# Multi-modal and multi-scale non-local means method to analyze spectroscopic datasets

**Niklas Mevenkamp[1], Katherine MacArthur[2], Vasiliki Tileli[3], Philipp Ebert[4], Leslie J. Allen[2,5], Benjamin Berkels[1], Martial Duchamp[2,6*]**

1 – AICES Graduate School, RWTH Aachen University, Aachen, Germany
2 – Ernst Ruska-Centrum, Forschungszentrum Jülich GmbH, Jülich, Germany
3 – Institute of Materials, École polytechnique fédérale de Lausanne, CH-1015 Lausanne, Switzerland.
4 – Peter Grünberg Institut, Forschungszentrum Jülich GmbH, Jülich, Germany
5 – School of Physics, University of Melbourne, Parkville, Victoria 3010, Australia
6 – School of Materials Science and Engineering, Nanyang Technological University, Singapore

* Corresponding author

**Abstract**

A multi-modal and multi-scale non-local means (M3S-NLM) method is proposed to extract atomically resolved spectroscopic maps from low signal-to-noise (SNR) datasets recorded with a transmission electron microscope. This method improves upon previously tested denoising techniques as it takes into account the correlation between the dark-field signal recorded simultaneously with the spectroscopic dataset without compromising on the spatial resolution. The M3S-NLM method was applied to electron energy dispersive X-ray and electron-energy-loss spectroscopy (EELS) datasets. We illustrate the retrieval of the atomic scale diffusion process in an $Al_{1-x}In_xN$ alloy grown on GaN and the surface oxidation state of perovskite nanocatalysts. The improved SNR of the EELS dataset also allows the retrieval of atomically resolved oxidation maps considering the fine structure absorption edge of $LaMnO_3$ nanoparticles.

## 1. Introduction

In the last decade, aberration correction in transmission electron microscopes has led to atomic resolution electron energy-loss spectroscopy (EELS) and energy-dispersive X-ray (EDX) mapping. In EELS atomically-resolved compositional[1] and valence mapping[2] as well as spectroscopic analysis at the single atom level[3] have been demonstrated. Atomic resolution EDX mapping was first demonstrated in 2010[4] and recently put on an absolute scale[5]. The low inelastic scattering cross-section of high-energy electrons with core electron atomic levels implies, however, that only a small fraction of the incoming electron beam current contributes to analytical signals inside a transmission electron microscope. The application of atomically resolved spectroscopic transmission electron microscopy (TEM) techniques to specimens that are electron-beam sensitive or change as a function of exposure time, especially for applications in the life sciences and for studies of soft matter, is still to be demonstrated.

Post-processing techniques based on multivariate statistical analysis (MVA) are used to improve the signal-to-noise ratio (SNR) of spectroscopic datasets by taking advantage of redundancy in the information. Starting with principal component analysis[6,7] in the late 90's, followed by more advanced decomposition techniques such as independent component analysis[8], non-negative matrix factorization[9-11] or geometric-based extraction methods like vertex component analysis[12-14] have greatly improved the exploitability of the spectroscopic data. However, Lichtert and Verbeeck have shown MVA-based techniques can introduce a significant bias, which may alter the conclusions drawn from noisy experimental data sets[15]. On the other hand, non-local denoising algorithms[16] such as non-local means[17], the block-matching and a 3D filtering algorithm[18], consider the entire image for averaging and select the set of pixels associated with a reference pixel based on the similarity of the image content within neighborhoods around the respective pixels (instead of their spatial vicinity). Given that objects of interest within the image actually occur multiple times, this approach has proven to offer a faithful reconstruction while sacrificing much less spatial resolution than with local

averaging, or none at all[19]. In particular, previous studies indicate that the non-local denoising approach enables a high quality reconstruction of low dose high-resolution scanning (S-)TEM images or EELS scans[20-22]. EELS and EDX datasets are commonly recorded simultaneously with an imaging STEM signal with a higher SNR. The temporal and spatial correlation between the imaging and analytical signals makes some information redundant. Estimating the genuine image intensity or spectrum contrast at a given reference pixel is typically done by averaging over a set of somehow associated values, e.g., over a local window centered around the reference pixel. Thereby, image details that are smaller than the window size are blurred, resulting in a loss of spatial resolution. Here, we present a denoising technique, namely the multi-modal and multi-scale non-local means (M3S-NLM) method. It combines the image information of jointly acquired scans of possibly different modalities (e.g. high-angle annular-dark-field (HAADF)-STEM and EELS or EDX signals) in order to select sets of associated pixels for averaging. In the rest of the manuscript, STEM, EELS and EDX refer to HAADF-STEM, EELS-STEM and EDX-STEM unless otherwise stated.

## 2. Materials and methods

### 2.1 Multi-modal and multi-scale non-local means method

#### 2.1.1 Non-local means method

A widely used technique for noise reduction is mean filtering. For a given window size (2s+1 by 2s+1) with a positive integer s, each pixel is replaced with the average of all pixels in the window:

$$I_{MEAN}(x,y)=\frac{1}{(2s+1)^2}\sum_{\Delta x=-s}^{s}\sum_{\Delta y=-s}^{s} I(x+\Delta x, y+\Delta y) \quad \text{eq. 1}$$

where $I(x,y)$ is the pixel intensity at coordinate $(x,y)$ .

The disadvantage of such a basic denoising approach is that it smears out fine details and thereby reduces the spatial resolution of the reconstructed images on the scale of the window size. Modern denoising methods try to address this problem by utilizing the structure of the given images. For instance, the NLM incorporates weights in the above average that measure how similar the structure surrounding the respective pixels is (within a given structure window of size (2s+1 by 2s+1)). The decision on the structure window size is balanced between robustness to noise and preserving fine features (the larger it is, the more robust to noise it becomes, but at the same time it becomes more unlikely to find a larger block which actually has the same structure). The search for similarities is performed over a so-called search window of size ($2W_x$ +1 by $2W_y$+1), ideally the full image. The intensity of the NLM denoised image is computed as:

$$I_{NLM}(x,y)=\frac{\sum_{\Delta w_x=-W_x}^{W_x}\sum_{\Delta w_y=-W_y}^{W_y} I(x+\Delta w_x, y+\Delta w_y)\, w(I,(x,y),(x+\Delta w_x, y+\Delta w_y))}{\sum_{\Delta w_x=-W_x}^{W_x}\sum_{\Delta w_y=-W_y}^{W_y} w(I,(x,y),(x+\Delta w_x, y+\Delta w_y))} \quad \text{eq. 2}$$

with weights

$$w(I,(x,y),(\hat{x},\hat{y}))=\exp\left[\frac{-d(I,(x,y),(\hat{x},\hat{y}))}{h^2}\right] \quad \text{eq. 3}$$

that represent the structure window similarity, which is defined by the structure window distance (eq. 4).

$$d(I,(x,y),(\hat{x},\hat{y}))=\sum_{\Delta x=-s}^{s}\sum_{\Delta y=-s}^{s}\left(I(x+\Delta x, y+\Delta y)-I(\hat{x}+\Delta x, \hat{y}+\Delta y)\right)^2 \quad \text{eq. 4}$$

The filter parameter $h$ is used to tweak the amount of denoising applied to the image, i.e. larger $h$ will result in stronger denoising and for $h \to 0$ less noise will be removed. $h$ is connected to the noise standard deviation (σ) of the image. If known, $h = \sqrt{2}\sigma$ is a good choice for the filter parameter. In this study, STEM images were transformed using the generalized Anscombe transformation and EDX datasets using the classical Anscombe transformation as the latter mainly contained Poisson noise. In such a case, we have $\sigma = 1$ and $h = \sqrt{2}$ . In eq. 2, the size of the search window is defined by $(W_x, W_y)$ . Ideally, for each reference pixel $(x, y)$ the entire image is searched for similar features and used to estimate the average pixel intensity $I_{NLM}(x, y)$ . However, in practice, this is typically too computationally intensive, since the effort then grows as the square of the number of pixels in contrast to mean filtering, which requires only linear effort in the number of pixels. Thus, $(W_x, W_y)$ are typically restricted to a reasonably sized search window. The similarity of two blocks (defined by the structure window) is represented by the values of the structure window distance, $d$ . A low value of $d$ means the block centered at $(x, y)$ is similar to the block centered at $(x + \Delta w_x, y + \Delta w_y)$ , a large value of $d$ means they are not similar. This reflects in the values of the weights (w), low $d$ means $w \to 1$ , large $d$ means $w \to 0$ . It turns out the double sum over the search window in the numerator of the eq. 2 will add pixels that have similar neighbors. The whole NLM procedure is repeated by shifting the reference block pixel by pixel and line by line across the entire scanned area resulting in a less noisy new image.

**2.1.2 Multi-modal and multi-scale non-local means (M3S-NLM)**

In the case of STEM images, non-local averaging algorithms have been found particularly suited to find sufficient similar features for averaging. Recently, a refined version of the NLM algorithm, namely the block matching and 3d filtering (BM3D) algorithm[18, 23] that is similar to the NLM algorithm, has been found to outperform the NLM algorithm for STEM images[20]. Therefore, the STEM images were reconstructed using the BM3D algorithm and the EDX/EELS datasets using the NLM algorithm. In the following, we describe how the NLM principle can be extended to multi-modal image acquisition, in particular in the context of STEM and EDX/EELS parallel imaging. A more detailed description is given in Ref. 24. For the STEM modality ( $I_{STEM}$ ), the structure window distance is defined as follows:

$$d_{STEM}(I, \vec{u}, \vec{v}) = d(I_{STEM}, \vec{u}, \vec{v}) \quad \text{eq. 5}$$

where $\vec{u} = (x, y)$ and $\vec{v} = (\hat{x}, \hat{y})$ . We will keep the vectorial notation in the rest of the manuscript.

For the EDX modality, we observed that the SNR is so low that a comparison of individual spectra at different positions would be meaningless. Thus, we need to define a structure window similarity that is more robust to extremely low SNRs. To this end, we propose to first down-sample the spectral channels and then to compare the two STEM and EDX structure windows. Thus, let $I_{EDX}$ denote an EDX image intensity with $n_c$ the number of spectral channels. Let $\hat{I}_{EDX}$ denote a resampled version of $I_{EDX}$ where each $r_c$ consecutive channels have been averaged together and $c$ being the channel of interest, i.e.,

$$\hat{I}_{EDX}(\vec{u}, c) = \frac{1}{r_c} \sum_{k=1}^{r_c} I_{EDX}(\vec{u}, (c-1) r_c + k) \quad \text{eq. 6}$$

Thus, $\hat{I}_{EDX}$ now has $n_c^r = \frac{n_r}{r_c} < n_c$ channels.

To further increase the SNR of the EDX dataset and prior applying the M3S-NLM procedure, we define the spatially averaged structure window distance for the EDX modality ( $d_{EDX}$ ), as the sum over the binned energy channels of the squared difference of the intensities averaged over a given window (2s+1 by 2s+1 pixels) between a set of two pixels $(\vec{u},\vec{v})$ and given by:

$$d_{EDX}(\vec{u},\vec{v}) = \sum_{c=1}^{n_c^r} (\hat{I}_{EDX}^{MEAN}(\vec{u},c) - \hat{I}_{EDX}^{MEAN}(\vec{v},c))^2 \quad \text{eq. 7}$$

where:

$$\hat{I}_{EDX}^{MEAN}(\vec{u},c) = \frac{1}{(2s+1)^2} \sum_{\Delta\vec{s}\in N} \hat{I}_{EDX}(\vec{u}+\Delta\vec{s},c) \quad \text{eq. 8}$$

and where $\Delta\vec{s} = (\Delta x, \Delta y)$ and $N$ being the structure window of size (2s+1 by 2s+1) centered at (0,0) used instead of the double sum eq. 4. $d_{EDX}(\vec{u},\vec{v})$ is a measure on how similar the EDX spectra at positions $\vec{u}$ and $\vec{v}$ are. Now, for sufficiently large windows sizes, the similarity measure for EDX structure windows is meaningful even for extremely low SNRs. However, it is also much less sensitive to spatial variations due to the spatial average performed via eq. 8. To alleviate this issue, we combine the advantages of high spatial resolution and fidelity encoded in the STEM structure window distances, $d_{STEM}$ , with the spectral resolution of the EDX structure window distances, $d_{EDX}$ . Due to the different modalities, these two structure window distances have completely different scales. To account for this, we normalize them with the reference structure window values:

$$d_{STEM}^{ref}(\vec{u}) = \sum_{\Delta\vec{w}\in M} (I_{STEM}(\vec{u}+\Delta\vec{w}))^2 \quad \text{eq. 9}$$

$$d_{EDX}^{ref}(\vec{u}) = \sum_{c=1}^{n_c^r} (\hat{I}_{EDX}^{MEAN}(\vec{u}+\Delta\vec{w},c))^2 \quad \text{eq. 10}$$

where $\Delta\vec{w} = (\Delta w_x, \Delta w_y)$ and $M$ being the search window of size ($2W_x$ +1 by $2W_y$+1).

Then, we define the joint similarity as the normalized sum of the similarities of all modalities, in our case the STEM image with EELS/EDX dataset. The combined multi-scale and multi-modal structure window distance is then defined as the following weighted normalized convex combination of the two individual structure window distances:

$$d_{M3S-NLM}(\vec{u},\vec{v}) = c_1 \frac{d_{STEM}(\vec{u},\vec{v})}{d_{STEM}^{ref}(\vec{u})} + c_2 \frac{d_{EDX}(\vec{u},\vec{v})}{d_{EDX}^{ref}(\vec{u})} \quad \text{eq. 11}$$

Here, $c_{1,}c_2$ in the range $[0,1]$ with $c_1+c_2=1$ are weights that can be used to trade-off the importance of the STEM and EDX modalities for block similarity recognition against each other. Unless one wants to put emphasis on one of the two modalities, the values $c_1=c_2=0.5$ can be chosen. Only if this sum is small enough, i.e. if the neighborhoods of the two pixels are similar with respect to each modality, the new pixel is considered for averaging. This allows the algorithm to differentiate between atomic columns with the same STEM signal but different EELS/EDX spectra. Note $d_{STEM}(\vec{u},\vec{v})$ and $d_{EDX}(\vec{u},\vec{v})$ are symmetric but $d_{M3S-NLM}(\vec{u},\vec{v})$ is not. The end results of the M3S-NLM procedure results in a new less noisy image. This entire matching and averaging procedure is repeated using every pixel in the image, aggregating the resulting denoised blocks, thereby producing a high quality

reconstruction for each pixel at full scan resolution. The schematic description of the procedure is given Fig. 1. As mentioned before, we suggest to use the NLM principle to compute the final EDX/EELS estimate, which is defined again as follows:

$$I_{EDX}^{M3S-NLM}(\vec{u},c)=\frac{\sum_{\Delta\vec{w}\in M} I_{EDX}(\vec{u}+\Delta\vec{w},c)\,w(I,\vec{u},\vec{u}+\Delta\vec{w})}{\sum_{\Delta\vec{w}\in M} w(I,\vec{u},\vec{u}+\Delta\vec{w})} \quad \text{eq. 12}$$

with weights defined as in eq. 3, but with the distance $d$ replaced by $d_{M3S-NLM}$ from eq. 11. Note that the weight w in eq. 12 does not depend on $c$, but is the same for every channel, since the distance $d_{EDX}$ in eq. 7 only depends on the spatial location.

For the STEM modality, we used the BM3D method to compute the final estimate. It is similar in structure to the NLM algorithm, but it does not simply average similar structure windows. Instead, it filters them collaboratively in a suitably transformed domain, such as Fourier or Wavelet spaces. Since the details of BM3D algorithm are already given in previous studies, we only described the main steps and refer to Refs. 18, 20, 23 for more details and exact formulas. The main steps of BM3D are as follows. For each pixel in the image, the first 256 most similar structure windows are grouped and stacked into a 3d tensor. Then, a domain transform is applied (e.g. Fourier or Wavelet) and the corresponding coefficients are shrinked by a suitable operator (e.g. hard-thresholding or Wiener filter). It follows, the inverse of the chosen domain transform is applied. This results in a group of denoised structure windows, which are again aggregated to form the reconstruction. In each pass, this process is repeated for every pixel to form an entire filtered image. BM3D algorithm comprises two such passes, using hard-thresholding for coefficient shrinkage in the first pass and Wiener filtering in the second pass (with coefficients based on the initial estimate from the first pass).

It is important to point out that the weighted average is computed using the full spatial and spectral resolution images. The reduction in spectral resolution is only encoded into the EDX similarity weights to account for the reduced SNR, which is mitigated by the preserved spatial resolution within the STEM similarity weights. In practice, to calculate $I_{EDX}^{M3S-NLM}$, we sum the 256 most similar structure windows, ranked by their $d_{M3S-NLM}$ values, to limit the contribution of non-similar structure windows on the denoised images. For the STEM modality, this restriction is inherently given by the grouping performed as part of the BM3D procedure.

However, at least one of the similarity measures (e.g., STEM, EELS or EDX signals) should operate at full spatial resolution. After a set of associated pixels has been found, the averaging is performed at the full resolution for all scans, yielding a high-resolution reconstruction of the entire dataset. In these studies, the spectral resolution of the EDX dataset is reduced by a factor of eight for evaluating the similarity measure, the spectral resolution of the EELS dataset is kept unchanged as the signal-to-noise was large enough to define the similarity windows.

Let us point out that decreasing the spatial resolution of any of the similarity measures can only be done without affecting the spatial resolution of the denoised dataset, as long as there are no two neighborhoods with the following properties: 1) different full resolution representation in the modality with the reduced similarity measure, 2) same reduced resolution representation in the modality with the reduced similarity measure, 3) same full resolution representation in all other modalities. This is not the case for any of the results presented here, and we consider it unlikely in practical settings. A detailed description of the algorithm is given in the Ref. 24.

### 2.2 Experimental setup

In order to directly map the compositional variations at atomic scale, we prepared cross-sectional STEM specimens by ion beam milling. The $Al_{1-x}In_xN$ sample is highly sensitive to ion milling and therefore the final cleaning was performed with low energy $Ar^+$ ion bombardment (0.5 eV) at liquid $N_2$ temperature using a Fischione Nanomill system. Structural investigations of the $GaN/Al_{1-x}In_xN$ specimen were performed at 200 kV using a thermofisher scientific Titan scanning transmission electron microscope equipped with a spherical aberration corrector at the condenser plane, resulting in a sub-Angstrom probe diameter with a convergence semi-angle of 20 mrad and a four quadrant EDX detector[25]. The STEM image and EDX dataset sizes 744 x 561 and 744 x 561 x 2048 pixels respectively and the energy per channel of 10 eV was used. $LaMnO_3$ particles were dispersed on a lacey carbon grid and investigated using a double corrected thermofisher scientific Titan TEM, with a Cc corrector in the image plane, at 300 kV[26]. The EEL spectrum image analyzed below was acquired at a collection semi-angle of 40 mrad and a probe convergence semi-angle of 20 mrad. The subpixel scanning mode was used to record the EELS dataset, i.e. the STEM signal was continuously scanned inside a given EELS pixel, thus one pixel of the EELS dataset corresponds to 16x16 pixels in the STEM image. This subpixel scanning approach was used to increase the SNR of the EELS spectra without compromising the STEM resolution. The STEM image and EELS dataset sizes are 368x1696 and 23x102x2048 pixels respectively and an energy dispersion of 0.1 eV/channel was used.

### 2.3 Generation of Synthetic EDX datasets

In order to support our claim, we constructed an artificial dataset where the ground truth STEM image consists of a periodic pattern of exactly identical Gaussian functions. As the intensities of both the STEM and EDX signals scale with the beam intensity for given set-up and specimen configuration, the ground truth EDX map was defined by multiplying the ground truth STEM intensities by the EDX to STEM experimental intensity ratio and assigning the two different spectra to groups of neighboring atomic columns in an alternating fashion. The artificial noisy measurements for this dataset were created by applying Poisson noise to the EDX map. Accordingly, the individual noisy spectra contain 1-3 counts in a few channels while all other channels are equal to zero. For the STEM image, we applied mixed Poisson-Gaussian noise and chose the parameters such that the resulting image looks similar to the experimental one in Fig. 2.

### 2.4 EDX and STEM simulations

For a given probe position, the fraction of incident electrons causing ionization events that can subsequently produce the X-ray peak under consideration, accounting for the dynamical elastic and thermal scattering of the electron probe was calculated based on the quantum excitation of phonons (QEP) model[27] as implemented in the μSTEM simulation suite[28] and as discussed further in Ref. 5.

## 3. Results

### 3.1 Application to atomic resolution EDX dataset

Epitaxially grown $Al_{1-x}In_xN$ alloys on GaN substrates have many interesting properties for optoelectronic and electronic applications due to the tunable band gap ranging from 0.7 eV (InN) to 6.2 eV(AlN), obtained by controlling the In to Al ratio[29]. The quality and the properties of the $Al_{1-x}$ $In_xN$ materials have been shown to depend on strain and compositional

fluctuations[30, 31]. Even small In fluctuations affect the local band gap, the photoluminescence, and carrier scattering, and are responsible for changing the properties of quantum heterostructures or high electron mobility transistors[32, 33]. It is therefore of high importance to identify the composition fluctuations in $Al_{1-x}In_xN$. The analysis of atomic scale interdiffusion in $GaN/Al_{1-x}In_xN$ interfaces at the atomic scale has been reported using high resolution (HR-)STEM[34]. The HR-STEM images were used to determine the absolute concentration of the heavy species by comparison of experimental data and simulations, but were not sufficient to recover the possible concentration fluctuation of the light species at atomic resolution[34, 35]. On the other hand, windowless EDX detectors are able to record elemental characteristic X-ray lines at atomic resolution for low atomic number elements. Unfortunately to date, due to the extreme sensitivity of this material to the electron beam irradiation, only low spatially resolved EDX datasets have been reported[37, 38]. Therefore, this system is ideal to apply the M3S-NLM method.

EDX maps extracted from raw and M3S-NLM denoised datasets recorded near the $GaN/Al_{1-x}In_xN$ interface are depicted in Fig. 2. The improvement in the SNR after applying the M3S-NLM method is significant compared to the raw dataset. Figure 3a shows a representative raw EDX spectrum of an individual pixel (labeled 'Single Raw') that contains only two pixels with non-zero values. Thus, due to the poor SNR of the EDX dataset, a direct integration of the intensities at the N Kα, Ga Kα, Al Kα and In Lα lines does not allow one to distinguish the atomic columns in the EDX maps (Fig. 2a). Despite the poor SNR of each individual spectrum, the raw average spectra (labeled 'Ave. Raw' in Fig.3a) across the entire dataset contains the characteristic X-ray lines of interest, including the N Kα line, demonstrating that the spectroscopic information is indeed present in the raw EDX dataset. Figure 3a also shows single (labeled 'Single Denoised') and averaged (labeled 'Ave. Denoised') EDX spectra after applying the M3S-NLM method. The 'single denoised' spectrum contains X-ray peaks of the elements of interest which were not visible in the 'single raw' spectrum extracted at the same position. The two averaged spectra are identical, their difference labeled ' Ave. difference' in Fig. 3a, indicating information was neither lost nor added in the spectral domain. The elemental EDX maps extracted after applying the M3S-NLM method contain the chemical information at atomic scale for the N Kα, Ga Kα, Al Kα and In Lα X-ray lines and allow a detailed description of the diffusion occurring at the interface (Fig. 2b). It is important to note the atomic information was present in the raw EDX dataset as a sub-angstrom probe was used to collect the spectra but it was previously mask by the very poor SNR of the dataset. Here, we have been able to separate the signal from the noise, revealing the atomic resolution information contained in the EDX dataset by making use of the atomic resolution information clearly visible in the ADF STEM image which was acquired in parallel.

Figure 3b shows the integrated elemental profiles parallel to the $GaN/Al_{1-x}In_xN$ interface for the raw dataset while Fig. 3c illustrates the same for the M3S-NLM denoised dataset. The diffusion of heavy species (Ga and In) over five atomic layers is expected for the observed contrast variation across the interface in the STEM image. The diffusion of Al inside the GaN layer is not visible in the STEM image but can be measured in the EDX Al maps. The diffusion length of the Ga, Al and In have been estimated using the Fick's law in one dimension (eq. 13) fitted on the NLM denoised dataset (the use of the denoised data allows one to improve the fit quality, i.e., smaller standard deviation as compared to fits performed on the raw dataset as the spread of the experimental values is reduced by the denoising procedure) and found to be 0.89, 0.71 and 0.99 nm respectively (Figs. 4). The Fick's law in one dimension was fitted using the optimize.curve_fit function of the scipy Python-based package according to the following formula:

$$n(x) = n_0\left(erfc\left(\frac{x - x_0}{b}\right)\right) + c \quad \text{eq. 13}$$

with $b = 2\sqrt{Dt}$ , D being the diffusion constant, t the time at the considered temperature to which the diffusion happens, $n_0$ the number of EDX counts at the interface, $x_0$ the position of the interface and c the concentration at the interface.

The M3S-NLM method also allows one to extract information at the N Kα edge that does not contribute to the STEM signal in a discernible way. The integrated N Kα signal (Fig. 2a) shows a noisy contrast as well as its intensity integrated parallel to the interface in Fig. 3b. On the denoised dataset, an increase of the N concentration is observed at the GaN/$Al_{1-x}In_xN$ interface (Fig. 3c). The integration of the N Kα signal parallel to the interface clearly shows the two first atomic columns of the GaN layer containing an excess of N of approximately 2 % as well as the three neighboring atomic columns in the $Al_{1-x}In_xN$ layer, albeit in a lesser extent. To confirm the fluctuations of the N Kα intensities are not due to the denoising procedure, we calculated the mean values of the X-ray counts at the N Kα peak (in the energy range 0.34-0.475 keV on the raw dataset) in three distinct areas delimited by the vertical dotted lines Figs. 3b,c. This gives the values of 0.134, 0.141 and 0.128 X-rays counts per pixel for the GaN, interface and $Al_{1-x}In_xN$ areas respectively.

The investigation of atomic diffusion at interfaces with TEM techniques is not directly interpretable since the broadening of the beam inside the specimen could result in apparent interdiffusion. To estimate this effect, we performed multislice simulations of an atomically flat interface and calculated the STEM and EDX integrated images. The results are presented in Fig. 5. Both, Ga K and In L maps show a few percent contribution across the interface, where these elements are not present, that is not likely to be visible experimentally. Therefore significant experimental signal beyond this is an indication of diffusion of Ga, In and Al across the interface. We have simulated the N contributions from the GaN layer (N1 K) and from the $Al_{1-x}In_xN$ layer (N2 K), as well as the total N contribution. Some "wrap around" effects can be seen for the N maps. It appears the N signal is present in both the N and the column with metal atoms, the brightest peaks being at the location of the N columns. The contribution at the columns with metal atoms is due to thermally scattered electrons when the probe is on the "adjacent" column of heavy atoms as demonstrated by the simulations that separate the elastic and thermal diffuse scattering (TDS) contributions to the total EDX signal, Fig. 6. Also of note is the visible difference in N intensities (N1 K and N2 K) in both materials, despite each material containing the same concentration of N atoms. In both the simulations and the denoised experimental N concentration profile (Fig. 3c), the averaged N signal intensity is larger in the GaN layer compared to the $Al_{1-x}In_xN$ layer, despite the same N atomic density for each of these stoichiometric compounds. Therefore, this effect should not be associated with a composition variation across the specimen but is due to a difference in the electron beam-matter interactions as a function of the local chemistry of the material. In the simulations, the value of x in the $Al_{1-x}In_xN$ layer is 0.2 resulting in an average atomic number at the metal atom site of 20.2 (average of In 49 and Al 13), which is significantly lower than the atomic number of Ga, 31. Heavier atoms are known to have a stronger channeling effect on the electron beam therefore this could may in part explain the increased intensity seen in the GaN layer.[39]

**3.2 Application to high resolution EELS dataset**

Perovskite oxide nanostructures are under intensive investigation using TEM techniques due to their wide range of possible applications and, in particular, due to their exceptional suitability for oxygen electrocatalysis that takes place on their surface or subsurface regions[40]. The knowledge of the surface polarity and oxidation state of the transition metals present in the perovskite structure is fundamental to a better understanding of the mechanism by which the

electrocatalysis proceeds. In the case of $LaMnO_3$, it was theoretically predicted that a change of the coordination number can affect the activity towards the oxygen reduction reaction (ORR). In addition to elemental mapping, EELS is able to deliver chemical information, such as oxidation state or local atomic coordination at atomic scale[2]. Density functional theory calculations for the Mn $L_3$ edge have shown that coordination and oxidation state changes at the surface result in a unique surface reconstruction that favours ORR[41, 42].

We used EELS, simultaneously with STEM, to record an atomically resolved spectrum image of $LaMnO_3$ particles probing the O K and Mn L edges at atomic resolution. The $L_3$ Mn edge has been shown to shift approximately 1.5 eV, to lower energies, when the oxidation state of the Mn changes from $Mn^{3+}$ (as is in the bulk state) to $Mn^{2+}$ [43, 44]. The experimental $L_3$ Mn edge has thus been integrated around (635 – 642 eV) and (642 – 648 eV) which correspond to the energy ranges of each oxidation state. The EELS spectra are rather noisy (typical individual spectra are shown in Fig. 8) and do not allow one to obtain atomic resolution integrated EELS intensity maps around the energy range of interest, as shown Fig. 7a. We applied the NLM method to the EELS dataset and integrated the energy ranges around the O K and Mn L edges. The results are given in Fig. 7c, a clear difference is observed between the surface and sub-surface region compare to the bulk region. This result indicates the presence of a different oxidation or atomic coordination between the two regions. One difficulty in extracting integrated maps from EELS datasets is the removal of the background contribution in front of the edges of interest before integrating over a given energy range. We noticed that the background subtraction gives better results after applying PCA to the raw dataset (Fig. 7b). We then used the same approach of applying PCA to the NLM denoised data. The integrated maps are presented in Figs. 7d. The atomic resolution is recovered for the three O maps while the resolution of the Mn maps is improved.

The O pre-peak in the 528 – 532 eV range is present in the bulk part and vanishes toward the surface as shown by the two extracted NLM denoised spectra Fig. 8. This pre-peak is linked to electron exchange between the Mn and O and it is a signature of the hybridization of the Mn 3d with the O 2p energy levels. Similarly, the Mn $L_3$ edge is shifted toward lower energy for the $Mn^{2+}$ oxidation state compared to the $Mn^{3+}$ and $Mn^{4+}$ states[43, 44].

**4. Discussion**

The previous two examples have shown the success of applying the (M3S-)NLM method to recover the atomically resolved EDX or EELS datasets for which the resolution of the reconstructed maps is limited by their poor SNR. The spatial averaging of the low SNR spectroscopic datasets is based on the similarities in the STEM and spectroscopic dataset modalities. The information of the HAADF-STEM image is sensitive to high atomic numbers while the EELS datasets are more sensitive to low atomic number elements and the EDX datasets contain information for both low and high atomic numbers. A naive picture would be to think that the local averaging of low atomic number elements cannot be done based on the STEM modality, but it has to be kept in mind that both modalities should match (through their respective structure windows distances) in order to be locally averaged. The success of this double block matching approach is demonstrated for the N map where the N atomic positions are found to be shifted compared to the metallic atomic columns and for the O K edge where fine structure information is recovered at atomic scale.

Another point to address is what minimum SNR value is required in order to extract meaningful information using the M3S-NLM method. We focus on the EDX spectroscopic dataset as their SNR is orders of magnitude lower compared to their EELS counterpart. Synthetic EDX datasets are generated based on the dataset of Fig. 2, with distinct regions of GaN and $Al_{1-x}In_xN$ (see method section 2.3). The average number of counts per energy channel in the EDX dataset

modalities is varied to $7.3x10^{-5}$, $1.4x10^{-4}$, $7.0x10^{-4}$ (which corresponds to the SNR of the EDX dataset of Fig. 2), and $3.6x10^{-3}$ counts per energy channel, which reflects an increasing electron dose to which the specimen is exposed. The cross-sections of the events leadings to the STEM and EDX signals are considered to be independent of the electron dose. Thus, the ratio of the average number of counts of the STEM to the EDX signals is kept constant. Figure 9 show the elemental maps extracted from the raw datasets (left columns) and after applying the M3S-NLM method (right columns). For the lowest number of counts, the local information is not recovered for both modalities. For the synthetic dataset (with $1.4x10^{-4}$ counts/energy channel, only the atomic resolution of the STEM image is recovered. For a low SNR dataset, if the similarity measurement is not able to distinguish the given regions, then information in the spectral domain may originate from another region. To check for the presence of this artifact, the average of each of the regions for the raw datasets and after M3S-NLM denoising are compared. An artifact free denoising dataset should lead to similar averaged spectra. For the EDX dataset with signal of $7.0x10^{-4}$ counts per energy channel and larger, the different elements are attributed to their respective regions at the atomic level.

**5. Conclusion**

We have successfully applied the (M3S-)NLM technique to EDX and EELS datasets to retrieve maximum information at atomic resolution for specimens that degrade under prolonged electron beam exposure or when the SNR is limited by the exposure time (that relates to the specimen drift). We confirmed the robustness of this methodology by applying it to low SNR EDX and low signal-to-background EELS. We show that the algorithm leads to important practical conclusions that were out-of-reach when using classical denoising techniques (MVA-based techniques). The method presented in this article leads to important practical implications for retrieval of maximum information in EDX and EELS datasets. We discuss the possible artifacts and demonstrate the method is robust against low correlation between the information contained in the STEM and the spectral modalities. The use of the M3S-NLM method is also expected to open-up the possibility of spectroscopic analysis of extremely beam sensitive biological specimens as no assumption on the need of atomically resolved STEM image is needed.

**Acknowledgments**

This work was supported by Ministry of Education Academic Research Fund Tier 1 grants RG101/17 (Nanyang Technological University). This research was supported under the Discovery Projects funding scheme (Project DP140102538) of the Australian Research Council and by the Alexander von Humboldt Foundation. The authors at AICES RWTH Aachen University were funded in part by the Excellence Initiative of the German Federal and State Governments through grant GSC 111. Katherine MacArthur acknowledges the financial support of the Helmholtz Association in the form of a post doctoral scholarship. We thank J.-F. Carlin, R. Butté, and N. Grandjean from EPFL for kindly providing the AlInN/GaN sample and Prof. Kelsey A. Stoerzinger from Oregon State University and Prof. Yang Shao-Horn from MIT for kindly providing the $LaMnO_3$ specimen.

Figures captions:

Figure 1: Illustration of the M3S-NLM denoising procedure. The entire procedure is repeated while comparing the reference block with all possible blocks in the dataset and by shifting the reference block pixel by pixel and line by line across the entire scan area. (a, b) Schematic to calculate the STEM structure window distances, $d_{STEM}$ , and (c, d) to

calculate EDX structure window distances, $d_{EDX}$ . (e) The EDX and STEM structure window distances are merged in a single $d_{M3S-NLM}$ distance that accounts for the similarities for both the EDX and EELS dataset as described in eq. 11. (f, g) Schematic of the reconstructed datasets, accounting for similarities between the imaging and spectroscopic datasets via $d_{M3S-NLM}$ , obtained from eqs. 12 and 13. The procedure is similar for EELS dataset where EDX is replaced by EELS.

Figure 2: HAADF-STEM images and EDX maps of the N Kα, Ga Kα, Al Kα and In Lα edges simultaneously recorded at the GaN/$Al_{1-x}In_xN$ interface for the (a) raw and (b) M3S-NLM datasets. The color-scale indicates the atomic percent concentration of the denoised maps. The field of view of each image is 9 nm by 6.8 nm.

Figure 3: (a) Averaged and single spectra extracted from the raw and denoised dataset. From the top to the bottom: 'Ave. Denoised' and 'Ave. Raw' are averaged over the entire EDX dataset for the NLM denoised and raw dataset, 'Ave. difference' is the difference between the two averaged spectra, and 'Single Denoised' and 'Single Raw' are single spectra extracted from the middle pixel of the dataset. 'Single Raw' contains only two pixels with non-zero values and is representative of the dataset. 'Single Denoised' is extracted from the same position as 'Single Raw' after applying the M3S-NLM denoising procedure. Note the difference in vertical scales between the spectra. (b-c) Integrated EDX intensity profiles parallel to the GaN/$Al_{1-x}In_xN$ interface extracted from the (b) raw and (c) M3S-NLM denoised EDX maps (Figs. 2).

Figure 4: EDX profiles after applying the M3S-NLM denoising technique fitted using Fick's law shown in eq. 13. The fitted coefficients are shown next to the profiles of the (a) Ga Kα, (b) Al Kα and (c) In Lα and described in the main text.

Figure 5: Simulated HAADF and EDX images of the GaN/$Al_{1-x}In_xN$ interface. The N1 K and N2 K maps are the integrated signals due to the N atomic columns of the GaN and $Al_{1-x}In_xN$ layers, respectively. The profiles show the integrated EDX and HAADF calculated signals integrated along the interface. To account for the finite source size, a blurring of 0.7 Å was applied. The K-series and L-series account for the all the transitions, i.e., from the 1s and 2s/2p respectively which differs from the experimental data which account only for the Kα and Lα transitions. The field of view of each image is 3.3 nm by 3.9 nm.

Figure 6: Simulation of the (a) elastic, (b) thermal diffuse scattering (TDS) contributions and of the (c) total X-ray signal, i.e., the sum of (a)+(b), to the N maps at the GaN/$Al_{1-x}In_xN$ interface. The field of view of each image is 3.9 nm by 3.3 nm.

Figure 7: (a) Integrated EELS-STEM dataset of $LaMnO_3$ at the O K and Mn L edges. The different peaks have been integrated to track the EEL fine structure changes across the specimen. The O K edge is integrated over the 528 – 532 eV, 532 – 536 eV and 536 – 543 eV ranges. The Mn L edge is integrated over the 635 – 642 eV and 642 – 648 eV ranges. The maps are integrated from (a) the raw dataset, (b) after PCA denoising, (c) after NLM denoisng and (c) after NLM denoising and followed by PCA. The field of view of each image is 1.8 nm by 8.5 nm.

Figure 8: Spectra extracted from the positions A and B marked in Fig. 7. The bottom two spectra are extracted from the subsurface region (marked B in Fig. 7) and the top two spectra are extracted from the bulk part (marked A in (Fig. 7)).

Figure 9: Synthetic EDX dataset used to test the robustness of the M3S-NLM denoising method against low SNR and against the none correlation between the HAADF-STEM and EDX-STEM modalities for the N Kα and Al Kα edges. (a, b) corresponds to an averaged signal of $7.3x10^{-5}$ count/pixel, (c, d) of $1.4x10^{-4}$ counts/pixel, (e,f) of $7x10^{-4}$ counts/pixel, i.e., the signal to noise of the EDX-STEM dataset of the Fig. 2 and (g, h) of $3.6x10^{-3}$ counts/pixel. (a, c, e, g) rows are

the reconstructed maps from the raw synthetic dataset while (b, d, f, h) are reconstructed from the M3S-NLM denoised dataset. From the top to the bottom, the rows correspond to the HAADF, N Kα, Ga Kα, Al Kα and In Lα reconstructed maps. The field of view of each image is 3.1 nm by 3.1 nm.

**References**


1. Bosman, M. *et al. Phys. Rev. Lett.* **99**, 086102 (2007).
2. Ramasse, Q. M. *et al. Nano. Lett.* **13**, 4989-4995 (2013).
3. Boothroyd, C. B. *et al. Ultramicroscopy* **145**, 66-73 (2014).
4. D'Alfonso, A. J., Freitag, B. Klenov, D. & Allen, L.J. *Phys. Rev. B* **81**, 100101(R) (2010).
5. Chen, Z., Weyland, M., Sang, X., Xu, W., Dycus, J.H., Lebeau, J.M., D'Alfonso, A.J., Allen, L.J. and Findlay, S.D., *Ultramicroscopy*, **168**, 7–16 (2016).
6. Trebbia, P. & Bonnet, N. *Ultramicroscopy* **34,** 165–178 (1990).
7. Trebbia, P. & Mory, C. *Ultramicroscopy* **34,** 179–203 (1990).
8. Rossouw, D. et al. *Nano Lett.* **15**, 2716-2720 (2015).
9. Lee, D. D. & Seung, H. S. *Nature* **401,** 788–791 (1999).
10. de la Peña, F. et al. *Ultramicroscopy* **111,** 169–176 (2011).11. Nicoletti, O. et al. *Nature* **502,** 80–84 (2013).
12. Nascimento, J. M. P. & Bioucas Dias, *IEEE Trans. Geosci. Remote Sens.* **43,** 898–910 (2005).
13. Dobigeon, N. & Brun, N. *Ultramicroscopy* **120,** 25–34 (2012).
14. Duchamp, M. et al. *Appl. Phys. Lett.* **102,** 133902 (2013).
15. Lichtert, S. & Verbeeck, *Ultramicroscopy* **125,** 35–42 (2013).
16. Buades, A., Coll, B. & Morel, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* **2,** 60–65 (2005).
17. Buades, A., Coll, B. & Morel, J.-M., SIAM Review, **52,** 113-147 (2010).
18. Dabov, K., Foi, A., Katkovnik, V. & Egiazarian, K. *IEEE Trans. Image Process.* **16,** 2080–2095 (2007).
19. Mairal, J., Bach, F., Ponce, J., Sapiro, G. & Zisserman, A. *IEEE 12th International Conference on Computer Vision* 2272–2279 (2009).
20. Mevenkamp, N. et al. *Adv. Struct. Chem. Imaging* **1,** 3 (2015).
21. Binev, P. et al. *Modeling Nanoscale Imaging in Electron Microscopy* (eds. Vogt, T., Dahmen, W. & Binev, P.) 127–145 (Springer US, 2012).
22. Jeong, J. S. & Mkhoyan, K. A. *Microscopy and Microanalysis* **22**, 536-543 (2016).
23. Dabov K. *et al.* Proceedings Volume 6064, Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning; 606414 (2006).
24. Mevenkamp, N., Phd dissertation: Non-local denoising and unsupervised quantitative analysis in scanning transmission electron microscopy, Aachen, Germany 2017, DOI: 10.18154/RWTH-2017-04917
25. Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons (ER-C) *et al.* (2016). FEI Titan G2 80-200 CREWLEY. Journal of large-scale research facilities, **2**, A43.
26. Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons (ER-C) *et al.* (2015). FEI Titan G3 50-300 PICO. Journal of large-scale research facilities, **1**, A34.
27. Forbes, B. D., Martin, A.V., Findlay, S.D., D'Alfonso, A.J. & Allen, L.J. *Phys. Rev. B* **82**, 104103 (2010).
28. Allen, L.J., D'Alfonso, A. J. & Findlay S.D. *Ultramicroscopy* **151**, 11–22 (2015).
29. Bvurgaftman I. et al. *J. Appl. Phys.* **94,** 3675–3696 (2003).
30. Cosenday G. et al. *Appl. Phys. Lett.* **98,** 181111 (2011).
31. Portz, V. et al. *Appl. Phys. Lett.* **109,** 132102 (2016).
32. Butté, R. et al. *J. Phys. Appl. Phys.* **40,** 6328 (2007).
33. Liuolia V. et al. *Appl. Phys. Lett.* **100,** 242104 (2012).
34. Lin, P.-Y., Chen, J.-Y., Shih, Y.-S. & Chang, L. *Nanoscale Res. Lett.* **9,** 628 (2014).
35. Rosenauer, A. et al. *Ultramicroscopy* **109,** 1171–1182 (2009).
36. Rosenauer, A. et al. *Ultramicroscopy* **111,** 1316–1327 (2011).
37. Chen, W.-C., Wu, Y.-H., Peng, C.-Y., Hsiao, C.-N. & Chang, L. *Nanoscale Res. Lett.* **9,** 204 (2014).
38. Wong, M. H., *et al. Appl. Phys. Lett.* **100,** 72107 (2012).
39. Rossouw, C.J., *et al.* Ultramicroscopy **96**, 299 (2003).
40. Grimaud, A. *et al. Nat. Energy* **2,** 16189 (2016).
41. Ahmad, E. A. *et al. Physical Review Letters* **108,** 259701 (2012).
42. Ignatans, R. *et al, J. Phys. Chem. C* (2019) *J. Phys. Chem.* **123**, 11621 (2019).
43. Kurata H. & Colliex C. *Phys. Rev. B* **48**, 2102 (1993).
44. Garvie L.A.J. & Craven A.J. *Phys. Chem. Minerals* **21**, 191-206 (1994).