UWM

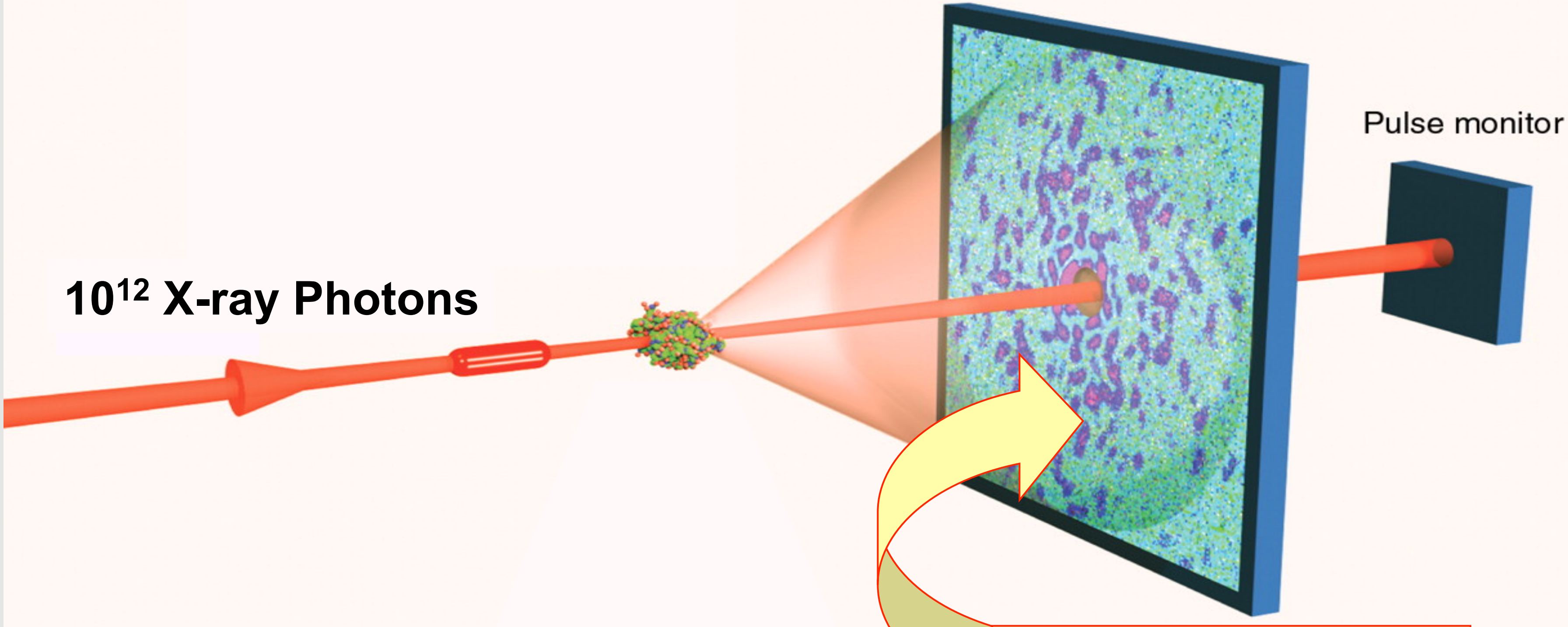
UNIVERSITY of WISCONSIN

Structure & Dynamics from Random Snapshots of Heterogeneous Ensembles

Abbas Ourmazd

ourmazd@uwm.edu

Single-particle Diffract & Destroy



10^{12} X-ray Photons

Total photons: ~ 100

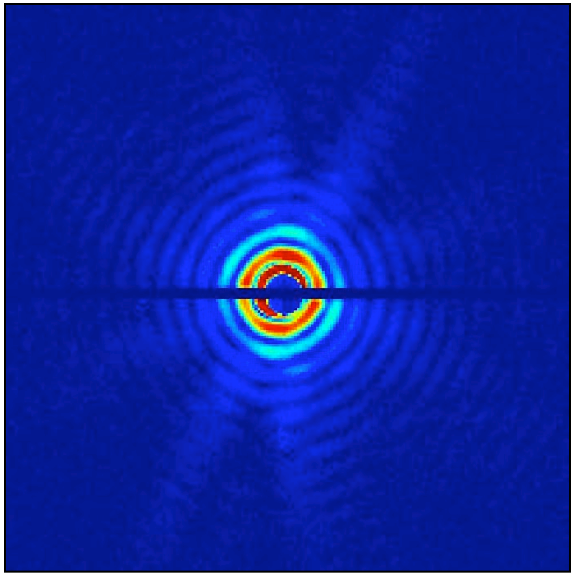
Photons/pixel: ~ 10^{-2}

Unknown orientation

Graphic from Gaffney & Chapman; Science, 316, 1444 (2007)

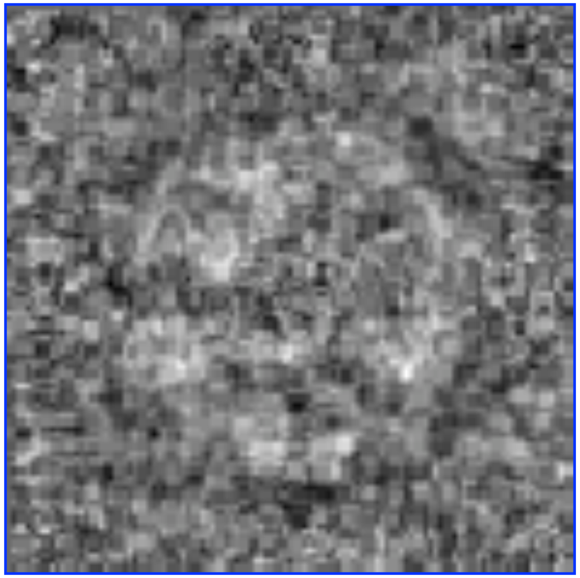
Single-particle Snapshots

XFEL



T4 Virus
Schlichting et al.

Cryo-EM



Chaperonin Molecule
Chiu et al.

Recover 3D structure from 2D snapshots

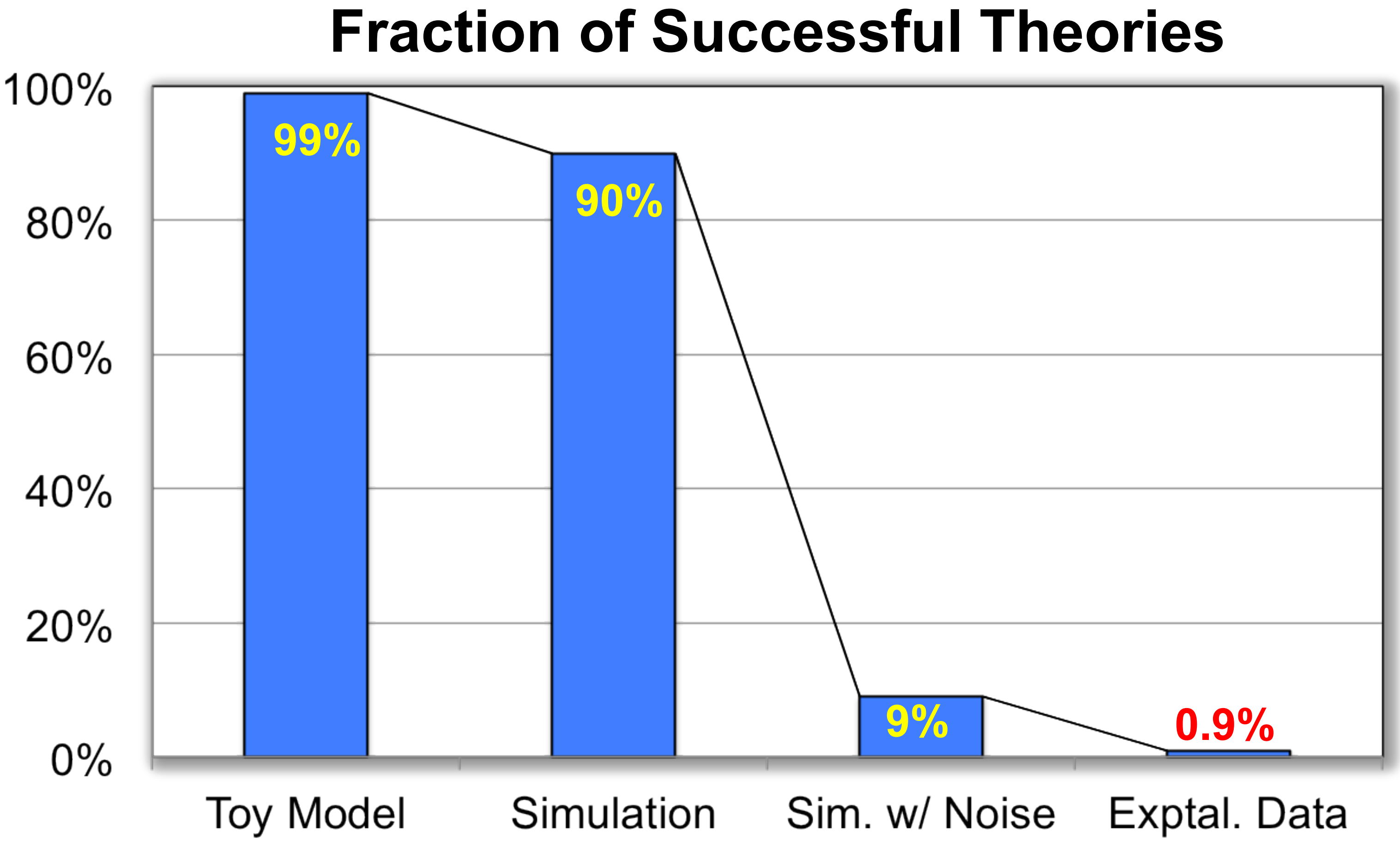
Structure vs. Nature

- Radiation damage limits resolution
  - Even in XFEL “scatter-then-destroy”
- Spread the dose over many objects
  - Crystallography, cryo-EM, XFEL
- Reconstruct assuming identical objects
  - If different, separate into sets of nearly identical objects
- Heterogeneity common & fundamental; limits resolution!
  - NOT low signal, or low scattering q
- Heterogeneity implies dynamics
  - In physics, chemistry, biology
- New approaches for recovering structure & dynamics
  - From random snapshots of heterogeneous ensembles

Outline

- Single particles
  - Reconstruct homogeneous ensembles
- Molecular machines
  - Map conformations
- Dynamics
  - Time-resolved, strong-field reactions
- Conclusions
  - The ultimate “3D movie”?
- Simulated and experimental data

Caveat Emptor!



Buyer beware!

Key Messages

- Identical ensembles (molecules, viruses, nanocrystals)
  - Can reconstruct at expected signal, or lower
  - Computation limits object size, heterogeneity limits resolution
  - Experimental issues with XFEL (intensity fluctuations, hit-rate, etc.)
  - 3Å resolution bar set by cryo-EM (*Science*, Aug. 27, 2010)
- Heterogeneous ensembles (molecular machines)
  - Can sort multiple discrete species and/or conformations
  - Way appears open to mapping conformations
  - *NJP* 12, 1 (2010); *Nature* 466, 329 (2010)
- Dynamics (ultrafast processes)
  - Reactions & their time evolution
- Key tool: Manifold mapping
  - Non-Bayesian graph theory

Identical Objects

3D structure from snapshots

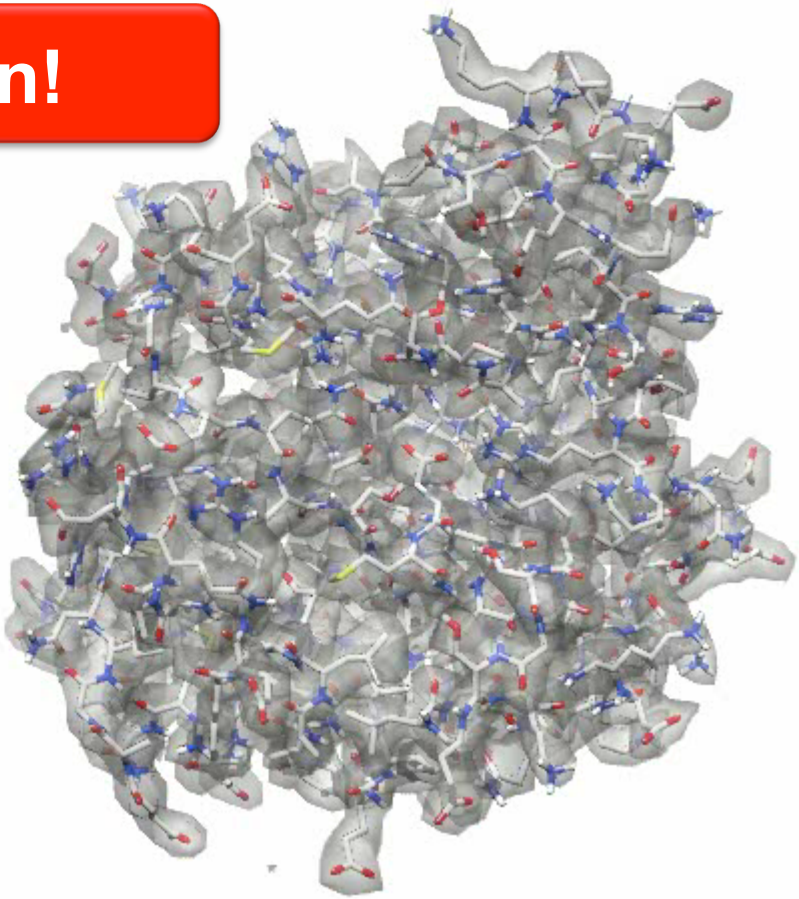
Unknown orientation

Ultra-low signal

E. Coli Adenylate Kinase

Reconstructed Electron Density

Simulation!

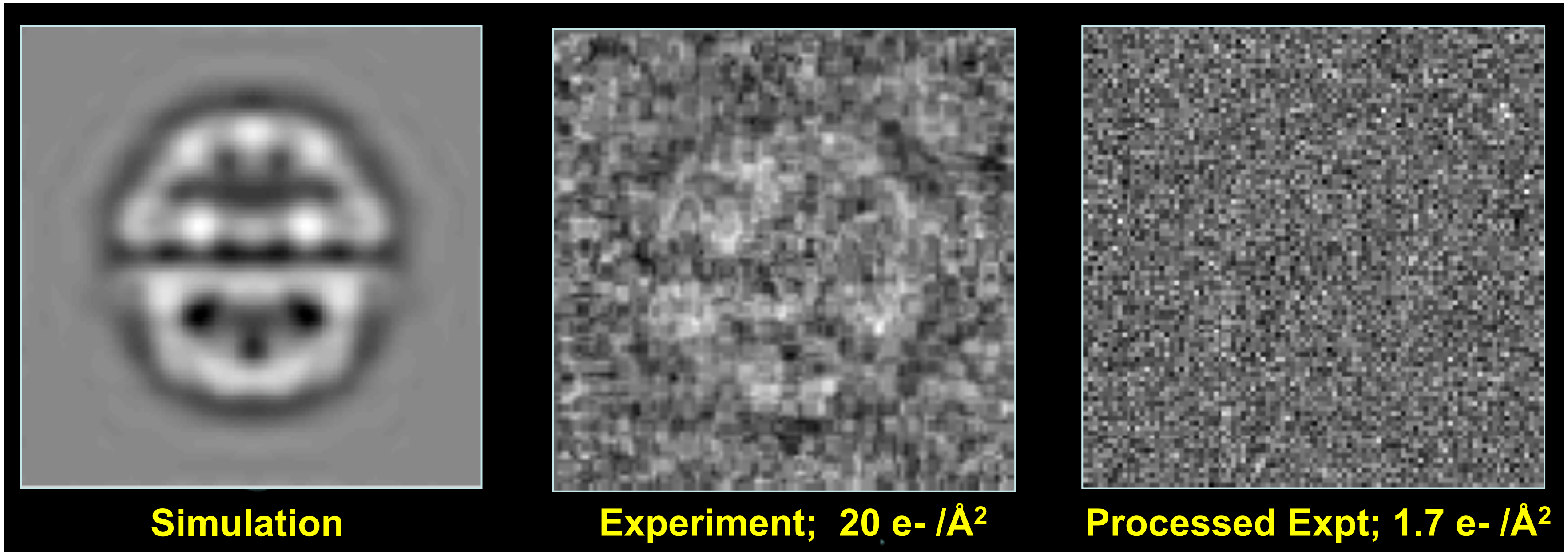


2 million diffraction snapshots

10^4x higher complexity than previously possible

Chaperonin Molecule

From Cryo-EM Snapshots


Simulation
Experiment; 20 e- /Å^2
Processed Expt; 1.7 e- /Å^2

3D reconstruction of single molecule at 12x lower dose than currently needed

Experiment!

Why Better?

- Reconstruct your head with snapshots from random angles
  - In a dark room, with flu

- Standard methods sort snapshots into bins & average
  - “Full frontals, not sneezing”
  - “Full frontal, sneezing a little”
  - “From the side, blowing nose” etc.
  - Lots of classes, each with few shots

- BUT picture from back of head has info about full-frontal
  - Shows where your ears are!

- Manifold techniques exploit this “mutual information”
  - Thrown away by classification methods

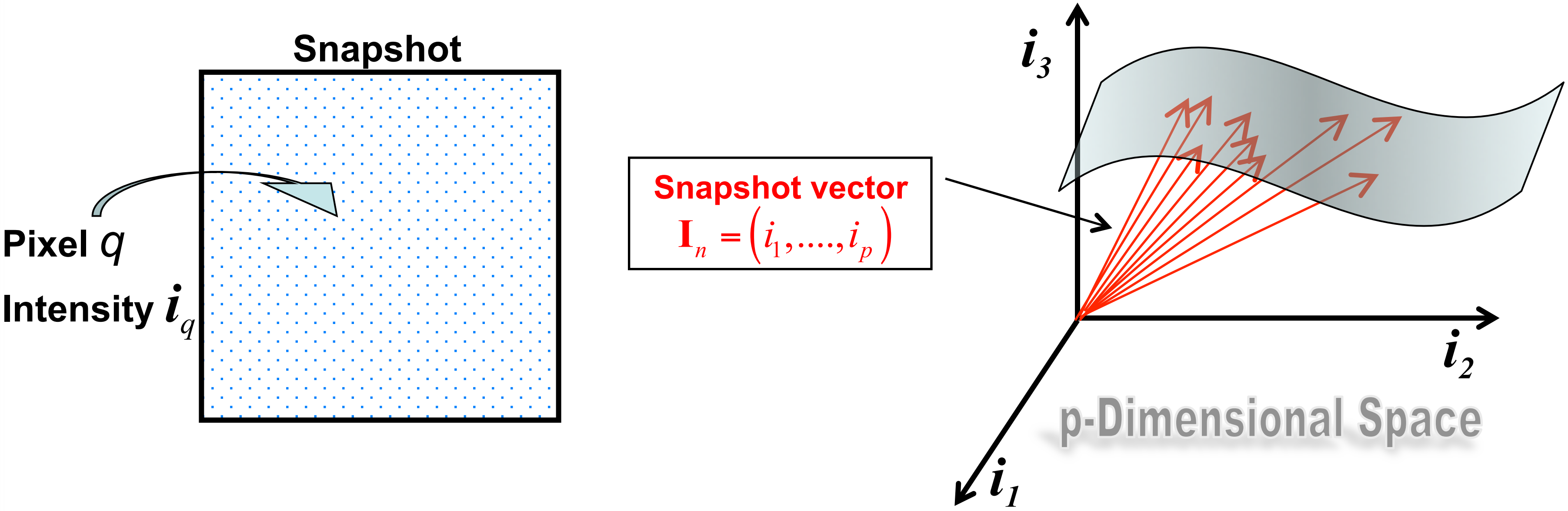
Ensemble of Non-identical Objects

Different Species and/or Distinct Conformations

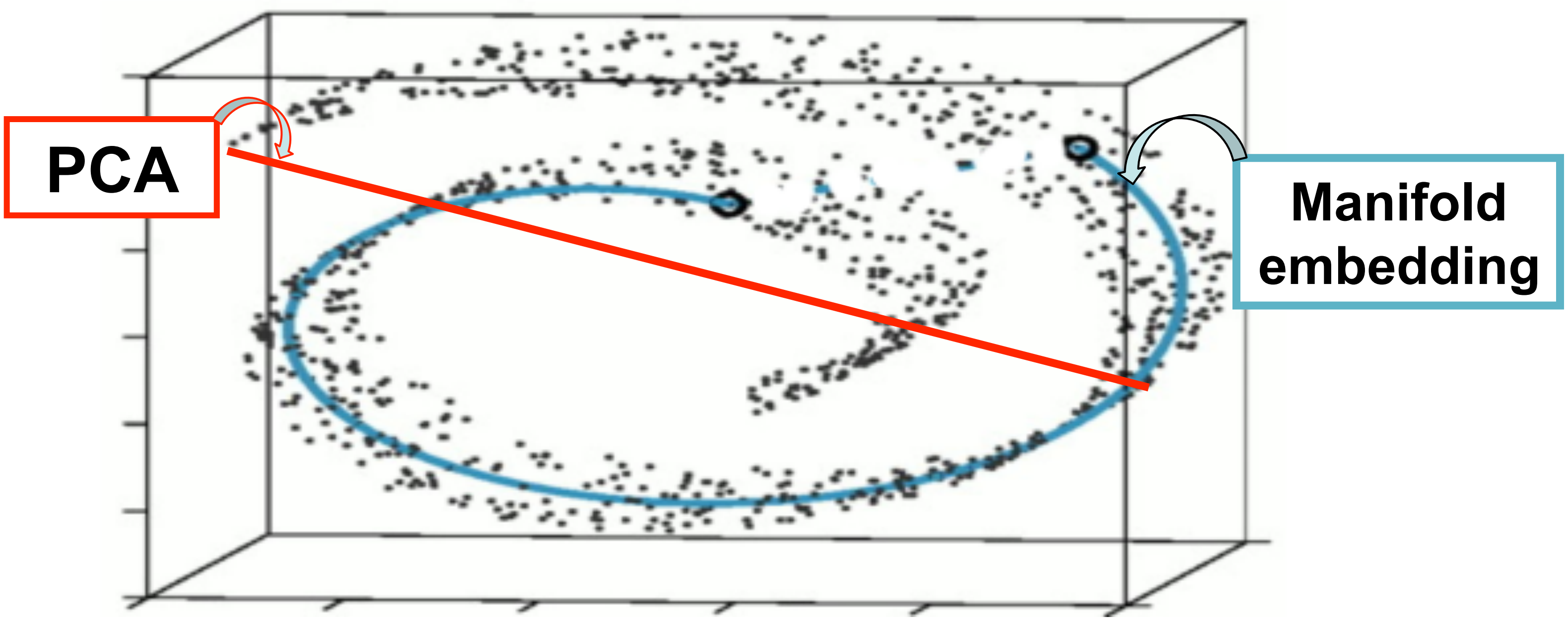
“Post facto Purification”

Correlations as Manifolds

- All we have is ensemble of diffracted intensities
  - A snapshot is $\mathbf{I}_n = \left(i_1, \ldots, i_p\right)$
  - Lives in p-dimensional "intensity space"
  - Correlations produce hypersurface (manifold)

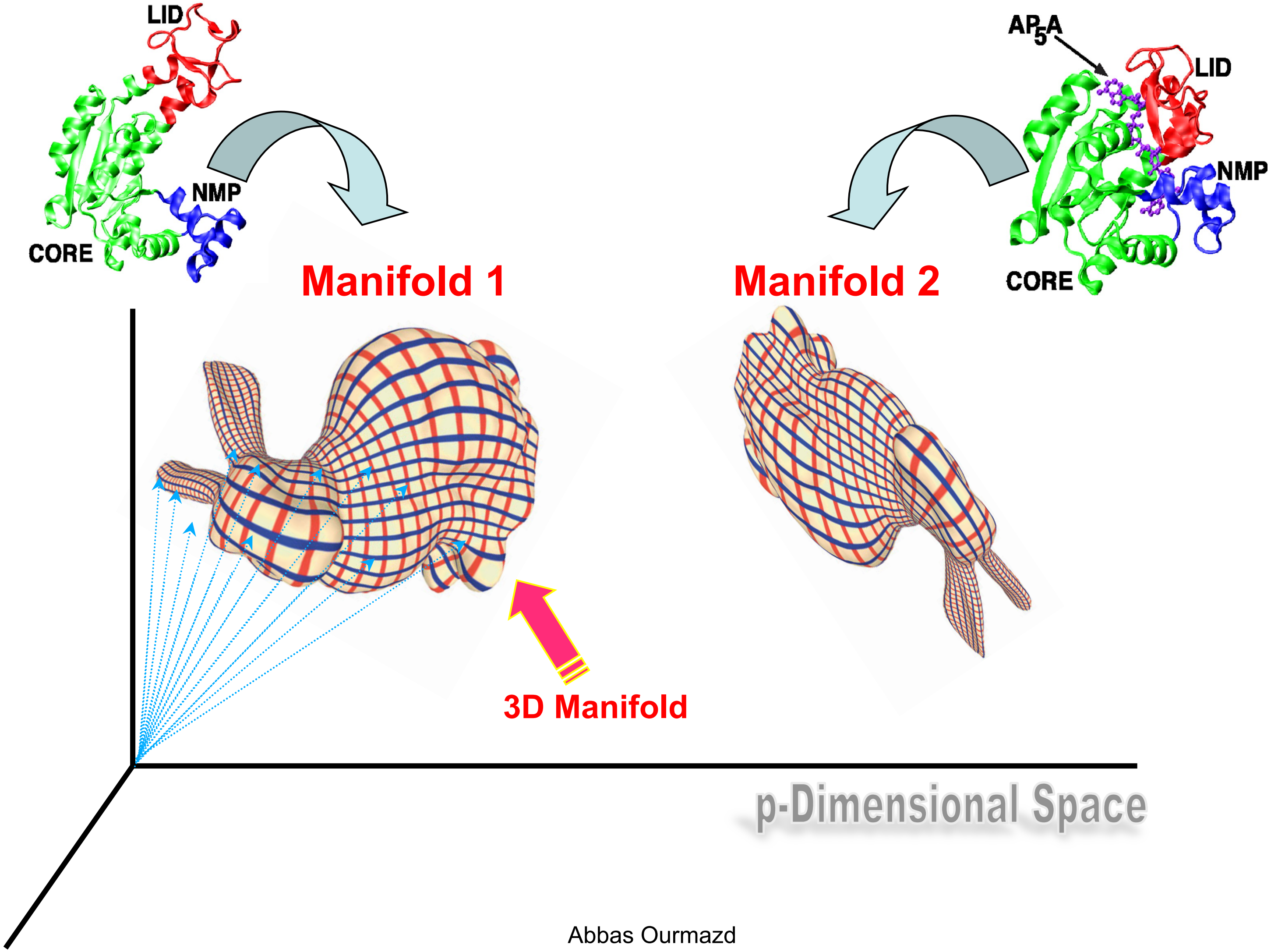

Snapshot
Pixel q
Intensity i_q
Snapshot vector
$\mathbf{I}_n = \left(i_1, \ldots, i_p\right)$
i_3
i_2
i_1
p-Dimensional Space

Manifold-embedding vs. PCA



- Reveals intrinsic, nonlinear structure of data
  - Inaccessible to PCA
  - Varieties: Isomap, Diffusion Map, Riemannian embedding, etc.
  - Flavors: Isotropic kernel, anisotropic kernel

Discrete Conformations


LID
NMP
CORE
AP_5A
LID
NMP
CORE

Manifold 1

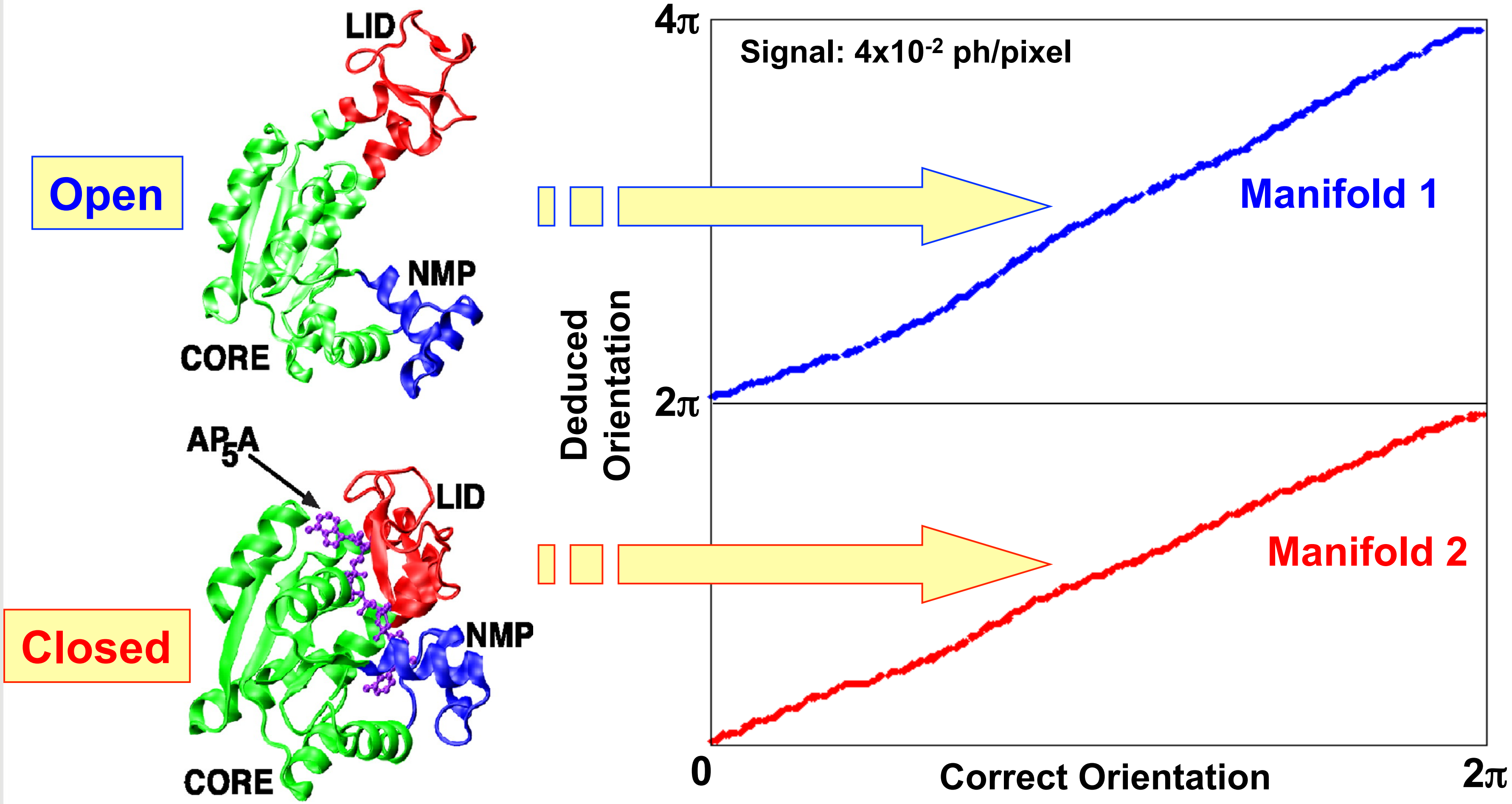
Manifold 2

3D Manifold

p-Dimensional Space

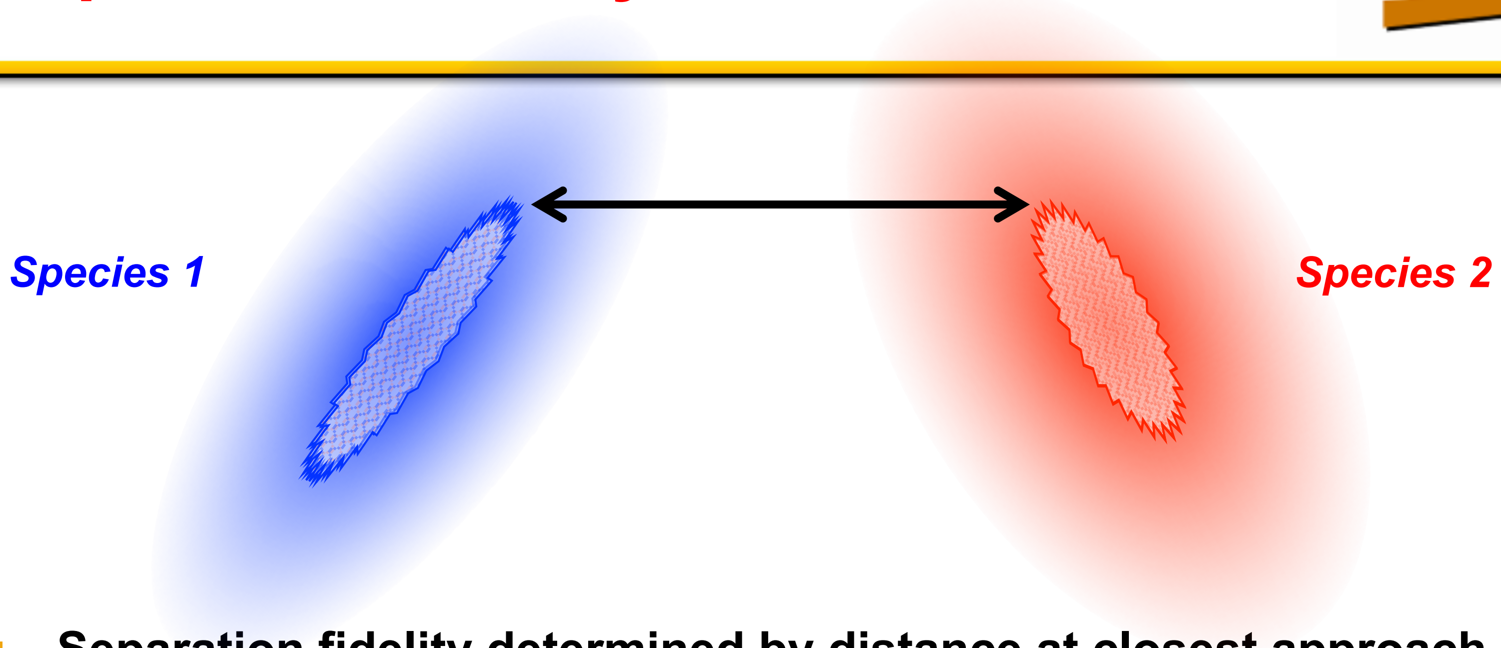
Ensemble of Two Conformations

ADK Closed and Open



Simulation!

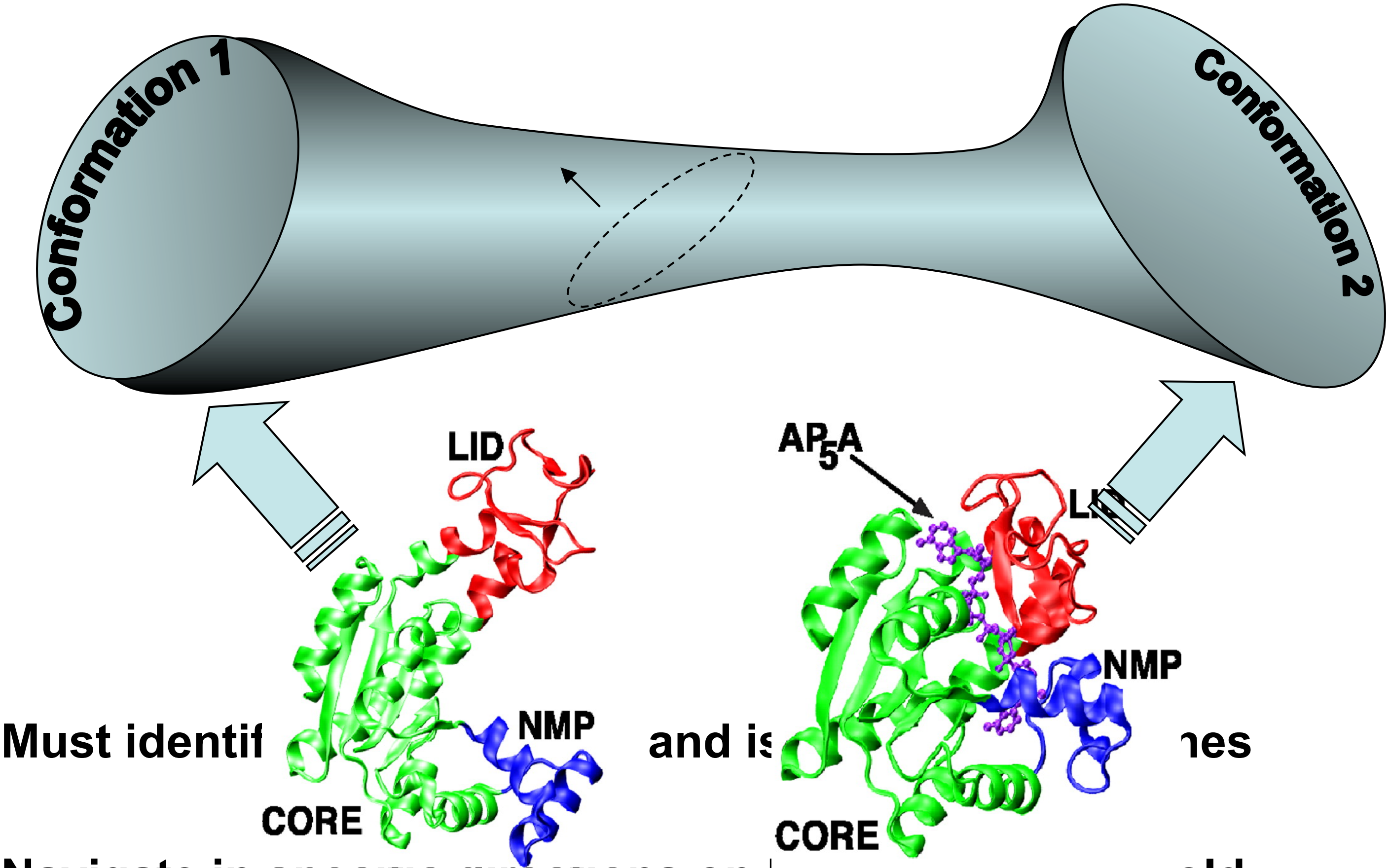
Separation Fidelity



- **Separation fidelity determined by distance at closest approach**
    - **In standard deviations σ ("width of data cloud")**

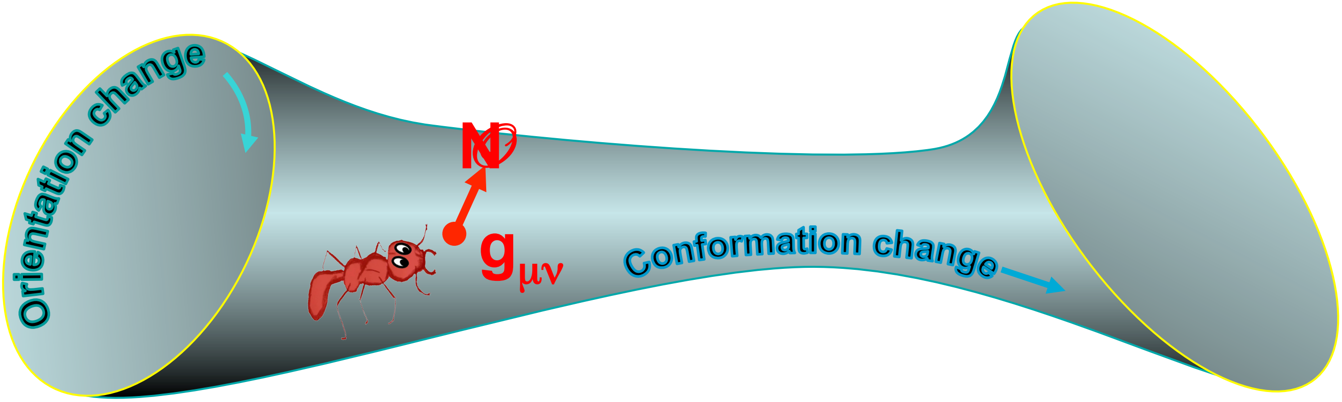
XFEL data	*Separation*	*Remarks*
Simulated	~ 8.5σ	0.04 photons/pixel
Experiment	~ 2σ	Viruses, nano-rice (Optics Exp. Aug 2011)

Mapping Conformations


Conformation 1
Conformation 2
LID
NMP
CORE
AP5A
LI
NMP
CORE

- Must identif and is nes
- Navigate in specific directions on Riemannian manifold

Navigating the Manifold


Orientation change
N
$g_{\mu\nu}$
Conformation change

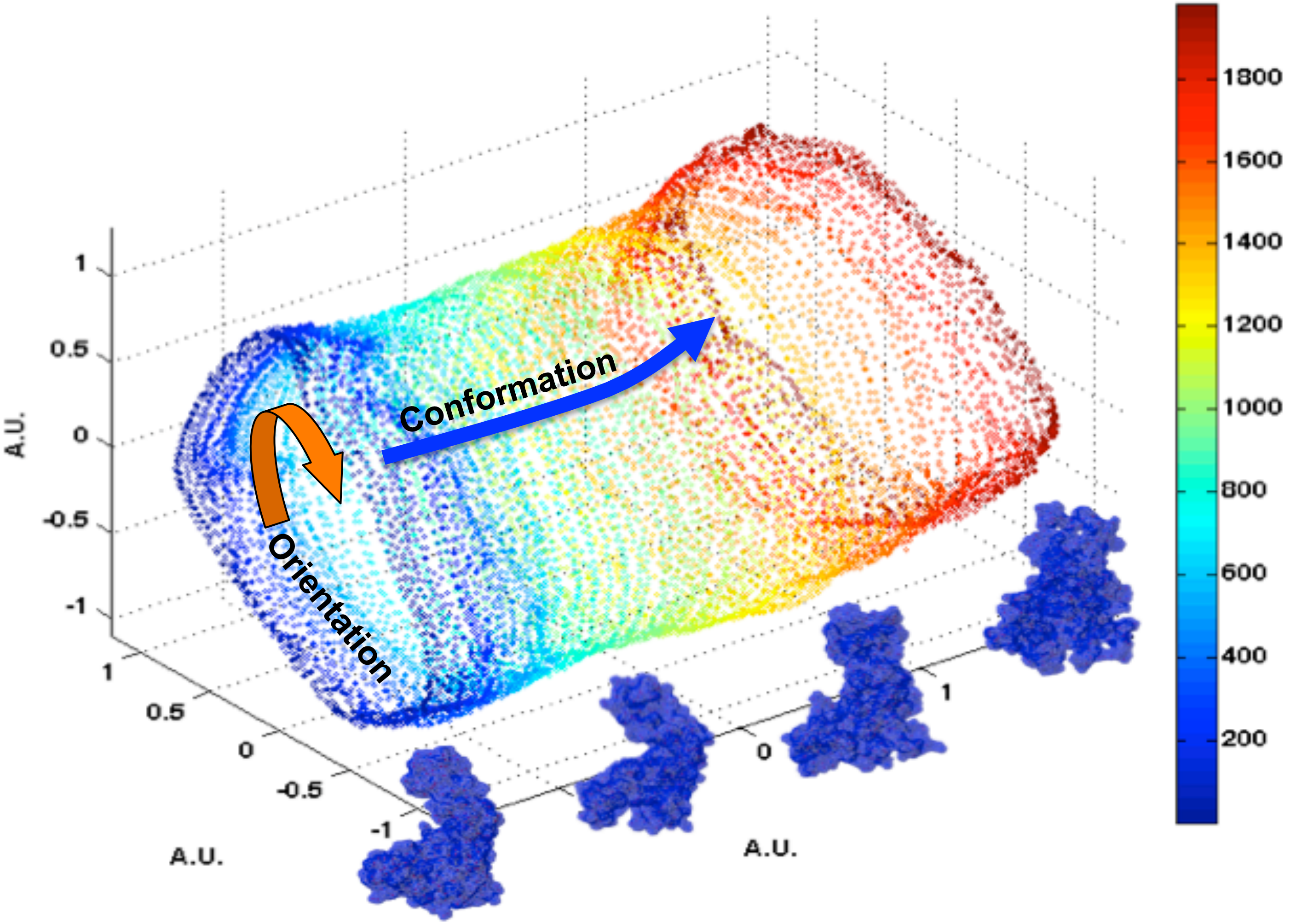
- I am an ant, crawling on the manifold
    - Riemannian approach: measure local rates of change

- Hitherto undiscovered symmetries provide compass
    - Symmetries of OPERATIONS in space, not objects!
    - Navigation = Perception!

- Applies to ALL manifolds produced by scattering
    - Giannakis, Schwander, Yoon, Ourmazd, http://arxiv.org/abs/1009.5035

Conformational Continua

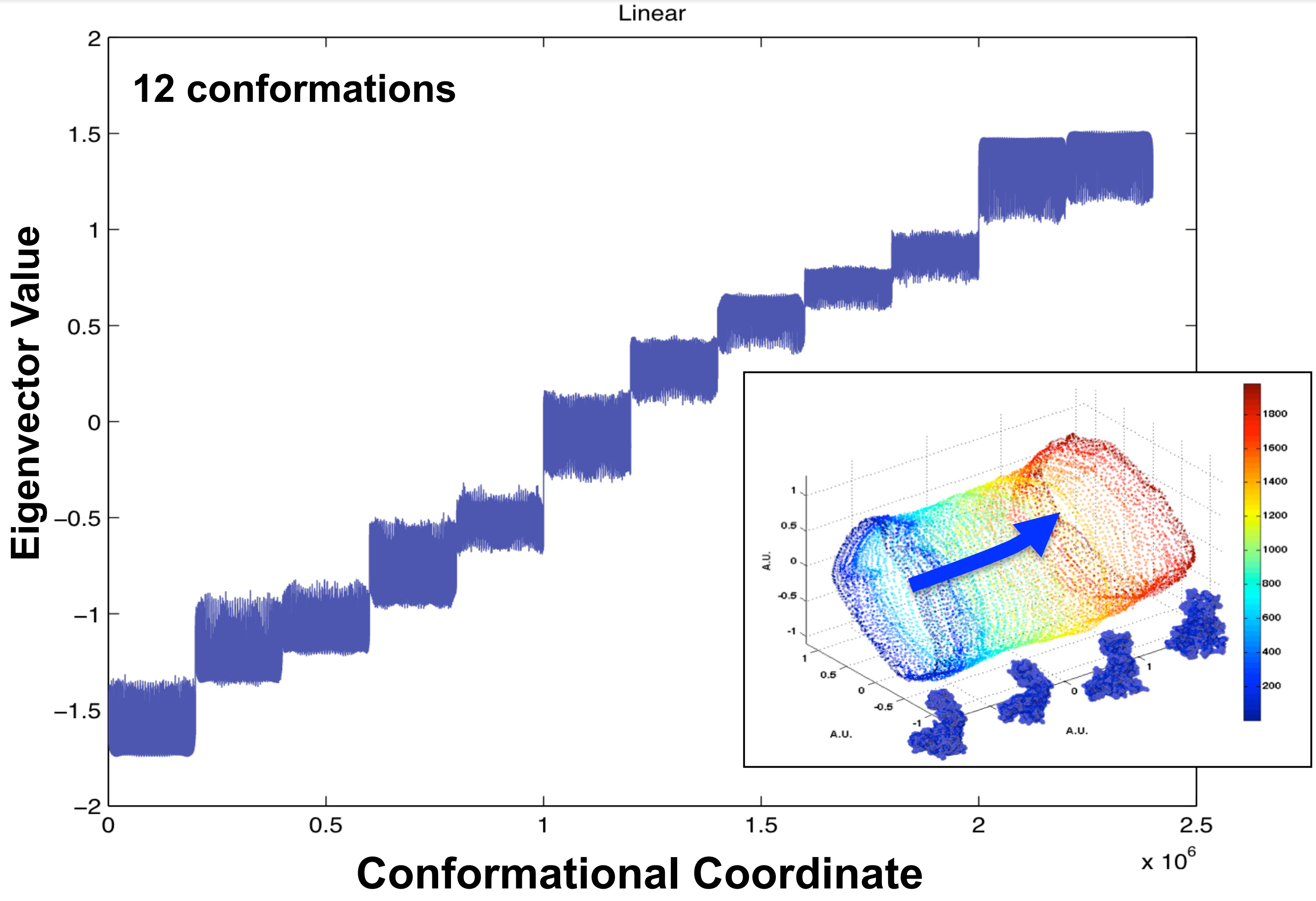
Unfolding of ADK: 850K, 5ns



Manifold formed by diffraction snapshots of melting ADK molecule

Conformational Continua

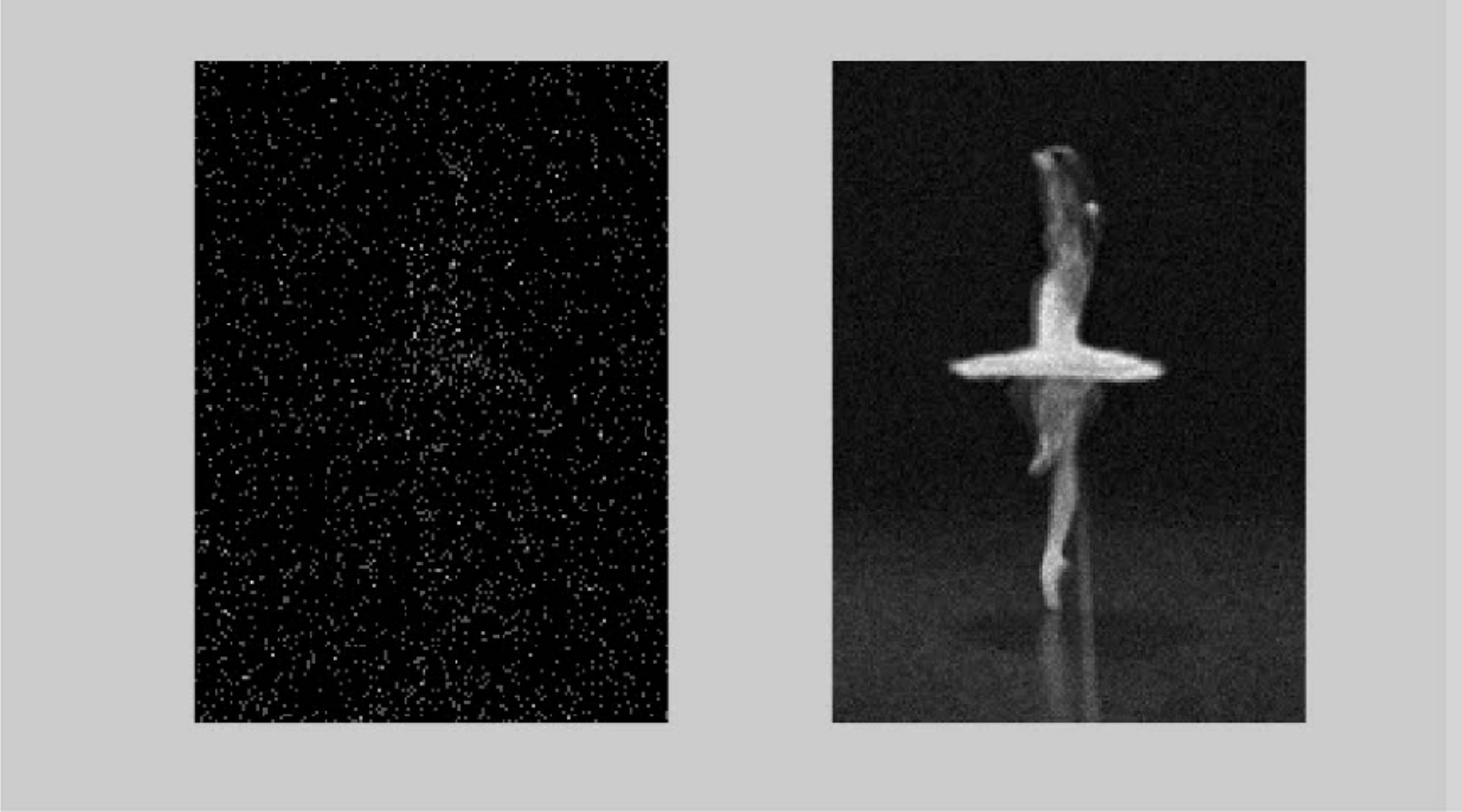
Melting ADK (Simulation)



Simulation; Work in progress!

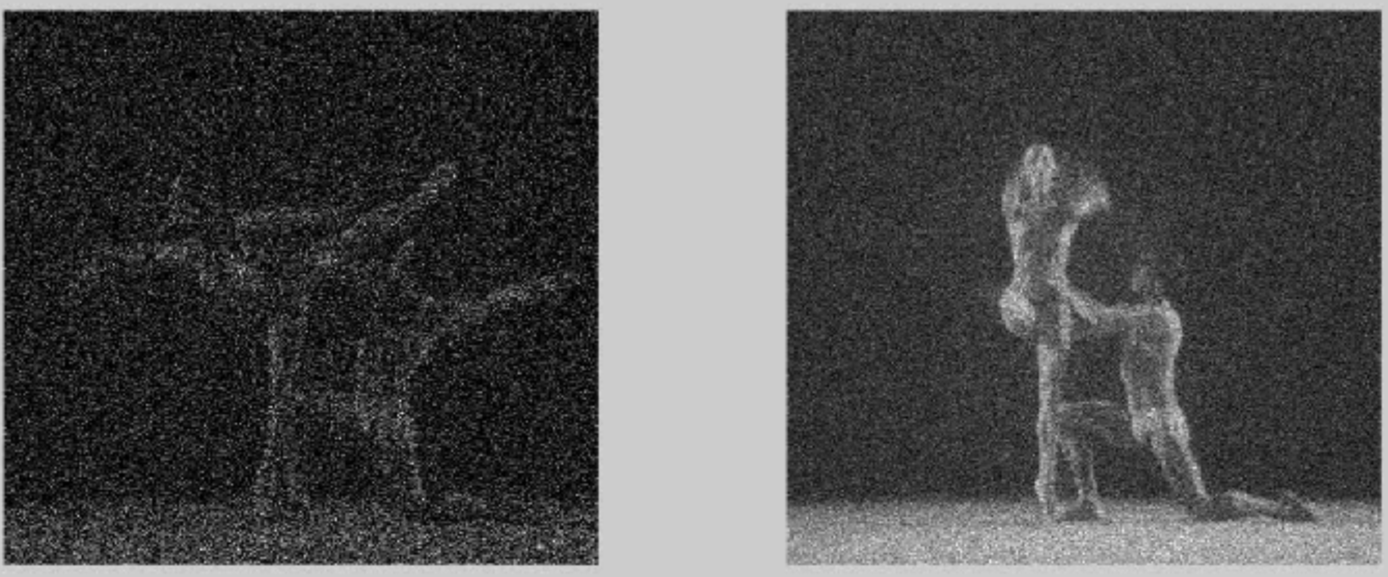
Time Evolution

Rotating Object, SNR: -21dB



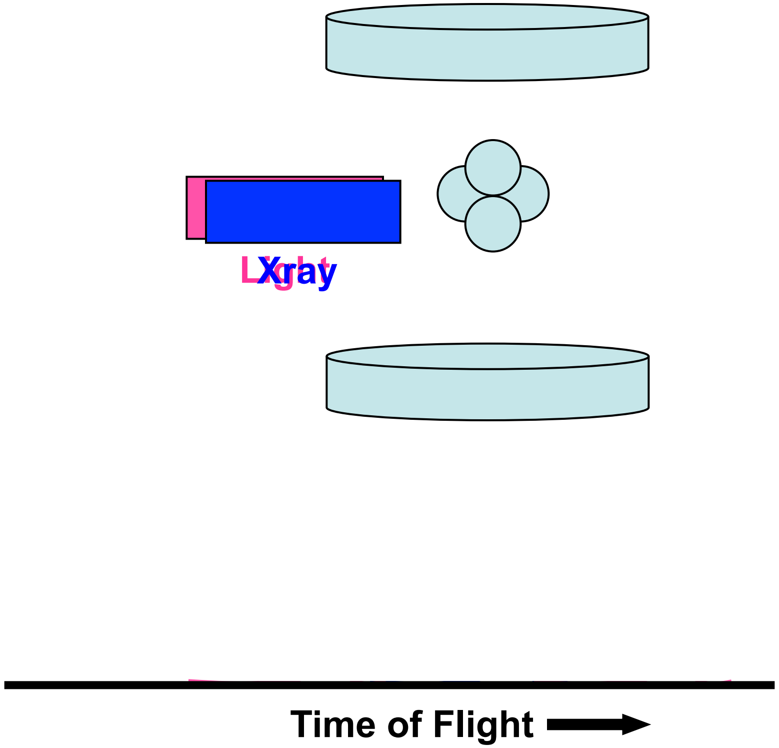
Time Evolution

Interacting Bodies, SNR:-11dB



Pump-probe at LCLS

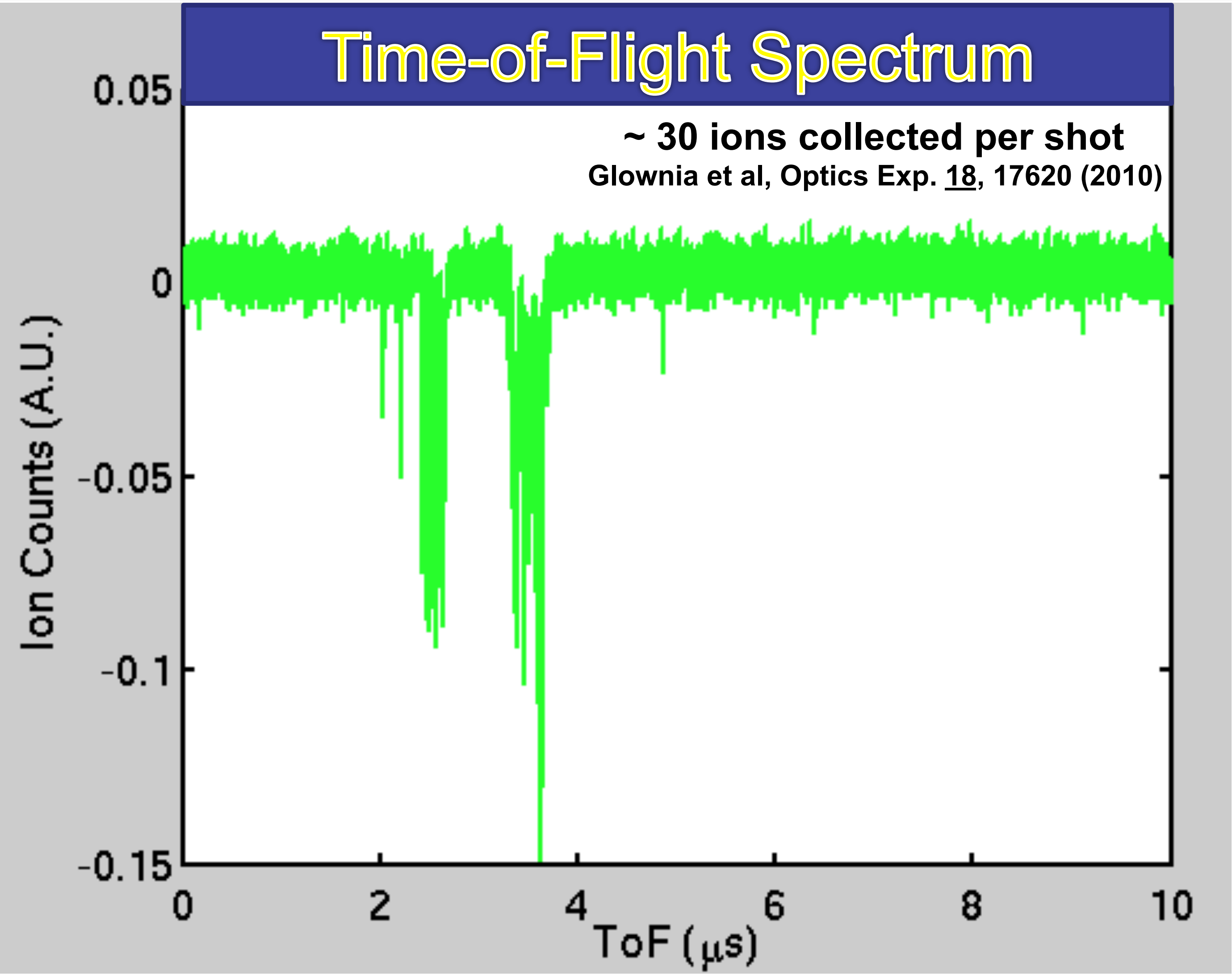
N_2 Coulomb Explosion


Light
Xray
Time of Flight

Timing jitter: 280fs FWHM

Glownia et al, Optics Exp. 18, 17620 (2010)

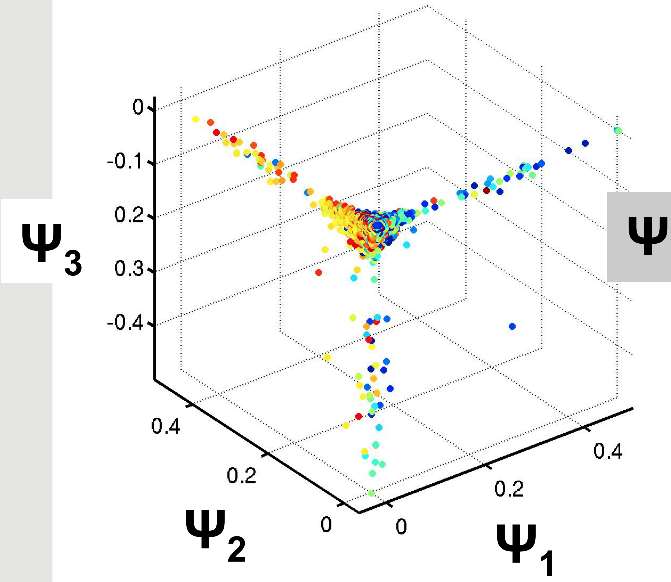
N_2 Coulomb Explosion



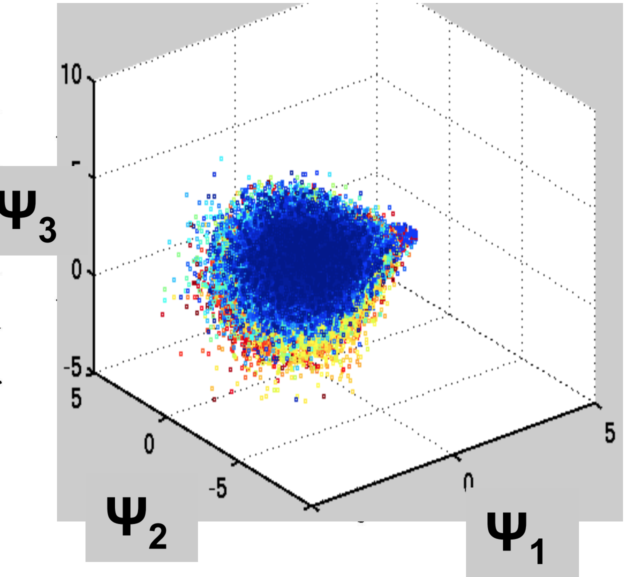
~ 30 ions collected per shot

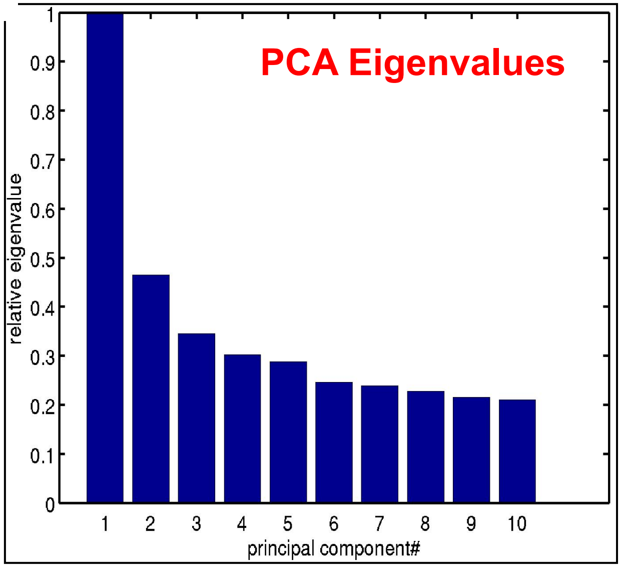
Glownia et al, Optics Exp. 18, 17620 (2010)

N_2 Coulomb Explosion
Manifold Analysis



Principle Comp. Analysis (PCA)



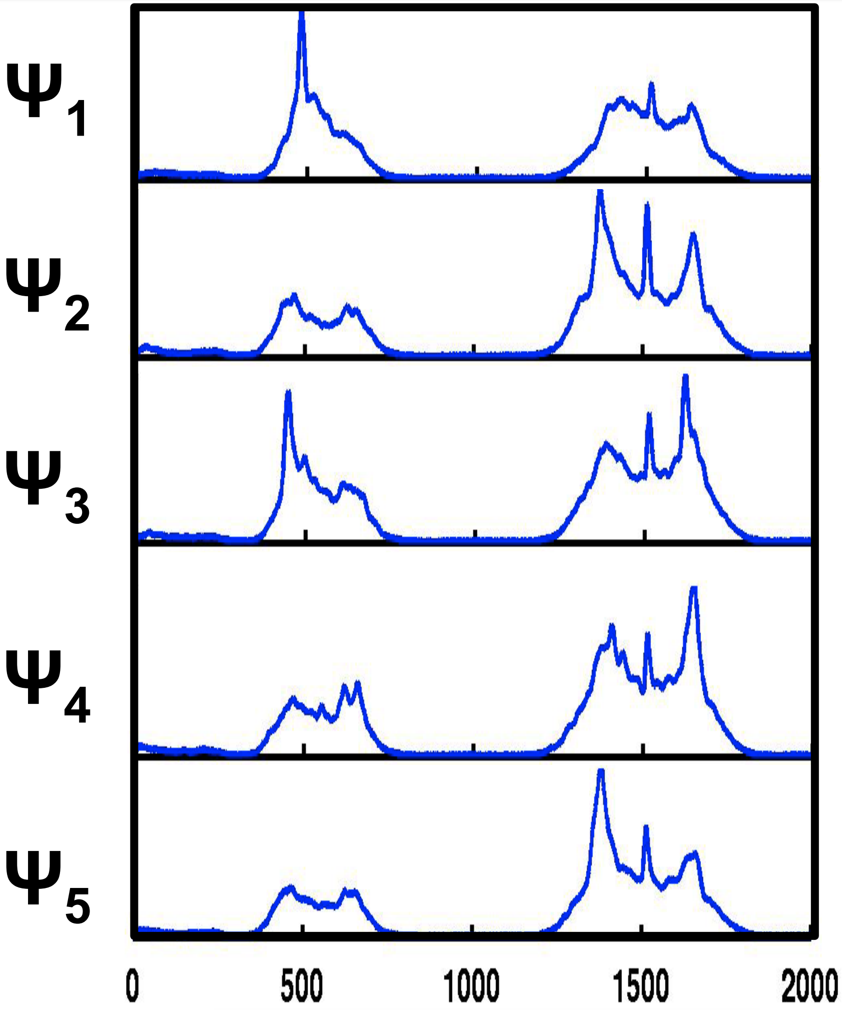


Experiment!

- **FIVE dimensions**
  - **From eigenvalue spectrum, manifold structure**
  - **Standard (linear) PCA produces featureless blob**
- **Each dimension an explosion channel**

N_2 Explosion Channels

ToF Spectra from Manifold



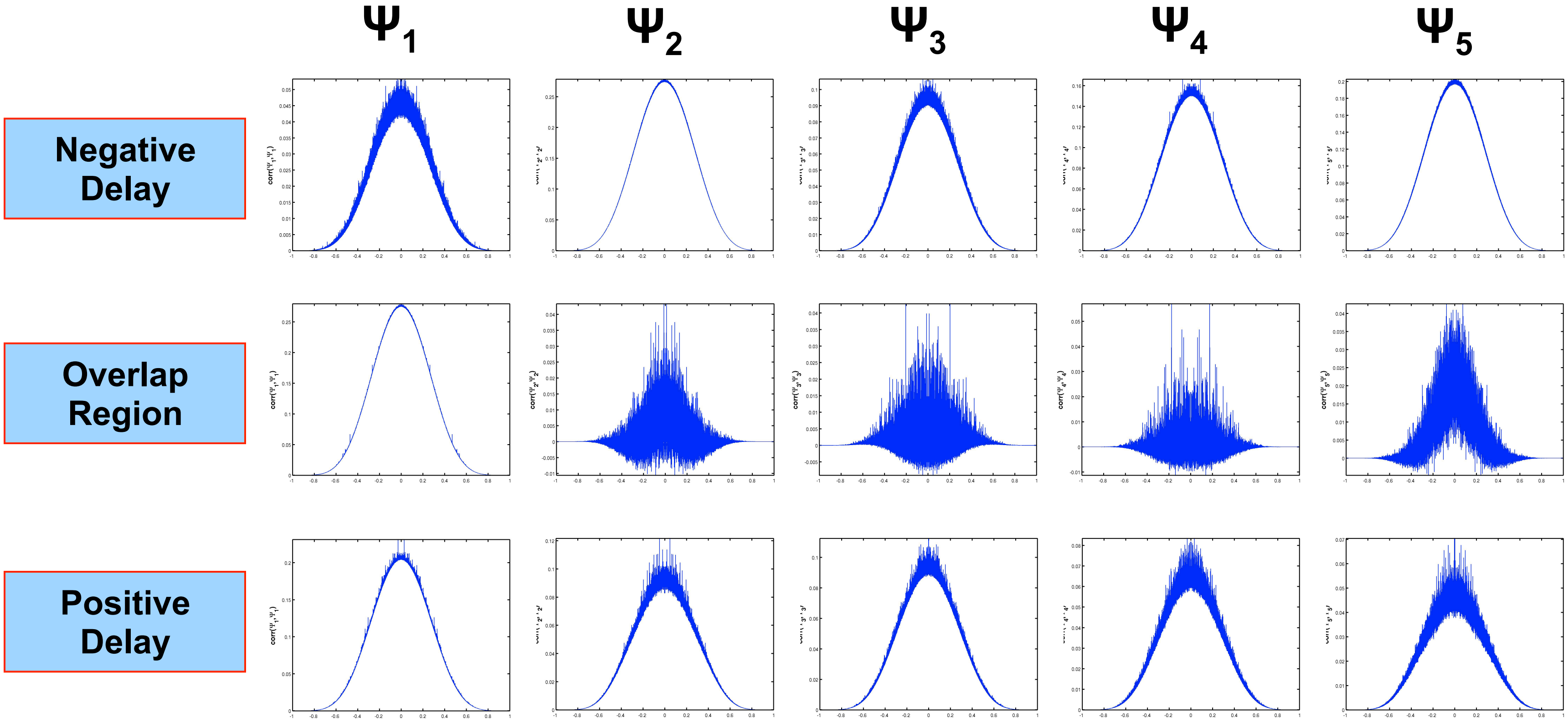
Experiment!

Fine structure NOT noise

Raw, no core, Z3(+), May 31, 11

Autocorrelations

Different Delay Regimes



Central peaks clipped

Pedestal: No correlated activity; Fine structure: Correlation

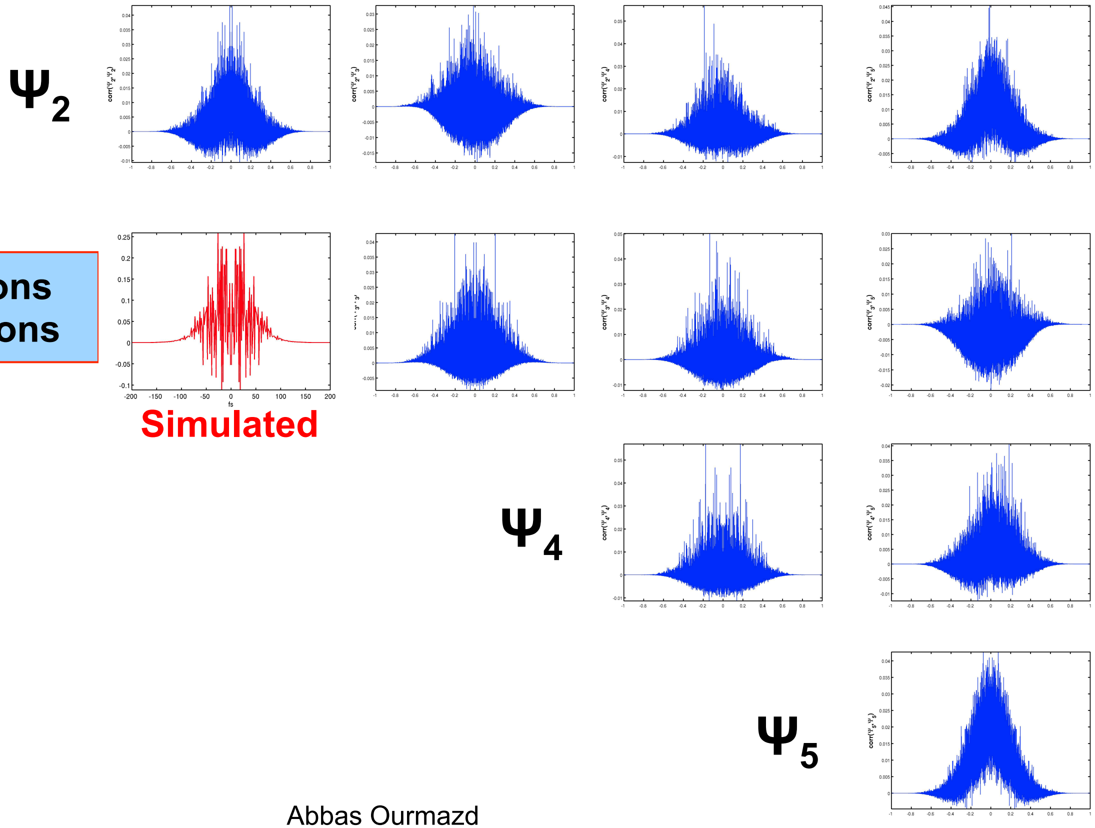
Pronounced fine structure when IR and X-ray pulses overlap

Correlation Matrix

Overlap Region

Femtosecond fine Structure Coherence between channels

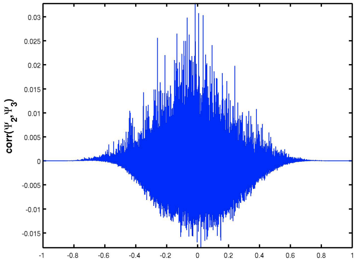
Diagonal: Autocorrelations
Off-diagonal: X-correlations



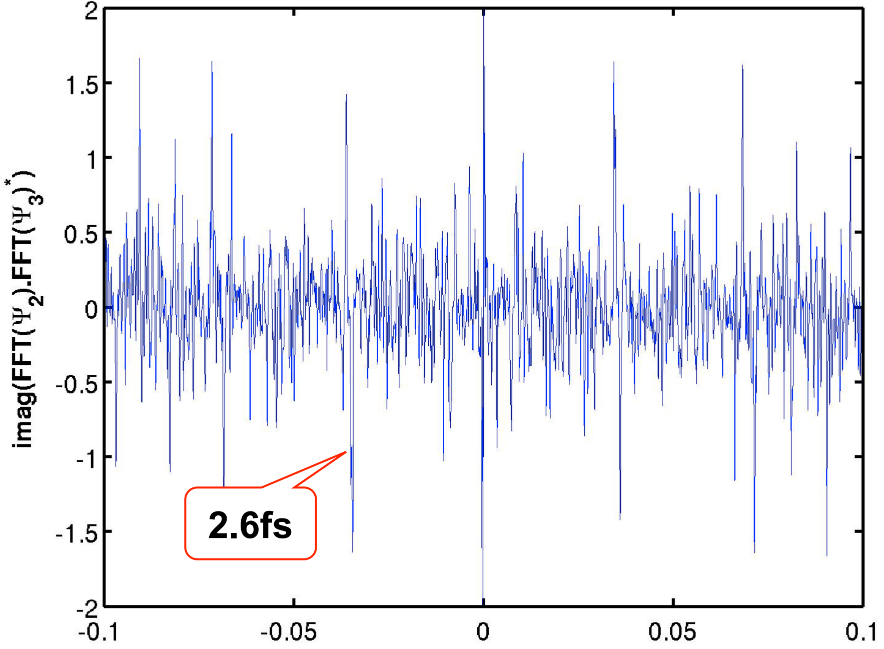
Coherence between Channels

$\Psi_2^* \Psi_3$

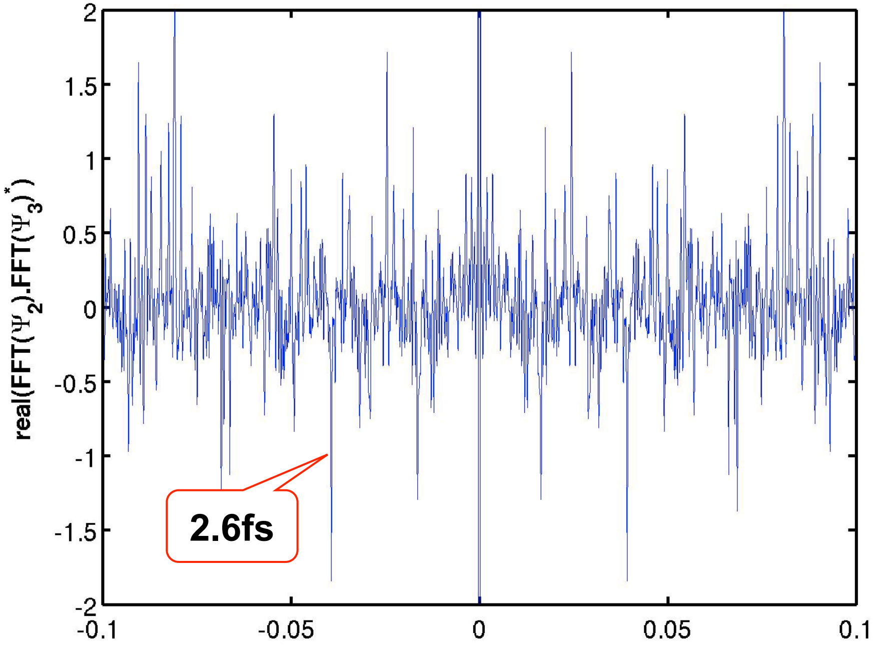
X-correlation



FFT (Imaginary Part)



FFT (Real Part)



Conclusions & Outlook

- Manifolds represent powerful new route to structure & dynamics
  - Capture entire information content, exploit symmetries
  - Remove need for “identical objects”

- Recover structure, separate species, map conformations
  - From ultra-low signal , random snapshots
  - Applications from molecules to people

- Femtosecond dynamics
  - Despite (overwhelming) timing jitter
  - Structure and dynamics of ultrafast reactions (“3D movies”)

- Potential route to new vistas
  - Machine learning & symmetries of perception
  - Platonic Forms

Acknowledgments

Collaborators:

Russell Fung
Dimitris Giannakis
Peter Schwander
Chunhong Yoon
Phil Bucksbaum
Wah Chiu
Ryan Coffee
James Cryan
Jim Glownia
George Phillips
+ 60 others

Financial Support

U.S. DEPARTMENT OF ENERGY Office of Science

DE-SC0002164

National Institutes of Health

The Nation's Medical Research Agency

Protein Structure Initiative GM074901