We would like to thank all reviewers for their constructive feedback and insightful comments. Here, our answers:

Comments to all reviewers First, we want to emphasize that our work is the first step towards the control of partial-observable continuous time discrete state space systems using modern machine learning methods. As the reviewers have clearly noticed and we also mentioned in the paper, the method in its current form is limited to low dimensional problems, since we solve the optimal non-linear filtering problem exactly. The toy problems function to explain the intuition behind the presented methodology and show that the proposed method works in principle. As the theory for the method in its current form is already quite complex, we think that further approximation methods would go way beyond the scope of a single paper. That being said, we will incorporate further explanations of the model and the underlying methodology to increase comprehensibility and to emphasize its applicability. This will also include discussing further related work areas such as stochastic hybrid systems and queueing networks, which are used, e.g., for TCP congestion control. We will add a short discussion on the control of partial observable continuous state space problems. The solution of these problems, however, involves even more difficulties since they require the stochastic optimal control framework under stochastic PDE dynamics and is thus left unsolved for future work.

Reviewer #1 Time discretization is definitively possible but would go along with the need for pseudo observations at time points at which no observations are emitted by the latent process. This matter even represents a modelling problem on its own. Furthermore, the main benefit of not having an a priori time discretization of the problem is that the analysis can be carried out in continuous time as described in the paper. This way, the numerical errors occurring due to time discretization on a digital computer are controlled by sophisticated numerical differential equation solvers, which control the error automatically, by, e.g., adaptive step size methods.

The gridworld problem is one of the standard benchmarks for discrete spaces. From our point of view, a continuous space relaxation of this problem would defeat its purpose. Even fully-observable continuous state and continuous time problems are historically solved by a state space discretization, as in the work of [1]. Additionally, the partial observable case involves even more difficulties, for further details see Reviewer #2.

In our opinion, a direct comparison to advanced discrete time POMDP solvers would not be fair. Our proposed method does not aim at outperforming discrete methods for small problems but provides a principled way to approach continuous time discrete space problems. The main advantage to discrete methods becomes substantial when further approximations of the filtering distributions are introduced and larger problems are tackled. As already described, we leave the introduction of these approximations for future work to limit complexity.

Reviewer #2 Sorry for the misnomer, we will scrap the word "countable" in "finite countable set". We just thought of the corresponding term for "countably infinite" which led to this accident.

As you already noted, continuous state space problems would involve, e.g., SDE dynamics of the latent process and the observation process. Therefore, the control of continuous state space problems under partial observability involves the control of the filtering distribution, which is described by a probability density. As the time evolution of the density is given by the Kushner-Stratonovich, a stochastic PDE, even simulating the system under a policy is a very hard problem. The main difficulty here is that the sufficient statistics of a general probability density are infinite dimensional. To approach this problem, one would have to resort to finding an efficient projection to a finite dimensional space by, e.g., looking at the first and second order moments. It is well known that for some special cases the filtering distribution has finite dimensional sufficient statistics. For example for a linear observation process with linear latent dynamics, the Kalman filter describes the evolution of the first and second order moments of the Gaussian filtering density. One then could use the finite dimensional mean and covariance inside the function approximator to estimate the value function even for arbitrary reward functions in contrast to the LQG method. However, for non-linear problems one would have to resort to approximations which find a finite dimensional projection, as it is done in, e.g., assumed density filtering.

Reviewer #3 We will use the extra page of the camera-ready version for additional explanations to make the intuition behind the methodology clear and increase comprehensibility. Regarding your comment on time discretization, several new difficulties would have to be discussed as mentioned in Reviewer #1.

Reviewer #4 The related work which you suggested, addresses the problem of time-discrete approximate optimal filtering in hybrid systems. In our paper, the equations for optimal filtering are by no means a contribution but originate from prior work which we referenced in the corresponding section. Instead, the focus lies on the problem of controlling these partial-observable systems. A further minor difference is that we approach the problem in continuous time as many original processes are likely to naturally evolve this way. However, the topic of approximate optimal filtering and hybrid systems seems a fruitful direction for future work when considering real world applications. We will therefore add relations to these topics as future work directions to the conclusion section.

References

[1] H. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24. Springer Science & Business Media, 2013.