## Appendix:

We sincerely request readers to refer to the link below for more visualization results: https://zhengmianlun.github.io/publications/deepEmulator.html.

### A.1. Dataset Information

In this paper, we trained our network on a sphere dataset but tested it on five character meshes from the Adobe's Mixamo dataset [12]. Table A.1 provides detailed information about the five character meshes, including the vertex number and the edge length on the original surface mesh as well as the corresponding uniform volumetric mesh.

In Figure A.1, we show how we set constraints for each of the meshes, from a side view. The red vertices are constrained to move based on the skinned animation and drive the free vertices to deform with secondary motion.

### A.2. Full Quantitative and Qualitative Results

In Tables A.2- A.16, we provide the quantitative results of our network tested on the five character meshes and 15 motions. The corresponding error plots are given in Figures A.2-A.16. We also provide the error plots for the compared methods. Across all the test cases, our method achieves the most stable rollout prediction with the lowest error.

In **DeepEmulator.html**, we provide animation sequences of our results as well as other comparison methods.

### A.3. Further Analysis of Baseline Performance

As introduced in Section 4.2, we adopted the implicit backward Euler approach (Equation 3) as ground truth and the faster explicit central differences integration (Equation 2) as the baseline. Although the baseline method is 10 times faster than the implicit integrator with the same time step (1/24 second), it explodes after a few frames. In order to achieve stable simulation results, we found that it requires at least 100 sub-steps ($\Delta t \leq 0.0004$). In Table A.17, we provide the per-frame running time of the explicit integration with 50 and 100 steps.

### A.4. Choice of the Training Dataset

In Section 5, we mentioned a future direction of expanding the training dataset beyond primitive-based datasets such as spheres. Here, we analyze an alternative training dataset, namely the "Ortiz Dataset", created by running our physically-based simulator on the volumetric mesh surrounding the Ortiz character (same mesh as in Table A.1), with motions acquired from Adobe's Mixamo. In both datasets, we use the same number of frames. We report our results in Table A.18 to A.22.

Our experiments show that the network trained on the Sphere Dataset in most cases (75%) outperforms the Ortiz Dataset. We think there are two reasons for this. First, the local patches in the sphere are general and not specific to any geometry, making the learned neural network more general and therefore more suitable for characters other than Ortiz. Second, the motions in the Ortiz Dataset were created by human artists, and as such these motions follow certain human-selected artistic patterns. The motions in the Sphere Dataset, however, consist of random translations and rotations, which provides a denser sampling of motions in the possible motion space, and therefore improves the robustness of the network.

### A.5. Analysis of the Local Patch Size

In the main paper, we show our network architecture for 1-ring local patches (Figure 4). Namely, in the main paper the MLP $f_{\beta}^{\text{internal\_force}}$ learns to predict the internal forces from the **1**-ring neighbors around the center vertex. Here, we present an ablation study whereby the network learns based on 2-ring local patches, and 3-ring local patches, respectively. For 2-ring local patches, we add an additional MLP $f_{\beta 2}^{\text{internal\_force}}$ that receives the inputs from the 2-ring neighbors of the center vertex. The output latent vector is concatenated to the input of the $g_{\gamma}$ MLP. Similar operation is adopted for the 3-ring local patch network by adding another MLP for the 3-ring internal forces.

For the training loss, the network achieves the RMSE of 0.00257, 0.00159 and 0.00146 for 1-ring, 2-ring and 3-ring local patches, respectively. In Table A.23, we provide the corresponding test results on the five characters. Overall, we didn't see obvious improvements by increasing the local patch size. This could be because 2-ring and 3-ring local patches exhibits larger variability of structure, different to the sphere mesh, particularly for a center vertex close to the boundary. Therefore, we adopt 1-ring local patches in our paper.

| Character | Vertex Number (surface mesh) | Edge Length (surface mesh) | Disconnected Components | Vertex Number (tet mesh) | Edge Length (tet mesh) |
|---|---|---|---|---|---|
| Big vegas | 3711 | $[0.0024, 0.46]$ | 8 | 1468 | $[0.20, 0.35]$ |
| Kaya | 4260 | $[0.0049, 0.42]$ | 4 | 1417 | $[0.20, 0.35]$ |
| Michelle | 14267 | $[0.00047, 0.27]$ | 1 | 1105 | $[0.20, 0.35]$ |
| Mousey | 6109 | $[0.0023, 0.37]$ | 1 | 2303 | $[0.20, 0.35]$ |
| Ortiz | 24798 | $[0.00087, 0.085]$ | 1 | 1258 | $[0.20, 0.35]$ |

Table A.1: Detailed information on the five test characters. Each character's surface mesh was re-scaled uniformly to lie exactly within a bounding box of dimensions 5×5×5.

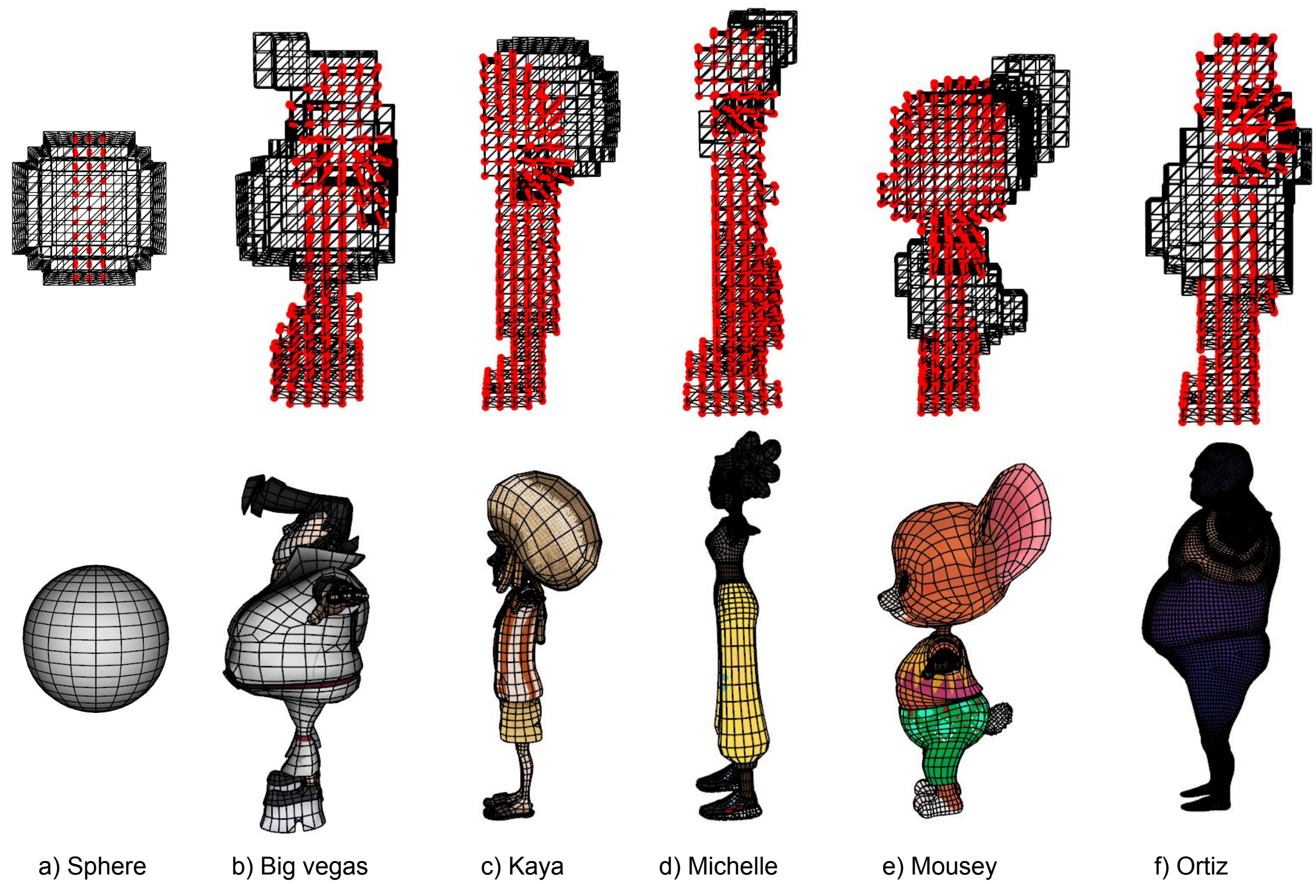


Figure A.1: The constraints (red vertices) set on the volumetric mesh surrounding the surface mesh.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground Truth | \ | \ | \ | \ | $[1.21E5, 2.76E5, 1.02E6]$ |
| Our Method | **0.0098** | **0.053** | **0.063** | **0.059** | $[\mathbf{1.56E5}, \mathbf{4.92E5}, \mathbf{2.83E6}]$ |
| Ours w/o ref. motion | 0.058 | 0.19 | 0.55 | 7.70 | $[2.76E5, 1.5E15, 6.5E15]$ |
| Baseline | \ | 8.23E120 | 1.07E121 | 1.57E121 | $[2.28E5, Nan, 4.71E165]$ |
| CFD-GCN [5] | 0.031 | 50.33 | 72.06 | 79.16 | $[2.60E5, 2.4E18, 9.9E18]$ |
| GNS [23] | 0.057 | 0.20 | 0.31 | 0.55 | $[3.24E5, 1.79E9, 6.83E9]$ |
| MeshGraphNets [21] | 0.058 | 0.14 | 0.10 | 2.85 | $[2.62E5, 3.9E12, 1.8E13]$ |

Table A.2: Quantitative results: Big vegas, 283-frame hip hop dancing 1.

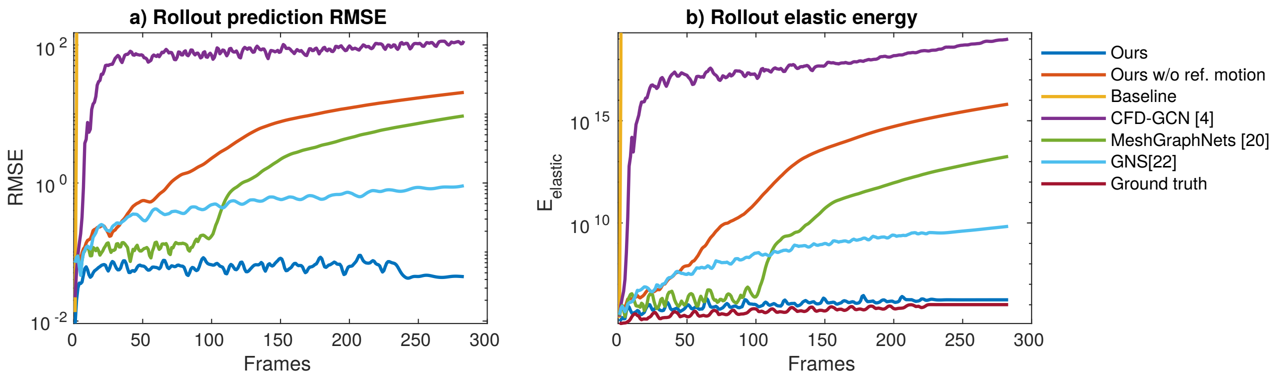


Figure A.2: Plot of the quantitative results: Big vegas, 283-frame hip hop dancing 1.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground Truth | \ | \ | \ | \ | $[2.42E5, 3.53E5, 1.77E5]$ |
| Our Method | **0.0093** | **0.066** | **0.074** | **0.073** | $[\mathbf{3.20E5}, \mathbf{8.18E5}, \mathbf{6.32E6}]$ |
| Ours w/o ref. motion | 0.061 | 0.31 | 0.87 | 14.19 | $[1.10E6, 2.6E16, 1.0E17]$ |
| Baseline | \ | 7.83E120 | 1.06E121 | 1.79E121 | $[3.17E5, Nan, 4.69E165]$ |
| CFD-GCN [5] | 0.031 | 67.91 | 110.82 | 591.46 | $[5.79E5, 3.8E22, 1.6E23]$ |
| GNS [23] | 0.060 | 0.32 | 0.50 | 0.68 | $[1.90E6, 5.64E9, 1.9E10]$ |
| MeshGraphNets [21] | 0.062 | 0.16 | 0.38 | 6.58 | $[7.58E5, 3.9E13, 1.6E14]$ |

Table A.3: Quantitative results: Big vegas, 366-frame hip hop dancing 2.

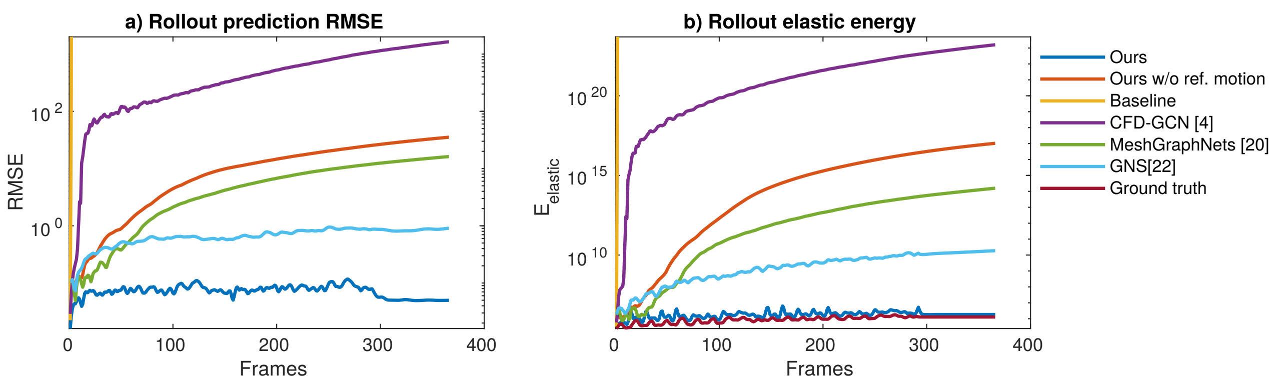


Figure A.3: Plot of the quantitative results: Big vegas, 366-frame hip hop dancing 2.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground Truth | \ | \ | \ | \ | $[3.82E4, 3.99E5, 1.50E6]$ |
| Our Method | **0.0062** | **0.050** | **0.057** | **0.065** | $[\mathbf{4.12E4}, \mathbf{6.68E5}, \mathbf{2.76E6}]$ |
| Ours w/o ref. motion | 0.047 | 0.15 | 0.36 | 21.37 | $[6.19E4, 8.1E16, 3.4E17]$ |
| Baseline | \ | 7.67E120 | 1.07E121 | 2.51E121 | $[4.40E3, Nan, 4.68E165]$ |
| CFD-GCN [5] | 0.027 | 47.90 | 76.76 | 110.26 | $[8.89E4, 3.8E17, 1.5E19]$ |
| GNS [23] | 0.046 | 0.17 | 0.32 | 0.56 | $[5.77E4, 4.20E9, 1.8E10]$ |
| MeshGraphNets [21] | 0.048 | 0.084 | 0.081 | 10.12 | $[6.23E4, 2.8E14, 1.1E15]$ |

Table A.4: Quantitative results: Big vegas, 594-frame samba dancing 1.

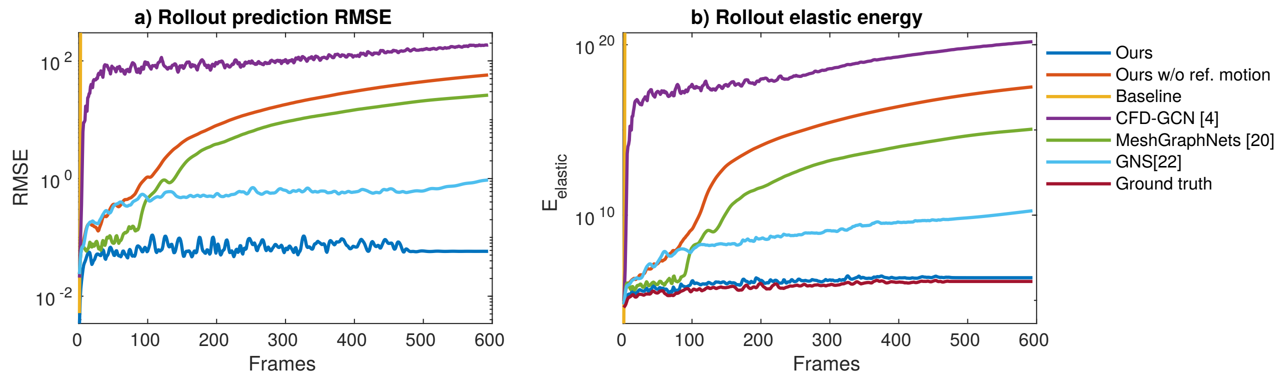


Figure A.4: Plot of the quantitative results: Big vegas, 594-frame samba dancing 1.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground Truth | \ | \ | \ | \ | $[8.02E4, 3.72E5, 1.74E6]$ |
| Our Method | **0.0058** | **0.052** | **0.050** | **0.064** | $[\mathbf{9.76E4}, \mathbf{6.35E5}, \mathbf{3.18E6}]$ |
| Ours w/o ref. motion | 0.038 | 0.13 | 0.40 | 13.88 | $[1.45E5, 1.8E16, 7.7E16]$ |
| Baseline | \ | 7.67E120 | 9.97E120 | 2.15E121 | $[9.58E3, Nan, 4.70E165]$ |
| CFD-GCN [5] | 0.027 | 75.12 | 86.49 | 101.63 | $[1.59E5, 1.9E19, 7.7E19]$ |
| GNS [23] | 0.037 | 0.17 | 0.30 | 0.62 | $[1.29E5, 5.85E8, 3.32E9]$ |
| MeshGraphNets [21] | 0.040 | 0.090 | 0.091 | 7.89 | $[1.36E5, 9.3E13, 3.8E14]$ |

Table A.5: Quantitative results: Big vegas, 493-frame samba dancing 2.

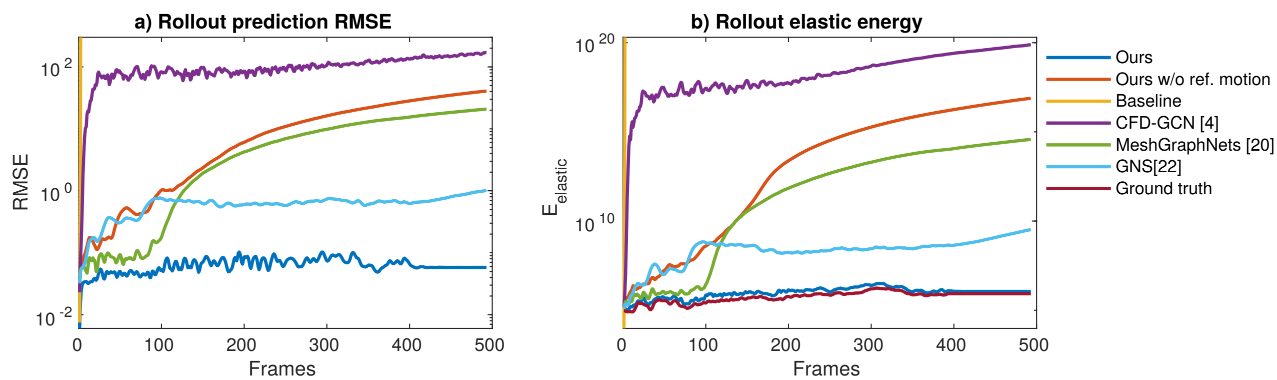


Figure A.5: Plot of the quantitative results: Big vegas, 493-frame samba dancing 2.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground Truth | \ | \ | \ | \ | $[2.24E5, 2.06E5, 1.34E6]$ |
| Our Method | **0.0065** | **0.062** | **0.054** | **0.070** | $[\mathbf{2.92E5}, \mathbf{4.31E5}, \mathbf{1.98E6}]$ |
| Ours w/o ref. motion | 0.040 | 0.35 | 0.90 | 12.82 | $[5.12E5, 1.6E16, 6.6E16]$ |
| Baseline | \ | 7.76E120 | 1.02E121 | 1.96E121 | $[4.48E5, Nan, 4.70E165]$ |
| CFD-GCN [5] | 0.083 | 50.17 | 71.88 | 79.87 | $[1.95E6, 1.1E17, 4.9E17]$ |
| GNS [23] | 0.040 | 0.18 | 0.31 | 0.50 | $[5.61E5, 2.85E9, 1.1E10]$ |
| MeshGraphNets [21] | 0.043 | 0.13 | 0.11 | 5.24 | $[4.09E5, 2.6E13, 1.1E15]$ |

Table A.6: Quantitative results: Big vegas, 399-frame samba dancing 3.

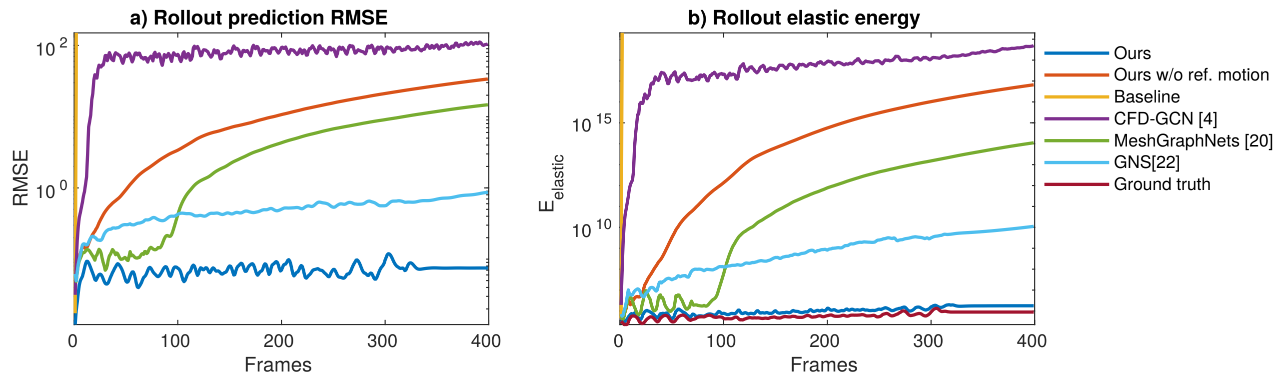


Figure A.6: Plot of the quantitative results: Big vegas, 399-frame samba dancing 3.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground Truth | \ | \ | \ | \ | $[5.47E4, 1.31E5, 6.40E51]$ |
| Our Method | **0.0067** | **0.054** | **0.058** | **0.041** | $[\mathbf{5.68E4}, \mathbf{1.50E5}, \mathbf{1.03E6}]$ |
| Ours w/o ref. motion | 0.042 | 0.097 | 0.15 | 20.02 | $[7.70E4, 1.1E16, 4.7E17]$ |
| Baseline | \ | 6.21E120 | 8.00E120 | 2.00E121 | $[1.41E4, Nan, 3.15E165]$ |
| CFD-GCN [5] | 0.016 | 0.22 | 72.15 | 69.87 | $[7.33E4, 3.6E17, 1.7E18]$ |
| GNS [23] | 0.041 | 0.15 | 0.28 | 0.47 | $[7.20E4, 8.63E8, 4.02E9]$ |
| MeshGraphNets [21] | 0.042 | 0.063 | 0.084 | 0.068 | $[7.27E4, 3.03E5, 2.12E6]$ |

Table A.7: Quantitative results: Kaya, 650-frame dancing running man.

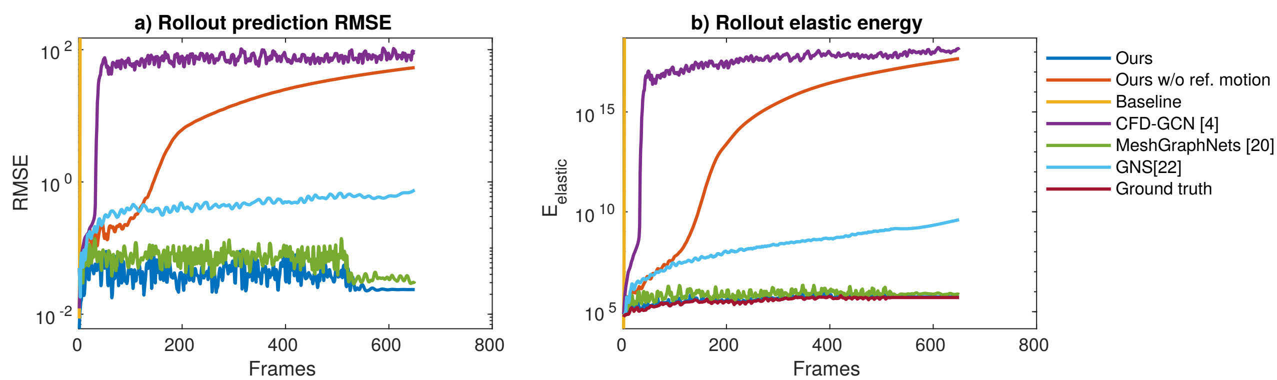


Figure A.7: Plot of the quantitative results: Kaya, 650-frame dancing running man.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[5.14E4, 7.09E4, 3.15E5]$ |
| Our Method | **0.0075** | **0.083** | **0.067** | **0.054** | $[\mathbf{5.17E4}, \mathbf{1.49E5}, \mathbf{7.76E5}]$ |
| Ours w/o ref. motion | 0.030 | 0.23 | 0.41 | 3.16 | $[5.87E4, 2.1E13, 8.4E13]$ |
| Baseline | \ | 5.66E120 | 7.98E120 | 8.99E120 | $[9.30E2, Nan, 3.15E165]$ |
| CFD-GCN [5] | 0.023 | 35.34 | 62.31 | 59.14 | $[7.21E4, 8.9E16, 4.0E17]$ |
| GNS [23] | 0.029 | 0.21 | 0.29 | 0.45 | $[5.83E4, 3.65E7, 1.16E8]$ |
| MeshGraphNets [21] | 0.03 | 0.11 | 0.15 | 2.44 | $[5.88E4, 5.3E10, 2.5E12]$ |

Table A.8: Quantitative results: Kaya, 167-frame zombie scream.

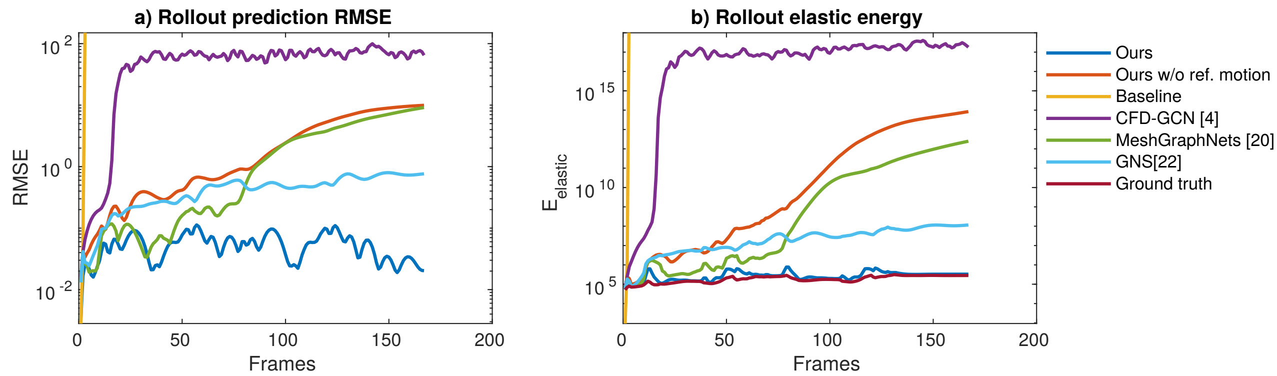


Figure A.8: Plot of the quantitative results: Kaya, 167-frame zombie scream.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[2.18E4, 1.12E5, 4.93E5]$ |
| Our Method | **0.0041** | **0.033** | **0.033** | **0.04** | $[\mathbf{2.20E4}, \mathbf{1.25E5}, \mathbf{6.45E5}]$ |
| Ours w/o ref. motion | 0.060 | 0.12 | 0.19 | 5.24 | $[2.27E4, 1.2E15, 5.5E15]$ |
| Baseline | \ | 5.81E120 | 7.86E120 | 1.52E121 | $[9.26E0, Nan, 1.27E165]$ |
| CFD-GCN [5] | 0.017 | 27.36 | 42.25 | 73.72 | $[4.58E4, 3.7E18, 1.4E19]$ |
| GNS [23] | 0.060 | 0.10 | 0.19 | 0.37 | $[2.26E4, 4.79E9, 1.9E10]$ |
| MeshGraphNets [21] | 0.06 | 0.082 | 0.11 | 0.079 | $[2.26E4, 1.76E5, 9.94E5]$ |

Table A.9: Quantitative results: Michelle, 371-frame gangnam style.

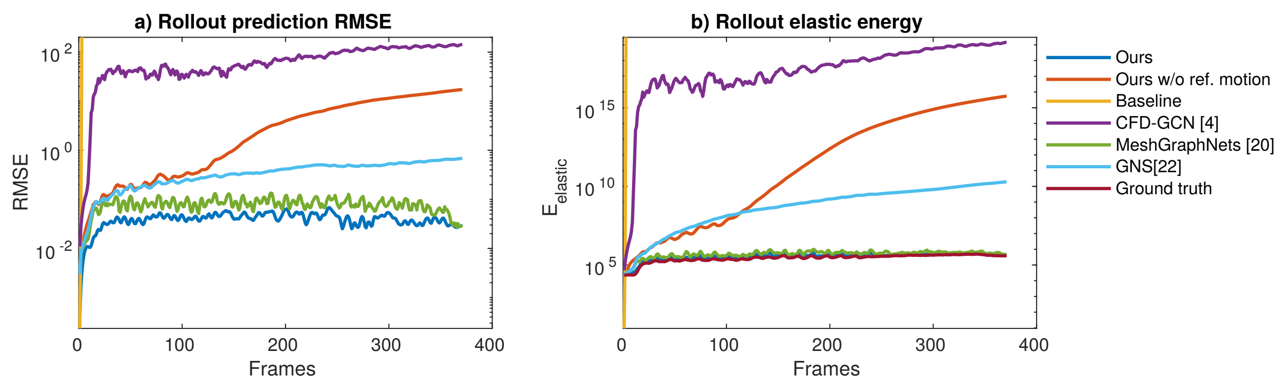


Figure A.9: Plot of the quantitative results: Michelle, 371-frame gangnam style.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[1.22836E5, 4.67E5, 2.60E6]$ |
| Our Method | **0.0056** | **0.025** | **0.024** | **0.047** | $[\mathbf{1.43E5}, \mathbf{5.06E5}, \mathbf{2.79E6}]$ |
| Ours w/o ref. motion | 0.082 | 0.13 | 0.15 | 15.20 | $[2.96E5, 5.3E16, 2.2E17]$ |
| Baseline | \ | 6.25E120 | 7.47E120 | 2.25E121 | $[1.25E5, Nan, 1.27E165]$ |
| CFD-GCN [5] | 0.019 | 36.06 | 36.80 | 64.16 | $[2.05E5, 1.5E19, 5.7E19]$ |
| GNS [23] | 0.082 | 0.12 | 0.14 | 0.48 | $[2.35E5, 5.1E10, 2.0E11]$ |
| MeshGraphNets [21] | 0.082 | 0.065 | 0.077 | 4.31 | $[2.06E5, 6.5E12, 2.2E13]$ |

Table A.10: Quantitative results: Michelle, 627-frame swing dancing 1.

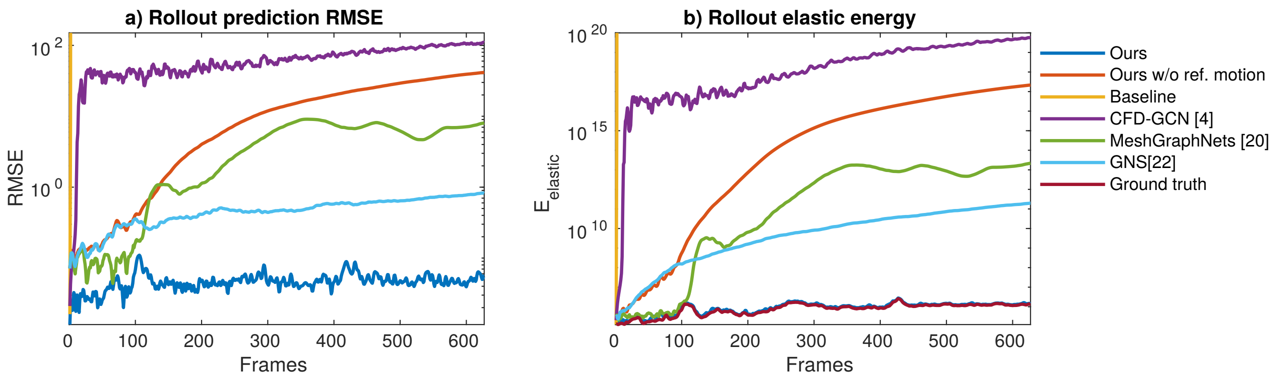


Figure A.10: Plot of the quantitative results: Michelle, 627-frame swing dancing 1.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[4.46E4, 4.18E5, 1.82E6]$ |
| Our Method | **0.0056** | **0.049** | **0.037** | **0.055** | $[\mathbf{4.54E4}, \mathbf{4.48E5}, \mathbf{2.01E6}]$ |
| Ours w/o ref. motion | 0.086 | 0.14 | 0.16 | 20.14 | $[5.66E4, 1.1E17, 4.2E17]$ |
| Baseline | \ | 5.71E120 | 7.69E120 | 2.35E121 | $[2.28E3, Nan, 1.27E165]$ |
| CFD-GCN [5] | 0.019 | 25.05 | 42.43 | 64.23 | $[5.58E4, 3.6E18, 1.6E19]$ |
| GNS [23] | 0.085 | 0.13 | 0.19 | 0.43 | $[5.31E4, 2.7E10, 1.0E14]$ |
| MeshGraphNets [21] | 0.086 | 0.11 | 0.094 | 0.11 | $[5.59E4, 4.61E5, 2.20E6]$ |

Table A.11: Quantitative results: Michelle, 699-frame swing dancing 2.

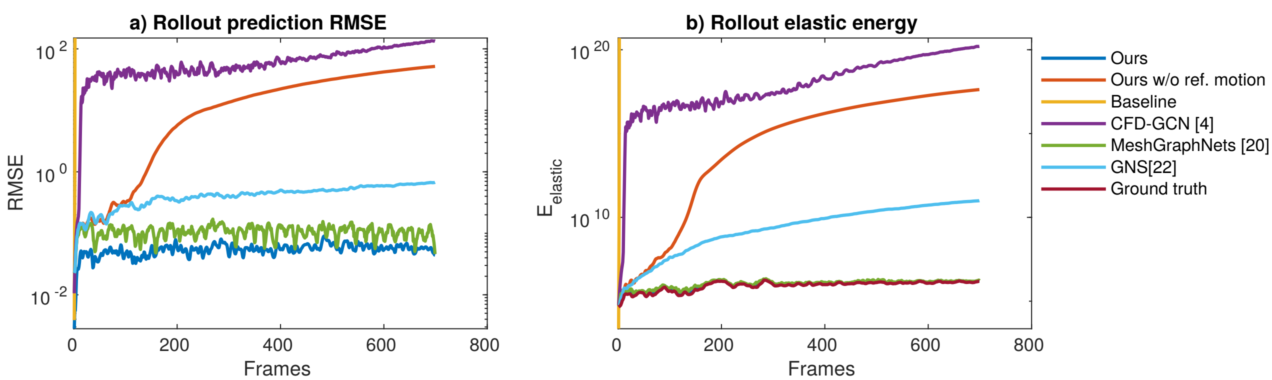


Figure A.11: Plot of the quantitative results: Michelle, 699-frame swing dancing 2.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[1.20E5, 8.97E4, 5.66E5]$ |
| Our Method | **0.0080** | **0.077** | **0.10** | **0.086** | $[\mathbf{2.45E5}, \mathbf{2.49E5}, \mathbf{1.60E6}]$ |
| Ours w/o ref. motion | 0.057 | 0.24 | 0.42 | 1.43 | $[1.27E6, 2.7E11, 1.7E13]$ |
| Baseline | \ | 7.78E120 | 1.07E121 | 1.24E121 | $[3.79E5, Nan, 7.19E165]$ |
| CFD-GCN [5] | 0.041 | 56.37 | 82.57 | 72.71 | $[4.10E5, 2.5E17, 1.1E18]$ |
| GNS [23] | 0.057 | 0.50 | 0.69 | 0.92 | $[1.02E6, 5.74E8, 1.76E9]$ |
| MeshGraphNets [21] | 0.057 | 0.15 | 1.30 | 3.48 | $[7.24E5, 4.3E10, 2.0E15]$ |

Table A.12: Quantitative results: Mousey, 158-frame dancing.

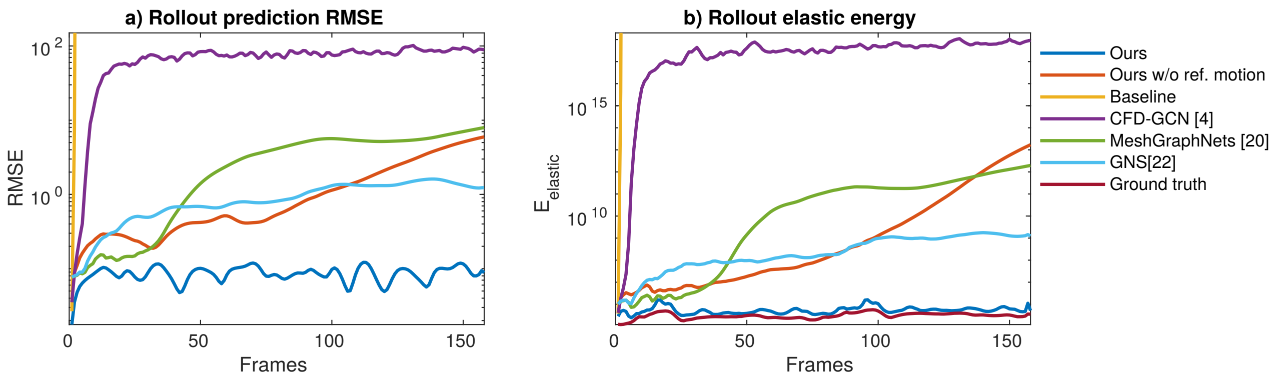


Figure A.12: Plot of the quantitative results: Mousey, 158-frame dancing.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[1.01E5, 1.14E5, 7.62E5]$ |
| Our Method | **0.0066** | **0.067** | **0.11** | **0.09** | $[\mathbf{1.04E5}, \mathbf{3.37E5}, \mathbf{2.12E6}]$ |
| Ours w/o ref. motion | 0.036 | 0.19 | 0.28 | 2.95 | $[1.43E5, 1.2E13, 6.3E14]$ |
| Baseline | \ | 8.29E120 | 1.12E121 | 1.51E121 | $[2.69E4, Nan, 7.23E165]$ |
| CFD-GCN [5] | 0.043 | 68.29 | 78.87 | 75.08 | $[2.42E5, 3.4E17, 1.3E18]$ |
| GNS [23] | 0.036 | 0.26 | 0.62 | 0.86 | $[1.37E5, 2.92E8, 1.31E9]$ |
| MeshGraphNets [21] | 0.037 | 0.15 | 1.61 | 8.63 | $[1.39E5, 1.1E15, 4.4E15]$ |

Table A.13: Quantitative results: Mousey, 255-frame shuffling.

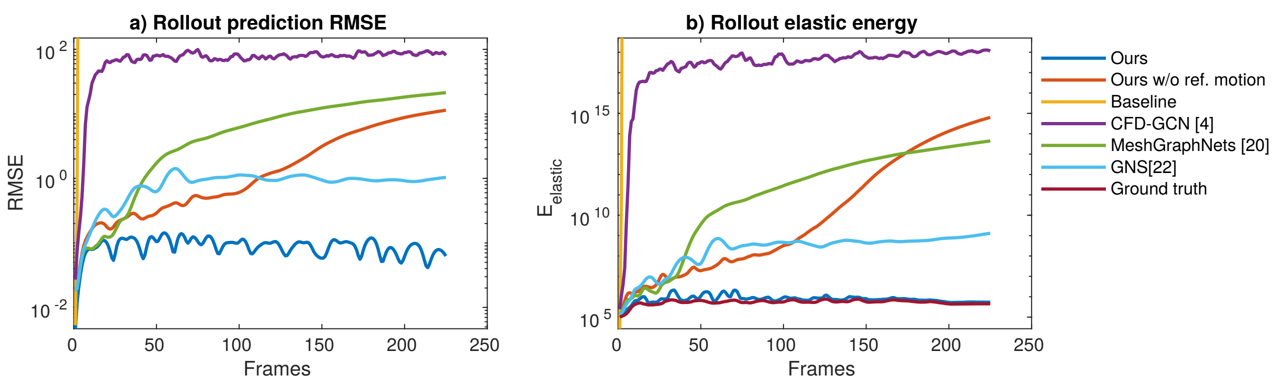


Figure A.13: Plot of the quantitative results: Mousey, 255-frame shuffling.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[4.00E5, 9.80E4, 8.39E5]$ |
| Our Method | **0.0090** | **0.087** | **0.10** | **0.10** | $[\mathbf{5.27E5}, \mathbf{3.23E5}, \mathbf{2.19E6}]$ |
| Ours w/o ref. motion | 0.039 | 0.36 | 0.35 | 7.86 | $[1.23E6, 6.7E15, 2.9E16]$ |
| Baseline | \ | 8.48E120 | 1.10E121 | 1.81E121 | $[3.12E5, Nan, 7.20E165]$ |
| CFD-GCN [5] | 0.17 | 66.19 | 73.93 | 78.95 | $[1.42E8, 1.5E17, 5.2E17]$ |
| GNS [23] | 0.039 | 0.28 | 0.50 | 0.63 | $[9.23E5, 8.74E8, 3.58E9]$ |
| MeshGraphNets [21] | 0.040 | 0.21 | 2.14 | 14.97 | $[7.86E5, 8.4E13, 3.2E14]$ |

Table A.14: Quantitative results: Mousey, 627-frame swing dancing.

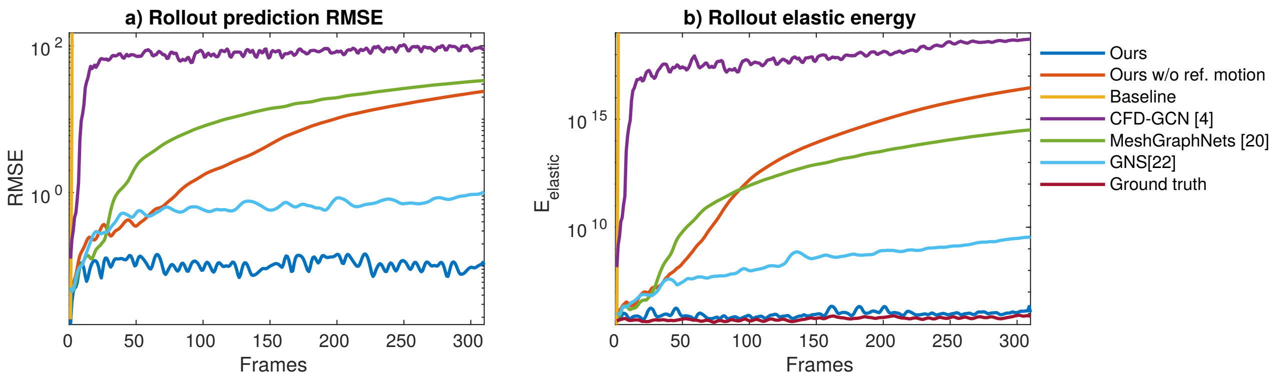


Figure A.14: Plot of the quantitative results: Mousey, 627-frame swing dancing.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[6.02E5, 3.24E4, 7.38E5]$ |
| Our Method | **0.0057** | **0.082** | **0.077** | **0.073** | $[\mathbf{6.16E5}, \mathbf{1.55E5}, \mathbf{1.35E6}]$ |
| Ours w/o ref. motion | 0.041 | 0.29 | 0.40 | 1.20 | $[8.01E5, 3.6E12, 2.1E13]$ |
| Baseline | \ | 8.15E120 | 1.09E121 | 1.09E121 | $[7.17E4, Nan, 4.00E165]$ |
| CFD-GCN [5] | 0.030 | 10.71 | 65.43 | 58.90 | $[9.47E5, 1.7E18, 7.0E18]$ |
| GNS [23] | 0.040 | 0.30 | 0.22 | 0.27 | $[7.46E5, 3.70E7, 1.31E8]$ |
| MeshGraphNets [21] | 0.042 | 0.090 | 0.088 | 0.096 | $[6.35E5, 2.67E5, 1.66E6]$ |

Table A.15: Quantitative results: Ortiz, 122-frame cross jumps rotation.

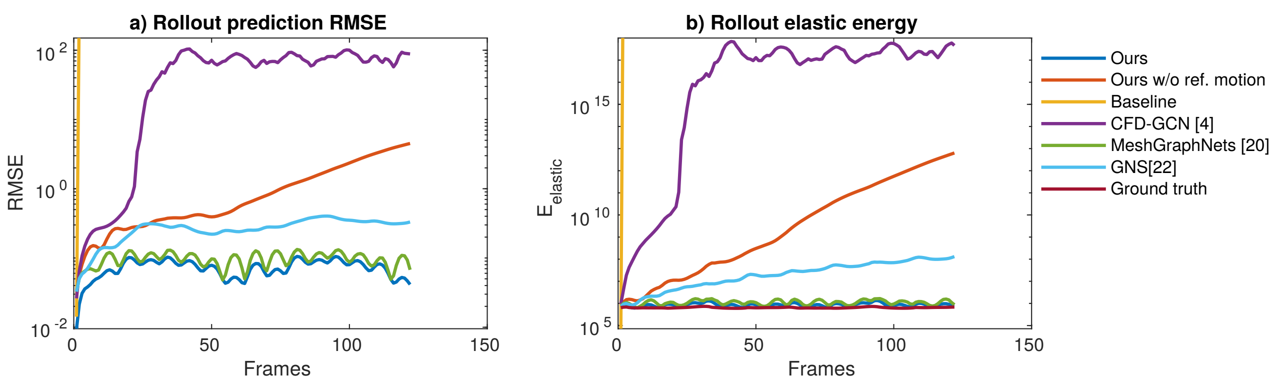


Figure A.15: Plot of the quantitative results: Ortiz, 122-frame cross jumps rotation.

| Methods | Single Frame | Rollout-24 | Rollout-48 | Rollout-All | $E_{elastic}$ $[min, stdev, max]$ |
|---|---|---|---|---|---|
| Ground truth | \ | \ | \ | \ | $[3.81E3, 4.86E4, 2.40E5]$ |
| Our Method | **0.0039** | **0.039** | **0.032** | **0.036** | $[\mathbf{4.84E3}, \mathbf{7.56E4}, \mathbf{4.02E5}]$ |
| Ours w/o ref. motion | 0.034 | 0.10 | 0.17 | 4.18 | $[1.62E4, 1.7E15, 7.9E15]$ |
| Baseline | \ | 7.40E120 | 9.48E120 | 1.62E121 | $[2.89E3, Nan, 3.99E165]$ |
| CFD-GCN [5] | 0.017 | 0.57 | 83.52 | 71.93 | $[3.96E4, 3.4E17, 1.9E18]$ |
| GNS [23] | 0.034 | 0.17 | 0.21 | 0.31 | $[1.09E4, 8.91E7, 3.37E8]$ |
| MeshGraphNets [21] | 0.034 | 0.065 | 0.071 | 0.064 | $[1.69E4, 1.40E5, 8.45E5]$ |

Table A.16: Quantitative results: Ortiz, 326-frame jazz dancing.

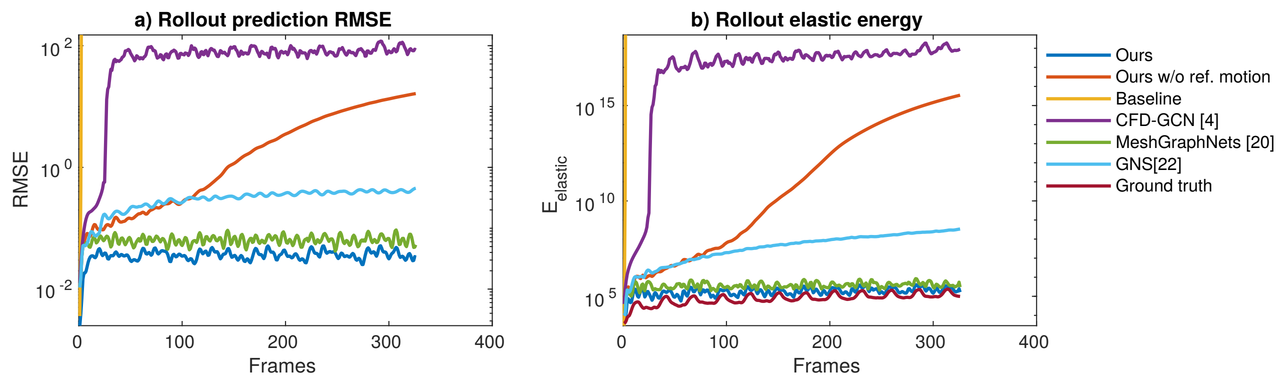


Figure A.16: Plot of the quantitative results: Ortiz, 326-frame jazz dancing.

| character | # vertices (tet mesh) | $t_{GT}$ s/frame 1 step / frame | $t_{BL}$ s/frame 1 step / frame | $t_{BL}$ s/frame 50 steps / frame | $t_{BL}$ s/frame 100 steps / frame | $t_{ours}$ s/frame 1 step / frame |
|---|---|---|---|---|---|---|
| Big vegas | 1468 | 0.58 | 0.056 | 2.57027 | 6.20967 | **0.017** |
| Kaya | 1417 | 0.52 | 0.052 | 2.42985 | 5.72762 | **0.015** |
| Michelle | 1105 | 0.33 | 0.032 | 1.52916 | 3.64744 | **0.013** |
| Mousey | 2303 | 0.83 | 0.084 | 3.90579 | 9.5897 | **0.018** |
| Ortiz | 1258 | 0.51 | 0.049 | 2.2496 | 5.16806 | **0.015** |

Table A.17: The running time of the ground truth, the baseline, and our method.

| Motion | Dataset | Single Frame | Rollout-24 | Rollout-48 | Rollout-All |
|---|---|---|---|---|---|
| Hip hop dancing 1 283 frames | Sphere Dataset | **0.0098** | 0.053 | **0.063** | **0.0591** |
| | Ortiz Dataset | 0.0111 | **0.0512** | 0.067 | 0.0969 |
| Hip hop dancing 2 366 frames | Sphere Dataset | **0.0093** | 0.0664 | 0.0744 | **0.0727** |
| | Ortiz Dataset | 0.0101 | **0.0564** | **0.0607** | 0.1241 |
| Samba dancing 1 594 frames | Sphere Dataset | **0.0062** | 0.0495 | **0.0571** | **0.0654** |
| | Ortiz Dataset | 0.0062 | **0.0481** | 0.061 | 0.143 |
| Samba dancing 2 493 frames | Sphere Dataset | **0.0058** | 0.0521 | 0.0496 | **0.0635** |
| | Ortiz Dataset | 0.0064 | **0.0367** | **0.0423** | 0.1331 |
| Samba dancing 3 399 frames | Sphere Dataset | **0.0065** | 0.0615 | **0.0537** | **0.0702** |
| | Ortiz Dataset | 0.0065 | **0.0383** | 0.0616 | 0.1282 |

Table A.18: Quantitative results: the network trained on different datasets and tested on the character Big vegas.

| Motion | Dataset | Single Frame | Rollout-24 | Rollout-48 | Rollout-All |
|---|---|---|---|---|---|
| Dancing running man 650 frames | Sphere Dataset | **0.0067** | **0.0544** | **0.0578** | **0.0411** |
| | Ortiz Dataset | 0.0113 | 0.075 | 0.0839 | 0.1176 |
| Zombie scream 167 frames | Sphere Dataset | **0.0075** | 0.0834 | 0.0666 | **0.0537** |
| | Ortiz Dataset | 0.0126 | **0.0749** | **0.0645** | 0.076 |

Table A.19: Quantitative results: the network trained on different datasets and tested on the character Kaya.

| Motion | Dataset | Single Frame | Rollout-24 | Rollout-48 | Rollout-All |
|---|---|---|---|---|---|
| Gangnam style | Sphere Dataset | **0.0041** | **0.0329** | 0.0332 | **0.0401** |
| 371 frames | Ortiz Dataset | 0.0059 | 0.0329 | **0.0319** | 0.0969 |
| Swing dancing 1 | Sphere Dataset | **0.0056** | **0.025** | **0.0236** | **0.0471** |
| 627 frames | Ortiz Dataset | 0.0079 | 0.0324 | 0.035 | 0.1213 |
| Swing dancing 2 | Sphere Dataset | **0.0056** | **0.0491** | **0.0373** | **0.0548** |
| 699 frames | Ortiz Dataset | 0.0067 | 0.058 | 0.04 | 0.1264 |

Table A.20: Quantitative results: the network trained on different datasets and tested on the character Michelle.

| Motion | Dataset | Single Frame | Rollout-24 | Rollout-48 | Rollout-All |
|---|---|---|---|---|---|
| Dancing | Sphere Dataset | **0.008** | **0.0771** | **0.1003** | **0.0858** |
| 158 frames | Ortiz Dataset | 0.0122 | 0.0871 | 0.1056 | 0.122 |
| Shuffling | Sphere Dataset | **0.0066** | **0.0666** | 0.1115 | **0.09** |
| 225 frames | Ortiz Dataset | 0.0115 | 0.0936 | **0.1072** | 0.1303 |
| Swing dancing | Sphere Dataset | **0.009** | **0.0871** | **0.1001** | **0.1042** |
| 627 frames | Ortiz Dataset | 0.0144 | 0.1236 | 0.1113 | 0.1594 |

Table A.21: Quantitative results: the network trained on different datasets and tested on the character Mousey.

| Motion | Dataset | Single Frame | Rollout-24 | Rollout-48 | Rollout-All |
|---|---|---|---|---|---|
| cross jumps rotation | Sphere Dataset | 0.0057 | 0.0819 | 0.0765 | 0.0726 |
| 122 frames | Ortiz Dataset | **0.0053** | **0.0719** | **0.0532** | **0.0702** |
| jazz dancing | Sphere Dataset | **0.0039** | **0.0391** | **0.0316** | **0.0363** |
| 326 frames | Ortiz Dataset | 0.0051 | 0.0365 | 0.0338 | 0.0946 |

Table A.22: Quantitative results: the network trained on different datasets and tested on the character Ortiz.

| Test Dataset | Patch size | Single Frame | Rollout-24 | Rollout-48 | Rollout-All |
|---|---|---|---|---|---|
| Big vegas | 1-ring | 0.0075 | 0.057 | 0.060 | 0.066 |
| | 2-ring | **0.0068** | 0.060 | 0.066 | 0.077 |
| | 3-ring | 0.0075 | **0.043** | **0.0418** | **0.059** |
| Kaya | 1-ring | 0.0071 | **0.069** | 0.062 | **0.047** |
| | 2-ring | **0.0060** | 0.078 | **0.060** | 0.072 |
| | 3-ring | 0.0064 | 0.099 | 0.095 | 0.092 |
| Michelle | 1-ring | 0.0051 | 0.036 | **0.031** | 0.047 |
| | 2-ring | **0.0045** | **0.033** | 0.033 | **0.047** |
| | 3-ring | 0.0047 | 0.043 | 0.042 | 0.059 |
| Mousey | 1-ring | 0.0079 | **0.077** | **0.10** | **0.093** |
| | 2-ring | **0.0069** | 0.11 | 0.19 | 0.14 |
| | 3-ring | 0.0075 | 0.11 | 0.21 | 0.16 |
| Ortiz | 1-ring | 0.0048 | **0.061** | **0.054** | **0.054** |
| | 2-ring | **0.0042** | 0.069 | 0.089 | 0.070 |
| | 3-ring | 0.0051 | 0.094 | 0.097 | 0.087 |

Table A.23: Quantitative results: the network trained on local patches of different patch sizes.