I am thankful to the reviewers for their careful reading of the paper and their helpful comments. I will fix/revise all minor issues and add an analysis of the function approximation error, which shows that the bounds are non-vacuous. I also emphasize the motivation of this work. Before answering some of the comments in detail, I would like to emphasize that this work opens up a new approach to represent uncertainty of the returns in RL. It provides the fundamental theoretical guarantees that one needs before developing sophisticated algorithms, and empirically evaluating them.

R4: Non-vacuousness of the bounds in Theorems 2 and 3?
A: The bounds are well-behaving under mild conditions for $p = 1$. It is true that $\|\tilde{\varepsilon}\|_{\infty,p}$ might be infinity for $p > 1$ unless we have restrictive conditions, but its behaviour is reasonable (to be specified) for $p = 1$. Thanks to your comment on this issue, I investigated the approximation error properties for some reasonable choices of $\mathcal{F}$. The result, briefly speaking, is that if the reward distribution is smooth, a band-limited function class $\mathcal{F}_b$ provides an approximation error that goes to zero as b increases. Furthermore, if the first s absolute moments of the reward distribution is finite (uniformly for all $x \in \mathcal{X}$), the CVF $\tilde{V}(\cdot; x)$ belongs to $C^s([-b, b]) \cap \mathcal{F}_b$. This leads to well-behaving covering number, which can be used to obtain a convergence rate for estimation error.

Let us define $\mathcal{F}_b$ as the space of CF with bandwidth of b, i.e., $\tilde{V}(\omega; x)$ is zero for $|\omega| > b$. Assume that the reward function is β-smooth in the sense that $c_0|\omega|^{-\beta} \le |\tilde{R}(\omega; x)| \le c_1|\omega|^{-\beta}$ for $|\omega|$ large enough (Jianqing Fan, Annals of Statistics, 1991), which is satisfied by exponential, uniform, gamma, etc. distributions. We can also define super-smooth distributions, with examples such as normal or Cauchy. Let us focus on the approximation error of solving the regression problem in Eq. (14). At each iteration we may pick $\tilde{V}_{k+1}(\omega; x) = (\tilde{T}^\pi \tilde{V}_k)(\omega; x)\mathbb{I}\{\omega \in [-b, +b]\}$. This function is in $\mathcal{F}_b$. Because of the β-smoothness of $\tilde{R}$, the function approximation error $\tilde{\varepsilon}_{k+1,\mathrm{AE}} = \tilde{V}_{k+1} - \tilde{T}^\pi \tilde{V}_k$ satisfies $\|\tilde{\varepsilon}_{k+1,\mathrm{AE}}\|_{\infty,1} \le c_1 b^{-(1+\beta)}$ (and faster for super-smooth distributions).

Providing a convergence rate for the estimation error requires some more (mild) assumptions. Let $\mathcal{F}^s_{b,r}$ be the subset of $\mathcal{F}_b$ with the additional condition that $\tilde{V}(\cdot; x) \in C^s([-b, +b])$ (for any fix $x \in \mathcal{X}$). The reasoning required to provide a covering number to be used by the estimation error analysis goes as follows: *(1)* If the reward has s-finite absolute moments, its CF $\tilde{R}(\cdot; x)$ is s-times differentiable (cf. Lemma 7). *(2)* $\tilde{R}$ can be approximated by a function within $\mathcal{F}^s_{b,r}$, with an error that depends on its β-smoothness and the choice of b (almost as before). *(3)* We can prove that if $\tilde{V}_k \in \mathcal{F}^s_{b,r}$, it stays in the same smoothness class after applying the Bellman operator (with possibly a larger norm r'). *(4)* The estimation error depends on the complexity of $\mathcal{F}^s_{b,r}$. This is a smoothness class, whose covering number is well behaving, i.e., $\log \mathcal{N}(\varepsilon, \mathcal{F}^s_{b,r}) \le cb(\frac{r}{\varepsilon})^{-1/s}$.

R1, R3: Motivation? Why not represent the distribution instead?
A: The first motivation is that a new representation opens up possibilities for designing new algorithms. A good example is in the field of control theory, where we have tools to analyze a dynamical system in either the time or frequency domain. Even though they are equivalent in many cases, designing a controller in the frequency domain might be easier. This work brings the frequency-based representation of uncertainty to DistRL. The second motivation is that estimating a probability distribution of returns with a parametric model by performing MLE is infeasible in general (due to the computational challenge of computing the partition function), whereas estimating CF is not (LL39-41).

R4, R3: How to solve in practice? How Eq. (14) can be solved? How to deal with the integral?
A: Performing ACVI requires us to solve a series of regression problems. Algorithmically the only difference here is that the input includes both state x and frequency ω. Eq. (14) is only one specific (ERM-based) approach, but is not the only one. Focusing on Eq. (14): This is similar to the usual Fitted Q-Iteration. The integral can be approximated numerically, for example by discretizing over various ω. As shown in the response to *R4: Non-vacuousness ...*, we can focus on a bounded domain for ω. I expect computing it analytically might not be possible for general parametrization of CVF, but one might be able to exploit the regularities of, say, a decision tree to compute it more efficiently (constancy of values within a leaf). Also note that estimating ECF has a long history in the statistics and econometrics literature, so it is possible to borrow methods studied there too (see references mentioned in LL36-39).

R4: Other Q&As. Q: Conditional independence without action? **A:** The current derivations are correct if the policy is deterministic, as the action is uniquely determined by the state and the policy. If π is stochastic, we need to condition on action too, as you mentioned. **Q:** Distribution without density (LL109-111). **A:** CF exists even if the density does not. **Q:** Correct use of Banach fixed point theorem? **A:** When the paper talks about the convergence (LL174-182), I am careful to ensure that we are talking about bounded terms. I will clarify this. **Q:** Missing π in Parseval? **A:** You are right! Equality is for a different convention for the Fourier transforms. The change in the final result is that π in the denominator becomes $\sqrt{\pi}$. **Q:** L212: Uniform weighting leads to a finite integral? **A:** If we limit the bandwidth to b, as discussed earlier, we do not need to be worried about the unboundedness of the integral. More generally, the finiteness seems to depends on the tail behaviour of $\tilde{T}^\pi \tilde{V}_k$, which for example is satisfied with β-smoothness with large enough β.