CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Automated baseline correction evaluation score for Raman spectroscopy
Authors	Innocente, Simone;Visentin, A.;Andersson-Engels, S.;Komolibus, K.;Gautam, R.
Publication date	2026-04-23
Original Citation	Innocente, S, Visentin, A, Andersson-Engels, S, Komolibus, K & Gautam, R 2026, 'Automated baseline correction evaluation score for Raman spectroscopy', ACS Omega, vol. 11, no. 17, pp. 25057-25068. https://doi.org/10.1021/acsomega.5c09870
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://www.scopus.com/pages/publications/105037855058 - 10.1021/acsomega.5c09870https://www.scopus.com/pages/publications/105037855058?origin=resultslist
Rights	Publisher Copyright: © 2026 The Authors. Published by American Chemical Society. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-11-14 16:23:26
Item downloaded from	https://hdl.handle.net/10468/18751

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Supporting Information

Automated Baseline Correction Evaluation Score for Raman Spectroscopy

Simone Innocente (1, 3)*, Andrea Visentin (3, 4), Stefan Andersson-Engels (1, 2), Katarzyna Komolibus (1), Rekha Gautam* (1)

1. Biophotonics@Tyndall, Tyndall National Institute, Dyke Parade, Cork, Ireland
2. School of Physics, University College Cork, College Road, Cork, Ireland
3. School of Computer Science and Information Technology, UCC, Cork, Ireland
4. Insight Centre for Data Analytics, UCC, Cork, Ireland

*Correspondence: Simone Innocente: simone.innocente@tyndall.ie, Rekha Gautam: rekha.gautam@tyndall.ie

Additional Experimental Details

To further assess the generality and applicability of the IS-Score, additional experiments were conducted on Raman spectroscopy datasets from various application domains, including food science, pharmaceuticals and mineralogy. For each dataset, three baseline correction algorithms were evaluated. Results are presented as boxplot of the IS-Score across all spectra, along with the average spectrum and corresponding baselines.

The datasets include:

- A Resonant Raman dataset from food science, comprising 30 spectra of extra virgin olive oil at different aging stages [1];
- A paracetamol dataset with 50 spectra from various manufacturers, labelled by brand codes to capture formulation differences [2];
- A single polyhalite spectrum from the RRUFF mineral Raman dataset (no boxplot shown due to limited sample size) [3].

The variability in the distribution of IS-Score values primarily reflects the intrinsic diversity of Raman spectra. Baseline correction evaluation remains a complex task due to the lack of a proper definition of "*good*" baseline among experts and subjective nature of task itself. Despite this challenge, the IS-Score reveals consistent trends that align with the visual quality of baseline correction, making it a useful tool to streamline and accelerate Raman data preprocessing workflows.

Food Science Dataset

The spectral range for this dataset is 501.82 to 547.50 cm^{-1} with 1022 Raman shifts in total.

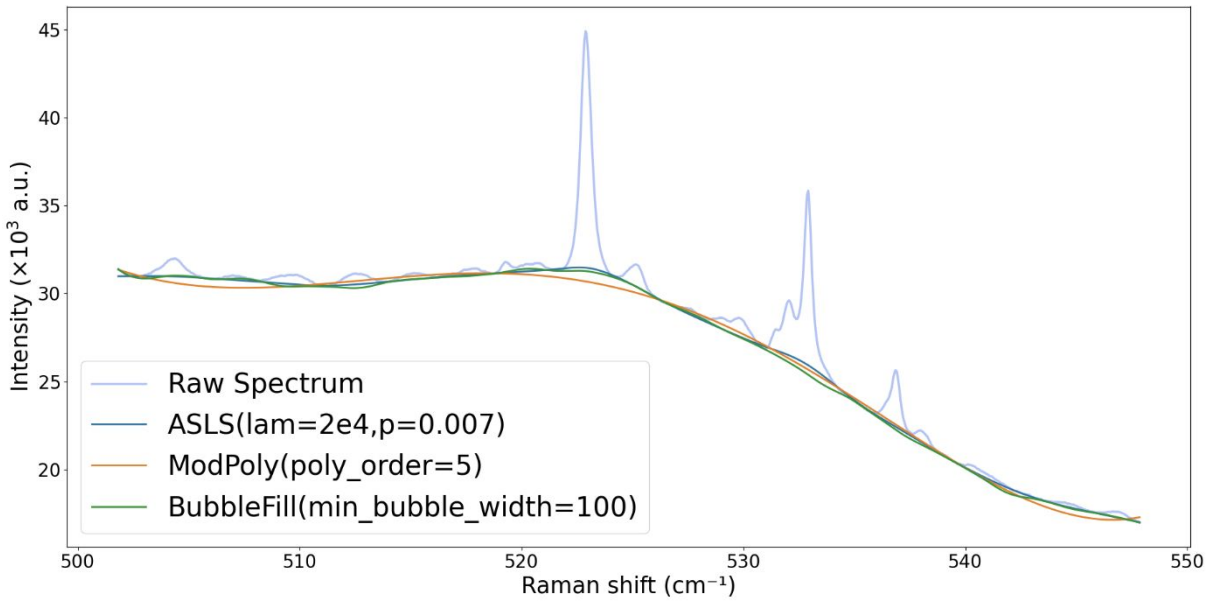


Figure S1: Mean spectrum from the food science dataset with the baselines evaluated.

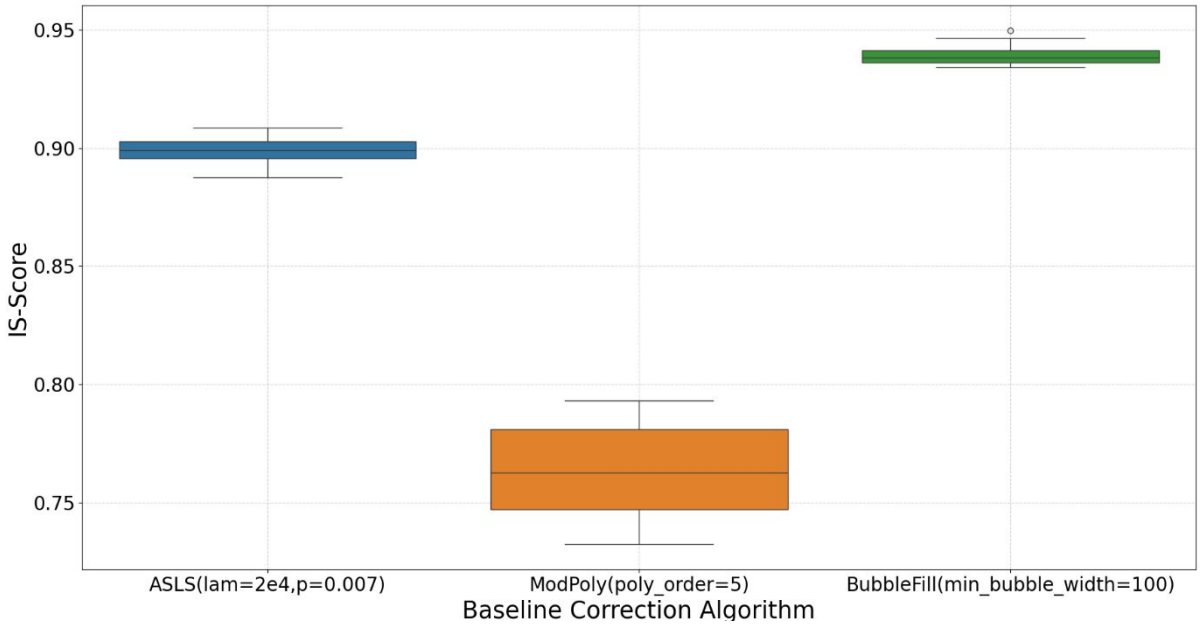


Figure S2: Boxplot of the IS-Score for the food science dataset for the three different baselines tested.

Baseline	Mean IS-Score
ASLS(lambda=2e4, p=0.007)	0.8990
ModPoly(poly_order=5)	0.7635
BubbleFill(min_bubble_width=100)	0.9394

Table S1: Average IS-Score for each baseline across the food science dataset.

Pharmaceutical Dataset

The spectral range for this dataset is 300.77 to 1850.54 cm^{-1} with 788 Raman shifts in total.

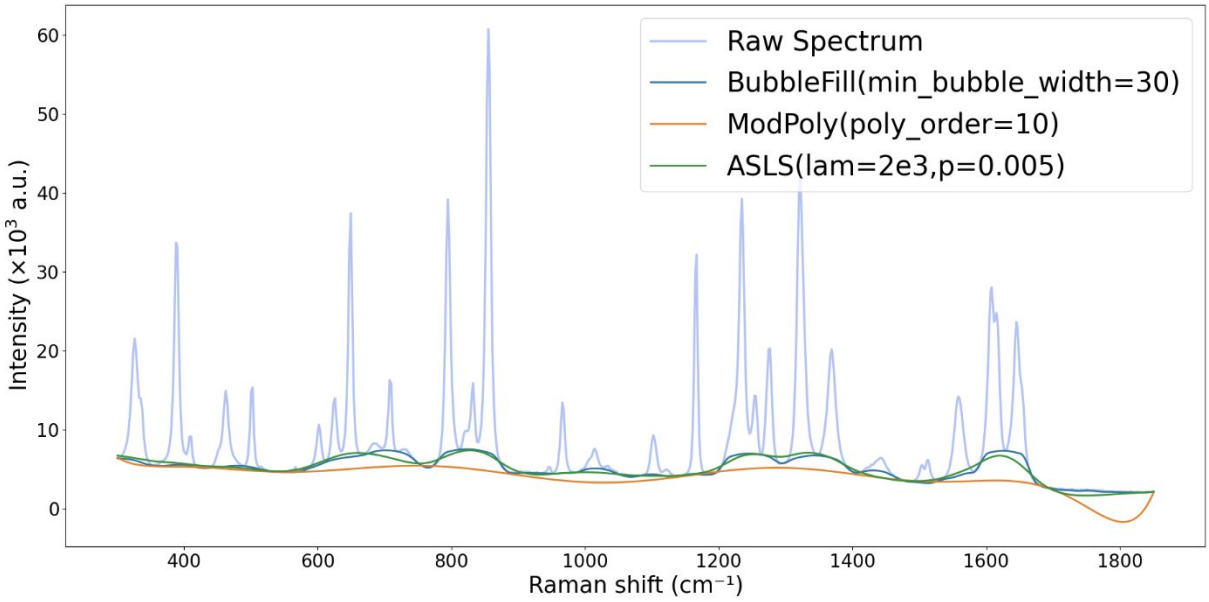


Figure S3: Mean spectrum from the pharmaceutical dataset with the baselines evaluated.

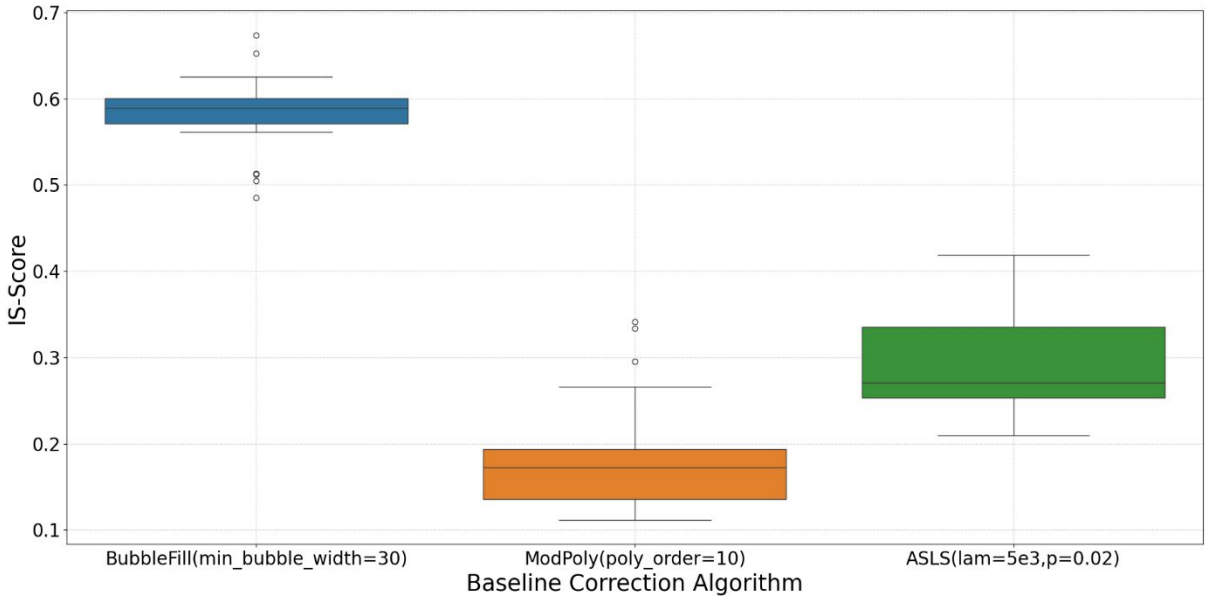


Figure S4: Boxplot of the IS-Score for the pharmaceutical dataset for the three different baselines tested.

Baseline	Mean IS-Score
BubbleFill(min_bubble_width=30)	0.5825
ModPoly(poly_order=10)	0.1762
ASLS(lambda=5e3, p=0.002)	0.2872

Table S2: Average IS-Score for each baseline across the pharmaceutical dataset.

Mineralogy Spectrum

The spectral range for this spectrum is 119.925 to 6578.36 cm^{-1} with 3350 Raman shifts in total.

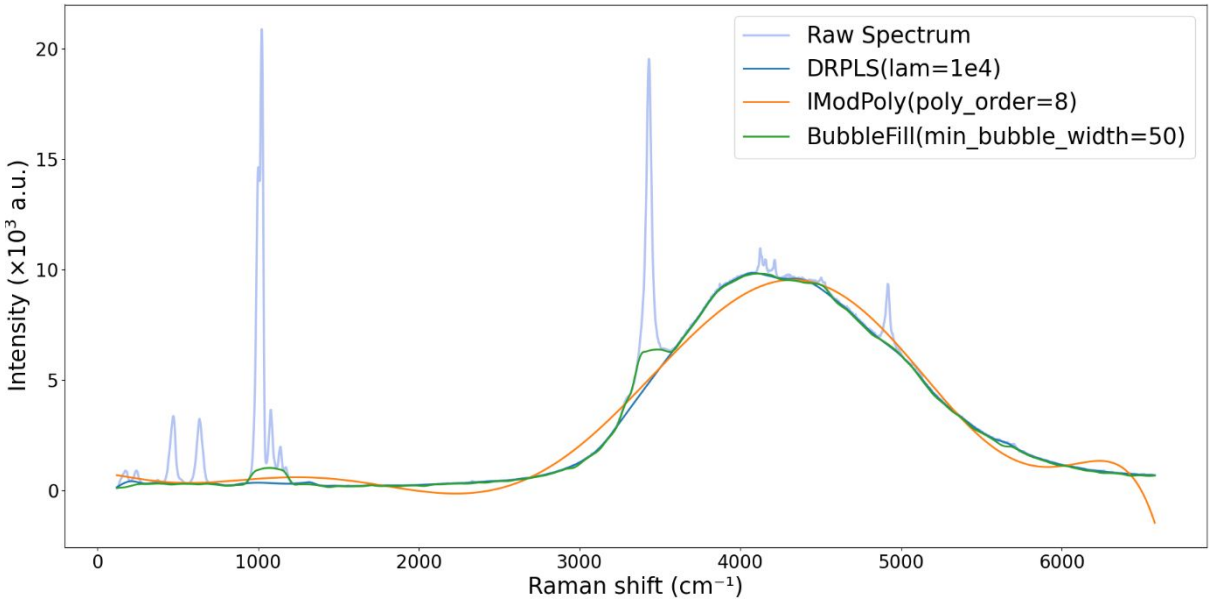


Figure S5: Mineral spectrum with the baseline evaluated.

Baseline	IS-Score
DRPLS(lam=1e4)	0.8971
IModPoly(poly_order=8)	0
BubbleFill(min_bubble_width=50)	0.6151

Table S3: IS-Score for each baseline for the mineral spectrum.

Sensitivity Parameter Analysis

To assess the robustness of the IS-Score with respect to the choice of internal thresholds, a parameter sensitivity analysis was conducted. Specifically, the analysis focused on several key thresholds employed in the algorithm, including (1) the 75% peak prominence, (2) the coefficient of the logarithmic term introduced in the single peak penalty, (3) the positive-negative AUC ratio in the AUC penalization and (4) the mean-ratio threshold. These parameters were selected through empirical testing on multiple representative Raman spectra; however, in this analysis, we systematically varied each parameter across a broad and plausible interval and quantified the resulting impact on the IS-Score. The datasets used in this analysis include two datasets employed in the main manuscript, as well as all additional datasets provided in this Supplementary Information.

1. **Prominence Threshold**

The prominence threshold was varied within the range **0.60 - 0.90** in steps of 0.05**.** Figure S6, S7, S8, S9, and S10 present the corresponding boxplots for each baseline across the five datasets: isolated red blood cells and polymer bag, red blood cells in polymer bag, food science, pharmaceutical and mineralogy. For the mineralogy dataset a scatter plot is presented because only one single spectrum is available. The prominence controls the likelihood that a peak will be penalized: the higher its value, the more likely a peak is to be penalized. Across the evaluated range for each threshold and dataset tested, the IS-Score remained stable, indicating that the metric is robust to change in this internal parameter. Although higher prominence values led to a modest decrease in the performance of the IS-Score, this effect does not impact the relative ranking of baseline methods under investigation. The value of 0.75 was selected as a balanced compromise, high enough to enforce meaningful penalties while preserving the accuracy of the IS-Score.

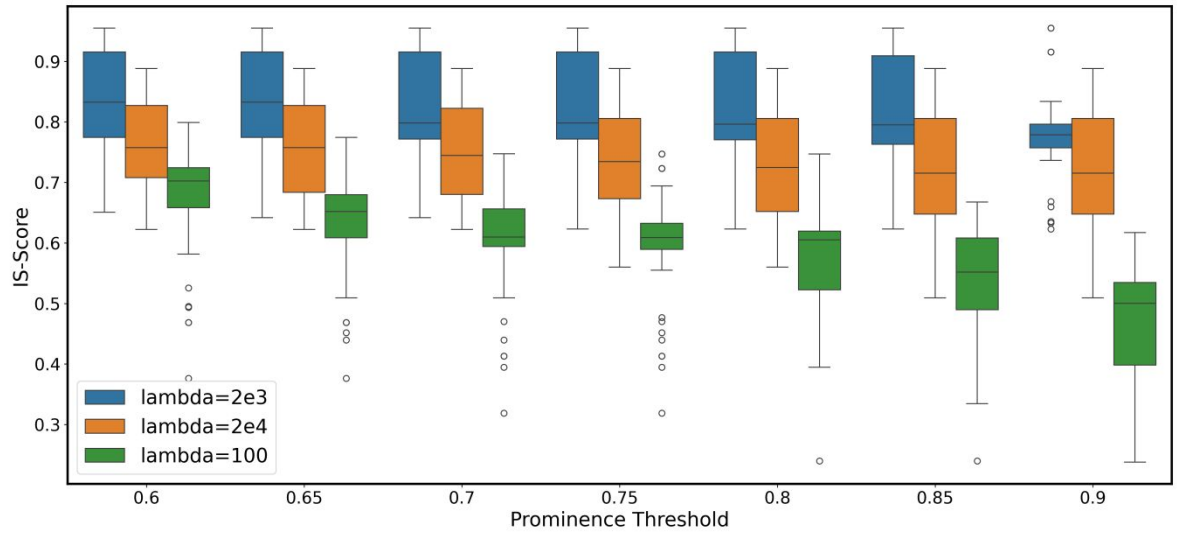


Figure S6: IS-Score obtained for prominence threshold between 0.60 and 0.90 for each lambda value of the Asymmetric Least Square (ASLS) baseline, evaluated on the ***isolated red blood cells and polymer bags*** *dataset.*

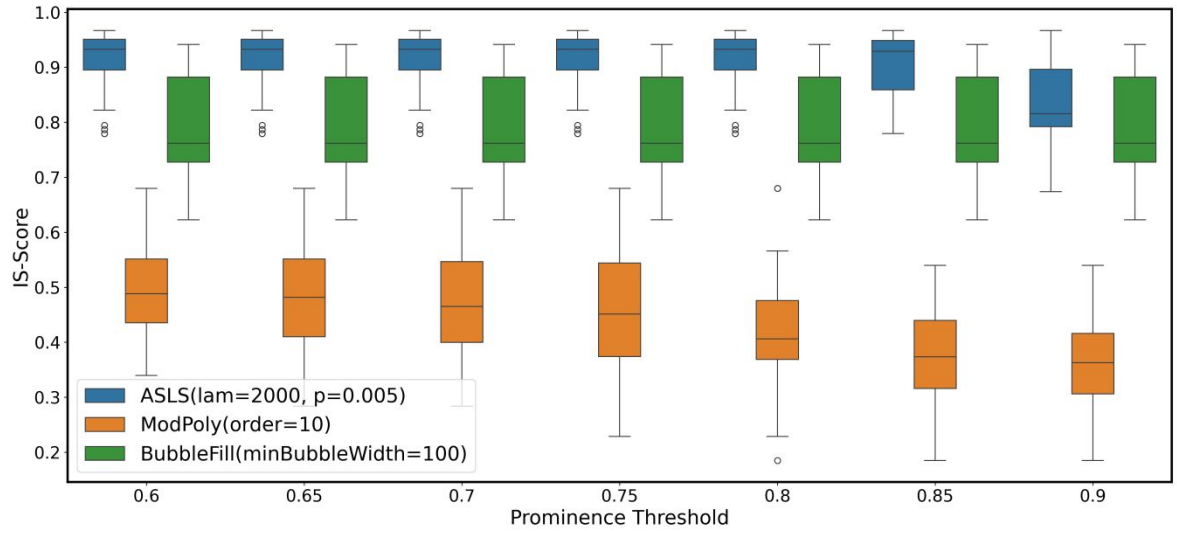


*Figure S7: IS-Score obtained for prominence threshold between 0.60 and 0.90 for each baseline tested for the **red blood cells in polymer bags** dataset.*

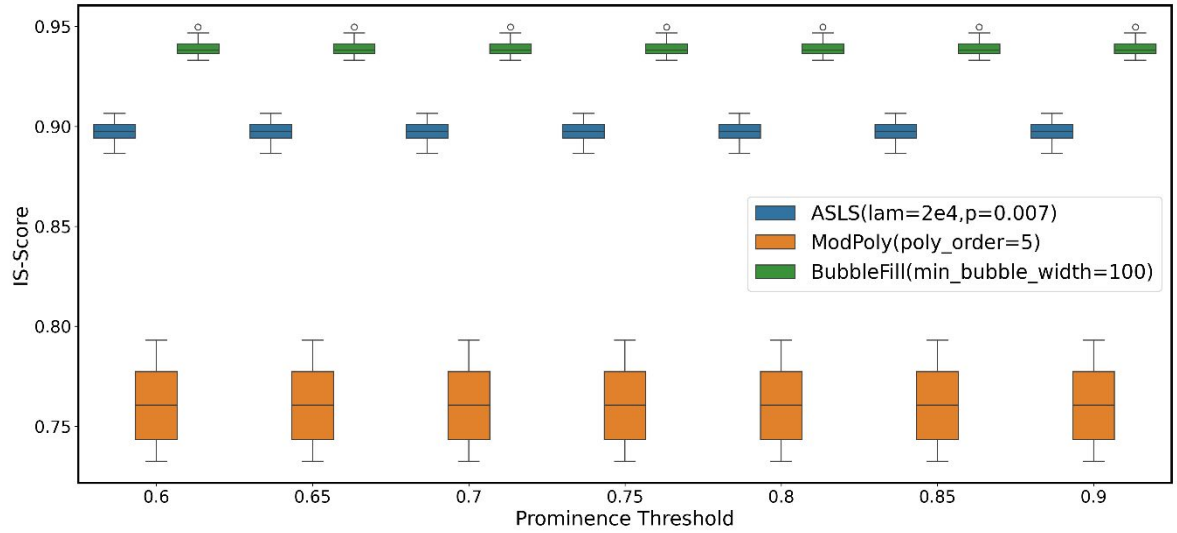


*Figure S8: IS-Score obtained for prominence threshold between 0.60 and 0.90 for each baseline tested for the **food science** dataset.*

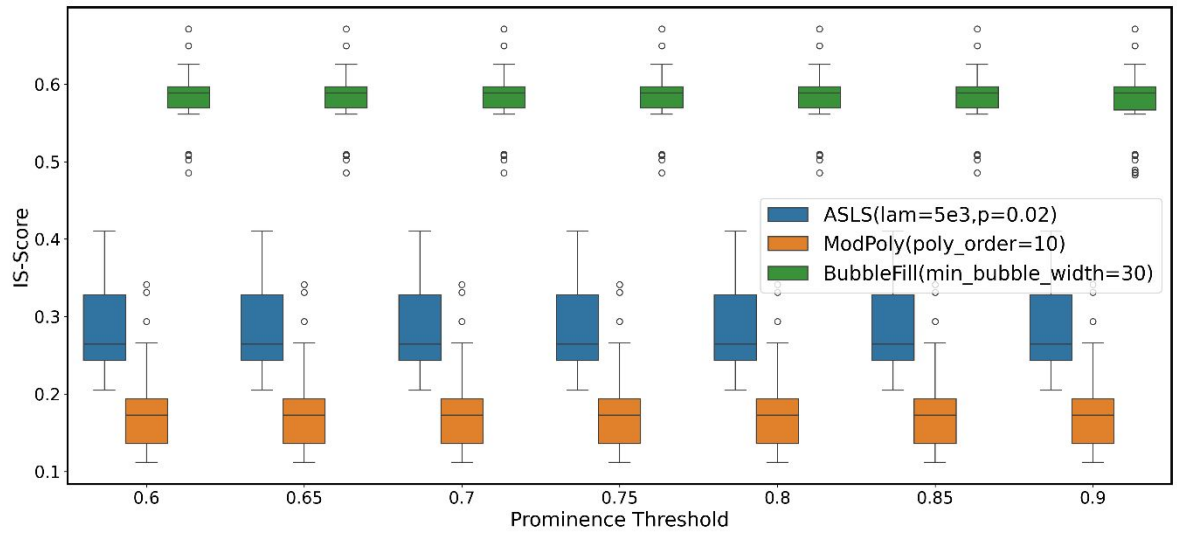


*Figure S9: IS-Score obtained for prominence threshold between 0.60 and 0.90 for each baseline tested for the **pharmaceutical** dataset.*

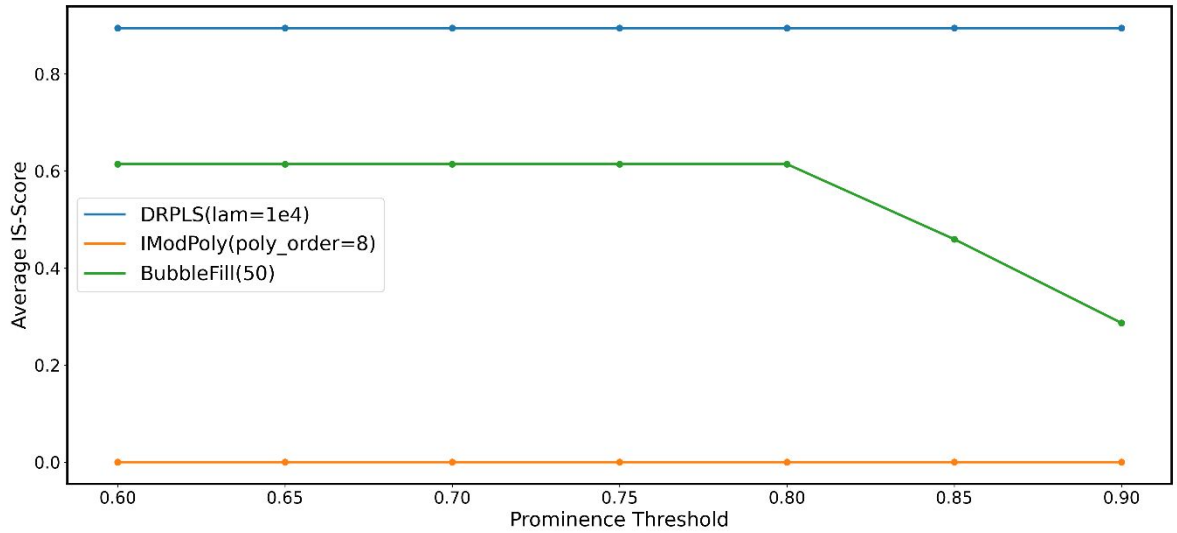


*Figure S10: IS-Score obtained for prominence threshold between 0.60 and 0.90 for each baseline tested for the **mineralogy** spectrum.*

2. Logarithmic Threshold

The log threshold was varied within the range **1.05 - 1.95,** in steps of 0.15**.** Figure S11, S12, S13, S14, and S15 present the corresponding boxplots for each baseline across the five datasets: isolated red blood cells and polymer bag, red blood cells in polymer bag, food science, pharmaceutical and mineralogy. For the mineralogy dataset a scatter plot is presented because only one single spectrum is available. This threshold controls the strength of the penalty applied in the single peak penalization block. Across all datasets and baseline, the IS-Score remains consistent, preserving both the relative raking and the values of the IS-Scores. The value of 1.5 was selected as a balanced choice, providing an effective penalization without excessively reducing the IS-Score, thereby ensuring stable and comparable performance across different conditions.

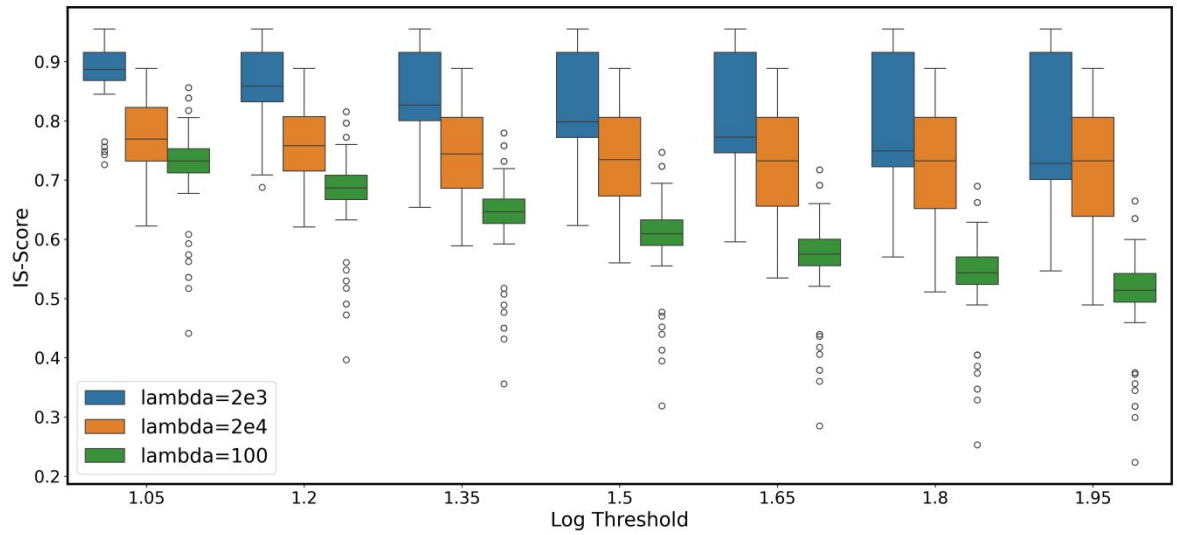


Figure S11: IS-Score obtained for logarithmic threshold between 1.05 and 1.95 for each lambda value of the Asymmetric Least Square (ASLS) baseline, evaluated on *the **isolated red blood cells and polymer bags** dataset.*

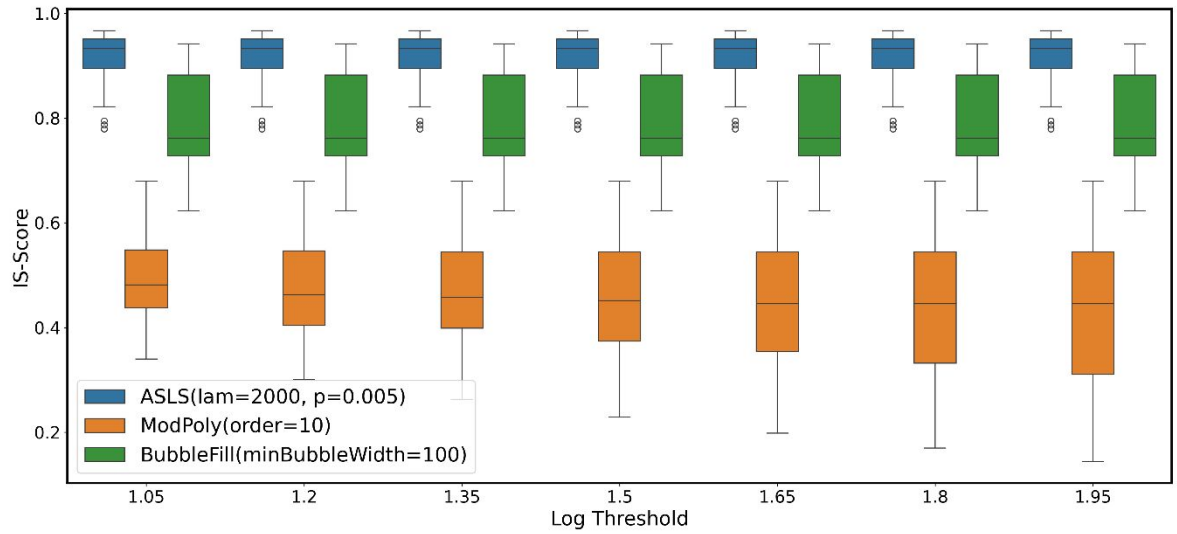


*Figure S12: IS-Score obtained for logarithmic threshold between 1.05 and 1.95 for each baseline tested for the **red blood cells in polymer bags** dataset.*

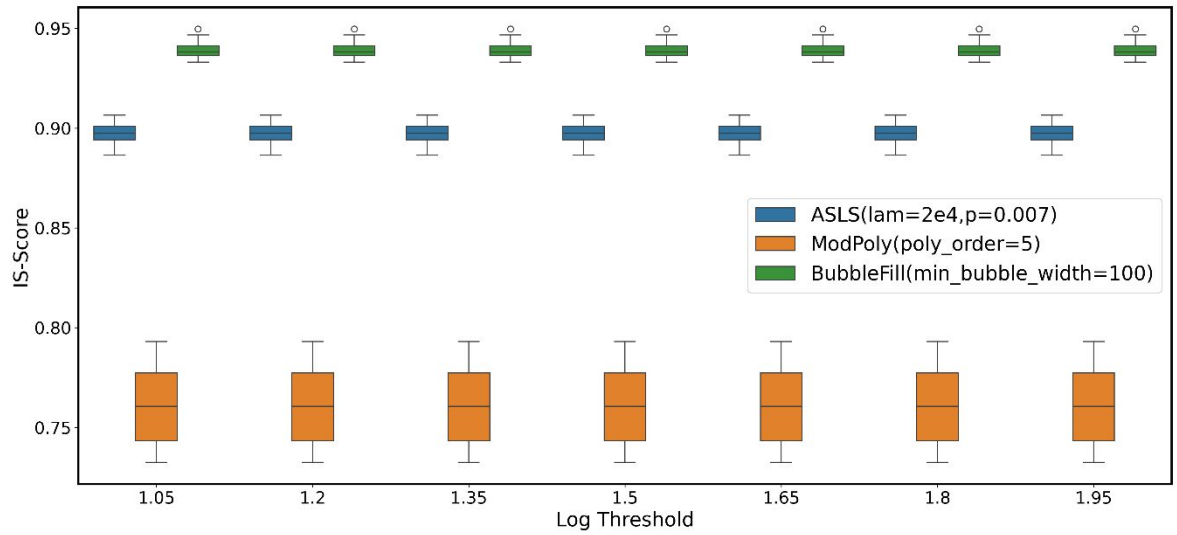


*Figure S13: IS-Score obtained for logarithmic threshold between 1.05 and 1.95 for each baseline tested for the **food science** dataset.*

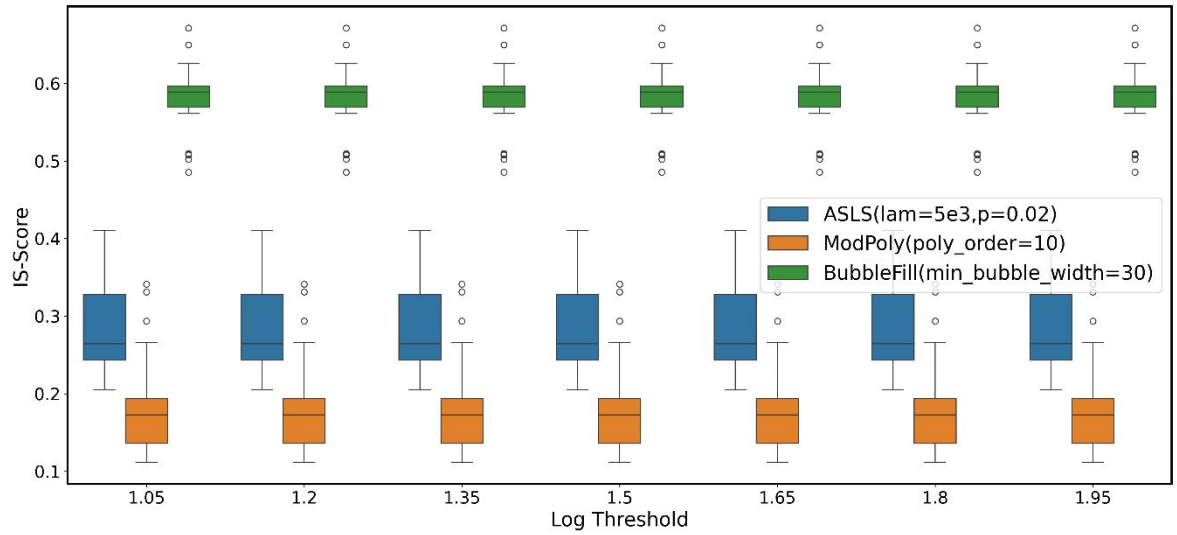


*Figure S14: IS-Score obtained for logarithmic threshold between 1.05 and 1.95 for each baseline tested for the **pharmaceutical** dataset.*

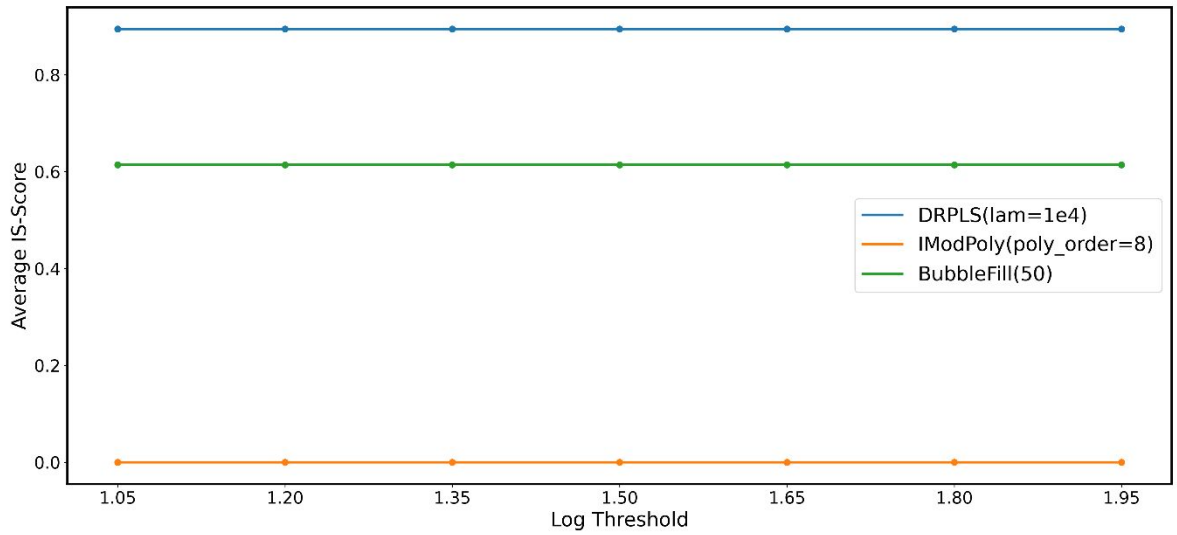


*Figure S15: IS-Score obtained for logarithmic threshold between 1.05 and 1.95 for each baseline tested for the **mineralogy** spectrum.*

Prominence-Log Threshold Heatmap

Prominence and logarithmic (log) thresholds belong to the same decision block within the IS-Score. To visualize potential interaction between these two parameters, we generated heatmaps showing the IS-Score value obtained for all combinations of the prominence and log thresholds, computed separately for each baseline correction method tested here. This multi-parameter analysis enables a thorough assessment of the stability of the IS-Score under different baseline correction configurations. Figure S16, S17, S18, S19, and S20 present the corresponding heatmaps for each baseline across the five datasets: isolated red blood cells and polymer bag, red blood cells in polymer bag, food science, pharmaceutical and mineralogy. For the food science dataset (Figure S18), the values remain unchanged, as neither the prominence nor the log coefficient affects the single peak penalty, resulting in stable scores across the tested ranges.

In general, when peaks are affected, higher values for both thresholds lead to reduced IS-Scores. This trend is expected, as increasing the thresholds raises the likelihood of penalizing each peak, while a higher log coefficient amplifies the associated penalty. Nevertheless, despite these changes in the average IS-Score, the results remain consistent, preserving the relative ranking of the scores across the different baseline tested. The prominence threshold of 0.75 and log coefficient of 1.5 were selected as a balanced compromise. They avoid excessive penalization that could obscure meaningful differences between baselines while maintaining sufficient sensitivity to distinguish medium- and high-quality corrections, ensuring robust IS-Score performance across diverse datasets.

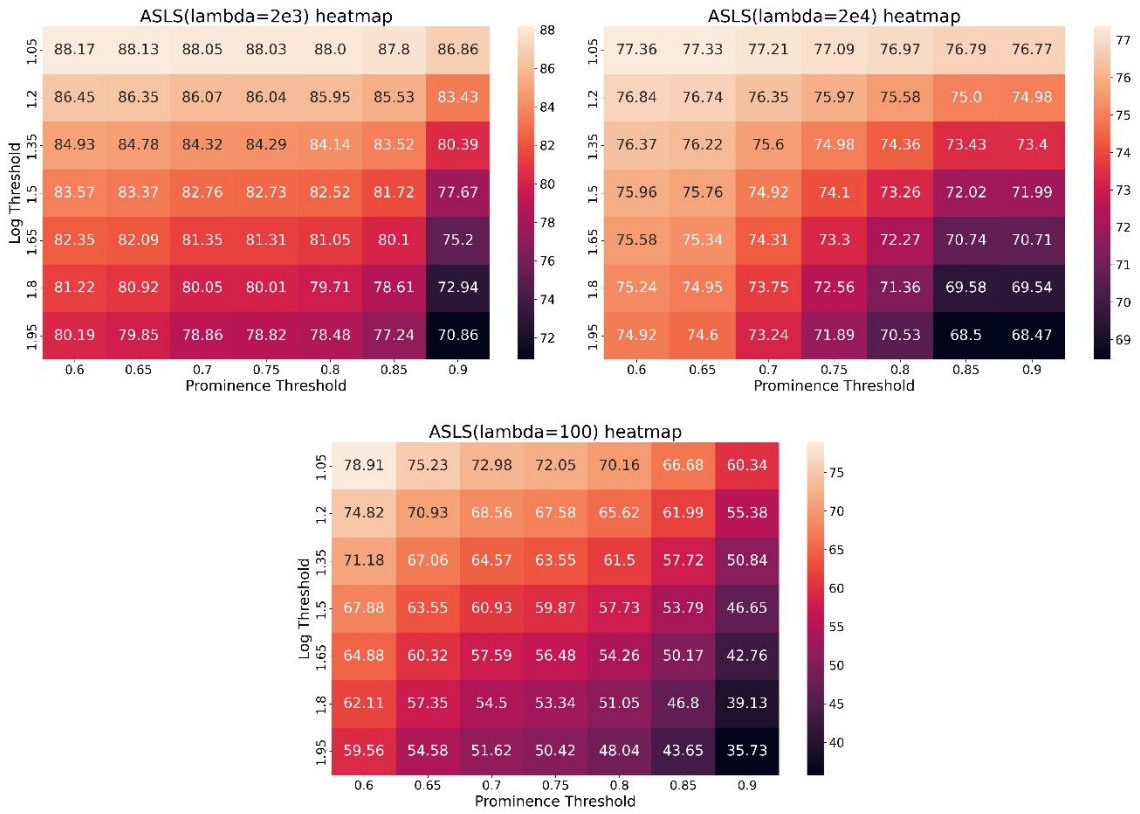


Figure S16: Average IS-Score (%) obtained by varying prominence and logarithmic for each lambda value of the Asymmetric Least Square (ASLS) baseline, evaluated on *the* ***isolated red blood cells and polymer bags*** *dataset.*

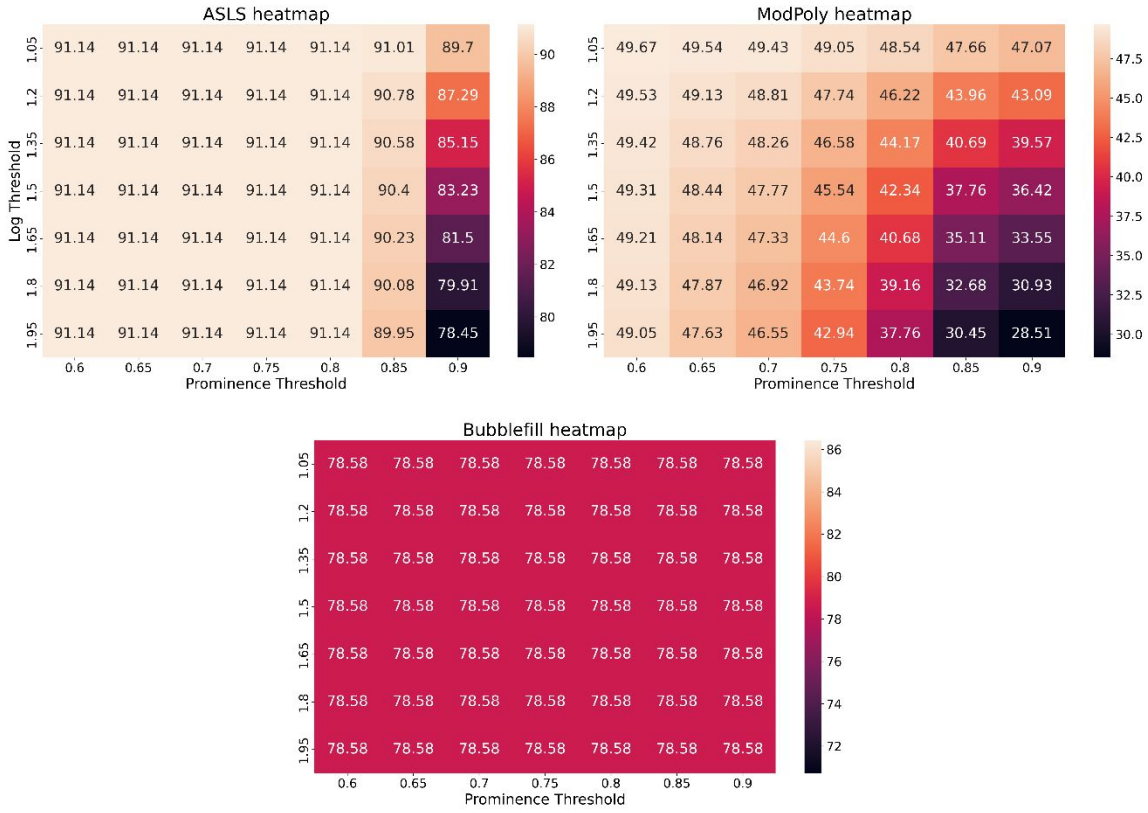


Figure S17: Average IS-Score (%) obtained by varying prominence and logarithmic for each baseline tested for the ***red blood cells in polymer bags*** *dataset.*

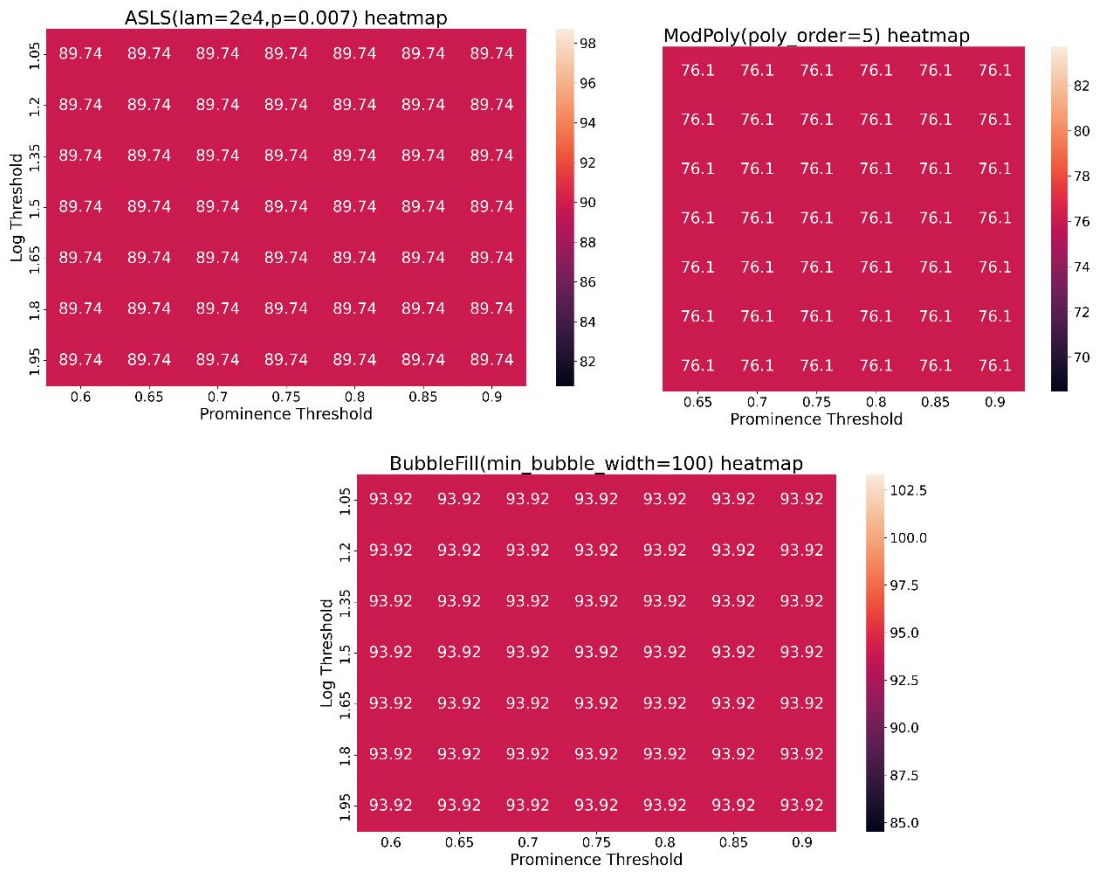


*Figure S18: Average IS-Score (%) obtained by varying prominence and logarithmic for each baseline tested for the **food science** dataset.*

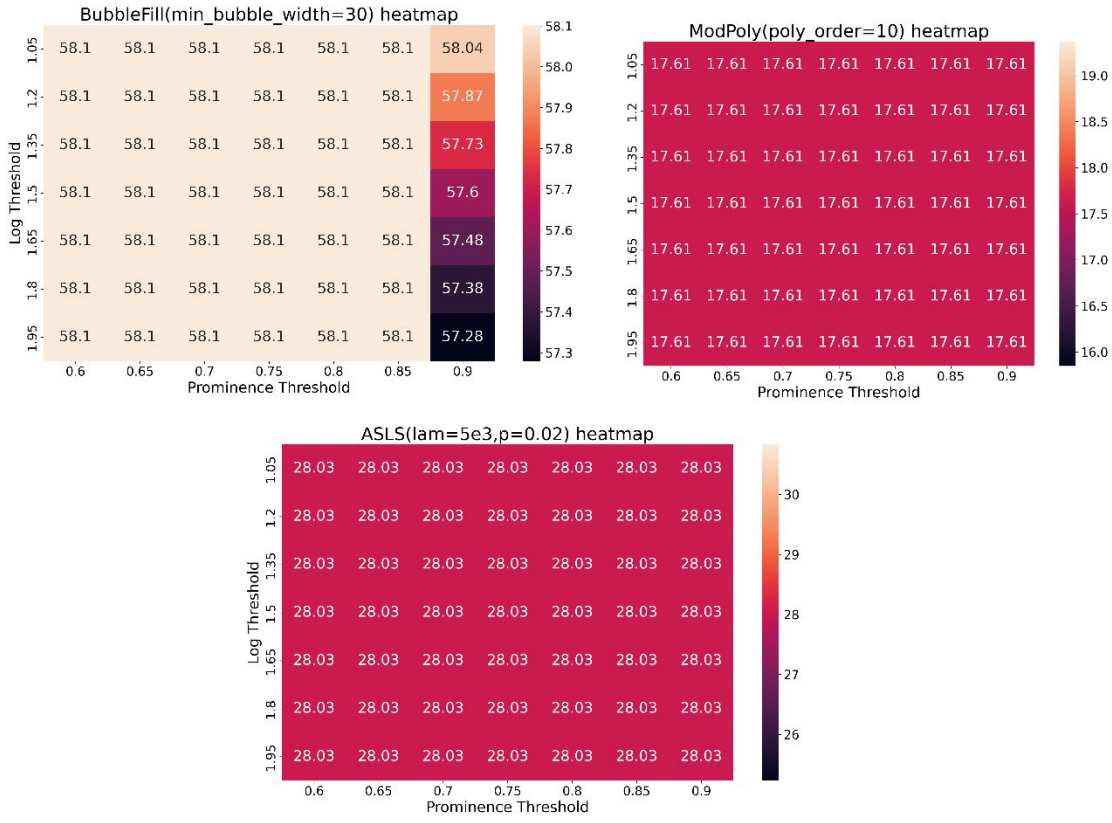


*Figure S19: Average IS-Score (%) obtained by varying prominence and logarithmic for each baseline tested for the **pharmaceutical** dataset.*

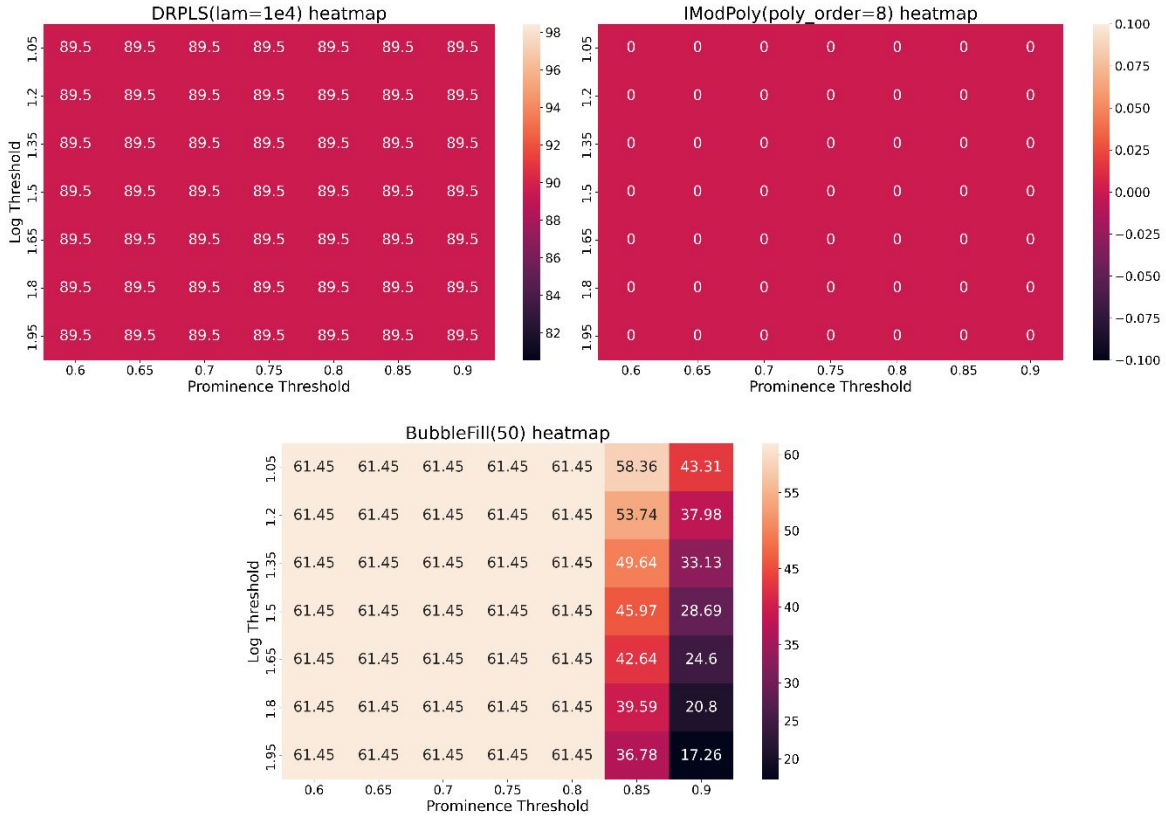


*Figure S20: IS-Score (%) obtained by varying prominence and logarithmic for each baseline tested for the **mineralogy** spectrum.*

3. Mean-Ratio Threshold

The mean-ratio threshold was varied from **1 - 10,** in steps of 1**.** Figure S21, S22, S23, S24, and S25 present the corresponding boxplots for each baseline across the five datasets: isolated red blood cells and polymer bag, red blood cells in polymer bag, food science, pharmaceutical and mineralogy. For the mineralogy dataset a scatter plot is presented because only one single spectrum is available. This threshold controls how sensitive the Mean-Ratio block of the IS-Score is to balance between properly positioned dips (below baseline) and improperly positioned dips (above baseline). Across the tested range of thresholds, the overall penalization behaviour remains consistent. The isolated red blood cells and polymer bag and mineralogy datasets remained stable across the entire range, whereas red blood cells in polymer bag, food science showed variability for threshold below 2 and pharmaceutical for threshold below 3. Threshold below 3 failed to detect subtle, but visually apparent, over correction, while threshold above 7 did not impact the final results. A value of 5 provide the most reliable balance, penalizing spectra when roughly ≥ 17% of dips lie above the baseline. This setting consistently captured meaningful over-correction without introducing excessive false positives. Although conceptually the penalty would be expected to trigger when more than half of the detected dips lie above the baseline, a more conservative threshold of 5 (i.e. allowing only 1 problematic dip per 5 dips below the baseline) was adopted. This choice provides a stricter control over this component of the algorithm while maintaining an appropriate balance with the other scoring blocks. Furthermore, this conservative threshold choice accounts also for the fact that dips may already be penalized by the single-band and region-band components of the IS-Score.

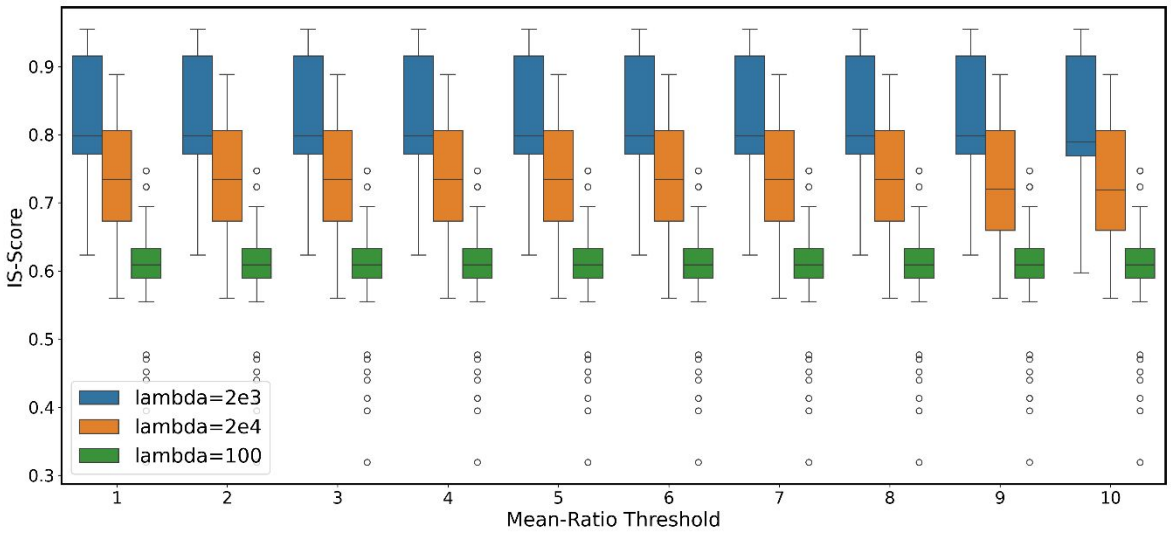


Figure S21: IS-Score obtained for Mean-Ratio threshold between 1 and 10 for each lambda value of the Asymmetric Least Square (ASLS) baseline, evaluated on *the **isolated red blood cells and polymer bags** dataset.*

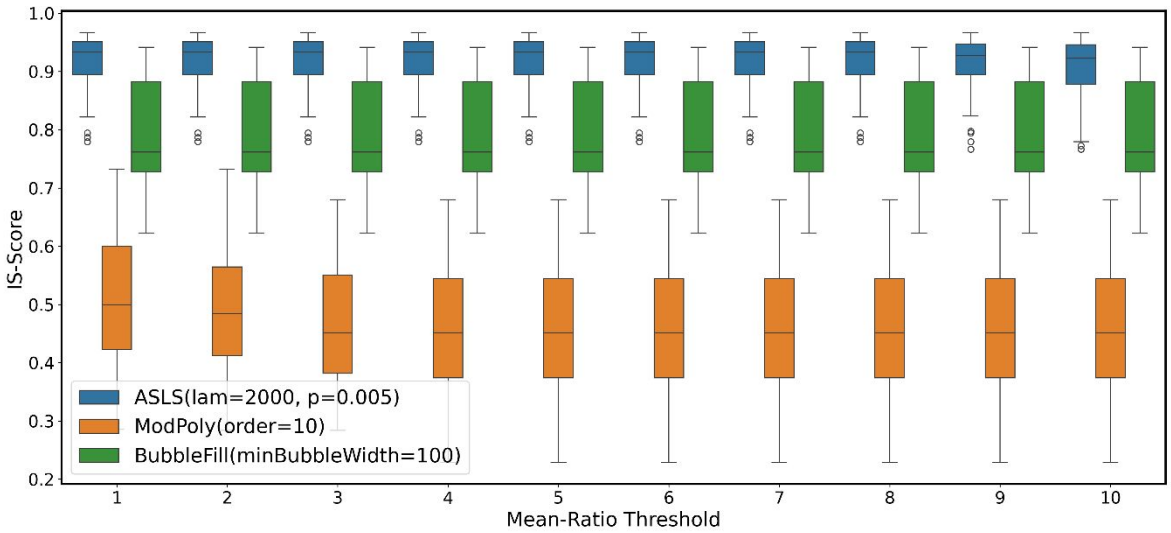


*Figure S22: IS-Score obtained for Mean-Ratio threshold between 1 and 10 for each baseline tested for the **red blood cells in polymer bags** dataset.*

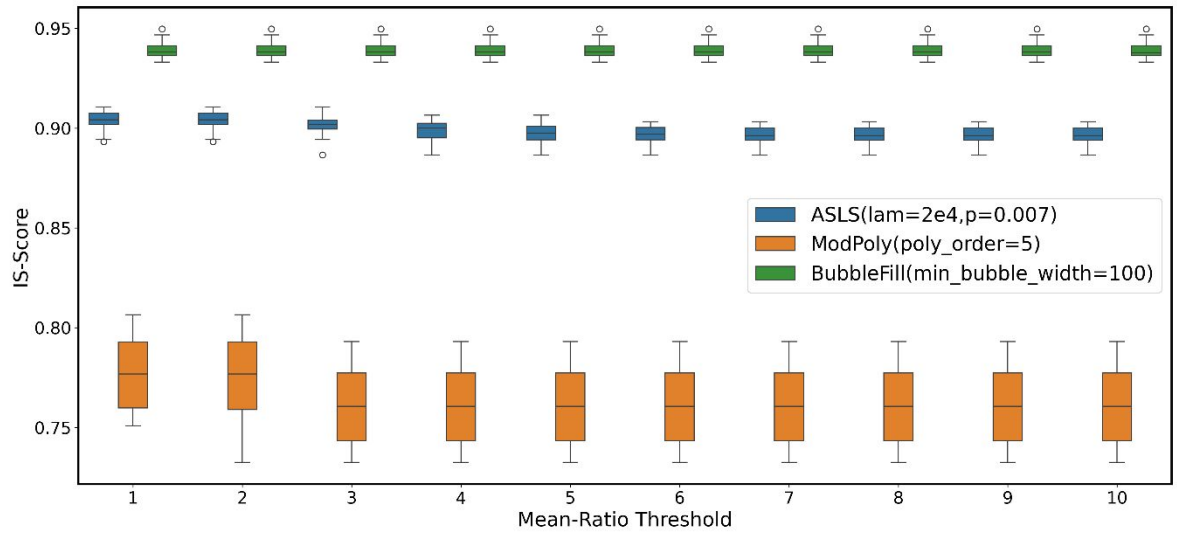


*Figure S23: IS-Score obtained for Mean-Ratio threshold between 1 and 10 for each baseline tested for the **food science** dataset.*

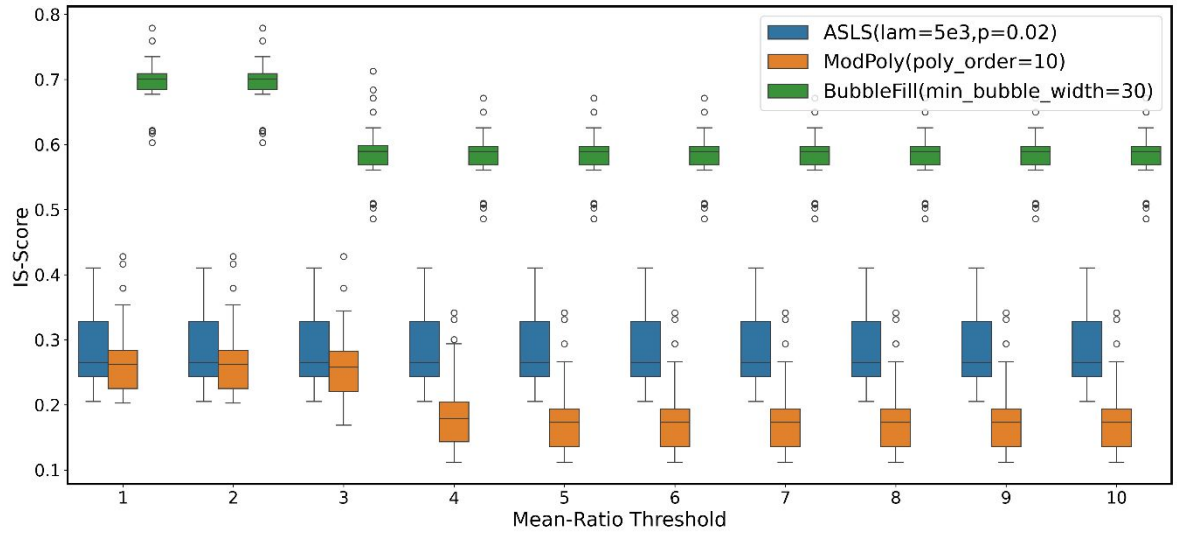


*Figure S24: IS-Score obtained for Mean-Ratio threshold between 1 and 10 for each baseline tested for the **pharmaceutical** dataset.*

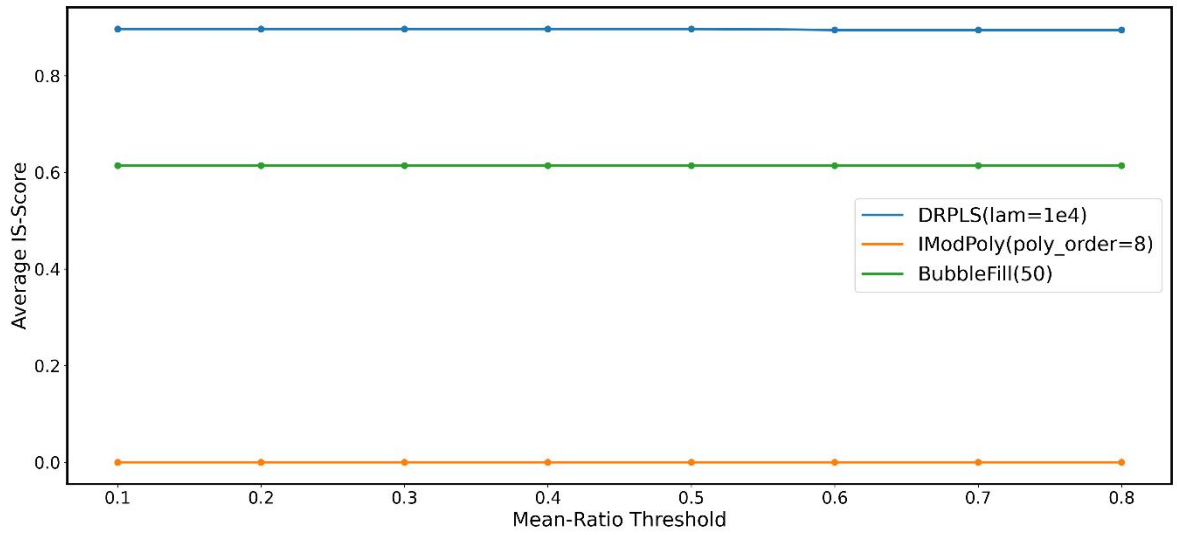


*Figure S25: IS-Score obtained for Mean-Ratio threshold between 1 and 10 for each baseline tested for the **mineral** spectrum.*

4. AUC Positive-Negative Threshold

The AUC threshold was varied from **0.10 – 0.80**, in steps of 0.10. Figure S26, S27, S28, S29, and S30 present the corresponding boxplots for each baseline across the five datasets: isolated red blood cells and polymer bag, red blood cells in polymer bag, food science, pharmaceutical and mineralogy. For the mineralogy dataset a scatter plot is presented because only one single spectrum is available. This threshold defines the sensitivity of the penalty applied in the AUC penalization block, by determining whether the ratio between the negative and positive areas is large enough to trigger a stronger penalty. Higher values of the thresholds make it less likely for the penalization condition to be triggered, resulting in a more conservative adjustment. Conversely, lower values increase the likelihood of applying a stronger penalty, making the method more sensitive. Across the tested range, the IS-Score remains generally stable, showing minimal variation and preserving the relative ranking of baseline correction methods. The threshold of 0.5 was selected to represent half of ratio between the negative and positive differences. In practice, when the ratio exceeds 0.5, at least half of the area deviation is negative, signalling that the baseline is likely overfitting. This choice provides a balanced criterion for applying the penalty, ensuring that only significant deviations trigger stronger penalization.

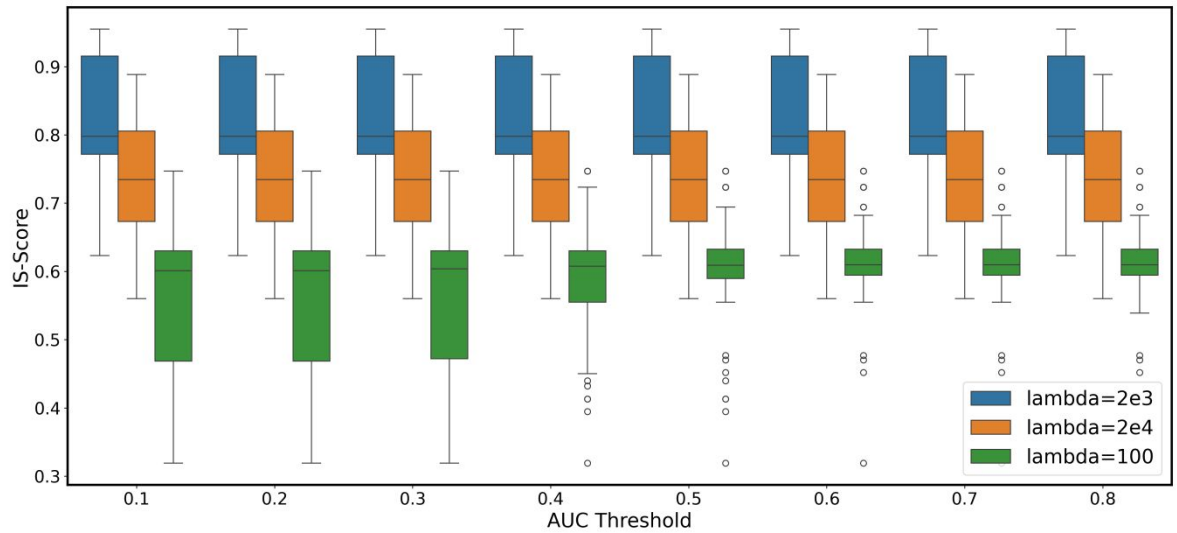


Figure S26: IS-Score obtained for AUC threshold between 0.1 and 0.8 for each lambda value of the Asymmetric Least Square (ASLS) baseline, evaluated on *the **isolated red blood cells and polymer bags** dataset.*

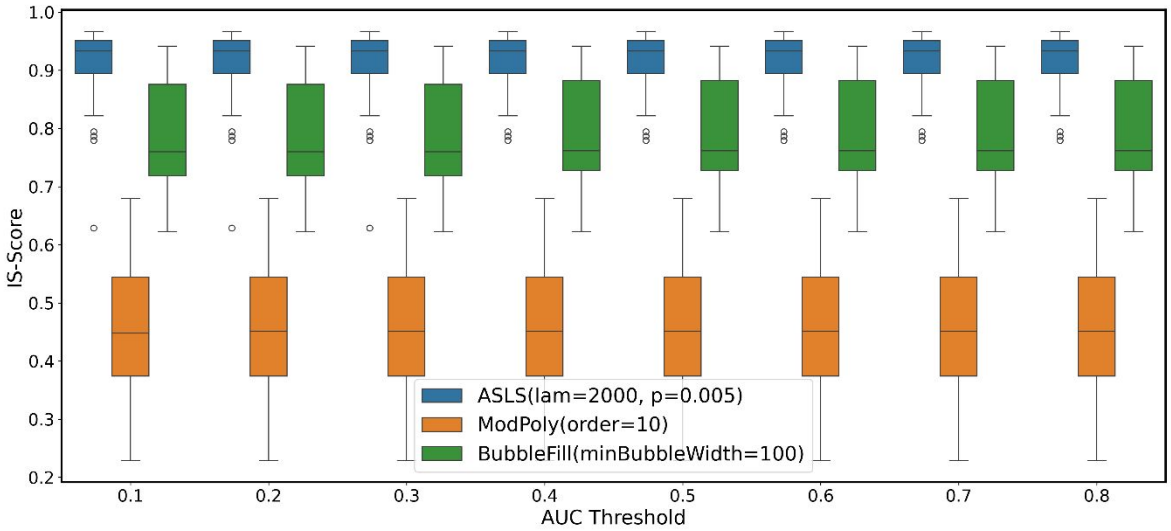


*Figure S27: IS-Score obtained for AUC threshold between 0.1 and 0.8 for each baseline tested for the **red blood cells in polymer bags** dataset.*

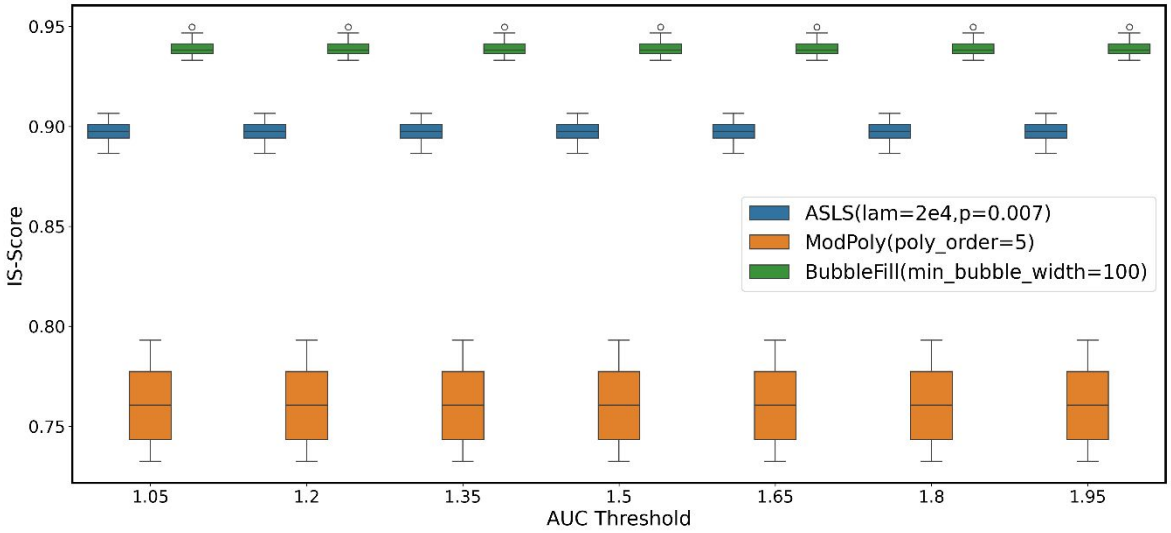


*Figure S28: IS-Score obtained for AUC threshold between 0.1 and 0.8 for each baseline tested for the **food science** dataset.*

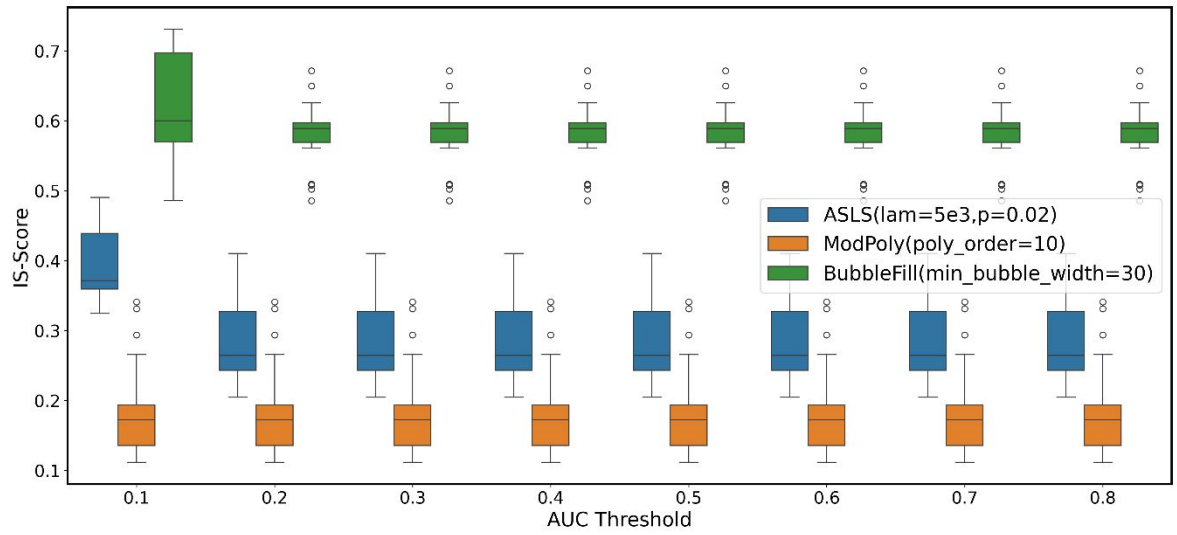


*Figure S29: IS-Score obtained for AUC threshold between 0.1 and 0.8 for each baseline tested for the **pharmaceutical** dataset*

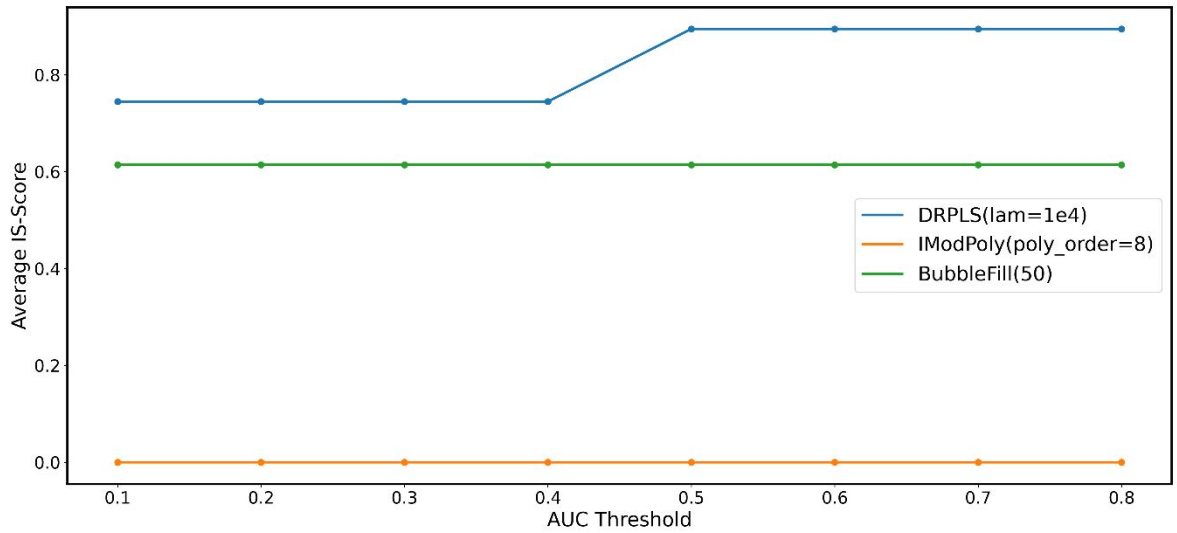


*Figure S30: IS-Score obtained for AUC threshold between 0.1 and 0.8 for each baseline tested for the **mineralogy** spectrum.*

Graphic User Interface (GUI)

To enhance usability and adoption, a GUI is provided alongside with the GitHub repository [4] containing the full IS-Score implementation. The GUI allows users to upload individual or multiple spectra, select from various baseline correction algorithms, and evaluate their performance using the IS-Score. Built on top of *RamanSPy* [5]- an open-source Python library for Raman spectroscopy analysis- the interface simplifies the integration of IS-Score existing workflows and promotes broader accessibility for non-expert users.

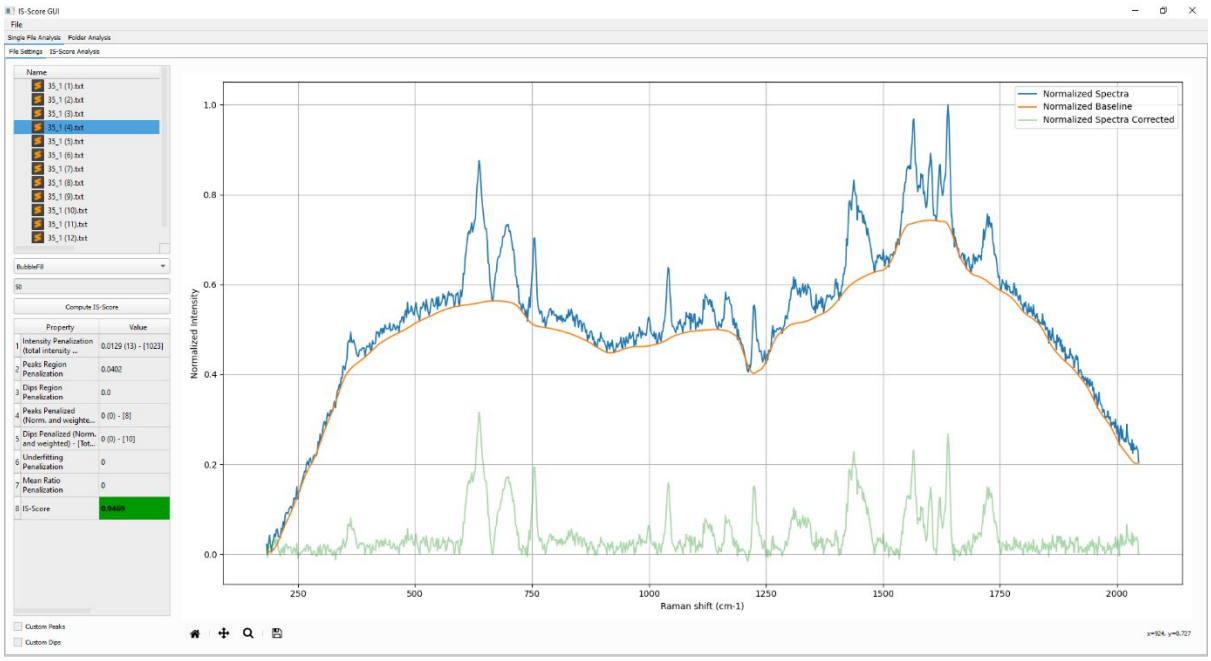


Figure S31: Main window of the GUI with a loaded spectrum and a baseline correction algorithm evaluated.

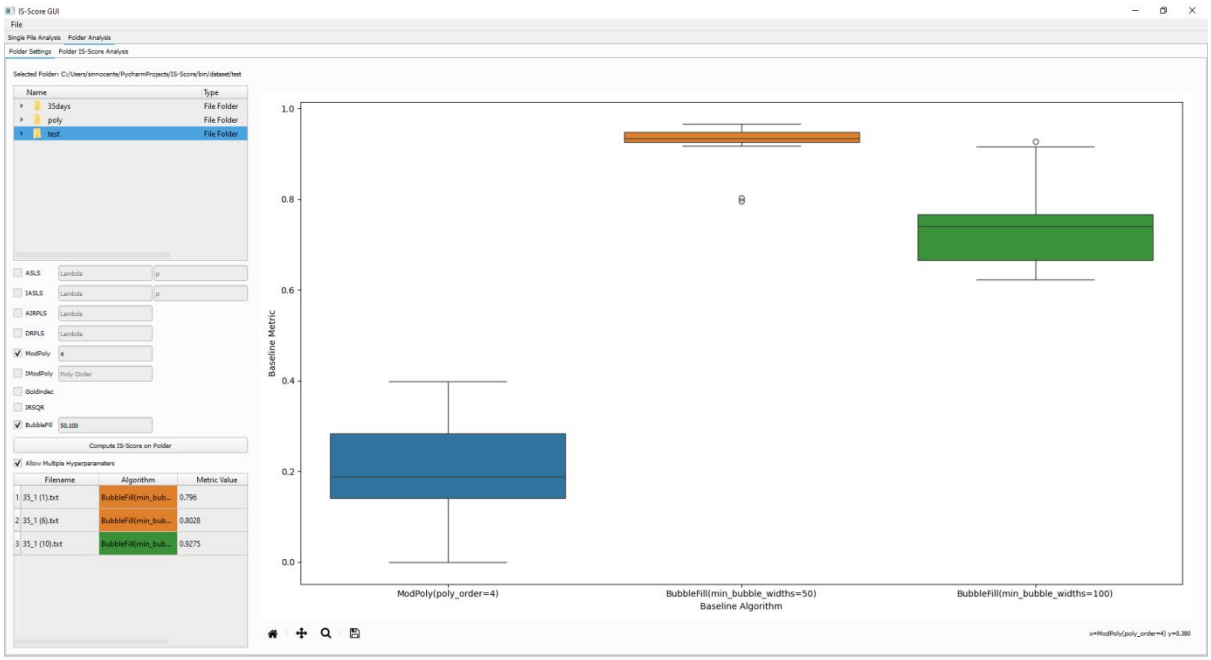


Figure S6: Boxplot of the IS-Score over a folder containing a set of spectra with three baseline correction algorithms tested.

References

1. Eggertson, E. C.; Venturini, F. Resonant Raman spectroscopy of carotenoids in aging of extra virgin olive oil. *Sensors* **2023,** 23(17), 7621
2. Rodas, J. R. R.; Ferreira, F.; Rivas-Martínez, G. I.; Melquiades, F. L.; Mello-Román, J. C.; Costa, A. M. D. S.; Amarilla, S.; Mendez, F.; Herrera, E. Z. Dataset of Raman Spectroscopy Responses for Over-The-Counter Drugs in Paraguay, Including Acetylsalicylic Acid, Paracetamol, and Ibuprofen. *Data in Brief* **2024**, *56*, 110844
3. RRUFF Project (2003–). RRUFF Database: Raman, X-ray, Infrared, and Chemistry Data. Available at: https://rruff.info/polyhalite/display=default
4. IS-Score GitHub repository: https://github.com/InnocenteSimone/IS_Score/
5. Georgiev, D.; Simon Vilms Pedersen; Xie, R.; Álvaro Fernández-Galiana; Stevens, M. M.; Barahona, M. RamanSPy: An Open-Source Python Package for Integrative Raman Spectroscopy Data Analysis. *Analytical chemistry* **2024**, *96(21)*, 8492-8500