

# Accurate reconstruction of EBSD datasets by a multimodal data approach using an evolutionary algorithm

Marie-Agathe Charpagne<sup>a,\*</sup>, Florian Strub<sup>b</sup>, Tresa M. Pollock<sup>a</sup>

<sup>a</sup>Materials Department, University of California, Santa Barbara, CA 93106, USA

<sup>b</sup>Univ. Lille, CNRS, Centrale Lille, Inria, UMR 9189 - CRISTAL, F-59000 Lille, France

---

## Abstract

A new method has been developed for the correction of the distortions and/or enhanced phase differentiation in Electron Backscatter Diffraction (EBSD) data. Using a multi-modal data approach, the method uses segmented images of the phase of interest (laths, precipitates, voids, inclusions) on images gathered by backscattered or secondary electrons of the same area as the EBSD map. The proposed approach then search for the best transformation to correct their relative distortions and recombines the data in a new EBSD file. Speckles of the features of interest are first segmented in both the EBSD and image data modes. The speckle extracted from the EBSD data is then meshed, and the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is implemented to distort the mesh until the speckles superimpose. The quality of the matching is quantified via a score that is linked to the number of overlapping pixels in the speckles. The locations of the points of the distorted mesh are compared to those of the initial positions to create pairs of matching points that are used to calculate the polynomial function that describes the distortion the best. This function is then applied to un-distort the EBSD data, and the phase information is inferred using the data of the segmented speckle. Fast and versatile, this method does not require any human annotation and can be applied to large datasets and wide areas. Besides, this method requires very few assumptions concerning the shape of the distortion function. It can be used for the single compensation of the distortions or combined with the phase differentiation. The accuracy of this method is of the order of the pixel size. Some application examples in multiphase materials with feature sizes down to 1  $\mu\text{m}$  are presented, including Ti-6Al-4V Titanium alloy, Rene 65 and additive manufactured Inconel 718 Nickel-base superalloys.

**Keywords:** EBSD, image segmentation, distortions, multi-modal data, CMA-ES

---

## 1. Introduction

### 2. 1.1. Limits of phase differentiation by EBSD

3     Electron Backscattered Diffraction (EBSD) is a powerful tool for gathering local crystallographic data in multi-  
4     phase materials. This technique enables a fast mapping of microstructures with a good angular resolution (typically  
5     0.5°), based on the collection of diffraction patterns and the analysis of Kikuchi bands on those patterns. Initially im-  
6     plemented in scanning electron microscopes (SEM), it has been adapted to transmission microscopy (TEM) under the  
7     name of Transmission Kikuchi Diffraction (TKD) [1, 2]. In the past years, tridimensional EBSD (3D EBSD) by serial  
8     sectioning has also emerged as a new tool to reconstruct and analyze materials in three dimensions [3–8]. To index  
9     the collected patterns, most commercial software use algorithms based on Hough transformations to detect the char-  
10   acteristic bands on the Kikuchi patterns. This process has the advantage of speed but suffers from some limitations.  
11   Because of the diffraction patterns they produce, some phases are intrinsically challenging to index with this method.  
12   This is the case for the finely distributed  $\beta$  phase in  $\alpha - \beta$  Titanium alloys, for example. Another well-known limita-  
13   tion of this method happens in the presence of two phases of similar structures. This is the case of  $\gamma$  and  $\gamma'$  phases in

---

\*corresponding author

Email address: marie.charpagne@engineering.ucsb.edu (Marie-Agathe Charpagne)

14 Nickel and Cobalt-base superalloys. The ordered  $L1_2$   $\gamma'$  phase produces extra very low-contrast bands compared to  
15 the Face Centered Cubic (FCC)  $\gamma$ , that are undetectable by Hough transforms. In the case of phases producing similar  
16 Kikuchi patterns but having significantly different chemistries, one can employ multi-modal data acquisition. Nowell  
17 and Wright [9] have proposed the simultaneous collection of X-ray energy dispersive spectroscopy (XEDS) data, along  
18 with EBSD data. This method has been shown to enable phase differentiation in complex alloys, and reveal unique  
19 crystallographic relationships between phases [10–12]. However, XEDS-EBSD requires a significant difference in the  
20 chemistry of the phases of interest and a minimum number of counts on the XEDS spectra, which can lower the speed  
21 of data collection. The acceleration voltage also has to be high enough to enable the detection of the characteristic  
22 X-rays emission of the elements of interest. Increasing the acceleration voltage would also result in more counts on  
23 the EDS spectra and increase the indexing speed, but is detrimental to the spatial resolution. Indeed, the XEDS inter-  
24 action volume is about one micrometer at 20 kV in Nickel alloys and increases greatly with the acceleration voltage.  
25 Using the different chemical composition of the phases, other methods involving multi-modal data combination have  
26 been implemented. Phases of different composition exhibit different intensities on back-scattered electrons (BSE)  
27 images, due to their different Z number. Payton and Nolze [13] have proposed a method that involves the acquisition  
28 of BSE images on a tilted sample, using specific BSE detectors located above the phosphorous screen of the EBSD  
29 detector. Those authors have shown that this method enables a more accurate phase differentiation than XEDS-EBSD,  
30 since the contrast on the back-scattered images is very sensitive to the variations in chemical composition. They also  
31 demonstrate the capability of that method for segmenting small features precisely, with typical dimensions of the order  
32 of a few micrometers. The limit of this technique is that not all the EBSD detectors are equipped with such built-in  
33 diode sensors. The collection of back-scattered data is usually made by relatively large retractable BSE detectors that  
34 have several quadrants, at a 0° tilt angle. On the other hand, EBSD data is collected at a 70° tilt angle. This, among  
35 other reasons, leads to complex distortions in the EBSD versus the BSE data, which makes their precise recombination  
36 very challenging.

### 37 1.2. Distortions in EBSD data

38 EBSD data is known to suffer from distortions that arise from many instrument and detector artifacts [14]. The  
39 following types of distortions can be distinguished:

- 40 • First order spatial distortions: those are induced when tilting the sample at 70°. An area that would be a square  
41 at 0° tilt turns into a trapezoid if the rotation axis of the stage is not perfectly parallel to the surface of the  
42 sample. The borders of the trapeze are also misaligned with those of the non-tilted area. Those distortions can  
43 be modeled by a first order affine transformation, involving a translation and a rotation.
- 44 • Second order spatial distortions: the electron beam is deflected by a set of lenses that induce a barrel distortion  
45 on the final image. This distortion is mostly visible at low magnification. It can be modeled by a second order  
46 polynomial function.
- 47 • Space-time distortions: when a sample is being exposed for some time to the electron beam, charging effects  
48 occur. The accumulation of charges on the surface of the sample leads to the deflection of the beam during  
49 scanning. This effect occurs in any scanning process. It is usually the most pronounced in the beginning of the  
50 scan and becomes more stable as the scan progresses. A good example is shown in [15]. The amount of drift  
51 is related to the conductivity of the sample itself, the conductivity of the mounting system that is being used,  
52 the voltage and aperture, as well as the cleanliness of the scanned surface. The amount of drift is also more  
53 important at high than low magnifications [16]. There is no simple physical model to describe the pattern of the  
54 drift distortions.

55 Given the complexity of the net resulting distortion, it cannot be easily calculated.

## 56 2. Method

57 In this context, we propose a method that enables the compensation of the distortions in EBSD data and adds  
58 phase differentiation if wanted, by accurately matching it with BSE images collected at a 0° tilt angle, at individual  
59 pixel precision and over broad areas. To do so, the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is  
60 used to calculate the relative distortion function between the EBSD data and the BSE image.

61 **2.1. Working with speckles**

62 With regard to the recombination of multi-modal data, one first intuitive approach is to use sets of matching points  
 63 over the area of interest in complementary images. Some methods aiming at correcting the distortions in EBSD  
 64 maps have been implemented using this principle [15]. They require the manual selection of a set of reference and  
 65 corresponding points in the BSE image and the EBSD data. The higher the number of pairs of points, the greater the  
 66 precision. However, this step can be time-consuming and does or made difficult by the lack of contrasted features  
 67 in both EBSD and BSE data. Furthermore, it does not scale up to large or tridimensional datasets, as it may require  
 68 human annotations for every slice in the dataset. In most 3D EBSD datasets, the process of data collection consists in  
 69 successive movements of the stage to positions that are practical for machining, imaging, and collection of diffraction  
 70 patterns. The placement of the stage from a slice to another is never exactly the same [17]. This modifies the set of  
 71 reference points on every slice.

72 Here, we propose to use speckles of similar features without any human annotations, allowing the recombination of  
 73 multi-modal data to scale up to large EBSD datasets. To do so, a high number of automatically generated reference  
 74 points are combined by using microstructural features as speckles, such as pores, a second-phase, or precipitates.  
 75 Since the goodness of the superimposition of the speckles can be quantified, it can be used as a parameter to maximize,  
 76 enabling an automated matching process. Therefore, quantifying the goodness of the superposition of the speckles is  
 77 critical.

78 **Defining a similarity measure between the speckles**

In order to superpose the EBSD and BSE speckles, a quantitative measurement of their similarity is needed. The Dice similarity metric [18] (also known as F1-score) describes the relative overlap between the segmented images. More precisely, this score measures the overlap between segmented pixels. Formally, given the binary EBSD speckle  $\mathcal{I}^{EBSD} \in \{0, 1\}^{I \times J}$  and the binary BSE image  $\mathcal{I}^{BSE} \in \{0, 1\}^{J \times J}$  where 1 encodes the segmented object and 0 the background image, and  $I \times J$  their domain of definition, the Dice similarity for is computed as follows:

$$\text{Similarity}(\mathcal{I}^{EBSD}, \mathcal{I}^{BSE}) = 2 \frac{|\mathcal{I}_{seg}^{EBSD} \cap \mathcal{I}_{seg}^{BSE}|}{|\mathcal{I}_{seg}^{EBSD}| + |\mathcal{I}_{seg}^{BSE}|} \quad (1)$$

79 where  $|\cdot|$  encodes the size of the underlying set and  $\mathcal{I}_{seg}$  only corresponds to the segmented pixels in the image. Thus,  
 80  $|\mathcal{I}_{seg}^{EBSD} \cap \mathcal{I}_{seg}^{BSE}|$  corresponds to the number of matching pixels and  $|\mathcal{I}_{seg}^{EBSD}|$  and  $|\mathcal{I}_{seg}^{BSE}|$  respectively correspond to the  
 81 total number of pixels in the EBSD and BSE images. This similarity measure counts the number of overlapping  
 82 segmented pixels normalized by the sum of segmented pixels. It ranges from 0 (no overlapping segmented pixels) to  
 83 1 (perfect matching). The value of the  $\text{Similarity function}(\mathcal{I}^{EBSD}, \mathcal{I}^{BSE})$  is referred to as the 'score' in the following.  
 84 As the Dice similarity is both intuitive, robust and easy to compute, it has been extensively used in several image  
 85 segmentation applications such as medicine [19].

86 **2.2. Correction work-flow**

The global process of the reconstruction is shown in figure 1. Starting from the initial EBSD data and the BSE image, two speckles are generated. In the first step (initial alignment), the BSE speckle is rescaled so its pixel size corresponds to that of the EBSD data. The rescaling is done using a nearest neighbor interpolation. The BSE speckle image is also rotated and translated so it is pre-aligned to the EBSD data. To do so, a grid-search over a set of affine transformations is performed, to distort the re-sized speckle over three parameters: two translations along the x and y directions,  $t_x, t_y$ , and a rotation of angle  $\alpha$ . The user is free to choose the widths of the translation and the rotation angle, as well as their increment size. The minimum increment for the translations is one pixel. The step size for the angle can be refined as much as wanted. All the possible translations and rotations within the provided ranges are tested. The transformation leading to the best score is saved and applied to the BSE speckle. The latter is then used in the second step, as the reference speckle for the compensation of the distortions. In the second step (CMA-ES on fig. 1), the EBSD speckle is meshed on a regular grid, and the grid is distorted using the CMA-ES optimizer, in an iterative process. The initial and new locations of the points of the grid are used as pairs of control points, to estimate the distortion function. At each iteration, the score produced by the distorted EBSD speckle is calculated. The distortion function corresponding to the distorted speckle that exhibits the best score is used to determine the new location of the points in the new EBSD file and the segmented BSE image is used to fill in the phase data (if needed). The parametrization and calculation of the distortion function are explained in the following.

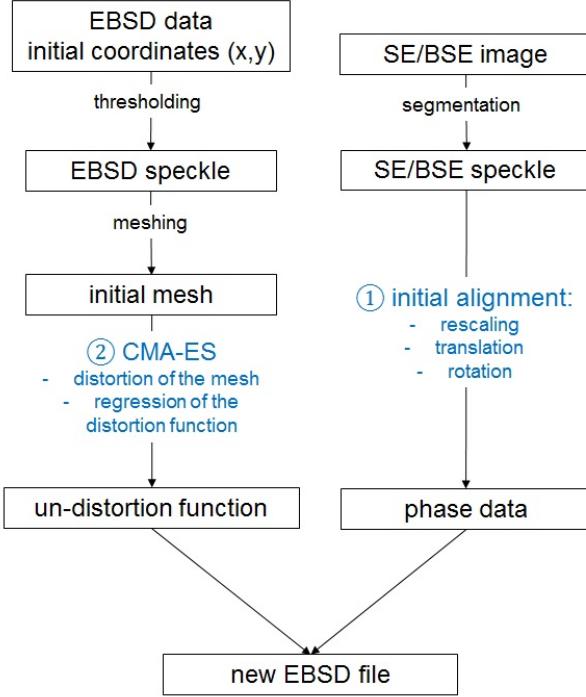


Figure 1: Schematic representation of the reconstruction algorithm.

103 2.3. *The distortion function*

Formally, the goal is to find the distortion function  $f$  that maximizes the similarity between the EBSD and the BSE speckles. This leads to the optimization procedure described in eq. 3:

$$f^* = \underset{f}{\operatorname{argmax}} (\text{Similarity}(\mathcal{I}^{EBSD}, \mathcal{I}^{BSE})) \quad (2)$$

$$= \underset{f}{\operatorname{argmax}} \left( 2 \frac{|f(\mathcal{I}^{EBSD})_{seg} \cap \mathcal{I}^{BSE}_{seg}|}{|f(\mathcal{I}^{EBSD})_{seg}| + |\mathcal{I}^{BSE}_{seg}|} \right) \quad (3)$$

104 As described in section 1, many overlapping distortion phenomena occur while collecting the EBSD data, leading  
 105 to a non-linear distortion. In the following, the nature of the distortion function is assumed to be polynomial. It  
 106 is a very weak assumption as polynomial functions can approximate a large range of complex functions. Besides,  
 107 the proposed method can be easily extended to other function approximators such as Gaussian processes, trees etc.  
 108 Formally, a polynomial distortion of degree  $P$  is defined as follows:

$$f(x, y) = \begin{cases} x' = \sum_{n=0}^P \sum_k c_{k,n-k}^x x^k y^{n-k} \\ y' = \sum_{n=0}^P \sum_k c_{k,n-k}^y x^k y^{n-k} \end{cases} \quad (4)$$

109 For instance, given a polynomial fit of degree 2, the coordinates  $(x, y)$  of a pixel are transformed into  $(x', y')$ :

$$f(x, y) = \begin{cases} x' = c_{0,0}^x + c_{1,0}^x x + c_{0,1}^x y + c_{1,1}^x xy + c_{2,0}^x x^2 + c_{0,2}^x y^2 \\ y' = c_{0,0}^y + c_{1,0}^y x + c_{0,1}^y y + c_{1,1}^y xy + c_{2,0}^y x^2 + c_{0,2}^y y^2 \end{cases} \quad (5)$$

110 Given a polynomial order, the goal is to find the optimal set of weights  $c^*$  that maximizes the similarity measure.  
 111 This set of weights  $c$  can be computed by two means: by directly looking into the space of parameters  $c$  or by  
 112 generating intermediate matching points that are used to perform a polynomial regression.

113 While the first approach may be more direct, the manifold of parameters  $c$  is very difficult to explore as a small  
 114 change of parameters can lead to drastically different distortions. Therefore, in the proposed method, a set of matching  
 115 points is generated, that is used to regress the polynomial coefficients  $c$ . As a small change in the location of  
 116 the matching points lead to a close distortion, this process allows a better granularity in exploring the space of the  
 117 distortions. More precisely, a regular grid and a distorted one are generated, and then the coefficients  $c$  are determined  
 118 by a least-squares fitting method that maps the distorted grid to the regular one. This function is described below:

- 119 • Generate a distorted mesh grid  $M = ((x_0, y_0), \dots, (x_N, y_N))$
- 120 • Generate a regular mesh grid  $M' = ((x'_0, y'_0), \dots, (x'_N, y'_N))$
- 121 • Find the distortion  $f$  parametrized by  $c$  such as  $c = \operatorname{argmin}_{\hat{c}} \sum_i^N \|(\frac{x'_i}{y'_i}) - f_{\hat{c}}(x_i, y_i)\|_2$

122 where,  $\|\cdot\|$  is the euclidean norm. Note that the dimension of the mesh grid is tuned by the user via the number  
 123 of points  $N$  and/or the step-size between points. The influence of that parameter is discussed in section 4. As a  
 124 result, finding the optimal distortion  $f^*$  requires to find the optimal set of matching points  $M^*$  that better describes the  
 125 underlying distortion. These matching points are then used to estimate the distortion itself. Counter-intuitively, the  
 126 distorted grid may not match physical reality as discussed further in section 4.4.

127 Finally, the *polynomial distortion* function is called from the python package *Sklearn* [20] and *Skimage* [21]. The  
 128 next section describes how the pairs of matching points are created, using the CMA-ES optimizer.

#### 129 2.4. Using CMA-ES as a randomized black-box optimization to correct the distortions

130 The goal of the optimization is to find the function  $f$  that maximizes the *Similarity* function between the speckles  
 131 by finding the best distorted mesh grid  $M$  (which is non-differentiable). To do so, a black-box optimizer method is  
 132 used. Black-box optimization algorithms are efficient at solving non-linear, non-convex optimization problems, in  
 133 continuous domains. They also overcome the deficiencies of the derivative-based methods in complex multidimen-  
 134 sional landscapes that are rugged, noisy, have outliers or local optima, or when no error-gradient is available. All  
 135 those features make the black-box optimizer relevant for the present problem, where the shape of the distortion is  
 136 unknown and the process of superposing speckles intrinsically implies the *Similarity* function to reach local maxima.  
 137 The influence of noise present on the EBSD speckle is discussed in section 4.5.

138 The black-box algorithm that is used in the present case is CMA-ES, which stands for Covariance Matrix Adap-  
 139 tation Evolutionary Strategy. CMA-ES has become a standard tool for continuous optimization and has been applied  
 140 in various fields of research, such as image recognition for biology [22, 23], energy [24], chemistry [25, 26]. Though,  
 141 to the authors knowledge, it has not been applied in the field of electron microscopy. CMA-ES belongs to the family  
 142 of evolutionary strategies (ESs), they are iterative algorithms, based on the principles of natural selection. CMA-ES  
 143 involves a parametrized distribution (a multivariate normal distribution) that evolves throughout the iterations [27].  
 144 ESs follow four distinct steps: initialization, sampling, evaluation and update. The initialization step consists in ini-  
 145 tializing the probability distribution parameters and sampling an initial population of individuals accordingly. At each  
 146 iteration, a new *generation* of individuals is created. Those individuals are candidate solutions to the optimization  
 147 problem, which goodness can be evaluated on a *fitness function*. After evaluation of the fitness of each individual in  
 148 that population, the statistics of the distribution are updated according to the algorithm at hand. A new population is  
 149 re-sampled according the new distribution and the process is repeated until a termination criterion is reached. Those  
 150 operations are described below:

151 Initialize the distribution parameters  $\theta$

152 For each generation  $g = 0, 1, \dots$  *number of iterations*:

153     Sample  $\lambda$  individuals  $x_1, x_2, \dots, x_\lambda \in \mathbb{R}^n$  according the probability distribution  $P_\theta(x)$

154     Evaluate them on the fitness function

155     Update  $\theta \leftarrow F(\theta, x_1, x_2, \dots, x_\lambda, \text{Similarity}(x_1), \text{Similarity}(x_2), \dots, \text{Similarity}(x_\lambda))$

156 where  $P_\theta$  is a probability distribution that describes where the good solutions are believed to be and  $F(\cdot)$  is some  
 157 update rule. In the CMA-ES,  $P_\theta$  is a multivariate normal distribution [27]. Those steps are detailed in the following.

158 *Initialization.* The distribution  $P_\theta$  encodes the distribution of the distorted spatial mesh grid. More precisely, the dis-  
 159 tribution is a multi-variate normal distribution where each dimension encodes either the x-coordinate or y-coordinate  
 160 of a given point of the grid. For example, if the mesh is a 4x4 grid, the multi-variate normal distribution has 32 di-  
 161 mensions (16 points of 2 coordinates each). The initial distribution is defined by centering the mesh distribution on a  
 162 regular grid. The standard deviation then encodes the initial acceptable moving distance of the matching points. Both  
 163 the spacing of the points of the grid and the standard deviation are chosen by the user. In practise, CMA-ES slowly  
 164 and iteratively distorts the regular grid, in order to retrieve a set of control (matching) points that encode the distortion.  
 165 The concept of mesh grid here is to understand as an array of points that do not have any topological connectivity. In  
 166 other words, the permutation of two points of the grid is allowed and does not affect the regression of the distortion  
 167 function. The properties of the mesh are discussed in more detail in section 4.4.

168 *Sampling.* New distorted meshes are generated following the distribution  $P_\theta$ . The coordinates are rounded, and  
 169 matching points outside the range of the speckle dimension are kept.

170 *Evaluation.* The fitness function is the *Similarity* described in eq. 1. The coefficients  $c$  of the distortion function are  
 171 first regressed by matching the distorted mesh grid to the regular one. The similarity measure is then computed and  
 172 returned as the fitness score.

*Update.* The basic CMA equation for sampling the search points at the generation  $g + 1$  is given in eq. 6.

$$x_k^{(g+1)} \leftarrow m^{(g)} + \sigma^{(g)} \mathcal{N}(0, C^{(g)}), \text{ for } k = 1, 2, \dots, \lambda. \quad (6)$$

173 where  $x_k^{(g+1)} \in \mathbb{R}^n$  is the  $k$ -th offspring of the generation  $g + 1$ .  $m^g \in \mathbb{R}^n$  is the mean value of the distribution at  
 174 the generation  $g$ .  $\sigma^{(g)} \in \mathbb{R}^+$  is the step size at the generation  $g$ .  $\mathcal{N}(0, C^{(g)})$  is a multivariate normal distribution with  
 175 a mean of zero and covariance matrix  $C^g$ .  $C^g \in \mathbb{R}^{n \times n}$  is the covariance matrix of the distribution at the generation  
 176  $g$ . It is symmetric definite positive and describes the geometrical shape of the distribution. The initial value of  $\sigma$ ,  
 177  $\sigma_0$ , is picked by the user and  $C_0 = I$ . They both evolve throughout the iterations, as the population evolves. The  
 178 self-adaptation of those parameters is the key point for the rapid convergence of the optimization[28]. The equations  
 179 governing the update of  $\sigma^{(g)}$ ,  $C^{(g)}$  and  $m^g$  are described in Appendix A. In the present case, this update shifts the  
 180 spatial distribution of the points of the distorted grid in order to search for the optimal distortion function.

181 *Termination.* This process is repeated until the termination criterion is reached. In the present case, a criterion based  
 182 on the maximum number of steps is used. More advanced methods can be used to detect when the optimization (and  
 183 thus the fitness score) starts plateauing. If several speckles are known to have the same distortion, they also can be used  
 184 as a validation criterion to avoid over-fitting. Once the CMA-ES algorithm is finished, the means of the distribution  
 185  $P_\theta$  are used as the final distorted mesh grid.

## 186 2.5. Generation of the new EBSD map

187 Once the maximum number of calls has been reached, the polynomial function leading to the best superposition  
 188 of the speckles is applied to the EBSD data. In this process, the original grid of the EBSD is kept and the Euler angles,  
 189 confidence indexes and other data is interpolated using the polynomial function. The points of the grid that end up  
 190 containing no data are filled with zero-values for the confidence indexes and the (0, 0, 0) triplet of Euler angles. The  
 191 rotated and translated segmented BSE image is used to fill-in the phase data of the new EBSD file. For that reason, and  
 192 as already mentioned, the segmentation of the BSE image has to be done as accurately as possible. Some application  
 193 examples are shown in section 3.

## 194 2.6. Source code and data availability

195 The full distortion pipeline was developed in Python. The code source, the hyperparameters and one data set are  
 196 available at <https://github.com/MLmicroscopy>.

197 **3. Results**

198 All materials have been characterized in a FEI Versa 3D SEM, equipped with a TSL EDAX EBSD Hikari Plus  
199 Camera with an indexing speed of 300 to 400 frames per second and a 4x4 binning. The computer for the distortion  
200 correction was equipped with a Quad Core processor, 4.2 GHz and 64 GB of RAM. Most computations required about  
201 10 to 15 minutes.

202 *3.1.  $\beta$  phase in  $\alpha - \beta$  Titanium alloy*

203 A sample of Ti-6Al-4V with an  $\alpha - \beta$  structure has been embedded in bakelite and electro-chemically polished.  
204 It was then mounted on the SEM stage and copper tape was used to enable electrical conductivity from the sample  
205 holder to the polished surface. An area of  $14 \times 20 \mu\text{m}$  has been characterized with a 20 kV acceleration voltage and a  
206 step size of 40 nm. One purpose of collecting this dataset is to better resolve the phase map, using the better quality  
207 of the segmentation of the BSE image, with fine details. This mounting system was purposely used in order to induce  
208 a large drift of the beam during the EBSD scan.

209 The orientation map for this sample is shown on fig. 2-a. The corresponding BSE image is shown on fig. 2-d,  
210 where the  $\beta$  phase exhibits a lighter contrast than the  $\alpha$  phase. Fig. 2-b shows the initial phase map with the  $\beta$  phase  
211 in white. This phase map is superimposed to the Index Quality map on fig. 2-c, where the  $\beta$  phase is colored in red.  
212 The comparison of fig. 2-b and -e shows that only a small fraction of the  $\beta$  phase was indeed identified as such on the  
213 initial EBSD data. Its initial area fraction is 3.4%, versus 9.9% in the segmented BSE image. The CMA-ES optimizer  
214 has been applied on a  $25 \times 25$  points mesh grid, with an initial standard deviation  $\sigma_0 = 20$  pixels, using a polynomial  
215 order of 3 and 5,000 iterations. Fig. 2-f shows the final phase map, colored as fig. 2-c. It superimposes well with  
216 the Index Quality map. The final area fraction of  $\beta$  phase is 9.9%, as in the BSE speckle. The contribution of the  
217 CMA-ES optimizer to compensate higher order distortions and match all the features of the speckles is discussed in  
218 section 4.1.

220 *3.2.  $\gamma$  and  $\gamma'$  phases in Rene 65 superalloy*

221 Rene 65 is a polycrystalline Nickel-based alloy that has been designed for turbine disk applications. Its chemical  
222 composition (wt%) is Ni-16 Cr 13 Co 3.7 Ti 2.1 Al 4 Mo 4 W 1 Fe 0.7 Nb 0.05 Zr 0.016 B [29]. The sample was  
223 extracted from a ring that was subjected to a sequence of forging operations followed by solution annealing and aging  
224 treatments, 1065°C for one hour and 760°C for 8 hours respectively. It was prepared using conventional metallography  
225 techniques, followed by vibratory polishing using a  $0.04 \mu\text{m}$   $\text{Al}_2\text{O}_3$  suspension. Its microstructure consists of fine  $\gamma$   
226 matrix grains, which average equivalent diameter is  $10-12 \mu\text{m}$ . Spherical primary  $\gamma'$  precipitates are located on the  
227 grain boundaries, with an equivalent diameter in the range of  $1-4 \mu\text{m}$ . EBSD maps and corresponding BSE images  
228 have been acquired using an acceleration voltage of 20 kV with a step size of  $0.1 \mu\text{m}$  over an area of  $150 \times 200 \mu\text{m}$ .  
229 Figure 3-a shows the initial EBSD map colored according to the orientation of the crystals projected along the normal  
230 to the polished surface. Figure 3-b shows the corresponding BSE image on which the precipitates exhibit a darker  
231 contrast. Figure 3-c and -d show the BSE and EBSD speckles, respectively, created by segmenting the  $\gamma'$  precipitates.  
232 Fig. 3-c has been created by segmenting the features which equivalent diameter is smaller than  $5 \mu\text{m}$ , considering the  
233 annealing twin boundaries as internal defects. This criterion enables a rough segmentation of the precipitates. Note  
234 that some precipitates may have been omitted from the segmentation and some small grains may have been mistakenly  
235 segmented as precipitates. The influence of the quality of the EBSD speckle on the goodness of the compensation of  
236 the distortions is discussed in section 4.5. This segmented image was then used as the EBSD speckle for the CMA-ES  
237 optimizer, along with a segmented BSE image of the same area. The CMA-ES optimizer was used on a mesh grid  
238 with a spacing of 24 pixels between consecutive points, an initial standard deviation of 5 pixels and a polynomial order  
239 of 3. The initial grid is presented on figure 3-e, the final grid is presented on figure 3-f. The superimposed speckles  
240 after applying the CMA-ES optimizer are shown on fig. 3-g. The distortion correction leads to a final score of 0.62  
241 after 1,800 iterations. Figure 3-h shows the final phase map with the precipitates colored in red, superimposed with  
242 the Image Quality map displayed in grey scale. The location of the precipitates corresponds well with the location  
243 of the grain boundaries on the Image Quality map, over most of the area of interest. The precision of the method is  
244 discussed in sections 4.6 and 4.4.

245    3.3. Pores in Additive Manufactured Inconel 718

246    The alloy Inconel 718 has been manufactured by the Electron Beam Melting (EBM) technique, as described in  
247 [30]. A sample was cut by Electrical Discharge Machining (EDM) and polished parallel the build direction, using  
248 conventional metallography processes followed by  $0.05 \mu\text{m}$  colloidal silica. Image and EBSD data have been acquired  
249 under a 20 kV acceleration voltage, using a  $0.5 \mu\text{m}$  step size over an area of  $300 \times 700 \mu\text{m}$ . Fig. 4-a shows the  
250 Secondary Electrons (SE) image that was used for the SE speckle. The beam contamination zone corresponding  
251 to the area of the EBSD map exhibits a trapezoid shape. A standard cleaning procedure involving slight cleaning  
252 of the EBSD data has been applied, consisting in removing isolated groups of less than 3 pixels having the same  
253 orientation, leading to a change of 1% of the data. Since the pores could not be indexed, the TSL software assigned  
254 random orientations to those pixels. The inverse pole figure map, projected along the build direction, is shown on  
255 fig. 4-b. Some of the grains have their  $<001>$  axis preferentially aligned with the build direction (vertical), a  
256 common feature of additive manufactured face-centered cubic materials. A partial EBSD speckle of the pores has  
257 been created by segmenting the pixels of lowest Confidence Indexes (CI). On the other hand, the pores have been  
258 accurately segmented from a SE image of the same area, in order to create the SE speckle. The CMA-ES strategy has  
259 been applied, using a  $65 \times 65$  points mesh grid, an initial standard deviation  $\sigma_0 = 20$  pixels, a maximum sampling  
260 of 5,000 iterations and a polynomial order of 3 for the distortion function. Fig. 4-c shows the superimposed speckles  
261 after the CMA-ES strategy. The EBSD speckle consists of light blue points on a transparent background, and the  
262 SE speckle consists of dark red points on a light red background. Thus, as shown in the insert on the figure: light  
263 red pixels (labeled "1" on the insert) correspond to the SE background, dark red pixels (labeled "2") correspond to  
264 pores of the SE speckle that did not match any pores of the EBSD speckle, light blue pixels (labeled "3") are pores of  
265 the EBSD speckle that did not match the SE speckle dark blue pixels (labeled "4") are pores that were successfully  
266 matched.<sup>1</sup> From this basis, the final EBSD map is shown on fig. 4-d. Only the points with a confidence index higher  
267 than 0.2 are plotted. The color code is the same as fig. 4-b and the pores are colored in black. Most of those pores  
268 are located on the grain and sub-grain boundaries. After applying the distortion, some data points ended up out of the  
269 initial grid, which leads to the black pixels on the edges of fig. 4-d. This dataset illustrates the capacity of this method  
270 to compensate the distortions over broad areas while still matching small features with a good precision.

271    4. Discussion

272    4.1. Versatility of the method

273    Several methods have been proposed in the literature, for correcting the distortions in EBSD data, using various  
274 algorithms to make up for the drift distortions. Zhang et al. [15] have shown that a thin plate spline function enables  
275 compensation of the distortions in EBSD data. Contrary to most methods, the use of the CMA-ES strategy does not  
276 assume any specific shape of the distortion function, nor constrain the order of the polynomial function. The use of  
277 speckles and a score to quantify the goodness of the superposition enables the matching of both images at the resolu-  
278 tion of the pixel. Fig. 5 shows a comparison between the superposition of the speckles in the Ti6Al4V dataset, after  
279 the compensation of affine distortions only (a) and after the CMA-ES procedure (b). On a blue background, the EBSD  
280 speckle is colored in red and the BSE speckle in white. Compensating the affine distortions enables a good overlap of  
281 the speckles. The CMA-ES optimizer then compensates the finer distortions, even very local, and enables to achieve  
282 a better matching of the speckles despite an important drift of the beam. Three examples are presented in the inserts  
283 of Fig. 5, where the laths are accurately superimposed.

284    Figure 6 shows the maps of the vertical and horizontal components of the distortion function in the Inconel 718  
285 material. The origin and x and y directions of the scan are indicated with arrows. The electron beam scanning pattern  
286 starts from the top left corner and goes line-by-line from top to bottom. A polynomial order of 3 has been used. Both  
287 maps show that the maximum error occurs on the first points of the scan. Indeed, drift phenomena are the highest in  
288 the beginning of the scan. The magnitude of both components then decrease along both x and y directions. Along  
289 the x direction,  $f_x$  decreases relatively more than  $f_y$ , which means that the error on the x' coordinate decreases along

1using a method based on the CI to segment the pores can lead to inaccuracies in the segmentation: some pixels may have been segmented as pores even though they are not (typically those pixels correspond to grain boundaries).

291 this direction, whereas the error on the y' coordinate remains relatively constant. From a line to another (along the y  
 292 direction), the magnitude of  $f_x$  decreases, which means that the error on the location of the x' coordinate decreases  
 293 from a line to another. A similar observation can be made about the vertical component  $f_y$ . Those trends are consistent  
 294 with most drift phenomena, where the first scanned pixels that are the most subjected to drift. Drift phenomena tend  
 295 to decrease as the scan proceeds: along a line, and from a line to another. A quantitative analysis of those trends can  
 296 be found in Appendix B.

#### 297 4.2. Sensitivity to the choice of the initial parameters in CMA-ES

298 The influence of the CMA-ES parameters on the goodness of the matching has been studied in the Ti-6Al-4V  
 299 dataset. Figure 7 shows the final scores obtained for different numbers of points per width and height of the image,  
 300 and initial step sizes. The obtained surface is convex, with a global maximum reached for (35,21), pointed by the red  
 301 arrow on the figure. For a fixed number of grid points, the score admits a maximum for a given initial step size  $\sigma_0$ .  
 302 This behavior is visible on the four points on the far right of the figure, on the part of the envelope highlighted by a  
 303 blue curve. For 55 grid points per length of the image, the score reaches a maximum of 0.5128 for an initial step size  
 304 of 27 pixels. A similar behavior occurs for any given initial step size but varying numbers of points in the mesh grid.  
 305 Note that a similar behavior is expected for all data sets, but for different values of the parameters.

#### 306 4.3. Precision of the reconstruction

307 The accuracy of the reconstruction of the phases can be quantified by comparing the location of the phase bound-  
 308 aries and grain boundaries on the reconstructed maps. Such measurements have been performed on the Rene 65  
 309 dataset and displayed on fig. 8. A subset of the corrected EBSD map has been selected where three line profiles  
 310 have been traced, crossing isolated or clusters of  $\gamma'$  precipitates. The point-to-point misorientation is plotted in black  
 311 and superimposed to the phase data, which is plotted in red. The first profile crosses a single precipitate. The phase  
 312 boundaries and grain boundaries "a" and "b" are indicated on the corresponding profile. The transition from a phase  
 313 to the other corresponds well with the crossing of high-angle boundaries. The second profile crosses a cluster of two  
 314 precipitates which were separated on the segmentation of the BSE image. The phase boundaries "c", "d" and "f" are  
 315 well superposed with the grain boundaries. Note that the boundary "d" consists of a single pixel. The profile also  
 316 crosses an annealing twin boundary ( $\Sigma_3$ , 60° rotation around a  $<111>$  axis) present in one of the precipitates at  
 317 the marker "e", characterized by a 60° misorientation. The third profile also crosses two precipitates which were not  
 318 separated on the segmented BSE image. The grain boundaries of the first precipitate are labeled "g" and "h" on the  
 319 profile. The second precipitate contains an annealing twin boundary "i", that is crossed right before the grain bound-  
 320 ary "j". This grain boundary is well superimposed with the phase boundary. The error on the location of the phase  
 321 boundaries has been measured on all the profiles, by comparing their location  $x_{Ph.B}$  to that of the grain boundaries  
 322  $x_{GB}$  through the parameter  $\Delta x$  of eq. 7. The results are displayed in table 1. The error is less than 0.2  $\mu m$  for all the  
 323 boundaries (i.e. 2 pixels), reaching  $\approx 0 \mu m$  for most of them. The error on the location of the two phase boundaries  
 324 that were not exactly correlated with the grain boundaries ("f" and "g") corresponds to a misplacement of 1 and 2  
 325 pixels respectively. This gives an indication of the precision achievable with this correction method, assuming a good  
 326 segmentation of the BSE image is achieved. A segmentation in which the boundaries of the objects have been eroded  
 327 or blurred impacts the quality of the reconstruction.

$$\Delta x = \|x_{Ph.B} - x_{GB}\| \quad (7)$$

#### 328 4.4. Convergence, stability and repeatability of the CMA-ES optimization

329 During the CMA-ES optimization procedure, new distorted mesh grids are generated to regress the polynomial  
 330 distortion. Counter-intuitively, these grids are not constrained to encode a realistic distortion; they are only optimized  
 331 such that the final distortion is meaningful. Thus, one must not use the distorted mesh grids outside the polynomial  
 332 regression step, as several different meshe grids can lead to the same distortion function. Indeed, there exists an  
 333 infinite number of sets of points that extrapolate to the same final function. Those points do not have to always be  
 334 on the curve itself. Even though phenomena of grid points swapping does not affect the calculation of the distortion

Table 1: Precision of the location of the phase boundaries versus grain boundaries in the profiles of Fig. 8, in the Rene 65 dataset

Boundary	$\Delta x (\mu m)$
a	0.0
b	0.0
c	0.0
d	0.0
f	0.1
g	0.2
j	0.0

function, the authors suggest that the user sets the  $\sigma_0$  parameter as less than the distance between two consecutive grid points, if the user wants to avoid this phenomenon. To a lesser extent, several polynomial coefficients may also lead to a very similar resulting distortion function, only differing around the border of the speckle.

Despite this apparent limitation, CMA-ES turns out to be highly reproducible in practice. For example, fig. 9 shows the typical pattern of convergence of the CMA-ES algorithm: fig. 9-a shows the evolution of the score as a function of the number of iterations in the CMA optimizer and 9-b shows the evolution of the step size throughout the iterations. Starting from a score value of about 0.4, the CMA optimizer enables a quick convergence towards a steady-state value of about 0.7. This steady state is reached after less than 5000 iterations. As the score increases, the step-size  $\sigma$  decreases with the number of iterations, indicating little variation around the mesh grid ground.

The CMA-ES strategy has been applied 100 times on the Rene 65 dataset. Each run consisted in 2000 iterations, with a step size of 75 points, an initial standard-deviation  $\sigma_0$  of 5 pixels. A polynomial order of 3 was used for the polynomial function. The CMA optimizer produced an average score of 0.6587 with an associated standard deviation of 0.0015. The similarity score between distorted speckles over the 100 runs was 0.9353 in average, indicating a good precision and repeatability of the process. Figure 10-a shows the evolution of the score throughout the iterations: the dark blue curve corresponds to the mean score over 100 runs. The light blue colored area around the curve corresponds to the lower and upper bounds of the score (mean - standard deviation, and mean + standard deviation, respectively). Those values show that the optimization is stable and repeatable over several runs. Some steps are clearly visible on the first 500 iterations and correspond to the new bounds determined by the creation of new individuals at each generation in CMA-ES. Based on those statistics, a heat map has also been generated, displayed on fig. 10-b. This map shows the segmented BSE speckle colored according to the number fraction of times that a pixel was present at a given ( $x'$ ,  $y'$ ) location. In other words, a pixel that appears in white (value 1.0) was always assigned to the  $\gamma'$  phase. On the opposite, a pixel that appears in black was never assigned to this phase. The more consistent and precise the optimization, the sharper the contrast on this map. The shape of the precipitates appears clearly on fig. 10-b, which is consistent with the good repeatability suggested by fig. 10-a. This map also gives information about the precision of the reconstruction over the whole map. The center of the precipitates usually appears with a value of 1.0, however their boundaries are usually less precise. The deviation on the reconstruction of the boundaries of the precipitates is illustrated on two examples, labeled "A" and "B". The precipitate "A" is close to the border of the image, and the precipitate "B" is in the center. The inserts on the side of the figure show that the boundaries of those two precipitates are not reconstructed with the same consistency over the 100 slices. There is only one layer of pixels having a value lower than 1.0 on the precipitate B, versus 2 to 3 rows on the boundary of the precipitate A. The boundaries of most of the precipitates located on the edges of the map have a similar coloring. This indicates that the CMA-ES optimizer leads to consistent results in the center of the map, with a deviation of about 1 pixel as the location of the phase boundaries are approached. At the borders of the map, the consistency on the location of the phase map is about 2 to 3 pixels. This corresponds, in this dataset, to an associated error of 0.2 to 0.3  $\mu m$  on the location of the phase boundaries. This error remains much smaller than the actual size of the features of interest. In other words, the reconstruction does not artificially create additional features (precipitates) nor assign the wrong phase to any feature.

Again, it should also be noted that different final grids can describe the same polynomial distortion function, since the calculation of the function is not affected by the permutation of the points of the mesh grid. Moreover, different polynomial functions can describe the same physical distortion. This phenomenon is enhanced if the points of the

374 speckle are sparsely distributed or mostly located in the center of the region of interest.

#### 375 4.5. Sensitivity to the accuracy of the EBSD speckle

376 As mentioned previously, the use of speckles eliminates the manual selection of reference points. Another advantage  
377 of using speckles and a score based on the number of superimposed pixels, is that the EBSD speckle can be only  
378 a small subset of the original image. This is the case in the Rene 65 dataset, where the BSE speckle contains all the  $\gamma'$   
379 precipitates but the EBSD speckle contains the smallest features in the dataset. Some precipitates are missing on this  
380 speckle, and small grains are mistakenly segmented as precipitates. This does deteriorate the theoretical maximum  
381 score achievable, but does not prevent a good superposition of the speckles, as already shown on fig. 3-c. The influence  
382 of the quality of the EBSD speckle has been studied on the Inconel 718 dataset: some noise was added (wrong  
383 features and random noise) and some pores were also removed from the EBSD speckle on purpose, in comparison  
384 with the one used on 3.3 (fig. 4). The amount of features added or deleted is quantified by the  $\Delta f$  parameter, which  
385 quantifies the percentage of pixels points added or deleted to the speckle containing the correct number of pores:  
386 
$$\Delta f = \frac{\text{number of pixels in wrong speckle}}{\text{number of pixels in right speckle}}$$
. Figure 11 shows the best score obtained as a function of the percentage of pixels added  
387 or deleted in the EBSD speckle. Empty and filled markers correspond to the initial and final scores in CMA-ES,  
388 respectively. The speckle at  $\Delta f = 0\%$  corresponds to the EBSD pattern in which only the pores were segmented. It  
389 leads to the best score achievable. The deletion of points leads to a drastic decrease of the score. The addition of  
390 "wrong" cavities, however, leads to a slight decrease of the score but stagnates to a roughly constant value, as the  
391 number of points added increases. This indicates that the calculation of the distortions is fairly tolerant to the presence  
392 of noise or wrong elements in the EBSD speckle. However, missing points are more detrimental.

#### 393 4.6. The case of $\gamma - \gamma'$ nickel-base superalloys: comparison with other mapping techniques

394 The  $\gamma$  and  $\gamma'$  phases of nickel-based superalloys are well known to be challenging to differentiate, on the single  
395 basis of their diffraction patterns. Methods like conventional EBSD with an indexation of phases based on the Hough  
396 transform are not efficient. Other indexation techniques have been proposed, such as the dictionary approach by De  
397 Graef [31]. In this approach, the collected patterns are compared to a dictionary of synthetically generated patterns.  
398 This technique enables the indexation of patterns with a better angular resolution [32] as well as the differentiation of  
399 phases in complex alloys [33] and geological materials [34]. However, it can be computationally expensive and -this  
400 far- has not been able to differentiate  $\gamma$  and  $\gamma'$  phases. Combined EDS-EBSD methods have been used in the last years  
401 but suffer from long indexing times (and so, increased beam drifting), post-processing times and limited resolution  
402 [12]. Using a completely different technique but a similar approach to dictionary indexing, with ion channeling  
403 imaging, the iCHORD method of Langlois et al. [35] has been shown to enable the differentiation of phases [36].  
404 The channeling contrast is used to obtain information about the orientation. On the other hand, the secondary ion  
405 signals from  $\gamma$  and  $\gamma'$  phases are sufficiently contrasted to enable the identification of phases. In a recent comparative  
406 study, Vernier et al. [37] have shown that this method enables a more accurate mapping than combined EDX-EBSD  
407 methods, since it enables to identify precipitates of size down to 150 nm. The scanning speed is also 3 times faster  
408 than EDS-EBSD methods. The time necessary for the acquisition and indexation of the patterns however, is typically  
409 similar than that of the combined EDS-EBSD method. With a post-processing time of about 15 to 30 minutes, the  
410 multimodal data recombination process using the CMA-ES optimizer appears to be a competitive alternative to the  
411 other methods proposed in the literature. The only constraint is the acquisition of a BSE image of sufficient quality  
412 (in terms of contrast), and its precise segmentation. Indeed, in this method, the size of the smallest objects resolved  
413 depends not only on the step size of the EBSD map but also directly on the quality of the segmentation of the BSE  
414 image. Assuming a good segmentation, the spatial resolution of this technique is about a few pixels, which is better  
415 than EDS-EBSD methods, and comparable to the iCHORD method. It can also be easily implemented in any SEM  
416 equipped with a conventional EBSD detector.

### 417 5. Conclusion and perspectives

418 A new method for the compensation of the distortions and improved phase differentiation in EBSD data has been  
419 developed. The principle consists in using an electron image of the same area than the EBSD data, taken at a 0° tilt  
420 angle, and using it as a reference. Similar features are segmented out of the EBSD data and electron image. Then, the

421 CMA Evolutionary Strategy is applied in order to match the speckles. The goodness of the superposition is measured  
422 by a score, which ranges from 0 to 1. Crystallographic and phase data are then recombined in a new EBSD file. This  
423 new file has the same grid than the initial EBSD file.

424 This method has been applied successfully to nickel-based superalloys and a titanium alloy. Assuming a precise  
425 segmentation of the phases on the electron image, this method can reach a precision of a couple of pixels over broad  
426 areas, despite important drift phenomena. It can be used on any EBSD dataset, as long as two speckles of similar  
427 features can be generated. It is relatively tolerant to inaccuracies and noise in the EBSD speckle. Full automation and  
428 short computation times make it a competitive post-processing method, when compared to other options proposed in  
429 the literature.

430 **Acknowledgements**

431 The authors gratefully acknowledge a Vannevar Bush Fellowship, ONR Grant N00014-18-1-3031.  
432 Henry Proudhon from Mines ParisTech is acknowledged for prolific conversations and exchange of ideas. Olivier  
433 Pietquin, Jérémie Mary and Philippe Preux from INRIA are acknowledged for insightful conversations. Xuedong  
434 Shang from INRIA is acknowledged for his various comments on CMA-ES. McLean Echlin from the University of  
435 Santa Barbara is acknowledged for providing the Ti-6Al-4V sample. Mickael Kirka from the Oak Ridge National  
436 Laboratory is acknowledged for providing the additive manufactured Inconel 718 sample. Tim Clark from Carlton  
437 Forge Works (PCC Corporation) is acknowledged for providing the Rene 65 material.

438 **References**

- [1] S. Suzuki, Features of Transmission EBSD and its Application, *JOM* 65 (9) (2013) 1254–1263, ISSN 1047-4838, doi:\let\@tempa\bibinfo{@X@doi10.1007/s11837-013-0700-6}.
- [2] M. P. Echlin, M. Straw, S. Randolph, J. Filevich, T. M. Pollock, The TriBeam system: Femtosecond laser ablation in situ SEM, *Materials Characterization* 100 (2015) 1–12, ISSN 1044-5803, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.MATCHAR.2014.10.023}.
- [3] D. Rowenhorst, A. Gupta, C. Feng, G. Spanos, 3D Crystallographic and morphological analysis of coarse martensite: Combining EBSD and serial sectioning, *Scripta Materialia* 55 (1) (2006) 11–16, ISSN 13596462, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.scriptamat.2005.12.061}, URL <http://linkinghub.elsevier.com/retrieve/pii/S135964620600039X>.
- [4] G. Spanos, D. Rowenhorst, A. Lewis, A. Geltmacher, Combining Serial Sectioning, EBSD Analysis, and Image-Based Finite Element Modeling, *MRS Bulletin* 33 (06) (2008) 597–602, ISSN 0883-7694, doi:\let\@tempa\bibinfo{@X@doi10.1557/mrs2008.124}.
- [5] F. Lin, A. Godfrey, D. J. Jensen, G. Winther, 3D EBSD characterization of deformation structures in commercial purity aluminum, *Materials Characterization* 61 (11) (2010) 1203–1210, ISSN 1044-5803, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.MATCHAR.2010.07.013}.
- [6] M. Calcagnotto, D. Ponge, E. Demir, D. Raabe, Orientation gradients and geometrically necessary dislocations in ultrafine grained dual-phase steels studied by 2D and 3D EBSD, *Materials Science and Engineering: A* 527 (10–11) (2010) 2738–2746, ISSN 0921-5093, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.MSEA.2010.01.004}.
- [7] L. Holzer, M. Cantoni, Review of FIB tomography, in: *Nanofabrication using focused ion and electron beams : principles and applications*, Oxford University Press, ISBN 9780199734214, 410–435, 2011.
- [8] A. T. Polonsky, M. P. Echlin, W. C. Lenthe, R. R. Dehoff, M. M. Kirka, T. M. Pollock, Defects and 3D structural inhomogeneity in electron beam additively manufactured Inconel 718, *Materials Characterization* ISSN 1044-5803, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.MATCHAR.2018.02.020}.
- [9] M. M. Nowell, S. I. Wright, Phase differentiation via combined EBSD and XEDS, *Journal of Microscopy* 213 (3) (2004) 296–305, ISSN 00222720, doi:\let\@tempa\bibinfo{@X@doi10.1111/j.0022-2720.2004.01299.x}.
- [10] G. West, R. Thomson, Combined EBSD/EDS tomography in a dual-beam FIB/FEG-SEM, *Journal of Microscopy* 233 (3) (2009) 442–450, ISSN 00222720, doi:\let\@tempa\bibinfo{@X@doi10.1111/j.1365-2818.2009.03138.x}.
- [11] D. Child, G. West, R. Thomson, The use of combined three-dimensional electron backscatter diffraction and energy dispersive X-ray analysis to assess the characteristics of the gamma/gamma-prime microstructure in alloy 720Li $\varnothing$ , *Ultramicroscopy* 114 (2012) 1–10, ISSN 03043991, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.ultramic.2011.11.003}.
- [12] M.-A. Charpagne, P. Vennegues, T. Billot, J.-M. Franchet, N. Bozzolo, Evidence of multimicrometric coherent  $\gamma$  precipitates in a hot-forged  $\gamma\gamma$  nickel-based superalloy, *Journal of Microscopy* 263 (1), ISSN 13652818, doi:\let\@tempa\bibinfo{@X@doi10.1111/jmi.12380}.
- [13] E. Payton, G. Nolze, The Backscatter Electron Signal as an Additional Tool for Phase Segmentation in Electron Backscatter Diffraction, *Microscopy and Microanalysis* 19 (04) (2013) 929–941, ISSN 1431-9276, doi:\let\@tempa\bibinfo{@X@doi10.1017/S1431927613000305}.
- [14] G. Nolze, Image distortions in SEM and their influences on EBSD measurements, *Ultramicroscopy* 107 (2-3) (2007) 172–183, ISSN 03043991, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.ultramic.2006.07.003}, URL <http://linkinghub.elsevier.com/retrieve/pii/S0304399106001483>.
- [15] Y. Zhang, A. Elbrond, F. Lin, A method to correct coordinate distortion in EBSD maps, *Materials Characterization* 96 (2014) 158–165, ISSN 1044-5803, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.MATCHAR.2014.08.003}.
- [16] A. D. Kammers, S. Daly, Digital Image Correlation under Scanning Electron Microscopy: Methodology and Validation, *Experimental Mechanics* 53 (9) (2013) 1743–1761, ISSN 0014-4851, doi:\let\@tempa\bibinfo{@X@doi10.1007/s11340-013-9782-x}.
- [17] K. Mingard, H. Jones, M. Gee, Metrological challenges for reconstruction of 3-D microstructures by focused ion beam tomography methods, *Journal of Microscopy* 253 (2) (2014) 93–108, ISSN 00222720, doi:\let\@tempa\bibinfo{@X@doi10.1111/jmi.12100}.
- [18] L. R. Dice, Measures of the amount of ecologic association between species, *Ecology*, Wiley Online Library 26 (3) (1945) 297–302.
- [19] A. A. Taha, A. Hanbury, Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, *BMC medical imaging* 15 (1) (2015) 29.
- [20] Scikit-learn, <http://scikit-learn.org/stable/>, [Online; accessed 23-October-2018], 2018.
- [21] Scikit-image, <https://scikit-image.org/>, [Online; accessed 23-October-2018], 2018.
- [22] O. Ibáñez, L. Ballerini, O. Cordón, S. Damas, J. Santamaría, An experimental study on the applicability of evolutionary algorithms to craniofacial superimposition in forensic identification, *Information Sciences* 179 (23) (2009) 3998–4028, ISSN 00200255, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.ins.2008.12.029}, URL <http://linkinghub.elsevier.com/retrieve/pii/S0020025509000085>.
- [23] A. Sisniega, J. W. Stayman, J. Yorkston, J. H. Siewersden, W. Zbijewski, Motion compensation in extremity cone-beam CT using a penalized image sharpness criterion, *Physics in Medicine and Biology* 62 (9) (2017) 3712–3734, ISSN 0031-9155, doi:\let\@tempa\bibinfo{@X@doi10.1088/1361-6560/aa6869}, URL <http://stacks.iop.org/0031-9155/62/i=9/a=3712?key=crossref.84caaff65ba9b2a3f6da8ee49365aca7>.
- [24] S. S. Reddy, B. Panigrahi, R. Kundu, R. Mukherjee, S. Debchoudhury, Energy and spinning reserve scheduling for a wind-thermal power system using CMA-ES with mean learning technique, *International Journal of Electrical Power & Energy Systems* 53 (2013) 113–122, ISSN 01420615, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.ijepes.2013.03.032}, URL <http://linkinghub.elsevier.com/retrieve/pii/S014206151300149X>.
- [25] S.-E. K. Fateen, A. Bonilla-Petriciolet, G. P. Rangaiah, Evaluation of Covariance Matrix Adaptation Evolution Strategy, Shuffled Complex Evolution and Firefly Algorithms for phase stability, phase equilibrium and chemical equilibrium problems, *Chemical Engineering Research and Design* 90 (12) (2012) 2051–2071, ISSN 02638762, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.cherd.2012.04.011}, URL <http://linkinghub.elsevier.com/retrieve/pii/S0263876212001700>.
- [26] A. C. J. Weber, E. E. Burnell, W. L. Meerts, C. A. de Lange, R. Y. Dong, L. Muccioli, A. Pizzirusso, C. Zannoni, Communication: Molecular dynamics and H-1 NMR of n-hexane in liquid crystals, *The Journal of Chemical Physics* 143 (1) (2015) 011103, doi:\let\@tempa\bibinfo{@X@doi10.1063/1.4923253}.
- [27] N. Hansen, A. Ostermeier, Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation, in: Pro-

- 502       ceedings of IEEE International Conference on Evolutionary Computation, IEEE, ISBN 0-7803-2902-3, 312–317, doi:\let\@tempa\bibinfo{@  
 503       X@doi10.1109/ICEC.1996.542381, 1996.
- 504       [28] N. Hansen, A. Ostermeier, Completely Derandomized Self-Adaptation in Evolution Strategies, *Evolutionary Computation* 9 (2) (2001) 159–  
 505       195, ISSN 1063-6560, doi:\let\@tempa\bibinfo{@X@doi10.1162/106365601750190398.
- 506       [29] C. M. Heaney, M. Lasonde, A. Powell, B. Bond, C. O'Brien, Development of a New Cast and Wrought Alloy (René 65) for High Temperature  
 507       Disk Applications, in: 8th International Symposium on Superalloy 718 and Derivatives, 67–77, 2014.
- 508       [30] M. M. Kirkka, Y. Lee, D. A. Greeley, A. Okello, M. J. Goin, M. T. Pearce, R. R. Dehoff, Strategy for Texture Management in Metals Additive  
 509       Manufacturing, *JOM* 69 (3) (2017) 523–531, ISSN 1047-4838, doi:\let\@tempa\bibinfo{@X@doi10.1007/s11837-017-2264-3.
- 510       [31] Y. H. Chen, S. U. Park, D. Wei, G. Newstadt, M. A. Jackson, J. P. Simmons, M. De Graef, A. O. Hero, A Dictionary Approach to Electron  
 511       Backscatter Diffraction Indexing, *Microscopy and Microanalysis* 21 (03) (2015) 739–752, ISSN 1431-9276, doi:\let\@tempa\bibinfo{@X@  
 512       doi10.1017/S1431927615000756.
- 513       [32] F. Ram, S. Wright, S. Singh, M. De Graef, Error analysis of the crystal orientations obtained by the dictionary approach to EBSD indexing,  
 514       *Ultramicroscopy* 181 (2017) 17–26, ISSN 03043991, doi:\let\@tempa\bibinfo{@X@doi10.1016/j.ultramic.2017.04.016.
- 515       [33] F. Ram, M. De Graef, Phase differentiation by electron backscatter diffraction using the dictionary indexing approach, *Acta Materialia* 144  
 516       (2018) 352–364, ISSN 1359-6454, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.ACTAMAT.2017.10.069.
- 517       [34] K. Marquardt, M. De Graef, S. Singh, H. Marquardt, A. Rosenthal, S. Koizumi, Quantitative electron backscatter diffraction (EBSD) data  
 518       analyses using the dictionary indexing (DI) approach: Overcoming indexing difficulties on geological materials, *American Mineralogist*  
 519       102 (9) (2017) 1843–1855, ISSN 0003-004X, doi:\let\@tempa\bibinfo{@X@doi10.2138/am-2017-6062.
- 520       [35] C. Langlois, T. Douillard, H. Yuan, N. Blanchard, A. Descamps-Mandine, B. Van de Moortèle, C. Rigotti, T. Epicier, Crystal orientation  
 521       mapping via ion channeling: An alternative to EBSD, *Ultramicroscopy* 157 (2015) 65–72, ISSN 0304-3991, doi:\let\@tempa\bibinfo{@X@  
 522       doi10.1016/J.ULTRAMIC.2015.05.023.
- 523       [36] C. Langlois, M.-A. Charpagne, S. Dubail, T. Douillard, N. Bozzolo, Ni-based superalloy: crystalline orientation mapping and gamma-gamma  
 524       phases discrimination with the iCHORD method, in: European Microscopy Congress 2016: Proceedings, Wiley-VCH Verlag GmbH & Co.  
 525       KGaA, Weinheim, Germany, 930–931, doi:\let\@tempa\bibinfo{@X@doi10.1002/9783527808465.EMC2016.6924, 2016.
- 526       [37] S. Vernier, J.-M. Franchet, M. Lesne, T. Douillard, J. Silvent, C. Langlois, N. Bozzolo, iCHORD-SI combination as an alternative to EDS-  
 527       EDS coupling for the characterization of  $\gamma$ - $\gamma$  nickel-based superalloy microstructures, *Materials Characterization* 142 (2018) 492–503,  
 528       ISSN 1044-5803, doi:\let\@tempa\bibinfo{@X@doi10.1016/J.MATCHAR.2018.06.015.

529    **Appendix A. Update of the parameters in CMA-ES**

530      Set $m \in \mathbb{R}^n$ , $\sigma \in \mathbb{R}_+$ , $\lambda$	Input
531 $C = I$ , $p_c = 0$ , $p_\sigma = 0$ ,	Initialize parameters
532 $c_c \approx 4/n$ , $c_\sigma \approx 4/n$ , $c_1 \approx 2/n^2$ , $c_\mu \approx \mu_w/n^2$ , $c_1 + c_\mu \leq 1$ , $d_\sigma \approx 1 + \sqrt{\frac{\mu_w}{n}}$ , $w_{i=1\dots\lambda}$ such that $\mu_w = \frac{1}{\sum_{i=1}^\mu w_i^2} \approx 0.3\lambda$	
533      While not terminate,	Sampling and variation
534 $x_i = m + \sigma y_i$ , where $y_i = \mathcal{N}_i(0, C)$ , for $i = 1, \dots, \lambda$	Update mean
535 $m \leftarrow \sum_{i=1}^\mu w_i x_{i:\lambda} = m + \sigma y_w$ , where $y_w = \sum_{i=1}^\mu w_i x_{i:\lambda}$	Cumulation for the calculation of $C$
536 $p_c \leftarrow (1 - c_c)p_c + \mathbb{1}_{\ p_\sigma\  < 1.5\sqrt{n}} \sqrt{1 - (1 - c_c^2)} \sqrt{\mu_w} y_w$	Cumulation for the calculation of $\sigma$
537 $p_\sigma \leftarrow (1 - c_\sigma)p_\sigma + \sqrt{1 - (1 - c_\sigma)^2} \sqrt{\mu_w} C^{-\frac{1}{2}} y_w$	Update of the covariance matrix $C$
538 $C \leftarrow (1 - c_1 - c_\mu)C + c_1 p_c p_c^T + c_\mu \sum_{i=1}^\mu w_i y_{i:\lambda} y_{i:\lambda}^T$ ,	Update of the step size $\sigma$ .
539 $\sigma \leftarrow \sigma \exp\left(\frac{c_\sigma}{d\sigma}\left(\frac{\ p_\sigma\ }{E\ \mathcal{N}(0,I)\ } - 1\right)\right)$	
540	

541    **Appendix B. Derivative of the vertical and horizontal components of the distortion function**

542    Figure B.12 shows 2 sets of maps of the partial derivatives of the vertical and horizontal components of the  
 543    distortion function reported in fig. 6. As discussed in section 4.1, the variation in color on all four components shows  
 544    that the error on the location ( $x', y'$ ) varies as the scan progresses. Fig. B.12 also shows that this variation is not  
 545    linear. The magnitude of  $\frac{\partial f_x}{\partial x}$  (B.12-a) increases from the left to the right of the map, which means that the error on the  
 546     $x'$  coordinate increases along each line. This gradient of color is the same along the  $y$  component of the map which  
 547    means that this error remains within the same magnitude from a line to another. The magnitude of this overall gradient  
 548    remains always much higher than that of  $\frac{\partial f_x}{\partial y}$  (fig. B.12-b), which means that most of the error on the  $x'$  coordinates  
 549    are accumulated along a line. However, the decreasing magnitude of  $\frac{\partial f_x}{\partial y}$  (fig. B.12-c) along the  $y$  component of the  
 550    map, indicates that the error on the  $x'$  location decreases from a line to another. this is consistent with a decreasing  
 551    drift effect as the scan proceeds. As concerns the derivatives of the vertical component  $f_y$ , one can see that most of  
 552    the error on the  $y'$  coordinate occurs along the vertical direction, from a line to another (fig. B.12-d). On the other  
 553    hand,  $\frac{\partial f_y}{\partial x}$  (fig. B.12-c) is fairly constant and low along a line and from a line to another. Those trends are consistent  
 554    with drift phenomena, where the drift is high in the beginning of the scan, and decreases along a line, and from a line  
 555    to another.

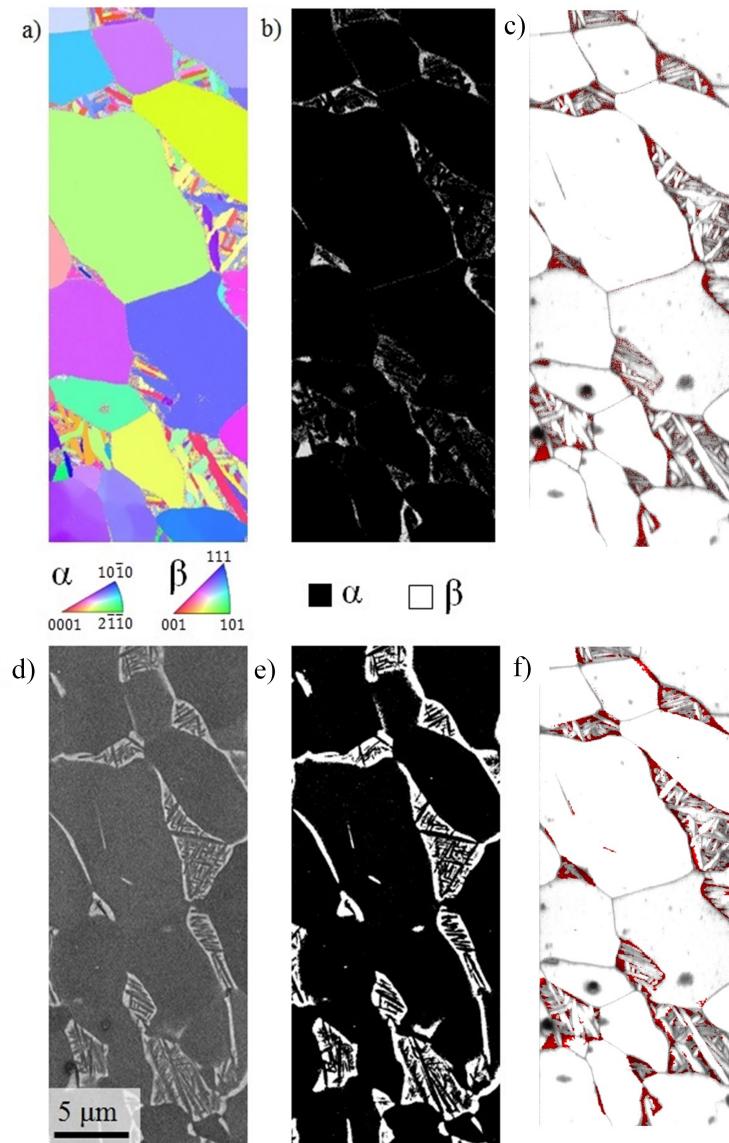


Figure 2: Microstructure of the Ti-64 sample: a) Inverse Pole Figure (projected along the normal to the sample surface), b) EBSD speckle segmented out of the initial phase map, with  $\alpha$  phase in black and  $\beta$  phase in white, c) Initial phase map superimposed to the Image Quality map, d) BSE image used for phase segmentation, e) Segmented BSE image used as a phase map in the un-distorted EBSD file. f) Final phase map colored as c).

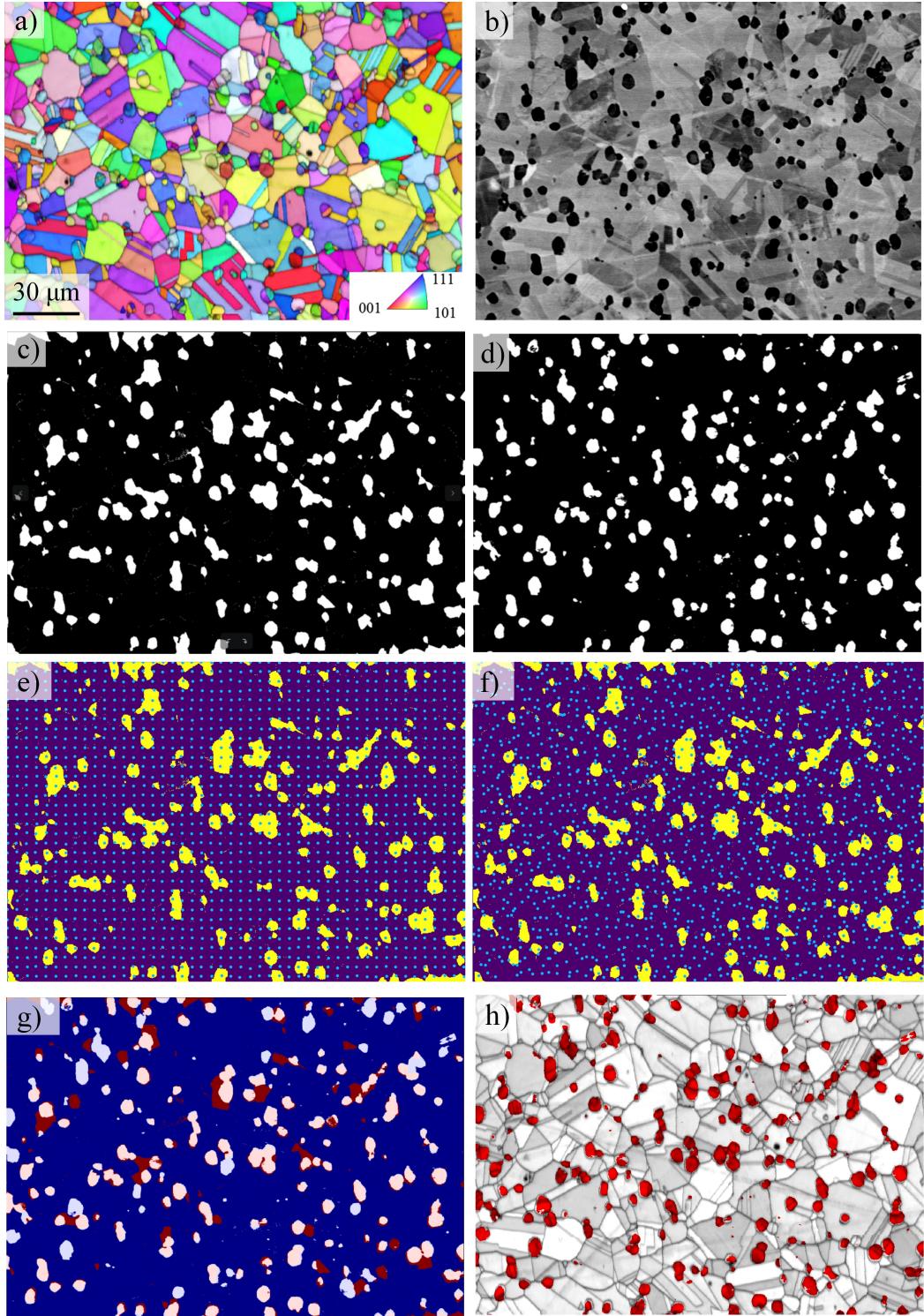


Figure 3: Microstructure of the solution annealed Rene 65 sample. a) EBSD map colored according to the Inverse pole figure projected along the normal direction to the surface of the sample, b) BSE image from which the precipitates have been segmented, c) BSE speckle, d) EBSD speckle: smallest features of the dataset segmented from the initial EBSD file, considering annealing twins as internal defects, e) initial mesh over the EBSD speckle, f) final mesh over the EBSD speckle, g) Superimposed EBSD (blue) and BSE (red) speckles, h) Phase map with the precipitates colored in red superimposed to the pattern quality map of the new EBSD file.

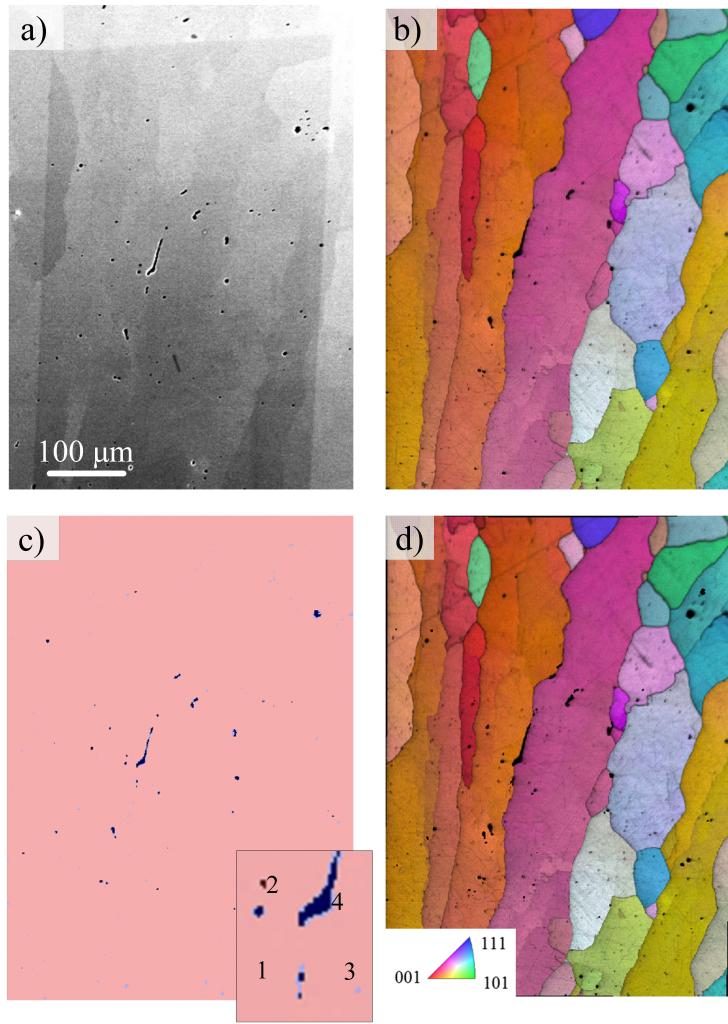


Figure 4: Microstructure of Additive Manufactured Inconel 718, a) SE image of used for the segmentation of the pores, the trapezoid contamination zone corresponding to the EBSD area is visible, b) Initial EBSD map colored according to the Inverse Pole Figure projected along the build direction superimposed to the Image Quality map in greyscale, c) Superimposed SE (red) and EBSD (blue) speckles after CMA-ES, d) Final EBSD map superimposed to the Image quality map in greyscale, with pores colored in black.

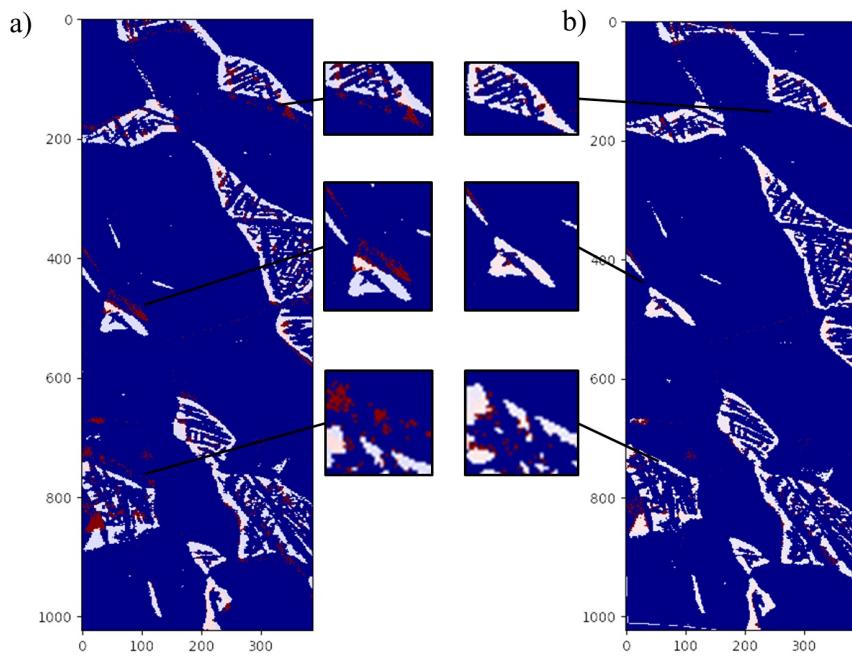


Figure 5: Speckle superposition after correction of a) Affine distortions only, b) All distortions. On a blue background, the EBSD speckle is colored in red and the BSE speckle in white.

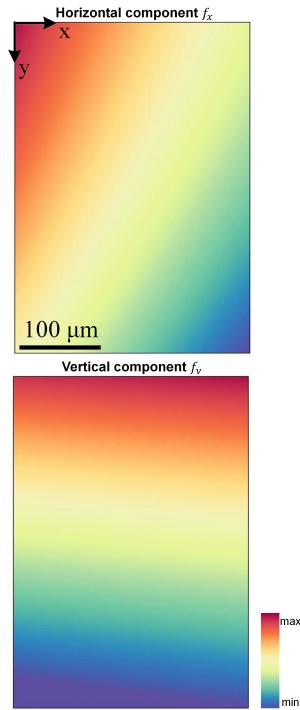


Figure 6: X and Y components of the distortion function in the Additive Manufactured Inconel 718 sample.

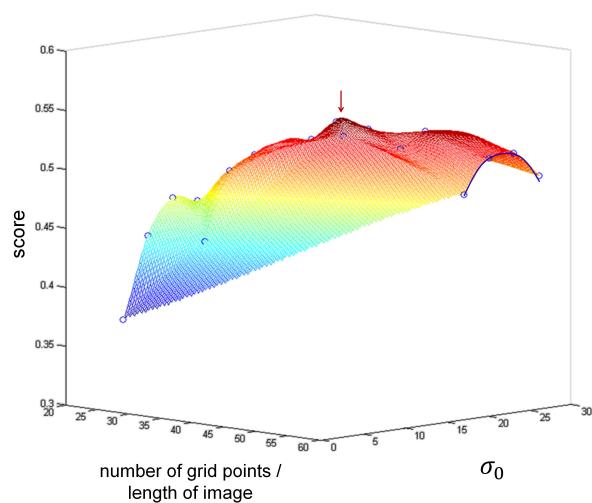


Figure 7: Influence of the initial CMA-ES parameters on the final score, in the Ti-6Al-4V dataset.

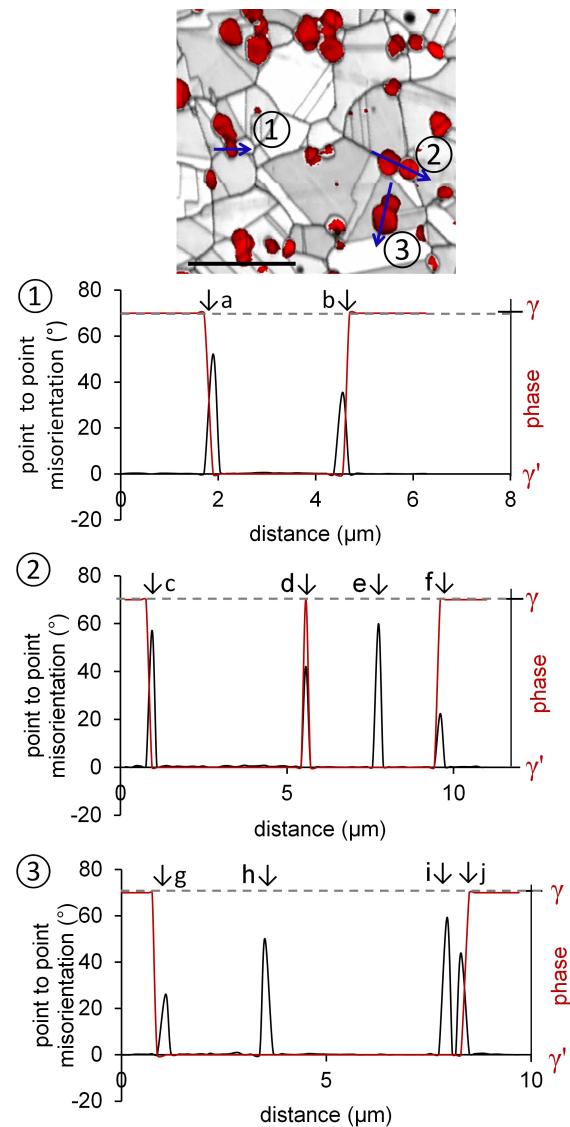


Figure 8: Comparison of the location of phase and grain boundaries in the reconstructed EBSD data of the Rene 65 alloy, on three misorientation profiles across  $\gamma'$  precipitates.

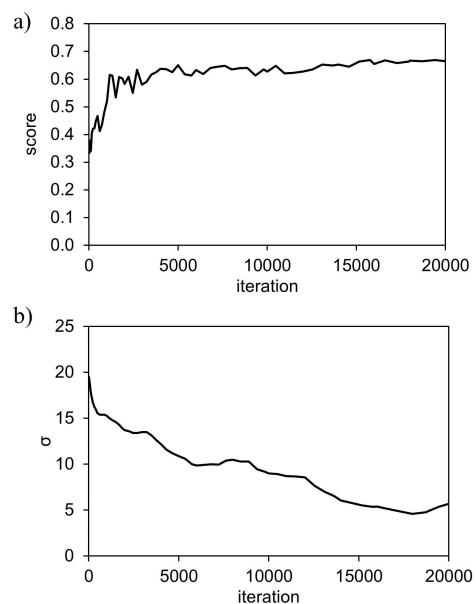


Figure 9: Typical optimization path with CMA-ES, as a function of the iterations, here on the AM Inconel 718 dataset: a) Value of the similarity function at the generation  $g$ , b)  $\sigma^{(g)}$ : step size at the generation  $g$ .

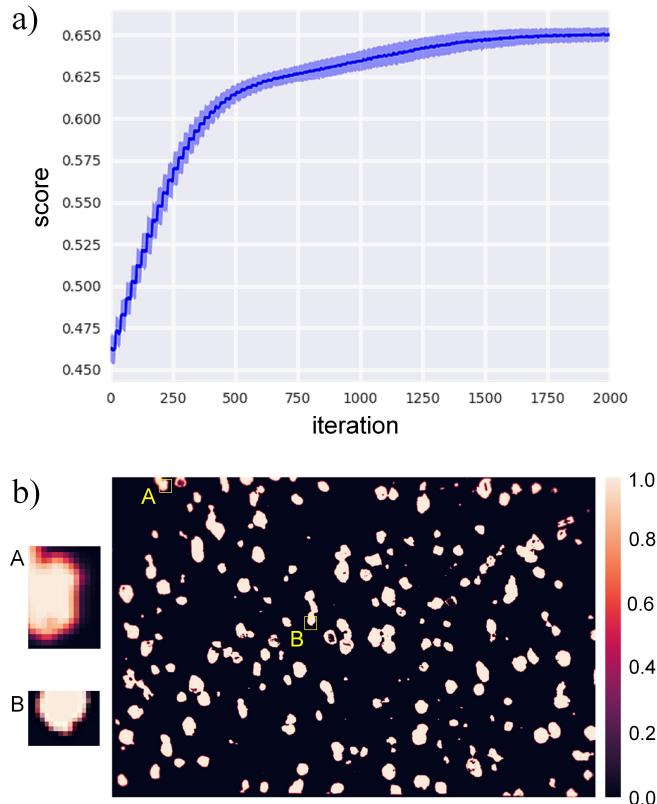


Figure 10: Convergence and stability of the CMA optimization over 100 runs in the Rene 65 dataset. a) Mean score (dark blue), minimum and maximum values (light blue), as a function of the number of iterations, b) Corresponding heatmap : the pixels are colored according to the number fraction of times they were identical among the 100 runs.

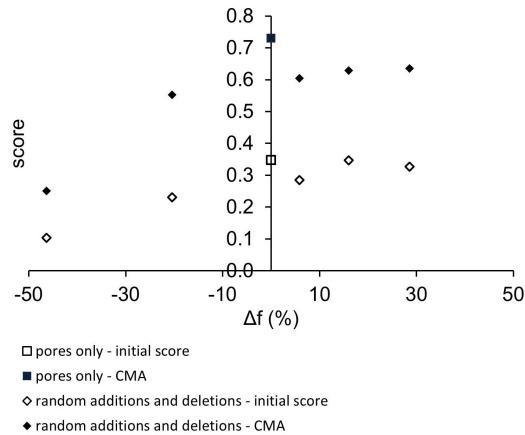


Figure 11: Influence of the addition or deletion of points on the EBSD speckle. Initial scores and final scores in CMA-ES as a function of the percent of points added or deleted.

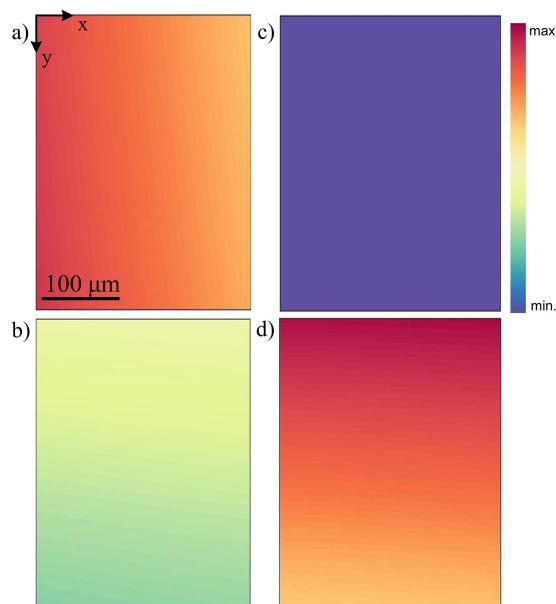


Figure B.12: Partial derivatives of the vertical and horizontal components of the distortion function along the X and Y coordinates, for the Additive Manufactured Inconel 718 material. a)  $\frac{\partial f_x}{\partial x}$ , b)  $\frac{\partial f_x}{\partial y}$ , c)  $\frac{\partial f_y}{\partial x}$ , d)  $\frac{\partial f_y}{\partial y}$ .