Comparing Supernova Models Using Past and Future Neutrino Data

Jackson Olsen

University of Minnesota

October 26, 2021

In recent years, CCSN simulations have grown sophisticated enough to cover the neutrino signal for ~ 10 s in detail

Is it feasible to distinguish between these models, using either SN 1987A neutrino data or data from a future supernova?

Neutrino Emission Models

We'll consider three models, generously provided by the Garching group

- z9.6-SFHo
- s20-SFHo
- s27-LS220

Three progenitor models, two nuclear equations of state

Neutrino Emission Phases

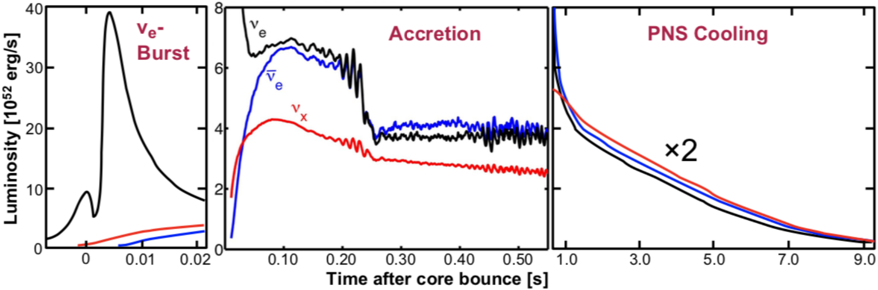


Figure from Janka (2017).

Neutrino emission occurs in three phases

- Electron neutrino breakout burst
- Accretion emission
- Proto-neutron star cooling

Predicted Neutrino Signal

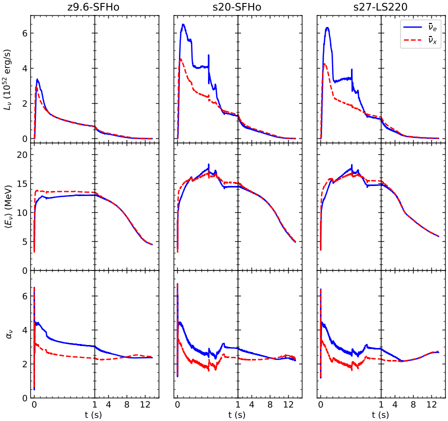
Neutrino spectrum:

$$f_{\nu_\beta}(E_\nu, t) = \frac{\left(\frac{E_\nu}{T_{\nu_\beta}}\right)^{\alpha_{\nu_\beta}} e^{-E_\nu/T_{\nu_\beta}}}{\Gamma(1+\alpha_{\nu_\beta})\, T_{\nu_\beta}}$$

with

$$\alpha_{\nu_\beta} = \frac{2\langle E_{\nu_\beta}\rangle^2 - \langle E^2_{\nu_\beta}\rangle}{\langle E^2_{\nu_\beta}\rangle - \langle E_{\nu_\beta}\rangle^2},$$

$$T_{\nu_\beta} = \frac{\langle E_{\nu_\beta}\rangle}{1+\alpha_{\nu_\beta}}$$



Past: SN 1987A

SN 1987A Data

Three detectors observed neutrinos from SN 1987A

- **Kamiokande II** (11 events)
- IMB (8 events)
- Baksan (5 events)

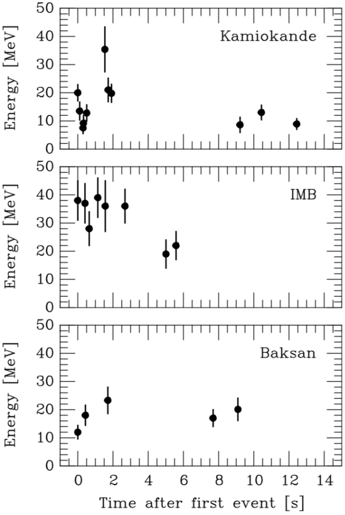


Figure from Janka (2017).

Kamiokande II

2.14 kton water Cherenkov detector

Best channel for SN neutrinos: inverse beta decay (IBD)
$\bar{\nu}_e + p \longrightarrow e^+ + n$

Only relevant species are $\bar{\nu}_e$ and $\bar{\nu}_x$

Only sensitive to accretion emission and PNS cooling

Time (s)	Energy (MeV)
0	20.0 ± 2.9
0.107	13.5 ± 3.2
0.303	7.5 ± 2.0
0.324	9.2 ± 2.7
0.507	12.8 ± 2.9
1.541	35.4 ± 8.0
1.728	21.0 ± 4.2
1.915	19.8 ± 3.2
9.219	8.6 ± 2.7
10.433	13.0 ± 2.6
12.439	8.9 ± 1.9

Predicted Event Rate

Predicted (unoscillated) flux at the detector:

$$F_{\nu_\beta}(E_\nu, t) = \frac{L_{\nu_\beta}}{4\pi d^2 \langle E_{\nu_\beta} \rangle} f_{\nu_\beta}(E_\nu, t)$$

Predicted event rate:

$$\frac{d^2 N}{dt dE} = B(E) + \int F_{\mathrm{det}}(E_e, t) \sigma(E_e) \frac{N_p \epsilon(E_e)}{\sigma_E \sqrt{2\pi}} \exp\left[-\frac{(E - E_e)^2}{2\sigma_E^2}\right] dE_e$$

where

- $B(E)$: Background rate
- N_p: Number of protons
- $\epsilon(E_e)$: Detector efficiency
- σ_E: Energy resolution

Oscillations

Incorporate oscillations with swapping fraction f:

$$F_{\mathrm{det}}(E_\nu, t) = f\, F_{\bar{\nu}_e}(E_\nu, t) + (1 - f) F_{\bar{\nu}_x}(E_\nu, t)$$

Three oscillations scenarios

- No oscillations (NO): $f = 1.0$
- MSW with normal mass hierarchy (NH): $f = 0.681$
- MSW with inverted mass hierarchy (IH): $f = 0.022$

Model Parameters

Time offset

- No way to determine when exactly the ν signal reaches the detector relative to the first event
- Therefore, syncing the model time with the time of the first detection is difficult

$$t = t_{\mathrm{det}} + t_{\mathrm{off}}$$

- Treat t_{off} as a model parameter

Supernova Distance

- Distance to SN 1987A has been measured to be $d = 51.4 \pm 1.2$ kpc
- We will also treat this as a parameter, with a restrictive prior

Bayesian Statistics

Bayesian Model Comparison

Bayes' Theorem:

$$P(M_\alpha|D) = \frac{P(D|M_\alpha)P(M_\alpha)}{P(D)}$$

- $P(M_\alpha)$ is the prior information about the models
- $P(D|M_\alpha)$ is the Bayesian evidence
- $P(D)$ is a normalization constant

Bayesian Model Comparison

Bayes' Theorem:

$$P(M_\alpha|D) = \frac{P(D|M_\alpha)P(M_\alpha)}{P(D)}$$

- $P(M_\alpha)$ is a uniform prior
- $P(D|M_\alpha) = \int d\vec{\theta}\,\mathcal{L}_\alpha(\vec{\theta})P(\vec{\theta}|M_\alpha)$ is the Bayesian evidence, $P(\vec{\theta}|M_\alpha) = P(t_{\rm off}|M_\alpha)P(d|M_\alpha)$ is the parameter prior
  - We take $P(t_{\rm off}|M_\alpha)$ to be uniform
  - $P(d|M_\alpha)$ is Gaussian, with $d = 51.4 \pm 1.2$ kpc
- $P(D) = \sum_\beta P(D|M_\beta)P(M_\beta)$ is the normalization constant

Bayes Factor

$B_{\alpha\beta}$	$\log B_{\alpha\beta}$	Strength of Evidence
1-3	0-1	Negligible
3-20	1-3	Positive
20-150	3-5	Strong
> 150	> 5	Very strong

To compare models, we use the Bayes factor

$$B_{\alpha\beta} = \frac{P(D|M_\alpha)}{P(D|M_\beta)} = \frac{P(M_\alpha|D)}{P(M_\beta|D)}\frac{P(M_\beta)}{P(M_\alpha)} = \frac{P(M_\alpha|D)}{P(M_\beta|D)}$$

Results of Bayesian Analysis

Model	d (kpc)	t_{off} (s)	$\langle N \rangle$	$P(M_\alpha \mid D)$
z9.6-SFHo (NO)	51.39	0.048	6.81	0.2807
z9.6-SFHo (NH)	51.39	0.036	7.17	0.2684
z9.6-SFHo (IH)	51.39	0.024	7.92	0.2037
s20-SFHo (NO)	51.45	0.054	19.5	0.0058
s20-SFHo (NH)	51.45	0.054	19.4	0.0060
s20-SFHo (IH)	51.45	0.026	19.3	0.0043
s27-LS220 (NO)	51.43	0.051	15.1	0.0913
s27-LS220 (NH)	51.43	0.051	15.1	0.0875
s27-LS220 (IH)	51.43	0.033	15.0	0.0523

$\frac{P(\mathrm{z9.6-SFHo(NO)}|D)}{P(\mathrm{s20-SFHo(NO)}|D)} \approx 48.4 \Rightarrow$ Strongly favored

$\frac{P(\mathrm{z9.6-SFHo(NO)}|D)}{P(\mathrm{s27-LS220(NO)}|D)} \approx 3.1 \Rightarrow$ Positively favored

$\frac{P(\mathrm{z9.6-SFHo(NO)}|D)}{P(\mathrm{z9.6-SFHo(IH)}|D)} \approx 1.4 \Rightarrow$ No preference among oscillation scenarios

SN 1987A Conclusions

We can draw limited conclusions about SN 1987A based on the data

Data seems to favor a model with a shorter accretion period

No insights about neutrino oscillations

Future Supernova

Prospects for Future SN Data

$\sim 1 - 3$ Galactic CCSN expected per century

Better detectors running now than in 1987, such as Super-K

We should see many more neutrinos from the next Galactic CCSN

Will it be enough to distinguish models?

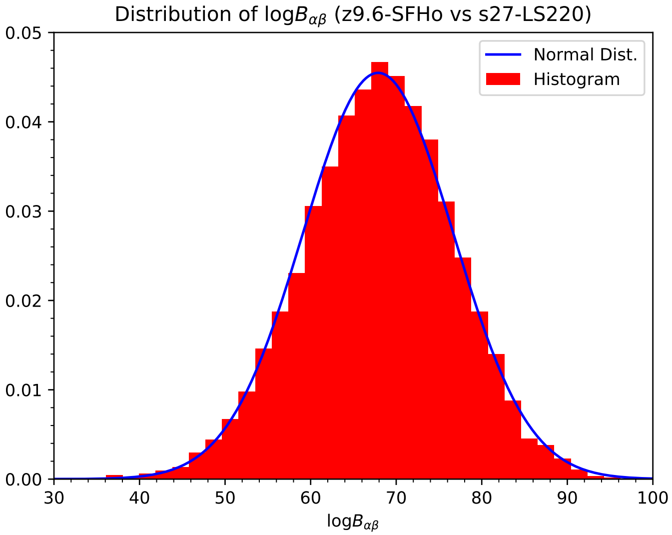
Example

Let's imagine we have an ideal water Cherenkov detector, $\mathcal{D}_I$

A SN occurs at 50 kpc, following z9.6-SFHo, with no oscillations

To simulate this, we can sample many possible datasets from the distribution for z9.6-SFHo
$\implies$ Compute the distribution of the Bayes factors

Distribution of log $B_{\alpha\beta}$



Example: Results

Comp. Model	$\log B_{\alpha\beta}$ ($t_{\rm off}$ unknown)	$\log B_{\alpha\beta}$ ($t_{\rm off}$ known)
s20-SFHo	113.7 ± 11.3	114.1 ± 11.3
s27-LS220	67.9 ± 8.8	68.2 ± 8.8

Take $\log B_{\alpha\beta} > 5.0$ to be our distinguishability threshold

We can distinguish models at a 95% confidence level with $\langle \log B_{\alpha\beta} \rangle - z\sigma[\log B_{\alpha\beta}] > 5.0$, with $z = 1.645$

$\Longrightarrow$ Both models can be distinguished from z9.6-SFHo at 50 kpc

We computed assuming both unknown and known time offsets, and find negligible difference

$\Longrightarrow$ For simplicity, take the time offset to be known

Super-Kamiokande

Large (32 kton) water Cherenkov detector

After spallation cut, background is negligible

Conservative efficiency: ~ 0.75

Imperfect energy resolution

Addition of Gd should allow for tagging of IBD events

Super-K Results

Model	z9.6-SFHo	s20-SFHo	s27-LS220
z9.6-SFHo	–	83.95 ± 9.66 75.8 ± 9.51 61.99 ± 9.0	50.34 ± 7.55 44.17 ± 7.39 34.68 ± 6.91
s20-SFHo	149.16 ± 22.93 124.56 ± 20.09 93.13 ± 16.63	–	22.9 ± 8.21 24.04 ± 8.49 26.88 ± 9.11
s27-LS220	88.6 ± 17.64 70.78 ± 15.0 50.02 ± 11.98	15.54 ± 4.6 15.98 ± 4.61 17.24 ± 4.69	–

Using Super-K rather than the ideal detector reduces $\log B_{\alpha\beta}$ values

All models are distinguishable at 50 kpc, regardless of the oscillation scenario

Conclusions

Applying Bayesian techniques, we find that SN 1987A ν data favors a short accretion period

These techniques could also be used in the event of a future Galactic CCSN to distinguish between state-of-the-art ν emission models

The three models could be distinguished up to 50 kpc, effectively anywhere in the Galaxy

Supported by US Department of Energy grant DE-FG02-87ER40328; calculations were carried out at the Minnesota Supercomputing Institute

Bayesian Parameter Estimation

Bayes' Theorem:

$$P(\vec{\theta}|D, M_\alpha) = \frac{P(D|\vec{\theta}, M_\alpha)P(\vec{\theta}|M_\alpha)}{P(D|M_\alpha)}$$

- $\vec{\theta}$ are the model parameters
- $P(\vec{\theta}|M_\alpha)$ is the prior
- $P(D|\vec{\theta}, M_\alpha)$ is the likelihood function
- $P(D|M_\alpha)$ is a normalization constant

Bayesian Parameter Estimation

Bayes' Theorem:

$$P(\vec{\theta}|D, M_\alpha) = \frac{P(D|\vec{\theta}, M_\alpha)P(\vec{\theta}|M_\alpha)}{P(D|M_\alpha)}$$

- In our case, $\vec{\theta} = (t_{\text{off}}, d)$
- $P(D|\vec{\theta}, M_\alpha) = \frac{\langle N \rangle^N e^{-\langle N \rangle}}{N!} \prod_{i=1}^{N} p(E_i, t_i|\vec{\theta}, M_\alpha)$ is the likelihood
- $P(\vec{\theta}|M_\alpha) = P(t_{\text{off}}|M_\alpha)P(d|M_\alpha)$ is the prior
  - We take $P(t_{\text{off}}|M_\alpha)$ to be uniform
  - $P(d|M_\alpha)$ is Gaussian, with $d = 51.4 \pm 1.2$ kpc
- $P(D|M_\alpha) = \int d\vec{\theta}\, P(D|\vec{\theta}, M_\alpha)P(\vec{\theta}|M_\alpha)$

Bayesian Model Comparison

Bayes' Theorem:

$$P(M_\alpha|D) = \frac{P(D|M_\alpha)P(M_\alpha)}{P(D)}$$

- $P(D|M_\alpha) = \int d\vec{\theta}\, P(D|\vec{\theta}, M_\alpha)P(\vec{\theta}|M_\alpha)$ is the Bayesian evidence
- $P(M_\alpha)$ is a (uniform) prior
- $P(D) = \sum_\beta P(D|M_\beta)P(M_\beta)$ is a normalization constant

Goodness-Of-Fit Intuition

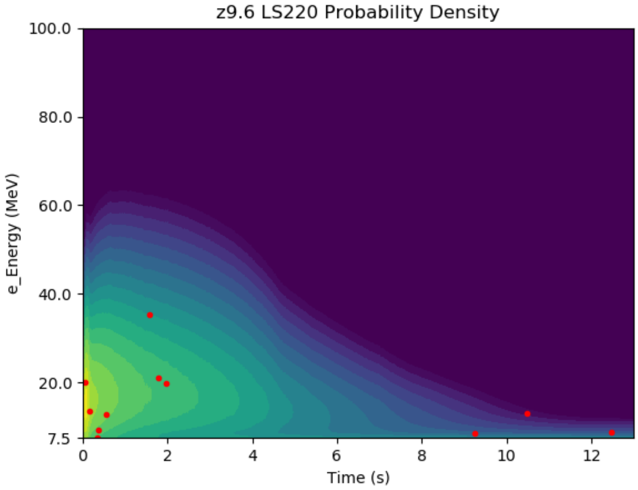


Figure 1: Plot of probability density and Kam II data.

Hypothesis Testing

- Specify null and alternative hypotheses H_0 and H_A
- Define a test statistic $\lambda \equiv \prod_i p(t_i, E_i)$
- Set threshold for and calculate p-value:

$$\text{p-value} \equiv \int_0^{\lambda^*} d\lambda \, \text{pdf}(\lambda)$$

- The p-value is the probability of observing a λ as likely or less likely than λ^*, assuming the null hypothesis is true
- If p-value $<$ threshold $= 0.050$, reject H_0 in favor of H_A

Normalized SN 1987A Comparison

Model	K	t_{off} (s)	$\langle N \rangle$	p-value	$p(M_i \| D)$
z9.6-SFHo (NO)	1.32	0.048	8.90	0.16	0.2638
z9.6-SFHo (NH)	1.26	0.036	8.96	0.20	0.2223
z9.6-SFHo (IH)	1.15	0.024	9.06	0.25	0.1359
s20-SFHo (NO)	0.46	0.054	9.14	0.19	0.0400
s20-SFHo (NH)	0.47	0.054	9.29	0.25	0.0385
s20-SFHo (IH)	0.47	0.026	9.21	0.39	0.0245
s27-LS220 (NO)	0.60	0.051	9.18	0.12	0.1108
s27-LS220 (NH)	0.60	0.051	9.16	0.17	0.1038
s27-LS220 (IH)	0.61	0.033	9.27	0.25	0.0603

Estimating the $\log B_{\alpha\beta}$ Distribution

Instead of Monte Carlo simulations, let's do some (mildly) clever mathematics

$$\langle \log B_{\alpha\beta} \rangle = \sum_N \mathrm{pois}(N; M_\alpha) \int d\vec{x}\; P(\vec{x}|M_\alpha) \log \frac{P(\vec{x}|M_\alpha)}{P(\vec{x}|M_\beta)}$$

After some tedium,

$$\langle \log B_{\alpha\beta} \rangle = \langle N \rangle_\alpha \log \tfrac{\langle N \rangle_\alpha}{\langle N \rangle_\beta} - \Delta_{\alpha\beta} + \langle N \rangle_\alpha \langle \log \tfrac{p(t,E|M_\alpha)}{p(t,E|M_\beta)} \rangle,$$

with $\Delta_{\alpha\beta} \equiv \langle N \rangle_\alpha - \langle N \rangle_\beta$

Can calculate $\sigma[\log B_{\alpha\beta}]$ similarly

Model Distributions

Previously, we included $\langle N \rangle$ predictions in our analysis

Let's now generalize our models to only the probability distributions $p(t, E|M_\alpha)$

Likelihood becomes $P(D|M_\alpha) = \prod_{i=1}^{N} p(E_i, t_i|M_\alpha)$

$\Longrightarrow$ All the model distributions are distinguishable at 10 kpc