

Yang You¹, Bokui Shen¹, Congyue Deng, Haoran Geng², He Wang², Leonidas Guibas¹

Check out our Experimental Results Problems with Deformable Manipulation project website Simulation Results How to generate good paths, with infinite degrees of freedom In general, long-horizon task Stage 1 planning is non-trivial **Rolling Pin** LLM generated Execution Method Overview subgoal Stage 2 LLM-guided multi-tool planning Gripper Stage 2: Create a hole Stage 1: Flatten dough **Real-robot execution** using Rolling pin "Make a **donut**' GPT-4 Tool name: [Rolling pin] Tool name: [Pole] LLM generated Shape Python Code: Shape Python Code: User input prompt Execution subgoal Real-robot Results **Tool Description:** Initial - Knife: it can cut the dough in Tool state half into two pieces, the cut is descriptions done on x axis... - Gripper: ... Stage 1 - Rolling pin: ... Cutter Diff. physics & dense corr. Note: Guiding Single-tool EMD - Use numpy to generate the rules point clouds for each subgoal. space planning Flatten the dough **Context prompt** Stage 2 **Rolling** Pin Top level: Multi-Tool Selection Bottom level: Single-Tool Planning Stage 3 with LLM Guidance in the EMD Space **Rolling Pin DiffPhysics-P2P:** Decide the next subgoal with We ask the LLM to output the following items: Execution $\mathbf{p}'_i = \mathbf{p}_i - \alpha \cdot \frac{\partial \operatorname{emd}(\{\mathbf{p}_i\}, \{\bar{\mathbf{p}}_i\})}{\alpha}$ • A one-line explanation of what this step is doing Ablation Studies **Initial position selection** $\mathbf{x}^* = \arg \max_{\mathbf{x}} \sum_{i} \frac{\|\mathbf{p}'_i - \mathbf{p}_i\|_1}{\mathrm{sdf}_{\mathbf{x}}(\mathbf{p}_i) + \delta}$ • Name of the tool to be used. • Python code to generate target point clouds Next reachable Target Current w/o LLM High-Level Planning point cloud point cloud • Variable names for the input and output. point cloud w/o Volume Preserving and Chain Reasoning $\{p_i'\} \quad p_i' = p_i - lpha rac{\partial \mathrm{emd}(\{p_i\}, \{ar{p_i}\})}{\partial p_i}$ - 8.0 G D 0.6 - $\{p_i\}$ $\{ar{p_i}\}$ Ours • Location of each piece in a dictionary format 0.501 σ 0.4 0.2 0.289 with a variable name as the key 0.001-0.001 0.002 0.001 • Volume of each piece is also in a dictionary -0.165 Action sequence -0.2 format with a variable name as the key Diff. physics from dense corr. w/o LLM High-Level Planning $L = \sum_i \|\mathbf{p}_i' - \mathbf{p}_i\|_1$ (%) **Additional guidelines in the prompt:** w/o Volume Preserving and Chain Reasoning 001 Rate Ours **Tool reset upon failures** Volume preserving 80 65 SS Get stuck! 🗙 Tool reset Lump removed ✔ 60 • Chain of reasoning nitial plan volume changes 8 40 25 Su 20 0 0 0 TwoPancakes Baguette Donut (a) Ablation results for multiple-tool experiments.







	Donut ↓	Baguette \downarrow	TwoPancakes↓
w/o Volume Preserving and Chain Reasoning	73.9%	42.5%	65.0%
Ours	9.8%	38.9%	0.0%

Make a Dount 😟 : Language-Guided Hierarchical EMD-Space Planning for

¹Stanford University ²Peking University













(b) Ablation results for single-tool experiments.





