Toward Scalable and Generalisable Garment Simulation with Graph Neural Networks

AORAN LIU

Master of Philosophy

Supervisor: Prof. Zhiyong Wang
Associate Supervisor: A/Prof. Weidong Cai

A thesis submitted in fulfilment of
the requirements for the degree of
Master of Philosophy

School of Computer Science
Faculty of Engineering
The University of Sydney
Australia

14 October 2025

To my parents.

For their endless love, support and encouragement.

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Signature:
School of Computer Science
Faculty of Engineering
The University of Sydney
20 June 2025

Authorship Attribution Statement

The publications associated with this thesis are as follows.

- Chapter 3 of this thesis is based on the published paper:
  Aoran Liu, Kun Hu, Clinton Mo, Changyang Li, and Zhiyong Wang. "Extended Short-and Long-Range Mesh Learning for Fast and Generalized Garment Simulation." In 2025 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2025.
  My contributions are: Design of the algorithm; Implementation of the algorithm; Conducting experiments; Writing the draft for the manuscript.
- Chapter 4 of this thesis is based on a paper that is currently under review:
  Aoran Liu, Kun Hu, Clinton Mo, Qiuxia Wu, Wenxiong Kang and Zhiyong Wang. "Pb4U-GNet: Resolution-Adaptive Garment Simulation via Propagation-before-Update Graph Network." Under review at AAAI 2026.
  My contributions are: Design of the algorithm; Implementation of the algorithm; Conducting experiments; Writing the draft for the manuscript.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Student Name: Aoran Liu
Signature:
Date: 20 June 2025

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Supervisor Name: Zhiyong Wang
Signature:
Date: 20 June 2025

Generative AI Attribution Statement

During the preparation of the thesis, the author used ChatGPT, Claude, and Gemini for the purposes of text enhancement. The use of these generative AI tools includes paraphrasing, grammar corrections, and spelling corrections. The author confirms that where text was modified by generative AI, the content was reviewed for possible errors, inaccuracies, and bias. The author takes full responsibility for the submitted thesis and ensures the work is their own and has used generative AI within the parameters of use.

Signature:
School of Computer Science
Faculty of Engineering
The University of Sydney
20 June 2025

Abstract

Realistic garment simulation is crucial for numerous multimedia and computer graphics applications, including film production, video games, virtual try-on systems, and fashion design. Physics-based methods dominate this field due to their strong generalisability and high-quality simulation results. However, these approaches typically require numerical optimisation to solve the garment's equilibrium state at each timestep, which is often computationally expensive and presents a significant barrier to interactive applications such as video games and virtual try-on.

The rapid advancement of artificial intelligence has introduced data-driven methods as a promising alternative for achieving fast garment simulation. Early data-driven methods primarily employed pose-driven strategies, using body pose sequences to predict corresponding garment motion through neural networks. However, pose-driven methods suffer from limited generalisability—they cannot adapt to different garment meshes without retraining, as the neural network learns mesh-specific patterns between body pose and garment behaviour.

To address these generalisation limitations, recent research has demonstrated the significant potential of Graph Neural Networks (GNNs) for garment simulation. GNNs directly capture local vertex interactions and predict per-vertex garment dynamics, enabling the network to focus on local behaviours whilst eliminating dependency on global garment shape or topology. This vertex-level approach not only achieves superior generalisation across different garment meshes but also delivers improved simulation quality compared to pose-driven methods.

Despite their promise, GNN-based methods face two key limitations. First, accurate prediction of per-vertex dynamics requires a large receptive field on each vertex, typically necessitating multiple rounds of message passing, which reduces computational efficiency. Second, these methods are sensitive to mesh resolution and struggle to generalise to resolutions

not seen during training. Together, these challenges limit the practicality of GNN-based garment simulation.

This thesis investigates solutions to these limitations to enhance the practicality and robustness of GNN-based garment simulation. Specifically, we present two studies that address these issues individually. The first study focuses on enhancing simulation efficiency by introducing two novel modules: the Laplacian-Smoothed Dual Message Passing (LSDMP) module and the Geodesic Self-Attention (GSA) module. LSDMP integrates Laplacian smoothing to efficiently propagate short-range features, while GSA captures global mesh information via self-attention. Together, these modules significantly expand each vertex's receptive field without requiring deep message passing, improving both efficiency and scalability.

The second study focuses on improving the generalisability of GNN-based methods to garment meshes with varying resolutions. We introduce an adaptive message passing scheme that dynamically adjusts the message passing depth according to the input mesh resolution. Additionally, we propose a resolution-aware scaling operation to ensure consistent physical behaviour of the garment across different resolutions. This approach significantly enhances the model's ability to generalise to unseen mesh resolutions during inference.

Extensive experiments are conducted in both studies to validate the effectiveness of the proposed methods. We also demonstrate their superior performance compared to existing state-of-the-art approaches. Collectively, the outcomes of these studies advance the practicality of GNN-based garment simulation and reinforce its potential for real-world deployment.

Acknowledgements

The journey of completing this research has been both challenging and rewarding. I would like to express my gratitude to everyone who has supported and encouraged me along the way. This thesis would not have been possible without their guidance and care.

First and foremost, I extend my sincere thanks to my supervisor, Professor Zhiyong Wang, for his unwavering support and mentorship throughout this journey. His insightful guidance at every stage of the research has been instrumental in shaping this work. I am especially grateful for his help in developing my critical thinking and presentation skills.

I am also deeply thankful to Dr. Kun Hu for his generous support and invaluable feedback. His suggestions were crucial during the early stages of idea formulation and in refining the manuscripts for publication.

My sincere thanks go to Mr. Clinton Mo, whose critical feedback and engaging discussions have profoundly shaped my research mindset. His insights were essential in the development of my research ideas.

I would like to thank all my friends and colleagues at the university and in the research lab for their kindness and encouragement: Dr. Mengwei He, Mr. Wenhua Wu, Mr. Lintao Wang, Mr. Zhuqiang Lu, Mr. Penghui Wen, Mr. Hengzhi Chen, Mr. Yilin Zhu, Mr. Zexin Hu, Mr. Hang Yu, Mr. Zhenfei Yin, Mr. Liqian Feng, Mr. Ivan Gao, Ms. Yu Luo, Mr. Jeffrey Lu, Ms. Anastasia Yi, Mr. Shiwen Zhao, Mr. Chuancheng Shi, Mr. Zifan Zheng, Mr. Kevin Li, Mr. Keshen Zhou, Mr. Zhaoqing Wang, Ms. Yaxuan Song, Mr. Yutong Wang, and Mr. Jiahao Chen. Their companionship has made this journey much more enjoyable and fulfilling.

A special note of thanks goes to my dear partner, Qinghua Liu, for her love and unwavering support. Her presence has helped me navigate the most stressful and difficult times.

Lastly, I am forever grateful to my parents and my sister, whose unconditional love and support have always been my greatest source of strength.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Garment Simulation

Garment simulation is the computational process of modelling the dynamic behaviour of garments in virtual environments, driven by the motion of an underlying avatar. It aims to accurately reproduce how garments deform and move in response to physical forces such as gravity, inertia, collisions, and friction. As a core topic in computer graphics, garment simulation plays a vital role in a range of multimedia applications, including film production, video games, virtual try-on platforms, and fashion design. Achieving visually realistic and physically plausible cloth behaviour is essential for enhancing both user experience and visual fidelity in these domains.

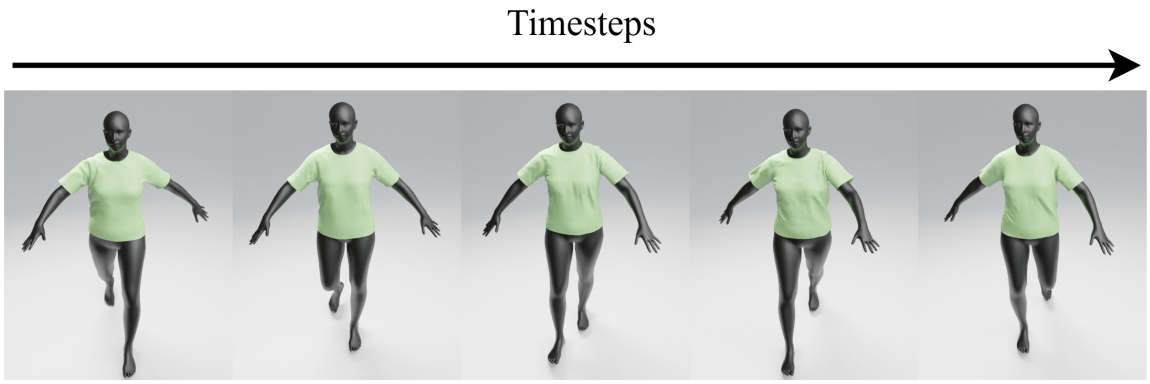

Figure 1.1. Garment simulation refers to the process of modelling the dynamic behaviour of garments, driven by the motion of an underlying avatar.

Depending on the application context, simulation methods must balance visual realism and computational efficiency. For example, in film production, simulations are typically performed offline, allowing for high computational costs to achieve cinematic quality. In

contrast, interactive applications such as video games or real-time virtual try-on require fast and responsive simulations, where computational efficiency often takes precedence over strict physical accuracy.

Traditionally, garment simulation has been achieved using physics-based approaches such as mass-spring systems (Provot et al. 1995; Baraff and Witkin 1998) and finite element methods (Liu et al. 2022). These methods are grounded in well-established physical principles and can produce highly accurate and realistic results. They provide excellent generalisability across different garment types and offer intuitive parameter control based on physical material properties. However, their reliance on numerical optimisation to solve the garment's equilibrium state at each timestep often leads to prohibitively high computational costs, making them unsuitable for real-time applications where latency is critical.

To address these computational limitations, recent research has increasingly turned to data-driven approaches that approximate physical behaviour using machine learning models. These approaches can be broadly categorised into two paradigms: pose-driven and mesh-driven methods.

Pose-driven methods train neural networks to directly map body poses to garment deformations. These methods are extremely lightweight and fast at inference since body pose data are typically low-dimensional. However, they suffer from severe generalisability limitations—the neural network learns fixed patterns between body poses and garment deformation that are highly garment-specific, making it difficult to adapt to new garment designs without retraining.

Mesh-driven methods, particularly those based on Graph Neural Networks (GNNs), address this limitation by explicitly utilizing mesh topology and geometry. Rather than learning global pose-to-deformation mappings, GNNs predict per-vertex deformations by modelling local interactions between vertices through message passing mechanisms. This approach enables generalisation to unseen garment meshes with different topologies while producing more realistic dynamics due to the explicit modelling of vertex interactions. Several recent works (Pfaff et al. 2020; Sanchez-Gonzalez et al. 2020; Grigorev et al. 2023) have demonstrated that GNN-based simulators can achieve considerable speedups over traditional physics-based

methods while maintaining comparable accuracy and exhibiting strong generalisability across different garment types and topologies.

1.2 Research Motivations

While GNN-based garment simulation represents a significant advancement over traditional physics-based methods, critical limitations remain that restrict practical deployment in real-world applications. Most existing GNN-based garment simulators build upon the MeshGraphNet framework (Pfaff et al. 2020). As illustrated in Figure 1.2, MeshGraphNet operates through three stages: first, each vertex encodes local features (velocity, mass, vertex normal, etc.) into a latent embedding; second, a series of message passing layers exchange information between neighboring vertices and update their embeddings; finally, these embeddings are decoded into predicted vertex accelerations for the current timestep.

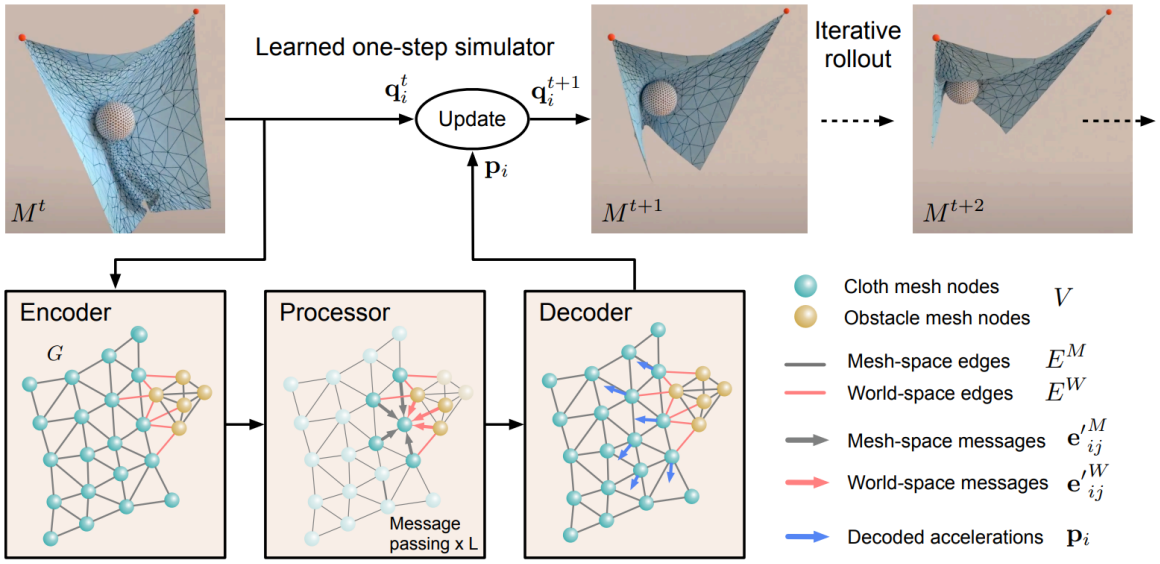

Figure 1.2. Overview of the MeshGraphNet (Pfaff et al. 2020) architecture, which propagates vertex features through a series of message passing layers.

Despite strong performance on mesh-based simulation, this framework suffers from two fundamental limitations that stem from its reliance on fixed-depth message passing:

Limited Scalability to High-Resolution Meshes. Accurate cloth simulation requires capturing long-range interactions such as elastic wave propagation. In MeshGraphNet, each message passing step expands a vertex's receptive field by one graph hop, meaning that k message passing layers yield a receptive field spanning k-hop neighbors. For high-resolution meshes where vertices are densely packed, achieving physically meaningful receptive fields demands dozens of message passing iterations. This creates a computational bottleneck: as mesh resolution increases, the cost of information propagation grows substantially, making the approach prohibitively expensive for real-time applications or detailed garment geometry. The fundamental issue is that receptive field growth is tied to the number of layers rather than the physical distance needed to capture cloth dynamics.

Poor Generalization Across Mesh Resolutions. The fixed number of message passing steps creates a tight coupling between model architecture and mesh resolution. A model trained with k message passing steps learns to operate within a specific receptive field size. When applied to higher-resolution meshes, the same k steps cover a smaller physical region of the garment, failing to capture necessary long-range interactions. Conversely, on lower-resolution meshes, excessive message passing can cause oversmoothing and degrade prediction quality. This resolution sensitivity is particularly problematic in production environments where garment meshes vary widely due to design requirements, rendering platforms, or computational constraints. Current methods like MeshGraphNets (Pfaff et al. 2020) and HOOD (Grigorev et al. 2023) require retraining or manual tuning of network depth for each target resolution, limiting their practical utility.

These two limitations are inherently interconnected: both arise from the fundamental design choice of using a fixed number of graph convolution layers to propagate information. This thesis addresses these challenges through two complementary research contributions that introduce novel architectural components to the simulation pipeline. Our goal is to bridge the gap between the theoretical promise of data-driven garment simulation and its practical deployment in applications demanding both computational efficiency and adaptability across diverse garment representations.

1.3 Summary of Contributions

As outlined in the previous section, this thesis presents two research studies aimed at improving the practicality of GNN-based garment simulation.

In our first study (Chapter 3), we address the scalability and efficiency limitations of existing GNN-based methods. Our focus is on enlarging the receptive field of each message passing step without introducing significant computational overhead. To achieve this, we develop two complementary modules: *Laplacian-Smoothed Dual Message Passing* module and *Geodesic Self-Attention* module. Laplacian smoothing, a lightweight operation traditionally used for mesh regularisation, enables vertex features to diffuse efficiently across wider neighbourhoods. Importantly, its attenuation effect is well aligned with the dissipative nature of elastic wave propagation, making it a natural choice for modelling fabric dynamics. In parallel, geodesic-aware self-attention constructs direct connections between distant vertices, allowing each vertex to access global mesh information without relying on many iterative message passing steps. By integrating these two modules, we formulate a novel GNN-based garment simulation pipeline that propagates information more effectively and at lower cost, thereby improving scalability to higher-resolution meshes.

In our second study (Chapter 4), we focus on improving the resolution robustness of GNN-based simulators. A fundamental issue of current pipelines is their reliance on a fixed receptive field, which restricts generalisability across different mesh resolutions. To overcome this, we introduce an adaptive message passing framework that decouples message propagation from feature updates through a propagate-before-update strategy. This design allows the receptive field to expand dynamically in response to the input mesh resolution. Furthermore, we incorporate a resolution-aware update scaling mechanism that adjusts network outputs according to mesh granularity, ensuring physically consistent predictions across varying garment resolutions.

In summary, the main contributions of this thesis are as follows:

- We improve the efficiency and scalability of GNN-based garment simulation by introducing a Laplacian smoothing-based message passing scheme and a geodesic-aware self-attention module, both of which expand the effective receptive field while reducing computational cost.
- We enhance the resolution generalisability of GNN-based models by proposing a propagate-before-update framework and a resolution-aware scaling mechanism, enabling consistent performance across meshes of varying resolution.
- We conduct extensive experiments to validate the effectiveness and robustness of our proposed methods, and provide a comprehensive analysis of their strengths and limitations.

1.4 Thesis Structure

The remainder of this thesis is structured as follows:

Chapter 2 reviews prior work in garment simulation, covering both physics-based and learning-based approaches, with an emphasis on recent learning-based methods.

Chapter 3 presents the first research study, which focuses on improving the simulation efficiency and scalability of GNN-based methods through a Laplacian smoothing-based message passing scheme and a geodesic-aware self-attention module.

Chapter 4 presents the second research study, which addresses resolution generalisability of GNN-based methods via an adaptive message passing framework and a resolution-aware scaling mechanism.

Chapter 5 concludes the thesis with a summary of findings, a discussion of limitations, and directions for future work.

CHAPTER 2

Literature review

Garment simulation has been a well-studied topic in computer graphics, with a wide range of methods proposed over the past decades. This chapter reviews the key developments in the field, beginning with traditional physics-based approaches and progressing to more recent learning-based methods. While both paradigms are discussed, particular emphasis is placed on learning-based techniques, as they form the foundation of this thesis.

We first review classical physics-based simulation methods, which serve as the backbone of high-fidelity garment modelling. We then discuss recent advances in machine learning for garment simulation, highlighting different architectures, input representations, and their respective strengths and limitations.

2.1 Physics-Based Methods

2.1.1 Classical Physics-Based Methods

Garment dynamics is characterised by highly non-linear behaviour, with small variations in force or configuration leading to significantly different outcomes. Prior works in cloth simulation (Terzopoulos et al. 1987; Baraff and Witkin 1998; Bridson et al. 2002) have highlighted the sensitivity and complexity of fabric motion, which arises from the interplay of stretching, bending, and collision forces. To address this complexity, Terzopoulos et al. (Terzopoulos et al. 1987) introduced the use of partial differential equations based on principles from elasticity theory and continuum mechanics to model dynamic material behaviour. This

formulation established a robust foundation for simulating realistic garment deformations, including stretching, bending, and other complex behaviours.

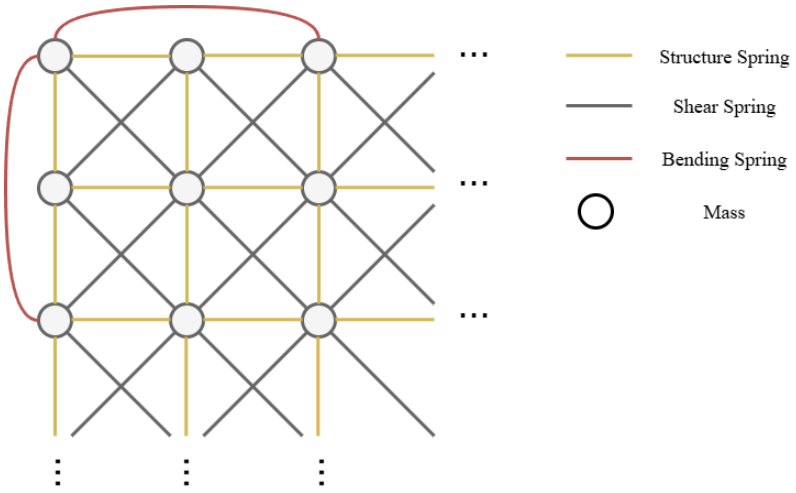

Figure 2.1. Visualisation of a mass-spring system.

Building upon this theoretical groundwork, the mass-spring method (Provot et al. 1995) emerged as a widely-adopted computational approach for fabric dynamics simulation. As shown in 2.1, the mass-spring method represents cloth as a discrete network of point masses m_i connected by springs, where each spring encodes local elastic properties of the fabric material.

In this formulation, the elastic force exerted by a spring connecting vertices i and j is computed using Hooke's law in vectorised form:

$$\mathbf{F}_{e,ij} = -k_{ij} \left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\| - L_{ij}}{\|\mathbf{x}_i - \mathbf{x}_j\|} \right) (\mathbf{x}_i - \mathbf{x}_j) \tag{2.1}$$

where L_{ij} is the rest length of the spring, $\|\mathbf{x}_i - \mathbf{x}_j\|$ is the current Euclidean distance between vertices, and k_{ij} is the spring stiffness constant. This formulation ensures the force acts along the spring direction with magnitude proportional to the displacement from equilibrium.

The net elastic force acting on each vertex i is computed by summing contributions from all connected springs in its neighborhood $\mathcal{N}(i)$:

$$\mathbf{F}_e^i = \sum_{j \in \mathcal{N}(i)} \mathbf{F}_{e,ij} \tag{2.2}$$

The vertex acceleration is then obtained through Newton's second law, incorporating both elastic forces and external forces (gravity, collision response, etc.):

$$\mathbf{a}_i = \frac{\mathbf{F}_e^i + \mathbf{F}_{ext}^i}{m_i} \tag{2.3}$$

where $\mathbf{F}_{ext}^i$ represents the sum of all external forces acting on vertex i. For temporal integration, simple forward Euler integration can be employed to update the vertex position:

$$\mathbf{v}_{t+1}^i = \mathbf{v}_t^i + \mathbf{a}_t^i \Delta t \tag{2.4}$$

$$\mathbf{x}_{t+1}^i = \mathbf{x}_t^i + \mathbf{v}_{t+1}^i \Delta t \tag{2.5}$$

However, it should be noted that this explicit forward integration scheme suffers from numerical instability for large time steps Δt, particularly with stiff springs. This is because the pipeline assumes that vertex velocities remain constant within each time step, an approximation that is physically inaccurate. In reality, vertex velocities change continuously, and ignoring this variation can introduce significant errors. When large time steps are used, these errors accumulate, leading to numerical instability and overshooting in the simulation. To reduce such errors, very small time steps are typically required, but this dramatically increases the computational cost, making the simulation inefficient for practical applications.

To address this challenge, later works (Baraff and Witkin 1998; Choi and Ko 2005; Volino et al. 2009) proposed applying implicit and semi-implicit Euler integration to cloth simulation,

marking another foundational advancement in garment dynamics. The key insight behind these methods is to solve for future states by considering the forces and accelerations that will exist at the end of the time step, rather than relying solely on current-state information. Instead of using the present state to directly predict the next state, implicit and semi-implicit integration formulate and solve a system of equations that accounts for how forces will change as vertices move to their new positions during the time step. These approaches provide significantly greater numerical stability and enable the use of larger time steps without experiencing energy blow-up or numerical instability.

In addition to the mass-spring method, the Finite Element Method (FEM) (Grinspun et al. 2003; Liu et al. 2022) represents another classical approach for garment simulation. While FEM follows a similar overall pipeline to mass-spring methods, it differs fundamentally in how forces are computed. Specifically, FEM treats each triangular face in the garment mesh as a finite element with well-defined material properties. For each element, FEM computes the deformation gradient by comparing the current configuration to the rest (undeformed) state, then evaluates the elastic energy using continuum mechanics-based constitutive models such as linear elasticity or hyperelastic materials. Forces acting on vertices are derived by taking the negative gradient of the total elastic energy with respect to vertex positions, and accelerations are subsequently computed using Newton's second law.

Beyond traditional mass–spring and finite-element approaches, researchers have also explored yarn-level simulation to capture cloth behaviour with greater physical fidelity (Kaldor et al. 2008; Kaldor et al. 2010; Cirio et al. 2014). Other complementary lines of work focus on material-specific modelling (Bhat et al. 2003; Miguel et al. 2012; Wang et al. 2011) and on more robust collision-handling schemes (Baraff and Witkin 1998; Bridson et al. 2002; Harmon et al. 2009; Tang et al. 2018; Li et al. 2020), all of which further improve the realism and stability of garment simulation.

In conclusion, classical physics-based methods form the foundation of garment simulation, offering physically interpretable and controllable frameworks. However, their computational demands, whether from requiring small time steps for stability or solving complex nonlinear systems in implicit schemes, can limit their efficiency for real-time applications.

2.1.2 Constraint based method

To achieve more efficient simulation, an alternative class of methods—constraint-based approaches—has been developed. These methods reformulate the simulation process as an optimisation problem that directly updates vertex positions to satisfy physical constraints, rather than explicitly computing velocities and performing time integration. For garment simulation, two prominent constraint-based methods are Position-Based Dynamics (PBD) and Projective Dynamics (PD).

2.1.2.1 Position-Based Dynamics

Position-Based Dynamics (PBD) (Müller et al. 2007; Bender et al. 2014) is widely adopted in real-time simulation engines due to its stability and computational efficiency. The PBD framework is driven by a set of constraints that encode the desired physical behaviour of the garment. For instance, when modelling cloth as a mass-spring system, distance constraints ensure that edges maintain approximately their rest lengths, reflecting the material's resistance to stretching. Additional constraints may include bending resistance, collision avoidance, and boundary conditions.

PBD operates through an iterative constraint satisfaction process. After applying external forces (such as gravity) and computing predicted positions using explicit integration, the algorithm iteratively projects vertices to satisfy constraints. In each iteration, constraint violations are computed and corrections are applied to vertex positions, gradually bringing the system closer to a physically plausible state. This projection-based approach avoids solving complex differential equations, making PBD particularly suitable for applications prioritizing computational efficiency and numerical stability. However, PBD's computational cost scales with mesh resolution, and achieving real-time performance typically requires reducing mesh detail or limiting iteration counts, which can compromise accuracy and visual quality.

2.1.2.2 Projective Dynamics

Projective Dynamics (PD) (Bouaziz et al. 2014; Liu et al. 2013) extends the PBD method by reformulating the iterative projection process as an energy optimisation scheme with local and global steps. In the local step, the method examines each small piece of the cloth individually, such as each garment edge, and determines what the ideal shape should be based on the physical constraints, effectively linearizing the nonlinear energy terms. In the global step, all these individual preferences are combined to find the best compromise that satisfies the overall structure of the entire garment. This is done by solving a sparse linear system derived from the quadratic energy minimisation objective.

The key advantage of this approach lies in the regular, predictable structure of the global step's linear system that we need to solve. The system matrix can be pre-factorised once at the beginning of the simulation and then efficiently reused across all time steps, dramatically reducing computational overhead. This mathematical efficiency enables Projective Dynamics to handle higher-resolution cloth meshes while maintaining real-time performance, making it well-suited for interactive applications.

2.2 Learning-Based Methods

Although Position-Based Dynamics (PBD) and Projective Dynamics (PD) provide efficient simulation solutions for low-to-medium resolution meshes, their computational cost increases significantly when simulating high-resolution garment meshes, limiting their real-time applicability. With recent advances in machine learning and artificial intelligence, learning-based methods (Santesteban et al. 2019; Holden et al. 2019; Zhang et al. 2022a; Santesteban et al. 2022b; Grigorev et al. 2023) have emerged as promising alternatives, demonstrating the ability to maintain computational efficiency while accurately reproducing high-resolution garment details, such as realistic wrinkles and folds.

2.2.1 Pose-Driven Methods

One prominent direction among learning-based approaches is pose-driven methods, which train neural networks to predict garment deformation directly from body pose sequences, as shown in Fig. 2.2. Due to the lightweight representation of body poses, these methods typically achieve high computational efficiency. In this section, we introduce the evolvement of pose-driven methods.

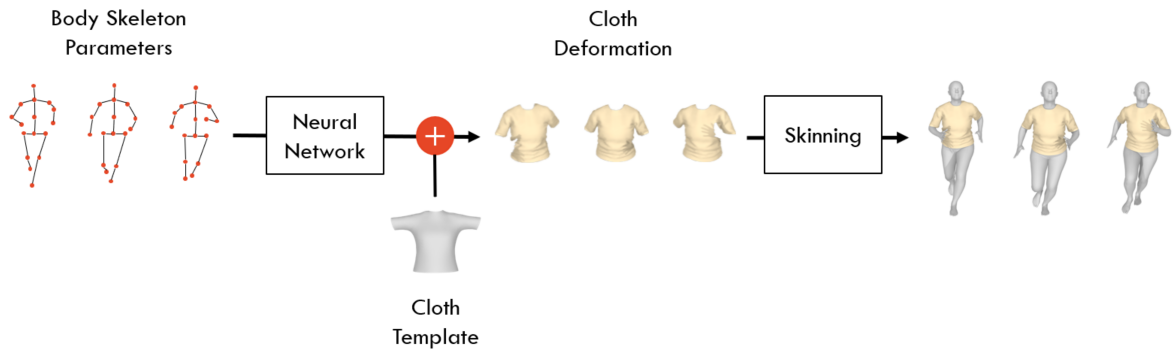

Figure 2.2. General pipeline of pose-driven methods.

2.2.1.1 Learning Pose-Based Deformation

To learn the mapping between body poses and garment deformations, early approaches (Wang et al. 2010; Xu et al. 2014) adopt example-based methods to synthesise fine wrinkle details. For instance, Wang's work. (Wang et al. 2010) construct a precomputed dataset containing garment simulations across diverse body poses. During inference, their method first obtains a coarse garment deformation using traditional physics-based simulation. Fine wrinkles are then generated by querying this dataset with the input body pose and linearly blending the retrieved samples. This approach avoids directly simulating high-resolution wrinkles, thus improving runtime efficiency. However, it requires expensive dataset precomputation and, because wrinkle synthesis relies entirely on data sampling, may produce artifacts or inaccuracies for unseen poses or motions.

To better capture the dynamic relationship between body poses and garment deformation, subsequent methods replaced the data sampling process with neural networks to generate

wrinkle details directly. For example, Lahner's work (Lahner et al. 2018) introduces a two-stage neural architecture for garment simulation. In the first stage, a neural network is trained to predict the coarse garment shape from input body poses. In the second stage, fine wrinkle details are synthesised using a generative adversarial network (GAN) (Goodfellow et al. 2020). Specifically, this stage is framed as an image enhancement task, where the UV map of the coarse prediction from the first stage is used as input to the GAN, which then outputs a refined UV map with high-frequency wrinkle details.

With a different strategy, Gundogdu et al. (Gundogdu et al. 2019) proposed GarNet, a learning-based framework that efficiently predicts garment deformations by leveraging both body and garment mesh features. By combining global and local information through a two-stream architecture and K-nearest neighbor (KNN)-based feature fusion, GarNet achieves impressive efficiency and accuracy, offering predictions up to 100 times faster than traditional physics-based methods.

Despite their advantages, a major limitation of early learning-based approaches is that each model is specifically trained for a fixed garment–body pair. As a result, a separate model must be retrained for every new combination of garment and body shape, which greatly limits scalability and flexibility. This constraint poses a significant challenge for practical deployment, especially in real-world applications where a wide variety of body shapes and garment types must be handled.

2.2.1.2 Learning Body Shape as a Variable

To enhance the flexibility of pose-driven garment simulation methods, later works start to incorporate body shape as an additional input variable alongside body pose. One early direction toward this goal involves retargeting techniques, which aim to transfer garment deformations to new body shapes and poses. For instance, Pons-Moll's work (Pons-Moll et al. 2017) proposed a method that, given a sequence of 3D scans of clothed humans represented as point clouds, extracts and segments garment meshes from the underlying body mesh. These extracted garments can then be fitted to new body shapes and poses using a deformation

transfer strategy. Similarly, some later works (Wang et al. 2018; Lahner et al. 2018; Bhatnagar et al. 2019) also adopt related retargeting approaches to handle variation in body shape.

The core idea behind these retargeting techniques is to compute the displacement between the garment mesh and the body surface in the original configuration, and then reapply this displacement to new bodies. While this method produces visually plausible results and generalises reasonably well across different body shapes, it suffers from key limitations. Notably, as also observed in Guan's work (Guan et al. 2012), the garment size tends to scale with body size due to the direct reuse of displacements. Moreover, because the wrinkles are determined solely by the original body shape, they remain unchanged regardless of variations in the new body shape—resulting in unrealistic outcomes.

In contrast, many recent approaches move beyond retargeting and instead build upon the SMPL model (Loper et al. 2015), a widely used parametric human body representation that encodes pose and shape as low-dimensional vectors. By leveraging this parameterisation, learning-based methods can explicitly model the joint influence of body pose and shape on garment deformation. This facilitates better disentanglement of the two factors and supports improved generalisation across diverse human body types.

For example, Santesteban et al. (Santesteban et al. 2019) propose a GRU-based framework for predicting garment deformations conditioned on SMPL body pose and shape parameters. The model first uses an MLP to fit the garment to the body shape, followed by a GRU network that captures pose- and shape-dependent deformations over time. To supervise training, they construct a dataset of ground-truth simulations spanning a wide range of SMPL pose and shape combinations.

While this method generalises well across different body shapes and achieves high simulation accuracy and efficiency, two key limitations remain. First, the predicted garment deformations often exhibit collisions, which require computationally expensive post-processing to resolve. Second, the model struggles to simulate loose-fitting garments such as dresses and skirts. This limitation arises from its use of linear blend skinning (Loper et al. 2015) to drape the predicted garment onto the body, where garment skinning weights are approximated from

the nearest body vertices. When garment vertices are far from the body, which is common in loose garments, this approximation introduces significant errors, leading to inaccurate deformations.

A follow-up work by Santesteban et al. (Santesteban et al. 2021) addresses both the collision and skinning limitations. The method achieves collision-free garment prediction by learning a compact, collision-free subspace using a variational autoencoder (VAE) (Pinheiro Cinelli et al. 2021). Ground-truth garment simulations are first projected into a canonical garment space and then reconstructed by the VAE, which is trained with a self-supervised collision loss that penalises interpenetration. This enables the VAE to encode garments into a collision-free latent space. A GRU-based regressor then predicts garment deformations within this subspace, eliminating the need for post-processing. To overcome the limitations of skinning, the method also introduces a learned diffused body model that extends skinning weights to any 3D point around the body, enabling more accurate deformation of loose-fitting garments.

Another notable work, VirtualBones (Pan et al. 2022), also addresses the skinning weight issue in loose-fitting garments by introducing a set of virtual bones for the garment. These virtual bones act as a dedicated skeletal structure, enabling the garment to deform using linear blend skinning (LBS) relative to this auxiliary skeleton. This approach allows for more accurate deformation, especially in regions where garment vertices are distant from the body. Additionally, Diao et al. (Diao et al. 2023) propose a method that leverages joint-specific feature learning through attention mechanisms, further enhancing simulation quality for loose-fitting garments.

2.2.1.3 Generalise to Different Garment

Another key limitation of existing pose-driven garment simulation methods is their limited generalisability to different garments. Most of these methods are trained on specific garment meshes and cannot easily adapt to new garment topologies or different material properties without retraining. To address this issue, several works have focused on improving the generalisation capabilities of learning-based simulation models.

TailorNet (Patel et al. 2020) is the first model introduced to enhance generalisability across garment variations. TailorNet introduces a set of controllable style parameters—such as sleeve length and collar size—that can be adjusted at inference time to produce garments with different designs, all without retraining the model. This approach marks a significant step toward garment-independent simulation. Furthermore, TailorNet decomposes garment deformation into low-frequency and high-frequency components, predicting them separately to achieve more realistic wrinkle detail.

In a different direction, Wang et al. (Wang et al. 2019) propose a learning-based model that accommodates garments with varying material properties. Their method accepts a sequence of body motions and at least one key frame of the corresponding garment animation, then predicts the remaining garment motion. The model's key innovation lies in its ability to implicitly infer latent material parameters from the input key frame, enabling adaptation to garments with different stiffness or elasticity properties. This capability allows the model to generalise across diverse intrinsic material settings during inference.

A significant limitation of this approach, however, is its dependence on specific garment-character pairs, necessitating independent training for each new combination. Despite this constraint, the concept of learning intrinsic material parameters from limited input data represents a promising research direction, as it suggests the potential for inferring fabric properties directly from garment meshes or animations.

Addressing similar challenges, Liu et al. (Liu et al. 2023) present a more recent material-aware model that dynamically adjusts network parameters based on input material settings to produce material-specific simulation results.

While these contributions advance the generalisation capabilities of learning-based garment simulation, the fundamental limitation persists: most existing models require retraining when confronted with novel garment topologies. This constraint primarily stems from the reliance on fully connected layers in previous architectures, which necessitate fixed-size input representations and are therefore incompatible with garments exhibiting varying mesh structures.

To address this limitation, Vidaurre et al. (Vidaurre et al. 2020) propose a novel architecture based on Graph Convolutional Networks (GCNs) (Zhang et al. 2019). The key advantage of GCNs lies in their ability to operate on irregular graph structures, making them inherently topology-agnostic. This enables the model to process garments with different mesh connectivity and geometric configurations without requiring architectural modifications or retraining. Additionally, the authors introduce a set of garment design parameters that control the shape of the input garment. These parameters can be modified at inference time, allowing users to customise garment shapes dynamically without model retraining.

Building on this foundation, other works such as DeePSD (Bertiche et al. 2021b) and GenSim (Tiwari and Bhowmick 2023) similarly integrate GCNs into their simulation pipelines. These methods encode node-wise features from the garment mesh and combine them with body pose features to predict garment deformation. By leveraging the topology-agnostic properties of GCNs, these approaches demonstrate the ability to generalise across diverse garment types and mesh structures, making them more suitable for real-world applications where garment variation is commonplace.

2.2.1.4 Physics-Informed Self-Supervised Learning

Early approaches to pose-driven garment simulation predominantly relied on supervised learning frameworks trained on ground-truth simulation data. These methods employed purely data-driven training without incorporating physics-based constraints or guidance. Consequently, the models learned to reproduce results that resemble the ground truth, rather than capturing the underlying physical laws governing garment dynamics. As a result, they often struggle to generate physically plausible outcomes, particularly when extrapolating beyond their training distributions. This limitation is in several existing data-driven methods (Santesteban et al. 2019; Santesteban et al. 2021; Patel et al. 2020). Moreover, the ground-truth data required for training these supervised models is typically generated through computationally expensive physics-based simulations, creating a significant bottleneck in the data acquisition process.

To address these fundamental challenges, Bertiche et al. (Bertiche et al. 2021a) introduce the first self-supervised learning framework for garment simulation, termed Physics-Based Neural Simulation (PBNS). The core innovation of their approach lies in the integration of physically-motivated loss functions that model the potential energy of garments. Specifically, they incorporate elastic and gravitational potential energy terms derived from classical mechanics principles.

The framework is grounded in a mass-spring model representation, where each vertex in the garment mesh is treated as a point mass connected to neighboring vertices through spring-like edges. The elastic potential energy component is computed based on the deviation of predicted edge lengths from those in the reference garment template. This formulation naturally penalises configurations where mesh edges undergo excessive stretching or compression, thereby encouraging physically consistent deformations. The gravitational potential energy follows the standard formulation $U = mgh$, ensuring that the model accounts for gravitational effects on garment behaviour. Additionally, the framework incorporates auxiliary loss terms to address practical simulation challenges. A collision loss term minimises interpenetration between the garment and body surfaces, while a pinning loss prevents unrealistic sliding of the garment due to gravitational forces. Through the optimisation of these physics-informed loss functions, the neural network learns to predict garment deformations that are both physically consistent and energetically favorable, eliminating the dependency on expensive ground truth simulation data while improving physical realism.

Despite these advances, the PBNS framework exhibits several limitations. Visual inspection of simulation results reveals that predicted garments appear overly smooth and lack realistic surface detail and wrinkle formation. Furthermore, the physics-based loss terms operate on individual frames independently, failing to capture temporal dependencies and motion continuity across sequential frames.

Building upon these insights, subsequent works have sought to enhance the self-supervised framework by incorporating temporal modelling capabilities. SNUG (Santesteban et al. 2022a) introduces several key improvements to address the limitations of frame-independent processing. The method incorporates a more sophisticated elastic potential energy formulation

based on the Saint Venant-Kirchhoff (StVK) constitutive model, which provides a more accurate representation of material behaviour under large deformations compared to the simple spring-based approach. Additionally, SNUG introduces an inertia loss term that models the temporal dynamics of garment motion, ensuring that predicted deformations respect the physical principle of momentum conservation. These enhancements result in substantially improved visual quality and more realistic garment behaviour.

Extending this trajectory, NCS (Bertiche et al. 2022) further advances dynamic modelling through a novel decomposition strategy. The method decouples garment deformation into two physically motivated components: static deformation, which represents the quasi-static equilibrium configuration under gravitational and contact forces for the current frame, and dynamic deformation, which captures the inertia-driven oscillatory behaviour and temporal fluctuations inherent to cloth dynamics. This decomposition strategy allows for independent modelling of equilibrium states and transient dynamics, enabling the framework to better capture the garment dynamics.

2.2.2 Mesh-Driven Methods

While body poses provide valuable guidance for garment deformation prediction, pose-driven methods fundamentally rely on learning static mappings between body configurations and corresponding garment responses. This assumption proves problematic when confronted with complex or previously unseen body poses that deviate significantly from the training distribution. The fixed nature of these learned mappings necessitates the development of dedicated models for each specific garment type, severely limiting the practicality. Although several existing works (Patel et al. 2020; Vidaurre et al. 2020) have attempted to mitigate these limitations, the resulting models often exhibit constrained flexibility and face trade-offs between simulation quality, computational efficiency, and generalisation capability.

Mesh-driven methods, on the other hand, fundamentally address this challenge by leveraging explicit mesh information to model vertex-level dynamics. As shown in Fig. 2.3, These methods treat each vertex in the garment mesh as an individual computational unit, employing

message-passing mechanisms that facilitate information exchange between spatially adjacent vertices. This vertex-centric paradigm enables the network to capture fine-grained local geometric interactions and neighborhood dependencies, allowing for the prediction of detailed per-vertex quantities such as accelerations, velocities, or displacements. By operating at the vertex level, mesh-driven methods demonstrate superior adaptability to varying mesh topologies while maintaining enhanced physical fidelity through more granular modelling of cloth dynamics.

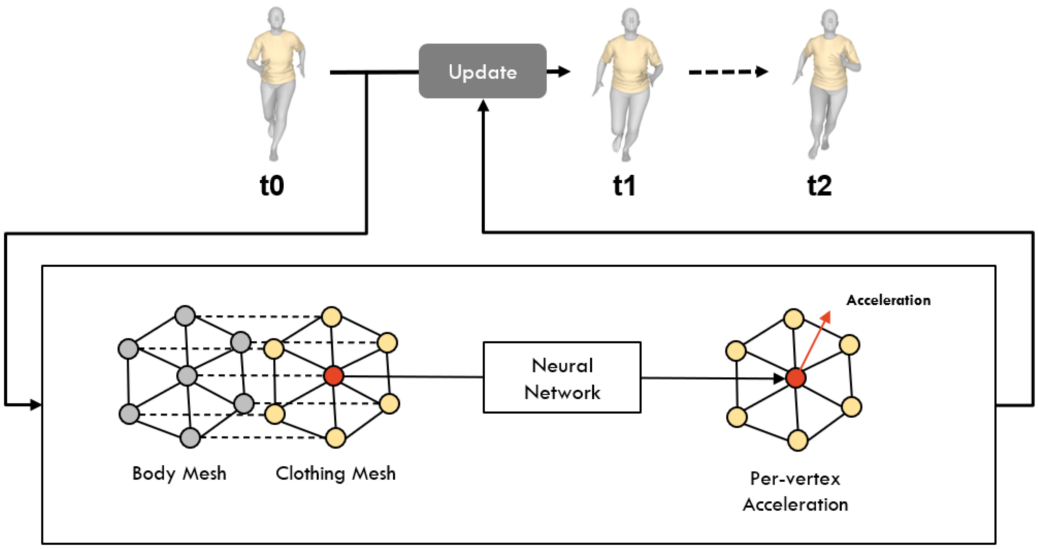

Figure 2.3. General pipeline of mesh-driven methods.

2.2.2.1 Vertex-Level Dynamics Learning

To enable vertex-level dynamics learning, Graph Neural Networks (GNNs) have been widely adopted. A representative mesh-driven approach is MeshGraphNet (Pfaff et al. 2020), a GNN-based simulation framework designed for mesh-based physics simulations. MeshGraphNet follows an encoder–processor–decoder architecture. The encoder transforms vertex-level dynamic inputs, such as velocity, into latent representations. These vertex embeddings are then updated through a message-passing module that facilitates information exchange among neighboring vertices. The updated latent features are decoded to predict the vertex-wise acceleration, which is subsequently used in a standard time integration scheme to update vertex positions.

Beyond GNNs, Transformer-based models (Wu et al. 2024; Li et al. 2024) have also been introduced for vertex-level feature learning. In these models, each vertex applies an attention mechanism to aggregate information from its neighbors. Unlike GNNs, which propagate information step-by-step through local message passing, Transformers can model multi-hop interactions directly via self-attention. However, the quadratic complexity of attention operations with respect to the number of vertices often demands optimisation to improve scalability. For example, Li's work (Li et al. 2024) restricts attention to geodesically close neighbors to reduce computational cost, while Transolver (Wu et al. 2024) proposes a slicing technique that projects mesh features onto structured feature maps before applying attention.

Compared to GNN-based approaches, Transformer-based models generally excel at capturing long-range dependencies but may struggle to represent fine-grained local details, such as wrinkles. Furthermore, they often suffer from high computational and memory costs due to their quadratic attention complexity. As a result, GNN-based methods remain widely adopted in garment simulation due to their strong inductive bias for mesh-based data and their compatibility with existing physical priors.

2.2.2.2 Towards Scalable and Generalisable GNN-Based Simulation

Since GNN-based methods perform vertex-wise feature learning to predict garment dynamics, they typically rely on iterative message passing to propagate information and compute accelerations for each vertex. While effective, this process becomes increasingly expensive as the resolution of the garment mesh grows, due to the larger number of vertices and local interactions that must be processed. Moreover, accurately capturing complex garment behaviours often requires a sufficiently large message-passing range, which further increases computational cost when more message-passing layers are added.

To address these scalability challenges, several works such as Multi-scale MeshGraphNet (Fortunato et al. 2022) and HOOD (Grigorev et al. 2023) introduce hierarchical graph structures to enable multi-level message passing. These methods apply graph pooling to the original mesh to generate coarser representations by selecting subsets of the original vertices. Pooling can be applied recursively to form multiple hierarchical levels. During message passing,

information flows not only across fine-grained vertices but also between higher-level vertices, allowing long-range dependencies to be captured more efficiently. This hierarchical aggregation effectively reduces computational cost while preserving or even enhancing simulation quality.

However, these hierarchical methods typically rely on predefined mesh decimation strategies to construct the graph hierarchy, which assumes the mesh can be meaningfully reduced to lower resolutions—an assumption that does not always hold. Additionally, building the hierarchical structure introduces extra edges into the graph, leading to increased memory consumption during training and inference.

Improving the message-passing efficiency of GNNs is not only a challenge in garment simulation but also a broader issue in the graph learning community. A common strategy in scalable graph learning is to reduce message-passing costs via importance sampling (Chen et al. 2018; Hamilton et al. 2017; Chiang et al. 2019). However, in garment simulation, accurate force and acceleration estimation require precise information from all neighboring vertices, making sampling-based methods less suitable. Another line of work precomputes a single message-passing step using static graph features (Zhang et al. 2022b; Wu et al. 2019), but this approach is also not applicable for garment simulation, where vertex features evolve dynamically over time.

Another key limitation that remains unresolved in GNN-based methods is their sensitivity to mesh resolution. GNNs often struggle to generalise to resolutions not encountered during training. Although MeshGraphNet (Pfaff et al. 2020) uses adaptive remeshing to enable resolution-agnostic simulation, it removes user control over the output resolution. Moreover, the remeshing process itself introduces significant computational overhead, further reducing the efficiency of the simulation.

2.3 Summary

Garment and fabric simulation has long been a central topic in computer graphics. Traditional methods are primarily physics-based, relying on physical laws and numerical solvers to produce realistic results. However, their high computational cost limits efficiency, especially in real-time applications where speed often comes at the expense of quality.

In recent years, machine learning has emerged as a promising alternative for garment simulation. Over the past decade, learning-based approaches have advanced considerably. Early models predicted garment deformations from body pose sequences, supervised by physics-based simulations. Later, self-supervised methods introduced physics-inspired loss functions, improving realism while removing the need for precomputed ground truth data. However, pose-driven models remain limited in generalisability, as they are typically trained for a single garment mesh and require retraining for new geometries.

To overcome these limitations, recent research has shifted towards Graph Neural Networks (GNNs), which have proven highly effective in learning-based physics simulation due to their natural ability to model mesh and particle structures. In garment simulation, GNNs have demonstrated impressive generalisability across different body motions and garment types, along with strong predictive performance. However, current studies reveal two persistent challenges. First, GNNs are computationally expensive and scale poorly to high-resolution meshes. Second, when applied to mesh resolutions unseen during training, existing GNN-based models often struggle to generalise effectively.

In summary, despite the progress of GNN-based methods, scalability and generalisability remain open challenges. More efficient message-passing schemes are required to improve simulation speed, and greater robustness to mesh resolution variations is necessary to extend applicability to unseen garment geometries. These limitations highlight the need for further research into efficient, resolution-agnostic GNN frameworks to make learning-based garment simulation broadly applicable in practice.

CHAPTER 3

Extended Short- and Long-Range Mesh Learning for Fast and Generalised Garment Simulation

---

In this chapter, we introduce our first research study, which addresses the scalability and efficiency challenges of GNN-based garment simulation. To this end, we introduce a novel simulation framework that incorporates two key modules designed to enhance message passing between vertices: the Laplacian-Smoothed Dual Message Passing (LSDMP) module and the Geodesic-Aware Self-Attention (GSA) module. The chapter is structured as follows: Section 3.1 provides an overview of the proposed approach; Section 3.2 details the technical design of the framework; Section 3.3 presents experimental results demonstrating the effectiveness of our method; Section 3.4 summarise the research work and dicuss the future directions.

3.1 Introduction

For garment simulation, classical approaches rely on mass-spring systems (Provot et al. 1995), which approximate the forces acting on the garment and distribute them across the vertices of the corresponding 3D garment mesh. However, the commonly used iterative optimisation-based methods for modelling such forces are computationally expensive, posing challenges for various scenarios such as real-time simulation.

In recent years, deep learning networks have emerged as potential alternatives for approximating garment behaviours with lower latency than classical approaches. Graph neural networks (GNNs) in particular have demonstrated remarkable capabilities in modelling complex physical systems, including garment simulation (Pfaff et al. 2020; Sanchez-Gonzalez et al. 2020). Nevertheless, in the context of garment simulation, GNNs do not inherently address the

iterative optimisation process, as their performances largely depend on their message-passing range, i.e. the geodesic distance over which vertex-wise information can propagate. To achieve high-quality garment behaviours with proper collision and elasticity handling, a sufficiently large message-passing range is essential. However, this often requires numerous GNN layers, which can be computationally inefficient. Therefore, a primary objective in GNN-based methods is to expand each vertex's influence range while minimising the number of message-passing layers. For example, recent studies have explored hierarchical graph structures, which enable parallel message-passing at multiple scales (Grigorev et al. 2023; Fortunato et al. 2022). Although such methods introduce long-range connections between vertices, they rely on predefined decimation strategies and require the cloth mesh to be reducible to lower resolutions, which is not always feasible. This heavily limits their utility when a variety of meshes are necessary, such as in fashion applications.

In this chaper, we propose a novel GNN-based mesh learning framework to enhance the message-passing range in garment simulation without requiring multi-scale mesh definitions. Our framework introduces two key modules as shown in Fig. 3.1: the *Laplacian-Smoothed Dual Message-Passing* (LSDMP) module for extended short-range modelling, and the *Geodesic Self-Attention* (GSA) module to facilitate long-range vertex connections. The LSDMP module employs a dual message-passing design that facilitates message exchange over a wider range. Specifically, each LSDMP layer extends conventional message-passing with a Laplacian-smoothing process for adjacent features aggregation, which efficiently propagates the impact of each vertex to nearby vertices. Complementing the short-range LSDMP module, the long-range GSA module captures global mesh information through attention mechanisms. GSA introduces geodesic distance embeddings to effectively represent the spatial relationships between vertices, enabling the module to focus on long-range vertex interactions.

With the proposed modules, our framework effectively captures both short- and long-range mesh patterns of garments. Comprehensive experiments demonstrate that our method adheres closely to physical laws, while enhancing parameter efficiency and inference latency compared to existing approaches.

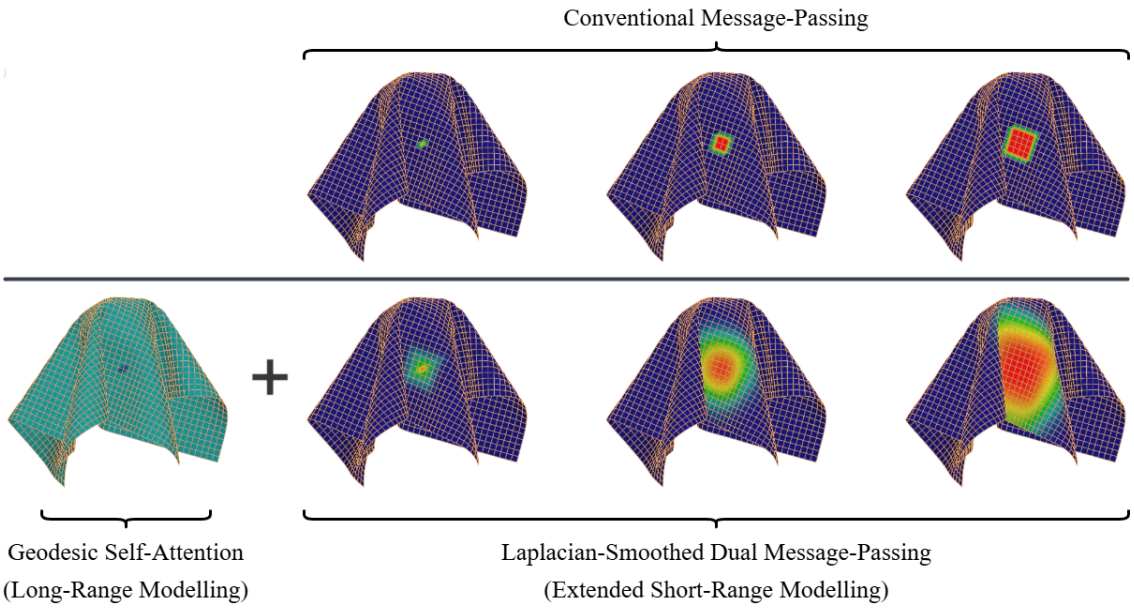

Figure 3.1. Laplacian smoothing-based propagation allows features to reach distant vertices much earlier in an attenuated fashion, and is vastly more efficient than conventional message-passing. Geodesic self-attention further enables graph-aware long-range connections without message-passing mechanisms.

In summary, our key contributions are as follows:

- We propose a novel GNN-based architecture for garment simulation, leveraging short- and long-range connections.
- We devise a *Laplacian-Smoothed Dual Message-Passing* module to extend message-passing of short-range garment features with adjacent vertex-based smoothing.
- We devise a *Geodesic Self-Attention* module to capture long-range features, with geodesic distance embeddings for vertex connections complementary to LSDMP.
- Comprehensive experiments demonstrate our method's superior performance against existing methods, using fewer layers and lower computational latency.

3.2 Methodology

Our method follows a three-stage pipeline consisting of an encoder, a message-passing module, and a decoder (Pfaff et al. 2020). We primarily enhance the message-passing module with two novel components: a *Laplacian Smoothed Dual Message-Passing* module and a

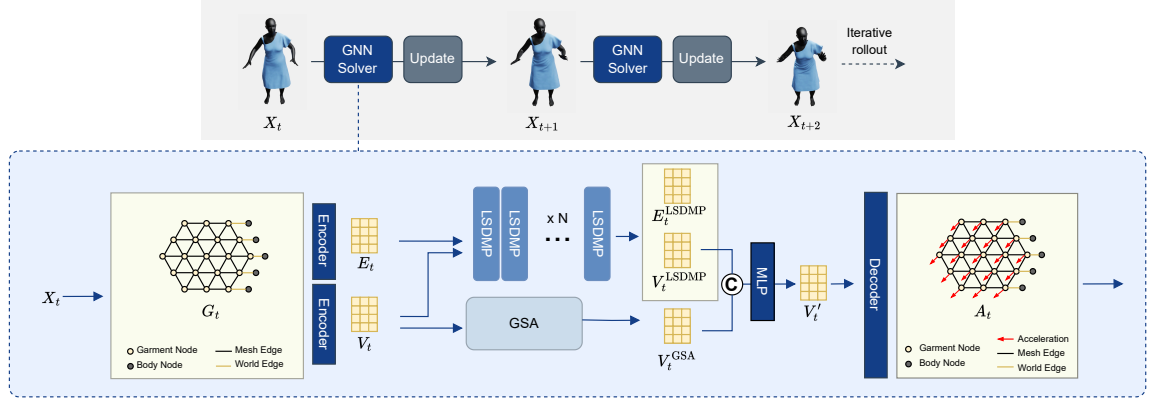

Figure 3.2. Overview of the proposed method for 3D garment simulation. It consists of two novel modules: LSDMP (Laplacian Smoothed Dual Message-Passing) and GSA (Geodesic Self-Attention), to process extended short and long-range mesh learning, respectively.

Geodesic Self-Attention module. In this section, we provide an overview of this pipeline, followed by a detailed explanation of the proposed components.

3.2.1 Garment Simulation Pipeline

Consider a garment mesh M_G with n_G vertices and a human body mesh M_B with n_B vertices in their undeformed states. Garment simulation can be modeled as an iterative process that estimates a sequence of deformed mesh states $X = \{X_0, X_1, ..., X_T\}$. At each timestep t, the state X_t is represented as $X_t = \{X_t^G, X_t^B\}$, where $X_t^G \in \mathbb{R}^{n_G \times 3}$ and $X_t^B \in \mathbb{R}^{n_B \times 3}$ represent the garment and human body mesh states, respectively, each containing their vertex coordinates.

Given the current state X_t, garment simulation estimates the vertex-wise acceleration of the garment mesh A_t, and then uses it to predict the next state of the garment $\hat{X}_{t+1}^G$ as an estimation of X_{t+1}^G. Particularly, X_t is viewed as a graph with its mesh vertices and edges. Similarly to MeshGraphNet (Pfaff et al. 2020), we add world edges between the body and the garment vertices to model collisions. The world edges are based on spatial proximity: if a body and garment vertex pair is within a given distance threshold, a world edge is added between them.

The vertices and edges of X_t are encoded into latent features V_t and E_t, respectively, with two individual encoders. The latent features are then fed into the LSDMP module to perform message propagation. In parallel, the GSA module operates on the latent vertex features for a global understanding in terms of the garment. The two modules produce updated vertex features V_t^{LSDMP} and V_t^{GSA}. These features are concatenated and processed by an MLP to yield the refined latent vertex features V_t'. Finally, the decoder uses V_t' to estimate the vertex acceleration A_t. An Euler integration is applied to update the per-vertex velocity of the garment $Q_{t+1}^G = Q_t^G + A_t \Delta t$, and subsequently, the overall graph for the timestep $t+1$ is obtained by: $X_{t+1}^G = X_t^G + Q_{t+1}^G \Delta t$, where Δt is the associated timestep size.

3.2.2 Laplacian Smoothed Dual Message-Passing

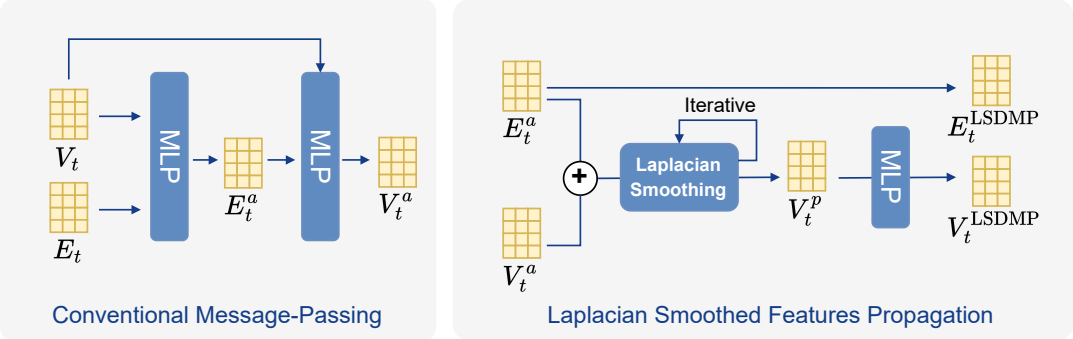

Figure 3.3. LSDMP extends conventional message-passing using Laplacian-smoothing to efficiently and emblematically propagate vertex features.

The Laplacian Smoothed Dual Message-Passing (LSDMP) module takes the edge features E_t and vertex features V_t as input and outputs the updated features V_t^{LSDMP} and E_t^{LSDMP}, respectively. The LSDMP module consists of a sequential of LSDMP layers. As shown in Fig. 3.3, the operations in each layer can be divided into two stages: a conventional message-passing stage (Pfaff et al. 2020) and a Laplacian features propagation stage.

In the message-passing stage, each edge first aggregates features from its two connected vertices, followed by an MLP-based update. For the edge connecting the i^{th} and j^{th} vertices,

the edge features e_{ij}^a can be obtained through:

$$e_{ij}^a = f_e(e_{ij}, v_i, v_j), \tag{3.1}$$

where f_e represents the MLP for the edge-wise update, and v_i and v_j are the corresponding latent vertex features. Next, each vertex compiles information from its connected edges. For the i^{th} vertex, its compiled edge features v_i^a are computed as:

$$v_i^a = f_v(v_i, \sum_j e_{ij}^{a,m}, \sum_j e_{ij}^{a,w}), \tag{3.2}$$

where $e_{ij}^{a,m}$ is the features of the connected mesh edges of the i^{th} vertex and $e_{ij}^{a,w}$ represents the features of the connected world edges. f_v is an MLP for the vertex-wise update.

In the Laplacian features propagation stage, we propagate vertex features to nearby vertices using iterative smoothing operations. Laplacian smoothing is computationally efficient, allowing receptive fields to expand with minimal latency. In each iteration, a vertex aggregates the features from its neighbouring vertices and the connected edges. For the i^{th} vertex, the aggregated features for v_i^p are inferred as:

$$v_i^p = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} (v_j^a + e_{ij}^a), \tag{3.3}$$

where $\mathcal{N}(i)$ is the index set for the vertices and edges connected with the i^{th} vertex. The set of aggregated features for all vertices at time t is denoted as $V_t^p = \{v_i^p \mid i \in \mathcal{V}\}$, where $\mathcal{V}$ represents the set of all vertex indices. Finally, the vertex features are obtained through an MLP f_v':

$$V_t^{\text{LSDMP}} = f_v'(V_t^p). \tag{3.4}$$

3.2.3 Self-Attention with Geodesic Distance Embedding

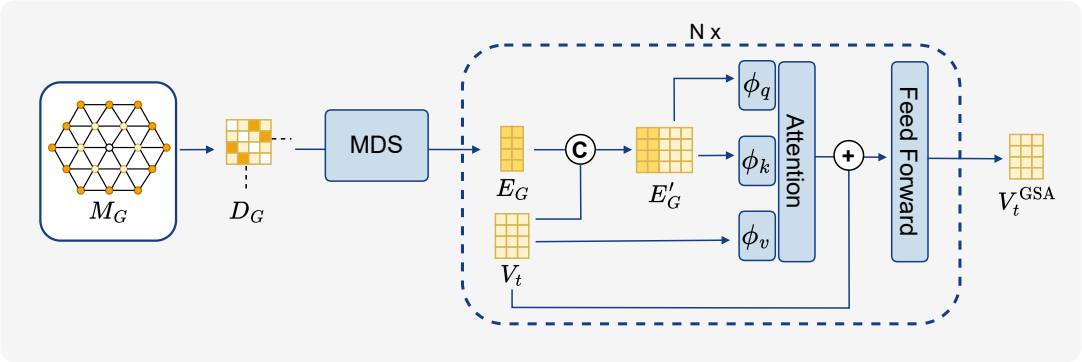

Figure 3.4. GSA injects geodesic information into vertex features by appending MDS-reduced coordinates derived from the geodesic distance matrix.

The LSDMP module enables each garment vertex to gather extended short-range information from its neighbouring vertices; however, capturing long-range dependencies and global context across the entire mesh is essential for producing high-quality garment patterns. To address this, the GSA module, based on self-attention, is designed to enable each vertex to efficiently capture global information as illustrated in Fig. 3.4.

For the self-attention mechanism to function effectively between any two vertices, each vertex must have an understanding of its geodesic relationship with the others. For this purpose, we introduce geodesic distance embeddings into the self-attention. Notably, the quadratic complexity of the standard self-attention mechanism is computationally expensive for garments due to the high vertex counts of high-resolution meshes. Inspired by GraphGPS (Rampášek et al. 2022), we adopt the linear attention approach (Choromanski et al. 2020) to reduce computational overhead, which re-expresses the attention operation using kernel feature maps. This reformulation allows the $n \times n$ attention matrix to be factorised into two smaller projections, reducing the overall complexity from quadratic $\mathcal{O}(n^2)$ to linear $\mathcal{O}(n)$. As a result, our approach maintains the expressive power of self-attention while significantly lowering computational overhead, enabling efficient application to high-resolution garment meshes.

However, linear attention achieves efficiency by avoiding explicit computation of the full pairwise attention matrix, and thus does not natively encode pairwise relationships between

vertices. To compensate for this limitation, we introduce a vertex-wise geodesic embedding that encodes intrinsic distances on the garment surface. By incorporating these embeddings into the attention mechanism, we restore pairwise distance-awareness, ensuring that the model can still capture the geometric context necessary for realistic cloth interactions.

Specifically, for the garment mesh, we pre-compute the geodesic distance matrix $D_G \in \mathbb{R}^{n_G \times n_G}$ of the garment, which contains the pairwise geodesic distances between vertices. To formulate D_G in a vertex-wise manner, we generate the per-vertex k-dimensional geodesic distance embeddings $E_G \in \mathbb{R}^{n_G \times k}$ by applying Multidimensional Scaling (MDS) (Abdi 2007). MDS minimises the following objective:

$$f(E_G) = \sum_{i \neq j} \left(d_{ij} - \|d_i - d_j\|_2\right)^2, \tag{3.5}$$

where d_i, d_j are the generated geodesic embeddings of the i^{th} and j^{th} vertices derived from E_G. d_{ij} represents the geodesic distance between the i^{th} and j^{th} vertices as given by D_G. This optimisation problem ensures that the pairwise Euclidean distances in the embedding space closely approximate the pairwise geodesic distances.

By incorporating geodesic embeddings into the queries and keys, the attention module becomes aware of the geodesic relationships between vertices. The GSA module can be stacked multiple times, with the resulting features denoted as V^{GSA}.

3.2.4 Optimisation

We follow an existing optimisation strategy from HOOD (Grigorev et al. 2023), which employs physical losses for unsupervised learning. Our approach incorporates a comprehensive set of loss functions designed to enforce fundamental physics principles including inertia, material deformation, bending stiffness, gravity, collision avoidance, and surface friction.

The total loss function combines all physical constraints:

$$\mathcal{L} = \mathcal{L}_{\text{inertia}} + \mathcal{L}_{\text{stretch}} + \mathcal{L}_{\text{bending}} + \mathcal{L}_{\text{gravity}} + \mathcal{L}_{\text{collision}} + \mathcal{L}_{\text{friction}}. \tag{3.6}$$

3.2.4.1 Inertia Loss

The inertia loss enforces Newton's first law by penalising deviations from expected motion based on previous velocity. For each garment vertex $\mathbf{w}_t^G \in X_t^G$ at time step t, we compare its actual position with the expected position $\hat{\mathbf{w}}_t^G$ assuming constant velocity:

$$\hat{\mathbf{w}}_t^G = \mathbf{w}_{t-1}^G + \mathbf{v}_{t-1}\Delta t \tag{3.7}$$

The inertia loss is then defined as:

$$\mathcal{L}_{\text{inertia}} = \sum_t \sum_{\mathbf{w}_t^G \in X_t^G} \frac{1}{2\Delta t^2} (\mathbf{w}_t^G - \hat{\mathbf{w}}_t^G)^\top m(\mathbf{w}_t^G)(\mathbf{w}_t^G - \hat{\mathbf{w}}_t^G) \tag{3.8}$$

where $m(\mathbf{w}_t^G)$ represents the mass associated with vertex $\mathbf{w}_t^G$.

3.2.4.2 Stretch Loss

The stretch loss models the elastic deformation energy to prevent unrealistic stretching and compression of the fabric material:

$$\mathcal{L}_{\text{stretch}} = \sum_t \sum_{\mathbf{w}_t^G \in X_t^G} \text{volume}(\mathbf{w}_t^G)\Psi(\mathbf{w}_t^G) \tag{3.9}$$

where $\Psi(\mathbf{w}_t^G)$ is the strain energy density computed using the St. Venant-KirchhoffSantesteban et al. 2022a elastic material model, and $\text{volume}(\mathbf{w}_t^G)$ represents the volume element associated with the vertex.

3.2.4.3 Bending Loss

The bending loss captures the resistance to out-of-plane deformation by modelling bending stiffness between adjacent mesh faces:

$$\mathcal{L}_{\text{bending}} = \sum_{t} \sum_{e \in E(X_t^G)} \frac{k_{\text{bending}}}{2} \theta_{e,t}^2 \tag{3.10}$$

where k_{bending} is the material bending stiffness parameter, $E(X_t^G)$ denotes the set of edges in the garment mesh, and $\theta_{e,t}$ represents the dihedral angle between adjacent faces sharing edge e at time t.

3.2.4.4 Gravity Loss

The gravity loss incorporates gravitational effects by modelling the gravitational potential energy of the garment:

$$\mathcal{L}_{\text{gravity}} = \sum_{t} \sum_{\mathbf{w}_t^G \in X_t^G} -m(\mathbf{w}_t^G) g \cdot z(\mathbf{w}_t^G) \tag{3.11}$$

where $z(\mathbf{w}_t^G)$ denotes the vertical (z-axis) coordinate of vertex $\mathbf{w}_t^G$, $m(\mathbf{w}_t^G)$ is the vertex mass, and g is the gravitational acceleration magnitude.

3.2.4.5 Collision Loss

The collision loss prevents penetration between garment vertices and the human body surface using a penalty-based approach:

$$\mathcal{L}_{\text{collision}} = \sum_{t} \sum_{\mathbf{w}_t^G \in X_t^G} k_{\text{collision}} \cdot \max(\epsilon - d(\mathbf{w}_t^G), 0)^3 \tag{3.12}$$

where $k_{\text{collision}}$ is the collision penalty weight, ϵ is a safety margin parameter, and $d(\mathbf{w}_t^G)$ represents the signed distance from vertex $\mathbf{w}_t^G$ to the nearest body surface (negative values indicate penetration).

3.2.4.6 Friction Loss

The friction loss models the resistance to sliding motion between the garment and body surface. We first define the expected vertex position assuming perfect coupling with the body surface motion:

$$\hat{\mathbf{w}}_t^G = \mathbf{w}_{t-1}^G + \big(\mathbf{p}(f_t(\mathbf{w}^G)) - \mathbf{p}(f_{t-1}(\mathbf{w}^G))\big) \tag{3.13}$$

where $f_t(\mathbf{w}^G)$ identifies the nearest body face to vertex $\mathbf{w}^G$ at time t, and $\mathbf{p}(f)$ returns the centroid position of face f.

The friction loss then penalises deviations from this expected motion:

$$\mathcal{L}_{\text{friction}} = \sum_t \sum_{\mathbf{w}_t^G \in X_t^G} \mu \cdot m(\mathbf{w}_t^G) \cdot g \cdot \cos(\bar{\theta}_t) \cdot \left\| \text{proj}_{\bar{\theta}_t}(\mathbf{w}_t^G - \hat{\mathbf{w}}_t^G) \right\| \tag{3.14}$$

where μ is the friction coefficient, $\bar{\theta}_t$ represents the average surface slope angle between faces $f_{t-1}(\mathbf{w}^G)$ and $f_t(\mathbf{w}^G)$, and $\text{proj}_{\bar{\theta}_t}(\cdot)$ projects the displacement vector onto the tangent plane defined by slope $\bar{\theta}_t$.

3.3 Experimental Results and Discussions

3.3.1 Dataset

For garment simulation experimentation, we employed the AMASS (Mahmood et al. 2019) dataset with its human motion for training and evaluation. In total, 56 motion sequences were selected, containing 6,465 frames. 4 of these sequences were used for evaluation and the rest were used for training. We also used the same garment template set as HOOD (Grigorev et al. 2023) for training, including a T-shirt, tank top, long-sleeved shirt, shorts, long pants and a dress. Additionally, for out-of-domain testing, we employed knee-length pants, a short-sleeve shirt, and a tight dress.

3.3.2 Evaluation Metrics

Following existing practices, we adopt the physics loss terms used in training for evaluation purposes as well. When used as evaluation metrics, the meaning of each term may slightly differ from their meaning in training, as they now serve to quantify the physical realism and accuracy of the simulation results rather than guide the optimisation process.

Inertia loss. The inertia loss measures deviations from expected motion trajectories based on previous velocities, evaluating how well momentum conservation is maintained throughout the simulation.

Stretch loss. The stretch loss is used to evaluate the stretch and compression behaviour of the garment material. Physically correct simulation results always tend to have low strain energy, as real fabrics naturally resist excessive deformation. A high strain energy usually indicates that the garment is either over-stretching or over-compressing beyond realistic material limits. Therefore, the stretch loss serves as an indicator of the physical correctness of the simulation result, with lower values generally corresponding to more plausible material behaviour.

Bending loss. The bending loss evaluates the bending stiffness and out-of-plane deformation characteristics of the simulated garment. This metric quantifies how well the simulation captures the natural resistance of fabric to folding and creasing. Lower bending loss values indicate that the garment maintains appropriate bending behaviour without excessive sharp folds or unrealistic smoothness.

Gravity loss. The gravity loss assesses how realistically the simulation incorporates gravitational effects on the garment. This metric measures the potential energy distribution of the cloth vertices, with lower values indicating that the garment settles into gravitationally stable configurations. Evaluation using this metric helps ensure that simulated garments exhibit natural draping behaviour and respond appropriately to gravitational forces without floating or exhibiting unphysical suspension. Since the loss is calculated as the sum of the vertical coordinates of all vertices (weighted by mass), its value may be negative depending on the position of the garment.

Collision loss. The collision loss serves as a critical metric for evaluating penetration avoidance between the garment and human body. During evaluation, this metric quantifies the extent of interpenetration violations, with higher values indicating more severe collision artifacts. A well-performing simulation should maintain consistently low collision loss, ensuring that clothing remains on the correct side of the body surface throughout the motion sequence.

Friction loss. The friction loss measures sliding behaviour between garment and body surfaces, evaluating the realism of surface interactions. This metric assesses whether the garment exhibits appropriate resistance to sliding motion, avoiding both unrealistic slipping and excessive adherence to body surfaces.

3.3.3 Implementation Details

Our model consists of separate stacks of LSDMP and GSA layers. The LSDMP module consists of 15 stacked layers, with each layer conducting a conventional message-passing step

followed by 3 Laplacian smoothing steps. The GSA module is designed to be lightweight, comprising 4 stacked attention blocks with one single-head self-attention layer. We trained the model over 100,000 iterations as in HOOD (Grigorev et al. 2023), on an RTX 4070 Ti 12GB GPU, completing the process in approximately 16 hours. The training process involves randomly sampling a body pose from a motion sequence and a garment template from the training template set to use as input for each iteration. We compute the linear blend skinning result of the garment for this pose, which serves as the initial garment state (Pavlakos et al. 2019; Grigorev et al. 2023).

3.3.4 Simulation Performance on Native Garments

To evaluate the effectiveness of our method, we provide a comparison to state-of-the-art methods including SSCH (Santesteban et al. 2021), MSN (Liu et al. 2023), MAT (Li et al. 2024), MeshGraphNet (Pfaff et al. 2020) and HOOD (Grigorev et al. 2023) regarding physical errors with the garment meshes in both training and evaluation. SSCH and MSN are pose-conditioned and garment-specific methods, requiring separate models to be trained for each garment. In contrast, MAT, MeshGraphNet and HOOD are generalised methods capable of simulating diverse garments with a single unified model. Note that all methods, including our LSDMP module, are configured with 15 layers, corresponding to a maximum conventional message-passing distance of 15 vertices.

Table 3.1 demonstrates that the proposed method achieves the lowest total loss values among all evaluated methods, highlighting its physical awareness. Specifically, our method shows a significant advantage in stretch energy, a critical metric for maintaining realistic deformation and simulation stability. A lower stretch energy indicates effective constraint modelling, which contributes to accurate and stable garment simulations. This improvement is enabled by the two proposed modules. The LSDMP module expand the message-passing range effectively to allow each vertex to propagate features across a wider range, aiding in the precise modelling on elastic behaviours. The GSA module, on the other hand, allows vertices to effectively capture the acceleration distribution of the entire mesh through the geodesic embeddings, resulting in more stable simulation results. In addition to stretch energy, our method also

Method↓	Stretch↓	Bending↓	Inertia↓	Collision↓	Friction↓	Gravity↓	Total↓
SSCH	2.09E+00	7.20E-03	9.33E-03	8.15E-02	3.14E-03	-2.15E-01	1.97E+00
MSN	1.92E-01	8.17E-03	5.50E-03	6.26E-02	2.82E-03	-2.37E-01	3.39E-02
MAT	1.07E+00	9.41E-03	**4.50E-03**	2.36E-02	2.84E-03	-1.59E-01	9.54E-01
MeshGraphNet	2.77E-01	6.81E-03	5.94E-03	6.08E-02	2.22E-03	**-3.06E-01**	4.68E-02
HOOD	1.42E-01	5.34E-03	5.20E-03	1.96E-02	2.03E-03	-2.84E-01	-1.10E-01
Ours	**9.00E-02**	**4.63E-03**	4.77E-03	**1.49E-02**	**1.79E-03**	-2.49E-01	**-1.33E-01**

Table 3.1. Physical simulation performance on native garments.

Method	Stretch↓	Bending↓	Inertia↓	Collision↓	Friction↓	Gravity↓	Total↓
Baseline	2.77E-01	6.81E-03	5.94E-03	6.08E-02	2.22E-03	**-3.06E-01**	4.68E-02
LSDMP	1.22E-01	4.74E-03	**4.51E-03**	1.50E-02	1.80E-03	-2.52E-01	-1.05E-01
GSA	1.22E-01	6.67E-03	4.58E-03	**1.37E-02**	1.88E-03	-2.51E-01	-1.02E-01
LSDMP + GSA	**9.00E-02**	**4.63E-03**	4.77E-03	1.49E-02	**1.79E-03**	-2.49E-01	**-1.33E-01**

Table 3.2. Ablation study on the LSDMP and GSA modules.

Method	Stretch↓	Bending↓	Inertia↓	Collision↓	Friction↓	Gravity↓	Total↓
Knee-length Pants							
MeshGraphNet	4.52E-01	4.05E-03	8.03E-03	5.50E+00	**2.55E-03**	**-8.50E-01**	5.12E+00
HOOD	4.97E-01	**3.85E-03**	7.83E-03	5.50E+00	2.71E-03	-8.49E-01	5.16E+00
Ours	**1.69E-01**	3.92E-03	**7.75E-03**	**5.41E+00**	2.59E-03	-8.28E-01	**4.76E+00**
Tight Dress							
MeshGraphNet	1.27E-01	6.18E-03	5.28E-03	3.19E-03	2.40E-03	**-2.31E-01**	-8.67E-02
HOOD	1.13E-01	6.20E-03	5.86E-03	**1.45E-03**	2.61E-03	-2.22E-01	-9.28E-02
Ours	**7.43E-02**	**5.47E-03**	**4.89E-03**	1.51E-03	**2.26E-03**	-1.98E-01	**-1.10E-01**
Short-sleeved Shirt							
MeshGraphNet	7.02E-02	3.92E-03	3.06E-03	1.53E-03	1.23E-03	**6.95E-02**	1.49E-01
HOOD	4.94E-02	3.84E-03	2.94E-03	**7.46E-04**	1.23E-03	7.62E-02	1.34E-01
Ours	**3.44E-02**	**3.52E-03**	**2.82E-03**	1.24E-03	**1.12E-03**	8.39E-02	**1.27E-01**

Table 3.3. Physical simulation performance on unseen garments.

achieves significant improvement on collision handling, due to the integration of the GSA module. By incorporating geodesic embeddings, vertices can instantly detect collisions in geodesically close regions and respond without the latency or attenuation typically introduced by message-passing. The ablation study in Table 3.2 affirms these observations. The model with both modules achieve the lowest physics loss, indicating the contributions of the two modules in producing physically accurate simulation results.

We present the qualitative examples in Fig. 4.4, which demonstrate that our method produces more realistic dynamics and finer wrinkle details in more accurate regions compared to other methods. In terms of garment dynamics, pose-conditioned models, such as SSCH and MSN, often cling too closely to the body shape and overreact to its movements, resulting in unrealistic garment dynamics and giving the appearance of garments being unnaturally stiff and thick. As a graph-based method, our method focuses more on the interaction between vertices rather than relying solely on body pose, which enables our method to produce more realistic dynamics. In terms of detail fineness, the results from SSCH, MAT, and HOOD appear overly smoothed, lacking fine-grained wrinkle details, whereas MeshGraphNet and MSN predict such details excessively. Our method finds a strong middle ground with wrinkle intensity, and more importantly, tends to keep wrinkle details contained exclusively within garment regions that should realistically have wrinkles.

We also visualise the stretch energy distribution of the garment mesh for selected frames in Fig. 3.6. It shows that our method consistently predicts lower stretch energy compared to other methods, suggesting more physically stable results. It should be noted, however, that body motions introduce entropy reductions to the cloth's energy state, which is why regions of frequent complex collisions, such as collars, show high stretch energy from physical-based algorithms.

3.3.5 Generalisability to Unseen Garments

We demonstrate our method's capability to adapt to arbitrary garment meshes by evaluating its performance on three unseen garment meshes, including a tight dress, a short-sleeved top, and knee-length pants. The accuracy is compared against MeshGraphNet and HOOD, both of which are capable of mesh-agnostic garment simulation. Table 3.3 shows that our method consistently achieves the lowest total physics loss across all unseen garments. Particularly, it outperforms the other two methods by a large margin in optimising the stretch energy. Furthermore, bending stiffness and friction also observe generally stronger results, with a notable exception being for the knee-length pants garment mesh, likely due to the different

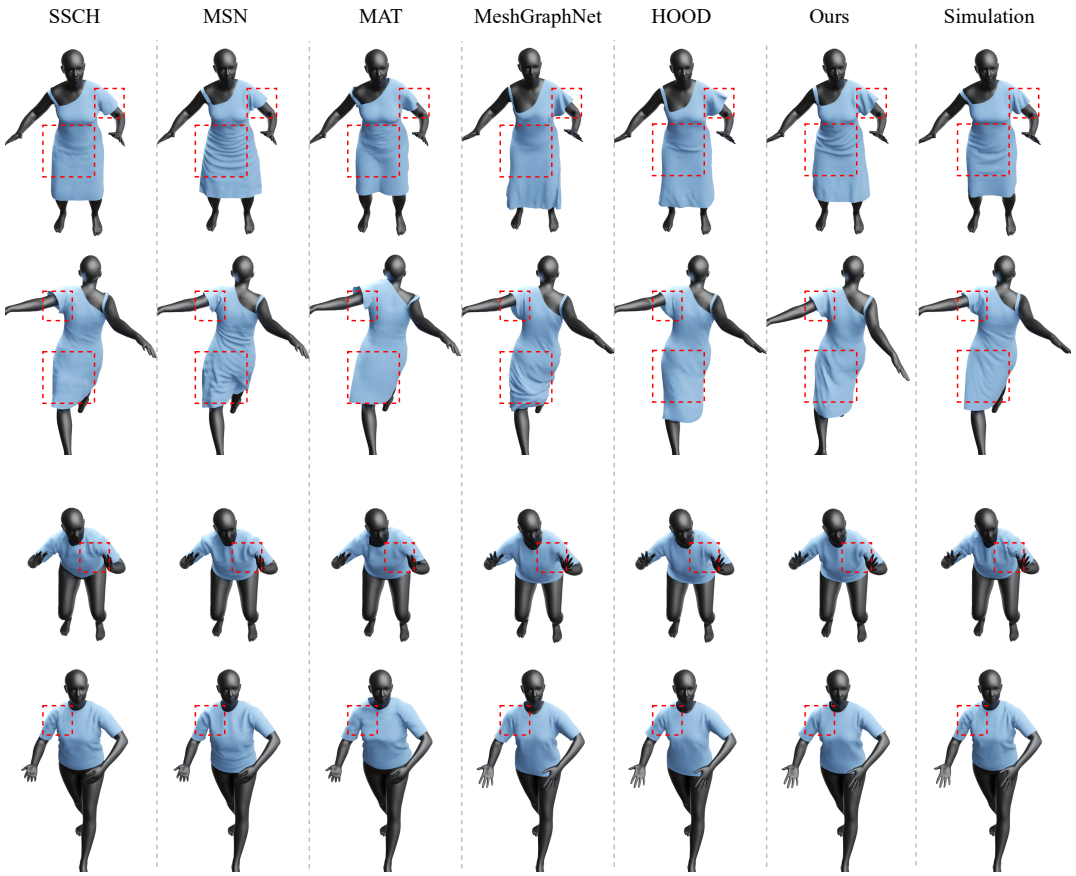

Figure 3.5. Qualitative comparison with SOTA methods. We further include a physical simulation method - ARCSim (Narain et al. 2012) for reference.

level of such forces in knee-length pants compared to the training meshes. This demonstrates the general effectiveness of the attenuative behaviours of LSDMP.

However, it appears that our method's improvements in collision are nullified with unseen meshes. This limitation stems from our geodesic self-attention (GSA) mechanism, which relies on surface-measured distances to guide vertex attention patterns. While this approach works effectively for garments similar to the training set, it struggles with novel geometries where geodesic distance patterns differ substantially. For instance, sleeves, skirt hems, or collars in unseen garments may exhibit very different shapes, lengths, or proportions than their training counterparts. These geometric variations cause the learned distance relationships to misalign with the new garment structure, disrupting the geodesic signal's transferability to unseen garments.

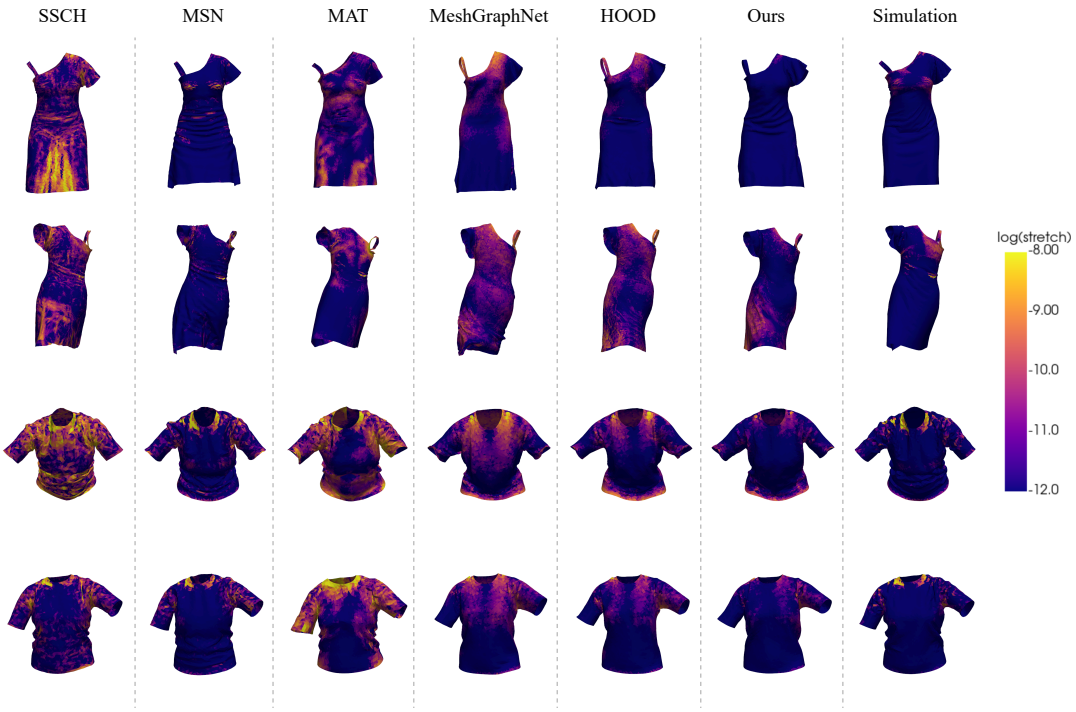

Figure 3.6. Stretch energy distribution.

3.3.6 Model Efficiency and Inference Latency

To demonstrate the layer/parameter efficiency of our method, we compare it against two baseline models based on the conventional message-passing proposed in MeshGraphNet . The baseline models are configured with 15 and 48 layers, whereas our method uses 10 and 15-layer setups. As shown in Table 3.4, both setups of our method significantly outperform the 15-layer baseline in terms of total physics loss, while achieving performance comparable to the 48-layer baseline. Furthermore, our method achieves these results with a substantially lower inference latency. Our 15-layer setup requires only $\sim 40\%$ of the parameters necessary for the 48-layer baseline, highlighting its parameter efficiency. Despite delivering quantitative results comparable to the 48-layer baseline model, our 15-layer method halves the inference latency and requires notably less memory during inference.

Model	# Layers	# Parameters	Total Loss↓	Latency (ms)↓	Memory (MB)↓
Baseline	48	12.0M	**-1.66E-01**	133.8	320
	15	3.85M	4.68E-02	57.8	290
Ours	15	4.75M	-1.40E-01	68.8	292
	10	3.23M	-1.21E-01	**54.0**	**286**

Table 3.4. Inference latency, parameter efficiency, and memory usage compared to conventional message passing.

3.4 Summary

In this chapter, we present a GNN-based garment simulation framework that enhances existing methods by incorporating two new modules: LSDMP and GSA. These modules extend the GNN's message-passing range, enabling the capture of both extended short- and long-range features with minimal computational overhead. Experiments have demonstrated that the self-attention and attenuation features of our method achieve superior simulation quality with a compact model size and low inference latency.

While the proposed method enhances the efficiency of GNN-based garment simulation, the computational cost of GNNs remains significant at ultra high mesh resolutions, potentially limiting its suitability for applications demanding strict real-time inference speeds. A promising direction for future work could involve applying super-resolution techniques to the outputs of low-resolution simulations, rather than directly simulating high-resolution meshes. Additionally, while our GSA module helps mitigate artifacts caused by garment self-collision, it does not fully address the issue, as no proper optimisation measures are currently in place for self-collision. An intriguing avenue for future research would be to incorporate a self-collision objective (Grigorev et al. 2024) to explore whether the GSA module can further minimise self-collision artifacts.

CHAPTER 4

Pb4U-GNet: Resolution-Agnostic Garment Simulation via Propagation-before-Update Graph Network

---

In this chapter, we introduce our second research study, which aims to enhance the generalisability of GNN-based garment simulation methods across varying mesh resolutions. To this end, we propose a novel graph-based framework, the Propagate-before-Update Graph Network (Pb4U-GNet), which enables resolution-agnostic simulation without the need for retraining. The chapter is structured as follows: Section 4.1 provides an overview of the proposed approach; Section 4.2 details the technical methodology; Section 4.3 presents comprehensive experimental results; and Section 4.4 concludes the work and discusses potential directions for future work.

4.1 Introduction

Realistic cloth and garment simulation play a crucial role in many computer vision and graphics applications, including virtual try-on, virtual-reality experiences and digital human modelling. Conventional methods model cloth with physics-based formulations, such as mass-spring systems (Provot et al. 1995), to approximate internal elastic forces and reproduce its dynamic behaviour. These approaches employ numerical integration to advance the simulation forward in time, while an iterative solver enforces physical constraints at each timestep to maintain equilibrium. However, for high-resolution meshes, these repeated constraint-solving iterations become computationally prohibitive, making physics-based techniques expensive for real-time applications.

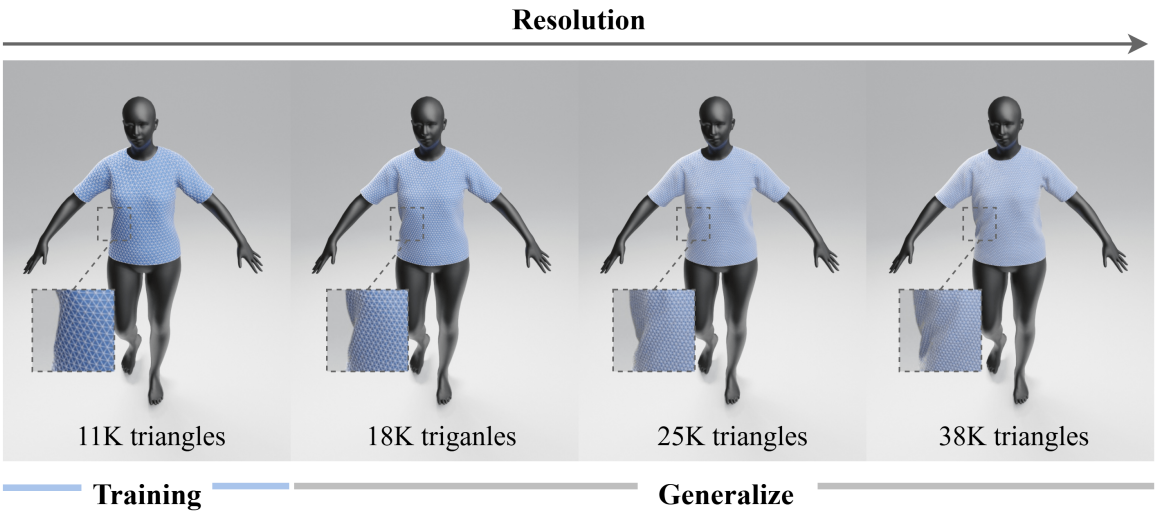

Figure 4.1. Sample results of Pb4U-GNet, a resolution-adaptive garment simulation framework based on graph neural networks. Trained on low resolution meshes with ∼11K triangles, Pb4U-GNet generalises effectively to significantly higher resolutions, producing stable and realistic simulation results without retraining.

To accelerate simulation, deep learning methods have been proposed as alternatives to physics-based solvers (Gundogdu et al. 2019; Santesteban et al. 2019; Pfaff et al. 2020). Among them, graph neural networks (GNNs) (Scarselli et al. 2008) have emerged as powerful predictors of garment dynamics, combining strong generalisability with visually realistic results. GNNs perform message passing: every garment vertex exchanges state information with its neighbours, updates its latent feature, and then predicts its next position via a learnable function.

Despite their promise, current GNN-based garment simulators exhibit poor generalisation across different mesh resolutions. Models trained on a specific mesh resolution frequently fail to produce realistic results when applied to meshes with different densities, particularly those with higher resolution than encountered during training. This fundamentally constrains practical deployment, where mesh resolution must adapt dynamically to varying computational budgets and application-specific requirements. Moreover, training directly on high-resolution meshes is often computationally prohibitive, creating a practical dilemma for achieving robust cross-resolution performance.

We identify two factors that drive this resolution-transfer failure. First, a fixed message-passing depth confines each vertex to a preset hop distance, so on a fine mesh the network sees too little context, while increasing the depth for a coarse mesh over-smooths the result. Second, deformation magnitude scales with mesh density: global motion is distributed across more vertices in finer meshes, leading to diminished per-vertex displacement. This mismatch introduces physically inconsistent predictions on unseen resolutions.

To address these challenges, we propose Propagation-before-Update Graph Network (Pb4U-GNet), a novel framework for resolution-adaptive garment simulation. Pb4U-GNet decouples message propagation from feature update in a two-stage *propagation-before-update* scheme: it first performs iterative message passing to flexibly control the receptive field, then updates all vertex features collectively. Building on this design, we introduce two complementary strategies. First, a *resolution-aware propagation control* strategy that dynamically adjusts the message passing depth based on mesh resolution, maintaining a consistent receptive field across varying discretisation. Second, a *resolution-aware update scaling* strategy rescales the predicted accelerations based on local geometric scale, ensuring physically consistent deformation across meshes of different densities. Extensive quantitative and qualitative evaluations demonstrate that Pb4U-GNet significantly outperforms existing methods in generalising across meshes of varying resolutions, even when trained exclusively on low-resolution inputs. In summary, our key contributions are as follows:

- We propose Pb4U-GNet, a novel Propagation-before-Update Graph Network for garment simulation, capable of generalising to unseen garment resolutions even when trained solely on low-resolution meshes.
- We introduce a resolution-aware propagation control strategy that dynamically adjusts the message passing depth based on mesh resolution, maintaining consistent spatial coverage across varying discretizations.
- We devise a resolution-aware scaling update strategy that normalises the predicted vertex acceleration according to their geometric scale, ensuring physically consistent deformation across meshes of different densities.

4.2 Methodology

As shown in Figure 4.2, the proposed Pb4U-GNet decouples message propagation from feature update in a *propagation-before-update* scheme: it first performs iterative message passing to flexibly control the receptive field, then updates features collectively. To enable resolution-adaptive modelling, a *resolution-aware propagation control* and a *resolution-aware update scaling* mechanisms are devised.

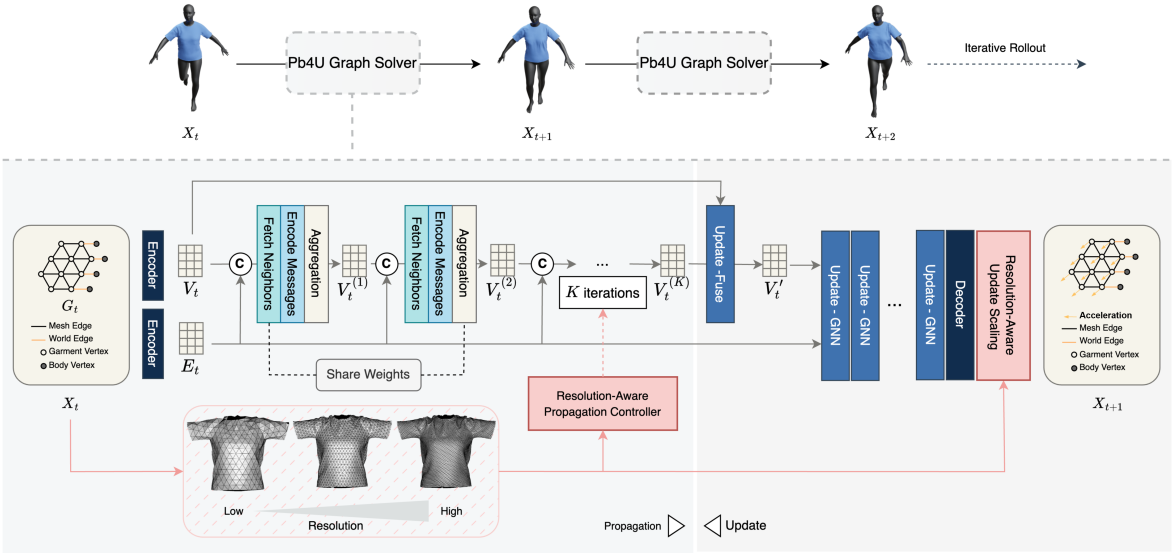

Figure 4.2. Illustration of the proposed proposed Pb4U-GNet, which decouples message propagation with a *propagation-before-update* scheme. With a *resolution-aware propagation control* and a *resolution-aware update scaling* design, it enables resolution-adaptive garment simulation.

4.2.1 Graph Representation

Let $\mathbf{X}_g \in \mathbb{R}^{n_g \times 3}$ and $\mathbf{X}_b \in \mathbb{R}^{n_b \times 3}$ represent the garment and body mesh vertices, with n_g and n_b vertices, respectively. Garment simulation can be formulated as an autoregressive prediction problem over a sequence of mesh states $\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_n)$, where each state $\mathbf{X}_t = (\mathbf{X}_{g,t}, \mathbf{X}_{b,t}, \mathcal{E}_t)$ represents the positions of both garment and body vertices at time step t. For vertex connectivity $\mathcal{E}_t$, in addition to the mesh topology, we follow the approach of Pfaff et al. 2020 and incorporate garment-body interactions by introducing *world edges* between

garment and body vertices based on spatial proximity. Specifically, a world edge is added between a garment vertex $x_i \in \mathbf{X}_{g,t}$ and a body vertex $x_j \in \mathbf{X}_{b,t}$ if the Euclidean distance between them is below a predefined threshold.

The vertex features include physical and geometric attributes such as velocity, mass, surface normals, material parameters, and a vertex-type indicator (garment or body). Notably, absolute spatial information, such as vertex positions, is excluded to ensure invariance to global translation. The edge features encode relative geometric relationships between connected vertices. These include: (1) the current relative direction vector between the two vertices; (2) the relative direction vector in the rest (undeformed) state; and (3) the relative edge length, defined as the ratio between the current and rest length. To promote resolution-adaptive modelling, absolute edge lengths are avoided. Instead, all edge features are formulated in relative terms, enabling the model to generalise across garment meshes with varying resolutions and edge densities.

4.2.2 Problem Formulation & Pb4U-GNet

Given the current mesh state $\mathbf{X}_t$, the objective is to predict the garment deformation at the next time step, denoted as $\hat{\mathbf{X}}_{g,t+1}$, which approximates the future garment state $\mathbf{X}_{g,t+1}$. This prediction can be modelled by a function f_θ with learnable parameters in θ:

$$\hat{\mathbf{X}}_{g,\,t+1} = f_\theta(\mathbf{X}_{g,\,t},\, \mathbf{X}_{b,\,t}). \tag{4.1}$$

Specifically, for $\mathbf{X}_t$, a vertex encoder first maps the raw vertex features into latent embeddings $\mathbf{V}_t$, while an edge encoder transforms edge features into latent embeddings $\mathbf{E}_t$. These encoded representations serve as the initial inputs to the message propagation and graph update. With the Pb4U design, each vertex aggregates information from its spatial and topological neighbours through message-passing. This process is performed iteratively for K steps, producing intermediate embeddings $\mathbf{V}'_t$. The number of message-passing steps K is dynamically determined by a resolution-aware propagation control mechanism: for high-resolution meshes, K is increased to ensure sufficient receptive field coverage; for

low-resolution meshes, it is reduced to improve computational efficiency and prevent over-smoothing from irrelevant distant information.

The intermediate embeddings $\mathbf{V}'_t$ are subsequently refined by a stack of conventional GNN layers (Pfaff et al. 2020), producing final latent representations $\mathbf{V}''_t$, which are then processed through a vertex decoder to predict vertex-wise garment accelerations $\tilde{\mathbf{A}}_{g,t}$. This acceleration is then scaled by a vertex-wise resolution-aware factor $\mathbf{S}$:

$$\mathbf{A}_{g,t} = \mathbf{S} \odot \tilde{\mathbf{A}}_{g,t} \tag{4.2}$$

Finally, the garment mesh state at the next time step, $\hat{\mathbf{X}}_{g,t+1}$, is obtained by applying forward Euler integration. Specifically, the vertex velocities $\mathbf{U}_g$ are first updated as $\mathbf{U}_{g,t+1} = \mathbf{U}_{g,t} + \mathbf{A}_{g,t}\Delta t$, where Δt is the time step size. Then, the vertex positions are computed as $\hat{\mathbf{X}}_{g,t+1} = \mathbf{X}_{g,t} + \mathbf{U}_{g,t+1}\Delta t$.

4.2.3 Propagation-before-Update

To support resolution-adaptive receptive fields, we propose a decoupled message-passing scheme that separates message propagation from feature updates. Given initial vertex and edge embeddings, messages are recursively accumulated over K steps to expand the receptive field. A single update is then applied using the accumulated messages. This design allows K to be adjusted by resolution: higher values for for high-resolution meshes to capture long-range dependencies, and smaller for coarse meshes to reduce overhead.

Formally, during the propagation stage, each vertex maintains an aggregated feature vector $\mathbf{h}_{t,i}$, which is initially set to the vertex feature embedding $\mathbf{v}_{t,i}$. At each aggregation step k, the i^{th} garment vertex aggregates information from its neighbouring vertices and edges to compute an intermediate aggregated feature vector $\tilde{\mathbf{h}}^k_{t,i}$, defined as:

$$\tilde{\mathbf{h}}^k_{t,i} = \text{LayerNorm}\left(\sum_{j \in \mathcal{N}(i)} f_m\left(\mathbf{h}^{k-1}_{t,i}, \mathbf{h}^{k-1}_{t,j}, \mathbf{e}_{t,ij}\right)\right), \tag{4.3}$$

where $\mathbf{h}_{t,i}^{k-1}$ and $\mathbf{h}_{t,j}^{k-1}$ are the previous aggregated feature embeddings of the i^{th} vertex and its neighbour; $\mathbf{e}_{t,ij}$ represents the edge features between them, and $\mathcal{N}(i)$ represents the set of vertex indices neighbouring the i^{th} vertex. The learnable message propagation function $f_m(\cdot)$, implemented as a multi-layer perceptron (MLP), encodes input features from each neighbour into a message vector. These vectors are summed and then normalised using LayerNorm to produce the intermediate aggregated representation $\tilde{\mathbf{h}}_{t,i}^{k}$. This is then combined with the previous aggregated feature vector via a decay-based accumulation:

$$\mathbf{h}_{t,i}^{k} = \gamma \cdot \mathbf{h}_{t,i}^{k-1} + \tilde{\mathbf{h}}_{t,i}^{k}, \tag{4.4}$$

where γ is a decay factor that controls the impact of earlier aggregated messages. The decay factor γ in equation 4.4 draws inspiration from cloth physics: as elastic waves travel through fabric, they naturally weaken due to damping and energy redistribution. We apply this same principle to our recursive message propagation, where γ gradually reduces the influence of older aggregated messages. This approach prevents messages from accumulating uncontrollably while maintaining important long-range contextual information. Through empirical testing, we set γ to 0.9, which provides an effective balance between system stability and information retention.

After K propagation steps, the accumulated message features for each vertex are combined with the original vertex embeddings using a learnable update function f_u, which fuses the original and propagated features into a unified latent representation $\mathbf{v}'_{t,i}$, defined as:

$$\mathbf{v}'_{t,i} = f_u(\mathbf{v}_{t,i}, \mathbf{h}_{t,i}^{K}), \tag{4.5}$$

where $\mathbf{v}_{t,i}$ denotes the original vertex embedding of the i^{th} vertex, and $\mathbf{h}_{t,i}^{K}$ is the aggregated feature vector after K iterations of message propagation. The update function $f_u(\cdot)$ is also implemented as an MLP, and the set of updated vertex features for all garment vertices is denoted as $\mathbf{V}'_t = \{\mathbf{v}'_{t,i} \mid i \in \mathcal{V}\}$ where $\mathcal{V}$ represents the set of all vertex indices.

The updated graph with $\mathbf{V}'_t$ is further processed by a graph neural network (GNN) using a fixed number of layers, which refines the vertex features based on mesh connectivity and

produces the final vertex embedding $\mathbf{V}_t''$. These final embeddings are then processed through a vertex decoder to predict vertex-wise garment accelerations.

By decoupling aggregation from update, we allow flexible control over the receptive field size without entangling it with update frequency, enabling better adaptation to a wide range of mesh topologies and resolutions.

4.2.4 Resolution-Aware Propagation Control

We define D as the effective physical propagation distance calibrated from the base resolution (the lowest-resolution meshes used in training). Concretely, we first determine the number of propagation steps K_{base} required for stable simulation at the base resolution. Given the average edge length $\overline{L}_{base}$ of the garment meshes at this resolution, we set

$$D = K_{base} \times \overline{L}_{base} \tag{4.6}$$

This establishes a physically consistent distance that is resolution-independent. For higher-resolution meshes, the number of propagation steps is then computed as

$$K = \lfloor D \times \bar{L}^{-1} \rfloor \tag{4.7}$$

where $\bar{L}$ is the average edge length of the input mesh in world space. As the mesh resolution increases (i.e., $\bar{L}$ decreases), this formulation proportionally increases K, thereby preserving a consistent physical receptive field across varying resolutions.

4.2.5 Resolution-Aware Update Scaling

Although the decoupled message propagation mechanism allows flexible control of the receptive field, it does not inherently ensure resolution-adaptive predictions. This is because physical quantities such as displacement or acceleration depend on mesh resolution: in high-resolution meshes, each vertex represents a smaller area and mass, resulting in smaller per-vertex accelerations under the same global deformation. As a result, models trained on coarse meshes may overestimate displacements when applied to finer meshes.

Therefore, we introduce a resolution-aware update scaling mechanism based on per-vertex edge length. Motivated by geometric similarity principles in continuum mechanics, where displacement fields scale linearly with element size, we compute a vertex-specific scaling factor as the average length of its connected edges:

$$\mathbf{s}_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} l_{ij}, \tag{4.8}$$

where $s_i \in \mathbf{S}$, which is used in formula 4.2; l_{ij} represents the Euclidean length of the edge connecting vertex i and the neighbour vertex with index j at rest state; $\mathcal{N}(i)$ represents the vertex indices neighbouring vertex i. This scaling restores physically consistent displacement magnitudes by compensating for resolution-dependent geometric variation, enabling the model to learn resolution-invariant deformation behaviour and generalise across meshes of varying densities.

It is worth mentioning that our update scaling is not intended as a full constitutive fabric model. Instead, it serves as a first-order approximation motivated by geometric similarity principles. The scaling module ensures resolution-consistent magnitudes, while the learned network accounts for material-specific nonlinearities.

4.2.6 Physics-Based Supervision

Similar to Grigorev et al. 2023, we train our model in a fully self-supervised manner using six physics-based loss terms, each governing a different aspect of garment dynamics: *stretch*, *bending*, *collision*, *gravity*, *friction*, and *inertia*. In detail, we have:

(1) Stretch loss $\mathcal{L}_{\text{stretch}}$ measures stretching and compression energy using the St. Venant–Kirchhoff model, encouraging the garment to maintain realistic material properties and avoid excessive deformation; (2) Bending loss $\mathcal{L}_{\text{bending}}$ penalises curvature between adjacent mesh faces, promoting appropriate stiffness and preventing unnatural folding; (3) Collision loss $\mathcal{L}_{\text{collision}}$ quantifies garment–body interpenetration as the sum of penetration depths across intersecting vertices, enforcing physical separation; (4) Gravity loss $\mathcal{L}_{\text{gravity}}$ encourages natural draping by penalising vertically raised vertices to simulate gravity; (5) Friction loss $\mathcal{L}_{\text{friction}}$

penalises tangential motion at garment–body contact points to reduce unrealistic sliding; and (6) Inertia loss $\mathcal{L}_{\text{inertia}}$ promotes temporal coherence by penalising abrupt velocity changes across timesteps, preserving physical momentum. The composite loss function is formulated as:

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{\text{stretch}} + \mathcal{L}_{\text{bending}} + \mathcal{L}_{\text{collision}} \\ + \mathcal{L}_{\text{gravity}} + \mathcal{L}_{\text{friction}} + \mathcal{L}_{\text{inertia}}.\end{aligned} \tag{4.9}$$

This allows the model to learn physically accurate garment dynamics without relying on ground-truth supervision. This physics-based formulation enables flexible training across diverse motion types and mesh resolutions.

4.3 Experiments

4.3.1 Experimental Setup

Dataset. We use the VTO dataset (Santesteban et al. 2019), which includes a diverse set of human motion sequences; four are held out for testing, and the remainder are used for training. The training set comprises four garment types (T-shirt, tank top, long-sleeve shirt, and long dress), each with five mesh resolutions ranging from 11K to 38K triangles. Only the lowest resolution is used for training, while higher resolutions are reserved for evaluation to assess resolution generalisation.

Evaluation Metric. Since our method is self-supervised, we assess the physical plausibility of the simulations using the same physics-based loss terms applied during training following the same setting in existing studies Grigorev et al. 2023.

Training. During training, we randomly sample individual frames from the motion sequences in the training set. For each frame, the garment mesh is first deformed with linear blend skinning (LBS) Santesteban et al. 2019 to match the corresponding SMPL body pose, providing a plausible but coarse initial state that our model subsequently refines. The mode is trained for 100,000 iterations, which takes $\sim$24 hours on an NVIDIA RTX 4070 Ti GPU.

Resolution	Model	Stretch	Bending	Collision	Inertia	Gravity	Friction	Total
Lv.1 (11K)	MGN	5.30E-02	3.39E-03	1.71E-02	2.02E-03	-7.20E-02	1.25E-03	4.70E-03
	HOOD	6.83E-02	2.34E-03	2.16E-02	1.80E-03	-8.59E-02	1.30E-03	9.45E-03
	ESLR	2.97E-02	1.81E-03	1.99E-04	1.72E-03	-6.01E-02	1.09E-03	**-2.56E-02**
	CCRAFT	1.22E-01	2.96E-03	1.14E-06	1.91E-03	-8.56E-02	1.19E-03	4.24E-02
	Pb4U-GNet (Ours)	3.06E-02	2.92E-03	1.62E-03	1.78E-03	-5.49E-02	1.34E-03	-1.66E-02
Lv.2 (18K)	MGN	4.13E-01	4.80E-03	1.12E-01	3.66E-03	-1.02E-01	1.36E-03	4.32E-01
	HOOD	3.56E-01	4.67E-03	1.13E-04	2.83E-03	-1.15E-01	1.36E-03	2.49E-01
	ESLR	8.32E-02	2.08E-03	5.60E-02	1.89E-03	-8.38E-02	1.17E-03	6.06E-02
	CCRAFT	1.89E-01	4.12E-03	6.10E-04	1.88E-03	-8.68E-02	1.20E-03	1.10E-01
	Pb4U-GNet (Ours)	4.17E-02	4.23E-03	1.21E-02	1.72E-03	-5.30E-02	1.43E-03	**8.13E-03**
Lv.3 (25K)	MGN	1.43E+03	1.23E-01	3.39E+00	1.08E-02	-3.66E-01	1.47E-03	1.44E+03
	HOOD	3.89E-01	4.88E-03	2.40E-04	2.93E-03	-1.20E-01	1.35E-03	2.78E-01
	ESLR	2.27E-01	2.52E-03	4.64E-02	2.10E-03	-1.07E-01	1.27E-03	1.73E-01
	CCRAFT	2.51E-01	4.59E-03	4.37E-06	1.82E-03	-8.68E-02	1.22E-03	1.72E-01
	Pb4U-GNet (Ours)	5.70E-02	5.83E-03	4.69E-02	1.70E-03	-4.95E-02	1.50E-03	**6.34E-02**
Lv.4 (38K)	MGN	1.24E+06	1.03E+01	8.22E+00	4.07E-02	-2.55E+00	1.70E-03	1.24E+06
	HOOD	2.52E+00	1.56E-02	1.97E-01	5.08E-03	-1.78E-01	1.65E-03	2.57E+00
	ESLR	1.07E+05	9.14E-01	2.46E+01	9.95E-03	-3.25E-02	1.66E-03	1.07E+05
	CCRAFT	3.60E-01	6.13E-03	3.84E-04	1.87E-03	-8.72E-02	1.24E-03	2.82E-01
	Pb4U-GNet (Ours)	1.30E-01	1.43E-02	1.03E-01	1.64E-03	-2.81E-02	1.68E-03	**2.22E-01**

Table 4.1. Performance comparison with state-of-the-art methods across different mesh resolutions, evaluated using physics-based loss metrics.

Model Implementation. Each vertex and edge in the input graph is first encoded to a 128-dimensional latent space, which serves as the initial feature representation. During message propagation, both the message and update functions are implemented as MLPs with two hidden layers of 128 units each. After propagation, the features are further processed by a GNN comprising 15 MeshGraphNet blocks Pfaff et al. 2020, which refine the aggregated vertex embeddings.

4.3.2 Generalisation to Unseen Resolutions

Quantitative Evaluation. To evaluate the generalisation ability of our model across varying mesh complexities, we assess simulation accuracy on four garment resolutions, with the average triangle count ranging from level 1 (lowest) to level 4 (highest). The corresponding mesh resolutions are approximately: Level 1 - 11K triangles, Level 2 - 18K, Level 3 - 25K and Level 4 - 38K. All models are trained exclusively on garments with the lowest resolution (11K), while the other resolutions remain unseen during training. Evaluation is conducted on held-out motion sequences to ensure an unbiased assessment of generalisation performance.

Table 4.1 summarises the evaluation results, reporting individual physics-based metrics along with the total loss. The physics loss metrics are treated as the residuals of the cloth energy minimisation during training. Lower values therefore indicate closer convergence to a physically valid state, providing a measure of physical plausibility. We compare our method against four state-of-the-art graph-based simulators: MGN (Pfaff et al. 2020), HOOD (Grigorev et al. 2023), ESLR (Liu et al. 2025) and CCRAFT (Grigorev et al. 2024). It can be observed that, at the lowest resolution (11K), our method performs comparably to existing approaches. However, as the resolution increases beyond the training level, it significantly outperforms the other methods. In particular, existing methods exhibit clear signs of divergence at high resolution, most notably in the stretch loss. These results highlight the robustness and scalability of our approach in handling diverse mesh resolutions, especially in challenging high-resolution scenarios.

Figure 4.3 demonstrates the temporal evolution of the log stretching energy loss for each method. The simulations are conducted on the test sequence 07_02, covering four different garment types under the highest resolution setting (38K). Low and stable stretch energy is essential for realistic garment simulation as it ensures the fabric behaves like real cloth, maintaining its original shape without unnatural stretching or shrinking over time. As shown, MGN and ESLR exhibit exploding stretch energy when simulating the dress and long-sleeve. Other methods remain more stable, while ours consistently achieves the lowest stretching energy across all garments, highlighting its robustness at high resolution.

Qualitative Evaluation. Rows (a)–(d) of Figure 4.4 show visual comparisons of different methods on various motion frames using trained garment types at the highest resolution setting (38K). Physics-based simulation results are included as reference. Existing methods often struggle to preserve realistic garment behaviour: MGN shows severe distortions and topological artefacts (e.g., tearing, collapsing); HOOD and ESLR preserve the structure better but often exhibit slipping and misalignment, especially around the shoulders; CCRAFT performs best among baselines but still produces overstretched garments, lacking fine details like wrinkles.

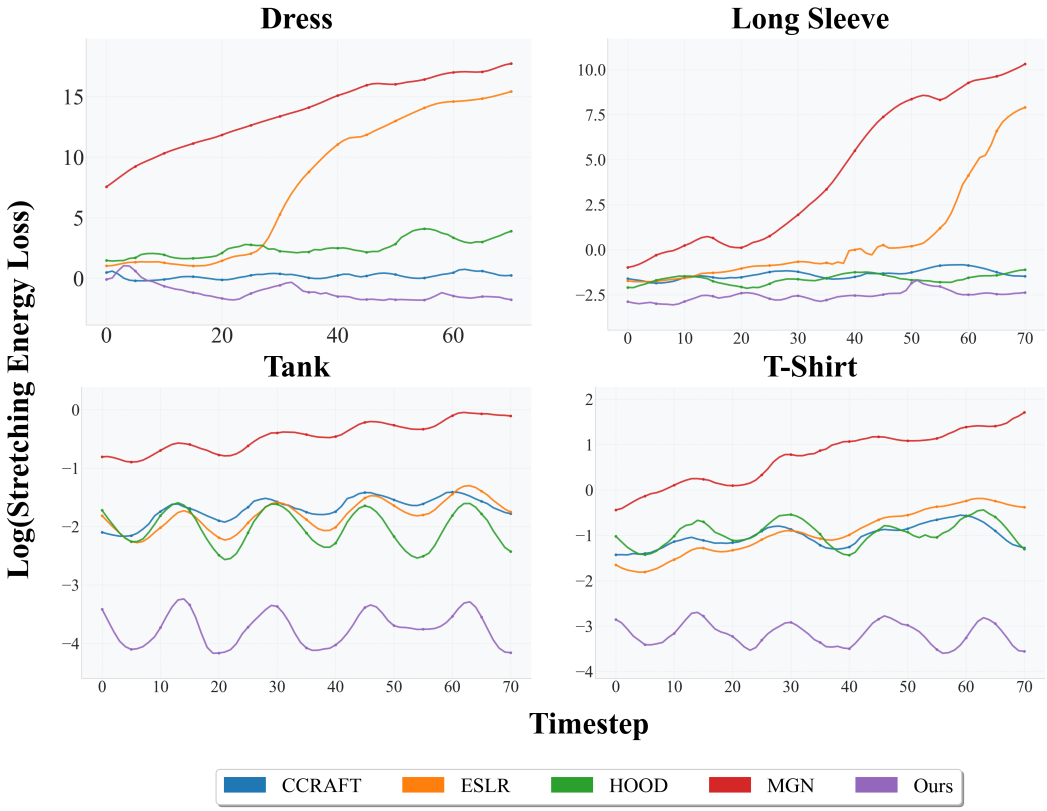

Figure 4.3. Stretch loss vs. time. The plots illustrate the temporal evolution of log stretching energy for each method on the test sequence 07_02. Our method consistently maintains the lowest stretch energy across the simulations, demonstrating better physics validity.

Model	Stretch	Bending	Collision	Inertia	Gravity	Friction	Total
MGN	7.16E+04	8.63E+00	3.40E+01	3.80E-02	-2.08E+00	2.52E-03	7.16E+04
ESLR	1.49E+00	1.23E-02	8.03E-05	2.15E-03	-1.21E-01	1.96E-03	1.39E+00
HOOD	1.87E+00	1.77E-02	2.06E-04	2.57E-03	-1.52E-01	1.75E-03	1.74E+00
CCRAFT	3.11E-01	1.03E-02	4.63E-04	1.88E-03	-6.16E-02	1.79E-03	2.64E-01
Ours	1.05E-01	2.13E-02	3.05E-02	1.57E-03	-5.33E-03	2.33E-03	**1.55E-01**

Table 4.2. Quantitative evaluation on unseen garments under the highest resolution setting.

In contrast, our method consistently generates realistic, physically plausible deformations that closely match the physics simulated results. Garments remain well-fitted, structurally intact, and exhibit realistic wrinkles, demonstrating strong robustness and generalisation under high-resolution, challenging scenarios.

4.3.3 Generalisation to Unseen Garments

To assess the generalisability of our model on unseen garment categories, we evaluated two novel garment types: a form-fitting dress and a cardigan. All evaluations were conducted at the highest resolution. Table 4.2 shows the quantitative comparison between our method and state-of-the-art approaches. Our method achieves the lowest physics loss, confirming its robust generalisability to novel garments.

We present the rendered outputs for both garments in rows (e) and (f) of Figure 4.4. Our method demonstrates better alignment with the physics simulation result, preserving realistic wrinkles and fabric dynamics. While competing methods display over-stretching and misalignment artefacts, as observed in the previous qualitative results.

4.3.4 Propagation Depth Matters: Adapting to Mesh Complexity

Figure 4.5 shows how total physics loss varies with the number of message propagation steps across three mesh resolutions (12K, 25K, 48K) using a dress template. Physics loss generally decreases with more steps, with higher-resolution meshes requiring more to converge, highlighting the link between resolution and receptive field size. Finer meshes need broader receptive fields to capture long-range interactions and maintain physical accuracy. Our method addresses this by adjusting the number of steps based on resolution, improving both efficiency and scalability.

In Table 4.3, we present the per-frame runtime efficiency of our method compared to baseline models that use a fixed number of message propagation steps. For low-resolution meshes, our method adaptively reduces the number of propagation steps, resulting in improved computational efficiency while maintaining comparable simulation accuracy. For higher-resolution meshes, the model increases the number of steps as needed to preserve accuracy, achieving the lowest physics loss among all methods. These results highlight the strength of our adaptive propagation module, which dynamically allocates computation based on mesh complexity and incurs additional cost only when necessary.

Resolution	Model	Physics Loss	Latency (ms)
12K	MGN	-4.22E-01	**46.4**
	HOOD	**-4.38E-01**	50.8
	ESLR	-4.33E-01	54.6
	CCRAFT	-3.53E-01	97.2
	Pb4U-GNet (Ours)	-4.11E-01	50.0
25K	MGN	-3.65E-01	81.8
	HOOD	-3.81E-01	**81.5**
	ESLR	-4.11E-01	98.2
	CCRAFT	-2.10E-01	171.4
	Pb4U-GNet (Ours)	**-3.84E-01**	105.3
48K	MGN	5.75E+03	144.5
	HOOD	6.40E-01	**141.1**
	ESLR	1.38E-01	184.8
	CCRAFT	2.13E-01	761,499.3
	Pb4U-GNet (Ours)	**-2.42E-01**	196.4

Table 4.3. Inference efficiency vs. simulation accuracy across mesh resolutions.

Figure 4.6 further illustrates this adaptive behaviour by plotting cumulative propagation time against step count across different mesh resolutions. Red markers denote the automatically selected step counts: 10 steps for 12K meshes, 15 steps for 25K meshes, and 20 steps for 48K meshes. This progressive scaling demonstrates how our adaptive module efficiently allocates computational resources proportional to mesh complexity.

4.3.5 Ablation Study

Model	Lv.1 (11K)	Lv.3 (25K)	Lv.4 (38K)
Pb4U-GNet (Ours)	**-1.66E-02**	**6.34E-02**	**2.22E-01**
w/o Propagation Control	-1.61E-03	1.08E+06	1.08E+09
w/o Update Scaling	-5.78E-03	1.55E+13	7.34E+13
w/o Both	4.70E-03	1.44E+03	1.24E+06

Table 4.4. Ablation study.

We conduct an ablation study to evaluate the impact of the resolution-aware propagation control and update scaling modules. Table 4.4 reports total physics loss across 11K, 25K, and 38K mesh resolutions. When the propagation control module is removed, performance remains stable at 11K resolution but leads to sharp performance drops at higher resolutions,

showing that fixed receptive fields fail to capture long-range dependencies in finer meshes. Similarly, disabling update scaling causes significant degradation at 25K and 38K, highlighting the need to adapt predicted dynamics to mesh resolution. These results confirm that both components are essential for robust, resolution-adaptive simulation.

4.4 Summary

We presented Pb4U-GNet, a novel framework for resolution-agnostic garment simulation. By decoupling message propagation from feature updates, and introducing resolution-aware propagation control and update scaling mechanisms, our method achieves physically consistent results across meshes of varying densities. Extensive experiments demonstrate that Pb4U-GNet outperforms existing approaches in both accuracy and generalisation, making it well-suited for real-world applications involving diverse garment resolutions.

While our method exhibits strong generalisation across different mesh resolutions, it may underperform in scenarios involving high-resolution garments and rapid body motion. This limitation likely arises from the absence of body information during message propagation, which can impair effective collision handling. Future work could address this by incorporating body vertices into the message-passing process to enhance robustness. Additionally, our current approach employs a predefined rule to control the number of message propagation steps, without accounting for the dynamic state of the mesh. A promising direction for future research is to develop a learnable controller that adaptively adjusts the receptive field based on both mesh resolution and deformation complexity - allocating additional propagation steps only when necessary.

Figure 4.4. Rendered simulation results on high-resolution garment meshes. Baseline methods often struggle to preserve realistic fabric stretch, leading to noticeable over-stretched artefacts. Baseline methods also fail to preserve realistic wrinkle details.

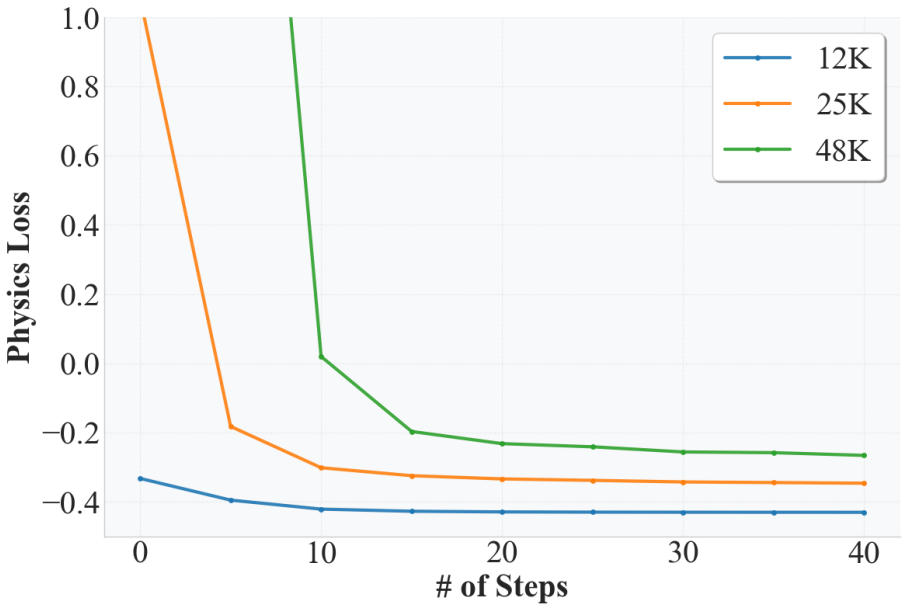

Figure 4.5. Physics loss vs. number of message propagation steps. Garments with different mesh resolutions require varying numbers of propagation steps to achieve stable and accurate simulation results.

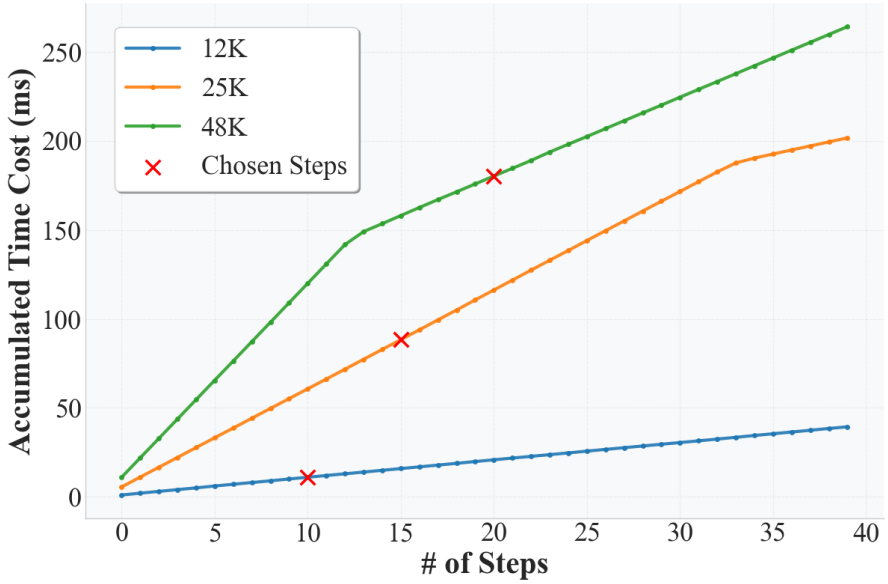

Figure 4.6. Per-step propagation time analysis across mesh resolutions. The cumulative time increases linearly with the number of steps, with steeper slopes for higher resolutions. Red markers denote the propagation steps selected by our adaptive module, demonstrating efficient, resolution-aware computational allocation.

CHAPTER 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, we focus on developing learning-based garment simulation methods as an efficient alternative to traditional physics-based solvers. Specifically, we adopt Graph Neural Networks (GNNs) as our baseline surrogate model and investigate strategies to enhance the practicality of GNN-based garment simulation. Our contributions extend the existing GNN-based simulation framework in terms of both scalability and generalisability.

For scalability, we revise the message-passing mechanism in GNNs to improve propagation efficiency, enabling vertex-wise features to be exchanged more effectively across the garment mesh. To this end, we propose a novel Laplacian-smoothing-based message passing scheme, which significantly broadens the receptive field of each propagation step with minimal computational overhead. Additionally, we introduce a geodesic self-attention module to capture global mesh features and support long-range interactions between distant vertices.

For generalisability, we address the challenge of generalising GNNs to garment meshes with varying resolutions. To achieve this, we propose a propagate-before-update strategy that allows the message-passing depth to adapt dynamically to the complexity of the input mesh. Furthermore, we introduce a resolution-aware update scaling scheme to ensure consistent predictions across meshes of different resolutions. Together, these designs enable the model to generalise effectively without requiring re-training for each new resolution.

With these improvements, our enhanced GNN framework can robustly simulate garments of diverse shapes, scales, and resolutions. This makes it far more practical for real-world applications, where garment mesh topologies can vary significantly.

5.2 Future Work

While our methods significantly improve the practicality of GNN-based garment simulation, several limitations remain.

First, in terms of scalability, our current approach still struggles with ultra-high-resolution garment meshes. Although the revised message-passing strategy improves propagation efficiency, its inherently iterative nature remains a bottleneck, making it computationally expensive for extremely fine meshes. Second, regarding generalisability, our model may still fail in challenging scenarios such as fast or complex body motions (e.g., running) or severe garment-body collisions. These situations can lead to inaccurate predictions or interpenetration due to insufficient modelling of collision dynamics. Moreover, the current framework has been primarily evaluated on regular or near-regular triangulations and does not yet generalise well to highly irregularly re-meshed garments, where large variations in local edge lengths and connectivity may compromise message propagation and scaling. Lastly, controllability remains a major limitation. Unlike traditional physics-based solvers, which allow fine-grained user control through physical parameters, learning-based methods often lack intuitive and robust external control mechanisms.

To address these limitations and further improve robustness, we identify several promising future directions:

- **Integration of Pb4U-GNet and ESLR modules**. While Pb4U-GNet (Chapter 4) focuses on resolution-adaptive receptive field control and ESLR (Chapter 3) introduces efficient range expansion via LSDMP and GSA, a direct substitution of Pb4U's propagation block with LSDMP led to degraded performance. We attribute this to conflicts between Pb4U's decoupled propagation–update design and LSDMP's

per-layer updates, which caused disrupted resolution-aware control. A promising future direction is to explore principled integration, such as adapting LSDMP as a propagation-only operator within Pb4U's framework, recalibrating resolution-aware depth control, and optionally incorporating GSA post-update. Such an approach could combine the scalability and robustness of Pb4U with the efficiency of ESLR, yielding a unified model for garment simulation.

- **Super-resolution for garment simulation**. Super-resolution techniques, widely studied in machine learning, upscale low-resolution inputs into high-resolution outputs using neural networks. A similar strategy can be applied to garments: simulate the garment at a coarse resolution and then refine it to a higher-resolution mesh. This avoids the computational burden of simulating high-resolution meshes directly while preserving fine details.
- **Irregular meshes**. Our current framework has been primarily evaluated on regular or near-regular triangulations, where edge lengths and connectivity remain relatively uniform. In practice, however, garment meshes often exhibit irregular structure, resulting in highly variable local edge distributions. Such irregular discretisations may degrade the performance of our message propagation and resolution-aware scaling modules, as these components implicitly assume approximate regularity in local neighbourhoods. Recent work, such as DiffusionNet (Sharp et al. 2022), has demonstrated that diffusion-based operators can achieve strong robustness to resolution, remeshing, and representation changes. Integrating similar discretisation-agnostic mechanisms into our framework represents a promising direction for future work, potentially enabling more reliable performance across diverse mesh structures encountered in real-world garment modelling.
- **Collision-aware message propagation**. Collision handling—both cloth-body and cloth-cloth (self-collision)—remains one of the most difficult challenges in garment simulation. Fast or complex body motions often lead to garment-body interpenetration, while self-collisions are expensive to detect and resolve. Although recent learning-based methods have begun to address self-collisions, they often incur high

computational costs. Developing a lightweight, high-speed collision resolver integrated into the message-passing process would be a critical step toward real-time simulation.

- **Improved external controllability**. Compared to traditional simulators, learning-based models currently offer limited control. While some methods support manipulation through material parameters, overall controllability remains coarse and unintuitive. Enabling fine-grained, user-defined control (e.g., through external fields, high-level garment behaviour directives, or intuitive latent parameters) could make ML-based simulators more practical for real-world design applications.
- **Hybrid simulation using GNNs and pose conditioning**. One key advantage of machine learning models is their ability to leverage known correlations between inputs and outputs. A promising future direction is to develop hybrid frameworks that combine pose-conditioned inputs with GNN-based simulations. This would allow the network to generalise better to diverse motions and garments while maintaining high efficiency and scalability.

Bibliography

Abdi, Hervé (2007). 'Metric Multidimensional Scaling (MDS): Analyzing Distance Matrices'. In: *Encyclopedia of Measurement and Statistics*, pp. 1–13.

Baraff, David and Andrew Witkin (1998). 'Large steps in cloth simulation'. In: *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: Association for Computing Machinery, pp. 43–54. ISBN: 0897919998.

Bender, Jan et al. (2014). 'A Survey on Position-Based Simulation Methods in Computer Graphics'. In: *Computer Graphics Forum* 33.6, pp. 228–251.

Bertiche, Hugo, Meysam Madadi and Sergio Escalera (Dec. 2021a). 'PBNS: Physically Based Neural Simulation for Unsupervised Garment Pose Space Deformation'. In: *ACM Transactions on Graphics (TOG)* 40.6.

— (2022). 'Neural cloth simulation'. In: *ACM Transactions on Graphics (TOG)* 41.6, pp. 1–14.

Bertiche, Hugo et al. (2021b). 'DeePSD: Automatic deep skinning and pose space deformation for 3D garment animation'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5471–5480.

Bhat, Kiran S et al. (2003). 'Estimating cloth simulation parameters from video'. In.

Bhatnagar, Bharat Lal et al. (2019). 'Multi-garment net: Learning to dress 3d people from images'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5420–5430.

Bouaziz, Sofien et al. (July 2014). 'Projective dynamics: fusing constraint projections for fast simulation'. In: *ACM Transactions on Graphics (TOG)* 33.4. ISSN: 0730-0301.

Bridson, Robert, Ronald Fedkiw and John Anderson (2002). 'Robust Treatment of Collisions, Contact and Friction for Cloth Animation'. In: *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 594–603.

Chen, Jie, Tengfei Ma and Cao Xiao (2018). 'Fastgcn: fast learning with graph convolutional networks via importance sampling'. In: *arXiv preprint arXiv:1801.10247*.

Chiang, Wei-Lin et al. (2019). 'Cluster-gcn: An efficient algorithm for training deep and large graph convolutional networks'. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, pp. 257–266.

Choi, Kwang-Jin and Hyeong-Seok Ko (2005). 'Stable but responsive cloth'. In: *ACM SIGGRAPH 2005 Courses*, 1–es.

Choromanski, Krzysztof et al. (2020). 'Rethinking attention with performers'. In: *arXiv preprint arXiv:2009.14794*.

Cirio, Gabriel et al. (2014). 'Yarn-level simulation of woven cloth'. In: *ACM Transactions on Graphics (TOG)* 33.6, pp. 1–11.

Diao, Junqi et al. (2023). 'Combating Spurious Correlations in Loose-fitting Garment Animation Through Joint-Specific Feature Learning'. In: *Computer Graphics Forum*. Vol. 42. 7. Wiley Online Library, e14939.

Fortunato, Meire et al. (2022). 'Multiscale meshgraphnets'. In: *arXiv preprint arXiv:2210.00612*.

Goodfellow, Ian et al. (2020). 'Generative adversarial networks'. In: *Communications of the ACM* 63.11, pp. 139–144.

Grigorev, Artur, Michael J. Black and Otmar Hilliges (2023). 'HOOD: Hierarchical Graphs for Generalized Modelling of Clothing Dynamics'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16965–16974.

Grigorev, Artur et al. (2024). 'ContourCraft: Learning to Resolve Intersections in Neural Multi-Garment Simulations'. In: *ACM SIGGRAPH 2024 Conference Papers*, pp. 1–10.

Grinspun, Eitan et al. (2003). 'Discrete Shells'. In: *Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 62–67.

Guan, Peng et al. (July 2012). 'DRAPE: DRessing Any PErson'. In: *ACM Transactions on Graphics (TOG)* 31.4. ISSN: 0730-0301.

Gundogdu, Erhan et al. (2019). 'Garnet: A two-stream network for fast and accurate 3d cloth draping'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 8739–8748.

Hamilton, Will, Zhitao Ying and Jure Leskovec (2017). 'Inductive representation learning on large graphs'. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)* 30.

Harmon, David et al. (2009). 'Asynchronous contact mechanics'. In: *ACM SIGGRAPH 2009 Conference Papers*, pp. 1–12.

Holden, Daniel et al. (2019). 'Subspace neural physics: Fast data-driven interactive simulation'. In: *Proceedings of the 18th annual ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 1–12.

Kaldor, Jonathan M., Doug L. James and Steve Marschner (2008). 'Simulating knitted cloth at the yarn level'. In: *ACM SIGGRAPH 2008 Conference Papers*. Los Angeles, California: Association for Computing Machinery. ISBN: 9781450301121.

— (July 2010). 'Efficient yarn-based cloth with adaptive contact linearization'. In: *ACM Transactions on Graphics (TOG)*.

Lahner, Zorah, Daniel Cremers and Tony Tung (2018). 'Deepwrinkles: Accurate and realistic clothing modeling'. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 667–684.

Li, Minchen, Danny M Kaufman and Chenfanfu Jiang (2020). 'Codimensional incremental potential contact'. In: *arXiv preprint arXiv:2012.04457*.

Li, Peizhuo et al. (2024). 'Neural Garment Dynamics via Manifold-Aware Transformers'. In: *Computer Graphics Forum (Proceedings of Eurographics 2024)* 43.2.

Liu, Aoran et al. (2023). 'Material-Aware Self-Supervised Network for Dynamic 3D Garment Simulation'. In: *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, pp. 630–635.

Liu, Aoran et al. (2025). 'Extended Short- and Long-Range Mesh Learning for Fast and Generalized Garment Simulation'. In: *arXiv preprint arXiv:2504.11763*.

Liu, Tiantian et al. (Nov. 2013). 'Fast simulation of mass-spring systems'. In: *ACM Transactions on Graphics (TOG)* 32.6. ISSN: 0730-0301.

Liu, Wing Kam, Shaofan Li and Harold S. Park (June 2022). 'Eighty Years of the Finite Element Method: Birth, Evolution, and Future'. In: *Archives of Computational Methods in Engineering* 29, pp. 4431–4453.

Loper, Matthew et al. (Oct. 2015). 'SMPL: a skinned multi-person linear model'. In: *ACM Transactions on Graphics (TOG)* 34.6. ISSN: 0730-0301.

Mahmood, Naureen et al. (Oct. 2019). 'AMASS: Archive of Motion Capture as Surface Shapes'. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5442–5451.

Miguel, Eder et al. (2012). 'Data-driven estimation of cloth simulation models'. In: *Computer Graphics Forum*. Vol. 31. 2pt2. Wiley Online Library, pp. 519–528.

Müller, Matthias et al. (Apr. 2007). 'Position based dynamics'. In: *Journal of Visual Communication and Image Representation* 18.2, pp. 109–118. ISSN: 1047-3203.

Narain, Rahul, Armin Samii and James F O'brien (2012). 'Adaptive anisotropic remeshing for cloth simulation'. In: *ACM Transactions on Graphics (TOG)* 31.6, pp. 1–10.

Pan, Xiaoyu et al. (2022). 'Predicting loose-fitting garment deformations using bone-driven motion networks'. In: *ACM SIGGRAPH 2022 Conference Paper*, pp. 1–10.

Patel, Chaitanya, 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 (CVPR)*, pp. 7365–7375.

Pavlakos, Georgios et al. (2019). 'Expressive body capture: 3d hands, face, and body from a single image'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10975–10985.

Pfaff, Tobias et al. (2020). 'Learning mesh-based simulation with graph networks'. In: *Proceedings of the International Conference on Learning Representations (ICLR)*.

Pinheiro Cinelli, Lucas et al. (2021). 'Variational autoencoder'. In: *Variational methods for machine learning with applications to deep networks*. Springer, pp. 111–149.

Pons-Moll, Gerard et al. (2017). 'ClothCap: Seamless 4D Clothing Capture and Retargeting'. In: *ACM Transactions on Graphics (TOG)* 36.4. Two first authors contributed equally.

Provot, Xavier et al. (1995). 'Deformation constraints in a mass-spring model to describe rigid cloth behaviour'. In: *Graphics interface*. Canadian Information Processing Society, pp. 147–147.

Rampášek, Ladislav et al. (2022). 'Recipe for a General, Powerful, Scalable Graph Transformer'. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 35, pp. 14501–14515.

Sanchez-Gonzalez, Alvaro et al. (2020). 'Learning to simulate complex physics with graph networks'. In: *Proceedings of the International Conference on Machine Learning (ICML)*. ICML'20. JMLR.org.

Santesteban, Igor, Miguel A Otaduy and Dan Casas (2019). 'Learning-based animation of clothing for virtual try-on'. In: *Computer Graphics Forum*. Vol. 38. 2. Wiley Online Library, pp. 355–366.

— (2022a). 'Snug: Self-supervised neural dynamic garments'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8140–8150.

Santesteban, Igor et al. (2021). 'Self-supervised collision handling via generative 3D garment models for virtual try-on'. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11763–11773.

Santesteban, Igor et al. (2022b). 'ULNeF: Untangled layered neural fields for mix-and-match virtual try-on'. In: *Advances in Neural Information Processing Systems* 35, pp. 12110–12125.

Scarselli, Franco et al. (2008). 'The graph neural network model'. In: *IEEE Transactions on Neural Networks* 20.1, pp. 61–80.

Sharp, Nicholas et al. (2022). 'DiffusionNet: Discretization Agnostic Learning on Surfaces'. In: *ACM Transactions on Graphics (TOG)* 41.3.

Tang, Min et al. (2018). 'I-Cloth: Incremental collision handling for GPU-based interactive cloth simulation'. In: *ACM Transactions on Graphics (TOG)* 37.6, pp. 1–10.

Terzopoulos, Demetri et al. (Aug. 1987). 'Elastically deformable models'. In: *ACM SIGGRAPH Computer Graphics* 21.4, pp. 205–214. ISSN: 0097-8930.

Tiwari, Lokender and Brojeshwar Bhowmick (2023). 'Garsim: Particle based neural garment simulator'. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 4472–4481.

Vidaurre, Raquel et al. (2020). 'Fully Convolutional Graph Neural Networks for Parametric Virtual Try-On'. In: *Computer Graphics Forum* 39.8, pp. 145–156.

Volino, Pascal, Nadia Magnenat-Thalmann and Francois Faure (2009). 'A simple approach to nonlinear tensile stiffness for accurate cloth simulation'. In: *ACM Transactions on Graphics (TOG)* 28.4, Article–No.

Wang, Huamin, James F O'Brien and Ravi Ramamoorthi (2011). 'Data-driven elastic models for cloth: modeling and measurement'. In: *ACM Transactions on Graphics (TOG)* 30.4, pp. 1–12.

Wang, Huamin et al. (2010). 'Example-based wrinkle synthesis for clothing animation'. In: *ACM SIGGRAPH 2010 Conference Papers*. Los Angeles, California: Association for Computing Machinery.

Wang, Tuanfeng Y. et al. (Dec. 2018). 'Learning a shared shape space for multimodal garment design'. In: *ACM Transactions on Graphics (TOG)* 37.6. ISSN: 0730-0301.

Wang, Tuanfeng Y. et al. (Nov. 2019). 'Learning an Intrinsic Garment Space for Interactive Authoring of Garment Animation'. In: *ACM Transactions on Graphics (TOG)* 38.6. ISSN: 0730-0301.

Wu, Felix et al. (2019). 'Simplifying graph convolutional networks'. In: *Proceedings of the 36th International Conference on Machine Learning (ICML)*. PMLR, pp. 6861–6871.

Wu, Haixu et al. (2024). 'Transolver: A fast transformer solver for pdes on general geometries'. In: *arXiv preprint arXiv:2402.02366*.

Xu, Weiwei et al. (July 2014). 'Sensitivity-Optimized Rigging for Example-Based Real-Time Clothing Synthesis'. In: *ACM Transactions on Graphics (TOG)* 33.4. ISSN: 0730-0301.

Zhang, Meng, Duygu Ceylan and Niloy J Mitra (2022a). 'Motion guided deep dynamic 3d garments'. In: *ACM Transactions on Graphics (TOG)* 41.6, pp. 1–12.

Zhang, Si et al. (2019). 'Graph convolutional networks: a comprehensive review'. In: *Computational Social Networks* 6.1, pp. 1–23.

Zhang, Wentao et al. (2022b). 'Graph attention multi-layer perceptron'. In: *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 4560–4570.