Outline of Five 1 Hour Lectures

- A Survey of My Work in SDE Modeling and Inference
- Refresher / Introduction to Ito's Formula and Its Applications in Statistical Inference (e.g. Testing a Fundamental Levy Process Assumption)
- Refresher/ Introduction to Quadratic Variation and Extensions to Account for High Frequency "Artifacts" (Market Microstructure)
- Details on Some Modern Goodness of Fit Tests and Open Problems with Illustrative Examples
- Highlight of Some of My Recent Works Combining Elements of the Above

Outline of Five 1 Hour Lectures

- A Survey of My Work in SDE Modeling and Inference
- Refresher / Introduction to Ito's Formula and Its Applications in Statistical Inference (e.g. Testing a Fundamental Levy Process Assumption)
- Refresher/ Introduction to Quadratic Variation and Extensions to Account for High Frequency "Artifacts" (Market Microstructure)
- Details on Some Modern Goodness of Fit Tests and Open Problems with Illustrative Examples
- Highlight of Some of My Recent Works Combining Elements of the Above

Outline of Five 1 Hour Lectures

In All Lectures, I Attempt to Provide "Something for Everyone" [from Beginning Graduate Student to Researchers in Time Series or SDEs]

However, the Middle 3 Lectures Are Targeted More at Fundamental Material [I Present Established Works of Others … Skewed, Of Course, By My Personal Interests]

Lecture I: Motivation and Some Overlapping Interests in Finance and Chemical Physics

Christopher P. Calderon

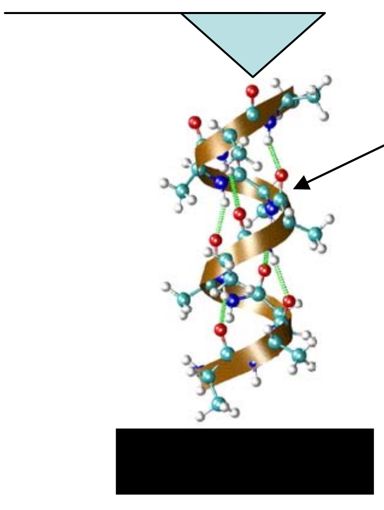
Rice University / Numerica Corporation

Research Scientist

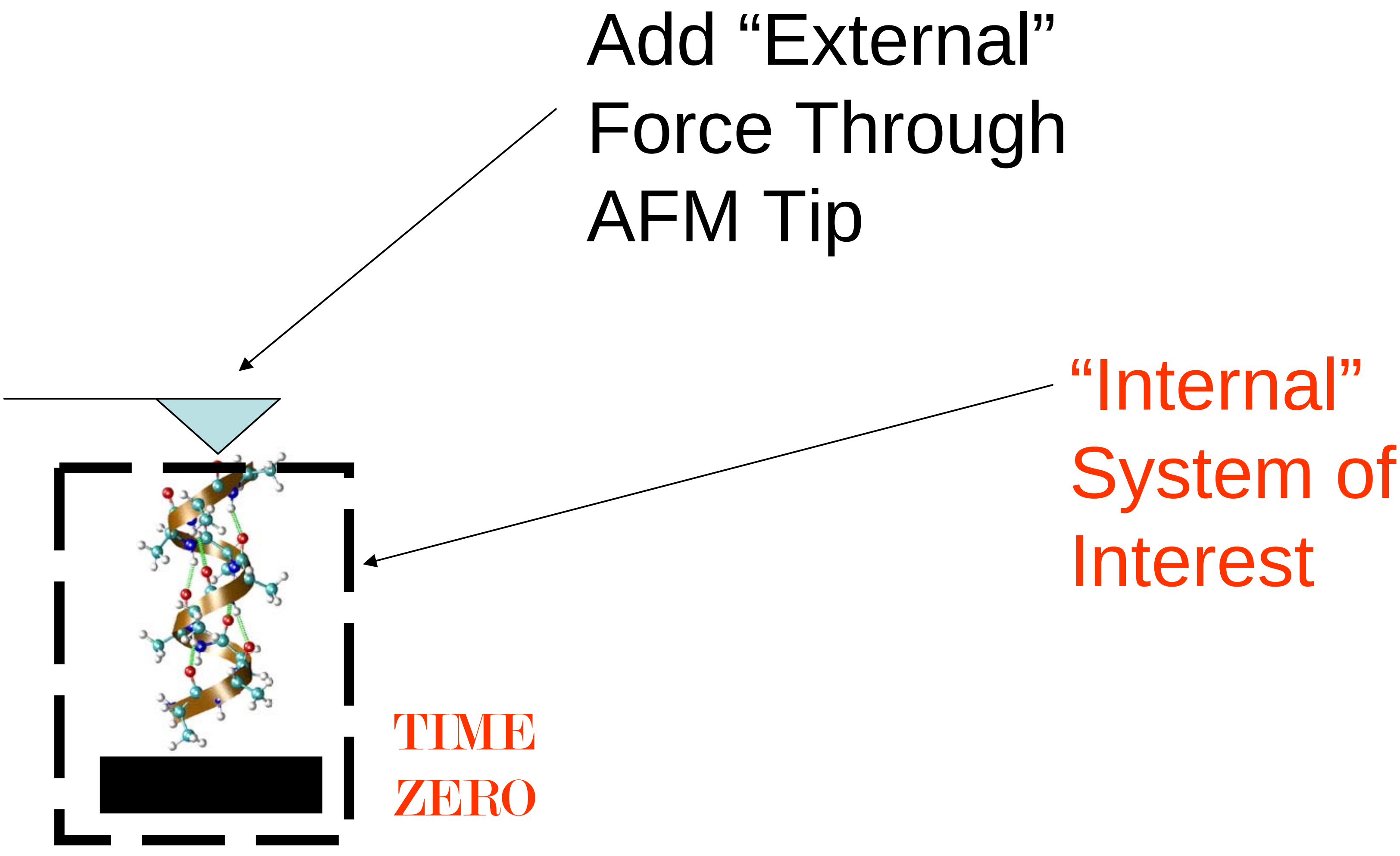
Single-Molecule Experiments

Atomic **F**orce **M**icroscope (AFM) Tip

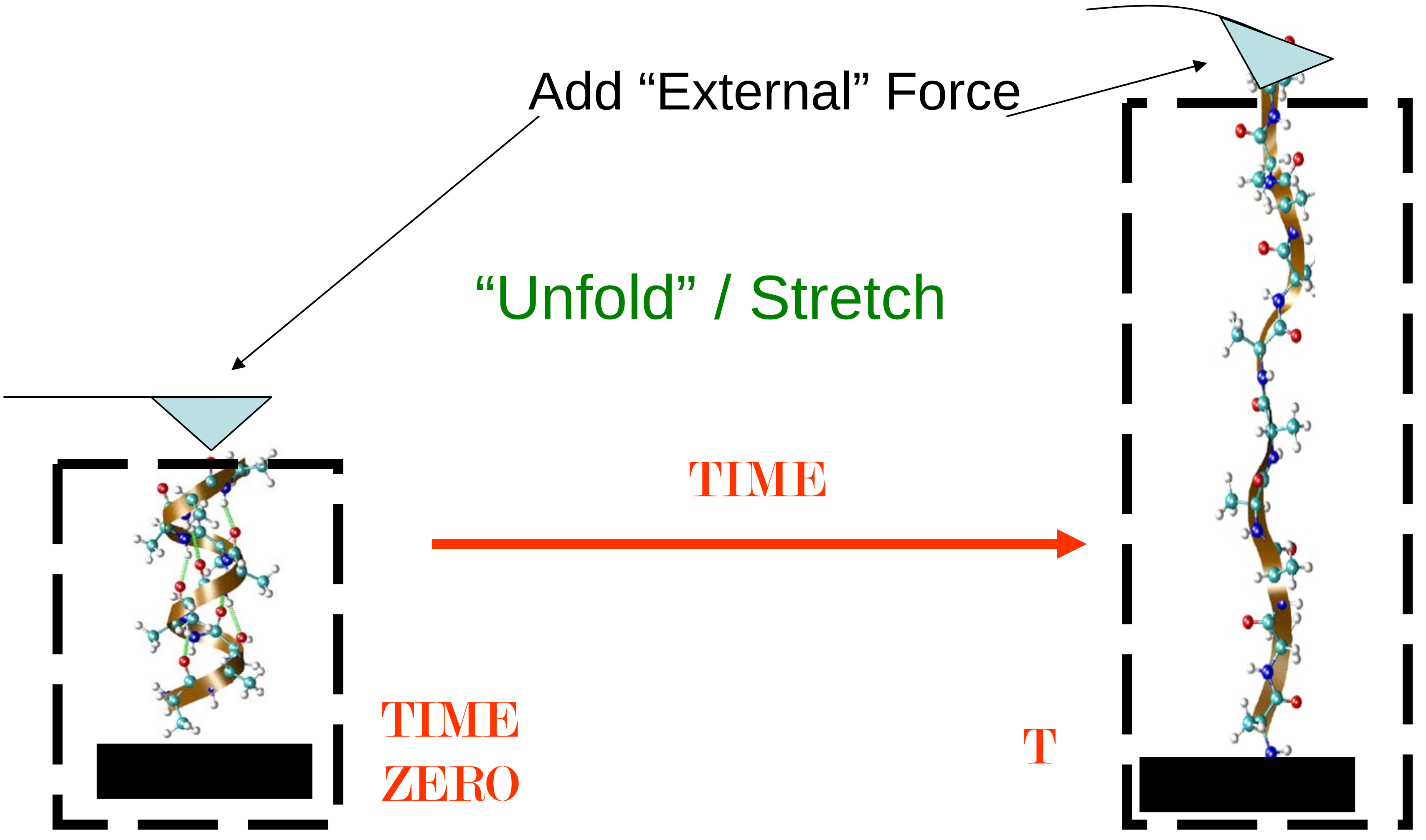
Single Biomolecule (Protein or DNA)



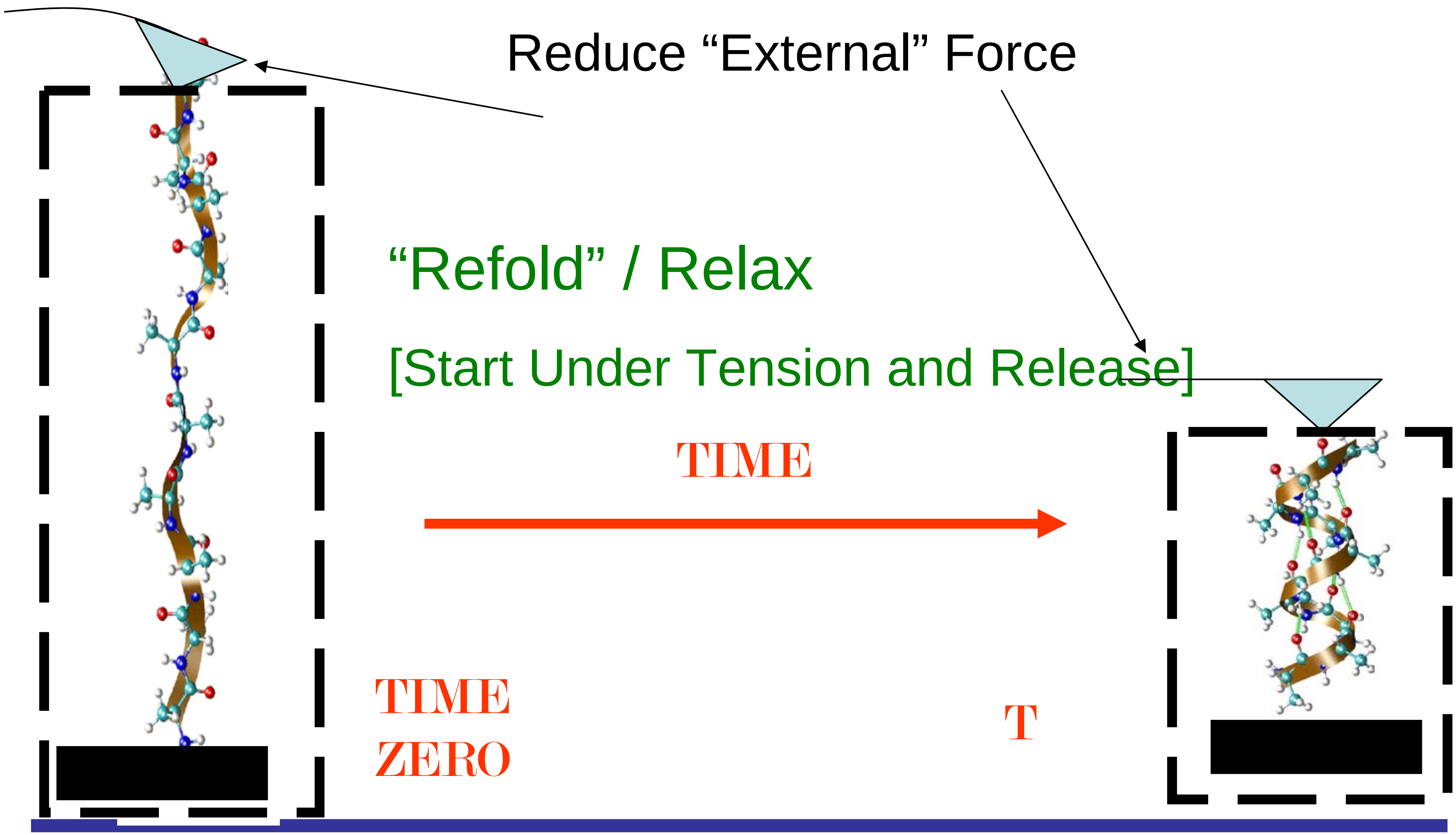
Mechanical Manipulation of Single-Molecules



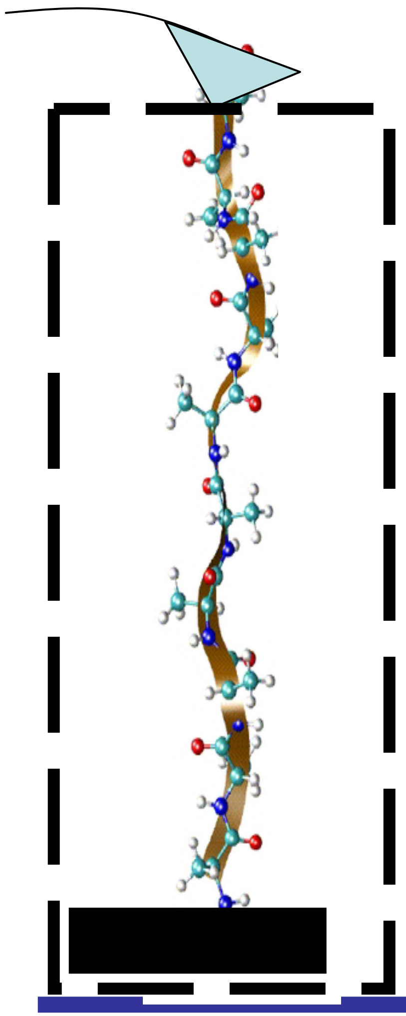
Mechanical Manipulation of Single-Molecules



Mechanical Manipulation of Single-Molecules

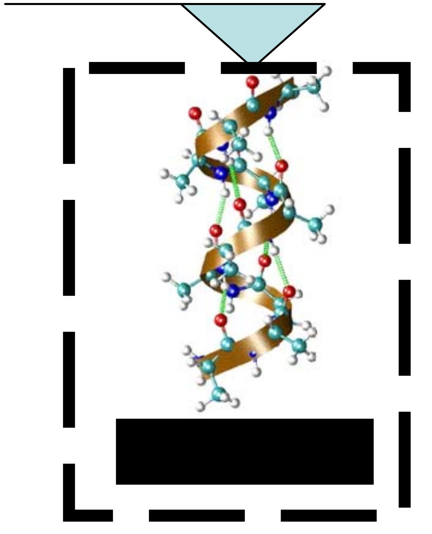


Mechanical Manipulation of Single-Molecules

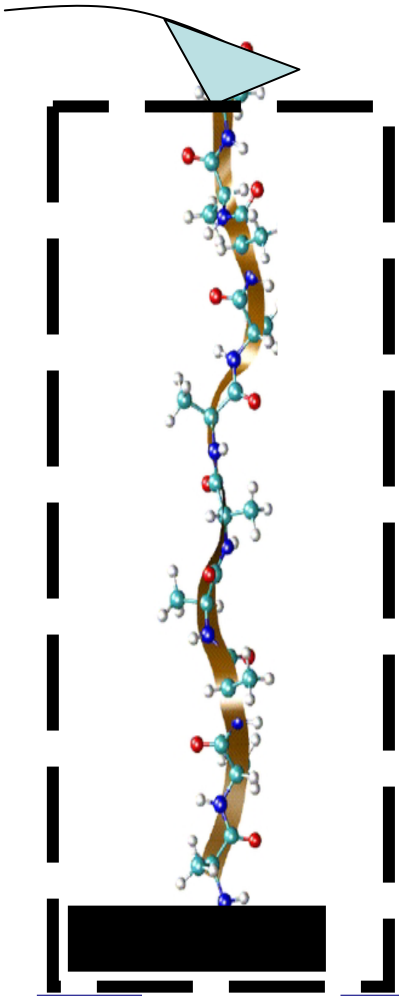


Experiments Allow for New Types of Kinetic Insights Into Complex Biological Processes

(Protein Folding, DNA Transcription, etc.)

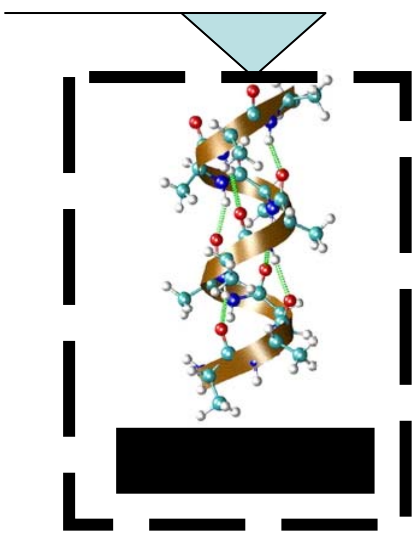


Mechanical Manipulation of Single-Molecules

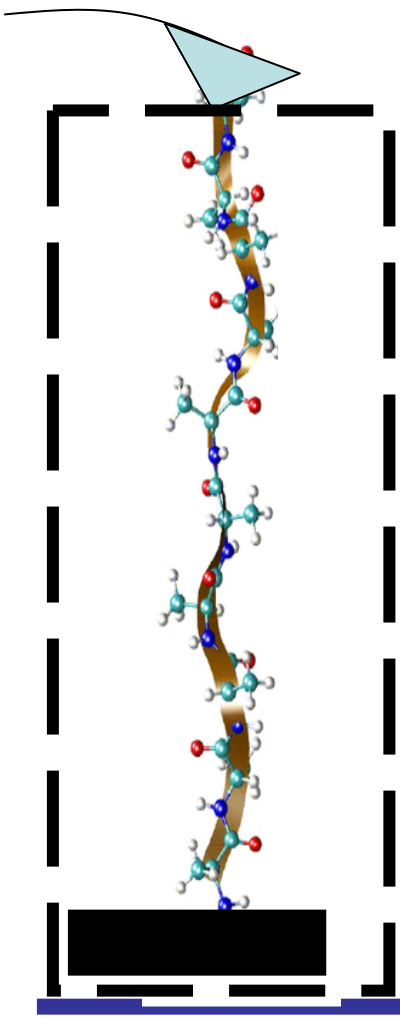


Experiments Allow for New Types of Kinetic Insights Into Complex Biological Processes
(Protein Folding, DNA Transcription, etc.)

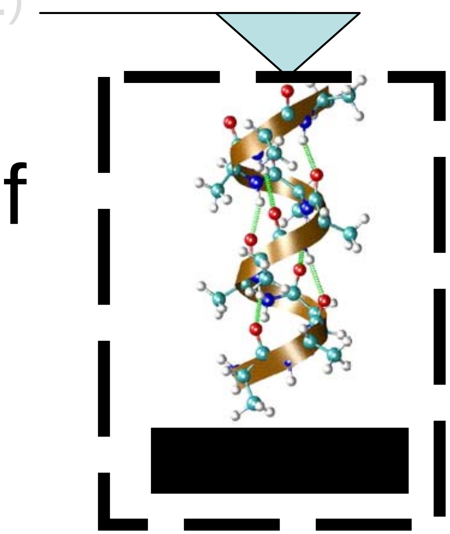
Though Thermal Fluctuations Strongly Influence the Dynamics at Single-Molecule Time/Length Scales



Mechanical Manipulation of Single-Molecules



Experiments Allow for New Types of Kinetic Insights Into Complex Biological Processes
(Protein Folding, DNA Transcription, etc.)



Physical Scientists are Realizing Importance of Stochastic Modeling Beyond Population Models

Assumption on Governing Dynamics

IF given:

1) Position (q)

2) Momentum (p)

of all N atoms and

3) Accurate Force Field (not currently a practical assumption)

Any physical property can be accurately predicted at desired resolution

$$q, p \in \mathbb{R}^{3N}$$

and

$$\mathcal{H}(\Gamma, t) \in \mathbb{R}$$

where N is the number of atoms and

$$\Gamma \equiv (p, q)$$

Assumption: Classical Statistical Mechanics Adequately Describes the Dynamics

$$\frac{dp}{dt} = -\frac{\partial \mathcal{H}}{\partial q}$$

$$\frac{dq}{dt} = \frac{p}{m} \equiv \frac{\partial \mathcal{H}}{\partial p}$$

Deterministic System
(Chaotic ODEs, e.g.
Nose-Hoover Dynamics)

$$p, q \in \mathbb{R}^{3N}$$

and

$$\mathcal{H}(\Gamma, t) \in \mathbb{R}$$

where N is the number of atoms and

$$\Gamma \equiv (p, q)$$

Assumption: Classical Statistical Mechanics Adequately Describes the Dynamics

$$\frac{dp}{dt} = -\frac{\partial \mathcal{H}}{\partial q}$$

$$\frac{dq}{dt} = \frac{p}{m} \equiv \frac{\partial \mathcal{H}}{\partial p}$$

MANY Timescales.
Numerically Difficult to Solve
Due to Constraints Imposed by
Numerical Stability

$$p, q \in \mathbb{R}^{3N}$$

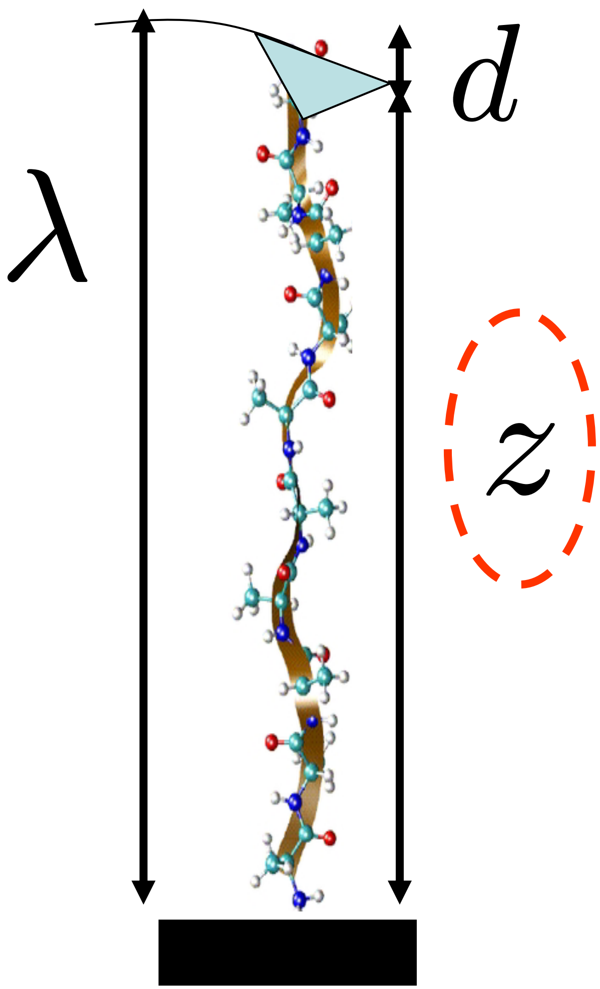
and

$$\mathcal{H}(\Gamma, t) \in \mathbb{R}$$

where N is the number of atoms and

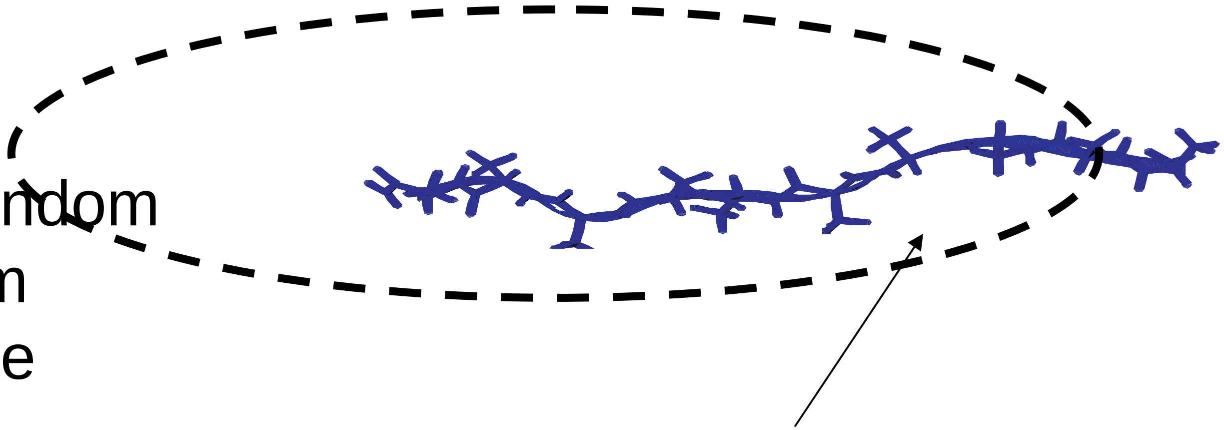
$$\Gamma \equiv (p, q)$$

Treating Limited Resolution



- "Z" is Technically a Function of Position and Momentum of All "N" Atoms. *(Not Yet Experimentally Accessible)*
- If "Z" Really Captures All of the Relevant Physics, How Can One Fully Utilize the Time Series Data?
- If "Z" is Only Part of Story (but Only Quantity Experimentally Accessible), How Can One Make Use of Time Series Data?

Illustration Using a Controlled MD Simulation



1) Draw Random Coil From Ensemble

Drawn from Ensemble of Conformations @ time 0

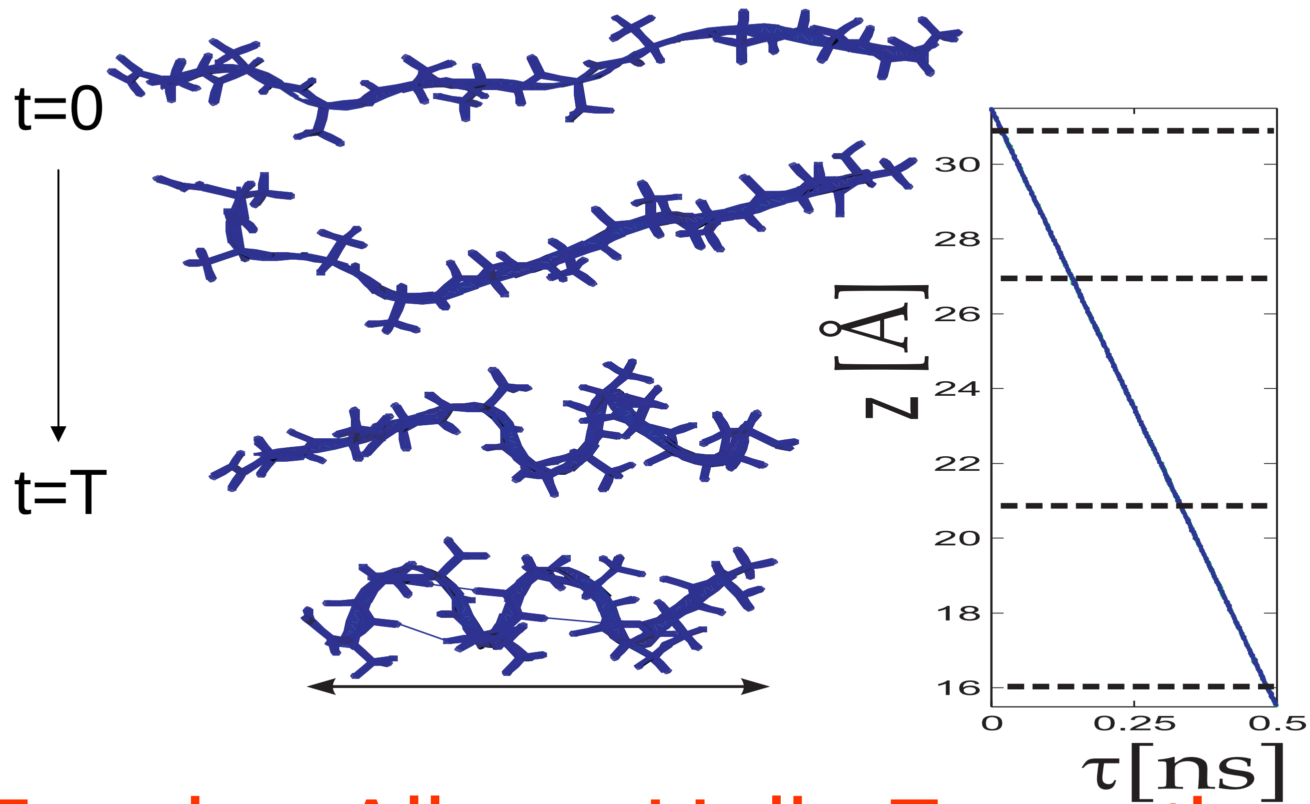
Park & Schulten, *J. Chem. Phys.* **120** (2004).
Calderon, *J. Chem. Phys.* **127** (2007).
Calderon & Chelli, *J. Chem. Phys.* **128** (2008).

Tracking the Snapshots

1) Draw Random Coil From Ensemble

2) Compress at Constant Rate

3) Record end-to-end Distance



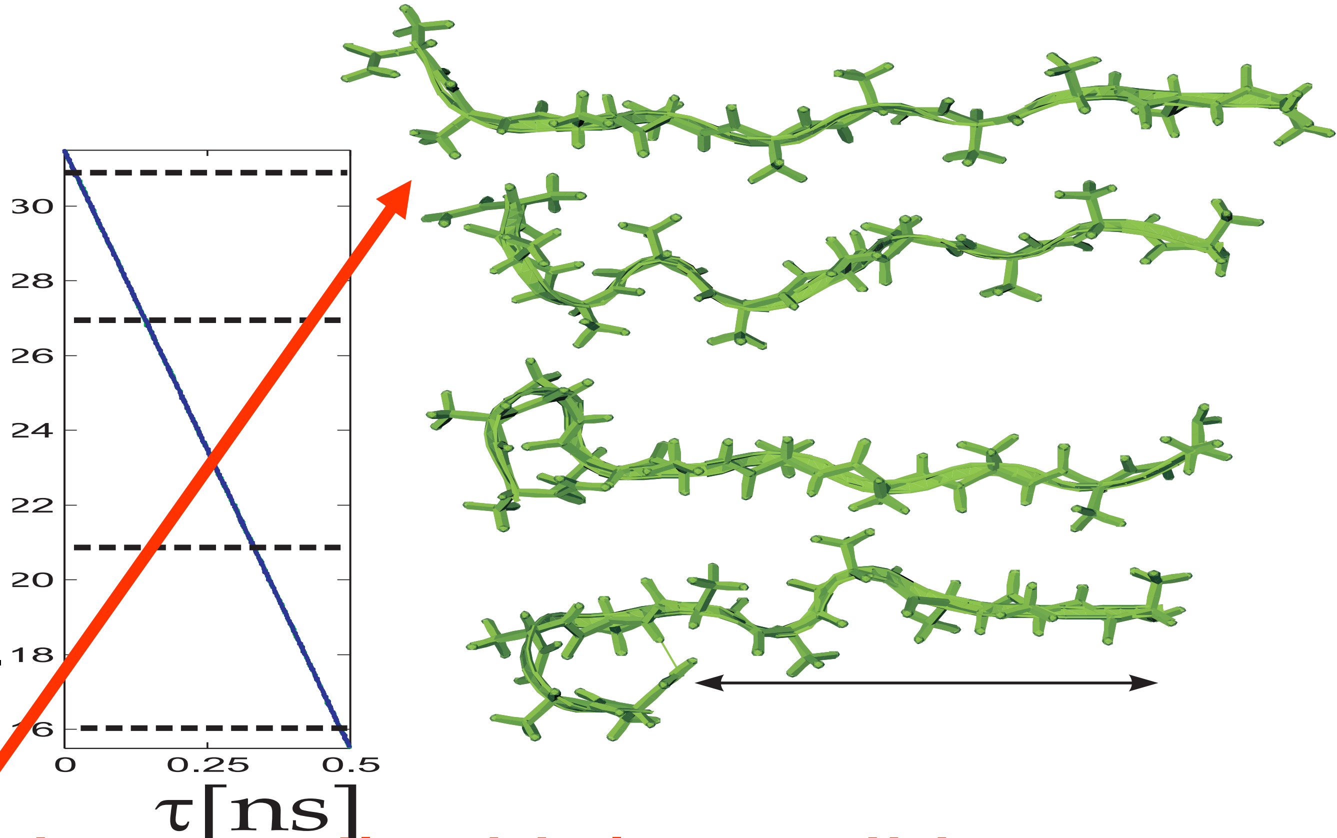
Releasing Tension Allows Helix Formation

Tracking the Snapshots

1) Draw Random Coil From Ensemble

2) Compress at Constant Rate

3) Record end-to-end Distance



New Random Coil Initial Condition

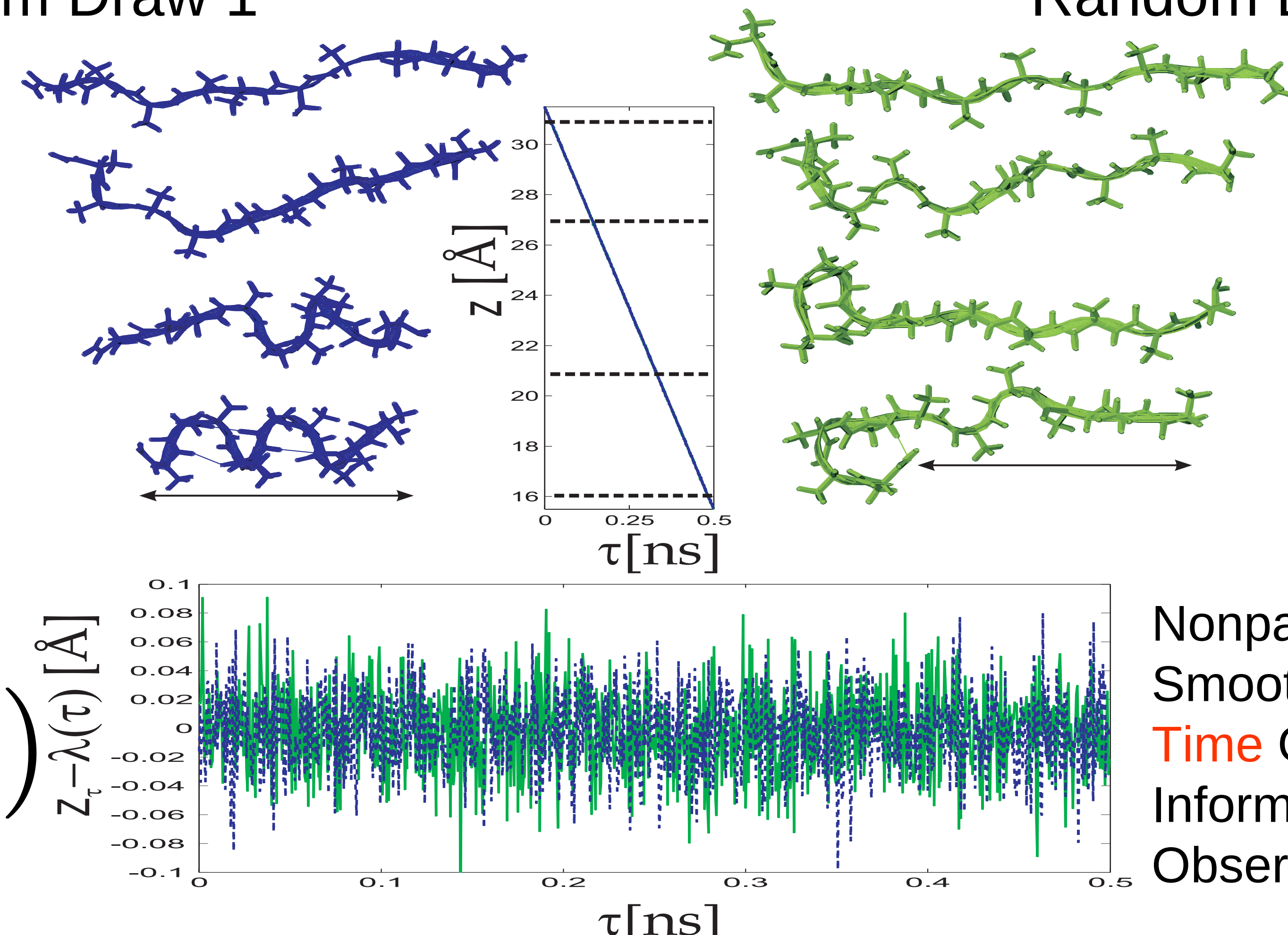
[same z, different “p and q”]

Limited Information in “z”

Random Draw 1 | Random Draw 2

t=0

t=T



$d(\tau)$

Nonparametric Smoothing in Time Gives Little Information About Observed Data

One Approach: Infer Stochastic Evolution Rules for Each Path

Stochastic Differential Equation (SDE):

$$dz_t = b(z_t, t; \theta)dt + \sigma(z_t; \theta)dW_t$$

"Pathwise" Approach Does Not Require Stationarity or Ergodicity. A "Local Diffusion Coeffcient" or "Spot Volatility" Can Be Fit

COLLECTION of SDEs → Approximate Complex Distributions AND Have Loose Physical Interpretation

Learn By Wrapping Structure Around Data:Time Dependent Diffusions

Stochastic Differential Equation (SDE):

dB_t, dW_t Standard Brownian Motion / Wiener Process

$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$

Drift Function

Comes from External Added Force ("Constant Velocity" Pulling)

Instantaneous Noise Amplitude

Not Surprisingly, Diffusive SDEs Are Popular in Mathematical Finance

$$dz_t = b(z_t, t; \theta)dt + \sigma(z_t, t; \theta)dW_t$$

Drift

E.g. Seasonal Trend

Fast Scale Noise

Assumes Levy Process Adequate (Hope Correlated Noise Like "Bid/Ask" Spread "Average Out")

Not Surprisingly, Popular in Finance

$$dz_t = b_t dt + \sigma_t dB_t$$

"Drift"

"Spot Volatility"

Also Stochastic Volatility Modeling Allows Processes Beyond Diffusions… But My Community Prefers Diffusions for Various Reasons

$$dz_t = b_t dt + \sigma_t dB_t + \lambda_t dN_t$$

"Jump Terms"

SDE Illustration

$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$

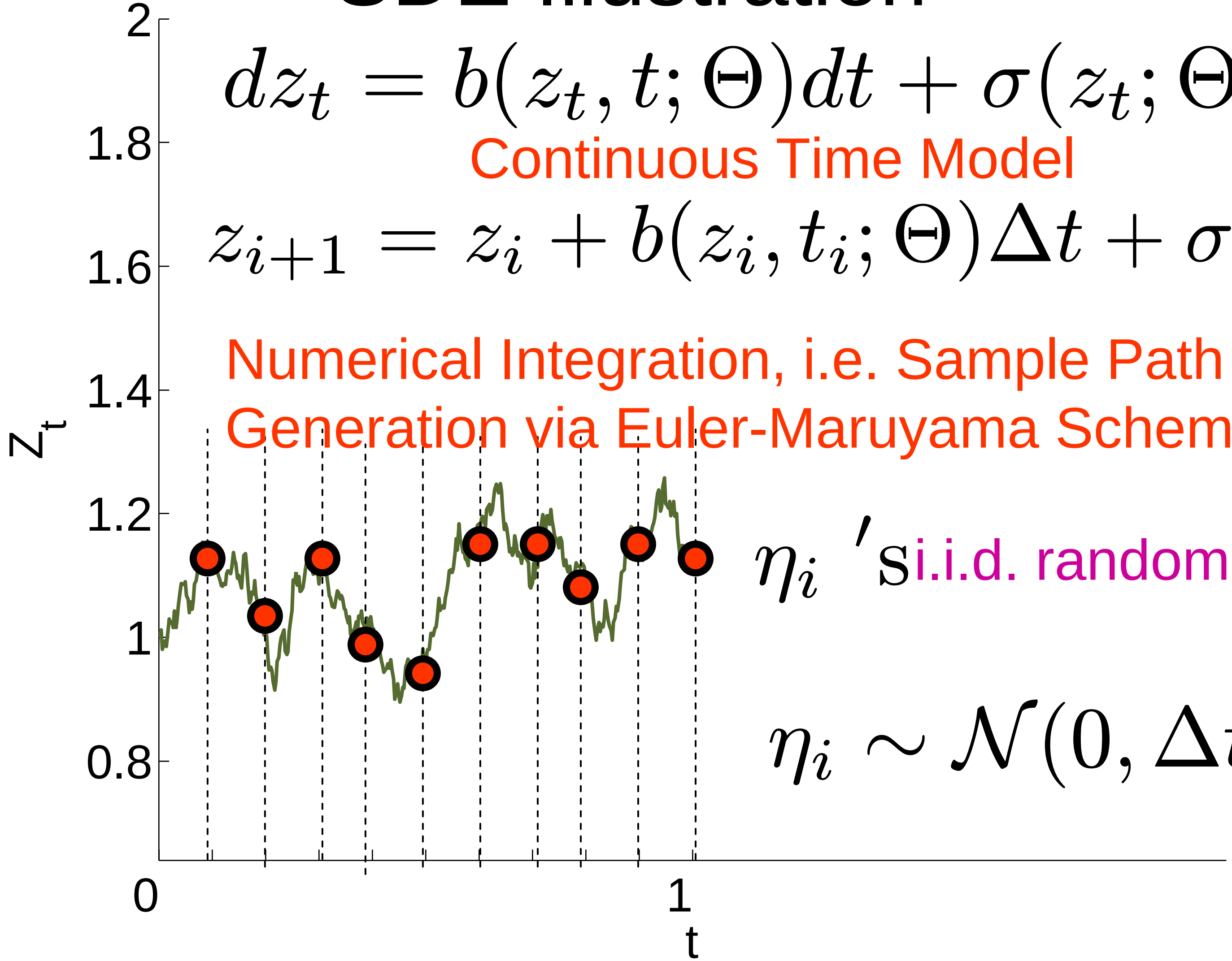
Continuous Time Model

$$z_{i+1} = z_i + b(z_i, t_i; \Theta)\Delta t + \sigma(z_i; \Theta)\eta_i$$

Numerical Integration, i.e. Sample Path Generation via Euler-Maruyama Scheme

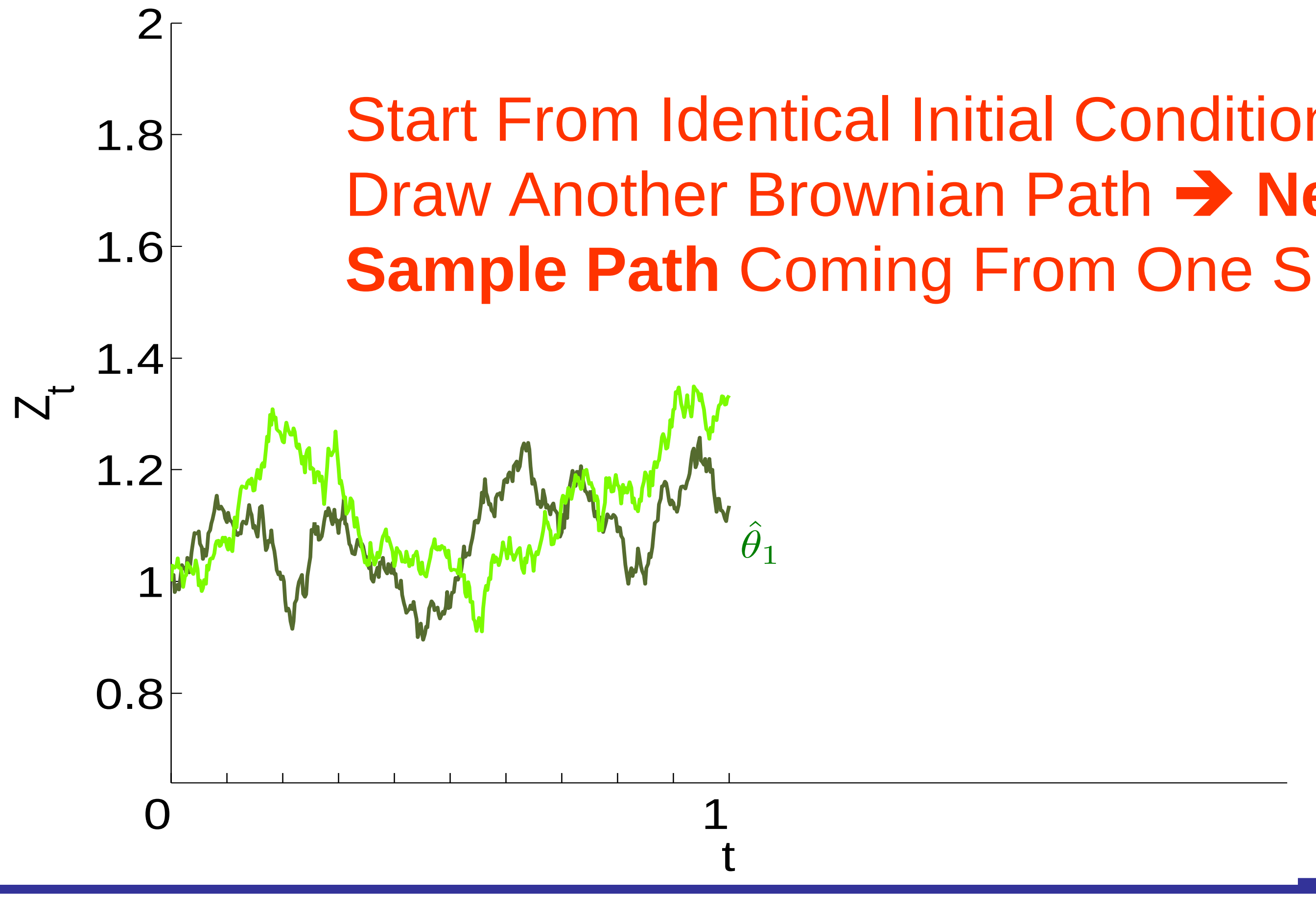
η_i 's i.i.d. random variables

$$\eta_i \sim \mathcal{N}(0, \Delta t)$$



Stochastic Process

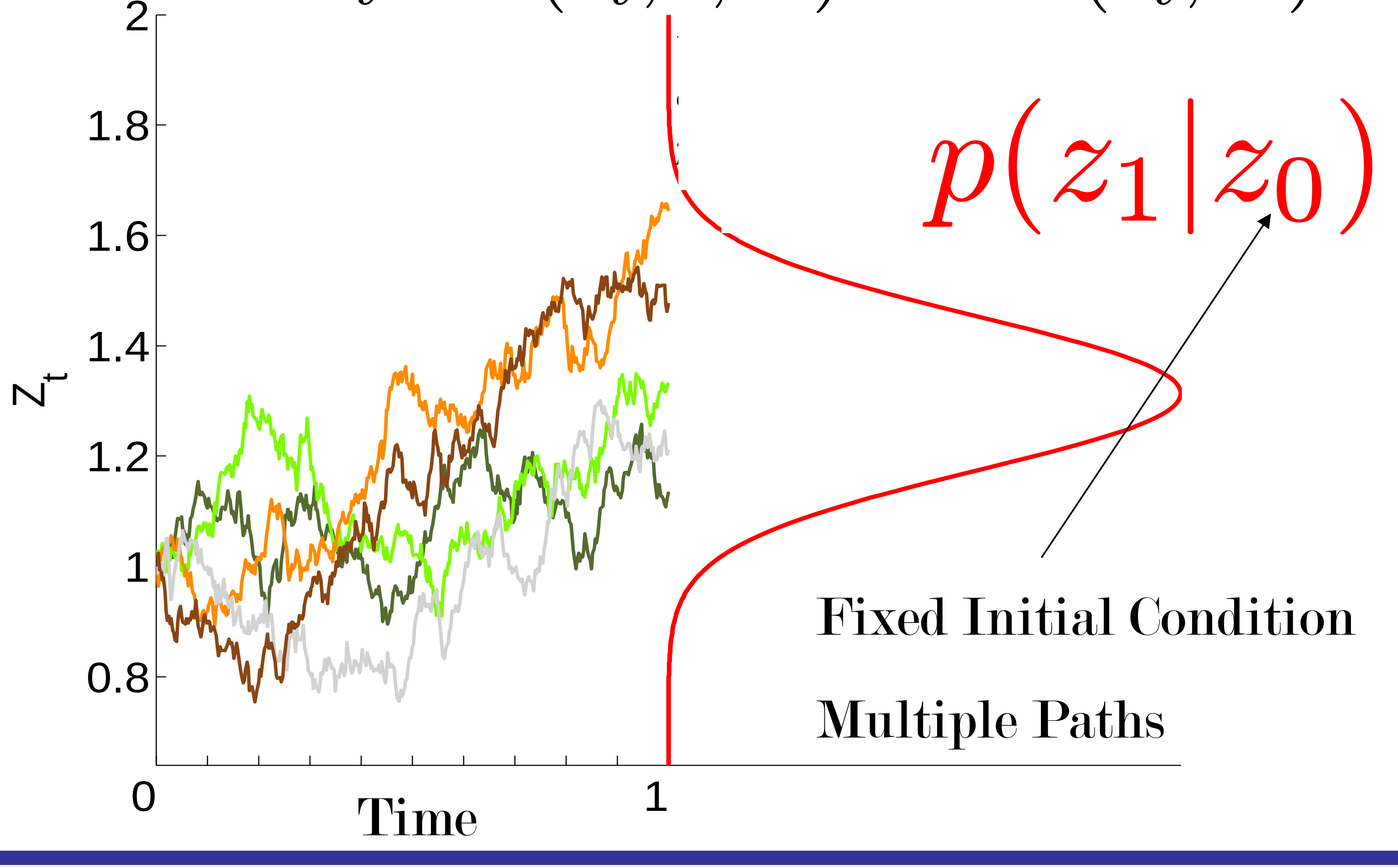
$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$



Start From Identical Initial Condition and Draw Another Brownian Path ➔ **New Sample Path** Coming From One SDE

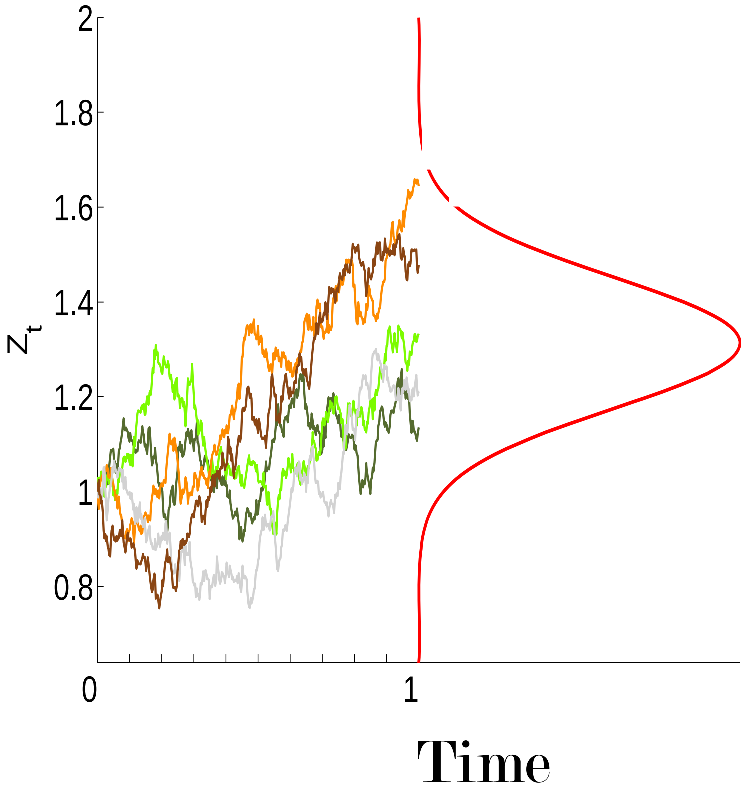
Each Single SDE Connected to a PDE

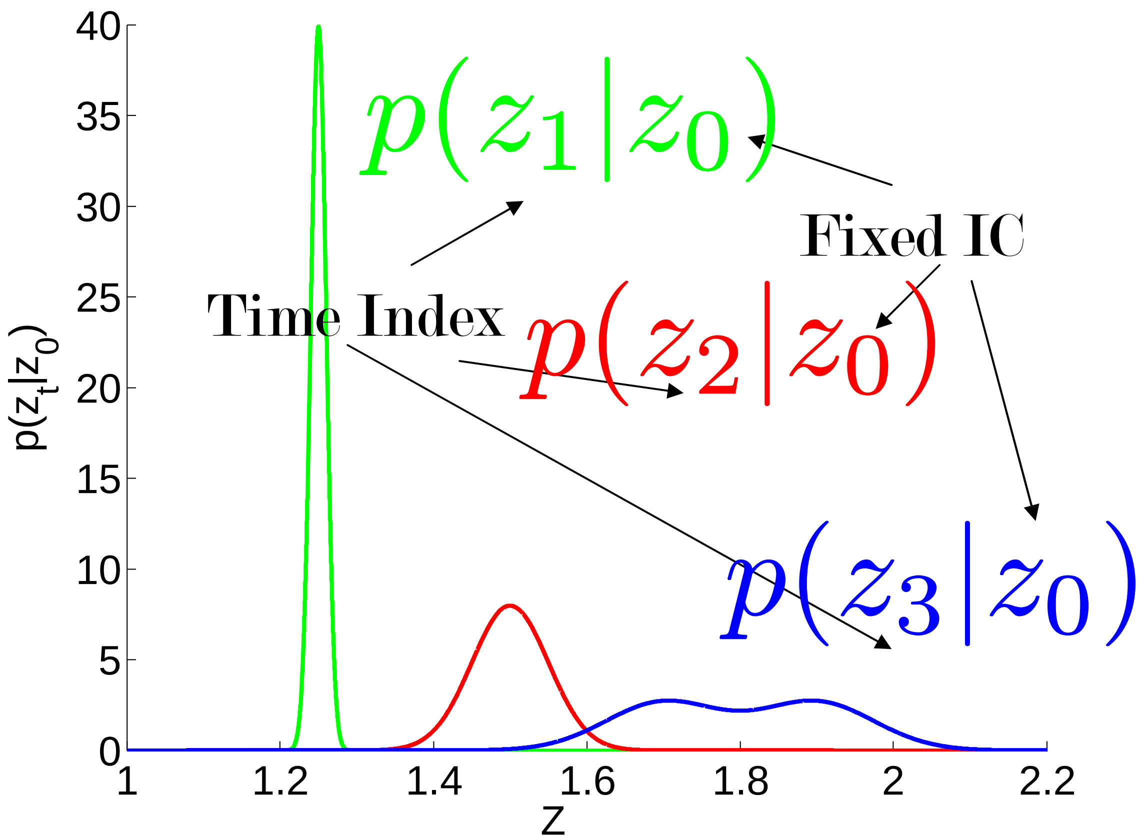
$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$



Each Single SDE Connected to a PDE

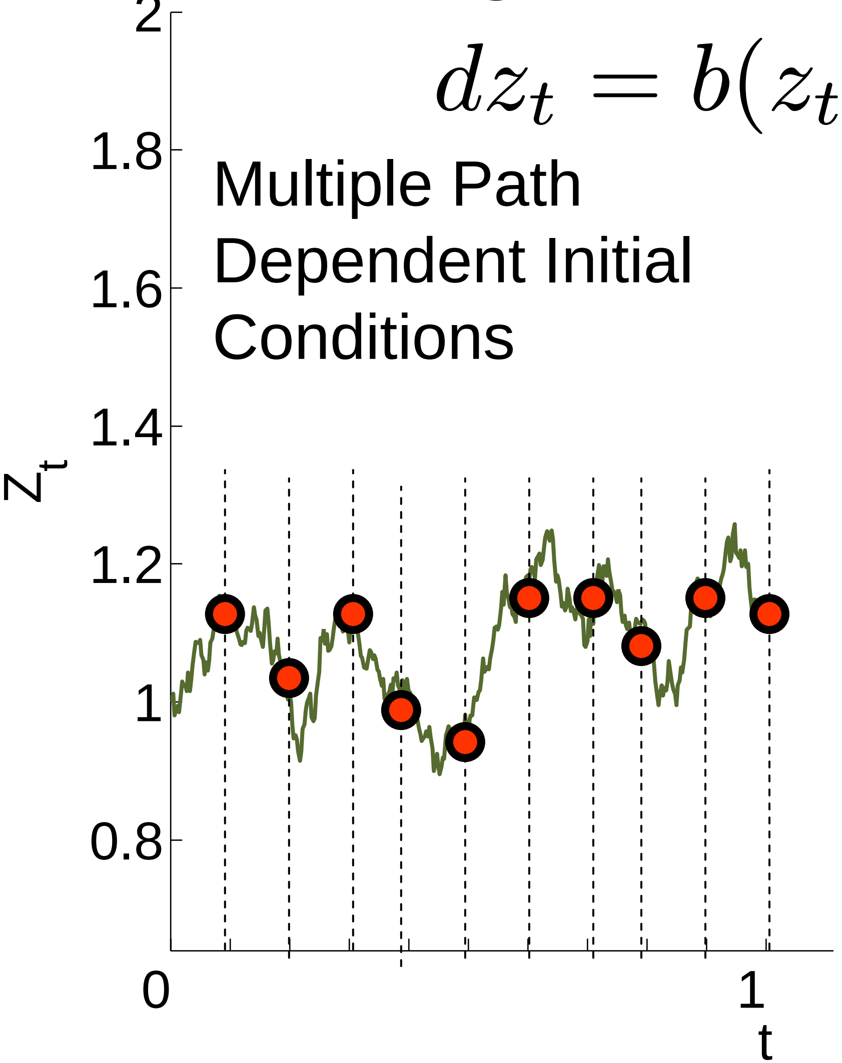
$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$

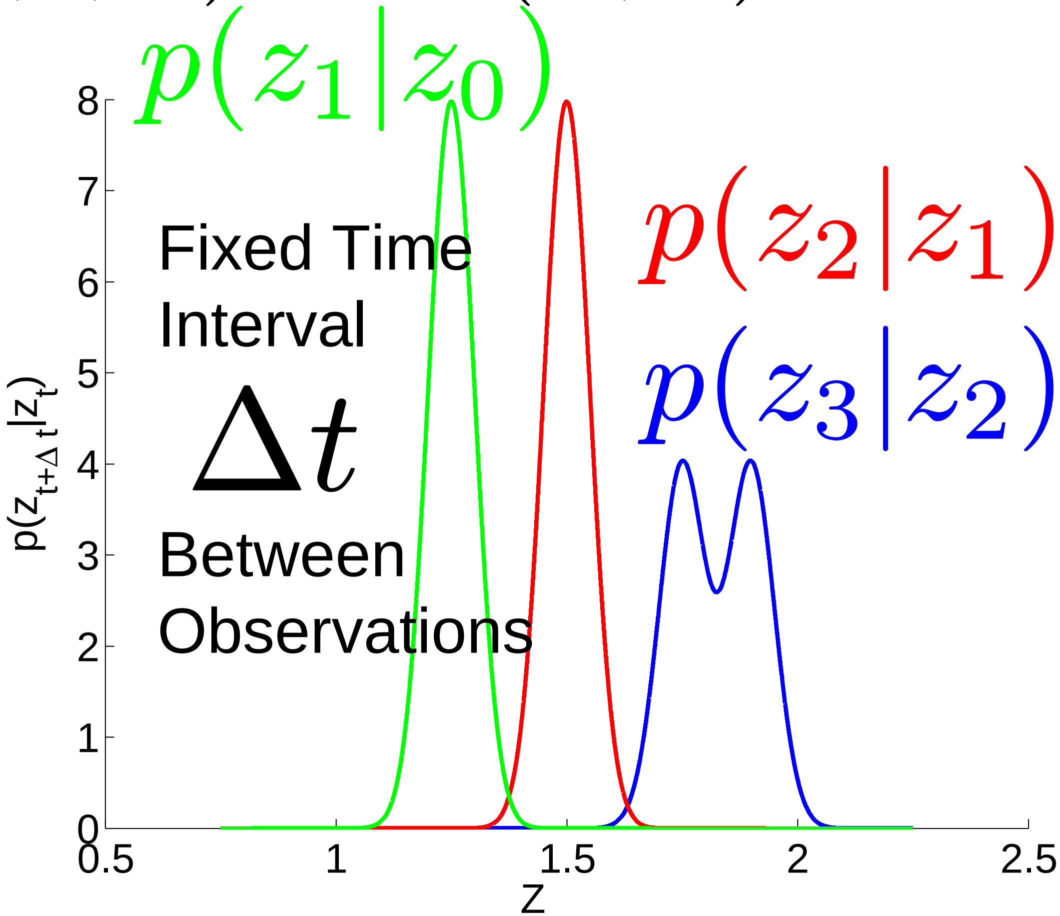




Each Single SDE Connected to a PDE

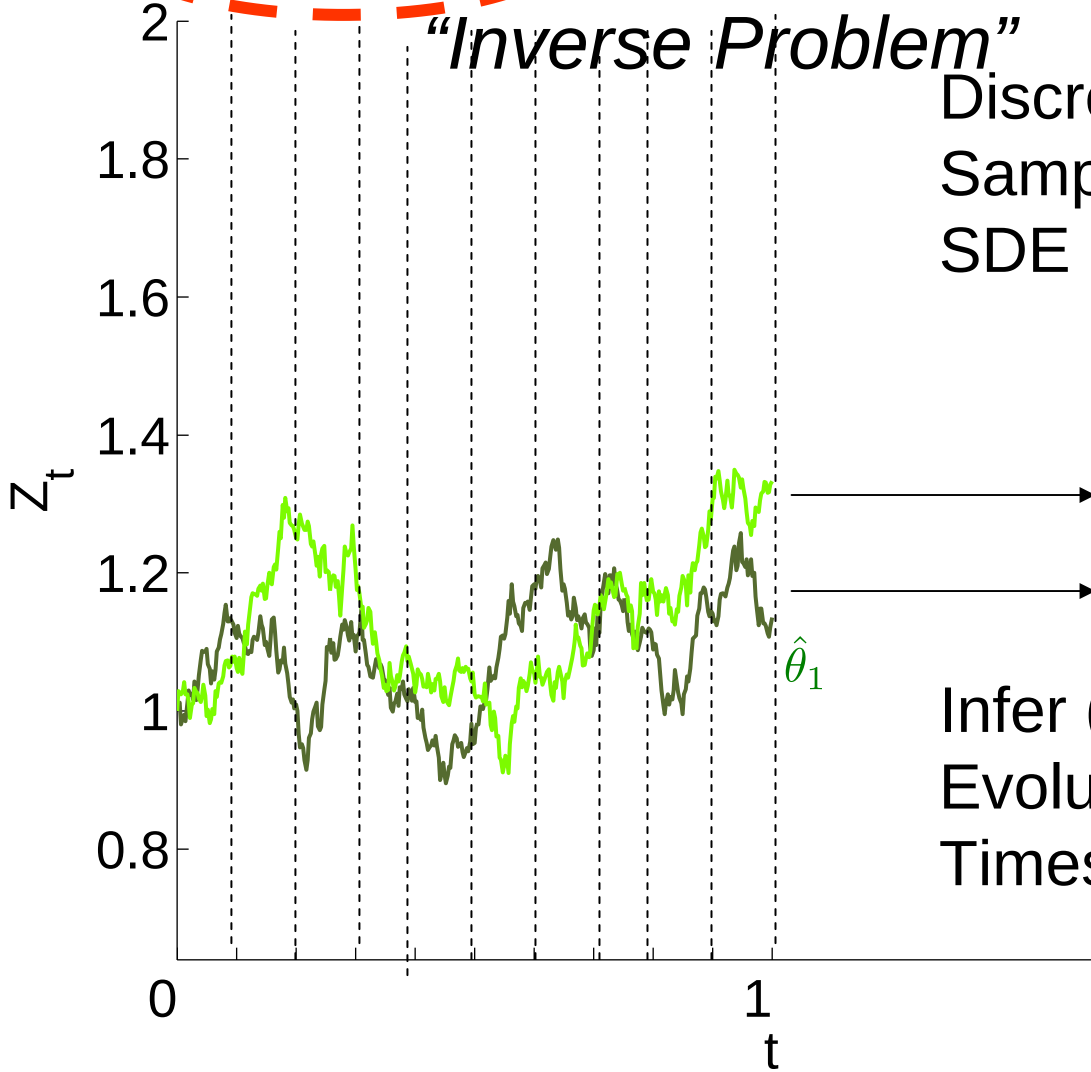
$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$





Pathwise Estimation View

"Inverse Problem"



Discrete Times Series Sample of Continuous SDE Paths

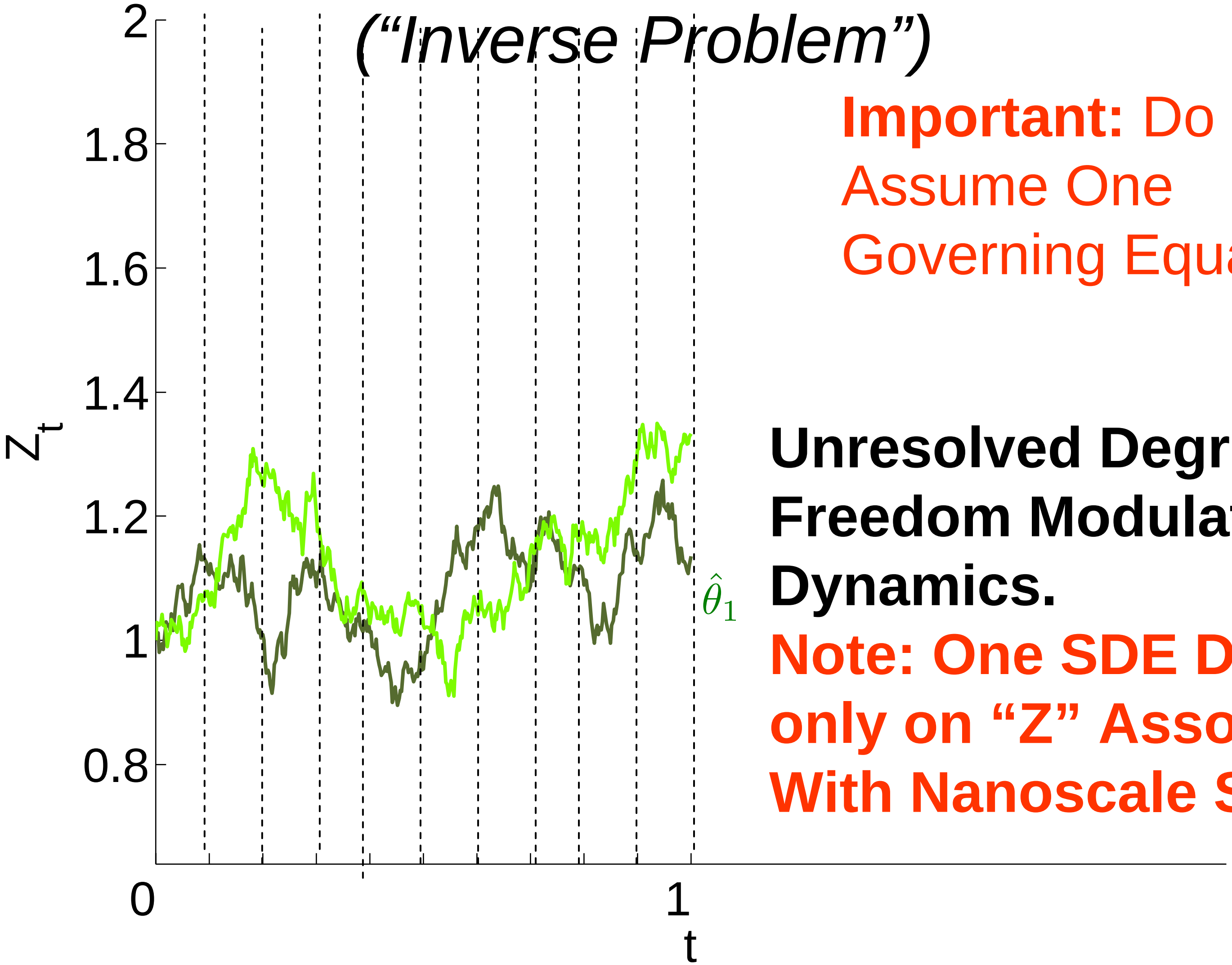
$$\left(z_0, \ldots, z_T\right)^{(2)}$$

$$\left(z_0, \ldots, z_T\right)^{(1)}$$

Infer (Stochastic) Evolution Rules from Times Series.

Pathwise Estimation View

("Inverse Problem")



Important: Do NOT Assume One Governing Equation

Unresolved Degrees of Freedom Modulate Dynamics.

Note: One SDE Depending only on "Z" Associated With Nanoscale System

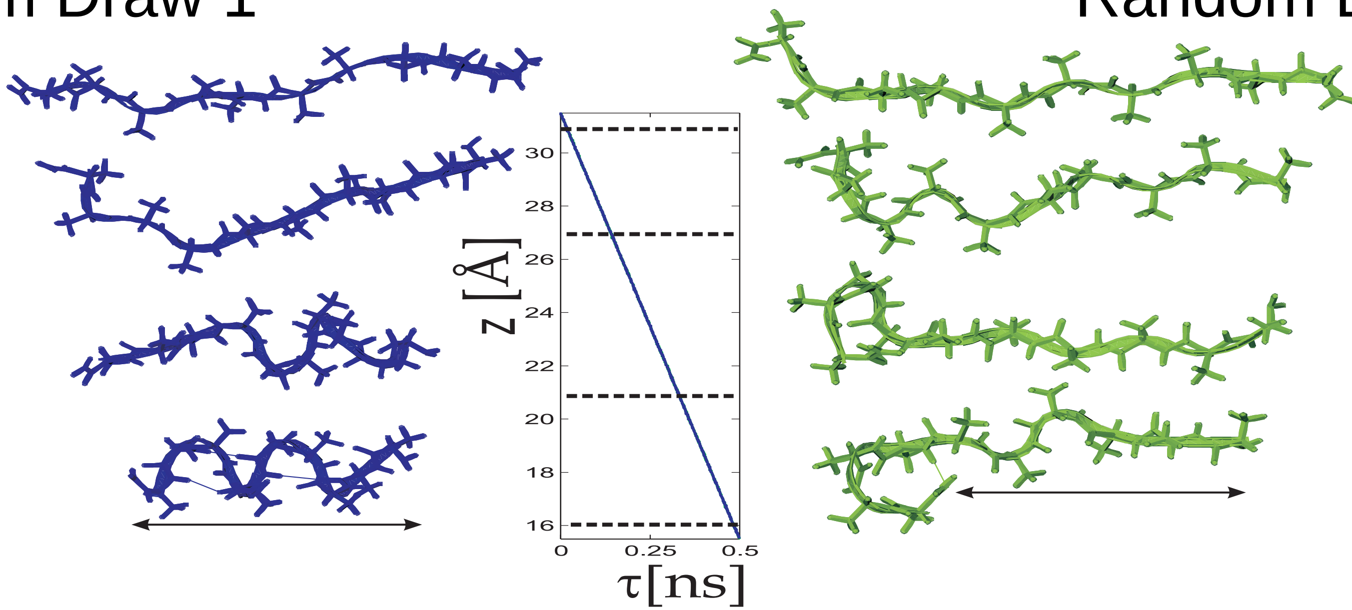
Recall the Previous Example

Random Draw 1

Random Draw 2

t=0

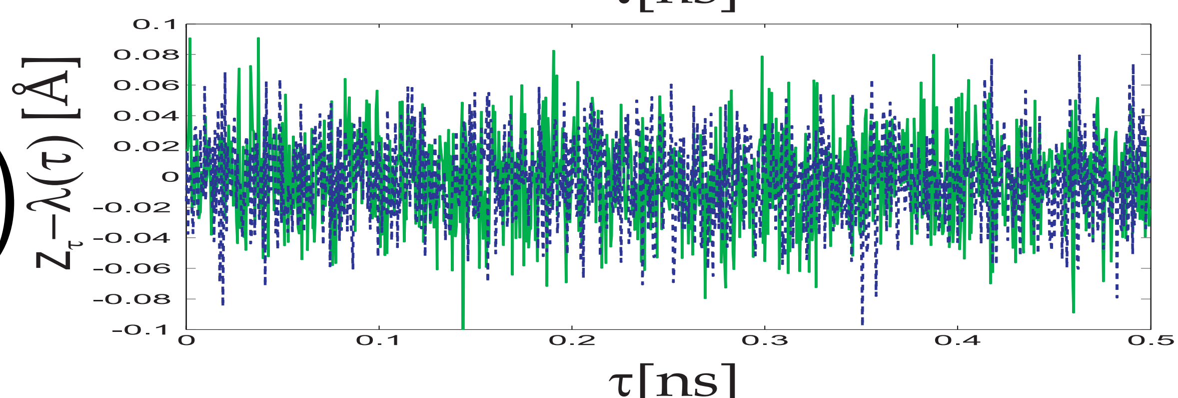
t=T



t=0

t=T

$$d(\tau)$$



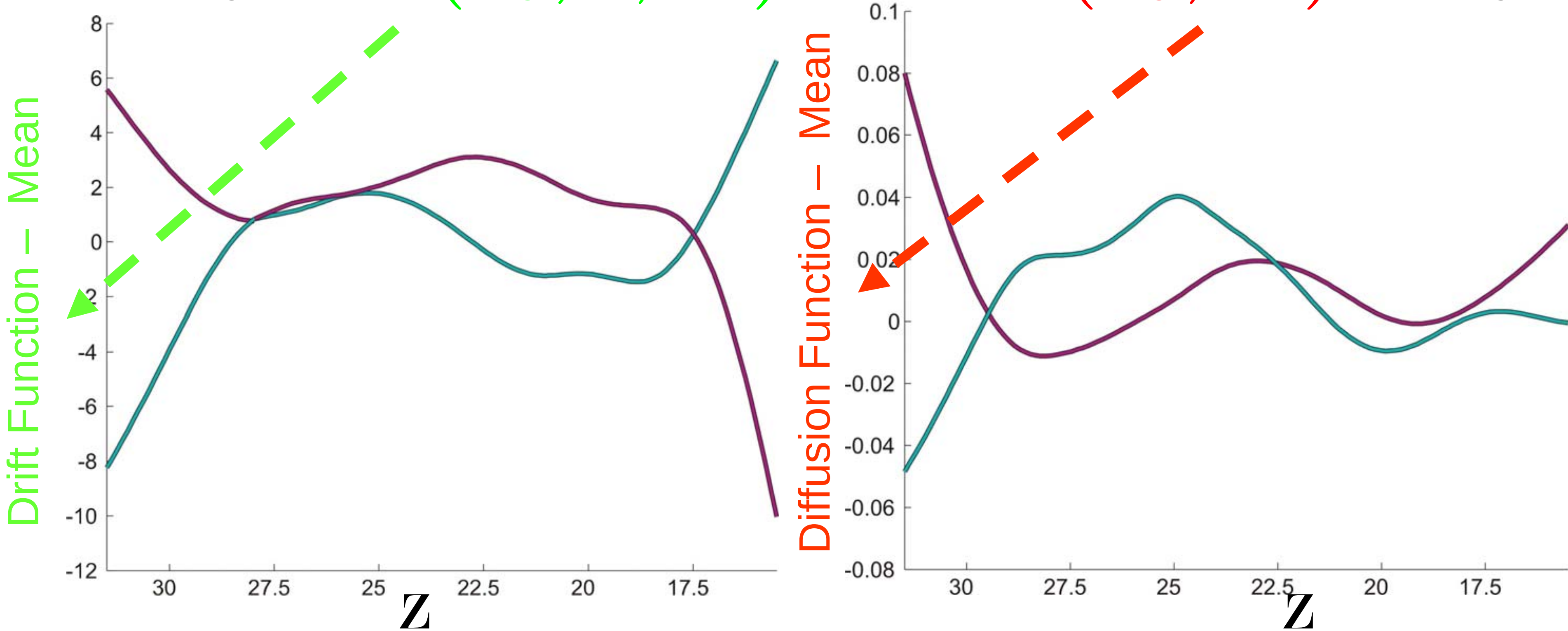
Nonparametric Smoothing in Time Gives Little Information About Observed Data

SDE Function Curves

Calderon, J. Chem. Phys. **127** (2007); Calderon & Chelli, *J. Chem. Phys.* **128** (2008).

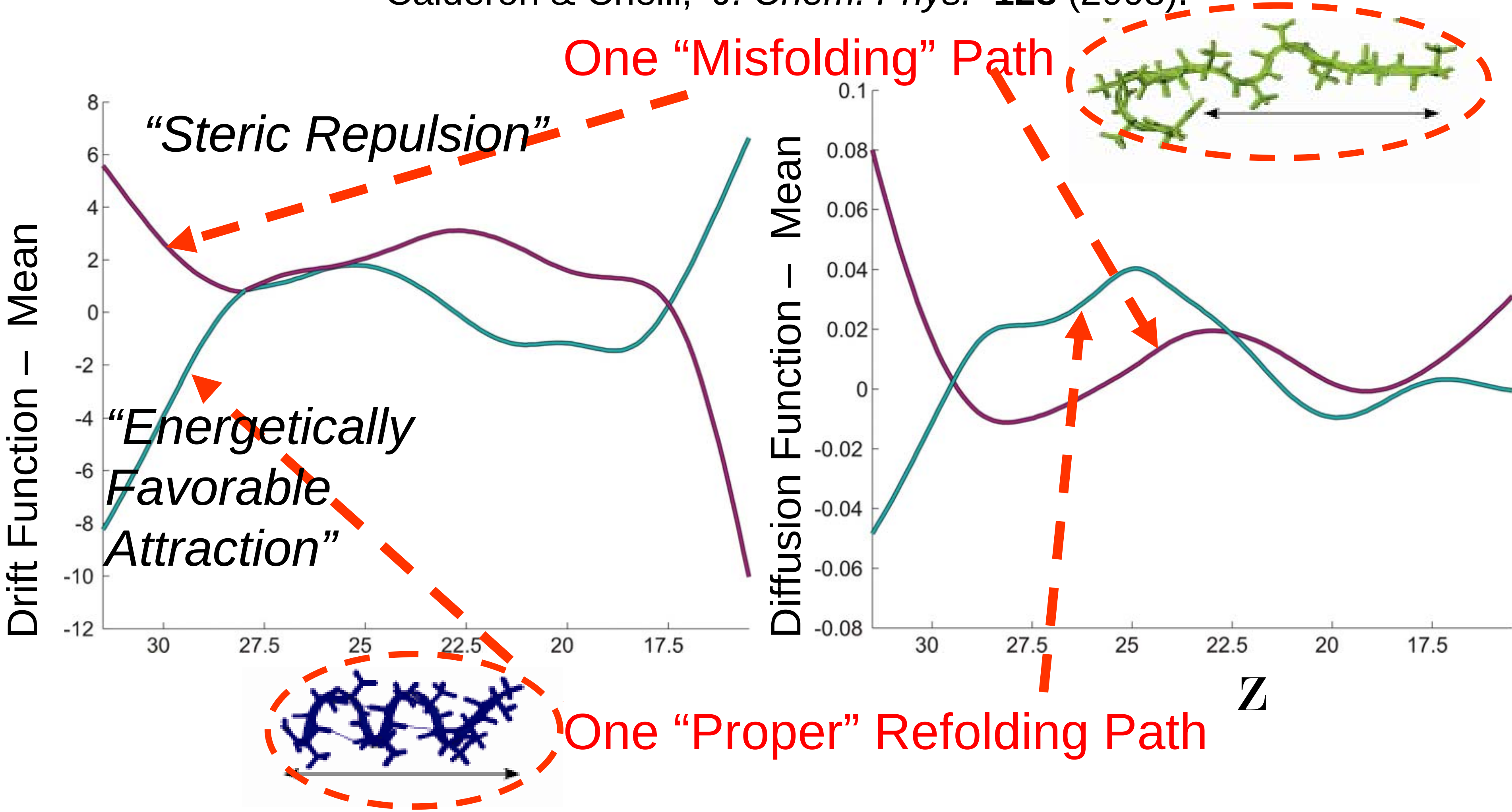
Summarize Each Pulling Experiment Information Into 2 Curves

$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$



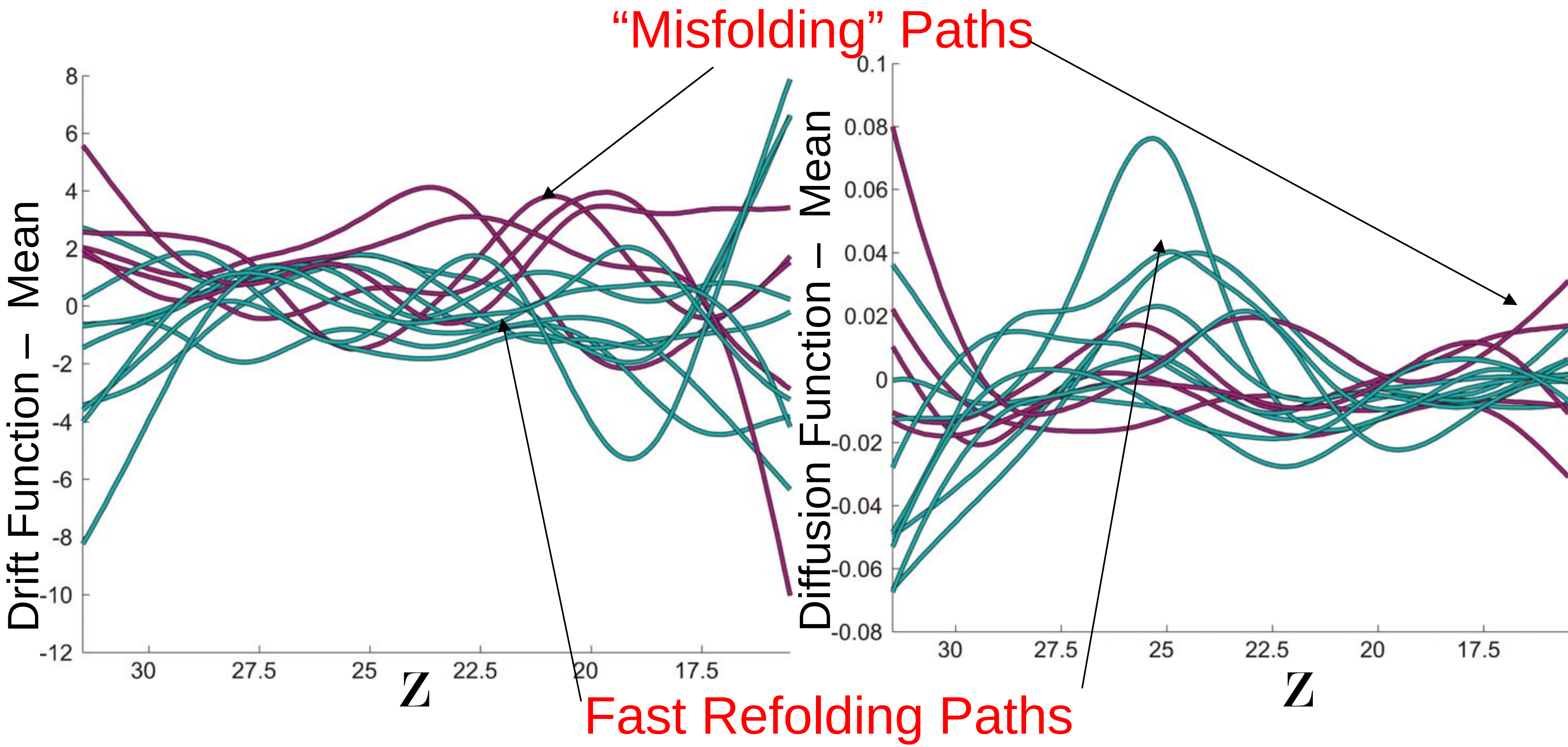
SDE Function Curves

Calderon & Chelli, *J. Chem. Phys.* **128** (2008).



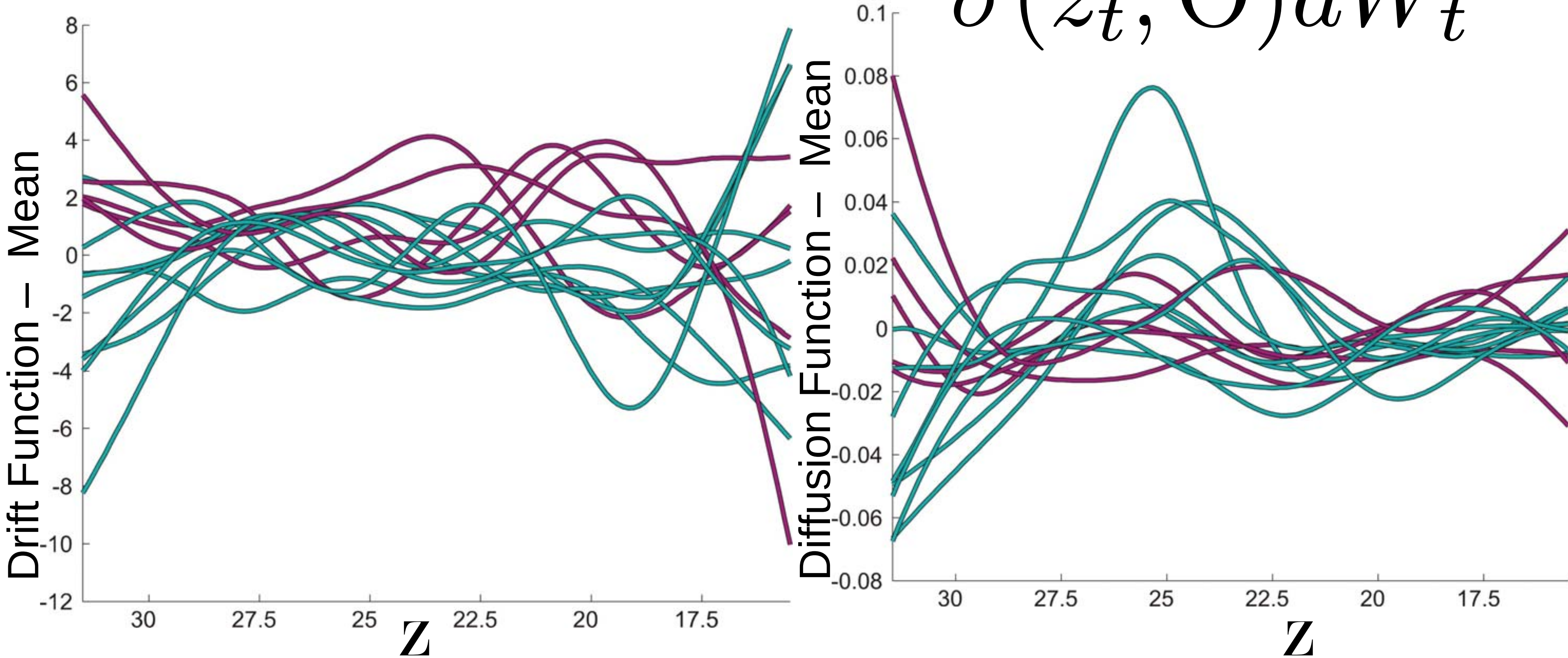
Population of SDE Curves: Functional Data Analysis

Calderon & Chelli, *J. Chem. Phys.* **128** (2008).



Fluctuations are Informative and Have Connection to Physical Phenomena. Quantified by Local Diffusion or "Spot Volatility":

$$\sigma(z_t;\Theta)dW_t$$

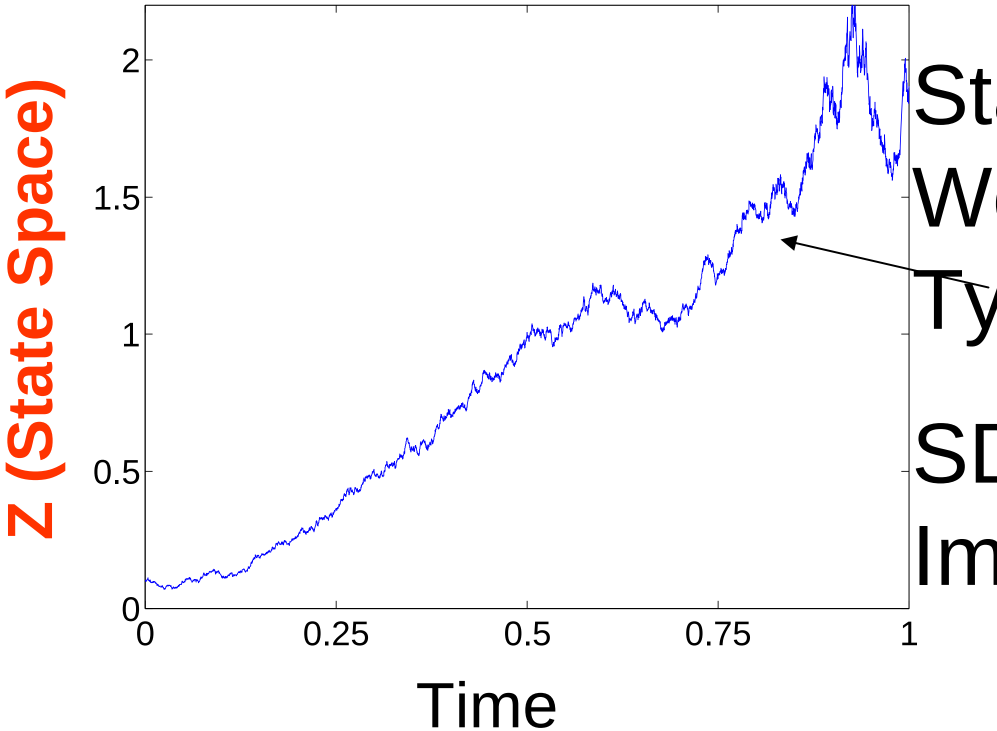


Importance, Complications, and Benefits Associated with Modeling Non-Trivial “Spot Volatility” Not Widely Appreciated in Statistical Mechanics and Chemical Physics

Reliable Statistical Inference Tools Hard to Come by In Non-stationary Diffusion Models with Nonconstant Local Diffusion (Spot Volatility Is a Function or Process)

Note: Noise Magnitude Depends on State

And the State is Evolving in a Non-Stationary Fashion



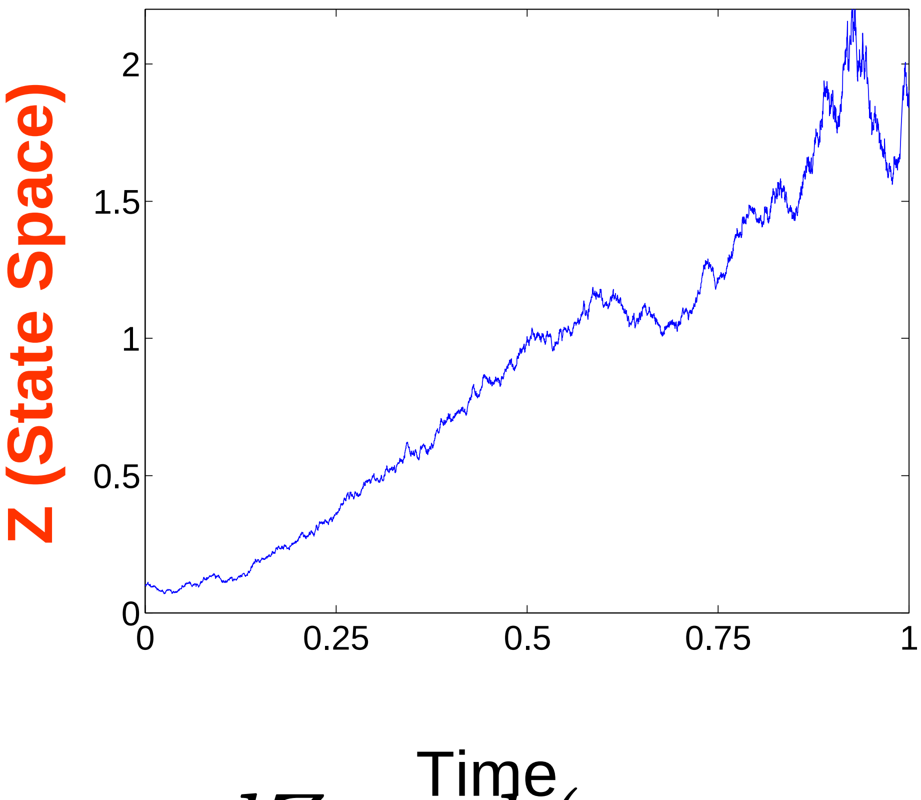
State Dependent Noise Would "Fool" Wavelet Type Methods

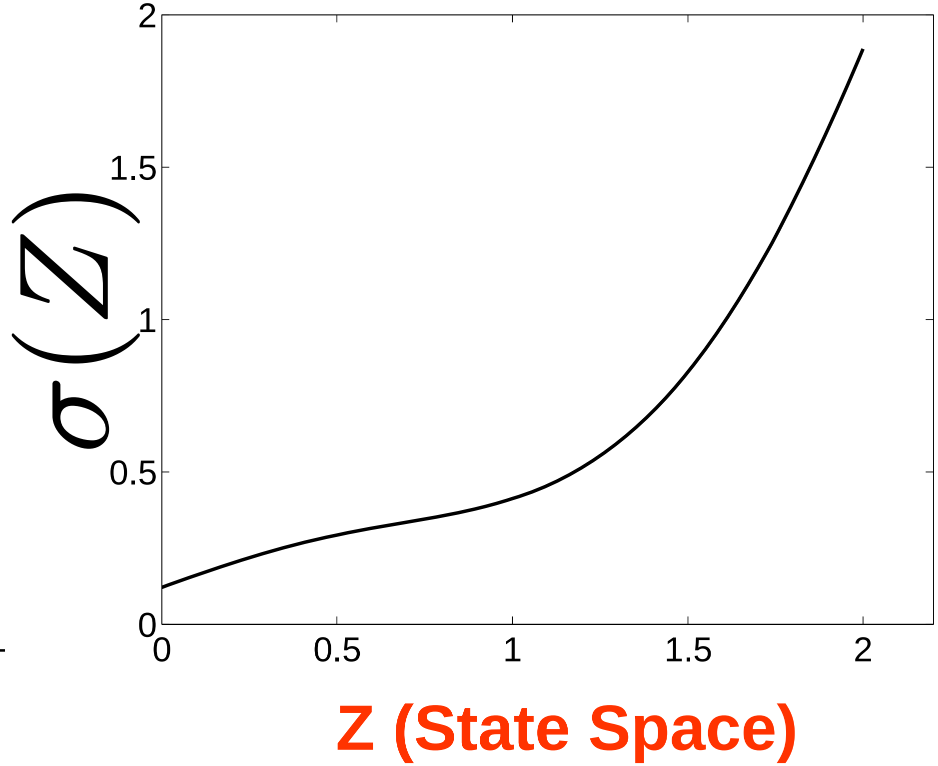
SDE Structure Helpful to Impose for EDA

$$dZ_t = k(v_{\text{pull}}t - Z_t)dt + \sigma(Z_t)dW_t$$

Note: Noise Magnitude Depends on State

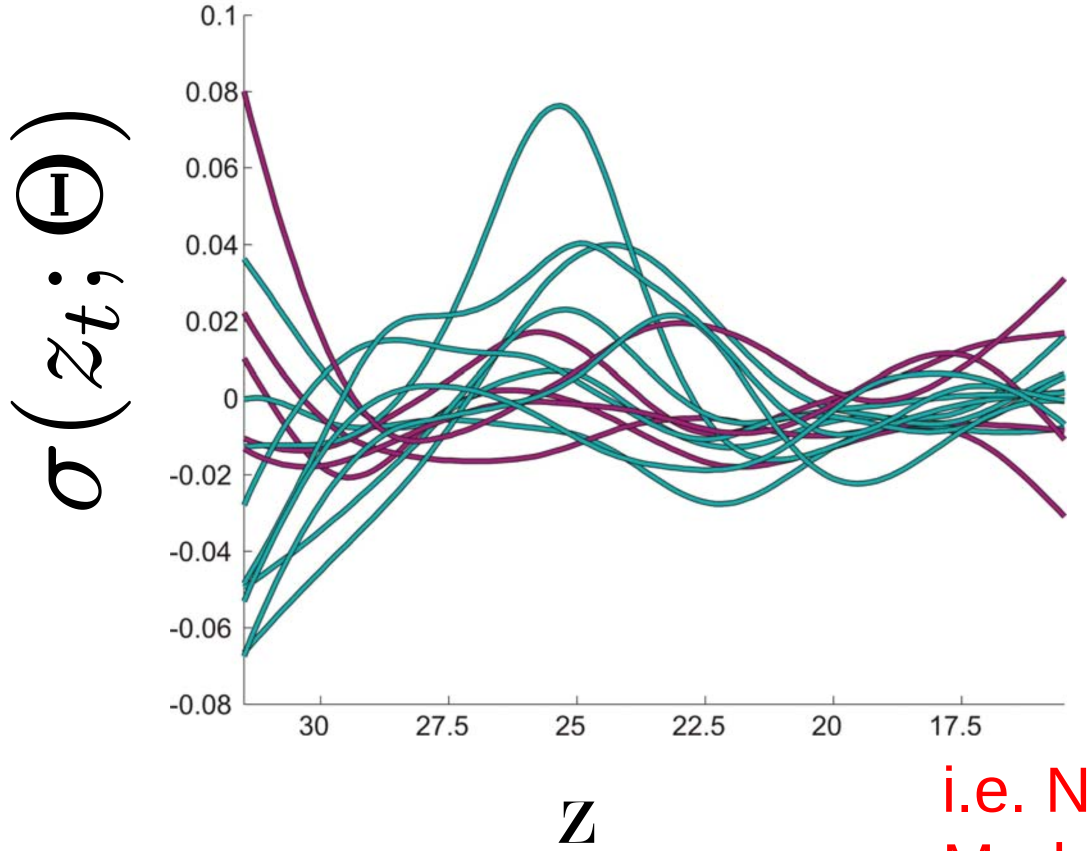
And the State is Evolving in a Non-Stationary Fashion





$$dZ_t = k(v_{\text{pull}}t - Z_t)dt + \sigma(Z_t)dW_t$$

Fluctuations are Informative and Have Connection to Physical Phenomena. Quantified by Diffusion Function:



Global Drift and Diffusion Functions Nonlinear and of *a priori* Unknown Functional Form

$$\mu(z_t, t; \Theta)$$
$$\sigma(z_t; \Theta)$$

i.e. No Explicit Parametric Model Gives These Functions

Fluctuations are Informative and Have Connection to Physical Phenomena. Quantified Partially by Diffusion (i.e. Spot Volatility):

How to Semiparmetrically Estimate These?

Global Drift and Diffusion Function Nonlinear and of *a priori* Unknown Functional Form

$$\mu(z_t, t; \Theta)$$
$$\sigma(z_t; \Theta)$$

Sketch of Some Mathematical Details in Later .....

i.e. No Explicit Parametric Model Gives These Functions

Fluctuations are Informative and Have Connection to Physical Phenomena. Quantified Partially by Diffusion (i.e. Spot Volatility):

How to Semiparmetrically Estimate These?

Sketch of Some Mathematical Details Later…..

Global Drift and Diffusion Function Nonlinear and of *a priori* Unknown Functional Form

$$\mu(z_t, t; \Theta)$$

$$\sigma(z_t; \Theta)$$

BUT First, One More Example Illustrating Variation Between These Functions Estimated Using Different Time Series are Informative

i.e. No Explicit Parametric Model Gives These Functions

Time Dependent Diffusions

Stochastic Differential Equation (SDE):

$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$

$$y_{t_i} = z_{t_i} + \epsilon_{t_i}$$ "Measurement" Noise

For Frequently Sampled Single-Molecule Experimental Data, "Physically Uninteresting Measurement Noise" Can Be Large Component of Signal.

Find/Approximate: $p(z_t, | y_s; \Theta)$

"Transition Density" / Conditional Probability Density

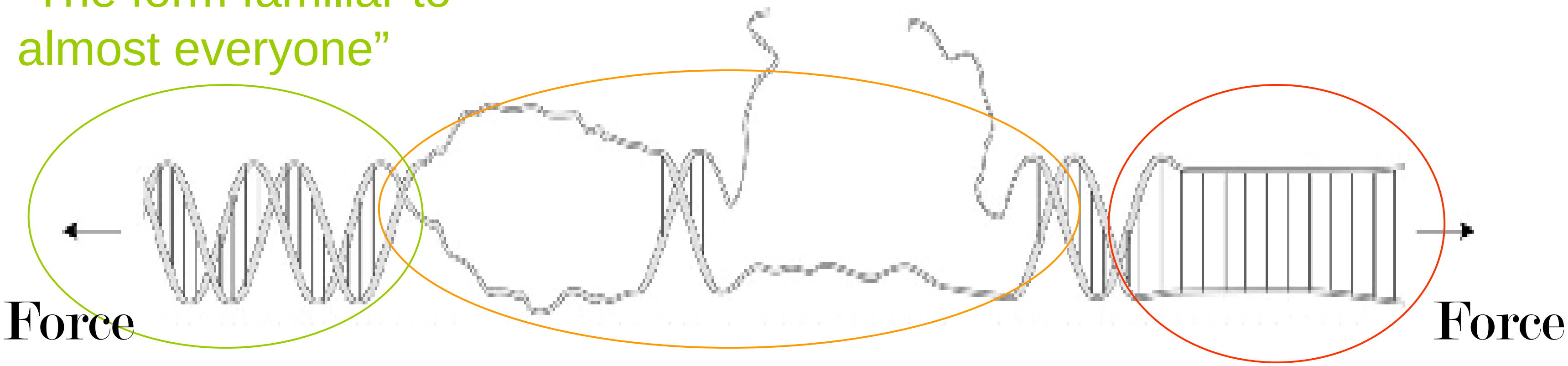
DNA Melting

B-DNA

“The form familiar to almost everyone”

“Molten” DNA

“S-DNA”

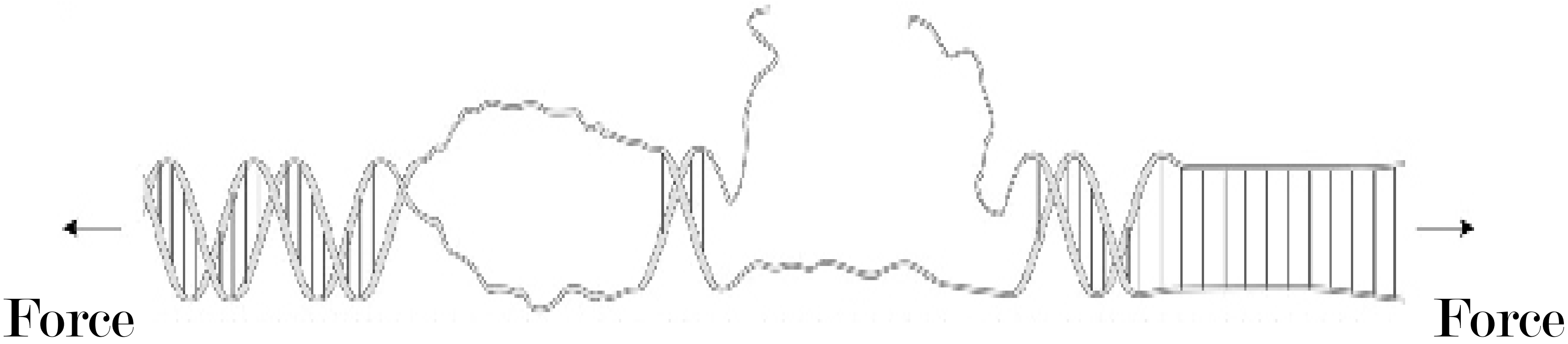


Bustamante, Bryant, & Smith, *Nature*, **421** (2003).

Cocco, Yan, Leger, Chatenay, & Marko, *PRE*, **70** (2004).

Whitelam, Pronk, & Geissler, *Biophys. J.,* **94** (2008).

DNA Melting



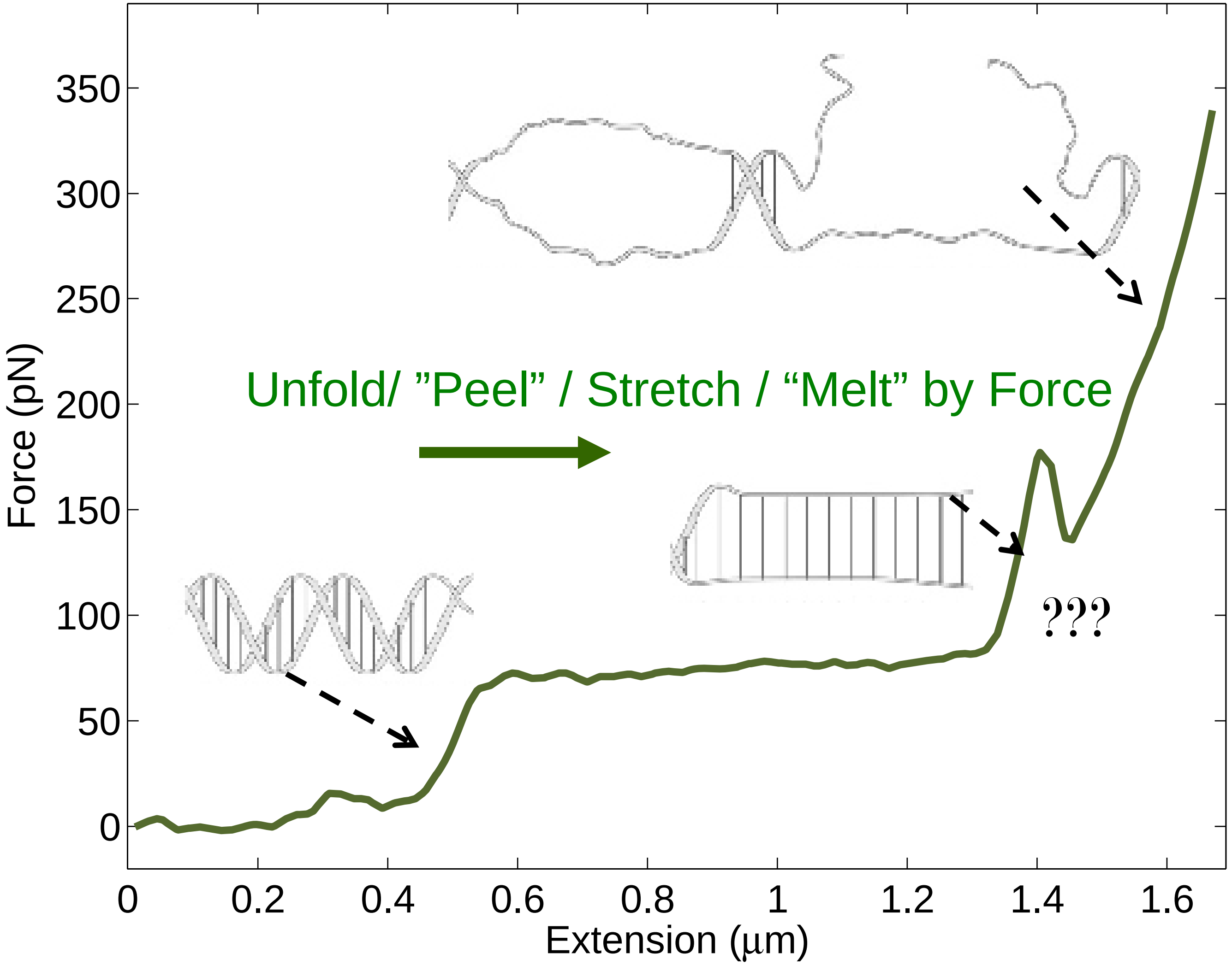
Tensions Important in Fundamental Life Processes Such As: DNA Repair and DNA Transcription

Single-Molecule Experiments Not Just a Neat Toy:

They Have Provided Insights Bulk Methods Cannot

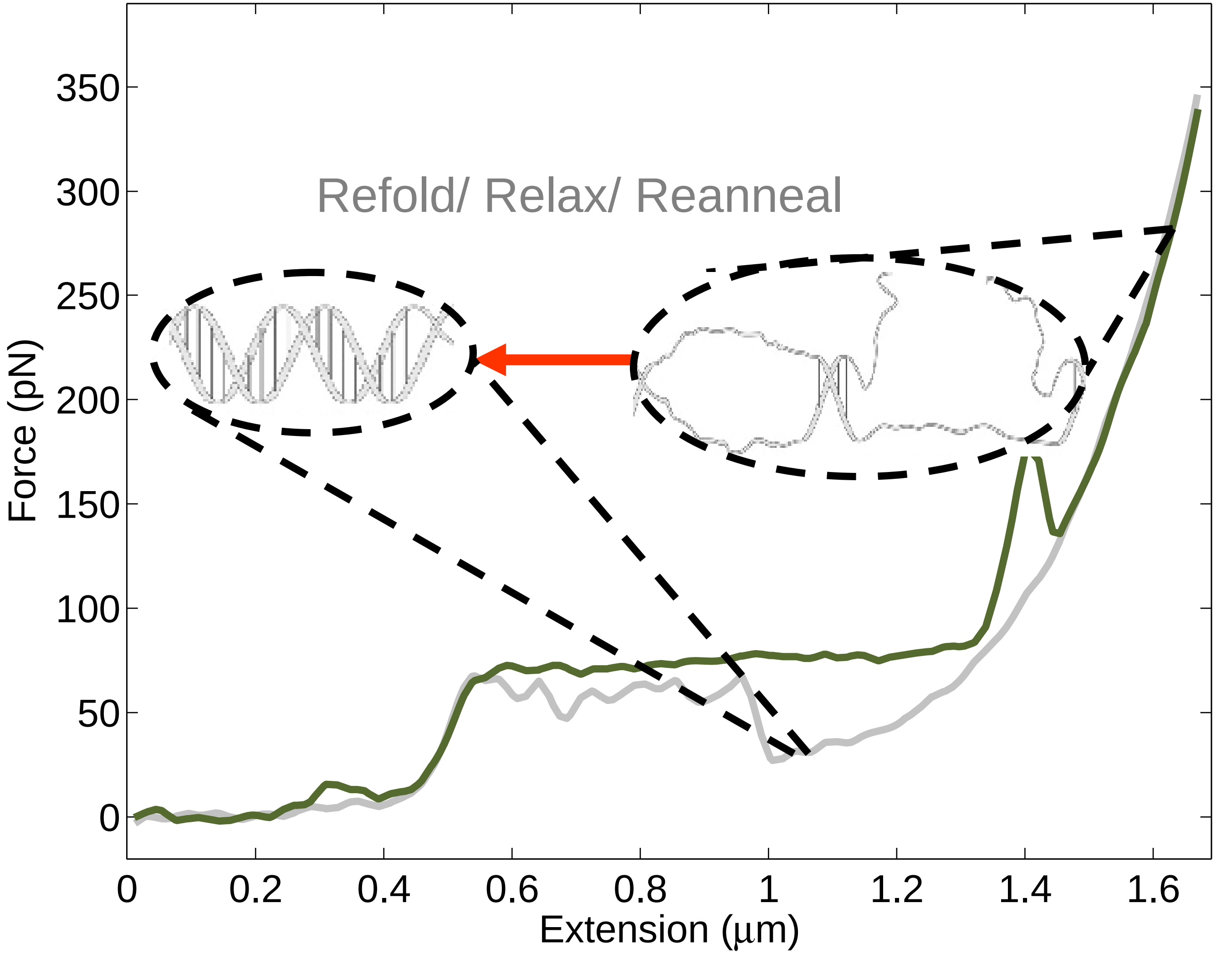
DNA Melting

Calderon, Chen, Lin, Harris, Kiang, *J. Phys.: Condensed Matter* (2009).



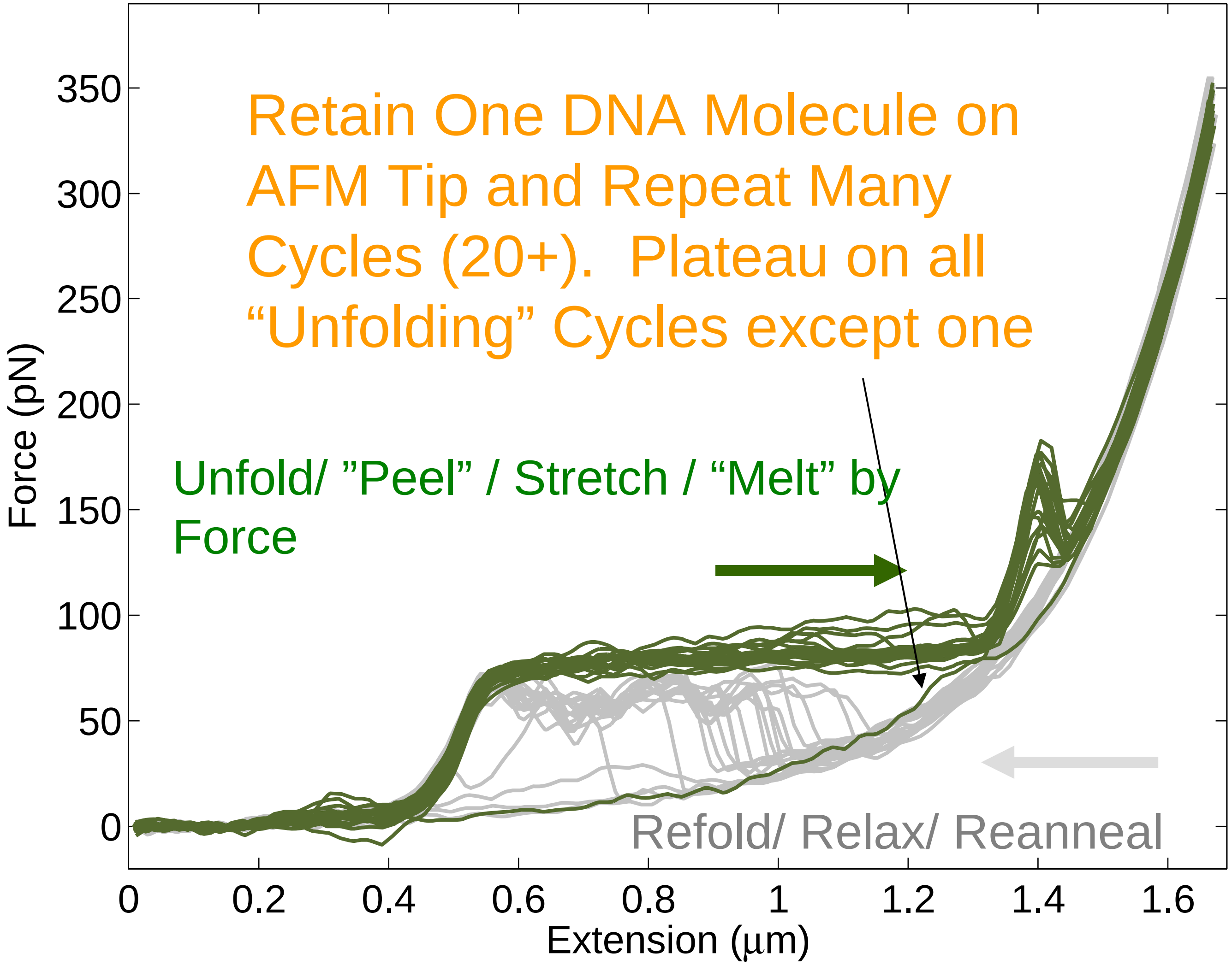
DNA Melting

Calderon, Chen, Lin, Harris, Kiang, *J. Phys.: Condensed Matter* (2009).



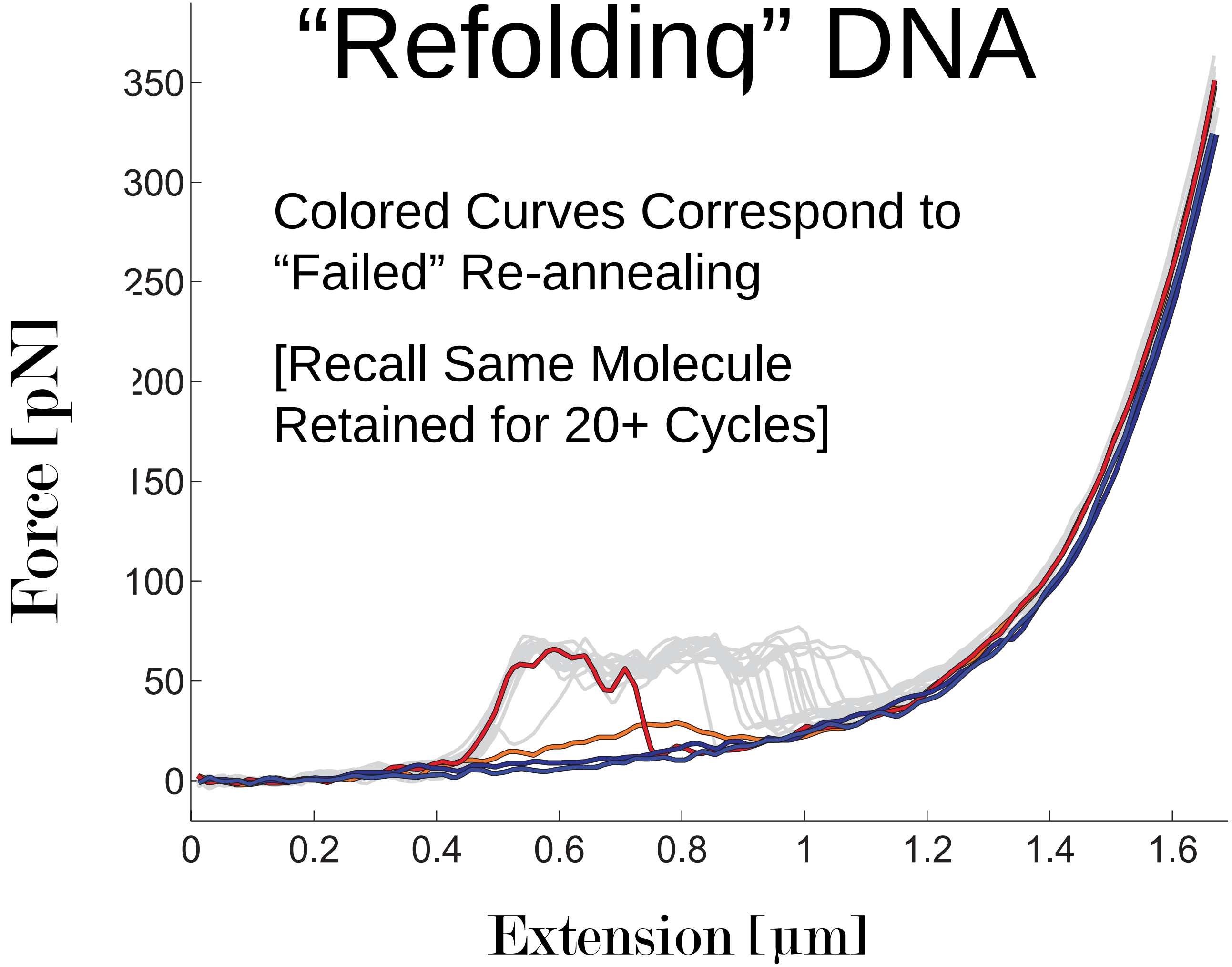
DNA Melting

Calderon, Chen, Lin, Harris, Kiang, *J. Phys.: Condensed Matter* (2009).



Significant Hysteresis Observed in “Refold/ Relaxation” Portion of Cycle

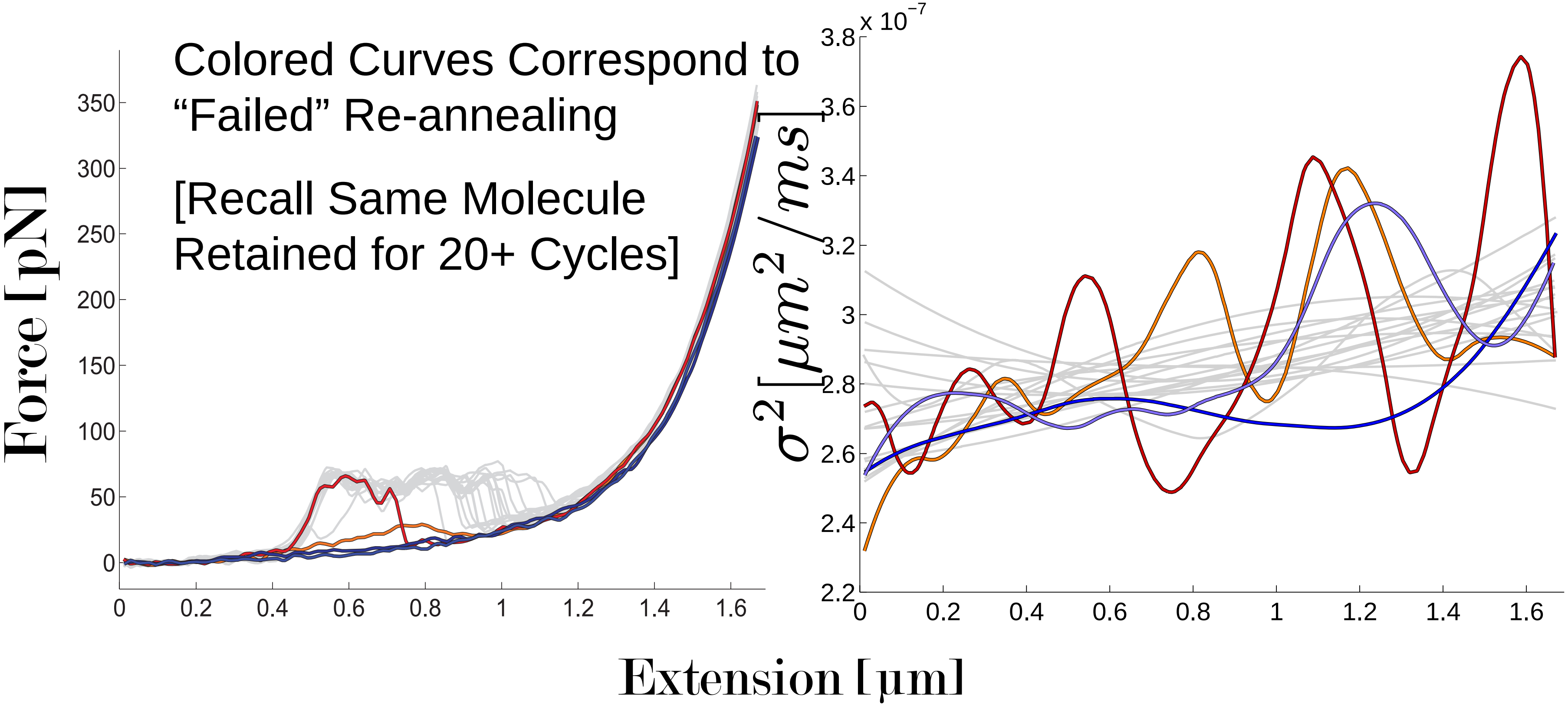
Kinetic Signatures Encountered in "Refolding" DNA



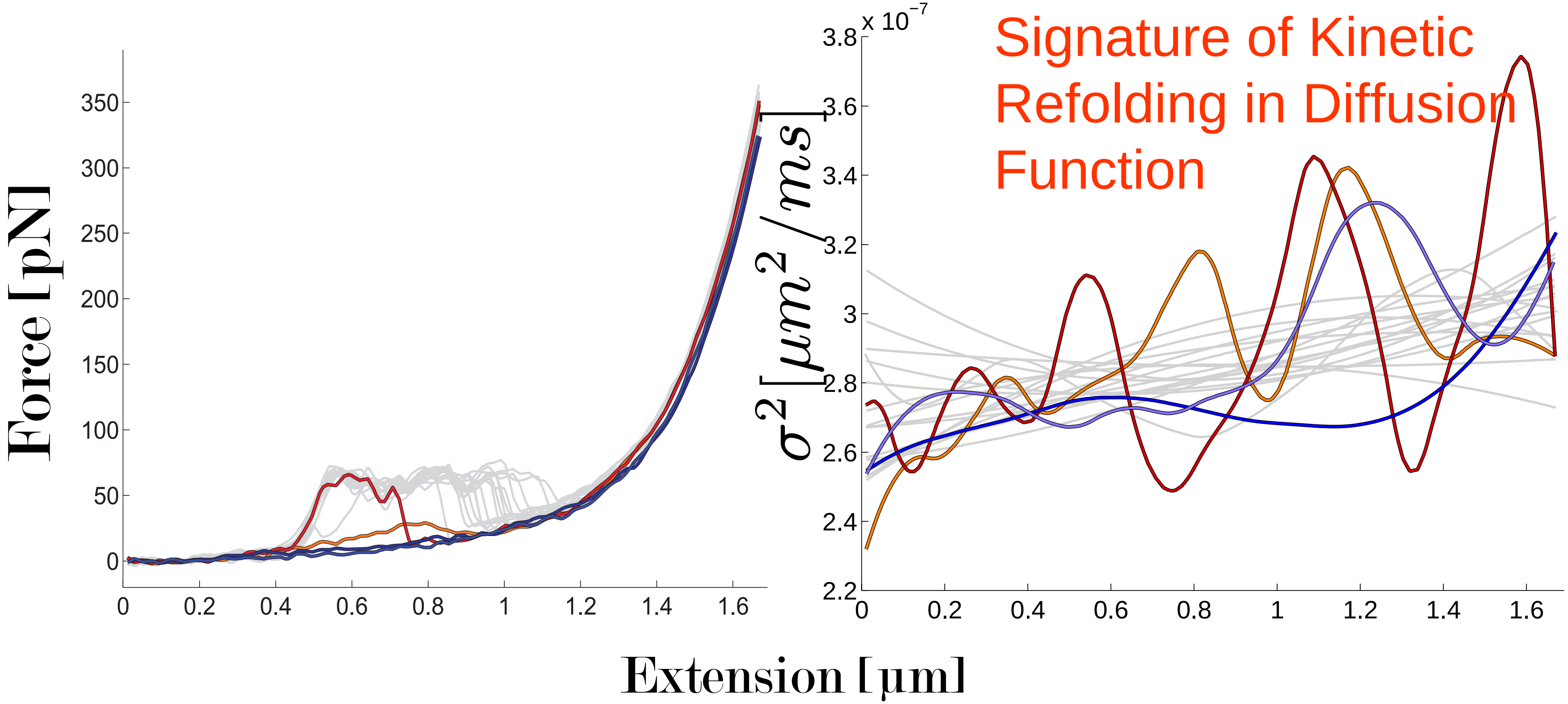
Colored Curves Correspond to "Failed" Re-annealing

[Recall Same Molecule Retained for 20+ Cycles]

Kinetic Signatures Encountered in "Refolding" DNA



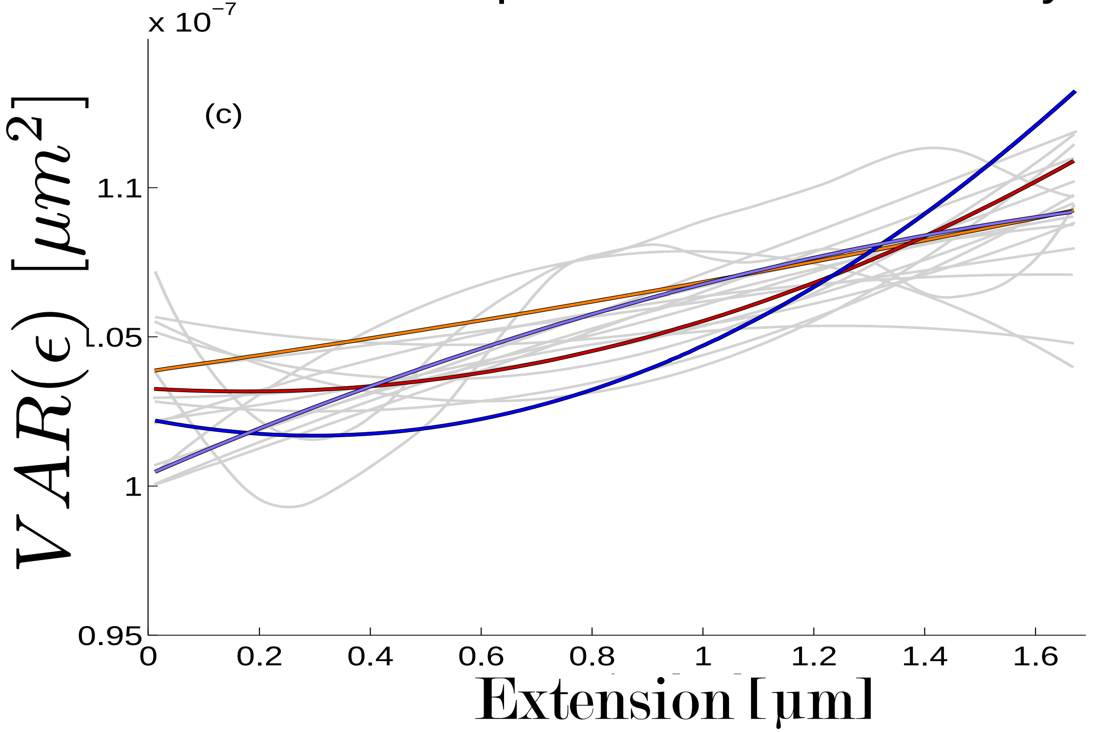
Kinetic Signatures Encountered in "Refolding" DNA



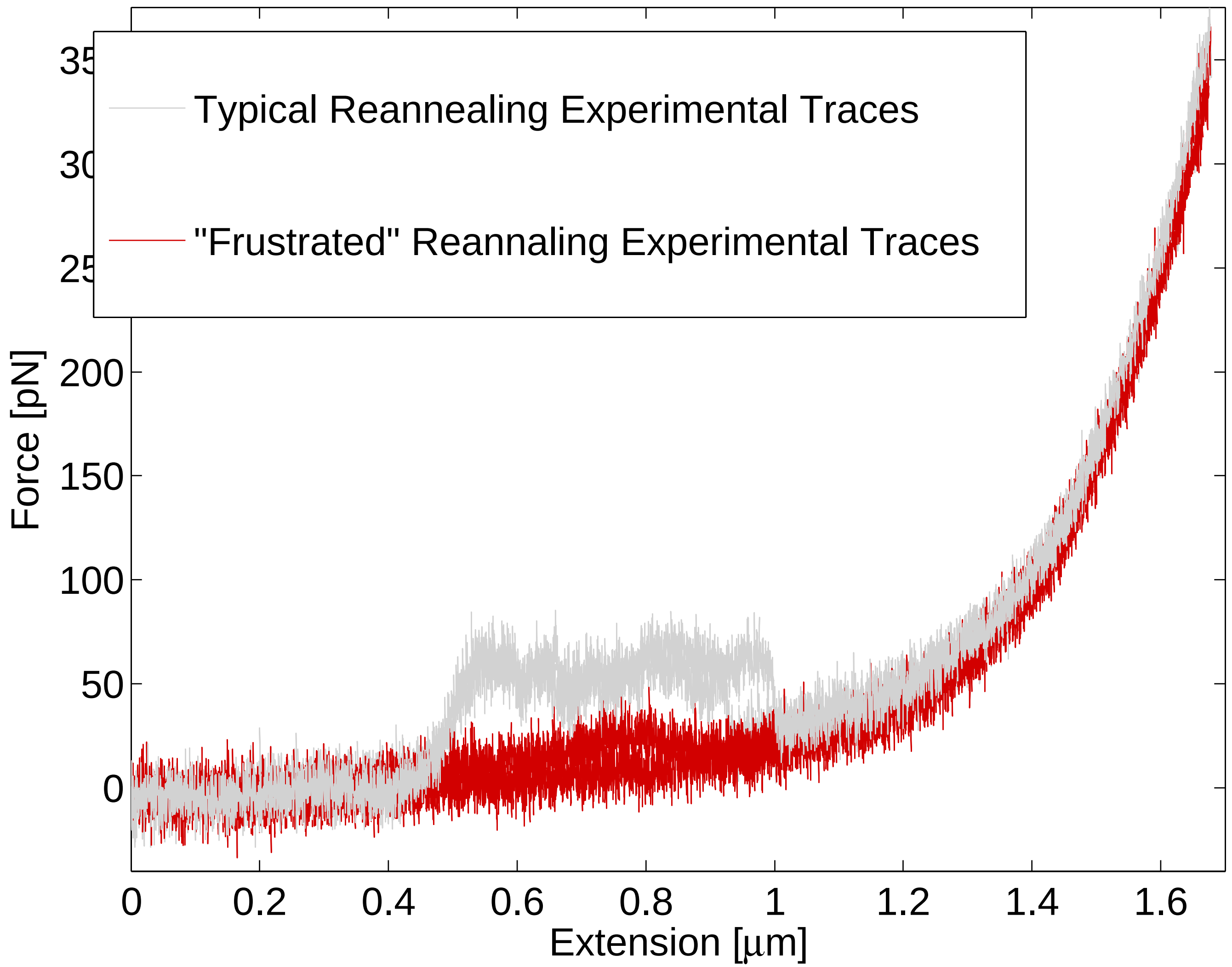
Decoupling Noise Sources

- In Experiments, Apparatus/Measurement Noise Dominates Frequently Sampled Data (Experimentalists Often Average Over Time)
- Nonstationary Nature of Time Series Complicates Statistical Analysis

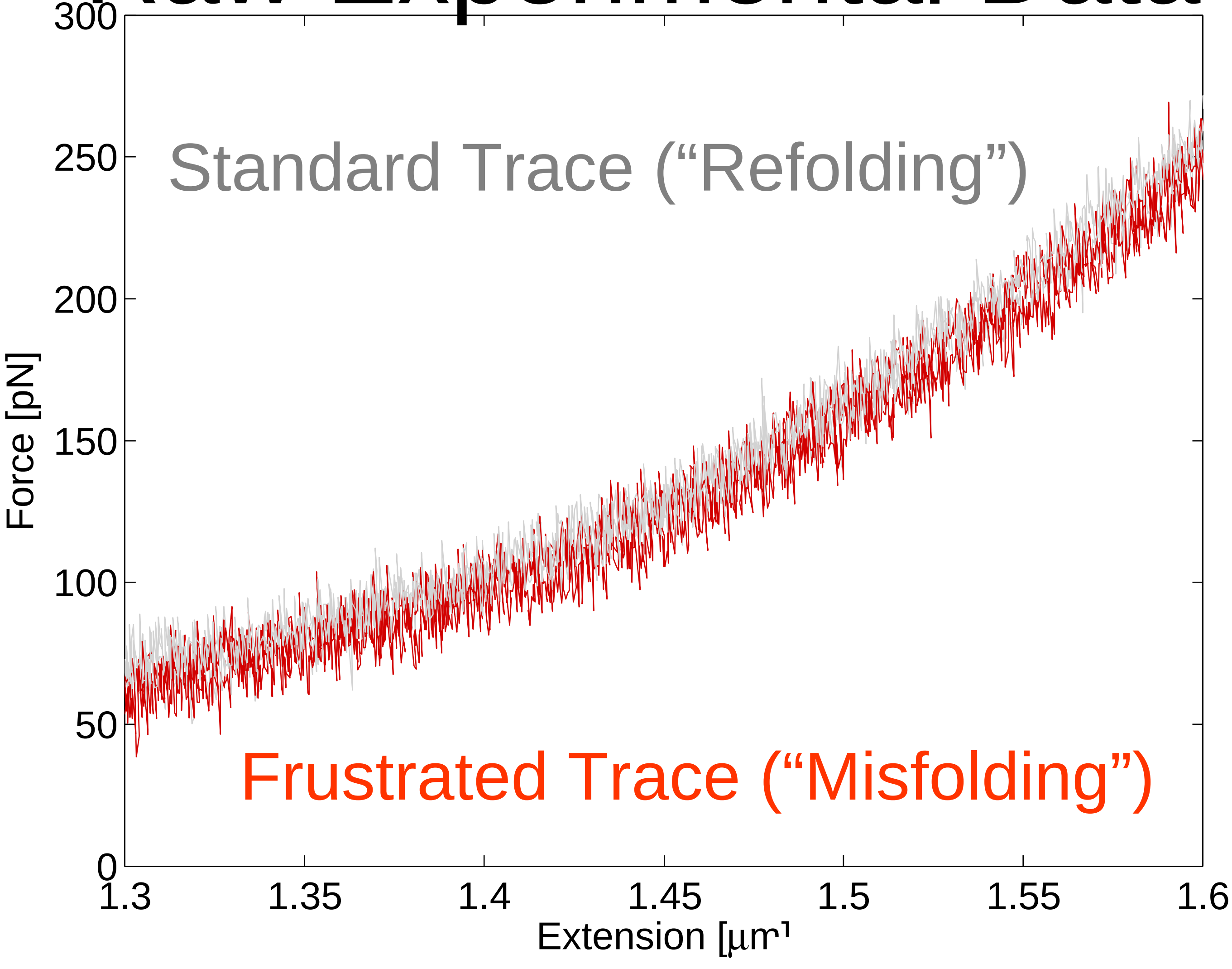
Without Structure of Diffusion Model Imposed "Standard" Variance Analysis Methods Fail to Distinguish Reannealing Failures.



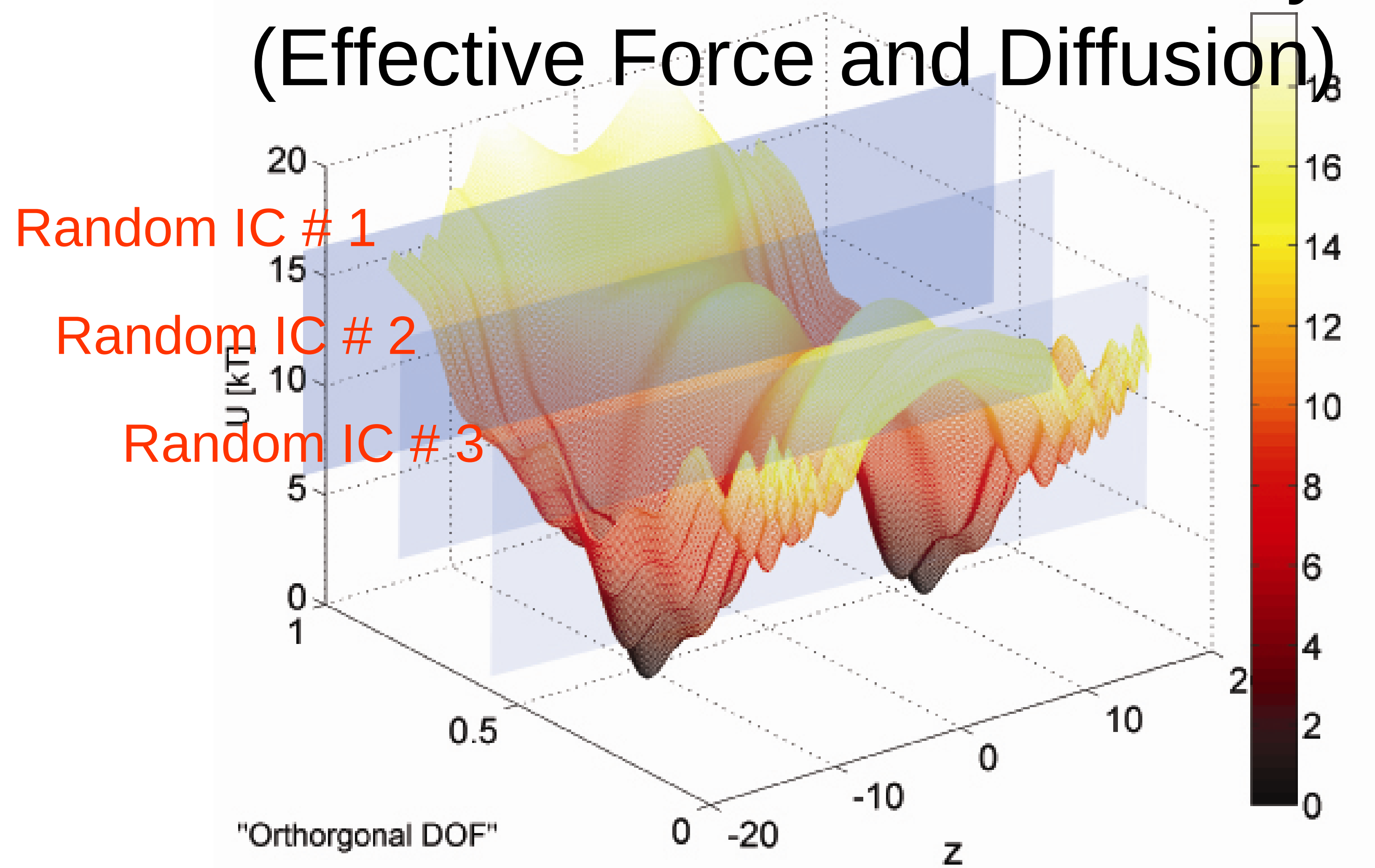
Raw Experimental Data



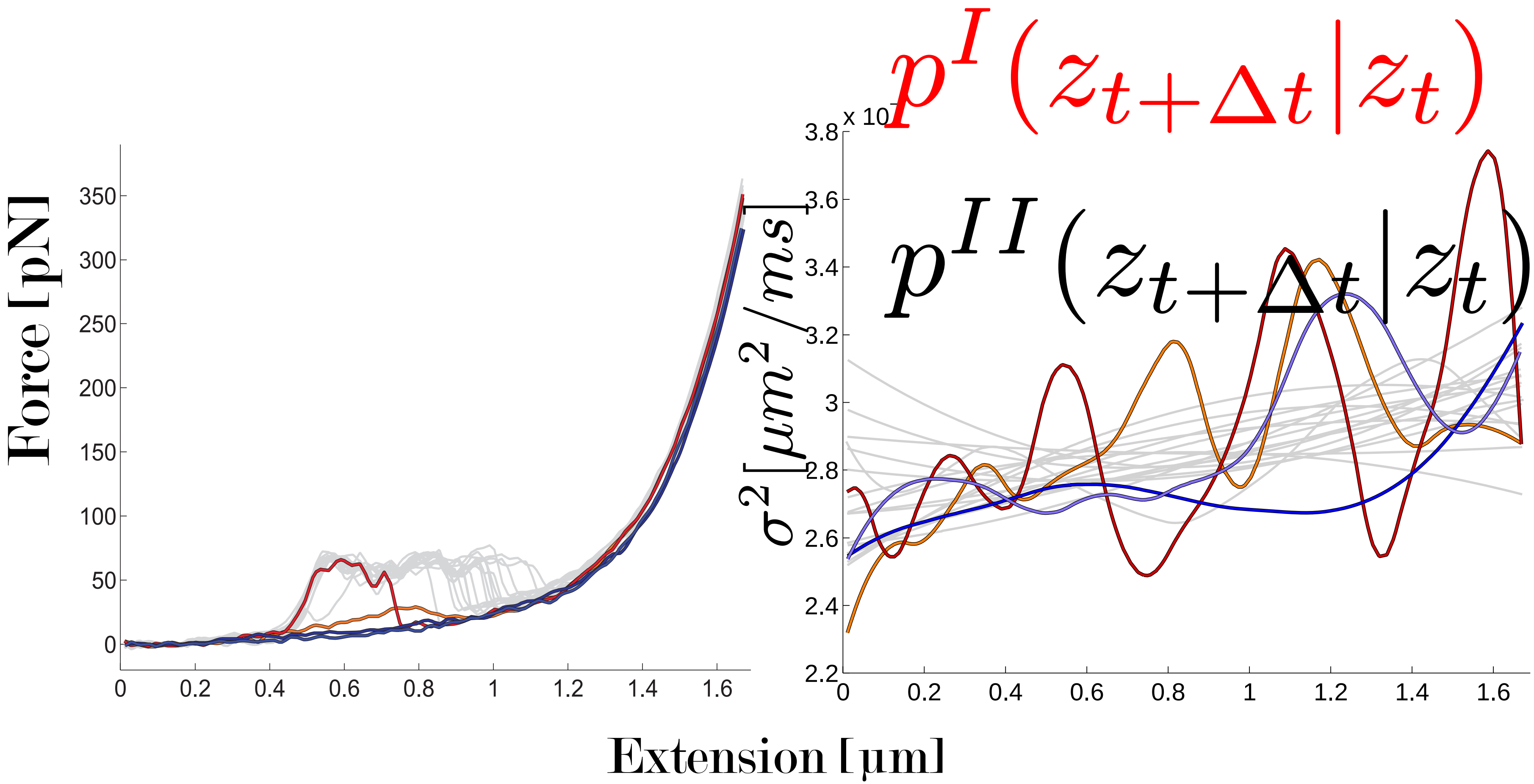
Raw Experimental Data



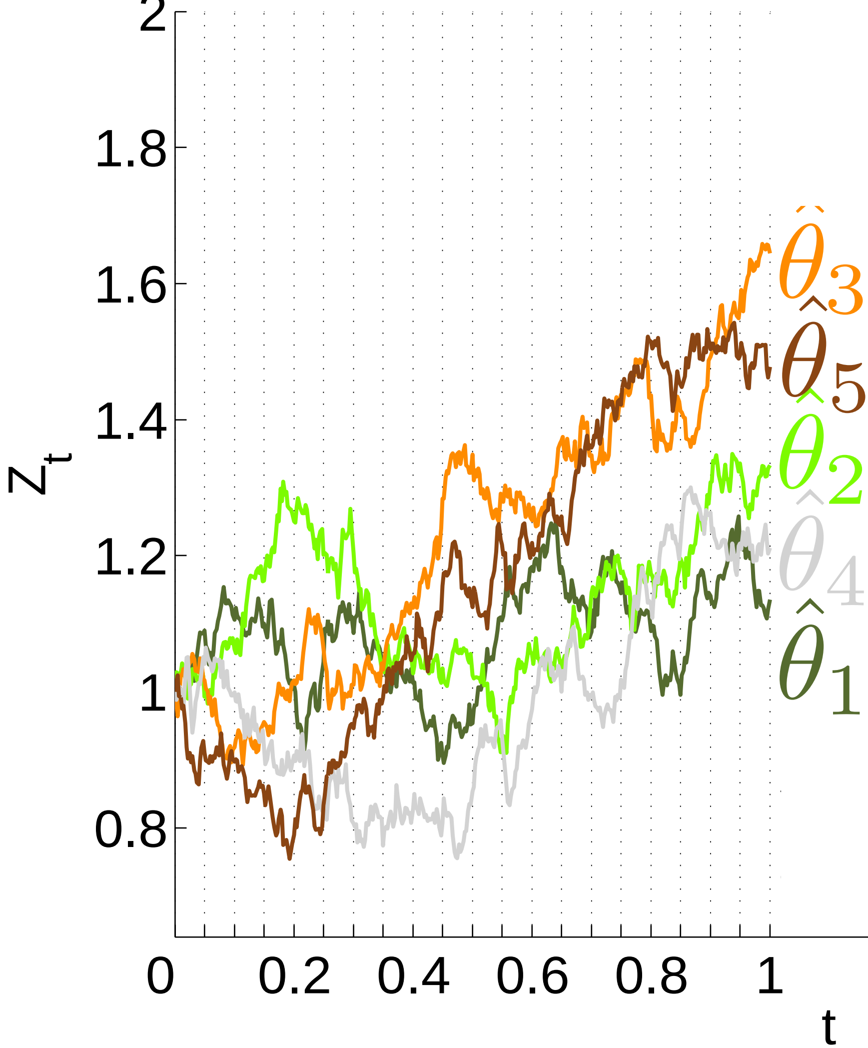
Random IC Modulates Stochastic Dynamics (Effective Force and Diffusion)



Rethink "Temporal Memory" vs. Unresolved Coordinates



Pathwise View

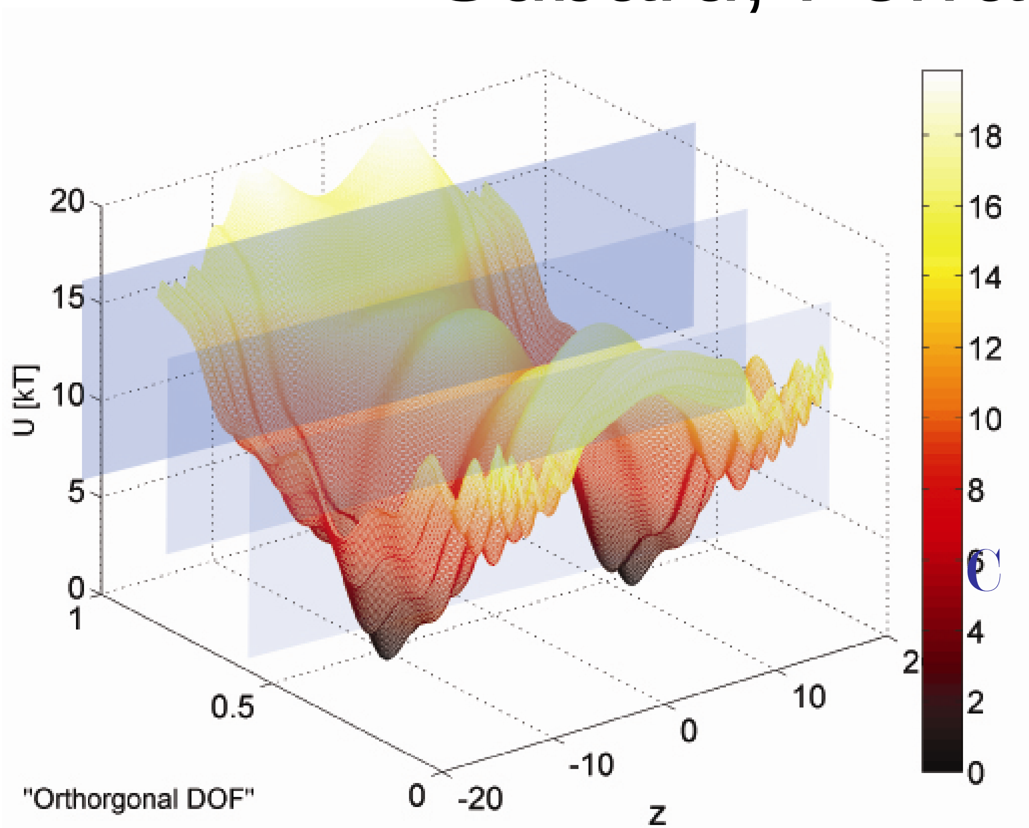


Ensemble Averaging Problematic Due to Nonergodic / Nonstationary Sampling

Collection of Models Another Way of Accounting for “Memory”

$$\langle \hat{\theta}_i \rangle$$

Same Sector, Different Players (e.g. Toyota, Subaru, Ferrari)



Globally Environment Has Common Shape, But Local Details Differ

Also Many More Factors Than Those Modeled Influence Shape of "Surface"

Why is This Potentially Relevant to Finance?

- Variation Between Different Sectors and Between Companies
- For a Relatively Fixed Environment/Operating Condition, Desirable to Understand Similarities and Differences Between Sectors and between Companies Same Sector

SDE Estimation and Inference: Bird's Eye View of Some Material To Come

(Time Domain Maximum Likelihood Methods)

Maximum Likelihood

For given model & discrete observations, find the maximum likelihood estimate:

$$\hat{\Theta} \equiv \max_{\Theta} \; p(z_0, \ldots, z_T; \Theta)$$

Special case of Markovian Dynamics:

$$\hat{\Theta} \equiv \max_{\Theta} \; p(z_0; \Theta) p(z_1 | z_0; \Theta) \ldots p(z_T | z_{T-1}; \Theta)$$

"Transition Density" (aka Conditional Probability Density)

Transition Density Approximations

SDE

$$dX_t = b(X_t, t; \Theta)dt + \sigma(X_t; \Theta)dW_t$$

Corresponding PDE (Backward Kolmogorov Eq.)

$$\frac{\partial f(X,t)}{\partial t} + b(X,t;\Theta)\frac{\partial f(X,t)}{\partial X} + \frac{1}{2}\sigma(X;\Theta)\sigma(X;\Theta)^T\frac{\partial^2 f(X,t)}{\partial X^2} = 0$$

Adjoint Solution →

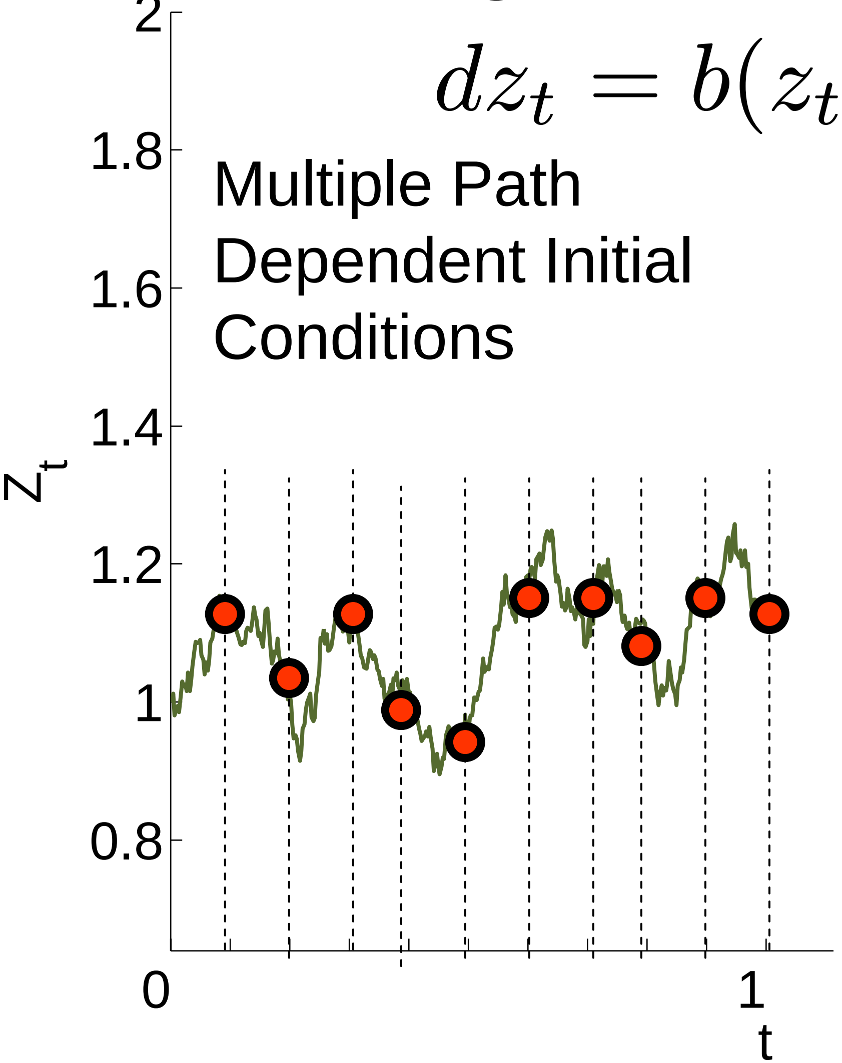
$$p(t+\Delta t, x_{t+\Delta t} | t, x_t; \Theta)$$

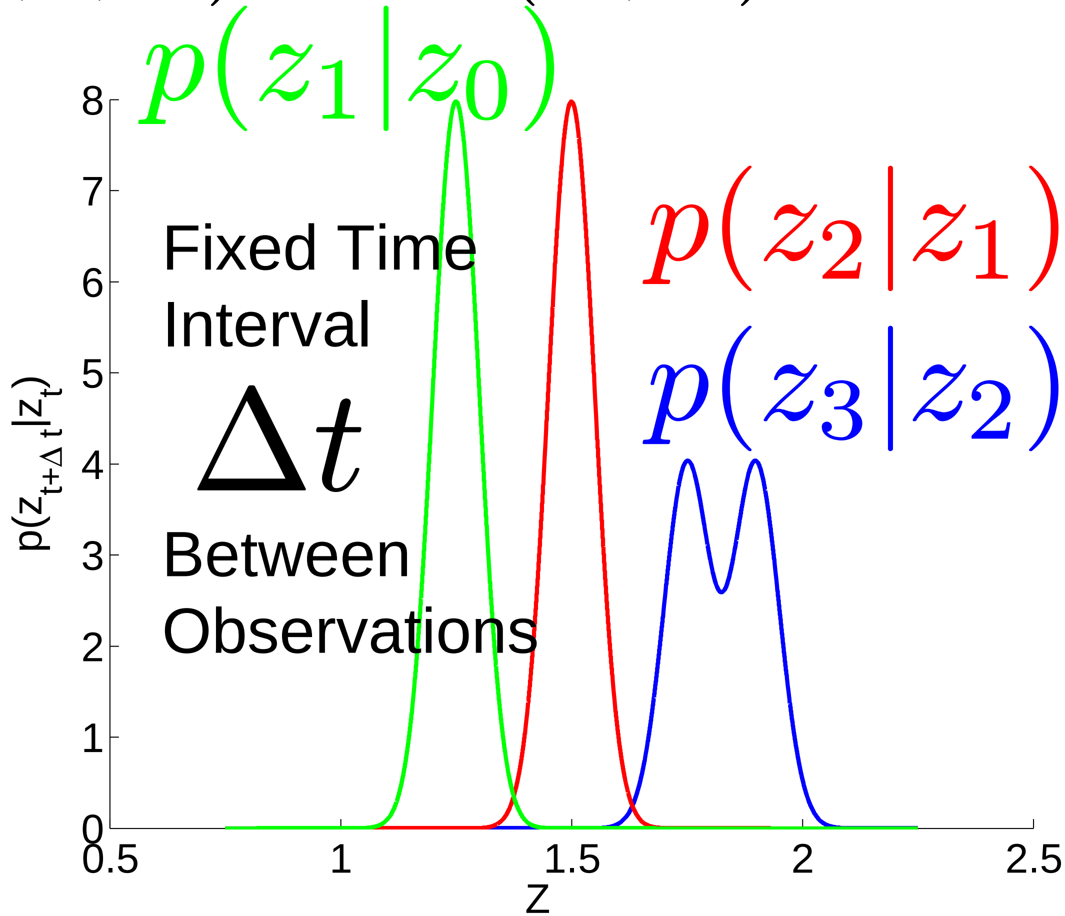
Solution Corresponding to Fokker-Planck Eq.

MANY approximations (both deterministic & stochastic)

Each Single SDE Connected to a PDE

$$dz_t = b(z_t, t; \Theta)dt + \sigma(z_t; \Theta)dW_t$$





“Velocity Correlation” and Levy Process Proxies

[Quantitative Criteria for Neglecting Inertia]

$$q_i(t+\delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma, t)dt$$

Is This A Levy Process at Our Time Resolution?

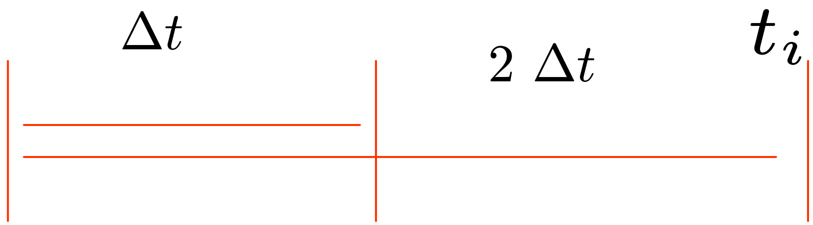
Any Model Can Be Fit, But Is It Statistically Justified Given Observational Data and an Assumed SDE Surrogate Structure?

Calderon, JCP (2007); Calderon & Chelli, JCP (2008); Calderon & Arora, JCTC (2009).

“Velocity Correlation” and Levy Process Proxies

[Quantitative Criteria for Neglecting Inertia]

$$q_i(t+\delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma,t)\,dt$$



$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

MANY O(fs) size “time steps” between

When is $0 \approx -\gamma v(\Gamma,t) + F(\Gamma,t)$

Calderon, JCP (2007); Calderon & Chelli, JCP (2008); Calderon & Arora, JCTC (2009).

Singular Perturbation / Model Reduction:

$$\frac{dq}{dt} = \frac{1}{m} p(\Gamma, t) = v(\Gamma, t)$$

$$m \frac{dv}{dt} = \frac{dp}{dt} = -\gamma v(\Gamma, t) + F(\Gamma, t)$$

$$0 \approx -\gamma v(\Gamma, t) + F(\Gamma, t)$$

Approximate Above By Ignoring "Inertia" and Modeling Momentum/Velocity as White Noise

$$dq_t = \frac{1}{\gamma} F(\Gamma_t, t)dt + \sigma(\Gamma_t)dW_t$$

Model Reduction

$$0 \approx -\gamma v(\Gamma, t) + F(\Gamma, t)$$

Approximate Above By Ignoring "Inertia" and Modeling Momentum/Velocity as a Levy Process ("White Noise in Stat. Phys. Terminology")

$$dq_t = \frac{1}{\gamma} F(\Gamma_t, t)dt + \sigma(\Gamma_t)dW_t$$

If One Waits "Long Enough", Hope is Fast Scale Noise Becomes Statistically Indistinguishable from a Levy Process

$$q_i(t + \delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma, t)dt$$

Δt

$2\ \Delta t$

$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

$$\Downarrow$$

$$(q_0, q_j, q_{2j}, \ldots)$$

$$q_i(t + \delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma, t)dt$$

Δt

$2\ \Delta t$

$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

$$(q_0, q_j, q_{2j}, \ldots)$$

Estimate SDE Parameter

$$\hat{\theta}$$

$$q_i(t+\delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma, t)\,dt$$

Δt

$2\ \Delta t$

$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

$$(q_0, q_j, q_{2j}, \ldots)$$

Estimate SDE then Use Inferred Parametric Transition Density to Apply Probability Integral (Rosenblatt) Transform

$$p(q_{2j} | q_j; \hat{\theta})$$

$$q_i(t+\delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma, t)\,dt$$

Δt $\quad$ $2\ \Delta t$

$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

$$(q_0, q_j, q_{2j}, \ldots)$$

Use SDE Transition Density to Transform

i.i.d. if Null Correct

$$(z_0, z_j, z_{2j}, \ldots)$$

Test Statistic (Null Dist. Computable in Finite Samples)

$$q_i(t+\delta t) = q_i(t) + \int_{t_i}^{t_i+\delta t} v(\Gamma, t)\,dt$$

Δt

$2\ \Delta t$

$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

$$(q_0, q_j, q_{2j}, \ldots)$$

$$z_{2j} = \int_{-\infty}^{q_{2j}} p(q|q_j;\theta_0)\,dq$$

i.i.d. if

Null Correct

$$(z_0, z_j, z_{2j}, \ldots)$$

Test Statistic (Null Dist. Computable in Finite Samples)

$$q_i(t + \delta t) = q_i(t) + \int_{t_i}^{t_i + \delta t} v(\Gamma, t)\, dt$$

Δt $\quad 2\ \Delta t$

$$(q_0, \ldots, q_j, \ \ldots, q_{2j}, \ \ldots, \ q_T)$$

$$(q_0, q_j, q_{2j}, \ldots)$$

Use “Omnibus” Test:

Simultaneously Check Shape and INDEPENDENCE Assumption

i.i.d. if

Null Correct

$$(z_0, z_j, z_{2j}, \ldots) \longrightarrow Q$$

Utility of Quantitative Goodness-of-Fit Tests in Non-stationary Regime:

Some Statistical Physics Folklore: Given Enough "Averaging Time" a Second Order Dynamical System , e.g.,

$$\frac{d^2\Gamma}{dt^2} + f(\Gamma)\frac{d\Gamma}{dt} = g(t)$$

Can be Approximated by a "First Order" System, e.g.,

$$\frac{d\Gamma}{dt} = h(\Gamma, t)$$

Goodness-of-Fit Using Accurate Transition Densities

Hong & Li, *Rev. Financial Studies*, **18** (2005). Ait-Sahalia, Fan, Peng, *JASA* (in press)

Both Tests Applied in: Calderon & Arora, *JCTC* **5** (2009).

$$\Delta t >> \tau \qquad \tau = \frac{m}{\gamma}$$

"Effective Mass" (m)

"Damping/Friction" (γ)

Above Problem Relatively Under Control (Though More Powerful/Specific Tests Would Be Nice).

Harder Inference Problem Detecting When Slowly Evolving Latent Process Substantially Modulates Dynamics

Acknowledgements

- Nolan Harris, Harry Chen, Ching-Hwa Kiang
  (DNA AFM Experimental Data)
- Dennis Cox, Josue Martinez, Raymond Carroll
  (Discussions on Local Estimation/Penalized Spline Issues)
- Dan Sorensen
  (Discussions on Numerics and Computation)
- Yannis Kevrekidis
  (Sparking Interest in Estimation)

Financial Support:
NIH Nanobiology Training Grant (T90 DK070121-04) and by the Director, Office of Advanced Scientific Computing Research of the D.O.E.

Fitting and Testing Nonlinear SDEs

Statistical Inference and Fitting with Approximate Conditional Densities

Calderon, *SIAM Mult. Mod. & Sim.* **6** (2007).

Fitting and Testing Time Inhomogeneous Diffusions (w/ State Dependent Noise)

Calderon, *J. Chem. Phys.* **126** (2007).

Calderon & Chelli, *J. Chem. Phys.* **128** (2008).

Calderon, Janosi & Kosztin, *J. Chem. Phys.* **130** (2009).

Calderon, *J. Phys. Chem. B* **in press** (2010).

Modeling Experimental Time Series (Measurement Noise + Thermal Noise)

Calderon, Chen, Lin, Harris, & Kiang, *J. Phys.: Condensed Matter* **21** (2009).

Calderon, Harris, Kiang, & Cox, *J. Phys. Chem. B* **113** (2009).

Calderon, Harris, Kiang, & Cox, *J. Mol. Recognit.* **22** (2009).

Global Models via Penalized Splines

Calderon, Martinez, Carroll, & Sorensen, submitted (2009).

Calderon, Martinez, Carroll, & Sorensen, submitted (2009).