

# OMNIPHYSGS: 3D CONSTITUTIVE GAUSSIANS FOR GENERAL PHYSICS-BASED DYNAMICS GENERATION

**Anonymous authors**

Paper under double-blind review

## REBUTTAL APPENDIX

We provide additional details and results to address the reviewers’ concerns and comments.

- **For Reviewer 7Xoc.** Section C provides experiments comparing our method to directly optimizing coefficients for different constitutive models, which demonstrates the difference and contribution of our work compared to related methods (Nagasawa et al., 2019; Su et al., 2023). Detailed analysis of these works is provided in the OpenReview comments. Section E provides more qualitative results, with prompts that do not describe motions, to demonstrate the robustness of our method.
- **For Reviewer h6D1.** Section B provides an analysis and improved results of the artifacts in the can-duck collision scene. Section E provides more qualitative results regarding the same material with different material parameters.
- **For Reviewer ktNn.** Section A provides the detailed results of the user study as a complement to the metrics in the main manuscript. Section B provides an analysis and improved results of the artifacts in the can-duck collision scene. Section C provides experiments comparing our method to directly optimizing a vector representing the probability of different pre-defined constitutive models. Section E provides more qualitative results regarding real-world scenes.
- **For Reviewer S54K.** Section E provides more qualitative results regarding both multiple objects and real-world scenes.
- **For Reviewer kAsd.** Section C provides experiments comparing our method to directly optimizing coefficients for different constitutive models, which demonstrates the difference and contribution of our work compared to the related method (Su et al., 2023). Detailed analysis of the work is provided in the OpenReview comments. Section D provides the revised version of the original Figure 3.

For other comments and concerns, we provide detailed responses in the OpenReview comments. We appreciate the reviewers’ valuable feedback and suggestions. Thank you for your time and input.

### A USER STUDY

As a complement to the metrics in the main manuscript, we conducted a user study among 20 participants to evaluate the quality of the generated dynamics by different methods. In the study, the participants were asked to rank the videos based on two criteria: the text alignment and physical plausibility of the dynamics. Table 1 shows the detailed results of the user study, where the numbers indicate the average ranking of each method. A lower ranking reflects better performance. The results indicate that our method achieves better performance in modeling various dynamics while the baseline methods achieve comparable performance only in the pure elasticity case. This conclusion aligns with the quantitative and qualitative findings discussed in the main manuscript. Figure 2 presents the user interface of our user study.

### B ANALYSIS OF ARTIFACTS IN COLLISION

We analyze the artifacts occurring in the generated dynamics, particularly those involving complex interactions such as collisions. Our findings indicate that the number of grids in the MPM simulator is a key factor influencing the quality of generated dynamics. Figure 4 (see the supplementary video for better visualization) compares the dynamics generated with varying grid resolutions. The results

054 demonstrate that increasing the number of grids effectively reduces aliasing artifacts and enhances  
055 the physical plausibility of the generated dynamics.  
056

## 057 C DISCUSSION ON THE NETWORK ARCHITECTURE 058

059 In the main manuscript, we implement the proposed physics-guided network with a neural network  
060 composed of two components, a 3D feature encoder and a physics-aware decoder. A naive way  
061 to simplify the network is to directly optimize a vector representing the probability of different  
062 pre-defined constitutive models for each particle. However, our preliminary experiments revealed  
063 that this straightforward method struggles with convergence and is prone to numerical instability.  
064 Figure 3 shows the predicted softmax probability during training of both the naive method and our  
065 proposed approach. we used the prompt “*a sand wolf collapsing*”, where the model is expected to  
066 converge to the non-elastic sand constitutive model. The results show that the naive method fails to  
067 converge, as the probabilities for both elastic and non-elastic models remain close to 0.5.  
068

## 069 D REVISION OF THE ORIGINAL FIGURE 3 070

071 We have revised the original Figure 3 to illustrate how the decoder assigns different materials to  
072 different parts of the scene. Figure 1 shows the revised version of the original Figure 3.  
073

## 074 E MORE EXPERIMENT RESULTS 075

076 We provide more qualitative results in this section, including the visualization of **the same material**  
077 **with different strengths** (Figure 5), **multiple objects** (Figure 6), and **real-world scene** generated  
078 using our method (Figure 7). In some cases, we remove the verb describing motions from prompts  
079 to demonstrate the robustness of our method when using material-only prompts. For better visual-  
080 ization, please refer to the supplementary video.  
081

## 082 REFERENCES 083

- 084 Kentaro Nagasawa, Takayuki Suzuki, Ryohei Seto, Masato Okada, and Yonghao Yue. Mixing  
085 sauces: a viscosity blending model for shear thinning fluids. *ACM Trans. Graph.*, 2019.  
086  
087 Haozhe Su, Xuan Li, Tao Xue, Chenfanfu Jiang, and Mridul Aanjaneya. A generalized constitutive  
088 model for versatile mpm simulation and inverse learning with differentiable physics. *Proceedings*  
089 *of the ACM on Computer Graphics and Interactive Techniques*, 2023.  
090  
091  
092  
093  
094  
095  
096  
097  
098  
099  
100  
101  
102  
103  
104  
105  
106  
107

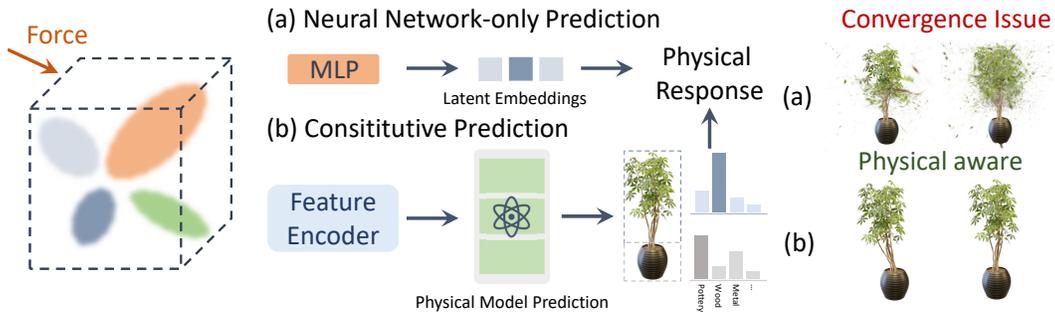


Figure 1: Constitutive Gaussian Network (Revised version).

Table 1: Quantitative results of the user study. The table shows the average ranking of each method for each scene and the overall average ranking. **Lower** values indicate better performance. The best results are highlighted in bold.

| Method       | Scene          |                  |               |               |            |               | Average       |
|--------------|----------------|------------------|---------------|---------------|------------|---------------|---------------|
|              | Swinging Ficus | Collapsing Ficus | Rubber Bear   | Sand Bear     | Jelly Cube | Water Cube    |               |
| PhysDreamer  | 3.1579         | 3.1250           | <b>2.0588</b> | 2.7692        | 2.1765     | <b>2.5385</b> | 2.6376        |
| Physics3D    | <b>2.1579</b>  | 2.7500           | 2.8824        | 3.0000        | 2.6471     | 3.0000        | 2.7396        |
| DreamPhysics | 2.4737         | 3.0000           | 2.2353        | 2.7692        | 2.3529     | 3.3077        | 2.6898        |
| <b>Ours</b>  | 2.2105         | <b>1.1250</b>    | 2.8235        | <b>1.4615</b> | 2.8235     | <b>1.1538</b> | <b>1.9330</b> |

| Method       | Scene           |               |                   |                 | Average       |
|--------------|-----------------|---------------|-------------------|-----------------|---------------|
|              | Rubber and Sand | Duck and Pile | Rubber hits Metal | Bear into Water |               |
| PhysDreamer  | 2.8000          | 3.1250        | 2.8125            | 3.0588          | 2.9491        |
| Physics3D    | 2.7333          | 2.7500        | 2.8750            | 2.8824          | 2.8102        |
| DreamPhysics | 3.0667          | 2.6875        | 3.0625            | 3.0588          | 2.9689        |
| <b>Ours</b>  | <b>1.4000</b>   | <b>1.4375</b> | <b>1.2500</b>     | <b>1.0000</b>   | <b>1.2719</b> |

## Evaluation of Physical Plausibility and Text Alignment

Figure 2: The user interface of our user study.

162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215

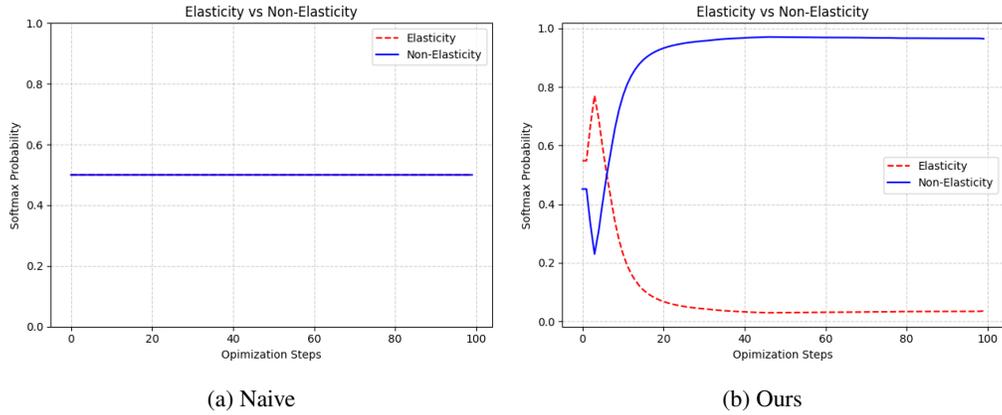
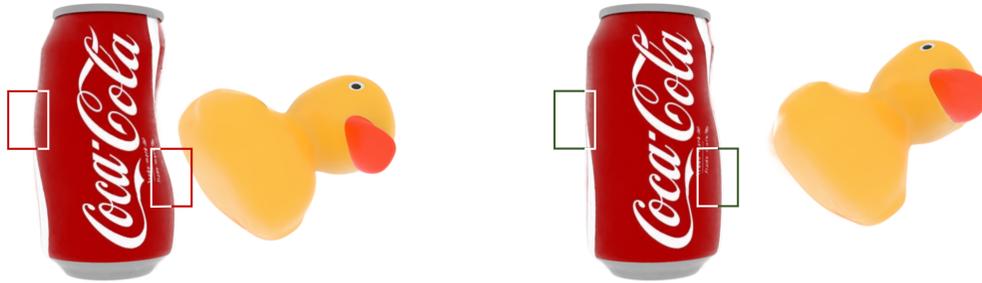


Figure 3: The predicted softmax probability of naive vector optimization and our method during the optimization process.



(a)  $25 \times 25 \times 25$  grids

(b)  $50 \times 50 \times 50$  grids



(c)  $75 \times 75 \times 75$  grids

(d)  $100 \times 100 \times 100$  grids

Figure 4: Comparison of the dynamics generated by different numbers of grids. We present the frame at the same time step for each grid size. Artifacts are highlighted in the red box.

216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269



(a) "A soft ficus swinging heavily in the wind."



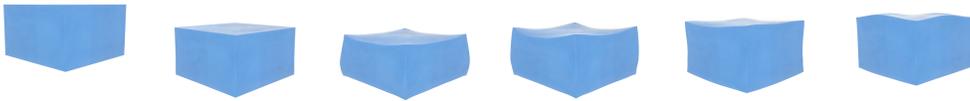
(b) "A ficus swinging slightly."



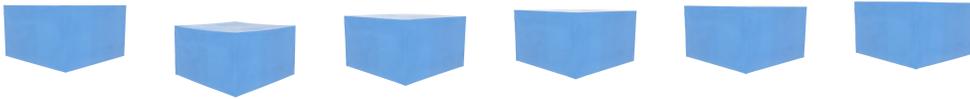
(a) "A very soft, bouncy rubber wolf."



(b) "A hard wolf."



(a) "A very tender jelly."



(b) "A firming jelly."

Figure 5: Qualitative visualizations of 3D dynamic synthesis for a single object in the same material but with different strength levels. We present the results of our method.

270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323



(a) "Different kinds of materials on the table"



(b) "Pillows falling into a basket."

Figure 6: Qualitative visualizations of 3D dynamic synthesis for more than 2 objects. We present the results of our method.



(a) "Flowers swinging gently."



(b) "A fox shaking its head."

Figure 7: Qualitative visualizations of 3D dynamic synthesis for real-world scenes. We present the results of our method.