

# Optimal control under probability constraint

P. CARPENTIER — *ENSTA ParisTech*

J.-P. CHANCELIER — *École des Ponts ParisTech*

G. COHEN — *École des Ponts ParisTech*



This work has been supported by  
**Thalès-Aléniya-Space** (Cannes) and **CNES** (Toulouse)

25<sup>th</sup> IFIP TC7 Conference

September 14, 2011

# Presentation outline

- 1 Problem formulation
- 2 Modeling improvement
- 3 Stochastic Arrow-Hurwicz algorithm
- 4 Numerical results

- 1 Problem formulation
  - Satellite model and deterministic optimization problem
  - Engine failure
  - Stochastic formulation
- 2 Modeling improvement
- 3 Stochastic Arrow-Hurwicz algorithm
- 4 Numerical results

## Satellite model

$$\frac{dr}{dt} = v, \quad \frac{dv}{dt} = -\mu \frac{r}{\|r\|^3} + \frac{F}{m} \kappa, \quad (1a)$$

$$\frac{dm}{dt} = -\frac{T}{g_0 I_{sp}} \delta. \quad (1b)$$

(1a) : 6-dimensional state vector (position  $r$  and velocity  $v$ ).

(1b) : 1-dimensional state vector (mass  $m$  including fuel).

$\kappa$  involves the direction cosines of the thrust and the on-off switch  $\delta$  of the engine (3 controls), and  $\mu, F, T, g_0, I_{sp}$  are constants.

The deterministic control problem is to drive the satellite from the initial condition at  $t_i$  to a known **final position  $r_f$  and velocity  $v_f$**  at  $t_f$  (given) while **minimizing fuel consumption  $m(t_i) - m(t_f)$** .

## Deterministic optimization problem

Using **equinoctial coordinates** for the position and velocity

$\leadsto$  **state vector**  $x \in \mathbb{R}^7$ ,

and **cartesian coordinates** for the thrust of the engine

$\leadsto$  **control vector**  $u \in \mathbb{R}^3$ ,

the deterministic optimization problem is written as follows:



## Engine failure

- Sometimes, the engine may **fail to work** when needed: the satellite **drift away** from the deterministic optimal trajectory. After the engine control is recovered, it is not always possible to **drive the satellite to the final target** at  $t_f$ .
- By **anticipating** such possible failures and by **modifying** the trajectory followed **before** any such failure occurs, one may **increase** the possibility of eventually reaching the target. But such a deviation from the deterministic optimal trajectory results in a **deterioration of the economic performance**.
- The problem is thus to **balance** the **increased probability** of eventually reaching the target despite possible failures against the **expected economic performance**, that is, to **quantify** the price of safety one is ready to pay for.

## Stochastic formulation (1)

A failure is modeled using two random variables:

- $t_p$  : random initial time of the failure,
- $t_d$  : random duration of the failure.

For every realization  $(t_p^\xi, t_d^\xi)$ :

- 1  $u(\cdot)$  denotes the control used prior to any failure  
 $\rightsquigarrow u$  is defined over  $[t_i, t_f]$  but implemented over  $[t_i, t_p^\xi]$   
 and corresponds to an **open-loop control**,
- 2 the control is 0 in  $[t_p^\xi, t_p^\xi + t_d^\xi]$ ,
- 3  $v^\xi(\cdot)$  denotes the control used after the end of the failure  
 $\rightsquigarrow v^\xi$  is defined over  $[t_p^\xi + t_d^\xi, t_f]$  (if nonempty)  
 and corresponds to a **closed-loop strategy v**.

The **satellite dynamics** in the stochastic formulation writes:

$$x^\xi(t_i) = x_i, \quad \dot{x}^\xi(t) = f^\xi(x^\xi(t), u(t), v^\xi(t)).$$

## Stochastic formulation (2)

The problem is to minimize the **expected cost** (fuel consumption)

- w.r.t. the open-loop control  $u$  and the closed-loop strategy  $\mathbf{v}$ ,
- the **probability to hit the target** at time  $t_f$  being at least  $p$ .

$$\min_{u(\cdot)} \min_{\mathbf{v}(\cdot)} \mathbb{E} \left( K(x^\xi(t_f)) \right) \quad (3a)$$

$$\min_{u(\cdot)} \min_{\mathbf{v}(\cdot)} \mathbb{E} \left( K(x^\xi(t_f)) \mid C(x^\xi(t_f)) = 0 \right) \quad (3b)$$

subject to:

$$x^\xi(t_i) = x_i, \quad \dot{x}^\xi(t) = f^\xi(x^\xi(t), u(t), \mathbf{v}^\xi(t)), \quad (3c)$$

$$\|u(t)\| \leq 1 \quad \forall t \in [t_i, t_f], \quad \|\mathbf{v}^\xi(t)\| \leq 1 \quad \forall t \in [t_p^\xi + t_d^\xi, t_f], \quad (3d)$$

$$\mathbb{P} \left( C(x^\xi(t_f)) = 0 \right) \geq p. \quad (3e)$$



- 1 Problem formulation
- 2 Modeling improvement**
  - Handling of probability and conditional expectation
  - Dealing with the ratio of expectations
- 3 Stochastic Arrow-Hurwicz algorithm
- 4 Numerical results

## Indicator function

Consider the real-valued **indicator function**  $I(y) = \begin{cases} 1 & \text{if } y = 0, \\ 0 & \text{otherwise.} \end{cases}$

Then

$$\mathbb{P}\left(C(x^\xi(t_f)) = 0\right) = \mathbb{E}\left(I(\|C(x^\xi(t_f))\|)\right),$$

and

$$\mathbb{E}\left(K(x^\xi(t_f)) \mid C(x^\xi(t_f)) = 0\right) = \frac{\mathbb{E}\left(K(x^\xi(t_f)) \times I(\|C(x^\xi(t_f))\|)\right)}{\mathbb{E}\left(I(\|C(x^\xi(t_f))\|)\right)}.$$

## Problem reformulation

The problem is (shortly) reformulated as

$$\min_{u(\cdot)} \min_{\mathbf{v}(\cdot)} \frac{\mathbb{E} \left( K(x^\xi(t_f)) \times \mathbb{I}(\|C(x^\xi(t_f))\|) \right)}{\mathbb{E} \left( \mathbb{I}(\|C(x^\xi(t_f))\|) \right)} \quad (4a)$$

$$\text{s.t. } \mathbb{E} \left( \mathbb{I}(\|C(x^\xi(t_f))\|) \right) \geq p. \quad (4b)$$

Such a formulation is however not well-suited for numerical implementation (e.g. Arrow-Hurwicz algorithm):

**a ratio of expectations is not an expectation!**

## An useful lemma

Using compact notation, Problem (4) is:

$$\min_{\mathbf{u}} \frac{J(\mathbf{u})}{\Theta(\mathbf{u})} \quad \text{s.t.} \quad \Theta(\mathbf{u}) \geq p, \quad (5)$$

in which  $J$  and  $\Theta$  assume positive values.

- ① If  $\mathbf{u}^\#$  is a solution of (5) and if  $\Theta(\mathbf{u}^\#) = p$ , then  $\mathbf{u}^\#$  is also a solution of

$$\min_{\mathbf{u}} J(\mathbf{u}) \quad \text{s.t.} \quad \Theta(\mathbf{u}) \geq p. \quad (6)$$

- ② Conversely, if  $\mathbf{u}^\#$  is a solution of (6), and if an optimal Kuhn-Tucker multiplier  $\beta^\#$  satisfies the condition

$$\beta^\# \geq \frac{J(\mathbf{u}^\#)}{\Theta(\mathbf{u}^\#)},$$

then  $\mathbf{u}^\#$  is also a solution of (5).

## Back to a cost in expectation

Finally, instead of (4) we aim at solving a problem in which the **cost and constraint** functions correspond to **expectations**.

$$\begin{aligned} & \min_{u(\cdot)} \min_{v(\cdot)} \mathbb{E} \left( K(x^\xi(t_f)) \times I(\|C(x^\xi(t_f))\|) \right) \\ \text{s.t.} \quad & \mathbb{E} \left( I(\|C(x^\xi(t_f))\|) \right) \geq p . \end{aligned}$$

or equivalently

$$\min_{u(\cdot)} \mathbb{E} \left( \min_{v^\xi(\cdot)} K(x^\xi(t_f)) \times I(\|C(x^\xi(t_f))\|) \right) \quad (7a)$$

$$\text{s.t.} \quad \mathbb{E} \left( I(\|C(x^\xi(t_f))\|) \right) \geq p . \quad (7b)$$

- 1 Problem formulation
- 2 Modeling improvement
- 3 Stochastic Arrow-Hurwicz algorithm
  - Algorithm overview
  - Smoothing function I
  - Solving the resulting approximated problem
  - Numerical implementation
- 4 Numerical results

# Lagrangian formulation

$$\begin{aligned} & \min_{u(\cdot)} \mathbb{E} \left( \min_{v^\xi(\cdot)} K(x^\xi(t_f)) \times I(\|C(x^\xi(t_f))\|) \right) \\ \text{s.t. } & p - \mathbb{E} \left( I(\|C(x^\xi(t_f))\|) \right) \leq 0 \quad \longleftrightarrow \quad \mu \end{aligned}$$

Assume there exists a saddle point for the associated Lagrangian.  
 In order to solve

$$\max_{\mu \geq 0} \min_{u(\cdot)} \left\{ \underbrace{\mu p + \mathbb{E} \left( \min_{v^\xi(\cdot)} (K(x^\xi(t_f)) - \mu) \times I(\|C(x^\xi(t_f))\|) \right)}_{W(u, \mu, \xi)} \right\}.$$

that is,

$$\max_{\mu \geq 0} \min_{u(\cdot)} \left\{ \mu p + \mathbb{E} (W(u, \mu, \xi)) \right\},$$

we use the **stochastic Arrow-Hurwicz algorithm** (see [2]–[3]).

# Algorithm overview

## Arrow-Hurwicz algorithm

At iteration  $k$ ,

- ① draw a failure  $\xi^k = (t_p^{\xi^k}, t_d^{\xi^k})$  according to its probability law,
- ② compute the gradient of  $W$  w.r.t.  $u$  and update  $u(\cdot)$ :

$$u^{k+1} = \Pi_{\mathcal{B}} \left( u^k - \varepsilon^k \nabla_u W(u^k, \mu^k, \xi^k) \right),$$

- ③ compute the gradient of  $W$  w.r.t.  $\mu$  and update  $\mu$ :

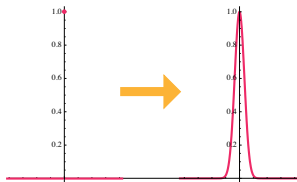
$$\mu^{k+1} = \max \left( 0, \mu^k + \rho^k (\rho + \nabla_\mu W(u^{k+1}, \mu^k, \xi^k)) \right).$$



## First difficulty: $I$ is not a smooth function

At every iteration  $k$ , we must evaluate function  $W$  as well as **its derivatives** w.r.t.  $u(\cdot)$  and  $\mu$ . **But  $W$  is not differentiable!**

$$I(y) = \begin{cases} 1 & \text{if } y = 0, \\ 0 & \text{otherwise,} \end{cases} \rightsquigarrow I_r(y) = \begin{cases} \left(1 - \frac{y^2}{r^2}\right)^2 & \text{if } y \in [-r, r], \\ 0 & \text{otherwise.} \end{cases}$$



There are specific rules to drive  $r$  to 0 as the iteration number  $k$  goes to infinity in order to obtain the best asymptotic **Mean Quadratic Error** of the gradient estimates (see [1]).

## Second difficulty: solving the inner problem

The approximated closed-loop problem to solve at each iteration is:

$$W_r(u^k, \xi^k, \mu^k) = \min_{v^\xi(\cdot)} \left\{ (K(x^\xi(t_f)) - \mu^k) \times \mathbf{I}_{r^k}(\|C(x^\xi(t_f))\|) \right\}.$$

In this setting, we have to check if the target is reached **up to  $r^k$** .  
 Different cases have to be considered:

- ① the target can be reached accurately,
- ② the target can be reached up to  $r^k$  only,
- ③ the target cannot be reached up to  $r^k$ .

If reaching the target is **possible** but **too expensive** (that is  $K(x^\xi(t_f)) \geq \mu^k$ ), the best thing to do is to **stop immediately**.

In practice, the solution of the approximated problem is derived from the resolution of two standard optimal control problems. . .

## Parameters tuning

**Gradient step length:**

$$\varepsilon^k = \frac{a}{b+k}, \quad \rho^k = \frac{c}{d+k},$$

↪ usual for a stochastic gradient algorithm.

**Smoothing parameter:**

$$r^k = \frac{\alpha}{\beta + k^{\frac{1}{3}}},$$

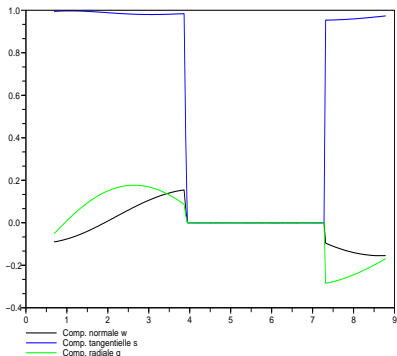
↪ MQE reduced by a factor 1000 in about 100.000 iterations.

- 1 Problem formulation
- 2 Modeling improvement
- 3 Stochastic Arrow-Hurwicz algorithm
- 4 Numerical results**
  - Brief description
  - Results for different values of  $p$
  - The price of safety

## Mission description

### Interplanetary mission (Earth-Mars trajectory):

- duration of the mission: 450 days,
- $t_p$ : exponential distribution s.t.  $\mathbb{P}(t_p \geq t_f) = \pi_f \approx 0.58$ ,
- $t_d$ : exponential distribution s.t.  $\mathbb{P}(2 \leq t_d \leq 7) \approx 0.80$ .



Using normalized units:

- $t_i = 0.69$  and  $t_f = 8.73$ .

The **deterministic optimal control** has a “bang–off–bang” shape. Along the **deterministic optimal path**, the probability to recover a failure is:

$$p^{\text{det}} \approx 0.94.$$

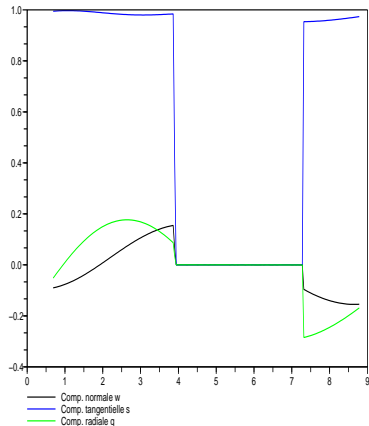
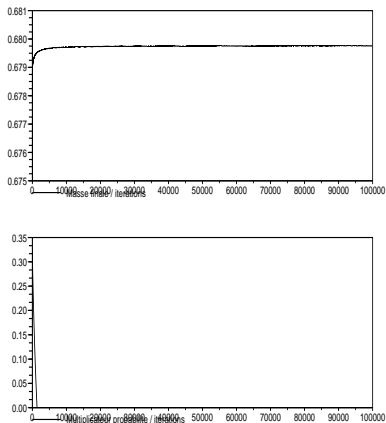


Figure: Probability level  $p < \pi_f$

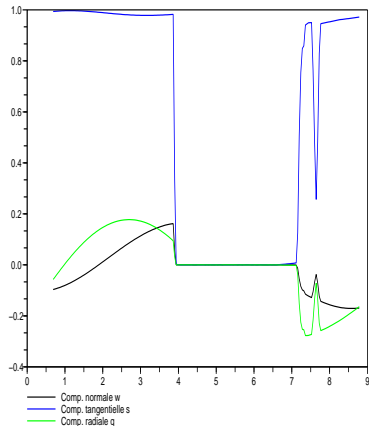
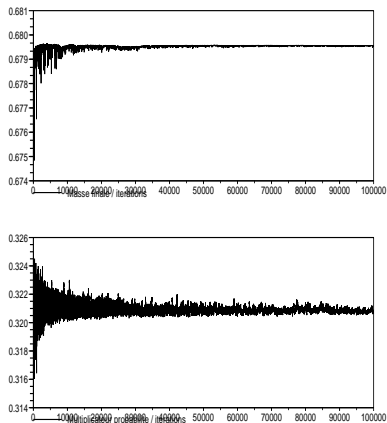


Figure: Probability level  $\rho = 0.750$

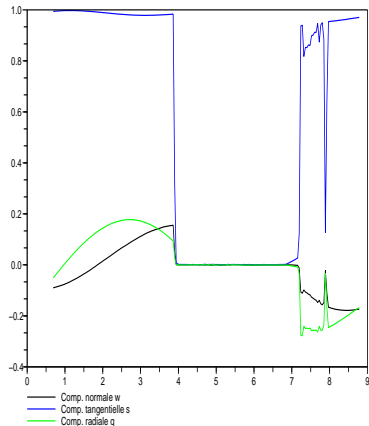


Figure: Probability level  $\rho = 0.960$



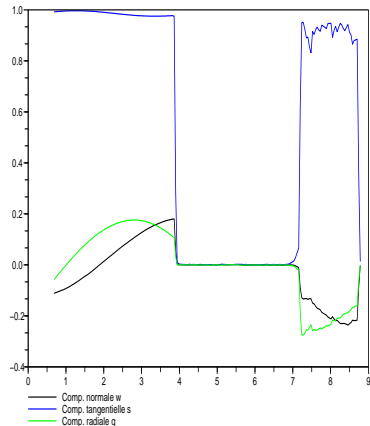
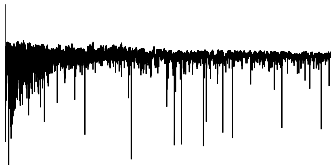
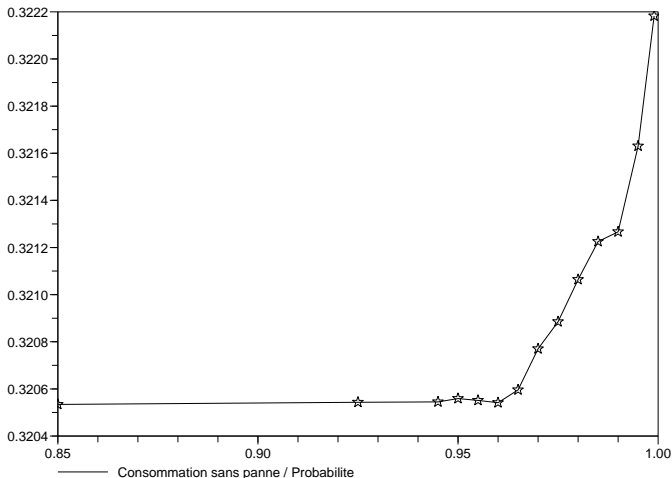


Figure: Probability level  $\rho = 0.990$

# Fuel consumption versus probability level



# Conclusion

## Main conclusion

We are able to deal with probability constraints in the optimal control framework.

## Future works

- *From the theoretical point of view:*
  - existence of a saddle point for the constrained problem,
  - smoothing process (results available only for inequality constraints).
- *From the numerical point of view:*
  - efficient solver for the downstream problem,
  - computer parallelization.



Laetitia Andrieu, Guy Cohen and Felisa J. Vázquez-Abad.

Gradient-based simulation optimization under probability constraints.

*European Journal of Operational Research*, 212, 345-351, 2011.



Jean-Christophe Culioli and Guy Cohen.

Decomposition/Coordination Algorithms in Stochastic Optimization.

*SIAM Journal on Control and Optimization*, 28, 1372-1403, 1990.



Jean-Christophe Culioli and Guy Cohen.

Optimisation stochastique sous contraintes en espérance.

*CRAS*, t.320, Série I, pp. 75-758, 1995.



Thierry Dargent.

Transfert d'orbite optimal par application du principe du minimum.

*Logiciel T\_3D*, version 2, 19 mars 2004.



András Prékopa.

*Stochastic Programming*.

Kluwer, Dordrecht, 1995.