Computational Unique Continuation

Erik Burman[1]

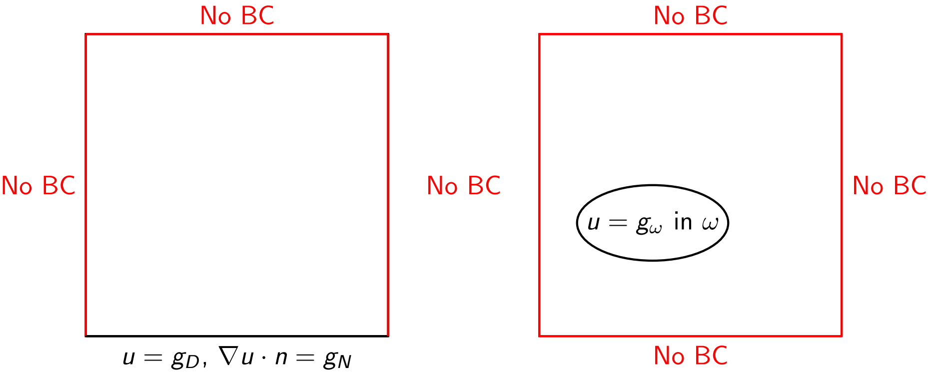
Department of Mathematics
University College London

Control of PDEs and related topics, Toulouse, 2025

[1] Joint work with Ali Feizmohammadi, Arnaud Münch, Mihai Nechita, Lauri Oksanen, Janosch Preuss, Mingfei Lu

Unique continuation

Find $u \in H^1(\Omega)$ such that $\Delta u = 0$ in Ω



- Left: continuation from the boundary: the elliptic Cauchy problem;
- Right: continuation from some subdomain $\omega \subset \Omega$;
- The data g_D, g_N and g_ω must be compatible. Uniqueness.

Ill-posed problems: Hadamard's example[2]

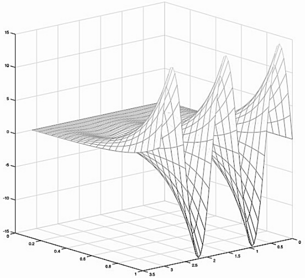
- The elliptic Cauchy problem

$$\begin{aligned} -\triangle u &= 0 \text{ in } [0,1]^2 \\ u(x,0) &= 0 \\ \frac{\partial u}{\partial y}(x,0) &= g. \end{aligned}$$

- $g = N^{-1}\sin(Nx)$ gives the solution

$$u_N(x,y) = \frac{1}{N^2}\sin(Nx)\sinh(Ny)$$

- For $N \to \infty$, $g \to 0$, but $u_N \to \infty$ for all most every (x,y) with $y > 0$.
- Additional boundedness assumption on u gives conditional stability



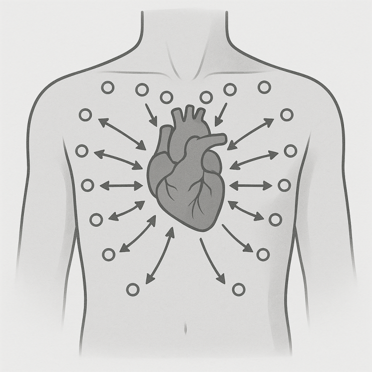
Example of Hadamard solution u_N

[2]Hadamard 1902

Application: Inverse problem of electrocardiography[3]

Given the potential and the current on the torso recover the potential ϕ_H on the heart.

- Equation: $\triangle\phi = 0$.
- Data on torso: $\phi = \phi_T$, $\nabla\phi \cdot n = 0$.
- Cauchy problem $\Rightarrow$ Unique continuation.



[3] Bartholomay, 1969; Colli-Magenes, 1980

Application: Acoustic Noise Cancellation[4]

- Wave Equation

$$\partial_{tt} u - \Delta u = 0 \quad \text{in } (0,T) \times \Omega,$$

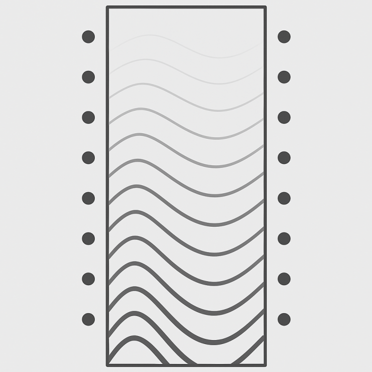
$$\nabla u \cdot n|_{\partial\Omega} = g,$$

$$u(0,x) = u_0(x),\, \partial_t u(0,x) = u_1(x).$$

- Find a control $g(t,x)$ such that:

$$u(T,x) = 0, \quad \partial_t u(T,x) = 0 \quad \text{in } \Omega.$$

- Under the Geometric Control Condition, null-controllability holds and g exists.
- Unique continuation on the adjoint equation.



[4] Glowinski and Lions 1995; Ervedoza and Zuazua 2013

Unique continuation, elliptic equation[5]

Consider three open, connected and non-empty sets $\omega \subset B \subset \Omega$ in $\mathbb{R}^n$.

Problem 1. Given $u|_\omega$ determine $u|_B$ for u satisfying the elliptic equation

$$\mathcal{P}u := \nabla \cdot \mu \nabla u + \beta \cdot \nabla u + \rho u = 0 \text{ in } \Omega.$$

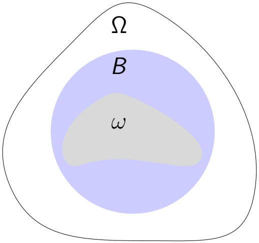
μ, ρ in $C^1(\bar{\Omega})$, $\mu > 0$, $\beta \in [C^1(\bar{\Omega})]^d$.

- B does not touch the boundary of Ω, problem conditionally ($\star$) Hölder stable [John'60]: There are $C > 0$ and $\alpha \in (0,1)$ such that (for all $u \in H^1(\Omega)$),

$$\|u\|_{H^1(B)} \leq \overbrace{C\|u\|^{1-\alpha}_{H^1(\Omega)}}^{(*)} \|u\|^\alpha_{UC}$$

  where $\|u\|_{UC} := \|u\|_{L^2(\omega)} + \|\mathcal{P}u\|_{H^{-1}(\Omega)}$.
- B touches the boundary, stability logarithmic:

$$\|u\|_{H^1(B)} \leq C\|u\|_{H^2(\Omega)} \log(\|u\|_{UC})^{-\alpha}$$

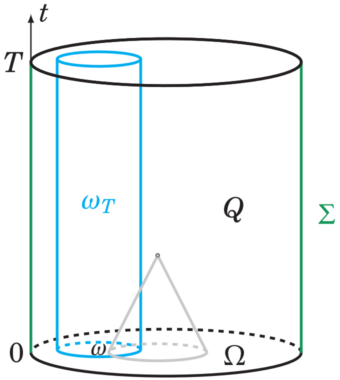


[5] Alessandrini et al. 2009; E.B., Nechita and Oksanen, 2020; E.B., Oksanen and Lu, 2024

Unique continuation, wave equation in $Q = (0, T) \times \Omega$

Problem 2. Given $u(t,x)|_{\omega_T} := u_\omega$ determine u solution to

$$\Box u := \partial_t^2 u - \triangle u = 0, \ (t,x) \in Q$$



Can we reconstruct u in all of Q?

Yes, we can if ...

- $\omega_T \subset Q$ fulfills the geometric control condition,
- $u|_\Sigma$ is given as data (say $u|_\Sigma = 0$).

Note: these are inherently space-time problems.

Lipschitz stability estimate

Thm.[E.B., Feizmohammadi, Münch & Oksanen, '23][a]

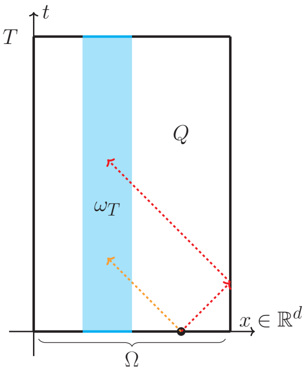
Assume that $\omega_T \subset Q$ fulfills the geometric control condition [Bardos, Lebeau & Rauch, '92]. Then for any $u \in H^1(Q)$

$$\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))} \lesssim \|u\|_{L^2(\omega_T)} + \|\Box u\|_{H^{-1}(Q)} + \|u\|_{L^2(\Sigma)}$$

I.e. for a solution u of wave eq. that vanishes on Σ:

$$\|u\|_{L^\infty(0,T;L^2(\Omega))} + \|\partial_t u\|_{L^2(0,T;H^{-1}(\Omega))} \lesssim \|u\|_{L^2(\omega_T)}.$$

Unknown bcs $\Rightarrow$ stability Hölder at best.



[a] Lasiecka et al. 1986; Le Rousseau et al. 2017

Traditional approach based on Tikhonov regularisation[7]

Problem 1. (simplified) Given $u|_\omega$ determine $u|_B$ for u satisfying $\Delta u + \rho u = 0$ in Ω.

Let $U \in H^1(\Omega)$ be an extension of the data $u|_\omega$ and set $f = (\Delta + \rho)U$. Replacing u by $u - U$ we get the equivalent problem:

Find $u \in D(\mathcal{P})$ such that $\mathcal{P}u = f$,

where $\mathcal{P} = \Delta + \rho$ and $D(\mathcal{P}) = \{u \in H^1(\Omega) : u = 0 \text{ in } \omega\}$.

Tikhonov regularised problem is obtained by minimising the functional

$$\mathcal{L}(u) = \underbrace{\|\mathcal{P}u - f\|^2}_{\text{data fitting}} + \underbrace{\epsilon\|u\|^2}_{\text{regularisation}}, \quad u \in D(\mathcal{P}).$$

Here $\|\cdot\|$ is the norm in $L^2(\Omega)$, but the norms should be chosen from the stability estimates! That typically means minimzation in dual norm.[6]

[6] Chung, Ito and Yamamoto, 2022; Dahmen, Monsuur and Stevenson, 2023.

[7] Tikhonov and Arsenin, -74

Quasireversibility method

The unique critical point of the functional

$$\mathcal{L}(u) = \|\mathcal{P}u - f\|^2 + \epsilon\|u\|^2, \quad u \in D(A),$$

is the solution to the following linear system for u.

$$(\mathcal{P}u, \mathcal{P}v) + \epsilon(u, v) = (f, \mathcal{P}v), \quad \forall v \in D(A). \tag{EL}$$

Here and below $(\cdot, \cdot)$ denotes the L^2-scalar product.

This system can be discretised, and a discrete approximation of u can be computed. This is the quasi-reversibility method [Lattès-Lions'67].

Quasireversibility method. First regularise, then discretise.

Stabilised finite element methods. First discretise, then regularise.

The system (EL) is a weak formulation of a 4th order equation. A mixed formulation of the quasi-reversibility method gives a system of 2nd order equations [Bourgeois'05].

Stabilised FEM: Continuum Lagrangian

Problem 1. Given $u|_\omega$ determine $u|_B$ for u satisfying $\mathcal{P}u = 0$ in Ω.

Notation:

- Write $g_\omega = u|_\omega$ for the data.
- Define the bilinear form $a(u,z) = (\mu \nabla u + \beta u, \nabla z) + (\rho u, z)$. Then the weak formulation of $\mathcal{P}u = 0$ reads

$$a(u,z) = 0, \quad \forall z \in H_0^1(\Omega).$$

We will use the following Lagrangian, with a Lagrange multiplier z,

$$\mathcal{L}(u,z) = \underbrace{\|u - g_\omega\|^2_{L^2(\omega)}}_{\text{data fitting}} + \underbrace{a(u,z)}_{\text{constraint}}, \quad u \in H^1(\Omega),\ z \in H_0^1(\Omega).$$

Unique continuation: the critical point is the unique minimiser

Stabilised FEM: Step 2. Discretisation

We will use a scale of finite element spaces to discretise the Lagrangian

$$\mathcal{L}(u, z) = \|u - g_\omega\|^2_{L^2(\omega)} + a(u, z), \quad u \in H^1(\Omega),\ z \in H^1_0(\Omega).$$

Suppose, for simplicity, that $\Omega \subset \mathbb{R}^2$ and that Ω is a polygon. Notation:

- $\mathcal{T}_h$ is a family of triangulations of Ω, indexed by the mesh size h, that is, the maximal diameter of the triangles in $\mathcal{T}_h$.
- V_h^p is the space of continuous, piecewise polynomials of degree $\leq p$ on $\mathcal{T}_h$.

Observe that V_h is a subspace of $H^1(\Omega)$.

The Lagrangian $\mathcal{L} : H^1(\Omega) \times H^1_0(\Omega) \to \mathbb{R}$ is discretised simply by restricting it on the subspace $V_h \times W_h$ where $W_h = V_h^m \cap H^1_0(\Omega)$ for some m to be determined.

Unique continuation: fails for finite elements, no uniqueness[8]

[8] Ben Belgacem, Jelassi and Girault, CMA, 2023; E.B., Oksanen and Ziyao, SINUM, 2025

Stabilised FEM: Step 3. Regularisation

- Discrete Lagrangian:

$$\mathcal{L}_h(u_h, z_h) = \|u_h - g_\omega\|^2_{L^2(\omega)} + a(u_h, z_h), \quad u_h \in V_h,\ z_h \in W_h.$$

- Add stabilising terms, s and s^* to improve stability.

$$\mathcal{L}_h(u_h, z_h) = \underbrace{\|u_h - g_\omega\|^2_{L^2(\omega)}}_{\text{data fitting}} + \underbrace{a(u_h, z_h)}_{\text{constraint}} + \underbrace{s(u_h, u_h) - s^*(z_h, z_h)}_{\text{regularisation}},\ u_h \in V_h,\ z_h \in W_h.$$

- Euler-Lagrange equations: find $u_h, z_h \in V_h \times W_h$ such that

$$\begin{aligned} a(u_h, w_h) - s^*(z_h, w_h) &= 0, \quad \forall w_h \in W_h &(1)\\ a(v_h, z_h) + s(u_h, v_h) + (u_h, v_h)_\omega &= (g_\omega, v_h)_\omega, \quad \forall v_h \in V_h. &(2) \end{aligned}$$

- Compact: find $u_h, z_h \in V_h \times W_h$ such that

$$\boxed{A_h[(u_h, z_h), (v_h, w_h)] = (q, v_h)_\omega \text{ for all } v_h, w_h \in V_h \times W_h}$$

$$A_h[(u_h, z_h), (v_h, w_h)] := a(u_h, w_h) - s^*(z_h, w_h) + a(v_h, z_h) + s(u_h, v_h) + (u_h, v_h)_\omega$$

Inf-sup stability of the form A_h I

- The problem is ill-posed so it can not be uniformly inf-sup stable in h in any norm on x_h.
- At best $\|\!|x_h, y_h|\!\|$ can be strong enough to bound the right hand side of the stability estimate, $\|x_h\|_\omega + \|\mathcal{P}x_h\|_{H^{-1}(\Omega)}$.
- Necessary inf-sup stability: there exists a positive constant c_s independent of h such that there holds

$$c_s \|\!|x_h, y_h|\!\| \leq \sup_{(v_h,w_h)\in V_h\times W_h} \frac{A_h[(x_h, y_h), (v_h, w_h)]}{\|\!|v_h, w_h|\!\|}$$

- Depends on the choices of s, s^*, V_h^p and W_h^m.

Inf-sup stability of the form A_h II

- $\mathcal{F}_h$ denotes interior faces of $\mathcal{T}_h$, $[\![\nabla v_h]\!]|_F$ the jump of ∇v_h over the face F.
- Example 1. Fix p and take $m = 1$.

$$s(u_h, v_h) = \sum_{K \in \mathcal{T}_h} (h^2 \mathcal{P} u_h, \mathcal{P} v_h)_K + \sum_{F \in \mathcal{F}_h} (h[\![\nabla u_h]\!], [\![\nabla v_h]\!])_F$$
$$s^*(z_h, w_h) = (\nabla z_h, \nabla w_h)_\Omega \text{ or } s^* = s + (h \nabla z_h \cdot n, \nabla w_h \cdot n)_{\partial\Omega}.$$

- Example 2[9]. Fix p, and take $m = p + d - 1$. If $\mu = 1$, $\beta = \rho = 0$,

$$s(u_h, v_h) = 0$$
$$\text{and } s^*(z_h, w_h) = (\nabla z_h, \nabla w_h)_\Omega$$
$$\text{or } s^* = \sum_K (h^2 \Delta z_h, \Delta w_h)_K + \sum_{F \in \mathcal{F}_h} (h[\![\nabla z_h]\!], [\![\nabla w_h]\!])_F + (h \nabla z_h \cdot n, \nabla w_h \cdot n)_{\partial\Omega}.$$

- In both cases $|\!|\!| v_h, w_h |\!|\!|^2 = \|v_h\|_\omega^2 + \|h \mathcal{P} v_h\|_{\mathcal{T}_h}^2 + \|h[\![\nabla v_h]\!]\|_{\mathcal{F}_h}^2 + s^*(w_h, w_h)$.
- For example 1: $|\!|\!| v_h, w_h |\!|\!|^2 = A_h[(v_h, w_h), (v_h, -w_h)]$, $\forall v_h, w_h \in V_h \times W_h$.

[9] E.B., Lu, arXiv:2409.20448, 2024

Stabilised FEM for the unique continuation problem[10]

- If the stability is conditional, $\|u\|_V \le C$, $s(u_h, v_h)$ must be augmented with

$$h^{2p}\|u\|^2_{H^1(\Omega)} \text{ to control } \underbrace{\|u\|^{1-\alpha}_{H^1(\Omega)}}_{\text{this factor}}(\|u\|_\omega + \|\mathcal{P}u\|_{H^{-1}(\Omega)}).$$

Lemma (Convergence of stabilising terms and data)

- *The inf-sup stable Lagrangian has a unique critical point* $(u_h, z_h) \in V_h \times W_h$.
- *This approximation satisfies the bound*

$$|||u - u_h, z_h||| + h^p\|u - u_h\|_V \le C\|u\|_{H^{p+1}(\Omega)}h^p.$$

- Uses that unique continuation holds in the sense that

$$\boxed{|||v_h, w_h||| = 0 \;\Rightarrow\; v_h = w_h \equiv 0}$$

The error estimate uses techniques of stabilised FEM (for well-posed problems).

[10]E.B. C.R.A.S. 2014, E.B., Nechita, Oksanen, JMPA, 2019, E.B., Oksanen, Lu, M2AN, 2025.

Convergence rate of $\|u - u_h\|_{H^1(B)}$ with noiseless data

Suppose that $u \in H^{p+1}(\Omega)$ where u is the solution of

$$\begin{cases} \mathcal{P}u = 0, & \text{in } \Omega, \\ u|_\omega = g_\omega, & \text{in } \omega. \end{cases}$$

Theorem Let $(u_h, z_h) \in V_h \times W_h$ be the unique critical point of the Lagrangian $\mathcal{L}_q$. Then

$$\|u - u_h\|_{H^1(B)} \lesssim h^{\alpha p} \|u\|_{H^{p+1}(\Omega)}$$

- By the Lemma $\|u_h\|_V \lesssim \|u\|_{H^{p+1}(\Omega)}$.
- By the definition of the dual norm and Galerkin orthogonality

$$\|\mathcal{P}(u - u_h)\|_{H^{-1}(\Omega)} \lesssim \|\!|u - u_h, z_h|\!\|.$$

- The claim then follows using the conditional stability estimate and the error bound on $\|\!|u - u_h, z_h|\!\|$.

Error estimate with noisy data

Suppose that $u \in H^{p+1}(\Omega)$ where u is the solution of

$$\begin{cases} \mathcal{P}u = 0, & \text{in } \Omega, \\ u|_\omega = g_\omega, & \text{in } \omega. \end{cases}$$

Consider the case that g_ω is known only up to an error $\delta g \in L^2(\omega)$. That is, we assume that $\tilde{g} = g_\omega + \delta g$ is known. Let $\delta = \|\delta g\|_{L^2(\omega)}$.

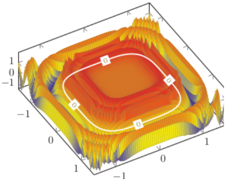
Theorem Let $(u_h, z_h) \in V_h \times W_h$ be the unique critical point of the Lagrangian $\mathcal{L}_h$. Then

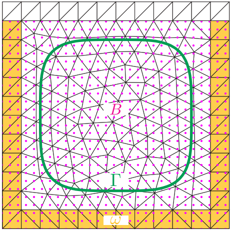
$$\|u - u_h\|_{H^1(B)} \lesssim h^{\alpha p}\|u\|_{H^{p+1}(\Omega)} + h^{-(1-\alpha)p}\delta.$$

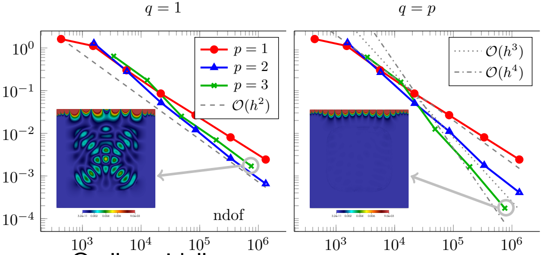
- Condition for error reduction: $(\delta/\|u\|_{H^{p+1}(\Omega)})^{1/p} \lesssim h$.
- When the fun stops, stop (or fix a lower bound of the regularization term).

Helmholtz problem, convex geometry[11], $\mu_2\rho_1/\mu_1\rho_2 = 128$

- Interface problem, jumping coefficients.
- Unfitted FEM, interface approximation q.
- Target domain B contained in convex hull of ω.
- Helmholtz problem with oscillatory solution.







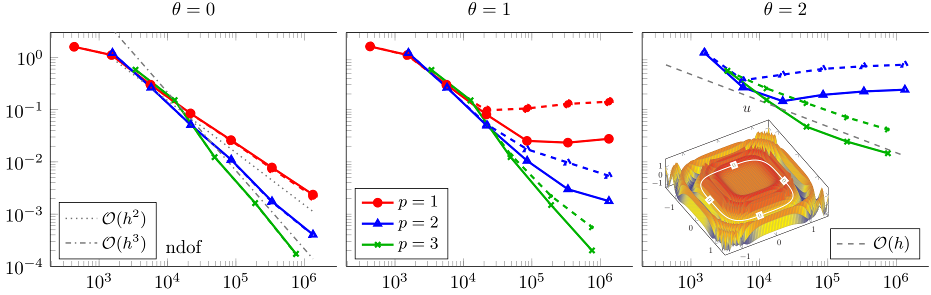
- Ω_1 "inside" green curve, Ω_2 "outside".
- For $q = p$ we observe: $\|u - u_h\|_{L^2(B)} \sim \mathcal{O}(h^{\alpha(q+1)})$ for $\alpha \approx 0.8$.

[11] E.B. and Preuss, SMAI, 2025

Data perturbations: $\delta = \tilde{\delta}_p h^{p-\theta}$ for $\theta \in \{0, 1, 2\}$, $q = p$.

$$\rightsquigarrow \|u - u_h \circ \Phi_h^{-1}\|_B \lesssim h^{\alpha p} \left(\|u\|_p + \tilde{\delta}_p h^{-\theta} \right)$$

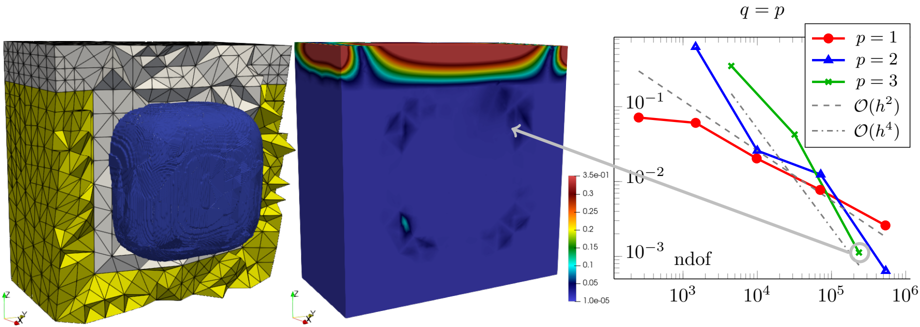
- For $\|u\|_p h^\theta > \tilde{\delta}_p$: converges as for exact data.
- For $\|u\|_p h^\theta < \tilde{\delta}_p$: rate deteriorates to $\mathcal{O}(h^{\alpha p - \theta})$.



Solid lines: $\tilde{\delta}_p \in [1, 8, 28]$ for $p \in [1, 2, 3]$, dashed lines $\tilde{\delta}_p \in [5, 24, 80]$.

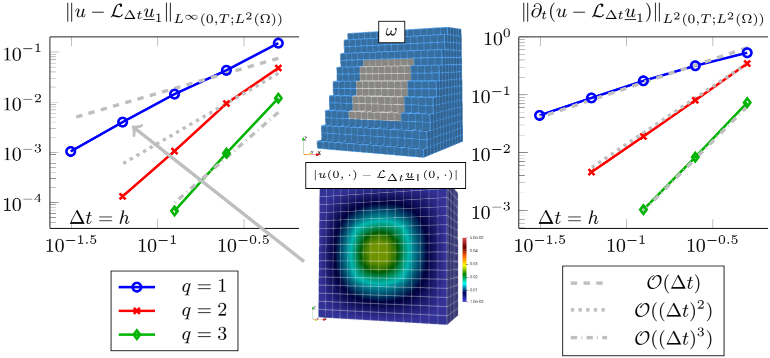
Purely elliptic problem in 3D

- We consider a pure diffusion ($\rho = 0$) interface problem.
- The interface is defined by the levelset function $\phi = \|x\|_4 - 1$.
- Unfitted finite element method, q order of approximation of interface.



- Left: mesh and configuration, ω yellow zone. B convex hull of ω.
- Middle: Contourlines of the error.
- Right: Convergence of $\|u - u_h\|_{L^2(B)}$.

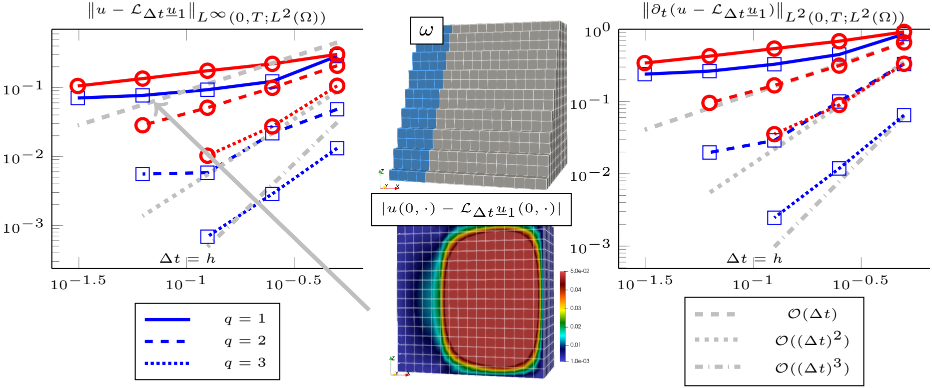
Wave equation in the cube: $\Omega = [0,1]^3$, $T = 1/2$[12]



- GCC fulfilled.
- Space-time FEM using DG in time, $\mathcal{L}_{\Delta t}$ is a C^0-reconstruction.
- Equal order polynomial order q in space and time.

[12] E.B. and Preuss, IMAJNA, 2025

What happens when GCC is violated



Control problem: direct formulation

By [Ervedoza–Zuazua'10] there is $(u, \phi) \in C^\infty((0,T) \times \Omega)^2$ solving

$$\begin{cases} \Box u = \chi \phi, \\ u|_{x \in \partial\Omega} = 0, \\ u|_{t=0} = u_0, \ \partial_t u|_{t=0} = u_1, \\ u|_{t=T} = 0, \ \partial_t u|_{t=0} = 0, \end{cases} \qquad \begin{cases} \Box \phi = 0, \\ \phi|_{x \in \partial\Omega} = 0. \end{cases}$$

Lemma. The solution to the above system is unique.

Proof. Let $(u_{(j)}, \phi_{(j)})$, $j = 1, 2$, be solutions and write $u = u_{(1)} - u_{(2)}$ and $\phi = \phi_{(1)} - \phi_{(2)}$. Then (u, ϕ) satisfies the system with $u_0 = u_1 = 0$, and

$$(\chi\phi, \phi)_{L^2((0,T)\times\Omega)} = (\Box u, \phi)_{L^2((0,T)\times\Omega)} = (u, \Box\phi)_{L^2((0,T)\times\Omega)} = 0.$$

The observability estimate implies $\phi = 0$, and $u = 0$ follows. □

Weak formulation of the control problem

- With $Q := (0, T) \times \Omega$ define the forms:

$$a(u, v) = -(\partial_t u, \partial_t v)_Q + (\nabla u, \nabla v)_{\mathcal{Q}} - (u, \partial_\nu v)_{L^2(\partial Q)} - (\partial_\nu u, v)_{L^2((0,T)\times\partial\Omega)},$$
$$L(v) = (u_1, v|_{t=0})_{L^2(\Omega)} - (u_0, \partial_t v|_{t=0})_{L^2(\Omega)},$$

  and $c(\phi, v) = (\chi\phi, v)_{L^2(Q)}$.
- If smooth (u, ϕ) solves

$$\begin{cases} \Box u = \chi\phi, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t=0} = u_0,\ \partial_t u|_{t=0} = u_1, \\ u|_{t=T} = 0,\ \partial_t u|_{t=0} = 0, \end{cases} \qquad \begin{cases} \Box\phi = 0, \\ \phi|_{x\in\partial\Omega} = 0. \end{cases}$$

- Then for all smooth enough ψ, v

$$a(u, \psi) = c(\phi, \psi) + L(\psi), \quad a(v, \phi) = 0.$$

Smoothness of minimal control

Theorem [Ervedoza–Zuazua'10]. Suppose GCC and $(u_0, u_1) \in C_0^\infty(\Omega)^2$. Then there is a solution $\phi \in C^\infty((0,T)\times\Omega)$ to the control problem s.t.

$$\begin{cases} \partial_t^2 \phi - \Delta \phi = 0, & \text{in } (0,T)\times\Omega, \\ \phi|_{x\in\partial\Omega} = 0, & \\ \phi|_{t=T} = \phi_0,\ \partial_t u|_{t=T} = \phi_1, & \end{cases}$$

where (ϕ_0, ϕ_1) is the unique minimizer over $L^2(\Omega)\times H^{-1}(\Omega)$ of

$$J(\phi_0,\phi_1) = \frac{1}{2}\int_0^T \int_\Omega \chi(t,x)|\phi(t,x)|^2 dxdt \\ + \langle u_0, \partial_t \phi|_{t=0}\rangle_{H_0^1(\Omega)\times H^{-1}(\Omega)} - (u_1, \phi|_{t=0})_{L^2(\Omega)}.$$

Stabilized FEM for the control problem[13]

We write $U_0 = (u_0, u_1)$ for the data and define the "energy"

$$E(U_0) = h^{-1}\|u_0\|^2_{L^2(\Omega)} + h\|u_1\|^2_{L^2(\Omega)}.$$

Our finite element method has the form: find the critical point of the Lagrangian $\mathcal{L}_C(u, \phi) : V_h^p \times V_h^q \to \mathbb{R}$,

$$\boxed{\mathcal{L}_C(u, \phi) = \frac{1}{2}E(U|_{t=0} - U_0) + \frac{1}{2}S(u, \phi) + \frac{1}{2}c(\phi, \phi) - a(u, \phi) + L(\phi),}$$

where $U = (u, \partial_t u)$ and S is a quadratic form giving the stabilization.

[13] E.B., Mohammadi, Münch and Oksanen, ESAIM COCV, 2023

Stabilization

We write $\mathcal{F}_h$ for the set of internal faces of the triangulation $\mathcal{T}_h$, and $[\![\cdot]\!]$ for the jump over $F \in \mathcal{F}_h$. The stabilization is given by

$$S(u,\phi) = \sum_{K \in \mathcal{T}_h} h^2 \|\Box u - \chi\phi\|^2_{L^2(K)} - \sum_{K \in \mathcal{T}_h} h^2 \|\Box \phi\|^2_{L^2(K)} + j(u,u) - j(\phi,\phi) + E(U|_{t=T}),$$

$$j(u,u) = \sum_{F \in \mathcal{F}_h} h \|[\![\partial_{n_F} u]\!]\|^2_{L^2(F)} + h^{-1}\|u\|^2_{L^2((0,T)\times\partial\Omega)}.$$

Observe that $S(u,\phi) = 0$ for a smooth solution (u,ϕ) to

$$\begin{cases} \Box u = \chi\phi, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t=0} = u_0,\ \partial_t u|_{t=0} = u_1, \\ u|_{t=T} = 0,\ \partial_t u|_{t=0} = 0, \end{cases} \qquad \begin{cases} \Box\phi = 0, \\ \phi|_{x\in\partial\Omega} = 0. \end{cases}$$

Error estimate (optimal control)

Theorem [E.B.-Feizmohammadi-Münch-Oksanen]. Suppose that the GCC holds. Let $p, q \geq 1$ and let $(u, \phi) \in H^{p+1}(Q) \times H^{q+1}(Q)$ be the unique solution to

$$\begin{cases} \Box u = \chi\phi, \\ u|_{x\in\partial\Omega} = 0, \\ u|_{t=0} = u_0, \ \partial_t u|_{t=0} = u_1, \\ u|_{t=T} = 0, \ \partial_t u|_{t=0} = 0, \end{cases} \qquad \begin{cases} \Box\phi = 0, \\ \phi|_{x\in\partial\Omega} = 0. \end{cases}$$

Then the Lagrangian $\mathcal{L}_C$ has a unique critical point $(u_h, \phi_h) \in V_h^p \times V_h^q$ and

$$\|\chi(\phi - \phi_h)\|_{L^2(M)} \lesssim h^p \|u\|_{H^{p+1}(Q)} + h^q \|\phi\|_{H^{q+1}(Q)}.$$

- Solutions with limited regularity (partial result for optimal control):
  Assume that $u_0 = 0$ and $u_1 \in L^2(\Omega)$ then

$$(u_h, \phi_h) \rightharpoonup (u, \phi) \text{ in } [L^2(Q)]^2.$$

Computational experiments: control problem, the weight

- Recall $\chi(x,t) = \chi_0(t)\chi_1(x)^2$
- Cut off functions $\chi_0 \in C_0^\infty([0,T])$ and $\chi_1 \in C_0^\infty([0,1])$

$$\chi_0(t) = \frac{e^{-\frac{1}{2t}} e^{-\frac{1}{2(T-t)}}}{e^{-\frac{1}{T}} e^{-\frac{1}{T}}}, \qquad \chi_1(x) = \frac{e^{-\frac{1}{5(x-a)}} e^{-\frac{1}{5(b-x)}}}{e^{-\frac{2}{5(b-a)}} e^{-\frac{2}{5(b-a)}}} 1_{[a,b]}(x) \tag{3}$$

for any $0 < a < b < 1$ and $T > 0$.

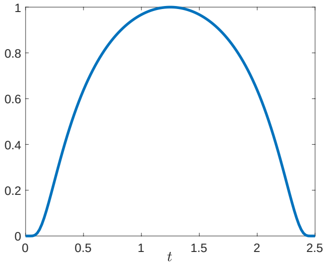


Figure: The $C_0^\infty([0,T])$ function $t \mapsto \chi_0(t)$, $t \in [0,T]$ with $T = 2.5$.

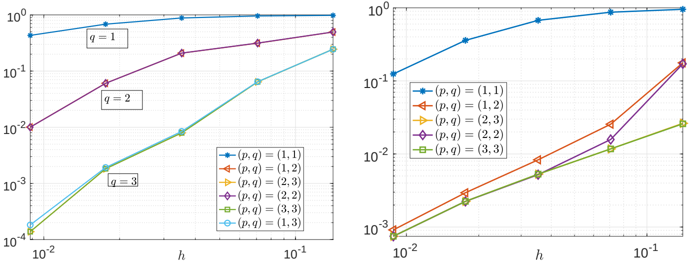
Computational experiments: control problem, setup

- $\Omega = (0, 1)$
- Initial data:

$$(u_0, u_1) = (\sin(\pi x), 0) \in H^{k+1}(\Omega) \times H^k(\Omega) \quad \forall k \in \mathbb{N}. \qquad \textbf{(Ex1)}$$

- We use the cut-off functions defined above, with $T = 2$, $a = 0.1$ and $b = 0.4$.
- The GCC holds true for this set of data.
- Reference solution computed on fine mesh, 409 000 triangles and $(u_h, \phi_h) \in V_h^p \times V_h^q$ with $(p, q) = (3, 3)$.

Computational experiments: control problem, $1+1d$



- $\|\chi(\phi - \phi_h)\|_{L^2(M)} / \|\chi\phi\|_{L^2(M)}$ vs. h;
- Left: $\chi_0(t)$ and $\chi_1(x)$ given by (3);
- Right: $\chi_0(t) = 1$ and $\chi_1(x) = 1_{(a,b)}(x)$.
- The smooth cut-off gives optimal convergence $O(h^{q+1})$
- The rough does not give optimal convergence for higher p, q.

Optimality of estimates I[14]

- Let $\{V_h\}$ be a family of finite dimensional spaces, with smallest length scale h.
- Assume that $\{u_h\}_h$ is a sequence of approximations in V_h to

$$\begin{cases} \Delta u = 0, & \text{in } \Omega, \\ u|_\omega = g_\omega, & \text{in } \omega, \end{cases}$$

  obtained using perturbed data $\tilde{g} := g_\omega + \delta g$, $\delta = \|\delta g\|_\omega$.
- Assume that the sequence $\{u_h\}$ satsifies
  1. $\|\Delta u_h\|_{H^{-2}(\Omega)} \lesssim h^{p+1} |u|_{H^{p+1}(\Omega)}$.
  2. $\|u_h - \tilde{g}\|_\omega \lesssim h^{p+1} |u|_{H^{p+1}(\Omega)} + \delta$.
  3. $\|u - u_h\|_\Omega \lesssim |u|_{H^{p+1}(\Omega)} + h^{-p-1}\delta$.

[14] E.B., Nechita and Oksanen, FoCM, 2025

Optimality of estimates II

- Estimate using the conditional stability[15]:

$$\|u-u_h\|_{L^2(B)} \lesssim (\|g_\omega - u_h\|_\omega + \|\Delta u_h\|_{H^{-2}(\Omega)})^\alpha (\|u-u_h\|_{L^2(\Omega)} + \|\Delta u_h\|_{H^{-2}(\Omega)})^{1-\alpha}. \quad (4)$$

- Using the properties 1.-3. of $\{u_h\}_h$ we get the bound[16]

$$\boxed{\|u-u_h\|_{L^2(B)} \lesssim h^{\alpha(p+1)} \|u\|_{H^{p+1}(\Omega)} + h^{-(1-\alpha)(p+1)}\delta.} \quad (5)$$

- A more general estimate with two parameters α_1 and α_2 would be:

$$\|u-u_h\|_{L^2(B)} \lesssim h^{\alpha_1(p+1)} \|u\|_{H^{p+1}(\Omega)} + h^{-(1-\alpha_2)(p+1)}\delta. \quad (6)$$

Theorem [E.B.–Nechita–Oksanen]. No approximation of the unique continuation problem in a subspace of H^1, having at most the rate of convergence of the assumptions 1. and 2. above can satisfy (6) for $\alpha_1, \alpha_2 \in [\alpha, 1)$, with $\alpha_1 > \alpha$ or $\alpha_2 > \alpha$.

---

[15] Stevenson and Mansuur, 2024

[16] Stevenson and Mansuur, 2024; E.B., Oksanen and Lu, 2025

Conclusions

- Finite element analysis for linear ill-posed problems with conditional stability.
- Stability estimates give guiding principle for method design.
- Error analysis based on (c.f. Lax equivalence Theorem):

conditional stability + residual stability + consistency $\rightarrow$ optimal convergence

- Results on unique continuation extends to control problems.
- Optimality with respect to stability and approximation.

Finite dimensionality of missing data, in 2D[17]

- Let $\mathcal{V}_N$ be a finite dimensional subspace of $H^1(\partial\Omega)$.
- Then if P is a projection on $\mathcal{V}_N$, the following stability estimate holds for $w \in H^1(\Omega)$,

$$\boxed{\|v\|_{H^1(\Omega)} \lesssim \|w\|_{L^2(\omega)} + \|(I-P)w\|_{H^{1/2}(\partial\Omega)}, + \|\Delta w\|_{H^{-1}(\Omega)}}$$

- Here the hidden constant grows exponentially in N.
- Relies on:
  1. Norm equivalence of discrete spaces
  2. Trace inequality for harmonic functions in weak norm [Grisvard, Thm 1.5.3.4]:

$$\|w\|_{H^{-\frac{1}{2}-\epsilon}(\partial\Omega)} \leq \|w\|_{L^2(\Omega)}, \quad \epsilon > 0.$$

  3. Logarithmic global stability, $\mu(s) = \log(s))^{-\alpha}$, $\alpha \in (0,1)$.
  4. Assume w harmonic. Let $\delta = \|Pw\|_{H^{1/2}(\partial\Omega)}$, $\eta = \|w\|_{L^2(\omega)}$, $E = \|w\|_{H^1(\Omega)}$.
  5. Then, since $E \lesssim \delta$ for harmonic functions,

$$\delta \underbrace{\lesssim}_{1.} \|w\|_{H^{-\frac{1}{2}-\epsilon}(\partial\Omega)} \underbrace{\lesssim}_{2.} \|w\|_{L^2(\Omega)} \underbrace{\lesssim}_{3.} \delta\mu\left(\frac{\eta}{E}\right) \Rightarrow 1 \lesssim \mu\left(\frac{\eta}{E}\right) \Rightarrow 1 \lesssim \frac{\eta}{E}$$

[17] E.B. and Oksanen, M3AS, 2024.

FEM of UC, finite dimensional trace

Problem 4. Given $u|_\omega$ determine u satisfying the Poisson equation $\Delta u = f$ in Ω.

Assume that u is close to $\mathcal{V}_N$ on the boundary, i.e. for some $\delta > 0$ there exists $y_{\partial,N} \in \mathcal{V}_N$ such that

$$\boxed{\|u - y_{\partial,N}\|_{H^{3/2}(\partial\Omega)} \leq \delta} \tag{7}$$

On $X_h = V_h \times W_h$ define the Lagrangian

$$\boxed{\begin{aligned} \mathcal{L}(u_h, z_h) &= \tfrac{1}{2}\|u_h - q\|^2_{L^2(\omega)} + \tfrac{1}{2}B(u_h) + \tfrac{1}{2}s(u_h, u_h) \\ &+ a(u_h, z_h) - (f, z_h) \end{aligned}} \tag{8}$$

where we assume $a(\cdot,\cdot)$ coercive on $H^1_0(\Omega) \times H^1_0(\Omega)$ and

$$B(u) := h^{-1}\|u - Pu\|^2_{L^2(\partial\Omega)} + h\|\nabla_\partial(u - Pu)\|^2_{L^2(\partial\Omega)},$$

with ∇_∂ the tangential gradient on $\partial\Omega$.

Finite element method and error estimate

Theorem [E.B.–Oksanen]. Suppose that $u \in H^2(\Omega)$, with Ω convex, is the solution of Problem 4 and $(u_h, z_h) \in X_h$ is the stationary point of (8). Then, when $p = m = 1$, there holds

$$\|u - u_h\|_{H^1(\Omega)} \leq C_N(h\|D^{k+1}u\|_{L^2(\Omega)} + \delta), \tag{9}$$

Convergence optimal up to the distance between $u|_{\partial\Omega}$ and $\mathcal{V}_N$

- When h is small enough no stabilization is necessary, but difficult to quantify.
- How can we find $\mathcal{V}_N$ and can we make N small enough?
- In a set of collective data identify $\mathcal{V}_N$ using PCA[18] or machine learning[19].

[18] Boulakia, Corrie and Lombardi, M2AN 2025
[19] E.B., Larson, Larsson, Larsson. CMAME, 2025

Observed Boundary Data

- Consider the problem to find $u \in H^1(\Omega)$ such that

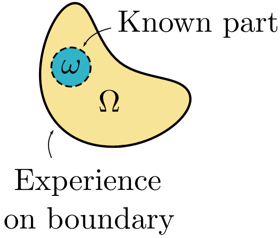
$$\boxed{\underbrace{\mathcal{P}u = f \text{ in } H^{-1}(\Omega)}_{\text{PDE (physics)}} \quad \text{and} \quad \underbrace{u|_\omega = q}_{\text{known part of solution}}}$$

  where $\Omega \subset \mathbb{R}^n$, $\omega \subset \Omega$ is a subset of Ω where $u = q$ with q known measured data, and $\mathcal{P} : H^1(\Omega) \to H^{-1}(\Omega)$ a second-order nonlinear pde-operator

- Instead of specified boundary conditions we have

$$\boxed{\underbrace{g \in \mathcal{G}_D = \{g_i \,|\, i \in I\} \subset H^{1/2}(\partial\Omega)}_{\text{experience}}}$$

  where $\mathcal{G}_D$ is a set of observed boundary data



General Minimization Idea Revisited

- We approach this problem by solving the constrained minimization problem

$$u = \underbrace{\operatorname{arginf}_{v \in V} \frac{1}{2} \|v - q\|^2_{M(\omega)}}_{\text{minimizing w.r.t. known part}} \text{ subject to: } \underbrace{\underbrace{\mathcal{P}(v) = f \text{ in } H^{-1}(\Omega)}_{\text{PDE}} \text{ and } \underbrace{v|_{\partial\Omega} \in \mathcal{G}_D}_{\text{experience}}}_{\text{constraints}}$$

- With access to an efficient solution operator $\phi_u : \mathcal{G}_D \ni g \mapsto u \in H^1(\Omega)$ for the PDE constraint, fulfilling $\mathcal{P}u = f$ and $u|_{\partial\Omega} = g$, the problem can be formulated as minimizing

$$\min_{g \in \mathcal{G}_D} \frac{1}{2} \|\phi_u(g) - q\|^2_{M(\omega)}$$

- The solution operator ϕ_u may be constructed either using a fast numerical solver, or a parametrization based on off-line computation/learning
- If we can identify a lower-dimensional structure in $\mathcal{G}_D$, we can further simplify the minimization by parametrizing the experience constraint

Machine Learning for Parametrizing the Constraints

We use tools from machine learning to parametrize a solution space satisfying the constraints:

- **Experience constraint:** Compressed using an autoencoder, essentially constructing a parametrization

$$\boxed{\phi_g : Z \to \mathcal{G}_D, \qquad \mathcal{G}_D \approx \phi_g(Z)}$$

  of the data as a function over a lower-dimensional latent space $Z = I^m \subset \mathbb{R}^m$

- **PDE constraint:** Resolved using operator learning, which enables representation of the mapping from boundary data to solution

$$\boxed{\phi_u : \mathcal{G}_D \ni g \to u \in H^1(\Omega)}$$

  which solves

$$\mathcal{P}(u) = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega$$

Representing the Boundary Dataset

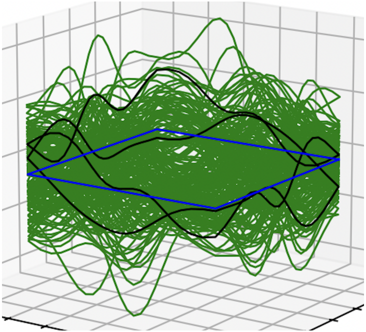
The boundary condition dataset $\mathcal{G}_D$ needs to be represented in some compute friendly way:

- We generate a synthetic data set $\mathcal{G}_D$ from the Fourier series

$$g(s) = g_0 + \sum_{n=1}^{(N-1)/2} g_{2n-1} \sin(2\pi ns/L) + g_{2n} \cos(2\pi ns/L)$$

  with coefficients $g_j = \hat{g}_j + \delta_j$, where $\hat{g}_j \sim \mathcal{U}(-1, 1)$ and $\delta_j \sim \mathcal{N}(0, 0.0225)$
- Then we compute POD boundary basis functions. Keep N significant ones
- Thus the data set $\mathcal{G}_D$ may be represented by N POD coefficients $a = (a_1, ..., a_N)$

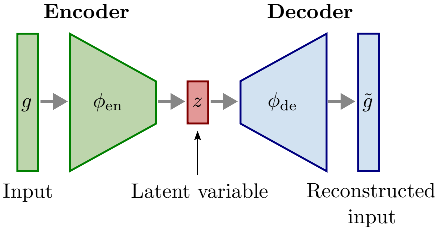
$$\boxed{F_N : \mathcal{G}_D \mapsto G_N = \{a_i \mid i \in I\} \subset \mathbb{R}^N}$$



Autoencoder Compression of the Expansion Coefficients

If there is a underlying smooth low-dimensional structure in G_N we may find and exploit it using an *autoencoder*

$$\boxed{\phi_{\text{auto}} = \phi_{\text{de}} \circ \phi_{\text{en}}}$$



- **Input:** N POD coefficients $(a_1, ..., a_N)$
- **Output:** Approximation of input
- **Latent variables:** Latent layer output, dim $<< N$
- **Layers:** Weights W, biases b, and activation σ as $\sigma(Wx + b)$
- **Activation:** Exponential Linear Unit (ELU)

Training: Loss function is squared l^2-error. The optimization problem during training is

$$\min_{\zeta} \mathbb{E}_{a \in G_N} \left[\| a - \phi_{\text{auto},\zeta}(a) \|_{l^2}^2 \right]$$

Operator Learning of the PDE

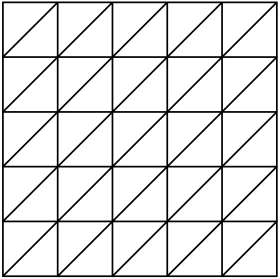
- We shall approximate the operator

$$\phi_u : \mathcal{G}_D \ni g \mapsto u \in H^1(\Omega)$$

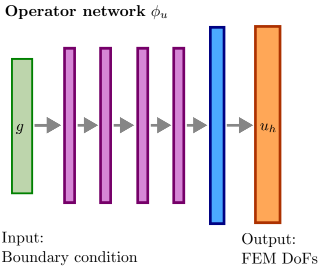
  solving the PDE with boundary conditions $v|_{\partial\Omega} = g$
- We discretize using a finite element space V_h and formulate the problem as an energy minimization problem (or using some other loss function)
- Using the representation of the boundary data we approximate ϕ_u using a neural network

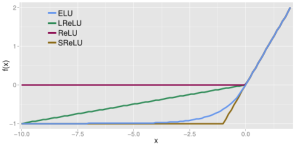
$$\phi_{u,N,h} : G_N \to V_h$$

  directly mapping the POD coefficients to the finite element solution



Operator Learning of the PDE





- **Input:** POD coefficients
- **Output:** FE-DoFs
- **Layers:** Weights W, biases b, and activation σ as $\sigma(Wx + b)$
- **Activation:** Exponential Linear Unit (ELU)

Training: The loss function is the energy functional of the PDE problem. For $v \in V_h$ s.t. $v|_{\partial\Omega} = g_N$,

$$E(v) = \int_\Omega \frac{1}{2}|\mathcal{P}^{1/2}v|^2 - fv \, \mathrm{d}x = \sum_{T \in \mathcal{T}_h} \int_T \frac{1}{2}|\mathcal{P}^{1/2}v|^2 - fv \, \mathrm{d}$$

The optimization problem during training is

$$\min_\theta \mathbb{E}_{a \sim \mathcal{N}} \left[E(\phi_{u,\theta,N,h}(a)) \right]$$

Minimization in Latent Space

Recall the inverse problem on constrained minimization form

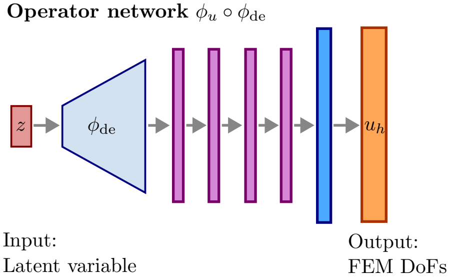
$$u = \operatorname{arginf}_{v \in V} \frac{1}{2} \|v - q\|^2_{M(\omega)} \text{ subject to: } \mathcal{P}v = f \text{ in } H^{-1}(\Omega) \text{ and } v|_{\partial\Omega} \in \mathcal{G}_D$$

- The problem may be formulated as minimization over *boundary data* G_N

$$\boxed{\min_{a \in G_N} \frac{1}{2} \|\phi_{u,\theta,N,h}(a) - q\|^2_{M(\omega)}}$$

- Further slim down by using decoder part $\phi_{\mathrm{de},\zeta}$ of $\phi_{\mathrm{auto},\zeta}$. Thus minimization over *lower-dimensional latent space Z*

$$\boxed{\min_{z \in Z} \frac{1}{2} \|\phi_{u,\theta,N,h} \circ \phi_{\mathrm{de},\zeta}(z) - q\|^2_{M(\omega)}}$$



Numerical Example

Operator learning
We consider triangle meshes of the unit square, standard P1-elements, and the nonlinear energy functional

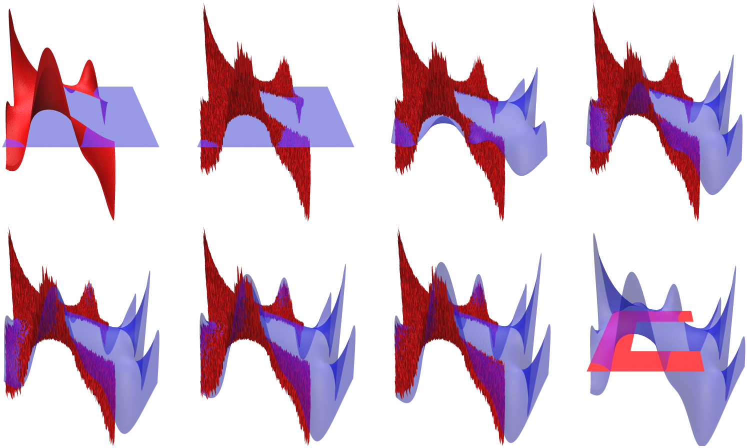
$$E(v) = \int_{(-0.5,0.5)^2} \frac{1}{2}(1+v^2)|\nabla v|^2 \, \mathrm{d}x$$

Experience
To construct boundary data with a low-dimensional structure, we choose coefficients a from L Gaussian bell curves dependent on n_Z latent variables z_k as

$$a_j = a_j(z) = \exp(-2(z_k - z_{0,l})^2) + \delta_j$$

Each a_j is assigned exactly one bell curve l with midpoint $z_{0,l}$ and one latent variable z_k according to $l = j \mod L$ and $k = j \mod n_Z$

Numerical example: Op/dec, $\delta_\omega \in \mathcal{U}(-0.05, 0.05)$



Conclusions

- Finite element analysis for linear ill-posed problems with conditional stability.
- Stability estimates give guiding principle for method design.
- Error analysis based on (c.f. Lax equivalence Theorem):

conditional stability $+$ residual stability $+$ consistency $\rightarrow$ optimal convergence

- Optimality with respect to stability and approximation.
- Finite dimensionality of missing data recovers Lipschitz stability.
- Collective data integration improves stability.
- Machine learning facilitates manifold and operator learning for data compression and rapid optimization.

Elements of proof of optimality, concentric balls

$0 < r_1 < r_2 < r_3 < R$

1. Sharp three ball's inequality for harmonic functions (α optimal)

$$\|u\|_{L^2(B(r_2))} \leq \|u\|^{\alpha}_{L^2(B(r_1))}\|u\|^{(1-\alpha)}_{L^2(B(r_3))}, \ \alpha = \frac{\beta}{1+\beta}, \ \beta = \frac{\log(r_3)-\log(r_2)}{\log(r_2)-\log(r_1)}. \quad (10)$$

2. Assume that there exists $\alpha_1, \alpha_2 \geq \alpha$, with one inequality strict so that

$$\|u-u_h\|_{L^2(B(r_2))} \lesssim h^{\alpha_1 p}\|u\|_{H^{p+1}(B(r_3))} + h^{-(1-\alpha_2)p}\|\delta q\|_{B(r_1)}. \quad (11)$$

3. Choosing $u = r^n \sin(n\theta)$, or $\delta q = r^n \sin(n\theta)$ satisfaction of (11) implies

$$\|u\|_{L^2(B(r_2))} \lesssim \|u\|^{\tilde{\alpha}}_{L^2(B(r_1))}\|u\|^{1-\tilde{\alpha}}_{H^{p+1}(B(r_3))}, \text{ with } \tilde{\alpha} = \frac{\alpha_1}{1+\alpha_1-\alpha_2}.$$

4. Show that $\tilde{\alpha} > \alpha$. Example: $\alpha_1 < \alpha_2$ then $\alpha_2 - \alpha_1 \in (0,1)$ and

$$\tilde{\alpha} \geq \frac{\alpha}{1-(\alpha_2-\alpha_1)} > \alpha.$$

5. Caccioppoli type inequality, with $R > r_3$, such that $\alpha(R) < \tilde{\alpha}$

$$\|u\|_{L^2(B(r_2))} \lesssim \|u\|^{\tilde{\alpha}}_{L^2(B(r_1))}\|u\|^{1-\tilde{\alpha}}_{L^2(B(R))}.$$

6. Contradicts the optimality of α!

Elements of proof, finite dimensional stability

- Assume w harmonic. Let $\delta = \|Pw\|_{H^{1/2}(\partial\Omega)}$, $\eta = \|w\|_{L^2(\omega)}$, $E = \|w\|_{H^1(\Omega)}$.
- Then, since $\delta \lesssim E$ for harmonic functions,

$$\delta \underbrace{\lesssim}_{\text{I}} \|w\|_{H^{-\frac{1}{2}-\epsilon}(\partial\Omega)} \underbrace{\lesssim}_{\text{II}} \|w\|_{L^2(\Omega)} \underbrace{\lesssim}_{\text{III}} \delta\mu\left(\frac{\eta}{E}\right) \to 1 \lesssim \mu\left(\frac{\eta}{E}\right) \to 1 \lesssim \frac{\eta}{E}$$

- I: norm equivalence on finite dimensional spaces.
- II: trace inequality.
- III: stability
- We have shown: $\|w\|_{H^1(\Omega)} \lesssim \|w\|_{L^2(\omega)} \qquad \text{(Lip)}$.
- Let $u \in H^1(\Omega)$ and $v \in H^1(\Omega)$ solution of

$$\Delta v = \Delta u, \quad v|_{\partial\Omega} = (I - P)u.$$

Then $w = u - v$ satisfies (Lip) and

$$\|u\|_{H^1(\Omega)} \leq \|w\|_{H^1(\Omega)} + \|v\|_{H^1(\Omega)} \underbrace{\lesssim}_{(Lip)} \|w\|_{L^2(\omega)} + \|v\|_{H^1(\Omega)}$$

$$\lesssim \|u\|_{L^2(\omega)} + \|\Delta u\|_{H^{-1}(\Omega)} + \|(I-P)u\|_{H^{\frac{1}{2}}(\partial\Omega)}.$$

Example : Autoencoder vs PCA Example

Simple non-linear example with $a \in \mathbb{R}^9$ and $z \in \mathbb{R}^3$ (entries $\sim \mathcal{U}(-2, 2)$) given by

$$a = Az + Bz^2 + \delta$$

Matrices $A, B \in \mathbb{R}^{9\times 3}$ (fixed entries $\sim \mathcal{U}(-1, 1)$), perturbation $\delta \in \mathbb{R}^9$ (entries $\sim \mathcal{N}(0, \mathrm{sd}^2)$)

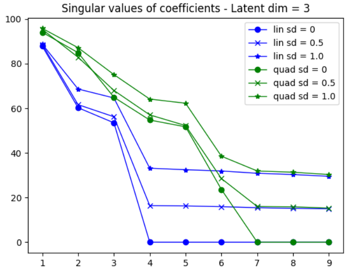


Figure: **PCA** - Linear ($B = 0$), Quadratic ($B \neq 0$)

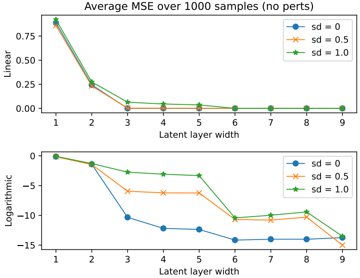


Figure: **Autoencoders** - Quadratic ($B \neq 0$)

PCA suggests latent dim = 3 for Linear and latent dim = 6 for Quadratic

Autoencoders suggest latent dim = 3 AND latent dim = 6 for Quadratic

Time (M1 CPU): PCA (1k samples) $\lesssim$ 1 s, training 9 AEs (tot 6 layers, ELU) $\approx$ 30 - 75 mins!

Example : Architecture and Training - Operator network

- 5 layers of widths X-X-X-X-O, ELU activation in the first four, input data width $N = 21$
- Training with Adam optimizer, 10^6 iterations, input batches, number of elements used in E
- Learning errors as measures of "well-trainedness"

Table: Architectures and batch sizes for various meshes

Mesh	DoFs (O)	Width (X)	Batch size
10x10	81	64	32
28x28	729	256	64
82x82	6561	512	64
244x244	59049	1024	96

Table: GPU (A100) training info and learning errors

Mesh	Els	Training time	GPU Util	Inference time	$E(u_{g=0})$	H_0^1-error 1k-avg (rel)	L^2-error 1k-avg (rel
10x10	All	339 s	47%	0.8 ms	8.9e-5	1.6e-2 (1.23%)	1.2e-3 (1.07
28x28	All	354 s	69%	0.8 ms	5.1e-6	5.5e-3 (0.73%)	2.9e-4 (0.44
82x82	All	617 s	100%	0.8 ms	3.7e-6	3.6e-3 (0.82%)	8.4e-5 (0.22
244x244	4k	2733 s	100%	0.8 ms	1.0e-4	2.5e-2 (9.8%)	1.6e-4 (0.72

Example : Architecture and Training - Autoencoder network

- For the POD coefficients bell curve latent dependence, we take

$$(n_Z, L) = (3, 7) \text{ with } x_0 = (0, 2, 4, 6, 8, 10, 12) \text{ and } z_k \sim \mathcal{U}(-2, 14)$$

- 5 layers of widths 64-64-Z-64-21, latent layer width Z varied, ELU activation in first four
- Training with Adam optimizer, 10^6 iterations, batch size $= 64$, training time $210 - 250$ s on Apple M1 CPU

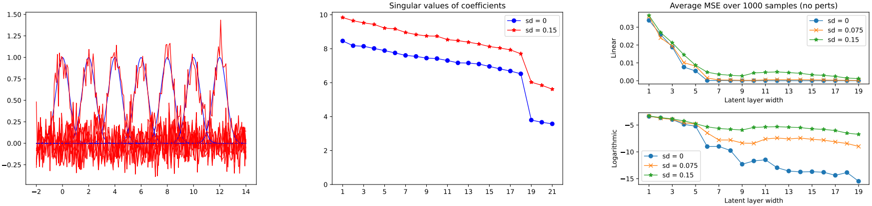


Figure: **Left:** Coefficient bell curves. **Middle:** PCA results. **Right:** Autoencoder results for three different perturbations used on data during training.

Autoencoders suggest a reduction to 9 dimensions could have same effect as down to 17

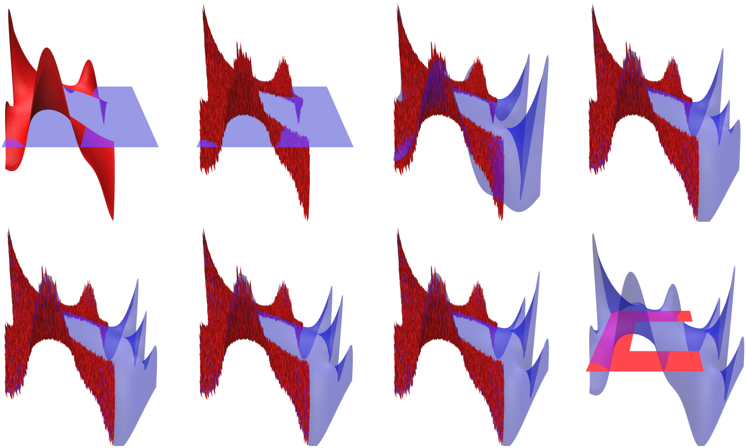
Example : Minimization

In practice, we take $q \in V_h$ and use MSE of DoFs belonging to ω in minimization

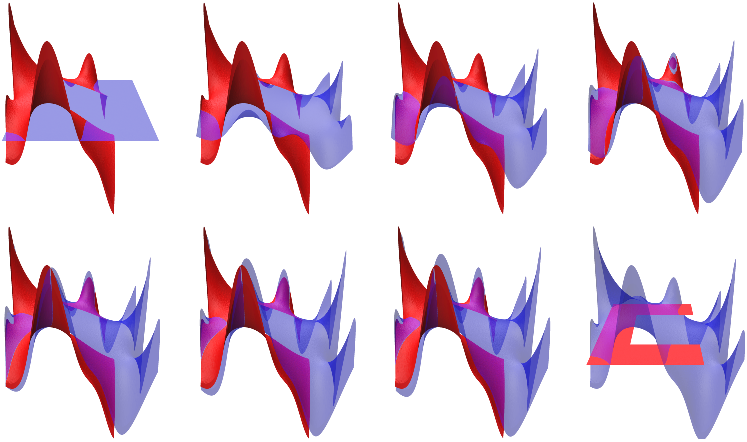
Table: Summary of optimization results for *nonlinear* inverse problems using an operator network (See subsequent figure series). All problems have the same reference solution and use the operator network for the 244x244 mesh (21 input coefficients, 59049 output DoFs). In the table, "Op" means the operator network, "dec" means decoder, "sd = x" means what perturbation was added to the training data for the decoder, and "δ_ω" means that noisy data was used for the inverse problem.

Configuration	Iterations	Avg iter time	MSE_ω	$MSE_{co(\omega)}$	MSE
Op, δ_ω	2843	4.93e-2 s	8.36e-4	8.38e-4	1.9
Op + dec "sd = 0"	1481	4.65e-2 s	2.22e-3	1.52e-3	4.83
Op + dec "sd = 0", δ_ω	1018	4.67e-2 s	3.07e-3	2.37e-3	5.25
Op + dec "sd = 0.15", δ_ω	212	5.19e-2 s	1.99e-2	1.44e-2	5.5

Example : Minimization: Op, δ_ω



Example : Minimization: Op + dec “sd = 0”



Example : Minimization: Op + dec “sd = 0.15”, δ_ω