AIMS Mathematics, 10(12): 29784–29814.
DOI: 10.3934/math.20251309
Received: 24 October 2025
Revised: 09 December 2025
Accepted: 12 December 2025
Published: 17 December 2025

https://www.aimspress.com/journal/Math

Research article

Deterministic value iteration for perpetual American put options

Eungpyo Kim and Jaegi Jeon*

Graduate School of Data Science, Chonnam National University, Gwangju 61186, Republic of Korea

* **Correspondence:** Email: jaegijeon@jnu.ac.kr; Tel: +82-62-530-5795.

Abstract: We introduce a deterministic, policy-targeted Bellman value-iteration framework for computing the optimal exercise boundary of perpetual American put options. Our method replaces path sampling in the Bellman operator with Gauss–Hermite quadrature and employs shape-preserving interpolation for off-grid evaluations, eliminating sampling noise and reducing computational cost. Under the Black–Scholes (BS) model, our approach recovers the analytic boundary with a mean absolute percentage error below 1.5% in approximately 19–56 seconds. The resulting policy values, evaluated via Monte Carlo simulation, deviate from the analytic benchmark by less than 0.07%. For the Heston model, where no closed-form solution exists, our method produces boundaries that differ from a high-resolution finite-difference benchmark by 1–5%. Despite these boundary deviations, the expected payoffs from the policies are remarkably close, with relative policy value gaps well below 0.2%. Notably, our method computes the boundary in about 127–180 seconds, a significant speedup compared to the 2,103–3,119 seconds required by the finite-difference method. This work presents a practical and robust alternative for optimal stopping problems, offering a compelling balance of speed and accuracy, particularly when partial differential equation (PDE) solvers are cumbersome or Monte Carlo simulation is prohibitively expensive.

Keywords: American options; perpetual option; optimal stopping; exercise boundary; value iteration; Gauss–Hermite quadrature; Black–Scholes model; Heston model
Mathematics Subject Classification: Primary 91G20, 60G40; Secondary 65D30, 65N06, 90C33, 35R35

1. Introduction

Derivatives, and options in particular, are essential tools for risk management and strategic decision-making. American-style options are especially important in practice due to their flexibility, which allows them to be exercised at any time before expiration. This early-exercise feature creates significant valuation challenges compared with their European counterparts, a challenge central to option pricing

theory [1–5]. A particularly informative benchmark in this context is the perpetual American put, whose time-homogeneous structure simplifies the optimal stopping problem to a stationary one. In many applications, however, practitioners care not only about the option's price but also about a stable and accurate estimate of the exercise boundary, as this boundary dictates the implementation of the exercise-versus-continuation decision rule. While finite-maturity American options are of primary interest in practice, we focus here on the perpetual case as a time-homogeneous benchmark that isolates the core policy-targeted iteration; a finite-horizon extension of the framework, based on a backward Gauss–Hermite quadrature value-iteration (GHQ–VI) recursion, is briefly discussed later. This paper develops a simple, deterministic, and accurate procedure to estimate the perpetual exercise boundary under both the Black–Scholes (BS) [1] and the Heston stochastic volatility [6] dynamics.

Existing literature predominantly adopts a price-first approach: Estimate the value function as accurately as possible and obtain the boundary as a by-product. Finite-difference solvers for the associated free-boundary problem, which often takes the form of a linear complementarity problem (LCP) solved with methods like projected successive over-relaxation (PSOR), can be accurate but are computationally demanding and sensitive to discretization choices in the 2D Heston settings [3, 7]. Simulation-based techniques, most notably the least-squares Monte Carlo (LSM) method, are flexible but inherit sampling noise and basis sensitivity [8–11]. Duality methods provide tight price bounds yet do not directly stabilize the policy itself [12, 13]. More recently, deep learning methods based on backward stochastic differential equations have broadened the high-dimensional toolkit, but this line of work remains value-centric rather than boundary-centric [14]. In practice, these choices can yield option prices with small errors while the implied boundary—especially near the kink induced by the max operator—remains comparatively volatile.

Closer to the present work, there are "boundary-first" finite-horizon approaches in which the early-exercise boundary (or an exercise region) is discretized on a time grid and then used as an input to a partial differential equation (PDE) or finite-difference solver [15, 16]. Related work develops a near-optimality framework for American options and shows that suitably defined families of exercise policies can produce option values that remain extremely close to the optimum [17]. Our contribution is complementary: We focus on the perpetual case and develop a quadrature-based value-iteration method that targets the stationary exercise boundary directly. In our numerical study, Monte Carlo policy evaluation shows that option values remain very stable even under moderate boundary perturbations, echoing the valuation robustness highlighted in [17].

To address the limitations of standard price-first approaches, we propose a policy-targeted value iteration procedure, hereafter GHQ–VI, which employs a deterministic Bellman update combining Gauss–Hermite quadrature (GHQ) with shape-preserving piecewise cubic Hermite interpolation (PCHIP). The expectation in the Bellman operator is evaluated without sampling via GHQ (cf. direct-integration GHQ schemes for option valuation in [18]; see also [19, 20]), while off-grid evaluations use PCHIP to suppress spurious oscillations near the exercise kink (in contrast to the cubic spline used in [18]). The method extends from the BS model (1D) to the Heston model (2D) by tensor-product GHQ with correlation injection and a simple nonnegativity projection for variance. Although we update the continuation value, we treat the boundary as a first-class output: At every iteration, we extract and monitor the exercise threshold and stop when the boundary stabilizes. In this sense, our approach remains value-based but is explicitly policy-targeted.

It is well understood in optimal control and optimal stopping that policy identification is typically

harder than value estimation [21]*. This difficulty arises because the value and payoff curves meet tangentially at the free boundary (a result of value matching and smooth pasting). Consequently, even a tiny vertical error ε in the value function can translate into a disproportionately larger horizontal shift $|\delta S|$ in the boundary, often on the order of $O(\sqrt{\varepsilon})$ [4, 23]. These facts motivate a design that reduces variance deterministically (GHQ for the expectation), controls interpolation artifacts precisely where the kink bites (shape-preserving PCHIP), and evaluates progress in boundary space (boundary-stability stopping rule) rather than only in value space.

We validate GHQ–VI under both the BS and Heston models. For the BS model, the ground-truth perpetual boundary is analytic; for the Heston model, we construct a high-resolution benchmark using the finite-difference method (FDM) to solve the stationary obstacle problem, which is formulated as an LCP and solved with a red-black PSOR (RB-PSOR) scheme (see Appendix A). Our protocol reports (i) boundary accuracy, measured by the mean absolute percentage error ($\mathrm{MAPE}_{\mathrm{bnd}}$); (ii) the realized policy value gap, measured as the relative difference $\hat{\Delta}_N/\hat{V}_{\mathrm{bench}}$ based on 10^5 common-random-number Monte Carlo paths, where each boundary is used as the exercise policy; and (iii) runtime to compute the boundary. This setup explicitly acknowledges that policy estimates can fluctuate more than prices, while realized value can remain robust (details in Section 5).

In the BS model, GHQ–VI recovers the analytic boundary with $\mathrm{MAPE}_{\mathrm{bnd}} \le 1.49\%$ across the parameter sets, while the relative policy value gaps $\hat{\Delta}_N/\hat{V}_{\mathrm{bench}}$ do not exceed about 0.07%; the runtime is roughly 19–56 s per configuration. In the Heston model, the GHQ–VI boundary preserves the expected monotone decline in v and shows relative deviations typically in the 1–5% range (with the largest deviations concentrated at lower r and higher long-run variance θ), yet the relative policy value gaps stay under roughly 0.2%. Computing the boundary takes about 127–180 s with GHQ–VI versus 2,103–3,119 s for the FDM-LCP benchmark on our setups. These patterns are consistent with the design rationale above: Boundaries are harder to stabilize than values, but realized performance is robust.

The main contributions of this paper can be summarized as follows:

- **Policy-targeted value iteration.** We introduce a deterministic Bellman update (using GHQ) combined with shape-preserving interpolation implemented via PCHIP, where the exercise boundary is treated as a primary output and monitored for convergence. The approach applies to the BS and Heston models with minimal model-specific machinery [1, 6].

- **Variance-free expectation in place of sampling.** Replacing Monte Carlo by GHQ removes sampling noise and yields stable iteration; prior evidence supports GHQ's speed/accuracy in option contexts [18–20].

- **Practical accuracy and efficiency.** On both models, the learned boundaries are accurate (with small relative boundary errors), and the associated relative policy value gaps remain numerically negligible. At the same time, GHQ–VI delivers more than an order-of-magnitude speedup in boundary computation compared with a high-resolution FDM-LCP benchmark.

- **Complementary to price-first methods.** Our procedure can serve as a computationally efficient boundary extractor when closed forms are unavailable, PDE solvers are cumbersome, or Monte Carlo is expensive [7–9, 12, 14].

*For seminal works on dynamic programming and optimal stopping, see, e.g., [5, 22].

The remainder of this paper is organized as follows: Section 2 presents the stationary Bellman formulation for perpetual puts. Section 3 details the proposed GHQ–VI procedure. Section 4 establishes convergence and error guarantees for the algorithm. Section 5 reports boundary and policy-value accuracy, along with runtime comparisons against benchmarks. Section 6 concludes.

2. Problem setup and stationary Bellman formulation

2.1. American options and optimal stopping strategy

American options may be exercised at any time prior to expiry, which makes their valuation significantly more complex than that of European options. This added flexibility naturally leads to an optimal stopping problem in which the holder must determine the optimal exercise time τ^* that maximizes the expected payoff.

The value of an American put option is fundamentally an optimal stopping problem [2, 23]. Its value is captured by a value function, $V(t, S)$, which is defined as the maximum expected payoff over all permissible exercise times τ for a given underlying price process $\{S_t\}_{t\geq 0}$, a strike price K, and a risk-free interest rate r

$$V(t, S) = \sup_{t\leq\tau\leq T} \mathbb{E}\left[e^{-r(\tau-t)}(K - S_\tau)^+ \,\middle|\, S_t = S\right],$$

where the term $(K - S_\tau)^+ = \max\{K - S_\tau, 0\}$ is the intrinsic value or immediate payoff from exercise at time τ. Hereafter, we denote the payoff function by $\Phi(S) = (K - S)^+$. In this formulation, the main difficulty lies in determining the optimal stopping time $\tau^*(t, S)$, or equivalently, the optimal exercise boundary $S^*(t)$, that separates the exercise and continuation regions.

In this setting, the continuation value at time t is defined as

$$C(t, S) = e^{-r\Delta t}\mathbb{E}\left[V(t + \Delta t, S_{t+\Delta t}) \mid S_t = S\right].$$

The value function V then satisfies the Bellman recursion

$$V(t, S) = \max\left\{\Phi(S),\ C(t, S)\right\}. \tag{2.1}$$

Accordingly, an optimal stopping time $\tau^*(t, S)$ can be characterized as the first time at which immediate exercise dominates continuation

$$\tau^*(t, S) = \inf\left\{\tau \in [t, T] : \Phi(S_\tau) \geq C(\tau, S_\tau)\right\}. \tag{2.2}$$

This condition implicitly defines the optimal exercise boundary, denoted $S^*(t)$, where the value of holding the option equals the value of exercising it.

The optimal stopping strategy is fully characterized by the optimal exercise boundary $S^*(t)$, which partitions the (t, S)-plane into regions for continuation and immediate exercise, as illustrated in Figure 1. Since this boundary is governed by the continuation value, the central computational task is to accurately and efficiently approximate $C(t, S)$ across the state space.

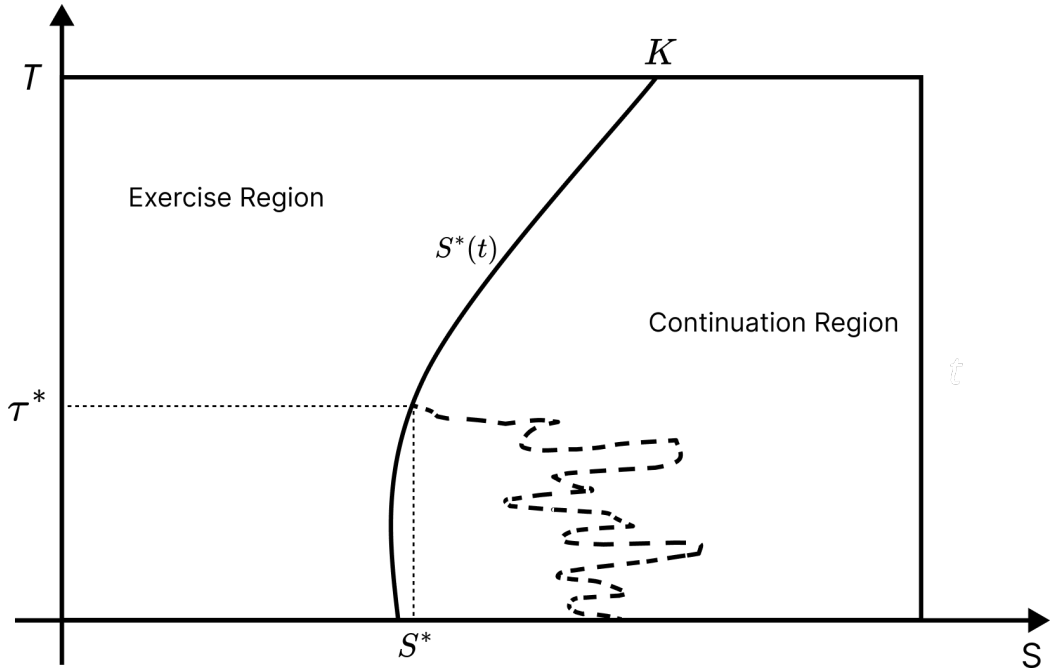

Figure 1. Optimal exercise boundary for a finite-maturity American put option. The boundary $S^*(t)$ divides the (t, S)-plane into an exercise region and a continuation region.

2.2. Perpetual reduction: Stationary Bellman

In the perpetual case, where the option has no expiry, time-homogeneity implies that the value and continuation functions depend only on the current state of the underlying asset.

Specifically, the continuation function C satisfies the stationary Bellman equation, a cornerstone of discounted dynamic programming [24, 25]

$$C^{(k+1)}(S) = e^{-r\Delta t}\mathbb{E}\Big[\max\{\Phi(S_{t+\Delta t}), C^{(k)}(S_{t+\Delta t})\} \mid S_t = S\Big], \tag{2.3}$$

where Δt is a discretization step in time and $\{C^{(k)}\}$ denotes the sequence of continuation value approximations at iteration k. The recursion thus defines an iterative scheme converging to the fixed point $C^*(S)$, the true continuation value.

In the stationary setting, the value is characterized by the relation (2.1)

$$V^*(S) = \max\{\Phi(S),\ C^*(S)\}.$$

The optimal exercise boundary S^* is characterized by the indifference condition

$$\Phi(S^*) = C^*(S^*). \tag{2.4}$$

Accordingly, the state space decomposes into the exercise set $\{S < S^*\}$, where immediate exercise is optimal, and the continuation set $\{S > S^*\}$, where holding is optimal. In the perpetual case, the free boundary S^* is time-invariant and therefore a constant threshold.

These stationary identities are model-agnostic: They depend only on the risk-neutral transition law of the state inside the conditional expectation in (2.3). In the next subsection, we specify the transition law for the BS (1D) and Heston (2D) models, which differ only in the form of the state evolution appearing in (2.3).

2.3. Models

To instantiate the stationary Bellman relation in Section 2.2, we now specify the risk-neutral dynamics of the underlying asset, given by stochastic differential equations (SDEs), in the two canonical option-pricing models, BS [1] and Heston [6]. The one-step transition over Δt that appears in (2.3) is induced by these dynamics and will be discretized deterministically in Section 3.

2.3.1. Black–Scholes model

Under the risk-neutral measure, the asset price S_t follows the SDE (geometric Brownian motion with constant volatility)

$$dS_t = rS_t\,dt + \sigma S_t\,dW_t, \tag{2.5}$$

where $r > 0$ is the risk-free rate, $\sigma > 0$ is the constant volatility, and W_t is a standard Brownian motion.

2.3.2. Heston model

To capture stochastic volatility, we augment the state with the instantaneous variance v_t and assume the Heston dynamics

$$dS_t = rS_t\,dt + \sqrt{v_t}\,S_t\,dW_t^S, \tag{2.6}$$

$$dv_t = \kappa(\theta - v_t)\,dt + \xi\sqrt{v_t}\,dW_t^v, \tag{2.7}$$

with correlation $dW_t^S\,dW_t^v = \rho\,dt$. Here, $\kappa > 0$ is the mean-reversion rate, $\theta > 0$ is the long-run variance level, $\xi > 0$ is the volatility of volatility, and $\rho \in (-1, 1)$ is the instantaneous correlation. The variance process (2.7) is a mean-reverting square-root diffusion, specifically the Cox–Ingersoll–Ross (CIR) process.[†]

2.4. Benchmarks used for validation

In the BS model, the perpetual American put admits a well-known closed-form solution for its constant exercise boundary [27], which serves as our analytic benchmark:

$$S^* = \frac{\beta}{\beta - 1}\,K,$$

where

$$\beta = \left(\frac{1}{2} - \frac{r}{\sigma^2}\right) - \sqrt{\left(\frac{1}{2} - \frac{r}{\sigma^2}\right)^2 + \frac{2r}{\sigma^2}},$$

which we use as an analytic reference.

Unlike in the BS model, no closed-form solution is known to exist for the perpetual American put under the Heston dynamics. As a numerical benchmark, we discretize the stationary variational inequality (obstacle problem) by FDM; the resulting discrete system takes the form of an LCP, which we solve with a RB-PSOR scheme. Full stencil coefficients, boundary conditions, and solver parameters are deferred to Appendix A; in Section 5, we compare the learned boundaries and prices against this FDM-LCP reference.

[†]Under the Feller condition $2\kappa\theta \geq \xi^2$, the CIR process (2.7) stays strictly positive [26].

3. Method

3.1. Gauss–Hermite quadrature

To evaluate the expectation in (2.3) deterministically, we use GHQ, a classical and highly accurate numerical integration technique [28]. GHQ replaces pathwise sampling with a weighted sum under the standard normal distribution. Unlike Monte Carlo, whose statistical error decays at $O(M^{-1/2})$, GHQ achieves high accuracy with a small, fixed number of nodes.

For a function $g(Z)$ with $Z \sim \mathcal{N}(0,1)$,

$$\mathbb{E}[g(Z)] = \int_{-\infty}^{\infty} g(z)\varphi(z)\,dz \approx \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} w_i\, g(\sqrt{2}\,x_i). \tag{3.1}$$

Here, $\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\{-z^2/2\}$ is the standard normal density, and (x_i, w_i) are the Gauss–Hermite nodes and weights.[‡] Instead of integrating over the entire real line, GHQ replaces the infinite integral by a weighted sum of function evaluations at carefully chosen points, providing a simple yet accurate numerical approximation. Figure 2 illustrates this idea: The expectation under the standard normal density $\varphi(z)$ is approximated by a weighted sum of function evaluations at the Gauss–Hermite nodes x_i with weights w_i.

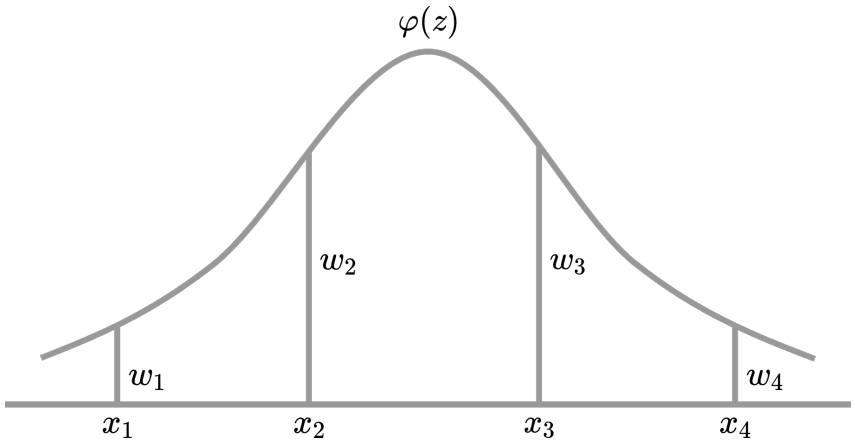

Figure 2. Conceptual illustration of a Bellman update using four-node GHQ.

3.1.1. Bellman update via 1D GHQ

In our framework, the GHQ rule is applied directly to the Bellman operator for the continuation value (2.3). Instead of sampling next-step paths, we evaluate the integrand at deterministically transformed future prices $S^{(i)}_{t+\Delta t}$ produced by the model's one-step transition and accumulate a weighted sum

$$C^{(k+1)}(S_t) \approx e^{-r\Delta t} \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} w_i \max\left\{\Phi(S^{(i)}_{t+\Delta t}), C^{(k)}(S^{(i)}_{t+\Delta t})\right\}. \tag{3.2}$$

[‡]We follow the physicists' convention for Hermite polynomials with weight e^{-x^2}. The mapping to the standard normal uses $z = \sqrt{2}\,x$ and introduces the factor $1/\sqrt{\pi}$, yielding $\mathbb{E}[g(Z)] \approx \frac{1}{\sqrt{\pi}} \sum_i w_i\, g(\sqrt{2}\,x_i)$.

3.1.2. Bellman update via 2D GHQ

In the Heston model, the Markov state is (S_t, v_t); hence, the stationary continuation value is $C^*(S, v)$, and the GHQ-based Bellman update operates in two dimensions as follows. For two Brownian increments (Z_1, Z_2) with correlation ρ, we form a product rule with independent GHQ nodes (x_i, x_j) and inject correlation by the linear map

$$Z_1 = \sqrt{2}\, x_i, \quad Z_2 = \sqrt{2}\left(\rho\, x_i + \sqrt{1-\rho^2}\, x_j\right). \tag{3.3}$$

Then

$$C^{(k+1)}(S_t, v_t) \approx e^{-r\Delta t}\frac{1}{\pi}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \max\left\{\Phi(S_{t+\Delta t}^{(i,j)}), C^{(k)}(S_{t+\Delta t}^{(i,j)}, v_{t+\Delta t}^{(i,j)})\right\}. \tag{3.4}$$

3.2. Model-specific one-step transitions

The GHQ nodes (x_i) (or (x_i, x_j)) are pushed through the model's one-step transition over Δt to produce off-grid evaluation points for the value/continuation function. In the subsequent equations, we use a superscript index, such as (i) or (i, j), to denote an outcome generated from the corresponding GHQ node.

3.2.1. Black–Scholes (1D)

Under (2.5), the log return over Δt is normal. Using GHQ nodes (x_i),

$$S_{t+\Delta t}^{(i)} = S_t \exp\left[(r - \tfrac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{2\Delta t}\, x_i\right]. \tag{3.5}$$

3.2.2. Heston (2D)

With dynamics (2.6)–(2.7) and correlation ρ, we employ a standard Euler–Maruyama discretization for the one-step transitions [29], using the correlated GHQ shocks (3.3),

$$S_{t+\Delta t}^{(i,j)} = S_t \exp\left[(r - \tfrac{1}{2}v_t)\Delta t + \sqrt{2\Delta t\, v_t}\, x_i\right], \tag{3.6}$$

$$v_{t+\Delta t}^{(i,j)} = v_t + \kappa(\theta - v_t)\Delta t + \xi\sqrt{2\Delta t\, v_t}\left(\rho x_i + \sqrt{1-\rho^2}\, x_j\right). \tag{3.7}$$

We enforce positivity of variance by $v_{t+\Delta t}^{(i,j)} \leftarrow \max\{0, v_{t+\Delta t}^{(i,j)}\}$.[§]

3.3. Deterministic value iteration (Black–Scholes)

Our algorithm computes the continuation value $C(S)$ on a discrete grid of asset prices, denoted by $\{S_\ell\}$. The core of the method is the iterative Bellman update performed at each grid node S_ℓ.

To maintain an intuitive link to the time-continuous dynamics, we will use the generic notation S_t to represent the current state (which corresponds to a grid node S_ℓ) and $S_{t+\Delta t}$ for the state after a single time step Δt. One GHQ-based Bellman update, starting from a given grid node $S_t \equiv S_\ell$, proceeds as follows (Steps 1–3 correspond to Figure 3 and Step 4 to Figure 4):

[§]More sophisticated positivity-preserving schemes for the variance process, such as full truncation or reflection, could also be used at the expense of additional implementation effort and runtime; we found the simple projection to be sufficient over the parameter ranges considered in our experiments.

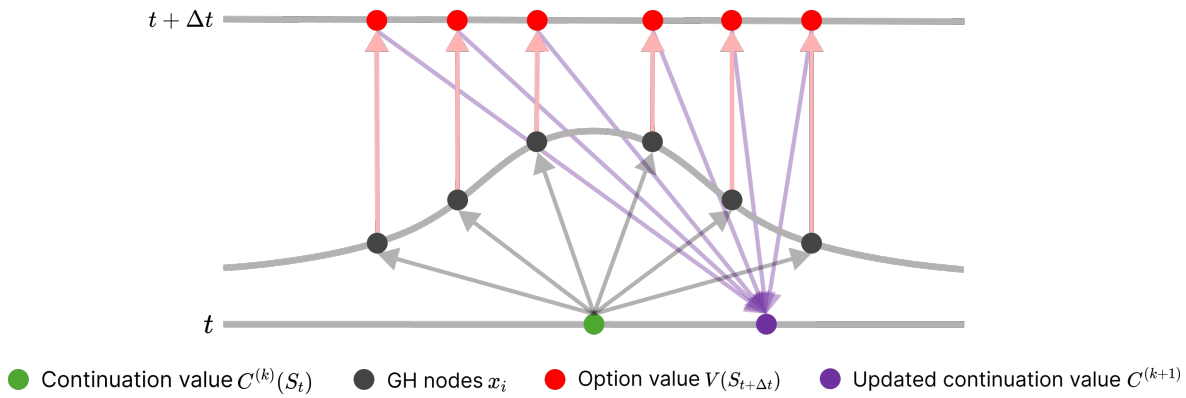

Figure 3. GHQ–VI continuation value update. Computation of $C^{(k+1)}(S_t)$ via GHQ over future asset values $S^{(i)}_{t+\Delta t}$ with weights w_i (see (3.8)).

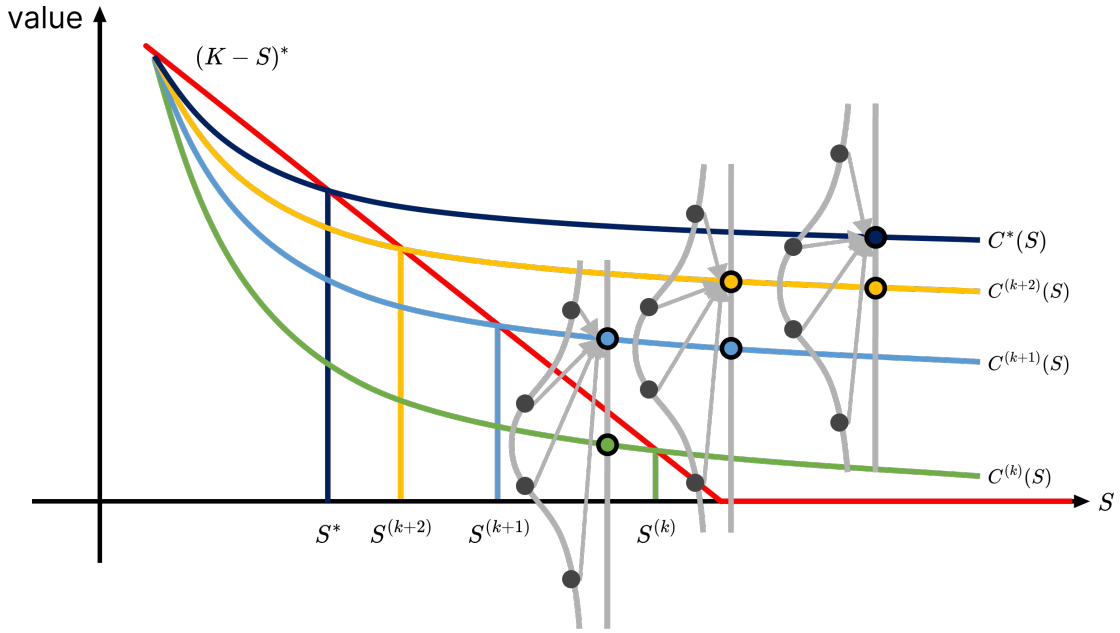

Figure 4. Convergence of continuation value functions $C^{(k)}(S)$ to the fixed point $C^*(S)$ and corresponding boundary S^*.

How one update is computed (at a single grid node S_ℓ).

1. **Map GHQ nodes to futures (Figure 3).** Using (3.5), compute the next-step asset prices $S^{(i)}_{t+\Delta t}$ for all GHQ nodes x_i. This maps standard-normal nodes to lognormal future prices under the BS model.

2. **Evaluate continuation and payoff (Figure 3).** Evaluate the continuation $C^{(k)}$ at each off-grid point $S^{(i)}_{t+\Delta t}$ via a shape-preserving, C^1 PCHIP along S [30]; this preserves local monotonicity and suppresses overshoot near the max-kink. Queries outside the grid are clamped to the nearest grid value. The payoff is $\Phi(S^{(i)}_{t+\Delta t})$, and the one-step value integrand is then

$$V^{(i)}_{\text{step}} \;=\; \max\{\Phi(S^{(i)}_{t+\Delta t}),\; C^{(k)}(S^{(i)}_{t+\Delta t})\},$$

which we distinguish from the payoff to avoid ambiguity.

3. **Aggregate via GHQ (Figure 3).** Form the discounted GHQ average to obtain the updated continuation, as mentioned in (3.2)

$$C^{(k+1)}(S_t) \;=\; e^{-r\Delta t}\,\frac{1}{\sqrt{\pi}}\sum_{i=1}^{n} w_i\, V_{\text{step}}^{(i)}. \tag{3.8}$$

   Consistent with the optimality (value-matching) condition in (2.4), the updated exercise boundary $S^{(k+1)}$ is defined as the (unique) root of

$$C^{(k+1)}(S) - \Phi(S) = 0.$$

4. **Repeat and extract the boundary (Figure 4).** Apply Steps 1–3 to all grid nodes S_ℓ, rebuild the interpolant for $C^{(k+1)}$, and iterate until convergence, i.e., until the boundary change $\left|S^{(k+1)} - S^{(k)}\right| \le \text{tol}_S$ for L-consecutive iterations. The optimal exercise boundary is finally extracted as the root of $C^*(S) - \Phi(S) = 0$.

For reference, a classical value iteration scheme would declare convergence as soon as the value-residual satisfies $\|C^{(k+1)} - C^{(k)}\|_\infty < \text{tol}$. In contrast, GHQ–VI updates the continuation function $C^{(k)}$ in the usual way, but terminates only when the exercise boundary has stabilized, i.e., $\Delta_{\max}^{(k+1)} = \max_v |S^{(k+1)} - S^{(k)}| \le \text{tol}_S$ for L-consecutive iterations. This boundary-based stopping rule is precisely what we mean by a "policy-targeted" value iteration.

3.4. Deterministic value iteration (Heston)

Relative to the BS case, the Heston case adds a variance state v, requiring the algorithm to operate on a two-dimensional grid of states, denoted as $\{(S_\ell, v_m)\}$. The GHQ-based Bellman update remains the same in spirit, but (i) uses the correlated node map (3.3); (ii) requires interpolation in two dimensions (shape-preserving PCHIP along S and linear along v); and (iii) aggregates with $\frac{1}{\pi}\sum_{i,j} w_i w_j$ as in (3.4). The optimal exercise policy is a curve $S^*(v)$ satisfying $C^*(S^*(v), v) = \Phi(S^*(v))$. The full procedure of our GHQ–VI algorithm is summarized in Algorithm 1.

How one update is computed (at a single grid node (S_ℓ, v_m)). For clarity and consistency with the time-continuous dynamics, we let the generic current state (S_t, v_t) be equivalent to the grid node for this step, i.e., $(S_t, v_t) \equiv (S_\ell, v_m)$. The update then proceeds in four steps:

1. **Map GHQ nodes and generate futures (correlation injection).** First, generate correlated standard normal shocks (Z_1, Z_2) by applying the linear map (3.3) to the independent GHQ nodes (x_i, x_j). Then compute the next-step states $(S_{t+\Delta t}^{(i,j)}, v_{t+\Delta t}^{(i,j)})$ using these shocks in the one-step transition dynamics defined by (3.6)–(3.7). Finally, enforce positivity of the variance $v_{t+\Delta t}^{(i,j)} \leftarrow \max\{0,\, v_{t+\Delta t}^{(i,j)}\}$.

2. **Evaluate continuation and payoff at off-grid points.** Evaluate $C^{(k)}$ at $(S_{t+\Delta t}^{(i,j)}, v_{t+\Delta t}^{(i,j)})$ using a tensor-product interpolant: shape-preserving, C^1-continuous PCHIP along S and linear interpolation along v. Queries outside the grid are clamped to the nearest grid value. The payoff is $\Phi(S_{t+\Delta t}^{(i,j)})$. The one-step value integrand is

$$V_{\text{step}}^{(i,j)} = \max\left\{\Phi(S_{t+\Delta t}^{(i,j)}),\; C^{(k)}(S_{t+\Delta t}^{(i,j)}, v_{t+\Delta t}^{(i,j)})\right\}.$$

3. **Aggregate via 2D GHQ.** Update the continuation value according to (3.4)

$$C^{(k+1)}(S_t, v_t) = e^{-r\Delta t} \frac{1}{\pi} \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j V_{\text{step}}^{(i,j)}. \tag{3.9}$$

Consistent with the optimality condition in (2.4), the updated exercise boundary $S^{(k+1)}(v_t)$ is a curve in (S, v) given, for each variance node v_t, by the (unique) root of the value-matching condition

$$C^{(k+1)}(S, v_t) - \Phi(S) = 0.$$

4. **Repeat over the grid and extract the boundary.** Apply Steps 1–3 to all (S_ℓ, v_m), rebuild the interpolant for $C^{(k+1)}$, and iterate until convergence, i.e., until the boundary change $\Delta_{\max}^{(k+1)} = \max_{v_t} \left| S_{v_t}^{(k+1)} - S_{v_t}^{(k)} \right| \le \text{tol}_S$ for L-consecutive iterations. Finally, for each variance slice v_t, extract $S^*(v_t)$ as the root of $C^*(S, v_t) - \Phi(S) = 0$.

Before turning to convergence analysis, we briefly comment on a finite-horizon extension of the GHQ–VI framework.

Remark (finite-horizon extension). The present work focuses on perpetual American puts, where time-homogeneity of the underlying model implies that the optimal exercise boundary is time-independent and can be computed as the fixed point of a stationary Bellman operator. Our GHQ–VI is tailored to this setting.

For a finite maturity T, a natural extension is to work on a discrete time grid $0 = t_0 < t_1 < \cdots < t_M = T$ with $\Delta t = t_{m+1} - t_m$ and to define

$$C_M(S) = \Phi(S), \qquad C_m(S) = \max\left\{\Phi(S),\ e^{-r\Delta t}\, \mathbb{E}[C_{m+1}(S_{t_{m+1}}) \mid S_{t_m} = S]\right\}, \quad m = M-1, \ldots, 0,$$

where $\Phi(S) = (K - S)^+$. The conditional expectation has exactly the same structure as in the perpetual case and can again be approximated at each time level by the GHQ-interpolation operator on a fixed spatial grid. Comparing payoff and continuation then yields a discrete sequence of thresholds $\{S_m^*\}_{m=0}^M$, providing a stepwise (Bermudan-type) approximation of the time-dependent exercise boundary $S^*(t)$. This is conceptually related to boundary-first finite-horizon approaches that approximate $S^*(t)$ on a time grid and solve the associated pricing PDE on the known exercise region (see, e.g., [15, 16]).

Unlike the perpetual case, however, the finite-horizon formulation does not lead to a single fixed-point problem: The collection $\{C_m\}_{m=0}^M$ is generated by a backward recursion. As a result, the overall accuracy depends on a trade-off between the time step Δt (a coarse grid produces a rough, stepwise boundary, especially in low-r or high-volatility regimes) and the accumulation of quadrature and interpolation errors when Δt is very small. A systematic implementation and error analysis of this finite-horizon GHQ–VI extension—including the choice of time discretization, quadrature order, and spatial resolution—would significantly broaden the scope of this paper and is therefore left for future work.

Algorithm 1. The GHQ–VI Algorithm: Deterministic Value Iteration for the Heston Model

Inputs: $K, r, \kappa, \theta, \xi, \rho, \Delta t$; grids $\mathcal{G}_S = \{S_\ell\}$, $\mathcal{G}_v = \{v_m\}$; GHQ nodes/weights $\{x_i, w_i\}_{i=1}^n$; convergence window L; boundary tolerance tol_S

Init: $C^{(0)}(S_\ell, v_m) = \Phi(S_\ell)$; precompute product weights $w_{ij} = w_i w_j$

for $k = 0, 1, \ldots$ **do**

  Build tensor-product interpolant for $C^{(k)}$ (PCHIP in S, linear in v; clamp outside domain)

  for each $(S_\ell, v_m) \in \mathcal{G}_S \times \mathcal{G}_v$ **do**

    for each GHQ pair (x_i, x_j) **do**

      Correlated node map: $Z_1 = \sqrt{2}\, x_i$, $Z_2 = \sqrt{2}\left(\rho x_i + \sqrt{1-\rho^2}\, x_j\right)$ ((3.3))

      Transition:

$$S' \leftarrow S_\ell \exp\left[(r - \tfrac{1}{2} v_m)\Delta t + \sqrt{\Delta t\, v_m}\, Z_1\right],$$
$$v' \leftarrow v_m + \kappa(\theta - v_m)\Delta t + \xi \sqrt{\Delta t\, v_m}\, Z_2$$

      Positivity: $v' \leftarrow \max\{0, v'\}$

      Interpolation: $\widehat{C} \leftarrow \widehat{\mathcal{I}}(C^{(k)}; S', v')$ ($\widehat{\mathcal{I}}$: PCHIP in S, linear in v, clamp)

      One-step value: $V^{(i,j)} \leftarrow \max\left(\Phi(S'), \widehat{C}\right)$

    end for

    Bellman update: $C^{(k+1)}(S_\ell, v_m) \leftarrow e^{-r\Delta t} \dfrac{1}{\pi} \sum_{i,j} w_{ij}\, V^{(i,j)}$

    Boundary update: for each $v_m \in \mathcal{G}_v$, solve $S^{(k+1)}(v_m)$ s.t. $C^{(k+1)}(S, v_m) = \Phi(S)$

  end for

  Convergence check: If $\Delta_{\max}^{(k+1)} = \max_m \left|S^{(k+1)}(v_m) - S^{(k)}(v_m)\right| \le \mathrm{tol}_S$ for L-consecutive iterations, **break**.

end for

Boundary extraction: for each v_m, find $S^*(v_m)$ solving $C^*(S, v_m) = \Phi(S)$

4. Convergence and error guarantees

We now establish the core theoretical guarantees for GHQ–VI. Our analysis is set in the Banach space $(C_b(\mathcal{X}), \|\cdot\|_\infty)$, where $\mathcal{X}$ denotes the state space and a generic state is written $x \in \mathcal{X}$ (while X_t denotes the Markov state process). The specific form of $\mathcal{X}$ depends on the model.

- For BS, $\mathcal{X} \subseteq \mathbb{R}^+$ and $X = S$.
- For Heston, $\mathcal{X} \subseteq \mathbb{R}^+ \times \mathbb{R}^+$ and $X = (S, v)$.

For numerical implementation, $\mathcal{X}$ is typically truncated to a compact domain (BS: $[S_{\min}, S_{\max}]$; Heston: $[S_{\min}, S_{\max}] \times [0, v_{\max}]$). The space $C_b(\mathcal{X})$ consists of all bounded, continuous functions on $\mathcal{X}$ endowed with the supremum norm $\|f\|_\infty := \sup_{x\in\mathcal{X}} |f(x)|$.

Let $\gamma := e^{-r\Delta t} \in (0, 1)$. The stationary Bellman operator $\mathcal{T} : C_b(\mathcal{X}) \to C_b(\mathcal{X})$ is defined, for any $C \in C_b(\mathcal{X})$ and $x \in \mathcal{X}$, by

$$(\mathcal{T}C)(x) = \gamma\, \mathbb{E}\left[\max\{\Phi(X_{t+\Delta t}), C(X_{t+\Delta t})\} | X_t = x\right]. \tag{4.1}$$

Standing assumptions.

1. $r > 0$.
2. (Feller) For every bounded continuous function f, the map $x \mapsto \mathbb{E}[\, f(X_{t+\Delta t}) \mid X_t = x\,]$ is continuous. (This holds for both the BS and Heston models.)
3. Φ is bounded and continuous except for the kink at $S = K$.
4. (Boundary regularity) For each variance slice $v \in [0, v_{\max}]$ in the Heston model, the section $S \mapsto (C^* - \Phi)(S, v)$ has a unique zero $S^*(v)$ with a strict sign change across it, and $v \mapsto S^*(v)$ is continuous.

Theorem 1 (Contraction and global convergence). On $(C_b(\mathcal{X}), \|\cdot\|_\infty)$, the Bellman operator $\mathcal{T}$ in (4.1) is a γ-contraction

$$\|\mathcal{T}C_1 - \mathcal{T}C_2\|_\infty \le \gamma \, \|C_1 - C_2\|_\infty.$$

Hence, $\mathcal{T}$ has a unique fixed point C^*, and for any $C^{(0)} \in C_b(\mathcal{X})$, the iteration $C^{(k+1)} = \mathcal{T}C^{(k)}$ satisfies

$$\|C^{(k)} - C^*\|_\infty \le \gamma^k \, \|C^{(0)} - C^*\|_\infty.$$

Moreover, if $C^{(0)} \le C^*$ (e.g., $C^{(0)} = \Phi$), then $C^{(k)} \nearrow C^*$.

Theorem 2 (Stability of the numerical fixed point). Let $\widehat{\mathcal{T}} : C_b(\mathcal{X}) \to C_b(\mathcal{X})$ denote the GHQ-based approximate Bellman operator with interpolation, formally defined for each model in Appendix B ((B.1)–(B.2)). Assume $\widehat{\mathcal{T}}$ is a γ-contraction in $\|\cdot\|_\infty$ and define the one-step operator discrepancy

$$\varepsilon := \sup_{C \in C_b(\mathcal{X})} \left\| \widehat{\mathcal{T}}C - \mathcal{T}C \right\|_\infty. \tag{4.2}$$

If $\widehat{C}^*$ and C^* denote the unique fixed points of $\widehat{\mathcal{T}}$ and $\mathcal{T}$, respectively, then

$$\begin{aligned} \|\widehat{C}^* - C^*\|_\infty &\le \frac{\varepsilon}{1-\gamma}, \\ \|C^{(k)} - \widehat{C}^*\|_\infty &\le \gamma^k \, \|C^{(0)} - \widehat{C}^*\|_\infty. \end{aligned}$$

Remark. The operator discrepancy ε is not a tunable parameter; it aggregates errors arising from quadrature, interpolation, time-discretization, and domain-truncation. While it can be reduced by measures such as increasing the GHQ order or refining the computational grid, it cannot be eliminated entirely.

Corollary (Convergence of the exercise boundary). Let $g^{(k)} := C^{(k)} - \Phi$ and $g^* := C^* - \Phi$. If $C^{(k)} \to C^*$ uniformly and the boundary regularity condition in Standing assumption 4 holds, then the exercise boundary extracted from $C^{(k)}$ converges slice-wise: $S^{(k)}(v) \to S^*(v)$ for each $v \in [0, v_{\max}]$ (v being absent in the BS case). Moreover, if the sign-change margin is uniform in v, the convergence becomes uniform in v as well. (Proof sketch in Appendix B.)

Remark (Stopping rule and accuracy certificate). In practice, the iteration is stopped once the boundary stabilizes, i.e., when

$$\Delta_{\max}^{(k+1)} := \max_v \, |S^{(k+1)}(v) - S^{(k)}(v)| \le \mathrm{tol}_S$$

for L-consecutive iterations, where tol_S is a user-defined stopping tolerance. This pragmatic rule is justified by a residual-based certificate for boundary accuracy. Letting

$$\Delta_C^{(k+1)} := \|C^{(k+1)} - C^{(k)}\|_\infty, \tag{4.3}$$

the residual-distance inequality (B.8) and the separation margin $\omega(\varepsilon)$ from (B.7) provide the following guarantee:

$$\Delta_C^{(k+1)} \;\le\; (1-\gamma)\,\tfrac{1}{2}\,\omega(\varepsilon) \Longrightarrow \sup_v |S^{(k)}(v) - S^*(v)| \;\le\; \varepsilon.$$

This provides a computable accuracy certificate: Once the Bellman residual drops below the threshold, the computed boundary $S^{(k)}(v)$ is guaranteed to be uniformly ε-close to the true boundary $S^*(v)$.

5. Experimental results

5.1. Experimental setup

We evaluate the proposed GHQ–VI under both the BS and Heston models. Our goals are threefold: (i) measure how accurately the learned optimal early-exercise boundary matches the analytic solution (BS) or FDM-LCP (Heston) benchmarks; (ii) quantify the impact of boundary discrepancies on realized option value under a consistent Monte Carlo protocol; and (iii) report elapsed time for the proposed method (and, for the Heston model, compare against an FDM baseline).

Bellman updates. We compute discounted expectations of $\max\{\Phi, C^{(k)}\}$ using the GHQ-based Bellman updates in (3.2) (BS) and (3.4) (Heston), with one-step transitions defined in section 3.2. For the number of quadrature nodes, we find different settings to be optimal for each model. In the one-dimensional BS case, we use $n = 9$ nodes. For the two-dimensional Heston model, we employ $n = 7$ nodes for each dimension, resulting in a 7×7 product grid of quadrature points (49 points in total). In our convergence study, we also tested $n = 9$ nodes for the Heston model but found no meaningful accuracy improvement compared with $n = 7$; however, the computational cost increased significantly due to the 9×9 quadrature grid (81 points versus 49), resulting in nearly twice the runtime.

Discretization. For the BS model, we use a uniform grid in S, while for the Heston model we use a tensor grid in (S, v). The asset price domain for both models is set to $[0, S_{\max}]$, with the upper bound $S_{\max}$ chosen sufficiently large and scaled relative to the strike (e.g., $S_{\max} = 3.3K$). This scaling ensures that the domain adequately covers the continuation region, particularly in low-interest-rate or high-volatility scenarios, so that the option value is already very close to zero near $S_{\max}$. In preliminary experiments, we also tested alternative values such as $S_{\max} = 3.0K$ and $S_{\max} = 3.6K$ and observed no meaningful differences in the reported boundary or value errors, so we fixed $S_{\max} = 3.3K$ as a reasonable compromise between truncation accuracy and computational cost. For the Heston model, the variance domain is $[0, v_{\max}]$, where $v_{\max}$ is set several standard deviations above the long-run mean θ (e.g., $v_{\max} = \theta + 3\sqrt{\xi^2\theta/2\kappa}$).

The number of grid points is $|\mathcal{G}_S| = 500$ for the BS model and $|\mathcal{G}_S| \times |\mathcal{G}_v| = 1000 \times 100$ for the Heston model. These choices were informed by a convergence study: Roughly doubling the number of grid points in each direction (e.g., $|\mathcal{G}_S| = 1000$ for BS and 1500×200 for Heston) produced boundary

and policy-value results that differed negligibly from those reported in Section 5, while the runtime increased substantially. The time step for one-step transitions is $\Delta t = 1/252$ (daily), and off-grid values are evaluated using the interpolation scheme in Section 3.

Stopping criteria and runtime accounting. For our experiments, we set the tolerance to $\mathrm{tol}_S = 10^{-4}$ and the convergence window to $L = 10{,}000$. Although this choice of L may appear large, the Bellman updates continue to adjust the continuation value $C^{(k)}$ while the payoff Φ is fixed, and changes in $C^{(k)} - \Phi$ near the free boundary become increasingly small and slow. In particular, for the Heston model the exercise boundary is a curve $S^*(v)$ over all variance grid points; we require that the sign of $C^{(k)}(S, v) - \Phi(S)$ stops changing across the entire (S, v)-grid. We therefore use a relatively long convergence window to ensure that no further sign changes occur along the boundary in any variance slice and that the estimated exercise boundary is genuinely stabilized. We also enforce a hard cap on iterations. For the BS model, we report the elapsed time of the proposed GHQ–VI value iteration only; it includes per-iteration costs (transition mapping, interpolation, GHQ aggregation, and interpolant rebuild). For the Heston model, we report both the GHQ–VI elapsed time (as above) and the FDM baseline elapsed time, which includes PDE assembly and RB-PSOR iterations until convergence.

Benchmarks and metrics. For the BS model, the ground truth is the analytic perpetual put boundary; for the Heston model, we use a stationary FDM-LCP reference solved by RB-PSOR (Appendix A). Hereafter, "prop" denotes the proposed GHQ–VI method, and "bench" denotes the analytic (BS) or FDM-LCP (Heston) benchmark; the corresponding quantities are written with the subscripts "prop" and "bench" as needed. For the optimal exercise boundary, we report the relative error

$$\mathrm{MAPE}_{\mathrm{bnd}} = \left| \frac{S^*_{\mathrm{prop}} - S^*_{\mathrm{bench}}}{S^*_{\mathrm{bench}}} \right|, \tag{5.1}$$

while policy value differences are assessed via Monte Carlo policy evaluation as described below and in Section 5.4.

Monte Carlo protocol (policy value impact). We simulate 10^5 risk-neutral paths per configuration with common random numbers (fixed seed). The stopping time τ is the first hitting time of the policy boundary (BS: $S_t \le S^*$; Heston: $S_t \le S^*(v_t)$), and the discounted payoff is $e^{-r\tau}\,\Phi(S_\tau)$. Paths that do not hit the boundary before the truncation time $T = 100$ years contribute a zero payoff. Using common paths isolates the effect of the policy choice [10]. These Monte Carlo runs are used only to assess policy value accuracy and are not included in the reported runtimes; the resulting value differences and confidence intervals are reported in Section 5.4.

Computing environment. All experiments were implemented in Python 3.12 and executed in a Google Colab environment equipped with an NVIDIA A100-SXM4-40GB GPU (40 GB device memory) and a 12-core Intel(R) Xeon(R) CPU @ 2.20 GHz.

5.2. Black–Scholes results

Table 1 summarizes the proposed (GHQ–VI) boundaries across a range of (K, σ, r) and compares them with the analytic solution. Accuracy is uniformly high: Absolute boundary errors are typically

well below one unit of underlying price, and the boundary MAPE remains below 1.49% across all entries (MAPE values are reported in percent), including low-r/high-σ regimes. Elapsed times range from about 19–56 seconds across the parameter sets.

These BS configurations are also used in the Monte Carlo policy evaluation of Section 5.4, where we compare the policies induced by the GHQ–VI and analytic boundaries under a common-random-number protocol. The analysis shows that, despite boundary deviations of up to about 1.5%, the resulting relative policy value gaps are numerically negligible, confirming that modest boundary misspecification has only a minor impact on the perpetual option value in the BS setting.

Table 1. Black–Scholes (BS): Benchmark (analytic) vs proposed (GHQ–VI) optimal exercise boundaries.

Parameters			**Boundary S^***			
K	σ	r	**Benchmark**	**Proposed**	**MAPE [%]**	**Elapsed time [s]**
80	0.2	0.02	40.00	39.41	1.49	56
80	0.2	0.03	48.00	47.67	0.69	40
80	0.2	0.04	53.33	53.20	0.25	36
80	0.4	0.02	16.00	15.90	0.60	19
80	0.4	0.03	21.82	21.60	0.98	27
80	0.4	0.04	26.67	26.39	1.04	26
100	0.2	0.02	50.00	49.26	1.49	54
100	0.2	0.03	60.00	59.59	0.69	40
100	0.2	0.04	66.67	66.50	0.25	36
100	0.4	0.02	20.00	19.88	0.60	19
100	0.4	0.03	27.27	26.93	1.26	34
100	0.4	0.04	33.33	32.98	1.05	24

Note: The boundary MAPE is computed according to (5.1). "Elapsed time" reports the time required by the GHQ–VI method to compute the optimal exercise boundary S^*. Benchmark corresponds to the closed-form analytic solution.

5.3. Heston results

Under the Heston model, where no closed-form boundary is available, we compare the GHQ–VI boundary $S^*_{\text{prop}}(v)$ to the two-dimensional FDM-LCP benchmark $S^*_{\text{bench}}(v)$ (see Appendix A for details). Figure 5 shows six representative configurations. In each panel, the top subplot displays the boundary curves across variance levels v, and the bottom subplot shows MAPE_{bnd} in (5.1). Across the tested settings, the GHQ–VI-based boundary preserves the qualitative shape: It is monotone decreasing in v (higher instantaneous variance leads to earlier exercise), and its curvature with respect to v tracks the FDM curve closely for most of the range. Quantitatively, the relative discrepancy is generally 1–5%, with the largest deviations concentrated at higher θ (long-term variance) and lower r. This increase in boundary discrepancies under low interest rates and high long-run variance can be explained both economically and numerically. When r is small, discounting is weak and the effective horizon of the perpetual put becomes very long, making the option value more sensitive to the volatility dynamics and increasing its curvature in the variance direction. Similarly, a larger long-run variance level θ

keeps the variance process in high-volatility states for longer, further amplifying the sensitivity of the continuation value to v. In such high-curvature regimes, small local approximation errors—stemming from fixed-order quadrature and interpolation on a finite (S, v) grid—can be magnified when translated into the free-boundary location, which explains why the GHQ–VI boundary deviates more visibly from the FDM-LCP benchmark. The impact of these boundary discrepancies on option values is assessed via Monte Carlo policy evaluation in Section 5.4.

For runtime comparison and scalability, the Heston FDM baseline requires solving a large-scale linear complementarity problem on a 2D grid each time parameters change. As reported in Table 2, this typically takes 2,103 − 3,119 seconds per configuration on our setup. In contrast, the proposed method converges in just 127−180 seconds per configuration, achieving a significant speedup of over 15 times. This efficiency arises from bypassing complex PDE assembly and iterative RB-PSOR sweeps. We also observe that GHQ–VI runtimes are less sensitive to parameter variations because the deterministic expectation operator has a predictable cost and the boundary stabilizes after a comparable number of iterations across cases. In larger grids or tighter tolerances, the difference becomes more pronounced as FDM scales with the 2D grid size, whereas the GHQ–VI loop primarily scales with the (1D or separable) quadrature and interpolation costs.

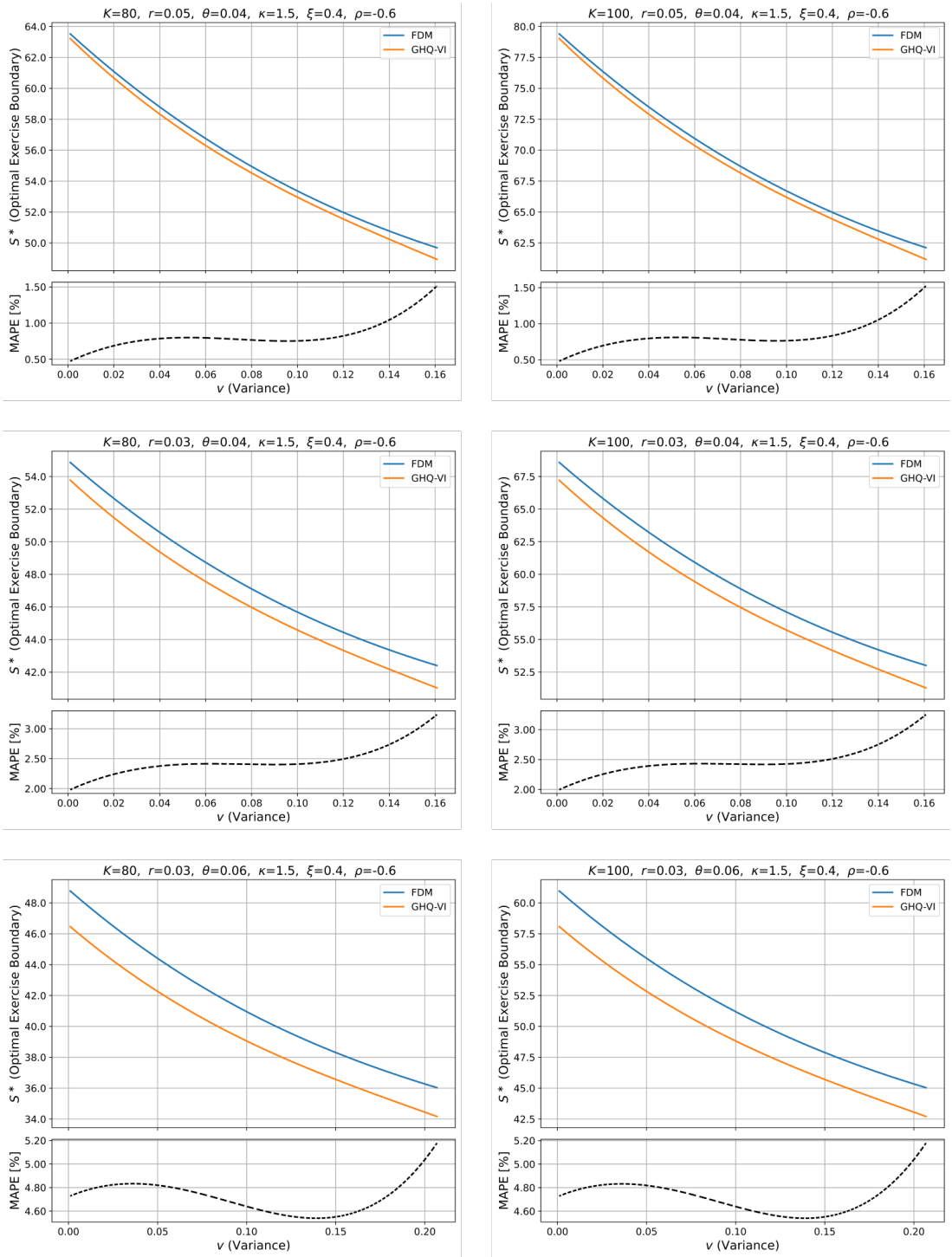

Figure 5. Comparison of optimal exercise boundaries under the Heston model for six different parameter settings. In each panel, the top subplot compares the FDM benchmark boundary with the one found by GHQ–VI (the curves are spline-smoothed for visual clarity only; all reported metrics are computed from the original unsmoothed data), while the bottom subplot shows the corresponding relative error, $\mathrm{MAPE}_{\mathrm{bnd}}(v)$.

Table 2. Heston: Benchmark (FDM-LCP) vs proposed (GHQ–VI) boundary computation times.

Parameters			**Elapsed time [s] to compute $S^*(v)$**	
K	r	θ	**Benchmark**	**Proposed**
80	0.05	0.04	2,103	127
80	0.03	0.04	3,027	172
80	0.03	0.06	2,618	179
100	0.05	0.04	2,180	136
100	0.03	0.04	3,119	179
100	0.03	0.06	2,678	180

Note: The FDM-LCP method (solved with RB-PSOR) serves as the deterministic benchmark for computing the optimal exercise boundary $S^*(v)$. Benchmark values therefore correspond to the high-resolution FDM-LCP solution. The reported runtimes reflect the total time required to compute the full boundary curve under each method. Throughout, the Heston parameters are fixed at $\kappa = 1.5$, $\xi = 0.4$, and $\rho = -0.6$, while K, r, and θ are varied as shown.

5.4. Monte Carlo policy evaluation

In addition to the deterministic comparisons against the analytic (BS) and finite-difference (Heston) benchmarks in Sections 5.2 and 5.3, we perform a Monte Carlo policy evaluation to quantify how boundary approximation errors translate into losses in the option value. For each model and parameter configuration, let τ^{bench} and τ^{prop} denote the stopping times induced by the reference exercise boundary (closed-form in the BS case, high-resolution FDM-LCP in the Heston case) and by the GHQ–VI boundary, respectively.

We follow the Monte Carlo protocol described in Section 5.1 ($N = 10^5$ daily steps up to $T = 100$ years with common random numbers). For each path $i = 1, \ldots, N$, we compute the discounted payoff under the two policies

$$P^{(i)}_{\text{bench}} = e^{-r\tau^{(i)}_{\text{bench}}}\,\Phi\left(S^{(i)}_{\tau^{(i)}_{\text{bench}}}\right), \qquad P^{(i)}_{\text{prop}} = e^{-r\tau^{(i)}_{\text{prop}}}\,\Phi\left(S^{(i)}_{\tau^{(i)}_{\text{prop}}}\right),$$

where paths that do not hit the boundary before T contribute a zero payoff. We then form the pathwise difference

$$D^{(i)} = P^{(i)}_{\text{prop}} - P^{(i)}_{\text{bench}}$$

and define the Monte Carlo estimates of the benchmark and GHQ–VI policy values and their difference as

$$\hat{V}_{\text{bench}} = \frac{1}{N}\sum_{i=1}^{N} P^{(i)}_{\text{bench}}, \qquad \hat{V}_{\text{prop}} = \frac{1}{N}\sum_{i=1}^{N} P^{(i)}_{\text{prop}}, \qquad \hat{\Delta}_N = \hat{V}_{\text{prop}} - \hat{V}_{\text{bench}} = \frac{1}{N}\sum_{i=1}^{N} D^{(i)}.$$

The sample variance of $D^{(i)}$ and the standard error of $\hat{\Delta}_N$ are given by

$$s_N^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(D^{(i)} - \hat{\Delta}_N\right)^2, \qquad \text{SE}(\hat{\Delta}_N) = \frac{s_N}{\sqrt{N}},$$

and we report the 95% confidence interval $\hat{\Delta}_N \pm 1.96\,\text{SE}(\hat{\Delta}_N)$ together with the relative policy value gap $\hat{\Delta}_N / \hat{V}_{\text{bench}}$.

In this section, both the benchmark value $\hat{V}_{\text{bench}}$ and the GHQ–VI value $\hat{V}_{\text{prop}}$ are therefore estimated by Monte Carlo, and the reported relative policy value gap $\hat{\Delta}_N/\hat{V}_{\text{bench}}$ is normalized by the Monte Carlo estimate of the benchmark policy value. This choice ensures that numerator and denominator are measured under exactly the same discrete-time simulation and truncation scheme (including the fact that paths which do not hit the boundary before T contribute a zero payoff), so that the reported ratios reflect pure policy effects rather than differences between the PDE and Monte Carlo discretizations. Deterministic benchmark values for comparison are still provided by the analytic solution (BS) and the high-resolution FDM-LCP scheme (Heston) in Sections 5.2 and 5.3.

For the BS model, the relative policy value gaps $\hat{\Delta}_N/\hat{V}_{\text{bench}}$ in Table 3(a) are extremely small: In absolute value, they are all below approximately 0.07% of the benchmark price. Several confidence intervals include zero, and even in the cases where the interval lies strictly below zero, the magnitude of the estimated loss remains negligible compared to the option value level. These results are consistent with Section 5.2, where boundary MAPE remains below 1.5%, and suggest that such modest boundary deviations have only a minor effect on the perpetual policy value in the BS setting within the precision of our experiments.

Across all configurations, the 95% confidence intervals for $\hat{\Delta}_N$ are extremely narrow relative to the corresponding option value levels, indicating that Monte Carlo sampling error is negligible compared with the discretization and boundary-approximation errors.

For the Heston model, Table 3(b) shows a similarly benign picture. Across all six configurations, the relative policy value gaps $\hat{\Delta}_N/\hat{V}_{\text{bench}}$ remain well below 0.2% in absolute value (the largest deviation is about 0.12%), and most confidence intervals contain zero. Figure 6 complements the table by plotting the empirical discounted payoff histograms: In each configuration, the benchmark (FDM-LCP) and GHQ–VI histograms almost overlap, with the GHQ–VI distribution exhibiting at most a very slight right shift. This confirms that, even though the Heston boundaries in Figure 5 display relative discrepancies of the order of 1–5% (especially in low-r or high-θ regimes), the induced policy values remain extremely close in the perpetual limit.

Overall, the proposed method (i) reproduces the analytic BS boundary with high fidelity and yields policy values that are almost indistinguishable from the benchmark, and (ii) in the Heston setting, although boundary discrepancies widen in low-r or high-θ regimes, the Monte Carlo policy evaluation indicates that the GHQ–VI policy values remain robust, with relative policy value gaps well below 0.2% in our experiments, while substantially reducing runtime and implementation complexity. These numerical findings are broadly consistent with theoretical results on ε-optimal exercise regions for American options, where families of near-optimal exercise policies are shown to produce prices confined to an explicitly bounded interval around the true American value [17]. Our Monte Carlo evidence can be viewed as an empirical instance of this value robustness in the perpetual BS and Heston settings.

Table 3. Discounted policy values under Black–Scholes (BS) and Heston models.

(a) Black–Scholes

K	σ	r	**Benchmark**	**Proposed**	$\hat{\Delta}_N$	SE($\hat{\Delta}_N$)	95% CI	$\hat{\Delta}_N/\hat{V}_{\text{bench}}$ [%]
80	0.2	0.02	19.68	19.67	-0.0131	0.0073	[-0.0275, 0.0012]	-0.0668
80	0.2	0.03	14.76	14.75	-0.0054	0.0051	[-0.0154, 0.0045]	-0.0367
80	0.2	0.04	11.78	11.78	-0.0012	0.0030	[-0.0070, 0.0046]	-0.0102
80	0.4	0.02	42.52	42.52	0.0023	0.0029	[-0.0034, 0.0079]	0.0053
80	0.4	0.03	35.55	35.54	-0.0135	0.0049	[-0.0231, -0.0039]	-0.0380
80	0.4	0.04	30.66	30.65	-0.0169	0.0056	[-0.0277, -0.0060]	-0.0550
100	0.2	0.02	24.60	24.58	-0.0164	0.0092	[-0.0344, 0.0016]	-0.0666
100	0.2	0.03	18.45	18.44	-0.0067	0.0063	[-0.0191, 0.0057]	-0.0362
100	0.2	0.04	14.72	14.72	-0.0016	0.0037	[-0.0089, 0.0057]	-0.0109
100	0.4	0.02	53.14	53.15	0.0029	0.0036	[-0.0042, 0.0100]	0.0055
100	0.4	0.03	44.44	44.42	-0.0199	0.0069	[-0.0335, -0.0063]	-0.0447
100	0.4	0.04	38.33	38.31	-0.0207	0.0070	[-0.0343, -0.0071]	-0.0540

(b) Heston

K	r	θ	**Benchmark**	**Proposed**	$\hat{\Delta}_N$	SE($\hat{\Delta}_N$)	95% CI	$\hat{\Delta}_N/\hat{V}_{\text{bench}}$ [%]
80	0.05	0.04	11.06	11.07	0.0086	0.0066	[-0.0044, 0.0215]	0.0773
80	0.03	0.04	16.08	16.08	0.0004	0.0105	[-0.0201, 0.0209]	0.0025
80	0.03	0.06	20.89	20.91	0.0260	0.0138	[-0.0010, 0.0530]	0.1243
100	0.05	0.04	13.83	13.84	0.0110	0.0083	[-0.0052, 0.0272]	0.0794
100	0.03	0.04	20.10	20.10	0.0005	0.0131	[-0.0251, 0.0262]	0.0027
100	0.03	0.06	26.11	26.14	0.0322	0.0172	[-0.0015, 0.0660]	0.1235

Note: $\hat{\Delta}_N$ denotes the Monte Carlo estimate of the value difference $\hat{V}_{\text{prop}} - \hat{V}_{\text{bench}}$ based on $N = 10^5$ paths with common random numbers, SE($\hat{\Delta}_N$) is its standard error, and the last column reports the relative policy value gap $\hat{\Delta}_N/\hat{V}_{\text{bench}}$ (in percent) with respect to the Monte Carlo benchmark value $\hat{V}_{\text{bench}}$. In panel (a), the "Benchmark" column reports the Monte Carlo estimate $\hat{V}_{\text{bench}}$ for the policy induced by the analytic BS perpetual put boundary, while in panel (b), it reports $\hat{V}_{\text{bench}}$ for the policy induced by the high-resolution FDM-LCP boundary.

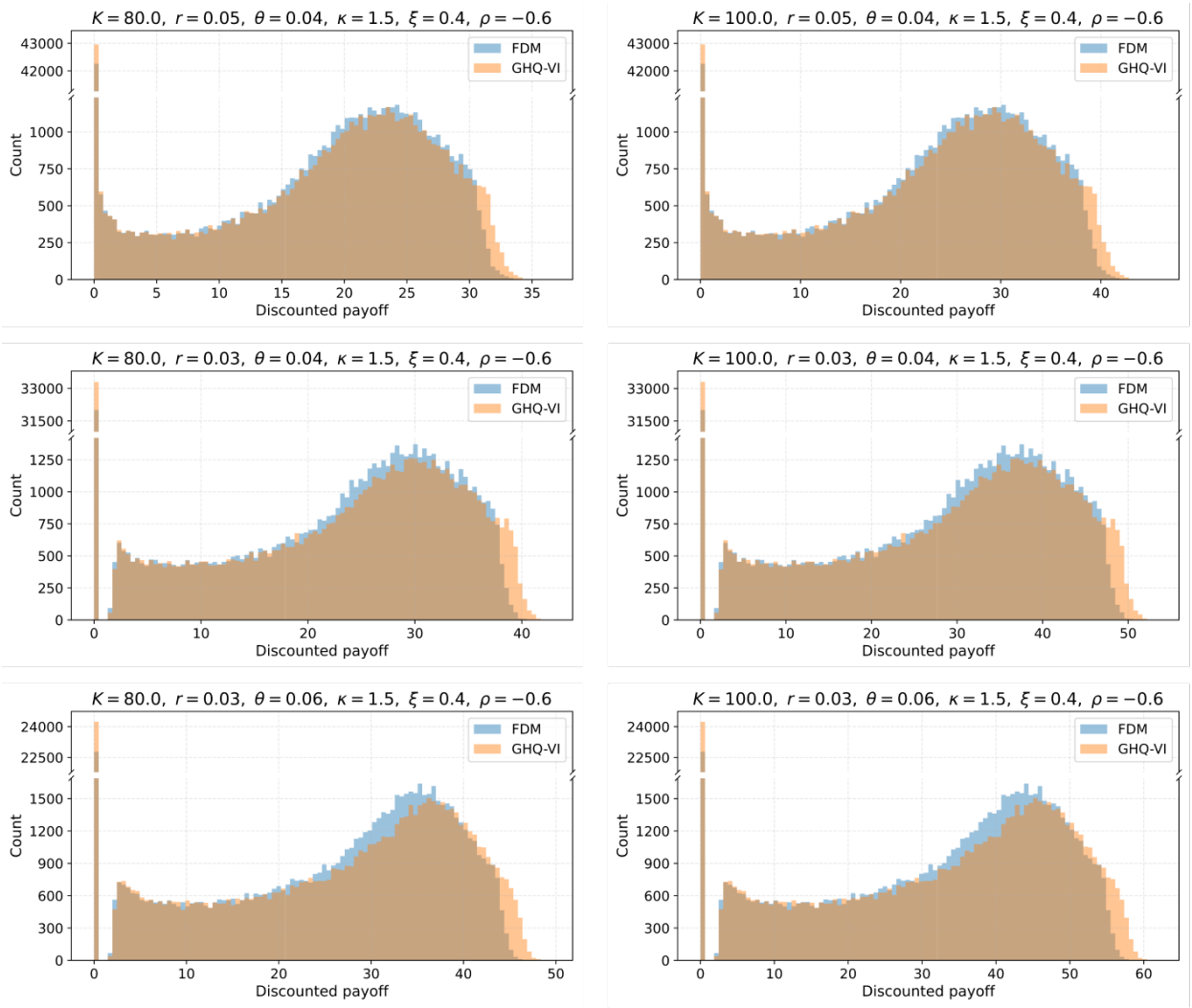

Figure 6. Discounted payoff histograms under the Heston model. Overlaid histograms for the benchmark (FDM-LCP, blue) and GHQ–VI (orange) policies, using common random numbers. The six panels correspond to the same parameter settings as in Figure 5.

6. Conclusions

We propose a GHQ–VI algorithm to estimate the optimal exercise boundary for perpetual American put options under both the BS and Heston models. The method's core approach is to replace path sampling with deterministic GHQ for evaluating the expectation in the Bellman operator, while using a stable, monotone PCHIP for robust off-grid evaluations. In combination, these choices substantially reduce runtime and computational cost relative to Monte Carlo while retaining high approximation accuracy.

Empirically, in the BS setting, our method closely recovers the theoretical boundary: Across a wide range of (K, σ, r) tuples, the boundary MAPE is $\leq$ 1.49%. When we use the GHQ–VI and analytic boundaries as policies in 10^5 Monte Carlo paths, the resulting policy values are almost indistinguishable from each other. The relative policy value gaps $\hat{\Delta}_N/\hat{V}_{\text{bench}}$ in our experiments are

all below about 0.07% in absolute value. Runtimes range from about 19–56 s, confirming that the proposed numerical procedure is both highly accurate and computationally efficient.

For the Heston model, where no closed form exists, we benchmark against a high-resolution FDM with a RB-PSOR solver for the associated LCP. The GHQ–VI boundary captures the expected monotone decline in the optimal threshold as variance increases, with relative boundary errors around 1–5% depending on parameters (discrepancy increases as interest rates fall and volatility rises). Crucially, when the two boundaries (FDM vs GHQ–VI) are each deployed as policies in 10^5 Monte Carlo paths, the expected payoff differences are small, with the relative policy value gaps generally at most about 0.2% in the common-random-number Monte Carlo evaluation.

Overall, the proposed GHQ–VI framework (i) enables accurate evaluation of expectations without path simulation; (ii) extends naturally from the 1D BS state space to the 2D Heston (S, v) grid via interpolation; and (iii) offers a reliable and efficient alternative when closed forms are unavailable, PDE-based approaches are complex to implement, or Monte Carlo is costly. In both models, the largest boundary discrepancies arise in low-r or high-θ regimes, where the value function is more curved and more sensitive to the variance state; refining the scheme specifically in these high-curvature regions (e.g., via adaptive quadrature or local mesh refinement) is an interesting direction for future work. More broadly, the challenge of improving boundary accuracy in these extreme regimes points to several additional refinements, including adaptive selection of GHQ order and nodes, boundary localization with mitigation of interpolation bias, and iteration acceleration (e.g., improved relaxation and post-projection strategies).

In summary, our framework provides a flexible and robust numerical approach for optimal stopping problems. By focusing directly on the optimal policy, it delivers robust performance from the perspective that matters most in practice: the final realized payoffs.

Author contributions

J. Jeon: Conceptualization, formal analysis, writing – review & editing, supervision, funding acquisition; E. Kim: Software, investigation, writing – original draft, visualization; J. Jeon, E. Kim: Methodology, validation. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors used ChatGPT solely for language polishing of several paragraphs; all content was reviewed and verified by the authors, who take full responsibility for the text and results.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Nos. RS-2023-00242528 and RS-2024-00355646).

This research was supported by Global-Learning Academic research institution for Master's · PhD students, and Postdocs (LAMP) Program of the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS-2024-00442775).

This result was supported by the "Regional Innovation System & Education (RISE)" through the Gwangju RISE Center, funded by the Ministry of Education (MOE) and the Gwangju Metropolitan Government, Republic of Korea (2025-RISE-05-011).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. F. Black, M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.*, **81** (1973), 637–654. https://doi.org/10.1086/260062
2. R. C. Merton, Theory of rational option pricing, *Bell J. Econ. Manag. Sci.*, **4** (1973), 141–183. https://doi.org/10.2307/3003143
3. M. J. Brennan, E. S. Schwartz, The valuation of American put options, *J. Financ.*, **32** (1977), 449–462. https://doi.org/10.2307/2326779
4. I. J. Kim, The analytic valuation of American options, *Rev. Financ. Stud.*, **3** (1990), 547–572. https://doi.org/10.1093/rfs/3.4.547
5. H. P. McKean Jr, A free boundary problem for the heat equation arising from a problem in mathematical economics, *Indust. Manag. Rev.*, **6** (1965), 32–39.
6. S. L. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Rev. Financ. Stud.*, **6** (1993), 327–343. https://doi.org/10.1093/rfs/6.2.327
7. S. Foulon, ADI finite difference schemes for option pricing in the heston model with correlation, *Int. J. Numer. Anal. Mod.*, **7** (2010), 303–320.
8. F. A. Longstaff, E. S. Schwartz, Valuing American options by simulation: A simple least-squares approach, *Rev. Financ. Stud.*, **14** (2001), 113–147. https://doi.org/10.1093/rfs/14.1.113
9. J. N. Tsitsiklis, B. Van Roy, Regression methods for pricing complex American-style options, *IEEE T. Neural Networ.*, **12** (2001), 694–703. https://doi.org/10.1109/72.935083
10. P. Glasserman, *Monte Carlo methods in financial engineering*, Springer, **53** (2004). https://doi.org/10.1007/978-0-387-21617-1
11. R. M. Reesor, L. Stentoft, X. Zhu, A critical analysis of the weighted least squares Monte Carlo method for pricing American options, *Financ. Res. Lett.*, **64** (2024), 105379. https://doi.org/10.1016/j.frl.2024.105379
12. M. B. Haugh, L. Kogan, Pricing American options: A duality approach, *Oper. Res.*, **52** (2004), 258–270. https://doi.org/10.1287/opre.1030.0070
13. L. Andersen, M. Broadie, Primal-dual simulation algorithm for pricing multidimensional american options, *Manag. Sci.*, **50** (2004), 1222–1234. https://doi.org/10.1287/mnsc.1040.0258
14. Y. Chen, J. W. Wan, Deep neural network framework based on backward stochastic differential equations for pricing and hedging American options in high dimensions, *Quant. Financ.*, **21** (2021), 45–67. https://doi.org/10.1080/14697688.2020.1788219

15. T. S. Zaevski, A new approach for pricing discounted american options, *Commun. Nonlinear Sci.*, **97** (2021), 105752.

16. T. S. Zaevski, H. Sariev, M. Savov, A fast and accurate numerical approach for pricing American-style power options, *Mathematics*, **13** (2025), 2031.

17. T. Zaevski, On the -optimality of american options, *China Financ. Rev. Int.*, **15** (2025), 688–714. https://doi.org/10.1108/CFRI-06-2024-0361

18. X. Luo, P. V. Shevchenko, Fast and simple method for pricing exotic options using Gauss-Hermite quadrature on a cubic spline interpolation, *J. Financ. Eng.*, **1** (2014), 1450033. https://doi.org/10.1142/S2345768614500330

19. L. Goudenège, A. Molent, A. Zanette, Moving average options: Machine learning and Gauss-Hermite quadrature for a double non-Markovian problem, *Eur. J. Oper. Res.*, **303** (2022), 958–974. https://doi.org/10.1016/j.ejor.2022.03.002

20. X. C. S. Lin, D. W. C. Miao, E. E. T. Chang, Testing the closed-form spread option pricing formula based on Gauss-Hermite quadrature for a jump-diffusion model, *Comput. Econ.*, **64** (2024), 2879–2908. https://doi.org/10.1007/s10614-023-10468-2

21. B. Hambly, R. Xu, H. Yang, Recent advances in reinforcement learning in finance, *Math. Financ.*, **33** (2023), 437–503. https://doi.org/10.1111/mafi.12382

22. R. Bellman, Dynamic programming, *Science*, **153** (1966), 34–37. https://doi.org/10.1126/science.153.3731.34

23. G. Peskir, A. Shiryaev, *Optimal stopping and free-boundary problems*, Springer, 2006. https://doi.org/10.1007/978-3-7643-7390-0

24. D. Blackwell, Discounted dynamic programming, *Ann. Math. Stat.*, **36** (1965), 226–235. https://doi.org/10.1214/aoms/1177700285

25. M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*, John Wiley & Sons, 2014.

26. J. C. Cox, J. E. Ingersoll, S. A. Ross, A theory of the term structure of interest rates, *Econometrica*, **53** (1985), 385–407. https://doi.org/10.2307/1911242

27. S. E. Shreve, *Stochastic calculus for finance II: Continuous-time models*, Springer, **11** (2004). https://doi.org/10.1007/978-1-4757-4296-1

28. W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing*, Cambridge university press, 2007.

29. R. Lord, R. Koekkoek, D. V. Dijk, A comparison of biased simulation schemes for stochastic volatility models, *Quant. Financ.*, **10** (2010), 177–194. https://doi.org/10.1080/14697680802392496

30. F. N. Fritsch, R. E. Carlson, Monotone piecewise cubic interpolation, *SIAM J. Numer. Anal.*, **17** (1980), 238–246. https://doi.org/10.1137/0717021

31. H. Pham, *Continuous-time stochastic control and optimization with financial applications*, Springer Science & Business Media, **61** (2009). https://doi.org/10.1007/978-3-540-89500-8

32. D. J. Duffy, *Finite difference methods in financial engineering: A partial differential equation approach*, John Wiley & Sons, 2013. https://doi.org/10.1002/9781118673447

33. P. Wilmott, *Paul Wilmott on quantitative finance*, John Wiley & Sons, 2013.

Appendix

A. Finite-difference LCP benchmark for the Heston model

This appendix outlines the FDM-LCP method used as the numerical benchmark for the Heston model in Section 5.

A.1. Continuous formulation

Inside the continuation region, where the option is held, its value $V(S, v)$ must satisfy the stationary Heston PDE

$$\mathcal{L}V - rV = 0, \tag{A.1}$$

where $\mathcal{L}$ is the Heston infinitesimal generator

$$\mathcal{L}V := rS\, V_S + \kappa(\theta - v)V_v + \tfrac{1}{2}vS^2\, V_{SS} + \rho\xi vS\, V_{Sv} + \tfrac{1}{2}\xi^2 v\, V_{vv}. \tag{A.2}$$

The possibility of early exercise, however, transforms this into a stationary variational inequality (obstacle problem) [23, 31], which combines the PDE with the exercise payoff $\Phi(S) := (K - S)^+$:

$$\max\{\mathcal{L}V - rV, \Phi(S) - V\} = 0. \tag{A.3}$$

The problem is solved on a truncated domain $\Omega = [0, S_{\max}] \times [0, v_{\max}]$ with the following boundary conditions:

$$V(0, v) = K, \quad V(S_{\max}, v) = 0, \tag{A.4}$$

$$V_v(S, 0) = 0, \quad V_v(S, v_{\max}) = 0. \tag{A.5}$$

A.2. Discretization and LCP formulation

We employ a uniform grid (S_i, v_j) with spacings $(\Delta S, \Delta v)$ and approximate $V(S_i, v_j)$ by $V_{i,j}$. (Note: For the description of the FDM benchmark, we adopt the standard convention of using subscripts i and j for the spatial grid indices. These should not be confused with the indices i and j used for the GHQ nodes in the main text.)

The generator $\mathcal{L}$ is discretized using second-order central differences. The resulting nine-point stencil defines the discrete linear operator A. Its coefficients are derived from the local generator coefficients

$$\alpha_{i,j} = \tfrac{1}{2}v_j S_i^2, \quad \beta_{i,j} = rS_i, \quad \gamma_{i,j} = \tfrac{1}{2}\xi^2 v_j, \quad \delta_{i,j} = \kappa(\theta - v_j), \quad \eta_{i,j} = \rho\xi v_j S_i.$$

The interior row of A at (i, j) is given by

$$(A\mathbf{v})_{i,j} = a^C_{i,j} V_{i,j} + \sum_{\nu \in \{W,E,S,N,SW,SE,NW,NE\}} a^{\nu}_{i,j} V_{\nu},$$

with the nine stencil coefficients

$$a_{i,j}^{C} = r + 2\frac{\alpha_{i,j}}{\Delta S^2} + 2\frac{\gamma_{i,j}}{\Delta v^2},$$
$$a_{i,j}^{W} = -\frac{\alpha_{i,j}}{\Delta S^2} + \frac{\beta_{i,j}}{2\Delta S}, \quad a_{i,j}^{E} = -\frac{\alpha_{i,j}}{\Delta S^2} - \frac{\beta_{i,j}}{2\Delta S}, \quad a_{i,j}^{S} = -\frac{\gamma_{i,j}}{\Delta v^2} + \frac{\delta_{i,j}}{2\Delta v}, \quad a_{i,j}^{N} = -\frac{\gamma_{i,j}}{\Delta v^2} - \frac{\delta_{i,j}}{2\Delta v},$$
$$a_{i,j}^{SW} = -\frac{\eta_{i,j}}{4\Delta S\,\Delta v}, \quad a_{i,j}^{SE} = +\frac{\eta_{i,j}}{4\Delta S\,\Delta v}, \quad a_{i,j}^{NW} = +\frac{\eta_{i,j}}{4\Delta S\,\Delta v}, \quad a_{i,j}^{NE} = -\frac{\eta_{i,j}}{4\Delta S\,\Delta v}.$$

With the operator A and the payoff array $\Phi_{i,j} = (K - S_i)^+$ (broadcast over j) collected into $\mathbf{\Phi}$, the discrete counterpart to (A.3) is the linear complementarity problem [3, 32]

$$A\mathbf{v} \geq \mathbf{0}, \quad \mathbf{v} \geq \mathbf{\Phi}, \quad (\mathbf{v} - \mathbf{\Phi})^\top (A\mathbf{v}) = 0. \tag{A.6}$$

The Dirichlet conditions (A.4) are imposed strongly, while the Neumann conditions (A.5) are handled using ghost-point elimination.

A.3. Solver and implementation details

The LCP (A.6) is solved using a RB-PSOR scheme, a standard and efficient iterative method for such problems [32, 33]. The grid is colored according to $c(i, j) = (i + j) \bmod 2$, and the iteration is initialized with the payoff, $V^{(0)} = \Phi$.

For each node (i, j) in a color sweep (e.g., all red nodes, then all black nodes), the update proceeds in three steps:

1. **Gauss–Seidel update:** Compute the tentative value $\widetilde{V}_{i,j}$ using the most recent neighbor values.
2. **Over-relaxation:** Calculate the relaxed value
$$V_{i,j}^{\text{tent}} = V_{i,j}^{(k)} + \omega(\widetilde{V}_{i,j} - V_{i,j}^{(k)}).$$
3. **Projection:** Enforce the obstacle constraint
$$V_{i,j}^{\text{new}} = \max(\Phi_{i,j},\, V_{i,j}^{\text{tent}}).$$

After each half-iteration (i.e., after all nodes of a single color are updated), the boundary conditions are re-applied.

Iteration terminates when the relative change in the solution vector falls below prescribed tolerances. The stopping conditions are

$$\frac{\|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\|_\infty}{\max(1, \|\mathbf{v}^{(k+1)}\|_\infty)} \leq \varepsilon_{\text{it}}. \tag{A.7}$$

Once converged, the optimal boundary $S^*(v_j)$ for each variance slice is extracted by finding the root of $V(S, v_j) - \Phi(S) = 0$ via linear interpolation.

Implementation constants (used in Section 5). The results in Section 5 were generated using the following parameters: grid size $(N_S, N_v) = (1200, 800)$, spatial domain bounds $S_{\max} = 3.3K$, and $v_{\max} = \theta + 3\sqrt{\xi^2\theta/(2\kappa)}$. We employ uniform grids in both dimensions, so that the mesh sizes are $\Delta S = S_{\max}/(N_S - 1)$ and $\Delta v = v_{\max}/(N_v - 1)$. The RB-PSOR solver uses a relaxation parameter of $\omega = 1.4$ and stopping tolerances of $\varepsilon_{\text{it}} = 10^{-8}$.

B. Proof sketches and auxiliary tools for Section IV

This appendix provides concise proof sketches for the theoretical results presented in Section 4.

Key quantities. We work on the Banach space $(C_b(\mathcal{X}), \|\cdot\|_\infty)$ with a discount factor of $\gamma = e^{-r\Delta t}$. The core operators and error metrics, defined in Section 4, are

$\mathcal{T}$: The exact stationary Bellman operator (4.1).

$\widehat{\mathcal{T}}$: The numerical Bellman operator based on GHQ and interpolation, defined by model in (B.1)–(B.2).

$C^*, \widehat{C}^*$: The unique fixed points of $\mathcal{T}$ and $\widehat{\mathcal{T}}$, respectively.

ε : The one-step operator discrepancy, defined in (4.2).

$\Delta_C^{(k+1)}$: The Bellman residual after iteration k, defined in (4.3).

$\omega(\varepsilon)$: The separation margin, defined in (B.7).

The numerical Bellman operator $\widehat{\mathcal{T}}$. The numerical operator $\widehat{\mathcal{T}} : C_b(\mathcal{X}) \to C_b(\mathcal{X})$ is the concrete implementation of the Bellman update rule described in Section 3. Its form depends on the model.
Black–Scholes (1D). For a function $C \in C_b(\mathcal{X})$ and a state $S \in \mathcal{X}$, the operator is defined as

$$(\widehat{\mathcal{T}}C)(S) := \frac{\gamma}{\sqrt{\pi}} \sum_{i=1}^{n} w_i \max\left\{\Phi(S^{(i)}), (\widehat{\mathcal{I}}C)(S^{(i)})\right\}, \tag{B.1}$$

where $S^{(i)}$ are the future asset prices generated from S and the GHQ nodes x_i via the one-step transition in Section 3.2, and $\widehat{\mathcal{I}}$ is the interpolation operator (PCHIP with clamping).
Heston (2D). For a function $C \in C_b(\mathcal{X})$ and a state $(S, v) \in \mathcal{X}$, the operator is defined as

$$(\widehat{\mathcal{T}}C)(S, v) := \frac{\gamma}{\pi} \sum_{i,j=1}^{n} w_i w_j \max\left\{\Phi(S^{(i,j)}), (\widehat{\mathcal{I}}C)(S^{(i,j)}, v^{(i,j)})\right\}, \tag{B.2}$$

where $(S^{(i,j)}, v^{(i,j)})$ are the future states generated from (S, v) and GHQ nodes (x_i, x_j) via the one-step transition in Section 3.2, and $\widehat{\mathcal{I}}$ is the interpolation operator (PCHIP in S, linear in v, with clamping).

B.1. Auxiliary lemmas

The proofs rely on the following standard nonexpansiveness properties (i.e., they are 1-Lipschitz with respect to the sup norm):

Fact B.1 (Max operator is nonexpansive)**.** *The mapping $h \mapsto \max\{\Phi, h\}$ is nonexpansive with respect to the sup norm*

$$\|\max\{\Phi, h_1\} - \max\{\Phi, h_2\}\|_\infty \le \|h_1 - h_2\|_\infty. \tag{B.3}$$

Fact B.2 (Conditional expectation is nonexpansive)**.** *The conditional expectation operator is nonexpansive in the sup norm*

$$\|\mathbb{E}[H_1|\mathcal{F}] - \mathbb{E}[H_2|\mathcal{F}]\|_\infty \le \|H_1 - H_2\|_\infty. \tag{B.4}$$

Fact B.3 (GHQ and interpolation are nonexpansive)**.** *Since the 1D/2D GHQ weights satisfy*

$$\frac{1}{\sqrt{\pi}}\sum_{i=1}^{n} w_i = 1, \quad \frac{1}{\pi}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j = 1, \tag{B.5}$$

respectively, GHQ aggregation is a convex average and nonexpansive. The interpolation scheme (PCHIP and linear) is also designed to be nonexpansive.

Fact B.4 (Continuity of the Bellman operator)**.** *Under the Feller continuity property (Standing assumption 2), the Bellman operator $\mathcal{T}$ defined in* (4.1) *is a self-map on $C_b(\mathcal{X})$. That is, if $C \in C_b(\mathcal{X})$, then its image $\mathcal{T}C$ is also in $C_b(\mathcal{X})$.*

B.2. Proof sketch of Theorem 1 (contraction and global convergence)

Recall the Bellman operator $\mathcal{T}$ in (4.1). By Fact B.4, $\mathcal{T}$ maps bounded continuous functions to bounded continuous functions. Moreover, for any C_1, C_2,

$$\begin{aligned}\|\mathcal{T}C_1 - \mathcal{T}C_2\|_\infty &= \gamma \,\|\,\mathbb{E}\,[\max\{\Phi, C_1\} - \max\{\Phi, C_2\}\,|\,X_t = \cdot]\|_\infty \\ &\overset{\text{(B.4)}}{\leq} \gamma\,\|\max\{\Phi, C_1\} - \max\{\Phi, C_2\}\|_\infty \\ &\overset{\text{(B.3)}}{\leq} \gamma\,\|C_1 - C_2\|_\infty.\end{aligned}$$

Therefore, $\mathcal{T}$ is a γ-contraction in $\|\cdot\|_\infty$. By Banach's fixed-point theorem, it follows that there exists a unique fixed point C^*. The existence of this fixed point, along with the geometric convergence of the value iteration sequence, are foundational results in dynamic programming (see, e.g., [25]):

$$\|C^{(k)} - C^*\|_\infty \;\leq\; \gamma^k\,\|C^{(0)} - C^*\|_\infty, \qquad k = 0, 1, \ldots$$

Finally, the monotonicity property ($C^{(0)} \leq C^* \Rightarrow C^{(k)} \nearrow C^*$) also follows, since both $C \mapsto \max\{\Phi, C\}$ and conditional expectation preserve order.

B.3. Proof sketch of Theorem 2 (stability of the numerical fixed point)

The proof proceeds in two steps: We first show that $\widehat{\mathcal{T}}$ is a γ-contraction, and then we bound the distance between the fixed points.

The numerical operator $\widehat{\mathcal{T}}$ is a composition of interpolation $\widehat{\mathcal{I}}$, the max operator, and GHQ aggregation. As shown in Facts B.1 and B.3, each of these mappings is nonexpansive. Therefore, their composition is also nonexpansive. The full operator includes scaling by γ, making it a γ-contraction. Formally, for any $C_1, C_2 \in C_b(\mathcal{X})$

$$\begin{aligned}\|\widehat{\mathcal{T}}C_1 - \widehat{\mathcal{T}}C_2\|_\infty &\leq \gamma\left\|\max\{\Phi, \widehat{\mathcal{I}}C_1\} - \max\{\Phi, \widehat{\mathcal{I}}C_2\}\right\|_\infty \\ &\overset{\text{(B.3)}}{\leq} \gamma\,\|\widehat{\mathcal{I}}C_1 - \widehat{\mathcal{I}}C_2\|_\infty \\ &\overset{\textit{Fact B.3}}{\leq} \gamma\,\|C_1 - C_2\|_\infty.\end{aligned}$$

Thus, $\widehat{\mathcal{T}}$ is a γ-contraction and admits a unique fixed point $\widehat{C}^*$.

Using the triangle inequality, the definition of the operator discrepancy ε (4.2), and the fact that $\mathcal{T}$ is a γ-contraction,

$$\|\widehat{C}^* - C^*\|_\infty = \|\widehat{\mathcal{T}}\widehat{C}^* - \mathcal{T}C^*\|_\infty \le \underbrace{\|\widehat{\mathcal{T}}\widehat{C}^* - \mathcal{T}\widehat{C}^*\|_\infty}_{\le \varepsilon} + \underbrace{\|\mathcal{T}\widehat{C}^* - \mathcal{T}C^*\|_\infty}_{\le \gamma \|\widehat{C}^* - C^*\|_\infty}.$$

Rearranging the inequality yields the desired result:

$$\|\widehat{C}^* - C^*\|_\infty \ \le \ \frac{\varepsilon}{1-\gamma}. \tag{B.6}$$

B.4. Proof sketch of the corollary (convergence of the exercise boundary)

Let $g^{(k)} = C^{(k)} - \Phi$ and $g^* := C^* - \Phi$. By Theorem 1, $C^{(k)} \to C^*$ uniformly; hence, $g^{(k)} \to g^*$ uniformly. Under the boundary regularity (unique zero with strict sign change; continuous dependence on v in Heston), for any fixed v and any $\eta > 0$, there exists $\delta(\eta, v) > 0$ such that

$$|S - S^*(v)| \ge \eta \quad \Longrightarrow \quad |g^*(S, v)| \ \ge \ \delta(\eta, v).$$

Therefore, by uniform convergence, we can choose k large enough such that $|g^{(k)}(S, v) - g^*(S, v)| \le \frac{1}{2}\delta(\eta, v)$ for all S. This forces $g^{(k)}$ to have the same sign as g^* at the points $S^*(v) \pm \eta$, which means $g^{(k)}(\cdot, v)$ must have opposite signs at these two points. By the intermediate value theorem, it possesses a zero $S^{(k)}(v)$ in $(S^*(v) - \eta, S^*(v) + \eta)$. Since $\eta > 0$ is arbitrary, $S^{(k)}(v) \to S^*(v)$ for each v. Moreover, if one has a uniform lower bound $\delta(\eta, v) \ge \underline{\delta}(\eta) > 0$ for $v \in [0, v_{\max}]$, then the convergence is uniform in v.

B.5. Residual-based certificate via a separation margin

The proof relies on the geometry of the continuation premium $g^* := C^* - \Phi$. The true boundary is the zero-level set of this function, which we formalize as $\Gamma^* := \{x \in \mathcal{X} : \ g^*(x) = 0\}$. We then define the slice-wise S-distance to this set:

$$d_S(x, \Gamma^*) \ := \ \begin{cases} |S - S^*|, & \text{BS } (\, x = S\,), \\ |S - S^*(v)|, & \text{Heston } (\, x = (S, v)). \end{cases}$$

For $\varepsilon > 0$, define the separation margin

$$\omega(\varepsilon) \ := \ \inf\{\, |g^*(x)| : \ x \in \mathcal{X}, \ d_S(x, \Gamma^*) \ge \varepsilon \,\} \ > 0, \tag{B.7}$$

which holds on the compact domain under the boundary regularity (Standing assumption (iv)). Since $\mathcal{T} : C_b(\mathcal{X}) \to C_b(\mathcal{X})$ is a γ-contraction, the residual-distance inequality

$$\|C - C^*\|_\infty \ \le \ \frac{\|C - \mathcal{T}C\|_\infty}{1-\gamma} \tag{B.8}$$

yields for the iterate $C^{(k)}$

$$\|C^{(k)} - C^*\|_\infty \ \le \ \frac{\|C^{(k)} - \mathcal{T}C^{(k)}\|_\infty}{1-\gamma} = \frac{\|C^{(k)} - C^{(k+1)}\|_\infty}{1-\gamma} \overset{(4.3)}{=} \frac{\Delta_C^{(k+1)}}{1-\gamma}. \tag{B.9}$$

Therefore, if

$$\Delta_C^{(k+1)} \;\le\; (1-\gamma)\,\tfrac{1}{2}\,\omega(\varepsilon),$$

then $\|C^{(k)} - C^*\|_\infty \le \frac{1}{2}\,\omega(\varepsilon)$, which implies the zero-level set of $g^{(k)} := C^{(k)} - \Phi$ stays inside the ε-tube around Γ^*:

$$\sup_v |S^{(k)}(v) - S^*(v)| \;\le\; \varepsilon.$$

Remark (implemented operator). When iterating $\widehat{\mathcal{T}}$, the same reasoning yields $\|C^{(k)} - \widehat{C}^*\|_\infty \le \Delta_C^{(k+1)}/(1-\gamma)$ with C^* replaced by $\widehat{C}^*$. Together with (B.6), i.e., $\|\widehat{C}^* - C^*\|_\infty \le \varepsilon/(1-\gamma)$, this adds an $O(\varepsilon/(1-\gamma))$ slack to the certificate relative to the true boundary.