Méthodes numériques pour les EDP.
Projet 2: bases réduites et inégalités variationnelles.

Ce projet consiste à implémenter une méthode de base réduite pour une inégalité variationnelle. On considère le problème **paramétré** :

Définition 1. *Étant donné $\mu \in \mathcal{P}$, trouver $(u(\mu), \lambda(\mu)) \in V \times M$ tel que*

$$\begin{aligned} a(u(\mu), v; \mu) + b(v, \lambda(\mu)) &= f(v; \mu), \qquad & v \in V \\ b(u(\mu), \eta - \lambda(\mu)) &\leq g(\eta - \lambda(\mu); \mu), \qquad & \eta \in M, \end{aligned}$$

où :

- $\mathcal{P}$ est un ensemble de paramètres,
- $u, v \in V$, V un espace de Hilbert,
- $\lambda \in M$, $M \subset W$ un cône d'un espace de Hilert W,
- a est une forme bilinéaire continue coercive, c'est-à-dire qu'il existe $\alpha(\mu)$ et $\gamma(\mu)$ telles que :

$$|a(u, v; \mu)| \leq \gamma(\mu)\|u\|.\|v\|,$$

$$\alpha(\mu)\|u\|^2 \leq |a(u, u; \mu)|,$$

- b est une forme bilinéaire continuer vérifiant une condition "de stabilité inf-sup", i.e. $\beta > 0$:

$$\inf_{\eta \in W} \sup_{v \in V} b(v, \eta)/(\|v\|_V \, \|\eta\|_W) \geq \beta$$

- f, g sont des formes linéaires continues sur V et W respectivement : $\|f(\cdot; \mu)\|_V \leq \gamma_f(\mu)$ et $\|g(\cdot; \mu)\|_W \leq \gamma_g(\mu)$.

1 Discrétisation par base réduite

Pour résoudre efficacement ce problème, on souhaite construire une méthode de base réduite. Pour ce faire, on considère $S = \{\mu_1, \ldots, \mu_N\} \subset \mathcal{P}$, un ensemble de jeu de paramètres et les espaces réduits suivants :

- $W_N := \mathrm{span}\{\lambda(\mu_i)\}_{i=1}^N \subset W$,
- $M_N := \mathrm{span}_+\{\lambda(\mu_i)\}_{i=1}^N := \left\{\sum_{i=1}^N \alpha_i \lambda(\mu_i) | \alpha_i \geq 0\right\} \subset M$,
- $V_N := \mathrm{span}\{u(\mu_i)\}_{i=1}^N \subset V$.

On introduit alors le problème réduit :

Définition 2. *Étant donné $\mu \in \mathcal{P}$, trouver $(u_N(\mu), \lambda_N(\mu)) \in V_N \times M_N$ tel que*

$$\begin{aligned} a(u_N(\mu), v_N; \mu) + b(v_N, \lambda_N(\mu)) &= f(v_N; \mu), \quad v_N \in V_N & (1) \\ b(u_N(\mu), \eta_N - \lambda_N(\mu)) &\leq g(\eta_N - \lambda_N(\mu); \mu), \quad \eta_N \in M_N. & (2) \end{aligned}$$

2 Travail théorique

On souhaite obtenir une estimation a posteriori. On introduit pour cela les résidus:

$$\begin{aligned} r(v;\mu) &:= f(v;\mu) - a(u_N(\mu), v;\mu) - b(v, \lambda_N(\mu)), \quad v \in V. \\ s(\eta;\mu) &:= b(u_N(\mu), \eta) - g(\eta;\mu), \quad \eta \in W. \end{aligned}$$

Notons que le résidu s n'est pas censé s'annuler, mais seulement etre négatif. On introduit également leur représentants de Riesz $v_r(\mu) \in V, \eta_s(\mu) \in W$, définis par

$$\begin{aligned} \langle v, v_r(\mu)\rangle_V &= r(v;\mu), \quad v \in V, \\ \langle \eta, \eta_s(\mu)\rangle_W &= s(\eta;\mu), \quad \eta \in W. \end{aligned}$$

Prouver les inégalités suivantes :

1. Pour tout $v \in V$ et $\mu \in \mathcal{P}$

$$r(v;\mu) = a(u(\mu) - u_N(\mu), v;\mu) + b(v, \lambda(\mu) - \lambda_N(\mu)).$$

Proof. L'erreur de consistence:

$$\begin{aligned} r(v) &= f(v) - a(u_N, v) - b(v, \lambda_N) \\ &= f(v) - a(u_N, v) - b(v, \lambda_N) - [f(v) - a(u, v) - b(v, \lambda)] \\ &= a(u - u_N, v) + b(v, \lambda - \lambda_N) \end{aligned}$$

□

2. Pour tout $\mu \in \mathcal{P}$

$$\|\lambda(\mu) - \lambda_N(\mu)\|_W \leq \frac{1}{\beta}\left(\|r(\cdot;\mu)\|_V + \gamma_a(\mu)\|u(\mu) - u_N(\mu)\|_V\right).$$

Proof. By the "inf-sup stability" of the bilinear form b, $\exists \beta$ s.t. $\inf_{\eta\in W}\sup_{v\in V}\frac{b(v,\eta)}{\|v\|_V\|\eta\|_W} \geq \beta$, which implies $\sup_{v\in V}\frac{b(v,\lambda-\lambda_N)}{\|v\|_V\|\lambda-\lambda_N\|_W} \geq \beta$. On the other hand, by continuity of a, we have $|a(u,v)| \leq \gamma_a\|u\|\|v\|$. Then,

$$\begin{aligned} \|\lambda - \lambda_N\|_W &\leq \frac{1}{\beta}\sup_v \frac{b(v, \lambda - \lambda_N)}{\|v\|} \\ &= \frac{1}{\beta}\sup_v \frac{r(v) - a(u - u_N, v)}{\|v\|} \\ &\leq \frac{1}{\beta}\sup_v \frac{\|r\|\|v\| - \gamma_a\|u - u_N\|\|v\|}{\|v\|} \\ &= \frac{1}{\beta}(\|r\| - \gamma_a\|u - u_N\|) \end{aligned}$$

□

3. On pose

$$\begin{aligned}
\delta_r(\mu) &:= \|r(\cdot;\mu)\|_{V'} = \|v_r(\mu)\|_V \\
\delta_{s1}(\mu) &:= \|\pi(\eta_s(\mu))\|_W \\
\delta_{s2}(\mu) &:= \langle \lambda_N(\mu), \pi(\eta_s(\mu))\rangle_W .
\end{aligned}$$

Expliquer chacune des étapes du calcul suivant :

$$\begin{aligned}
\alpha \|u-u_N\|_V^2 &\leq a(u-u_N, u-u_N) \ \text{by coercivity} \\
&= r(u-u_N) - b(u-u_N, \lambda-\lambda_N) \ \text{by question (1)} \\
&\leq \delta_r \|u-u_N\|_V + b(u, \lambda_N - \lambda) + b(u_N, \lambda-\lambda_N) \ \text{by continuity of r}
\end{aligned}$$

Since (u, λ) is a solution of the system, we know that $g(\lambda_N - \lambda) \geq b(u, \lambda_N - \lambda)$, then by definition of s we have

$$\begin{aligned}
&\leq \delta_r \|u-u_N\|_V + g(\lambda_N - \lambda) + s(\lambda-\lambda_N) + g(\lambda - \lambda_N) \\
&= \delta_r \|u-u_N\|_V + s(\lambda-\lambda_N)
\end{aligned}$$

Now, we claim that $s(\lambda_N) = 0$. In fact, we know that $b(u_N, \eta_N - \lambda_N) \leq g(\eta_N - \lambda_N)$ for any $\eta_N \in M_N$. In Particular, taking $\eta_N = 0$ and $\eta_N = 2\lambda_N$ we obtain $b(u_N, \lambda_N) \geq g(\lambda_N)$ and $b(u_N, \lambda_N) \leq g(\lambda_N)$ respectively. This implies

$$\begin{aligned}
\alpha \|u-u_N\|_V^2 &\leq \delta_r \|u-u_N\|_V + s(\lambda) \\
&= \delta_r \|u-u_N\|_V + \langle \lambda, \pi(\eta_s)\rangle_W + \langle \lambda, \eta_s - \pi(\eta_s)\rangle_W
\end{aligned}$$

Here we assume that the coordinate coefficients λ with respect to the "base" M are positive, since it's a positive cone. To do this in a more rigorous way, we suggest to use the definition given in [3] : π is a projection from W to M, in the sense of standard orthogonal projection. Consequently, $\langle \lambda, \eta_s - \pi(\eta_s)\rangle_W \leq 0$.

$$\begin{aligned}
\alpha \|u-u_N\|_V^2 &\leq \delta_r \|u-u_N\|_V + \langle \lambda, \pi(\eta_s)\rangle_W + \langle \lambda, \eta_s - \pi(\eta_s)\rangle_W \\
&\leq \delta_r \|u-u_N\|_V + \langle \lambda, \pi(\eta_s)\rangle_W \\
&= \delta_r \|u-u_N\|_V + \langle \lambda - \lambda_N, \pi(\eta_s)\rangle_W + \delta_{s2} \\
&\leq \delta_r \|u-u_N\|_V + \|\lambda-\lambda_N\|_W \delta_{s1} + \delta_{s2} \ \text{by Cauchy–Schwarz inequality,}
\end{aligned}$$

où π la partie positive : une composante de $\pi(v)$ est soit 0 si la composante de v est négative, soit la composante elle-même si elle est positive.

4. En déduire:

$$\begin{aligned}
\|u(\mu) - u_N(\mu)\|_V &\leq \Delta_u(\mu) := c_1(\mu) + \sqrt{c_1(\mu)^2 + c_2(\mu)}, \\
\|\lambda(\mu) - \lambda_N(\mu)\|_W &\leq \Delta_\lambda(\mu) := \frac{1}{\beta}\left(\delta_r(\mu) + \gamma_a(\mu)\Delta_u(\mu)\right),
\end{aligned}$$

avec les constantes :

$$c_1(\mu) := \frac{1}{2\alpha(\mu)}\left(\delta_r(\mu) + \frac{\delta_{s1}(\mu)\gamma_a(\mu)}{\beta}\right), \quad c_2(\mu) := \frac{1}{\alpha(\mu)}\left(\frac{\delta_{s1}(\mu)\delta_r(\mu)}{\beta} + \delta_{s2}(\mu)\right).$$

Proof. From 2 and 3 we conclude that

$$\alpha\|u-u_N\|_V^2 \le \delta_r\,\|u-u_N\|_V + \frac{\delta_{s1}}{\beta}(\delta_r+\gamma_a\,\|u-u_N\|_V)+\delta_{s2}.$$

Which implies

$$\alpha\,\|u-u_N\|_V^2 - (\delta_r+\frac{\gamma_a\delta_{s1}}{\beta})\,\|u-u_N\|_V + \frac{\delta_r\delta_{s1}}{\beta}+\delta_{s2}\le 0$$

and consequently

$$\begin{aligned}
\|u(\mu)-u_N(\mu)\|_V &\le \frac{\delta_r+\frac{\gamma_a\delta_{s1}}{\beta}+\sqrt{\left(\delta_r+\frac{\gamma_a\delta_{s1}}{\beta}\right)^2+4\alpha\left(\frac{\delta_r\delta_{s1}}{\beta}+\delta_{s2}\right)}}{2\alpha}\\
&= c_1(\mu)+\sqrt{c_1(\mu)^2+c_2(\mu)}\\
&= \Delta_u(\mu).
\end{aligned}$$

The second inequality is just consequence of 2 and the previous inequality. □

5. Montrer que les estimateurs a posteriori trouvés s'annulent si $(u_N(\mu),\lambda_N(\mu))=(u(\mu),\lambda(\mu))$.

   Proof. Let $(u_N(\mu),\lambda_N(\mu))=(u(\mu),\lambda(\mu))$ then by 1. we have

   $$r(v;\mu)=a(u(\mu)-u_N(\mu),v;\mu)-b(v,\lambda(\mu)-\lambda_N(\mu))=0.$$

   Which implies $\delta_r=0$.

   On the other hand, $\langle\eta_s,\eta\rangle_W=b(u,\eta)-g(\eta)\le 0,\ \forall\eta\in M$. Where follows that $\pi(\eta_s)=0$, so $\delta_{s1}=\delta_{s2}=0$. Then, we conclude that $\Delta_u(\mu)=\Delta_\lambda(\mu)=0$. □

3 Travail pratique

On considère les données suivantes, correspondant à un fil élastique suspendu entre les deux points $(0,0)$ et $(1,0)$. Le domaine $\Omega=(0,1)$ est discrétisé à l'aide d'un maillage uniforme de pas $\Delta x := 1/K$ for $K\in\mathbb{N}$. L'espace V est celui correspondant à une discrétisation en élément fini P^1 :

$$V := \{v\in H_0^1(\Omega)|v_{|[x_k,x_{k+1}]}\in P_1, k=0,\dots,K-1\}$$

de dimension $H_V=H_W=H:=K-1=201$ avec $x_k:=k\Delta x$. On choisit la base standard $\psi_i\in V$ de noeuds $x_i\in\Omega$, i.e., $\psi_i(x_j)=\delta_{ij}, i,j=1,\dots,H$. Le cône est défini par $M:=\text{span}_+\{\chi_i\}_{i=1}^H$, où la famille $(\chi_i)_{i=1}^H$ est choisie telle que $\underline{B}=(b(\psi_i,\chi_j))_{i,j=1}^{H,H}$ correspondant à $b(\cdot,\cdot)$ soit la matrice identité.
Les formes bilinéaires a et b sont données par :

$$\begin{aligned}
a(u,v;\mu) &:= \int_\Omega \nu(\mu)(x)\nabla u(x)\cdot\nabla v(x)dx\ , \quad u,v\in V\\
b(u,\eta) &:= -\eta(u), \quad u\in V,\eta\in W
\end{aligned}$$

avec $\nu(\mu)(x)=\mu_1 Ind_{[0,1/2]}(x)+\nu_0 Ind_{[1/2,1]}(x)$, $\nu_0=0.15$, qui caractérise l'élasticité du fil. La fonction obstacle est

$$h(x;\mu)=-0.2\,(\sin(\pi x)-\sin(3\pi x))-0.5+\mu_2(x-0.5).$$

On fixe $\mathcal{P} := [0.05, 0.25] \times [-0.05, 0.5] \subset \mathbb{R}^2$. Les formes linéaires f et g sont données par :

$$f(v;\mu) = f(v) := -\int_\Omega v(x)dx, \quad v \in V.$$

$$g(\eta;\mu) = \sum_{i=1}^{H} \underline{\eta}_i h(x_i;\mu) \quad \text{for} \quad \eta = \sum_{i=1}^{H} \underline{\eta}_i \chi_i \in W$$

1. Faire un code pour le problème élément fini (de dimension H).

2. Expliquer pourquoi après discrétisation, le système se met sous la forme :

$$\begin{aligned} \overline{A}_N(\mu)\overline{u}_N(\mu) + \overline{B}_N\overline{\lambda}_N(\mu) &= \overline{f}_N(\mu) \\ \overline{\lambda}_N(\mu) &\geq 0 \\ \overline{g}_N(\mu) - \overline{B}_N^T\overline{u}_N(\mu) &\geq 0 \\ \overline{\lambda}_N(\mu)^T(\overline{g}_N(\mu) - \overline{B}_N^T\overline{u}_N(\mu)) &= 0. \end{aligned}$$

   On donnera la définition des matrices et vecteurs impliqués dans cette formule.

   Proof. Let $\{u_n\}_{n=1}^N$ be the reduce base associated to the space V_N, $\{\lambda_n\}_{n=1}^N$ be the reduce base associated to the space W_N and $(u_N, \lambda_N) \in V_N \times W_N$ the reduced solution, we write $\overline{u_N}^T = (\alpha_1, \ldots, \alpha_N)$ and $\overline{\lambda_N}^T = (\beta_1, \ldots, \beta_N)$ such that

$$u_N = \sum_i \alpha_i u_i; \;\; \lambda_N = \sum_i \beta_i \lambda_i.$$

   Since (u_N, λ_N) satisfies 1, we know

$$a(\sum_i \alpha_i u_i, u_j) + b(u_j, \sum_i \beta_i \lambda_i) = f(u_j)$$
$$\sum_i \alpha_i a(u_i, u_j) + \sum_i \beta_i b(u_j, \lambda_i) = f(u_j),$$

   which can be written in the form

$$\overline{A}_N(\mu)\overline{u}_N(\mu) + \overline{B}_N\overline{\lambda}_N(\mu) = \overline{f}_N(\mu)$$

   where $\overline{A}_N(\mu) = (a_{ij} = a(u_j, u_i, \mu))_{N\times N}$, $\overline{B}_N = (b_{ij} = b(u_i, \lambda_j))_{N\times N}$ and $\overline{f}_N = (f(u_i, \mu))_{N\times 1}$. The vector $\overline{\lambda}_N(\mu) \geq 0$ since W_N is a cone. On the other hand, as is seen above, the solution must satisfy $b(u_N, \lambda_N) = g(\lambda_N; \mu)$, which implies $b(u_N, \eta) = g(\eta, \mu)$ for any $\eta \in M_N$. From what we obtain

$$b(\sum_i \alpha_i u_i, \sum_j \beta_j \lambda_j) = g(\sum_j \beta_j \lambda_j; \mu)$$

$$\sum_j \beta_j (\sum_i \alpha_i b(u_i, \lambda_j) - g(\lambda_j)) = 0$$

$$\overline{\lambda}_N^T(\overline{g}_N - \overline{B}_N^T\overline{u}_N) = 0$$

   and taking $\eta = \lambda_j$ we get

$$\overline{g}_N - \overline{B}_N^T\overline{u}_N \geq 0$$

   where $\overline{g}_N = (g_i) = g(\lambda_i; \mu))_{N\times 1}$. □

3. À l'aide de la fonction octave 'qp', faire un code qui résoud ce problème.

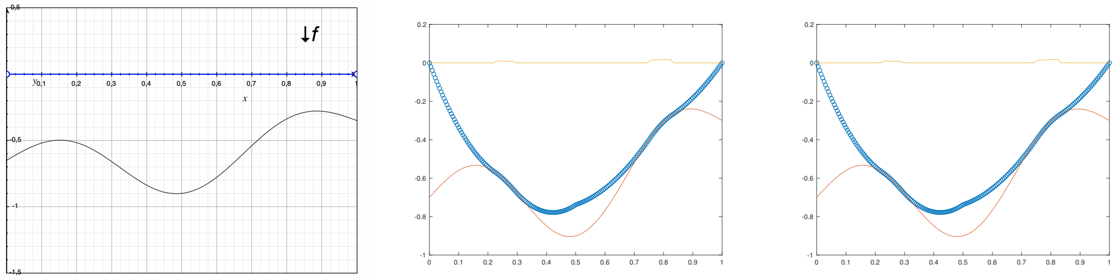

Figure 1: Left: an edge-fixed rod under gravity. Middle: sample solution with active-set method (implemented with Matlab@). Right: sample solution with inner point method (Matlab@ built-in function). Blue: primal solution. Yellow: dual solution. Red: supporting plane

4. Coder deux générateurs de bases réduites[1] pour ce problème : l'un basé sur la véritable erreur, le dernier avec les estimateurs a posteriori (on pourra prendre la somme de $\Delta_u(\mu)$ et $\Delta_\lambda(\mu)$).

5. Tester les deux bases : on prendra comme critère l'erreur maximum observée sur un ensemble random dans $\mathcal{P}$.

[1]Le terme de "base" est abusif pour ce qui concerne l'espace M_N, car c'est un cône...

4 Numerical Experiments

4.1 Physical Background

Consider an ideal inhomogeneous rod placed horizontally, with position $u(x)$ and elasticity ν, under uniform gravity$f = -1$, fixed at both ends, and supported by an absolutely smooth surface h below (Shown in Figure (1)). This is a constrained non-divergent-form elliptic PDE with Dirichlet boundary condition:

$$\begin{cases} -\nabla \cdot (\nu \nabla u) = f & \text{in } (0,1) \\ u = 0 & \text{on } \{0,1\} \\ u \geq h & \end{cases}$$

Our first approach is to formulate it as a variational problem, writing the Dirichlet functional as

$$\inf_{u \in H_0^1(\Omega)} \quad J(u) = \tfrac{1}{2} \int_0^1 \nu |\nabla u|^2 \mathrm{d}x - \int_0^1 f \mathrm{d}x \tag{3}$$

$$\text{s.t.} \qquad u \geq h \tag{4}$$

and introducing functional space $W = H_0^1(\Omega)$ with $\Omega = (0,1)$, which characterizes Dirichlet condition. With the Lagrangian multiplier, we translate this problem into a saddle-point form. (One can refer to e.g. [1])

$$\inf_{u \in H_0^1(\Omega)} \sup_{\begin{cases} \lambda \in H^{-1}(\Omega) \\ \lambda \geq 0 \end{cases}} \frac{1}{2} \int_0^1 \nu |\nabla u|^2 \mathrm{d}x + \int_0^1 \lambda (h - u) - \int_0^1 f(u)$$

where λ in the positive cone $M = \{v \in H^{-1}(\Omega) | v \geq 0\}$. Note that surface h and elasticity ν are parameterized with μ. In our setting, conditions and parameters is given by $\nu(x) = \mu_1 1_{(0,\frac{1}{2})} + \nu_0 1_{(\frac{1}{2},1)}$ (elasticity at the mid point is $\frac{\mu_0+\nu_1}{2}$), and $h(x;\mu) = -0.2\,(\sin(\pi x) - \sin(3\pi x)) - 0.5 + \mu_2(x - 0.5)$.

In order to compute the problem efficiently, we introduce our second approach, Reduced Basis Method (RB), so that offline pre-computations are conducted to reduce time during the online phase. The reduced form of algebraic saddle point problem is described in section (3.2), and is proven before. Note that the solution $\bar{u_N}$ (coordinate coefficient) is not unique. However the reduced solution $u_N = \sum_{i=1}^N \bar{u_N} \varphi_i$ is unique, where (φ_i) is the reduced basis. Another way to deduce the reduced problem (algebraically) is shown below:

First we write the primal-dual formulation of QP using the KKT condition of the Lagrangance:

$$\begin{cases} A\bar{u} + B\bar{\lambda} = \bar{f} \\ \bar{u} - h \geq 0 \\ \bar{\lambda} > 0 \\ \lambda^T(\bar{u} - h) = 0 \end{cases}$$

Setting $\bar{u} = U_N * \overline{u_N}$, and $\bar{\lambda} = \Lambda_N * \overline{\lambda_N}$ where $U_N = [\bar{u_1}, \ldots, \bar{u_N}]$ is the primal reduced basis, $\Lambda_N = [\bar{\lambda_1}, \ldots, \bar{\lambda_N}]$ is the dual reduced basis, and $\overline{u_N} = (\alpha_1, \ldots, \alpha_N)^T$. Multiplying both sides with Λ_N^T or U_N^T, we conclude with the reduced formulation:

$$\begin{cases} U_N^T A U_N \overline{u_N} + U_N^T B \Lambda_N \overline{\lambda_N} = \overline{f} \\ -\Lambda_N^T h - (U_N^T B \Lambda_N)^T \overline{u_N} \geq 0 \\ \overline{\lambda_N} > 0 \\ \overline{\lambda_N}^T \big(- \Lambda_N^T h - (U_N^T B \Lambda_N)^T \overline{u_N} \big) = 0 \end{cases}$$

5 Outcome

We find that

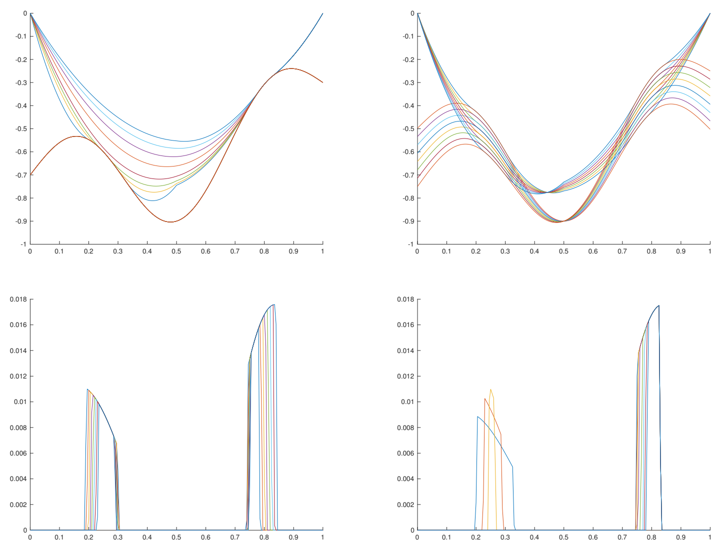

Figure 2: Left: fix μ_2 and adjust μ_1 to be evenly distributed over parameter space. Right: fix μ_1 and adjust μ_2. Up: primal solution. Bottom: dual solution.

6 Computational Aspects

6.1 Different QP solvers

We have tested both active set method (AS) and interior point method (IP), and find that AS is more efficient in low dimension, while IP has better scalability. This coincides with the result given in [4].

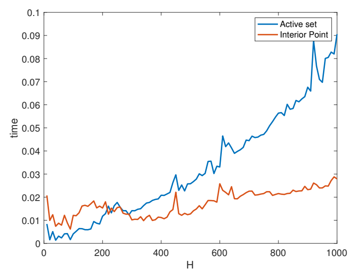

Figure 3: Comparison between Active Set method (implemented) and Interior Point method (built-in)

6.2 Evaluation of Sobolev norm

By definition of δ_r,δ_{s1} and δ_{s2}, we have to compute numerically the Sobolev norm and the dual Sobolev norm. As is suggested in [3] (or [2] for a detailed discussion), we used

$$\|u\|_{H_0^1} \simeq \overline{u}^T R_{H_0^1} \bar{u}$$

$$\|u\|_{H^{-1}} \simeq \bar{u}^T R_{H_0^1}^{-1} \bar{u}$$

with $R_{H^1} = K + M$, $R_{H_0^1} = K$, where K is the stiffness matrix, and M is the mass matrix.

References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[2] Carsten Burstedde. *On the numerical evaluation of fractional Sobolev norms.* Citeseer, 2006.

[3] Bernard Haasdonk, Julien Salomon, and Barbara Wohlmuth. A reduced basis method for parametrized variational inequalities. *SIAM Journal on Numerical Analysis*, 50(5):2656–2676, 2012.

[4] Mark SK Lau, Siew Peng Yue, Keck Voon Ling, and Jan M Maciejowski. A comparison of interior point and active set methods for fpga implementation of model predictive control. In *2009 European Control Conference (ECC)*, pages 156–161. IEEE, 2009.