Physics-informed Bayesian Optimization of an Electron Microscope

Desheng Ma[1], Chenyu Zhang[1], Yu-Tsun Shao[1], Zhaslan Baraissov[1], Cameron Duncan[2], Adi Hanuka[3], Auralee Edelen[3], Jared Maxson[2], David A. Muller[1,4]

1. School of Applied and Engineering Physics, Cornell University, Ithaca NY 2. Department of Physics, Cornell University, Ithaca NY 3. SLAC National Accelerator Laboratory, Menlo Park CA 4. Kavli Institute for Nanoscale Science, Cornell University, Ithaca, NY

Abstract

In this project, we design a new scheme targeting fully automated aberration corrector tuning, achieving greater speed and less human bias. Specifically, we solve three problems in aberration correction from the perspective of accelerator physics via machine learning.

- We derive the fundamental connection between the aberration function and the delivered beam emittance in phase space via Wigner distribution.
- We demonstrate a customized convolutional neural network can accurately predict emittance from experimentally accessible electron Ronchigrams.
- We design a Bayesian approach to optimize for the aberration corrector parameters for minimum emittance growth with uncertainty quantification.

From Aberration Coefficients to Beam Emittance

The standard expression of aberrations in electron microscopy is the aberration function, which characterizes the electron beam in its wave function.

$$\chi(\alpha,\phi) = \frac{2\pi}{\lambda}\sum_{n,m}\frac{C_{n,m}\alpha^{n+1}\cos\left(m(\phi-\phi_{n,m})\right)}{n+1} \overset{or}{=} \frac{2\pi}{\lambda}\sum_{n,m}\frac{\alpha^{n+1}}{n+1}\left|\tilde{C}^*_{n,m}e^{im(\phi-\phi_{n,m})}\right|$$

$$\psi_{abr}(\vec{\alpha}) = \psi_0(\vec{\alpha})e^{i\chi(\vec{\alpha})}$$

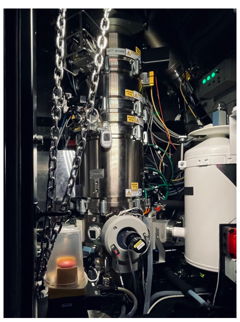
Wigner-Weyl transform of the electron wave function approximates the shift of the beam to the gradient of the aberration function.

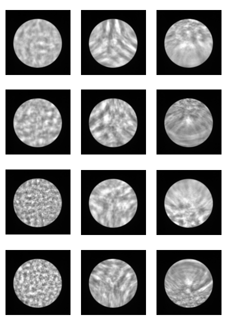
$$W(\vec{\xi},\vec{\alpha}) = \frac{1}{4\pi^2}\int d\alpha'^2\tilde{\psi}^*\left(\vec{\alpha}-\frac{1}{2}\vec{\alpha}'\right)\tilde{\psi}\left(\vec{\alpha}+\frac{1}{2}\vec{\alpha}'\right)e^{i\vec{\xi}\cdot\vec{\alpha}'}$$

$$\Rightarrow \quad \vec{\xi}(\vec{\alpha}) \approx \vec{\xi}(x',y') \approx \nabla\chi(x',y')$$

Therefore, we can define the root-mean-square (RMS) emittance of the electron beam as,

$$\varepsilon_\chi = \sqrt{\langle\nabla\chi^2\rangle\langle\vec{\alpha}^2\rangle - \langle\nabla\chi\cdot\vec{\alpha}\rangle^2} = \sqrt{\det\left(\mathrm{var}(\nabla\chi,\vec{\alpha})\right)}$$



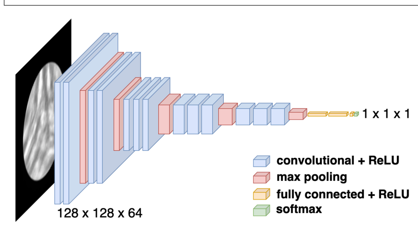


A1	A2*	B2**
50 nm	250 nm	500 nm
-50 nm	-250 nm	-500 nm
100 nm	500 nm	1000 nm
-100 nm	-500 nm	-1000 nm

Fig. 1. Spectra 300 installed in Duffield Hall, Cornell University.
Fig. 2. Ronchigrams collected from the Spectra 300 microscope with varying aberration coefficients.

Measurement-free Measurement via Deep Learning

We leverage the expressive power of state-of-the-art deep learning models to extract the abundant phase space information from electron Ronchigrams. Here we design a customized convolutional neural network (CNN) based on the VGG architecture commonly used in image recognition. With a training data set of 25,000 simulated Ronchigrams under various types of aberrations labeled by their ground-truth beam emittance, the CNN can serve as a rapid beam quality measurement tool tool of emittance without the need of extra measurement.



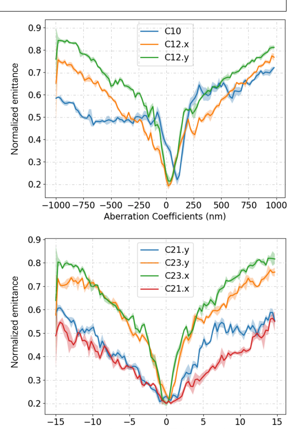


Fig. 3. Left: VGG-16 architecture to map Ronchigrams to beam emittance. Right: CNN response to change in single aberration coefficients from Nion UltraSTEM line scans (acquisition time 200 ms). Note that though defocus $C_{1,0}$ should not affect emittance, variation is present in the upper plot due to the hysteresis and coupling between lenses.

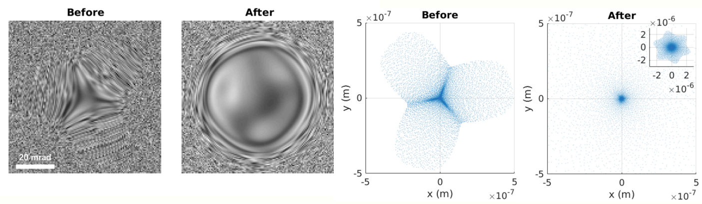


Fig. 4. Left: Ronchigrams from the beam before and after optimization.
Right: Transverse distribution of electrons at the center of the objective lens.
The comparison indicates that optimization of aberration correctors according to the minimization of beam emittance growth can effectively eliminate lower order aberrations.

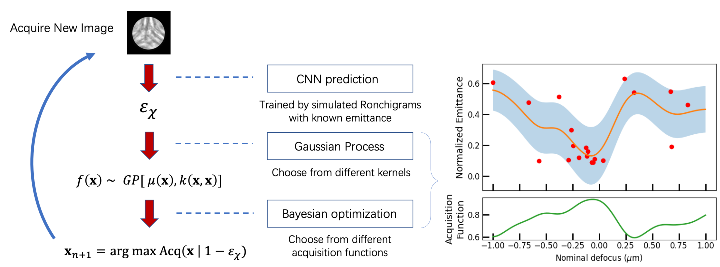


Fig. 5. Workflow of online Bayesian optimization of an electron microscope. The right illustrates the process in 1D.

Physics-informed Bayesian Optimization

The mapping from Ronchigrams to emittance enabled by deep learning gives the black-box function to optimize for aberration correction. Here we adopt Bayesian optimization, a sample efficient and gradient free method, for faster convergence and uncertainty quantification, where a gaussian processes (GP) model is applied as the surrogate model. The modelling accuracy largely relies on the kernel, which captures the correlations between different observations and among different dimensions. While most generic kernels are isotropic (e.g., RBF, Matern), we can further implement a deep kernel, built upon a simple deep neural network, to improve the optimization performance by effectively learning the correlations between input dimensions.

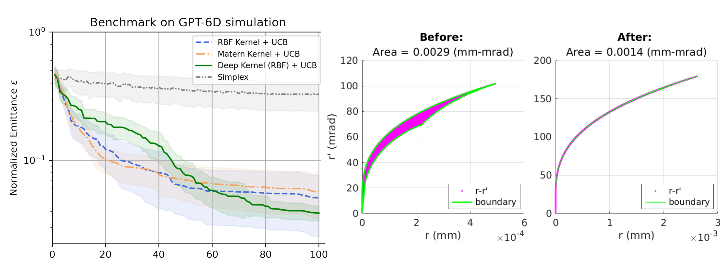


Fig. 6. Left: Benchmark results of deep kernel Bayesian optimization v.s. generic Bayesian optimization v.s. Simplex method on the GPT-6D simulation.
Right: Phase space distribution of the electrons at the center of the objective lens in polar coordinates. The area of the cloud represents beam emittance and is calculated from the nonconvex polygon boundary marked in green.

Acknowledgement