Ag Surface Segregation in Sub-10 nm Bimetallic AuAg Nanoparticles Quantified by STEM-EDS and Machine Learning: Implications for Fine-Tuning Physico-Chemical Properties for Plasmonics and Catalysis Applications

Murilo Moreira, Matthias Hillenkamp, Varlei Rodrigues, Daniel Ugarte

Ag Surface Segregation in Sub-10 nm Bimetallic AuAg Nanoparticles Quantified by STEM-EDS and Machine Learning: Implications for Fine-Tuning Physico-Chemical Properties for Plasmonics and Catalysis Applications

Murilo Moreira,[†,‡] Matthias Hillenkamp,[†,‡] Varlei Rodrigues,[*,†] and Daniel Ugarte[*,†]

†*Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin, Campinas, SP 13083-859, Brazil*

‡*Institute of Light and Matter, University Lyon, Université Claude Bernard Lyon 1, CNRS, UMR5306, Villeurbanne F-69622, France*

E-mail: varlei@ifi.unicamp.br; dmugarte@ifi.unicamp.br

Abstract

Mono- and multimetallic nanoparticles are extensively studied in various fields due to their tunable physicochemical properties and potential for replacing expensive metals with more abundant and affordable ones. The chemical structure, i.e. the spatial distribution of elements inside nanoparticles, plays a crucial role in defining their properties, particularly in catalytic processes. However, accurately determining the spatial chemical distribution within sub-10 nm bimetallic nanoparticles remains a challenge. In this

study, we have used scanning transmission electron microscopy associated with energy-dispersive spectroscopy to acquire hyperspectral images of gold-silver alloy nanoparticles in the 3-10 nm size range. We have quantified the chemical composition as a function of radial position; Ag enrichment towards the nanoparticle surface is robustly confirmed by statistical analysis, error bars, and non-overlapping 3-sigma uncertainty intervals at the nanoparticle center and surface. Two complementary machine learning analyses (Principal Component Analysis, PCA and, Non-negative Matrix Factorization, NMF) reveal that our experiments contain latent information on subtle composition variations inside the particles. The proposed data analysis procedures have also been validated by simulated datasets. These findings pave the way for more precise structural and chemical investigations of alloys on the nanoscale.

Keywords

bi-metallic nanoparticles, nanoalloys, chemical composition, scanning transmission electron microscopy, energy-dispersive spectroscopy, machine learning

1 Introduction

Mono- and Multimetallic nanoparticles play an important role in fundamental and applied research. The possibility of combining several metallic elements permits the fine-tuning of physicochemical properties and, at the same time, the replacement of sparse and expensive metals (Pt, Pd, Rh, Au...) with more abundant and cheaper ones (Cu, Ag...).[1–3] This has been extensively studied in plasmonics, where nanoalloys allow controlling the Surface Plasmon Resonance over a wide spectral range[4,5] or in catalysis, where alloying can significantly increase selectivity, sensitivity and activity of the catalysts.[6–8]

One additional parameter that has to be taken into account to optimize nanoparticle properties is the chemical structure, i.e. the spatial distribution of elements inside each

nanoparticle. Their physicochemical properties depend, sometimes strongly, on whether the metals mix randomly or segregate into abrupt/gradual core-shell or Janus-type structures, and if chemically ordered structures are adopted.[1–3] This is particularly important in catalytic behavior, where adsorption and bond activation depend not only on surface and subsurface elements but also on those further away from the surface.[9] It is thus of high importance to be able to experimentally determine, on a quantitative level, the chemical structure inside bimetallic nanoparticles (BNPs). This has already been achieved for comparably large particles,[10,11] but not yet in the size range below 10 nm. Such small BNPs are of particular interest for their (quantum) size effects and increased catalytic activity, but the low signal level represents a serious challenge to assess quantitative chemical composition with high spatial resolution.[12] Detailed analyses concerning thermodynamic properties such as miscibility and diffusion or catalytic processes require quantitative and spatially resolved information.[13,14]

Scanning transmission electron microscopes (STEMs) represent one of the most powerful and popular tools to gather information on the physical and chemical properties of individual nanosystems on the nanometer and atomic level.[15] Concerning chemical analysis, X-ray energy-dispersive spectroscopy (EDS) is an easy-to-use method to reveal chemical elements present in the sample.[16] The association of STEM nm-wide electron probe and EDS allows the recording of so-called hyperspectral images,[17] where an entire analytical X-ray spectrum can be acquired at each image pixel. The HSI approach generates huge 3D or 4D datasets[18] that are a very rich source of information. However, it is often extremely difficult to disentangle the spatial and spectral contributions of the different physical and chemical factors.

Retrieving a maximum of information from such an entangled data set requires the use of advanced statistical algorithms, such as machine learning (ML) tools, which exploit redundancies and can reveal non-apparent correlations;[19] in this way, they can extract information that is inaccessible by traditional methods.[20]

Here we will tackle the challenging case of the chemical characterization of sub-10 nm nanoparticles made from a bimetallic alloy. The full understanding of the system requires the determination of chemical composition as a function of NP diameter as a) the increased surface-to-volume ratio may influence the equilibrium structures of the particles leading to size-dependent chemical properties and b) the predominant role of the surface in nanoparticles might induce complex phenomena such as chemical gradients which are not possible in macroscopic equilibrium phase diagrams of an alloy. This last issue can be induced for ground state structures either by surface energy minimization or by enhanced chemical reactivity with the environment. Meta-stable BNPs moreover display various chemical structures, strongly dependent on their formation and/or reaction history.

One of the most widely studied bimetallic systems is AuAg, due to its suitability for fundamental studies with the high resistivity against oxidation and good signal level in imaging techniques.[11,13,14,21] At the same time, its physico-chemical properties make it particularly interesting for plasmonics, bio-applications[22] or catalysis.[23–25]

We have studied chemical composition gradients inside individual AuAg alloy nanoparticles in the 3-10 nm diameter range using quantitative chemical composition measurement based on STEM-EDS. The particles have been produced by physical methods (gas aggregation, figure 1a))[26,27] and they are surfactant-free. The quantitative chemical composition of each individual BNP was obtained by applying the Cliff-Lorimer[28] approach and standard error propagation[29] on EDS counts obtained by azimuthal integration. Our results reveal an Ag surface enrichment as a result of measured radial gradients. The statistical validity of the detected composition modifications has been confirmed using two machine learning tools (Principal Component Analysis - PCA[30,31] and Non-negative Matrix Factorization – NMF[32,33]). We emphasize that, in this work, ML methods have been exploited to reveal that experiments carry latent information on chemical composition variation inside BNPs, which confirms the statistical significance and sensitivity of the experimental measurements and described data analysis.

2 Results and Discussions

Figure 1b) shows EDS data and Figure 1c) displays typical high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) results from bimetallic NPs generated in a gas aggregation source (a typical size distribution can be found in previous work on this system[34]). The atomic structure of AuAg NPs obtained under the very same conditions is already described elsewhere[35] and here we will focus only on their chemical composition gradient studied by STEM EDS. The HAADF-STEM contrast is related to the atomic number (Z) and thickness of the material traversed by the electron beam. This imaging mode is extremely useful because Au and Ag have significant atomic number differences ($Z_{Au} = 79$ and $Z_{Ag} = 47$), indicating that brighter areas should show higher Au content, while slightly darker areas suggest Ag richer regions, see figure 1c). The NP images in Figure 1 suggest that the alloy particles are not chemically homogeneous and should be formed by an Ag-rich surface due to a core-shell structure or a radial chemical gradient.

The chemical maps for Au and Ag derived from the HSIs are presented in figure 2; it is easily observable by the eye that most of the particles show a diameter that is slightly bigger in the Ag chemical map than in the Au one (see EDS line profiles in Supporting Figure S1). In general terms, the EDS counting level is rather low (few EDS counts per pixel), preventing any precise and accurate quantification by conventional EDS-STEM approaches.

At this point, we must remember that the full understanding of nanomaterials and the correct modeling of their chemical and physical properties requires quantitative and reproducible interpretation of imaging, diffraction, or spectroscopic studies,[20,36,37] following rigorous error analysis to generate meaningful and reproducible measurements and interpretations.[38–40] With this idea in mind, we have carefully designed an experimental study to get a very high counting level (several thousand counts) from individual sub-10 nm NPs. We will dedicate the next paragraph to evaluating the quantitative chemical composition of individual NPs and of eventual segregation or chemical gradients inside them. The EDS counting level of our experiments attained $I_{Au} \sim 3000$ for a ~ 6 nm diameter NP, which

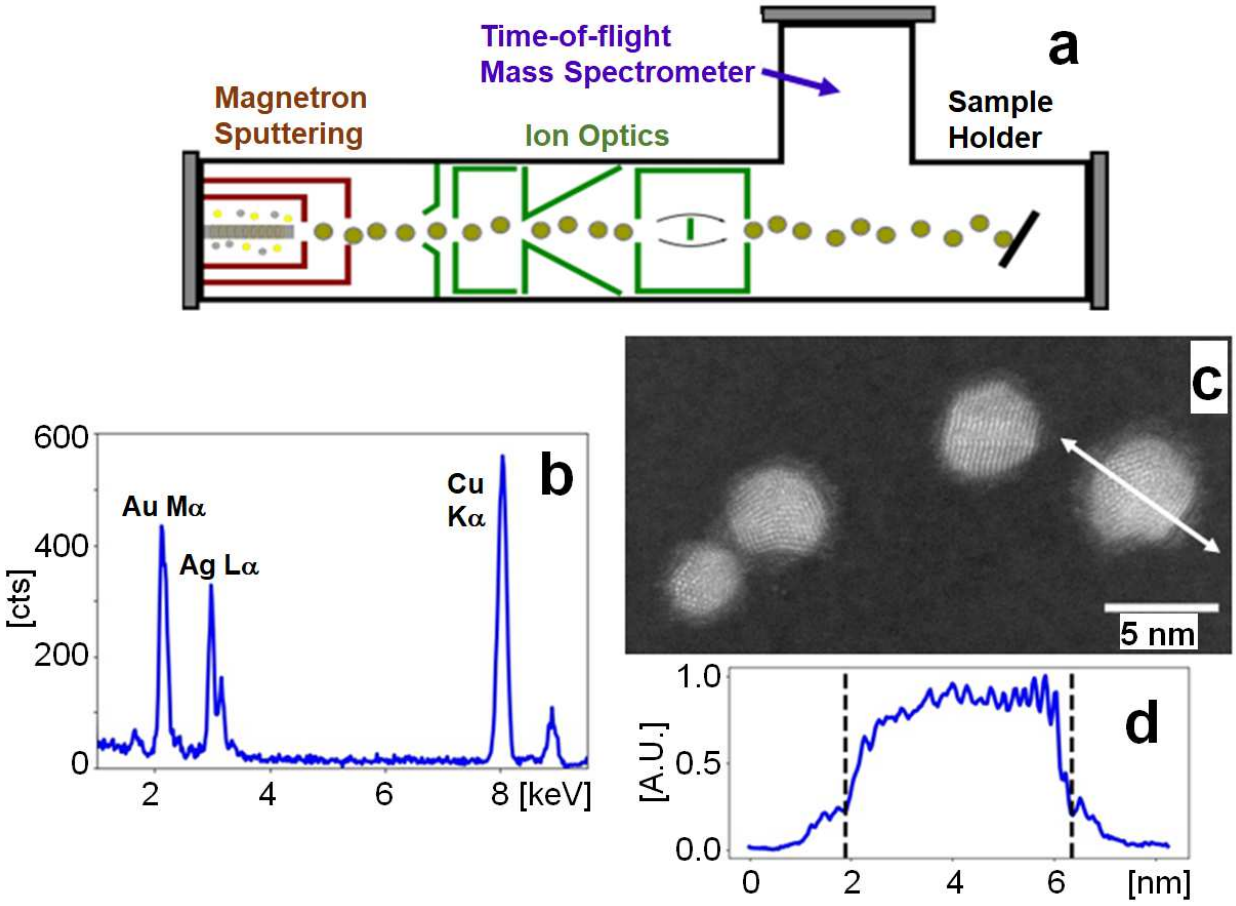


Figure 1: a) Schematic illustration of the gas aggregation cluster source, see Materials and Methods for details. Typical STEM experimental results obtained for the AuAg BNPs: b) EDS spectrum from a ∼5 nm NP generated by the addition of all the spectra in the NP region. Note that the Au peak maximum exceeds 400 counts (integrated total intensity $I_{\mathrm{Au,\ NP}} \sim$ 2000 cts). c) HAADF image of particles around 5 nm in diameter, showing a brighter core and lower intensity at the outer shell, suggesting an Ag-rich surface. d) Intensity profile obtained along the line in the right BNP in c). The dashed lines help identify the brighter-darker contrast areas in the BNP image.

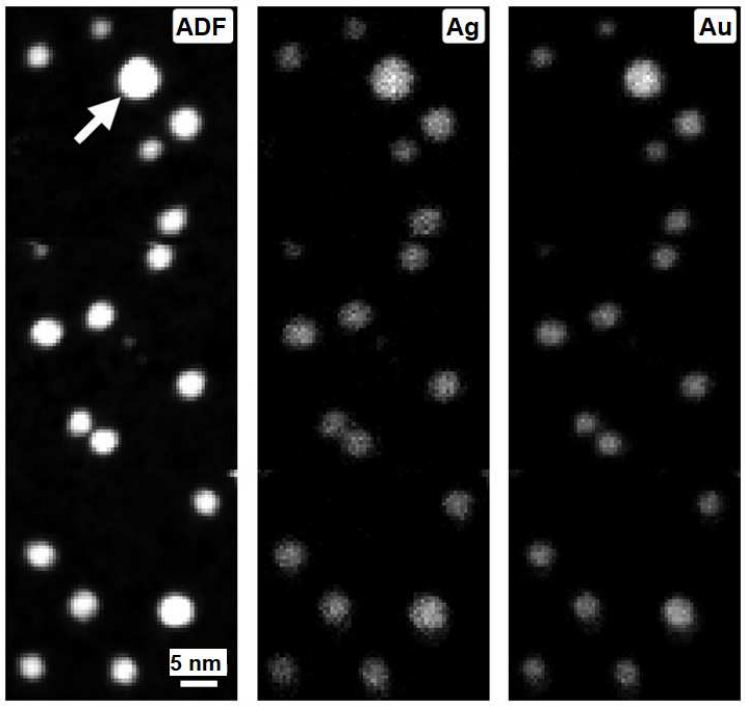


Figure 2: Comparison of ADF-STEM and elemental Ag and Au maps of different nanoparticles (maps include non-zero pixels with net counts > 5). Note that most of the NPs display a larger size in the Ag map than in the Au one, suggesting a tendency to surface Ag enrichment (see line profiles in Supporting Figure S1).

yields a composition error bar around $\sigma_{C_{\mathrm{Au}}} \sim 0.01$ in atom fraction. Here probe size and pixel size have been optimized for the aim of the paper, obtaining chemical concentrations with small uncertainty intervals in few-nanometer-size nanoparticles. Experiments must fulfill a compromise of beam size, beam current, sample drift, and dwell time to get a high x-ray counting level and avoid sample damage. The ensemble of these limiting conditions has led us to experiments with a 0.59 nm pixel size, so no atomic details or atom column EDS is contained in the data set.

The chemical maps shown in Figure 2 have been numerically processed to automatically identify NPs and quantify their EDS signal. First, we have determined the NP center by using a cross-correlation between each ADF image (or EDS map) and an ideal circular NP contrast (top-hat type, 1 inside the disk, 0 outside it, see NP identification results in figure S2). The local maxima of the correlation function can be automatically indexed in the image, and they represent the NP centers. This procedure is similar to the ones utilized in the detection and evaluation of diffraction spots in electron diffraction studies.[15,41] It is important to mention that identical results are obtained on particle centers when we use the total x-ray count image (simple addition of Au and Ag elemental maps).

The next processing step defined a threshold value of the ADF intensity to determine the particle radius, assuming a circular shape of the spheroidal BNPs in the 2D projection. The chemical composition of individual BNPs has been evaluated by adding the EDS counts from all pixels inside the determined circular region around the BNP center in order to perform a full chemical composition analysis with error bars. The HSI used for the present study (Figure 2) contains 19 particles and the corresponding relative chemical composition as a function of size is shown in Figure S3, with a mean composition of $C_{\mathrm{Ag}} = 0.513 \pm 0.003$. Note that the fact that the smallest BNP displays a significantly higher proportion of Ag is consistent with our previous work[34] and can be attributed to the BNP formation in the source. In this study, however, we will not address this point in detail and we will focus on the distribution of chemical elements inside individual BNPs.

The challenging step of obtaining reliable measurements of possible composition variations inside the BNPs is motivated by ADF images and EDS maps Figures 1, 2 and intensity line profiles in Supporting Figure S1), which suggest an Ag surface enrichment. However, the analysis of intensity profiles of only a few pixels in width (as usually extracted from chemical maps) will hardly attain the minimal counting level for a reliable assessment of quantitative chemical composition variations.

To verify the radial symmetry of the elemental distribution in the NP (or, in other words, the spheroidal homogeneous morphology of BNPs), we have calculated the center of mass of Au and of Ag EDS intensity for each NP region. These centers of mass and the geometrical center derived previously by cross-correlation show an excellent agreement. They differ by less than a tenth of a pixel for the majority of NPs and less than a third of a pixel for the smaller ones. This analysis confirms unambiguously that the 2D projection of the BNPs shows circularly symmetrical Au and Ag EDS signal distributions for each particle and that structures such as Janus can be excluded.

Based on this radial symmetry, we can now perform an azimuthal integration in order to improve the signal-to-noise level. For this, we have integrated EDS counts for all pixels at the same distance from the BNP center in order to generate radial intensity profiles (the profiles have been generated as a histogram of EDS counts with histogram step 0.5 nm). Note that even though the projected metal thickness close to the BNP surface is small, the higher number of pixels at large radii compensates efficiently for this decrease in signal per pixel.

Figure 3a displays the result of applying this procedure to the biggest NP in the HSI ($\sim$7 nm in diameter, marked with an arrow in Figure 2, and indexed as NP#0 in Figure S2). We can clearly observe that the central region of the NP shows an Au intensity which is higher than the one for Ag. In contrast, the Ag intensity becomes slightly larger than Au close to the NP surface, confirming qualitatively a surface Ag segregation. We can then normalize the radial intensity by dividing it by the number of pixels at each radial value,

thus generating the average spectral counts for each radial distance, Figure 3b. The shape of the radial intensity profiles is consistent with a spherical shape for the analyzed NP. We note that the generated EDS counting level at all radial positions is actually very good for such a small BNP,[36,37] as counting values enable us to get an error bar for (C_{Ag} or C_{Au}) of approximately 0.01 in atomic fraction for each radial distance r.

It is essential to emphasize that we have attained an analytical quality of EDS that yields a measurable chemical composition difference between the center and surface of the NP. This conclusion is physically valid and robust because it has been achieved following strict measurement methods and error criteria ($C_{\mathrm{Ag,\ center}} = 0.46 \pm 0.01$, $C_{\mathrm{Ag,\ surface}} = 0.60 \pm 0.02$, for this 7 nm alloy NP; see Supporting Table S1 for the complete list of radial concentration values). In fact, the 3-sigma uncertainty intervals of both measurements do not overlap.

To provide a statistically valid analysis, we have measured the chemical composition profiles for all particles in the HSI. We have observed a similar tendency of Ag enrichment going from the NP center to the edge confirming surface segregation for the ensemble of analyzed NPs (see Figure S4). However, for smaller NPs, the uncertainty intervals between concentration at the center and surface may superpose due to the larger error bars related to their lower signal levels.

There are many examples in the literature where Ag enrichment at the surface of AuAg alloy NPs is reported. Whether this partial segregation is intrinsic to the alloy at the nanoscale or induced by chemical interface effects, notably oxidation, is still highly controversial.[13,42–44]

The quantitative assessment of composition variations in nanomaterials by STEM-EDS is quite a challenging task, not only due to low signal levels but also due to assumptions included in models (ex. neglecting absorption effects, etc.)[15,16] and possible sample modifications during electron beam irradiation.[45,46] The measurement of chemical gradients is associated with subtle intensity changes in pixel intensity inside the NP.

To verify the robustness of our results on sub-10 nm BNPs, we have applied unsupervised machine learning (ML) methods (PCA[30,31] and NMF[32,33]) to verify if the algorithms can de-

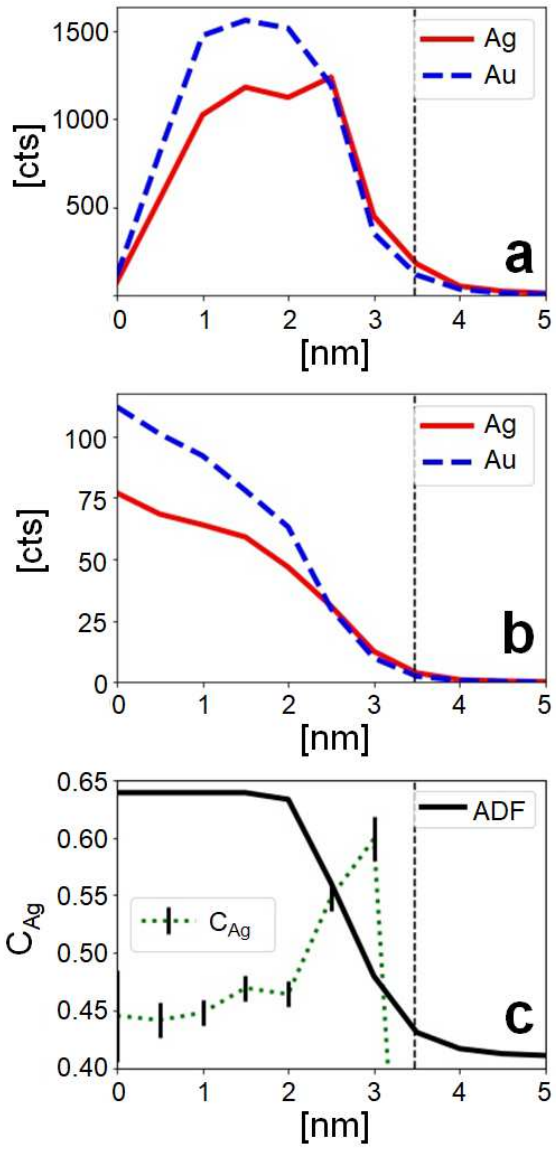


Figure 3: a) Azimuthal integration of EDS intensity for one individual nanoparticle (marked with an arrow in Figure 2). The plot shows the total elemental signal for each radial position (0.5 nm step size). b) Radial intensity profile normalized by the number of pixels contributing to each radius. c) Quantitative chemical composition analysis (including error bars) showing the Ag surface segregation; the azimuthally integrated radial ADF intensity profile is also displayed, saturated in the central region of the nanoparticle. The vertical dashed line indicates the estimated NP radius.

tect latent information on radial composition variations in our data. A positive answer to this question provides a qualitative, but nevertheless strong and direct statistical verification that obtained radial composition changes represent actual information contained in the EDS spectra. This work does not aim to compare quantification procedures between the conventional Cliff-Lorimer approach and Machine Learning tools used for denoising. It is very important to note that in denoised reconstruction, the spectra in the dataset do not represent independent measurements and are fully correlated measurements; this leads to a complex evaluation of error bars and accuracy.[36,37] There is no well-accepted procedure to calculate uncertainty after denoising, so we have decided to derive a quantitative chemical analysis using only the raw data without any denoising and applying well-established and widely recognized methods to evaluate error bars using partial derivatives.

A detailed study of PCA processing of EDS HSI of bimetallic AuAg particles has been recently reported by our group.[36,37] In PCA, the raw EDS data are decomposed into a number of orthogonal eigenvectors or components, referred to as loadings, that encode the data information. The correlation between the spectra associated with a certain pixel and a particular component shows up as a significant amplitude (score) for that component. For this rather simple case, it is possible to interpret the information carried by the relevant principal components.[36,37,47] A maximum variance, associated with the first loading (PCA#1), represents the background and spurious peaks (ex. copper grid, carbon substrate, etc.) in the EDS spectra. This component shows negligible intensity at the spectral Au or Ag peak positions and it is straightforward to conclude that it does not carry any information about the BNPs. The second loading (PCA#2) represents the mean EDS intensity correlated with the BNPs and it provides the average chemical composition of the particles. The third loading (PCA#3) yields an unphysical profile where Au and Ag EDS peaks show signals of different signs (Figure 4a). It thus describes composition variations within the analyzed region. This is expressed in Equations 1 and 2, where $s_{i,n}$ represent the score from component PCA#n at pixel I, and I_X^n represent the intensity of EDS peaks for element X at pixel i. In other

words, we must observe an anti-correlation between Au and Ag counts. This explains why PCA#3 displays a negative Ag peak and provides a tool to reproduce changes in the $I_{\mathrm{Ag}}/I_{\mathrm{Au}}$ intensity ratio and the chemical composition directly deduced from it.

$$I^{\mathrm{PCA}}_{\mathrm{Au},i} = s_{i,\mathrm{PCA\#2}}\ I^{\mathrm{PCA\#2}}_{\mathrm{Au}} + s_{i,\mathrm{PCA\#3}}\ |I^{\mathrm{PCA\#3}}_{\mathrm{Au}}| \tag{1}$$

$$I^{\mathrm{PCA}}_{\mathrm{Ag},i} = s_{i,\mathrm{PCA\#2}}\ I^{\mathrm{PCA\#2}}_{\mathrm{Ag}} - s_{i,\mathrm{PCA\#3}}\ |I^{\mathrm{PCA\#3}}_{\mathrm{Ag}}| \tag{2}$$

The efficient detection of the PCA#3 loading in our EDS STEM experiment is a very important fact because it is a direct confirmation that latent information on composition variation is present and detectable in our data, i.e. that our BNPs cannot be homogeneously alloyed. We must emphasize that PCA#3 detection is not attained very easily when studying NPs; this is strongly related to the high quality of our EDS measurements, mainly attributed to a very high counting level, much higher than frequently obtained for qualitative chemical maps. Higher order components than PCA#3 carry only noise, as already discussed elsewhere.[36,37]

The above discussion has addressed the information contained in PCA loadings through their spectral profile (ex. EDS peak occurrence and sign). It is important to emphasize that the resulting component scores are worthy of careful analysis and may contain a wealth of useful information. Potapov & Lubk[47] inferred the amount of information carried by PCA components by generating scatter plots relating scores from all image pixels. When the point clouds display some structure or anisotropy, the components carry potentially meaningful information, while a symmetric round cloud reveals that these components contain mostly uncorrelated noise (see Moreira et al[36,37] for PCA score plots from EDS HSI measured for AuAg BNPs).

The strength and efficiency of this approach can be clearly visualized in Figure 4a, where we display the scatter plot correlating scores from components PCA#2 (average composition) and PCA#3 (composition variation). The horizontal axis ($s_{i,\mathrm{PCA\#2}}$) indicates the intensity of

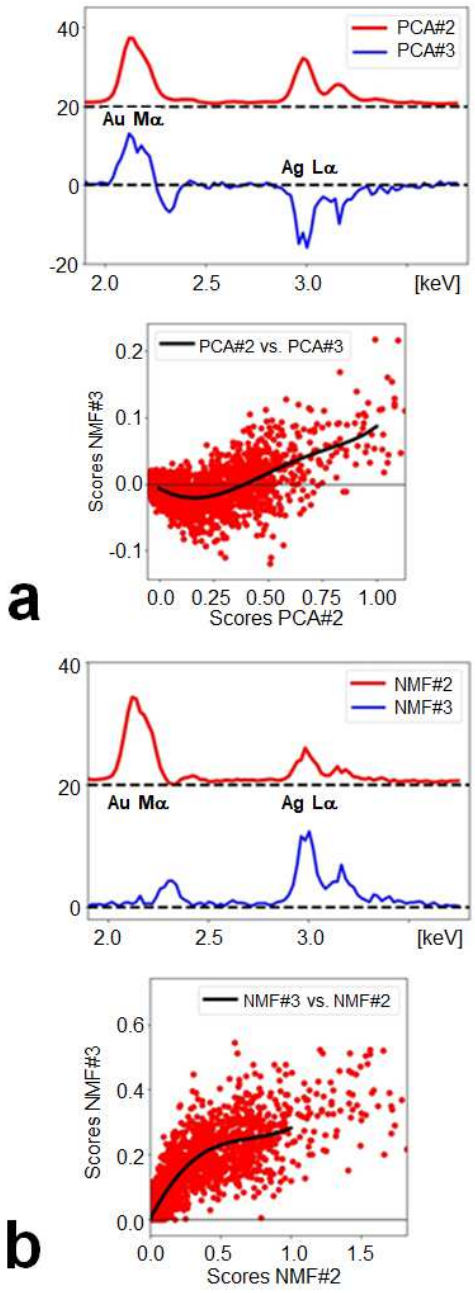


Figure 4: Machine Learning analysis of the HSI of bimetallic nanoparticles. a) Comparison of Au and Ag EDS peak profiles from PCA components #2 and #3 (remark that the signal of Ag EDS peak is negative for PCA#3); at the bottom, a scatter plot shows the correlation of scores. b) Comparison of components NMF#2 and NMF#3 in the Au and Ag EDS peak regions and the scatter plot correlating scores of NMF#2 and NMF#3. The elongated and curved shape of scatter clouds indicates a nonlinear relation between scores in vertical and horizontal axes. Both ML results directly suggest an Ag enrichment for the thinnest sample regions (NP edges or surfaces,) or smaller particles (see text for explanations). The black line in the scatter plots represents a polynomial fit to experimental points, in order to allow an easy eye recognition of the average non-linear relation between the scores. In a), the curve corresponding to PCA#2 is shifted upwards to facilitate visualization. The dashed lines are the zeros for each component. The same goes for b), where NMF#2 is also shifted.

both I_{Ag} and I_{Au}, so we can associate this axis, in a first approximation, to sample thickness at the pixel position. The vertical axis (PCA#3) represents the composition variation (Au enrichment in upward direction; note that for PCA#3 the Ag peak is negative). The cloud displays a curved shape displaying a gradual upward inclination as a function of increasing PCA#2 scores. An increasing PCA#3 score ($s_{i,\mathrm{PCA\#3}}$) means a stronger Au signal and less Ag counts (Equations 1 and 2, see Figure 4a). In addition, negative PCA#3 scores enhance Ag intensity (signal at contribution of Ag peak, product $s_{i,\mathrm{PCA\#3}}\, I_{\mathrm{Ag}}^{\mathrm{PCA\#3}}$ becomes positive) and diminish the Au signal (product $s_{i,\mathrm{PCA\#3}}\, I_{\mathrm{Au}}^{\mathrm{PCA\#3}}$) becomes negative. This occurs for the smaller PCA#2 scores (thinner region) or located at the particle border (see PCA#2 and PCA#3 score radial profiles in Figure S5, which confirm that negative PCA#2 scores occur close to NP surface). This simple analysis of pixel-by-pixel score correlations immediately unveils the information carried by the PCA components: Ag enrichment for thinnest sample regions (pixels located at BNP edges in their 2D projection).

There are several different methods for unsupervised machine learning, each with its advantages and drawbacks. PCA, as described above, is well suited for first data treatment, as its algorithm easily converges. It is useful for the establishment of the number of independent components necessary for the correct description of the data and for a recently suggested alternative error propagation after the denoising process.[36,37] A second, widely used technique is Non-negative Matrix Factorization (NMF), which is especially attractive for blind-source-separation, as all loadings and scores will be non-negative, which in principle represents better most physical measurements and also it should help to generate a better interpretation of latent factors.[48]

The application of NMF processing requires that the number of components be defined in advance; based on PCA analysis, we conclude three necessary NMF components. The loading information bears some resemblance to PCA results, with the exception that all loadings (spectral information) and scores will be positive, as expected for this method (see Figure 4b). The intensity estimation involves just products or additions of positive numbers

(Eqs. 3 and 4). In this case, the equations for loadings and scores are written as:

$$I^{\mathrm{NMF}}_{\mathrm{Au},i} = s_{i,\mathrm{NMF\#2}}\ I^{\mathrm{NMF\#2}}_{\mathrm{Au}} + s_{i,\mathrm{NMF\#3}}\ I^{\mathrm{NMF\#3}}_{\mathrm{Au}} \tag{3}$$

$$I^{\mathrm{NMF}}_{\mathrm{Ag},i} = s_{i,\mathrm{NMF\#2}}\ I^{\mathrm{NMF\#2}}_{\mathrm{Ag}} + s_{i,\mathrm{NMF\#3}}\ I^{\mathrm{NMF\#3}}_{\mathrm{Ag}} \tag{4}$$

Analyzing the information content of different NMF components, we find that the first one (NMF#1) again contains negligible information on the particles. The second NMF loading (NMF#2) contains a major Au peak and a minor Ag one, while NMF#3 includes a very strong dominant Ag peak. The NMF#3 component describes changes in the $I_{\mathrm{Ag}}/I_{\mathrm{Au}}$ ratio by mainly adding Ag counts, with negligible effect on the Au intensity. The ML procedures reveal a small peak in the 2.3 keV region, close to the Au Mα line (it appears positive in PCA#3 and negative in NMF#3 components). Neither PCA nor NMF consider EDS peaks in a specific way such as energy position, width, or relative intensity, so these unexpected ML results are not associated with specific EDS data treatments or fitting. It is important to note that this energy region is difficult to interpret from the x-ray spectroscopy perspective.[49] The main peak from sulfur (S) is expected at this precise energy, and it could be expected in chemically synthesized NPs using thiol as a passivating agent. This explanation can be ruled out because our samples have been produced by physical synthesis, without chemical passivation. Another possible explanation, which we consider to be the best to explain the ML anomalous peak at 2.3 KeV, is a sum peak due to the simultaneous detection of Si and O photons (1.74 keV + 0.53 keV $\sim$ 2.3 keV). In this case, the Si photon is due to a process known as self-fluorescence in the EDS detector made from a silicon crystal.[49] How this spurious EDS signal is recognized and revealed by ML tools is somewhat difficult to predict. Therefore, these peaks do not have a physical origin in the nanoparticle sample but instead are due to the EDS detection process.

An identical analysis of score correlation (Figure 4b) can also be applied to the NMF data. Again, we observe a nonlinear behavior (elongated and curved point cloud). The

relative weight of the NMF#3 component (additional Ag signal) diminishes gradually as the NMF#2 score (or Au signal) increases. If the NMF#2 scores (horizontal axis) are interpreted as before (thickness or NP diameter indicator), this again shows that the relative weight of the Ag signal decreases for increasing Au intensity (left to right in NMF#2 scores) and it points out an Ag depletion for thicker regions (Au enrichment at the NP core or, in other terms Ag enrichment at the surface; this might also point out Ag enrichment for smaller particles). As for PCA, the NMF results provide qualitative support and statistical validity of conclusions on chemical composition variations revealed by the quantitative EDS chemical analysis.

The two blind source separation methods, PCA and NMF, translate the information content in the EDS HSI into two quite different encoding procedures but the two different ML approaches have both detected and identified the same tendency or latent information in the EDS signals. It is important to mention that the input matrix of ML data represents just a list of the spectra, pixel by pixel, and absolutely no information about their spatial location. In this way, there is no possibility of any biased interpretation of the spectral characteristics of pixels located at the center or close to the surface of the nanoparticle. In strict terms, the scatter plots in Figure 4 confirm the statistical validity in relation to $I_{\mathrm{Ag}}/I_{\mathrm{Au}}$ and its variation pixel by pixel without any assumption or model on the numerical algorithms used for nanoparticle identification or NP diameter estimation or chemical composition calculation.

In order to improve both, our understanding of actual BNP chemical structure and also of the validity of the ML results, we have also analyzed simulated data sets. The aim is to observe if BNPs with controlled Ag surface segregation will generate similar information as the one derived from the experimental data. We have constructed 6 nm diameter BNPs formed of a 4 nm core of composition $C_{\mathrm{Ag}} = 0.3$, and a 1 nm-thick shell with composition of $C_{\mathrm{Ag}} = 0.6$. This configuration guarantees that the BNP chemical composition as a whole is $C_{\mathrm{Ag,NP}} \sim 0.5$, as observed in our experiments. Figures 5 and S6 show that the results of ML processing of the simulated HSI are in full agreement with our experimental results. Our

simulated HSI contains an ensemble of particles of identical radius (no size distribution), so the analysis of score scatter plots as presented in Figure 5 can be directly associated with sample thickness at the analyzed pixel, simplifying the interpretation. Since the particles have been simulated with an abrupt core-shell interface and a homogeneous shell of $C_{\mathrm{Ag}} = 0.6$ composition, the derived composition profile correctly displays a clear plateau of chemical composition at the NP surface in the radial profiles. This is, however, in contrast with the experimental profiles, where the surface shows a constantly rising profile, suggesting that the bimetallic particles contain a composition gradient across the BNP.

3 Conclusions

We have experimentally measured radial chemical gradients inside individual bimetallic AuAg alloy NPs between 3 and 7 nm diameter using EDS-STEM and, for most of the BNPs the uncertainty intervals from chemical concentration particle center and surface do not overlap considering 3-sigma criteria, confirming the robust assessment of Ag surface enrichment. Machine Learning tools (PCA and NMF) have been used to verify if latent information on radial composition variations is carried by our dataset. This analysis shows that experimental data contains latent information concerning Ag enrichment in the thinnest sample regions (i.e. BNP surface), confirming qualitatively the detection of chemical concentration variations in the sample.

To achieve this result, the thorough planning of proposed procedures considers the data collection, careful EDS intensity analysis associated with rigorous error bar calculations, and the analysis of raw data and its statistical validation through unsupervised machine learning procedures and simulated datasets. Furthermore, the proposed procedures and interpretations are tested with simulated data. These results open the way for more accurate structural and chemical studies of alloy nanosystems, in particular considering the constant progress of EDS detector technology and the recent increase of detection solid angles to more

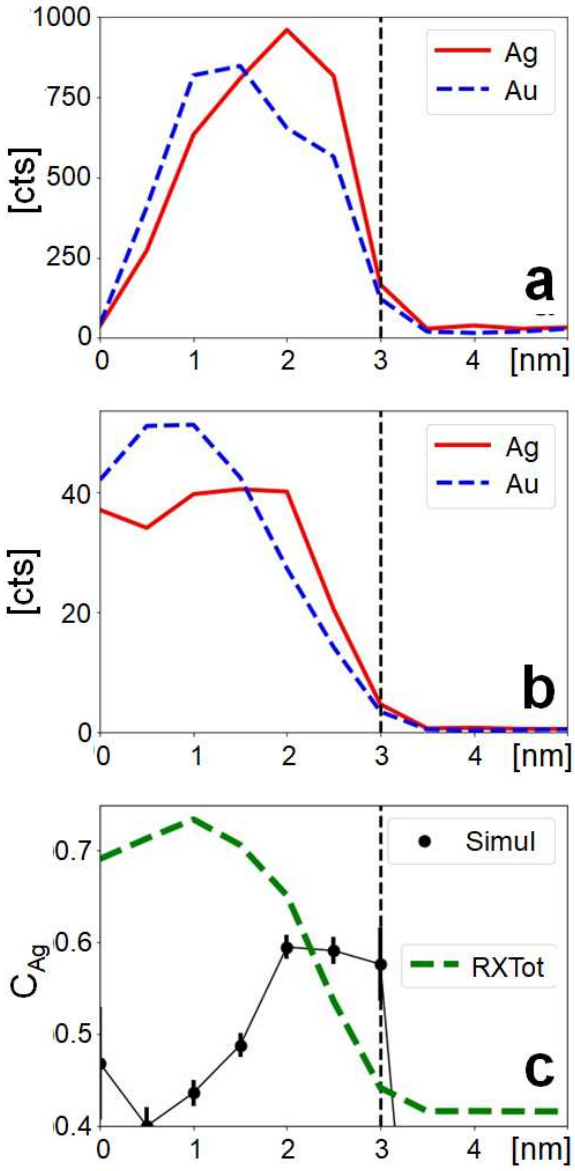


Figure 5: Chemical analysis results obtained from a simulated HSI (core-shell 6 nm NP). a) Azimuthal integration of X-ray intensity for an individual nanoparticle showing the total EDS signal for each radial position. b) Radial intensity profile normalized by the number of pixels at each radius. c) Quantitative chemical composition analysis (including error bars) showing the unambiguous detection of an enriched shell. In (c), the total elemental EDS signal (Au + Ag, RXTot) intensity profile is also displayed, analogous to the ADF profile in Figure 3c; the vertical dashed line indicates the BNP radius.

than 4 sR.[50] Whether the observed Ag enrichment at the BNP surface is intrinsic to the nanoalloy or due to oxidation is beyond the scope of this article and will be addressed in a future publication.

Our study has revealed that ML methods must be carefully chosen as a function of the underlying scientific question. PCA is usually considered not well suited for EDS due to the occurrence of unphysical loading profiles (EDS spectra) including negative intensities. However, in our case study, the physical meaning of the principal components is very clear and the information contained is extremely useful (average composition from PCA#2 and composition variations from PCA#3).

NMF processing is more easily accepted due to its "non-negativity", where all outputs, loadings, and scores, must be positive so that they always represent a physical spectrum. Our NMF decomposition of EDS spectra from AuAg BNPs shows that the intensity expected in each pixel is related to the coupled contribution of both NMF#2 and NMF#3, both in the core and close to the surface. This means that none of the NMF components carries discernible physical information (not even NP average composition, as contained in PCA#2). A very optimistic expectation of NMF decomposition could be that NMF components would discriminate the existence of two different metal phases (core and shell) in the BNP. Our simulation clearly shows that no such discrimination was observed. Briefly, it is always important to consider that, depending on the technical/scientific question to be answered, sometimes PCA may be preferable to NMF to analyze EDS HSI, while sometimes the complementarity of both reveals an optimum of information.

4 Materials and Methods

4.1 Nanoparticle production

Binary alloy (Au_xAg_{1-x}) NPs have been generated using a homemade gas aggregation source described in detail in references,[26,27] figure 1a). Briefly, a hollow cylindrical magnetron is

used to sputter atoms from a target of a target located on the axis of the cylinder and made of twisted Au and Ag wires. In-situ time-of-flight mass spectrometry has been used to measure the BNP mass distribution which follows a log-normal function with $\sim$ 4 nm mean average diameter and 3 nm in width.[36] Using ion optics, the particles were deposited in "soft landing" mode on the TEM grid (kinetic energy per atom $\sim$ 0.05 eV for a 4 nm NP). The average Au composition in an atomic fraction of the sample is $C_{\mathrm{Au}} = 0.51 \pm 0.01$, as obtained by EDS-TEM from an ensemble of NPs using a large open electron beam (diameter several microns).

4.2 Electron microscopy: data acquisition and processing

We have used a STEM with 4 silicon drift detectors (SDD) (Titan Themis, Super X Quad SDD, 0.8 sR, LNNANO-Campinas-Brazil) operating at 80 keV. The EDS HSI acquisition parameters were 0.59 nm for the pixel size, and a dwell time of $\sim$200 ms per pixel; the image size was set to 60×60 pixels to guarantee the operation of automatic drift correction when realigning scan regions. To minimize the dose rate, a series of images of the same region was acquired, and subsequently, individual frames were added (10 scans at 20 ms dwell time).[51] For all experiments we have used a low background Be sample holder. Raw EDS spectra were binned to 512 energy channels to increase the signal-to-noise ratio (SNR), leading to an EDS channel width of 20 eV. The integrated signal from the Au-Mα and Ag-Lα EDS peaks has been used to perform chemical composition calculations.

The quantitative analysis of the individual NP chemical composition followed the Cliff-Lorimer approach[28] using experimentally measured K_{AB} factors (Eq. 5, C_{y} and $I_{\mathrm{y,\ NP}}$, atomic percentage and X-ray intensity of element y per NP respectively).

$$\frac{C_{\mathrm{A}}}{C_{\mathrm{B}}} = K_{\mathrm{AB}} \frac{I_{\mathrm{A,\ NP}}}{I_{\mathrm{B,\ NP}}} \tag{5}$$

The Cliff-Lorimer factor (K_{AB}) has been measured using a thin film of known compo-

sition, $K_{\mathrm{AgAu}} = 1.17 \pm 0.01$. A multilayer thin film was generated by thermal evaporation onto a substrate using a quartz balance to control the atomic ratio through the film thickness. Estimations of chemical composition and their error bars have taken into account all quantitative EDS analysis steps applied on a fitted curve of the experimental EDS spectra (background removal, X-ray peak integration, and composition determination). To calculate error bars, we have considered Poisson noise for all EDS count measurements (EDS peaks counts and the substracted background); the other source of uncertainty was the Cliff-Lorimer factor; absorption has been neglected). All EDS, and ML processing steps have been performed using the open-source Hyperspy Python library.[52]

4.3 Simulated Datasets

Synthetic HSI data must be a realistic description of actual experimental situations. Different factors induce modifications of the overall EDS SNR for each pixel, which determines the ML output.

Firstly, it is important to generate bimetallic NPs (6 nm in diameter) that display the surface segregation of one chemical element. This has been modeled using a core-shell particle with abrupt interface, where the center is silver depleted to $C_{\mathrm{Ag}} \sim 0.3$ and the surface layer (1 nm thick) is Ag enriched ($C_{\mathrm{Ag}} \sim 0.6$). A simple calculation shows that this geometrical core-shell structure reproduces the experimental BNP composition as $C_{\mathrm{Ag}} \sim 0.5$. Secondly, the HSI should not contain identical particles, where all pixels inside the NP show the same noise-free EDS signal before the addition of Poisson noise. This configuration may generate a different SNR than in an experiment. To avoid this, we have generated an ensemble of different core-shell NPs with a random distribution of the alloy atoms inside the particles. In practice, we have considered a spherical volume with fcc atomic arrangement, and the elemental occupation was determined with a random number generator considering a binomial probability of occupation at each pixel region based on chemical composition (example 70% and 30% for Au and Ag atom at core). This yields the number of Au or

Ag atom in each HSI pixel, and immediately account for variations of thickness (or of the electron beam path due to spherical shape) in the NP, such that counts diminish from center to surface pixels. As the simulated HSI pixel step of 0.5 nm is very close to the Au/Ag lattice parameter of ∼0.4 nm, we have noted that a well-oriented crystal (zone axis along the electron beam direction) generates a well-defined pattern of atom distribution in the 2D NP projection. As this pattern could strongly influence ML analysis, we have added a random 2-axes rotation to the NP crystal before calculating its 2D projection. Subsequently, we have defined the total number of counts per NP and distributed them inside pixels. The EDS peaks were defined in accordance with experimental width and total counts per pixel. EDS background was added by the conventional formula for EDS continuous background (Kramers formula $I_{cm} \approx i_p Z \left[\left(E_0 - E_\nu \right) / E_\nu \right]$, where I_{cm} is the x-ray continuum intensity background at the at an energy E_ν, i_p is the incident beam current and Z is the atomic number).[53] The background was scaled to get a Peak-to-Background level in agreement with the experiments. The particles have been assumed to be deposited on an a-C film and spurious peaks from the sample environment (C, Fe, Ni, etc.) have been also incorporated into the EDS spectrum (peak intensities heights were set such that they reproduce measurements). Finally, random Poisson noise was added to the noise-free X-ray spectra. For the sake of simplicity, our simulations have employed $K_{\mathrm{AgAu}} = 1$.

4.4 PCA & NMF Processing.

Before processing, we performed a variance stabilization of data through rescaling.[54–56] The so-called scree-plot[30,31,57] is used to compare the covariance of each eigenspectum and to identify the most relevant loading. In the studies reported here, the scree-plots show a clear profile displaying a well-defined kink allowing the clear identification of three components whose data variance is higher than noise (see example in Figure 1 in reference[36]). For the NMF treatment, three components have been used in the reconstruction, in accordance with the number of components derived from the PCA scree-plot. It is important to consider that

the NMF algorithm is non-convex, so the minimization procedures may get trapped in local minima,[58] with each run possibly yielding a different decomposition result as demonstrated for the analysis of diffraction pattern in 4D-STEM HSI.[41] It is always recommendable to test the algorithm convergence to the same solution by realizing different runs with different starting values. For the experiments addressed here, the number of components is rather low (3 NMF components), so we have not found solution variations between runs; however, this may be critical when many NMF components must be recovered.[41]

Acknowledgement

D.U. acknowledges financial support from the Brazilian Agencies FAPESP (No. 2014/01045-0), CNPq (402571/2016-9, 306513/2017-0 and 402676/2021-1) and FAEPEX-UNICAMP (2632/17). V. R. thanks funding from FAPESP (2007/01722-9, 2020/15497-1 and 2022/03869-7) and CNPq (306787/2020, 555647/2006-4 and 577046/2008-0). M.M. thanks funding from CNPq (No. 162541/2018-0). Access to the FEG-TEM/STEM from the Brazilian Nanotechnology National Laboratory is acknowledged (LNNANO, grant No. ME - 22329). M.H. acknowledges financial support from FAPESP (No. 2016/12807-4) and the French National Research Agency (ANR) via the project 'SchNAPSS', ANR-21-CE09-0021. Further support by the French Campus France Eiffel Excellency Scholarship Program (EIFFEL-DOCTORAT 2020/P760767) for M.M. is acknowledged.

4.5 Author Contributions

M.M. performed the experiments. Data processing, and simulations were realized by M.M., V.R. and D.U. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Supporting Information Available

NPs EDS count profile; NP localization algorithm; chemical composition of AuAg NP; radial evolution of AuAg NP chemical composition; analysis of the EDS signal for all NPs detected in the reported experiment; machine learning results obtained by processing the experimental STEM-EDS HSI; machine learning results obtained by processing a simulated BNPs EDS data.

References

(1) Alloyeau, D., Mottet, C., Ricolleau, C., Eds. *Nanoalloys: Synthesis, Structure and Properties*; Springer-Verlag London, 2012.

(2) Ferrando, R. *Structure and Properties of Nanoalloys*; Frontiers of Nanoscience; Elsevier, 2017; Vol. 10.

(3) Calvo, F., Ed. *Nanoalloys: From Fundamentals to Emergent Applications*, second edition ed.; Elsevier, 2020.

(4) Cottancin, E.; Broyer, M.; Lermé, J.; Pellarin, M. In *Handbook of Nanophysics: Nanoelectronics and Nanophotonics*; Sattler, K., Ed.; CRC Press: Boca Raton, FL, USA, 2011; Chapter Optical Properties of Metal Clusters and Nanoparticles.

(5) Camus, E.; Pellarin, M.; Blanchard, N.; Boisron, O.; Hillenkamp, M.; Roiban, L.; Andreazza, P.; Cottancin, E. Structural and Optical Characterization of Nanoalloys Mixing Gold or Silver With Aluminium or Indium: Evolution Under Various Reactive Environments. *Faraday Discuss.* **2022**, https://doi.org/10.1039/D2FD00109H.

(6) García, S.; Zhang, L.; Piburn, G. W.; Henkelman, G.; Humphrey, S. M. Microwave Synthesis of Classically Immiscible Rhodium–Silver and Rhodium–Gold Alloy Nanoparticles: Highly Active Hydrogenation Catalysts. *ACS Nano* **2014**, *8*, 11512.

(7) Zhou, L.; Swearer, D. F.; Zhang, C.; Robatjazi, H.; Zhao, H.; Henderson, L.; Dong, L.; Christopher, P.; Carter, E. A.; Nordlander, P.; Halas, N. J. Quantifying hot carrier and thermal contributions in plasmonic photocatalysis. *Science* **2018**, *362*, 69–72.

(8) Huang, B.; Kobayashi, H.; Yamamoto, T.; Toriyama, T.; Matsumura, S.; Nishida, Y.; Sato, K.; Nagaoka, K.; Haneda, M.; Xie, W.; Nanba, Y.; Koyama, M.; Wang, F.; Kawaguchi, S.; Kubota, Y.; Kitagawa, H. A CO Adsorption Site Change Induced by Copper Substitution in a Ruthenium Catalyst for Enhanced CO Oxidation Activity. *Angew. Chem. Int. Ed.* **2019**, *58*, 2230–2235.

(9) Xie, C.; Niu, Z.; Kim, D.; Li, M.; Yang, P. Surface and Interface Control in Nanoparticle Catalysis. *Chem. Rev.* **2020**, *120*, 1184–1249.

(10) Slater, T. J. A.; Macedo, A.; Schroeder, S. L. M.; Burke, M. G.; O'Brien, P.; Camargo, P. H. C.; Haigh, S. J. Correlating Catalytic Activity of Ag–Au Nanoparticles with 3D Compositional Variations. *Nano Lett.* **2014**, *14*, 1921–1926.

(11) Skorikov, A.; Albrecht, W.; Bladt, E.; Xie, X.; van der Hoeven, J. E. S.; van Blaaderen, A.; Van Aert, S.; Bals, S. Quantitative 3D Characterization of Elemental Diffusion Dynamics in Individual Ag@Au Nanoparticles with Different Shapes. *ACS Nano* **2019**, *13*, 13421–13429.

(12) Chen, P.-C.; Liu, X.; Hedrick, J. L.; Xie, Z.; Wang, S.; Lin, Q.-Y.; Hersam, M. C.; Dravid, V. P.; Mirkin, C. A. Polyelemental nanoparticle libraries. *Science* **2016**, *352*, 1565–1569.

(13) Guisbiers, G.; Mendoza-Cruz, R.; Bazán-Díaz, L.; Velázquez-Salazar, J. J.; Mendoza-Perez, R.; Robledo-Torres, J. A.; Rodriguez-Lopez, J.-L.; Montejano-Carrizales, J. M.; Whetten, R. L.; José-Yacamán, M. Electrum, the Gold-Silver Alloy, from the Bulk Scale to the Nanoscale: Synthesis, Properties, and Segregation Rules. *ACS Nano* **2016**, *10*, 188–198.

(14) Lasserus, M.; Schnedlitz, M.; Knez, D.; Messner, R.; Schiffmann, A.; Lackner, F.; Hauser, A. W.; Hofer, F.; Ernst, W. E. Thermally induced alloying processes in a bimetallic system at the nanoscale: AgAu sub-5 nm core–shell particles studied at atomic resolution. *Nanoscale* **2018**, *10*, 2017–2024.

(15) Carter, C. B., Williams, D. B., Eds. *Transmission Electron Microscopy*; Springer International Publishing: Cham, 2016.

(16) Goldstein, J. I.; Newbury, D. E.; Michael, J. R.; Ritchie, N. W.; Scott, J. H. J.; Joy, D. C. *Scanning Electron Microscopy and X-Ray Microanalysis*; Springer New York: New York, NY, 2018.

(17) Pennycook, S. J., Nellist, P. D., Eds. *Scanning Transmission Electron Microscopy*; Springer New York, NY, 2011.

(18) Thomas, J. M.; Leary, R. K.; Eggeman, A. S.; Midgley, P. A. The rapidly changing face of electron microscopy. *Chem. Phys. Lett.* **2015**, *631-632*, 103–113.

(19) Brown, K. A.; Brittman, S.; Maccaferri, N.; Jariwala, D.; Celano, U. Machine Learning in Nanoscience: Big Data at Small Scales. *Nano Lett.* **2020**, *20*, 2–10.

(20) Spurgeon, S. R.; Ophus, C.; Jones, L.; Petford-Long, A.; Kalinin, S. V.; Olszta, M. J.; Dunin-Borkowski, R. E.; Salmon, N.; Hattar, K.; Yang, W.-C. D.; Sharma, R.; Du, Y.; Chiaramonti, A.; Zheng, H.; Buck, E. C.; Kovarik, L.; Penn, R. L.; Li, D.; Zhang, X.; Murayama, M.; Taheri, M. L. Towards data-driven next-generation transmission electron microscopy. *Nature Materials* **2021**, *20*, 274.

(21) Liao, T.-W.; Yadav, A.; Hu, K.-J.; van der Tol, J.; Cosentino, S.; D'Acapito, F.; Palmer, R. E.; Lenardi, C.; Ferrando, R.; Grandjean, D.; Lievens, P. Unravelling the nucleation mechanism of bimetallic nanoparticles with composition-tunable core–shell arrangement. *Nanoscale* **2018**, *10*, 6684–6694.

(22) Sotiriou, G. A.; Etterlin, G. D.; Spyrogianni, A.; Krumeich, F.; Leroux, J.-C.; Pratsinis, S. E. Plasmonic biocompatible silver–gold alloyed nanoparticles. *Chem. Commun.* **2014**, *50*, 13559–13562.

(23) Liu, J.-H.; Wang, A.-Q.; Chi, Y.-S.; Lin, H.-P.; Mou, C.-Y. Synergistic effect in an Au-Ag alloy nanocatalyst: CO oxidation. *J. Phys. Chem. B* **2005**, *109*, 40–43.

(24) Tokonami, S.; Morita, N.; Takasaki, K.; Toshima, N. Novel Synthesis, Structure, and Oxidation Catalysis of Ag/Au Bimetallic Nanoparticles. *J. Phys. Chem. C* **2010**, *114*, 10336–10341.

(25) Wang, J. L.; Ando, R. A.; Camargo, P. H. C. Investigating the Plasmon-Mediated Catalytic Activity of AgAu Nanoparticles as a Function of Composition: Are Two Metals Better than One? *ACS Catal.* **2014**, *4*, 3815–3819.

(26) de Sá, A. D. T.; Oiko, V. T. A.; di Domenicantonio, G.; Rodrigues, V. New experimental setup for metallic clusters production based on hollow cylindrical magnetron sputtering. *J. Vac. Sci. Technol. B* **2014**, *32*, 061804.

(27) Oiko, V. T. A.; de Sá, A. D. T.; Rodrigues, V. *Gas-Phase Synthesis of Nanoparticles*; John Wiley & Sons, Ltd, 2017; Chapter 7, pp 123–136.

(28) Cliff, G.; Lorimer, G. W. The quantitative analysis of thin specimens. *Journal of Microscopy* **1975**, *103*, 203–207.

(29) Bevington, P. R.; Robinson, D. K. *Data reduction and error analysis for the physical sciences; 3rd ed.*; McGraw-Hill: New York, NY, 2003.

(30) Jolliffe, I. *Principal Component Analysis*; Springer New York, 2002.

(31) Jolliffe, I. T.; Cadima, J. Principal component analysis: a review and recent developments. *Philos. Trans. Royal Soc. A* **2016**, *374*, 20150202.

(32) Lee, D. D.; Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791.

(33) Lin, C.-J. Projected Gradient Methods for Nonnegative Matrix Factorization. *Neural Computation* **2007**, *19*, 2756–2779.

(34) Moreira, M.; Felix, L. C.; Cottancin, E.; Pellarin, M.; Ugarte, D.; Hillenkamp, M.; Galvao, D. S.; Rodrigues, V. Influence of Cluster Sources on the Growth Mechanisms and Chemical Composition of Bimetallic Nanoparticles. *The Journal of Physical Chemistry C* **2023**, *127*, 1944–1954.

(35) Corrêa, L. M.; Moreira, M.; Rodrigues, V.; Ugarte, D. Quantitative Structural Analysis of AuAg Nanoparticles Using a Pair Distribution Function Based on Precession Electron Diffraction: Implications for Catalysis. *ACS Applied Nano Materials* **2021**, *4*, 12541–12551.

(36) Moreira, M.; Hillenkamp, M.; Divitini, G.; Tizei, L. H. G.; Ducati, C.; Cotta, M. A.; Rodrigues, V.; Ugarte, D. Improving Quantitative EDS Chemical Analysis of Alloy Nanoparticles by PCA Denoising: Part I, Reducing Reconstruction Bias. *Microscopy and Microanalysis* **2022**, *28*, 338–349.

(37) Moreira, M.; Hillenkamp, M.; Divitini, G.; Tizei, L. H. G.; Ducati, C.; Cotta, M. A.; Rodrigues, V.; Ugarte, D. Improving Quantitative EDS Chemical Analysis of Alloy Nanoparticles by PCA Denoising: Part II. Uncertainty Intervals. *Microscopy and Microanalysis* **2022**, *28*, 723–731.

(38) Bevington, P.; Robinson, D. *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed.; McGraw-Hill Education, 2003.

(39) Hughes, I.; Hase, T. *Measurements and their Uncertainties: A practical guide to modern error analysis*; Oxford University Press, 2010.

(40) Drosg, M. *Dealing with Uncertainties: A Guide to Error Analysis*; Springer Berlin, Heidelberg, 2009.

(41) Uesugi, F.; Koshiya, S.; Kikkawa, J.; Nagai, T.; Mitsuishi, K.; Kimoto, K. Non-negative matrix factorization for mining big data obtained using four-dimensional scanning transmission electron microscopy. *Ultramicroscopy* **2021**, *221*, 113168.

(42) Belić, D.; Chantry, R. L.; Li, Z. Y.; Brown, S. A. Ag-Au nanoclusters: Structure and phase segregation. *Appl. Phys. Lett.* **2011**, *99*, 171914.

(43) Eom, N.; Messing, M. E.; Johansson, J.; Deppert, K. General Trends in Core-Shell Preferences for Bimetallic Nanoparticles. *ACS Nano* **2021**, *15*, 8883–8895.

(44) Nelli, D.; Mottet, C.; Ferrando, R. Interplay between interdiffusion and shape transformations in nanoalloys evolving from core–shell to intermixed structures. *Faraday Discuss.* **2023**, *242*, 52–68.

(45) Egerton, R. F.; McLeod, R.; Wang, F.; Malac, M. Basic questions related to electron-induced sputtering in the TEM. *Ultramicroscopy* **2010**, *110*, 991–997.

(46) Braidy, N.; Jakubek, Z. J.; Simard, B.; Botton, G. A. Quantitative energy dispersive X-ray microanalysis of electron beam-sensitive alloyed nanoparticles. *Microscopy and microanalysi* **2008**, *14*, 166–175.

(47) Potapov, P.; Lubk, A. Optimal principal component analysis of STEM XEDS spectrum images. *Adv. Struct. Chem. Imag.* **2019**, *5*, 4.

(48) Jany, B. R.; Janas, A.; Krok, F. Retrieving the Quantitative Chemical Information at Nanoscale from Scanning Electron Microscope Energy Dispersive X-ray Measurements by Machine Learning. *Nano Letters* **2017**, *17*, 6520–6525.

(49) Goldstein, J. I.; Newbury, D. E.; Michael, J. R.; Ritchie, N. W. M.; Scott, J. H. J.;

Joy, D. C. *Scanning Electron Microscopy and X-Ray Microanalysis*; Springer New York: New York, NY, 2018; Chapter 18.

(50) Zaluzec, N. J. X-ray Spectrometry in the Era of Aberration-Corrected Electron Optical Beam Lines. *Microscopy and Microanalysis* **2023**, *29*, 334–340.

(51) Jones, L.; Varambhia, A.; Beanland, R.; Kepaptsoglou, D.; Griffiths, I.; Ishizuka, A.; Azough, F.; Freer, R.; Ishizuka, K.; Cherns, D.; Ramasse, Q. M.; Lozano-Perez, S.; Nellist, P. D. Managing dose-, damage- and data-rates in multi-frame spectrum-imaging. *Microscopy* **2018**, *67*, i98–i113.

(52) De La Peña, F.; Ostasevicius, T.; Fauske, V. T.; Burdet, P.; Jokubauskas, P.; Nord, M.; Sarahan, M.; Prestat, E.; Johnstone, D. N.; Taillon, J.; Caron, J.; Furnival, T.; Macarthur, K. E.; Eljarrat, A.; Mazzucco, S.; Migunov, V.; Aarholt, T.; Walls, M.; Winkler, F.; Donval, G.; Martineau, B.; Garmannslund, A.; Zagonel, L.-F.; Iyengar, I. Electron Microscopy (Big and Small) Data Analysis With the Open Source Software Package HyperSpy. *Microscopy and Microanalysis* **2017**, *23*, 214–215.

(53) Goldstein, J. I.; Newbury, D. E.; Michael, J. R.; Ritchie, N. W. M.; Scott, J. H. J.; Joy, D. C. *Scanning Electron Microscopy and X-Ray Microanalysis*; Springer New York: New York, NY, 2018; p 49.

(54) Keenan, M. R.; Kotula, P. G. Accounting for Poisson noise in the multivariate analysis of ToF-SIMS spectrum images. *Surf. Interface Anal.* **2004**, *36*, 203–212.

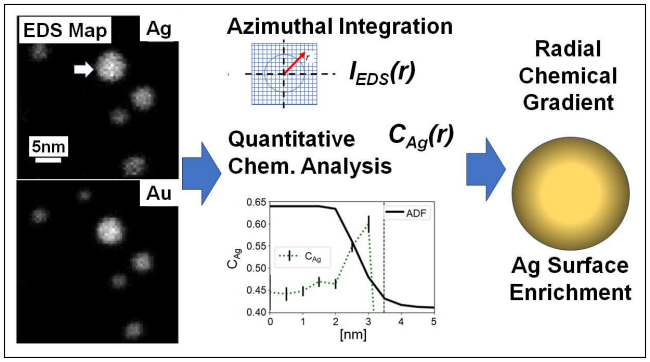
(55) Kotula, P. G.; Keenan, M. R. Application of Multivariate Statistical Analysis to STEM X-ray Spectral Images: Interfacial Analysis in Microelectronics. *Microsc. Microanal.* **2006**, *12*, 538–544.

(56) Kotula, P. G.; Benthem, M. H. V. Revisiting Noise Scaling for Multivariate Statistical Analysis. *Microsc. Microanal.* **2015**, *21*, 1423–1424.

(57) Titchmarsh, J. M. EDX spectrum modelling and multivariate analysis of sub-nanometer segregation. *Micron* **1999**, *30*, 159–171.

(58) Martineau, B. H.; Johnstone, D. N.; van Helvoort, A. T. J.; Midgley, P. A.; Eggeman, A. S. Unsupervised machine learning applied to scanning precession electron diffraction data. *Advanced structural and chemical imaging* **2019**, *5*, 3.

TOC Graphic


EDS Map
Ag
5nm
Au
Azimuthal Integration
$I_{EDS}(r)$
Quantitative
Chem. Analysis
$C_{Ag}(r)$
ADF
C_{Ag}
[nm]
Radial
Chemical
Gradient
Ag Surface
Enrichment

Supporting Information: Ag Surface Segregation in Sub-10 nm Bimetallic AuAg Nanoparticles Quantified by STEM-EDS and Machine Learning: Implications for Fine-Tuning Physico-Chemical Properties for Plasmonics and Catalysis Applications

Murilo Moreira,[†,‡] Matthias Hillenkamp,[†,‡] Varlei Rodrigues,[*,†] and Daniel Ugarte[*,†]

†*Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin, Campinas, SP 13083-859, Brazil*

‡*Institute of Light and Matter, University Lyon, Université Claude Bernard Lyon 1, CNRS, UMR5306, Villeurbanne F-69622, France*

E-mail: varlei@ifi.unicamp.br; dmugarte@ifi.unicamp.br

Supporting Information Available

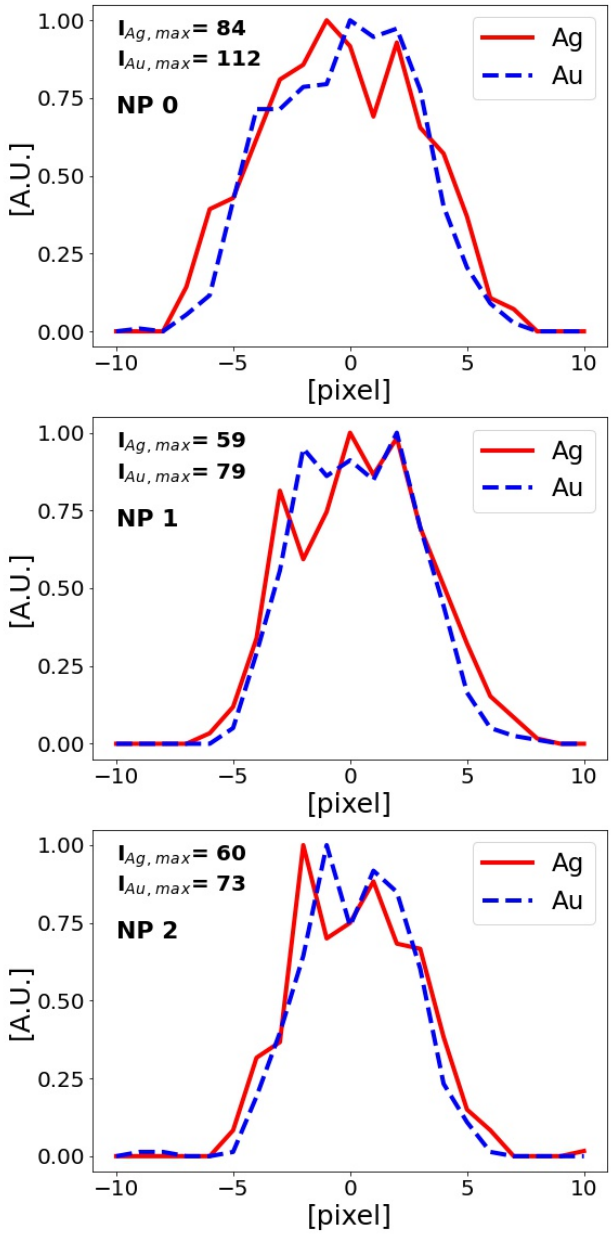


Figure S1: EDS intensity profiles from the 3 larger NPs in the dataset (horizontal line of pixels crossing through NP center). Note that in all cases, the Ag intensity at the BNP border is slightly higher than for Au. Intensity variations in the BNP center are attributed to Poisson noise (profiles have been renormalized using their maximum count value, indicated at the top left in each plot).

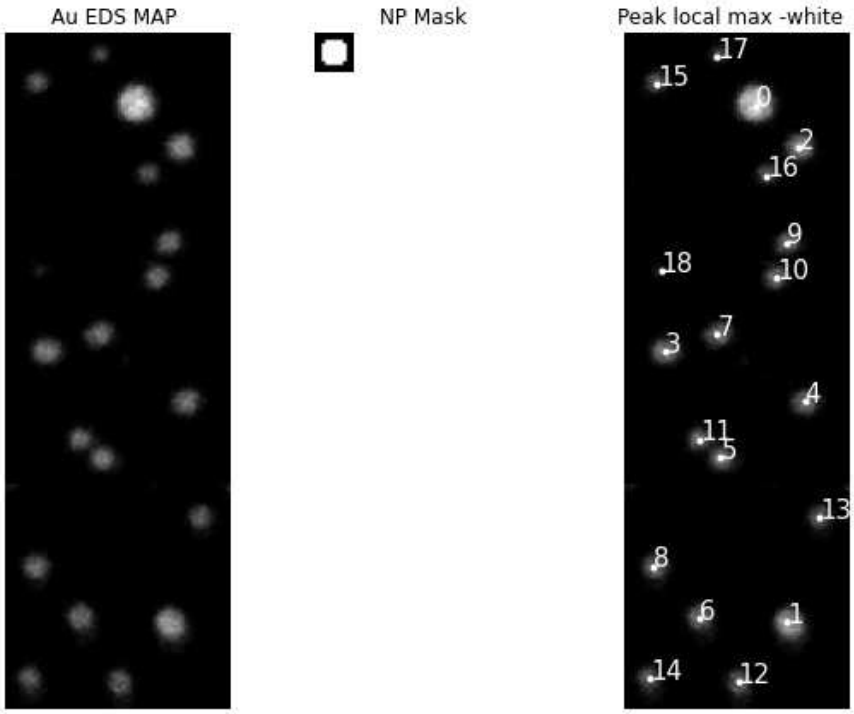


Figure S2: NP localization algorithm. A binarized disk (center) corresponding to a circular particle image is used to perform a cross-correlation with the EDS signal map ((Au + Au) signal). NP centers are easily identified by a local-maxima finding algorithm applied to the cross-correlation result. In the right image, the numbers indicate NP indexing and they have been sorted as a function of cross-correlation function decreasing value (or in other terms, NP decreasing maximum intensity which we assume a directly correlated with NP diameter).

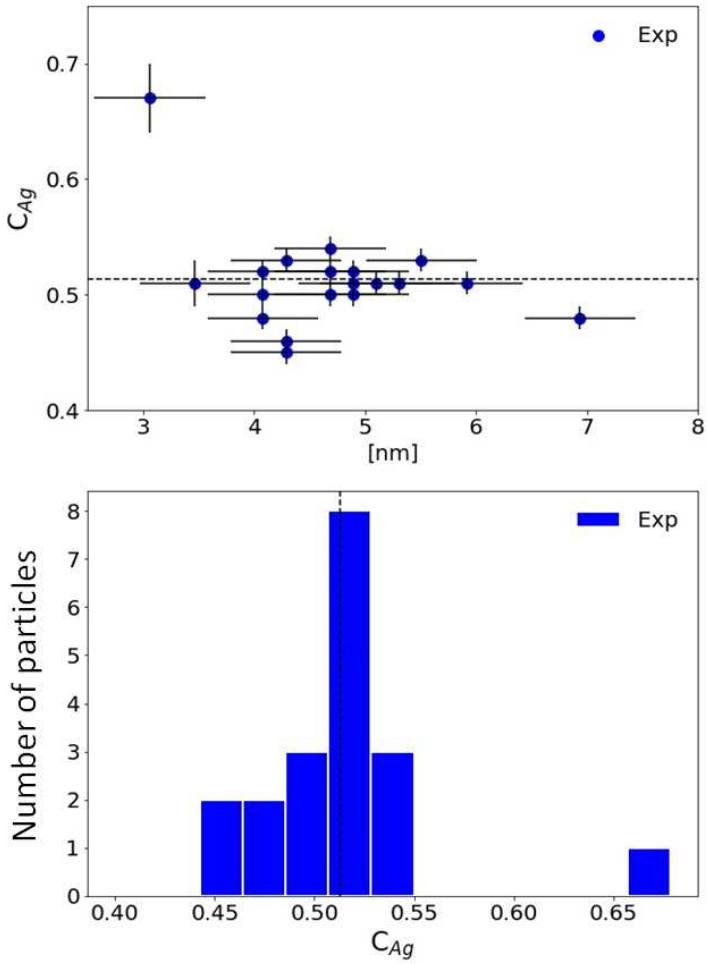


Figure S3: Quantitative analysis of the chemical composition of bimetallic alloy NPs as a function of diameter and composition histogram for the HSI displayed in Figure 2 (total of 19 NPs).

Table S1: Radial Evolution of chemical composition deduced by quantitative EDS analysis. Data correspond to the particle indicated as NP#0 in Figure S2 (for this particle the measured chemical composition was $C_{\mathrm{Ag}} = 0.49 \pm 0.01$. EDS intensity and other EDS results on NP#0 are shown in Figure 3).

Radial position [nm]	C_{Ag}	Error Bar C_{Ag}
0.0	0.45	0.04
0.5	0.45	0.02
1.0	0.45	0.01
1.5	0.47	0.01
2.0	0.46	0.01
2.5	0.55	0.01
3.0	0.60	0.02

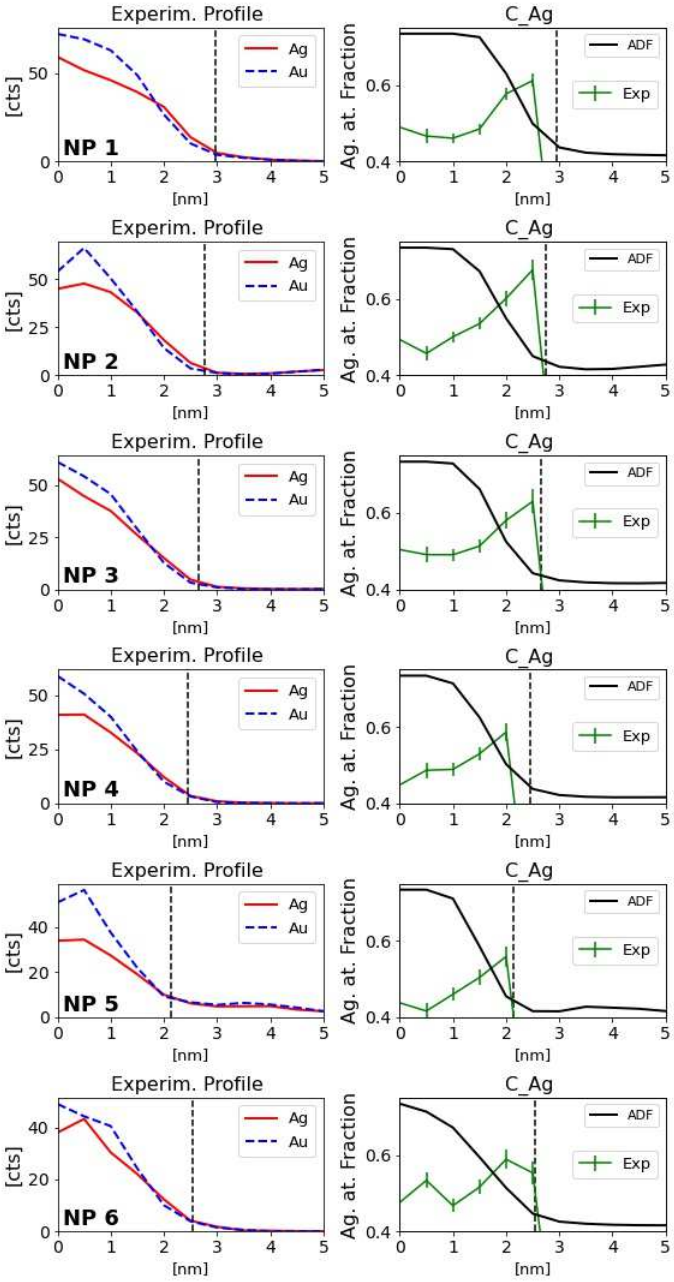

Experim. Profile
C_Ag
Ag
Au
ADF
Exp
[cts]
Ag. at. Fraction
[nm]
NP 1
NP 2
NP 3
NP 4
NP 5
NP 6

Figure S4

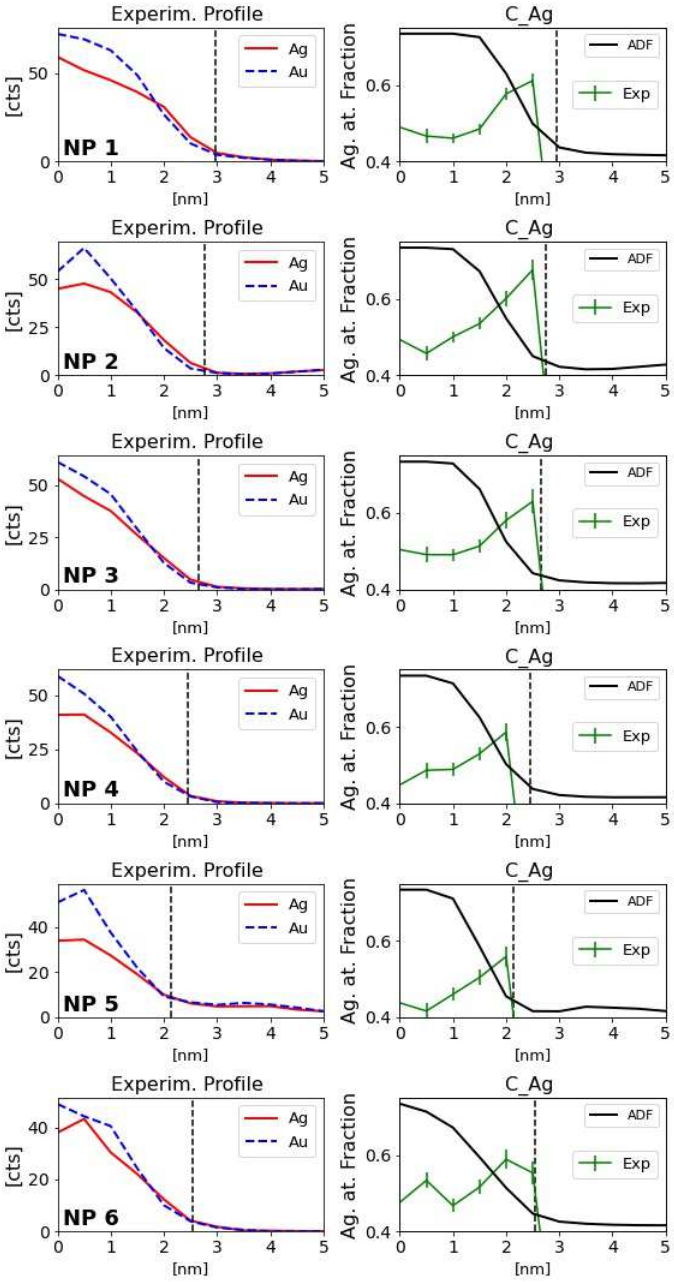

Experim. Profile
C_Ag
Ag
Au
ADF
Exp
[cts]
Ag. at. Fraction
[nm]
NP 1
NP 2
NP 3
NP 4
NP 5
NP 6

Figure S4

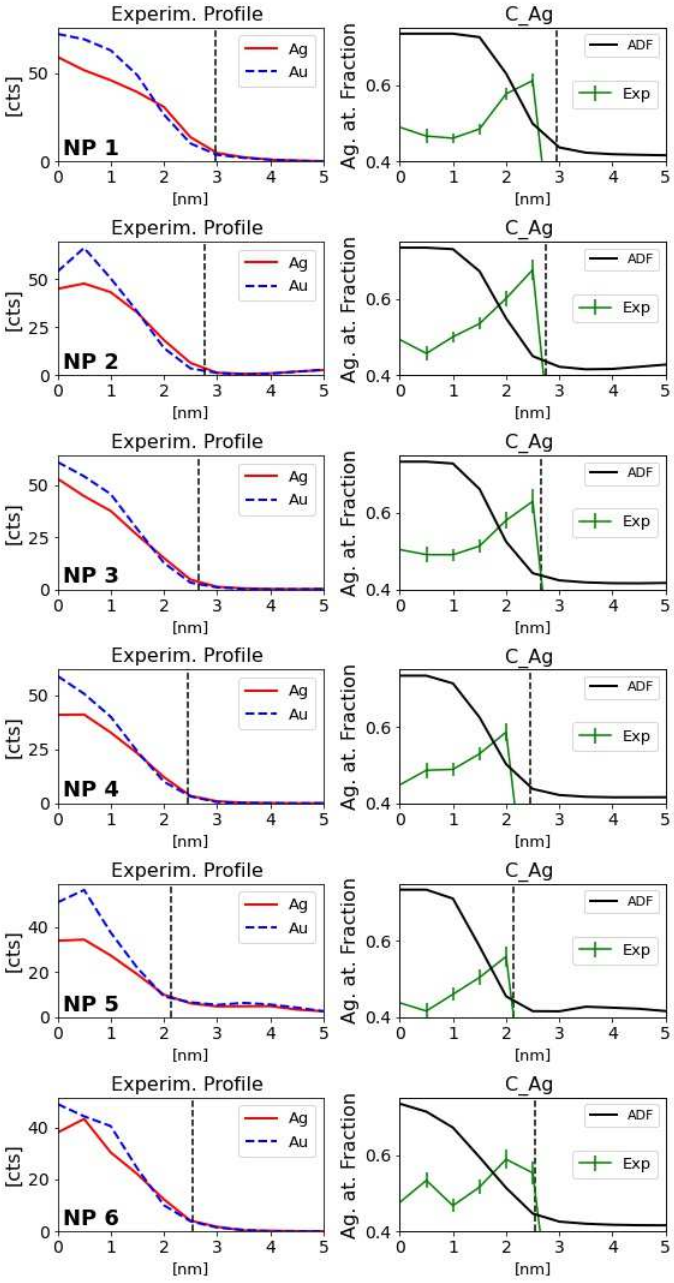


Figure S4: Processing of EDS signal for all NPs detected in the experiment (NP#0 profiles have been displayed in Figure 3, main text). Radial EDS elemental intensities for Au and Ag obtained by the azimuthal integration of X-ray intensity are shown on the right; the radial intensity profiles have been normalized by the number of pixels contributing to each radius. On the left, we display the radial dependency of measured chemical composition C_{Ag}, (including error bars). We note that all NPs show an increase of Ag content towards the surface. The black solid line indicates the azimuthally integrated and normalized radial ADF intensity; the vertical dashed line indicates the estimated NP radius from the analysis of ADF intensity. For larger particles, the ADF intensity is saturated in the central region of the nanoparticle.

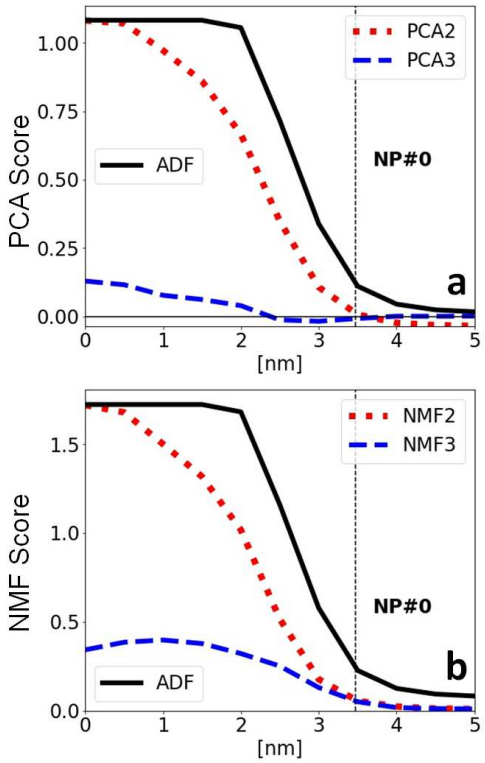


Figure S5: Machine learning results obtained by processing the experimental STEM-EDS HSI; in particular we present results for NP#0 (7.0 ± 0.5 nm in diameter, chemical composition $C_{\mathrm{Ag,\ Exp}} = 0.48 \pm 0.01$). a) comparison of radial profiles of scores from the two PCA components carrying information on the BNP chemical composition; note that PCA#3 score becomes negative close to NP surface, generating an enhancement of Ag content (negative EDS Ag peak in PCA#3, see Figure 4 in main text). b) Similar plot to (a) for the NMF analysis; close to the BNP surface the NMF#2 contribution (this component carries mainly Au counts) decreases drastically. The black solid line indicates the azimuthally integrated and normalized radial ADF intensity; the vertical dashed line indicates the estimated NP radius from the analysis of ADF intensity. Here, the ADF intensity is saturated in the central region of the nanoparticle.

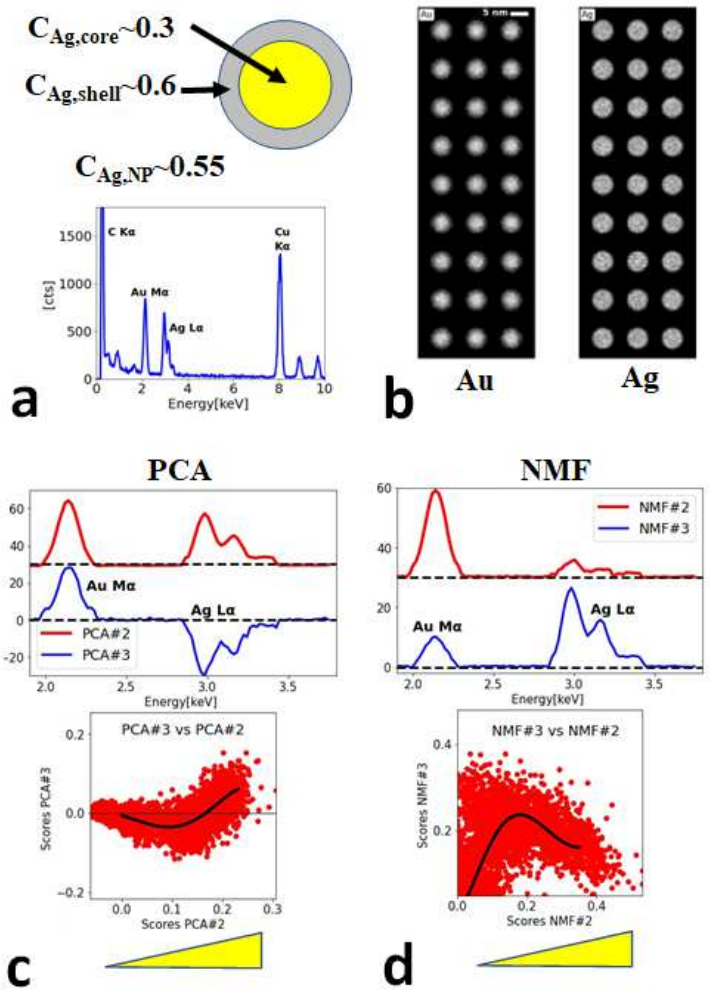


Figure S6: Machine learning results obtained by processing a simulated data set of 27 bimetallic NPs (6 nm in diameter and with a nominal Ag concentration of $C_{\mathrm{Ag}} = 0.55$, ∼3300 EDS counts for each element Au or Ag). a) BNP were assumed to have a core-shell structure with the abrupt interface, where the shell is enriched in Ag; a typical simulated EDS spectrum adding all pixels inside a BNP is displayed at the bottom. b) Map of EDS intensities for Au Mα and Ag Lα intensities; note that NPs display a larger size in the Ag map. c) Comparison of PCA components carrying chemical information on the particles (PCA#2 and PCA#3); scatter plot showing the correlation of components on the simulated HSI. (weight in the reconstruction). d) NMF results comparing components and scores for the simulated HSI. ML results show that both PCA and NMF are able to reveal that pixels with maximum Au signal (corresponding to NP central positions) tend to deplete the contribution of component carrying Ag contribution; conversely, the NP shells contain an enhanced contribution of Ag counts. This is in full agreement with experimental results. Note that simulated HSIs do not include a BNP size distribution, the Au intensity can be more easily related to sample thickness at the analyzed pixel. The black line in the scatter plots (c) and (d) represents a polynomial fit in order to allow an easy eye recognition of the average non-linear relation between the scores.