The latent variable proximal point algorithm for variational problems with inequality constraints

John Papadopoulos[1], Jørgen Dokken[2], Patrick Farrell[3],
Brendan Keith[4], Thomas Surowiec[2]

[1]Weierstrass Institute Berlin, [2]Simula, [3]Oxford, [4]Brown

June 26, 2025, 30th Biennial Numerical Analysis Meeting

Introduction

Pointwise constraints appear everywhere, e.g. contact mechanics (non-penetration), stress constraints in elasticity, sandpile growth, financial mathematics, pattern formation, engineering design, biological models...

Optimization problem

$$\min_{u \in U} J(u) \quad \text{subject to} \quad Bu(x) \in C(x) \text{ for almost every } x \in \Omega.$$

- primal-dual active set, multigrid, finite-dimensional constrained optimizers (often mesh dependent, confined to low-order[1]).
- penalty methods (infeasible solutions, suboptimal for high-order, ill-conditioning).

[1] With notable exceptions in Kirby & Shapero (2024) and Banz & Schröder (2015).

Introduction

Pointwise constraints appear everywhere, e.g. contact mechanics (non-penetration), stress constraints in elasticity, sandpile growth, financial mathematics, pattern formation, engineering design, biological models...

Optimization problem

$$\min_{u \in U} J(u) \quad \text{subject to} \quad Bu(x) \in C(x) \text{ for almost every } x \in \Omega.$$

- primal-dual active set, multigrid, finite-dimensional constrained optimizers (often mesh dependent, confined to low-order[1]).
- penalty methods (infeasible solutions, suboptimal for high-order, ill-conditioning).

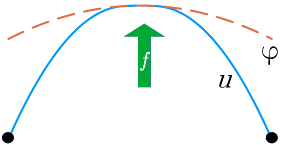
[1] With notable exceptions in Kirby & Shapero (2024) and Banz & Schröder (2015).

Introduction

Pointwise constraints appear everywhere, e.g. contact mechanics (non-penetration), stress constraints in elasticity, sandpile growth, financial mathematics, pattern formation, engineering design, biological models...

Optimization problem

$$\min_{u \in U} J(u) \text{ subject to } Bu(x) \in C(x) \text{ for almost every } x \in \Omega.$$



1D obstacle problem

- primal-dual active set, multigrid, finite-dimensional constrained optimizers (often mesh dependent, confined to low-order[1]).
- penalty methods (infeasible solutions, suboptimal for high-order, ill-conditioning).

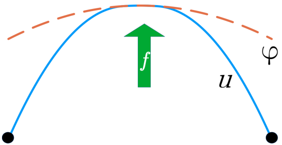
[1] With notable exceptions in Kirby & Shapero (2024) and Banz & Schröder (2015).

Introduction

Pointwise constraints appear everywhere, e.g. contact mechanics (non-penetration), stress constraints in elasticity, sandpile growth, financial mathematics, pattern formation, engineering design, biological models...

Optimization problem

$$\min_{u \in U} J(u) \text{ subject to } Bu(x) \in C(x) \text{ for almost every } x \in \Omega.$$



1D obstacle problem

- primal-dual active set, multigrid, finite-dimensional constrained optimizers (often mesh dependent, confined to low-order[1]).
- penalty methods (infeasible solutions, suboptimal for high-order, ill-conditioning).

[1] With notable exceptions in Kirby & Shapero (2024) and Banz & Schröder (2015).

Problems of interest

Examples

- (Obstacle problem.) Find $u : \Omega \to \mathbb{R}$,

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \quad \text{subject to } u(x) \leq \varphi(x).$$

- (Elastic-plastic torsion.) Find $u : \Omega \to \mathbb{R}$,

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \quad \text{subject to } |\nabla u|(x) \leq \varphi(x).$$

- (Signorini.) Find $u : \Omega \to \mathbb{R}^d$,

$$\min_{u \in H_g^1(\Omega)^d} \int_\Omega \frac{1}{2}(\mathbf{C}\varepsilon(u)) : \varepsilon(u) - f \cdot u \,\mathrm{d}x \quad \text{subject to } u \cdot \tilde{n} \geq 0 \text{ on } \Gamma_T.$$

Problems of interest

Examples

- (Obstacle problem.) Find $u : \Omega \to \mathbb{R}$,

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \;\text{ subject to } u(x) \leq \varphi(x).$$

- (Elastic-plastic torsion.) Find $u : \Omega \to \mathbb{R}$,

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \;\text{ subject to } |\nabla u|(x) \leq \varphi(x).$$

- (Signorini.) Find $u : \Omega \to \mathbb{R}^d$,

$$\min_{u \in H_g^1(\Omega)^d} \int_\Omega \frac{1}{2}(\mathbf{C}\varepsilon(u)) : \varepsilon(u) - f \cdot u \,\mathrm{d}x \;\text{ subject to } u \cdot \tilde{n} \geq 0 \text{ on } \Gamma_T.$$

Problems of interest

Examples

- (Obstacle problem.) Find $u : \Omega \to \mathbb{R}$,

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \quad \text{subject to } u(x) \leq \varphi(x).$$

- (Elastic-plastic torsion.) Find $u : \Omega \to \mathbb{R}$,

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \quad \text{subject to } |\nabla u|(x) \leq \varphi(x).$$

- (Signorini.) Find $u : \Omega \to \mathbb{R}^d$,

$$\min_{u \in H_g^1(\Omega)^d} \int_\Omega \frac{1}{2}(\mathbf{C}\varepsilon(u)) : \varepsilon(u) - f \cdot u \,\mathrm{d}x \quad \text{subject to } u \cdot \tilde{n} \geq 0 \text{ on } \Gamma_T.$$

Latent variable proximal point (LVPP) blueprint

LVPP is a new and powerful framework for solving variational problems with pointwise constraints.

Variational problem with inequality constraints

Apply LVPP: sequence of nonlinear systems of PDEs

Discretize: sequence of nonlinear systems of equations

Newton solver

Latent variable proximal point (LVPP) blueprint

LVPP is a new and powerful framework for solving variational problems with pointwise constraints.

- Variational problem with inequality constraints
  - Apply LVPP: sequence of nonlinear systems of PDEs
    - Discretize: sequence of nonlinear systems of equations
      - Newton solver

Deriving the LVPP algorithm

$$\min_{u \in U} J(u) \text{ subject to } u \in K := \{v : (Bv)(x) \in C(x) \text{ for a.e. } x \in \Omega\}.$$

Bregman proximal point

First regularize the optimization problem via a *Bregman divergence*:

$$\min_{u \in U} J(u) + \frac{1}{\alpha} \int_{\Omega_d} R(Bu) - R(Bu^{k-1}) - \nabla R(Bu^{k-1})(Bu - Bu^{k-1}) \, \mathrm{d}\mathcal{H}_d \quad \text{(BD)}$$

The (classical) Bregman proximal point algorithm seeks $u^k \in K$ satisfying the *smooth* PDE:

$$\alpha_k \langle J'(u^k), v\rangle + \langle \nabla R(Bu^k) - \nabla R(Bu^{k-1}), Bv\rangle = 0 \quad \forall v \in U. \quad \text{(BPP)}$$

Numerically solving (BPP) via Newton's method performs poorly.

Deriving the LVPP algorithm

$$\min_{u \in U} J(u) \ \text{ subject to } \ u \in K := \{v : (Bv)(x) \in C(x) \text{ for a.e. } x \in \Omega\}.$$

Bregman proximal point

First regularize the optimization problem via a *Bregman divergence*:

$$\min_{u \in U} J(u) + \frac{1}{\alpha} \int_{\Omega_d} R(Bu) - R(Bu^{k-1}) - \nabla R(Bu^{k-1})(Bu - Bu^{k-1}) \, \mathrm{d}\mathcal{H}_d \quad \text{(BD)}$$

The (classical) Bregman proximal point algorithm seeks $u^k \in K$ satisfying the *smooth* PDE:

$$\alpha_k \langle J'(u^k), v \rangle + \langle \nabla R(Bu^k) - \nabla R(Bu^{k-1}), Bv \rangle = 0 \ \ \forall v \in U. \quad \text{(BPP)}$$

Numerically solving (BPP) via Newton's method performs poorly.

Leibniz Association

Deriving the LVPP algorithm

$$\min_{u \in U} J(u) \;\text{ subject to }\; u \in K := \{v : (Bv)(x) \in C(x) \text{ for a.e. } x \in \Omega\}.$$

Bregman proximal point

First regularize the optimization problem via a *Bregman divergence*:

$$\min_{u \in U} J(u) + \frac{1}{\alpha} \int_{\Omega_d} R(Bu) - R(Bu^{k-1}) - \nabla R(Bu^{k-1})(Bu - Bu^{k-1}) \, \mathrm{d}\mathcal{H}_d \quad \text{(BD)}$$

The (classical) Bregman proximal point algorithm seeks $u^k \in K$ satisfying the *smooth* PDE:

$$\alpha_k \langle J'(u^k), v \rangle + \langle \nabla R(Bu^k) - \nabla R(Bu^{k-1}), Bv \rangle = 0 \;\; \forall v \in U. \qquad \text{(BPP)}$$

Numerically solving (BPP) via Newton's method performs poorly.

Deriving the LVPP algorithm

$$\min_{u\in U} J(u) \;\text{ subject to }\; u \in K := \{v : (Bv)(x) \in C(x) \text{ for a.e. } x \in \Omega\}.$$

Bregman proximal point

First regularize the optimization problem via a *Bregman divergence*:

$$\min_{u\in U} J(u) + \frac{1}{\alpha}\int_{\Omega_d} R(Bu) - R(Bu^{k-1}) - \nabla R(Bu^{k-1})(Bu - Bu^{k-1})\,\mathrm{d}\mathcal{H}_d \quad \text{(BD)}$$

The (classical) Bregman proximal point algorithm seeks $u^k \in K$ satisfying the *smooth* PDE:

$$\alpha_k\langle J'(u^k), v\rangle + \langle \nabla R(Bu^k) - \nabla R(Bu^{k-1}), Bv\rangle = 0 \;\; \forall v \in U. \qquad \text{(BPP)}$$

Numerically solving (BPP) via Newton's method performs poorly.

The LVPP algorithm

Introduce a latent variable $\psi = \nabla R(Bu)$ and reformulate the primal equation (BPP) as a saddle point system.

The LVPP subproblem

Given ψ^{k-1}, for $k = 1, 2, \ldots,$ we seek (u^k, ψ^k) satisfying

$$\alpha_k \langle J'(u^k), v\rangle + \langle \psi^k, Bv\rangle = \langle \psi^{k-1}, Bv\rangle \quad \forall v \in U,$$
$$Bu^k - (\nabla R)^{-1}(\psi^k) = 0 \text{ a.e.},$$

- Pick proximal parameters α_k such that $\sum_{j=1}^{k} \alpha_j \to \infty$.
- Pick pointwise operator $(\nabla R)^{-1}$ such that $\nabla R(Bu)(x) \to \infty$ as $Bu(x) \to \partial C(x)$.

Generates two distinct approximations for Bu: Bu^k and $(\nabla R)^{-1}(\psi^k)$ (always feasible even after discretization).

The LVPP algorithm

Introduce a latent variable $\psi = \nabla R(Bu)$ and reformulate the primal equation (BPP) as a saddle point system.

The LVPP subproblem

Given ψ^{k-1}, for $k = 1, 2, \dots$, we seek (u^k, ψ^k) satisfying

$$\alpha_k \langle J'(u^k), v\rangle + \langle \psi^k, Bv\rangle = \langle \psi^{k-1}, Bv\rangle \quad \forall v \in U,$$
$$Bu^k - (\nabla R)^{-1}(\psi^k) = 0 \text{ a.e.},$$

- Pick proximal parameters α_k such that $\sum_{j=1}^{k} \alpha_j \to \infty$.
- Pick pointwise operator $(\nabla R)^{-1}$ such that $\nabla R(Bu)(x) \to \infty$ as $Bu(x) \to \partial C(x)$.

Generates two distinct approximations for Bu: Bu^k and $(\nabla R)^{-1}(\psi^k)$ (always feasible even after discretization).

The LVPP algorithm

Introduce a latent variable $\psi = \nabla R(Bu)$ and reformulate the primal equation (BPP) as a saddle point system.

The LVPP subproblem

Given ψ^{k-1}, for $k = 1, 2, \dots$, we seek (u^k, ψ^k) satisfying

$$\alpha_k \langle J'(u^k), v\rangle + \langle \psi^k, Bv\rangle = \langle \psi^{k-1}, Bv\rangle \quad \forall v \in U,$$
$$Bu^k - (\nabla R)^{-1}(\psi^k) = 0 \text{ a.e.},$$

- Pick proximal parameters α_k such that $\sum_{j=1}^{k} \alpha_j \to \infty$.
- Pick pointwise operator $(\nabla R)^{-1}$ such that $\nabla R(Bu)(x) \to \infty$ as $Bu(x) \to \partial C(x)$.

Generates two distinct approximations for Bu: Bu^k and $(\nabla R)^{-1}(\psi^k)$ (always feasible even after discretization).

Choices for $(\nabla R)^{-1}$

Feasible set K	B	$(\nabla R)^{-1}(\psi)$
$\{u \geq \phi\}$	id	$\phi + \exp\psi$
$\{\phi_1 \leq u \leq \phi_2\}$	id	$\dfrac{\phi_1 + \phi_2 \exp\psi}{1 + \exp\psi}$
$\{\operatorname{tr} u \geq \phi\}$	tr	$\phi + \exp\psi$
$\{(\operatorname{tr} u) \cdot n \leq \phi\}$	$\operatorname{tr}(\cdot) \cdot n$	$\phi - \exp(-\psi)$
$\{\lvert\nabla u\rvert \leq \phi\}$	∇	$\dfrac{\phi\psi}{\sqrt{1 + \lvert\psi\rvert^2}}$
$\{u \geq 0,\ \sum_i u_i = 1\}$	id	$\dfrac{\exp\psi}{\sum_i \exp\psi_i}$
$\{\det(\nabla^2 u) \geq 0\}$	∇^2	$\exp\psi$

LVPP for the obstacle problem

Obstacle problem: weak formulation of LVPP

$U = H_0^1(\Omega)$, $B = B^* = \mathrm{id}$, $J' = -\Delta - f$, and $(\nabla R)^{-1}(\psi) = \varphi - \mathrm{e}^{-\psi}$.

Given $\psi^{k-1} \in L^\infty(\Omega)$, for $k = 1, 2, \ldots$, we seek (u^k, ψ^k) satisfying, for all $(v, q) \in H_0^1(\Omega) \times L^\infty(\Omega)$,

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v),$$

$$(u^k, q) + (\mathrm{e}^{-\psi^k}, q) = (\varphi, q).$$

Theorem (B. Keith, T. Surowiec, FoCM, 2024)

Suppose that Ω is an open, bounded and Lipschitz domain, $f \in L^\infty(\Omega)$ and $\varphi \in \{\phi \in H^1(\Omega) \cap C(\bar{\Omega}) : \Delta\phi \in L^\infty(\Omega)\}$, then

$$\|u^* - u^k\|_{H^1(\Omega)} \lesssim \left(\sum_{j=1}^{k} \alpha_j\right)^{-1/2}.$$

Note that $u^k \to u^*$ in $H^1(\Omega)$ even if $\alpha_k = 1$ for all $k \in \mathbb{N}$.

LVPP for the obstacle problem

Obstacle problem: weak formulation of LVPP

$U = H_0^1(\Omega)$, $B = B^* = \mathrm{id}$, $J' = -\Delta - f$, and $(\nabla R)^{-1}(\psi) = \varphi - \mathrm{e}^{-\psi}$.
Given $\psi^{k-1} \in L^\infty(\Omega)$, for $k = 1, 2, \ldots$, we seek (u^k, ψ^k) satisfying, for all $(v, q) \in H_0^1(\Omega) \times L^\infty(\Omega)$,

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v),$$
$$(u^k, q) + (\mathrm{e}^{-\psi^k}, q) = (\varphi, q).$$

Theorem (B. Keith, T. Surowiec, FoCM, 2024)

Suppose that Ω is an open, bounded and Lipschitz domain, $f \in L^\infty(\Omega)$ and $\varphi \in \{\phi \in H^1(\Omega) \cap C(\bar{\Omega}) : \Delta\phi \in L^\infty(\Omega)\}$, then

$$\|u^* - u^k\|_{H^1(\Omega)} \lesssim \left(\sum_{j=1}^{k} \alpha_j\right)^{-1/2}.$$

Note that $u^k \to u^*$ in $H^1(\Omega)$ even if $\alpha_k = 1$ for all $k \in \mathbb{N}$.

LVPP for the obstacle problem

Obstacle problem: weak formulation of LVPP

$U = H_0^1(\Omega)$, $B = B^* = \mathrm{id}$, $J' = -\Delta - f$, and $(\nabla R)^{-1}(\psi) = \varphi - \mathrm{e}^{-\psi}$.
Given $\psi^{k-1} \in L^\infty(\Omega)$, for $k = 1, 2, \ldots$, we seek (u^k, ψ^k) satisfying, for all $(v, q) \in H_0^1(\Omega) \times L^\infty(\Omega)$,

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v),$$
$$(u^k, q) + (\mathrm{e}^{-\psi^k}, q) = (\varphi, q).$$

Theorem (B. Keith, T. Surowiec, FoCM, 2024)

Suppose that Ω is an open, bounded and Lipschitz domain, $f \in L^\infty(\Omega)$ and $\varphi \in \{\phi \in H^1(\Omega) \cap C(\bar{\Omega}) : \Delta\phi \in L^\infty(\Omega)\}$, then

$$\|u^* - u^k\|_{H^1(\Omega)} \lesssim \left(\sum_{j=1}^{k} \alpha_j\right)^{-1/2}.$$

Note that $u^k \to u^*$ in $H^1(\Omega)$ even if $\alpha_k = 1$ for all $k \in \mathbb{N}$.

LVPP for the obstacle problem

Obstacle problem: weak formulation of LVPP

$U = H_0^1(\Omega)$, $B = B^* = \mathrm{id}$, $J' = -\Delta - f$, and $(\nabla R)^{-1}(\psi) = \varphi - \mathrm{e}^{-\psi}$.
Given $\psi^{k-1} \in L^\infty(\Omega)$, for $k = 1, 2, \ldots$, we seek (u^k, ψ^k) satisfying, for all $(v, q) \in H_0^1(\Omega) \times L^\infty(\Omega)$,

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v),$$
$$(u^k, q) + (\mathrm{e}^{-\psi^k}, q) = (\varphi, q).$$

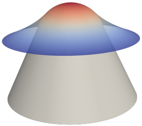
Theorem (B. Keith, T. Surowiec, FoCM, 2024)

Suppose that Ω is an open, bounded and Lipschitz domain, $f \in L^\infty(\Omega)$ and $\varphi \in \{\phi \in H^1(\Omega) \cap C(\bar{\Omega}) : \Delta\phi \in L^\infty(\Omega)\}$, then

$$\|u^* - u^k\|_{H^1(\Omega)} \lesssim \left(\sum_{j=1}^{k} \alpha_j\right)^{-1/2}.$$

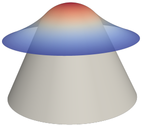
Note that $u^k \to u^*$ in $H^1(\Omega)$ even if $\alpha_k = 1$ for all $k \in \mathbb{N}$.

Examples

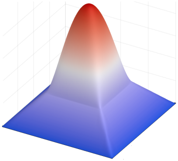


Obstacle, $u \leq \varphi$.

Examples

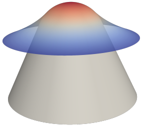


Obstacle, $u \leq \varphi$.

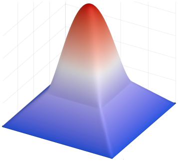


Gradient-type, $|\nabla u| \leq \varphi$.

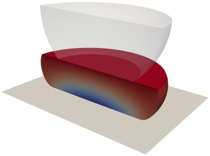
Examples



Obstacle, $u \leq \varphi$.

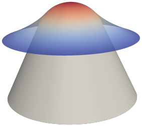


Gradient-type, $|\nabla u| \leq \varphi$.

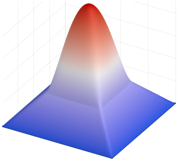


Signorini, $(u \cdot n)|_\Gamma \geq 0$.

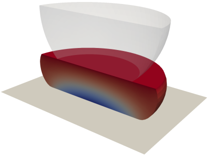
Examples



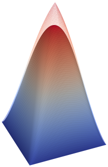
Obstacle, $u \leq \varphi$.



Gradient-type, $|\nabla u| \leq \varphi$.

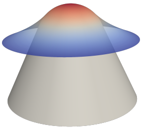


Signorini, $(u \cdot n)|_\Gamma \geq 0$.

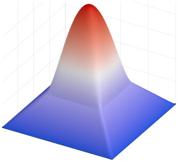


Thermoforming, $u \leq \varphi(u)$.

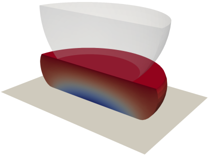
Examples



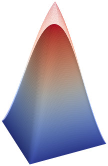
Obstacle, $u \leq \varphi$.



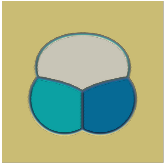
Gradient-type, $|\nabla u| \leq \varphi$.



Signorini, $(u \cdot n)|_\Gamma \geq 0$.

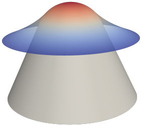


Thermoforming, $u \leq \varphi(u)$.

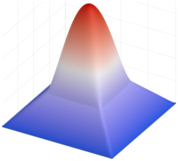


Cahn–Hilliard,
$u_i \geq 0$, $\sum_i u_i = 1$.

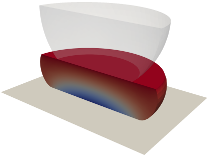
Examples



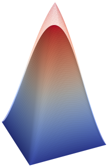
Obstacle, $u \leq \varphi$.



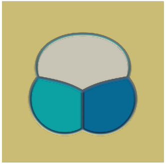
Gradient-type, $|\nabla u| \leq \varphi$.



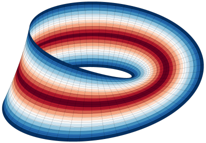
Signorini, $(u \cdot n)|_\Gamma \geq 0$.



Thermoforming, $u \leq \varphi(u)$.

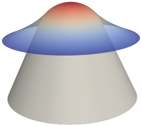


Cahn–Hilliard,
$u_i \geq 0, \sum_i u_i = 1$.

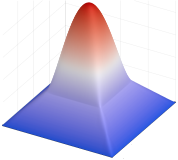


Eikonal, $|\nabla u| = 1$.

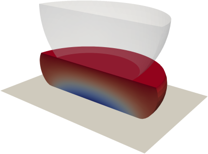
Examples



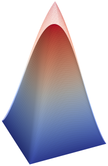
Obstacle, $u \leq \varphi$.



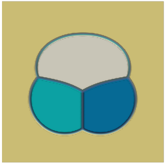
Gradient-type, $|\nabla u| \leq \varphi$.



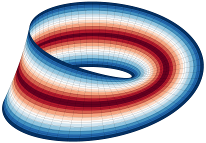
Signorini, $(u \cdot n)|_\Gamma \geq 0$.



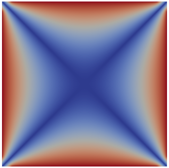
Thermoforming, $u \leq \varphi(u)$.



Cahn–Hilliard,
$u_i \geq 0$, $\sum_i u_i = 1$.

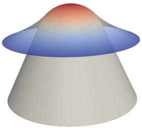


Eikonal, $|\nabla u| = 1$.

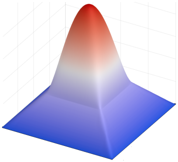


Nematic liquid crystals,
$|\lambda_i(Q)| \leq 1/2$.

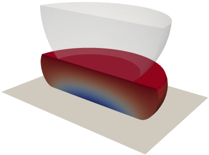
Examples



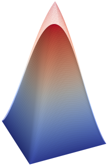
Obstacle, $u \leq \varphi$.



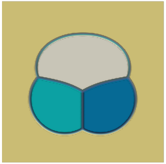
Gradient-type, $|\nabla u| \leq \varphi$.



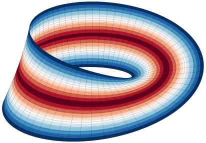
Signorini, $(u \cdot n)|_\Gamma \geq 0$.



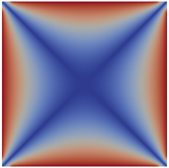
Thermoforming, $u \leq \varphi(u)$.



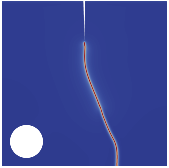
Cahn–Hilliard,
$u_i \geq 0$, $\sum_i u_i = 1$.



Eikonal, $|\nabla u| = 1$.



Nematic liquid crystals,
$|\lambda_i(Q)| \leq 1/2$.



Fracture.

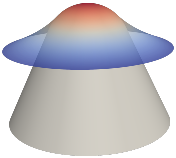
Contact problems

Contact problems

Obstacle problem solver comparisons

	Degree $p=1$			Degree $p=2$		
Method	h	$h/2$	$h/4$	h	$h/2$	$h/4$
LVPP	15	13	12	15	16	12
Active Set (PETSc)	11	16	25	Not bound preserving		
Trust-Region (Galahad)	6	12	19			
Interior Point (IPOPT)	9	9	8			
IPOPT without Hessian	90	260	500			

(a) Number of linear system solves for popular solvers using various mesh sizes h.

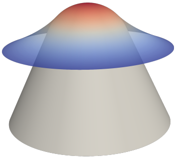


(b) Obstacle ϕ (grey) and membrane u (red/blue).

Obstacle problem solver comparisons

	Degree $p = 1$			Degree $p = 2$		
Method	h	$h/2$	$h/4$	h	$h/2$	$h/4$
LVPP	15	13	12	15	16	12
Active Set (PETSc)	11	16	25	Not bound preserving		
Trust-Region (Galahad)	6	12	19			
Interior Point (IPOPT)	9	9	8			
IPOPT without Hessian	90	260	500			

(a) Number of linear system solves for popular solvers using various mesh sizes h.



(b) Obstacle ϕ (grey) and membrane u (red/blue).

Mesh size h	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
Finite Difference	10	15	13	15	16	16
Degree p	8	16	24	32	40	48
Spectral Method	16	17	16	16	16	15

(c) Number of linear system solves for the proximal finite difference and spectral methods.

Thermoforming solver comparison

The thermoforming quasi-variational inequality seeks $u : \Omega \to \mathbb{R}$ minimizing

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \text{ subject to } u \leq \varphi(T) := \Phi_0 + \xi T, \tag{1a}$$

where Φ_0 and ξ are given and T satisfies

$$-\Delta T + \beta T = g(\Phi_0 + \xi T - u), \quad \partial_\nu T = 0 \text{ on } \partial\Omega. \tag{1b}$$

LVPP subproblem

Given ψ^{k-1}, we seek (u^k, T^k, ψ^k) satisfying for all $(v, q, w) \in H_0^1(\Omega) \times L^\infty(\Omega) \times H^1(\Omega)$

$$(\nabla T^k, \nabla q) + \beta(T^k, q) = (g(\mathrm{e}^{-\psi^k}), q), \tag{2a}$$

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v), \tag{2b}$$

$$(u^k, w) + (\mathrm{e}^{-\psi^k}, w) = (\Phi_0 + \xi T^k, w). \tag{2c}$$

Thermoforming solver comparison

The thermoforming quasi-variational inequality seeks $u : \Omega \to \mathbb{R}$ minimizing

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \text{ subject to } u \le \varphi(T) := \Phi_0 + \xi T, \tag{1a}$$

where Φ_0 and ξ are given and T satisfies

$$-\Delta T + \beta T = g(\Phi_0 + \xi T - u), \quad \partial_\nu T = 0 \text{ on } \partial\Omega. \tag{1b}$$

LVPP subproblem

Given ψ^{k-1}, we seek (u^k, T^k, ψ^k) satisfying for all $(v, q, w) \in H_0^1(\Omega) \times L^\infty(\Omega) \times H^1(\Omega)$

$$(\nabla T^k, \nabla q) + \beta(T^k, q) = (g(\mathrm{e}^{-\psi^k}), q), \tag{2a}$$

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v), \tag{2b}$$

$$(u^k, w) + (\mathrm{e}^{-\psi^k}, w) = (\Phi_0 + \xi T^k, w). \tag{2c}$$

Thermoforming solver comparison

The thermoforming quasi-variational inequality seeks $u : \Omega \to \mathbb{R}$ minimizing

$$\min_{u \in H_0^1(\Omega)} \int_\Omega \frac{1}{2}|\nabla u|^2 - fu \,\mathrm{d}x \text{ subject to } u \le \varphi(T) := \Phi_0 + \xi T, \tag{1a}$$

where Φ_0 and ξ are given and T satisfies

$$-\Delta T + \beta T = g(\Phi_0 + \xi T - u), \quad \partial_\nu T = 0 \text{ on } \partial\Omega. \tag{1b}$$

LVPP subproblem

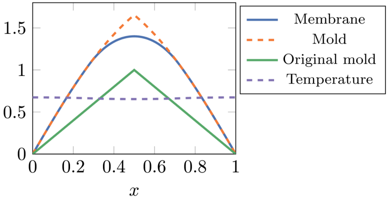
Given ψ^{k-1}, we seek (u^k, T^k, ψ^k) satisfying for all $(v, q, w) \in H_0^1(\Omega) \times L^\infty(\Omega) \times H^1(\Omega)$

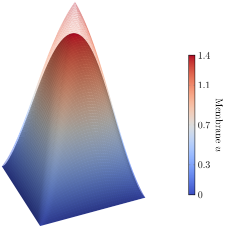
$$(\nabla T^k, \nabla q) + \beta(T^k, q) = (g(\mathrm{e}^{-\psi^k}), q), \tag{2a}$$

$$\alpha_k(\nabla u^k, \nabla v) + (\psi^k, v) = \alpha_k(f, v) + (\psi^{k-1}, v), \tag{2b}$$

$$(u^k, w) + (\mathrm{e}^{-\psi^k}, w) = (\Phi_0 + \xi T^k, w). \tag{2c}$$

Thermoforming solver comparison





Solver	Outer loop	Linear system solves	Run time (s)
LVPP	13	20	61.70
Moreau–Yosida Penalty	14	51	78.01
Semismooth Active Set	7	236	112.60
Fixed Point	164	8493	3633.72

The performance of four solvers, terminating when $\|u^k - u^{k-1}\|_{H^1(\Omega)} \leq 10^{-5}$.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, `https://arxiv.org/abs/2503.05672`.
`https://github.com/METHODS-Group/ProximalGalerkin`

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., `https://arxiv.org/abs/2412.13733`.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, `https://arxiv.org/abs/2503.05672`.
`https://github.com/METHODS-Group/ProximalGalerkin`

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., `https://arxiv.org/abs/2412.13733`.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, https://arxiv.org/abs/2503.05672.

https://github.com/METHODS-Group/ProximalGalerkin

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., https://arxiv.org/abs/2412.13733.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, https://arxiv.org/abs/2503.05672.
https://github.com/METHODS-Group/ProximalGalerkin

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., https://arxiv.org/abs/2412.13733.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, https://arxiv.org/abs/2503.05672.
https://github.com/METHODS-Group/ProximalGalerkin

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., https://arxiv.org/abs/2412.13733.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, https://arxiv.org/abs/2503.05672.

https://github.com/METHODS-Group/ProximalGalerkin

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., https://arxiv.org/abs/2412.13733.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, `https://arxiv.org/abs/2503.05672`.

`https://github.com/METHODS-Group/ProximalGalerkin`

Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., `https://arxiv.org/abs/2412.13733`.

Conclusions

- Many pointwise constraints can be effectively handled by LVPP resulting in a nonlinear system of smooth PDEs.
- LVPP is discretization agnostic.
- Observed discretization-independent number of linear system solves.
- LVPP has a simple mechanism for enforcing pointwise constraints on the discrete level (without the need for a projection).
- Ease of implementation — the algorithm reduces to the repeated solve of a smooth nonlinear system of PDEs *without requiring specialized discretizations*.
- Robust numerical performance since convergence occurs even α_k is kept small.

The latent variable proximal point algorithm for variational problems with inequality constraints, *to appear in CMAME* (2025)

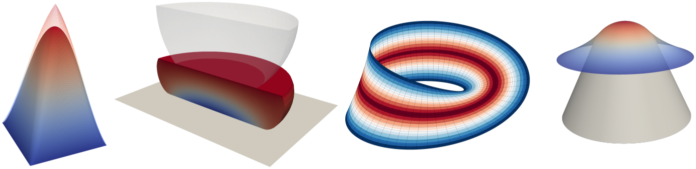
J. Dokken, P. Farrell, B. Keith, I. P., T. Surowiec, https://arxiv.org/abs/2503.05672.
https://github.com/METHODS-Group/ProximalGalerkin

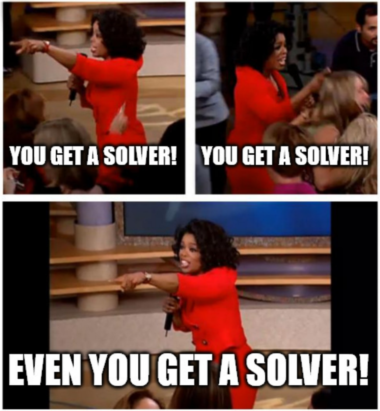
Hierarchical proximal Galerkin: a fast *hp*-FEM solver for variational problems with pointwise inequality constraints (2024)

I. P., https://arxiv.org/abs/2412.13733.

Thank you for listening!

✉ papadopoulos@wias-berlin.de



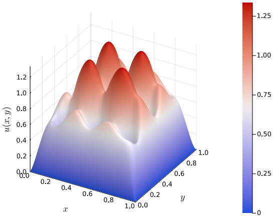

YOU GET A SOLVER!
YOU GET A SOLVER!
EVEN YOU GET A SOLVER!

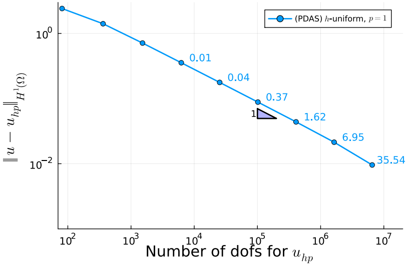

EVERYONE GETS A SOLVER!!!!

High-order FEM discretizations

Observations

1. LVPP is discretization agnostic $\to$ use sparsity-preserving high-order FEM.
2. After a Newton linearization & FEM discretization we are solving linear saddle point systems.
3. These admit block preconditioners with sparse Schur complement approximations.

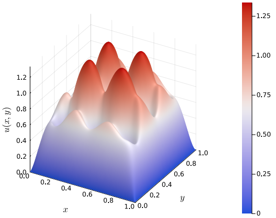


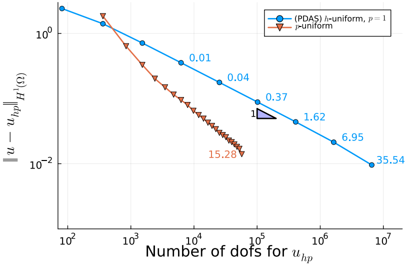


High-order FEM discretizations

Observations

1. LVPP is discretization agnostic $\rightarrow$ use sparsity-preserving high-order FEM.
2. After a Newton linearization & FEM discretization we are solving linear saddle point systems.
3. These admit block preconditioners with sparse Schur complement approximations.

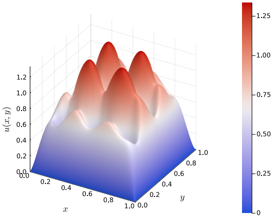


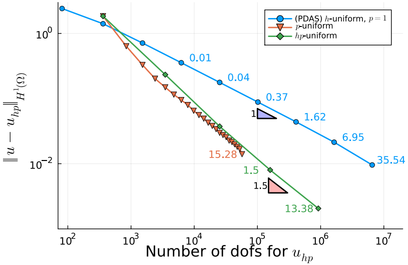


High-order FEM discretizations

Observations

1. LVPP is discretization agnostic $\rightarrow$ use sparsity-preserving high-order FEM.
2. After a Newton linearization & FEM discretization we are solving linear saddle point systems.
3. These admit block preconditioners with sparse Schur complement approximations.





High-order FEM (thermoforming)

		Obstacle subsolve for u		Nonlinear subsolve for T	
p	Fixed point	Avg. Newton	Avg. GMRES	Avg. Newton	Avg. GMRES
6	4	15.00	11.00	1.50	2.83
12	4	15.25	15.85	2.00	3.13
22	4	16.00	19.36	2.00	3.00
32	4	16.00	21.09	2.00	3.00
42	4	15.75	21.75	2.25	3.11
52	4	15.00	22.40	2.00	3.00
62	4	15.00	21.90	2.00	3.00
72	4	15.00	21.90	2.00	3.00
82	4	15.25	21.61	2.00	3.00

p-independent Newton and preconditioned GMRES iteration counts to solve the thermoforming problem. Unbelievable!

High-order FEM (thermoforming)

		Obstacle subsolve for u		Nonlinear subsolve for T	
p	Fixed point	Avg. Newton	Avg. GMRES	Avg. Newton	Avg. GMRES
6	4	15.00	11.00	1.50	2.83
12	4	15.25	15.85	2.00	3.13
22	4	16.00	19.36	2.00	3.00
32	4	16.00	21.09	2.00	3.00
42	4	15.75	21.75	2.25	3.11
52	4	15.00	22.40	2.00	3.00
62	4	15.00	21.90	2.00	3.00
72	4	15.00	21.90	2.00	3.00
82	4	15.25	21.61	2.00	3.00

Partial degree

p-independent Newton and preconditioned GMRES iteration counts to solve the thermoforming problem. Unbelievable!

High-order FEM (thermoforming)

		Obstacle subsolve for u		Nonlinear subsolve for T	
p	Fixed point	Avg. Newton	Avg. GMRES	Avg. Newton	Avg. GMRES
6	4	15.00	11.00	1.50	2.83
12	4	15.25	15.85	2.00	3.13
22	4	16.00	19.36	2.00	3.00
32	4	16.00	21.09	2.00	3.00
42	4	15.75	21.75	2.25	3.11
52	4	15.00	22.40	2.00	3.00
62	4	15.00	21.90	2.00	3.00
72	4	15.00	21.90	2.00	3.00
82	4	15.25	21.61	2.00	3.00

Partial degree

Outer loop

p-independent Newton and preconditioned GMRES iteration counts to solve the thermoforming problem. Unbelievable!

High-order FEM (thermoforming)

		Obstacle subsolve for u		Nonlinear subsolve for T	
p	Fixed point	Avg. Newton	Avg. GMRES	Avg. Newton	Avg. GMRES
6	4	15.00	11.00	1.50	2.83
12	4	15.25	15.85	2.00	3.13
22	4	16.00	19.36	2.00	3.00
32	4	16.00	21.09	2.00	3.00
42	4	15.75	21.75	2.25	3.11
52	4	15.00	22.40	2.00	3.00
62	4	15.00	21.90	2.00	3.00
72	4	15.00	21.90	2.00	3.00
82	4	15.25	21.61	2.00	3.00

Partial degree

Outer loop

Average Newton steps to solve an obstacle subproblem

Average preconditioned GMRES iterations per Newton step

p-independent Newton and preconditioned GMRES iteration counts to solve the thermoforming problem. <u>Unbelievable!</u>

High-order FEM (thermoforming)

		Obstacle subsolve for u		Nonlinear subsolve for T	
p	Fixed point	Avg. Newton	Avg. GMRES	Avg. Newton	Avg. GMRES
6	4	15.00	11.00	1.50	2.83
12	4	15.25	15.85	2.00	3.13
22	4	16.00	19.36	2.00	3.00
32	4	16.00	21.09	2.00	3.00
42	4	15.75	21.75	2.25	3.11
52	4	15.00	22.40	2.00	3.00
62	4	15.00	21.90	2.00	3.00
72	4	15.00	21.90	2.00	3.00
82	4	15.25	21.61	2.00	3.00

Partial degree

Outer loop

Average Newton steps to solve an obstacle subproblem

Average preconditioned GMRES iterations per Newton step

Average Newton steps to solve a temperature PDE subproblem

Average preconditioned GMRES iterations per Newton step

p-independent Newton and preconditioned GMRES iteration counts to solve the thermoforming problem. Unbelievable!