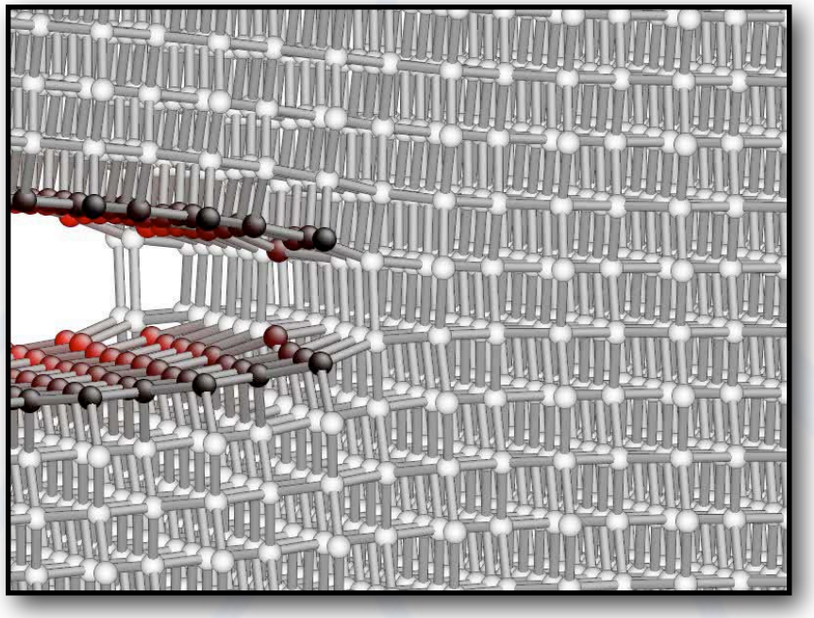
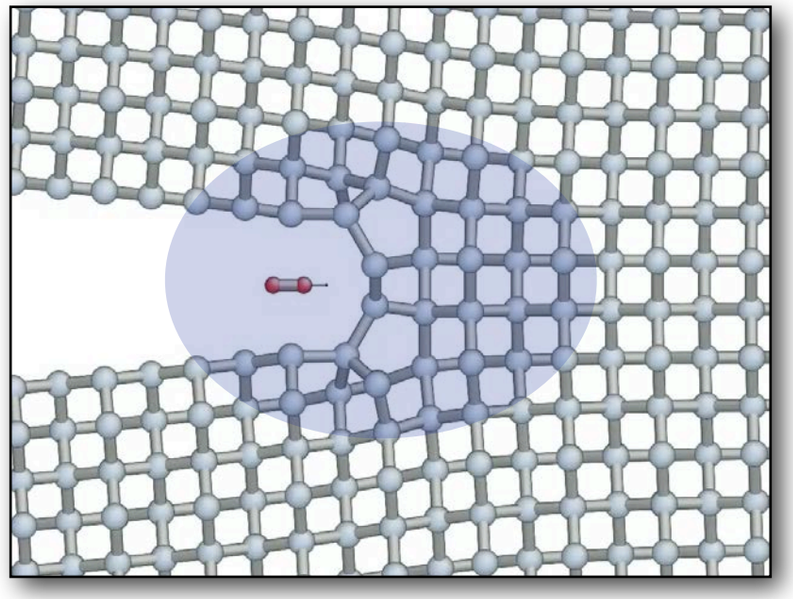
Complex chemistry & realistic systems…

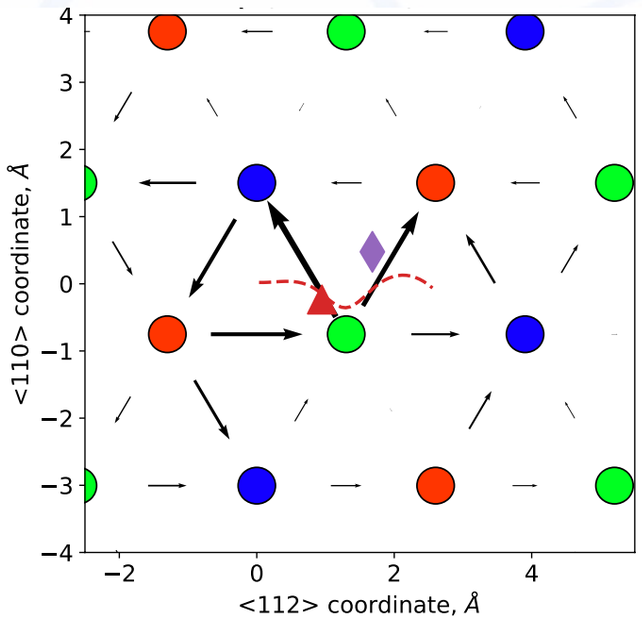
… require large systems, long timescales and quantification of uncertainty



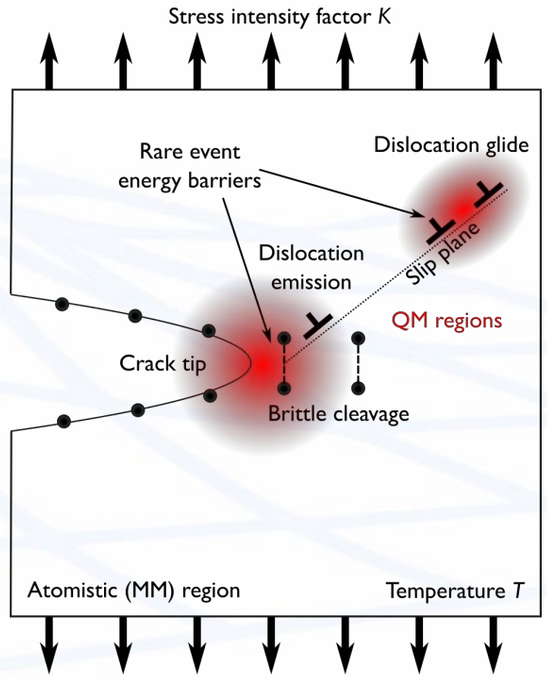
Three Dimensional Systems



Stress Corrosion Cracking



Dislocation/impurity interaction



Cleavage vs. Emission

General-purpose machine learning potential for silicon

Gaussian Approximation Potential (GAP) framework – data-driven Gaussian Process model, trained from DFT data via SOAP representation of atomic environments
(~2.5k configs, ~170k atomic enviroments, sparsified to 9k)

$$E = \sum_{i<j} V^{(2)}(r_{ij}) + \sum_i \sum_s^M \alpha_s K(\mathcal{R}_i, \mathcal{R}_s),$$

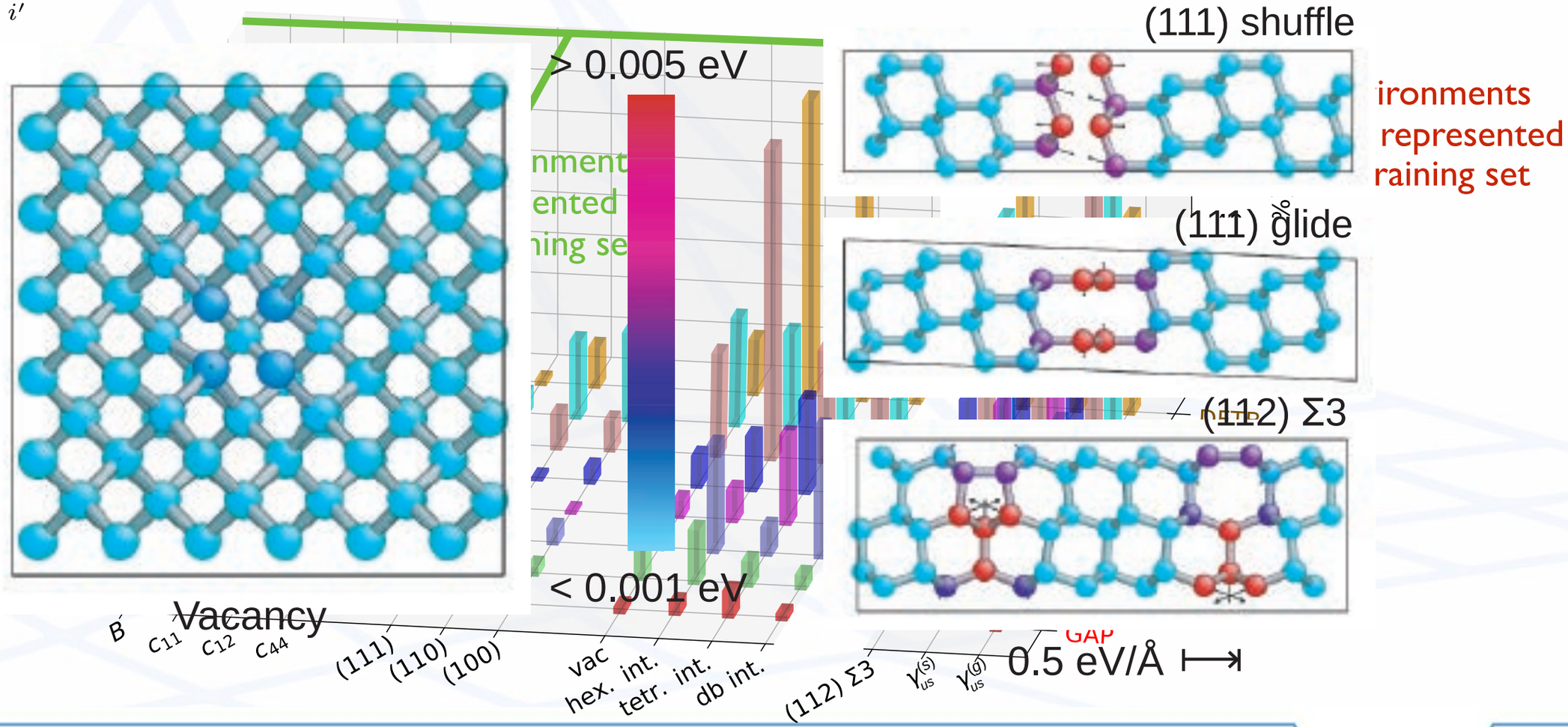
$$\tilde{K}(\mathcal{R}_i, \mathcal{R}_j) = \int_{\hat{R}\in SO_3} d\hat{R} \left| \int d\mathbf{r} \rho_i(\mathbf{r}) \rho_j(\hat{R}\mathbf{r}) \right|^2$$

$$\rho_i(\mathbf{r}) = \sum_{i'} f_{\text{cut}}(r_{ii'}) e^{-(\mathbf{r}-\mathbf{r}_{ii'})/2\sigma^2_{\text{atom}}}$$

$$\boldsymbol{\alpha} = \left[\mathbf{K}_{MM} + \mathbf{K}_{MN}\mathbf{L}\boldsymbol{\Lambda}^{-1}\mathbf{L}^{\mathrm{T}}\mathbf{K}_{NM}\right]^{-1}\mathbf{K}_{MN}\mathbf{L}\boldsymbol{\Lambda}^{-1}\boldsymbol{y},$$

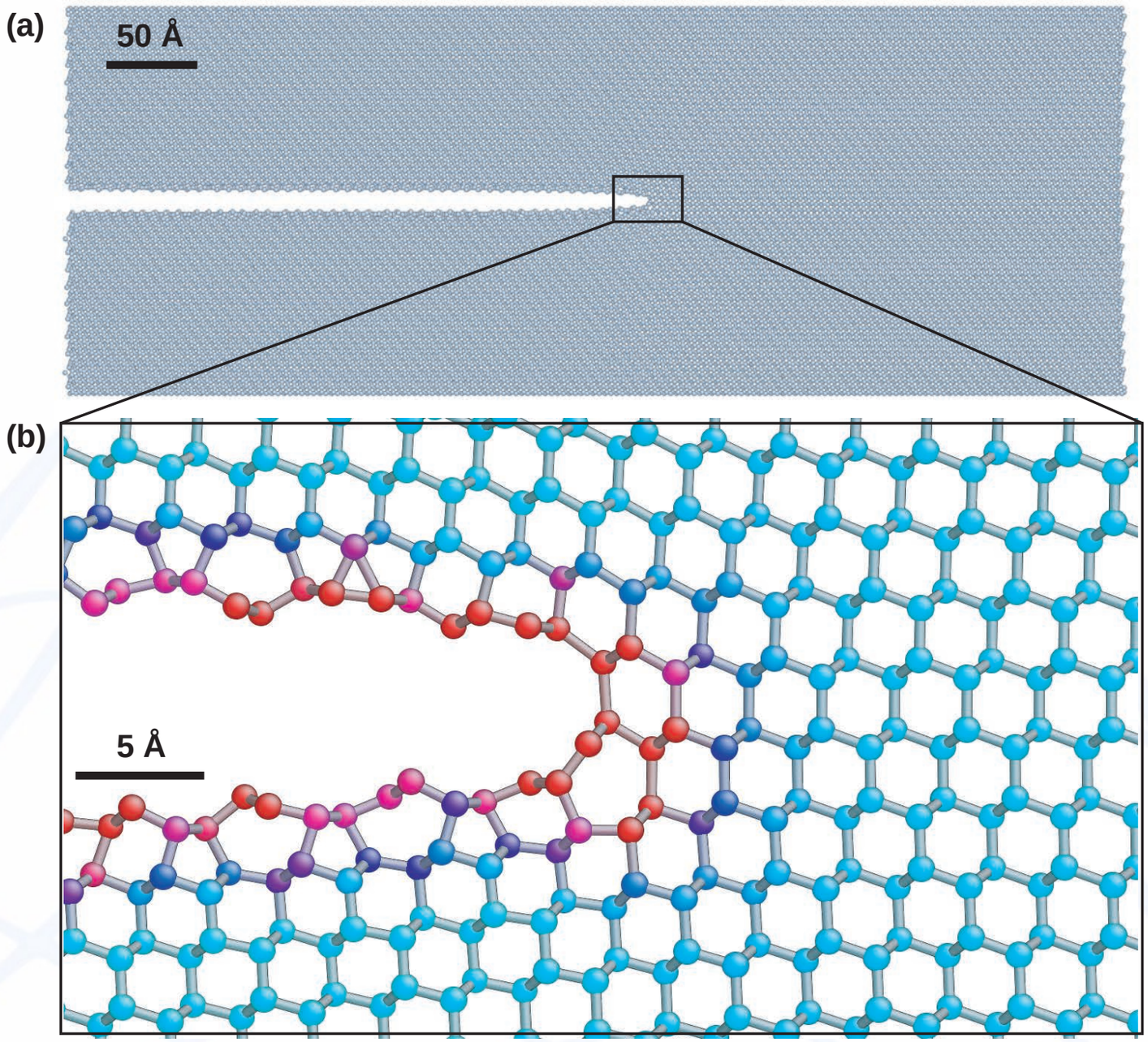
Posterior variance (over an effective ensemble of potentials trained on same data) is analytic:

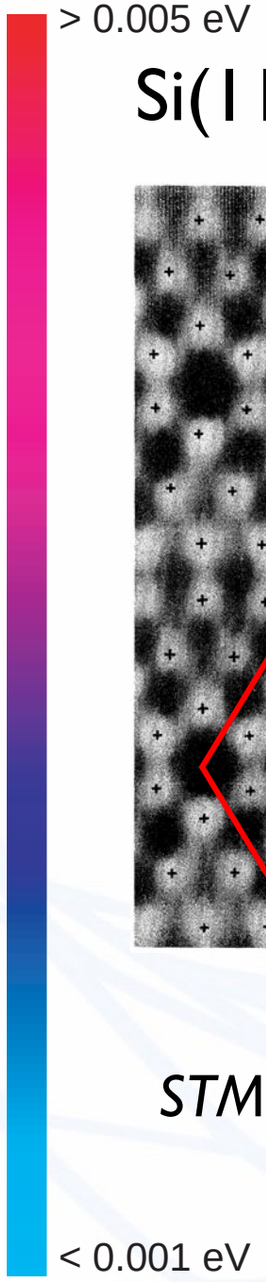
$$K(\mathcal{R}_i, \mathcal{R}_i) - \mathbf{k}^T(\mathbf{K}_{MM} + \sigma_e \mathbf{I})^{-1}\mathbf{k}$$



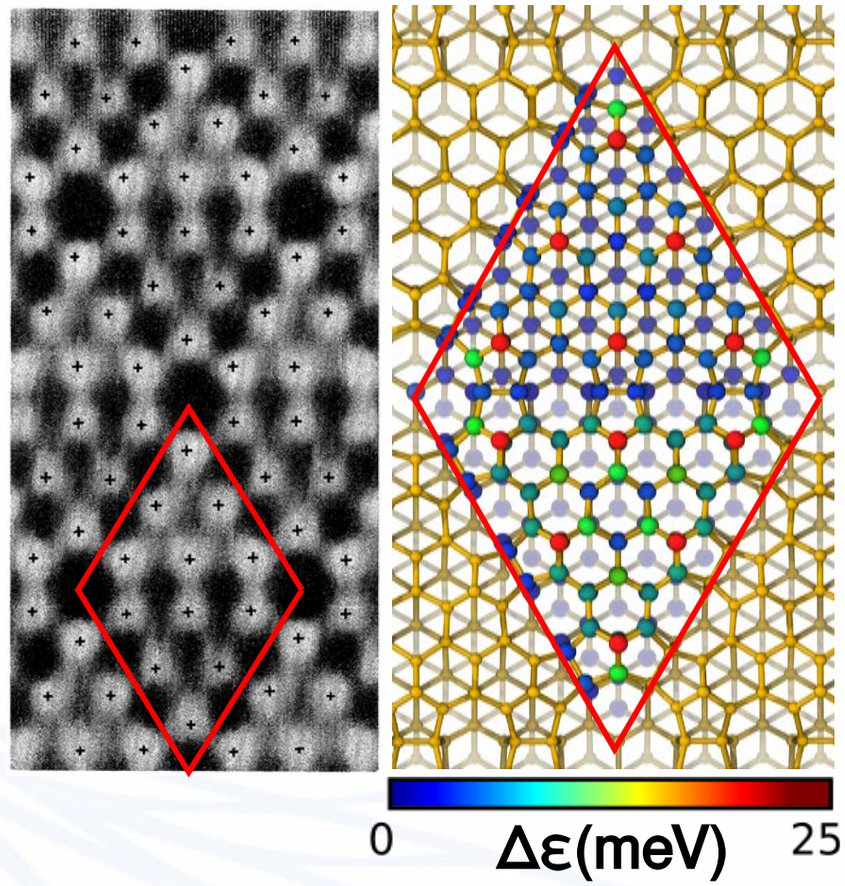
A. P. Bartok et al. *Science Advances* **3**, e1701816 (2017)
A. P. Bartok, JRK, N. Bernstein and G. Csanyi, PRX **8**, 041048 (2018)

Uncertainty Quantification for the Silicon GAP model





Si(111) 7x7 reconstruction



STM image

GAP trained without adatoms, error per atom

Gaussian Approximation Potential (GAP) for silicon – data-driven model, with per-atom predicted errors from variance of posterior probability distribution

A. P. Bartok et al. *Science Advances* **3**, e1701816 (2017)
A. P. Bartok, JRK, N. Bernstein and G. Csanyi, PRX **8**, 041048 (2018)