doi:10.1017/S1431927618001605
Microsc. Microanal. 24 (Suppl 1), 2018

GPU-Accelerated Matrix Exponentiation for 5-D STEM-DCI Simulations

Saransh Singh[1], Marc De Graef[1]

[1] Dept. of Materials Science and Engineering, Carnegie Mellon Univ., Pittsburgh, PA, USA

First order coupled partial differential equations are ubiquitous in science and engineering and form the basis for a number of mathematical models. Examples range from economic modeling, predator-prey models and supersymmetric and string theory. In the context of high energy electron microscopy, these equations describe the dynamical scattering of amplitudes between different reciprocal lattice vectors and are known as the Darwin-Howie-Whelan (DHW) equations. These equations have a well known solu-tion which involves the calculation of a matrix exponential. With continual improvements in computer hardware, it has been possible to compute these exponentials for a large number of scattered beams resulting in very accurate $n-$beam dynamical diffraction simulations. This has found application in various microscopy techniques such as Electron Back-Scatter Diffraction (EBSD), Electron Channeling Contrast Imaging (ECCI), phase contrast microscopy etc. However, with improvements in the microscope hardware, new modalities, such as $4-$D Scanning Transmission Electron Microscopy Diffraction Contrast Imaging (STEM-DCI), have emerged as new state-of-the-art tools in the field. These techniques routinely generate enormous amounts of data and require the availability of fast and efficient forward models to interpret contrast features seen in real complex microstructures. This necessitates the need to speed up the key step in the dynamical diffraction computations, namely the matrix exponentiation. The speed-up in computation time allows simulations for addition parameters to the $4-$D STEM, such as time series, load series etc. In this contribution, we discuss a GPU based matrix exponentiation algorithm, targeted mainly for the STEM-DCI modality, but easily extendable to other diffraction modalities both in the backscatter and forescatter geometries. Comparisons with existing CPU based algorithms are made showing impressive speed-ups. Finally, we will present results of STEM-DCI image calculations for a complex two phase $\gamma - \gamma'$ microstructure with misfit displacements at the interfaces.

The $n-$beam DHW equations for a single phase microstructure in the presence of lattice defects are given by a set of coupled first order differential equations, where the number n denotes the number of reflections. This equation is usually written in the concise matrix notation shown in eq. 1, where the modified wave function, $\Phi = [\phi_{\mathbf{g}_1}\, \phi_{\mathbf{g}_2} \cdots]^{\mathrm{T}}$ is a column vector with the amplitudes of the electron wave functions for each reciprocal vector; the matrix $\mathcal{A}$ contains the geometry of the diffraction (diagonal $s_{\mathbf{g}}$ terms) and coupling coefficients (off-diagonal $q_{\mathbf{g}}$ terms). The true wavefunction is related to the modi-fied wavefunction through a simple Hadamard product (denoted by $\circ$) with a column vector phase factor, $\Theta = [\theta_{\mathbf{g}_1}\ \ \theta_{\mathbf{g}_2} \cdots]^{\mathrm{T}}$. These coupled differential equations have a well known solution involving the exponential of the structure matrix and the wavefunction at the crystal-vacuum boundary ($\Phi(0)$) as shown in eq. 1.

$$\frac{\mathrm{d}\Phi}{\mathrm{d}z} = \mathrm{i}\mathcal{A}\Phi; \quad \mathcal{A}_{\mathbf{g}\mathbf{g}'} = \begin{cases} 2\pi s_{\mathbf{g}} & : \mathbf{g} = \mathbf{g}' \\ \frac{\mathrm{e}^{\mathrm{i}\alpha_{\mathbf{g}-\mathbf{g}'}}}{q_{\mathbf{g}-\mathbf{g}'}} & : \mathbf{g} \neq \mathbf{g}' \end{cases} ; \quad \Psi = \Phi \circ \Theta; \quad \Phi(z) = \mathrm{e}^{\mathrm{i}\mathcal{A}z}\Phi(0) \tag{1}$$

For parallel illumination, which is the case in a conventional transmission electron microscope (CTEM) mode, the wavefunction at the boundary is given by $\Psi(0) = \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$. Thus the boundary condition for the parallel illumination is given by $\Psi \equiv [1\ \ 0 \cdots]^{\mathrm{T}}$. For the converged probe in the STEM mode, the electron wave functions in the crystal is given by a sum of plane waves with slightly different wave vectors; $\Psi(0) = \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}}$. Therefore, the electron wave function is computed for a set of different $\mathbf{k}$ vectors. For accurate computations, the number of $\mathbf{k}$ vectors for the computations is of the order, $n_{\mathbf{k}} \approx 400$. A naive implementation of the STEM modality would increase the computation time to $n_{\mathbf{k}}$ times compared to the CTEM modality. This necessitates the case for smart approximations as well as massive parallellization

of the exponential algorithm.

The structure matrix is split up as a sum of a wavevector dependent diagonal matrix, $\mathcal{A}_d^k$ and a "*Hollow*" off-diagonal matrix, $\mathcal{A}_H$ insensitive to the beam direction. For a small slice thickness, ϵ, the exponential of the sum of these matrices can be approximated by neglecting the higher order terms in the *Zassenhaus* expansion. These expressions are given by the following equations:

$$\mathcal{A}^{\mathbf{k}} = \mathcal{A}_d^{\mathbf{k}} + \mathcal{A}_H; \qquad \mathrm{e}^{\mathrm{i}\mathcal{A}^{\mathbf{k}}\epsilon} = \mathrm{e}^{\mathrm{i}\mathcal{A}_d^{\mathbf{k}}\frac{\epsilon}{2}}\,\mathrm{e}^{\mathrm{i}\mathcal{A}_H\epsilon}\,\mathrm{e}^{\mathrm{i}\mathcal{A}_d^{\mathbf{k}}\frac{\epsilon}{2}}. \tag{2}$$

Thus, the exponential for different $\mathbf{k}$ vectors involves evaluating the exponential of off diagonal components only once. The exponential of the diagonal matrix is trivial to evaluate. In our implementation, the matrix exponential is calculated using the 9^{th} order Taylor's series expansion. The 9^{th} degree polynomial is split into three 3^{rd} order terms with complex coefficients. Thus, only $\mathcal{A}^2$ and $\mathcal{A}^3$ need to be evaluated which involves only 4 matrix products. These equations have been listed in [3]. In addition to the Taylor series approximation, scaling and squaring is used to make the method numerically stable. Since matrix multiplication is highly parallellizable on the GPU, the exponential was implemented as a series of matrix multiplications.

$$\mathrm{e}^{\mathrm{i}\mathcal{A}\epsilon} = \mathrm{I} + \sum_{n=1}^{9}\frac{\mathcal{A}^n}{n!} = \prod_{n=0}^{2}\left(\mathcal{A}^3 + c_{3n+1}\mathcal{A}^2 + c_{3n+2}\mathcal{A} + c_{3n+3}\mathrm{I}\right). \tag{3}$$

Figure 1(a) shows the initial $\gamma - \gamma'$ microstructure with 50% contour surfaces for the four translational variants in different color. The bright field image is shown in Fig. 1(b). 1(c) shows the annular dark field image with the detector geometry shown in the figure. Finally, Fig. 1(d)-(f) show the full convergent beam diffraction (CBED) pattern at location marked in 1(b).

References:

[1] S. Singh and M. De Graef, Microscopy and Microanalysis, **21**(S3), p. 2213.

[2] R.S. Pennington, F. Wang, C.T. Koch, Ultramicroscopy, **141**, p. 32.

[3] The authors acknowledge R. Shi (LLNL) for providing the phase field data set, ONR Vannevar Bush Faculty Fellowship support (N00014-16-1-2821) as well as the computational resources of the Materials Characterization Facility at CMU, grant MCF-677785.

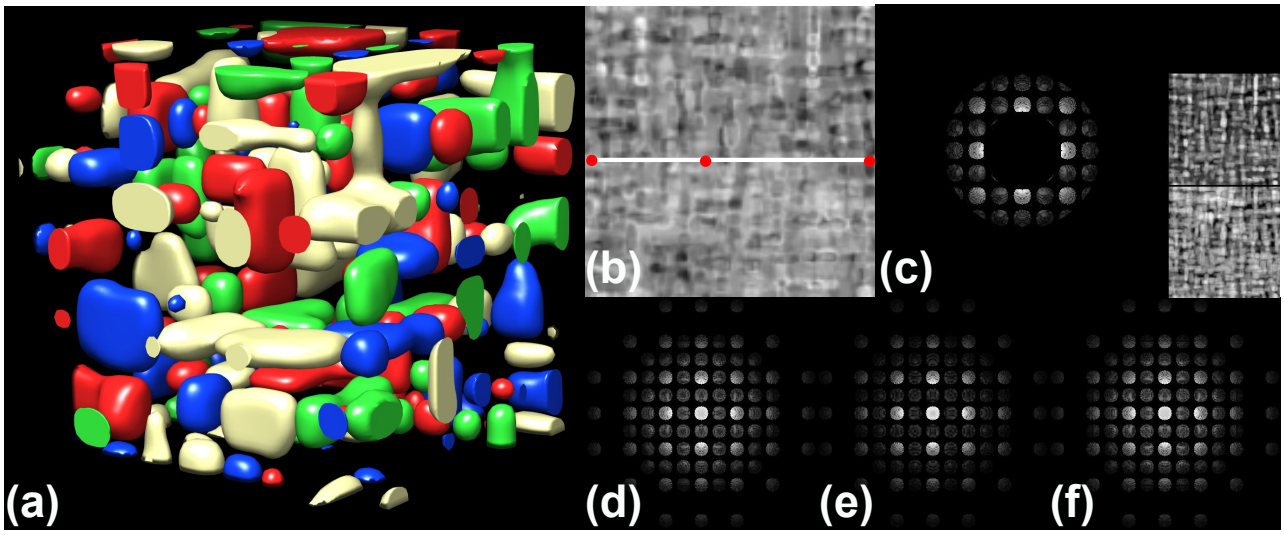

Figure 1. (a) Initial microstructure (b) bright field image (c) Annular dark field images with detector geometry and (d)-(f) full CBED patterns from locations marked in (b).

https://doi.org/10.1017/S1431927618001605 Published online by Cambridge University Press