HAL open science

The Parareal Algorithm for American Options

Gilles Pagès, Olivier Pironneau, Guillaume Sall

To cite this version:

Gilles Pagès, Olivier Pironneau, Guillaume Sall. The Parareal Algorithm for American Options. Comptes Rendus. Mathématique, 2016, Série I Mathémtiques, 354 (11), pp.1132-1138. hal-01320331

HAL Id: hal-01320331

https://hal.sorbonne-universite.fr/hal-01320331v1

Submitted on 23 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Analyse Numérique

The Parareal Algorithm for American Options

La méthode pararéelle pour les options américaines

Gilles Pagès [a], Olivier Pironneau [b] Guillaume Sall [a]

[a] *Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, Case 188, 4 pl. de Jussieu, F-75252 Paris Cedex 5.*
[b] *Laboratoire Jacques Louis Lions, UMR 7598, Case 187, 4 pl. de Jussieu, F-75252 Paris Cedex 5.*

Reçu le 22 Mi 2016 ; accepté après révision le +++++

Présenté par ALain Bensoussan

Résumé

La méthode pararéelle pour les options américaines. Dans cette note la méthode pararéelle est introduite pour l'algorithme LSMC de Longstaff-Schartz pour calculer des options américaines sur une machine parallèle. Dans une section numérique les performances de la méthode sont données dans le cas scalaire à deux niveaux d'abord puis multi-niveaux. Un théorème de convergence est aussi énoncé lorsque la méthode d'Euler explicite est utilisée avec un pas de temps $\Delta t > \delta t$ le pas de temps de la grille fine. Une estimation est obtenue qui permet d'analyser la méthode pararéelle multi-niveaux. *Pour citer cet article : G. Pagès, O. Pironneau, G. Sall, C. R. Acad. Sci. Paris, Ser. I ? ? ? (2016).*

Abstract

The Parareal Algorithm for American Options. This note provides a description of the parareal method, a numerical section to assess the performance of the method for American contracts in the scalar case computed by LSMC and parallelized by parareal time decomposition with two or more levels. It contains also a convergence proof for the two levels parareal Monte-Carlo method when the coarse grid solution is computed by an Euler explicit scheme with time step $\Delta t > \delta t$, the time step used for the Euler scheme at the fine grid level. Hence the theorem provides a tool to analyze also the multilevel parareal method. *To cite this article: G. Pagès, O. Pironneau, G. Sall, C. R. Acad. Sci. Paris, Ser. I ??? (2016).*

Keywords Financial securities, risk assessment, American options, LSMC, Parareal. AMDS: 37M25, 65N99

1. Introduction

In quantitative finance risk assessment is computer intensive and expensive and there is a market for cheaper and faster methods as seen from the large literature on parallelism and GPU implementation of numerical methods for option pricing [1,6,7,10,11,13,14,15,20].

American contracts are not easy to compute on a parallel computer; even if a large number of them have to be computed at once, an embarrassingly parallel problem, still the cost of the transfer of data makes parallelism at the level of one contract attractive. But the task is not easy, especially when the number of underlying assets is large [3,21], ruling out the PDE approach [2]. Furthermore the most popular sequential algorithm is the Least Square Monte-Carlo (LSMC) method of Longstaff and Schwartz [18]. Exploiting parallelism by allocating blocks of Monte-Carlo paths to different processors is not convincingly efficient [7] because the backward regression is essentially sequential and needs all Monte-Carlo paths in the same processor.

In this note we investigate the parareal method, introduced in [17], for the task. An earlier study by Bal and Maday [4] has paved the way but it is restricted to Stochastic Differential Equations (SDE) without LSMC. Yet it contains a convergence proof for the two levels method in the restricted case where the solution is computed exactly at the lowest level [4].

This note provides a description of the method, a numerical section to assess the performance of the method for American contracts in the scalar case computed by LSMC and parallelized by parareal time decomposition. It contains also a convergence proof for the two levels parareal Monte-Carlo method when the coarse grid solution is computed by an Euler explicit scheme with time step $\Delta t > \delta t$, the time step used for the Euler scheme at the fine grid level. Hence the theorem provides a tool to analyze also the multilevels parareal method.

Convergence of LSMC for American contracts has been proved by Clement, Lamberton and Protter [9]; it is not unreasonable to expect an extension of their estimates for the parareal method but this note does not contain such a result, only a numerical assessment.

2. The Problem

With the usual notations [16] consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and functions $b, \sigma, f : [0,T] \times \mathbb{R} \mapsto \mathbb{R}$, uniformly Lipschitz continuous in x, t.

Let $W = (W_t)_{t\in[0,T]}$ be a standard Brownian motion on $(\Omega, \mathcal{A}, \mathbb{P})$. Let $X = (X_t)_{t\in[0,T]},\ X_t \in \mathbb{R}$, be a diffusion process, strong solution of the SDE

$$dX_t = b\,(t, X_t)\,dt + \sigma(t, X_t)dW_t, \quad X(0) = X_0 \in \mathbb{R}. \tag{1}$$

A (vanilla) European contract on X is defined by its maturity T and its payoff $\mathbb{E}[f(T, X_T)]$, typically $f(t,x) = e^{r(T-t)}(\kappa - x)^+$ in the case of a put of strike price κ and interest rate r. An American style contract allows the owner to claim the payoff $f(t, X_t)$ at any time $\in [0,T]$. So a rational strategy to

Email addresses: `gilles.pages@upmc.fr` (Gilles Pagès), `olivier.pironneau@upmc.fr` (Olivier Pironneau), `guillaume.sall@upmc.fr` (Guillaume Sall).

maximize the average profit $\mathcal{V}$ at time t is to find the $[t,T]$-valued $\mathcal{F}$-stopping time solution of the*Snell envelope* problem:

$$\mathcal{V}(t,X_t) := \mathbb{E}[e^{-r(\tau_t-t)}f(\tau_t,X_{\tau_t})|\mathcal{F}_t] = \mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}_t^{\mathcal{F}}}\mathbb{E}[e^{-r(\tau-t)}f(\tau,X_\tau)|\mathcal{F}_t]$$

where $\mathcal{F} = (\mathcal{F}_t)_{t\in(0,T)}$ is the (augmented) filtration of W and $\mathcal{T}_t^{\mathcal{F}}$ denotes the setof $[t,T]$-valued $\mathcal{F}$-stopping times. Such an optimal stopping times exists (see [8]).We do not specify b, σ or f to stay in a general Optimal Stopping framework. In practice American style options are replaced by so-called *Bermuda* options where the exercise instants are restricted to a time grid $t_k = kh$, $k = 0 : K$ where $h = \frac{T}{K}$ ($K \in \mathbb{N}^*$). Owing to the Markov property of $\{X_{t_k}\}_{k=0}^K$, the corresponding Snell envelope reads $(V(t_k,X_{t_k}))_{k=0:K}$ and satisfies a Backward Dynamic Programming recursion on k:

$$V(T,X_T) = f(T,X_T),\ \ V(t_k,X_{t_k}) = \max\{f(t_k,X_{t_k}), e^{-rh}\mathbb{E}[V(t_{k+1},X_{t_{k+1}})|X_{t_k}]\},\ k = K-1,\dots,0. \quad (2)$$

The optimal stopping times τ_k (from time t_k) are given by a similar backward recursion:

$$\tau_K = T,\ \tau_k = t_k \text{ if } f(t_k,X_{t_k}) > e^{-rh}\mathbb{E}[V(t_{k+1},X_{t_{k+1}})|X_{t_k}],\ \tau_k = \tau_{k+1} \text{ otherwise},\quad k = K-1,\dots,0 \quad (3)$$

When $(X_{t_k})_{k=0:K}$ cannot be simulated at a reasonable computational cost, it can be approximated by the Euler scheme with step h, denoted $(\bar{X}^h_{t_k})_{k=0:K}$, which is a simulable Markov chain recursively defined by

$$\bar{X}^h_{t_{k+1}} = \bar{X}^h_{t_k} + b(t_k,\bar{X}^h_{t_k})h + \sigma(t_k,\bar{X}^h_{t_k})\Delta W,\quad \bar{X}^h_0 = X_0,\ k = 0,\dots,K-1, \quad (4)$$

where $\Delta W := W_{t_{k+1}} - W_{t_k} = \sqrt{h}\,Z_k$ so that $\{Z_k\}_{k=0}^{K-1}$ are i.i.d. $\mathcal{N}(0,1)$-distributed random variables. From now on we switch to the Euler scheme, its Snell envelope, etc.

In LSMC for each k, the conditional expectation $\mathbb{E}[V(t_{k+1},\bar{X}^h_{t_{k+1}})|\bar{X}^h_{t_k}]$ as a function of $x = \bar{X}^h_{t_k}$, is approximated by its projection on the linear space spanned by the monomials $\{x^p\}_{p=0}^P$ from the values $\{e^{-rh}V(t_{k+1},\bar{X}^{h,(m)}_{t_{k+1}})\}_{m=1}^M$ generated by M Monte-Carlo paths using (4); then each path has its own optimal stopping time at each $k\in\{0,\dots,K-1\}$ given by (for the stopping problem starting at k)

$$\tau_K^{(m)} = T,\ \ \tau_k^{(m)} = t_k \text{ if } f(t_k,\bar{X}^{h,(m)}_{t_k}) > e^{-rh}\sum_0^P \bar{a}_k^p\,(\bar{X}^{h,(m)}_{t_k})^p,\ \ \tau_k^{(m)} = \tau_{k+1}^{(m)} \text{ otherwise}$$

where

$$\{\bar{a}_k^o,\dots,\bar{a}_k^P\} = \arg\min_{\{(a^0,\dots,a^P)\in\mathbb{R}^{P+1}\}}\sum_1^M\Big(V\big(t_{k+1},\bar{X}^{h,(m)}_{t_{k+1}}\big) - \sum_0^P a_p\big(\bar{X}^{h,(m)}_{t_k}\big)^p\Big)^2.$$

Finally the price of the American contract is computed by

$$V(0,X_0) \approx \max\{f(0,X_0), \frac{1}{M}\sum_{m=1}^M e^{-r\tau_1^{(m)}}f(\tau_1^{(m)},\bar{X}^{h,(m)}_{\tau_1^{(m)}})\}.$$

Note that $\sum_0^P \bar{a}_k^p\,(\bar{X}^h_{t_k})^p$ is the best approximation of $\mathbb{E}[V(t_{k+1},\bar{X}^h_{t_{k+1}})|\bar{X}^h_{t_k}]$ in least square sense in the vector subspace$\langle(\bar{X}^h_{t_k})^p,\ p = 0 : P\rangle$ of $L^2(\mathbb{P})$.

3. A Two Level Parareal Algorithm

3.1. *The Parareal Method*

Consider an ODE

$$\dot{x} = f(x,t),\ x(0) = x_0,\ t\in[t_0,t_K] = \cup_0^{K-1}[t_k,t_{k+1}].$$

Assume that $G_\delta(x_k, t_k)$ is a high precision integrator which computes x at t_{k+1} from x_k at t_k. Assume G_Δ is a similar integrator but of low precision. The parareal algorithm is an iterative process with $n = 0, \ldots, N-1$ above a forward loop in time, $k = 0, \ldots, K-1$

$$x_{k+1}^{n+1} = G_\Delta(x_k^{n+1}, t_k) + G_\delta(x_k^n, t_k) - G_\Delta(x_k^n, t_k). \tag{5}$$

So the coarse grid solution is corrected by the difference between the fine grid prediction computed from the old value on that interval and the coarse grid old solution. In this analysis G_δ and G_Δ are Euler explicit schemes with time step $\delta t < \Delta t$ respectively.

The same method can be applied to an SDO in the context of the Monte-Carlo method provided the random variables $\{Z_{k,j}^m\}_{j:=1:J-1,k:=1,K}^{m=1:M}$ defining ΔW in (4) are sampled once and for all in the initial phase of the algorithm and reused for all n (see the initializaton step in algorithm 3.2 for the notations).

The iterative process (5) is applied on each sample path with G_Δ a single step of (4) with $h = \Delta t$ and G_δ the result of J steps of (4) with $h = \delta t$. An error analysis is available in [4] for the stochastic case in the limit case $\delta t = 0$, i.e. when the fine integrator is the exact solution. For other problems with parareal see [17] and [12]. In this note we also extend the result of [4] to the case $0 < \delta t < \Delta t$.

Theorem 3.1 *Assume $b, \sigma : [0,T] \times \mathbb{R}$ continuous, C^2 in x with spatial derivatives uniformly Lipschitz in $t \in [0,T]$. Then there exist C, independent of $k, \Delta t$ and n, such that for $k = 0:K,\ n = 0:N$*

$$\|\hat{X}_{t_k}^n - \bar{X}_{t_k}^\delta\|_{L^2(\mathbb{P})} \le (C\Delta t)^n \sqrt{\binom{k}{n}} \|\bar{X}_{t_k}^\Delta - \bar{X}_{t_k}^\delta\|_{L^2(\mathbb{P})} \le (C\Delta t)^n \sqrt{\binom{k}{n}} \sqrt{\Delta t}. \tag{6}$$

Furthermore $\hat{X}_{t_k}^n = \bar{X}_{t_k}^\delta$ for all $n \ge k$.

Corollary 3.2 *For a fixed δt and n parareal iterations, the final and uniform errors satisfy*

$$\|\hat{X}_T^n - \bar{X}_T^\delta\|_{L^2(\mathbb{P})} \le (C\Delta t)^{\frac{n}{2}} \sqrt{\frac{\Delta t}{n!}} \quad \text{and} \quad \Big\| \max_{k=0:K} |\hat{X}_{t_k}^n - \bar{X}_{t_k}^\delta| \Big\|_{L^2(\mathbb{P})} \le \frac{(C\Delta t)^{\frac{n}{2}}}{\sqrt{(n+1)!}} \tag{7}$$

respectively where C only depends on the Lipschitz constants and norms of $b, b', b'', \sigma, \sigma', \sigma''$ and on T.

This estimate shows that when Δt is smaller than C the method converges exponentially in n and geometrically in Δt.

Remark 1 The estimate (3.1) indicates that a recursive use of parareal with each sub-interval redivided into $J = O(\Delta t^{-1})$ smaller intervals, the so-called multilevels parareal, is better than many iterations at the second level only. Indeed, as the error decreases proportionally to $(\Delta t)^{\frac{n}{2}}$ at each level and as Δt becomes Δt^2 at the next grid level, the error after L levels is decreased by $(\Delta t)^{\frac{nL}{2}}$.

3.2. *Algorithm*

We denote by V_k a realization of $V(t_k, X_{t_k}),\ k = 0 : K = \frac{T}{\Delta t}$; consider a refinement of each interval $(t_k, t_{t_{k+1}})$ by a uniform sub-partition of time step $\delta t = \frac{\Delta t}{J}$, for some integer $J > 1$.Then

$$[t_k, t_{k+1}] = \cup_{j=0}^{J-1} [t_{k,j}, t_{k,j+1}] \text{ with } t_{k,j+1} = t_{k,j} + \delta t,\ j = 0, \ldots, J-1, \text{ so that } t_k = t_{k,0} = t_{k-1,J}.$$

Denote by $\mathfrak{P}f$ the projection of f on the monomials $1, x, \ldots, x^P$.

Let $n = 0, \ldots, N-1$ be the iteration index of the parareal algorithm.

Initialization Generate $\{Z_{k,j}^m\}_{k=1:K,j=1:J}^{m=1:M}$ for the M paths of the Monte-Carlo method with the coarse and fine mesh.

Compute recursively forward all Monte-Carlo paths $\{\hat{X}^0_{t_k}(\omega^m)\}_{m=1}^M$ from $\hat{X}^0_0 = X_0$ by using (4) with $h = \Delta t$ and then recursively backward $\hat{V}^0_k = \max\{f(t_k, \hat{X}^0_{t_k}), e^{-r\Delta t}\mathfrak{P}\,\mathbb{E}[\hat{V}^0_{k+1}|\hat{X}^0_{t_k}]\}$, $k = K-1:0$ from $\hat{V}^0_K(\omega^m) = e^{-rT}f(T, \hat{X}^0_T(\omega^m))$, $m = 1:M$.

for n=0:N-1

for all M paths,

for k=0:K-1 (forward loop):

(i) Compute the fine grid solution $\{\tilde{X}^{\delta,n}_{t_{k,j}}\}_{j=0}^J$ of (4) with refined step $h = \delta t = \frac{\Delta t}{J}$, started at $t_{k,0} = t_k$ from $\hat{X}^n_{t_k}$.

(ii) Compute the coarse grid solution at t_{k+1}: $\bar{X}^{\Delta}_{t_{k+1}} = \hat{X}^{n+1}_{t_k} + b(t_k, \hat{X}^{n+1}_{t_k})\Delta t + \sigma(t_k, \hat{X}^{n+1}_{t_k})\Delta W$.

(iii) Set $\hat{X}^{n+1}_{t_{k+1}} = \bar{X}^{\Delta}_{t_{k+1}} + \tilde{X}^{\delta,n}_{t_{k,J}} - \hat{X}^n_{t_{k+1}}$.

end k-loop

end M-loop.

initialization : Compute $\bar{V}^{n+1}_K = \hat{V}^{n+1}_K = f(T, \hat{X}^{n+1}_T)$

for k=K-1,...,0 (backward loop):

(i) On each (t_k, t_{k+1}), from $\tilde{V}^{\delta,n}_{k,J} = \mathfrak{P}\,\mathbb{E}(\hat{V}^n_{k+1} \,|\, \tilde{X}^{\delta,n}_{k,J})$, compute by a backward loop in j

$$\tilde{V}^{\delta,n}_{k,j} = \max\left\{f(t_{k,j}, \tilde{X}^{\delta,n}_{t_{k,j}}), e^{-r\delta t}\mathfrak{P}\,\mathbb{E}[\tilde{V}^{\delta,n}_{k,j+1}|\tilde{X}^{\delta,n}_{t_{k,j}}]\right\},\ j = J-1:0.$$

(ii) Compute $\bar{V}^{n+1}_k = \max\left\{f(t_k, \hat{X}^{n+1}_{t_k}), e^{-r\Delta t}\mathfrak{P}\,\mathbb{E}[\bar{V}_{k+1}|\hat{X}^{n+1}_{t_k}]\right\}$.

(iii) Set $\hat{V}^{n+1}_k = \bar{V}^{n+1}_k + \tilde{V}^{\delta,n}_{k,0} - \hat{V}^n_k$.

end backward k-loop

end n-loop

Remark 2 Note that all fine grid computations are local and can be allocated to a separate processor for each k, for parallelization;

The following partial results can be established for algorithm 3.2bis obtained from 3.2 by changing the first step into $\tilde{V}^{\delta,n}_{k,J} = \mathfrak{P}\,\mathbb{E}(\bar{V}^n_{k+1} \,|\, \tilde{X}^{\delta,n}_{k,J})$ and the last step into: $\hat{V}^{n+1}_k = \bar{V}^{n+1}_k + \tilde{V}^{\delta,n}_{k,0} - \bar{V}^n_k$.

Proposition 3.3 (a) *Let*

$$\tilde{V}^{\Delta,n}_{t_k} = \mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\mathbb{E}[e^{-r(\tau-t_k)}f(\tau, \hat{X}^n_\tau)|\mathcal{F}_{t_k}], \quad \bar{V}^{\Delta,\delta}_{t_k} = \mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\mathbb{E}[e^{-r(\tau-t_k)}f(\tau, \bar{X}^\delta_\tau)|\mathcal{F}_{t_k}]$$

where $\mathcal{T}^{\mathcal{F}}_{t_k}$ denotes the set of $\{t_k, t_{k+1}, \ldots, t_K\}$-valued $\mathcal{F}$-stopping times. Then, for some constant C,

$$\left\|\max_{k=0:K}\left|\tilde{V}^{\Delta,n}_{t_k} - \bar{V}^{\Delta,\delta}_{t_k}\right|\right\|_{L^2(\mathbb{P})} \le [f]_{\text{Lip}}\frac{(C\Delta t)^{\frac{n}{2}}}{\sqrt{(n+1)!}}.$$

(Note that $(\bar{V}^{\Delta,\delta}_{t_k})_{k=0:K}$ is but the coarse Snell envelope of the refined Euler scheme). At a fixed time t_k we have the better estimate

$$\left\|\tilde{V}^{\Delta,n}_{t_k} - \bar{V}^{\Delta,\delta}_{t_k}\right\|_2 \le [f]_{Lip}(C\Delta t)^{n+\frac{1}{2}}\sqrt{\binom{K+1}{n+1} - \binom{k}{n+1}}. \tag{8}$$

(b) *Let $(\bar{V}^\delta_{t_k})_{k=0,K}$ denote the "fine" Snell envelope of the refined Euler scheme at times t_k defined by*

$$\tilde{V}^\delta_{t_k} = \mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\mathbb{E}[e^{-r(\tau-t_k)}f(\tau, \bar{X}^\delta_\tau)|\mathcal{F}_{t_k}]$$

Then, for some constant C',

$$\left\|\tilde{V}^{\delta,n}_{t_k,0} - \bar{V}^\delta_{t_k}\right\|_{L^2(\mathbb{P})} \le C\sqrt{\Delta t}.$$

Remark 3 A result similar to (a) can be obtained for $(\bar{V}^n_{t_k})_{k=0:K}$ i.e. when $\mathcal{F}_{t_k}$ is replaced by $\hat{X}^n_{t_k}$ in the expectation defining $V^n_{t_k}$ at the cost of losing a $\sqrt{\Delta t}$ in the error estimate. Both quantities $\tilde{V}^{\Delta,n}_{t_k}$ and $\bar{V}^n_{t_k}$ do not coincide as $\hat{X}^n$ is not Markovian (it also depends on $\hat{X}^{n-1}$).

4. Numerical Tests

We have chosen an underlying asset which satisfies the Black-Scholes SDE, i.e. $b = rX_t$ and $\sigma = \sigma_0 X_t$ The payoff is $f(t,x) = e^{r(T-t)}(\kappa - x)^+$ with $X_0 = 36,\ \kappa = 40,\ \sigma_0 = 0.2,\ r = 0.05,\ b = r,\ T = 1, M = 100.000$ as in Longstaff-Schwartz [18]. The interpolation used in the LSMC is on the basis $\{1\ x\ x^2\}$, i.e. $P = 2$. The American payoff is then 4.478 at an early exercise $\tau = 0.634$.

4.1. *Convergence of the Parareal Algorithm*

We have chosen a fine grid with $\delta t = T/32$. The free parameters are Δt which governs the number of points on the coarse grid and n the number of parareal algorithm. The error between the American payoff computed on the fine grid by LSMC and the same computed by the parareal algorithm is displayed on Table 1 for both algorithms 3.2 and 3.2bis.

K	J	Δt	$n=1$	$n=2$	$n=3$	$n=4$
2	16	0.666667	0.60338	0.152339	0.0171122	0.000833293
4	8	0.4	0.237451	0.0437726	0.00217885	0.000725382
8	4	0.222222	0.0854814	0.0156243	0.000735309	0.000515332
16	2	0.117647	0.0257407	0.00120513	0.000439038	0.000262921
2	16	0.666667	0.5912463	0.1434691	0.0418341	0.0414722
4	8	0.4	0.2245711	0.0743709	0.0225051	0.0224303
8	4	0.222222	0.0740923	0.0205441	0.0072178	0.0072066
16	2	0.117647	0.0194701	0.0021758	0.0021592	0.0021509

Table 1
Absolute error from the American payoff computed on the fine grid by a sequential LSMC standard algorithm and the same computed using the parareal iterative algorithms 3.2 and 3.2bis. The coarse grid has K intervals; the coarse time step is $\Delta t/K$; the fine grid has a fixed number of points hence each interval $(t_k, t_{t_{k+1}})$ it has J sub-intervals. The top 4 lines of numbers corresponds to Algorithm 3.2 while the last 4 lines correspond to Algorithm 3.2bis for which a partial convergence estimate can be obtained but which does not work as well numerically.

The same information about convergence is now displayed in the two graphs on figure 1 for the errors versus Δt and the errors versus n.

4.2. *Multilevels Parareal Algorithm*

The previous construction being recursive one can again apply the two-levels parareal algorithm to LSMC on each interval $[t_k, t_{t_{k+1}}]$. The problem of finding the optimal strategy for parallelism and computing time is complex, because of there are so many parameters; in what follows the number of levels is $L = 4$; furthermore, when an interval with $J + 1$ points is divided into sub-intervals each is endowed with a partition using $J + 1$ points as well. So if the coarse grid has K intervals, the 4th grid has K^4

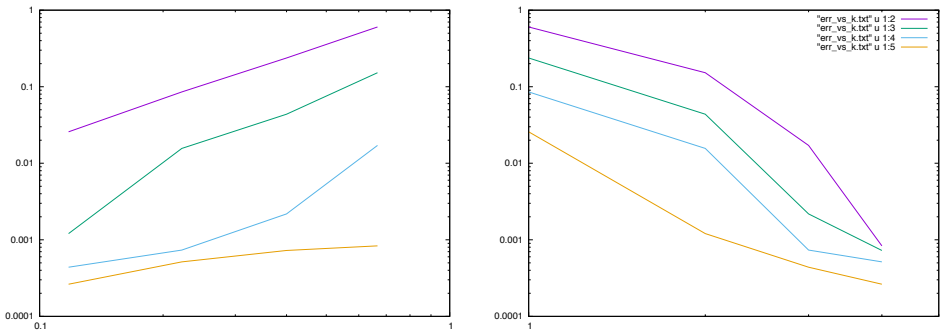

Figure 1. *Errors on the payoff versus Δt on the left for several values of n and versus n on the right for several values of Δt. Both graphs are for Algorithm 3.2 in log-log scales and indicate a general behavior of the error ϵ not incompatible with (3.1). The difficulty is that the method converges very fast making a numerical asymptotic error analysis hard.*

intervals. The results are compared with the reference value of Longstaff-Schwartz, 4.478, and shown on table 2 and figure 2.

The number of parareal iterations is 4 but the error is displayed at each n. We have also carried out

K	K^4	$n=1$	$n=2$	$n=3$	$n=4$
2	16	0.353319	0.118118	0.0477764	0.0276106
3	81	0.208997	0.0390674	0.0225218	0.0186619
4	256	0.141692	0.0259727	0.0192504	0.0136437
5	625	0.105196	0.0220218	0.0179706	0.0129702

Table 2

Absolute error between the computed payoff with the multilevels parareal method and the reference value of Longstaff-Schwartz. The number of levels is L=4, each level is subdivided into K intervals; K^4 is the number of intervals at the deepest level.

some tests with sub-partitions using $J \neq K$. Thus each level has its own number of points per intervals, J_l. These errors are also shown on Figure 2 for $n = 2$. It seems to be $O(K^4)$ for K small and $O(K^2)$ for K

Time-step				Absolute-error				
J_1	J_2	J_3	J_4	Total	$n=1$	$n=2$	$n=3$	$n=4$
6	5	4	3	360	0.108593	0.0305688	0.0202016	0.0142071
3	4	5	6	360	0.35365	0.0316707	0.0167488	0.0135151
20	2	2	2	160	0.0231221	0.0163731	0.0155624	0.013314
2	20	2	2	160	0.354477	0.0835047	0.0231243	0.0121775
2	2	20	2	160	0.351854	0.115285	0.015826	0.0137373
2	2	2	20	160	0.355166	0.119577	0.0444797	0.0110232

Table 3

Absolute error between the computed payoff with the multilevel parareal method and the reference value of Longstaff&Schwartz. There are $L = 4$ levels; at level $l-1$ to obtain level l each interval is divided into J_l intervals. The errors are given versus the number of parareal iterations $n = 1, 2, 3, 4$. Note that all subdivisions give more or less the same precision; computationally and for parallelism the last one is the best.

bigger. The method was implemented in parallel; each interval is allocated to a processor, at each level in

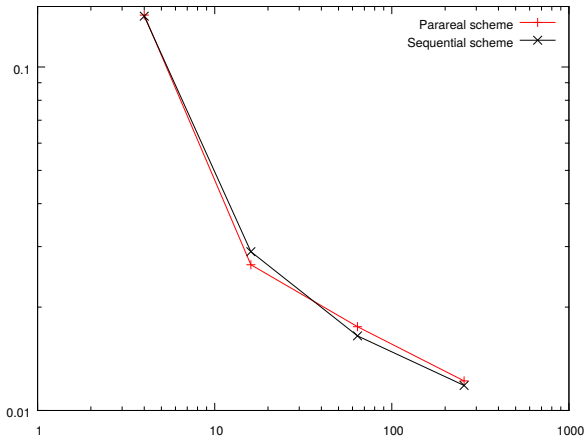

Figure 2. *Comparison between a standard LSMC solution and the parareal solution for the same number of time intervals at the finest level. The 4 points have respectively 1,2,3,4 levels; the first data point has one level and 4 intervals, the second has 2 levels and 16 intervals, the third 3 levels and 64 intervals, the fourth 4 levels and 256 intervals . The total number of time steps is on the horizontal axis, in log scale and the error at* $n = 2$ *is on the vertical axis in log scale as well.*

a round-robin order. Almost perfect parallelism is obtained in our tests on a machine with 32 processors, as shown on Figure 3.

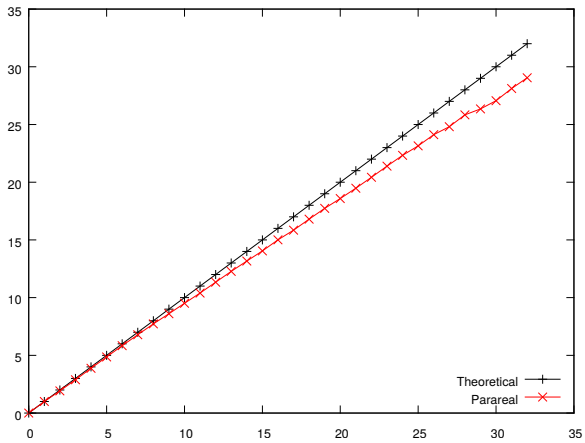

Figure 3. *Speed-up versus the number of processors, i.e. the parareal CPU time on a parallel machine divided by the parareal CPU time on the same machine but running on one processor. There are two levels only; the parameters are* $K = 1, 2.., 32$, $n = 2$ *and* $J = 100$ *so as to keep each processor fully busy.*

References

[1] L. Abbas-Turki and B. Lapeyre. American option pricing on multi-core graphic cards. *Business Intelligence and Financial Engineering*, 2009. International Conference, BIFE 09.

[2] Y. Achdou and O. Pironneau. *Computation methods for option pricing.* Frontiers in Applied Mathematics. SIAM, Philadelphia, 2005. xviii+297 pp., ISBN 0-89871-573-3.

[3] F. Aitsahlia and P. Carr. American options: A comparison of numerical methods. *Numerical Methods in Finance*, pages 67–87, 1997.

[4] G. Bal and Y. Maday. A "parareal" time discretization for non-linear pde's with application to the pricing of an american put. *Recent developments in domain decomposition methods (Zürich, 2001)*, 23:189–202, 2002.

[5] V. Bally and G. Pagès. A quantization algorithm for solving discrete time multidimensional optimal stopping problems. *Bernoulli*, 6:1003–1049, 2003.

[6] M. Benguigui. *Valorisation d'options américaines et Value-At-Risk de portefeuille sur cluster de GPUs\CPUs*. PhD thesis, Université Nice Sophia Antipolis, France, 2015.

[7] M. Benguigui and F. Baude. Towards parallel and distributed computing on GPU for American basket option pricing. 2012. International Workshop on GPU Computing in Cloud in conjunction with 4th IEEE international conference on Cloud Computing Technology and Science, Taipei.

[8] A. P. Caverhill and N. Webber. American options: theory and numerical analysis. In S. Hodges, editor, *Options: recent advances in theory and practise*, pages 80–94. Manchester University Press, Manchester, 1990.

[9] E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression method for american option pricing. *Finance Stochastics*, 6:449–471, 2002.

[10] D. M. Dang, C. C. Christara, and K. R. Jackson. An efficient graphics processing unit-based parallel algorithm for pricing multi-asset American options. *Concurrency and Computation: Practice and Experience*, 24(18):849–866, 2012.

[11] M. Fatica and E. Phillips. Pricing American options with least squares Monte carlo on GPUs. 2013. 6th Workshop on High Performance Computational Finance.

[12] M. Gander and S. Vandewalle. Analysis of the parareal time-parallel, time-integration method. *SIAM J. Sci. Comput.*, 29(2):556–578, 2007.

[13] J.-Y. Girard and I. M. Toke. Monte Carlo valuation of multidimensional American option through grid computing. *Lecture note in computer science*, 3743:462–469, 2006.

[14] Y. Hu, Q. Lu, Z. Cao, and J. Wang. Parallel simuation of high-dimensional American option pricing based on CPU versus MIC. *Concurrency and Computation: Practice and Experence*, 27(15):1110–1121, 2015.

[15] L. Khodja, J.-Y. Girard, R. Couturier, and P. Spitéri. Parallel solution of American option derivatives on GPU clusters. *Computers and Mathematics with Applications*, 65(111):1830–1848, 2013.

[16] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics*. Springer Verlag, Berlin, 1999.

[17] J.-L. Lions, Y. Maday, and G. Turinici. Résolution d'edp par un schéma en temps "pararéel". *C.R. Acad. Sci. Paris sér. I Math.*, 332:661–668, 2000.

[18] F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: a simple least squares approach. *Review of Financial Studies*, 14:113–148, 2001.

[19] G. Pagès. *Introduction to numerical probability for finance*. 2016. Available at `http://www.proba.jussieu.fr/dw/lib/exe/fetch.php?media=users:pages:probnum_gilp_pf16.pdf`, 354p.

[20] G. Pagès and B. Wilbertz. GPGPUs in computational finance: massive parallel computing for American style options. *Concurrency and Computation: Practice and Experience*, 24(18):837–848, 2012.

[21] J. Wan, K. Lai, A. Kolkiewicz, and K. Tan. A parallel quasi-Monte Carlo approach to pricing American options on multiple assets. *International Journal of High Performance Computing and Networking*, 4:321–330, 2006.

5. Proofs for the reviewers. Will not be part of the Compte-Rendu.

5.1. *Proof of Theorem 3.1 and Corollary 3.2*

For the sake of simplicity we detail the proof in the homogeneous case $b(t,x) \equiv b(x)$ and $\sigma(t,x) = \sigma(x)$ and we assume that b and σ are bounded. Let $(X_t)_{t\in[0,T]}$ be the solution of the diffusion process

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 \in \mathbb{R}. \tag{9}$$

We assume that b and σ are uniformly Lipschitz in x. The dependency on t of b and σ has been dropped for clarity. Consider the Euler scheme with a coarse time step $\Delta t = \frac{T}{K}$ and a fine time step $\delta t = \frac{\Delta t}{J} = \frac{T}{KJ}$

$$\begin{aligned}
\bar{X}^{\Delta}_{t_{k+1}} &= \bar{X}^{\Delta}_{t_k} + \Delta t\, b(\bar{X}^{\Delta}_{t_k}) + \sigma(\bar{X}^{\Delta}_{t_k})(W_{t_{k+1}} - W_{t_k}), \quad k = 0,\dots,K, \quad t_k = k\Delta t \\
\bar{X}^{\delta}_{t_{k,j+1}} &= \bar{X}^{\delta}_{t_{k,j}} + \delta t\, b(\bar{X}^{\delta}_{t_{k,j}}) + \sigma(\bar{X}^{\delta}_{t_{k,j}})(W_{t_{k,j+1}} - W_{t_{k,j}}), \quad j = 0,\dots,J, \quad t_{t_{k,j}} = t_k + j\delta t.
\end{aligned} \tag{10}$$

We associate to these their continuous counterpart:

$$\begin{aligned}
d\bar{X}^{\Delta}_t &= b(\bar{X}^{\Delta}_{\underline{t}_\delta})dt + \sigma(\bar{X}^{\Delta}_{\underline{t}_\Delta})dW_t, \qquad \underline{t}_\Delta = \lfloor \frac{t}{\Delta t} \rfloor \Delta t, \\
d\bar{X}^{\delta}_t &= b(\bar{X}^{\delta}_{\underline{t}_\delta})dt + \sigma(\bar{X}^{\delta}_{\underline{t}_\delta})dW_t, \ \underline{t}_\delta = \lfloor \frac{t}{\delta t} \rfloor \delta t.
\end{aligned} \tag{11}$$

The parareal scheme is: $\hat{X}^0_{t_k} = \bar{X}^{\Delta}_{t_k}, \quad k = 1,\dots,K,$

$$\begin{aligned}
\hat{X}^{n+1}_{t_{k+1}} &= G_\Delta\Big(\hat{X}^{n+1}_{t_k}, \frac{1}{\sqrt{\Delta t}}(W_{t_{k+1}} - W_{t_k})\Big) \\
&\quad + G^{(J)}_\delta\Big(\hat{X}^{n}_{t_k}, \{\frac{1}{\sqrt{\delta t}}(W_{t_{k,j+1}} - W_{t_{k,j}})\}_{j=0}^{J-1}\Big) - G_\Delta\Big(\hat{X}_{t_k}, \frac{1}{\sqrt{\Delta t}}(W_{t_{k+1}} - W_{t_k})\Big)
\end{aligned} \tag{12}$$

where

$$G_\Delta(x,z) = \Delta t b(x) + \sigma(x)\sqrt{\Delta t}\, z \quad \text{and} \quad G_\delta(x,z) = \delta t b(x) + \sigma(x)\sqrt{\delta t}\, z$$

We denote by $\mathbf{Z}_{k+1} = \{Z_{k,j}\}_{j=1}^J$ a J-dimensional white noise with an $\mathcal{N}(0, I_J)$ distribution. As a consequence, setting $Z_{k,j} = \frac{1}{\sqrt{\delta t}}(W_{t_{k,j+1}} - W_{t_{k,j}})$ and $Z_{k+1} = \sum_{j=1}^J Z_{k,j} = W_{t_{k+1}} - W_{t_k}$, we have

$$\begin{aligned}
\hat{X}^{n+1}_{t_{k+1}} - \bar{X}^{\delta}_{t_{k+1}} &= G_\Delta(\hat{X}^{n+1}_{t_k}, Z_{k+1}) - G_\Delta(\bar{X}^{\delta}_{t_k}, Z_{k+1}) + G^{(J)}_\delta(\hat{X}^{n}_{t_k}, \mathbf{Z}_{k+1}) - G^{(J)}_\delta(\bar{X}^{\delta}_{t_k}, \mathbf{Z}_{k+1}) \\
&= G_\Delta(\hat{X}^{n+1}_{t_k}, Z_{k+1}) - G_\Delta(\bar{X}^{\delta}_{t_k}, Z_{k+1}) + \phi_{\Delta,\delta}(\hat{X}^{n}_{t_k}, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(\bar{X}^{\delta}_{t_k}, \mathbf{Z}_{k+1})
\end{aligned} \tag{13}$$

with

$$\phi_{\Delta,\delta}(x, \mathbf{Z}_{k+1}) = G^{(J)}_\delta(x, \mathbf{Z}_{k+1}) - G_\Delta(x, Z_{k+1}) = \bar{X}^{\delta,x}_{\Delta t} - \bar{X}^{\Delta,x}_{\Delta t} \tag{14}$$

where $\bar{X}^{\delta,x}_t$ denotes the solution of the δ-Euler scheme at t starting from x at 0, and similarly with $\bar{X}^{\Delta}$.

Our aim is to establish an induction propeerty for $\|\hat{X}^{n+1}_{t_{k+1}} - \bar{X}^{\delta}_{t_{k+1}}\|_2^2$. To thsi end we first deal with teh last two terms of (13). Note that

$$\begin{aligned}
\phi_{\Delta,\delta}(x, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(y, \mathbf{Z}_{k+1}) &= \bar{X}^{\delta,x}_{\Delta t} - \bar{X}^{\Delta,x}_{\Delta t} - (\bar{X}^{\delta,y}_{\Delta t} - \bar{X}^{\Delta,y}_{\Delta t}) \\
&= -\big(b(x)\Delta t + \sigma(x)\Delta W\big) + \int_0^{\Delta t} b(\bar{X}^{\delta,x}_{\underline{s}_\delta})ds + \int_0^{\Delta t} \sigma(\bar{X}^{\delta,x}_{\underline{s}_\delta})dW_s \\
&\quad + \big(b(y)\Delta t + \sigma(y)\Delta W\big) - \int_0^{\Delta t} b(\bar{X}^{\delta,y}_{\underline{s}_\delta})ds - \int_0^{\Delta t} \sigma(\bar{X}^{\delta,y}_{\underline{s}_\delta})dW_s
\end{aligned}$$

$$= -\int_0^{\Delta t} \Big(b(x) - b(y) - \big(b(\bar{X}^{\delta,x}_{\underline{s}_\delta} - b(\bar{X}^{\delta,y}_{\underline{s}_\delta}) \big) \Big) ds$$
$$- \int_0^{\Delta t} \Big(\sigma(x) - \sigma(y) - \big(\sigma(\bar{X}^{\delta,x}_{\underline{s}_\delta} - \sigma(\bar{X}^{\delta,y}_{\underline{s}_\delta}) \big) \Big) dW_s \tag{15}$$

The first integral will be called A. The last integral B can be bounded as follows

$$\mathbb{E}[B^2] = \int_0^{\Delta t} \mathbb{E}\Big[\Big(\sigma(x) - \sigma(y) - \big(\sigma(\bar{X}^{\delta,x}_{\underline{s}_\delta} - \sigma(\bar{X}^{\delta,y}_{\underline{s}_\delta}) \big) \Big)^2 \Big] ds \tag{16}$$

Applying Ito's formula to $\bar{X}^{\delta,x}$ yields for any $s \in [0, \Delta t]$

$$\sigma(\bar{X}^{\delta,x}_s) = \sigma(x) + \int_0^s \sigma'(\bar{X}^{\delta,x}_u)\sigma(\bar{X}^{\delta,x}_{\underline{u}_\delta}) dW_u$$
$$+ \int_0^s \Big(\sigma'(\bar{X}^{\delta,x}_u) b(X^{\delta,x}_{\underline{u}_\delta}) + \frac{1}{2}\sigma"(X^{\delta,x}_u)\sigma^2(X^{\delta,x}_{\underline{u}_\delta}) \Big) du \tag{17}$$

The same holds with y instead of x and so

$$\sigma(\bar{X}^{\delta,x}_s) - \sigma(\bar{X}^{\delta,y}_s) - (\sigma(x) - \sigma(y)) = \int_0^s S^{x,y}_u du + \int_0^s H^{x,y}_u dW_u \tag{18}$$

with

$$S^{x,y}_u = \frac{1}{2}\Big[\sigma''(X^{\delta,x}_u)\sigma^2(X^{\delta,x}_{\underline{u}_\delta}) - \sigma''(X^{\delta,y}_u)\sigma^2(X^{\delta,y}_{\underline{u}_\delta}) \Big] + \sigma'(\bar{X}^{\delta,x}_u) b(X^{\delta,x}_{\underline{u}_\delta}) - \sigma'(\bar{X}^{\delta,y}_u) b(X^{\delta,y}_{\underline{u}_\delta})$$
$$H^{x,y}_u = \sigma'(X^{\delta,x}_u)\sigma(X^{\delta,x}_{\underline{u}_\delta}) - \sigma'(X^{\delta,y}_u)\sigma(X^{\delta,y}_{\underline{u}_\delta}). \tag{19}$$

Hence, denoting $\|f\|_2 := \|f\|_{L^2(\mathbb{P})}$ and using general Minkowski and Doob inequalities, we get

$$\|\sigma(\bar{X}^{\delta,x}_s) - \sigma(\bar{X}^{\delta,y}_s) - (\sigma(x) - \sigma(y))\|_2 \le \int_0^s \|S^{x,y}_u\|_2 du + \Big\| \int_0^s H^{x,y}_u dW_u \Big\|_2$$
$$\le \int_0^s \|S^{x,y}_u\|_2 du + \Big[\mathbb{E}[\int_0^s (H^{x,y}_u)^2 du] \Big]^{\frac{1}{2}} \le \int_0^s \|S^{x,y}_u\|_2 du + \Big[\int_0^s (\|H^{x,y}_u)\|_2^2 du] \Big]^{\frac{1}{2}}. \tag{20}$$

Now σ' bounded and σ Lipschitz leads to

$$\|H^{x,y}_u\|_2 \le \|\sigma'\|_\infty [\sigma]_{Lip} \|\bar{X}^{\delta,x}_{\underline{u}_\delta} - \bar{X}^{\delta,y}_{\underline{u}_\delta}\|_2 + \|\sigma\|_\infty [\sigma']_{Lip}) \|\bar{X}^{\delta,x}_u - \bar{X}^{\delta,y}_u\|_2. \tag{21}$$

A classical result (see *e.g.* [19]) on the Euler scheme says that for all $v \in [0, T]$, uniformly in δ,

$$\|\bar{X}^{\delta,x}_v - \bar{X}^{\delta,y}_v\|_2 \le \sup_{t \in [0,T]} \|\bar{X}^{\delta,x}_t - \bar{X}^{\delta,y}_t\|_2 \le C_{b,\sigma}|x - y|. \tag{22}$$

Consequently

$$\sup_{u \in [0,T]} \|H^{x,y}_u\|_2 \le C_{b,\sigma,\sigma'}|x - y|. \tag{23}$$

As for $S^{x,y}_u$, assuming σ'' Lipschitz, a similar computation leads to

$$\sup_{u \in [0,T]} \|S^{x,y}_u\|_2 \le C_{b,\sigma,\sigma',\sigma''}|x - y|. \tag{24}$$

Note by the way that all these terms vanish if σ is constant.

Plugging these bounds in (20) leads to

$$\begin{aligned}&\|\sigma(\bar{X}_s^{\delta,x}) - \sigma(\bar{X}_s^{\delta,y}) - (\sigma(x) - \sigma(y))\|_2 \\ &\qquad \leq \max\{C_{b,\sigma,\sigma'}, C_{b,\sigma,\sigma',\sigma''}\}(\sqrt{s} + s)|x - y] \leq \tilde{C}\sqrt{s}|x - y|\end{aligned} \tag{25}$$

which implies in turn,

$$\mathbb{E}[B]^2 \leq \int_0^{\Delta t} \tilde{C}^2 s|x - y|^2 ds = \frac{1}{2}\tilde{C}^2(\Delta t)^2|x - y|. \tag{26}$$

The term A in (15) can be treated likewise:

$$b(\bar{X}_s^{\delta,x}) - b(x) = \int_0^s b'(\bar{X}_u^{\delta,x})\sigma(\bar{X}_{\underline{u}_\delta}^{\delta,x})dW_u + \int_0^s \Big[b'(\bar{X}_u^{\delta,x})b(\bar{X}_{\underline{u}_\delta}^{\delta,x}) + \frac{1}{2}b''((\bar{X}_u^{\delta,x}))\sigma^2((\bar{X}_{\underline{u}_\delta}^{\delta,x})\Big]du. \tag{27}$$

In the end, provided b'' is Lipschitz continuous,

$$\begin{aligned}\mathbb{E}[A^2] &\leq \Delta t \int_0^{\Delta t} (\mathbb{E}[b(x) - b(y) - [b(\bar{X}_{\underline{u}_\delta}^{\delta,x}) - b(\bar{X}_{\underline{u}_\delta}^{\delta,y})]])^2 du \\ &\leq \Delta t \frac{1}{2}\tilde{C}_{b,b',b'',\sigma}(\Delta t)^2|x - y|^2 = \frac{1}{2}\tilde{C}_{b,b',b'',\sigma}(\Delta t)^3|x - y|^2.\end{aligned} \tag{28}$$

So we have proved that

$$\|\phi_{\Delta,\delta}(x, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(y, \mathbf{Z}_{k+1})\|_2^2 \leq C(\Delta t)^\alpha|x - y|^2 \text{ with } \alpha = 2 \text{ or } 3 \text{ if } \sigma \text{ constant} \tag{29}$$

Let us bound the two other terms in (13)

$$\begin{aligned}\|G_\Delta(x, Z) - G_\Delta(y, Z)\|_2^2 &\leq \Big(1 + \Delta t([b]_{Lip} + \frac{1}{2}[\sigma]_{Lip}^2) + \frac{\Delta t^2}{2}[b]_{Lip}^2\Big)|x - y|^2 \\ &\leq \big(1 + C'_{b,\sigma}\Delta t\big)^2|x - y|^2.\end{aligned} \tag{30}$$

As one must raise (13) to the square, a cross term appears,

$$\begin{aligned}C &:= \mathbb{E}\Big[\Big(G_\Delta(\hat{X}_{t_k}^{n+1}, Z_{k+1}) - G_\Delta(\bar{X}_{t_k}^\delta, Z_{k+1})\Big)\Big(\phi_{\Delta,\delta}(\hat{X}_{t_k}^n, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(\bar{X}_{t_k}^\delta, \mathbf{Z}_{k+1})\Big)\Big] \\ &= \mathbb{E}\Big[\Big((x - y) + \Delta t(b(x) - b(y)) + (\sigma(x) - \sigma(y))\Delta W\Big)\Big(A' + B'\Big)\Big]\end{aligned} \tag{31}$$

where A' and B' are as in the last two integrals in (15) except that x is changed to x'.

Note that $\mathbb{E}[B'] = 0$ because it is a stochastic integral, hence

$$C = [(x - y) + \Delta t(b(x) - b(y))]\mathbb{E}[A'] + (\sigma(x) - \sigma(y))\mathbb{E}[(A' + B')\Delta W].$$

Now, using Schwartz' inequality,

$$\begin{aligned}|\mathbb{E}[C]| &\leq (1 + \Delta t[b]_{Lip})\|A'\|_2|x - y| + [\sigma]_{Lip}|x - y|\|\Delta W\|_2\|A' + B'\|_2 \\ &\leq (1 + \Delta t[b]_{Lip})\|A'\|_2|x - y| + [\sigma]_{Lip}|x - y|\sqrt{\Delta t}\big(\|A'\|_2 + \|B'\|_2\big).\end{aligned} \tag{32}$$

We recall our previous bounds on A and B,

$$\|B'\|_2 \leq \tilde{C}\Delta t|x' - y|, \quad \|A'\|_2 \leq \tilde{C}(\Delta t)^{\frac{3}{2}}|x' - y|? \tag{33}$$

Consequently, with $2\bar{C} = \tilde{C}(1 + \Delta t[b]_{Lip} + [\sigma]_{Lip})$,

$$|\mathbb{E}[C]| \leq 2\bar{C}(\Delta t)^{\frac{3}{2}}|x'-y||x-y|. \tag{34}$$

We are now in a position to patch the pieces together; C denotes a generic constant:

$$\begin{aligned}
&\mathbb{E}\Big[\Big(G_\Delta(\hat{X}^{n+1}_{t_k}, Z_{k+1}) - G_\Delta(\bar{X}^\delta_{t_k}, Z_{k+1}) + \phi_{\Delta,\delta}(\hat{X}^n_{t_k}, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(\bar{X}^\delta_{t_k}, \mathbf{Z}_{k+1})\Big)^2\Big] \\
&= \mathbb{E}\Big[\Big(G_\Delta(\hat{X}^{n+1}_{t_k}, Z_{k+1}) - G_\Delta(\bar{X}^\delta_{t_k}, Z_{k+1})\Big)^2\Big] + \mathbb{E}\Big[\Big(\phi_{\Delta,\delta}(\hat{X}^n_{t_k}, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(\bar{X}^\delta_{t_k}, \mathbf{Z}_{k+1})\Big)^2\Big] \\
&\quad +2\mathbb{E}\Big[\Big(G_\Delta(\hat{X}^{n+1}_{t_k}, Z_{k+1}) - G_\Delta(\bar{X}^\delta_{t_k}, Z_{k+1})\Big)\Big(\phi_{\Delta,\delta}(\hat{X}^n_{t_k}, \mathbf{Z}_{k+1}) - \phi_{\Delta,\delta}(\bar{X}^\delta_{t_k}, \mathbf{Z}_{k+1})\Big)\Big] \\
&\leq (1+C\Delta t)^2|x-y|^2 + C\Delta t^2|x'-y|^2 + 2C(\Delta t)^{\frac{3}{2}}|x-y||x'-y| \\
&\leq (1+C\Delta t)^2|x-y|^2 + C\Delta t^2|x'-y|^2,
\end{aligned} \tag{35}$$

since $2(\Delta t)^{\frac{3}{2}}|x-y||x'-y| \leq \Delta t|x-y|^2 + \Delta t^2|x'-y|^2$.

As $x = \hat{X}^{n+1}_{t_k}, x' = \hat{X}^n_{t_k}, y = \bar{X}^\delta_{t_k}$ are independant of $\mathbf{Z}_{k+1}$, integrating (35) with respect to the distribution of the triplet $\mathbb{P}_{(\hat{X}^{n+1}_{t_k}, \hat{X}^n_{t_k}, \bar{X}^\delta_{t_k})}(dx, dx', dy)$ yields by Fubini's theorem

$$\|\hat{X}^{n+1}_{t_{k+1}} - \bar{X}^\delta_{t_{k+1}}\|_2^2 \leq (1+C'\Delta t)\|\hat{X}^{n+1}_{t_k} - \bar{X}^\delta_{t_k}\|_2^2 + C'\Delta t^2\|\hat{X}^n_{t_k} - \bar{X}^\delta_{t_k}\|_2^2. \tag{36}$$

Let $\epsilon^n_k := \|\hat{X}^n_{t_k} - \bar{X}^\delta_{t_k}\|_2^2$. Then,

$$\epsilon^{n+1}_{k+1} \leq (1+C'\Delta t)\epsilon^{n+1}_k + C'\Delta t^2\epsilon^n_k. \tag{37}$$

We will now show that

Lemma 5.1 *The recurrence (37) together with $\epsilon^n_0 = 0$ implies, for some constant C,*

$$\epsilon^n_k \leq CC'^n \binom{k}{n} (\Delta t)^{2n+1}, \quad k = 0, \ldots, K,\ n = 0, \ldots, N \tag{38}$$

and also $\epsilon^n_k = 0,\ \forall n \geq k$.

Proof. Introduce $\tilde{\epsilon}^n_k = (1+C'\Delta t)^{n-k}C'^{-n}(\Delta t)^{-2n}\epsilon^n_k$ It is easy to see that it satisfies

$$\tilde{\epsilon}^{n+1}_{k+1} \leq \tilde{\epsilon}^{n+1}_k + \tilde{\epsilon}^n_k. \tag{39}$$

Notice now that by the convergence estimates for the Euler scheme,

$$\begin{aligned}
\tilde{\epsilon}^0_k \leq \epsilon^0_k &= \|\hat{X}^\Delta_{t_k} - \bar{X}^\delta_{t_k}\|^2 \leq \|\hat{X}^\Delta_{t_k} - X_{t_k}\|^2 + \|X_{t_k} - \bar{X}^\delta_{t_k}\|^2 \\
&\leq (\sqrt{\Delta t} + \sqrt{\delta t})^2(1+|X_0|)^2 C_{b,\sigma,T} \leq C_1\Delta t.
\end{aligned} \tag{40}$$

It implies in turn that $\tilde{\epsilon}^n_k = 0,\ \forall n \geq k$. Indeed by (39)

$$\tilde{\epsilon}^{n+1}_{k+1} \leq \tilde{\epsilon}^{n+1}_k \leq \ldots \leq \tilde{\epsilon}^{n+1}_0 = 0.$$

Finally if the lemma holds for k, n it holds for $k+1, n+1$ because by (37) and (40)

$$\tilde{\epsilon}^{n+1}_{k+1} \leq C_1\Delta t\Big[\binom{k}{n+1} + \binom{k}{n}\Big] = C_1\Delta t\binom{k+1}{n+1}. \tag{41}$$

5.2. *Proof of Proposition 3.3*

For every $k \in \{0, \dots, K\}$,

$$\begin{aligned}
|\tilde{V}^{\Delta,n}_{t_k} - \bar{V}^{\Delta,\delta}_{t_k}| &= \left|\mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\mathbb{E}[e^{-r(\tau-t_k)}f(\tau,\hat{X}^n_\tau)|\mathcal{F}_{t_k}] - \mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\mathbb{E}[e^{-r(\tau-t_k)}f(\tau,\bar{X}^\delta_\tau)|\mathcal{F}_{t_k}]\right| \\
&\le \mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\left|\mathbb{E}\Big[e^{-r(\tau-t_k)}f(\tau,\hat{X}^n_\tau) - e^{-r(\tau-t_k)}f(\tau,\bar{X}^\delta_\tau)|\mathcal{F}_{t_k}\Big]\right| \\
&\le \mathbb{E}\left[\mathbb{P}\text{-ess sup}_{\tau\in\mathcal{T}^{\mathcal{F}}_{t_k}}\left|f(\tau,\hat{X}^n_\tau) - f(\tau,\bar{X}^\delta_\tau)\right| \,\middle|\, \mathcal{F}_{t_k}\right] \\
&\le [f]_{\text{Lip}}\mathbb{E}\left[\max_{l=k,\dots,K}\left|\hat{X}^n_{t_l} - \bar{X}^\delta_{t_l}\right| \,\middle|\, \mathcal{F}_{t_k}\right]. \qquad (42)
\end{aligned}$$

Consequently, owing to Doob's inequality

$$\begin{aligned}
\left\|\max_{k=0:K}|\tilde{V}^{\Delta,n}_{t_k} - \bar{V}^{\Delta,\delta}_{t_k}|\right\|_2^2 &\le 4[f]^2_{Lip}\left\|\max_{l=0:K}|\hat{X}^n_{t_l} - \bar{X}^\delta_{t_l}|\right\|_2^2 \le 4[f]^2_{Lip}\sum_{l=0}^{K}\|\hat{X}^n_{t_l} - \bar{X}^\delta_{t_l}\|_2^2 \\
&\le 4[f]^2_{Lip}(C\Delta t)^{2n+1}\sum_{l=n}^{K}\binom{l}{n} \\
&= 4[f]^2_{Lip}(C\Delta t)^{2n+1}\binom{K+1}{n+1} \le \left[2[f]_{Lip}\frac{(C'\Delta t)^{\frac{n}{2}}}{\sqrt{(n+1)!}}\right]^2 \qquad (43)
\end{aligned}$$

where we used in the second line that $X^n_{t_l} = \bar{X}^\delta_{t_l}$ for $l \le n-1$.

If we are only interested in $\|\tilde{V}^{\Delta,n}_{t_k} - \bar{V}^{\Delta,\delta}_{t_k}\|_2^2$, we obtain starting again from (42)

$$\begin{aligned}
\|\tilde{V}^{\Delta,n}_{t_k} - \bar{V}^{\Delta,\delta}_{t_k}\|_2^2 &\le [f]^2_{Lip}\sum_{l=k\vee n}^{K}\|\hat{X}^n_{t_l} - \bar{X}^\delta_{t_l}\|_2^2 \le [f]^2_{Lip}(C\Delta t)^{2n+1}\sum_{l=k\vee n}^{K}\binom{l}{n} \\
&= [f]^2_{Lip}(C\Delta t)^{2n+1}\left[\binom{K+1}{n+1} - \binom{k}{n+1}\right] \qquad (44)
\end{aligned}$$

with the usual convention on the binomial coefficient when $k \le n$.

5.3. *Correcting Markovian deficiency*

In fact the LSMC method approximates by regression $\mathbb{E}\big(f(\hat{X}^n_{t_{k+1}})|\hat{X}^n_{t_k}\big)$, not $\mathbb{E}\big(f(\hat{X}^n_{t_{k+1}})|\mathcal{F}_{t_k}\big)$, since $(\hat{X}_{t_k})_{k=0:K}$ is not a Markov chain as emphasized by its very definition. The quantity of interest is in fat $\tilde{V}^{reg}_{t_k}$ satisfying the Backward Dynamical Programming formula:

$$\tilde{V}^{reg}_{t_k} = f(T,\hat{X}^n_T),\ \tilde{V}^{reg}_{t_k} = \max\Big(f(t_k,\hat{X}^n_{t_k}), \mathbb{E}\big(\tilde{V}^{reg}_{t_k}|\hat{X}^n_{t_k}\big)\Big),\ k = 0\dots,K-1.$$

What is the error induced by considering $\tilde{V}^{reg}$ rather than $\tilde{V}$?

Proposition 5.2 *Let $(\tilde{V}^{reg}_{t_k})_{k=0:K}$ be defined as above. Then there exists a real constant $C > 0$ only depending on b, σ, $[f]_{\text{Lip}}$ and T such that*

$$\max_{k=0:K}\|\tilde{V}^{reg}_{t_k} - \tilde{V}_{t_k}\|_2 \le C^{n-1}\frac{(\Delta t)^{\frac{n-1}{2}}}{\sqrt{(n+1)!}}.$$

Proof. Starting from both backward definitions we get

$$|\tilde{V}^{reg}_{t_k} - \tilde{V}_{t_k}| \leq \Big|\mathbb{E}\big(\tilde{V}^{reg}_{t_{k+1}}\big|\hat{X}^n_{t_k}\big) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\Big|.$$

Now, using that $\hat{X}^n_{t_k}$ is $\mathcal{F}_{t_k}$-measurable, one has

$$\mathbb{E}\big(\tilde{V}^{reg}_{t_{k+1}}\big|\hat{X}^n_{t_k}\big) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big) = \mathbb{E}\big(\tilde{V}^{reg}_{t_{k+1}} - \tilde{V}_{t_{k+1}}\big|\hat{X}^n_{t_k}\big) \overset{\perp}{+} \mathbb{E}\big(\mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big|\hat{X}^n_{t_k}\big)\big) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)$$

where $\overset{\perp}{+}$ denotes orthogonality in $L^2(\mathbb{P})$. Consequently, also using that conditional expectations an L^2-contraction

$$\big\|\tilde{V}^{reg}_{t_k} - \tilde{V}_{t_k}\big\|_2^2 \leq \big\|\tilde{V}^{reg}_{t_{k+1}} - \tilde{V}_{t_{k+1}}\big\|_2^2 + \big\|\mathbb{E}\big(\mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big|\hat{X}^n_{t_k}\big)\big) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big\|_2^2.$$

Conditional expectation being an orthogonal projector, for every Borel function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\varphi(\hat{X}^n_{t_k}) \in L^2(\mathbb{P})$

$$\big\|\mathbb{E}\big(\mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big|\hat{X}^n_{t_k}\big)\big) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big\|_2^2 \leq \big\|\varphi(\hat{X}^n_{t_k}) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big\|_2^2.$$

Now, let us consider the Snell envelope of the payoff $\big(f(t_{k,j}, \bar{X}^\delta_{k,j})\big)_{k,j}$ associated to the global refined Euler scheme with step δt, denoted $(\bar{V}^\delta_{t_{k,j}})_{k,j}$.

$$\big\|\varphi(\hat{X}^n_{t_k}) - \mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big\|_2^2 \leq \Big(\big\|\tilde{V}_{t_{k+1}} - \bar{V}^\delta_{t_{k+1}}\ \big\|_2 + \big\|\varphi(\hat{X}^n_{t_k}) - \mathbb{E}\big(\bar{V}^\delta_{t_{k+1}}\big|\mathcal{F}_{t_k}\big)\big\|_2\Big)^2.$$

As the Euler scheme is a Markov chain, we now that $\bar{V}^\delta_{t_{k,j}} = v_{k,j}(\bar{X}^\delta_{t_{k,j}})$ (with $v_{K,J} = f(T,.)$ and $v_{k,J} = v_{k+1,0}$) and, see [5], the propagation of Lipschitz continuity holds so that $[v_{k,j}]_{\mathrm{Lip}} \leq C = C_{b,\sigma,f,T}$, $k = 0, \ldots, K$, $j = 0, \ldots, J$. Consequently

$$\mathbb{E}\big(\tilde{V}_{t_{k+1}}\big|\mathcal{F}_{t_k}\big) = \mathbb{E}\big(v_{t_{k+1,0}}(\bar{X}^\delta_{t_{k+1}})\big|\mathcal{F}_{t_k}\big) = \mathbb{E}\big(v_{t_{k+1,0}} \circ G^{(J)}_\delta(\bar{X}^\delta_{t_k}, \mathbf{Z}_{k+1})\big|\bar{X}^\delta_{t_k}\big) = w_k(\bar{X}^\delta_{t_k})$$

so that $[w_k]_{\mathrm{Lip}} \leq (1+C'\delta t)^J [v_{k+1,0}]_{\mathrm{Lip}} \leq (1+C'\delta t)C$. Setting $\varphi = w_k$ finally yields, owing to Theorem 3.1 and Equation (8) in the proof of Corollary 3.2, we get the recursion

$$\begin{aligned}
\big\|\tilde{V}^{reg}_{t_k} - \tilde{V}_{t_k}\big\|_2^2 &\leq \big\|\tilde{V}^{reg}_{t_{k+1}} - \tilde{V}_{t_{k+1}}\big\|_2^2 + 2\left[[f]^2_{\mathrm{Lip}}(C\Delta t)^{2n+1}\left[\binom{K+1}{n+1} - \binom{k}{n+1}\right] + [w_k]^2_{\mathrm{Lip}} C^{2n}\binom{k}{n}(\Delta t)^{2n+1}\right] \\
&\leq \big\|\tilde{V}^{reg}_{t_{k+1}} - \tilde{V}_{t_{k+1}}\big\|_2^2 + C'\,(C\Delta t)^{2n+1}\left[\binom{K+1}{n+1} - \binom{k}{n+1} + \binom{k}{n}\right] \\
&\leq \big\|\tilde{V}^{reg}_{t_{k+1}} - \tilde{V}_{t_{k+1}}\big\|_2^2 + C'\,(C\Delta t)^{2n+1}\binom{K+1}{n+1}.
\end{aligned}$$

Having in mind that $\tilde{V}^{reg}_T = \tilde{V}_T = f(T, \hat{X}^n_T)$, we derive

$$\big\|\tilde{V}^{reg}_{t_k} - \tilde{V}_{t_k}\big\|_2^2 \leq C'\,(C\Delta t)^{2n+1}(K-k)\binom{K+1}{n+1} \leq (C'')^{n-1}\frac{(\Delta t)^{n-1}}{(n+1)!}. \tag{45}$$

6. With a modified algorithm

If we change the algorithm, we may have a theoretcial convergence result. The amendment is as follows :

6.1. *Modified algorithm*

One modifies the first and the last line backward loop **for k=K-1,. . .,0** (backward loop):

(i) On each (t_k, t_{k+1}), from $\tilde{V}^{\delta,n}_{k,J} = \mathfrak{P}\,\mathbb{E}(\bar{V}^n_{k+1} \mid \tilde{X}^{\delta,n}_{k,J})$, compute by a backward loop in j

$$\tilde{V}^{\delta,n}_{k,j} = \max\left\{f(t_{k,j}, \tilde{X}^{\delta,n}_{t_{k,j}}), e^{-r\delta t}\mathfrak{P}\,\mathbb{E}[\tilde{V}^{\delta,n}_{k,j+1}|\tilde{X}^{\delta,n}_{t_{k,j}}]\right\},\ j = J-1:0.$$

(ii) Compute $\bar{V}^{n+1}_k = \max\left\{f(t_k, \hat{X}^{n+1}_{t_k}), e^{-r\Delta t}\mathfrak{P}\,\mathbb{E}[\bar{V}_{k+1}|\hat{X}^{n+1}_{t_k}]\right\}$.

(iii) Set $\hat{V}^{n+1}_k = \bar{V}^{n+1}_k + \tilde{V}^{\delta,n}_{k,0} - \bar{V}^n_k$.

We assuem oin what follwos taht $\mathfrak{P} = Id$.

6.2. *Statements and proofs*

First, we introduce our "target object": let $(\bar{V}^\delta_{t_{k,j}})_{k,j}$ be the $(\mathbb{P}, \mathcal{F}_{t_{k,j}})_{k=0:K,j=0:J}$-Snell envelope of $\big(f(t_{k,j}, \bar{X}^\delta_{t_{k,j}})\big)_{k,j}$ where $(\bar{X}^\delta_{t_{k,j}})_{k=0:K,j=0:J}$ denotes the (global) Euler scheme with step δt (starting from x_0 at time $t = 0$).

Then, we recall that $(\tilde{X}^{\delta,n}_{t_{k,j}})_{j=0,\dots,J}$ denotes the Euler scheme with step δ starting from $\widehat{X}^n_{t_k}$ at time $t_k = t_{k,0}$. Let $\big(\tilde{V}^{\delta,n}_{t_{k,J}}\big)_{j=0,J}$ denote the $(\mathbb{P}, \mathcal{F}_{t_{k,j}})_{j=0,\dots,K}$-Snell envelope of

$$\tilde{f}_{t_{k,J}} = \tilde{V}^{\delta,n}_{t_{k,J}} \quad \text{and} \quad \tilde{f}_{t_{k,\ell}} = f(t_{k,\ell}, \tilde{X}^{\delta,n}_{t_{k,\ell}}),\ \ell = 0,\dots,J-1.$$

where

$$\tilde{V}^{\delta,n}_{t_{k,J}} = \mathbb{E}\big(\bar{V}^n_{k+1} \mid \tilde{X}^{\delta,n}_{t_{k,J}}\big).$$

Having in mind that, owing to the "pre-conditioning", the Markov property implies by induction that $\tilde{V}^{\delta,n}_{k,j+1}$ is a function of $\tilde{X}^{\delta,n}_{t_{k,j+1}}$ so that

$$\mathbb{E}\big(\tilde{V}^{\delta,n}_{k,j+1}|\mathcal{F}_{t_{k,j}}\big) = \mathbb{E}\big(\tilde{V}^{\delta,n}_{k,j+1}|\tilde{X}^{\delta,n}_{t_{k,j}}\big),\ j = J-1,\dots,0.$$

This second case justifies the use of LSMC in the sense that there is no Markov default in that phase like for the coarse component of the parareal scheme.

Proposition 6.1 *When* $\tilde{V}^{\delta,n}_{t_{k,J}} = \mathbb{E}\big(\bar{V}^n_{k+1} \mid \tilde{X}^{n,\delta}_{t_{k,J}}\big)$.

$$\Big\|\max_{j=0,\dots,J}\big|\bar{V}^\delta_{t_{k,j}} - \tilde{V}^{\delta,n}_{t_{k,j}}\big|\Big\|_2 \le 2[f]_{\mathrm{Lip}}C\sqrt{\Delta t}$$

Proof. (a) As the $\mathbb{P}$-Snell envelope with horizon Δt of $(\tilde{f}_{t_{k,j}})_j$, starting at t_k with timestep δt, the sequence $(\tilde{V}_{t_{k,j}})_j$ reads

$$\tilde{V}^{\delta,n}_{t_{k,j}} = \mathbb{P}\text{-ess sup}\Big\{\mathbb{E}\big(\tilde{f}_\theta|\mathcal{F}_{t_{k,j}}\big),\ \theta \in \mathcal{T}^{\mathcal{F},\delta}_{t_{k,j},t_{k,J}}\Big\}$$

where $\mathcal{T}^{\mathcal{F},\delta}_{t_{k,j},t_{k,J}}$ denotes the set of $\{t_{k,\ell},\ \ell = j,\dots,J\}$-valued $(\mathcal{F}_{t_{k,\ell}})_\ell$-stopping times and

$$\tilde{f}_{t_{k,J}} = \mathbb{E}\big(\bar{V}^n_{k+1} \mid \tilde{X}^{n,\delta}_{t_{k,J}}\big) \quad \text{and} \quad \tilde{f}_{t_{k,\ell}} = f(t_{k,\ell}, \tilde{X}^{n,\delta}_{t_{k,\ell}}),\ \ell = 0,\dots,J-1.$$

The Markov property shared by the Euler scheme $(\bar{X}^\delta_{t_{k,j}})_{k,j}$ implies that $(\bar{V}^\delta_{t_{k,j}})_{k,j}$ classically satisfies a Backward Dynamic Programing formula, which in turn implies the following local representation

$$\bar{V}^\delta_{t_{k,j}} = \mathbb{P}\text{-ess sup}\Big\{\mathbb{E}\big(\bar{f}_\theta|\mathcal{F}_{t_{k,j}}\big),\ \theta \in \mathcal{T}^{\mathcal{F},\delta}_{t_{k,j},t_{k,J}}\Big\}$$

where $\bar{f}_{t_{k,J}} = \bar{V}^\delta_{t_{k,J}}$ and $\bar{f}_{t_{k,\ell}} = f(t_{k,\ell}, \bar{X}^\delta_{t_{k,\ell}})$, $\ell = 0,\dots,J-1$.

Consequently, as seen in the proof of the former proposition,

$$\left|\bar{V}^{\delta}_{t_{k,j}} - \tilde{V}^{\delta,n}_{t_{k,j}}\right| \le \mathbb{E}\left[\max_{\ell=j,\dots,J}\left|\tilde{f}_{t_{k,\ell}} - \bar{f}_{t_{k,\ell}}\right| \Big| \mathcal{F}_{t_{k,j}}\right]$$

so that, owing to conditional Jensen's Inequality

$$\left|\bar{V}^{\delta}_{t_{k,j}} - \tilde{V}^{\delta,n}_{t_{k,j}}\right|^2 \le \mathbb{E}\left[\max_{\ell=0,\dots,J}\left|\tilde{f}_{t_{k,\ell}} - \bar{f}_{t_{k,\ell}}\right|^2 \Big| \mathcal{F}_{t_{k,j}}\right].$$

It follows form the conditional Doob's Inequality that

$$\mathbb{E}\left[\max_{j=0,\dots,J}\left|\bar{V}^{\delta}_{t_{k,j}} - \tilde{V}^{\delta,n}_{t_{k,j}}\right|^2 \Big| \mathcal{F}_{t_{k,0}}\right] \le 4\mathbb{E}\left[\max_{\ell=0,\dots,J}\left|\tilde{f}_{t_{k,\ell}} - \bar{f}_{t_{k,\ell}}\right|^2 \Big| \mathcal{F}_{t_{k,0}}\right].$$

Now, f being Lipschitz continuous in x uniformly in $t\in[0,T]$, we get

$$\begin{aligned}\max_{\ell=0,\dots,J}\left|\tilde{f}_{t_{k,\ell}} - \bar{f}_{t_{k,\ell}}\right|^2 &\le |\mathbb{E}(\bar{V}^n_{k+1} \mid \tilde{X}^n_{t_{k,J}}) - \bar{V}^{\delta}_{t_{k,J}}|^2 \vee \left[[f]^2_{\mathrm{Lip}} \max_{j=0,\dots,J-1}|\tilde{X}^n_{t_{k,j}} - \bar{X}^{\delta}_{t_{k,j}}|^2\right] \\ &\le \left|\mathbb{E}(\bar{V}^n_{k+1} \mid \tilde{X}^n_{t_{k,J}}) - \bar{V}^{\delta}_{t_{k,J}}\right|^2 + [f]^2_{\mathrm{Lip}} \max_{j=0,\dots,J-1}\left|\tilde{X}^n_{t_{k,j}} - \bar{X}^{\delta}_{t_{k,j}}\right|^2. \end{aligned} \tag{46}$$

Conditional expectation given $\tilde{X}^n_{t_{k,J}}$ being an orthogonal projector, one has by the Pythagoras Theorem

$$\left\|\mathbb{E}(\bar{V}^n_{k+1} \mid \tilde{X}^n_{t_{k,J}}) - \bar{V}^{\delta}_{t_{k,J}}\right\|^2 = \left\|\mathbb{E}(\bar{V}^n_{k+1} - \bar{V}^{\delta}_{t_{k,J}} \mid \tilde{X}^n_{t_{k,J}})\right\|^2 + \left\|\mathbb{E}(\bar{V}^{\delta}_{t_{k,J}} \mid \tilde{X}^n_{t_{k,J}}) - \bar{V}^{\delta}_{t_{k,J}}\right\|^2.$$

The functions $f(t_{k,j},.)$ being $[f]_{\mathrm{Lip}}$-Lipschitz, we derive from the Markov property that $\bar{V}^{\delta}_{t_{k,J}} = v_{k,J}(\bar{X}^{\delta}_{t_{k,J}})$ where (see [5]) $v_{k,J}$ is Lipschitz continous with $[v_{k,J}]_{\mathrm{Lip}} \le C = C_{b,\sigma,T}$. Conditional expectation given $\tilde{X}^n_{t_{k,J}}$ is the best quadratic approximation by a function of $\tilde{X}^n_{t_{k,J}}$, consequently

$$\left\|\bar{V}^{\delta}_{t_{k,J}} - \mathbb{E}(\bar{V}^{\delta}_{t_{k,J}} \mid \tilde{X}^n_{t_{k,J}})\right\|^2 \le [v_{k,J}]_{\mathrm{Lip}}\left\|\bar{X}^{\delta}_{t_{k,J}} - \tilde{X}^n_{t_{k,J}}\right\|_2^2 \le C\left\|\bar{X}^{\delta}_{t_{k,J}} - \tilde{X}^n_{t_{k,J}}\right\|_2^2$$

whereas, as an L^2-contraction,

$$\left\|\mathbb{E}(\bar{V}^n_{k+1} - \bar{V}^{\delta}_{t_{k,J}} \mid \tilde{X}^n_{t_{k,J}})\right\|^2 \le \left\|\bar{V}^n_{k+1} - \bar{V}^{\delta}_{t_{k,J}}\right\|^2 = \left\|\bar{V}^n_{k+1} - \bar{V}^{\delta}_{t_{k+1}}\right\|^2.$$

On the other hand, the (pseudo-)flow property for the Euler scheme yields

$$\mathbb{E}\left[\max_{j=0,\dots,J}\left|\tilde{X}^n_{t_{k,j}} - \bar{X}^{\delta}_{t_{k,j}}\right|^2 \Big| \mathcal{F}_{t_{k,0}}\right] \le C\left|\hat{X}^n_{t_k} - \bar{X}^{\delta}_{t_k}\right|^2.$$

still for a real constant $C = C_{b,\sigma,T}$. Taking the expectation in (46) and in the above equation, then using the bounds established in (6) of Theorem 3.1 and in (8), we derive that

$$\begin{aligned}\mathbb{E}\left[\max_{j=0,\dots,J}\left|\bar{V}^{\delta}_{t_{k,j}} - \tilde{V}^{\delta,n}_{t_{k,j}}\right|^2\right] &\le 4\left[[f]^2_{\mathrm{Lip}} C_{b,\sigma,T}\mathbb{E}\left[\left|\hat{X}^n_{t_k} - \bar{X}^{\delta}_{t_k}\right|^2\right] + \mathbb{E}\left[\left|\bar{V}^{\delta}_{t_{k+1}} - \bar{V}^n_{k+1}\right|^2\right]\right] \\ &\le 4\left[[f]^2_{\mathrm{Lip}} C^{2n+1}\binom{k}{n}(\Delta t)^{2n+1} + \mathbb{E}\left[\left|\bar{V}^{\delta}_{t_{k+1}} - \bar{V}^n_{k+1}\right|^2\right]\right]\end{aligned}$$

by Theorem 3.1. Now, if we denote by $(\bar{V}^{\delta}_{t_k})_{k=0:K}$ the $(\mathbb{P},\mathcal{F}_{t_k})$-Snell envelope of the Euler scheme of $f(t_k,\bar{X}^{\delta}_{t_k})_{k=0:K}$ (exercise are possible only attires t_k), then

$$\left\|\bar{V}^{\delta}_{t_{k+1}} - \bar{V}^n_{k+1}\right\|_2 \le \left\|\bar{V}^{\delta}_{t_{k+1}} - \bar{V}^{\Delta,\delta}_{k+1}\right\|_2 + \left\|\bar{V}^{\Delta,\delta}_{k+1} - \bar{V}^n_{k+1}\right\|_2.$$

It is elementary to show on the very definition of Snell envelope, the uniform Lipschitz continuity of the functions $f(t_k, .)$ and the fact that the Euler scheme is an Itô process that

$$\left\|\bar{V}^{\delta}_{t_{k+1}} - \bar{V}^{\Delta,\delta}_{k+1}\right\|_2 \leq C_{b,\sigma,T}\sqrt{\Delta t}.$$

On the other hand, by Proposition 3.3, as $(\bar{V}_k)_k = \tilde{V}^{\delta,n}_{t_k})_k$ coincide, we derive that

$$\left\|\bar{V}^{\Delta,\delta}_{k+1} - \bar{V}^{n}_{k+1}\right\|_2 \ \leq [f]_{\mathrm{Lip}} \frac{(C'\Delta t)^{\frac{n}{2}}}{\sqrt{(n+1)!}}.$$

So, we have convergence but we are note able to highlight the parareal speeding up of the procedure.