Bayesian inference and deterministic anisotropy for molecular geometry retrieval in gas-phase diffraction experiments

ULiTiMA: 3/13/2023

Kareem Hegazy, Varun Makhija, Phil Bucksbaum, Jeff Corbett, James Cryan, Nick Hartmann, Markus Ilchen, Keith Jobe, Renkai Li, Igor Makasyuk, Xiaozhe Shen, Xijie Wang, Stephen Weathersby, Jie Yang, **Ryan Coffee**

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Presentation Objective and Contents

SLAC

We introduce a mathematically rigorous data driven approach to retrieve molecular frame geometries while reducing simulation requirements.

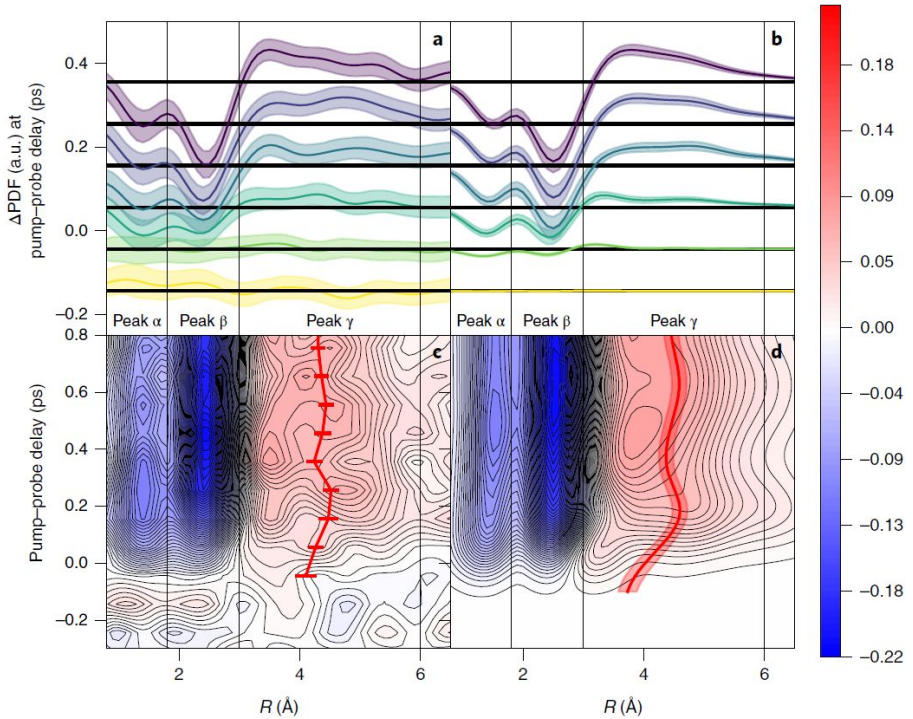
1. Current approaches and the curse of dimensionality
2. Anisotropy reveals the molecular frame (MF)
3. Bayesian Inference
   a. Overcoming the curse of dimensionality
   b. Metropolis Hastings Algorithm
   c. Results
4. Future Application to Excited State Geometric Dynamics
5. Summary

Current Molecular Geometry Retrieval Methods
Traditional and New Methods

SLAC

Traditional Method

- Pair distribution function (PDF)
- Compare PDF, or diffraction, with theory to interpret the molecular geometry transience

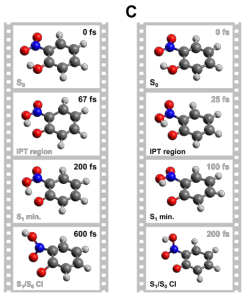


Wolf, T J A et al. "The photochemical ring-opening of 1,3-cyclohexadiene imaged by ultrafast electron diffraction." *Nature chemistry* vol. 11,6 (2019): 504-509. doi:10.1038/s41557-019-0252-7

Data Focused/Driven Method

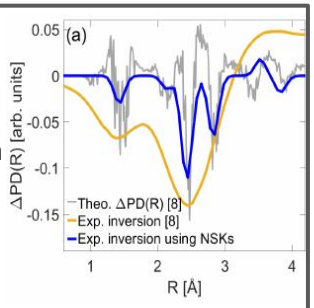
- Use ML to optimize primary features
- Employ kernel transformations from diffraction to PDF
- Use many simulated geometries to statistically improve precision

Nune et al. retrieved the primary molecular geometry changes with a genetic algorithm. Each atom was allowed to move within intuitive constraints.



Nunes, J P F et al. "Imaging proton transfer dynamics in o-nitrophenol by Ultrafast Electron Diffraction" Submitted to Journal

Natan employs a kernel to transform diffraction in PDFs and accounting for detector artifacts



Natan, Adi, "Real-Space Inversion and Super-Resolution of Ultrafast Scattering using Natural Scattering Kernels" *Arxiv*, 2021, doi: 10.48550/ARXIV.2107.05576

Stankus et al. compared excited state measurements to many simulations to form a distribution of likely atomic distances and angles.

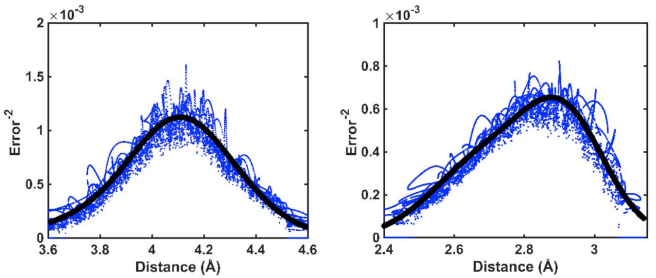


Fig. 4: The inverse of the least-squares fitting error as a function of the given structural parameter and the fits as described in the text. Left: O-C5 distance at 0.05 ps delay time. Right: O-N distance at 2.65 ps delay time.

Stankus, B., Yong, H., Zotev, N. et al. Ultrafast X-ray scattering reveals vibrational coherence following Rydberg excitation. Nat. Chem. 11, 716–721 (2019). https://doi.org/10.1038/s41557-019-0291-0

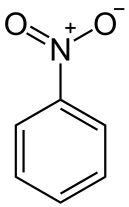
Current Molecular Geometry Retrieval Methods
Curse of Dimensionality

The number of geometries to sample (S) grows exponentially, for a grid search

$$S = R^{3N-6}$$

R: Number of sample points per dimension
N: Number of atoms
V: Percent of volume

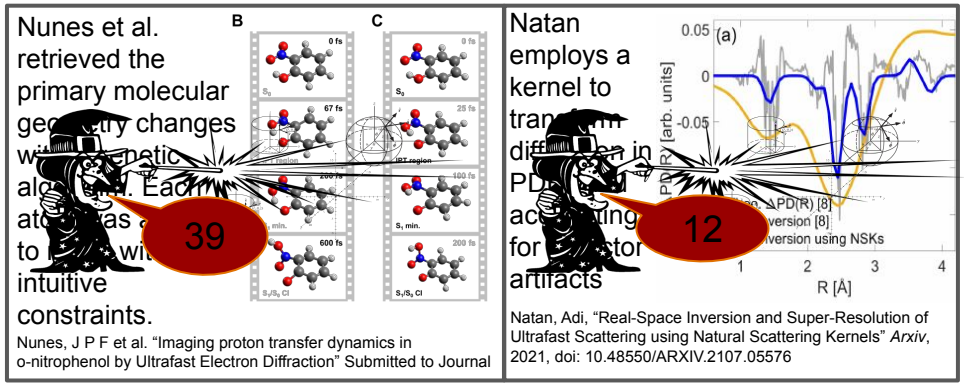
Nitrobenzene: 21 dimensions (ignoring H)

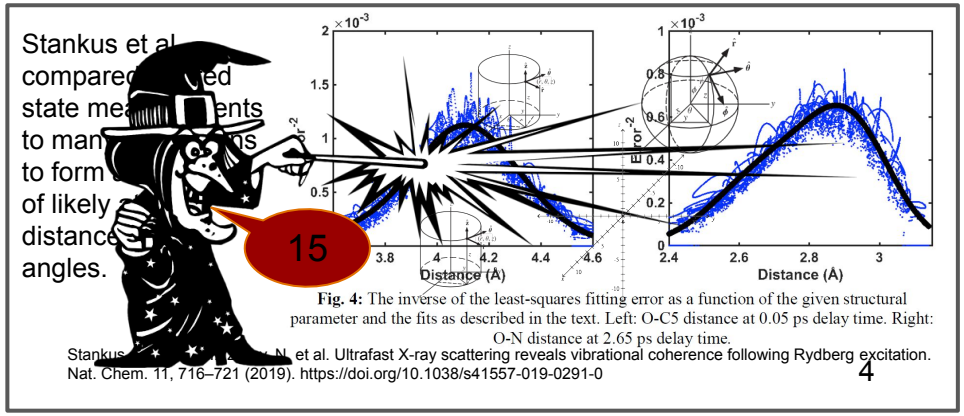


S	R
10^6	1.93
10^9	2.68
10^{21}	10

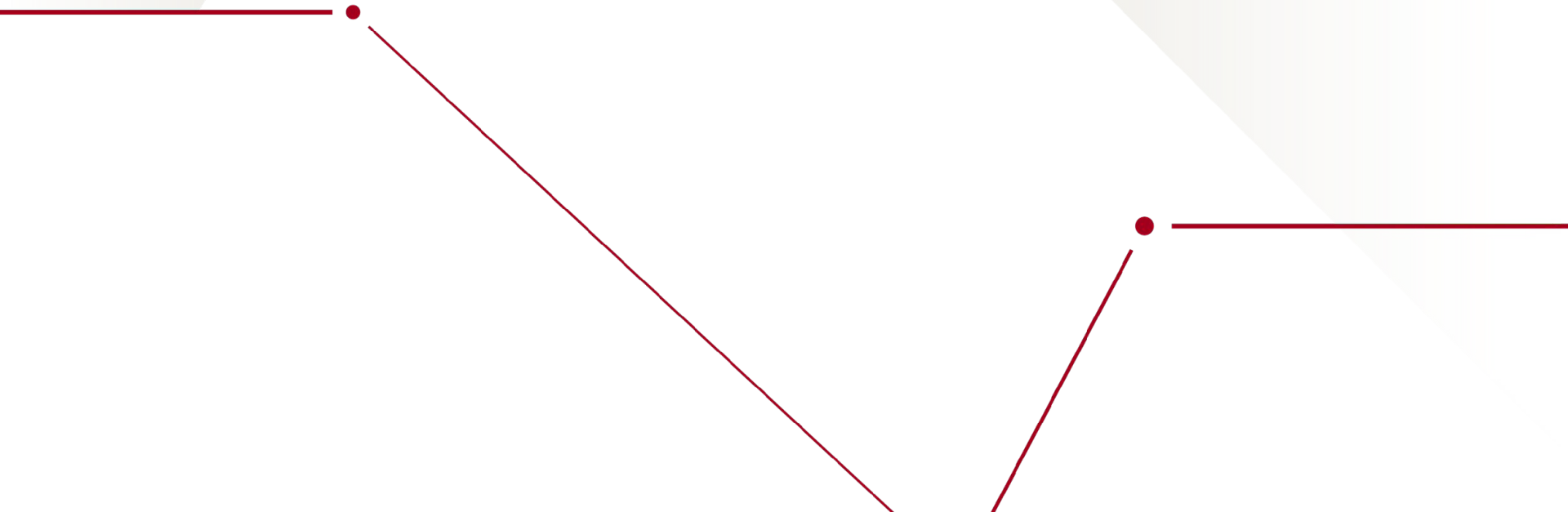
Data Focused/Driven Method

- Use ML to optimize primary features
- Employ kernel transformations from diffraction to PDF
- Use many simulated geometries to statistically improve precision





Anisotropy Reveals the Molecular Frame (MF)



Anisotropy Reveals the Molecular Frame (MF)

- Anisotropy provides constraints on molecular frame (MF) degrees of freedom.
- Combining many measurements (constraints) allows one to retrieve the MF.
- Application in photo-electron spectroscopy
  - V. Makhija, et. al., (2016), arXiv:1611.06476 [physics.atom-ph]
  - C. Marceau, et. al., Phys. Rev. Lett. 119, 083401 (2017)
  - M. Gregory, et. al., (2020), arXiv:2012.04561 [physics.chem-ph]
- Use a stretched NO_2: an asymmetric top

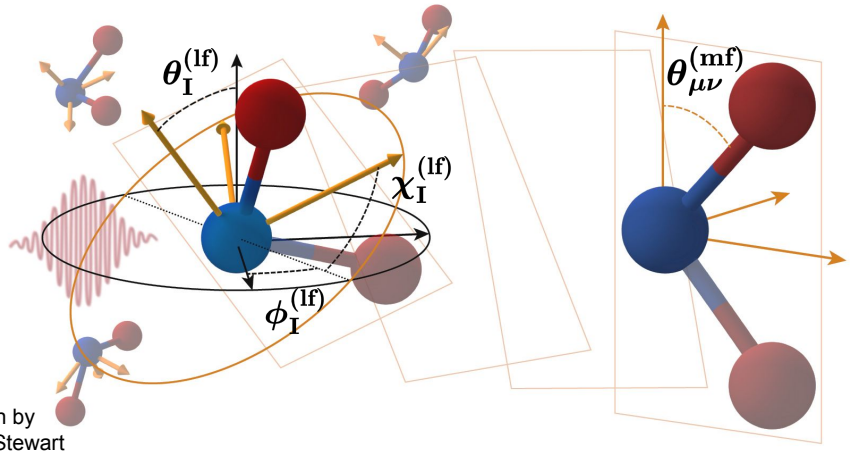
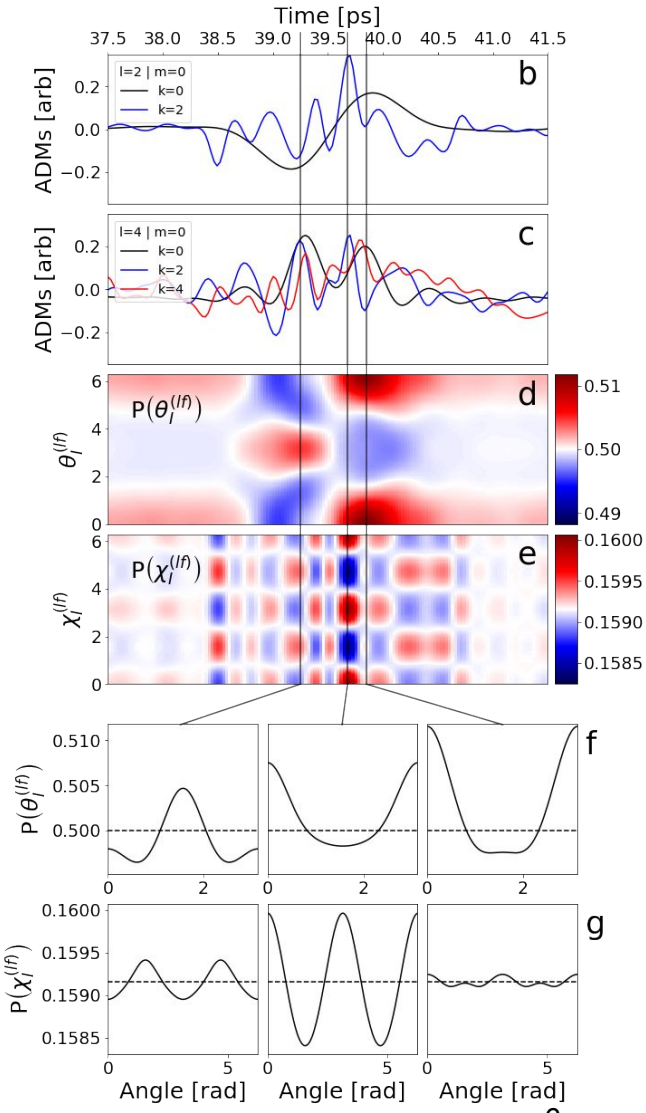


Illustration by Gregory Stewart



NO_2 rotational wavepacket anisotropy signature

Accessing the MF via Deterministic Anisotropy

SLAC

$$\langle I(\mathbf{q})\rangle(t) = \mathcal{I}\Bigg(\sum_{\mu} |f_{\mu}^{*}(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Bigg\{ f_{\mu}(q) f_{\nu}^{*}(q) \sum_{l} i^{l} 8\pi^2 \sqrt{4\pi\,(2l+1)} \times \sum_{m,k} (-1)^{k-m} \underbrace{Y_{l}^{-m}\left(\theta_{q}^{(\mathrm{lf})}, \phi_{q}^{(\mathrm{lf})}\right)}_{\text{Lab Frame}} \langle \Psi(t) | \underbrace{D_{mk}^{l}\left(\phi_{\mathrm{I}}^{(\mathrm{lf})}, \theta_{\mathrm{I}}^{(\mathrm{lf})}, \chi_{\mathrm{I}}^{(\mathrm{lf})}\right)}_{\text{Ensemble Anisotropy}} \underbrace{j_{l}(q\Delta r_{\mu\nu}) Y_{l}^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)}_{\text{Molecular Frame Geometry}} | \Psi(t) \rangle \Bigg\} \Bigg)$$

Independent atom approximation

- **Measurement:** Lab frame anisotropy
- **Simulation:** Ensemble anisotropy
- **Result:** Molecular frame geometry
  - Explicit dependence on MF geometric angles
  - Each pair-wise contribution is labelled

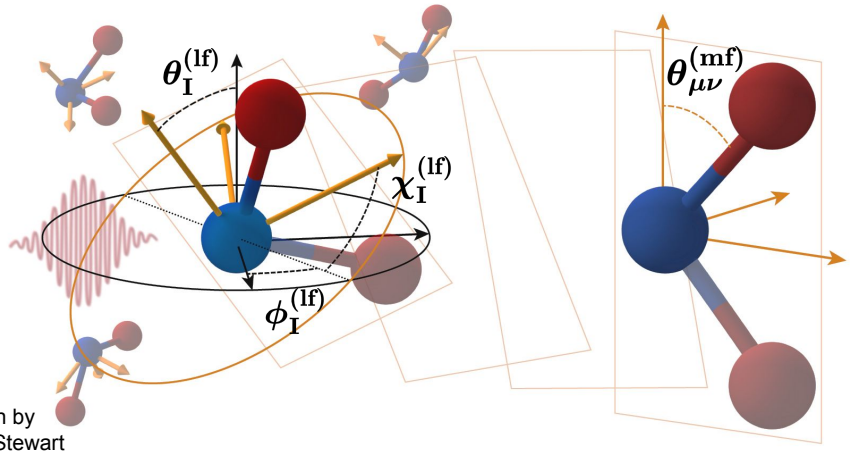


Illustration by Gregory Stewart

Accessing the MF via Deterministic Anisotropy

SLAC

Independent atom approximation

$$\langle I(\mathbf{q})\rangle(t) = \mathcal{I}\Bigg(\sum_{\mu} |f_{\mu}^{*}(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Bigg\{ f_{\mu}(q) f_{\nu}^{*}(q) \sum_{l} i^{l} 8\pi^2 \sqrt{4\pi\,(2l+1)} \times \sum_{m,k} (-1)^{k-m} \underbrace{Y_{l}^{-m}\left(\theta_{q}^{(\mathrm{lf})}, \phi_{q}^{(\mathrm{lf})}\right)}_{\text{Lab Frame}} \langle \Psi(t) | \underbrace{D_{mk}^{l}\left(\phi_{\mathrm{I}}^{(\mathrm{lf})}, \theta_{\mathrm{I}}^{(\mathrm{lf})}, \chi_{\mathrm{I}}^{(\mathrm{lf})}\right)}_{\text{Ensemble Anisotropy}} \underbrace{j_{l}(q\Delta r_{\mu\nu}) Y_{l}^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)}_{\text{Molecular Frame Geometry}} | \Psi(t) \rangle \Bigg\} \Bigg)$$

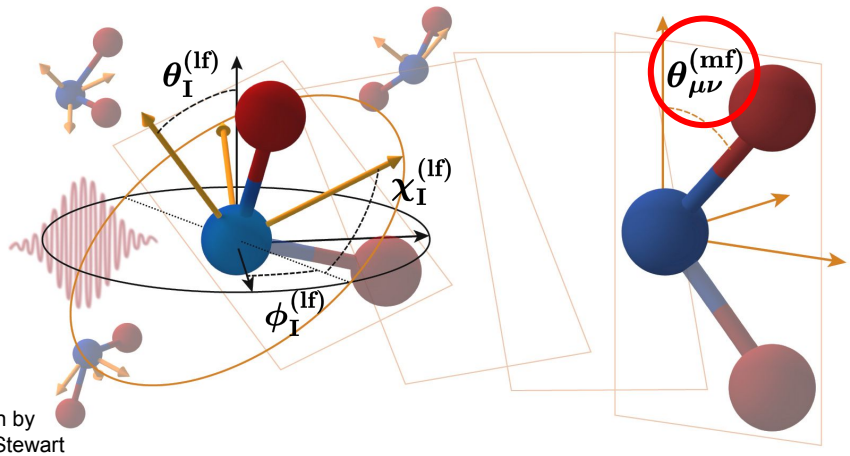


Illustration by Gregory Stewart

Accessing the MF via Deterministic Anisotropy

SLAC

$$\langle I(\mathbf{q})\rangle(t) = \mathcal{I}\Bigg(\sum_{\mu} |f_{\mu}^{*}(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Bigg\{ f_{\mu}(q) f_{\nu}^{*}(q) \sum_{l} i^{l} 8\pi^2 \sqrt{4\pi\,(2l+1)} \times \sum_{m,k} (-1)^{k-m} \underbrace{Y_{l}^{-m}\left(\theta_{q}^{(\mathrm{lf})}, \phi_{q}^{(\mathrm{lf})}\right)}_{\text{Lab Frame}} \langle \Psi(t)| \underbrace{D_{mk}^{l}\left(\phi_{\mathrm{I}}^{(\mathrm{lf})}, \theta_{\mathrm{I}}^{(\mathrm{lf})}, \chi_{\mathrm{I}}^{(\mathrm{lf})}\right)}_{\text{Ensemble Anisotropy}} \underbrace{j_{l}(q\Delta r_{\mu\nu}) Y_{l}^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)}_{\text{Molecular Frame Geometry}} |\Psi(t)\rangle \Bigg\} \Bigg)$$

Independent atom approximation

Explicit dependence on molecular frame angles provides an opportunity to develop methods that extract these angles, and hence a unique geometry, from data alone.

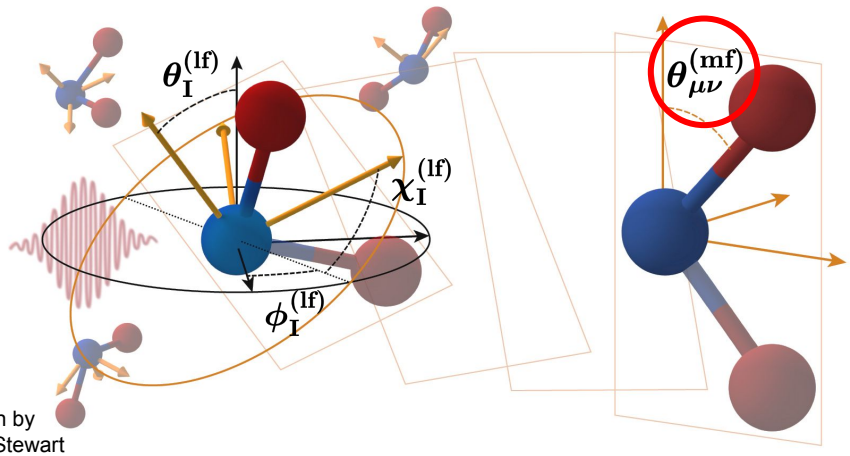
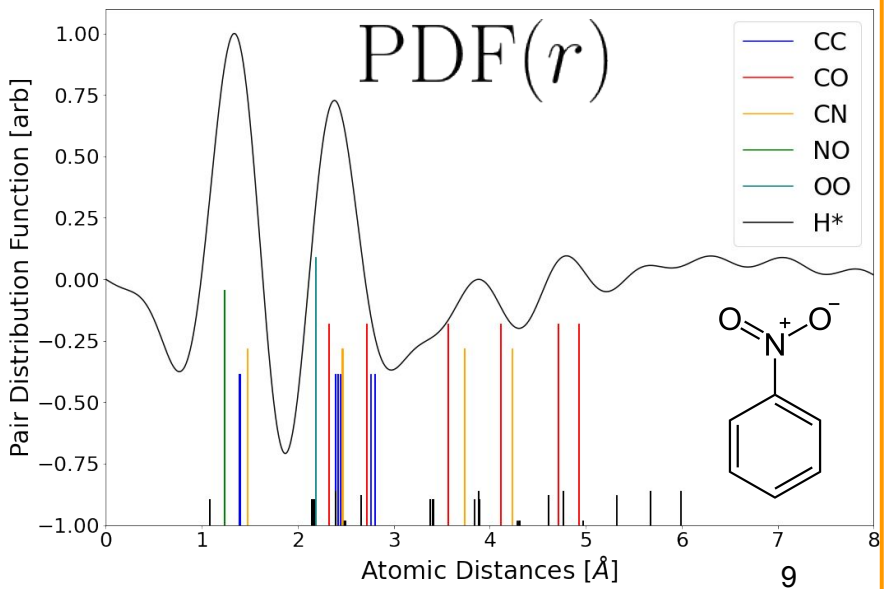


Illustration by Gregory Stewart

PDF access distances only



Accessing the MF via Deterministic Anisotropy

SLAC

Independent atom approximation

$$\langle I(\mathbf{q})\rangle(t) = \mathcal{I}\Bigg(\sum_{\mu} |f_{\mu}^{*}(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Bigg\{ f_{\mu}(q) f_{\nu}^{*}(q) \sum_{l} i^{l} 8\pi^2 \sqrt{4\pi\,(2l+1)} \times \sum_{m,k} (-1)^{k-m} \underbrace{Y_{l}^{-m}\left(\theta_{q}^{(\mathrm{lf})}, \phi_{q}^{(\mathrm{lf})}\right)}_{\text{Lab Frame}} \langle \Psi(t)| \underbrace{D_{mk}^{l}\left(\phi_{\mathrm{I}}^{(\mathrm{lf})}, \theta_{\mathrm{I}}^{(\mathrm{lf})}, \chi_{\mathrm{I}}^{(\mathrm{lf})}\right)}_{\text{Ensemble Anisotropy}} \underbrace{j_{l}(q\Delta r_{\mu\nu}) Y_{l}^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)}_{\text{Molecular Frame Geometry}} |\Psi(t)\rangle \Bigg\} \Bigg)$$

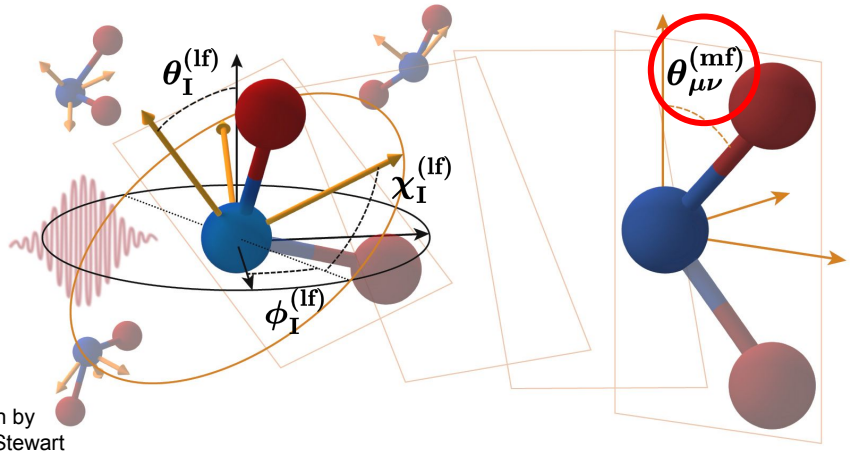


Illustration by Gregory Stewart

Must know the ground state rotational constants and static polarizability to simulate the ensemble anisotropy

Accessing the MF via Deterministic Anisotropy
Rigid Rotor

SLAC

Rigid Rotor Approximation

$$\langle I(\mathbf{q})\rangle_{\text{rigid}}(t) = \mathcal{I}\Bigg(\sum_{\mu}|f_\mu(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu}\text{Re}\Bigg\{f_\mu(q)f_\nu^*(q)\sum_l i^l 8\pi^2\sqrt{\frac{4\pi}{(2l+1)}} \times \sum_{m,k}(-1)^{k-m}\underbrace{Y_l^{-m}\left(\theta_q^{(\text{lf})},\phi_q^{(\text{lf})}\right)}_{\text{Lab Frame}}\underbrace{\langle\Psi(0)|j_l(q\Delta r_{\mu\nu})Y_l^{-k}\left(\theta_{\mu\nu}^{(\text{mf})},\phi_{\mu\nu}^{(\text{mf})}\right)|\Psi(0)\rangle}_{\text{Molecular Frame Geometry}}\underbrace{\mathcal{A}_{mk}^l(t)\big|_{\text{rigid}}}_{\text{Anisotropy}}\Bigg\}\Bigg)$$

Remove LF Y_l^m by fitting measured anisotropy

$$B_l^m(q,t) = \mathcal{I}\sum_{\mu,\nu:\mu\neq\nu}\text{Re}\Bigg\{f_\mu(q)f_\nu^*(q)i^l 8\pi^2\sqrt{\frac{4\pi}{(2l+1)}} \times \sum_k(-1)^{k-m}\underbrace{\left\langle j_l(q\Delta r_{\mu\nu})Y_l^{-k}\left(\theta_{\mu\nu}^{(\text{mf})},\phi_{\mu\nu}^{(\text{mf})}\right)\right\rangle}_{\text{Molecular Frame Geometry}}\underbrace{\mathcal{A}_{mk}^l(t)\big|_{\text{rigid}}}_{\text{Anisotropy}}\Bigg\}$$

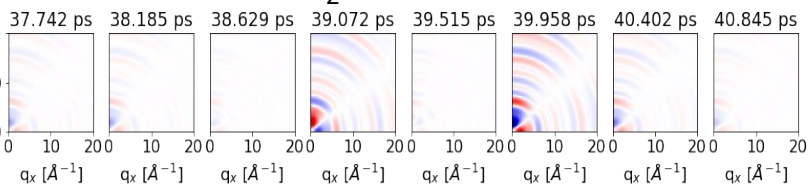
Remove time dependence by fitting ADMs

$$C_{lmk}(q) = \mathcal{I}\sum_{\mu,\nu:\mu\neq\nu}\text{Re}\Bigg\{f_\mu(q)f_\nu^*(q)(-1)^{k-m}i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}}\underbrace{\left\langle j_l(q\Delta r_{\mu\nu})Y_l^{-k}\left(\theta_{\mu\nu}^{(\text{mf})},\phi_{\mu\nu}^{(\text{mf})}\right)\right\rangle}_{\text{Molecular Frame Geometry}}\Bigg\}$$

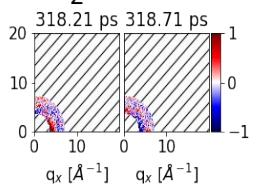
$$M_{lmk}(q) = \frac{C_{lmk}(q)}{\sum_\mu|f_\mu(q)|^2}$$

$\Delta\langle I(q)\rangle(t)$

NO_2 Simulation

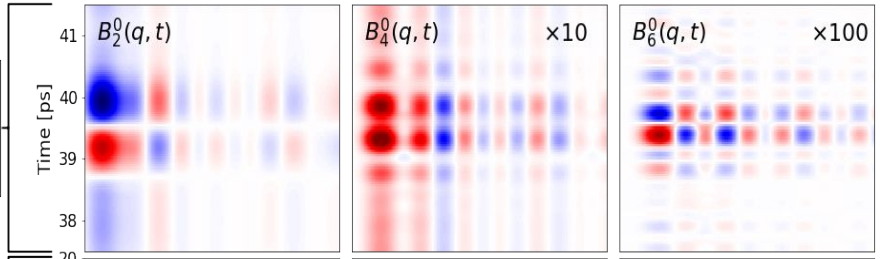


N_2O Data

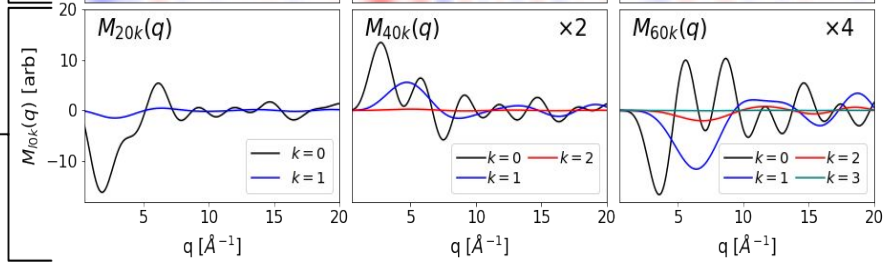


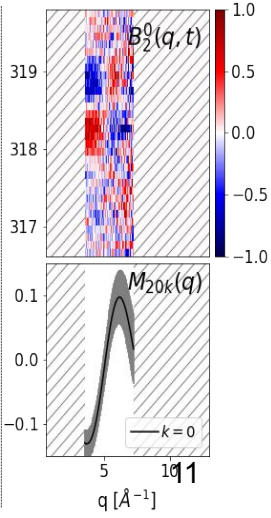
Extracting Molecular Frame Information

$B_l^0(q,t)$

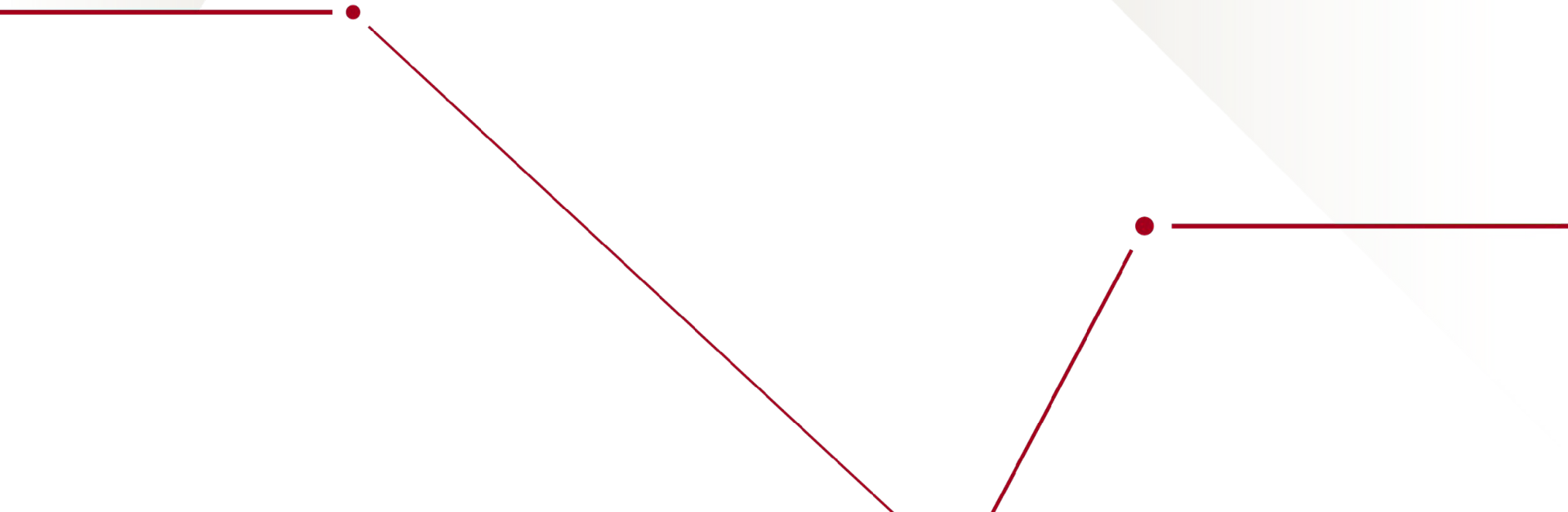


$M_{l0k}(q)$





Bayesian Inference



Modeling $|\Psi(\boldsymbol{r}^{(mf)})|^2$ and Search Parameters

SLAC

- **What we have**
  - Relation between measurement and molecular frame geometry
- **What we want**
  - Invert $C_{lmk}(q)$ for $|\Psi(\boldsymbol{r}^{(mf)})|^2$
  - ~10 eqn / ~10 terms / ~300 variables / embedded in 3N-6 dims
- **How we do it**
  - Model $|\Psi(\boldsymbol{r}^{(mf)})|^2 \approx P(\boldsymbol{r}^{(mf)}|\boldsymbol{\theta},C)$
  - Retrieve $P(\boldsymbol{\theta}|C)$

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \boxed{|\Psi(\mathbf{r})|^2} d\mathbf{r}}_{\text{Molecular Frame Geometry}} \Big\}$$

$$C_{lmk}^{(\mathrm{calc})}(q,\boldsymbol{\Theta}) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \boxed{P(\mathbf{r},\boldsymbol{\Theta}|C)} d\mathbf{r}}_{\text{Molecular Frame Geometry}} \Big\}$$

$$P(\mathbf{r},\boldsymbol{\Theta}|\,C) \approx |\Psi(\mathbf{r})|^2$$

$$P(\boldsymbol{\Theta}|C) = \int P(\mathbf{r},\boldsymbol{\Theta}|\,C)\, d\mathbf{r}$$

Delta Distribution:

$$P^{(\delta)}(\mathbf{r},\boldsymbol{\Theta}|\,C) = \delta\left(\boldsymbol{\Theta}^{(\mathrm{delta})} - \mathbf{r}\right)$$

$$\boldsymbol{\Theta}^{(\mathrm{delta})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \langle \mathrm{NO}^{(2)}\rangle, \langle \angle \mathrm{ONO}\rangle\right]$$

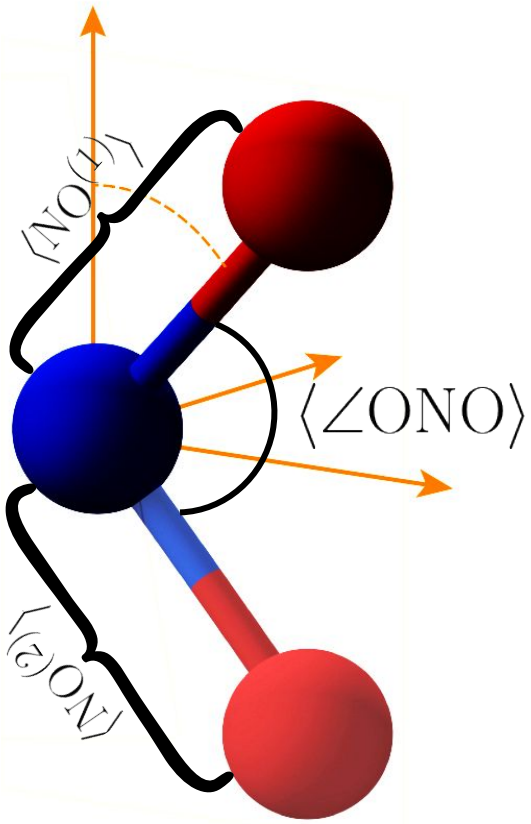
Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r},\boldsymbol{\Theta}|\,C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\mathrm{gauss})}} \exp\left\{ \frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\mathrm{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\mathrm{gauss})}} \right)^2 \right\}$$

$$\boldsymbol{\Theta}^{(\mathrm{gauss})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \sigma\left(\mathrm{NO}^{(1)}\right), \langle \mathrm{NO}^{(2)}\rangle, \sigma\left(\mathrm{NO}^{(2)}\right), \langle \angle \mathrm{ONO}\rangle, \sigma(\angle \mathrm{ONO})\right]$$

Modeling $|\Psi(r^{(mf)})|^2$ and Search Parameters

- The **θ** parameterization allows for novel measurements
  - Degrees of freedom to specify a unique geometry
  - Bond distances and angles
  - Width of the wave packet σ(⟨…⟩)



$$P(\mathbf{r}, \mathbf{\Theta}|\, C) \approx |\Psi(\mathbf{r})|^2$$

$$P(\mathbf{\Theta}|C) = \int P(\mathbf{r}, \mathbf{\Theta}|\, C)\, d\mathbf{r}$$

Delta Distribution:

$$P^{(\delta)}(\mathbf{r}, \mathbf{\Theta}|\, C) = \delta\left(\mathbf{\Theta}^{(\text{delta})} - \mathbf{r}\right)$$

$$\mathbf{\Theta}^{(\text{delta})} = \left[\langle \text{NO}^{(1)}\rangle, \langle \text{NO}^{(2)}\rangle, \langle \angle \text{ONO}\rangle\right]$$

Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r}, \mathbf{\Theta}|\, C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\text{gauss})}} \exp\left\{\frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\text{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\text{gauss})}}\right)^2\right\}$$

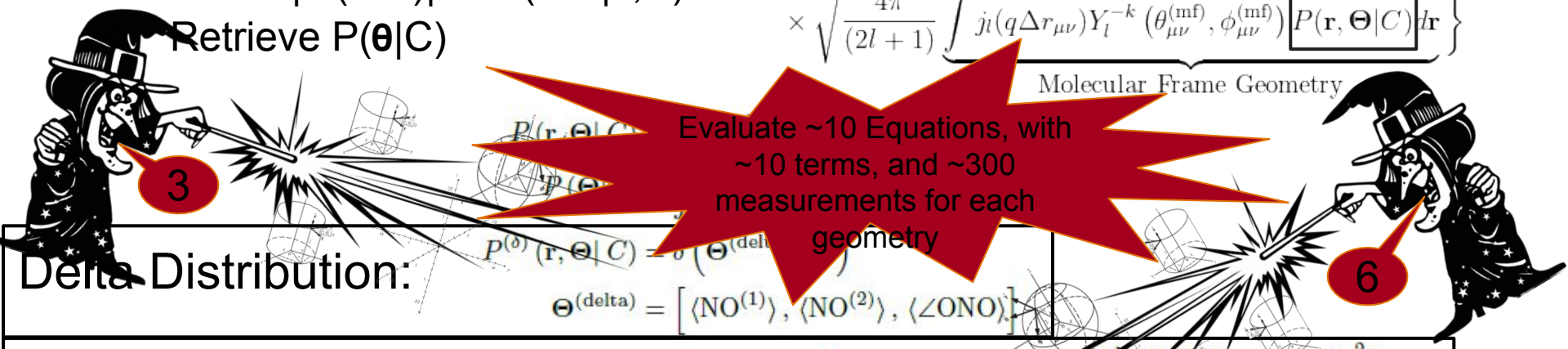
$$\mathbf{\Theta}^{(\text{gauss})} = \left[\langle \text{NO}^{(1)}\rangle, \sigma\left(\text{NO}^{(1)}\right), \langle \text{NO}^{(2)}\rangle, \sigma\left(\text{NO}^{(2)}\right), \langle \angle \text{ONO}\rangle, \sigma\left(\angle \text{ONO}\right)\right]$$

Modeling $|\Psi(r^{(mf)})|^2$ and Search Parameters

- **What we have**
  - Relation between measurement and molecular frame geometry
- **What we want**
  - Invert $C_{lmk}(q)$ for $|\Psi(r^{(mf)})|^2$
  - ~10 eqn / ~10 terms / ~300 variables / embedded in 3N-6 dims
- **How we do it**
  - Model $|\Psi(r^{(mf)})|^2 \approx P(r^{(mf)}|\theta,C)$
  - Retrieve $P(\theta|C)$

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \boxed{|\Psi(\mathbf{r})|^2} d\mathbf{r}}_{\text{Molecular Frame Geometry}} \Big\}$$

$$C_{lmk}^{(\mathrm{calc})}(q,\Theta) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \boxed{P(\mathbf{r},\Theta|C)} d\mathbf{r}}_{\text{Molecular Frame Geometry}} \Big\}$$



Delta Distribution:

$$P^{(\delta)}(\mathbf{r},\Theta|C) = \delta\left(\Theta^{(\mathrm{delta})}\right)$$

$$\Theta^{(\mathrm{delta})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \langle \mathrm{NO}^{(2)}\rangle, \langle \angle \mathrm{ONO}\rangle\right]$$

Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r},\Theta|C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\mathrm{gauss})}} \exp\left\{ \frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\mathrm{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\mathrm{gauss})}} \right)^2 \right\}$$

$$\Theta^{(\mathrm{gauss})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \sigma\left(\mathrm{NO}^{(1)}\right), \langle \mathrm{NO}^{(2)}\rangle, \sigma\left(\mathrm{NO}^{(2)}\right), \langle \angle \mathrm{ONO}\rangle, \sigma\left(\angle \mathrm{ONO}\right)\right]$$

Metropolis-Hastings Algorithm

SLAC

Efficient, unbiased, data-driven sampling of **Θ**

$$P(\boldsymbol{\Theta}|C) = [\boldsymbol{\Theta}_0, \cdots, \boldsymbol{\Theta}_N]$$

$$P(\boldsymbol{\Theta}'|\boldsymbol{\Theta}) = \frac{P(\boldsymbol{\Theta}'|C)}{P(\boldsymbol{\Theta}|C)} = \min\left[\frac{P(C|\boldsymbol{\Theta}')P(\boldsymbol{\Theta}')}{P(C|\boldsymbol{\Theta})P(\boldsymbol{\Theta})}, 1\right]$$

Select **Θ'** nearby **Θ** and calculate expected $C_{lmk}(q)$

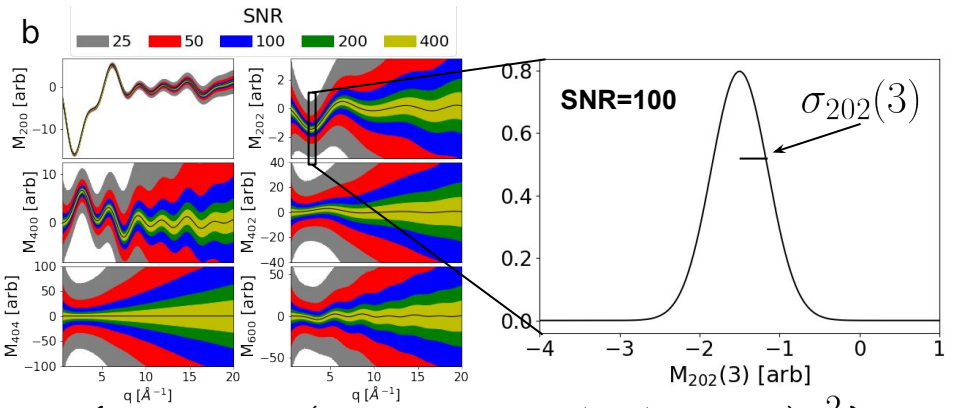
$$P^{(\mathcal{N})}(\mathbf{r}, \boldsymbol{\Theta}|\, C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\text{gauss})}} \exp\left\{\frac{-1}{2}\sum_{i=0}^{i<N_{dof}}\left(\frac{\Theta_{2i}^{(\text{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\text{gauss})}}\right)^2\right\}$$

$$\boldsymbol{\Theta}^{(\text{gauss})} = \left[\langle \text{NO}^{(1)}\rangle, \sigma\left(\text{NO}^{(1)}\right), \langle \text{NO}^{(2)}\rangle, \sigma\left(\text{NO}^{(2)}\right), \langle \angle\text{ONO}\rangle, \sigma\left(\angle\text{ONO}\right)\right]$$

$$C_{lmk}^{(\text{calc})}(q, \boldsymbol{\Theta}) = \mathcal{I}\sum_{\mu,\nu,\mu\neq\nu} \text{Re}\left\{ f_\mu(q) f_\nu^*(q)(-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\text{mf})}, \phi_{\mu\nu}^{(\text{mf})}\right) P(\mathbf{r}, \boldsymbol{\Theta}|C) d\mathbf{r}}_{\text{Molecular Frame Geometry}} \right\}$$

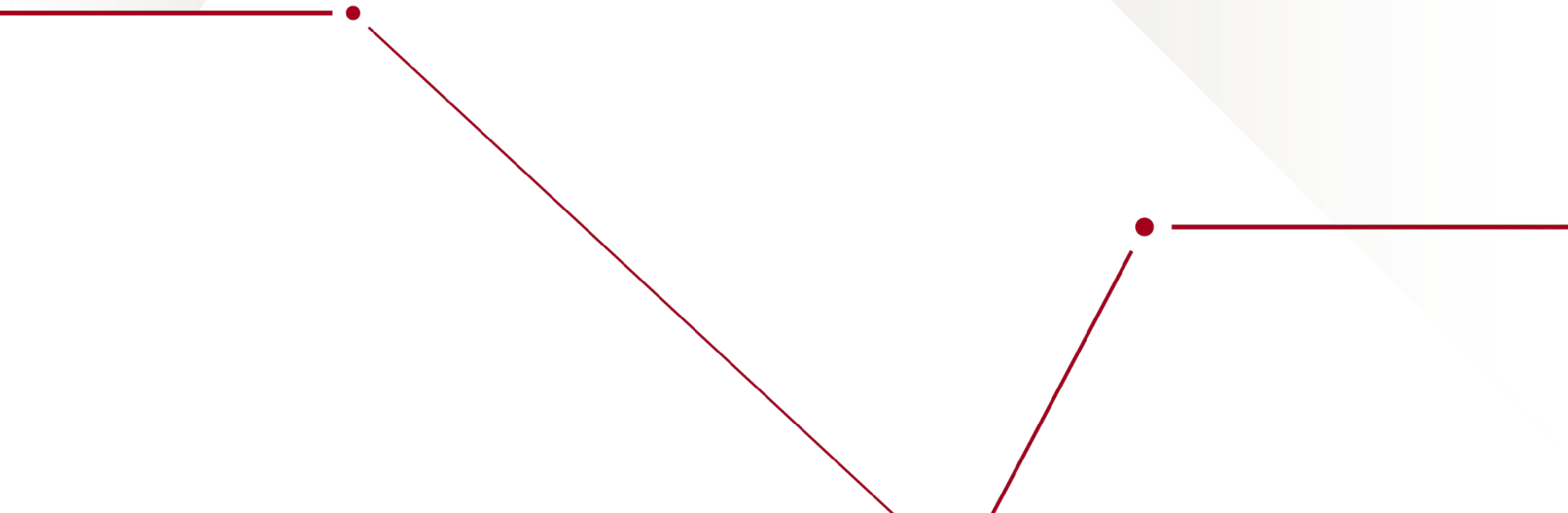
Compute probabilistic formulation of data and compare with probabilistic measurement.

- Exploits mean measurement
- Exploits experimental variation



$$P(C|\boldsymbol{\Theta}) = \left[\prod_{lmk,q} \frac{1}{\sigma_{lmk}(q)\sqrt{2\pi}}\right] \exp\left\{\frac{-1}{2}\sum_{lmk,q}\left(\frac{C_{lmk}(q) - C_{lmk}^{(\text{calc})}(q, \boldsymbol{\Theta})}{\sigma_{lmk}(q)}\right)^2\right\}$$

Application to simulated NO_2 and measured N_2O



Retrieving the MF Geometry Probability Distribution
Algorithm Input

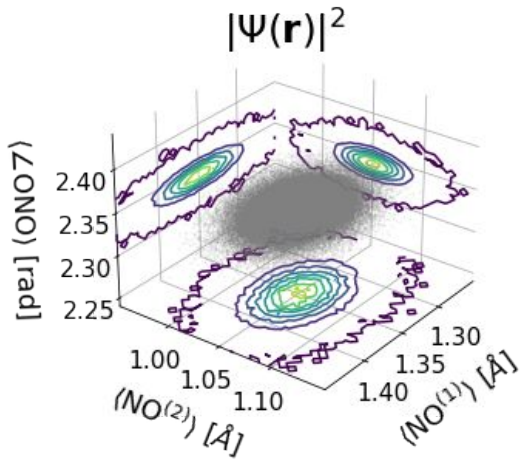
SLAC

Simulated NO_2

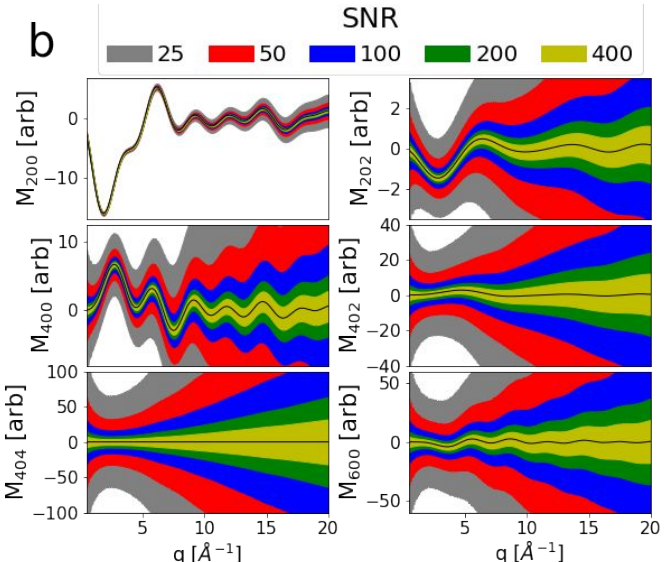
Measured N_2O

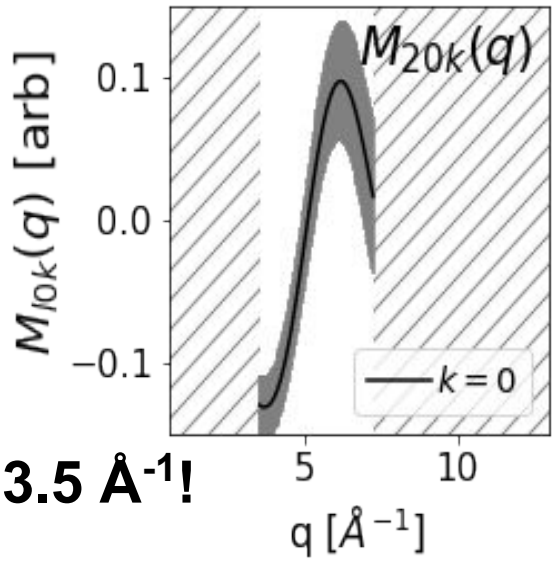
$|\Psi(\boldsymbol{r}^{(mf)})|^2$

Data



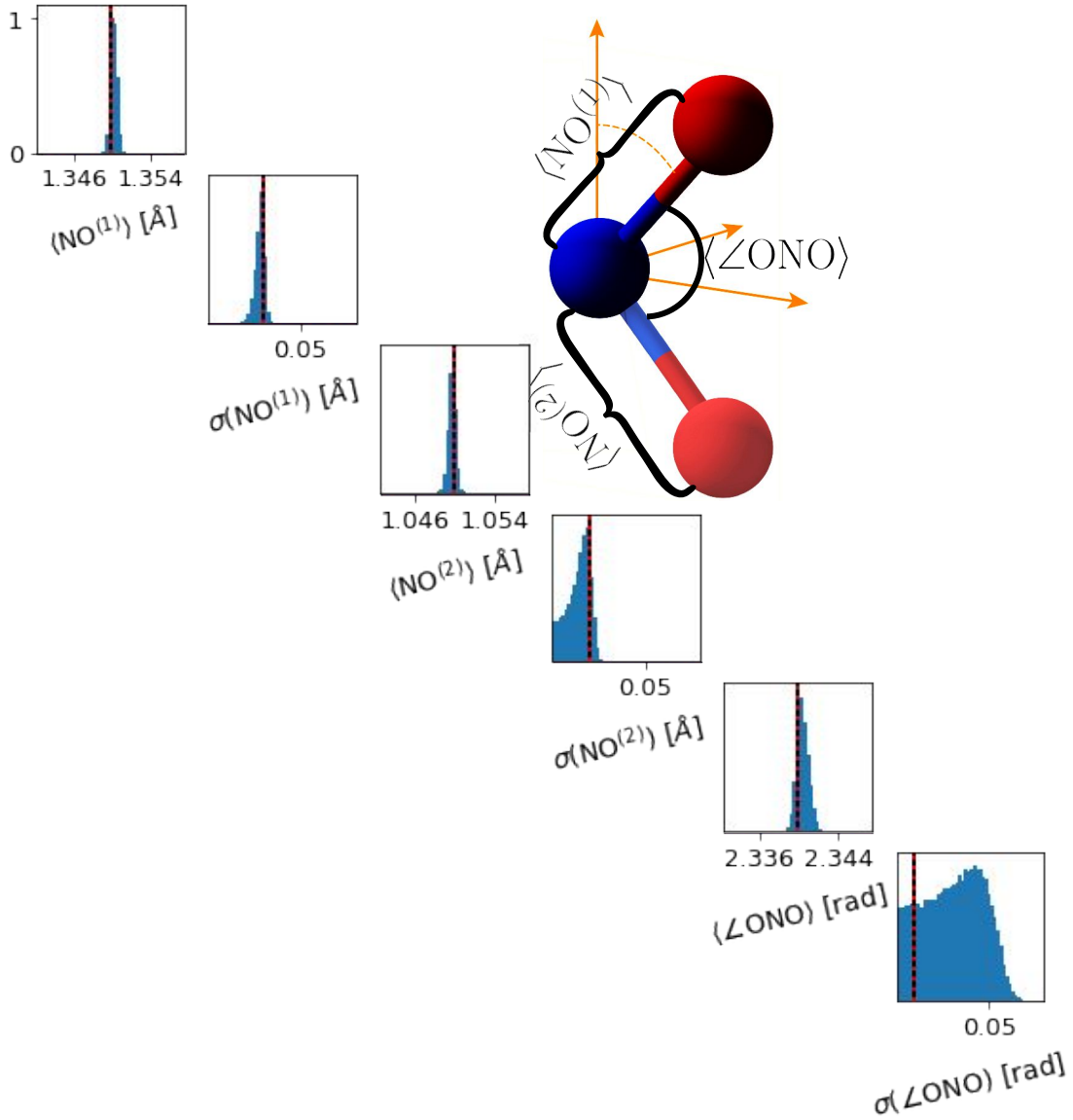
?





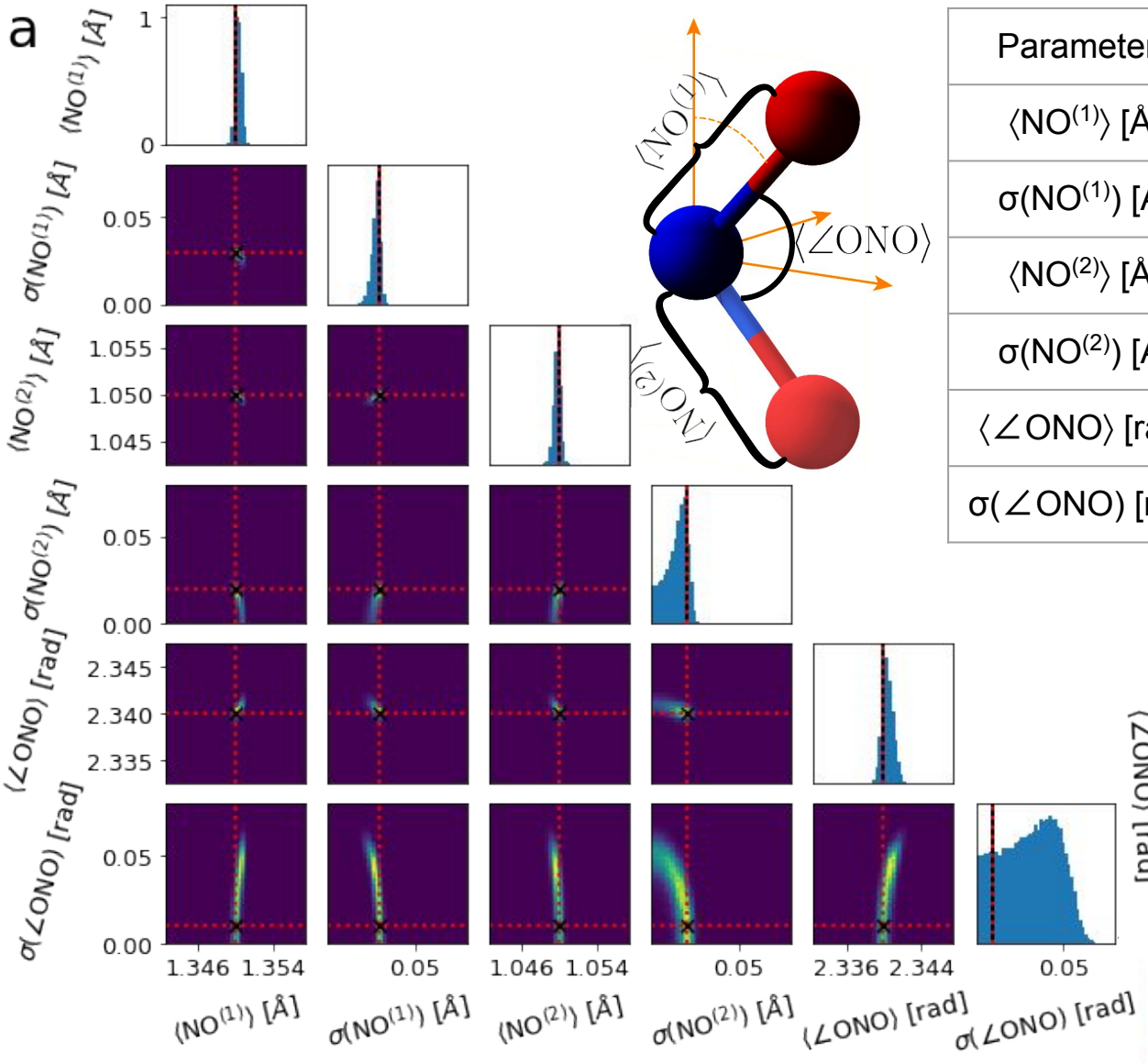
Only 3.5 Å^{-1}!

Retrieving the MF Geometry Probability Distribution Results: P(θ|C)

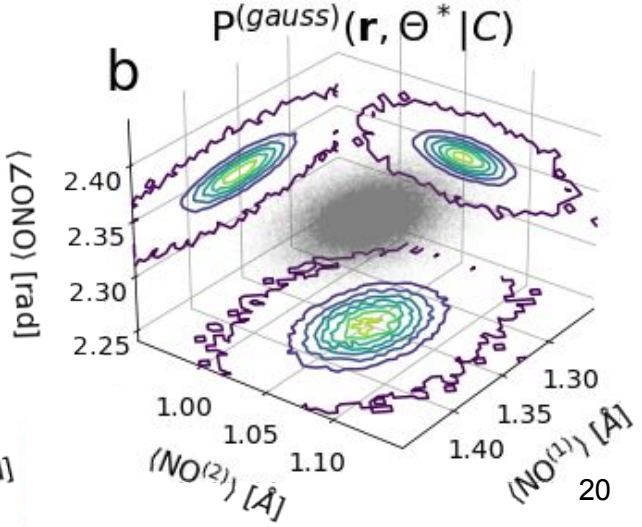


Retrieving the MF Geometry Probability Distribution Results: P(θ|C)

SLAC

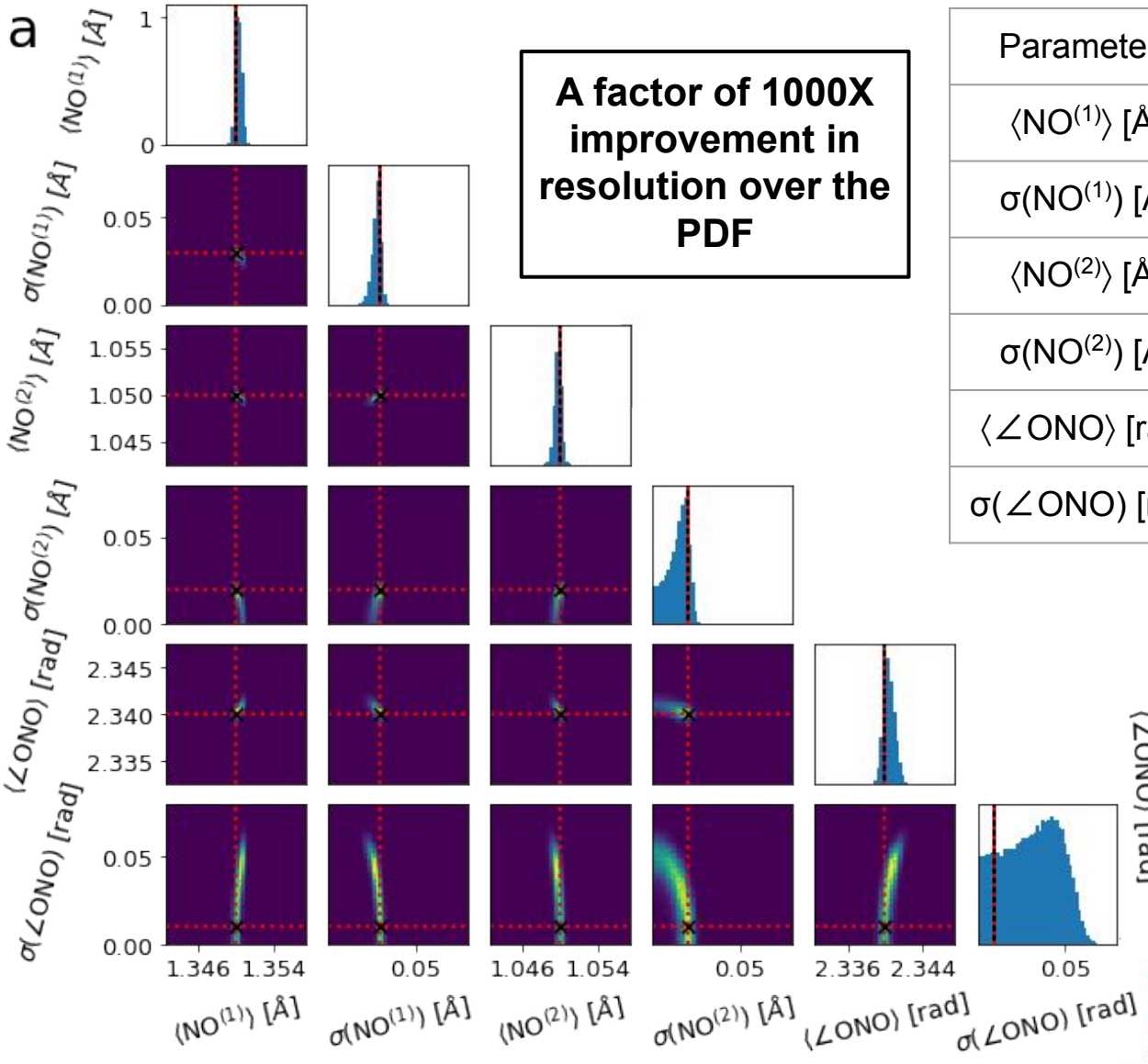


Parameters	Input	Retrieved	Resolution
⟨NO$^{(1)}$⟩ [Å]	1.35	1.35	0.00029
σ(NO$^{(1)}$) [Å]	0.03	0.03	0.0019
⟨NO$^{(2)}$⟩ [Å]	1.05	1.05	0.00029
σ(NO$^{(2)}$) [Å]	0.02	0.02	0.0054
⟨∠ONO⟩ [rad]	2.34	2.34	0.00047
σ(∠ONO) [rad]	0.01	0.0101	0.015



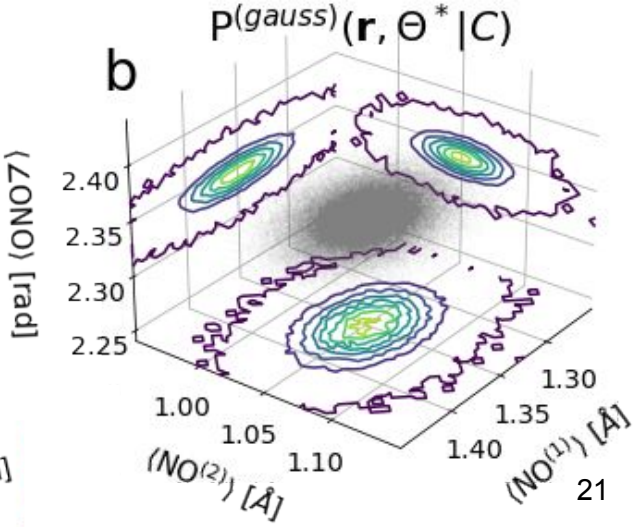
Retrieving the MF Geometry Probability Distribution Results: P(θ|C)

SLAC



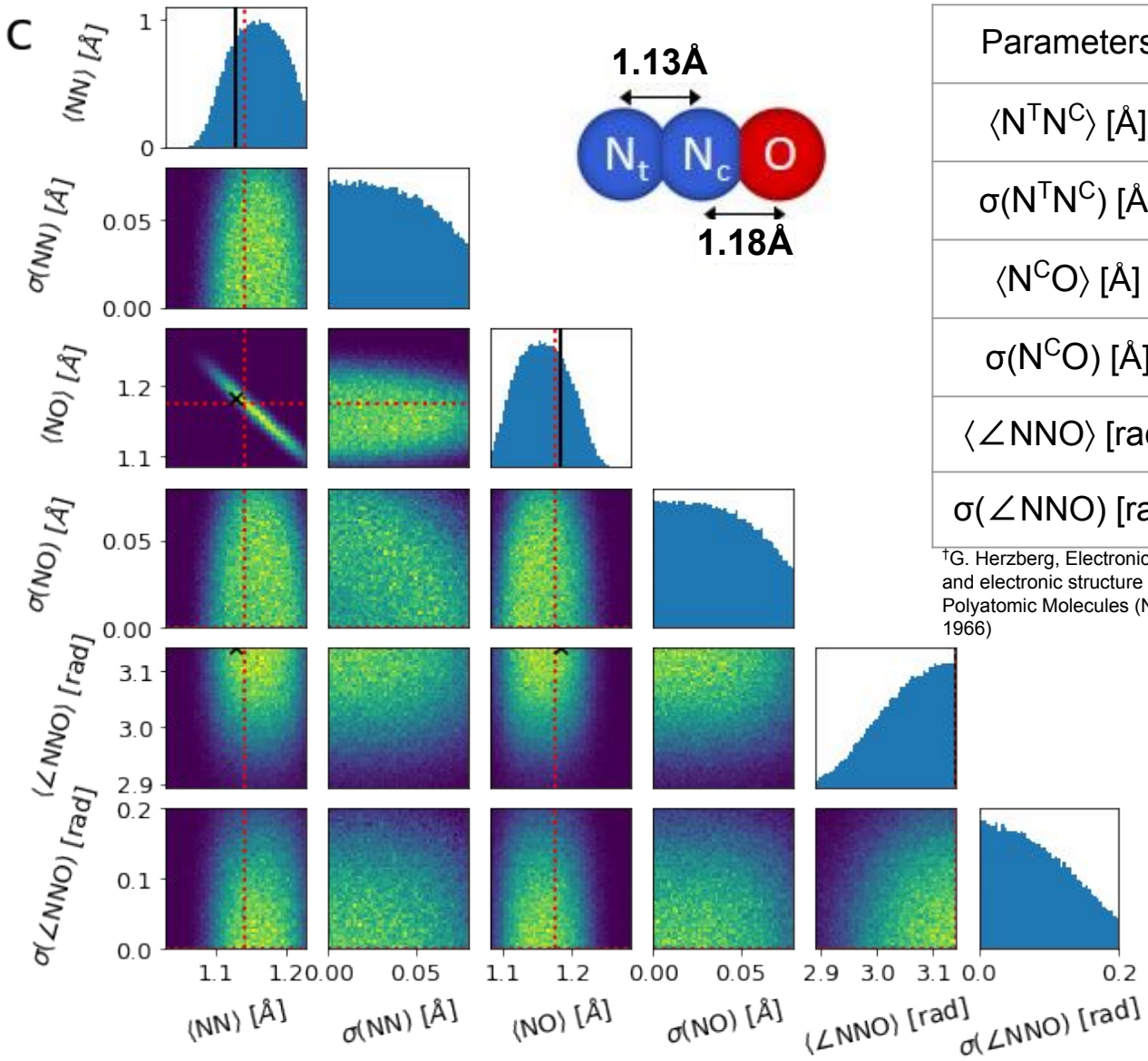
A factor of 1000X improvement in resolution over the PDF

Parameters	Input	Retrieved	Resolution
$\langle NO^{(1)} \rangle$ [Å]	1.35	1.35	0.00029
$\sigma(NO^{(1)})$ [Å]	0.03	0.03	0.0019
$\langle NO^{(2)} \rangle$ [Å]	1.05	1.05	0.00029
$\sigma(NO^{(2)})$ [Å]	0.02	0.02	0.0054
$\langle \angle ONO \rangle$ [rad]	2.34	2.34	0.00047
$\sigma(\angle ONO)$ [rad]	0.01	0.0101	0.015



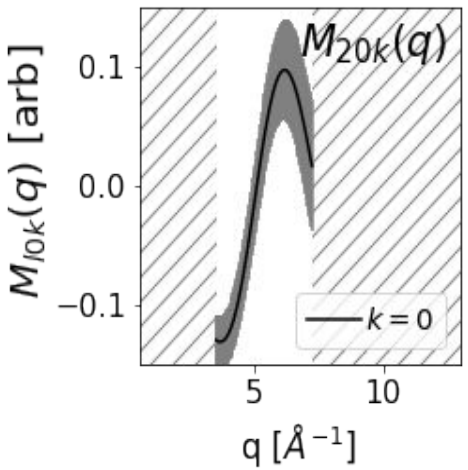
Retrieving the MF Geometry Probability Distribution
Results: P(θ|C)

SLAC



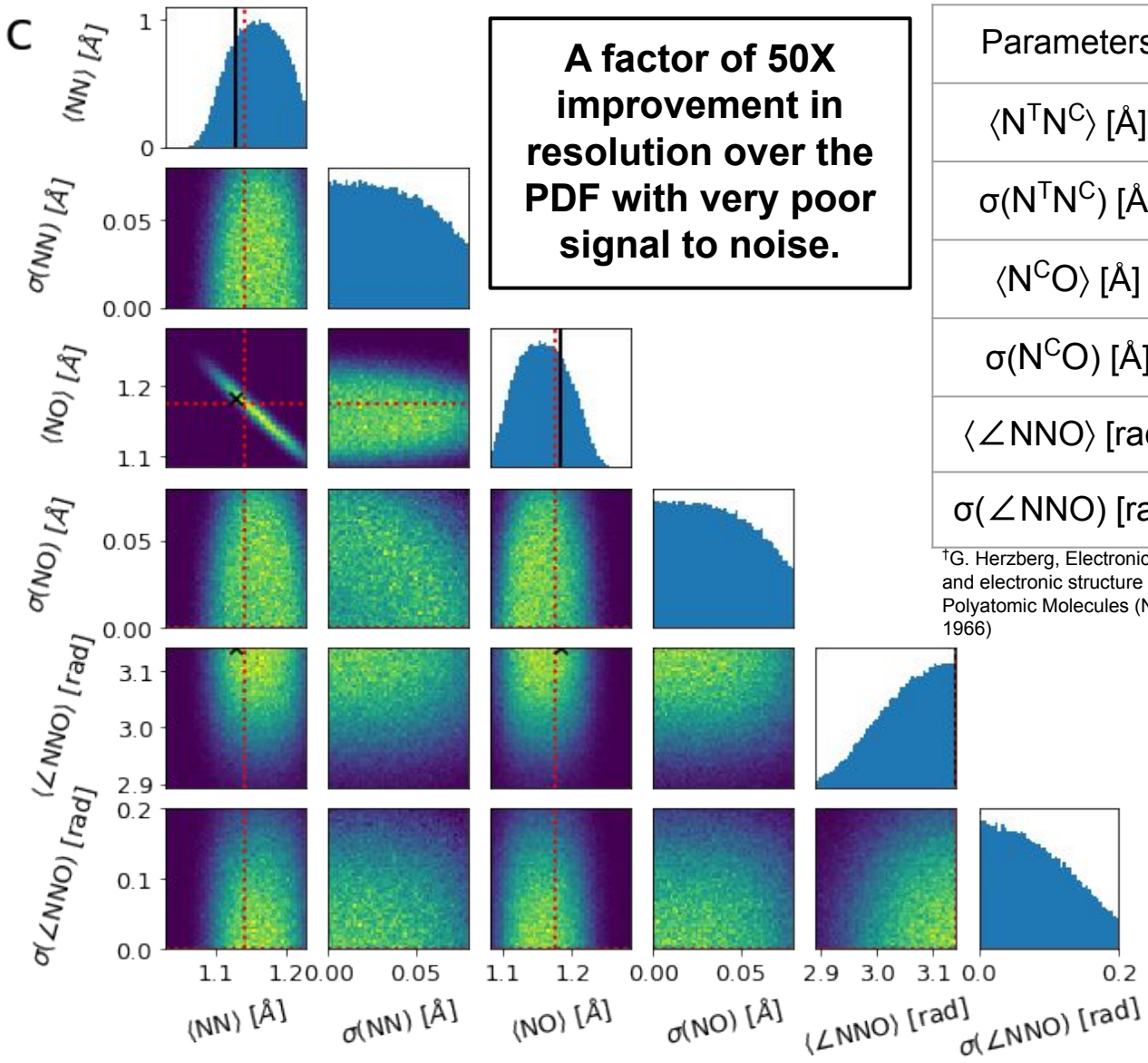
Parameters	Lit.†	Retrieved	Resolution
$\langle N^T N^C \rangle$ [Å]	1.128	1.143	0.039
$\sigma(N^T N^C)$ [Å]		0.081	0.028
$\langle N^C O \rangle$ [Å]	1.184	1.175	0.036
$\sigma(N^C O)$ [Å]		$3.08*10^{-8}$	0.027
$\langle \angle NNO \rangle$ [rad]	3.142	3.142	0.061
$\sigma(\angle NNO)$ [rad]		$5.5*10^{-12}$	0.062

†G. Herzberg, Electronic Spectra and electronic structure of Polyatomic Molecules (Nostrand, 1966)



Retrieving the MF Geometry Probability Distribution Results: P(θ|C)

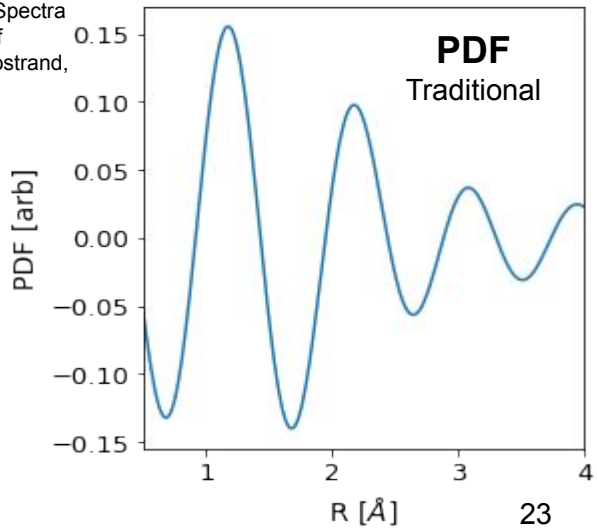
SLAC



A factor of 50X improvement in resolution over the PDF with very poor signal to noise.

Parameters	Lit.†	Retrieved	Resolution
$\langle N^T N^C \rangle$ [Å]	1.128	1.143	0.039
$\sigma(N^T N^C)$ [Å]		0.081	0.028
$\langle N^C O \rangle$ [Å]	1.184	1.175	0.036
$\sigma(N^C O)$ [Å]		$3.08*10^{-8}$	0.027
$\langle \angle NNO \rangle$ [rad]	3.142	3.142	0.061
$\sigma(\angle NNO)$ [rad]		$5.5*10^{-12}$	0.062

†G. Herzberg, Electronic Spectra and electronic structure of Polyatomic Molecules (Nostrand, 1966)



Retrieving Excited State Dynamics

SLAC

$$\langle I(\mathbf{q})\rangle^{(2)}_{\text{sep}}(t,\tau) = \mathcal{I}\Bigg(\sum_{\mu} |f_\mu(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \text{Re}\Bigg\{ f_\mu(q) f^*_\nu(q) \sum_l i^l 8\pi^2 \sqrt{4\pi\,(2l+1)} \sum_{m_1,m_2} (-1)^{m_1-m_2} \times Y_l^{-m_2}\left(\theta_q^{(\text{lf})}, \phi_q^{(\text{lf})}\right) \tilde{\mathcal{A}}^l_{m_1 m_2}(n,n';\tau) \left\langle \psi^{n'}_{\text{el-vib}}(t) \middle| j_l(q\Delta r_{\mu\nu}) Y_l^{-m_1}\left(\theta^{(\text{mf})}_{\mu\nu}, \phi^{(\text{mf})}_{\mu\nu}\right) \middle| \psi^{n}_{\text{el-vib}}(t) \right\rangle \Bigg\} \Bigg)$$

- Separation of rotational and vibrational time scales
- Independent atom approximation
- **Measurement:** Lab frame anisotropy
- **Simulation:** Ensemble anisotropy
- **Result:** Molecular frame geometry
  - Explicit dependence on MF geometric angles
  - Each pair-wise contribution is labelled

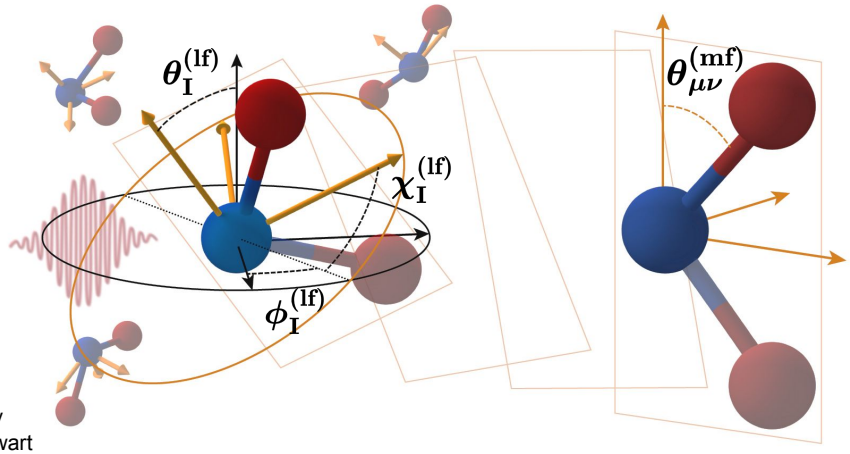


Illustration by Gregory Stewart

Rotational dynamics is of order a few to 10 picoseconds and vibrational or isomerization dynamics of interest are often on the femtosecond timescale. If sufficient rotation occurs outside of the ground state geometry one can use $C_{000}(q,t)$ which is independent of anisotropy but does not have an explicit dependence on the molecular frame angles.

Summary

Khegazy@stanford.edu

SLAC

- Directly probe $|\Psi(\boldsymbol{r}^{(mf)})|^2$ in a high dimensional space
  - Rigorously retrieve distribution of geometry parameters in high dimensions
  - **MHA: efficient, unconstrained, and unbiased geometric search**
- Generally applicable to current experiments and (potentially) excited state dynamics
  - **Choice of model (Normal, …) allows novel measurements like width**
  - Can use induced anisotropy from excitation dipole and $C_{000}(q,t)$
- High resolution (roughly 100-1000 times better than PDF)
  - **Simulation resolutions distances/angles: Order 1 - 0.1 pm / ~1 mrad**
  - **Data with LMK=[2,0,0] and q=[3.5,7] Å^{-1} distances/angles: ~4 pm / 6 mrad**
  - High signal to noise is more important than q range
- **Potentially turn ultrafast gas phase diffraction into a discovery oriented technique without requiring excited state simulations**
  - This analysis only requires knowledge of the ground state geometry and ensemble anisotropy simulations which are much more tractible than excited state dynamics

Ryan Coffee

Varun Makhija

SLAC UED Team

U.S. DEPARTMENT OF ENERGY | Office of Science

Backup Slides

Accessing the MF via Deterministic Anisotropy

SLAC

$$\langle I(\mathbf{q})\rangle(t) = \mathcal{I}\Bigg(\sum_{\mu} |f_{\mu}^{*}(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Bigg\{ f_{\mu}(q) f_{\nu}^{*}(q) \sum_{l} i^{l} 8\pi^2 \sqrt{4\pi\,(2l+1)}$$

$$\times \sum_{m,k} (-1)^{k-m} \underbrace{Y_{l}^{-m}\left(\theta_{q}^{(\mathrm{lf})}, \phi_{q}^{(\mathrm{lf})}\right)}_{\text{Lab Frame}} \langle \Psi(t)| \underbrace{D_{mk}^{l}\left(\phi_{\mathrm{I}}^{(\mathrm{lf})}, \theta_{\mathrm{I}}^{(\mathrm{lf})}, \chi_{\mathrm{I}}^{(\mathrm{lf})}\right)}_{\text{Ensemble Anisotropy}} \underbrace{j_{l}(q\Delta r_{\mu\nu}) Y_{l}^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)}_{\text{Molecular Frame Geometry}} |\Psi(t)\rangle \Bigg\} \Bigg)$$

Independent atom approximation

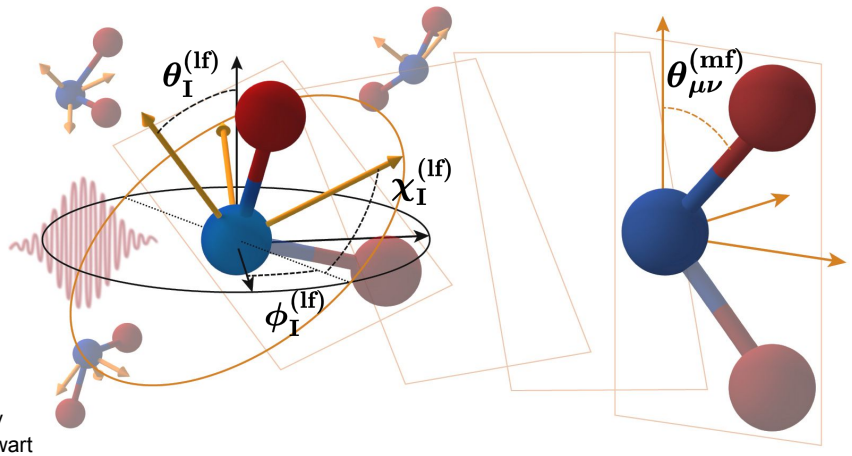
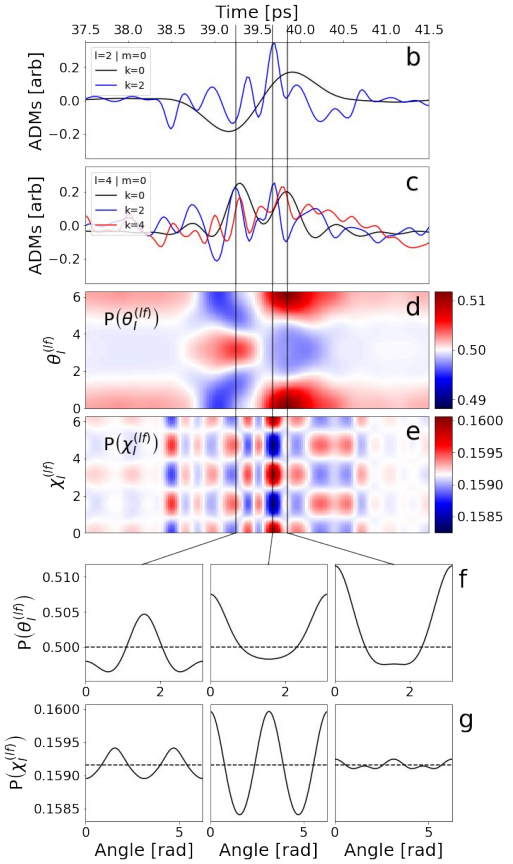


Illustration by
Gregory Stewart



Anisotropy Reveals the Molecular Frame (MF)

- Anisotropy provides constraints on molecular frame (MF) degrees of freedom.
- Combining many measurements (constraints) allows one to retrieve the MF.
- Application in photo-electron spectroscopy
  - V. Makhija, et. al., (2016), arXiv:1611.06476 [physics.atom-ph]
  - C. Marceau, et. al., Phys. Rev. Lett. 119, 083401 (2017)
  - M. Gregory, et. al., (2020), arXiv:2012.04561 [physics.chem-ph]
- Use a stretched NO_2: an asymmetric top

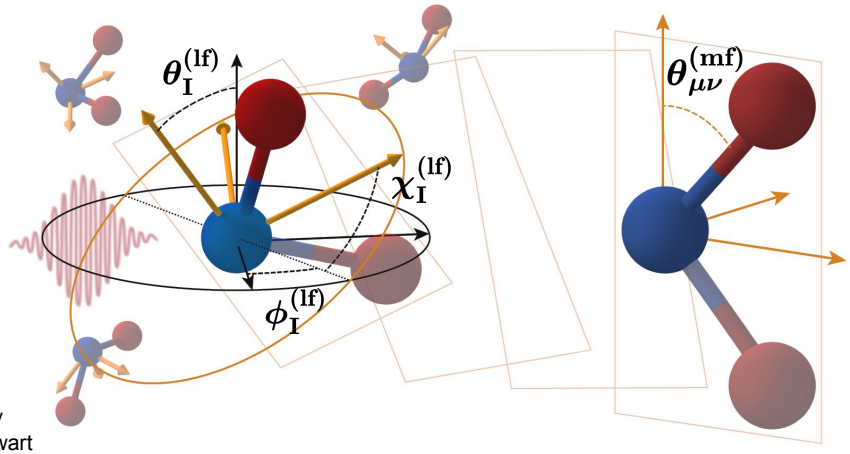
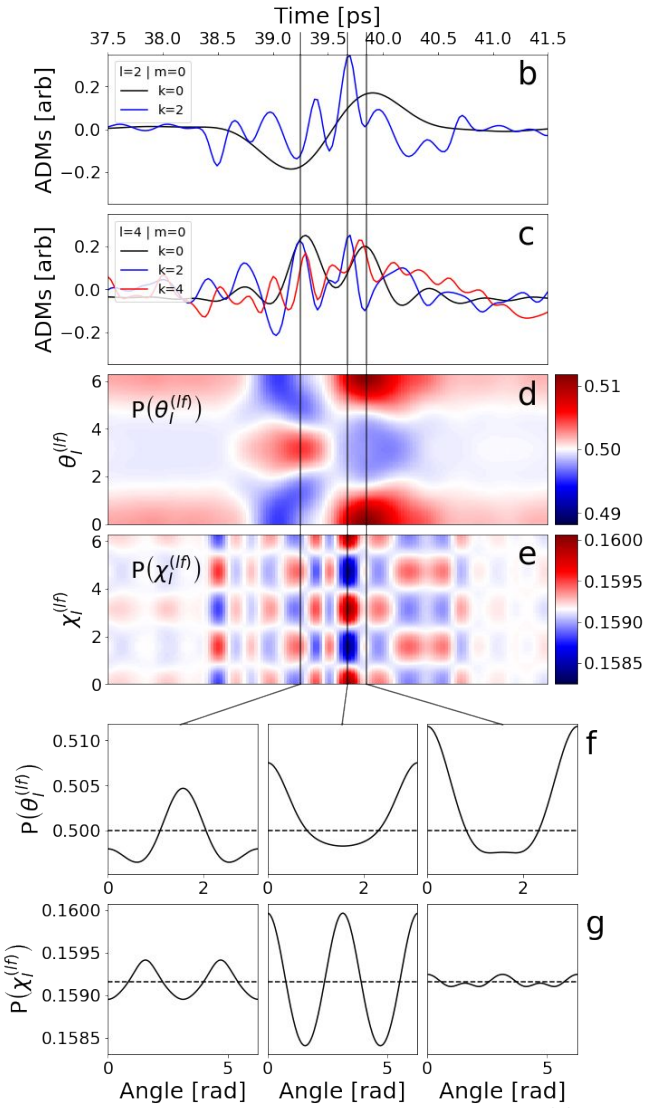


Illustration by Gregory Stewart



NO_2 rotational wavepacket anisotropy signature

Accessing the MF via Deterministic Anisotropy

SLAC

$$\langle I(\mathbf{q})\rangle(t) = \mathcal{I}\Bigg(\sum_{\mu} |f_{\mu}^{*}(q)|^2 + \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Bigg\{ f_{\mu}(q) f_{\nu}^{*}(q) \sum_{l} i^{l} 8\pi^2 \sqrt{4\pi\,(2l+1)} \times \sum_{m,k} (-1)^{k-m} \underbrace{Y_{l}^{-m}\left(\theta_{q}^{(\mathrm{lf})}, \phi_{q}^{(\mathrm{lf})}\right)}_{\text{Lab Frame}} \langle \Psi(t) | \underbrace{D_{mk}^{l}\left(\phi_{\mathrm{I}}^{(\mathrm{lf})}, \theta_{\mathrm{I}}^{(\mathrm{lf})}, \chi_{\mathrm{I}}^{(\mathrm{lf})}\right)}_{\text{Ensemble Anisotropy}} \underbrace{j_{l}(q\Delta r_{\mu\nu}) Y_{l}^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)}_{\text{Molecular Frame Geometry}} | \Psi(t) \rangle \Bigg\} \Bigg)$$

- Independent atom approximation
- **Measurement:** Lab frame anisotropy
- **Simulation:** Ensemble anisotropy
- **Result:** Molecular frame geometry
  - Explicit dependence on MF geometric angles
  - Each pair-wise contribution is labelled

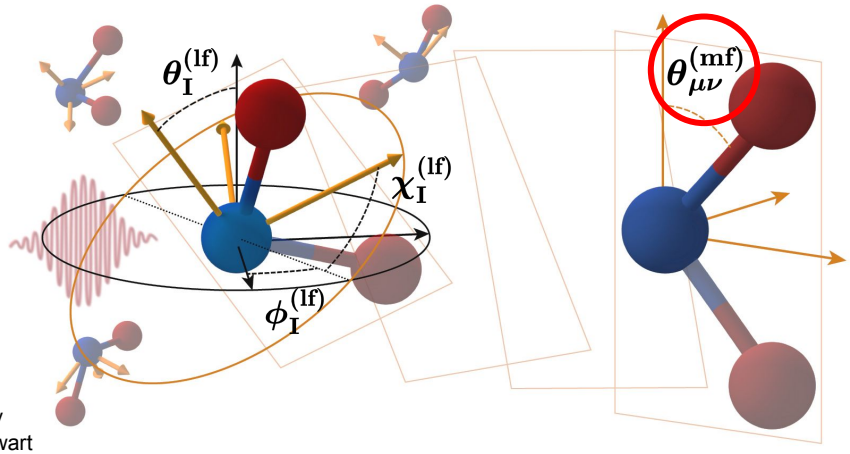


Illustration by Gregory Stewart

Explicit dependence on molecular frame angles provides an opportunity to develop methods that extract these angles, and hence a unique geometry, from data alone.

Retrieving the MF Geometry Probability Distribution
Delta Distribution and Systematic Error

SLAC

Delta Distribution Systematic Error

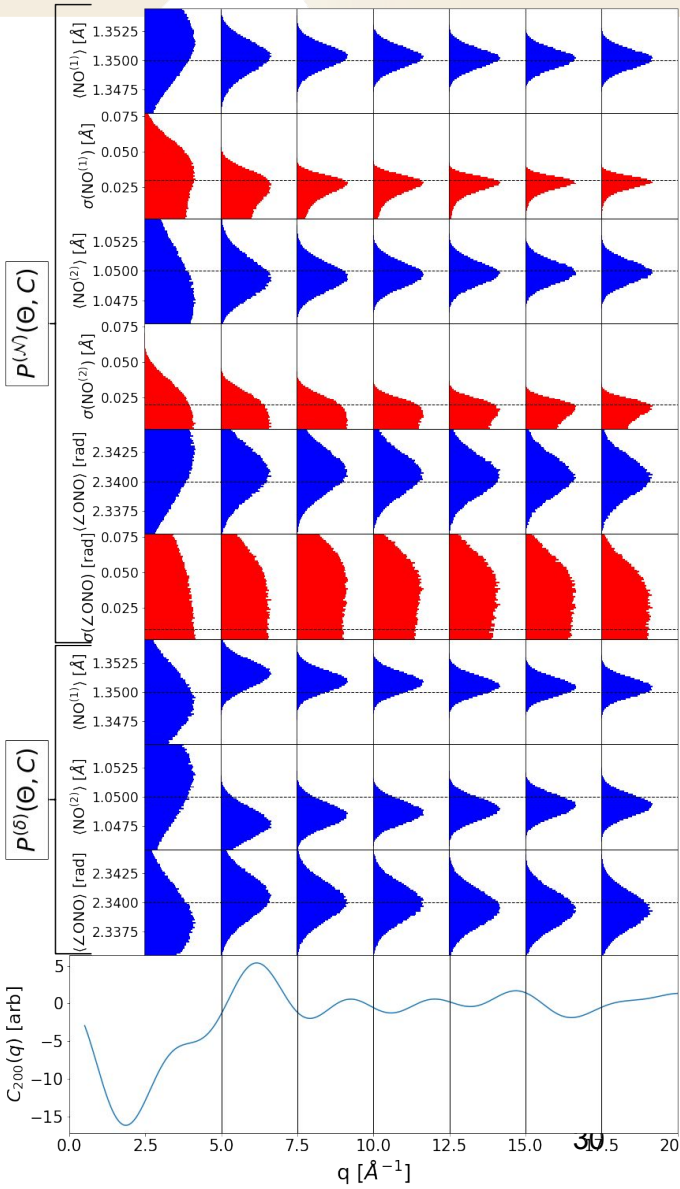
- Assumes signal from ensembles and a single geometry are comparable
- Accuracy is roughly 1000 times worse
- At ≤picometer resolution, the correct value can be 2-3 standard deviations of P(**θ**|C)

Normal Distribution Mitigates this Systematic

- Normal distribution P(**θ**|C) distributions are closely centered around the expected value
- The 1d mode does not change with q range

Why use the delta distribution?

- It is of order 100 times faster
- Provides sufficient accuracy for debugging analysis



Convergence: Autocorrelation Time (τ)

SLAC

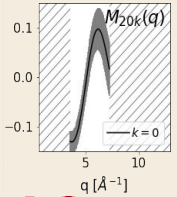
Criteria for each chain (1000 chains are used)

- At least 100 τ in length
- $\Delta\tau/\tau$ < 0.01

Autocorrelation Time

- Number of steps needed for geometries to be uncorrelated
- Expectations: $E_{p(\theta)}\left[f(\theta)\right] \approx \frac{1}{N}\sum_{n}^{N} f\left(\theta^{(n)}\right)$ $\qquad \mathrm{SEM} = \sqrt{\frac{\mathrm{Var}_{p(\theta)}[f(\theta)]}{N}}$
- Correlation Effects: $\mathrm{SEM} = \sqrt{\frac{\tau}{N}\mathrm{Var}_{p(\theta)}[f(\theta)]}$

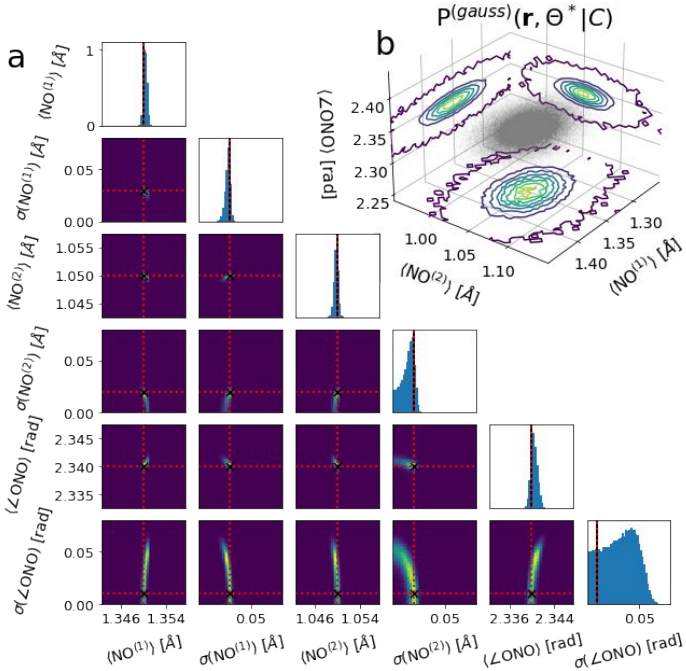
Retrieving the MF Geometry Probability Distribution Results



SLAC

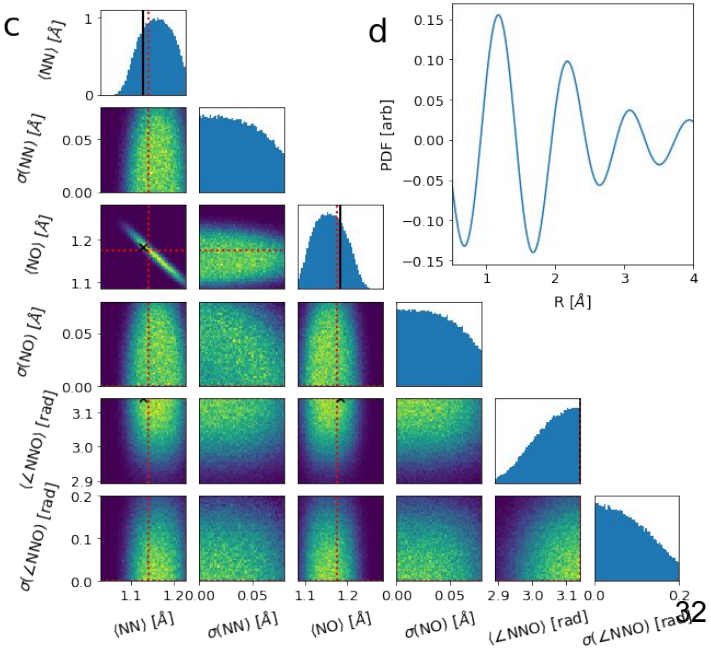
Simulated NO_2

	Input	Θ^*	σ^Θ
$\langle NO^{(1)} \rangle$ [Å]	1.35	1.3500	0.00029
$\sigma\left(NO^{(1)}\right)$ [Å]	0.03	0.03000	0.0019
$\langle NO^{(2)} \rangle$ [Å]	1.05	1.0500	0.00029
$\sigma\left(NO^{(2)}\right)$ [Å]	0.02	0.02000	0.0054
$\langle \angle ONO \rangle$ [rad]	2.34	2.340	0.00047
$\sigma\left(\angle ONO\right)$ [rad]	0.01	0.01010	0.015



Measured N_2O

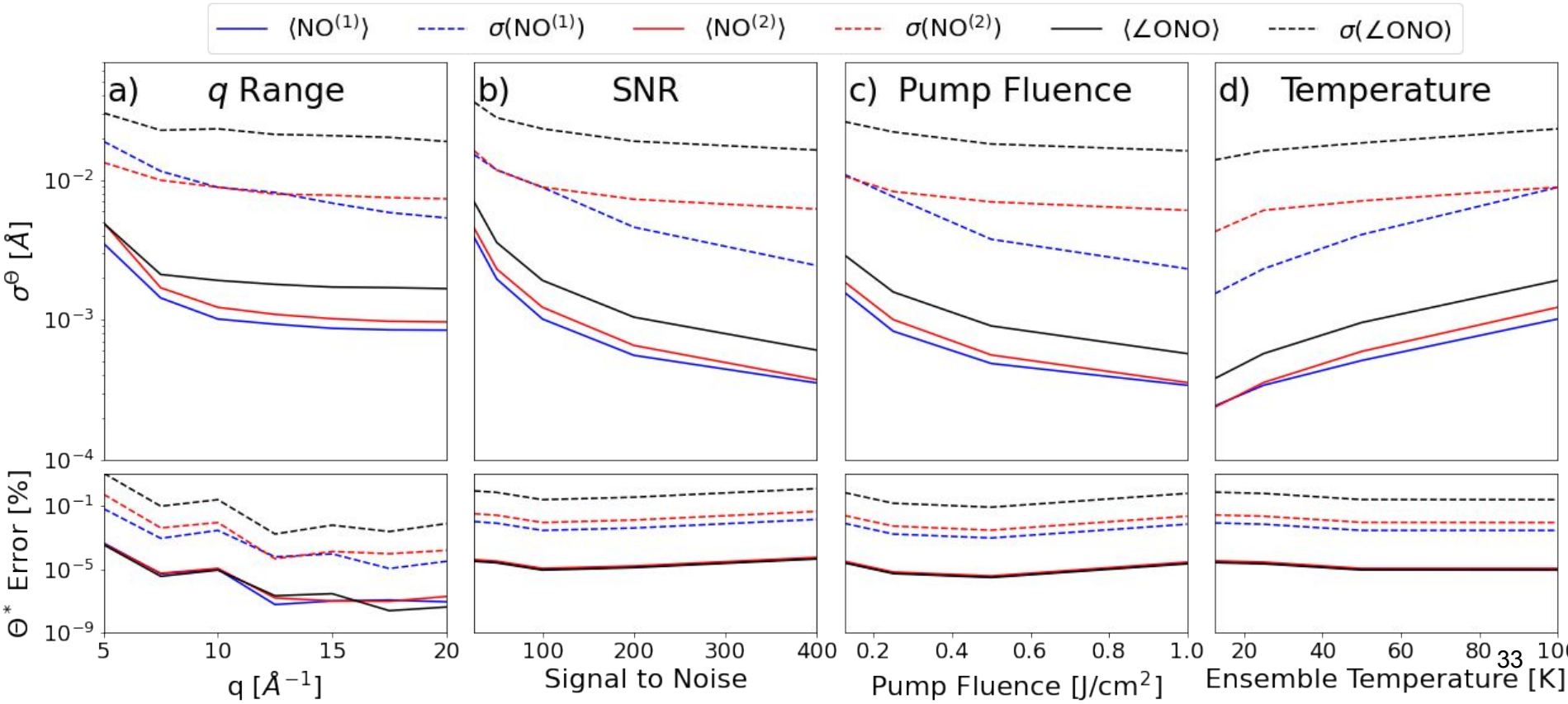
	$\Theta^*_{\text{Literature}}$	Θ^*	σ^Θ
$\langle N^T N^C \rangle$ [Å]	1.128	1.142	0.039
σ (NN) [Å]		0.081	0.028
$\langle NO \rangle$ [Å]	1.184	1.175	0.036
σ (NO) [Å]		3.08×10^{-8}	0.027
$\langle \angle NNO \rangle$ [rad]	3.142	3.142	0.061
σ ($\angle$NNO) [rad]		5.5×10^{-12}	0.062



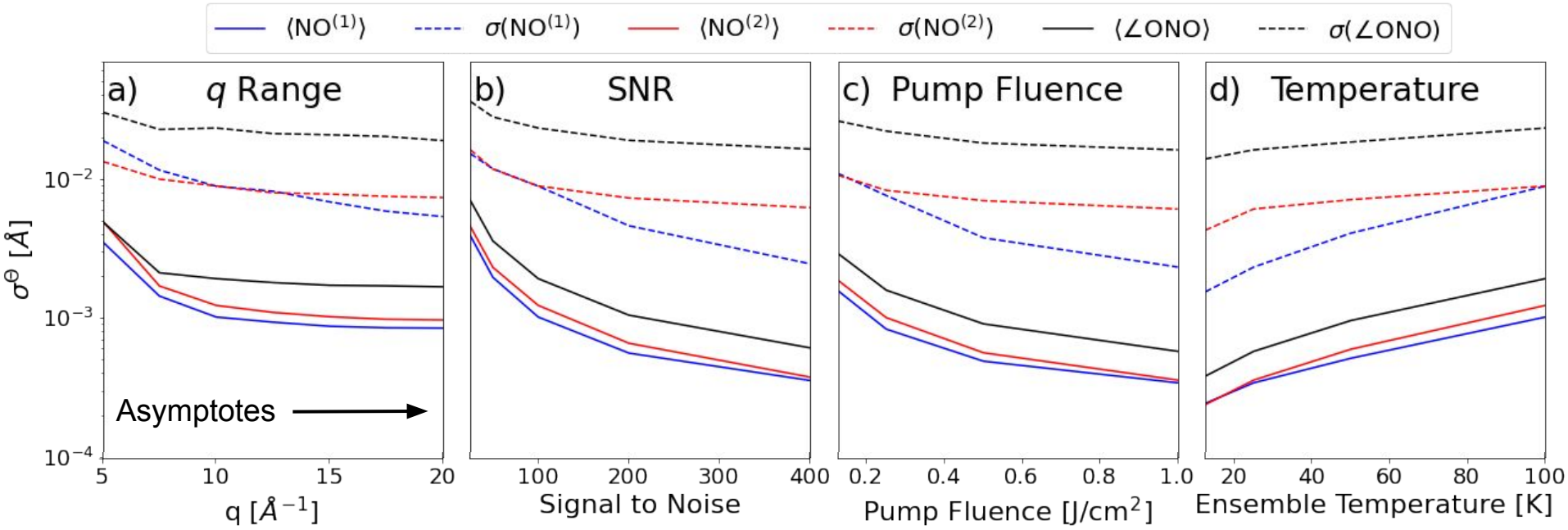
Retrieving the MF Geometry Probability Distribution
Experimental Parameters

SLAC

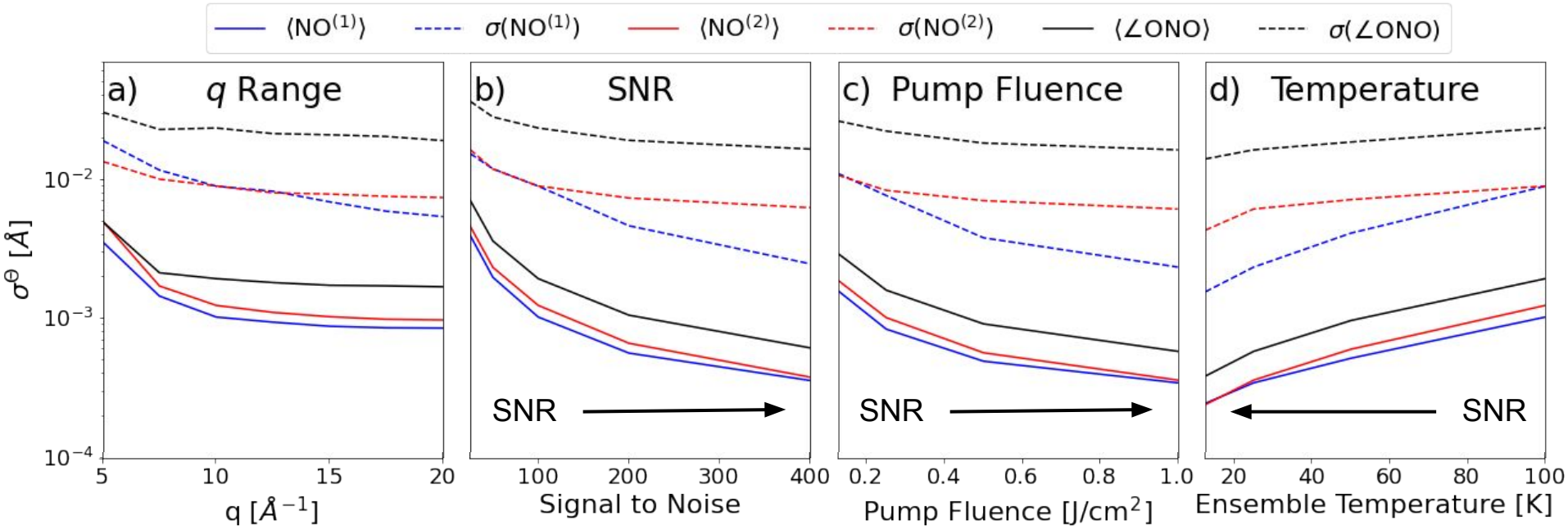
a) q Range	b) SNR	c) Pump Fluence	d) Temperature
q : 0.5 - [5, 20] Å^{-1} SNR: 100 Fluence: 1 J/cm^2 Temp: 100 K	q : 0.5 - 10 Å^{-1} SNR: [25, 400] Fluence: 1 J/cm^2 Temp: 100 K	q : 0.5 - 10 Å^{-1} SNR: 100 Fluence: [0.12, 1] J/cm^2 Temp: 100 K	q : 0.5 - 10 Å^{-1} SNR: 100 Fluence: 1 J/cm^2 Temp: [12.5, 100] K



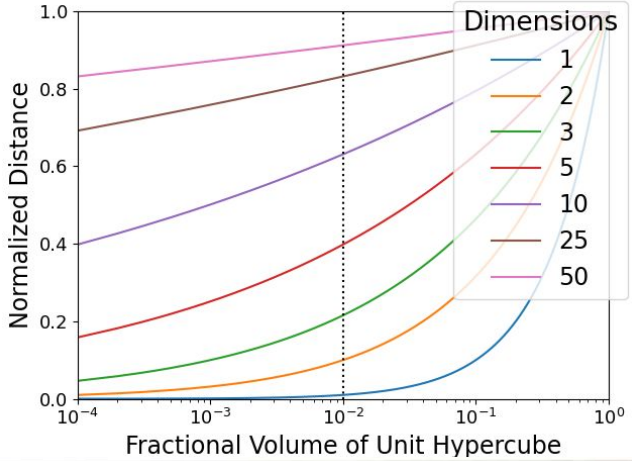
Retrieving the MF Geometry Probability Distribution
Experimental Parameters



Retrieving the MF Geometry Probability Distribution
Experimental Parameters



Curse of Dimensionality

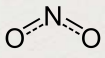


Assumptions

1. Grid Search
2. Ignore hydrogens
3. Know atom pair-wise distances within 1 Å

Degrees of freedom: $N_{dof} = 3N_{atoms} - 6$
Number of samples: N_S
Number of sample per dimension: N_S/D
Sampled volume given SS step size: $V_S|SS$

NO_2



N_{dof}= 3

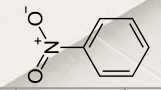
N_S	N_S/D
10^6	100
10^7	215
10^8	464

Cyclohexadiene

N_{dof}= 12

N_S	N_S/D	$V_S\|0.1$	$V_S\|0.25$
10^6	3.16	10^{-5}%	5.96%
10^7	3.83	10^{-4}%	59.6%
10^8	4.64	10^{-3}%	>100%

Nitrobenzene



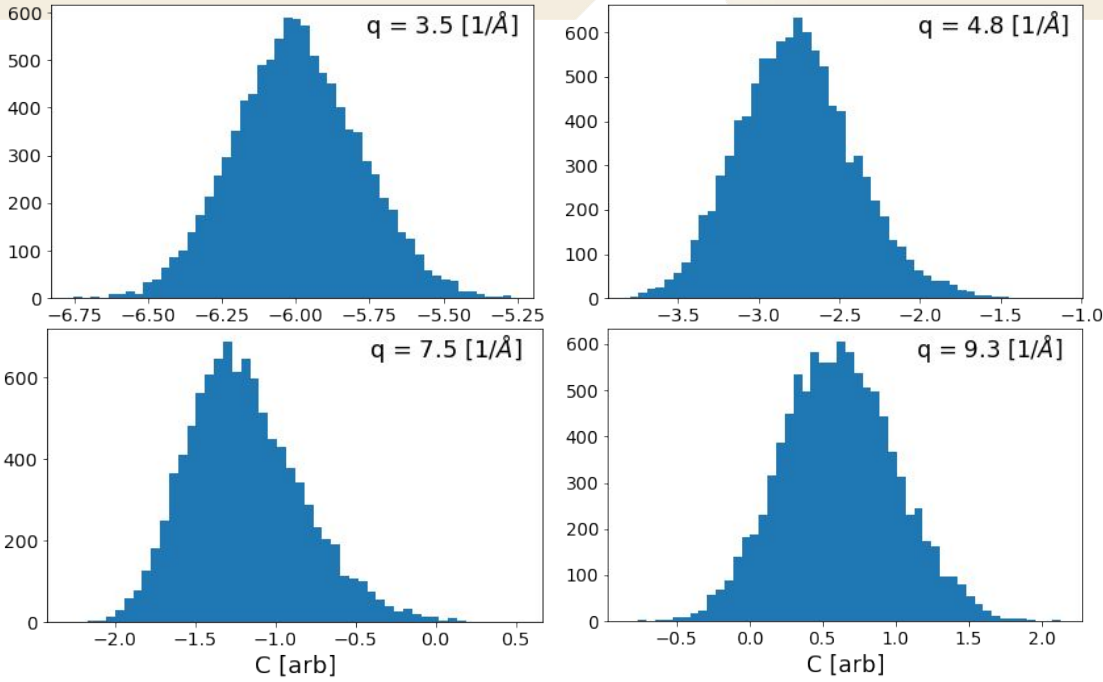
N_{dof}= 21

N_S	N_S/D	$V_S\|0.1$	$V_S\|0.25$
10^6	1.93	10^{-14}%	2.3×10^{-5}%
10^7	2.15	10^{-13}%	2.3×10^{-4}%
10^8	2.40	10^{-12}%	2.3×10^{-3}%

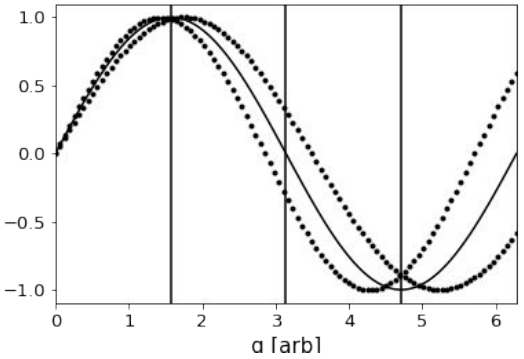
Systematic Errors in Retrieving Geometric Parameters

SLAC

- Systematic errors are caused by non-gaussian $C_{lmk}(q)$ distributions
- Gaussian distributions of geometries lead to non-gaussian distributions due to sinusoids.



$$P(C|\mathbf{r}^{(mf)}) = \prod_{lmk,q} \frac{1}{\sigma_{lmk}(q)\sqrt{2\pi}} \exp\left\{\frac{-1}{2}\left(\frac{C_{lmk}^{(data)}(q) - C_{lmk}^{(calc)}\left(q, \mathbf{r}^{(mf)}\right)}{\sigma_{lmk}(q)}\right)^2\right\}$$



Bayesian Inferencing
Reframing the Problem

SLAC

What we have: Molecular frame representations for the nuclear geometry with explicit dependence on pair-wise distances and angles

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\left\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\left\langle j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)\right\rangle}_{\text{Molecular Frame Geometry}} \right\}$$

$$\longrightarrow \quad C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\left\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \left|\Psi(\mathbf{r}^{(\mathrm{mf})})\right|^2 d\mathbf{r}^{(\mathrm{mf})}}_{\text{Molecular Frame Geometry}} \right\}$$

What we want: Invert the integral equation for $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$

How we do it: Approximate $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$ with a chosen distribution and solve for the model parameter ($\boldsymbol{\theta}$) distribution P($\boldsymbol{\theta}$|C)

$$P(\mathbf{r}, \boldsymbol{\Theta}|\, C) \approx |\Psi(\mathbf{r})|^2$$

$$P(\boldsymbol{\Theta}|C) = \int P(\mathbf{r}, \boldsymbol{\Theta}|\, C)\, d\mathbf{r}$$

Delta Distribution:

$$P^{(\delta)}(\mathbf{r}, \boldsymbol{\Theta}|\, C) = \delta\left(\boldsymbol{\Theta}^{(\mathrm{delta})} - \mathbf{r}\right)$$

$$\boldsymbol{\Theta}^{(\mathrm{delta})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \langle \mathrm{NO}^{(2)}\rangle, \langle \angle \mathrm{ONO}\rangle\right]$$

Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r}, \boldsymbol{\Theta}|\, C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\mathrm{gauss})}} \exp\left\{ \frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\mathrm{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\mathrm{gauss})}} \right)^2 \right\}$$

$$\boldsymbol{\Theta}^{(\mathrm{gauss})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \sigma\left(\mathrm{NO}^{(1)}\right), \langle \mathrm{NO}^{(2)}\rangle, \sigma\left(\mathrm{NO}^{(2)}\right), \langle \angle \mathrm{ONO}\rangle, \sigma(\angle \mathrm{ONO})\right]$$

Bayesian Inferencing
Reframing the Problem

SLAC

What we have: Molecular frame representations for the nuclear geometry with explicit dependence on pair-wise distances and angles

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\left\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\left\langle j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)\right\rangle}_{\text{Molecular Frame Geometry}} \right\}$$

$$\rightarrow \quad C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\left\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \left|\Psi(\mathbf{r}^{(\mathrm{mf})})\right|^2 d\mathbf{r}^{(\mathrm{mf})}}_{\text{Molecular Frame Geometry}} \right\}$$

What we want: Invert the integral equation for $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$

How we do it: Approximate $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$ with a chosen distribution and solve for the model parameter ($\boldsymbol{\theta}$) distribution P($\boldsymbol{\theta}$|C)

$$P(\mathbf{r}, \boldsymbol{\Theta}|\, C) \approx |\Psi(\mathbf{r})|^2$$

$$P(\boldsymbol{\Theta}|C) = \int P(\mathbf{r}, \boldsymbol{\Theta}|\, C)\, d\mathbf{r}$$

When finding $\boldsymbol{\theta}$ one can make novel measurements

Delta Distribution:

$$P^{(\delta)}(\mathbf{r}, \boldsymbol{\Theta}|\, C) = \delta\left(\boldsymbol{\Theta}^{(\mathrm{delta})} - \mathbf{r}\right)$$

$$\boldsymbol{\Theta}^{(\mathrm{delta})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \langle \mathrm{NO}^{(2)}\rangle, \langle \angle \mathrm{ONO}\rangle\right]$$

Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r}, \boldsymbol{\Theta}|\, C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\mathrm{gauss})}} \exp\left\{ \frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\mathrm{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\mathrm{gauss})}} \right)^2 \right\}$$

$$\boldsymbol{\Theta}^{(\mathrm{gauss})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \sigma\left(\mathrm{NO}^{(1)}\right), \langle \mathrm{NO}^{(2)}\rangle, \sigma\left(\mathrm{NO}^{(2)}\right), \langle \angle \mathrm{ONO}\rangle, \sigma(\angle \mathrm{ONO})\right]$$

Bayesian Inferencing
Reframing the Problem

SLAC

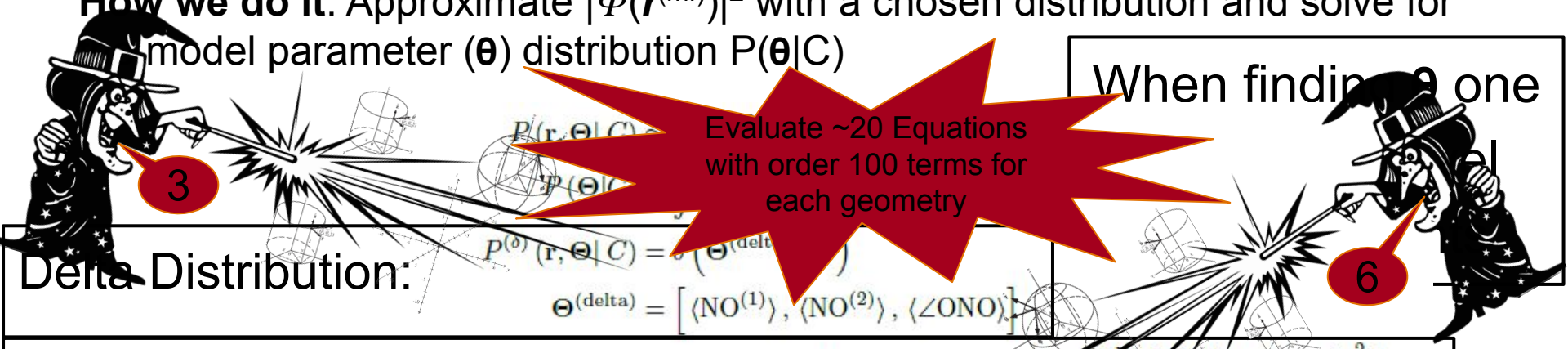
What we have: Molecular frame representations for the nuclear geometry with explicit dependence on pair-wise distances and angles

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\left\langle j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right)\right\rangle}_{\text{Molecular Frame Geometry}} \Big\}$$

$$\longrightarrow \quad C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) |\Psi(\mathbf{r}^{(\mathrm{mf})})|^2 d\mathbf{r}^{(\mathrm{mf})}}_{\text{Molecular Frame Geometry}} \Big\}$$

What we want: Invert the integral equation for $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$

How we do it: Approximate $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$ with a chosen distribution and solve for model parameter ($\boldsymbol{\theta}$) distribution P($\boldsymbol{\theta}$|C)



Delta Distribution:

$$P^{(\delta)}(\mathbf{r}, \Theta|\, C) = \delta\left(\Theta^{(\mathrm{delta})}\right)$$

$$\Theta^{(\mathrm{delta})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \langle \mathrm{NO}^{(2)}\rangle, \langle \angle \mathrm{ONO}\rangle\right]$$

Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r}, \Theta|\, C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\mathrm{gauss})}} \exp\left\{ \frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\mathrm{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\mathrm{gauss})}} \right)^2 \right\}$$

$$\Theta^{(\mathrm{gauss})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \sigma\left(\mathrm{NO}^{(1)}\right), \langle \mathrm{NO}^{(2)}\rangle, \sigma\left(\mathrm{NO}^{(2)}\right), \langle \angle \mathrm{ONO}\rangle, \sigma\left(\angle \mathrm{ONO}\right)\right]$$

Bayesian Inferencing
Reframing the Problem

What we have: Molecular frame representations for the nuclear geometry with explicit dependence on pair-wise distances and angles

$C_{lmk}(q)$ ... $(q) f_\nu^*(q)(-1)^{k-m} i^l 8\pi^2$ ... $\phi_{\mu\nu}^{(\mathrm{mf})} \rangle \}$ ... metry

→

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q)(-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})} \ldots^{(\mathrm{mf})}\right) \Big\}$$

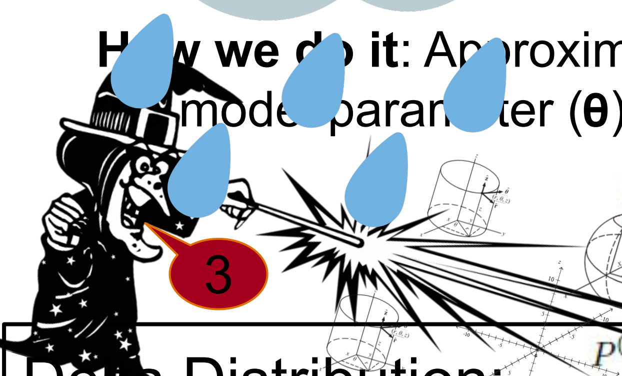
Molecu...

Metropolis Hastings Algorithm

Metropolis Hastings Algorithm

Wh... ...nvert the integral equation for $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$

H...**w we d**...**o it**: Ap...proximate $|\Psi(\boldsymbol{r}^{(\mathrm{mf})})|^2$ with a chosen distribution... mode... paran...ter ($\boldsymbol{\theta}$) distribution P($\boldsymbol{\theta}$|C)



Evaluate ~20 Equations with order 100 terms for each geometry

$P(\mathbf{r}, \Theta | C)$ ...

... $P(\Theta | C$ ...

Wh...n findi... θ one ...

Delta Distribution: $P^{(\delta)}(\mathbf{r}, \Theta | C) = \delta\left(\Theta^{(\mathrm{delt}\ldots}\ldots\right)$

$$\Theta^{(\mathrm{delta})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \langle \mathrm{NO}^{(2)}\rangle, \langle \angle \mathrm{ONO}\rangle\right]$$

Normal Distribution:

$$P^{(\mathcal{N})}(\mathbf{r}, \Theta | C) = \frac{1}{\sqrt{2\pi}^{N_{dof}} \prod_{i=0}^{i<N_{dof}} \Theta_{2i+1}^{(\mathrm{gauss})}} \exp\left\{ \frac{-1}{2} \sum_{i=0}^{i<N_{dof}} \left(\frac{\Theta_{2i}^{(\mathrm{gauss})} - \mathbf{r}_i}{\Theta_{2i+1}^{(\mathrm{gauss})}} \right)^2 \right\}$$

$$\Theta^{(\mathrm{gauss})} = \left[\langle \mathrm{NO}^{(1)}\rangle, \sigma\left(\mathrm{NO}^{(1)}\right), \langle \mathrm{NO}^{(2)}\rangle, \sigma\left(\mathrm{NO}^{(2)}\right), \langle \angle \mathrm{ONO}\rangle, \sigma\left(\angle \mathrm{ONO}\right)\right]$$

Bayesian Inferencing
Metropolis Hastings Algorithm

SLAC

- Inverts the system of equations to solve for the joint P(**θ**|C) distribution
- **Unbiased** sampling method designed for high dimensional spaces

$$C_{lmk}(q) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \boxed{|\Psi(\mathbf{r})|^2} d\mathbf{r}}_{\text{Molecular Frame Geometry}} \Big\}$$

Bayesian Inferencing

$$C_{lmk}^{(\mathrm{calc})}(q,\Theta) = \mathcal{I} \sum_{\mu,\nu:\mu\neq\nu} \mathrm{Re}\Big\{ f_\mu(q) f_\nu^*(q) (-1)^{k-m} i^l 8\pi^2 \times \sqrt{\frac{4\pi}{(2l+1)}} \underbrace{\int j_l(q\Delta r_{\mu\nu}) Y_l^{-k}\left(\theta_{\mu\nu}^{(\mathrm{mf})}, \phi_{\mu\nu}^{(\mathrm{mf})}\right) \boxed{P(\mathbf{r},\Theta|C)} d\mathbf{r}}_{\text{Molecular Frame Geometry}} \Big\}$$

Method

1. Select neighboring **θ**
2. Add new **θ** to P(**θ**|C) with probability $P(C|\boldsymbol{\theta})_{\mathrm{New}}/P(C|\boldsymbol{\theta})_{\mathrm{Prev}}$

$$P\left(C|\boldsymbol{\Theta}\right) = e^{L(\boldsymbol{\Theta})} \left[\prod_{lmk,q} \frac{1}{\sigma_{lmk}(q)\sqrt{2\pi}}\right] \exp\left\{\frac{-1}{2} \sum_{lmk,q} \left(\frac{C_{lmk}(q) - C_{lmk}^{(\mathrm{calc})}(q,\boldsymbol{\Theta})}{\sigma_{lmk}(q)}\right)^2\right\}$$

SLAC National Accelerator Laboratory