Cloth-Splatting: 3D Cloth State Estimation from RGB Supervision

Alberta Longhini*[1], Marcel Büsching*[1], Bardienus P. Duisterhof[2], Jens Lundell[1], Jeffrey Ichnowski[2], Mårten Björkman[1], Danica Kragic[1]

[1]KTH Royal Institute of Technology [2]Carnegie Mellon University

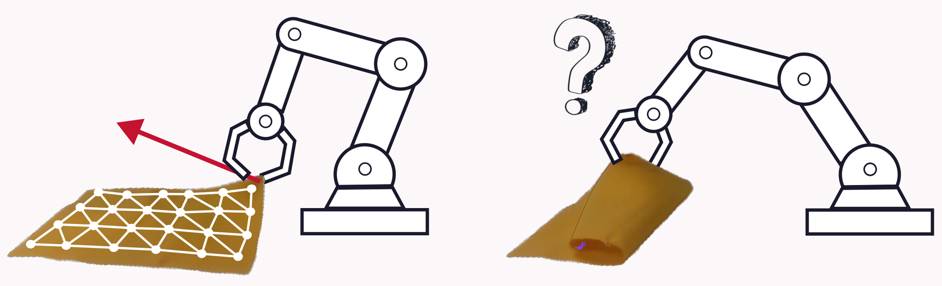
Introduction

Estimating the 3D state of deformable objects is a long-standing challenge in robotic manipulation.

Current solutions either rely exclusively on 2D images, overlooking the cloth's 3D structure, or use 3D representations that neglect valuable information in RGB observations.

To overcome these limitations we propose **combining 3DGS with an action-conditioned dynamics model for state estimation.**

Aim

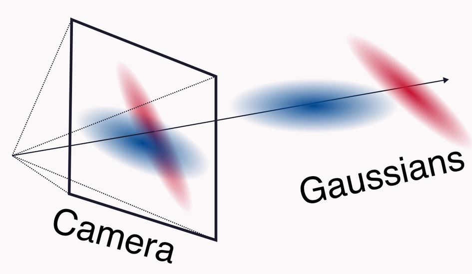


Estimation of 3D cloth states from RGB observation for robotic manipulation.

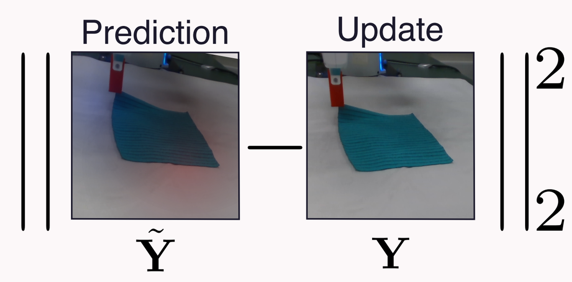
Background

3D Gaussian Splatting (3DGS) is a novel method for fast 3D reconstruction and image synthesis. The scene is represented by 3D Gaussians defined by position, size, color, and opacity.

Rendering
Projecting Gaussians to the image plane

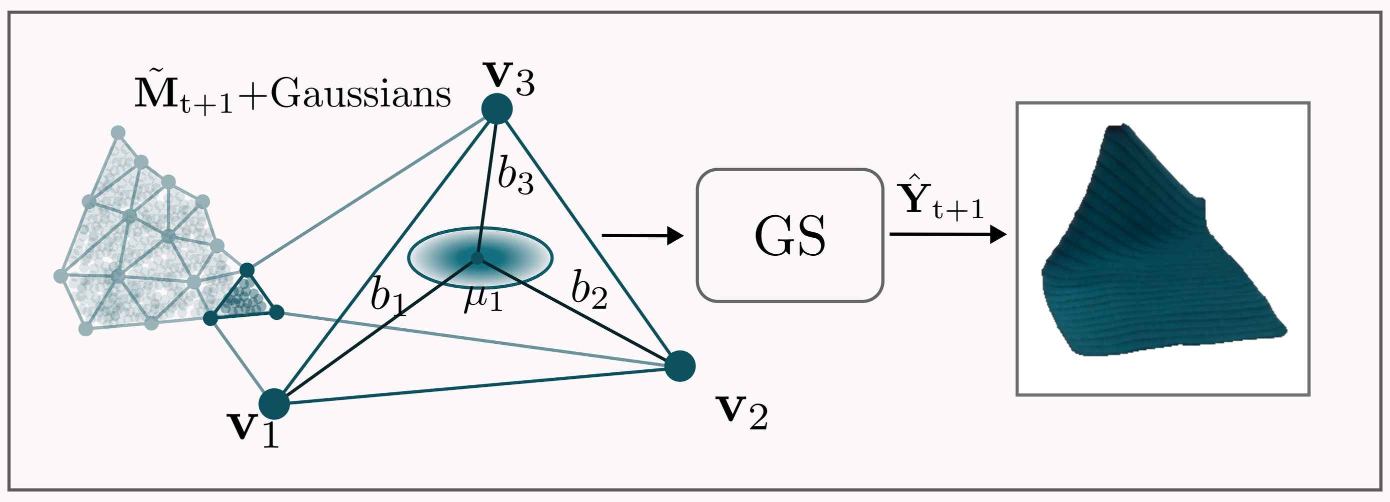


Optimization
Minimizing RGB loss



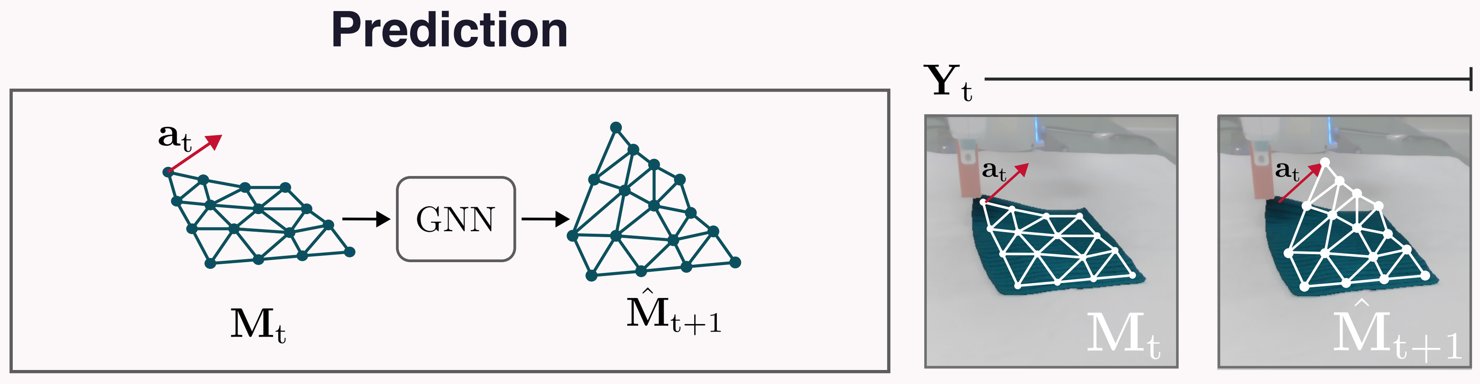
Method

Mesh-constrained 3DGS (h_{GS})

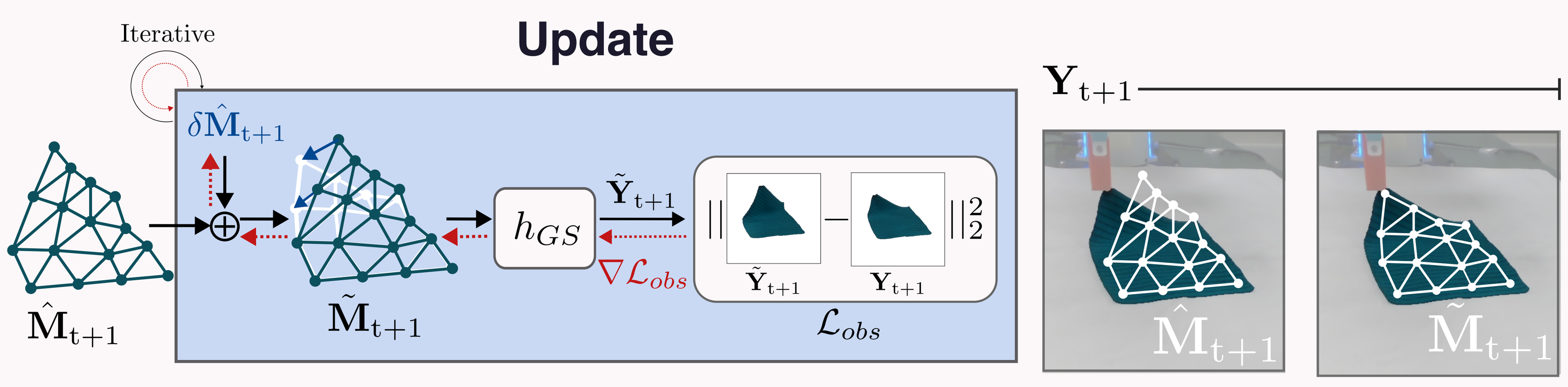


The **cloth geometry is represented as mesh $\mathrm{M_t}$** while the **cloth appearance is modelled via 3D Gaussians** placed on the faces of the Mesh. This creates a **differentiable mapping between cloth state space and rhe space of the observations $\mathrm{Y_t}$.**

Prediction



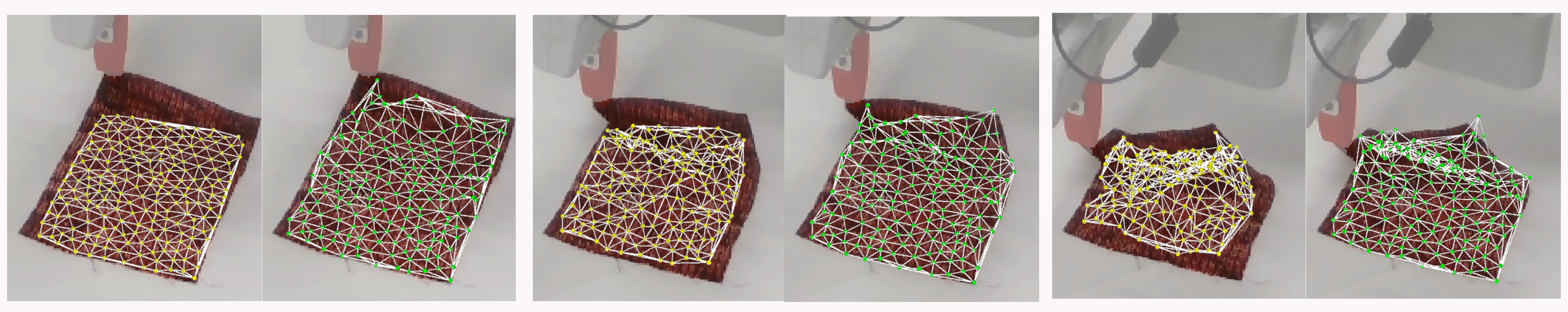
Update



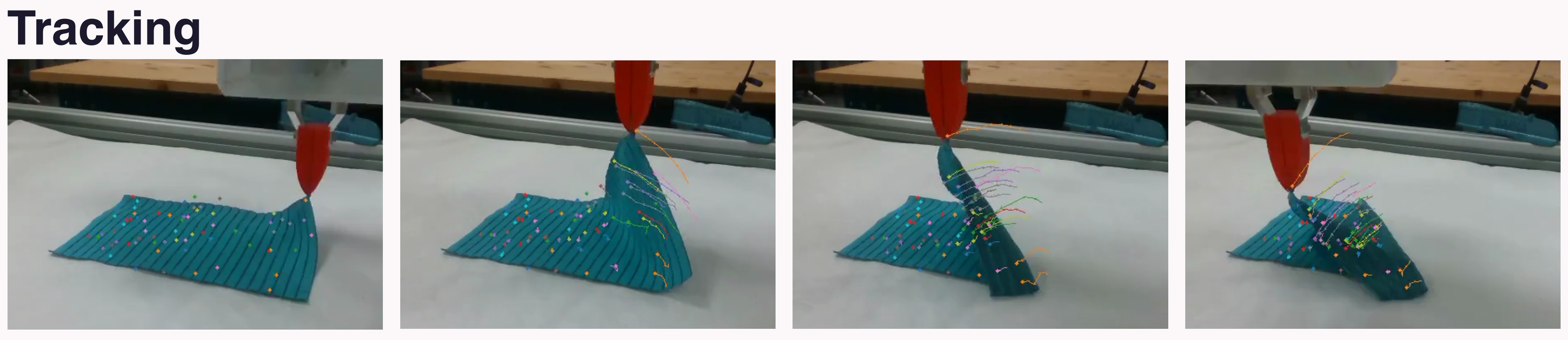
The **prediction-update framework**, integrates a pre-trained action-conditioned dynamics model GNN with the differentiable map h_{GS}. This allows for updating the mesh state via gradient based optimization of a rendering loss $\mathcal{L}_{obs}$.

Results

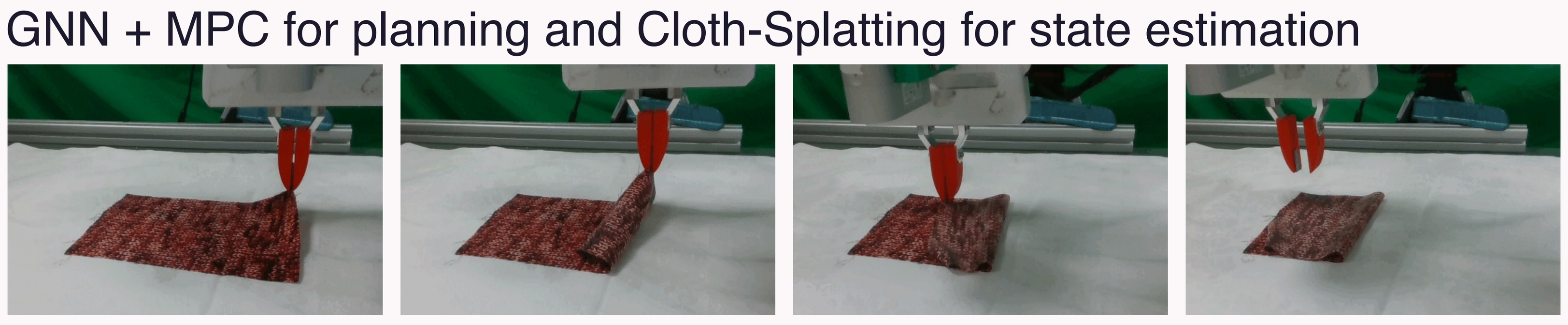
Mesh representation
Prediction (left, yellow) and update (right, green)



Tracking



Manipulation
GNN + MPC for planning and Cloth-Splatting for state estimation



Model Ablation
Evaluated on simulated data

Ablation		3D MTE [mm]
(A1)	Only GNN	21.552
(A2)	No GNN	16.135
(A3)	No $\mathcal{L}_{reg}$	15.772
(A4)	1 view	16.525
(A5)	2 views	9.535
(A6)	3 views	9.004
Full (4 views)		**8.923**

Tracking (accuracy)
Evaluated on simulated data

Metric	Method	Mean
3D MTE ↓ [mm]	DynaGS	10.924 ± 11.246
	MD-Splatting	3.635 ± 6.235
	GNN	13.853 ± 12.674
	Cloth-Splating	**3.284** ± 3.722
3D δ_{avg} ↑	DynaGS	0.804 ± 0.105
	MD-Splatting	0.847 ± 0.083
	GNN	0.747 ± 0.121
	Cloth-Splatting	**0.862** ± 0.072
Survial rate ↑	DynaGS	0.835 ± 0.113
	MD-Splatting	0.887 ± 0.077
	GNN	0.800 ± 0.128
	Cloth-Splatting	**0.910** ± 0.072

Conclusion

Cloth-Splatting improves tracking accuracy and reduces convergence time by ~85 % over current baselines.

Future work will focus on broader categories of defomable objects, real-time perfomance and robust intialization.

References

3D Gaussian Splatting for Real-Time Radiance Field Rendering, Kerbl et al. SIGGRAPH 2023

MD-Splatting: Learning Metric Deformation from 4D Gaussians in Highly Deformable Scenes, Duisterhof et al. arXiv 2023

Dynamic 3D Gaussians: Tracking by persistent dynamic view synthesis, Luiten et al. 3DV 2024