Interpolation of Supernova Time Series Spectra with Optimal Transport

Jacob Krag Nørgaard (GXV668), Gustav Gudmandsen (RZD536), Anna Sophia Maxen (DHV202) and Lucas Hallgren (ZCB218)

March 2025

1 Introduction

For our presentation and write-up, we chose the article "A novel optimal transport-based approach for interpolating spectral time series" by [Ramirez et al., 2024] on using optimal transport (OT) theory for interpolating between temporally separated supernovae (SNe) spectra. More than 10^7 SNe are expected to be observed by LSST in its ten year operation time. In order to classify these many SNe, machine learning techniques have to be developed and trained on existing SNe spectra of different classes and temporal evolutions. The production of synthetic SNe spectra for temporal evolution, becomes a way to cut down on the observations needed for us to get an insight into the behaviour and properties of these SNe.

SNe are important in astronomy because they are contributors to galaxy evolution, dynamical and chemical, and they are being used as standard candles in distance measurements throughout the Universe. Classification of SNe (sub)-types and determining the properties of supernova spectra and light curves for large SN samples are therefore crucial to develop these fields.

2 Method

Optimal Transport (OT) is a method within "Transport Theory" focused on determining the most efficient way to move "mass" from one distribution to another. "Mass" in this context is understood as the representation of resources, probabilities, or in our case; flux.

Given a set of locations describing the initial distribution $(x_1, x_2, ...x_n)$, the goal of OT is to determine an optimal transport plan, $T(x_n)$, to move the mass of the locations to a new set of locations $(y_1, y_2, ...y_n)$ as

$$T(x_i) = y_i \tag{1}$$

The optimal transport plan is the one that minimizes the total cost function C_T defined

$$C_T = \sum_i^n c(x_i, T(x_i)) \tag{2}$$

Wherein c is the cost function of moving one point to another $x_i \to y_i$. The cost function between two points, $c(x_i, y_i) \geq 0$, is usually just the squared Euclidean distance between the points, but it can be more complex. Therefore, the cost of the transport is larger if the distance is larger. Through minimization we can determine the OT plan and construct the so-called Wasserstein distance, defined as the quantified metric between the two distributions.

An interpolation can be obtained by mapping the OT. Creating the OT plan between two known distributions, 1 initial and final, are weighted with weights [1 - α, α], where $\alpha \in [0, 1]$ depends on their proximity with the interpolation. An array of different α's leads to different interpolated spectra corresponding to different timestamps between the initial and final distributions.

In their paper, this entire process is computed through the Python Optimal Transport (POT) library, which provides a function for finding the Wasserstein barycenter between each pair of distributions.

3 Results

Fig. 1 shows an example of using OT to do a weighted interpolation between two distributions. We see a comparison to a standard linear interpolator, which for every point in the domain just is the average of the two distributions. As they describe it in the paper: "(Finding the Wasserstein barycenter) is like finding a middle point, not in terms of physical distance, but in terms of how much you would have to change each distribution to reach this middle point." Ramirez et al. [2024]

What has been described so far is where two spectra (distributions) have been used in the OT interpolation machinery, but in the paper a more complex model is set up. Using several phase spectra to generate several pairs of interpolations for a missing phase spectrum, they end with a final phase spectrum as a weighted average of the individual interpolated phase spectra. Meaning that, instead of two mother distributions and one interpolation pr. time step, they have many known distributions, and do interpolations between all possible pairs of spectra for all desired time steps. Multiple interpolations are then weighted for each time step, to get the final prediction.

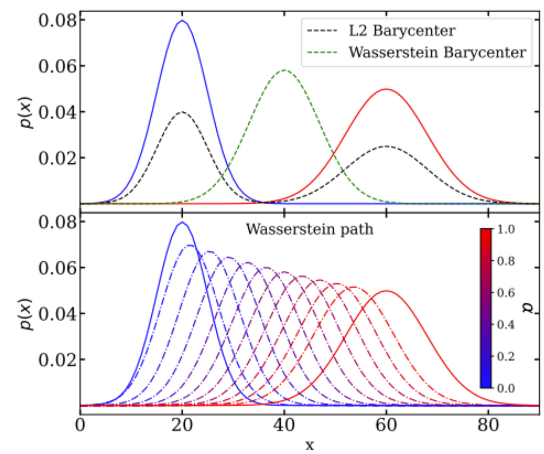

Figure 1: Plots illustrating the OT method with Wasserstein distances. *Top*: Comparison between the Wasserstein Barycenter interpolation and the linear interpolator $L2$. *Bottom*: Wasserstein distances from the initial (blue, $\alpha = 0$) distribution to the final (red, $\alpha = 1$) one with different weights illustrating the temporal evolution. Figure from [Ramirez et al., 2024].

Relative spectral residuals, ϵ, can be calculated between an interpolated spectrum and a known model spectrum to the same phase (time). This quantifies how good the interpolation is, and ϵ is included when more interpolations are averaged weighted to one.

4 Conclusion

OT is a viable and efficient method for interpolating between large amounts of data separated by differences in distributions. Through tools such as the Python Optimal Transport (POT) package, it can be implemented in a wide variety of software problems. This paper highlighted just one method in which OT and Transportation Theory can be utilized to interpolate between data. It also has appliances in economics, and several topics in ML, such as generative modeling and image-to-image translation.

References

Mauricio Ramirez, Giuliano Pignata, Francisco Förster, Santiago González-Gaitán, Claudia P. Gutiérrez, Bastian Ayala, Guillermo Cabrera-Vives, Márcio Catelan, Alejandra M. Muñoz Arancibia, and Jonathan Pineda-García. A novel optimal transport-based approach for interpolating spectral time series: Paving the way for photometric classification of supernovae. , 691:A33, November 2024. doi: 10.1051/0004-6361/202449170.