Data-Driven Nonparametric Robust Control under Dependence Uncertainty

Erhan Bayraktar[*]   Tao Chen[†]

This Version: September 11, 2022

ABSTRACT: We consider a multi-period stochastic control problem where the multivariate driving stochastic factor of the system has known marginal distributions but uncertain dependence structure. To solve the problem, we propose to implement the nonparametric adaptive robust control framework. We aim to find the optimal control against the worst-case copulae in a sequence of shrinking uncertainty sets which are generated from continuously observing the data. Then, we use a stochastic gradient descent ascent algorithm to numerically handle the corresponding high dimensional dynamic inf-sup optimization problem. We present the numerical results in the context of utility maximization and show that the controller benefits from knowing more information about the uncertain model.

1 Introduction

In this paper we propose a nonparametric approach for solving a stochastic Markovian control problem in discrete time under a special type of uncertainty. We assume that the multivariate driving random factor of the underlying stochastic process has known marginals and unknown dependence. Naturally, to deal with such kind of uncertainty, one can choose a parametric family of copula functions and learn the unknown parameter from the observed data. In this work, we avoid making such postulation to prevent model misspecification. When it comes to handling model uncertainty, there are different approaches, parametric and nonparametric, developed in the past decades to incorporate learning into solving control problems with unknown system models (cf. [KV15], [CG91], [Rie75], [CM20]). However, earlier studies show that a pure learning approach without awareness of the model risk is prone to risk caused by estimation error and often leads to overly aggressive controls and system outcomes with high variances. On the other hand, the central idea of robust control goes back to [GS89]. A large body of research have been devoted to this area since then, and produced fruitful results. Robust techniques are extremely successful in dealing with model risk but if the learning phase is lacking in the framework, corresponding controls can be overly conservative and even trivial. Our work aims to address all the issues mentioned above when handling dependence uncertainty through applying the nonparametric adaptive robust

methodology proposed in [BC22] and developing an efficient numerical scheme for implementing such method. The main idea of adaptive robust control is to consider a specific type of sequential games: the controller first constructs a sequence of uncertainty sets by continuously observing the system data; then, at each time step the nature chooses one model from the time-t uncertainty set that is the worst for the controller; finally, the controller will apply an optimal control in response, and both players (the controller and the nature) repeat the same process at latter time steps. Note that such framework is data-driven and the learning phase is integrated into the optimization phase. Regarding constructing the desired uncertainty sets, we refer to [BCC17] and [BC21] for the parametric case in discrete time and continuous time, respectively. Discussion of constructing time dependent uncertainty sets in discrete time in the nonparametric case can be found in our earlier work [BC22]. In summary, to deal with the dependence uncertainty, we use the method described in the aforementioned papers to build uncertainty sets in terms of the Wasserstein ball centered at the nonparametric estimators of the copula by utilizing the concentration results for the empirical distribution and the Wasserstein distance (cf. [DBGM99], [FG15]), and formulate the corresponding robust control problem. Practically, numerical search of the worst-case copula in a Wasserstein ball is extremely difficult. One can utilize the results presented in [GK17] to obtain a duality result for solving the relevant inf-sup optimization problem. However, as we will see in the sequel, the Bellman equation derived in this framework has unavoidable high dimension. Thus, solving such optimization problem using brute force is virtually impossible. Towards this end, we design a stochastic gradient descent ascent algorithm to mitigate the computational burden. On top of that, we use the Gaussian process regression model to approximate the corresponding value function. The gradient of such approximation can be explicitly computed, and this fact increases the numerical efficiency in the dynamic programming backward recursion.

The rest of the paper is organized as follows. We begin Section 2 with setting up the model and then discuss the stochastic process that represents the learning of the unknown dependence, and the construction of the uncertainty sets which contain all the considered copulae at different time steps. Section 3 is dedicated to the formulation of the nonparametric adaptive robust control framework. We investigate the solution of the nonparametric adaptive robust control problem in Section 4. In this section, we prove the Bellman principle of optimality for the problem and show the existence of measurable worst-case model selector as well as the existence of measurable optimal control. We also discuss the comparison among several different methods that can be used to handle the stochastic control problem under dependence uncertainty. Finally, in Section 5 we consider an illustrative example. Namely, the uncertain utility maximization problem where the investor needs to allocate the wealth among the money market account and several risky assets without knowing the true dependence of return processes of the risky assets. We apply the nonparametric adaptive robust control approach to such problem and provide a numerical solver by using machine learning techniques such as stochastic gradient descent ascent and the Gaussian process. Numerical results presented in this section show that knowing more about the distribution of the underlying stochastic process leads to optimal trading strategy that is more balanced between profit seeking and risk aversion.

2 Nonparametric Adaptive Robust Control Problem Subject to Dependence Uncertainty

Let $(\Omega, \mathscr{F})$ be a measurable space, and $T \in \mathbb{N}$ be a fixed time horizon. Let $\mathcal{T} = \{0, 1, 2, \ldots, T\}$, $\mathcal{T}' = \{0, 1, 2, \ldots, T-1\}$, and $\mathcal{T}'' = \{1, 2, \ldots, T\}$. On the space $(\Omega, \mathscr{F})$ we consider a controlled

stochastic process $X = \{X_t,\ t \in \mathcal{T}\}$ taking values in $\mathbb{R}^d$ with dynamics

$$X_{t+1} = S(t, X_t, \varphi_t, Z_{t+1}), \quad t \in \mathcal{T}'. \tag{2.1}$$

The above $Z_t = (Z_t^{(1)}, \ldots, Z_t^{(n)})$, $t \in \mathcal{T}''$, is an n-dimensional i.i.d. sequence adapted to the natural filtration of the process X which is denoted by $\mathbb{F} = \{\mathcal{F}_t, t \in \mathcal{T}\}$. The process $\varphi = \{\varphi_t, t \in \mathcal{T}'\}$ is also $\mathbb{F}$-adapted and takes values in a compact subset $A \subset \mathbb{R}^m$. For every $t \in \mathcal{T}'$, the function $S(t, \cdot, \cdot, \cdot) : \mathbb{R}^d \times A \times \mathbb{R}^n \to \mathbb{R}^d$ is deterministic and continuous. In addition, we denote by $\mathcal{A}_t$ the set of all processes that take values in A and are adapted to the filtration $\mathbb{F}_t := \{\mathcal{F}_s, t \leq s \leq T - 1\}$. Each element in $\mathcal{A}_t$ is called an admissible control starting at time t, and we use the convention $\mathcal{A} = \mathcal{A}_0$. In this work, we assume that the process Z is observable, and for every fixed $t \in \mathcal{T}''$, the marginal distributions of Z_t are known to the controller, but the dependence structure among marginals is unknown and needs to be learned. We write $\mathcal{P}(\mathbb{R}^n)$ as the set of all distributions on $\mathbb{R}^n$. Finally, let $\ell : \mathbb{R}^d \to \mathbb{R}$ be continuous, bounded from above, and playing the role of the loss function. In this work, denote by F_* the true distribution of Z_t, $t \in \mathcal{T}''$, we will formulate and solve a robust optimization problem aiming to minimize the expected loss with the restriction that the worst-case distribution of Z_t, $t \in \mathcal{T}''$, matches the marginals of F_*. To proceed, we specify the following notations which will be used throughout.

- For any $z \in \mathbb{R}^n$, $z^{(i)}$, $i \in \mathcal{N} := \{1, \ldots, n\}$, is the ith component of z.

- For any CDF F of a n-dimensional random variable, $F^{(i)}$, $i \in \mathcal{N}$, is the ith marginal CDF.

- For any CDF F of a n-dimensional random variable, $F^{-1} = (F^{(1),-1}, \ldots, F^{(n),-1})$.

- For any CDF F of a n-dimensional random variable and $z \in \mathbb{R}^n$,

$$\begin{aligned} F \circ z =& (F^{(1)}(z^{(1)}), \ldots, F^{(n)}(z^{(n)})), \\ F^{-1} \circ z =& (F^{(1),-1}(z^{(1)}), \ldots, F^{(n),-1}(z^{(n)})). \end{aligned}$$

- For any function f on $\mathbb{R}^n$, and CDF F of a n-dimensional random variable,

$$\begin{aligned} f \circ F =& f(F^{(1)}, \ldots, F^{(n)}), \\ f \circ F^{-1} =& f(F^{(1),-1}, \ldots, F^{(n),-1}). \end{aligned}$$

In order to deal with the dependence uncertainty, we will take the copula perspective and note that

$$F_*(z) = C_* \left(F_*^{(1)} \left(z^{(1)} \right), \ldots, F_*^{(n)} \left(z^{(n)} \right) \right), \tag{2.2}$$

where C_* is the unknown true copula function of Z_t. Here, we make a standing assumption that $F_*^{(i)}$, $i \in \mathcal{N}$, is continuous and strictly increasing. Consequently, the copula function C_* is unique. In the sequel, we will write $\mathbb{P}_F$ and $\mathbb{E}_F$ as the probability measure and expectation, respectively, associated to the distribution function F.

In the first step, having in mind the possibility that the parametric family which C_* belongs to could be unknown, we need to define an estimator for C_* based on the observed data of Z in a nonparametric manner. There are essentially two different ways to achieve this. One is to use the up-to-time t information of Z to construct a multi-dimensional empirical distribution $\widehat{F}_t$ for F_*, and define an estimator of C_* via (2.2). The other is to consider a kernel based estimator of F_* which is a smoothed version of $\widehat{F}_t$ instead. One example in such family of methods is the so-called perturbed

empirical distribution. Both approaches have their advantages: using the empirical distribution will simplify the numerical computation; while the kernel based method offers a better estimator that is more in line with the assumption that $F_*^{(i)}$, $i \in \mathcal{N}$, are continuous. To this end, with slight abuse of notations, we choose the avenue of the first method mentioned earlier and denote by $\widehat{F}_t$ the empirical distribution of F_* based on information up to time t, where $\widehat{F}_0$ is constructed by observing historical data of Z with sample size t_0. It is not hard to see that $\widehat{F}_t$ satisfies the following updating rule

$$\widehat{F}_{t+1}(z) = \frac{(t_0 + t)\widehat{F}_t(z) + \prod_{i \in \mathcal{N}} \mathbb{1}_{Z_{t+1}^{(i)} \leq z^{(i)}}}{t_0 + t + 1}, \quad z \in [0,1]^n. \tag{2.3}$$

Next, we define

$$\widehat{C}_t(u) = \widehat{F}_t(F_*^{-1} \circ u), \quad u \in [0,1]^n, \tag{2.4}$$

and $\widehat{C}_t$ is an estimator of the copula C_*. Moreover, we have

$$\widehat{C}_{t+1}(u) = \frac{(t_0 + t)\widehat{C}_t(u) + \prod_{i \in \mathcal{N}} \mathbb{1}_{[0,u_i]}\left(F_*^{(i)}\left(Z_{t+1}^{(i)}\right)\right)}{t_0 + t + 1} =: R(t, \widehat{C}_t, Z_{t+1}), \quad u \in [0,1]^n. \tag{2.5}$$

Remark 2.1. We want to stress that even though F_* is uncertain, but we assume that the marginals $F_*^{(i)}$, $i \in \mathcal{N}$, are known, and therefore due to our definition F_*^{-1} is also well-known to the controller.

In many works, regardless of knowing $F_*^{(i)}$, $i \in \mathcal{N}$, or not, the empirical copula is defined as $\widehat{C}_t(u) = \widehat{F}_t(\widehat{F}_t^{-1} \circ u)$ which means using the empirical marginals instead of the true ones to construct the estimator for the copula. Note that in such construction, due to the discontinuity of $\widehat{F}_t$, the resulting $\widehat{C}_t$ is not necessarily a copula. On the other hand, the estimator of C_* defined in (2.4) is not a copula either. Hence, any application or computation purely based on either of these two estimators could be problematic. We argue that, for this reason, one should consider a robust framework to mitigate the inherent risk comes from the statistical estimation procedure.

In rest of the paper, we write

$$\widehat{F}_{t+1}(z) = \frac{(t_0 + t)\widehat{F}_t(z) + H_{t+1}(z - Z_{t+1})}{t_0 + t + 1}, \quad z \in \mathbb{R}^n,$$

where H_t is a sequence of continuous distribution function such that $\lim_{t \to \infty} H_t = \mathbb{1}_{[0,\infty)^n}$ if t is allowed to go to infinity, or $H_t \equiv \mathbb{1}_{[0,\infty)^n}$ for any considered t. The former choice corresponds to constructing $\widehat{F}$ as the so-called perturbed empirical distribution and the latter is the case of choosing $\widehat{F}$ to be the regular empirical distribution. We will prove all the technical results for both choices as our theory works for either one of the two possible frameworks.

To proceed, we fix the notations $\widehat{F}$ and $\widehat{C}$ for estimators defined as in (2.3) and (2.4), and assume that, for any $0 \leq \alpha \leq 1$, there exists a deterministic function $r(\alpha, t_0, t)$ decreasing in $t \in \mathcal{T}'$ such that

$$\mathbb{P}(d_{W,p}(\widehat{C}_t, C_*) \leq r(\alpha, t_0, t)) \geq 1 - \alpha, \tag{2.6}$$

where $d_{W,p}$ is the Wasserstein distance between distributions in $\mathcal{P}([0,1]^n)$ of order $p \geq 1$. In application, regarding construction of $r(\alpha, t_0, t)$, we will use the concentration results for the empirical distribution in [FG15, Theorem 2] by imposing necessary integrability assumptions on C_*. The take-away is that there exists some function $r(\alpha, t_0, t)$ satisfying (2.6).

Next, we model the uncertainty of C_*, by using (2.6), in terms of the confidence region as

$$\mathcal{C}_t^\alpha(\widehat{C}_t) = \left\{ C \in \mathcal{P}([0,1]^n) : d_{W,p}(\widehat{C}_t, C) \le r(\alpha, t_0, t) \right\}, \tag{2.7}$$

and we call $\mathcal{C}_t^\alpha$ the uncertainty set of C_*. Since $[0,1]^n$ is a compact subset of $\mathbb{R}^n$, then $d_{W,p}$ is a metric on $\mathcal{P}([0,1]^n)$. Throughout, we consider the metric space $(\mathcal{P}([0,1]^n), d_{W,p})$. As mentioned earlier, we will adopt the adaptive robust control framework to our setup in this project. Hence, we are going to find the worst-case copula function in $\mathcal{C}_t^\alpha$ and search for the optimal control in reaction to such worst-case copula. One needs to realize that $\mathcal{C}_t^\alpha$ includes all types of distributions on $[0,1]^n$ and in general these distributions do not necessarily have uniform marginals. In other words, not all $C \in \mathcal{C}_t^\alpha$ are copulae, and this issue will be addressed in the sequel.

Before we discuss and formulate the nonparametric adaptive robust control problem under dependence uncertainty, we first provide the following technical results for preparation.

Lemma 2.2. *For fixed $t \in \mathcal{T}'$, the mapping $R(t, \cdot, \cdot) : \mathcal{P}([0,1]^n) \times \mathbb{R} \to \mathcal{P}([0,1]^n)$ is continuous.*

Proof. We will prove the result for H_t, $t \in \mathcal{T}'$, being continuous distribution functions. Regarding the case of $H_t \equiv \mathbb{1}_{[0,\infty)}$, we refer the readers to [CM20] or Lemma 3.1 below.

Assume that $(C_i, z_i) \to (C, z)$, where $C_i, C \in \mathcal{P}([0,1]^n)$, $z_i, z \in \mathbb{R}$, $i = 1, 2, \ldots$. Then, $d_{W,p}(C_i, C) \to 0$ and $z_i \to z$. Denote $\mu_{C_i,z_i} = R(t, C_i, z_i)$ and $\mu_{C,z} = R(t, C, z)$, and $\mathcal{M} := \{f : |f(z) - f(z')| \le \|z - z'\|, \ z, z' \in [0,1]^n\}$. Note that $[0,1]^n$ is compact, so for any $p \ge 1$ there exists a constant b_p such that $d_{W,p}(F, G) \le b_p d_{W,1}(F, G)$ for any $F, G \in \mathcal{P}([0,1]^n)$. Hence, we have that

$$d_{W,p}(\mu_{C_i,z_i}, \mu_{C,z}) \le b_p d_{W,1}(\mu_{C_i,z_i}, \mu_{C,z}) = b_p \sup \left\{ \int_{[0,1]^n} f d\mu_{C_i,z_i} - \int_{[0,1]^n} f d\mu_{C,z} : f \in \mathcal{M}. \right\}$$

By the construction of the function R, we obtain

$$\sup \left\{ \int_{[0,1]^n} f d(\mu_{C_i,z_i} - \mu_{C,z}) : f \in \mathcal{M}. \right\} = \sup \left\{ \frac{t_0 + t}{t_0 + t + 1} \int_{[0,1]^n} f d(C_i - C) \right. $$
$$\left. + \frac{\int_{[0,1]^n} f d(H_{t+1}(F_*^{-1} \circ u - z_i) - H_{t+1}(F_*^{-1} \circ u - z))}{t + t_0 + 1} : \ f \in \mathcal{M} \right\}.$$

Note that

$$H_{t+1}(F_*^{-1} \circ u - z_i) = \mathbb{P}(\widetilde{H} \le F_*^{-1} \circ u - z_i)$$
$$= \mathbb{P}\left(\left(F_*^{(1)}(\widetilde{H}^{(1)} + z_i^{(1)}), \ldots, F_*^{(n)}(\widetilde{H}^{(n)} + z_i^{(n)}) \right) \le u \right),$$

where $\widetilde{H}$ is a multi-dimensional random variable with distribution H_{t+1}. Then, we get that

$$\sup \left\{ \int_{[0,1]^n} f d(H_{t+1}(F_*^{-1} \circ u - z_i) - H_{t+1}(F_*^{-1} \circ u - z)) : \ f \in \mathcal{M} \right\}$$
$$= \sup \left\{ \mathbb{E}\left[f\left(F_*^{(1)}(\widetilde{H}^{(1)} + z_i^{(1)}), \ldots, F_*^{(n)}(\widetilde{H}^{(n)} + z_i^{(n)}) \right) \right] \right.$$
$$\left. - \mathbb{E}\left[f\left(F_*^{(1)}(\widetilde{H}^{(1)} + z^{(1)}), \ldots, F_*^{(n)}(\widetilde{H}^{(n)} + z^{(n)}) \right) \right] : f \in \mathcal{M} \right\}$$
$$\le \mathbb{E}\left[\left\| \left(F_*^{(1)}(\widetilde{H}^{(1)} + z_i^{(1)}), \ldots, F_*^{(n)}(\widetilde{H}^{(n)} + z_i^{(n)}) \right) - \left(F_*^{(1)}(\widetilde{H}^{(1)} + z^{(1)}), \ldots, F_*^{(n)}(\widetilde{H}^{(n)} + z^{(n)}) \right) \right\| \right].$$

Since F_*^j, $j \in \mathcal{N}$, are all bounded continuous functions, thus by dominated convergence theorem,

$$\sup \left\{ \frac{\int_{[0,1]^n} f d(H_{t+1}(F_*^{-1} \circ u - z_i) - H_{t+1}(F_*^{-1} \circ u - z))}{t + t_0 + 1} : f \in \mathcal{M} \right\} \to 0$$

as $z_i \to z$. On the other hand,

$$\sup \left\{ \frac{t_0 + t}{t_0 + t + 1} \int_{[0,1]^n} f d(C_i - C) : f \in \mathcal{M} \right\} = \frac{t_0 + t}{t_0 + t + 1} d_{W,1}(C_i, C).$$

In view that $d_{W,1}(C_i, C) \leq d_{W,p}(C_i, C)$ for $p \geq 1$, we have that $d_{W,p}(C_i, C) \to 0$ implies that the above term converges to 0 as well. Therefore, $d_{W,p}(\mu_{C_i, z_i}, \mu_{C,z}) \to 0$ and the mapping $R(t, \cdot, \cdot)$ is continuous. $\qquad \square$

Another important property that $\mathcal{C}_t^\alpha$ satisfies as a set valued function is upper hemi-continuity (u.h.c.) and it is stated below.

Lemma 2.3. *For every $t \in \mathcal{T}'$, the set valued function $\mathcal{C}_t^\alpha$ is upper hemi-continuous.*

Proof. We need to show that for any $C \in \mathcal{P}([0,1]^n)$ and any open set E such that $\mathcal{C}_t^\alpha(C) \subset E \subset \mathcal{P}([0,1]^n)$, there exists a neighborhood D of C such that for all $C' \in D$, $\mathcal{C}_t^\alpha(C') \subset E$.

To this end, let $\varepsilon = \text{dist}(C, \partial E) - r(\alpha, t_0, t)$ where $\text{dist}(C, \partial E)$ is the shortest Wasserstein-p distance from C to the boundary of E. Since E is an open set, then $\varepsilon > 0$. Hence, we take D as the Wasserstein-p ball centered at C with radius ε. It is clear that for any $C' \in D$, we have $\mathcal{C}_t^\alpha(C') \subset E$. Therefore, for fixed $t \in \mathcal{T}'$, $\mathcal{C}_t^\alpha$ is u.h.c.. $\qquad \square$

3　Nonparametric Adaptive Robust Control Problem

Now we proceed to formulate the nonparametric adaptive robust control problem assuming the true copula function C_* is unknown. Recall that we consider $\mathcal{P}([0,1]^n)$ with the metric $d_{W,p}$. Obviously, $(\mathcal{P}([0,1]^n, d_{W,p})$ is separable and complete. Hence, it is a Polish space and thus a Borel space. We define the augmented state process $Y = \{Y_t = (X_t, \widehat{C}_t), t \in \mathcal{T}\}$, and such process has the following dynamics

$$Y_{t+1} = \mathbf{G}(t, X_t, \varphi_t, Z_{t+1}) := (S(t, X_t, \varphi_t, Z_{t+1}), R(t, \widehat{C}_t, Z_{t+1})), \quad t \in \mathcal{T}'. \tag{3.1}$$

According to the assumption that S is continuous and Lemma 2.2, we get that $\mathbf{G}$ is continuous and therefore Borel measurable.

3.1　Transition Kernel of the State Process

Through the rest of this paper, we will write $\mathcal{C}_t^\alpha(Y_t) := \mathcal{C}_t^\alpha(\widehat{C}_t)$ for $Y_t = (X_t, \widehat{C}_t)$. For $C \in \mathcal{P}([0,1]^n)$, one would like to define the transition kernel of Y as

$$Q_t(dy'|y, a, F) = \mathbb{P}_{C \circ F^*}(\mathbf{G}(t, y, a, Z_{t+1}) \in dy').$$

However, $C \circ F^*$ is a well-defined distribution on $\mathbb{R}^n$ only when $C^{(i)}$, $i \in \mathcal{N}$, is left continuous at 1 and

$$\lim_{u \to 0} C^{(i)}(u) = 0.$$

Distribution $C \in \mathcal{P}([0,1]^n)$ does not satisfies the above conditions will result in a distribution $C \circ F^*$ putting mass on infinity. To overcome this problem, we consider $\overline{\overline{\mathbb{R}}} = \mathbb{R} \bigcup \{-\infty, \infty\}$ with Borel σ-algebra $\mathcal{B}(\overline{\mathbb{R}}) = \sigma (\mathcal{B}(\mathbb{R}) \bigcup \{[-\infty, x), (x, \infty], x \in \mathbb{R}\})$. We also consider a metric on $\overline{\mathbb{R}}$ consistent with $\mathcal{B}(\overline{\mathbb{R}})$. Consequently, we have $\overline{\mathbb{R}}^n$ equipped with the product topology and it is a Polish space as well as a Borel space. In addition, we consider $\mathcal{P}(\overline{\mathbb{R}}^n)$ the set of all probability distributions on $\overline{\mathbb{R}}^n$ with topology consistent with weak convergence. For any $F \in \mathcal{P}(\mathbb{R}^n)$, we denote by $\overline{F}$ its extension in $\mathcal{P}(\overline{\mathbb{R}}^n)$ as

$$\overline{F}(x) = \begin{cases} F(x), & x \in \mathbb{R}^n, \\ 0, & -\infty \in \{x^{(i)}, i \in \mathcal{N}\}, \\ 1, & \infty \in \{x^{(i)}, i \in \mathcal{N}\}. \end{cases}$$

Naturally, we change the state space of X to $\overline{\mathbb{R}}^n$, and define the augmented state space $E_Y = \overline{\mathbb{R}}^n \times \mathcal{P}([0,1]^n)$ with product topology. It is then a Borel space and the Borel σ-algebra $\mathcal{E}_Y$ coincides with the product σ-algebra. In addition, for any relevant real valued function f, we write

$$f(\infty) = \lim_{x \to \infty} f(x).$$

To proceed, for any $t \in \mathcal{T}'$, $(y, a) \in E_Y \times A$, and distribution $C \in \mathcal{P}([0,1]^n)$, Q_t is a probability measure on $\mathcal{E}_Y$ such that

$$Q_t(D|y, a, C) = \mathbb{P}_{C \circ F_*}(\mathbf{G}(t, y, a, Z_{t+1}) \in D), \quad D \in \mathcal{E}_Y.$$

In particular, when C is a copula, $C \circ F_*$ is a n-dimensional probability distribution on $\mathbb{R}^n$ with marginals $F_*^{(1)}, \ldots, F_*^{(n)}$. One important property of Q_t is that it is Borel measurable for fixed $t \in \mathcal{T}'$. Such property is crucial for proving the existence of measurable optimal control and worst-case copula. To this end, we extend Lemma 2.2 to the general case and provide the following technical results.

Lemma 3.1. *For every fixed $t \in \mathcal{T}'$, the mapping*

$$\overline{R}(t, F, z') = \frac{(t_0 + t)F(z) + \overline{H}_{t+1}(z - z')}{t_0 + t + 1}, \quad F \in \mathcal{P}(\overline{\mathbb{R}}^n), \; z, z' \in \overline{\mathbb{R}}^n. \tag{3.2}$$

is continuous on $\mathcal{P}(\overline{\mathbb{R}}^n) \times \overline{\mathbb{R}}^n$.

Proof. Assume that $(F_i, z_i') \to (F, z_0')$ and it implies that $F_i \to F$ in distribution and $z_i' \to z_0'$. Take any continuity set B of $\overline{R}(t, F_0, z_0')$. That is for the boundary ∂B of B, the probability $\mathbb{P}_{\overline{R}(t, F_0, z_0')}(\partial B)$ is equal to 0. Therefore, we get that B is also a continuity set of F and $\overline{H}_{t+1}(z - z_0')$.

Now, since $F_i \to F_0$ in distribution, then by the Portmanteau lemma, we know

$$\lim_{i \to \infty} \mathbb{P}_{F_i}(B) \to \mathbb{P}_{F_0}(B).$$

On the other hand, for any $z' \in \overline{\mathbb{R}}^n$, $\overline{H}_{t+1}(z - z') = \mathbb{P}(\widetilde{H} \le z - z') = \mathbb{P}(\widetilde{H} + z' \le z)$, where $\widetilde{H}$ is the random variable with distribution $\overline{H}_{t+1}$. Since z_i' converges to the constant z_0', then $\widetilde{H} + z_i' \to \widetilde{H} + z_0'$ in distribution. In turn, we get

$$\lim_{i \to \infty} \mathbb{P}_{\overline{H}_{t+1}(z - z_i')}(B) \to \mathbb{P}_{\overline{H}_{t+1}(z - z_0')}(B).$$

To summarize, we have that

$$\lim_{i \to \infty} \mathbb{P}_{\overline{R}(t, F_i, z_i')}(B) = \mathbb{P}_{\overline{R}(t, F_0, z_0')}(B),$$

and $\overline{R}(t, F, z')$ is continuous on $\mathcal{P}(\overline{R}^n) \times \overline{\mathbb{R}}^n$. $\square$

With this result in hand, we define the transition kernel on the product space $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n)$ with product topology. That is, fix $t \in \mathcal{T}'$, for any $(x, F, a, G) \in \overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n) \times A \times \mathcal{P}(\overline{\mathbb{R}}^n)$, and any set D belongs to the Borel σ-algebra of $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n)$, we write

$$\overline{Q}_t(D|x, F, a, G) = \mathbb{P}_G((S(t, x, a, Z_{t+1}), \overline{R}(t, F, Z_{t+1})) \in D).$$

Then, we show the measurability of $\overline{Q}_t$, $t \in \mathcal{T}'$.

Lemma 3.2. *For every fixed $t \in \mathcal{T}'$, $\overline{Q}_t$ is a continuous stochastic kernel on $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n)$ given $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n) \times A \times \mathcal{P}(\overline{\mathbb{R}}^n)$.*

Proof. It is enough to prove that for any bounded continuous function f on $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n)$ we have

$$\int_{\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n)} f(y) \overline{Q}_t(dy|x, F, a, G) \tag{3.3}$$

is continuous w.r.t. (x, F, a, G). To this end, we view that

$$\int_{\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n)} f(y) \overline{Q}_t(dy|x, F, a, G) = \int_{\overline{\mathbb{R}}^n} f((S(t, x, a, z), \overline{R}(t, F, z)) dG(z).$$

By assumption on S and Lemma 3.1, for fixed $t \in \mathcal{T}'$, S and $\overline{R}$ are continuous w.r.t. other variables. Since f is also continuous, then $f((S(t, x, a, z), \overline{R}(t, F, z))$ is continuous w.r.t. (x, F, a, z). In addition, $G(z)$ can be seen as a continuous stochastic kernel on $\overline{\mathbb{R}}^n$ given (x, F, a, G) because it does not depend on (x, F, a). Hence, by [BS78, Proposition 7.30], we get that

$$\int_{\overline{\mathbb{R}}^n} f((S(t, x, a, z), \overline{R}(t, F, z)) dG(z)$$

is continuous w.r.t. (x, F, a, G). We immediately conclude that (3.3) is continuous w.r.t. (x, F, a, G). $\square$

Next, we define the following mappings between $\mathcal{P}([0, 1]^n)$ and $\mathcal{P}(\overline{\mathbb{R}}^n)$ similarly to (2.4): for any $F \in \mathcal{P}(\overline{\mathbb{R}}^n)$,

$$C(u) := F(F_*^{-1} \circ u), \quad u \in [0, 1]^n, \tag{3.4}$$

and for any $C \in \mathcal{P}([0, 1]^n)$,

$$F(x) := C(F_* \circ x), \quad x \in \overline{\mathbb{R}}^n. \tag{3.5}$$

We realize that due to our assumptions on the marginals $F_*^{(i)}$, $i \in \mathcal{N}$, the mapping

$$u \to F_* \circ u, \quad u \in [0, 1]^n,$$

defines a homeomorphism between $[0, 1]^n$ and $\overline{\mathbb{R}}^n$. By using such observation, we show the main result of this section as follows.

Proposition 3.3. *For each $t \in \mathcal{T}'$, the probability $Q_t(\,\cdot\,|y, a, C)$ is a continuous stochastic kernel on E_Y given $E_Y \times A \times \mathcal{P}([0, 1]^n)$. Moreover, it is a Borel measurable stochastic kernel on E_Y given $E_Y \times A \times \mathcal{P}([0, 1]^n)$.*

Proof. Through the proof, all the convergence for probability distributions are understood in the weak sense. We will prove the statements in several steps.

8

1. Establish a mapping M_1 from $E_Y \times A \times \mathcal{P}([0,1]^n)$ to $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n) \times A \times \mathcal{P}(\overline{\mathbb{R}}^n)$ that is continuous.

2. Use Lemma 3.2 to obtain the continuous mapping $\overline{Q}_t$ from $\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n) \times A \times \mathcal{P}(\overline{\mathbb{R}}^n)$ to $\mathcal{P}(\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n))$.

3. Construct a mapping M_2 from $\mathcal{P}(\overline{\mathbb{R}}^n \times \mathcal{P}(\overline{\mathbb{R}}^n))$ to $\mathcal{P}(E_Y)$, which is restricted to the family of $\overline{Q}_t(\,\cdot\,|x, F, a, G)$, and satisfies some proper form of continuity.

We also need M_1 and M_2 to satisfy $Q_t = M_2 \circ \overline{Q}_t \circ M_1$. To this end, for the first step, we take $M_1 : (x, \widetilde{C}, a, C) \mapsto (x, \widetilde{C} \circ F_*, a, C \circ F_*)$, $F, G \in \mathcal{P}([0,1]^n)$. Assume that the sequences $\{\widetilde{C}_i \in \mathcal{P}([0,1]^n), i = 1, 2, \ldots\}$ and $\{C_i \in \mathcal{P}([0,1]^n), i = 1, 2, \ldots\}$ converge to $\widetilde{C}_0 \in \mathcal{P}([0,1]^n)$ and $C_0 \in \mathcal{P}([0,1]^n)$ in distribution, respectively. It implies that for any continuity point u of $\widetilde{C}_0$, we have $\widetilde{C}_i(u) \to \widetilde{C}_0(u)$ as $i \to \infty$. In view of the homeomorphism

$$u \to F_*^{-1} \circ u, \quad u \in [0,1]^n,$$

if z is a continuity point of $\widetilde{C}_0 \circ F_*$, then $\left(F_*^{(1)}(z^{(1)}), \ldots, F_*^{(n)}(z^{(n)})\right)$ is a continuity point of $\widetilde{C}_0$. Recall that $\widetilde{C}_i$ converges to $\widetilde{C}_0$ in distribution, we get

$$\lim_{i \to \infty} \widetilde{C}_i \circ F_*(z) = \lim_{i \to \infty} \widetilde{C}_i(F_* \circ z) = \lim_{i \to \infty} \widetilde{C}_i \left(F_*^{(1)}(z^{(1)}), \ldots, F_*^{(n)}(z^{(n)})\right)$$
$$= \widetilde{C}_0 \left(F_*^{(1)}(z^{(1)}), \ldots, F_*^{(n)}(z^{(n)})\right) = \widetilde{C}_0 \circ F_*(z),$$

which implies that $\widetilde{C}_i \circ F_*$ converges to $\widetilde{C}_0 \circ F_*$ in distribution. Similarly, $C_i \circ F_*$ converges to $C_0 \circ F_*$ in distribution, and step 1 is done. Also note that step 2 is fulfilled according to Lemma 3.2.

Next, define $M_2 : \overline{Q}_t(\,\cdot\,|x, F, a, G) \mapsto Q_t(\,\cdot\,|x, F \circ F_*^{-1}, a, G \circ F_*^{-1})$, $F, G \in \mathcal{P}(\overline{\mathbb{R}}^n)$. We will show that Q_t is continuous w.r.t. (x, F, a, G). In turn, the mapping M_2 will be continuous according to the definition of continuous parametric mapping. Let $D \in \mathcal{E}_Y$ be a closed set, and let (F_i, G_i) converges to (F_0, G_0). Similarly to the discussion in step 1, we have $F_i \circ F_*^{-1}$ and $G_i \circ F_*^{-1}$ converge to $F_0 \circ F_*^{-1}$ and $G_0 \circ F_*^{-1}$, respectively. Thus, by Lemma 2.2 and the fact that continuity in distribution and continuity in the Wasserstein sense are equivalent on $\mathcal{P}([0,1]^n)$, the function $\mathbf{G}(t, x, F \circ F_*^{-1}, a, z)$ is continuous w.r.t. (x, F, a, z). Consider any bounded continuous function f on E_Y, we get that

$$\int_{E_Y} f(y) Q_t(dy|x, F \circ F_*^{-1}, a, G \circ F_*^{-1}) = \int_{\mathbb{R}^n} f(\mathbf{G}(t, x, F \circ F_*^{-1}, a, z)) dG(F_*^{-1} \circ z). \qquad (3.6)$$

The integrand $f(\mathbf{G}(t, x, F \circ F_*^{-1}, a, z))$ is continuous w.r.t. (x, F, a, z). Also, we view the $G \circ F_*^{-1}$ as a continuous stochastic kernel w.r.t. (x, F, a, G). By [BS78, Proposition 7.30] again, the integral

$$\int_{\overline{\mathbb{R}}^n} f(\mathbf{G}(t, x, F \circ F_*^{-1}, a, z)) dG(F_*^{-1} \circ z)$$

is continuous in (x, F, a, G), and so is

$$\int_{E_Y} f(y) Q_t(dy|x, F \circ F_*^{-1}, a, G \circ F_*^{-1})$$

due to (3.6). As a result, Q_t is continuous in (x, F, a, G), and step 3 is complete.

Finally, we note that $Q_t = M_2 \circ \overline{Q}_t \circ M_1$. According to the above discussion and equivalence between continuity in distribution and in the Wasserstein sense on $\mathcal{P}([0,1]^n)$, Q_t is a continuous stochastic kernel. Hence, it is also a Borel measurable stochastic kernel. $\qquad \square$

3.2 Formulation of Adaptive Robust Control Problem

In this work, we will formulate and solve a closed loop feedback control problem. To this end, with slight abuse of notations, we say that a control process φ is Markovian if for every $t \in \mathcal{T}'$

$$\varphi_t = \varphi_t(Y(t)),$$

such that $\varphi_t(\cdot) : E_Y \to A$ is a measurable mapping. Similarly, a process ψ is called a Markovian model selector if

$$\psi_t = \psi_t(Y(t)),$$

where $\psi_t(\cdot) : E_Y \to \mathcal{P}([0,1]^n)$ is measurable. In the adaptive robust framework, we consider the Markovian control processes and Markovian model selectors such that $\psi_t(y) \in \mathcal{C}_t^\alpha(y)$ and $\psi_t(y)$ is restricted to be a copula for any $y \in E_Y$. For every $t \in \mathcal{T}'$, any time-t state $y_t \in E_Y$, and control process $\varphi \in \mathcal{A}_t$, we define

$$\mathbf{\Psi}_t^{y_t} = \left\{ \psi : \psi_s(y_s) \in \mathcal{C}_s^\alpha(y_s), \ t \le s \le T - 1 \right\}.$$

In addition, let $\varphi \in \mathcal{A}_t$, and δ_y be the Dirac probability measure assigns the mass at $y \in E_Y$. Then, at time t, for the current state $Y_t = y$, define the probability measure $\mathbb{Q}_{y,t}^{\varphi,\psi}$ on the concatenated canonical space $\mathsf{X}_{s=t}^T E_Y$ as

$$\mathbb{Q}_{y,t}^{\varphi,\psi}(B_t \times \cdots \times B_T) = \int_{B_t} \int_{B_{t+1}} \cdots \int_{B_T} \prod_{s=t}^{T-1} Q_s(dy_{s+1}|y_s, \varphi(s,y_s), \psi(s,y_s)) \delta_y(dy_t).$$

Accordingly, we consider the family of probability measures $\mathcal{Q}_{y,t}^\varphi = \left\{ \mathbb{Q}_{y,t}^{\varphi,\psi}, \ \psi \in \mathbf{\Psi}_t^y \right\}$, and we write, for simplicity, $\mathcal{Q}_y^\varphi = \mathcal{Q}_{y,0}^\varphi$. Finally, for $Y_0 = y \in E_Y$, the nonparametric adaptive robust control problem in this work is formulated as

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q}^\psi \in \mathcal{Q}_y^\varphi} \mathbb{E}_{\mathbb{Q}^\psi}[\ell(X_T)], \tag{3.7}$$

$$\text{s.t.} \quad \psi_t \text{ is a copula}, \ t \in \mathcal{T}'.$$

Note that for every $t \in \mathcal{T}'$, any $y \in E_Y$, search for the worst-case copula in $\mathcal{C}_t^\alpha(y)$ is practically impossible. To overcome such difficulty, we adopt the method of penalty functions and use it to reformulate (3.7). Ultimately, for every fixed $t \in \mathcal{T}'$ and any fixed $(y,a) \in E_Y \times A$, we want to find the optimizer $\psi_t(y,a)$ such that $\psi_t(y,a) \circ F_*$ has marginals $F_*^{(i)}$, $i \in \mathcal{N}$. Hence, we denote by $C_b(\mathbb{R})$ the set of all continuous, bounded, real valued functions, and consider the penalty function

$$\rho(\psi) = \sum_{t \in \mathcal{T}'} \rho_t(\psi_t) := \sum_{t \in \mathcal{T}'} \sum_{i \in \mathcal{N}} \sup_{f_i \in C_b(\mathbb{R})} \left(\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f_i] - \mathbb{E}_{F_*^{(i)}}[f_i] \right)$$

The choice of the penalty function is inspired by the criterion that characterize the weak convergence of probability distributions, and it is similar to the penalty function used in [GK17]. Note that for such penalty, we have

Lemma 3.4. *For any fixed $t \in \mathcal{T}'$, ψ_t is a copula if and only if $\rho_t(\psi_t) = 0$.*

Proof. If ψ_t is a copula, then it is obvious that $\rho_t(\psi_t) = 0$. On the other hand, note that for any $i \in \mathcal{N}$,

$$\sup_{f \in C_b(\overline{\mathbb{R}})} \left(\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f] - \mathbb{E}_{F_*^{(i)}}[f] \right) \geq \mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[0] - \mathbb{E}_{F_*^{(i)}}[0] = 0.$$

Hence, if $\rho_t(\psi_t) = 0$, we obtain that for any $i \in \mathcal{N}$

$$\sup_{f \in \mathcal{C}_b(\overline{\mathbb{R}})} \left(\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f] - \mathbb{E}_{F_*^{(i)}}[f] \right) = 0.$$

Moreover, we get that for any $i \in \mathcal{N}$, and any $f \in C_b(\overline{\mathbb{R}})$

$$\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f] - \mathbb{E}_{F_*^{(i)}}[f] = 0. \tag{3.8}$$

Otherwise, there exists some $i \in \mathcal{N}$ and $f_0 \in C_b(\overline{\mathbb{R}})$ such that

$$\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f_0] - \mathbb{E}_{F_*^{(i)}}[f_0] < 0.$$

In turn, for such i, we have

$$\sup_{f \in C_b(\overline{\mathbb{R}})} \left(\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f] - \mathbb{E}_{F_*^{(i)}}[f] \right) \geq \mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[-f_0] - \mathbb{E}_{F_*^{(i)}}[-f_0]$$

$$= - \left(\mathbb{E}_{(\psi_t \circ F_*)^{(i)}}[f_0] - \mathbb{E}_{F_*^{(i)}}[f_0] \right) > 0,$$

and $\rho_t(\psi_t) > 0$ as a result which is a contradiction. Therefore, equality 3.8 holds true. Consequently, by the criterion of weak convergence, we get that ψ_t is a copula. $\square$

The above result justifies the choice of ρ. Another advantage of using the proposed penalty function is that for any fixed $f \in C_b(\overline{\mathbb{R}})$, the linearity in $\psi_t \circ F_*$ allows us to obtain a tractable numerical computation scheme (cf. Section 5.1) for the value functions to be defined later. To proceed, we provide the following technical result on the regularity property of ρ_t, $t \in \mathcal{T}'$, which plays an important role in proving the existence of optimal control.

Lemma 3.5. *For every $t \in \mathcal{T}'$, the penalty term ρ_t is lower semi-continuous (l.s.c.).*

Proof. We will show that $\liminf_{C \to C_0} \rho_t(C) \geq \rho_t(C_0)$ for any $C_0 \in \mathcal{P}([0,1]^n)$. To this end, let C_j be any sequence that converges to C_0. If $\rho_t(C_0) = +\infty$, then, for any $b \in \mathbb{R}$, there exists a family $\{f_i, i \in \mathcal{N}\} \subset C_b(\overline{\mathbb{R}})$ such that

$$\sum_{i \in \mathcal{N}} \left(\mathbb{E}_{(C_0 \circ F_*)^{(i)}}[f_i] - \mathbb{E}_{F_*^{(i)}}[f_i] \right) > b.$$

Since $C_j \to C_0$, then there exists a $n_0 > 0$ such that for any $j \geq n_0$ we get

$$\sum_{i \in \mathcal{N}} \left(\mathbb{E}_{(C_j \circ F_*)^{(i)}}[f_i] - \mathbb{E}_{F_*^{(i)}}[f_i] \right) > \frac{b}{2},$$

where continuity of $\mathbb{E}_{(C \circ F_*)^{(i)}}[f_i]$ in C is obtained according to the proof of Proposition 3.3. Thus, we have that

$$\rho_t(C_j) > \frac{b}{2}, \quad j \geq n_0,$$

11

and moreover

$$\lim_{j \to \infty} \rho_t(C_j) = +\infty.$$

If $\rho_t(C_0) < +\infty$, then for any $\varepsilon > 0$, there exists a family $\{f_i, i \in \mathcal{N}\} \subset C_b(\overline{\mathbb{R}})$ such that

$$\sum_{i \in \mathcal{N}} \left(\mathbb{E}_{(C_0 \circ F_*)^{(i)}}[f_i] - \mathbb{E}_{F_*^{(i)}}[f_i] \right) > \rho_t(C_0) - \varepsilon.$$

Hence, from the convergence of C_j to C_0, there exists $n_0 > 0$ and for any $j \geq n_0$ the following inequality holds true

$$\sum_{i \in \mathcal{N}} \left(\mathbb{E}_{(C_j \circ F_*)^{(i)}}[f_i] - \mathbb{E}_{F_*^{(i)}}[f_i] \right) > \rho_t(C_0) - 2\varepsilon.$$

Thus, $\rho_t(C_j) \geq \rho_t(C_0) - 2\varepsilon$ for any $j \geq n_0$, and $\lim_{j \to \infty} \rho_t(C_j) \geq \rho_t(C_0) - 2\varepsilon$. Because ε is arbitrary, we get that $\lim_{j \to \infty} \rho_t(C_j) \geq \rho_t(C_0)$. In summary, we conclude that $\liminf_{C \to C_0} \rho_t(C) \geq \rho_t(C_0)$ and ρ_t is l.s.c.. $\qquad \square$

In the sequel, we will consider a reformulation of (3.7) as follows

$$\inf_{\varphi \in \mathcal{A}} \sup_{\mathbb{Q}^\psi \in \mathcal{Q}_y^\varphi} \left(\mathbb{E}_{\mathbb{Q}^\psi}[\ell(X_T)] - \rho(\psi) \right). \tag{3.9}$$

Before discussing the solution of (3.9), we want to remark that, in this work, the uncertainty set is some set of measures on $\mathcal{P}([0,1]^n)$ centered at the empirical copula. Another approach is to consider the worst-case model of Z_t, $t \in \mathcal{T}''$, chosen from some uncertainty set centered around the empirical distribution of Z_t. By imposing a penalty function similar to ρ, we can ensure that the worst-case model matches F_* in terms of the marginals. Thus, we are able to set up a problem that is similar to (3.9) without considering the copula. Regarding the comparison between such approach and our method, we refer to Section 4.3 below. In case that the marginals are also uncertain, one can extend the framework in [BC22] to the multi-dimensional case via the multivariate empirical distribution.

4 Solution of the Nonparametric Adaptive Robust Control Problem

In this section, we show that the solution of the nonparametric adaptive robust control problem (3.9) is given by solving the following adaptive robust Bellman equations

$$V_T(y) = \ell(x), \quad y \in E_Y,$$
$$V_t(y) = \inf_{a \in A} \sup_{C \in \mathcal{C}_t^\alpha(y)} \left(\mathbb{E}_{C \circ F_*}[V_{t+1}(\mathbf{G}(t, y, a, Z_{t+1}))] - \rho_t(C) \right), \quad y \in E_Y, \ t \in \mathcal{T}'. \tag{4.1}$$

4.1 Existence of Measurable Optimal Control

To show existence of measurable optimal control to the problem (4.1), we denote

$$v_t(y, a, C) = \mathbb{E}_{C \circ F_*}[V_{t+1}(\mathbf{G}(t, y, a, Z_{t+1}))] - \rho_t(C), \quad t \in \mathcal{T}',$$

and provide the following preliminary results.

Proposition 4.1. *For every $t \in \mathcal{T}$, the functions v_t and V_t are upper semi-continuous (u.s.c.). Moreover, there exists a Borel measurable function ψ_t^* takes values in the set of copulae such that for any fixed (y, a)*

$$\sup_{C \in \mathcal{C}_t^\alpha(y)} v_t(y, a, C) = v_t(y, a, \psi_t^*(y, a)). \tag{4.2}$$

Proof. By our assumption, the function $V_T(y) = \ell(x)$ is continuous and hence u.s.c.. For any $C \in \mathcal{P}([0,1]^n)$, we write

$$\mathbb{E}_{C \circ F_*}[V_T(\mathbf{G}(T-1, y, a, Z_T))] = \int_{E_Y} V_T(y_T) dQ_{T-1}(y_T | y, a, C).$$

According to Proposition 3.3, Q_{T-1} is a continuous stochastic kernel. Then, in view of the assumption that ℓ is bounded from above and [BS78, Proposition 7.31], we get that

$$\mathbb{E}_{C \circ F_*}[V_T(\mathbf{G}(T-1, y, a, Z_T))]$$

is u.s.c.. On the other hand, by Lemma 3.5, $-\rho_{T-1}(C)$ is u.s.c., and in turn, v_{T-1} are u.s.c..

Next, since that $[0,1]^n$ is compact, then $\mathcal{P}([0,1]^n)$ (equipped with metric $d_{W,p}$) is compact. Let $D = \bigcup_{(y,a) \in E_Y \times A} (\{(y, a)\} \times \mathcal{C}_t^\alpha(y))$ which can be viewed as the graph of the set valued function $\mathcal{C}_t^\alpha(y)$. Note that $\mathcal{C}_t^\alpha(y)$ is closed and, by Lemma 2.3, is u.h.c.. Hence, the set D is closed. According to [BS78, Proposition 7.33], the function $\check{V}_{T-1}(y, a) = \sup_{C \in \mathcal{C}_t^\alpha(y)} v_{T-1}(y, a, C)$ is u.s.c.. As a result, $V_{T-1}(y) = \inf_{a \in A} \check{V}_{T-1}(y, a)$ is u.s.c. as well. Moreover, for any fixed (y, a), there exists a Borel measurable function ψ^* such that (4.2) holds true.

Last but not the least, for any $C \in \mathcal{C}_{T-1}^\alpha(y)$ that is not a copula, it is not hard to see that $\rho_{T-1}(C) = +\infty$. Consequently, for such C, $v_{T-1}(y, a, C) = -\infty$. For any C that is a copula, we have

$$v_{T-1}(y, a, C) = \mathbb{E}_{C \circ F_*}[V_T(\mathbf{G}(t, y, a, Z_T))] \geq -\infty.$$

This indicates that for any (y, a), $\psi_{T-1}^*(y, a)$ is a copula. The rest of the proof follows analogously. $\square$

The above proposition proves the existence of measurable worst-case copula selector. Regarding the optimal control, we use the semi-analyticity property of the value function V_t as follows.

Proposition 4.2. *For every $t \in \mathcal{T}$, the function V_t is lower semi-analytic (l.s.a.). Moreover, for any $\varepsilon > 0$, there exists an analytically measurable function $\varphi_t^{*,\varepsilon}$ such that*

$$\sup_{C \in \mathcal{C}_t^\alpha(y)} v_t(y, \varphi_t^{*,\varepsilon}(y), C) = \begin{cases} V_t(y) + \varepsilon & \text{if } V_t(y) > -\infty, \\ -1/\varepsilon & \text{if } V_t(y) = -\infty. \end{cases} \tag{4.3}$$

In other words, $\varphi_t^{*,\varepsilon}$ is the ε-optimal control.

Proof. Since $V_T(y) = \ell(x)$ is continuous by assumption, then V_T is l.s.a.. Recall

$$v_{T-1}(y, a, C, \lambda_{T-1}) = \int_{E_Y} V_T(y_T) Q_{T-1}(dy_T | y, a, C) - \rho_{T-1}(C).$$

The first term on the RHS is l.s.a. as Q_{T-1} is a Borel measurable stochastic kernel by Proposition 3.3, and the second term $-\rho_{T-1}(C)$ is l.s.a. because it is u.s.c.. Therefore, the function $v_{T-1}(y, a, C)$ is l.s.a..

In view of (4.2) in Proposition 4.1 where ψ_{T-1}^* is Borel measurable, we deduce that

$$\widehat{V}_t(y,a) = \sup_{C \in \mathcal{C}_{T-1}^\alpha(y)} v_{T-1}(y,a,C)$$

is l.s.a. according to part (3) of [BS78, Lemma 7.30].

Next, the set A is also Borel measurable and in turn analytic. Hence, by [BS78, Proposition 7.50], $V_{T-1}(y)$ is l.s.a., and (4.3) holds true for $t = T - 1$. The rest of the proof follows analogously. □

Remark 4.3. We want to stress that in general the result in Lemma 4.2 cannot be improved in the sense that one is unable to obtain the measurable optimal control instead of ε-optimal control. Following [BS78, Proposition 7.50], on the set

$$I = \left\{ y \in E_Y : \exists a_y \in A \text{ such that } \widehat{V}_t(y,a_y) = V_t(y) \right\}$$

one will get a universally measurable optimal control. In our setup, all the continuous points of V_t belong to I. However, due to the presence of the penalty term, it is impossible to show the lower semi-continuity of V_t to the best of our knowledge.

Now, we proceed to show that the problem (3.9) satisfies the Bellman principle and is solved by (4.1).

4.2 Dynamic Programming

Here we will use Proposition 4.1 and Proposition 4.2 to show that problem (3.9) is solved by the Bellman equation (4.1). Denote

$$\rho_{t:T-1}(\psi) = \sum_{s=t}^{T-1} \rho_s(\psi_s),$$

then we have the main result of this section as follows.

Theorem 4.4. *For every $t \in \mathcal{T}'$, and any $y \in E_Y$, we have*

$$V_t(y) = \inf_{\varphi \in \mathcal{A}_t} \sup_{\mathbb{Q} \in \mathcal{Q}_{y,t}^\varphi} \left(\mathbb{E}_{\mathbb{Q}}[\ell(X_T)] - \rho_{t:T-1}(\psi) \right).$$

Proof. We will prove the statement via backward induction in $t = T-1, \ldots, 1, 0$.

When $t = T - 1$ and $y \in E_Y$, we have

$$\inf_{\varphi \in \mathcal{A}_{T-1}} \sup_{\mathbb{Q}^\psi \in \mathcal{Q}_{y,T-1}^\varphi} \left(\mathbb{E}_{\mathbb{Q}^\psi}[\ell(X_T)] - \rho_{T-1}(\psi) \right)$$

$$= \inf_{a \in A} \sup_{C \in \mathcal{C}_{T-1}^\alpha(y)} \left(\int_{E_Y} V_T(y') Q_{T-1}(y'|y,a,C) - \rho_{T-1}(C) \right) = V_{T-1}(y).$$

Next, for $t = T - 2, \ldots, 0$ and $y \in E_Y$, denote $\varphi^t = \{\varphi_s, s = t, \ldots, T-1\}$ and $\psi^t = \{\psi_s, s = t, \ldots, T-1\}$, then

$$\inf_{\varphi \in \mathcal{A}_t} \sup_{\mathbb{Q}^\psi \in \mathcal{Q}_{y,t}^\varphi} \left(\mathbb{E}_{\mathbb{Q}^\psi}[\ell(X_T)] - \rho_{t:T-1}(\psi) \right)$$

$$= \inf_{(a,\varphi^{t+1}) \in \mathcal{A}_t} \sup_{C \in \mathcal{C}_t^\alpha(y)} \int_{E_Y} \sup_{\mathbb{Q}^{\psi^{t+1}} \in \mathcal{Q}_{y',t+1}^{\varphi^{t+1}}} \left(\mathbb{E}_{\mathbb{Q}^{\psi^{t+1}}}[\ell(X_T)] - \rho_{t:T-1}((C, \psi^{t+1})) \right) Q_t(dy'|y,a,C),$$

where by induction

$$\sup_{\mathbb{Q}^{\psi^{t+1}} \in \mathcal{Q}^{\varphi^{t+1}}_{y',t+1}} \left(\mathbb{E}_{\mathbb{Q}^{\psi^{t+1}}}[\ell(X_T)] - \rho_{t:T-1}((C,\psi^{t+1})) \right)$$

$$= \sup_{\mathbb{Q}^{\psi^{t+1}} \in \mathcal{Q}^{\varphi^{t+1}}_{y',t+1}} \left(\mathbb{E}_{\mathbb{Q}^{\psi^{t+1}}}[\ell(X_T)] - \rho_{t+1:T-1}(\psi^{t+1}) - \lambda_t \rho_t(C) \right)$$

$$\geq V_{t+1}(y') - \rho_t(C).$$

Hence, we get

$$\inf_{\varphi \in \mathcal{A}_t} \sup_{\mathbb{Q}^{\psi} \in \mathcal{Q}^{\varphi}_{y,t}} \left(\mathbb{E}_{\mathbb{Q}^{\psi}}[\ell(X_T)] - \rho_{t:T-1}(\psi) \right)$$

$$\geq \inf_{(a,\varphi^{t+1}) \in \mathcal{A}_t} \sup_{C \in \mathcal{C}^{\alpha}_t(y)} \int_{E_Y} \left(V_{t+1}(y') - \rho_t(C) \right) Q_t(dy'|y,a,C)$$

$$= \inf_{a \in A} \sup_{C \in \mathcal{C}^{\alpha}_t(y)} \left(\int_{E_Y} V_{t+1}(y') Q_t(dy'|y,a,C) - \rho_t(C) \right) = V_t(y). \tag{4.4}$$

On the other hand, for any $\varepsilon > 0$, let $\varphi^{t+1,\varepsilon} \in \mathcal{A}_{t+1}$ be an ε-optimal control starting at time $t+1$. We obtain that

$$\sup_{\mathbb{Q}^{\psi^{t+1}} \in \mathcal{Q}^{\varphi^{t+1,\varepsilon}}_{y',t+1}} \left(\mathbb{E}_{\mathbb{Q}^{\psi^{t+1}}}[\ell(X_T)] - \rho_{t:T-1}((C,\psi^{t+1})) \right)$$

$$\leq \inf_{\varphi^{t+1} \in \mathcal{A}_{t+1}} \sup_{\mathbb{Q}^{\psi^{t+1}} \in \mathcal{Q}^{\varphi^{t+1}}_{y',t+1}} \left(\mathbb{E}_{\mathbb{Q}^{\psi^{t+1}}}[\ell(X_T)] - \rho_{t:T-1}((C,\psi^{t+1})) \right) + \varepsilon$$

$$= \left(V_{t+1}(y') - \lambda_t \rho_t(C) \right) + \varepsilon.$$

In what follows, we have

$$\inf_{\varphi \in \mathcal{A}_t} \sup_{\mathbb{Q}^{\psi} \in \mathcal{Q}^{\varphi}_{y,t}} \left(\mathbb{E}_{\mathbb{Q}^{\psi}}[\ell(X_T)] - \rho_{t:T-1}(\psi) \right)$$

$$\leq \inf_{a \in A} \sup_{\mathbb{Q}^{\psi} \in \mathcal{Q}^{(a,\varphi^{t+1,\varepsilon})}_{y',t+1}} \left(\mathbb{E}_{\mathbb{Q}^{\psi}}[\ell(X_T)] - \rho_{t:T-1}((C,\psi^{t+1})) \right)$$

$$= \inf_{a \in A} \sup_{C \in \mathcal{C}^{\alpha}_t(y)} \int_{E_Y} \sup_{\mathbb{Q}^{\psi^{t+1}} \in \mathcal{Q}^{\varphi^{t+1,\varepsilon}}_{y',t+1}} \left(\mathbb{E}_{\mathbb{Q}^{\psi^{t+1}}}[\ell(X_T)] - \rho_{t:T-1}((C,\psi^{t+1})) \right) Q_t(dy'|y,a,C)$$

$$\leq \inf_{a \in A} \sup_{C \in \mathcal{C}^{\alpha}_t(y)} \int_{E_Y} \left(V_{t+1}(y') - \rho_t(C) \right) Q_t(dy'|y,a,C) + \varepsilon$$

$$= \inf_{a \in A} \sup_{C \in \mathcal{C}^{\alpha}_t(y)} \left(\int_{E_Y} V_{t+1}(y') Q_t(dy'|y,a,C) - \rho_t(C) \right) + \varepsilon$$

$$= V_t(y) + \varepsilon.$$

Since ε is arbitrary, as a consequence,

$$\inf_{\varphi \in \mathcal{A}_t} \sup_{\mathbb{Q}^{\psi} \in \mathcal{Q}^{\varphi}_{y,t}} \left(\mathbb{E}_{\mathbb{Q}^{\psi}}[\ell(X_T)] - \rho_{t:T-1}(\psi) \right) \leq V_t(y). \tag{4.5}$$

15

Combine (4.4) and (4.5), we conclude that

$$\inf_{\varphi \in \mathcal{A}_t} \sup_{\mathbb{Q}^\psi \in \mathcal{Q}^\varphi_{y,t}} \left(\mathbb{E}_{\mathbb{Q}^\psi}[\ell(X_T)] - \rho_{t:T-1}(\psi) \right) = V_t(y).$$

$\square$

The above theorem shows that the problem (3.9) is solved via the Bellman equation (4.1). In the next section, we will compare the approach (3.9) to some different methods that can be used to handle the optimization problem under dependence uncertainty.

4.3 Comparison to Robust Control Problem via Empirical Distribution

In this section, we will compare three different setups that can be used for solving the proposed problem. One of other approaches that is viable in this setup is, as we mentioned earlier, the multidimensional version of the methodology discussed in [BC22]. Briefly speaking, one can build confidence regions in terms of Wasserstein balls centered at the empirical distribution $\widehat{F}_t$, say, denoted as $\mathcal{C}^{\alpha,e}_t(\widehat{F}_t)$. Then, formulate the adaptive robust problem that optimizes the expected loss/utility against the worst-case model in $\mathcal{C}^{\alpha,e}_t(\widehat{F}_t)$. Obviously, doing so will ignore the available information of the marginals $F^{(i)}_*$, $i \in \mathcal{N}$. However, for large value of T, such method will have a decent performance as the corresponding value function will get quite closed to the true one when t approaches T (cf. [BC22]). The associated Bellman equation is given as follows

$$V^e_t(y') = \inf_{a \in A} \sup_{F \in \mathcal{C}^{\alpha,e}_t(y')} \mathbb{E}_F[V^e_{t+1}(\mathbf{G}(t, y', a, Z_{t+1}))], \quad t \in \mathcal{T}', \tag{4.6}$$

where y' is the state of the process $Y'_t = (X_t, \widehat{F}_t)$, $t \in \mathcal{T}'$. The third method is somewhat similar to the one we just described but utilizes the information of the known marginals. One can impose essentially the same penalty function as in this work and force the worst-case distribution to have correct marginals. To this end, we define the following one-step penalty function

$$\rho^m_t(F) = \sum_{i \in \mathcal{N}} \sup_{f_i \in C_b(\mathbb{S}^{(i)})} \left(\mathbb{E}_{F^{(i)}}[f_i] - \mathbb{E}_{F^{(i)}_*}[f_i] \right),$$

for any $F \in \mathcal{C}^{\alpha,m}_t(\widehat{F}_t)$, which is the Wasserstein ball centered at $\widehat{F}_t$ that might have a different radius compared to $\mathcal{C}^{\alpha,e}_t(\widehat{F}_t)$. Also, the above $\mathbb{S}^{(i)}$ is the support of $Z^{(i)}_{t+1}$, $i \in \mathcal{N}$. The corresponding Bellman equation is

$$V^m_t(y') = \inf_{a \in A} \sup_{F \in \mathcal{C}^{\alpha,m}_t(y')} \left(\mathbb{E}_F[V^m_{t+1}(\mathbf{G}(t, y', a, Z_{t+1}))] - \rho_t(F) \right), \quad t \in \mathcal{T}', \tag{4.7}$$

where y' is defined the same way as in (4.6).

Now the question is that which framework is the best in what situation, and can the methods with imposed penalty functions take advantage of the available information of the marginals and outperform the approach that ignores it? First off, note that one can always take $\mathcal{C}^{\alpha,m}_t = \mathcal{C}^{\alpha,e}_t$ backed by the concentration result. Then, due to the presence of the penalty function $\rho_t(F)$, framework (4.7) is essentially searching for the worst-case distribution in a smaller set compared to (4.6). Therefore, we have

$$V^m_t(y') \leq V^e_t(y').$$

In other words, (4.7) yields smaller loss which is meant to be minimized. In this sense, the solution corresponding to (4.7) is better than the one corresponding to (4.6).

The comparison between (3.9) and (4.7) is much more complicated, and it turns out that the preference depends on additional properties of F^*. To this end, we first discuss the relationship between the uncertainty sets $\mathcal{C}_t^\alpha(y)$ and $\mathcal{C}_t^{\alpha,e}(y')$ where $y = (x, \widehat{C})$ and $y' = (x, \widehat{F})$ such that $\widehat{C} = \widehat{F} \circ F_*^{-1}$. This means that both methods use the exact same information at every time step $t \in \mathcal{T}'$. In the first case, when there exists a constant $B > 1$ such that

$$\|F_* \circ z_1 - F_* \circ z_2\| \leq B\|z_1 - z_2\|, \tag{4.8}$$

where

$$F_* \circ z = \left(F_*^{(1)}(z^{(1)}), \ldots, F_*^{(n)}(z^{(n)}) \right).$$

Hence, for any f_1, f_2 on $[0, 1]^n$, such that

$$|f_1(u_1) - f_2(u_2)| \leq \|u_1 - u_2\|^p, \quad u_1, u_2 \in [0, 1]^n,$$

we get

$$|f_1(F_* \circ z_1) - f_2(F_* \circ z_2)| \leq \|F_* \circ z_1 - F_* \circ z_2\|^p \leq B^p \|z_1 - z_2\|^p, \quad z_1, z_2 \in \mathbb{R}^n.$$

Following such obervation, we have for any C

$$
\begin{aligned}
d_{W,p}^p(\widehat{C}, C) &= \sup_{f_1, f_2} \left(\int_{[0,1]^n} f_1(u) d\widehat{C}(u) - \int_{[0,1]^n} f_2(u) dC(u) \right) \\
&= \sup_{f_1, f_2} \left(\int_{\mathbb{R}^n} f_1(F_* \circ z) d\widehat{C}(F_* \circ z) - \int_{\mathbb{R}^n} f_2(F_* \circ z) dC(F_* \circ z) \right) \\
&\leq B^p \sup_{g_1, g_2} \left(\int_{\mathbb{R}^n} g_1(z) d\widehat{C}(F_* \circ z) - \int_{\mathbb{R}^n} g_2(z) dC(F_* \circ z) \right) \\
&= B^p d_{W,p}^p(\widehat{F}, C \circ F_*), \tag{4.9}
\end{aligned}
$$

where $|g_1(z_1) - g_2(z_2)| \leq \|z_1 - z_2\|^p$ for any $z_1, z_2 \in \mathbb{R}^n$. Thus, for the radius $r^m(\alpha, t_0, t)$ of $\mathcal{C}_t^{\alpha,m}(y')$, it suffices to take $r^m(\alpha, t_0, t) = r(\alpha, t_0, t)/B$. Note that, in theory one should choose $r^e(\alpha, t_0, t)$, which is the radius of $C_t^{\alpha,e}$, as $r^e(\alpha, t_0, t) = r(\alpha, t_0, t)$. Hence, in this case, (4.7) will perform better than (4.6) because it has a smaller associated Wasserstein ball.

On the other hand, between (3.9) and (4.7), in view of (4.9), we obtain that for any $F \in \mathcal{C}_t^{\alpha,m}(y')$, there exists a corresponding $C \in \mathcal{C}_t^\alpha(y)$. Consequently, we can view $\mathcal{C}_t^{\alpha,m}$ as a "smaller" set than $\mathcal{C}_t^\alpha$. Moreover, we immediately get that

$$V_t^m(y') \leq V_t(y),$$

and (4.7) outperforms (3.9). One should note that, in this case, it is unclear which one is better between (3.9) and (4.7).

The other case is that each $\mathbb{S}^{(i)}$, $i \in \mathcal{N}$, is a compact subset of $\mathbb{R}$, and we suppose that there exists a constant $B' > 0$ such that

$$\|F_* \circ z_1' - F_* \circ z_2'\| \geq B'\|z_1' - z_2'\|, \tag{4.10}$$

for any $z_1', z_2' \in \prod_{i \in \mathcal{N}} \mathbb{S}^{(i)}$. Under these assumptions, for any g_1, g_2 on $\mathbb{R}^n$, such that

$$|g_1(z_1') - g_2(z_2')| \leq \|z_1' - z_2'\|^p, \quad z_1', z_2' \in \prod_{i \in \mathcal{N}} \mathbb{S}^{(i)},$$

we obtain

$$|g_1(F_*^{-1} \circ u_1) - g_2(F_*^{-1} \circ u_2)| \leq \|F_*^{-1} \circ u_1 - F_*^{-1} \circ u_2\|^p \leq \frac{1}{B'}\|u_1 - u_2\|, \quad u_1, u_2 \in [0, 1]^n.$$

Then, for any F whose associated random variable has support $\prod_{i \in \mathcal{N}} \mathbb{S}^{(i)}$, and $F = C \circ F_*$ for some copula C, we have

$$
\begin{aligned}
d_{W,p}^p(\widehat{F}, F) &= \sup_{g_1, g_2} \left(\int_{\mathbb{R}^n} g_1(z) d\widehat{F}(z) - \int_{\mathbb{R}^n} g_2(z) dF(z) \right) \\
&= \sup_{g_1, g_2} \left(\int_{\prod_{i \in \mathcal{N}} \mathbb{S}^{(i)}} g_1(z) d\widehat{F}(z) - \int_{\prod_{i \in \mathcal{N}} \mathbb{S}^{(i)}} g_2(z) dF(z) \right) \\
&= \sup_{g_1, g_2} \left(\int_{\prod_{i \in \mathcal{N}} \mathbb{S}^{(i)}} g_1(z) d\widehat{C}(F_* \circ z) - \int_{\prod_{i \in \mathcal{N}} \mathbb{S}^{(i)}} g_2(z) dC(F_* \circ z) \right) \\
&= \sup_{g_1, g_2} \left(\int_{[0,1]^n} g_1(F_*^{-1}(u) d\widehat{C}(u) - \int_{[0,1]^n} g_2(F_*^{-1}(u) dC(u) \right) \\
&\leq \frac{1}{B'} \sup_{f_1, f_2} \left(\int_{[0,1]^n} f_1(u) d\widehat{C}(u) - \int_{[0,1]^n} f_2(u) dC(u) \right) \\
&= \frac{1}{B'} d_{W,p}^p(\widehat{C}, C),
\end{aligned}
$$

where $|f_1(u_1) - f_2(u_2)| \leq \|u_1 - u_2\|$ for any $u_1, u_2 \in [0, 1]^n$. As a result, similarly to the above discussion, we view $\mathcal{C}_t^\alpha$ as a "smaller" set than $\mathcal{C}_t^{\alpha, m}$, and we conclude

$$V_t(y) \leq V_t^m(y') \leq V_t^e(y').$$

In other words, the framework proposed in this work will perform the best among all three mentioned approaches.

Finally, we remark that when $\mathbb{S}^{(i)}$, $i \in \mathcal{N}$, are compact, and both (4.8) and (4.10) hold true, the two frameworks (3.9) and (4.7) are equivalent. If neither (4.8) or (4.10) is satisfied, it is unclear which one between (3.9) and (4.7) is more preferable.

5 Uncertain Utility Maximization with Known Marginal Distributions

In this section, we will apply the adaptive robust control methodology to the uncertain utility maximization problem where the marginal distributions of the underlying random noise are known. Note that even though our theory applies to both empirical distribution and perturbed empirical distribution, we will focus only on the case of empirical distribution in this section. The advantage of using the perturbed empirical distribution is that the resulting $\widehat{C}$ is a copula. This could be a more sounding approach in some applications, however in this work, it will increase the computation burden which is already heavy. Hence, we leave the corresponding investigating for potential future studies.

Towards this end, we first discuss the algorithm that we will implement.

5.1 Numerical Algorithm

From the practical point of view, one major difficulty in solving a robust control problem with Wasserstein uncertainty set such as $\mathcal{C}_t^\alpha$ is that direct searching through $\mathcal{C}_t^\alpha(y)$ to find the optimizer $\psi_t^*(y)$ is impossible. Typically, one uses a duality argument to re-write such problem as a scalar optimization problem. In [EK18], the authors provide a rather comprehensive discussion on the tractable reformulations of an optimization problem over the Wasserstein ball. With some postulation on the loss function, the authors achieve a convex reduction of the worst-case expectation problems. However, results in [EK18] only apply to the case of optimizing the expected loss while in this work we deal with a different problem as in (4.1)

$$\sup_{C \in \mathcal{C}_t^\alpha(y)} \left(\mathbb{E}_{C \circ F_*}[V_{t+1}(\mathbf{G}(t, y, a, Z_{t+1}))] - \rho_t(C) \right). \tag{5.1}$$

Due to the non-linearity of the penalty function ρ_t, one cannot directly adopt the idea in [EK18] to reformulate problem (4.1) for the numerical purpose.

To this end, we recall the paper [GK17] which considers the numerical solution of a 1-period robust optimization problem with Wasserstein uncertainty set centered at the estimated copula which is similar to our worst-case copula optimization problem. Let $\{z_{-t_0+1}, \ldots, z_t\}$ be the historical data of Z up to time $t \in \mathcal{T}'$, the strong duality result in [GK17] implies that problem (5.1) can be solved as

$$\inf_{\gamma \in \mathbb{R}_+, f_i \in C_b(\mathbb{R})} \left\{ \gamma r(\alpha, t_0, t)^p + \sum_{i=1}^n \int_{\mathbb{R}} f_i(z) dF_*^{(i)}(z) + \frac{1}{t_0 + t} \sum_{j=-t_0+1}^t V_{t+1}^\gamma(\mathbf{G}(t, y, a, z_j)) \right\}, \tag{5.2}$$

where

$$V_{t+1}^\gamma(\mathbf{G}(t, y, a, z_j)) = \sup_{z \in \mathbb{R}^n} \left(V_{t+1}(\mathbf{G}(t, y, a, z)) - \sum_{i=1}^n f_i(z^{(i)}) - \gamma d_{F_*}^p(z, z_j) \right), \tag{5.3}$$

and

$$d_{F_*}(\xi, \zeta) = d\left(\left(F_*^{(1)}(\xi^{(1)}), \ldots, F_*^{(n)}(\xi^{(n)}) \right), \left(F_*^{(1)}(\zeta^{(1)}), \ldots, F_*^{(n)}(\zeta^{(n)}) \right) \right). \tag{5.4}$$

Remark 5.1. In general, the premetric d_{F_*} is defined as

$$d_{F_*}(\xi, \zeta) = \liminf_{d(\xi_i, \xi), d(\zeta_i, \zeta) \to 0} d\left(\left(F_*^{(1)}(\xi_i^{(1)}), \ldots, F_*^{(n)}(\xi_i^{(n)}) \right), \left(F_*^{(1)}(\zeta_i^{(1)}), \ldots, F_*^{(n)}(\zeta_i^{(n)}) \right) \right).$$

Due to our assumption that $F_*^{(i)}$, $i \in \mathcal{N}$, are continuous, so d_{F_*} becomes (5.4).

In [GK17], to use the strong duality result in practice, the authors essentially assume that the range of the random variable Z_t, $t \in \mathcal{T}$, is a finite set. Such assumption is obviously not true in this work, and hence we will use the following treatment to overcome such obstacle. Towards this end, denote by $C([0, 1])$ the set of all continuous functions on $[0, 1]$, we first apply the change of variable to (5.2) – (5.3) as follows

$$\inf_{\gamma \in \mathbb{R}_+, f_i \in C([0,1])} \left\{ \gamma r(\alpha, t_0, t)^p + \sum_{i=1}^n \int_0^1 f_i\left(F_*^{(i),-1}(u) \right) du + \frac{1}{t_0 + t} \sum_{j=-t_0+1}^t V_{t+1}^\gamma\left(\mathbf{G}\left(t, y, a, F_*^{-1}(u_j)\right) \right) \right\},$$

where

$$V_{t+1}^{\gamma}\left(\mathbf{G}\left(t,y,a,F_*^{-1}(u_j)\right)\right) = \sup_{u\in[0,1]^n}\left(V_{t+1}\left(\mathbf{G}\left(t,y,a,F_*^{-1}(u)\right)\right) - \sum_{i=1}^{n} f_i\left(F_*^{(i),-1}(u^{(i)})\right) - \gamma d^p(u,u_j)\right).$$

Next, we recall the Weierstrass approximation theorem which indicates that any $f_i \circ F_*^{(i),-1} \in C([0,1])$ can be approximated by polynomials. In addition, the Bernstein polynomials

$$\beta_{k,K}(u) = \binom{K}{k}u^k(1-u)^{K-k}, \quad u\in[0,1], \ k=0,\dots,K,$$

form a basis for the space of polynomials of degree K. Also, note that for fixed K and any $k=0,\dots,K$,

$$\int_{[0,1]} \beta_{k,K}(u)du = \frac{1}{K+1}.$$

Hence, the above optimization problem is approximated as

$$\inf_{\gamma\in\mathbb{R}_+, g_{i,k}\in\mathbb{R}}\left\{\gamma r(\alpha,t_0,t)^p + \sum_{i=1}^{n}\sum_{k=0}^{K}\frac{g_{i,k}}{K+1} + \frac{1}{t_0+t}\sum_{j=-t_0+1}^{t} V_{t+1}^{\gamma}(\mathbf{G}(t,y,a,F_*^{-1}(u_j)))\right\},$$

where

$$V_{t+1}^{\gamma}(\mathbf{G}(t,y,a,F_*^{-1}(u_j))) = \sup_{u\in[0,1]^n}\left(V_{t+1}(\mathbf{G}(t,y,a,F_*^{-1}(u))) - \sum_{i=1}^{n}\sum_{k=0}^{K} g_{i,k}\beta_{k,K}(u^{(i)}) - \gamma d^p(u,u_j)\right).$$

By including the optimization over feasible controls, the inf-sup problem we need to solve is

$$\inf_{a\in A,\gamma\in\mathbb{R}_+, g_{i,k}\in\mathbb{R}}\left\{\gamma r(\alpha,t_0,t)^p + \sum_{i=1}^{n}\sum_{k=0}^{K}\frac{g_{i,k}}{K+1} + \frac{1}{t_0+t}\sum_{j=-t_0+1}^{t} V_{t+1}^{\gamma}(\mathbf{G}(t,y,a,F_*^{-1}(u_j)))\right\}, \quad (5.5)$$

where

$$V_{t+1}^{\gamma}(\mathbf{G}(t,y,a,F_*^{-1}(u_j))) = \sup_{u\in[0,1]^n}\left(V_{t+1}(\mathbf{G}(t,y,a,F_*^{-1}(u))) - \sum_{i=1}^{n}\sum_{k=0}^{K} g_{i,k}\beta_{k,K}(u^{(i)}) - \gamma d^p(u,u_j)\right).$$
$$(5.6)$$

With regard to designing a numerical algorithm to solve (5.5) – (5.6), we first note that the inf problem in (5.5) and the sup problem in (5.6) have dimensions $m + n*(K+1) + 1$ and n, respectively. Hence, we have a non-trivial high dimensional inf-sup optimization problem for which a solver using brute force will be extremely inefficient and virtually impossible. On top of the inherent high dimension in optimization, it also requires large amount of computation cost to evaluate $\frac{1}{t_0+t}\sum_{j=-t_0+1}^{t} V_{t+1}^{\gamma}(\mathbf{G}(t,y,a,F_*^{-1}(u_j)))$ given that t_0 is decently big. In view of these difficulties, we propose to use the stochastic gradient descent ascent (SGDA) method to overcome the obstacles.

To explain the idea of SGDA in our work, for any fixed $t\in\mathcal{T}'$, $y\in E_Y$, we denote

$$\hat{v}_t(a,\gamma,g,u;\hat{u}) = \gamma(r(\alpha,t_0,t)^p - d^p(u,\hat{u})) + \sum_{i=1}^{n}\sum_{k=0}^{K}\left(\frac{g_{i,k}}{K+1} - g_{i,k}\beta_{k,K}(u^{(i)})\right)$$
$$+ V_{t+1}(\mathbf{G}(t,y,a,F^{-1}(u))),$$

where $\hat{u}\in\{u_j, j=-t_0+1,\dots,t\}$. Such formulation comes from the considered loop as follows. Given some $(a(l),\gamma(l),g(l),u(l))$, $l=0,1,\dots$,

1. uniformly simulate $\hat{u}(l) \in \{u_j, j = -t_0 + 1, \ldots, t\}$;

2. update $(a(l), \gamma(l), g(l), u(l))$ as

$$a(l+1) = a(l) - \eta(l)\frac{\partial}{\partial a}\hat{v}_t(a(l), \gamma(l), g(l), u(l); \hat{u}(l)),$$

$$\gamma(l+1) = \gamma(l) - \eta(l)\frac{\partial}{\partial \gamma}\hat{v}_t(a(l), \gamma(l), g(l), u(l); \hat{u}(l)),$$

$$g(l+1) = g(l) - \eta(l)\frac{\partial}{\partial g}\hat{v}_t(a(l), \gamma(l), g(l), u(l); \hat{u}(l)),$$

$$u(l+1) = u(l) + \eta(l)\frac{\partial}{\partial u}\hat{v}_t(a(l), \gamma(l), g(l), u(l); \hat{u}(l)),$$

where $\eta(l)$ denote the step size;

3. Goto 1.

The above loop will stop when there is no improvement of $(a(l), \gamma(l), g(l), u(l))$ anymore, and we denote the corresponding number l of iterations as l^*.

To apply SGDA, one needs to have a functional representation of V_t, $t = 1, \ldots, T-1$, such that the above gradients of $\hat{v}_t$ can be quickly computed. To construct such functional approximation, we first choose N so-called design points y_t^j, $j = 1, \ldots, N$, at which we solve for $V_t(y_t^j)$, $j = 1, \ldots, N$, and then build a regression model $\widetilde{V}_t$ based on the training data $(y_t^j, V_t(y_t^j))$, $j = 1, \ldots, N$. For the choice of the regression model, we will use the Gaussian process (GP) with kernel function $\mathbf{k}(\cdot, \cdot)$ of the Matern-3/2 type. For detailed discussion of GP, one can refer to [RW06]. We also summarize the basic idea of GP in earlier works (cf. [CL21], [BCC20], [CM20], [BC22]).

When it comes to regression, one needs to note that the second component $\widehat{C}$ of the state variable y is in general a distribution and hence infinitely dimensional, which cannot be regressed against numerically. To such end, we will approximate each $\widehat{C}$ with its first m moments of marginals and covariance among marginals. By doing so, the dimension of the state variable y is reduced to $mn + \binom{n}{2} + 1$ and we denote the approximating vector by $\tilde{y}$.

The rationale behind choosing GP with the Matern-3/2 kernel is twofold. On one hand, it is a regression model sophisticated enough to capture the complicated structure of V_t with minimal assumptions, as Matern-3/2 kernel is used for regressing the function that has first order derivative. On the other hand, the gradients of $\widetilde{V}_t$, and in turn of $\hat{v}_t$, can be analytically computed which is of particular importance to our implementation. To see this, recall that given y_t^j, $j = 1, \ldots, N$, the resulting regression model has the following representation

$$\widetilde{V}_t(\tilde{y}) = (\mathbf{k}(\tilde{y}, \tilde{y}_t^1), \ldots, \mathbf{k}(\tilde{y}, \tilde{y}_t^N))[\mathbf{K} + \varepsilon^2\mathbf{I}]^{-1}(V_t(y_t^1), \ldots, V_t(y_t^N))^\top. \qquad (5.7)$$

The above $\varepsilon^2\mathbf{I}$ with $\mathbf{I}$ being the identity matrix is the perturbation added to the matrix $\mathbf{K}$ to ensure that the latter is invertible. The matrix $\mathbf{K}$ satisfies that $\mathbf{K}_{ij} = \mathbf{k}(\tilde{y}_t^i, \tilde{y}_t^j)$ and

$$\mathbf{k}(\tilde{y}^1, \tilde{y}^2) = \left(1 + \sqrt{3}d_s(\tilde{y}^1, \tilde{y}^2)\right)\exp\left(-\sqrt{3}d_s(\tilde{y}^1, \tilde{y}^2)\right),$$

where d_s is the scaled Euclidean distance between y^1 and y^2 with parameter s:

$$d\left(\left(\frac{\tilde{y}^{1,(1)}}{s^{(1)}}, \ldots, \frac{\tilde{y}^{1,(n)}}{s^{(n)}}\right), \left(\frac{\tilde{y}^{2,(1)}}{s^{(1)}}, \ldots, \frac{\tilde{y}^{2,(n)}}{s^{(n)}}\right)\right),$$

where $d(\cdot, \cdot)$ is the usual Euclidean distance. The vector s is called the length scale and fitting a GP model means to use the maximum likelihood method to find the optimal value of s. After $\widetilde{V}_t$ is fitted, the product $[\mathbf{K} + \varepsilon^2 \mathbf{I}]^{-1} (V_t(y_t^1), \ldots, V_t(y_t^N))^\top$ in (5.7) is a fixed constant vector $(\nu_t^1, \ldots, \nu_t^N)$. Therefore, (5.7) can be rewritten as $\widetilde{V}_t(\tilde{y}) = \sum_{j=1}^N \nu_t^j \mathbf{k}(\tilde{y}, \tilde{y}_t^j)$. As a result, gradient of $\widetilde{V}_t$ is a linear combination of the gradients of $\mathbf{k}(\tilde{y}, \tilde{y}_t^j)$ and with the explicit formula of $\mathbf{k}$ being available, its gradient can therefore be computed analytically. There is no need to stress how important the capability of computing the gradient without any numerical approximation is in a gradient based optimization algorithm. Our solver for problem (5.5) – (5.6) is summarized as follows. For $t = T - 2, \ldots, 0$,

1. Assume that $V_{t+1}(y_{t+1}^i)$ and $\varphi_{t+1}^{*,\varepsilon}(y_{t+1}^i)$, $i = 1, \ldots, N$, are computed. Fit the GP models $\widetilde{V}_{t+1}$ and $\widetilde{\varphi}_{t+1}^{*,\varepsilon}$ by using the training data $(y_{t+1}^i, V_{t+1}(y_{t+1}^i))$, and $(y_{t+1}^i, \varphi_{t+1}^{*,\varepsilon}(y_{t+1}^i))$, $i = 1, \ldots, N$, respectively.

2. Choose $y_t^i \in E_Y$, $i = 1, \ldots, N$.

3. For each y_t^i, choose initial guesses $(a(0), \gamma(0), g(0), u(0))$, and use SGDA to compute $(a(l^*), \gamma(l^*), g(l^*), u(l^*))$. Set $V_t(y_t^i) = \hat{v}_t(a(l^*), \gamma(l^*), g(l^*), u(l^*); \hat{u}(l^*))$, and $\varphi_t^{*,\varepsilon}(y_t^i) = a(l^*)$.

4. Goto 1: start the next recursion for $t - 1$.

To analyze the performance of the computed optimal control $\varphi^{*,\varepsilon}$, we use forward simulation to estimate the mean terminal loss over the out-of-sample paths. For $t = 0, \ldots, T - 1$,

1. Draw $N' > 0$ i.i.d. $Z_{t+}^1, \ldots, Z_{t+1}^{N'}$ from the true distribution corresponding to F_*.

2. All paths start from the initial y_0. The state along path i is updated according to $\mathbf{G}(t, y_t^i, \widetilde{\varphi}_t^{*,\varepsilon}(y_t^i), Z_{t+1}^i)$, $i = 1, \ldots, N'$.

3. Obtain the terminal state X_T^i, $i = 1, \ldots, N'$, and compute

$$V^c := \frac{1}{N'} \sum_{i=1}^{N'} \ell(X_T^i)$$

as the estimated mean terminal loss.

In Section 4.3, we know that there is no complete theoretical result of the comparison between our adaptive robust method based on copula and the approach of using empirical distribution without marginal information. Hence, we will focus on such comparison when presenting our numerical results.
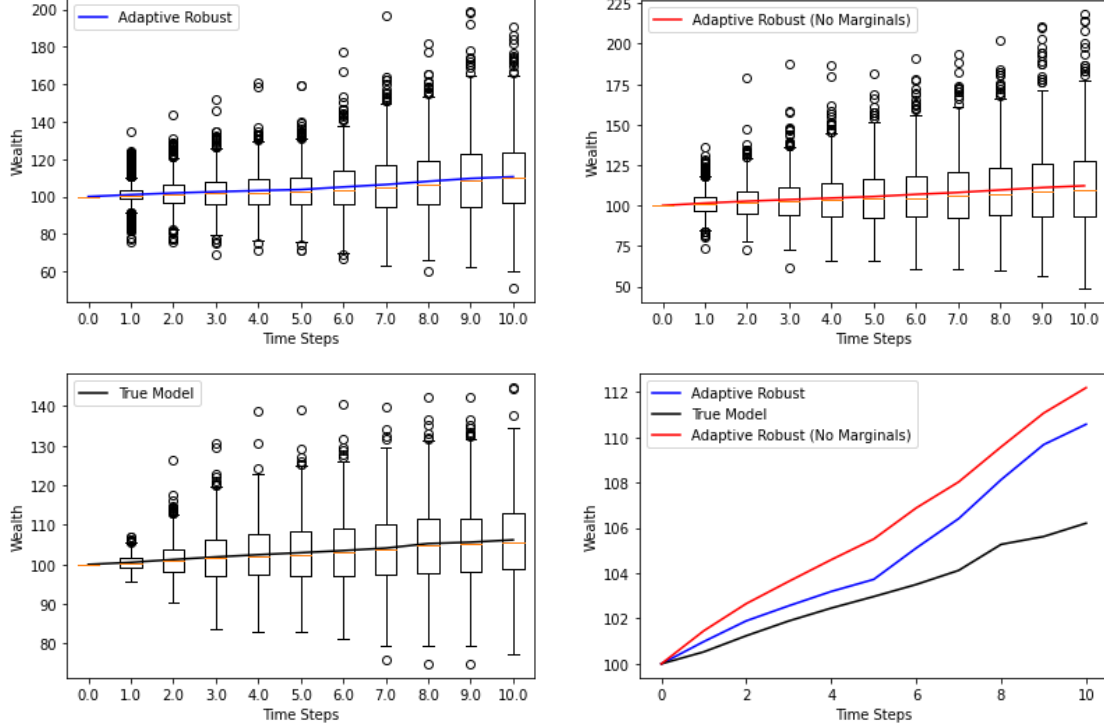
Figure 1: *Evolution of the wealth along out-of-sample paths. Upper left: box-plot of the wealth generated by $\varphi^{*,\varepsilon}$; Upper right: box-plot of the wealth generated by φ^e; Bottom left: box-plot of the wealth generated by φ^{tr}; Bottom right: comparison of mean wealth.*

5.2 Numerical Results

In this section, we apply the algorithm described above to the uncertain utility maximization problem. We take $d = 2$ and consider two different stocks where Z_{t+1} is the vector of log-returns from time t to $t+1$, which follows a multi-dimensional normal distribution. Correspondingly, let X_t be the wealth at time t with the dynamics

$$X_t = X_{t-1}\left((1 - \varphi_t^{(1)} - \varphi_t^{(2)})(1 + r) + \varphi_t^{(1)}e^{Z_t^{(1)}} + \varphi_t^{(2)}e^{Z_t^{(2)}}\right).$$

The above $\varphi_t^{(i)}$, $i = 1, 2$, is the proportion of wealth invested in stock i. The constant r and $Z_t^{(i)}$, $i = 1, 2$, are the interest rate and log-return of stock i, respectively. In addition, we choose the loss function ℓ such that $-\ell = U$ which is the exponential utility function. Then, we are effectively facing a utility maximization problem involving two different stocks under model uncertainty.

We compare three different types of strategies: 1. adaptive robust control method with known marginals; 2. adaptive robust control method based on empirical distribution without using marginal information; 3. optimal control without uncertainty. We denote by $\varphi^{*,\varepsilon}$, φ^e, and φ^{tr} the corresponding strategies, respectively. We also use V^c, V^e, and V^{tr} to denote the corresponding estimated expected utility, respectively. In Section 4.3, we find that it is unclear which of $\varphi^{*,\varepsilon}$ and φ^e will perform better when the marginal CDFs are Lipschitz continuous. Hence, we will compare these two strategies for Lipschitz marginals by using numerical results presented in the sequel.

To proceed, we take the following values for our parameters: $t_0 = 400$, $T = 10$, $N = N' = 1000$, $\alpha = 0.1$, $p = 2$, and $m = 2$. Regarding the true distribution F_*, it is bivariate normal with

mean $\mu = (0.09, 0.13)$ and covariance matrix $\Sigma = \left(\begin{smallmatrix} 0.25^2 & 0.85 \cdot 0.25 \cdot 0.4 \\ 0.85 \cdot 0.25 \cdot 0.4 & 0.4^2 \end{smallmatrix} \right)$ where all numbers are annualized. Finally, the annual interest rate is $r = 0.02$. With $T = 10$, it means that we consider 10 trading periods in a year.

	AR	AR (No Marginals)	TR
V	19.8726	19.8659	19.8857
$\text{var}(X_T)$	441.7334	628.7182	118.3159
$q_{0.30}(X_T)$	92.4007	90.4748	96.3813
$q_{0.90}(X_T)$	136.7146	144.7677	121.2001
$\max(X_T)$	190.4936	190.4936	144.6299
$\min(X_T)$	51.2641	48.7990	77.4599

Table 1: *Expected utility, variance, 30%-quantile, 90%-quantile, maximum, and minimum of the out-of-sample terminal wealth. AR: adaptive robust; TR: no uncertainty.*

Table 1 shows that, in terms of the expected utility, adaptive robust using marginals performs better than the approach without using information of marginals, but worse than the case of knowing the true model. Similar comparison results hold true if we look at the 30% quantile and the minimum value of the terminal wealth. On the contrary, φ^e produces the highest variance, 90% quantile, and maximum of the terminal wealth, while φ^{tr} generates the lowest values in these aspects. These observations can be obtained by viewing the box plots in Figure 1. In addition, adaptive robust without information of marginals has the highest mean wealth while knowing the true model gives the lowest value.

As we can see, when the investor knows more about the true model, the resulting terminal wealth excels in the risk management aspects such as variance, 30% quantile, and minimum value. Knowing less information of the true model will produce higher mean, 90% quantile, and maximum of the wealth which signal a strategy that is relatively more aggressive and profit seeking. By looking at the expected utility which takes into account both profit seeking and risk aversion, we notice that the more information we know about the model the higher mean utility we obtain. With such numerical results, we conclude that, in general, knowing more about true model leads to a more balanced and less risky optimal strategy.

References

[BC21]   T. Bhudisaksang and A. Cartea. Online drift estimation for jump-diffusion processes. Bernoulli - Journal of the Bernoulli Society, 27(4):2494–2518, 2021.

[BC22]   E. Bayraktar and T. Chen. Nonparametric adaptive robust control under model uncertainty. Preprint, 2022.

[BCC17]   T. Bielecki, T. Chen, and I. Cialenco. Recursive construction of confidence regions. Electron. J. Statist., 11(2):4674–4700, 2017.

[BCC20]   T. R. Bielecki, T. Chen, and I. Cialenco. Time-inconsistent markovian control problems under model uncertainty with application to the mean-variance portfolio selection. Submitted for publication, 2020.

[BS78]   D. P. Bertsekas and S. Shreve. Stochastic Optimal Control: The Discrete-Time Case. Academic Press, 1978.

[CG91]   H. F. Chen and L. Guo. Identification and stochastic adaptive control. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., 1991.

[CL21]   T. Chen and M. Ludkovski. A machine learning approach to adaptive robust utility maximization and hedging. SIAM Journal on Financial Mathematics, 3(12):1226–1256, 2021.

[CM20]   T. Chen and J. Myung. Nonparametric adaptive bayesian stochastic control under model uncertainty. Preprint, 2020.

[DBGM99] E. Del Barrio, E. Giné, and C. Matrán. Central limit theorems for the wasserstein distance between the empirical andthe true distributions. The Annals of Probability, 27(2):1009–1071, 1999.

[EK18]   P. M. Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations. Mathematical Programming, 171:115–166, 2018.

[FG15]   N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. Probability Theory and Related Fields, 162:707–738, 2015.

[GK17]   R. Gao and A. Kleywegt. Data-driven robust optimization with known marginal distributions. In Technical Report. Georgia Institute of Technology, 2017.

[GS89]   I. Gilboa and D. Schmeidler. Maxmin expected utility with nonunique prior. J. Math. Econom., 18(2):141–153, 1989.

[KV15]   P. R. Kumar and P. Varaiya. Stochastic Systems: Estimation, Identification, and Adaptive Control. Prentice Hall, Inc., 2015.

[Rie75]   U. Rieder. Bayesian dynamic programming. Adv. Appl. Prob., 7:330–348, 1975.

[RW06]   C. E. Rasmussen and C. K. I. Williams. Gaussian Processes for Machine Learning. The MIT Press, 2006.