

FoldBack: Self-Correcting Masked Generative Policy for Long-Horizon Garment Folding

Lipeng Zhuang, Shiyu Fan, Yingdong Ru, Zhuo He,
Florent P. Audonnet, Paul Henderson, Gerardo Aragon Camarasa

Abstract—We present FoldBack, a self-correcting masked generative policy for long-horizon garment folding. Existing long-trajectory policies may continue after a missed or slipped grasp even when the garment has not reached the intended configuration. We structure FoldBack’s recovery mechanisms around three inference-time decisions: *when to refine and verify, how to roll back, and where and how to retry*. FoldBack aligns refinement and grasp verification with pick-and-place events, returns the robot to a retryable pre-grasp configuration while preserving successful grasps, and selectively regenerates the failed segment and selected future actions while avoiding previous failed grasp locations. To our knowledge, FoldBack is the first editable full-trajectory policy to unify these decisions, enabling failed interactions to be detected, undone, and repaired before execution continues, without recovery demonstrations or base-policy retraining. Across 33 real garments from six categories, FoldBack achieves 75.2% final folding success and 0.837 final-mask IoU, versus 45.7% and 0.689 for the strongest prior baseline.

I. INTRODUCTION

Garment folding is a long-horizon manipulation task that requires executing a sequence of precise single-arm or bimanual pick-and-place actions. Each action changes the garment configuration, which in turn determines the conditions for subsequent folding actions. A missed or slipped grasp therefore invalidates subsequent folding actions [1]. For example, a robot may close its gripper and lift without successfully moving the garment, leaving it in a configuration incompatible with the next folding action. Reliable folding not only requires long-horizon action generation but also a mechanism for detecting grasp failures and recovering from them before execution advances.

Visuomotor imitation learning enables robots to learn complex behaviors from demonstrations [1]–[3]. Short-horizon policies generate action chunks myopically given recent observations, enabling reactive execution [4]–[7]. However, grasp failures can produce out-of-distribution observations that cause subsequent predictions to stall, drift, or lose coherence with the overall folding strategy. Long-horizon generative policies generate complete, coherent action sequences and can refine future action tokens as observations arrive [8], [9]; yet executed action tokens remain fixed in the trajectory history without verifying the intended manipulation outcome. Thus, execution may advance to the next folding stage without revisiting the failed interaction. Runtime monitors can detect such failures but do not themselves revise the policy trajectory [10]–[12], while recovery-aware methods

often rely on additional recovery data, policy adaptation, or large-model reasoning [1], [13]–[16].

Motivated by these limitations, we introduce FoldBack, an outcome-gated masked generative policy for long-horizon garment folding that supports recovery from grasp failures. FoldBack represents the full action trajectory as discrete tokens and maintains a grasp-verified history, a tentative grasp-critical window, and an editable future. Executing a grasp-critical window does not immediately add it to history. The window is added only after all required grasps are verified; otherwise, the history remains unchanged and the failed window stays editable (Fig. 1).

FoldBack is then organized around three inference-time decisions which constitute the contributions of this paper:

- 1) *when* to refine and verify (Sec. III-A),
- 2) *how* to roll back after a failure (Sec. III-B), and
- 3) *where and how* to retry (Sec. III-C).

This enables recovery without recovery demonstrations or base-policy retraining. Indeed, across 33 garments from 6 categories, FoldBack achieves 75.2% final folding success and a mean final-mask IoU of 0.837 across single-arm and bimanual folding tasks, outperforming the strongest prior baseline by 29.5 percentage points in final folding success and 0.148 in mean final-mask IoU. FoldBack also generalizes to unseen garment instances, initial orientations, and non-canonical configurations as shown in Sec. IV.

II. RELATED WORK

Learning-based garment manipulation. Garment-folding policies typically compose prescribed primitives or generate local action segments [17]–[19]. SpeedFolding [20] and UniFolding [2] execute parameterized primitives, while BiFold [21] predicts a language-conditioned pick-and-place action at each step. Model-based methods learn cloth dynamics for planning [22], [23], while UniClothDiff [24] uses diffusion models for generative state estimation and dynamics prediction within model-predictive control. Other generative policies model distributions over action primitives (DeformPAM [25]), closed-loop or local manipulation trajectories (FoldNet [1] and HALO [26]), or stage-wise point-cloud trajectories (MetaFold [27]). These methods are effective for local or stage-wise manipulation, but long-horizon folding requires consistent execution across multiple strategies with both single- and bimanual actions. FoldBack addresses this with full-trajectory masked generation for an editable plan (Sec. III), Key-Action-Aligned Inference for event-aligned refinement and grasp verification (Sec. III-A),

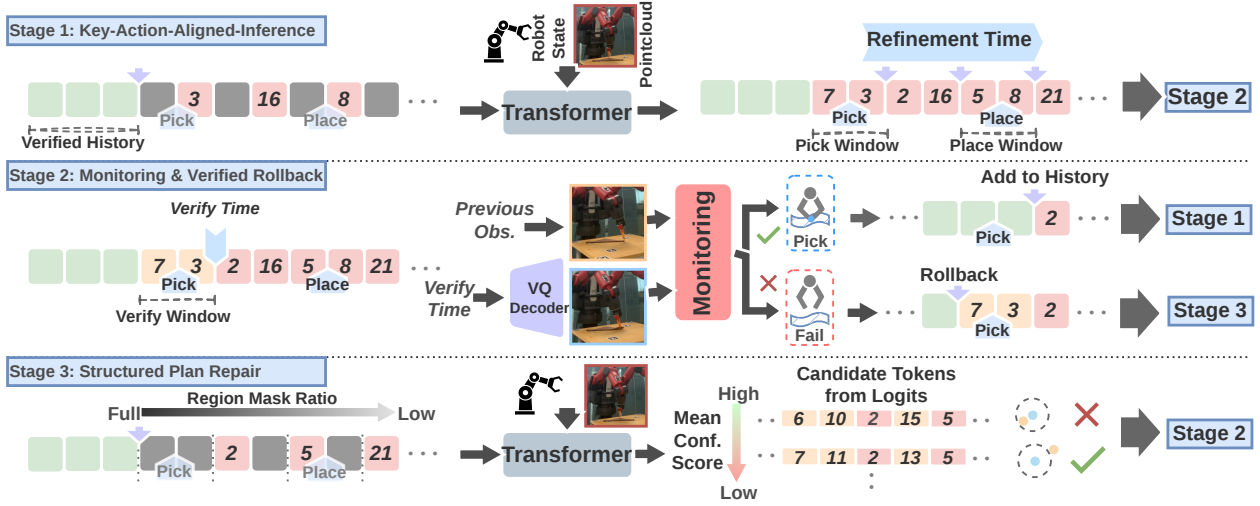


Fig. 2. Overview of FoldBack. Key-Action-Aligned Inference aligns trajectory refinement with pick-and-place events; grasp monitoring and verified rollback determine whether a grasp-critical window is added to the verified trajectory history or kept editable for recovery; structured plan repair then remarks and resamples the editable trajectory after failure.

A. Key-Action-Aligned Inference (KAAI)

KAAI ensures trajectory refinement and grasp verification happen after pick-and-place transitions of garment folding rather than at uniform execution intervals. Since pick and place events mark the main physical transitions in folding, after which the garment configuration can change substantially, this effectively enables the robot to perform visual servoing at key moments in the trajectory. Following initial trajectory generation, FoldBack extracts these key actions from the decoded gripper-state sequences $(g_1^r, \dots, g_T^r)$. An open-to-closed transition of the gripper defines a PICK, whereas a closed-to-open transition defines a PLACE. This yields J key actions at times $t_1 < \dots < t_J$.

Each key action k defines an action window $\mathcal{W}_k = \{t_k - 4, \dots, t_k + 3\}$. Let $\mathcal{I}_k \subseteq \{1, \dots, N\}$ denote the smallest set of token indices whose decoded actions cover $\mathcal{W}_k$. In bimanual manipulation, the two arm-specific PICK transitions are treated as a single key action, whose action window spans from the first gripper-closing event to the end of the corresponding joint lift. A window whose key action is a PICK is *grasp-critical* because it establishes a grasp that the rest of the folding stage depends on. Before and after executing each key-action window $\mathcal{W}_k$, FoldBack refines the remaining unexecuted tokens using the latest observation. After each refinement, FoldBack re-decodes the gripper-state sequence and recomputes the transition time t_j , action window $\mathcal{W}_j$, and token set $\mathcal{I}_j$ for each unexecuted key action j . The PICK/PLACE order and arm assignment remain fixed, while their timing may shift as the trajectory is refined. We use two MGP mask-refinement iterations at each update.

B. Grasp-outcome monitoring and Verified Rollback

FoldBack verifies each required grasp after lifting, and then adds the corresponding grasp-critical window to the trajectory history only when all grasps succeed. In sequential pick-and-place manipulation, each grasp is verified after its own lift. In bimanual manipulation, the first arm maintains its grasp while the second performs its pick; both arms then

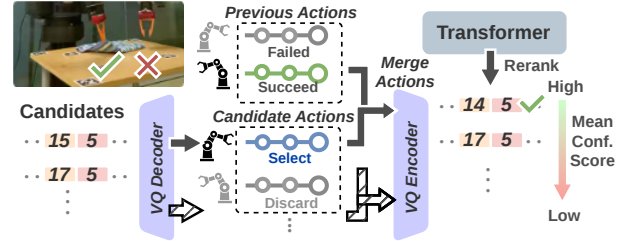


Fig. 3. Repair under an asymmetric bimanual grasp failure.

lift, and both grasps are verified. If any grasp fails, FoldBack initiates the verified rollback procedure described below.

a) *Local Grasp-Outcome Monitor*: At verification point k , FoldBack evaluates each arm r whose grasp must be verified. Using the gripper poses, we extract pre-pick and post-lift local point clouds, $\mathbf{P}_{k,r}^{\text{pre}}$ and $\mathbf{P}_{k,r}^{\text{lift}}$, from a fixed region below the fingertips before grasp closure and after lifting, respectively. The two point clouds are expressed in their respective gripper frames. This representation focuses the monitor on changes in local garment geometry, largely excludes the gripper, and requires no garment segmentation. Thus, a PointNet-based [38] binary classifier f_ψ predicts

$$p_{k,r} = f_\psi(\mathbf{P}_{k,r}^{\text{pre}}, \mathbf{P}_{k,r}^{\text{lift}}) \in [0, 1], \quad (1)$$

where ψ denotes the monitor parameters and $p_{k,r}$ is the predicted grasp-success probability. The grasp is verified as successful when $p_{k,r} \geq \tau_{\text{mon}}$, with $\tau_{\text{mon}} = 0.50$ selected on the validation set and fixed for all evaluations.

We train f_ψ using binary cross-entropy on successful Flat'n'Fold demonstrations. A positive pair contains $\mathbf{P}_{k,r}^{\text{pre}}$ and its corresponding $\mathbf{P}_{k,r}^{\text{lift}}$. A negative pair replaces $\mathbf{P}_{k,r}^{\text{lift}}$ with a local point cloud from a nearby pre-pick frame at the location where the post-lift point cloud would be taken. Because the garment remains on the table in that frame, the resulting pair approximates a missed grasp without requiring physical failure trajectories. Training pairs are constructed independently for each arm, and one classifier is shared across single-arm and bimanual stages. The trained policy

and monitor remain fixed during deployment.

b) Grasp-Verified Trajectory History: FoldBack does not immediately add the executed tokens of a grasp-critical window to the trajectory history. Instead, it keeps the token subsequence $\mathbf{y}_{\mathcal{I}_k}$ tentative until post-lift verification is complete, while the unexecuted future trajectory remains editable. Let $v_k \in \{0, 1\}$ denote the grasp-outcome monitor decision for $\mathcal{W}_k$, where $v_k = 1$ indicates that all associated grasps are verified as successful and $v_k = 0$ otherwise. If $v_k = 1$, FoldBack appends $\mathbf{y}_{\mathcal{I}_k}$ to the grasp-verified trajectory history $\mathcal{H}^{\text{ver}}$; otherwise, $\mathcal{H}^{\text{ver}}$ remains unchanged and the failed window remains editable for plan repair; neither the failed attempt nor the rollback motion is added to the history. Tokens executed outside the current grasp-critical window remain fixed as trajectory history. For a bimanual manipulation window, the joint token subsequence $\mathbf{y}_{\mathcal{I}_k}$ is saved to the trajectory history only if both arms' grasps are verified as successful.

c) Verified rollback: FoldBack records the executed end-effector motion from the pre-grasp to the post-lift configuration. When $v_k = 0$, each failed gripper opens and the recorded motion is executed in reverse, restoring the pre-grasp configuration. In a bimanual manipulation, both arms return in coordination to their pre-grasp configurations; if only one grasp fails, the successful gripper remains closed to preserve its grasp. At the restored configuration, FoldBack captures an updated point cloud and robot state. The procedure is policy-agnostic and requires no base-policy re-training. After rollback, action-chunking policies replan from the updated observation, whereas masked long-trajectory policies can additionally retain grasp-verified history and regenerate affected editable tokens.

C. Structured Trajectory Plan Repair

Verified rollback restores a retryable configuration but does not revise the trajectory. This is because the unexecuted actions were generated under the assumption that the failed grasp would succeed. When a grasp-critical window $\mathcal{W}_k$ fails verification, FoldBack repairs the trajectory in two steps.

a) Distance-guided remasking: The unexecuted token positions after $\mathcal{I}_k$ are grouped into segments $\mathcal{S}_1, \dots, \mathcal{S}_{J-k}$, one for each subsequent key-action window. $\mathcal{S}_1$ contains the actions immediately after the failed window up to the next key-action window; each subsequent segment spans between consecutive key-action windows. The final segment additionally includes any remaining actions after the last key-action window. FoldBack fully remasks $\mathcal{I}_k$ and assigns each segment $\mathcal{S}_\ell$ the masking ratio, defined as the fraction of action tokens masked and regenerated:

$$\rho_\ell = \rho_{\text{base}} + (\rho_{\text{near}} - \rho_{\text{base}})\alpha^{\ell-1}, \quad (2)$$

where ρ_{base} is the masking ratio used by MGP-Long during trajectory refinement, $\rho_{\text{near}} > \rho_{\text{base}}$ is the ratio applied to $\mathcal{S}_1$, and $\alpha \in (0, 1)$ controls the rate of decay. The masking ratio therefore decreases toward ρ_{base} as the segment distance from the failure increases.

Within each segment $\mathcal{S}_\ell$, FoldBack remasks the ρ_ℓ -fraction of tokens with the lowest stored confidence γ_i^{last} , defined as the probability assigned to the current token y_i when it was most recently generated. We denote the resulting set of remasked positions, including $\mathcal{I}_k$, by $\mathcal{M}$. FoldBack regenerates the tokens at these positions using the observation captured after the robot returns to the pre-grasp configuration, while the grasp-verified prefix and unselected future tokens remain fixed as anchors. During subsequent mask-refinement iterations, the stored confidence of each regenerated token is updated with its newly predicted probability. We use $\rho_{\text{near}} = 0.80$, $\rho_{\text{base}} = 0.30$, and $\alpha = 0.50$ in all experiments.

b) Failure-aware candidate resampling: A regenerated trajectory plan may repeat the grasp that just failed. FoldBack therefore draws eight repair candidates for $\mathcal{M}$ from the post-rollback observation. Each arm maintains an episode-level memory $\mathcal{F}_r$ of its failed table-plane pick locations, to which the current failure is added before sampling. A candidate is rejected if its proposed pick for an arm that just failed lies within an exclusion radius $\epsilon = 0.04$ m of any location in $\mathcal{F}_r$. Under an asymmetric bimanual failure, only candidates for the failed arm are rejected based on proximity to previous failed pick locations, while the successful arm retains its grasp and is not resampled. FoldBack samples up to three candidate batches to obtain a valid retry trajectory. If no valid candidate is found, the recovery episode is marked as unsuccessful and the manipulation terminates at the restored pre-grasp configuration.

Let $\gamma_i^{(q)}$ denote the probability assigned to candidate ($q \in \{1, \dots, 8\}$) at regenerated position $i \in \mathcal{M}$. Each candidate is scored by its mean confidence over the regenerated positions:

$$C^{(q)} = \frac{1}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} \gamma_i^{(q)}. \quad (3)$$

For single-arm failures or bimanual failures of both grasps, FoldBack executes the valid candidate with the highest $C^{(q)}$.

For an asymmetric bimanual failure (Fig. 3), FoldBack constructs a time-aligned window for each valid resampled motion. Each candidate combines the stored motion of the successful arm, whose gripper remains closed, with the regenerated motion of the failed arm. Candidates that fail inverse kinematics or violate joint limits are discarded. Each feasible candidate is re-encoded by the frozen tokenizer and evaluated in one batched MGT pass, with $\mathcal{I}_k$ masked and all other candidate tokens fixed.

Let $\bar{\gamma}_i^{(q)}$ denote the probability assigned to the re-encoded bimanual token at $i \in \mathcal{I}_k$. The candidate score is

$$\bar{C}^{(q)} = \frac{1}{|\mathcal{M}|} \left(\sum_{i \in \mathcal{M} \setminus \mathcal{I}_k} \gamma_i^{(q)} + \sum_{i \in \mathcal{I}_k} \bar{\gamma}_i^{(q)} \right), \quad (4)$$

which replaces the original confidence terms for the failed window with those of the executable bimanual candidate. FoldBack executes the feasible candidate with the highest $\bar{C}^{(q)}$. If no feasible candidate remains, it samples another candidate batch, subject to the same three-batch limit. Once all required grasps are verified after the retry, the re-encoded

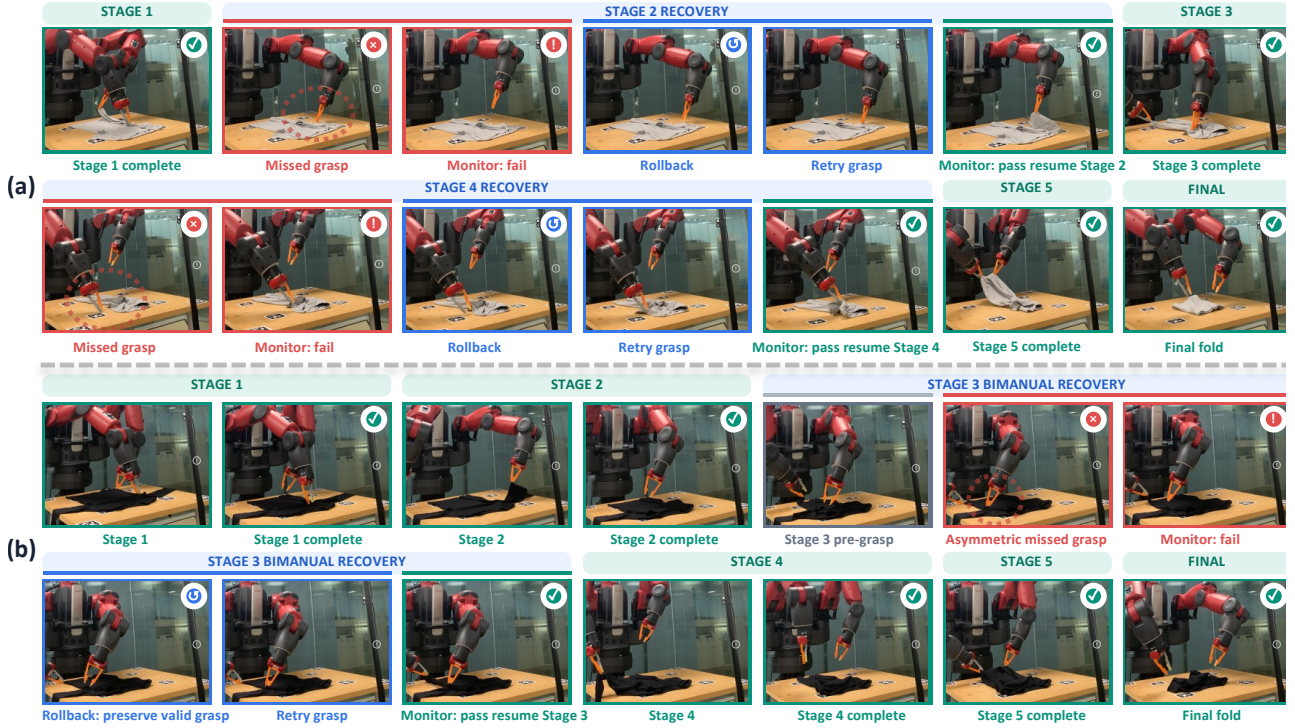


Fig. 4. Real-world FoldBack recovery. (a) Short-sleeved T-shirt: recovery from repeated grasp failures during multi-stage folding. (b) Long-sleeved T-shirt: recovery from an asymmetric bimanual failure while preserving the successful grasp.

bimanual subsequence at $\mathcal{I}_k$ is appended to $\mathcal{H}^{\text{ver}}$, the corresponding stored confidences are updated, and the remaining trajectory is refined using the new post-lift observation.

IV. EXPERIMENTS

We evaluate FoldBack in the real world, aiming to answer four questions: (1) *Can it improve the execution of long-horizon garment-folding strategies?* (Sec. IV-A), (2) *Can verified rollback and trajectory plan repair recover from controlled one- and two-corner grasp slips?* (Sec. IV-B), (3) *How well does the policy generalize to unseen garments and initial configurations?* (Sec. IV-C), (4) *How do FoldBack’s inference-time design choices affect execution and recovery?* (Sec. IV-D). All policies are trained on Flat’n’Fold robot demonstrations.

a) Real-world setup: We follow the Flat’n’Fold setup which includes a Baxter dual-arm robot and a single ZED2 RGB-D camera. Each point cloud is cropped to the workspace and downsampled to 4,096 points. From the point cloud and robot proprioception, the policy predicts a 16-D action vector as defined in Sec. III. Unless stated otherwise, each trial starts with a flattened garment in the workspace. We evaluate on 33 garments from six categories, including napkins of two sizes, three towel instances, pants, short-sleeved T-shirts, long-sleeved T-shirts, and short-sleeved shirts. The garment instances vary in size. The evaluated multi-stage strategies may alternate between single-arm and bimanual manipulation. Some categories use multiple strategies; for towels, the initial orientation determines whether the policy performs a two- or three-stage fold. We train one

policy per category jointly on all demonstrated strategies for that category.

b) Model variants: We evaluate *FoldBack w/o Rollback-and-Repair*, which retains KAAI but disables grasp-outcome verification, rollback, and trajectory plan repair, and *FoldBack w/o Structured Repair*, which retains KAAI, grasp-verified history, and rollback but replaces structured repair with MGP-Long’s [9] confidence-based remasking of the failed window and remaining future, followed by a single repair candidate.

c) Baselines: We compare against the action-chunking policies 3D Diffusion Policy (DP3) [7] and SKIL [4], and the full-trajectory generative policy MGP-Long [9]. We also evaluate *DP3 + Verified Rollback*, which equips DP3 with the same grasp-outcome monitor and verified rollback as FoldBack without modifying or retraining the base policy. During normal execution, these baselines perform inference every 12 executed actions, whereas FoldBack refines its trajectory at key-action boundaries. Upon a detected failure, the DP3 variant returns to the pre-grasp configuration, discards the unexecuted actions in the current action chunk, and generates a new action chunk from the resulting observation. DP3, SKIL, and MGP-Long do not explicitly detect grasp failures or initiate retries, and therefore continue execution after a failed grasp.

d) Evaluation metrics: A recovery episode starts when a grasp failure is first detected in a folding stage and allows up to three rollback-and-retry attempts; further failures during those retries belong to the same episode. The episode ends when the stage completes, and the trial terminates if the stage is still incomplete after the third retry. *Recovery success*

TABLE I
REAL-WORLD LONG-HORIZON GARMENT-FOLDING PERFORMANCE. ‘VR’, ‘R&R’, AND ‘SR’ DENOTE VERIFIED ROLLBACK, ROLLBACK-AND-REPAIR, AND STRUCTURED REPAIR, RESPECTIVELY.

Method	Napkin				Towel				Pants			
	SC	RS	FS	IoU	SC	RS	FS	IoU	SC	RS	FS	IoU
DP3	71.7	–	56.7	0.717	60.5	–	53.3	0.703	66.7	–	60.0	0.706
SKIL	75.0	–	63.3	0.755	65.8	–	60.0	0.743	60.0	–	53.3	0.731
MGP-Long	78.3	–	70.0	0.786	73.7	–	66.7	0.791	73.3	–	66.7	0.764
FoldBack w/o R&R	83.3	–	73.3	0.828	81.6	–	73.3	0.815	80.0	–	66.7	0.786
DP3 + VR	85.0	53.8	76.7	0.823	78.9	60.0	73.3	0.814	76.7	55.6	66.7	0.788
FoldBack w/o SR	90.0	70.0	83.3	0.853	86.8	66.7	80.0	0.846	83.3	60.0	73.3	0.826
FoldBack(ours)	93.3	90.9	90.0	0.872	94.7	85.7	93.3	0.876	86.7	83.3	80.0	0.851

Method	Short-sleeved T-Shirt				Short-sleeved Shirt				Long-sleeved T-Shirt				Average			
	SC	RS	FS	IoU	SC	RS	FS	IoU	SC	RS	FS	IoU	SC	RS	FS	IoU
DP3	34.8	–	6.7	0.502	30.4	–	6.7	0.469	26.7	–	6.7	0.458	44.3	–	35.2	0.610
SKIL	42.0	–	13.3	0.558	33.3	–	13.3	0.501	34.7	–	6.7	0.548	48.7	–	39.0	0.656
MGP-Long	50.7	–	20.0	0.605	43.5	–	13.3	0.551	40.0	–	13.3	0.540	56.3	–	45.7	0.689
FoldBack w/o R&R	63.8	–	33.3	0.692	56.5	–	26.7	0.648	54.7	–	26.7	0.652	67.2	–	53.3	0.750
DP3 + VR	65.2	42.1	26.7	0.674	55.1	47.1	20.0	0.627	52.0	45.0	20.0	0.619	66.3	48.9	51.4	0.738
FoldBack w/o SR	75.4	61.9	46.7	0.761	66.7	55.0	33.3	0.716	66.7	63.6	40.0	0.711	76.2	61.9	62.9	0.795
FoldBack(ours)	82.6	79.2	60.0	0.811	76.8	78.3	53.3	0.782	77.3	77.8	60.0	0.793	83.9	80.6	75.2	0.837

rate (RS) is the fraction of episodes whose stage completes within this budget, and *first-retry success rate* (FRS) is the fraction of episodes that complete on the first retry. *Stage completion rate* (SC) is the fraction of required folding stages completed across all trials. We take a reference mask from the final top-down frame of a successful demonstration of the same garment and folding strategy, and we compute the IoU in image coordinates. Final-mask IoU is averaged over all trials, including incomplete ones. A trial counts as a *final folding success* (FS) if every stage completes and the final IoU is at least 0.80.

A. Long-Horizon Garment Folding

We conduct 30 trials on napkins (15 trials per napkin size) and 15 trials for each other category (i.e. pants, t-shirts, etc.).

Table I shows that MGP-Long is the strongest baseline without recovery. KAAI alone improves average final success from 45.7% to 53.3%, while full FoldBack reaches 75.2% final success and 0.837 final-mask IoU, achieving the best final success in all six categories; representative recovery executions are shown in Fig. 4. Structured repair further improves final success from 62.9% to 75.2%. Moreover, the gains are larger on the evaluated five-stage strategies. Against MGP-Long, the stage-completion gains are 13.4–15.0 points on the evaluated two-stage folds and 31.9–37.3 points on the three five-stage garment categories.

B. Recovery under Controlled Grasp-Slip Perturbations

Natural failures vary across trials, so we inject controlled one- or two-corner slips at a predetermined pick event during bimanual towel folding. During lifting, an operator dislodges either one corner (producing one failed and one successful grasp) or both corners (causing both grasps to fail). Each trial contains one perturbation and no subsequent intervention. We conduct 15 trials for each evaluated method–condition pair. *Failure-detection rate* (Det.) is the fraction of injected perturbations detected by the monitor, while *recovery success*

TABLE II
RECOVERY FROM CONTROLLED GRASP-SLIP PERTURBATIONS ON TOWEL FOLDING. ‘FB’, ‘VR’, ‘SR’, AND ‘FF’ DENOTE FOLDBACK, VERIFIED ROLLBACK, STRUCTURED REPAIR, AND FIXED FAILED WINDOW.

Method	One-Corner Slip				Two-Corner Slip		
	Det.↑	Pres. ↑	RS ↑	FS ↑	Det.↑	RS ↑	FS ↑
DP3	N/A	N/A	N/A	6.7	N/A	N/A	0.0
DP3 + VR	86.7	38.5	53.8	33.3	86.7	46.2	20.0
MGP-Long	N/A	N/A	N/A	13.3	N/A	N/A	6.7
FB w/o Rollback	93.3	92.9	57.1	40.0	93.3	42.9	13.3
FB w/o SR + FF	93.3	64.3	50.0	26.7	100.0	40.0	20.0
FB w/o SR	93.3	71.4	64.3	53.3	100.0	53.3	33.3
FoldBack (ours)	93.3	100.0	85.7	73.3	100.0	80.0	66.7

rate is the fraction of detected recovery episodes whose affected stage is completed within three recovery attempts. For one-corner slips, *successful-grasp preservation rate* (Pres.) is the fraction of detected asymmetric failures in which the successful grasp remains held from failure detection until recovery or termination. To isolate the effect of grasp-verified history, we compare FoldBack without Structured Repair with *FoldBack without Structured Repair with a Fixed Failed Window*, which treats the failed window as executed history and keeps it fixed after rollback, refining only the remaining future trajectory. We also evaluate *FoldBack without Rollback*, which uses the same grasp-outcome monitoring, grasp-verified history, and plan repair but repairs the trajectory directly from the failure configuration without returning to the pre-grasp state. We further evaluate the frozen monitor on held-out garments, using its first prediction after each lift to classify the grasp outcome. Labels are manually annotated independently of the monitor output, and the threshold is selected on the validation set and fixed for testing. Across 96 failed and 352 successful grasps, the monitor achieves 92.7% failed-grasp recall and an arm-wise false-positive rate of 4.5%. This indicates that the monitor detects most grasp

TABLE III

GENERALIZATION PERFORMANCE, REPORTED AS FINAL FOLDING SUCCESS RATE (%) / FINAL-MASK IOU. S-T, L-T, AND S-S DENOTE SHORT-SLEEVE T-SHIRT, LONG-SLEEVE T-SHIRT, AND SHORT-SLEEVE SHIRT, RESPECTIVELY.

Method	Initial Orientation (Pants)			Unseen Instance						Non-Canonical		
	Seen	Mod.	Large	Nap.	Tow.	Pants	S-T	L-T	S-S	Avg.	Nap.	Tow.
DP3	60.0/0.706	46.7/0.661	33.3/0.603	60.0/0.704	40.0/0.642	50.0/0.671	10.0/0.486	0.0/0.421	10.0/0.458	28.3/0.564	40.0/0.648	33.3/0.612
MGP-Long	66.7/0.764	60.0/0.724	40.0/0.642	70.0/0.774	50.0/0.710	60.0/0.731	20.0/0.596	10.0/0.521	10.0/0.530	36.7/0.644	53.3/0.704	46.7/0.674
FoldBack (ours)	80.0/0.851	73.3/0.817	60.0/0.752	90.0/0.863	80.0/0.833	70.0/0.812	60.0/0.780	50.0/0.756	50.0/0.742	66.7/0.798	80.0/0.826	73.3/0.798

failures while rarely triggering unnecessary recovery after a successful grasp.

Verified rollback improves DP3’s final folding success under both perturbations (Table II). Under one- and two-corner slips, keeping the failed window editable rather than fixing it after execution raises recovery success by 14.3 and 13.3 percentage points, respectively, while physical rollback yields corresponding gains of 28.6 and 37.1 points over FoldBack without Rollback.

Full FoldBack achieves recovery success of 85.7% and 80.0% and improves final folding success over FoldBack without Structured Repair by 20.0 and 33.3 points, respectively. Full FoldBack preserves the successful grasp in 100% of detected asymmetric failures because it keeps the successful gripper closed and reuses that arm’s recorded motion while regenerating only the failed arm. FoldBack without Structured Repair regenerates both arms jointly and preserves the successful grasp in only 71.4% of detected asymmetric failures.

C. Generalization

Table III compares DP3, MGP-Long, and FoldBack on unseen initial orientations, held-out garment instances, and non-canonical initial configurations. Qualitative examples are provided in the accompanying video.

Initial-orientation generalization. Using pants, we evaluate at the training orientation, where the garment is placed diagonally on the table as in the demonstrations, and two unseen rotation ranges relative to this orientation: moderate (5° – 20°) and large (20° – 30°). Trials at the training orientation reuse the pants trials from Table I, while each unseen rotation range contains 15 trials per method. FoldBack achieves final success rates of 80.0%, 73.3%, and 60.0%, compared with 66.7%, 60.0%, and 40.0% for MGP-Long.

Unseen garment instances. For each category, we hold out one physical garment from training and retrain each method on the remaining demonstrations. Ten trials per held-out instance yield 60 trials per method. FoldBack achieves 66.7% average final success, compared with 36.7% for MGP-Long and 28.3% for DP3, and performs best in all six categories.

Non-canonical initial configurations. We initialize napkins and towels in near-flat configurations with local wrinkles or a partially folded corner, requiring the robot to flatten before folding. This evaluates whether the learned flatten-and-fold behavior transfers to unseen non-canonical initial configurations. Across 15 trials per category, FoldBack achieves 80.0% and 73.3% final success on napkins and

TABLE IV
THREE ABLATIONS ON TOWEL FOLDING.

(a) Schedule			(b) Remasking			(c) Resampling			
Variant	SC $\uparrow$	FS $\uparrow$	Variant	RS $\uparrow$	FS $\uparrow$	Variant	FRS $\uparrow$	RS $\uparrow$	FS $\uparrow$
Fixed-12	73.7	66.7	Full future	75.0	80.0	Single	55.6	77.8	80.0
Uniform	76.3	66.7	Failed window	71.4	80.0	Conf.-ranked	66.7	83.3	86.7
KAAI	81.6	73.3	Dist.-guided	85.7	93.3	Failure-aware	71.4	85.7	93.3

towels, respectively, versus 53.3% and 46.7% for MGP-Long. Examples are provided in the accompanying video.

D. Ablation Studies

We conduct three ablations on towel-folding with 15 trials per variant (Table IV). To isolate event alignment from update frequency, we compare three refinement strategies using the same trained policy with grasp monitoring and rollback-and-repair disabled. *Fixed-12* refines every 12 executed actions, *Uniform* distributes the same number of refinements as KAAI uniformly over the horizon, and *KAAI* refines before and after each key-action window. KAAI achieves the highest stage completion and final success; because Uniform uses the same number of refinements, this comparison isolates the benefit of event alignment. We next compare three remasking strategies. *Full future* regenerates the entire unexecuted trajectory, whereas *Failed window* regenerates only the failed key-action window. *Distance-guided* fully remasks the failed window and progressively lowers the masking ratio over subsequent segments, prioritizing low-confidence tokens. It achieves the highest recovery success (85.7%) and final folding success (93.3%). Finally, we compare three candidate-selection strategies. *Single* draws one repair candidate, *Confidence ranking* selects the highest-confidence candidate among $Q = 8$, and *Failure-aware* additionally rejects candidates whose failed-arm pick lies within the exclusion region around previous failed grasps. Failure-aware achieves the highest first-retry, recovery, and final success rates: 71.4%, 85.7%, and 93.3%. Batched resampling takes 58 ms per recovery event versus 37 ms for a single candidate, adding little inference latency.

V. CONCLUSION

We introduced FoldBack, a self-correcting masked generative framework for long-horizon garment folding. FoldBack maintains a grasp-verified trajectory history so that failed grasp windows can be updated rather than becoming fixed trajectory history. Upon failure, it returns the robot to a retryable pre-grasp configuration, selectively repairs the trajectory with distance-guided remasking, and resamples

candidates while avoiding previously failed grasp locations. Across six garment categories, FoldBack improves stage completion, recovery success, and final folding performance over action-chunking and full-trajectory baselines. Future work will extend FoldBack to more challenging initial configurations that require a separate flattening phase before folding can begin.

REFERENCES

- [1] Y. Chen, B. Xiao, and H. Wang, “Foldnet: Learning generalizable closed-loop policy for garment folding via keypoint-driven asset and demonstration synthesis,” *IEEE Robotics Autom. Lett.*, vol. 11, no. 3, pp. 3820–3827, 2026.
- [2] H. Xue, Y. Li, W. Xu, H. Li, D. Zheng, and C. Lu, “Unifolding: Towards sample-efficient, scalable, and generalizable robotic garment folding,” in *Conference on Robot Learning (CoRL)*, vol. 229, 2023, pp. 3321–3341.
- [3] L. Zhuang, S. Fan, Y. Ru, F. P. Audonnet, P. Henderson, and G. Aragon-Camarasa, “Flat’n’fold: A diverse multi-modal dataset for garment perception and manipulation,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2025, pp. 7937–7944.
- [4] S. Wang, J. You, Y. Hu, J. Li, and Y. Gao, “SKIL: semantic keypoint imitation learning for generalizable data-efficient manipulation,” *CoRR*, 2025.
- [5] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, “Learning fine-grained bimanual manipulation with low-cost hardware,” in *Robotics: Science and Systems (RSS)*, 2023.
- [6] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, “Diffusion policy: Visuomotor policy learning via action diffusion,” *Int. J. Robotics Res.*, vol. 44, no. 10-11, pp. 1684–1704, 2025.
- [7] Y. Ze, G. Zhang, K. Zhang, C. Hu, M. Wang, and H. Xu, “3d diffusion policy: Generalizable visuomotor policy learning via simple 3d representations,” in *Robotics: Science and Systems (RSS)*, 2024.
- [8] Z. Xian, N. Gkanatsios, T. Gervet, T. Ke, and K. Fragkiadaki, “Chaineddiffuser: Unifying trajectory diffusion and keypose prediction for robotic manipulation,” in *Conference on Robot Learning (CoRL)*, vol. 229, 2023, pp. 2323–2339.
- [9] L. Zhuang, S. Fan, F. P. Audonnet, Y. Ru, G. Aragon-Camarasa, and P. Henderson, “Masked generative policy for robotic control,” *CoRR*, vol. abs/2512.09101, 2025.
- [10] C. Agia, R. Sinha, J. Yang, Z.-a. Cao, R. Antonova, M. Pavone, and J. Bohg, “Unpacking failure modes of generative policies: Runtime monitoring of consistency and progress,” *arXiv preprint arXiv:2410.04640*, 2024.
- [11] C. Xu, T. K. Nguyen, E. Dixon, C. Rodriguez, P. Miller, R. Lee, P. Shah, R. Ambrus, H. Nishimura, and M. Itkina, “Can we detect failures without failure data? uncertainty-aware runtime failure detection for imitation learning policies,” *arXiv preprint arXiv:2503.08558*, 2025.
- [12] C. Gokmen, D. Ho, and M. Khansari, “Asking for help: Failure prediction in behavioral cloning through value approximation,” in *IEEE International Conference on Robotics and Automation, ICRA*, 2023, pp. 5821–5828.
- [13] Y. Dai, J. Lee, N. Fazeli, and J. Chai, “Racer: Rich language-guided failure recovery policies for imitation learning,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2025, pp. 15 657–15 664.
- [14] Z. Liu, A. Bahety, and S. Song, “Reflect: Summarizing robot experiences for failure explanation and correction,” *arXiv preprint arXiv:2306.15724*, 2023.
- [15] C. Ma, K. Lu, G. Yang, J. Liu, S. Xu, B. Byrne, I. Havoutis, N. Trigoni, and A. Markham, “Cyclevla: Proactive self-correcting vision-language-action models via subtask backtracking and minimum bayes risk decoding,” *arXiv preprint arXiv:2601.02295*, 2026.
- [16] H. Hao, S. N. Syed, J. Ichnowski, and J. Schneider, “Far: Failure-aware retry for test-time recovery and continual policy improvement,” *arXiv preprint arXiv:2607.01111*, 2026.
- [17] K. Mo, C. Xia, X. Wang, Y. Deng, X. Gao, and B. Liang, “Foldsformer: Learning sequential multi-step cloth manipulation with space-time attention,” *IEEE Robotics Autom. Lett.*, vol. 8, no. 2, pp. 760–767, 2023.
- [18] A. Canberk, C. Chi, H. Ha, B. Burchfiel, E. Cousineau, S. Feng, and S. Song, “Cloth funnels: Canonicalized-alignment for multi-purpose garment manipulation,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 5872–5879.
- [19] T. Weng, S. M. Bajracharya, Y. Wang, K. Agrawal, and D. Held, “Fabricflownet: Bimanual cloth manipulation with a flow-based policy,” in *Conference on Robot Learning (CoRL)*, vol. 164, 2021, pp. 192–202.
- [20] Y. Avigal, L. Berscheid, T. Asfour, T. Kröger, and K. Goldberg, “Speedfolding: Learning efficient bimanual folding of garments,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 1–8.
- [21] O. Barbany, A. Colomé, and C. Torras, “Bifold: Bimanual cloth folding with language guidance,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2025, pp. 5652–5659.
- [22] X. Lin, Y. Wang, Z. Huang, and D. Held, “Learning visible connectivity dynamics for cloth smoothing,” in *Conference on Robot Learning (CoRL)*, vol. 164, 2021, pp. 256–266.
- [23] Z. Huang, X. Lin, and D. Held, “Mesh-based dynamics with occlusion reasoning for cloth manipulation,” in *Robotics: Science and Systems (RSS)*, 2022.
- [24] T. Tian, H. Li, B. Ai, X. Yuan, Z. Huang, and H. Su, “Diffusion dynamics models with generative state estimation for cloth manipulation,” *CoRR*, vol. abs/2503.11999, 2025.
- [25] W. Chen, H. Xue, F. Zhou, Y. Fang, and C. Lu, “Deformpam: Data-efficient learning for long-horizon deformable object manipulation via preference-based action alignment,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2025, pp. 6896–6903.
- [26] Y. Wang, R. Wu, Y. Chen, J. Wang, J. Liang, Z. Zhu, H. Geng, J. Malik, P. Abbeel, and H. Dong, “Dexgarmentlab: Dexterous garment manipulation environment with generalizable policy,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 38, 2025.
- [27] H. Chen, J. Li, R. Wu, Y. Liu, Y. Hou, Z. Xu, J. Guo, C. Gao, Z. Wei, S. Xu, J. Huang, and L. Shao, “Metafold: Language-guided multi-category garment folding framework via trajectory generation and foundation model,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025, pp. 4339–4346.
- [28] T.-W. Ke, N. Gkanatsios, and K. Fragkiadaki, “3D diffuser actor: Policy diffusion with 3D scene representations,” *arXiv preprint arXiv:2402.10885*, 2024.
- [29] N. M. M. Shafiullah, Z. J. Cui, A. Altanzaya, and L. Pinto, “Behavior transformers: Cloning k modes with one stone,” *arXiv preprint arXiv:2206.11251*, 2022.
- [30] S. Lee, Y. Wang, H. Etukuru, H. J. Kim, N. M. M. Shafiullah, and L. Pinto, “Behavior generation with latent actions,” *arXiv preprint arXiv:2403.03181*, 2024.
- [31] A. Prasad, K. Lin, J. Wu, L. Zhou, and J. Bohg, “Consistency policy: Accelerated visuomotor policies via consistency distillation,” in *Robotics: Science and Systems*, 2024.
- [32] Q. Zhang, Z. Liu, H. Fan, G. Liu, B. Zeng, and S. Liu, “Flowpolicy: Enabling fast and robust 3d flow-based policy via consistency flow matching for robot manipulation,” in *AAAI Conference on Artificial Intelligence (AAAI)*, vol. 39, no. 14, 2025, pp. 14 754–14 762.
- [33] K. Black, A. Z. Ren, M. Equi, and S. Levine, “Training-time action conditioning for efficient real-time chunking,” *arXiv preprint arXiv:2512.05964*, 2025.
- [34] K. Black, M. Y. Galliker, and S. Levine, “Real-time execution of action chunking flow policies,” in *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- [35] M. Du, A. Khazatsky, T. Gerstenberg, and C. Finn, “To err is robotic: Rapid value-based trial-and-error during deployment,” *CoRR*, vol. abs/2406.15917, 2024.
- [36] G. Zheng, S. Seenivasan, M. Johnson-Roberson, and W. Zhi, “Rewind-il: Online failure detection and state respawning for imitation learning,” *CoRR*, vol. abs/2604.16683, 2026.
- [37] A. Van Den Oord, O. Vinyals *et al.*, “Neural discrete representation learning,” *Advances in neural information processing systems*, vol. 30, 2017.
- [38] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 652–660.