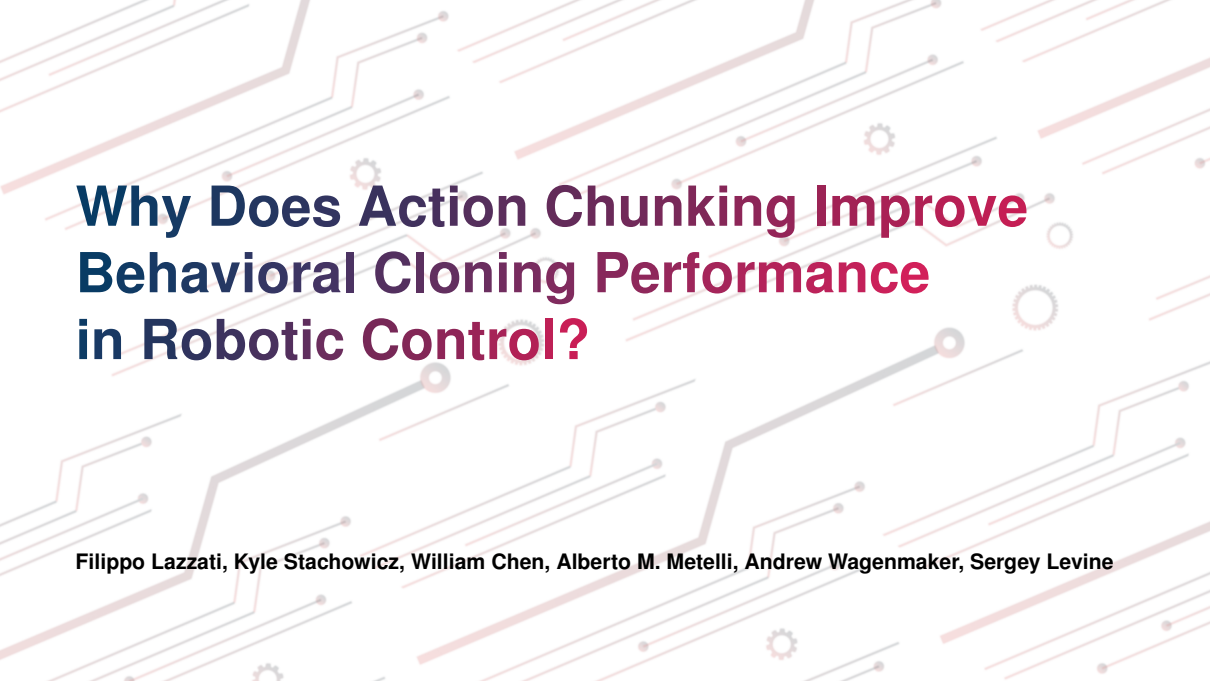




Why Does Action Chunking Improve Behavioral Cloning Performance in Robotic Control?

Filippo Lazzati, Kyle Stachowicz, William Chen, Alberto M. Metelli, Andrew Wagenmaker, Sergey Levine

Introduction

- **Action Chunking** → predicting and executing a *sequence* of actions rather than a single action

Introduction

- **Action Chunking** → predicting and executing a *sequence* of actions rather than a single action
- Virtually all state-of-the-art robot policies use some form of action chunking

Introduction

- **Action Chunking** → predicting and executing a *sequence* of actions rather than a single action
- Virtually all state-of-the-art robot policies use some form of action chunking
- In this work, we
 1. Investigate common hypothesis
 2. Find a new mechanism “implicit ensembles”
 3. Propose a new method based on it

Existing Hypothesis

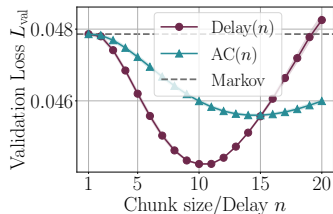
- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)

Existing Hypothesis

- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)
- We compared the validation loss of a Markov policy $a_h|s_h$ with those of delayed $a_h|s_{h-n}$ and chunked policies $a_{h:h+n}|s_h$

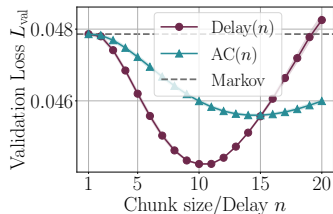
Existing Hypothesis

- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)
- We compared the validation loss of a Markov policy $a_h|s_h$ with those of delayed $a_h|s_{h-n}$ and chunked policies $a_{h:h+n}|s_h$



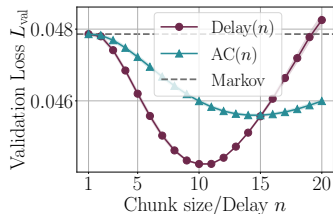
Existing Hypothesis

- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)
- We compared the validation loss of a Markov policy $a_h|s_h$ with those of delayed $a_h|s_{h-n}$ and chunked policies $a_{h:h+n}|s_h$
- **H2: Horizon reduction**
“Action chunks mitigate compounding error by reducing the effective horizon” (Zhang et al. 2025)

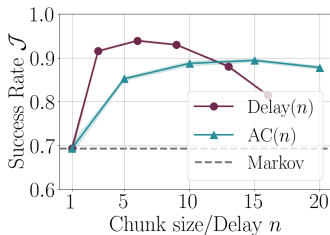


Existing Hypothesis

- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)
- We compared the validation loss of a Markov policy $a_h|s_h$ with those of delayed $a_h|s_{h-n}$ and chunked policies $a_{h:h+n}|s_h$

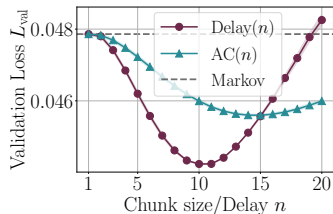


- **H2: Horizon reduction**
“Action chunks mitigate compounding error by reducing the effective horizon” (Zhang et al. 2025)

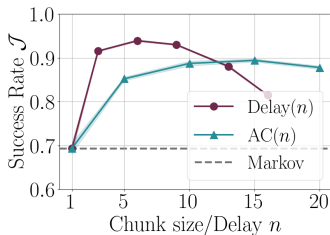


Existing Hypothesis

- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)
- We compared the validation loss of a Markov policy $a_h|s_h$ with those of delayed $a_h|s_{h-n}$ and chunked policies $a_{h:h+n}|s_h$



- **H2: Horizon reduction**
“Action chunks mitigate compounding error by reducing the effective horizon” (Zhang et al. 2025)



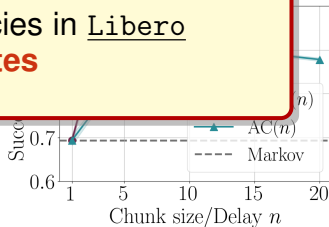
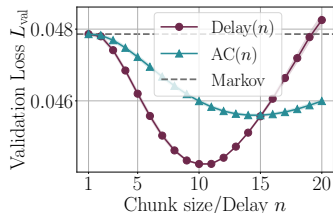
- Markov: $\Omega(2^H \cdot \epsilon)$
Chunked/delayed: $\mathcal{O}((n+1)^{H/n} \cdot \epsilon)$

Existing Hypothesis

- **H1: Non-Markovian expressivity**
“Action-chunked policies better capture non-Markov human behaviors” (Zhao et al. 2023)
- **H2: Horizon reduction**
“Action chunks mitigate compounding error by reducing the effective horizon” (Zhang et al. 2025)

- We compare a Markov delayed policies and

Delayed policies perform as well as **action chunking** policies in Libero
→ **past states**



- Markov: $\Omega(2^H \cdot \epsilon)$
Chunked/delayed: $\mathcal{O}((n+1)^{H/n} \cdot \epsilon)$

Existing Hypothesis

- **H1: Non-Markovian expressivity**

“Action-chunked policies better

capture non-Markovian dependencies

behaviors

- **H2: Horizon reduction**

“Action chunks mitigate

reducing the

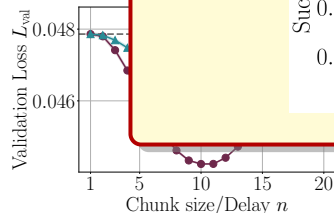
ang et al.

- We compare

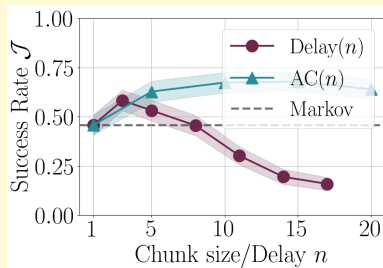
a Markov

delayed action

policies and



However, **action chunking**
is still **better** in Robomimic



- Markov: $\Omega(2^H \cdot \epsilon)$

Chunked/delayed: $\mathcal{O}((n+1)^{H/n} \cdot \epsilon)$

A New Mechanism

- **Implicit Ensembles:**

“Action chunking implicitly learns and executes an ensemble of predictors

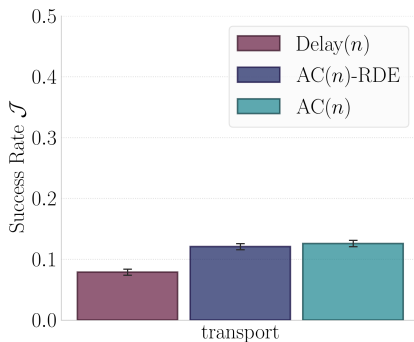
$a_h \mid o_h, a_h \mid o_{h-1}, \dots, a_h \mid o_{h-k}$ instead of a single one $a_h \mid (o_h, o_{h-1}, \dots, o_1)$ ”

A New Mechanism

- **Implicit Ensembles:**

“Action chunking implicitly learns and executes an ensemble of predictors

$a_h \mid o_h, a_h \mid o_{h-1}, \dots, a_h \mid o_{h-k}$ instead of a single one $a_h \mid (o_h, o_{h-1}, \dots, o_1)$ ”



A New Mechanism

- **Implicit**

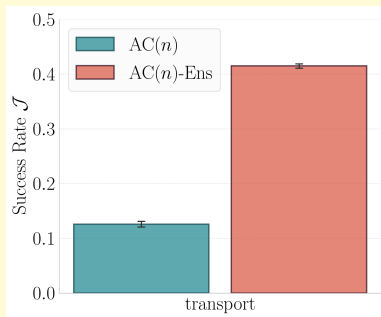
“Action c

$a_h \mid o_h, a$

Explicit ensembles of chunked/delayed policies are even *better*

predictors

$o_{t-1}, \dots, o_1)$ ”



Thank you for your attention!

References |

-  Zhang, Thomas T. et al. (2025). *Action Chunking and Exploratory Data Collection Yield Exponential Improvements in Behavior Cloning for Continuous Control*. arXiv: 2507.09061 [cs.LG]. URL: <https://arxiv.org/abs/2507.09061>.
-  Zhao, Tony Z. et al. (2023). *Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware*. arXiv: 2304.13705 [cs.RO]. URL: <https://arxiv.org/abs/2304.13705>.