Available online at www.sciencedirect.com
ScienceDirect
Mathematics and Computers in Simulation 143 (2018) 125–137
MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Original Articles

Stratified regression-based variance reduction approach for weak approximation schemes

D. Belomestny[a,b,*], S. Häfner[c], M. Urusov[a]

[a] *Duisburg–Essen University, Essen, Germany*
[b] *National Research University Higher School of Economics, Moscow, Russia*
[c] *Pricewaterhouse Coopers GmbH, Frankfurt, Germany*

Received 22 January 2016; received in revised form 19 February 2017; accepted 13 May 2017
Available online 25 May 2017

Abstract

In this paper we suggest a modification of the regression-based variance reduction approach recently proposed in Belomestny et al. [1]. This modification is based on the stratification technique and allows for a further significant variance reduction. The performance of the proposed approach is illustrated by several numerical examples.

Keywords: Control variates; Stratification; Monte Carlo methods; Weak schemes; Regression

1. Introduction

Let $T > 0$ be a fixed time horizon. Consider a d-dimensional diffusion process $(X_t)_{t\in[0,T]}$ defined by the Itô stochastic differential equation

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{1}$$

for Lipschitz continuous functions $\mu : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$, where $(W_t)_{t\in[0,T]}$ is a standard m-dimensional Brownian motion. Suppose we want to compute the expectation

$$u(t,x) := \mathbb{E}[f(X_T^{t,x})], \tag{2}$$

where $X^{t,x}$ denotes the solution to (1) started at time t in point x. The standard Monte Carlo (SMC) approach for computing $u(0,x)$ at a fixed point $x \in \mathbb{R}^d$ basically consists of three steps. First, an approximation $\overline{X}_T$ for $X_T^{0,x}$ is constructed via a time discretisation in Eq. (1) (we refer to [4] for a nice overview of various discretisation schemes).

* Corresponding author at: Duisburg–Essen University, Essen, Germany.
E-mail addresses: denis.belomestny@uni-due.de (D. Belomestny), stefan.haefner@de.pwc.com (S. Häfner), mikhail.urusov@uni-due.de (M. Urusov).

http://dx.doi.org/10.1016/j.matcom.2017.05.003

Next, N_0 independent copies of the approximation $\overline{X}_T$ are generated, and, finally, a Monte Carlo estimate V_{N_0} is defined as the average of the values of f at simulated points:

$$V_{N_0} := \frac{1}{N_0}\sum_{i=1}^{N_0} f\left(\overline{X}_T^{(i)}\right). \tag{3}$$

In the computation of $u(0,x) = \mathbb{E}[f(X_T^{0,x})]$ by the SMC approach there are two types of error inherent: the discretisation error $\mathbb{E}[f(X_T^{0,x})] - \mathbb{E}[f(\overline{X}_T)]$ and the Monte Carlo (statistical) error, which result from the substitution of $\mathbb{E}[f(\overline{X}_T)]$ with the sample average V_{N_0}. The aim of variance reduction methods is to reduce the statistical error. For example, in the so-called control variate variance reduction approach one looks for a random variable ξ with $\mathbb{E}\xi = 0$, which can be simulated, such that the variance of the difference $f(\overline{X}_T) - \xi$ is minimised, that is,

$$\mathrm{Var}[f(\overline{X}_T) - \xi] \to \min \text{ under } \mathbb{E}\xi = 0.$$

Then one uses the sample average

$$V_{N_0}^{CV} := \frac{1}{N_0}\sum_{i=1}^{N_0}\left[f\left(\overline{X}_T^{(i)}\right) - \xi^{(i)}\right] \tag{4}$$

instead of (3) to approximate $\mathbb{E}[f(\overline{X}_T)]$. The use of control variates for computing expectations of functionals of diffusion processes via Monte Carlo was initiated by Newton [7] and further developed in Milstein and Tretyakov [6]. Heath and Platen [3] use the integral representation to construct unbiased variance-reduced estimators. In Belomestny et al. [1] a novel regression-based approach for the construction of control variates, which reduces the variance of the approximated functional $f(\overline{X}_T)$ was proposed. As shown in [1], the "Monte Carlo approach with the Regression-based Control Variate" (abbreviated below as "RCV approach") as well as its enhancement, called "recursive RCV (RRCV) approach", is able to achieve a higher order convergence of the resulting variance to zero, which in turn leads to a significant complexity reduction as compared to the SMC algorithm. The RCV approaches become especially simple in the case of the so-called weak approximation schemes, i.e., the schemes, where simple random variables are used in place of Brownian increments, and which became quite popular in recent years. In this paper we further enhance the performance of the RRCV algorithm by combining it with stratification. The idea of the resulting *stratified RCV (SRCV)* algorithm is based on partitioning of the state space into a collection of sets $\mathscr{A}_1, \dots, \mathscr{A}_p$ and then performing conditional regressions separately on each set. It turns out that by choosing $\mathscr{A}_1, \dots, \mathscr{A}_p$ to be the level sets of the discrete-valued random variables used in the weak approximation scheme, we can achieve a further variance reduction effect as compared to the original approach in [1]. The paper is organised as follows. In Section 2, the SRCV algorithm is introduced and compared with the RCV and RRCV ones. The complexity analysis of the SRCV algorithm is conducted in Section 3. Section 4 is devoted to the simulation study. Necessary proofs are collected in Section 5.

2. SRCV approach and its differences with RCV and RRCV ones

In what follows $J \in \mathbb{N}$ denotes the time discretisation parameter. We set $\Delta := T/J$ and consider discretisation schemes denoted by $(X_{\Delta,j\Delta})_{j=0,\dots,J}$, which are defined on the grid $\{j\Delta : j = 0,\dots,J\}$. In Sections 2.1 and 2.2 we consider weak schemes of order 1. In this setting we recall the RCV and RRCV algorithms, introduce the SRCV algorithm and explain how it compares to the RCV and RRCV ones. In Section 2.3 we briefly discuss the case of weak schemes of order 2.

2.1. RCV algorithm for first order schemes

Let us consider a weak scheme of order 1, where d-dimensional approximations $X_{\Delta,j\Delta}$, $j = 0,\dots,J$, satisfy $X_{\Delta,0} = x_0$ and

$$X_{\Delta,j\Delta} = \Phi_\Delta(X_{\Delta,(j-1)\Delta}, \xi_j), \quad j = 1,\dots,J, \tag{5}$$

for some functions $\Phi_\Delta : \mathbb{R}^{d+m} \to \mathbb{R}^d$, with $\xi_j = (\xi_j^1,\dots,\xi_j^m)$, $j = 1,\dots,J$, being m-dimensional i.i.d. random vectors with i.i.d. coordinates satisfying

$$\mathbb{P}\left(\xi_j^k = \pm 1\right) = \frac{1}{2}, \quad k = 1,\dots,m.$$