

# Deep Optimal Stopping and Bermudan Options Pricing

Probabilités numériques et méthodes de Monte Carlo en finance

---

**Réalisé par :**

Philippe Yao

Georges Rolland

Master 2 Probabilités et Finance

**Encadrants :**

Gilles Pagès

Vincent Lemaire

May 7, 2025

# Contents

<b>1</b>	<b>Options Bermudéennes et problème d'arrêt optimal</b>	<b>2</b>
<b>2</b>	<b>Approximation par réseaux de neurones</b>	<b>4</b>
2.1	Formulation du problème . . . . .	4
2.2	Structure des réseaux de neurones . . . . .	5
2.3	Résultat de convergence . . . . .	6
<b>3</b>	<b>Bornes de prix et intervalle de confiance</b>	<b>6</b>
3.1	Borne inférieure . . . . .	6
3.2	Borne supérieure et approche duale . . . . .	7
3.2.1	Lien avec la décomposition de Doob-Meyer . . . . .	7
3.2.2	Martingale issue de l'apprentissage . . . . .	7
3.2.3	Formulation duale de Rogers (2002) et martingales lagrangiennes . . . . .	8
3.2.4	Cas de l'option bermudéenne max call . . . . .	9
3.3	Intervalles de confiance et estimation ponctuelle . . . . .	11
<b>4</b>	<b>Analyse de robustesse et réduction de variance</b>	<b>12</b>
4.1	Suivi de l'apprentissage : évolution de la loss . . . . .	12
4.2	Impact de la profondeur du réseau sur la précision de la borne inférieure . . . . .	13
4.3	Impact du nombre d'étapes de discrétisation $N$ et de $S_0$ . . . . .	14
4.4	Effet des techniques de réduction de variance . . . . .	15
<b>5</b>	<b>Conclusion</b>	<b>16</b>

## Introduction

Ce projet porte sur la valorisation d'options Bermudéennes, qui donnent au détenteur la possibilité d'exercer à des dates discrètes prédéfinies. Leur évaluation repose sur un problème d'arrêt optimal, souvent difficile à résoudre numériquement en grande dimension. Nous nous appuyons sur deux articles complémentaires pour approcher ce problème :

- **Becker, Cheridito, Jentzen (2019)** introduisent une méthode d'apprentissage par réseau de neurones pour approximer la règle d'exercice optimale, fournissant ainsi une **borne inférieure** du prix.
- **Rogers (2002)** propose une formulation duale permettant la construction d'une **borne supérieure** à l'aide de martingales surmajorantes.

L'objectif du projet est d'implémenter ces deux estimateurs pour des options Bermudéennes, d'analyser leur robustesse numérique, ainsi que d'évaluer l'impact de techniques de réduction de variance sur la précision des bornes obtenues.

## 1 Options Bermudéennes et problème d'arrêt optimal

Les options Bermudéennes constituent une classe intermédiaire entre les options européennes (exerçables uniquement à maturité) et les options américaines (exerçables à tout moment avant maturité). Elles ne peuvent être exercées qu'à certaines dates discrètes prédéterminées  $0 = t_0 < t_1 < \dots < t_N = T$ . À chaque date d'exercice  $t_n$ , l'investisseur peut décider soit de continuer, soit d'exercer immédiatement.

### Modélisation et cadre probabiliste

Soit  $(X_n)_{n=0,\dots,N}$  un processus de Markov à valeurs dans  $R^d$  sur un espace de probabilité  $(\Omega, \mathcal{F}, P)$ , avec  $X_n \in R^d$  représentant l'état du système (par exemple, le prix du sous-jacent) à l'instant  $t_n$ . Le problème d'arrêt optimal consiste à déterminer une variable d'arrêt  $\tau$  prenant des valeurs dans  $\{0, 1, \dots, N\}$  qui maximise l'espérance d'une récompense :

$$\sup_{\tau \in \mathcal{T}} E[g(\tau, X_\tau)]$$

où  $g(n, x)$  est la fonction de gain, et  $\mathcal{T}$  est l'ensemble des temps d'arrêt admissibles.

Pour garantir l'existence d'une solution optimale, on suppose que la fonction de gain  $g$  satisfait les conditions d'intégrabilité suivantes :

$$E[|g(n, X_n)|] < \infty, \quad \forall n = 0, \dots, N$$

et pour pouvoir construire un intervalle de confiance autour de la valeur optimale, une hypothèse plus forte est considérée :

$$E[g(n, X_n)^2] < \infty, \quad \forall n = 0, \dots, N$$

## Formulation du problème d'arrêt

Le processus décisionnel est exprimé à travers une suite de fonctions  $f_n : R^d \rightarrow \{0, 1\}$ , indiquant s'il est optimal d'arrêter au temps  $n$  (valeur 1) ou de continuer (valeur 0). Une politique d'arrêt peut alors être représentée par la formule suivante :

$$\tau = \sum_{n=1}^N n f_n(X_n) \prod_{j=0}^{n-1} (1 - f_j(X_j))$$

Cette expression permet de définir le fait que l'arrêt intervient à la première date  $n$  telle que  $f_n(X_n) = 1$ , i.e., lorsque la politique d'arrêt déclenche une exécution.

La stratégie optimale est construite rétroactivement en partant de  $f_N \equiv 1$  (exercice obligatoire à maturité) : à chaque date  $n$ , la décision d'exercer ou de continuer est obtenue en comparant la valeur d'exercice immédiat et la valeur espérée de continuation, conformément au principe d'optimalité.

## Auxiliaires et valeur optimale

Pour chaque  $n = 0, \dots, N$ , on introduit les problèmes d'arrêt auxiliaires :

$$V_n = \sup_{\tau \in \mathcal{T}_n} E[g(\tau, X_\tau)]$$

où  $\mathcal{T}_n$  désigne les temps d'arrêt postérieurs ou égaux à  $n$ . Grâce à un raisonnement récursif, le théorème principal montre qu'il existe une politique optimale  $(f_n)_n$  telle que  $f_N \equiv 1$  (exercice obligatoire à maturité), et que la stratégie optimale est déterminée rétroactivement via l'inégalité :

$$E[g(\tau_n, X_{\tau_n})] \geq V_n - (V_{n+1} - E[g(\tau_{n+1}, X_{\tau_{n+1}})])$$

Ce cadre mathématique est la base de la méthode d'apprentissage proposée dans l'article [?], exploitée dans la suite du projet via un entraînement neuronal.

## Modèle de marché

Dans le cadre de la valorisation d'options Bermudéennes, on considère un modèle multidimensionnel de Black-Scholes discrétisé.

On suppose qu'il existe  $d$  actifs sous-jacents, dont les prix  $(S_n^1, \dots, S_n^d)$  évoluent selon une dynamique log-normale couplée. Plus précisément, sous la probabilité risque-neutre  $Q$ , les processus  $(S_n^i)_{n=0, \dots, N}$  suivent :

$$S_{n+1}^i = S_n^i \cdot \exp \left( \left( r - \delta_i - \frac{1}{2} \sigma_i^2 \right) \Delta t + \sigma_i \sqrt{\Delta t} \cdot Z_{n+1}^i \right), \quad i = 1, \dots, d$$

où :

- $r$  est le taux d'intérêt sans risque,
- $\delta_i$  est le taux de dividende continu de l'actif  $i$ ,
- $\sigma_i$  est la volatilité de l'actif  $i$ ,
- $(Z_n^1, \dots, Z_n^d)$  sont des vecteurs de variables gaussiennes normales centrées réduites, corrélées selon une matrice de corrélation  $\rho$ .

Les vecteurs  $(Z_n)_{n=1, \dots, N}$  sont supposés i.i.d., avec :

$$E[Z_n] = 0, \quad \text{et} \quad \text{Var}(Z_n) = 1$$

où  $\rho$  est une matrice symétrique définie positive de taille  $d \times d$ . Cependant dans la suite, pour des raisons de simplification, nous supposerons dans nos études des actifs décorrelés ( $\rho = 0$ )

La simulation des trajectoires  $(S_n)_{n=0, \dots, N}$  repose sur cette discrétisation du processus, avec  $\Delta t = T/N$  l'intervalle de temps constant entre deux dates d'exercice.

Ce modèle constitue la base sur laquelle sont entraînés les réseaux neuronaux pour approximer la politique d'arrêt optimale, ainsi que pour estimer les bornes de valorisation duales.

## 2 Approximation par réseaux de neurones

Pour résoudre le problème d'arrêt optimal associé aux options Bermudéennes, l'approche *Deep Optimal Stopping* repose sur une approximation récursive des décisions d'arrêt  $\{f_n\}_{n=0}^{N-1}$  via des réseaux de neurones.

### 2.1 Formulation du problème

On cherche à approximer l'expression suivante :

$$\sup_{\tau \in \mathcal{T}} E[g(\tau, X_\tau)],$$

où  $\mathcal{T}$  désigne l'ensemble des temps d'arrêt  $\mathcal{F}_n$ -adaptés et  $g(n, X_n)$  est une fonction mesurable (généralement le payoff actualisé au temps  $n$ ).

Pour rendre le problème accessible à l'optimisation par gradient, une version intermédiaire  $F^\theta : R^d \rightarrow (0, 1)$  est introduite. Il s'agit d'un réseau de neurones produisant une probabilité d'arrêt continue, utilisée pour

approximer la décision binaire. La version "dure"  $f_n^\theta$  sera ensuite déduite par seuillage.

Le temps d'arrêt optimal est exprimé à l'aide d'une suite de fonctions d'arrêt  $\{f_n^\theta\}_{n=0}^{N-1}$  telles que

$$\tau = \sum_{n=1}^N n f_n^\theta(X_n) \prod_{j=0}^{n-1} (1 - f_j^\theta(X_j)),$$

avec  $f_N \equiv 1$ . La stratégie consiste à approximer les fonctions  $f_n^\theta$  via des réseaux de neurones entraînés de manière récursive, en backward induction, du pas  $n = N - 1$  vers  $n = 0$ .

## 2.2 Structure des réseaux de neurones

Chaque  $f_n^\theta$  est approché par un réseau de neurones de la forme :

$$F^\theta = \psi \circ a_I^\theta \circ \varphi_{q_{I-1}} \circ a_{I-1}^\theta \circ \dots \circ \varphi_{q_1} \circ a_1^\theta,$$

où :

- $a_i^\theta(x) = A_i x + b_i$  est une couche affine ;
- $\varphi_j(x_1, \dots, x_j) = (x_1^+, \dots, x_j^+)$  est la fonction ReLU appliquée composante par composante ;
- $\psi(x) = \frac{e^x}{1+e^x}$  est la fonction logistique qui permet d'interpréter la sortie comme une probabilité d'arrêt.

La dimension  $q$  de l'espace des paramètres  $\theta$  dépend de l'architecture du réseau. Pour une profondeur  $I \geq 2$  avec des tailles de couches cachées  $q_1, \dots, q_{I-1}$ , on a :

$$q = 1 + q_1 + \dots + q_{I-1} + dq_1 + q_1 q_2 + \dots + q_{I-2} q_{I-1} + q_{I-1},$$

ce qui correspond au nombre total de poids et biais des couches affines.

La version "dure" de la décision est alors définie par :

$$f_n^\theta(x) = 1_{[0, \frac{1}{2}]}(F^\theta(x)),$$

c'est-à-dire une décision binaire d'arrêt si la probabilité de sortie est supérieure à 1/2.

## Optimisation des paramètres

Les paramètres  $\theta_n$  du réseau  $f_n^\theta$  sont obtenus en maximisant une espérance de type :

$$E[g(n, X_n) F^\theta(X_n) + g(\tau_{n+1}, X_{\tau_{n+1}})(1 - F^\theta(X_n))],$$

où  $\tau_{n+1}$  est la stratégie d'arrêt définie à partir des réseaux entraînés aux pas ultérieurs  $\{f_m^\theta\}_{m=n+1}^N$  et  $F^\theta$  est la version souple du réseau de neurones produisant une probabilité d'arrêt.

Cette optimisation est réalisée à l'aide de la méthode du gradient stochastique sur un échantillon de trajectoires simulées  $\{X_n^k\}_{n=0}^N$ , avec les techniques suivantes : Xavier initialization (Glorot et Bengio, 2010), Batch Normalization (Ioffe et Szegedy, 2015), et l'algorithme d'Adam pour la mise à jour des poids du réseau (Kingma et Ba, 2015).

## 2.3 Résultat de convergence

Le résultat de convergence fondamental de [?] établit que les réseaux de neurones utilisés dans cette méthode sont suffisamment expressifs pour approximer toute fonction d'arrêt mesurable avec une erreur arbitrairement petite.

Plus précisément, pour tout pas de temps  $n \in \{0, 1, \dots, N-1\}$  et tout temps d'arrêt  $\tau_{n+1} \in \mathcal{T}_{n+1}$  fixé, il existe pour toute profondeur  $I \geq 2$  et tout  $\varepsilon > 0$  une architecture de réseau (définie par des entiers  $q_1, \dots, q_{I-1}$ ) telle que :

$$\begin{aligned} \sup_{\theta \in R^q} E \left[ g(n, X_n) F^\theta(X_n) + g(\tau_{n+1}, X_{\tau_{n+1}}) (1 - F^\theta(X_n)) \right] \\ \geq \sup_{f \in \mathcal{D}} E \left[ g(n, X_n) f(X_n) + g(\tau_{n+1}, X_{\tau_{n+1}}) (1 - f(X_n)) \right] - \varepsilon \end{aligned}$$

où  $\mathcal{D}$  désigne l'ensemble des fonctions mesurables  $f : R^d \rightarrow \{0, 1\}$ .

Cette inégalité montre que l'approximation réalisée par réseau de neurones est universelle pour ce type de problème d'arrêt optimal. Cette capacité d'approximation permet alors de construire efficacement des stratégies d'exercice sous-optimales, à partir desquelles nous pourrions dériver des estimations précises des bornes inférieure et supérieure du prix de l'option.

## 3 Bornes de prix et intervalle de confiance

### 3.1 Borne inférieure

Une fois les décisions d'arrêt approximativement optimales  $f_n^\theta$  apprises via les réseaux de neurones, on peut obtenir une estimation de borne inférieure  $L$  pour la valeur optimale  $V_0 = \sup_{\tau \in \mathcal{T}} E[g(\tau, X_\tau)]$ .

Pour cela, on génère un nouvel échantillon indépendant de trajectoires  $(y_n^k)_{n=0}^N$  pour  $k = 1, \dots, K_L$ , sur lesquelles on applique la règle d'arrêt  $\tau^\Theta$  associée aux réseaux entraînés. Le temps d'arrêt  $\tau^\Theta$  est une

fonction mesurable de la trajectoire complète :

$$\tau^\Theta = \sum_{n=1}^N n f_n^\theta(y_n^k) \prod_{j=0}^{n-1} (1 - f_j^\theta(y_j^k)).$$

L'estimateur Monte Carlo de la borne inférieure est alors donné par :

$$\hat{L} = \frac{1}{K_L} \sum_{k=1}^{K_L} g(\tau_k^\Theta, y_{\tau_k^\Theta}^k),$$

où  $g(n, x)$  désigne la récompense reçue si l'on exerce à l'instant  $n$  en état  $x$ .

Cet estimateur  $\hat{L}$  est un estimateur sans biais de la borne  $L$ , et converge vers  $L$  lorsque  $K_L \rightarrow \infty$  en vertu de la loi forte des grands nombres.

La borne inférieure ainsi obtenue représente la performance moyenne de la politique d'arrêt apprise. Elle constitue un indicateur fondamental pour juger de la qualité de l'approximation apportée par l'architecture neuronale.

## 3.2 Borne supérieure et approche duale

### 3.2.1 Lien avec la décomposition de Doob-Meyer

Dans le cadre du problème d'arrêt optimal, la **valeur optimale**  $V_n$  peut être exprimée comme :

$$V_n = \text{ess sup}_{\tau \in \mathcal{T}_n} E[g(\tau, X_\tau) \mid \mathcal{F}_n],$$

où  $(\mathcal{F}_n)$  est la filtration naturelle.

La suite  $(V_n)_{n=0, \dots, N}$  forme une **surmartingale**, et par la **décomposition de Doob-Meyer**, il existe une martingale  $(M_n^H)$  et un processus croissant  $(A_n^H)$  tels que :

$$V_n = M_n^H - A_n^H.$$

La borne supérieure optimale est atteinte en utilisant cette martingale  $(M_n^H)$ . En pratique, cependant, cette martingale théorique est souvent inconnue : l'approche duale consiste donc à **approximer** une bonne martingale  $(M_n^\Theta)$  pour construire une borne supérieure praticable.

### 3.2.2 Martingale issue de l'apprentissage

À partir des réseaux de neurones ayant approximé les décisions d'arrêt  $\{f_n^\theta\}_{n=0}^{N-1}$ , il est possible de construire une martingale candidate  $(M_n^\Theta)_{n=0, \dots, N}$ . Elle repose sur la continuation estimée après chaque étape :



$$C_n^\Theta = E[g(\tau_{n+1}^\Theta, X_{\tau_{n+1}^\Theta}) \mid \mathcal{F}_n],$$

et les incréments sont donnés par :

$$\Delta M_n^k = f^{\theta^n}(z_n^k)g(n, z_n^k) + (1 - f^{\theta^n}(z_n^k))C_n^k - C_{n-1}^k,$$

où  $z_n^k$  désigne une trajectoire simulée et  $C_n^k$  est estimé via Monte Carlo conditionnel.

La martingale approchée est alors reconstruite par sommation :

$$M_n^k = \sum_{m=1}^n \Delta M_m^k, \quad M_0^k = 0.$$

Cependant, l'estimation de la martingale issue directement de l'approximation du stopping optimal est plus coûteuse en calculs et elle hérite des erreurs d'approximation des réseaux de neurones, entraînant une propagation de l'erreur sur toute la trajectoire. L'approche proposée par Rogers permet au contraire de construire des martingales lagrangiennes qui ne sont les plus optimales mais qui sont plus simples à implémenter, offrant des bornes supérieures raisonnables à un coût numérique bien plus faible.

### 3.2.3 Formulation duale de Rogers (2002) et martingales lagrangiennes

L'approche de Rogers reformule la borne supérieure comme un problème de minimisation convexe sur l'ensemble des martingales adaptées nulles à l'origine. La valeur optimale  $V_0$  satisfait :

$$V_0 = \inf_{M \in \mathcal{M}} E \left[ \sup_{0 \leq n \leq N} (g(n, X_n) - M_n) \right],$$

où  $\mathcal{M}$  désigne l'ensemble des  $\mathcal{F}_n$ -martingales vérifiant  $M_0 = 0$ .

L'idée pratique est de construire une famille de martingales candidates  $(M^{(i)})_{i=1}^L$  relativement simples, puis de chercher la meilleure combinaison linéaire :

$$M^\lambda = \sum_{i=1}^L \lambda_i M^{(i)}, \quad \text{avec } \lambda = (\lambda_1, \dots, \lambda_L)^T.$$

Le vecteur  $\lambda$  est optimisé numériquement pour minimiser :

$$E \left[ \sup_{0 \leq n \leq N} (g(n, X_n) - M_n^\lambda) \right].$$

L'optimisation est effectuée en deux étapes :

- D'abord sur un nombre modéré de trajectoires (de l'ordre de quelques centaines) pour déterminer  $\lambda^*$ .

- Puis sur un grand échantillon de trajectoires (typiquement  $10^4$  ou plus) pour estimer de manière précise la borne finale.

Notons que la phase d'optimisation de  $\lambda^*$  constitue la partie la plus coûteuse du processus. Toutefois, Rogers souligne que la précision de l'optimisation n'est pas cruciale : dans le cadre d'un problème convexe avec une fonction différentiable au minimum, de légères erreurs sur la valeur optimale de  $\lambda^*$  entraînent des variations négligeables sur la borne supérieure estimée. Aucun effort spécifique n'a été entrepris pour accélérer cette étape, mais il serait intéressant d'explorer si un nombre plus réduit de trajectoires, ou un meilleur choix initial, permettrait d'obtenir des résultats comparables à moindre coût.

Enfin, pour améliorer encore l'estimation sans doubler le coût de simulation, une extrapolation de Richardson est proposée. Elle repose sur le fait de générer des trajectoires fines avec  $2n$  dates, puis de les sous-échantillonner pour obtenir une version plus grossière avec  $n$  dates. On calcule ensuite la borne supérieure sur ces deux grilles, et on effectue une extrapolation pour annuler le biais d'erreur principal dû à la discrétisation. Cette méthode permet de corriger efficacement l'erreur sans accroître significativement le coût numérique.

### 3.2.4 Cas de l'option bermudéenne max call

On considère une option Bermudéenne de type call sur le maximum de  $d$  actifs, avec une échéance finale  $T$  et un prix d'exercice  $K$ . Son payoff (fonction de gain) à une date  $t$  quelconque est donné par :

$$H(S_t) = \left( \max_{1 \leq i \leq d} S_i(t) - K \right)^+$$

L'option peut être exercée aux dates discrètes  $0 = t_0 < t_1 < \dots < t_N = T$ .

Pour appliquer l'approche duale de Rogers (2002) à ce produit, on cherche une martingale  $M_t$  (sous la mesure risque-neutre, avec  $M_0 = 0$ ) dont les écarts  $M_{t_{k+1}} - M_{t_k}$  sont choisis astucieusement à chaque intervalle  $[t_k, t_{k+1}]$ .

On décrit ci-dessous deux processus martingales candidates  $M^{(1)}$  et  $M^{(2)}$  spécifiquement conçus pour le max-call multi-actifs, puis on expliquera comment les combiner linéairement pour améliorer la borne duale.

**Martingale candidate 1 :** valeur actualisée de l'option d'échange entre les deux meilleurs actifs.

À chaque date  $t_k$ , on identifie les deux actifs les plus chers, notés  $S_k^{(1)}$  et  $S_k^{(2)}$ , avec  $S_k^{(1)} = \max_i S_i(t_k)$  et  $S_k^{(2)}$  le second maximum. On considère une **option européenne d'échange** à maturité  $T$  permettant

d'échanger  $S^{(2)}$  contre  $S^{(1)}$ . Son prix est donné par la formule de Margrabe :

$$\text{Ex}(t_k) = S_k^{(1)} e^{-\delta_1(T-t_k)} \Phi(d_1) - S_k^{(2)} e^{-\delta_2(T-t_k)} \Phi(d_2),$$

avec

$$d_1 = \frac{\log(S_k^{(1)}/S_k^{(2)}) + (q_2 - q_1 + \frac{1}{2}\sigma^2)(T - t_k)}{\sigma\sqrt{T - t_k}}, \quad d_2 = d_1 - \sigma\sqrt{T - t_k},$$

et  $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$  la variance relative.

Le processus martingale  $M^{(1)}$  est construit comme la somme des accroissements de cette valeur actualisée, activée uniquement lorsque l'option est dans la monnaie :

$$\Delta M_{t_k}^{(1)} = \begin{cases} \text{Ex}(t_k) - \text{Ex}(t_{k-1}) & \text{si } \max_i S_i(t_k) > K, \\ 0 & \text{sinon.} \end{cases}$$

**Martingale candidate 2** : Valeur actualisée du call européen sur l'actif leader.

À chaque date  $t_k$ , on identifie l'actif leader  $i_k^*$  tel que  $S_{i_k^*}(t_k) = \max_i S_i(t_k)$ . On considère alors une **option call européenne** sur  $S_{i_k^*}$  de maturité  $T$  et de strike  $K$ , évaluée via la formule de Black-Scholes avec dividende :

$$C_{i_k^*}(t_k) = S_{i_k^*}(t_k) e^{-q_{i_k^*}(T-t_k)} \Phi(d_1) - K e^{-r(T-t_k)} \Phi(d_2),$$

avec

$$d_1 = \frac{\log\left(\frac{S_{i_k^*}(t_k)}{K}\right) + \left(r - q_{i_k^*} + \frac{1}{2}\sigma_{i_k^*}^2\right)(T - t_k)}{\sigma_{i_k^*}\sqrt{T - t_k}}, \quad d_2 = d_1 - \sigma_{i_k^*}\sqrt{T - t_k}.$$

On définit alors une martingale discrète  $M^{(2)}$  à partir des variations du prix de cette option européenne, mais activée uniquement lorsque l'actif leader dépasse le strike :

$$\Delta M_{t_k}^{(2)} = \begin{cases} e^{-r\Delta t} C_{i_k^*}(t_{k+1}) - C_{i_k^*}(t_k), & \text{si } S_{i_k^*}(t_k) > K, \\ 0, & \text{sinon.} \end{cases}$$

avec  $M_{t_0}^{(2)} = 0$ .

Cette condition d'activation reflète l'idée que le call sur l'actif leader n'entre réellement en jeu que lorsque cet actif devient "exerçable", c'est-à-dire in-the-money. Ce choix permet de réduire le bruit inutile dans les incréments et de concentrer la martingale sur les zones pertinentes du payoff.

### Combinaison linéaire optimisée et estimation finale de la borne supérieure

Les deux martingales candidates précédentes capturent des comportements complémentaires du payoff. Étant toutes deux des martingales nulles en 0, toute combinaison linéaire

$$M_t = \alpha M_t^{(1)} + \beta M_t^{(2)}$$

reste une martingale valide.

Dans la pratique, on ajuste les coefficients  $(\alpha, \beta)$  *a posteriori* par descente de gradient, en minimisant la borne supérieure estimée :

$$U(\alpha, \beta) = E \left[ \max_{0 \leq k \leq N} \left( e^{-rt_k} g(S_{t_k}) - M_{t_k} \right) \right].$$

Après optimisation, l'estimateur final de la borne supérieure est donné par :

$$\hat{U} = \frac{1}{K_U} \sum_{k=1}^{K_U} \sup_{0 \leq n \leq N} \left( g(n, z_n^k) - M_n^{\alpha^*, \beta^*}(z^k) \right),$$

où  $(\alpha^*, \beta^*)$  sont les coefficients optimaux trouvés.

Cette approche permet d'obtenir une borne duale resserrée, tout en conservant la garantie théorique d'encadrement du prix optimal sans dépendre directement de la précision du réseau de neurones initial.

### 3.3 Intervalles de confiance et estimation ponctuelle

Une fois les bornes inférieure  $\hat{L}$  et supérieure  $\hat{U}$  estimées, une estimation ponctuelle naturelle de la valeur optimale  $V_0$  est donnée par la moyenne :

$$\hat{V}_0 = \frac{\hat{L} + \hat{U}}{2}.$$

Pour évaluer la précision de ces bornes, on utilise un théorème central limite appliqué aux estimateurs Monte Carlo. Sous des hypothèses standards d'intégrabilité quadratique sur le payoff  $g(n, X_n)$ , les estimateurs  $\hat{L}$  et  $\hat{U}$  sont asymptotiquement normaux lorsque le nombre de trajectoires  $K_L, K_U$  tend vers l'infini.

Les variances asymptotiques sont estimées par :

$$\hat{\sigma}_L^2 = \frac{1}{K_L - 1} \sum_{k=1}^{K_L} \left( g(\tau^k, y_{\tau^k}^k) - \hat{L} \right)^2, \quad \hat{\sigma}_U^2 = \frac{1}{K_U - 1} \sum_{k=1}^{K_U} \left( \sup_{0 \leq n \leq N} \left( g(n, z_n^k) - M_n^{\alpha^*, \beta^*}(z^k) \right) - \hat{U} \right)^2,$$

où  $\tau^k$  est le temps d'arrêt proposé par les réseaux de neurones sur le chemin  $y^k$ , et  $z^k$  désigne un chemin indépendant pour l'évaluation de la borne supérieure.

Ainsi, pour un niveau de confiance  $1 - \alpha$ , on construit les intervalles asymptotiques :

$$\left[ \hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}, \infty \right) \quad \text{et} \quad \left( -\infty, \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}} \right],$$

où  $z_{\alpha/2}$  est le quantile d'ordre  $1 - \alpha/2$  de la loi normale standard.

Un intervalle de confiance valide pour  $V_0$  est alors donné par :

$$\left[ \hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}, \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}} \right],$$

avec une couverture asymptotique d'au moins  $1 - \alpha$ .

Cette procédure fournit une évaluation robuste de la valeur de l'option Bermudéenne ainsi qu'une quantification explicite de l'incertitude liée à l'approximation numérique.

## 4 Analyse de robustesse et réduction de variance

### 4.1 Suivi de l'apprentissage : évolution de la loss

Pour un ensemble fixe de paramètres du marché et d'entraînement :

- Taux d'intérêt :  $r = 0.05$
- Volatilité :  $\sigma = 0.2$
- Taux de dividende :  $\delta = 0.1$
- Prix initial des actifs :  $S_0 = 100$
- Strike :  $K = 100$
- Corrélation entre actifs :  $\rho = 0$
- Maturité :  $T = 3$  ans
- Nombre de dates d'exercice :  $N = 9$
- Nombre d'actifs :  $d = 5$
- Nombre de trajectoires d'entraînement :  $M = 8192$
- Taux d'apprentissage (learning rate =  $10^{-3}$ )
- Nombre de couches cachées : 2

nous représentons ci-dessous l'évolution de la fonction de perte (*loss*) au fil des epochs, pour chaque date d'exercice.

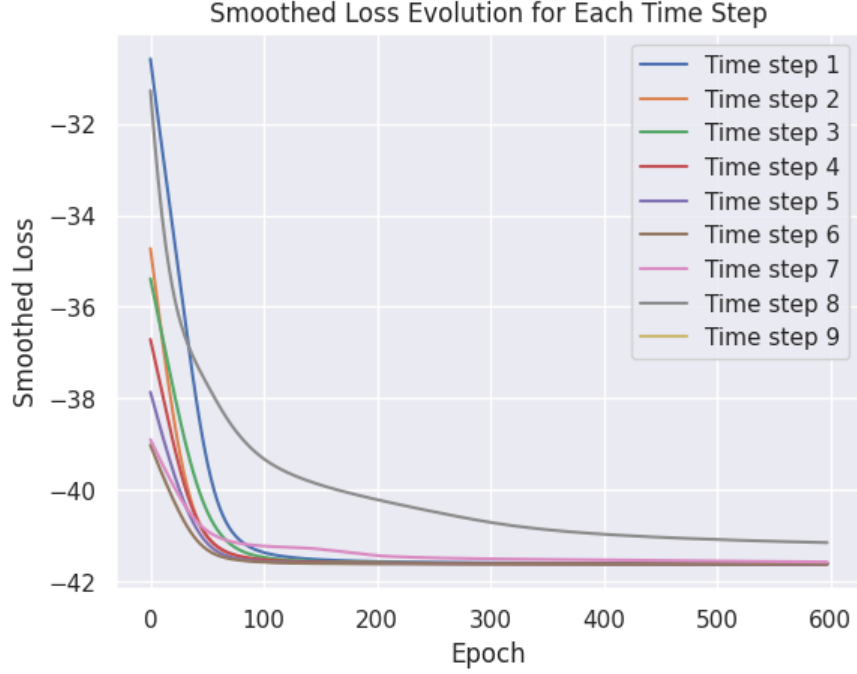


Figure 1: Évolution lissée de la loss par epoch pour chaque étape de temps ( $t_1$  à  $t_9$ ).

La figure 1 montre une décroissance régulière de la fonction de perte pour chaque pas de temps, jusqu'à atteindre un plateau indiquant la convergence du réseau. On observe que les dates proches de l'échéance (typiquement  $t_8$  ou  $t_9$ ) présentent une loss plus élevée au début, traduisant une difficulté initiale à approximer le gain marginal proche de maturité. Cela s'atténue rapidement avec l'entraînement.

Cette visualisation valide empiriquement le bon fonctionnement du schéma d'entraînement backward, avec un comportement cohérent de la dynamique d'optimisation sur l'ensemble des pas de temps.

#### 4.2 Impact de la profondeur du réseau sur la précision de la borne inférieure

Pour cette étude, on utilise toujours les mêmes paramètres  $S_0 = 100$ ,  $K = 100$ ,  $\rho = 0$ ,  $\sigma = 0.2$ ,  $d = 5$ ,  $T = 3$  ans et  $N = 9$ . On observe les valeurs de la borne inférieure ainsi que sa variance pour différentes valeurs de la profondeur du réseau  $I$ .

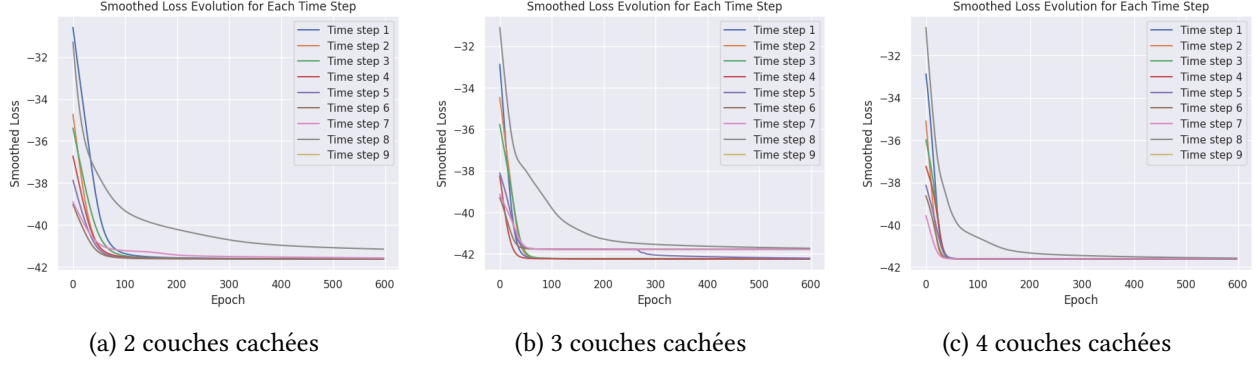


Figure 2: Évolution de la loss selon la profondeur du réseau

Table 1: Borne inférieure en fonction de la profondeur

$I$	$\hat{L}$	$\sigma_L^2$
2	26.31	656.67
3	26.57	624.23
4	26.51	651.51

On observe que la profondeur  $I = 3$  fournit la meilleure estimation : la borne inférieure  $\hat{L}$  est la plus élevée (26.57) et la variance associée  $\hat{\sigma}_L^2$  est la plus faible (624.23). Cela suggère qu'un réseau avec trois couches cachées constitue un bon compromis entre expressivité et stabilité pour ce problème, par rapport aux architectures plus simples ( $I = 2$ ) ou plus profondes ( $I = 4$ ). On retient donc un réseau à 3 couches cachées pour la suite de nos travaux.

### 4.3 Impact du nombre d'étapes de discrétisation $N$ et de $S_0$

Nous étudions ensuite l'influence du nombre d'actifs  $d$  et du niveau initial  $S_0$  sur les estimations de prix et la qualité des bornes inférieure et supérieure.

Pour chaque couple  $(N, S_0)$ , nous reportons :

- La borne inférieure estimée  $\hat{L}$ ,
- La borne supérieure estimée  $\hat{U}$ ,
- Le point estimate  $\hat{V}_0 = (\hat{L} + \hat{U})/2$ ,
- L'intervalle de confiance à 95% associé.

Table 2: Estimation des bornes de prix selon  $N$  et  $S_0$  (avec  $S_0 = 100$ ,  $r = 0,05$ ,  $\rho = 0$ ,  $\sigma = 0,2$ ,  $\delta = 0,1$ ,  $d = 5$ )

$N$	$S_0$	$\hat{L}$	$\hat{U}$	$\hat{V}_0$	95% CI
6	100	26.35	26.65	26.50	[26.33, 26.67]
6	110	37.41	38.24	37.83	[37.39, 38.26]
6	120	49.11	50.56	49.83	[49.08, 50.58]
9	100	26.31	27.83	27.07	[26.28, 27.85]
9	110	37.41	39.41	38.41	[37.39, 39.43]
9	120	49.28	51.86	50.57	[49.27, 51.86]

Ce tableau présente les estimations des bornes de prix pour une option Bermudéenne max-call, en fonction du nombre de dates d'exercice  $N$  et du niveau initial  $S_0$ , dans un cadre à  $d = 5$  actifs indépendants.

- On observe que lorsque  $S_0$  augmente, les bornes inférieure  $\hat{L}$  et supérieure  $\hat{U}$  augmentent également, ce qui est cohérent avec l'intuition : plus l'actif est initialement élevé, plus l'option a de chances d'être dans la monnaie.
- Dans l'ensemble, les estimations sont cohérentes. Toutefois, dans certains cas, la borne supérieure  $\hat{U}$  s'éloigne davantage de la borne inférieure  $\hat{L}$ , indiquant une approximation potentiellement moins précise, ce qui est attendu car les martingales utilisées pour l'optimisation ne sont pas optimales mais servent uniquement à obtenir des résultats raisonnables à moindre coût computationnel.

#### 4.4 Effet des techniques de réduction de variance

Dans cette section, nous analysons l'impact de techniques de réduction de variance sur l'estimateur de la borne supérieure  $\hat{U}$ , en conservant les mêmes paramètres de marché que dans les sections précédentes. Bien que ces méthodes soient également applicables à la borne inférieure  $\hat{L}$ , nous concentrons notre étude sur  $\hat{U}$ , car l'estimateur dual implique un maximum sur les dates d'exercice, ce qui le rend souvent plus instable numériquement.

Nous mettons en œuvre la **méthode de la variable de contrôle**, qui consiste à introduire une variable auxiliaire corrélée avec l'estimateur initial, dans le but de compenser sa variance. Plus précisément, nous utilisons comme variable de contrôle le *payoff européen* associé à l'option max-call, calculé à l'échéance  $T$  :

$$\text{Control} = \left( \max_{1 \leq i \leq d} S_i(T) - K \right)^+.$$

L'idée est d'ajuster l'estimateur brut  $\hat{U}$  en y soustrayant un multiple optimal (au sens de la réduction de



variance) d'une variable auxiliaire centrée :

$$\hat{U}_{\text{corrigée}} = \hat{U} - \beta \cdot (Y - E[Y]),$$

où  $Y$  désigne la variable de contrôle, ici le payoff européen, et  $\beta$  est un coefficient de régression calculé de manière empirique. Plus précisément,  $\beta$  est choisi pour minimiser la variance de l'estimateur corrigé. Cela conduit à l'expression analytique suivante :

$$\beta^* = \frac{\text{Cov}(\hat{U}, Y)}{\text{Var}(Y)},$$

qui est estimée sur l'échantillon de simulation par :

$$\hat{\beta} = \frac{\frac{1}{M} \sum_{i=1}^M (U_i - \bar{U})(Y_i - \bar{Y})}{\frac{1}{M} \sum_{i=1}^M (Y_i - \bar{Y})^2},$$

avec  $(U_i)$  les réalisations de l'estimateur dual sur chaque trajectoire et  $(Y_i)$  les valeurs correspondantes de la variable de contrôle. Ce choix garantit que la variance de  $\hat{U}_{\text{corrigée}}$  est minimisée parmi toutes les combinaisons linéaires possibles entre  $\hat{U}$  et  $Y$ .

Cette méthode s'avère particulièrement efficace dans notre cadre, car le payoff européen est naturellement corrélé au payoff Bermudéen maximal, tout en étant facile à simuler et à centrer. Nous montrons numériquement qu'elle permet une réduction significative de la variance de l'estimateur dual, tout en conservant un biais nul.

Table 3: Impact des techniques de réduction de variance sur l'estimation de la borne inférieure

Méthode	Estimation $\hat{U}$	Var( $\hat{U}$ )
Sans réduction	27.8825	233.012787
variable de contrôle	27.8825	130.405182

La table ci-dessus met en évidence l'effet de la méthode de variable de contrôle sur l'estimation de la borne supérieure  $\hat{U}$ . Si la valeur de l'estimateur reste inchangée, la variance est réduite de plus de 44%, ce qui confirme l'efficacité de la correction pour stabiliser l'estimation sans introduire de biais. Cette amélioration est d'autant plus intéressante qu'elle repose sur une variable facile à simuler et naturellement corrélée au payoff.

## 5 Conclusion

Dans ce projet, nous avons étudié la valorisation des options Bermudéennes via une approche d'arrêt optimal, en combinant deux méthodes complémentaires issues de la littérature récente : une borne inférieure construite par apprentissage profond, et une borne supérieure obtenue par la formulation duale de Rogers.

L'approche neuronale permet d'approximer efficacement la politique d'arrêt, tandis que la méthode duale assure un encadrement rigoureux de la valeur optimale.

Nous avons implémenté ces méthodes dans un cadre de simulation Monte Carlo haute dimension, en testant différents paramètres de marché et structures de réseaux. Nos expériences montrent que la borne inférieure bénéficie grandement d'un choix adapté de profondeur du réseau, tandis que la borne supérieure peut être estimée de manière stable et rapide grâce à des martingales lagrangiennes bien choisies.

L'analyse de robustesse a également mis en lumière l'apport significatif des techniques de réduction de variance, en particulier la méthode des variables de contrôle, qui permet une réduction notable de la variance sans affecter l'estimateur.

### **Critique et perspectives.**

Bien que la méthode duale de Rogers utilisant des martingales lagrangiennes offre des résultats raisonnables avec un coût numérique maîtrisé, elle repose sur des martingales construites de manière heuristique. L'efficacité de ces martingales pourrait être étudiée plus finement dans le cas d'autres types d'options (put, options sur panier, options lookback, etc.), afin d'évaluer la généralité et les limites de cette méthode. Par ailleurs, le recours à des méthodes plus sophistiquées de sélection ou d'apprentissage de martingales (par exemple via reinforcement learning) pourrait ouvrir de nouvelles perspectives pour affiner les bornes supérieures.

Enfin, un prolongement naturel serait l'étude de versions plus avancées du cadre actuel, intégrant par exemple des frictions de marché, des coûts de transaction, ou encore des dynamiques non log-normales plus proches des données de marché réelles.