Deriving an Analytical Expression for the Rise of Type Ia Supernovae with Symbolic Regression and High Cadence data

Mark Kanex [1] **Qinan Wang** [2]

Abstract

Early excess emission in Type Ia supernovae light curves provides a direct clue to the progenitor systems and explosion mechanism of these standardizable candles. Yet the traditional baseline model, i.e., the empirical 'power-law' model developed from low cadence data, is increasingly limited for modern high cadence data due to its restricted time coverage and sensitivity to the chosen fitting window. Here we use symbolic regression to search for compact analytic descriptions of sub-hour-cadence SNe Ia light curves observed by *TESS* and *Kepler*, and examine whether events with early excess emission occupy a distinct region of the learned parameter space. Our approach is able to discover functions with inherent invariance and separate different light curve components. We find that a model combining quadratic and sinusoidal terms provides the best description of the observed rise. These preliminary results demonstrate the potential of symbolic regression as an interpretable, data-driven approach for discovering empirical structure in astronomical data.

1. Introduction

Type Ia supernovae (SNe Ia) are widely believed to originate from the thermonuclear explosion of a white dwarf (WD) star when they are approaching or exceeding their mass limit, the so-called Chandrasekhar mass. With their standardizable light curves at optical and NIR wavelengths (e.g. Phillips, 1993), they provided the primary evidence for the universe's accelerating expansion (Riess et al., 1998; Perlmutter et al., 1999) and remain a cornerstone of the standard ΛCDM model. Nevertheless, the exact explosion mechanism and progenitor system of SNe Ia remain as a significant question in transient astronomy, as well as how these uncertainties might influence cosmological measurements and interpretations (see summary by Maoz et al., 2014).

Interest in the early-time behavior of SNe Ia has grown significantly following the companion interaction model suggested by Kasen (2010) that could provide hint on the progenitor properties. The canonical 'expanding fireball' model predicts that the early light curve of SNe Ia is expected to follow a power-law close to t^2 with minimal color evolution under simplified assumptions (Arnett, 1982). Extensive observations have also shown that a power-law model $L \propto t^{\alpha}$ with index $\alpha \sim 2$ can serve as a good approximation to the early optical light curve of most SNe Ia. This holds true for both well-sampled individual events like SN 2009ig (Foley et al., 2012), SN 2011fe (Nugent et al., 2011) and ASASSN-14lp (Shappee et al., 2016), as well as broader population of normal SNe Ia (e.g., Riess et al., 1999; Aldering et al., 2000; Goldhaber et al., 2001; Strovink, 2007; Garg et al., 2007; Hayden et al., 2010; Olling et al., 2015; Miller et al., 2020b). However, the power-law model has some inherent arbitrarity as the applicable range of the time or flux is not well-defined, and change in the fitting time range can result in significantly different results (Wang et al., 2021).

While the power-law model can well approximate majority of SNe Ia in early days, a handful of SNe Ia have been discovered with clear early excess features above the power-law rise (e.g. Hosseinzadeh et al., 2017; Dimitriadis et al., 2019; De et al., 2019; Miller et al., 2020a; Burke et al., 2021; Jiang et al., 2021; Dimitriadis et al., 2023; Wang et al., 2023), especially with the high cadence data coming from surveys such as *Kepler* and *TESS*. Such early excess features can be explained by multiple mechanisms including the interaction between the supernova ejecta and the non-degenerate companion in the SD scenario (Kasen, 2010), the surface detonation (Shen et al., 2018; Polin et al., 2019; Perets et al., 2019) or the shallow distribution of Ni^{56} due to mixing (Piro & Nakar, 2013; Magee et al., 2020). Distinguishing excess signals from different models require high fidelity analysis

[1]Department of Physics, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.. Correspondence to: Mark Kanex <inffiah@mit.edu>.

Proceedings of the AI4Physics Workshop at the 43rd International Conference on Machine Learning (AI4Physics@ICML 2026), Seoul, South Korea. 2026. Copyright 2026 by the author(s).

on the early rise of SNe Ia, which is severely hampered by the inaccuracy and arbitratity of the power-law model and potential deviation in the rise of 'normal' SNe Ia from the simple power-law model. This issue becomes a potential factor in the deviation in early excess rate between different statistical samples, which can range from a few percents (Fausnaugh et al., 2023) to up to $\sim 30\%$ (Deckers et al., 2022). Therefore, establishing a solid empirical model for the rise of SNe Ia and a golden criteria between SNe Ia with and without early excess is crucial for subsequent progenitor studies and cosmological analysis. The effort in improving the early light curve has been limited by the scarcity of photometric data at early phase, due to the required rapid response time and high cadence capability. With their wider field-of-view and sub-hour cadence, *Kepler* and *TESS* has been providing the most complete early photometry of various types of transients to-date, enabling robust modeling with high precision.

In search of a fiducial analytic model for SNe Ia rise, we turn to a symbolic regression algorithm (Cranmer et al., 2023), which is designed to derive an analytical and interpretable function between multiple observables and variables, rather than black-box models built from other machine learning techniques. Historically, astronomers have been uncovering empirical laws from astronomical data since the era of Ptolemy, and symbolic regression can extend that effort in a more systematic way: it does not pre-specify a certain functional form, but instead searches a broad functional space for the best-match mathematical expressions. In this way, it can reduce human bias toward favored analytical forms or oversimplified equations, and thus has the potential to discover higher order structure in the correlations. This technique has already been used in multiple astronomical studies as emulators for numerical simulations (e.g., de Soto & Villar), though few studies so far have used it to discover new laws from large data sets. Our model may further provide a rule of thumb for distinguishing the SNe Ia with and without an extra component, and for exploring whether they form two distinct subpopulations or instead lie along a continuum.

2. Data and sample

We retrieved the light curves of SNe Ia from the full *Kepler* and K2 mission as well as *TESS* Sector 1–90 from the Mikulski Archive for Space Telescopes (MAST). The photometry were reduced following the data reduction routine described in Dimitriadis et al. (2019); Shaya et al. (2015); Wang et al. (2021) and Dimitriadis et al. (2023); Wang et al. (2023); Fausnaugh et al. (2023) for *Kepler* and *TESS* data, respectively. We then selected a 'golden sample' consisting of 10 SNe Ia with redshift $z \lesssim 0.02$, sufficient S/N, and no prominent instrumental artifacts, such as the scattered light from the Earth and Moon, background noise, or major downlink breaks around the time of explosion. The light curves of our SNe sample are shown in Figure 1. The remaining data are used in later validation. The data are binned to 6-hour cadence for robust photometric accuracy and uncertainty estimation with minimal impact on the later analysis. Among the objects we choose, we caution that SN 2023bee miss critical data within ~ 2 days after the time of explosion (Wang et al., 2023), so the estimation of t_0 and baseline flux could be biased due to that. Despite that, we still like to include it in our training sample as it is the closest SNe Ia observed in *TESS* and has a unique marginally confirmed early excess. One photometric data point in SN 2020tld (TJD $\approx$ 59107.876) was identified as an outlier that, when included, forces the regression into rapid artificial oscillation; this point is excluded from all regression steps but retained in figures.

3. Methods and Analysis

3.1. Overview

Symbolic regression is a machine-learning algorithm that searches over the space of possible mathematical expressions to discover symbolic expressions that describe data (Cranmer et al., 2023). We designed a framework based on state-of-art symbolic regression algorithm, `PySR`[1], for modeling early supernovae light curves in a consistent, stable, and interpretable way. Our framework uses a residual symbolic regression approach, invariance-based data augmentations, and bootstrapping for model selection to identify symbolic form(s) that can universally describe or distinguish the sub-population of the broader family of normal SNe Ia light curves analyzed in this paper.

3.2. Residual Symbolic Regression

We use residual symbolic regression to represent each light curve as a sum of interpretable symbolic components, rather than performing symbolic regression once and forcing a single expression to explain all data. Our implementation first fits the data and then recursively fits the residual expression with restricted expression complexity until the residuals fall below critical value, and the final expression is the summation of all the resulting expression. The residual is then updated by subtracting this component. This process allows for a consistent decomposition of the major and additive minor components in the expression, and thus can serve as an approach to identify the extra additive component on top of the canonical rise part of a normal SNe Ia light curve. To minimize edge effects, we perform symbolic regression over an extended time window by adding an additional day on each side of the fitting interval of interest, while only reporting results within the original window of interest.

[1] https://github.com/MilesCranmer/PySR

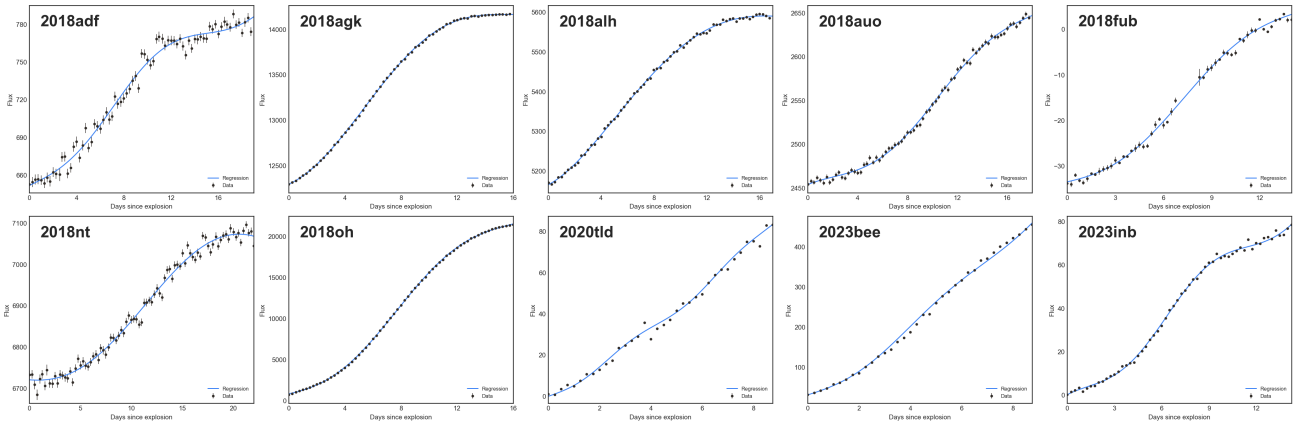

Figure 1. High S/N complete *Kepler* and *TESS* light curves of SNe Ia used for training the light-curve model, together with the corresponding best-fit symbolic regression models.

We perform symbolic regression using `PySR`, allowing the binary operators $\{+, -, \times, \div, \wedge\}$ and the unary operators $\{\sin, \exp\}$. We assign operator complexities of 1 to addition and subtraction, 2 to multiplication, 3 to division, 4 to exponentiation, and 5 to the sine and exponential functions. We enforce constraints that prevent the unary functions, $\sin$ and $\exp$, from nesting within one another (e.g., $\sin(\exp(x_0)))$, which is indicative of overfitting behavior as rapid oscillations can be added to an otherwise smooth regression curve to indiscriminately capture arbitrary minor residual errors without reflecting meaningful signal. The maximum expression complexity for each component is restricted to 25, so that the resulting components remain simple.

To prevent overfitting, we limit the number of components by allowing symbolic regression to run at most twice, and retaining the second component only if it improves a modified Bayesian Information Criterion, BIC. Here, the penalty term in the standard BIC formula, k, is the total symbolic complexity of the expression rather than the number of free parameters. In practice, adding more than two components does not produce any meaningful change in the fit, as additional terms tend to have very small amplitudes and primarily capture noise rather than real features in data.

3.3. Invariance-Based Data Transformation

Physical laws often possess some invariant properties, and incorporating known physical constraints can improve regression quality. In our case, the supernova light curve should be: (1) scale invariant, i.e. the amplitude can vary arbitrarily as it mostly depends on the distance to the supernovae rather than the supernovae physics; and (2) independent of t_0 explosion time which is itself uncertain. These invariances may also help us identify characteristic quantities that correlates with the shape of light curve, e.g. characteristic rise time if it exists. In addition, the regression is sensitive to the choice of fitting window, which is set to be between the calculated explosion time and end time. In reality it is influenced largely by temporal coverage and instrumental systematics in individual light curves rather than supernova physics. Thus, we introduce an invariance-based data augmentation scheme to improve robustness and interpretability while reducing the risk of overfitting or unphysical results.

The augmentation is created by generating multiple shifted and scaled versions of the same light curve within the original regression window. When conducting model selection, for each light curve, to minimize edge effects, we first extend the time window by one day before the time of explosion and after the peak time, and apply random small variations in the window size drawn from a Gaussian distribution with standard deviation $\sigma = 5\%$, clipped to a hard limit of $\pm 10\%$. We also randomly apply a time shift t_0 drawn from a Gaussian distribution with standard deviation $\sigma = 5\%$ of the window size, clipped to a hard limit of $\pm 10\%$ of the window size, along with random uniformly selected amplitude scaling factors A. That is, we transform the data as

$$(t, f) \;\longrightarrow\; (t - t_0,\; A\, f,\; t_0,\; A)\,.$$

thereby changing the regression problem from modeling a single-variable function to a multivariate regression problem, with flux depending on time, the applied time shift, and the amplitude scaling factor. We make it so the transformed points remain within the same time window as the original data for regression and plotting. This creates multiple snapshots of the same data with different starting and ending points along the curve, as well as supernovae exploding at different distance and time. Additionally, the *Kepler* and *TESS* data are sampled with fixed cadence, so the algorithm occasionally encounters Nyquist-related bias. To fix that, we repeatedly rerun the residual symbolic regression on 'resampled' versions of the same light curves. Each run

generates a marginally perturbed duplicate of light curve by slightly shifting each data point in time axis along the quadratic interpolation defined by the adjacent points. We then apply symbolic regression in this higher-dimensional space using the same residual procedure as before. The data augmentation process encourages it to learn a single symbolic form that is invariant to time shifts in the bounds and amplitude scalings, and robust against Nyquist-related bias and randomness unrelated to supernovae physics.

3.4. Model Discovery and Selection

We generated 100 augmented, transformed and resampled replicas of each light curves as described in Section 3.3. All of these replicas are fitted at once with `PySR`. Each run produces an expression for each light curve. We then rewrite it into a common general form by replacing numerical constants with free parameters, so that expressions with the same structure but different constants are treated to be the same model. For example, regressions with a linear component and a sine component would map to

$$a\,t + b + \sin(c\,t + d).$$

For each light curve, we count how often each general form appears across all runs and set the lower bar on occurrence rate of $\sqrt{N}$ for acceptable models, where $N = 100$ is the total number of bootstrap runs for that light curve. For the remaining candidate models passing this threshold, we compute a simple heuristic score based on how well the model fits the data, how often it appears, and the complexity of the equation. We expand on this in the appendix Section A. We then choose the final model for all light curves from the highest-scoring candidates and select an overall general form that can represent these models across all light curves, ensuring that it contains the key terms identified by this procedure. This criteria picks the robust forms against the noise and perturbations and while getting rid of the randomness in individual case. The remaining expressions and their associated scores are listed in Table 1.

3.5. Discovered Symbolic Forms

As can be seen from Table 1, across all the supernovae, the rise light curve functions that score the highest and appear the most often follow a common structure consisting of a smooth linear or quadratic component with a single sinusoidal term added. Expressions containing multiple sinusoidal terms, exponential expressions, and cross terms, such as $x_0 \sin(x_0)$, occasionally occur. The inclusion of those extra terms accounting for subtle structures in data could modestly increase R^2, but also penalized by dramatically increasing complexity. Given both their low recurrence in the symbolic regression results and their more complicated functional form, these expressions are not well motivated as preferred models for individual light curves in our score system. They are even less suitable as candidates for a single general form to describe all of the light curves together.

As we observe that the supernovae consistently favor either a quadratic model, a linear model with a sinusoidal component, or a quadratic model with a sinusoidal component, we conclude that they are generally well described by a low-order polynomial rise supplemented by a single sinusoidal component. We therefore choose

$$at^2 + bt + c + d\sin(et + f) \tag{1}$$

as a common model. This expression captures the main behavior of the light curves in our training sample without unnecessary complexity.

It is reassuring that the selected model is consistent with previous early-time power-law descriptions of Type Ia supernova light curves, which typically have a power-law index ~ 2 within first $\sim$ week after explosion, and afterwards gradually flattens within around ~ 18 days (Fausnaugh et al., 2023). In our case, the derived model can be approximated by a second order polynomial around the time of explosion when t is small and thus in agreement with the early quadratic or power-law approximation, and can naturally flattens out near peak with $(e\,t + f) \pmod{2\pi} > \pi/2$ (e.g. see different components in Figure 4). Moreover, the traditional power-law model requires an additional Gaussian or skewed-Gaussian component to fit the SNe Ia with an early excess. For example, both SN 2018oh and SN 2023bee has been modeled as a quadratic rise plus a skewed or normal Gaussian component up to around half of the maximum luminosity (Dimitriadis et al., 2019; Wang et al., 2023). Our results agree with this approach in general, but place it into a single general form that applies across multiple light curves at once as the change in sinusoidal phase f can mimic the occurrence of the extra component of the early time for both SN 2018oh and SN 2023bee. Whether this is a complete description for all early excess feature for SNe Ia, however, requires more samples for validation.

3.6. Validation with Low Signal-to-Noise *TESS* Supernovae sample

With our general expression for early-time Type Ia supernova light curves, we next validate it by fitting the complete *Kepler* and *TESS* light curve sample with lower S/N and incomplete data coverage (Fausnaugh et al., 2023) to test whether the derived symbolic model can serve as a reasonable analytic model among the general Type Ia supernova population, especially in the low S/N limit. Similar to training sample, the data are also binned into 6-hour interval for robust photometric accuracy uncertainty estimation. *TESS*

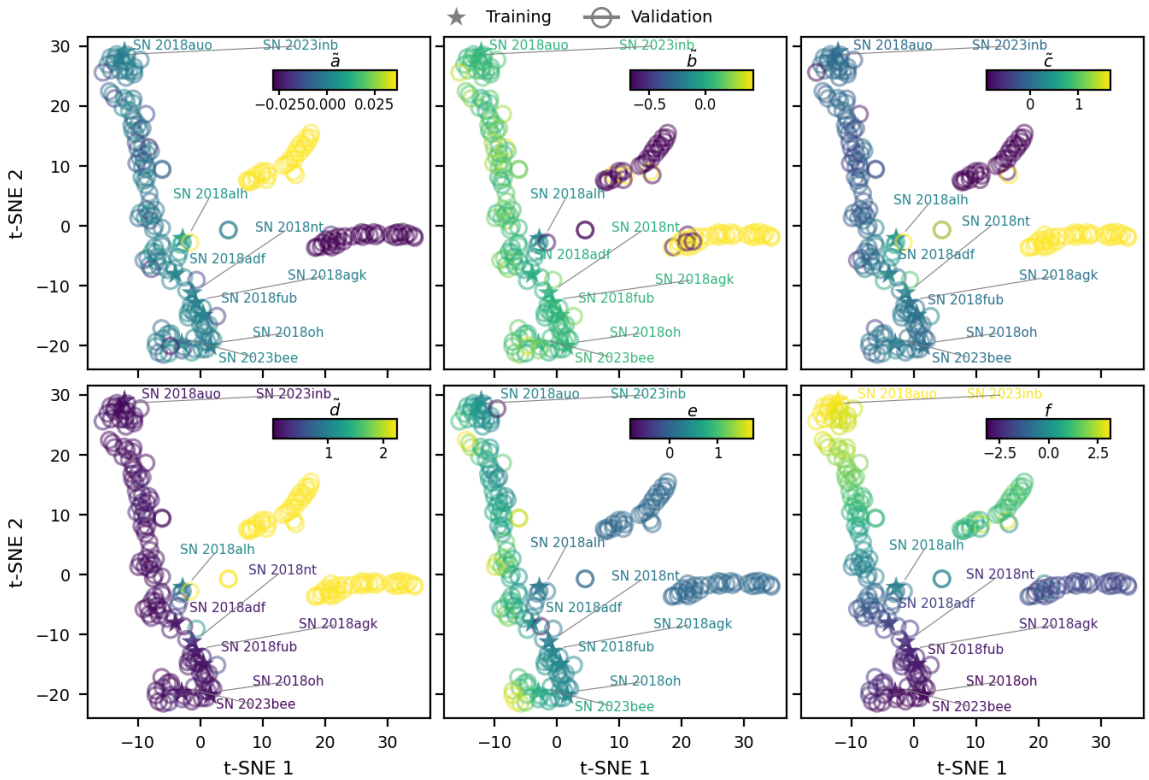

Figure 2. The t-SNE visualization of the best-fit parameter distribution for both training and validation sample, with each panel showing the correlation of t-SNE vectors to free parameters. The clusterings at t-SNE 1$\gtrsim 5$ in validation sample result from the poor fit to the data due to low S/N and the existence of data gap within *TESS* sample.

data sometimes suffer from data gap due to inter-sector downlink gap, scattered light from Earth and Moon (Wang et al., 2023), or other types of instrumental systematics, and relative short duration of *TESS* sectors (~ 28 days) makes it difficult to cover. Thus, a data cut is also applied to cut supernovae with < 30 data points after binning. We then fit the final validation sample which contain 195 supernovae after cut, as well as the ten supernovae in training sample. The best-fit results for training sample is shown in Figure 1.

Since different supernovae span different time ranges and flux levels, we rescale the individual light curves using the flux range of each object with method described in (Fausnaugh et al., 2023). This places all of the parameters on the same footing and allows them to be compared directly across SNe Ia with different apparent luminosity due to different distance, intrinsic luminosity, background flux and drift in instrumental zero-point. Fitting to the normalized light curves gives a dimensionless parameter vector $(\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d}, e, f)$.

The fitting results on the large sample enables demographic analysis on the light curve properties and parameter distributions. Figure 2 shows the visualization of best fit parameters of individual event and parameter range projected in t-SNE space. It is worth noting that all the SNe Ia in training sample with high fidelity light curve locates on the major branch at t-SNE $1 < 5$ along with the majority of the validation sample, while two minor clusters at t-SNE $1 > 5$ represents the bad fit due to the low S/N or instrumental systematics like the large data gap around the time of explosion or varying background. Because of that, we exclude the objects in the minor clusters in the following analysis, while keeping the remaining 131 supernovae on the major branch. Figure 5 in Appendix shows the correlated distribution of each parameters of samples on the major branch.

With the fitting results we also utilize the kernel density estimation (KDE) for a most representative model which corresponds to the peak density in the parameter space. The best-fit results for individual supernovae and the KDE peak model are shown in Figure 3. In addition to Figure 1 showing the best-fit results to all training samples, we highlight the fit to SN 2018oh and SN 2018agk in Figure 4, which are by far the most well studied SNe Ia with and without early excess respectively. The polynomial and sinusoidal components are also shown in Figure 4 separately. This suggests that the selected symbolic form is flexible enough to describe all rise light curve of SNe Ia in *TESS* without the need to add extra components, no matter whether they

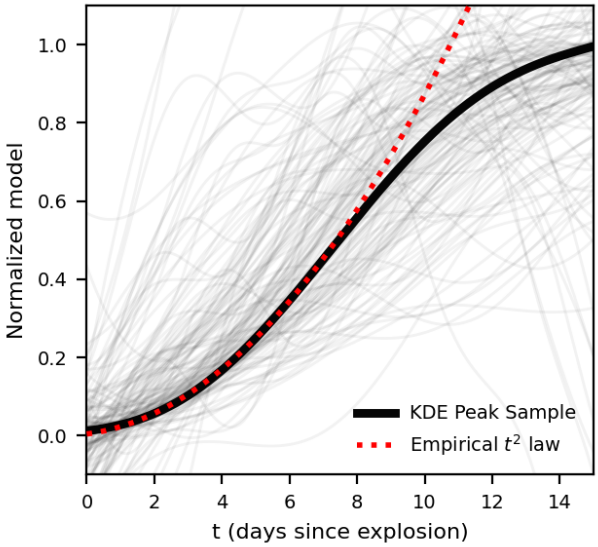

Figure 3. The best-fit models to individual supernovae in training and validation sample, with the most representative model of SNe Ia rise defined by the KDE peak in parameter space. Red dotted line shows the empirical power-law model which can only approximate the rise for the first ~ 8 days.

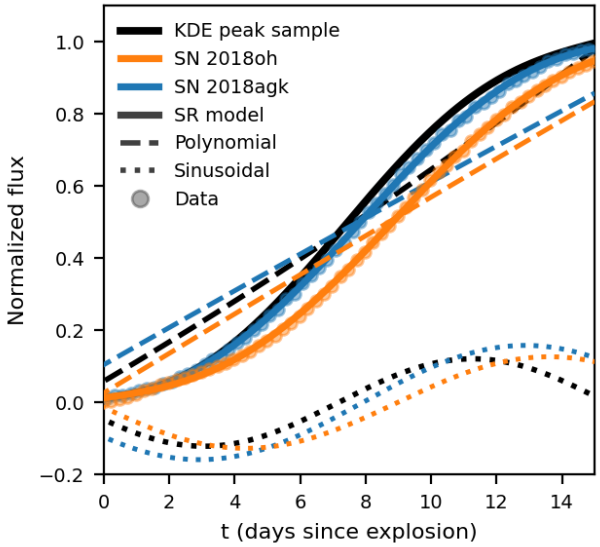

Figure 4. The KDE peak model of SNe Ia rise, in comparison with two representative SNe Ia in training sample, SN 2018oh, an SNe Ia confirmed with an early excess, and SN 2018agk, an SNe Ia confirmed with a smooth power-law. The model derived by SR and two major components are plotted in different line styles. Notice that despite its relatively long rise, the flattening at the beginning of SN 2018oh's light curve indicates an early excess with amplitude of $\sim 3\%$ of its peak luminosity.

were previously identified with SNe Ia with early excess component or not.

4. Results and Conclusion

We find a combined quadratic plus sinusoidal function can robustly fit the single-band high cadence light curve of SNe Ia from the time of explosion to peak uniquely with symbolic regression algorithm. The new model is compatible with the widely accepted empirical power-law model within first few days after explosion, and has extra advantage as it covers the whole rise light curve and does not have the arbitrarity in the fitting range. Our study has demonstrated the potential of symbolic regression algorithm in identifying the hidden analytical relationship in large astronomical data set without the need for human observations and experience.

When mapping individual supernova in parameter space, we find no obvious evidence of multiple clustering besides the ones resulting from instrumental systematics, and supernovae previous confirmed with and without early excess features fall onto the same major branch, with only minor difference in the parameter range. Such result could indicate that there is no clear distinction between the SNe Ia subpopulation with and without the early excess feature, or the symbolic regression method we adopt is not sufficiently sensitive to distinguish the two class. Our analysis is largely limited by the number of SNe Ia that are confirmed with early excess features, and more high fidelity light curve sample will help further resolve the related problem.

References

Aldering, G., Knop, R., and Nugent, P. The Rise Times of High- and Low-Redshift Type IA Supernovae Are Consistent. , 119(5):2110–2117, May 2000. doi: 10.1086/301344.

Arnett, W. D. Type i supernovae. i-analytic solutions for the early part of the light curve. *The Astrophysical Journal*, 253:785–797, 1982.

Burke, J., Howell, D. A., Sarbadhicary, S. K., Sand, D. J., Amaro, R. C., Hiramatsu, D., McCully, C., Pellegrino, C., Andrews, J. E., Brown, P. J., Itagaki, K., Shahbandeh, M., Bostroem, K. A., Chomiuk, L., Hsiao, E. Y., Smith, N., and Valenti, S. A Bright Ultraviolet Excess in the Transitional 02es-like Type Ia Supernova 2019yvq. , 919 (2):142, October 2021. doi: 10.3847/1538-4357/ac126b.

Cranmer, M. et al. Interpretable machine learning for science with symbolic regression. *arXiv preprint arXiv:2305.01582*, 2023.

De, K., Kasliwal, M. M., Polin, A., Nugent, P. E., Bildsten, L., Adams, S. M., Bellm, E. C., Blagorodnova, N., Burdge, K. B., Cannella, C., Cenko, S. B., Dekany, R. G., Feeney, M., Hale, D., Fremling, U. C., Graham, M. J., Ho, A. Y. Q., Jencson, J. E., Kulkarni,

S. R., Laher, R. R., Masci, F. J., Miller, A. A., Patterson, M. T., Rebbapragada, U., Riddle, R. L., Shupe, D. L., and Smith, R. M. ZTF 18aaqeasu (SN2018byg): A Massive Helium-shell Double Detonation on a Sub-Chandrasekhar-mass White Dwarf. , 873(2):L18, March 2019. doi: 10.3847/2041-8213/ab0aec.

de Soto, K. and Villar, V. A. Representing core-collapse supernova light curves analytically with symbolic regression. In *NeurIPS 2023 AI for Science Workshop*.

Deckers, M., Maguire, K., Magee, M. R., Dimitriadis, G., Smith, M., Sainz de Murieta, A., Miller, A. A., Goobar, A., Nordin, J., Rigault, M., Bellm, E. C., Coughlin, M., Laher, R. R., Shupe, D. L., Graham, M., Kasliwal, M., and Walters, R. Constraining Type Ia supernova explosions and early flux excesses with the Zwicky Transient Factory. , 512(1):1317–1340, May 2022. doi: 10.1093/mnras/stac558.

Dimitriadis, G., Foley, R. J., Rest, A., Kasen, D., Piro, A. L., Polin, A., Jones, D. O., Villar, A., Narayan, G., Coulter, D. A., Kilpatrick, C. D., Pan, Y. C., Rojas-Bravo, C., Fox, O. D., Jha, S. W., Nugent, P. E., Riess, A. G., Scolnic, D., Drout, M. R., K2 Mission Team, Barentsen, G., Dotson, J., Gully-Santiago, M., Hedges, C., Cody, A. M., Barclay, T., Howell, S., KEGS, Garnavich, P., Tucker, B. E., Shaya, E., Mushotzky, R., Olling, R. P., Margheim, S., Zenteno, A., Kepler spacecraft Team, Coughlin, J., Van Cleve, J. E., Cardoso, J. V. d. M., Larson, K. A., McCalmont-Everton, K. M., Peterson, C. A., Ross, S. E., Reedy, L. H., Osborne, D., McGinn, C., Kohnert, L., Migliorini, L., Wheaton, A., Spencer, B., Labonde, C., Castillo, G., Beerman, G., Steward, K., Hanley, M., Larsen, R., Gangopadhyay, R., Kloetzel, R., Weschler, T., Nystrom, V., Moffatt, J., Redick, M., Griest, K., Packard, M., Muszynski, M., Kampmeier, J., Bjella, R., Flynn, S., Elsaesser, B., Pan-STARRS, Chambers, K. C., Flewelling, H. A., Huber, M. E., Magnier, E. A., Waters, C. Z., Schultz, A. S. B., Bulger, J., Lowe, T. B., Willman, M., Smartt, S. J., Smith, K. W., DECam, Points, S., Strampelli, G. M., ASAS-SN, Brimacombe, J., Chen, P., Muñoz, J. A., Mutel, R. L., Shields, J., Vallely, P. J., Villanueva, S., J., PTSS/TNTS, Li, W., Wang, X., Zhang, J., Lin, H., Mo, J., Zhao, X., Sai, H., Zhang, X., Zhang, K., Zhang, T., Wang, L., Zhang, J., Baron, E., DerKacy, J. M., Li, L., Chen, Z., Xiang, D., Rui, L., Wang, L., Huang, F., Li, X., Cumbres Observatory, L., Hosseinzadeh, G., Howell, D. A., Arcavi, I., Hiramatsu, D., Burke, J., Valenti, S., ATLAS, Tonry, J. L., Denneau, L., Heinze, A. N., Weiland, H., Stalder, B., Konkoly, Vinkó, J., Sárneczky, K., Pál, A., Bódi, A., Bognár, Z., Csák, B., Cseh, B., Csörnyei, G., Hanyecz, O., Ignácz, B., Kalup, C., Könyves-Tóth, R., Kriskovics, L., Ordasi, A., Rajmon, I., Sódor, A., Szabó, R., Szakáts, R., Zsidi, G., ePESSTO, Williams, S. C., Nordin, J., Cartier, R., Frohmaier, C., Galbany, L., Gutiérrez, C. P., Hook, I., Inserra, C., Smith, M., Arizona, U. o., Sand, D. J., Andrews, J. E., Smith, N., and Bilinski, C. K2 Observations of SN 2018oh Reveal a Two-component Rising Light Curve for a Type Ia Supernova. , 870(1):L1, January 2019. doi: 10.3847/2041-8213/aaedb0.

Dimitriadis, G., Maguire, K., Karambelkar, V. R., Lebron, R. J., Liu, C., Kozyreva, A., Miller, A. A., Ridden-Harper, R., Anderson, J. P., Chen, T.-W., Coughlin, M., Della Valle, M., Drake, A., Galbany, L., Gromadzki, M., Groom, S. L., Gutiérrez, C. P., Ihanec, N., Inserra, C., Johansson, J., Müller-Bravo, T. E., Nicholl, M., Polin, A., Rusholme, B., Schulze, S., Sollerman, J., Srivastav, S., Taggart, K., Wang, Q., Yang, Y., and Young, D. R. SN 2021zny: An early flux excess combined with late-time oxygen emission suggests a double white dwarf merger event. , February 2023. doi: 10.1093/mnras/stad536.

Fausnaugh, M. M., Vallely, P. J., Tucker, M. A., Kochanek, C. S., Shappee, B. J., Stanek, K. Z., Ricker, G. R., Vanderspek, R., Agarwal, M., Daylan, T., Jayaraman, R., Hounsell, R., and Muthukrishna, D. Four years of Type Ia Supernovae Observed by TESS: Early Time Light Curve Shapes and Constraints on Companion Interaction Models. *arXiv e-prints*, art. arXiv:2307.11815, July 2023. doi: 10.48550/arXiv.2307.11815.

Foley, R. J., Challis, P. J., Filippenko, A. V., Ganeshalingam, M., Landsman, W., Li, W., Marion, G. H., Silverman, J. M., Beaton, R. L., Bennert, V. N., Cenko, S. B., Childress, M., Guhathakurta, P., Jiang, L., Kalirai, J. S., Kirshner, R. P., Stockton, A., Tollerud, E. J., Vinkó, J., Wheeler, J. C., and Woo, J. H. Very Early Ultraviolet and Optical Observations of the Type Ia Supernova 2009ig. , 744(1): 38, January 2012. doi: 10.1088/0004-637X/744/1/38.

Garg, A., Stubbs, C. W., Challis, P., Wood-Vasey, W. M., Blondin, S., Huber, M. E., Cook, K., Nikolaev, S., Rest, A., Smith, R. C., Olsen, K., Suntzeff, N. B., Aguilera, C., Prieto, J. L., Becker, A., Miceli, A., Miknaitis, G., Clocchiatti, A., Minniti, D., Morelli, L., and Welch, D. L. Light Curves of Type Ia Supernovae from Near the Time of Explosion. , 133(2):403–419, February 2007. doi: 10.1086/510118.

Goldhaber, G., Groom, D. E., Kim, A., Aldering, G., Astier, P., Conley, A., Deustua, S. E., Ellis, R., Fabbro, S., Fruchter, A. S., Goobar, A., Hook, I., Irwin, M., Kim, M., Knop, R. A., Lidman, C., McMahon, R., Nugent, P. E., Pain, R., Panagia, N., Pennypacker, C. R., Perlmutter, S., Ruiz-Lapuente, P., Schaefer, B., Walton, N. A., and York, T. Timescale Stretch Parameterization of Type Ia Supernova B-Band Light Curves. , 558(1):359–368, September 2001. doi: 10.1086/322460.

Hayden, B. T., Garnavich, P. M., Kessler, R., Frieman, J. A., Jha, S. W., Bassett, B., Cinabro, D., Dilday, B., Kasen, D., Marriner, J., et al. The rise and fall of type ia supernova light curves in the sdss-ii supernova survey. *The Astrophysical Journal*, 712(1):350, 2010.

Hosseinzadeh, G., Sand, D. J., Valenti, S., Brown, P., Howell, D. A., McCully, C., Kasen, D., Arcavi, I., Bostroem, K. A., Tartaglia, L., Hsiao, E. Y., Davis, S., Shahbandeh, M., and Stritzinger, M. D. Early Blue Excess from the Type Ia Supernova 2017cbv and Implications for Its Progenitor. , 845(2):L11, August 2017. doi: 10.3847/2041-8213/aa8402.

Jiang, J.-a., Maeda, K., Kawabata, M., Doi, M., Shigeyama, T., Tanaka, M., Tominaga, N., Nomoto, K., Niino, Y., Sako, S., Ohsawa, R., Schramm, M., Yamanaka, M., Kobayashi, N., Takahashi, H., Nakaoka, T., Kawabata, K. S., Isogai, K., Aoki, T., Kondo, S., Mori, Y., Arimatsu, K., Kasuga, T., Okumura, S.-i., Urakawa, S., Reichart, D. E., Taguchi, K., Arima, N., Beniyama, J., Uno, K., and Hamada, T. Discovery of the Fastest Early Optical Emission from Overluminous SN Ia 2020hvf: A Thermonuclear Explosion within a Dense Circumstellar Environment. , 923(1):L8, December 2021. doi: 10.3847/2041-8213/ac375f.

Kasen, D. Seeing the Collision of a Supernova with Its Companion Star. , 708(2):1025–1031, January 2010. doi: 10.1088/0004-637X/708/2/1025.

Magee, M. R., Maguire, K., Kotak, R., Sim, S. A., Gillanders, J. H., Prentice, S. J., and Skillen, K. Determining the ^{56}Ni distribution of type Ia supernovae from observations within days of explosion. , 634:A37, February 2020. doi: 10.1051/0004-6361/201936684.

Maoz, D., Mannucci, F., and Nelemans, G. Observational clues to the progenitors of type ia supernovae. *Annual Review of Astronomy and Astrophysics*, 52:107–170, 2014.

Miller, A. A., Magee, M. R., Polin, A., Maguire, K., Zimmerman, E., Yao, Y., Sollerman, J., Schulze, S., Perley, D. A., Kromer, M., Dhawan, S., Bulla, M., Andreoni, I., Bellm, E. C., De, K., Dekany, R., Delacroix, A., Fremling, C., Gal-Yam, A., Goldstein, D. A., Golkhou, V. Z., Goobar, A., Graham, M. J., Irani, I., Kasliwal, M. M., Kaye, S., Kim, Y. L., Laher, R. R., Mahabal, A. A., Masci, F. J., Nugent, P. E., Ofek, E., Phinney, E. S., Prentice, S. J., Riddle, R., Rigault, M., Rusholme, B., Schweyer, T., Shupe, D. L., Soumagnac, M. T., Terreran, G., Walters, R., Yan, L., Zolkower, J., and Kulkarni, S. R. The Spectacular Ultraviolet Flash from the Peculiar Type Ia Supernova 2019yvq. , 898(1):56, July 2020a. doi: 10.3847/1538-4357/ab9e05.

Miller, A. A., Yao, Y., Bulla, M., Pankow, C., Bellm, E. C., Cenko, S. B., Dekany, R., Fremling, C., Graham, M. J., Kupfer, T., Laher, R. R., Mahabal, A. A., Masci, F. J., Nugent, P. E., Riddle, R., Rusholme, B., Smith, R. M., Shupe, D. L., van Roestel, J., and Kulkarni, S. R. ZTF Early Observations of Type Ia Supernovae. II. First Light, the Initial Rise, and Time to Reach Maximum Brightness. , 902(1):47, October 2020b. doi: 10.3847/1538-4357/abb13b.

Nugent, P. E., Sullivan, M., Cenko, S. B., Thomas, R. C., Kasen, D., Howell, D. A., Bersier, D., Bloom, J. S., Kulkarni, S. R., Kandrashoff, M. T., Filippenko, A. V., Silverman, J. M., Marcy, G. W., Howard, A. W., Isaacson, H. T., Maguire, K., Suzuki, N., Tarlton, J. E., Pan, Y.-C., Bildsten, L., Fulton, B. J., Parrent, J. T., Sand, D., Podsiadlowski, P., Bianco, F. B., Dilday, B., Graham, M. L., Lyman, J., James, P., Kasliwal, M. M., Law, N. M., Quimby, R. M., Hook, I. M., Walker, E. S., Mazzali, P., Pian, E., Ofek, E. O., Gal-Yam, A., and Poznanski, D. Supernova SN 2011fe from an exploding carbon-oxygen white dwarf star. , 480(7377):344–347, December 2011. doi: 10.1038/nature10644.

Olling, R. P., Mushotzky, R., Shaya, E. J., Rest, A., Garnavich, P. M., Tucker, B. E., Kasen, D., Margheim, S., and Filippenko, A. V. No signature of ejecta interaction with a stellar companion in three type ia supernovae. *Nature*, 521(7552):332, 2015.

Perets, H. B., Zenati, Y., Toonen, S., and Bobrick, A. Normal type Ia supernovae from disruptions of hybrid He-CO white-dwarfs by CO white-dwarfs. *arXiv e-prints*, art. arXiv:1910.07532, October 2019. doi: 10.48550/arXiv.1910.07532.

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R., Nugent, P., Castro, P., Deustua, S., Fabbro, S., Goobar, A., Groom, D., et al. Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517 (2):565, 1999.

Phillips, M. M. The Absolute Magnitudes of Type IA Supernovae. , 413:L105, August 1993. doi: 10.1086/186970.

Piro, A. L. and Nakar, E. What can we Learn from the Rising Light Curves of Radioactively Powered Supernovae? , 769(1):67, May 2013. doi: 10.1088/0004-637X/769/1/67.

Polin, A., Nugent, P., and Kasen, D. Observational Predictions for Sub-Chandrasekhar Mass Explosions: Further Evidence for Multiple Progenitor Systems for Type Ia Supernovae. , 873(1):84, March 2019. doi: 10.3847/1538-4357/aafb6a.

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan,

C. J., Jha, S., Kirshner, R. P., et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116 (3):1009, 1998.

Riess, A. G., Filippenko, A. V., Li, W., Treffers, R. R., Schmidt, B. P., Qiu, Y., Hu, J., Armstrong, M., Faranda, C., Thouvenot, E., and Buil, C. The Rise Time of Nearby Type IA Supernovae. , 118(6):2675–2688, December 1999. doi: 10.1086/301143.

Shappee, B. J., Piro, A. L., Holoien, T. W. S., Prieto, J. L., Contreras, C., Itagaki, K., Burns, C. R., Kochanek, C. S., Stanek, K. Z., Alper, E., Basu, U., Beacom, J. F., Bersier, D., Brimacombe, J., Conseil, E., Danilet, A. B., Dong, S., Falco, E., Grupe, D., Hsiao, E. Y., Kiyota, S., Morrell, N., Nicolas, J., Phillips, M. M., Pojmanski, G., Simonian, G., Stritzinger, M., Szczygieł, D. M., Taddia, F., Thompson, T. A., Thorstensen, J., Wagner, M. R., and Woźniak, P. R. The Young and Bright Type Ia Supernova ASASSN-14lp: Discovery, Early-time Observations, First-light Time, Distance to NGC 4666, and Progenitor Constraints. , 826(2): 144, August 2016. doi: 10.3847/0004-637X/826/2/144.

Shaya, E. J., Olling, R., and Mushotzky, R. Active Galactic Nuclei Discovered in the Kepler Mission. , 150(6):188, December 2015. doi: 10.1088/0004-6256/150/6/188.

Shen, K. J., Kasen, D., Miles, B. J., and Townsley, D. M. Sub-Chandrasekhar-mass White Dwarf Detonations Revisited. , 854(1):52, February 2018. doi: 10.3847/1538-4357/aaa8de.

Strovink, M. Diversity of Decline Rate-corrected Type Ia Supernova Rise Times: One Mode or Two? , 671(2): 1084–1097, December 2007. doi: 10.1086/523089.

Wang, Q., Rest, A., Zenati, Y., Ridden-Harper, R., Dimitriadis, G., Narayan, G., Villar, V. A., Magee, M. R., Foley, R. J., Shaya, E. J., Garnavich, P., Wang, L., Hu, L., Bódi, A., Armstrong, P., Auchettl, K., Barclay, T., Barentsen, G., Bognár, Z., Brimacombe, J., Bulger, J., Burke, J., Challis, P., Chambers, K., Coulter, D. A., Csörnyei, G., Cseh, B., Deckers, M., Dotson, J. L., Galbany, L., González-Gaitán, S., Gromadzki, M., Gully-Santiago, M., Hanyecz, O., Hedges, C., Hiramatsu, D., Hosseinzadeh, G., Howell, D. A., Howell, S. B., Huber, M. E., Jha, S. W., Jones, D. O., Könyves-Tóth, R., Kalup, C., Kilpatrick, C. D., Kriskovics, L., Li, W., Lowe, T. B., Margheim, S., McCully, C., Mitra, A., Muñoz, J. A., Nicholl, M., Nordin, J., Pál, A., Pan, Y.-C., Piro, A. L., Rest, S., Rino-Silvestre, J., Rojas-Bravo, C., Sárneczky, K., Siebert, M. R., Smartt, S. J., Smith, K., Sódor, Á., Stritzinger, M. D., Szabó, R., Szakáts, R., Tucker, B. E., Vinkó, J., Wang, X., Wheeler, J. C., Young, D. R., Zenteno, A., Zhang, K., and Zsidi, G. SN 2018agk: A Prototypical Type Ia Supernova with a Smooth Power-law Rise in Kepler (K2). , 923(2):167, December 2021. doi: 10.3847/1538-4357/ac2c84.

Wang, Q., Rest, A., Dimitriadis, G., Ridden-Harper, R., Siebert, M. R., Magee, M., Angus, C. R., Auchettl, K., Davis, K. W., Foley, R. J., Fox, O. D., Gomez, S., Jencson, J. E., Jones, D. O., Kilpatrick, C. D., Pierel, J. D. R., Piro, A. L., Polin, A., Politsch, C. A., Rojas-Bravo, C., Shahbandeh, M., Villar, V. A., Zenati, Y., Ashall, C., Chambers, K. C., Coulter, D. A., de Boer, T., DiLullo, N., Gall, C., Gao, H., Hsiao, E. Y., Huber, M. E., Izzo, L., Khetan, N., LeBaron, N., Magnier, E. A., Mandel, K. S., McGill, P., Miao, H.-Y., Pan, Y.-C., Stevens, C. P., Swift, J. J., Taggart, K., and Yang, G. Flight of the Bumblebee: the Early Excess Flux of Type Ia Supernova 2023bee Revealed by TESS, Swift, and Young Supernova Experiment Observations. , 962(1):17, February 2023. doi: 10.3847/1538-4357/ad0edb.

A. Score Function

We define a heuristic score to compare candidate symbolic expressions as a weighted combination of three terms: goodness of fit, consistency across bootstrap runs, and model complexity,

$$w_{R^2}\left[-\log_{10}(1-R^2)\right] + w_{\mathrm{occ}}\,\frac{n_{\mathrm{occ}}}{N} - w_{\mathrm{comp}}\,C. \tag{2}$$

The first term rewards models that fit the data well, with $-\log_{10}(1-R^2)$ emphasizing differences when R^2 is close to 1. The second term favors expressions that appear consistently across bootstrap runs, helping to filter out unstable fits that depend on small perturbations of the data. The final term penalizes model complexity through the expression complexity C, discouraging unnecessarily complicated expressions.

This score is used to rank candidate symbolic forms that pass the occurrence threshold (i.e., appearing in at least 10 bootstrap realizations). When R^2 values are similar, the occurrence and complexity terms help distinguish between models, favoring consistent and simple expressions. The score is not a formal information criterion, but a practical way to compare models using relative differences. In practice, candidates with substantially lower scores are effectively disregarded, as they typically correspond to severely overfit or underfit, unnecessarily complex, or unstable expressions, allowing us to focus on a smaller set of robust models.

We set $w_{R^2} = 1$, $w_{\mathrm{occ}} = 1$, and $w_{\mathrm{comp}} = 0.10$, and take $N = 100$ to be the total number of bootstrap runs, where each run fits a single PySR model to 100 symmetry-informed duplicates of the light curve generated as described in Section 3.3.

B. Fitting results to all SNe Ia sample in *Kepler* and *TESS*

Table 1. Derived symbolic forms for individual SNe Ia and their scores.

Object	General Symbolic Form	Complexity	Occurrences	R^2	Score
SN 2018auo	$a + bx_0^c + d\sin(e + fx_0)$	26	11	0.99789	0.1864
	$a + bx_0 + cx_0^2 + d\sin(e + fx_0)$	29	12	0.99786	-0.1111
SN 2018agk	$a + bx_0 + cx_0^2$	14	16	0.98839	0.6953
	$a + bx_0 + cx_0^2 + dx_0^3$	25	19	0.99864	0.5567
	$a + bx_0 + cx_0^2 + d\sin(e + fx_0)$	29	21	0.99024	-0.6793
SN 2020tld	$a + bx_0 + cx_0^2$	14	58	0.99503	1.4839
	$a + bx_0 + cx_0^2 + \frac{d}{e + x_0}$	22	14	0.99522	0.2609
SN 2018fub	$a + bx_0 + c\sin(d + ex_0)$	21	53	0.99630	0.8621
	$a + bx_0 + c\sin(d + ex_0) + f\sin(gx_0)$	34	14	0.99640	-0.8163
	$a + bx_0 + c\sin(d + ex_0) + f\sin(g + hx_0)$	36	13	0.99655	-1.0079
SN 2018nt	$a + bx_0$	6	35	0.96079	1.1566
	$a + bx_0 + c\sin(d + ex_0)$	21	48	0.98856	0.3215
SN 2023bee	$a + bx_0 + cx_0^2 + d\sin(e + x_0)$	26	29	0.99931	0.8491
	$a + bx_0 + c\sin(d + x_0) + e\exp(-x_0)$	31	10	0.99951	0.3100
SN 2018alh	$a + bx_0 + cx_0^2 + dx_0^3$	25	10	0.99868	0.4798
	$a + bx_0 + cx_0^2 + d\sin(e + fx_0)$	29	41	0.99390	-0.2756
SN 2018adf	$a + bx_0 + cx_0^2$	14	18	0.96467	0.2318
	$a + bx_0 + cx_0^2 + d\sin(e + fx_0)$	29	22	0.98320	-0.9054
SN 2023inb	$a + bx_0 + c\sin(d + ex_0)$	21	27	0.99851	0.9966
	$a + bx_0 + c\sin(d + ex_0) + f\sin(gx_0)$	34	14	0.99848	-0.4412
	$a + bx_0 + c\sin(d + ex_0) + f\sin(g + hx_0)$	36	13	0.99874	-0.5708
SN 2018oh	$a + bx_0 + c\sin(d + ex_0)$	21	17	0.99994	2.2685
	$a + bx_0 + cx_0^2 + dx_0^3$	25	42	0.99955	1.2683

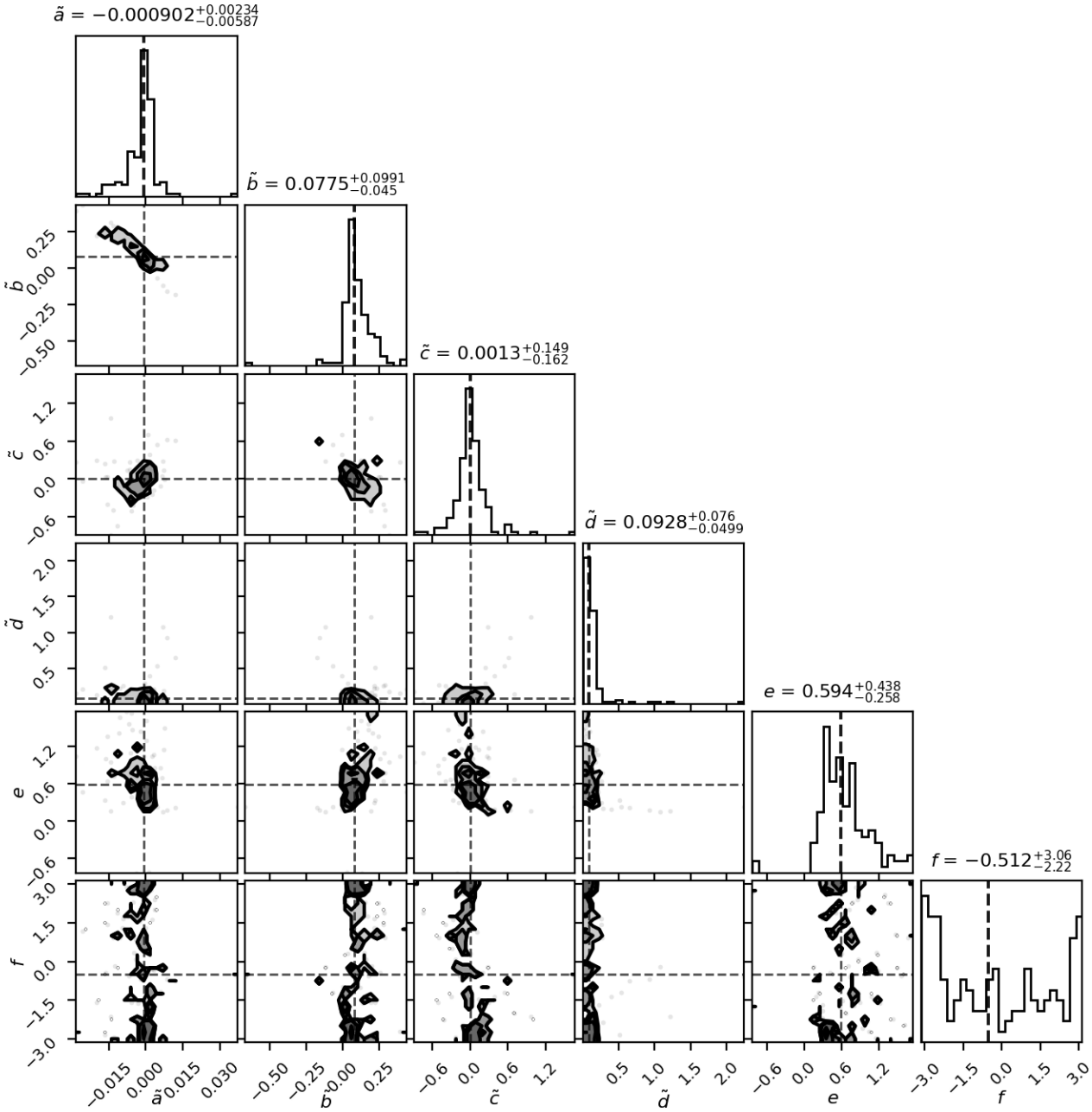

Figure 5. The distribution of the normalzied best-fit parameters to the and *TESS* SNe Ia light curve sample. To avoid degeneracy with the sinusoidal function, the range of $\tilde{d}$ and f are restricted to be $[0, \infty]$ and $[-\pi, \pi]$ respectively. The bad fit results due to gap in data or data quality have been excluded. The values on top of each plot represent the median and the $16\% - 84\%$ of each parameter distribution.