15-869
Lecture 12
Body Pose and Dynamics

Yaser Sheikh

Human Motion Modeling and Analysis

Fall 2012

Debate Advice

Affirmative Team

- **Thesis Statement**: Every (good) paper has a thesis. What is the most provocative statement the paper is trying to make?
- **Spill the beans early**: Explain the key insight of the paper as soon as possible. Like on Slide 1.
- **Teaser**: Show an example result first so that it's clear to everyone what the goal is. Also mention the input, clearly.
- **Equations**: Explain them carefully or don't include them.

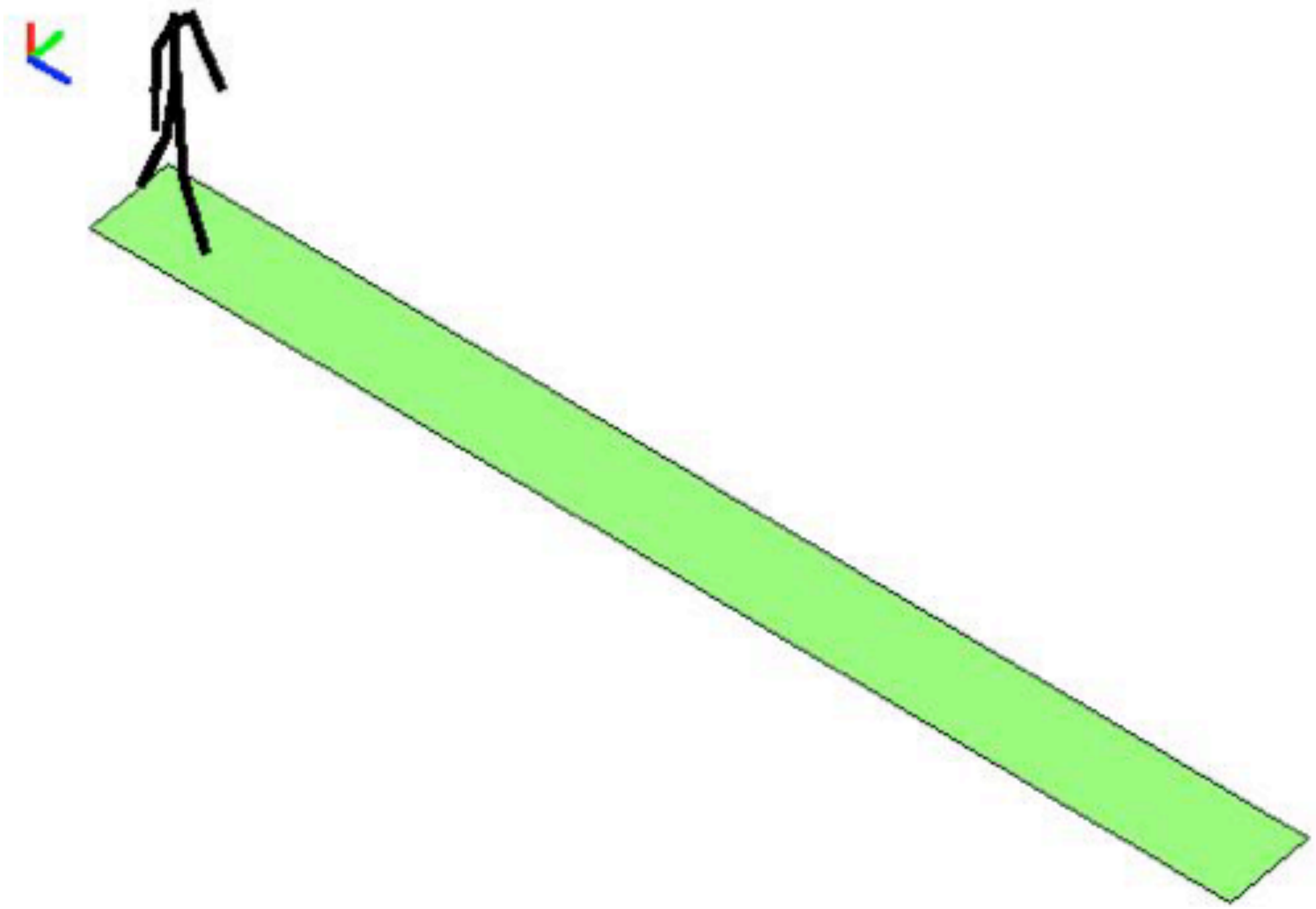
Debate Advice

Improvement Team

- **Examine the Thesis**: Is it sound? Does it over-reach? Is it practically useful?
- **State a Core Objection**: State a Core Objection early in the process.
- **Explain limitations**: Usually in the paper. Assumptions, model restrictions, computation, etc.

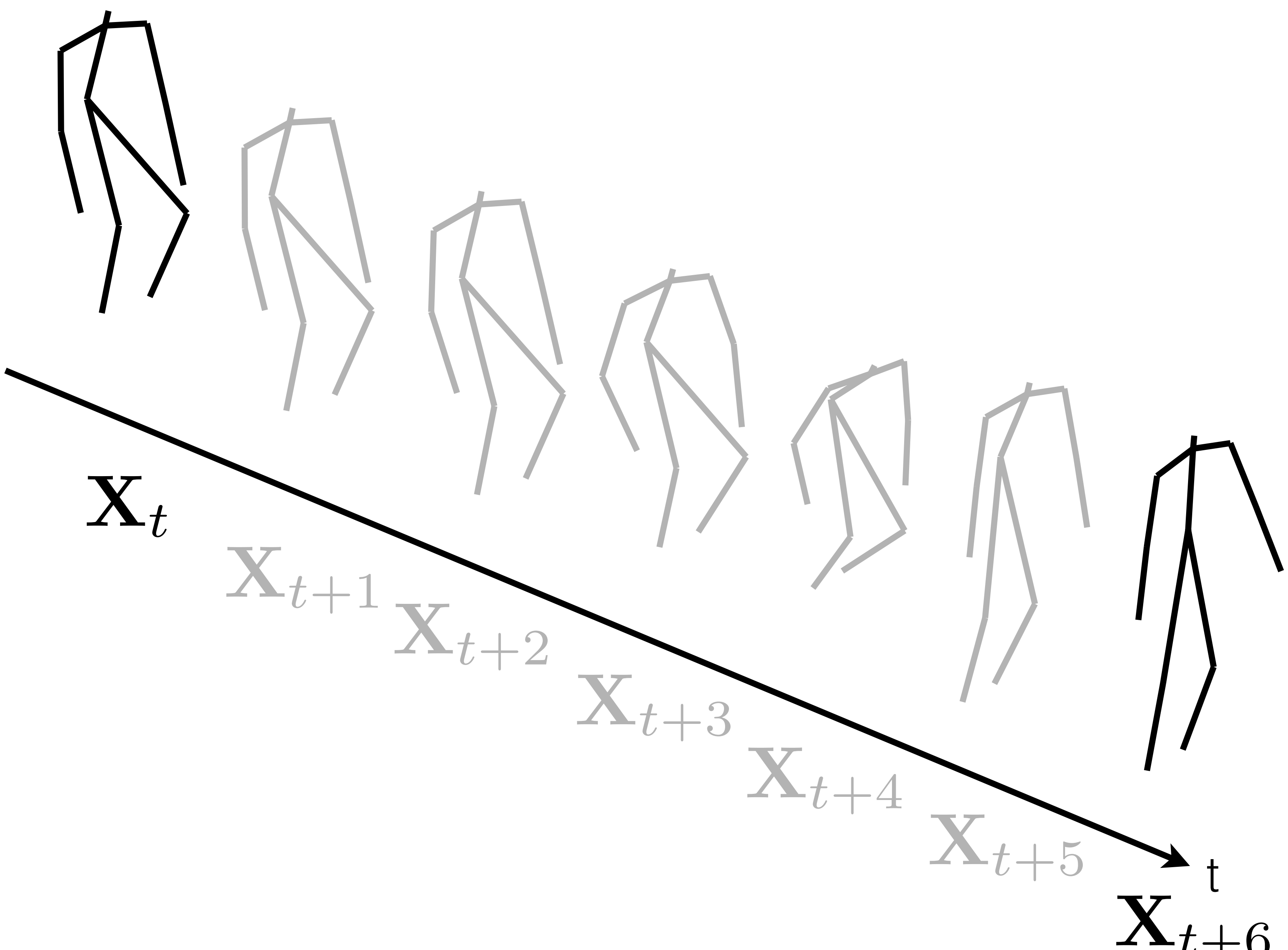
Human Motion

How can we model human motion?



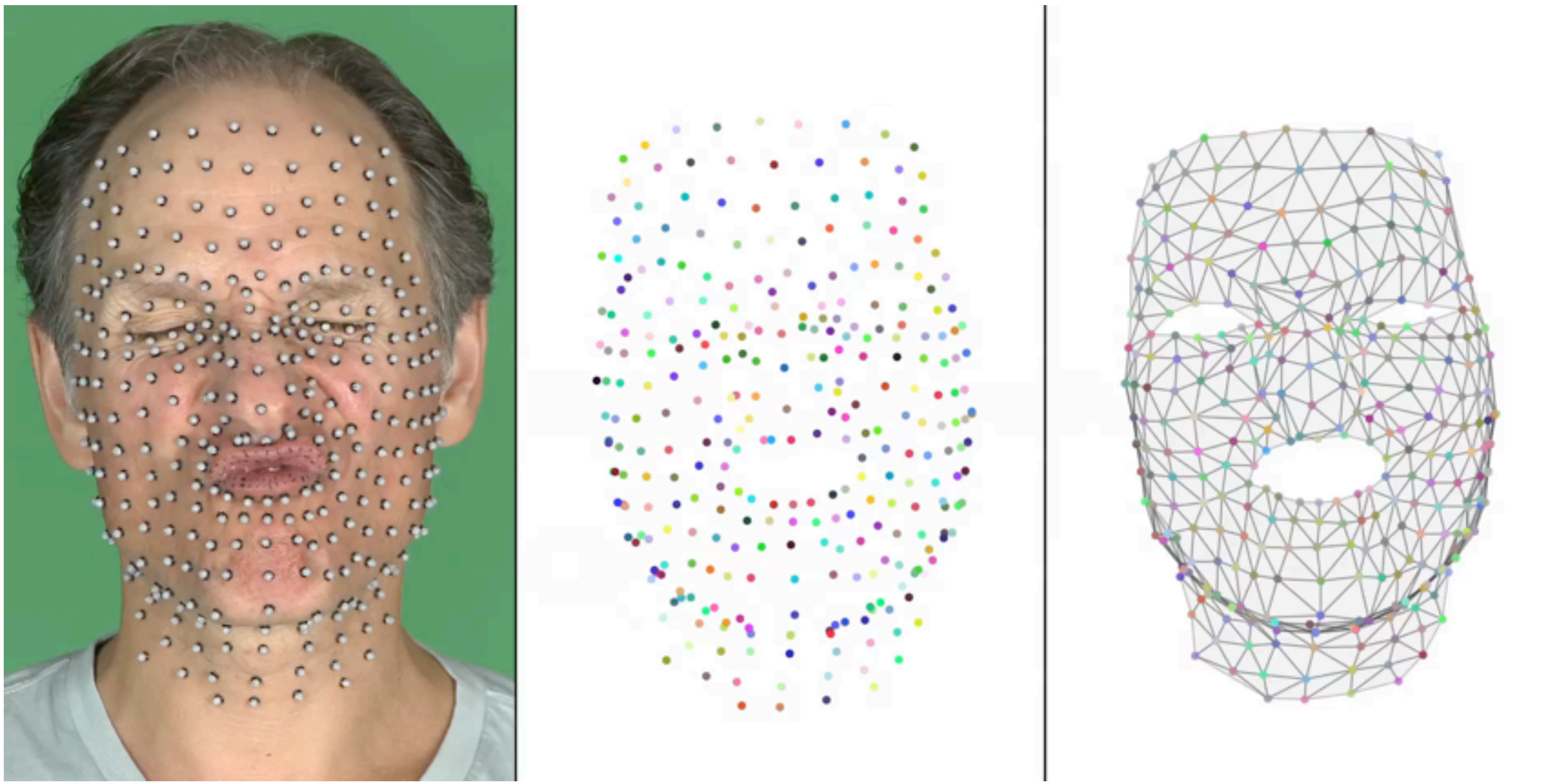
Why Model Motion?

Keyframing



Why Model Motion?

Labeling

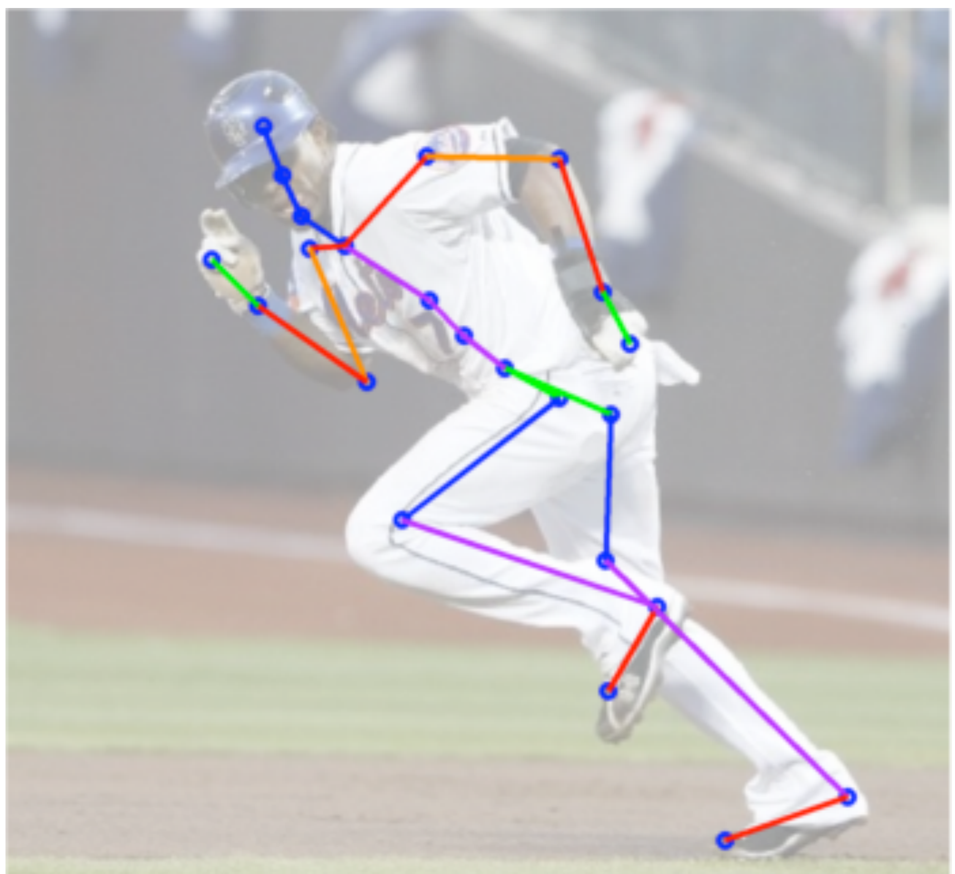


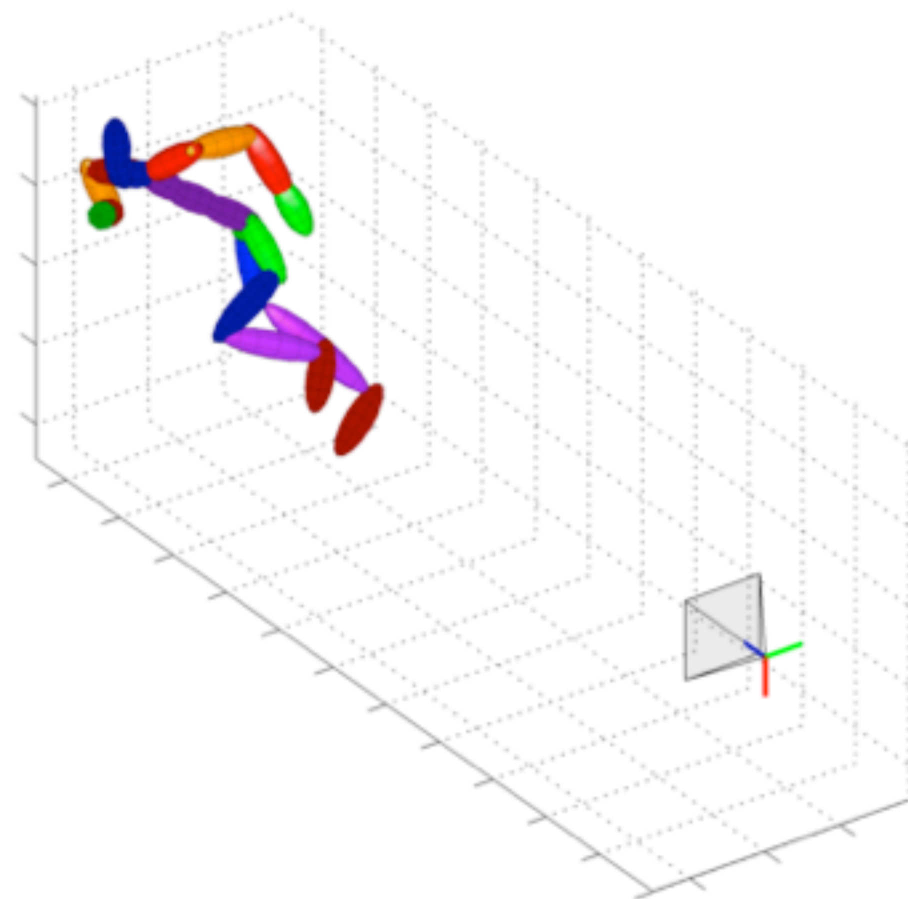
$\mathbf{X}_t^i$: Point Position at time t

$\mathbf{Z}_t^i$: Point Label at time t

Why Model Motion?

3D Reconstruction

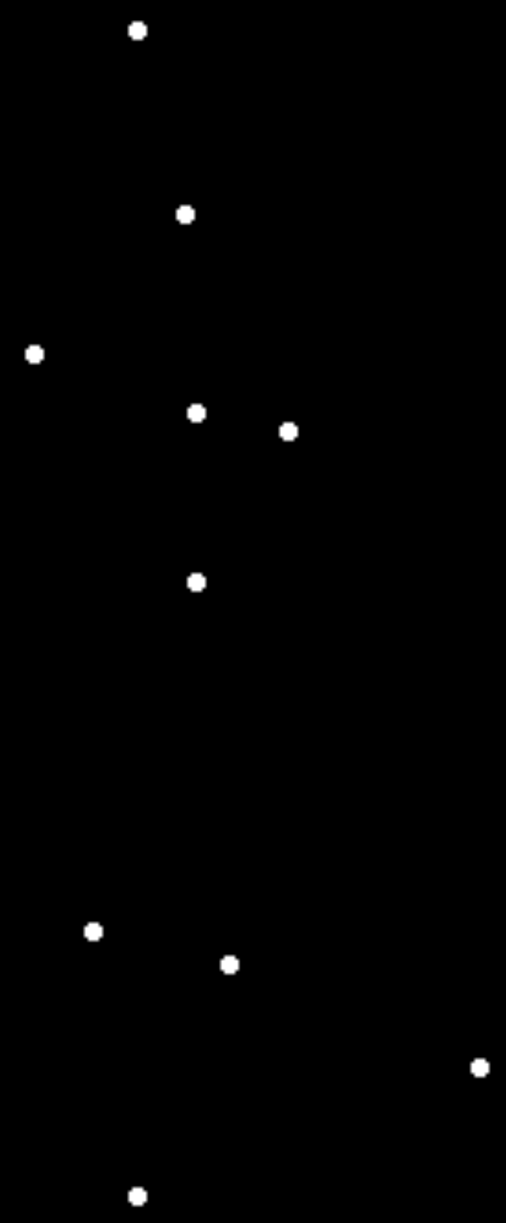




$$\mathbf{X}_t = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & & X_P \\ Y_1 & Y_2 & Y_3 & Y_4 & \cdots & Y_P \\ & & & & & \end{bmatrix}$$

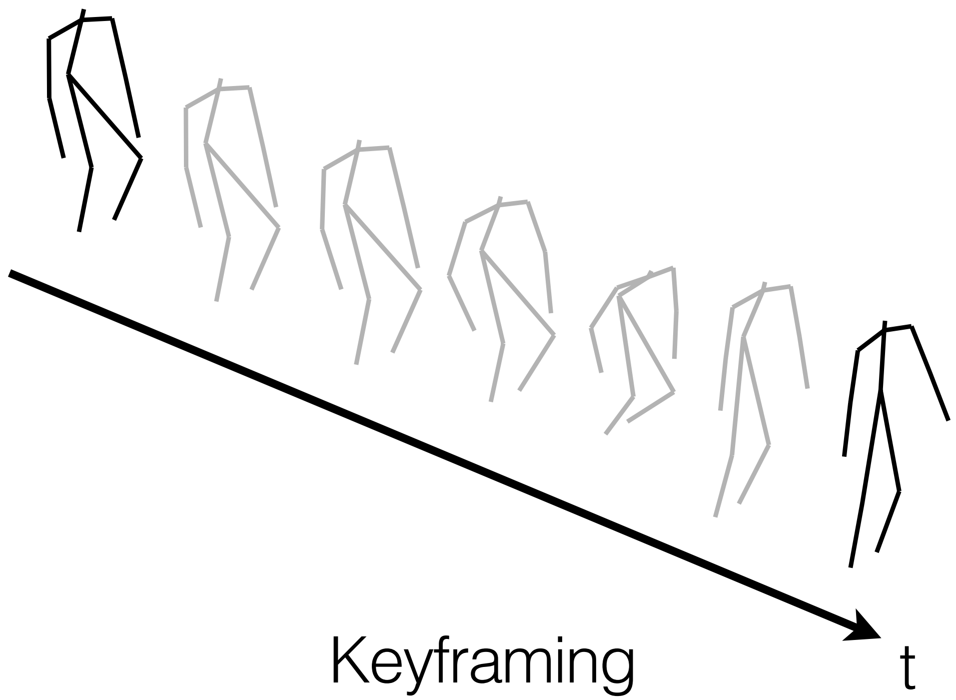
Why Model Motion?

Action Recognition

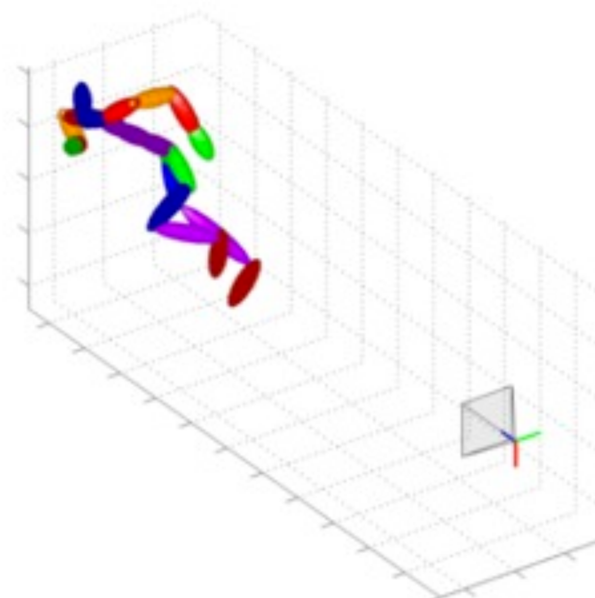


Knowns and Unknowns

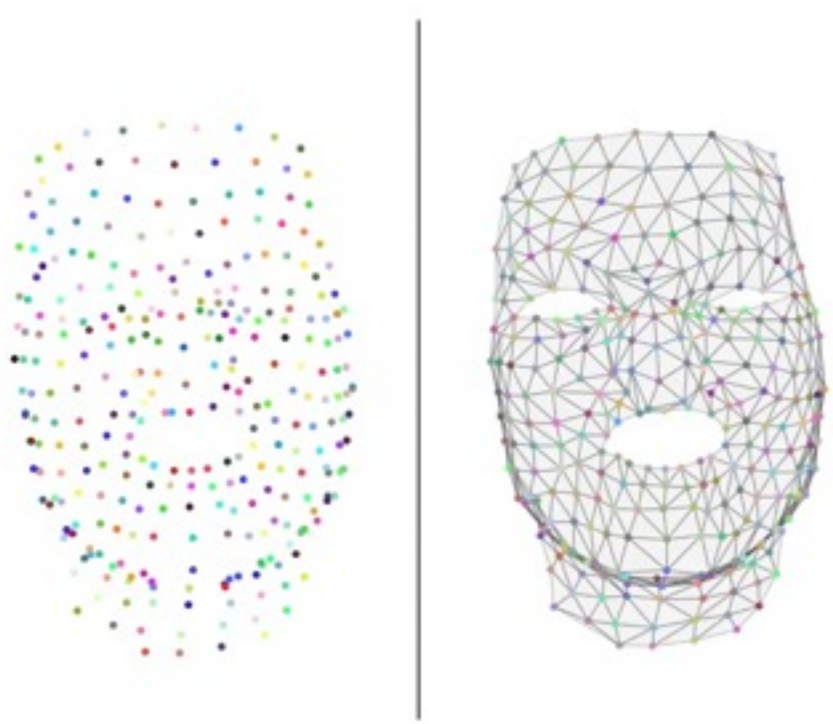
There is structure in human pose and motion



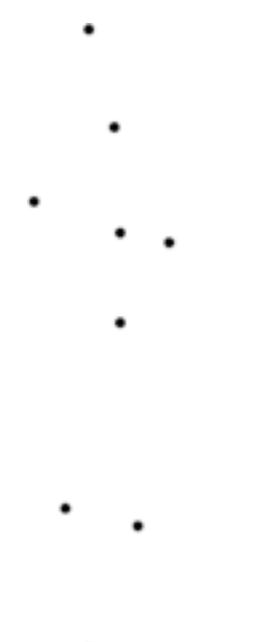
Keyframing



Reconstruction



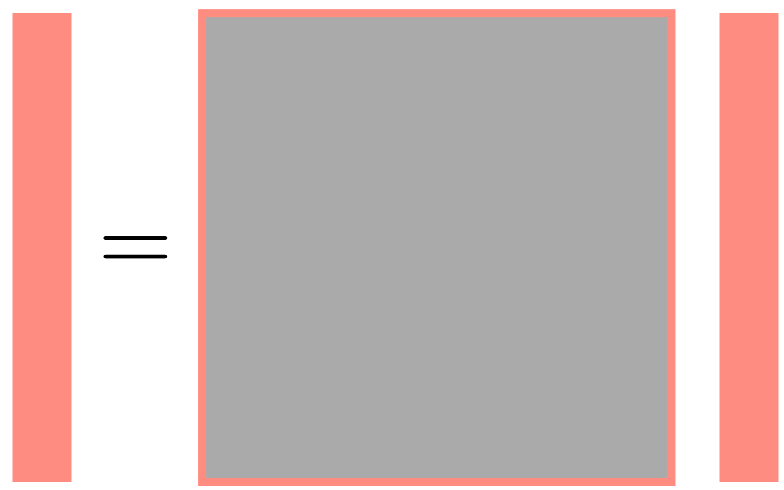
Labeling



Recognition

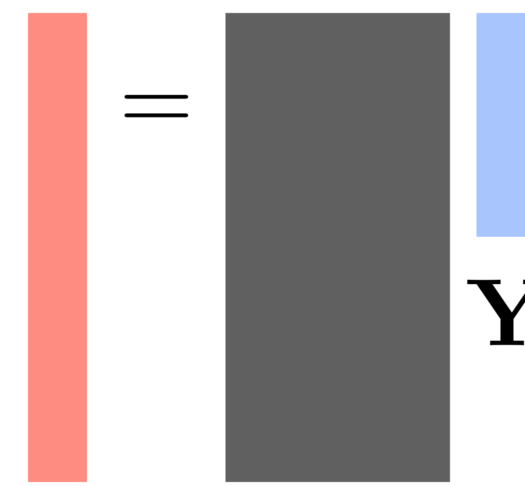
Lecture in One Slide

Notation



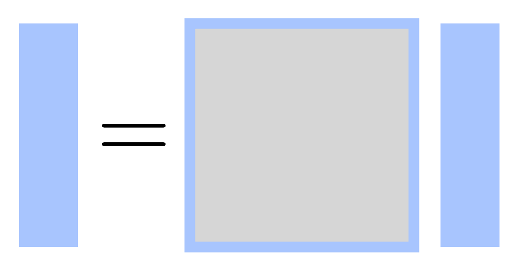
$$\mathbf{X}_{t+1} = \mathbf{D}\,\mathbf{X}_t$$

Dynamics



$$\mathbf{X}_t = \mathbf{B}\,\mathbf{Y}_t$$

Latent Variable Model



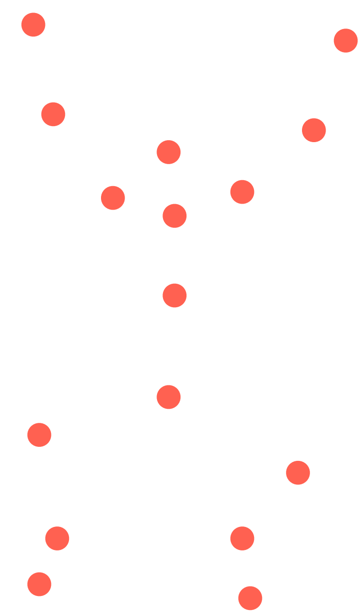
$$\mathbf{Y}_{t+1} = \mathbf{G}\,\mathbf{Y}_t$$

Latent Dynamics

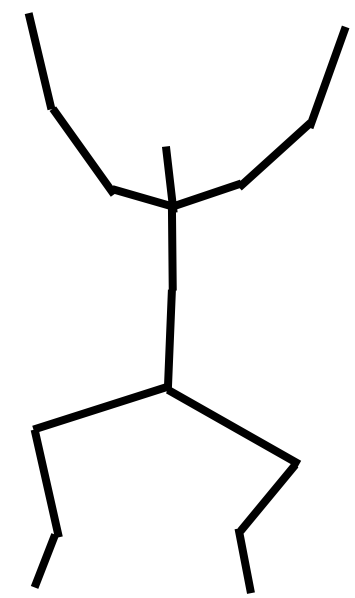
Representing Pose

Configuration Space

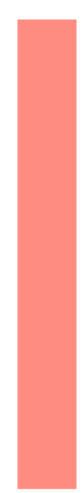
D Features



D Points



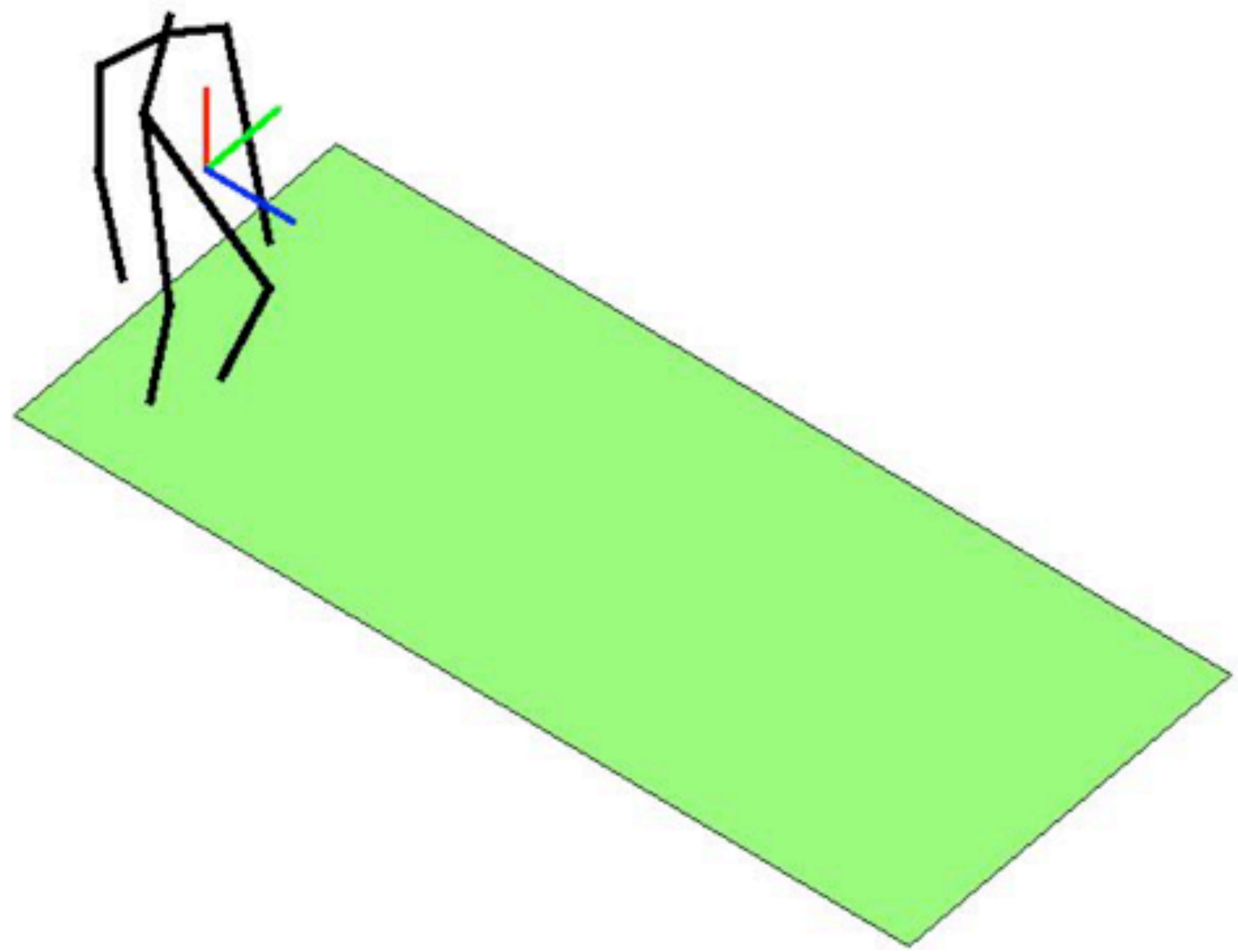
D Joint Angles



Observation

$$\mathbf{X}_t \in \mathbb{R}^D$$

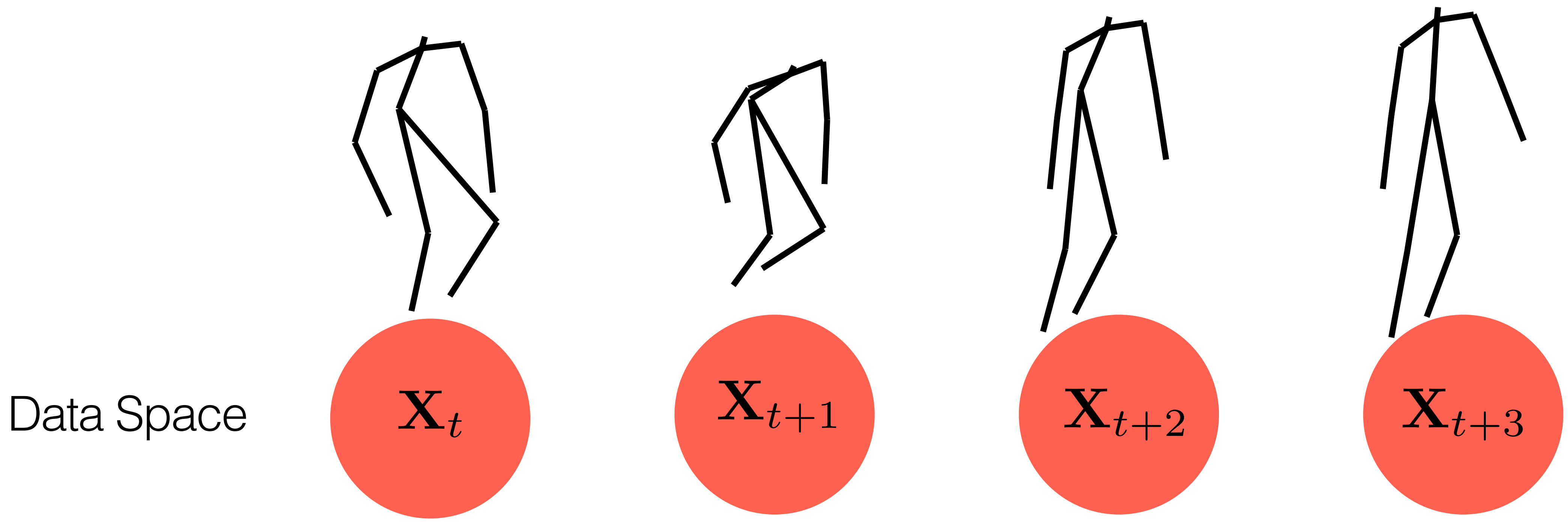
The Space of Actions



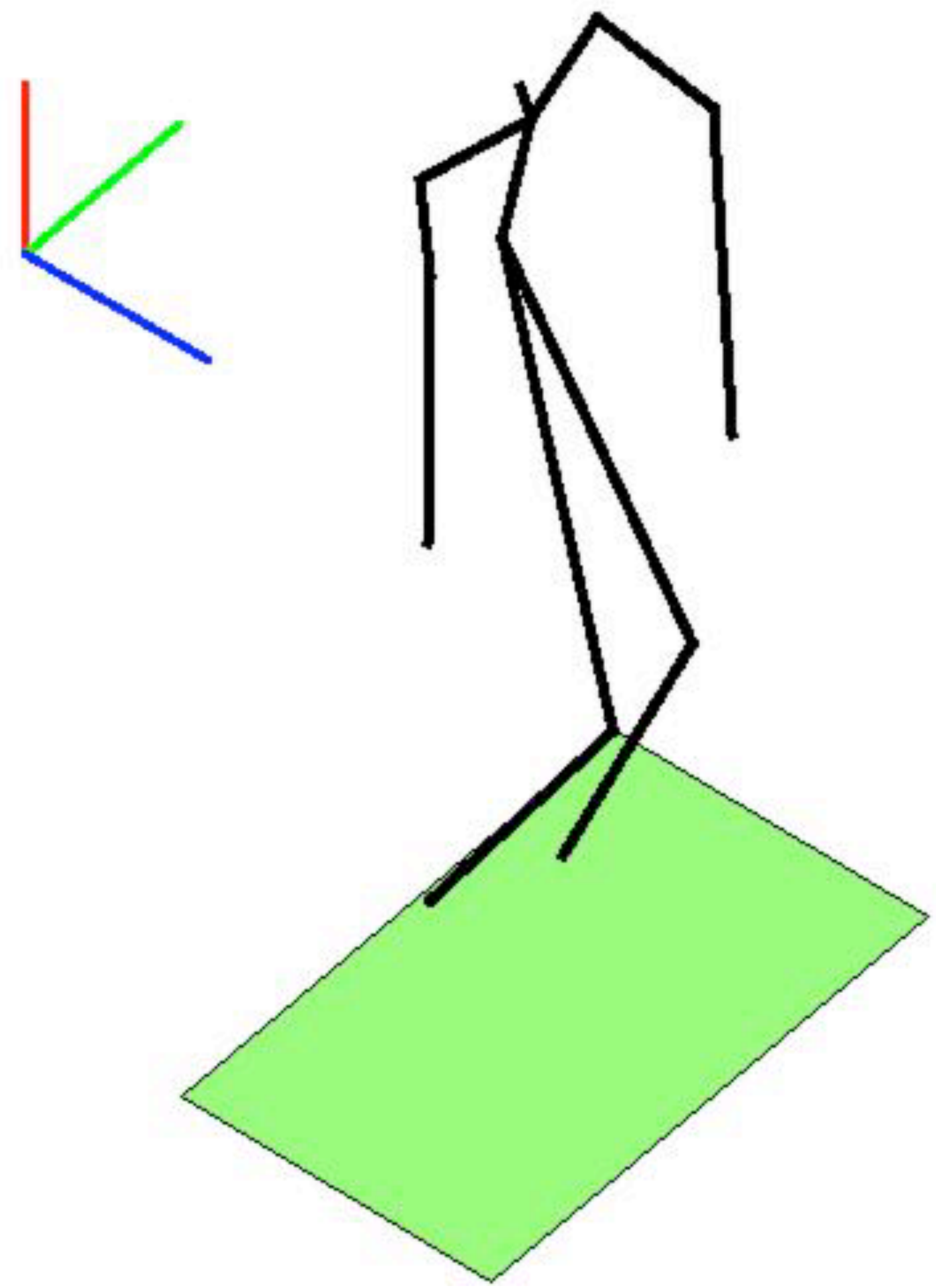
$$p(\mathbf{X}_1, \cdots, \mathbf{X}_F)$$

Human Motion

Frame-wise Independent



$$p(\mathbf{X}_1, \cdots, \mathbf{X}_F) = p(\mathbf{X}_1)p(\mathbf{X}_2)\cdots p(\mathbf{X}_F)$$



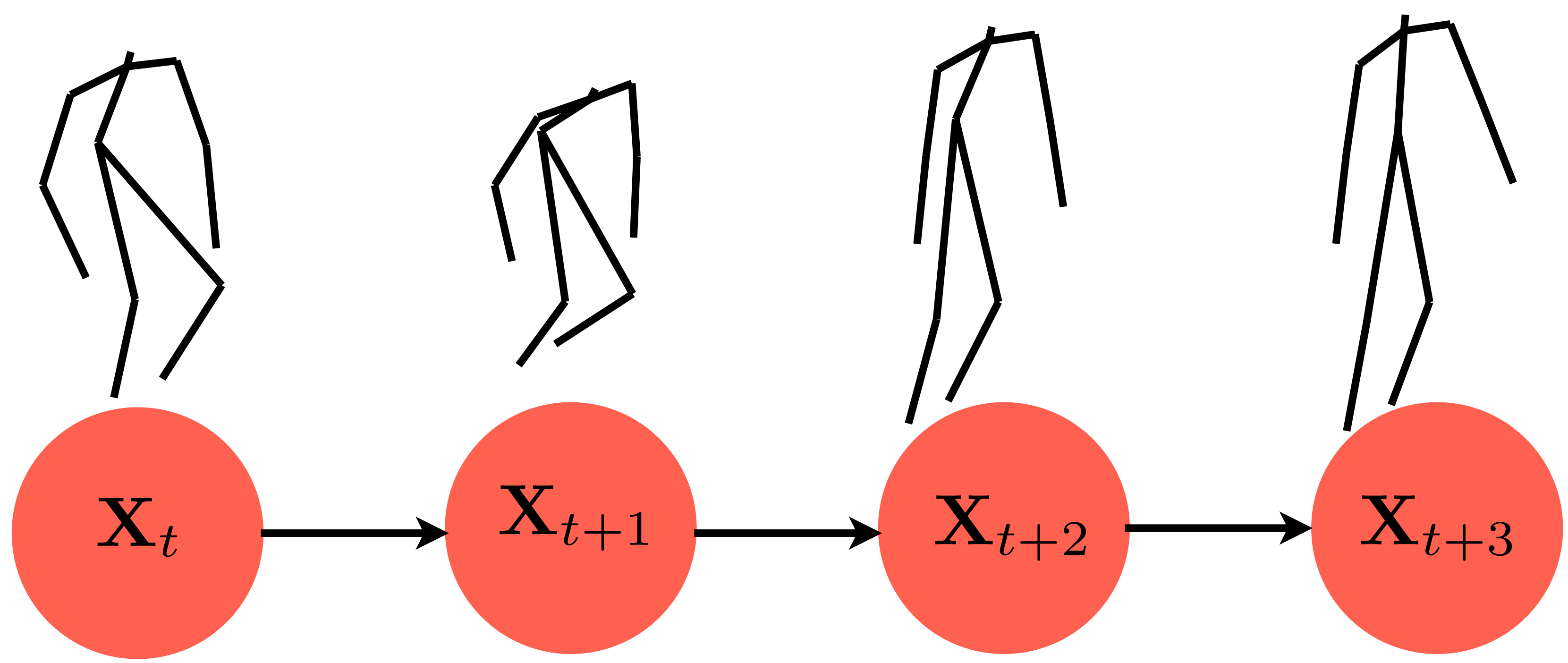
$$p(\mathbf{X}_1)p(\mathbf{X}_2)\cdots p(\mathbf{X}_F)$$

$$p(\mathbf{X}_t) = \mathcal{N}(\mathbf{X}_t|\mu, \Sigma)$$

Autoregressive Models

First-order Markov Model

Data Space



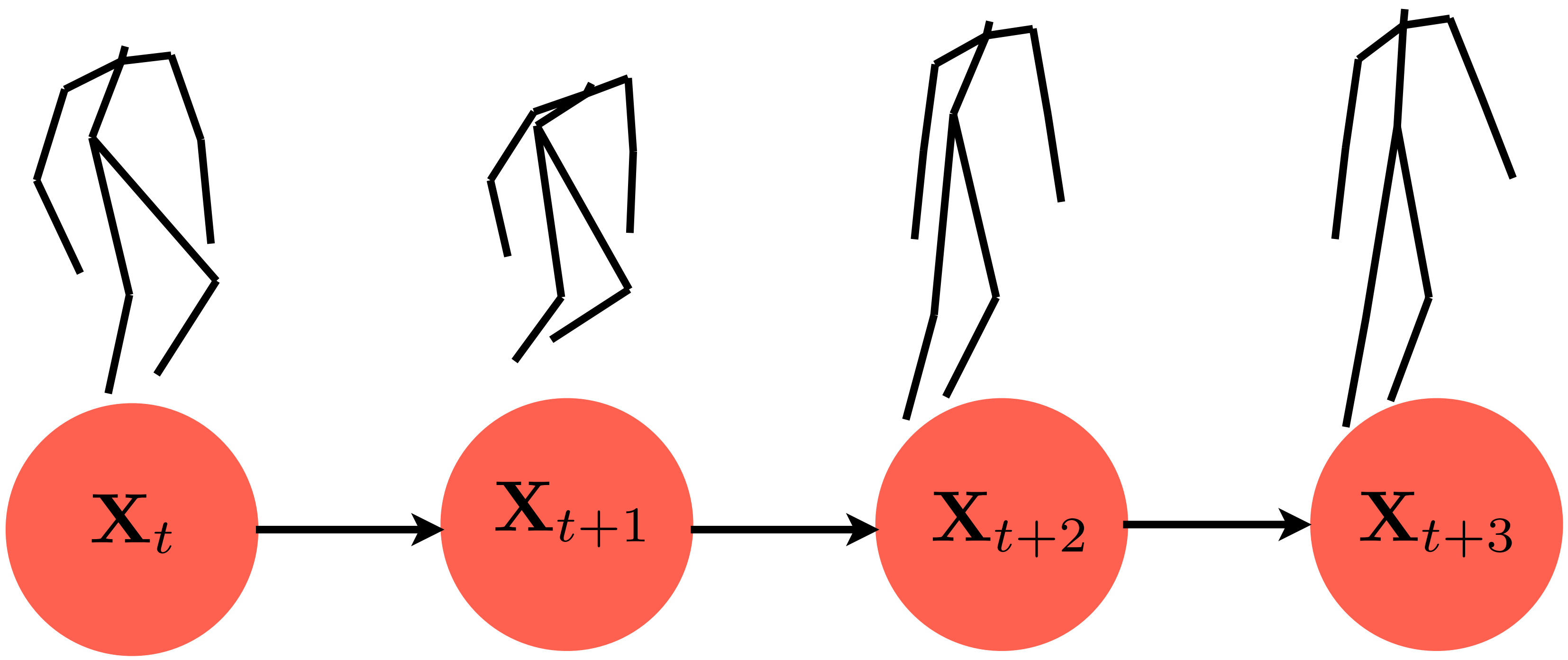
$$p(\mathbf{X}_1, \cdots, \mathbf{X}_F) = p(\mathbf{X}_1) \prod_{t=2}^{F} p(\mathbf{X}_t | \mathbf{X}_{t-1})$$

$$p(\mathbf{X}_t | \mathbf{X}_1, \cdots, \mathbf{X}_{t-1}) = p(\mathbf{X}_t | \mathbf{X}_{t-1})$$

First-Order Markov Models

Linear Dynamics

Data Space



$$p(\mathbf{X}_t | \mathbf{X}_{t-1})$$

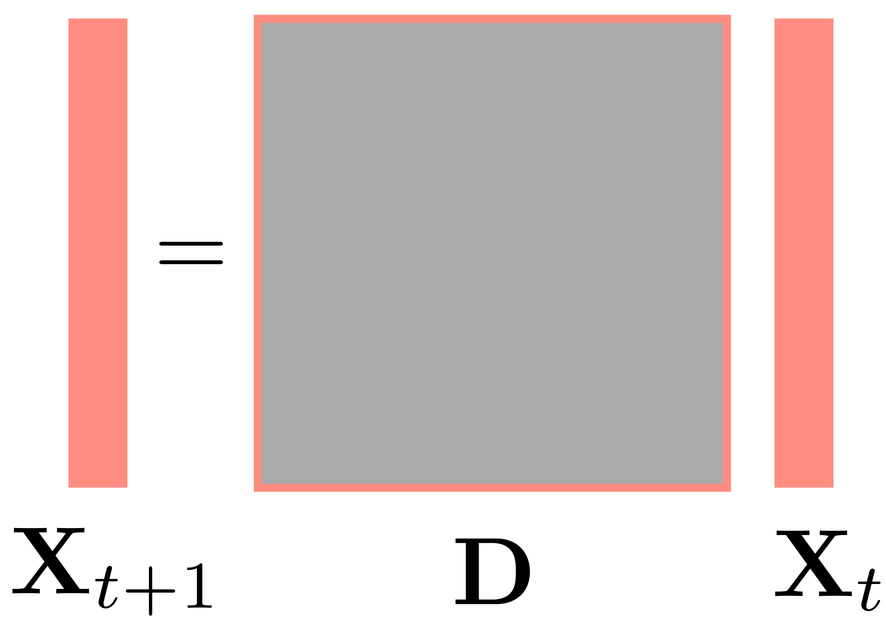
$$\mathbf{X}_t = f(\mathbf{X}_{t-1})$$ Dynamical Function

$$\mathbf{X}_{t+1} = \mathbf{D}\mathbf{X}_t$$ Linear Dynamics

Markov Models

First Order

$$\mathbf{X}_{t+1} = \mathbf{D}\mathbf{X}_t \qquad \mathbf{D} \in \mathbb{R}^{D \times D}$$



$$\mathbf{X}_{t+n} = \mathbf{D}^n\mathbf{X}_t$$

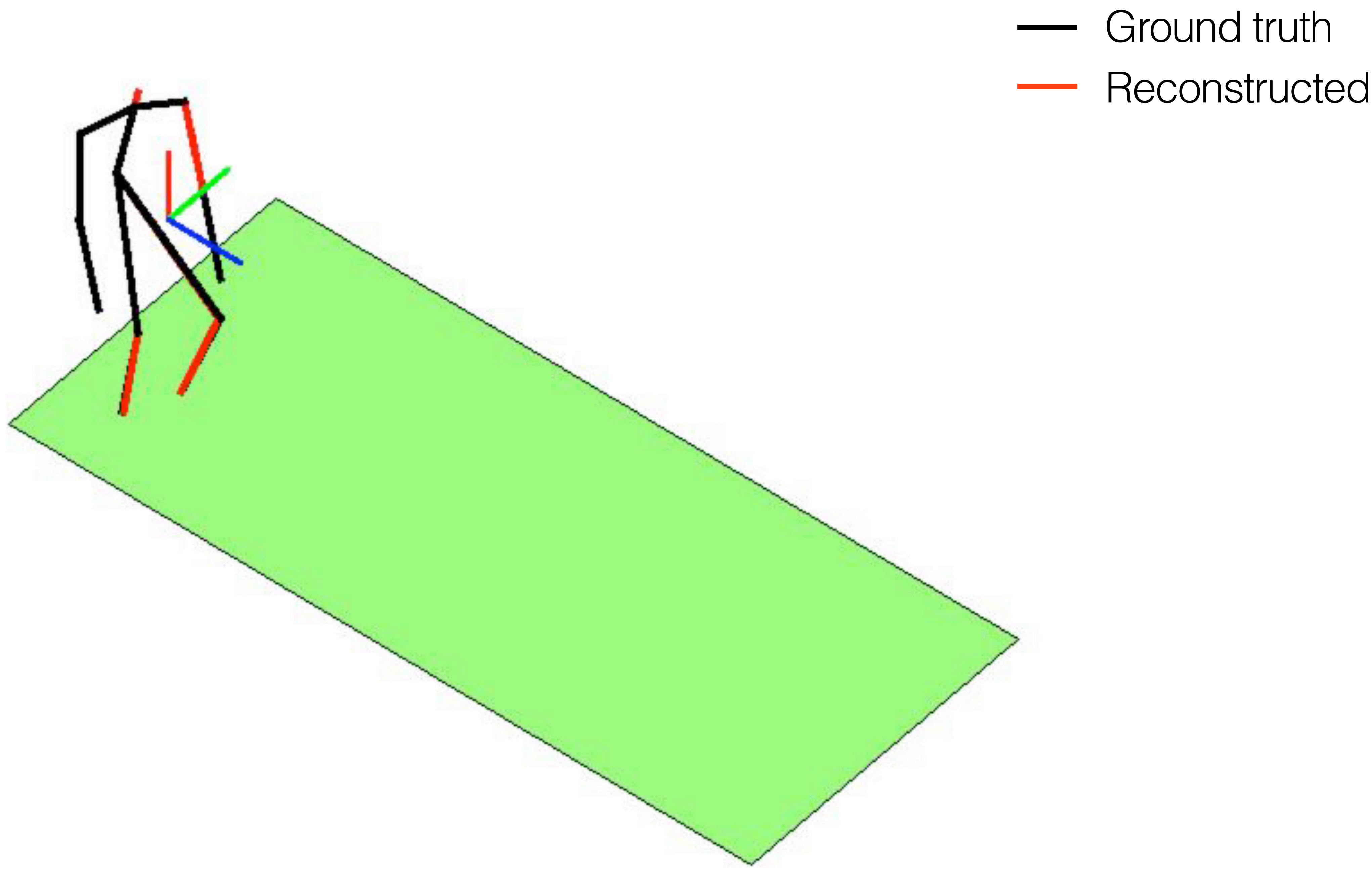
Markov Model

Reconstruction

$$\begin{bmatrix} \bar{\mathbf{X}}_{t+1} \\ \bar{\mathbf{X}}_{t+2} \\ \bar{\mathbf{X}}_{t+3} \\ \vdots \\ \bar{\mathbf{X}}_{t+F+1} \end{bmatrix} = \mathbf{D} \begin{bmatrix} \mathbf{X}_{t} \\ \mathbf{X}_{t+1} \\ \mathbf{X}_{t+2} \\ \vdots \\ \mathbf{X}_{t+F} \end{bmatrix}$$

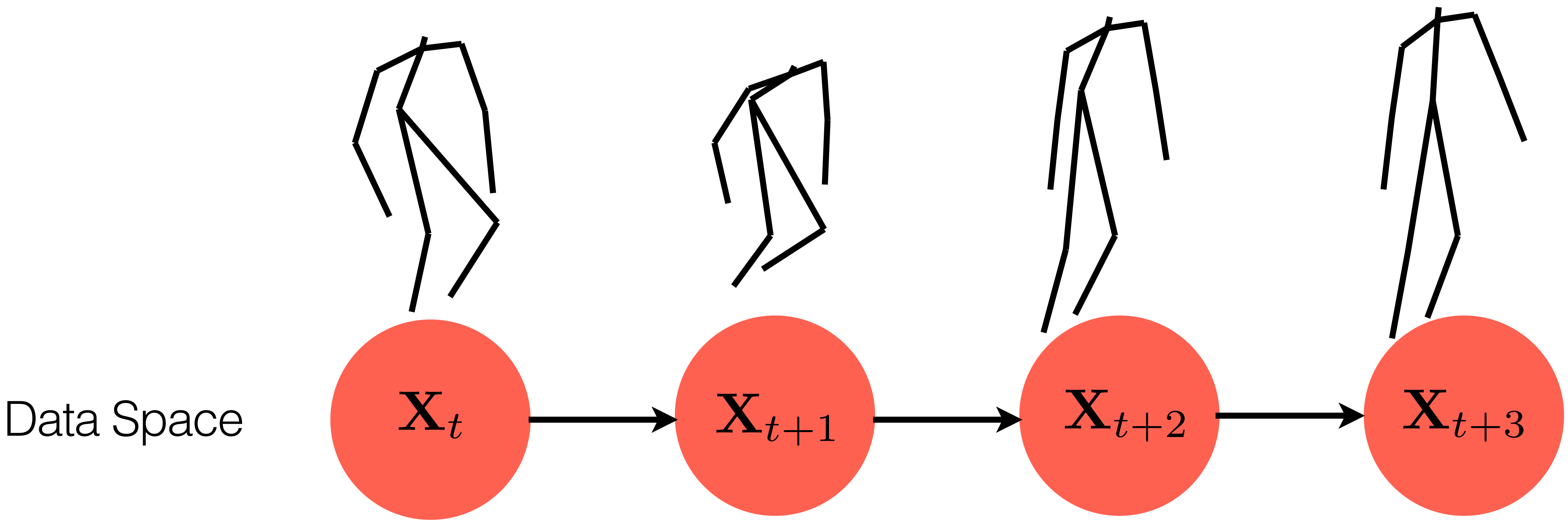
First-order AR

Reconstruction



Autoregressive Models

First-order Markov Model



$$p(\mathbf{X}_1, \cdots, \mathbf{X}_F) = p(\mathbf{X}_1) \prod_{t=2}^{F} p(\mathbf{X}_t | \mathbf{X}_{t-1})$$

$$p(\mathbf{X}_t | \mathbf{X}_{t-1}) = \mathcal{N}(\mathbf{X}_t | \mathbf{D}\mathbf{X}_{t-1}, \Sigma)$$

First-Order AR

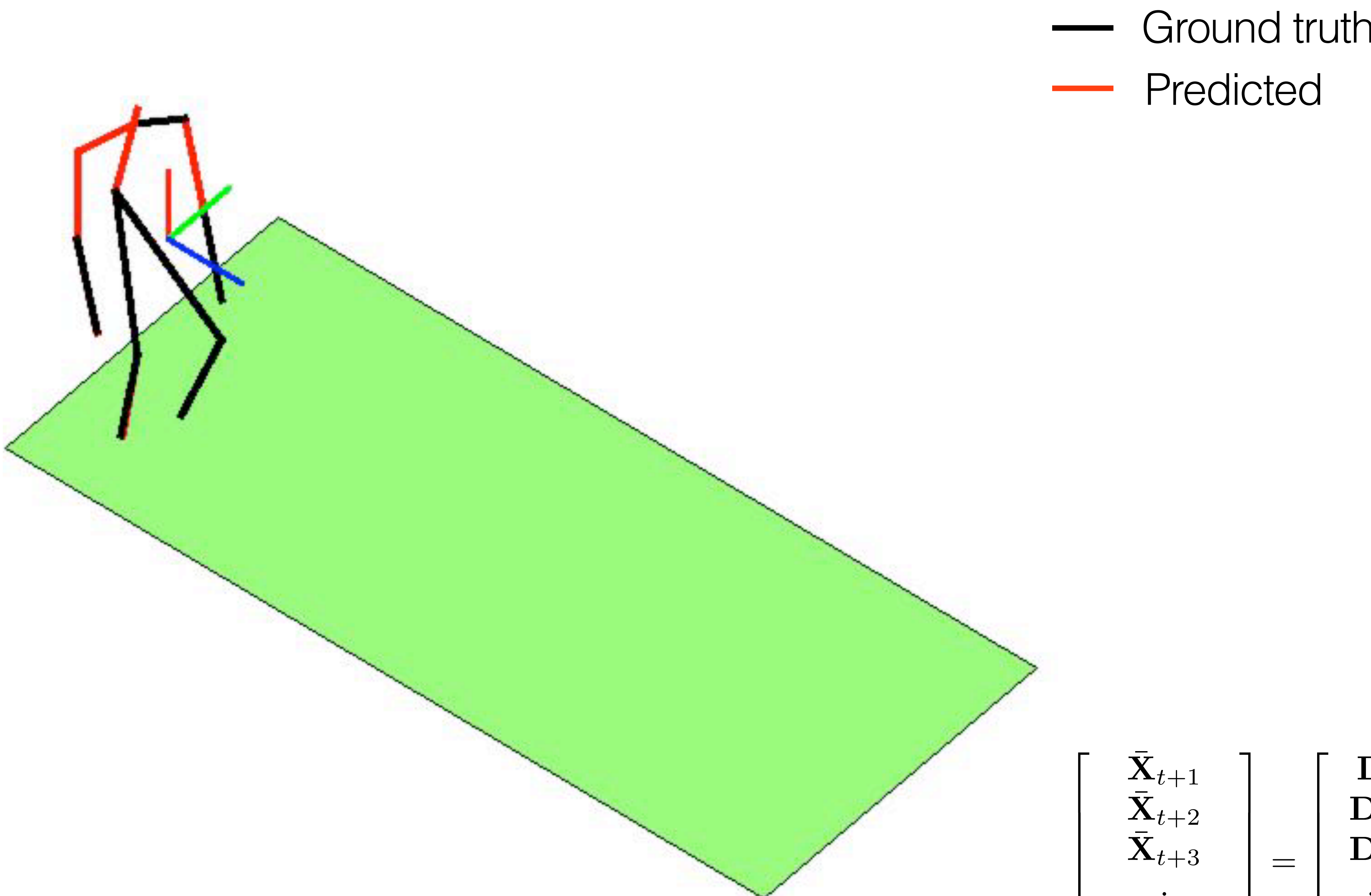
Prediction

$$\begin{bmatrix} \bar{\mathbf{X}}_{t+1} \\ \bar{\mathbf{X}}_{t+2} \\ \bar{\mathbf{X}}_{t+3} \\ \vdots \\ \bar{\mathbf{X}}_{t+F+1} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \\ \mathbf{D}^2 \\ \mathbf{D}^3 \\ \vdots \\ \mathbf{D}^F \end{bmatrix} \mathbf{X}_t$$

Observability Matrix

First-Order AR

Linear Prediction



$$\begin{bmatrix} \bar{\mathbf{X}}_{t+1} \\ \bar{\mathbf{X}}_{t+2} \\ \bar{\mathbf{X}}_{t+3} \\ \vdots \\ \bar{\mathbf{X}}_{t+F+1} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \\ \mathbf{D}^2 \\ \mathbf{D}^3 \\ \vdots \\ \mathbf{D}^F \end{bmatrix} \mathbf{X}_t$$

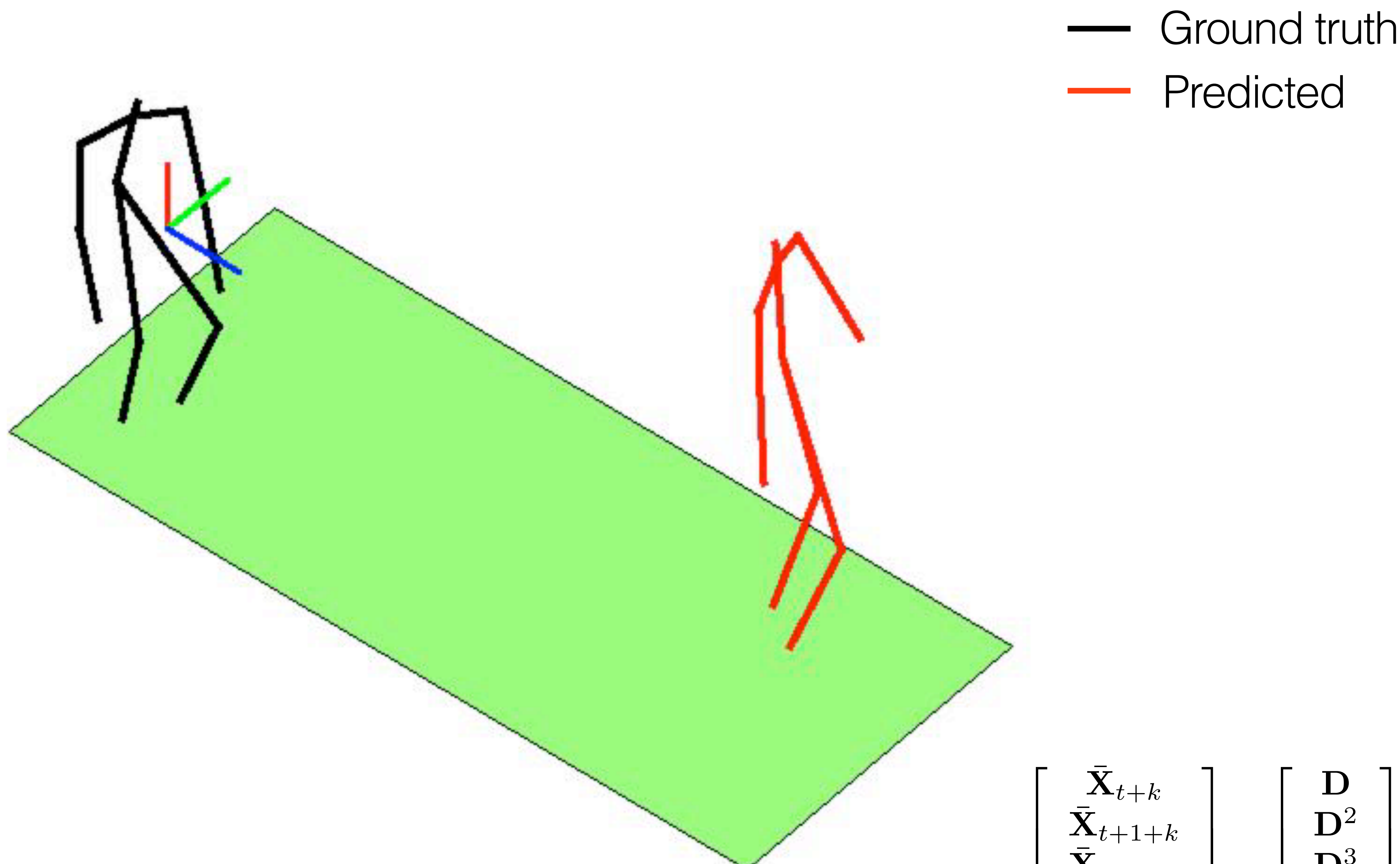
First-Order AR

Prediction Generalization

$$\begin{bmatrix} \bar{\mathbf{X}}_{t+k} \\ \bar{\mathbf{X}}_{t+1+k} \\ \bar{\mathbf{X}}_{t+2+k} \\ \vdots \\ \bar{\mathbf{X}}_{t+F+k} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \\ \mathbf{D}^2 \\ \mathbf{D}^3 \\ \vdots \\ \mathbf{D}^F \end{bmatrix} \mathbf{X}_{t+k}$$

First Order AR

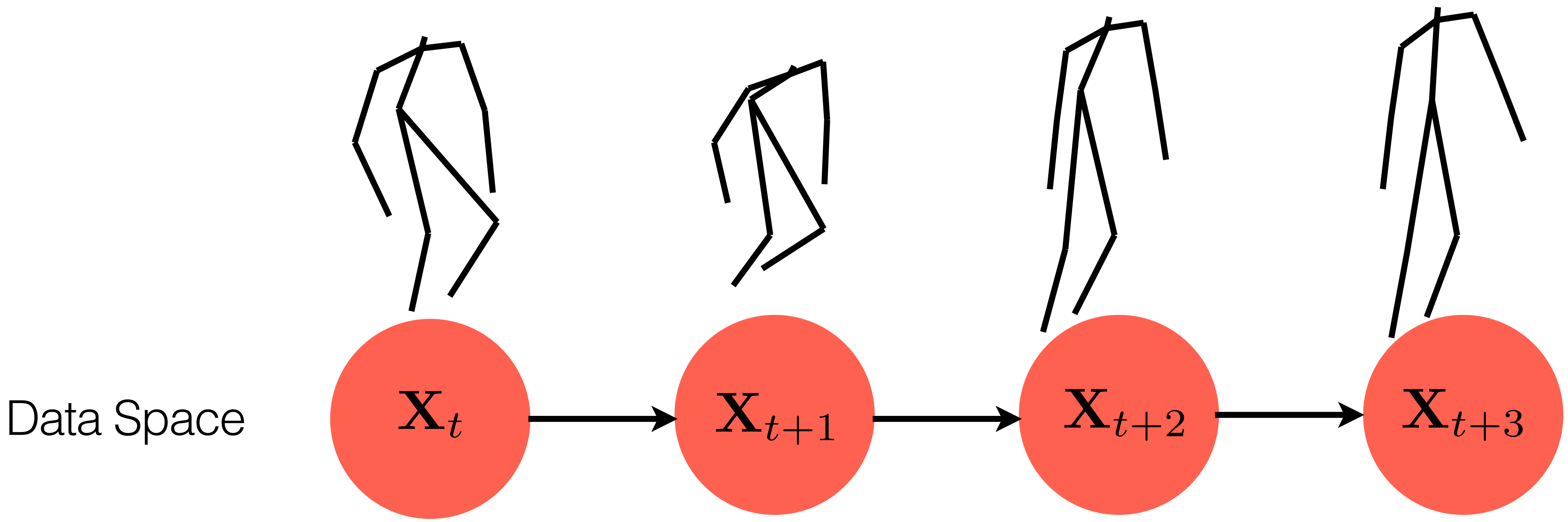
Prediction Generalization



$$\begin{bmatrix} \bar{\mathbf{X}}_{t+k} \\ \bar{\mathbf{X}}_{t+1+k} \\ \bar{\mathbf{X}}_{t+2+k} \\ \vdots \\ \bar{\mathbf{X}}_{t+F+k} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \\ \mathbf{D}^2 \\ \mathbf{D}^3 \\ \vdots \\ \mathbf{D}^F \end{bmatrix} \mathbf{X}_{t+k}$$

Autoregressive Models

First-order Markov Model

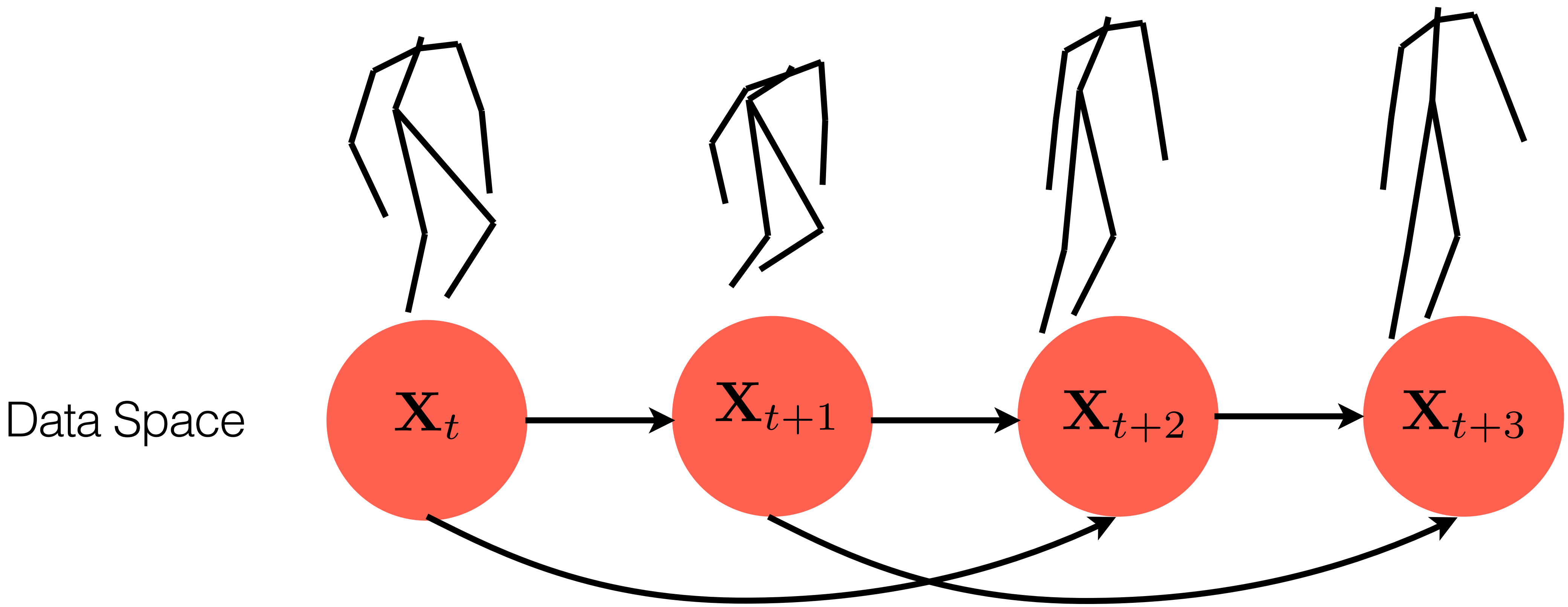


$$p(\mathbf{X}_t|\mathbf{X}_{t-1}) = \mathcal{N}(\mathbf{X}_t|\mathbf{D}\mathbf{X}_{t-1}, \Sigma)$$

Ideas?

Autoregressive Models

Second-Order Markov Model

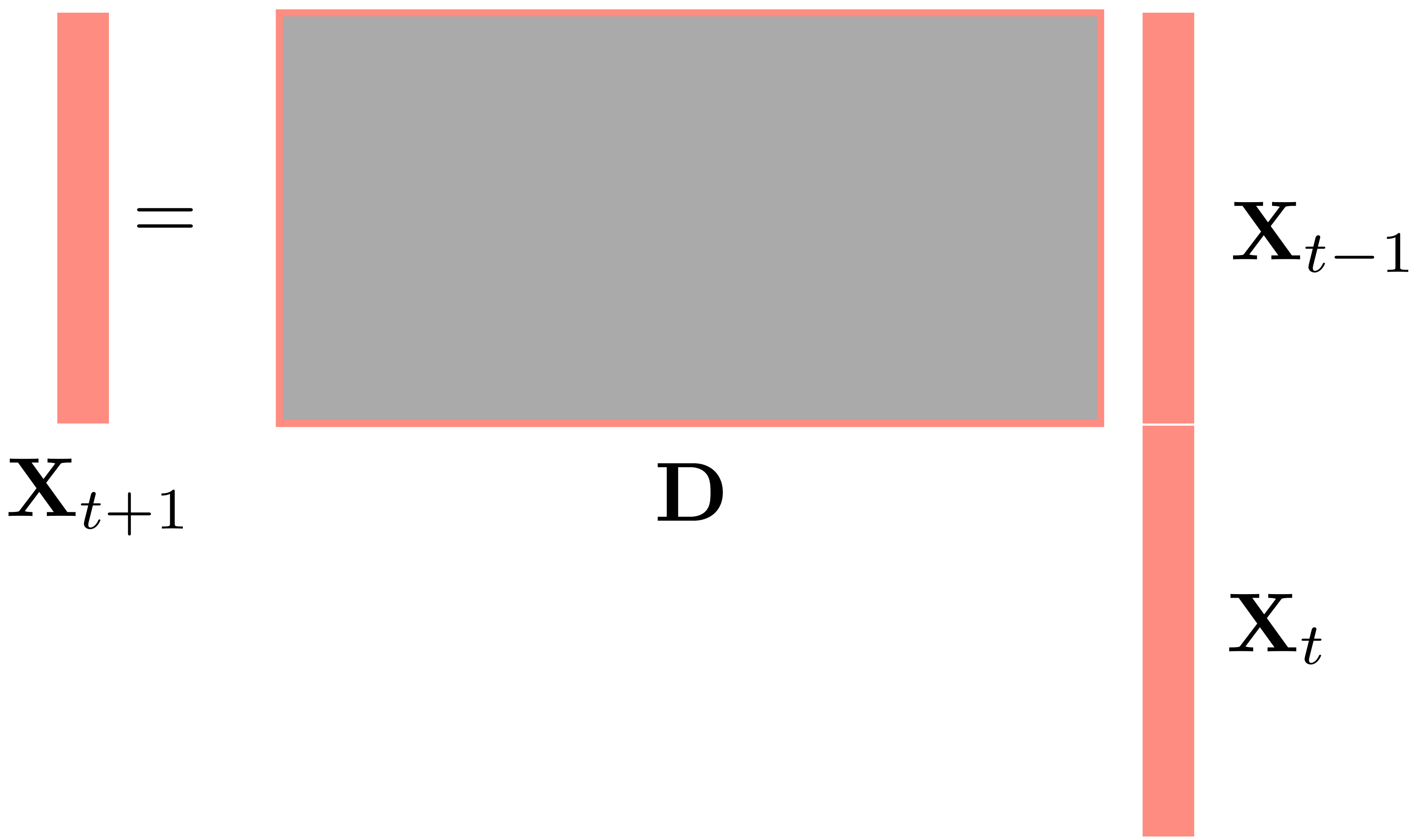


$$p(\mathbf{X}_1, \cdots, \mathbf{X}_F) = p(\mathbf{X}_1)p(\mathbf{X}_2|\mathbf{X}_1)\prod_{t=1}^{F} p(\mathbf{X}_t|\mathbf{X}_{t-1}, \mathbf{X}_{t-2})$$

AR Models

Second-Order Systems

$$\mathbf{X}_{t+1} = \mathbf{D} \begin{bmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_t \end{bmatrix} \qquad \mathbf{D} \in \mathbb{R}^{D \times 2D}$$



Considerations

AR systems

- **Complexity**: Curse of dimensionality, computation, compaction?
- **Predictive Precision**: How accurately does the model predict observations?
- **Generalization Ability**: How well does the model generalize to new data?

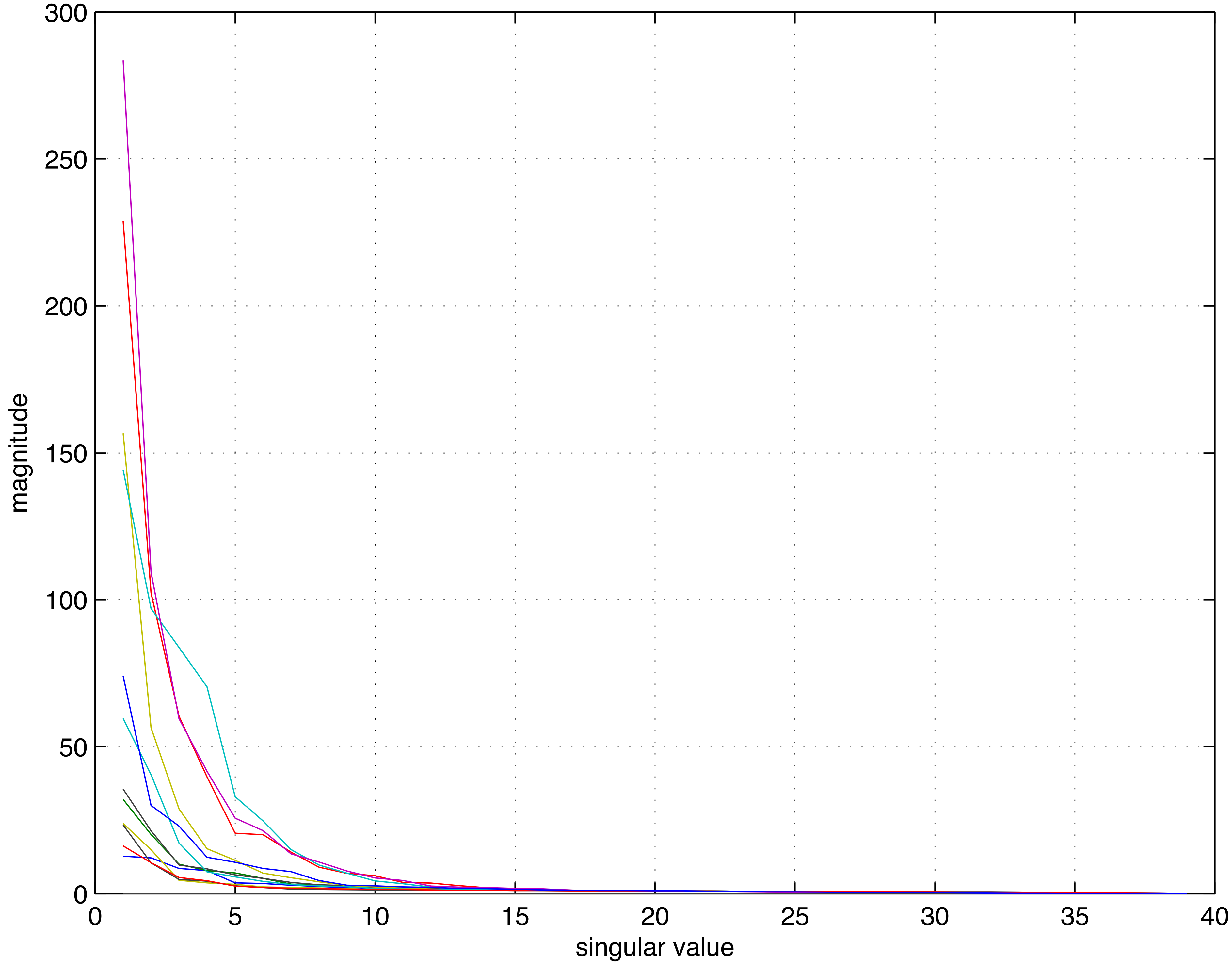
Considerations

Trade-offs

- **Memory**: How far back should you look? How much is it worth in extra dimensionality?
- **Linearity**: How much of a limitation is the linearity of the dynamical system?

Singular Values of D

Compaction across Actions (walk, sit, fall, etc.)



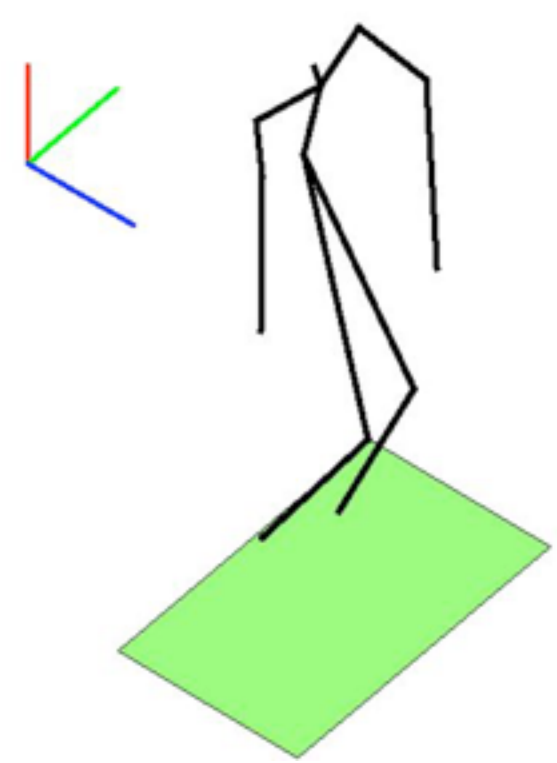
Ideas?

Pose Correlations

Latent Variable Models

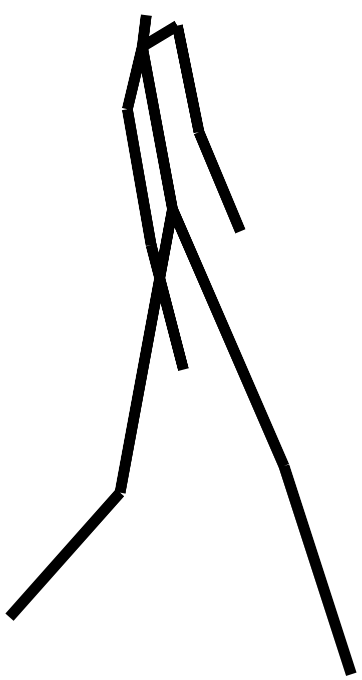
- How many degrees of freedom are there *really*?

$$\mathbf{X}_t \sim \mathcal{N}(\mu, \Sigma)$$



Latent Variable Models

Linear Models



$= \quad w_1$

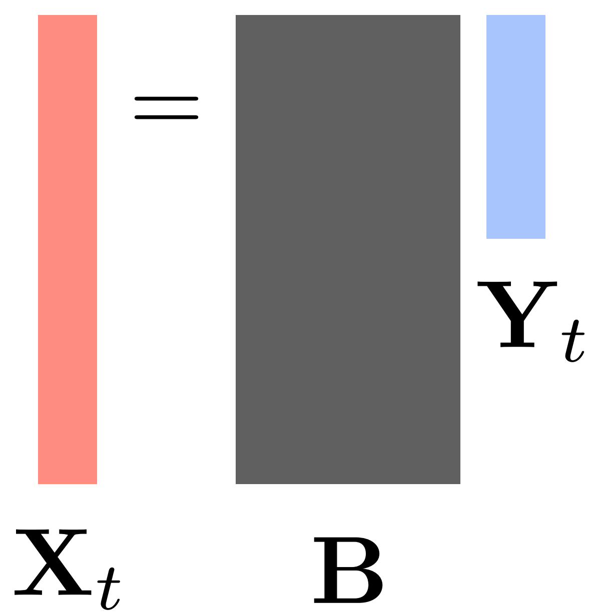
$+ w_2$

$+ w_3$

$\cdots$

Linear Projection

Principal Component Analysis



Latent Variable Model

$\{\mathbf{X}_n\}$: Training Data $\qquad \mathbf{X}_t \in \mathbb{R}^D$

$$\mathbf{Y}_t \in \mathbb{R}^M \quad M < D$$

Probabilistic PCA

Distribution

$$p(\mathbf{X}_t|\mathbf{Y}_t)$$

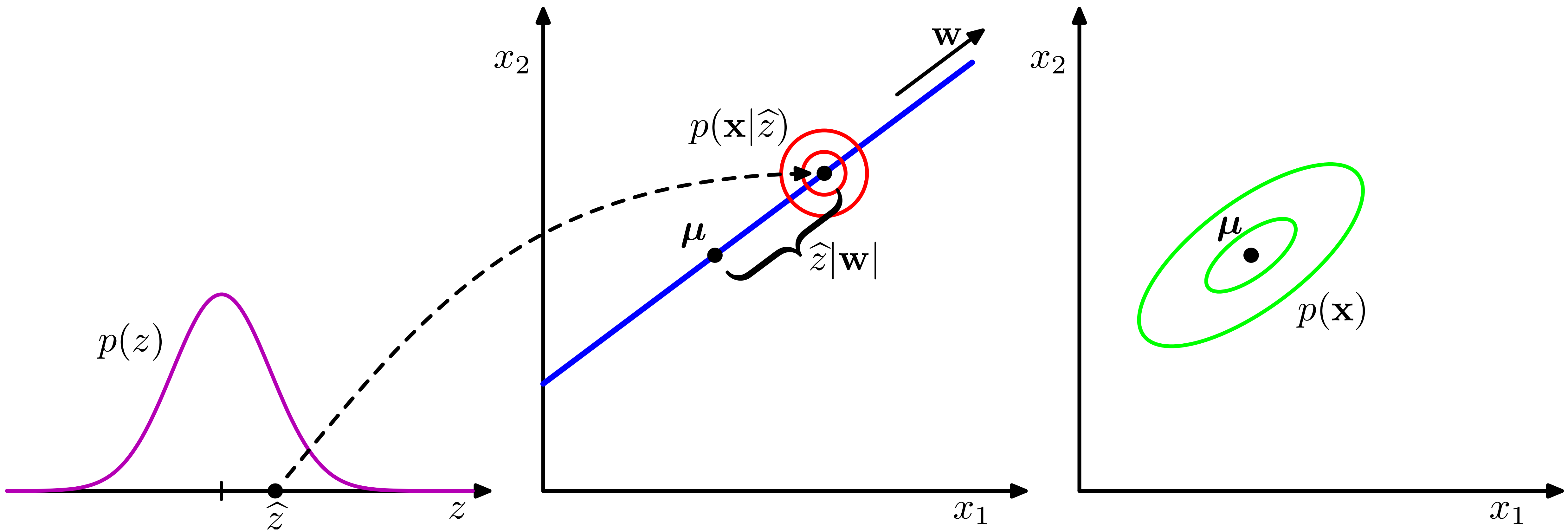
$$\mathbf{X}_t = \mathbf{B}\mathbf{Y}_t + \mu + \epsilon$$

$$p(\mathbf{Y}_t) = \mathcal{N}(\mathbf{z}|0, \mathbf{I})$$

$$p(\mathbf{X}_t|\mathbf{Y}_t) = \mathcal{N}(\mathbf{X}_t|\mathbf{B}\mathbf{Y}_t + \mu, \sigma^2\mathbf{I})$$

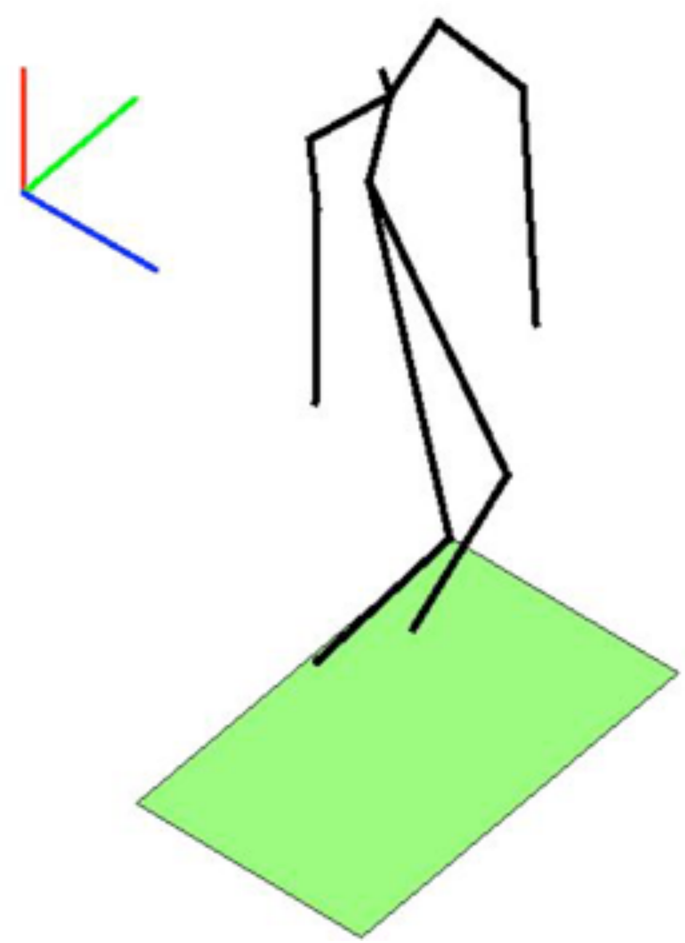
Probabilistic PCA

Generative View of PCA



Probabilistic PCA

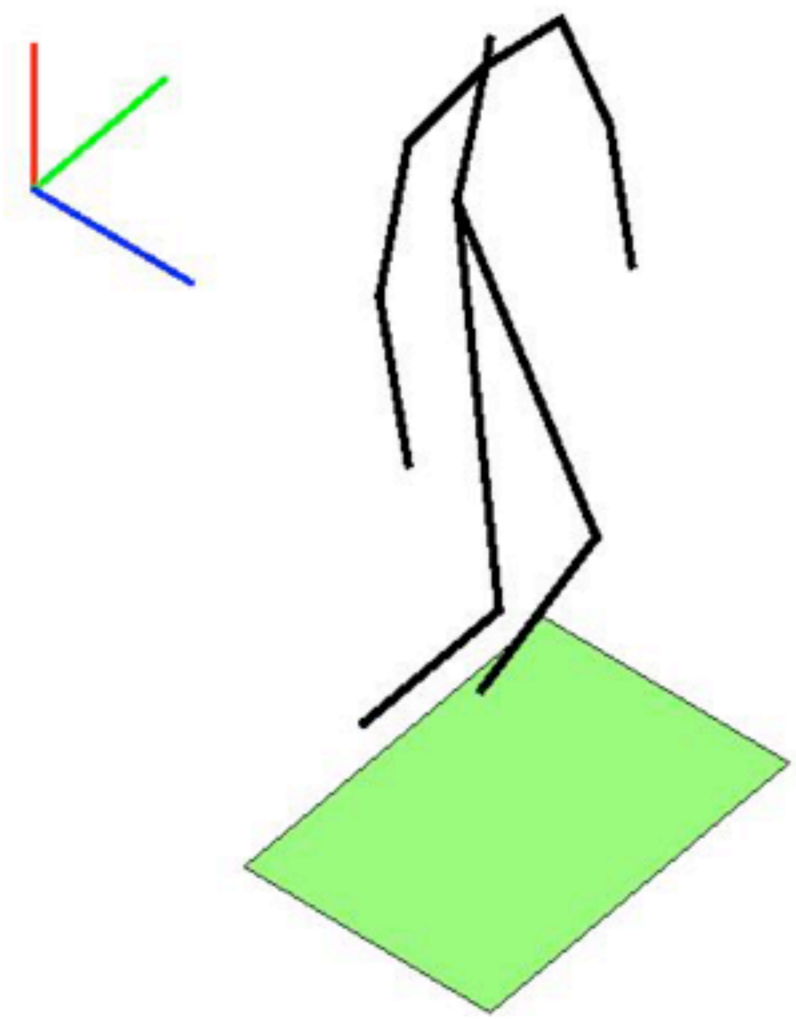
Sampling Standing Up



$$\mathbf{X}_t \sim \mathcal{N}(\mu, \Sigma)$$

Probabilistic PCA

Sampling Standing Up

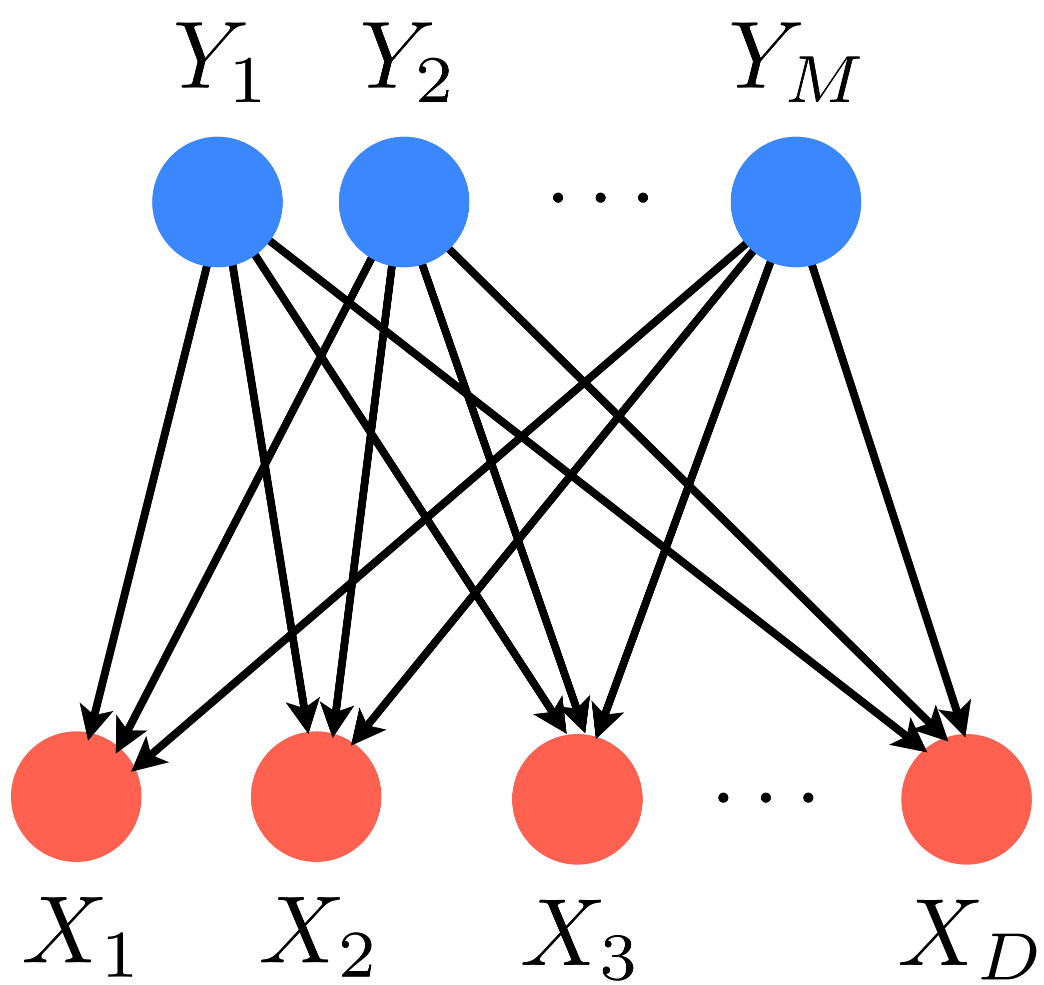


$$p(\mathbf{X}_t|\mathbf{Y}_t) = \mathcal{N}(\mathbf{X}_t|\mathbf{B}\mathbf{Y}_t + \mu, \sigma^2\mathbf{I})$$

Graphical Model

Component Analysis

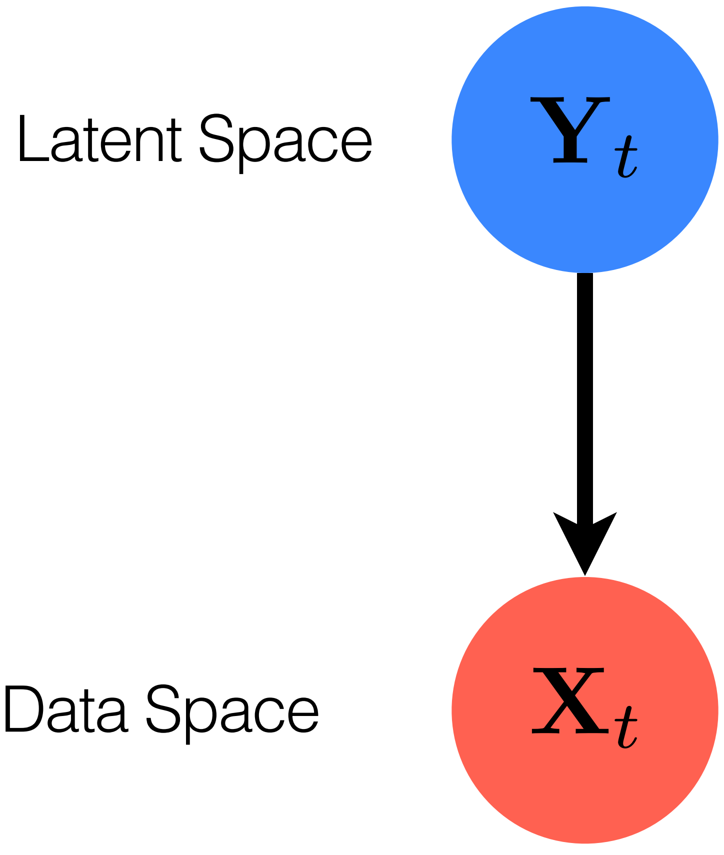
$$\mathbf{X}_t = \mathbf{B}\mathbf{Y}_t + \mu + \epsilon$$



Graphical Model

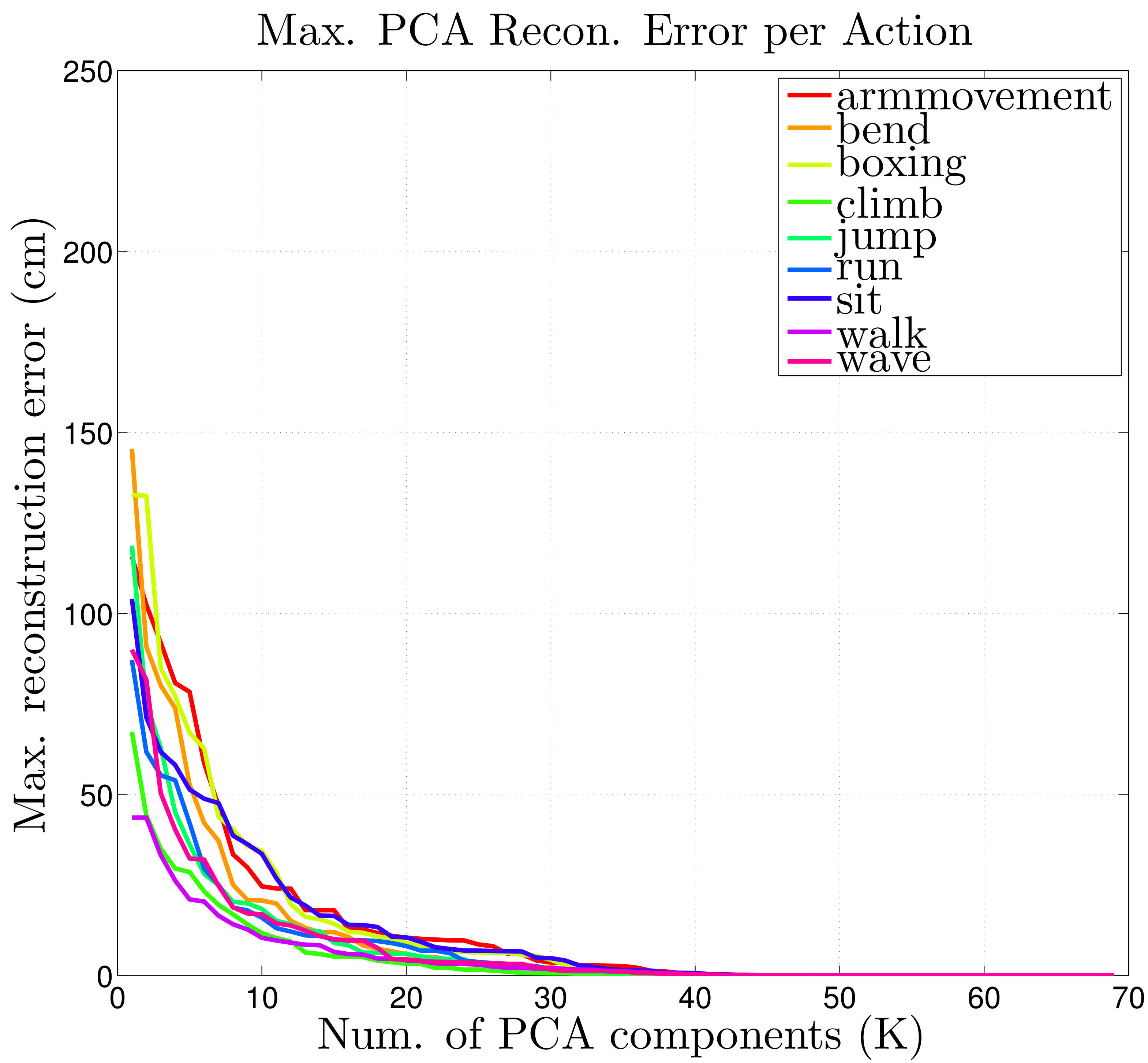
Component Analysis

$$\mathbf{X}_t = \mathbf{B}\mathbf{Y}_t + \mu + \epsilon$$



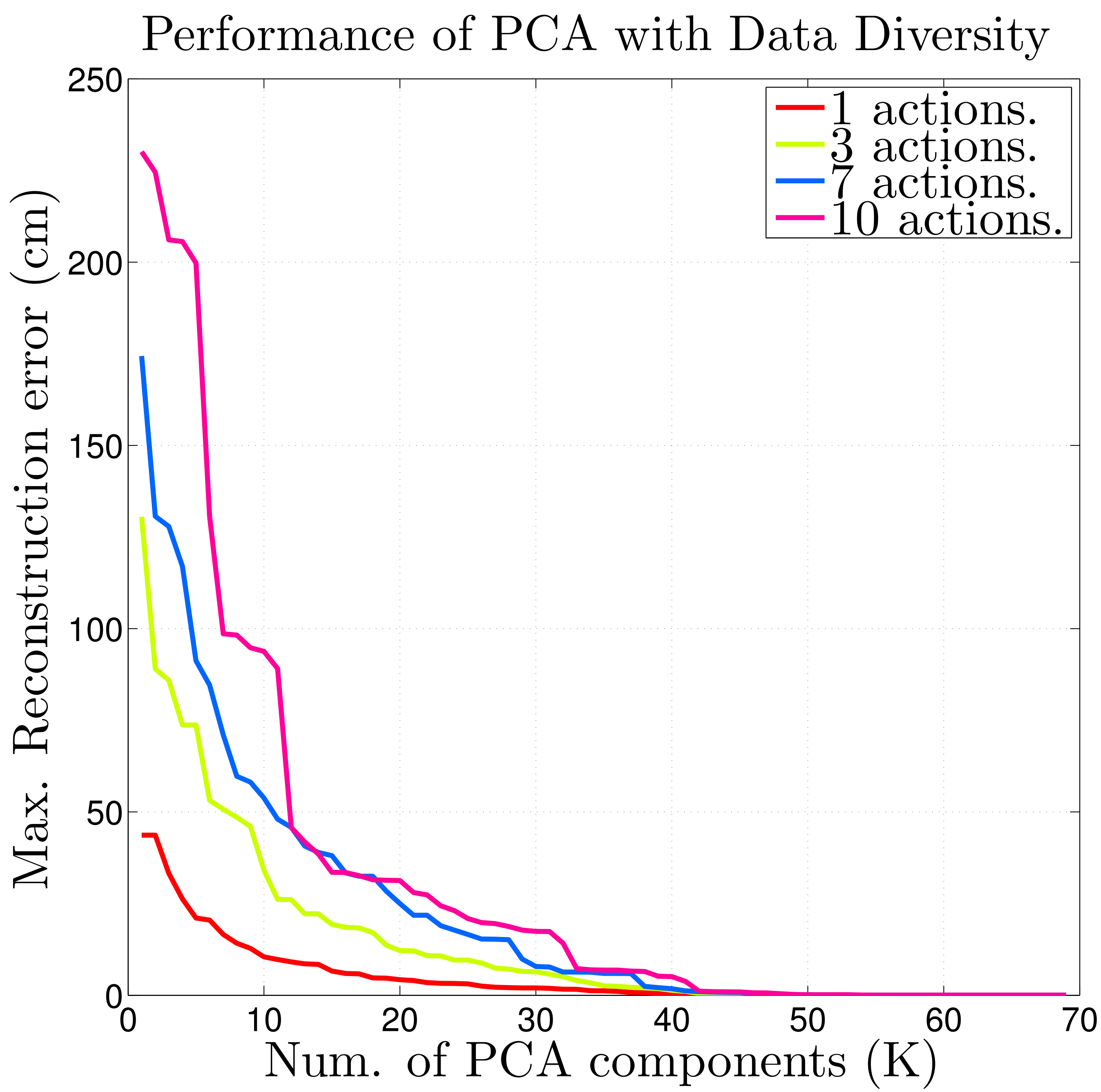
PCA Works!

Less than 5cm max-error with <30 components



Well...Somewhat...

PCA works when Action is Known

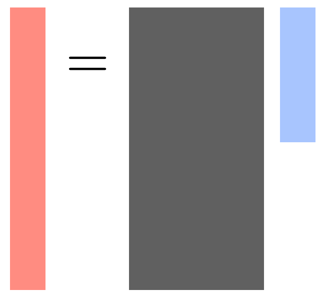


Overcomplete Dictionaries

Ramakrishna et al. 2012

$$\arg\min_{\mathbf{Y}_t} \|\mathbf{X}_t - \mathbf{B}\mathbf{Y}_t\|_2$$

$$\mathbf{X}_t = \mathbf{B}\mathbf{Y}_t$$



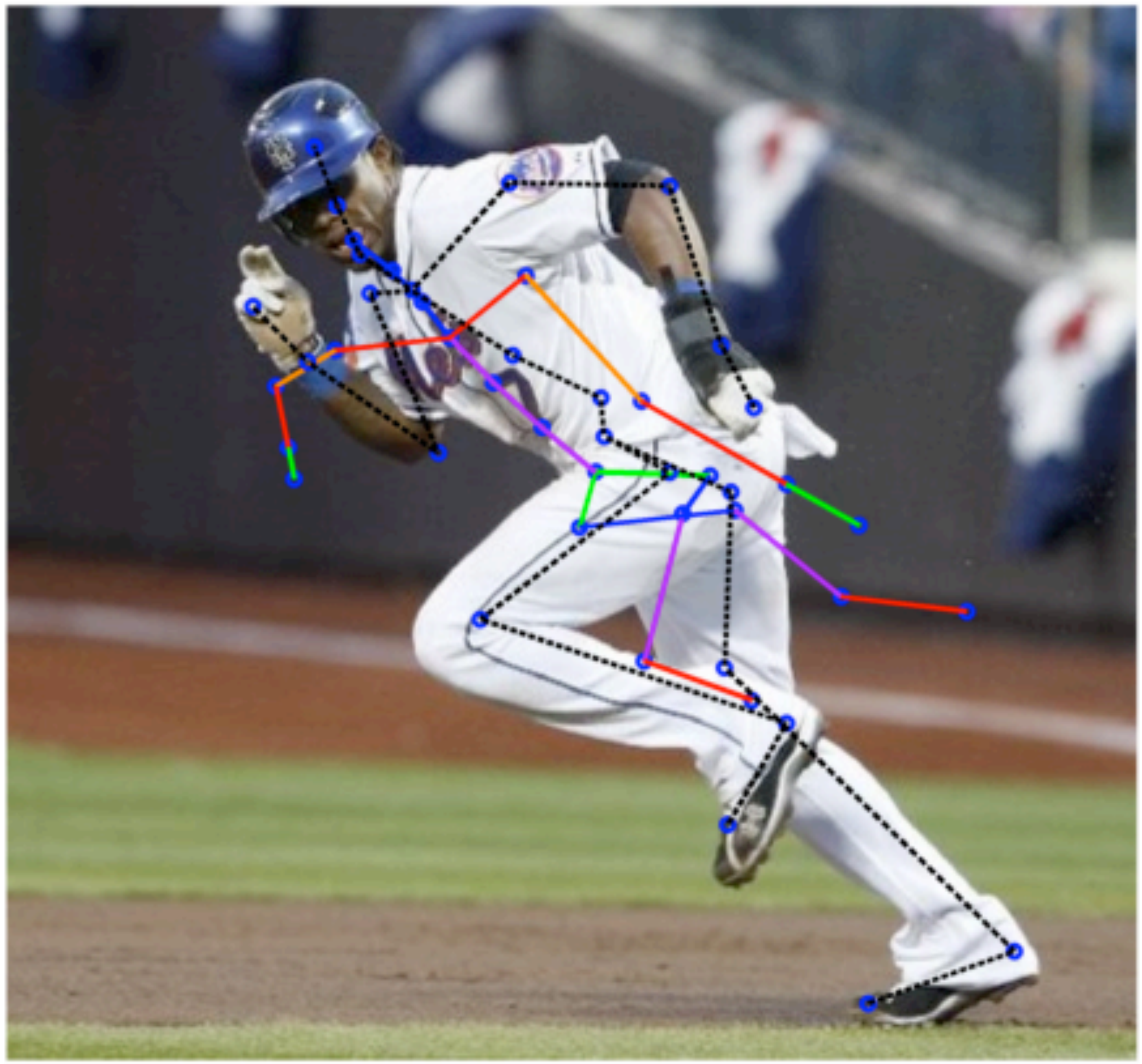
$$\mathbf{X}_t = \mathbf{B}\mathbf{Y}_t$$

L1-norm “encourages” sparsity in **Y**

3D Reconstruction

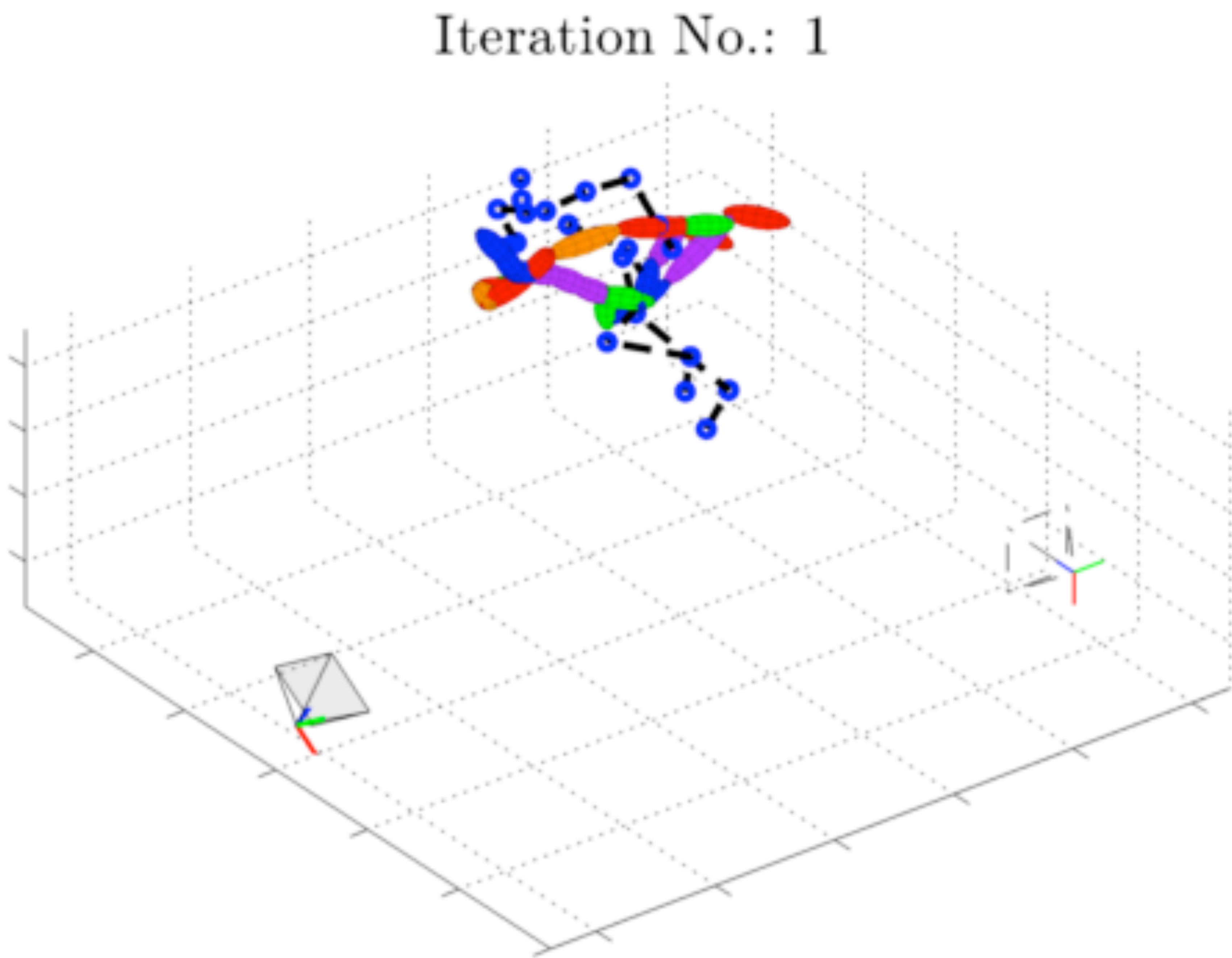
Reprojection Error Decreases at Each Iteration

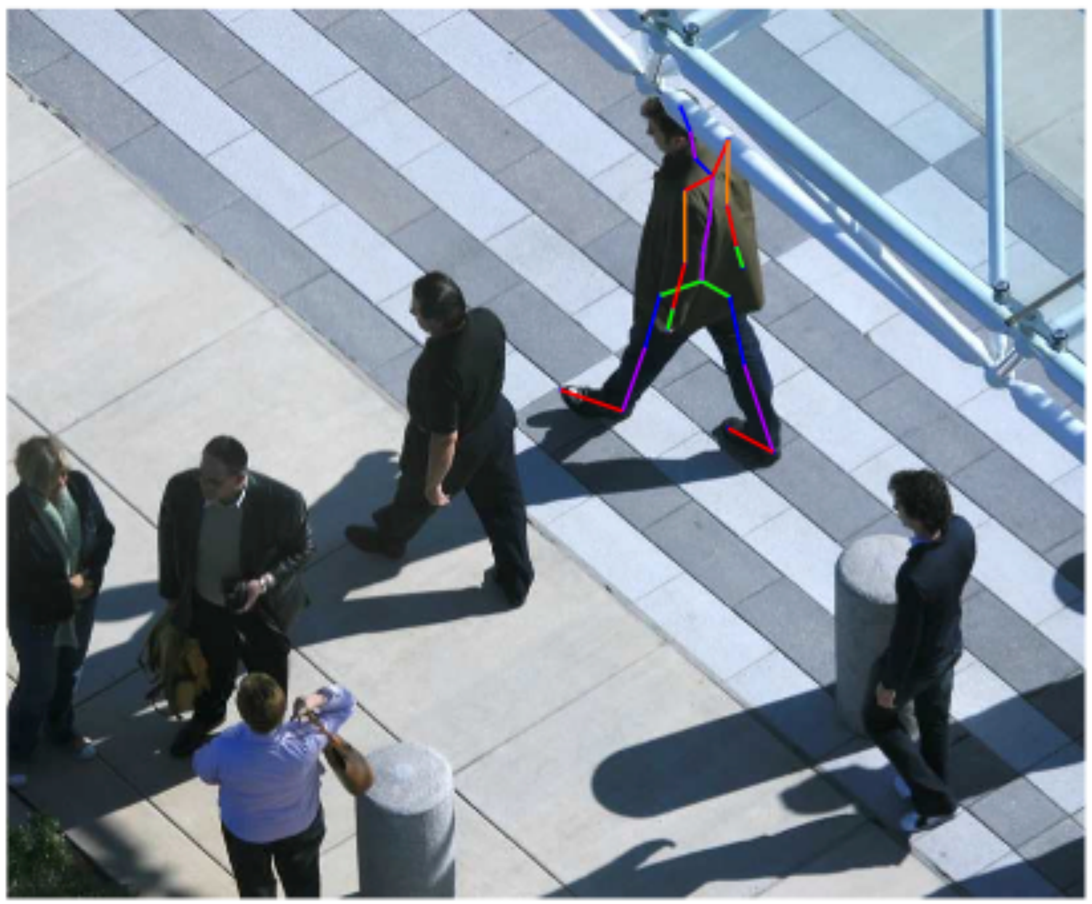
Iteration No.: 1

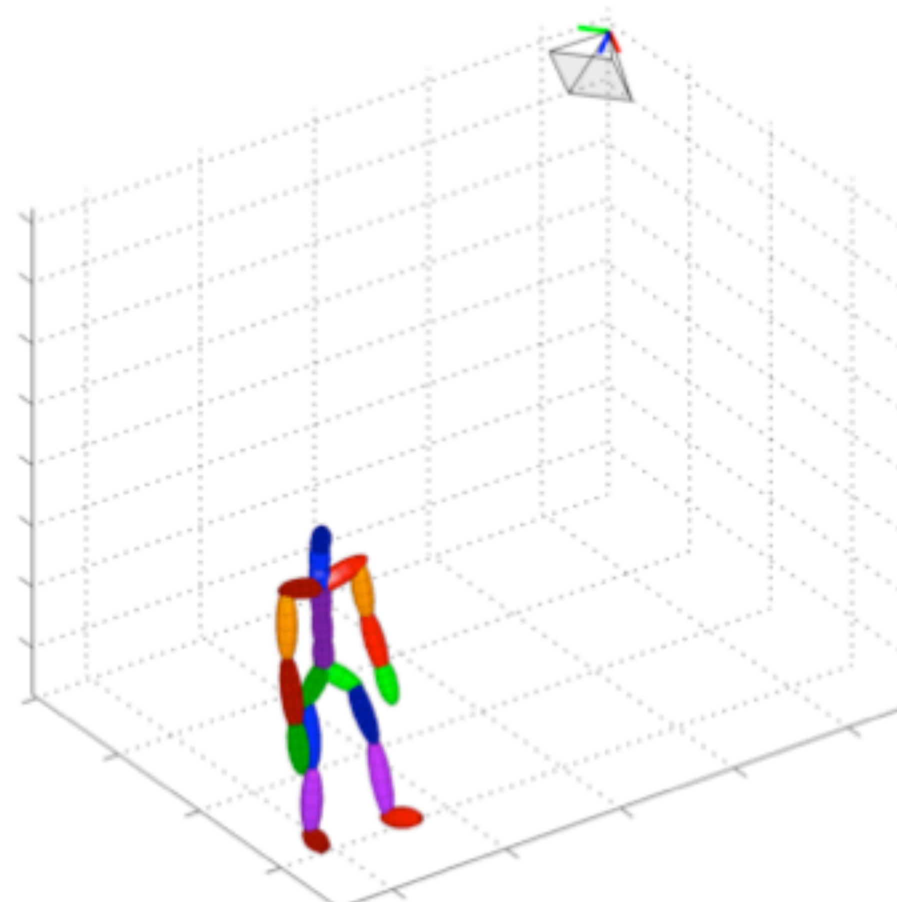


3D Reconstruction

3D Pose and Camera

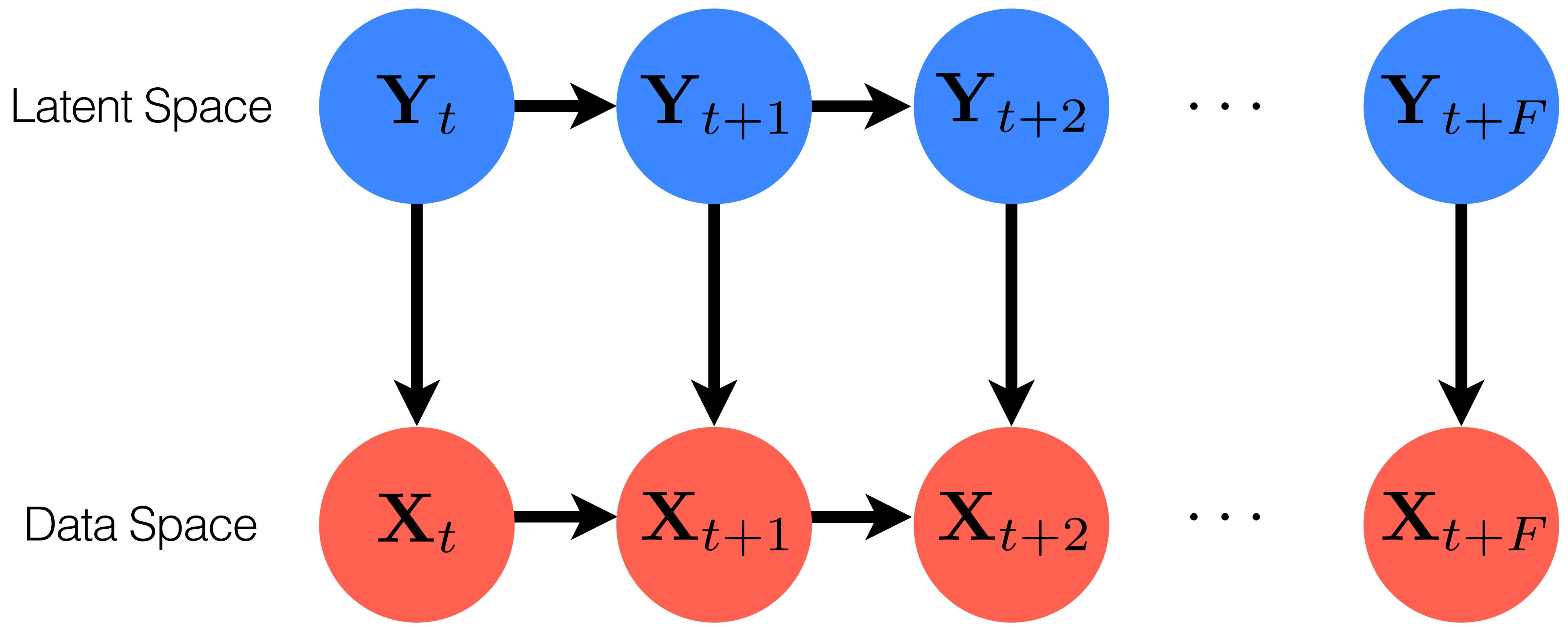






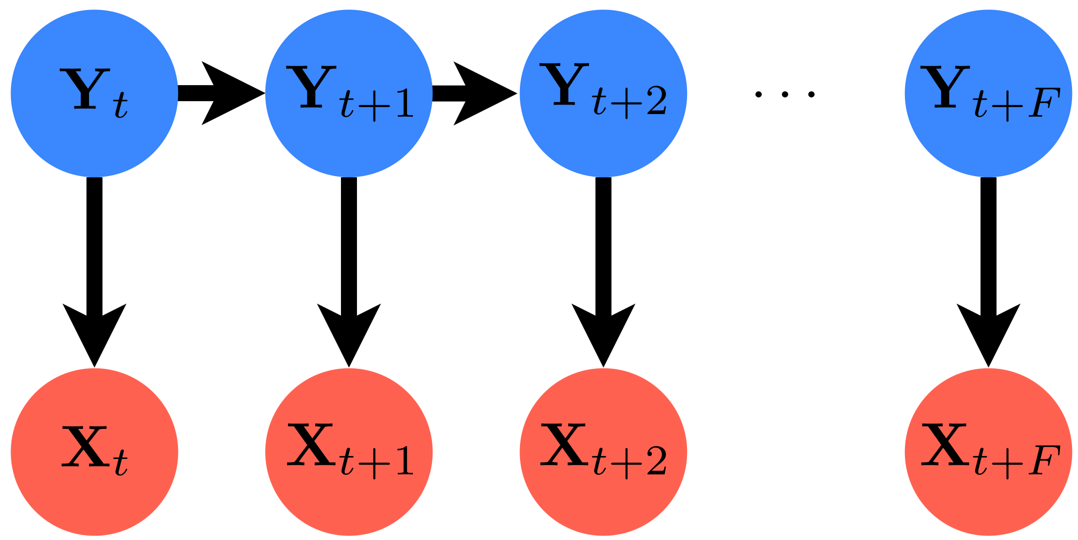
Dynamical Models

Graphical Models

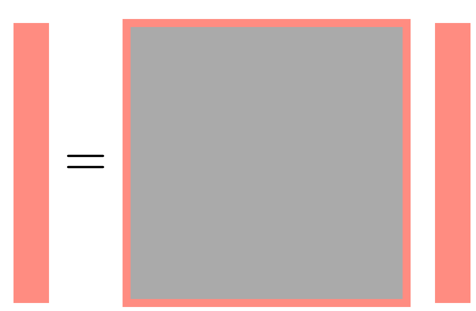


Linear Dynamical System

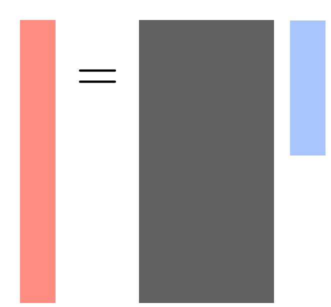
Graphical Summary



$$\mathbf{X}_{t+1} = \mathbf{D}\mathbf{X}_t \qquad \mathbf{X}_t = \mathbf{B}\mathbf{Y}_t$$



Linear Dynamics

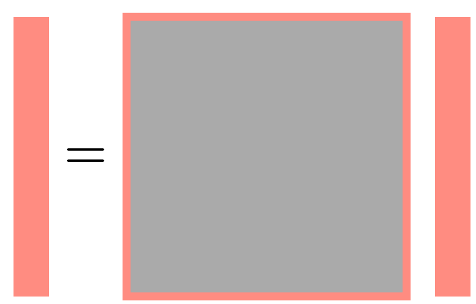


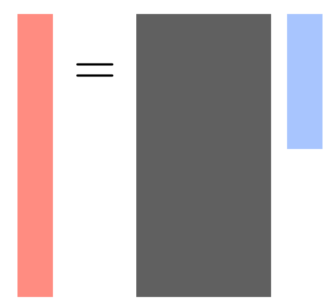
Latent Variable Model

Model Reduction

Dynamics in the Latent Space

$$\mathbf{X}_{t+1} = \mathbf{D}\mathbf{X}_t \qquad \mathbf{X}_t = \mathbf{B}\mathbf{Y}_t$$

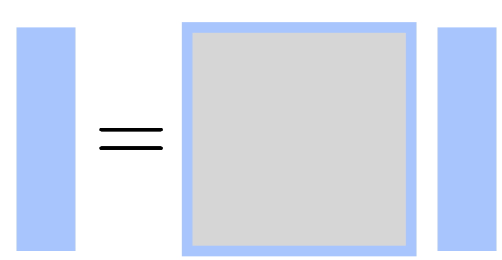




$$\mathbf{B}\mathbf{Y}_t = \mathbf{D}\mathbf{B}\mathbf{Y}_t$$

$$\mathbf{Y}_t = \mathbf{B}^T\mathbf{D}\mathbf{B}\mathbf{Y}_t$$

$$\mathbf{Y}_t = \mathbf{G}\mathbf{Y}_t$$

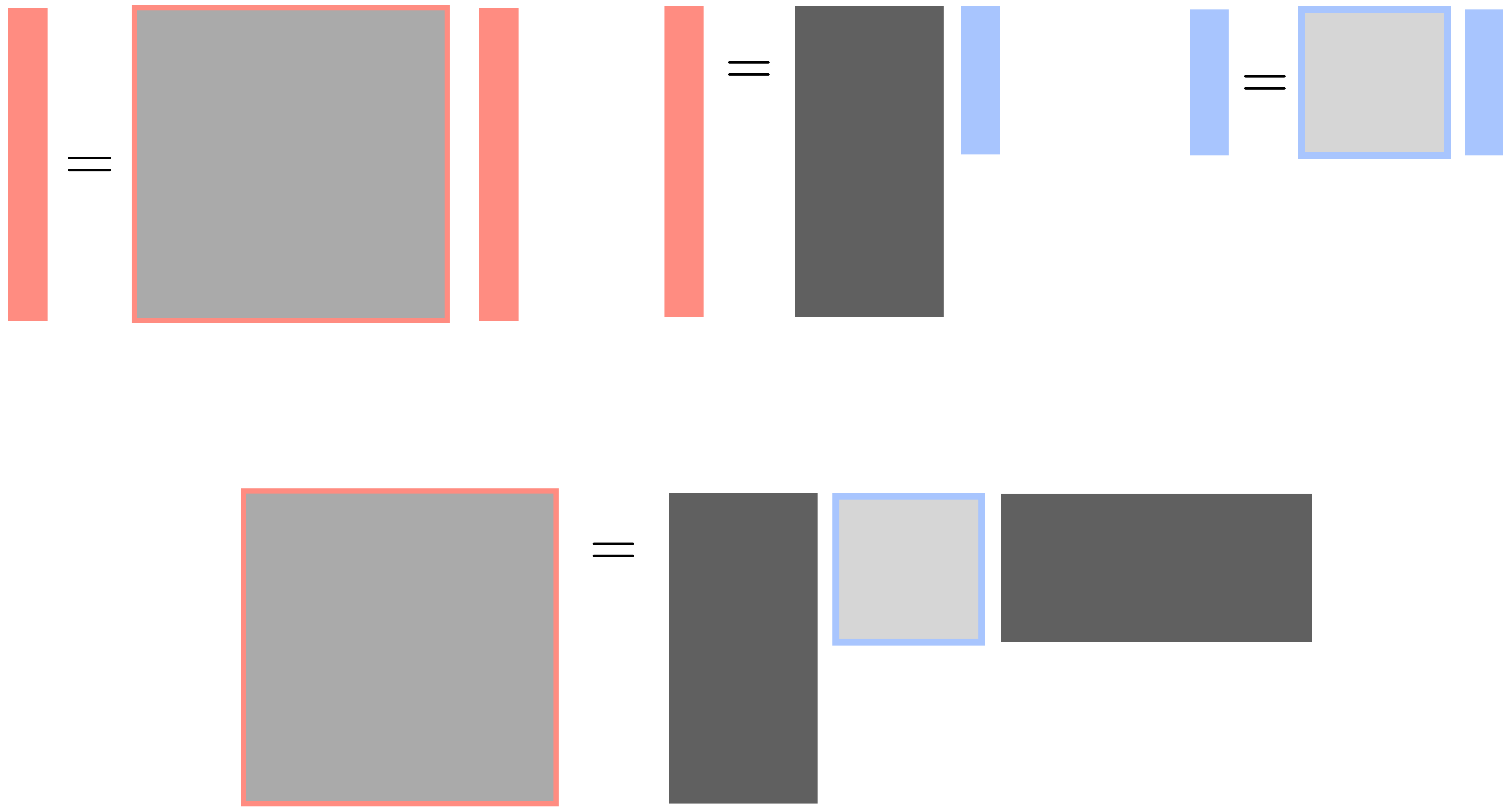


$$\mathbf{G} = \mathbf{B}^T\mathbf{D}\mathbf{B}$$

Model Reduction

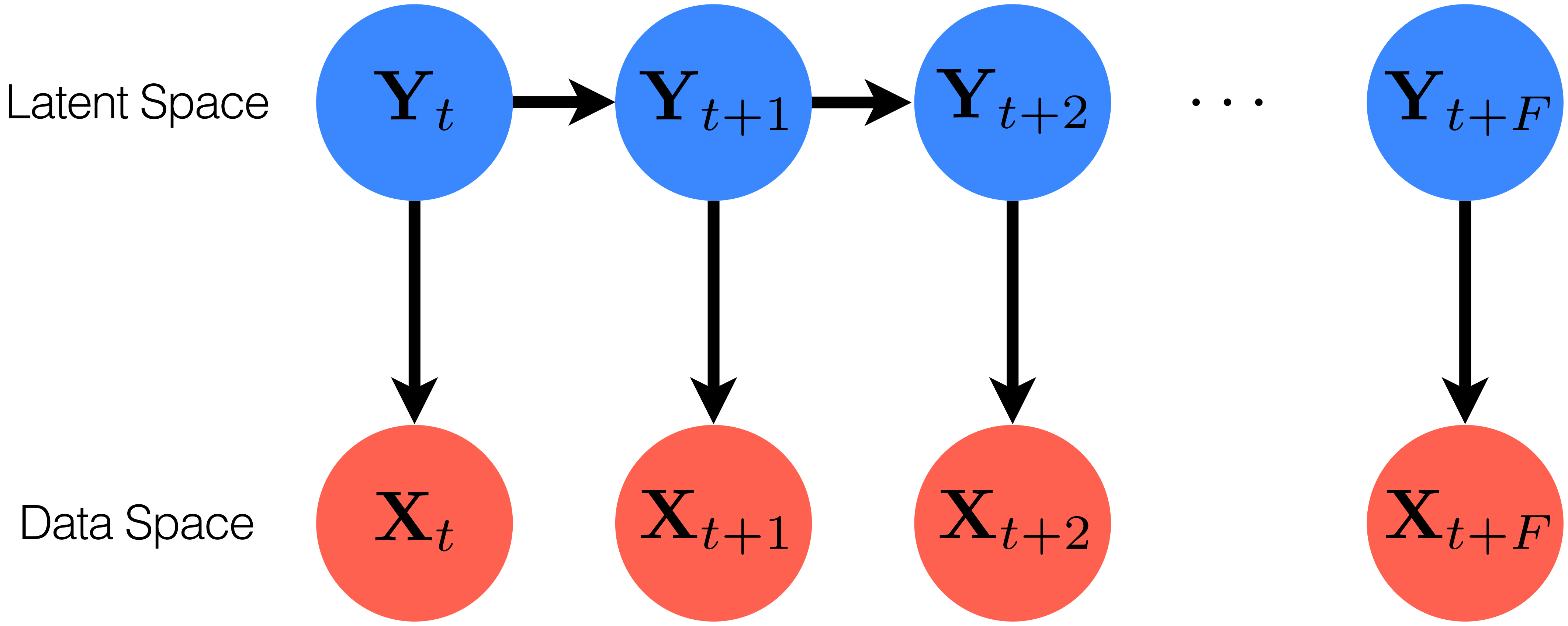
Projected Dynamics

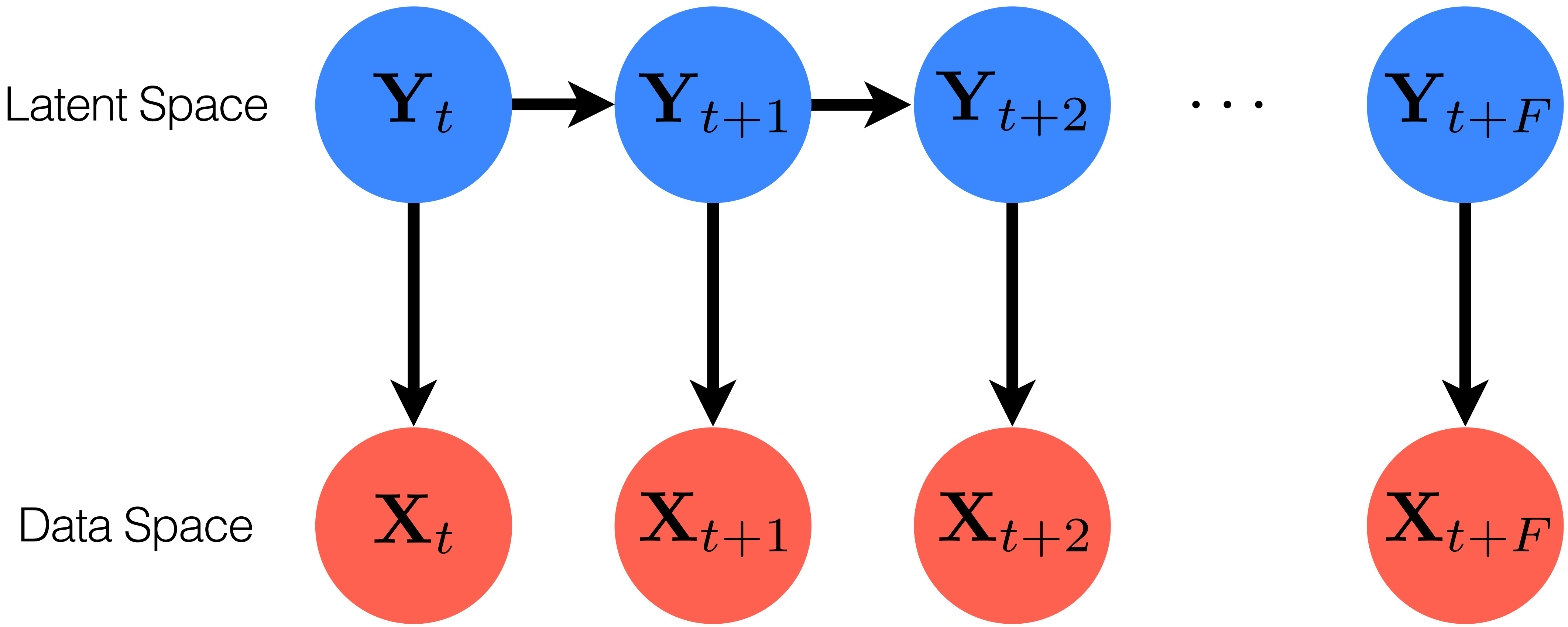
$$\mathbf{G} = \mathbf{B}^T\mathbf{D}\mathbf{B}$$



Dynamical Models

Graphical Models

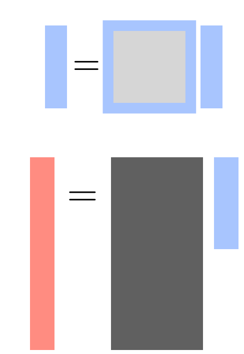




$$p(\mathbf{X}_1, \cdots, \mathbf{X}_F) = \int p(\mathbf{X}_1, \cdots, \mathbf{X}_F, \mathbf{Y}_1, \cdots \mathbf{Y}_F) d\mathcal{Y}$$

$$p(\mathbf{Y}_t | \mathbf{Y}_{t-1}) = \mathcal{N}(\mathbf{Y}_t | \mathbf{G}\mathbf{Y}_{t-1}, \Gamma)$$

$$p(\mathbf{X}_t | \mathbf{Y}_t) = \mathcal{N}(\mathbf{X}_t | \mathbf{B}\mathbf{Y}_t, \Sigma)$$



$$p(\mathbf{Y}_1) = \mathcal{N}(\mathbf{Y}_1 | \mu, \mathbf{V})$$

Nonlinear Dynamical Models?

- Linear-Gaussian Models work for individual activities
- Nonlinear Latent Variable Models:
  - Density Networks
  - Generative Topographic Mapping
  - Kernel PCA
  - Gaussian Process Latent Variable Models
- Nonlinear Dynamics:
  - Switching Linear Dynamical Models
  - Gaussian Process Dynamical Models
  - Sampling-based methods (i.e., Particle Filters)

Reading List

- Pavlovic et al. Learning Switching Linear Models of Human Motion
- Lawrewnce et al. Gaussian Process Latent Variable Models for Visualisation of High Dimensional Data
- Fleet, Motion Models for People Tracking.
- Ramakrishna et al., Reconstructing 3D Human Pose from 2D Image Landmarks, 2012.