Moment-SOS hierarchy and exit time of stochastic processes

Didier Henrion, Mauricio Junca, Mauricio Velasco

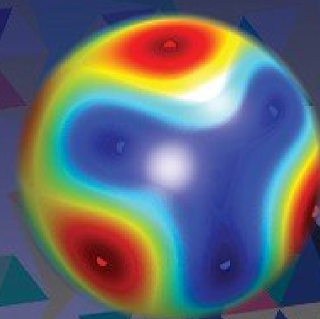
arXiv:2101.06009, hal:03110452

2 February 2021

Series on Optimization and its Applications – Vol. 4

The Moment-SOS Hierarchy

Lectures in Probability, Statistics, Computational Geometry, Control and Nonlinear PDEs



Didier Henrion
Milan Korda
Jean B. Lasserre

World Scientific

Main steps of the moment-SOS aka Lasserre hierarchy

Given a **nonlinear nonconvex** problem:

1. Reformulate it as a **linear** problem (at the price of enlarging or changing the space of solutions);

2. Solve approximately the linear problem with a hierarchy of tractable **convex** relaxations (of increasing size);

3. Ensure **convergence**: either the original problem is solved at a finite relaxation size, or its solution is approximated with increasing quality.

At each step, **conic duality** is an essential ingredient.

In this talk we follow this program for evaluating functionals of nonlinear stochastic differential equations

The main contribution is the elementary proof of convergence using elementary results from PDE analysis

Stochastic process

Stochastic differential equation (SDE)

$$d\mathbf{X} = \mathbf{b}(\mathbf{X})dt + \mathbf{B}(\mathbf{X})d\mathbf{W}, \quad \mathbf{X}(0) = x$$

driven by m-dimensional Brownian motion $\mathbf{W}$ and initialized at x in a given open set $\mathscr{X}$ of $\mathbb{R}^n$ with smooth boundary

Drift $\mathbf{b} = (b_i)_i$ and diffusion $\mathbf{B} = (b_{ij})_{i,j}$ coefficients are given

Assume $\mathbf{B}$ has full rank, so that matrix $\mathbf{A} = (a_{ij} := \frac{1}{2}\sum_k b_{ik}b_{jk})_{i,j}$ is positive definite

Assume $\mathbf{b}$ and $\mathbf{B}$ are continuous and growing at most linearly outside of $\mathscr{X}$, so that there is a unique solution to the SDE, the **stochastic process** $\mathbf{X}(t)$

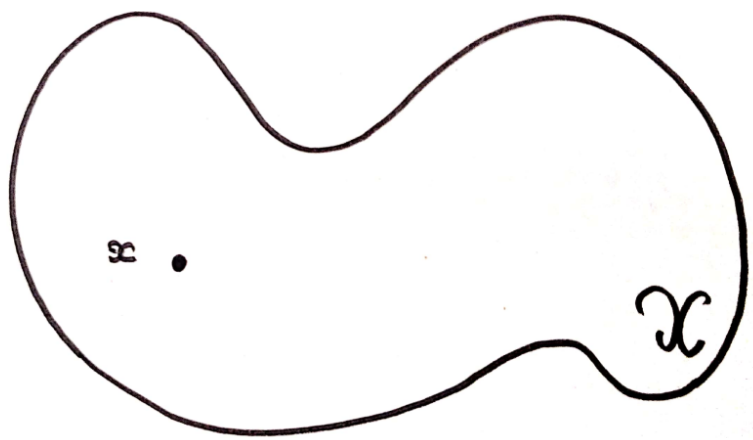
Exit time problem

Let $g : \partial \mathscr{X} \to \mathbb{R}$ be a given continuous function

We want to evaluate the function

$$v^*(x) := E[g(\mathbf{X}(\tau_x))]$$

where τ_x is the first time $\mathbf{X}(.)$ hits $\partial \mathscr{X}$



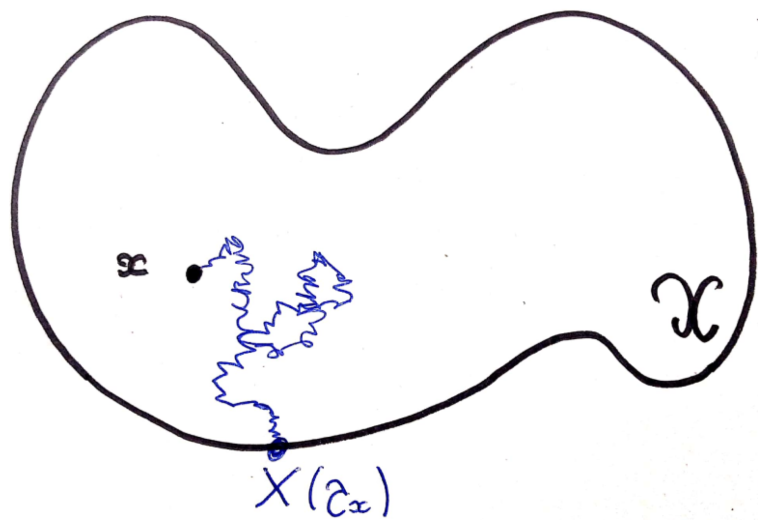
Exit time problem

Let $g : \partial\mathscr{X} \to \mathbb{R}$ be a given continuous function

We want to evaluate the function

$$v^*(x) := E[g(\mathbf{X}(\tau_x))]$$

where τ_x is the first time $\mathbf{X}(.)$ hits $\partial\mathscr{X}$



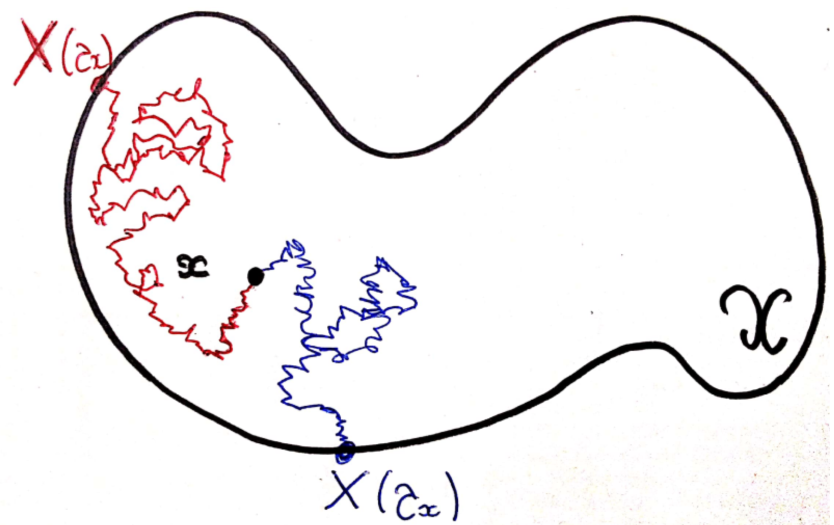
Exit time problem

Let $g : \partial\mathscr{X} \to \mathbb{R}$ be a given continuous function

We want to evaluate the function

$$v^*(x) := E[g(\mathbf{X}(\tau_x))]$$

where τ_x is the first time $\mathbf{X}(.)$ hits $\partial\mathscr{X}$



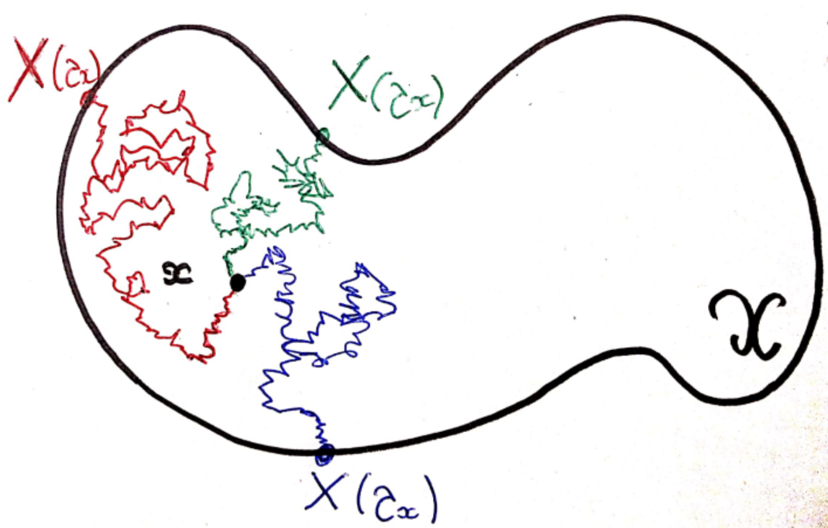
Exit time problem

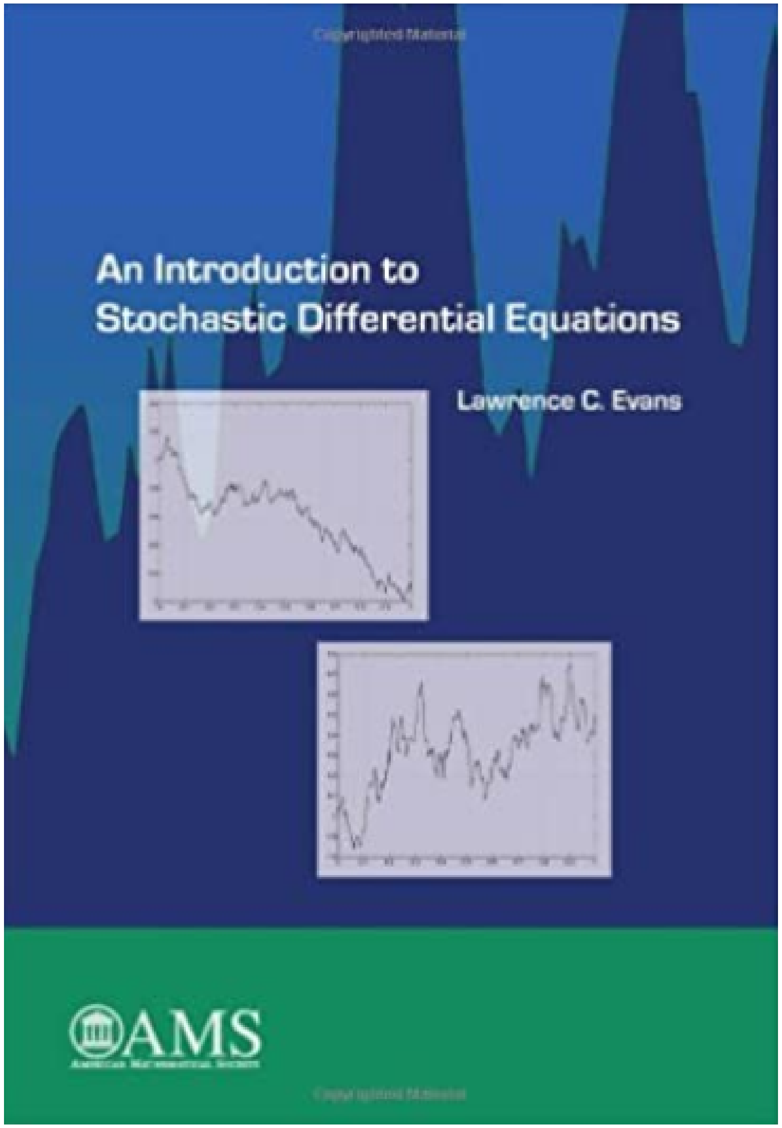
Let $g : \partial\mathscr{X} \to \mathbb{R}$ be a given continuous function

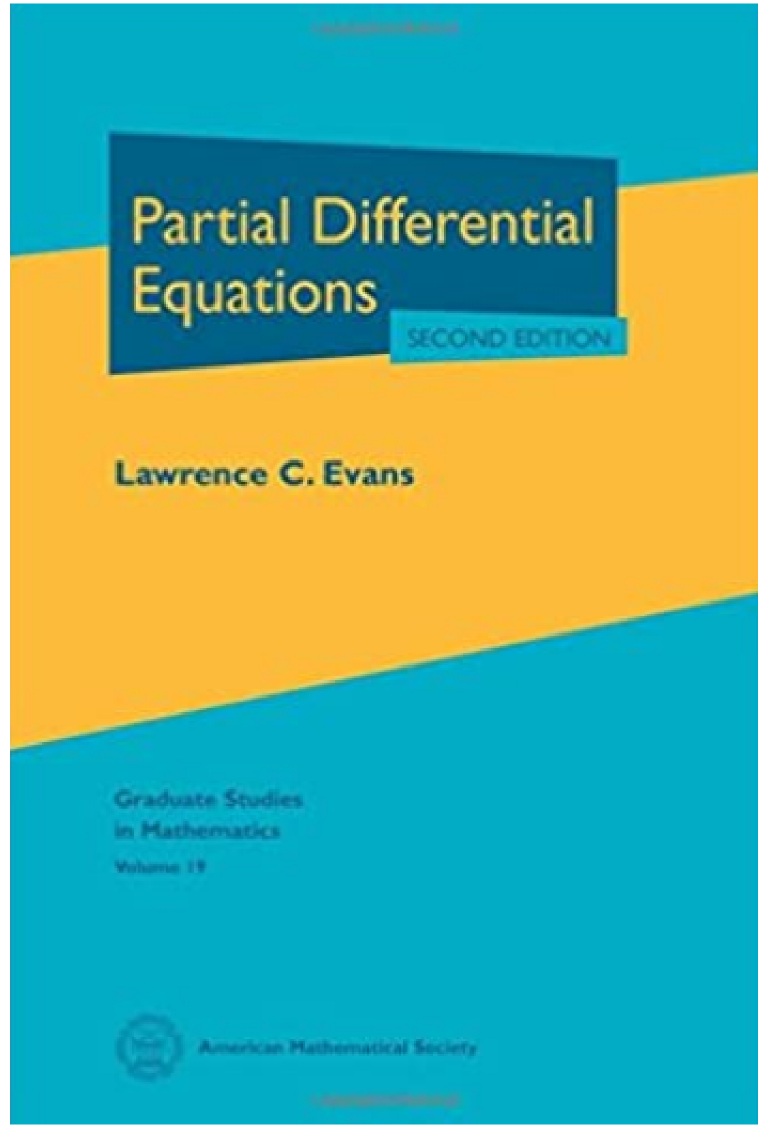
We want to evaluate the function

$$v^*(x) := E[g(\mathbf{X}(\tau_x))]$$

where τ_x is the first time $\mathbf{X}(.)$ hits $\partial\mathscr{X}$




An Introduction to
Stochastic Differential Equations
Lawrence C. Evans
AMS


Partial Differential
Equations
SECOND EDITION
Lawrence C. Evans
Graduate Studies
in Mathematics
Volume 19
American Mathematical Society

Stochastic calculus

Generator of stochastic process

$$-Lf := \sum_{i,j} a_{ij}\partial_i\partial_j f + \sum_i b_i \partial_i f$$

Since $\mathbf{A}$ is positive definite, linear operator L is **uniformly elliptic**

Given $f \in C^2(\overline{\mathscr{X}})$, Itô's chain rule implies that

$$f(\mathbf{X}(\tau_x)) = f(\mathbf{X}(0)) - \int_0^{\tau_x} Lf\,ds + \int_0^{\tau_x} \sum_k \sum_i \partial_i f b_{ik}\, dW_k$$

and taking the expected value yields Dynkin's formula

$$E[f(\mathbf{X}(\tau_x))] = \underbrace{E[f(\mathbf{X}(0))]}_{f(x)} - E\left[\int_0^{\tau_x} Lf\,ds\right]$$

Linear reformulation

Dynkin's formula

$$E[f(\mathbf{X}(\tau_x))] + E\left[\int_0^{\tau_x} Lf ds\right] = f(x)$$

then becomes

$$\langle f, \nu \rangle + \langle Lf, \mu \rangle = f(x)$$

upon defining the **exit location measure** ν on $\partial\mathscr{X}$ and the **expected occupation measure** μ on $\mathscr{X}$ so that

$$\langle f, \nu \rangle := E\left[f(\mathbf{X}(\tau_x))\right], \quad \langle f, \mu \rangle := E\left[\int_0^{\tau_x} f(\mathbf{X}(s))ds\right]$$

for every test function f

For L' adjoint to L, in the sense of distributions it holds

$$\nu + L'\mu = \delta_x$$

Lower and upper bounds

Recall that we want to evaluate the function

$$v^*(x) := E[g(\mathbf{X}(\tau_x))] = \langle g, \nu \rangle$$

so consider the LP

$$v_{\text{min}}(x) := \min_{\mu,\nu} \langle g, \nu \rangle \text{ s.t. } \nu + L'\mu = \delta_x$$

and

$$v_{\text{max}}(x) := \max_{\mu,\nu} \langle g, \nu \rangle \text{ s.t. } \nu + L'\mu = \delta_x$$

which satisfy by construction

$$v_{\text{min}}(x) \leq v^*(x) \leq v_{\text{max}}(x)$$

for each $x \in \mathscr{X}$

No relaxation gap

Theorem: $v_{\text{min}}(x) = v^*(x) = v_{\text{max}}(x)$ for each $x \in \mathscr{X}$

Since the criterion $\langle g, \nu \rangle$ depends only on ν, the proof consists of proving that there is a unique ν solving $\langle f, \nu \rangle + \langle Lf, \mu \rangle = f(x)$

Given any $p \in C(\partial\mathscr{X})$, let $f_p \in C^2(\mathscr{X}) \cup C(\partial\mathscr{X})$ be the **unique** solution to the **elliptic** linear PDE

$$\begin{aligned} Lf &= 0 \quad \text{in } \mathscr{X} \\ f &= p \quad \text{on } \partial\mathscr{X} \end{aligned}$$

Plugging f_p into $\langle f, \nu \rangle + \langle Lf, \mu \rangle = f(x)$ yields $\langle p, \nu \rangle = f_p(x)$

Since $\partial\mathscr{X}$ is compact, the space $C(\partial\mathscr{X})$ is separable and by choosing countably many functions $p \in C(\partial\mathscr{X})$ we can generate countably many linear relations that uniquely specify ν, QED

Duality

Dual to the LP

$$\min_{\mu,\nu} \langle g, \nu \rangle \text{ s.t. } \nu + L'\mu = \delta_x$$

is the LP

$$\max_{v} v(x) \text{ s.t. } Lv \leq 0 \text{ in } \mathscr{X}, \; v \leq g \text{ on } \partial\mathscr{X}$$

where the maximization is with respect to functions $v \in C^2(\mathscr{X})$

There is **no duality gap**, i.e. both LP have the same value

Subsolutions

It is well-known that the value function v^* solves the elliptic PDE

$$\begin{aligned} Lv = 0 & \quad \text{in } \mathscr{X} \\ v = g & \quad \text{on } \partial\mathscr{X} \end{aligned}$$

Any admissible function v for the dual LP

$$\max_v v(x) \text{ s.t. } Lv \le 0 \text{ in } \mathscr{X}, \; v \le g \text{ on } \partial\mathscr{X}$$

is a **subsolution** of the elliptic PDE, i.e. $v^* \ge v$ on $\overline{\mathscr{X}}$

The LP selects a subsolution that touches v^* from below at x

Random initial condition

All the above developments generalize readily to the case that the initial condition $\mathbf{X}(0)$ in the SDE is a random variable whose law is a given probability measure ξ on $\mathscr{X}$

The previous results can then be retrieved with the particular choice $\xi = \delta_x$ for a given $x \in \mathscr{X}$

The quantity to be evaluated is averaged

$$\int_{\mathscr{X}} E[g(\mathbf{X}(\tau_x))] d\xi(x)$$

and the linear PDE becomes

$$\nu + L'\mu = \xi$$

Scalar example

Let us apply the moment-SOS hierarchy to the LPs for an elementary exit time problem

The SDE is $d\mathbf{X}_t = (1 + 2\mathbf{X}_t)dt + \sqrt{2}\mathbf{X}_t d\mathbf{W}_t$ on the domain $\mathscr{X} := (0, 1)$ with initial condition $x = 1/2$

This process always exits at $\{1\}$, so $\nu = \delta_1$ and $v^*(x) = g(1)$

For the choice $g(z) = z^2$ we report the values of the lower and upper bounds obtained by solving the moment relaxations for increasing relaxation degrees

Scalar example - bounds for increasing relaxation degrees

relaxation degree	2	4	6	8	10
lower bound	0.65000	0.92157	0.98118	0.99503	0.99827
upper bound	1.00000	1.00000	1.00000	1.00000	1.00000

The degree 10 relaxation was solved with MOSEK in 0.15s

The corresponding GloptiPoly script is given next

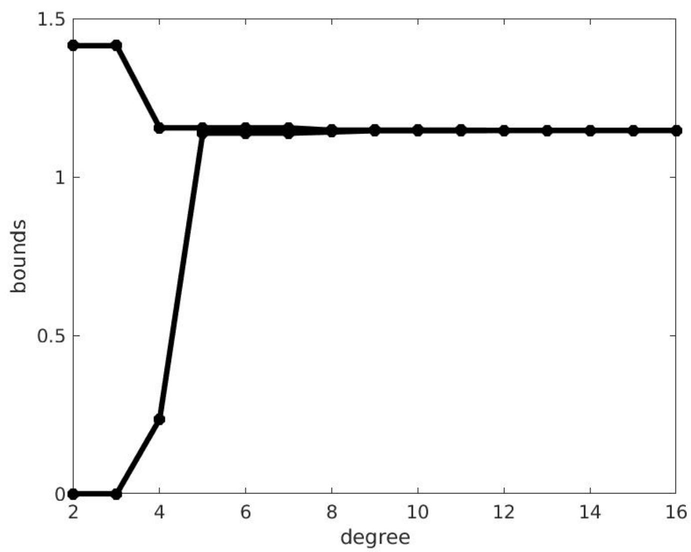
```
dmax = 10; % relaxation degree
mpol xmu xnu
mu = meas(xmu); % expected occupation measure
nu = meas(xnu); % exit location measure

x0 = 0.5; % initial condition
momeqs = []; % linear moment equations
for d = 0:dmax
Lfmu = 0;
if d > 0, Lfmu = Lfmu - mom((1+2*xmu)*(d*xmu^(d-1))); end
if d > 1, Lfmu = Lfmu - mom(xmu^2*(d*(d-1)*xmu^(d-2))); end
if d > 0, fnu = mom(xnu^d); else fnu = mass(nu); end
momeqs = [momeqs; fnu+Lfmu == x0^d];
end
```

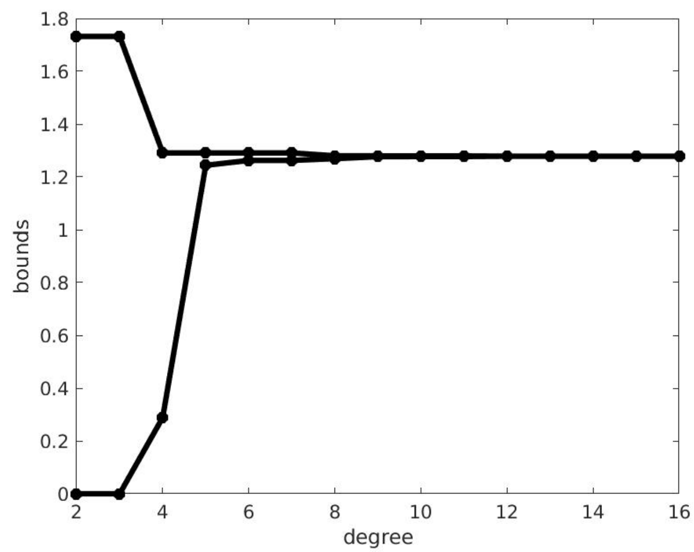
```
g = xnu^2; % functional

% construct moment relaxation
P = msdp(min(g), momeqs, xmu*(1-xmu)>=0, xnu*(1-xnu)==0);

% solve SDP problem
msol(P);

% bound
double(g)

% approximate mass of the occupation measure
double(mass(mu))
```

Multivariate example

Now let $g(z) := \sum_{k=1}^{n} z_k^2$ for the n-dim. Brownian $\mathbf{X}_t = \mathbf{W}_t$ in the convex semi-algebraic domain $\mathscr{X} := \{z \in \mathbb{R}^n : \sum_{k=1}^{n} z_k^4 \leq 1\}$ with initial condition $x = 0$

Here are the bounds for $n = 2$ (left) and $n = 3$ (right)





Here we report the number of moments and the computational time required to solve the moment relaxation of degree 8, for increasing values of the dimension n

For this relaxation degree the gap between the lower and upper bounds on the functional is less than 2%

dimension n	2	3	4	5	6	7	8
number of moments	73	249	705	1749	3927	8151	15873
time (seconds)	0.15	0.59	1.9	5.6	19	51	195