We thank all reviewers for their constructive and helpful comments that will allow us to better shape this paper.

1 Reviewer I:

We are very thankful for the review and will definitely increase our plot sizes in the final version in case of acceptance. Also, thank you for pointing out real-world experiments. In fact, we plan to take our approach to robotics in the future. We believe self-driving cars present an ideal test-bed for our algorithm.

2 Reviewer II:

Thank you for the constructive and informative review. We will make sure to increase the figure sizes.

On Few-Shot Learning We achieve few-shot learning through meta-learning. Concretely, our model learns shared dynamics across different views and could be used as the transition model for learning from a new-coming view.

On PILCO These result are shown on both the Cart-Pole (Table I in the submission) and Hopper (Fig. (4) in the submission). We also attempted PILCO on Race-Car and the parking environments. Unfortunately, we did not get good performance on these tasks. We think this is due to the very sparse reward signals in these problems – a setting less suitable for PILCO and can potentially be remedied by shaping the reward function. We chose to stick standard rewards and plan to tackle complete model-based RL in the future. We also found that the combination of model-based and model-free is more suitable for general tasks to reach higher rewards with fewer training data. We will make sure to clarify these in the main paper as well.

3 Reviewer IV:

Thank you for the feedback that will help us improve our paper. We will make sure to fix the typos noted by the reviewer.

On the distribution of i_t We will clarify in the original paper. i_t is allowed to vary from one to N with being the different types of observations. Each i_t is dependent on the observation function (availability of specific views) and conditioned on the environment's state.

On X-Axis and Log-Scale We use Log-Scale to compare the performance of model-based (MB) and model-free (MF) in a single plot (e.g., PILCO and PPO), since MB methods require much fewer samples than MF. We follow a standard protocol for stopping criteria, where training is halted when the difference of prediction loss p (and reconstruction loss r) between two training steps is less than a threshold (4) for a continuous of 5000 steps. Please note that the Y-Axis of Fig. (2) is log loss, we will change it to original loss with log-scale.

On Comparison Against PPO We respectfully disagree with the reviewer and firmly believe that our comparison against PPO is fair. Namely, we compare against PPO by using the same amount of data from the same views. We showed that our model outperforms PPO in view-average return because it learns the shared latent state and joint transitions. Auxiliary tasks comparison can be interesting direction in the future but we think current results clearly demonstrate our method is superior to PPO in a fair manner (both using same amount of data).

On View Choices for Modeling Experiments Our modeling experiments are designed to illustrate the capability of our model learning the shared latent states and dynamics across different views. We showed that our model could effectively learn underlying shared dynamics in contrast to the state-of-the-art world model.

On Toy Problem and More Experiments We thank the reviewer for pointing this out. We plan to have a toy problem to better illustrate the performance. Also, we plan to improve the figure sizes. We also ran multiple runs and have variance plots. We will add in the main paper. We refrain from showing them here so as to keep with the fairness and page limit.