raccoon: A Python package for removing wiggle artifacts in the JWST NIRSpec integral field spectroscopy

Anowar J. Shajib [ORCID] [1,2,3]

1 Department of Astronomy & Astrophysics, University of Chicago, Chicago, IL 60637, USA [ROR] **2** Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA **3** Center for Astronomy, Space Science and Astrophysics, Independent University, Bangladesh, Dhaka 1229, Bangladesh

Summary

raccoon is a Python package for removing resampling noise – commonly referred to as "wiggles" – from spaxel-level spectra in datacubes obtained from the JWST Near Infrared Spectrograph's (NIRSpec) integral field spectroscopy (IFS) mode. These wiggles arise as artifacts during resampling of the 2D raw data into 3D datacubes, due to the point spread function (PSF) being undersampled. The standard JWST data reduction pipeline does not correct for this noise. The wiggle artifacts can significantly degrade the scientific usability of the data, particularly at the spaxel level, undermining the exquisite spatial resolution of NIRSpec. raccoon provides an empirical correction by modeling and removing these artifacts, thereby restoring the fidelity of the extracted spectra. raccoon forward-models the wiggles as a chirp function impacting one or more template spectra that are directly fit to the original data across the entire wavelength range. The best-fit wiggle model is then used to clean the data while propagating the associated uncertainties.

Statement of need

The JWST NIRSpec's IFS mode (Böker et al., 2023) enables spatially resolved infrared spectroscopy of astronomical sources with an unprecedented combination of signal-to-noise ratio, redshift coverage, and spatial resolution. However, it suffers from heavy undersampling of the point-spread function, leading to resampling noise that manifests as low-frequency sinusoidal artifacts known as "wiggles." These artifacts can significantly distort the overall spectral shape, bias line measurements, and compromise kinematic analyses at the single-spaxel level, thereby limiting the scientific potential of the NIRSpec IFU data. raccoon provides a user-friendly, robust, and computationally efficient solution to identify and remove these artifacts, enabling precise studies of galaxy kinematics and early-Universe phenomena that would otherwise be hindered by the presence of wiggles. raccoon has already been used in scientific publications to facilitate robust measurements of stellar kinematics from JWST NIRSpec spectra (Shajib et al., 2025; TDCOSMO Collaboration, 2025).

Functionality

Since the wiggles in the single-spaxel spectra in the reduced JWST NIRSpec datacube are artifacts or resampling noise arising due to undersampling the PSF, dithering can, in principle, mitigate the issue of PSF undersampling and thus the wiggles. However, the commonly adopted 4-point dither pattern is still insufficient to recover the optimal sampling and completely

40 eliminate the wiggles in the rectified datacube (Law et al., 2023). The wiggles are washed out
41 in the spectra summed from multiple spaxels within a sufficiently large aperture (illustrated in
42 Figure 1). This provides a basis for making an empirical correction for the wiggles by comparing
43 the single-spaxel spectrum to the one summed within an aperture around it. This principle was
44 also used by previous corrective algorithms, such as in Perna et al. (2023) and in the Python
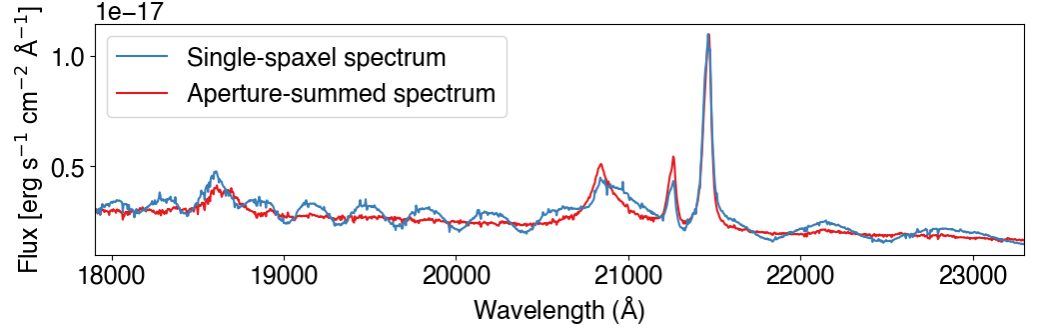45 routine WICKED (Dumont et al., 2025).

Figure 1: Illustration of the wiggles in the single-spaxel spectrum (blue), which is a manifestation of the resampling noise in the standard-pipeline-reduced datacube due to PSF undersampling. The red spectrum shows the aperture-summed spectrum within a 4-spaxel radius, normalized to match its maximum to that of the single-spaxel one. The illustrated data are of an active galactic nucleus (AGN) from Perna et al. (2023).

46 In raccoon, the wiggle is modeled as a sinusoidal chirp function

$$W(\lambda) = 1 + A(\lambda) \left[\sin \phi(\lambda) + a_1 \sin^2 \phi(\lambda) + a_2 \sin(3\phi(\lambda)) \right], \tag{1}$$

47 where $A(\lambda)$ is the wavelength-dependent amplitude and $\phi(\lambda) = \lambda\, k(\lambda) + \phi_0$ is the phase that
48 depends on a wavelength-dependent wavenumber $k(\lambda)$. The wiggle's peaks and troughs can
49 be asymmetrically and symmetrically sharpened (or de-sharpened) by varying the a_1 and a_2
50 parameters, respectively.

51 Given this parametrization for the wiggle, a single-spaxel spectrum $D(\lambda)$ is modeled with

$$M(\lambda) = W(\lambda)\, T(\lambda), \tag{2}$$

52 where $T(\lambda)$ is a template for the correct spectrum devoid of the wiggles. In raccoon, this
53 template is constructed based on the circular-aperture-summed spectrum $C(\lambda)$. The template
54 can also optionally include a spectrum $S(\lambda)$ summed from a shell or annulus centered on the
55 corresponding spaxel, following Dumont et al. (2025). Including the shell-summed spectrum
56 in the template can account for changes in the line shape between the single-spaxel spectrum
57 and the aperture-summed one, for example, in the case of lines broadened by stellar kinematics
58 that can vary across the galaxy (Dumont et al., 2025). The aperture radius and the inner and
59 outer radii of the shell are to be adjusted by the user, as the appropriate values can depend on
60 the source morphology and the astronomical scene. The template is constructed as

$$T(\lambda) = c_1\, C(\lambda) + c_2\, S(\lambda) + c_3\, \lambda^b + \sum_{n=0}^{N} p_n\, \lambda^n, \tag{3}$$

61 where $c_1, c_2, c_3, p_0, \dots, p_N$ are linear coefficients. Here, the power-law plus polynomial term
62 (i.e., $c_3\, \lambda^b + \sum_{n=0}^{N} p_n\, \lambda^n$) on the right-hand side models the difference in the continuum
63 between the single-spaxel spectrum and $c_1\, C(\lambda) + c_2\, S(\lambda)$.

64 The functions $A(\lambda)$ and $k(\lambda)$ are modeled with B-splines, with the number of knots adjustable
65 by the user. raccoon provides functionality for the user to determine the appropriate number

66 of knots using model selection criteria based on the Bayesian information criterion (BIC) or the
67 minimum *a posteriori* chi-square metric (χ^2_{MAP}). The best-fit values for the linear coefficients
68 (i.e., $c_1, c_2, c_3, p_0, ..., p_N$) and non-linear parameters (i.e., a_1, a_2, b, and the coefficients of
69 the B-spline basis functions) are determined by minimizing

$$\chi^2 = \sum_i \frac{(D_i - M_i)^2}{\sigma_i^2}, \tag{4}$$

70 where the index i runs across the wavelength pixels and σ_i is the associated noise level. Figure 2
71 illustrates an example of the fitted model to a given spectrum (of an AGN). In this example,
72 raccoon's robust performance is demonstrated in fitting the given spectrum while modeling
73 the wiggle signal (Figure 3). The user can mask out regions of the spectrum that have strong
74 features potentially impacting the quality of the fit, or optionally adopt outlier rejection in the
75 fit using the false discovery rate method (Benjamini & Hochberg, 1995) or sigma-clipping.
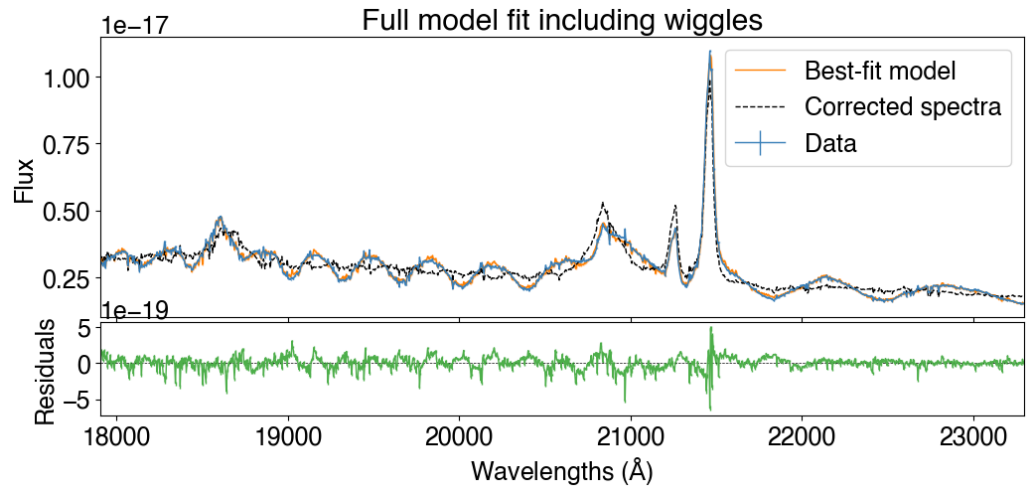

Figure 2: Modeling of the full spectrum (blue in the top panel) based on the template spectra and the wiggles impacting it. The illustrated spectrum is the same one from Figure 1. The best-fit model is shown in orange and the wiggle-corrected spectrum is shown in black. The bottom panel illustrates the residuals (green) between the original data and the best-fit model.
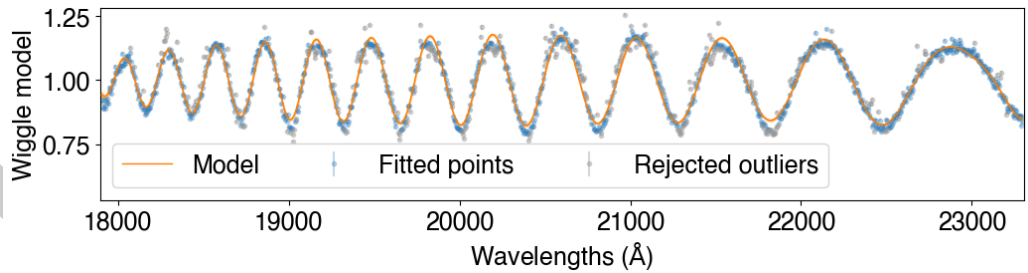

Figure 3: Illustration of the extracted and modeled wiggle signal. The illustrated data points represent $D(\lambda)/M(\lambda)$ based on the best-fit $M(\lambda)$, and the orange line illustrates the best-fit wiggle model $W(\lambda)$. The grey points are rejected outliers using the false discovery rate method, and the blue points mark the wavelengths where the data were fit to the model.

76 raccoon enables users to efficiently process multiple spaxels in a datacube, either across the
77 entire field or within a specified spatial region. A user-defined detection threshold for the
78 wiggle signal can be set, ensuring that corrections are only applied to spectra where significant
79 wiggles are present.

raccoon offers several advantages over previously available scripts and routines. Earlier tools (Dumont et al., 2025; Perna et al., 2023) extract the wiggle signal by subtracting the single-spaxel spectrum from a template spectrum (similar to or the same as that in Equation 3) fitted to the data without accounting for wiggles. This approach can result in the loss of a fraction of the wiggle signal during numerical fitting. In contrast, raccoon fits the template spectrum and the parametric wiggle model $W(\lambda)$ simultaneously to the data, eliminating the need for an intermediate extraction step and preserving the full wiggle signal. While previous scripts perform fitting separately in multiple wavelength slices, raccoon fits across the entire wavelength range at once. This makes raccoon more robust to gaps in the fitted spectrum, whether excluded by manual masking or outlier rejection. Another important distinction is that earlier tools apply an additive correction, whereas raccoon uses a multiplicative correction. Since both approaches derive the correction empirically from the data, this difference should not affect performance in practice. However, as the wiggles manifest as a multiplicative effect on the spectra (Law et al., 2023), raccoon's implementation is physically motivated. Additionally, raccoon is uniquely installable via the `pip` command, allowing more flexible import into Python scripts or notebooks and making it more user-friendly and portable than its peers.

Acknowledgements

References

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Böker, T., Beck, T. L., Birkmann, S. M., Giardino, G., Keyes, C., Kumari, N., Muzerolle, J., Rawle, T., Zeidler, P., Abul-Huda, Y., Alves de Oliveira, C., Arribas, S., Bechtold, K., Bhatawdekar, R., Bonaventura, N., Bunker, A. J., Cameron, A. J., Carniani, S., Charlot, S., … Willott, C. J. (2023). In-orbit Performance of the Near-infrared Spectrograph NIRSpec on the James Webb Space Telescope. *Publications of the Astronomical Society of the Pacific*, *135*, 038001. https://doi.org/10.1088/1538-3873/acb846

Dumont, A., Neumayer, N., Seth, A. C., Böker, T., Eracleous, M., Goold, K., Greene, J. E., Gültekin, K., Ho, L. C., Walsh, J. L., & Lützgendorf, N. (2025). WIggle Corrector Kit for NIRSpEc Data: WICKED. *arXiv e-Prints*, arXiv:2503.09697. https://doi.org/10.48550/arXiv.2503.09697

Law, D. R., E. Morrison, J., Argyriou, I., Patapis, P., Álvarez-Márquez, J., Labiano, A., & Vandenbussche, B. (2023). A 3D Drizzle Algorithm for JWST and Practical Application to the MIRI Medium Resolution Spectrometer. *The Astronomical Journal*, *166*, 45. https://doi.org/10.3847/1538-3881/acdddc

Perna, M., Arribas, S., Marshall, M., D'Eugenio, F., Übler, H., Bunker, A., Charlot, S., Carniani, S., Jakobsen, P., Maiolino, R., Rodríguez Del Pino, B., Willott, C. J., Böker, T., Circosta, C., Cresci, G., Curti, M., Husemann, B., Kumari, N., Lamperti, I., … Scholtz, J. (2023). GA-NIFS: The ultra-dense, interacting environment of a dual AGN at z $\sim$ 3.3 revealed by JWST/NIRSpec IFS. *679*, A89. https://doi.org/10.1051/0004-6361/202346649

Shajib, A. J., Treu, T., Suyu, S. H., Law, D., Yıldırım, A., Cappellari, M., Galan, A., Knabel, S., Wang, H., Birrer, S., Courbin, F., Fassnacht, C. D., Frieman, J. A., Melo, A., Morishita, T., Mozumdar, P., Sluse, D., & Stiavelli, M. (2025). TDCOSMO XXIII. First spatially

resolved kinematics of the lens galaxy obtained using JWST-NIRSpec to improve time-delay cosmography. *arXiv e-Prints*, arXiv:2506.21665. https://arxiv.org/abs/2506.21665

TDCOSMO Collaboration. (2025). TDCOSMO 2025: Cosmological constraints from strong lensing time delays. *arXiv e-Prints*, arXiv:2506.03023. https://doi.org/10.48550/arXiv.2506.03023