gravity (Marek et al. 2006).

### 3.2. Event Generation

We have developed a new supernova neutrino event generator called sntools (snt 2020). It is open source, written in Python and makes heavy use of the NumPy (van der Walt et al. 2011) and SciPy (Virtanen et al. 2020) libraries. In the following, we briefly discuss the neutrino interaction cross sections and treatment of flavour conversion implemented in sntools, before describing the generated data sets.

#### 3.2.1. Cross Sections

sntools implements modern, high-precision cross-sections for the interaction channels described in section 2.2. For inverse beta decay, it implements the full result from Strumia & Vissani (2003), including radiative corrections based on the approximation in Kurylov et al. (2003). For neutrino-electron scattering, it implements the result from Bahcall et al. (1995), which includes one-loop QCD and electroweak corrections as well as QED radiative corrections. For charged-current interactions of $\nu_e$ and $\overline{\nu}_e$, it implements a four-group fit (Nakazato et al. 2018) based on a recent shell model calculation (Suzuki et al. 2018).

#### 3.2.2. Treatment of Neutrino Flavour Conversion

As neutrinos produced inside the supernova traverse a smoothly varying density profile while exiting the star, they experience adiabatic flavour conversion via the MSW effect. Afterwards, they propagate in a mass eigenstate until they interact. The neutrino fluxes $\Phi_{\nu_e}$ observed in a detector are therefore linear combinations of the initial fluxes $\Phi_{\nu^0}$, predicted by the supernova simulation. For normal mass ordering, this relation is given by (Dighe & Smirnov 2000)

$$\Phi_{\nu_e} = \sin^2 \theta_{13} \cdot \Phi_{\nu_e}^0 + \cos^2 \theta_{13} \cdot \Phi_{\bar{\nu}_e}^0 \quad (1)$$

$$\Phi_{\bar{\nu}_e} = \cos^2 \theta_{12} \cdot \sin^2 \theta_{13} \cdot \Phi_{\nu_e}^0 + (1 - \cos^2 \theta_{12} \cos^2 \theta_{13}) \cdot \Phi_{\bar{\nu}_e}^0 \quad (2)$$

$$2\Phi_{\nu_e} = \cos^2 \theta_{13} \cdot \Phi_{\nu_e}^0 + (1 + \sin^2 \theta_{13}) \cdot \Phi_{\bar{\nu}_e}^0 \quad (3)$$

$$2\Phi_{\bar{\nu}_e} = (1 - \cos^2 \theta_{12} \cos^2 \theta_{13}) \cdot \Phi_{\nu_e}^0 + (1 + \cos^2 \theta_{12} \cos^2 \theta_{13}) \cdot \Phi_{\bar{\nu}_e}^0 \quad (4)$$

while for inverted mass ordering, it is given by

$$\Phi_{\nu_e} = \sin^2 \theta_{12} \cos^2 \theta_{13} \cdot \Phi_{\nu_e}^0 + (1 - \sin^2 \theta_{12} \cos^2 \theta_{13}) \cdot \Phi_{\bar{\nu}_e}^0 \quad (5)$$

$$\Phi_{\bar{\nu}_e} = \sin^2 \theta_{13} \cdot \Phi_{\nu_e}^0 + \cos^2 \theta_{13} \cdot \Phi_{\bar{\nu}_e}^0 \quad (6)$$

$$2\Phi_{\nu_e} = (1 - \sin^2 \theta_{12} \cos^2 \theta_{13}) \cdot \Phi_{\nu_e}^0 + (1 + \sin^2 \theta_{12} \cos^2 \theta_{13}) \cdot \Phi_{\bar{\nu}_e}^0 \quad (7)$$

$$2\Phi_{\bar{\nu}_e} = \cos^2 \theta_{13} \cdot \Phi_{\nu_e}^0 + (1 + \sin^2 \theta_{13}) \cdot \Phi_{\bar{\nu}_e}^0 \quad (8)$$
### Table 1.

Number of events expected during the time interval of 20 ms to 520 ms for a supernova at the fiducial distance of 10 kpc ($N_{10\text{ kpc}}$) and the distances at which 100 or 300 events are expected in the inner detector of Hyper-Kamiokande ($d_{100}$ and $d_{300}$, respectively) for the five supernova models considered in this work and for both normal and inverted mass ordering.

| Model    | $N_{10\text{ kpc}}$ | $d_{100}$ | $d_{300}$ | $N_{10\text{ kpc}}$ | $d_{100}$ | $d_{300}$ |
|----------|----------------------|-----------|-----------|----------------------|-----------|-----------|
| Totani   | 20 021               | 141 kpc   | 82 kpc    | 22 717               | 151 kpc   | 87 kpc    |
| Nakazato | 17 978               | 134 kpc   | 77 kpc    | 16 005               | 127 kpc   | 73 kpc    |
| Couch    | 27 539               | 166 kpc   | 96 kpc    | 24 983               | 158 kpc   | 91 kpc    |
| Vartanyan| 10 372               | 102 kpc   | 59 kpc    | 94 000               | 97 kpc    | 56 kpc    |
| Tamborra | 25 025               | 158 kpc   | 91 kpc    | 20 274               | 142 kpc   | 82 kpc    |

In both cases, the factor of 2 in the last two equations accounts for the fact that we use $\nu_x$ (or $\bar{\nu}_x$) to refer to either $\nu_\mu$ or $\nu_\tau$ (either $\bar{\nu}_\mu$ or $\bar{\nu}_\tau$), not to their sum. These equations assume purely adiabatic transition (corresponding to $P_H = 0$ in Dighe & Smirnov (2000); Fogli et al. (2005)), which is appropriate during the early part of the neutrino emission we consider here (Fogli et al. 2005). For $\theta_{12}$ and $\theta_{13}$ we use values from the Particle Data Group (Tanabashi et al. 2018). The effect of the uncertainty in both quantities on the generated data sets is much smaller than the random fluctuations between data sets generated from the same neutrino flux.

Neutrino self-interactions near the centre of the supernova could induce additional flavour conversion (Duan et al. 2006a,b). While these collective effects are the subject of intense theoretical study, no clear picture has yet emerged of how these effects will manifest in a given supernova (Chakraborty et al. 2016) and they are therefore not considered here.

### 3.2.3. Data Sets

Using sntools, we have generated data sets for the supernova models described in section 3.1, for both normal and inverted mass ordering and for two different event counts per data set, as described below. For every combination of these parameters, we have generated 1000 data sets in order to determine how accurately Hyper-Kamiokande is able to identify the true model despite random variations in each data set. All events were distributed randomly within the inner detector of Hyper-Kamiokande.

All data sets cover the time interval from 20 ms to 520 ms after core-bounce. It contains the shock stagnation and accretion phase, which shows the largest differences between models. The earlier (neutronization burst; see Kachelrieß et al. (2005)) and later (proto-neutron star cooling; see Suwa et al. (2019)) phases of neutrino emission are much better understood and exhibit only minor variations between models, making them less relevant for model discrimination. Furthermore, due to the limited computing time available, many simulations—including the Couch, Vartanyan and Tamborra models used here—focus on the accretion phase and don’t include the full cooling phase. Accordingly, by considering only this 500 ms time interval we are able to include a wider range of models.

We have chosen the number of events per data set to be either 100 or 300. The lower data set size was chosen in order to determine the lowest number of events needed to separate the models. As table 1 shows, depending on the supernova model this correspond to a distance of at least 102 kpc (97 kpc) for normal (inverted) mass ordering. The larger size was chosen in order to demonstrate the increase in accuracy offered by a moderate increase in statistics. 300 events correspond to a supernova distance of at least 59 kpc (56 kpc) for normal (inverted) mass ordering, and is thus representative of a supernova in the Large or Small Magellenic Cloud at a distance of 50 kpc (Pietrzyński et al. 2013) or 61 kpc (Hilditch et al. 2005), respectively. A much closer supernova, i.e. within the Milky Way, would offer a much higher event rate and thus even more accurate model discrimination.

Throughout this analysis we assume that the distance to the supernova—and thus the normalization of the neutrino flux—is completely unknown; we only use the time and energy structure to distinguish between models. If additional distance information is available—e.g. because an optical counterpart is identified—this could in principle be used to further improve the model discrimination accuracy.

### 3.3. Detector Simulation and Event Reconstruction

To simulate events in Hyper-Kamiokande, we use WCSim (WCS 2020), a package for simulating water Cherenkov detectors that is based on the physics simulation framework GEANT4 (Agostinelli et al. 2003) and the data analysis.
framework ROOT (Brun & Rademakers 1997). The vertex, direction and energy reconstruction follows the same approach developed by the Super-Kamiokande collaboration (Abe et al. 2011) and is based on the BONSAI code (Smy 2007).

As discussed in section 2.1, the exact photosensor configuration of Hyper-Kamiokande has not yet been determined and we use a very conservative configuration with a 20% photocoverage with 50 cm PMTs.

3.4. Data Reduction

After reconstruction, we apply two cuts to all reconstructed events: an energy cut, which removes all events with a reconstructed kinetic energy less than 5 MeV, and a fiducial volume cut, which removes all events whose reconstructed vertex is less than 1.5 m away from the top, bottom or side walls of the inner detector. The resulting fiducial mass is 187 kt.

These cuts are intended to eliminate low-energy background from accidental coincidences of dark noise as well as radioactive decays in the detector. Analogous cuts are also used for the solar neutrino analysis in Super-Kamiokande (Abe et al. 2016b). While several more advanced cuts used in that analysis, which rely on a comparison between MC simulations and observations, cannot currently be applied to the analysis presented here, the more stringent energy cut together with the much higher event rate (10^2 Hz to 10^3 Hz for the distant supernova bursts considered here, compared to about 10^-4 Hz for solar neutrinos in Super-Kamiokande) result in an effectively background-free data set. Other backgrounds, including muon-induced spallation events or atmospheric neutrinos, occur at a much lower rate and are thus negligible during the single 500 ms time interval considered here.

Once Hyper-Kamiokande is operating and the low-energy backgrounds are characterized in detail, it will likely be possible to develop more targeted cuts that allow us to include more low-energy events and extend the fiducial volume while remaining effectively background-free.

The fiducial volume cut described above removes about 13% of all events in the inner detector. The effect of the energy cut is generally small, though it depends on the energy spectrum of the initial neutrino flux and therefore on the supernova model and the mass ordering. As an example, figure 2 shows the energy spectra in different interaction channels for the Totani model. Due to the strong energy-dependence of the cross sections, three of the four interaction channels produce almost no events at 5 MeV or below. Only elastic $\nu_e$-scattering—a subdominant channel which contributes about 5% of all events—has a significant contribution at energies below 5 MeV. Overall, out of the initial 100 or 300 events per data set more than 80% typically remain after applying these cuts.

4. RESULTS

4.1. Log-Likelihood Function

After the cuts described above, we apply a log-likelihood function to the reconstructed times and energies of the remaining events in each data set to determine how well that data set matches each of the supernova models.

This log-likelihood function is similar to one that was originally derived for analysis of SN1987A taking into account only the main interaction channel, inverse beta decay (Loredo & Lamb 1989). However, the function used here includes all interaction channels. It is derived in appendix A and given by

$$L = \ln \mathcal{L} = \sum_{i=1}^{N_{\text{obs}}} \ln \left( \sum_{\alpha} N_{i,\alpha} \right),$$

(9)

where the index $i$ runs over the $N_{\text{obs}}$ events remaining in the data set and $N_{i,\alpha}$ is the number of events predicted by a given supernova model in the interaction channel $\alpha$ in an infinitesimally small bin around the reconstructed time and energy of event $i$.

By using infinitesimally small bins in time and energy, this likelihood function makes optimal use of all available information. In contrast, using a binned chi-squared test to compare observation with models requires a sufficiently large number of events per bin to be accurate. Especially in the case of a distant supernova, where only hundreds or thousands of events may be observed in Hyper-Kamiokande, two-dimensional binning in time and energy would only be possible in very coarse bins, which would lose a lot of the available information.

The absolute numerical values of this likelihood function depend on the bin size chosen and are therefore not physically meaningful. However, when calculating likelihood ratios for different models (i.e. differences in the log-likelihood, $\Delta L = L_A - L_B$), this dependence cancels out and the ratio describes whether model A or B is more likely to produce a given data set. We will therefore exclusively use likelihood ratios to compare different models.
Figure 4. Histograms showing the distribution of $\Delta L = L_{\text{black}} - L_{\text{red}}$ for all pairs of supernova models considered here, for 100 events per data set, normal mass ordering and 20% photocoverage. The purple vertical line in each panel indicates $\Delta L = 0$.

4.2. $N=100$ Events per Data Set

Figure 4 shows pairwise comparisons of the five different models described in section 3.1 for normal mass ordering. For example, the top right panel shows a comparison of the Couch model (black histogram) and Nakazato model (red histogram). For most data sets generated from the Couch (Nakazato) model, $\Delta L = L_{\text{Couch}} - L_{\text{Nakazato}}$ is positive (negative), indicating that this method is generally able to identify the true model. However, the overlap of both histograms indicates that misidentification sometimes occurs because of random fluctuations in the data sets. Other model pairs in the figure show a similarly clear separation with only minor overlap around $\Delta L = 0$. The largest overlap is seen between the Couch and Tamborra models, indicating that these models are most similar and hardest to distinguish.

This can be seen more clearly in the top half of table 2, which compares all five supernova models simultaneously by determining which of them produces the highest likelihood for a given data set. For each model, the respective row indicates what fraction of the 1000 generated data sets were identified as which model. For example, 85.3% of Tamborra data sets were identified correctly, while 8.4% were misidentified as corresponding to the Couch model. For the Couch model, almost 80% of data sets were identified correctly, with most of the remaining data sets being misidentified as the Tamborra model. Finally, the three other models are identified correctly in over 95% of all cases. The bottom half of table 2 shows results for the inverted mass ordering. In this scenario, the largest overlap is observed between the Tamborra and Vartanyan models, with an 85–90% chance of identifying those data sets correctly and a chance of just over 10% of confusing these models for one another. As for the normal mass ordering, the other three models are identified correctly in over 95% of all cases. Histograms showing one-to-one comparisons for each pair of models can be found in appendix B.

4.3. $N=300$ Events per Data Set

When considering larger data sets, the effect of random fluctuations between individual data sets will decrease. As a result, the accuracy of model identification is expected to increase significantly.
### Table 2.

Accuracy with which the true model can be identified, for 100 events per data set. Each line shows what fraction (in %) of the 1000 data sets generated for a given model (left column) were identified as each of the five models. Correctly identified models are highlighted. Top: Normal mass ordering. Bottom: Inverted mass ordering.

| True Model | Reconstructed Model | Couch | Nakazato | Tamborra | Totani | Vartanyan |
|------------|---------------------|-------|----------|----------|--------|-----------|
| Normal     | Couch               | 79.5  | 5.7      | 12.2     | 1.2    | 1.4       |
|            | Nakazato            | 3.3   | 96.1     | 0.3      | 0.1    | 0.2       |
|            | Tamborra            | 8.4   | 0.0      | 85.3     | 3.3    | 3.0       |
|            | Totani              | 0.4   | 0.0      | 1.6      | 97.9   | 0.1       |
|            | Vartanyan           | 0.0   | 0.1      | 1.7      | 0.3    | 97.9      |

| True Model | Reconstructed Model | Couch | Nakazato | Tamborra | Totani | Vartanyan |
|------------|---------------------|-------|----------|----------|--------|-----------|
| Inverted   | Couch               | 96.0  | 3.5      | 0.4      | 0.1    | 0.0       |
|            | Nakazato            | 0.8   | 99.2     | 0.0      | 0.0    | 0.0       |
|            | Tamborra            | 0.0   | 0.1      | 85.8     | 2.1    | 12.0      |
|            | Totani              | 0.3   | 0.0      | 2.0      | 97.7   | 0.0       |
|            | Vartanyan           | 0.0   | 0.2      | 10.5     | 0.1    | 97.9      |

### Table 3.

Same as table 2 but for 300 events per data set. Top: Normal mass ordering. Bottom: Inverted mass ordering.

| True Model | Reconstructed Model | Couch | Nakazato | Tamborra | Totani | Vartanyan |
|------------|---------------------|-------|----------|----------|--------|-----------|
| Normal     | Couch               | 98.2  | 0.2      | 1.6      | 0.0    | 0.0       |
|            | Nakazato            | 0.1   | 99.9     | 0.0      | 0.0    | 0.0       |
|            | Tamborra            | 1.6   | 0.0      | 98.0     | 0.2    | 0.2       |
|            | Totani              | 0.0   | 0.0      | 0.0      | 100.0  | 0.0       |
|            | Vartanyan           | 0.0   | 0.0      | 0.0      | 0.0    | 100.0     |

| True Model | Reconstructed Model | Couch | Nakazato | Tamborra | Totani | Vartanyan |
|------------|---------------------|-------|----------|----------|--------|-----------|
| Inverted   | Couch               | 99.9  | 0.1      | 0.0      | 0.0    | 0.0       |
|            | Nakazato            | 0.0   | 100.0    | 0.0      | 0.0    | 0.0       |
|            | Tamborra            | 0.0   | 0.0      | 97.4     | 0.1    | 2.5       |
|            | Totani              | 0.0   | 0.0      | 0.0      | 100.0  | 0.0       |
|            | Vartanyan           | 0.0   | 0.0      | 0.8      | 0.0    | 99.2      |

Table 3 shows results of the model identification for 300 events per data set, which are consistent with this expectation. The top half shows results for normal mass ordering. The Couch and Tamborra models—which are most likely to be confused for each other in normal mass ordering—are now identified correctly with about 98% accuracy and the probability of misidentifying one for the other is just 1.6%. The bottom half shows results for inverted mass ordering. The Tamborra and Vartanyan models—which are most likely to be confused for each other in inverted mass ordering—are now identified correctly with over 97% accuracy.

For both normal and inverted mass ordering, the other three models are identified correctly with at least 99.9% accuracy. Histograms showing one-to-one comparisons for each pair of models can be found in appendix B.

#### 4.4. Observation of an Actual Supernova Neutrino Burst

Above, we have answered the following question: Assuming that model X describes the true neutrino fluxes from a supernova, how likely are we to correctly identify X when comparing it with a range of other models? This lets us identify which models are more or less similar to each other and assess Hyper-Kamiokande’s model discrimination.
Table 4. Interpretation of Bayes factor when comparing two models A and B. Adapted from Kass & Raftery (1995).

| $\ln B_{AB}$ | $B_{AB}$ | Evidence for model A over model B |
|--------------|---------|-----------------------------------|
| 0 to 1       | 1 to 3  | Negligible                         |
| 1 to 3       | 3 to 20 | Positive                           |
| 3 to 5       | 20 to 150 | Strong                           |
| $> 5$        | $> 150$ | Very strong                       |

The Hyper-Kamiokande Collaboration

In a Bayesian interpretation (Loredo & Lamb 2002), the ratio of likelihoods for two models A and B is equal to the Bayes factor $B_{AB}$ and equivalently, the difference in log-likelihoods is $\Delta L = \ln B_{AB}$. If there is no a priori reason to prefer one model over the other, this can be used to exclude disfavoured models beyond a certain threshold.

A suggested interpretation of Bayes factors is listed in table 4. Looking at the pairwise model comparison in figure 4, we see that this interpretation matches our intuition: The range from $\Delta L = -5$ to 5 contains almost the complete overlap between both histograms, where misidentification of data sets may occur, indicating that requiring $\Delta L \geq -5$ is unlikely to wrongly exclude the true model. At the same time, most data sets based on the wrong model are correctly excluded by this criterion. Once we observe an actual supernova neutrino burst, this criterion will therefore allow us to narrow down the list of supernova models that are compatible with the observed signal.

Since only likelihood ratios are physically meaningful as discussed in section 4.1, this will determine which model fits the observed events better than all other models. To determine whether the preferred model is actually compatible with the data, a separate goodness-of-fit test is required.

5. SUMMARY AND DISCUSSION

In this study, we introduced a likelihood function to determine how well neutrino fluxes predicted by a supernova model match an observed set of events. It makes optimal use of the timing and energy information of every reconstructed event and includes four interaction channels relevant for water Cherenkov detectors. This method is highly versatile and can in principle be used to determine any factor that affects the neutrino flux from a supernova.

As a proof of principle, we selected five different supernova models and generated data sets of 100 (300) events in Hyper-Kamiokande, corresponding to a supernova distance of at least 102 kpc (59 kpc) for normal mass ordering or 97 kpc (56 kpc) for inverted mass ordering. We then simulated and reconstructed these events with the experiment’s official software toolchain. For both normal and inverted mass ordering, using this method lets us identify the correct supernova model with high accuracy.

When the next supernova happens in the Milky Way or one of the nearby dwarf galaxies, Hyper-Kamiokande will thus be able to reliably identify a small number of supernova models that best match the observed neutrino burst. This will be a powerful tool for guiding models towards a precise reproduction of the explosion mechanism observed in nature.

Throughout this study, we have assumed a very conservative detector configuration with a 20% photocoverage. An increased photocoverage would improve the detector performance, particularly at low energies, which may allow us to introduce more targeted cuts and include more signal events while remaining effectively background-free. With improved low-energy performance, it would also be possible to tag neutrino interactions producing neutrons by detecting 2.2 MeV gamma-rays from neutron capture on hydrogen. This would let us distinguish between different interaction channels on an event-by-event basis and determine the fluxes of neutrinos and antineutrinos separately, which could further improve our model discrimination accuracy (Nikrant et al. 2018).

We have also assumed that the distance of the supernova is unknown, which leaves the normalization of the supernova fluxes open. If the distance of the supernova can be determined to sufficient accuracy—e. g. if an optical counter-

Note that we show $B_{ij}$ here, whereas the original paper lists $2B_{ij}$ due to its similarity with the more familiar $\Delta \chi^2$ values. (Kass & Raftery 1995; Ianni et al. 2009)
part is visible—this would fix that normalization factor and further help distinguish between models that predict a
different number of events at a fixed distance. Other neutrino experiments—particularly those employing different
and complementary detection techniques—could apply an analogous likelihood method to their own observations; the
combined likelihood of a given model would then simply be the product of the likelihoods calculated by each individual
experiment. Near-future gravitational wave detectors are expected to be sensitive to supernovae at distances of up
to a few tens of kpc (Abbott et al. 2018), making it possible in principle to include the gravitational wave signal to
improve the model identification further.

ACKNOWLEDGMENTS

We thank MacKenzie Warren, Ken’ichiro Nakazato, Tomonori Totani, Adam Burrows, David Vartanyan and Irene
Tamborra for access to the supernova models used in this work and for answering various related questions.

This work was supported by MEXT Grant-in-Aid for Scientific Research on Innovative Areas titled “Exploration of
Particle Physics and Cosmology with Neutrinos” under Grants No. 18H05535, No. 18H05536 and No. 18H5537. In
addition, participation of individual researchers has been further supported by funds from JSPS, Japan; the European
Union’s Horizon 2020 Research and Innovation Programme H2020 grant numbers RISE-GA822070-JENNIFIER2 2020
and RISE-GA872549-SK2HK; SSTF-BA1402-06, NRF grants No. 2009-0083526, NRF-2015R1A2A05001869, NRF-
2016R1D1A1A02936965, NRF-2016R1D1A3B02010606 and NRF-2017R1A2B4102757 funded by the Korean govern-
ment (MSIP); JSPS-RFBR Grant #20-52-50010/20 and the Ministry of Science and Higher Education under contract
#075-15-2020-778, Russia; Brazilian Funding agencies, CNPq and CAPES; STFC ST/R00031X/2, ST/T002891/1,
ST/V002872/1, Consolidated Grants, UKRI MR/S032843/1 and MR/S034102/1, UK.

Software: BONSAI (Smy 2007), sntools (snt 2020), WCSim (WCS 2020), matplotlib (Hunter 2007), NumPy (van der
Walt et al. 2011), SciPy (Virtanen et al. 2020)

APPENDIX

A. DERIVATION OF LIKELIHOOD FUNCTION

In this section, we derive the likelihood function introduced in section 4.1. It is based on a likelihood function
derived by Loredo & Lamb (1989) to analyse events from SN1987A, but we extend it to account for multiple interaction
channels.

We start by considering bins in time and observed energy, where the bin size $\Delta t \cdot \Delta E$ is sufficiently small that the
expected number of events per bin,

$$ N_i = \frac{d^2 N(E_i, t_i)}{dE dt} \Delta E \Delta t, $$

(A1)

is much smaller than 1. Here, $N(E, t)$ is the observed event rate as a function of time and energy as predicted by a
supernova model.

Assuming a Poisson distribution, the probability of observing 0 events in a single interaction channel in a bin around
time $t_i$ and energy $E_i$ is

$$ P_{0,\alpha} = \exp (-N_{i,\alpha}). $$

(A2)

When considering multiple interaction channels, the probability of observing 0 events is simply the product of the
probabilities of observing 0 events in every single interaction channel, i.e.

$$ P_0 = \prod_{\alpha} P_{0,\alpha} = \prod_{\alpha} \exp (-N_{i,\alpha}). $$

(A3)

In any given interaction channel the probability of observing exactly one event is

$$ P_{1,\alpha} = N_{i,\alpha} \exp (-N_{i,\alpha}) $$

(A4)

4 Throughout this appendix, we will use greek letters to refer to interaction channels.
and the total probability of observing exactly one event is

\[
P_1 = \sum_{\alpha} \left( P_{1,\alpha} \cdot \prod_{\beta \neq \alpha} P_{0,\beta} \right) \tag{A5}\]

\[
= \sum_{\alpha} \left[ N_{i,\alpha} \exp(-N_{i,\alpha}) \cdot \prod_{\beta \neq \alpha} \exp(-N_{i,\beta}) \right] \tag{A6}\]

\[
= \sum_{\alpha} \left[ N_{i,\alpha} \cdot \prod_{\beta} \exp(-N_{i,\beta}) \right] \tag{A7}\]

\[
= P_0 \cdot \sum_{\alpha} N_{i,\alpha}. \tag{A8}\]

The bin size was chosen such that the probability of observing more than one event in a bin is negligible. The likelihood of observing a set of events \( (E_i, t_i) \) with \( i = 1, \ldots, N_{\text{obs}} \) is then given by

\[
L = \prod_{i=1}^{N_{\text{obs}}} P_1(E_i, t_i) \cdot \prod_{i \neq j} P_0(E_j, t_j) \tag{A9}\]

\[
= \prod_{i=1}^{N_{\text{obs}}} \left( \sum_{\alpha} N_{i,\alpha} \right) P_0(E_i, t_i) \cdot \prod_{j \neq i} P_0(E_j, t_j) \tag{A10}\]

\[
= \prod_{i=1}^{N_{\text{obs}}} \sum_{\alpha} N_{i,\alpha} \cdot \prod_{j} P_0(E_j, t_j), \tag{A11}\]

where the products over \( j \neq i \) include only bins that do not contain an event, while products over \( j \) include all bins.

For simplicity, we consider the log-likelihood \( L = \ln L \). Using \( \ln(a \cdot b) = \ln(a) + \ln(b) \), the log-likelihood function is

\[
L = \sum_{i=1}^{N_{\text{obs}}} \ln \left( \sum_{\alpha} N_{i,\alpha} \right) + \sum_{j} \ln \left( P_0(E_j, t_j) \right), \tag{A12}\]

where the second term simplifies to

\[
\sum_{j} \ln \left( P_0(E_j, t_j) \right) = \sum_{j} \ln \left[ \prod_{\alpha} \exp(-N_{j,\alpha}) \right] \tag{A13}\]

\[
= \sum_{j} \sum_{\alpha} \ln \left[ \exp(-N_{j,\alpha}) \right] \tag{A14}\]

\[
= - \sum_{j} \sum_{\alpha} N_{j,\alpha} \tag{A15}\]

\[
= - N_{\text{exp}} \tag{A16}\]

and since we assume in this paper that the distance to the supernova is unknown, we normalize the event rate so as to reproduce the observed number of events. \( N_{\text{exp}} \) is therefore model-independent and since we only consider likelihood ratios of different models A and B, \( \Delta L = L_A - L_B \), this part cancels out. The final likelihood function we use is therefore given by

\[
L = \sum_{i=1}^{N_{\text{obs}}} \ln \left( \sum_{\alpha} N_{i,\alpha} \right). \tag{A17}\]

B. PAIRWISE MODEL COMPARISONS

This appendix contains figures showing pairwise model comparisons.

REFERENCES

2020, sn tools: Event generator for supernova burst neutrinos. https://github.com/JostMigenda/sn tools
2020, WCSim: The WCSim GEANT4 application. https://github.com/WCSim/WCSim
Figure 5. Histograms showing the distribution of $\Delta L = L_{\text{black}} - L_{\text{red}}$ for all pairs of supernova models considered here, for 100 events per data set and inverted mass ordering. The purple vertical line in each panel indicates $\Delta L = 0$.

Figure 6. Histograms showing the distribution of $\Delta L = L_{\text{black}} - L_{\text{red}}$ for all pairs of supernova models considered here, for 300 events per data set and normal mass ordering. The purple vertical line in each panel indicates $\Delta L = 0$. 
Figure 7. Histograms showing the distribution of $\Delta L = L_{\text{black}} - L_{\text{red}}$ for all pairs of supernova models considered here, for 300 events per data set and inverted mass ordering. The purple vertical line in each panel indicates $\Delta L = 0$. 
