Embedding Supernova Evolution in a Metric Space

Ofek Bengiat

Department of Particle Physics and Astrophysics, Gal-Yam Group
Weizmann Institute of Science, Israel

Supernova Working Group Telecon, April 2021

Overview

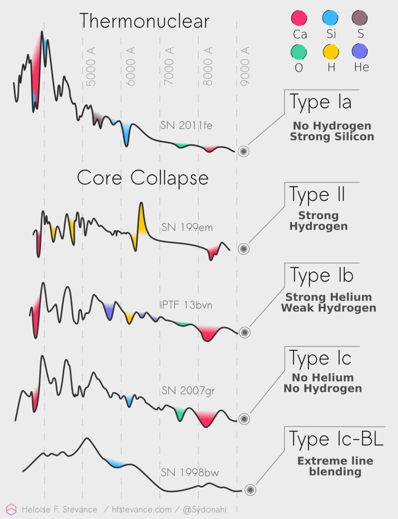
SN Spectroscopic Classification Today

- Usually relies on a single point in time, i.e. $f(t = t_0, \lambda)$.
- Mutually exclusive & exhaustive to at least some extent (many SNe labeled "peculiar")
- Detects standard candles (SNe Ia)

Problem

disagreement between times

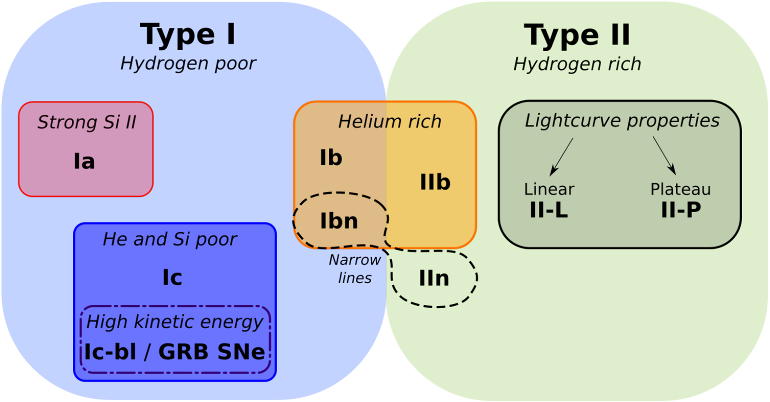
Information from other times is disregarded



Flux rescaled + const. →

Rest Wavelength [Å] →

Current Classification is Old



Problem

Not quantitative definitions, negative definitions exist.
Type continua not addressed (e.g. Ib-IIb, Ib-Ic, Ic-IcBL)

OBJECTIVE OF THIS WORK:

Devise a quantitative and time-domain-aware method to compare between/classify SNe.

Creating a Data Point: Interpolation

PyCoCo: a tool developed in Vincenzi et al. (2019) for SN spectro-photometric data interpolation & more.
Data: 84 Well-Observed, typical SNe, mostly from Vincenzi et al. (2019).
Quality cuts based on data distribution (in progress).

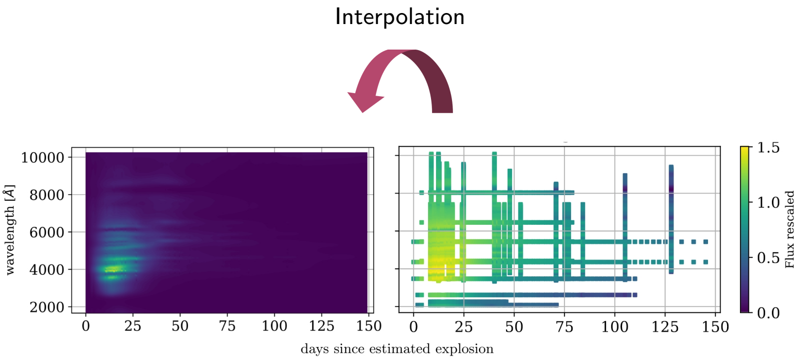
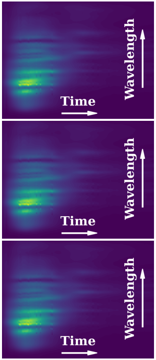


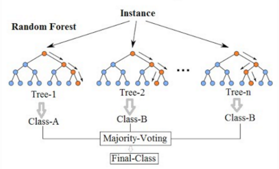
Figure: Left: A data point – $f(t, \lambda)$. Right: Its Raw Input. (SN 2007af)

A Metric: Unsupervised Random Forest

- Trained on the data alone, **unaware** of Types.
- Learns which **correlations** are important.
- Output: a **dissimilarity matrix**



$\xrightarrow{\text{training}}$



$\rightarrow$ a function $d : \mathbb{SN} \times \mathbb{SN} \rightarrow [0, \infty)$, which provides a **dissimilarity matrix**

Unsupervised Random Forest: More Details

- Create a synthetic data set:
  - same number of features ($N_{\mathrm{ft}} = 5151$)
  - same marginal distributions
  - but, zero correlations
- Train RF classifier on discriminating real data from synthetic data $\rightarrow$ RF learns the important correlations
- Similarity between two objects = # trees with exact same path for them

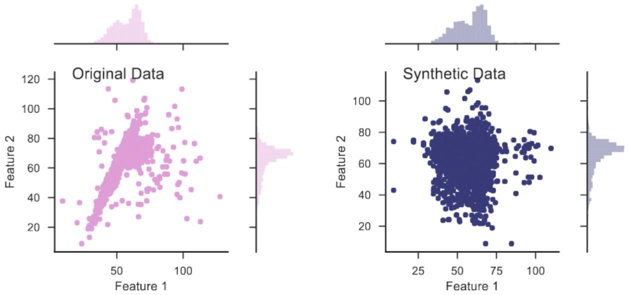


Figure: Synthetic data, illustration for 2 features. Graph (and idea) from Baron (2019).

How to Visualize a Dissimilarity Matrix

- Have: the result, in the form of a dissimilarity matrix
- Want: to **see** it
- Solution: an **embedding** (tSNE, Maaten and Hinton, 2008): construct an **abstract 3D Euclidean Space**, where the Euclidean distances are close to the ones from the matrix.

a dissimilarity matrix:

$$\begin{pmatrix} & & \\ d(a,b) = 12 & & \\ d(a,c) = 22 & d(b,c) = 27 & \end{pmatrix} \xrightarrow{\text{embedding}}$$

We have NO coordinates

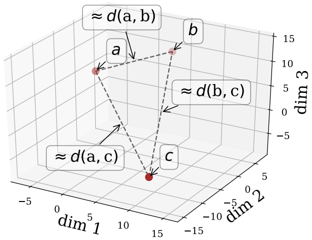


Figure: Axes are some coordinates of the abstract Euclidean Space constructed.

Agreement With Current Scheme

- Space is unaware of the Types (colors), still produces agreement.
- Type Ib is split
- Contaminations vs. continua

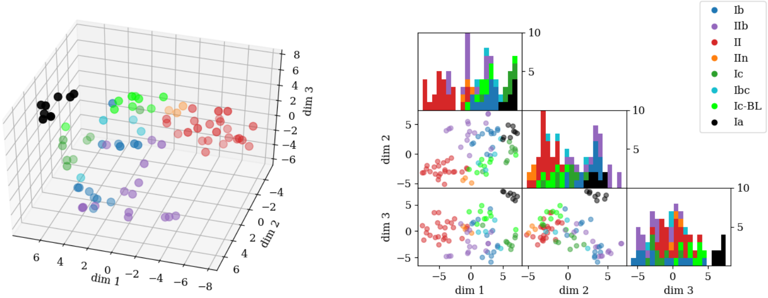


Figure: A tSNE visualization of the resulting dissimilarity matrix as an abstract 3D Euclidean Space.

Example: Query SN2005cp (IIn)

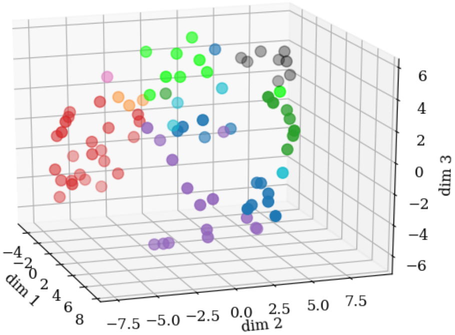
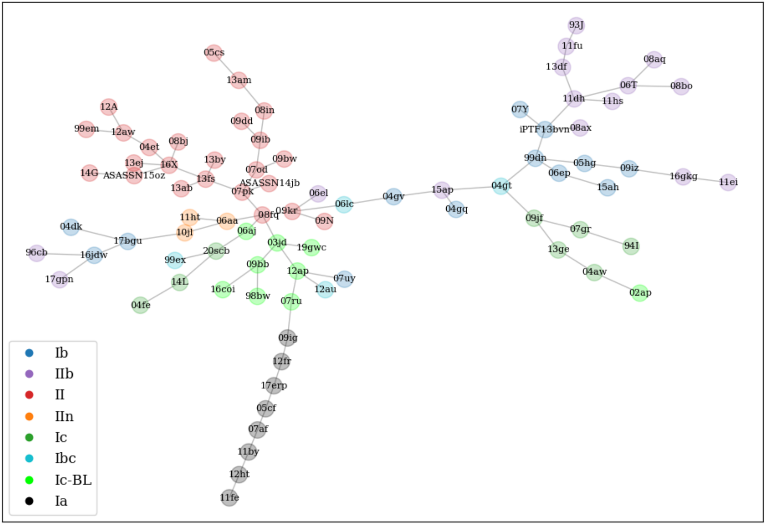


Figure: SN2005cp (1 spectrum + BVRI) in pink is placed in the IIn cluster.

Minimal Spanning Tree: trying to find some hierarchy



Example: Examining the Split of Type Ib

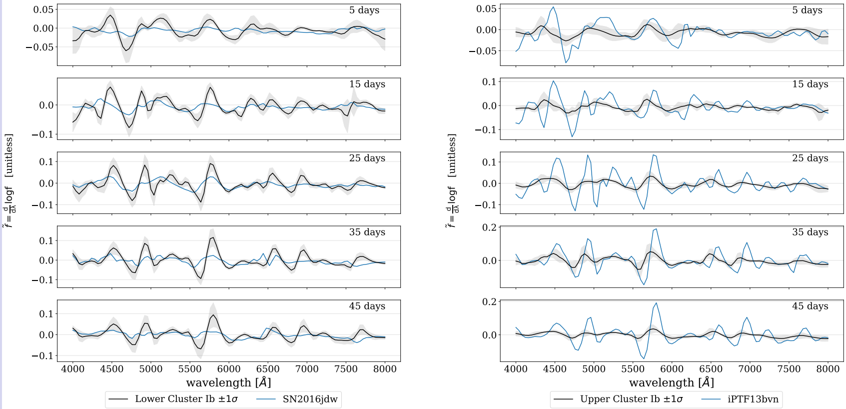


Figure: Comparison of averages of the two Ib clusters (black lines $\pm 1\sigma$) with a typical SN of the other cluster (blue lines). The upper cluster is characterized by less prominent lines than in the lower cluster.

Summary

- Unsupervised, time-domain aware algorithm reproduces known types for standard SNe
- Could ultimately be a basis for a quantitative time-domain-aware classification
- Further understanding of types and continua – could tell one where to look
- Use on young/ambiguous SNe, solving dilemmas, archives & surveys when complete
- Ongoing work: expanding data set + post-embedding analysis

Thank you for listening.
Questions? Ideas?

References

D. Baron, Machine Learning in Astronomy: a practical overview, 2019.

J. Guillochon et al., *Astrophysical Journal* **2017**, *835*, 64.

L. van der Maaten, G. Hinton, *Journal of Machine Learning Research* **2008**, *9*, 2579–2605.

M Vincenzi et al., *Monthly Notices of the Royal Astronomical Society* **2019**, *489*, 5802–5821.

O. Yaron, A. Gal-Yam, *Publications of the Astronomical Society of the Pacific* **2012**, *124*, 668–681.

Image Credits:
NASA/ESA, The Hubble Key Project Team and The High-Z Supernova Search Team
H.F. Stevance / @Sydonahi / hfstevance.com
NASA, ESA, HEIC, and The Hubble HeritageTeam (STScI/AURA)