Taking uncertainties into account in numerical shape optimization.

Marc Dambrine

joint works with Charles Dapogny (Grenoble), Helmut Harbrecht, Michael Peters (Basel) and Bénédicte Puig (Pau)

November 25, 2016

Why should we consider uncertainties in shape optimization ?

- Mechanical systems rely on data, e.g. the loads, the properties of a constituent material, or the geometry of the system itself.

- In concrete situations, such data are plagued with uncertainties because:

    - they may be available only through (error-prone) measurements,
    - they may be altered with time (wear) and conditions of the ambient medium.

- The performances of structures are very sensitive to small perturbations of data.

Need to somehow anticipate uncertainties when designing and optimizing shape

Modelisation choice : uncertainties are seen as a random process.

$\rightsquigarrow$ we face a familly of shape functionals $(J(\omega, D))_{\omega \in \Omega}$ for domains D in an admissible class $\mathcal{A}$.

How to take these uncertainties into account in a viable numerical optimization strategy ?

Modelisation choice : uncertainties are seen as a random process.

How to take these uncertainties into account in a viable numerical optimization strategy ?

A priori two strategies can be explored in order to compute such a shape:

1. *solve* each problem

$$D^*(\omega) = \operatorname{Argmin}_{D \in \mathcal{A}} J(\omega, D)$$

   and take the *expectation* of D^*

Modelisation choice : uncertainties are seen as a random process.

How to take these uncertainties into account in a viable numerical optimization strategy ?

A priori two strategies can be explored in order to compute such a shape:

1. *solve* each problem

$$D^*(\omega) = \operatorname{Argmin}_{D\in\mathcal{A}} J(\omega, D)$$

   and take the *expectation* of D^*
   two difficulties : how to define an average shape ? how to compute it ?

Modelisation choice : uncertainties are seen as a random process.

How to take these uncertainties into account in a viable numerical optimization strategy ?

A priori two strategies can be explored in order to compute such a shape:

1. *solve* each problem

$$D^*(\omega) = \mathrm{Argmin}_{D \in \mathcal{A}} \, J(\omega, D)$$

   and take the *expectation* of D^*
   two difficulties : how to define an average shape ? how to compute it ?

2. consider an averaged objective

$$\mathbb{J}_\sigma(D) = \mathbb{E}\left[J(., D)\right] + \sigma \mathbb{V}ar\left[J(., D)\right]$$

   for $\sigma \geq 0$ and solve

$$D^* = \mathrm{Argmin}_{D \in \mathcal{A}} \, \mathbb{J}_\sigma(D)$$

Modelisation choice : uncertainties are seen as a random process.

How to take these uncertainties into account in a viable numerical optimization strategy ?

A priori two strategies can be explored in order to compute such a shape:

1. *solve* each problem

$$D^*(\omega) = \text{Argmin}_{D\in\mathcal{A}}\, J(\omega, D)$$

   and take the *expectation* of D^*
   two difficulties : how to define an average shape ? how to compute it ?

2. consider an averaged objective

$$\mathbb{J}_\sigma(D) = \mathbb{E}\left[J(., D)\right] + \sigma \mathbb{V}ar\left[J(., D)\right]$$

   for $\sigma \geq 0$ and solve

$$D^* = \text{Argmin}_{D\in\mathcal{A}}\, \mathbb{J}_\sigma(D)$$

   difficulty: how to compute $\mathbb{J}_\sigma$ and its shape gradient ?

Exploring the first strategy: the Vorob'ev expectation.

Theory of random sets, I. Molchanov, Springer serie Probability and its applications, 2005

- Taking expectation of parametrizations leads to non intrinsic notions of expectation of random sets
- A convenient notion: Vorobe'v expectation.
  - first introduce the *coverage function* p of a random set $D(.)$

$$\forall \mathsf{x} \in \mathbb{R}^d, \quad p(\mathsf{x}) = \mathbb{E}\left[x \in D(.)\right].$$

  - idea: The Vorob'ev expectation $\mathbb{E}_{\mathcal{V}}[D]$ of $D(\omega)$ is then defined as a quantile of $D(\omega)$ such that its volume is $\mathbb{E}[\mathcal{L}^d(D)]$.

Definition (Vorob'ev expectation)

The Vorob'ev expectation $\mathbb{E}_{\mathcal{V}}[D]$ of $D(\omega)$ is defined as the set $\{\mathsf{x} \in \mathbb{R}^2 : p(\mathsf{x}) \geq \mu\}$ for $\mu \in [0,1]$ which is determined from the condition

$$\mathcal{L}^d(\{\mathsf{x} \in \mathbb{R}^d : p(\mathsf{x}) \geq \lambda\}) \leq \mathbb{E}[\mathcal{L}^d(D)] \leq \mathcal{L}^d(\{\mathsf{x} \in \mathbb{R}^d : p(\mathsf{x}) \geq \mu\})$$

for all $\lambda > \mu$.

- Drawback: there is no set-valued notion of correlation but notions of scalar deviation like:

$$\mathcal{D}_{\mathcal{V}}[D] = \mathbb{E}\left[\mathcal{L}^d(D(.)\Delta\mathbb{E}_{\mathcal{V}}[D]\right] \text{ or } \mathcal{D}_{\mathcal{H}}[D] = \mathbb{E}\left[d_{\mathcal{H}}(D(.),\mathbb{E}_{\mathcal{V}}[D])\right]$$

Hence, we get only unsatisfactory Bienaymé-Tchebychev like inequalities.

- Drawback: there is no set-valued notion of correlation but notions of scalar deviation like:

$$\mathcal{D}_{\mathcal{V}}[D] = \mathbb{E}\left[\mathcal{L}^d(D(.)\Delta\mathbb{E}_{\mathcal{V}}[D]\right] \text{ or } \mathcal{D}_{\mathcal{H}}[D] = \mathbb{E}\left[d_{\mathcal{H}}(D(.),\mathbb{E}_{\mathcal{V}}[D])\right]$$

  Hence, we get only unsatisfactory Bienaymé-Tchebychev like inequalities.

- Advantage: this is computable. The idea is to build an estimator for p with i.i.d. copies D_i of D for $1 \leq i \leq M$, by the empirical mean

$$p_M(x) = \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}_{D_i}(x).$$

  P. Heinrich, R. S. Stoica, and V. C. Tran. Level sets estimation and Vorob'ev expectation of random compact sets. Spatial Statistics, 2(1):47–61, 2012.

- Drawback: there is no set-valued notion of correlation but notions of scalar deviation like:

$$\mathcal{D}_{\mathcal{V}}[D] = \mathbb{E}\left[\mathcal{L}^d(D(.)\Delta\mathbb{E}_{\mathcal{V}}[D]\right] \text{ or } \mathcal{D}_{\mathcal{H}}[D] = \mathbb{E}\left[d_{\mathcal{H}}(D(.), \mathbb{E}_{\mathcal{V}}[D])\right]$$

Hence, we get only unsatisfactory Bienaymé-Tchebychev like inequalities.

- Advantage: this is computable. The idea is to build an estimator for p with i.i.d. copies D_i of D for $1 \leq i \leq M$, by the empirical mean

$$p_M(x) = \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}_{D_i}(x).$$

P. Heinrich, R. S. Stoica, and V. C. Tran. Level sets estimation and Vorob'ev expectation of random compact sets. Spatial Statistics, 2(1):47–61, 2012.

- Drawback : this is extremely costly, computing each D_i means *solving a shape optimization problem*

Illustration: exterior Bernoulli free boundary problem

Solution of a free boundary problem in presence of geometric uncertainties, M.D., H. Harbrecht, M.Peters and B. Puig, in redaction

We consider the free boundary problem

$$\begin{cases} \Delta u = 0 & \text{in } D, \\ \|\nabla u\| = f & \text{on } \Gamma, \\ u = 0 & \text{on } \Gamma, \\ u = 1 & \text{on } \Sigma, \end{cases}$$

in the case that the interior boundary is uncertain, i.e., if $\Sigma = \Sigma(\omega)$ with an additional parameter $\omega \in \Omega$.

First example: the setting

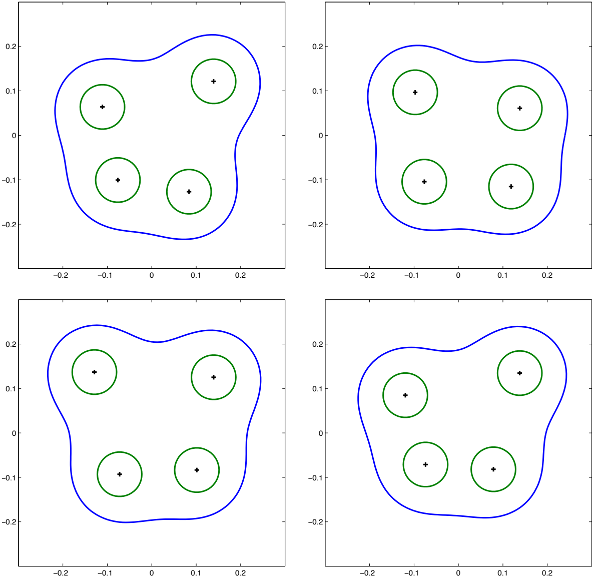
The inner boundary is the union of four circles of radius 0.05 and random centers

$$C_{i,j}(\omega) = \left(\frac{(-1)^i}{10} + 0.04 X_{2(2i+j)}(\omega), \frac{(-1)^i}{10} + 0.04 X_{2(2i+j)+1}(\omega) \right) \text{ for } i,j = 0,1.$$

The random variables $X_1, \ldots, X_8$ are independants, uniform on $[-1, 1]$

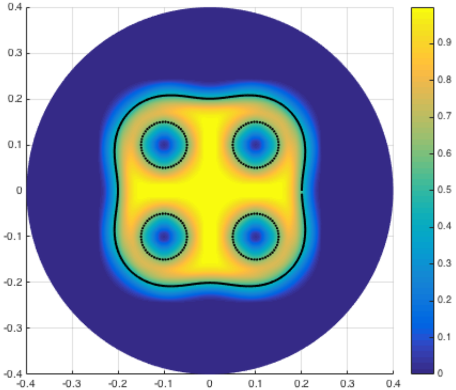
First example: some realisations

Simulations by Michael PETERS



First example: the coverage function and the Vorob'ev expectation

10^6 computations of a free boundary problem: around ten days of computation on a big laptop



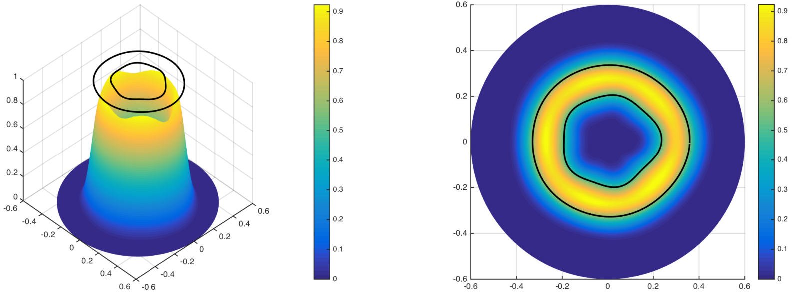
The inner boundary is the curve parametrized in polar coordinates by

$$r(\theta,\omega) = 0.2 + 0.01 f(\theta) + \sum_{k=1}^{10} \frac{\sqrt{2}}{k} \left[X_{2k-1} \cos k\theta + X_{2k} \sin k\theta\right]$$

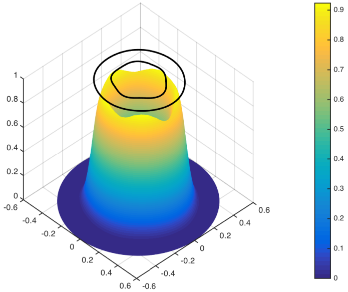
- the random variables $X_1, \ldots, X_8$ are independants, uniform on $[-1, 1]$
- f is the trigonometric polynomial of coefficients

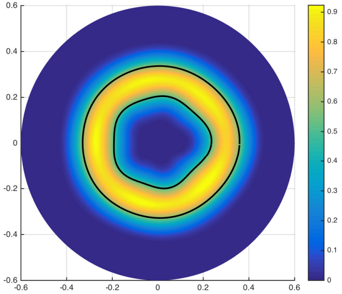
  $$[a_5, \ldots, a_0, b_1, \ldots, b_4] = [0.33, 0.26, 0.51, 0.70, 0.89, 0.48, 0.55, 0.14, 0.15, 0.26]$$

Second example: the coverage function and the Vorob'ev expectation



Second example: the coverage function and the Vorob'ev expectation





the problem is extremely stable $\rightsquigarrow$ it will be interesting to perform similar simulation for a less stable problem but computational cost increases

Exploring the second strategy: using an averaged objective.

Our objective: derive a deterministic expression for $\mathbb{J}$ and its shape gradient to avoid Monte-Carlo method to obtain reasonable computational times

there is one significant case where we know how to do that.

Shape optimization for quadratic functionals and states with random right-hand sides M.D, C. Dapogny and H. Harbrecht. SIAM Control and Optimization 53 (2015), no. 5, 3081–3103

A non trivial case.

- Consider an linear elliptic state equation with random right-hand side. For example, the equations of linear elasticity with random forcing

$$\begin{cases} -\mathrm{div}(Ae(u)) = 0 & \text{in } D, \\ u = 0 & \text{on } \Gamma_D, \\ Ae(u)n = g & \text{on } \Gamma_N, \\ Ae(u)n = 0 & \text{on } \Gamma, \end{cases}$$

where $e(u) = (\nabla u + \nabla u^T)/2$ and with Hooke's law A given by

$$\forall e \in \mathcal{S}(\mathbb{R}^d), \ \ Ae = 2\mu e + \lambda \, \mathrm{Tr} e \, I.$$

A non trivial case.

- Consider an linear elliptic state equation with random right-hand side. For example, the equations of linear elasticity with random forcing

$$\begin{cases} -\mathrm{div}(Ae(u)) = 0 & \text{in } D, \\ u = 0 & \text{on } \Gamma_D, \\ Ae(u)n = g & \text{on } \Gamma_N, \\ Ae(u)n = 0 & \text{on } \Gamma, \end{cases}$$

where $e(u) = (\nabla u + \nabla u^T)/2$ and with Hooke's law A given by

$$\forall e \in \mathcal{S}(\mathbb{R}^d), \;\; Ae = 2\mu e + \lambda \,\mathrm{Tr} e\, I.$$

- Consider a quadratic functional in the state: for example *compliance* of shapes

$$\mathcal{C}(D, \omega) = \int_D Ae(u_D)(x, \omega) : e(u_D)(x, \omega)\, dx = \int_D g(x, \omega) \cdot u_D(x, \omega)\, ds(x),$$

The idea in finite dimension: the setting.

- Consider a finite dimensional space $\mathcal{H}$ of designs h. The cost function is

$$C(h,\omega) =< Bu(h,\omega), u(h,\omega) >$$

where the state $u(h,\omega)$ is defined as the solution of the linear system with random right-hand side

$$A(h)u(h,\omega) = f(\omega)$$

- The averaged objective is

$$\mathcal{M}(h) = \mathbb{E}[C(h,.)] = \int_\Omega \mathcal{C}(h,\omega)\mathbb{P}(d\omega)$$

The idea in finite dimension: doubling the variables to transform quadratic quantities in linear ones.

Sparse finite elements for elliptic problems with stochastic loading, Numerische Mathematik, C. Schwab, R.Todor, 95/4 (2003), pp. 707-734

- Rewrite cost function as

$$C(h,\omega) =< Bu(h,\omega), u(h,\omega) >= \mathcal{B} : (u(h,\omega) \otimes u(h,\omega))$$

$v \otimes w$ of two vectors $v, w \in \mathbb{R}^N$ is the $(N \times N)$-matrix with entries $(v \otimes w)_{i,j} = v_i w_j$, $i, j = 1, \ldots, N$,

: stands for the Frobenius inner product over matrices.

- The averaged objective becomes

$$\mathcal{M}(h) = \mathbb{E}[C(h,.)] = \int_\Omega \mathcal{C}(h,\omega)\mathbb{P}(d\omega) = \mathcal{B} : \mathrm{Cov}(u,u)(h)$$

$\mathsf{Cov}(u,v)(h)$ is the *covariance matrix* (of size N^2) of $u(h,\omega)$, whose entries read

$$\mathsf{Cov}(u,v)(h)_{i,j} = \int_\Omega u_i(h,\omega)v_j(h,\omega)\,\mathbb{P}(d\omega), \quad i,j = 1,\ldots,N.$$

Notation: $\mathsf{Cov}(u) = \mathsf{Cov}(u,u)$

- $\mathrm{Cov}(u)(h)$ can be directly computed: it solves the (N^2)-dimensional system
$$(\mathcal{A}(h) \otimes \mathcal{A}(h))\,\mathrm{Cov}(u)(h) = \mathrm{Cov}(f).$$
- we can derive with respect to h in the direction $\widehat{h}$:
$$\mathcal{M}'(h)(\widehat{h}) = \left(\mathcal{A}'(h)(\widehat{h}) \otimes I\right)\mathrm{Cov}(u,p)(h).$$
where $\mathrm{Cov}(u,p)(h)$ solves
$$\left(\mathcal{A}(h) \otimes \mathcal{A}(h)^T\right)\mathrm{Cov}(u,p)(h) = -\,(\mathcal{A}(h) \otimes \mathcal{B})\,\mathrm{Cov}(u)(h).$$

Conclusion

Both, the objective function $\mathcal{M}(h)$ and its gradient, can be exactly calculated from the sole datum of the covariance matrix of f (and not of its law!).

Going back to the compliance case

Following the previous idea, $\mathcal{M}(D)$ can be rewritten

$$\mathcal{M}(D) = \int_D ((Ae_x : e_y)\mathrm{Cov}(u))(x,x)\,dx,$$

$(Ae_x : e_y) : [H^1_{\Gamma_D}(D)]^d \otimes [H^1_{\Gamma_D}(D)]^d \to L^2(D) \otimes L^2(D)$ is the linear operator induced from the bilinear mapping

$$(u,v) \mapsto Ae(u) : e(v).$$

its derivative reads

$$\forall \theta \in \Theta_{ad},\ \ \mathcal{M}'(D)(\theta) = -\int_\Gamma ((Ae_x : e_y)\mathrm{Cov}(u))(x,x)(\theta \cdot n)(x)\,ds(x).$$

It remains to compute $\mathrm{Cov}(u)$

$\mathrm{Cov}(u) \in [H^1_{\Gamma_D}(D)]^d \otimes [H^1_{\Gamma_D}(D)]^d$ is the unique solution to the following boundary value problem:

$$(\mathrm{div}_x \otimes \mathrm{div}_y)(Ae_x \otimes Ae_y)\,\mathrm{Cov}(u) = 0 \text{ in } D \times D,$$

with boundary conditions

$$\begin{cases} \mathrm{Cov}(u) = 0 & \text{on } \Gamma_D \times \Gamma_D, \\ (\mathrm{div}_x \otimes I_y)(Ae_x \otimes I_y)\,\mathrm{Cov}(u) = 0 & \text{on } D \times \Gamma_D, \\ (I_x \otimes \mathrm{div}_y)(I_x \otimes Ae_y)\,\mathrm{Cov}(u) = 0 & \text{on } \Gamma_D \times D, \\ (Ae_x \otimes Ae_y)\,\mathrm{Cov}(u)(n_x \otimes n_y) = \mathrm{Cov}(g) & \text{on } \Gamma_N \times \Gamma_N, \\ (\mathrm{div}_x \otimes I_y)(Ae_x \otimes Ae_y)\,\mathrm{Cov}(u)(I_x \otimes n_y) = 0 & \text{on } D \times (\Gamma_N \cup \Gamma), \\ (I_x \otimes \mathrm{div}_y)(Ae_x \otimes Ae_y)\,\mathrm{Cov}(u)(n_x \otimes I_y) = 0 & \text{on } (\Gamma_N \cup \Gamma_N) \times D, \\ (Ae_x \otimes Ae_y)\,\mathrm{Cov}(u)(n_x \otimes n_y) = 0 & \text{on } \big((\Gamma_N \cup \Gamma) \times (\Gamma_N \cup \Gamma)\big) \setminus (\Gamma_N \times \Gamma_N), \\ (Ae_x \otimes I_y)\,\mathrm{Cov}(u)(n_x \otimes I_y) = 0 & \text{on } (\Gamma_N \times \Gamma) \times \Gamma_D, \\ (I_x \otimes Ae_y)\,\mathrm{Cov}(u)(I_x \otimes n_y) = 0 & \text{on } \Gamma_D \times (\Gamma_N \times \Gamma). \end{cases}$$

Computing $\mathrm{Cov}(u)(h)$*: low-rank approximation of* $\mathrm{Cov}(f)$*.*

The idea in finite dimension

If

$$\mathrm{Cov}(f) \approx \sum_{i=1}^{m} f_i \otimes f_i,$$

Then,

$$\mathrm{Cov}(u)(h) \approx \sum_{i=1}^{m} u_i(h) \otimes u_i(h), \text{ and } \mathrm{Cov}(u,p)(h) \approx \sum_{i=1}^{m} u_i(h) \otimes p_i(h)$$

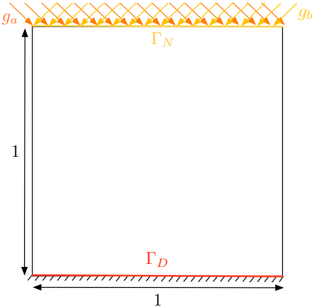
where $u_i(h)$ (resp. $p_i(h)$) solves

$$\mathcal{A}(h) u_i(h) = f_i \text{ resp. } \mathcal{A}(h)^T p_i(h) = -\mathcal{B}^T u_i(h).$$

Cost of the computation

Computing and its gradient cost m usual systems for the state and m for the adjoint.

Example: a robust bridge



The loadings are

$$g(x,\omega) = \xi_1(\omega) g_a(x) + \xi_2(\omega) g_b(x),$$

with

$g_a = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $g_b = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$,

ξ_1 and ξ_2 are centered of variance 1.

Our goal: minimize the averaged compliance under the volume constraint $Vol(D) = 0.35$ enforced owing to a standard Augmented Lagrangian procedure.

Simulations obtained by C. Dapogny - each 12 min on a MacBook air

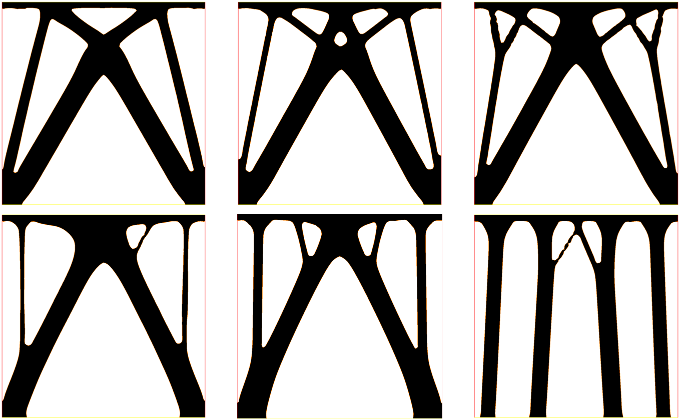
Set $\alpha := \int_\Omega \xi_1 \xi_2 \, \mathbb{P}(d\omega)$ so that

$$\mathrm{Cor}(g) = g_a \otimes g_a + g_b \otimes g_b + \alpha \left(g_a \otimes g_b + g_b \otimes g_a \right).$$

Simulations obtained by C. Dapogny - each 12 min on a MacBook air

Set $\alpha := \int_\Omega \xi_1 \xi_2 \, \mathbb{P}(d\omega)$ so that

$$\mathrm{Cor}(g) = g_a \otimes g_a + g_b \otimes g_b + \alpha \left(g_a \otimes g_b + g_b \otimes g_a \right).$$



Optimal shapes obtained associated to degrees of correlation $\alpha = -1, -0.7, 0, 0.5, 0.8, 1$ (from left to right, top to bottom).