A. Appendix Section

Due to page limitations in the main text, we have included additional supplementary materials in the appendix. This section primarily provides the following:

1. A comprehensive introduction to our newly proposed dataset (**Sec. A.1**).
2. A detailed description of the **implementation** (**Sec. A.2**).
3. A qualitative comparison and analysis of **inverse skinning** using neural blend skinning weights (**Sec. A.3**).
4. Setup for User Study (**Sec. A.4**).
5. Visualization of Material Prototype Centers (**Fig.7**).

A.1. Dataset

To evaluate PhysRig and compare its performance against traditional Linear Blend Skinning (LBS) methods, we construct a diverse simulation dataset tailored for the inverse skinning task. This dataset enables a comprehensive analysis of PhysRig's ability to recover underlying motion parameters and material properties under various challenging conditions.

We curate 17 structurally distinct objects from Objaverse, Mixamo, and The Amazing Animals Zoo, ensuring broad coverage of different articulation types and deformation patterns. These objects are categorized into three groups: (i) Humanoid Characters (5 objects), (ii) Quadruped Animals (6 objects): leopard, mammoth, stego, krin, cow, and raccoon, and (iii) Other Entities (6 objects): t-rex, pterosaur, whale, angelfish, cobra, and shark Each object is associated with 1 to 4 motion sequences, resulting in a total of 40 motion sequences, with each sequence containing 20 to 100 frames. This setup ensures a diverse range of temporal dynamics, allowing us to evaluate PhysRig's generalization across various topologies and articulation mechanisms.

To further assess PhysRig's ability to learn material properties, we provide two different material configurations for each of the 40 motion sequences, resulting in a total of 80 cases: (i) Homogeneous-material objects, where the entire structure exhibits a uniform material property. (ii) Heterogeneous-material objects, where different regions are assigned distinct material properties, simulating realistic soft-tissue variations and composite structures. In total, our dataset consists of 120 cases, including 40 cases with original motion sequences and 80 cases with different material configurations. By systematically introducing controlled material variations, our dataset enables a fine-grained evaluation of PhysRig's capability to recover Young's modulus, Poisson's ratio, and skeletal motion parameters across diverse material configurations. Given that the objects in our dataset primarily consist of animals and humans, which typically exhibit similar Poisson's ratios across different tissues, we assume a homogeneous Poisson's ratio for all objects. This dataset serves as a quantitative benchmark, evaluating both motion accuracy and material property estimation.

A.2. Implementation Details

In our experiments, the mesh object consists of approximately 2000 to 50000 vertices. We discretize the simulation field into a 100^3 grid for simulation. To reduce computational complexity, the number of material prototypes is set to be 25–200, which are uniformly distributed within the volume. For accurate motion modeling, we employ 100 substeps between successive frames (25 FPS), corresponding to a duration of:

$$\Delta t = 4 \times 10^{-4} \text{ seconds per sub-step.}$$

We adopt an alternating optimization strategy to separately optimize the material properties and velocity. Specifically, the material parameters are optimized using the AdamW optimizer, while the velocity parameters are optimized using SGD. To achieve efficient convergence, the initial learning rate for material training is set to be 20 times that of velocity training. The initial learning rate for velocity optimization is adjusted based on different scenarios, ranging from 5×10^{-3} to 2×10^{-2}, with commonly used values of 0.008 and 0.01. The material learning rate is set accordingly as 20 times the velocity learning rate. All optimization processes employ linear learning rate decay, with the learning rate reset to its initial value at the beginning of each alternating phase. During training, each scene undergoes three alternating optimization cycles for material and velocity, where the material is trained for 20 iterations per cycle, while the velocity is optimized for 30 iterations per frame. This iterative alternating optimization strategy gradually reduces the coupling between velocity and material, leading to more stable parameter learning.

Voxel-Based Adaptive Sampling. To achieve a comprehensive volumetric representation, we first load the input triangular mesh and extract its vertices and surface sample points. We then adaptively partition the space into multiple voxels based on the bounding box of the mesh and a predefined resolution. For each voxel, we evaluate its center point to determine whether it lies inside the mesh. If the center is within the mesh interior, we randomly generate a set number of sampling points inside the voxel. Each of these points is then checked to verify whether it remains inside the mesh. Finally, we aggregate all valid interior sampling points, along with surface points and original vertices, into a unified complete point cloud. This ensures a dense and well-distributed volumetric representation of the input mesh, which is subsequently outputted for further processing.

A.3. Qualitative Comparisons on Inverse Skinning

Although the LBS-based method is initialized with ground truth skinning weights, we observe that when jointly optimizing skinning weights and bone transformations, the optimization process can fall into a suboptimal local minimum due to

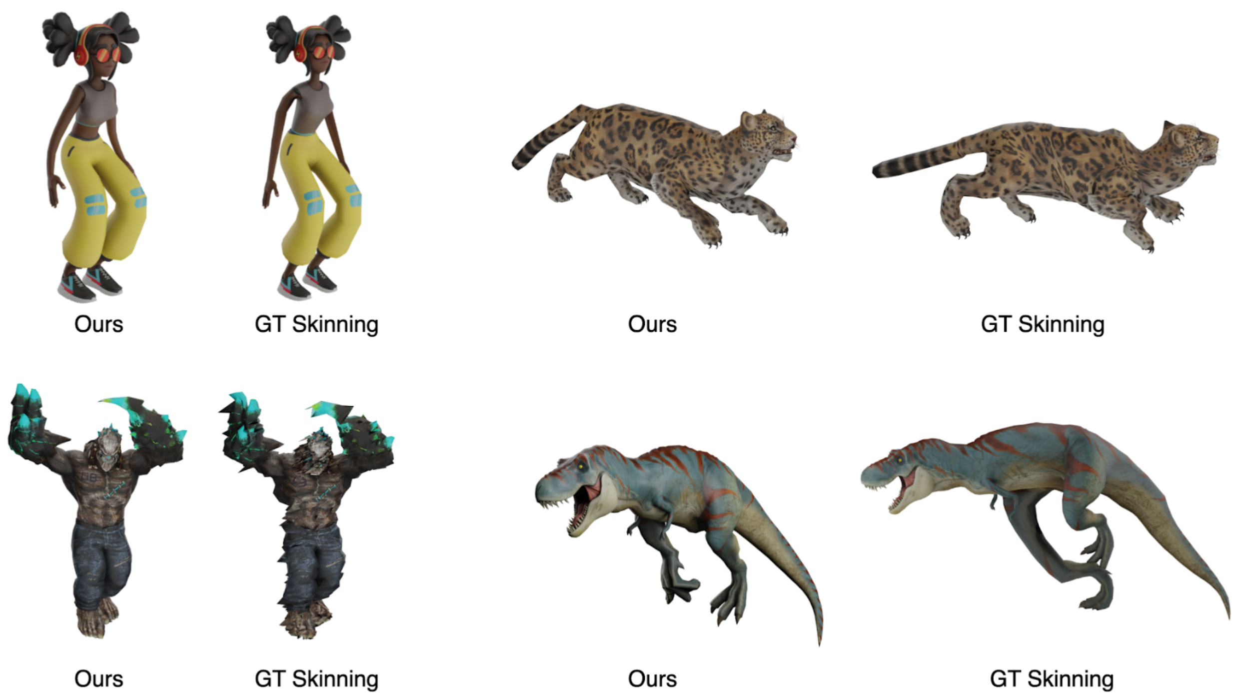

Figure 6. Qualitative Comparisons of **PhysRig** and neural linear blend skinning using ground truth skinning weights as initialization.

incorrect bone transformations, ultimately leading to unsatisfactory results. Additionally, despite incorporating the least motion loss, LBS often exhibits noticeable frame-to-frame jitter or abrupt changes in motion.

Moreover, LBS can suffer from various artifacts, such as unrealistic folding (e.g., in the leopard case), incorrect changes in volume size (e.g., in the Michelle case), or distorting the wrong body parts in an attempt to minimize loss, leading to severe deformations (e.g., in the T-Rex case).

In contrast, PhysRig, being grounded in a physics-based simulator, produces significantly more realistic and physically plausible results.

A.4. User Study

We provide a user study for comparison between PhysRig and neural linear blend skinning on inverse skinning. We asked 50 lay participants from online platforms for user studies, to rate the quality of **120** optimized mesh sequences from LBS-1, 2, 3, and PhysRig on a scale of 0 (low) to 5 (high). The participants are paid at an hourly rate of 16 USD. The results from different methods are anonymized as A to E, and the order is randomized. We provide the video visualization of these comparisons.

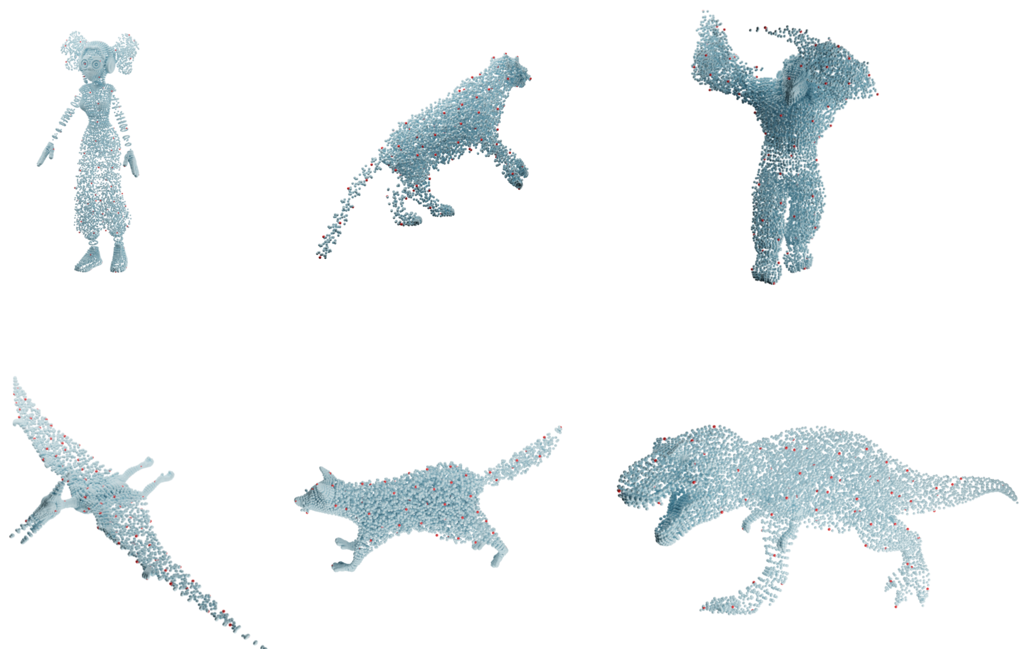

Figure 7. Visualization of Material Prototype Centers.