Numerical methods for optimal decumulation of a defined contribution pension plan

Peter A. Forsyth and
George Labahn

Abstract The decumulation of a defined contribution (DC) pension plan is well known to be one of the hardest problems in finance. We model this decumulation challenge as an optimal stochastic control problem. The control problem is solved, at each rebalancing date, by alternatively solving a linear partial-integro differential equation (PIDE) followed by an optimization step. We solve the PIDE by using a δ-monotone Fourier method, which ensures that monotonicity holds to $O(\delta)$. We allow for the use of leverage (i.e. borrowing to invest in stocks), as well as minimum constraints on bond holdings. We pay particular attention to minimizing wrap-around error, an issue which is endemic for Fourier methods and central to the effective use of these methods for optimal control problems. Rather unexpectedly, we find that restricting the portfolio equity fraction to a maximum of 50% does not reduce portfolio efficiency noticeably. This may be a useful strategy for risk-averse retirees.

Keywords: decumulation, stochastic control, risk
JEL codes: G11, G22
AMS codes: 91G, 65N06, 65N12, 35Q93

1 Introduction

In developed countries, there is a strong and growing shift from Defined Benefit (DB) pension plans to Defined Contribution (DC) plans. The private sector, and increasingly government institutions, are unwilling to take on the balance sheet liabilities of a DB plan (Thinking Ahead Institute, 2024).

Under a DC plan, the employer and employee contribute to a (usually) tax advantaged account, which is typically invested in a mix of stocks and bonds. Upon retirement, the employee is now responsible for both choosing an investment strategy and deciding on a withdrawal schedule. Although it is often advocated that retirees should buy annuities, these are unpopular with DC plan holders (Peijnenburg et al., 2016) for numerous valid reasons (MacDonald et al., 2013). In particular, annuities are generally overpriced, lack true inflation protection and retirees have no access to capital in the event of emergencies.

Peter A. Forsyth
David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1, e-mail: paforsyt@uwaterloo.ca

George Labahn
David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1, e-mail: glabahn@uwaterloo.ca

Since it is usually challenging to fund a reasonable lifestyle investing solely in riskless assets, the retiree is now exposed to both investment risk and longevity risk. This problem of funding retirement from accumulated DC capital, also known as the decumulation problem, has been termed *"the nastiest, hardest problem in finance,"* by Nobel Laureate William Sharpe (Ritholz, 2017).

There have been several suggestions for DC plan *spending rules* (Anarkulova et al., 2025). Probably the most ubiquitous rule is the *four per cent rule* (Bengen, 1994). This rule suggests that retirees can withdraw four per cent of their initial capital (at age 65) annually, adjusted for inflation. Based on historical US data, an investor who invested in a portfolio of 50% bonds and 50% stocks (rebalanced annually), and followed the four per cent rule, would have never run out of savings over any thirty year historical period.

However, the four per cent rule has many critics. For example, rolling thirty year periods have high correlations. In addition, a fixed withdrawal schedule, as well as a constant weight stock allocation is somewhat simplistic. In order to better stress test spending rules, Anarkulova et al. (2025) use block bootstrap resampling of historical data, to generate a distribution of outcomes. We will also use block bootstrap resampling to test our results in this paper.

Decumulation strategies are naturally posed as a problem in optimal stochastic control. The controls in this case are the stock/bond split on each rebalancing date and the (real) annual withdrawal amounts. In our case we impose maximum and minimum constraints on the withdrawal amounts and also impose constraints on the maximum stock allocation.

1.1 Our Contributions

In this paper, we consider the same scenario posed in Bengen (1994). Our objective function is to maximize total withdrawals over a thirty year period (with minimum/maximum annual constraints) and minimize the risk of running out of savings before year thirty.

An optimal stochastic control problem used for decumulation is solved at each rebalancing date, by alternatively solving a linear partial-integro differential equation (PIDE) followed by an optimization step. The PIDE is solved by using a δ-monotone Fourier method, which ensures that monotonicity holds to $O(\delta)$, an important property for control problems. The $\delta-$monotone technique generalizes the method described in Forsyth and Labahn (2019). Our basic dynamic programming technique is similar to the high-level description in Forsyth (2022), but here we delve much deeper in describing the numerical methods. In particular, we focus on such issues as wrap-around errors in Fourier methods or Partial-Integro Differential Equations (PIDEs), and the 2-d δ-monotone Green's function approximation.

Previous studies for decumulation have imposed no-shorting and no-leverage constraints on stock investments. In this paper, we relax the no-leverage constraint on the stock investment and also investigate more restrictive investments in stocks. For example, a risk averse retiree may wish to restrict the maximum fraction invested in equities to be significantly less than one.

Rather surprisingly, we find that restricting the maximum fraction in stocks to be 50% does not lower the efficient frontier by a large amount. The most significant effect of the 50% restriction, compared to allowing a maximum stock fraction of up to 130% (leverage) is that the time for the median withdrawals to reach the maximum is delayed by one year. This may be an acceptable trade-off for risk averse investors.

From the numerical point of view, when solving optimal control problems using Fourier methods, wrap-around error can be significant, due to the possible instantaneous movement to domain boundaries at reallocation times (Lippa, 2013; Ignatieva et al., 2018; Alonso-Garcca et al., 2018). We show that our technique reduces wrap-around error to almost machine level precision.

1.2 Organization of this Chapter

The rest of this paper is organized as follows. In the next section we state our problem setting followed by a mathematical formulation of our model in Section 3. The latter section includes information on the stochastic processes followed by bonds and stocks along with a description of our control set. Section 4 describes our conflicting risk and reward measures, Expected Shortfall and Expected Withdrawals, while Section 5 looks at how one maximizes these conflicting measures, something which is initially set up as a pre-commitment problem which is not implementable. Section 6 then discusses how one can formulate our problem into an (implementable) dynamic program, one which is solved by a PIDE between withdrawal time intervals. Section 7 then shows how one can solve these PIDEs using Green's functions and Fourier analysis. The wrap around problem that comes with using Fourier analysis and periodicity is addressed in the next section. That section also includes the resulting numerical algorithm for solving our optimal control decumulation problem. Section 9 then presents an example along with its numerical solution followed in the next section by tests of robustness of our numerical example. The paper ends with a conclusion along with an Appendix providing extra details regarding the wraparound error discussed earlier.

2 Problem Setting

As noted in Anarkulova et al. (2025), retirees have a revealed preference for spending rules, such as that advocated by Bengen (1994). Spending rules (such as the four per cent rule) are popular with retirees as they are simple to implement, have an intuitive interpretation, and the rolling 30 year backtests in Bengen (1994) are easy to understand.

In this paper we restrict attention to the scenario outlined in Bengen (1994), that is, we consider a 65-year old retiree who desires fixed minimum annual (real) cash flows over a 30 year time horizon. We also impose an upper limit (a cap) on maximum withdrawals in any year. From the CPM2014 table from the Canadian Institute of Actuaries[1], the probability that a 65-year old Canadian male attains the age of 95 is about 0.13. In spite of this relatively low probability of achieving the age of 95, assuming a 30 year time horizon is considered a prudent approach for minimizing the risk of exhausting savings. In addition, observe that we will not mortality weight future cash flows, as is done when averaging over a population for pricing annuities. Mortality weighting may not be meaningful for an individual investor, since it reflects population averages, rather than the circumstances of an individual retiree.

Since we allow investing in risky assets, with a minimum cash withdrawal each year, it is possible to exhaust savings.[2] In this case, we continue to withdraw funds from the portfolio, which is equivalent to borrowing cash. In such circumstances, we assume that withdrawals continue, something which can be viewed as borrowing against other resources. This debt accumulates at the borrowing rate. Essentially, we are assuming that the investor has other assets, for example real estate, which can be used as a hedge of last resort. In this case, this debt could then be funded using a reverse mortgage, with real estate as collateral (Pfeiffer et al., 2013).

---

[1] `www.cia-ica.ca/docs/default-source/2014/214013e.pdf`.

[2] Ignoring the trivial and unlikely case where investing in riskless assets can fund 30 year minimum withdrawals.

It is important to note that an implicit assumption here is that real estate is not fungible with other financial assets, except as a last resort. This mental separation of assets is a common aspect of behavioral finance (Shefrin and Thaler, 1988). Real estate in this case has a dual role: if market returns are exceptionally strong, or the retiree passes away earlier then expected, this property can be a bequest. If market returns are poor, or in the case of extreme longevity, then real estate can be used to fund expenses if the retirement account is exhausted.

3 Mathematical Formulation

We assume that the investor has access to two funds: a broad market stock index fund and a constant maturity bond index fund. Let $S(t)$ and $B(t)$ denote the real (inflation adjusted) *amounts* invested in the stock index and the bond index, respectively, at time t with T being the investment horizon. These amounts will evolve over time, depending on the investor's asset allocation, and changes in the real unit prices of the assets. In the absence of an investor determined control such as cash withdrawals or rebalancing, any changes in $S(t)$ and $B(t)$ result from asset price evolution. We model the stock index as following a jump diffusion, which permits modelling both continuous stochastic processes as well as market jumps.

We model the real returns of a constant maturity bond index as a stochastic process (see for example, Lin et al., 2015; MacMinn et al., 2014). This avoids the need to model bond prices and inflation separately. As done in MacMinn et al. (2014), we assume that the constant maturity bond index follows a jump diffusion process, with justification being found in Forsyth et al. (2022, Appendix A).

3.1 The Diffusion Processes.

For any time t let

$$S(t^-) \equiv \lim_{\epsilon \to 0^+} S(t-\epsilon) \quad \text{and} \quad S(t^+) \equiv \lim_{\epsilon \to 0^+} S(t+\epsilon)$$

denote the values of S the instant before t^- and the instant after t^+ time t. Such notation will also be used for any other time dependent function. We let ξ^s be a random number representing a jump multiplier, that is, when a jump occurs, $S(t) = \xi^s S(t^-)$. The use of a jump process accounts for non-normal asset returns. As in (Kou, 2002; Kou and Wang, 2004), we assume that $\log(\xi^s)$ follows a double exponential distribution with u^s the probability of an upward jump and $1 - u^s$ the probability of a downward jump. The density function for $y = \log \xi^s$ is

$$f^s(y) = u^s \eta_1^s e^{-\eta_1^s y} \mathbf{1}_{y \geq 0} + (1-u^s)\eta_2^s e^{\eta_2^s y} \mathbf{1}_{y<0} \,. \tag{1}$$

We also define

$$\gamma_\xi^s = E[\xi^s - 1] = \frac{u^s \eta_1^s}{\eta_1^s - 1} + \frac{(1-u^s)\eta_2^s}{\eta_2^s + 1} - 1 \,, \tag{2}$$

where $E[\cdot]$ is the expectation. In the absence of control, $S(t)$ evolves according to

$$\frac{dS(t)}{S(t^-)} = \left(\mu^s - \lambda_\xi^s \gamma_\xi^s\right) \, dt + \sigma^s \, dZ^s + d\left(\sum_{i=1}^{\pi_t^s} (\xi_i^s - 1)\right), \tag{3}$$

where μ^s is the (uncompensated) drift rate, σ^s is the volatility, dZ^s is the increment of a Wiener process, π_t^s is a Poisson process with positive intensity parameter λ_ξ^s, and the ξ_i^s are i.i.d. positive random variables having distribution (1). Moreover, ξ_i^s, π_t^s, and Z^s are assumed to all be mutually independent.

Similarly, let the amount in the bond index the instant before t be $B(t^-)$. In the absence of investor intervention, $B(t)$ evolves as

$$\frac{dB(t)}{B(t^-)} = \left(\mu^b - \lambda_\xi^b \gamma_\xi^b + \mu_c^b \mathbf{1}_{\{B(t^-)<0\}}\right) dt + \sigma^b \, dZ^b + d\left(\sum_{i=1}^{\pi_t^b} (\xi_i^b - 1)\right), \tag{4}$$

where the terms in equation (4) are defined analogously to equation (3). In particular, π_t^b is a Poisson process with positive intensity parameter λ_ξ^b, and the density function for $y = \log \xi^b$ is

$$f^b(y) = u^b \eta_1^b e^{-\eta_1^b y} \mathbf{1}_{y \geq 0} + (1 - u^b) \eta_2^b e^{\eta_2^b y} \mathbf{1}_{y<0} \ ,$$

and $\gamma_\xi^b = E[\xi^b - 1]$. Again ξ_i^b, π_t^b, and Z^b are assumed to all be mutually independent. The term $\mu_c^b \mathbf{1}_{\{B(t^-)<0\}}$ in equation (4) represents the extra cost of borrowing, that is, the spread.

The diffusion processes are correlated, that is, $dZ^s \cdot dZ^b = \rho_{sb} \, dt$. However, contrary to common belief, an analysis of historical data suggests that the stock and bond jump processes are essentially uncorrelated (see Forsyth (2020b) for empirical justification). We make this assumption in this work.

We define the investor's total wealth at time t as

$$\text{Total wealth } \equiv W(t) = S(t) + B(t).$$

In case of insolvency, the portfolio is liquidated, trading stops and any outstanding debt accrues interest at the borrowing rate.

3.2 The Set of Controls.

Consider a set of discrete times

$$\mathcal{T} = \{t_0 = 0 < t_1 < t_2 < \ldots < t_M = T\}$$

where we assume that $t_i - t_{i-1} = \Delta t = T/M$ is constant. Here $t_0 = 0$ is the inception time of the investment and $\mathcal{T}$ is the set of withdrawal/rebalancing times, as defined in equation (3.2). At each rebalancing time t_i, $i = 0, 1, \ldots, M-1$, the investor first withdraws an amount of cash q_i from the portfolio, and then afterwards rebalances the portfolio. At $t_M = T$ the portfolio is liquidated and no cash flow occurs. This is enforced by specifying $q_M = 0$.

Let

$$W(t_i^-) = S(t_i^-) + B(t_i^-)$$

denote the instant before withdrawals and rebalancing at t_i, we then have that $W(t_i^+)$ is given by

$$W(t_i^+) = W(t_i^-) - q_i \ ; \ i \in \mathcal{T} \ ,$$

with $W(t_M^+) = W(t_M^-)$ since $q_M \equiv 0$.

Typically, DC plan savings are held in a tax-advantaged account, with no taxes triggered by rebalancing. With infrequent (e.g. yearly) rebalancing, we also expect other transaction costs, to be small, and hence can be ignored. It is possible to include transaction costs, but at the expense of increased computational cost (van Staden et al., 2018).

Let $X(t) = (S(t), B(t))$, $t \in [0,T]$ denote the multi-dimensional controlled underlying process and $x = (s,b)$ the realized state of the system. The rebalancing control $p_i(\cdot)$ is the fraction invested in the stock index at the rebalancing date t_i, that is,

$$p_i\left(X(t_i^-)\right) = p\left(X(t_i^-), t_i\right) = \frac{S(t_i^+)}{S(t_i^+) + B(t_i^+)} .$$

Let $q_i(\cdot)$ be the amount withdrawn at time t_i, that is, $q_i\left(X(t_i^-)\right) = q\left(X(t_i^-), t_i\right)$. Formally, the controls depend on the state of the investment portfolio, before the rebalancing occurs, that is, $p_i(\cdot) = p\left(X(t_i^-), t_i\right)$ and $q_i(\cdot) = q\left(X(t_i^-), t_i\right)$ $t_i \in \mathcal{T}$, where $\mathcal{T}$ is the set of rebalancing times. However, it will be convenient to note that in our case, we find the optimal control $p_i(\cdot)$ amongst all strategies with constant wealth (after withdrawal of cash). Hence, with some abuse of notation, we will now consider $p_i(\cdot)$ to be function of wealth after withdrawal of cash, where we use the shorthand notation W_i^- and W_i^+ for the variables representing wealth the instant before and instant after t_i. Using a similar shorthand notation for S_i and B_i we then have

$$\begin{aligned} p_i(\cdot) &= p(W(t_i^+), t_i) = p_i(W_i^+) \\ S_i^+ &= p_i(W_i^+)\ W_i^+ \\ B_i^+ &= (1 - p_i(W_i^+))\ W_i^+ . \end{aligned}$$

Note that the control for $p_i(\cdot)$ depends only W_i^+. Since $p_i(\cdot) = p_i(W_i^- - q_i)$, it follows that

$$q_i(\cdot) = q_i(W_i^-) ,$$

which we discuss further in Section 6.

A control at time t_i is then given by the pair $(q_i(\cdot), p_i(\cdot))$ where the notation $(\cdot)$ denotes that the control is a function of the state. Let $\mathcal{Z}$ represent the set of admissible values of the controls $(q_i(\cdot), p_i(\cdot))$. We impose no-shorting, bounded leverage constraints (assuming solvency) along with maximum and minimum values for the withdrawals. In addition we apply the constraint that in the event of insolvency due to withdrawals ($W(t_i^+) < 0$), or in the case of leverage, trading ceases and debt (negative wealth) accumulates at the appropriate borrowing rate of return (that is, a spread over the bond rate). We also specify that the stock assets are liquidated at $t = t_M$.

We can then define our controls by

$$\mathcal{Z}_q(W_i^-, t_i) = \begin{cases} [q_{\min}, q_{\max}] & t_i \in \mathcal{T}\ ;\ t_i \neq t_M\ ;\ W_i^- \geq q_{\max} \\ [q_{\min}, \max(q_{\min}, W_i^-)] & t_i \in \mathcal{T}\ ;\ t \neq t_M\ ;\ W_i^- < q_{\max} \\ \{0\} & t_i = t_M \end{cases} , \tag{5}$$

$$\mathcal{Z}_p(W_i^+, t_i) = \begin{cases} [0, p_{\max}] & t_i \in \mathcal{T}\ ;\ t_i \neq t_M\ ;\ W_i^+ > 0 \\ \{0\} & t_i \in \mathcal{T}\ ;\ t_i \neq t_M\ ;\ W_i^+ \leq 0 \\ \{0\} & t_i = t_M \end{cases} . \tag{6}$$

The rather complicated expression in equation (5) imposes the assumption that as wealth becomes small, the retiree first tries to avoid insolvency as much as possible and in the event of insolvency, withdraws only $q_{\min}$.

The set of admissible values for $(q_i, p_i), t_i \in \mathcal{T}$, can then be written as

$$(q_i, p_i) \in \mathcal{Z}(W_i^-, W_i^+, t_i) = \mathcal{Z}_q(W_i^-, t_i) \times \mathcal{Z}_p(W_i^+, t_i) . \tag{7}$$

For implementation purposes, we have written equation (7) in terms of the wealth after withdrawal of cash. However, we remind the reader that since $W_i^+ = W_i^- - q_i$, the controls are formally a function of the state $X(t_i^-)$ before the control is applied.

The admissible control set $\mathcal{A}$ can then be written as

$$\mathcal{A} = \left\{ (q_i, p_i)_{0 \le i \le M} : (p_i, q_i) \in \mathcal{Z}(W_i^-, W_i^+, t_i) \right\}$$

with an admissible control $\mathcal{P} \in \mathcal{A}$ written as

$$\mathcal{P} = \{ (q_i(\cdot), p_i(\cdot)) \ : \ i = 0, \ldots, M \} \ .$$

We also define $\mathcal{P}_n \equiv \mathcal{P}_{t_n} \subset \mathcal{P}$ as the tail of the set of controls in $[t_n, t_{n+1}, \ldots, t_M]$, that is,

$$\mathcal{P}_n = \{ (q_n(\cdot), p_n(\cdot)), \ldots, (q_M(\cdot), p_M(\cdot)) \}$$

and $\mathcal{A}_n$, the tail of the admissible control set, as

$$\mathcal{A}_n = \left\{ (q_i, p_i)_{n \le i \le M} : (q_i, p_i) \in \mathcal{Z}(W_i^-, W_i^+, t_i) \right\} ,$$

so that $\mathcal{P}_n \in \mathcal{A}_n$.

4 Risk and Reward

In this section we consider our measures of risk and reward for our retiree. In this case the retiree is primarily concerned with the risk of depleting savings while at the same time hoping to maximize the cash withdrawals from his plan.

4.1 Risk: Expected Shortfall (ES)

Two typical measures of financial tail risk of an investment portfolio are Value at Risk (VAR) and Conditional Value of risk (CVAR), with the latter also known as Expected Shortfall. Expected Shortfall is an alternative to value at risk that is more sensitive to the shape of the tail of the loss distribution.

Suppose

$$\int_{-\infty}^{W_\alpha^*} g(W_T) \ dW_T = \alpha,$$

where $g(W_T)$ be the probability density function of wealth W_T at $t = T$. Then W_α^* can be viewed as the Value at Risk (VAR) at level α since the probability of wealth being more than W_α^* is $1 - \alpha$. For example, if $\alpha = .01$, then 99% of the outcomes have $W_T > W_\alpha^*$. If W_α^* is sufficiently large, this suggests very low risk of running out of savings. The Expected Shortfall (ES) at level α is then

$$\text{ES}_\alpha = \frac{1}{\alpha} \int_{-\infty}^{W_\alpha^*} W_T \ g(W_T) \ dW_T \ , \tag{8}$$

which is the mean of the worst α fraction of outcomes. Typically $\alpha \in \{.01, .05\}$. The definition of ES in equation (8) uses the probability density of the final wealth distribution, not the density of *loss*. Hence in our case, a larger value of ES, that is, a larger value of average worst case terminal wealth, is desired. [3]

Define $X_0^- = X(t_0^-)$. Given an expectation under control $\mathcal{P}$, $E_{\mathcal{P}}[\cdot]$, then as noted by Rockafellar and Uryasev (2000), the expected shortfall at level α has the alternate formulation as

$$\mathrm{ES}_\alpha(X_0^-, t_0^-) = \sup_{W^*} E_{\mathcal{P}_0}^{X_0^-, t_0^-}\left[W^* + \frac{1}{\alpha}\min(W_T - W^*, 0) \right] . \tag{9}$$

The admissible set for W^* in equation (9) is over the set of possible values for W_T.

The notation $\mathrm{ES}_\alpha(X_0^-, t_0^-)$ emphasizes that ES_α is as seen at (X_0^-, t_0^-), that is, the pre-commitment ES_α. A strategy based purely on optimizing the pre-commitment value of ES_α at time zero is *time-inconsistent*. As such it has been termed by many as *non-implementable*, since the investor has an incentive to deviate from the time zero pre-commitment strategy at $t > 0$. However we consider the pre-commitment strategy merely as a device to determine an appropriate level of W^* in equation (9). If we fix W^* for all positive t, then this strategy is the induced time-consistent strategy (Strub et al., 2019; Forsyth, 2020a; Cui et al., 2022) and hence is implementable. For further discussion of the relationship between time consistent and pre-commitment strategies, see (Vigna, 2014; Menoncin and Vigna, 2017; Vigna, 2017; Strub et al., 2019; Forsyth, 2020a; Bjork et al., 2021; Cui et al., 2022). In particular, see Forsyth (2020a) for discussion of the induced time consistent policy resulting from the use of ES risk.

An alternative measure of risk could be based on variability of withdrawals (Forsyth et al., 2020). However, we note that we have constraints on the minimum and maximum withdrawals, so that variability is mitigated. We also assume that given these constraints, the retiree is primarily concerned with the risk of depleting savings, something which is well measured by ES.

4.2 Reward: Expected Total Withdrawals (EW)

We will use expected total withdrawals as a measure of reward in the following. More precisely, we define EW (expected withdrawals) as

$$\mathrm{EW}(X_0^-, t_0^-) = E_{\mathcal{P}_0}^{X_0^+, t_0^+}\left[\sum_{i=0}^{M} q_i \right] , \tag{10}$$

where we assume that the investor survives for the entire decumulation period. This is consistent with the scenario in Bengen (1994) .

We remark that there is no discounting term in equation (10) as all quantities are real, that is, inflation-adjusted. It is straightforward to introduce discounting, but we view setting the real discount rate to zero to be a reasonable and conservative choice. See Forsyth (2022) for further comments.

5 Maximizing Conflicting Measures: Problem EW-ES

Expected withdrawals (EW) and expected shortfall (ES) are two conflicting measures. We handle this by using a scalarization technique to find the Pareto points for this bi-objective optimization problem, that is, the set of points where one objective cannot be

[3] In practice, the negative of W_α^* is often the reported VAR. and the negative of ES is commonly referred to as Conditional Value at Risk (CVAR).

improved without worsening the other. For a given scalarization parameter $\kappa > 0$, the goal is to then find the control $\mathcal{P}_0$ that maximizes

$$\mathrm{EW}(X_0^-, t_0^-) + \kappa\, \mathrm{ES}_\alpha(X_0^-, t_0^-) \ .$$

More precisely, we define the pre-commitment EW-ES problem $(PCES_{t_0}(\kappa))$ problem in terms of the value function $J(s, b, t_0^-)$ derived from (9) and (10):

$$J\left(s, b, t_0^-\right) = \sup_{\mathcal{P}_0 \in \mathcal{A}} \sup_{W^*} \left\{ E_{\mathcal{P}_0}^{X_0^-, t_0^-} \left[\sum_{i=0}^{M} q_i + \kappa \left(W^* + \frac{1}{\alpha} \min(W_T - W^*, 0) \right) + \ \epsilon W_T \middle| X(t_0^-) = (s, b) \ \right] \right\} \tag{11}$$

$$\text{subject to} \begin{cases} (S(t), B(t)) \text{ follows processes (3) and (4);} \quad t \notin \mathcal{T} \\ W_\ell^+ = W_\ell^- - q_\ell \ ; \ X_\ell^+ = (S_\ell^+, B_\ell^+) \\ W_\ell^- = \left(S_\ell^- + B_\ell^- \right) \\ S_\ell^+ = p_\ell(\cdot) W_\ell^+ \ ; \ B_\ell^+ = (1 - p_\ell(\cdot)) W_\ell^+ \\ (q_\ell(\cdot), p_\ell(\cdot)) \in \mathcal{Z}(W_\ell^-, W_\ell^+, t_\ell) \\ \ell = 0, \ldots, M \ ; \ t_\ell \in \mathcal{T} \end{cases} . \tag{12}$$

Note that we have added an extra term $E_{\mathcal{P}_0}^{X_0^-, t_0^-}[\epsilon W_T]$ to equation (11). If we have a maximum withdrawal constraint, and if $W_t \gg W^*$ as $t \longrightarrow T$, then the controls become ill-posed. In this fortunate state for the investor, we can break investment policy ties either by setting $\epsilon < 0$, which will force investments in bonds, or by setting $\epsilon > 0$, which will force investments into stocks. Choosing $|\epsilon| \ll 1$ ensures that this term only has an effect if $W_t \gg W^*$ and $t \longrightarrow T$. See Forsyth (2022) for more discussion of this.

We can interchange the $\sup \sup(\cdot)$ [4] in equation (11) to represent the value function as

$$J\left(s, b, t_0^-\right) = \sup_{W^*} \sup_{\mathcal{P}_0 \in \mathcal{A}} \left\{ E_{\mathcal{P}_0}^{X_0^+, t_0^+} \left[\sum_{i=0}^{M} \mathfrak{q}_i + \kappa \left(W^* + \frac{1}{\alpha} \min(W_T - W^*, 0) \right) + \epsilon W_T \middle| X(t_0^-) = (s, b) \right] \right\} . \tag{13}$$

Since the inner supremum in equation (13) is a continuous function of W^* and the optimal value of W^* in equation (13) is bounded, [5] we can define

$$\mathcal{W}^*(s,b) = \arg\max_{W^*} \left\{ \sup_{\mathcal{P}_0 \in \mathcal{A}} \left\{ E_{\mathcal{P}_0}^{X_0^+, t_0^+} \left[\ \sum_{i=0}^{M} \mathfrak{q}_i + \kappa \left(W^* + \frac{1}{\alpha} \min(W_T - W^*, 0) \right) + \epsilon W_T \middle| X(t_0^-) = (s, b) \right] \right\} \right. . \tag{14}$$

Given $\mathcal{W}^*(s,b)$ from equation (14), then the optimal control $\mathcal{P}^*$ which solves Problem (13) is the same control which solves

[4] Let $F = \sup_{(a,b) \in A \times B} f(a, b)$, then $\forall \epsilon > 0$, $\exists (a^*, b^*) \in A \times B$, s.t. $f(a^*, b^*) > F - \epsilon$. Then $F \geq \sup_{a \in A} \sup_{b \in B} f(a, b) \geq \sup_{b \in b} f(a^*, b) \geq f(a^*, b^*) > F - \epsilon$. Hence, $\epsilon \to 0$ implies $\sup_{a \in A} \sup_{b \in B} f(a, b) = F$. Similarly $\sup_{b \in B} \sup_{a \in A} f(a, b) = F$.

[5] This is the same as noting that a finite value at risk exists.

$$\hat{J}\left(s,b,t_0^-\right) = \sup_{\mathcal{P}_0 \in \mathcal{A}} \left\{ E_{\mathcal{P}_0}^{X_0^+,t_0^+} \left[\sum_{i=0}^{M} \mathfrak{q}_i + \frac{\kappa}{\alpha} \min(W_T - \mathcal{W}^*(s,b), 0) + \epsilon W_T \middle| X(t_0^-) = (s,b) \right] \right\}. \tag{15}$$

Hence Problem (15) is the induced time consistent policy for Problem (13). We refer the reader to Forsyth (2020a) for an extensive discussion concerning pre-commitment and time consistent expected shortfall strategies.

6 Formulation as a Dynamic Program

We can use the method in Forsyth (2020a) to determine our value function. Write (13) as

$$J(s,b,t_0^-) = \sup_{W^*} \tilde{v}(s,b,W^*,0^-) ,$$

where $\tilde{v}(s,b,W^*,t)$ is defined as

$$\tilde{v}(s,b,W^*,t_n^-) = \sup_{\mathcal{P}_n \in \mathcal{A}_n} \left\{ E_{\mathcal{P}_n}^{\hat{X}_n^+,t_n^+} \left[\sum_{i=n}^{M} q_i + \kappa \left(W^* + \frac{1}{\alpha} \min((W_T - W^*),0) \right) + \epsilon W_T \;\middle|\; \hat{X}(t_n^-) = (s,b,W^*) \right] \right\} \tag{16}$$

$$\text{subject to} \begin{cases} (S(t),B(t)) \text{ follows processes (3) and (4);} \quad t \notin \mathcal{T} \\ W_\ell^+ = W_\ell^- - q_\ell \; ; \; X_\ell^+ = (S_\ell^+, B_\ell^+, W^*) \\ W_\ell^- = \left(S_\ell^- + B_\ell^- \right) \\ S_\ell^+ = p_\ell(\cdot) W_\ell^+ \; ; \; B_\ell^+ = (1 - p_\ell(\cdot)) W_\ell^+ \\ (q_\ell(\cdot), p_\ell(\cdot)) \in \mathcal{Z}(W_\ell^-, W_\ell^+, t_\ell) \\ \ell = n, \ldots, M \; ; \; t_\ell \in \mathcal{T} \end{cases} . \tag{17}$$

The original problem (13) has therefore been decomposed into two steps:

1. For given initial cash W_0, and a fixed value of W^*, solve problem (16) using dynamic programming in order to determine $\tilde{v}(0,W_0,W^*,0^-)$.
2. Solve problem (13) by maximizing over W^*

$$J(0,W_0,0^-) = \sup_{W^*} v(0,W_0,W^*,0^-) . \tag{18}$$

6.1 Dynamic Programming Solution of Problem (16)

We give a brief overview of the method described in detail in (Forsyth, 2022). We apply the dynamic programming principle to $t_n \in \mathcal{T}$

$$\tilde{v}(s,b,W^*,t_n^-) = \sup_{q \in \mathcal{Z}_q(w^-,t_n)} \left\{ \sup_{p \in \mathcal{Z}_p(w^- - q,t_n)} \left[q + \tilde{v}((w^- - q)p, (w^- - q)(1-p), W^*, t_n^+) \right] \right\}$$
$$\text{where } w^- = (s+b) . \tag{19}$$

If we set

$$h(w, t_n, W^*) = \left[\sup_{p \in \mathcal{Z}_p(w,t_n)} \tilde{v}(wp, w(1-p), W^*, t_n^+) \right] \tag{20}$$

then equation (19) becomes

$$\tilde{v}(s, b, W^*, t_n^-) = \sup_{q \in \mathcal{Z}_q(w^-,t_n)} \left\{ q + \left[h((w^- - q), W^*, t_n^+) \right] \right\}$$
$$\text{with } w^- = (s+b). \tag{21}$$

This approach effectively replaces a two dimensional optimization for (q_n, p_n), by two sequential one dimensional optimizations. From equations (20) and (21), the optimal pair (q_n, p_n) satisfies

$$q_n = q_n(w^-, W^*) \qquad \text{where} \qquad w^- = (s+b)$$
$$p_n = p_n(w, W^*) \qquad \text{where} \qquad w = w^- - q_n \ .$$

In other words, the optimal withdrawal control q_n is only a function of total wealth before withdrawals while the optimal control p_n is a function only of total wealth after withdrawals. If a withdrawal results in $W^+ > 0$, then the optimal control for p is determined along lines of constant wealth in the (s, b) plane while if a withdrawal results in $W^+ < 0$, then stocks are liquidated ($p = 0$). This is illustrated in Figure 1.

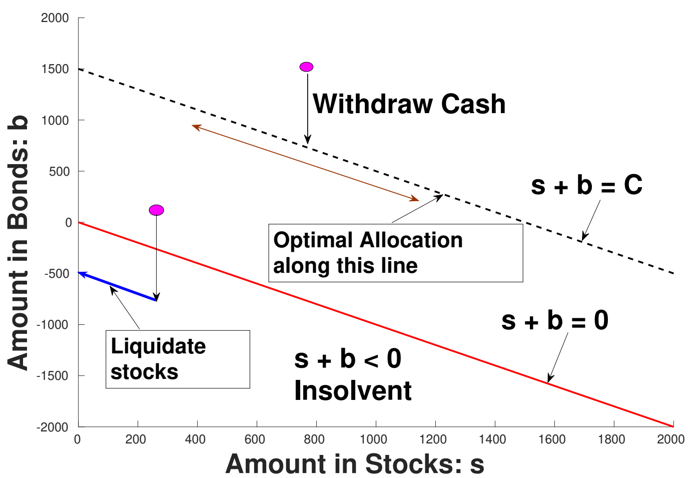


Fig. 1: *Schematic of withdrawal controls.*

At $t = T$, we have

$$\tilde{v}(s, b, W^*, T^+) = \kappa \left(W^* + \frac{\min((s+b-W^*), 0)}{\alpha} \right) + \epsilon(s+b)$$

while at points in between rebalancing times, that is when $t \notin \mathcal{T}$, standard arguments from SDEs (3-4), and Forsyth (2022) give

$$\tilde{v}_t + \frac{(\sigma^s)^2 s^2}{2} \tilde{v}_{ss} + (\mu^s - \lambda_\xi^s \gamma_\xi^s) s \tilde{v}_s + \lambda_\xi^s \int_{-\infty}^{+\infty} \tilde{v}(e^y s, b, t) f^s(y)\, dy + \frac{(\sigma^b)^2 b^2}{2} \tilde{v}_{bb}$$
$$+ (\mu^b + \mu_c^b \mathbf{1}_{\{b<0\}} - \lambda_\xi^b \gamma_\xi^b) b \tilde{v}_b + \lambda_\xi^b \int_{-\infty}^{+\infty} \tilde{v}(s, e^y b, t) f^b(y)\, dy - (\lambda_\xi^s + \lambda_\xi^b)\tilde{v} + \rho_{sb}\sigma^s\sigma^b sb\tilde{v}_{sb} = 0 \ . \tag{22}$$

It is convenient to consider the two cases $b \geq 0$ and $b \leq 0$ separately. For $b \geq 0$, we solve

$$\begin{aligned}
&\tilde{v}_t + \frac{(\sigma^s)^2 s^2}{2}\tilde{v}_{ss} + (\mu^s - \lambda_\xi^s \gamma_\xi^s)s\tilde{v}_s + \lambda_\xi^s \int_{-\infty}^{+\infty} \tilde{v}(e^y s, b, t) f^s(y)\ dy + \frac{(\sigma^b)^2 b^2}{2}\tilde{v}_{bb} \\
&\quad + (\mu^b - \lambda_\xi^b \gamma_\xi^b) b\tilde{v}_b + \lambda_\xi^b \int_{-\infty}^{+\infty} \tilde{v}(s, e^y b, t) f^b(y)\ dy - (\lambda_\xi^s + \lambda_\xi^b)\tilde{v} + \rho_{sb}\sigma^s\sigma^b sb\tilde{v}_{sb} = 0\ , \\
&\qquad b \geq 0, s \geq 0\ . \qquad (23)
\end{aligned}$$

When $b < 0$, it is convenient to re-write equation (22) in terms of debt $\hat{b} = -b$. Letting $\hat{v}(s, \hat{b}, t) = \tilde{v}(s, b, t), b < 0, \hat{b} = -b$ in equation (22) we obtain

$$\begin{aligned}
&\hat{v}_t + \frac{(\sigma^s)^2 s^2}{2}\hat{v}_{ss} + (\mu^s - \lambda_\xi^s \gamma_\xi^s)s\hat{v}_s + \lambda_\xi^s \int_{-\infty}^{+\infty} \hat{v}(e^y s, \hat{b}, t) f^s(y)\ dy + \frac{(\sigma^b)^2 \hat{b}^2}{2}\hat{v}_{\hat{b}\hat{b}} \\
&\quad + (\mu^b + \mu_c^b - \lambda_\xi^b \gamma_\xi^b)\hat{b}\hat{v}_{\hat{b}} + \lambda_\xi^b \int_{-\infty}^{+\infty} \hat{v}(s, e^y \hat{b}, t) f^b(y)\ dy - (\lambda_\xi^s + \lambda_\xi^b)\hat{v} + \rho_{sb}\sigma^s\sigma^b s\hat{b}\hat{v}_{s\hat{b}} = 0\ , \\
&\qquad \hat{b} > 0, s \geq 0\ . \qquad (24)
\end{aligned}$$

Note that equation (24) is now amenable to a transformation of the form $\hat{x} = \log \hat{b}$ since $\hat{b} > 0$, something required when using a Fourier method (Forsyth and Labahn, 2019; Forsyth, 2022) to solve equation (24).

After rebalancing, if $b \geq 0$, then b cannot become negative, since $b = 0$ is a barrier in equation (22). However, b can become negative after withdrawals, in which case b remains in the state $b < 0$. There equation (24) applies, unless there is an injection of cash to move to a state with $b > 0$. The terminal condition for equation (24) is

$$\hat{v}(s, \hat{b}, W^*, T^+) = \kappa\left(W^* + \frac{\min((s - \hat{b} - W^*), 0)}{\alpha}\right) + \epsilon(s - \hat{b})\ ;\ \hat{b} > 0\ .$$

7 The δ-Monotone Fourier Method

The δ-monotone Fourier method originated with the work of Forsyth and Labahn (2019), where it was applied to one dimensional problems in stochastic control. In the following, we give a sketch showing how to extend those methods to work for two dimensional problems.

We illustrate this method by focusing on equation (23). Since $b > 0, s > 0$ we can let $x_1 = \log s, x_2 = \log b, \tau = T - t$ with $v(x_1, x_2, \tau) = \tilde{v}(e^{x_1}, e^{x_2}, T - \tau)$. Then (23) converts to

$$\begin{aligned}
v_\tau &= \frac{(\sigma^s)^2}{2} v_{x_1 x_1} + (\mu^s - \lambda_\xi^s \gamma_\xi^s) v_{x_1} + \lambda_\xi^s \int_{-\infty}^{\infty} v(x_1 + y, x_2, \tau) f^s(y)\ dy + \frac{(\sigma^b)^2}{2} v_{x_2 x_2} \\
&\quad + (\mu^b - \lambda_\xi^b \gamma_\xi^b) v_{x_2} + \lambda_\xi^b \int_{-\infty}^{\infty} v(x_1, x_2 + y, t) f^b(y)\ dy - (\lambda_\xi^s + \lambda_\xi^b) v + \rho_{sb}\sigma^s\sigma^b v_{x_1 x_2}. \qquad (25)
\end{aligned}$$

The exact solution of equation (25) can be written as

$$v(x_1, x_2, \tau + \Delta\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1 - z_1, x_2 - z_2, \Delta\tau)\ v(z_1, z_2, \tau)\ dz_1\ dz_2\ , \qquad (26)$$

where $g(x_1, x_2, \Delta\tau)$ is the Green's function of equation (23) (Garroni and Menaldi, 1992).

Remark 1 (Form of Green's function) Note that the Green's function for equation (23) is of the form $g(x_1, x_2, z_1, z_2) = g(x_1 - z_1, x_2 - z_2)$, that is a function of the differences. Intuitively, we can view the Green's function as a scaled probability density f, and this means that $f(x_1, x_2)|(z_1, z_2)$ depends only on the difference, $f(x_1, x_2)|(z_1, z_2) = f(x_1-z_1, x_2-z_2)$. See Forsyth and Labahn (2019) for more discussion of this observation.

The Fourier transform pair for the Green's function is given as

$$
\begin{aligned}
G(\omega_1, \omega_2, \Delta\tau) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1, x_2, \Delta\tau) e^{-2\pi i w_1 x_1} e^{-2\pi i w_2 x_2}\, dx_1\, dx_2 \\
g(x_1, x_2, \Delta\tau) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(\omega_1, \omega_2, \Delta\tau) e^{2\pi i w_1 x_1} e^{2\pi i w_2 x_2}\, d\omega_1\, d\omega_2\ .
\end{aligned} \tag{27}
$$

where $i = \sqrt{-1}$. Standard techniques then give the Fourier Transform of the Green's function $G(\omega_1, \omega_2, \Delta\tau)$ for equation (25) as

$$
G(\omega_1, \omega_2, \Delta\tau) = e^{\Psi(\omega_1, \omega_2)\Delta\tau}\ , \tag{28}
$$

where

$$
\begin{aligned}
\Psi(\omega_1, \omega_2) = &-\frac{(\sigma^s)^2}{2}(2\pi\omega_1)^2 + \left(\mu^s - \lambda_\xi^s \gamma_\xi^s - \frac{(\sigma^s)^2}{2}\right)(2\pi i\omega_1) + \lambda_\xi^s \overline{F^s}(\omega_1) \\
&-\frac{(\sigma^b)^2}{2}(2\pi\omega_2)^2 + \left(\mu^b - \lambda_\xi^b \gamma_\xi^b - \frac{(\sigma^b)^2}{2}\right)(2\pi i\omega_2) + \lambda_\xi^b \overline{F^b}(\omega_1) \\
&-(\lambda_\xi^s + \lambda_\xi^b) - \rho_{sb}\sigma^s\sigma^b(2\pi\omega_1)(2\pi\omega_2).
\end{aligned} \tag{29}
$$

Here $\overline{F^b}(\omega_1), \overline{F^s}(\omega_1)$ are the complex conjugates of the Fourier transforms of the density functions f^s, f^b:

$$
\begin{aligned}
\overline{F^s}(\omega_1) &= \frac{u^s}{1 - \frac{2\pi i\omega_1}{\eta_1^s}} + \frac{1 - u^s}{1 + \frac{2\pi i\omega_1}{\eta_2^s}} \\
\overline{F^b}(\omega_2) &= \frac{u^b}{1 - \frac{2\pi i\omega_2}{\eta_1^b}} + \frac{1 - u^b}{1 + \frac{2\pi i\omega_2}{\eta_2^b}}\ .
\end{aligned}
$$

7.1 Localization

Define

$$
\Omega = [(x_1)_{\min}, (x_1)_{\max}] \times [(x_2)_{\min}, (x_2)_{\max}]\ .
$$

As is typically the case, we assume that the Green's function $g(x_1, x_2, \Delta\tau)$ decays to zero as $|x_1|, |x_2| \to \infty$. More precisely, we assume that

Assumption 1 (Decay of Green's Function)**.** *For some $A > 0$, $g(x_1, x_2, \Delta\tau)$ is negligible if* $\min(|x_1|, |x_2|) > A$.

We choose our region Ω so that we can assume that the Green's function is negligible for $(x_1,x_2) \notin \Omega$. As such we replace the Fourier transform pair (27) by their Fourier series equivalent

$$G(\omega_1^k, \omega_2^j, \Delta\tau) \simeq \int_{(x_1)_{\min}}^{(x_1)_{\max}} \int_{(x_2)_{\min}}^{(x_2)_{\max}} g(x_1, x_2, \Delta\tau)\; e^{-2\pi i \omega_1^k x_1}\; e^{-2\pi i \omega_2^j x_2}\; dx_1\; dx_2$$

$$g(x_1, x_2, \Delta\tau) = \frac{1}{P_1 P_2} \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} G(\omega_1^k, \omega_2^j, \Delta\tau)\; e^{2\pi i \omega_1^k x_1}\; e^{2\pi i \omega_2^j x_2} \tag{30}$$

with

$$P_1 = (x_1)_{\max} - (x_1)_{\min}\ ;\ P_2 = (x_2)_{\max} - (x_2)_{\min}; \quad \omega_1^k = \frac{k}{P_1}\ ;\ \omega_2^j = \frac{j}{P_2}$$

along with a localized version of equation (26)

$$v(x_1, x_2, \tau + \Delta\tau) = \int_{(x_1)_{\min}}^{(x_1)_{\max}} \int_{(x_2)_{\min}}^{(x_2)_{\max}} g(x_1 - z_1, x_2 - z_2, \Delta\tau)\; v(z_1, z_2, \tau)\; dz_1\; dz_2\ . \tag{31}$$

Scaling factors in equation (30) are selected to correspond to a finite domain form of equation (27).

7.2 Periodic Extension

We have informally derived the expression for the Green's function by localizing the infinite domain Green's function. Localizing the problem on Ω, and using a Fourier series representation of the Green's function makes the implicit assumption of periodic extension.

Assumption 2. *Periodicity*

1. *The control problem is defined on the finite domain* Ω*.*
2. *The solution* $v(x_1, x_2, \tau)$ *is extended periodically*

$$v(x_1 \pm P_1, x_2, \tau) = v(x_1, x_2 \pm P_2, \tau) = v(x_1, x_2, \tau)\ . \tag{32}$$

3. *The jump size density functions* $f^s(y_1), f^b(y_2)$ *are defined on* $y_1 \in [(x_1)_{\min}, (x_1)_{\max}]$ *and* $y_2 \in [(x_2)_{\min}, (x_2)_{\max}]$*, with periodic extension*

$$f^s(y_1 + P_1) = f^s(y_1) \quad \textit{and} \quad f^b(y_2 + P_2) = f^b(y_2)\ . \tag{33}$$

4. *Periodic boundary conditions (32)-33) are specified for the PIDE problem (25).*
5. *From equation (30) we can also see that*

$$g(x_1 \pm P_1, x_2, \tau) = g(x_1, x_2 \pm P_2, \tau) = g(x_1, x_2, \tau).$$

It is more rigorous to make Assumption 2, and then derive the Green's function. It is straightforward to verify that this yields the Fourier series representation (30), with $G(\omega_1^k, \omega_2^j)$ given by equations (28)-(29). In particular, we have that (Garroni and Menaldi, 1992)

$$g(x_1, x_2, \Delta\tau) \geq 0$$

$$\int_\Omega g(x_1, x_2, \Delta\tau) dx_1\; dx_2 = 1\ .$$

7.3 Discretization

We discretize equation (31) on a grid $\{(x_1, x_2)_{j,k}\}, \{(z_1, z_2)_{j,k}\}$, by setting

$$
\begin{aligned}
x_1^j = \hat{x}_1^0 + j\Delta x_1 \ ; \ x_2^k = \hat{x}_2^0 + k\Delta x_2 \ ; \ j = -\frac{N_1}{2}, \ldots \frac{N_1}{2} - 1 \ ; \ k = -\frac{N_2}{2}, \ldots \frac{N_2}{2} - 1 \\
\Delta x_1 = \frac{P_1}{N_1} \ ; \ \Delta x_2 = \frac{P_2}{N_2} \\
P_1 = (x_1)_{\max} - (x_1)_{\min} \ ; \ P_2 = (x_2)_{\max} - (x_2)_{\min} \\
v_{j,k}(\tau) \equiv v(x_1^j, x_2^k, \tau).
\end{aligned}
$$

Define linear basis functions as

$$
\phi_1^j(x_1) = \begin{cases} \frac{(x_1 - (x_1^j - \Delta x_1))}{\Delta x_1} & x_1^j - \Delta x_1 \leq x_1 \leq x_1^j \\ \frac{(x_1^j + \Delta x_1 - x_1)}{\Delta x_1} & x_1^j \leq x_1 \leq x_1^j + \Delta x_1 \\ 0 & \text{otherwise} \end{cases} .
$$

with a similar definition for $\phi_2^k(x_2)$. We can then represent the solution $v(x_1, x_2, \tau)$ as a linear combination of the basis functions

$$
v(x_1, x_2, \tau) \simeq \sum_{j=-N_1/2}^{N_1/2-1} \sum_{k=-N_2/2}^{N_2/2-1} \phi_1^j(x_1)\phi_2^k(x_2)v_{j,k}(\tau) \ . \tag{34}
$$

Substituting equation (34) into equation (31) then gives

$$
\begin{aligned}
& v_{\ell,m}(\tau + \Delta\tau) \\
&= \sum_{j=-N_1/2}^{N_1/2-1} \sum_{k=-N_2/2}^{N_2/2-1} v_{j,k}(\tau) \int_{(x_1)_{\min}}^{(x_1)_{\max}} \int_{(x_2)_{\min}}^{(x_2)_{\max}} \phi_1^j(x_1) \ \phi_2^k(x_2) \ g(x_1^\ell - x_1, x_2^m - x_2, \Delta\tau) \ dx_1 \ dx_2 \ .
\end{aligned} \tag{35}
$$

For a given pair (j, k) the double integrals then simplify to

$$
\begin{aligned}
& \int_{(x_1)_{\min}}^{(x_1)_{\max}} \int_{(x_2)_{\min}}^{(x_2)_{\max}} \phi_1^j(x_1) \ \phi_2^k(x_2) \ g(x_1^\ell - x_1, x_2^m - x_2, \Delta\tau) \ dx_1 \ dx_2 \\
&\qquad = \int_{x_1^j - \Delta x_1}^{x_1^j + \Delta x_1} \int_{x_2^k - \Delta x_2}^{x_2^k + \Delta x_2} \phi_1^j(x_1) \ \phi_2^k(x_2) \ g(x_1^\ell - x_1, x_2^m - x_2, \Delta\tau) \ dx_1 \ dx_2 \\
&\qquad = \int_{x_1^\ell - x_1^j - \Delta x_1}^{x_1^\ell - x_1^j + \Delta x_1} \int_{x_2^m - x_2^k - \Delta x_2}^{x_2^m - x_2^k + \Delta x_2} \phi_1^j(x_1^\ell - x_1) \ \phi_2^k(x_2^m - x_2) \ g(x_1, x_2, \Delta\tau) \ dx_1 \ dx_2 \\
&\qquad = \int_{x_1^\ell - x_1^j - \Delta x_1}^{x_1^\ell - x_1^j + \Delta x_1} \int_{x_2^m - x_2^k - \Delta x_2}^{x_2^m - x_2^k + \Delta x_2} \phi_1^{\ell - j}(x_1)\phi_2^{m-k}(x_2) \ g(x_1, x_2, \Delta\tau) \ dx_1 \ dx_2 \ .
\end{aligned}
$$

Here the last line follows from the property of linear basis functions,

$$
\phi_1^j(x_1^\ell - x_1) = \phi_1^{\ell - j}(x_1), \quad \phi_2^k(x_2^k - x_2) = \phi_2^{m-k}(x_2).
$$

Defining

$$
\tilde{g}_{\ell - j, m-k}(\Delta\tau) \equiv \frac{1}{\Delta x_1 \Delta x_2} \int_{x_1^\ell - x_1^j - \Delta x_1}^{x_1^\ell - x_1^j + \Delta x_1} \int_{x_2^m - x_2^k - \Delta x_2}^{x_2^m - x_2^k + \Delta x_2} \phi_1^{\ell - j}(x_1)\phi_2^{m-k}(x_2) \ g(x_1, x_2, \Delta\tau) \ dx_1 \ dx_2,
$$

then implies that for each $\ell = -N_1/2, \ldots, N_1/2-1$, $m = -N_2/2, \ldots, N_2/2-1$, the convolution (35) can be written as

$$v_{\ell,m}(\tau+\Delta\tau) = \sum_{j=-N_1/2}^{N_1/2-1} \sum_{k=-N_2/2}^{N_2/2-1} v_{j,k}(\tau)\ \tilde{g}_{\ell-j,m-k}(\tau)\ \Delta x_1\ \Delta x_2. \tag{36}$$

Furthermore using the Fourier series form for $g(\cdot)$ from equation (30) we have

$$\begin{aligned}
&\tilde{g}_{\ell-j,m-k}(\Delta\tau)\\
&= \frac{1}{\Delta x_1\ \Delta x_2} \int_{x_1^\ell - x_1^j - \Delta x_1}^{x_1^\ell - x_1^j + \Delta x_1} \int_{x_2^m - x_2^k - \Delta x_2}^{x_2^m - x_2^k + \Delta x_2} \phi_1^{\ell-j}(x_1)\ \phi_2^{m-k}(x_2)\ g(x_1,x_2,\tau)\ dx_1\ dx_2\\
&= \frac{1}{P_1 P_2} \sum_{u=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \left(\frac{\sin^2 \pi\omega_1^u \Delta x_1}{(\pi\omega_1^u \Delta x_1)^2}\right) \left(\frac{\sin^2 \pi\omega_2^p \Delta x_2}{(\pi\omega_2^p \Delta x_2)^2}\right) G(\omega_1^u, \omega_2^p, \Delta\tau)\ e^{2\pi i \omega_1^u (x_1^\ell - x_1^j)}\ e^{2\pi i \omega_2^p (x_2^m - x_2^k)}\\
&= \frac{1}{P_1 P_2} \sum_{u=-\infty}^{\infty} \sum_{p=-\infty}^{\infty} \left(\frac{\sin^2 \pi\omega_1^u \Delta x_1}{(\pi\omega_1^u \Delta x_1)^2}\right) \left(\frac{\sin^2 \pi\omega_2^p \Delta x_2}{(\pi\omega_2^p \Delta x_2)^2}\right) G(\omega_1^u, \omega_2^p, \Delta\tau)\ e^{2\pi i u(\ell-j)/N_1}\ e^{2\pi i p(m-k)/N_2}.
\end{aligned} \tag{37}$$

Here for the inside integrals in (37) we have used the following formula for linear basis functions

$$\frac{1}{\Delta x}\int_{y-\Delta x}^{y+\Delta x} e^{2\pi i\omega x}\phi(x)dx = e^{2\pi i\omega y}\frac{\sin^2 \pi\omega\Delta x}{(\pi\omega\Delta x)^2}$$

applied separately for ϕ_1 and ϕ_2. Note that equation (37) is independent of $\hat{x}_1^0, \hat{x}_2^0$.[6]

For future reference, we note the DFT pair for any grid function[7] $h(x_1^j, x_2^k) = h_{jk}$ is given by

$$H(\omega_1^\ell, \omega_2^m, \Delta\tau) = \frac{P_1}{N_1}\frac{P_2}{N_2}\sum_{j,k} e^{-2\pi i\ell j/N_1} e^{-2\pi i m k/N_2} h_{jk} \tag{38}$$

$$h_{pn}(\Delta\tau) = \frac{1}{P_1 P_2}\sum_{\ell,m} e^{2\pi i\ell p/N_1} e^{2\pi i m n/N_2} H(\omega_1^\ell, \omega_2^m, \Delta\tau) \tag{39}$$

which can be verified by substituting equation (38) into equation (39)

$$h_{pn}(\Delta\tau) = \frac{1}{N_1 N_2}\sum_{j,k} h_{jk} \sum_{\ell,m} e^{2\pi i\ell(p-j)/N_1} e^{2\pi i m(n-k)/N_2} = h_{pn}\ ,$$

since

$$\sum_{\ell,m} e^{2\pi i\ell(p-j)/N_1} e^{2\pi i m(n-k)/N_2} = \begin{cases} N_1 N_2 & p = j \text{ and } n = k \\ 0 & \text{otherwise} \end{cases}. \tag{40}$$

We denote this pair by $H = DFT(h)$, $h = IDFT(H)$.

7.3.1 Discrete Index Domain

Although we have defined the solution on the physical domain Ω, note that from equation (37) $\tilde{g}_{\ell-j,m-n}$ is independent of $(\hat{x}_1^0, \hat{x}_2^0)$. In addition, the DFT pair are defined indepen-

[6] For the case of $u = 0$, $p = 0$ in equation (37) we take the limiting value as $\omega_1^0 \to 0$, $\omega_2^0 \to 0$.

[7] Note that the DFT here is independent of any physical coordinates.

dently of the original grid. Consequently, in order to avoid confusion concerning the various physical and grid domains, we will refer only to discrete index domains. Define

$$\mathcal{D} = \left\{ (\ell,j) \Big| -\frac{N_1}{2} \le \ell \le \frac{N_1}{2} - 1 \; , \; -\frac{N_2}{2} \le j \le \frac{N_2}{2} - 1 \right\} .$$

Then $\tilde{g}_{\ell,j}, v_{\ell,j}$ are both in the index domain $\mathcal{D}$.

7.4 Efficient Computation of the Discrete Convolution

The discrete convolution (36) is performed at each rebalancing date, hence an efficient evaluation is desirable. The convolution can be written as

$$\begin{aligned} v_{\ell,m}(\tau + \Delta\tau) &= \sum_{j=-N_1/2}^{N_1/2-1} \sum_{k=-N_2/2}^{N_2/2-1} v_{j,k}(\tau) \tilde{g}_{\ell-j,m-k}(\Delta\tau) \; \Delta x_1 \; \Delta x_2 \\ &= \frac{P_1}{N_1} \frac{P_2}{N_2} \sum_{j,k} \tilde{g}_{\ell-j,m-k}(\Delta\tau) v_{j,k}(\tau) \; . \end{aligned} \tag{41}$$

Using equations (38) and (39) to write $\tilde{g}(\Delta\tau) = IDFT(\tilde{G}(\Delta\tau)), v = IDFT(V)$ in equation (41) gives

$$\begin{aligned} v_{\ell,m}(\tau + \Delta\tau) &= \frac{1}{N_1 P_1} \frac{1}{N_2 P_2} \sum_{j,k} \sum_{p,u} \tilde{G}_{pu}(\Delta\tau) e^{2\pi i p(\ell-j)/N_1} e^{2\pi i u(m-k)/N_2} \sum_{q,r} V_{q,r} e^{2\pi i q j/N_1} e^{2\pi i r k/N_2} \\ &= \frac{1}{N_1 P_1} \frac{1}{N_2 P_2} \sum_{p,u} \sum_{q,r} \tilde{G}_{pu}(\Delta\tau) V_{qr} e^{2\pi i p\ell/N_1} e^{2\pi i u m/N_2} \sum_{j,k} e^{2\pi i j(q-p)/N_1} e^{2\pi i k(r-u)/N_2} \\ &= \frac{1}{P_1} \frac{1}{P_2} \sum_{p,u} \tilde{G}_{pu}(\Delta\tau) V_{pu} e^{2\pi i p\ell/N_1} e^{2\pi i u m/N_2} \end{aligned}$$

where we have used equation (40) in the last step. More compactly we can write

$$\tilde{v}(\tau + \Delta\tau) = IDFT\big(\tilde{G}(\Delta\tau) \circ DFT\,(v(\tau)) \; \big) \, , \tag{42}$$

where $y \circ z$ is the Hadamard product of vectors y, z.

7.5 δ-Monotonicity

Using the definition for $\tilde{g}$ from equation (37), then this represents an exact integration over linear basis functions of the exact Green's function (with periodic boundary conditions) on Ω. Hence

$$\tilde{g}_{j,k} \ge 0 \; ; \; \forall (j,k) \in \mathcal{D} \tag{43}$$

This gives rise to an important *monotonicity* property (Forsyth and Labahn, 2019). Suppose we have two discrete solutions $v_{\ell,m}(\tau), u_{\ell,m}(\tau)$ such that

$$v_{\ell,m}(\tau) > u_{\ell,m}(\tau) \; ; \; \forall (\ell,m) \in \mathcal{D} \; . \tag{44}$$

Then, using the time advance algorithm (36), it follows that the *discrete comparison principle* holds:

$$v_{\ell,m}(\tau+\Delta\tau) > u_{\ell,m}(\tau+\Delta\tau) \; ; \; \forall(\ell,m) \in \mathcal{D}.$$

The positivity condition (43) ensures that the order property (44) is preserved by the discretized algorithm. This is, of course, a property of the exact solution. This property is important since we compare the values of the discrete solution in order to determine the optimal control. See Forsyth and Labahn (2019) for more discussion of this.

However, note that the dimension of $\tilde{g}$ is (N_1, N_2), but we need to sum an infinite series to determine the discrete values of $\tilde{g}_{m,u}$ from equation (36). In practice, we truncate the series in equation (37), that is,

$$\begin{aligned} &\tilde{g}_{\ell-j,m-k}(\Delta\tau) \\ &\simeq \frac{1}{P_1 P_2} \sum_{u=-N_1^g/2}^{N_1^g/2-1} \sum_{p=-N_2^g/2}^{N_2^g/2-1} \left(\frac{\sin^2 \pi\omega_1^u \Delta x_1}{(\pi\omega_1^u \Delta x_1)^2} \right) \left(\frac{\sin^2 \pi\omega_2^p \Delta x_2}{(\pi\omega_2^p \Delta x_2)^2} \right) G(\omega_1^u, \omega_2^p, \Delta\tau)\, e^{2\pi i \frac{u(\ell-j)}{N_1}} \; e^{2\pi i \frac{p(m-k)}{N_2}} \end{aligned} \tag{45}$$

where the grid sizes (N_1^g, N_2^g) can be chosen independently from (N_1, N_2). However, truncating the Fourier series means that the positivity condition (43) may not hold any longer. In Forsyth and Labahn (2019), it is suggested that (N_1^g, N_2^g) be selected so that

$$\sum_{m=-N_1^g/2}^{N_1^g/2+1} \sum_{n=-N_2^g/2}^{N_2^g/2-1} \Delta x_1 \Delta x_2 \; |\min(\tilde{g}_{m,n}, 0)| < \delta \frac{\Delta\tau}{T} \; . \tag{46}$$

This ensures that the cumulative effect (after $(T/(\Delta\tau)$ steps) of non-positivity means that the discrete comparison principle holds to $O(\delta)$. Note that if the timestep $(\Delta\tau)$ is constant (which is usually the case), then equation (45) needs to be computed only once. Hence it is inexpensive to select sizes (N_1^g, N_2^g) which satisfy condition (46). Equation (45) can also be computed efficiently using FFTs, assuming $N_1^g/N_1, N_2^g/N_2$ are powers of two (see Forsyth and Labahn (2019).

8 Periodicity and the problem of Wrap-around error

Localization of the Green's function effectively implies a periodic extension of the Green's function and the solution (see Assumption 2). In option pricing applications, this *wrap-around* error does not cause difficulty. However, in control applications, impulse controls (such as rebalancing) may require extensive use of solution values near the edges of the grid. For example, derisking by investing in an all bond portfolio results in an impulse control which drives $s \to 0$ instantaneously. This can cause significant wrap-around pollution (see (Lippa, 2013; Ruijter et al., 2013; Ignatieva et al., 2018)).

Given the localized problem defined on $\Omega = [(x_1)_{\min}, (x_1)_{\max}] \times [(x_2)_{\min}, (x_2)_{\max}]$, with widths

$$P_1 = (x_1)_{\max} - (x_1)_{\min} \quad \text{and} \quad P_2 = (x_2)_{\max} - (x_2)_{\min}$$

we construct an auxiliary grid with $N_1^\dagger = 2N_1$ nodes in the x_1 direction, and $N_2^\dagger = 2N_2$ nodes in the x_2 direction, on the domain $\Omega^\dagger = [(x_1)_{\min}^\dagger, (x_1)_{\max}^\dagger] \times [(x_2)_{\min}^\dagger, (x_2)_{\max}^\dagger]$, with

$$\begin{aligned} (x_1)_{\min}^\dagger &= (x_1)_{\min} - \frac{P_1}{2} \; ; \quad (x_1)_{\max}^\dagger = (x_1)_{\max} + \frac{P_1}{2} \\ (x_2)_{\min}^\dagger &= (x_2)_{\min} - \frac{P_2}{2} \; ; \quad (x_2)_{\max}^\dagger = (x_2)_{\max} + \frac{P_2}{2} \end{aligned}$$

$$P_1^\dagger = (x_1)_{\max}^\dagger - (x_1)_{\min}^\dagger = 2P_1 \; ; \quad P_2^\dagger = (x_2)_{\max}^\dagger - (x_2)_{\min}^\dagger = 2P_2$$

Although we have increased the size of the physical domain Ω, we remind the reader of the discussion in Section 7.3.1. It will be more convenient in the following to refer to the extended index domain

$$\mathcal{D}^\dagger = \left\{(k,j) \;\middle|\; -N_1^\dagger/2 \le k \le N_1^\dagger/2 - 1 \;,\; -N_2^\dagger/2 \le j \le N_2^\dagger/2 - 1\right\} .$$

Denote the projected Green's function for grid indexes in $\mathcal{D}^\dagger$ by $\tilde{g}^\dagger$. Note from equation (37) that $\tilde{g}^\dagger$ does not depend on the actual physical domain, but only on $(\Delta x_1, \Delta x_2)$. We construct and store the DFT of the projection of the Green's function $\tilde{G}^\dagger(\Delta\tau) = DFT(\tilde{g}^\dagger(\Delta\tau))$ on this auxiliary index grid $\mathcal{D}^\dagger$.

Before applying the time advance algorithm (42) we form the padded array $v^\dagger(\tau)$:

$$\begin{aligned}
v^\dagger(x_1^m, x_2^k, \tau) &= v(x_1^m, x_2^k, \tau) \;, && (m,k) \in \mathcal{D} \\
&= v(x_1^{-N_1/2}, x_2^k, \tau) \;, && m \in [-N_1^\dagger/2, -N_1/2 - 1];\;\; k \in [-N_2/2, N_2 - 1] \\
&= v(x_1^{-N_1/2}, x_2^{-N_2/2}, \tau), && m \in [-N_1^\dagger/2, -N_1/2 - 1];\;\; k \in [-N_2^\dagger, -N_2 - 1] \\
&= v(x_1^m, x_2^{-N_2/2}, \tau) \;, && m \in [-N_1/2, N_1/2 - 1] \;;\;\; k \in [-N_1^\dagger/2, -N_2/2 - 1] \\
&= A(x_1^m, x_2^k, \tau) \;, && m > N_1/2 - 1 \quad \text{or} \quad k > N_2/2 - 1 \;,
\end{aligned} \tag{47}$$

where $A(x_1^m, x_2^k, \tau)$ is an asymptotic form of the solution which we assume to be available from financial reasoning. On the auxiliary grid, the points where $x_1 < (x_1)_{min}$ or $x_2 < (x_2)_{\min}$ correspond to the points where $s \to 0$ or $b \to 0$, with very small grid spacing. Hence extending the solution by constant values to the left and bottom is expected to generate a very small error.

We then modify the time advance algorithm (42) as shown in Algorithm 1. Note that the step (3) discards the computed solution in the padded areas $\{v^\dagger(\tau + \Delta\tau)_{i,j} \big| (i,j) \in \mathcal{D}^\dagger - \mathcal{D}\}$, since these points may be contaminated by wrap-around errors.

Require: $v(\tau)$; $\widetilde{G}^\dagger$
1: Form $v^\dagger(\tau)$ using (47)
2: $v^\dagger(\tau + \Delta\tau) = IDFT\big(\tilde{G}^\dagger \circ DFT\left(v^\dagger(\tau)\right)\big)$ {IDFT (Hadamard product)}
3: $v_{k,m}(\tau + \Delta\tau) = v^\dagger_{k,m}(\tau + \Delta\tau)$, $(k, m) \in \mathcal{D}$ {Discard values in $\mathcal{D}^\dagger - \mathcal{D}$.}

Algorithm 1: *Advance time $v(\tau) \to v(\tau + \Delta\tau)$.*

8.1 Error Due to Wrap-around

Let $\tau^n = n\Delta\tau$ and $\mathcal{N}\Delta\tau = T$, with $\Delta\tau$ fixed. We can then write the discrete convolution (41) in the physical domain, compactly as

$$v_{\ell,m}^{n+1} = \Delta x_1 \Delta x_2 \sum_{(j,k)\in\mathcal{D}^\dagger} \tilde{g}^\dagger_{\ell-j,m-k} v^\dagger_{j,k}(\tau^n) \;,\; (\ell,m) \in \mathcal{D} \;. \tag{48}$$

Note that the right hand side of equation (48) uses the padded values for $v^\dagger_{j,k}$, but generates the unpadded $v_{\ell,m}^{n+1}$ on the left hand side. In addition, $\tilde{g}^\dagger$ is periodic with period $P_1^\dagger$ in the x_1 direction, and period $P_2^\dagger$ in the x_2 direction.

The use of Algorithm 1 means that the periodic extension of $\tilde{g}$ is used in equation (48) for any terms of the sum such that

$$(\ell - j, m - k) \notin \mathcal{D}^\dagger \ .$$

This periodic extension *wraps around* the solution domain. For example, points with indexes $\ell - j < -N_1^\dagger/2$ reference points with index $\ell - j + N_1^\dagger$, which is at the opposite side of the grid, and potentially generates wrap-around error. See also Appendix 12 for further discussion of this.

Definition 1 (Wraparound error.) Assume $\tilde{g}^\dagger_{\ell,m}$ for $(\ell, m) \in \mathcal{D}^\dagger$ is periodic

$$\tilde{g}^\dagger(\ell \pm N_1^\dagger, m) = \tilde{g}^\dagger(\ell, m \pm N_2^\dagger) = \tilde{g}^\dagger(\ell, m) \ . \tag{49}$$

Suppose $v^\dagger_{j,k}(\tau^n)$ is determined by boundary data for $(x_1^j, x_2^m) \in (\Omega^\dagger - \Omega)$, which we assume to be exact,[8] with $(N_1^\dagger, N_2^\dagger) = (2N_1, 2N_2)$. Then, the wraparound error for convolution (48) evaluated using DFTs, at timestep $n\Delta\tau$, denoted by ϵ^n_{wrap} is

$$\varepsilon^n_{wrap} = \Delta x_1 \Delta x_2 \max_{(\ell,m)\in\mathcal{D}} \sum_{(j,k)\in\mathcal{D}^\dagger} \left| \tilde{g}^\dagger_{\ell-j,m-k} v^\dagger_{j,k}(\tau^n) \right| \mathbf{1}_{(\ell-j,m-k)\notin\mathcal{D}^\dagger} \ . \tag{50}$$

We now state a theorem on the effectiveness of our padding method. See Appendix 12 for a proof.

Theorem 1 (Wraparound error) *Suppose $\tilde{g}^\dagger_{\ell,m}$ for $(\ell, m) \in \mathcal{D}^\dagger$ is periodic as defined in equation (49), and that $v^\dagger_{\ell,m}(\tau^n)$ be determined by boundary data for $(\ell, m) \in (\mathcal{D}^\dagger - \mathcal{D})$, with $(N_1^\dagger, N_2^\dagger) = 2(N_1, N_2)$. Assume further that there exists a constant C such that*

$$\left| v^\dagger_{\ell,m}(\tau^n) \right| \leq C \ , \quad (\ell, m) \in \mathcal{D}^\dagger \ , \quad \forall\, n \ \text{ such that } \ (n\Delta\tau) \leq T \ ,$$

and that we choose (N_1, N_2) sufficiently large so that

$$\left(\Delta x_1 \Delta x_2 \sum_{(j,k)\in\mathcal{D}^\dagger - \mathcal{D}} |\tilde{g}^\dagger_{j,k}| \right) \leq \varepsilon_b \ \Delta\tau \ . \tag{51}$$

Then

$$\varepsilon^n_{wrap} \leq C\varepsilon_b \ \Delta\tau \ , \tag{52}$$

and the wraparound error after $\mathcal{N}$ steps is bounded by

$$\varepsilon^{\mathcal{N}} \leq T\epsilon_b C \quad \text{where } T = \mathcal{N}\Delta\tau \ .$$

Remark 2 (Asymptotic Form of Green's Function) For a pure diffusion in one dimension, the Green's function for the Black-Scholes equation behaves like

$$g(x - x', \Delta\tau) = O\left(e^{-(x-x')^2/(\sigma^2\Delta\tau)} \right) \ \text{as } |x - x'| \to \infty \ .$$

Similar asymptotic exponential decay as $\max(|x_1|,|x_2|) \to \infty$ is also seen for the full 2-d Green's function(Garroni and Menaldi, 1992). Hence ensuring that condition (51) holds is not onerous.

[8] In practice, we use the method in equation (47), which we expect will generate a small error.

8.2 Numerical Method for Optimal Control

As mentioned in Section 16, we need the optimal pair (q, p) at each rebalancing date. The optimal controls are obtained by first discretizing the controls and thereafter determining the controls by exhaustive search. Recall that the PIDE grid spacing is $\Delta x_1, \Delta x_2$. Let N_q and N_p denote the number of points in the discretizations of q and p, respectively. We will also need to discretize the possible values of total wealth $w = s + b$, with N_w denoting the number of discrete wealth values. This approach (discretization and search), is then guaranteed to converge to the viscosity solution of the optimal control problem in the limit as

$$\Delta x_1, \Delta x_2 \to 0 \ ; \ N_q, N_p, N_w \to \infty \ ; \ \delta \to 0 \ ,$$

provided that the complete numerical scheme is $\delta-$monotone, consistent and ℓ_∞ stable, see (Barles and Souganidis, 1991; Forsyth and Labahn, 2019).

Given an array of discrete values, that is, $v_{j,k}(\tau_n^-)$, we denote the linear interpolant of the grid values at an arbitrary point (z_1, z_2) by

$$v(\mathcal{I}(z_1, z_2), \tau_n^-) \ .$$

The numerical analogue of equation (20) requires a temporary array, $h(w_k), k = 1, \ldots, N_w$, where $w_k = s + b$, the total wealth. This array is determined by

$$\begin{aligned}
\underbrace{p^*(w_k)}_{k=1,\ldots,N_w} &= \underset{\substack{j=1,\ldots,N_p \\ p_j \in \mathcal{Z}_p(w_k, \tau_n^-)}}{\arg\max} \ v\Big(\mathcal{I}\ \big(\ \log(w_k p_j)\ ,\ \log(w_k(1-p_j))\ \big)\ , \tau_n^- \Big) \\
\underbrace{h(w_k)}_{k=1,\ldots,N_w} &= v\Big(\mathcal{I}\Big(\ \log(\ w_k p^*(w_k)\ )\ ,\ \log(\ w_k(1-p^*(w_k))\ )\ \Big), \tau_n^-\Big) \ .
\end{aligned} \qquad (53)$$

In order to compute (21), we break this down into two steps, first using

$$\underbrace{q^*(w_k)}_{k=1,\ldots,N_w} = \underset{\substack{j=1,\ldots,N_q \\ q_j \in \mathcal{Z}_q(w_k, \tau_n)}}{\arg\max} \ \Big\{ \ q_j + h\ \big(\ \mathcal{I}(w_k - q_j)\ \big)\ \Big\} \qquad (54)$$

where $h(\mathcal{I}(\cdot))$ refers to one dimensional interpolation for the one dimensional array $h(w_k)$. Then

$$\underbrace{v(x_1^j, x_2^k, \tau_n^+)}_{\substack{j=-N_1/2,\ldots,N_1/2-1 \\ k=-N_2/2,\ldots,N_2/2-1 \\ w_{j,k}=e^{x_1^j}+e^{x_2^k}}} = q^*\ \big(\ \mathcal{I}(w_{j,k})\ \big) + h\Big(\ \mathcal{I}\ \big(\ w_{j,k} - q^*\ \big(\ \mathcal{I}(w_{j,k})\ \big)\ \big)\ \Big) \ . \qquad (55)$$

8.3 Summary: Algorithm for Problem (16).

Algorithm 2 summarizes the final algorithm. The initial set-up cost is computing the Fourier weights on the extended domain $\widetilde{G}^\dagger$. Each timestep consists of first determining the optimal control, then followed by advancing to the next rebalancing time using the precomputed weights $\widetilde{G}^\dagger$.

Remark 3 (Use of linear interpolation) Equations (53-55) make heavy use of linear interpolation. As discussed in Forsyth and Labahn (2019)) this is the only interpolation

Require: Weights $\widetilde{G}^{\dagger}$ in padded Fourier domain satisfying tolerance δ (see equations (45-46))

Input: number of timesteps $\mathcal{N}$, initial solution $(v^0)^-$
for $n = 1, \ldots, \mathcal{N}$ **do** {Timestep loop}
  Optimal control $v(\tau_{n-1}^-) \rightarrow v(\tau_{n-1}^+)$ equations (53-55)
  Advance time $v(\tau_{n-1}^+) \rightarrow v(\tau_n^-)$ using Algorithm 1
end for{End timestep loop}

Algorithm 2: *Monotone Fourier method.*

method (in general) which will preserve δ-monotonicity. This means that the convergence rate cannot exceed second order in the grid spacing. This is in contrast to the methods in (Fang and Oosterlee, 2008; 2009; Ruijter et al., 2013), where, in some circumstances, high order rates of convergence can be achieved, but at the cost of possible non-monotonicity.

Remark 4 (Insolvency Between Rebalancing Times) For $b > 0$, insolvency cannot occur between rebalancing dates. However, for $b < 0$, it is possible that a large drop in the value of stocks could result in $s + b < 0$ between rebalancing times. Numerical tests using subtimesteps between rebalancing times (and checking for insolvency) did not show any significant changes to the final values at $t = T$. Hence, in all our reported numerical tests, we only check for insolvency at rebalancing times.

8.4 Final Algorithm: Problem EW-ES (18).

The final step for the complete solution of the EW-ES problem is the optimization step (18). We solve Problem EW-ES on a sequence of grids, with increasing number of grid nodes. On the coarsest grid, we discretize W^* and find the maximum of equation (18) by exhaustive search. Note that each evaluation of the objective function in (18) requires a solution of Problem 16. Using the coarse grid value of W^* as a starting value, we use a one-dimensional optimization algorithm to maximize the objective function (18) on each finer grid.

9 Numerical Example and Results

We consider a 65 year old retiree who has \$1,000K (one million) in pension savings. The retiree needs to withdraw a minimum of 30K per year, but has no use for more than 60K per year. All amounts are inflation adjusted. We assume that the retiree has other sources of income (work pension, government benefits) which, when added to the 30K per year of withdrawals, accumulate to a satisfactory income level.

We consider a 30 year time horizon, since this is consistent with the Bengen (1994) scenario. Recall that the probability of a 65 year old Canadian male attaining the age of 95 is about 0.13, so this is a fairly conservative assumption. Similar probabilities are also true for other western countries.

The pension savings are assumed to be held in a tax advantaged account, so that there are no tax consequences for rebalancing. We consider the withdrawals to be the desired amount before any income taxes.

Note that we also consider that the retiree has mortgage free real estate worth 400K. We can regard this as a hedge of last resort.

Remark 5 (Reverse Mortgage Hedge) If we assume that a reverse mortgage can be used to obtain a non-recourse loan of half the value of the real estate ($200K$), then from a risk management point of view, any strategy which results in an expected shortfall $> -200K$ is probably acceptable. All quantities are assumed inflation adjusted.

Other scenario details are listed in Table 1. As a point of comparison, the Bengen (1994) strategy would withdraw a fixed amount of 40K (real) per year (4% of the initial 1,000K), and rebalance to a constant weight of 50% in stocks at each rebalancing date.

Investment horizon T (years)	30.0
Equity market index	CRSP Cap-weighted index (real)
Bond index	30-day T-bill (US) (real)
Initial portfolio value W_0	1000
Mortgage free real estate	400
Cash withdrawal/rebalancing times	$t = 0, 1.0, 2.0, \ldots, 29.0$
Maximum withdrawal (per year)	$q_{\max} = 60$
Minimum withdrawal (per year)	$q_{\min} = 30$
Equity fraction range	$[0, p_{\max}]$
	$p_{\max} = 0.5, 0.8, 1.0, 1.3$
Borrowing spread μ_c^b	0.03
Rebalancing interval (years)	1.0
α (Expected shortfall parameter)	.05
Stabilization ϵ (see equation (11))	-10^{-4}
Market parameters	See Table 2

Table 1: *Input data for examples. Monetary units: thousands of dollars. All amounts inflation adjusted.*

For the computational study in this paper, we use data from the Center for Research in Security Prices (CRSP) on a monthly basis from 1926:1 to 2024:12.[9] The specific indices used are the CRSP 30 day U.S. T-bill index for the bond asset, and the CRSP cap-weighted total return index for the stock asset[10]. Retirees are, naturally, concerned with preserving real (not nominal) spending power. Hence, we use the US CPI index (from CSRP) to adjust these indexes for inflation. We use the above market data in two different ways in subsequent investigations.

We use the threshold technique (Mancini, 2009; Cont and Mancini, 2011; Dang and Forsyth, 2016) to estimate the parameters for the parametric stochastic process models. Since the index data is in real terms, all parameters reflect real returns. Table 2 shows the results of calibrating the models to the historical data. The correlation ρ_{sb} is computed by removing any returns which occur at times corresponding to jumps in either series, and then using the sample covariance. Further discussion of the validity of assuming that the stock and bond jumps are independent is given in Forsyth (2020b).

We first compute and store the optimal controls in the *synthetic market*, that is, the stock market processes follow the parametric models (3) and (4). We then use Monte Carlo simulation, coupled with the stored optimal controls, to analyze the properties of the optimal allocation/withdrawal.

[9] More specifically, results presented here were calculated based on data from Historical Indexes, ©2024 Center for Research in Security Prices (CRSP), The University of Chicago Booth School of Business. Wharton Research Data Services was used in preparing this article. This service and the data available thereon constitute valuable intellectual property and trade secrets of WRDS and/or its third-party suppliers.

[10] The stock index includes all distributions for all domestic stocks trading on major U.S. exchanges.

CRSP	μ^s	σ^s	λ^s	u^s	η_1^s	η_2^s	ρ_{sb}
	0.088241	0.147361	0.31313	0.22581	4.3608	5.5309	0.096279
30-day T-bill	μ^b	σ^b	λ^b	u^b	η_1^b	η_2^b	ρ_{sb}
	0.0034	0.0139	0.3838	0.3947	61.510	53.356	0.096279

Table 2: *Parameters for parametric market models (3) and (4), fit to CRSP data (inflation adjusted) for 1926:1 to 2024:12.*

Finally, as a check on robustness, we test the controls (computed in the synthetic market), in the *historical market*, using block bootstrap resampled historical data.

9.1 Results: Convergence

We first check on the convergence of our method. The numerical parameters are given in Table 3.

Grid sizes ($b > 0$)	512×512
	1024×1024
	2048×2048
Extended grid $(N_1^\dagger, N_2^\dagger)$	$(2N_1, 2N_2)$
$\hat{x}_1^0, \hat{x}_2^0$	$\log(100)$
$(x_1)_{\min}, (x_2)_{\min}$	-7.5 + $\hat{x}_1^0$
$(x_1)_{\max}, (x_2)_{\max}$	+10 + $\hat{x}_2^0$
Monotonicity condition δ (equation(46))	10^{-6}
Number of points in w grid N_w (equation (53))	$4N_1$
Number of points in p grid N_p (equation (53))	$N_1/10$
Number of points in q grid N_q (equation (54))	$N_1/10$

Table 3: *Numerical parameters*

Figure 2 shows the EW-ES efficient frontiers, computed using various grid sizes. The grid here refers to the grid for $b > 0$. There is an additional grid (of the same size) for $b < 0$. The optimal control is computed and stored, using the δ-monotone PIDE method. Statistics are then generated using Monte Carlo (MC) simulations, using the stored optimal controls, with 2.56×10^6 MC simulations being used. The curves for all grid sizes essentially overlap, indicating convergence for practical purposes. All subsequent results use the 2048×2048 grid. For comparison, we also show the results for the Bengen (1994) policy, which is clearly much less efficient than the optimal policy.

Table 4 shows detailed convergence results for a single point on the EW-ES frontier ($p_{\max} = 1.0, \kappa = 0.866$). The optimal controls computed using Algorithm 2 are stored, and then used in Monte Carlo (MC) simulations. The value function appears to converge smoothly (for the PIDE method).

It is also interesting to note that, in all our tests using the parameters in Table 3, we find that the wraparound error condition 51) is bounded by

$$\left(\Delta x_1 \Delta x_2 \sum_{(j,k) \in \mathcal{D}^\dagger - \mathcal{D}} |\tilde{g}^\dagger_{j,k}| \right) < 10^{-14} \; . \tag{56}$$

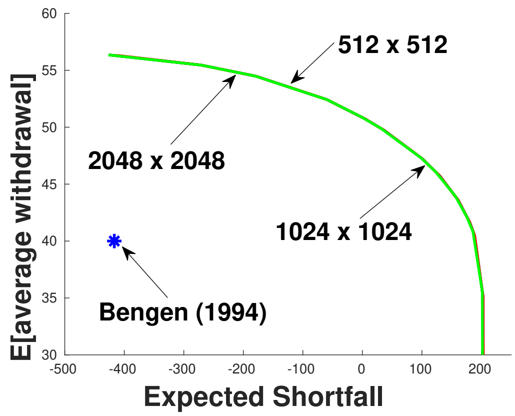


Fig. 2: *EW-ES convergence test. Real stock index: deflated real capitalization weighted CRSP, real bond index: deflated 30 day T-bills. Scenario in Table 1. Parameters in Table 2. The optimal control is determined by solving the using the method in Algorithm 2. Grid refers to the grid used in the PIDE solve, $n_s \times n_b$, where n_s is the number of nodes in the* $\log s$ *direction, and n_b is the number of nodes in the* $\log b$ *direction. Units: thousands of dollars (real). The controls are stored, and then the final results are obtained using a Monte Carlo method, with* 2.56×10^6 *simulations. Maximum fraction in stocks* $p_{\max} = 1.0$. *Bengen (1994) refers to a constant weight in stocks* $p = 0.5$, *rebalanced annually, and constant yearly withdrawals of* 40 *per year. All amounts are inflation adjusted.*

	Algorithm in 2			Monte Carlo	
Grid	ES	$E[\sum_i q_i]/M$	Value Function (13)	ES	$E[\sum_i q_i]/M$
512×512	-14.6176	51.1162	1520.941	-10.302	51.1507
1024×1024	-9.16685	51.0706	1524.251	-7.5173	51.0781
2048×2048	-4.84764	50.9780	1525.179	-3.8866	50.9762

Table 4: *EW-ES convergence test. Real stock index: deflated real capitalization weighted CRSP, real bond index: deflated 30 day T-bills. Scenario in Table 1. Parameters in Table 2. The optimal control is determined by solving the using the method in Algorithm 2. Grid refers to the grid used in the PIDE solve, $n_s \times n_b$, where n_s is the number of nodes in the* $\log s$ *direction, and n_b is the number of nodes in the* $\log b$ *direction. Units: thousands of dollars (real). The controls are stored, and then the results are verified using a Monte Carlo method, with* 2.56×10^6 *simulations. Maximum fraction in stocks* $p_{\max} = 1.0$. $\kappa = 0.8660$.

Note that

$$\left(\Delta x_1 \Delta x_2 \sum_{(j,k) \in \mathcal{D}^\dagger} |\tilde{g}^\dagger_{j,k}| \right) \leq 1 + \delta \simeq 1$$

with δ given in Table 3. Consequently, the numerical result (56) indicates that the wraparound error is only slightly larger than double precision machine epsilon.

9.2 Efficient Frontiers: Synthetic market

Figure 3 shows the efficient frontiers computed using various value of $p_{\max}$ (labelled $pmax$ on the Figure), the maximum values of the fraction in stocks. As must be true mathematically, the curves with higher values of $p_{\max}$ plot above curves with smaller

$p_{\max}$. However, a striking result is that the curves are all fairly tightly clustered, in comparison to the Bengen (1994) strategy. In particular, it seems reasonable to target an $ES \simeq 0$. The expected annual withdrawal for all values of $p_{\max}$ are very close for $ES = 0$.[11] However, this is not quite a free lunch, as we will see when we examine the results more closely.

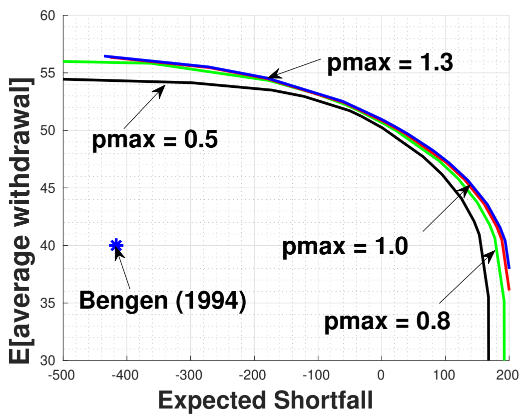


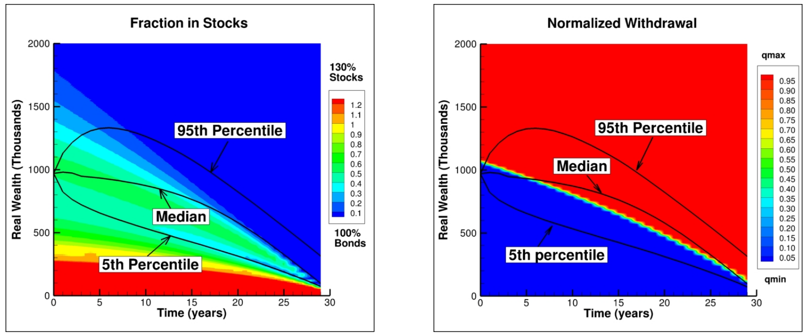
Fig. 3: *Comparison of maximum leverage constraint* $p_{\max}$*, synthetic market. Real stock index: deflated real capitalization weighted CRSP, real bond index: deflated 30 day T-bills. Scenario in Table 1. Parameters in Table 2. The optimal control is determined using Algorithm 2. The controls are stored (using* 2048×2048 *grid), and then the final results are obtained using a Monte Carlo method, with* 2.56×10^6 *simulations. Bengen (1994) refers to a constant weight in stocks* $p = 0.5$*, rebalanced annually, and constant yearly withdrawals of* $40K$ *per year. All amounts are inflation adjusted.*

9.3 Heat Maps: Synthetic Market

In order to gain some insight into the optimal controls, we plot the heat maps of the optimal asset allocation and the optimal withdrawals, for the cases $p_{\max} = 1.3, 1.0, 0.5$ in Figures 4, 5, 6. For each case, we choose the point on the efficient frontier so that $ES \simeq 0$, since this is an *interesting* point on the efficient frontier.

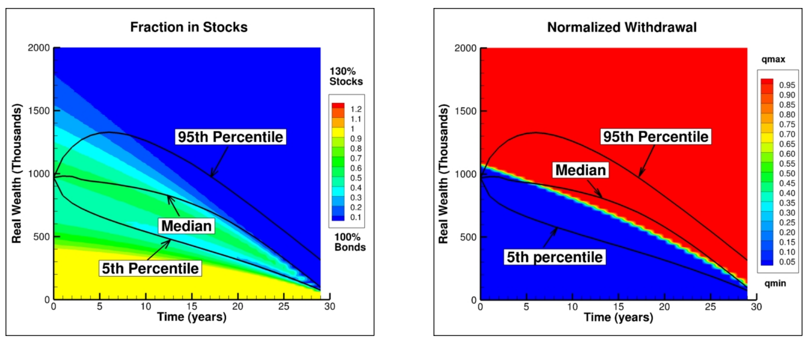
First, note that Figures 4(b), 5(b) and 6(b) show that the optimal withdrawal strategies for a wide range of $p_{\max}$ are very similar. In all cases, the withdrawal control is approximately *bang-bang*, that is, the optimal policy is only to withdraw the maximum or minimum amounts. For an explanation of this (the bang-bang property) see Forsyth (2022). The left hand plots of Figures 4, 5, 6 show the optimal allocation strategies. We use the same colour scale for all plots, with the maximum allocation to stocks being 130%. It is interesting to see that the allocation strategies are quite similar as long as we restrict attention to the areas in the heat maps above the 5th wealth percentile. Large differences in the allocation appear below the 5th wealth percentile.

[11] Although from Remark 5, even $ES \simeq -200K$ is also acceptable. Hence a target of $ES = 0$ is very conservative.



(a) *Fraction in stocks* (b) *Withdrawals*

FIG. 4: $p_{\max} = 1.3$. *Heat map of controls: fraction in stocks and withdrawals, computed using Algorithm 2. Real capitalization weighted CRSP index, and real 30-day T-bills. Scenario given in Table 1. Control computed and stored from the Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year).* $\kappa = 0.8583, EW \simeq 50.9,$ $ES \simeq 0.96$. *Percentiles from bootstrapped historical market. Normalized withdrawal* $(q - q_{\min})/(q_{\max} - q_{\min})$. *Units: thousands of dollars.*



(a) *Fraction in stocks* (b) *Withdrawals*

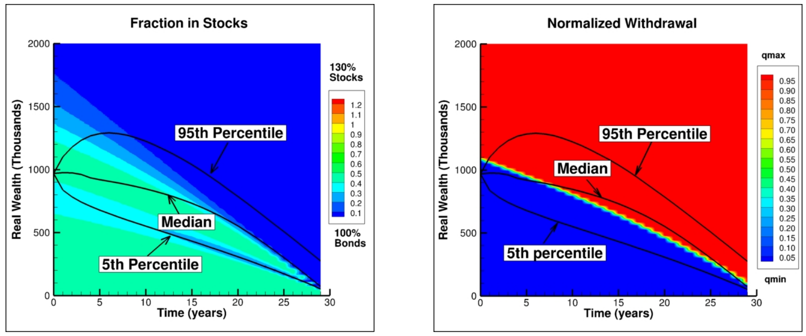
FIG. 5: $p_{\max} = 1.0$. *Heat map of controls: fraction in stocks and withdrawals, computed using Algorithm 2. Real capitalization weighted CRSP index, and real 30-day T-bills. Scenario given in Table 1. Control computed and stored using Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year).* $\kappa = 0.8860, EW = 50.7,$ $ES = 4.6$. *Percentiles from bootstrapped historical market. Normalized withdrawal* $(q - q_{\min})/(q_{\max} - q_{\min})$. *Units: thousands of dollars.*

9.4 Percentiles Versus Time: Synthetic Market

Figures 7. 8 and 9 show the percentiles of fraction in stocks, wealth, and withdrawals versus time. We select the point on the efficient frontier for each case so that $ES \simeq 0$.

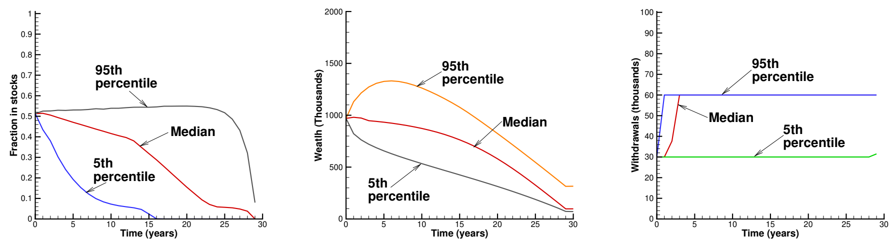
Rather surprisingly, the percentiles fraction in stocks and percentiles wealth are very similar, for all values of $p_{\max}$. However, we do see some differences in the withdrawal percentiles (the rightmost panel in each plot). The median withdrawal is slower to increase to the maximum for $p_{\max} = 0.5$ compared to the case with $p_{\max} = 1.0$.

In addition, we can see that even for $p_{\max} = 1.3$, the 95th percentile for stock allocation is less than 0.6 (Figure 7(a)), suggesting that with this level of ES risk, the optimal policy rarely allocates a large fraction to stocks.



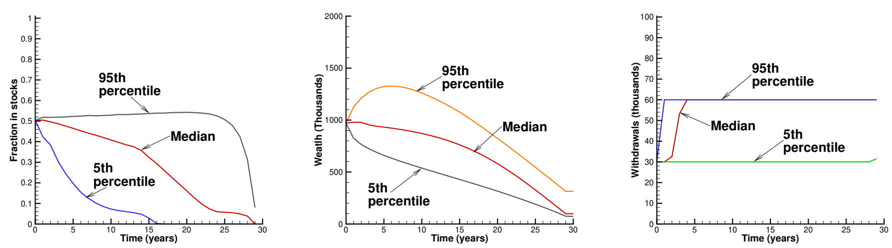
(a) *Fraction in stocks* (b) *Withdrawals*

FIG. 6: $p_{\max} = 0.5$. *Heat map of controls: fraction in stocks and withdrawals, computed using Algorithm 2. Real capitalization weighted CRSP index, and real 30-day T-bills. Scenario given in Table 1. Control computed and stored using Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year).* $\kappa = 1.0, EW \simeq 50.2$, $ES = 1.25$. *Percentiles from bootstrapped historical market. Normalized withdrawal* $(q - q_{\min})/(q_{\max} - q_{\min})$. *Units: thousands of dollars.*



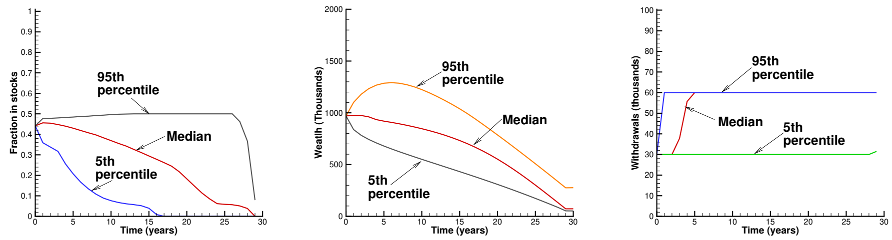
(a) *Percentiles fraction in stocks* (b) *Percentiles wealth* (c) *Percentiles withdrawals*

FIG. 7: *Scenario in Table 1 Strategy computed in synthetic market.* $p_{\max} = 1.3$. *Parameters based on the real CRSP index, and real 30-day T-bills (see Table 2). Control computed and stored from the Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year),* $EW \simeq 50.9$, $ES = 0.96$ *(*$\kappa = 0.8583$*). Units: thousands of dollars.*



(a) *Percentiles fraction in stocks* (b) *Percentiles wealth* (c) *Percentiles withdrawals*

FIG. 8: *Scenario in Table 1 Strategy computed in synthetic market.* $p_{\max} = 1.0$. *Parameters based on the real CRSP index, and real 30-day T-bills (see Table 2). Control computed and stored from Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year),* $EW \simeq 50.7$, $ES = 4.6$ *(*$\kappa = 0.8860$*). Units: thousands of dollars.*



(a) *Percentiles fraction in stocks*

(b) *Percentiles wealth*

(c) *Percentiles withdrawals*

FIG. 9: *Scenario in Table 1 Strategy computed in synthetic market.* $p_{\max} = 0.5$. *Parameters based on the real CRSP index, and real 30-day T-bills (see Table 2). Control computed and stored using Algorithm 2 in the synthetic market.* $q_{min} = 30$, $q_{\max} = 60$ *(per year),* $EW = 50.2$, $ES = 1.25$ *(*$\kappa = 1.0$*). Units: thousands of dollars.*

10 Bootstrap results

10.1 Historical Market

In order to check on the robustness of the above results, we proceed as follows. We compute and store the optimal controls based on the parametric model (3-4) as for the synthetic market case. However, we compute statistical quantities using the stored controls, but using bootstrapped historical return data directly. In this case, we make no assumptions concerning the stochastic processes followed by the stock and bond indices. We remind the reader that all returns are inflation-adjusted. We use the stationary block bootstrap method (Politis and Romano, 1994; Politis and White, 2004; Patton et al., 2009; Cogneau and Zakalmouline, 2013; Dichtl et al., 2016; Cavaglia et al., 2022; Simonian and Martirosyan, 2022; Anarkulova et al., 2022).

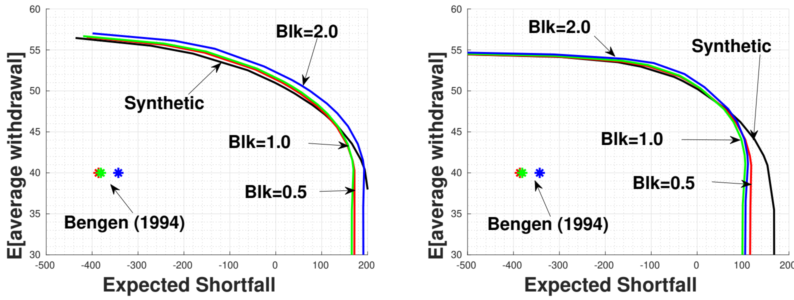
A key parameter is the expected blocksize. Sampling the data in blocks accounts for serial correlation in the data series. We use the algorithm in Patton et al. (2009) to determine the optimal blocksize for the bond and stock returns separately, see Table 5. However, in our simulations, we use a paired sampling approach to simultaneously draw returns from both time series. In this case, a reasonable estimate for the blocksize for the paired resampling algorithm would be about 2.0 years. We will give results for a range of blocksizes as a check on the robustness of the bootstrap results. Detailed pseudo-code for block bootstrap resampling is given in Forsyth and Vetzal (2019).

Optimal expected block size for bootstrap resampling historical data

Data	Optimal expected block size $\hat{b}$ (months)
30-day T-bill	51
CRSP cap weighted index	4.0

TABLE 5: *Optimal expected blocksize* $\hat{b} = 1/v$*, from Patton et al. (2009). Range of historical data is between 1926:1 and 2024:12. The blocksize is a draw from a geometric distribution with* $Pr(b = k) = (1 - v)^{k-1} v$*. Expected blocksize estimate algorithm from Politis and White (2004); Patton et al. (2009) using market data from CRSP.*

Figure 10 plots (i) the synthetic efficient frontier and (ii) bootstrap efficient frontier, computed using the synthetic market controls, tested in the historical market, for various



(a) $p_{\max} = 1.3$ (b) $p_{\max} = 0.5$

Fig. 10: *Optimal controls computed using the synthetic market model. These controls tested using bootstrapped historical data. Expected blocksizes (years) shown.* 10^6 *bootstrap resamples. Real stock index: deflated real capitalization weighted CRSP, real bond index: deflated 30 day T-bills. Scenario in Table 1. Parameters in Table 2. The Bengen control withdraws 40 per year, and rebalances annually to 50% bonds and 50% stocks. The Bengen results are also shown for expected blocksizes of* $0.5, 1.0, 2.0$ *years.*

blocksizes. Figure 10(a) shows the results for $p_{\max} = 1.3$. In this case, all the frontiers are quite close, especially compared to the Bengen (1994) strategy, tested in the historical market. The effect of using different expected blocksizes in the block bootstrap algorithm is quite small. Overall, this plot suggests that the optimal controls in this case are robust to model parameter misspecification.

The historical test for $p_{\max} = 0.5$ is shown in Figure 10(b). Again, the historical frontiers using different blocksizes are quite close. However, the historical efficient frontiers do deviate somewhat from the synthetic market frontier, for values of $ES > 100$. This indicates that there is some effect of model parameter misspecification, for large values of ES in the case of $p_{\max} = 0.5$. However, for values of $ES < 100$, the synthetic and historical market frontiers are very close. So, unless the retiree is extremely risk averse, this may not be a problem of practical concern.

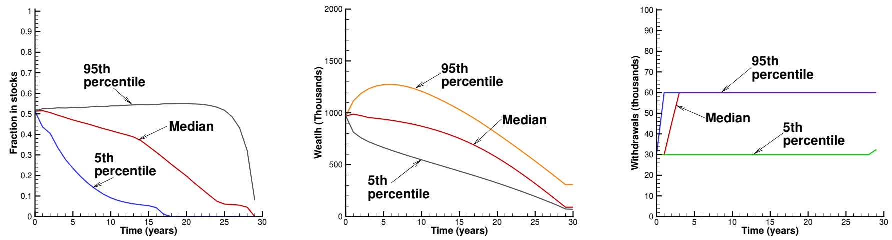
Figures 11 and 12 show the percentiles of fraction in stocks, wealth, and optimal withdrawals, tested in the historical market. The two extreme cases: $p_{max} = 1.3$ and $p_{\max} = 0.5$ are shown.

The percentiles in stocks and wealth are fairly close, for both cases. However, a comparison of Figures 11(c) and Figure 12(c) indicates that the median withdrawal increases from the minimum to the maximum over about three years for the case with $p_{max} = 1.3$. The comparable time frame for $p_{\max} = 0.5$ is about four years.

Although most retirees would prefer to larger withdrawals during the early years of retirement, slightly lower initial withdrawals in exchange for never having more than 50% stock allocation might be seen as a reasonable tradeoff.

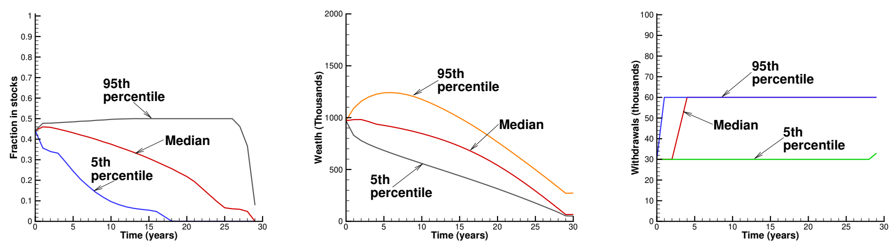
11 Conclusions

In this paper, we have developed a technique for solving the optimal control problem associated with decumulation of a defined benefit (DC) pension plan. The basic control algorithm at each rebalancing time consists of (i) solution of an optimization problem and (ii) advancing the solution to the next rebalancing time (going backwards). We use a δ-monotone scheme, based on Fourier methods, for step (ii). This method preserves order relations to $O(\delta)$, which is an important property for numerical solution of optimal control. We pay particular attention to ensuring that Fourier wraparound error can be made arbitrarily small.



(a) *Percentiles fraction in stocks*　(b) *Percentiles wealth*　(c) *Percentiles withdrawals*

Fig. 11: *Scenario in Table 1 Strategy computed in synthetic market.* $p_{\max} = 1.3$. *Tested in the bootstrapped historical market, expected blocksize 1.0 years.* 10^6 *bootstrap simulations. Bootstrap date based on the real CRSP index, and real 30-day T-bills 1926:1-2024:12. Control computed and stored using Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year),* $EW = 51.7$, $ES = 19$ *(*$\kappa = 0.8583$*). Units: thousands of dollars. Median withdrawal at* $q_{\max}$ *at year 3.0*



(a) *Percentiles fraction in stocks*　(b) *Percentiles wealth*　(c) *Percentiles withdrawals*

Fig. 12: *Scenario in Table 1 Strategy computed in synthetic market.* $p_{\max} = 0.5$. *Tested in the bootstrapped historical market, expected blocksize 1.0 years.* 10^6 *bootstrap stimulations. Bootstrap date based on the real CRSP index, and real 30-day T-bills 1926:1-2024:12. Control computed and stored using Algorithm 2 in the synthetic market.* $q_{min} = 30, q_{\max} = 60$ *(per year),* $EW = 50.3$, $ES = 6.7$ *(*$\kappa = 1.0$*). Units: thousands of dollars. Median withdrawal at* $q_{\max}$ *at year 4.0.*

From a practical point of view, we have verified that the controls are robust by testing the controls using block bootstrap resampling of historical data, which makes no assumptions about stochastic processes for the underlying stock and bond indexes.

Perhaps the most interesting result, in terms of investment strategies for retirees, is the following. If the maximum fraction in equities is restricted to be less than 50%, this strategy is only slightly less efficient than allowing a maximum fraction of 130% (that is, using leverage). This suggests that for most retirees, there is no need to undertake risky strategies during the decumulation of a DC account.

These optimal control strategies are far more efficient than the typical four per cent rule (Bengen, 1994). The optimal policy has an expected average withdrawal (real) of more than 5% of initial capital over thirty years, with approximately zero expected shortfall at age 95 (based on historical data). In contrast, the strategy suggested in (Bengen, 1994), withdraws 4% of initial capital annually, but has an expected shortfall of more than 35% of initial capital at age 95.

Finally, we note that solving the optimal control problem for decumulation based on solving PIDEs is limited to three dimensions (i.e. three assets) or less. However, problems with more assets can be solved using a machine learning approach ((Li and Forsyth, 2019; Van Staden et al., 2023; 2024; Chen et al., 2025)).

Appendices

12 Wrap Around Error: Details

To gain some insight into the wrap-around problem, we consider a highly simplified, one dimensional problem. To avoid subscript clutter, in this section, we use the notation

$$\tilde{g}(m-\ell) \equiv \tilde{g}_{m-\ell} \; ; \; u^n(m) \equiv u^n_m .$$

Using the above notation, consider the discrete time advance convolution

$$u^n(m) = \Delta x \sum_{\ell=-N^\dagger/2}^{N^\dagger/2-1} \tilde{g}(m-\ell)\, u^{n-1}(\ell) \;,$$
$$m = -N/2, \ldots, N/2-1 \;,\; N \leq N^\dagger \;, \tag{57}$$

where we assume periodic extension of $\tilde{g}$

$$\tilde{g}(\ell \pm N^\dagger) = \tilde{g}(\ell) \; .$$

In the above, if $N^\dagger > N$, nodes corresponding to $m < -N/2$ and $m > N/2-1$ are the padded nodes. In this case, we assume that the padded nodal values $u^n(m)$, $m < -N/2$ and $m > N/2-1$, are determined by boundary data.

As an example of wrap-around error, we examine a worst case term in equation (57). Consider the term in (57) corresponding to $m = -N/2$, and $l = N^\dagger/2-1$, namely

$$\Delta x \; \tilde{g}(-N/2 - N^\dagger/2 + 1)\; u^{n-1}(N^\dagger/2-1). \tag{58}$$

By periodic extension, we shift the argument of $\tilde{g}(\cdot)$ by $N^\dagger$, resulting in

$$\tilde{g}(-N/2 - N^\dagger/2 + 1) = \tilde{g}(-N/2 - N^\dagger/2 + 1 + N^\dagger) = \tilde{g}(-N/2 + N^\dagger/2 + 1), \tag{59}$$

and hence, the term (58) becomes

$$\Delta x \; \tilde{g}(-N/2 + N^\dagger/2 + 1)\; u^{n-1}(N^\dagger/2-1). \tag{60}$$

Hence, in this extreme case, equation (57) becomes

$$u^n(-N/2) \;=\; \Delta x \; \tilde{g}(-N/2 + N^\dagger/2 + 1)\; u^{n-1}(N^\dagger/2-1) \;+\; \sum_{l=-N^\dagger/2}^{N^\dagger/2-2} (\text{ remaining terms }). \tag{61}$$

Example 1 (No padding: $N^\dagger = N$) Suppose we do not use any padding, so that $N^\dagger = N$. In this case, equation (61) becomes

$$u^n(-N/2) \;=\; \Delta w \; \tilde{g}(1)\; u^{n-1}(N/2-1) \;+\; \sum_{l=-N/2}^{N/2-2} (\text{ remaining terms }).$$

Since, in general, $\tilde{g}(1)$ is not small, we can see that the term $u^{n-1}(N/2-1)$ has a considerable effect on $u^n(-N/2)$, which should not be the case. We can see here that the periodic extension of $\tilde{g}$ causes a *wrap-around* effect.

Example 2 (Padding: $N^\dagger = 2N$) If $N^\dagger = 2N$, then equation (61) becomes

$$u^n(-N/2) \;=\; \Delta w \; \tilde{g}(N/2+1)\; u^{n-1}(N^\dagger/2-1) \;+\; \sum_{l=-N^\dagger/2}^{N^\dagger/2-2} (\text{ other terms }). \qquad (62)$$

If we select N sufficiently large so that $\tilde{g}(\pm N/2) \simeq 0$, then the leading term in equation (62) is small, and hence, wrap-around error is reduced.

12.1 Wrap-around: a formal result

From our previous examples, we can see that wrap-around may occur in equation (57) if

$$(m-\ell) < -N^\dagger/2 \;\text{ or }\; (m-\ell) > N^\dagger/2 - 1.$$

This leads us to the following formal definition of wrap-around error. We show that, with $N^\dagger = 2N$, wrap-around error is sufficiently reduced.

Definition 2 (Wrap-around error) Assume $\{\tilde{g}(\cdot)\}$ is periodic with period $N^\dagger$ and $u^n(m)$, for $m < -N/2$ or $m > N/2-1$, are determined by boundary data with $N^\dagger = 2N$. Then the wrap-around error for equation (57), at timestep n, denoted by e^n_{wrap}, is

$$e^n_{\text{wrap}} = \max_m \left\{ \Delta w \sum_{\ell \in \mathbb{N}^\dagger} \left| \tilde{g}(m-\ell)\; u^{n-1}(l) \right| \left(\mathbf{1}_{\{(m-\ell)<-N^\dagger/2\}} + \mathbf{1}_{\{(m-\ell)>N^\dagger/2-1\}} \right) \right\}. \qquad (63)$$

We can now state the following result.

Theorem 2 *Let $\tilde{g}(\cdot)$ be periodic with period $N^\dagger$ and $u^n(m)$, for $m < -N/2$ or $m > N/2-1$, be determined by boundary data with $N^\dagger = 2N$. Assume further that $u^n(m)$ is bounded, so for $0 \le n \le M$, there exists a positive constant C such that*[12]

$$|u^n(m)| \le C, \quad m \in \mathbb{N}^\dagger \;, \forall n \,. \qquad (64)$$

If N is selected sufficiently large so that

$$\Delta x \sum_{\ell=-N^\dagger/2}^{-N/2-1} |\tilde{g}(\ell)| + \Delta x \sum_{\ell=N/2}^{N^\dagger/2-1} |\tilde{g}(\ell)| \;\le\; \epsilon_e \Delta\tau \qquad (65)$$

then the wrap-around error after N steps is bounded by $TC\epsilon_e$.

Proof. Applying property (64) to equation (63) gives

$$e^n_{\text{wrap}} \le C \max_m \left\{ \Delta x \sum_{\ell=-N^\dagger/2}^{N^\dagger/2-1} |\tilde{g}(m-\ell)| \left(\mathbf{1}_{\{(m-\ell)<-N^\dagger/2\}} + \mathbf{1}_{\{(m-\ell)>N^\dagger/2-1\}} \right) \right\}. \qquad (66)$$

Recall that $m \in \{-N/2, \ldots, N/2-1\}$, hence the worst case values of m on the right hand side of equation (66) are $m = -N/2$ and $m = N/2-1$. Thus equation (66) gives

[12] This is essentially a stability condition. See Forsyth and Labahn (2019) for a proof of stability for the δ-monotone method.

$$
\begin{aligned}
e^n_{\text{wrap}} \le C\Delta x &\sum_{\ell=-N^\dagger/2}^{N^\dagger/2-1} |\tilde{g}(N/2-1-\ell)| \mathbf{1}_{\{(N/2-1-\ell) > N^\dagger/2-1\}} \\
&+ C\Delta x \sum_{\ell=-N^\dagger/2}^{N^\dagger/2-1} |\tilde{g}(-N/2-\ell)|\, \mathbf{1}_{\{(-N/2-\ell) < -N^\dagger/2\}}.
\end{aligned} \tag{67}
$$

Also, since $N = N^\dagger/2$ equation (67) becomes

$$
\begin{aligned}
e^n_{\text{wrap}} \le C\Delta x &\sum_{\ell=-N^\dagger/2}^{N^\dagger/2-1} |\tilde{g}(N^\dagger/4-1-\ell)|\, \mathbf{1}_{\{(N^\dagger/4-1-\ell) > N^\dagger/2-1\}} \\
&+ C\Delta x \sum_{\ell=-N^\dagger/2}^{N^\dagger/2-1} |\tilde{g}(-N^\dagger/4-l)|\, \mathbf{1}_{\{(-N^\dagger/4-\ell) < -N^\dagger/2\}},
\end{aligned} \tag{68}
$$

and eliminating the indicator functions gives

$$
e^n_{\text{wrap}} \le C\Delta x \sum_{\ell=-N^\dagger/2}^{-N^\dagger/4-1} |\tilde{g}(N^\dagger/4-1-\ell)| + C\Delta x \sum_{\ell=N^\dagger/4+1}^{N^\dagger/2-1} |\tilde{g}(-N^\dagger/4-\ell)|.
$$

Shifting $\tilde{g}(\cdot)$ by $\pm N^\dagger$ so that the argument of $\tilde{g}(\cdot)$ is in the range $[-N^\dagger/2, N^\dagger/2-1]$, implies

$$
\begin{aligned}
e^n_{\text{wrap}} &\le C\Delta x \sum_{\ell=-N^\dagger/2}^{-N^\dagger/4-1} |\tilde{g}(N^\dagger/4-1-\ell-N^\dagger)| + C\Delta x \sum_{\ell=N^\dagger/4+1}^{N^\dagger/2-1} |\tilde{g}(-N^\dagger/4-\ell+N^\dagger)| \\
&= C\Delta x \sum_{\ell=-N^\dagger/2}^{-N^\dagger/4-1} |\tilde{g}(-3N^\dagger/4-1-\ell)| + C\Delta x \sum_{\ell=N^\dagger/4+1}^{N^\dagger/2-1} |\tilde{g}(3N^\dagger/4-\ell)|.
\end{aligned} \tag{69}
$$

Rearranging the indices, gives

$$
e^n_{\text{wrap}} \le C\Delta x \sum_{\ell=-N^\dagger/2}^{-N^\dagger/4-1} |\tilde{g}(\ell)| + C\Delta x \sum_{\ell=N^\dagger/4+1}^{N^\dagger/2-1} |\tilde{g}(\ell)|, \tag{70}
$$

which, since $N = N^\dagger/2$, implies that equation (70) satisfies

$$
e^n_{\text{wrap}} \le C\Delta x \sum_{\ell=-N^\dagger/2}^{-N/2-1} |\tilde{g}(\ell)| + C\Delta x \sum_{\ell=N/2+1}^{N^\dagger/2-1} |\tilde{g}(\ell)|. \tag{71}
$$

Since

$$
\Delta x \sum_{\ell=N/2+1}^{N^\dagger/2-1} |\tilde{g}(\ell)| \le \Delta x \sum_{\ell=N/2}^{N^\dagger/2-1} |\tilde{g}(\ell)| \tag{72}
$$

then

$$
\begin{aligned}
e^n_{\text{wrap}} &\le C\Delta x \sum_{\ell=-N^\dagger/2}^{-N/2-1} |\tilde{g}(\ell)| + C\Delta x \sum_{\ell=N/2}^{N^\dagger/2-1} |\tilde{g}(\ell)| && (73) \\
&= C\epsilon_e \Delta\tau && (74)
\end{aligned}
$$

where the last step follows from condition (65). Applying equation (74) recursively gives the bound $TC\epsilon_e$. □

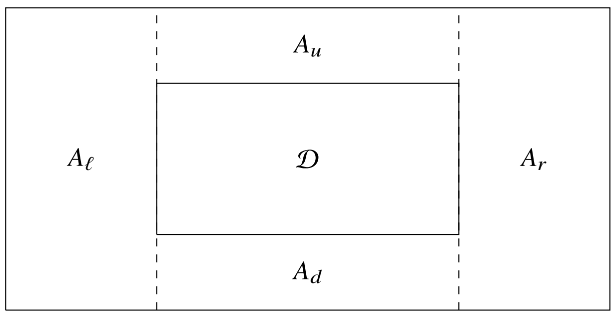
It might at first sight seem odd to weaken the error bound using equation (72). However, this makes the final result easy to interpret. Let

$$\mathcal{S}^\dagger = \{-N^\dagger/2, \ldots, +N^\dagger/2 - 1\} \; ; \; \mathcal{S} = \{-N/2, \ldots, +N/2 - 1\} \tag{75}$$

so that $\mathcal{S}^\dagger$ is the set of node indexes for the padded domain, and $\mathcal{S}$ is the set of node indices for the original domain. Consequently, condition (73) can be written compactly as

$$e^n_{\text{wrap}} \leq C\left(\Delta x \sum_{\ell \in (\mathcal{S}^\dagger - \mathcal{S})} |\tilde{g}(\ell)|\right) \;\leq\; C\epsilon_e \Delta\tau \; . \tag{76}$$

Proving Theorem 1 basically follows the previous proof of the simpler case. Divide the region $\mathcal{D}' - \mathcal{D}$ outside $\mathcal{D}$ into right and left A_r, A_ℓ and upper and lower A_u, A_d regions as follows.



Our goal is to give an upper bound to

$$\varepsilon^n_{wrap} = \left(\Delta x_1 \Delta x_2 \max_{(\ell,m)\in\mathcal{D}} \sum_{(j,k)\in\mathcal{D}^\dagger} \left|\tilde{g}^\dagger_{\ell-j,m-k} v^\dagger_{j,k}(\tau^n)\right| \mathbf{1}_{(\ell-j,m-k)\notin\mathcal{D}^\dagger}\right)$$

which by the boundedness assumptions of Theorem 1 is the same as bounding

$$\left(\Delta x_1 \Delta x_2 \max_{(\ell,m)\in\mathcal{D}} \sum_{(j,k)\in\mathcal{D}^\dagger} \left|\tilde{g}^\dagger_{\ell-j,m-k}\right| \mathbf{1}_{(\ell-j,m-k)\notin\mathcal{D}^\dagger}\right) .$$

As in the 1d wraparound error, our worse case values occur at borders, in this case the four corners of $\mathcal{D}$. In the horizontal direction when the first components are $-N_1/2$ and $N_1/2 - 1$ we can use the same manipulations as done in the 1d case to obtain the wraparound error contribution from A_ℓ and A_r and use an identical argument when the second argument is $-N_2/2$ and $N_2/2 - 1$ to get the wraparound error contribution from A_u and A_d. These manipulations reduce to

$$e^n_{\text{wrap}} \leq C\left(\Delta x_1 \Delta x_2 \sum_{(\ell,m)\in(\mathcal{D}^\dagger - \mathcal{D})} |\tilde{g}(\ell, m)|\right) \;\leq\; C\epsilon_e \Delta\tau$$

for each time step and hence give the error bound $e^{\mathcal{N}}_{\text{wrap}} \leq T\epsilon_e C$ after $\mathcal{N}$ steps.

References

Alonso-Garcca, J., O. Wood, and J. Ziveyi (2018). Pricing and hedging guaranteed minimum withdrawal benefits under a general Lévy framework using the COS method. *Quantitative Finance 18*, 1049–1075.

Anarkulova, A., S. Cederburg, and M. S. O'Doherty (2022). Stocks for the long run? Evidence from a broad sample of developed markets. *Journal of Financial Economics 143:1*, 409–433.

Anarkulova, A., S. Cederburg, M. S. O'Doherty, and R. W. Sias (2025). The safe withdrawal rate: Evidence from a broad sample of developed markets. *Journal of Pension Economics and Finance*. To appear.

Barles, G. and P. Souganidis (1991). Convergence of approximation schemes for fully nonlinear equations. *Asymptotic Analysis 4*, 271–283.

Bengen, W. (1994). Determining withdrawal rates using historical data. *Journal of Financial Planning 7*, 171–180.

Bjork, T., M. Khapko, and A. Murgoci (2021). *Time inconsistent control theory with finance applications*. New York: Springer Finance.

Cavaglia, S., L. Scott, K. Blay, and S. Hixon (2022). Multi-asset class factor premia: A strategic asset allocation perspective. *The Journal of Portfolio Management 48:9*, 14–32.

Chen, M., M. Shirazi, P. A. Forsyth, and Y. Li (2025). Machine learning and Hamilton-Jacobi-Bellman equation for optimal decumulation: a comparison study. *Journal of Computational Finance 29:1*, 77–118.

Cogneau, P. and V. Zakalmouline (2013). Block bootstrap methods and the choice of stocks for the long run. *Quantitative Finance 13*, 1443–1457.

Cont, R. and C. Mancini (2011). Nonparametric tests for pathwise properties of semimartingales. *Bernoulli 17*, 781–813.

Cui, X., D. Li, X. Qiao, and M. Strub (2022). Risk and potential: An asset allocation framework with applications to robo-advising. *Journal of the Operations Research Society of China 10*, 529–558.

Dang, D.-M. and P. A. Forsyth (2016). Better than pre-commitment mean-variance portfolio allocation strategies: a semi-self-financing Hamilton-Jacobi-Bellman equation approach. *European Journal of Operational Research 250*, 827–841.

Dichtl, H., W. Drobetz, and M. Wambach (2016). Testing rebalancing strategies for stock-bond portfolios across different asset allocations. *Applied Economics 48*, 772–788.

Fang, F. and C. Oosterlee (2008). A novel pricing method for European options based on Fourier cosine series expansions. *SIAM Journal on Scientific Computing 31*, 826–848.

Fang, F. and C. Oosterlee (2009). Pricing early-exercise and discrete barrier options by Fourier cosine series expansions. *Numerische Mathematik 114*, 27–62.

Forsyth, P. and G. Labahn (2019). ϵ-Monotone Fourier methods for optimal stochastic control in finance. *Journal of Computational Finance 22:4*, 25–71.

Forsyth, P. A. (2020a). Multi-period mean CVAR asset allocation: Is it advantageous to be time consistent? *SIAM Journal on Financial Mathematics 11:2*, 358–384.

Forsyth, P. A. (2020b). Optimal dynamic asset allocation for DC plan accumulation/decumulation: Ambition-CVAR. *Insurance: Mathematics and Economics 93*, 230–245.

Forsyth, P. A. (2022). A stochastic control approach to defined contribution plan decumulation: "The Nastiest, Hardest Problem in Finance". *North American Actuarial Journal 26:2*, 227–252.

Forsyth, P. A. and K. R. Vetzal (2019). Optimal asset allocation for retirement savings: deterministic vs. time consistent adaptive strategies. *Applied Mathematical Finance 26:1*, 1–37.

Forsyth, P. A., K. R. Vetzal, and G. Westmacott (2020). Optimal asset allocation for a DC pension decumulation with a variable spending rule. *ASTIN Bulletin 50:2*, 419–447.

Forsyth, P. A., K. R. Vetzal, and G. Westmacott (2022). Optimal control of the decumulation of a retirement portfolio with variable spending and dynamic asset allocation. *ASTIN Bulletin 51:3*, 905–938.

Garroni, M. G. and J. L. Menaldi (1992). *Green functions for second order parabolic integro-differential problems*. New York: Longman Scientific.

Ignatieva, K., A. Song, and J. Ziveyi (2018). Fourier space timestepping algorithm for valuing guaranteed minimum withdrawal benefits in variable annuities under regime switching and stochastic mortality. *ASTIN Bulletin 48*, 139–169.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science 48*, 1086–1101.

Kou, S. G. and H. Wang (2004). Option pricing under a double exponential jump diffusion model. *Management Science 50*, 1178–1192.

Li, Y. and P. Forsyth (2019). A data-driven neural network approach to optimal asset allocation for target based defined contribution pension plans. *Insurance: Mathematics and Economics 86*, 189–204.

Lin, Y., R. MacMinn, and R. Tian (2015). De-risking defined benefit plans. *Insurance: Mathematics and Economics 63*, 52–65.

Lippa, J. (2013). A fourier space time-stepping approach to problems in finance. MMath thesis, University of Waterloo.

MacDonald, B.-J., B. Jones, R. J. Morrison, R. L. Brown, and M. Hardy (2013). Research and reality: A literature review on drawing down retirement financial savings. *North American Actuarial Journal 17*, 181–215.

MacMinn, R., P. Brockett, J. Wang, Y. Lin, and R. Tian (2014). The securitization of longevity risk and its implications for retirement security. In O. S. Mitchell, R. Maurer, and P. B. Hammond (Eds.), *Recreating Sustainable Retirement*, pp. 134–160. Oxford: Oxford University Press.

Mancini, C. (2009). Non-parametric threshold estimation models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*, 270–296.

Menoncin, F. and E. Vigna (2017). Mean-variance target based optimisation for defined contribution pension schemes in a stochastic framework. *Insurance: Mathematics and Economics 76*, 172–184.

Patton, A., D. Politis, and H. White (2009). Correction to: automatic block-length selection for the dependent bootstrap. *Econometric Reviews 28*, 372–375.

Peijnenburg, K., T. Nijman, and B. J. Werker (2016). The annuity puzzle remains a puzzle. *Journal of Economic Dynamics and Control 70*, 18–35.

Pfeiffer, S., J. R. Salter, and H. E. Evensky (2013). Increasing the sustainable withdrawal rate using the standby reverse mortgage. *Journal of Financial Planning 26:12*, 55–62.

Politis, D. and J. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*, 1303–1313.

Politis, D. and H. White (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews 23*, 53–70.

Ritholz, B. (2017). Tackling the 'nastiest, hardest problem in finance'. `www.bloomberg.com/view/articles/2017-06-05/tackling-the-nastiest-hardest-problem-in-finance`.

Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*, 21–42.

Ruijter, M. J., C. W. Oosterlee, and R. Albers (2013). On the Fourier cosine series expansion method for stochastic control problems. *Numerical Linear Algebra and Applications 20*, 598–625.

Shefrin, H. M. and R. H. Thaler (1988). The behavioral life-cycle hypothesis. *Economic Inquiry 26*, 609–643.

Simonian, J. and A. Martirosyan (2022). Sharpe parity redux. *The Journal of Portfolio Management 48:9*, 183–193.

Strub, M., D. Li, X. Cui, and J. Gao (2019). Discrete-time mean-CVaR portfolio selection and time-consistency induced term structure of the CVaR. *Journal of Economic Dynamics and Control 108*. Article 103751 (electronic).

Thinking Ahead Institute (2024). Global pension assets study 2024. https://www.thinkingaheadinstitute.org/content/uploads/2024/02/GPAS-2024.pdf.

van Staden, P. M., D.-M. Dang, and P. Forsyth (2018). Time-consistent mean-variance portfolio optimization: a numerical impulse control approach. *Insurance: Mathematics and Economics 83*, 9–28.

Van Staden, P. M., P. A. Forsyth, and Y. Li (2023). Beating a benchmark: dynamic programming may not be the right numerical approach. *SIAM Journal on Financial Mathematics 14*(2).

Van Staden, P. M., P. A. Forsyth, and Y. Li (2024). A global-in-time neural network approach to dynamic portfolio optimization. *Applied Mathematical Finance 31:3*, 131–163.

Vigna, E. (2014). On efficiency of mean-variance based portfolio selection in defined contribution pension schemes. *Quantitative Finance 14*, 237–258.

Vigna, E. (2017). Tail optimality and preferences consistency for intertemporal optimization problems. Working paper no. 502 , Collegio Carlo Alberto, Università Degli Studi di Torino.