

Swing Contract Pricing: A Parametric Approach with Adjoint Automatic Differentiation and Neural Networks

Vincent Lemaire¹, Gilles Pagès¹, and Christian Yeo^{1,2}

¹Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation, UMR 8001, case 158, 4, pl. Jussieu, F-75252 Paris Cedex 5, France

²Engie Global Markets, 1 place Samuel Champlain, 92400 Courbevoie, France

Abstract

We propose two parametric approaches to price swing contracts with firm constraints. Our objective is to create approximations for the optimal control, which represents the amounts of energy purchased throughout the contract. The first approach involves explicitly defining a parametric function to model the optimal control, and the parameters using stochastic gradient descent-based algorithms. The second approach builds on the first one, replacing the parameters with neural networks. Our numerical experiments demonstrate that by using Langevin-based algorithms, both parameterizations provide, in a short computation time, better prices compared to state-of-the-art methods (like the one given by Longstaff and Schwartz).

Keywords - *Swing contracts, stochastic control, stochastic approximation, neural network, Langevin dynamics, greeks.*

Introduction

With the energy market becoming increasingly deregulated, various derivative products have emerged, offering flexibility in delivery dates and purchased energy amounts. “Swing” contracts, also known as “Take-or-Pay” contracts (see [38] for more details), are among the most widely traded contracts in gas and power markets. These contracts allow their holders to purchase amounts of energy on fixed exercise dates, subject to certain constraints. There exist two types of constraints. In the firm constraints setting, the contract holder is not allowed to violate the constraints. Beside this case, there exists an alternative setting where the holder can violate them but will face a penalty proportional to the default. Our paper focuses on the valuation of swing contracts in the firm constraints case.

The valuation of swing contracts is a more challenging task than that of classic American-style contracts [9, 19, 27, 36], mainly due to the presence of constraints related to time and volume. From a probabilistic standpoint, the valuation of swing contracts is related to a Stochastic Optimal Control (SOC) problem. Here, the control represents the vector of amounts of energy that can be purchased at each exercise date. To solve this SOC problem, two primary groups of methods have been developed.

The first group concerns methods that are based on the so called “Backward Dynamic Programming Principle” (BDPP). In this group, the price of the swing contract can be written as a solution of a dynamic programming equation (see [1, 2, 3, 20, 24, 38]). More precisely, at each exercise date, given

the underlying price and the cumulative consumption, the contract value is given by the maximum over the possible consumptions of the immediate cash flow plus the (conditional) expected value of future cash flows. The latter amount (which is in fact a conditional expectation) is often called the “continuation value” and its evaluation is the main difficulty when consider BDPP-based approaches. The mostly used method for computing the continuation value is the Longstaff and Schwartz one (see [27]); where the continuation value is approximated as an orthogonal projection over a subspace spanned by a finite number of squared integrable random variables (see [3, 27]). Another method is based on the so called optimal quantization (see [30]) where the underlying (continuous) random variable is approximated by its discrete version which in turns is defined by a “projection” on nearest neighbors or Voronoï cells (see [2]). Thus the continuation value reduces to a conditional expectation of a discrete random variable. One of the major problems with BDPP’s approaches is the following. To compute the value of the contract at each exercise date, the maximum is taken over a geometric interval. In practice the latter interval needs to be discretized leading to a precision/complexity tradeoff. Indeed, achieving a high level of accuracy in pricing the contract requires a finer discretization of the geometric interval, which in turn increases computation time. Additionally, the Longstaff-Schwartz method faces a storage challenge, as regression coefficients for each simulation and admissible cumulative consumption must be stored at each exercise date to compute the continuation value. This often leads to a memory overflow when a large number of simulations is required to obtain a price with a tight confidence interval. Meanwhile, the optimal quantization method suffers from the “curse of dimensionality” because it is well known that the rate of convergence of this method is of order $\mathcal{O}(N^{-\frac{1}{d}})$ (where N is the number of simulations and d is the problem dimension).

An alternative approach to BDPP-based methods is to consider the valuation of swing contracts as a global optimization problem. As mentioned, the valuation of swing contract is equivalent to solve a SOC problem where the objective is to find a vector of purchase amounts that maximizes the expected value of cumulative cash flows up to the expiry of the contract. This leads to a stochastic optimization problem where the control is generally substituted by a parametric function, and optimal parameters are identified using Stochastic Gradient Descent (SGD) algorithms. The use of parametric functions to approximate the optimal control reduces the control problem to a parametric optimization problem. However, the success of such approaches is contingent upon finding a suitable parametric function, which often requires a solid understanding and intuition of the optimal control’s behavior. Note that, solving SOC problem with SGD based algorithms have been considered in [17] for general SOC problems and in [3] for the swing contracts case. To the best of our knowledge, there has been limited exploration of such approaches in the literature to price swing options. Thus, in this paper, we propose two parametric methods for global optimization and compare them with the Longstaff-Schwartz method. We optimize both parameterizations using two algorithms: Adaptive Moment Estimation (Adam) and (Preconditioned) Stochastic Gradient Langevin (PSGLD). While Adam is widely used in stochastic approximation, PSGLD has demonstrated effectiveness in recent studies for Bayesian learning. Our results show that using Langevin-type algorithms can accelerate the training of our parameterizations and leads to better prices compared to the state-of-the-art methods.

Our paper is organized as follows. Section 1. We describe swing contracts and recall the pricing framework. Section 2. We present our contribution by proposing two numerical methods for the pricing of swing contracts. Section 3. We present generalities about stochastic approximation and study two optimization algorithms to optimize our numerical solutions. Finally in section 4 we present some additional results about the effectiveness of our methods.

1 On swing contracts

In this section, we will provide an overview of swing contracts and explain how to build the swing volume grid defining the space of volume constraints. In a firm constraints setting, this space is a critical component to price swing contracts, and this will be discussed towards the end of this section.

1.1 Description

Swing option is a commonly encountered derivative on energy (gas and power) markets which allows its holder to buy amounts of energy q_k at times t_k , $k = 0, \dots, n-1$ (called exercise dates) until the contract maturity $t_n = T$. At each exercise date t_k , the purchase price (or strike price of the contract) is denoted K_k and can be constant (i.e. $K_k = K$, $k = 0, \dots, n-1$) or indexed on a formula. In this paper, we focus on the fixed strike price case but the case of indexed strike can be treated likewise. In this latter case the strike at a time t_k is often calculated as the average of the underlying prices over a period which is already completed. Therefore the indexed strike price case only adds a computational complexity rather than a theoretical one.

One important feature on swing option are volume constraints. Swing option gives its holder a flexibility on the amount of energy he is allowed to purchase and this flexibility is subject to some (firm) constraints

- **Local constraints:** at each exercise date t_k , the holder of the swing contract is allowed to buy at least q_{\min} and at most q_{\max} i.e.,

$$q_{\min} \leq q_k \leq q_{\max}, \quad 0 \leq k \leq n-1. \quad (1.1)$$

- **Global constraints:** the cumulative consumption up to the maturity of the contract must be not lower than Q_{\min} and not greater than Q_{\max} i.e.,

$$Q_n = \sum_{k=0}^{n-1} q_k \in [Q_{\min}, Q_{\max}], \quad \text{with } Q_0 = 0 \text{ and } 0 \leq Q_{\min} \leq Q_{\max} \leq +\infty. \quad (1.2)$$

Note that in this paper we only consider the **firm constraints** case which means that the holder of the contract is not allowed to violate the constraints. This case is often released in the literature by adding a penalty term whenever the global constraints are violated (see [2, 3]). In penalty setting, the existence of an optimal Markov consumption had been proved in [3] using a smooth penalty function. A similar result has been established in [1] for the firm constraints case.

Now let us focus on the firm constraints case in which one of the main components is the physical space. That is the area of possible actions taking into account the volume constraints.

1.2 Physical space

At each exercise date, through the firm constraints, the achievable cumulative volumes are delimited by boundaries represented by the functions $t_k \mapsto Q^{\text{down}}(t_k)$ and $t_k \mapsto Q^{\text{up}}(t_k)$

$$\begin{cases} Q^{\text{down}}(t_0) = 0 \\ Q^{\text{down}}(t_k) = \max(0, Q_{\min} - (n-k) \cdot q_{\max}) \quad , \quad k \in \{1, \dots, n-1\} \\ Q^{\text{down}}(t_n) = Q_{\min} \end{cases} \quad (1.3)$$

$$\begin{cases} Q^{up}(t_0) = 0 \\ Q^{up}(t_k) = \min(k \cdot q_{\max}, Q_{\max}) \quad , k \in \{1, \dots, n-1\} \\ Q^{up}(t_n) = Q_{\max} \end{cases} \quad (1.4)$$

where $Q^{down}(t_k)$ and $Q^{up}(t_k)$ denote respectively the lower and the upper bound of the cumulative consumption at time t_k . These boundaries lead to the physical space of the swing contract, as drawn in Figure 1 and represents at each exercise date, the range of attainable cumulative consumptions.

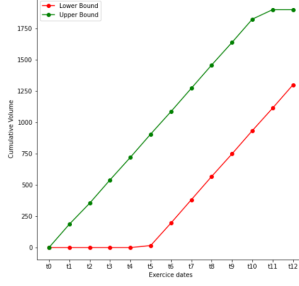


Figure 1: Illustration of swing physical space with $q_{\min} = 0, q_{\max} = 180, Q_{\min} = 1300, Q_{\max} = 1900$ and 12 exercise dates.

In Figure 1 we assumed that $q_{\min} = 0$. This assumption can be made without loss of generality because, as shown in [2], the general case can always be reduced to the case where $q_{\min} = 0$ and $q_{\max} = 1$ (see appendix A for the proof). Besides, if at an exercise date t_k we have bought the amounts q_0, \dots, q_{k-1} leading to a cumulative consumption $Q_k = \sum_{i=0}^{k-1} q_i$, then due to local constraints and the cumulative consumption borders, the actual range of attainable cumulative consumption at time t_{k+1} is the following

$$Q_k + q_k \in \left[\underbrace{\max(Q^{down}(t_{k+1}), Q_k + q_{\min})}_{:=L_{k+1}(Q_k)}, \underbrace{\min(Q^{up}(k+1), Q_k + q_{\max})}_{:=U_{k+1}(Q_k)} \right]. \quad (1.5)$$

That is, starting with a cumulative consumption Q_k , at time t_k , the range of admissible volume is

$$q_k \in \left[\underbrace{L_{k+1}(Q_k) - Q_k}_{:=A_k^-(Q_k)}, \underbrace{U_{k+1}(Q_k) - Q_k}_{:=A_k^+(Q_k)} \right]. \quad (1.6)$$

With theses basic blocks, we can now set the theoretical framework for the pricing of swing contracts.

1.3 Pricing and sensitivity calculus

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space. The price at time t of the forward contract delivered at the maturity T is denoted by $F_{t,T}$. In this paper (as in [2, 3]) we consider contract on the spot price ($S_t = F_{t,t}$) even if in practice the spot is not a tradable instrument on market. Instead, the most encountered gas contract is the day-ahead forward contract whose price is $F_{t,t+1}$. But this case can be treated in the same way.

The decision process $(q_k)_{0 \leq k \leq n-1}$ is defined on the same probability space and is supposed $\mathcal{F}_{t_k}^S$ -adapted. At each exercise date t_k , by buying a certain volume q_k , the holder of the contract makes a profit (or loss)

$$\psi_k(q_k, S_{t_k}) := q_k \cdot (S_{t_k} - K). \quad (1.7)$$

Let $Q \in \mathbb{R}_+$. Given a cumulative consumption Q , an admissible strategy at time t_k is a vector (q_k, \dots, q_{n-1}) lying within the following set

$$\mathcal{A}_{k,Q}^{Q_{\min}, Q_{\max}} = \left\{ (q_\ell)_{k \leq \ell \leq n-1}, q_\ell : (\Omega, \mathcal{F}_{t_\ell}^S, \mathbb{P}) \mapsto [q_{\min}, q_{\max}], \sum_{\ell=k}^{n-1} q_\ell \in [(Q_{\min} - Q)_+, Q_{\max} - Q] \right\}.$$

Note that using equation (1.6), the preceding set reads,

$$\mathcal{A}_{k,Q}^{Q_{\min}, Q_{\max}} = \left\{ (q_\ell)_{k \leq \ell \leq n-1}, q_\ell : (\Omega, \mathcal{F}_{t_\ell}^S, \mathbb{P}) \mapsto [A_\ell^-(Q_\ell), A_\ell^+(Q_\ell)], \text{ where } Q_\ell = Q + \sum_{i=k}^{\ell-1} q_i \right\}. \quad (1.8)$$

with the convention $\sum_{i=k}^{k-1} q_i = 0$. Then for every non negative $\mathcal{F}_{t_{k-1}}^S$ -measurable random variable Q , the price of the swing option at time t_k starting from a cumulative consumption Q is given by

$$P_k(S_{t_k}, Q) = \operatorname{ess\,sup}_{(q_\ell)_{k \leq \ell \leq n-1} \in \mathcal{A}_{k,Q}^{Q_{\min}, Q_{\max}}} \mathbb{E} \left(\sum_{\ell=k}^{n-1} e^{-r_\ell(t_\ell - t_k)} \psi_\ell(q_\ell, S_{t_\ell}) | \mathcal{F}_{t_k}^S \right), \quad (1.9)$$

where the expectation is taken under the risk-neutral probability and r_ℓ are interest rates over the period $[t_0, t_{n-1}]$ that we will assume to be zero throughout this paper. Then the price of the swing contract is given by (S_{t_0} is assumed to be deterministic)

$$P_0 := P_0(S_{t_0}, 0) = \sup_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}} \mathcal{J}(q_0, \dots, q_{n-1}). \quad (1.10)$$

where, given an admissible strategy (q_0, \dots, q_{n-1}) , the reward function \mathcal{J} is defined as the expected value of cumulative future cash flows up to the expiry

$$\mathcal{J}(q_0, \dots, q_{n-1}) := \mathbb{E} \left(\sum_{\ell=0}^{n-1} \psi_\ell(q_\ell, S_{t_\ell}) \right). \quad (1.11)$$

The latter problem appears to be a constrained stochastic control problem in which the aim is to find an admissible strategy that maximizes the reward function \mathcal{J} . As mentioned in the introduction, there exists in the literature two groups of pricing methods to solve this optimization problem. The first group of methods is based on the “backward dynamic programming principle” and the second on a global optimization approach. In this paper, we propose two numerical solutions to solve problem (1.10) based on this latter group.

It is important to note the “bang-bang” feature, which implies that under certain conditions, at each exercise date, the optimal consumption is given by one of two values: A_k^- or A_k^+ (defined in equation (1.6)). This feature has been proven for the case of firm constraints and when global constraints are whole numbers by Pagès et al. [1]. For the penalty case, a proof can be found in Gobet et al. [3]. The “bang-bang” feature is particularly valuable since it allows to significantly reduce the computation time it.

In this paper and unless otherwise stated, we use a one-factor model as in [3, 20]. That is,

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma e^{-\alpha(T-t)} dW_t, \quad t \leq T \quad (1.12)$$

where W is a standard Brownian motion. As mentioned before we deal with the spot price whose price is given by a straightforward application of Itô formula

$$S_t = F_{0,t} \cdot \exp\left(\sigma X_t - \frac{1}{2}\lambda_t^2\right), \quad X_t = \int_0^t e^{-\alpha(t-s)} dW_s \quad \text{and} \quad \lambda_t^2 = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}). \quad (1.13)$$

Along with this model, we will use two settings (presented below). In each case we set $\alpha = 4, \sigma = 0.7$.

Case 1

$$31 \text{ exercise dates} \quad q_{\min} = 0 \quad q_{\max} = 6 \quad Q_{\min} = 140 \quad Q_{\max} = 200.$$

Case 2 (as in [2, 3])

$$365 \text{ exercise dates} \quad q_{\min} = 0 \quad q_{\max} = 6 \quad Q_{\min} = 1300 \quad Q_{\max} = 1900.$$

The computer that has been used has the following characteristics: *Processor: Intel(R) Core(TM) i7-1185G7 @ 3.00GHz 3.00 GHz, 32Go of RAM, Microsoft Windows 10 Enterprise.* Deep learning part had been implemented using PyTorch toolbox. The GPU device used is: Nvidia A100-PCIE 40GB.

It is worth noting that for practitioners, the prices of derivatives are closely linked to their sensitivities with respect to market data. These sensitivities are essential for hedging purposes, as they indicate how the price of a derivative product changes with the market. Computation of sensitivities involves calculating derivatives, and in our case, we need to differentiate a price with respect to some parameters, where the price is a solution to the stochastic control problem in equation (1.10). To achieve this, we rely on the so called “envelope theorem.”

Let $f(x, \alpha)$ and $g_j(x, \alpha), j = 1, 2, \dots, m$ be real-valued continuously differentiable functions on $\mathbb{R}^{n+\ell}$, where $x \in \mathbb{R}^n$ are some variables and $\alpha \in \mathbb{R}^\ell$ are parameters, and consider the constrained optimization problem

$$\begin{cases} \max_x f(x, \alpha) \\ \text{subject to } g_j(x, \alpha) \geq 0, \quad 1 \leq j \leq m \end{cases}$$

We introduce the Lagrangian function,

$$\mathcal{L}(x, \lambda, \alpha) = f(x, \alpha) + \langle \lambda, g(x, \alpha) \rangle$$

where $\lambda \in \mathbb{R}^m$ is the Lagrange multipliers, $\langle \cdot, \cdot \rangle$ is the Euclidean inner-product and $g = (g_1, \dots, g_m)^\top$. Then we define the value function $V(\alpha) = f(x^*(\alpha), \alpha)$ where $x^*(\alpha)$ is a solution that maximizes the function $f(\cdot, \alpha)$. The following theorem gives the derivative of the value function V in case it is differentiable.

Theorem 1 (Envelope theorem) *Assume that V and \mathcal{L} are continuously differentiable. Then,*

$$\frac{\partial V(\alpha)}{\partial \alpha_k} = \frac{\partial \mathcal{L}(x^*(\alpha), \lambda^*(\alpha), \alpha)}{\partial \alpha_k} \quad k = 1, \dots, \ell$$

where $\frac{\partial \mathcal{L}}{\partial \alpha_k} = \frac{\partial f}{\partial \alpha_k} + \langle \lambda, \frac{\partial g}{\partial \alpha_k} \rangle$.

The envelope theorem states that, under some regularity conditions, in a SOC problem, for example our problem (1.10), the derivative of the optimal objective function \mathcal{J} is given by differentiating the cash flows along with the optimal control. In our model (1.12), we deduce the following proposition which is a corollary application of Theorem 1.

Proposition 1 *Let $(q_k^*)_{0 \leq k \leq n-1}$ be a solution of the problem (1.10). Assume that the function*

$$(F_{0,t_k})_{0 \leq k \leq n-1} \mapsto \mathbb{E} \left(\sum_{k=0}^{n-1} q_k^* \cdot \left(F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K \right) \right)$$

is continuously differentiable. Let P_0 be the price of the swing contract (1.10). Note that P_0 is a function of $(S_{t_k})_{0 \leq k \leq n-1}$ which in turns depends on $(F_{0,t_k})_{0 \leq k \leq n-1}$ in the model (1.12). Then for all $k = 0, \dots, n-1$, the delta (sensitivity of swing price with respect to the initial forward price) is given by

$$\frac{\partial P_0}{\partial F_{0,t_k}} = \mathbb{E} \left(q_k^* \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} \right).$$

PROOF We define the following functions:

$$f((q_k)_{0 \leq k \leq n-1}, (F_{0,t_k})_{0 \leq k \leq n-1}) := \mathbb{E} \left(\sum_{k=0}^{n-1} q_k \cdot \left(F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K \right) \right)$$

$$g_1((q_k)_{0 \leq k \leq n-1}, (F_{0,t_k})_{0 \leq k \leq n-1}) := \sum_{k=0}^{n-1} q_k - Q_{\min}$$

$$g_2((q_k)_{0 \leq k \leq n-1}, (F_{0,t_k})_{0 \leq k \leq n-1}) := Q_{\max} - \sum_{k=0}^{n-1} q_k$$

and for all $k = 0, \dots, n-1$:

$$g_{2k+3}((q_k)_{0 \leq k \leq n-1}, (F_{0,t_k})_{0 \leq k \leq n-1}) := q_k - q_{\min}$$

$$g_{2k+4}((q_k)_{0 \leq k \leq n-1}, (F_{0,t_k})_{0 \leq k \leq n-1}) := q_{\max} - q_k$$

Then it suffices to prove that functions $f, g_1, g_2, \dots, g_{2n+2}$ are continuously differentiable in order to use the envelope theorem. This holds for functions $g_1, g_2, \dots, g_{2n+2}$. It remains to prove it for function f . The latter reduces to show that function f is continuously differentiable in each of its components. For any $k = 0, \dots, n-1$, the random variable $q_k \cdot (F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K)$ is integrable since

$$\left| q_k \cdot (F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K) \right| \leq q_{\max} \cdot (F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} + K) \in \mathbb{L}_{\mathbb{R}}^1(\mathbb{P}).$$

Moreover, the function $F_{0,t_k} \mapsto q_k \cdot (F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K)$ is differentiable and its derivative does not depends on F_{0,t_k} . Furthermore,

$$\left| \frac{\partial}{\partial F_{0,t_k}} \left(q_k \cdot (F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K) \right) \right| = \left| q_k \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} \right| \leq q_{\max} \cdot e^{\sigma X_{t_k}} \in \mathbb{L}_{\mathbb{R}}^1(\mathbb{P}).$$

Then thanks to Lebesgue theorem to interchange derivation and integral, the function f is continuously differentiable in all F_{0,t_k} . Likewise one may also show that for all $k = 0, \dots, n-1$ the function f is continuously differentiable in q_k , and one may use the envelope theorem. For any $k = 0, \dots, n-1$ we introduce the Lagrangian of the problem (1.10)

$$\mathcal{L}\left((q_k)_{0 \leq k \leq n-1}, \lambda, (F_{0,t_k})_{0 \leq k \leq n-1}\right) = \mathbb{E}\left(\sum_{k=0}^{n-1} q_k \cdot \left(F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K\right)\right) + \langle \lambda, g((q_k)_{0 \leq k \leq n-1}, (F_{0,t_k})_{0 \leq k \leq n-1}) \rangle,$$

where $g = (g_1, \dots, g_{2n+2})^\top$ and $\lambda \in \mathbb{R}^{2n+2}$. Then it follows from envelope theorem that for any $k = 0, \dots, n-1$:

$$\frac{\partial P_0}{\partial F_{0,t_k}} = \mathbb{E}\left(\sum_{k=0}^{n-1} q_k^* \cdot \frac{\partial}{\partial F_{0,t_k}} \left(F_{0,t_k} \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}} - K\right)\right) = \mathbb{E}\left(\sum_{k=0}^{n-1} q_k^* \cdot e^{\sigma X_{t_k} - \frac{1}{2} \Lambda_{t_k}}\right).$$

This completes the proof.

We have outlined the theoretical framework for pricing swing contracts and computing their associated sensitivities. To obtain practical numerical solutions, we introduce parametric methods in this paper. Specifically, we replace the optimal control variable q_k at a time t_k with a parametric function $q_k(I_k; \theta)$, where I_k represents information needed to determine the volume to purchase, and θ is a parameter belonging to a parameter space Θ that depends on the chosen parameterization. Thus, the optimization problem in equation (1.10) can be expressed as,

$$\sup_{\theta \in \Theta} \mathcal{J}\left(q_0(I_0; \theta), \dots, q_{n-1}(I_{n-1}; \theta)\right). \quad (1.14)$$

Note that the parameter space has to be designed to guarantee that, almost surely, for $\theta \in \Theta$, $(q_0(I_0; \theta), \dots, q_{n-1}(I_{n-1}; \theta))$ is an admissible strategy. That is, it lies within $\mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}$ which is defined through (1.8) and where the range of admissible volumes at each time t_k depends on purchased volumes up to that time.

2 Swing pricing: A global optimization approach

In this paper, we propose two parametric methods for approximating the optimal control (i.e., the volume to purchase at each exercise date) in a swing pricing context. It is worth noting that the use of parametric exercise strategies for pricing swing contracts is not a new approach and has demonstrated its advantages when compared to classical methods such as the Longstaff and Schwartz method (see [3, 27]). For instance, Gobet et al. [3] developed two parametric methods, in the context of swing contracts with penalties, which provide satisfactory prices when compared to their benchmark (obtained using the forest of trees method [24]). The first method is based on a neural network that approximates a $[0, 1]$ -valued function f , which is then used to determine the volume to purchase in the range $[q_{\min}, q_{\max}]$ (recall that they studied a swing contract with penalties. Thus the range of admissible volume is not constrained as in our case through amounts A^- and A^+ (1.6)) using the simple transformation $q_{\min} + (q_{\max} - q_{\min})f$. In addition to the previous parameterization, the authors proposed another one based on the heuristic that a

higher gas price usually results in a higher consumption at a fixed strike price (though this may not always be true due to global constraints). This heuristic suggests the existence of a threshold beyond which the highest possible volume should be purchased. We studied this heuristic using a neural network to replace the volume q_k with a nonlinear parameterization $q_k := q(t_k, S_{t_k} - K; \theta)$, where θ are the parameters of the neural network trained to solve the swing pricing problem (1.14) (see section 3 for details on the training process). The results showed that when the payoff $S_{t_k} - K$ is below a certain threshold β_1^k , the optimal consumption tends to be A_k^- (as defined in equation (1.6)), while for payoffs above a certain threshold $\beta_2^k (> \beta_1^k)$, the optimal consumption tends to be A_k^+ (also defined in equation (1.6)). Therefore, the optimal consumption profile exhibits a similar behavior as that depicted in Figure 2.

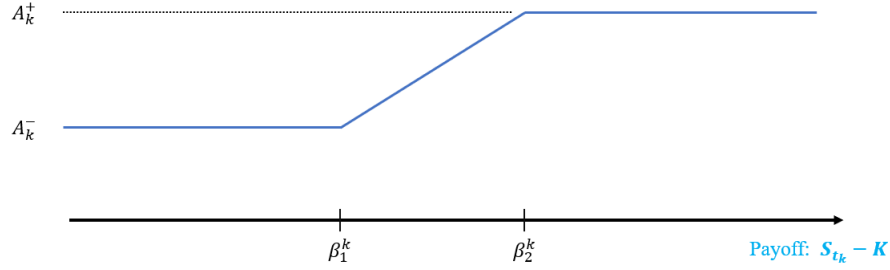


Figure 2: Optimal consumption profile as a function of the payoff.

Based on this observation, we aim to find a parametric function that reproduces the consumption profile shown in Figure 2, which depends on two thresholds β_1 and β_2 (that are assumed to be constant over time in a first step). To achieve this, we define a parametric “decision” function

$$I_k \mapsto f_k(I_k; \theta) := f(I_k; \theta) = \mathbf{1}_{\{I_k > \beta_2\}} + \frac{I_k - \beta_1}{\beta_2 - \beta_1} \mathbf{1}_{\{\beta_1 \leq I_k \leq \beta_2\}} \quad (2.1)$$

where $\mathbf{1}$ denotes the indicator function, $\theta = (\beta_1, \beta_2)^\top \in \mathbb{R}^2$ and $I_k = S_{t_k} - K$ is the payoff. Note that the thresholds and therefore the parameterization do not depend on time. The preceding “decision” function, once defined, allows to choose an admissible control using the following transformation

$$q_k(I_k; \theta) = A_k^-(Q_k^\theta) + \left(A_k^+(Q_k^\theta) - A_k^-(Q_k^\theta) \right) \cdot f_k(I_k; \theta) \quad (2.2)$$

with $Q_0^\theta = 0$ and $Q_k^\theta = \sum_{i=0}^{k-1} q_i(I_i; \theta)$ for all $1 \leq k \leq n-1$. Then we aim at solving the parametric optimization problem (1.14) with the preceding choice of parameterization (2.2) and with $\Theta = \{(x, y) \in \mathbb{R}^2; x < y\}$. To achieve this, we look for values of the thresholds $\theta = (\beta_1, \beta_2)$ which maximize the resulting reward function \mathcal{J} (using strategy given by (2.1)). To implement this approach, we simulate 10^5 independent realizations of the payoff $I_k = S_{t_k} - K$ and compute the Monte-Carlo price given by strategy (2.2). The values of β_1, β_2 used to compute this maximum are the 50 evenly distributed within the interval $[-10, 10]$, for the contract setting of case 2 (as described in section 1.3).

The optimal values for β_1 and β_2 that lead to the maximum price of 2611.7 are obtained as -2.24 and -1.02 , respectively. This simple parametric strategy which has no flexibility (the strategy is the same for all exercise dates since thresholds are constants) and does not depend on the cumulative volume (which is obviously essential when deciding which volume to buy at each exercise dates) gives a satisfying

price when compared to the more elaborated methods in [3]¹. Following the promising results obtained by this straightforward and simple strategy, we introduce two improvements leading to two different parameterizations.

- **Explicit Payoff-Volume parameterization:** we suppose the optimal consumption behaves like profile 2 and define a (smooth) parametric function reproducing this profile. The final idea is to use stochastic gradient descent based optimization algorithms to find the best parameters. In this parameterization, we allow the strategy to depend not only on time but also on cumulative consumption. This latter is useful and gives flexibility to our strategy.
- **Neural network parameterization:** this is a variant of the preceding parameterization. Here we replace coefficients which define the first parameterization by the output of a neural network. Once the neural network is trained, we get the parameters and define the same strategy.

To illustrate both points of view, we use the settings of case 1 in the following sections.

2.1 Explicit Payoff-Volume parameterization (*PV strat*)

We aim to find a parameterization that replicates the target profile in Figure 2, and that is also sufficiently regular to allow for optimization using Stochastic Gradient Descent methods. The parameterization needs to be a sufficiently regular function to ensure that the resulting reward function \mathcal{J} is differentiable. To achieve this, we select the logistic function, denoted by $\sigma(x) := \frac{1}{1+e^{-x}}$, which is infinitely differentiable on \mathbb{R} and has a shape close enough to that of the target profile. Additionally, to incorporate flexibility into our exercise strategy, we no longer assume that the parameters are constant over exercise dates. We also need to take into account for the dependence of the strategy on cumulative consumption. Hence, we propose the following parameterization.

$$f_k(I_k; \theta) = \sigma(\langle \theta_k, I_k \rangle) \quad \text{with} \quad I_k = (S_{t_k} - K, M(Q_k), 1)^\top \in \mathbb{R}^3 \quad (2.3)$$

where for some vector $\theta = (\theta_1, \dots, \theta_n) \in \Theta = \mathbb{R}^{3n}$, $\theta_k \in \mathbb{R}^3$ ($1 \leq k \leq n$) is the subvector of θ made of the k^{th} , $(k+1)^{th}$, $(k+2)^{th}$ components of θ . Thus vector θ embed all parameters used to define the strategy for all exercise dates. In other words, each \mathbb{R}^3 -valued component θ_k represents coefficients that drive the optimal decision of *PV strat* at the corresponding time t_k . The function M is chosen as the (normalized) margin or remaining purchasing capacity defined by: $M(Q) := \frac{Q - Q_{\min}}{Q_{\max} - Q_{\min}}$. The importance of this choice is discussed in Remark 1 and 2. In case $Q_{\min} = Q_{\max}$, one may use $M(Q) = \frac{Q - Q_{\min}}{Q_{\min}}$. Then the problem to solve is the same as in (1.14) where the parameters space is $\Theta = \mathbb{R}^{3n}$ and the control $q_k(I_k; \theta)$ in (2.2) is defined using the strategy (2.3).

Remark 1 (Importance of function M) At first glance, the function M may seem to be a simple normalization of the cumulative consumption. However, it should be noted that several other normalizations were tested and did not lead to a good strategy. We tested the following normalizations: $M(Q_k) = Q_k$, $\frac{Q_k - Q_{\min}}{q_{\max} - q_{\min}}$, $\frac{k \cdot q_{\max} - Q_k}{q_{\max} - q_{\min}}$ and $\frac{Q_k - Q_{\text{down}}(t_{k+1})}{Q_{\text{up}}(t_{k+1}) - Q_{\text{down}}(t_{k+1})}$.

Remark 2 (Remaining capacity *Versus* Cumulative Consumption) We observed that the remaining purchasing capacity is a more crucial factor than the current cumulative consumption when determining

¹Note the relative error between this price and the best of theirs is roughly 2% and they studied the penalty case whose theoretical price is greater than that of the firm constraints case we study in this paper.

the volume to buy at each exercise date. This contrasts with the neural network approach tested in [3], which uses the current cumulative consumption. When using $M(Q) = Q$, our algorithm tended to purchase the maximum possible volume at the contract's outset, as long as the payoff was positive. This behavior suggests that the algorithm did not take into account the fact that purchasing capacity diminishes as the contract approaches its expiration, due to global constraints.

Remarks 1 and 2 explain the reason for which function M yields desirable results. In addition, this choice has a rational interpretation for the optimal control behavior (see appendix D).

2.2 Neural Network parameterization (*NN strat*)

Our second parameterization is based on neural networks. The goal of a neural network is to approximate some mapping $x \mapsto \Phi^*(x)$ by a parametric one (neural network) $x \mapsto \Phi(x; \theta)$ where θ is a parameter (or weights) of the neural network that has to be optimized such in order to give a “good” approximation. A Neural Network can approximate a wide class of complicated (linear or non-linear) relationships (For more details, see Universal Approximation Theorem [18]) between the inputs and outputs by composing linear functions and non-linear thresholds. More precisely, note that a neural network is made of nodes connected to one another where a column of nodes forms a layer (when there are more than one layer in the neural network architecture, we speak of deep neural network). The outermost (see diagram 3) are the input and output layers and all those in between are called the hidden layers. The connection between the input and output layers through hidden layer is made by means of linear functions and activation functions (non-linear ones):

- **Linear functions:** a node is connected to another with an associated weight (w) and bias (b) so that it will receive a weighted value ($wx + b$) from a node in the previous layer. $\theta = (w, b)$ forms parameters of the neural network.
- **Non-linear functions:** each node contains an activation function that activates the node (so the input can pass through the node) when the input of the node is above a certain threshold.

Training a neural network is performed through two steps: **forward propagation** and **backward propagation**. The forward propagation leads the predicted output \hat{y} starting from the input x layer-by-layer. More formally, the forward propagation comes down to the evaluation of the following function:

$$x \in \mathbb{R}^d \mapsto \Phi(x; \theta) := a_I^{\theta_I} \circ \phi_{q_{I-1}} \circ a_{I-1}^{\theta_{I-1}} \circ \dots \circ \phi_{q_1} \circ a_1^{\theta_1}(x) \in \mathbb{R}^\ell \quad (2.4)$$

where

▷ I, q_I, \dots, q_1 are positive integers specifying the depth of the network and the number of nodes for each hidden layer.

▷ $a_1^{\theta_1} : \mathbb{R}^d \rightarrow \mathbb{R}^{q_1}, \dots, a_{I-1}^{\theta_{I-1}} : \mathbb{R}^{q_{I-2}} \rightarrow \mathbb{R}^{q_{I-1}}$ and $a_I^{\theta_I} : \mathbb{R}^{q_{I-1}} \rightarrow \mathbb{R}^\ell$ are affine functions. As explained above, these functions are of the form: $a(x) = Wx + b$ where W is a matrix of weights and b is a vector of bias.

▷ For $j \in \mathbb{N}$, $\phi_j : \mathbb{R}^j \rightarrow \mathbb{R}^j$ are the activation functions. In this paper we only use ReLU (Rectified Linear Unit) function i.e, $x \mapsto \max(x, 0)$.

In our case, the neural network is designed/trained to provide a “decision” function which maximizes the reward function \mathcal{J} . To achieve this, we design a \mathbb{R}^3 -valued neural network $\Phi(\cdot; \theta)$ (thus $\ell = 3$, see

diagram 3) of form (2.4) and which depends on the time t_k , the payoff $S_{t_k} - K$ and the margin $M(Q_k)$ defined as in (2.3). Thus $d = 3$ and parameter θ of the neural network $\Phi(\cdot; \theta)$ lies within

$$\Theta = \mathbb{R}^{q_1} \times \mathbb{R}^{3 \times q_1} \times \left(\prod_{i=2}^{I-1} \mathbb{R}^{q_i} \times \mathbb{R}^{q_i \times q_{i-1}} \right) \times \mathbb{R}^3 \times \mathbb{R}^{3 \times q_{I-1}}. \quad (2.5)$$

As for *PV strat*, we define recursively the “decision” function by $f_0(I_0; \theta) = \sigma(\langle \Phi(t_0, S_{t_0} - K, M(0); \theta), I_0 \rangle)$ and for all $1 \leq k \leq n - 1$

$$f_k(I_k; \theta) = \sigma(\langle \Phi(t_k, S_{t_k} - K, M(Q_k^\theta); \theta), I_k \rangle) \quad \text{with} \quad Q_k^\theta = \sum_{i=0}^{k-1} q_i(I_i; \theta) \quad (2.6)$$

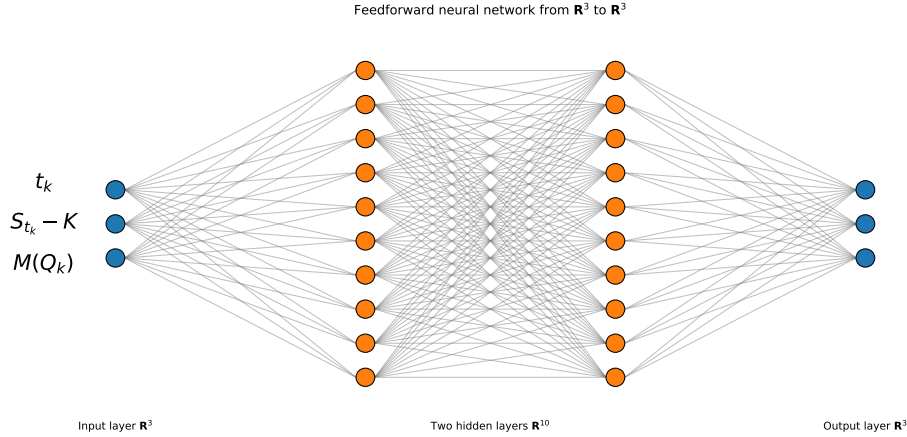


Figure 3: Illustration of (deep) neural network architecture.

where information I_k is the same as in (2.3). Then the parametric optimization problem to solve is the same as defined in equation (1.14) with the preceding choice of Θ given by equation (2.5) and where the control is defined using the strategy (2.6) combined with (2.2).

Remark 3 (*PV strat Versus NN strat*) Note that in *NN strat*, the parameter $\theta \in \Theta$ (as defined in (2.5)) replaces all vectors $\theta_k \in \mathbb{R}^3$ in equation (2.3), which drives the strategy in *PV strat*. Unlike the latter, in *NN strat*, the same set of parameters defines the strategy for all exercise dates, but the strategy is different according to the exercise date since time is integrated as input. In contrast, the size of the parameter vector in *PV strat* increases linearly with the number of exercise dates. Additionally, in section 3.4, we will, numerically, demonstrate that even for a high number of exercise dates, a relatively small neural network architecture is still sufficient. This is one of the advantages of *NN strat* over *PV strat*.

Our neural network parameterization is a more robust one compared to the one performed by Gobet et Al. [3]. Indeed, in their article they had chosen, for the decision function, a non linear $[0, 1]$ -valued parameterization. Our neural network parameterization is an improvement of theirs in the sense that we first impose a particular parametric shape. Then the parameters that drive the latter shape are chosen as

an output of a neural network. This approach not only provides flexibility to the strategy but also helps the algorithm to “learn” the optimal control by imposing a specific shape.

All the strategies presented above require training to find the best parameters. To achieve this, we rely on stochastic approximation theory.

3 Stochastic optimization

In this section we aim at optimizing a \mathbb{R} -valued function $h : y \in \mathbb{R}^q \mapsto \mathbb{E}(H(y, Z))$ with H being a \mathbb{R} -valued function defined on $\mathbb{R}^q \times \mathbb{R}^d$ and where Z is a d -dimensional random vector. Like a probabilistic extension of the classic Newton-Raphson procedure, stochastic approximation [23, 34] suggests the following procedure to find zeros of function h :

$$\forall n \in \mathbb{N}, \quad y_{n+1} = y_n - \gamma_{n+1} \cdot h(y_n) \quad , 0 < \gamma_n \leq \gamma_0. \quad (3.1)$$

When h is the gradient of another function i.e, $h = \nabla V$, we speak about **stochastic gradient descent** (SGD) which is an extension of gradient descent (see [37]) to stochastic optimization. Hereafter we consider a SGD setting i.e, $h = \nabla V$. The sequence $(\gamma_n)_{n \in \mathbb{N}}$ is called the algorithm step or learning rate and has to be chosen carefully. In fact if the step is too large, then the procedure (3.1) can overshoot and completely miss the global optimum of the function V . If the step is too small, the procedure may take a long time to converge as it will be taking small steps towards the global minimum. Thus to ensure convergence (which is widely discussed in the literature [5, 10, 35]), a classical requirement is the “decreasing step” assumption. This means that the step sequence is non-increasing and satisfy the following conditions

$$\sum_{n \geq 0} \gamma_n = +\infty \quad \text{and} \quad \sum_{n \geq 0} \gamma_n^2 < +\infty. \quad (3.2)$$

From a practical point of view, to implement iteration (3.1), one uses mini-batch procedure. That is, rather than computing gradient over a single observation, leading to one iteration in procedure (3.1), we use small batches. We divide a sample of size M into B samples of size $L := M/B$. Then we perform B iterations using

$$y_{n+1} = y_n - \frac{\gamma_{n+1}}{L} \sum_{\ell=1}^L H(y_n, Z_{n+1}^{b, [\ell]}), \quad 0 \leq n \leq N. \quad (3.3)$$

where $Z_{n+1}^{b, [\ell]}$ are independent copies of Z . This alternative increases update frequency compared to vanilla SGD which allows a more robust convergence, avoiding local minima/maxima. It should be noted that in the SGD framework, H often represents the gradient of a multivariate function, which makes it difficult or impossible to implement the SGD procedure by hand. To address this issue, we use a technique called “Adjoint Algorithm Differentiation” (AAD) as described in [4, 32]. This technique involves a program that computes a value and automatically computes derivatives of that value by combining the derivatives of several simple arithmetic expressions.

Vanilla SGD can often perform poorly due to various reasons. One of the well-known reasons is its inability to converge or a slow convergence to a local minima/maxima when the objective function has saddle points. A saddle point is a critical point on the surface of the function’s graph that is not a local extremum. Saddle points are often found in regions that look like plateaus, as shown in Figure 4. It should be noted that saddle points are notorious in high-dimensional spaces, especially in cases involving deep neural networks, which often have non-convex objective functions (see [13, 26]). In fact, as the

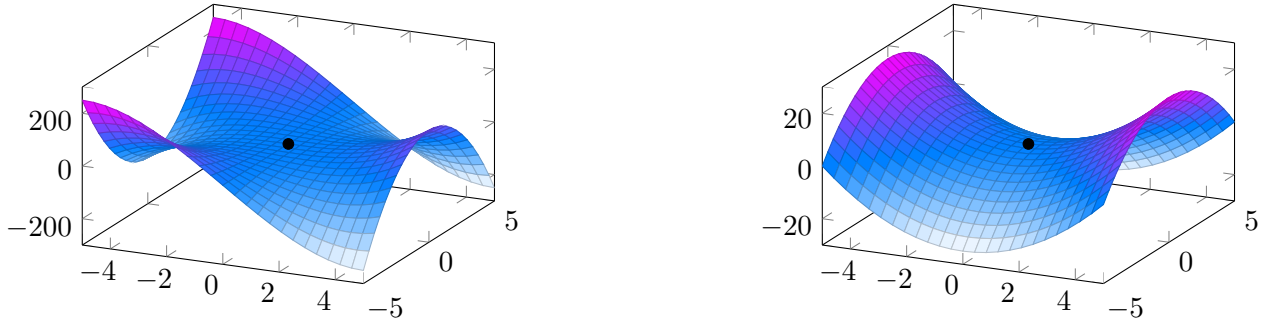


Figure 4: Saddle points illustration. $(x, y) \mapsto x^3 - 3xy^2$ (on left) and $(x, y) \mapsto x^2 - y^2$ (on right). The point $(0, 0)$ is a saddle point.

dimensionality of the problem increases, the number of saddle points has been shown to increase exponentially compared to the number of local minima/maxima. For both *PV strat* and *NN strat*, the number of parameters to optimize is high, especially in *NN strat*. Our experiments have shown that vanilla SGD does not perform well in these high-dimensional spaces. Therefore, in this paper, we propose two alternative optimization algorithms, Adaptive Moment Estimation (Adam) and Stochastic Langevin Dynamics (SGLD), and compare their performance.

3.1 Adaptive Moment Estimation (Adam)

Adam [22] is an efficient stochastic optimization method that only requires first-order gradients with little memory requirement. The method computes individual adaptive learning rates for different parameters from estimates of first gradient moments. Practically, this is done by means of a preconditioning matrix P (see updating (3.4)). Adam can be seen as a combination of RMSprop² and Stochastic Gradient Descent with momentum [33]. The procedure in Adam updating is the following

$$y_{n+1} = y_n - \gamma_{n+1} P_{n+1} \cdot \widehat{\nabla V}(y_n) \quad (3.4)$$

where $\widehat{\nabla V}$ is a (scaled) estimation of the gradient (see Algorithm 1) computed as a moving average of the estimated gradient in the spirit of momentum technique. $(P_n)_n$ is a sequence of preconditioning matrix. In Adam updating the latter matrix is a diagonal matrix whose diagonal components depend on the squared (elementwise) of the gradient g_{n+1} estimated after iteration n . More precisely, given a gradient $g = (g_1, \dots, g_q) \in \mathbb{R}^q$, the square elementwise of this gradient denoted $g \odot g$ is given by

$$g \odot g = (g_1^2, \dots, g_q^2). \quad (3.5)$$

In the same spirit, for some matrix $A = (a_{i,j})_{1 \leq i,j \leq q}$, $B = (b_{i,j})_{1 \leq i,j \leq q}$ where $b_{i,j} \neq 0$ for all $1 \leq i, j \leq q$ we define elementwise division as follows

$$A \oslash B = (c_{i,j})_{1 \leq i,j \leq q} \quad \text{where } c_{i,j} = a_{i,j}/b_{i,j}. \quad (3.6)$$

Adam algorithm is described in 1.

²Unpublished optimization algorithm designed for neural networks, first proposed by Geoffrey Hinton in a Coursera course https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf

Algorithm 1 Adam updating. λ is a small correction term to avoid division by zero. \odot denotes element-wise multiplication and \oslash denotes elementwise division defined in (3.5) and (3.6). Id_q represents the $q \times q$ identity matrix.

Require:

- ▷ γ : Step size
- ▷ $\mu_1, \mu_2 \in [0, 1)$: Exponential decay rates for the moment estimates
- ▷ $V(y)$: Stochastic objective function with parameters y . Note that V is computed by its Monte-Carlo version using (3.3)
- ▷ y_0 : Initial parameter vector

$M_0 = 0$ (Initialize 1^{st} moment vector)
 $MS_0 = 0$ (Initialize 2^{nd} moment vector)

```

1: while  $y_n$  not converged do
2:    $g_{n+1} \leftarrow \nabla_y V(y_n)$                                 ▷ Compute gradient w.r.t to parameters
3:    $M_{n+1} \leftarrow \mu_1 \cdot M_n + (1 - \mu_1) \cdot g_{n+1}$       ▷ Update biased first moment estimate
4:    $MS_{n+1} \leftarrow \mu_2 \cdot MS_n + (1 - \mu_2) \cdot g_{n+1} \odot g_{n+1}$   ▷ Update biased second raw moment estimate
5:    $\widehat{M}_{n+1} \leftarrow M_{n+1} / (1 - \mu_1^{n+1})$                 ▷ Compute bias-corrected first moment estimate
6:    $\widehat{MS}_{n+1} \leftarrow MS_{n+1} / (1 - \mu_2^{n+1})$               ▷ Compute bias-corrected second raw moment estimate
7:    $P_{n+1} = \text{diag} \left( Id_q \oslash (\lambda \cdot Id_q + \sqrt{\widehat{MS}_{n+1}}) \right)$   ▷ Compute preconditioning matrix
8:    $y_{n+1} \leftarrow y_n - \gamma \cdot P_{n+1} \cdot \widehat{M}_{n+1}$       ▷ Update parameters
9: end while

```

Adam uses the squared gradients to scale the learning rate like RMSprop and it takes advantage of momentum by using the moving average (driven by μ_1, μ_2) of the gradient instead of the gradient itself like SGD with momentum. The scaling of the learning rate by the squared gradient allows to solve gradient magnitude problem. Otherwise, sometimes the gradients may be huge and sometimes small which intricate the choice of a learning rate. The exponential moving average of the gradient allows to “de-noise” the estimation of the gradient. Indeed, it is recommended to choose $\mu_1, \mu_2 \approx 1$ so that the contribution of the current estimation of the gradient is small since we give a higher weight to previous data points and lower one to current data points. This allows to smooth subsequent estimation of gradient.

Adam has become the default algorithm for stochastic optimization and especially for neural networks training. Its theoretical convergence had been widely studied in the literature. First, the authors of the algorithm analyzed the convergence in a convex framework with bounded gradient assumption. For non-convex cases, one may refers to [14, 40].

In this paper, we do not limit our study to Adam updating. We propose to explore other algorithm based on Langevin dynamics and which shares, as Adam, good properties when dealing with saddle points.

3.2 Preconditioned Stochastic Gradient Langevin Dynamics (PSGLD)

Stochastic Gradient Langevin Dynamics (SGLD) has been introduced for bayesian learning [39] to approximate a posterior distribution given some data. SGLD combines Robbins-Monro type algorithm with Langevin dynamics which injects noise into the parameter updates such that the trajectory of the parameters converges to the posterior distribution. To be more precise, it has been shown that, under mild conditions, the latter posterior distribution is the unique invariant probability measure of the Langevin

Stochastic Differential Equation (SDE):

$$dy_t = -\nabla V(y_t)dt + \sqrt{2}dW_t, \quad (3.7)$$

where W is a q -dimensional Brownian motion. Practically, the posterior distribution is approximated using an Euler discretization of the Langevin SDE

$$y_{n+1} = y_n - \gamma \nabla V(y_n) + \sqrt{2\gamma} Z_{n+1}, \quad (3.8)$$

where $(Z_n)_n$ is a sequence of i.i.d. standard q -dimensional Gaussian vectors.

From equations (3.1) (with $h = \nabla V$) and (3.8), SGLD appears to be an extension of vanilla SGD by adding an exogenous white noise to the gradient descent. This allows to regularize the problem and escape from traps (see [28]). The use of Langevin dynamics based algorithms for optimization is justified by the fact that recent studies (see [11, 12]) showed that sampling from a distribution which concentrates around the global minimum of V is a similar task as minimizing V via certain optimization algorithms. But, under some conditions, it can be shown (see [29]) that if $(Y_t^y)_{t \geq 0}$ is a \mathbb{R}^d -valued random vector solution of the following SDE

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t, \quad Y_0 = y \quad (3.9)$$

with W a standard q -dimensional Brownian motion and a drift term of the form

$$b := -\frac{1}{2} \left((\sigma \sigma^\top) \nabla V - \left[\sum_{j=1}^d \partial_{y_j} (\sigma \sigma^\top)_{ij} \right]_{i=1:d} \right). \quad (3.10)$$

Then, the distribution

$$\nu_V(dy) = C_V e^{-V(y)} \cdot \lambda_d(dy) \quad (3.11)$$

is the unique invariant distribution of the SDE (3.9). In particular if $\sigma = \sqrt{2} \cdot Id_d$ then get the langevin SDE (3.7). Which shows that the latter SDE converges to its stationary distribution, namely the Gibbs measure $\propto \exp(-V(y))$ which, in fact, concentrates on the global minimum of V . All this motivates the use of Langevin dynamics based algorithms for stochastic optimization (see [6, 8, 21]). In this paper, as in [8], we consider the application of Langevin algorithms in a non-bayesian setting and we will consider the preconditioned version of SGLD which leads to Preconditioned SGLD (PSGLD, see [25]). The latter choice is due to the fact that in its standard version, SGLD (without noise) acts like SGD by updating all parameters with the same step size leading to slow mixing when the components of y have different curvature. Given a sequence of preconditioning matrix $(P_n)_n$, the updating procedure (3.8), reads in PSGLD setting

$$y_{n+1} = y_n - \gamma_{n+1} P_{n+1} \cdot \nabla V(y_n) + \sigma_{n+1} \sqrt{\gamma_{n+1}} \mathcal{N}(0, P_{n+1}) \quad (3.12)$$

where $(\sigma_n)_n$ is a constant or non-increasing sequence controlling the amount of injected noise. As in Adam updating the preconditioning matrix sequence $(P_n)_n$ is a diagonal matrix whose components depend on the squared (elementwise) of the gradient estimation. The general procedure is presented in Algorithm 2.

Algorithm 2 PSGLD updating. We take $\sigma_n = \sigma$ and $\gamma_n = \gamma$. \odot and \oslash are defined in (3.5) and (3.6)

Require:

- ▷ γ : Step size
- ▷ $\mu \in [0, 1)$: Exponential decay rate for the moment squared estimates
- ▷ $V(y)$: Stochastic objective function with parameters y . Note that V is computed by its Monte-Carlo version using (3.3)
- ▷ y_0 : Initial parameter vector

$MS_0 = 0$ (Initialize 2^{nd} moment vector)

```

1: while  $y_n$  not converged do
2:    $g_{n+1} \leftarrow \nabla_y V(y_n)$                                 ▷ Compute gradient w.r.t to parameters
3:    $MS_{n+1} = \mu \cdot MS_n + (1 - \mu) \cdot g_{n+1} \odot g_{n+1}$       ▷ Update biased second raw moment estimate
4:    $P_{n+1} = \text{diag} \left( Id_q \oslash (\lambda \cdot Id_q + \sqrt{MS_{n+1}}) \right)$   ▷ Compute preconditioning matrix
5:    $y_{n+1} \leftarrow y_n - \gamma P_{n+1} \cdot g_{n+1} + \sigma \sqrt{\gamma} \cdot \mathcal{N}(0, P_{n+1})$   ▷ Update parameters
6: end while

```

The convergence of Langevin algorithm has been widely studied. One may refers to [15, 16] when the sequence $(\sigma_n)_n$ is constant and to [7, 29] when the sequence is decreasing.

In the next section we implement *PV strat* and *NN strat* with settings of case 1 and optimize them using both Adam and PSGLD presented above.

3.3 Adam versus PSGLD

We implement *PV strat* and *NN strat* using either Adam or PSGLD optimization algorithm with case 1 setting. For this case the Longstaff-Schwartz method, with canonical polynomial functions of degree 3 provides a price of 65.14 with a 95 % confidence interval: [65.08, 65.21] and in 4 minutes and 40 second in average. Note that with case 1 (even for case 2) setting, regardless the chosen methods, the resulting Monte-Carlo estimator of the swing price exhibits high variance due to the high variance of the underlying. Thus, in the following table and just for this section, prices are computed by averaging *100 replications* of prices. Each of the 100 prices is obtained using $B = 1$ batch of size $L = M = 2^{14}$ and a learning rate $\gamma = 0.1$. In what follows we denote by N the number of iterations in the stochastic procedure (see (3.3)), by \widehat{P}_0 the (raw) price resulting from *PV strat* or *NN strat*. And \widehat{P}_0^{BG} denotes the price given by forcing the decision function (either (2.3) for *PV strat* or (2.6) for *NN strat*) to be strictly bang-bang. Results are recorded in Tables 1 and 2 for *PV strat* and in Tables 2 and 4 for *NN strat*.

N	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
3000	65.11 ([65.04, 65.18])	65.15 ([65.08, 65.22])	68.4
1000	65.02 ([64.95, 65.08])	65.18 ([65.12, 65.24])	21.3

Table 1: Results for *PV strat* using Adam optimization algorithm. Values in brackets are confidence intervals (95%). Column “time” includes both training and valuation times.

σ	β	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
$1 \cdot e^{-6}$	0.8	65.19 ([65.12, 65.26])	65.21 ([65.13, 65.28])	21.3
$1 \cdot e^{-6}$	0.9	65.21 ([65.15, 65.28])	65.23 ([65.17, 65.29])	21.3
$1 \cdot e^{-5}$	0.9	65.22 ([65.16, 65.28])	65.23 ([65.18, 65.29])	21.6

Table 2: Results for *PV strat* using PSGLD optimization algorithm. Values in brackets are confidence intervals (95%). Column “time” includes both training and valuation times. We used $N = 1000$ iterations and $\lambda = 1 \cdot e^{-10}$.

Tables 1 and 2 suggest that Adam requires several iterations and, as a result, increases computation time when using a learning rate of 0.1. Whereas with only 1000 iterations, using PSGLD optimization algorithm, leads to a better price. This fact is also observable in *NN strat* where PSGLD optimization algorithm allows to achieve a better price than Adam.

N	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
3000	65.22 ([65.16, 65.28])	65.23 ([65.17, 65.29])	150.4
1000	65.13 ([65.05, 65.20])	65.22 ([65.15, 65.29])	53.2

Table 3: Results for *NN strat* using Adam optimization algorithm. For the neural network architecture, we used 2 hidden layers ($I = 2$) with 10 units per layer ($q_1 = 10, q_2 = 10$). Values in brackets are confidence intervals (95%). Columns “time” includes the training and the valuation times.

σ	β	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
$1 \cdot e^{-6}$	0.9	65.23 ([65.16, 65.30])	65.21 ([65.14, 65.28])	53.4
$1 \cdot e^{-5}$	0.9	65.23 ([65.14, 65.31])	65.22 ([65.13, 65.30])	52.3
$1 \cdot e^{-6}$	0.8	65.27 ([65.20, 65.35])	65.26 ([65.18, 65.33])	51.3

Table 4: Results for *NN strat* using PSGLD optimization algorithm. For the neural network architecture, we used 2 hidden layers ($I = 2$) with 10 units per layer ($q_1 = 10, q_2 = 10$). Values in brackets are confidence intervals (95%). Columns “time” includes the training and the valuation times. We used $N = 1000$ iterations and $\lambda = 1 \cdot e^{-10}$.

Tables 3 and 4 show that prices obtained by *NN strat* are better than that of *PV strat* and, as already mentioned, PSGLD optimization algorithm provides a better exercise strategy. This means that adding some noise in the optimization procedure helps to converge quickly. Following the performance of PSGLD updating, we consider this algorithm in the remainder along with the following configuration: $\sigma = 1 \cdot e^{-6}, \beta = 0.8, \lambda = 1 \cdot e^{-10}$. With the latter setting, one may compute sensitivity of the swing price with respect to the initial forward price (see Figure 5).

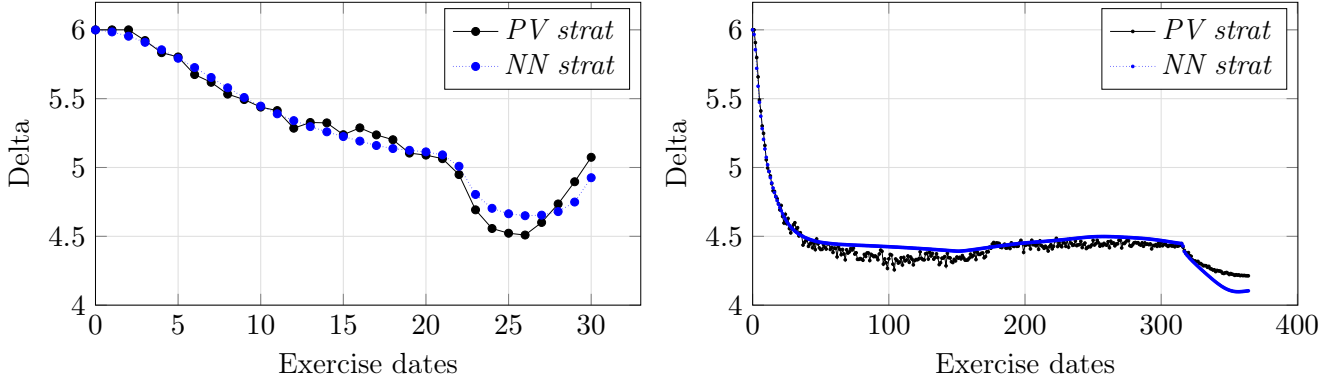


Figure 5: Delta forward for setting of case 1 (left) and case 2 (right) using *PV strat* and *NN strat*.

Note that for both optimization algorithms (Adam and PSGLD) different other hyperparameters had been tested and results are reported in Appendix E.

3.4 Practitioner’s corner: Transfer learning

As mentioned, in practice, training both parameterizations we proposed in this paper is allowed by AAD. However, the latter can be time consuming especially in the case of a swing contract having several exercise dates and using more than one batch per iteration (see Appendix E). For instance, let us consider a swing contract with maturity one year and daily exercise (365 exercise dates). Alongside we use the diffusion model (1.12) with the setting of case 2. The results are reported in the table below. As observed in [3] the Longstaff-Schwartz method generates numerical instabilities. Moreover, note that we get better prices than optimal quantization method like in [2].

	\widehat{P}_0	\widehat{P}_0^{BG}	(T_{train}, T_{eval})
<i>PV strat</i>	2690.10 ([2689.20, 2691.00])	2692.59 ([2691.68, 2693.49])	(274.1, 74.7)
<i>NN strat</i>	2693.74 ([2692.83, 2694.64])	2694.12 ([2693.21, 2695.02])	(680.9, 398.3)

Table 5: Results for a one-year swing contract. Values in brackets are confidence intervals (95%). The valuation was performed with a sample of size $1 \cdot e^8$. For the training we used $N = 1000$ iterations. T_{train} denotes the training time and T_{eval} the valuation time.

From Table 5 one may notice that the computation time is not negligible even if, compared to other methods (like Longstaff-Schwartz) it remains reasonable. To reduce it, we propose a method based on the so called **transfer learning** (for details see [31]). Transfer learning refers to a machine learning method where a model developed for a task is reused as the starting point for a model on a second and similar task. This method may accelerates the evaluation of swing contract with several exercise dates. In our case, we perform “transfer learning” as follows. For a swing contract with several exercise dates, we first consider an aggregated version. That means that we consider a contract with less exercise dates and where the local constraints are aggregated. For instance for a swing contract over one year with daily exercise and with $q_{\min} = 0, q_{\max} = 6$, we may consider a contract with one exercise per month (the middle of each month) and where for example, for months with 30 days, the local constraints become $q_{\min} = 0 \times 30 = 0, q_{\max} = 30 \times 6 = 180$. Then we run few iterations to optimize the pricing of the

aggregated contract. The resulting parameters are used as an initial guess for the actual pricing problem. To extend the parameters obtained in the problem with 12 exercise dates to the problem with 365 exercise dates, we assume that all days within the same month behave the same. Therefore the initial guess parameters for all days within the same month are those obtained in the aggregated problem for this month. It turns out that this allows to reach more quickly a descent optimum.

	\widehat{P}_0 \widehat{P}_0^{BG}	T_{agg}	(T_{train}, T_{eval})
PV strat	2688.73 ([2687.82, 2689.63]) 2692.81 ([2691.90, 2693.71])	4.8	(83.2, 76.3)
NN strat	2693.22 ([2692.31, 2694.13]) 2694.12 ([2693.21, 2695.02])	6.9	(206.5, 398.1)

Table 6: Results for a one-year swing contract using transfer learning. We used 500 iterations for the aggregated contract and 300 iterations for the actual one-year contract. For the valuation, we used a sample of size $1 \cdot e^8$. T_{agg} denotes the training time for the aggregated contract, T_{train} and T_{eval} denote respectively the training time and the valuation time for the actual contract.

We can notice from Table 6 that by means of transfer learning the computation time is drastically reduced. With our adaptation of transfer learning, the training time is reduced by a factor of 3 without degrading the accuracy. Besides, always with in mind the aim of reducing the computation time, transfer learning can be used differently. In a few words, by means of transfer learning, when the market data change or when the contracts settings change we do not need to start all over again. To illustrate this point let us consider below three cases: M1, M2, M3, representing some shifts of market/contract settings. Recall that the baseline settings are those consider above (case 2 setting).

Case M1	Case M2	Case M3
$F_{0,t_k} = 22$	$(Q_{\min}, Q_{\max}) = (1400, 2000)$	$K = 18$

Table 7: Market data move scenarios

In Table 7 we consider three different market move scenarios described as follows. In the first case we bumped the initial forward curve from 20 to 22. In the second case we change the global constraints from $Q_{\min} = 1300$ and $Q_{\max} = 1900$ to $Q_{\min} = 1400$ and $Q_{\max} = 2000$. In the final case we reduced the strike price from 20 to 18. We could have changed the local constraints but as we already pointed out, swing pricing can always reduce to the case where $q_{\min} = 0, q_{\max} = 1$ (see Appendix A). Now the question is the following: Could our baseline model give accurate prices without training again ? Or could it help us to get accurate prices quickly ? The answers to theses questions are recorded in the following tables.

Cases	Re-use	Re-train	(T_{train}, T_{eval})
Case M1	5942.65 ([5941.64, 5943.65])	6005.75 ([6004.70, 6006.79])	(85.4, 77.2)
	5944.27 ([5943.26, 5945.27])	6008.54 ([6007.50, 6009.58])	
Case M2	2514.86 ([2513.90, 2515.81])	2516.99 ([2516.04, 2517.93])	(85.0, 76.9)
	2517.26 ([2516.30, 2518.21])	2518.73 ([2517.78, 2519.67])	
Case M3	5655.98 ([5655.06, 5656.89])	5744.91 ([5743.95, 5745.86])	(85.9, 77.05)
	5657.67 ([5656.75, 5658.58])	5747.39 ([5746.43, 5748.34])	

Table 8: Results for *PV strat.* Column “Re-use” provides results when the baseline model parameters are reused as is. Column “Re-train” gives results when we re-train with only 300 iterations with the baseline model parameters used as starting values. T_{train} denotes the training time and T_{eval} the valuation time. The testing computation time is the same when we re-use the model as when we re-train.

Cases	Re-use	Re-train	(T_{train}, T_{eval})
Case M1	5981.17 ([5980.14, 5982.19])	6011.08 ([6010.03, 6012.12])	(203.1, 397.8)
	5982.57 ([5981.54, 5983.59])	6011.62 ([6010.57, 6012.66])	
Case M2	2509.97 ([2509.01, 2510.90])	2516.92 ([2515.96, 2517.87])	(203.7, 397.5)
	2511.08 ([2510.12, 2512.03])	2517.58 ([2516.62, 2518.53])	
Case M3	5713.80 ([5712.85, 5714.72])	5749.97 ([5749.01, 5750.92])	(203.4, 397.6)
	5715.24 ([5714.34, 5716.17])	5750.51 ([5749.55, 5751.46])	

Table 9: Results for *NN strat.* Column “Re-use” provides results when the baseline model parameters are reused as is. Column “Re-train” gives results when we re-train with only 300 iterations with the baseline model parameters used as starting values. T_{train} denotes the training time and T_{eval} the valuation time. The testing computation time is the same when we re-use the model as when we re-train.

Tables 8 and 9 demonstrate that in the case where global constraints change (M2), using the baseline parameters without additional training gives prices that are comparable to those obtained by training with a few iterations starting from the baseline parameters. Furthermore, it is important to note that increasing the number of iterations beyond 300 (we tested up to 1000 iterations in our experiments) does not lead to better prices than those obtained after 300 iterations (see columns labeled “Re-train”). This suggests that reusing the baseline parameters can speed up convergence. This observation remains true regardless of the case (M1, M2, or M3).

4 Numerical experiments

We now perform additional simulations to demonstrate the effectiveness of the two parameterizations that we proposed in this paper.

4.1 Three factor model

We first consider a three factor model whose dynamics is given by

$$\frac{dF_{t,T}}{F_{t,T}} = \sigma_1 e^{-\alpha_1(T-t)} dW_t^1 + \sigma_2 e^{-\alpha_2(T-t)} dW_t^2 + \sigma_3 e^{-\alpha_3(T-t)} dW_t^3, \quad (4.1)$$

where for all $1 \leq i, j \leq 3$, the instantaneous correlation is given by

$$\langle dW^i, dW^j \rangle_t = \begin{cases} dt & \text{if } i = j \\ \rho_{i,j} \cdot dt & \text{if } i \neq j \end{cases}$$

In this model, the spot price is given by

$$S_t = F_{0,t} \cdot \exp \left(\langle \sigma, X_t \rangle - \frac{1}{2} \lambda_t^2 \right),$$

where $\sigma = (\sigma_1, \sigma_2, \sigma_3)^\top$, $X_t = (X_t^1, X_t^2, X_t^3)^\top$ and for all $1 \leq i \leq 3$,

$$X_t^i = \int_0^t e^{-\alpha_i(t-s)} dW_s^i \quad \text{and} \quad \lambda_t^2 = \frac{1}{2} \sum_{i=1}^3 \frac{\sigma_i^2}{\alpha_i} (1 - e^{-2\alpha_i t}) + \sum_{i \neq j} \rho_{i,j} \frac{\sigma_i \sigma_j}{\alpha_i + \alpha_j} (1 - e^{-(\alpha_i + \alpha_j)t}).$$

We fix the following configuration $\sigma_i = \sigma = 0.7, \alpha_i = \alpha = 1.5, \rho_{i,j} = \rho \in [-1, 1]$. The swing contract setting corresponds to case 1. We use $N = 1000$ iterations and a learning rate $\gamma = 0.1$. Results are recorded in Tables 10 and 11.

ρ	\widehat{P}_0	\widehat{P}_0^{BG}
0.6	172.71 ([172.51, 172.90])	172.72 ([172.52, 172.92])
0.3	147.79 ([147.63, 147.95])	147.78 ([147.62, 147.95])
-0.2	91.02 ([90.92, 91.12])	91.02 ([90.92, 91.12])

Table 10: Results using *PV strat.* Values in brackets are confidence intervals (95%). The valuation had been performed with a sample of size $1 \cdot e^8$. For each result, the execution time (training plus testing) is roughly equal to 22s.

ρ	\widehat{P}_0	\widehat{P}_0^{BG}
0.6	173.05 ([172.85, 173.24])	172.98 ([178.78, 173.17])
0.3	147.97 ([147.81, 148.14])	147.92 ([147.75, 148.08])
-0.2	91.18 ([91.08, 91.28])	91.13 ([91.03, 91.23])

Table 11: Results using *NN strat.* Values in brackets are confidence intervals (95%). The valuation had been performed with a sample of size $1 \cdot e^8$. For the neural network architecture we used $I = 2$ layers with $q_1 = q_2 = 10$ units. For each result, the execution time (training plus testing) is roughly equal to 45s.

It should be noted that there is a different approach to implement *NN strat.*, and taking into account the structure of forward prices. Specifically, in the three-factor framework (4.1), the forward price depends on state variables that are components of the \mathbb{R}^3 -valued random vector X_t . In this context, one can include X_{t_k}

as an additional input to the neural network described in (2.4), in addition to time t_k , payoff $S_{t_k} - K$, and volume normalization $M(Q_k)$. That is, using the above notation, $I_k = (t_k, S_{t_k} - K, X_{t_k}, M(Q_k)) \in \mathbb{R}^{d+3}$, where d is the dimension of X_{t_k} (in the three-factor model, $d = 3$). This approach can help to capture the correlation structure in a multi-factor framework. The prices resulting from this alternative approach are presented in Table 12.

ρ	\widehat{P}_0	\widehat{P}_0^{BG}
0.6	172.76 ([172.56, 172.95])	172.70 ([172.50, 172.89])
0.3	147.95 ([147.78, 148.11])	147.91 ([147.75, 148.07])
-0.2	91.11 ([91.01, 91.20])	91.07 ([90.97, 91.17])

Table 12: Results using *NN strat* including state variables. Values in brackets are confidence intervals (95%). The valuation had been performed with a sample of size $1 \cdot e^8$. For the neural network architecture we used $I = 2$ layers with $q_1 = q_2 = 10$ units. For each result, the execution time (training plus testing) is roughly equal to 50s.

It appears that, regardless of whether the model is multi-factor or not, both strategies (*PV strat* and *NN strat*) can determine the optimal volume to purchase at each exercise date, based only on the payoff and the cumulative volume. In the following section, we will evaluate the performance of our strategies on a more complex diffusion model.

4.2 Multi-curve forward diffusion

We finally consider a multi-curve model. At a certain valuation date t we consider p risk factors. Each risk factor $i \in \{1, \dots, p\}$, denoted by F_{t,T_i} , is a forward contract observed at date t and expiring at date T_i . It is modeled through an instantaneous volatility function $\sigma_t(T_i)$ with a dynamics given by,

$$\frac{dF_{t,T_i}}{F_{t,T_i}} = \sigma_t(T_i) dW_t(T_i) \quad (4.2)$$

where $(W_t(T_i))_{t \geq 0}$ is a standard Brownian motion. The instantaneous correlation between Brownian motions are given by

$$\langle dW.(T_i), dW.(T_j) \rangle_t = \rho_{i,j} dt$$

The model (4.2) implies to diffuse one curve per risk factor; F_{t,T_i} defined by its maturity date T_i . Likewise, the spot at a given date generates its own risk factor. Thus if we consider a swing contract on the day-ahead where we can exercise on all working days of a year, then we have about 365 delivery dates. Therefore 365 curves to diffuse. But due to high correlation (at any date, two risk factors are very correlated when maturities are very close), one can see that only a few risk factors really impact the price. In practice forward contracts delivering in the same month are highly correlated. Thus we will consider there are as many factor as months. For example, for a contract on a whole year $p = 12$; therefore 12 (correlated) Brownian motions per exercise date. In this model we consider a valuation date given by 17th march 2021 and a swing contract delivering on the period January 2022-December 2022. The swing constraints are: $q_{\min} = 0, q_{\max} = 1, Q_{\min} = 180, Q_{\max} = 270$. For the diffusion model, the correlation matrix is (*in percentage*)

100	99.62	98.61	91.12	91.12	91.12	90.24	90.24	90.24	87.2	87.2	87.2
99.62	100	98.91	91.65	91.65	91.65	90.68	90.68	90.68	87.8	87.8	87.8
98.61	98.61	100	93.89	93.89	93.89	92.79	92.79	92.79	89.94	89.94	89.94
91.12	91.65	93.89	100	100	100	99.47	99.47	99.47	97.06	97.06	97.06
91.12	91.65	93.89	100	100	100	99.47	99.47	99.47	97.06	97.06	97.06
91.12	91.65	93.89	100	100	100	99.47	99.47	99.47	97.06	97.06	97.06
90.24	90.68	92.79	99.4	99.47	99.4	100	100	100	97.32	97.32	97.32
90.24	90.68	92.79	99.47	99.47	99.47	100	100	100	97.32	97.32	97.32
90.24	90.68	92.79	99.47	99.47	99.47	100	100	100	97.32	97.32	97.32
87.2	87.8	89.94	97.06	97.06	97.06	97.32	97.32	97.32	100	100	100
87.2	87.8	89.94	97.06	97.06	97.06	97.32	97.32	97.32	100	100	100
87.2	87.8	89.94	97.06	97.06	97.06	97.32	97.32	97.32	100	100	100

Volatilities and initial forward prices are assumed to be constant per month. For months in 2022, the square of the instantaneous volatility $\sigma_t^2(T_i)$ is given by Figure 6 and the initial forward curve is given by Figure 7.

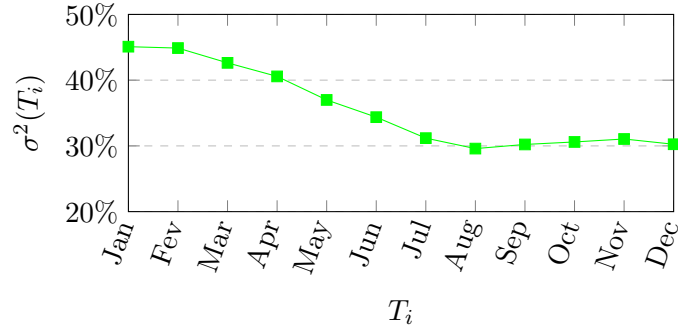


Figure 6: Square of instantaneous volatility per risk factor observed on 17-March-2021. Values are: (45.09%, 44.88%, 42.63%, 40.57%, 36.99%, 34.37%, 31.18%, 29.59%, 30.22%, 30.6%, 31.05%, 30.25%)

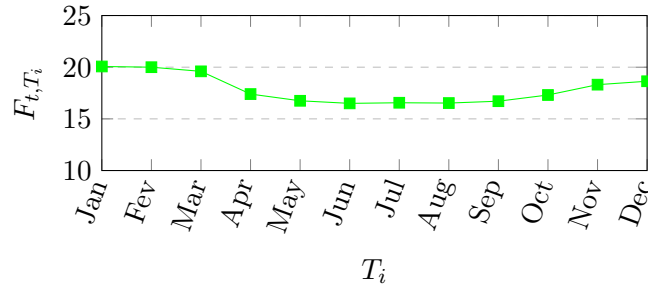


Figure 7: Forward prices per risk factor observed on 17-March-2021. Prices per risk factor are: (20.07, 20, 19.6, 17.4, 16.75, 16.5, 16.56, 16.53, 16.71, 17.31, 18.31, 18.64)

Finally, the strike price corresponds to the average of all initial forward prices; which gives 17.865. For this diffusion model, *PV strat* does not perform well. This is probably due to the multi-curve framework

which is more complex than a multi factor model. However modifying *NN strat* as in the preceding section by adding, in neural network inputs, Brownian motions $(W_{t_k}(T_i))_{1 \leq i \leq p}$ at each exercise date t_k allows to achieve very good prices. We compare the latter implementation of *NN strat* with a method based on a static replication of swing contracts by means of spread options. The latter method aims at approximating swing options prices by selecting a combination of spread options (*SO*) where the underlying is driven by the same dynamics (4.2). The selection is done through a linear programming method. Results are recorded in Table 13.

\widehat{P}_0	\widehat{P}_0^{BG}	<i>SO</i>
708.46 ([705.76, 711.16])	711.32 ([708.61, 714.03])	645.3

Table 13: Results using *NN strat*. Values in brackets are confidence intervals (95%). The valuation had been performed with a sample of size 5×10^6 . For the neural network architecture we used $I = 2$ layers with $q_1 = q_2 = 50$ units.

Conclusion

We introduced two parametric approaches for pricing swing contracts. The first method involves a direct parameterization of the optimal control based on some heuristics, while the second method improves upon the first by using neural networks. We conducted numerical experiments to compare two optimization algorithms (Adam and PSGLD). Our results demonstrate that using Langevin-based optimization algorithms allows us to achieve better prices with short computation time. We also found that our methods outperform the state-of-the-art methods in terms of accuracy. Additionally, we tested our neural network parameterization in various complex diffusion models and demonstrating its robustness. We can hence conclude that our neural network-based method is well-suited to the pricing swing options.

Acknowledgments

The first author is grateful for the financial support provided by Engie Global Markets via CIFRE agreement and would like to thank Asma Meziou and Frederic Muller for throughout insights on swing contracts.

A Swing contract decomposition

We aim to prove that any swing contract can be reduced to a normalized contract, namely a swing contract with local constraints $q_{\min} = 0, q_{\max} = 1$. The swing contract price is given by

$$P_0 = \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} q_\ell \times (S_{t_\ell} - K) \right)$$

where

$$\mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}} = \left\{ (q_\ell)_{0 \leq \ell \leq n-1}, q_\ell : (\Omega, \mathcal{F}_{t_\ell}^S, \mathbb{P}) \mapsto [q_{\min}, q_{\max}], \sum_{\ell=0}^{n-1} q_\ell \in [Q_{\min}, Q_{\max}] \right\}.$$

It follows from the linearity of the expectation that,

$$P_0 = q_{\min} \times \mathbb{E} \left(\sum_{\ell=0}^{n-1} (S_{t_\ell} - K) \right) + (q_{\max} - q_{\min}) \times \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} \frac{q_\ell - q_{\min}}{q_{\max} - q_{\min}} \times (S_{t_\ell} - K) \right).$$

The first term in the last equality is given by

$$\mathbb{E} \left(\sum_{\ell=0}^{n-1} (S_{t_\ell} - K) \right) = \sum_{\ell=0}^{n-1} \mathbb{E} (S_{t_\ell}) - n \cdot K$$

and can be easily computed using either closed formula (depending on the underlying diffusion model) or Monte-Carlo method. Let us consider the second term. Let $(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}$ and define $\tilde{q}_\ell = \frac{q_\ell - q_{\min}}{q_{\max} - q_{\min}}$. Note that $(\tilde{q}_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}}$ where

$$\mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}} = \left\{ (q_\ell)_{0 \leq \ell \leq n-1}, q_\ell : (\Omega, \mathcal{F}_{t_\ell}, \mathbb{P}) \mapsto [0, 1], \sum_{\ell=0}^{n-1} q_\ell \in [\tilde{Q}_{\min}, \tilde{Q}_{\max}] \right\}.$$

and

$$\tilde{Q}_{\min} = \frac{(Q_{\min} - n \cdot q_{\min})_+}{q_{\max} - q_{\min}}, \quad \tilde{Q}_{\max} = \frac{(Q_{\max} - n \cdot q_{\min})_+}{q_{\max} - q_{\min}}.$$

Thus,

$$\begin{aligned} \mathbb{E} \left(\sum_{\ell=0}^{n-1} \frac{q_\ell - q_{\min}}{q_{\max} - q_{\min}} \times (S_{t_\ell} - K) \right) &= \mathbb{E} \left(\sum_{\ell=0}^{n-1} \tilde{q}_\ell \times (S_{t_\ell} - K) \right) \\ &\leq \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} q_\ell \times (S_{t_\ell} - K) \right). \end{aligned}$$

Therefore taking the supremum yields,

$$\operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} \frac{q_\ell - q_{\min}}{q_{\max} - q_{\min}} \times (S_{t_\ell} - K) \right) \leq \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} q_\ell \times (S_{t_\ell} - K) \right).$$

Conversely let $(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}}$ and define $\tilde{q}_\ell = q_{\min} + (q_{\max} - q_{\min}) \cdot q_\ell \in [q_{\min}, q_{\max}]$. It follows $\sum_{\ell=0}^{n-1} q_\ell \in [\tilde{Q}_{\min}, \tilde{Q}_{\max}]$ so that $(\tilde{q}_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}$. Thus,

$$\begin{aligned} \mathbb{E} \left(\sum_{\ell=0}^{n-1} q_\ell \times (S_{t_\ell} - K) \right) &= \mathbb{E} \left(\sum_{\ell=0}^{n-1} \frac{\tilde{q}_\ell - q_{\min}}{q_{\max} - q_{\min}} \times (S_{t_\ell} - K) \right) \\ &\leq \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} \frac{q_\ell - q_{\min}}{q_{\max} - q_{\min}} \times (S_{t_\ell} - K) \right). \end{aligned}$$

Taking the supremum, we get,

$$\operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} q_\ell \times (S_{t_\ell} - K) \right) \leq \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{Q_{\min}, Q_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} \frac{q_\ell - q_{\min}}{q_{\max} - q_{\min}} \times (S_{t_\ell} - K) \right).$$

Therefore,

$$P_0 = q_{\min} \times \mathbb{E} \left(\sum_{\ell=0}^{n-1} (S_{t_\ell} - K) \right) + (q_{\max} - q_{\min}) \times \operatorname{ess\,sup}_{(q_\ell)_{0 \leq \ell \leq n-1} \in \mathcal{A}_{0,0}^{\tilde{Q}_{\min}, \tilde{Q}_{\max}}} \mathbb{E} \left(\sum_{\ell=0}^{n-1} q_\ell \times (S_{t_\ell} - K) \right).$$

B On the rate of convergence

This section aims to estimate, at least numerically, the rate of convergence. We denote by U_n the price obtained after the n^{th} iteration in the training step. We make the assumption that for some constant $C > 0$

$$U_n = U_\infty + \frac{C}{n^\alpha}$$

where U_∞ is the limit of the stochastic procedure that we assume to exist. Thus,

$$\log(|U_{2n} - U_n|) = K_\alpha + \alpha \log\left(\frac{1}{n}\right), \quad \text{where } K_\alpha = \log(|C(2^{-\alpha} - 1)|).$$

Therefore the coefficient α (representing the rate of convergence following the assumption) appears to be the slope in the log-log regression of $|U_{2n} - U_n|$ against $\frac{1}{n}$.

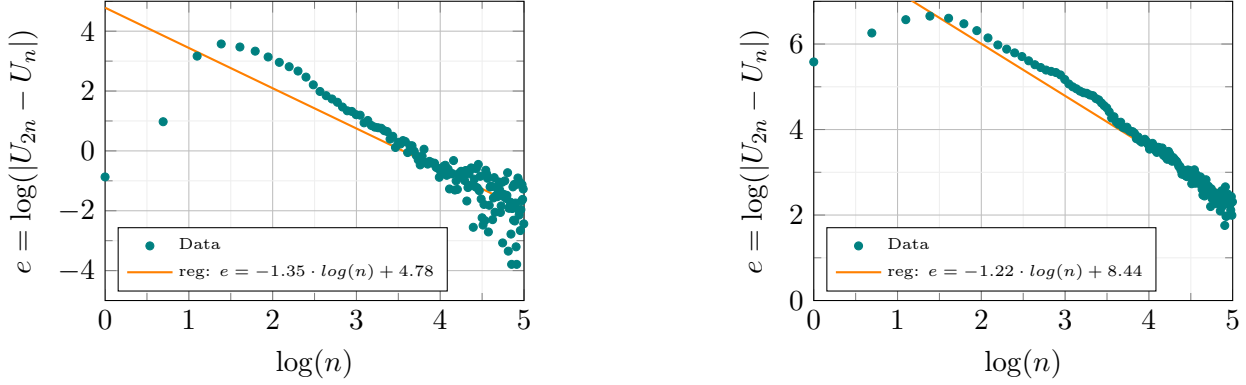


Figure 8: Numerical Convergence (logarithmic scale) for *PV strat*. Case 1 (left), case 2 (right)

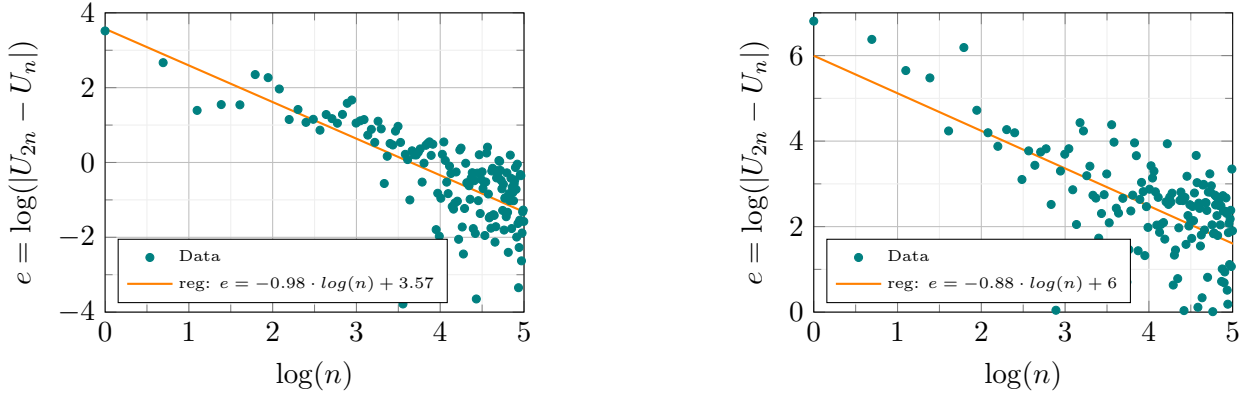


Figure 9: Numerical Convergence (logarithmic scale) *NN strat*. Case 1 (left), case 2 (right)

From figures 8 and 9 it suggests that both parameterizations give a rate of convergence of order $\mathcal{O}(\frac{1}{N})$ (where N is the number of iterations). This convergence rate is much faster than Monte-Carlo, and suggests that our parameterizations are efficient alternatives to Longstaff-Schwartz method for this problem.

C Estimator variance and computation time

In this section we illustrate the high variance phenomenon which may appear when pricing swing option. To this end, we compute 100 realizations of swing contract price with case 1 setting for the three methods: *PV strat*, *NN strat* and Longstaff-Schwartz. The distributions of prices are represented in Figures 10, 11, 12.

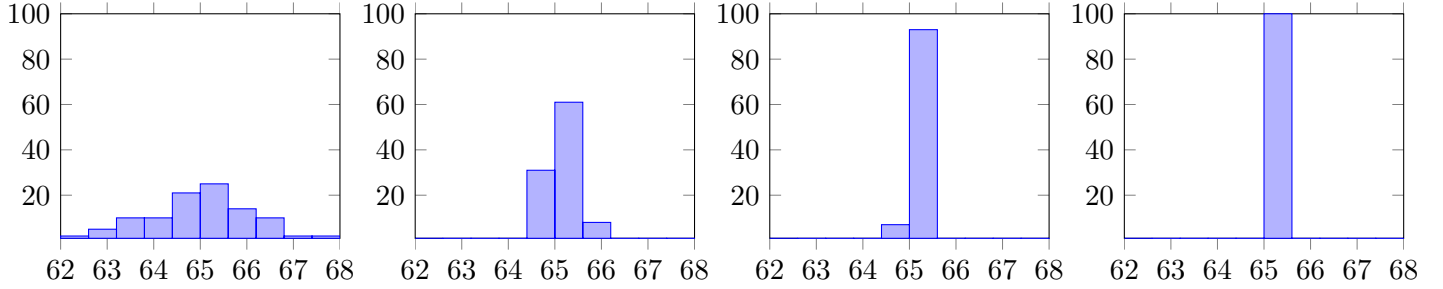


Figure 10: Distribution of swing prices using *PV strat*. From left to right we used successively $10^5, 10^6, 10^7, 5 \times 10^8$ simulations.

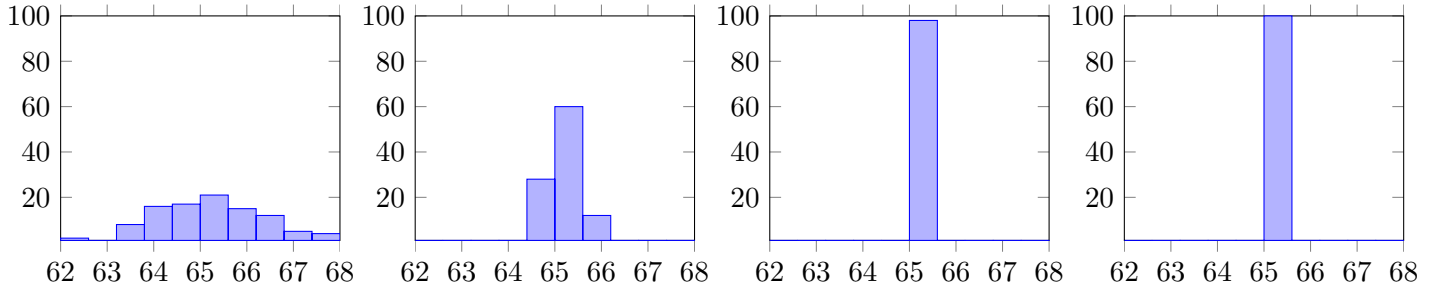


Figure 11: Distribution of swing prices using *NN strat*. From left to right we used successively $10^5, 10^6, 10^7, 5 \times 10^8$ simulations.

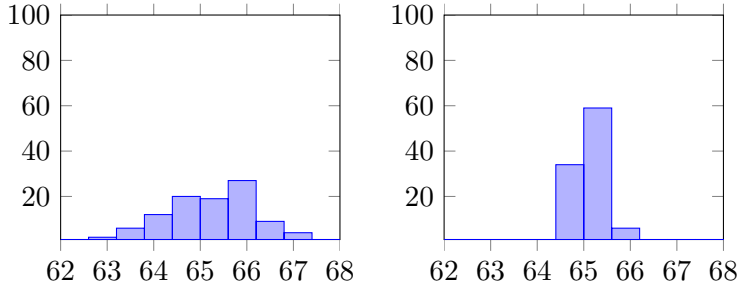


Figure 12: Distribution of swing prices using Longstaff-Schwartz method. From left to right we used successively $10^5, 10^6$ simulations. Higher number of simulations leads to memory overflow.

Figures 10, 11, and 12 demonstrate that prices can significantly fluctuate, regardless of the pricing method. Using 1000000 simulations is insufficient to obtain a reliable price estimate. Storage limitations prevent the Longstaff-Schwartz method from exceeding this number of simulations, and this limit is even lower when the number of exercise dates is high. However, our proposed methods allow to increase the number of simulations as needed once the strategy trained. This is possible because we can evaluate our strategies on mini-batches sequentially instead of using the entire test set. For instance, to compute a price with a sample size of 10^8 , we estimate 100 prices over 100 mini-batches, each with a sample size of 10^6 . The final price is then the average of the 100 estimated prices.

Hereafter (see Figure 13) we present CPU time for *PV strat* and *NN strat*. We use one batch with size of 2^{14} and the same PSGLD setting.

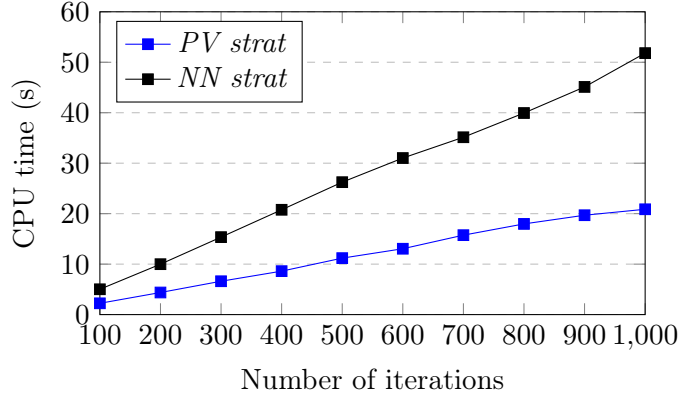


Figure 13: CPU time (in seconds) as a function of the number of iterations.

D *PV strat* coefficients

In this appendix we present the coefficients obtained with *PV strat* and either Adam or PSGLD updating. For that purpose, we used $N = 1000$ iterations, a learning rate of 0.1 and the following configurations. For Adam, we use $B = 4$ batches of size 2^{12} . For the PSGLD updating, we use $\sigma = 1 \cdot e^{-6}$, $\beta = 0.8$, $\lambda = 1 \cdot e^{-10}$ and $B = 1$ batch of size 2^{14} . In the following graphics, coefficients a_k, b_k represent the coefficients which multiply respectively the payoff $S_{t_k} - K$, the margin of cumulative consumption $M(Q_k)$. The coefficient c_k is the constant (see (2.3)). For both optimization algorithms, the strategy is estimated with the same sample.

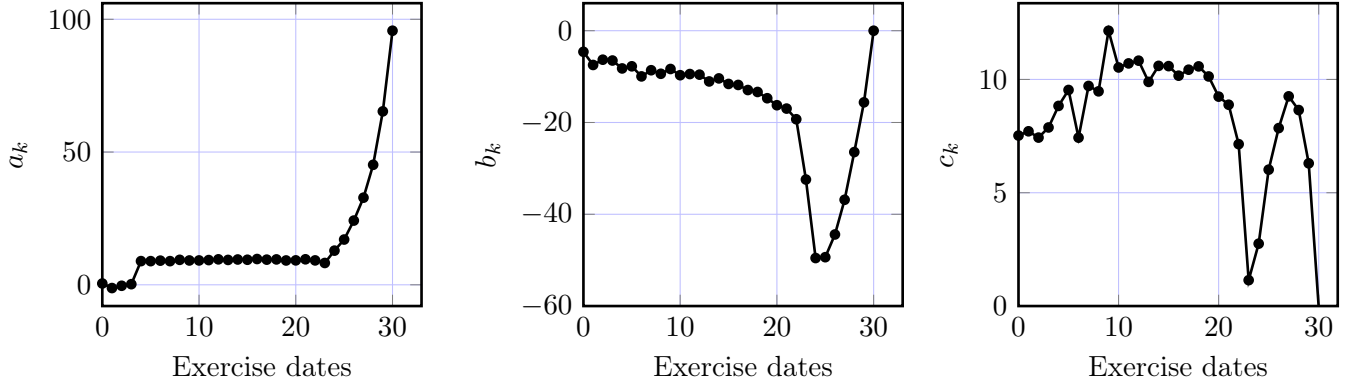


Figure 14: Coefficients of *PV strat* using PSGLD updating.

In general, it can be observed that the coefficients estimated using the PSGLD algorithm are smoother across the exercise dates compared to those estimated with the Adam updating method. Additionally, *PV strat* results in interpretable coefficients. Specifically, the coefficient a_k tends to increase globally, with a rapid growth over the last exercise dates. This is because, as we move forward in exercise dates, the minimal global constraint is gradually fulfilled, giving us more flexibility on the control or allowed volumes to purchase. In this particular example, the maximum global constraint is always fulfilled, so towards the last exercise dates, we can decide to buy the minimum possible amount if the payoff is negative and buy the maximum possible amount if the payoff is positive (i.e., the parameterization (2.3) approaches 0 or 1).

On the other hand, the coefficient b_k generally decreases at the beginning of the contract and starts increasing towards zero over the last exercise dates. This indicates that the remaining exercise capacity has less influence on the strategy towards the end of the contract. This is consistent with the claim made earlier, which stated that towards the end of the contract, the optimal strategy depends mainly on the payoff and less on the volume constraints.

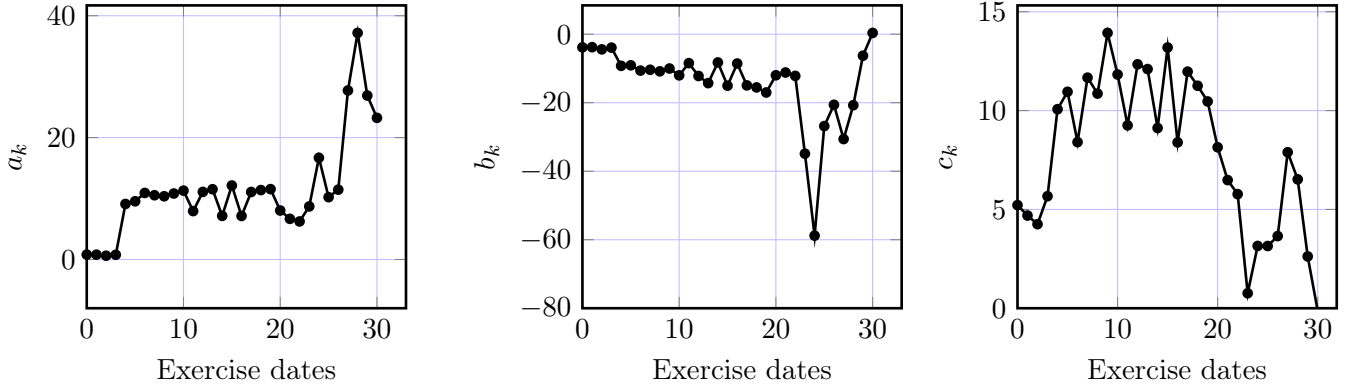


Figure 15: Coefficients of *PV strat* using Adam updating.

E Summary tables: Adam versus PSGLD

Hereafter are recorded all the results obtained using Adam and PSGLD updatings with different hyperparameters combinations. Prices are computed by averaging 100 replications of prices; each obtained using a sample size of 10^6 . We consider a swing contract with case 1 setting. Results are recorded in tables 14, 15, 16, 17.

$L \times B$	γ	N	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
$2^{14} \times 1$	0.1	3000	65.11 ([65.04, 65.18])	65.15 ([65.08, 65.22])	68.4
$2^{14} \times 1$	0.1	1000	65.02 ([64.95, 65.08])	65.18 ([65.12, 65.24])	21.3
$2^{12} \times 4$	0.1	3000	65.14 ([65.08, 65.21])	65.13 ([65.07, 65.20])	261.2
$2^{12} \times 4$	0.1	1000	65.15 ([65.08, 65.22])	65.17 ([65.11, 65.24])	88.8
$2^{14} \times 1$	0.01	3000	64.73 ([64.63, 64.79])	65.05 ([64.98, 65.11])	64.4
$2^{14} \times 1$	0.01	1000	63.91 ([63.83, 63.98])	64.70 ([64.62, 64.77])	22.1
$2^{12} \times 4$	0.01	3000	65.11 ([65.04, 65.18])	65.15 ([65.08, 65.22])	258.2
$2^{12} \times 4$	0.01	1000	64.81 ([64.74, 64.88])	65.03 ([64.98, 65.12])	88.3

Table 14: Summary table for *PV strat* using Adam. The values in brackets are confidence intervals (95%). The column “time” includes both training and valuation time.

$L \times B$	γ	N	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
$2^{14} \times 1$	0.1	3000	65.22 ([65.16, 65.28])	65.23 ([65.17, 65.29])	150.4
$2^{14} \times 1$	0.1	1000	65.13 ([65.05, 65.20])	65.22 ([65.15, 65.29])	53.2
$2^{12} \times 4$	0.1	3000	65.26 ([65.21, 65.32])	65.23 ([65.17, 65.29])	603.4
$2^{12} \times 4$	0.1	1000	65.20 ([65.14, 65.26])	65.20 ([65.14, 65.26])	133.5
$2^{14} \times 1$	0.01	3000	64.78 ([62.65, 64.90])	65.06 ([64.98, 65.14])	157.1
$2^{14} \times 1$	0.01	1000	64.81 ([64.74, 64.88])	65.05 ([64.98, 65.12])	55.1
$2^{12} \times 4$	0.01	3000	65.20 ([65.15, 65.26])	65.20 ([65.14, 65.26])	607.3
$2^{12} \times 4$	0.01	1000	65.02 ([64.92, 65.11])	65.17 ([65.09, 65.24])	135.4

Table 15: Summary table for *NN strat* using Adam. We used a neural network architecture as follows: 2 hidden layers ($I = 2$) and 10 units per layer ($q_1 = 10, q_2 = 10$). The values in brackets are confidence intervals (95%). The time includes the training and the valuation time.

σ	β	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
$1e^{-5}$	0.8	65.15 ([65.09, 65.22])	65.17 ([65.10, 65.23])	22.6
$1e^{-5}$	0.7	65.13 ([65.05, 65.21])	65.15 ([65.07, 65.22])	22.1
$1e^{-5}$	0.9	65.22 ([65.16, 65.28])	65.23 ([65.18, 65.29])	21.6
$1e^{-6}$	0.8	65.19 ([65.12, 65.26])	65.21 ([65.13, 65.28])	21.3
$1e^{-6}$	0.9	65.21 ([65.15, 65.28])	65.23 ([65.17, 65.29])	21.4

Table 16: Summary table for *PV strat* using PSGLD. The values in brackets are confidence intervals (95%). The column “time” includes both training and valuation time. We used a learning rate equal to 0.1, one batch of size 2^{14} , $\lambda = 1 \cdot e^{-10}$ and 1000 iterations.

σ	β	\widehat{P}_0	\widehat{P}_0^{BG}	Time (s)
$1e^{-5}$	0.8	65.18 ([65.12, 65.24])	65.16 ([65.10, 65.22])	53.6
$1e^{-5}$	0.7	65.22 ([65.15, 65.30])	65.20 ([65.13, 65.27])	52.1
$1e^{-5}$	0.9	65.23 ([65.14, 65.31])	65.22 ([65.13, 65.30])	52.3
$1e^{-6}$	0.8	65.27 ([65.20, 65.35])	65.26 ([65.18, 65.33])	51.3
$1e^{-6}$	0.9	65.23 ([65.16, 65.30])	65.21 ([65.14, 65.28])	53.4

Table 17: Summary table for *NN strat* using PSGLD. We used $I = 2$ layers with $q_1 = q_2 = 10$ units. The values in brackets are confidence intervals (95%). The column “time” includes both training and valuation time. We used a learning rate equal to 0.1, one batch of size 2^{14} , $\lambda = 1 \cdot e^{-10}$ and 1000 iterations.

References

- [1] Olivier Bardou, Sandrine Bouthemy, and Gilles Pagès. When are swing options bang-bang and how to use it. *arXiv: Probability*, 2007.
- [2] Olivier Bardou, Sandrine Bouthemy, and Gilles Pagès. Optimal quantization for the pricing of swing options. *Applied Mathematical Finance*, 16:183 – 217, 2009.
- [3] Christophe Barrera-Esteve, Florent Bergeret, Charles Dossal, Emmanuel Gobet, Asma Meziou, Rémi Munos, and Damien Reboul-Salze. Numerical methods for the pricing of swing options: A stochastic control approach. *Methodology and Computing in Applied Probability*, 8:517–540, 2006.
- [4] Atilim Gunes Baydin, Barak A. Pearlmutter, Alexey Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *J. Mach. Learn. Res.*, 18:153:1–153:43, 2017.
- [5] L Eon Bottou. Online learning and stochastic approximations. 1998.
- [6] Pierre Bras. Langevin algorithms for very deep neural networks with application to image classification, 2022.
- [7] Pierre Bras and Gilles Pagès. Convergence of Langevin-Simulated Annealing algorithms with multiplicative noise. 31 pages + Supplementary Material (6 pages), September 2021.
- [8] Pierre Bras and Gilles Pagès. Langevin algorithms for Markovian Neural Networks and Deep Stochastic control. *arXiv e-prints*, page arXiv:2212.12018, December 2022.
- [9] Mark Broadie and Paul Glasserman. A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, 7:35–72, 2004.
- [10] Jerry Chee and Panos Toulis. Convergence diagnostics for stochastic gradient descent with constant learning rate. In Amos Storkey and Fernando Perez-Cruz, editors, *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pages 1476–1485. PMLR, 09–11 Apr 2018.
- [11] Arnak S. Dalalyan. Theoretical guarantees for approximate sampling from smooth and log-concave densities. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79, 2014.
- [12] Arnak S. Dalalyan. Further and stronger analogy between sampling and optimization: Langevin monte carlo and gradient descent. In *Annual Conference Computational Learning Theory*, 2017.
- [13] Yann Dauphin, Razvan Pascanu, Çağlar Gülçehre, Kyunghyun Cho, Surya Ganguli, and Yoshua Bengio. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. In *NIPS*, 2014.
- [14] Alexandre Défossez, Leon Bottou, Francis Bach, and Nicolas Usunier. A simple convergence proof of adam and adagrad. *Transactions on Machine Learning Research*, 2022.
- [15] Alain Durmus and Éric Moulines. Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *The Annals of Applied Probability*, 27(3):1551 – 1587, 2017.
- [16] Alain Durmus and Eric Moulines. High-dimensional bayesian inference via the unadjusted langevin algorithm, 2018.
- [17] Emmanuel Gobet and Remi Munos. Sensitivity analysis using itô–malliavin calculus and martingales, and application to stochastic optimal control. *SIAM J. Control and Optimization*, 43:1676–1713, 01 2005.
- [18] Kurt Hornik, Maxwell B. Stinchcombe, and Halbert L. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.
- [19] Patrick Jaillet, Damien Lambertson, and Bernard Lapeyre. Variational inequalities and the pricing of american options. *Acta Applicandae Mathematica*, 21:263–289, 1990.
- [20] Patrick Jaillet, Ehud I. Ronn, and Stathis Tompaidis. Valuation of commodity-based swing options. *Manag. Sci.*, 50:909–921, 2004.

- [21] Laurent Valentin Jospin, Hamid Laga, Farid Boussaid, Wray Buntine, and Mohammed Bennis. Hands-on bayesian neural networks—a tutorial for deep learning users. *IEEE Computational Intelligence Magazine*, 17(2):29–48, may 2022.
- [22] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2015.
- [23] Harold J. Kushner and George Yin. Stochastic approximation and recursive algorithms and applications. 2003.
- [24] Ali Lari-Lavassani and Mohamadreza Simchi. A discrete valuation of swing options. 2002.
- [25] Chunyuan Li, Changyou Chen, David Edwin Carlson, and Lawrence Carin. Preconditioned stochastic gradient langevin dynamics for deep neural networks. In *AAAI*, 2016.
- [26] Ziyin Liu, Botao Li, and Masahito Ueda. SGD may never escape saddle points. *CoRR*, abs/2107.11774, 2021.
- [27] Francis A. Longstaff and Eduardo S. Schwartz. Valuing american options by simulation: A simple least-squares approach. *The Finance*, 2001.
- [28] Arvind Neelakantan, Luke Vilnis, Quoc V. Le, Ilya Sutskever, Lukasz Kaiser, Karol Kurach, and James Martens. Adding gradient noise improves learning for very deep networks, 2015.
- [29] Gilles Pagès and Fabien Panloup. Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. *Ann. Appl. Probab.*, 33(1):726–779, 2023.
- [30] Gilles Pagès, Huyen Pham, and Jacques Printems. Optimal quantization methods and applications to numerical problems in finance. 2004.
- [31] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22:1345–1359, 2010.
- [32] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zach DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.
- [33] Ning Qian. On the momentum term in gradient descent learning algorithms. *Neural networks : the official journal of the International Neural Network Society*, 12:145–151, 02 1999.
- [34] Herbert E. Robbins. A stochastic approximation method. *Annals of Mathematical Statistics*, 22:400–407, 2007.
- [35] Herbert E. Robbins and David O. Siegmund. A convergence theorem for non negative almost supermartingales and some applications**research supported by nih grant 5-r01-gm-16895-03 and onr grant n00014-67-a-0108-0018. 1971.
- [36] L. C. G. Rogers. Monte carlo valuation of american options. *Mathematical Finance*, 12, 2002.
- [37] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning internal representations by error propagation. 1986.
- [38] Andrew Carl Thompson. Valuation of path-dependent contingent claims with multiple exercise decisions over time: The case of take-or-pay. *Journal of Financial and Quantitative Analysis*, 30:271 – 293, 1995.
- [39] Max Welling and Yee Whye Teh. Bayesian learning via stochastic gradient langevin dynamics. In *International Conference on Machine Learning*, 2011.
- [40] Fangyu Zou, Li Shen, Zequn Jie, Weizhong Zhang, and Wei Liu. A sufficient condition for convergences of adam and rmsprop. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11119–11127, 2018.