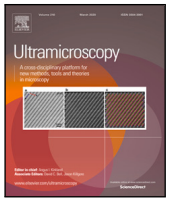
Ultramicroscopy 233 (2022) 113407

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic



Orientation, pattern center refinement and deformation state extraction through global optimization algorithms

Chaoyi Zhu [a,*], Christian Kurniawan [a], Marcus Ochsendorf [a], Dayong An [b], Stefan Zaefferer [b], Marc De Graef [a]

[a] Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, PA, USA
[b] Max-Planck-Institut für Eisenforschung, Max-Planck-Str. 1, 40237, Düsseldorf, Germany

ARTICLE INFO

Keywords:
EBSD
Global optimization
Orientation refinement
Pattern center refinement
Deformation tensor inference

ABSTRACT

Global optimization algorithms have been adopted to the simultaneously refinement of orientation and pattern center for electron backscatter diffraction patterns as well as deformation state extraction. The hyperparameter space and mutation schemes of differential evolution (DE) algorithm has been thoroughly investigated and showed to be a more efficient algorithm than the particle swarm optimization (PSO) algorithm. The optimal hyperparameters for DE generally depend on conditions such as the number of variables to be optimized and the size of bounded search space but reasonably close initial values for crossover probability is 0.9, mutation factor is 0.5, population size is ten times the number of variables, and number of iterations is at least 100. Validation on a set of simulated undeformed single crystal nickel patterns reveals a mean accuracy of ≈ 0.03° and ≈0.01% detector width across a large field of view. In addition, validation using noisy simulated deformed patterns with known deformation state and pattern center shows that the mean accuracy of shear strain and rotation components is ≈0.001 and for the normal strain ≈0.002.

1. Introduction

Fully automated electron backscatter diffraction (EBSD) in a scanning electron microscope [1–4] has enabled many aspects of materials characterization, such as spatially resolved crystal orientation and phase mapping, texture analysis, and deformation studies [5]. In the past, efforts in the EBSD community have been focused on reducing the measurement error through improved information retrieval algorithms or advanced calibration methods [6]; for instance, the angular resolution of the modern Hough-based EBSD system is around 0.2-1° and the absolute orientation error could reach about 1-2° [7,8]. The angular resolution (i.e., the precision of the EBSD measurement) is typically associated with the retrieval method, i.e., the Hough transform to locate Kikuchi bands and the localization of the pattern center [6]. The absolute orientation error (i.e., the accuracy of the EBSD measurement) is attributed to poor system calibration, e.g., an unaccounted for inclination of the sample surface [9]. Since the true orientation is not known for real patterns, the true orientation error is often hard to determine. From the von Mises–Fisher distribution, one can statistically infer that the true orientation accuracy is limited by the confidence region (0.3°) of the retrieval algorithm in which the angular distance between detector plane normal vectors and recalculated plane normal vectors is minimized. To improve the angular resolution, methods such as the 3D-Hough transform [10], iterative indexing [11], pattern comparison method [12], and HR-EBSD [13] have been previously developed.

Additionally, the accuracy of orientation determination also relies on an accurate localization of the pattern center [6]. One common method of refining the pattern center of a single pattern is to use the iterative fitting method by minimizing an objective function related to the pattern center coordinates [14], resulting in an accuracy of around 0.5% of the detector width. Other more sophisticated methods have also been developed based on improved retrieval algorithms [10], a moving detector screen [15,16], shadow-casting [17], and a single crystal calibrant [18,19].

More recently, the dynamical scattering of BSEs on their way out of the sample has been modeled starting from the Darwin–Howie–Whelan dynamic equations [20] in the Bloch-wave formalism [21]. In addition, the stochastic nature of scattering can be merged with the deterministic process using a Monte Carlo simulation approach [22] that allows for simulated patterns to be compared directly with experimental patterns with encouraging indexing precision [23–26]. For

* Corresponding author.
E-mail addresses: zchaoyi@berkeley.edu (C. Zhu), christian.paryoto@gmail.com (C. Kurniawan), mochsend@andrew.cmu.edu (M. Ochsendorf), andyallone@nwpu.edu.cn (D. An), s.zaefferer@mpie.de (S. Zaefferer), degraef@cmu.edu (M. De Graef).

https://doi.org/10.1016/j.ultramic.2021.113407
Received 30 December 2020; Received in revised form 26 September 2021; Accepted 3 October 2021
Available online 11 November 2021

instance, Singh et al. have shown that the dictionary indexing (DI) approach, which is based on normalized dot products between simulated EBSD patterns and experimental patterns, has an improved sensitivity compared to the conventional Hough-based indexing for poor quality diffraction patterns and overlapping patterns [27]. Recently, Ding et al. have also shown that a uniformly sampled dictionary of patterns can be effectively used as a training data set for a convolutional neural network (CNN) to index experimental patterns [28]. However, indexing of lower symmetry crystals necessarily increases the required size of the dictionary since the dictionary size is proportional to the size of fundamental zone (FZ) of the corresponding crystal. Alternatively, forward model based orientation indexing can be conveniently accomplished with a single simulated spherical Kikuchi sphere of the corresponding crystal structure using the spherical harmonic transform (SHT) [26,29], significantly reducing the computation time for lower symmetry crystals.

With the increasing demand of using EBSD to determine strain [30] or to resolve pseudosymmetry issues [31,32], a convenient way of simultaneously refining the orientation and the pattern center has become increasingly attractive. For instance, the pattern matching method has been incorporated recently into an optimization algorithm to obtain a refined orientation and pattern center. Different optimization algorithms have been explored by Winkelmann et al. (NMS: multi-start Nelder–Mead simplex) [33]; Singh et al. (BOBYQA: bound optimization by quadratic approximation) [34]; Pang et al. (SNOBFIT: stable noisy optimization by branch and fit) [35], and Tanaka et al. (DE: differential evolution) [36] to simultaneously refine pattern center and orientation. Some recent applications of this technique include the spatial resolved mapping c/a ratio of martensites [37, 38]. These optimization methods mainly fall into three categories: (1) *sampling methods* or *direct search methods* that require a rather significant number of function evaluations, e.g., DE; (2) *modeling methods* that approximate the function over a region by some model function, e.g., BOBYQA; and (3) a hybrid approach of (1) and (2) like the multi-start NMS and SNOBFIT. In comparison, *sampling methods* usually provide a more accurate picture of the optimization landscape to locate the global minimum whereas the *modeling methods* are less accurate in terms of the global picture of optimization but are computationally cheaper. However, direct implementation of a *sampling method* such as the Nelder–Mean simplex alone has been found to be inefficient [33,35] due to the sloppy optimization landscape when orientation and pattern center were simultaneously refined. As shown by Pang et al. [35] and Singh et al. [34], the small orientation changes can be compensated by shifts in pattern center positions and vice versa, resulting in a shallow optimization landscape and a slow convergence along the sloppy direction. A multi-start NMS has been shown to help with this convergence problem [33] with a search space of the optimization limited to 1° in misorientation and ±2% in detector width for the pattern center. Similarly, SNOBFIT also provides excellent refinement accuracy within a limited search space of 1° in misorientation and ±1% in detector width for the pattern center. More specifically, the SNOBFIT method uses branching and local stochastic model fitting (linear or quadratic) to simultaneously search for multiple local minima: a surrogate function based derivative-free approach with multiple search branches to explore the global optimization landscape simultaneously. Since the surrogate functions are fitted to the evaluated function points instead of directly interpolated, it is robust for noisy function values. However, these hybrid methods rely on the search spaces containing the ground truth to be sufficiently narrow to allow for an accurate convergence.

To overcome the limited range of search space and the sloppy optimization landscape, global optimization methods, such as DE, have been shown to work rather well for a search space of 10° in misorientation and ±5% in detector width for the pattern center using noisy simulated patterns [36]. Despite the success of their work, some questions were left unanswered, for instance, the justification of using DE over other global optimization methods such as particle swarm optimization; the optimal range of optimization parameters and choice mutation scheme; the effect of changing the search space, and so on. More specifically, one concern was related to the optimization of the pattern center in the presence of significant deformation, which has been shown to cause a bias in the pattern center. Therefore, a method of incorporating the deformation tensor into the optimization routine is needed.

Experimental measurement of absolute strain and rotation has been previously explored by many others [39–42] using relatively strain free reference patterns from the same material. Furthermore, simulation based HR-EBSD has been previously studies by Villert et al. [39] using both kinematically and dynamically simulated strain free reference patterns. Another iterative scheme developed by Kacher et al. [19] used kinematical pattern to be simulated several times until convergence to be used as strain free reference pattern. However, calibrant sample is needed for Kacher's approach to accurately locate the pattern center so that the comparison of simulated pattern and experimental becomes meaningful. The present study aims to provide an accurate determination of the pattern center and deformation tensor via simulated based optimization approach without a calibrant sample. The caveat of simultaneous determination of pattern center and deformation tensor has been previously discussed by Alkorta [43]. To a first order approximation, the error in pattern center can be compensated by a phantom deformation gradient tensor. The polar decomposition of this phantom deformation gradient tensor reveals that it consists of two lattice rotation components ω_{31} and ω_{23} that compensates for the pattern center shifts in the y and x direction respectively (in the detector frame). This also helps to explain the sloppy optimization problem observed by others [34,35], which is the motivation of the current work.

In the present work, evolutionary algorithms, such as particle swarm optimization (PSO) and differential evolution (DE), are explored together and compared. Both methods can be parallelized easily over multiple CPU threads using openMP. In addition, NMS has also been implemented after global optimization to further refine the solution, i.e., a hybrid optimization strategy. These global (hybrid) optimization algorithms (DE and PSO) have been tested on the problem of pattern center and orientation refinement using simulated data with a known ground truth. Since the performance of these algorithms relies on an appropriate selection of the optimization parameters, a parametric study has been done to fine-tune these parameters. In addition, direct inference of the deformation state through pattern matching using a forward model based global optimization is also revisited in this study, based on recent work in our group [44]. Previously, we have demonstrated that an approximate model for the inclusion of deformation into simulated EBSPs can be utilized to accurately infer the deformation gradient tensor (eight degrees of freedom) and the pattern center parameters (three degrees of freedom) with the PSO+NMS method [44], while keeping the Euler angles known and fixed. The accuracy of the deformation tensor was found to be $\approx 10^{-4}$ on noise free simulated patterns and $\approx 10^{-3}$ for noisy patterns. However, the spatial variation of the absolute error using noisy simulated patterns based on Yoffe–Shaibani–Hazzledine (YSH) [45,46] dislocation displacement field data was unclear and no experimental patterns were included.

We will present in this paper the efficient use of hybrid optimization (global algorithm then followed by local algorithm) to simultaneously refine pattern/orientation and infer the deformation tensor. For the first time, the selection of hyperparameters for the global optimization algorithm has been studied in details for pattern matching problem. More importantly, combined with HR-EBSD, spatially resolved strain maps closer to the absolute strain maps can be obtained after the strain state has been corrected for the reference pattern. We would also like to point out that, since the strain states are not known for experimental patterns used in this study, the "absolute" strain mapping could still be subject to some unknown systematic error. The accuracy measure presented in this paper has been derived from simulated patterns with

known strain states. Of course, there are still work left to be pursued in terms of experimental validation and improvement of the current algorithm. Nevertheless, the current method has clearly extended the conventional HR-EBSD method and led to promising new research directions in the understanding the relationship between mechanical behavior of materials and microstructure, particularly, when a strain free reference pattern is not readily available.

2. Methodology

2.1. Optimization problem setup

Solving an optimization problem is the process of maximizing or minimizing an objective function defined with respect to a feasible set of variables. In the context of diffraction pattern optimization, the optimization problem is designed to search for a set of variables that generates a simulated pattern that best matches the target pattern based on the formal definition below:

$$\max_{\mathbf{x}\in\mathbf{S}} f(\mathbf{x}) = -\min_{\mathbf{x}\in\mathbf{S}}(-f(\mathbf{x})), \tag{1}$$

where $\mathbf{x}$ is a vector containing n simulation variables (pattern center, deformation tensor, orientation, etc.), and $\mathbf{S}$ is a feasible subset of $\mathbb{R}^n$, $f(\mathbf{x})$ is the image similarity measure between the simulated pattern and target pattern (normalized dot product, cross-correlation coefficient, etc.). Global optimization specifically seeks for the global minimizer $\mathbf{x}^*$ which gives the optimal value $f(\mathbf{x}^*)$ at which

$$f(\mathbf{x}^*) \le f(\mathbf{x}) \quad \forall\, \mathbf{x} \in \mathbf{S}. \tag{2}$$

From the available experimental data set determined through Hough transform indexing, an initial estimate of the solution to the orientation as well as pattern center coordinates is known prior to the optimization. Hence, the feasible solution set can reasonably be further constrained down, i.e., bounded optimization. For a well calibrated system, this range of search space is typically small as discussed in the introduction section. If the EBSD system is poorly calibrated by the vendor's technician or being improperly operated by user, the idea of implementing the global optimization becomes more attractive. It can tolerate a larger error relative to the initial estimate by expanding the search space.

The orientation search space is confined by the misorientation angle relative to an initial estimate in the back end. Each rotation relative to the initial orientation is described by a neo-Eulerian representation i.e., a stereographic vector $\mathbf{s}$, also known as the modified Rodrigues parameter, [47] given by:

$$\mathbf{s} = \mathbf{n}\tan\omega/4, \tag{3}$$

where $\mathbf{n}$ is the unit vector describing the axis of rotation and ω is the angle of rotation about $\mathbf{n}$. Additionally, the pattern center search space is defined in terms of percent of the detector width, consistent with other literature in the community. Finally, the elastic deformation gradient tensor components are bounded relative to the identity matrix which includes both rotation and stretch of the lattice. Therefore, the options for optimization are: 1) orientation and pattern center, 2) deformation gradient tensor and pattern center. Details regarding how to use the global optimization program can be found in the supplementary material.

2.2. Evolutionary algorithm: Differential evolution

Differential evolution involves an evolving population of search points based on a bio-inspired evolutionary information to allow the population of search points to converge to the global minimum, i.e., an evolutionary algorithm. The basic differential evolution algorithm, following the study by Storn and Price [48], consists of several critical steps: first, the initialization step is a random sampling of target vectors $\mathbf{x}_{i,g}$ within the prescribed range with respect to the initial solution (vectors outside the prescribed range are placed on the boundary). In this case, the initial solution is determined by Hough indexing or other available indexing methods.

$$\mathbf{x}_{i,g} = x_{j,i,g}, \quad j \in \{1,\dots,D\}, i \in \{1,\dots,N_p\}, g \in \{1,\dots,g_{max}\}, \tag{4}$$

where D is the number of parameters to be optimized, N_p is the population size for target vectors of a generation and g is the maximum number of generations or iterations. For each target vector in the generation, a mutation vector $\mathbf{v}_{i,g+1}$ is generated to evolve the search space through different mutation scheme, e.g., DE/rand/1, DE/rand/2, DE/rand-to-best/1, DE/best/1, DE/best/2; these five mutation schemes correspond to the following mutation equations:

$$\mathbf{v}_{i,g+1} = \mathbf{x}_{r_1,g} + F(\mathbf{x}_{r_2,g} - \mathbf{x}_{r_3,g}), \tag{5a}$$

$$\mathbf{v}_{i,g+1} = \mathbf{x}_{r_1,g} + F(\mathbf{x}_{r_2,g} + \mathbf{x}_{r_3,g} - \mathbf{x}_{r_4,g} - \mathbf{x}_{r_5,g}), \tag{5b}$$

$$\mathbf{v}_{i,g+1} = \mathbf{x}_{i,g} + F(\mathbf{x}_{\text{best},g} - \mathbf{x}_{i,g} + \mathbf{x}_{r_1,g} - \mathbf{x}_{r_2,g}), \tag{5c}$$

$$\mathbf{v}_{i,g+1} = \mathbf{x}_{\text{best},g} + F(\mathbf{x}_{r_1,g} - \mathbf{x}_{r_2,g}), \tag{5d}$$

$$\mathbf{v}_{i,g+1} = \mathbf{x}_{\text{best},g} + F(\mathbf{x}_{r_1,g} + \mathbf{x}_{r_2,g} - \mathbf{x}_{r_3,g} - \mathbf{x}_{r_4,g}), \tag{5e}$$

where $r_1, r_2, \dots, r_5$ are random integers $\in \{1,\dots,N_p\}$, $\mathbf{x}_{\text{best},g}$ is the target vector with the lowest objective function value, and F (>0) is a user defined mutation factor (typically $F \in [0,2]$).

Each trial vector $\mathbf{u}_{i,g+1}$ is further diversified by implementing a binomial crossover step to the mutation vectors and target vectors.

$$\mathbf{u}_{i,g+1} = u_{j,i,g+1} \begin{cases} v_{j,i,g+1} & \text{if } rand_j \le C_r \text{ or } j = j_{rand}, \\ x_{j,i,g} & \text{otherwise,} \end{cases} \tag{6}$$

where $rand_j \in [0,1]$ is a random constant for the jth evaluation, $j_{rand} \in \{1,\dots,D\}$ is a random integer index, and $C_r \in [0,1]$ is a user defined crossover probability. Finally, a selection based on objective function values of trial vectors and target vectors is performed to obtain the next generation of target vectors $\mathbf{x}_{i,g+1}$.

$$\mathbf{x}_{i,g+1} = \begin{cases} \mathbf{u}_{i,g+1} & \text{if } f(\mathbf{u}_{i,g+1}) \le f(\mathbf{x}_{i,g}) \\ \mathbf{x}_{i,g} & \text{otherwise.} \end{cases} \tag{7}$$

In Section 3.2, fine-tuning of the optimization parameters is carried out to achieve a balance of computational cost and accurate optimization results. The effect of employing different mutation strategies has also been explored by varying the amount of added noise and changing the range of search spaces.

2.3. Evolutionary algorithm: Particle swarm optimization

Particle swarm optimization is also an evolutionary global search algorithm, originally developed by Eberhart and Kennedy [49]. The implementation of PSO in this study adopted some modifications suggested by others [50,51]. It was inspired by the foraging behavior of flocks of birds and schools of fish by formulating a population of particles each having its own position ($\mathbf{x}_{i,g}$ where $i \in \{1,\dots,N_p\}, g \in 1,\dots,g_{max}$) and velocity ($\mathbf{v}_{i,g}$); N_p is the total number of particles and g_{max} is the maximum number of iterations.

The population of particles ($\mathbf{x}_{i,g}$) is first randomly initialized from a uniform probability distribution function within a bounded domain and the initial velocity of each particle initialized to zero. During the optimization process, the position and the velocity of each particle is updated according to information ($f(\mathbf{x}_{i,g})$: objective function values) shared among all particles in the current iteration. The velocities to update the position of the current iteration of particles are given by the following equation:

$$\mathbf{v}_{i,g+1} = w_g\mathbf{v}_{i,g} + c_1 r_1(\mathbf{x}_{pb,g} - \mathbf{x}_{i,g}) + c_2 r_2(\mathbf{x}_{gb,g} - \mathbf{x}_{i,g}) \tag{8}$$

where $w_g = w_0(w_{\text{damp}})^{g-1}$ is the inertia weight updated for each iteration, with w_0 being the initial inertia weight and w_{damp} being the inertia weight damping coefficient; c_1 is the self-cognition acceleration

coefficient, c_2 is the social cognition acceleration coefficient and r_1, r_2 are uniformly distributed random numbers in the range $[0, 1]$. $\mathbf{x}_{pb,g}$ is the personal best particle, i.e., the particle with the lowest personal objective function value and the $\mathbf{x}_{gb,g}$ is the global best particle, i.e., the particle with the lowest global objective function value. The positions of particles in the next iteration are therefore obtained by

$$\mathbf{x}_{i,g+1} = \mathbf{x}_{i,g} + \mathbf{v}_{i,g+1} \tag{9}$$

To confine the particles in the case when the actual global minimum is close to the boundary of the search space, particles sitting outside the defined search space after each iteration are placed exactly on the corresponding dimension's boundary.

In the present work, the PSO is used in comparison to the DE method. The default particle swarm optimization parameters in this study are adjusted based on our previous work [44]: $w_0 = 1, w_{\text{damp}} = 0.9, c_1 = 2, c_2 = 2$. While fine-tuning of optimization parameters is also possible, it has been found in this work and previous work that the mutation of the velocities based on the best performing personal/global particles sometimes drive the population towards a local minimum for large search spaces.

2.4. Nelder–mead simplex

The Nelder–Mead Simplex algorithm is a widely used derivative free direct search algorithm first proposed by Nelder and Mead (1965) [52]. A simplex is a geometric figure in an n-dimensional space which has $n + 1$ vertices. Based on subsequent reflection, expansion, contraction, and shrinkage of the simplex, the shape of the simplex adapts to the optimization landscape and converges to a local minimum. Rigorous proof of the convergence behavior of the Nelder–Mead simplex is still an active field of research, which in some cases has already been proven to be non-convergent for standard NMS [53]. In addition, standard NMS often becomes stagnant and inefficient in high dimensional problems [54]. In this study, it has also been found that direct implementation of standard Nelder–Mean simplex tends to get stuck (extremely slow convergence) within the sloppy part of the optimization landscape, i.e., an inefficient convergence. Appropriate parameter scaling could possibly improve the convergence [35] or potentially using a more advanced variant of NMS with adaptive parameters [55] and a multi-start procedure [33]. In the present work, the standard NMS has been implemented, following O'Neill's work [56], only as an option to further refine the globally optimized solution to circumvent these common issues related to NMS.

2.5. Objective function and rotation correction

The choice of objective function is crucial to achieve convergence in any evolutionary algorithm. Indeed, prior-knowledge of the target problem to be optimized can inform the selection of an appropriate objective function. In this study, the forward model based optimization demands that the objective function be an effective image similarity metric between simulated pattern and target pattern. From a previous study by Pang et al. [35], the optimization landscape for pattern center and orientation has been found to be very sloppy for the normalized dot product metric, meaning that similar objective function values exist for a wide range of input parameters.

A rotation correction method is proposed here along the orientation space in order to form a less sloppy optimization landscape while the feature space for the pattern center is kept unchanged [34]. For a deviation $(\triangle PC_x, \triangle PC_y)$ of the pattern center from the initial pattern center calibration of the EBSD system, assuming that the initial calibration is precisely known, a rotation correction term given by the axis–angle pair $(\mathbf{n}, \cos\omega)$ can be calculated to scale the orientation space according to:

$$\mathbf{n} = \frac{1}{\rho_c}(-\triangle PC_x \cos\alpha, \triangle PC_y \cos 2\alpha, \triangle PC_x \sin\alpha), \tag{10a}$$

$$\cos\omega = \frac{DD + \triangle PC_y}{[DD^2 + 2 \cdot DD \cdot \triangle PC_y \delta \sin 2\alpha + (\triangle PC_x \delta)^2 + (\triangle PC_y \delta)^2]^{1/2}}, \tag{10b}$$

where α is the tilt angle between the detector screen and sample surface, DD is the detector distance, ρ_c ensures that $\mathbf{n}$ is a unit vector, and δ is the detector pixel size; the rotation correction undoes the apparent pattern rotation when the pattern center is modified. In the rotation correction formula, α, DD, δ are fixed constants. The $(\triangle PC_x, \triangle PC_y)$ values changes during the optimization as defined by the distance between trial pattern center solution and initial calibration of PC. The coordinates of a search point x_{trail} during the optimization step is therefore defined by the following equation:

$$x_{trail} = [L_{correction}(\mathbf{s}_{\mathbf{trial}}), PC_{trial}] \tag{11}$$

where the rotation correction operator $L_{correction}$ is calculated from $(\mathbf{n}, \cos\omega)$ to rotate the trial stereographic vector. The trial pattern center PC_{trial} values do not change in this step. We would like to mention here again that this correction method assumes that the initial calibration of pattern center is not significantly off, which leads to some limitations that will be further discussed.

3. Results

3.1. Optimization landscapes

Simulated patterns are used for the visualization of optimization landscapes and for the parametric study, noise sensitivity test and deformation tensor inference accuracy test. Details regarding the simulation parameter are listed in the supplementary materials.

With the inclusion of rotation correction in the orientation space, as discussed in Section 2.5, optimization landscapes of four image similarity metrics including normalized dot product (NDP), cross-correlation coefficient (XCF), root mean square error (RMSE) and mutual information (MI) are depicted in Fig. 1. The formal definitions of these similarity metrics are stated in Appendix A. In these plots, only two variables, Φ and PC_y, are adjusted to map out a 2D slice of the image similarity landscape since a six-dimensional space is difficult to visualize. The pattern shift due to Φ rotation is often confused with the pattern center shift in the y direction. The x $(\triangle\Phi)$ and y $(\triangle PC_y)$ axes correspond to the differences relative to the ground truth ($\Phi = 0°$ and $PC_y = 0$ pixels). These normalized landscapes reveal that the NDP and XCF landscapes are very similar to each other with and without rotation correction. This sloppy feature is also observed in the RMSE landscape but it is narrower compared to NDP and XCF. For the same initial population, the initial convergence to the sloppy feature is likely faster than NDP and XCF. The only concern for RMSE is that it is intrinsically more sensitive to contrast and brightness changes. In addition, the MI landscape has a very narrow and deep valley close to the ground truth. The lack of gradient information in general does not favor fast convergence compared to others.

It is clear that the rotation correction slightly removes sloppiness of all the objective functions' optimization landscapes but the benefit in terms of accuracy improvement will be limited. It has been found in this study that the rotation correction does indeed slightly improve the convergence speed for the global optimization part. However, the rotation correction needs to be switched off for the NMS in the hybrid method because it will incorrectly bias the Nelder–Mead simplex algorithm. In the following test studies, we used the rotation correction only for the global optimization step.

3.2. Differential evolution: Rules of thumb and hyperparameter tuning

During the global optimization process for DE, a few optimization parameters are usually set prior to execution: N_p population size, g number of generations, C_r crossover probability, F mutation factor.

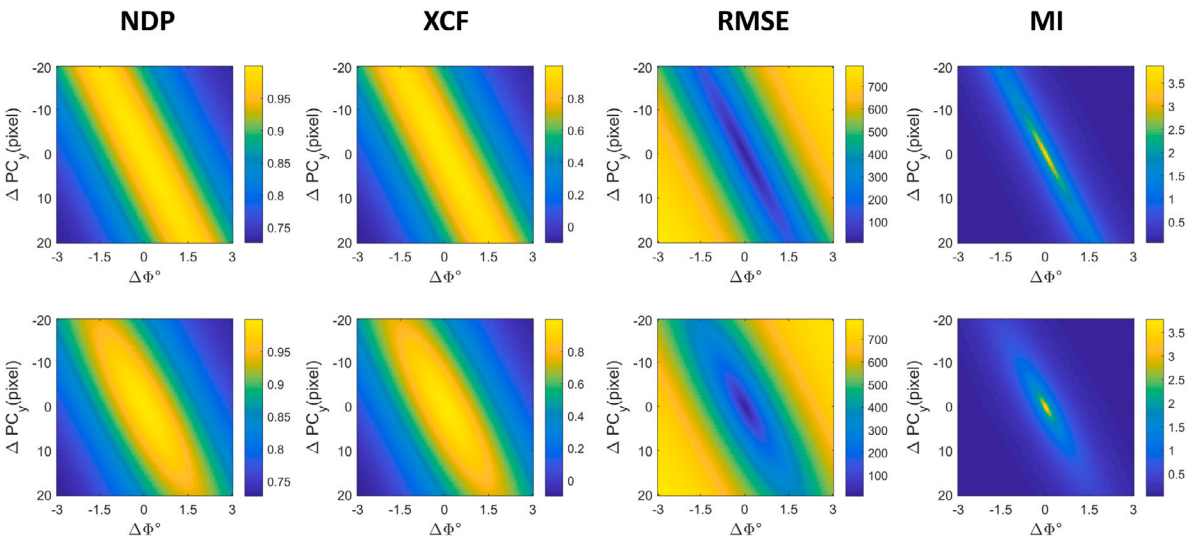


Fig. 1. Optimization landscapes using objective functions (from left to right) normalized dot product (NDP), cross-correlation coefficient (XCF), root mean square error (RMSE), mutual information (MI) (top row) without rotation correction (bottom row) with rotation correction.

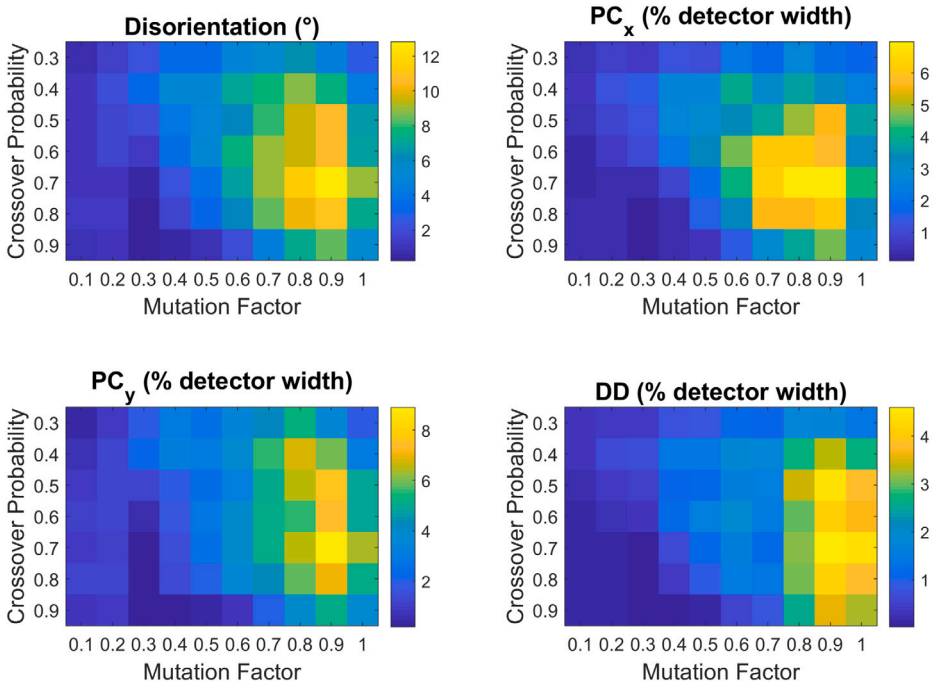


Fig. 2. Orientation and pattern center errors in relation to the crossover probability and mutation factor for the optimization (mutation scheme: DE/rand/1) performed at $N_p = 70$ and $g = 100$ within search space (±20°, ±10% detector width).

The initial population of size N_p is randomly selected from a bounded region of parameter space (the maximum range of interest) with a uniform probability distribution. For many engineering applications, N_p is typically equal to 10 times the number of parameters to be optimized (D) to sample a sufficiently large parameter space. The crossover probability $C_r \in [0, 1]$ is often lower than 0.3 but increasing it to above 0.8 could help diversify the trial vector and achieve convergence. In addition, the mutation factor F is usually $\in [0.5, 1]$ to allow sufficient mutation to cover as much optimization landscape as possible. The number of generations set the terminating criterion which depends on the required level of accuracy for a given search space. If the search space is increased significantly, the number of generations required to achieve a similar level of accuracy as the smaller search space will need to be increased or the population size needs to be increased accordingly.

For optimal performance in this particular problem of orientation and pattern center refinement ($D = 6$), the above rules of thumb provide an enormously large hyperparameter space that could be fine tuned. Since the convergence accuracy would generally improve with a sizable population and a large number of generations, the effect of crossover probability and mutation factor was first tested with a relatively large population size (N_p=70), number of generations (g = 100) and search space (±20°, ±10% detector width). For calibrated EBSD system, the orientation and pattern center accuracy will be at least an order of magnitude better than the search space used here. The large range of search space is used here primarily as a proof of concept study for testing the optimization algorithm. The default mutation scheme DE/rand/1 was used here and the effect of other mutation schemes will be discussed in the next section. Due to the stochastic nature of the optimization, ten different orientations were randomly selected

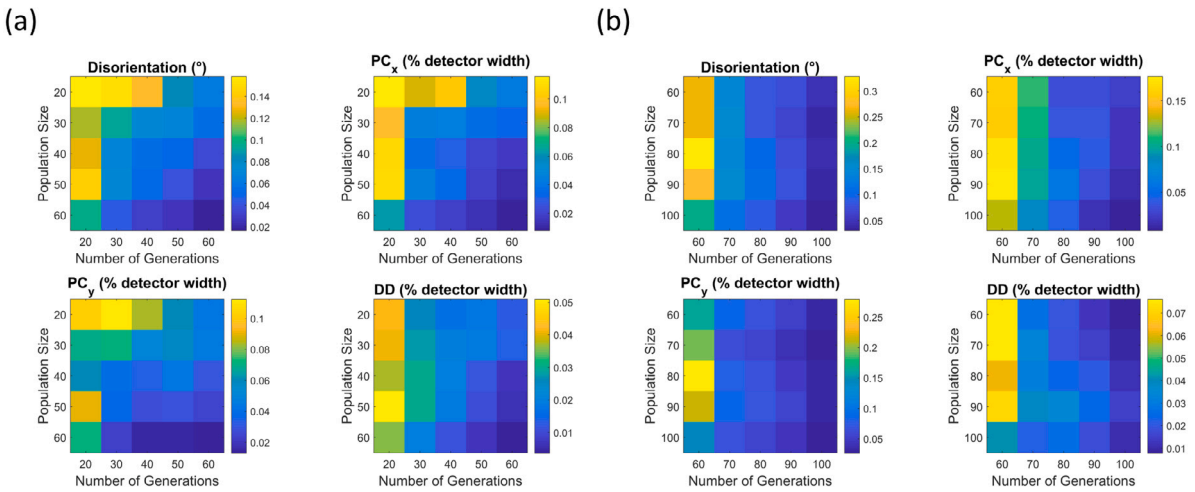


Fig. 3. Orientation and pattern center errors in relation to the population size and number of generations (mutation scheme: DE/rand/1) performed at $C_r = 0.9$ and $F = 0.5$ within search space (a) ±0.5°, ±0.5% detector width and (b) ±10°, ±5% detector width.

to run ten times and then their error values were averaged for the hyperparameter tuning. In Fig. 2, the orientation error is represented as the disorientation with regard to the ground truth orientation, and the pattern center errors are expressed as percentage of detector width. Fig. 2 indicates that the approximate range for crossover probability is $C_r \in [0.7, 0.9]$ and the mutation factor $F \in [0.3, 0.5]$ to reach a disorientation error $< 0.1°$ and all pattern center coordinates error $< 0.1\%$.

To test the appropriate number of generations and population size, $C_r = 0.9$ and $F = 0.5$ were selected from the ranges determined in the previous test. This test was conducted in two separate search spaces: (1) Hough error space (±0.5°, ±0.5% detector width), (2) an extended error space (±10°, ±5% detector width), using the default DE/rand/1 mutation scheme. In Fig. 3(a), the optimal combination to reach a disorientation error $< 0.1°$ and all pattern center coordinate errors $< 0.1\%$ has been found to be achieved when $N_p > 30$ and $g > 30$ for the narrower search space. However, Fig. 3(b) suggests that $g = 70$ is at least required to achieve similar level of accuracy for the extended search space. In addition, increasing the number of generations has been found to be more efficient way of improving the convergence. Overall, the trend in Fig. 3 agrees with the expectation that larger population size and generations in general improves convergence accuracy. In addition, the values for N_p and g are dependent on the range of search space in order to achieve a similar level of accuracy. Hence, the computational time required to optimize orientation and pattern center can be adjusted according to the size of the search space.

Additionally, a parameter study has also been carried out for the PSO method regarding the initial inertia weight w_0 (test range: $[0.1, 1.5]$), inertia weight damping coefficient w_{damp} (test range: $[0.1, 1]$), self-cognition acceleration coefficient c_1 (test range: $[0.1, 2]$), and social cognition acceleration coefficient c_2 (test range: $[0.1, 2]$). However, our test with a search space of ±10° and ±5% detector width did not show clear improvement for the PSO method for all the different combinations of parameters while the population size (80) and the maximum number of iterations (50) were kept unchanged. The underlying reason for the relatively poor performance of PSO method will be discussed later.

3.3. Differential evolution: Mutation schemes and search space

The various mutation schemes can be categorized into three groups: (1) those that will be attracted towards the best population vector, e.g., DE/best/1, DE/best/2; (2) those that have a completely random mutation, e.g., DE/rand/1, DE/rand/2; and (3) one that has a mixed characteristics from (1) and (2), e.g., DE/rand-to-best/1. Here, five types of mutations schemes are directly compared with each other as a function of the number of generations using the optimization parameters selected in Section 3.2: $N_p = 60$, $C_r = 0.9$ and $F = 0.5$. To consider more realistic intensity distribution collected on a commercial EBSD system, several filters were applied to the simulated patterns: high-pass filter (0.05), adaptive histogram equalization with 10 regions and Gaussian white noise with peak signal-to-noise ratio of 11.72 dB. Tests were done with three different search spaces: Fig. 4(b) ±0.5°, ±0.5% detector width, Fig. 4(c) ±10°, ±5% detector width and Fig. 4(d) ±20°, ±10% detector width.

Within the Hough error search space, all mutation schemes seem to perform better than Hough indexing as shown in Fig. 4(b) with an average accuracy for orientation slightly below 0.01° and pattern center accuracy in the range between 0.01% to 0.1% detector width. For the intermediate search space as shown in Fig. 4(c), the accuracy starts to become more dependent on the mutation scheme since the speed at which each method converges while using the same optimization parameters varies. Overall, mutation schemes involving some component of the best solution converge faster than the random mutation methods. In the largest search space, as shown in Fig. 4(d), the mutation scheme involving only the best solution (DE/best/1/bin and DE/best/2/bin) tend to bias the overall population towards a local minimum whereas the mutation schemes with a certain random character (DE/rand/1/bin, DE/best-to-rand/1/bin) are more likely to converge to the correct ground truth, except that the DE/rand/2/bin mutation scheme does not converge within 100 generations. The optimal combination according to the test is then DE/best-to-rand/1/bin since it converges faster than the DE/rand/1/bin method to the correct global minimum.

Additional insight has also been obtained regarding the effect of noise on the accuracy of the optimization. As shown in Fig. 5, the amount of noise present does not affect the accuracy of the DE optimization results for both orientation and pattern center, showing excellent robustness against noise. From Fig. 5(e), the orientation error is around 0.02° to 0.04° and the pattern center errors vary quite substantially between 0.001% detector width and 0.04% detector width. In particular, the detector distance has a higher accuracy compared to the other components. The variation in the accuracy of the results is likely due to the random sampling in the initial population as well as the slight modification introduced to the optimization landscape by the added noise.

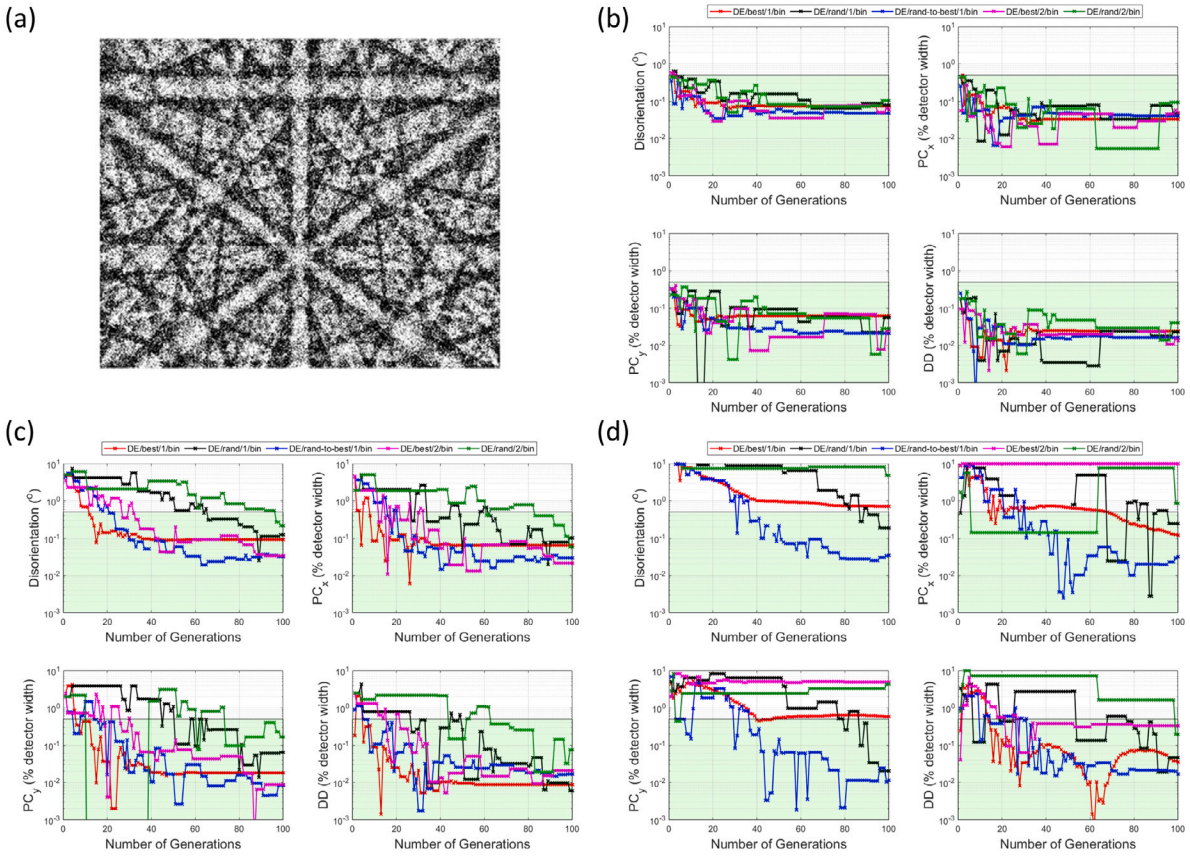


Fig. 4. (a) Simulated EBSP for Ni at 10 kV with high pass filter, adaptive histogram equalization and Gaussian white noise with peak signal-to-noise ratio of 11.72 dB; orientation and pattern center errors using different mutation schemes for the optimization: DE/rand/1, DE/rand/2, DE/best/1, DE/best/2, DE/rand-to-best/1) performed at N_p=60, C_r=0.9 and $F = 0.5$ within search space (b) ±0.5°, ±0.5% detector width, (c) ±10°, ±5% detector width and (d) ±20°, ±10% detector width.

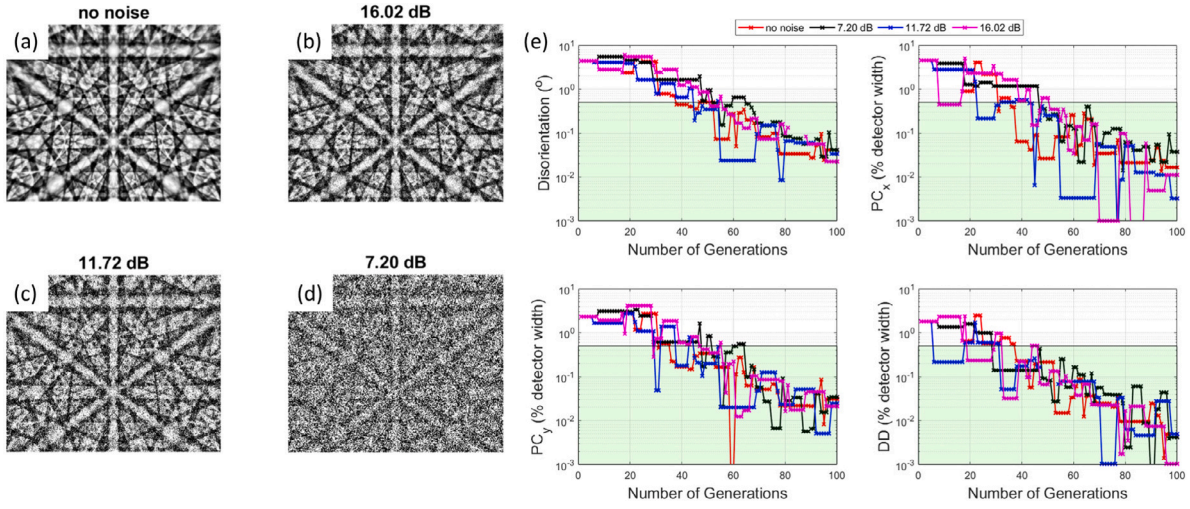


Fig. 5. (a–d) Simulated EBSPs for Ni at 10 kV with high pass filter, adaptive histogram equalization and Gaussian white noise with peak signal-to-noise ratio varying from 7.2 dB to 16.02 dB. Other optimization parameters include mutation scheme DE/rand/1, maximum number of iterations = 100, N_p = 60, C_r = 0.9 and $F = 0.5$ within search space of ±10°, ±5% detector width; (e) error values for optimized orientation and pattern center.

3.4. Large field of view line scan on single crystal nickel: hybrid optimization methods compared

Combining either particle swarm optimization or differential evolution with the Nelder–Mead simplex, two hybrid optimization methods (1. PSO + NMS; 2. DE + NMS) can be tested. In this study, the Nelder–Mean simplex allowed restart to improve convergence. Moreover, the stopping criterion for NMS was set to 300 objective function evaluations, which was a relatively early termination to compensate for the larger computational cost associated with the NMS part not being

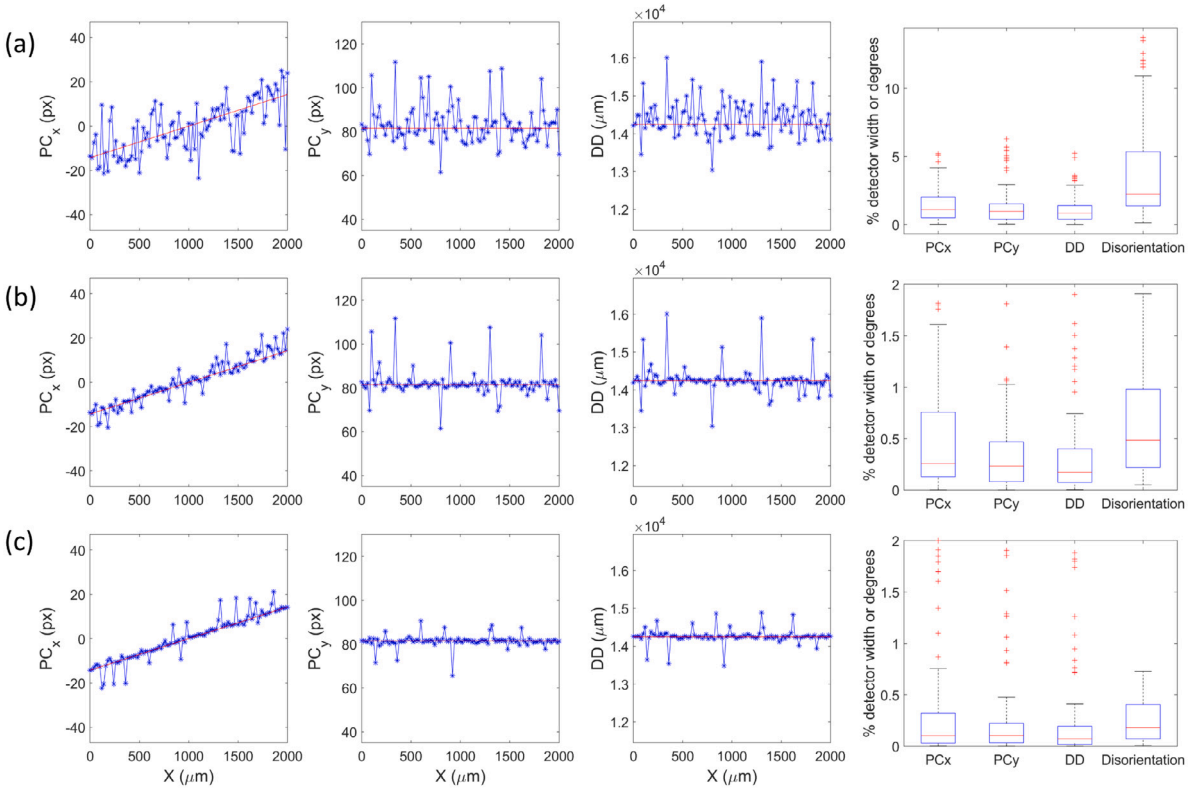


Fig. 6. (a) PSO only with a number of iterations=50; (b) PSO with a number of iterations=50 and solution refinement with NMS; (c) PSO with a number of iterations = 100 and solution refinement with NMS. Red lines are expected positions of the pattern center using EMsoft convention. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

parallelized. The convergence of the function evaluations was checked every 10 iterations to allow early termination of NMS i.e. prior to 300 function evaluations.

To compare the efficacy of these different hybrid optimization methods, a simulated data set was generated from a field of view of 2×3 mm at a sampling step size of 20 μm in both scan directions with known orientation and pattern center positions. The pattern center was updated realistically according to the geometry of the scan whereas the orientation was fixed; details regarding the simulation of this data set can be found in [34]. Optimization of the first row of the data was carried out with the initial pattern center position taken from the pattern at the center of the simulated data set. Hence, the pattern center positions of the entire row taken into optimization deviate substantially from the pattern center position of the center pattern. To make the optimization problem a bit more challenging, the initial orientation was changed arbitrarily by 5° from the true orientation.

The search space for both hybrid methods was set to be ±10°, ±5% detector width with 80 as the population (particle) number and 50 or 100 iterations for the global optimization part. For the PSO, optimization parameters were set to be: $w_0 = 1, w_{\text{damp}} = 0.9, c_1 = 2, c_2 = 2$. The DE used DE/1/best-to-rand/bin mutation scheme, crossover probability 0.9 and mutation factor 0.5.

After initial DE global optimization (100 iterations), the deviation from the ground truth has been found to be around 0.2% detector width and 0.4°. The deviation after PSO optimization (100 iterations) was on average 0.5% detector width and 0.8° with a much larger standard deviation due to the presence of outliers. The NMS step began once the global optimization has been completed after a set number of iterations. On average, the number of restarts was close to four times following PSO and only about one time following the DE optimization. Fig. 6(a–c) indicates that the use of PSO will sometimes bias the population towards a local minimum in the given search space, as shown by the local deviation from the expected red line. These outliers can be slightly refined with NMS with restarts but cannot be completely eliminated. An increase in the number of iterations of PSO from 50 to 100 still cannot completely remove the number of outliers. The hybrid method involving DE, as shown in Fig. 7(a–c), clearly shows a reduced number of outliers, particularly when the iterations for DE are increased to 100. The box plot in Fig. 7(c) shows that the average pattern center error is around 0.02% and the disorientation is around 0.04° with the DE+NMS hybrid approach for the first row of data from the large single crystal nickel sample.

3.5. *Inference of the deformation state: Simulated patterns*

To further explore the usage of the global search method, the inference of the deformation state has also been tested with the hybrid DE+NMS method proposed in this study with an objective function based on the NDP between patterns. Based on the recent study by Kurniawan et al. the same approximate model for the interpolation of simulated deformed EBSPs has been adopted [44,57]. The complexity in the optimization of deformation lies in the simultaneous inclusion of pattern center coordinates, which flattens out the optimization landscape. This has been previously found to lower the accuracy of the inferred deformation state using the PSO+NMS technique [44]. More specifically, the previous work suggested that the mapping of the spatial distribution of deformation field was almost impossible due to the overwhelmingly large relative error. This means that the sensitivity of the pattern matching approach for determining the deformation state from noisy simulated patterns is likely only suitable when a given deformation is larger than the deformation field surrounding a single edge dislocation. In this part of the study, a similar edge dislocation data set from [57] was used (without the Monte Carlo background intensity) but the overall magnitude of each component of the distortion has been increased by a factor of 100 and the step size has also been scaled by 100. Additionally, gamma correction ($\gamma = 0.333$)

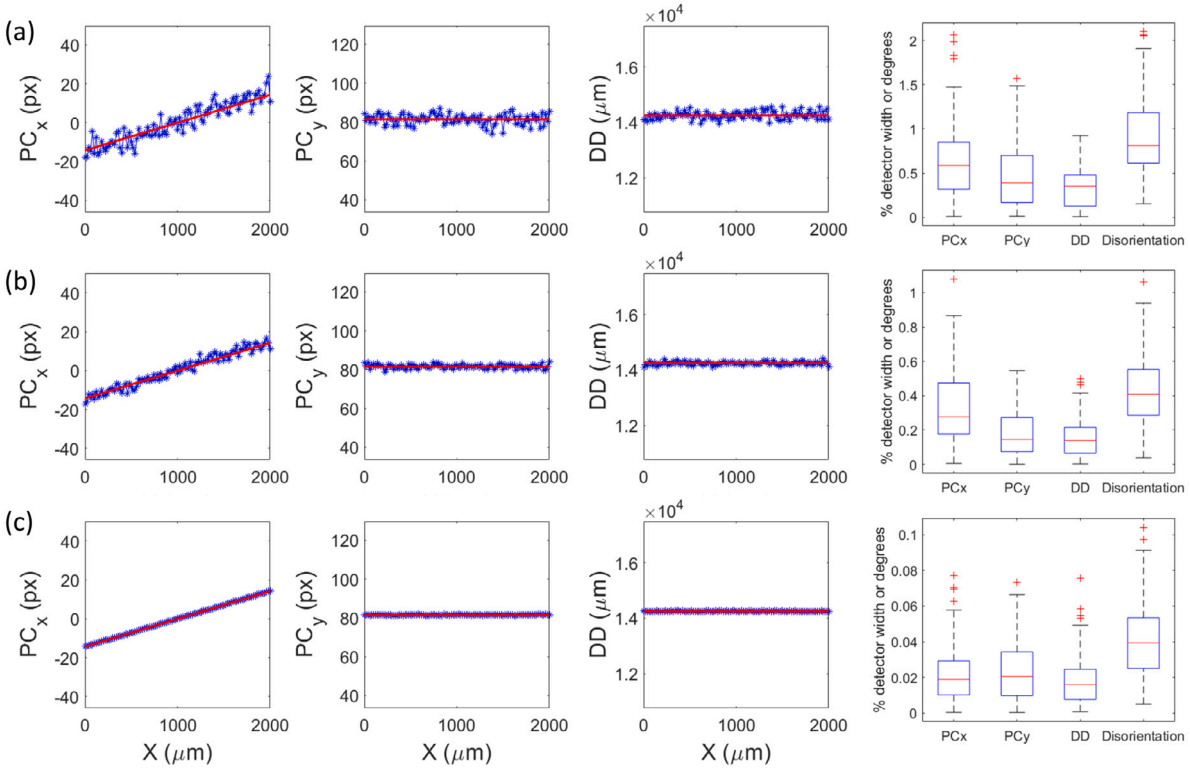


Fig. 7. (a) DE only with a number of iterations = 50; (b) DE with a number of iterations = 50 and solution refinement with NMS; (c) DE with a number of iterations = 100 and solution refinement with NMS. Red lines are expected positions of the pattern center using EMsoft convention. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and Gaussian white noise (16.02 dB) were applied to the simulated patterns. Furthermore, realistic pattern center coordinates have been incorporated into the simulation of these diffraction patterns according to the step size and the spatial coordinates. It is important to realize that both the magnitude of the distortion field and the step size are not realistic values for a single edge dislocation but these values have been chosen for the sake of expanding both the numerical range of deformation state and pattern center in order to obtain a realistic estimate of the accuracy of the proposed deformation inference method. Details regarding how the surface dislocation field is generated can be found in the supplementary information of a recent publication by Zhu et al. [57]. The expected strain and rotation field are shown in Fig. 8(a) where the extra-half plane of atom creates a compressive strain ϵ_{11} near the top half of panel. Although the data set was taken from a free surface ($z = 0$), the presence of non-zero ϵ_{33} components was indicative of the specific Poisson effect for the edge dislocation [58].

The DE part uses the DE/1/best-to-rand/bin mutation scheme, N_p=120, C_r = 0.9, F = 0.5, and g = 150. In addition, the NMS refinement limit was set to 300 objective function evaluations. The search space for the distortion tensor components ranges from −0.05 to 0.05 and the search space for the pattern center varies from ±0.5% detector width. The traction free boundary condition was applied post-optimization for the diagonal distortion tensor components, following Zhu et al. [57]. As shown in Fig. 8(b), the distribution of inferred strain and rotation field using hybrid optimization closely match the expected values given in Fig. 8(a). The most accurate components being inferred are the shear strain components and the rotation components (mean error ≈ 0.001) whereas the diagonal normal strain terms have rather poor accuracy (mean error ≈ 0.002). In addition, the accuracy of the pattern center coordinates is around 0.1% of the detector width.

3.6. Inference of deformation state: Experimental patterns

The experimental data set was collected at an accelerating voltage of 20 kV from a TRIP steel with composition Fe–17Mn-1.5Al-0.3C after fatigue cycling [59]. The elastic constants used in this study correspond to austenite with C_{11} = 206 GPa, C_{12} = 133 GPa and C_{44} = 119 GPa when we imposed the traction free boundary condition [59]. The experimental patterns were background corrected high quality patterns with a pattern size of 956 by 956 pixels. The scan was specifically focused on a large grain with clear dislocation structures, as shown in Fig. 11(b), which was then analyzed in this work with conventional HR-EBSD method using in-house code [57,60]. To address the problem of the unknown deformation state of the reference pattern, a three-step procedure is proposed to overcome this challenge: (1) optimization of Euler angles and pattern center determined from Hough indexing for the selected reference pattern; (2) inference of deformation state and pattern center with fixed Euler angles from (1) for the selected reference pattern; (3) application of cross-correlation based HR-EBSD to determine the deformation state of all the other points with respect to the reference pattern used in (2) and application of the corresponding strain state correction.

Since only the variation of lattice rotation represents useful deformation information, e.g., geometrically necessary dislocations, the lattice rotation field can be determined entirely from the HR-EBSD analysis in (3) without further refinement from the optimization routine. Therefore, the correction is only applied to the strain field in which the absolute values of the strain maps are unavailable with the conventional HR-EBSD technique. Prior to the three-step procedure, appropriate pattern filtering parameters have been determined, including adaptive histogram equalization parameter and high pass filter parameter. The high pass filter parameter removes low frequency background intensity gradients and enhances the band features. The adaptive histogram equalization routine adjusts the overall contrast of experimental to match with the simulated pattern in order to prevent spurious effect in the similarity metric. The filtering parameters can be selected with the *EMEBSDDIpreview* function implemented in EMsoft as part of the utility functions for dictionary indexing, shown in Fig. 9(a). An example of the simulated pattern of reference pattern P is given

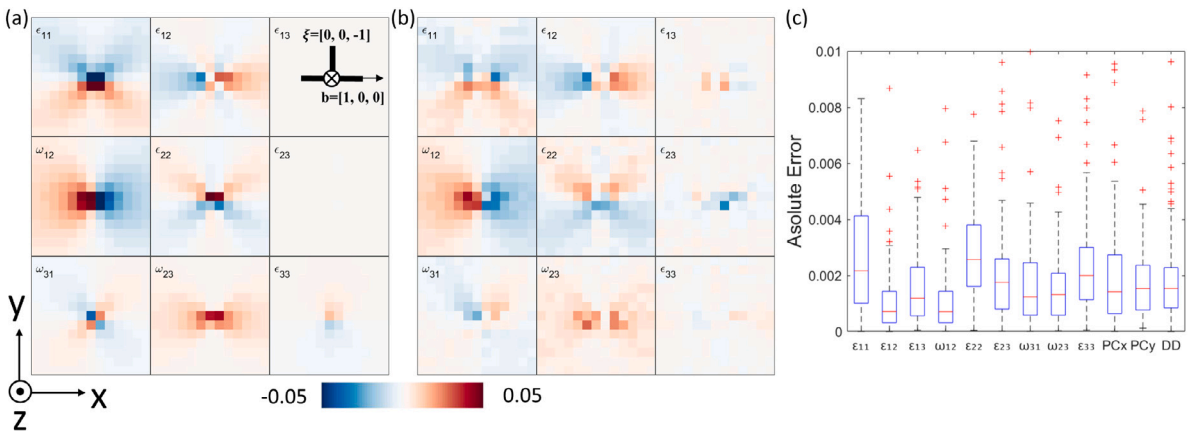


Fig. 8. Deformation tensor extraction and pattern center optimization accuracy test on a set of simulated diffraction patterns around a surface edge dislocation. (a) Expected surface ($z = 0$) strain and rotation field around the right-handed positive edge dislocation. Schematic of the dislocation configuration is overlaid on top of the ϵ_{13} component; (b) inferred surface ($z = 0$) strain and rotation field using hybrid optimization (DE+NMS); (c) absolute error of the strain, rotation, pattern center coordinates.

in Fig. 9(c) with optimized orientation and pattern center coordinates. In Fig. 9(c), 20 kV of primary beam energy is used for the Monte Carlo simulation of the backscattering energy spectrum, as described in more details by Callahan and DeGraef [22]. The agreement of the intensity distribution of simulated Kikuchi pattern to the experimental pattern shown in Fig. 9(c) is clear. The optimized parameters are used to re-simulate the same pattern P with an improved similarity metric compared to the unrefined simulated pattern (NDP improves from 0.89 to 0.93).

To further validate this approach, the following test has been carried out. The idea is that if the strain correction for the reference pattern were applied correctly, then the strain field determined from different reference patterns would be similar. Hence, the same TRIP steel data set was analyzed using the proposed three-step approach with two different reference patterns P and Q as shown in Fig. 10. In comparison, the uncorrected strain fields determined from reference patterns P and Q are relatively uniform across the grain, indicating rather small plastic deformation incurred during fatigue testing (hence smaller elastic strain to maintain compatibility). One specific component that is obviously different is ϵ_{12}. After strain state correction for the reference patterns, the ϵ_{12} components highlighted in Fig. 10 match up. This indicates that the inference of the deformation state of reference pattern will bring the strain maps determined from different reference patterns closer to the absolute strain states. However, this does not mean that these strain maps are equivalent to absolute strain maps since the deviation of the determined value to the ground truth is still impossible to determine. Due to this reason, the validation method is still prone to systemic error i.e. an unknown amount of shift to the true values.

Furthermore, the inference of the deformation states at P and Q both reveal non-zero surface shear strain ϵ_{13}, which is likely associated with the uncertainty in the geometry of the scan. Furthermore, the lattice rotation fields obtained based on two different reference patterns are clearly different in terms of their absolute values but their corresponding gradients in the lattice rotation fields are almost identical, i.e., identical distributions of geometrically necessary dislocations. In Fig. 11(c), the dislocation structures has been clearly mapped using the orientation gradients computation with significantly better sensitivity. To test the precision of the orientation mapping of experimental data in comparison of other techniques, grain reference orientation distribution (GROD: deviation of the local misorientation angle to the mean orientation in the grain) maps can be calculated and compared. Compared with the GROD maps generated by EDAX and refined dictionary indexing method, as shown in Fig. 12, the smoothly varying spatial distribution of GROD and the histogram of the global optimization based GROD map shows better precision compared to Hough transform and similar performance to the refined dictionary indexing method.

4. Discussion

In this study, a rotation correction step has been implemented to improve the optimization landscape along the orientation space. With the implementation of the rotation correction method, it was found that this rotation correction approach only slightly improves the rate of convergence with limited benefits for the accuracy of both DE and PSO methods. In addition, it also seems to incorrectly bias NMS away from the ground truth when the initial pattern center is far away from the ground truth. Therefore, inclusion of the rotation correction should only be applied within the global optimization part.

While the global optimization does not require explicit knowledge regarding the computation of derivatives in the optimization landscape, the convergence behavior heavily relies on the combination of optimization parameters. In this study, we have explored some important aspects of using global optimization that have not bee previously studies by Tanaka et al.[36]: (1) effect of changing optimization parameters on the global optimization, (2) the use of different mutation schemes, (3) direct inference of the deformation tensor. While the performance of refinement results largely depend on many factors such as the optimization parameters, the amount of noise level added, the accuracy of the initial estimate, etc, the overall orientation and pattern center results showed similar accuracy compared with other studies [33–36].

More specifically, larger search spaces require a larger population size to sample enough optimization landscape and more iterations to converge to the global minimum. Similarly, the performance of mutation schemes also depends on the range of the search space. For a small search space, $\pm 0.5°$ and $\pm 0.5\%$ detector width, these five mutation schemes have similar convergence rates. With large search spaces, the mutation scheme involving the best performing vector become noticeably faster than other schemes. However, these faster mutation schemes, i.e., DE/best/1/bin and DE/best/2/bin, are more prone to get stuck in local minima, as shown clearly in Fig. 4. Therefore, the DE/rand-to-best/1/bin is likely the most efficient mutation scheme compared to the others with regard to the convergence rate and accuracy.

For the PSO method, our hyperparameter tuning did not reveal clear improvement of the results. In addition, it was found that if the random initialization of the initial population was re-seeded several times, the global minimum can sometimes be obtained but not always guaranteed. Based on these observations, the PSO method for a large search space potentially requires a much larger population size to efficiently converge to the global minimum. This drawback of the PSO method is likely related to the velocity formulation which heavily weights the best performing particles. Based on our studies of the different evolutionary based global optimization algorithms, we have found that care must

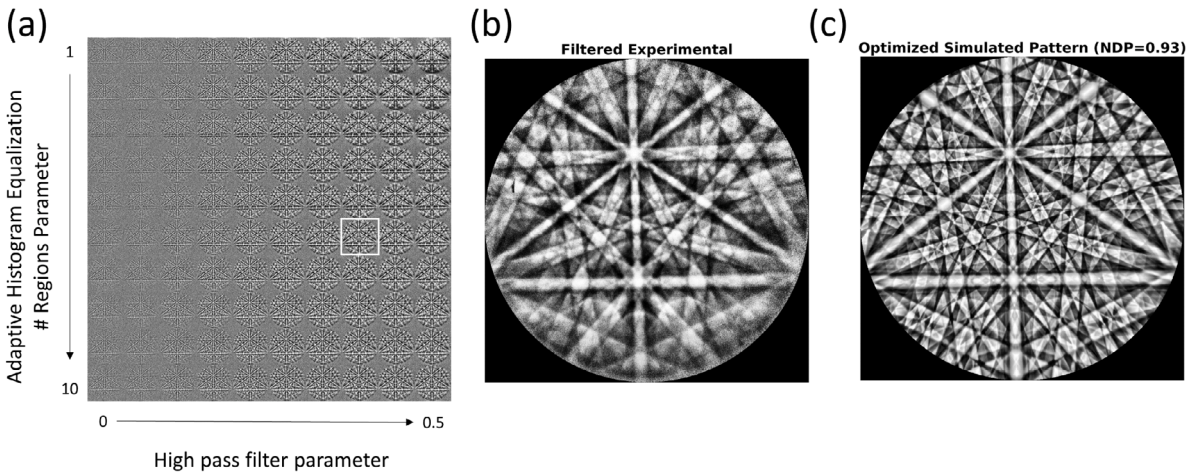


Fig. 9. (a) A matrix array of pre-processed patterns for the reference pattern P selected from the TRIP steel data set; the high-pass parameter varies from left to right, the number of regions in the adaptive histogram equalization from bottom to top. (b) Filtered experimental pattern, outlined by the white box in (a). (c) Simulated pattern for TRIP steel at 20 kV of P based on the filtering parameters from (b) with orientation and pattern center valued optimized using hybrid optimization (DE+NMS).

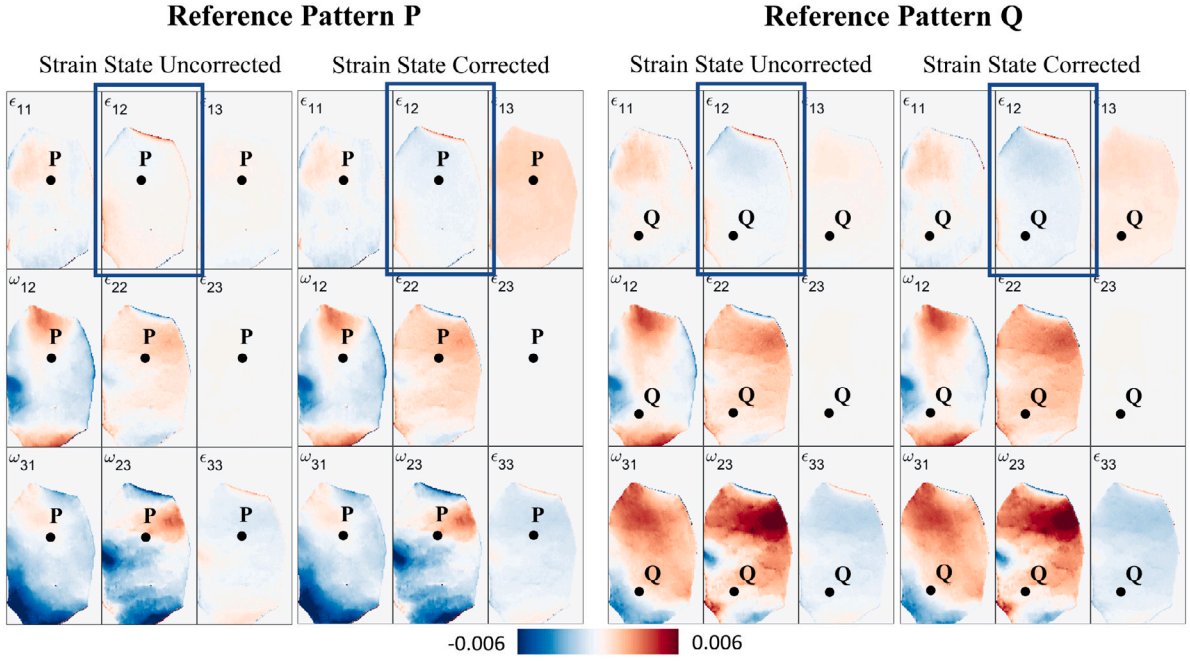


Fig. 10. HR-EBSD analysis of TRIP steel sample directly using reference pattern P and Q (strain state uncorrected) or combined with hybrid optimization for reference pattern's deformation state correction (strain state corrected).

be taken to select the optimization parameters, preferably via some hyperparameter tuning, as well as the appropriate choice of heuristic rules. A comprehensive understanding of how these global optimization algorithms perform for specific problems of interest will help to more effectively navigate the search points.

For enhanced performance, hybrid optimization of orientation and pattern center using DE+NMS and PSO+NMS within a search space of $\pm 10°$ and $\pm 5\%$ detector width with the same population size (80) and number of iterations (50 or 100) have been tested on a line scan from a simulated single crystal nickel data set with realistic pattern center variation across the large field of view. The outliers in the PSO+NMS, shown in Fig. 6(c), are due to the random trapping of the entire population into local minima during the PSO step. In comparison, DE+NMS can achieve an average accuracy of 0.02% detector width and orientation error of $0.04°$ without any outliers. Therefore, the efficient convergence of global optimization is crucial prior to the NMS refinement step, which demands an accurate initial simplex.

For the determination of deformation tensor, the accuracy of the optimization tested on the simulated YSH dislocation data set indicates a clear discrepancy in the accuracy of different components of the deformation tensor. For example, the shear strain and rotation components show much a higher accuracy ≈ 0.001 compared to the normal strains ≈ 0.002 when the search space of pattern center is around $\pm 0.5\%$ detector width. In a previous study, we have shown that direct inference of deformation can achieve extremely high accuracy ($\approx 10^{-4}$) without the inclusion of a pattern center refinement [44]. Clearly, the inclusion of a pattern center results in an extremely sloppy optimization landscape that is challenging even for the global optimization algorithm. Therefore, an additional optimization step for the pattern center and orientation prior to extraction of the deformation state is added in the test for the experimental data set.

Moreover, we have tested the hybrid optimization method on a TRIP steel experimental data set. The optimization method has been combined with the HR-EBSD method. It involved first optimizing the

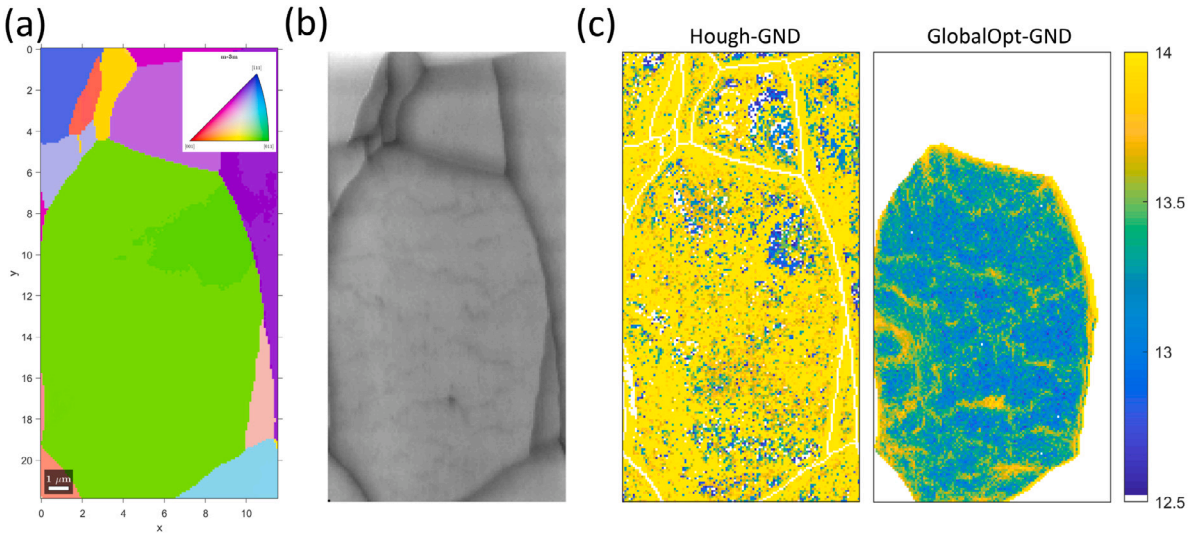


Fig. 11. (a) EBSD orientation map (IPF-Z), (b) image quality map, (c) HR-EBSD based GND density map (unit: log10 GND density).

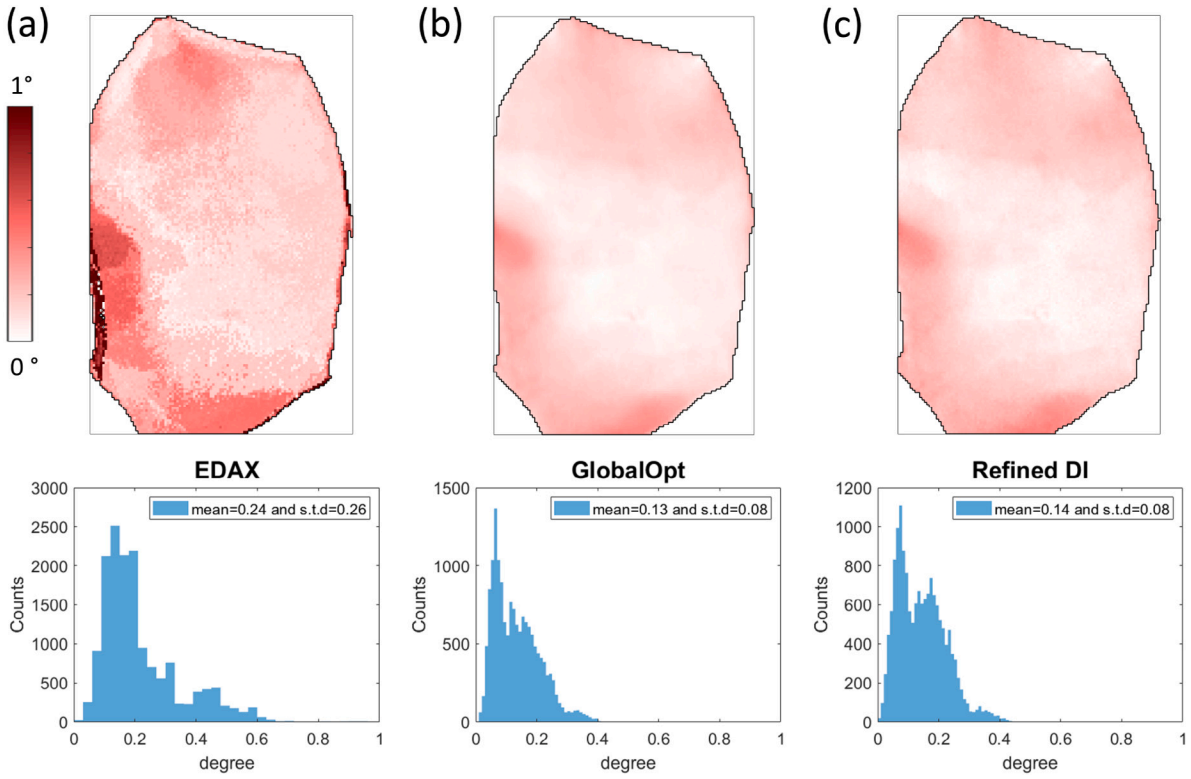


Fig. 12. Grain Reference Orientation Deviation (GROD) maps generated from (a) EDAX Hough indexing, (b) global optimization, (c) refined dictionary indexing.

orientation and pattern center of a reference pattern. Then, optimized Euler angles were fixed in the process when we inferred the deformation tensor and pattern center. Note that the deformation tensor contains a rotational part that also changes the orientation of the crystal. Therefore, the first step can probably be skipped for most well-calibrated system, which has accurate initial values of the pattern center and orientation. The second step is the crucial step to reduce error associated with pattern center due to elastic strain. The inferred strain state could then be used to correct HR-EBSD derived data with respect to the inferred reference pattern. Using the DE+NMS hybrid method to infer the strain states from different reference patterns P and Q in the TRIP steel data set, reference pattern strain states corrected strain maps indicate clear agreement in the shear strain ϵ_{12} component, as shown in Fig. 10. Nevertheless, future work is still required to further improve the sensitivity of the pattern matching approach to simultaneous deformation state extraction and pattern center refinement since the elastic strain limit of most metallic materials is about 0.002 [61]. Moreover, the experimental data used here had an unknown strain state so the demonstration here only served as an indirect approach to assess the accuracy of this "absolute" strain method. The difference between the strain state corrected strain maps using reference patterns P and Q showed mean strain errors about 0.001. Testing the efficacy of this method on a sample with known strain states, e.g., an epitaxial Si_xGe_{1-x} thin film deposited on Si substrate, is currently in progress, which will help to understand the accuracy of the inferred deformation

state and the variation of inferred pattern center across a large field of view.

Last but not least, the absolute orientation error of this technique has not been tested experimentally since perfect setup geometrically is a non-trivial exercise. Simultaneous optimization of orientation and pattern center using the incorrect sample-detector orientation relationship results in inaccurate orientation matrix with disorientation error equal to the tilt error, as shown in Fig. A.1 in the Appendix A. The pattern center remains quite accurate i.e. less than 0.05% detector with this level of uncertainty in the tilt angle, which is generally acceptable for elastic strain measurement [30,62]. In addition, the uncertainty in the tilt will also slightly bias the traction free boundary condition but the effect of such bias will be generally small on the elastic strain tensor given that the tilt error is not significant and only the normal stress is set to zero [63]. An additional complexity arose when we tested the orientation dependency of the error in the deformation tensor using simulated patterns; it is not clear why certain orientations had a slightly larger error in the deformation tensor. Future work is needed to further explore and address these issues.

5. Summary

In this study, global optimization methods such as DE and PSO have been implemented to refine orientation and pattern center for EBSD data so as to address the sloppy optimization problem. The direct inference of the deformation state is also made possible with the introduction of an approximate model for deformation tensor inclusion in the pattern simulation. It has been found that the DE algorithm is a more efficient algorithm compared to PSO based on hyperparameter tuning studies. In Section 3.2, it has been found that the reasonable ranges for the optimization parameters of DE are crossover probability $C_r \in [0.7, 0.9]$ and mutation factor $F \in [0.3, 0.5]$. In addition, the effect of using different mutation schemes and search spaces on DE has also been explored, which suggests that DE/best-to-rand/1/bin is the optimal strategy. As a general rule of thumb, a population size $N_p = 10D$ (D being the dimension of the problem) and a number of generations or iterations g=100 are sufficiently large for deformation tensor inference. For orientation and pattern center refinement, validation test on a set of realistically simulated undeformed single crystal nickel patterns indicates a mean accuracy of $\approx 0.03°$ and ≈0.01% detector width across a large field of view. For deformation inference, a hybrid strategy has been adopted by adding the NMS step after the global optimization to further refine our solution. The validation using noisy simulated deformed patterns with known deformation state and pattern center shows that the mean accuracy of shear strain and rotation components is ≈0.001 and for the normal strain ≈0.002. An indirect experimental validation of the deformation tensor was conducted on the TRIP steel sample undergone low cycle fatigue test, showing nice agreement of elastic strain fields from HR-EBSD after strain state correction.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

C. Zhu and M. De Graef acknowledge funding from a DoD Vannevar-Bush Faculty Fellowship (N00014-16-1-2821), as well as the computational facilities of the Materials Characterization Facility at CMU under grant # MCF-677785. C. Kurniawan and M. Ochsendorf, acknowledge financial support from the National Science Foundation, United States under award # DMR-1904629. D. An and S. Zaefferer acknowledge funding by the German Research Foundation [Deutsche Forschungsgemeinschaft (DFG)] through SFB 761 “Steel ab Initio”.

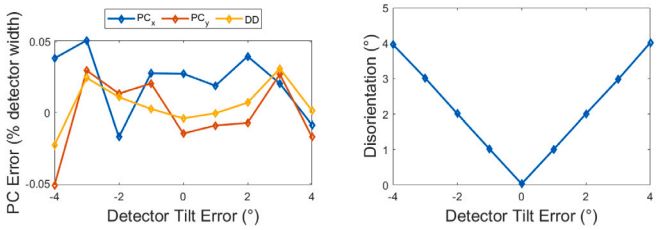


Fig. A.1. Effect of detector tilt error on the accuracy of (left) pattern center error (right) orientation error.

Appendix A

Definition of similarity metrics

The four types of image similarity metrics used in 3.1 are defined in the following. Suppose we have an image **A** and an image **B** with the same dimensions of i by j.

The normalized dot product (NDP) is defined as:

$$NDP = \frac{\sum_{i,j} A_{i,j} B_{i,j}}{\|\mathbf{A}\|\|\mathbf{B}\|}, \tag{12}$$

where $\|\mathbf{A}\| = \sqrt{\sum_{i,j} A_{i,j}^2}$.

The cross-correlation coefficient (XCF) is defined as:

$$XCF = \frac{\sum_{i,j}(A_{i,j} - \overline{\mathbf{A}})(B_{i,j} - \overline{\mathbf{B}})}{\|\mathbf{A} - \overline{\mathbf{A}}\|\|\mathbf{B} - \overline{\mathbf{B}}\|}, \tag{13}$$

where $\|\mathbf{A} - \overline{\mathbf{A}}\| = \sqrt{\sum_{i,j}(A_{i,j} - \overline{\mathbf{A}})^2}$ and $\overline{\mathbf{A}}$ is the mean of **A**.

The root-mean-square-error (RMSE) is defined as:

$$RMSE = \sqrt{\frac{\sum_{i,j}(A_{i,j} - B_{i,j})^2}{N}}, \tag{14}$$

where N is the number of pixels of image **A** or **B**.

The mutual information involves the use of Shannon entropy:

$$H = -\sum_i p_i \log p_i, \tag{15}$$

where p_i is the probability distribution of the intensity $i \in [0, \ldots, 255]$ of a gray-scale image between. Similarly, the Shannon entropy of a joint distribution (a measure of image registration) is defined as:

$$H = -\sum_{i,j} p_{i,j} \log p_{i,j}. \tag{16}$$

The mutual information MI(A,B) between images A and B can be established as the sum of Shannon entropy values of the two images minus the joint entropy.

$$MI(A, B) = H(A) + H(B) - H(A, B). \tag{17}$$

It is clear that the optimization of these objective function values require either maximizing (NDP, XCF, MI) or minimizing (RMSE) the corresponding metric for the pattern matching problem.

Effect of tilt error on pattern center and orientation

See Fig. A.1.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ultramic.2021.113407. Brief introduction about how to use the global optimization program implemented in EMsoft. A new Python interface (pyEMsoft) has also been developed to allow easier access to some of the EMsoft’s Fortran routines.

References

[1] D.J. Dingley, On-line determination of crystal orientation and texture determination in an SEM, Proc. R. Microsc. Soc. 19 (1984) 74–75.

[2] D.J. Dingley, M. Longden, J. Weinbren, J. Alderman, Online analysis of electron back scatter diffraction patterns. 1. texture analysis of zone refined polysilicon, Scanning Microscopy 1 (2) (1987) 451–456.

[3] N.C. Krieger Lassen, Image processing procedures for analysis of electron backscattering patterns, Scanning Microscopy 6 (1992) 115–121.

[4] S.I. Wright, B.L. Adams, Automatic analysis of electron backscatter diffraction patterns, Metallurg. Trans. A 23 (3) (1992) 759–767, http://dx.doi.org/10.1007/BF02675553.

[5] A.J. Schwartz, M. Kumar, B.L. Adams, D.P. Field, Electron backscatter diffraction in materials science, Electron Backscatter Diffraction in Materials Science, Kluwer Academic, New York, 2000, http://dx.doi.org/10.1007/978-0-387-88136-2.

[6] F. Ram, S. Zaefferer, T. Jäpel, D. Raabe, Error analysis of the crystal orientations and disorientations obtained by the classical electron backscatter diffraction technique, J. Appl. Crystallogr. 48 (3) (2015) 797–813, http://dx.doi.org/10.1107/S1600576715005762.

[7] F.J. Humphreys, Quantitative metallography by electron backscattered diffraction, J. Microsc. 195 (3) (1999) 170–185.

[8] S.I. Wright, M.M. Nowell, D.P. Field, A review of strain analysis using electron backscatter diffraction, Microsc. Microanal. 17 (3) (2011) 316–329, http://dx.doi.org/10.1017/S1431927611000055.

[9] G. Nolze, Image distortions in SEM and their influences on EBSD measurements, Ultramicroscopy 107 (2) (2007) 172–183, http://dx.doi.org/10.1016/j.ultramic.2006.07.003.

[10] C. Maurice, R. Fortunier, A 3D Hough transform for indexing EBSD and Kossel patterns, J. Microsc. 230 (3) (2008) 520–529, http://dx.doi.org/10.1111/j.1365-2818.2008.02045.x.

[11] K. Thomsen, N.H. Schmidt, A. Bewick, K. Larsen, J. Goulden, Improving the accuracy of orientation measurements using EBSD, Microsc. Microanal. 19 (S2) (2013) 724–725.

[12] I. Brough, P.S. Bate, F.J. Humphreys, Optimising the angular resolution of EBSD, Mater. Sci. Technol. 22 (11) (2006) 1279–1286, http://dx.doi.org/10.1039/an9830800615.

[13] A.J. Wilkinson, G. Meaden, D.J. Dingley, High-resolution elastic strain measurement from electron backscatter diffraction patterns: New levels of sensitivity, Ultramicroscopy 106 (4–5) (2006) 307–313, http://dx.doi.org/10.1016/j.ultramic.2005.10.001.

[14] N.C.K. Lassen, J.B. Bilde-Sø rensen, Calibration of an electron back-scattering pattern set-up, J. Microsc. 170 (2) (1993) 125–129.

[15] C. Maurice, K. Dzieciol, R. Fortunier, A method for accurate localisation of EBSD pattern centres, Ultramicroscopy 111 (2) (2011) 140–148, http://dx.doi.org/10.1016/j.ultramic.2010.10.007.

[16] D.A. Carpenter, J.L. Pugh, G.D. Richardson, L.R. Mooney, Determination of pattern centre in EBSD using the moving-screen technique, J. Microsc. 227 (3) (2007) 246–247, http://dx.doi.org/10.1111/j.1365-2818.2007.01807.x.

[17] K. Mingard, A. Day, C. Maurice, P. Quested, Towards high accuracy calibration of electron backscatter diffraction systems, Ultramicroscopy 111 (5) (2011) 320–329.

[18] D.J. Dingley, K. Baba-Kishi, Use of electron backscatter diffraction patterns for determination of crystal symmetry elements, Scanning Electron. Microscopy 2 (1986) 383–391.

[19] J. Kacher, C. Landon, B.L. Adams, D. Fullwood, Bragg's law diffraction simulations for electron backscatter diffraction analysis, Ultramicroscopy 109 (9) (2009) 1148–1156, http://dx.doi.org/10.1016/j.ultramic.2009.04.007.

[20] A. Howie, M.J. Whelan, Diffraction contrast of electron microscope images of crystal lattice defects-II. The development of a dynamical theory, Proc. R. Soc. Lond. Ser. A 263 (1313) (1961) 217–237.

[21] A. Winkelmann, C. Trager-Cowan, F. Sweeney, A.P. Day, P. Parbrook, Many-beam dynamical simulation of electron backscatter diffraction patterns, Ultramicroscopy 107 (4–5) (2007) 414–421, http://dx.doi.org/10.1016/J.ULTRAMIC.2006.10.006.

[22] P.G. Callahan, M. De Graef, Dynamical electron backscatter diffraction patterns. Part I: Pattern simulations, Microsc. Microanal. 19 (5) (2013) 1255–1265, http://dx.doi.org/10.1017/S1431927613001840.

[23] 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 backscatter diffraction indexing, Microsc. Microanal. 21 (3) (2015) 739–752, http://dx.doi.org/10.1017/S1431927615000756.

[24] A. Foden, D.M. Collins, A.J. Wilkinson, T.B. Britton, Indexing electron backscatter diffraction patterns with a refined template matching approach, Ultramicroscopy 207 (2019) 112845.

[25] A.J. Wilkinson, D.M. Collins, Y. Zayachuk, R. Korla, A. Vilalta-Clemente, Applications of multivariate statistical methods and simulation libraries to analysis of electron backscatter diffraction and transmission kikuchi diffraction datasets, Ultramicroscopy 196 (2018) 88–98, http://dx.doi.org/10.1016/j.ultramic.2018.09.011.

[26] W.C. Lenthe, S. Singh, M. De Graef, A spherical harmonic transform approach to the indexing of electron back-scattered diffraction patterns, Ultramicroscopy 207 (2019) 112841.

[27] S. Singh, Y. Guo, B. Winiarski, T.L. Burnett, P.J. Withers, M. De Graef, High resolution low kV EBSD of heavily deformed and nanocrystalline Aluminium by dictionary-based indexing, Sci. Rep. 10991 (8) (2018) http://dx.doi.org/10.1038/s41598-018-29315-8.

[28] Z. Ding, E. Pascal, M. De Graef, Indexing of electron back-scatter diffraction patterns using a convolutional neural network, Acta Mater. (2020).

[29] R. Hielscher, F. Bartel, T.B. Britton, Gazing at crystal balls – Electron backscatter diffraction indexing and cross correlation on a sphere, 2019/08/05, Microsc. Microanal. 25 (S2) (2019) 1954–1955, http://dx.doi.org/10.1017/S143192761901050X.

[30] T.B. Britton, C. Maurice, R. Fortunier, J.H. Driver, A.P. Day, G. Meaden, D.J. Dingley, K. Mingard, A.J. Wilkinson, Factors affecting the accuracy of high resolution electron backscatter diffraction when using simulated patterns, Ultramicroscopy 110 (12) (2010) 1443–1453, http://dx.doi.org/10.1016/j.ultramic.2010.08.001.

[31] W. Lenthe, S. Singh, M. De Graef, Prediction of potential pseudo-symmetry issues in the indexing of electron backscatter diffraction patterns, J. Appl. Crystallogr. 52 (5) (2019).

[32] E.L. Pang, P.M. Larsen, C.A. Schuh, Resolving pseudosymmetry in tetragonal ZrO2 using electron backscatter diffraction with a modified dictionary indexing approach, J. Appl. Crystallogr. 53 (4) (2020) 1060–1072.

[33] A. Winkelmann, G. Nolze, G. Cios, T. Tokarski, P. Bała, Refined calibration model for improving the orientation precision of electron backscatter diffraction maps, Materials 13 (12) (2020) 2816.

[34] S. Singh, F. Ram, M. De Graef, Application of forward models to crystal orientation refinement, J. Appl. Crystallogr. 50 (6) (2017) 1664–1676.

[35] E.L. Pang, P.M. Larsen, C.A. Schuh, Global optimization for accurate determination of EBSD pattern centers, Ultramicroscopy 209 (2020) 112876.

[36] T. Tanaka, A.J. Wilkinson, Pattern matching analysis of electron backscatter diffraction patterns for pattern centre, crystal orientation and absolute elastic strain determination–accuracy and precision assessment, Ultramicroscopy 202 (2019) 87–99.

[37] T. Tanaka, N. Maruyama, N. Nakamura, A.J. Wilkinson, Tetragonality of Fe-C martensite–a pattern matching electron backscatter diffraction analysis compared to X-ray diffraction, Acta Mater. (2020).

[38] A. Winkelmann, G. Nolze, G. Cios, T. Tokarski, Mapping of local lattice parameter ratios by projective Kikuchi pattern matching, Physical Review Materials 2 (12) (2018) 123803.

[39] S. Villert, C. Maurice, C. Wyon, R. Fortunier, Accuracy assessment of elastic strain measurement by EBSD, J. Microsc. 233 (2) (2009) 290–301, http://dx.doi.org/10.1111/j.1365-2818.2009.03120.x.

[40] M.D. Vaudin, Y.B. Gerbig, S.J. Stranick, R.F. Cook, Comparison of nanoscale measurements of strain and stress using electron back scattered diffraction and confocal Raman microscopy, Appl. Phys. Lett. 93 (19) (2008) 193116.

[41] G. Miyamoto, A. Shibata, T. Maki, T. Furuhara, Precise measurement of strain accommodation in austenite matrix surrounding martensite in ferrous alloys by electron backscatter diffraction analysis, Acta Mater. 57 (4) (2009) 1120–1131.

[42] A.J. Wilkinson, D. Randman, Determination of elastic strain fields and geometrically necessary dislocation distributions near nanoindents using electron back scatter diffraction, Phil. Mag. 90 (9) (2010) 1159–1177, http://dx.doi.org/10.1080/14786430903304145.

[43] J. Alkorta, Limits of simulation based high resolution EBSD, Ultramicroscopy 131 (2013) 33–38.

[44] C. Kurniawan, C. Zhu, M. DeGraef, Deformation state extraction from electron backscatter diffraction patterns via simulation-based pattern-matching, Scr. Mater. 190 (2021) 147–152, http://dx.doi.org/10.1016/j.scriptamat.2020.09.004.

[45] E.H. Yoffe, A dislocation at a free surface, Phil. Mag. 6 (69) (1961) 1147–1155.

[46] S.J. Shaibani, P.M. Hazzledine, The displacement and stress fields of a general dislocation close to a free surface of an isotropic solid, Phil. Mag. A 44 (3) (1981) 657–665.

[47] A. Morawiec, Orientations and Rotations, Springer, 2003.

[48] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, J. Global Optim. 11 (4) (1997) 341–359.

[49] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Ieee, 1995, pp. 39–43.

[50] E. Mezura-Montes, C.A.C. Coello, Constraint-handling in nature-inspired numerical optimization: past, present and future, Swarm Evol. Comput. 1 (4) (2011) 173–194.

[51] M.E.H. Pedersen, Good Parameters for Particle Swarm Optimization, Hvass Lab., Copenhagen, Denmark, Tech. Rep. HL1001, 2010, pp. 1551–3203.

[52] J.A. Nelder, R. Mead, A simplex method for function minimization, Comput. J. 7 (4) (1965) 308–313.

[53] K.I.M. McKinnon, Convergence of the Nelder–Mead simplex method to a nonstationary point, SIAM J. Optim. 9 (1) (1998) 148–158.

[54] L. Han, M. Neumann, Effect of dimensionality on the Nelder–Mead simplex method, Optim. Methods Softw. 21 (1) (2006) 1–16.

[55] F. Gao, L. Han, Implementing the nelder-mead simplex algorithm with adaptive parameters, Comput. Optim. Appl. 51 (1) (2012) 259–277.
[56] R.t. O'Neill, Algorithm AS 47: function minimization using a simplex procedure, J. R. Stat. Soc. C 20 (3) (1971) 338–345.
[57] C. Zhu, M. De Graef, EBSD Pattern simulations for an interaction volume containing lattice defects, Ultramicroscopy (2020) 113088.
[58] D. Hull, D.J. Bacon, Introduction to Dislocations, fifth ed., Elsevier, 2011.
[59] D. An, S. Zaefferer, Formation mechanism of dislocation patterns under low cycle fatigue of a high-manganese austenitic TRIP steel with dominating planar slip mode, Int. J. Plast. 121 (2019) 244–260.
[60] C. Zhu, K. Kaufmann, K.S. Vecchio, Novel remapping approach for HR-EBSD based on demons registration, Ultramicroscopy (2019) 112851.
[61] C. Maurice, J.H. Driver, R. Fortunier, On solving the orientation gradient dependency of high angular resolution EBSD, Ultramicroscopy 113 (2012) 171–181, http://dx.doi.org/10.1016/j.ultramic.2011.10.013.
[62] J. Basinger, D. Fullwood, J. Kacher, B. Adams, Pattern center determination in electron backscatter diffraction microscopy, Microsc. Microanal. 17 (3) (2011) 330–340, http://dx.doi.org/10.1017/S1431927611000389.
[63] T.J. Hardin, T.J. Ruggles, D.P. Koch, S.R. Niezgoda, D.T. Fullwood, E.R. Homer, Analysis of traction-free assumption in high-resolution ebsd measurements, J. Microsc. 260 (1) (2015) 73–85.