

ABLATION AND SENSITIVITY ANALYSIS

Domain	Task	TD-M(PC) <sup>2</sup> Critic + RIC (Ours)	TD-M(PC) <sup>2</sup> (World Model, no BC)	IQL (No World Model, no BC)	IQL Critic + Ours (Ours, No World Model)
D3IL	Sorting (2)	0.78 ± 0.05	0.53 ± 0.04	0.48 ± 0.15	0.91 ± 0.05
PushT	PushT	0.92 ± 0.07	0.62 ± 0.06	0.11 ± 0.02	0.85 ± 0.06

Table 4: Qualitative ablation results for D3IL and PushT, isolating the contributions of the world model and critic type. Each column shows the mean ± std success rate for the specified configuration.

Task	$\alpha = 0.0$	0.25	0.5	1.0	2.0
PushT	0.55 ± 0.05	0.63 ± 0.07	0.74 ± 0.05	0.85 ± 0.06	0.85 ± 0.07

Table 5: Sensitivity analysis of PushT task success rate as a function of the RL loss coefficient  $\lambda_{RL}$ . Each column shows the mean ± std success rate for a different value of  $\lambda_{RL}$ .

A EXPERIMENTAL DETAILS AND HYPERPARAMETERS

We detail the core training, network architecture, and algorithm-specific hyperparameters for all evaluated methods and tasks. Tables 6, 7, and 8 summarize the settings used for each algorithm and environment.

Table 6: Core Training Hyperparameters

Algorithm	
<b>VQBET</b>	
Learning Rate	1e-4 (main)
Training Steps	2,000,000
Batch Size	64
<b>TDMPC<sup>2</sup></b>	
Learning Rate	3e-4
Training Steps	500,000
Batch Size	256
<b>Diffusion</b>	
Learning Rate	1e-5
Training Steps	2,000,000
Batch Size	64
<b>RIC</b>	
Learning Rate	1e-4
Training Steps	2,000,000
Batch Size	64

Table 7: Network Architecture Parameters

Algorithm		
<b>VQBET</b>		
GPT Layers	8	
GPT Hidden Dim	512	
MLP Hidden Dim	1024	
VQVAE Embedding Dim	256	
<b>TDMPC<sup>2</sup></b>		
State Encoder Hidden	256	
Latent Dim	512	
Task Dim	96	
MLP Dim	512	
Q Ensemble Size	5	
<b>Diffusion</b>		
Down Dims	512,	1024,
	2048	
Diffusion Embed Dim	128	
Kernel Size	5	
<b>RIC</b>		
Task Dim	32	
MLP Dim	256	
Latent Dim	256	
Image Encoder Hidden	256	

Table 8: Algorithm-Specific Parameters

Algorithm	
<b>VQBET</b>	
N Obs Steps	5
Action Pred Tokens	3
Action Chunk Size	2
Block Size	500
<b>TDMPC<sup>2</sup> (22)</b>	
Horizon	10
N Action Steps	10
N Gaussian Samples	512
CEM Iterations	6
N Elites	64
Reward Coef.	0.1
Value Coef.	0.1
Entropy Coef. ( $\alpha$ )	1e-4
Scale threshold ( $s$ )	2.0
Prior Constraint ( $\lambda$ )	1.0
<b>Diffusion</b>	
N Obs Steps	2
Horizon	8
N Action Steps	2
Timesteps	100
Inference Steps	16
<b>RIC</b>	
N Obs Steps	5
Action Pred Tokens	3
Action Chunk Size	2
Num Bins	100

## NEURIPS PAPER CHECKLIST

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and follow the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer , , or .
- means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1–2 sentence) justification right after your answer (even for NA).

**The checklist answers are an integral part of your paper submission.** They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While " " is generally preferable to " ", it is perfectly acceptable to answer " "

provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering "" or "" is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading “NeurIPS Paper Checklist”,**
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers.**

### 1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper’s contributions and scope?

Answer:

Justification: The abstract highlights the contribution that our method can improve the quality of learned policies by as much as 20% in the presence of sub-optimal demonstration data. This result is justified by our results which show that when averaged across a set of randomly seeded runs of our method, our method achieves the Stacking and Push-T tasks 17% and 20% better with noisy demonstrations more than the closest baseline.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

### 2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer:

Justification: The effectiveness of RIC depends on the quality and diversity of the demonstrations provided, to ensure that the task is still solvable in spite of the suboptimal demonstrations provided. This is so that the learned model of the environment can be useful to the critic when estimating the expected return. We note that when using a model-based critic (TDMPC2) when the number of failed demonstrations is high, the action values are overestimated, leading to suboptimal performance when using the critic guidance term. This indicates that our algorithm is sensitive to the distribution of noise and failure modes in the data. Furthermore, we note that the critic training can be unstable and difficult to optimize, resulting in a critic guidance parameter to tune. Finally, we do not fully examine the computational efficiency or scalability of RIC, specifically choosing limited, smaller datasets in simulation.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

### 3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer:

Justification: We do not include a theoretical result.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

### 4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer:

Justification: Our method is reproducible by following the procedure for the algorithm as shown in the algorithm block and detailed by the methods, and following the hyperparameter settings included in the tables. We do not leverage any closed source models or datasets.

756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

## 5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer:

Justification: The code is pending for clarity and will be made available through the associated anonymized website.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.

- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

## 6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer:

Justification: Provided with the code, and in tables.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

## 7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer:

Justification: We provide error bars with 5 random seeds per run.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

## 8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer:

864 Justification: We provide details on our compute resources in the experiments  
865 section.

866 Guidelines:

- 867 • The answer NA means that the paper does not include experiments.
- 868 • The paper should indicate the type of compute workers CPU or GPU, internal  
869 cluster, or cloud provider, including relevant memory and storage.
- 870 • The paper should provide the amount of compute required for each of the  
871 individual experimental runs as well as estimate the total compute.
- 872 • The paper should disclose whether the full research project required more  
873 compute than the experiments reported in the paper (e.g., preliminary or failed  
874 experiments that didn't make it into the paper).
- 875

## 876 9. Code of ethics

877 Question: Does the research conducted in the paper conform, in every respect, with  
878 the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines?>

879 Answer:

880 Justification: We reviewed and abide by the ethics guidelines.

881 Guidelines:

- 882 • The answer NA means that the authors have not reviewed the NeurIPS Code  
883 of Ethics.
- 884 • If the authors answer No, they should explain the special circumstances that  
885 require a deviation from the Code of Ethics.
- 886 • The authors should make sure to preserve anonymity (e.g., if there is a special  
887 consideration due to laws or regulations in their jurisdiction).
- 888

## 889 10. Broader impacts

890 Question: Does the paper discuss both potential positive societal impacts and  
891 negative societal impacts of the work performed?

892 Answer:

893 Justification: Our work does not cause societal impact.

894 Guidelines:

- 895 • The answer NA means that there is no societal impact of the work performed.
- 896 • If the authors answer NA or No, they should explain why their work has no  
897 societal impact or why the paper does not address societal impact.
- 898 • Examples of negative societal impacts include potential malicious or unintended  
899 uses (e.g., disinformation, generating fake profiles, surveillance), fairness consid-  
900 erations (e.g., deployment of technologies that could make decisions that unfairly  
901 impact specific groups), privacy considerations, and security considerations.
- 902 • The conference expects that many papers will be foundational research and  
903 not tied to particular applications, let alone deployments. However, if there  
904 is a direct path to any negative applications, the authors should point it out.  
905 For example, it is legitimate to point out that an improvement in the quality  
906 of generative models could be used to generate deepfakes for disinformation.  
907 On the other hand, it is not needed to point out that a generic algorithm for  
908 optimizing neural networks could enable people to train models that generate  
909 Deepfakes faster.
- 910 • The authors should consider possible harms that could arise when the technology  
911 is being used as intended and functioning correctly, harms that could arise when  
912 the technology is being used as intended but gives incorrect results, and harms  
913 following from (intentional or unintentional) misuse of the technology.
- 914 • If there are negative societal impacts, the authors could also discuss possible  
915 mitigation strategies (e.g., gated release of models, providing defenses in addition  
916 to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a  
917 system learns from feedback over time, improving the efficiency and accessibility  
of ML).

918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971

## 11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer:

Justification: No risks caused by our model.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

## 12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer:

Justification: We use public open source data and tools

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, [paperswithcode.com/datasets](https://paperswithcode.com/datasets) has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

## 13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer:

Justification: No new assets

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.

- 972
- The paper should discuss whether and how consent was obtained from people whose asset is used.
  - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

973

974

975

976

977

14. **Crowdsourcing and research with human subjects**

978 Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

979

980

981 Answer:

982 Justification: No human subjects are used.

983 Guidelines:

- 984
- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
  - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
  - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

985

986

987

988

989

990

991

992

993

994

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

995 Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

996

997

998

999 Answer:

1000 Justification: No IRB approvals required

1001 Guidelines:

- 1002
- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
  - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
  - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
  - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

16. **Declaration of LLM usage**

1014 Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

1015

1016

1017

1018

1019 Answer:

1020 Justification: No usage of LLMs

1021 Guidelines:

- 1022
- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
  - Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.
- 1023
- 1024
- 1025