



HostSub_GP: Precise Galaxy Background Subtraction in Transient Long-slit Spectroscopy with Gaussian Processes

Chang Liu (刘畅)^{1,2,3}  and Adam A. Miller^{1,2,3} 

¹ Department of Physics and Astronomy, Northwestern University, 2145 Sheridan Rd, Evanston, IL 60208, USA; ptg.cliu@u.northwestern.edu

² Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA

³ NSF-Simons AI Institute for the Sky (SkAI), 172 E. Chestnut St., Chicago, IL 60611, USA

Received 2025 September 8; revised 2025 December 12; accepted 2026 January 23; published 2026 February 10

Abstract

We present a novel host galaxy subtraction technique in long-slit spectroscopy for extragalactic transients. Unlike classic methods which generally estimate the background using simple interpolation of local galaxy flux in the 2D spectrum, our approach leverages multi-band archival images of the host galaxies to model the background emission from the galaxy in the 2D spectrum. Such imaging encodes the wavelength-dependent galaxy profile along the slit, and is readily accessible through wide-field imaging surveys. We construct a smooth prior for the 2D galaxy profile with a Gaussian process (GP) based on these reference images, and use another GP to model the correlated deviations from the prior in the observed spectrum. This enables accurate inference of the galaxy flux blended with the transient. On synthetic long-slit data of a spiral galaxy extracted from a Multi Unit Spectroscopic Explorer hyper-spectral cube, the GP method remains robust as long as the host galaxy is spatially resolved and consistently outperforms classic methods. We apply the method to archival Keck spectra of two real transients, SN 2019eix and AT 2019qiz, to further demonstrate how the method uniquely recovers weak spectral features amid strong galaxy contamination, enabling refined constraints on the properties of both transients. We have released the software implementation, HostSub_GP, a scalable toolkit that leverages JAX, with an MIT license.

Unified Astronomy Thesaurus concepts: [Spectroscopy \(1558\)](#); [Astronomy data reduction \(1861\)](#); [Gaussian Processes regression \(1930\)](#); [Time domain astronomy \(2109\)](#)

1. Introduction

Deblending the photons of extragalactic transients from their host galaxies is a fundamental challenge in observational time-domain astronomy. Precise host subtraction in spectroscopy is critical in both classifying new transients and follow-up studies. Host contamination dilutes the spectral features from the transients, making robust classification particularly difficult for transients overlapping a high background (e.g., galaxy nuclei). For example, among the Type Ia supernovae (SNe Ia) discovered by the Zwicky Transient Facility, ambiguous subtyping is prevailing among those near the galaxy cores (Dimitriadis et al. 2025). A biased classification rate towards high background regions adds complexity to the rate studies on certain transient types. While fitting the joint spectral energy distribution (SED) of the transient and its host (e.g., Howell et al. 2005; Muthukrishna et al. 2019; Goldwasser et al. 2022) improves the classification accuracy on host contaminated spectra (Milligan et al. 2025), these methods still struggle

when the galaxy light dominates. Host contamination also impacts the quantitative measurement of spectral features. Even for SNe near maximum luminosity which outshine their background, host contamination can bias line velocity measurements (Burgaz et al. 2025). At later phases, when SNe are generally faint compared to their background, the effects of host contamination become even more pronounced. Many existing spectroscopic samples of late-time SNe are limited to objects with large offsets from their host galaxies (e.g., Graham et al. 2017; Maguire et al. 2018), which introduces substantial selection biases in the local host environment. Other samples inevitably include objects with significant residual galaxy contamination (e.g., Maguire et al. 2012; Tucker et al. 2020; Fang et al. 2022; Das et al. 2023), which compromise quantitative inferences of the SN properties, such progenitor masses of both SNe Ia (Flörs et al. 2018, 2020) and core-collapse SNe (Dessart & Hillier 2020; Fang et al. 2025a, 2025b).

Precise transient flux measurements in photometry often leverage archival reference images of the host galaxies. As a routine in modern time-domain surveys, image subtraction techniques (e.g., Alard & Lupton 1998), which match the point spread functions (PSFs) between the new and the reference

images to robustly subtract the background flux, play a vital role in transient discovery and producing reliable flux measurements (e.g., Shappee et al. 2014; Masci et al. 2019). In contrast, direct image subtraction is far more difficult in spectroscopy: (i) pre-transient archival spectroscopic observations suitable for subtraction rarely, if ever, exist; (ii) post-transient spectroscopic reference observations are unavailable until the transient has fully faded, and are significantly more resource-intensive (e.g., requiring longer exposure times on larger telescopes) than photometric references; (iii) slit losses (due to the wavelength dependence of the seeing) and atmospheric extinction are typically different when obtaining reference spectra and (iv) matching PSFs between new and reference observations is tricky—while PSF reconstruction on hyper-spectral data obtained with integral field spectrographs (IFS), which preserve two-dimensional (2D) spatial information, is feasible (Courbin et al. 2000; Bongard et al. 2011; Lezmy et al. 2022), it becomes extremely challenging in long-slit spectroscopy, the most widely used technique, where one of the two spatial dimensions is inherently lost.

Classic long-slit data reduction process usually treats transients as isolated sources. The background—comprising both sky emission (e.g., atmospheric emission, solar system and extragalactic background) and host galaxy light—is assumed to be smooth near the transient with little fluctuation and estimated by interpolating the flux measured beside the transient’s trace in the 2D spectrum, using regions outside the trace that are selected around it. However, the smoothness assumption no longer holds when the host environments have steep light profiles (e.g., galactic nuclei or bulges) or complex substructures (e.g., spiral arms or clumpy star-forming regions), common for extragalactic transients like SNe. Depending on the local light profile and regions chosen for background estimation, the classic technique can lead to either over- or under-subtraction of the host galaxy light. This introduces unpredictable systematic biases into the extracted transient spectrum, which cannot be reliably captured by standard uncertainty estimates. More advanced techniques can provide cleaner host galaxy subtraction in long-slit spectroscopy, but they typically require either firm knowledge of PSFs (e.g., Lucy & Walsh 2003; Blondin et al. 2005), or concurrent multi-band photometry with the spectrum, as well as host galaxy SED models (e.g., Ellis et al. 2008; Foley et al. 2008). In practice, these requirements are often difficult to meet.

Motivated by the success of image subtraction technique in photometry, we present a novel approach for host subtraction in long-slit spectroscopy. It utilizes archival images from all-sky imaging surveys that span the entire optical wavelength range, including the Sloan Digital Sky Survey (SDSS; York et al. 2000, *ugriz*), the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS, or PS; Chambers et al. 2016, *grizy*), the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys (LS; Dey et al. 2019, *griz*), and in the

near future, the Legacy Survey of Space and Time (Ivezić et al. 2019, *ugrizy*). From these multi-band images, we can construct a reasonable prior of the galaxy profile in a data-driven way via Gaussian processes (GPs) to avoid expensive SED fitting. Our software implementation, `HostSub_GP`, is a scalable toolkit built upon `tinygp` (Foreman-Mackey et al. 2024), leveraging the just-in-time compilation and automatic differentiation in `JAX` (Bradbury et al. 2018). The open source software is available at https://github.com/slowdivePTG/HostSub_GP with an MIT license.

In Section 2, we introduce the methodology of 2D galaxy spectrum modeling with GPs, followed by an elaboration of the pipeline modules in Section 3. We show that the GP method significantly outperforms the classic method based on linear interpolation in Section 4 by testing it on a synthetic dataset (Section 4.1) and two real transients (Sections 4.2 and 4.3). In Section 5, we summarize the recommended use cases for applying `HostSub_GP` following the main assumptions we have made, before presenting conclusions in Section 6.

2. Methodology: Modeling the 2D Spectrum with GPs

The ultimate goal of `HostSub_GP` is to estimate the galaxy light at each pixel of a 2D spectrum. The observed flux⁴ $f_{\text{obs}}(x, \lambda)$ is the function of a spatial coordinate x and the wavelength λ of each pixel, which consists of the contribution from the source $f_{\text{src}}(x, \lambda)$, its host galaxy $f_{\text{host}}(x, \lambda)$, and the sky background $f_{\text{sky}}(\lambda)$. Here we assume a constant sky background along the slit (typically $\lesssim 1'$ long). In a 2D spectrum with flat, bias, and dark fields calibrated, f_{sky} is only dependent on wavelength. We also assume that the atmospheric differential refraction has been corrected during the observation, either by observing at the parallactic angle or using an atmospheric dispersion corrector.

The transient, as a point source, contributes negligibly beyond a few PSF widths from its center. To estimate $f_{\text{host}} + f_{\text{sky}}$, we mask the transient’s trace and model the host and sky contributions in the surrounding regions. This allows us to infer the background flux blended with the transient.

Inspired by Lucy & Walsh (2003), f_{host} could be modeled as two semi-separable parts,

$$f_{\text{host}}(x, \lambda) = F_{\text{host}}(\lambda) \cdot \xi_{\text{host}}(x, \lambda). \quad (1)$$

Here, $F_{\text{host}}(\lambda)$, which is effectively a 1D spectrum of the galaxy, only has wavelength dependence, and $\xi_{\text{host}}(x, \lambda)$ is a normalized observed spatial profile, i.e., the intrinsic galaxy profile convolved with the PSF. We do not require the galaxy background to be homogeneous, i.e., ξ_{host} can be wavelength dependent, as the stellar population could be different across the region covered by the slit. Nevertheless, we expect ξ_{host} to vary slowly along the wavelength axis, whereas F_{host} can vary

⁴ Throughout this paper, uppercase F denotes the flux in a 1D spectrum, while lowercase f denotes the flux in a 2D spectrum (spectrogram).

on shorter scales (typically the instrumental resolution). Hence we need orders-of-magnitude fewer data points along the wavelength direction to model ξ_{host} . In addition, multi-band pre-transient imaging from archival surveys provides critical prior knowledge of ξ_{host} that we will integrate into our model.⁵

In practice, neither f_{host} nor ξ_{host} is directly measurable from the 2D spectrum, which is also blended with the sky background. To remove the sky emission, we define the global-sky-subtracted 2D spectrum f_{sub} as

$$\begin{aligned} f_{\text{sub}}(x, \lambda) &\equiv f_{\text{obs}}(x, \lambda) - \langle f_{\text{obs}} \rangle_{\text{sky}} \\ &= f_{\text{src}}(x, \lambda) + F_{\text{sub}}(\lambda) \cdot \xi_{\text{sub}}(x, \lambda), \end{aligned} \quad (2)$$

where f_{src} is considered negligible outside the trace mask, and

$$\xi_{\text{sub}}(x, \lambda) \equiv \xi_{\text{host}}(x, \lambda) - \langle \xi_{\text{host}} \rangle_{\text{sky}}(\lambda). \quad (3)$$

Here $\langle \cdot \rangle_{\text{sky}}$ denotes averaging the 2D quantity along the spatial axis over what we define as the global sky region with a width of L_{sky} , e.g., for ξ_{host}

$$\langle \xi_{\text{host}} \rangle_{\text{sky}}(\lambda) = \frac{1}{L_{\text{sky}}} \int_{\text{sky}} \xi_{\text{host}}(x, \lambda) dx, \quad (4)$$

resulting in a function of wavelength. Instead of modeling ξ_{host} directly, we estimate ξ_{sub} . For simplicity, we force ξ_{sub} to be normalized over what we define as the host region,

$$\int_{\text{host}} \xi_{\text{sub}}(x, \lambda) dx = 1, \quad (5)$$

so the physical significance of F_{sub} is the 1D spectrum extracted from a specific host region on the sky-subtracted 2D spectrum f_{sub} . It can be either positive or negative, depending on the relative brightness of the host and the sky regions. The selection of the two regions is very flexible, as long as (i) both regions are outside the trace mask; (ii) the average flux in the sky region is significantly different (either higher or lower) than the host region at all wavelengths; otherwise, ξ_{sub} will diverge when F_{sub} approaches 0 at any wavelength.

To bridge the archival images with the spectroscopic observation, we model both F_{sub} and ξ_{sub} with GPs, and implement the prior derived from the archival images as the mean function. In practice, we build a smooth model for F_{sub} using a 1D GP (notations with hats denote estimators),

$$\widehat{F}_{\text{sub}}(\lambda) \sim \mathcal{GP}(\mu_{1\text{D}}, K_1(\lambda, \lambda^*)), \quad (6)$$

which is conditioned on the observed values of F_{sub} . Here $\mu_{1\text{D}}$ is a mean value of the GP and K_1 is the kernel which determines the covariance between different wavelengths. Both $\mu_{1\text{D}}$ and the kernel parameters of K_1 are hyperparameters of the GP model. Separately, we build a smooth model for ξ_{sub}

with a 2D GP,

$$\widehat{\xi}_{\text{sub}}(x, \lambda) \sim \mathcal{GP}(\xi_{\text{sub, img}}(x, \lambda), K_2((x, \lambda), (x^*, \lambda^*))), \quad (7)$$

which is conditioned on the measured ξ_{sub} in the host region. Unlike the 1D GP which adopts a single, constant mean function $\mu_{1\text{D}}$, the mean function of the 2D GP is derived from the (multi-band) archival images (see Section 3.2), which is a function of both the spatial coordinates and wavelength. This mean function serves as a reasonable prior of the galaxy light profile along the slit, and, as we demonstrate below, will dramatically improve the precision in estimating the host galaxy background. K_2 is the corresponding kernel which determines the covariance between (x, λ) pairs.

Finally, we derive a model of f_{sub} as the product of \widehat{F}_{sub} and $\widehat{\xi}_{\text{sub}}$, so the estimated 2D source flux is simply

$$\widehat{f}_{\text{sub}}(x, \lambda) = \widehat{F}_{\text{sub}}(\lambda) \cdot \widehat{\xi}_{\text{sub}}(x, \lambda). \quad (8)$$

Comparing \widehat{f}_{sub} to f_{sub} yields the model likelihood, and the hyperparameters of the GP ($\mu_{1\text{D}}$ and kernel parameters) are optimized by maximizing the posterior. The implementation details of the methodology are elaborated in Section 3.3.

The output of `HostSub_GP` is the host galaxy subtracted 2D spectrum,

$$\widehat{f}_{\text{src}}(x, \lambda) = f_{\text{sub}}(x, \lambda) - \widehat{f}_{\text{sub}}(x, \lambda). \quad (9)$$

The 1D spectrum can be extracted from the 2D trace using existing pipelines, such as `iraf` (Tody 1986) and `PyPeIt` (Prochaska et al. 2020).

3. Pipeline Overview

In this section, we provide an overview of the `HostSub_GP` framework and, for each of the modules, we elaborate the details of the data reduction and modeling procedures. Figure 1 is a high-level summary of the pipeline.

3.1. Preprocessing

As input `HostSub_GP` requires a fully calibrated 2D spectrum and a configuration file containing user-defined parameters. Basic image calibrations, including the flat, bias, and dark field corrections, bad pixel rejection, cosmic ray (CR) removal, and 2D wavelength solution (i.e., the corresponding wavelength at each given pixel in the 2D spectrum), should be done in the upstream data reduction pipeline (e.g., `PyPeIt`). Flux calibration is not a requirement, because `HostSub_GP` does not model the SED of the host galaxies.

Additionally, it is expected that a first round of object finding and tracing has been carried out in the upstream pipeline, so `HostSub_GP` can locate the trace of the transient at each wavelength $(x_{\text{tr}}(\lambda), y_{\text{tr}}(\lambda))$. Here x and y roughly correspond to the spatial and spectral axes of the 2D spectrum, although the actual coordinate system is slightly curved. The

⁵ The resolving power of archival multi-band imaging is too low to serve as a useful prior for F_{sub} without SED modeling.

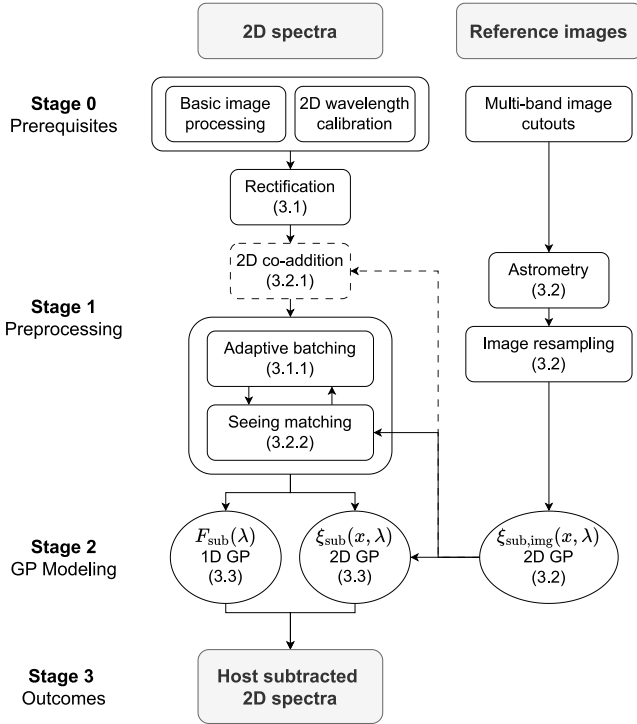


Figure 1. The data reduction framework in `HostSub_GP`. The 2D co-addition will only be performed if multiple exposures exist.

estimate can be further fine tuned by aligning the galaxy light in the spectrum to the archival galaxy images, as detailed in Section 3.2.1. The pipeline then calculates the sky separation from each pixel to the trace at the corresponding wavelength. Finally, the calibrated 2D spectrum (both the flux and the inverse variance of the measurements) is mapped onto a rectified grid of spatial (the sky separation from the transient in arcsecond) and spectral coordinates (the wavelength in Angstrom), to eliminate the curvature of the trace.

On the rectified spectrum, as illustrated in Section 2, users will need to specify three regions, each corresponding to a range of spatial distances to the trace of the transient: (i) a central mask, which contains the transient and is excluded in the modeling; (ii) the sky region, which is used to calculate the global sky $\langle f_{\text{obs}} \rangle_{\text{sky}}$ for each wavelength; and (iii) the host galaxy region, which is used to derive the 1D quantity F_{sub} and the 2D quantity ξ_{sub} that `HostSub_GP` actually models. Following rectification, each row of pixels corresponds to a fixed wavelength λ_j , so $\langle f_{\text{obs}} \rangle_{\text{sky}}(\lambda_j)$ is calculated as the mean flux in the sky region. After subtracting $\langle f_{\text{obs}} \rangle_{\text{sky}}$, $F_{\text{sub}}(\lambda_j)$ can be calculated as the sum of the flux in the row within the host galaxy region. Eventually, for every individual pixel i in the host galaxy region, $\xi_{\text{sub}}(x_i, \lambda_j)$ is simply the ratio of $f_{\text{sub}}(x_i, \lambda_j)$ and $F_{\text{sub}}(\lambda_j)$, as defined in Equation (3) and (5). An example of the 2D spectrum preprocessed by `HostSub_GP` is displayed in Figure 2.

Once F_{sub} and ξ_{sub} are separated, we note again that ξ_{sub} typically varies slowly along the spectral direction, i.e., the scale length is much greater than the instrumental spectral resolution. By default, `HostSub_GP` combines the pixels along the spectral direction by a factor of 64 (which can be modified by users in the configuration file) before ξ_{sub} is evaluated. The coarse binning (bin widths $\gtrsim 100 \text{ \AA}$) enhances the signal-to-noise ratio, while preserving the essential information, as the slow spectral variation of ξ_{sub} ensures that no significant details are lost.

3.1.1. Galaxy Emission Line Searching and Adaptive Batching

A notable exception where ξ_{sub} can fluctuate rapidly on the scale of the instrumental resolution occurs near the nebular emission lines of star-forming galaxies. The emission lines trace the youngest and most massive stars, whereas the spectral continuum and most absorption lines are shaped predominantly by the older stellar population. Since younger and older populations often show different distributions in galaxies, the spatial distribution of flux within the wavelength range of strong emission lines can be fairly different from the surrounding continuum. As a result, the emission lines require finer wavelength bins, whose widths correspond to the resolution of the spectrograph.

`HostSub_GP` automatically searches for prominent host galaxy emission lines and adjusts the bin size in an adaptive way. A candidate emission line will be labeled at wavelength λ_j if (i) $F_{\text{sub}}(\lambda_j)$ peaks significantly above the continuum; and (ii) ξ_{sub} calculated specifically at λ_j (i.e., without binning) deviates significantly from the surroundings. The line candidates are cross matched to a galaxy emission line library to determine the precise redshift and eliminate false positives resulting from CRs and/or sky lines which are imperfectly removed.

Once the locations of prominent galaxy lines are determined, the spectral bins are assigned following a few strategies: (i) the maximum bin size is specified by the user (or 64 by default); (ii) the minimum bin size is limited by the spectral resolution $\Delta\lambda$; (iii) the finest bins are centered on each galaxy emission line, with bin sizes increasing progressively farther from the lines; and (iv) the widths of adjacent bins cannot differ by more than a factor of two. An example of the adaptive binning is shown in the middle panel of Figure 2.

3.2. Building Priors from Archival Images

Wide-field imaging surveys have visited almost every corner in the sky with multiple filters. `HostSub_GP` uses these images to build a smooth model of the host galaxy light distribution along the slit, which serves as the prior in modeling the 2D spectrum.

`HostSub_GP` automatically downloads image cutouts covering the transient and its host galaxy from a couple of

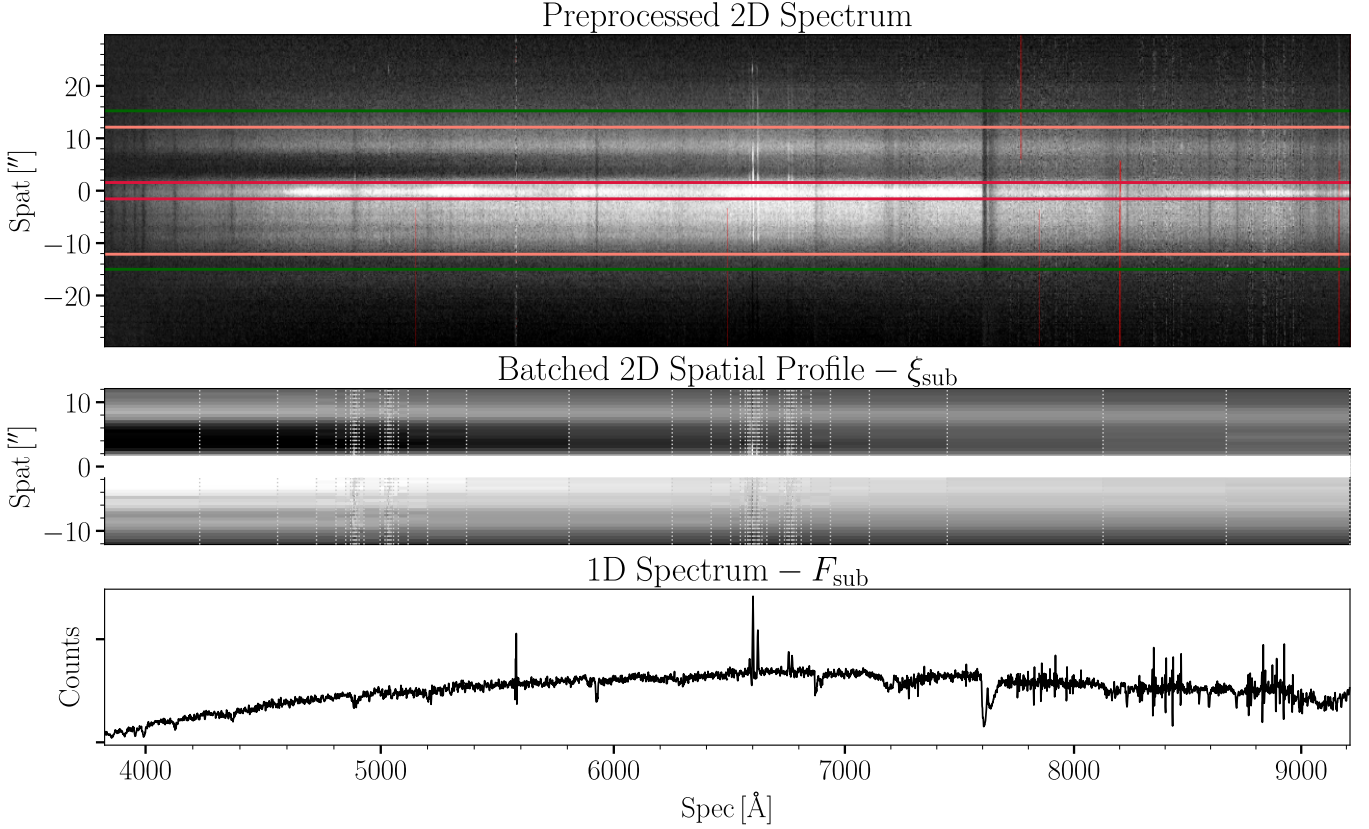


Figure 2. An example of the derived F_{sub} and ξ_{sub} from the preprocessed data with adaptive binning around galaxy emission lines. Upper panel: the example preprocessed 2D spectrum after co-addition, rectification, and global background subtraction. The edges of the sky regions (green), host regions (salmon), and transient mask (red) are displayed. The bad pixels are indicated with red colors. Middle panel: the batched 2D galaxy profile ξ_{sub} . Dashed gray lines indicate batch edges. Bottom panel: the 1D galaxy spectrum F_{sub} used to normalize the galaxy profile.

different surveys, including PS, LS, and the shallower SDSS. Users are also encouraged to load their own higher-quality images. Images are processed via Astrometry.net (Lang et al. 2010) for precise astrometric calibration. Then they are rotated and resampled onto a grid centered at the location of the transient and oriented along the same direction of the slit when the spectrum is taken. The resampling is conducted using the `reproject_adaptive` function of the `reproject` package (Robitaille et al. 2024), which carries out anti-aliased reprojection with the DeForest (2004) algorithm. Using pixels covered by the slit of a specific width, `HostSub_GP` derives the mock spatial distribution of the galaxy flux along the slit. And by subtracting the mean flux in the same sky region defined in preprocessing the 2D spectrum (Equation (3)) and normalizing the flux residual (Equation (5)), `HostSub_GP` obtains values of $\xi_{\text{sub,img}}$ at a few grid points (x_i, λ_j) , effectively a transmission-weighted average of the spatial profile ξ_{sub} , serving as an approximation to ξ_{sub} over the wavelength range covered by the filters. Notably, the spatial coordinates x_i depend on the pixel scales in the archival images, and we adopt the effective wavelength of each filter as a representative λ_j . As there are at

most a handful of broad-band filters available, the grid is expected to be coarse.

To build a smooth galaxy profile model out of the coarse grid, we leverage the 2D GP by assuming

$$\xi_{\text{sub,img}} \sim \mathcal{GP}(\mu_{\text{img}}, K_{\text{img}}((x, \lambda), (x^*, \lambda^*))), \quad (10)$$

where μ_{img} is an unknown constant mean function, and K_{img} is a 2D kernel parameterized as the combination of Matérn-3/2 kernel (K_{M32}) and exponential squared kernel (K_{SE}). With the same length scale l , the Matérn-3/2 covariance declines faster than the exponential squared covariance, so qualitatively the Matérn-3/2 kernel is better at capturing smaller scale fluctuations, whereas the exponential squared kernel better models any smooth, long term trend. The composite kernel is defined as⁶

$$K_{\text{img}} = (A_{x,r}K_{\text{M32}}(x, x^*; l_{x,r}) + A_{x,s}K_{\text{SE}}(x, x^*; l_{x,s})) \\ \times (A_{\lambda,r}K_{\text{M32}}(\lambda, \lambda^*; l_{\lambda,r}) + A_{\lambda,s}K_{\text{SE}}(\lambda, \lambda^*; l_{\lambda,s})), \quad (11)$$

⁶ In each bracket, the parameters before the semicolon denote the coordinates, while those after the semicolon denote the kernel hyperparameters.

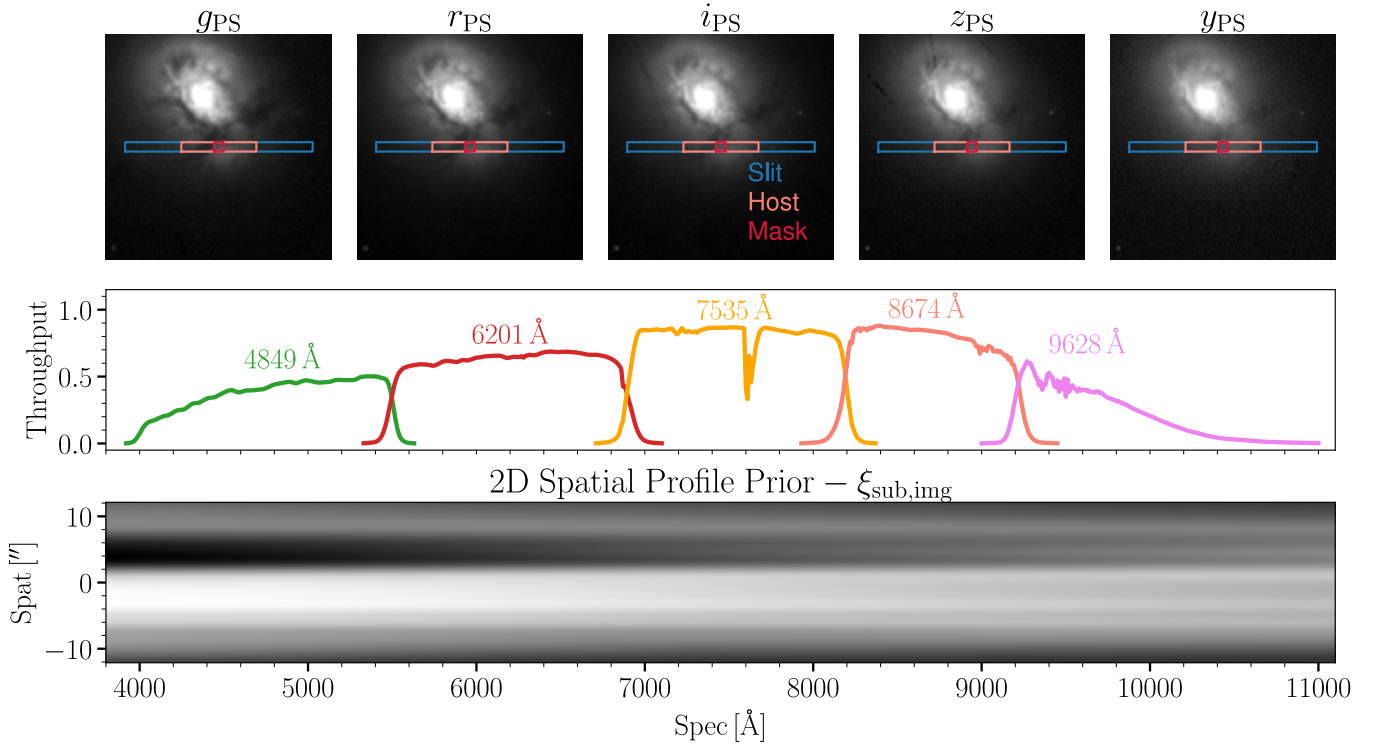


Figure 3. An example of constructing $\xi_{\text{sub,img}}$, which we use as the mean function in modeling ξ_{sub} with a GP (Equation (7)). Top panels: example archival PS images in *grizy* filters. The position of a $1'$ -long slit is displayed as the blue patch. A $24''$ -wide host region and a transient mask with a width of twice the seeing are indicated as the salmon and red patches, respectively. Middle panel: the throughput function of each filter and the corresponding effective wavelength. Bottom panel: the smooth $\xi_{\text{sub,img}}$ built with the galaxy light profiles (within the host region) at five discrete wavelengths with a 2D GP.

where $l_{x,r}$ and $l_{x,s}$ are the length scales on the spatial axis, $l_{\lambda,r}$ and $l_{\lambda,s}$ are the length scales on the spectral axis, while $A_{x,r}$, $A_{x,s}$, $A_{\lambda,r}$, and $A_{\lambda,s}$ parameterize the amplitudes and relative strength of the kernels.⁷ Typically, we expect $l_{x,r} \ll l_{x,s}$ and $l_{\lambda,r} \ll l_{\lambda,s}$, so both the rapid (r) and slow (s) fluctuations on both the spatial and spectral directions can be well monitored.

`HostSub_GP` allows users to impose reasonable constraints on the GP hyperparameters, particularly the length scale. While the initial values and allowed ranges of all these kernel parameters could be specified in the configuration file, by default, we initialize $l_{\lambda,r}$ and $l_{\lambda,s}$ to be 10^4 \AA and 10^5 \AA , respectively, both with a lower bound of 10^3 \AA (which could be adjusted to match the wavelength range of the broad-band filters). Both $l_{x,r}$ and $l_{x,s}$ are initialized to be the typical seeing of the archival images, as any real flux fluctuation along the spatial direction with a length scale shorter than the spatial resolution Δx , determined by the PSF, would have been smeared out and could not be physically meaningful. In ground-based observations, the PSF is primarily determined by atmospheric seeing. By default, `HostSub_GP` enforces a lower bound on the kernel length scale, requiring

⁷ For simplicity, we omit the subscript “img” from all kernel parameters and adopt the same convention for the kernels defined in Equations (6) and (7); note that these parameters are independent of one another.

$l_{x,r}, l_{x,s} > \Delta x/2.355$, and an additional upper bound on $l_{x,r}$ of $2\Delta x$. The factor of 2.355 converts the full width at half maximum (FWHM) of the PSF, the typical astronomical measure of seeing, into the standard deviation of a Gaussian profile. This constraint effectively prevents the model from interpreting random noise as high-frequency spatial fluctuations. Figure 3 shows an example of constructing $\xi_{\text{sub,img}}$ from archival PS images.

`HostSub_GP` implements the GP framework of `tinygp` (Foreman-Mackey et al. 2024), a lightweight and flexible GP library built on `JAX` (Bradbury et al. 2018), which accelerates numerical computations on Python arrays through just-in-time compilation, automatic differentiation, and GPU acceleration. The model hyperparameters are inferred by maximizing the marginal log-likelihood, corresponding to the log-probability of the GP conditioned on the discrete measurements of $\xi_{\text{sub,img}}(x_i, \lambda_j)$, with gradient descent powered by `JAXopt` (Blondel et al. 2021).

3.2.1. Trace Alignment and 2D Co-addition

For faint transients embedded within bright host galaxies, direct object tracing in the original 2D spectrum is often unreliable. To address this, `HostSub_GP` may also take the trace of a bright, isolated reference source (e.g., a standard

star), placed at the same location within the slit and observed with the same instrumental configuration, as an initial estimate of the trace. However, residual misalignment up to a few pixels can remain. `HostSub_GP` therefore refines the trace position by aligning the 2D spectrum to the spatial profile of the galaxy as modeled from archival imaging. On the raw 2D spectrum, the spatial profile is calculated as the flux-weighted, sigma-clipped average of the observed flux across a series of spatial bins in the host region. The pipeline then evaluates the imaging prior $\xi_{\text{sub,img}}(x - x_{\text{off}}, \bar{\lambda})$ at the flux-weighted mean wavelength $\bar{\lambda}$ of the 2D spectrum and searches a fine grid of trial offsets x_{off} for the value that maximizes the Pearson correlation between the two profiles. The resulting $x_{\text{off,opt}}$ is applied as the trace correction, yielding sub-pixel alignment which substantially improves subsequent host modeling.

The pipeline provides an option to co-add multiple rectified 2D spectra from individual exposures, as requested by users, to enhance signal-to-noise ratio. Prior to the co-addition, the median value of each row is subtracted from every 2D spectrum as a preliminary global sky subtraction, which helps to minimize background variations between exposures. The pixel-wise CR rejection is also refined by contrasting multiple frames. The co-addition is performed as an inverse-variance weighted mean at each pixel, using only valid, unmasked data. The propagated error is computed accordingly from the weighted variances.

3.2.2. Seeing Matching

To match the 2D spectrum and the galaxy profile prior even more precisely, we also need to ensure that the spatial resolutions, determined by the seeing during the observations, are comparable. This is similar to the standard image subtraction, where the PSFs of the (new) science image and the reference image need to be precisely matched, typically by convolving an appropriate kernel (with an FWHM of Δx_{kernel}) to degrade the spatial resolution of the sharper image (e.g., Alard & Lupton 1998),

$$\Delta x_{\text{final}}^2 = \Delta x_{\text{kernel}}^2 + \Delta x_{\text{init}}^2. \quad (12)$$

Here Δx_{init} and Δx_{final} denote the initial and final seeing of the images, respectively. Unlike imaging data, a 2D spectrum contains only one spatial dimension which can be used to decode the shape of the PSF. The PSF size also has a substantial negative wavelength dependence (Boyd 1978). Recovering the exact shape of the PSF in a 2D spectrum is particularly difficult. In this section, we show how the pipeline empirically matches the seeing of the 2D spectrum to that of the priors built with the archival images by minimizing the difference between the spectrum and the image.

If the seeing in the archival images is worse than in the spectroscopic data, the seeing matching is achieved by convolving the observed spectrum with 1D Gaussian kernels

with a range of widths corresponding to possible seeing differences. For each case, the length scale of the kernel has a negative dependence on the wavelength, which can be approximated as a power law, $\sigma_{\text{kernel}} \propto \lambda^{-\alpha}$.⁸ The host profile in the convolved spectrum is then compared to the imaging prior, yielding the χ^2 statistic, equivalent to the negative Gaussian log-likelihood. The optimal σ_{kernel} and α are inferred by minimizing the χ^2 , and then applied to the 2D spectrum. The allowed range of α is between 0.2 and 0.5. After the matching process, ξ_{sub} is re-evaluated on the convolved spectrum, and the pipeline proceeds to the main GP modeling.

Conversely, if the archival images have higher spatial resolution than the spectroscopic data, we match the seeing by degrading the archival images with 2D Gaussian kernels, again parameterized by σ_{kernel} and α . For a $(\sigma_{\text{kernel}}, \alpha)$ pair, kernel widths are computed for each image based on their effective wavelength, and the images are convolved with 2D Gaussian kernels of the corresponding widths. The pipeline then reconstructs the imaging model as described in Section 3.2, and similarly, the optimal σ_{kernel} and α are determined by minimizing the χ^2 .

We note that our seeing matching process is likely an oversimplification. For example, since the spectroscopic and imaging data are usually collected at different observatories, their wavelength dependencies of seeing (i.e., α) may differ. The goal of the seeing matching is not to build a perfect prior that exactly reproduces the spatial profile of the 2D spectrum, but rather to make the prior sufficiently similar at a reasonable computational cost. Our GP framework, illustrated in Section 3.3, is then well suited for modeling the residuals, which are usually highly correlated.

3.3. Main Process: GP Modeling

In the main module of the pipeline, we model F_{sub} and ξ_{sub} in a sequential way. In the first stage, `HostSub_GP` builds a preliminary model for F_{sub} with a 1D GP following Equation (6). Between a pair of wavelengths (λ, λ^*) , the kernel is defined as a combination of a Matérn-3/2 kernel and a Matérn-5/2 kernel,

$$K_{\text{l}} = A_{\text{r}} K_{\text{M}3/2}(\lambda, \lambda^*; l_{\lambda,\text{r}}) + A_{\text{s}} K_{\text{M}5/2}(\lambda, \lambda^*; l_{\lambda,\text{s}}), \quad (13)$$

which captures both small-scale fluctuations near narrow nebular emission and absorption lines, as well as large-scale trends across the continuum. To enable scalable GP inference, we adopt quasi-separable kernels implemented in `tinygp`,⁹

⁸ We expect $\alpha = 0.2$ if the atmospheric blurring is dominated by Kolmogorov turbulence (Fried 1965; Woolf 1982), but steeper dependencies can arise due to deviations from the Kolmogorov theory (e.g., local turbulence near the dome).

⁹ The exponential squared kernel is not quasi-separable, so we use the Matérn-5/2 kernel implementation in the `tinygp.quasiseq` package, which is smoother than the Matérn-3/2 kernel, to characterize large-scale variations.

reducing the computational complexity from $\mathcal{O}(N^3)$ to $\mathcal{O}(N)$ on 1D datasets (Eidelman & Gohberg 1999; Foreman-Mackey et al. 2017). By default, both $l_{\lambda,r}$ and $l_{\lambda,s}$ are forced to be greater than the spectral resolution $\Delta\lambda/2.355$. For the more rapidly varying kernel, an upper bound of $2\Delta\lambda$ is imposed on $l_{\lambda,r}$. The point estimates of the 1D GP hyperparameters are then obtained by maximizing the marginal likelihood.

Once a reasonable model for F_{sub} is obtained, the pipeline proceeds to a second stage where both the 1D F_{sub} and the 2D ξ_{sub} are modeled jointly. The optimal hyperparameters derived in the first stage are adopted as initial values for this round. The mean function $\xi_{\text{sub,img}}$ in Equation (7) has been described in detail in Section 3.2. The covariance matrix is constructed using a 2D Matérn-3/2 kernel

$$K_2 = A_{M32} K_{M32}((x, \lambda), (x^*, \lambda^*); l_x, l_\lambda). \quad (14)$$

As in the previous GP modeling step, `HostSub_GP` imposes lower bounds on the kernel length scales, requiring $l_x > \Delta x/2.355$ to prevent the model from fitting noise fluctuations that occur below the spatial resolution limits. Similar as that in defining the prior kernel in Section 3.2, by default, we set the lower bound of l_λ to be 10^3 \AA .

As discussed in Sections 3.1.1, $\xi_{\text{sub,img}}$ is typically not an adequate approximation of ξ_{sub} in the vicinity the nebular emission lines from the host galaxy, where abrupt jumps in ξ_{sub} can occur along the spectral axis (see also Figure 2). In such regions, the covariance structure of ξ_{sub} should account for the local discontinuities, with reduced correlation across wavelengths that straddle an emission line feature. Conversely, when both pixels lie within the vicinity of the same emission line, the covariance between their corresponding values of ξ_{sub} may be enhanced. This reflects the empirical observation that the spatial structure of the galaxy profile is often more variable near nebular emission lines, owing to the clumpy and irregular distribution of star-forming regions compared to the smoother profile associated with older stellar populations. To account for these discontinuities, we introduce a custom multiplicative kernel term k_{corr} . For each emission line centered at λ_i , we define a proximity function using a hyperbolic tangent window,

$$w_i(\lambda; ,l_\lambda) = \frac{1}{2} \left[\tanh\left(\frac{2(\lambda - \lambda_i + l_\lambda/2)}{l_\lambda}\right) - \tanh\left(\frac{2(\lambda - \lambda_i - l_\lambda/2)}{l_\lambda}\right) \right], \quad (15)$$

which serves as a differentiable approximation to the indicator function $\mathbf{1}[|\lambda - \lambda_i| < l_\lambda/2]$. This formulation enables the use of automatic differentiation for optimizing the width parameter l_λ , which sets the effective width influenced by the emission line. In the optimization, the width parameter is also required to be greater than $\Delta\lambda$. We then define two auxiliary functions

W_{one} and W_{both} :

$$W_{\text{one},i} = w_i(\lambda; l_\lambda) \cdot [1 - w_i(\lambda^*; l_\lambda)] + [1 - w_i(\lambda; l_\lambda)] \cdot w_i(\lambda^*; l_\lambda) \quad (16)$$

which attains values close to 1 if and only if exactly one of the two points lies near the emission line; and

$$W_{\text{both},i} = w_i(\lambda; l_\lambda) \cdot w_i(\lambda^*; l_\lambda) \quad (17)$$

which approaches 1 only when both points are within the vicinity of the same line. The emission line correction factor is defined as

$$k_{\text{corr}} = \prod_i [1 - W_{\text{one},i}(\lambda, \lambda^*; l_\lambda) + A_1 \cdot W_{\text{both},i}(\lambda, \lambda^*; l_\lambda)], \quad (18)$$

where $A_1 > 0$ controls the amplitude of the enhancement when both points are near the same emission feature. For all emission lines, A_1 is assumed to be the same. Eventually, the 2D covariance kernel is corrected as

$$K'_2 = k_{\text{corr}}(\lambda, \lambda^*; A_1, l_\lambda) \cdot K_2. \quad (19)$$

The correction term k_{corr} , as well as the adaptive batching strategy, effectively isolates the modeling of ξ_{sub} inside and outside the emission line regions, and helps to prevent the GP from being influenced by discontinuities at emission lines when modeling the smoother continuum regions.

The total log-likelihood to be maximized consists of the marginal log-likelihood functions from the GPs, along with a Gaussian log-likelihood term based on the residuals between the modeled and observed flux, \hat{f}_{sub} and f_{sub} , summed over all pixels

$$\ln \mathcal{L} = \ln \mathcal{L}_{GP,1} + \ln \mathcal{L}_{GP,2} - \sum_{x_i, \lambda_j} \frac{(\hat{f}_{\text{sub}} - f_{\text{sub}})^2}{2\sigma_{\text{sub}}^2}, \quad (20)$$

where $\mathcal{L}_{GP,1}$ and $\mathcal{L}_{GP,2}$ correspond to the marginal likelihood functions given the observed F_{sub} and ξ_{sub} under the GP priors (Aigrain & Foreman-Mackey 2023), as characterized by Equations (6) and (7), respectively, and σ_{sub} represents the per-pixel uncertainty in f_{sub} , assuming independent Gaussian noise. Since the log-likelihood is maximized under parameter constraints (e.g., lower bounds on kernel length scales and any other customized constraints imposed by the users), the resulting point estimates of kernel hyperparameters correspond to maximum a posteriori estimators under implicit priors.

Now that both the global and local sky background $\langle f_{\text{obs}} \rangle_{\text{sky}}$ and f_{sub} have been derived for the rectified 2D spectrum, the ultimate sky model $\langle f_{\text{obs}} \rangle_{\text{sky}} + f_{\text{sub}}$ is evaluated on each pixel on the original 2D spectrum as the final output of the pipeline. Downstream tasks, including 1D spectrum extraction and fluxing, may then be performed with external pipelines.

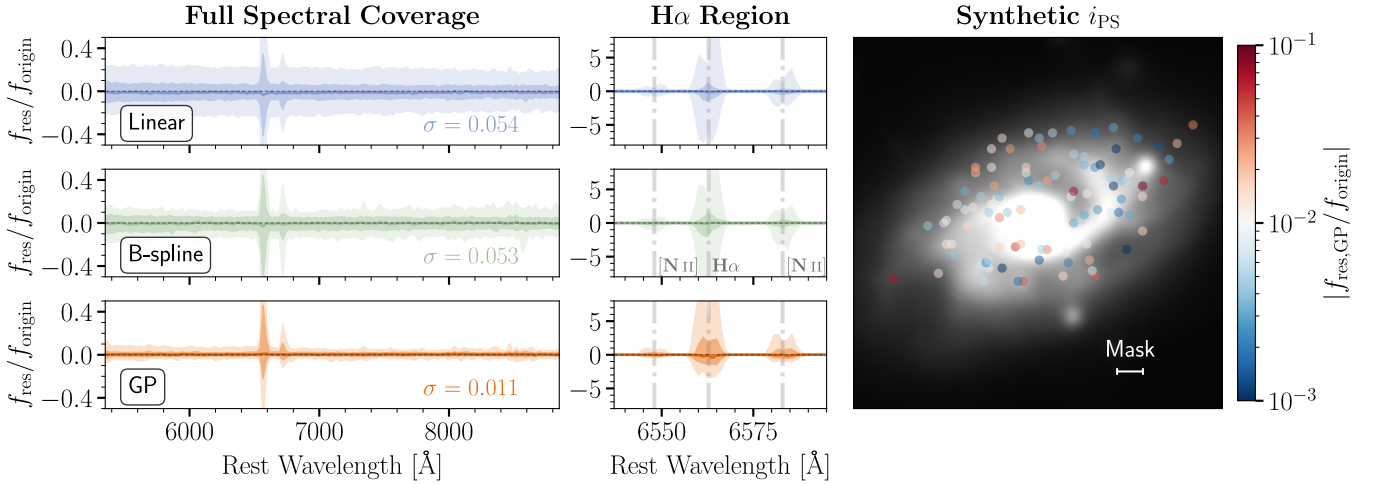


Figure 4. The GP method systematically improves the galaxy background estimates on the MUSE synthetic dataset of the spiral galaxy LCRS B110709.2–121854. The first two columns illustrate the distribution of fractional flux residuals (residuals normalized by the original galaxy flux in the PS i filters at randomly drawn slit locations) as a function of wavelengths in each of the 25 Å bins, resulting from the classic linear and B-spline interpolation methods, and the GP method. The first column covers the full wavelength range, whereas the second column zooms in on the region containing the strong H α and [N II] emission lines. Solid lines denote the median residuals across 100 random draws; the darker and lighter shaded regions represent the 68% and 95% percentiles, respectively. As a comparison, the black dashed line denotes the zero residual level. The σ values in the legends quantify the standard deviation of the 100 median fractional residuals, each computed across the full wavelength range for a single random draw. The rightmost panel shows the synthetic image in the i band convolved with a Gaussian kernel to downgrade the spatial resolution from 0.68 to 1.2. Overlaid are the locations of the randomly drawn slits, color coded by the corresponding absolute values of fractional residuals (median across all wavelengths) from the GP method. The typical aperture mask width (twice the seeing, 2.4) is indicated, though the mask width is determined case by case in the seeing matching.

4. Results

4.1. Test On a Synthetic Dataset

To evaluate the effectiveness of this new technique in improving host galaxy subtraction relative to the classic aperture spectroscopic method, we validate our pipeline using a large set of 2D synthetic galaxy spectra. Since these spectra do not include any transient, an ideal host subtraction method should produce zero flux residuals at all wavelengths. The synthetic 2D spectra are generated from a 3D hyper-spectral data cube acquired with the Multi Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010), which encodes the galaxy spectrum at every pixel within the field of view. By placing an artificial “slit” on the data cube, we could construct a 2D spectrum by aligning the 1D spectra from pixels on the slit.

We test our package on the archival MUSE data of LCRS B110709.2–121854 (PI: Tanvir), a galaxy containing a bright, red bulge (FWHM ~ 2.5) and well resolved spiral arms. We randomly place an artificial slit at 100 locations, with a width of 1". For simplicity, we only draw slits along the columns of the data cubes. The 2D spectrum is the column-average of the 1D spectrum for each row. Then a mask with an aperture size of twice the seeing placed on the 2D spectrum, where the mask center is also drawn randomly. Surrounding the mask is the host region to be modeled which is 10" wide. Note that the MUSE field is slightly distorted, and hence the alignment with the real archival images is non-trivial. We therefore build our

priors using the synthetic photometry in the PS r , i , and z filters by convolving the 3D data cube with the response curves of PS. The MUSE observations were conducted with a seeing of $\Delta x = 0.68$, much better than the typical seeing in PS1. To mimic the common situation where the seeing in the archival images is worse than that in the spectroscopic observations, we convolve a wavelength-dependent Gaussian kernel with the synthetic images. The FWHM of the Gaussian kernel is $\Delta x_{\text{kernel}} = 1''$ in the PS r band, and the corresponding downgraded seeing is ~ 1.2 . Adopting a wavelength dependence of $\lambda^{-0.2}$, the FWHMs of the kernels applied to the i and z bands are 0.96 and 0.93, respectively.

In Figure 4 we show the distribution of the median galaxy flux residuals as a function of wavelength in each of the 25 Å bins. The residual (f_{res}) is normalized by the galaxy background in the synthetic PS i band (f_{origin}), and should be considered as the fractional flux residual. Across the entire wavelength range, the GP method (orange) significantly improves the subtraction of the host galaxy continuum compared to the two classic methods, which estimates the galaxy background by interpolating the flux in the local sky regions surrounding the mask, either linearly (blue) or with a B-spline (green). In the linear-interpolation case, each part of the local sky region spans 1–2 Δx from the trace center. In the B-spline approach, we fit a third-order B-spline to a broader region, each part spanning 1–6 Δx from the trace center (or up to the slit edge), and set knot intervals to be 2 Δx . Statistically,

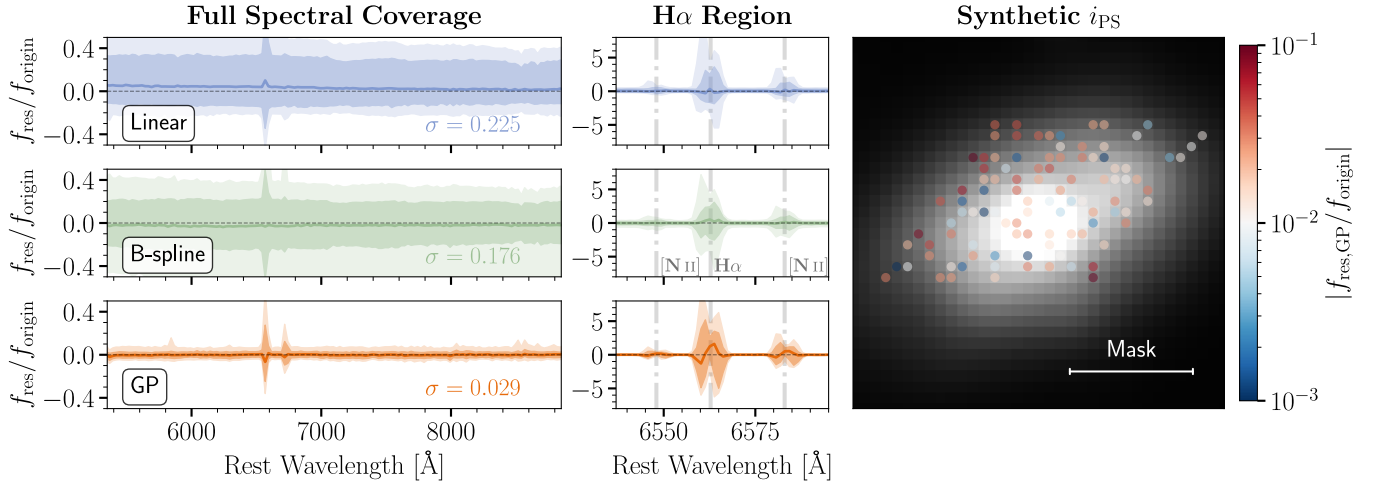


Figure 5. Same as Figure 4, but showing that the GP method still outperforms the classic methods in both the precision and accuracy on the synthetic dataset where the MUSE hyper-spectral data cube has been spatially binned by a factor of 4, simulating observations of a less resolved galaxy scaled to 25% of the angular size under the same seeing conditions. Again, seeings in the synthetic photometry on the rightmost panels have been downgraded to $\sim 1''.2$.

the GP method reduces the scatter of the fractional residuals by a factor of ~ 5 , and only 5% of the experiments have fractional residuals larger than 5%, all without custom hyperparameter tuning. In contrast, with the classic linear/B-spline subtraction method, 45%/37% of the experiments yield fractional residuals exceeding 5%. On the other hand, the GP method still cannot remove narrow galaxy emission lines (second column), but the strong over-/under-subtraction is limited to the immediate vicinity of the lines and does not affect the continuum modeling. The rightmost panel displays the locations of the slit centers, color coded with the corresponding absolute values of fractional flux residuals from the GP method. The fractional flux residuals exhibit no significant spatial pattern, whereas the absolute flux residuals are elevated near the galaxy center with the highest background.

We further assess the model performance when the galaxy is less resolved. In practice, we reduce the spatial resolution by performing 4×4 binning to the data cube, but retain the original pixel scale of 0.2 pixel^{-1} . Again, we convolve Gaussian kernels with the synthetic *riz* images to worsen the seeings to $\sim 1''.2$. The results are displayed in Figure 5. Effectively, in this test the galaxy spans only 25% of its original angular size. The bulge and spiral arms are no longer resolved and the FWHM of the galaxy outskirts is $\sim 2''$, only marginally resolved with the worsened seeing. We find that the scatter of the fractional residuals increases as the galaxy gets less resolved, but the GP method outperforms the classic methods even more significantly in this case, reducing the scatter by a factor of ~ 6 – 8 . More notably, the linear method yields substantial biases in flux residuals, with the median values tending to be positive, indicating systematic under-subtraction of galaxy background due to the overall curvature of the galaxy profile. In comparison, the GP method

effectively corrects for this bias. Eventually, the GP method hits the limit when the physical angular size of the galaxy is comparable with the seeing.

Following the success on the synthetic dataset, in the next sections we will apply our pipeline to two real transients observed with the Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995) on Keck, in which host galaxy removal critically affects our science interpretations of the spectra.

4.2. Revisiting a Nebular-phase SN: SN 2019eix

Hundreds of days after an SN explodes, the ejecta become optically thin. In this nebular phase, the spectra are dominated by the forbidden emission lines of the ions in the ejecta interior. Nebular-phase emission features uniquely probe the nucleosynthesis yields, ionization states, and kinematics in the innermost region of the SN ejecta, which are sensitive to the explosion channels and progenitor properties of both core-collapse and thermonuclear SNe.

However, SNe are naturally very faint in the nebular phase meaning they are often outshone by host galaxy light even in low-contamination environments. Here we revisit SN 2019eix, a peculiar SN with a vague classification (Padilla Gonzalez et al. 2023), and show that a nebular-phase spectrum with precise host galaxy subtraction provides smoking-gun evidence of its explosion mechanism and progenitor nature.

SN 2019eix is a hydrogen-poor SN that exploded near the bulge of NGC6695 ($z = 0.0183$) on its bar. SN 2019eix was originally classified as an SN Ic (Dimitriadis 2019), i.e., the explosion of a massive star with an envelope that has been stripped of hydrogen and helium. Yet Padilla Gonzalez et al. (2023) demonstrate that the spectral sequence of SN 2019eix are also consistent with some subluminous SNe Ia as outcomes

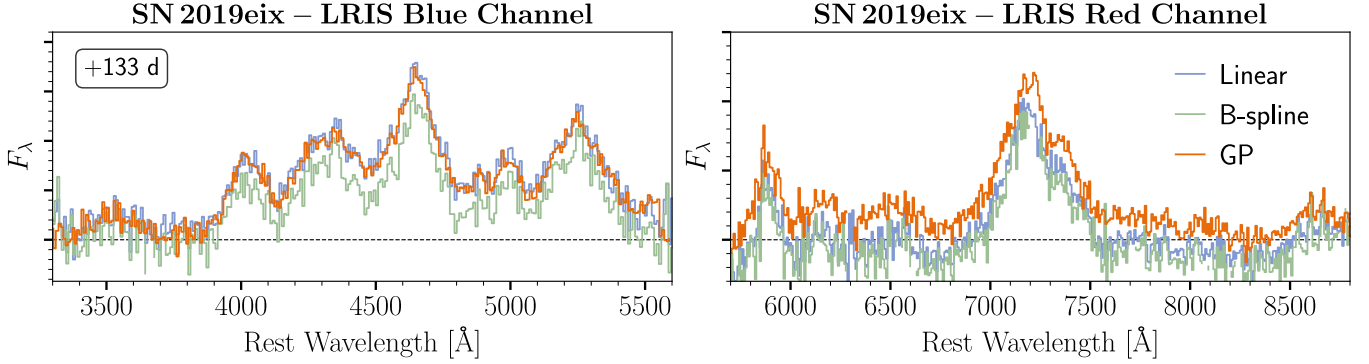


Figure 6. The GP method (orange) effectively removes the host galaxy contamination in the nebular-phase spectrum of SN 2019eix, whereas the spectrum extracted with the classic methods (blue and green) suffers an over-subtraction, which hinders detailed emission line modeling. All spectra have been smoothed and binned to a 7.5 \AA resolution. The horizontal dashed line indicates zero flux level. The flux scales are different between the two panels.

of an exploding white dwarf (WD), based on a dataset up to ~ 30 days after maximum luminosity. As is shown in the Figure 7 in Padilla Gonzalez et al. (2023), around the maximum luminosity of SN 2019eix, its spectral shape resembles both particular subluminous SNe Ia (e.g., SN 1991bg; Filippenko et al. 1992, SN 2016hmk; Galbany et al. 2019; Jacobson-Galán et al. 2020, and SN 2018byg; De et al. 2019), and SNe Ic (e.g., PTF12gzk; Ben-Ami et al. 2012). At this phase, differences between these peculiar events are subtle—on average SNe Ic exhibit stronger Si II/O I ratios than SNe Ia (Gal-Yam 2017; Sun & Gal-Yam 2017), but the distribution of this diagnostic is a continuum and typing can be uncertain for objects near the boundary.

About 133 days after maximum light, SN 2019eix was observed by Keck/LRIS (PI: Fremling). We show the spectrum in Figure 6, which is dominated by broad emission features. Notably, using classic interpolation-based methods¹⁰ (linear and B-spline), the galaxy background in the LRIS red channel is overestimated, so a substantial fraction of the spectrum shows negative flux.

In Figure 7 we show a comparison of SN 2019eix with a few other SNe I in the nebular phase. Despite their similarity near peak to SNe Ic, the nebular-phase spectra of these SNe Ia serve as a clear discriminant between the two classes. In both normal and subluminous SNe Ia, such as SN 2011fe and SN 1991bg, the nebular-phase spectra are dominated by forbidden emission lines of iron, nickel, and cobalt. On the contrary, the spectra of hydrogen-poor core-collapse SNe, such as PTF12gzk, are dominated by the [O I] doublet. As for SN 2019eix, prominent iron, nickel, and cobalt emission lines are revealed in the nebular-phase spectrum (+133 days) after careful host galaxy subtraction with the GP method, while there is no detectable

[O I], as expected in most SNe Ia. Additionally, the prominent line complex between 4000 and 5000 Å is another hallmark of SNe Ia. This emission complex, which is dominated by [Fe III], is not nearly as strong in core-collapse SNe (Figure 6). This confirms that SN 2019eix originates from a thermonuclear explosion of a WD.

While the classic subtraction methods can already reveal the major emission lines, the over-subtraction of galaxy light hinders accurate line intensity measurements, and more precise host galaxy subtraction is required to place stringent constraints on the progenitor properties. From the early-time observables, SN 2019eix has been suggested to be a peculiar SN Ia arising from the helium-shell double detonation of a sub-Chandrasekhar-mass WD (Padilla Gonzalez et al. 2023), which is independently supported by the low [Ni II]/[Fe II] ratio in the 7300 Å line complex (Maguire et al. 2018; Flörs et al. 2020; Blondin et al. 2022) from the host-subtracted nebular-phase spectrum presented here (Figure 7).

4.3. Revisiting a Nuclear Transient: AT2019qiz

Tidal disruption events (TDEs) originate from the disintegration of stars grazing the tidal radius of a super-massive black hole (SMBH) in their strong tidal field. They are unique probes of both the demographics of the SMBH properties (e.g., mass) and the nuclear stellar population. TDEs spotted by all-sky time-domain surveys are UV-bright transients featured by a blue continuum and broad hydrogen, helium, and nitrogen emission lines in the optical spectra, though a puzzling subclass of TDEs appear to be featureless throughout their evolution. Yet as nuclear transients, host galaxy subtraction is extremely challenging for spectroscopic observations of TDEs. Traditionally, a spectrum containing light from the TDE and the galaxy nucleus is first extracted, and then the host contamination is subtracted in the post-processing by fitting the joint spectrum as a combination of the galaxy spectrum (a rescaled archival spectrum of the galaxy, if it exists, or a synthetic spectral model), and a TDE continuum model

¹⁰ Pipelines such as `PyPeIt` also employ B-splines to model the local sky, and the results can be sensitive to the specific implementation, including how the local sky region is defined for the fit and how many knots are used, so it is tricky to recreate the methodology in other pipelines. Throughout this paper, we use the same implementations of both the linear and B-spline methods described in Section 4.1 for all comparisons.

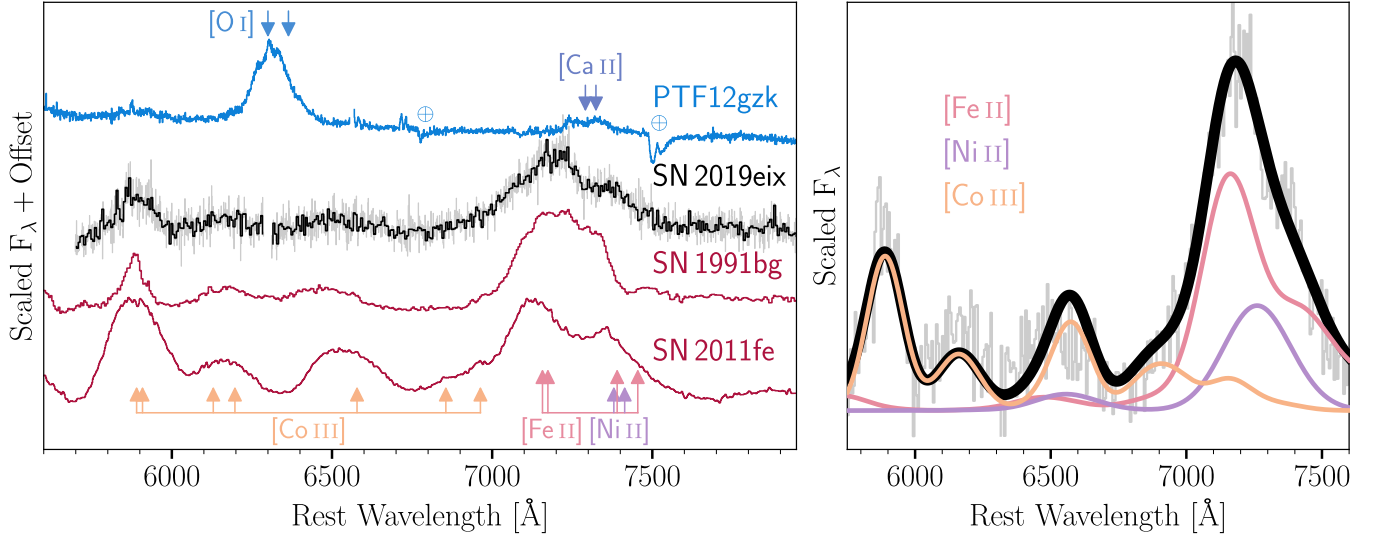


Figure 7. The nebular-phase spectrum of SN 2019eix reveals its sub-Chandrasekhar-mass WD progenitor. Left: the late-time spectrum of SN 2019eix highly resembles those of the normal SN Ia 2011fe (Stahl et al. 2020) and the sub-luminous SN Ia 1991bg (Filippenko et al. 1992) (both in maroon). The gray spectrum is the raw SN 2019eix spectrum with the host galaxy subtracted using the GP method, and the black spectrum is smoothed to the resolution of 5 Å. Within this wavelength range, spectra of SNe Ia are dominated by numerous [Fe II], [Ni II], [Co III] lines, where the strongest lines are indicated. On the other hand, the spectrum of the SN Ic, PTF12gzk (Ben-Ami et al. 2012) (blue) only shows [O I] $\lambda\lambda$ 6300, 6364 and [Ca II] $\lambda\lambda$ 7291, 7324. The telluric features in the PTF12gzk are indicated with \oplus . Right: modeling the 7300 Å complex of SN 2019eix, following the methodology in Liu et al. (2023a, 2023b), yields a low [Ni II]/[Fe II] ratio (~ 0.5), suggesting a sub-Chandrasekhar-mass WD origin. For a clean visualization we only display the smoothed spectrum in gray. The solid black curve is the joint model, whereas each colored curve corresponds to the contribution of each ion.

(blackbody or low-order polynomial). Empirical free parameters are introduced to handle the systematics due to the unknown slit loss in the long-slit spectroscopy, host galaxy extinction, etc. Our technique provides an alternative, self-consistent way of removing the host contamination for TDEs in nearby galaxies. Here we revisit the spectroscopy of AT 2019qiz to show how we are able to recover weak emission features of a TDE after careful host subtraction.

AT 2019qiz is a H+He TDE with prominent broad hydrogen and helium emission lines (Nicholl et al. 2020; Hung et al. 2021). Its optical spectral sequence is featured by the delayed appearance of narrow nebular lines of H and He II, as well as coronal lines of highly ionized Fe (Short et al. 2023). Hosted by a nearby galaxy ($z = 0.0151$) whose nucleus (FWHM $\sim 2''$) can be marginally resolved by ground-based facilities, AT 2019eix is an ideal object to test `HostSub_GP`. We re-reduced two Keck/LRIS spectra (PI: Foley) in Hung et al. (2021) obtained 48 days and 340 days after the maximum luminosity of the TDE, both displayed in Figure 8. The 2020 spectrum was obtained with a single, long (900 s) exposure and is heavily contaminated by CRs, so in Figure 8 we have omitted the affected pixels.¹¹

¹¹ Even if the trace of the transient is not directly impacted by CRs, if a significant portion of pixels in its vicinity is contaminated, the GP method will struggle to model the neighboring 2D spectrum. Therefore `HostSub_GP` is more sensitive to CRs than classic methods.

With the GP method, we dramatically reduce the level of host contamination compared to classic linear subtraction technique. In the 2019 spectrum, the Balmer break and the strong Ca H&K absorption lines are much better removed, reshaping the continuum between ~ 4000 and 5000 Å, where multiple broad H and He emission lines are present. This simplifies the measurement of the line flux, as an external galaxy spectral model is no longer required. In the 2020 spectrum, we uniquely detect (i) persistent broad He II $\lambda 4686$ emission; (ii) the emergence of narrow He II and H β emission—notably, using the classic subtraction technique, the narrow He II is not detected until the last epoch of the VLT/X-shooter spectrum from Short et al. (2023); (iii) the appearance of [Fe VII] $\lambda 6087$, and unambiguous non-detections for other coronal lines of higher ionization states (see Figure 9), by removing prominent line features from the galaxy and recovering smooth continua around lines of interest. It has been shown that these recombination lines keep brightening over time (Short et al. 2023), as the ionizing photons from the initial outburst reach material at larger radii. The fluxes of these recombination lines uniquely probe the gas properties surrounding the SMBH as well as the SED of the TDE outburst in the EUV and soft X-ray (~ 0.05 – 0.5 keV) range. With the GP method, we achieve more secure detections of these key recombination lines, enabling robust flux measurements (including tight upper limits for non-detections) that were previously hindered by host galaxy contamination. This

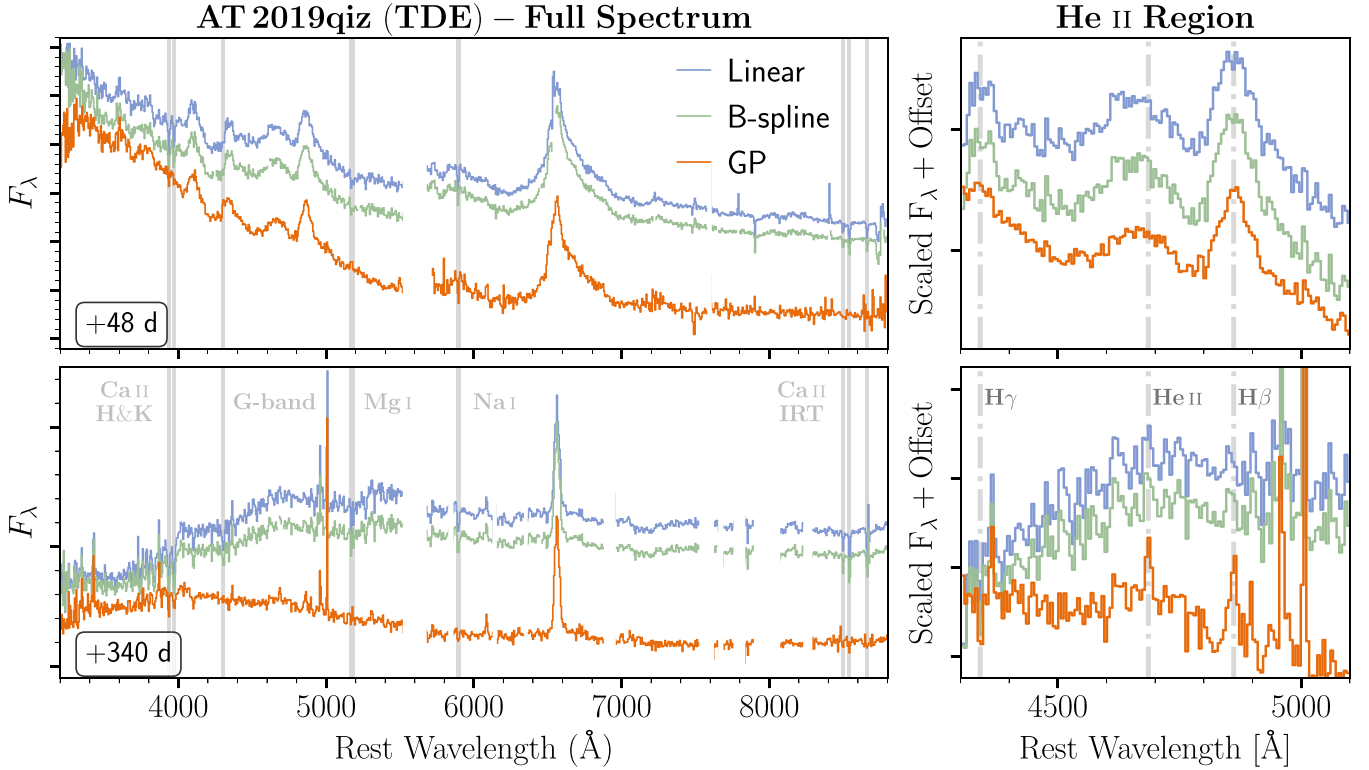


Figure 8. The GP method effectively removes the host galaxy contamination for AT 2019qiz, a bright TDE at the nucleus of a nearby galaxy. The left panel shows the full spectra (binned into 5 \AA), with some common galaxy absorption lines—the Ca II H&K doublet, G-band, the Na I D doublet, the Mg I $\lambda\lambda 5167, 5173, 5184$ triplet, and the Ca infrared triplet (IRT)—labeled in gray, all of which are better removed with the GP method. The right panel zooms in around the broad He II $\lambda 4686$ emission. With careful host galaxy subtraction using the GP method, we fix the underlying continuum beneath the He II $\lambda 4686$ and $H\beta$ emission in the 2019 spectrum (+48 days), critically for estimating their relative intensity, and uniquely show that they still persist in the 2020 spectrum (+340 days). We also probe the emerging narrow He II and $H\beta$ emission (dash-dotted lines) in the 2020 spectrum that is not obvious in the spectra reduced by the linear or B-spline method. The same galaxy absorption lines from Figure 6 are shown as vertical gray lines. Note that this spectrum consists of a single 900 s exposure, so the red side is heavily contaminated by CRs; we therefore manually masked the affected pixels. Emission lines of [O III] at 4363, 4959, and 5007 \AA are detectable in the 2020 spectrum, but are not labeled for clarity.

demonstrates a clear advance over standard host subtraction techniques for probing the TDE outburst and its host environment.

5. Application Scenarios

In this section, we review the key assumptions we have made in the host subtraction process, and provide recommendations for the practical application of `HostSub_GP`.

To estimate the galaxy background blended with a transient, we effectively model the ratio of the flux in the vicinity of the transient (the user defined host region) and the flux within the mask on the transient (Figure 2). This is achieved by separating the 2D spectrum as the product of a 1D spectrum (F_{sub}) and a 2D spatial profile (ξ_{sub}). If the host galaxy region does not contain sufficient photons, the F_{sub} extracted from this region will be very noisy and might even change sign across the wavelength range. This could introduce singularities in ξ_{sub} , leading to unpredictable biases in the host-subtraction results. The seeing matching process also requires a clean

galaxy spatial profile in the 2D spectrum. Consequently, `HostSub_GP` only applies when the host galaxy is spatially resolved, i.e., the angular size to be at least a few times greater than the seeing.

We further assume that the 2D profile ξ_{sub} varies slowly on both the spatial and spectral directions and that the galaxy profile in archival broad-band images should act as a good prior of ξ_{sub} . However, near narrow galaxy emission lines from H II regions, ξ_{sub} could fluctuate abruptly along the spectral direction, as the spatial distribution of H II regions (producing emission lines) is often different from that of the older stellar population (dominating the continuum). For most cases in the local universe, the broad-band fluxes, and thus our priors on the 2D galaxy profile, are dominated by the older stellar populations. As a result, ξ_{sub} typically deviates significantly from the prior derived from broad-band images, so there is no guarantee that the adopted prior is better than a non-informative one. Nevertheless, by assigning finer batches to the emission lines (Section 3.1.1) and the customized

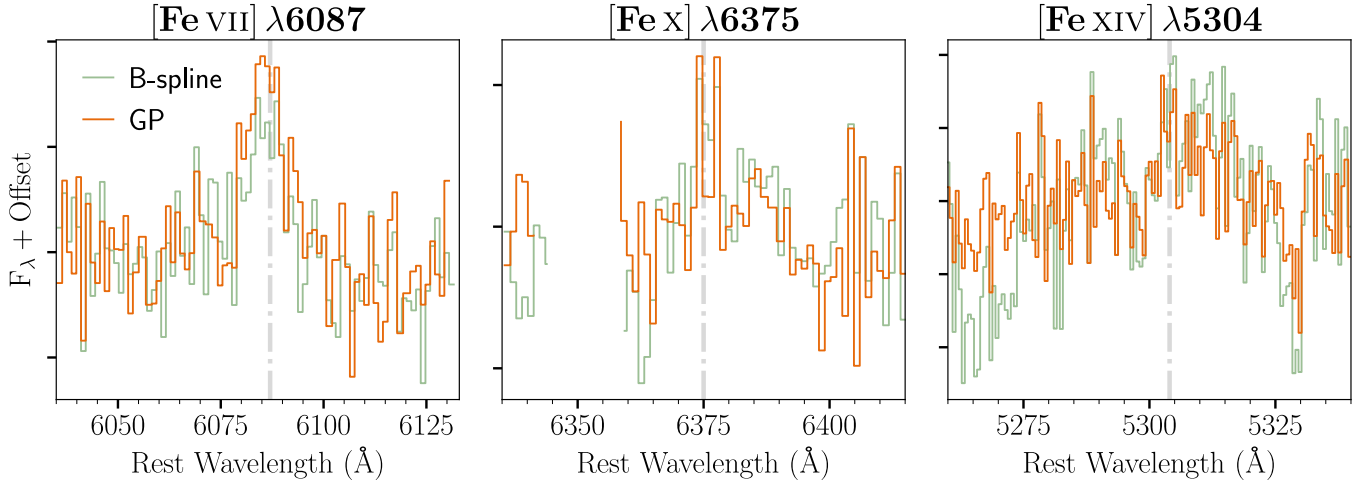


Figure 9. The GP method recovers smooth TDE continua near Fe coronal lines, as later identified in Short et al. (2023), from the 2020 LRIS spectrum of AT 2019qiz. This enables robust flux measurements of the lines from the source. Spectra reduced with both the classic B-spline and the GP method are shown with an added offset for alignment, but fluxes are not rescaled. Central wavelengths of these coronal lines are labeled as the dash-dotted gray lines. The [Fe VII] $\lambda 6087$ emission, isolated from strong galaxy features, appears consistent in both reductions. However, [Fe X] $\lambda 6375$ and [Fe XIV] $\lambda 5304$ lie near prominent galaxy absorption features, and only the GP method recovers the smooth continua. The region around [Fe XI] $\lambda 7892$ is not displayed due to severe CR contamination.

covariance kernel (Section 3.3), the ξ_{sub} in the narrow line regions will not affect the modeling of ξ_{sub} elsewhere.

We assume that the galaxy’s spectral line features remain at fixed wavelengths along the slit, effectively neglecting any radial velocity dispersion within the host galaxy when modeling the 2D spectrum. This assumption is valid when the velocity gradient along the slit is small compared to the spectral resolution. However, for rapidly rotating disk galaxies, it may be necessary to adjust the spectral resolution in the GP model to account for the line tilting in the 2D spectrum.

Another case where archival images fail to approximate the galaxy light profile in the 2D spectrum is when the transient overlaps a variable source, such as an active galactic nucleus. Additionally, when the wavelength range in the spectrum exceeds the reference image coverage (see Figure 3), the prior function must be extrapolated beyond the photometric coverage, in which case it may be less reliable.

Once the primary assumptions are satisfied—namely, when the host background of the transient is spatially resolved, non-variable, and the archival images provide a complete wavelength coverage—`HostSub_GP` can be universally applied to all long-slit observations of transients, irrespective to the transient types, phases, locations, or the host galaxy properties, in which cases it should provide a more robust host contamination removal than classic methods.

6. Summary and Conclusions

We have presented a novel technique for precise host galaxy subtraction in the long-slit spectroscopy of extragalactic transients. We extract knowledge of the galaxy light

distribution from multi-band reference images in archival imaging surveys, and inject the knowledge into the 2D spectrum modeling process with GPs. When the host galaxy is spatially resolved in the 2D spectrum, this GP method robustly evaluates the galaxy continuum blended with the transient spectrum. For transients emerging at complex host environments, it consistently outperforms the classic host subtraction methods, which assumes the galaxy background is an interpolation (with a simple analytical form, e.g., a linear or B-spline interpolation) of the local galaxy flux centered at the transient’s trace, on both the synthetic data and real long-slit observation.

Using the software implementation of this GP method, `HostSub_GP` (available at https://github.com/slowdivePTG/HostSub_GP), we have substantially improved the data quality of the archival Keck/LRIS spectra for (i) a nebular-phase SN, SN 2019eix, and (ii) a nuclear transient, AT 2019qiz. The re-reduced spectrum of SN 2019eix provides smoking-gun evidence on its nature as the thermonuclear explosion of a sub-Chandrasekhar-mass WD. For AT 2019qiz, we set tighter constraints on the delayed emergence of narrow nebular emission (He II and H β) and coronal emission lines (highly ionized Fe).

The utility of `HostSub_GP` is not limited to these two specific cases—instead it can be applied to all long-slit observations of transients with a well resolved, stationary host background. Other examples in the literature include the extraction of the late-time (+1200 days) spectrum of SN 2021aefx obtained with VLT/FORS2, in which a clean host subtraction reveals that the SN emission is dominated by a light echo (Kwok et al. 2025). `HostSub_GP` has also been applied to the real-time monitoring of AT 2025ulz (Kasliwal

et al. 2025), the potential electromagnetic counterpart of the sub-threshold gravitational wave trigger S250818k (LIGO Scientific Collaboration et al. 2025). By removing the strong galaxy background, `HostSub_GP` reveals a hydrogen P-Cygni profile in AT2025ulz, which links the transient to an SN origin.

Finally, designed to be run in a semi-automated fashion, `HostSub_GP` can be naturally implemented in the data reduction and classification pipeline for new transients. Developed upon `JAX` and `tinygp`, `HostSub_GP` enables efficient data-reduction on a personal laptop within a couple of minutes. Future implementation of GPU acceleration under the `JAX` framework will allow near real-time reduction of large spectroscopic datasets.

Acknowledgments

We thank the anonymous referee for the comments that significantly improved this paper. We thank Erik Tollerud, J. Xavier Prochaska, Brad Holden, Kyle Westfall, Ryan Foley, Greg Aldering, and Peter Nugent for fruitful discussions. C.L. and A.A.M. are supported by DoE award #DE-SC0025599, while A.A.M. is also supported by Cottrell Scholar Award #CS-CSA-2025-059 from Research Corporation for Science Advancement.

This work used computing resources provided by Northwestern University and the Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA). This research was supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration.

Software: `Astropy` (Astropy Collaboration et al. 2013, 2018), `JAX` (Bradbury et al. 2018), `JAXopt` (Blondel et al. 2021), `tinygp` (Foreman-Mackey et al. 2024), `PyPeIt` (Prochaska et al. 2020).

ORCID iDs

Chang Liu (刘畅)  <https://orcid.org/0000-0002-7866-4531>
Adam A. Miller  <https://orcid.org/0000-0001-9515-478X>

References

Aigrain, S., & Foreman-Mackey, D. 2023, *ARA&A*, 61, 329
Alard, C., & Lupton, R. H. 1998, *ApJ*, 503, 325
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
Bacon, R., Accardo, M., Adjali, L., et al. 2010, *SPiE*, 7735, 773508
Ben-Ami, S., Gal-Yam, A., Filippenko, A. V., et al. 2012, *ApJL*, 760, L33
Blondel, M., Berthet, Q., Cuturi, M., et al. 2021, arXiv:2105.15183
Blondin, S., Bravo, E., Timmes, F. X., Dessart, L., & Hillier, D. J. 2022, *A&A*, 660, A96
Blondin, S., Walsh, J. R., Leibundgut, B., & Sinton, G. 2005, *A&A*, 431, 757
Bongard, S., Soulez, F., Thiébaud, É., & Pecontal, É. 2011, *MNRAS*, 418, 258
Boyd, R. W. 1978, *JOSA*, 68, 877
Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: composable transformations of Python+NumPy programs, 0.3.13, <http://github.com/jax-ml/jax>
Burgaz, U., Maguire, K., Dimitriadis, G., et al. 2025, *A&A*, 694, A13
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Courbin, F., Magain, P., Kirkove, M., & Sohy, S. 2000, *ApJ*, 529, 1136
Das, K. K., Kasliwal, M. M., Fremling, C., et al. 2023, *ApJ*, 959, 12
De, K., Kasliwal, M. M., Polin, A., et al. 2019, *ApJ*, 873, L18
DeForest, C. E. 2004, *SoPh*, 219, 3
Dessart, L., & Hillier, D. J. 2020, *A&A*, 642, A33
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, 157, 168
Dimitriadis, G. 2019, Transient Name Server Classification Report, 2019-728, 1
Dimitriadis, G., Burgaz, U., Deckers, M., et al. 2025, *A&A*, 694, A10
Eidelman, Y., & Gohberg, I. 1999, *Integral Equations and Operator Theory*, 34, 293
Ellis, R. S., Sullivan, M., Nugent, P. E., et al. 2008, *ApJ*, 674, 51
Fang, Q., Moriya, T. J., Maeda, K., et al. 2025a, *ApJ*, 986, 39
Fang, Q., Nagakura, H., & Moriya, T. J. 2025b, *ApJ*, 994, 49
Fang, Q., Maeda, M., Kuncarayakti, H., et al. 2022, *ApJ*, 928, 151
Filippenko, A. V., Richmond, M. W., Branch, D., et al. 1992, *AJ*, 104, 1543
Flörs, A., Spyromilio, J., Maguire, K., et al. 2018, *A&A*, 620, A200
Flörs, A., Spyromilio, J., Taubenberger, S., et al. 2020, *MNRAS*, 491, 2902
Foley, R. J., Filippenko, A. V., Aguilera, C., et al. 2008, *ApJ*, 684, 68
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *AJ*, 154, 220
Foreman-Mackey, D., Yu, W., Yadav, S., et al. 2024, dfm/tinygp: The tiniest of Gaussian Process libraries, v0.3.0, Zenodo, doi:10.5281/zenodo.6389737
Fried, D. L. 1965, *JOSA*, 55, 1427
Gal-Yam, A. 2017, in *Handbook of Supernovae*, ed. A. W. Alsabti & P. Murdin (Springer International Publishing AG), 195
Galbany, L., Ashall, C., Höflich, P., et al. 2019, *A&A*, 630, A76
Goldwasser, S., Yaron, O., Sass, A., et al. 2022, *TNSAN*, 191, 1
Graham, M. L., Kumar, S., Hosseinzadeh, G., et al. 2017, *MNRAS*, 472, 3437
Howell, D. A., Sullivan, M., Perrett, K., et al. 2005, *ApJ*, 634, 1190
Hung, T., Foley, R. J., Veilleux, S., et al. 2021, *ApJ*, 917, 9
Ivezic, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, 873, 111
Jacobson-Galán, W. V., Polin, A., Foley, R. J., et al. 2020, *ApJ*, 896, 165
Kasliwal, M. M., Ahumada, T., Stein, R., et al. 2025, *ApJL*, 995, L59
Kwok, L. A., Schulze, S., Liu, C., et al. 2025, *TNSAN*, 129, 1
Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, *AJ*, 139, 1782
Lezmy, J., Copin, Y., Rigault, M., Smith, M., & Neill, J. D. 2022, *A&A*, 668, A43
LIGO Scientific CollaborationVIRGO CollaborationKagra Collaboration 2025, GCN, 41437, 1
Liu, C., Miller, A. A., Polin, A., et al. 2023a, *ApJ*, 946, 83
Liu, C., Miller, A. A., Boos, S. J., et al. 2023b, *ApJ*, 958, 178
Lucy, L. B., & Walsh, J. R. 2003, *AJ*, 125, 2266
Maguire, K., Jerkstrand, A., Smartt, S. J., et al. 2012, *MNRAS*, 420, 3451
Maguire, K., Sim, S. A., Shingles, L., et al. 2018, *MNRAS*, 477, 3567
Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, *PASP*, 131, 018003
Milligan, A., Hook, I., Frohmaier, C., et al. 2025, *MNRAS*, 543, 247
Muthukrishna, D., Parkinson, D., & Tucker, B. E. 2019, *ApJ*, 885, 85
Nicholl, M., Wevers, T., Oates, S. R., et al. 2020, *MNRAS*, 499, 482
Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, *PASP*, 107, 375
Padilla Gonzalez, E., Howell, D. A., Burke, J., et al. 2023, *ApJ*, 953, 25
Prochaska, J. X., Hennawi, J. F., Westfall, K. B., et al. 2020, *JOSS*, 5, 2308

Robitaille, T., Ginsburg, A., Mumford, S., et al. 2024, *astropy/reproject*:
v0.13.1, Zenodo, doi:[10.5281/zenodo.10931886](https://doi.org/10.5281/zenodo.10931886)
Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, *ApJ*, **788**, 48
Short, P., Lawrence, A., Nicholl, M., et al. 2023, *MNRAS*, **525**, 1568
Stahl, B. E., Zheng, W., de Jaeger, T., et al. 2020, *MNRAS*, **492**, 4325

Sun, F., & Gal-Yam, A. 2017, arXiv:[1707.02543](https://arxiv.org/abs/1707.02543)
Tody, D. 1986, *SPIE*, **627**, 733
Tucker, M. A., Shappee, B. J., Valley, P. J., et al. 2020, *MNRAS*, **493**, 1044
Woolf, N. J. 1982, *ARA&A*, **20**, 367
York, D. G., Adelman, J., Anderson, J. E. J., et al. 2000, *AJ*, **120**, 1579