Monte Carlo Pricing of Bermudan Options:

Correction of super-optimal and sub-optimal exercise

Christian Fries
12.07.2006
(Version 1.2)

www.christian-fries.de/finmath/talks/2006foresightbias

Agenda

- Monte-Carlo Pricing: Review and Notation
  - Monte-Carlo Simulation, Monte-Carlo Pricing
  - Some Key Problems in Modeling and Monte-Carlo Pricing
- Monte-Carlo Pricing of Bermudan Options
  - Definition
  - Valuation: Optimal Stopping Formulation
  - Valuation: The Backward Algorithm
  - Estimation of the Exercise Criteria
- Monte-Carlo Conditional Expectation Estimators
  - Full Re-Simulation
  - Perfect Foresight
  - Binning
  - Regression
- The Foresight Bias
  - Definition
  - Numerical Removal of the Foresight Bias
  - Is Foresight Bias Negligible?
  - Parallelization of Monte-Carlo Pricing
  - Analytic Calculation of the Foresight Bias
  - Sub-Optimality Bias
  - Analytic Correction of Foresight and Sub-Optimality in a Backward Algorithm
- Numerical Results
- References

Monte Carlo Pricing

Review and Notation

Monte-Carlo Simulation:

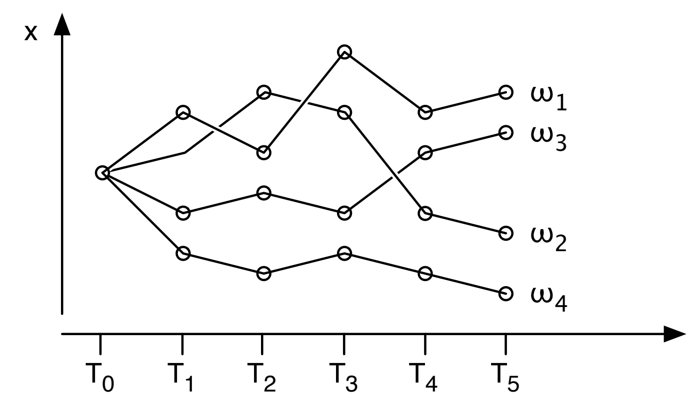
Given a stochastic process for some model primitives $X = (X_1, ..., X_d)$, for example an Itô process defined on some filtered probability space $(\Omega, \mathbb{P}, \{\mathcal{F}_t\})$

$$\mathrm{d}X = \mu \mathrm{d}t + \sigma \mathrm{d}W(t) \qquad X(0) = X_0.$$

Define a time discretization $0 = t_0 < t_1 < \ldots$ and apply a time discretization scheme, for example an Euler scheme

$$X(t_{i+1}) = X(t_i) + \mu(t_i)\Delta t_i + \sigma(t_i)\Delta W(t_i) \qquad X(t_0) = X_0.$$

Draw some simulation paths $\omega_1, \omega_2, \ldots, \omega_n$, i.e. draw random numbers $\Delta W(t_i, \omega_j)$ generating realizations $X(t_{i+1}, \omega_j)$.



Monte-Carlo Pricing:

Assume that the Numéraire $N(t)$ is a function of X and we have calculated $N(t_i, \omega_i)$. Assume that the time $T = t_k$ value $V(t_k)$ of a derivative product is a known function of $X(t_k)$. Then calculate $V(t_k)$ from $X(t_k)$ and approximate the expectation operator

$$\frac{V(t_0)}{N(t_0)} = \mathrm{E}\left(\frac{V(t_k)}{N(t_k)} \middle| \mathcal{F}_{t_0}\right) \approx \sum_{j=1}^{n} \frac{V(t_k, \omega_j)}{N(t_k, \omega_j)} \cdot \underbrace{\frac{1}{n}}_{= p(\omega_i)}$$

Monte Carlo Pricing: Review and Notation

Examples:

Black-Scholes Model for a Stock

$$X_1 := S \qquad \mathrm{d}S = \mu S\,\mathrm{d}t + \sigma S\,\mathrm{d}W \qquad \text{stock}$$

$$X_2 := B \qquad \mathrm{d}B = rB\mathrm{d}t \qquad \text{risk free money market account}$$

Choice of Numéraire $N := B \quad \Rightarrow \quad$ pricing measure dynamics: $r = \mu$.

LIBOR Market Model

$$X_i := L_i = L(T_i, T_{i+1}) \qquad \mathrm{d}L_i = \mu_i L_i \mathrm{d}t + \sigma_i L_i \mathrm{d}W_i \qquad \text{forward rate for } [T_i, T_{i+1}]$$

for $i = 0, \ldots, m-1$.

Choice of Numéraire $N := P(T_m) \quad \Rightarrow \quad$ pricing measure dynamics: $\mu_j(t) = -\sum_{\substack{l \geq j+1 \\ l \leq m-1}} \frac{\delta_l L_l(t) \cdot \sigma_j(t) \sigma_l(t) \rho_{j,l}(t)}{(1+\delta_l L_l(t))}$

where $P(T_m)$ denotes the zero coupon bond with maturity T_m.

Some Key Problems in Modeling and Monte-Carlo Pricing:

- **Calibration:** In contrast to ”implied modeling” (Dupire) which is usually done in connection with a lattice implementation (PDE / tree), the calibration of process parameters is difficult (complex inverse problem).
  ⇒ See, e.g., the talk of Piterbarg.

- **Sensitivities:** The Monte-Carlo calculation of a sensitivity, i.e. the partial derivative of a price, e.g. with finite differences, has short comings. Monte-Carlo pricing suffers from poor resolution of local properties, thus sensitivities of discontinuous payouts tend to be inaccurate.
  ⇒ See, e.g., likelihood ratio method, pathwise method, etc. [Gl03] and proxy simulation scheme method [FK05].

- **Pricing of Bermudan Options:**
  Subject of this talk
  ⇒ See, e.g., references [BG97], [F06], [LS01], [Pi03], [F05], [F06].

Monte Carlo Pricing of Bermudan Options

Monte Carlo Pricing of Bermudan Options

Bermudan Option on Underlyings $U(T_i)$

Given multiple exercise dates $T_1 < T_2 < T_3 < \ldots < T_n$ at each time T_i the holder has the choice between

[exercise] - choose the value $U(T_i)$ of some underlying financial product

[hold] - choose to exercise later, ie. an Bermudan Option with exercise dates $\{T_{i+1}, \ldots, T_n\}$.

$\rightarrow$ Option on option ... on option.

Let $V_{\{T_i,\ldots,T_n\}}$ denote the value process of a Bermudan option with exercise dates $\{T_i, \ldots, T_n\}$.

Value of Bermudan Option according to optimal exercise

$$V_{\{T_i,\ldots,T_n\}}(T_i) = \max\{U(T_i)\ ,\ V_{\{T_{i+1},\ldots,T_n\}}(T_i)\},$$

where

Conditional expectation at some future time

$$V_{\{T_{i+1},\ldots,T_n\}}(T_i) = N(T_i) \cdot \mathrm{E}^{\mathbb{Q}^N}\left(\frac{V_{\{T_{i+1},\ldots,T_n\}}(T_{i+1})}{N(T_{i+1})} \,\middle|\, \mathcal{F}_{T_i}\right)$$

is the value of the option $V_{\{T_{i+1},\ldots,T_n\}}$ conditioned on the time T_i states ($\mathcal{F}_{T_i}$).

$\Rightarrow$ Requires the calculation of a conditional expectation (difficult in Monte Carlo).

Bermudan Option on Underlyings $U(T_i)$

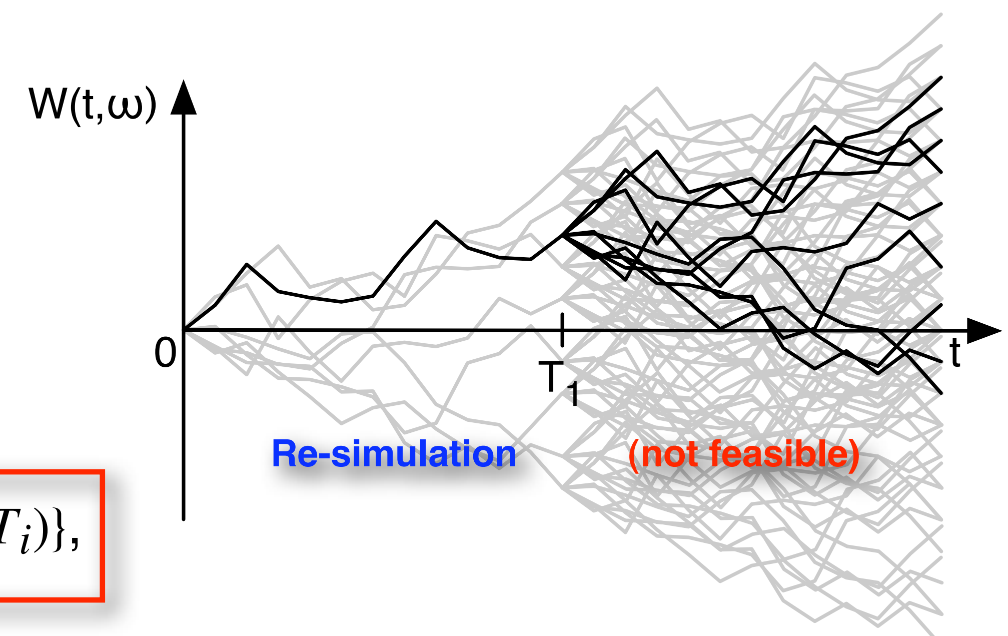
Given multiple exercise dates $T_1 < T_2 < T_3 < \ldots < T_n$ at each time T_i the holder has the choice between

[exercise] - choose the value $U(T_i)$ of some underlying financial product

[hold] - choose to exercise later, ie. an Bermudan Option with exercise dates $\{T_{i+1}, \ldots, T_n\}$.

→ Option on option ... on option.

Let $V_{\{T_i,\ldots,T_n\}}$ denote the value process of a Bermudan option with exercise dates $\{T_i, \ldots, T_n\}$.



Value of Bermudan Option according to optimal exercise

$$V_{\{T_i,\ldots,T_n\}}(T_i) = \max\{U(T_i)\ ,\ V_{\{T_{i+1},\ldots,T_n\}}(T_i)\},$$

where

Conditional expectation at some future time

$$V_{\{T_{i+1},\ldots,T_n\}}(T_i) = N(T_i) \cdot \mathrm{E}^{\mathbb{Q}^N}\left(\frac{V_{\{T_{i+1},\ldots,T_n\}}(T_{i+1})}{N(T_{i+1})} \;\middle|\; \mathcal{F}_{T_i} \right)$$

is the value of the option $V_{\{T_{i+1},\ldots,T_n\}}$ conditioned on the time T_i states ($\mathcal{F}_{T_i}$).

⇒ Requires the calculation of a conditional expectation (difficult in Monte Carlo).

Value of Bermudan Option

Consider relative prices: $\tilde{V} := \frac{V}{N}$ and $\tilde{U} := \frac{U}{N}$. The (Numéraire-relative) value of the Bermudan option is

$$\tilde{V}_{\{T_i,\ldots,T_n\}}(T_i) \;=\; \max\{\tilde{U}(T_i)\;,\;\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_i)\},$$

where

$$\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_i) \;=\; \mathrm{E}^{\mathbb{Q}^N}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \,\middle|\, \mathcal{F}_{T_i}\right)$$

is the (Numéraire-relative) value of the option $\tilde{V}_{\{T_{i+1},\ldots,T_n\}}$ conditioned on the time T_i states ($\mathcal{F}_{T_i}$).

Value of Bermudan Option: Optimal Stopping Formulation

For a given path $\omega \in \Omega$ let

$$\tau(\omega) := \min\{T_i : \underbrace{\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_i) < \tilde{U}(T_i)}_{\text{exercise criteria}}\}.$$

$\tau(\omega)$ is the optimal admissible exercise time on a given path ω.
Note: τ is a stopping time, i.e. $\{T \leq T_k\} \subset \mathcal{F}_{T_k}$.

This allows to express the Bermudan option value as a single (unconditioned) expectation:

$$\tilde{V}_{\{T_1,\ldots,T_n\}}(T_0) = \mathrm{E}^{\mathbb{Q}}(\tilde{U}(\tau) \mid \mathcal{F}_{T_0}).$$

Here $\tilde{U}(\tau)[\omega] := \tilde{U}(\tau(\omega), \omega)$ is the value realized on path ω by exercising (optimal) in $\tau(\omega)$.

This is just an equivalent formulation. It remains to calculate the stopping time through the optimal exercise criteria.

Next: Calculate the random variable $\tilde{U}(\tau)$ directly. → Backward Algorithm.

Value of Bermudan Option: The Backward Algorithm

Recursively define the values $\tilde{V}_i$ by *induction backward in time*:

Induction start: If the Bermudan is not exercised on the last exercise date T_n it's value is 0:

$$\tilde{V}_{n+1} \equiv 0$$

Induction step $i+1 \rightarrow i$ for $i = n, \ldots, 1$:

$$\tilde{V}_i = \begin{cases} \tilde{V}_{i+1} & \text{if } \tilde{U}(T_i) < \mathrm{E}^{\mathbb{Q}}(\tilde{V}_{i+1} | \mathcal{F}_{T_i}) \\ \tilde{U}(T_i) & \text{else.} \end{cases}$$

Clearly, the value $\tilde{V}_1$ coincides with the optimal exercise value $\tilde{U}(\tau)$ and we have

$$\tilde{V}_{\{T_1,\ldots,T_n\}}(T_0) \ = \ \mathrm{E}^{\mathbb{Q}}(\tilde{V}_1 \mid \mathcal{F}_{T_0}) \ \approx \ \sum_{j=1}^{n} \tilde{V}_1(\omega_j) \cdot \frac{1}{n}.$$

With the backward algorithm, the whole problem of pricing Bermudan problems has been moved to the estimation of the exercise criteria.

Value of Bermudan Option: The Backward Algorithm

Recursively define the values $\tilde{V}_i$ by *induction backward in time*:

Induction start: If the Bermudan is not exercised on the last exercise date T_n it's value is 0:

$$\tilde{V}_{n+1} \equiv 0$$

Induction step $i+1 \rightarrow i$ for $i = n, \ldots, 1$:

$$\tilde{V}_i = \begin{cases} \tilde{V}_{i+1} & \text{if } \tilde{U}(T_i) < \mathrm{E}^{\mathbb{Q}}(\tilde{V}_{i+1} | \mathcal{F}_{T_i}) \\ \tilde{U}(T_i) & \text{else.} \end{cases}$$

The difficulty to calculate conditional expectation is now restricted to the determination of the exercise criteria

Clearly, the value $\tilde{V}_1$ coincides with the optimal exercise value $\tilde{U}(\tau)$ and we have

$$\tilde{V}_{\{T_1,\ldots,T_n\}}(T_0) \;=\; \mathrm{E}^{\mathbb{Q}}(\tilde{V}_1 \,\big|\, \mathcal{F}_{T_0}) \;\approx\; \sum_{j=1}^{n} \tilde{V}_1(\omega_j) \cdot \frac{1}{n}.$$

With the backward algorithm, the whole problem of pricing Bermudan problems has been moved to the estimation of the exercise criteria.

Monte Carlo Pricing of Bermudan Options

Value of Bermudan Option: The Backward Algorithm

Recursively define the values $\tilde{V}_i$ by *induction backward in time*:

Induction start: If the Bermudan is not exercised on the last exercise date T_n it's value is 0:

$$\tilde{V}_{n+1} \equiv 0$$

Induction step $i+1 \rightarrow i$ for $i = n, \ldots, 1$:

$$\tilde{V}_i = \begin{cases} \tilde{V}_{i+1} & \text{if } \tilde{U}(T_i) < \mathrm{E}^{\mathbb{Q}}(\tilde{V}_{i+1} | \mathcal{F}_{T_i}) \\ \tilde{U}(T_i) & \text{else.} \end{cases}$$

The difficulty to calculate conditional expectation is now restricted to the determination of the exercise criteria

Clearly, the value $\tilde{V}_1$ coincides with the optimal exercise value $\tilde{U}(\tau)$ and we have

$$\tilde{V}_{\{T_1,\ldots,T_n\}}(T_0) \;=\; \mathrm{E}^{\mathbb{Q}}(\tilde{V}_1 \,|\, \mathcal{F}_{T_0}) \;\approx\; \sum_{j=1}^{n} \tilde{V}_1(\omega_j) \cdot \frac{1}{n}.$$

(unconditional expectation)

With the backward algorithm, the whole problem of pricing Bermudan problems has been moved to the estimation of the exercise criteria.

Monte Carlo Pricing of Bermudan Options

Problem: How to Estimate the Exercise Criteria:

The true exercise criteria is

$$\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_i) < \tilde{U}(T_i) \qquad \text{i.e.} \qquad \mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid \mathcal{F}_{T_i}\right) < \tilde{U}(T_i)$$

Solution: Tools for Bermudan Pricing in Monte Carlo:

- Optimization of Exercise Criteria (Andersen [An99])
  Choose some parametrized exercise criteria. Maximize the option price as a function of the parameters. Note: Suboptimal exercise criteria will lead to smaller option value. See references.

- Estimation of Conditional Expectation - Subject of this talk, see [F05].

  – Binning (see also [F06]).

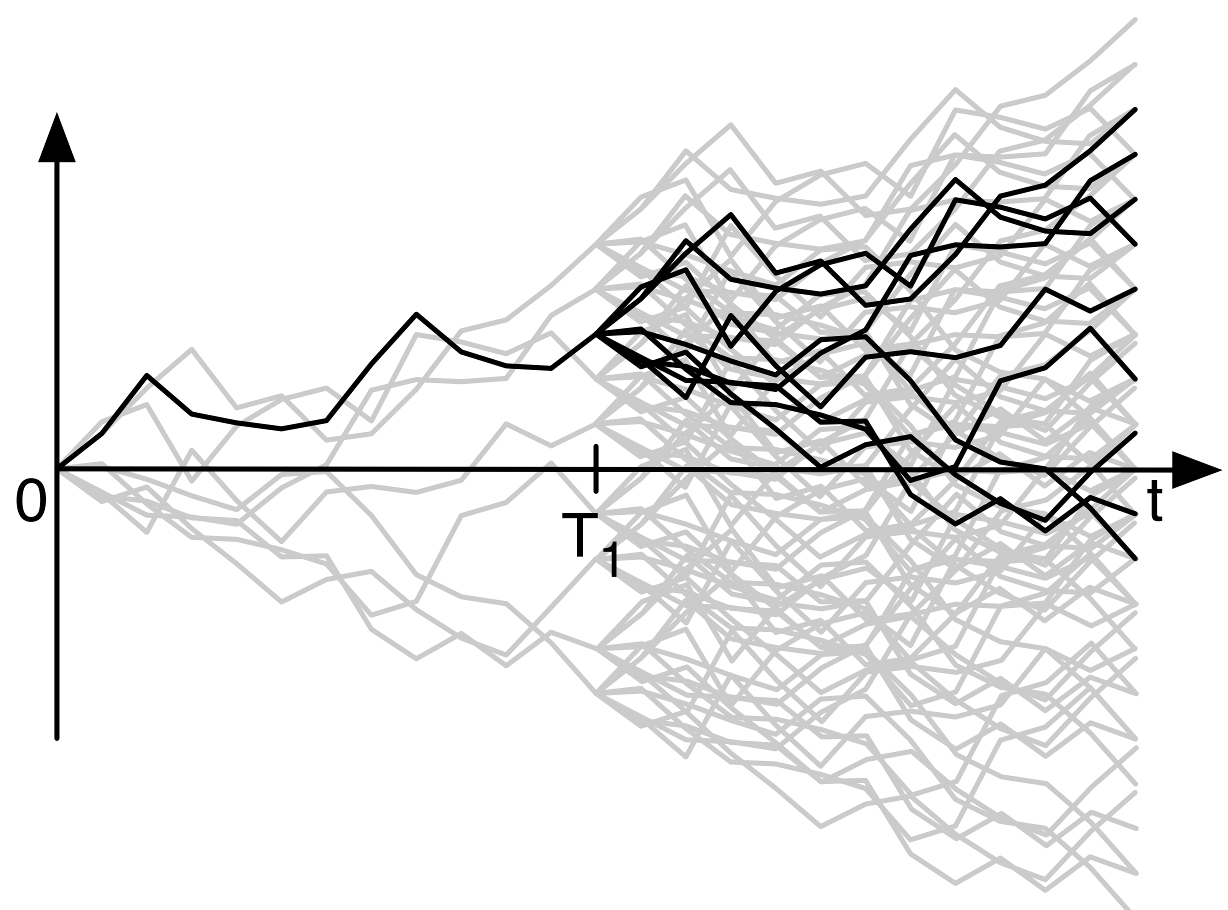
  – Least-Square Regression (Carriere [Ca96]aka. Longstaff-Schwarz [LS01], see also [CLP0

- Dual Method / Primal-Dual Method / Dual Problem (Davis & Karatzas [DK94], Rogers 2001 [Ro01])
  Optimizes the stopping time. See references.

Condition Expectation Estimators

Full Re-simulation

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid \mathcal{F}_{T_i}\right)[\omega_j] \approx \frac{1}{n_i}\sum_{k=1}^{n_j} \tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1})[\omega_{j,k}]$$



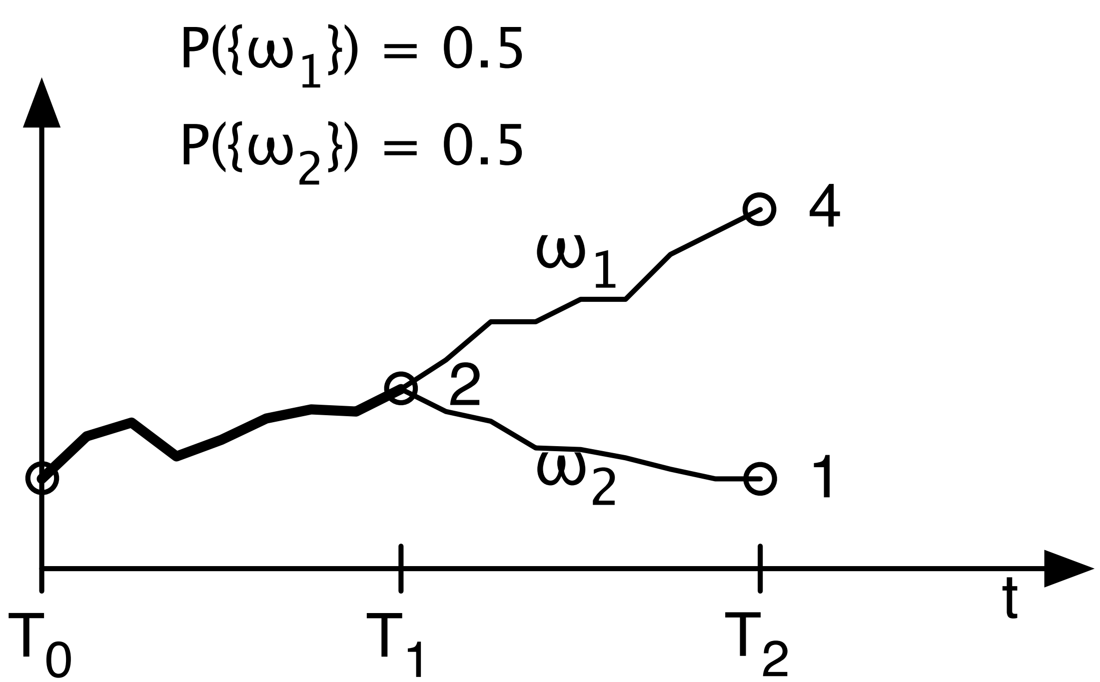
Full re-simulation is not feasible.

The computational cost grow exponentially in the number of exercise dates.

Perfect Foresight

Estimate the expectation by taking the path on which we are on:

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid \mathcal{F}_{T_i}\right)[\omega] \;\approx\; \tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1};\omega)$$



Example:

Optimal (admissible) exercise strategy:
$\tau(\omega_1) = \tau(\omega_2) = T_2$
Average value realized = 0.5 • 1 + 0.5 • 4 = 2.5
(optimal exercise)

Perfect foresight exercise strategy:
$\tau(\omega_1) = T_2$; $\tau(\omega_2) = T_1$
Average value realized = 0.5 • 2 + 0.5 • 4 = 3
(super-optimal exercise)

Perfect foresight largely over-estimates the option value. (More on this later).

Conditional Expectation as Functional Dependence

The filtration $\mathcal{F}_{T_1}$ represents the information known up to T_1. If we consider our Monte-Carlo simulation of our model primitives X, we see that all that is known up to T_1 is the finite set of random variables

$$Z := (X(t_0), X(t_1), X(t_2), \ldots, X(T_1))$$

The conditional expectation is a function of the $\mathcal{F}_{T_1}$-measurable random variable Z:

$$\mathrm{E}^{\mathbb{Q}^N}\left(\tilde{V}(T_2) \mid \mathcal{F}_{T_1}\right) = \mathrm{E}^{\mathbb{Q}^N}\left(\tilde{V}(T_2) \mid Z\right) = f(Z).$$

Depending on the product, the expectation will depend only on a very few random variables. E.g. if the product is not path dependent it will depend only on

$$Z := X(T_1).$$

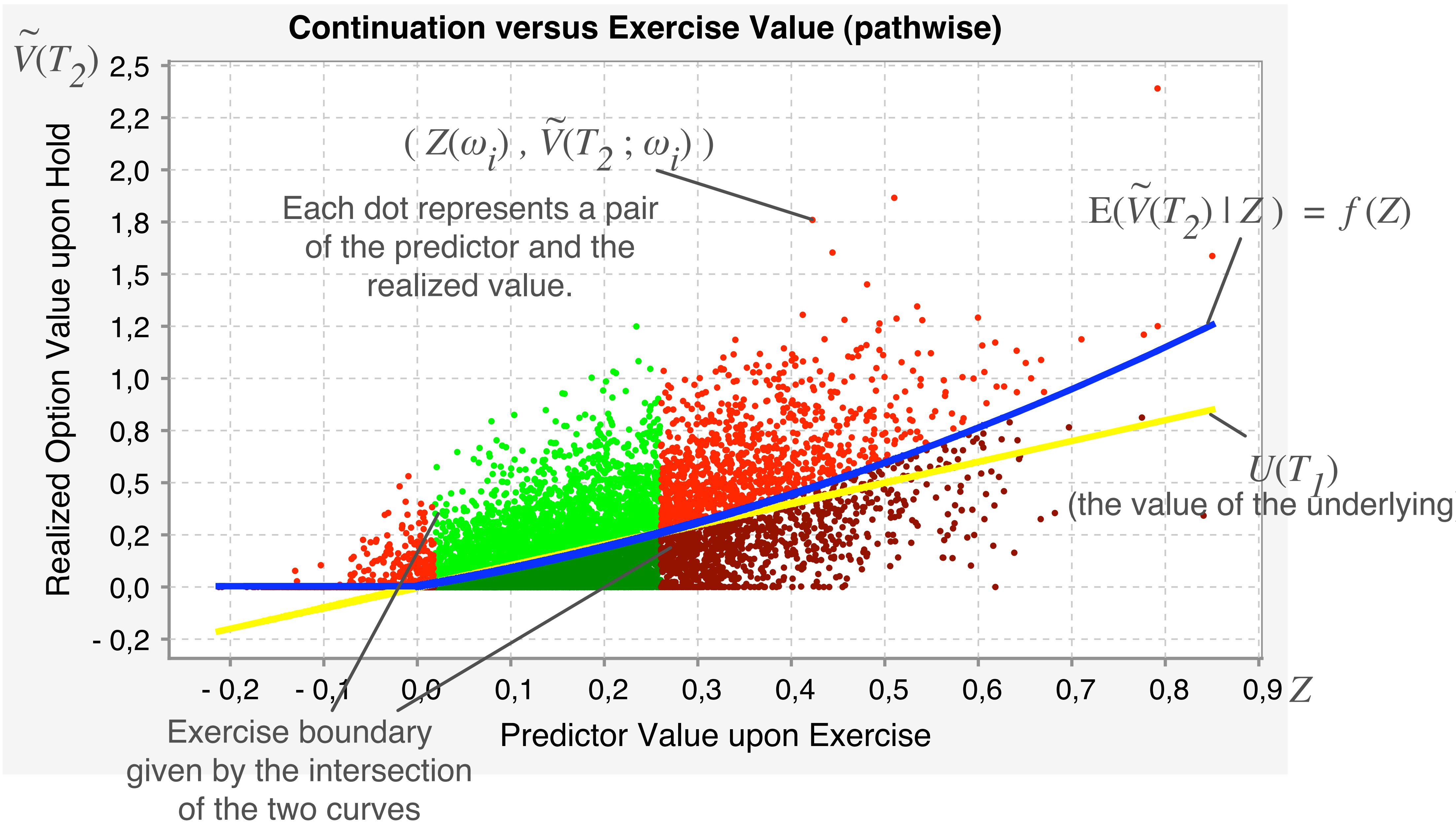
With this notation the backward algorithm's induction step $i+1 \rightarrow i$ becomes:

$$\tilde{V}_i = \begin{cases} \tilde{V}_{i+1} & \text{if } \tilde{U}(T_i) < f(Z(T_i)) \\ \tilde{U}(T_i) & \text{else.} \end{cases}$$

Estimation of Exercise Criteria $\Leftrightarrow$ Estimation of Cond. Expectation $\Leftrightarrow$ Estimation of $z \mapsto f(z)$

Monte-Carlo Conditional Expectation Estimators

Illustration



Remark: One method to estimate the conditional expectation is to estimate the function $z \mapsto f(z)$ as a regression polynomial.

Binning

Estimate the expectation by taking paths which are "nearby´´:

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; \mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z \in U_\epsilon(Z(\omega))\right),$$

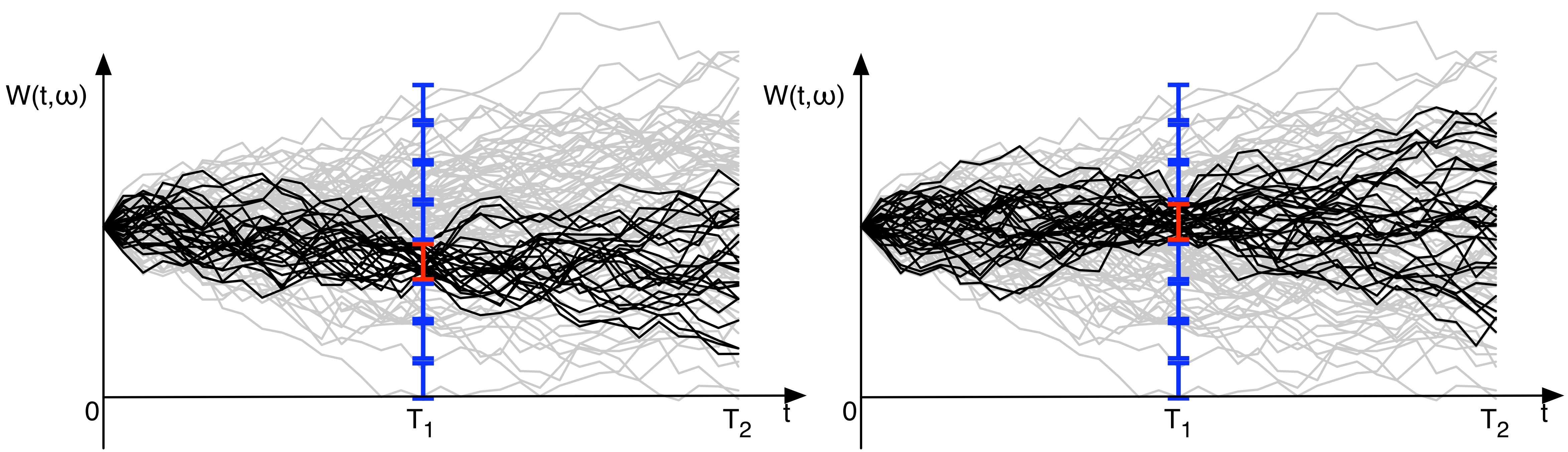
where $U_\epsilon(Z(\omega)) := \{z \mid \|Z(\omega) - z\| < \epsilon\}$ (set of paths which are *near* ω).

Instead of defining a bin $U_\epsilon(Z(\omega))$ for each path ω it is more efficient to start with a partition of $Z(\Omega)$:

Let $Z(\Omega) = \bigcup_k Z_k$, with Z_k disjoint. The binning approximation of the conditional expectation is

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; \mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z \in Z_k\right) \;=:\; H_k,$$

where Z_k denote the set with $Z(\omega) \in Z_k$.



Monte-Carlo Conditional Expectation Estimators

Binning

Let $Z(\Omega) = \cup_k Z_k$, with Z_k disjoint. Estimate the expectation by taking paths which are "nearby´´:

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; \mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z \in Z_k\right) \;=:\; H_k,$$

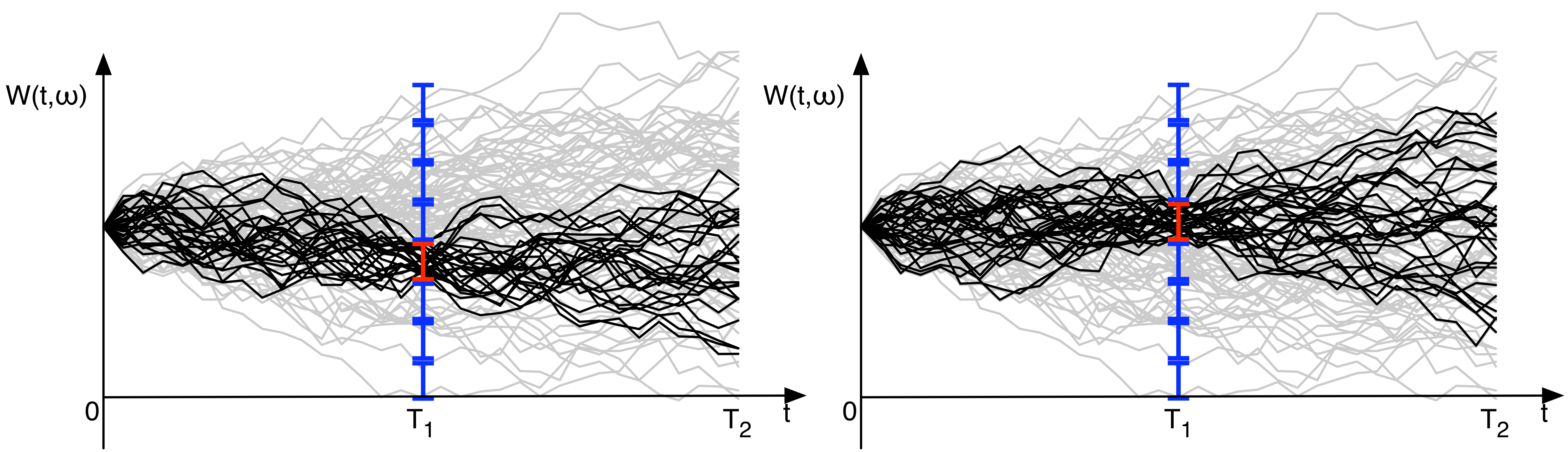
where Z_k denote the set with $Z(\omega) \in Z_k$.

Remark: Binning estimates the conditional expectation as a piecewise constant function

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; f(Z(\omega)),$$

where

$$f(Z(\omega)) = H_k \quad \text{for } k \text{ such that } Z(\omega) \in Z_k.$$

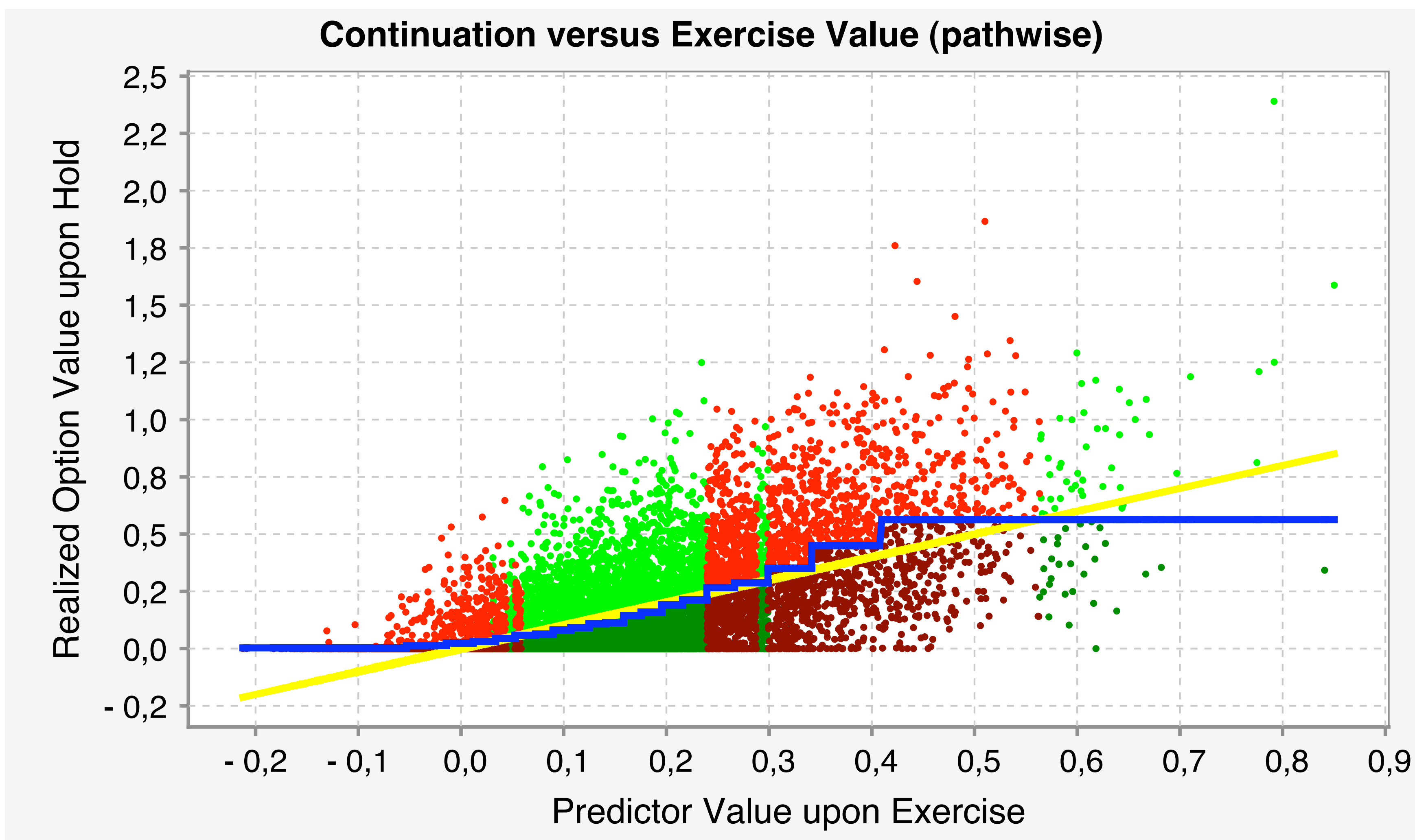


Binning

Let $Z(\Omega) = \cup_k Z_k$, with Z_k disjoint. Estimate the expectation by taking paths which are "nearby´´:

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; \mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z \in Z_k\right) \;=:\; H_k,$$

where Z_k denote the set with $Z(\omega) \in Z_k$.



Monte-Carlo Conditional Expectation Estimators

Regression

Estimate the conditional expectation as the best fit $\mathcal{F}_{T_i}$ function to $\mathcal{F}_{T_{i+1}}$ data:

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; f(Z(\omega), \alpha^*),$$

where

$$\alpha^* = \operatorname*{argmin}_{\alpha} \|\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) - f(Z(\omega), \alpha)\|.$$

Lemma (Linear Regression): Let $\Omega^* = \{\omega_1, \ldots, \omega_n\}$ be a given sample space, $V : \Omega^* \to \mathbb{R}$ and $Y := (Y_1, \ldots, Y_p) : \Omega^* \to \mathbb{R}^p$ given random variables. Furthermore let

$$f(y_1, \ldots, y_p, \alpha_1, \ldots, \alpha_p) := \sum \alpha_i y_i.$$

Then we have for any α^* with $X^{\mathsf{T}} X \alpha^* = X^{\mathsf{T}} v$

$$\|V - f(Y, \alpha^*)\|_{L_2(\Omega^*)} = \min_{\alpha} \|V - f(Y, \alpha)\|_{L_2(\Omega^*)},$$

where

$$X := \begin{pmatrix} Y_1(\omega_1) & \ldots & Y_p(\omega_1) \\ \vdots & & \vdots \\ Y_1(\omega_n) & \ldots & Y_p(\omega_n) \end{pmatrix}, \qquad v := \begin{pmatrix} V(\omega_1) \\ \vdots \\ V(\omega_n) \end{pmatrix}.$$

If $(X^{\mathsf{T}} X)^{-1}$ exists then $\alpha^* := (X^{\mathsf{T}} X)^{-1} X^{\mathsf{T}} v$.

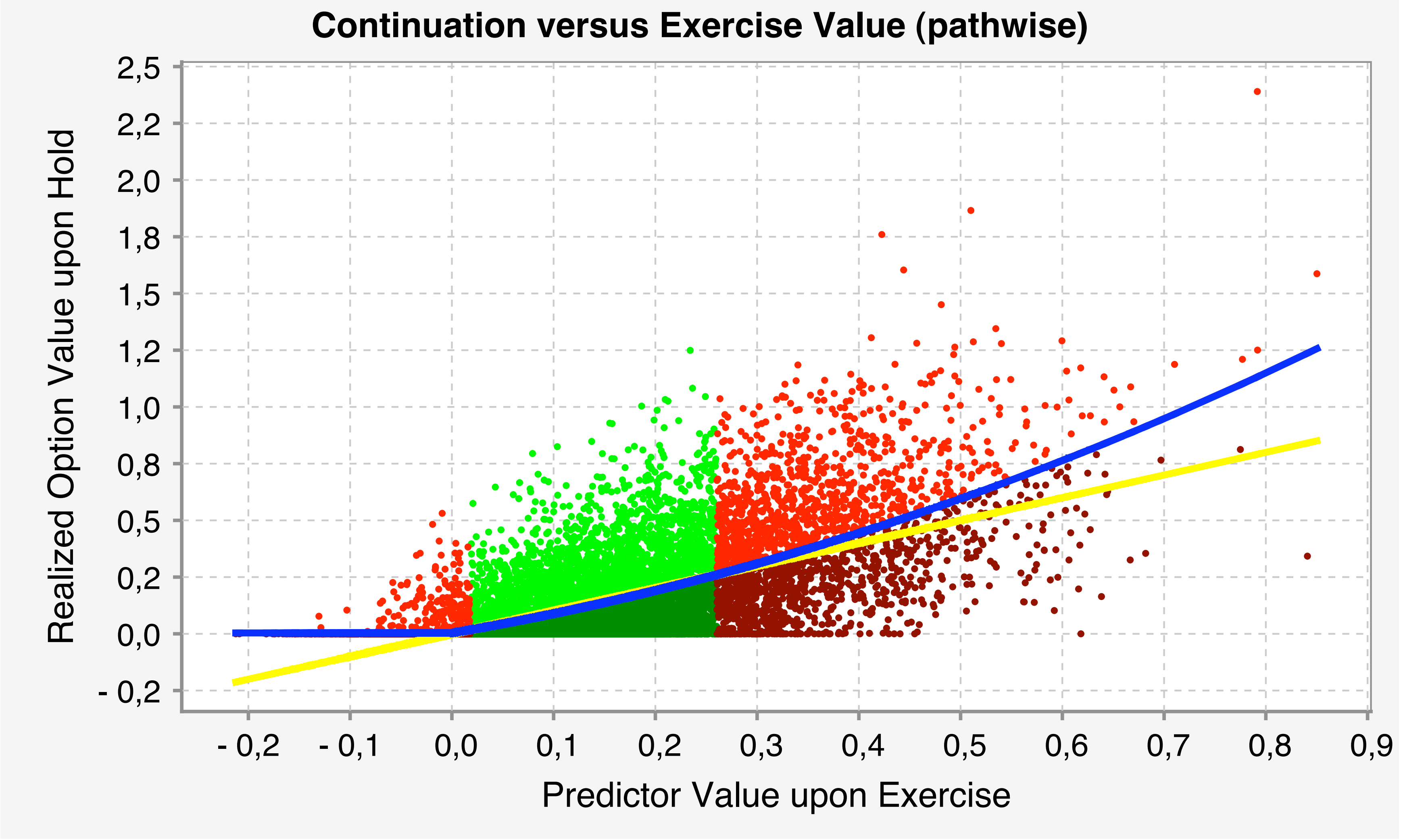
Regression

Estimate the conditional expectation as the best fit $\mathcal{F}_{T_i}$ function to $\mathcal{F}_{T_{i+1}}$ data:

$$\mathrm{E}^{\mathbb{Q}}\left(\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) \mid Z\right)[\omega] \;\approx\; f(Z(\omega), \alpha^*),$$

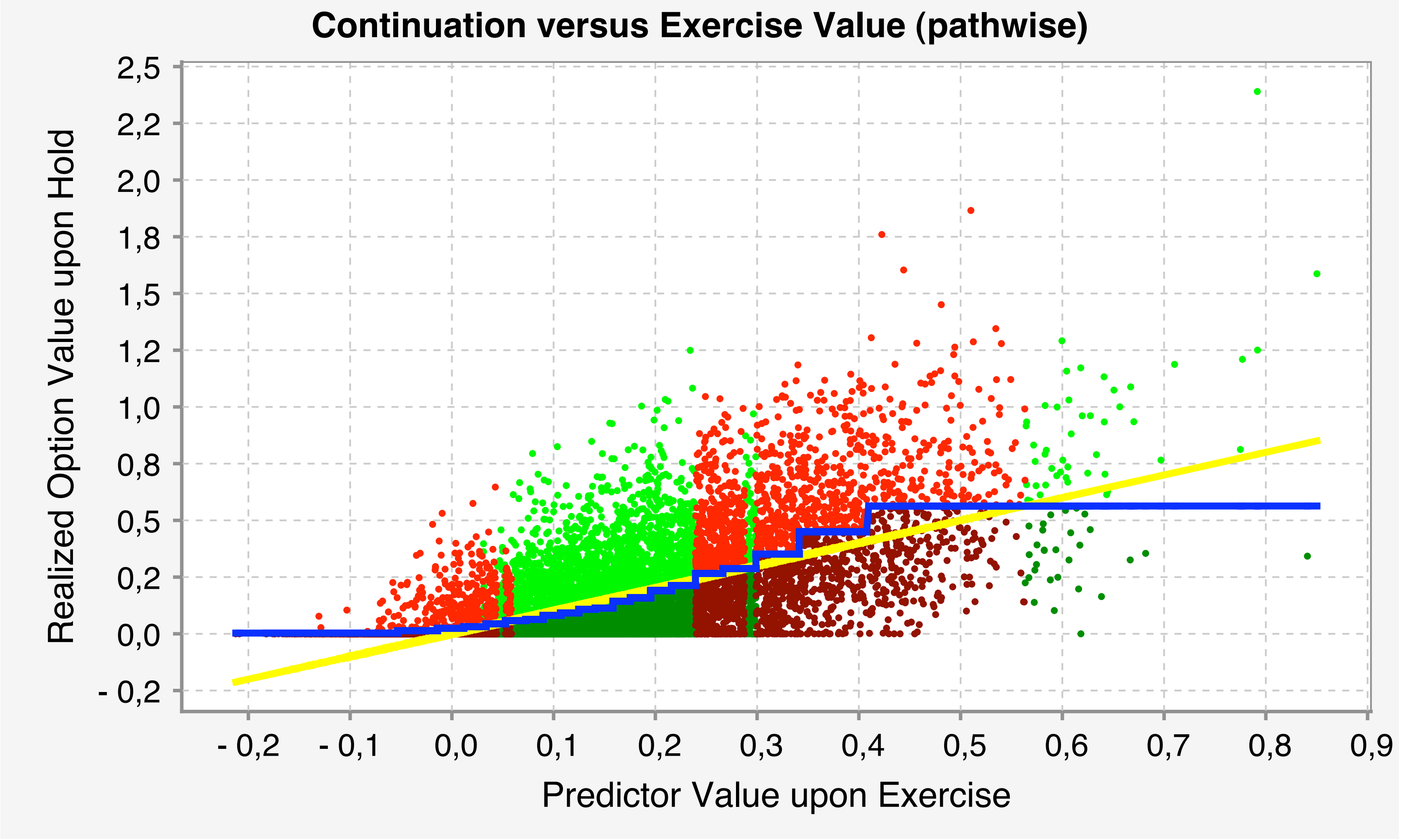
where

$$\alpha^* = \underset{\alpha}{\operatorname{argmin}} \|\tilde{V}_{\{T_{i+1},\ldots,T_n\}}(T_{i+1}) - f(Z(\omega), \alpha)\|.$$



Binning (revisited)

Binning is equivalent to a least square regression with piecewise constant basis functions. See [F06].



The Foresight Bias

Foresight Bias: Classification

Notation

Here and in the following we will consider the exercise criteria $\max(U, \mathrm{E}(\tilde{V} \mid Z))$, i.e. $\tilde{V}$ stands for $\tilde{V}_{i+1}$ and U stands for $U(T_i)$, for some i.

$$f(Z) = \mathrm{E}^{\mathrm{est}}(\tilde{V} \mid Z) = \mathrm{E}(\tilde{V} \mid Z) + \underbrace{\epsilon}_{\text{Monte Carlo error}} .$$

The Foresight Bias

Consider the optimal exercise value $\max(U, \mathrm{E}(\tilde{V} \mid Z))$ where the conditional expectation estimator has a Monte Carlo error which we denote by ϵ. Then the foresight bias is given by:

$$\mathrm{E}\big(\max(U, \mathrm{E}(\tilde{V} \mid Z) + \epsilon) \,\big|\, Z\big) \;=\; \max(U, \mathrm{E}(\tilde{V} \mid Z)) + \text{foresightbias}.$$

Note that

$$\mathrm{E}\big(\max(U, \mathrm{E}(\tilde{V} \mid Z) \,\big|\, Z\big) = \mathrm{E}\big(\max(U_{|Z}, \mathrm{E}(\tilde{V} \mid Z) \,\big|\, Z\big)$$

Conditional to Z the underlying U is a constant: Z contains all information $\mathcal{F}_{T_i}$ and U is $\mathcal{F}_{T_i}$-measurable. We therefore write $K := U_{|Z}$ and consider

$$\mathrm{E}\big(\max(K, \mathrm{E}(\tilde{V} \mid Z) + \epsilon \,\big|\, Z\big)$$

Doesn't that look familiar?

The foresight bias is the value of the option on the Monte-Carlo error.

Numerical Removal of the Foresight Bias

The standard approach to remove the foresight bias is to use two independent Monte-Carlo simulations. One will be used to estimate the exercise criteria (as a functional dependence on some state variable), the other will be used to calculate the payouts (using the backward algorithm).

The numerical removal of the foresight bias has two disadvantages:

- Numerical removal of the foresight bias slows down the pricing. Two independent Monte-Carlo simulations of the stochastic processes have to be generated. For some models (e.g. high dimensional interest rate models like the LIBOR Market Model) the generation of the Monte-Carlo paths is relatively time consuming.

- Numerical removal of the foresight bias makes the code of the implementation cumbersome. It is a desired design pattern to separate the stochastic process model and the generation of the Monte-Carlo paths from product pricing. The structure of the code will likely become less clear if a second independent simulation has to be created.

Removing foresight bias numerically, the Monte-Carlo error on the conditional expectation estimator will lead to **sub-optimal exercise**.

$\Rightarrow$ The Bermudan option price will be biased low.

Foresight Bias: Classification

Let $\epsilon_1(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$, i.e.

$$f(Z) + \epsilon_1(Z) \;:=\; \mathrm{E}(V \mid Z) + \epsilon_1(Z) \;=\; \mathrm{E}(V + \epsilon_1(Z) \mid Z).$$

Let $\epsilon_2(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$ in an independent simulation.

Foresight Biased Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_1(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_1(Z) \end{cases}$$

Option on the Monte-Carlo error

Numerical Removal of Foresight Bias:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Optimal Exercise with Monte-Carlo Error in Payout:

$$\begin{cases} U & U > f(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) \end{cases}$$

Desired exercise in Monte-Carlo pricing

Sub-Optimal Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Foresight Bias: Classification

Let $\epsilon_1(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$, i.e.

$$f(Z) + \epsilon_1(Z) \ := \ \mathrm{E}(V \mid Z) + \epsilon_1(Z) \ = \ \mathrm{E}(V + \epsilon_1(Z) \mid Z).$$

Let $\epsilon_2(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$ in an independent simulation.

Foresight Biased Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_1(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_1(Z) \end{cases}$$

Option on the Monte-Carlo error

Numerical Removal of Foresight Bias:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Optimal Exercise with Monte-Carlo Error in Payout:

$$\begin{cases} U & U > f(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) \end{cases}$$

Desired exercise in Monte-Carlo pricing

Sub-Optimal Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Foresight Bias: Classification

Let $\epsilon_1(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$, i.e.

$$f(Z) + \epsilon_1(Z) \ := \ \mathrm{E}(V \mid Z) + \epsilon_1(Z) \ = \ \mathrm{E}(V + \epsilon_1(Z) \mid Z).$$

Let $\epsilon_2(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$ in an independent simulation.

Foresight Biased Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_1(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_1(Z) \end{cases}$$

Option on the Monte-Carlo error

Numerical Removal of Foresight Bias:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Optimal Exercise with Monte-Carlo Error in Payout:

$$\begin{cases} U & U > f(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) \end{cases}$$

Desired exercise in Monte-Carlo pricing

Sub-Optimal Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Foresight Bias: Classification

Let $\epsilon_1(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$, i.e.

$$f(Z) + \epsilon_1(Z) \;:=\; \mathrm{E}(V \mid Z) + \epsilon_1(Z) \;=\; \mathrm{E}(V + \epsilon_1(Z) \mid Z).$$

Let $\epsilon_2(Z)$ denote the Monte-Carlo error of $\mathrm{E}(V \mid Z)$ in an independent simulation.

Foresight Biased Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_1(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_1(Z) \end{cases}$$

Option on the Monte-Carlo error

Numerical Removal of Foresight Bias:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Optimal Exercise with Monte-Carlo Error in Payout:

$$\begin{cases} U & U > f(Z) \\ V + \epsilon_1(Z) & U \leq f(Z) \end{cases}$$

Desired exercise in Monte-Carlo pricing

Sub-Optimal Exercise:

$$\begin{cases} U & U > f(Z) + \epsilon_2(Z) \\ V & U \leq f(Z) + \epsilon_2(Z) \end{cases}$$

Sub-optimal exercise

Foresight Bias

Is the Foresight Bias Negligible?

Is the Foresight Bias negligible?

An alternative to the numerical removal of the foresight bias is to not remove the foresight bias at all. This approach may be justified by the fact that the foresight bias will tend to zero as the number of paths tends to infinity. In addition the foresight bias is rather small, usually it is within Monte-Carlo errors.

However neglecting foresight bias may create larger relative errors when considering multiple exercise dates, a book of multiple options with foresight bias, or the aggregation of prices from independent Monte-Carlo simulation.

- For a single Bermudan option with few exercise dates the foresight bias is of the order of the Monte-Carlo error of the option itself.
- For a single Bermudan option with many exercise dates, the foresight biases induced at each exercise date may add up, the bias is still of the order of the Monte-Carlo error.
- For a large portfolio the foresight bias may become significant.
- For the aggregation of prices from independent Monte-Carlo simulation the foresight bias may become significant. (Parallelization Problem)

Aggregating Foresight Biased Options

However, summing up different options - each with a foresight bias and a Monte-Carlo error - may change the picture. If two options differ in strike or maturities their Monte-Carlo errors may become more and more independent.* Consider a book of n options (compound or Bermudan). If the n options have independent Monte-Carlo errors with standard deviation σ the Monte-Carlo error for the portfolio will be $\sqrt{n} \cdot \sigma$. But since the foresight bias is a systematic error it will grow linearly in n, i.e. if the n options have a foresight bias β the book will exhibit a foresight bias of $n \cdot \beta$. Assuming that for a family of options β and σ are of the same size we could say that: only if the Monte-Carlo errors of the single product prices are perfectly correlated we would have that the ratio of foresight bias to Monte-Carlo error $\frac{\beta}{\sigma}$ of a portfolio does not grow with the portfolio size.

This is also obvious from the interpretation of the foresight bias as an option on the (individual) Monte-Carlo error. The book will contain n such options.† In the end we have that the foresight bias may likely become significant.‡

*As example consider the two payouts $\min(\max(S(T), a_1), b_1)$ and $\min(\max(S(T), a_2), b_2)$ (i.e. $S(T)$ capped and floored). If (a_1, b_1) and (a_2, b_2) are disjoint a sampling of $S(T)$ will (in general) generate independent Monte-Carlo errors for the two payouts.

†Of course foresight bias may cancel if one averages short options with long options.

‡Our test case in [F05] exhibited a foresight bias 0.5 of the Monte-Carlo error. Pricing a book of 16 options may result in a foresight bias around 2 standard deviations (the 95% quantile).

Parallelization of Monte-Carlo Pricing of Bermudan Options

Lemma (Parallelization Lemma)
For the pricing of European products (i.e. products that do not involve an optionality with conditional expectation estimator) we have that the pricing error of the average price of two (independent) Monte-Carlo simulations with n paths is equal to the pricing error of a single Monte-Carlo simulation with $2 \cdot n$ paths.

Thus: We may parallelize the pricing of non-Bermudan products.

k (independent) simulations with n paths = 1 simulation with $k \cdot n$ paths

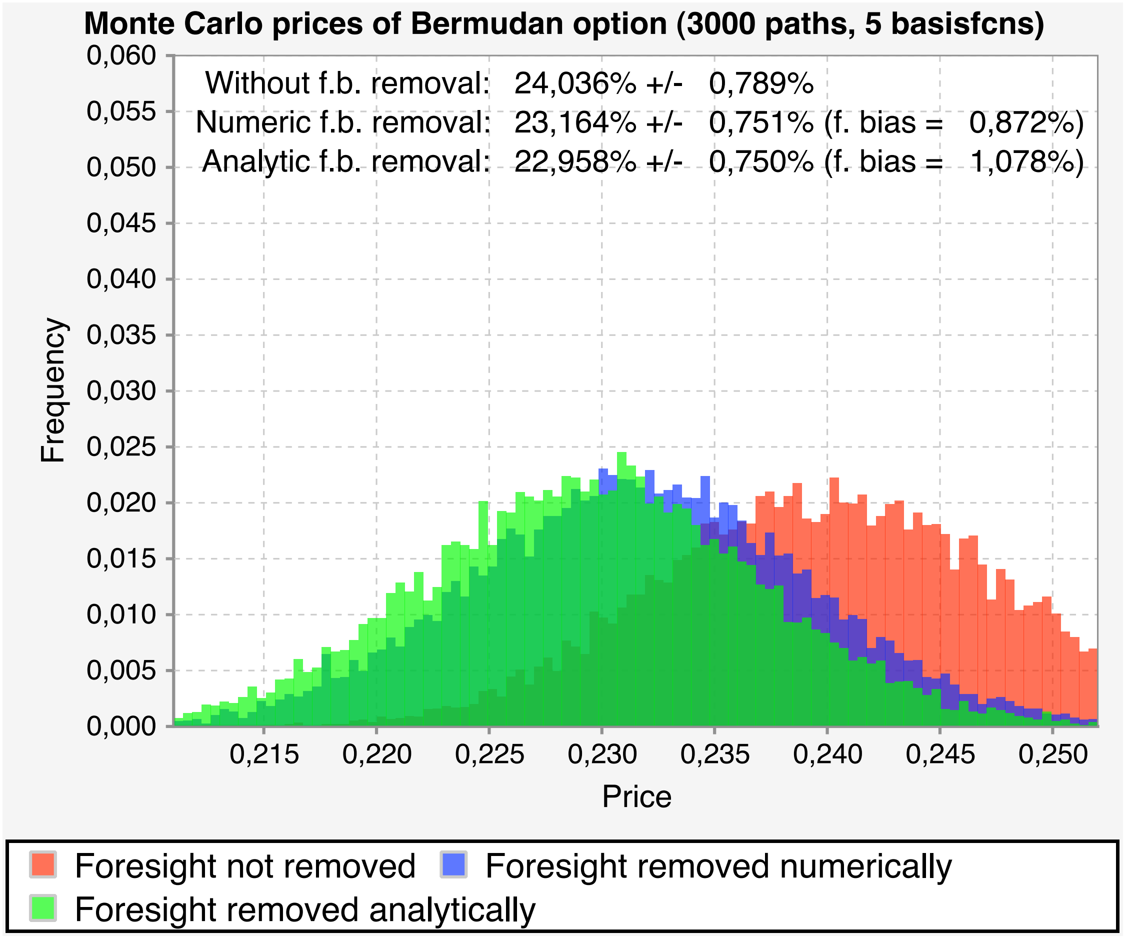
Remark: This lemma does not hold for Bermudan options.

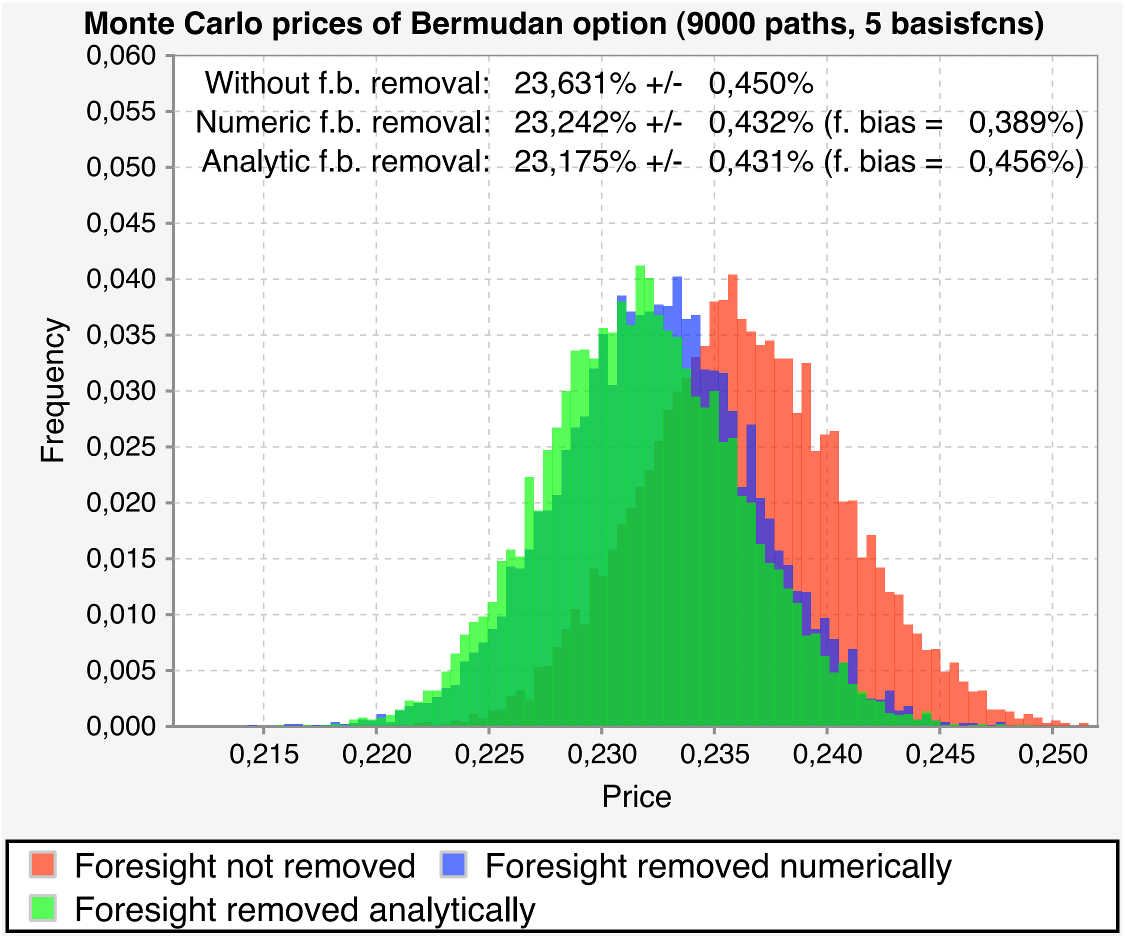
The Monte-Carlo price of a Bermudan option exhibits either a bias high due to the foresight bias or, if foresight is removed, a bias low due to the sub-optimal exercise induced by the Monte-Carlo error on the exercise criteria.

Averaging Bermudan prices of independent Monte-Carlo simulations
will not reduce the systematic error of foresight or sub-optimal exercise.

Numerical Result: Parallelization of Monte-Carlo Simulation

Repeated pricing with independent Monte-Carlo simulations (different random number seed) shows the distribution of the Monte-Carlo error.
The foresight bias is a systematic error, it corresponds to a shift of the mean.





Foresight Bias

Analytic Calculation

Estimation of the Foresight Bias

We want to asses the foresight bias induced by a Monte-Carlo error ϵ of the conditional expectation estimator $\mathrm{E}(\tilde{V} \mid Z)$, i.e. we consider the optimal exercise criteria

$$\max(U, \mathrm{E}(\tilde{V} \mid Z) + \epsilon).$$

Conditioned on a given $Z = z^*$ we assume that ϵ has normal distribution with mean 0 and standard deviation σ for fixed $\mathrm{E}(\tilde{V} \mid Z)$. Then we have the following result for the foresight bias:

Lemma 1: (Estimation of Foresight Bias)

Given a conditional expectation estimator of $\mathrm{E}(\tilde{V}|Z)$ with (conditional) Monte-Carlo error ϵ having normal distribution with mean 0 and standard deviation σ will result in a bias of the conditional mean of $\max(K, \mathrm{E}(\tilde{V}|Z) + \epsilon)$ given by

$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}, \qquad (1)$$

where $\mu := \mathrm{E}(\tilde{V}|Z)$, $\phi(x) := \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2)$ and $\Phi(x) := \int_{-\infty}^{x} \phi(\xi)\, \mathrm{d}\xi$.

Proof:

Let ϵ have Normal distribution with mean 0 and standard deviation σ. For $a, b \in \mathbb{R}$ we have with $\mu^* := b - a$

$$\begin{aligned}
\mathrm{E}(\max(a, b+\epsilon)) &= \mathrm{E}(\max(0, b-a+\epsilon)) + a = \mathrm{E}(\max(0, \mu^* + \epsilon)) + a \\
&= \frac{1}{\sigma}\int_0^\infty x \cdot \phi(\frac{x-\mu^*}{\sigma})\,\mathrm{d}x + a = \frac{1}{\sigma}\int_{-\mu^*}^\infty (x+\mu^*) \cdot \phi(\frac{x}{\sigma})\,\mathrm{d}x + a \\
&= \int_{-\frac{\mu^*}{\sigma}}^\infty (\sigma \cdot x + \mu^*) \cdot \phi(x)\,\mathrm{d}x + a \\
&= \sigma \cdot \phi(\frac{\mu^*}{\sigma}) + \mu^* \cdot (1 - \Phi(-\frac{\mu^*}{\sigma})) + a,
\end{aligned}$$

where we used $\int x\phi(x)\,\mathrm{d}x = \phi(x)$.

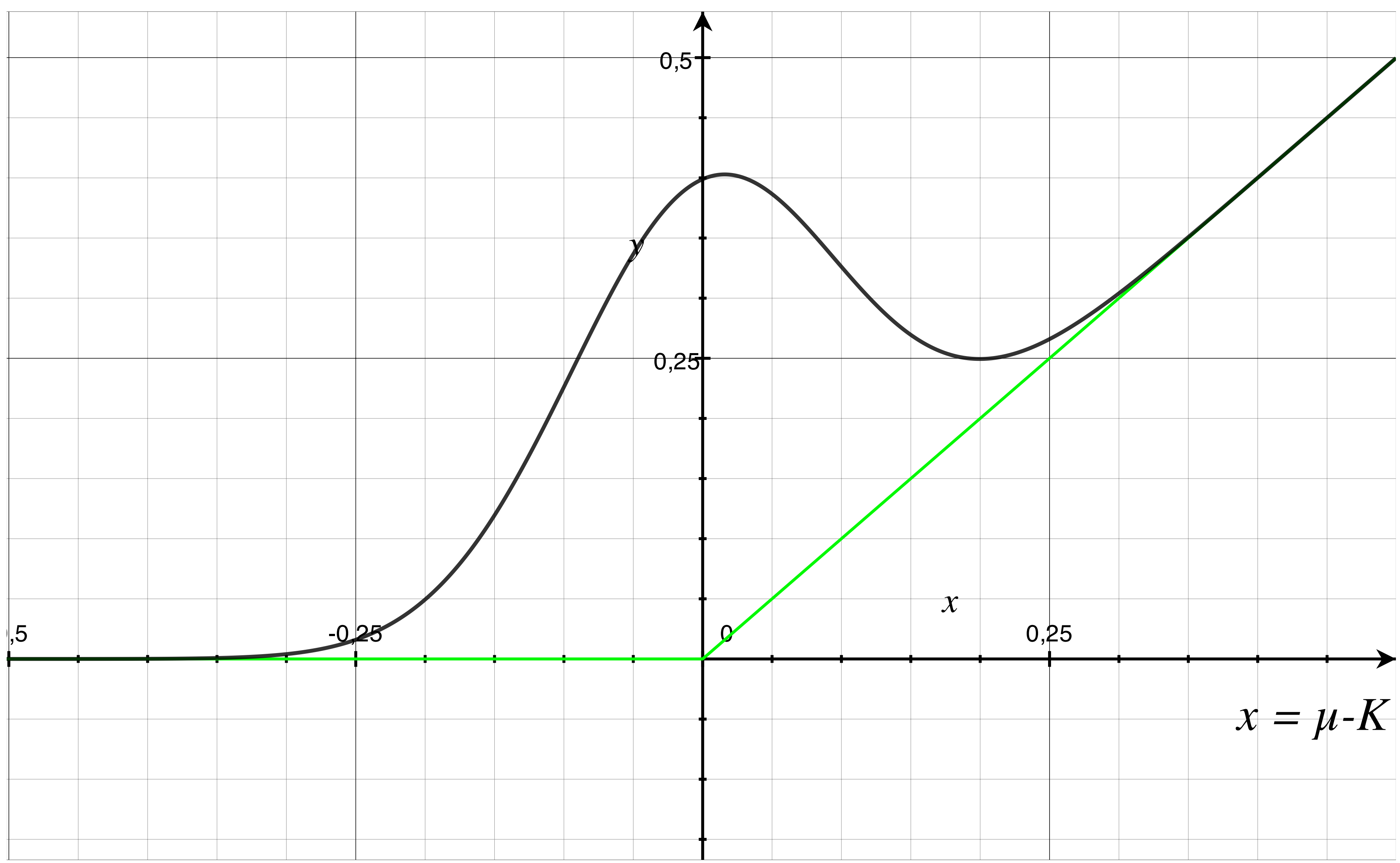
The result follows with $b = \mathrm{E}(\tilde{V}|Z)$, $a := K$, i.e. $\mu^* = \mu - K$.

Remark

The bias induced by the Monte-Carlo error of the conditional expectation estimator consists of two parts: The first part in (1) consists of the systematic one sided bias resulting from the non linearity of the $\max(a, b+x)$ function. The second part is a diffusion of the original payoff function. The Monte-Carlo error smears out the original payoff. The first part should be attributed to super-optimal exercise due to foresight, the second part to sub-optimal exercise due to Monte-Carlo uncertainty.

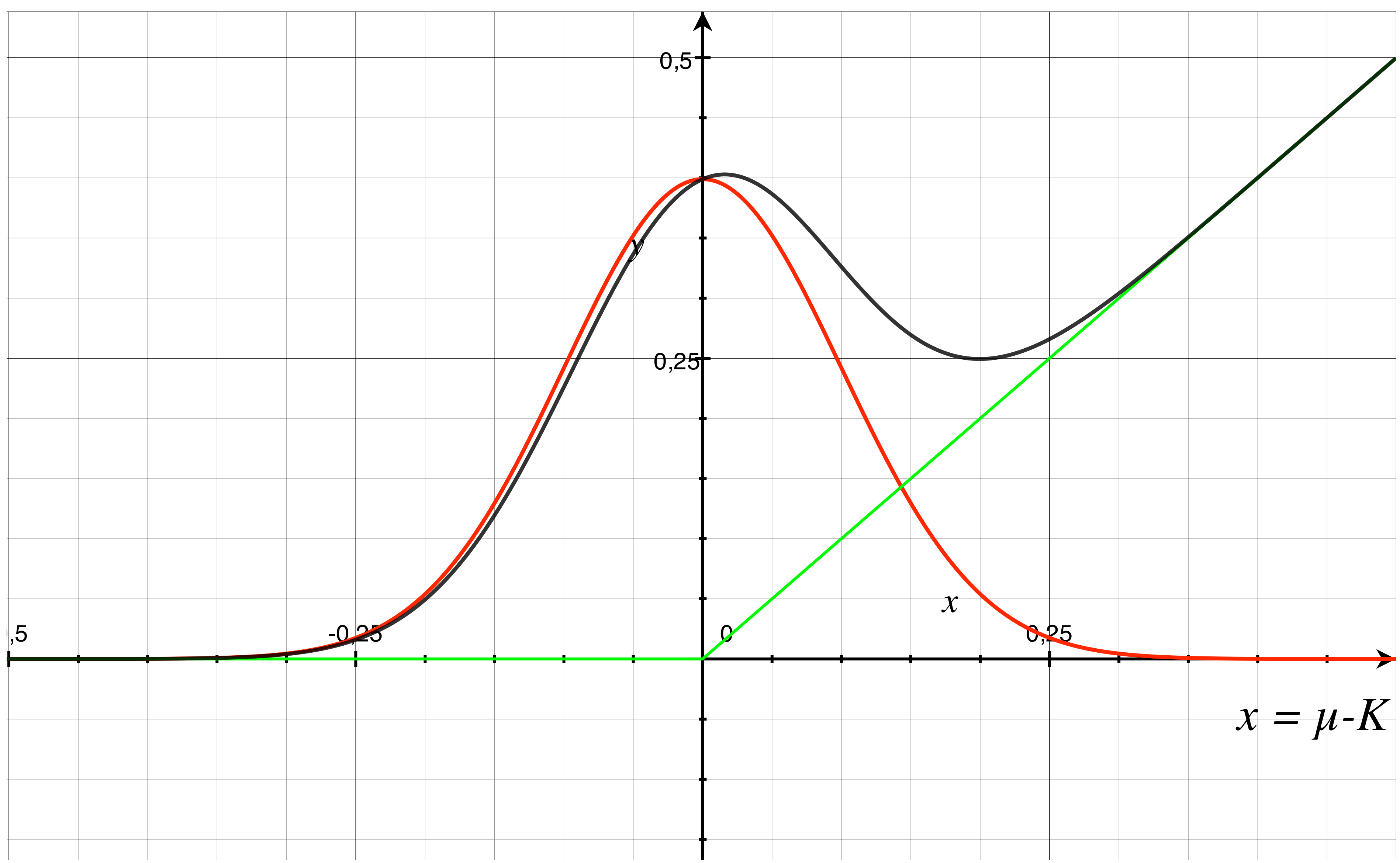
Foresight Biased Payout Function: Interpretation

$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}$$



Foresight Biased Payout Function: Interpretation

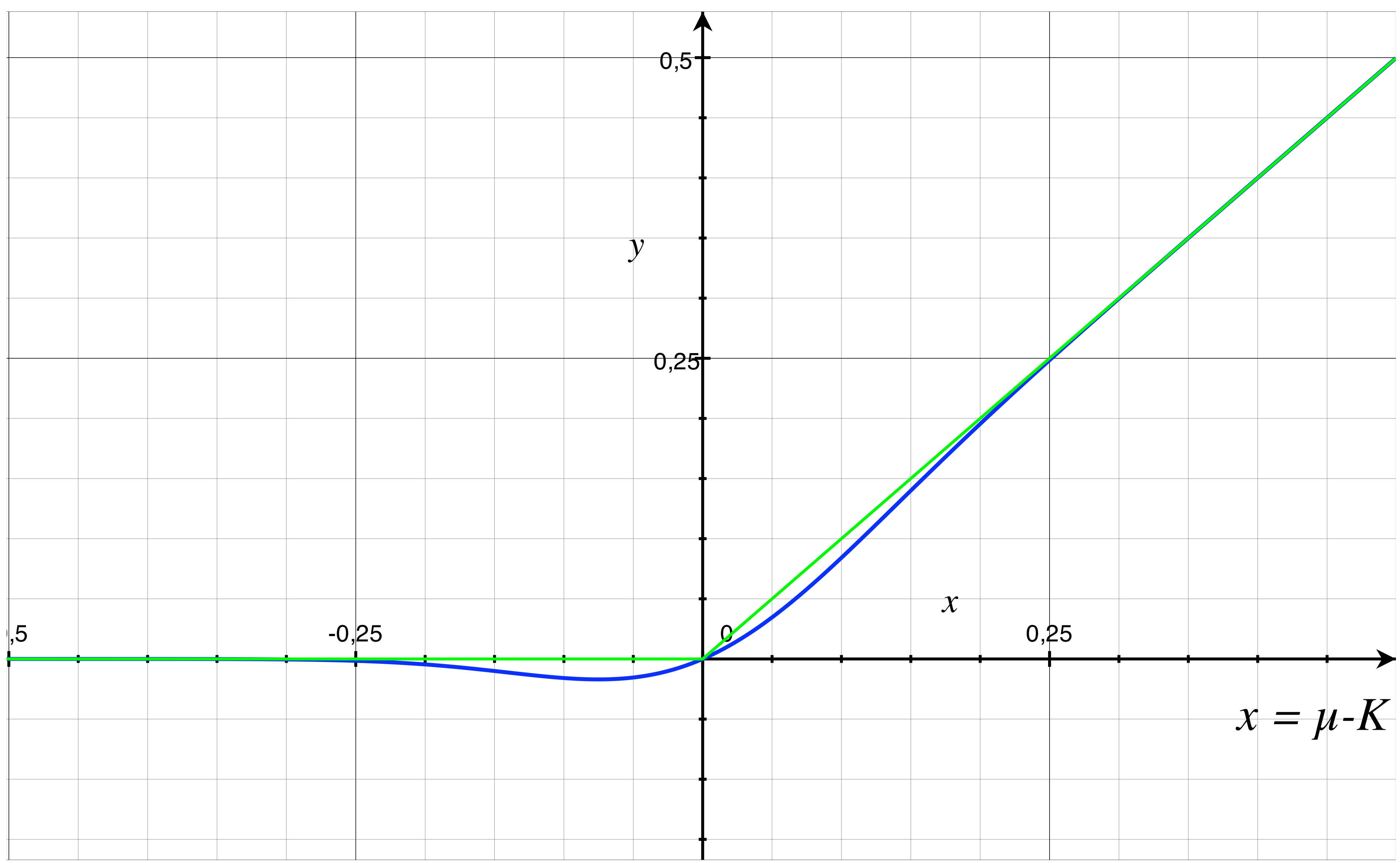
$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}$$



Foresight Bias: Analytic Calculation

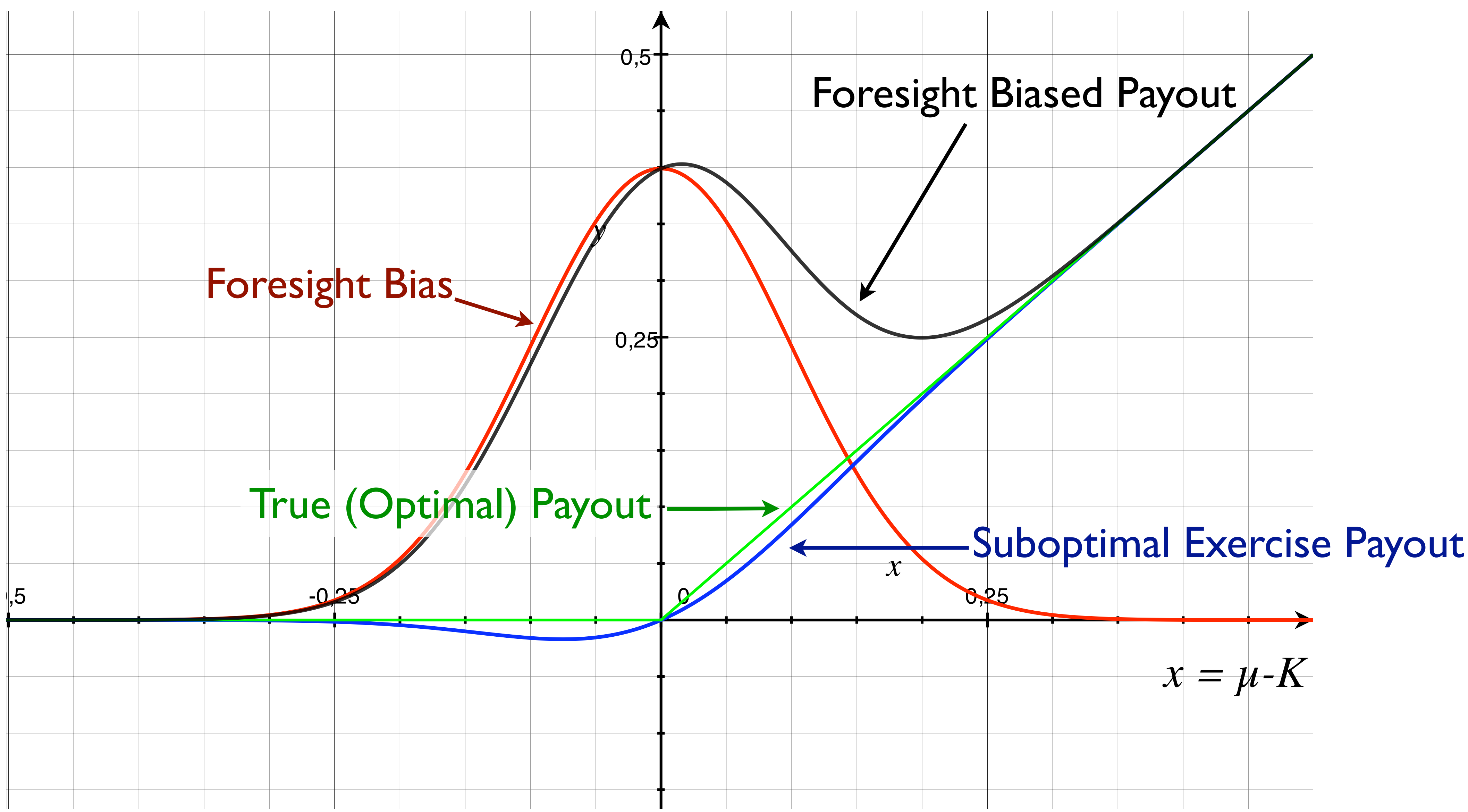
Foresight Biased Payout Function: Interpretation

$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}$$



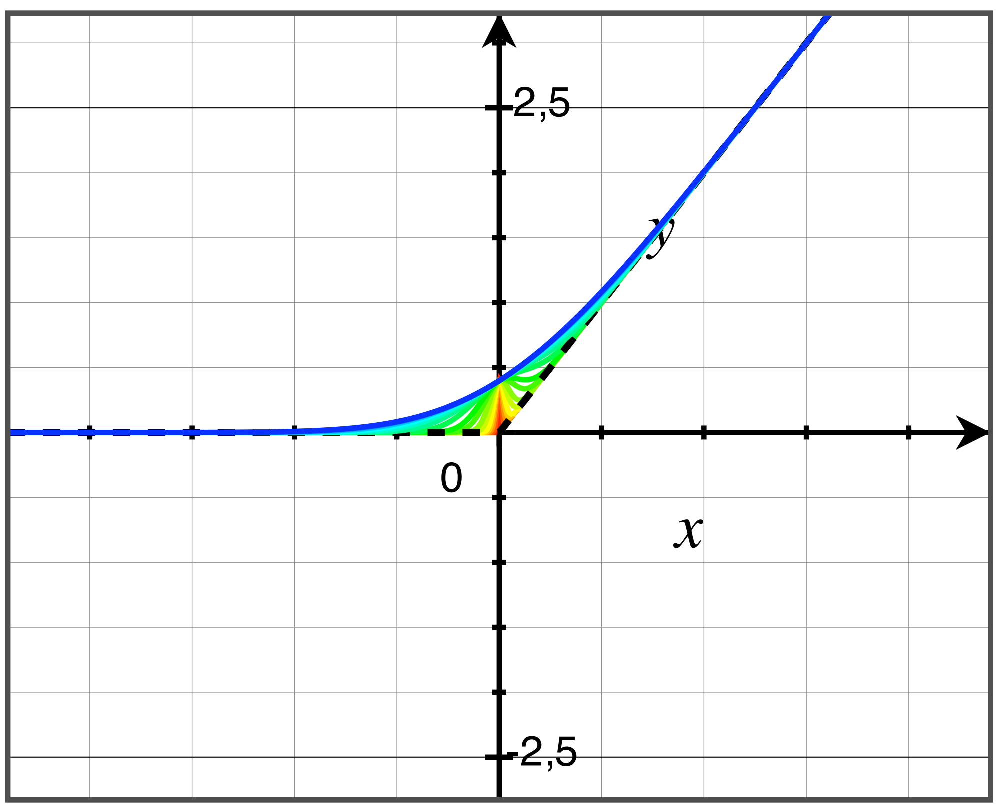
Foresight Biased Payout Function: Interpretation

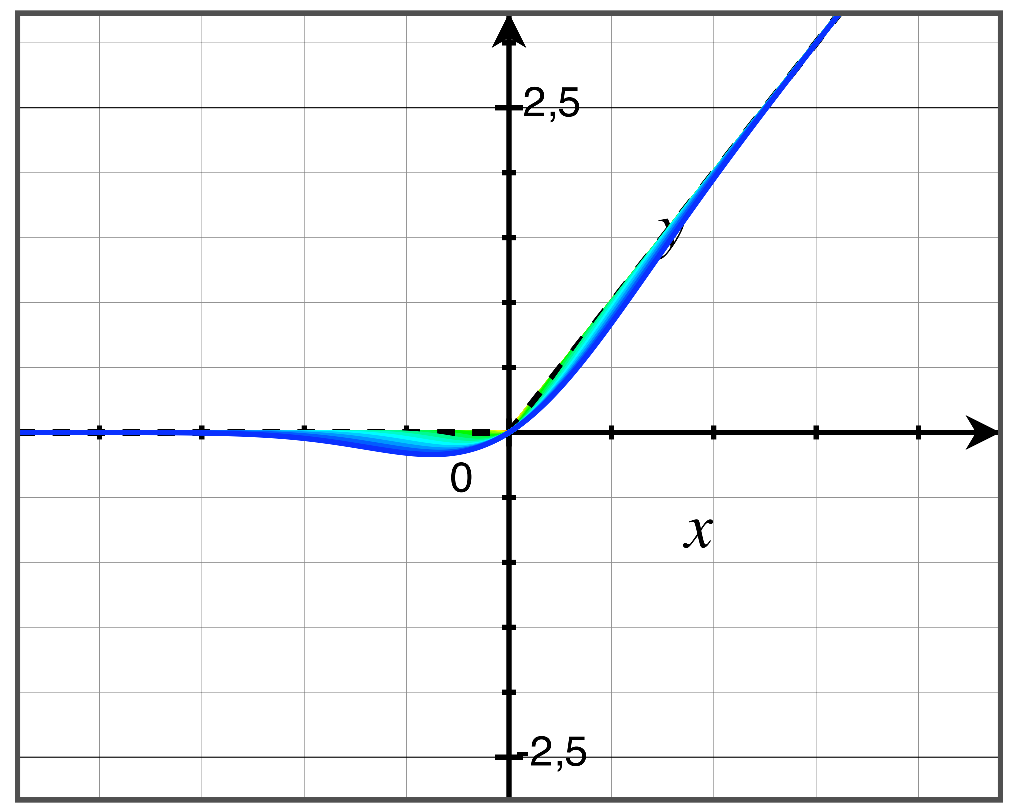
$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}$$



Foresight Biased Payout Function: Interpretation

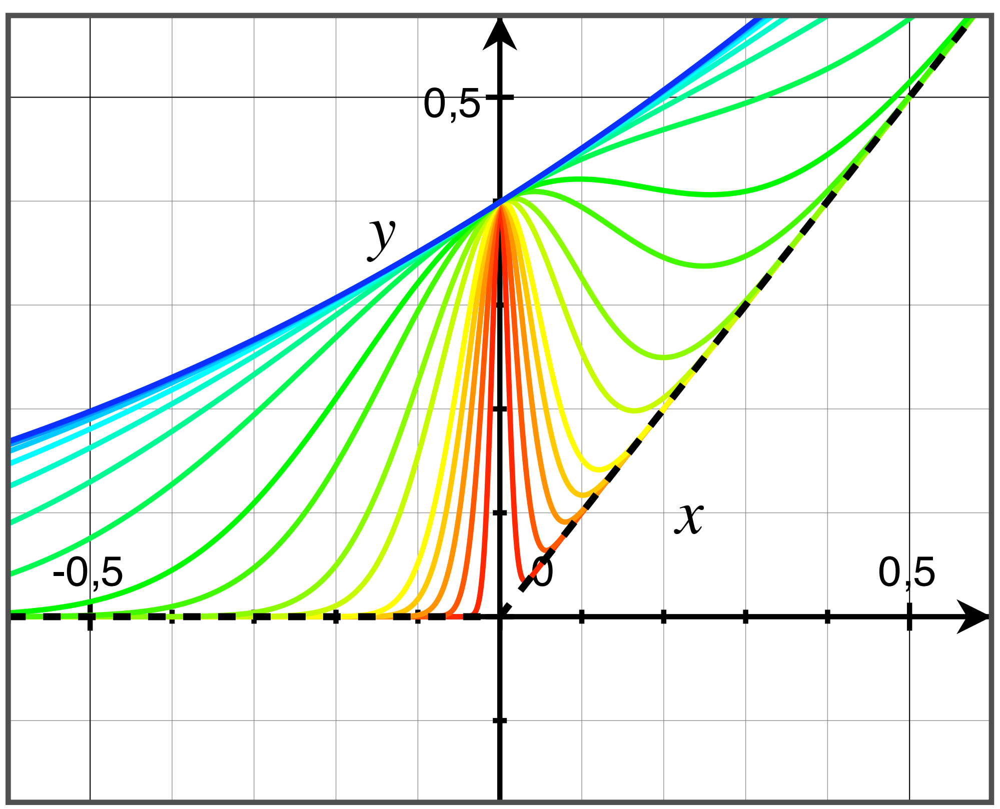
$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}$$

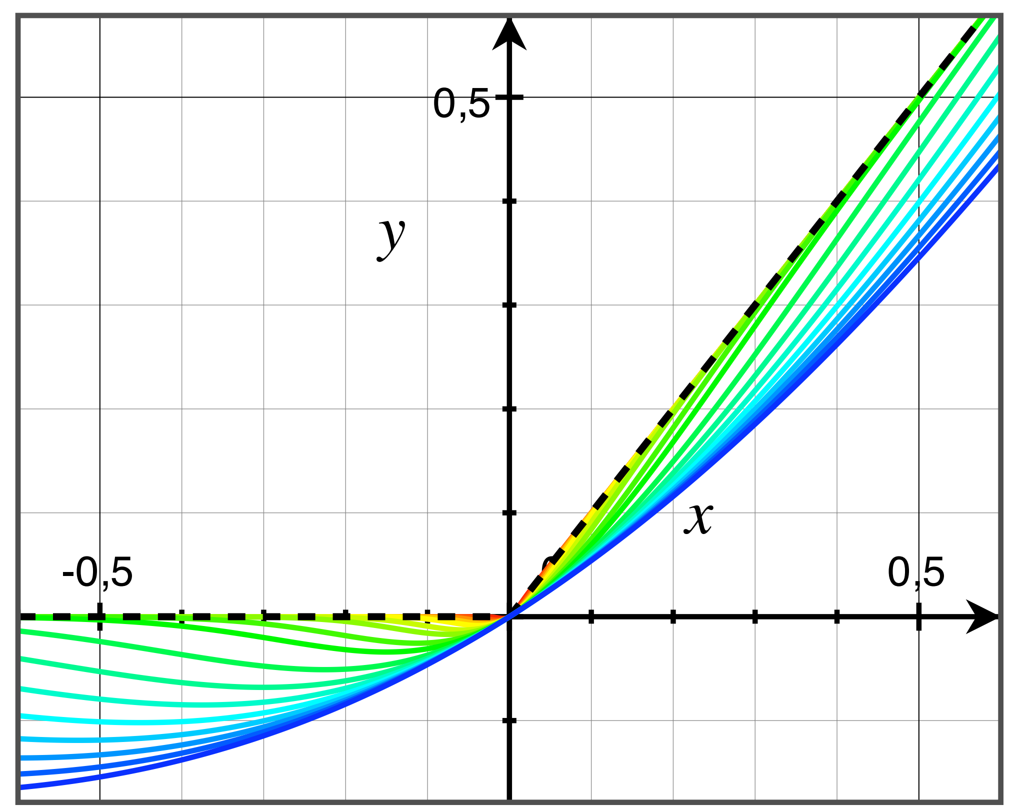




Foresight Biased Payout Function: Interpretation

$$\underbrace{\underbrace{\sigma \cdot \phi(-\frac{\mu - K}{\sigma})}_{\text{foresight bias}}}_{\text{biased high}} + \underbrace{\underbrace{(\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) + K}_{\text{smoothed payout}} - \underbrace{\max(K, \mathrm{E}(\tilde{V}|Z))}_{\text{true payout}}}_{\text{diffusive part, biased low}}$$





Foresight Bias

Analytic Correction in a Backward Algorithm

Foresight Bias Correction

With the notation as in Lemma 1 we define

$$\beta := \sigma \cdot \phi(-\frac{\mu - K}{\sigma})$$

as the *foresight bias correction* of the optimal exercise criteria

$$\max(K, \mathrm{E}(\tilde{V}|Z)),$$

where $\mu := \mathrm{E}(\tilde{V}|Z)$ and σ^2 is the variance of the Monte-Carlo error ϵ of the estimator μ.

Sub-optimality Correction

With the notation as in Lemma 1 we define

$$\gamma := (\mu - K) \cdot (1 - \Phi(-\frac{\mu - K}{\sigma})) - \max(0, \mu - K)$$

as the *suboptimal exercise correction* of the optimal exercise criteria

$$\max(K, \mathrm{E}(\tilde{V}|Z)),$$

where $\mu := \mathrm{E}(\tilde{V}|Z)$ and σ^2 is the variance of the Monte-Carlo error ϵ of the estimator μ.

Numerical Implementation

Estimation of foresight and sub-optimality correction:
In each iteration step of the backward algorithm we estimate the foresight bias correction β^{est} and the sub-optimality correction γ^{est} as follows:

- Calculate $\mu^{\mathrm{est}} = \mathrm{E}^{\mathrm{est}}(\tilde{V} \mid Z)$ using your favored conditional expectation estimator.
- Estimate the Monte-Carlo error σ^{est}.
- The exercise boundary is $K = U_{|Z}$
- $\beta^{\mathrm{est}} := \sigma^{\mathrm{est}} \cdot \phi(\frac{\mu^{\mathrm{est}} - K}{\sigma^{\mathrm{est}}})$
- $\gamma^{\mathrm{est}} := (\mu^{\mathrm{est}} - K) \cdot (1 - \Phi(-\frac{\mu^{\mathrm{est}} - K}{\sigma^{\mathrm{est}}})) - \max(0, \mu^{\mathrm{est}} - K)$

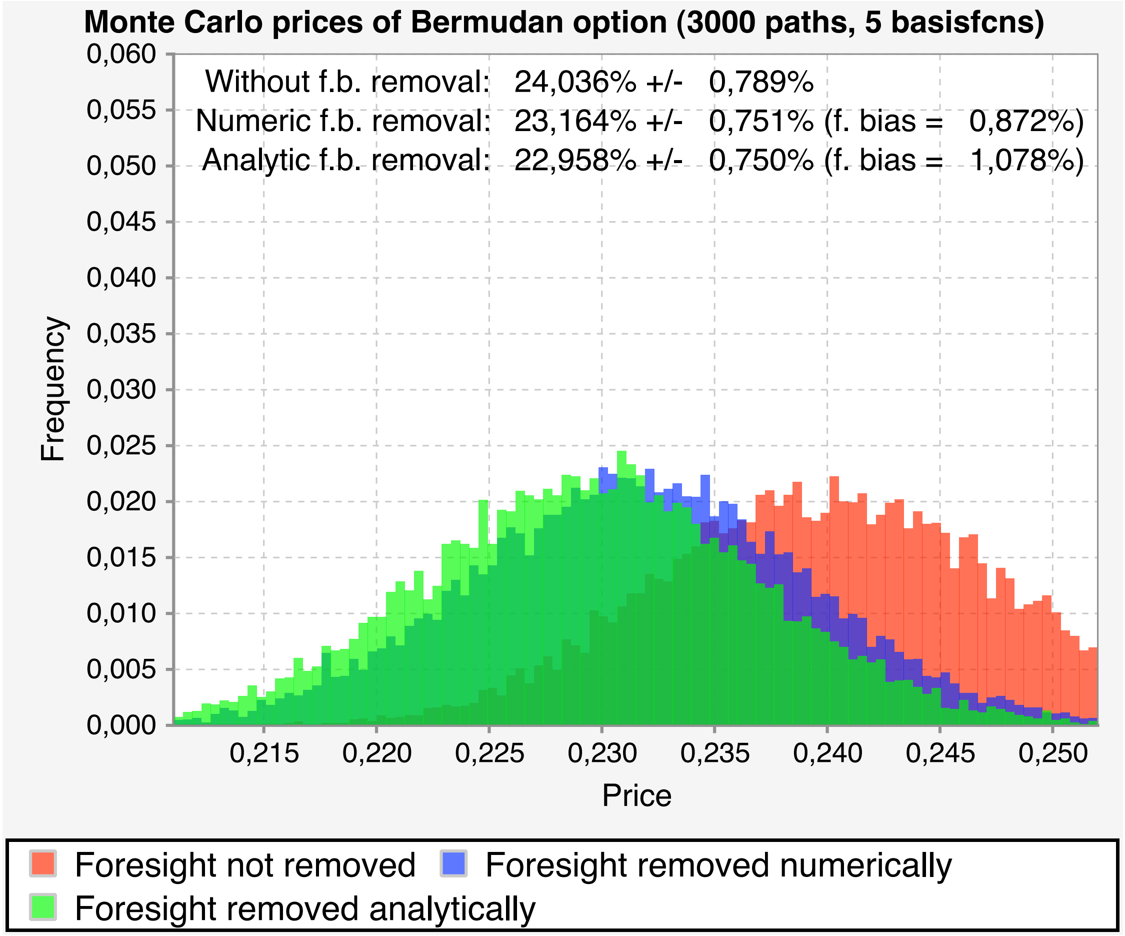
Modified backward algorithm induction step $i+1 \rightarrow i$ for $i = n, \ldots, 1$:

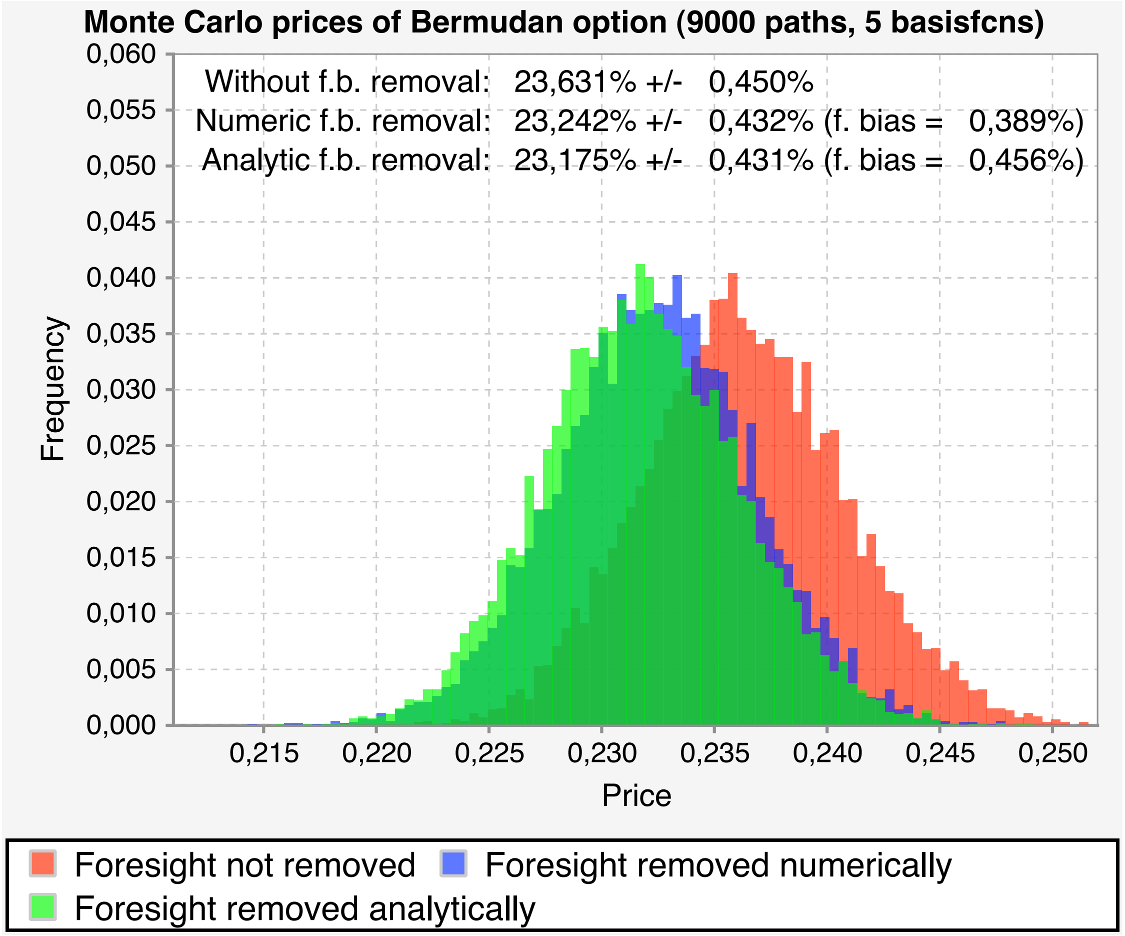
$$\tilde{V}_i := -\beta^{\mathrm{est}} - \gamma^{\mathrm{est}} + \begin{cases} \tilde{V}_{i+1} & \text{if } \tilde{U}(T_i) < \mathrm{E}^{\mathbb{Q}}(\tilde{V}_{i+1} | \mathcal{F}_{T_i}) \\ \tilde{U}(T_i) & \text{else.} \end{cases}$$

Numerical Results

Numerical Result: Parallelization of Monte-Carlo Simulation

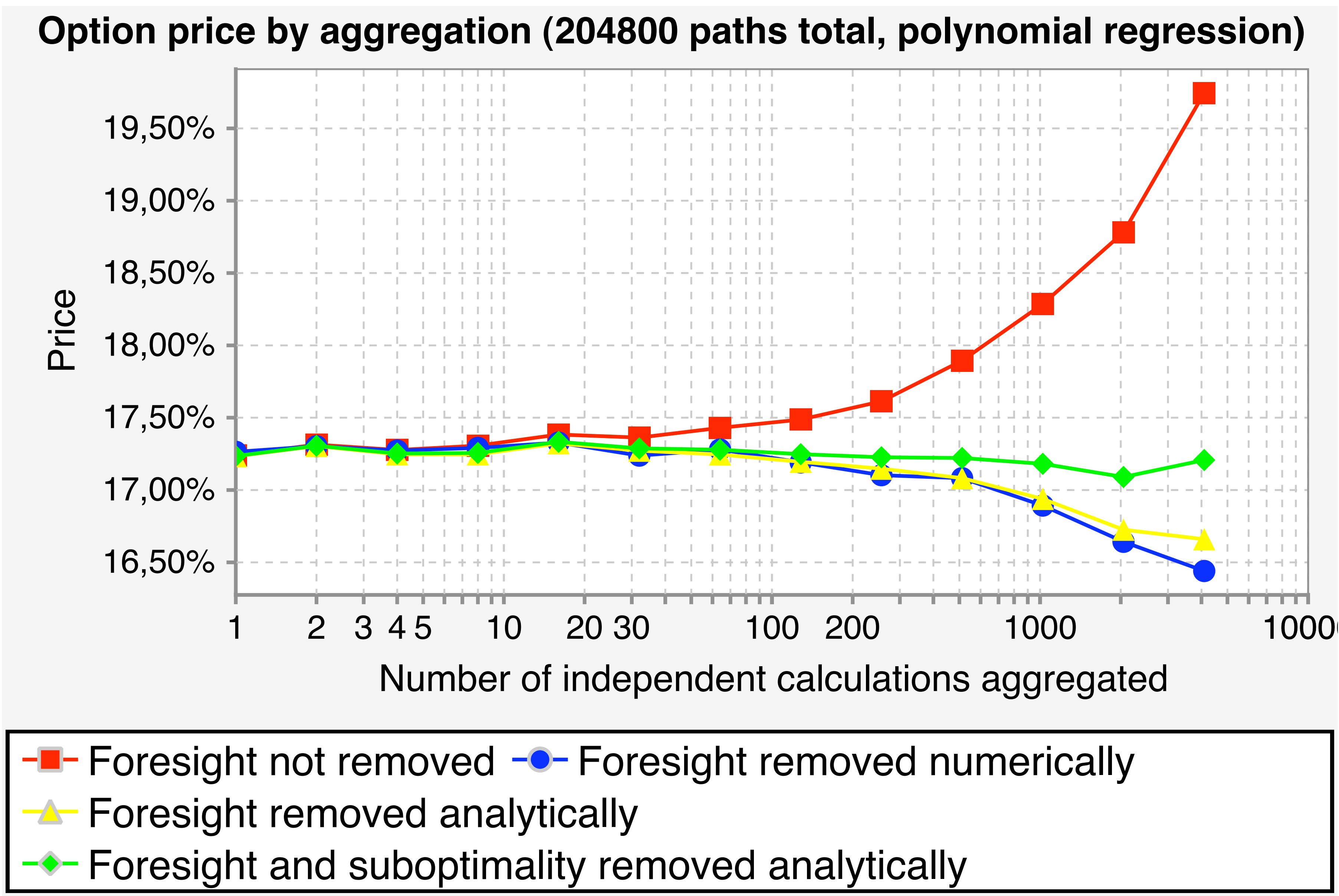
Repeated pricing with independent Monte-Carlo simulations (different random number seed) shows the distribution of the Monte-Carlo error.
The foresight bias is a systematic error, it corresponds to a shift of the mean.





Numerical Result: Parallelization of Monte-Carlo Simulation

We compare the aggregation k independent Monte-Carlo simulations with n/k paths for $k = 1, 2, 4, 8, 16, \ldots$. The foresight biased price grows with k. If foresight is removed, the sub-optimality biased price decays with k. With analytic foresight bias and sub-optimality correction the price is almost independent of k.



References

References

[An99] ANDERSEN, LEIF: A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor market model. Working paper. General Re Financial Products, 1999.

[BG97] BROADIE, MARK; GLASSERMAN, PAUL: Pricing American-Style Securities by Simulation. *J. Econom. Dynam. Control*, 1997, Vol. 21, 1323-1352.

[Ca96] CARRIERE, JACQUES F.: Valuation of Early-Exercise Price of Options Using Simulations and Non-parametric Regression. *Insurance: Mathematics and Economics* 19, 19-30, 1996.

[CLP01] CLÉMENT, EMMANUELLE; LAMBERTON, DAMIEN; PROTTER, PHILIP: An analysis of a least squares regression method for American option pricing. *Finance and Stochastics* 6, 449-471, 2002.

[DK94] DAVIS, MARK; KARATZAS, IOANNIS: A Deterministic Approach to Optimal Stopping, with Applications. In: Whittle, Peter (Ed.): Probability, Statistics and Optimization: A Tribute to Peter Whittle, 455-466, 1994. John Wiley & Sons, New York and Chichester, 1994. http://www.math.columbia.edu/~ik/DavisKaratzas.pdf

[F05] FRIES, CHRISTIAN P.: Foresight Bias and Sub-optimality Correction in Monte-Carlo Pricing of Options with Early Exercise: Classification, Calculation and Removal. (2005). http://www.christian-fries.de/finmath/foresightbias.

[F06] FRIES, CHRISTIAN P.: Mathematical Finance: Theory, Modeling, Implementation (lecture notes). http://www.christian-fries.de/finmath/book/index_en.html.

[FK05] FRIES, CHRISTIAN P.; KAMPEN, JÖRG: Proxy Simulation Schemes for generic robust Monte-Carlo sensitivities and high accuracy drift approximation (with applications to the LIBOR Market Model). 2005. http://www.christian-fries.de/finmath/proxyscheme

[Gl03] GLASSERMAN, PAUL: Monte Carlo Methods in Financial Engineering. 596 Pages. Springer, 2003. ISBN 0-387-00451-3.

[LS01] LONGSTAFF, FRANCIS A.; SCHWARTZ EDUARDO S.: Valuing American Options by Simulation: A Simple Least-Square Approach. *Review of Financial Studies* 14:1, 113-147, 2001.

References

[Pi03] PITERBARG, VLADIMIR V.: A practitioner's guide to pricing and hedging callable LIBOR exotics in forward LIBOR models, *Preprint*. 2003.

[Ro01] ROGERS, L. C. G.: Monte Carlo valuation of American options, *Preprint*. 2001.

please check www.christian-fries.de/finmath for updates