Supporting Information for "Detecting Sub-Micron Space Weathering Effects in Lunar Grains with Synchrotron Infrared Nanospectroscopy"

K. L. Utt[1], R. C. Ogliore[1], H. A.Bechtel[2], J. J. Gillis-Davis[1], and

B. L. Jolliff[3,4]

[1]Department of Physics, Washington University in St. Louis, St. Louis, MO, USA

[2]Advanced Light Source Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

[3]Department of Earth and Planetary Sciences, Washington University in St. Louis, St. Louis, MO, USA

[4]McDonnell Center for the Space Sciences, Washington University in St. Louis, St. Louis, MO, USA

Contents of this file

1. Text S1

Additional Supporting Information (Files uploaded separately)

1. Captions for Datasets S1 to S6

———

Corresponding author: Kainen L. Utt (k.l.utt@wustl.edu)

June 24, 2021, 3:00pm

Text S1. We used an iterative, non-linear, and robust peak fitting (deconvolution) procedure to determine peak parameters for SINS phase spectra. The procedure uses a trust-region minimization algorithm to fit a linear combination of Lorentzians with a local linear background (see (1)) to the data. Robustness is ensured by minimizing the residuals' summed square and using bi-square weighting to reduce the impact of outliers. The general equation for this fit is given by

$$f(\tilde{\nu}) = \left(p_1 \left(\tilde{\nu} - p_3\right) + p_2\right) + \sum_{i=1}^{n} \frac{a_i}{\pi} \left(\frac{c_i}{\left(b_i - \tilde{\nu}\right)^2 + c_i^2}\right), \tag{1}$$

where p_1, p_2, and p_3 are the coefficients of the local linear baseline. The remaining coefficients — a_i, b_i, and c_i — relate to the amplitude, position, and width of the relevant peak. The number of peaks (n) required to reliably produce a fit with sufficiently random residuals varied between samples. However, we found that $n = 12$ peaks was generally sufficient for stable convergence to a near-optimal solution in the lunar plagioclase samples. Fewer peaks were required to fit the spectra from the pyroxene sample.

This approach was chosen to accommodate the relatively low signal-to-noise ratio in some spectra; iterative models are more robust than classical methods when the signal is noisy (Gautam et al., 2015). To reduce the number of iterations required for the model fit to converge to a near-optimum solution, we introduced physically-motivated constraints on the coefficient values and initialized the algorithm with reasonable starting guesses for each parameter's values. These constraints are included alongside other model specifications in Data Set S3.

Note that the spectra were not smoothed before performing the fitting procedure — to do so would result in potentially severe underestimation errors when propagating uncertainties and estimating sample statistics (O'Haver, 2020).

Captions for Data Sets (Files uploaded separately)

Data Set S1. Raw SINS spectra for studied samples (sections 1–4 and the terrestrial anorthite standard). The relevant folders additionally contain the background spectral files (.CSV) that were used. These files are organized by section and scan. The original files, in proprietary .SPA format, are excluded from this Data Set due to file-size constraints. Copies of the .SPA files can be made available to interested parties upon request.

Data Set S2. This dataset is comprised of background-referenced and Fourier-transformed SINS spectra (.CSV). Note that both the amplitude and phase signals are included here. These files are organized by section and scan.

Data Set S3. The compressed folder contains the initial conditions for the iterative fitting model as well as the peak parameters calculated thereby. These data are formatted as .CSV files.

Data Set S4. The compressed folder contains 16 animated .GIF files (one for each linescan) displaying the fitted, un-renormalized phase spectra from each point in the linescan. The general darkening trend can be seen in spite of the noise.

Data Set S5. The individual frames of the animations included in Data Set S4. Files are saved in .PNG format with a name in the style of:

`\S[Sample #]_LS[Scan #]_[Depth]nm_[Phase/Amp].png`

Data Set S6. This dataset contains the fitted spectra (.CSV) and 2σ confidence intervals (CI). The files contained in this dataset are formatted such that the first column contains the wavenumbers, the second column contains the fitted spectra evaluated at that wavenumber, and the third and fourth columns contain the CI for the fit.

References

Gautam, R., Vanga, S., Ariese, F., & Umapathy, S. (2015). Review of multidimensional data processing approaches for raman and infrared spectroscopy. *EPJ Techniques and Instrumentation*, *2*(1), 1–38.

O'Haver, T. (2020). *A Pragmatic Introduction to Signal Processing*. University of Maryland at College Park. Retrieved from `https://terpconnect.umd.edu/~toh/spectrum/`