Optimisation of Inelastic X-ray Scattering imaging parameters for the study of artists' pigments

Laure CAZALS[1] Loïc BERTRAND[1] Agnès DESOLNEUX[2]

[1]Université Paris-Saclay, ENS Paris-Saclay, CNRS, PPSM, 91190 Gif-sur-Yvette, France.

[2]Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, 91190 Gif-sur-Yvette, France.

Résumé – La diffusion inélastique des rayons X (IXS) est une méthode puissante pour étudier les matériaux. Il est préférable de réduire le temps d'irradiation afin de préserver au maximum les échantillons, ce qui peut affecter l'interprétation des données. Nous proposons une approche statistique pour optimiser les paramètres expérimentaux afin de tirer le meilleur parti d'un temps de collecte donné.

Abstract – Inelastic X-ray scattering (IXS) is a powerful method for studying materials. It is preferable to reduce the irradiation time in order to preserve the samples as much as possible, which can affect the interpretation of the data. We propose a statistical approach to optimise the experimental parameters in order to make the most of a given collection time.

1 Introduction

In recent years, high-brightness X-ray sources have been increasingly used to study cultural heritage materials for their unique ability to probe their chemical properties [1]. New spectral imaging approaches may require long exposure to irradiation, which constrains or sometimes hampers the study of precious and/or sensitive samples. Inelastic X-ray scattering (IXS) is a clear illustration of this difficulty. IXS imaging has emerged as a promising technique for characterising cultural heritage artifacts, in particular artists' pigments [4]. However, its unique capability to speciate low atomic number ('light') elements is counterbalanced by the very low probability of interaction between the probe and matter, leading to long irradiation time to collect interpretable data [2, 5].

In IXS spectral images, each pixel value is a spectrum, that is, a number of photons per unit time as a function of energy. Let T be the maximum time that one could spend on one pixel, *i.e.* on one spectrum, denominated *time budget*. T can be broken into $T = M \times t$ with M the number of energy points and t the acquisition time per energy point. Given the preciousness of cultural heritage objects, experimenters strive to reduce T as much as possible [3]. To achieve this, a compromise is often qualitatively set between experimental parameters to make the best use of T based on compositional priors about the sample. Here, we propose a statistical method for rigorously selecting the optimal parameters (t, M). Synthetic spectral images are modelled for different (t, M). By solving an inverse problem, we recover the species distribution from the ground truth. The optimal experimental parameters (t, M) are then obtained by minimizing the mean squared error between the ground truth and the results of the inverse problem, under the constraint that $T = t \times M$ is bounded. Beyond the optimal number M of energy points to use, we also provide an optimal set of M energy points from an exhaustive energy grid collected at $N \geqslant M$ points.

2 Statistical framework

2.1 Ground truth

Artists' pigments are generally mixtures of distinct minerals resulting from raw materials. The resulting 3D IXS image can be described by K (mineral) phases, defined here as groups of pixels sharing the same chemical composition. Each phase is characterized by a pure spectrum, called *phase spectrum* I_k^0 with $k \in [\![1, K]\!]$. The I_k^0's are the number of photons per unit time as a function of a discrete energy vector $\{e_n\}_{n=1,\dots,N}$, with N the total number of energy points. The ground truth of an IXS image is an image where each pixel is associated with a combination of phase spectra, so as to generate a piecewise linear image. Partial volume effects are generated at the phase boundaries resulting from signal sampling on pixelated detectors [8]. The signal (spectrum) for each pixel $\mathbf{x}$ is then formulated as:

$$I^0(\mathbf{x}, e_n) = \sum_{k=1}^{K} p_k^0(\mathbf{x}) I_k^0(e_n), \tag{1}$$

with $p_k^0(\mathbf{x})$ the factorization coefficients representing the weights of each phase k in pixel $\mathbf{x}$.

2.2 Observed signals

The signal is captured by photon-counting pixel detector systems producing a shot noise. For each energy point e_n and each pixel $\mathbf{x}$, the observed signal is a discrete random variable following a Poisson distribution:

$$S(\mathbf{x}, e_n) \sim \mathcal{P}(t \times I^0(\mathbf{x}, e_n)), \tag{2}$$

with t the acquisition time per energy point. All random variables $S(\mathbf{x}, e_n)$ are independant.

2.3 Solving the inverse problem

To retrieve the ground truth, *i.e.* the $p_k^0(\mathbf{x})$, from the noisy observations $S(\mathbf{x}, e_n)$, we are facing an inverse problem that

we solve using a least squares estimator— although it does not maximize the likelihood function, it is unbiased and allows an explicit solution unlike Poisson regression. More precisely, for each pixel $S(\mathbf{x}, e_n)$, denoting $\mathbf{S}(\mathbf{x}) = (S(\mathbf{x}, e_n))_n \in \mathbb{R}^N$ the observed signal, we are looking for:

$$\hat{\mathbf{p}}(\mathbf{x}) = \underset{\mathbf{p}\in\mathbb{R}^K}{\operatorname{argmin}} ||\mathbf{S}(\mathbf{x}) - \mathbf{A}\mathbf{p}||^2, \tag{3}$$

where the matrix $\mathbf{A} \in \mathbb{R}^{N\times K}$ is the linear measurement operator given by

$$\mathbf{A} = t \times \mathbf{I}^0 = t \times \begin{pmatrix} I_1^0(e_1) & \cdots & I_K^0(e_1) \\ \vdots & \ddots & \vdots \\ I_1^0(e_N) & \dots & I_K^0(e_N) \end{pmatrix}. \tag{4}$$

We get the following solution: $\mathbf{A}^{\mathrm{T}}\mathbf{A}\hat{\mathbf{p}} = \mathbf{A}^{\mathrm{T}}\mathbf{S}$. Since the phase spectrum are non-collinear, the matrix $\mathbf{A}$ has rank $K \leq N$, $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is invertible, and we obtain a unique solution:

$$\hat{\mathbf{p}}(\mathbf{x}) = \frac{1}{t}(\mathbf{G}^0)^{-1}(\mathbf{I}^0)^{\mathrm{T}}\mathbf{S}(\mathbf{x}) \tag{5}$$

with $\mathbf{G}^0 = (\mathbf{I}^0)^{\mathrm{T}}\mathbf{I}^0$, the Gram matrix of the I_k^0's, that is a symmetric positive $K \times K$ matrix. The least squares estimator is indeed unbiased, since we have

$$\mathbb{E}(\hat{\mathbf{p}}(\mathbf{x})) = \frac{1}{t}(\mathbf{G}^0)^{-1}(\mathbf{I}^0)^{\mathrm{T}}\mathbb{E}(\mathbf{S}(\mathbf{x})) = \mathbf{p}^0(\mathbf{x}).$$

2.4 Optimisation of experimental parameters

We have already seen that the estimated factorization coefficients $\hat{\mathbf{p}}$ are unbiased, but we also want to control their variance that is the mean-squared error of the phase composition of the material. The covariance of the estimated factorization coefficients is given by

$$\mathrm{Cov}(\hat{\mathbf{p}}(\mathbf{x})) = \frac{1}{t^2}(\mathbf{G}^0)^{-1}(\mathbf{I}^0)^{\mathrm{T}}\mathrm{Cov}(\mathbf{S}(\mathbf{x}))\mathbf{I}^0(\mathbf{G}^0)^{-1}, \tag{6}$$

with

$$\mathrm{Cov}(\mathbf{S}(\mathbf{x})) = \mathrm{Diag}(t\mathbf{I}^0(\mathbf{x}, e_n)), \tag{7}$$

as $\{S(\mathbf{x}, e_n) \sim \mathcal{P}(t \times I^0(\mathbf{x}, e_n))\}_{n=1,\dots,N}$ are independent. We have the following inequality between positive symmetric matrices: $\mathrm{Cov}(\mathbf{S}(\mathbf{x})) \leqslant tCI_N$, with $C = \max_{\mathbf{x},e_n} \mathbf{I}^0(\mathbf{x}, e_n)$ a constant and I_N the identity matrix of $\mathbb{R}^N$. We thus get:

$$\mathrm{Cov}(\hat{\mathbf{p}}) \leq \frac{C}{t} \times (\mathbf{G}^0)^{-1}. \tag{8}$$

And therefore

$$\mathbb{E}\left(\|\hat{\mathbf{p}}(\mathbf{x}) - \mathbf{p}^0(\mathbf{x})\|^2\right) \leq \frac{C}{t} \times \mathrm{Tr}\left((\mathbf{G}^0)^{-1}\right). \tag{9}$$

To include the selection of M energy points among the N at a fixed t, we define the function

$$\mathrm{F}(\boldsymbol{\omega}) = \mathrm{Tr}(\mathbf{G}^0(\boldsymbol{\omega})^{-1}), \tag{10}$$

where the vector $\boldsymbol{\omega} \in \Omega_M = \{(\omega_1, \dots, \omega_N) \in \{0,1\}^N$ with $\sum_{n=1}^N \omega_n = M\}$ is the energy points selection vector, and the matrix $\mathbf{G}^0(\boldsymbol{\omega})$ is the Gram matrix restricted to the selected energy points:

$$\mathbf{G}^0_{k,l}(\boldsymbol{\omega}) = \sum_{n=1}^{N} \omega_n I_k^0(e_n) I_l^0(e_n), \tag{11}$$

with $k, l \in [\![1, K]\!]$. For a fixed t and M, the goal is to identify the $\boldsymbol{\omega} \in \Omega_M$ that minimizes $\mathrm{F}(\boldsymbol{\omega})$, resulting in an optimal selection of energy points. Evaluating all possible $\boldsymbol{\omega} \in \Omega_M$ energy subsets for each M is computationally intensive. Therefore, we reformulate the problem in a continuous relaxed convex form, allowing the use of an optimization approach based on the Frank-Wolfe algorithm [6]. The minimization of the function $\mathrm{F}(\boldsymbol{\omega}) = \mathrm{Tr}(\mathbf{G}^0(\boldsymbol{\omega})^{-1})$ (that can be shown to be a convex function) is performed on the convex set $\boldsymbol{\omega} \in \Omega'_M = \{(\omega_1, \dots, \omega_N) \in [0,1]^N$ with $\sum_{n=1}^N \omega_n \leq M\}$. At each iteration J of the Frank–Wolfe algorithm, we minimize the linear approximation of the problem by selecting the M smallest values of $\nabla\mathrm{F}$ — for which we have an explicit formula with negative components. We denote the result of the Frank-Wolfe algorithm as $\boldsymbol{\omega}^*_M$ (Algorithm 1).

Algorithm 1: Frank-Wolfe algorithm

```
Initialization ω := {M/N, ..., M/N}
while J < J_max do
    Compute the gradient ∇F(ω)
    Assign the M smallest values of ∇F to 1, and 0
      otherwise, in z.
    Update ω ⟵ ω + 2/(J+M) (z − ω)
    Update J ⟵ J + 1
end
Sort ω
return Indices of the M largest values of ω
```

From Equation (9), we see that, as the exposure time per energy point t increases, the MSE on the obtained factorization coefficients will decrease. This is explained by the fact that the physical model is a Poisson noise. From Equation (10), and computing the gradient of F, we observe that it is a decreasing function of $\boldsymbol{\omega}$. Therefore, increasing the number of energy points decreases the MSE. This is explained by the fact that adding more information in the inverse problem allows reducing the uncertainty. Now, when $T = t \times M$ is fixed, it is not clear what the optimal choice of (t, M) is since both should be as large as possible to minimize the MSE. We solve this problem in the next section using a digital twin approach.

3 Simulation results

3.1 Synthetic and experimental data

IXS spectra were experimentally collected on $K = 5$ artists' pigments at the GALAXIES beamline of the SOLEIL synchrotron facility [7] over $N = 124$ energy points, covering $\approx$ 24–76 eV [4]. The spectra were smoothed and assigned to I_k^0. The phase spectra were labelled: *1-light turquoise*, *2-zinc white*, *3-burnt umber*, *4-indigo veritable* and *5-indian yellow* (Figure 1A). A 2D synthetic image was produced using drawing software to represent the pure and mixed areas composed

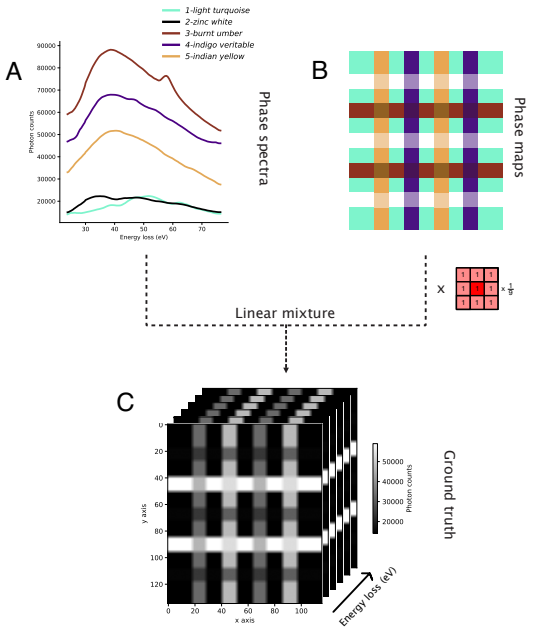

Figure 1 – **Construction of the ground truth as a combination of experimental and numerical data.** (1) The phase spectra I_k^0 are experimentally collected on raw materials and denoised. (2) A phase map is produced synthetically. (3) The hybrid datacube is created by injecting the I_k^0's into the phase map $p_k^0(\mathbf{x})$, yielding the ground truth.

by the five artists' pigments. Phase maps are obtained by (1) assigning $p_k^0 = 1$ where the pigments are pure and $p_k^0 = 0.5$ where the lines cross (Figure 1B) — in accordance with the most common case studies, (2) convolving the resulting phase map with a kernel of size 3×3 to model the partial volume effect at the shapes boundaries. By injecting the experimental spectra of the pure phases I_k^0 in the synthetic phase map, a hybrid ground truth is obtained (Figure 1C).

3.2 Comparing different energy grids

We applied the digital twin, keeping T fixed and varying M between $K = 5$ and $N = 124$, using either a regular grid of M energy points or the optimised grid obtained with the Frank-Wolfe algorithm. The noisy images were factorized using the least squares estimator method by injecting the I_k^0's subsampled at the corresponding energy grid point. To evaluate the best distribution of energy points, the mean squared error (MSE) is defined as a function of M:

$$\mathrm{MSE}(M) = \frac{1}{L_x \times L_y} \sum_{\mathbf{x}} \|\hat{\mathbf{p}}(\mathbf{x}) - \mathbf{p}^0(\mathbf{x})\|^2, \qquad (12)$$

where $\hat{\mathbf{p}}(\mathbf{x})$ is in fact a function of $\mathbf{G}^0$, that is itself a function of M. We note $L_x \times L_y$, the image size.

For different regular grids of M points, the MSE is calculated (Figure 2, red curve). An unstable behavior is observed for low M, attributed to the inability of the regular grid to capture discriminating spectral features, such as the peak at ≈ 60 eV in the *3-burn umber* spectrum, when collecting at sparse energy points. As M increases, the MSE tends to converge towards a constant value, demonstrating the equivalence

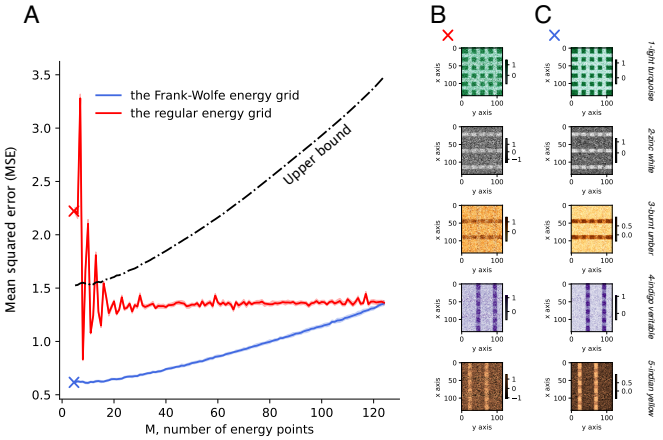

Figure 2 – **Mean squared error as a function of M at T fixed.** (A) Experimental MSE of data projected onto the regular energy grid and the Frank–Wolfe grid, along with the upper bound of MSE at energy points selected by the Frank–Wolfe algorithm as a function of M. Experimental curves represent the mean of 15 numerical experiments, with the corresponding 95% confidence interval. The resulting phase maps for $M = K$ for data projected onto the regular grid (B) and onto the Frank-Wolfe grid (C).

of *hyperspectral* (low t, high M) and *multispectral* (high t, low M) collection modalities when $M \gtrsim 40$

The Frank-Wolfe algorithm was applied to the phase spectra to calculate an optimal energy grid minimising the MSE (Figure 3). We observe that when a point is selected, it has a high chance of being kept when M increases. The resulting grid was evaluated using the digital twin (Figure 2, blue curve). Compared to experimental practice, our optimal approach is more efficient. In particular, at $M = K$, the difference in the ability of the two grids to recover the phase distribution is clear when we look at the phase maps obtained (Figure 2B, C). In addition, the trend in experimental MSE derived from the Frank-Wolfe optimised grid shows that the scenario in which M is minimal (and therefore t maximal) is the one that minimises the MSE. We note that the upper bound (Equation 9) for high M appears to be less optimal than the experimental MSE derived from the regular grid. Although the Frank-Wolfe optimisation seems successful, the factor used in the upper bound is rather coarse.

4 Conclusion

Compared with traditional experimental scenarios, the optimal selection of (t, M) parameters considerably improves data processing when irradiation time is limited. It is interesting to note that the best compromise is to collect low M at high t (that is, the multispectral modality rather than the hyperspectral one).The guidelines extracted from this work make it possible to safely collect IXS images of artists' pigments. We believe that such approaches could be extended to other imaging devices and samples and integrated into routine pre-run phases.

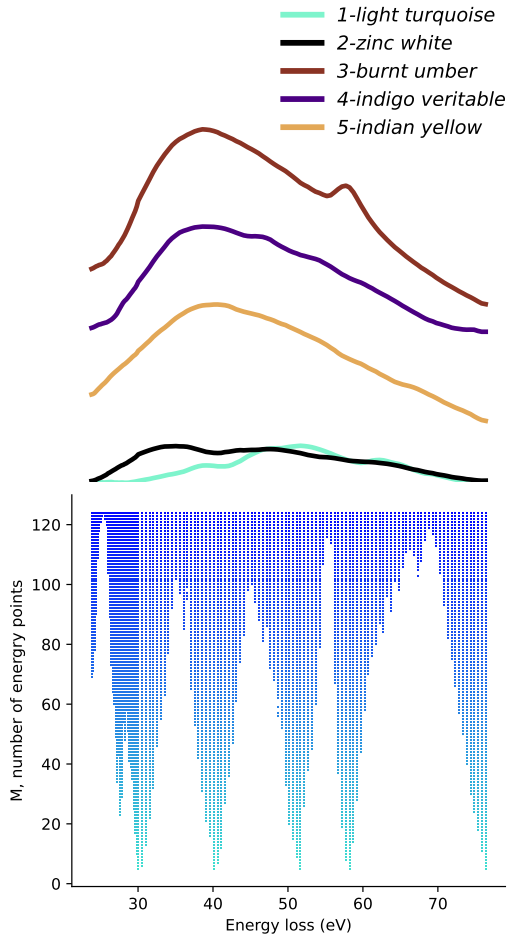

Figure 3 – **Optimal selection of energy points.** The selected optimal energy points, $\boldsymbol{\omega}_M^*$, are highlighted with markers. The Frank-Wolfe algorithm with 1,000 iterations was used for each M based on the I_k^0's. The marker density appears higher for energy points below 30 eV due to the smaller energy step size in the corresponding range. We also show the spectra above to visualize the position of the selected energy points.

Acknowledgements

We would like to thank Jean-Pascal Rueff for his support at the GALAXIES beamline (SOLEIL synchrotron, Saint-Aubin), and Lauren Dalecky (PPSM) for providing the IXS spectra. We thank Serge X. Cohen (IPANEMA) for valuable discussion about this work.

References

[1] Uwe Bergmann, Loïc Bertrand, Nicholas P. Edwards, Phillip L. Manning, and Roy A. Wogelius. Chemical mapping of ancient artifacts and fossils with x-ray spectroscopy. In Eberhard Jaeschke, Shaukat Khan, Jochen R. Schneider, and Jerome B. Hastings, editors, *Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications*, Springer Reference. Springer Nature Switzerland, 2019.

[2] Uwe Bergmann, Pieter Glatzel, and Stephen P. Cramer. Bulk-sensitive XAS characterization of light elements: from X-ray Raman scattering to X-ray Raman spectroscopy. *Microchemical Journal*, 71(2):221 – 230, 2002.

[3] Loïc Bertrand, Sebastian Schöder, Ineke Joosten, Samuel M. Webb, Mathieu Thoury, Thomas Calligaro, Étienne Anheim, and Aliz Simon. Practical advances towards safer analysis of heritage samples and objects. *TrAC Trends in Analytical Chemistry*, 164, 2023.

[4] Lauren Dalecky, Francesco Sottile, Linda Hung, Laure Cazals, Agnès Desolneux, Aurélia Chevalier, Jean-Pascal Rueff, and Loïc Bertrand. Non-resonant inelastic X-ray scattering for discrimination of pigments. *Physical Chemistry Chemical Physics*, 26(5):4363–4371, 2024.

[5] Rafaella Georgiou, Christoph J. Sahle, Dimosthenis Sokaras, Sylvain Bernard, Uwe Bergmann, Jean-Pascal Rueff, and Loïc Bertrand. X-ray Raman Scattering: A Hard X-ray Probe of Complex Organic Systems. *Chemical Reviews*, 122(15):12977–13005, August 2022.

[6] Martin Jaggi. Revisiting frank-wolfe: Projection-free sparse convex optimization. In *International conference on machine learning*, pages 427–435. PMLR, 2013.

[7] J.-P. Rueff, J. M. Ablett, D. Céolin, D. Prieur, Th. Moreno, V. Balédent, B. Lassalle-Kaiser, J. E. Rault, M. Simon, and A. Shukla. The GALAXIES beamline at the SOLEIL synchrotron: inelastic X-ray scattering and photoelectron spectroscopy in the hard X-ray range. *Journal of Synchrotron Radiation*, 22(1):175–179, January 2015.

[8] Philip J Withers, Charles Bouman, Simone Carmignato, Veerle Cnudde, David Grimaldi, Charlotte K Hagen, Eric Maire, Marena Manley, Anton Du Plessis, and Stuart R Stock. X-ray computed tomography. *Nature Reviews Methods Primers*, 1(1):18, 2021.