# CONVERGENCE OF A NUMERICAL SCHEME FOR OPTIMAL STOPPING OVER A FINITE TIME-HORIZON OF A DIFFUSION WITH SINGULAR BEHAVIOR

by

Steffen Domke

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Mathematics

at

The University of Wisconsin–Milwaukee

August 2025

## ABSTRACT

CONVERGENCE OF A NUMERICAL SCHEME FOR OPTIMAL STOPPING OVER A FINITE TIME-HORIZON OF A DIFFUSION WITH SINGULAR BEHAVIOR

by

Steffen Domke

The University of Wisconsin–Milwaukee, 2025
Under the Supervision of Professor Richard H. Stockbridge

This dissertation establishes an approximation scheme for finite time-horizon stopping problems involving a singular stochastic process on a compact state space, characterized by a singular martingale problem. The stopping problem is converted to a linear program (LP) with infinitely many constraints and variables having infinite degrees of freedom.

To obtain a numerical solution, the infinite-dimensional LP is converted into a finite LP. The original LP is approximated by a sequence of finite LPs, limiting to both a finite set of constraints and a finite-dimensional solution space. The value of an optimal approximate solution is shown to be arbitrarily close to the optimal value of original LP, and hence of the stopping probem, with increasing refinement of the approximation. Feasibility of the approximate solutions is guaranteed due weak convergence of measures, but only in the limit. The problem of pricing an American floating strike lookback call option can be reformulated to fit the models covered by this dissertation. The price and the stopping boundary can therefore be approximated using this scheme.

I would like to express my deepest gratitude to my advisor Prof. Richard Stockbridge for his guidance, dedication, and his professional and personal support in the course of this Ph.D. project. I learned so much about probability theory, research and project management from his teachings, his advice and his example. Thank you.

Many thanks to the members of my doctoral committee, Professors David Spade, Lei Wang, Jeb Willenbring, and Chao Zhu.

I would like to extend a special thank you to Prof. Vincent Larson for providing a research assistantship during my time in Milwaukee, as well as many hours of interesting conversations and personal support.

I am very grateful to Professor Gerhard Dikta and the Mathematical Sciences department at UWM for the opportunity to achieve my Master's and doctoral degrees.

And last, but not least, I am forever indebted to Gitte Kremling and Kristin Maul, and especially my parents, Gertrud and Manfred Domke, for their continued friendship and support. I would not be where I am without you. Thank you.

Steffen Domke

# TABLE OF CONTENTS

# LIST OF FIGURES

# 1 INTRODUCTION

Many processes of interest in applications (see, for example, the survey paper by Shreve (1988)) can be modelled as solutions to a stochastic differential equation of the form

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t) + m(X(t-))d\xi(t) \quad (1.1)$$

where $X$ is the state process with $E = \mathbb{R}^d$, $\xi$ is a nondecreasing process arising from the boundary behavior of $X$ (e.g., the local time on the boundary for a reflecting diffusion), and $W$ is a Brownian motion. Processes in which the set of times of increase of $\xi$ has Lebesgue measure 0 are called singular stochastic processes. This dissertation considers optimal stopping problems involving singular processes.

The purpose of the dissertation is to develop approximate numerical solution to these optimal stopping problems over a finite time horizon. It builds upon the formulation and results in Kurtz and Stockbridge (2001), in which the singular stochastic process is characterized by a singular martingale problem for its generators and is equivalently characterized by a system of linear equations over spaces of measures.

The use of occupation measures to solve optimization problems began with Young (1933) in the context of calculus of variations. The present work is an outgrowth of linear programming formulations for stochastic control that was initiated by Manne (1960) in discrete time and which has been developed for Markov decision problems (see e.g. Hernandez-Lerma and Lasserre (1996), Hernández-Lerma and Lasserre (1998), and Hernandez-Lerma and Lasserre (1999)). Linear programming for the optimal stopping of stochastic processes in continuous time has been established under very general conditions in Cho and Stockbridge (2002) for processes only having abolutely continuous behavior in time, and for processes having singular behavior in Helmes and Stockbridge (2007).

The resulting linear programs are typically infinite-dimensional whose variables are the occupation measures of the processes according to both regular time and singular time and the distribution of the state at the stopping time. This dissertation establishes the convergence of the solutions of a sequence of approximate finite-dimensional LPs to the optimal solution of the infinite-dimensional problem.

This dissertation considers an optimal stopping problem of the form

$$\mathbb{E}\left[e^{-\lambda\tau}\left(X(\tau) - Y(\tau)\right)\right] \tag{1.2}$$

in which the state process $X$ denotes the price of an asset, $Y$ denotes the running minimum process of the asset price, $\lambda > 0$ gives the discount rate and $\tau$ denotes the option holder's exercise time. The quantity used for illustration is an American lookback call option. This problem is very similar to the Russian option studied by Shepp and Shiryaev (1995), Graversen and Peskir (1998) and Peskir (2005). Chapter 3 shows how this specific stopping problem can be adapted to the form required for the scheme discussed in this dissertation.

The next section formulates the singular martingale problem and highlights the equivalence of the linear programming formulation with the original stochastic process formulation for the optimal stopping problem. This will be followed by a section giving an introduction to weak convergence of finite measures, the mode of convergence that will be necessary in Chapter 2 to show convergence of solutions. This chapter concludes with a description of the approximation techniques that are applied in this dissertation, introducing the grid of points and the finite-dimensional function spaces employed to reduce the number of constraint equations and the dimensionality of the solution spaces of measures.

The main body of the dissertation is in Chapter 2, in which we present proofs of convergence for the solutions of the finite-dimensional linear program to solutions of the infinite-dimensional linear program, modifying the approach that Vieten (2018) devised for optimal control problems. We first introduce a semi-infinite linear program and discuss convergence of values and solutions to the corresponding quantities for the infinite LP. Then we approximate the semi-infinite linear program by the finite-dimensional LP and, again, discuss convergence of values and solutions to the quantities for the semi-infinite LP.

In Chapter 3, we introduce a formulation for the American floating strike lookback option that matches numerical scheme can be applied to it. A similar option, called the "Russian option", was first devised by Shepp and Shiryaev (1993) for an infinite horizon and translated to finite time horizon by Peskir (2005). Lutz (2007) provided a first, similar numerical appproach to American floating strike lookback options without a proof of convergence. Numerical solutions are not presented in this dissertation, but are part of future work.

Finally, Chapter 4 lists a number of directions the presented research offers for future work, including numerical results for the example introduced previously in several ways.

## 1.1 Stochastic Formulation and LP Reformulation

We utilize the formulation and results of Helmes and Stockbridge (2007), which we will now present. Note that, in addition to the inherited assumptions, our approximation results require the state space $E$ to be compact.

### 1.1.1 Formulation of the Martingale Problem

For a complete, separable, metric space $S$, we define $M(S)$ to be the space of Borel measurable functions on $S$, $B(S)$ to be the space of bounded, measurable functions on $S$, $C(S)$ to be the space of continuous functions on $S$, $\overline{C}(S)$ to be the space of bounded, continuous functions on $S$, $\mathcal{M}(S)$ to be the space of finite Borel measures on $S$, and $\mathcal{P}(S)$ to be the space of probability measures on $S$. $\mathcal{M}(S)$ and $\mathcal{P}(S)$ are topologized by weak convergence.

Let $\mathcal{L}_t(S) = \mathcal{M}(S \times [0,t])$. We define $\mathcal{L}(S)$ to be the space of measures $\xi$ on $S \times [0,\infty)$ such that $\xi(S \times [0,t]) < \infty$, for each $t$, and topologized so that $\xi_n \to \xi$ if and only if $\int f d\xi_n \to \int f d\xi$, for every $f \in \overline{C}(S \times [0,\infty))$ with $\text{supp}(f) \subset S \times [0, t_f]$ for some $t_f < \infty$. Let $\xi_t \in \mathcal{L}_t(S)$ denote the restriction of $\xi$ to $S \times [0,t]$. Note that a sequence $\{\xi^n\} \subset \mathcal{L}(S)$ converges to a $\xi \in \mathcal{L}(S)$ if and only if there exists a sequence $\{t_k\}$, with $t_k \to \infty$, such that, for each $t_k$, $\xi^n_{t_k}$ converges weakly to $\xi_{t_k}$, which in turn implies $\xi^n_t$ converges weakly to $\xi_t$ for each $t$ satisfying $\xi(S \times \{t\}) = 0$. Finally, we define $\mathcal{L}^{(m)}(S) \subset \mathcal{L}(S)$ to be the set of $\xi$ such that $\xi(S \times [0,t]) = t$ for each $t > 0$. Throughout, we will assume that the state space $E$ is a complete, separable, metric space.

Let $A, B : \mathcal{D} \subset \overline{C}(E) \to C(E)$ and $\nu_0 \in \mathcal{P}(E)$. Let $X$ be an $E$-valued process and $\Gamma$ be an $\mathcal{L}(E)$-valued random variable. Let $\Gamma_t$ denote the $\mathcal{L}(E)$-random variable which, for each realization, is defined by restricting $\Gamma$ to $E \times [0,t]$. Then $(X, \Gamma)$ is a solution of the *singular martingale problem* for $(A, B, \nu_0)$ if there exists a filtration $\{\mathcal{F}_t\}$ such that $(X, \Gamma_t)$ is $\{\mathcal{F}_t\}$-progressive, $X(0)$ has distribution $\nu_0$, and for every $f \in \mathcal{D}$,

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))\,ds - \int_{E\times[0,t]} Bf(x)\Gamma(dx \times ds) \qquad (1.3)$$

is an $\{\mathcal{F}_t\}$-martingale.

Note we refer to $A$ as being the absolutely continuous generator of $X$ and $B$ as the singular generator. This is an intuitive labelling. Strictly speaking, the random measure $\Gamma$ may have an absolutely continuous part.

### 1.1.2 Conditions on $A$ and $B$

We assume that the absolutely continuous generator $A$ and the singular generator $B$ have the following properties.

**Condition 1.1.**

*i)* $A, B : \mathcal{D} \subset \overline{C}(E) \to C(E)$, $1 \in \mathcal{D}$, *and* $A1 = 0, B1 = 0$.

*ii) Defining* $(A_0, B_0) = \{(f, Af, Bf) : f \in \mathcal{D}\}$, $(A_0, B_0)$ *is* separable *in the sense that there exists a countable collection* $\{g_k\} \subset \mathcal{D}$ *such that* $(A_0, B_0)$ *is contained in the bounded, pointwise closure of the linear span of* $\{(g_k, Ag_k, Bg_k)\}$.

*iii)* $\mathcal{D}$ *is closed under multiplication and separates points.*

**Example 1.2.** Reflecting diffusion processes.

The most familiar class of processes of the kind we consider are reflecting diffusion processes satisfying equations of the form

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds + \int_0^t m(X(s))d\xi(s),$$

where $X$ is required to remain in the closure of a domain $D$ (assumed smooth for the moment) and $\xi$ increases only when $X$ is on the boundary of $D$. Then

$$Af(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + b(x)\cdot\nabla f(x),$$

where $a(x) = ((a_{ij}(x))) = \sigma(x)\sigma(x)^T$. In addition, under reasonable conditions $\xi$ will be continuous, so

$$Bf(x) = m(x)\cdot\nabla f(x).$$

Suppose that $(X,\Gamma)$ is a solution of the singular martingale problem for $(A, B, \nu_0)$ and let $\tau$ be a stopping time satisfying the conditions of the optional sampling theorem (see Ethier and Kurtz (1986, Theorem 2.2.13)). It follows from $(X,\Gamma)$ satisfying (1.3) that when we augment a time component to the state space by considering the process $(t, X)$,

$$\begin{aligned} \gamma(t)f(X(t)) - \gamma(0)f(X(0)) &- \int_0^t [\gamma(s)Af(X(s)) + \gamma'(s)f(X(s))]\,ds \\ &- \int_{\mathbb{R}^+\times E\times[0,t]} \gamma(s)Bf(x)\Gamma(ds\times dx\times dr) \end{aligned} \tag{1.4}$$

is also an $\{\mathcal{F}_t\}$-martingale for $\gamma \in \widehat{C}(\mathbb{R}^+)$ and $f \in \mathcal{D}$. It then follows from the optional sampling theorem that for each $\gamma \in \widehat{C}(\mathbb{R}^+)$ and $f \in \mathcal{D}$,

$$\begin{aligned} \mathbb{E}\Bigg[\gamma(\tau)f(X(\tau)) - \gamma(0)f(X(0)) &- \int_0^\tau \tilde{A}[\gamma f](s, X(s))\,ds \\ &- \int_{E\times[0,\tau]} \tilde{B}[\gamma f](s,x)\,\Gamma(ds\times dx\times dr)\Bigg] = 0, \end{aligned} \tag{1.5}$$

where $\tilde{A}[\gamma f] = \gamma Af + \gamma' f$ and $\tilde{B}[\gamma f] = \gamma Bf$. Let $\nu_\tau$ denote the distribution of $(\tau, X(\tau))$ and define the expected "occupation measures" on $\mathbb{R}^+ \times E$ by

$$\begin{aligned} \mu_0(G) &= \mathbb{E}\left[\int_0^\tau I_G(s, X(s))\,ds\right], \qquad \forall G \in \mathcal{B}(\mathbb{R}^+\times E), \\ \mu_1(G) &= \mathbb{E}\left[\int_{\mathbb{R}^+\times E\times[0,\tau]} I_G(s,x)\,\Gamma(ds\times dx\times dr)\right], \qquad \forall G \in \mathcal{B}(\mathbb{R}^+\times E). \end{aligned}$$

Notice that $\mu_0$ is the occupation measure of the process according to "regular" time (Lebesgue measure on $\mathbb{R}^+$) and $\mu_1$ is the occupation measure of $X$ according to the singular set of times at which $\Gamma$ increases. We refer to $\mu_1$ as the "singular" occupation measure. It then follows immediately from (1.5) that the measures $\nu_\tau$, $\mu_0$ and $\mu_1$ satisfy

$$\begin{aligned} 0 = \int \gamma(t) f(x)\, \nu_\tau(dx) - \gamma(0) \int f(x)\, \nu_0(dx) \; &- \int \tilde{A}[\gamma f](r,x)\, \mu_0(dr \times dx) \\ &- \int \tilde{B}[\gamma f](x)\, \mu_1(dr \times dx), \quad \forall f \in \mathcal{D}. \end{aligned} \tag{1.6}$$

**Remark 1.3.** For later reference, we observe that for any optimal stopping problem with finite horizon $T$, $\mu_0([0,T] \times E) \leq T$ since $\tau \leq T$.

Thus given any solution $(X, \Gamma)$ to the singular martingale problem for $(A, B)$ and any sufficiently nice $\tau$, the measures $\nu_\tau$, $\mu_0$ and $\mu_1$ will satisfy (1.6).

Let $c_0, c_1, c_2 \in M(\mathbb{R}^+ \times E)$ be measurable and bounded below, and represent the time-dependent running cost of the process according to regular time, the running cost according to the singular time and the stopping cost, respectively. The goal of the decision maker is to select a stopping rule $\tau$ so as to minimize the expect cost of the process up to time $\tau$:

$$\mathbb{E}\left[\int_0^\tau c_0(s, X(s))\, ds + \int_{\mathbb{R}^+ \times E \times [0,\tau)} c_1(s,x)\, \Gamma(ds \times dx \times dv) + c_2(\tau, X(\tau))\right]. \tag{1.7}$$

The following theorem from Helmes and Stockbridge (2007) establishes the equivalence of the optimal stopping problem with the linear program.

**Theorem 1.4.** *The optimal stopping problem of selecting a stopping time $\tau^*$ so as to minimize*

$$\mathbb{E}\left[\int_0^\tau c_0(s, X(s))\, ds + \int_{\mathbb{R}^+ \times E \times [0,\tau)} c_1(s,x)\, \Gamma(dx \times ds \times dv) + c_2(\tau, X(\tau))\right]$$

*subject to $(X, \Gamma)$ being a solution of the singular martingale problem for $(A, B, \nu_0)$ is equiv-*

*alent to the infinite dimensional linear program*

$$\textit{Minimize} \qquad \int c_0\, d\mu_0 + \int c_1\, d\mu_1 + \int c_2\, d\nu_\tau$$

$$\begin{aligned}\textit{Subject to} \qquad & 0 = \int \gamma f\, d\nu_\tau - \gamma(0) \int f\, d\nu_0 - \int \tilde{A}[\gamma f]\, d\mu_0 - \int \tilde{B}[\gamma f]\, d\mu_1, \qquad \forall \gamma f \in \mathcal{D}_1 \\ & \nu_\tau \in \mathcal{P}(\mathbb{R}^+ \times E), \\ & \mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^+ \times E).\end{aligned}$$

**Remark 1.5.** We make the following adjustments to this problem in order to guarantee convergence for our approximations:

1. We consider finite horizon optimal stopping problems with time horizon $T$, and so we require $\tau \leq T$
2. We require the state space $E$ to be compact, and hence the product space $[0, T] \times E$ is also compact.
3. We require that $c_0, c_1, c_2 \in C([0, T] \times E)$.
4. We assume that the collection of local time measures $\{\mu_1\}$ from feasible solutions $(\mu_0, \mu_1, \nu_\tau)$ are uniformly bounded. This is a condition for weak convergence. See Section 1.2.
5. For the construction of a feasible solution to the finite linear program, we additionally require all three measures to have densities with respect to some Lebesgue measure, and the density of the of the stopping measure $\nu_\tau$ to be almost surely strictly positive. This last assumption will aid in the construction of a feasible solution to the finite-dimensional linear program.

**Remark 1.6.** Since the underlying space $[0, T] \times E$ is compact, a sufficient condition for uniform boundedness of the collection $\{\mu_1\}$ mentioned above is that there exists an $f \in$

$C([0,T] \times E)$ satisfying $Bf \equiv 1$. To see this, take $\gamma \equiv 1$. Then, feasibility implies

$$\mu_1([0,T] \times E) = \int f d\nu_\tau - \int f d\nu_0 - \int Af d\mu_0.$$

Each integrand and each measure on the right-hand side of the equality is uniformly bounded, and therefore we can find a uniform upper bound for $\mu_1([0,T] \times E)$.

We are interested in maximizing the optimal payoff. So, we consider $c_0, c_1, c_2$ as reward functions. Those same functions and the test functions $\gamma f$, being continuous on a compact space, are bounded and uniformly continuous. As a consequence, $\overline{C}([0,T] \times E) = C([0,T] \times E)$. We take advantage of this structure to use weak convergence of measures. As a result, by minimizing the negative of the objective function, the Theorem 3.2 also applies for maximization problems, of which our example is one.

## 1.2 Weak Convergence of Measures

In order to show convergence of the solutions of the linear programs and their value functions, we will to employ the concept of weak convergence of measures. An introduction into the topic is provided by Billingsley (1999). In particular, we will work with weak convergence of finite Borel measures over metric spaces as discussed by Bogachev (2007). In this section, let $(S,d)$ be a metric space with Borel $\sigma$-algebra $\mathcal{F}$, giving us the measurable space $(S,\mathcal{F})$ and $\mathcal{M}(S)$ the space of all finite measure on $(S,\mathcal{F})$. In addition, let $C_b(S)$ be the space of bounded, continuous functions from $S$ to $\mathbb{R}$ and $C_b^u(S) \subset C_b(S)$ subspace of bounded, uniformly continuous functions. First, let us define the concept of weak convergence.

**Definition 1.7.** Let $\{\mu_n\}_{n\in\mathbb{N}}$ a sequence of finite measures in $\mathcal{M}(S)$ and $\mu \in \mathcal{M}(S)$. We say that $\mu_n$ *converges weakly* to $\mu$, or $\mu_n \Rightarrow \mu$ as $n \to \infty$, if for all bounded, continuous functions $f \in C_b(S)$

$$\lim_{n\to\infty} \int_S f(x)\mu_n(dx) = \int_S f(x)\mu(dx)$$

In order to utilize weak convergence for solutions of linear programs, we will use the following theorem which connects tightness to weak convergence, in a similar manner to compactness for sequences in topological spaces. In words, a tight measure has most of its mass is concentrated on a compact set. For an in depth discussion see Bogachev (2007).

**Definition 1.8.** A measure $\mu$ on the measurable space $(S, \mathcal{F})$ is called *tight* on $\mathcal{F}$ if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset S$ such that

$$\mu(K_\varepsilon^c) < \varepsilon.$$

We will need to extend this concept to families of measures.

**Definition 1.9.** A family of measures $\{\mu_\alpha\}_{\alpha \in \mathcal{A}}$ on $(S, \mathcal{F})$ is called *uniformly tight* if for every $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset S$ such that for all measures of the family

$$\mu_\alpha(K_\varepsilon^c) < \varepsilon.$$

For the central Theorem of this section we will also need the families of measures to be uniformly bounded.

**Definition 1.10.** A family of finite measures $\{\mu_\alpha\}_{\alpha \in \mathcal{A}}$ on $(S, \mathcal{F})$ is called uniformly bounded if there is a $l \geq 0$ such that

$$\mu_\alpha(S) \leq l \quad \forall \alpha \in \mathcal{A}.$$

The following is the central theorem we will use to show weak convergence later on.

**Theorem 1.11.** *Let $\{\mu_n\}_{n \in \mathbb{N}}$ be a sequence of finite measures on $(S, \mathcal{F})$. Then the following two statements are equivalent.*

1. *$\{\mu_n\}_{n \in \mathbb{N}}$ contains a weakly converging subsequence.*

2. *$\{\mu_n\}_{n \in \mathbb{N}}$ is uniformly tight and uniformly bounded.*

For a proof see Bogachev (2007, Theorem 8.6.2). Note that this, in particular holds for a sequence of probability measures.

## 1.3 Discretization

In this section, we will discuss the steps and assumptions to approximate the infinite-dimensional linear program resulting from an optimal stopping problem to a finite-dimensional LP that can be solved numerically.

Let the time space be the closed bounded interval $[0, T] \subset \mathbb{R}$ and the state space be the closed bounded interval $E = [e_l, e_r] \subset \mathbb{R}$. The domain of the augmented time-space-process is then $[0, T] \times E$. We will denote the set of finite measures on $[0, T] \times E$ as $\mathcal{M}([0, T] \times E)$, and the set of probability measures as $\mathcal{P}([0, T] \times E)$. The objective function depends on the three non-negative continuous payoff functions $c_0, c_1, c_2 : [0, T] \times E \to \mathbb{R}_0^+$, which describe the continuous payoff, the singular payoff, and the payoff at the final time $\tau$, respectively. For the example of the American floating strike lookback call option described in Chapter 3, we will only have a payoff at final time, i.e. $c_0 \equiv 0$ and $c_1 \equiv 0$. For generality purposes, we will keep those functions in the following treatment.

Recall that the general linear program, that we seek to solve is

$$
\begin{aligned}
\text{Maximize} \quad & \int\limits_{[0,T]\times E} c_0(t,x)\,\mu_0(dt \times dx) + \int\limits_{[0,T]\times E} c_1(t,x)\,\mu_1(dt \times dx) + \int\limits_{[0,T]\times E} c_2(t,x)\,\nu_\tau(dt \times dx) \\
\text{Subject to} \quad & Rf = \int\limits_{[0,T]\times E} f(t,x)\,\nu_\tau(dt \times dx) - \int\limits_{[0,T]\times E} \tilde{A}f(t,x)\,\mu_0(dt \times dx) \\
& \qquad - \int\limits_{[0,T]\times E} \tilde{B}f(t,x)\,\mu_1(dt \times dx) \qquad \forall f \in \tilde{\mathcal{D}} \\
& \nu_\tau \in \mathcal{P}([0,T] \times E), \quad \mu_0, \mu_1 \in \mathcal{M}([0,T] \times E),
\end{aligned}
$$

where $\mathcal{D}_\infty \subset C^{(1,2)}([0, T] \times E)$ is the domain of the augmented operators $\tilde{A}$ and $\tilde{B}$.

Since $\mu_1$ and $\tilde{B}$ govern the singular behavior, we can set both to zero in case of a problem without singular points. If the problem includes singular behavior, the singular time measure $\mu_1$ is concentrated on the set of singular behavior. We will assume there to be a time-independent and finite number of singular points $\{s_i\} \subset E$ in state space. In general, this will be a subset of the boundary points of the state space, $\{e_l, e_r\}$, and in our specific example of an American floating strike lookback option, just the upper boundary, $e_r = 1$. Thus, the set of singular behavior is represented by $[0, T] \times \{s_i\}$. For the general infinite-dimensional linear program, we can now define the feasible set

$$\mathcal{M}_\infty = \Bigg\{ (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}([0,T] \times E) \times \mathcal{M}([0,T] \times E) \times \mathcal{P}([0,T] \times E) \Big| \\ Rf = \int f d\nu_\tau - \int Af d\mu_0 - \int Bf d\mu_1 \quad \forall f \in \mathcal{D}_\infty, \\ \mu_0([0,T] \times E) \leq T, \\ \mu_1([0,T] \times E) \leq l \Bigg\}.$$

Recall that we showed in Remark 1.6 based on the structure of the optimal stopping problems we are discussing, that we can find an upper bound $l \in \mathbb{R}^+$ of the singular time measure. The objective function $J : \mathcal{M}([0,T] \times E) \times \mathcal{M}([0,T] \times E) \times \mathcal{P}([0,T] \times E) \to \mathbb{R}_0^+$ is defined by

$$J(\mu_0, \mu_1, \nu_\tau) = \int_{[0,T]\times E} c_0 d\mu_0 + \int_{[0,T]\times E} c_1 d\mu_1 + \int_{[0,T]\times E} c_2 d\nu_\tau.$$

Using these definitions, we can set up the infinite-dimensional linear program.

**Definition 1.12.** The *infinite-dimensional linear program*:
Find

$$\max \left\{ J(\mu_0, \mu_1, \nu_\tau) \,\middle|\, (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_\infty \right\}.$$

We will not be able to solve this linear program numerically, since both the number of constraints and the dimension of the space of variable, i.e. the space of measures we optimize over, are infinite. In this section, we will describe a number of approximations, that will eventually yield a finite-dimensional problem, that is solvable numerically.

In two separate steps, one for the space of test function and the other for the solution space of measures, we will use the existence of a countable basis, which we will then truncate to construct a finite-dimensional space. In Chapter 2, we will then use the finite-dimensional spaces to approximate the original infinite-dimensional linear program by a finite-dimensional linear program, and later show that optimal solutions are sufficiently close.

First, we will limit for the number of constraints, which is governed by the number of dimensions of the space of test functions. And in the second step, we will construct a finite-dimensional space of measures, which we can then use to introduce a finite number of parameters. In this step of the process, we will not specify the bases being used since the choice does not affect the general contruction of the approximations. In the following subsections the specific choices will be introduced.

Consider the infinite-dimensional space of test functions $\mathcal{D}_\infty \subset C^{(1,2)}([0,T] \times E)$. Recall that $\mathcal{D}_\infty$ is separable, i.e. we can find a countable basis. Let $\{f_k\}_{k\in\mathbb{N}}$ be such a countable basis of $\mathcal{D}_\infty$. A possible example is the set of B-spline basis functions.

Let $n \in \mathbb{N}$ and $\{f_k\}_{k\in\mathbb{N}}$ be a countable basis of $\mathcal{D}_\infty$. Now, let $\mathcal{D}_n$ be the $n$-dimensional subspace we get by truncating the basis at $n$

$$\mathcal{D}_n = (\mathrm{span}(f_1, \ldots, f_n), \|\cdot\|_{\mathcal{D}}),$$

and define the feasible set for the $(n, \infty)$-dimensional linear problem

$$\begin{aligned}\mathcal{M}_{(n,\infty)} = \Bigg\{ &(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}([0,T] \times E) \times \mathcal{M}([0,T] \times E) \times \mathcal{P}([0,T] \times E) \Big| \\ &Rf = \int f d\nu_\tau - \int Af d\mu_0 - \int Bf d\mu_1 \quad \forall f \in \mathcal{D}_n, \\ &\mu_0([0,T] \times E) \leq T, \\ &\mu_1([0,T] \times E) \leq l \Bigg\}.\end{aligned}$$

Note that the number of constraints is still infinite in this formulation. However, we can rewrite the definition using the linearity of the integral and the operators, thus limiting the number of constraints to $n$, one for each basis function

$$\mathcal{M}_{(n,\infty)} = \Big\{ (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}([0,T] \times E) \times \mathcal{M}([0,T] \times E) \times \mathcal{P}([0,T] \times E) \Big| \\ Rf_k = \int f_k d\nu_\tau - \int Af_k d\mu_0 - \int Bf_k d\mu_1 \quad k = 1, \dots, n, \\ \mu_0([0,T] \times E) \leq T, \\ \mu_1([0,T] \times E) \leq l \Big\}.$$

The corresponding linear program is then

**Definition 1.13.** The *semi-infinite* $(n, \infty)$*-dimensional linear program* is to find

$$\max \left\{ J\left(\mu_0, \mu_1, \nu_\tau\right) \big| \left(\mu_0, \mu_1, \nu_\tau\right) \in \mathcal{M}_{(n,\infty)} \right\}.$$

We will use this definition to show in Chapter 2, that an $\varepsilon$-optimal solution is still close enough for the original infinite-dimensional linear program.

In the second step, we will now approximate the space of measures in order to find a finite set of optimization parameters. As we already stated in Remark 1.5, we assume that all three measures $\mu_0$, $\mu_1$, and $\nu_\tau$ are absolutely continuous with respect to the Lebesgue measure on the appropriate space. Thus, we can define $L^1$-densities for each, $p_0 \in L^1([0,T] \times E)$, $p_1 \in L^1(S)$, and $p_\tau \in L^1(l)$, where $l \subset [0,T] \times E$ is the stopping boundary and $S$ is the set of singular behavior.

In order to find suitable finite-dimensional subspaces, we will now construct a common grid which can be used to define finite-dimensional subspaces of these $L^1$-spaces. Let $m \in \mathbb{N}$ be a fixed number. For $m$, construct a dyadic set of grid points.

$$T^{(m)} = \{t_i = 0 + \frac{T - 0}{2^m} \cdot i, \quad i = 0, \dots, 2^m\},$$
$$E^{(m)} = \{e_j = e_l + \frac{e_r - e_l}{2^m} \cdot j, \quad j = 0, \dots, 2^m\},$$
$$G^{(m)} = \{(t_i, e_j), \quad i, j = 0, \dots, 2^m\}.$$

Note that $|G^{(m)}| = (2^m + 1)^2$, for $m' < m''$, $G^{(m')} \subset G^{(m'')}$ and the union over all $m$ is a dense subset of $[0,T] \times E$.

The approximations of the measures will then be linear combinations of basis functions defined on this grid

$$\tilde{p}_\alpha(t,x) = \sum_{i=0}^{4^m-1} \gamma_{\alpha,i} p_i(t,x), \quad \alpha = 0, \tau,$$

where $\mathfrak{P} = \{p_i\}_{i\in\mathbb{N}}$ is the countable basis for $L^1([0,T]\times E)$ and $\mathfrak{P}^{(m)} = \{p_i\}_{i=0,\dots,4^m-1}$ is the truncated set of basis basis functions. Let us call the sets of all measures with such densities

$$\mathcal{M}_0^{(m)}([0,T]\times E) = \left\{\mu \in \mathcal{M}([0,T]\times E) \middle| \mu(dt\times dx) = \tilde{p}_0(t,x)\lambda(dt\times dx), \mu([0,T]\times E) \le T\right\}$$
$$\mathcal{M}_1^{(m)}([0,T]\times E) = \left\{\mu \in \mathcal{M}([0,T]\times E) \middle| \mu(dt\times dx) = \tilde{p}_1(t,x)\lambda(dt\times dx), \mu([0,T]\times E) \le l\right\}$$
$$\mathcal{M}_\tau^{(m)}([0,T]\times E) = \left\{\nu \in \mathcal{M}([0,T]\times E) \middle| \nu(dt\times dx) = \tilde{p}_\tau(t,x)\lambda(dt\times dx), \nu([0,T]\times E) = 1\right\}.$$

To account for the simple strategy of "Stop immediately", we add the point mass at $(0, x_0)$ to the set of measures $\mathcal{M}_\tau^{(m)}$. This also guarantees that there exists at least one feasible element. Based on this method of approximation, a straight forward definition of the feasible set for the $(n,m)$-dimensional linear program is

$$\mathcal{M}'_{(n,m)} = \Bigg\{ (\mu_0,\mu_1,\nu_\tau) \in \mathcal{M}_0^{(m)}([0,T]\times E) \times \mathcal{M}_1^{(m)}([0,T]\times E) \times \mathcal{M}_\tau^{(m)}([0,T]\times E) \Big| \\ Rf_k = \int f_k d\nu_\tau - \int Af_k d\mu_0 - \int Bf_k d\mu_1, \quad k = 1,\dots,n \Bigg\}.$$

However, in order to construct a feasible approximation for an arbitrary triple $(\mu_{0,n}, \mu_{1,n}, \nu_{\tau,n})$, we need to modify the constraint equations. Let $f_k$ be any basis test function corresponding to one of the constraint equations of the semi-infinite LP and $G^{(m)}$ be the grid of points $(t_i, e_j)$ at refinement level $m$. Let $\Pi_m = \{R_i\}$ be the partition of rectangle using these grid points. Then define

$$f_{k,i}^{(m)} = \inf_{(t,x)\in R_i}\{f_k(t,x)\}, \qquad k = 1,\dots,n, \quad i = 0,\dots,4^m-1,$$
$$\widehat{f}_k^{(m)} = \sum_{i=0}^{4^m-1} f_{k,i}^{(m)} \mathbb{1}_{R_i}, \qquad k = 1,\dots,n, \tag{1.8}$$

and use this piecewise constant approximation of the test functions to define the modified set of feasible solutions to the $(n,m)$-dimensional linear program

$$\mathcal{M}_{(n,m)} = \Big\{ (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_0^{(m)}([0,T]\times E) \times \mathcal{M}_1^{(m)}([0,T]\times E) \times \mathcal{M}_\tau^{(m)}([0,T]\times E) \Big| \\ Rf_k = \int \widehat{f}_k^{(m)} d\nu_\tau - \int Af_k d\mu_0 - \int Bf_k d\mu_1, \quad k = 1,\dots,n \Big\}.$$

Since, for any sequence of solutions $\{(\mu_{0,m}, \mu_{1,m}, \nu_{\tau,m})\}_{m\in\mathbb{N}}$ with $(\mu_{0,m}, \mu_{1,m}, \nu_{\tau,m}) \in \mathcal{M}_{(n,m)}$, the sequence of functions $\{\widehat{f}_k^{(m)}\}_{m\in\mathbb{N}}$ is uniformly bounded and therefore uniformly $\{\nu_{\tau,m}\}$-integrable, we can apply Theorem 3.5 in Serfozo (1982) to prove that if the sequence $\{(\mu_{0,m}, \mu_{1,m}, \nu_{\tau,m})\}_{m\in\mathbb{N}}$ converges weakly, the weak limits are solutions to the semi-infinite linear program. The detailed proof can be found in Section 2.2.

Note that the upper bound on the approximated measures will give us restrictions on the parameters for the linear program in the form of upper bounds on their values.

We have thus constructed a finite-dimensional feasible set. The exact number of parameters will be calculated in the following subsections. The corresponding $(n,m)$-dimensional linear program is then

**Definition 1.14.** The $(n,m)$-*dimensional linear program*:

Find

$$\max \left\{ J(\mu_0, \mu_1, \nu_\tau) \,\middle|\, (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_{(n,m)} \right\}.$$

The strategy of stopping immediately gives us a feasible solution in $\mu_0 \equiv 0$, $\mu_1 \equiv 0$, and $\nu_\tau = \delta_{(0,x_0)}$. Since we assumed a compact space, hence ensuring that the continuous payoff functions are bounded, the $(n,m)$-dimensional linear program is therefore bounded. And thus, by the theory of finite-dimensional linear programming, a solution exists. An overview can be found in Vanderbei (2014).

In Section 2.2, we will show that for large enough $m$, an optimal solution for the $(n,m)$-dimensional, now finite stopping problem is $\varepsilon$-optimal for the $(n,\infty)$-dimensional linear optimal stopping problem.

At this point, we can calculate the coefficients of the finite-dimensional linear program, determine the dependence of the solution on the parameters, and figure out if we need any additional restrictions on the parameter values to ensure that the resulting measures are non-negative and bounded.

### 1.3.1 B-Splines as Test Functions

Combining the methods of Vieten (2018) and Lutz (2007), we will use a finite set of cubic B-Spline basis functions to contruct a finite-dimensional subspace of $C^2([0,T] \times E)$. In contrast to Vieten's treatment of optimal control problems, for finite horizon optimal stopping problem, we will have to use test functions that are defined on time and state space dimensions, but also at least $C^{1,2}([0,T] \times E)$ in order for them to be elements of the domain of the extended generators of the augmented time-space-process $\tilde{A}$ and $\tilde{B}$ that are discussed in Section 1.1.

Let the state space $E = [e_l, e_r] \subset \mathbb{R}$ be a closed interval, and similarly $[0,T] \subset \mathbb{R}$. We now define a grid, i.e. a finite set of points $\{e_i\}_{i=0,\dots,n_E}$ in $[e_l, e_r]$ where $e_{i+1} > e_i$ for all $i$ and similarly on $T$, $\{t_j\}_{j=0,\dots,n_T}$, as well as a corresponding set of values of the target function $f : [0,T] \times E \to \mathbb{R} \times \mathbb{R}$, $\{f_{i,j}\}_{i=0,\dots,n_E, j=0,\dots,n_T} = \{f(e_i, t_j)\}_{i=0,\dots,n_E, j=0,\dots,n_T}$, giving us $N = n_T \cdot n_E$ grid points. In order to approximate or interpolate a function that hits all the points $(t_j, e_i, f_{i,j}), i = 0, \dots, n_E, j = 0, \dots, n_T$ there is a wealth of theory to draw from, most of which use a finite basis of functions to generate a linear combination that best approximates the original function, i.e. we are projecting the original function into the finite-dimensional subspace spanned by the given basis. We will use specifically a basis of cubic B-spline functions as described by De Boor (1978).

Note that De Boor (1978, Theorem XVII.9) provides the groundwork for products of B-splines being bases. Note also, that here, we are deviating from Lutz's choice of test functions (as well as base functions for the densities, which we will discuss in the next section). Lutz, in both instances, used hat functions, i.e. linear B-splines, in the space dimension with a discontinuous discretization in the time dimension, accepting the fact that the regularity is not enough for either dimension, which led him to having to apply the methods of integration by parts and finite differences in order to arrive at workable computations.

Instead, we want to guarantee enough regularity in all derivatives involved, which leads us

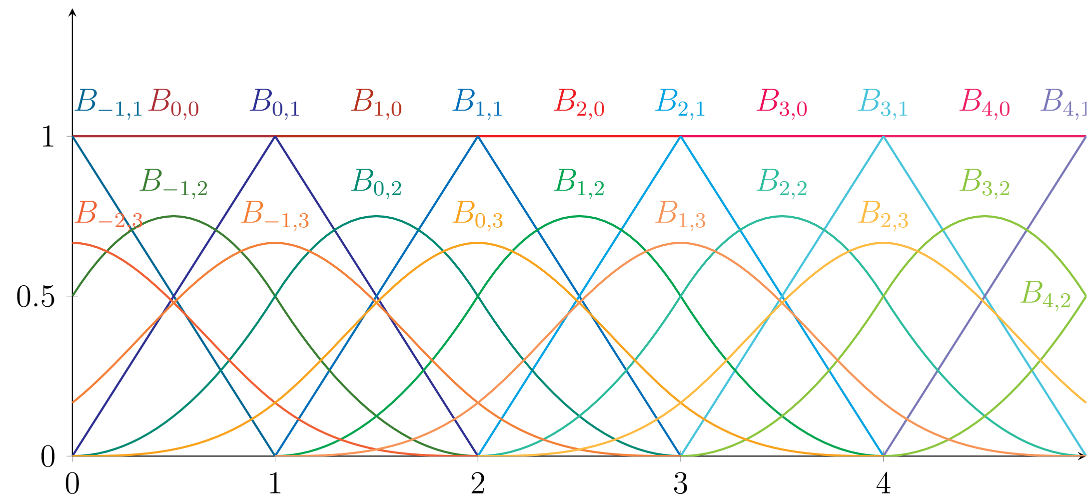


Figure 1: B-splines basis functions order 0 (constant) through 3 (cubic)

to using tensor products of B-splines as described in De Boor (1978, Theorem XVII.9). As Vieten did, we will take cubic B-spline in space direction, but in contrast, we will need continuously differentiable behavior in time direction to satisfy conditions imposed by the augmented continuous generator $\tilde{A}$. Therefore, our choice of B-spline basis will be $B^{(n_T,n_E)} = B_3^{(n_T)} \otimes B_3^{(n_E)}$ where $B_3^{(n_T)} = \{B_{i,3}^{(n_T)}; i = -3, \ldots, n_T - 1\}$ and $B_3^{(n_E)} = \{B_{j,3}^{(n_E)}; j = -3, \ldots, n_E - 1\}$, and

$$B_{j,3}^{(m)}(x) = (e_{j+4} - e_j) \sum_{i=j}^{j+4} \frac{[(e_i - x)^3]^+}{\Psi'_{j,3}(e_i)}, \quad j = -3, \ldots, m-1$$

where

$$\Psi_{j,3}(x) = \prod_{i=j}^{j+4} (x - e_i),$$

and therefore

$$\Psi'_{j,3}(e_i) = \sum_{k=j}^{j+4} \prod_{l=j, l\neq k}^{j+4} (e_i - e_l) = \prod_{l=j, l\neq i}^{j+4} (e_i - e_l).$$

A basis element in $B_{i,j} \in B^{(n_T,n_E)}$ will therefore be of the form $B_{i,j}(t,x) = B_{3,i}^{(n_T)}(t) \cdot B_{3,j}^{(n_E)}(x)$. For a regular grid of step size in time $\delta$ and a step size in state space of $h$, this gives us the formula

$$B_{l,3}^{(m)}(x) = \frac{1}{6h^3} \sum_{k=0}^{4} (-1)^k \binom{4}{k} (x - (l+k)h)^3 \mathbb{1}_{[(l+k)h,\infty)}(x)$$

Images of the one-dimensional B-spline basis functions are shown in Figure 1, while a two-dimensional B-spline basis function is shown in Figure 2. Hall and Meyer (1976, p.1-2) discuss uniform convergence of B-spline approximations in $C^2$.

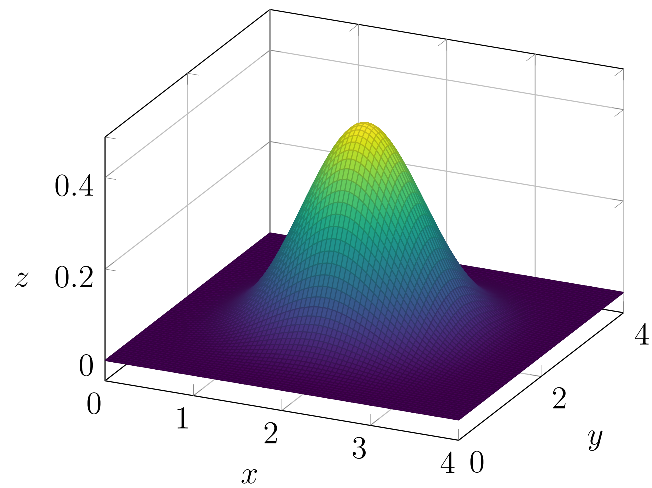


Figure 2: Two-dimensional cubic B-spline base function $B^{(0,0)}(x,y)$ with step on regular grid with step size $h = 1$ on its domain$[0;4] \times [0,4]$

### 1.3.2 Approximating Measures

As discussed in Section 1.1, we will need to approximate three measures: the expected occupation measure $\mu_0$, the local time measure for singular behavior $\mu_1$, and finally the stopping measure $\nu_\tau$, i.e. the distribution of the process at the stopping time. By construction, $\nu_\tau$ is a probability measure, while the other two are just assumed to be finite. We are treating all three as measures on $[0,T] \times E$, although they are concentrated in different subsets. In contrast to the control case, we will not go the route of regular conditional distributions in order to separate the approximation of the different dimensions. There is no equivalent to feedback controls to justify this step. Therefore, we will approximate two-dimensional measures. To keep the complexity of the linear programs low, we will approximate all three measures on the same mesh, which will be defined below.

As in the previous section, we start by defining a grid of discrete points that will be the basis for the approximation of the measures. Note that the different grids for test functions and measures are independent from each other, since they will be approximated in different steps. Once more, the state space $E = [e_l, e_r] \subset \mathbb{R}$ be a closed interval, and similarly $\mathcal{T} = [0, T] \subset \mathbb{R}$. We now define a grid, i.e. a finite set of points $\{e_i\}_{i=0,\dots,m_E}$ in $[e_l, e_r]$ where $e_{i+1} > e_i$ for all $i$ and similarly on $T$, $\{t_j\}_{j=0,\dots,m_T}$, as well as a corresponding set of values of the target function $f : [0, T] \times E \to \mathbb{R} \times \mathbb{R}$, $\{f_{i,j}\}_{i=0,\dots,m_E, j=0,\dots,m_T} = \{f(e_i, t_j)\}_{i=0,\dots,m_E, j=0,\dots,m_T}$, giving us $M = m_T \cdot m_E$ grid points.

We assume that the local time measure $\mu_1$ is concentrated on the finite and constant set of singular points in state space $\{s_1\}$, or in our example with only one point of reflection $s_1 = 1$. Thus, the measure is concentrated on a union of straight lines $[0, T] \times \{s_i\} \subset [0, T] \times E$.

The expected occupation measure $\mu_0$ is concentrated on the continuation region $C \subset [0, T] \times E$, i.e. for any set $A \subset C^c$ $\mu_0(A) = 0$.

We will approximate the three measures $\mu_0$, $\mu_1$, and $\nu_\tau$ by piecewise constants on the rectangles $\{R_i\}$ spanned by the grid points $\{(t_i, e_j) : i = 0, \dots, m_T, j = 0, \dots, m_E\}$. The major focus will be put on the approximation of the stopping measure $\nu_\tau$ since it has a non-trivial support and will need special consideration, as described in Section 2.2. Thus, the approximating densities will be of the form

$$\tilde{p}_0 = \sum_{i=0}^{M} w_{0,i} p_i, \quad \tilde{p}_1 = \sum_{i=0}^{M} w_{1,i} p_i, \quad \tilde{p}_\tau = \sum_{i=0}^{M} w_{\tau,i} p_i,$$

where $p_i = \mathbb{1}_{R_i}$ are the function of the chosen basis based on the rectangles $R_i$, and the weight vectors are the variables to be optimized in the linear program. The approximating measures are then defined as

$$d\tilde{\mu}_0 = \tilde{p}_0 d\lambda, \quad d\tilde{\mu}_1 = \tilde{p}_1 d\lambda, \quad d\tilde{\nu}_\tau = \tilde{p}_\tau d\lambda.$$

# 2 CONVERGENCE

Note that, for the set of problems we are considering, both of the expected occupation measures $\mu_0$ and $\mu_1$ are uniformly bounded, a key condition for the following proofs. For any optimal stopping problem that cannot guarantee uniform bounds in that way, we refer to the additional approximation step imposing an upper bound $l$ for the singular occupation measure $\mu_1$ in Vieten (2018, Section III.2). A similar treatment for the regular time occupation measure is not part of this disseration.

The proof of the convergence of the sequence of optimal values and solutions of the finite-dimensional LPs to an optimal value and solution of the original infinite-dimensional LP is broken into several parts. The first part establishes this convergence for the semi-infinite $(n, \infty)$-dimensional LPs to the infinite-dimensional LP. This convergence is established using the following steps:

1. The limit of a weakly convergent sequence of $\mathcal{M}_{(n,\infty)}$-feasible measures is $\mathcal{M}_\infty$-feasible.

2. The limit of a weakly convergent sequence of $\varepsilon$-optimal $\mathcal{M}_{(n,\infty)}$-feasible measures is $\mathcal{M}_\infty$-feasible.

3. For a sequence of $\varepsilon$-optimal $\mathcal{M}_{(n,\infty)}$-feasible measures the value for any element from the tail of that sequence will eventually be contained in a small enough neighborhood.

4. For any $\delta > 0$ and large enough $n$, the value of an $\varepsilon$-optimal $\mathcal{M}_{(n,\infty)}$-feasible solution is at most $2\varepsilon + \delta$ away from the optimal value of the infinite-dimensional problem.

The second part uses a similar structure to prove convergence of the sequence of optimal values and solutions of the finite $(n, m)$-dimensional LP to the semi-infinite $(n, \infty)$-dimensional LP.

## 2.1 Convergence of the Semi-Infinite LP Values and Solutions

We will first outline the path to proving that the value of an $\varepsilon$-optimal solution to the $(n, \infty)$-dimensional problem is $\varepsilon'$-close to the optimal value for the infinite-dimensional problem. Recall that we cannot guarantee that the approximating solutions for a specific $n$ are feasible for the infinite-dimensional LP. The first proposition shows that limiting measures of a sequence of feasible measures for the $(n, \infty)$-dimensional LP, as $n \to \infty$, exist and are feasible for the infinite-dimensional LP.

**Proposition 2.1.** *Let* $\{(\mu_{0,n}, \mu_{1,n}, \nu_{\tau,n})\}_{n\in\mathbb{N}}$ *be a sequence with* $(\mu_{0,n}, \mu_{1,n}, \nu_{\tau,n}) \in \mathcal{M}_{(n,\infty)}$ *for all* $n \in \mathbb{N}$*. Then there exist* $\nu_\tau \in \mathcal{P}([0, T] \times E)$*,* $\mu_0 \in \mathcal{M}([0, T] \times E)$*, and* $\mu_1 \in \mathcal{M}([0, T] \times E)$ *such that* $\nu_{\tau,n} \Rightarrow \nu_\tau$*,* $\mu_{0,n} \Rightarrow \mu_0$*, and* $\mu_{1,n} \Rightarrow \mu_1$ *along a subsequence and* $(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_\infty$*.*

*Proof.* For any $n > m \in \mathbb{N}$ and $g_m \in \mathcal{D}_m$ we have

$$\int g_m d\nu_{\tau,n} - \int Ag_m d\mu_{0,n} - \int Bg_m d\mu_{1,n} = Rg_m$$

Observe that $g_m$, $Ag_m$, and $Bg_m$ are bounded and uniformly continuous, since they are continuous and the domain $[0, T] \times E$ is compact. Furthermore, compactness of $[0, T] \times E$ along with uniform boundedness implies that the sequence $(\mu_{0,n}, \mu_{1,n}, \nu_{\tau,n})$ is tight. Thus, there exists a weakly convergent subsequence. Without loss of generality, we assume the entire sequence is weakly convergent. In other words, there are measures $\nu_\tau$, $\mu_0$, and $\mu_1$ satisfying $\nu_{\tau,n} \Rightarrow \nu_\tau$, $\mu_{0,n} \Rightarrow \mu_0$, and $\mu_{1,n} \Rightarrow \mu_1$. Hence,

$$\begin{aligned} Rg_m = \lim_{n\to\infty} Rg_m &= \lim_{n\to\infty} \left( \int g_m d\nu_{\tau,n} - \int Ag_m d\mu_{0,n} - \int Bg_m d\mu_{1,n} \right) \\ &= \int g_m d\nu_\tau - \int Ag_m d\mu_0 - \int Bg_m d\mu_1. \end{aligned}$$

Now, let $g \in \mathcal{D}_\infty$ and pick $g_m \in \mathcal{D}_m$ for all $m \in \mathbb{N}$ such that $g_m \to g$ as $m \to \infty$. By continuity

$$\lim_{m\to\infty} Ag_m = Ag, \quad \lim_{m\to\infty} Bg_m = Bg, \text{ and } \quad \lim_{m\to\infty} Rg_m = Rg$$

In addition, since those sequences are convergent in $C^{(1,2)}([0,T] \times E, \|\cdot\|_\infty)$ they are uniformly bounded (see Vieten (2018, Prop. II.1.18)). Therefore, we can apply the dominated convergence theorem:

$$\int g d\nu_\tau - \int A g d\mu_0 - \int B g d\mu_1 = \lim_{m\to\infty} \left( \int g_m d\nu_\tau - \int A g_m d\mu_0 - \int B g_m d\mu_1 \right)$$

This gives us

$$\begin{aligned} \int g d\nu_\tau - \int A g d\mu_0 - \int B g d\mu_1 &= \lim_{m\to\infty} \left( \int g_m d\nu_\tau - \int A g_m d\mu_0 - \int B g_m d\mu_1 \right) \\ &= \lim_{m\to\infty} R g_m = R g \end{aligned}$$

proving that the limit measures are feasible for the infinite-dimensional problem.
Let $l$ denote the uniform bound on all $\mu_{1,n}([0,T]\times E)$ over all $n$. Using the constant function 1, weak convergence implies

$$\mu_1([0,T]\times E) = \int 1 \; d\mu_1 = \lim_{n\to\infty} \int 1 \; d\mu_{1,n} = \lim_{n\to\infty} \mu_{1,n}([0,T]\times E) \le l$$

This establishes that $(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_\infty$. □

The next lemma shows that the weak limits of $\varepsilon$-optimal solutions of the $(n,\infty)$-dimensional LPs comprise an $\varepsilon$-optimal solution to the infinite-dimensional LP.

**Lemma 2.2.** *Let $\varepsilon > 0$ and suppose that, for each $n \in \mathbb{N}$, $(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n}) \in \mathcal{M}_{(n,\infty)}$ is an $\varepsilon$-optimal solution. Let $(\mu^\varepsilon_0, \mu^\varepsilon_1, \nu^\varepsilon_\tau)$ be a weak limit of $(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n})$. Then $(\mu^\varepsilon_0, \mu^\varepsilon_1, \nu^\varepsilon_\tau)$ is an $\varepsilon$-optimal solution for the infinite-dimensional LP.*

*Proof.* By the proof of the previous proposition, we know that there exists a weak limit $(\mu^\varepsilon_0, \mu^\varepsilon_1, \nu^\varepsilon_\tau)$ and it is feasible for $\mathcal{M}_\infty$.
Now, assume that it is not $\varepsilon$-optimal. Then there is a feasible solution $(\mu'_0, \mu'_1, \nu'_\tau) \in \mathcal{M}_\infty$ with

$$J(\mu'_0, \mu'_1, \nu'_\tau) > J(\mu^\varepsilon_0, \mu^\varepsilon_1, \nu^\varepsilon_\tau) + \varepsilon.$$

Let $\delta = J(\mu_0', \mu_1', \nu_\tau') - J(\mu_0^\varepsilon, \mu_1^\varepsilon, \nu_\tau^\varepsilon) - \varepsilon > 0$. Using the definition of the objective function $J$, we get

$$\int c_2 d\nu_\tau' + \int c_0 d\mu_0' + \int c_1 d\mu_1' > \int c_2 d\nu_\tau^\varepsilon + \int c_0 d\mu_0^\varepsilon + \int c_1 d\mu_1^\varepsilon + \varepsilon$$

However, since $c_0$, $c_1$, and $c_2$ are continuous functions on a compact support, weak convergence implies

$$\int c_2 d\nu_{\tau,n}^\varepsilon \to \int c_2 d\nu_\tau^\varepsilon, \quad \int c_0 d\mu_{0,n}^\varepsilon \to \int c_0 d\mu_0^\varepsilon, \text{ and } \quad \int c_1 d\mu_{1,n}^\varepsilon \to \int c_1 d\mu_1^\varepsilon$$

as $n \to \infty$. Hence, we can find an $N \in \mathbb{N}$ such that for all $n \geq N$

$$|J(\mu_0^\varepsilon, \mu_1^\varepsilon, \nu_\tau^\varepsilon) - J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon)| < \frac{\delta}{2}.$$

So we can calculate the difference between $J(\mu_0', \mu_1', \nu_\tau')$ and $J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon)$

$$\begin{aligned}
J(\mu_0', \mu_1', \nu_\tau') &= \delta + \varepsilon + J(\mu_0^\varepsilon, \mu_1^\varepsilon, \nu_\tau^\varepsilon) \\
&= \delta + \varepsilon + J(\mu_0^\varepsilon, \mu_1^\varepsilon, \nu_\tau^\varepsilon) - J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon) + J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon) \\
&> \delta + \varepsilon + J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon) - \frac{\delta}{2} \\
&= J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon) + \varepsilon + \frac{\delta}{2} \\
&> J(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon) + \varepsilon
\end{aligned}$$

Recall that $\mathcal{M}_\infty \subset \mathcal{M}_{(n,\infty)}$ due to the restriction of the constraint space. Thus, the above inequality contradicts the $\varepsilon$-optimality of $(\mu_{0,n}^\varepsilon, \mu_{1,n}^\varepsilon, \nu_{\tau,n}^\varepsilon)$. Therefore, we have

$$J(\mu_0^\varepsilon, \mu_1^\varepsilon, \nu_\tau^\varepsilon) \leq J(\mu_0', \mu_1', \nu_\tau') + \varepsilon \quad \forall (\mu_0', \mu_1', \nu_\tau') \in \mathcal{M}_\infty,$$

giving us $\varepsilon$-optimality for the weak limit. □

At this point, we veer away from the treatment of weak limits, instead focusing on the behavior for large $n$. With the next few lemmata and the closing theorem of this section, we aim to show, that the value of $\varepsilon$-optimal solutions to the $(n, \infty)$-dimensional LP is eventually close enough to the optimal value of the infinite-dimensional LP.

**Lemma 2.3.** *For each $n \in \mathbb{N}$, let $(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) \in \mathcal{M}_{(n,\infty)}$ be an $\varepsilon$-optimal solution. Then, for $\delta > 0$, there are a $z \in \mathbb{R}$ and an $N(\delta) \in \mathbb{N}$ such that for the values of the objective function we have*

$$J(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) \in \Big(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\Big) \quad \forall n \geq N(\delta)$$

*Proof.* Pick an arbitrary $\delta > 0$. Since $[0,T] \times E$ is compact, all three sequences of measures are tight. Furthermore, by assumption, they are uniformly bounded. As as result, this sequence has weak limits. Now, consider two sets of weakly convergent subsequences $(\mu^{\varepsilon}_{0,n_m}, \mu^{\varepsilon}_{1,n_m}, \nu^{\varepsilon}_{\tau,n_m})$ and $(\mu^{\varepsilon}_{0,n_{m'}}, \mu^{\varepsilon}_{1,n_{m'}}, \nu^{\varepsilon}_{\tau,n_{m'}})$, and assume that they have different limits:

$$\nu^{\varepsilon}_{\tau,n_m} \Rightarrow \tilde{\nu}_{\tau},\ \mu^{\varepsilon}_{0,n_m} \Rightarrow \tilde{\mu}_0,\ \mu^{\varepsilon}_{1,n_m} \Rightarrow \tilde{\mu}_1$$

and

$$\nu^{\varepsilon}_{\tau,n_{m'}} \Rightarrow \widehat{\nu}_{\tau},\ \mu^{\varepsilon}_{0,n_{m'}} \Rightarrow \widehat{\mu}_0,\ \mu^{\varepsilon}_{1,n_{m'}} \Rightarrow \widehat{\mu}_1$$

Now, let us assume that

$$\int c_2 d\tilde{\nu}_{\tau} + \int c_0 d\tilde{\mu}_0 + \int c_1 d\tilde{\mu}_1 < \int c_2 d\widehat{\nu}_{\tau} + \int c_0 d\widehat{\mu}_0 + \int c_1 d\widehat{\mu}_1 - \varepsilon$$

Using the same convergence argument as in the previous proof, choose $N$ large so that for any $m \geq N$

$$\int c_2 d\nu^{\varepsilon}_{\tau,n_m} + \int c_0 d\mu^{\varepsilon}_{0,n_m} + \int c_1 d\mu^{\varepsilon}_{1,n_m} < \int c_2 d\widehat{\nu}_{\tau} + \int c_0 d\widehat{\mu}_0 + \int c_1 d\widehat{\mu}_1 - \varepsilon$$

However, since $(\widehat{\mu}_0, \widehat{\mu}_1, \widehat{\nu}_{\tau}) \in \mathcal{M}_{\infty} \subset \mathcal{M}_{(n_m,\infty)}$, this contradicts the supposition of $\varepsilon$-optimality of $(\mu^{\varepsilon}_{0,n_m}, \mu^{\varepsilon}_{1,n_m}, \nu^{\varepsilon}_{\tau,n_m})$. Therefore, we have

$$\int c_2 d\tilde{\nu}_{\tau} + \int c_0 d\tilde{\mu}_0 + \int c_1 d\tilde{\mu}_1 \geq \int c_2 d\widehat{\nu}_{\tau} + \int c_0 d\widehat{\mu}_0 + \int c_1 d\widehat{\mu}_1 - \varepsilon.$$

And since we can similarly show that

$$\int c_2 d\widehat{\nu}_{\tau} + \int c_0 d\widehat{\mu}_0 + \int c_1 d\widehat{\mu}_1 \geq \int c_2 d\tilde{\nu}_{\tau} + \int c_0 d\tilde{\mu}_0 + \int c_1 d\tilde{\mu}_1 - \varepsilon,$$

we get

$$\left|\int d\tilde{\nu}_\tau - \int c_0 d\tilde{\mu}_0 - \int c_1 d\tilde{\mu}_1 - \left(\int d\widehat{\nu}_\tau - \int c_0 d\widehat{\mu}_0 - \int c_1 d\widehat{\mu}_1\right)\right| \leq \varepsilon.$$

To summarize, the limits of any two convergent subsequences of $\{(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n})\}_{n\in\mathbb{N}}$ are at most $\varepsilon$ apart. So we can find a number $z \in \mathbb{R}$ such that

$$J(\tilde{\mu}_0, \tilde{\mu}_1, \tilde{\nu}_\tau) \in \left[z - \frac{\varepsilon}{2}; z + \frac{\varepsilon}{2}\right]$$

for any weak limit of a convergent subsequence of $\{(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n})\}_{n\in\mathbb{N}}$. Thus, if the sequence $\{(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n})\}_{n\in\mathbb{N}}$ converges weakly, we can per definitionem find an $N(\delta)$ for any $\delta > 0$ such that

$$J(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n}) \in \left(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\right).$$

Now, assume $\{(\mu^\varepsilon_{0,n}, \mu^\varepsilon_{1,n}, \nu^\varepsilon_{\tau,n})\}_{n\in\mathbb{N}}$ does not converge, and that for any $N \in \mathbb{N}$ there is a $k \geq N$ such that

$$J(\mu^\varepsilon_{0,k}, \mu^\varepsilon_{1,k}, \nu^\varepsilon_{\tau,k}) \notin \left(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\right). \tag{2.1}$$

Thus, we can construct a subsequence $\{(\mu^\varepsilon_{0,n_k}, \mu^\varepsilon_{1,n_k}, \nu^\varepsilon_{\tau,n_k})\}_{k\in\mathbb{N}}$ for which the value of every element is outside of the interval. Since this sequence is still tight and uniformly bounded, we can find a convergent sub-subsequence $\{(\mu^\varepsilon_{0,n_{k_l}}, \mu^\varepsilon_{1,n_{k_l}}, \nu^\varepsilon_{\tau,n_{k_l}})\}_{l\in\mathbb{N}}$ with weak limits $(\overline{\mu}_0, \overline{\mu}_1, \overline{\nu}_\tau)$. However, since this is a subsequence of the original sequence, the value of the weak limits satisfies

$$J(\overline{\mu}_0, \overline{\mu}_1, \overline{\nu}_\tau) \in \left[z - \frac{\varepsilon}{2}; z + \frac{\varepsilon}{2}\right].$$

Consequently, there is a large $N(\delta)$ such that for any $l \geq N(\delta)$ we have

$$J(\mu^\varepsilon_{0,n_{k_l}}, \mu^\varepsilon_{1,n_{k_l}}, \nu^\varepsilon_{\tau,n_{k_l}}) \in \left(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\right)$$

This, however, contradicts the construction of the subsequence, and therefore the assumption (2.1). I.e., if the original sequence does not converge, we can have at most finitely many elements whose value is outside of $\left(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\right)$. In other words, there is an $N(\delta)$ such that for any $n \geq N(\delta)$

$$J(\mu_{0,n}^{\varepsilon}, \mu_{1,n}^{\varepsilon}, \nu_{\tau,n}^{\varepsilon}) \in \left(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\right).$$

We have now shown the conclusion whether or not the original sequence converges. Thus, we are done. □

Now that the value of $\varepsilon$-optimal solutions for large enough $n$ is confined in a bounded interval, we show that the optimal value for the infinite-dimensional LP is also bounded in a related interval. We will then combine the intervals in the result of this section's theorem.

**Proposition 2.4.** *Let $z$, $\varepsilon$ be as in the previous lemma and define the optimal value for the infinite-dimensional problem*

$$J^* = \sup\left\{J(\mu_0, \mu_1, \nu_\tau) : (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_\infty\right\}.$$

*Then*

$$z - \frac{\varepsilon}{2} \leq J^* \leq z + \frac{3\varepsilon}{2}.$$

*Proof.* Again, consider a sequence $\{(\mu_{0,n}^{\varepsilon}, \mu_{1,n}^{\varepsilon}, \nu_{\tau,n}^{\varepsilon})\}_{n\in\mathbb{N}}$ of $\varepsilon$-optimal solutions in $\mathcal{M}_{(n,\infty)}$ and an arbitrary convergent subsequence $\{(\mu_{0,n_k}^{\varepsilon}, \mu_{1,n_k}^{\varepsilon}, \nu_{\tau,n_k}^{\varepsilon})\}_{k\in\mathbb{N}}$ with limits $(\nu_\tau^{\varepsilon}, \mu_0^{\varepsilon}, \mu_1^{\varepsilon})$. We know from Lemma 2.2 and Lemma 2.3 that the limit is $\epsilon$-optimal and its value contained in the closed interval $\left[z - \frac{\varepsilon}{2}; z + \frac{\varepsilon}{2}\right]$. Therefore, for the optimal value $J^*$, we can find the following bounds

$$J^* \geq J(\mu_0^{\varepsilon}, \mu_1^{\varepsilon}, \nu_\tau^{\varepsilon}) \geq z - \frac{\varepsilon}{2}$$

and

$$J^* - \varepsilon \leq J(\mu_0^{\varepsilon}, \mu_1^{\varepsilon}, \nu_\tau^{\varepsilon}) \leq z + \frac{\varepsilon}{2}$$

Showing the upper bound $J^* \leq z + \frac{3\varepsilon}{2}$. □

The next theorem shows that the values of a sequence of $\varepsilon$-optimal solutions to the semi-infinite LP will eventually be $\varepsilon$-close to the optimal value of the infinite-dimensional LP.

**Theorem 2.5.** *Assume that, for each $n \in \mathbb{N}$, $(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) \in \mathcal{M}_{(n,\infty)}$ is an $\varepsilon$-optimal solution. Then, for $\delta > 0$, there is an $N(\delta) \in \mathbb{N}$ such that for all $n \geq N(\delta)$*

$$|J(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) - J^*| \leq 2\varepsilon + \delta.$$

*Proof.* By Lemma 2.3, choose $N$ large enough that for any $n \geq N$

$$J(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) \in \left(z - \frac{\varepsilon}{2} - \delta; z + \frac{\varepsilon}{2} + \delta\right).$$

By Proposition 2.4, we get the following bounds

$$J(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) - J^* < z + \frac{\varepsilon}{2} + \delta - \left(z - \frac{\varepsilon}{2}\right) = \varepsilon + \delta$$

and

$$J^* - J(\mu^{\varepsilon}_{0,n}, \mu^{\varepsilon}_{1,n}, \nu^{\varepsilon}_{\tau,n}) < z + \frac{3\varepsilon}{2} - \left(z - \frac{\varepsilon}{2} - \delta\right) = 2\varepsilon + \delta$$

This proves the conclusion. ☐

## 2.2 Convergence of the Finite LP Values and Solutions

In this section, we will let $n \in \mathbb{N}$ be fixed and $\mathcal{M}_{n,\infty}$ be the set of feasible solutions to the semi-infinite linear program for constraint refinement level $n$. Recall that the set of feasible solutions to the $(n, m)$-dimensional linear program is

$$\mathcal{M}_{(n,m)} = \Big\{ (\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}^{(m)}_0([0,T] \times E) \times \mathcal{M}^{(m)}_1([0,T] \times E) \times \mathcal{M}^{(m)}_\tau([0,T] \times E) \Big| \\ Rf_k = \int \widehat{f}^{(m)}_k d\nu_\tau - \int Af_k d\mu_0 - \int Bf_k d\mu_1, \quad k = 1, \ldots, n \Big\},$$

where $\widehat{f}^{(m)}_k$ is the piecewise constant approximation from below of the $k$-th test function as defined in (1.8)

$$\begin{aligned} f^{(m)}_{k,i} &= \inf_{(t,x)\in R_i} \{f_k(t,x)\}, && k = 1, \ldots, n, \quad i = 0, \ldots, 4^m - 1, \\ \widehat{f}^{(m)}_k &= \sum_{i=0}^{4^m-1} f^{(m)}_{k,i} \mathbb{1}_{R_i}, && k = 1, \ldots, n. \end{aligned}$$

The goal of this section is to relate the solutions of the finite LP to solutions of the semi-infinite LP. In the first half, we will solely focus on arbitrarily close approximations of given densities, while in the second half, we will loosely follow the structure of the previous section, showing that, for large enough $m$, optimal solutions of the finite LP are $\varepsilon$-optimal for the semi-infinite LP. Recall that we require several assumptions on the measures in the feasible set of the semi-infinite LP:

1. All three sets of measures in $\mathcal{M}_{(n,\infty)}$ need to be uniformly bounded. In fact, we can guarantee this as long as we are dealing with finite horizon optimal stopping problems. In that case, any feasible regular time occupation measure $\mu_0$ is bounded above by the time horizon $T$, and any feasible local time occupation measure $\mu_1$ is bounded above by a constant $l > 0$, as described in Remarks 1.3 and 1.6, respectively.

2. All three measures are absolutely continuous with respect to some Lebesgue measure, and thus have $L^1$-density functions.

3. We also require the density $p_\tau$ of the stopping measure $\nu_\tau$ to be Lebesgue-almost surely strictly positive. This will be needed to construct a feasible solution to the finite linear program.

The first lemma shows that we can approximate any measure with density by a measure having piecewise constant Radon-Nikodym derivative over a finite partition. This lemma will be applied to all three measures $\mu_0$, $\mu_1$, and with adjustments, that are described hereafter, to $\nu_\tau$.

**Lemma 2.6.** *Let $(S, \mathcal{F}, \lambda, d)$ be a measure space with metric $d$, $\sigma$-algebra $\mathcal{F}$. Let $\mu \ll \lambda$ be a finite measure on $S$ with Radon-Nikodym derivative $p$, i.e. $\mu(dx) = p(x)\lambda(dx)$. Let $f : S \to \mathbb{R}$ be a uniformly continuous and bounded function. Then for any $\varepsilon > 0$ there is a $\delta$ such that for any finite partition $\Pi = \{R_i\}$ which satisfies $\max_i\{diam(R_i)\} < \delta$, the piecewise* *constant approximation* $\tilde{p}_\Pi : S \to \mathbb{R}_0^+$ *of* $p$ *given by*

$$\tilde{p}_\Pi(x) = \sum_i \operatorname{ess\,inf}\{p(x)|x \in R_i\}\, \mathbb{1}_{R_i}(x)$$

*satisfies*

$$\|p - \tilde{p}_\Pi\|_{L^1(S)} < \varepsilon \quad \textit{and} \quad \left| \int_S f(x)p(x)\lambda(dx) - \int_S f(x)\tilde{p}_\Pi(x)\lambda(dx) \right| < \varepsilon.$$

*Proof.* First note, that the construction of $\tilde{p}_\Pi$ can be applied to any partition $\Pi$ of S, and that we have almost sure pointwise convergence with increasing refinement:

$$\lim_{\delta \to 0} p(x) - \tilde{p}_\Pi(x) = 0 \text{ for } \lambda\text{-a.e. } x \in S.$$

By the dominated convergence theorem, we get $L^1$-convergence, proving the first inequality for some $\delta_1$.

In this first part, we assume that $f$ is non-negative. By uniform continuity of $f$, we can now choose a $\delta_2$ such that $|f(x) - f(y)| < \frac{\varepsilon}{4\mu(X)}$ for $d(x,y) < \delta_2$. Due to the construction and the $L^1$-convergence of $\tilde{p}_\Pi$, we can choose a $\delta_3$ such that $\int_S p(x) - \tilde{p}_\Pi(x)\lambda(dx) < \frac{\varepsilon}{4F}$ where $F = \max_{x \in S}\{f(x)\} < \infty$ whenever $\max_i\{\text{diam}(R_i)\} < \delta_3$. Define $\delta = \min\{\delta_1, \delta_2, \delta_3\}$ and let $\Pi = \{R_i\}$ be a partition with $\max_i\{\text{diam}(R_i)\} < \delta$.

First, note that $p \geq \tilde{p}_\Pi$ implies the difference of the two integrals is non-negative. Recalling the definition of the lower Riemann integral approximation, we can now calculate an upper bound for the error of the approximated integral over the uniformly continuous function $f$.

Define for each $i$, $f_i = \operatorname*{ess\,inf}_{x \in R_i} \{f(x)\} \geq 0$ on each set $R_i$ of the partition $\Pi$.

$$
\begin{aligned}
&\int_S f(x)\ \mu(dx) - \int_S f(x)\ \tilde{p}_\Pi(x)\lambda(dx) \\
=&\sum_i \int_{R_i} f(x)\ p(x)\lambda(dx) - \sum_i \int_{R_i} f(x)\ \tilde{p}_\Pi(x)\lambda(dx) \\
=&\sum_i \left\{ \left[ \int_{R_i} f(x)\ p(x)\lambda(dx) - \int_{R_i} f(x)\ \tilde{p}_\Pi(x)\lambda(dx) \right] + \left[ \int_{R_i} f_i\ p(x)\lambda(dx) - \int_{R_i} f_i\ p(x)\lambda(dx) \right] \right. \\
&\left. + \left[ \int_{R_i} f_i\ \tilde{p}_\Pi(x)\lambda(dx) - \int_{R_i} f_i\ \tilde{p}_\Pi(x)\lambda(dx) \right] \right\} \\
=&\sum_i \left\{ \left[ \int_{R_i} (f(x) - f_i)\ p(x)\lambda(dx) \right] + \left[ \int_{R_i} (f_i - f(x))\ \tilde{p}_\Pi(x)\lambda(dx) \right] \right. \\
&\left. + \left[ \int_{R_i} f_i\ (p(x) - \tilde{p}_\Pi(x))\lambda(dx) \right] \right\}
\end{aligned}
$$

Note that the first and third integral are non-negative and the second integral is non-positive due to our choices of $f_i$ and our construction of $\tilde{p}_\Pi$.

We can now show the bounds of the above expression.

$$
\begin{aligned}
0 &\leq \int_S f(x)\ \mu(dx) - \int_S f(x)\ \tilde{p}_\Pi(x)\lambda(dx) \\
&= \sum_i \left\{ \left[ \int_{R_i} (f(x) - f_i)\ p(x)\lambda(dx) \right] + \left[ \int_{R_i} (f_i - f(x))\ \tilde{p}_\Pi(x)\lambda(dx) \right] \right. \\
&\qquad \left. + \left[ \int_{R_i} f_i\ (p(x) - \tilde{p}_\Pi(x))\lambda(dx) \right] \right\} \\
&< \sum_i \left\{ \left[ \int_{R_i} \frac{\varepsilon}{\mu(S)} p(x)\lambda(dx) \right] + 0 + f_i \cdot \left[ \int_{R_i} p(x) - \tilde{p}_\Pi(x)\lambda(dx) \right] \right\} \\
&\leq \frac{\varepsilon}{4\mu(S)}\mu(S) + F \left\| p - \tilde{p}_\Pi \right\|_{L^1(S)} \\
&< \frac{\varepsilon}{4\mu(S)}\mu(S) + F\frac{\varepsilon}{4F} \\
&\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2},
\end{aligned}
$$

Since the expression is non-negative, we now have

$$
\left| \int_S f(x)\ \mu(dx) - \int_S f(x)\ \tilde{p}_\Pi(x)\lambda(dx) \right| < \frac{\varepsilon}{2},
$$

thus completing the proof for non-negative functions $f$.

To get the full result for an arbitrary continuous function $f$, we split $f$ up into the positive part $f^+ = \max(f, 0)$ and the negative part $f^- = \max(-f, 0)$, apply the above to find $m^+ \in \mathbb{N}$ and $m^- \in \mathbb{N}$, and take the larger of the two $m = \max(m^+, m^-)$ to get

$$
\begin{aligned}
&\left|\int_S f(x)\ \mu(dx) - \int_S f(x)\ \tilde{p}_\Pi(x)\lambda(dx)\right| \\
=&\left|\int_S (f^+(x) - f^-(x))\ p(x)\lambda(dx) - \int_S (f^+(x) - f^-(x))\tilde{p}_\Pi(x)\lambda(dx)\right| \\
=&\left|\int_S f^+(x)\ p(x)\lambda(dx) - \int_S f^-(x)\ p(x)\lambda(dx) - \int_S f^+(x)\tilde{p}_\Pi(x)\lambda(dx) + \int_S f^-(x)\tilde{p}_\Pi(x)\lambda(dx)\right| \\
\leq&\left|\int_S f^+(x)\ p(x)\lambda(dx) - \int_S f^+(x)\tilde{p}_\Pi(x)\lambda(dx)\right| - \left|\int_S f^-(x)\ p(x)\lambda(dx) - \int_S f^-(x)\tilde{p}_\Pi(x)\lambda(dx)\right| \\
<&\frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,
\end{aligned}
$$

completing the proof. □

**Remark 2.7.** Lemma 2.6 was kept general but can be applied to the set of partitions and the truncated basis described in Section 1.3. The $\delta$ provided by the lemma gives us a lower bound for the level of refinement $m$ of the mesh. Recall that for a given $m$, the diameter of each rectangle $R_i$ of the partition is $\operatorname{diam}(R_i) = \frac{\sqrt{T^2+(e_r-e_l)^2}}{2^m}$. Therefore, if we require $\operatorname{diam}(R_i) < \delta$, then $m > \log_2\left(\frac{\sqrt{T^2+(e_r-e_l)^2}}{\delta}\right)$.

The next two lemmas establish a more specific approximation for the stopping measure $\nu_\tau$. We will first use Lemma 2.6 to find a sufficiently close approximation on the stopping boundary $l \subset [0, T] \times E$, and then define a corresponding density approximation that is piecewise constant on the usual rectangular dyadic partition $\Pi_m = \{R_i\}$ as defined in Section 1.3. This extended density will be the basis for constructing an approximate feasible solution in $\mathcal{M}_{(n,m)}$.

**Lemma 2.8.** *Let $\varepsilon > 0$, $D > 0$, and $\nu_\tau$ be a stopping measure on $[0, T] \times E$ as described above with $\lambda_l(\{p_\tau = 0\}) = 0$. Let $f : [0, T] \times E \to \mathbb{R}$ be a uniformly continuous and bounded function. Then there exists an $0 < \widehat{\varepsilon} < \varepsilon$, $m_0 \in \mathbb{N}$ such that for every $m \geq m_0$ there is $\delta > 0$ such that for any finite partition $\Pi_l = \{L_i\}$ of the stopping boundary $l$ satisfying*

$\max_i\{diam(L_i)\} < \delta$ *the piecewise constant approximation* $\tilde{p}_{\Pi_l} : l \to \mathbb{R}_0^+$ *given by*

$$\tilde{p}_{\Pi_l}(x) = \sum_i \operatorname{ess\,inf}\{p_\tau(x)|x \in L_i\}\, \mathbb{1}_{L_i}(x)$$

*satisfies* $\tilde{p}_{\Pi_l} \geq \widehat{\varepsilon}$ *as well as*

$$\|p_\tau - \tilde{p}_{\Pi_l}\|_{L^1(l)} < \frac{\widehat{\varepsilon}}{D} \quad \text{and} \quad \left|\int_l f(x)p_\tau(x)\lambda_l(dx) - \int_l f(x)\tilde{p}_{\Pi_l}(x)\lambda_l(dx)\right| < \frac{\widehat{\varepsilon}}{D}. \tag{2.2}$$

*Proof.* Let $\varepsilon > 0$, $D > 0$, $f$ be a uniformly continuous and bounded function, and $\nu_\tau$ and $p_\tau$ be given such that $d\nu_\tau = p_\tau d\lambda_l$. Define $F = \max(1, \|f\|_\infty) < \infty$.

As a first step, using continuity from above for measures, we can find an $0 < \widehat{\varepsilon} < \varepsilon$ such that $\lambda_l(\{p_\tau \leq \widehat{\varepsilon}\}) < \frac{1}{4DF}$. Define $\overline{p} \in L^1(l)$ as

$$\overline{p}(t,x) = \begin{cases} p_\tau(t,x) & p_\tau(t,x) \geq \widehat{\varepsilon} \\ \widehat{\varepsilon} & p_\tau(t,x) < \widehat{\varepsilon}. \end{cases}$$

Thus, $\|p_\tau - \overline{p}\|_{L^1} \leq \widehat{\varepsilon} \cdot \lambda_l(\{p_\tau \leq \widehat{\varepsilon}\}) < \frac{\widehat{\varepsilon}}{4DF}$. Now, apply Lemma 2.6 on the subspace $(l, \lambda_l)$ of $[0,T] \times E$ with the lifted density $\overline{p}$ and $\varepsilon = \frac{\widehat{\varepsilon}}{4DF} < \frac{\widehat{\varepsilon}}{D}$. The resulting approximation $\tilde{p}_{\tau,\Pi_l}$ satisifes

$$\|p_\tau - \tilde{p}_{\tau,\Pi_l}\|_{L^1(l)} \leq \|p_\tau - \overline{p}\|_{L^1(l)} + \|\overline{p} - \tilde{p}_{\tau,\Pi_l}\|_{L^1(l)} < \frac{\widehat{\varepsilon}}{4DF} + \frac{\widehat{\varepsilon}}{4DF} = \frac{\widehat{\varepsilon}}{2DF} < \frac{\widehat{\varepsilon}}{D}.$$

Finally, we show that the integrals are equally close.

$$\begin{aligned}
&\left|\int_l f(x)\ p_\tau(x)\lambda_l(dx) - \int_l f(x)\ \tilde{p}_{\tau,\Pi_l}(x)\lambda_l(dx)\right| \\
\leq &\left|\int_l f(x)\ p_\tau(x)\lambda_l(dx) - \int_l f(x)\ \overline{p}(x)\lambda_l(dx)\right| + \left|\int_l f(x)\ \overline{p}(x)\lambda_l(dx) - \int_l f(x)\ \tilde{p}_{\tau,\Pi_l}(x)\lambda_l(dx)\right| \\
< &\frac{\widehat{\varepsilon}}{4DF} + \|f\|_\infty \|\overline{p} - \tilde{p}_{\tau,\Pi_l}\|_{L^1(l)} = \frac{\widehat{\varepsilon}}{4DF} + \frac{\widehat{\varepsilon}}{2D} < \frac{\widehat{\varepsilon}}{D},
\end{aligned}$$

finishing the proof. □

**Remark 2.9.** Using the usual rectangular dyadic partitions $\Pi_m = \{R_i\}$, we now define a partition on the stopping boundary

$$\Pi_{l,m} = \{L_i | L_i = R_i \cap l, R_i \in \Pi_m, L_i \neq \emptyset\},$$

which satisfies $\max_i\{\text{diam}(L_i)\} < \delta$ for $m$ large enough. Note that the index set changes. However, for the sake of consistency and to emphasize the relationship with the original partition, we will keep the index $i$. As in Remark 2.7, the $\delta$ provided by Lemma 2.8 will give us a refinement level $m$ for the rectangular partition $\Pi_m$, that is fine enough. Thus, using Lemma 2.8 we can find an $m$ and a corresponding approximation

$$\tilde{p}_{\tau,m} = \sum_i w_{\tau,i} p_i^{(l)}. \tag{2.3}$$

where $p_i^{(l)} = \mathbb{1}_{L_i}$ are the constant basis functions on $l$ and consequently $\tilde{p}_{\tau,m}$ satisifies the conclusions of Lemma 2.8.

Note that, by using a sufficiently fine partition, we can guarantee that there are as many sets $L_i$ of the with positive mass with respect to $\lambda_l$ partition as we wish.

**Remark 2.10.** Having constructed this approximation $\tilde{p}_{\tau,m}$ on the stopping boundary $l$ with weights $w_{\tau,i}$, we will now in a sense extend the approximated density $\tilde{p}_{\tau,m} : l \to \mathbb{R}_0^+$ to a piecewise constant function $\overline{p}_{\tau,m} : [0,T] \times E \to \mathbb{R}_0^+$ using the set of basis functions $\mathfrak{P}^{(m)}$ based on the rectangular partition $\Pi_m = \{R_i\}$. Recall that $L_i = R_i \cap l$, the intersection of each rectangle of the partition $R_i$ with the stopping boundary $l$. Now, define

$$\overline{p}_{\tau,m}(t,x) = \sum_{R_i \in \Pi_m} w_{\tau,i} \frac{\lambda_l(L_i)}{\lambda_2(R_i)} \mathbb{1}_{R_i}. \tag{2.4}$$

Note that the weight $w_{\tau,i}\frac{\lambda_l(L_i)}{\lambda_2(R_i)}$ is positive on the rectangle $R_i$ given that $\lambda_l(L_i) > 0$.

The following lemma shows that the approximation on the rectangles is still close enough.

**Lemma 2.11.** *Let $\nu_\tau$ and $f$ be as in Lemma 2.8. Then for $\varepsilon > 0$ and $D > 0$ there exists an $\widehat{\varepsilon} < \varepsilon$ and an $m_0 \in \mathbb{N}$ such that for every $m \geq m_0$ there are functions $\tilde{p}_{\tau,m}$ and $\overline{p}_{\tau,m}$ given by* (2.3) *and* (2.4)*, respectively, with*

$$\left\|\tilde{p}_{\tau,m}\right\|_{L^1(l)} = \left\|\overline{p}_{\tau,m}\right\|_{L^1([0,T]\times E)},$$
$$\left|\int_l f(t,x)p_\tau(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} f(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| < \frac{\widehat{\varepsilon}}{D}.$$

*Proof.* An application of Lemma 2.8 with $\varepsilon$ and $D' = 5D$ shows the existence of an $m_0$ such that for each $m \geq m_0$ there is a $\tilde{p}_{\tau,m}$ of the form (2.3) satisfying the conclusions (2.2). Define $\varepsilon' = \min(1, \frac{\widehat{\varepsilon}}{D'})$ and, if necessary, increase $m_0$ such that uniform continuity of $f$ implies $\sup_{(t,x)\in R_i} f(t,x) - \inf_{(t,x)\in R_i} f(t,x) < \varepsilon'$ on each rectangle $R_i$. Fix $m \geq m_0$. Define $\overline{p}_{\tau,m}$ as in (2.4) using $\tilde{p}_{\tau,m}$. Then

$$\begin{aligned} \left\|\overline{p}_{\tau,m}\right\|_{L^1([0,T]\times E)} &= \int_{[0,T]\times E} \overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx) \\ &= \int_{[0,T]\times E} \sum_{i:R_i\in\Pi_m} w_{\tau,i} \frac{\lambda_l(L_i)}{\lambda_2(R_i)} \mathbb{1}_{R_i}(t,x)\lambda_2(dt \times dx) \\ &= \sum_{i:R_i\in\Pi_m} w_{\tau,i} \frac{\lambda_l(L_i)}{\lambda_2(R_i)} \int_{R_i} \lambda_2(dt \times dx) \\ &= \sum_{i:R_i\in\Pi_m} w_{\tau,i}\lambda_l(L_i) \qquad (2.5) \\ &= \int_l \sum_{i:R_i\in\Pi_m} w_{\tau,i} p_i^{(l)}(t,x)\lambda_l(dt \times dx) = \|\tilde{p}_{\tau,m}\|_{L^1(l)} . \end{aligned}$$

In addition

$$\begin{aligned} &\left| \int_l f(t,x)p_\tau(t,x)\lambda_l(dt \times dx) - \int_{[0,T]\times E} f(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx) \right| \\ &\leq \left| \int_l f(t,x)p_\tau(t,x)\lambda_l(dt \times dx) - \int_l f(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt \times dx) \right| \\ &+ \left| \int_l f(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt \times dx) - \int_{[0,T]\times E} f(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx) \right| . \end{aligned}$$

Observe that the first term is less than $\frac{\widehat{\varepsilon}}{5D}$ due to Lemma 2.8.

We examine the second term. Define the function $\widehat{f}^{(m)}$ as in (1.8). Then we have

$$\begin{aligned}
&\left|\int_l f(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} f(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| \\
&\leq \left|\int_l f(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt\times dx) - \int_l \widehat{f}^{(m)}(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt\times dx)\right| \\
&\quad + \left|\int_l \widehat{f}^{(m)}(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} \widehat{f}^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| \\
&\quad + \left|\int_{[0,T]\times E} \widehat{f}^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx) - \int_{[0,T]\times E} f(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| \\
&\leq \int_l \left|f(t,x) - \widehat{f}^{(m)}(t,x)\right| \tilde{p}_{\tau,m}(t,x)\lambda_l(dt\times dx) \\
&\quad + \left|\int_l \widehat{f}^{(m)}(t,x)\tilde{p}_{\tau,m}(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} \widehat{f}^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| \\
&\quad + \int_{[0,T]\times E} \left|\widehat{f}^{(m)}(t,x) - f(t,x)\right| \overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx).
\end{aligned}$$

The sum of the first and third term is bounded above by $\frac{\widehat{\varepsilon}}{2D}\cdot\|\tilde{p}_{\tau,m}\|_{L^1(l)}$. Since $\widehat{f}^{(m)}$, $\tilde{p}_{\tau,m}$, and $\overline{p}_{\tau,m}$ are constant on each of the rectangles $R_i$, a similar argument as in (2.5) gives that the second term is equal to zero. Combining all the terms establishes that

$$\left|\int_l f(t,x)p_\tau(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} f(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| < \frac{\widehat{\varepsilon}}{D},$$

finishing the proof. □

**Remark 2.12.** A careful analysis of the proof of Lemma 2.11 also establishes that

$$\left|\int_l f(t,x)p_\tau(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} \widehat{f}^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| < \frac{\widehat{\varepsilon}}{D}.$$

Note that the choice of $\widehat{f}^{(m)}$ is not unique. In each rectangle $R_i$, any evaluation of the function inside the rectangle is a valid choice for the constant $f_i^{(m)}$. We chose the infimum for clarity of presentation. Thus, this flexibility in choice of $\widehat{f}^{(m)}$ simplifies the implementation of the numerical scheme since it is not necessary to find the infimum over each rectangle.

We now turn to the convergence of optimal solutions of the finite LP to optimal solutions of the semi-infinite LP. The structure of the proof involves showing the existence of arbitrarily close approximations of the density functions with $m$ large enough. Then, establishing the existence of such approximations that are feasible. Next, we prove the weak convergence of optimal solutions in $\mathcal{M}_{(n,m)}$ to an optimal solution in $\mathcal{M}_{(n,\infty)}$, as $m \to \infty$. Finally, we will show $\varepsilon$-optimality in $\mathcal{M}_{(n,\infty)}$ for the optimal solutions of $\mathcal{M}_{(n,m)}$ for large enough $m$.

In order to show the existence of a feasible approximation to a given density, we will need to define the constraint matrix and constraint error. We are now restricting our choice of partitions to the dyadic partition $\Pi_m = \{R_j, j = 0, \ldots, 4^m - 1\}$ ($m \in \mathbb{N}$) for the measures and the truncated test function space $\mathcal{D}_n = \operatorname{span}(\{f_k\}_{k=1}^n)$. Recall that $\lambda_l$ and $\lambda_2$ are the Lebesgue measures on the stopping boundary $l$ and the underlying space $[0,T] \times E$, respectively. In accordance with the constraint equations for the $(n,m)$-dimensional linear program

$$Rf_k = \int \widehat{f}_k^{(m)} d\nu_\tau - \int Af_k d\mu_0 - \int Bf_k d\mu_1, \quad k = 1, \ldots, n,$$

where $\widehat{f}_k^{(m)}$ is the piecewise constant approximation from below of the $k$-th test function

$$\begin{aligned} f_{k,i}^{(m)} &= \inf_{(t,x)\in R_i} \{f_k(t,x)\}, && k = 1, \ldots, n, \quad i = 0, \ldots, 4^m - 1, \\ \widehat{f}_k^{(m)} &= \sum_{i=0}^{4^m-1} f_{k,i}^{(m)} \mathbb{1}_{R_i}, && k = 1, \ldots, n, \end{aligned}$$

we now define the constraint matrix $C^{(m)}$ using integrals on the stopping boundary $l$ and the constraint error $d$ using integrals on $[0,T] \times E$.

**Definition 2.13.** For $n, m \in \mathbb{N}$, define the constraint matrix for an optimal stopping problem $C^{(m)} \in \mathbb{R}^{n+1,4^m}$ by

$$\begin{aligned} C_{k,j}^{(m)} &= \int_l \widehat{f}_k^{(m)}(t,x) p_j^{(l)}(t,x) \lambda_l(dt \times dx), \quad \text{for } j = 0, 1, \ldots, 4^m - 1, \text{ for } k = 1, 2, \ldots, n, \\ C_{n+1,j}^{(m)} &= \int_l p_j^{(l)}(t,x) \lambda_l(dt \times dx) = \lambda_l(L_j), \quad \text{for } j = 0, 1, \ldots, 4^m - 1, \end{aligned}$$

where $p_j^{(l)}(t,x) = \mathbb{1}_{L_j}(t,x)$ is the $j$-th piecewise constant basis element for the chosen grid.

**Definition 2.14.** For $m \in \mathbb{N}$, and given piecewise constant densities $\tilde{p}_{0,m}$, $\tilde{p}_{1,m}$, and $\overline{p}_{\tau,m}$ on $[T,0] \times E$, where $\overline{p}_{\tau,m}$ is the approximation of $p_\tau$ as described in Remark 2.10. Define the constraint error $d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m}) =: d^{(m)} \in \mathbb{R}^{n+1}$

$$
\begin{aligned}
d_k^{(m)} &= Rf_k - \int_{[0,T]\times E} \widehat{f}_k^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx) + \int_{[0,T]\times E} Af_k(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt\times dx) \\
&\quad + \int_{[0,T]\times E} Bf_k(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt\times dx) \qquad \text{for } k=1,\dots,n \\
d_{n+1}^{(m)} &= 1 - \int_{[0,T]\times E} \overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx) = 1 - \left\|\overline{p}_{\tau,m}\right\|_{L^1([0,T]\times E)}.
\end{aligned}
$$

Note that in both definitions the integrals with respect to the stopping measure can be simplifed. For $k = 1,\dots,n$, we have

$$
\begin{aligned}
C_{k,j}^{(m)} = \int_l \widehat{f}_k^{(m)}(t,x)p_j^{(l)}(t,x)\lambda_l(dt\times dx) &= \int_l \sum_{i=0}^{4^m-1} f_{k,i}^{(m)} \mathbb{1}_{R_i}(t,x)p_j^{(l)}(t,x)\lambda_l(dt\times dx) \\
&= \int_l f_{k,j}^{(m)} p_j^{(l)}(t,x)\lambda_l(dt\times dx) \\
&= f_{k,j}^{(m)} \int_l p_j^{(l)}(t,x)\lambda_l(dt\times dx) \\
&= f_{k,j}^{(m)} \int_l \mathbb{1}_{L_j}(t,x)\lambda_l(dt\times dx) \\
&= f_{k,j}^{(m)}\lambda_l(L_j).
\end{aligned}
$$

Recall the density $\tilde{p}_{\tau,m}$ defined on the stopping boundary $l$ has the representation

$$
\tilde{p}_{\tau,m}(t,x) = \sum_{j=0}^{4^m-1} w_{\tau,j} p_j^{(l)}(t,x),
$$

and the density $\overline{p}_{\tau,m}$ defined on the entire space $[0,T]\times E$ using $p_j = \mathbb{1}_{R_j}$ $(j = 0,\dots,4^m-1)$ is

$$
\overline{p}_{\tau,m}(t,x) = \sum_{j=0}^{4^m-1} w_{\tau,j} \frac{\lambda_l(L_j)}{\lambda_2(R_j)} p_j(t,x).
$$

Hence, we can rewrite

$$\begin{aligned}\int_{[0,T]\times E} \widehat{f}_k^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx) &= \sum_{j=0}^{4^m-1}\int_{R_j} f_{k,j}^{(m)} w_{\tau,j}\frac{\lambda_l(L_j)}{\lambda_2(R_j)} p_j(t,x)\lambda_2(dt\times dx)\\ &= \sum_{j=0}^{4^m-1} w_{\tau,j} f_{k,j}^{(m)}\frac{\lambda_l(L_j)}{\lambda_2(R_j)}\int_{R_j} p_j(t,x)\lambda_2(dt\times dx)\\ &= \sum_{j=0}^{4^m-1} w_{\tau,j} f_{k,j}^{(m)}\lambda_l(L_j)\\ &= (C^{(m)}w_\tau)_k,\end{aligned}$$

where $w_\tau = (w_{\tau,0},\dots,w_{\tau,4^m-1})^T$ is the weight vector for $\tilde{p}_{\tau,m}$. And finally, according to Lemma 2.11,

$$\begin{aligned}d_{n+1}^{(m)} &= 1 - \int_{[0,T]\times E}\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\\ &= 1 - \left\|\overline{p}_{\tau,m}\right\|_{L^1([0,T]\times E)} = 1 - \|\tilde{p}_{\tau,m}\|_{L^1(l)} = 1-(C^{(m)}w_\tau)_{n+1}.\end{aligned}$$

**Remark 2.15.** Note that for a given $m\in\mathbb{N}$, a triple of measures $(\tilde{\mu}_{0,m},\tilde{\mu}_{1,m},\overline{\nu}_{\tau,m})$ is feasible for the $(n,m)$-dimensional if for its corresponding densities $(\tilde{p}_{0,m},\tilde{p}_{1,m},\overline{p}_{\tau,m})$ we have $\left\|d(\tilde{p}_{0,m},\tilde{p}_{1,m},\overline{p}_{\tau,m})\right\|_\infty = 0$.

Now, let $(\mu_{0,n},\mu_{1,n},\nu_{\tau,n})\in\mathcal{M}_{(n,\infty)}$. Since $\mathcal{D}_n$ has $n$ basis functions $\{f_1,\dots,f_n\}$, we have a finite number of integrals to approximate in the semi-infinite LP. We can therefore use Lemmas 2.6 and 2.8 to find a large enough $m$ such that the the approximations for each of the three densities are good enough. For a given $\varepsilon$, and by picking an appropriate $\varepsilon'$ for each integral, the lemmas establish the existence of a largest refinement level $m$, the corresponding partition $\Pi_m$, and approximating measures $(\tilde{\mu}_{0,m},\tilde{\mu}_{1,m},\overline{\nu}_{\tau,m})$ such that $\|d^{(m)}\|_\infty<\varepsilon$ holds. The next lemma proves this observation. To simplify notation, we will for the rest of the section drop the index $n$ from the notations for the measures in $\mathcal{M}_{(n,\infty)}$.

**Lemma 2.16.** *Let $(\mu_0,\mu_1,\nu_\tau)\in\mathcal{M}_{(n,\infty)}$ with densities $p_0,p_1,$ and $p_\tau$. Then for any $\varepsilon>0$ and $D>0$ there is an $0<\overline{\varepsilon}<\varepsilon$ and $m_0\in\mathbb{N}$ such that for each $m\ge m_0$, there are piecewise* *constant densities* $\tilde{p}_{0,m}$*,* $\tilde{p}_{1,m}$*, and* $\tilde{p}_{\tau,m}$ *as defined above satisfying* $\tilde{p}_{\tau,m} \geq D \cdot \overline{\varepsilon}$ *and*

$$\|d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})\|_\infty < \overline{\varepsilon},$$

$$|J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m})| < \overline{\varepsilon}.$$

*Proof.* Apply Lemmas 2.8 and 2.11 with $\varepsilon$ and $D' = 3D$, the given $\nu_\tau$, the payoff function at the final time $c_2$, and the set of basis functions for $\mathcal{D}_n$, $\{f_k\}_{k=1}^n$, to find an $0 < \widehat{\varepsilon} < \varepsilon$ and an $m_1 \in \mathbb{N}$. Set $\overline{\varepsilon} = \frac{\widehat{\varepsilon}}{D}$. Hence, $\tilde{p}_{\tau,m} \geq D \cdot \varepsilon$. Next, apply Lemma 2.6 with $\frac{\overline{\varepsilon}}{3}$, once for $\mu_0$ with the running payoff function $c_0$ and the set of functions $\{Af_k\}_{k=1}^n$, and second for $\mu_1$ with the singular payoff function $c_1$ and the set of function $\{Bf_k\}_{k=1}^n$ to find an $m_2 \geq m_1$. Then set $m_0 = m_2$ and choose $m \geq m_0$ with the corresponding approximating measures $\tilde{\mu}_{0,m}$, $\tilde{\mu}_{1,m}$, and $\tilde{\nu}_{\tau,m}$, as well as the corresponding $\overline{\nu}_{\tau,m}$ having density $\overline{p}_{\tau,m}$. Thus we have for all $k = 1, \ldots, n$

$$\begin{aligned}
\left| \int_{[0,T]\times E} Af_k(t,x)p_0(t,x)\lambda_2(dt \times dx) - \int_{[0,T]\times E} Af_k(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt \times dx) \right| &< \frac{\overline{\varepsilon}}{3}, \\
\left| \int_{[0,T]\times E} Bf_k(t,x)p_1(t,x)\lambda_1(dt \times dx) - \int_{[0,T]\times E} Bf_k(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt \times dx) \right| &< \frac{\overline{\varepsilon}}{3}, \\
\left| \int_l f_k(t,x)p_\tau(t,x)\lambda_l(dt \times dx) - \int_{[0,T]\times E} \widehat{f}_k^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx) \right| &< \frac{\overline{\varepsilon}}{3},
\end{aligned} \tag{2.6}$$

as well as

$$\begin{aligned}
\left| \int_{[0,T]\times E} c_0(t,x)p_0(t,x)\lambda_2(dt \times dx) - \int_{[0,T]\times E} c_0(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt \times dx) \right| &< \frac{\overline{\varepsilon}}{3}, \\
\left| \int_{[0,T]\times E} c_1(t,x)p_1(t,x)\lambda_1(dt \times dx) - \int_{[0,T]\times E} c_1(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt \times dx) \right| &< \frac{\overline{\varepsilon}}{3}, \\
\left| \int_l c_2(t,x)p_\tau(t,x)\lambda_l(dt \times dx) - \int_{[0,T]\times E} c_2(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx) \right| &< \frac{\overline{\varepsilon}}{3}.
\end{aligned} \tag{2.7}$$

From (2.7), it follows directly, that

$$\begin{aligned}
&|J(\mu_0,\mu_1,\nu_\tau) - J(\tilde{\mu}_{0,m},\tilde{\mu}_{1,m},\overline{\nu}_{\tau,m})| \\
&\leq \left|\int_{[0,T]\times E} c_0(t,x)p_0(t,x)\lambda_2(dt\times dx) - \int_{[0,T]\times E} c_0(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt\times dx)\right| \\
&+ \left|\int_{[0,T]\times E} c_1(t,x)p_1(t,x)\lambda_1(dt\times dx) - \int_{[0,T]\times E} c_1(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt\times dx)\right| \\
&+ \left|\int_l c_2(t,x)p_\tau(t,x)\lambda_l(dt\times dx) - \int_{[0,T]\times E} c_2(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)\right| \\
&< \frac{\overline{\varepsilon}}{3} + \frac{\overline{\varepsilon}}{3} + \frac{\overline{\varepsilon}}{3} = \overline{\varepsilon}.
\end{aligned}$$

It remains to show that the constraint error is small. Recall the definition of the constraint error $d^{(m)} = d(\tilde{p}_{0,m},\tilde{p}_{1,m},\overline{p}_{\tau,m})$

$$\begin{aligned}
d_k^{(m)} &= Rf_k - \int_{[0,T]\times E} \widehat{f}_k^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx) + \int_{[0,T]\times E} Af_k(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt\times dx) \\
&\quad + \int_{[0,T]\times E} Bf_k(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt\times dx) \qquad \text{for } k=1,\dots,n \\
d_{n+1}^{(m)} &= 1 - \int_{[0,T]\times E} \overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx)
\end{aligned}$$

First, we show

$$d_{n+1}^{(m)} = 1 - \left\|\overline{p}_{\tau,m}\right\|_{L^1([0,T]\times E)} = 1 - \|\tilde{p}_{\tau,m}\|_{L^1(l)} \geq 1 - \|p_\tau\|_{L^1(l)} - \|p_\tau - \tilde{p}_{\tau,m}\|_{L^1(l)} > -\overline{\varepsilon},$$

and using the reverse triangle inequality,

$$d_{n+1}^{(m)} = 1 - \|\tilde{p}_{\tau,m}\|_{L^1(l)} \leq 1 - \|p_\tau\|_{L^1(l)} + \|p_\tau - \tilde{p}_{\tau,m}\|_{L^1(l)} < \overline{\varepsilon},$$

proving that $\left|d_{n+1}^{(m)}\right| < \overline{\varepsilon}$. Since $(\mu_0,\mu_1,\nu_\tau) \in \mathcal{M}_{(n,\infty)}$ for the semi-infinite linear program, these measures satisfy the constraint equations for $k = 1,\dots,n$

$$Rf_k = \int f_k d\nu_\tau - \int Af_k d\mu_0 - \int Bf_k d\mu_1.$$

And hence, we have for the absolute value of $k$-th constraint error $d_k^{(m)}$ $(k = 1, \dots, n)$

$$\begin{aligned}\left|d_k^{(m)}\right| &= \Big|Rf_k - \int_{[0,T]\times E} \widehat{f}_k^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx) \\ &\quad + \int_{[0,T]\times E} Af_k(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt \times dx) + \int_{[0,T]\times E} Bf_k(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt \times dx)\Big| \\ &\leq \left|\int_{[0,T]\times E} f_k(t,x)p_\tau(t,x)\lambda_l(dt \times dx) - \int_{[0,T]\times E} \widehat{f}_k^{(m)}(t,x)\overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx)\right| \\ &\quad + \left|\int_{[0,T]\times E} Af_k(t,x)p_0(t,x)\lambda_2(dt \times dx) - \int_{[0,T]\times E} Af_k(t,x)\tilde{p}_{0,m}(t,x)\lambda_2(dt \times dx)\right| \\ &\quad + \left|\int_{[0,T]\times E} Bf_k(t,x)p_1(t,x)\lambda_1(dt \times dx) - \int_{[0,T]\times E} Bf_k(t,x)\tilde{p}_{1,m}(t,x)\lambda_1(dt \times dx)\right|.\end{aligned}$$

It follows by (2.6) for $k = 1, \dots, n$

$$\left|d_k(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})\right| < \frac{\overline{\varepsilon}}{3} + \frac{\overline{\varepsilon}}{3} + \frac{\overline{\varepsilon}}{3} = \overline{\varepsilon}.$$

Thus we have shown that $\left\|d^{(m)}\right\|_\infty < \overline{\varepsilon}$, finishing the proof. □

Now, we prove the existence of densities satisfying the conclusions of Lemma 2.16, which also guarantee the existence of a solution to an important linear equation involving the constraint matrix and constraint error. This will help us construct a feasible solution to the finite-dimensional linear program from the set of $\varepsilon$-close approximations.

**Lemma 2.17.** *For any $\varepsilon > 0$ we can find an $0 < \tilde{\varepsilon} < \varepsilon$ and an $m_0 \in \mathbb{N}$ large enough such that for any $m \geq m_0$, there are densities $\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m} \in span\{p_0, p_1, \dots, p_{4^m-1}\}$ on $[T, 0] \times E$ satisfying the conclusions of Lemma 2.16, as well as providing the existence of a solution $\tilde{y}$ to the equation $C^{(m)}y = -d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})$ satisfying $\|\tilde{y}\|_\infty < \tilde{\varepsilon}$.*

*Proof.* Recall that for any given $n$, we can find an $m_1 \in \mathbb{N}$ large enough, that for all $m \geq m_1$ $C^{(m)}$ has full rank of $n + 1$, i.e. we can find $n + 1$ linearly independent columns. In fact, we require a further restriction on the choice of columns. Recall that each column of the constraint matrix $C^{(m)}$ corresponds to a rectangle of the grid. Thus, for the given stopping boundary $l$, we require a choice of $n + 1$ independent columns with corresponding rectangles of the partition whose intersections with the stopping boundary $L_i$ have positive

$\lambda_l$-Lebesgue measure. This guarantees that the corresponding $n+1$ weights $w_{\tau,i}\frac{\lambda_l(L_i)}{\lambda_2(R_i)}$ of $\overline{p}_{\tau,m}$ are greater than zero. Choose $m_1$ and define $\overline{C}^{(m_1)} \in \mathbb{R}^{n+1,n+1}$ to be the matrix comprised of these $n+1$ independent columns. Apply Lemma 2.16 with $\varepsilon_d = \frac{\varepsilon}{\max\left\{1, \left\|\left(\overline{C}^{(m_1)}\right)^{-1}\right\|_\infty\right\}}$ and $D = \left\|\left(\overline{C}^{(m_1)}\right)^{-1}\right\|_\infty$ to find $m_2 \geq m_1$ and $0 < \tilde{\varepsilon}_d < \varepsilon_d$ such that for all $m \geq m_2$ we can find densities $\tilde{p}_{0,m}, \tilde{p}_{1,m} \in \operatorname{span}\{p_0, p_1, \ldots, p_{4^m-1}\}$ and $\tilde{p}_{\tau,m} \in \operatorname{span}(p_0^{(l)}, p_1^{(l)}, p_{4^m-1}^{(l)})$ with the corresponding $\overline{p}_{\tau,m}$ as described in (2.4) satisfying

$$\|d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})\|_\infty < \tilde{\varepsilon}_d,$$

$$\left|J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})\right| < \tilde{\varepsilon}_d,$$

as well as $\tilde{p}_{\tau,m} \geq \tilde{\varepsilon}_d$. Set $\tilde{\varepsilon} = \left\|\left(\overline{C}^{(m_1)}\right)^{-1}\right\|_\infty \cdot \tilde{\varepsilon}_d < \varepsilon$ and define the solution $y^{(1)} = -\left(\overline{C}^{(m_1)}\right)^{-1} \cdot d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})$. Define a larger vector $y^{(2)} \in \mathbb{R}^{4^{m_1}}$ by filling up $y^{(1)}$ with zeros for all the columns of $C^{(m_1)}$ that are missing from $\overline{C}^{(m_1)}$. Hence,

$$-d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m}) = \overline{C}^{(m_1)} y^{(1)} = C^{(m_1)} y^{(2)}.$$

Those solutions $y^{(1)}$ and $y^{(2)}$ satisfy

$$\begin{aligned}\left\|y^{(2)}\right\|_\infty = \left\|y^{(1)}\right\|_\infty &= \left\|(\overline{C}^{(m_1)})^{-1} \cdot d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})\right\|_\infty \\ &\leq \left\|(\overline{C}^{(m_1)})^{-1}\right\|_\infty \cdot \left\|d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})\right\|_\infty \\ &< \left\|(\overline{C}^{(m_1)})^{-1}\right\|_\infty \cdot \tilde{\varepsilon}_d = \tilde{\varepsilon}.\end{aligned}$$

From this result, we now construct a solution $\tilde{y} \in \mathbb{R}^{4^m}$ to the extended linear equation $C^{(m)}y = -d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})$ for any $m \geq m_2 \geq m_1$ satisfying $\|\tilde{y}\|_\infty < \tilde{\varepsilon}$. Considering the dyadic decomposition of a one-dimensional grid, the constraint matrix in that case satisfies for any refinement level $m$

$$C_{k,2i-1}^{(m+1)} + C_{k,2i}^{(m+1)} = C_{k,i}^{(m)}$$

for any $k = 1, \ldots, n+1$ and $i = 0, \ldots, 2^m - 1$. Thus, if $y$ is a solution to $C^{(m)}y = -d$, then the vector $y' \in \mathbb{R}^{2^{m+1}}$ whose entries are defined by

$$y'_{2i-1} = y'_{2i} = y_i$$

for $i = 0, \ldots, 2^m - 1$ satisifes $C^{(m+1)} y' = -d$, as well as $\|y'\| = \|y\| < \tilde{\varepsilon}$. With each increment of $m$, we bisect the one-dimensional intervals, and thus quarter the rectangles of the two-dimensional grid. Hence, in a similar manner, we can extend the solution from $m$ to $m+1$ for the two-dimensional grid. Iterating this construction of solutions from $m_1$ to the desired $m$ completes the proof. □

The previous lemmas established approximations of the given solutions $(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_{(n,\infty)}$. However, we have not yet discussed the feasibility of the approximations for the finite-dimensional LP. The next lemma adjusts the $\tilde{\nu}_{\tau,m}$ measure and the corresponding $\overline{\nu}_{\tau,m}$ to obtain a feasible solution to the finite-dimensional LP while essentially maintaining the approximation results. Similar to $\tilde{\nu}_{\tau,m}$, $\overline{\nu}_{\tau,m}$ is defined by $d\overline{\nu}_{\tau,m} = \overline{p}_{\tau,m} d\lambda_2$.

**Lemma 2.18.** *For any $\varepsilon > 0$ and a given element $(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_{(n,\infty)}$ we can find an $m_0 \in \mathbb{N}$ such that for each $m \geq m_0$ there is an element $(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m}) \in \mathcal{M}_{(n,m)}$ satisfying*

$$|J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m})| < \varepsilon$$

*Proof.* Apply Lemma 2.17 with $\varepsilon' = \min(\frac{\varepsilon}{2}, \frac{\varepsilon}{2\|c_2\|_\infty \lambda_l(l)})$ to find an $\widehat{\varepsilon} < \varepsilon'$, an $\overline{m} \in \mathbb{N}$, and measures $\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m}$ defined as

$$\tilde{\mu}_{0,m}(dt \times dx) = \tilde{p}_{0,m}(t,x)\lambda_2(dt \times dx),$$

$$\tilde{\mu}_{1,m}(dt \times dx) = \tilde{p}_{1,m}(t,x)\lambda_1(dt \times dx),$$

$$\overline{\nu}_{\tau,m}(dt \times dx) = \overline{p}_{\tau,m}(t,x)\lambda_2(dt \times dx),$$

such that for all $m \geq \overline{m}$ we have $|J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m})| < \widehat{\varepsilon}$. and we have a solution $\tilde{y}$ to the linear equation $C^{(m)} y = -d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m})$ such that $\|\tilde{y}\|_\infty < \widehat{\varepsilon}$. It remains to show that we can find a feasible solution whose value is $\varepsilon$-close to the value of the given solution to the semi-finite LP. To achieve that, we will adjust the weights of $\overline{p}_{\tau,m}$ such that the constraint error $\left\|d^{(m)}\right\|_\infty = 0$. Recall that the approximating density $\tilde{p}_{\tau,m}$ is of the form

$$\tilde{p}_{\tau,m} = \sum_i \tilde{w}_{\tau,i} p_i,$$

where $p_i = \mathbb{1}_{R_i}$ and each $\tilde{w}_{\tau,i} \geq \widehat{\varepsilon}$. $\tilde{p}_{0,m}$ and $\tilde{p}_{1,m}$ have a similar form but need not be adjusted. Note that at this point, we can not guarantee feasibility of the approximations $\overline{p}_{\tau,m}$, $\tilde{p}_{0,m}$, and $\tilde{p}_{1,m}$. Thus, we construct a modified set of densities $\tilde{p}^*_{0,m}, \tilde{p}^*_{1,m}, \overline{p}^*_{\tau,m}$ that is feasible and whose value function is still $\varepsilon$-close with respect to the given measures $\mu_0, \mu_1, \nu_\tau$. It will be enough to modify the approximation of the stopping measure $\overline{p}_{\tau,m}$, while keeping the other two approximations. To this end, define for each $i$, $w^*_{\tau,i} = \tilde{w}_{\tau,i} - \tilde{y}_i$ and a new density

$$\tilde{p}^*_{\tau,m} = \sum_i w^*_{\tau,i} p_i,$$

as well as the corresponding $\overline{p}^*_{\tau,m}$ and its measure $\overline{\nu}^*_{\tau,m}$ as described in Remark 2.10. To finish the proof, we need to show the following

1. For the triple of measures $(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m})$, the value function is close to the value function of the original triple of measures $\mu_0, \mu_1, \nu_\tau$, i.e.

$$\left| J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m}) \right| < \varepsilon.$$

2. All three approximated densities $(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}^*_{\tau,m})$ are positive.

3. The triple $(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m})$ is feasible, i.e. $\left\| d(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}^*_{\tau,m}) \right\|_\infty = 0$.

4. All three approximated measures are bounded: $\tilde{\mu}_{0,m}([0,T] \times E) \leq T$, $\tilde{\mu}_{1,m}([0,T] \times E) \leq l$, and $\overline{\nu}^*_{\tau,m}([0,T] \times E) = 1$.

For part 1 we have

$$\begin{aligned}
&\left| J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m}) \right| \\
&= \left| \int c_0 d\mu_0 - \int c_0 d\tilde{\mu}_{0,m} + \int c_1 d\mu_1 - \int c_1 d\tilde{\mu}_{1,m} + \int c_2 d\nu_\tau - \int c_2 d\overline{\nu}^*_{\tau,m} \right| \\
&\leq \left| \int c_0 d\mu_0 - \int c_0 d\tilde{\mu}_{0,m} + \int c_1 d\mu_1 - \int c_1 d\tilde{\mu}_{1,m} + \int c_2 d\nu_\tau - \int c_2 d\overline{\nu}_{\tau,m} \right| \\
&\quad + \left| \int c_2 \sum_i \tilde{y}_i \frac{\lambda_l(L_i)}{\lambda_2(R_i)} p_i d\lambda_2 \right| \\
&= \left| J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m}) \right| + \left| \int c_2 \sum_i \tilde{y}_i \frac{\lambda_l(L_i)}{\lambda_2(R_i)} p_i d\lambda_2 \right|
\end{aligned}$$

Lemma 2.17 guarantees that

$$|J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m})| < \widehat{\varepsilon} < \frac{\varepsilon}{2}.$$

Consider the second term:

$$\left| \int c_2 \sum_i \tilde{y}_i \frac{\lambda_l(L_i)}{\lambda_2(R_i)} p_i d\lambda_2 \right| \leq \|c_2\|_\infty \lambda_l(l) \|\tilde{y}\|_\infty < \|c_2\|_\infty \lambda_l(l) \widehat{\varepsilon} < \|c_2\|_\infty \lambda_l(l) \frac{\varepsilon}{2 \|c_2\|_\infty \lambda_l(l)} = \frac{\varepsilon}{2}.$$

Thus, we have shown

$$\left| J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m}) \right| < \varepsilon.$$

For part 2, we already know, that by construction, $\tilde{p}_{0,m} \geq 0$ and $\tilde{p}_{1,m} \geq 0$. It remains to show that for each $i$, $w^*_{\tau,i} \geq 0$. Recall that according to Lemma 2.17, $\tilde{p}_{\tau,m} \geq \widehat{\varepsilon}$, i.e. $\min_i \tilde{w}_{\tau,i} \geq \widehat{\varepsilon}$, as well as $\|\tilde{y}\|_\infty \leq \widehat{\varepsilon}$. Thus, it follows for every $i = 0, \ldots, 4^m - 1$ that

$$w^*_{\tau,i} = \tilde{w}_{\tau,i} - \tilde{y}_i \geq \widehat{\varepsilon} - \widehat{\varepsilon} = 0,$$

and due to the choice of columns and the construction of $\overline{p}^*_{\tau,m}$, $\overline{p}^*_{\tau,m} \geq 0$.

For part 3 and $k = n + 1$, we have

$$\begin{aligned} d_{n+1}(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}^*_{\tau,m}) &= 1 - (C^{(m)} w^*_\tau)_{n+1} \\ &= (1 - (C^{(m)} \tilde{w}_\tau)_{n+1}) + (C^{(m)} \tilde{y})_{n+1} \\ &= d_{n+1}(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m}) - d_{n+1}(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m}) = 0. \end{aligned}$$

For $k = 1, \ldots, n$, on the other hand, we have

$$\begin{aligned} d_k(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}^*_{\tau,m}) &= Rf_k - (C^{(m)}w^*_\tau)_k + \int_{[0,T]\times E} Af_k\tilde{p}_{0,m}(t,x)\lambda_2(dt\times dx) \\ &\quad + \int_{[0,T]\times E} Bf_k\tilde{p}_{1,m}(t,x)\lambda_1(dt\times dx) \\ &= Rf_k - (C^{(m)}(\tilde{w}_\tau - \tilde{y})_k + \int_{[0,T]\times E} Af_k\tilde{p}_{0,m}(t,x)\lambda_2(dt\times dx) \\ &\quad + \int_{[0,T]\times E} Bf_k\tilde{p}_{1,m}(t,x)\lambda_1(dt\times dx) \\ &= Rf_k - \int_{[0,T]\times E} \widehat{f}^{(m)}_k\overline{p}_{\tau,m}(t,x)\lambda_2(dt\times dx) + \int_{[0,T]\times E} Af_k\tilde{p}_{0,m}(t,x)\lambda_2(dt\times dx) \\ &\quad + \int_{[0,T]\times E} Bf_k\tilde{p}_{1,m}(t,x)\lambda_1(dt\times dx) + (C^{(m)}\tilde{y})_k \\ &= d_k(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m}) - d_k(\tilde{p}_{0,m}, \tilde{p}_{1,m}, \overline{p}_{\tau,m}) = 0 \end{aligned}$$

And finally, for part 4, by construction, $\tilde{p}_{0,m} \le p_0$ and $\tilde{p}_{1,m} \le p_1$, and therefore $\tilde{\mu}_{0,m}([0,T]\times E) \le \mu_0([0,T]\times E) \le T$ and $\tilde{\mu}_{1,m}([0,T]\times E) \le \mu_1([0,T]\times E) \le l$. And since we have already shown the approximating measures to be feasible, we have $\overline{\nu}^*_{\tau,m}([0,T]\times E) = 1$. Thus, we have shown that we can find an $m \in \mathbb{N}$ and a triple of measures $(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m}) \in \mathcal{M}_{(n,m)}$ such that

$$\left|J(\mu_0,\mu_1,\nu_\tau) - J(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}^*_{\tau,m})\right| < \varepsilon$$

□

From here on out, we mimic the approximation results from the infinite-dimensional LP by the semi-infinite LP. We start by showing that the limit of a weakly converging sequence of solutions to the finite LP is a solution for the semi-infinite LP.

**Lemma 2.19.** *Let $\{\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m}\}$ be a sequence of measures with $(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m}) \in \mathcal{M}_{(n,m)}$ for all $m \in \mathbb{N}$ and assume there are measures $\mu_0, \mu_1, \nu_\tau$ such that $\tilde{\mu}_{0,m} \Rightarrow \mu_0, \tilde{\mu}_{1,m} \Rightarrow \mu_1, \overline{\nu}_{\tau,m} \Rightarrow \nu_\tau$ as $m \to \infty$. Then $(\mu_0,\mu_1,\nu_\tau) \in \mathcal{M}_{(n,\infty)}$.*

*Proof.* Let $f_n \in \mathcal{D}_n$ be fixed and $\widehat{f}_n^{(m)}$ be the approximation of $f_n$ as defined in (1.8). Since for each $m \in \mathbb{N}$ $(\tilde{\mu}_{0,m}, \tilde{\mu}_{1,m}, \overline{\nu}_{\tau,m}) \in \mathcal{M}_{(n,m)}$, we have

$$Rf_n = \int \widehat{f}_n^{(m)} d\overline{\nu}_{\tau,m} - \int Af_n d\tilde{\mu}_{0,m} - \int Bf_n d\tilde{\mu}_{1,m}$$

Due to $Af_n$, and $Bf_n$ being bounded and uniformly continuous, weak convergence of the measures implies

$$\lim_{m\to\infty} \int Af_n d\tilde{\mu}_{0,m} = \int Af_n d\mu_0,$$
$$\lim_{m\to\infty} \int Bf_n d\tilde{\mu}_{1,m} = \int Bf_n d\mu_1.$$

Similarly we can apply Theorem 3.5 in Serfozo (1982) to get

$$\lim_{m\to\infty} \int \widehat{f}_n^{(m)} d\overline{\nu}_{\tau,m} = \int f_n d\nu_\tau.$$

Therefore,

$$\begin{aligned}\int f_n d\nu_\tau - \int Af_n d\mu_0 - \int Bf_n d\mu_1 &= \lim_{m\to\infty} \left( \int \widehat{f}_n^{(m)} d\overline{\nu}_{\tau,m} - \int Af_n d\tilde{\mu}_{0,m} - \int Bf_n d\tilde{\mu}_{1,m} \right) \\ &= \lim_{m\to\infty} Rf_n = Rf_n.\end{aligned}$$

Since the constant function 1 is bounded and uniformly continuous, we also have

$$\begin{aligned}\mu_1([0,T]\times E) &= \int_{[0,T]\times E} 1 d\mu_1 = \lim_{m\to\infty} \int_{[0,T]\times E} d\tilde{\mu}_{1,m} = \lim_{m\to\infty} \tilde{\mu}_{1,m}([0,T]\times E) \le l, \\ \mu_0([0,T]\times E) &= \int_{[0,T]\times E} 1 d\mu_0 = \lim_{m\to\infty} \int_{[0,T]\times E} d\tilde{\mu}_{0,m} = \lim_{m\to\infty} \tilde{\mu}_{0,m}([0,T]\times E) \le T, \\ \nu_\tau([0,T]\times E) &= \int_{[0,T]\times E} 1 d\nu_\tau = \lim_{m\to\infty} \int_{[0,T]\times E} d\overline{\nu}_{\tau,m} = \lim_{m\to\infty} \overline{\nu}_{\tau,m}([0,T]\times E) = 1.\end{aligned}$$

Hence, the triple of the weak limits $(\mu_0, \mu_1, \nu_\tau)$ is an element of $\mathcal{M}_{(n,\infty)}$, completing the proof. □

To streamline notation, we will hereafter refer to $\overline{\nu}_{\tau,m}$ as $\tilde{\nu}_{\tau,m}$. The following lemma shows that the limit of a weakly converging sequence of optimal solutions to the finite LPs is an optimal solution of the semi-infinite LP.

**Proposition 2.20.** *Let* $\{(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})\}_{m=1}^{\infty}$ *be a sequence of optimal measures with* $(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}) \in \mathcal{M}_{(n,m)}$ *for all* $m \in \mathbb{N}$ *and assume there are measures* $(\mu^*_0, \mu^*_1, \nu^*_\tau) \in \mathcal{M}_{(n,\infty)}$ *such that* $\tilde{\mu}^*_{0,m} \Rightarrow \mu^*_0, \tilde{\mu}^*_{1,m} \Rightarrow \mu^*_1, \tilde{\nu}^*_{\tau,m} \Rightarrow \nu^*_\tau$ *as* $m \to \infty$. *Then*

$$J(\mu^*_0, \mu^*_1, \nu^*_\tau) = \sup_{\mu_0,\mu_1,\nu_\tau \in \mathcal{M}_{(n,\infty)}} J(\mu_0, \mu_1, \nu_\tau) =: J^*.$$

*Proof.* Assume this is not the case. Since Lemma 2.19 establishes $(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}) \in \mathcal{M}_{(n,\infty)}$, the assumption implies

$$J(\tilde{\mu}^*_0, \tilde{\mu}^*_1, \tilde{\nu}^*_\tau) < J^*.$$

Hence, we can find a fixed $\varepsilon > 0$ and $(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_{(n,\infty)}$ such that

$$J(\tilde{\mu}^*_0, \tilde{\mu}^*_1, \tilde{\nu}^*_\tau) + \varepsilon \leq J(\mu_0, \mu_1, \nu_\tau).$$

Weak convergence of $(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})$ to $(\tilde{\mu}^*_0, \tilde{\mu}^*_1, \tilde{\nu}^*_\tau)$ implies that for some $m_0 \in \mathbb{N}$ large enough,

$$J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}) + \frac{\varepsilon}{2} < J(\mu_0, \mu_1, \nu_\tau) \quad \text{for } m \geq m_0. \tag{2.8}$$

However, by Proposition 2.18, for this specific choice of $\varepsilon$ and $(\mu_0, \mu_1, \nu_\tau)$, we can find an $m_1 \geq m_0$ and a solution $(\tilde{\mu}_{0,m_1}, \tilde{\mu}_{1,m_1}, \tilde{\nu}_{\tau,m_1}) \in \mathcal{M}_{(n,m_1)}$ satisfying

$$|J(\mu_0, \mu_1, \nu_\tau) - J(\tilde{\mu}_{0,m_1}, \tilde{\mu}_{1,m_1}, \tilde{\nu}_{\tau,m_1})| < \frac{\varepsilon}{2}. \tag{2.9}$$

Combining (2.8) and (2.9), we obtain using $m = m_1$

$$J(\tilde{\mu}^*_{0,m_1}, \tilde{\mu}^*_{1,m_1}, \tilde{\nu}^*_{\tau,m_1}) < J(\mu_0, \mu_1, \nu_\tau) - \frac{\varepsilon}{2} < J(\tilde{\mu}_{0,m_1}, \tilde{\mu}_{1,m_1}, \tilde{\nu}_{\tau,m_1}),$$

contradicting the optimality of $(\tilde{\mu}^*_{0,m_1}, \tilde{\mu}^*_{1,m_1}, \tilde{\nu}^*_{\tau,m_1})$. Thus the assumption

$$J(\tilde{\mu}^*_0, \tilde{\mu}^*_1, \tilde{\nu}^*_\tau) < J^*$$

is false, completing the proof. □

From here on out, we omit the assumption of weak convergence. We first show that the values of any sequence of optimal solutions to the finite LPs converge to the optimal value of the semi-infinite LP, followed by the central theorem for our treatment of the finite LPs.

**Proposition 2.21.** *Let* $\{\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}\}_{m\in\mathbb{N}}$ *be a sequence of optimal measures such that* $(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}) \in \mathcal{M}_{(n,m)}$ *for all* $m \in \mathbb{N}$. *Then the sequence of values* $\{J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})\}_{m\in\mathbb{N}}$ *converges to* $J^* := \sup_{\mu_0,\mu_1,\nu_\tau\in\mathcal{M}_{(n,\infty)}} J(\mu_0, \mu_1, \nu_\tau)$.

*Proof.* Recall that $\{J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})\}_{m\in\mathbb{N}}$ is increasing and bounded above. Thus it converges. Since $\{\tilde{\mu}^*_{0,m}\}_{m\in\mathbb{N}}$, $\{\tilde{\mu}^*_{1,m}\}_{m\in\mathbb{N}}$, and $\{\tilde{\nu}^*_{\tau,m}\}_{m\in\mathbb{N}}$ are tight and uniformly bounded sequences of measures, we can find a converging subsequence $\{(\tilde{\mu}^*_{0,m_j}, \tilde{\mu}^*_{1,m_j}, \tilde{\nu}^*_{\tau,m_j})\}_{j\in\mathbb{N}}$ and the corresponding limits $(\tilde{\mu}^*_0, \tilde{\mu}^*_1, \tilde{\nu}^*_\tau) \in \mathcal{M}_{(n,\infty)}$, i.e.

$$\tilde{\mu}^*_{0,m_j} \Rightarrow \tilde{\mu}^*_0, \qquad \tilde{\mu}^*_{1,m_j} \Rightarrow \tilde{\mu}^*_1, \qquad \tilde{\nu}^*_{\tau,m_j} \Rightarrow \tilde{\nu}^*_\tau,$$

as $j \to \infty$. By Proposition 2.20, and under the assumption that $c_0$, $c_1$, and $c_2$ are bounded and uniformly continuous functions,

$$\begin{aligned} J^* = J(\tilde{\mu}^*_0, \tilde{\mu}^*_1, \tilde{\nu}^*_\tau) &= \int c_2 d\tilde{\nu}^*_\tau + \int c_0 d\tilde{\mu}^*_0 + \int c_1 d\tilde{\mu}^*_1 \\ &= \lim_{j\to\infty} \int c_2 d\tilde{\nu}^*_{\tau,m_j} + \int c_0 d\tilde{\mu}^*_{0,m_j} + \int c_1 d\tilde{\mu}^*_{1,m_j} \\ &= \lim_{j\to\infty} J(\tilde{\mu}^*_{0,m_j}, \tilde{\mu}^*_{1,m_j}, \tilde{\nu}^*_{\tau,m_j}). \end{aligned}$$

Since we established above that $\{J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})\}_{m\in\mathbb{N}}$ converges and its subsequence $\{J(\tilde{\mu}^*_{0,m_j}, \tilde{\mu}^*_{1,m_j}, \tilde{\nu}^*_{\tau,m_j})\}_{j\in\mathbb{N}}$ converges to $J^*$, the main sequence converges to the same limit, i.e.

$$\lim_{m\to\infty} J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}) = J^*.$$

□

We now additionally omit the concept of convergence completely and show in the central theorem, that for any $\varepsilon > 0$ we can find an $m$ large enough such that optimal solutions of the finite-dimensional LP are $\varepsilon$-optimal for the semi-infinite LP.

**Theorem 2.22.** *For fixed $n \in \mathbb{N}$ and $\varepsilon > 0$ there is an $\overline{m} \in \mathbb{N}$ such that for all $m \geq \overline{m}$, if $(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})$ is an optimal solution for the $(n,m)$-dimensional linear program, then $(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})$ is $\varepsilon$-optimal for the $(n,\infty)$-dimensional linear program.*

*Proof.* Recall that for each $m \in \mathbb{N}$, $\mathcal{M}_{(n,m)} \subset \mathcal{M}_{(n,\infty)}$. And since, by Proposition 2.21, we have established that $J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m})$ is an increasing sequence converging to $J^* = \sup_{(\mu_0,\mu_1,\nu_\tau)\in\mathcal{M}_{(n,\infty)}} J(\mu_0, \mu_1, \nu_\tau)$, we can find for any fixed $\varepsilon > 0$ an $\overline{m} \in \mathbb{N}$ such that for any $m \geq \overline{m}$

$$0 \leq J(\tilde{\mu}^*_{0,m}, \tilde{\mu}^*_{1,m}, \tilde{\nu}^*_{\tau,m}) - J^* < \varepsilon.$$

□

The following remarks describe the the course of action for finding an $\varepsilon$-optimal value and detail the issues with the inclusion characteristics of the intermediate feasible sets.

**Remark 2.23.** For a fixed $\varepsilon > 0$, in order to find an $\varepsilon$-optimal solution for the original linear program $\mathcal{M}$, which has an infite amount of constraint equations and an infinite-dimensional space of measures, follow the structure of convergence laid out in this work:

1. Choose $\tilde{\varepsilon} > 0$ and $\tilde{\delta} > 0$ such that $2\tilde{\varepsilon} + \tilde{\delta} < \varepsilon$ to find $n \in \mathbb{N}$ large enough such that, by Theorem 2.5, the value of any $\tilde{\varepsilon}$-optimal solution in $\mathcal{M}_{(n,\infty)}$ is within $2\tilde{\varepsilon} + \tilde{\delta}$ of the optimal value $J^*$ for the infinite-dimensional linear program.

2. Finally, find an $m \in \mathbb{N}$ large enough such that any optimal solution in $\mathcal{M}_{(n,m)}$ is $\tilde{\varepsilon}$-optimal in $\mathcal{M}_{(n,\infty)}$.

**Remark 2.24.** Recall, that by construction, we get the following inclusions for the feasible sets of the various linear programs we discussed

1. $\mathcal{M}_\infty \subset \mathcal{M}_{(n,\infty)}$ since we are reducing the number of constraints for $\mathcal{M}_{(n,\infty)}$. In other words, not every solution to the semi-finite linear program is a solution to the inifite-dimensional linear program.

2. Due to the approximation of the constraint equations, there is no inclusion relation between $\mathcal{M}_{(n,\infty)}$ and $\mathcal{M}_{(n,m)}$. However, weak limits of feasible measures in $\mathcal{M}_{(n,m)}$ (as $m \to \infty$) will be feasible measures in $\mathcal{M}_{(n,\infty)}$.

Therefore, we can not be certain that the approximate solutions will solve the original linear program. In other words, we don't know if the stopping strategy we find numerically will be a valid strategy for the original optimal stopping problem.

# 3 AMERICAN FLOATING STRIKE LOOKBACK OPTIONS

In this chapter, we will describe the optimal stopping problem that determines the price of an American floating strike lookback option. The market consists of a riskless bond and a risky asset, a stock. We describe the mathematical structure of the market and identify the optimal stopping problem for the price of the option, which is governed by the expected payout in order to avoid arbitrage. Using changes of measure, we re-frame the problem in terms of a single diffusion with reflection. Using the equivalence proven in 1.4, we display the infinite dimensional linear program for the price of the option. We are going to focus on the call option, because it has an inherently compact domain.

## 3.1 Mathematical Framework

We consider the following setup or the American floating strike lookback options.

- **Filtered probability space**: $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0})$, where $\mathcal{F}_t = \sigma(W_s^{\mathbb{P}} : 0 \geq s \geq t)$ is the filtration and $W^{\mathbb{P}}$ is a $\mathbb{P}$-Brownian motion.

- **Bond Price**: The bond price process $B$ satisfies $dB_t = rB_t dt$ or

$$B_t = B_0 e^{rt},\ t \geq 0,$$

  where $B_0 > 0$ is the initial bond price and $r > 0$ the constant risk-free rate of return.

- **Stock Price**: The stock price process $S$ is a geometric Browian motion satisfying

$$dS_t = S_t(\mu\, dt + \sigma\, dW_t^{\mathbb{P}}),\ S_0 = s_0 > 0,\ t \geq 0,$$

  where $\mu \in \mathbb{R}$ is the local mean rate of return and $\sigma > 0$ the instantaneous volatility.

- **Finite horizon**: The American option has a finite horizon $T > 0$.

- **Exercise time**: The exercise or strike time is $0 \leq \tau \leq T$, where $\tau$ is an $\mathbb{F}$-stopping time.

The lookback nature of the option means the option value will be determined by either the running maximum or the running minimum of the stock price process. We concentrate on the floating strike lookback call option, whose payoff depends on the running minimum process $m$ satisfying

$$m_t = s_0 \wedge \min_{0 \leq s \leq t} S_s, \ \ t \geq 0.$$

(A similar analysis applies to the corresponding put option using the running maximum process.)

An American option can be exercised at any time before the finite horizon $T$. We call the time of exercise $0 \leq \tau \leq T$, the strike time. For the American floating strike lookback call option, the payoff at strike time $\tau$ is the difference of the stock price and the minimum price

$$S_\tau - m_\tau.$$

It is well know from the theory of mathematical finance that the price of the American floating strike lookback put option is the maximal expected present-value payoff under the risk-neutral measure $\mathbb{Q}$

$$\max_{\tau \in \mathcal{T}} \mathbb{E}_{\mathbb{Q}}[e^{-r\tau}(S_\tau - m_\tau)],$$

where $\mathcal{T}$ is the set of all stopping times less than or equal to $T$. Under the risk-neutral measure $\mathbb{Q}$, the stock price process $S$ also behaves like a geometric Brownian motion, but the rate of change is now $r$, the mean rate of return for the bond $B$ instead of $\mu$, and can be described by the stochastic differential equation

$$dS_t = S_t(r\,dt + \sigma\,dW_t^{\mathbb{Q}}), \ \ S_0 = s_0 > 0,$$

where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion. The behavior of the process therefore follows $S_t =$

$s_0 \exp\left\{\left(r - \frac{\sigma^2}{2}\right) t + \sigma W_t^{\mathbb{Q}}\right\}$ for $t \geq 0$. The expected payoff under $\mathbb{Q}$ can be written as

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[e^{-r\tau}(S_\tau - m_\tau)] &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r\tau} S_\tau \left(1 - \frac{m_\tau}{S_\tau}\right)\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{-r\tau}\left(s_0 e^{\left(r - \frac{\sigma^2}{2}\right)\tau + \sigma W_t^{\mathbb{Q}}}\right)\left(1 - \frac{m_\tau}{S_\tau}\right)\right] \\ &= s_0 \mathbb{E}_{\mathbb{Q}}\left[e^{-\frac{\sigma^2}{2}\tau + \sigma W_\tau^{\mathbb{Q}}}\left(1 - \frac{m_\tau}{S_\tau}\right)\right]\end{aligned}$$

Under a change of measure using the new measure $\tilde{\mathbb{P}}$ whose Radon-Nikodym density with respect to $\mathbb{Q}$ we define as

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{Q}} = e^{-\frac{\sigma^2}{2}\tau + \sigma W_\tau^{\mathbb{Q}}},$$

we can rewrite the expected payoff

$$\mathbb{E}_{\mathbb{Q}}[e^{-r\tau}(S_\tau - m_\tau)] = s_0 \mathbb{E}_{\tilde{\mathbb{P}}}\left[1 - \frac{m_\tau}{S_\tau}\right].$$

Shepp and Shiryaev (1993) invented the "Russian option", which is similar to the floating strike lookback put option. In a subsequent paper (Shepp and Shiryaev (1995)), they showed that the resulting quotient process is a geometric Brownian motion on $[1, \infty)$ with reflection at 1. Lutz (2007, Theorem 4.1) derived a similar proof for the floating strike lookback call option and found that, under a measure $\tilde{\mathbb{P}}$, the quotient process $X_t = \frac{m_t}{S_t}$ is a geometric Brownian motion on $(0, 1]$ with reflection at the boundary 1 whose stochastic differential equation is

$$dX_t = -X_t(r dt + \sigma dW_t^{\tilde{\mathbb{P}}}) - d\lambda_t,$$

where $\lambda$ is the local time of the process $X$ at 1.

The continuous and singular generators $A$ and $B$ of the process $X$ are therefore of the form

$$\begin{aligned} Af(x) &= -rxf'(x) + \frac{\sigma^2}{2}x^2 f''(x) \text{ and} \\ Bf(x) &= -f'(x). \end{aligned} \tag{3.1}$$

Note that the domain of the generators is $\mathcal{D} = C^2$ here.

The expected optimal payoff for the transformed optimal stopping problem and the process starting at $x$ can then be written as

$$V(x) = \max_{\tau \in \mathcal{T}} s_0 \mathbb{E}^x_{\tilde{\mathbb{P}}}[1 - X_\tau].$$

Hence, the optimal stopping problem for the American floating strike lookback call option, which depends on the stock price process $S$ and the running minimum price process $m$ has been reduced to an optimal stopping problem which only depends on the quotient process $X$:

$$\begin{aligned} \text{Maximize} \quad & s_0 \mathbb{E}_{\tilde{\mathbb{P}}}[1 - X_\tau] \\ \text{such that} \quad & dX_t = -X_t(rdt + \sigma dW^{\tilde{\mathbb{P}}}_t) - d\lambda_t, \\ & 0 \le t \le T, \\ & \tau \in \mathcal{T}. \end{aligned} \tag{3.2}$$

## 3.2 Equivalent Linear Program

In this section, we apply Theorem 1.4 to the example of the finite horizon American floating strike lookback call option.

Let the state space be $E = [0, 1]$, the compactification of $(0, 1]$, the time dimension $[0, T]$, and the resulting product space $[0, T] \times E$ as the domain of the augmented time-space process $(t, X_t)$. The continuous and singular generators of the process $X$ are given in (3.1). The payoff functions are $c_0 \equiv 0$, $c_1 \equiv 0$, and $c_2(x) = 1 - x$. Note that our payoffs are not time-dependent here.

Given the optimal stopping problem (3.2), the equivalent infinite-dimensional linear program is:

$$\begin{aligned} \text{Maximize} \quad & \int c_2 \, d\nu_\tau \\ \text{Subject to} \quad & 0 = \int \gamma f \, d\nu_\tau - \gamma(0) \int f \, d\nu_0 - \int \tilde{A} \left[\gamma f\right] \, d\mu_0 - \int \tilde{B} \left[\gamma f\right] \, d\mu_1 \\ & \forall \gamma \in C^1([0,T]), \quad \forall f \in \mathcal{D} \\ & \nu_\tau \in \mathcal{P}([0,T] \times E), \quad \mu_0, \mu_1 \in \mathcal{M}([0,T] \times E), \end{aligned}$$

with $\tilde{A}[\gamma f] = \gamma A f + \gamma' f$, $\tilde{B}[\gamma f] = \gamma B f$. Observe that $R[\gamma f] = \gamma(0) \int f d\nu_0$ here.

We now show that the collection of $\mu_1$ measures in the feasible set of the infinite-dimensional linear program is uniformly bounded.

**Proposition 3.1.** *Let $(\mu_0, \mu_1, \nu_\tau) \in \mathcal{M}_\infty$ be feasible. Then $\mu_1([0,T] \times E) \leq x_0$.*

*Proof.* Choose $\gamma \equiv 1$ and $f(x) = x$. Observe that $\tilde{A}[\gamma f](t,x) = Af(x) = -rx$ and $\tilde{B}[\gamma f](t,x) = -1$. Since $(\mu_0, \mu_1, \nu_\tau)$ is feasible, the constraint equation holds for the chosen $\gamma$ and $f$. Hence,

$$\begin{aligned} 0 = & \int \gamma(t) f(x) \, \nu_\tau(dt \times dx) - \gamma(0) \int f(x) \, \nu_0(dx) - \int \tilde{A} \left[\gamma f\right](t,x) \, \mu_0(dt \times dx) \\ & - \int \tilde{B} \left[\gamma f\right](t,x) \, \mu_1(dt \times dx) \\ = & \int x \, \nu_\tau(dt \times dx) - f(x_0) + \int rx \, \mu_0(dt \times dx) + \int 1 \, \mu_1(dt \times dx). \end{aligned}$$

Rearranging the terms gives

$$\mu_1([0,T] \times [0,1]) = x_0 - \int x \nu_\tau(dt \times dx) - \int rx \mu_0(dt \times dx) \leq x_0.$$

The last inequality holds since both integrals are non-negative and bounded above. □

# 4 OUTLOOK

We plan to apply the numerical scheme presented in this dissertation to the American floating strike lookback call option. Lutz (2007) showed numerical results for the a number of American floating strike lookback options with an approximation scheme using finite differences in the time-dimension and a basis constisting of piecewise linear functions, i.e. first order B-splines, on a shared grid for both the test functions and the densities, but did not provide a proof of convergence for the results. Applying the new scheme to this problem will provide a way of comparing the results, while also backing up the numerical solution with a proof of convergence. Lutz's work also provides results for the corresponding put option as well as the lookback call option with a dividend-paying stock. In the latter case, they find that the optimal strategy changes from "Always exercise at $T$" to a time- and value-dependent strategy. This can be verified with the existing scheme as presented here.

The quotient process for the put option, as shown by Lutz and Shepp and Shiryaev (1995) will evolve on the interval $[0, \infty)$, and hence have an unbounded state space. In order to apply the new numerical scheme to the corresponding optimal stopping problem, the scheme will have to be extended to unbounded state spaces. Vieten (2018) and Lutz (2007) provide an approach when applied to optimal control problems and optimal stopping, respectively, by introducing an upper bound. Vieten showed, that even with the added approximation step, the optimal solution to the finite linear program will still come arbitrarily close to the solution of the original LP by choosing an appropriate bound of the state space and sufficient refinements for the constraint space and solution space.

Another potential for application to examples is optimal stopping problems with state spaces in more than one dimension. The approximation scheme would not need to be modified much since the lemma guaranteeing sufficient approximation is formulated in a general enough manner to make the transition possible. A host of examples is given in Christensen et al. (2019).

In the course of the development of this approximation scheme, it is natural to treat the time dimension differently than the state space dimension. We will explore mixing a finite difference scheme in time with the finite element approach in space to improve numerical stability and efficiency.

In this dissertation, we restricted our investigation to optimal stopping problems on a rectangular compact space. A lot of the time, singular behavior will only occur on fairly simple boundaries, simplifying the approximation of the singular occupation measure. In order to expand the range of application, at least for the case of optimal stopping in a time-state space setting, we will explore problems, in which the process is reflected on a time-space boundary. A possible example is once more the American floating strike lookback option. The minimum process is a singular process. Thus, the combined two-dimensional process consisting of the stock process and the minimum process will be reflected on the diagonal line. We plan to compare the numerical results on this problem without dimension reduction to those applied to the problem as described in Chapter 3, which are using the dimension reduction.

In the course of the development of the approximation scheme, the variable nature of the stopping boundary proved to be a hurdle when it comes to finding a workable approximation scheme, and even more so for comparing singular measures. A more rigorous look at metrics, such as the Earth Mover's distance or the more general Wasserstein distances, on spaces of measures might yield a different view on the approximation problem.

# BIBLIOGRAPHY


Billingsley, Patrick (1999). *Convergence of probability measures*. Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. New York: John Wiley & Sons Inc., pp. x+277.

Bogachev, V. I. (2007). *Measure theory. Vol. I, II*. Berlin: Springer-Verlag, Vol. I: xviii+500 pp., Vol. II: xiv+575. DOI: 10.1007/978-3-540-34514-5.

Cho, Moon Jung and Richard H. Stockbridge (2002). "Linear Programming Formulation for Optimal Stopping Problems". In: *SIAM Journal on Control and Optimization* 40.6, pp. 1965–1982. DOI: 10.1137/S0363012900377663.

Christensen, Sören et al. (2019). "On optimal stopping of multidimensional diffusions". In: *Stochastic Processes and their Applications* 129.7, pp. 2561–2581. DOI: https://doi.org/10.1016/j.spa.2018.07.014.

De Boor, Carl (1978). *A practical guide to splines*. Vol. 27. springer-verlag New York.

Ethier, Stewart N. and Thomas G. Kurtz (1986). *Markov processes – characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons Inc.

Graversen, S. E. and Goran Peskir (1998). "On the Russian Option: The Expected Waiting Time". In: *Theory of Probability & Its Applications* 42.3, pp. 416–425. DOI: 10.1137/S0040585X97976295.

Hall, Charles A and W.Weston Meyer (1976). "Optimal error bounds for cubic spline interpolation". In: *Journal of Approximation Theory* 16.2, pp. 105–122. DOI: https://doi.org/10.1016/0021-9045(76)90040-X.

Helmes, Kurt and Richard H Stockbridge (2007). "Linear programming approach to the optimal stopping of singular stochastic processes". In: *Stochastics An International Journal of Probability and Stochastic Processes* 79.3-4, pp. 309–335.

Hernandez-Lerma, O. and Jean B. Lasserre (1996). *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Stochastic Modelling and Applied Probability. Springer New York, NY.

Hernandez-Lerma, O. and Jean B. Lasserre (1999). *Further Topics on Discrete-Time Markov Control Processes*. Stochastic Modelling and Applied Probability. Springer New York, NY.

Hernández-Lerma, Onésimo and Jean B. Lasserre (1998). "Approximation Schemes for Infinite Linear Programs". In: *SIAM Journal on Optimization* 8.4, pp. 973–988. DOI: 10.1137/S1052623497315768.

Kurtz, Thomas G. and Richard H. Stockbridge (2001). "Stationary Solutions and Forward Equations for Controlled and Singular Martingale Problems". In: *Electronic Journal of Probability* 6. Paper No. 14, pp. 1–52. DOI: 10.1214/EJP.v6-90.

Lutz, Matthias (2007). "A linear programming approach for optimal exercise and valuation of American lookback options". PhD thesis. University of Wisconsin–Milwaukee.

Manne, Alan S. (1960). "Linear Programming and Sequential Decisions". In: *Management Science* 6.3, pp. 259–267. (Visited on 05/24/2025).

Peskir, Goran (2005). "The Russian Option: Finite Horizon". In: *Finance and Stochastics* 9, pp. 251–267. DOI: 10.1007/s00780-004-0133-8.

Serfozo, Richard (1982). “Convergence of Lebesgue Integrals with Varying Measures”. In: *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* 44.3, pp. 380–402.

Shepp, Larry and A. N. Shiryaev (1993). “The Russian Option: Reduced Regret”. In: *The Annals of Applied Probability* 3.3, pp. 631–640. DOI: 10.1214/aoap/1177005355.

Shepp, Larry A. and Albert N. Shiryaev (1995). “A New Look at Pricing of the ”Russian Option“”. In: *Theory of Probability & Its Applications* 39.1, pp. 103–119. DOI: 10.1137/1139004.

Shreve, Steven E. (1988). “An Introduction to Singular Stochastic Control”. In: *Stochastic Differential Systems, Stochastic Control Theory and Applications*. Ed. by Wendell Fleming and Pierre-Louis Lions. Vol. 10. IMA Volumes in Mathematics and Its Applications. New York-Berlin: Springer New York, pp. 513–528.

Vanderbei, Robert J. (2014). *Linear Programming: Foundations and Extensions*. 4th ed. 2014. International Series in Operations Research & Management Science, 196. New York, NY: Springer US. DOI: 10.1007/978-1-4614-7630-6.

Vieten, Martin G (2018). “Numerical Solution of Stochastic Control Problems Using the Finite Element Method”. PhD thesis. The University of Wisconsin-Milwaukee.

Young, L. C. (1933). “On Approximation by Polygons in the Calculus of Variations”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 141.844, pp. 325–341.