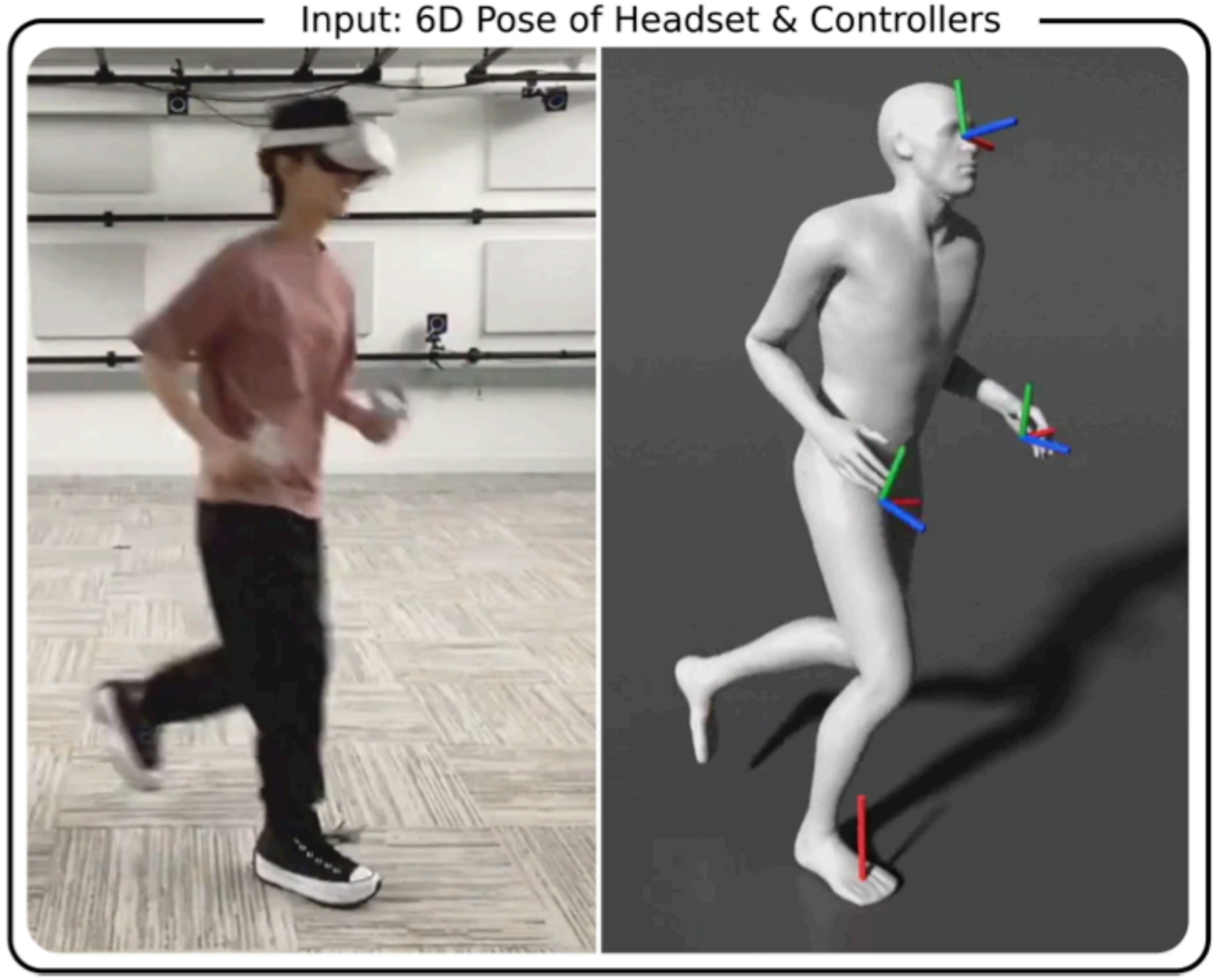
Physical Digital Humans in the Era of GenAI

May 2 @ GAMES Seminar

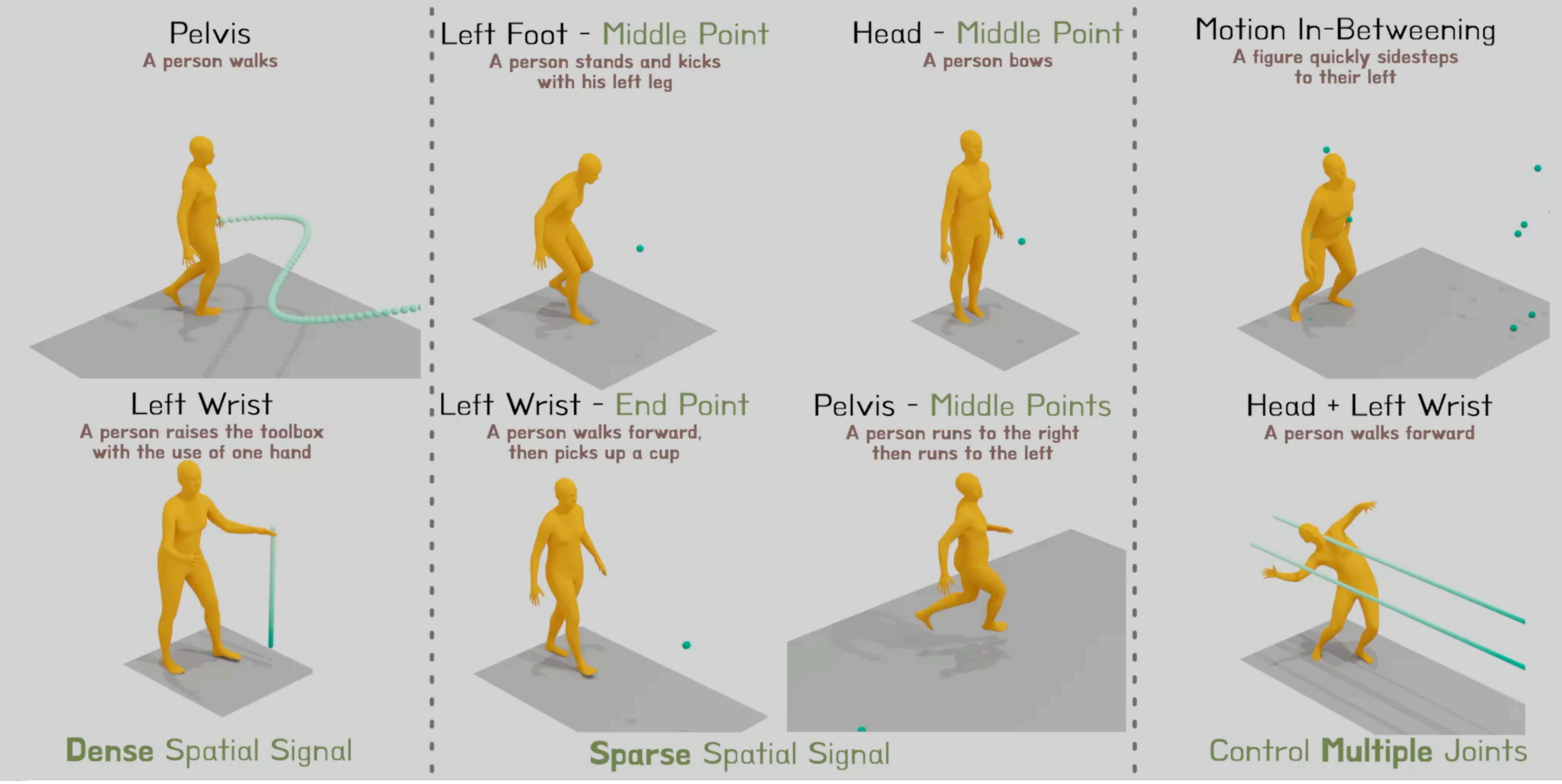
Yifeng Jiang
Ph.D. Candidate
Stanford University

2D Generative AI (of humans)

3D GenAI for XR/Spatial Computing/Simulation

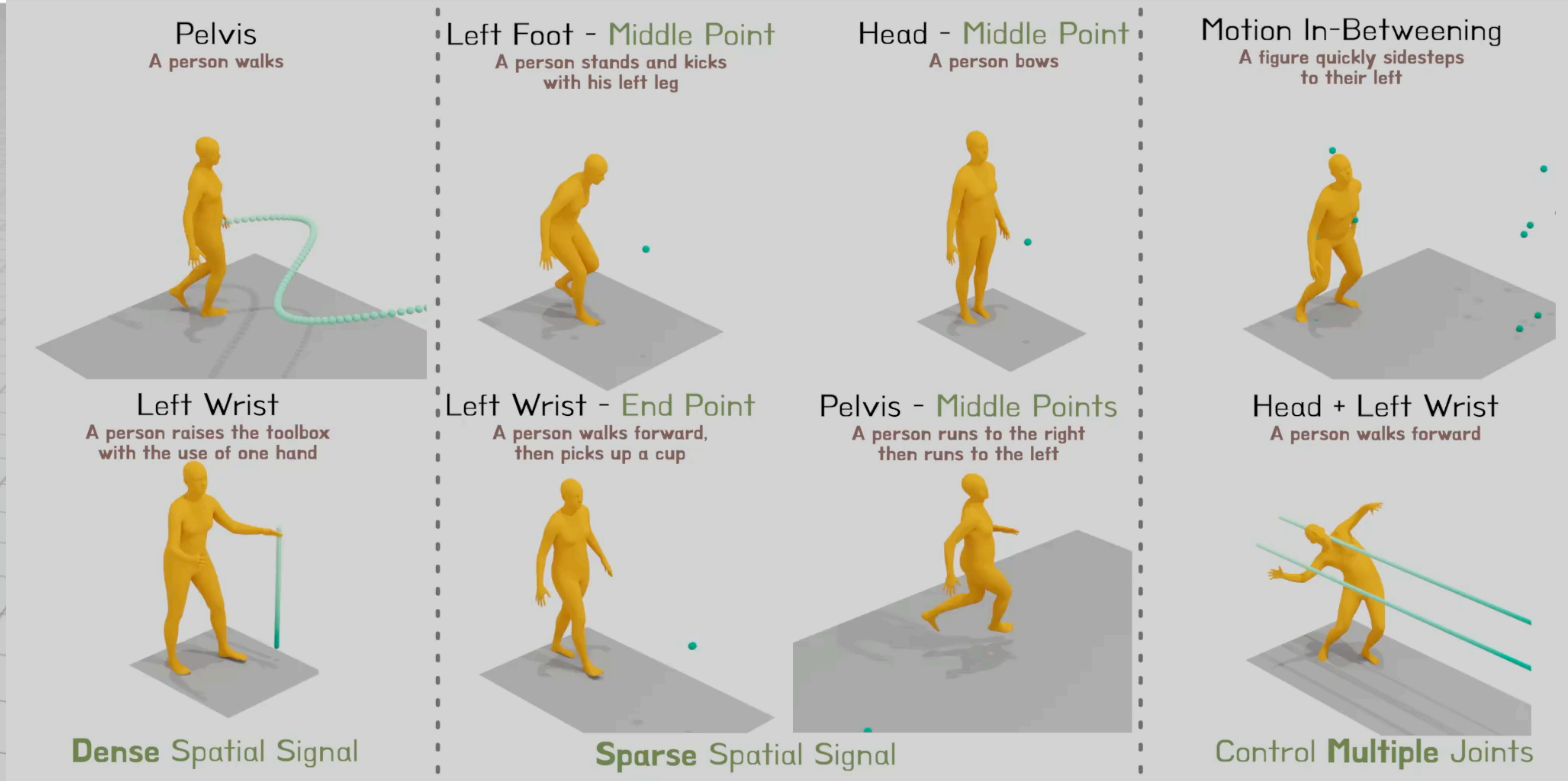


3D GenAI Digital Humans



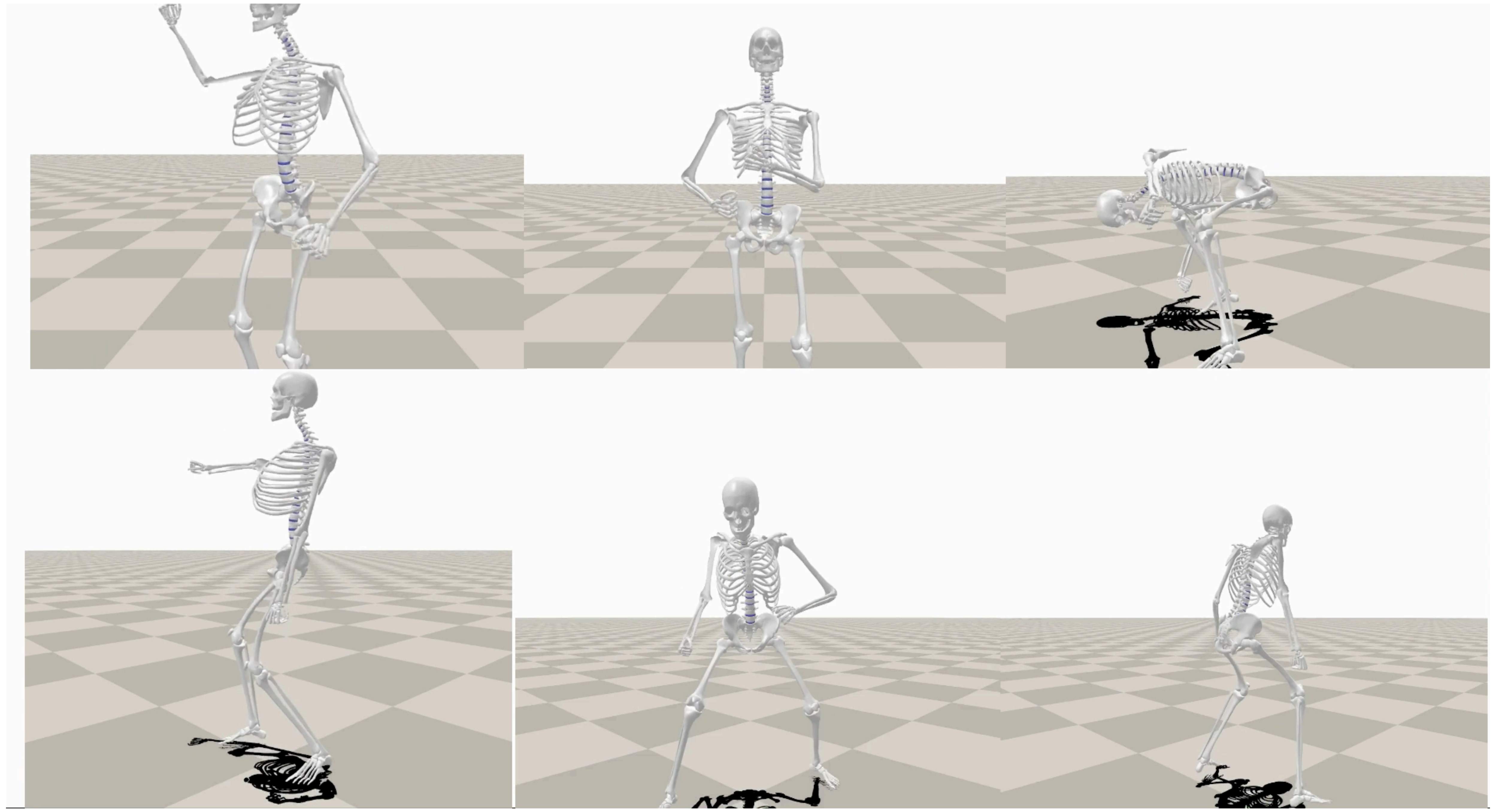
Full controllability of appearance and motions

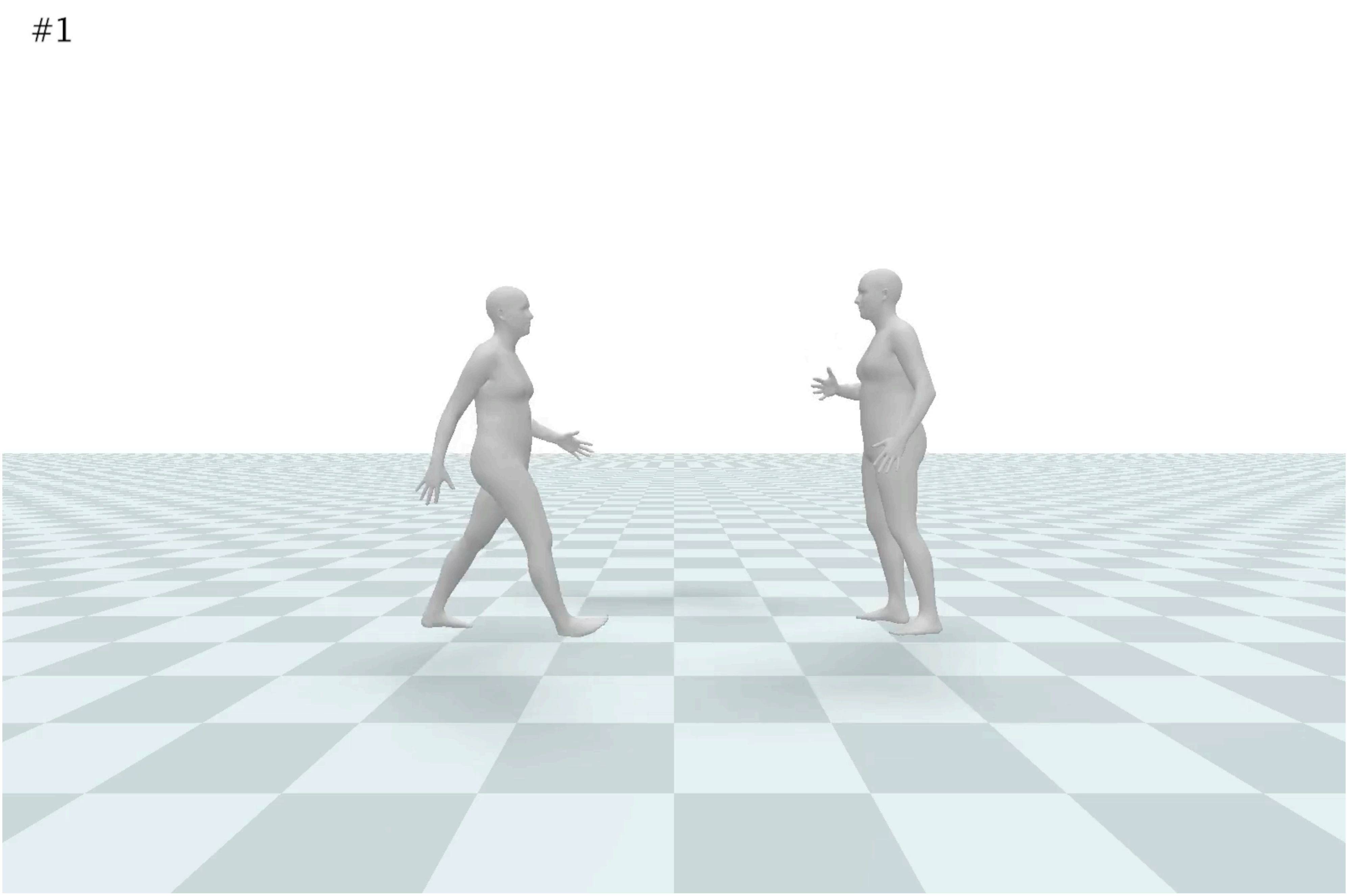
3D GenAI Digital Humans



Full controllability of appearance and motions

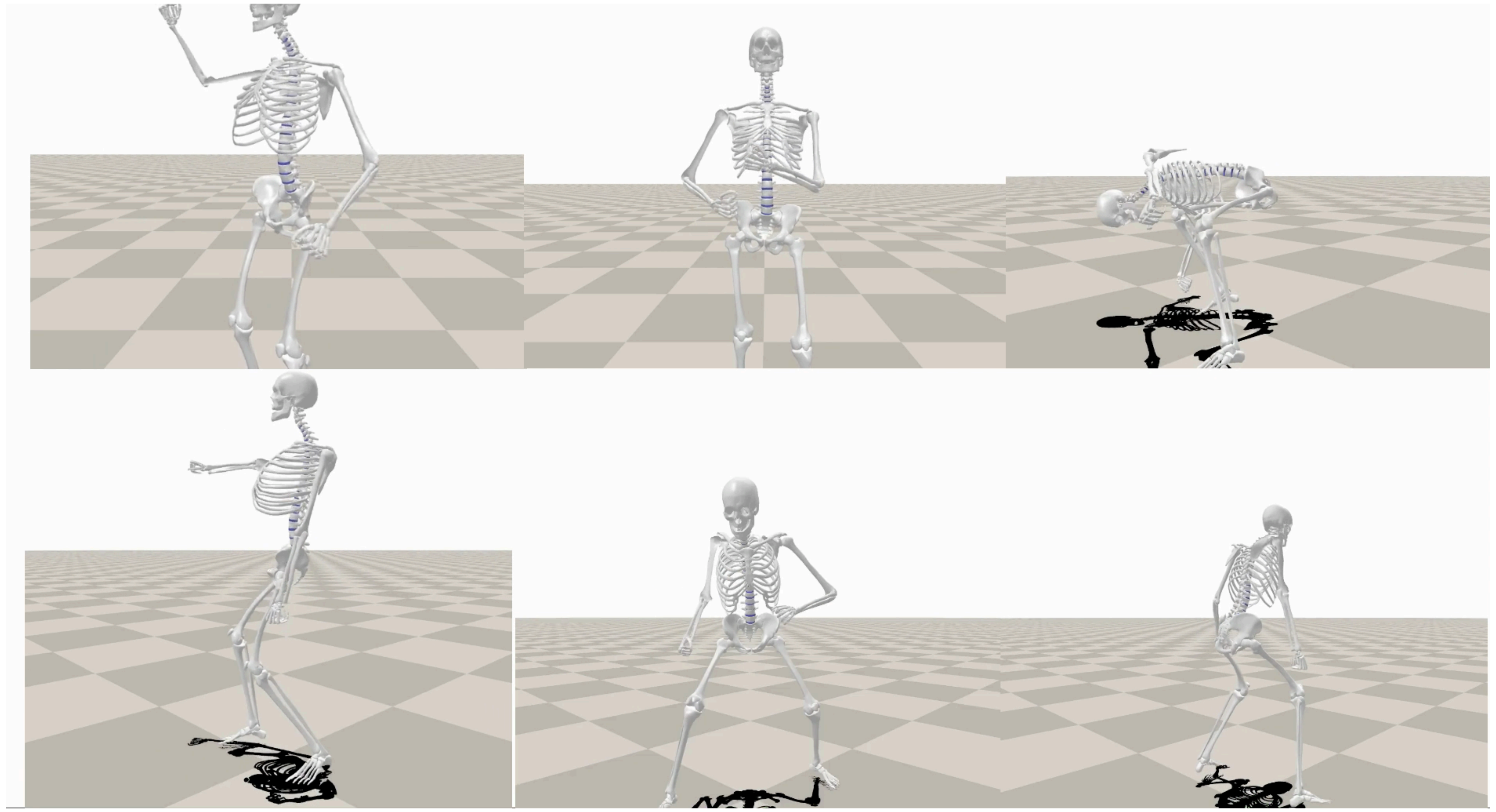
Next frontier: GenAI for 3D Physical Humans

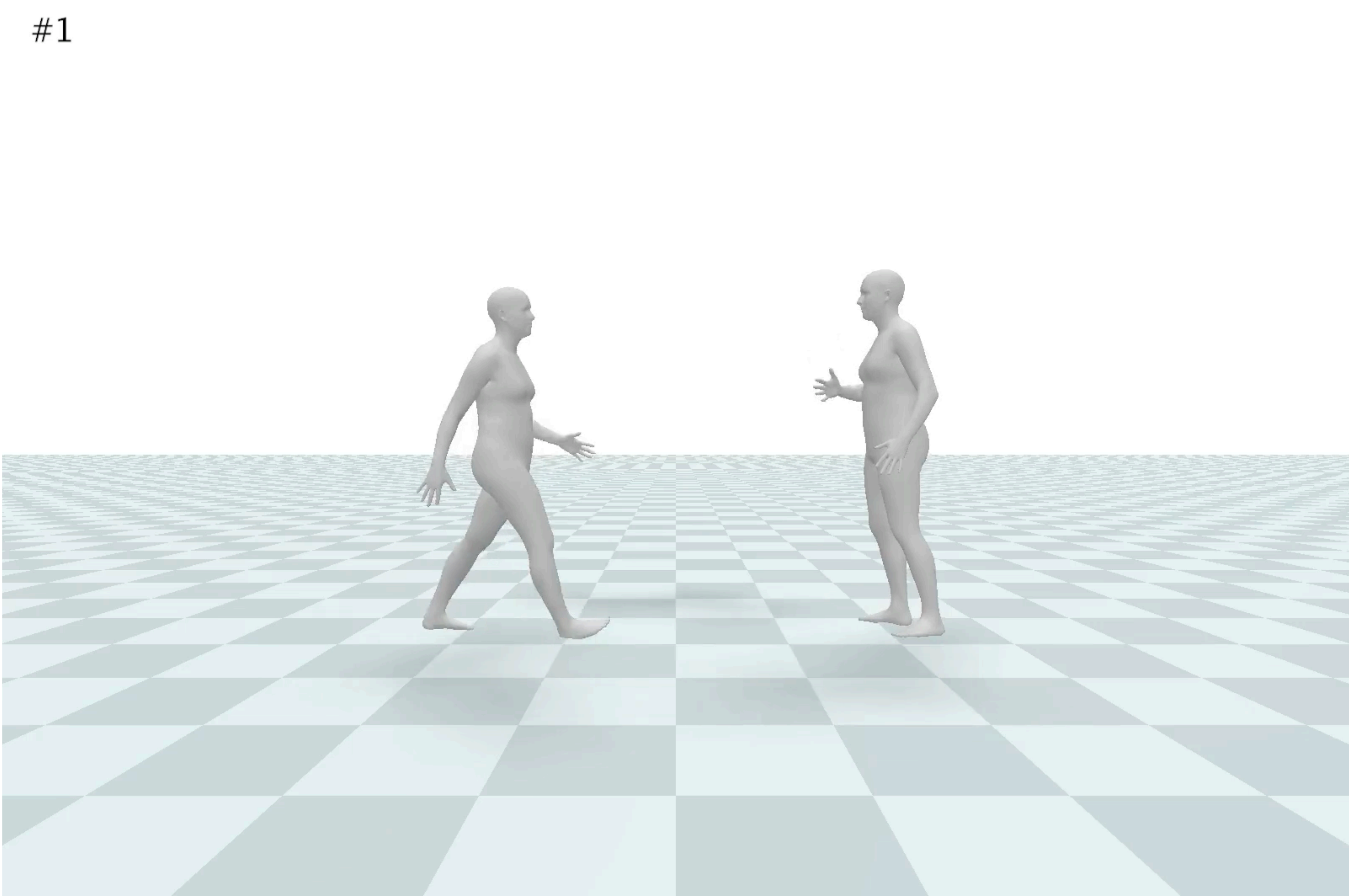




Bio & Physics modeling can augment detailed realism of generation

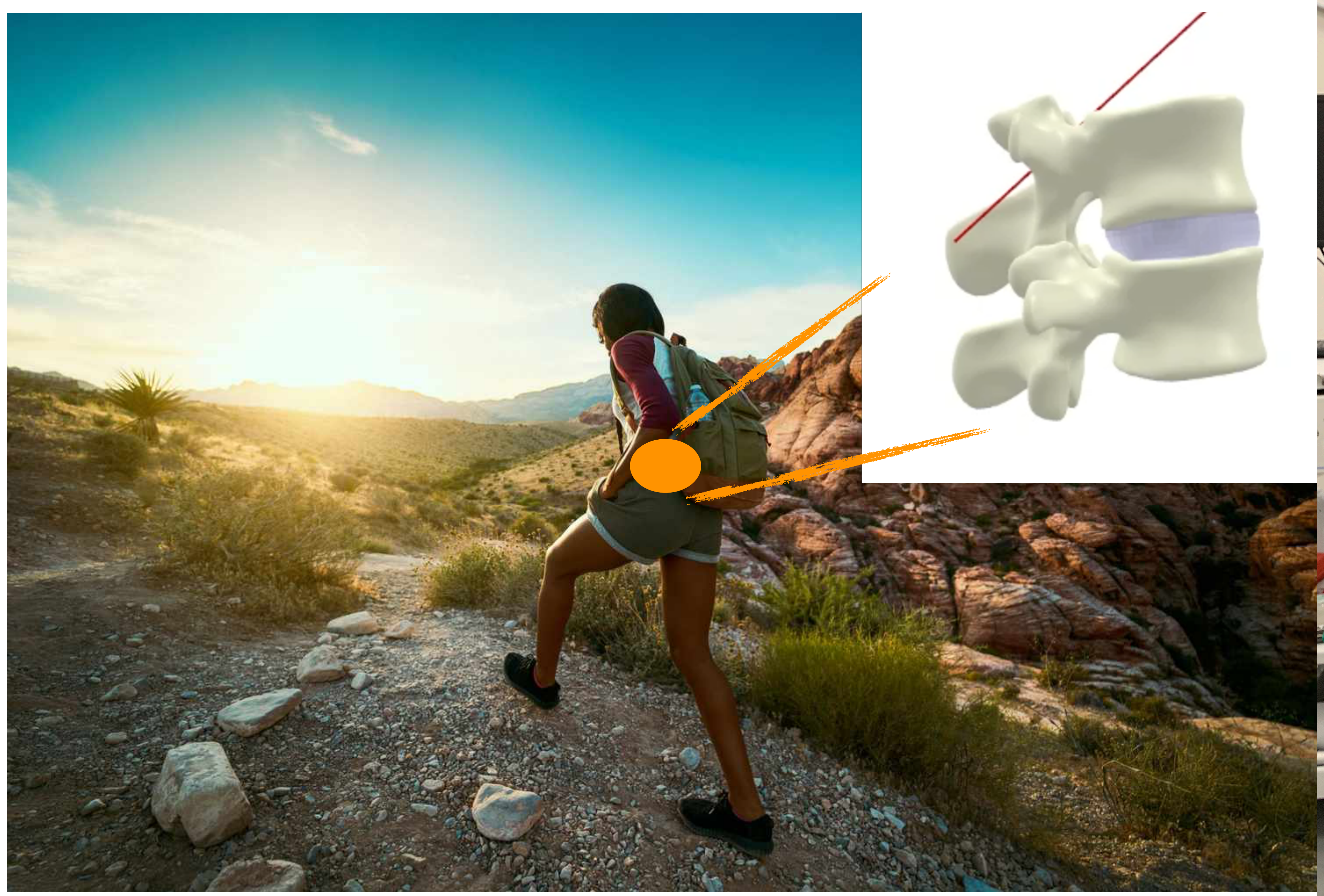
Next frontier: GenAI for 3D Physical Humans





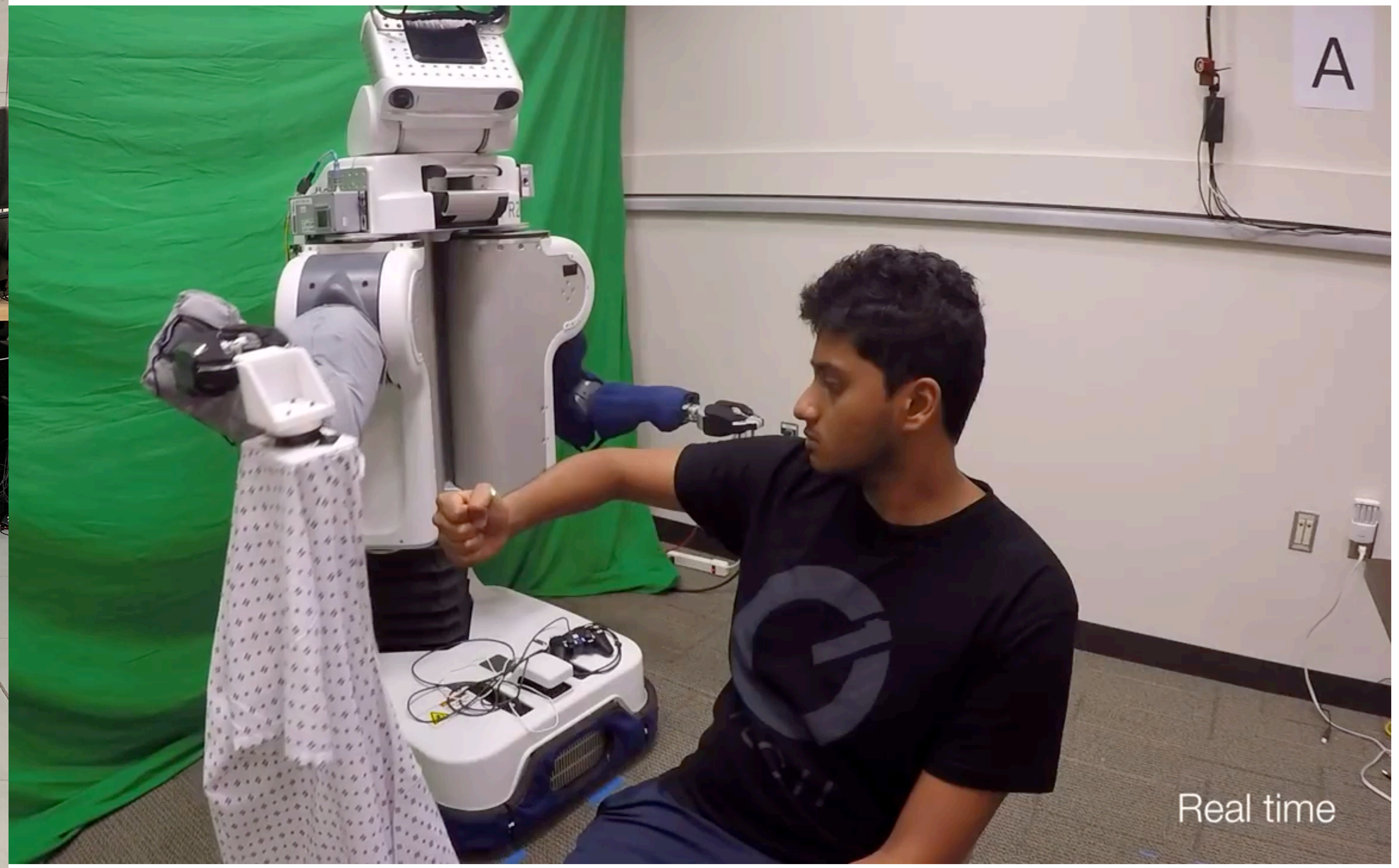
Bio & Physics modeling can augment detailed realism of generation

GenAI for Physical Humans: Also Many Real-world Applications



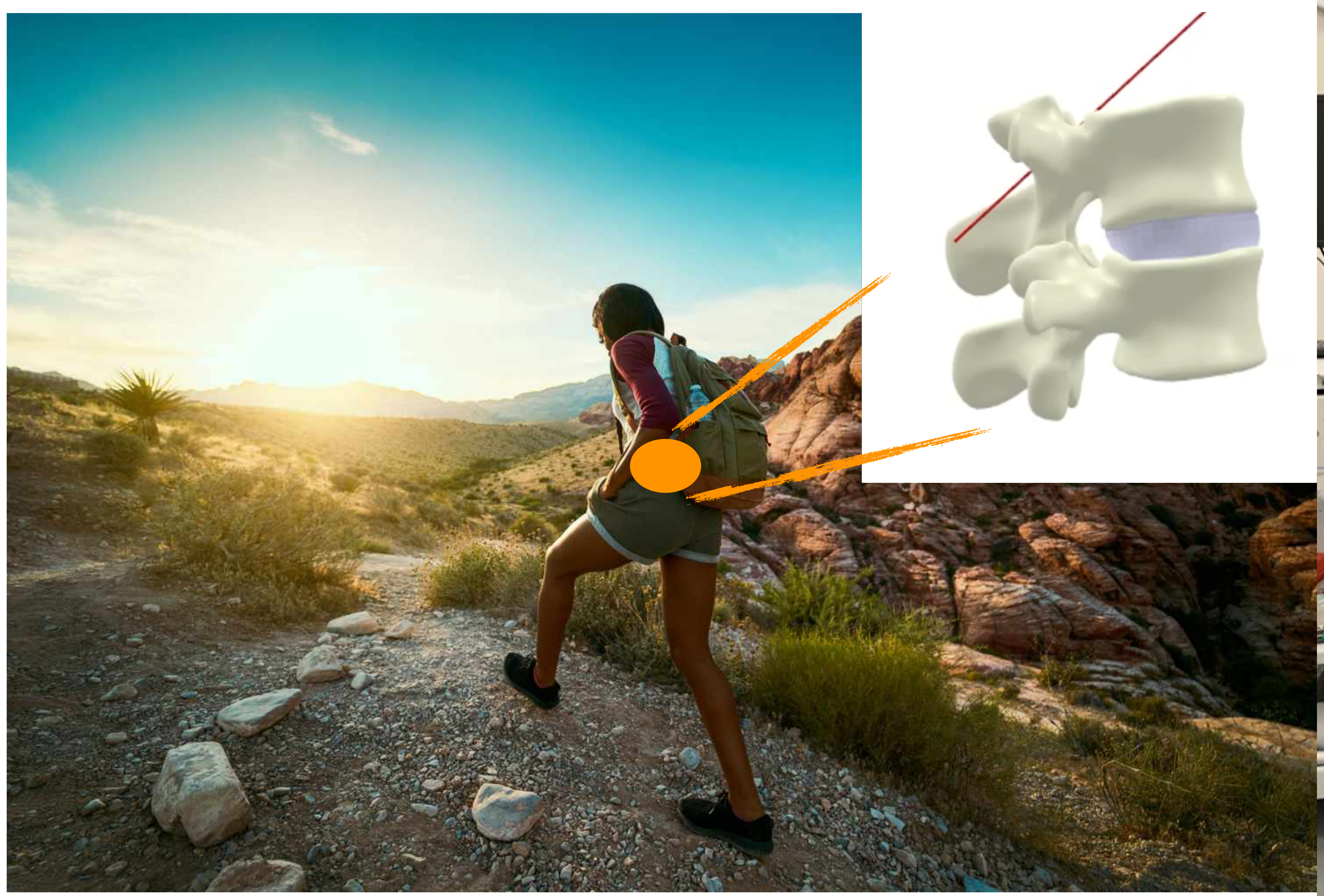
Disc force for injury prevention

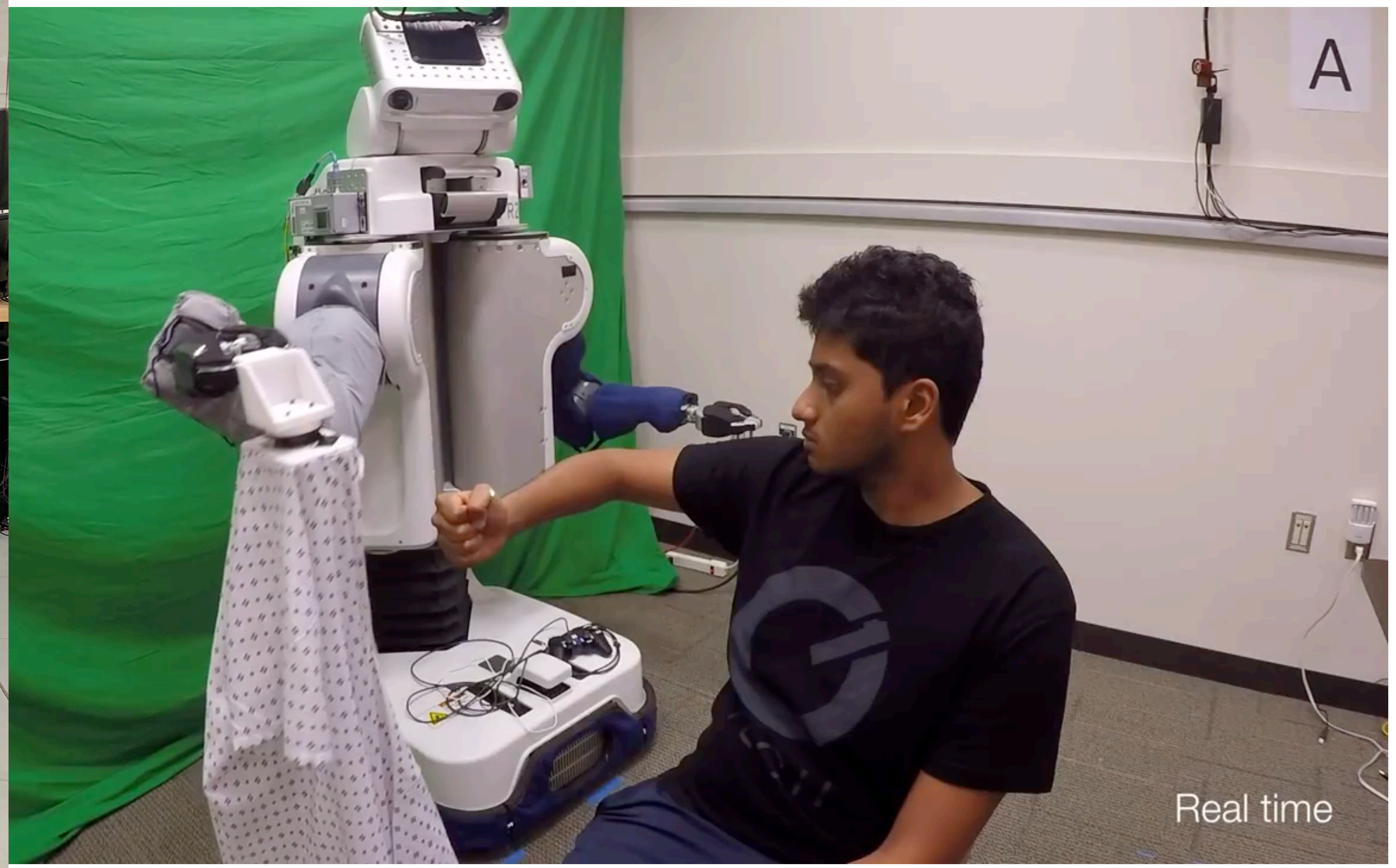
Knee load for Exoskeleton



Comfort level during dressing

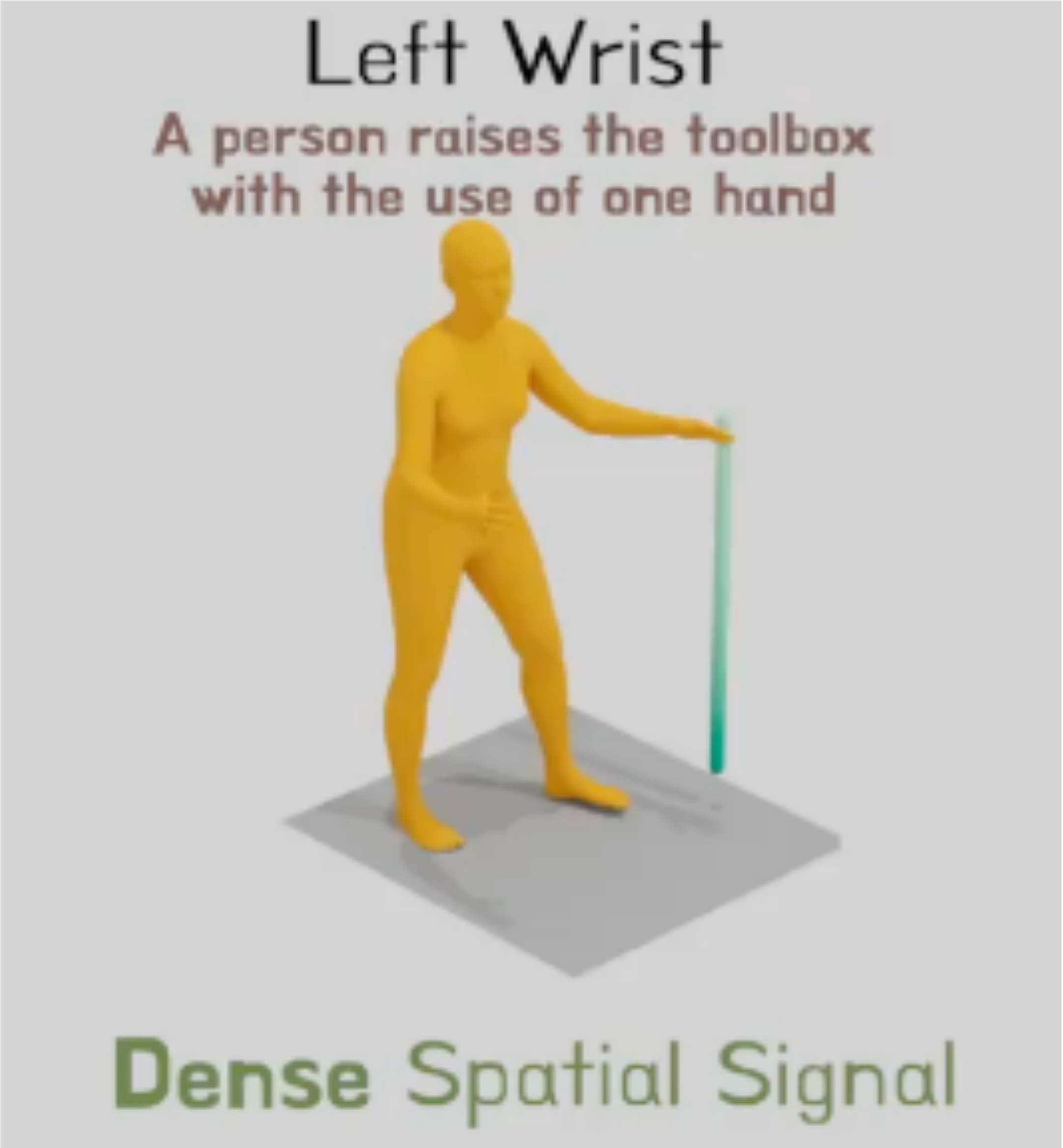
GenAI for Physical Humans: Also Many Real-world Applications





Disc force for injury prevention **Knee load for Exoskeleton** **Comfort level during dressing**

2D Gen-AI

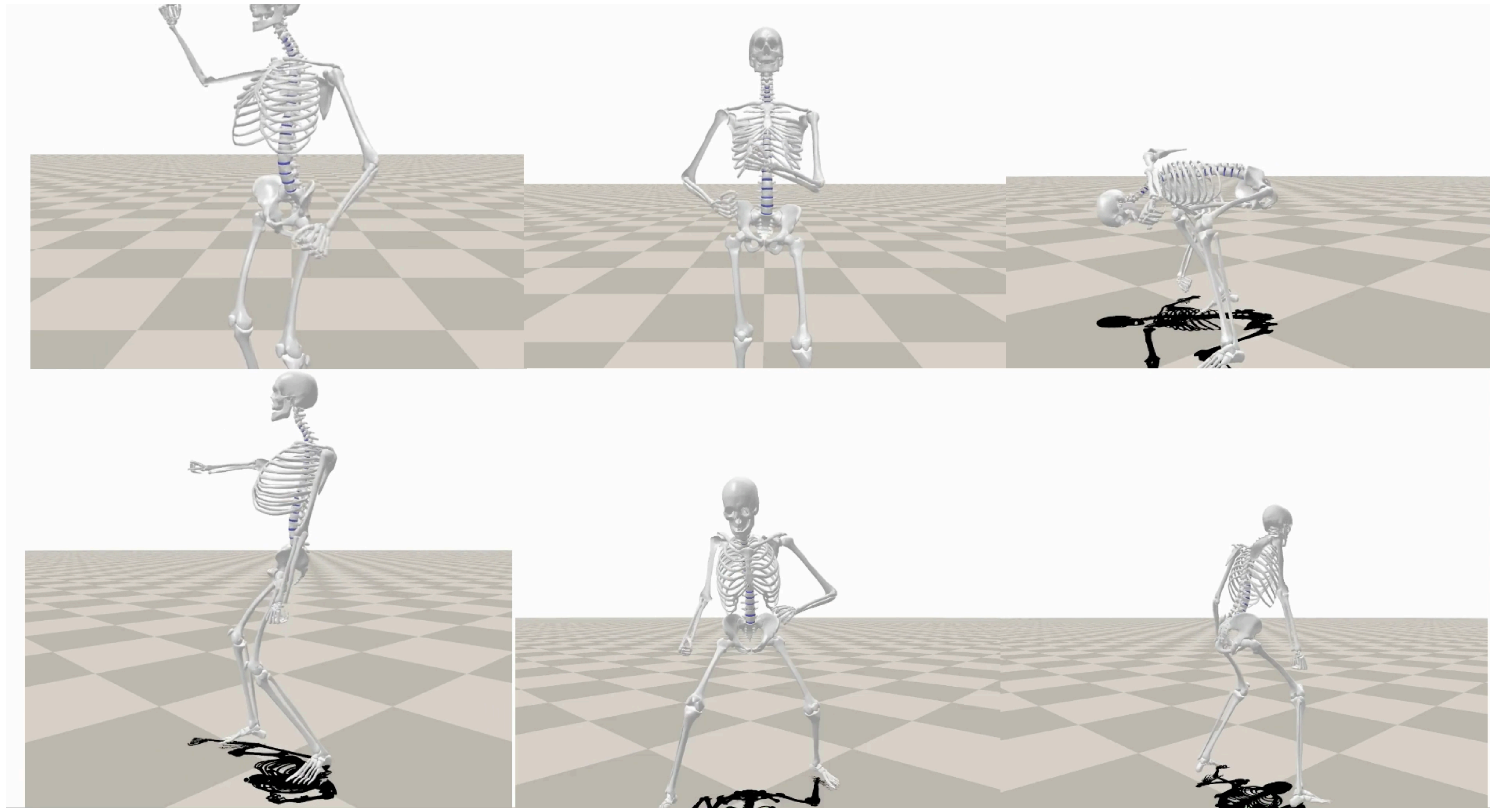


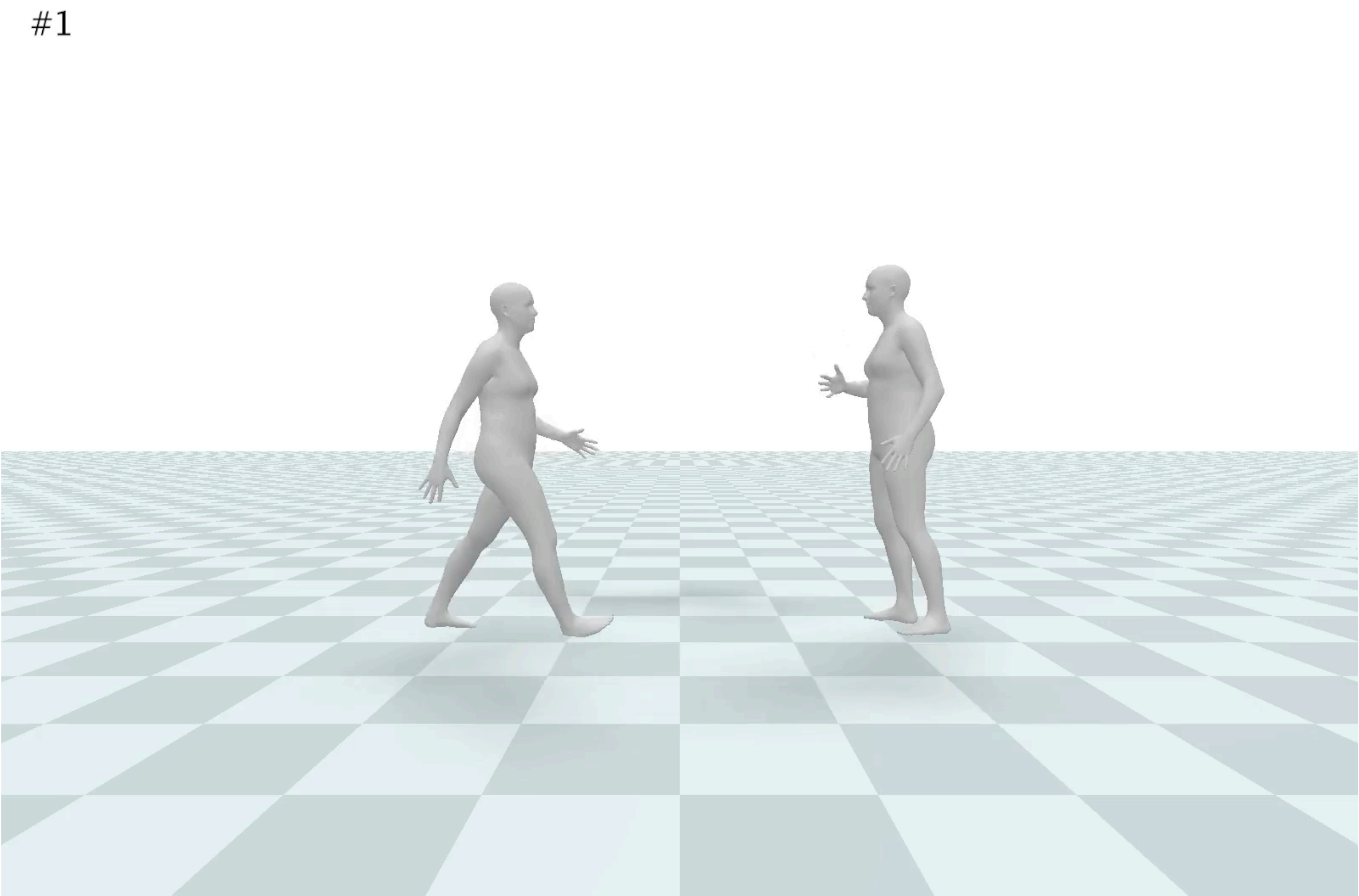
3D Gen-AI

Physical 3D Gen-AI

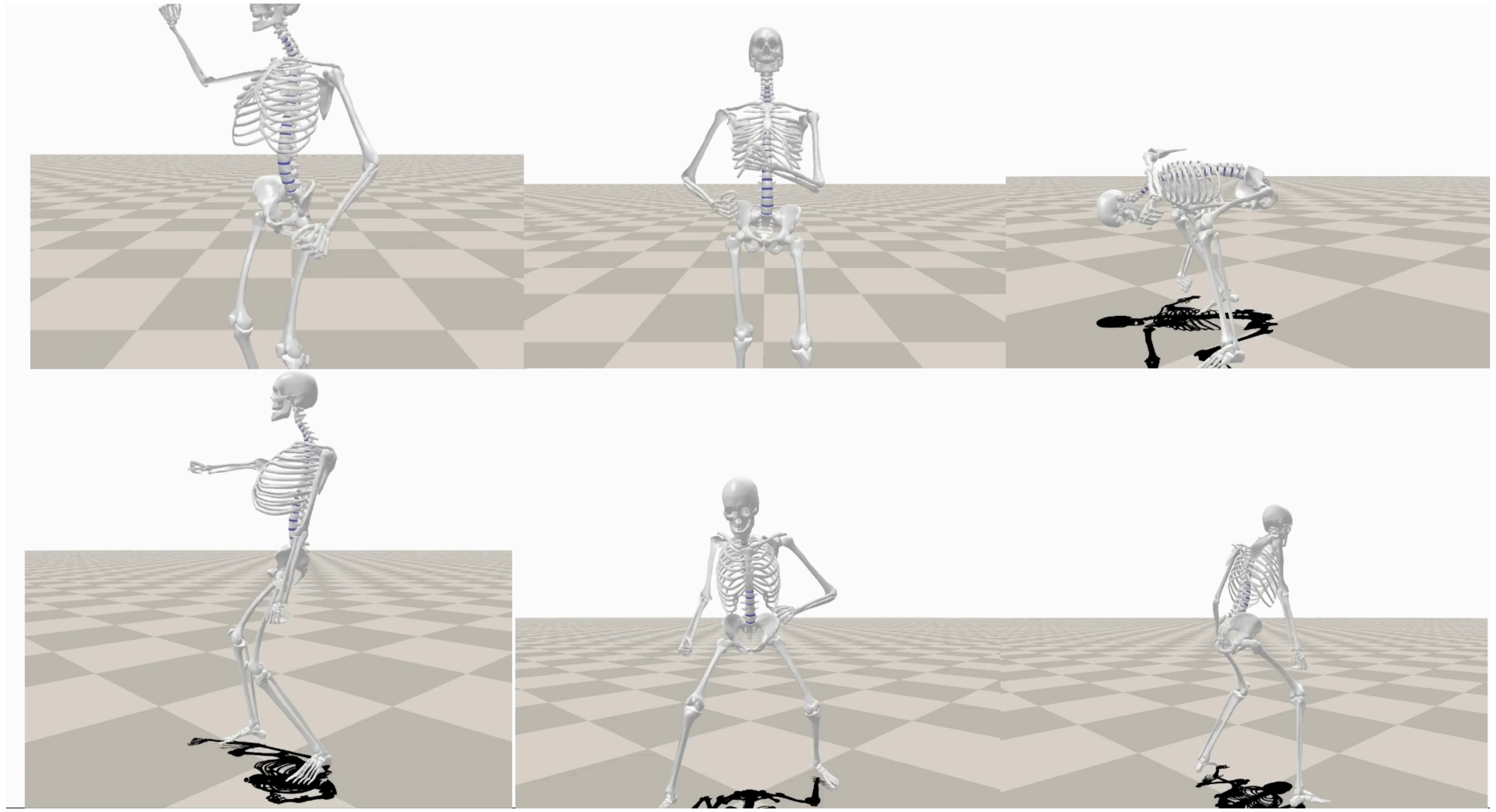
From Digital to Physical-world Applications

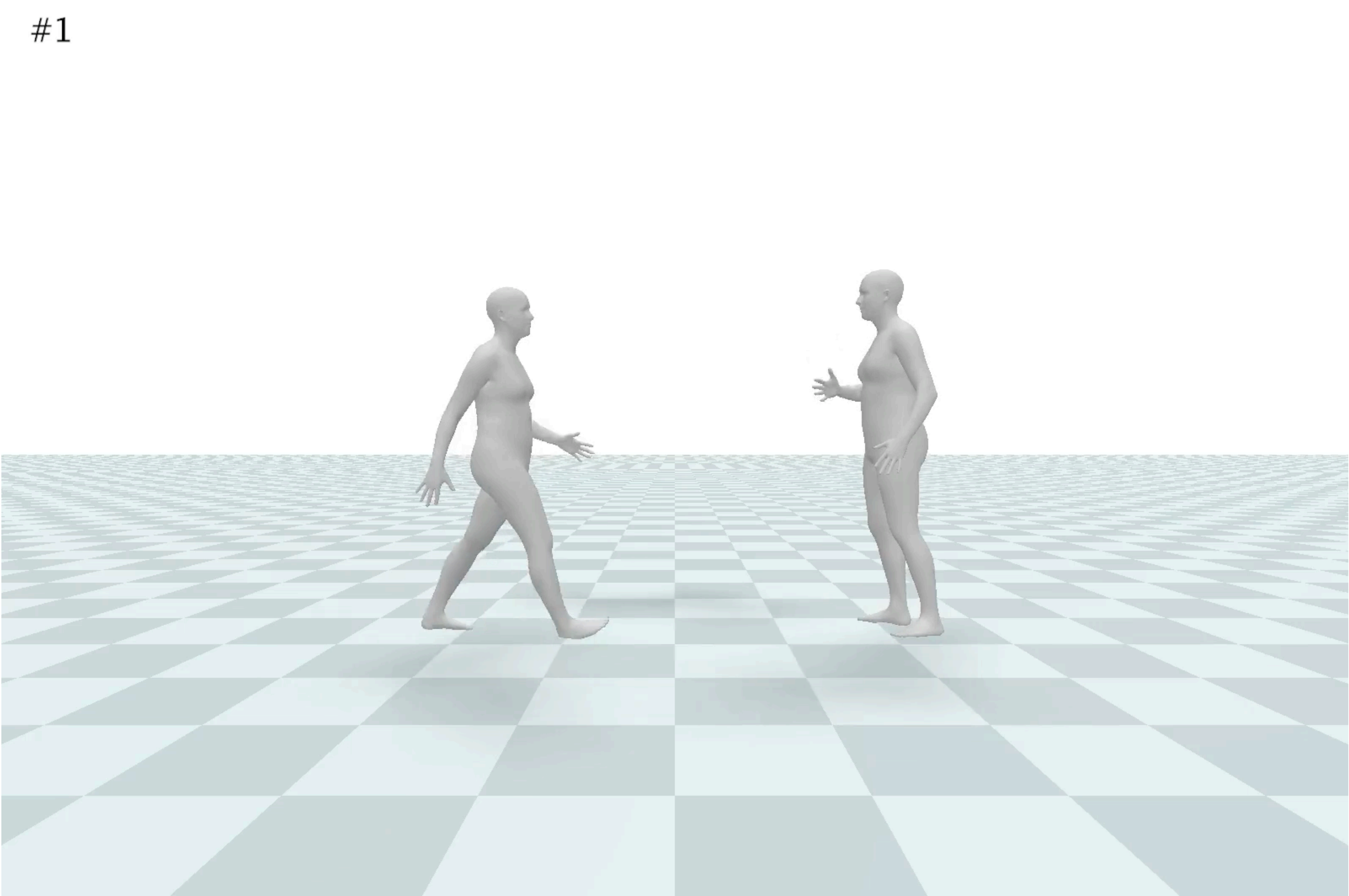
More challenging to obtain large-scale high-quality data





Data can be partially observable, scarce, expensive/unsafe to capture





Data can be partially observable, scarce, expensive/unsafe to capture

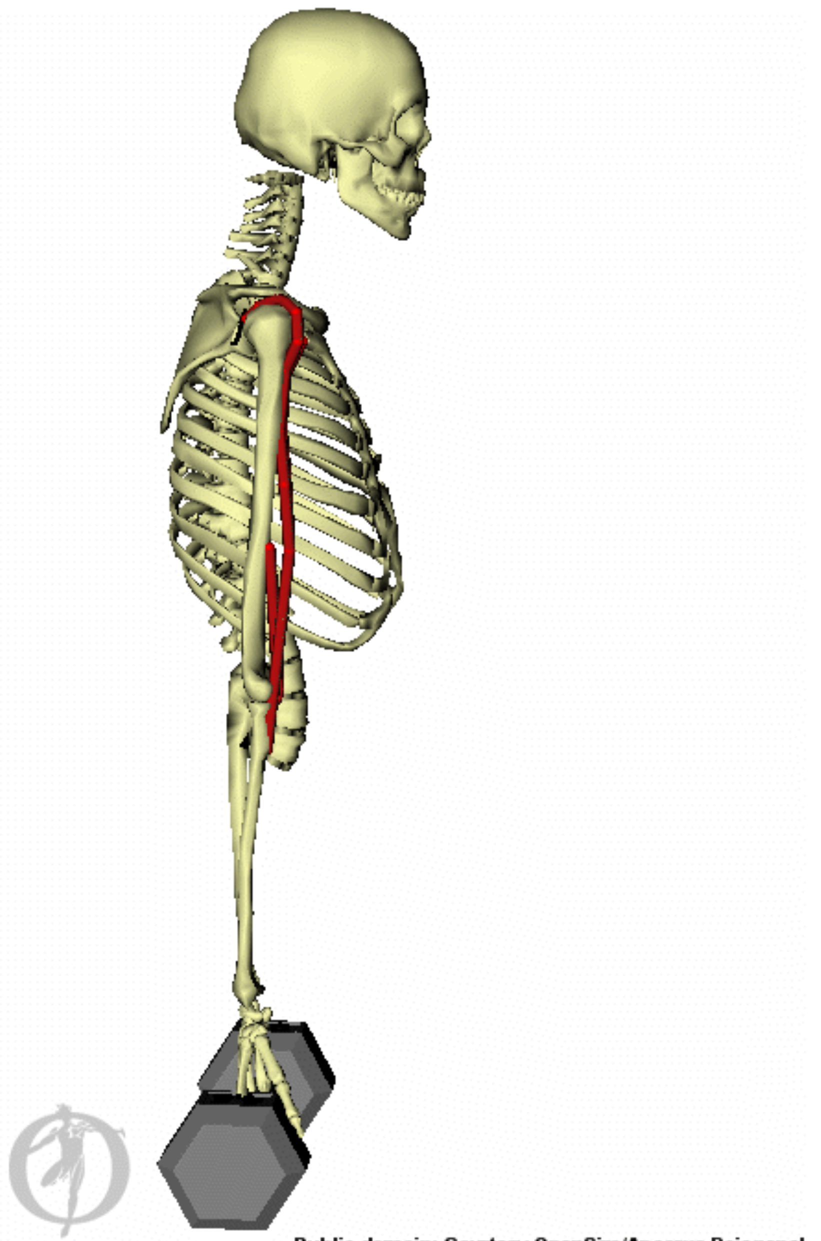
 

Physical Digital
Human

Real Human Data

Synthesized Human Data

Physical Digital Human

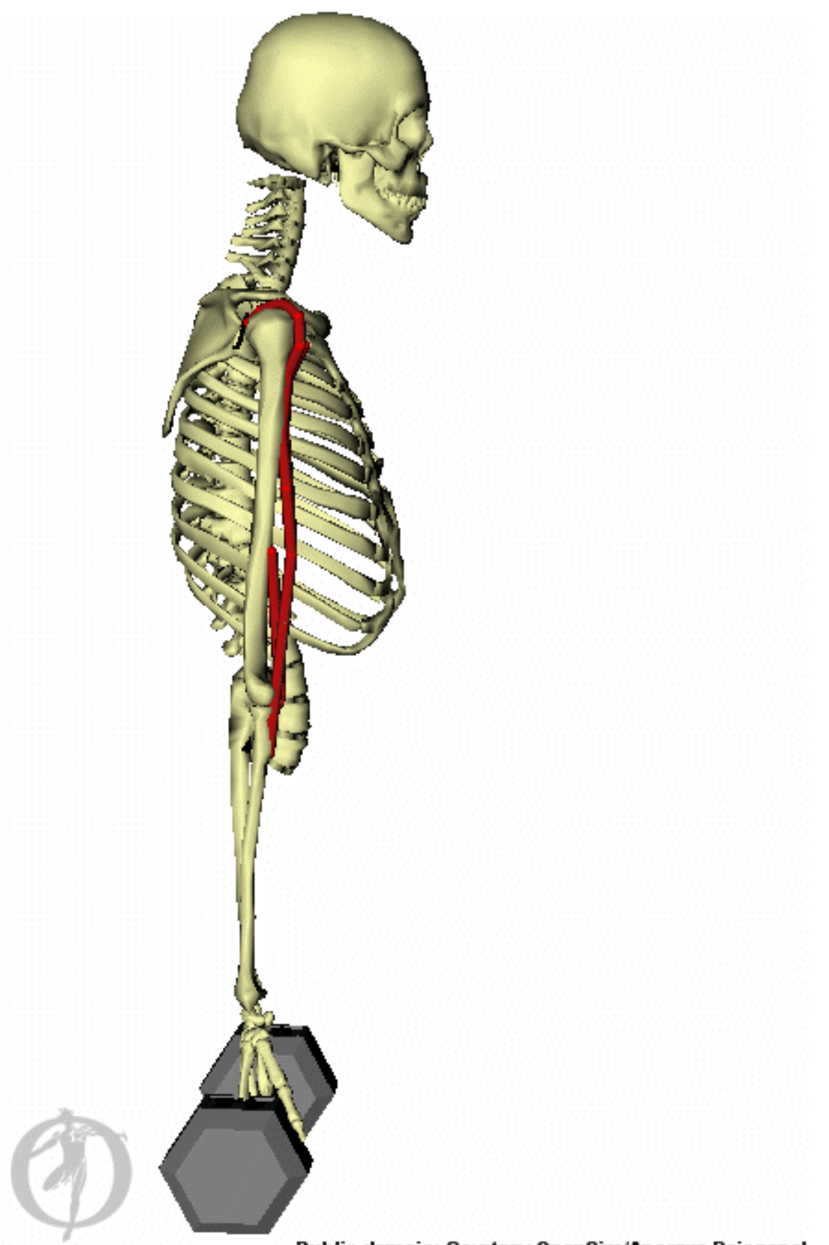


Public domain: Courtesy OpenSim/Apoorva Rajagopal

Real Human Data

Synthesized Human Data

Physical Digital Human



Public domain: Courtesy OpenSim/Apoorva Rajagopal

Modern Deep Learning

Physics Simulation

Physical Digital
Human

Modern Deep Learning

Physics Simulation

Part 1: Scalable Human Simulation with Learned Components

Modern Deep Learning ← Physics Simulation

Part 1: Scalable Human Simulation with Learned Components

Part 2: Simulation-augmented Generative Motion Model

Part 3: Scalable Physical Human Data Capture

Scalable Human Simulation with Learned Components

— How to accurately simulate human without explicit anatomy details

[Jiang et al] SIGGRAPH'19

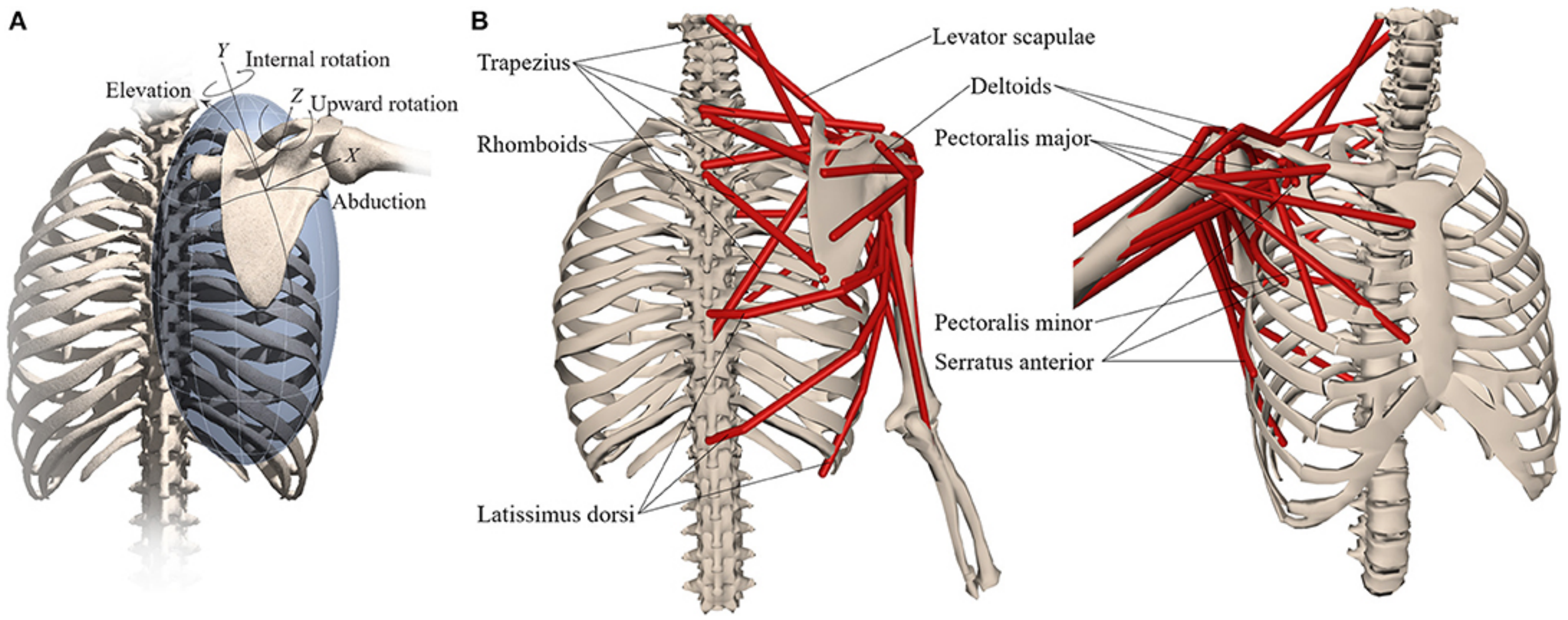
Standard Simulation Model

e.g. SMPL

23 ball-and-socket joints

Easy to simulate,
but not biomechanically accurate

Detailed Biomechanics Models & Simulations



Not fast & robust enough for large-scale training & synthetic data generation

https://www.frontiersin.org/files/Articles/469805/fnbot-13-00090-HTML/image_m/fnbot-13-00090-g001.jpg

The Tale of Two Simulation Spaces

Detailed, Anatomical

Simple, abstract

The Tale of Two Simulation Spaces

Detailed, Anatomical **Simple, abstract**

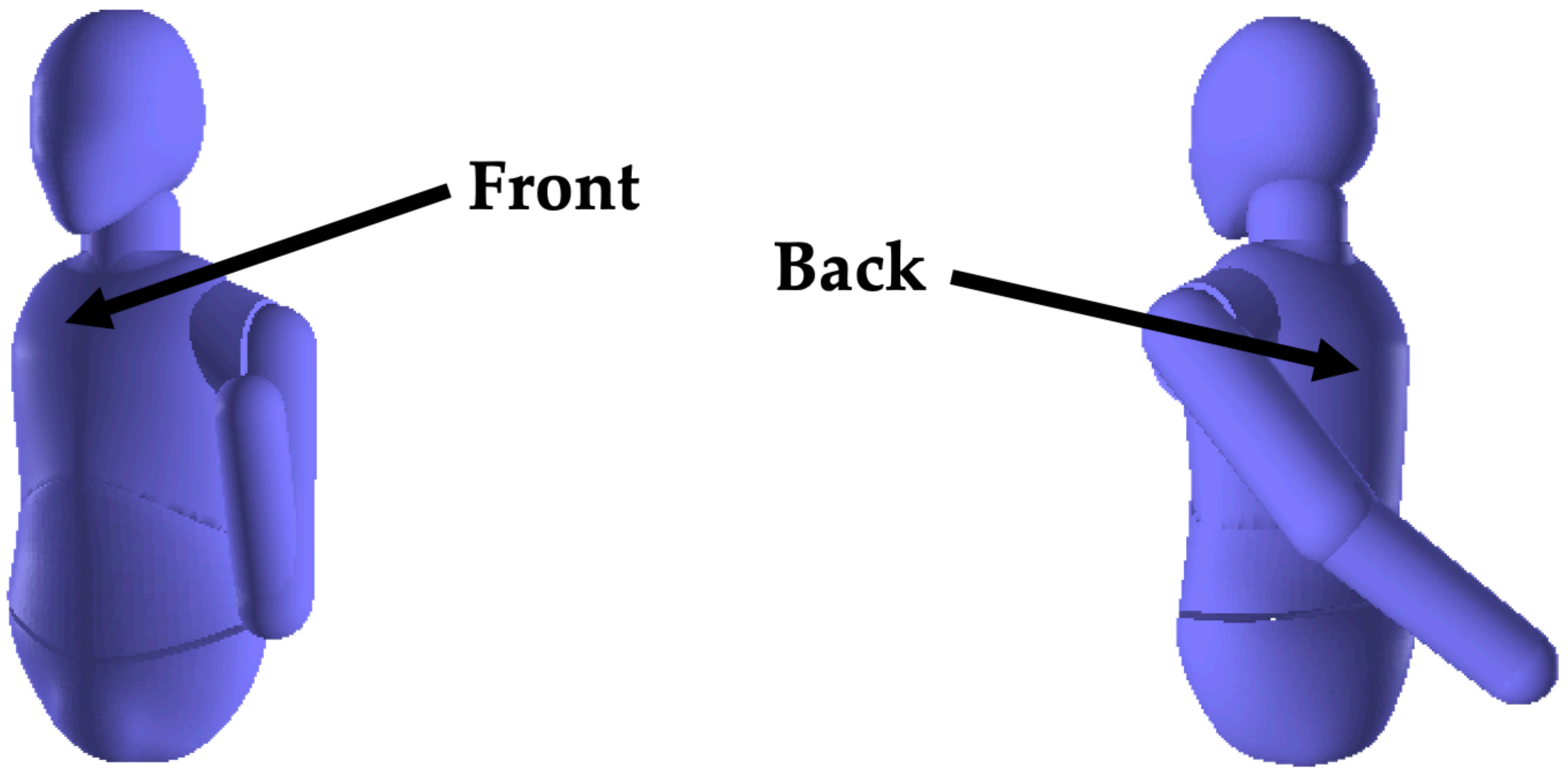
The Tale of Two Simulation Spaces

Detailed, Anatomical **Simple, abstract**

Current discrepancies?

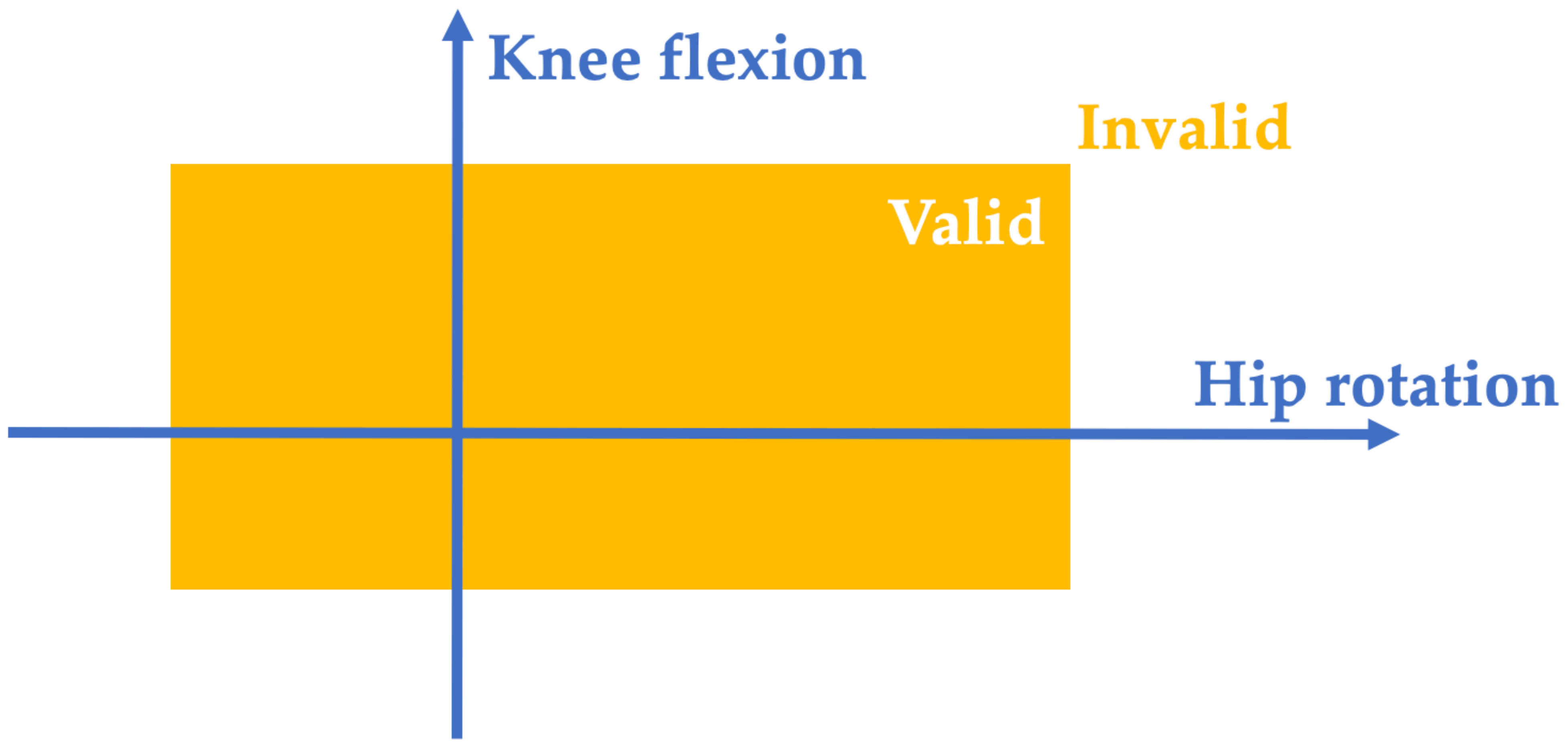
Example #1: joint limit (RoM) depends on other joints

Smaller elbow range when the arm is behind the back.



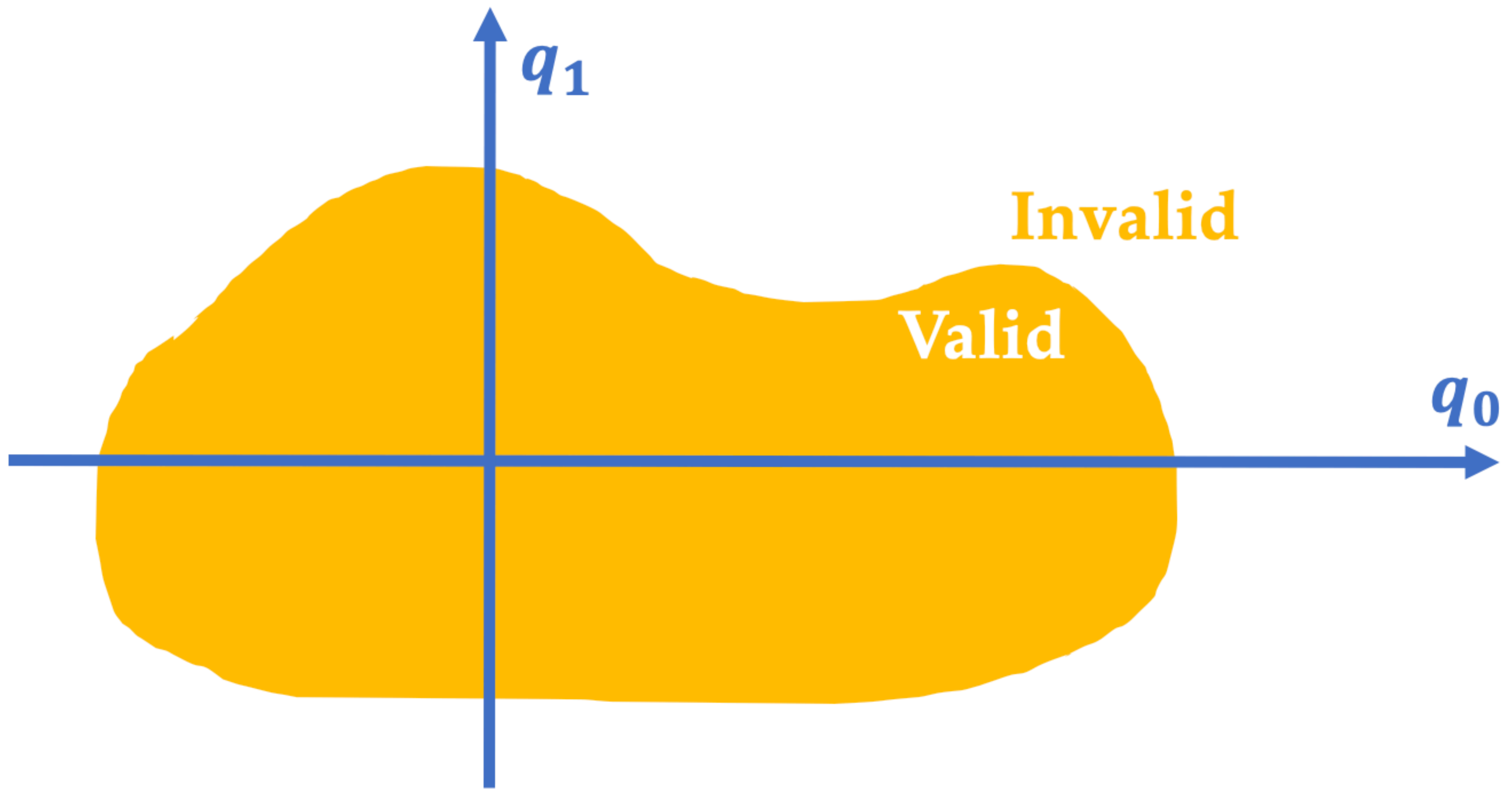
Example #1: joint limit (RoM) depends on other joints

Heuristic Boxed Limits



$$\boldsymbol{q}_{low} \le \boldsymbol{q} \le \boldsymbol{q}_{high}$$

Realistic "state-dependent" Joint Limits



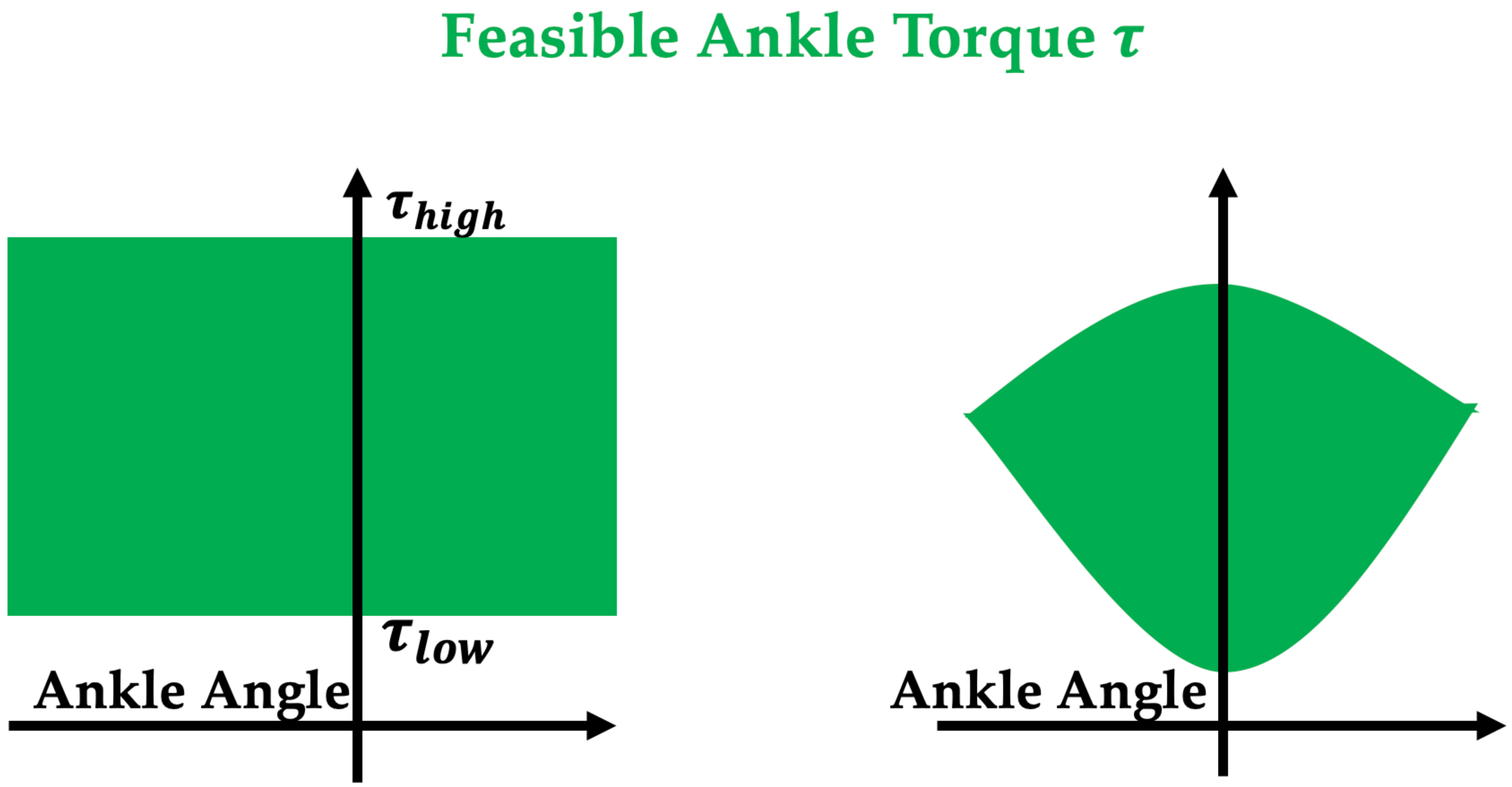
$$f_{bone-ligaments}(\boldsymbol{q}, \boldsymbol{l})$$

Example #2: torque capability is state-dependent

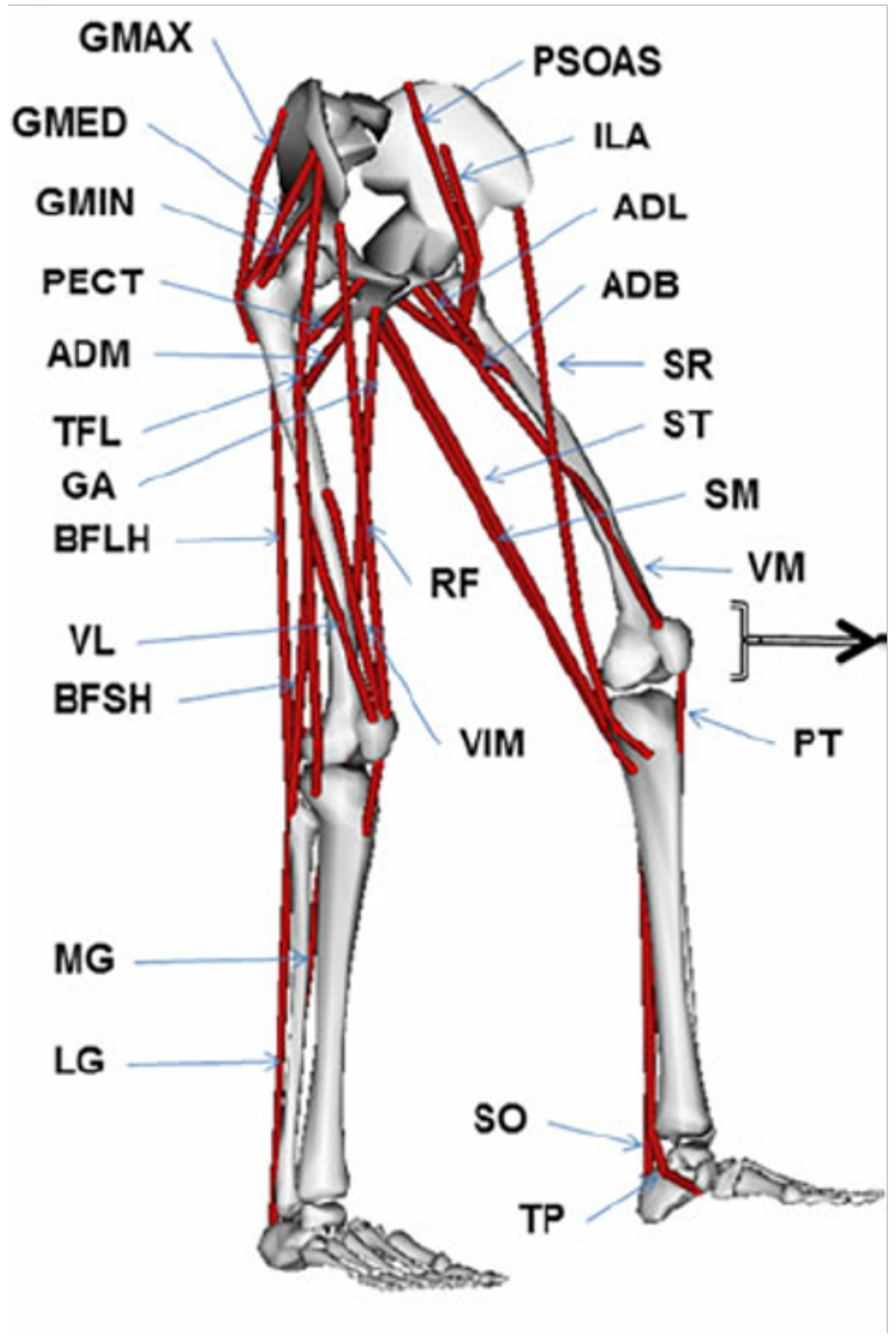
Self-defense

Heuristic Boxed Limits

State-dependent Joint Limits



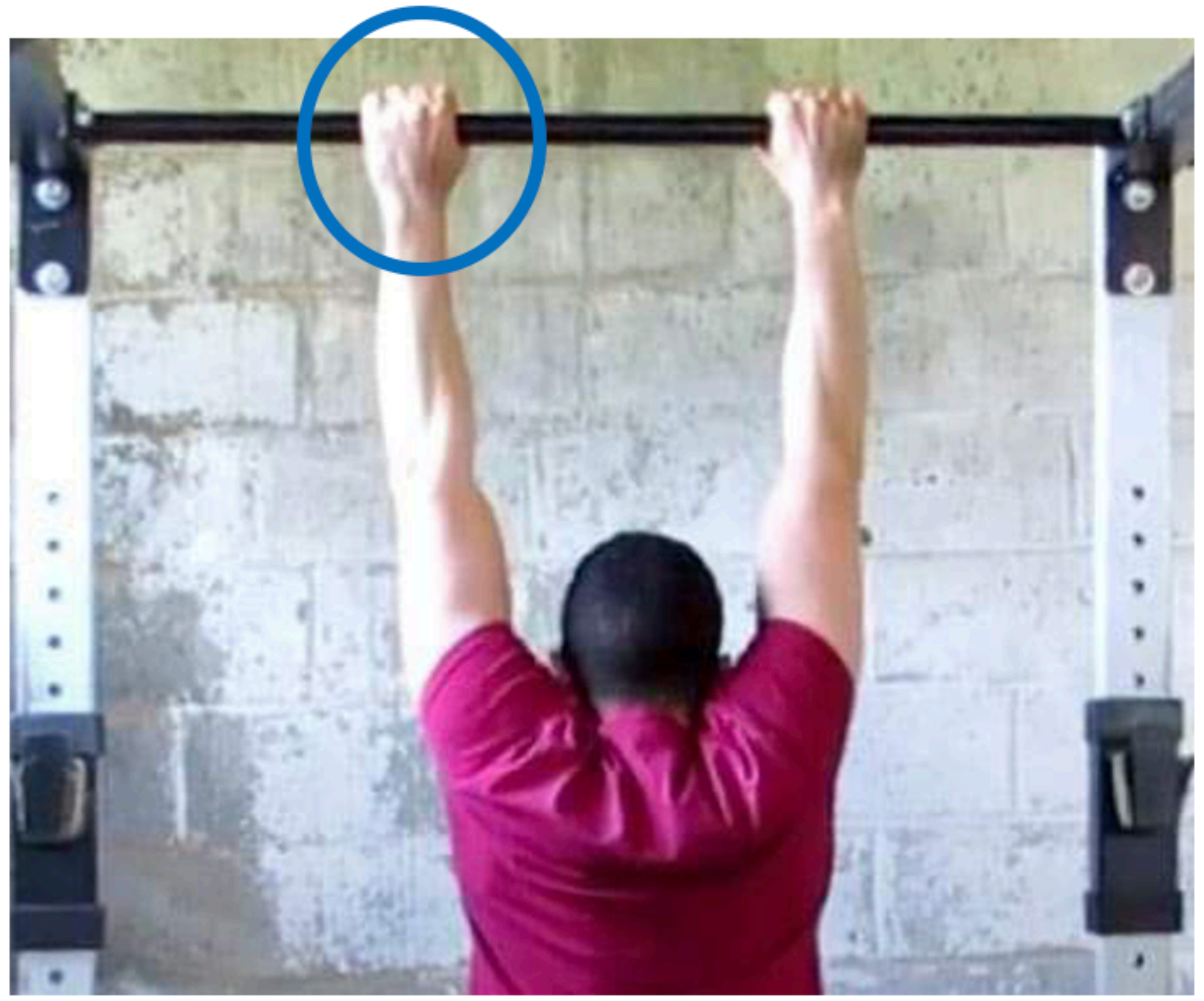
Example #2: torque capacity also depends on other joints



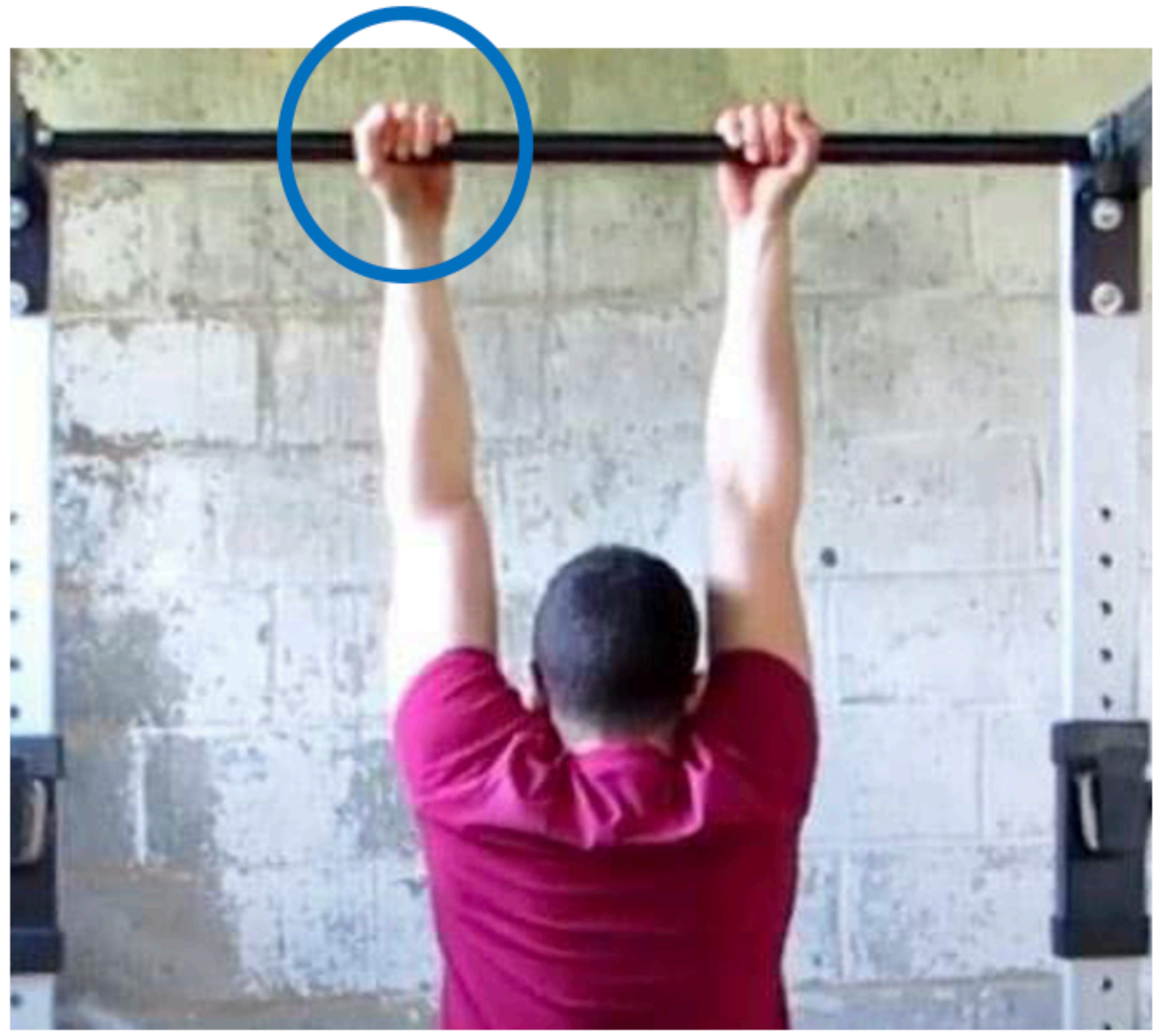
Each muscle spans multiple joints, and multiple muscles interplay at each joint

Example #3: metabolic rate is state dependent

"Same torque, different effort"



Pull-up



Chin-up

Standard Motion Control Formulation in "SMPL" Space

General to any task and task objective c_{task}

$\boldsymbol{\tau}$: Joint Torques

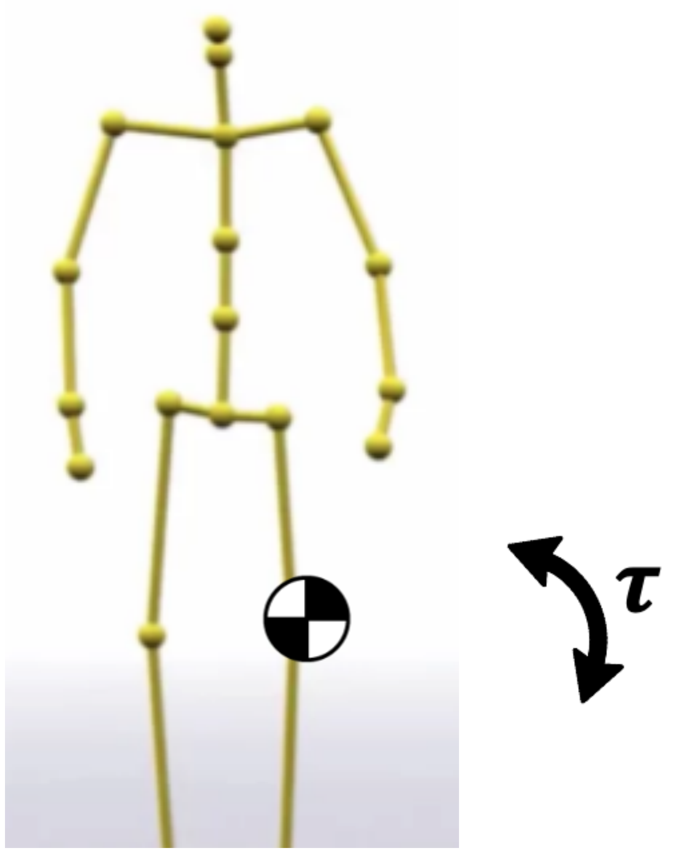
$$\min_{\boldsymbol{\tau}} \quad \sum \tau^2 + \boxed{c_{task}(q)}$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boldsymbol{\tau}_{low} \leq \boldsymbol{\tau} \leq \boldsymbol{\tau}_{high}$$

$$\boldsymbol{q}_{low} \leq \boldsymbol{q} \leq \boldsymbol{q}_{high}$$



Standard Motion Control Formulation in "SMPL" Space

Control / Energy

Regularization

$\boldsymbol{\tau}$: Joint Torques

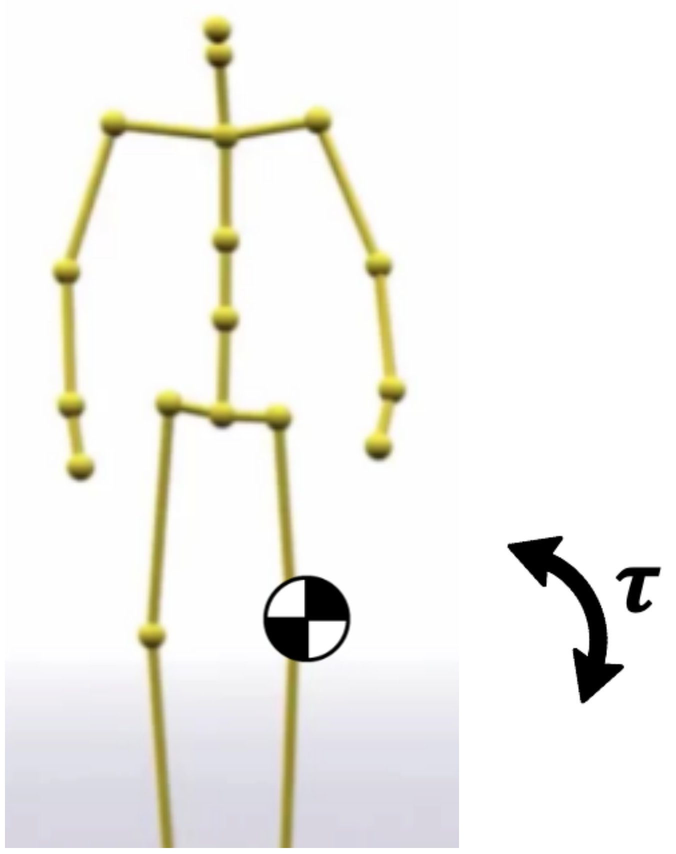
$$\min_{\boldsymbol{\tau}} \quad \boxed{\sum \tau^2} + c_{task}(q)$$

subject to

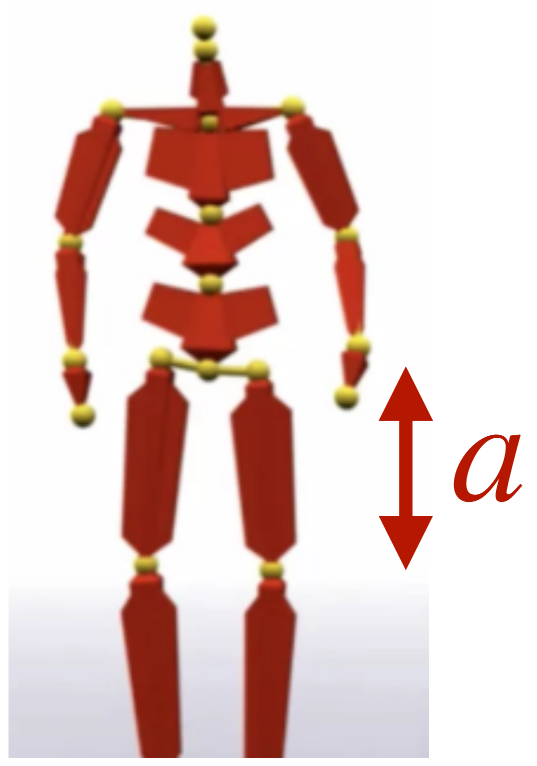
$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boldsymbol{\tau}_{low} \leq \boldsymbol{\tau} \leq \boldsymbol{\tau}_{high}$$

$$\boldsymbol{q}_{low} \leq \boldsymbol{q} \leq \boldsymbol{q}_{high}$$



In Comparison to Detailed Anatomical Simulation



$\boldsymbol{a}$: Muscle Activations

$$\min_{\boldsymbol{a}} \quad \boxed{\sum a^2} + c_{task}(q)$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boxed{\begin{array}{l} f_{muscle-dynamics}(\boldsymbol{a}, \boldsymbol{l}, \dot{\boldsymbol{l}}) \\ \mathbf{0} \le \boldsymbol{a} \le \mathbf{1} \\ f_{bone-ligaments}(\boldsymbol{q}, \boldsymbol{l}) \end{array}}$$

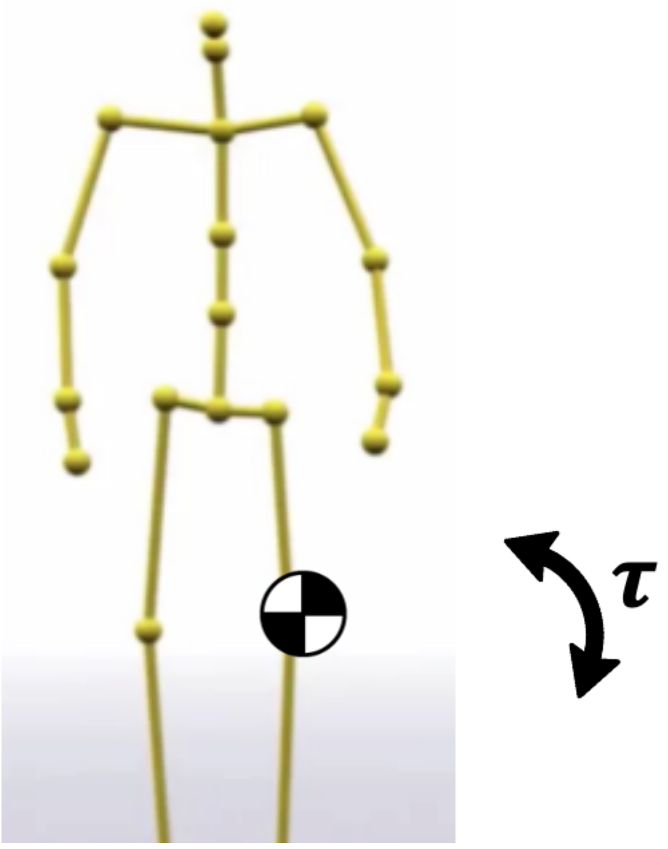
$\boldsymbol{\tau}$: Joint Torques

$$\min_{\boldsymbol{\tau}} \quad \boxed{\sum \tau^2} + c_{task}(q)$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boxed{\begin{array}{l} \boldsymbol{\tau}_{low} \le \boldsymbol{\tau} \le \boldsymbol{\tau}_{high} \\ \boldsymbol{q}_{low} \le \boldsymbol{q} \le \boldsymbol{q}_{high} \end{array}}$$



Expectedly, discrepancies in defining energy cost and constraints (e.g. capability limits)

Why Learning? A "Lift-up" in Simulation Space

Simpler Abstract Space

ML to supply "compressed" anatomical details

Faster to simulate &

Easier to solve control

Detailed Anatomical Space

Intuition: why simple sim can be as accurate as detailed sim?

Simpler Abstract Space

If final output is still skeletal motion

Anatomical space is redundant

- 90 leg muscles -> 10 DoFs
- Many bones -> a few DoFs at shoulder

Detailed Anatomical Space

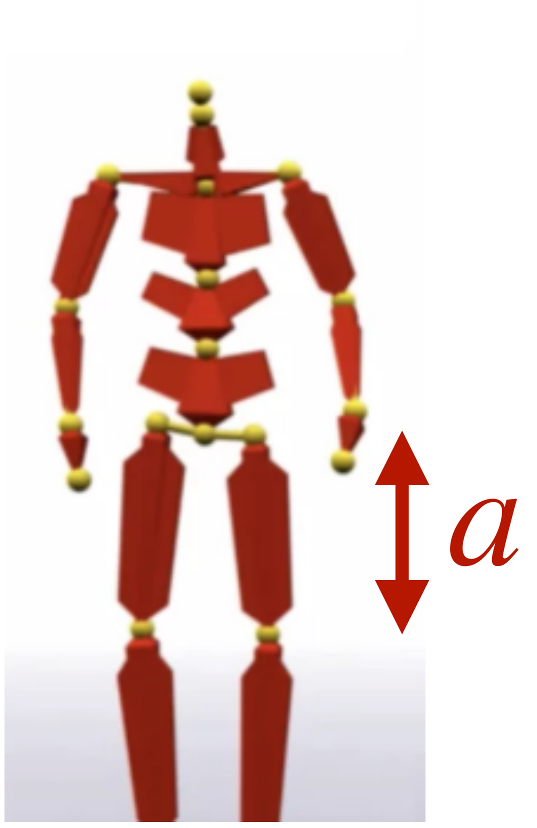
“State-dependency” to Bridge Simulation Spaces

Simpler Abstract Space

Learning “state-dependent” functions

Detailed Anatomical Space

Learned RoM, Torque limit, Metabolic energy Functions



$\boldsymbol{a}$: Muscle Activations

$$\min_{\boldsymbol{a}} \quad \sum a^2 + c_{task}(q)$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$f_{muscle-dynamics}(\boldsymbol{a}, \boldsymbol{l}, \dot{\boldsymbol{l}})$$

$$\boldsymbol{0} \leq \boldsymbol{a} \leq \boldsymbol{1}$$

$$f_{bone-ligaments}(\boldsymbol{q}, \boldsymbol{l})$$

$\boldsymbol{\tau}$: Joint Torques

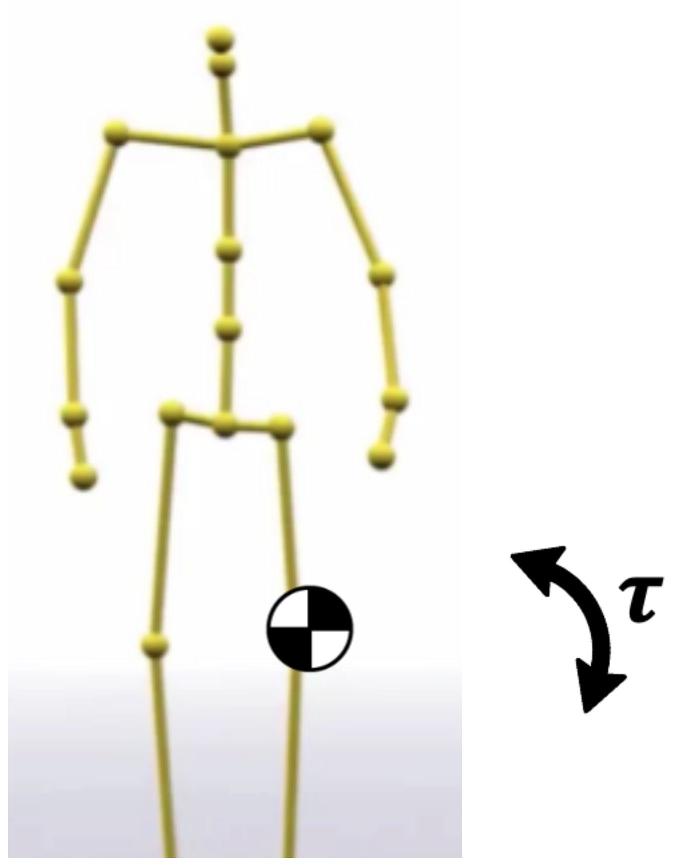
$$\min_{\boldsymbol{\tau}} \quad \sum \tau^2 + c_{task}(q)$$

subject to

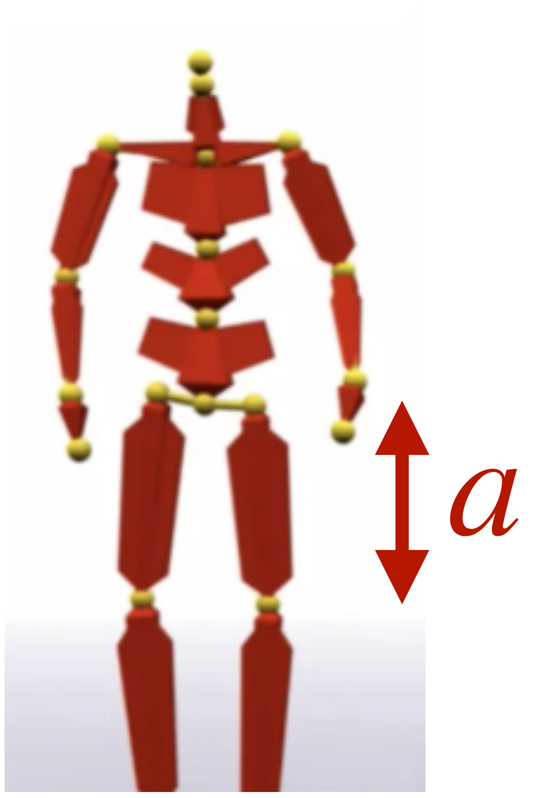
$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boldsymbol{\tau_{low}} \leq \boldsymbol{\tau} \leq \boldsymbol{\tau_{high}}$$

$$\boldsymbol{q_{low}} \leq \boldsymbol{q} \leq \boldsymbol{q_{high}}$$



Learned RoM, Torque limit, Metabolic energy Functions



$\boldsymbol{a}$: Muscle Activations

$$\min_{\boldsymbol{a}} \quad \sum a^2 + c_{task}(q)$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$f_{muscle-dynamics}(\boldsymbol{a}, \boldsymbol{l}, \dot{\boldsymbol{l}})$$

$$\boldsymbol{0} \leq \boldsymbol{a} \leq \boldsymbol{1}$$

$$f_{bone-ligaments}(\boldsymbol{q}, \boldsymbol{l})$$

$\boldsymbol{\tau}$: Joint Torques

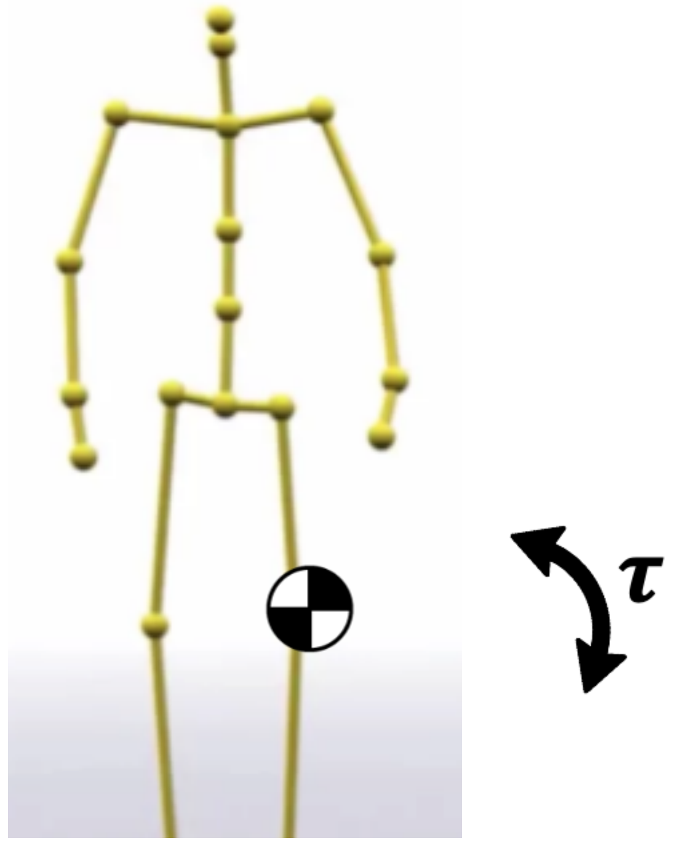
$$\min_{\boldsymbol{\tau}} \quad \sum \tau^2 + c_{task}(q)$$

subject to

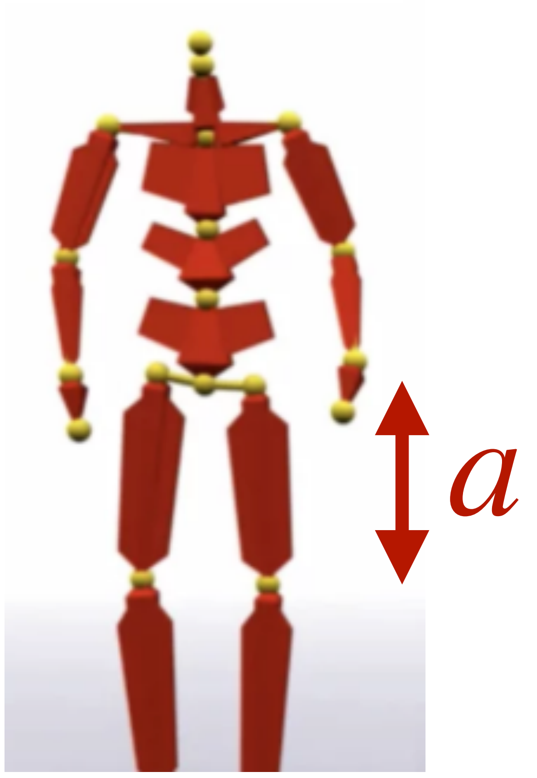
$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boldsymbol{\tau_{low}} \leq \boldsymbol{\tau} \leq \boldsymbol{\tau_{high}}$$

$$\mathrm{L}(\mathbf{q}) > \mathbf{0}$$



Learned RoM, Torque limit, Metabolic energy Functions



$\boldsymbol{a}$: Muscle Activations

$$\min_{\boldsymbol{a}} \quad \sum a^2 + c_{task}(q)$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$f_{muscle-dynamics}(\boldsymbol{a}, \boldsymbol{l}, \dot{\boldsymbol{l}})$$

$$\boldsymbol{0} \leq \boldsymbol{a} \leq \boldsymbol{1}$$

$$f_{bone-ligaments}(\boldsymbol{q}, \boldsymbol{l})$$

$\boldsymbol{\tau}$: Joint Torques

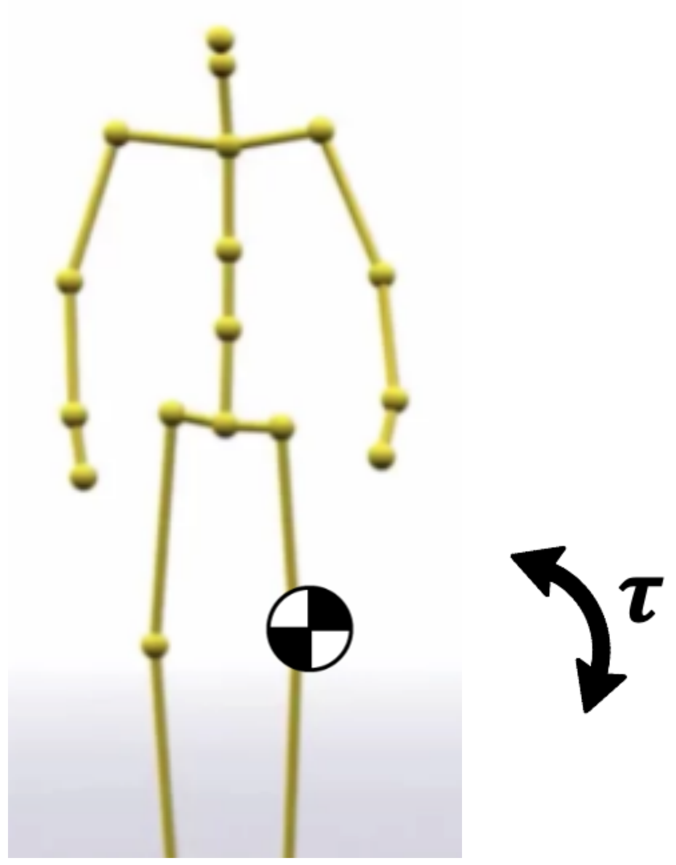
$$\min_{\boldsymbol{\tau}} E(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{\tau}) + c_{task}(q)$$

subject to

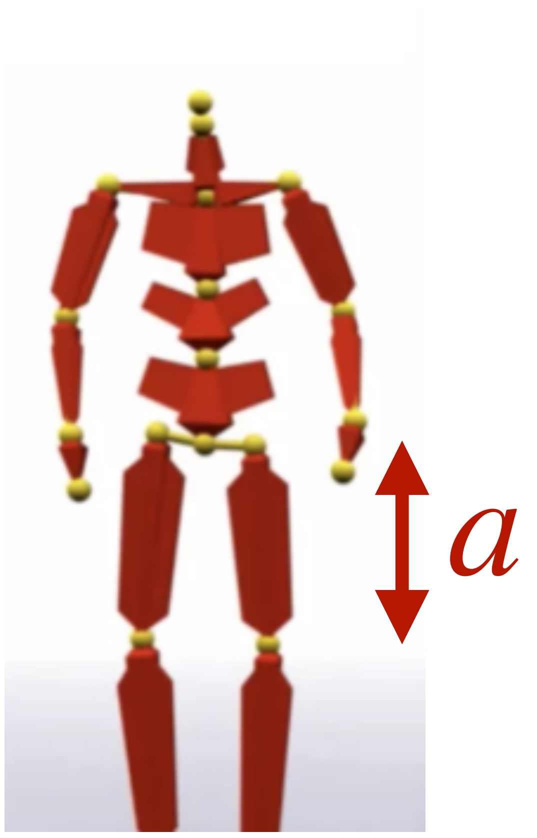
$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\boldsymbol{\tau}_{low} \leq \boldsymbol{\tau} \leq \boldsymbol{\tau}_{high}$$

$$\mathrm{L}(\mathbf{q}) > \mathbf{0}$$



Learned RoM, Torque limit, Metabolic energy Functions



$\boldsymbol{a}$: Muscle Activations

$$\min_{\boldsymbol{a}} \quad \sum a^2 + c_{task}(q)$$

subject to

$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$f_{muscle-dynamics}(\boldsymbol{a}, \boldsymbol{l}, \dot{\boldsymbol{l}})$$

$$\boldsymbol{0} \leq \boldsymbol{a} \leq \boldsymbol{1}$$

$$f_{bone-ligaments}(\boldsymbol{q}, \boldsymbol{l})$$

$\boldsymbol{\tau}$: Joint Torques

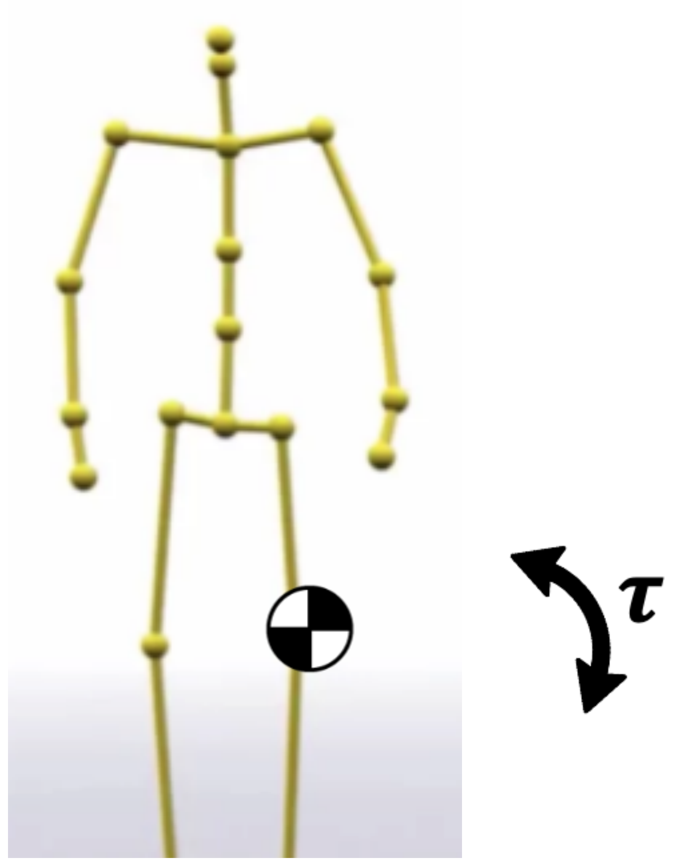
$$\min_{\boldsymbol{\tau}} E(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{\tau}) + c_{task}(q)$$

subject to

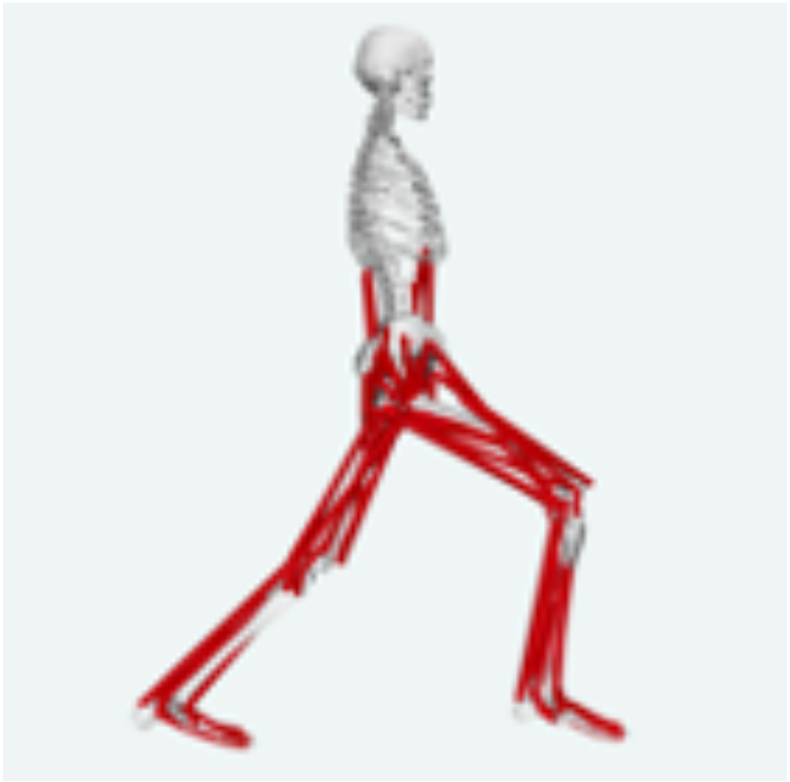
$$\ddot{\boldsymbol{q}} = f_{skel-dynamics}(\boldsymbol{q}, \dot{\boldsymbol{q}})$$

$$\mathrm{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{\tau}) \leq 0$$

$$\mathrm{L}(\mathrm{q}) > \boldsymbol{0}$$



Learned RoM, Torque limit, Metabolic energy Functions



Learn from detailed muscle simulator

[Acktar, Black CVPR'15]

Learn from real data

Learned RoM, Torque limit, Metabolic energy Functions

Learn from detailed muscle simulator

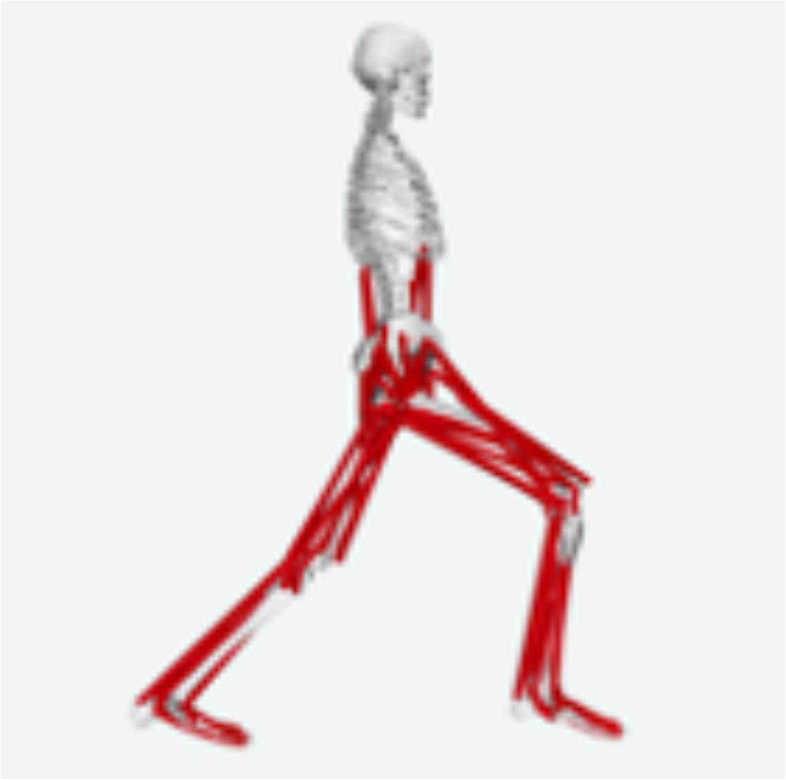
$$L(\mathbf{q}) > 0$$

[Acktar, Black CVPR'15]

Learn from real data

Learned RoM, Torque limit, Metabolic energy Functions

$$E(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{\tau})$$



Learn from detailed muscle simulator

$$\mathrm{L}(\mathbf{q}) > 0$$

[Acktar, Black CVPR'15]

Learn from real data

Learned RoM, Torque limit, Metabolic energy Functions

$$E(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{\tau})$$

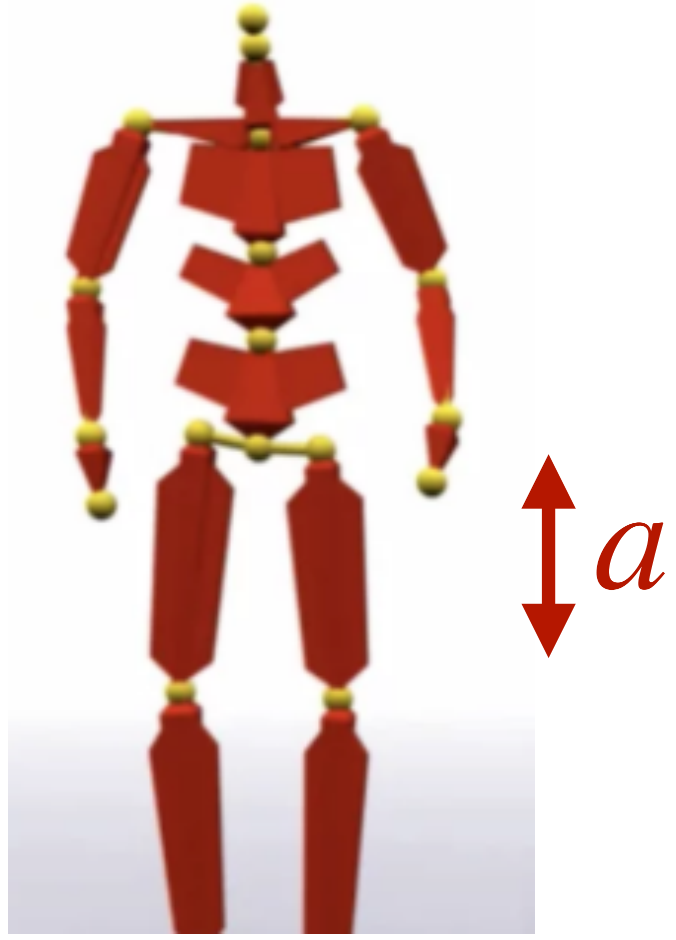
Learn from detailed muscle simulator

$$\mathbf{C}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{\tau}) \leq 0$$

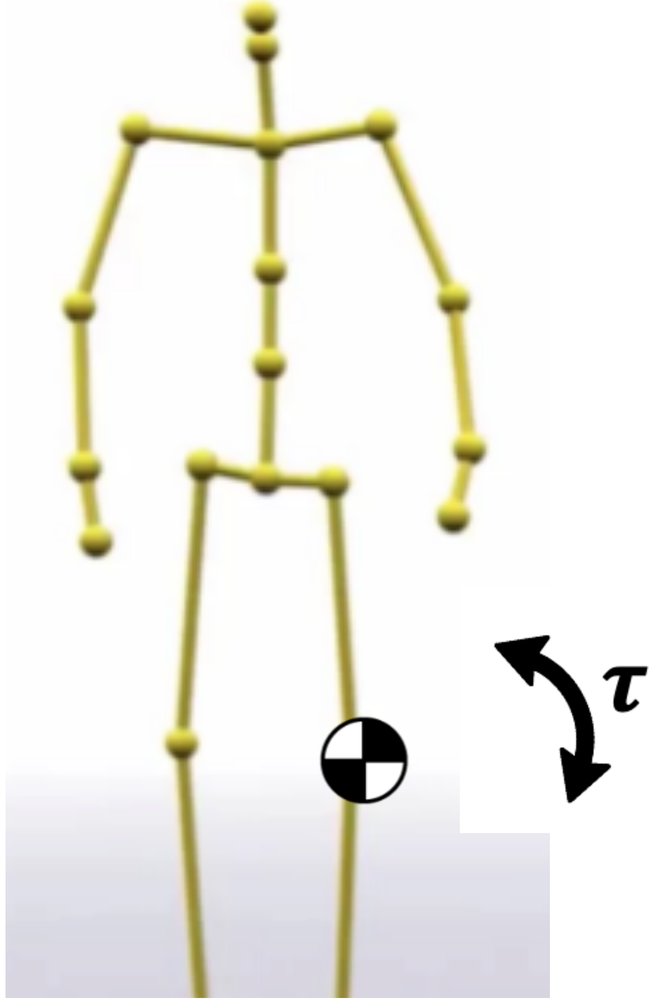
$$\mathrm{L}(\mathbf{q}) > \mathbf{0}$$

[Acktar, Black CVPR'15]

Learn from real data



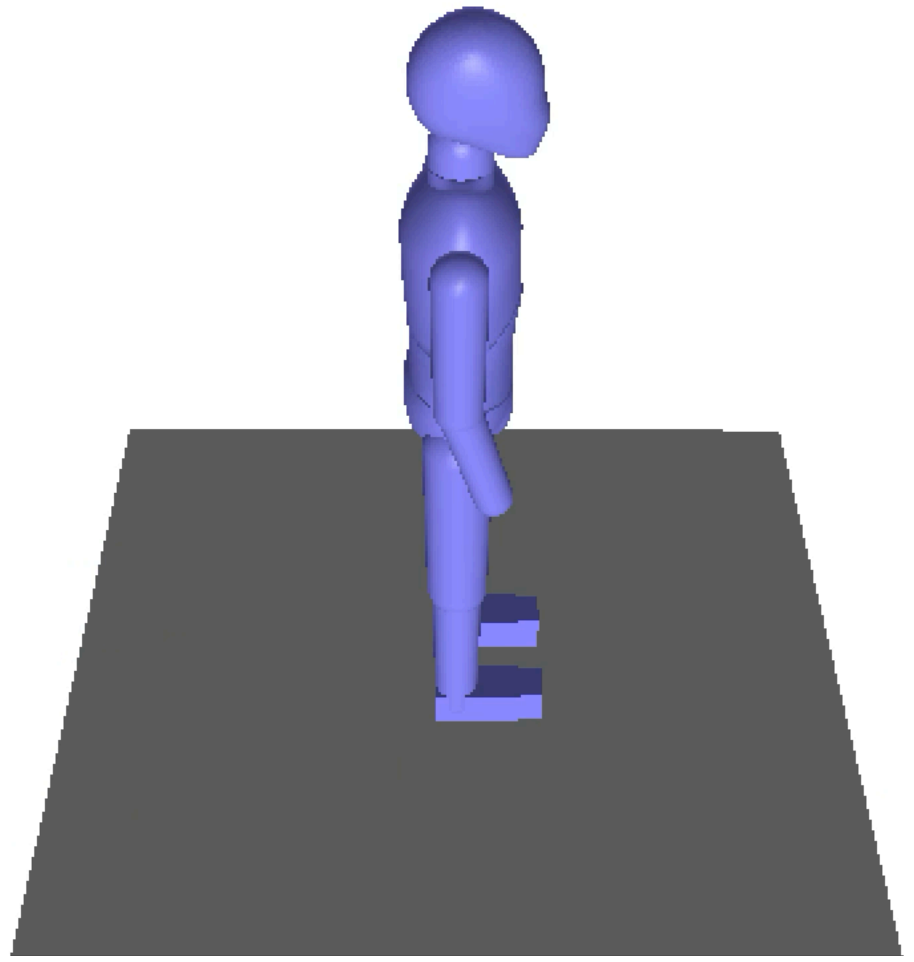
Augmented with learned state-dependent functions



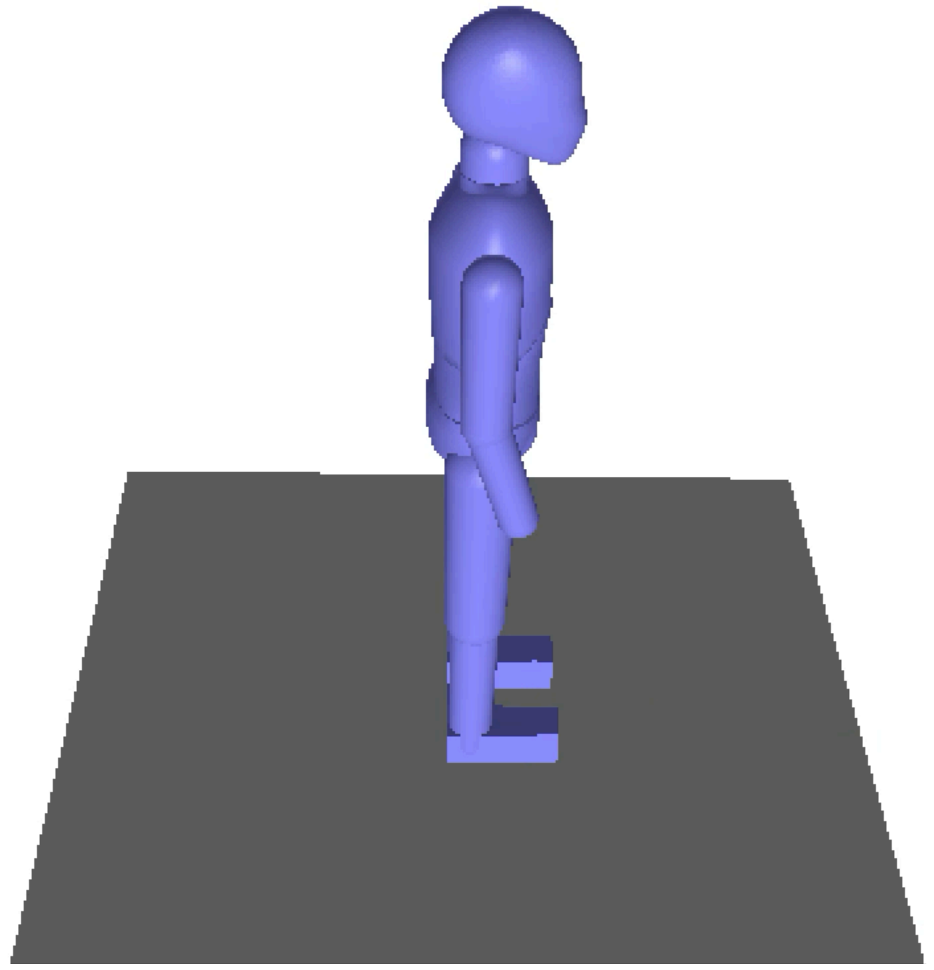
We can prove both control problems now have a same optimal value (equivalency)

Results

No Motion Control, Free-fall Simulation

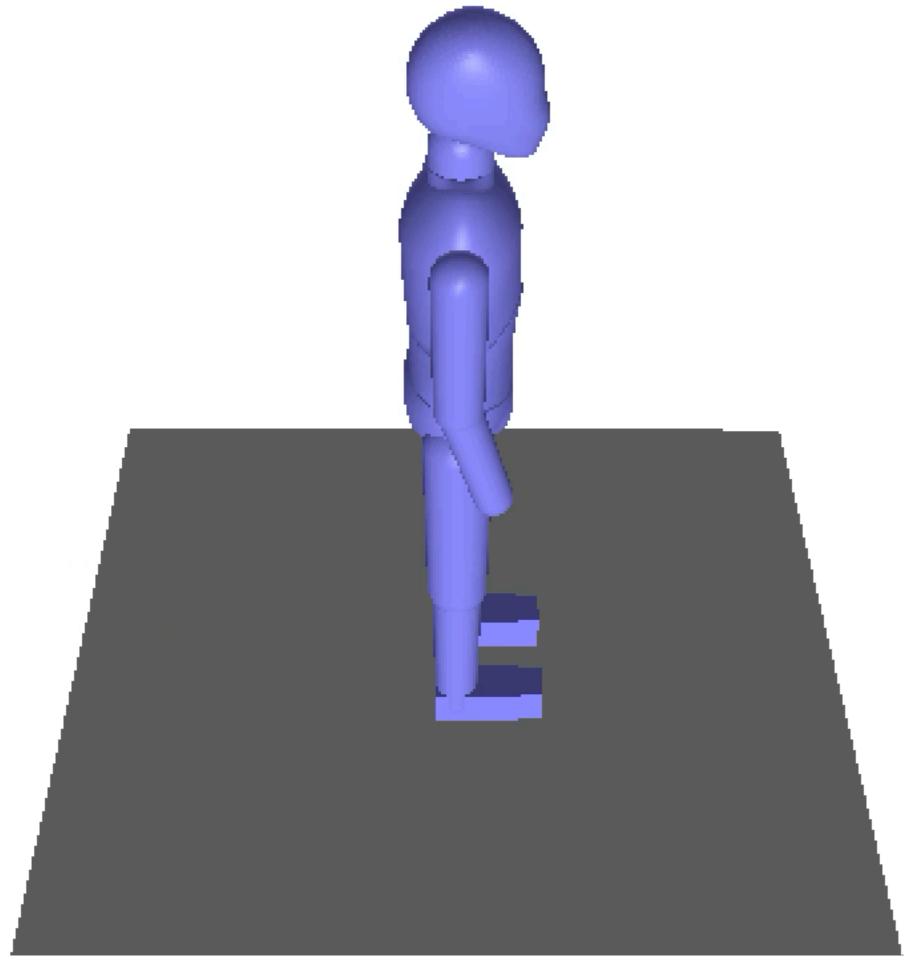


With learned $L(q) > 0$

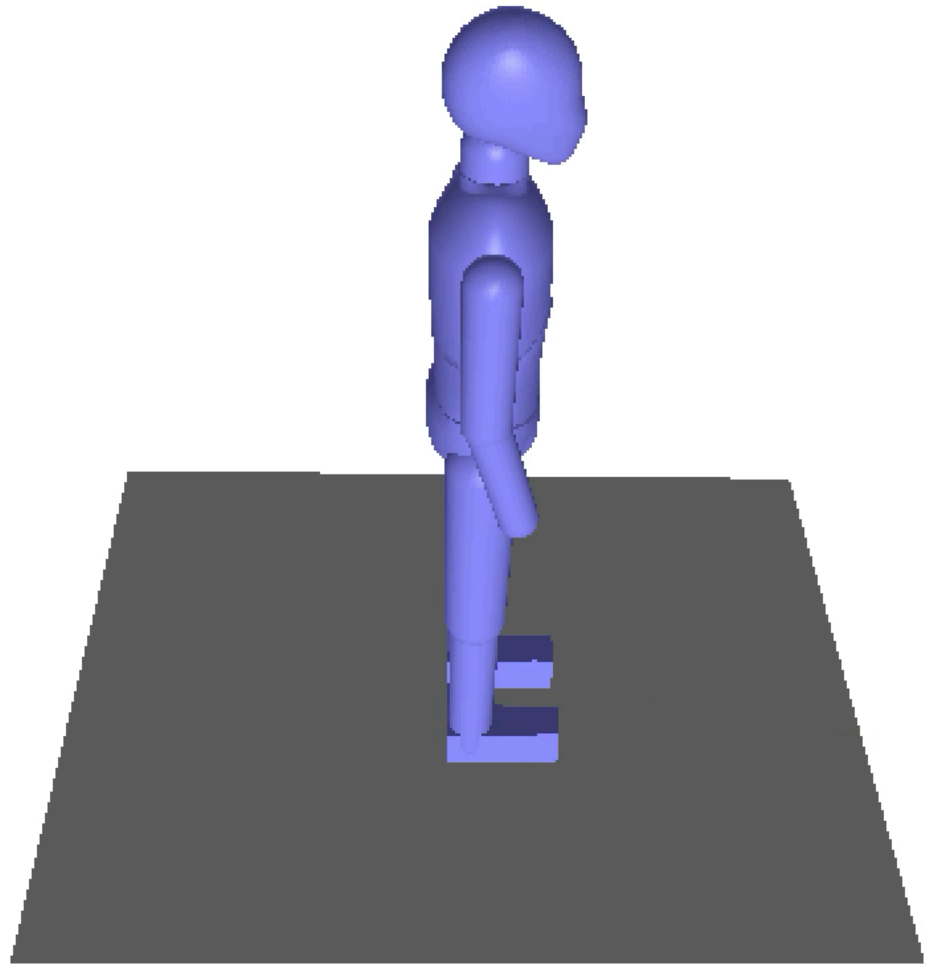


Without learned $L(q) > 0$

No Motion Control, Free-fall Simulation

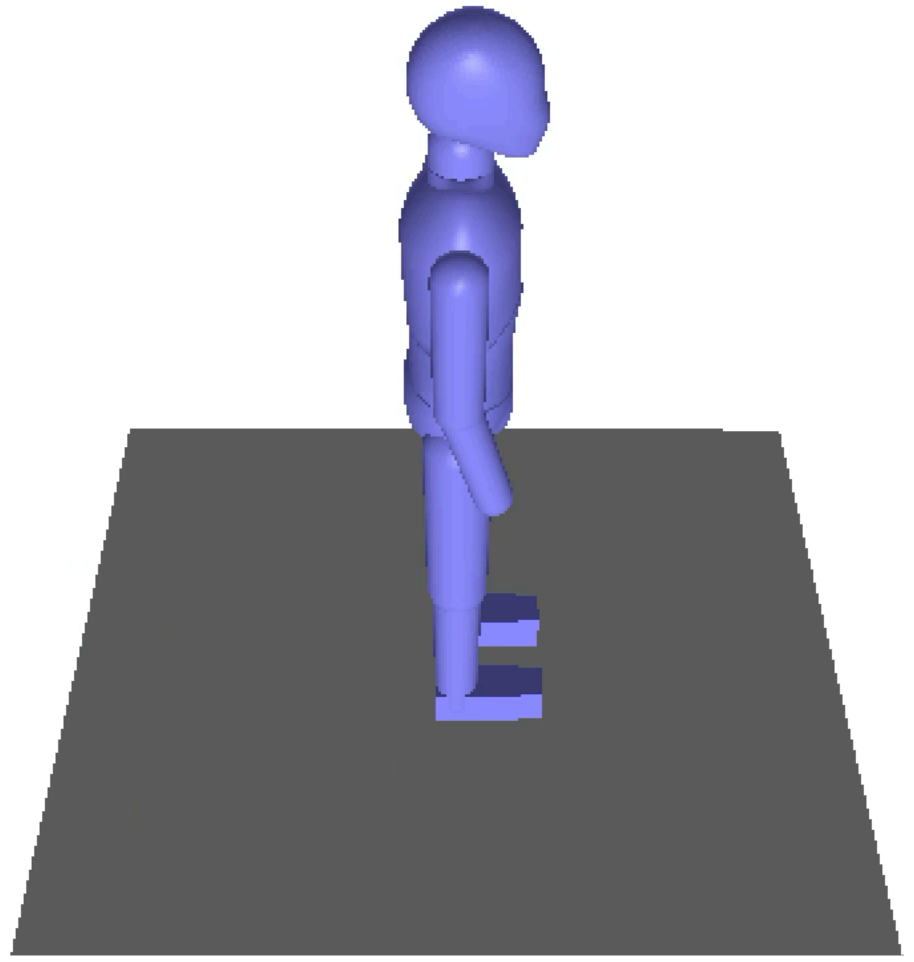


With learned $L(q) > 0$

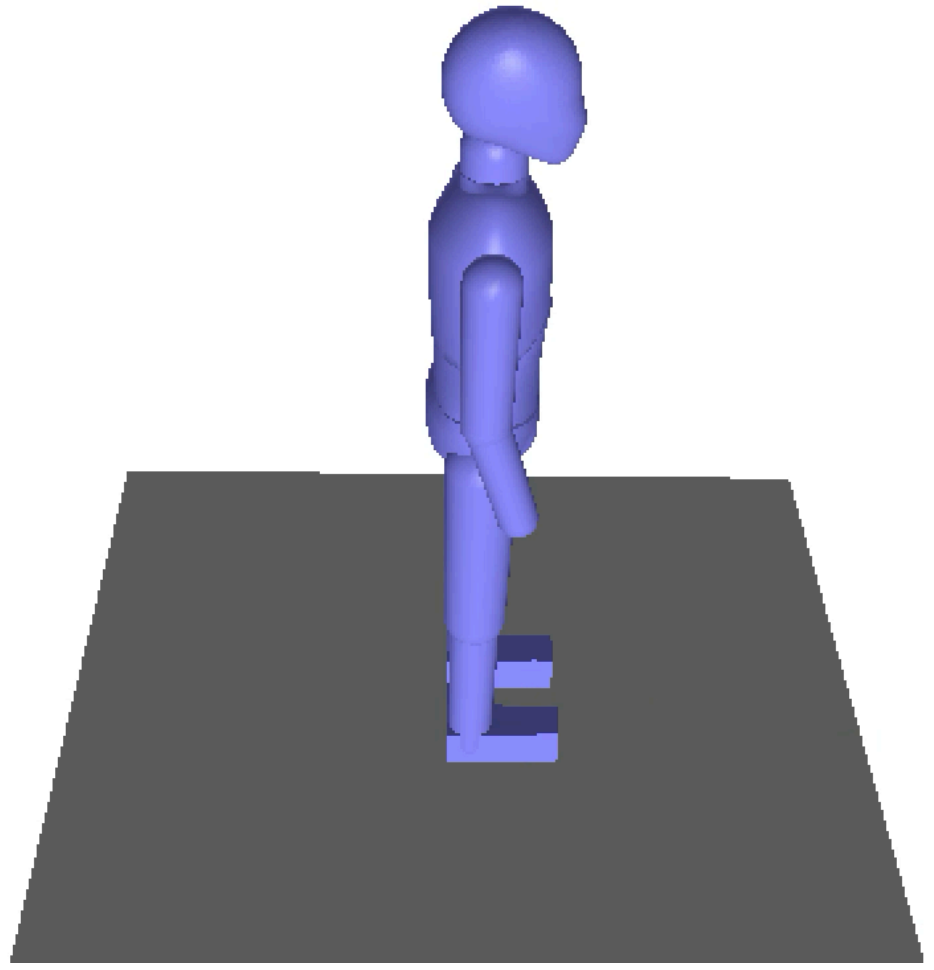


Without learned $L(q) > 0$

No Motion Control, Free-fall Simulation



With learned $L(q) > 0$



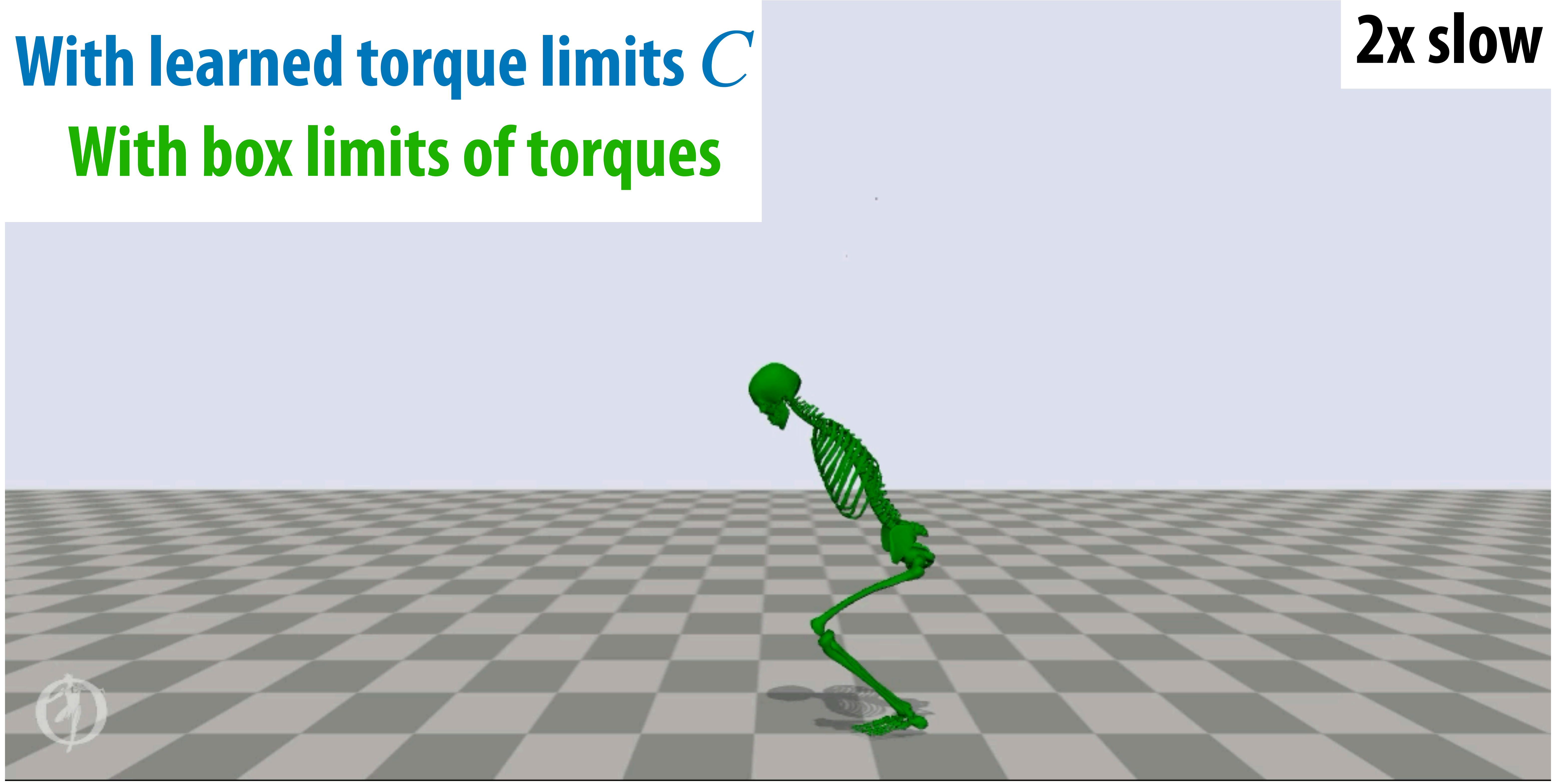
Without learned $L(q) > 0$

Motion Control: Jump as High as You Can

With learned torque limits C

With box limits of torques

2x slow

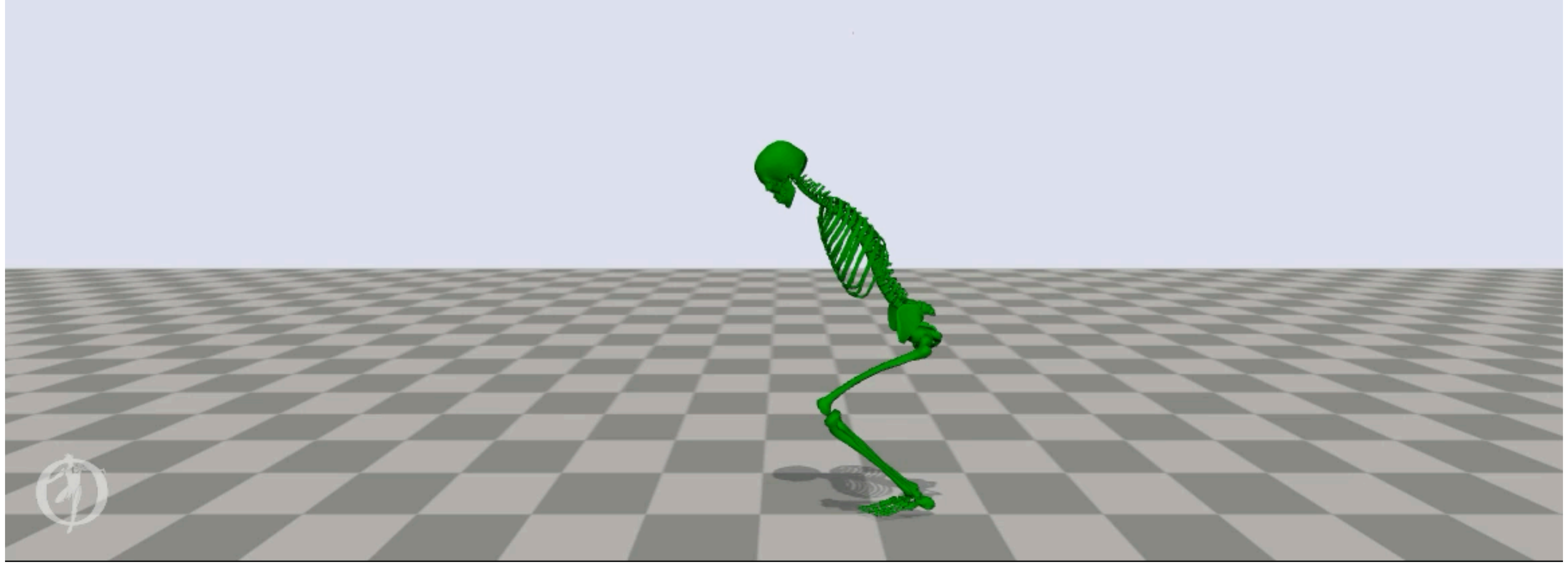


Motion Control: Jump as High as You Can

With learned torque limits C

With box limits of torques

2x slow

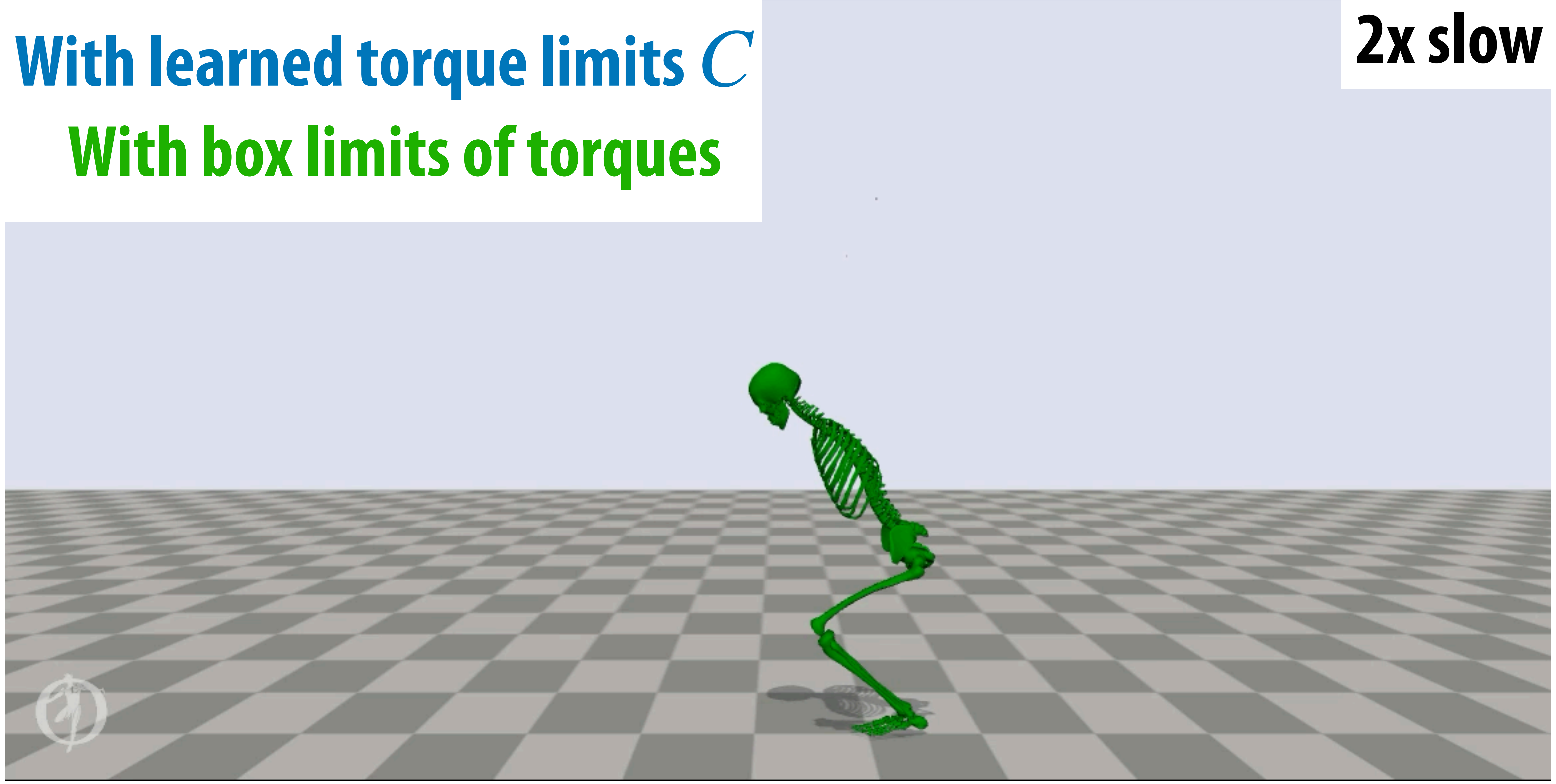


Motion Control: Jump as High as You Can

With learned torque limits C

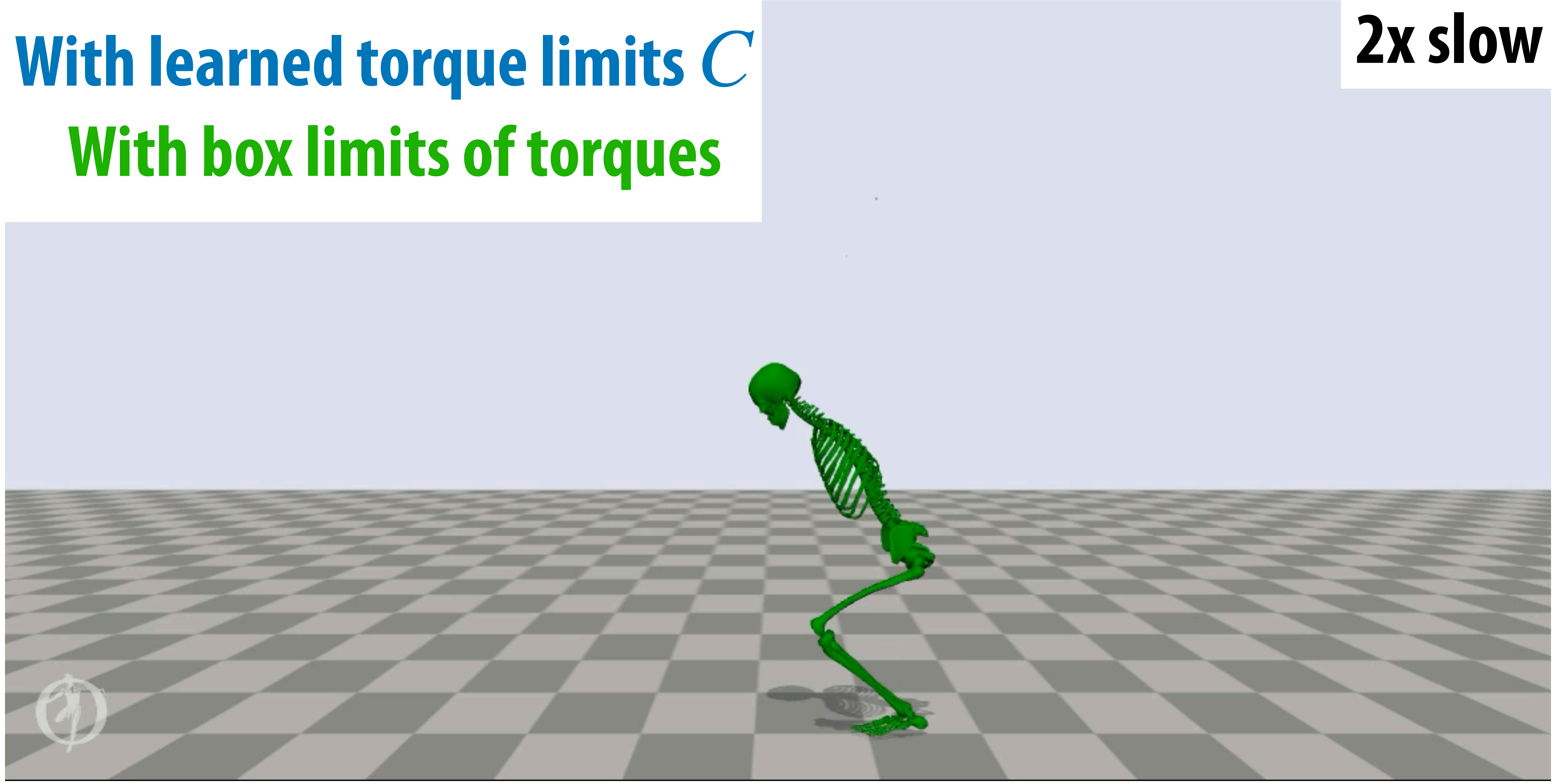
With box limits of torques

2x slow



Can jump higher if bends down more

Motion Control: Jump as High as You Can

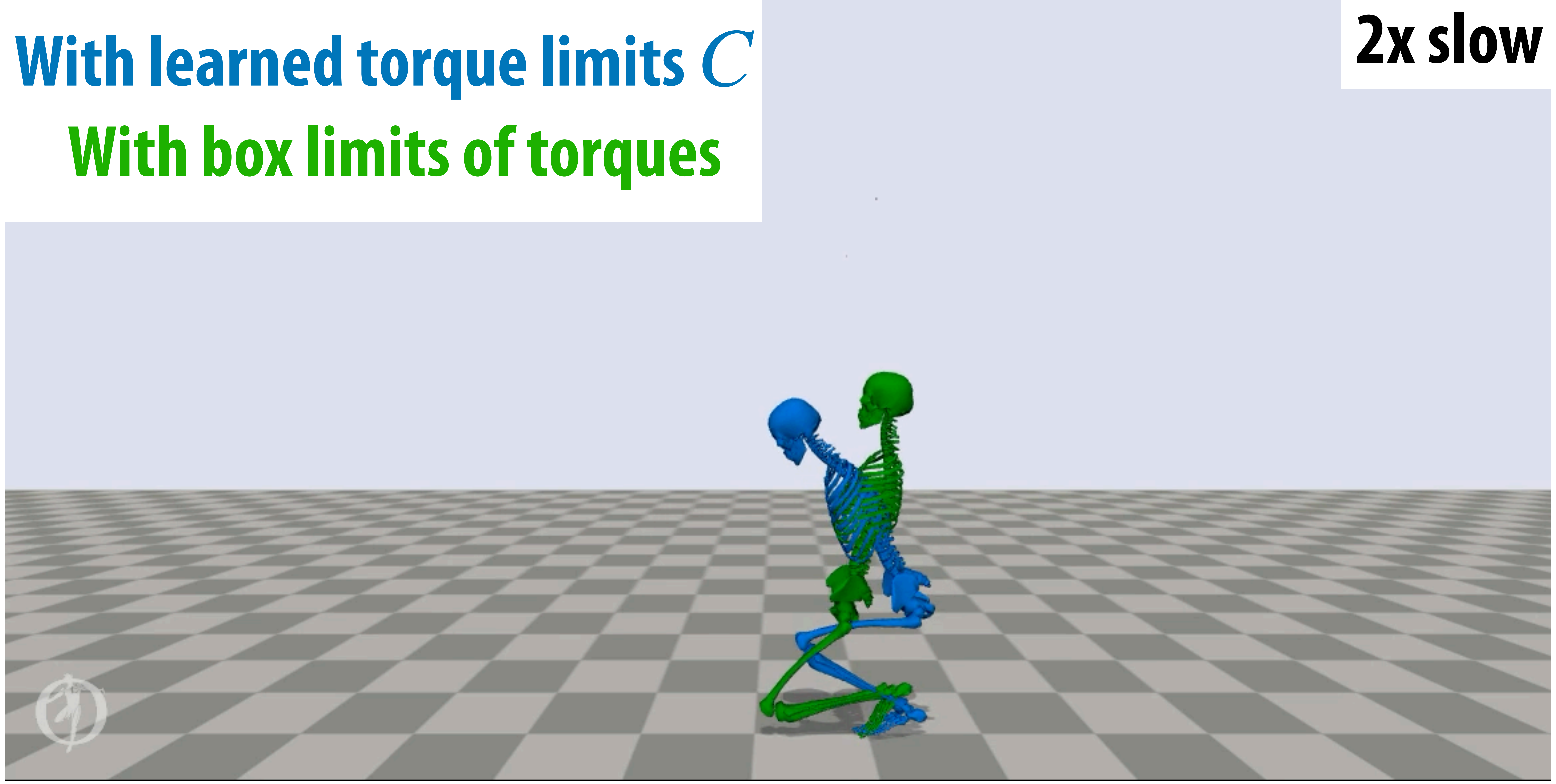


Can jump higher if bends down more

Motion Control: Jump as High as You Can

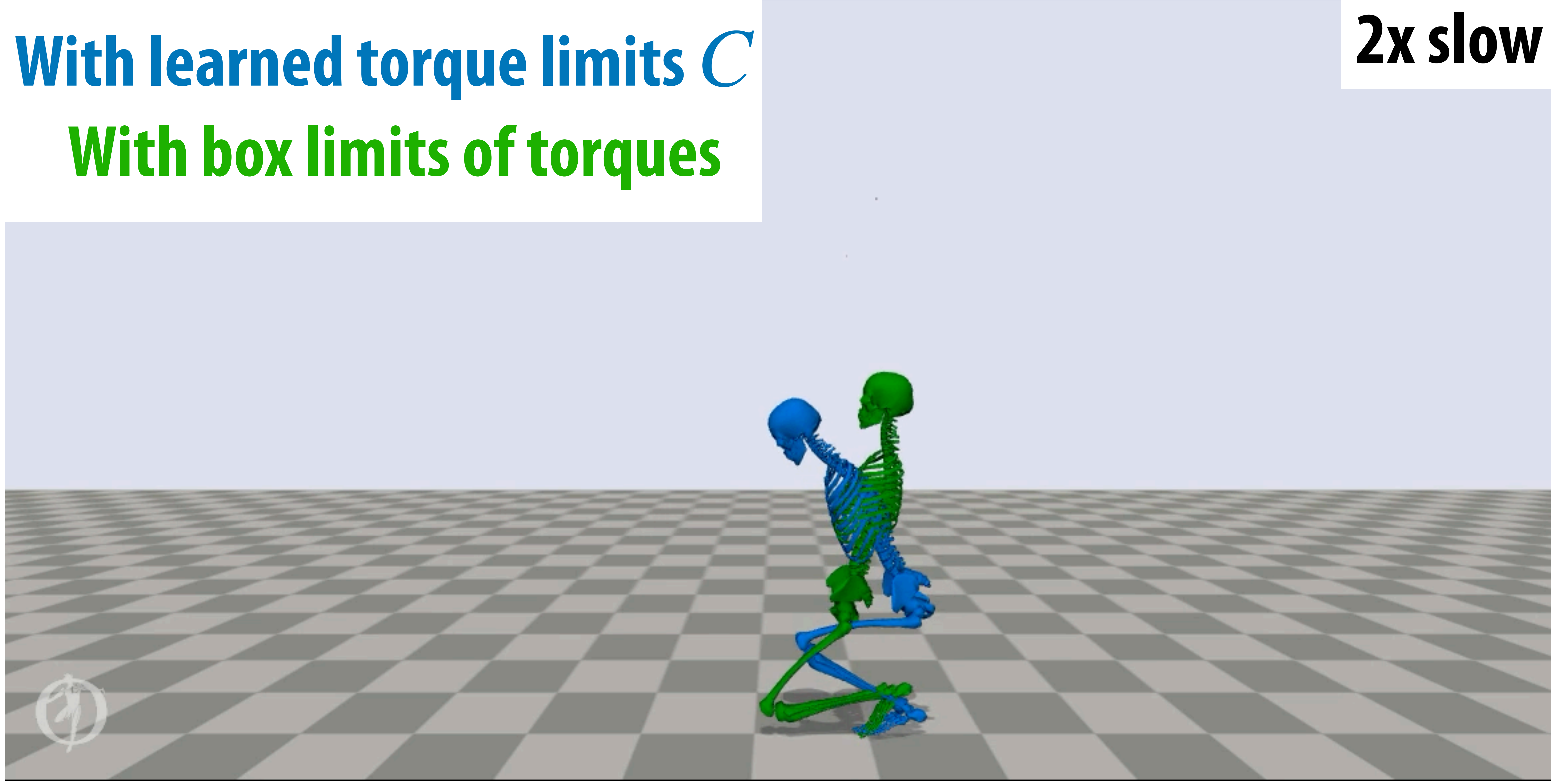
With learned torque limits C

With box limits of torques



Humans don't do that because small torque limit during hyper-flexion

Motion Control: Jump as High as You Can



Humans don't do that because small torque limit during hyper-flexion

Motion Control 2: Swing as Far as You Can

With learned torque limits C

With box limits of torques

Similarly, ours don't hyper-flex

Motion Control 2: Swing as Far as You Can

With learned torque limits C

With box limits of torques

Similarly, ours don't hyper-flex

Motion Control 2: Swing as Far as You Can

Ours
Detailed muscle models

Almost identical solution compared with detailed muscle simulation
Ours use 70% less computation & fewer iterations

Motion Control 2: Swing as Far as You Can

Ours

Detailed muscle models

Almost identical solution compared with detailed muscle simulation

Ours use 70% less computation & fewer iterations

Recap

Biomechanically accurate, fast, and easier for solving control

Facilitate large-scale simulations, for training / synthetic data generation

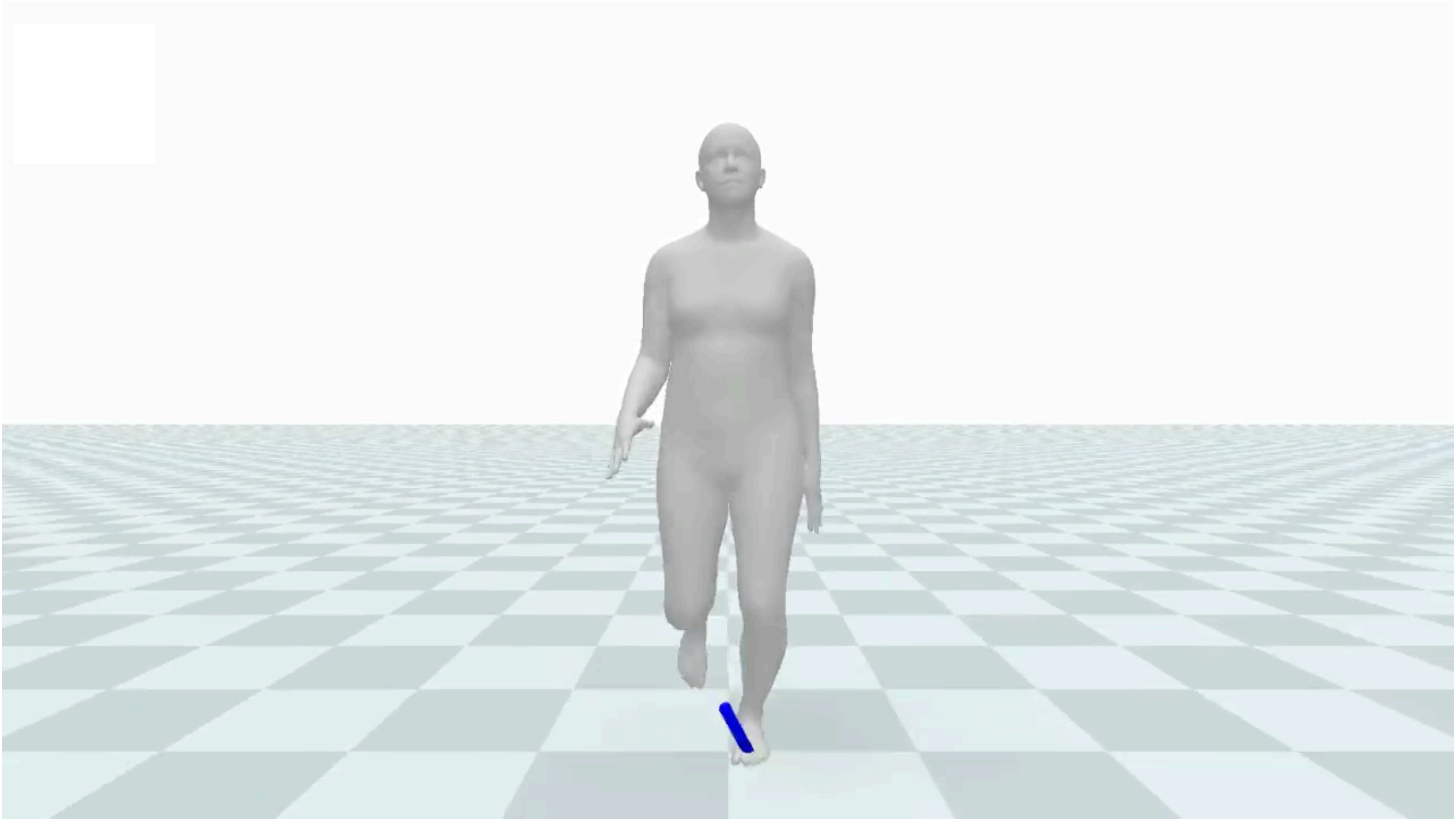
Learned anatomical functions to provably "compress" biomechanics knowledge

Simulation-augmented Generative Motion Model

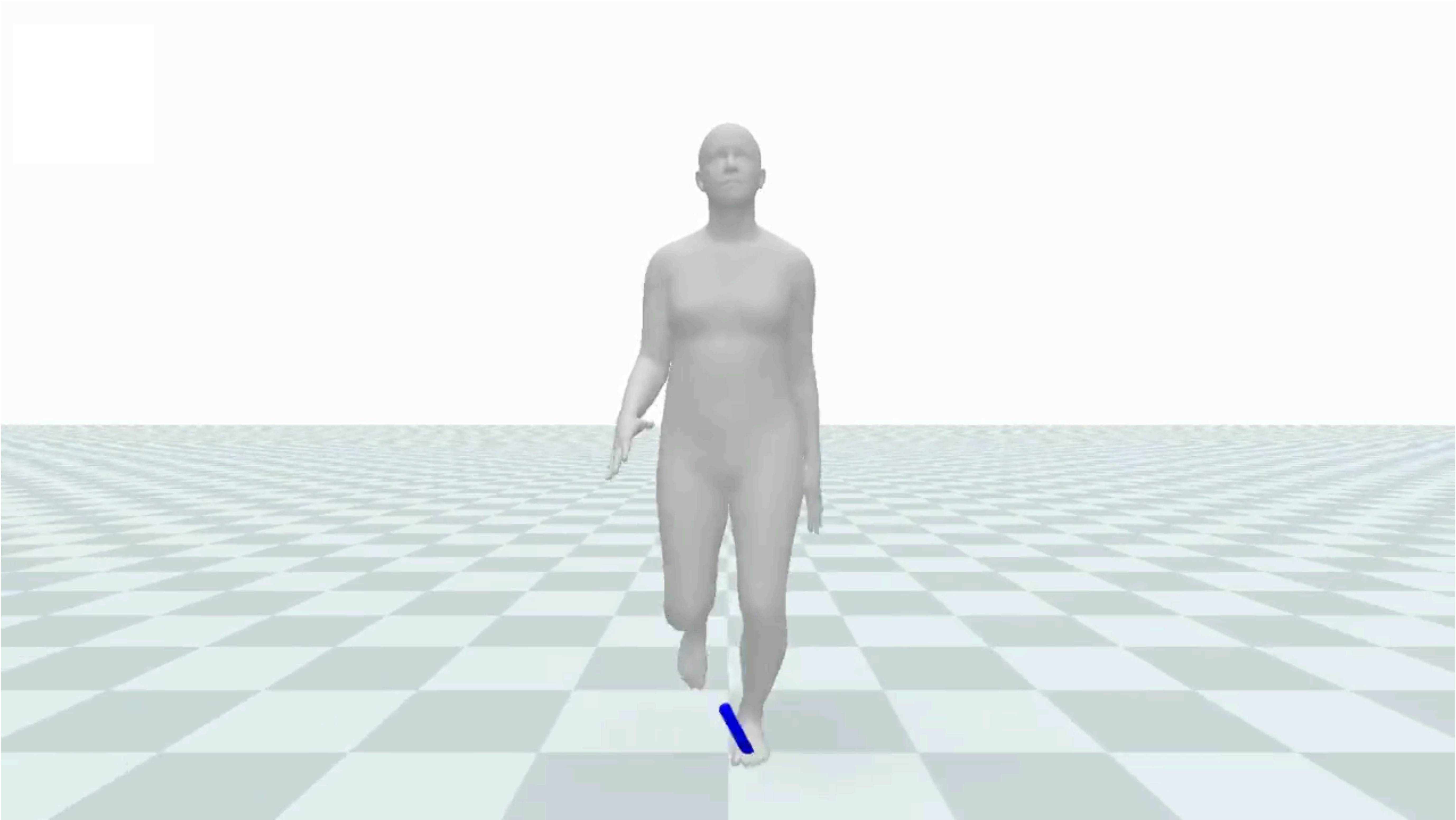
— How to build GenAI motion models that interactively reacts to physics

[Jiang et al] SIGGRAPH Asia '23

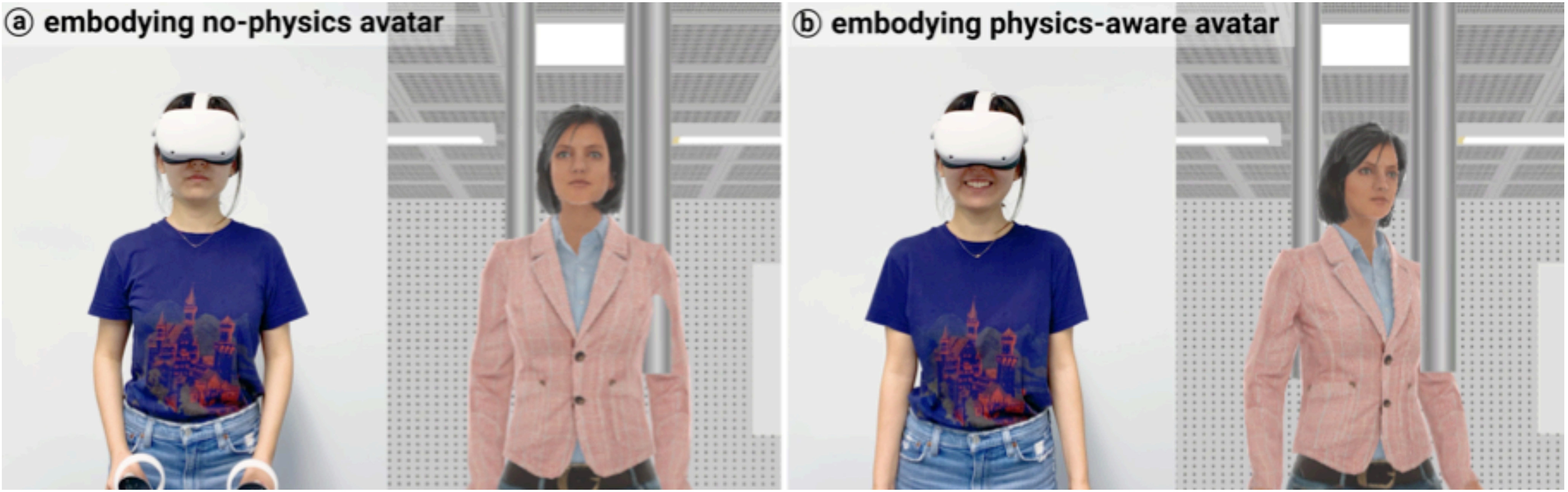
Digital Humans that Understands and Responds to Intuitive Physics



Digital Humans that Understands and Responds to Intuitive Physics



Physics-aware Digital Humans Can:



Improve immersion in AR/VR

Tao et al CHI'23

Physics-aware Digital Humans Can:

Help train robots / embodied AI agents in simulation

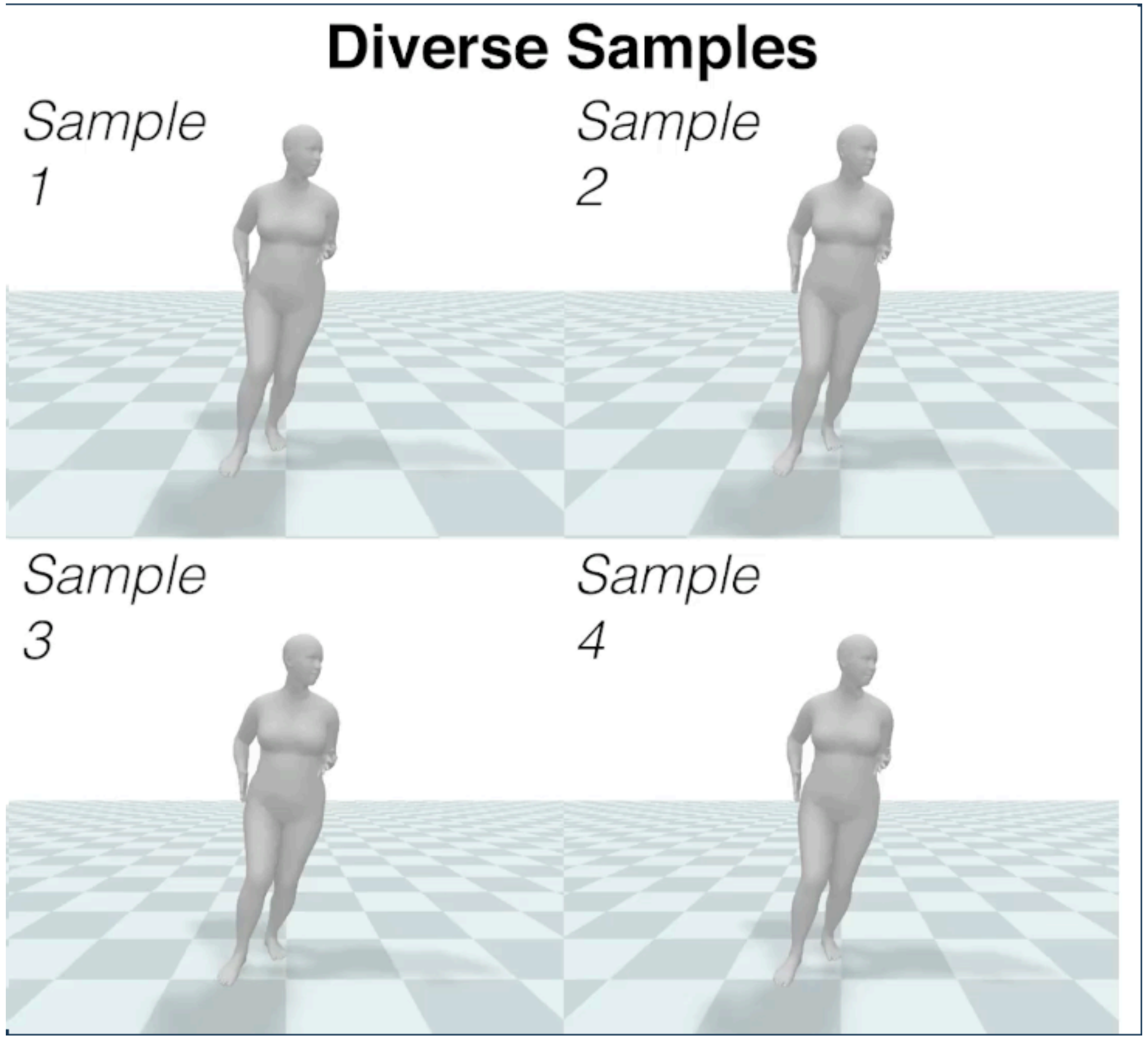
Habitat 3.0, 2023

Physics-aware Digital Humans Can:

Help train robots / embodied AI agents in simulation

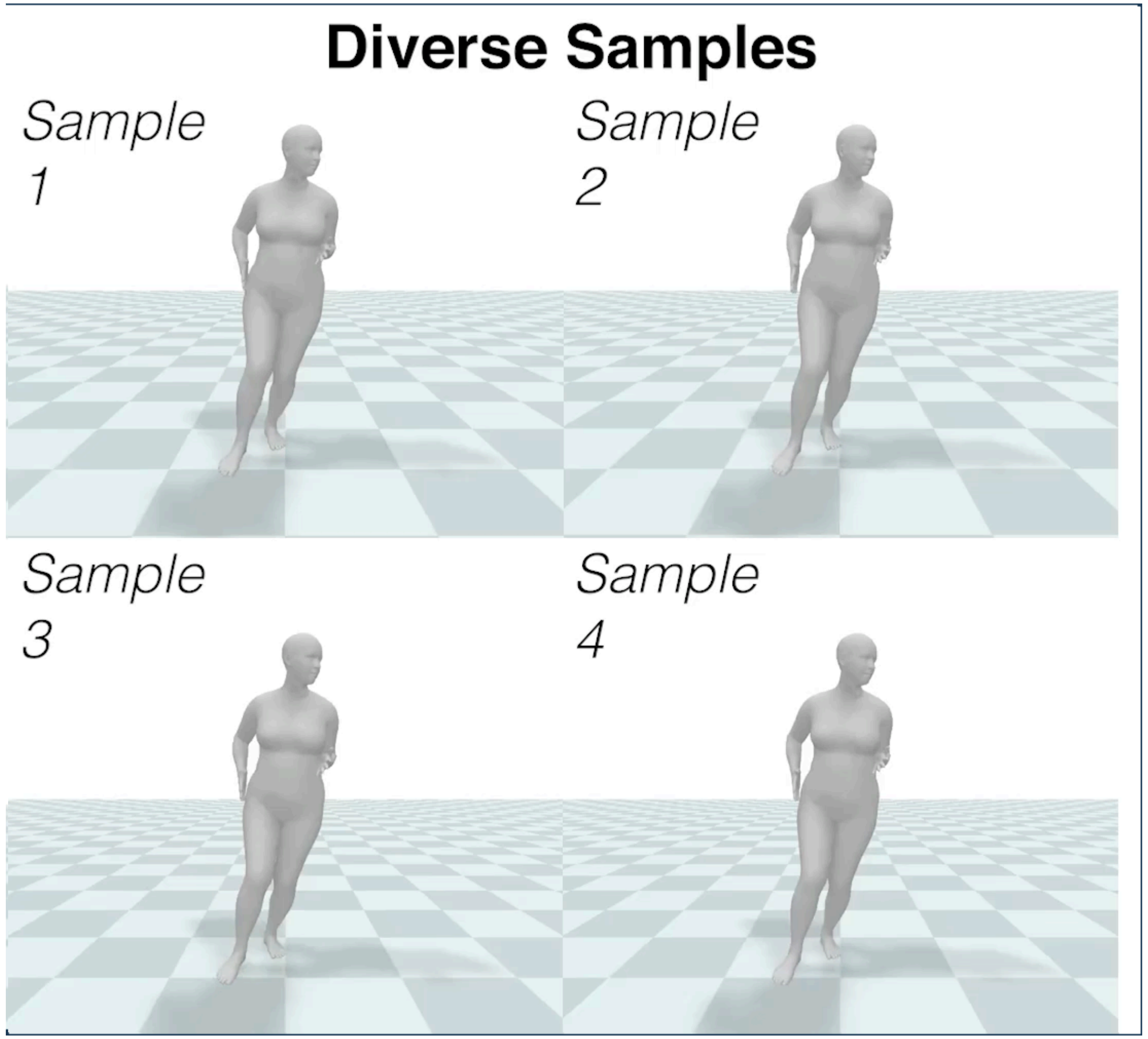
Habitat 3.0, 2023

Generative Models, for Motion



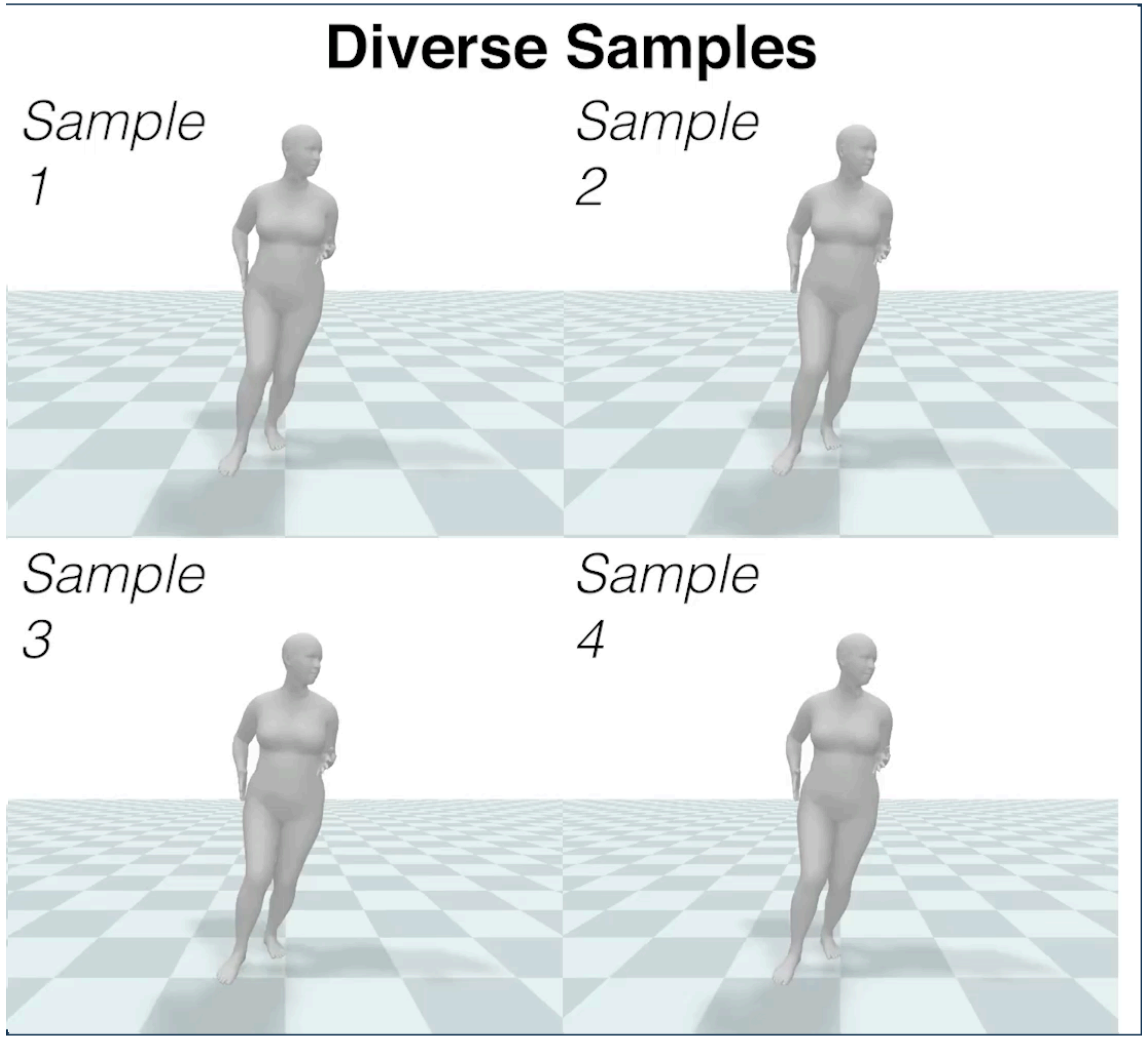
Rempe et al ICCV'21

Same Input, Diverse Output



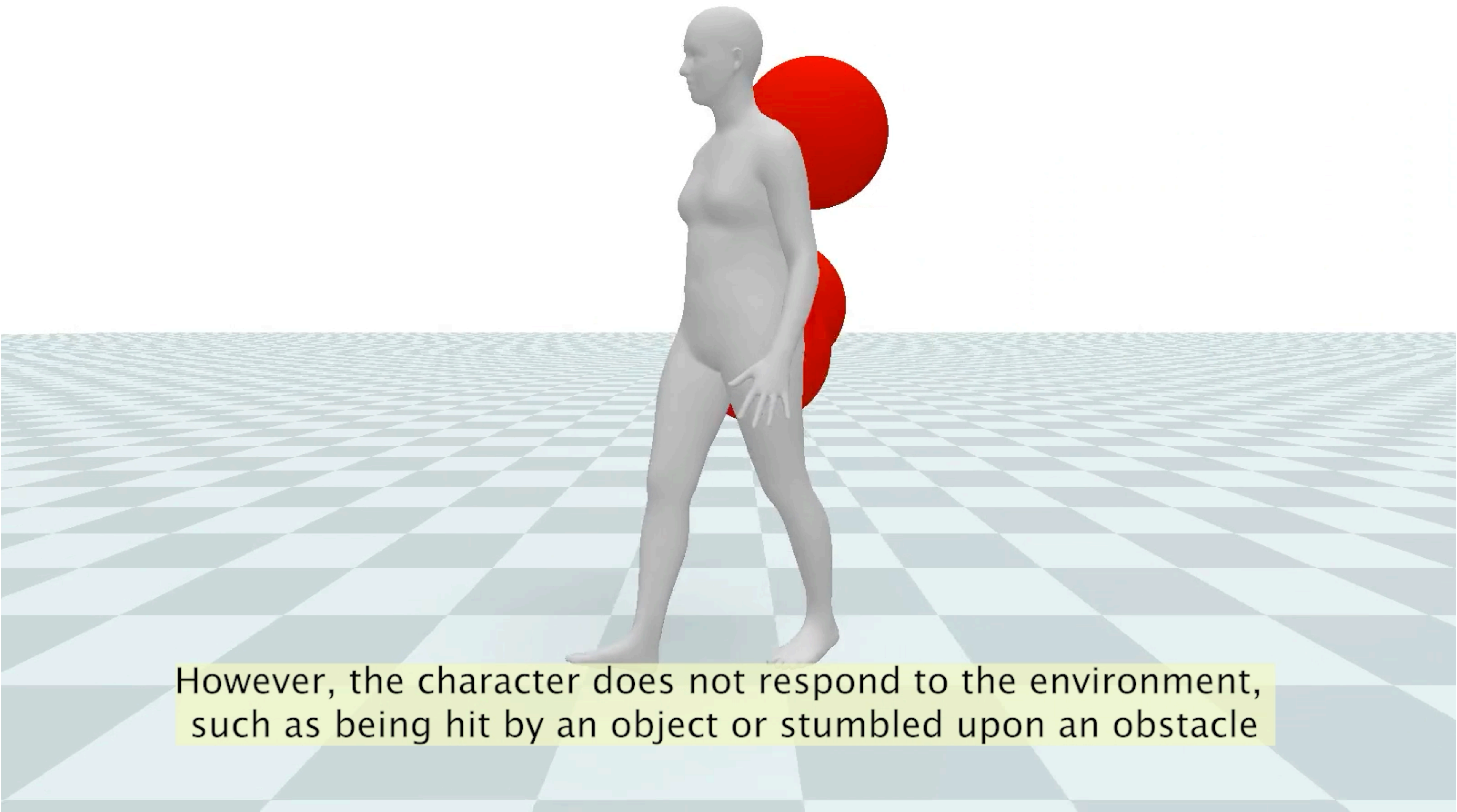
Rempe et al ICCV'21

Same Input, Diverse Output

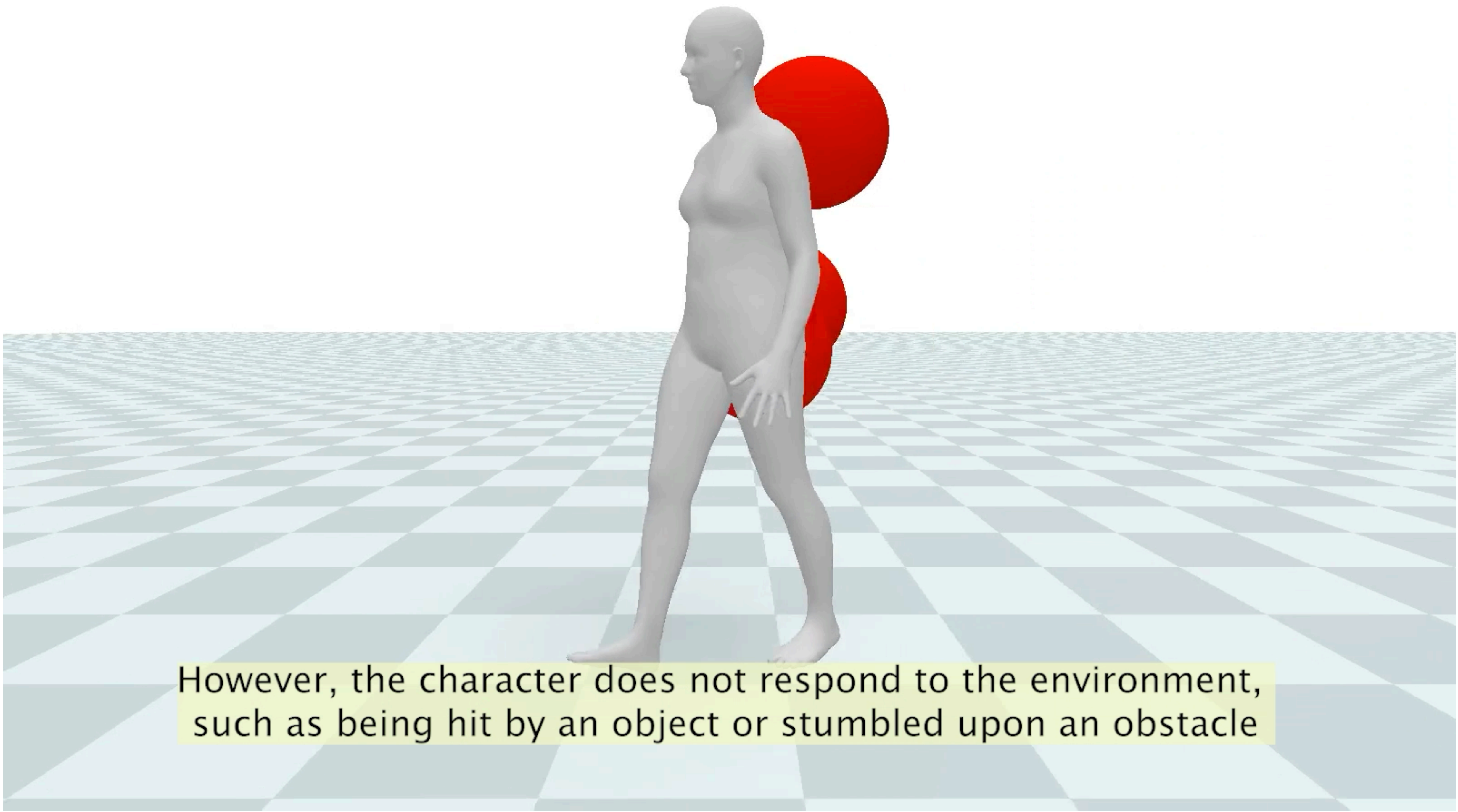


Rempe et al ICCV'21

Yes, but does not respond to physical events

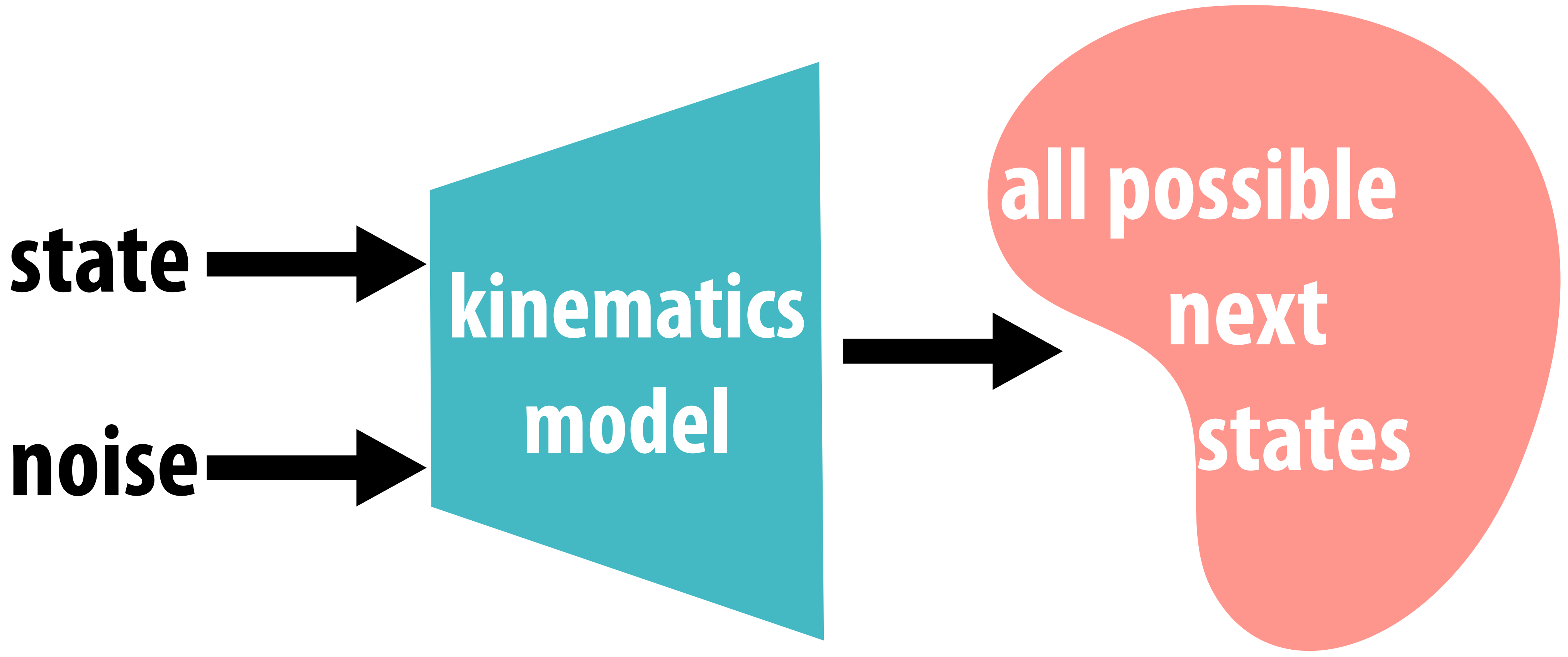


Yes, but does not respond to physical events



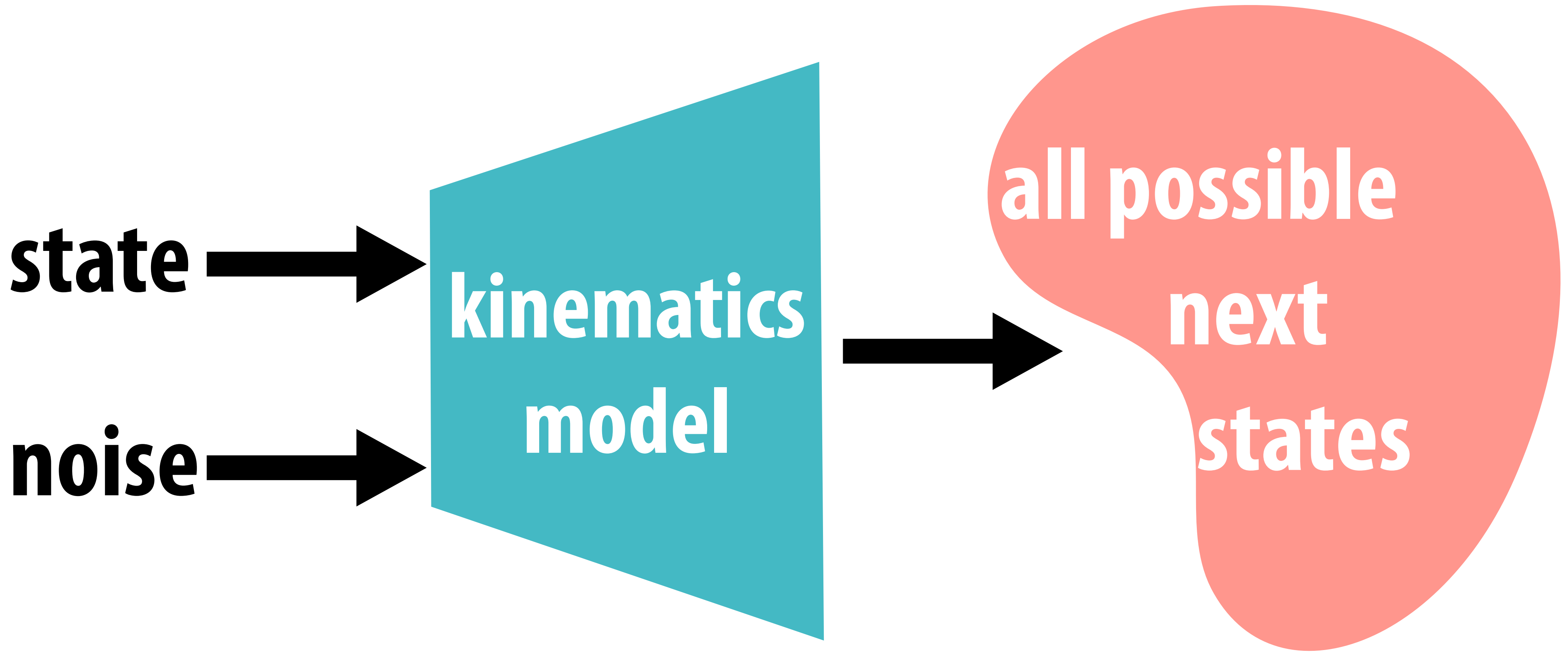
Challenges

1. Formulation does not consider physics



Challenges

1. **Formulation does not consider physics**



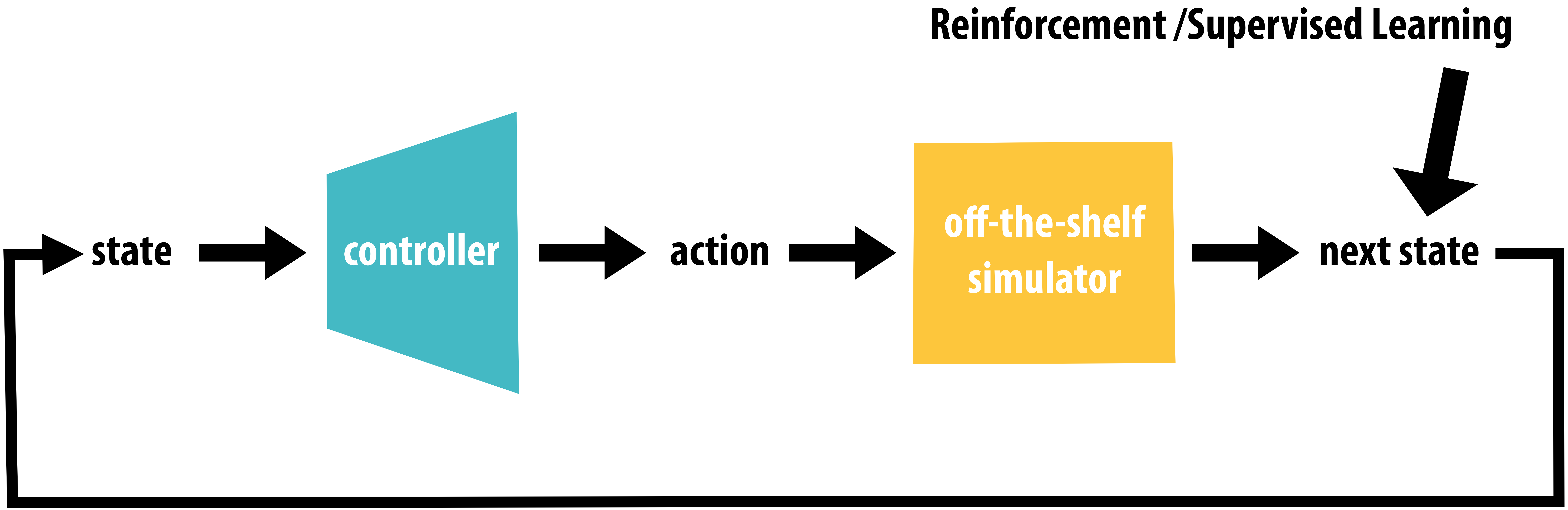
2. **Physical responses data unsafe to capture**

Challenges

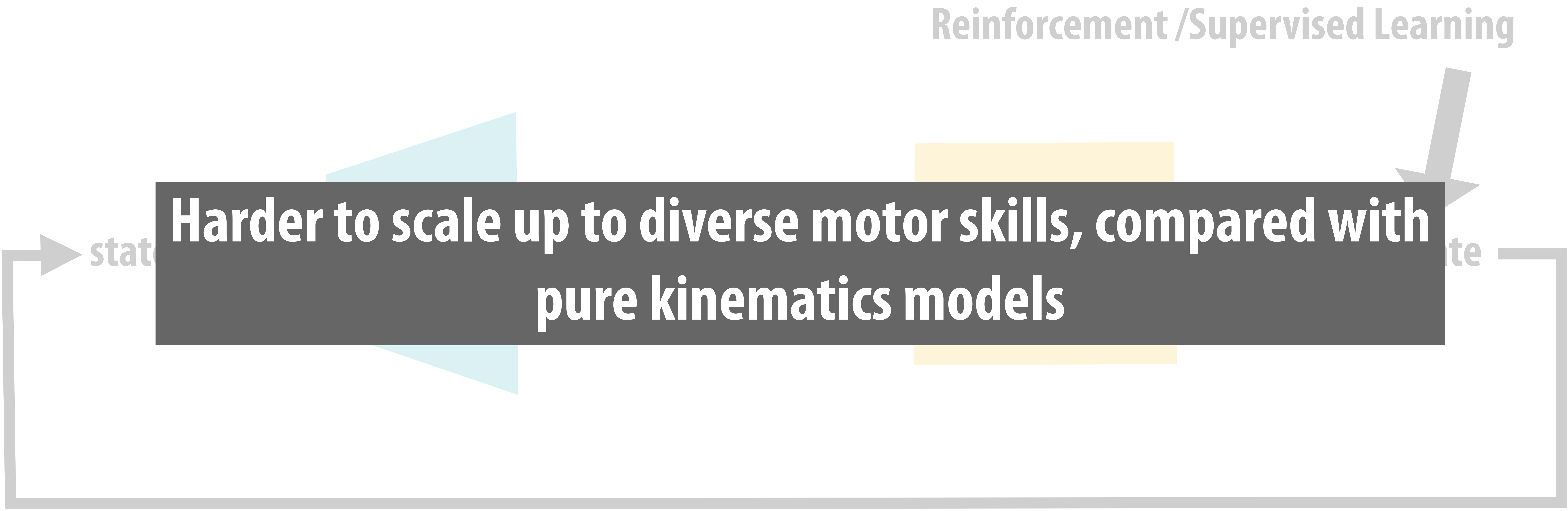
1. **Formulation does not consider physics**
2. **Physical responses data unsafe to capture**

Simulation could help with both!

Commonly, Off-the-shelf Simulation in Training Loop



Commonly, Off-the-shelf Simulation in Training Loop



Harder to scale up to diverse motor skills, compared with pure kinematics models

Physics plugin so that no further training is needed?

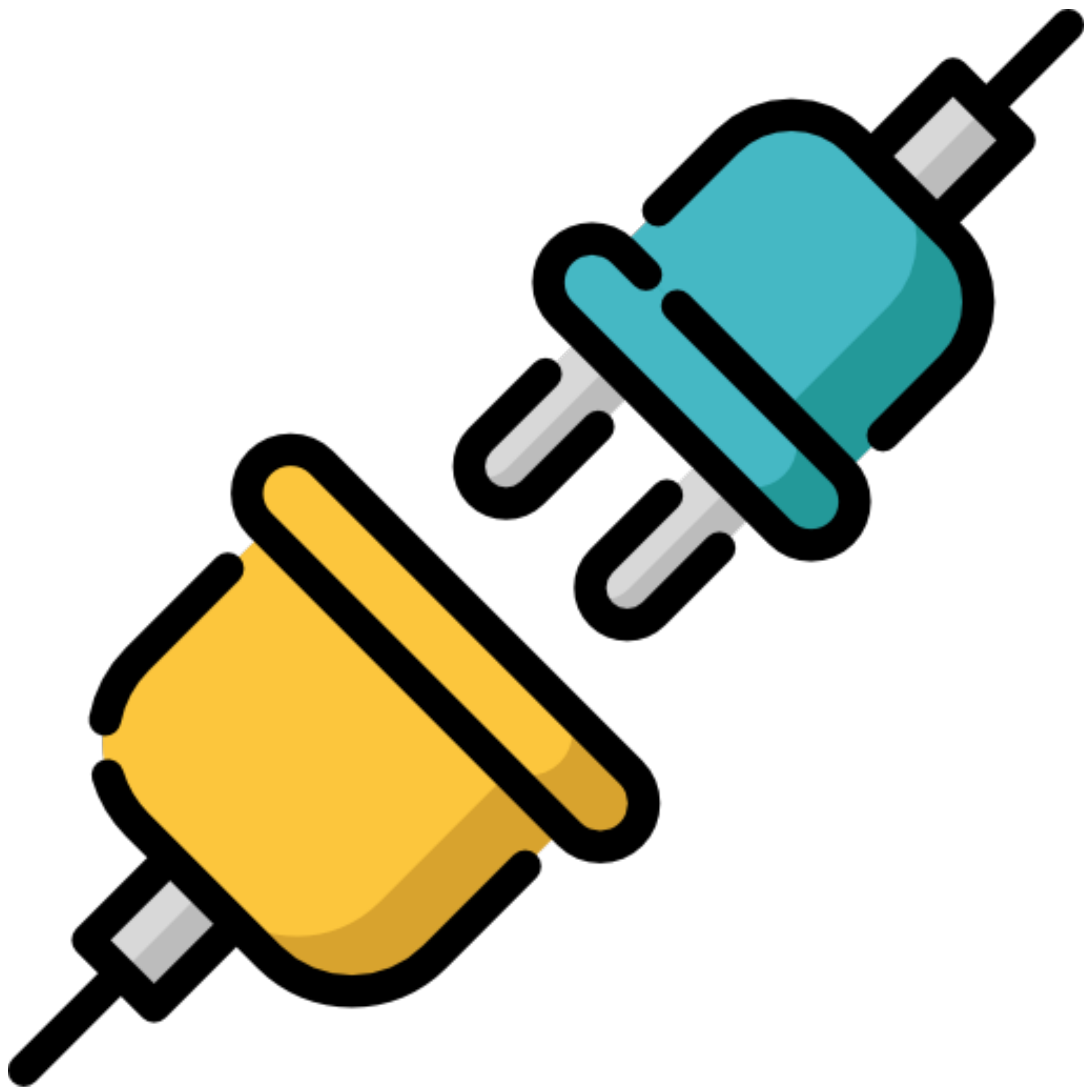
Physics plugin so that no further training is needed?

Pre-trained Kinematics
Generative Model

Physics plugin so that no further training is needed?

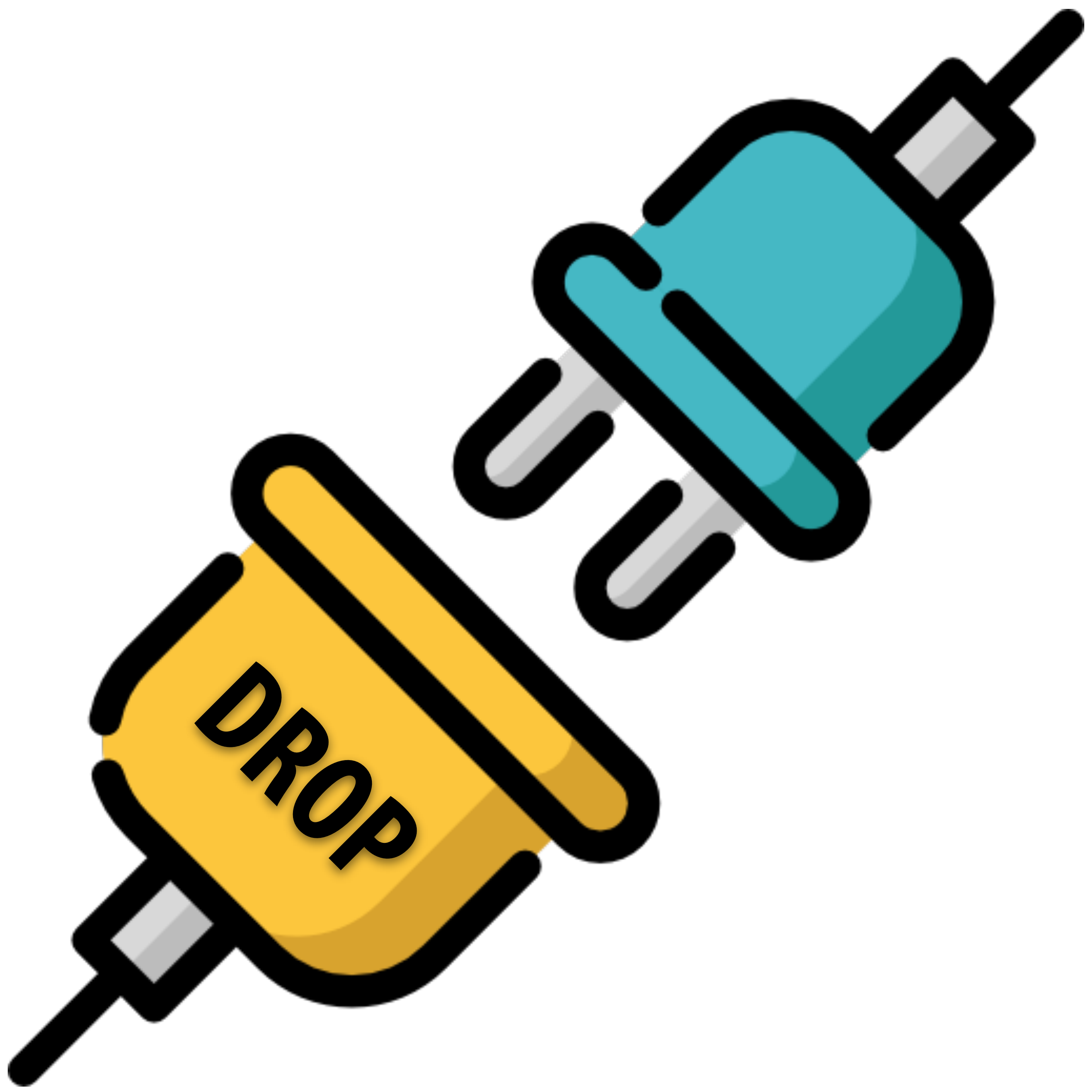
Pre-trained Kinematics
Generative Model

Introducing DROP



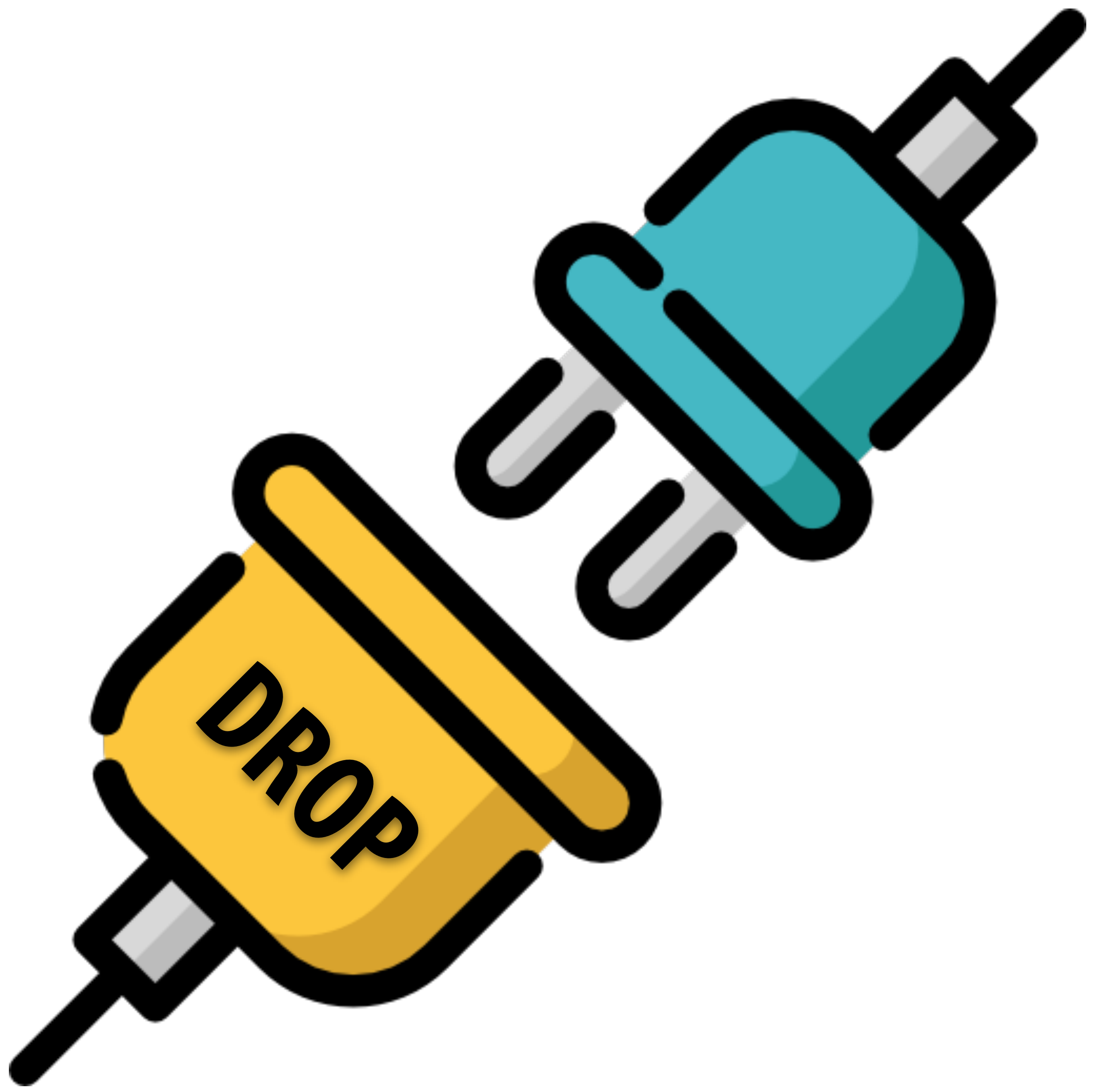
Introducing DROP

Pre-trained Kinematics
Generative Model



Introducing DROP

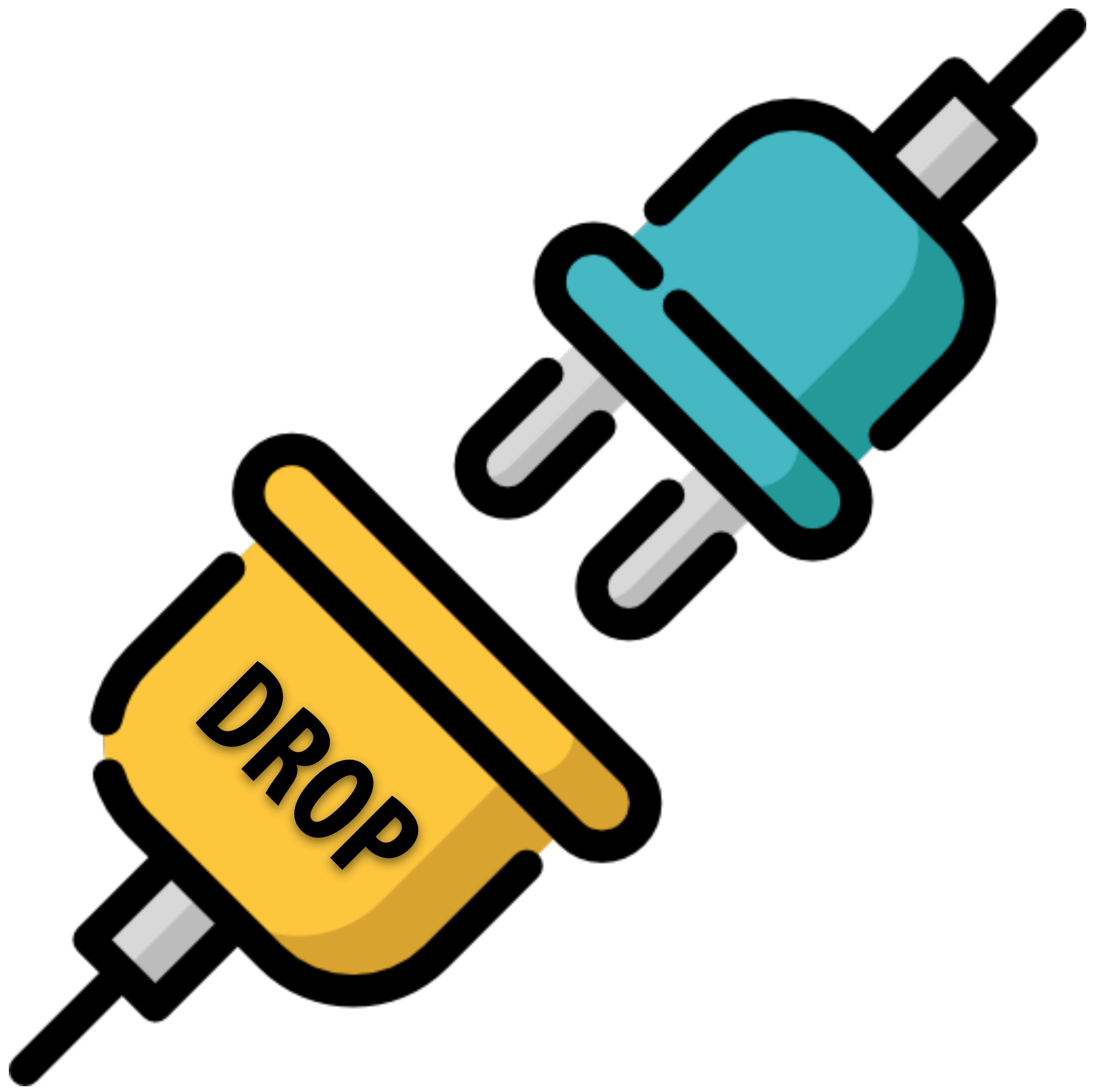
Pre-trained
Generative Model



Introducing DROP

Minimal Sim designed to fit Generative Models

Pre-trained
Generative Model

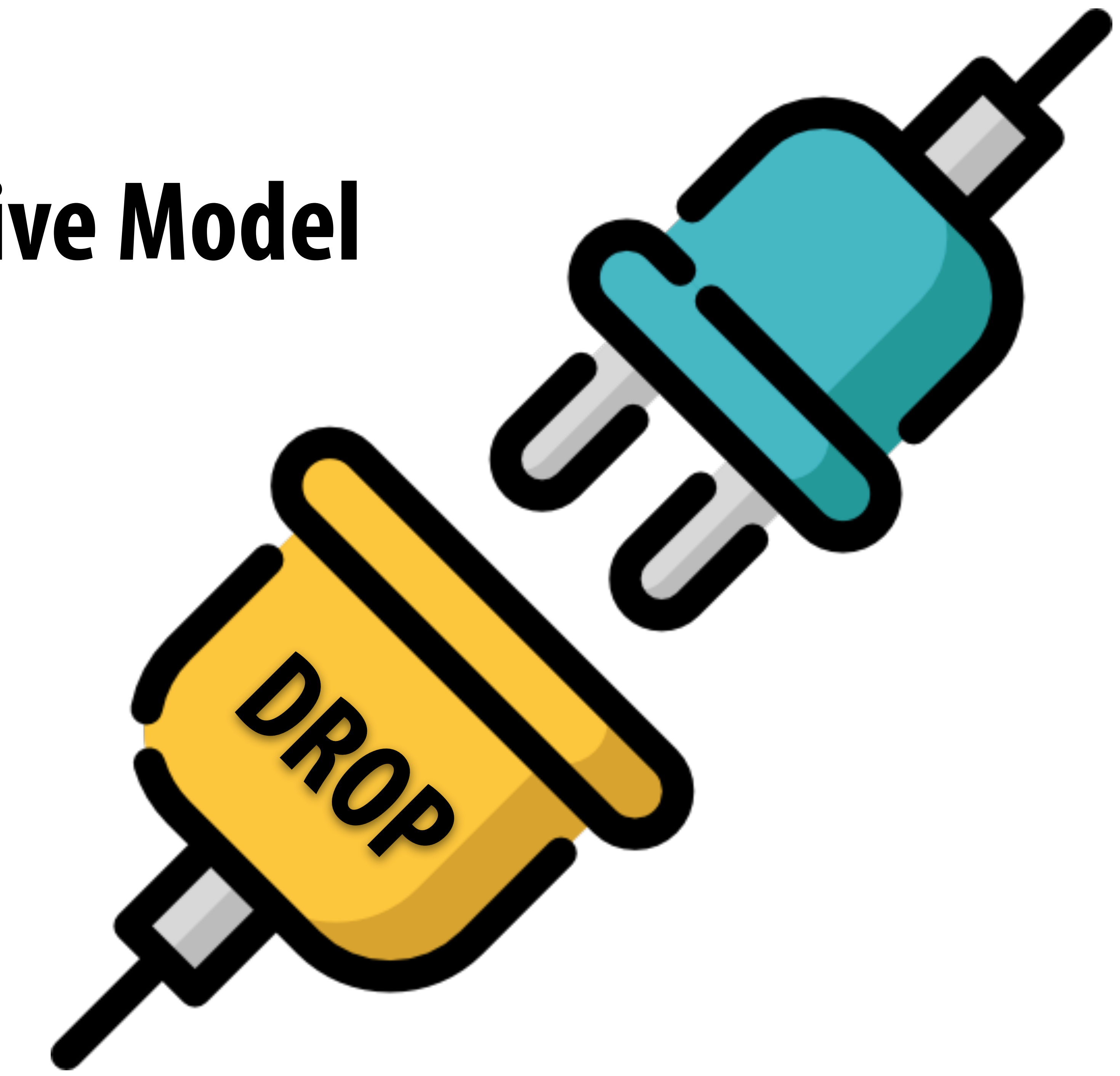


Introducing DROP

Minimal Sim designed to fit Generative Models

Plug in any pre-trained autoregressive Generative Model

Pre-trained
Generative Model



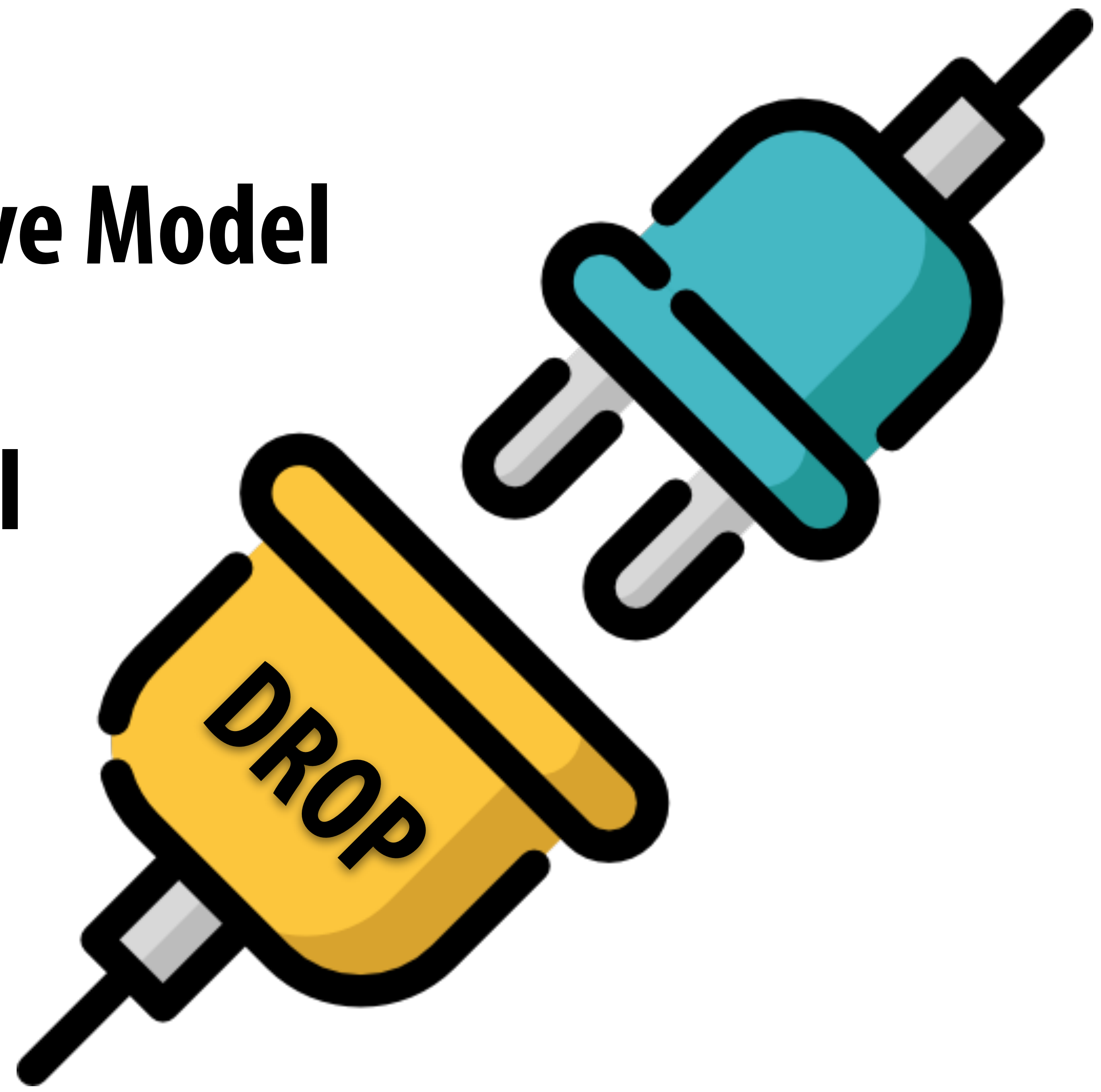
Introducing DROP

Minimal Sim designed to fit Generative Models

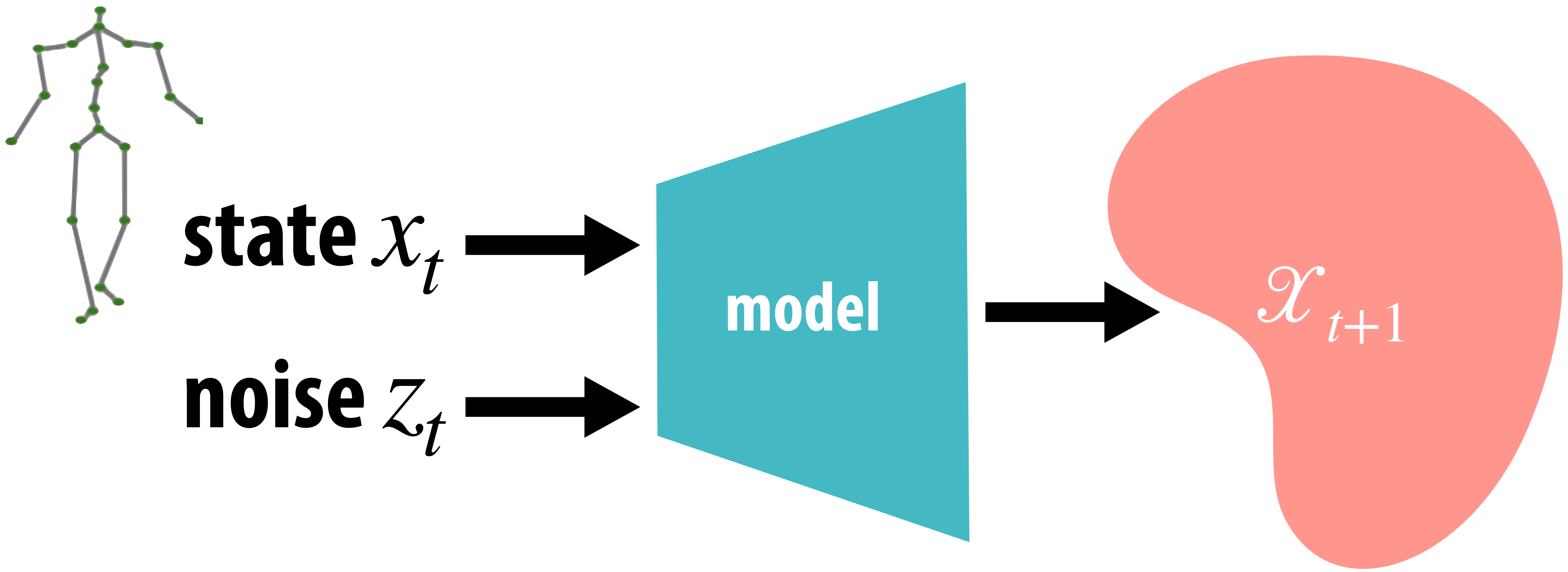
Plug in any pre-trained autoregressive Generative Model

Scalability fully inherited from Generative Model

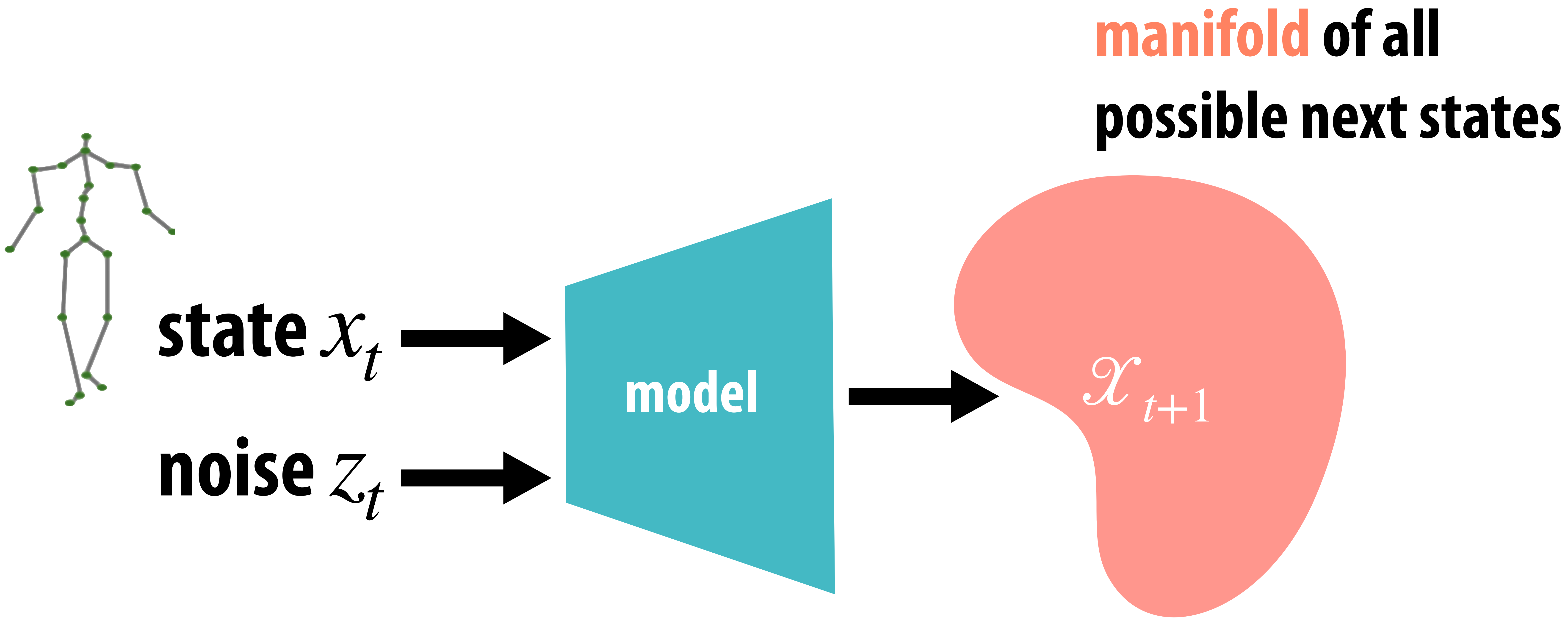
Pre-trained
Generative Model



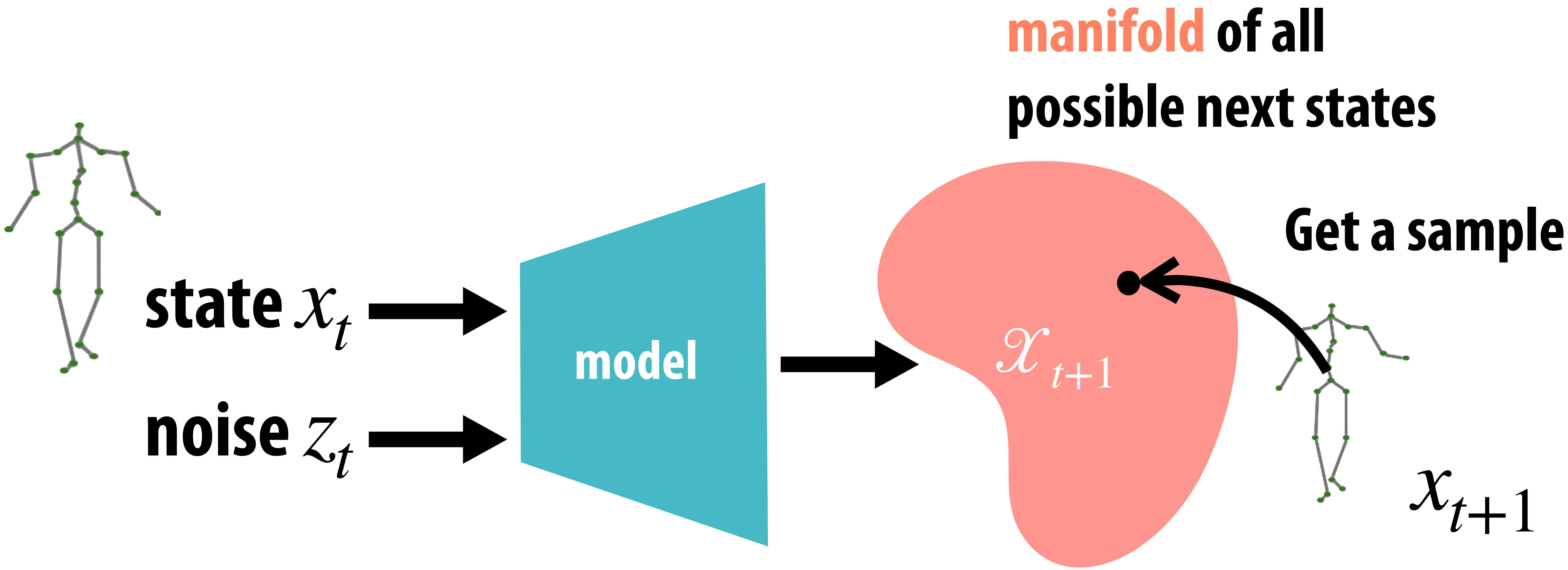
Pre-trained Autoregressive Generative Model



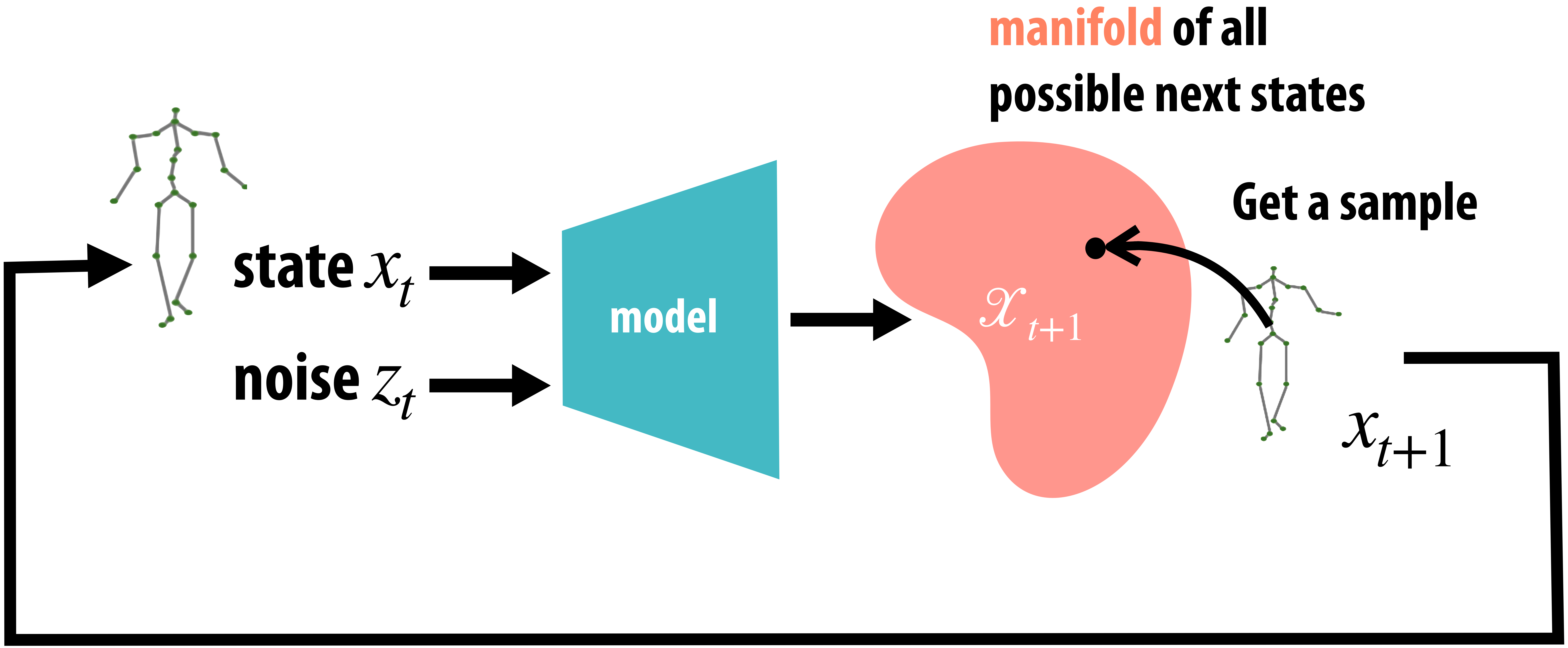
Pre-trained Autoregressive Generative Model



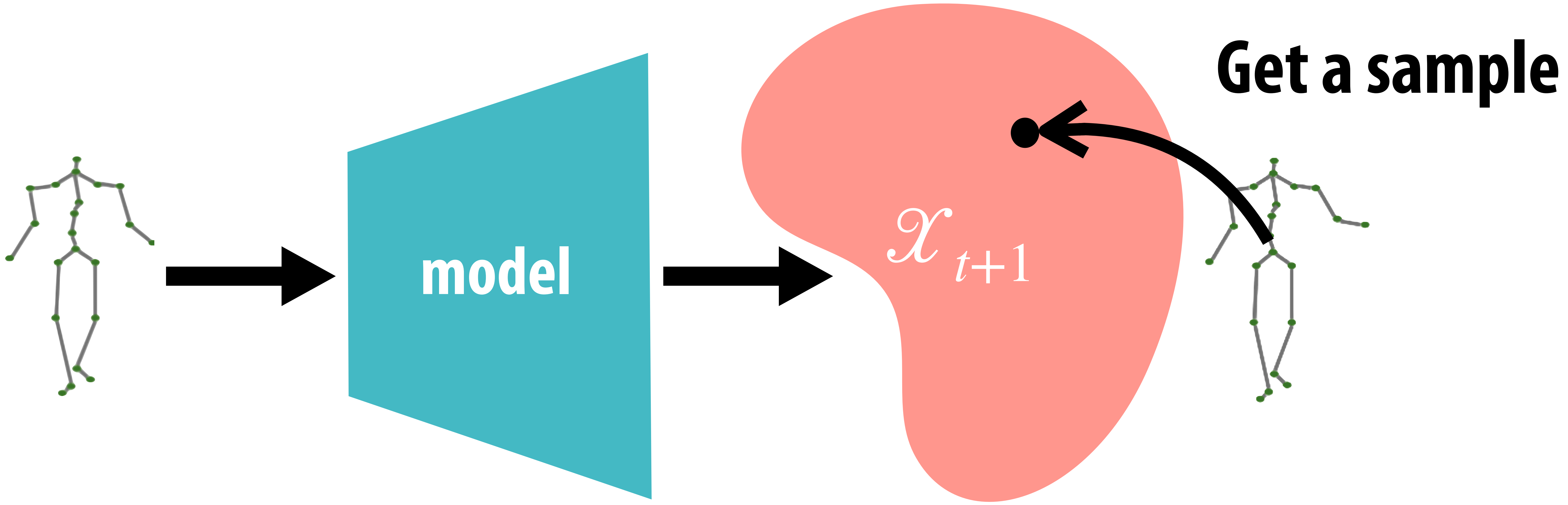
Pre-trained Autoregressive Generative Model



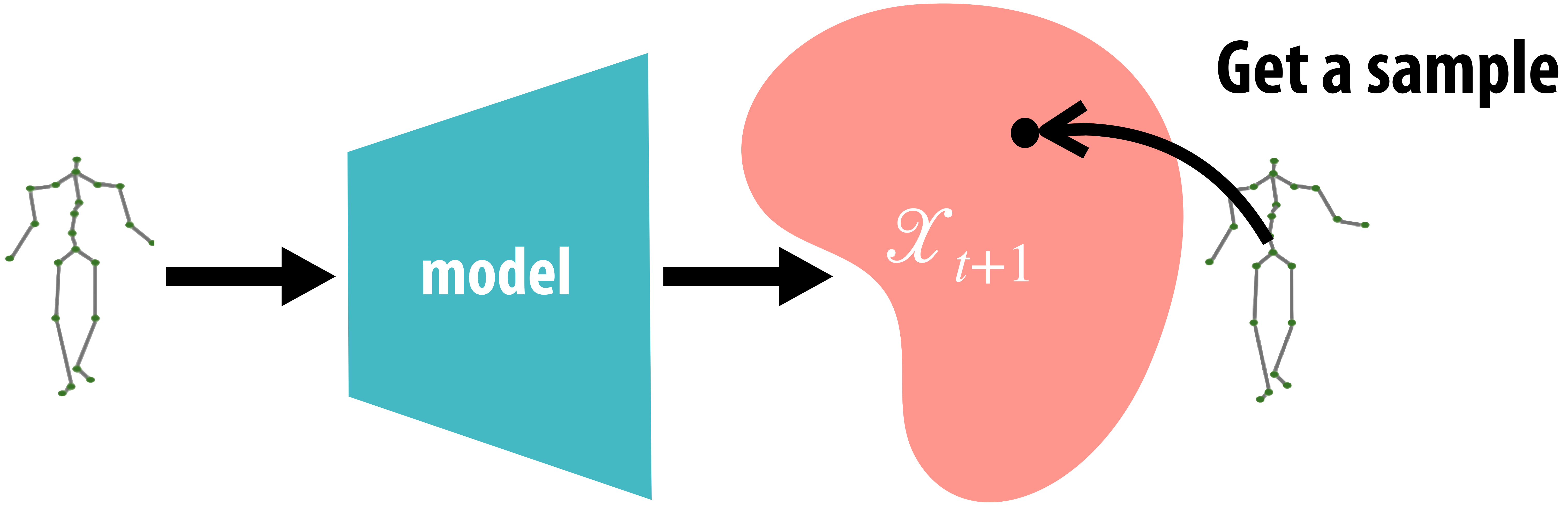
Pre-trained Autoregressive Generative Model

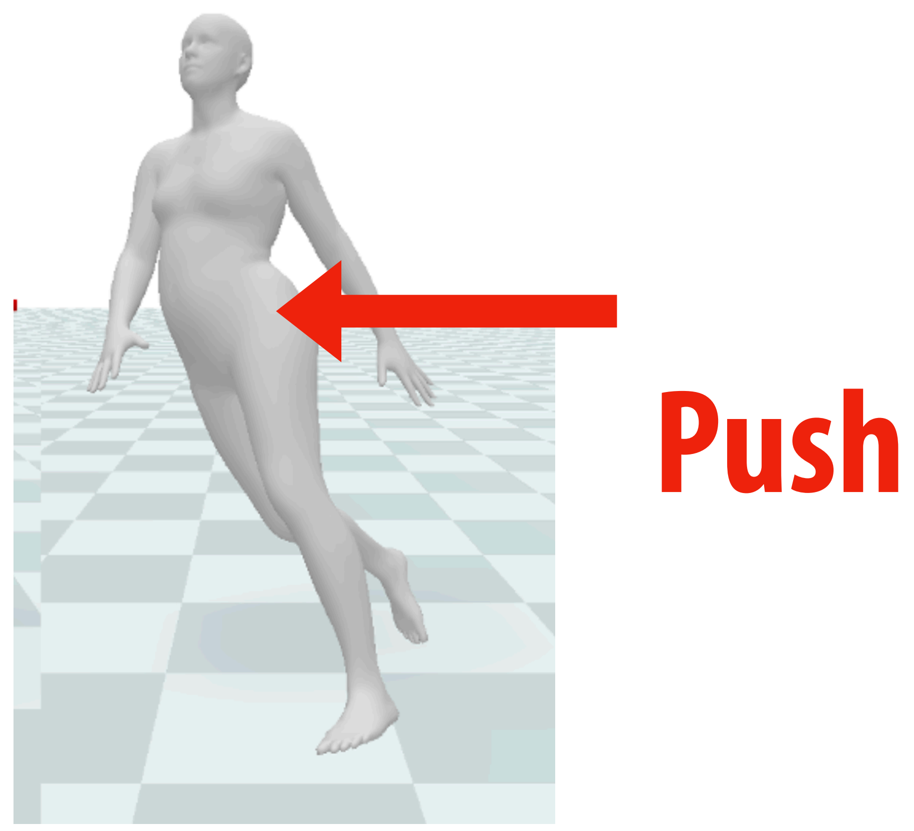


Naively, Physics as Post-processing…

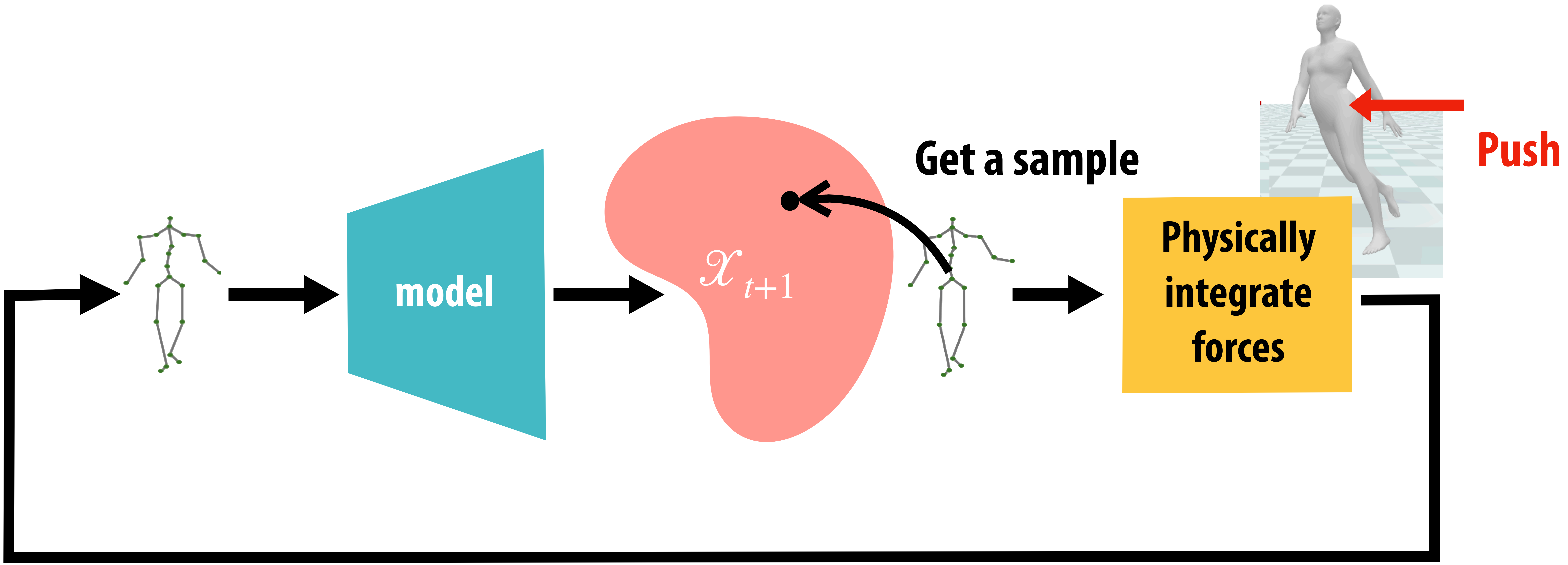


Naively, Physics as Post-processing…

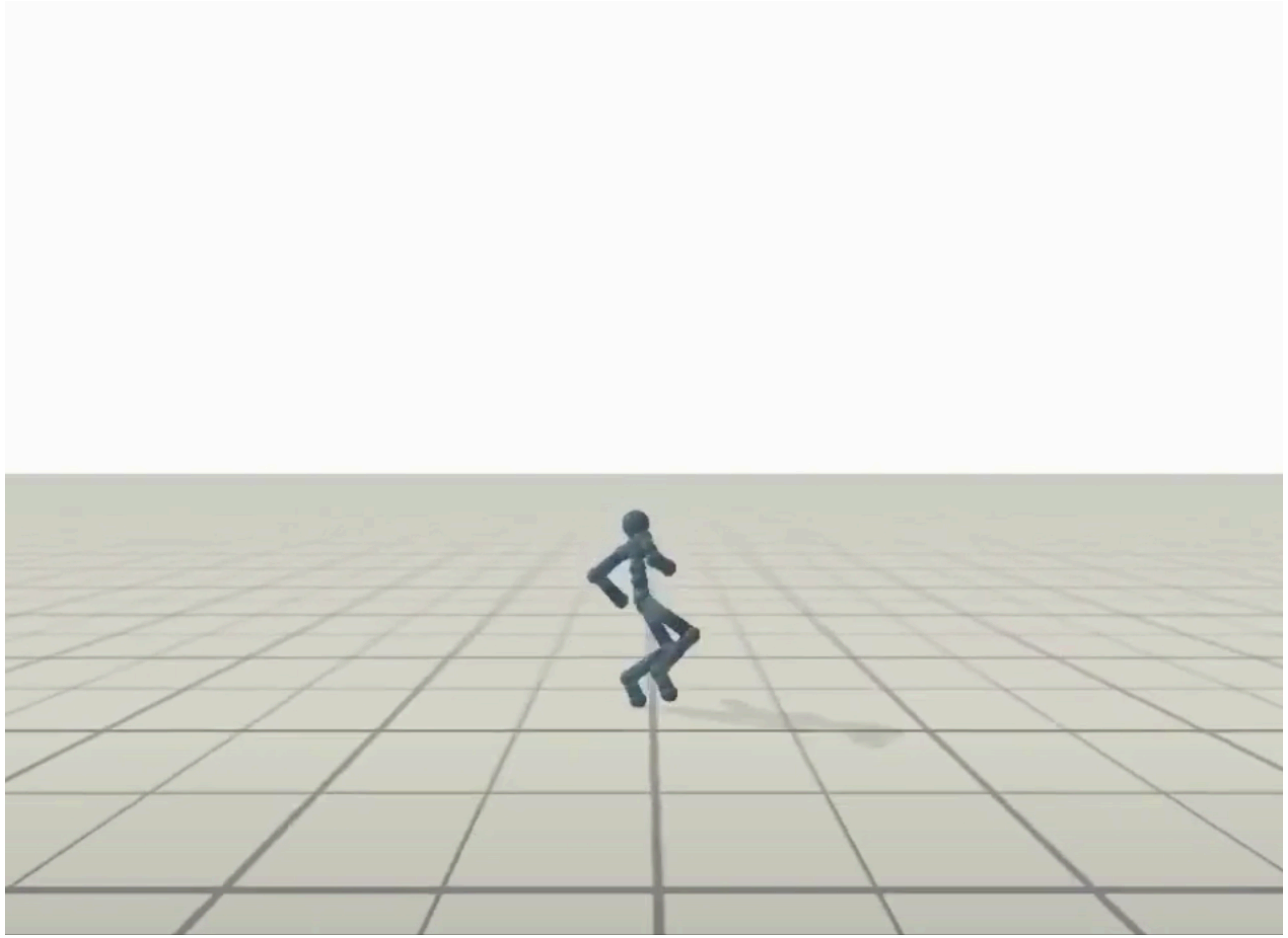




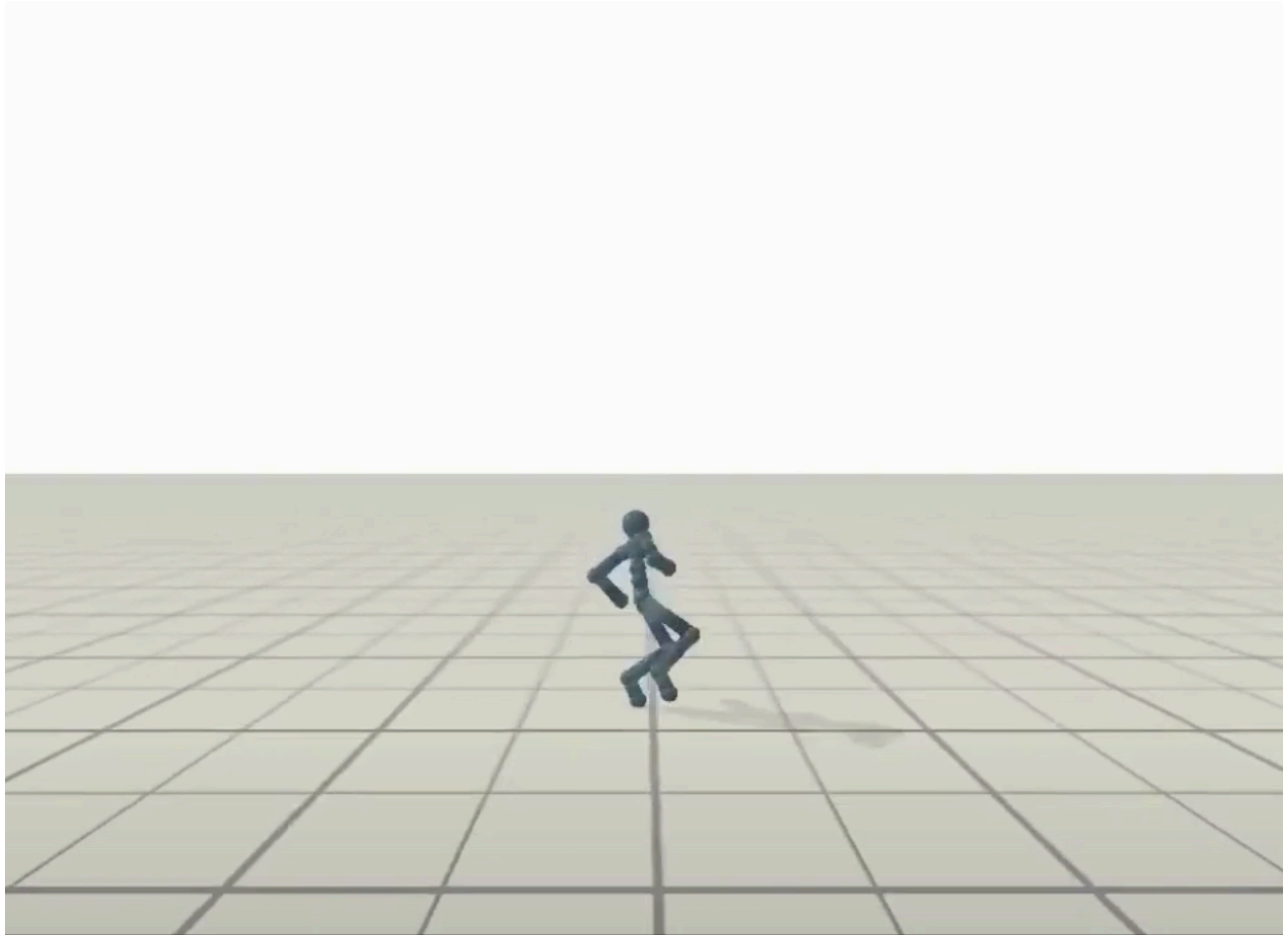
Naively, Physics as Post-processing…



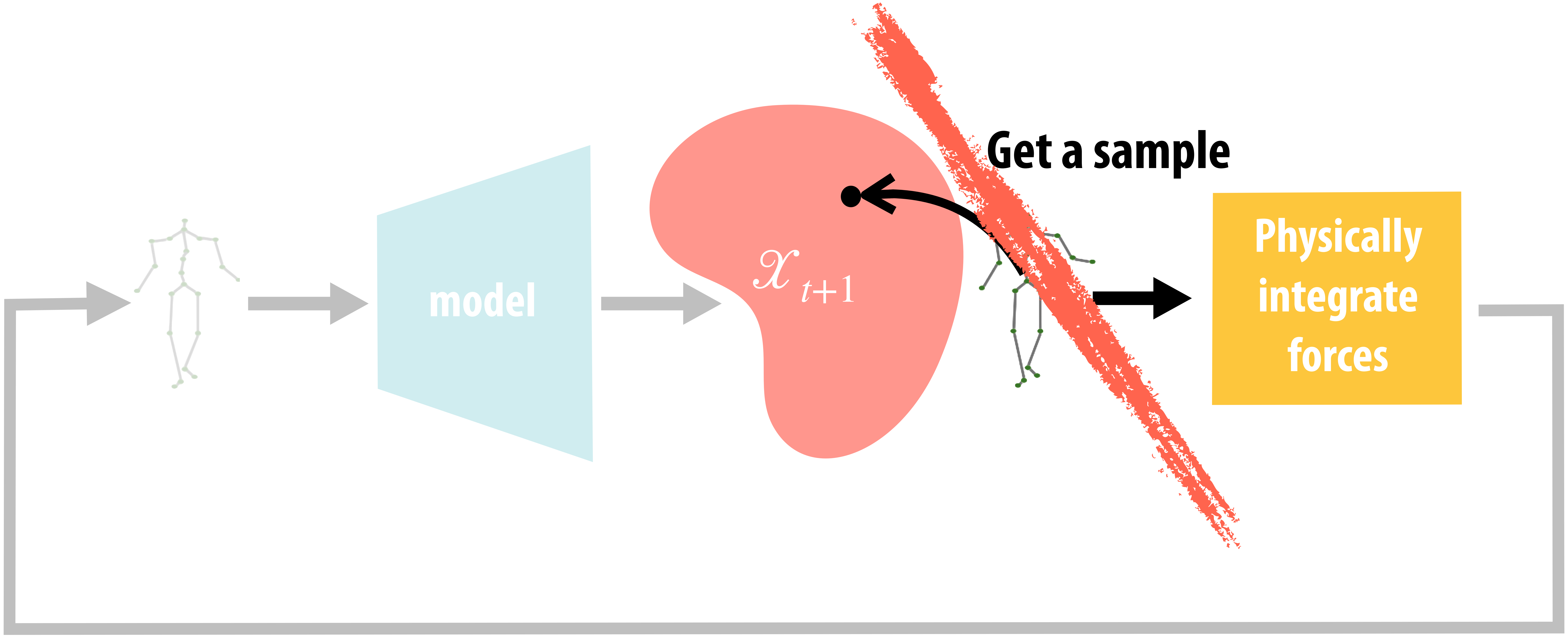
Can Lead to Model Drifting Out of Distribution



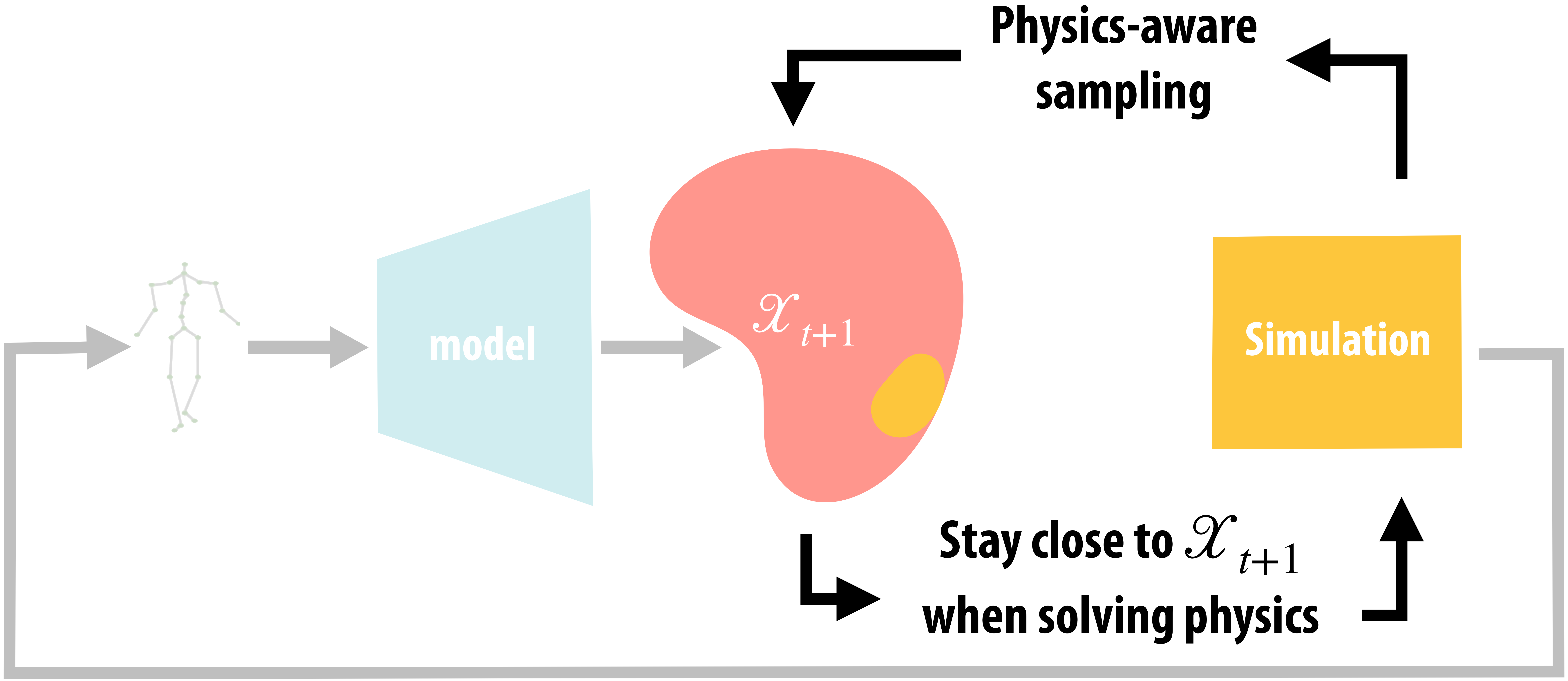
Can Lead to Model Drifting Out of Distribution



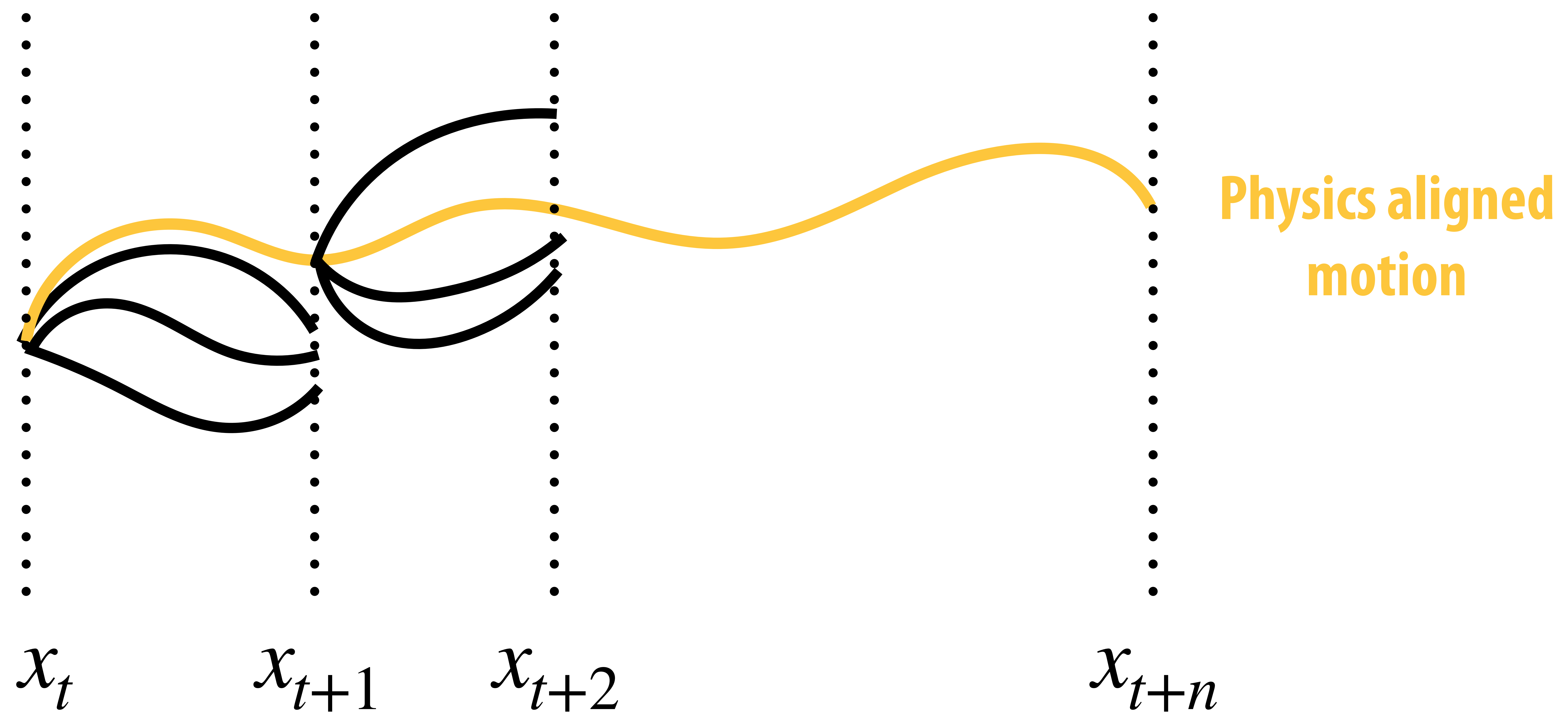
Instead of Isolated Sampling and Physics Post-processing



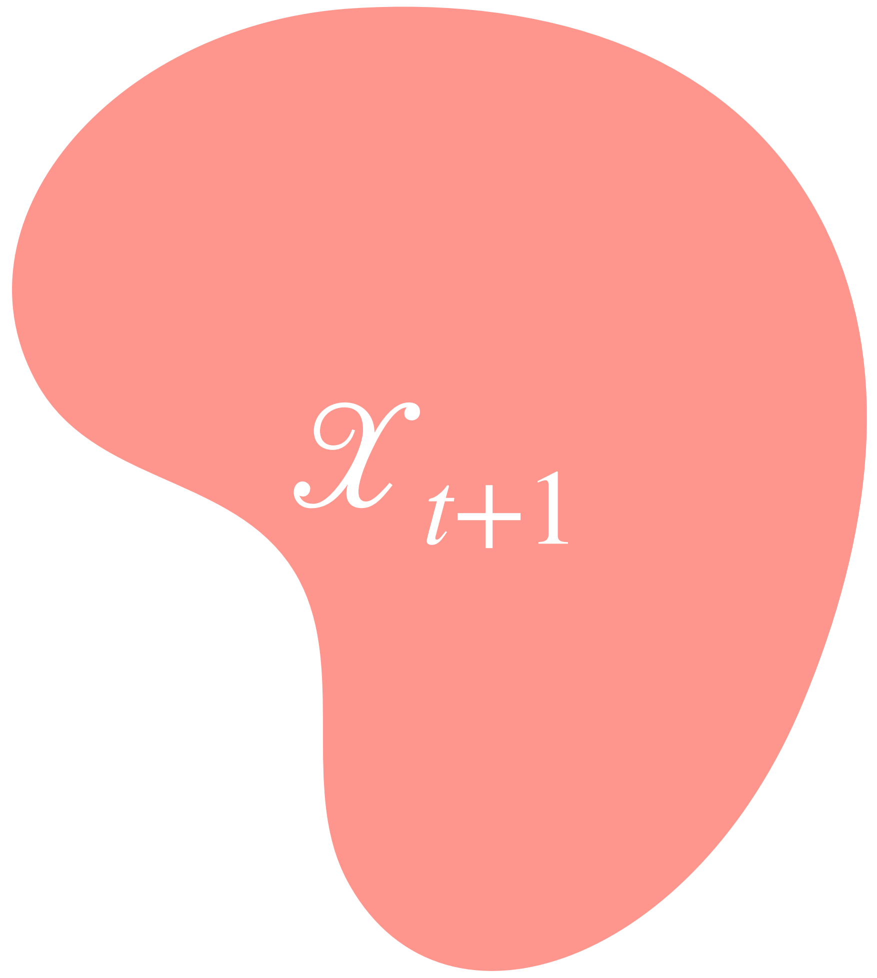
Manifold-aware Simulation



Intuitively, Need to "Align" Model Generation to Physics

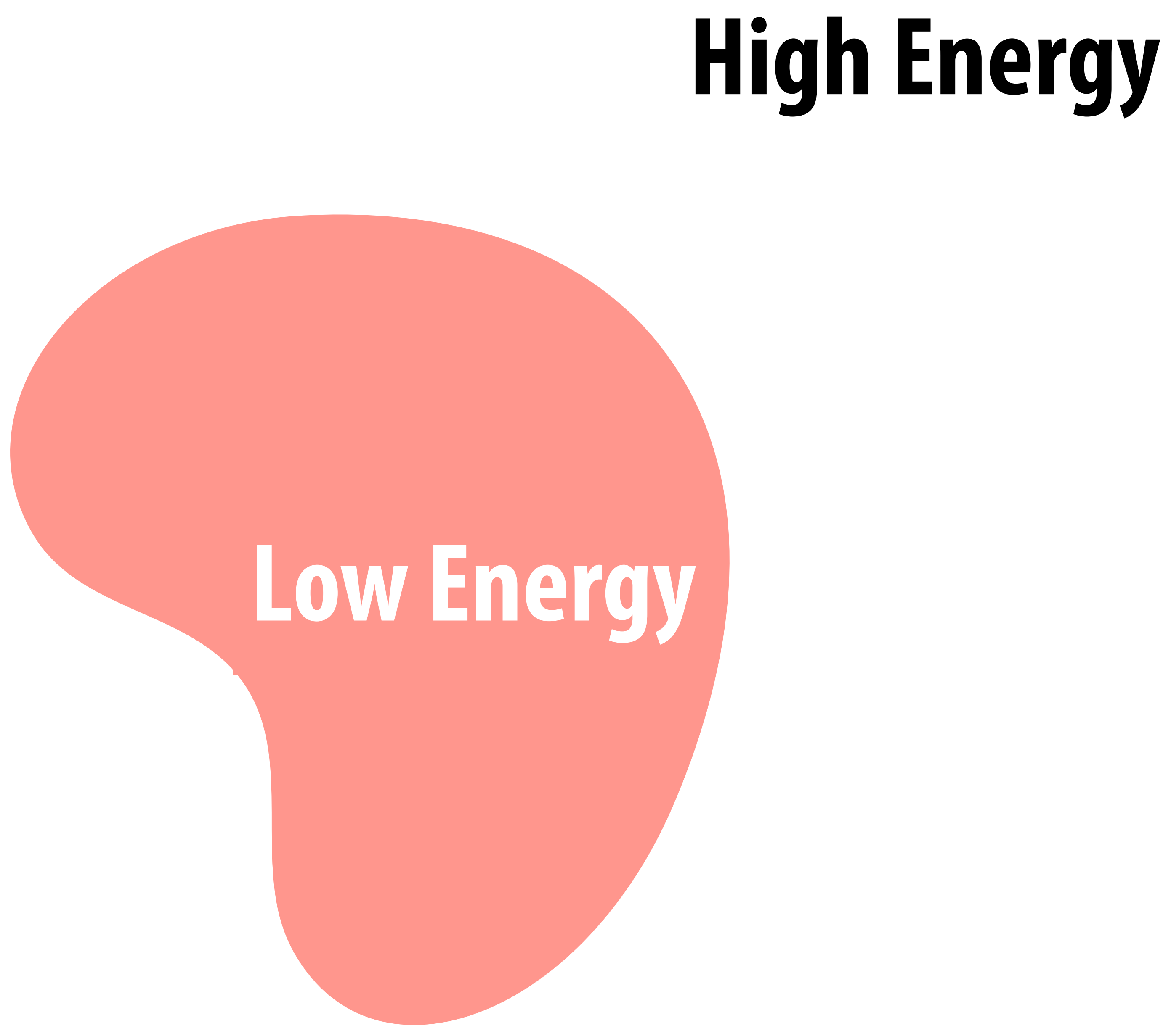


Energy-based Formulation for Model & Simulation

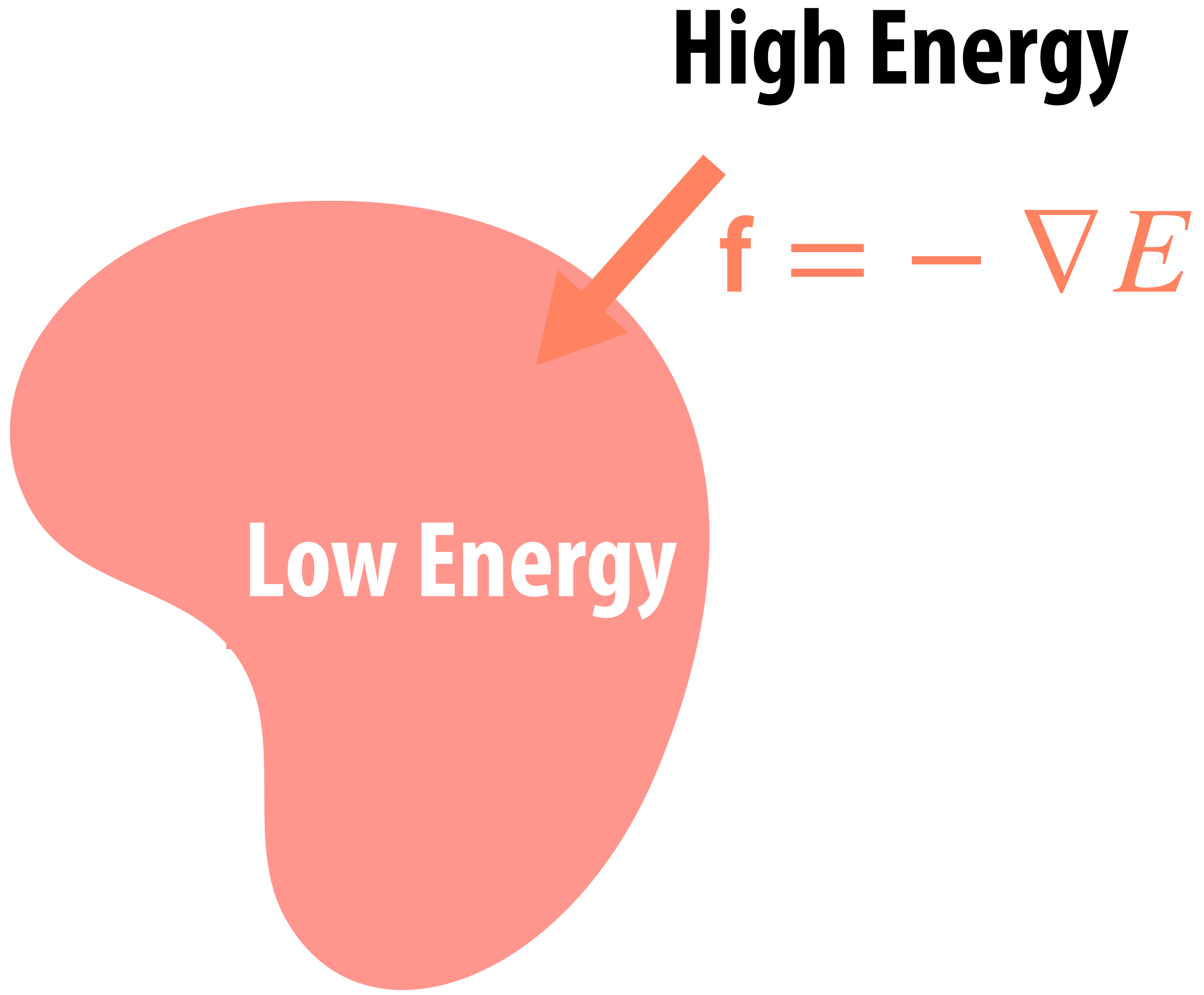


Energy-based Formulation for Model & Simulation

Energy-based Formulation for Model & Simulation

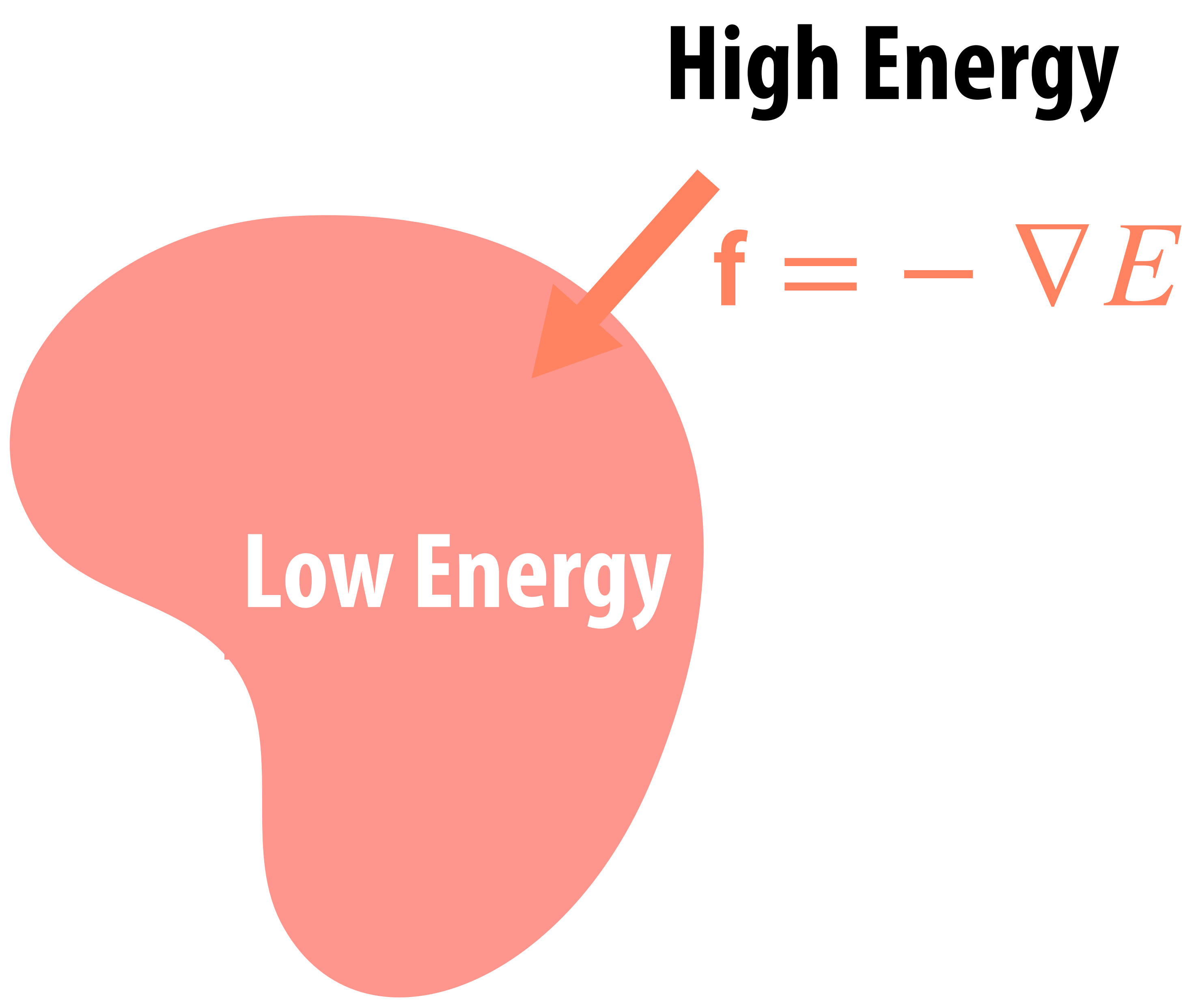


Energy-based Formulation for Model & Simulation



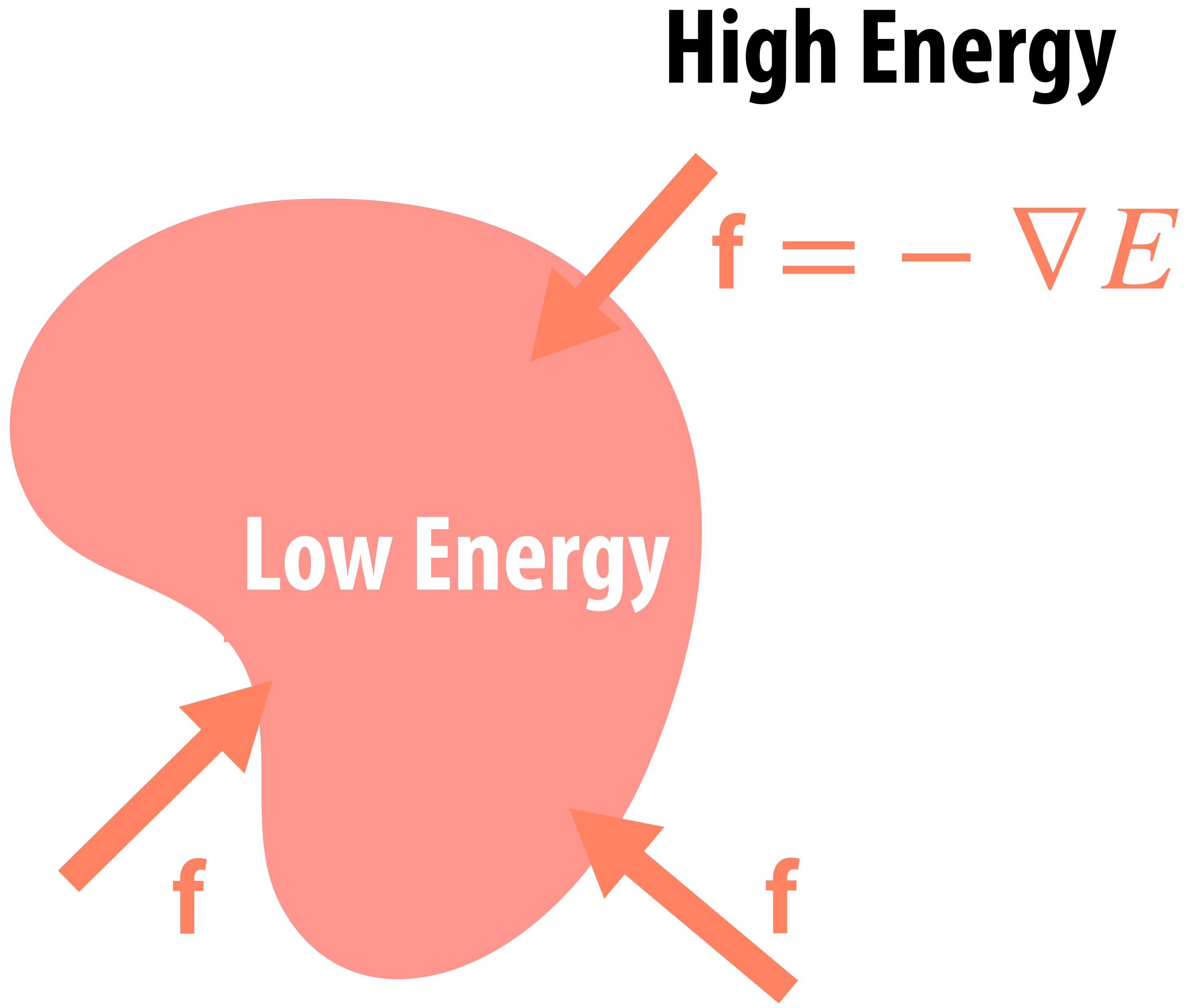
Energy-based Formulation for Model & Simulation

Akin to a control force from Generative Model

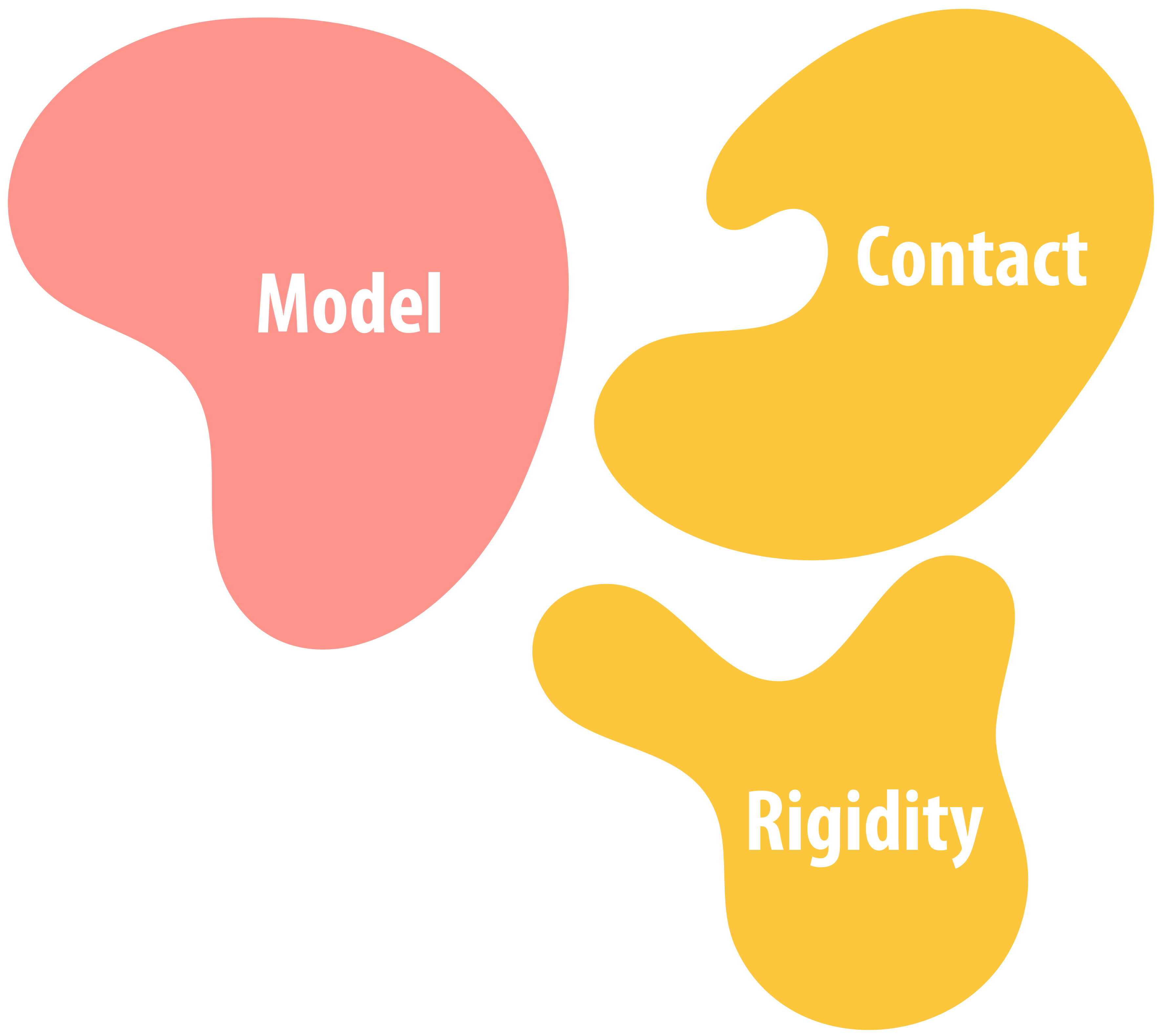


Energy-based Formulation for Model & Simulation

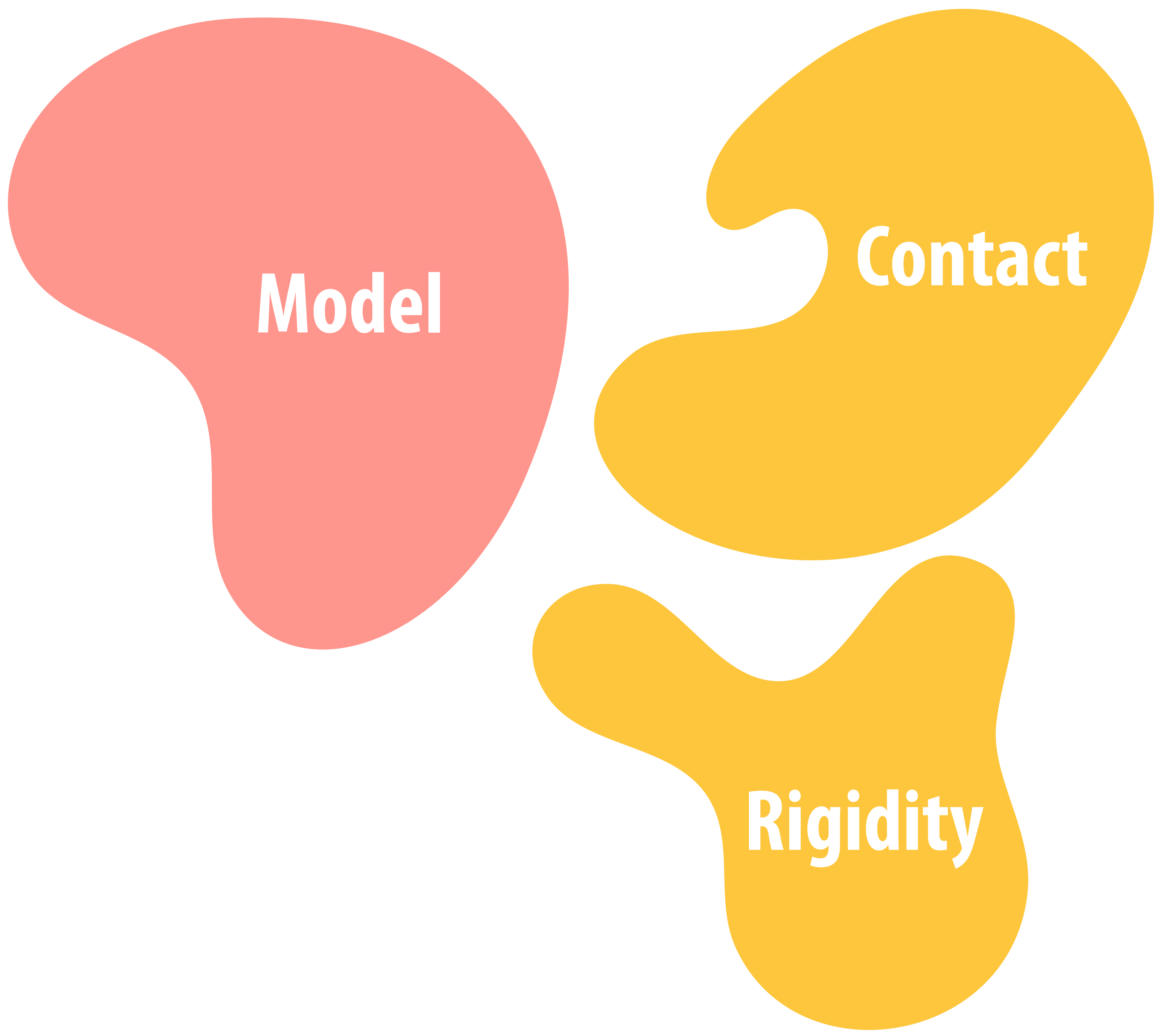
Akin to a control force from Generative Model



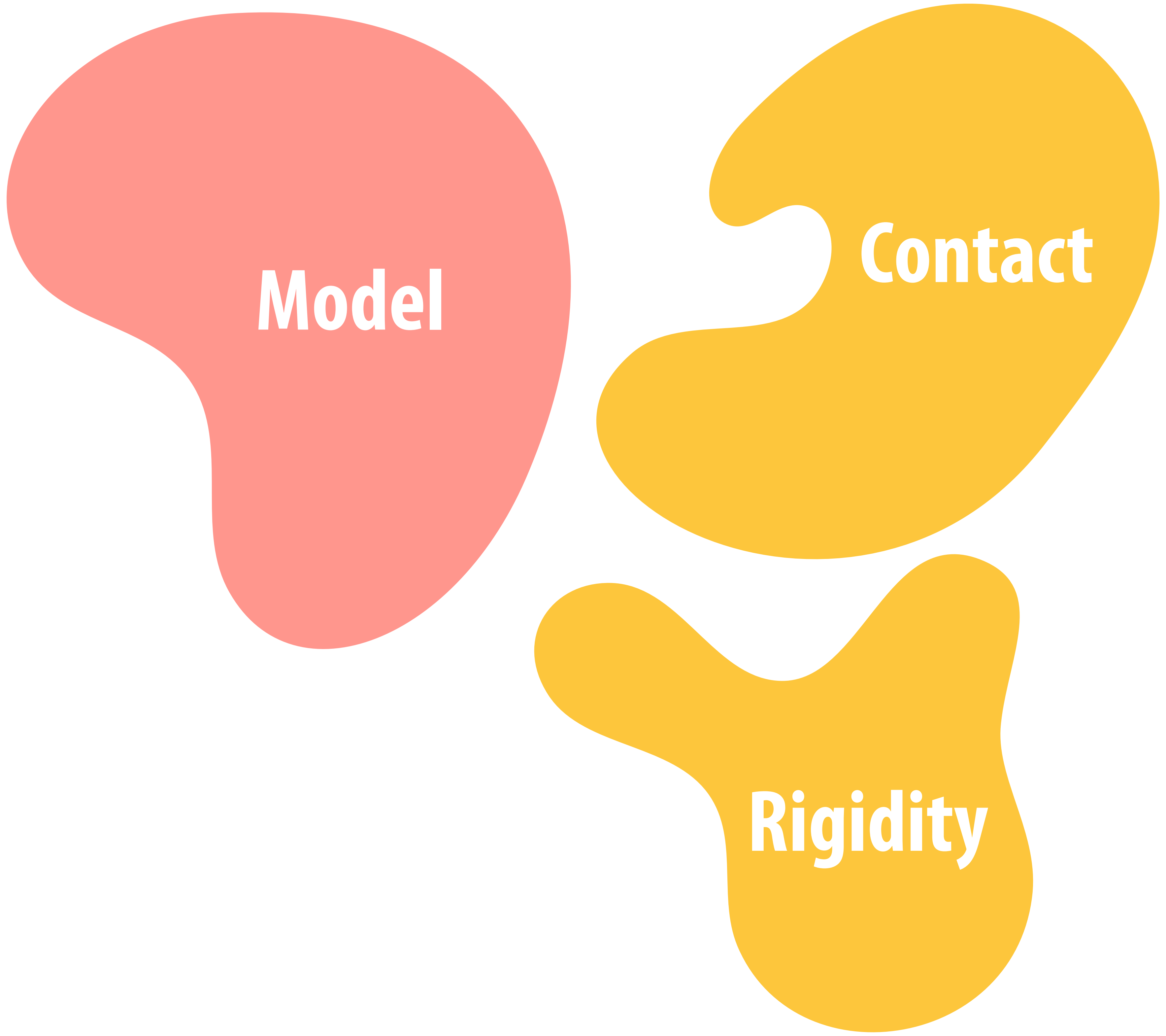
Other Energies to Align Model Generations to Physics



Other Energies to Align Model Generations to Physics

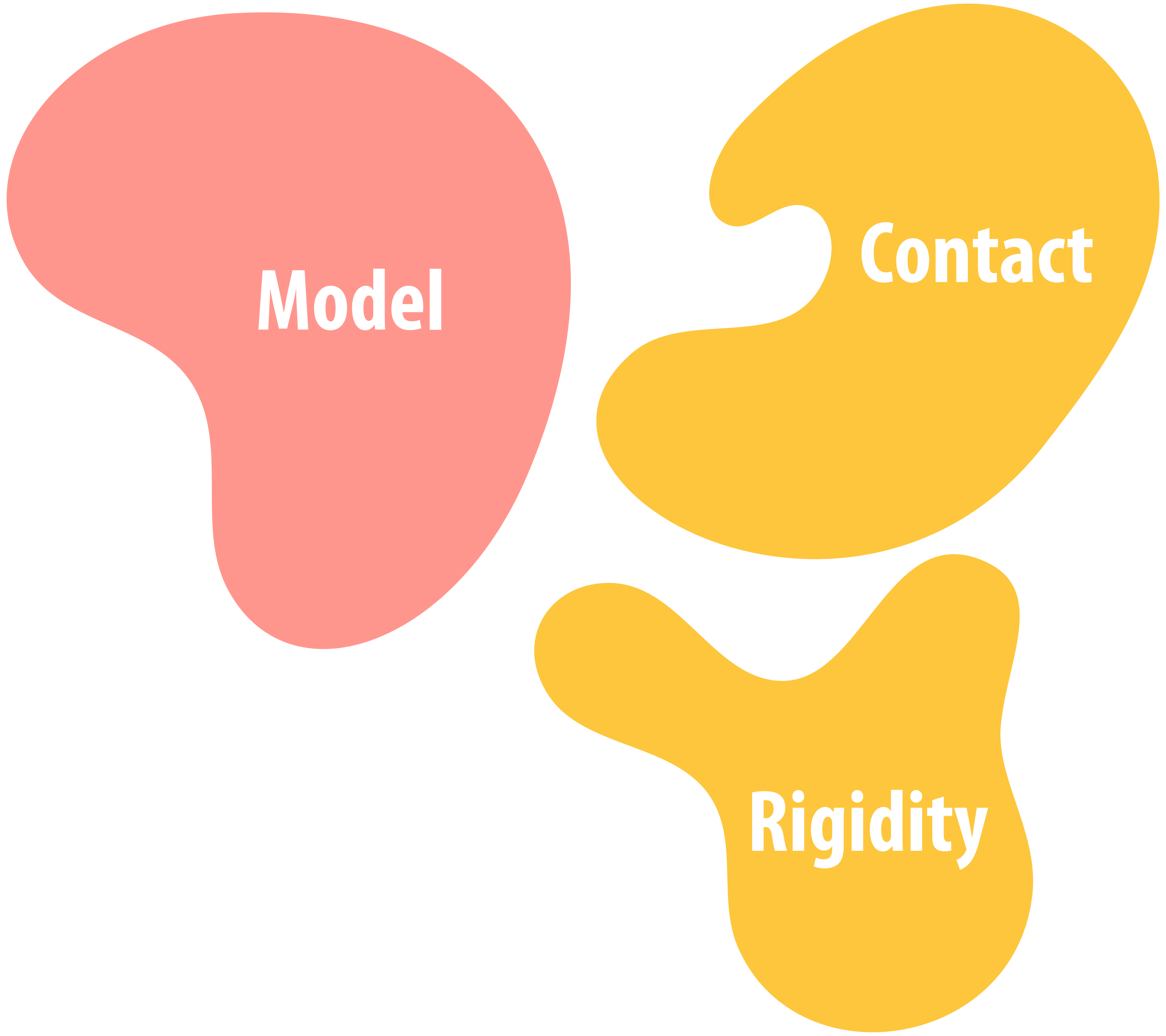


Other Energies to Align Model Generations to Physics



Other Energies to Align Model Generations to Physics

See paper all energy terms



Projective Dynamics for Simulation [Bouaziz 14]

Optimization-based (Variational) Integration:

$$x_{t+1} = \mathbf{argmin}_x \, E_{kin} + \sum_i E_i + E_m$$

Generative Model   Physical alignment   Implicit Euler integration

Projective Dynamics for Simulation [Bouaziz 14]

Optimization-based (Variational) Integration:

$$x_{t+1} = \mathbf{argmin}_x \; E_{kin} + \sum_i E_i + E_m$$

Generative Model   Physical alignment   Implicit Euler integration

Projective Dynamics for Simulation [Bouaziz 14]

Optimization-based (Variational) Integration:

$$x_{t+1} = \mathbf{argmin}_x \, E_{kin} + \sum_i E_i + E_m$$

Generative Model Physical alignment Implicit Euler integration

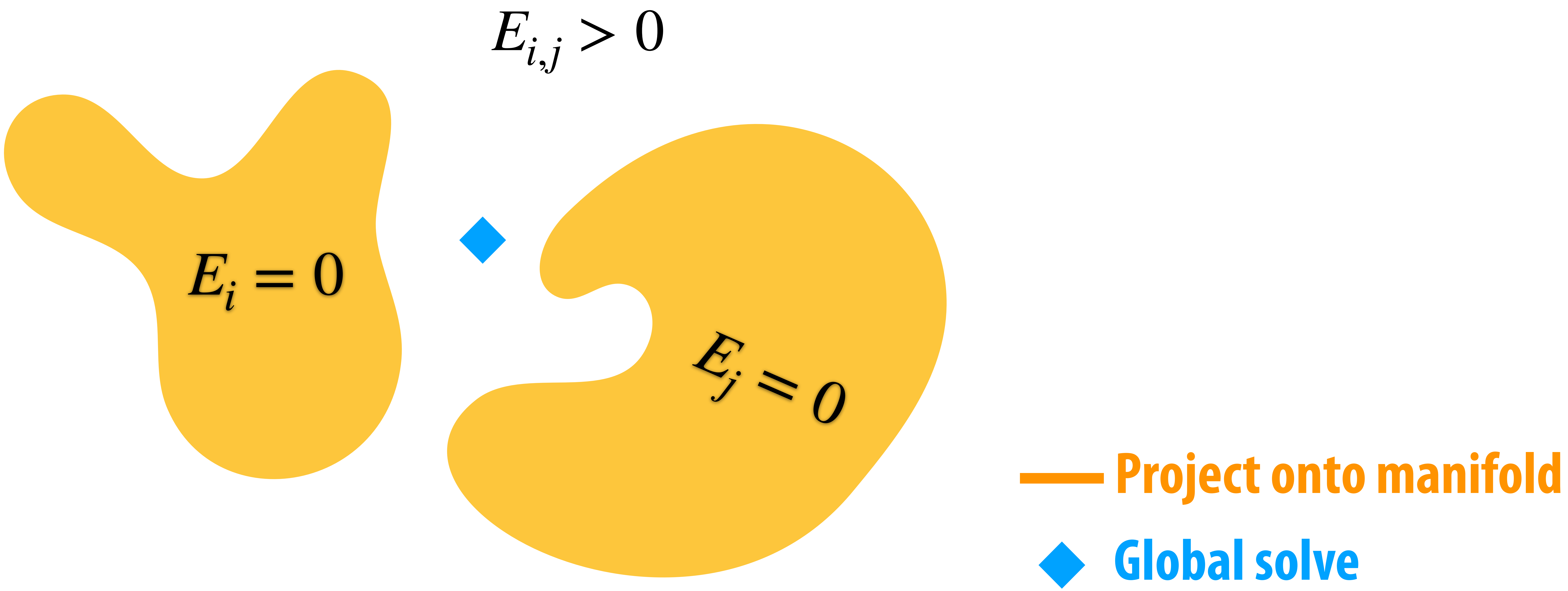
Projective Dynamics for Simulation [Bouaziz 14]

Optimization-based (Variational) Integration:

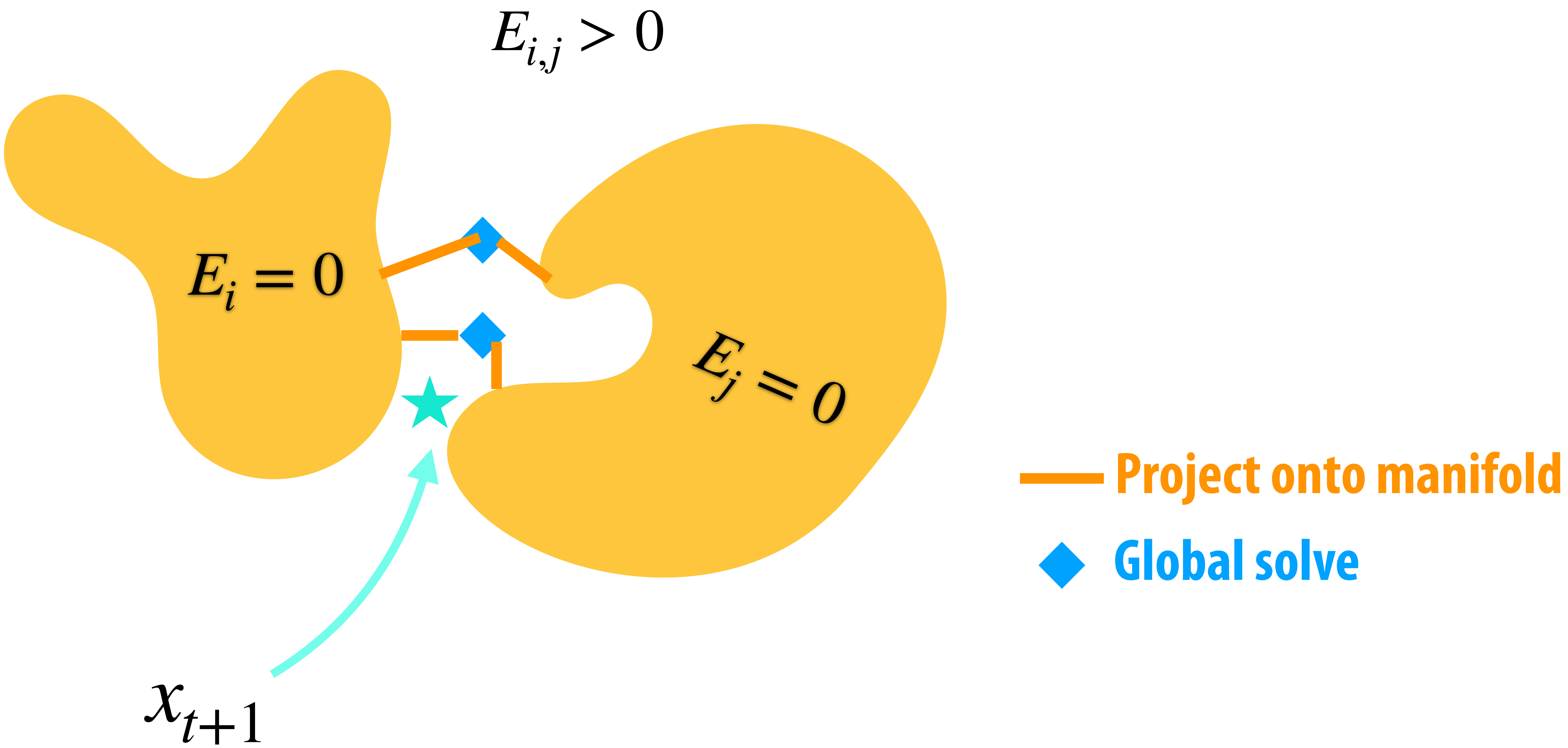
$$x_{t+1} = \textbf{argmin}_x \, E_{kin} + \sum_i E_i + E_m$$

Generative Model   Physical alignment   Implicit Euler integration

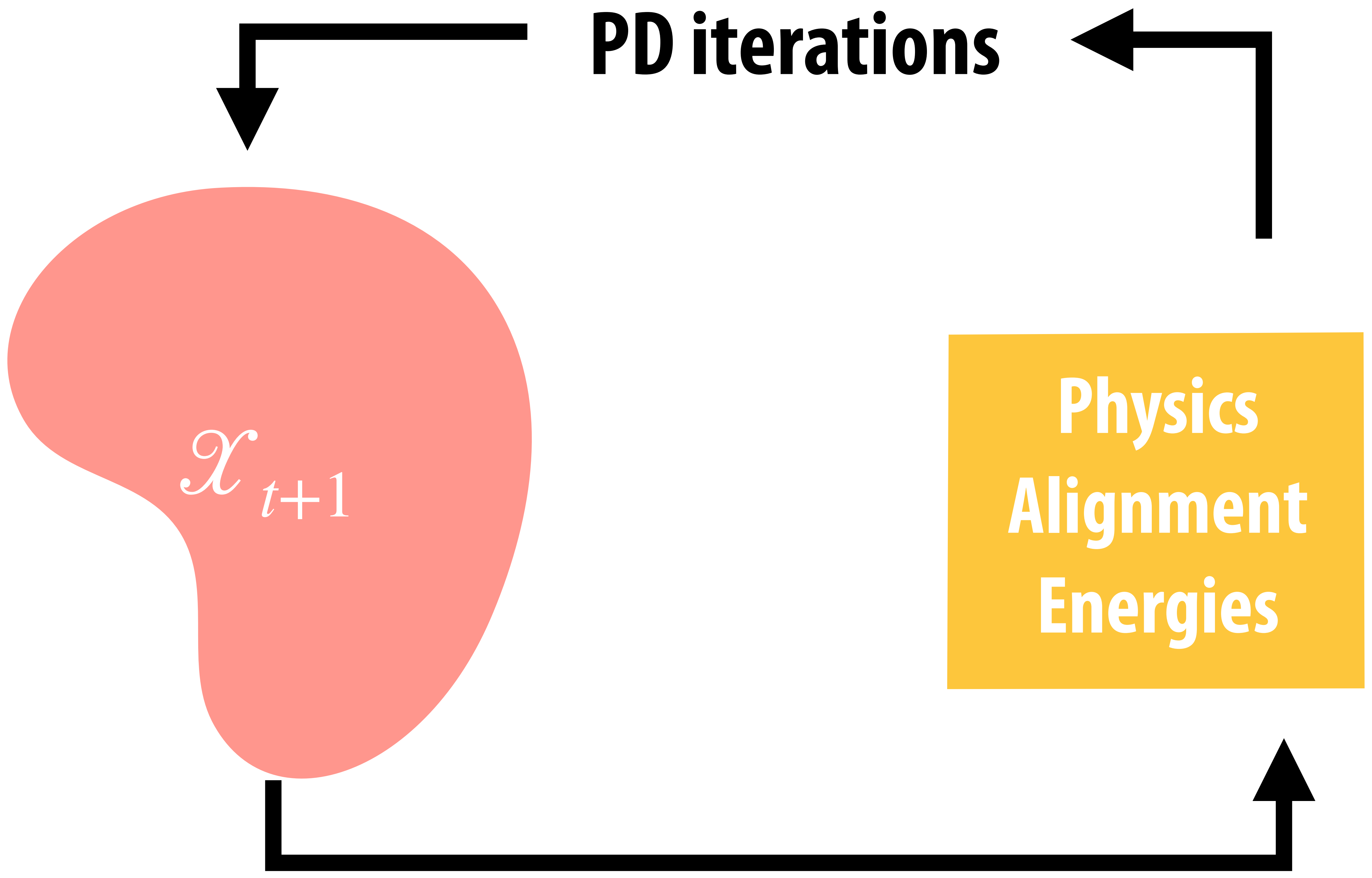
Projective Dynamics (PD) Naturally Support Manifolds



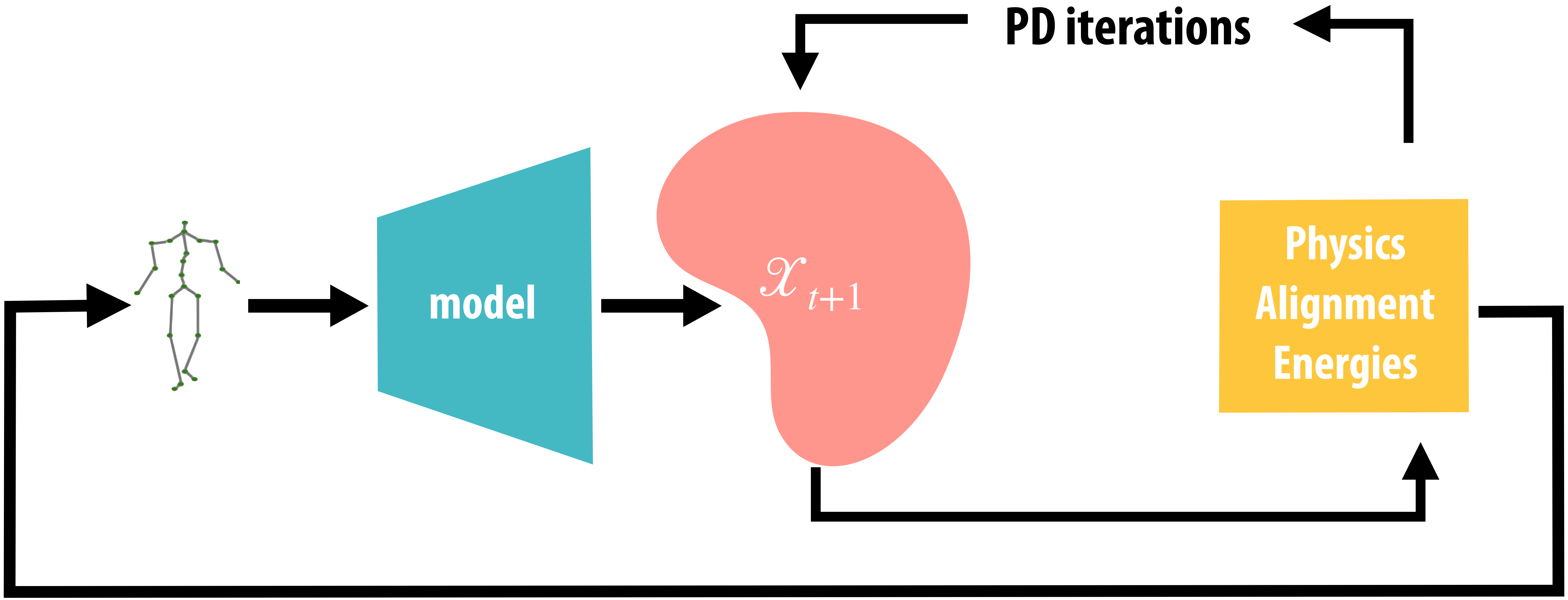
Projective Dynamics (PD) Naturally Support Manifolds



Putting Things Together



Putting Things Together



Results

Setup

Setup

Generative Model: HuMoR (ICCV'21) — trained on ~40h AMASS motion data

- Other models should work as well

Setup

Generative Model: HuMoR (ICCV'21) — trained on ~40h AMASS motion data

- Other models should work as well

Focus on showcasing dynamic responses

- That is, all demos are stress testing the low-data cases

Setup

Generative Model: HuMoR (ICCV'21) — trained on ~40h AMASS motion data

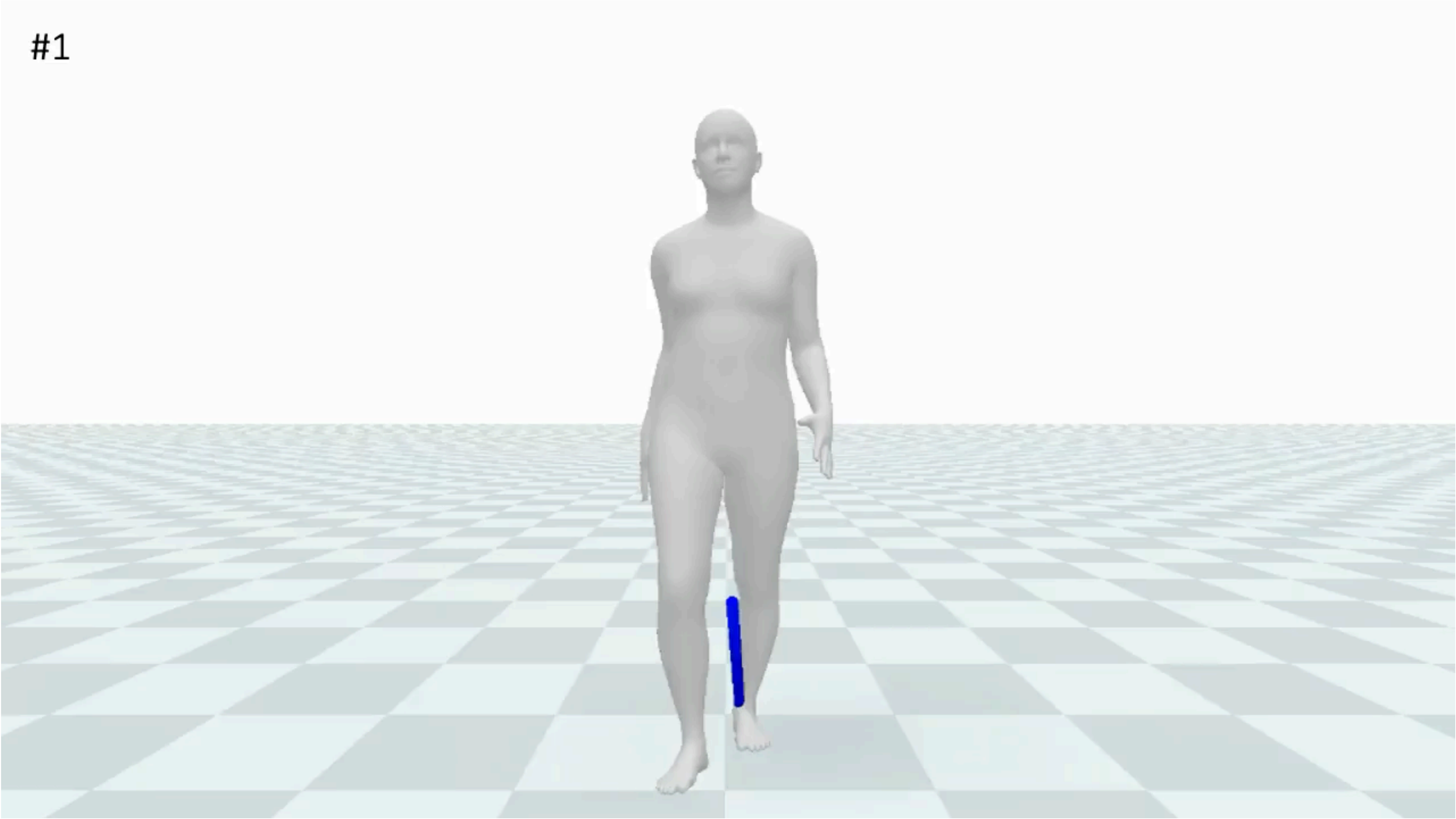
- Other models should work as well

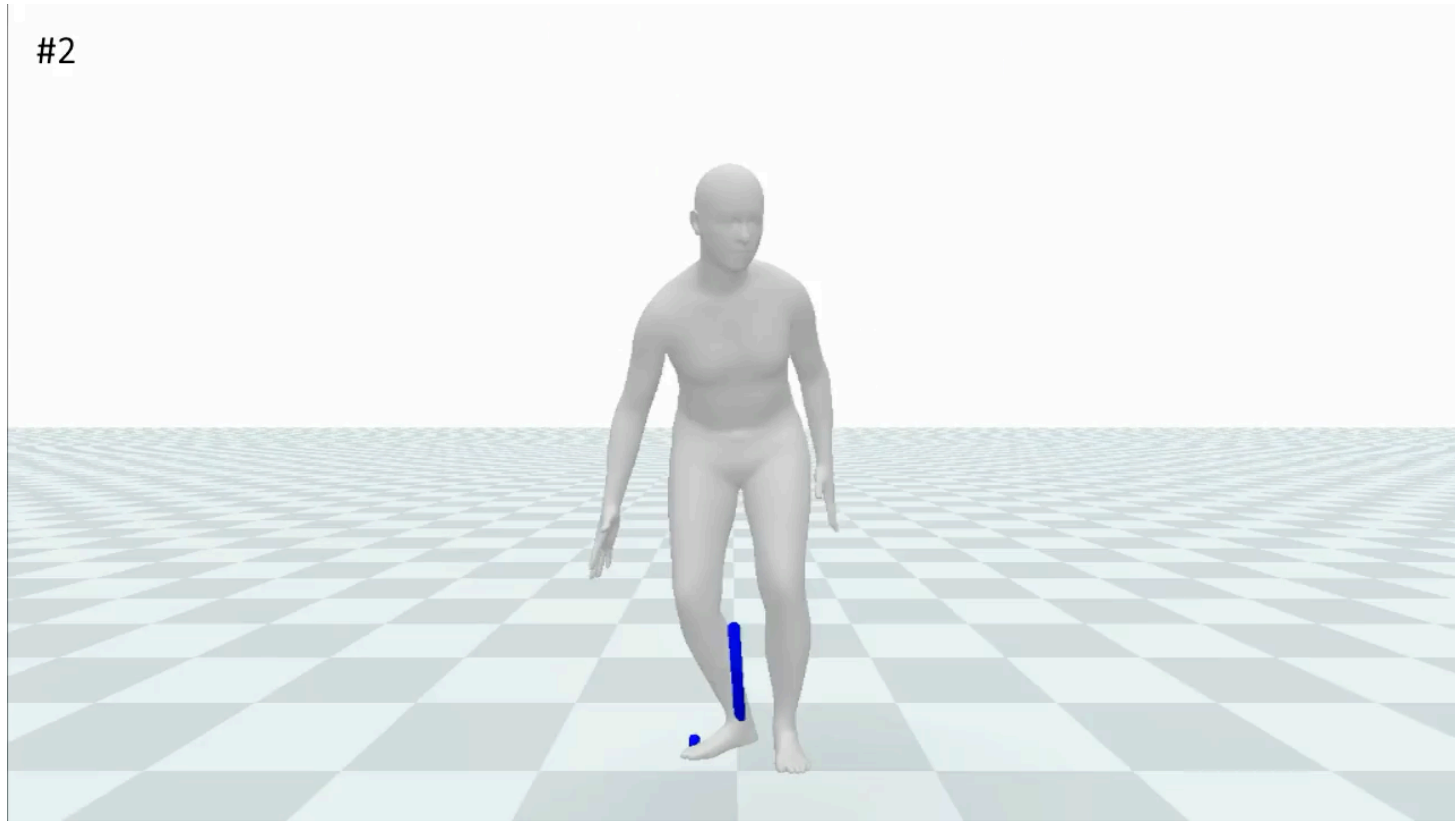
Focus on showcasing dynamic responses

- That is, all demos are stress testing the low-data cases

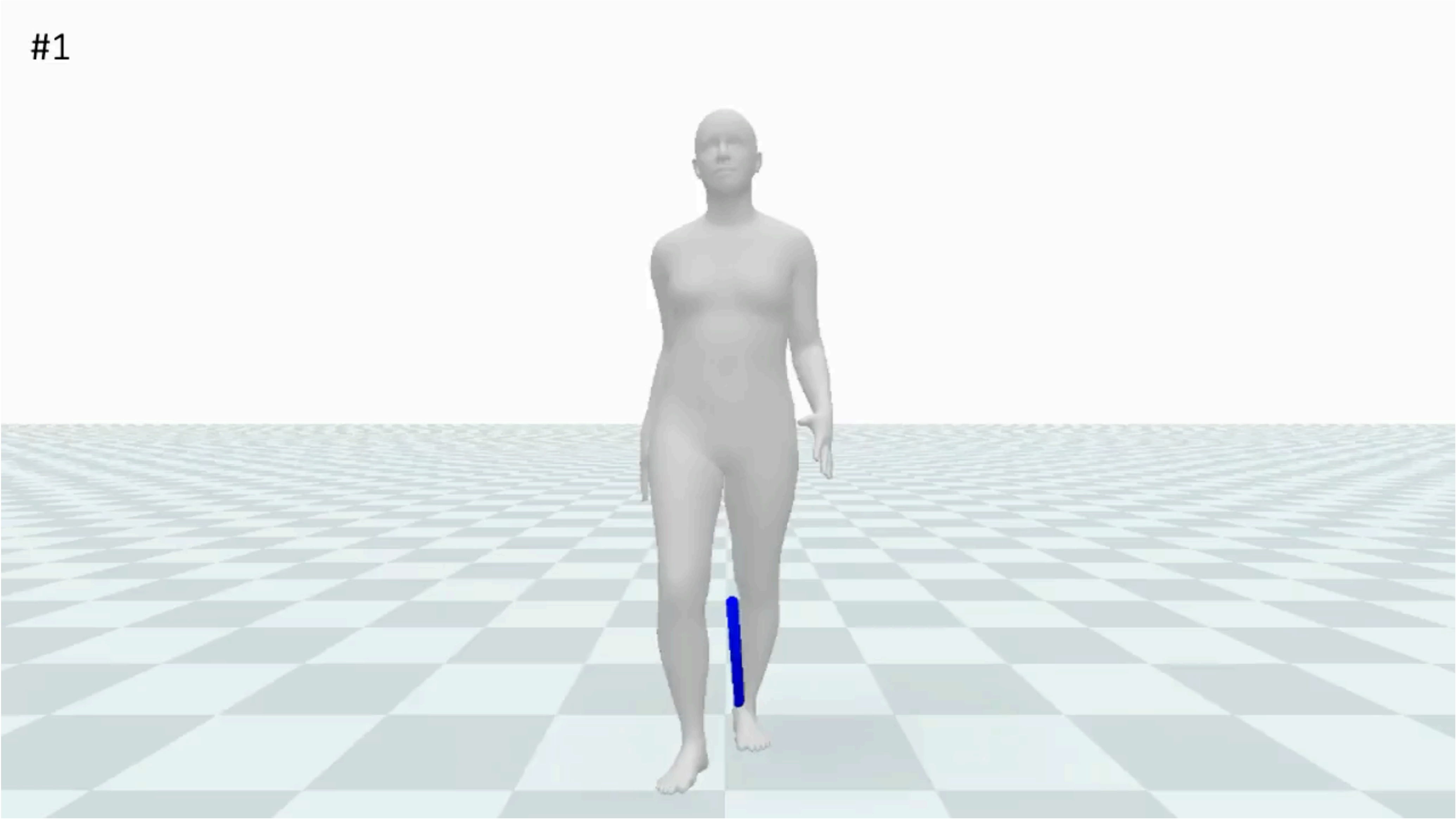
All demos are stochastically created without high-level motion planning

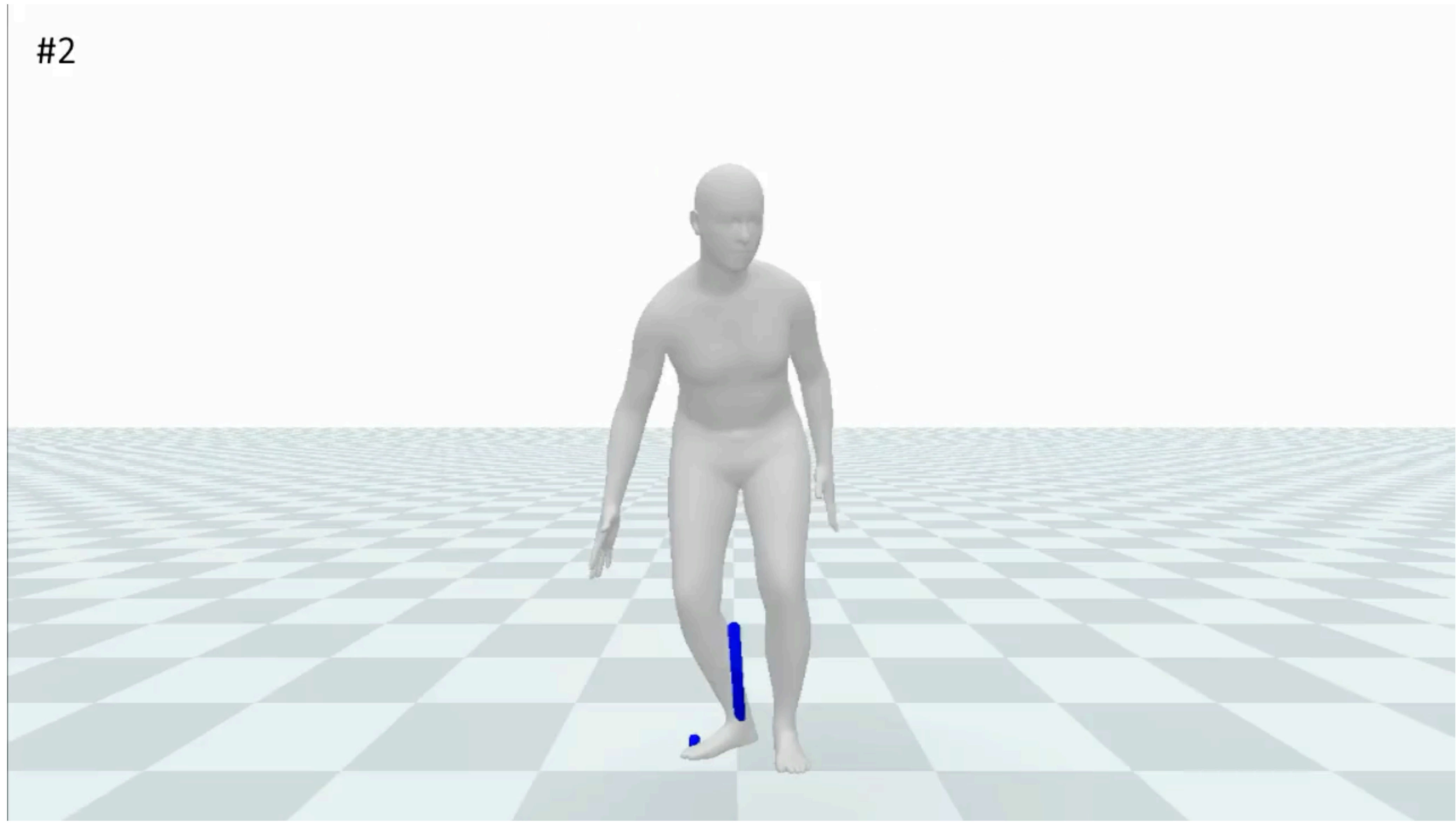
Being Thrown with Objects



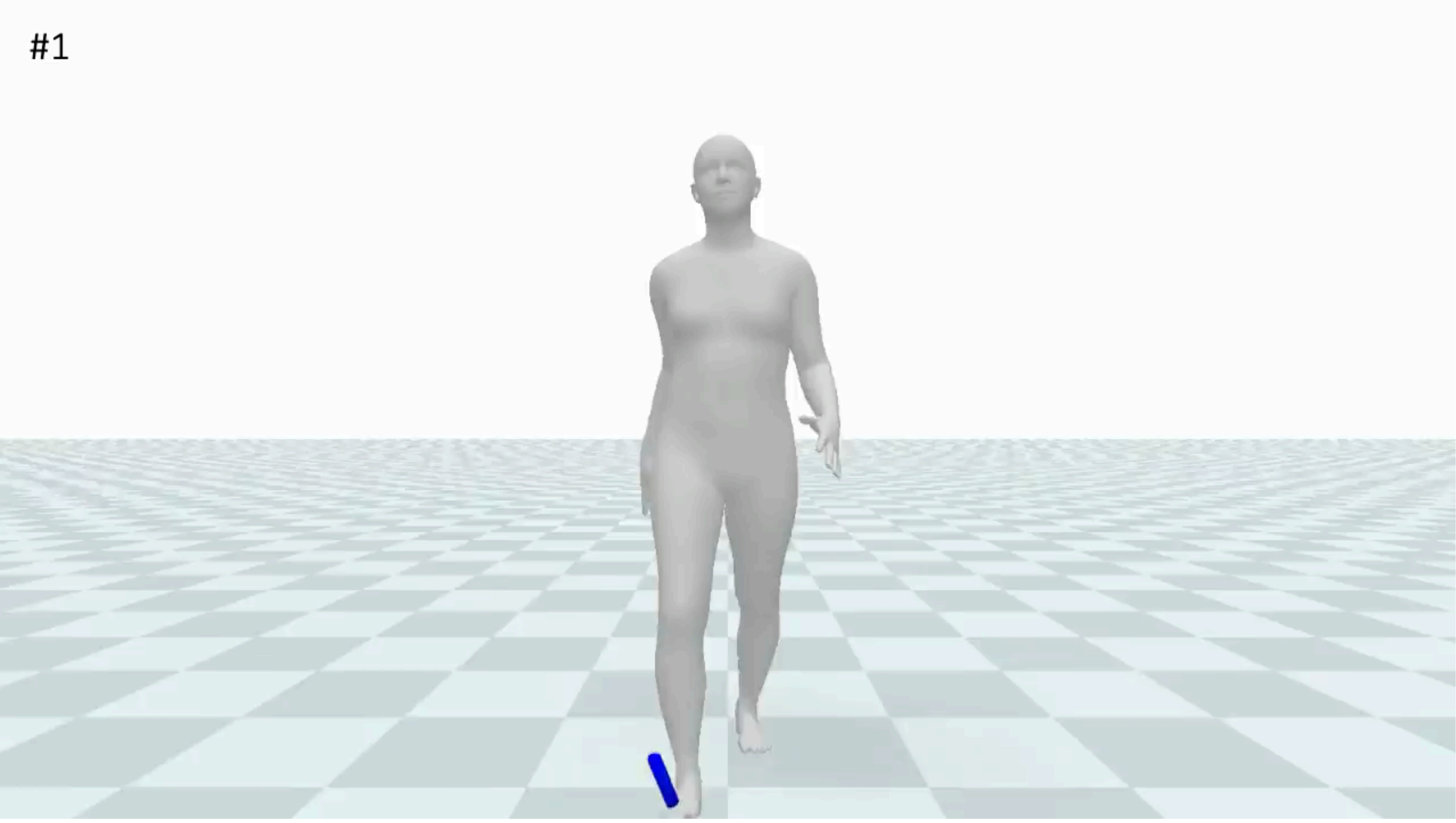


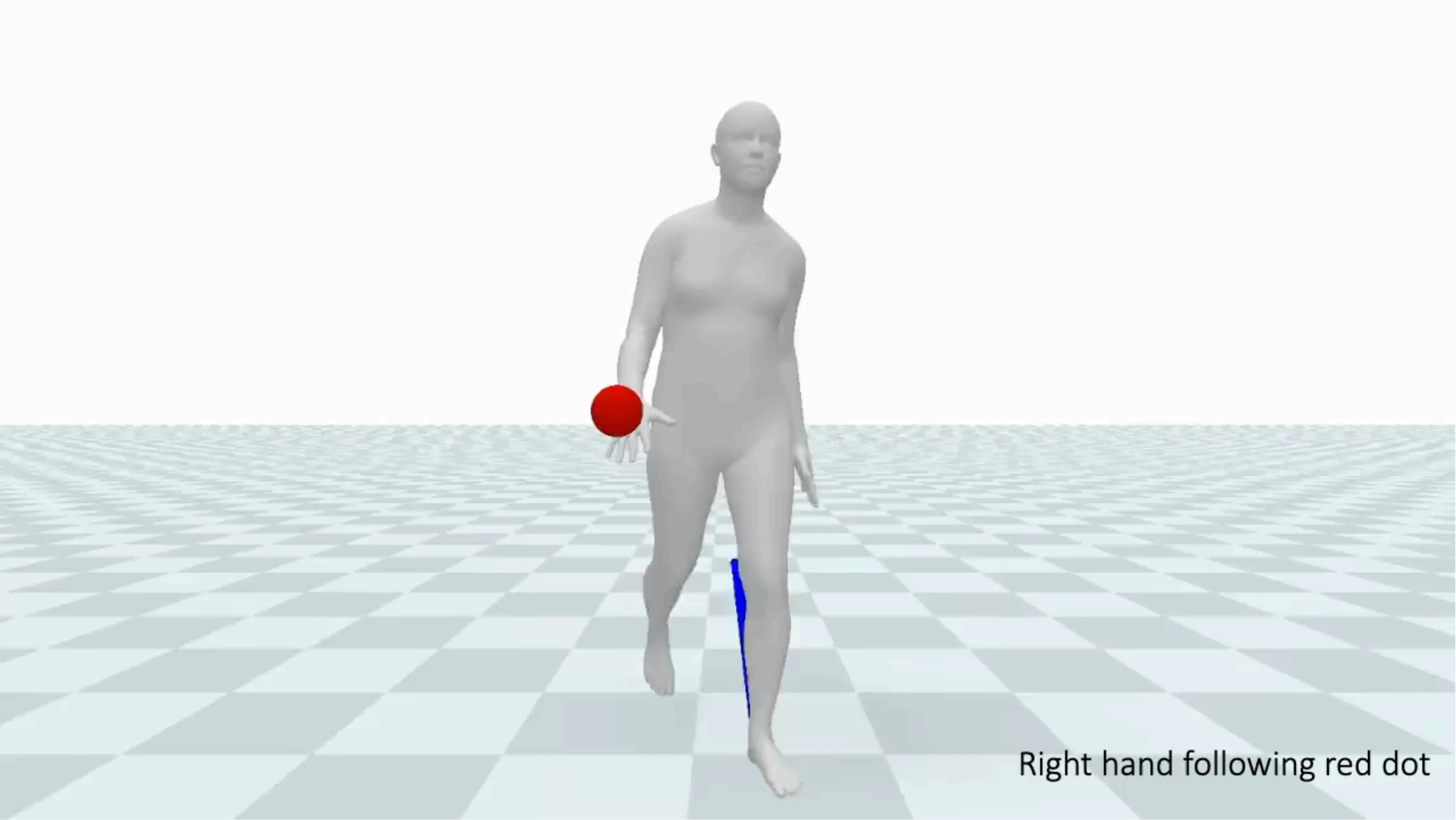
Being Thrown with Objects



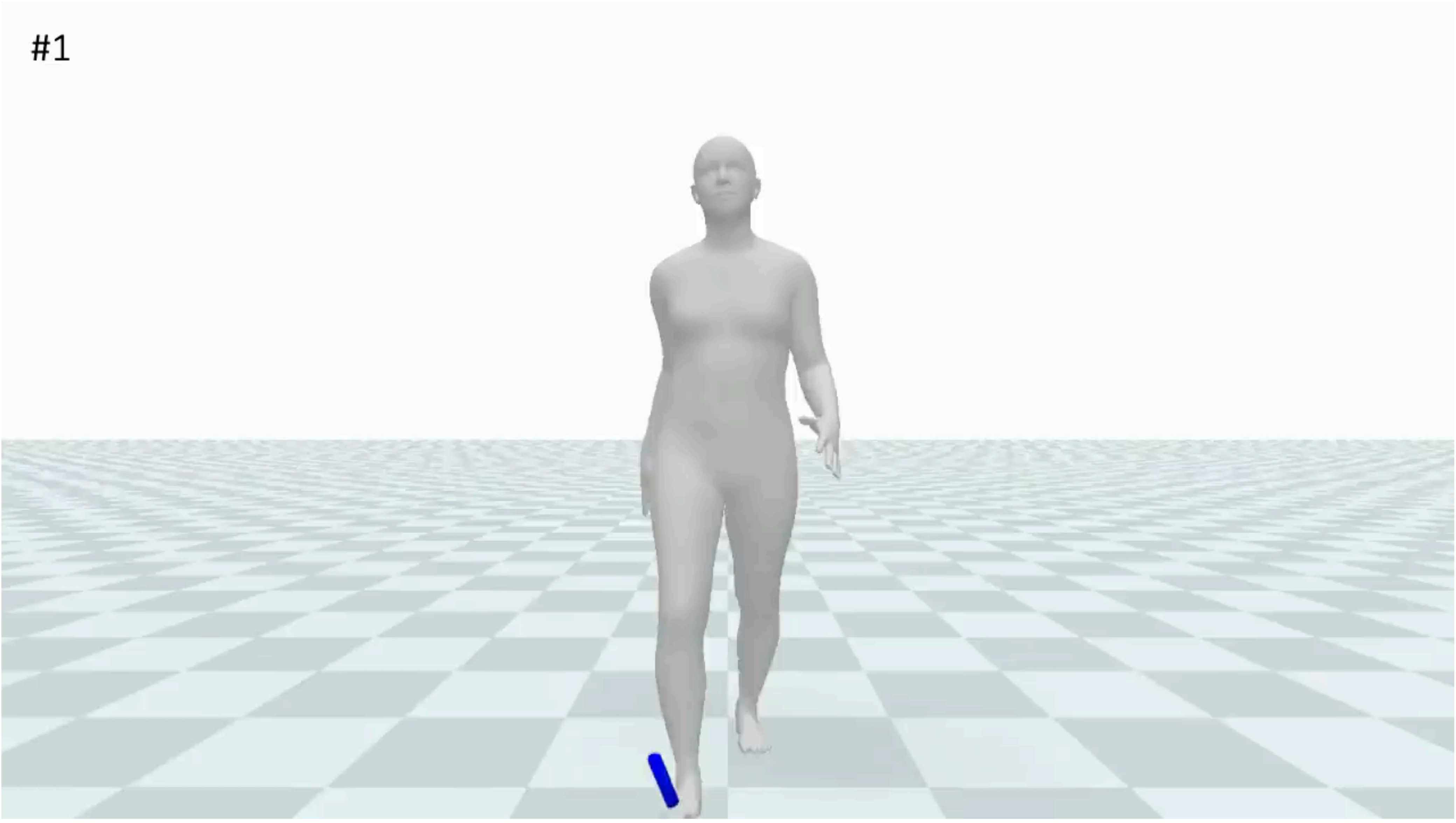


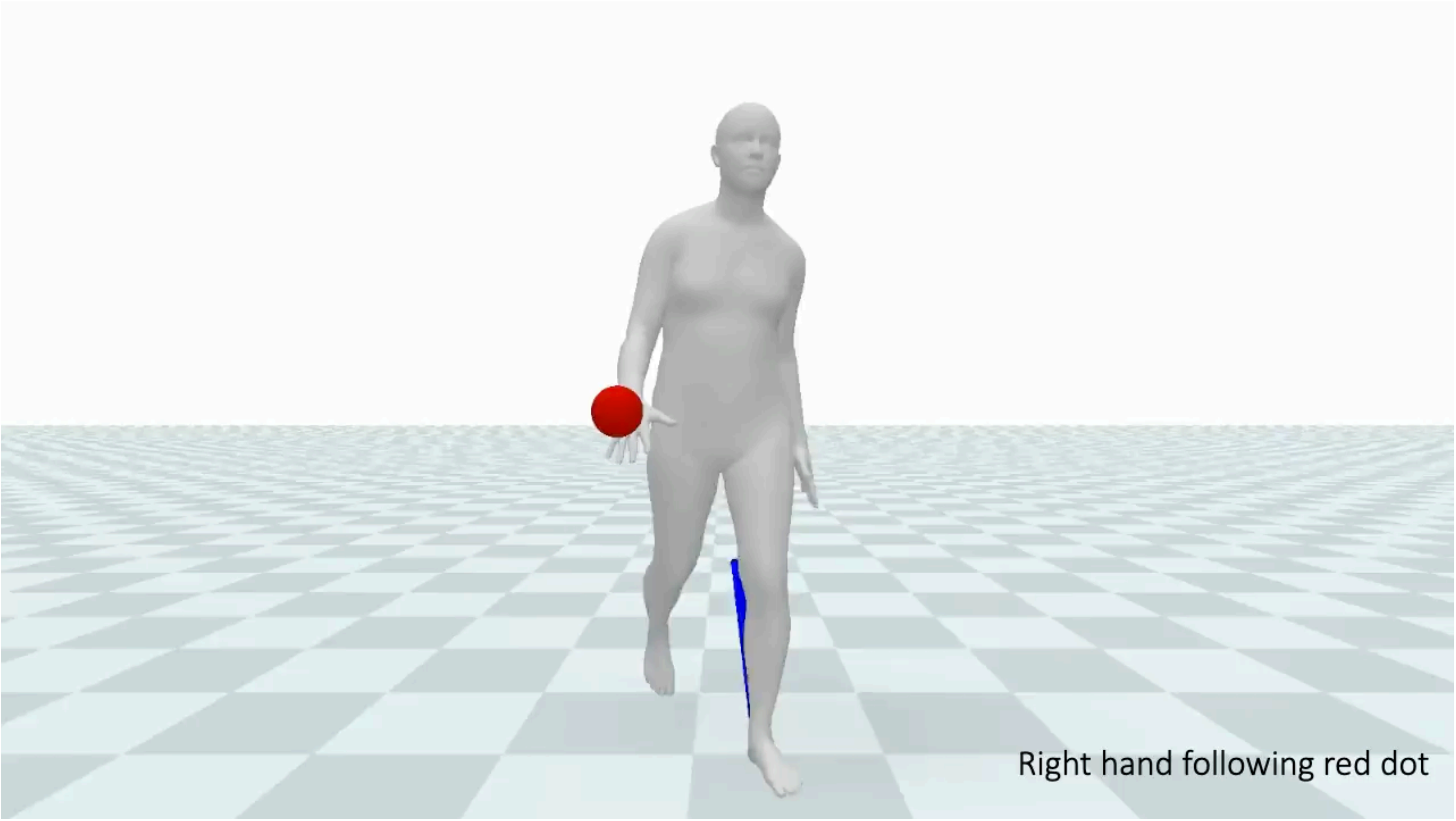
Flexible Framework Enabling Diverse Downstream Tasks



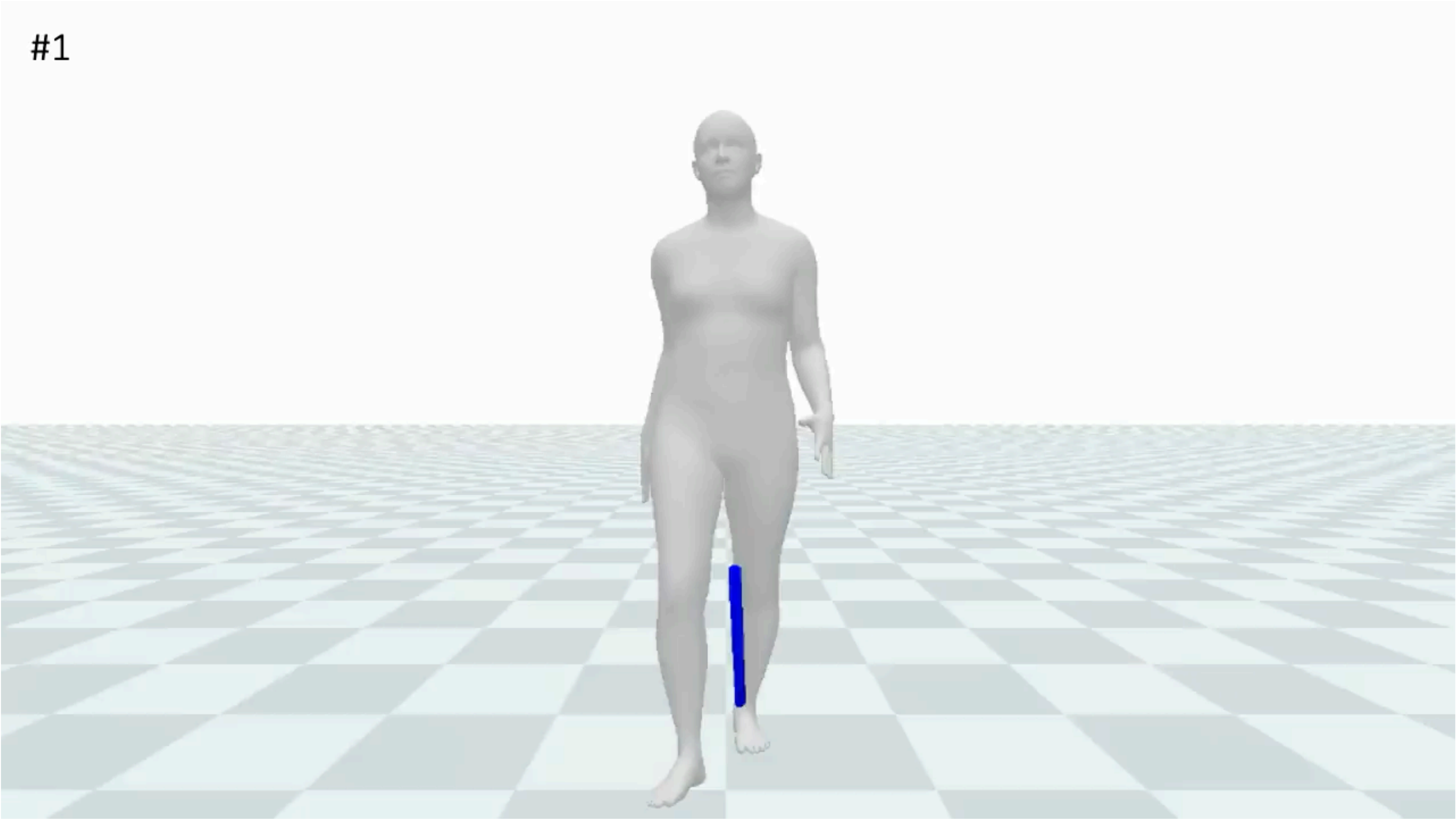


Flexible Framework Enabling Diverse Downstream Tasks

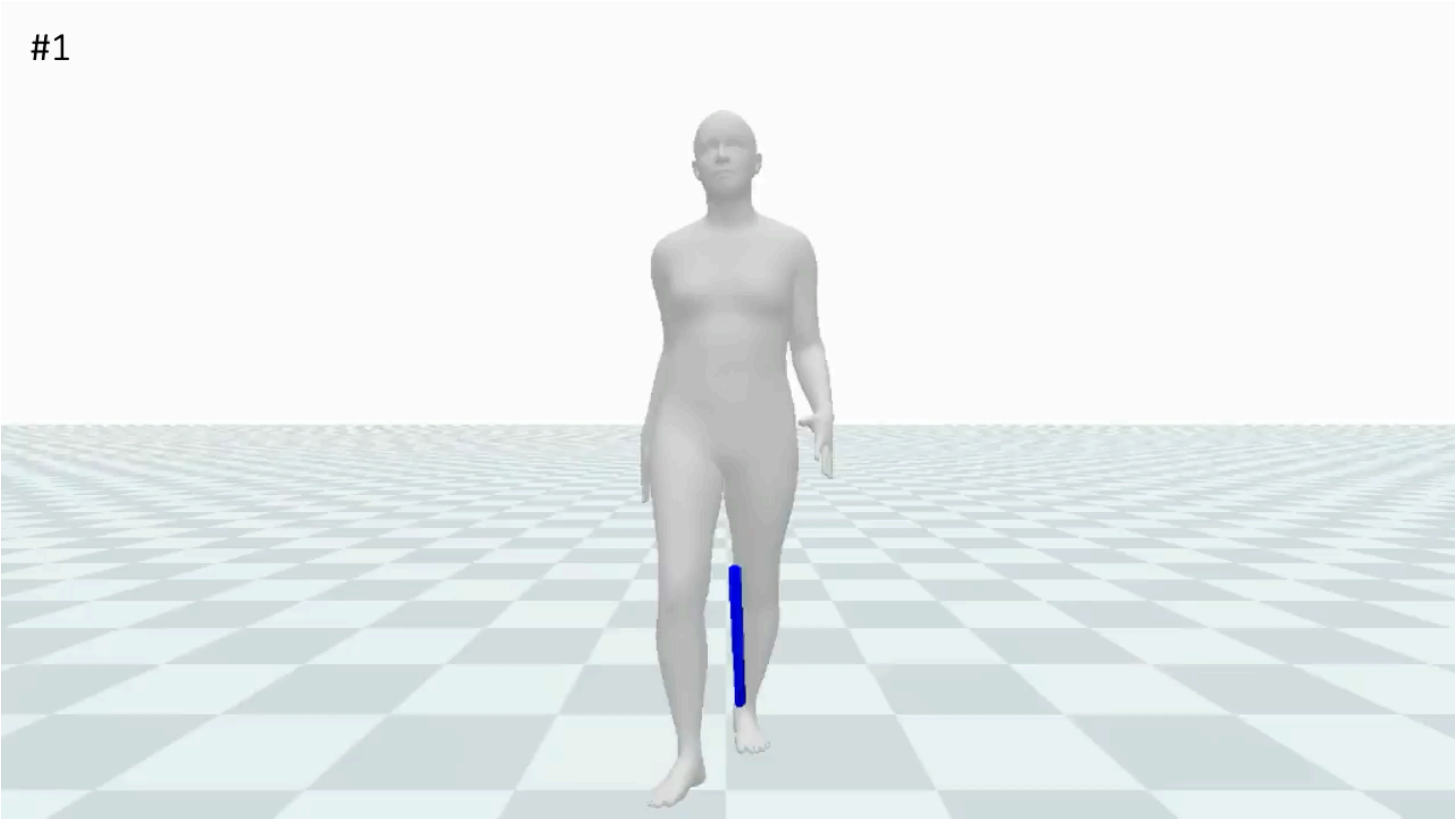




Emergent Behavior

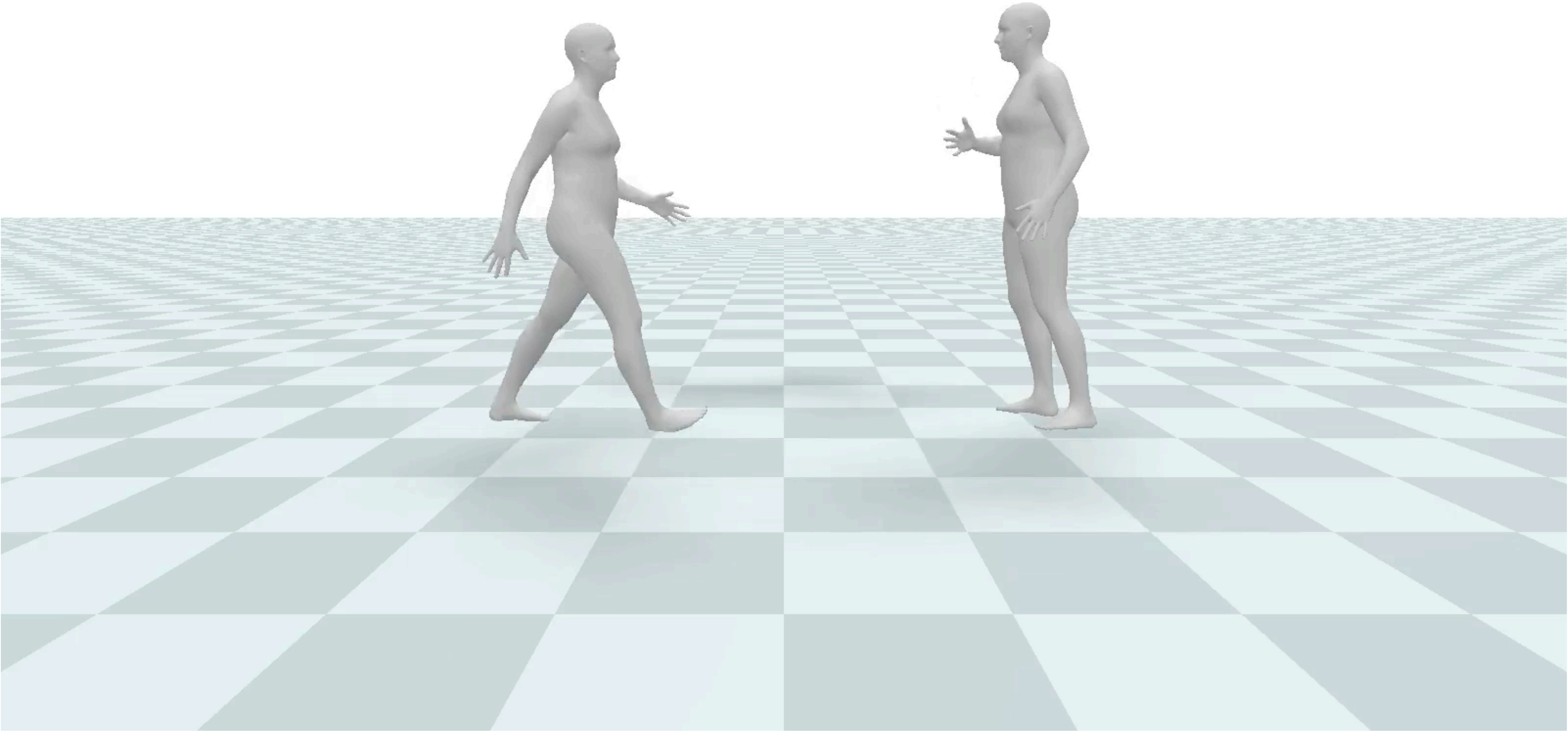


Emergent Behavior



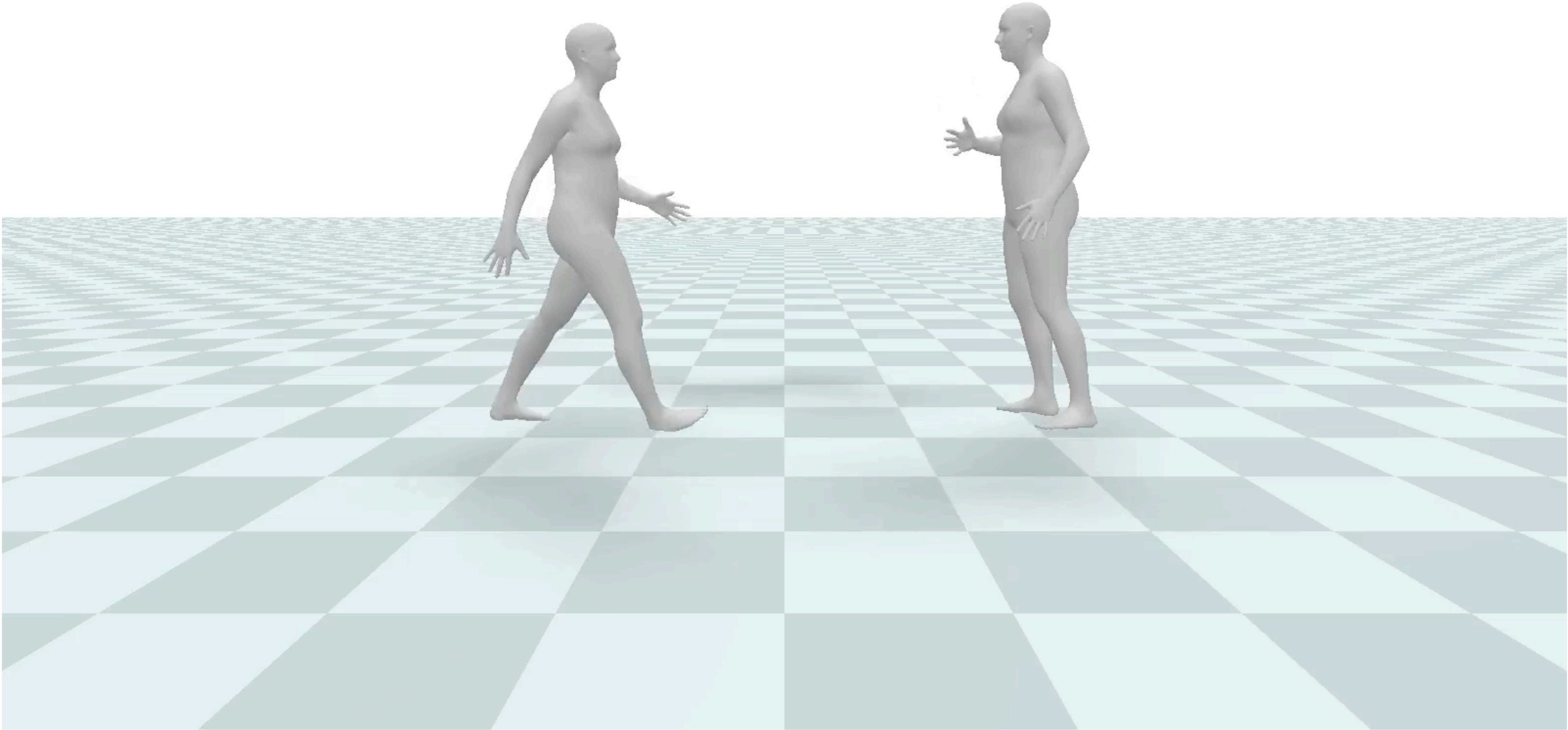
Two-character Interactions

#1



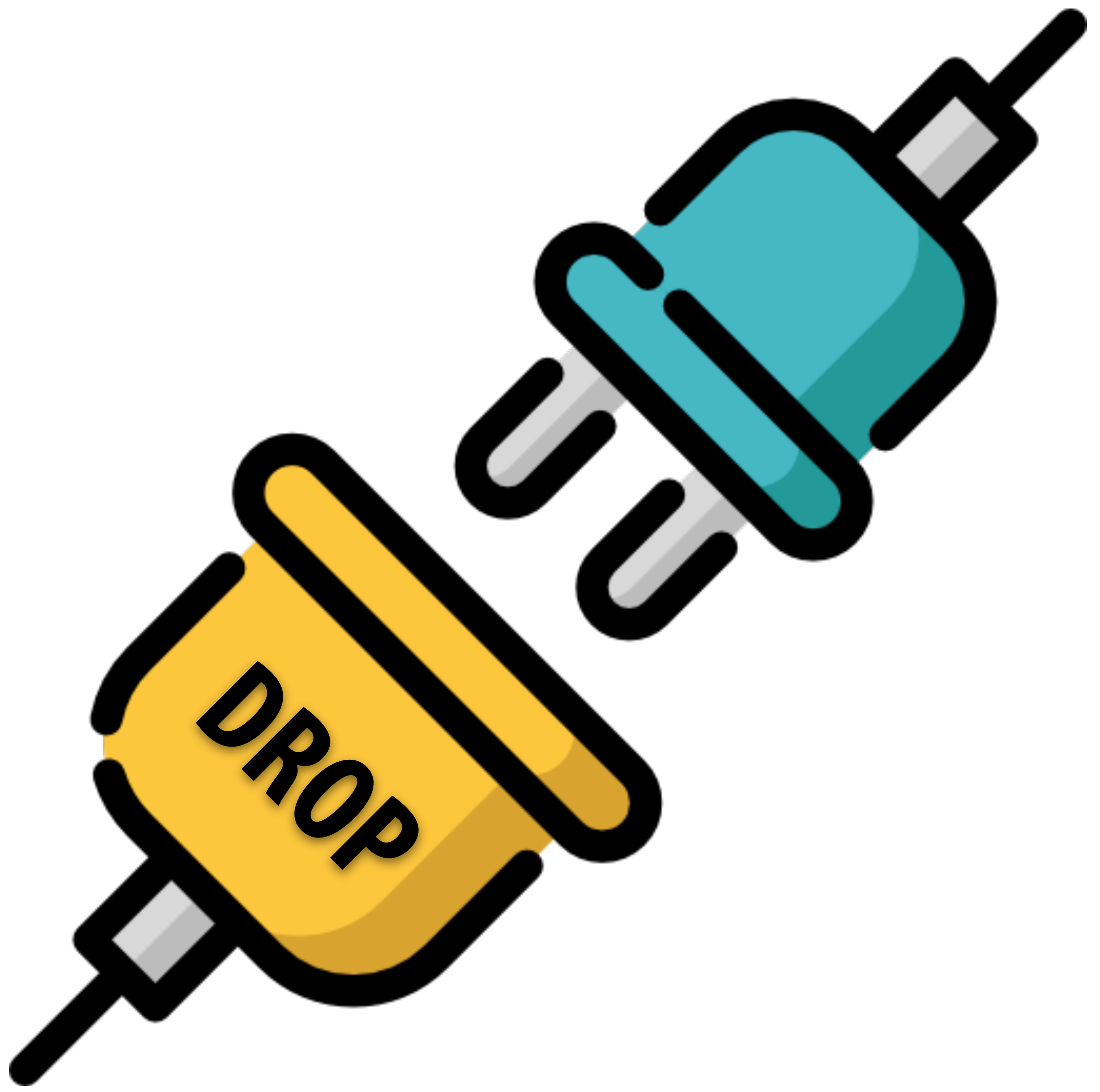
Two-character Interactions

#1



Recap

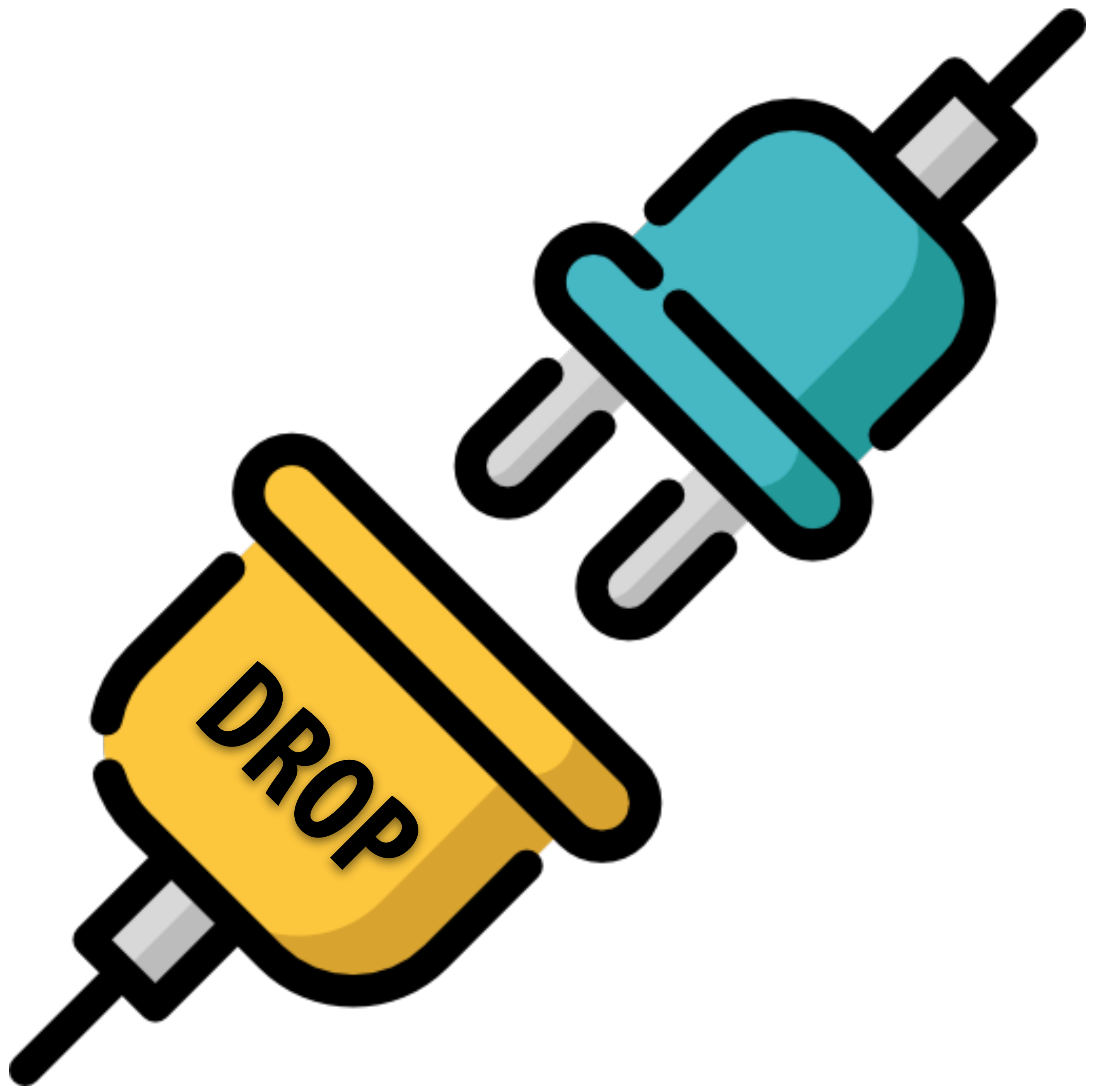
Pre-trained
Generative Model



Recap

Minimal Sim designed to fit Generative Models

Pre-trained
Generative Model

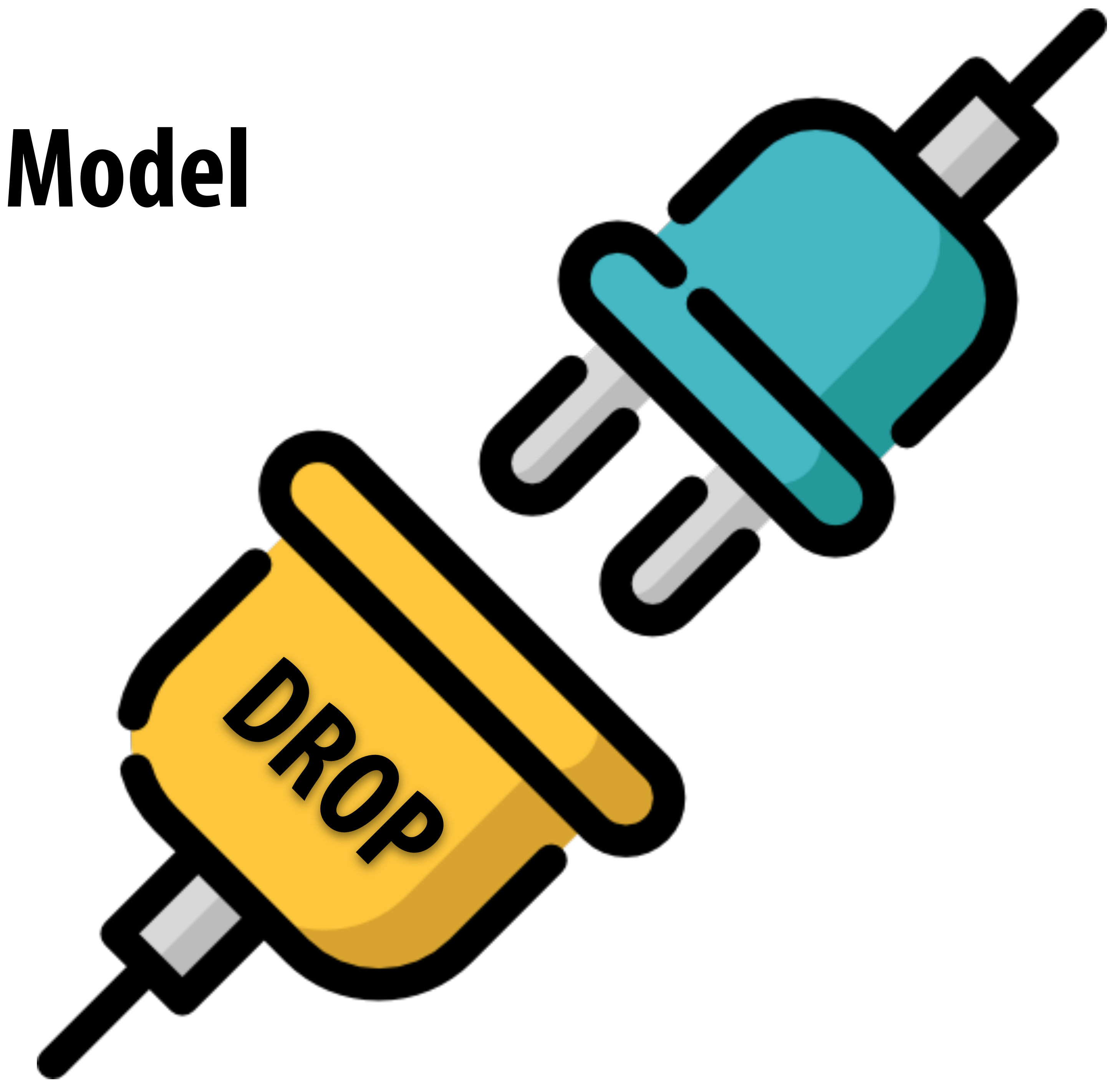


Recap

Minimal Sim designed to fit Generative Models

Plug in any pre-trained autoregressive Generative Model

Pre-trained
Generative Model



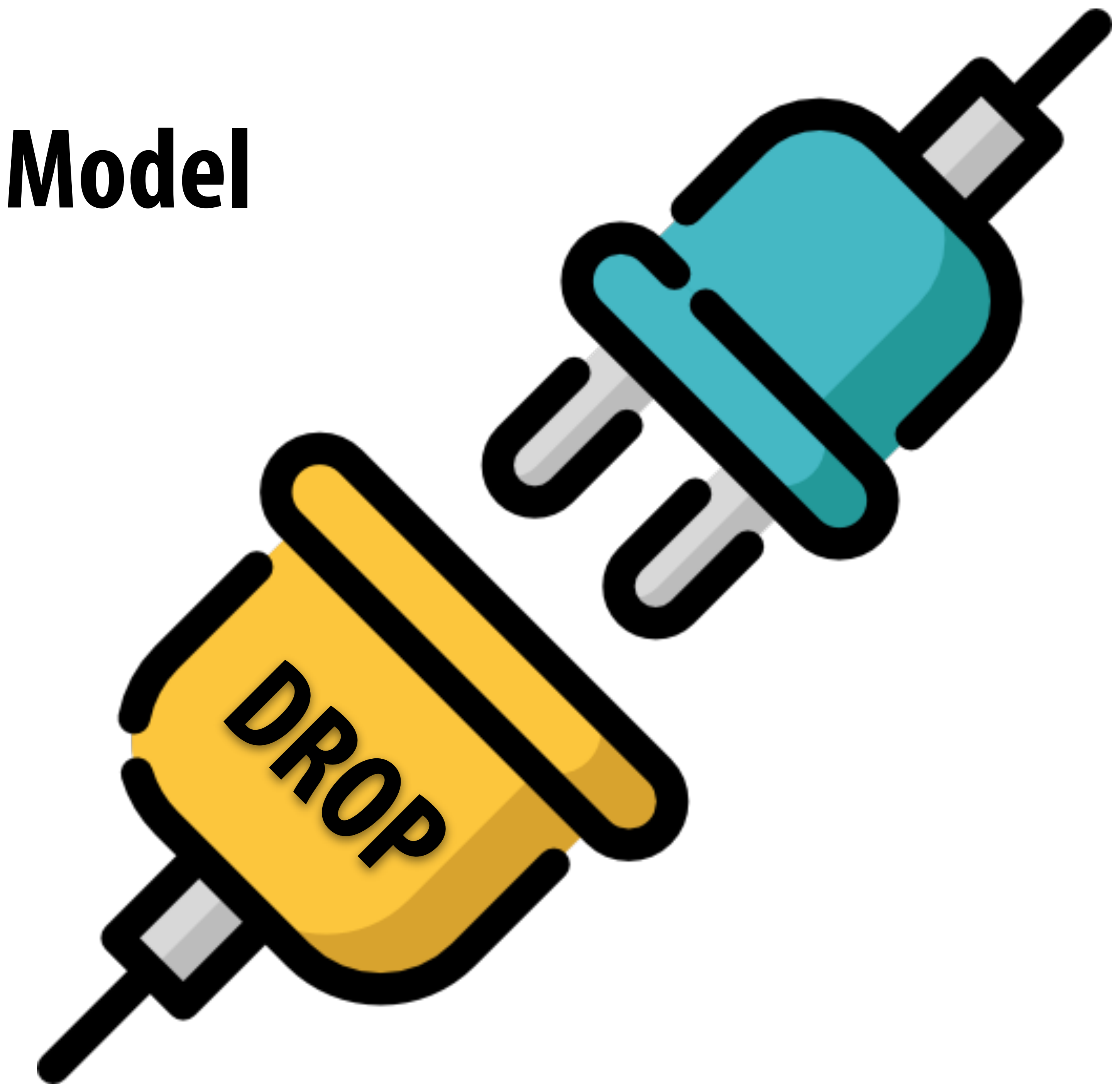
Recap

Minimal Sim designed to fit Generative Models

Plug in any pre-trained autoregressive Generative Model

Diverse physical motions at scale

Pre-trained
Generative Model

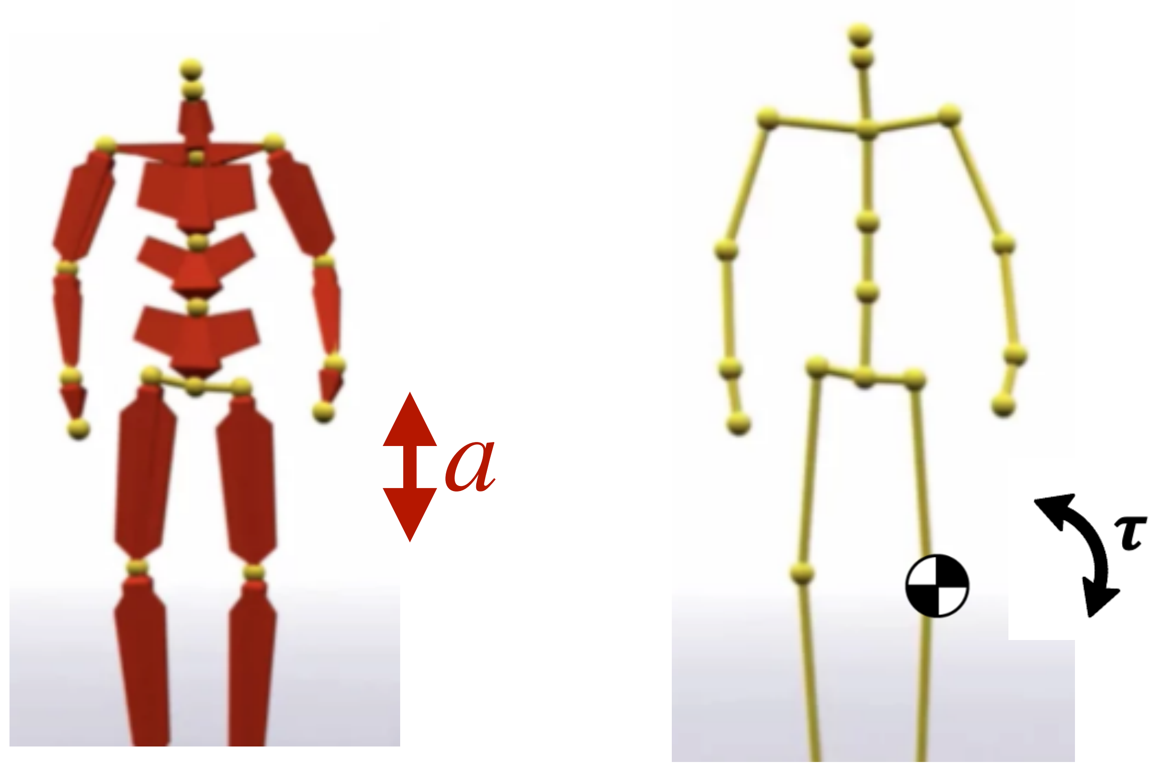


Scalable Physical Human Data Capture

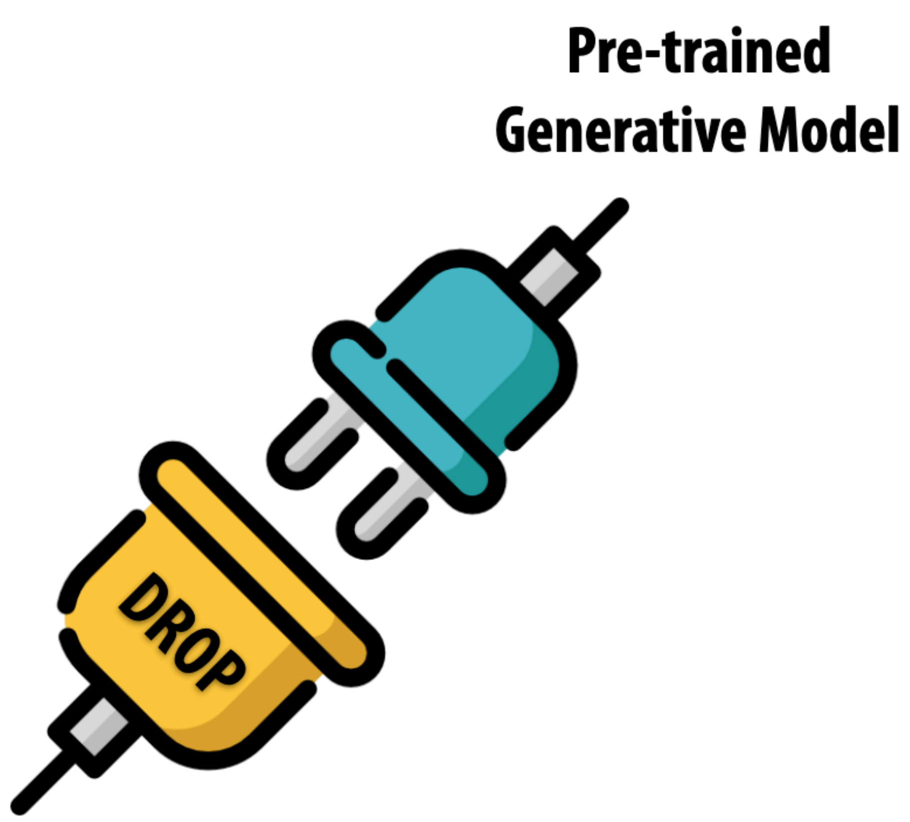
— How motion & physics prior can help scale up human data

[Jiang et al] SIGGRAPH Asia'22,
[Lee, Jiang, Liu] SIGGRAPH'23

So far…

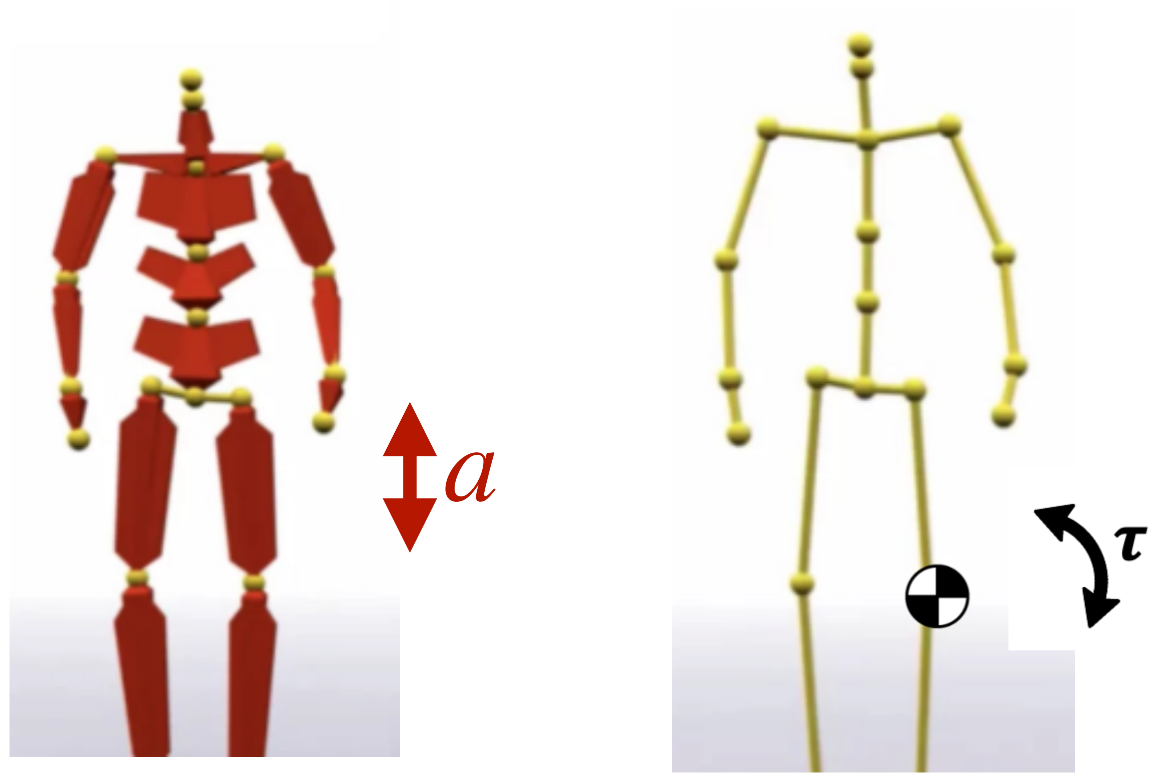


Data-driven Human Sim



Sim-augmented GenAI model

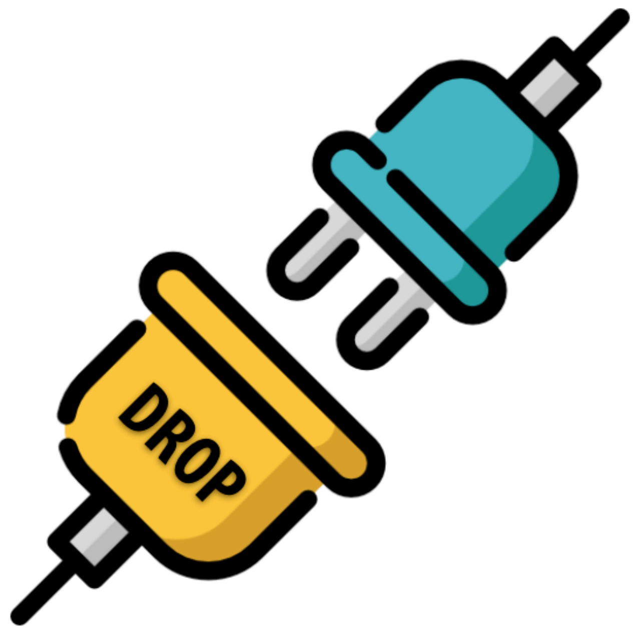
So far…



Data-driven Human Sim

Motion Data Engine

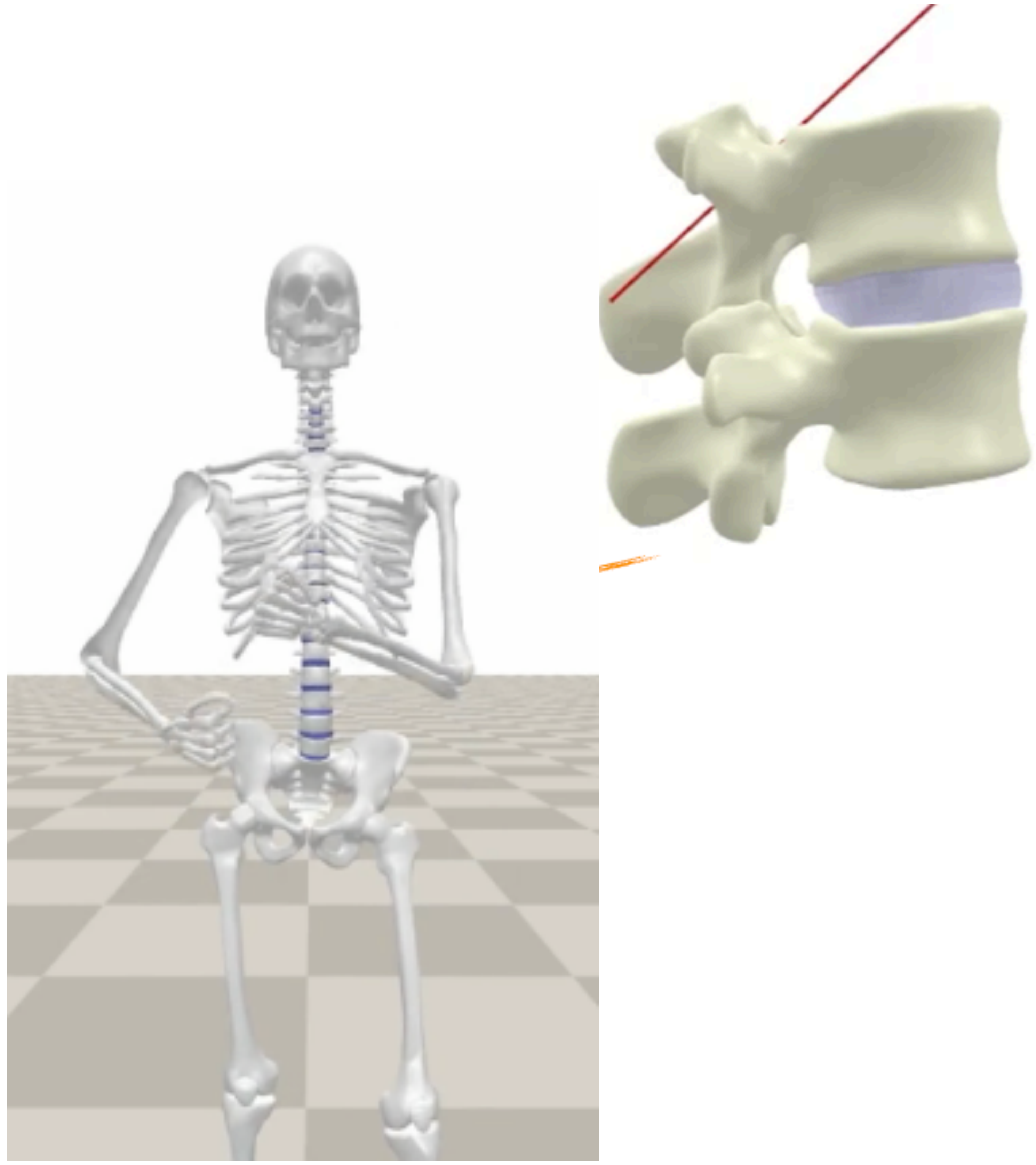
Pre-trained
Generative Model



Sim-augmented GenAI model

Motion capture can be tedious

Cannot Fully Observe All Quantities



e.g. detailed shoulder and spine movements

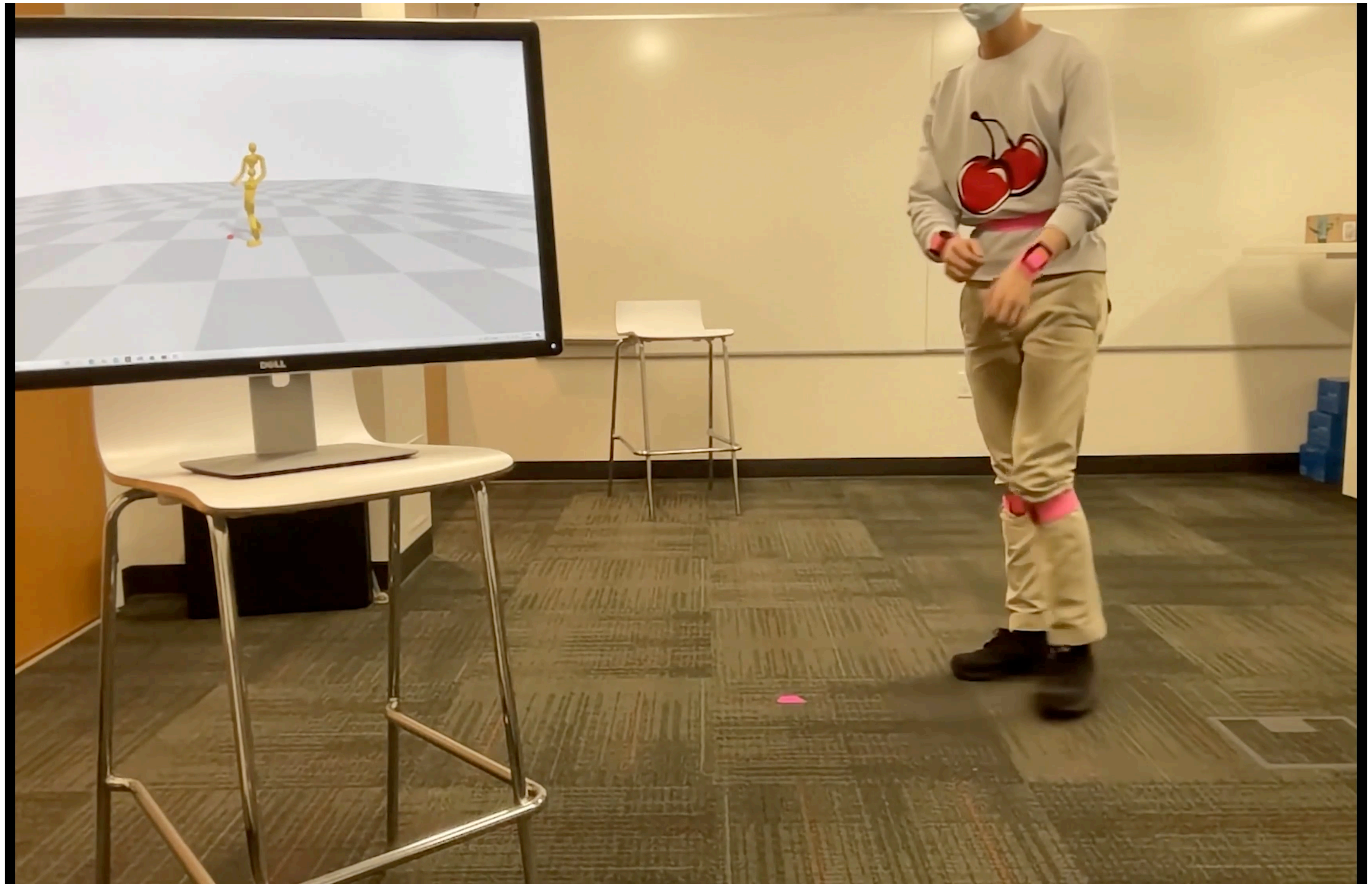
First, how might we capture human data cost-effectively, to scale up the process?

Wearable IMUs for Inexpensive Motion Capture

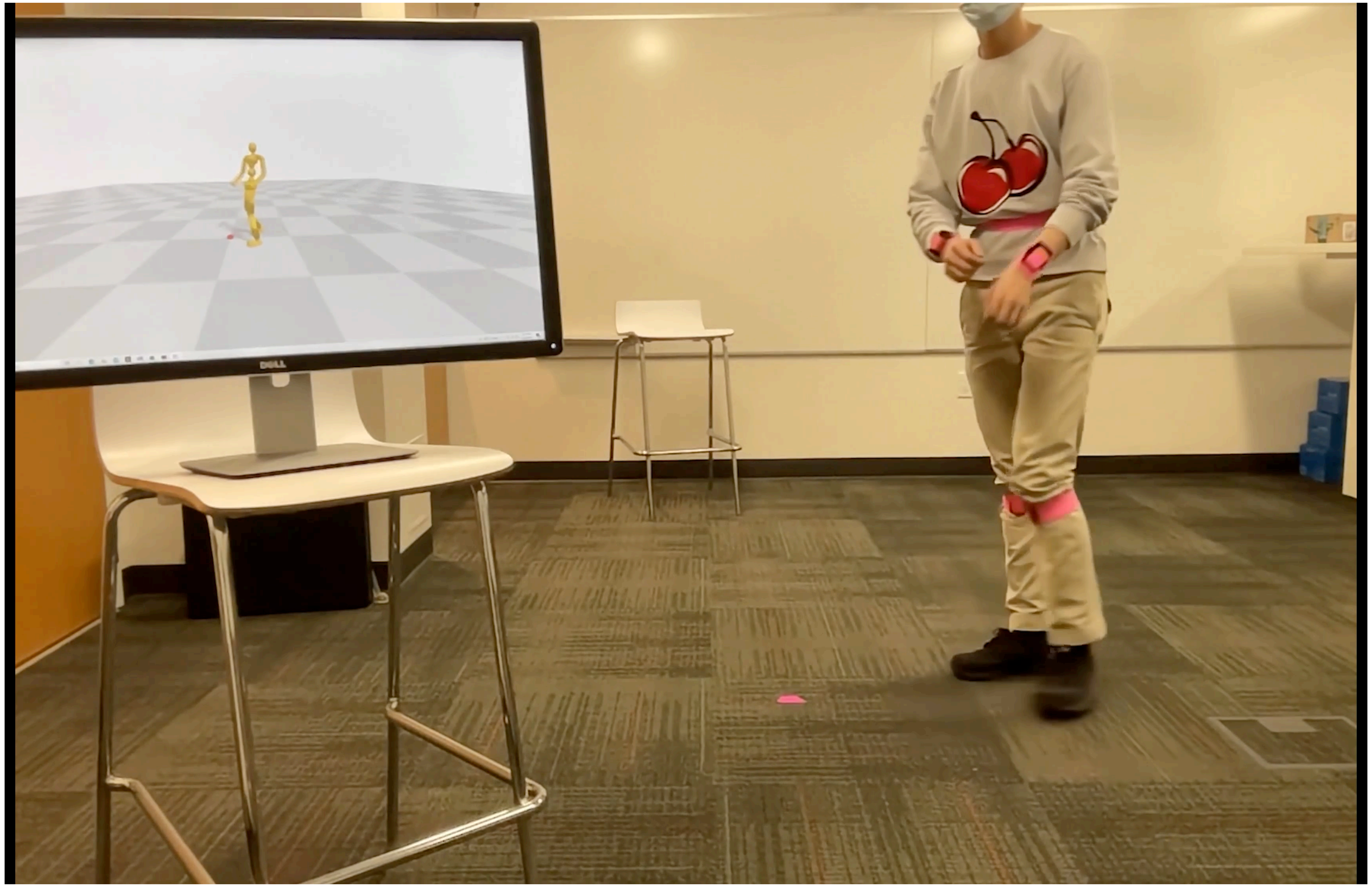
Xsens Awinda (17 IMUs) https://www.xsens.com/

Apple Airpods https://twitter.com/ConcreteSciFi/status/1311332262131113984

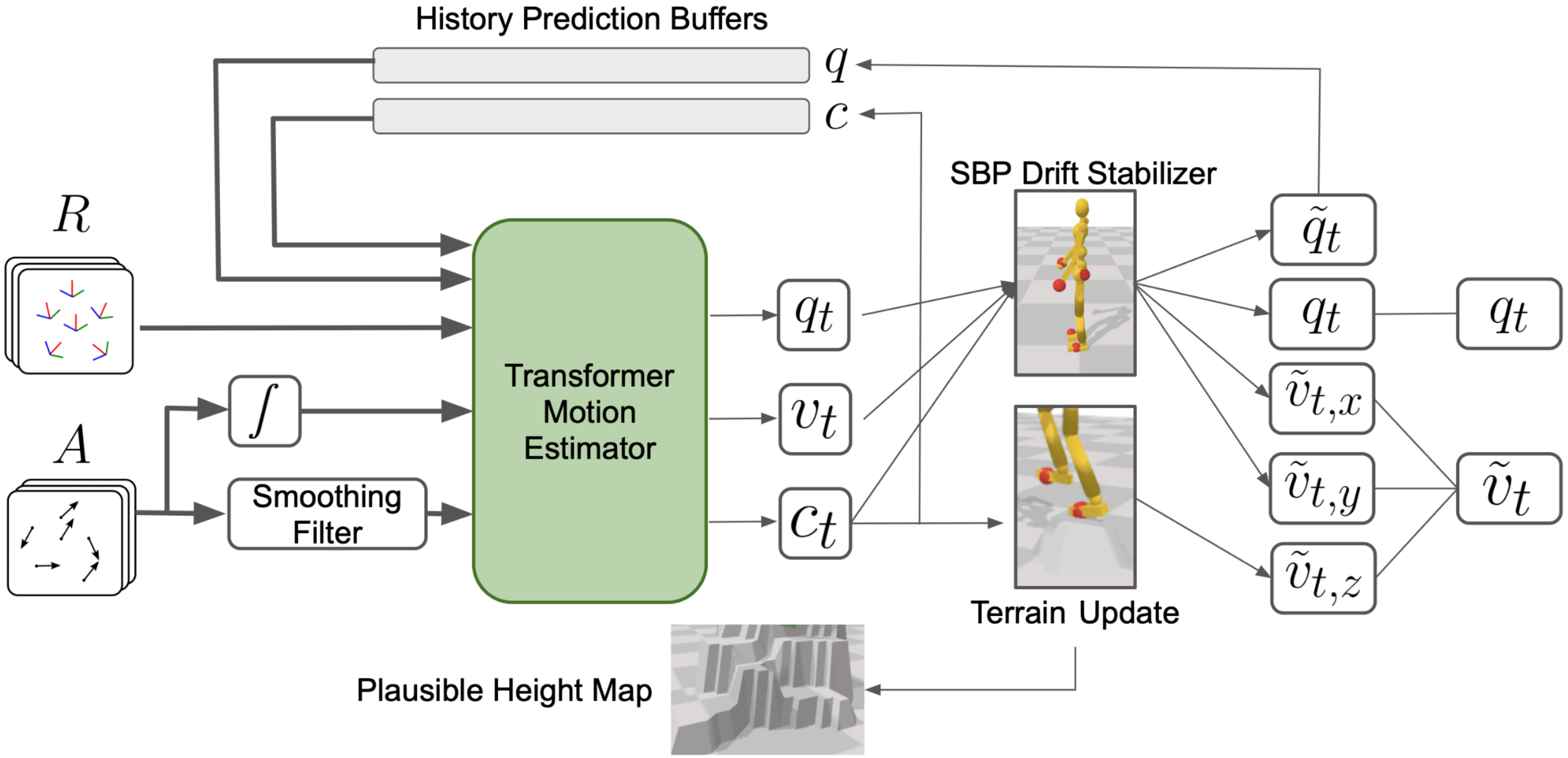
Only 6 Sparse IMUs — Minimized User Friction



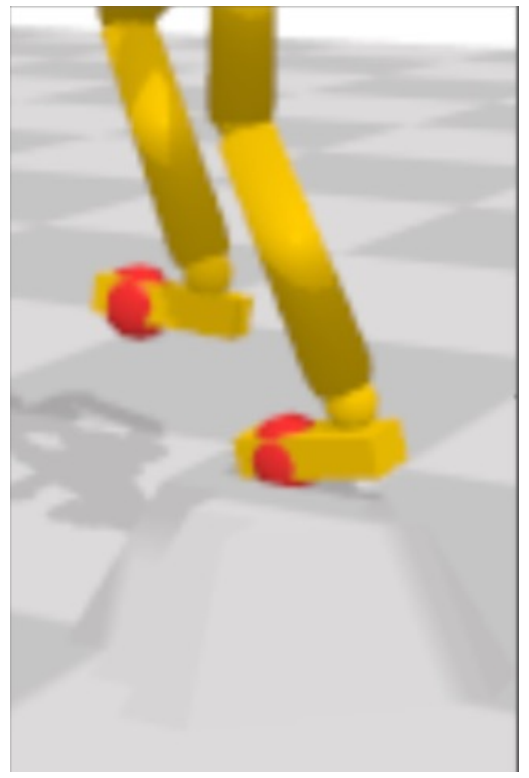
Only 6 Sparse IMUs — Minimized User Friction



Transformer-Decoder Based Model, Pretrained on Large Motion Data



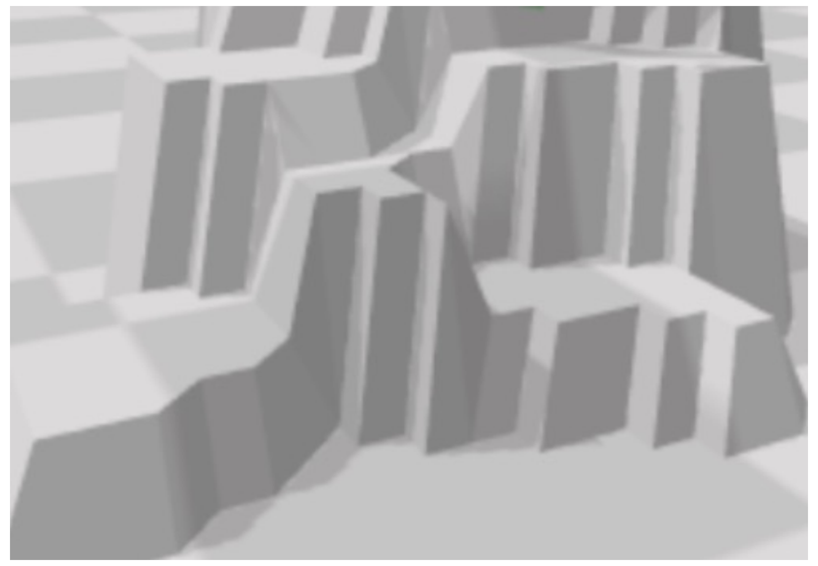
Simultaneous Terrain Map Generation



Correct slow drift

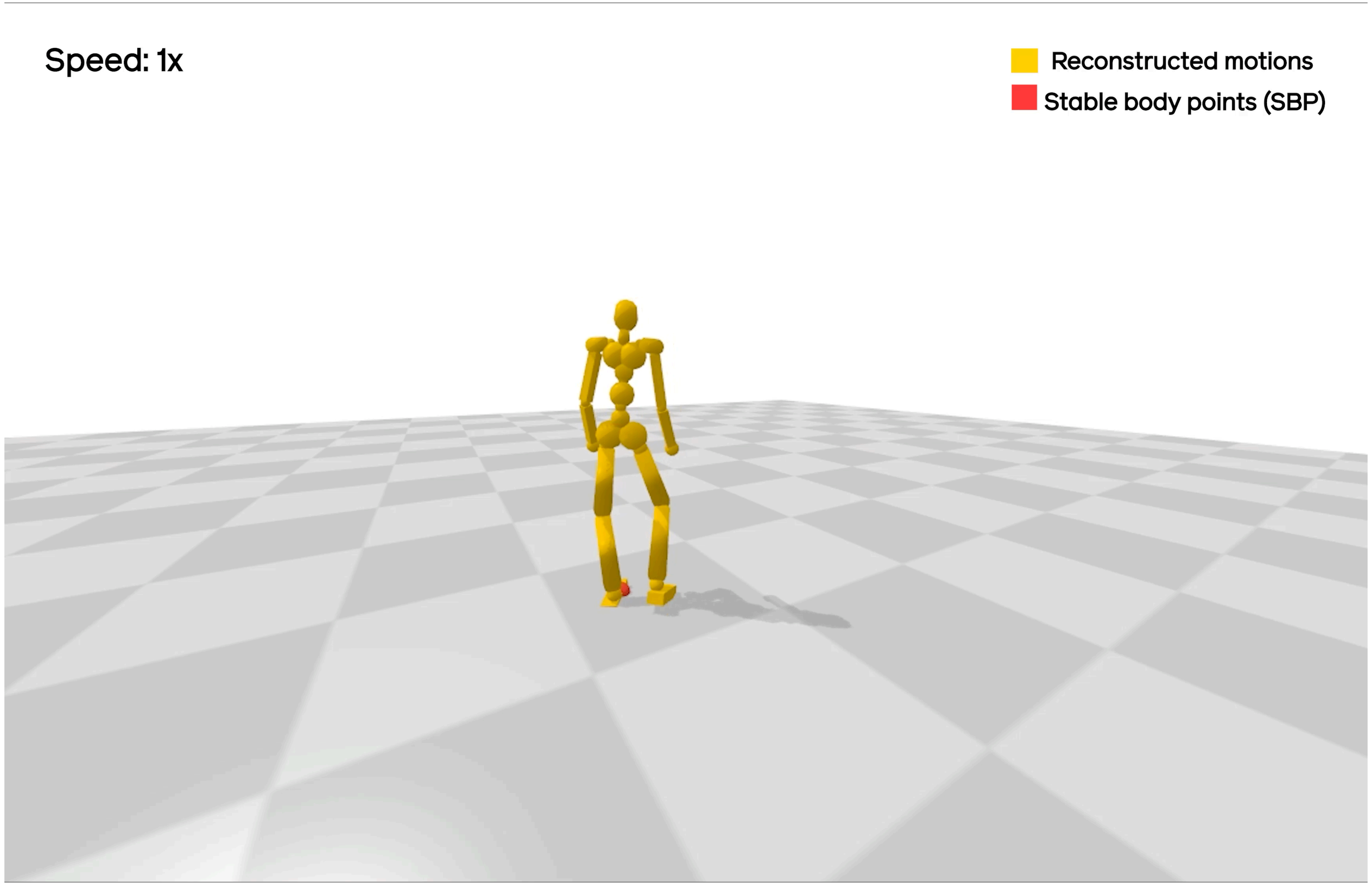
Predict plausible terrains

Predicted Motion

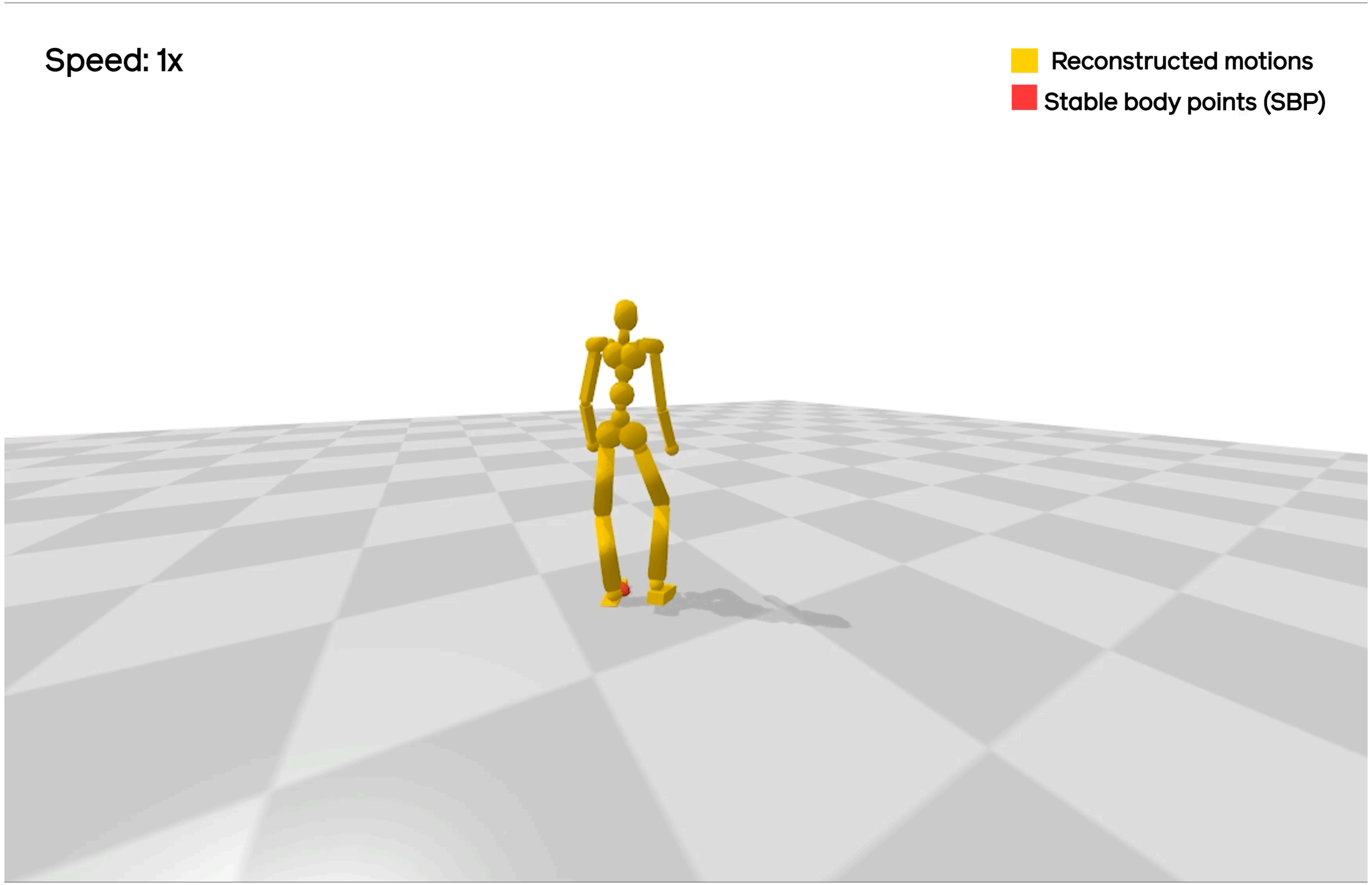


Height Map

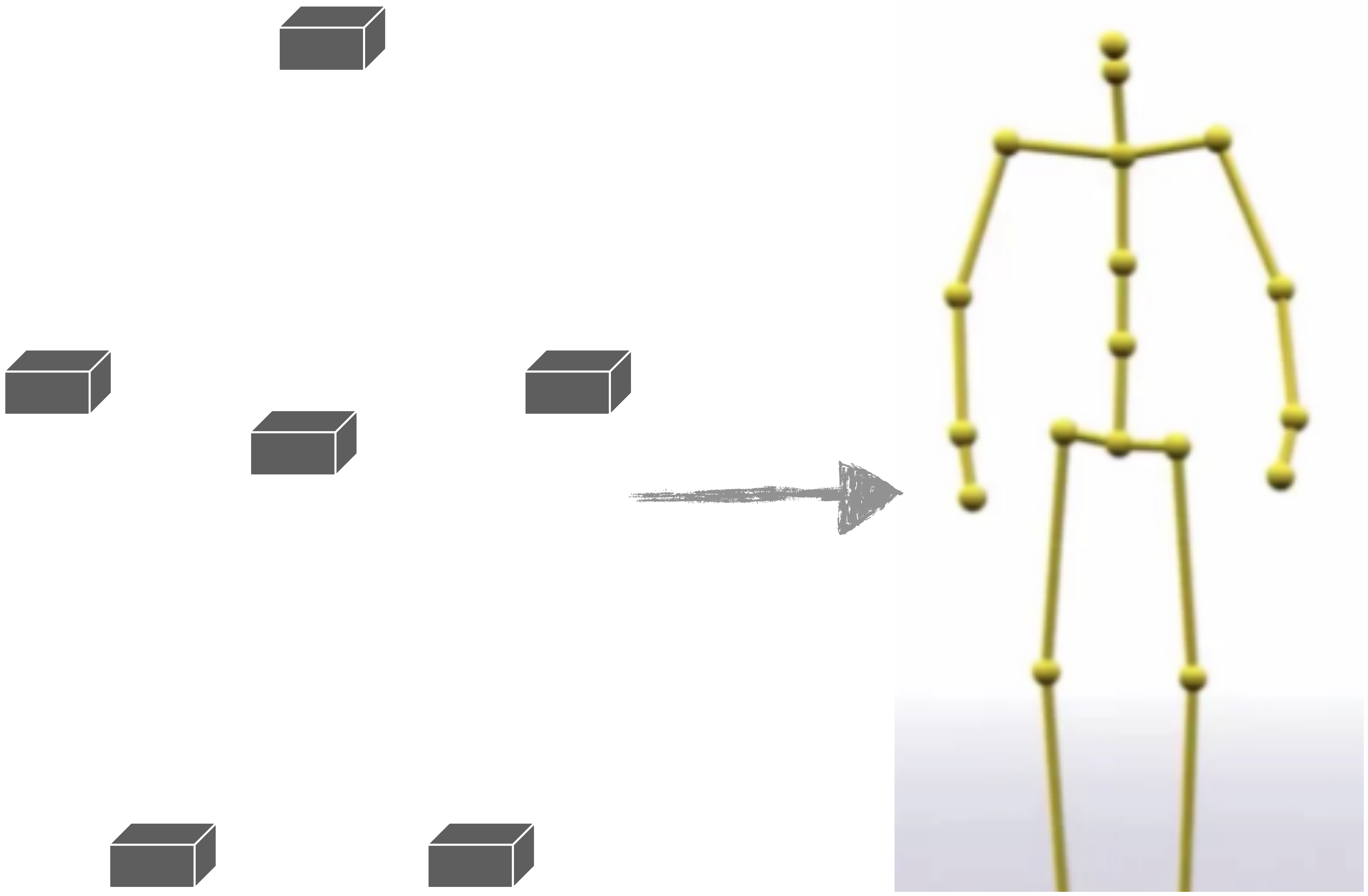
Results: Terrain Being One of the Infinitely Many Possibilities



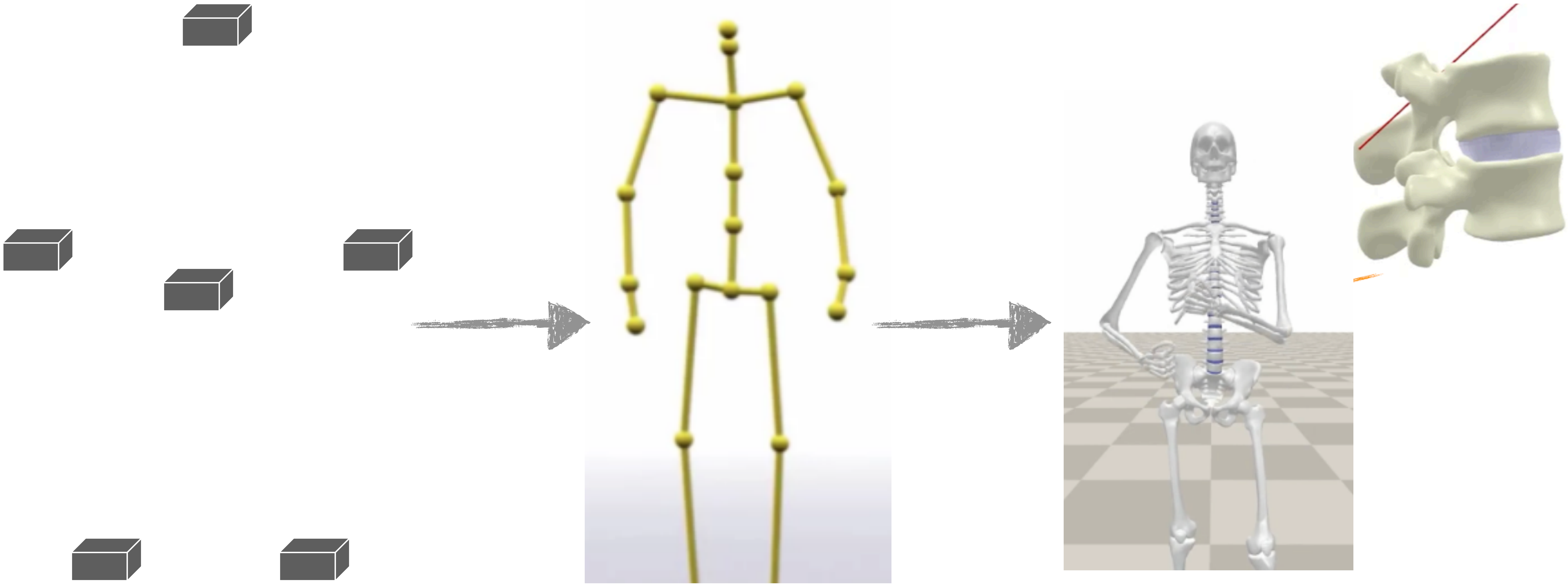
Results: Terrain Being One of the Infinitely Many Possibilities



If we can collect full-body motion data at scale,
what more could we do?



Sparse Sensors **Full-body Motion Estimates**

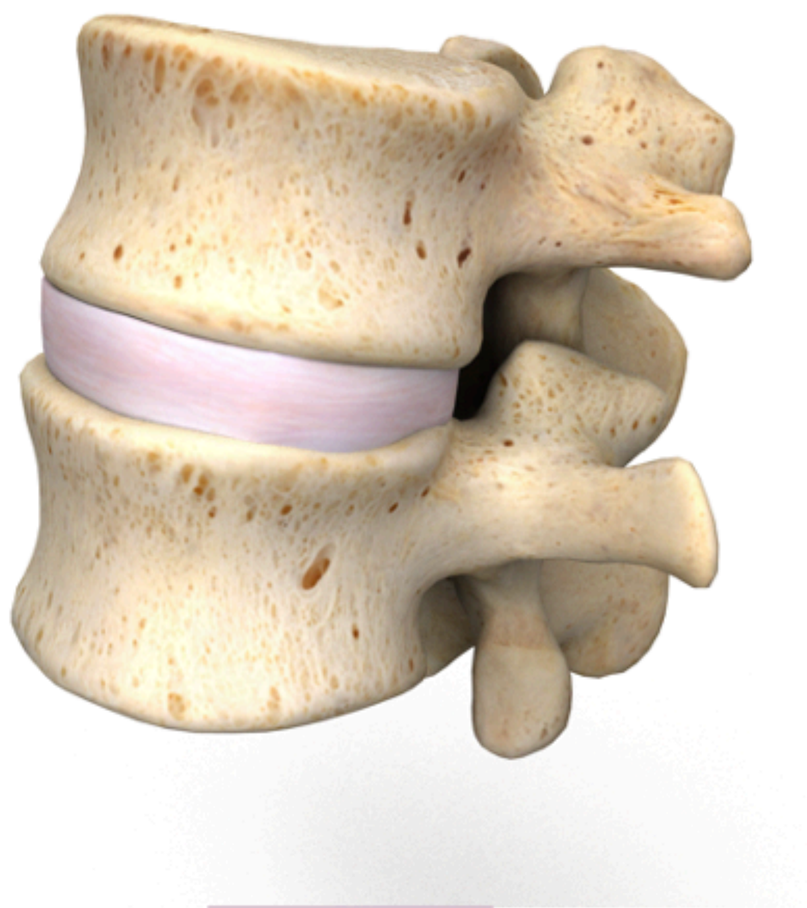


Sparse Sensors

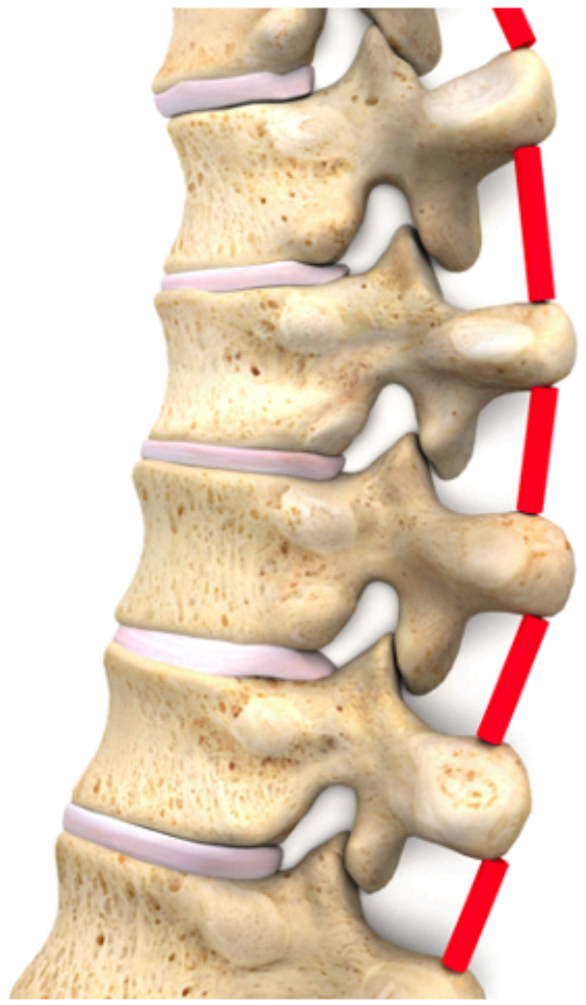
Full-body Motion Estimates

Detailed Spine Motion?

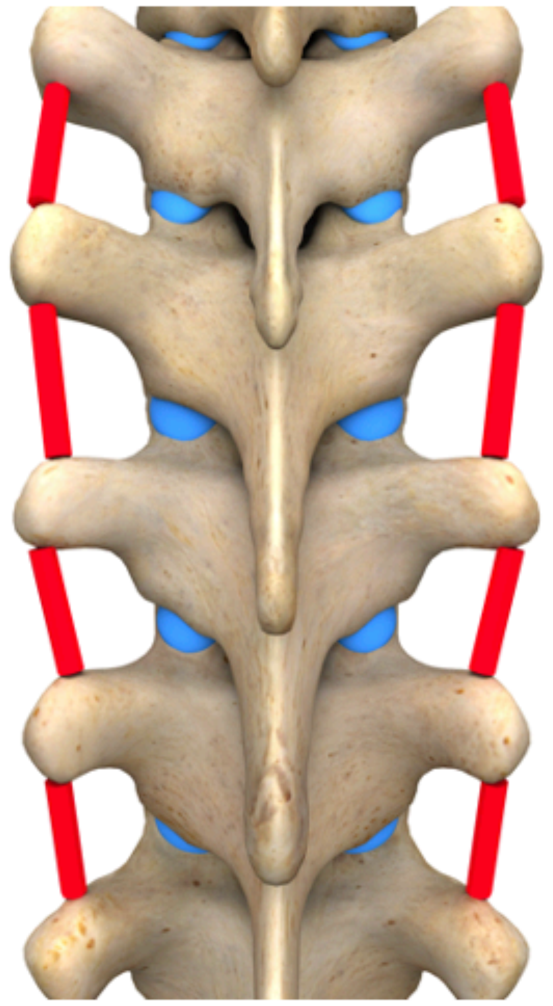
First, we built a detailed torso simulator



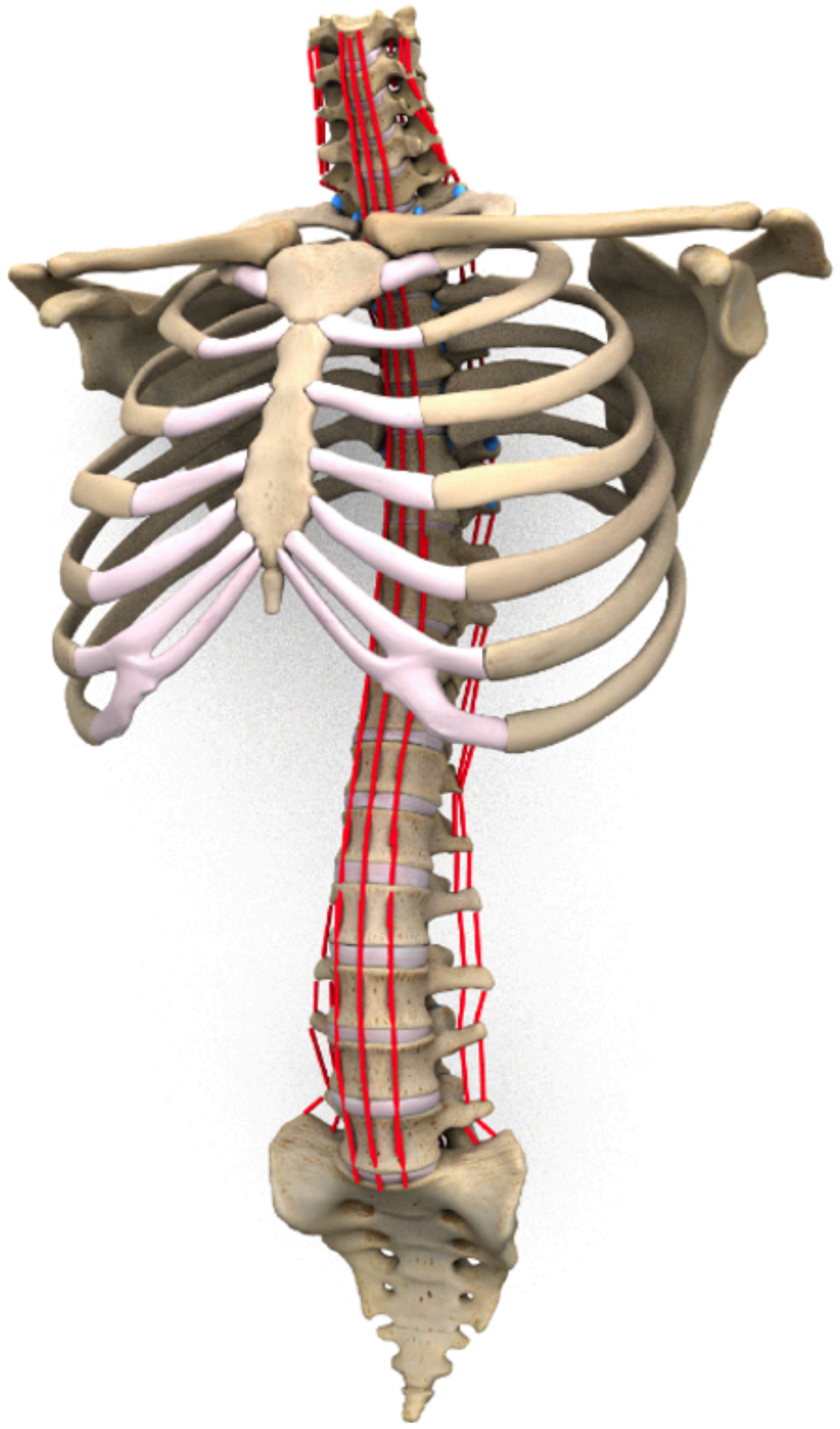
Discs



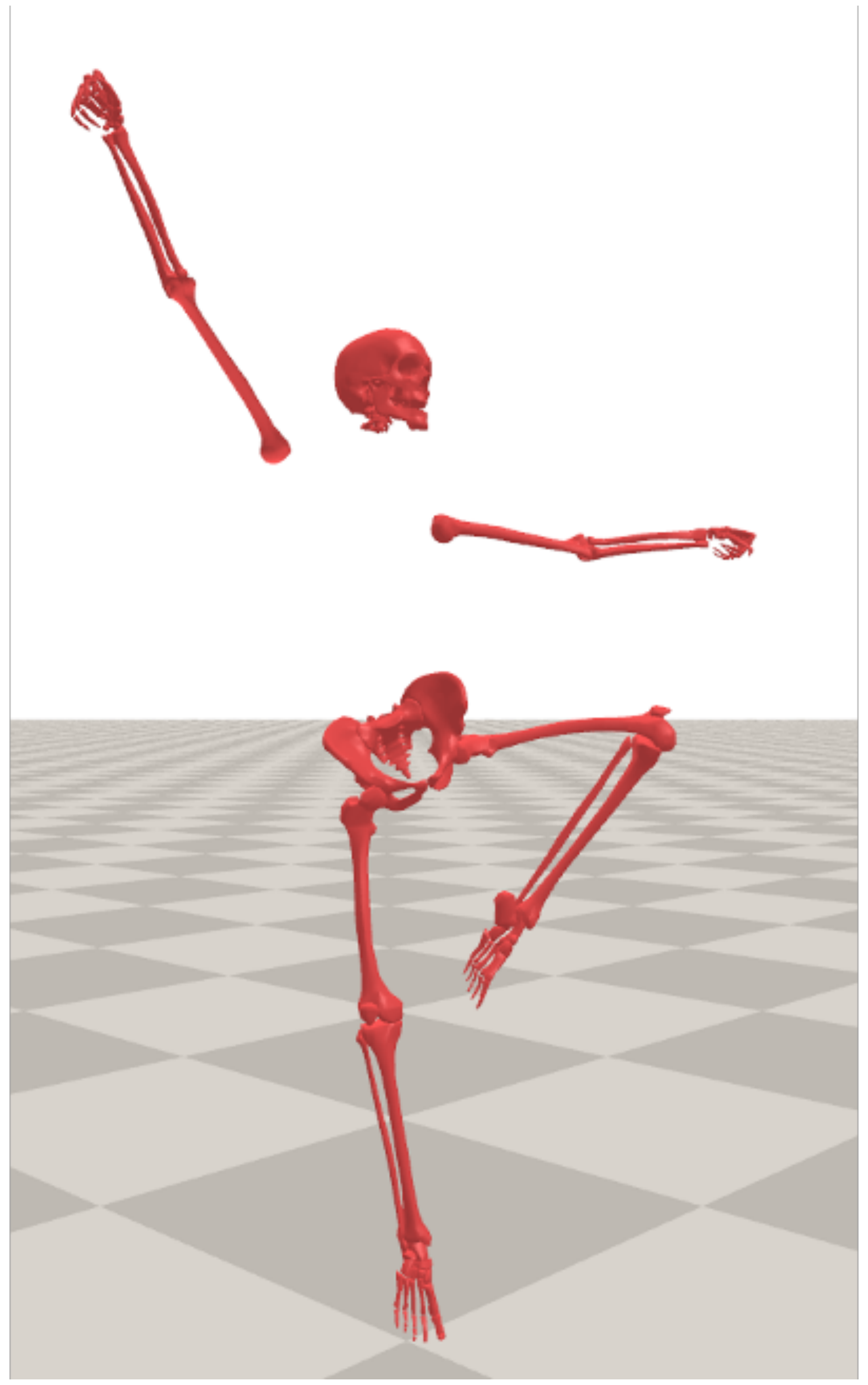
Ligaments



Facets



Use Simulator to "In-paint" Unobserved Spine Movements

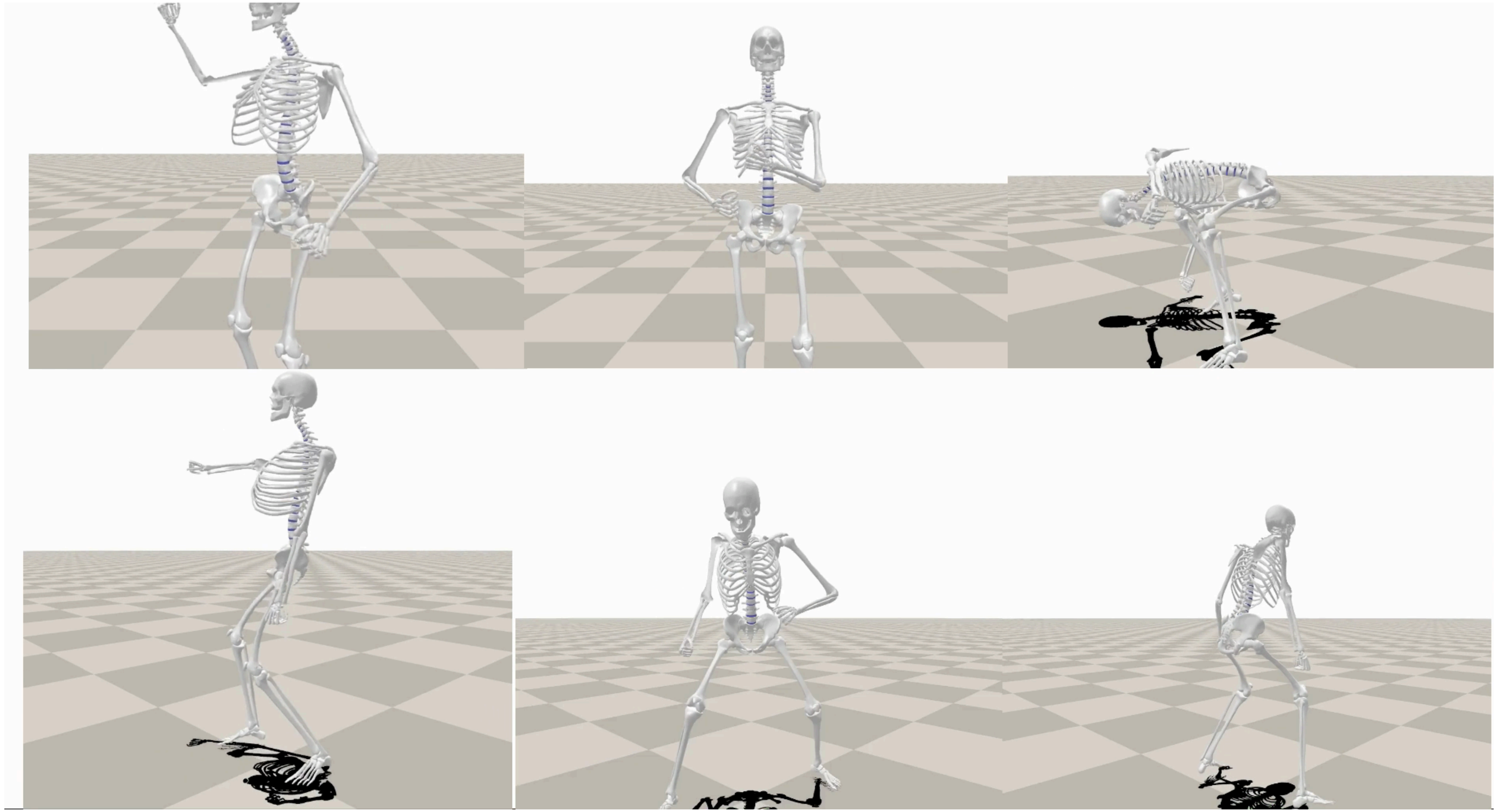


Given sparse locations of head, humerus, pelvis

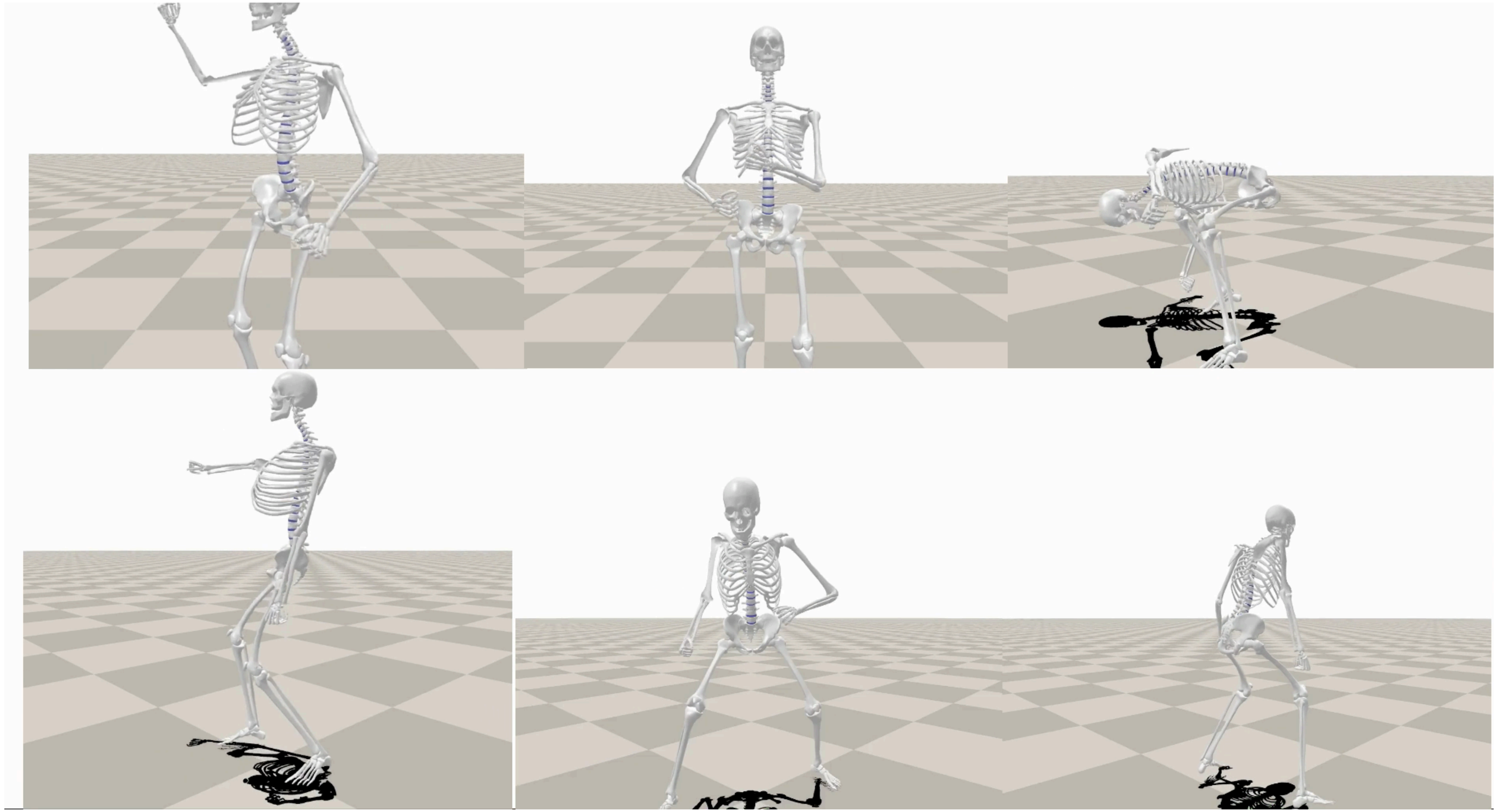
Simulate to static equilibrium

Detailed torso states

Results: In-painting a Large Dataset without Detailed Spines



Results: In-painting a Large Dataset without Detailed Spines



Recap

Cost-effective, scalable motion capture from IMUs and Smart Glasses

Augmenting coarse motion data with fine-grained spine movements

Theme: Motion Prior (Transformer, Diffusion, etc.) and Biophysical Prior help bridging

the gap between insufficient sensing and detailed human states

Concluding Thoughts

Modern Deep Learning ⇄ Bio/Physics Simulation

Physical Digital Human & World

The role of scalable simulation is irreplaceable for GenAI to continue to scale up:

Prior knowledge of physics/experts are very more dense in information

Simulation (synthesized data) brings expert knowledge to GenAI systems

Thank you!