

Neural Garment Dynamic Super-Resolution

MENG ZHANG*, School of Computer Science and Engineering, Nanjing University of Science and Technology, China
JUN LI, School of Computer Science and Engineering, Nanjing University of Science and Technology, China

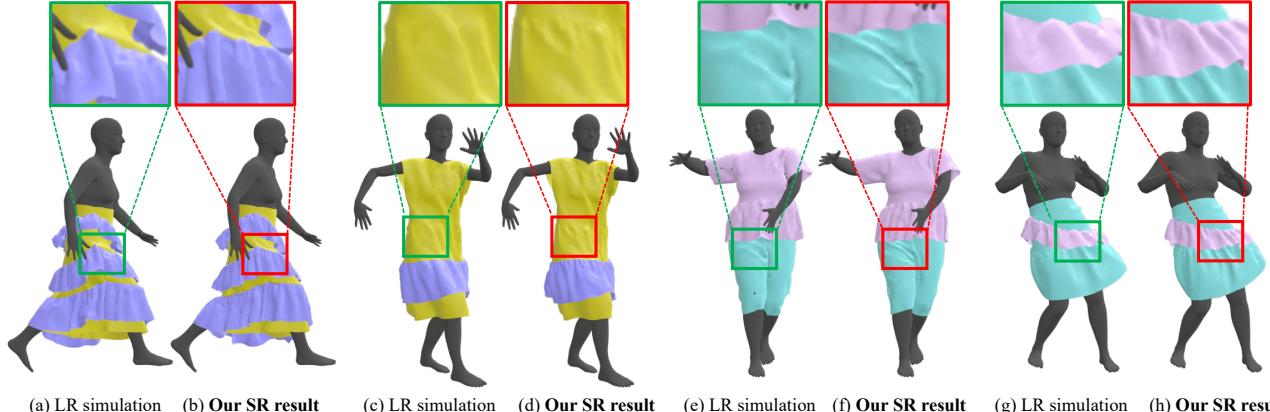


Fig. 1. Given low-resolution (LR) garment simulation as input (a,c,e,g), we present a learning-based method of *garment dynamic super resolution* to construct high-resolution garment dynamic geometry (b,d,f,h). We learn super resolution (SR) features from the coarse garment dynamics and the garment-body interactions, to produce high-resolution fine-grained and plausible wrinkle details to enhance the low-resolution geometry of various garment types, and enable the generalization ability of our method across unseen body motions (b,d,f,h), unseen body shape (d,f), and unseen garment types (f,h).

Achieving efficient, high-fidelity, high-resolution garment simulation is challenging due to its computational demands. Conversely, low-resolution garment simulation is more accessible and ideal for low-budget devices like smartphones. In this paper, we introduce a lightweight, learning-based method for garment dynamic super-resolution, designed to efficiently enhance high-resolution, high-frequency details in low-resolution garment simulations. Starting with low-resolution garment simulation and underlying body motion, we utilize a mesh-graph-net to compute super-resolution features based on coarse garment dynamics and garment-body interactions. These features are then used by a hyper-net to construct an implicit function of detailed wrinkle residuals for each coarse mesh triangle. Considering the influence of coarse garment shapes on detailed wrinkle performance, we correct the coarse garment shape and predict detailed wrinkle residuals using these implicit functions. Finally, we generate detailed high-resolution garment geometry by applying the detailed wrinkle residuals to the corrected coarse garment. Our method enables roll-out prediction by iteratively using its predictions as input for subsequent frames, producing fine-grained wrinkle details to enhance the low-resolution simulation. Despite training on a small dataset, our network robustly generalizes to different body shapes, motions, and garment types not present in the training data. We demonstrate significant improvements over state-of-the-art alternatives, particularly in enhancing the quality of high-frequency, fine-grained wrinkle

*M. Zhang is the corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SA Conference Papers '24, December 3–6, 2024, Tokyo, Japan

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1131-2/24/12...\$15.00
<https://doi.org/10.1145/3680528.3687610>

details. Code and data is released in <https://github.com/MengZephyr/Neural-Garment-Dynamic-Super-resolution/>

CCS Concepts: • Computing methodologies → Physical simulation; Mesh geometry models; • Networks;

Additional Key Words and Phrases: garment dynamics, super resolution, high frequency detail, animation, generalization

ACM Reference Format:

Meng Zhang and Jun Li. 2024. Neural Garment Dynamic Super-Resolution. In *SIGGRAPH Asia 2024 Conference Papers (SA Conference Papers '24)*, December 3–6, 2024, Tokyo, Japan. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3680528.3687610>

1 INTRODUCTION

Realistic garment behavior is crucial across various applications, including virtual try-on, 3D character design for games, movies, and VR/AR. To reproduce real garments in virtual world, physics-based simulation of garment dynamics has been extensively studied in the graphic community for decades, leading to substantial and promising progress [Li et al. 2021, 2023a]. One key challenge here is achieving high-resolution results and reproducing the intricate wrinkle details caused by complex human-garment interactions, which are essential for high visual quality. Since garments are typically represented as meshes in simulations, increasing resolution helps to model garment geometry with fine details [Wang 2021]. However, this approach results in computationally expensive simulations and difficulties in efficient data transmission and storage for data sequences. Although advancements in graphics hardware have significantly improved the run-time efficiency and fidelity of physics-based simulations, these improvements are not designed for low-budget application scenarios.

Recently, deep-learning-based algorithms have emerged as efficient alternatives for producing garment dynamics, owing to their relatively lower computational complexity. A popular approach involves decomposing garment deformation into high- and low-frequency components [Bertiche et al. 2022; Pan et al. 2022; Patel et al. 2020; Santesteban et al. 2019]. Several studies further enhance this method by using normal maps as intermediate representations to augment up-sampled low-frequency garment dynamic meshes [Lahner et al. 2018; Zhang et al. 2021a]. Halimi et al. [2023] and Chen et al. [2023] leverage mesh-based graph neural networks to refine the high-resolution meshes, up-sampled from coarse garments, by adding fine wrinkle details. However, these methods still encounter challenges related to the time consumption and geometric complexity of processing deformations on high-resolution garment meshes.

In this work, we formulate efficiently producing high-fidelity high-resolution garment geometry as a super-resolution task, aimed at either compressing high-resolution garment data or speeding up simulations for fine wrinkle details. This involves reconstructing high-frequency wrinkle details based on the coarse garment geometry, as low-resolution simulation is easily accessible and suitable for the application requirement of low-budget scenarios. To this end, we propose neural garment dynamic super-resolution (*GDSR*). Using low-resolution garment simulations as input, our method processes dynamic garment geometry from low to high resolution in an end-to-end manner, bypassing the need for intermediate normal maps used in other approaches (e.g., [Lahner et al. 2018; Zhang et al. 2021a]), which introduce additional computation for normal-guided deformation to generate high-resolution geometry.

Specifically, *GDSR* uses a mesh-graph-net to extract super resolution features from the coarse garment dynamics, as well as the garment interactions both with the body and among fabric layers. It further exploits the super-resolution features, using a hyper-net to construct continuous up-sampling fields to generate fine detailed high-resolution garment geometry. Considering the impact of coarse garment shape on the detail wrinkle performance, *GDSR* effectively corrects the low-resolution garment geometry, by decoding the super-resolution features, while simultaneously predicting high-resolution wrinkle details. To ensure temporal coherence in wrinkle details, *GDSR* performs the iterative roll-out prediction with the resultant corrected coarse garment as a historical reference.

Our method generalizes across various motion dynamics, body shapes, and garment types, even those not present in the training data. We demonstrate this generalization ability in Section 5. Comparing our approach to [Halimi et al. 2023; Zhang et al. 2021a], we show that our method enhances garment dynamics with more vivid geometric wrinkle details. The *GDSR* network is lightweight, with a storage size of only 65MB, independent of the garment mesh size. It achieves high running efficiency, potentially enabling fine-detailed high-resolution garment simulation in real-time. Computation times are discussed in Section 5.2 and the supplementary material.

Our key contributions are:

- *dynamic super-resolution features* to encode low-resolution garment dynamics and garment-body interactions.

- *wrinkle detail inference* to synthesize high-resolution wrinkle details while correcting low-resolution garment shape.
- *generalization* across various body motions, body shapes, and garment types.

2 RELATED WORK

Physics-based optimization. Physics-based garment simulation has been a fascinating topic [Baraff and Witkin 1998; Choi and Ko 2002, 2005; Cirio et al. 2014; Kim and Eberle 2020; Liu et al. 2013; Long et al. 2011; Nealen et al. 2006] in the graphics community for decades. But computation cost and stability concern is still the major issue, especially in the high-resolution garment simulation. Alternative approaches have been explored, to enhance fine wrinkle details by optimizing the coarse garment geometry based on physics constraint [Bridson et al. 2005; Müller and Chentanez 2010], operating on the garment mesh with stretch tensors [Chen et al. 2021a; Gillette et al. 2015; Li et al. 2018; Rohmer et al. 2010], quadratic shape functions [Rémillard and Kry 2013], grid structure [Wang 2021], and customized hierarchical prolongation operators [Zhang et al. 2023]. Notably, Wang et al. [2021] achieved detail synthesis at sub-millimeter levels based on a computing environment equipped with powerful GPUs. However, these approaches still remain computationally intensive.

Data-driven synthesis. Data-driven techniques are naturally explored to synthesize thin shell deformations and wrinkle details [Fröhlich and Botsch 2011; Grinspun et al. 2003]. Example-based method is applied to synthesize wrinkle details by blending the nearest wrinkle samples, searching based on body features that affect the garment [Wang et al. 2010; Xu et al. 2014], or garment geometry local features [Feng et al. 2010; Zurdo et al. 2012]. DRAPE [Guan et al. 2012] learns linear subspace models of garment deformations with respect to the body pose variation. The following work [Hahn et al. 2014] uses adaptive bases to achieve richer wrinkle details. Kavan et al. [2011] learned linear up-sampling operators, however, which is limited to the case of simple coarse cloth. Our work is loosely related to the idea of learning up-sampling operator. In contrast to previous works, we achieve efficient wrinkle detail enhancement on complex stylized garment types.

Neural-based prediction. Recent research focuses on using neural networks for predicting garment deformations, bridging garment dynamics and body motions [Bertiche et al. 2022; D. Li et al. 2022; Gundogdu et al. 2019; Li et al. 2023b; Pan et al. 2022; Wang et al. 2019], where many studies rely on the SMPL model [Loper et al. 2015]. A common approach decomposes garment deformation into low- and high-frequency components, with low-frequency driven by static body poses and high-frequency as displacement predictions. Mesh-based graph networks [Pfaff et al. 2020] and their successors [Grigorev et al. 2023; Libao et al. 2023] use physics-based constraints for improved accuracy. UV parameterization in 2D space is also explored [Shao et al. 2023; Shen et al. 2020; Zhang et al. 2022, 2021b]. Despite advancements, these methods struggle with high-resolution garment deformations due to time consumption and geometric complexity. Recent efforts include untangling multi-layered garments via untangled fields [Santesteban et al. 2022] and GNN-based models

for fabric layer interactions [Lee et al. 2023], though they remain limited to static body poses.

Since expensive the high-resolution garment simulation is, several research works mainly focus on generating high-resolution wrinkle details to enhance the coarse garment geometry. Later et al. [2018] employed a conditional generative adversarial network to add high-frequency details to the normal maps of the garments. Zhang et al. [2021a] cast the detail enhancement task as a style transfer of the normal maps. Both of those works adopt the normal map to represent the garment geometry at a cost of time, that spends on normal map rendering and normal guided geometry deformation. The architecture of mesh-based graph net is applied in [Halimi et al. 2023] to optimize the high-resolution wrinkle detail geometry, by iterative running the graph network to solve the physically-based potential energy. Chen et al. [2023] trained garment-specific neural networks to translate the detailed garment dynamics from the coarse garment sequence. To our knowledge, none of those previous works produce garments with high-frequency details as the pleated lace and skirt shown in Figure 1.

Super resolution. Techniques of deep learning have made significant progress for the super-resolution task of image and video [Chen et al. 2021b; Chu et al. 2020; Haris et al. 2019; Ledig et al. 2017; Sajjadi et al. 2018; Zhang et al. 2018]. 2D convolution neural networks (CNN) are commonly used, which are naturally compatible with the data structure of images, and help achieve a hierarchical feature encoding and decoding. Chen et al. [2021c] adopted the CNN-based super-resolution technique to synthesize cloth wrinkles by recording multiple garment geometry features in 2D image space, where the detail representation capability is limited by the image size. In contrast, we feature our garment dynamic super-resolution in the original low-resolution garment mesh with no conversion between the geometry and image space, and construct continuous up-sampling fields locally for coarse mesh triangles to produce high-resolution detailed garment geometry. Additionally, different from the image/video-based super-resolution operating on static low-resolution content, our work further explores to correct the coarse content (garment), jointly with the high-frequency detail reconstruction.

3 OVERVIEW

Starting from the low-resolution garment dynamics $\{C_t\}$ and the underlying body motions $\{B_t\}$, our *neural garment dynamic super-resolution (GDSR)* adeptly generates a high-resolution garment geometry sequence G_t imbued with vibrant fine wrinkle details.

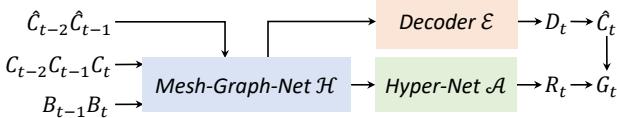


Fig. 2. **GDSR overview.** With the low-resolution garment dynamics C_{t-2} , C_{t-1} and C_t , the underlying body motion B_{t-1} , B_t , and the previous prediction \hat{C}_{t-2} , \hat{C}_{t-1} , our *GDSR* predicts the displacements D_t to compute the corrected coarse garment shape \hat{C}_t , and jointly generates detail wrinkle residual R to produce the final high-resolution garment geometry G_t .

We introduce *GDSR* to predict the displacement D_t to correct the coarse garment shape of the low-resolution simulation, and meanwhile, synthesize wrinkle detail residuals R_t . As shown in Figure 2, the network of *GDSR* comprises three integral network blocks: Mesh-Graph-Net \mathcal{H} (refer to Section 4.2), Decoder \mathcal{E} (refer to Section 4.3), and Mesh-Hyper-Net \mathcal{A} (refer to Section 4.4). The super-resolution features for all low-resolution garment vertices are computed using \mathcal{H} . Recognizing the significant impact of coarse garment shape on wrinkle detail performance, we rectify the coarse garment geometry at the low-resolution scale by using \mathcal{E} to decode the displacement D_t . We add the displacement d_t^i to each coarse garment vertex c_t^i as $\hat{c}_t^i = c_t^i + d_t^i$, resulting in the corrected coarse garment shape \hat{C}_t . Simultaneously, Mesh-Hyper-Net \mathcal{A} constructs continuous up-sampling fields for triangles of the coarse garment mesh, where wrinkle detail residuals R_t are decoded in terms of the local coordinate of each triangle. Finally, the high-resolution, finely detailed garment G_t is computed by combining the up-sampling of the \hat{C}_t with the wrinkle detail residuals R_t . To account for the historical dynamics influencing the current state of the garment, we run our method in a roll-out prediction manner, incorporating \hat{C}_{t-1} and \hat{C}_{t-2} down-sampled from previous predictions. The detailed architecture, illustrated in Figure 3, is described in the next section.

4 ALGORITHM

Given the high-resolution G and the low-resolution C of a garment, we observe a commonality in their semantic 2D patterns within the UV space. Leveraging the alignment in the 2D UV space, we can establish a correspondence between the low- and high-resolution garment meshes. Each interior point within a mesh triangle f can be expressed as a sampling S through a linear combination of the barycentric coordinates (u, v) and the three triangle vertices c^0, c^1, c^2 . Consequently, the up-sampling of a high-resolution garment mesh from the low-resolution C involves representing every vertex of the up-sampled mesh as:

$$S[C, f, u, v] = (1 - u - v)c^0 + uc^1 + vc^2, \quad (1)$$

where f is derived from the low-resolution mesh C . This geometric parametrization is established in the canonical space once a garment is set up and remains universally applicable across various sequences of garment dynamics.

4.1 Graph Feature Initialization.

Given the low-resolution garment dynamics C_t , C_{t-1} , C_{t-2} , and the underlying body motion B_t , B_{t-1} , we conceptualize a graph (Q_t, E_t) at current time t , where we assign the coarse garment node features \hat{q}^i and edge features \hat{e}^{ij} to represent garments based on the geometric mesh topology, where i and j are garment nodes connected by an edge. For multi-layered garments, we introduce additional edge features \hat{e}^{il} to describe the interaction between two garment nodes, i and l , each node located in a different cloth layer.

Node feature initialization. Considering the information of garment's geometry and dynamics, we initialize the graph node feature \hat{q}_t^i with the vertex normal n_t^i , velocity \dot{c}_t^i and acceleration \ddot{c}_t^i of the coarse garment vertex c_t^i . Additionally, in contrast to the

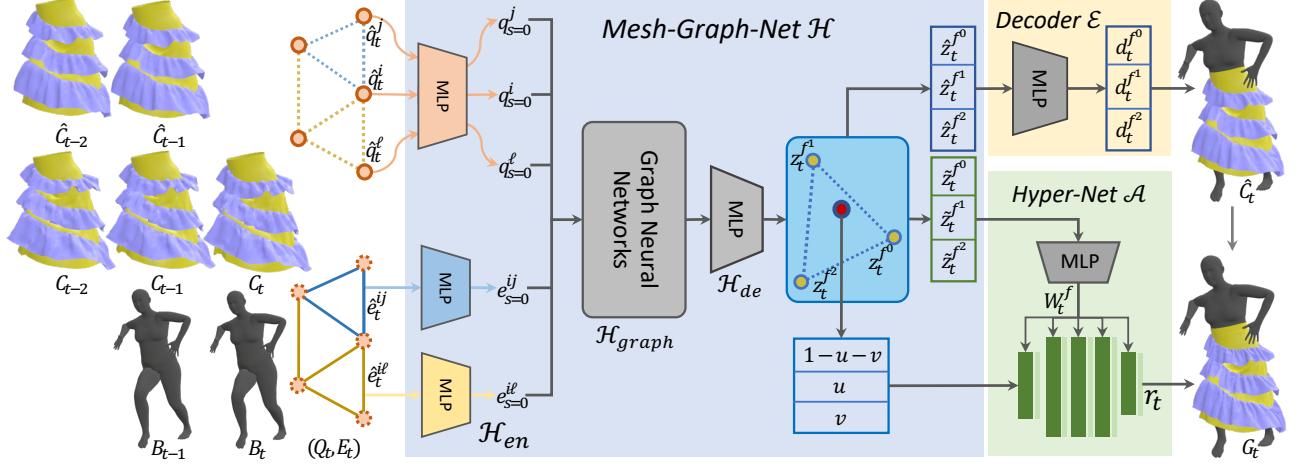


Fig. 3. Garment dynamic super-resolution architecture. Our method operates in a roll-out prediction manner. We represent the inputs of low-resolution simulations, body motions, and downsampled previous predictions as a graph with node features Q_t and edge features E_t . First, we employ *Mesh-Graph-Net* \mathcal{H} to compute super-resolution features $z_t := [\hat{z}_t | \tilde{z}_t]$ for coarse garment mesh. Considering the impact of coarse garment shape on detail wrinkle performance, *Decoder* \mathcal{E} decodes features \hat{z}_t to correct the low-resolution garment shape, resulting in \hat{C}_t . Simultaneously, *Hyper-Net* \mathcal{A} exploits super-resolution features \tilde{z}_t to construct an implicit function W_t^f of detail wrinkle residuals for each mesh triangle. Finally, we generate the high-resolution garment G_t by applying the detail wrinkle residuals on the corrected coarse garment.

normal-based methods [Lahner et al. 2018; Zhang et al. 2021a] of detail enhancement, we observe that the garment-and-body interactions significantly influence the generation of garment folds and fine-grained wrinkles. Thus, we detect the garment-and-body interaction by searching the nearest body vertex b_t for each coarse garment vertex c_t^i using the signed distance field estimated through ray casting [Roth 1982]. We define the interaction vector p_t^i as:

$$p_t^i := \text{ReLU}[\sigma_b - (c_t^i - b_t) \cdot n_t^b], \quad (2)$$

where n_t^b is the normal of body vertex b_t , ReLU is the rectified linear unit function [Nair and Hinton 2010], and σ_b is a predefined constant value. We also consider the interaction between the coarse garment vertex c_{t-1}^i and the body vertex b_t which potentially drives the garment deformation at time $t-1$, and define the interaction vector \hat{p}_t^i as:

$$\hat{p}_t^i := \text{ReLU}[\sigma_b - (c_{t-1}^i - b_t) \cdot n_t^b]. \quad (3)$$

Considering the history state of the garment dynamics influencing on the current state of the garment, we propagate the displacement $d_{t-1}^i = \hat{c}_{t-1}^i - c_{t-1}^i$ and the velocity $\dot{\hat{c}}_{t-1}^i$ involving the coarse garment geometries \hat{C}_{t-1} and \hat{C}_{t-2} , down-sampled from the history high-resolution garment predictions, to perform the iterative prediction at current time t . Consequently, we initialize the coarse garment node feature \hat{q}_t^i as:

$$\hat{q}_t^i := [n_t^i, \hat{c}_t^i, \dot{\hat{c}}_t^i, p_t^i, \|p_t^i\|, \hat{p}_t^i, \|\hat{p}_t^i\|, d_{t-1}^i, \dot{\hat{c}}_{t-1}^i], \quad (4)$$

where $\|p_t^i\|$ and $\|\hat{p}_t^i\|$ is the length of p_t^i and \hat{p}_t^i .

Edge feature initialization. The garment edge feature e_t^{ij} between node i and node j , is defined by the edge vector $c^j - c^i$ and edge length $\|c^j - c^i\|$ at the current and previous time $t, t-1$, and the

edge length $\|c_r^i - c_r^j\|$ at the rest state o :

$$\hat{e}_t^{ij} := [c_t^j - c_t^i, \|c_t^j - c_t^i\|, c_{t-1}^j - c_{t-1}^i, \|c_{t-1}^j - c_{t-1}^i\|, \|c_o^j - c_o^i\|]. \quad (5)$$

For garment nodes i and ℓ located in different cloth layers, we incorporate an extra edge feature $\hat{e}_t^{i\ell}$ when $\|c_t^i - c_t^\ell\| < \sigma_\ell$. This feature is defined similarly to the mesh topology-based edge feature \hat{e}_t^{ij} , but replaces the rest state edge length $\|c_o^i - c_o^\ell\|$ with the predefined constant value σ_ℓ .

4.2 Garment Super Resolution Feature Using \mathcal{H}

With the initialized graph as input, we train *Mesh-Graph-Net* \mathcal{H} to encode super-resolution features on the low-resolution garment mesh through three successive phases: Feature Encoder \mathcal{H}_{en} , Graph Process \mathcal{H}_{graph} and Feature Decoder \mathcal{H}_{de} .

Feature Encoder \mathcal{H}_{en} . With the initialized graph (Q, E) (temporal notation t is omitted for a brevity of description), we employ three blocks of multi-layer perceptrons (MLP) within the Feature Encoder \mathcal{H}_{en} to project the features of garment nodes $\{\hat{q}_t^i\}$, garment edges $\{\hat{e}_t^{ij}\}$, and cloth layered edges $\hat{e}_t^{i\ell}$ into three respective latent spaces, denoted as $\{q_{s=0}^i\}$, $\{e_{s=0}^{ij}\}$, and $\{e_{s=0}^{i\ell}\}$, each having the same dimension. These encoded features serve as the initialized state before undergoing the subsequent graph processing network. In our experiments, we have chosen a graph latent space dimension of 128.

Graph Process \mathcal{H}_{graph} . We adopt a fixed set of steps, S , for message passing in the Graph Process network \mathcal{H}_{graph} . In each block, we leverage MLPs with residual connection to update the graph features. Following [Pfaff et al. 2020], we first update the edge features by passing the node features to their connected edges:

$$e_s^{ij} = \mathcal{F}_{q \rightarrow e}[e_{s-1}^{ij}, q_{s-1}^i, q_{s-1}^j] + e_{s-1}^{ij}, \quad (6)$$

where $\mathcal{F}_{q \rightarrow e}$ represents an MLP. Second, we proceed to update the node features by passing the edges features to their corresponding adjacent nodes:

$$q_s^i = \mathcal{F}_{e \rightarrow q}[q_{s-1}^i, \sum_j e_s^{ij}, \sum_\ell e_s^{i\ell}] + q_{s-1}^i, \quad (7)$$

where $\mathcal{F}_{e \rightarrow q}$ denotes another MLP. If there is no layered edge connection with garment node i , we regard $\sum_\ell e_s^{i\ell}$ to be 0. After completing S steps of message passing within \mathcal{H}_{graph} , we obtain the output node features denoted as $\{q_S^i\}$.

To expedite the propagation of features across the garment mesh, we further extend Graph Process Network \mathcal{H}_{graph} to incorporate hierarchical message passing, as proposed in [Grigorev et al. 2023]. In our experiments, we set the number of message passing steps S to be 6, and establish a 2-ring hierarchical graph construction. For more details of the hierarchical graph process, please refer to [Grigorev et al. 2023].

Feature Decoder \mathcal{H}_{de} . Utilizing the node features $\{q_S^i\}$ as input, we apply Feature Decoder \mathcal{H}_{de} to decode the ultimate super-resolution features $\{z^i\}$ for every vertex on the low-resolution garment mesh. This is achieved by employing an MLP to increase the dimensionality of the node features from 128 to 640. Subsequently, we segment the super resolution feature $z^i = [\hat{z}^i | \tilde{z}^i]$ into displacement super-resolution feature \hat{z}^i and detail super-resolution feature \tilde{z}^i with the dimensions of 128 and 512, respectively.

4.3 Coarse Garment Correction Using \mathcal{E}

To rectify the coarse garment shape, we employ a Decoder \mathcal{E} to compute the displacements D with respect to the vertices of low-resolution mesh. Using three displacement super-resolution features $\hat{z}^{f^0}, \hat{z}^{f^1}, \hat{z}^{f^2}$ from a low-resolution mesh triangle, we decode the displacements of the three triangle vertices under the local coordinate H^f :

$$[\tilde{d}^{f^0}, \tilde{d}^{f^1}, \tilde{d}^{f^2}] = \mathcal{E}[\hat{z}^{f^0}, \hat{z}^{f^1}, \hat{z}^{f^2}], \quad (8)$$

where \mathcal{E} is an MLP. Then we compute the displacement vector d of a low-resolution garment vertex c , by averaging the displacements of the corresponding vertices f^c in the adjacent triangles f :

$$d = \frac{1}{|f|} \sum_f H^f[\tilde{d}^{f^c}], \quad (9)$$

where $|f|$ denotes the number of adjacent triangles, H^f is the local coordinate of the triangle f in the coarse garment C . Finally, we compute the corrected low-resolution garment mesh \hat{C} by adding the displacement D to the coarse garment input C , that is: $\hat{C} = C + D$.

4.4 Detail Wrinkle Residual Using \mathcal{A}

We locally construct the continuous up-sampling detail wrinkle field \mathcal{W}^f for each coarse garment triangle f . Employing a hyper-network \mathcal{A} , we decode the parameters $\{w\}^f$ of the implicit function \mathcal{W}^f using the detail super-resolution features $\hat{z}^{f^0}, \hat{z}^{f^1}, \hat{z}^{f^2}$ corresponding to the three vertices of the coarse mesh triangle f :

$$\{w\}^f = \mathcal{A}[\hat{z}^{f^0}, \hat{z}^{f^1}, \hat{z}^{f^2}], \quad (10)$$

where \mathcal{A} is an MLP network. Each up-sample vertex k of the high-resolution mesh possesses a barycentric coordinate (u^k, v^k) with respect to its corresponding triangle f in the coarse mesh topology. We apply the constructed \mathcal{W}^f to decode the fine detail wrinkle residual r^k with the barycentric coordinate (u^k, v^k) as input:

$$r^k = \mathcal{W}^f[1 - u^k - v^k, u^k, v^k]. \quad (11)$$

Then, we compute the position of any high-resolution garment mesh vertex g^k as:

$$g^k = \mathcal{S}[\hat{C}, f, u^k, v^k] + \hat{H}^f[r^k], \quad (12)$$

where \hat{C} is the corrected coarse garment shape as described in Section 4.3, and \hat{H}^f denotes the local coordinate of the triangle f in \hat{C} .

Specifically, \mathcal{W}^f is implemented as a complex MLP. We configure the number of layers in \mathcal{W}^f to be 5, with perceptrons x activated by the complex Gabor wavelet activation function $\text{Wire}(x, \omega_0, \alpha_0)$ except for the last layer of output. This choice exhibits superior performance in high-frequency feature reconstruction. We set the frequency control ω_0 and the spread control α_0 to be 5 and 10, respectively. For additional insights, please refer to [Saragadam et al. 2023].

4.5 Loss Function

As L1 loss demonstrates superior performance in super resolution tasks [Zhao et al. 2017], we train our *GDSR* using objective L1 loss functions to ensure the accuracy of both coarse garment correction and wrinkle detail generation.

Coarse garment correction. We first consider L1 loss with respect to the corrected low-resolution garment geometry \hat{C}_t and its normal \hat{N}_t to enforce a similarity to the ground truth \hat{C}_t^* , as follows:

$$L_{c_geo}(\hat{C}_t) = \|\hat{C}_t - \hat{C}_t^*\|_1 + 0.01\|\hat{N}_t - \hat{N}_t^*\|_1, \quad (13)$$

where \hat{C}_t^* is predefined before training by down-sample the ground truth of high-resolution garment mesh G_t , and \hat{N}_t^* denotes the normal of \hat{C}_t^* . Additionally, we enforce the stretching and shearing of \hat{C}_t to be similar as the ground truth \hat{C}_t^* . We approximate the mapping matrix:

$$w(\hat{C}_t) := \left(w_u(\hat{C}_t) | w_v(\hat{C}_t) \right)_{3 \times 2}, \quad (14)$$

over all triangles from the 2D cloth plane coordinate (u, v) to the garment \hat{C}_t in the world space. For more insights of the mapping matrix $w(\hat{C}_t)$, please refer to [Baraff and Witkin 2023; Kim and Eberle 2020]. Thus, we define the objective loss function as:

$$L_{c_def}(\hat{C}_t) = 10^4 \|\Delta \hat{C}_t\|^2 + 0.1 \|\theta(\hat{C}_t) - \theta(\hat{C}_t^*)\| \quad (15)$$

$$+ \|\|w_u(\hat{C}_t)\| - \|w_u(\hat{C}_t^*)\|\|_1 \quad (16)$$

$$+ \|\|w_v(\hat{C}_t)\| - \|w_v(\hat{C}_t^*)\|\|_1 \quad (17)$$

$$+ \|w_u(\hat{C}_t) \cdot w_v(\hat{C}_t) - w_u(\hat{C}_t^*) \cdot w_v(\hat{C}_t^*)\|_1, \quad (18)$$

where Δ is the Laplacian operator and θ denotes dihedral angles between adjacent mesh triangles of \hat{C} . The final loss function for correcting the low-resolution garment mesh is a linear combination of the terms:

$$L_{coa}(\hat{C}_t) = L_{c_geo}(\hat{C}_t) + 0.01 L_{c_def}(\hat{C}_t). \quad (19)$$

Patch-based learning wrinkle detail generation. We adopt a patch-based strategy to learn wrinkle detail generation. The high-resolution patches, denoted as P , are cropped based on the garment's 2D UV texture image, which maintains a consistent unit scale between pixels and 3D space, ensuring each mesh triangle corresponds to at least one pixel. As the normal map \mathcal{N} sensitively describes wrinkle details, we render the normal map for both the predicted patch P_t and the ground truth patch P_t^* of high-resolution garment simulation. We define a reconstruction loss concerning vertex positions and multi-layer features of the pre-trained VGG network [Simonyan and Zisserman 2015]:

$$L_{w_geo}(P_t) = \|P_t - P_t^*\|_1 + 0.01 \sum_{\ell} \|\text{VGG}^\ell[\mathcal{N}_t^P] - \text{VGG}^\ell[\mathcal{N}_t^{P^*}]\|_1. \quad (20)$$

As seen in the inset figure, utilizing VGG features (/w VGG) helps capture more reasonable high-frequency wrinkle patterns, in comparison to the result without using VGG (/wo VGG). To constraint the deformation of each triangle on the wrinkle detailed patches, we derive $w_u(P_t)$ and $w_v(P_t)$ from the estimation of mapping matrix $w(P_t)$ to define the deformation loss as:

$$L_{w_def}(P_t) = \|\Delta P_t\|^2 + \||w_u(P_t)| - |w_u(P_t^*)|\|_1 \quad (21)$$

$$+ \||w_v(P_t)| - |w_v(P_t^*)|\|_1 \quad (22)$$

$$+ \|w_u(P_t) \cdot w_v(P_t) - w_u(P_t^*) \cdot w_v(P_t^*)\|_1. \quad (23)$$

The final loss function for the wrinkle detail generation is defined as:

$$L_{wri} = \frac{1}{|P|} \sum_P L_{w_geo}(P_t) + 0.01 L_{w_def}(P_t), \quad (24)$$

where $|P|$ is the number of sampled patches.

Training. We jointly train the three network blocks in *GDSR*, which consist of Mesh-Graph-Net \mathcal{H} , Decoder \mathcal{E} and Mesh-HyperNet \mathcal{A} , to enhance high-resolution wrinkle details while correcting the low-resolution garment geometry. Therefore, we define the final loss function for *GDSR* as:

$$L_{GDSR} = \lambda_c L_{coa}(\hat{C}_t) + \lambda_w L_{wri}(\{P_t\}). \quad (25)$$

Initially, we set $\lambda_c = 1.9$ and $\lambda_w = 0.1$ since coarse garment correction is a prerequisite for wrinkle detail generation. During training, λ_c is decreased and λ_w is increased by 0.2 every 10 epochs until both reach 1. We then continue training *GDSR* with $\lambda_c = 1$ and $\lambda_w = 1$ until the loss converges. We use AdamW Optimizer with an initial learning rate of 1×10^{-4} , halving it every 50 epochs. Convergence typically occurs after around 300 epochs. Due to the varied mesh topology of the different garment types in our training data, we set the batch size to 1 for computing L_{coa} , while randomly cropping 8 normal patches from the high-resolution mesh for L_{wri} .

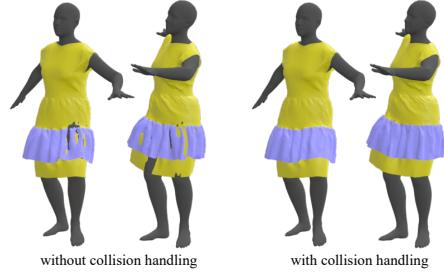
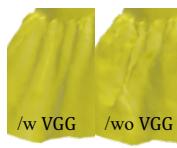


Fig. 4. **Explicit collision handling.** We detect the garment collisions with the body and that between the fabric layers at a coarse level of garment geometry, and propagate the detection to efficiently resolve the collisions for high resolution garment.

4.6 Explicit Collision Handling

We observe that garment collisions with the body and between different fabric layers are inevitable, especially in the case of unseen garments and body motions. Efficient collision handling alongside high-resolution garment dynamics is too expensive for constrained budgets. To address this, we detect potential collisions on the low-resolution mesh and propagate this information to the high-resolution garment. When running *GDSR*, we identify potential collisions between the corrected coarse garment \hat{C}_t and the body B_t . For a coarse garment vertex \hat{c}_t^i , a potential collision with a body vertex b_t is identified if $\sigma_b - (\hat{c}_t^i - b_t) \cdot n_t^b > 0$, where σ_b is a predefined constant, and n_t^b is the tangent plane normal vector of b_t . Subsequently, a high-resolution garment vertex g_t^k , where \hat{c}_i is the nearest low-resolution garment vertex to g^k in the 2D UV parameterization space, is adjusted to a new position:

$$g_t^k = g_t^k + \max[\sigma_b - (g_t^k - b_t) \cdot n_t^b, 0] n_t^b. \quad (26)$$

For a coarse garment vertex $\hat{c}_t^{\ell_e}$ in the exterior layer, we detect a collision with an interior layered garment vertex $\hat{c}t^{\ell_i}$ if $\sigma_\ell - (\hat{c}_t^{\ell_e} - \hat{c}_t^{\ell_i}) \cdot n_t^{\ell_i} > 0$ and $|\hat{c}_t^{\ell_e} - \hat{c}_t^{\ell_i}| < \gamma$, as the garment mesh is an open surface. Here σ_ℓ and γ are predefined constants. We resolve the collision between fabric layers of a high-resolution garment by adjusting the exterior layer vertex $g_t^{\ell_e}$ to account for the collision of the nearest coarse garment vertex $\hat{c}_t^{\ell_e}$ with the interior garment vertex $\hat{c}_t^{\ell_i}$:

$$g_t^{\ell_e} = g_t^{\ell_e} + \max[\sigma_\ell - (g_t^{\ell_e} - \hat{c}_t^{\ell_i}) \cdot n_t^{\ell_i}, 0] n_t^{\ell_i}. \quad (27)$$

Finally, we resolve garment collisions with the body and between fabric layers by updating the high-resolution garment G_t . The effect of this explicit collision handling is shown in Figure 4.

5 EXPERIMENTS AND RESULTS

Using two dancing sequences (swing and house) on five garment outfits (colored yellow or purple) with a fixed body shape, we train *GDSR* network with garment dynamic history states, derived from high-resolution garment ground truth. The first row in Figure 5 shows the results of our method testing on seen motions with the

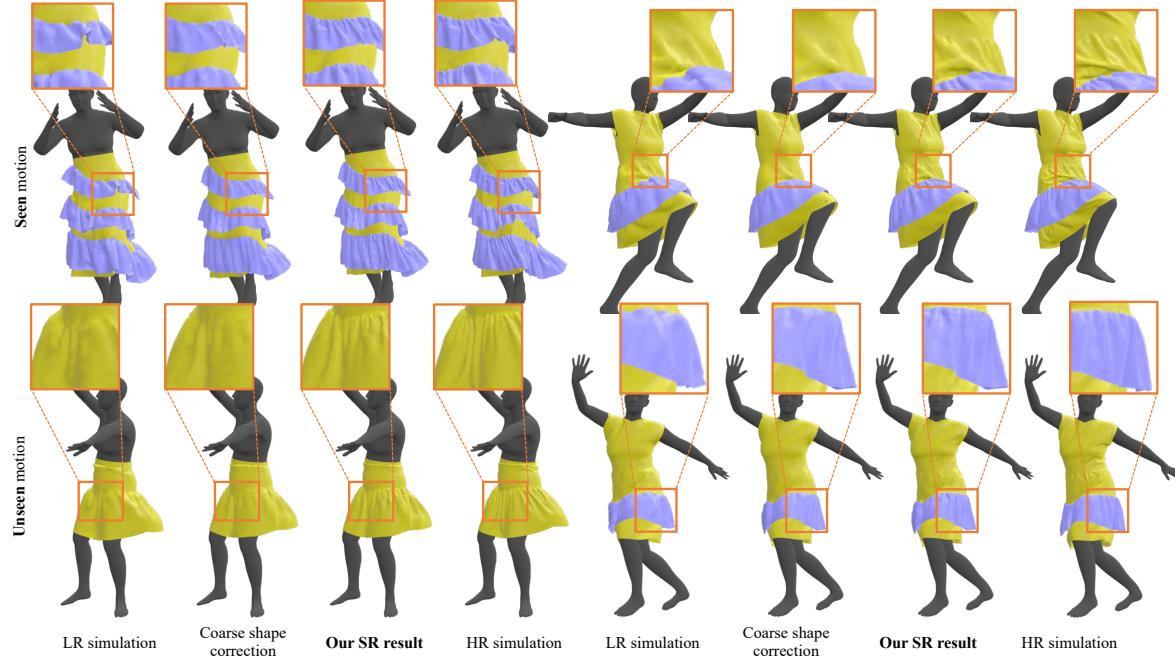


Fig. 5. We run our method by roll-out prediction on both seen (1st row) and unseen (2nd row) motions with the training garment types. For each example, we show input of low resolution simulation, results of coarse shape correction and detailed high resolution results, as well as the reference of high resolution simulation.

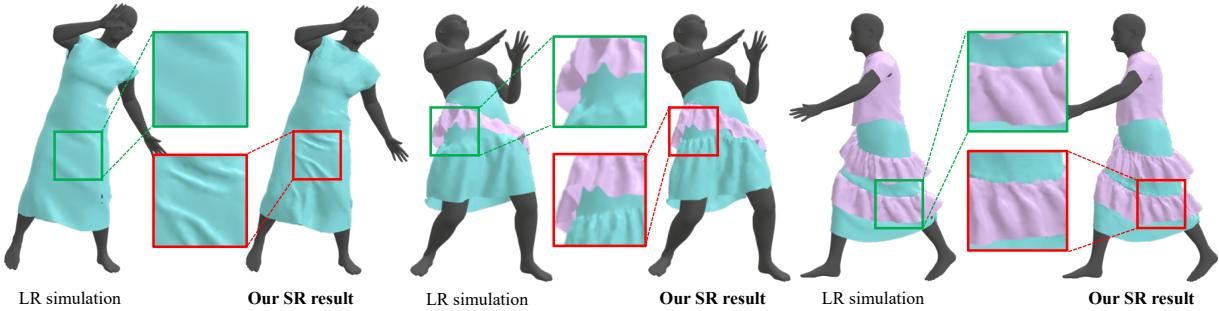


Fig. 6. We evaluate the generalization of our method on motions and garment types that are both out of the training data. Our method can synthesize fine-grained wrinkle details for either single- or multi-layer garments.

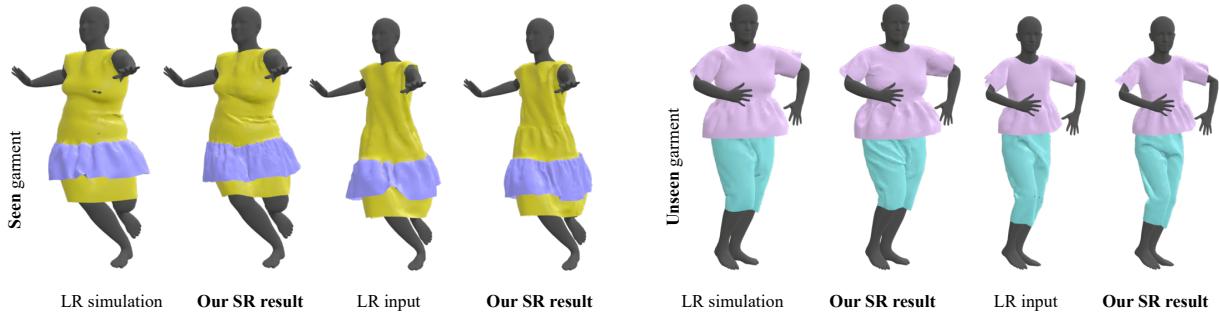


Fig. 7. We evaluate the generalization of our method across different body shapes on out-of-training motions with both seen and unseen garment types. Although trained with a fixed body shape, our method can synthesize detailed high-resolution garment geometry fitting with varying body shapes.

training garment types in a roll-out prediction manner. During iterative processing on the previous predictions and low-resolution simulation inputs, our method robustly produces fine-grained garment wrinkle details resembling high-resolution simulation references.

To test our method's generalization ability, we generate seven additional motion sequences and five garment outfits (colored teal or pink). The differences between the training and testing data are illustrated in the supplementary material. Our method successfully generalizes across various body motions, garment types, and body shapes not present in the training data. Sampled frames of our results are shown here. For a better evaluation of high-resolution garment details, please refer to the supplementary video.

Generalization to unseen motions. Instead of using body motions as global signals to drive garment dynamics, our network focuses on local garment dynamics and interactions with the underlying body to learn local displacement vectors and wrinkle detail residuals. The second row in Figure 5 demonstrates our method's generalization ability to unseen motion sequences (samba dancing and samba variant). Our method synthesizes fine-grained wrinkle details, producing wrinkle patterns similar to those in high-resolution simulations.

Generalization to unseen garments. Based on our training data, which includes multiple garment types, we generate new garments by removing or adding lace embellishments, modifying garment lengths, or combining different garments. Our method achieves remarkable generalization by using Mesh-Graph-Net, which is not restricted to any specific mesh topology, and Hyper-Net to construct continuous up-sampling fields for any local mesh triangle. Figure 6 showcases the results of our method on body motions and garment types that are not part of the training data. Our results demonstrate that our method can synthesize fine-grained wrinkle details for both single- and multi-layered garments.

Generalization to different body shapes. Our method is trained using a fixed body shape. In Figure 7, we first test its generalization to different body shapes with a garment in training data. The varying wrinkle patterns on the garment for different body shapes demonstrate that our method can synthesize detailed, high-resolution garment geometry that fits varying body shapes. Next, we design a combination of a lace t-shirt and pants in two different sizes to fit two body shapes. The results in Figure 7 show that our method performs plausibly on unseen garments with different body shapes.

Robustness of iterative roll-out prediction. Our method is trained with two separate dance motion sequences: *House* (714 frames) and *Swing* (794 frames). We evaluate the robustness of iterative roll-out prediction on walking sequences up to 1500 frames, testing on both seen and unseen garment types. As shown in Figure 8, our network uses its predictions as input for subsequent frames, producing reasonable predictions for long testing sequences. The wrinkle details synthesized by our method maintain plausible structural coherence between consecutive frames. We evaluate the stability of long roll-out prediction quantitatively in the supplemental material.

Effect of coarse shape correction. To evaluate the effect of coarse garment shape correction, we design two variants of GDSR by

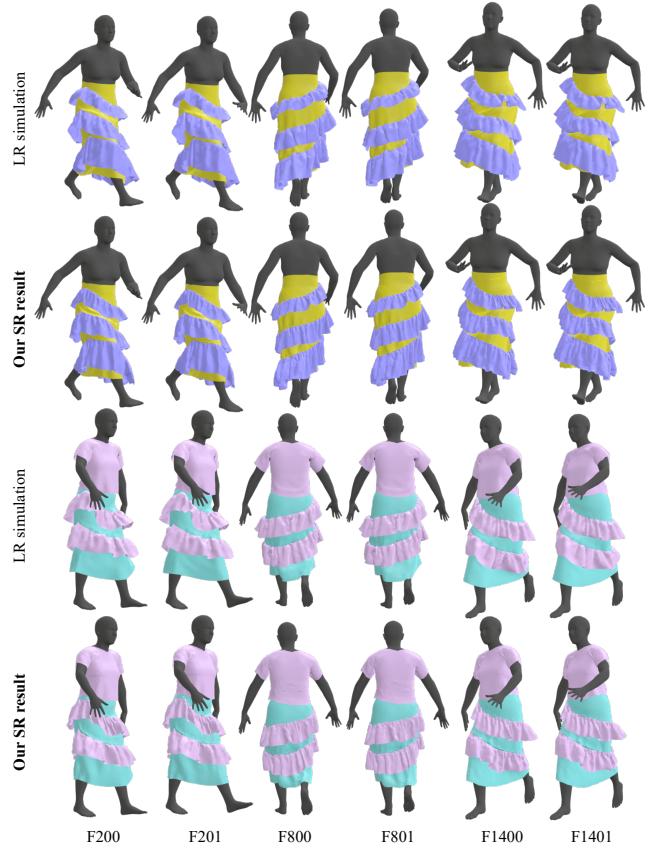


Fig. 8. We evaluate the robustness of our method under long roll-out predictions up to 1500 frames, using a seen garment (in purple and yellow) and an unseen garment (in teal and pink) with happy walking and normal walking motion sequences, respectively. Our method uses low-resolution simulation and its predictions as input for subsequent frames to produce reasonable wrinkle details while maintaining plausible temporal smoothness between frames.

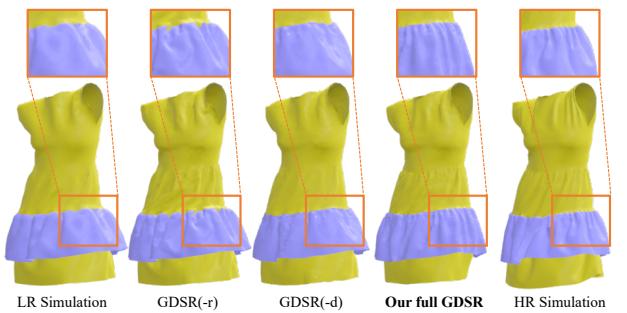


Fig. 9. To evaluate the effect of coarse garment correction, we designed two variants of GDSR by removing the *Decoder E*, named GDSR(-r) and GDSR(-d) respectively. Our full model of GDSR produces remarkably fine-grained high-resolution garment geometry with wrinkle detail patterns similar to the reference of high-resolution simulation, compared to GDSR(-r), which does not consider the influence of coarse garment shape, and GDSR(-d), which relies on roll-out prediction but without correcting the garment shape at coarse level.

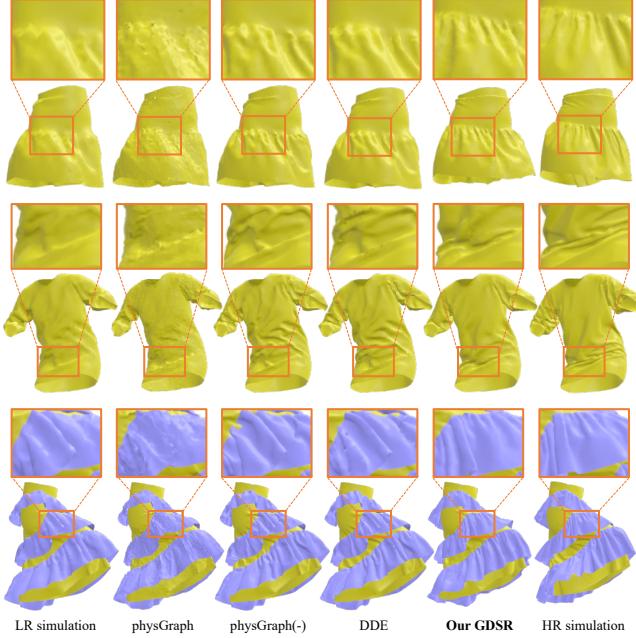


Fig. 10. Given the low-resolution simulation garment geometry, our GDSR constructs the high-resolution garment geometry by synthesizing fine detail wrinkles, applied on the corrected coarse garment mesh. Comparing to PhysGraph [Halimi et al. 2023] and DDE[Zhang et al. 2021a], GDSR captures fine-grained wrinkle patterns more similar to the reference high-resolution simulation.

removing the *Decoder* \mathcal{E} , named GDSR(-r) and GDSR(-d) respectively. Without considering the influence of coarse garment shape on the dynamics of wrinkle details, we train GDSR(-r) using a down-sampled coarse garment \hat{C} from the high-resolution simulation as input. This is done by initializing the input of node features with vectors of normal, velocity, and acceleration computed on the down-sampled geometry \hat{C} , and setting the displacements d_{t-1} to 0. As shown in Figure 9, with low-resolution simulation as input, GDSR(-r) only emphasizes the original wrinkle details present on the coarse garment geometry. Meanwhile, GDSR(-d) involves the history of high-resolution detail predictions in its subsequent predictions but, without the Decoder \mathcal{E} to correct the coarse garment shape at the coarse level, it fails to capture clear high-frequency details. In comparison, our full model of GDSR excels in capturing fine-grained wrinkle detail patterns similar to the reference of high-resolution simulation with coarse garment shape correction.

5.1 Comparisons

We evaluate our approach by comparing it with two recent techniques for enhancing garment details: DDE [Zhang et al. 2021a] and PhysGraph [Halimi et al. 2023]. DDE is an image-based method that enhances wrinkles by applying style transfer to normal images. We train DDE using downsampled garment geometry from high-resolution simulations. However, some high-frequency details are not captured because DDE does not consider the impact of coarse garment shape on garment wrinkles, and some details

are smoothed out by the subsequent normal-guided deformation. PhysGraph uses a mesh-based graph network but operates on high-resolution garment meshes up-sampled from low-resolution simulations. PhysGraph fails to converge effectively, resulting in noisy garment meshes when trained with the same data as GDSR. To address this, we train a variant (PhysGraph(-)) using down-sampled garment geometry from high-resolution simulations, similar to DDE. Without correcting the coarse garment shape, PhysGraph(-) achieves only comparable quality to DDE. In contrast, our GDSR captures fine wrinkle details that closely resemble the high-resolution simulation reference (see Figure 10). We report the quantitative comparisons with style improvement score in the supplemental material.

5.2 Running Efficiency

Our lightweight network, with a storage size of 65MB, is potentially compatible with consumer devices like smartphones. Despite unoptimized PyTorch code and low-resolution simulation efficiency, we believe our method has the potential to enable real-time, high-resolution clothing simulations. For instance, it takes only 0.115 seconds to synthesize one frame of detailed high-resolution garments with 43,257 vertices from coarse geometry with 5,010 vertices, including 0.083 seconds for low-resolution simulation in Marvelous Designer (MD). The supplemental material provides more details on computation time evaluation of GDSR and comparisons with physics-based simulations in MD and two baseline methods [Halimi et al. 2023; Zhang et al. 2021a].

6 LIMITATIONS & FUTURE WORK

Our work has several limitations. As shown in Figure 11, GDSR fails to capture wrinkle details influenced by fabric friction, as it lacks graph features to describe this friction. Additionally, due to low-resolution collision detection, some collisions remain unresolved, especially in multi-layered garments with intersecting vertices beyond the predefined threshold. Existing works [Lee et al. 2023; Santesteban et al. 2022] untangle the intersection of multi-layered garments specific for static body poses but not for dynamic motions. Future research could explore collision resolution for dynamic multi-layered garments using signed distance fields for local meshes. Furthermore, we aim to extend garment dynamic super-resolution to integrate wrinkle synthesis and garment appearance rendering, achieving better structural consistency across time and space than previous works [Zhang et al. 2021b]. Finally, as GDSR is trained on a single fabric material, we could introduce conditioned instance normalization to adjust latent features for different materials [Zhang et al. 2021a], or explore meta-learning for high-resolution garment simulations to meet the apparel industry's customization needs.

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their constructive comments; Kaizhang Kang for his generous helping with text, graphics, and video polishing; Mixamo for the motion sequences. This work was partially supported by the National Science Fund of China (No. 62472223 and No. 62072242).

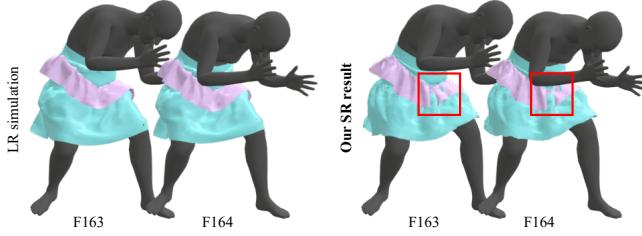


Fig. 11. Limitation. Our method fails to capture garment details influenced by friction between fabric layers. Although our explicit collision handling resolves most intersections between fabric layers, some collisions remain because we cannot detect intersections when the distances between intersecting vertices exceed the predefined threshold.

REFERENCES

- David Baraff and Andrew Witkin. 1998. Large Steps in Cloth Simulation. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*. Association for Computing Machinery, New York, NY, USA, 43–54. <https://doi.org/10.1145/280814.280821>
- David Baraff and Andrew Witkin. 2023. Large steps in cloth simulation. In *Seamless Graphics Papers: Pushing the Boundaries, Volume 2*. 767–778.
- Hugo Bertete, Meysam Madadi, and Sergio Escalera. 2022. Neural Cloth Simulation. *ACM Transactions on Graphics (TOG)* 41, 6 (2022), 1–14.
- R. Bridson, S. Marino, and R. Fedkiw. 2005. Simulation of clothing with folds and wrinkles. In *ACM SIGGRAPH 2005 Courses (Los Angeles, California) (SIGGRAPH '05)*. Association for Computing Machinery, New York, NY, USA, 3–es. <https://doi.org/10.1145/1198555.1198573>
- Lan Chen, Lin Gao, Jie Yang, Shibiao Xu, Juntao Ye, Xiaopeng Zhang, and Yu-Kun Lai. 2023. Deep deformation detail synthesis for thin shell models. In *Computer Graphics Forum*, Vol. 42. Wiley Online Library, e14903.
- Lan Chen, Juntao Ye, and Xiaopeng Zhang. 2021c. Multi-feature super-resolution network for cloth wrinkle synthesis. *Journal of Computer Science and Technology* 36, 3 (2021), 478–493.
- Yinbo Chen, Sifei Liu, and Xiaolong Wang. 2021b. Learning Continuous Image Representation with Local Implicit Image Function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 8628–8638.
- Zhen Chen, Hsiao-Yu Chen, Danny M Kaufman, Mélina Skouras, and Etienne Vouga. 2021a. Fine wrinkling on coarsely meshed thin shells. *ACM Transactions on Graphics (TOG)* 40, 5 (2021), 1–32.
- Kwang-Jin Choi and Hyeong-Seok Ko. 2002. Stable but Responsive Cloth. *ACM Trans. Graph.* 21, 3 (jul 2002), 604–611.
- Kwang-Jin Choi and Hyeong-Seok Ko. 2005. Research problems in clothing simulation. *Computer-aided design* 37, 6 (2005), 585–592.
- Mengyu Chu, You Xie, Jonas Mayer, Laura Leal-Taixé, and Nils Thuerey. 2020. Learning temporal coherence via self-supervision for GAN-based video generation. *ACM Transactions on Graphics (TOG)* 39, 4 (2020), 75–1.
- Gabriel Cirio, Jorge Lopez-Moreno, David Miraut, and Miguel A Otaduy. 2014. Yarn-level simulation of woven cloth. *ACM Transactions on Graphics (TOG)* 33, 6 (2014), 1–11.
- Y D. Li, Min Tang, Yun Yang, Zi Huang, R F. Tong, Shuang Cai Yang, Yao Li, and Dinesh Manocha. 2022. N-Cloth: Predicting 3D Cloth Deformation with Mesh-Based Networks. In *Computer Graphics Forum*, Vol. 41. Wiley Online Library, 547–558.
- Wei-Wen Feng, Yizhou Yu, and Byung-Uck Kim. 2010. A Deformation Transformer for Real-Time Cloth Animation. *ACM Trans. Graph.* 29, 4, Article 108 (July 2010), 9 pages.
- Stefan Fröhlich and Mario Botsch. 2011. Example-driven deformations based on discrete shells. In *Computer graphics forum*, Vol. 30. Wiley Online Library, 2246–2257.
- Russell Gillette, Craig Peters, Nicholas Vining, Essex Edwards, and Alla Sheffer. 2015. Real-time dynamic wrinkling of coarse animated cloth. In *Proceedings of the 14th ACM SIGGRAPH/eurographics symposium on computer animation*. 17–26.
- Artur Grigorev, Bernhard Thomaszewski, Michael J. Black, and Otmar Hilliges. 2023. HOOD: Hierarchical Graphs for Generalized Modelling of Clothing Dynamics. *Computer Vision and Pattern Recognition (CVPR)*.
- Eitan Grinspun, Anil N Hirani, Mathieu Desbrun, and Peter Schröder. 2003. Discrete shells. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation*. Citeseer, 62–67.
- Peng Guan, Loretta Reiss, David A Hirshberg, Alexander Weiss, and Michael J Black. 2012. Drape: Dressing any person. *ACM Transactions on Graphics (TOG)* 31, 4 (2012), 1–10.
- Erhan Gundogdu, Victor Constantin, Amrollah Seifoddini, Minh Dang, Mathieu Salzmann, and Pascal Fua. 2019. Garnet: A two-stream network for fast and accurate 3d cloth draping. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 8739–8748.
- Fabian Hahn, Bernhard Thomaszewski, Stelian Coros, Robert W Sumner, Forrester Cole, Mark Meyer, Tony DeRose, and Markus Gross. 2014. Subspace clothing simulation using adaptive bases. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 1–9.
- Oshri Halimi, Egor Larionov, Zohar Barzelay, Philipp Herholz, and Tuur Stuyck. 2023. PhysGraph: Physics-Based Integration Using Graph Neural Networks. *arXiv preprint arXiv:2301.11841* (2023).
- Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. 2019. Recurrent back-projection network for video super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 3897–3906.
- Ladislav Kavan, Dan Gerszewski, Adam W. Bargteil, and Peter-Pike Sloan. 2011. Physics-Inspired Upsampling for Cloth Simulation in Games. *ACM Trans. Graph.* 30, 4, Article 93 (July 2011), 10 pages.
- Theodore Kim and David Eberle. 2020. Dynamic deformables: implementation and production practicalities. In *ACM SIGGRAPH 2020 Courses*. 1–182.
- Zorah Lahner, Daniel Cremers, and Tony Tung. 2018. Deepwrinkles: Accurate and realistic clothing modeling. In *Proceedings of the European conference on computer vision (ECCV)*. 667–684.
- Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4681–4690.
- Dohae Lee, Hyun Kang, and In-Kwon Lee. 2023. ClothCombo: Modeling Inter-Cloth Interaction for Draping Multi-Layered Clothes. *ACM Transactions on Graphics (TOG)* 42, 6 (2023), 1–13.
- Minchen Li, Danny M. Kaufman, and Chenfanfu Jiang. 2021. Codimensional Incremental Potential Contact. *ACM Trans. Graph.* 40, 4, Article 170 (jul 2021), 24 pages.
- Minchen Li, Alla Sheffer, Eitan Grinspun, and Nicholas Vining. 2018. Foldsketch: Enriching garments with physically reproducible folds. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–13.
- Xuan Li, Yu Fang, Lei Lan, Huamin Wang, Yin Yang, Minchen Li, and Chenfanfu Jiang. 2023a. Subspace-Preconditioned GPU Projective Dynamics with Contact for Cloth Simulation. In *SIGGRAPH Asia 2023 Conference Papers* (Sydney, NSW, Australia,) (SA '23). Association for Computing Machinery, New York, NY, USA, Article 1, 12 pages.
- Yudi Li, Min Tang, Yun Yang, Ruofeng Tong, Shuangcui Yang, Yao Li, Bailin An, and Qilong Kou. 2023b. CTSN: Predicting Cloth Deformation for Skeleton-based Characters with a Two-stream Skinning Network. *arXiv preprint arXiv:2305.18808* (2023).
- Emmanuel Ian Libao, Myeongjin Lee, Sumin Kim, and Sung-Hee Lee. 2023. MeshGraph-NetRP: Improving Generalization of GNN-based Cloth Simulation. In *Proceedings of the 16th ACM SIGGRAPH Conference on Motion, Interaction and Games*. 1–7.
- Tiantian Liu, Adam W Bargteil, James F O'Brien, and Ladislav Kavan. 2013. Fast simulation of mass-spring systems. *ACM Transactions on Graphics (TOG)* 32, 6 (2013), 1–7.
- James Long, Katherine Burns, and Jingzhou (James) Yang. 2011. Cloth Modeling and Simulation: A Literature Survey. In *Digital Human Modeling*. Springer Berlin Heidelberg, Berlin, Heidelberg, 312–320.
- Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. 2015. SMPL: A Skinned Multi-Person Linear Model. *ACM Trans. Graph.* 34, 6, Article 248 (oct 2015), 16 pages.
- Matthias Müller and Nuttappong Chentanez. 2010. Wrinkle Meshes. In *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, MZoran Popovic and Miguel Otaduy (Eds.). The Eurographics Association.
- Vinod Nair and Geoffrey E Hinton. 2010. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*. 807–814.
- Andrew Nealen, Matthias Müller, Richard Keiser, Eddy Boxerman, and Mark Carlson. 2006. Physically based deformable models in computer graphics. In *Computer graphics forum*, Vol. 25. Wiley Online Library, 809–836.
- Xiaoyu Pan, Jiaming Mai, Xinwei Jiang, Dongxue Tang, Jingxiang Li, Tianjia Shao, Kun Zhou, Xiaogang Jin, and Dinesh Manocha. 2022. Predicting Loose-Fitting Garment Deformations Using Bone-Driven Motion Networks. In *ACM SIGGRAPH 2022 Conference Proceedings* (Vancouver, BC, Canada) (SIGGRAPH '22). Association for Computing Machinery, New York, NY, USA, Article 11, 10 pages.
- Chaitanya Patel, Zhouyingcheng Liao, and Gerard Pons-Moll. 2020. Tailornet: Predicting clothing in 3d as a function of human pose, shape and garment style. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 7365–7375.
- Tobias Pfaff, Meir Fortunato, Alvaro Sanchez-Gonzalez, and Peter Battaglia. 2020. Learning Mesh-Based Simulation with Graph Networks. In *International Conference on Learning Representations*.
- Oliver Rémiillard and Paul G Kry. 2013. Embedded thin shells for wrinkle simulation. *ACM Transactions on Graphics (TOG)* 32, 4 (2013), 1–8.

- Damien Rohmer, Tiberiu Popa, Marie-Paule Cani, Stefanie Hahmann, and Alla Sheffer. 2010. Animation wrinkling: augmenting coarse cloth simulations with realistic-looking wrinkles. *ACM Transactions on Graphics (ToG)* 29, 6 (2010), 1–8.
- Scott D Roth. 1982. Ray casting for modeling solids. *Computer graphics and image processing* 18, 2 (1982), 109–144.
- Mehdi SM Sajjadi, Raviteja Vemulapalli, and Matthew Brown. 2018. Frame-recurrent video super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 6626–6634.
- Igor Santesteban, Miguel Otaduy, Nils Thuerey, and Dan Casas. 2022. ULNeF: untangled layered neural fields for mix-and-match virtual try-on. *Advances in Neural Information Processing Systems* 35 (2022), 12110–12125.
- Igor Santesteban, Miguel A Otaduy, and Dan Casas. 2019. Learning-based animation of clothing for virtual try-on. In *Computer Graphics Forum*, Vol. 38. Wiley Online Library, 355–366.
- Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G Baraniuk. 2023. Wire: Wavelet implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18507–18516.
- Yidi Shao, Chen Change Loy, and Bo Dai. 2023. Towards Multi-Layered 3D Garments Animation. *arXiv preprint arXiv:2305.10418* (2023).
- Yu Shen, Junbang Liang, and Ming C Lin. 2020. Gan-based garment generation using sewing pattern images. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII* 16. Springer, 225–247.
- Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *International Conference on Learning Representations*.
- Huamin Wang. 2021. GPU-Based Simulation of Cloth Wrinkles at Submillimeter Levels. *ACM Trans. Graph.* 40, 4, Article 169 (jul 2021), 14 pages.
- Huamin Wang, Florian Hecht, Ravi Ramamoorthi, and James F. O'Brien. 2010. Example-Based Wrinkle Synthesis for Clothing Animation. *ACM Trans. Graph.* 29, 4, Article 107 (July 2010), 8 pages.
- Tuanfeng Y Wang, Tianjia Shao, Kai Fu, and Niloy J Mitra. 2019. Learning an intrinsic garment space for interactive authoring of garment animation. *ACM Transactions on Graphics (TOG)* 38, 6 (2019), 1–12.
- Weiwei Xu, Nobuyuki Umetani, Qianwen Chao, Jie Mao, Xiaogang Jin, and Xin Tong. 2014. Sensitivity-optimized rigging for example-based real-time clothing synthesis. *ACM Trans. Graph.* 33, 4 (2014), 107–1.
- Jiayi Eris Zhang, Jérémie Dumas, Yun Fei, Alec Jacobson, Doug L James, and Danny M Kaufman. 2023. Progressive Shell Qasistics for Unstructured Meshes. *ACM Transactions on Graphics (TOG)* 42, 6 (2023), 1–17.
- Meng Zhang, Duygu Ceylan, and Niloy J. Mitra. 2022. Motion Guided Deep Dynamic 3D Garments. *ACM Trans. Graph.* 41, 6, Article 219 (nov 2022), 12 pages.
- Meng Zhang, Tuanfeng Wang, Duygu Ceylan, and Niloy J. Mitra. 2021a. Deep detail enhancement for any garment. In *Computer Graphics Forum*, Vol. 40. Wiley Online Library, 399–411.
- Meng Zhang, Tuanfeng Y. Wang, Duygu Ceylan, and Niloy J. Mitra. 2021b. Dynamic Neural Garments. *ACM Trans. Graph.* 40, 6, Article 235 (dec 2021), 15 pages.
- Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. 2018. Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2472–2481.
- Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. 2017. Loss Functions for Image Restoration with Neural Networks. *IEEE Transactions on Computational Imaging* (2017).
- Javier S Zurdo, Juan P Brito, and Miguel A Otaduy. 2012. Animating wrinkles by example on non-skinned cloth. *IEEE Transactions on Visualization and Computer Graphics* 19, 1 (2012), 149–158.