Automatica 130 (2021) 109641

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Technical communique

Dynamic programming with value convexity[☆]

Guanlong Ren, John Stachurski *

Research School of Economics, Australian National University, ACT, Australia

ARTICLE INFO

Article history:
Received 7 October 2020
Received in revised form 27 December 2020
Accepted 1 March 2021
Available online 12 April 2021

Keywords:
Dynamic programming
Recursive preferences
Ambiguity

ABSTRACT

Many recent dynamic programming specifications fail to satisfy traditional contractivity conditions, which are a cornerstone of the standard optimality theory for infinite horizon problems in discrete time. We formulate alternative conditions based around monotonicity and "value" convexity. These conditions lead to an optimality theory that is as strong as the contractive case. Several applications are provided.

1. Introduction

Markov decision processes (MDPs) play a central role in operations research, economics, finance, engineering and computer science (Bertsekas, 2018; Kochenderfer, 2015). In recent years there has been rising interest in extensions to the standard model that can handle sophisticated preference and information structures, such as desire for robustness, risk sensitivity, narrow framing, uncertainty aversion, ambiguity aversion, and separation of atemporal risk aversion and intertemporal substitution (see, e.g., Chen and Sun (2012), Di Masi and Stettner (2007), Epstein and Zin (1989), Ju and Miao (2012), Lin et al. (2018), Ruszczyński (2010), Shen et al. (2013), or Bäuerle and Jaśkiewicz (2018)).

Under most of these extensions, aggregation of rewards over time becomes nonlinear, and the standard contractivity condition, on which the traditional theory rests, no longer holds.[1] In such settings, either optimality theory is lacking or the conclusions are degraded relative to the contractive case. To alleviate these shortcomings, we explore an alternative approach and provide a new set of conditions based around monotonicity and either convexity or concavity. We show that their implications are as strong as the contractive case. We also show that these conditions are satisfied in a range of models that fail to be contractive.[2]

Our study builds on earlier work analyzing growth models with recursive utility, which used monotonicity and concavity properties to show that the Bellman operator has a unique and globally attracting solution within a given class (Bloise & Vailakis, 2018; Marinacci & Montrucchio, 2010). We extend these ideas by providing a full set of optimality results, including identification of the value function with the unique fixed point of the Bellman operator, existence of optimal policies, and the validity of Bellman's principle of optimality. This is achieved by exploiting a fixed point theorem for monotone operators due to Du (1990).

2. General results

2.1. Preliminaries

Let $\mathbb{R}^{\mathsf{X}}$ be all functions from some metric space X to $\mathbb{R}$, let $b\mathsf{X}$ be the bounded Borel measurable functions in $\mathbb{R}^{\mathsf{X}}$ and let $bc\mathsf{X}$ be the continuous functions in $b\mathsf{X}$. Let $\|\cdot\|$ denote the supremum norm on $b\mathsf{X}$. For f and g in $\mathbb{R}^{\mathsf{X}}$, the statement $f \leqslant g$ means $f(x) \leqslant g(x)$ for all $x \in \mathsf{X}$, while $f \ll g$ means that $f \leqslant g - \varepsilon$ for some positive constant ε. Given $a, b \in \mathcal{F} \subset b\mathsf{X}$, the order interval $I := [a, b]$ is all f in $\mathcal{F}$ with $a \leqslant f \leqslant b$. We call $S: I \to I$ *geometrically stable* on I if S has a unique fixed point v^* in I and, for each $v \in I$, we can find constants $\lambda \in (0, 1)$ and $K \in \mathbb{R}$ such that $\|S^n v - v^*\| \leqslant \lambda^n K$ for all $n \in \mathbb{N}$. S is called *monotone increasing* if $Sv \leqslant Sv'$ whenever $v, v' \in I$ with $v \leqslant v'$; and *convex* if $S(\lambda v + (1-\lambda)v') \leqslant \lambda Sv + (1-\lambda)Sv'$ whenever $v, v' \in I$ and

☆ Financial support from ARC, Australia grant FT160100423 is gratefully acknowledged. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Valery Ugrinovskii under the direction of Editor André L. Tits.

* Corresponding author.
E-mail addresses: guanlong_ren89@163.com (G. Ren), john.stachurski@anu.edu.au (J. Stachurski).

[1] Chapter 2 of Bertsekas (2018) provides an exposition of the standard theory, while Bloise and Vailakis (2018) and Marinacci and Montrucchio (2010) discuss failure of contractivity in recursive preference models.

[2] It is worth noting that standard additively separable MDPs also satisfy our conditions. Hence our conditions subsume optimality theory for both standard and more sophisticated models.

https://doi.org/10.1016/j.automatica.2021.109641
0005-1098/

$0 \leqslant \lambda \leqslant 1$. S is called *concave* if $-S$ is convex. We will use a theorem of Du (1990) that, specialized to the current setting, states the following:

Theorem 1 (*Du*)**.** *Let $I := [a, b]$ be an order interval in either $b\mathsf{X}$ or $bc\mathsf{X}$ and let $S: I \to I$ be monotone increasing. If either (i) S is convex on I and $Sb \ll b$, or (ii) S is concave on I and $Sa \gg a$, then S is geometrically stable on I.*

2.2. A dynamic decision problem

Let X and A be metric spaces, called the *state* and *action space* respectively. Let Γ be a correspondence from X to A, called the *feasible correspondence*, with $\Gamma(x)$ represents actions available to the controller in state x. We call $\mathbb{G} := \{(x, a) \in \mathsf{X} \times \mathsf{A} : a \in \Gamma(x)\}$ the *feasible state–action pairs*. A *state–action aggregator* H maps feasible state–action pairs (x, a) and functions v in $b\mathsf{X}$ into real values $H(x, a, v)$ representing lifetime rewards, contingent on current action a, current state x and the use of v to evaluate future states. Traditional additively separable MDPs are implemented by setting

$$H(x, a, v) = r(x, a) + \beta \int v(x')P(x, a, \mathrm{d}x') \tag{1}$$

for some discount factor β, reward function r and transition function P.[3] More sophisticated applications are discussed in Section 3.

Fix w_1, w_2 in $bc\mathsf{X}$ with $w_1 \leqslant w_2$ and set $\mathcal{V} := [w_1, w_2]$ in $b\mathsf{X}$. Let $\mathcal{C}$ be the continuous functions in $\mathcal{V}$. We assume that **(A1)** the feasible correspondence Γ is nonempty, compact valued and continuous, **(A2)** the map $(x, a) \mapsto H(x, a, v)$ is Borel measurable on $\mathbb{G}$ whenever $v \in \mathcal{V}$ and continuous on $\mathbb{G}$ whenever $v \in \mathcal{C}$, and, for all (x, a) in $\mathbb{G}$, **(A3)** the state–action aggregator satisfies $H(x, a, v) \leqslant H(x, a, v')$ whenever $v \leqslant v'$ and **(A4)** the functions w_1 and w_2 satisfy $w_1(x) \leqslant H(x, a, w_1)$ and $H(x, a, w_2) \leqslant w_2(x)$.

If X and A are discrete we adopt the discrete topology, so (A1)–(A2) are always satisfied when $\Gamma(x)$ is finite for each x. (A3) is the monotonicity condition of Bertsekas (2018), which is standard, while (A4) allows w_1 and w_2 to act as lower and upper bounds for lifetime value.

We call H *value-convex* if, for all $(x, a) \in \mathbb{G}$, $\lambda \in [0, 1]$ and v, w in $\mathcal{V}$, we have $H(x, a, \lambda v + (1 - \lambda)w) \leqslant \lambda H(x, a, v) + (1 - \lambda)H(x, a, w)$. We call H *value-concave* if $-H$ is value-convex. The next two assumptions are used when maximizing and minimizing respectively:

Assumption 2.1 (*Convex Program*)**.** H is value-convex and there exists an $\varepsilon > 0$ such that $H(x, a, w_2) \leqslant w_2(x) - \varepsilon$ for all $(x, a) \in \mathbb{G}$.

Assumption 2.2 (*Concave Program*)**.** H is value-concave and there exists an $\varepsilon > 0$ such that $H(x, a, w_1) \geqslant w_1(x) + \varepsilon$ for all $(x, a) \in \mathbb{G}$.

2.3. Policies

Let Σ be all maps from X to A such that each $\sigma \in \Sigma$ is Borel measurable and satisfies $\sigma(x) \in \Gamma(x)$ for all $x \in \mathsf{X}$. We call Σ the *feasible policies*. For each $\sigma \in \Sigma$, we define the σ*-value operator* T_σ on $\mathcal{V}$ by

$$T_\sigma v(x) := H(x, \sigma(x), v) \qquad (x \in \mathsf{X},\ v \in \mathcal{V}). \tag{2}$$

It follows from assumptions (A2) and (A4) that each T_σ is a well defined self-map on $\mathcal{V}$. A fixed point $v_\sigma \in \mathcal{V}$ of T_σ is called a σ*-value function*.

[3] See Bertsekas (2018) for details. Additive separability refers to the fact that current rewards and continuation values are combined by addition.

Proposition 2. *If either Assumptions 2.1 or 2.2 holds, then T_σ is geometrically stable on $\mathcal{V}$ for each σ in Σ.*

Proof. Fix $\sigma \in \Sigma$. The map T_σ is monotone increasing function from $\mathcal{V}$ to itself by (A3) and (A4). If Assumption 2.1 holds, then T_σ is a convex operator on $\mathcal{V}$, as follows immediately from the definitions of T_σ and value-convexity of H. From Assumption 2.1 we also have $T_\sigma w_2 \ll w_2$, so Du's Theorem applies and the claim is confirmed. If Assumption 2.2 holds, then similar arguments show that T_σ is concave and satisfies $Tw_1 \gg w_1$. Again, Du's Theorem applies.

It follows from Proposition 2 that, for each $\sigma \in \Sigma$, the set $\mathcal{V}$ contains exactly one σ-value function v_σ. The value $v_\sigma(x)$ can be interpreted as the lifetime value of following policy σ over an infinite horizon.[4]

2.4. Maximization

With Assumption 2.1 in force, a policy $\sigma^* \in \Sigma$ is called *optimal* if $v_{\sigma^*}(x) \geqslant v_\sigma(x)$ for all $\sigma \in \Sigma$ and all $x \in \mathsf{X}$. The *value function* is defined at $x \in \mathsf{X}$ by $v^*(x) = \sup_{\sigma \in \Sigma} v_\sigma(x)$. Clearly $w_1 \leqslant v^* \leqslant w_2$. A function $v \in \mathcal{V}$ is said to satisfy the *Bellman equation* if

$$v(x) = \max_{a \in \Gamma(x)} H(x, a, v) \quad \text{for all } x \in \mathsf{X}. \tag{3}$$

Given $v \in \mathcal{C}$, a policy σ in Σ is called v*-greedy* if $\sigma(x) \in \operatorname{argmax}_{a \in \Gamma(x)} H(x, a, v)$ for all $x \in \mathsf{X}$. The *Bellman operator* T is a map sending v in $\mathcal{C}$ into

$$Tv(x) = \max_{a \in \Gamma(x)} H(x, a, v). \tag{4}$$

Existence of the maximum is guaranteed by (A1)–(A2).

2.5. Minimization

In the minimization setting, a policy $\sigma^* \in \Sigma$ is called *optimal* if $v_{\sigma^*}(x) \leqslant v_\sigma(x)$ for all $\sigma \in \Sigma$ and $x \in \mathsf{X}$. The *value function* associated with this planning problem is the function v^* defined at $x \in \mathsf{X}$ by $v^*(x) = \inf_{\sigma \in \Sigma} v_\sigma(x)$. A function $v \in \mathcal{V}$ is said to satisfy the *Bellman equation* if (3) holds with max replaced by min. The *Bellman operator* is defined by (4), after replacing max with min, while the definition of a v-greedy policy is as above, after swapping argmax for argmin.

2.6. Main result

We can now state our main result. In stating it, we will say that *Bellman's principle of optimality* holds if the set of optimal policies in Σ coincides with the v^*-greedy policies.

Theorem 3. *If Assumption 2.1 holds (maximization case) or Assumption 2.2 holds (minimization case), then*

(a) The Bellman operator is geometrically stable on $\mathcal{C}$.
(b) The Bellman equation has exactly one solution in $\mathcal{C}$ and that solution is v^*.
(c) Bellman's principle of optimality holds and at least one optimal policy exists.

[4] See, for example, Bertsekas (2018). While we focus here on stationary Markov policies, in the sense that each policy σ depends only on the current state and is invariant over time, it can be shown that, under the full set of assumptions introduced below, the resulting values weakly dominate the values obtained by optimizing with respect to the class of all nonstationary policies. The arguments are almost identical to those presented in the discussion of nonstationary policies in Section 2.1 of Bertsekas (2018) and the details are omitted.

Proof. We begin with the maximization case, holding Assumption 2.1 true. Regarding (a), our aim is to apply Du's Theorem. T is a self-map on $\mathcal{C}$ by (A1)–(A2) and Berge's theorem of the maximum. It remains to show that T is monotone increasing and convex on $\mathcal{C}$ with $Tw_2 \ll w_2$. The monotonicity of T on $\mathcal{C}$ is immediate from (A3), which yields $\max_{a\in\Gamma(x)} H(x,a,v) \leqslant \max_{a\in\Gamma(x)} H(x,a,v')$ for all $x \in \mathsf{X}$ whenever $v \leqslant v'$. To show convexity of T, fix $v, v' \in \mathcal{C}$ and $\lambda \in [0,1]$. For any given $(x,a) \in \mathbb{G}$, we have, by value-convexity,

$$H(x,a,\lambda v + (1-\lambda)v') \leqslant \lambda H(x,a,v) + (1-\lambda)H(x,a,v') \leqslant \lambda Tv(x) + (1-\lambda)Tv'(x).$$

Since $(x,a) \in \mathbb{G}$ was arbitrary, the above inequality implies $\max_{a\in\Gamma(x)} H(x,a,\lambda v + (1-\lambda)v') \leqslant \lambda Tv(x) + (1-\lambda)Tv'(x)$ for each $x \in \mathsf{X}$, which in turn means that $T[\lambda v + (1-\lambda)v'] \leqslant \lambda Tv + (1-\lambda)Tv'$. We also have $Tw_2 \ll w_2$, since, by Assumption 2.1, there is an $\varepsilon > 0$ such that, for each $x \in \mathsf{X}$, we have $Tw_2(x) = \max_{a\in\Gamma(x)} H(x,a,w_2) \leqslant w_2(x) - \varepsilon$. The proof of (a) is now complete.

For the proof of (b) in the maximization case, let v^* be the value function and let $\bar v$ be the unique fixed point of T in $\mathcal{C}$. To see that $\bar v = v^*$, first observe that $\bar v \in \mathcal{C}$ and hence a $\bar v$-greedy policy σ exists. For this policy we have, by definition, $T_\sigma \bar v(x) = T\bar v(x)$ at each x, from which it follows that $\bar v = T\bar v = T_\sigma \bar v$. Since T_σ is geometrically stable on $\mathcal{V}$, we know that its unique fixed point is v_σ, so $\bar v = v_\sigma$. But then $\bar v \leqslant v^*$, by the definition of v^*. To see that the reverse inequality holds, pick any $\sigma \in \Sigma$. We have $T_\sigma \bar v \leqslant T\bar v = \bar v$. Iterating on this inequality and using the monotonicity of T_σ gives $T_\sigma^k \bar v \leqslant \bar v$ for all k. Taking the limit with respect to k and using the stability of T_σ again gives $v_\sigma \leqslant \bar v$. Hence $v^* \leqslant \bar v$, and we can now conclude that $\bar v = v^*$.

Since $\bar v \in \mathcal{C}$, we have $v^* \in \mathcal{C}$. It follows that v^* is the unique solution to the Bellman maximization equation in $\mathcal{C}$. Part (b) of Theorem 3 is now established. Regarding part (c), by the definition of greedy policies and the value function v^*, we have that σ is v^*-greedyv^* if and only if $H(x,\sigma(x),v^*) = v^*(x)$ for all $x \in \mathsf{X}$. By Proposition 2, the second statement is equivalent to $v^* = v_\sigma$. Hence, by this chain of logic and the definition of optimality, σ is v^*-greedy $\iff v^* = v_\sigma \iff \sigma$ is optimal. Moreover, the fact that v^* is in $\mathcal{C}$ assures us that at least one v^*-greedy policy exists. Each such policy is optimal, so the set of optimal policies is nonempty.

The above reasoning completes the proof of the maximization case. The minimization results can be proved from the maximization results and the fact that $-f$ is convex whenever f is concave. (This is why the minimization case requires concavity rather than convexity in Assumption 2.2.) Full details can be found in the online supplement (Ren & Stachurski, 2020).

3. Applications

Theorem 3 can be applied to a range of discrete time dynamic programs that fail to satisfy the standard contractivity conditions. Examples include dynamic programs with Epstein–Zin preferences, ambiguity aversion, and narrow framing. Two examples are now given.

3.1. Epstein–Zin preferences

Epstein–Zin preferences provide the ability to separately control preferences over atemporal risk aversion and intertemporal substitution. The have been applied to a diverse set of problems, including asset pricing, fiscal and monetary policy, resource management and epidemiology (see, e.g., Bansal and Yaron (2004), Epstein and Zin (1989), or Augeraud-Véron et al. (2020)). Optimality results are challenging, since, under empirically plausible parameterizations, these preferences fail to satisfy contractivity.

Under Epstein–Zin preferences, the Bellman equation takes the form

$$v(x) = \max_{a\in\Gamma(x)} \{r(x,a)^\kappa + \beta[Rv(x,a)]^\kappa\}^{1/\kappa} \tag{5}$$

where R is the *certainty equivalent operator* defined by

$$Rv(x,a) := \left[\int v(x')^\eta P(x,a,\mathrm{d}x')\right]^{1/\eta}. \tag{6}$$

The expression in (6) matches the continuation value on the right hand side of (1) when $\eta = 1$. Under these preferences, κ governs elasticity of intertemporal substitution and γ governs atemporal risk aversion. We focus on the most empirically relevant case, which is $\eta < 0 < \kappa < 1$.[5]

It is convenient to apply the transformation $\hat v := v^\eta$ to the Bellman equation (5). Since $\eta < 0$, this leads to the minimization problem $\hat v(x) = \min_{a\in\Gamma(x)} H(x,a,\hat v)$ where

$$H(x,a,v) = \left\{ r(x,a)^\kappa + \beta\left[\int \hat v(x')P(x,a,\mathrm{d}x')\right]^{1/\theta}\right\}^\theta$$

with $\theta := \eta/\kappa < 0$. We assume that $m := \inf r(x,a)^\kappa > 0$ and $M := \sup r(x,a)^\kappa < \infty$. The feasible correspondence is assumed to satisfy (A1). Under mild conditions on the primitives (A2) also holds. (A3) is clearly satisfied. After setting set $w_1 := [(M+\delta)/(1-\beta)]^\theta$ and $w_2 := [m/(1-\beta)]^\theta$, where δ is a positive constant, straightforward manipulations show that (A4) holds. For example, we have

$$H(x,a,w_1) \geqslant \left\{ M + \beta\frac{M+\delta}{1-\beta}\right\}^\theta > w_1$$

for any $(x,a) \in \mathbb{G}$. In fact this last bound shows that w_1 also satisfies the strict inequality in Assumption 2.2, so it only remains to check the value-concavity of H. But this follows directly from the concavity of the function ψ defined for $b, t \geqslant 0$ by $\psi(t) := (b + \beta t^{1/\theta})^\theta$, as implied by $\theta < 0$. Hence the conditions of Theorem 3 hold and its conclusions are all valid.

3.2. Ambiguity aversion

Some recent studies consider ambiguity with respect to the laws of the system on the part of the controller and allow for ambiguity aversion. One foundational study is Klibanoff et al. (2009) and applications to asset pricing and financial decisions can be found in Berger and Eeckhoudt (2020) and Ju and Miao (2012).[6]

A generic version of the problem studied in Ju and Miao (2012) requires minimization with respect to the aggregator

$$H(s,z,a,v) = \{r(s,a,z)^\kappa + \beta[Nv(z,a)]^\kappa\}^{1/\kappa} \tag{7}$$

where

$$Nv(z,a) := \int \left[\int v(a,z')^\xi \pi_\theta(z,\mathrm{d}z')\right]^{\frac{1}{\xi}} \mu(z,\mathrm{d}\theta). \tag{8}$$

Here s and z are state variables, taking values in compact metric spaces, while κ and ξ are composite parameters. (They are a composite of three parameters, which separately control elasticity of intertemporal substitution, atemporal risk aversion and ambiguity aversion.) The transition probability function π_θ is indexed on

[5] See, for example, Schorfheide et al. (2018). Other parameterizations are treated in Ren and Stachurski (2020).

[6] Ambiguity aversion has also found applications in psychology, neuroscience, climate change, management science and other fields. See, e.g., Bayraktar and Zhang (2015) and Trautmann et al. (2011), or Olijslagers and van Wijnbergen (2019).

a vector of parameters $\theta \in \Theta$ that represent model uncertainty, while $\mu(z, \cdot)$ represents beliefs over these parameters conditional on the current exogenous state z. The parameter space Θ is a Borel subset of $\mathbb{R}^k$.

Consider the case where $\xi \in (0, 1)$ and $\kappa < 0$.[7] Since we are minimizing, Assumption 2.2 needs to be verified in order to apply Theorem 3. It can be shown that if we fix $\delta > 0$ and set $w_1 := [(M + \delta)/(1 - \beta)]^{1/\kappa}$ and $w_2 := [m/(1 - \beta)]^{1/\kappa}$, where $m := \inf r(s, a, z)^\kappa > 0$ and $M := \sup r(s, a, z)^\kappa < \infty$, then there exists an $\varepsilon > 0$ such that $H(s, z, a, w_1) \geqslant w_1 + \varepsilon$ for all feasible state–action pairs $((s, z), a)$. In addition, $H(s, z, a, w_2) \leqslant w_2(s, z)$ for all $((s, z), a)$.

The validity of value concavity, which is the remaining part of Assumption 2.2, depends on the parameters κ and ξ. In the supplement (Ren & Stachurski, 2020) we show that value concavity holds at the parameters setting chosen in the empirical component of Ju and Miao (2012). The argument is similar to that provided in Section 3.1.

4. Conclusion

We constructed an optimality theory for discrete time dynamic programs with features such as risk sensitivity, narrow framing, and ambiguity aversion, showing that monotonicity and convexity properties can substitute for the standard contractivity condition assumed in traditional MDPs. Extensions to the continuous time case are left for future research.

Acknowledgments

We thank Dimitri Bersekas and Quentin Batista for valuable comments, as well as the editors and three referees.

References

Augeraud-Véron, E., Fabbri, G., & Schubert, K. (2020). *Prevention and Mitigation of Epidemics: Biodiversity Conservation and Confinement Policies*: *Technical report CESifo Working Paper No. 8506*.

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, *59*(4), 1481–1509.

Bäuerle, N., & Jaśkiewicz, A. (2018). Stochastic optimal growth model with risk sensitive preferences. *Journal of Economic Theory*, *173*, 181–200.

Bayraktar, E., & Zhang, Y. (2015). Minimizing the probability of lifetime ruin under ambiguity aversion. *SIAM Journal on Control and Optimization*, *53*(1), 58–90.

Berger, L., & Eeckhoudt, L. (2020). *Risk, Ambiguity, And The Value Of Diversification*: *Technical report*, HAL archives.

Bertsekas, D. P. (2018). *Abstract dynamic programming* (2nd ed.). Athena Scientific.

Bloise, G., & Vailakis, Y. (2018). Convex dynamic programming with (bounded) recursive utility. *Journal of Economic Theory*, *173*, 118–141.

Chen, X., & Sun, P. (2012). Optimal structural policies for ambiguity and risk averse inventory and pricing models. *SIAM Journal on Control and Optimization*, *50*(1), 133–146.

Di Masi, G. B., & Stettner, Ł. (2007). Infinite horizon risk sensitive control of discrete time Markov processes under minorization property. *SIAM Journal on Control and Optimization*, *46*(1), 231–252.

Du, Y. (1990). Fixed points of increasing operators in ordered banach spaces and applications. *Applicable Analysis*, *38*(2), 1–20.

Epstein, L. G., & Zin, S. E. (1989). Risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, *57*(4), 937–969.

Ju, N., & Miao, J. (2012). Ambiguity, learning, and asset returns. *Econometrica*, *80*(2), 559–591.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory*, *144*(3), 930–976.

Kochenderfer, M. J. (2015). *Decision making under uncertainty: theory and application*. MIT press.

Lin, K., Jie, C., & Marcus, S. I. (2018). Probabilistically distorted risk-sensitive infinite-horizon dynamic programming. *Automatica*, *97*, 1–6.

Marinacci, M., & Montrucchio, L. (2010). Unique solutions for stochastic recursive utilities. *Journal of Economic Theory*, *145*(5), 1776–1804.

Olijslagers, S., & van Wijnbergen, S. (2019). *Discounting the future: on climate change, ambiguity aversion and Epstein-Zin preferences*: *Technical report*, CEPR Discussion Paper No. DP13708.

Ren, G., & Stachurski, J. (2020). Dynamic programming with recursive preferences: Optimality and applications (arxiv supplement). arXiv:1812.05748.

Ruszczyński, A. (2010). Risk-averse dynamic programming for Markov decision processes. *Mathematical Programming*, *125*, 235–261.

Schorfheide, F., Song, D., & Yaron, A. (2018). Identifying long-run risks: A Bayesian mixed-frequency approach. *Econometrica*, *86*(2), 617–654.

Shen, Y., Stannat, W., & Obermayer, K. (2013). Risk-sensitive Markov control processes. *SIAM Journal on Control and Optimization*, *51*(5), 3652–3672.

Trautmann, S. T., Vieider, F. M., & Wakker, P. P. (2011). Preference reversals for ambiguity aversion. *Management Science*, *57*(7), 1320–1333.

[7] Other cases are treated in Ren and Stachurski (2020).