

# Advanced Monte Carlo and Optimization Methods for Optimal Stopping Problems: Part II

**Denis Belomestny**

Premolab, Duisburg-Essen University

18.09.2012

# Optimal Stopping Problems

- $(X_j)_{j \geq 0}$  is a **Markov chain**
  - on a **filtered probability space**  $(\Omega, \mathcal{F}, (\mathcal{F}_j)_{j \geq 0}, P_x)$
  - with values in  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$
  - starting at  $x$  under  $P_x$  for some  $x \in \mathbb{R}^d$
- $G_j : \mathbb{R}^d \rightarrow \mathbb{R}, j = 0, \dots, \mathcal{J}$ , is a set of measurable functions that fulfill

$$E_x \left[ \sup_{0 \leq j \leq \mathcal{J}} |G_j(X_j)| \right] < \infty$$

# Optimal Stopping Problems

Consider the following discrete time **optimal stopping problem**:

$$Y_0^* = \sup_{\tau \in \{0, \dots, \mathcal{J}\}} E_x [G_\tau(X_\tau)],$$

where

- $\tau$  is a  $(\mathcal{F}_j)$ -stopping time with values in  $\{0, \dots, \mathcal{J}\}$ , i.e.  $\{\tau = j\} \in \mathcal{F}_j$

## Question

*How to approximate  $Y_0^*$  in the case when the expectation  $E[G_j(X_\tau)]$  cannot be computed in a closed form ?*

# Simulation-Based Optimization Algorithms

- Simulate a set of trajectories:  $X_0^{(m)}, \dots, X_{\mathcal{T}}^{(m)}$ ,  $m = 1, \dots, M$ .
- Fix a parametric set of stopping rules:  $\tau = \tau(\theta)$ ,  $\theta \in \Theta$ .
- Compute

$$\theta_M = \arg \sup_{\theta \in \Theta} \left\{ \frac{1}{M} \sum_{m=1}^M G_{\tau^{(m)}(\theta)}(X_{\tau^{(m)}(\theta)}) \right\}.$$

- Simulate a new set of trajectories:  $X_0^{(M+n)}, \dots, X_{\mathcal{T}}^{(M+n)}$ ,  $n = 1, \dots, N$ .
- Define an estimate for  $Y_0^*$  via

$$Y_{M,N} = \frac{1}{N} \sum_{n=1}^N G_{\tau^{(n)}(\theta_M)}(X_{\tau^{(n)}(\theta_M)}).$$

# Simulation-Based Optimization Algorithms

SBO algorithms are popular among practitioners (in finance and insurance) but there are some open questions:

- ▶ How fast does  $Y_{M,N}$  converge to  $Y_0^*$  as  $M, N \rightarrow \infty$  ?
- ▶ What is the optimal relation between  $M$ ,  $N$  and  $\Theta$  that minimizes the computational costs ?
- ▶ How to choose a parametric family of stopping times  $\tau(\theta)$ ,  $\theta \in \Theta$  ?

# Main Setup

## ► Snell-Envelope Process

$$Y_j^*(X_j) = \sup_{\tau \in \{j, \dots, \mathcal{T}\}} \mathbb{E} [G_\tau(X_\tau) | X_j]$$

## ► Continuation values

$$C_j^*(x) := \mathbb{E}[Y_{j+1}^*(X_{j+1}) | X_j = x], \quad j = 0, \dots, \mathcal{T} - 1$$

# Main Setup

- ▶ Introduce the **stopping region**  $\mathcal{S}^* = \mathcal{S}_1^* \times \dots \times \mathcal{S}_{\mathcal{J}}^*$  with  $\mathcal{S}_{\mathcal{J}}^* = \mathbb{R}^d$  by definition, and

$$\mathcal{S}_j^* = \left\{ x \in \mathbb{R}^d : Y_j^*(x) \leq G_j(x) \right\}, \quad j = 1, \dots, \mathcal{J} - 1.$$

- ▶ Introduce the **first entry times**  $\tau_j^*$  into  $\mathcal{S}^*$  by setting

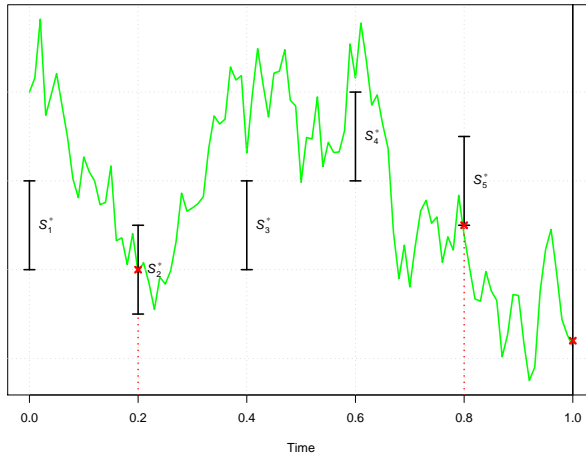
$$\tau_j^* = \tau_j(\mathcal{S}^*) := \min \{ j \leq l \leq \mathcal{J} : X_l \in \mathcal{S}_l \}.$$

## Theorem

*The stopping times  $\tau_j^*$  are optimal, i.e.,*

$$Y_j^*(x) = \mathbb{E} \left[ G_{\tau_j^*}(X_{\tau_j^*}) | X_j = x \right], \quad j = 1, \dots, \mathcal{J}.$$

# Main Setup





# Main Setup

- ▶  $(X_j^{(m)})_{j=0,\dots,\mathcal{J}}, m = 1, \dots, M$ , are  $M$  independent Markov chains with the same distribution as  $X$ .
- ▶  $\mathfrak{B}$  is a collection of sets from

$$\mathcal{B}^{\mathcal{J}} := \underbrace{\mathcal{B} \otimes \dots \otimes \mathcal{B}}_{\mathcal{J}}$$

containing all sets  $\mathcal{S} \in \mathcal{B}^{\mathcal{J}}$  of the form  $\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_{\mathcal{J}-1} \times \mathbb{R}^d$  with  $\mathcal{S}_j \in \mathcal{B}, j = 1, \dots, \mathcal{J} - 1$ .

Define

$$\mathcal{S}_M := \arg \sup_{\mathcal{S} \in \mathfrak{G}} \left\{ \frac{1}{M} \sum_{m=1}^M G_{\tau_1^{(m)}}(\mathcal{S}) \left( X_{\tau_1^{(m)}}^{(m)}(\mathcal{S}) \right) \right\},$$

where  $\mathfrak{G}$  is a subset of  $\mathfrak{B}$

# Main Setup

## ► Compute

$$Y_{M,N} = \frac{1}{N} \sum_{n=1}^N G_{\tau_M^{(n)}}(X_{\tau_M^{(n)}}^{(M+n)})$$

with

$$\tau_M^{(n)} = \min \left\{ 1 \leq j \leq \mathcal{J} : X_j^{(M+n)} \in \mathcal{S}_{M,j} \right\}, \quad n = 1, \dots, N.$$

## Observation

$Y_{M,N}$  is low biased, i.e., it fulfills  $Y_M := \mathbb{E}_x [Y_{M,N} | X^{(1)}, \dots, X^{(M)}] \leq Y^*$ .

# $\delta$ -entropy

► Let

$$d_X(G_1 \times \dots \times G_{\mathcal{J}}, G'_1 \times \dots \times G'_{\mathcal{J}}) = \sum_{j=1}^{\mathcal{J}} P_X(X_j \in G_j \Delta G'_j),$$

where  $\{G_j\}$  and  $\{G'_j\}$  are subsets of  $\mathbb{R}^d$ .

►  $N(\delta, \mathfrak{G}, d_X)$  is the smallest value  $n$  for which there exist pairs of sets

$$(G_{k,1}^L \times \dots \times G_{k,\mathcal{J}}^L, G_{k,1}^U \times \dots \times G_{k,\mathcal{J}}^U), \quad k = 1, \dots, n,$$

such that  $d_X(G_{k,1}^L \times \dots \times G_{k,\mathcal{J}}^L, G_{k,1}^U \times \dots \times G_{k,\mathcal{J}}^U) \leq \delta$  for all  $k = 1, \dots, n$ , and for any  $G \in \mathfrak{G}$  there exists  $j(G) \in \{1, \dots, n\}$  for which

$$G_{k(G),j}^L \subseteq G_j \subseteq G_{k(G),j}^U, \quad j = 1, \dots, \mathcal{J}.$$

# Entropy Assumption

Assume that the family of stopping regions  $\mathfrak{S}$  is such that

$$\mathcal{H}(\delta, \mathfrak{S}, d_X) := \log \{N(\delta, \mathfrak{S}, d_X)\} \leq A\delta^{-\varrho}$$

for some constant  $A > 0$ , any  $0 < \delta < 1$  and some  $\varrho > 0$ .

# Entropy Assumption

## Example (sets with Hölder continuous boundaries)

Consider the class

$$\mathfrak{S}_\gamma = \{S_{b_1} \times \dots \times S_{b_{J-1}} \times E : b_1, \dots, b_{J-1} \in \Sigma(\gamma, H)\}$$

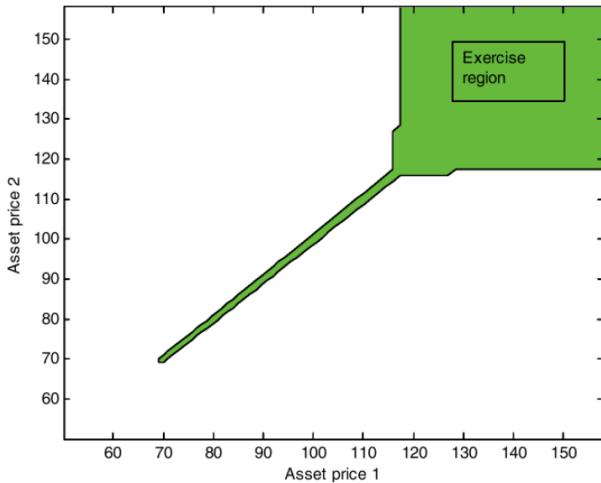
where

$$S_b := \{(x_1, \dots, x_d) \in \mathbb{R}^d : 0 \leq x_d \leq b(x_1, \dots, x_{d-1})\}.$$

## Proposition

The class  $\mathfrak{S}_\gamma$  satisfies the entropy assumption with  $\varrho = (d-1)/\gamma$ .

# Entropy Assumption



# Upper Bound

## Theorem

Assume that

$$Y_0^* - \sup_{\mathcal{S} \in \mathcal{G}} \mathbb{E}[G_{\tau_1}(\mathcal{S})(X_{\tau_1}(\mathcal{S}))] \leq DM^{-1/(1+\varrho)}$$

and the *margin condition* (MC):

$$\mathbb{P}_x(|G_j(X_j) - C_j^*(X_j)| \leq \delta) \leq A_{0,k} \delta^\alpha, \quad \delta < \delta_0$$

holds for  $j = 1, \dots, \mathcal{J}$ . Then for any  $U > U_0$  and  $M > M_0$

$$\mathbb{P}_x^{\otimes M} \left( Y_0^* - Y_M \geq (U/M)^{\frac{1+\alpha}{2+\alpha(1+\varrho)}} \right) \leq C \exp(-\sqrt{U}/B).$$

# Margin Condition

- $X_1, \dots, X_{\mathcal{J}}$  is a time homogenous Markov chain with the state space  $\mathbb{R}_+$  and a transition density  $p(y|x) = x^{-1} \bar{p}(y/x)$ .
- The function  $\bar{p}(z)$  stays positive on  $(0, \infty)$  and satisfies  $\bar{p}(z) \lesssim z^{-3/2}, z \rightarrow +\infty$ .
- Assume that  $G_j(x) = a_j(\kappa - x)^+$  where  $a_j, j = 1, \dots, \mathcal{J}$ , is a decreasing sequence of positive numbers.

## Observation

*The MC condition is fulfilled with  $\alpha \geq 1/2$ .*



# Implications

- It follows from our results that

$$Y_0^* - Y_M = O_P \left( M^{-\frac{1+\alpha}{2+\alpha(1+\varrho)}} \right) = o_P(M^{-1/2})$$

as long as  $\alpha > 0$ .

Hence

$$Y_0^* - Y_{M,N} = O_P \left( M^{-\frac{1+\alpha}{2+\alpha(1+\varrho)}} + N^{-\frac{1}{2}} \right).$$

# Implications

- Suppose that the computational cost of the optimization step is proportional to  $M$ .
- Given the overall computational cost  $\mathcal{C}$ , a reasonable choice of  $M$  can be found as a solution of the optimization problem:

$$M^{-\frac{1+\alpha}{2+\alpha(1+\varrho)}} + N^{-\frac{1}{2}} \rightarrow \min, \quad \text{s.t.} \quad M + N = \mathcal{C}.$$

- Solution:

$$M \asymp \mathcal{C}^{\frac{3}{2(1+\beta)}}, \quad \beta = \frac{1+\alpha}{2+\alpha(1+\varrho)}.$$

# Implications

- ▶ There exists a parametric family of stopping regions, i.e.,  $\varrho$  can be taken arbitrary small
- ▶ The functions  $G_j(x) - C_j^*(x)$  are Lipschitz continuous for all  $j = 0, \dots, \mathcal{J}$ , i.e.,  $\alpha = 1$

$$M \asymp C^{9/10}, \quad N \asymp C.$$

# Lower Bound

## Theorem

Let  $\tau \in \{1, 2\}$  and  $\mathcal{P}_{\alpha, \gamma}$  be a class of measures such that

- ▶ **MC** is fulfilled with some  $\alpha > 0$
- ▶ for any  $P \in \mathcal{P}_{\alpha, \gamma}$ , we have  $\mathcal{S}^* = \mathcal{S}^*(P) \in \mathfrak{G}_\gamma$  (**Hölder continuous boundaries**)

Then for **any stopping time**  $\tau_M \in \{1, 2\}$  measurable w.r.t.  $\mathcal{F}^{\otimes M}$ , it holds

$$\inf_{P \in \mathcal{P}_{\alpha, \gamma}} P^{\otimes M} \left( Y_0^* - Y_M \geq CM^{-\frac{1+\alpha}{2+\alpha(1+(d-1)/\gamma)}} \right) > 0$$

with  $Y_M = E_P[G_{\tau_M}(X_{\tau_M})]$ .

# Numerical example: Bermudan max-call option

- ▶ The risk-neutral dynamic of the asset  $X(t) = (X^1(t), \dots, X^d(t))$  :

$$\frac{dX^l(t)}{X^l(t)} = (r - \delta)dt + \sigma dW^l(t), \quad X^l(0) = x_0, \quad l = 1, \dots, d.$$

- ▶ At any time  $t \in \{t_1, \dots, t_{\mathcal{J}}\}$  the holder of the option may exercise it and receive the payoff

$$G_j(X_j) := \left( \max \left( X_j^1, \dots, X_j^d \right) - \kappa \right)^+,$$

where  $X_j := X(t_j)$  for  $j = 1, \dots, \mathcal{J}$ .

# Numerical example: Bermudan max-call option

Consider a parametric family of stopping regions:

$$\mathcal{S}_j(\theta_j) := \{(x_1, x_2) : (\max(x_1, x_2) - K)^+ > \theta_j^1; |x_1 - x_2| > \theta_j^2\},$$

where  $\theta_j \in \Theta$ ,  $j = 1, \dots, \mathcal{J}$ , and  $\Theta$  is a compact subset of  $\mathbb{R}^2$ .

## Remark

*We simplify the corresponding optimization problem by setting  $\theta_1 = \dots = \theta_{\mathcal{J}}$ .*

# Numerical example: Bermudan max-call option

- ▶ Simulate  $L$  independent sets of trajectories of the process  $(X_j)$ :

$$(X_1^{(l,m)}, \dots, X_{\mathcal{J}}^{(l,m)}), \quad m = 1, \dots, M,$$

where  $l = 1, \dots, L$ .

- ▶ Compute estimates  $\theta_M^{(1)}, \dots, \theta_M^{(L)}$  via

$$\theta_M^{(l)} := \arg \max_{\theta \in \Theta} \left\{ \frac{1}{M} \sum_{m=1}^M G_{\tau_1(\mathcal{S}(\theta))} \left( X_{\tau_1(\mathcal{S}(\theta))}^{(l,m)} \right) \right\}.$$

- ▶ Simulate a new set of trajectories:

$$(\tilde{X}_1^{(n)}, \dots, \tilde{X}_{\mathcal{J}}^{(n)}), \quad n = 1, \dots, N.$$

# Numerical example: Bermudan max-call option

- Compute  $L$  estimates for the optimal value function  $Y_1^*$

$$Y_{M,N}^{(l)} := \frac{1}{N} \sum_{n=1}^N G_{\tau_{M,l,n}} \left( \tilde{X}_{\tau_{M,l,n}}^{(n)} \right), \quad l = 1, \dots, L,$$

with

$$\tau_{M,l,n} := \min \left\{ 1 \leq j \leq \mathcal{J} : \tilde{X}_j^{(n)} \in \mathcal{S}_j \left( \theta_M^{(l)} \right) \right\}, \quad n = 1, \dots, N.$$

Denote by  $\sigma_{M,N,l}$  the standard deviation computed from the sample  $(G_{\tau_{M,l,n}}(\cdot), n = 1, \dots, N)$  and set  $\sigma_{M,N} = \min_l \sigma_{M,N,l}$ .

- Compute

$$\mu_{M,N,L} := \frac{1}{L} \sum_{l=1}^L Y_{M,N}^{(l)}, \quad \vartheta_{M,N,L} := \sqrt{\frac{1}{L-1} \sum_{l=1}^L \left( Y_{M,N}^{(l)} - \mu_{M,N,L} \right)^2}.$$



# Numerical example: Bermudan max-call option

We have

$$\bar{Y} - Y_{M,N} = \underbrace{(\bar{Y} - \mathbb{E}_{\mathbb{P}^{\otimes M}}[Y_M])}_I + \underbrace{(\mathbb{E}_{\mathbb{P}^{\otimes M}}[Y_M] - Y_M)}_{II} + \underbrace{Y_M - Y_{M,N}}_{III},$$

where

$$Y_{M,N} := \frac{1}{N} \sum_{n=1}^N G_{\tau_{M,n}} \left( \tilde{X}_{\tau_{M,n}}^{(n)} \right)$$

and

$$\bar{Y} := \max_{\theta \in \Theta} \mathbb{E}[G_{\tau_1}(\mathbf{s}(\theta))(X_{\tau_1}(\mathbf{s}(\theta)))].$$

# Numerical example: Bermudan max-call option

We use the following approximations:

► First term:

$$I \approx Q_1(M) := \mu_{M^*, N^*, L^*} - \mu_{M, N^*, L^*}, \quad N^*, L^* \gg 1$$

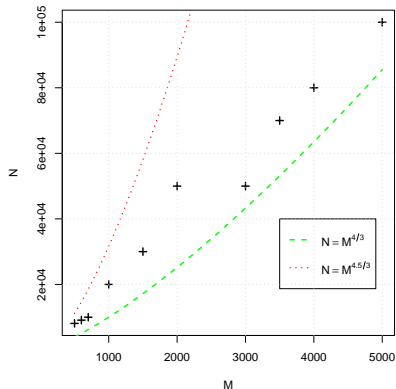
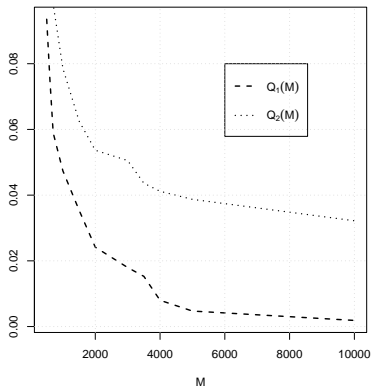
► Second term:

$$\text{sd}(II) = \sqrt{\text{Var}_{\mathbf{p} \otimes M} [Y_M]} \approx Q_2(M) := \sqrt{\vartheta_{M, N^*, L^*}}, \quad N^*, L^* \gg 1$$

► Third term:

$$\text{sd}(III) = \sqrt{\text{Var}_{\mathbf{p} \otimes N} [Y_{M, N}]} \approx Q_3(N) := \sigma_{M^*, N} / \sqrt{N}, \quad M^* \gg 1$$

# Numerical example: Bermudan max-call option



# Numerical example: Bermudan max-call option

Since  $Q_2(M)$  **dominates**  $Q_1(M)$ , by solving the equation

$$Q_2(M) = Q_3(N),$$

one can infer on the optimal relation between  $M$  and  $N$ .

## Conclusion

*The choice  $M = N^{3/4}$  is sufficient in this situation since it always leads to the inequality  $Q_1(M) + \sigma Q_2(M) \leq \sigma Q_3(N)$  for any  $\sigma > 1$ .*

# Tools for the proof

## Theorem

*It holds*

$$\begin{aligned} & Y_j^*(X_j) - Y_j(X_j) \\ &= \mathbb{E} \left[ \sum_{l=j}^{\mathcal{J}-1} |G_l(X_l) - \mathbb{E}[Y_{l+1}^*(X_{l+1})|\mathcal{F}_l]| \mathbf{1}_{\{X_l \in (S_l^* \triangle S_l) \setminus (\cap_{l'=l}^{\mathcal{J}-1} S_{l'})\}} \middle| \mathcal{F}_j \right] \end{aligned}$$

*for  $j = 1, \dots, \mathcal{J} - 1$ , where*

$$Y_j(X_j) := \mathbb{E} \left[ G_{T_j}(\mathbf{s})(X_{T_j}(\mathbf{s})) | \mathcal{F}_j \right], \quad j = 1, \dots, \mathcal{J}.$$

# Tools for the proof

- ▶ For any  $\mathcal{S} \in \mathfrak{S}$  define **empirical process**

$$\nu_M(\mathcal{S}) = \sqrt{M} \int g_{\mathcal{S}} d(P_X^{\otimes M} - P_X)$$

- ▶ Assume that

$$\mathcal{H}_B(\delta, \mathfrak{S}, d_X) \leq A\delta^{-\kappa}$$

- ▶ Then

$$P \left( \sup_{\mathcal{S} \in \mathfrak{S}, \|g_{\mathcal{S}} - g_{\mathcal{S}_0}\|_{L_2(P_X)} > \varepsilon} \frac{|\nu_M(\mathcal{S}) - \nu_M(\mathcal{S}_0)|}{\|g_{\mathcal{S}} - g_{\mathcal{S}_0}\|_{L_2(P_X)}^{1-\kappa/2}} > U \right) \leq C \exp(-U/C^2)$$



Belomestny, D. (2010).

On the rates of convergence of simulation-based optimization algorithms for optimal stopping problems. *Annals of Applied Probability*, **21**(1), 215–239.