A Dynamical Systems Approach to Identifying and Classifying Variable and Transient Phenomena

JARED HAND[1]

[1] *Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC)*
Department of Physics and Astronomy,
University of Pittsburgh,
Pittsburgh, PA 15260, USA

1. PAST AND PRESENT RESEARCH

My first research project at the University of Pittsburgh sought to identify any potential dependence of measured type Ia supernova (SN Ia) host galaxy biases on observation method (photometry versus spectroscopy, for example) and galaxy property fitting technique. SN Ia are standardizable candles, where correlation between observed properties, such as color and duration (or stretch), and peak SN Ia brightness are exploited to reduce measured peak brightness dispersion. Indeed, exploiting these empirical relations led directly to the discovery of dark energy as the dominant cosmological component driving the Universe's current accelerated expansion and its ultimate fate (Riess et al. 1998; Perlmutter et al. 1999). Ideally, correlations between SN Ia and intrinsic or extrinsic modes of spectral variation are accounted for during a standardization procedure. Unfortunately, established empirical relationships between host galaxy stellar mass, stellar age, and star formation rates, and peak SN Ia brightness exist even after standardization, leading to the now-ubiquitous mass step bias (Sullivan et al. 2010; Rigault et al. 2020). Specifically, the difference between predicted peak brightness ($\mu_{mod}(z_i)$) and standardized brightness (μ_i) for some SN Ia indexed here by i, referred to as the Hubble residual $\Delta\mu_i = \mu_{mod}(z_i) - \mu_i$, co-varies with host properties in a step-like fashion. The host bias's sign flips post-standardization so that, in the case of host galaxy mass, intrinsically dimmer SNe Ia occur in lower mass hosts.

We do not know the origin of these SN Ia hosts relationship, and a natural place to search for its source is in systematics arising from observation methodology choice or form derivation of host properties. In Hand et al. (2022), we used a PISCO integral field spectroscopy (IFS) host observation subsample (Galbany et al. 2018) with SDSS and GALEX photometric coverage (Alam et al. 2015; Morrissey et al. 2007) to derive independent estimates of stellar mass and star formation rate host mass estimates. We find no evidence that observation technique or fitting method biases, let alone

produces, the SN Ia host bias and subsequent mass step, with our results also being consistent with past comparisons of galaxy property estimators (see Conley 2013 for a review).

In August 2019 I relocated to the Lawrence Berkeley National Laboratory to work under Dr. Alex Kim developing a new empirical SN Ia model. This project was funded by a Department of Energy Office of Science SCGSR Award I received. SN Ia theory is insufficiently developed to parametrically model intrinsic SN Ia spectral variation, so the SN Ia research community instead develops increasingly sophisticated empirical models trained on ever-growing photometric and spectroscopy observation sets. Extrinsic SN Ia spectral variability is largely sourced from foreground dust attenuation. Dust sourced variability is phase-independent, and most SN Ia empirical models therefore assume all phase-independent spectral variation can be modeled with a single-component effective attenuation law (Jha et al. 2007; Burns et al. 2011; Mandel et al. 2022). Those that do not assume an underlying dust model in their architecture group all phase-independent variation into a single, physics-agnostic color law (Guy et al. 2007; Kenworthy et al. 2021). Neither approach is satisfactory. These effective dust attenuation models or agnostic color laws ignore per-SN variability in attenuation curve slope, an artificial constraint that may contribute to the mass step's size (Brout et al. 2019; Popovic et al. 2021). Physical consideration alone requires at least two dust attenuation model components (Weingartner & Draine 2001), and ignoring or grouping together potentially intrinsic, phase-independent variation is a shared weakness for all current SN Ia models. Separating and appropriately modeling intrinsic and extrinsic phase-independent variation will improve SN Ia sample standardization by improving light curve fitting. Such improvements are paramount for current and future SN Ia survey utility since uncertainty in Hubble constant estimates derived from SN Ia distance ladders is now dominated by SN Ia model systematics.

Dr. Kim and I have developed a new model that fits simultaneously two physics-agnostic phase independent color templates and a single phase-dependent variation template. This model is implemented with Stan (Carpenter et al. 2017) and trained using rest frame SNFactory spectrophotometric time series from (Aldering et al. 2002). We make novel use of bivectors to determine our best fit model, recognizing that our two color templates form the basis of an oriented structure within a two-dimensional vector subspace. We further exploit the innate geometry of our model to decompose our two-dimensional color model into consistent, physically interpretable results with a first of its kind application of geometric algebra to calculate ten-dimensional rotations, projections, and intersections. Instead of extracting two dust-sourced color templates, we recover one dust-like component and one component that is a mixture of extrinsic and intrinsic features. This phase-independent intrinsic variability is most noticeable around the Ca H%K spectral feature, a feature that exhibit relatively little change during the photospheric phase of an SN Ia explosion. Importantly, we find that phase-independent SN Ia variation modeled from observation cannot be ex-

plained by dust alone and largely negating the narrative that host relationship step biases arise entirely from extrinsic sources (Brout et al. 2019; Popovic et al. 2021)

These two extracted color templates provide unprecedented capacity for us to explore dust-source variation between SNe Ia, something that we do with a hierarchical Bayesian dust model also implemented with Stan. We find that SN Ia with phase-independent variation dominated by intrinsic modes (or intrinsic-dominated SNe Ia) systematically fit for lower total-to-selective extinction R_V values than their extrinsic-dominated counterparts. This intrinsic-dominated subsample includes both nominally dust free SNe Ia and SNe Ia with negative fit 'absolute extinction' values Past estimates of SN Ia sample R_V values have consistently been lower than the Milky Way average of $R_V = 3.1$, and our results purport that including such intrinsic-dominated SN Ia can bias one's effective R_V estimate low. More importantly, we recover different, non-Gaussian R_V distributions for our both intrinsic and extrinsic-dominated subpopulations.

A paper summarizing our methodology and the above results is currently undergoing an internal review within the Nearby Supernova Factory group, but Dr. Kim and I are already developing improved model versions. Our ten-band binned spectrophotometry currently used is too low resolution to quantify spectral features, so we will be training our next model with a 25-band binned spectrophotometry. Furthermore, it is unclear if our previously recovered intrinsic, phase-independent variation captures spectral variation that is indeed effectively phase independent, or is instead only a time-averaged mode of phase-dependent spectral variation. A comparison of our results to those of the nonlinear Twins framework and probabilistic autoencoder of Boone et al. (2021) and Stein et al. (2022), respectively, hints at our recovered intrinsic, phase-independent variational mode being the former. Nonetheless, two new model versions will clarify this ambiguity: a version with two phase-dependent components, and another with one phase-dependent and one phase-independent components. As we clearly extracted dust-source extrinsic variation, we will also incorporate dust attenuation recipes directly into these new models' architecture.

Once these new models are trained, we will compare per-SN parameter samples to host properties. Of particular interest will be a comparison per-SN fit R_V values to their host's fit global R_V. We will also analyze per-SN parameter clustering and compare to SN Ia peculiar spectral subclass membership labels that are known a priori. It will be interesting to see if per-SN dust parameters cluster correspond with SN Ia spectral subtypes, since that would entail intrinsic SN Ia variation is bleeding into purported extrinsic parameters values. Not all SNe Ia are standardizable candles and identifying these subtypes while simultaneously fitting light curves would be a useful tool for future surveys. Indeed, accurate and efficient identification and classification of variable or transient events (not just SN Ia subtypes) is a massively important prerequisite to any science deliverables This is especially true in an era of

big data astronomy and with the upcoming Legacy Survey of Space and Time (LSST) from the Vera C. Rubin Observatory (Ivezić et al. 2019).

2. MODELING OBSERVATIONS AS DYNAMICAL SYSTEMS

Time domain (TD) astronomy will be revolutionized by LSST's predicted $\approx 10^7$ transient/variable observations *per night* (Narayan et al. 2018). TD phenomena inform near-countless physical processes, such as variable stars and core collapse SNe (type Ib/c or II) constraining stellar interior and end-of-life stellar evolution models, respectively. Section 4.3 of Ivezić et al. (2019) summarizes many of the potential projects to be enabled by LSST TD data products. An alert-broker workflow characterizing (identifying transient versus variable events) and classifying (identifying physical or empirical subtypes), followed by prompt, user-friendly discriminating and distributing of labeled TD events to scientists is a primary objective of the LSST TD working group (see Narayan 2018 for an example architecture). Therefore, a fundamental component of LSST's alert-broker pipeline is an accurate and efficient classification procedure. Past and current surveys rely on spectroscopic followup or the human eye for classification, an approach made untenable by ≈ 20 TB/night of raw data expected from LSST. This leave the LSST alert-broker dependent on an algorithm-based approach utilizing machine learning.

Algorithmic TD classification from photometry relies on temporal evolution of broad-band photometric features to classify, requiring much larger training sets than spectroscopic or by-eye classification. Characterization and classification is currently done separately, such as the ANTARES alert-broker framework first characterizing if an object is a transient or variable event before performing classification (Narayan et al. 2018). Simultaneous characterization and classification would be preferred, simplifying one's pipeline and error propagation. Measurement observation heteroskedasticity (arising from more distant objects being more difficult to observe) and missed observations (due to weather or technical issues, for example) will complicate LSST photometric time series' data structure. Traditional machine learning classifiers cannot easily account for these technicalities. This is currently circumvented by dimensionality reduction of photometric time series via template fitting or wavelet decomposition. These reduced representations then become inputs for classifiers such as boosted decision trees and neural networks (see Lochner 2016 and Malz 2019 for SN classification details). This separate fitting/decomposition and classification procedure again complicates the alert-broker architecture and error propagation.

Template fittings compare TD observations against SN Ia empirical models that themselves are incomplete while being also sensitive to template training data (which, right now, is necessarily different from LSST's eventual SN Ia coverage). Furthermore, these current classification schemes struggle to accurately identify stripped envelope core collapse SNe, leading to false positive/negative errors that degrade LSST data product utility. For example, core collapse SN contamination in cosmology-ready

photometric SN Ia samples reduces per-SN statistical weight in deriving cosmology parameter estimation. Wavelet decomposition, although physics agnostic, is consistently less accurate than template-based algorithms (Malz et al. 2019).

More fundamentally, these staggered workflows ignore the dynamics of the light curve. TD objects have nonzero flux 'velocity' which is effectively ignored in current alert-broker pipelines. My research proposal will rectify this issue in a unified characterization/classification framework by modeling TD flux evolution as a dynamical system. The dynamics of different transient or variables will be functionally distinguishable once discovered using machine learning. Moreover, such a framework naturally quantifies different timescales, even in the case of unsupervised learning.

Two simultaneous models need to be trained in dynamical systems discover: a trained representation of one's dynamical system observations $\mathcal{O}$, and a trained representation of the vector field $\mathcal{V}$ inscribing the discovered dynamics. Both $\mathcal{O}$ and $\mathcal{V}$ should be trained simultaneously, informing the other. An added structure labeling each event needs to be incorporated as well. Furthermore, the data representation $\mathcal{O}$ must center observations to a temporal zero-point. For example, SN observation series should be centered to an agreed phase, such as the phase of maximum brightness. We faced a similar issue

REFERENCES

Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12, doi: 10.1088/0067-0049/219/1/12

Aldering, G., Adam, G., Antilogus, P., et al. 2002, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 61–72, doi: 10.1117/12.458107

Boone, K., Aldering, G., Antilogus, P., et al. 2021, ApJ, 912, 70, doi: 10.3847/1538-4357/abec3c

Brout, D., Scolnic, D., Kessler, R., et al. 2019, ApJ, 874, 150, doi: 10.3847/1538-4357/ab08a0

Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, AJ, 141, 19, doi: 10.1088/0004-6256/141/1/19

Carpenter, B., Gelman, A., Hoffman, M. D., et al. 2017, Journal of statistical software, 76

Conroy, C. 2013, ARA&A, 51, 393, doi: 10.1146/annurev-astro-082812-141017

Galbany, L., Anderson, J. P., Sánchez, S. F., et al. 2018, ApJ, 855, 107, doi: 10.3847/1538-4357/aaaf20

Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11, doi: 10.1051/0004-6361:20066930

Hand, J., Liu, S., Galbany, L., et al. 2022, ApJ, 925, 115, doi: 10.3847/1538-4357/ac389f

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Jha, S., Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122, doi: 10.1086/512054

Kenworthy, W. D., Jones, D. O., Dai, M., et al. 2021, ApJ, 923, 265, doi: 10.3847/1538-4357/ac30d8

Lochner, M., McEwen, J. D., Peiris, H. V., Lahav, O., & Winter, M. K. 2016, ApJS, 225, 31, doi: 10.3847/0067-0049/225/2/31

Malz, A. I., Hložek, R., Allam, T., J., et al. 2019, AJ, 158, 171, doi: 10.3847/1538-3881/ab3a2f

Mandel, K. S., Thorp, S., Narayan, G.,
    Friedman, A. S., & Avelino, A. 2022,
    MNRAS, 510, 3939,
    doi: 10.1093/mnras/stab3496

Morrissey, P., Conrow, T., Barlow, T. A.,
    et al. 2007, ApJS, 173, 682,
    doi: 10.1086/520512

Narayan, G., Zaidi, T., Soraisam, M. D.,
    et al. 2018, ApJS, 236, 9,
    doi: 10.3847/1538-4365/aab781

Perlmutter, S., Aldering, G., Goldhaber,
    G., et al. 1999, ApJ, 517, 565,
    doi: 10.1086/307221

Popovic, B., Brout, D., Kessler, R.,
    Scolnic, D., & Lu, L. 2021, arXiv
    e-prints, arXiv:2102.01776.
    https://arxiv.org/abs/2102.01776

Riess, A. G., Nugent, P., Filippenko,
    A. V., Kirshner, R. P., & Perlmutter, S.
    1998, ApJ, 504, 935,
    doi: 10.1086/306106

Rigault, M., Brinnel, V., Aldering, G.,
    et al. 2020, A&A, 644, A176,
    doi: 10.1051/0004-6361/201730404

Stein, G., Seljak, U., Bohm, V., et al.
    2022, arXiv e-prints, arXiv:2207.07645.
    https://arxiv.org/abs/2207.07645

Sullivan, M., Conley, A., Howell, D. A.,
    et al. 2010, Monthly Notices of the
    Royal Astronomical Society, 406, 782,
    doi: 10.1111/j.1365-2966.2010.16731.x

Weingartner, J. C., & Draine, B. T. 2001,
    ApJ, 548, 296, doi: 10.1086/318651