
Hierarchical Simulation-Based Inference of Supernova Power Sources and their Physical Properties

Edgar P. Vidal

Department of Physics and Astronomy
Tufts University
Medford, MA 02155, USA
edgar.vidal@tufts.edu

Alexander T. Gagliano

Institute for AI and Fundamental Interactions
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
gaglian2@mit.edu

Carolina Cuesta-Lazaro

Flatiron Institute
Institute for Advanced Studies
ccuesta-lazaro@flatironinstitute.org

Abstract

Time domain surveys such as the Vera C. Rubin Observatory are projected to annually discover millions of astronomical transients. This and complementary programs demand fast, automated methods to constrain the physical properties of the most interesting objects for spectroscopic follow up. Traditional approaches to likelihood-based inference are computationally expensive and ignore the multi-component energy sources powering astrophysical phenomena. In this work, we present a hierarchical simulation-based inference model for multi-band light curves that 1) identifies the energy sources powering an event of interest, 2) infers the physical properties of each subclass, and 3) separates physical anomalies in the learned embedding space. Our architecture consists of a transformer-based light curve summarizer coupled to a flow-matching regression module and a categorical classifier for the physical components. We train and test our model on $\sim 150k$ synthetic light curves generated with MOSFiT. Our network achieves a 90% classification accuracy at identifying energy sources, yields well-calibrated posteriors for all active components, and detects rare anomalies such as tidal disruption events (TDEs) through the learned latent space. This work demonstrates a scalable joint framework for population studies of known transients and the discovery of novel populations in the era of Rubin.

1 Introduction

A central theme of time-domain astrophysics lies in recovering the physical properties of an astrophysical system from photometric time-series observations alone. This is particularly relevant for the explosions of stars as supernovae (SNe). While the majority of observed supernovae are powered by reprocessing of high-energy photons produced by the radioactive decay of ^{56}Ni synthesized in the explosion, the observational taxonomy has expanded to include shock-powered SNe whose emission is dominated by interaction between the expanding ejecta and surrounding circumstellar material [1] and superluminous SNe (SLSNe) powered by a compact central engine [2].

While the physical signatures of these energy sources imprint themselves on SN observations, significant degeneracies exist at the level of optical photometry; worse, supernova emission can include contributions from multiple physical mechanisms simultaneously [3]. Existing inference techniques [4, 5] typically constrain the physical properties of a single energy source whose energy

input is assumed to dominate during specific phases of explosion (e.g., during the photospheric or nebular phase).

A complementary objective in time-domain astrophysics is the automated detection of anomalous events. Existing approaches flag supernovae whose light curve deviates from an empirical model [6] or from those observed in the broader population [7]. The detection of SNe whose emission deviates from expected *physical* models, however, is critical to constrain the diversity of emission mechanisms and for discovering entirely novel phenomena in upcoming large-scale photometric surveys such as the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 8). In this work, we simultaneously address both objectives with a hierarchical simulation-based inference framework capable of constraining both the primary energy sources associated with supernova emission *and* the physical properties of those sources.

To generate a realistically diverse sample of physical phenomena, we generate synthetic *ugrizy* LSST light curves from the Modular Open Source Fitter for Transients [MOSFiT; 9]. MOSFiT couples semi-analytic physical models for an event’s spectral energy distribution to modules that modulate this emission, as with diffusion through explosion ejecta or a viscous delay in the accretion of matter onto a black hole. We consider three energy sources in MOSFiT: the radioactive decay of ^{56}Ni [10], interaction with circumstellar material [CSM; 11–13], and spin-down by a central magnetar [14]. We define seven total SN models consisting of all combinations of these three components, and list our priors for the associated physical parameters in Table 1 in the Appendix.

We generate 20,000 light curves from each SN model, with 100 observations in each filter uniformly sampled in time from explosion to 200 days following explosion. We further split our dataset into fractions of 80%/10%/10% for training, testing, and validation, respectively.

2 Hierarchical Simulation-Based Inference

Our goal is to model the joint posterior distribution, $p(\mathcal{S}, \theta|x)$ over active energy source combinations \mathcal{S} and their associated physical properties, parametrized as θ , given the observed light curve x . We learn optimal summary statistics, $x_s = f_\psi(x)$, from the simulated samples, where f_ψ is the shared summarizer trained jointly with the neural posterior model.

As in [15], we model the joint posterior hierarchically by learning the conditionals $p(\mathcal{S}, \theta|x_s) = p(\mathcal{S}|x_s)p(\theta|\mathcal{S}, x_s)$, that we refer to as the source posterior, $p(\mathcal{S}|x_s)$, and the parameter posterior $p(\theta|\mathcal{S}, x_s)$. The light curve summarizer is shared between the two components.

The light curve shared summarizer. To address the sparsity and irregular sampling of supernova light curves, we encode our MOSFiT light curves using the multi-band transformer model presented in [16]. Observations are provided as input in the format $\{t, m\}$ after normalization, where t represents the number of days from first observation and m represents the event’s apparent magnitude. We project the times t using sinusoidal positional encodings, concatenate all brightness measurements m across all bands and add a one-hot-encoded vector representing the photometric passband of the observation (one of LSST-*ugrizy*). We use multi-head attention with 2 attention heads, and aggregate the outputs using a learnable query vector to produce a single weighted representation, $x_s \in \mathbb{R}^{d_s}$. The dimensionality of the summary statistics is set to $d_s = 64$.

Modeling the source posterior. We introduce a categorical network¹ to model the posterior over potential supernova energy sources, $p(\mathcal{S}|x_s)$, given a light curve representation x_s output by the summarizer. Let \mathcal{C} be the set of source components, and let $\mathcal{S} \subseteq \mathcal{C}$ denote the composite class (an element of the power set of \mathcal{C}) describing all active sources. We model a categorical distribution over all non-empty subsets, i.e., over $2^{|\mathcal{C}|} - 1$ classes, which allows for multiple sources components while keeping a single target per transient. Training is performed using the standard negative log likelihood loss function, $\mathcal{L}_{\text{source}}(\psi) = -\log q_\psi(\mathcal{S}|x_s)$, where $q(\mathcal{S}|x_s)$ denotes the model distribution approximating the true posterior $p(\mathcal{S}|x_s)$, and ψ are the weights of the network. The categorical network has 4 hidden layers with 128 hidden dimensions, uses sigmoid activations, and a final softmax output layer produces the predicted probabilities for each combination of source components.

¹[15] introduce a mixture of multivariate binary Grassmann distributions to predict model components in a similar hierarchical model; we find our training is more stable with the categorical network.

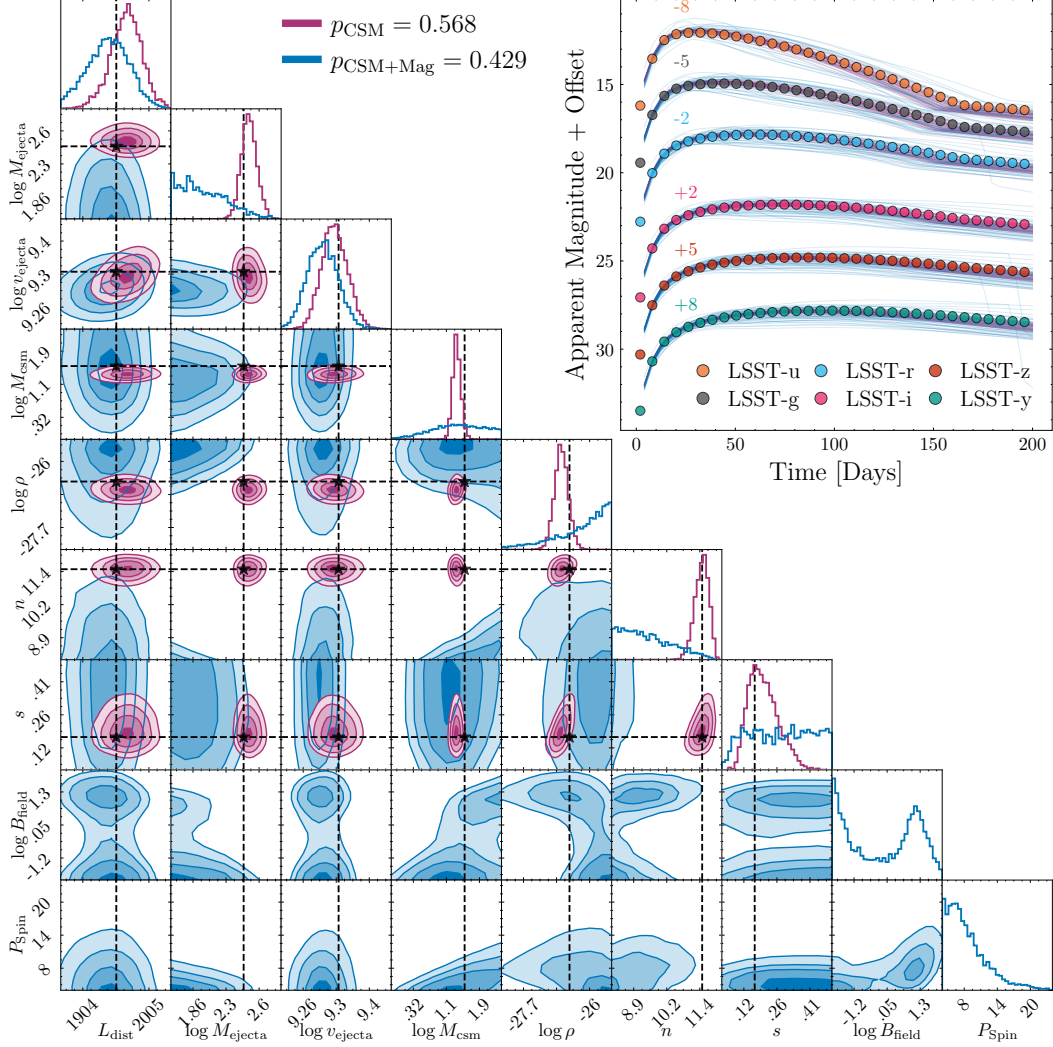


Figure 1: Example joint inference of power sources and their physical properties for a supernova light curve. On the left, we show the posterior samples for the two most likely source combinations powering the light curve, CSM (pink) and CSM+Mag (blue). We show that the presence of the magnetar engine can significantly alter the posteriors for the CSM parameters. In the upper right corner, we show the observed light curve and overlay the light curves drawn from the posteriors of the two power sources.

Modeling the parameter posterior. We model the posterior over physical parameters, $p(\theta|\mathcal{S}, x_s)$, using flow matching neural posterior estimation (FMPE) [17]. Flow matching learns to transform samples from a base distribution, here a normal distribution, to the target posterior by predicting velocity fields. In [17], the authors demonstrated that flow matching can enhance the flexibility of the more commonly used posterior estimation models, normalizing flows, whilst improving accuracy in data scarce scenarios.

In this work, we introduce a transformer as the velocity prediction network, conditioned on both the light curve summaries statistics x_s and the energy source combination \mathcal{S} . To handle varying parameter dimensionalities across different energy sources, we mask unused parameters in the transformer. During training, we condition and mask based on the true energy sources; during inference, we use outputs from the source posterior network. We train both components jointly: $\mathcal{L} = \mathcal{L}_\theta(\phi) + \lambda \mathcal{L}_{\text{source}}(\psi)$, where $\mathcal{L}_\theta(\phi)$ is the FMPE loss, and ϕ are the weights of the transformer velocity model. When $\lambda = 1$, it can be shown that this loss function minimizes the expected Kullback-Leibler divergence between the true joint posterior $p(\mathcal{S}, \theta|x_s)$ and the approximate posterior

$q_\phi(\mathcal{S}|x_s)q_\psi(\theta|x_s, \mathcal{S})$, under the assumption of regularity on v_t . Empirically, we find that $\lambda = 0.83$ gives the best approximation to the true posterior, as shown in Section 3 below.

3 Results

3.1 Sampling the joint posterior

In Figure 1, we show the joint inference of model components and their physical parameters for a light curve with degeneracy between energy sources predicted by the categorical network. The corner plot contains posterior samples of the two source combinations with highest probability: CSM and CSM with a magnetar engine, having probabilities $p_{\text{CSM}} = 0.568$ and $p_{\text{CSM+Mag}} = 0.439$, respectively. The presence of the magnetar engine can shift the posteriors of the CSM parameters ρ and n substantially, impacting the inferred properties of both the pre-explosion mass-loss history and the nature of the progenitor star (from a luminous blue variable or Wolf-Rayet-like progenitor with $7 < n < 10$ in the CSM+Mag model to a red supergiant-like progenitor with $n = 12$ in the CSM-only model; Matzner and McKee 18).

In the upper right corner of Figure 1, we show the light curve together with posterior resimulations with and without the magnetar engine. Whilst the CSM model fits the data well, we find that the magnetar model can also produce posterior samples that reproduce the observations. There remains active debate surrounding the dominant energy sources of superluminous supernovae; events with smooth photometric evolution and a lack of narrow spectral features have been argued to be powered by a magnetar engine [19], while the photometric diversity of the class may be more consistent with CSM interaction differing in e.g., geometry [20]. While we caution that MOSFiT is unable to simulate complex aspherical CSM morphologies or interaction beneath the explosion photosphere, the degeneracies in Figure 1 suggest that the presence of a magnetar engine may be hidden for months (as in, e.g., SN 2020wnt; Tinyanont et al. 21), particularly when inferred from sparse and noisy light curves for events at high redshift.

We show a quantitative assessment of the learned posteriors in Appendix C. Figure 3 shows the results of a TARP coverage test [22] demonstrating that all posteriors for all sources are well calibrated.

3.2 Detection of physical anomalies

Next, we explore our model’s capacity to identify novel time-domain phenomena. We simulate 10,000 Tidal Disruption Events (TDEs) in MOSFiT [23], events which are powered by the accretion of tidally-stripped stellar material onto a supermassive black hole, as our anomalous class. Our priors for the model are given in Table 1 in the Appendix.

We use the trained hierarchical SBI model to produce fixed-length representations of both our SN and TDE light curves, and further project them into a two-dimensional space for visualization with t-distributed Stochastic Neighbor Embedding [tSNE; 24]. Our results are shown in Figure 2.

We find strong separation between individual model components and physically meaningful overlap between models with multiple energy sources. Most overlap occurs between a subset of CSM+Mag+Ni, CSM+Mag, and Mag-only light curves, suggesting that for these events the contributions from the central magnetar dominates the light curve evolution. We also

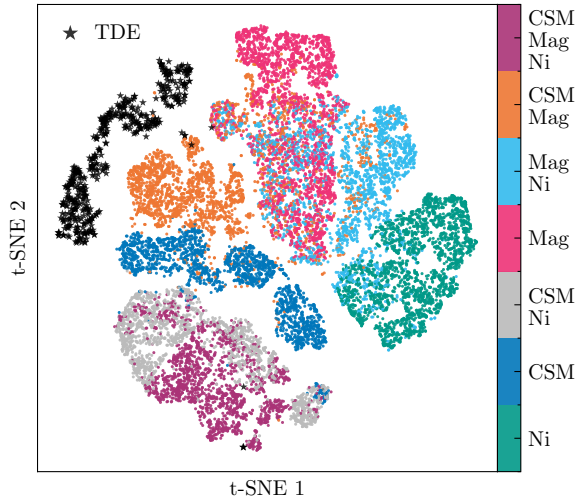


Figure 2: tSNE embedding of the latent features of all simulated SN light curves, colored by their power energy sources. TDE light curves are shown as black stars.

find a clean separation between most simulated TDEs and SNe despite their photometric similarities; a few events are embedded in similar positions as the CSM+Mag and Mag light curves, which may reflect the similarities in the MOSFiT prescriptions for black hole-accretion power (for TDEs) and neutron-star spindown power (for magnetars). These results suggest that our hierarchical SBI model may be able to identify physical analogs to novel populations of transients from LSST light curves.

4 Conclusions

We have shown that a single hierarchical SBI model can simultaneously identify the energy sources and physical parameters associated with optical time-domain phenomena. This work lays the foundation for scalable population-level studies of transients discovered by existing and upcoming time-domain surveys. In future work, we will re-train our model with more realistic LSST light curves using LSST Wide-Fast-Deep simulations.² We will also investigate the parameter posteriors for partial-phase light curves, which will allow astronomers to rapidly prioritize the most physically-interesting phenomena for real-time follow-up observations. Finally, while Figure 2 shows that the hierarchical model can be used to identify anomalous phenomena, additional work is needed to formalize this approach using density-estimation techniques applied to the learned latent space.

5 Acknowledgments and Disclosure of Funding

This work is supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, <http://iaifi.org/>). The research reported in this paper was done on the Tufts University High Performance Computing Cluster (<https://it.tufts.edu/high-performance-computing>).

References

- [1] Nathan Smith. Interacting Supernovae: Types II_n and Ib_n. In Athem W. Alsabti and Paul Murdin, editors, *Handbook of Supernovae*, page 403. 2017. doi: 10.1007/978-3-319-21846-5_38.
- [2] Takashi J. Moriya, Elena I. Sorokina, and Roger A. Chevalier. Superluminous Supernovae. *Space Sci. Rev.*, 214(2):59, March 2018. doi: 10.1007/s11214-018-0493-6.
- [3] Takashi J. Moriya and Keiichi Maeda. A Dip after the Early Emission of Superluminous Supernovae: A Signature of Shock Breakout within Dense Circumstellar Media. *ApJ*, 756(1): L22, September 2012. doi: 10.1088/2041-8205/756/1/L22.
- [4] V. Ashley Villar. Amortized Bayesian Inference for Supernovae in the Era of the Vera Rubin Observatory Using Normalizing Flows. *arXiv e-prints*, art. arXiv:2211.04480, November 2022. doi: 10.48550/arXiv.2211.04480.
- [5] Braden Garretson, Dan Milisavljevic, Bhagya Subrayan, and Danielle Dickinson. Rapid Supernova Model Inference using Amortized Posterior Estimation. In *American Astronomical Society Meeting Abstracts #243*, volume 243 of *American Astronomical Society Meeting Abstracts*, page 260.38, February 2024.
- [6] Daniel Muthukrishna, Kaisey S. Mandel, Michelle Lochner, Sara Webb, and Gautham Narayan. Real-time detection of anomalies in large-scale transient surveys. *MNRAS*, 517(1):393–419, November 2022. doi: 10.1093/mnras/stac2582.
- [7] Rithwik Gupta, Daniel Muthukrishna, and Michelle Lochner. A classifier-based approach to multiclass anomaly detection for astronomical transients. *RAS Techniques and Instruments*, 4: rzae054, January 2025. doi: 10.1093/rasti/rzae054.
- [8] Željko Ivezić, Steven M. Kahn, J. Anthony Tyson, Bob Abel, Emily Acosta, Robyn Allsman, David Alonso, Yusra AlSayyad, Scott F. Anderson, John Andrew, James Roger P. Angel, George Z. Angeli, Reza Ansari, Pierre Antilogus, Constanza Araujo, Robert Armstrong, Kirk T. Arndt, Pierre Astier, Éric Aubourg, Nicole Auza, Tim S. Axelrod, Deborah J. Bard, Jeff D. Barr,

²<https://community.lsst.org/t/release-of-v3-4-simulations/8548/7>

- Aurelian Barrau, James G. Bartlett, Amanda E. Bauer, Brian J. Bauman, Sylvain Baumont, Ellen Bechtol, Keith Bechtol, Andrew C. Becker, Jacek Becla, Cristina Beldica, Steve Bellavia, Federica B. Bianco, Rahul Biswas, Guillaume Blanc, Jonathan Blazek, Roger D. Blandford, Josh S. Bloom, Joanne Bogart, Tim W. Bond, Michael T. Booth, Anders W. Borgland, Kirk Borne, James F. Bosch, Dominique Boutigny, Craig A. Brackett, Andrew Bradshaw, William Nielsen Brandt, Michael E. Brown, James S. Bullock, Patricia Burchat, David L. Burke, Gianpietro Cagnoli, Daniel Calabrese, Shawn Callahan, Alice L. Callen, Jeffrey L. Carlin, Erin L. Carlson, Srinivasan Chandrasekharan, Glenaver Charles-Emerson, Steve Chesley, Elliott C. Cheu, Hsin-Fang Chiang, James Chiang, Carol Chirino, Derek Chow, David R. Ciardi, Charles F. Claver, Johann Cohen-Tanugi, Joseph J. Cockrum, Rebecca Coles, Andrew J. Connolly, Kem H. Cook, Asantha Cooray, Kevin R. Covey, Chris Cribbs, Wei Cui, Roc Cutri, Philip N. Daly, Scott F. Daniel, Felipe Daruich, Guillaume Daubard, Greg Daues, William Dawson, Francisco Delgado, Alfred Dellapenna, Robert de Peyster, Miguel de Val-Borro, Seth W. Digel, Peter Doherty, Richard Dubois, Gregory P. Dubois-Felsmann, Josef Durech, Frossie Economou, Tim Eifler, Michael Eracleous, Benjamin L. Emmons, Angelo Fausti Neto, Henry Ferguson, Enrique Figueroa, Merlin Fisher-Levine, Warren Focke, Michael D. Foss, James Frank, Michael D. Freeman, Emmanuel Gangler, Eric Gawiser, John C. Geary, Perry Gee, Marla Geha, Charles J. B. Gessner, Robert R. Gibson, D. Kirk Gilmore, Thomas Glanzman, William Glick, Tatiana Goldina, Daniel A. Goldstein, Iain Goodenow, Melissa L. Graham, William J. Gressler, Philippe Gris, Leanne P. Guy, Augustin Guyonnet, Gunther Haller, Ron Harris, Patrick A. Hascall, Justine Haupt, Fabio Hernandez, Sven Herrmann, Edward Hileman, Joshua Hoblitt, John A. Hodgson, Craig Hogan, James D. Howard, Dajun Huang, Michael E. Huffer, Patrick Ingraham, Walter R. Innes, Suzanne H. Jacoby, Bhuvnesh Jain, Fabrice Jammes, M. James Jee, Tim Jenness, Garrett Jernigan, Darko Jevremović, Kenneth Johns, Anthony S. Johnson, Margaret W. G. Johnson, R. Lynne Jones, Claire Juramy-Gilles, Mario Jurić, Jason S. Kalirai, Nitya J. Kallivayalil, Bryce Kalmbach, Jeffrey P. Kantor, Pierre Karst, Mansi M. Kasliwal, Heather Kelly, Richard Kessler, Veronica Kinnison, David Kirkby, Lloyd Knox, Ivan V. Kotov, Victor L. Krabbendam, K. Simon Krughoff, Petr Kubánek, John Kuczewski, Shri Kulkarni, John Ku, Nadine R. Kurita, Craig S. Lage, Ron Lambert, Travis Lange, J. Brian Langton, Laurent Le Guillou, Deborah Levine, Ming Liang, Kian-Tat Lim, Chris J. Lintott, Kevin E. Long, Margaux Lopez, Paul J. Lotz, Robert H. Lupton, Nate B. Lust, Lauren A. MacArthur, Ashish Mahabal, Rachel Mandelbaum, Thomas W. Markiewicz, Darren S. Marsh, Philip J. Marshall, Stuart Marshall, Morgan May, Robert McKercher, Michelle McQueen, Joshua Meyers, Myriam Migliore, Michelle Miller, and David J. Mills. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *ApJ*, 873(2):111, March 2019. doi: 10.3847/1538-4357/ab042c.
- [9] James Guillochon, Matt Nicholl, V. Ashley Villar, Brenna Mockler, Gautham Narayan, Kaisey S. Mandel, Edo Berger, and Peter K. G. Williams. MOSFiT: Modular Open Source Fitter for Transients. *ApJS*, 236(1):6, May 2018. doi: 10.3847/1538-4365/aab761.
- [10] D. K. Nadyozhin. The Properties of NI CO Fe Decay. *ApJS*, 92:527, June 1994. doi: 10.1086/192008.
- [11] E. Chatzopoulos, J. Craig Wheeler, J. Vinko, Z. L. Horvath, and A. Nagy. Analytical Light Curve Models of Superluminous Supernovae: χ^2 -minimization of Parameter Fits. *ApJ*, 773(1): 76, August 2013. doi: 10.1088/0004-637X/773/1/76.
- [12] V. Ashley Villar, Edo Berger, Brian D. Metzger, and James Guillochon. Theoretical Models of Optical Transients. I. A Broad Exploration of the Duration-Luminosity Phase Space. *ApJ*, 849 (1):70, November 2017. doi: 10.3847/1538-4357/aa8fcb.
- [13] Brighten Jiang, Shuai Jiang, and V. Ashley Villar. Extended Self-similar Solution for Circumstellar Material-supernova Ejecta Interaction. *Research Notes of the American Astronomical Society*, 4(1):16, January 2020. doi: 10.3847/2515-5172/ab7128.
- [14] Matt Nicholl, James Guillochon, and Edo Berger. The Magnetar Model for Type I Superluminous Supernovae. I. Bayesian Analysis of the Full Multicolor Light-curve Sample with MOSFiT. *ApJ*, 850(1):55, November 2017. doi: 10.3847/1538-4357/aa9334.
- [15] Cornelius Schröder and Jakob H. Macke. Simultaneous identification of models and parameters of scientific simulators. *arXiv e-prints*, art. arXiv:2305.15174, May 2023. doi: 10.48550/arXiv.2305.15174.

- [16] Gemma Zhang, Thomas Helfer, Alexander T. Gagliano, Siddharth Mishra-Sharma, and V. Ashley Villar. Maven: a multimodal foundation model for supernova science. *Machine Learning: Science and Technology*, 5(4):045069, December 2024. doi: 10.1088/2632-2153/ad990d.
- [17] Jonas Bernhard Wildberger, Maximilian Dax, Simon Buchholz, Stephen R. Green, Jakob Macke, and Bernhard Schölkopf. Flow Matching for Scalable Simulation-Based Inference. In *Machine Learning for Astrophysics*, page 34, July 2023. doi: 10.48550/arXiv.2305.17161.
- [18] Christopher D. Matzner and Christopher F. McKee. The Expulsion of Stellar Envelopes in Core-Collapse Supernovae. *ApJ*, 510(1):379–403, January 1999. doi: 10.1086/306571.
- [19] Sebastian Gomez, Matt Nicholl, Edo Berger, Peter K. Blanchard, V. Ashley Villar, Sofia Rest, Griffin Hosseinzadeh, Aysha Aamer, Yukta Ajay, Wasundara Athukoralalage, David C. Coulter, Tarraneh Eftekhari, Achille Fiore, Noah Franz, Ori Fox, Alexander Gagliano, Daichi Hiramatsu, D. Andrew Howell, Brian Hsu, Mitchell Karmen, Matthew R. Siebert, Réka Könyves-Tóth, Harsh Kumar, Curtis McCully, Craig Pellegrino, Justin Pierel, Armin Rest, and Qinan Wang. The Type I superluminous supernova catalogue I: light-curve properties, models, and catalogue description. *MNRAS*, 535(1):471–515, November 2024. doi: 10.1093/mnras/stae2270.
- [20] P. J. Pessi, R. Lunnan, J. Sollerman, S. Schulze, A. Gkini, A. Gangopadhyay, L. Yan, A. Gal-Yam, D. A. Perley, T. W. Chen, K. R. Hinds, S. J. Brennan, Y. Hu, A. Singh, I. Andreoni, D. O. Cook, C. Fremling, A. Y. Q. Ho, Y. Sharma, S. van Velzen, T. Kangas, A. Wold, E. C. Bellm, J. S. Bloom, M. J. Graham, M. M. Kasliwal, S. R. Kulkarni, R. Riddle, and B. Rusholme. Sample of hydrogen-rich superluminous supernovae from the Zwicky Transient Facility. *A&A*, 695:A142, March 2025. doi: 10.1051/0004-6361/202452014.
- [21] Samaporn Tinyanont, Stan E. Woosley, Kirsty Taggart, Ryan J. Foley, Lin Yan, Ragnhild Lunnan, Kyle W. Davis, Charles D. Kilpatrick, Matthew R. Siebert, Steve Schulze, Chris Ashall, Ting-Wan Chen, Kishalay De, Georgios Dimitriadis, Dillon Z. Dong, Christoffer Fremling, Alexander Gagliano, Saurabh W. Jha, David O. Jones, Mansi M. Kasliwal, Hao-Yu Miao, Yen-Chen Pan, Daniel A. Perley, Vikram Ravi, César Rojas-Bravo, Itai Sfaradi, Jesper Sollerman, Vanessa Alarcon, Rodrigo Angulo, Karoli E. Clever, Payton Crawford, Cirilla Couch, Srujan Dandu, Atirath Dhara, Jessica Johnson, Zhisen Lai, and Carli Smith. Supernova 2020wnt: An Atypical Superluminous Supernova with a Hidden Central Engine. *ApJ*, 951(1):34, July 2023. doi: 10.3847/1538-4357/acc6c3.
- [22] Pablo Lemos, Adam Coogan, Yashar Hezaveh, and Laurence Perreault-Levasseur. Sampling-based accuracy testing of posterior estimators for general inference, 2023. URL <https://arxiv.org/abs/2302.03026>.
- [23] Brenna Mockler, James Guillochon, and Enrico Ramirez-Ruiz. Weighing Black Holes Using Tidal Disruption Events. *ApJ*, 872(2):151, February 2019. doi: 10.3847/1538-4357/ab010f.
- [24] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.
- [25] Lukas Biewald. Experiment tracking with weights and biases, 2020. URL <https://www.wandb.com/>. Software available from wandb.com.

A Supernova Multi-Component Priors

We provide the priors for the parameters of our multi-component physical models below. Where values are not listed, the model defaults are used.

Model	Parameter	Description	Units	Prior
–	M_{ej}	Ejecta mass	M_{\odot}	$\log U(1, 50)$
	v_{ej}	Ejecta velocity	km s^{-1}	$\log U(2 \times 10^3, 2 \times 10^4)$
	L_{dist}	Luminosity distance	Mpc	$\log U(10, 2 \times 10^3)$
Ni	$f_{^{56}\text{Ni}}$	^{56}Ni Fraction	–	$\log U(10^{-3}, 10^{-1})$
Mag	P_{spin}	Pulsar spin period	ms	$U(0.7, 30)$
	B_{field}	Magnetic field strength	10^{14} Gauss	$\log U(0.1, 15)$
	M_{NS}^3	Neutron star mass	M_{\odot}	$\mathcal{N}(1.7, 0.2)$
CSM	n	Ejecta density profile index	–	$U(7, 12)$
	s	CSM density profile index	–	$U(0.1, 2.0)$
	M_{CSM}	CSM mass	M_{\odot}	$\log U(0.1, 50)$
TDE	ρ_0	CSM density profile scale	g cm^{-3}	$\log U(10^{-15}, 10^{-11})$
	T_v	Viscous timescale	days	$\log U(10^{-3}, 10^2)$
	M_*	Star mass	M_{\odot}	$U(10^{-1}, 5)$
	b	Scaled impact parameter	–	$U(0, 2)$
	M_{BH}	Black hole mass	M_{\odot}	$U(10^6, 10^8)$
	$l_{\text{ph},0}$	Photosphere power-law exponent	–	$U(0, 2)$
	$R_{\text{ph},0}$	Photosphere power-law constant	–	$\log U(10^{-4}, 10^4)$

Table 1: Priors for the physical parameters varied for each energy source modeled in this work.

B Loss Curves

We show the resulting training and validation loss curves in Figure 3, split by source component loss (labeled as Comp) and parameter posterior (Flow).

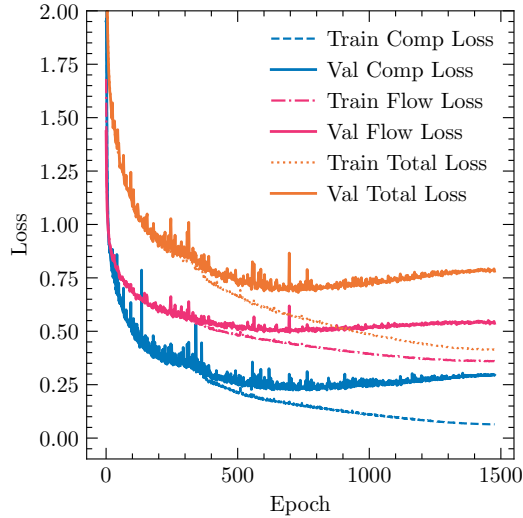


Figure 3: Loss curves for the flow network (pink), categorical network (blue), and combined model (orange) on the training (dashed line) and validation set (solid line).

³We consider the neutron star mass a nuisance parameter, and marginalize over it when sampling from the posterior of models containing a magnetar engine.

C Posterior Coverage Tests

The posterior coverage test is shown in Figure 4 for each of the different source components, estimated using our test set with TARP [22].

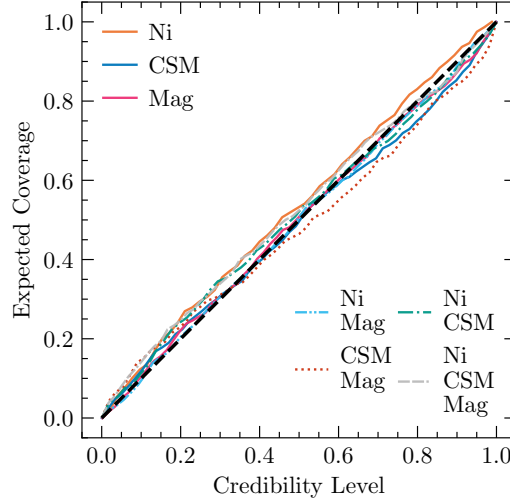


Figure 4: The expected coverage probability compared to the credibility level for the primary components (Ni, CSM, and Mag) and all possible combinations. The black dashed diagonal line indicates perfect calibration.

D Hyperparameter Tuning

We provide the hyperparameters used in our model, obtained through a Bayesian optimization sweep implemented with wandb [25].

Table 2: Hyperparameter sweep configuration and selected optimal hyperparameter values.

Parameter	Search Values / Range	Selected Value
<i>Summarizer hyperparameters</i>		
hidden dimension	[32, 64, 128, 256]	64
latent dimension	[16, 32, 64, 128]	64
transformer heads	[2, 4]	2
transformer depth	[2, 4, 8]	4
dropout rate	[0.0, 0.01, 0.02]	0.0
<i>Velocity network hyperparameters</i>		
hidden dimension	[64, 128, 256]	64
latent dimension	[32, 64, 128]	128
transformer heads	[2, 4]	4
transformer depth	[2, 4, 8]	8
<i>Categorical network hyperparameters</i>		
hidden dimension	[32, 64, 128]	128
number of layers	[2, 4]	4
<i>Joint training hyperparameters</i>		
classification weight	U(0.5, 1.2)	0.8264
<i>Optimization</i>		
learning rate	[1e-4, 3e-4]	1e-4
weight decay	[0.0, 1e-5, 1e-4]	0.0
batch size	[256]	256