Neural Control of Discrete Weak Formulations of PDEs

Ignacio Brevis[a], Ignacio Muga[a,c], Kris Van der Zee[b]
[a] IMA, Pontificia Universidad Católica de Valparaíso, Chile.
[b] School of Mathematical Sciences, The University of Nottingham, UK.
[c] Basque Center for Applied Mathematics (BCAM), Spain.

@Workshop on Minimum Residual & Least-Squares Finite Element Methods, PUC, Chile 6th October, 2022

Outline

Motivational Examples

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

$$\begin{cases} -u'' = \delta_\lambda \quad \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} \quad u(x_0) \end{cases}$$

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

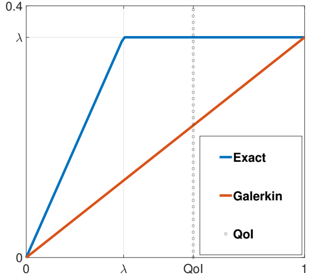
$$\begin{cases} -u'' = \delta_\lambda \quad \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} \quad u(x_0) \end{cases} \qquad \begin{cases} \text{Find } u \in H^1_{(0}(0,1) \text{ s.t.} \\ \displaystyle\int_0^1 u'v' = v(\lambda) \qquad \forall v \in H^1_{(0}(0,1) \end{cases}$$

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

$$\begin{cases} -u'' = \delta_\lambda \quad \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} \quad u(x_0) \end{cases}$$

$$\begin{cases} \text{Find } u \in H^1_{(0}(0,1) \text{ s.t.} \\ \int_0^1 u'v' = v(\lambda) \qquad \forall v \in H^1_{(0}(0,1) \end{cases}$$

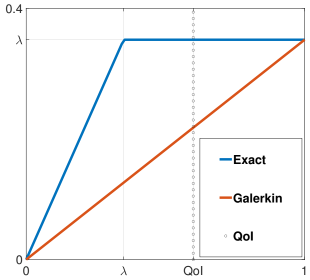


Trial space $\mathbb{U}_h = \text{span}\{x\}$

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

$$\begin{cases} -u'' = \delta_\lambda \quad \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} \quad u(x_0) \end{cases} \qquad \begin{cases} \text{Find } u \in H^1_{(0}(0,1) \text{ s.t.} \\ \displaystyle\int_0^1 u'v' = v(\lambda) \qquad \forall v \in H^1_{(0}(0,1) \end{cases}$$



Trial space $\mathbb{U}_h = \text{span}\{x\}$

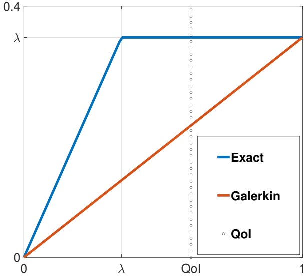
What if we use a parametrized test function

$$v(x) = \theta_1 x + e^{-\theta_2}(1 - e^{-\theta_1 x}) \quad ?$$

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

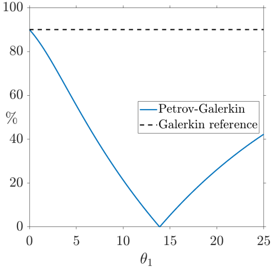
$$\begin{cases} -u'' = \delta_\lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} & u(x_0) \end{cases} \qquad \begin{cases} \text{Find } u \in H^1_{(0}(0,1) \text{ s.t.} \\ \int_0^1 u'v' = v(\lambda) \qquad \forall v \in H^1_{(0}(0,1) \end{cases}$$



Trial space $\mathbb{U}_h = \text{span}\{x\}$

What if we use a parametrized test function

$$v(x) = \theta_1 x + e^{-\theta_2}(1 - e^{-\theta_1 x}) \quad ?$$

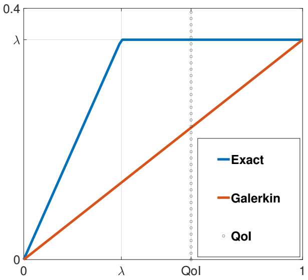


Relative error (θ_2 is fixed)

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

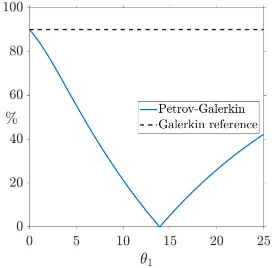
$$\begin{cases} -u'' = \delta_\lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} & u(x_0) \end{cases} \qquad \begin{cases} \text{Find } u \in H^1_{(0}(0,1) \text{ s.t.} \\ \int_0^1 u'v' = v(\lambda) \qquad \forall v \in H^1_{(0}(0,1) \end{cases}$$



Trial space $\mathbb{U}_h = \text{span}\{x\}$

What if we use a parametrized test function

$$v(x) = \theta_1 x + e^{-\theta_2}(1 - e^{-\theta_1 x}) \quad ?$$



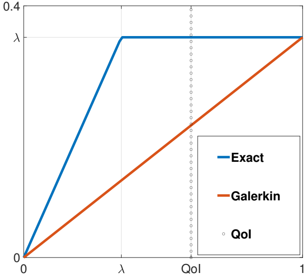
Relative error (θ_2 is fixed)

- Is the test function v trainable to reduce the error in known data ?

Motivating example 1: incorporation of data.

Let $\lambda \in (0,1)$ and consider the differential equation:

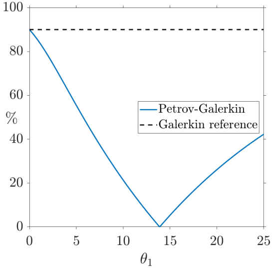
$$\begin{cases} -u'' = \delta_\lambda \quad \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{data:} \quad u(x_0) \end{cases} \qquad \begin{cases} \text{Find } u \in H^1_{(0}(0,1) \text{ s.t.} \\ \displaystyle\int_0^1 u'v' = v(\lambda) \qquad \forall v \in H^1_{(0}(0,1) \end{cases}$$



Trial space $\mathbb{U}_h = \text{span}\{x\}$

What if we use a parametrized test function

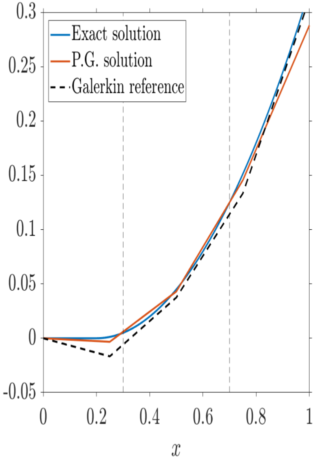
$$v(x) = \theta_1 x + e^{-\theta_2}(1 - e^{-\theta_1 x}) \quad ?$$



Relative error (θ_2 is fixed)

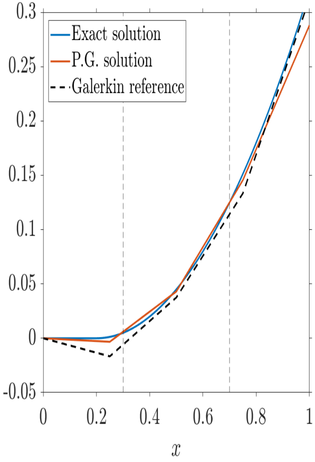
- Is the test function v trainable to reduce the error in known data ?
- How to come up with a practical family of trainable test functions?

Motivating example 1: incorporation of data.



How to obtain the red solution?

Motivating example 1: incorporation of data.

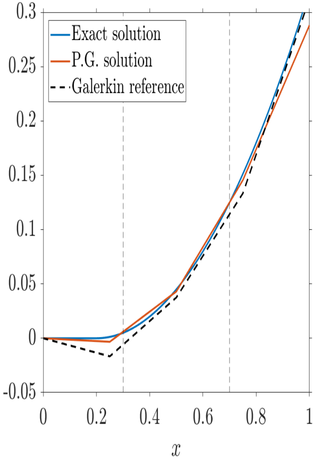


How to obtain the red solution?

IDEAS:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.

Motivating example 1: incorporation of data.

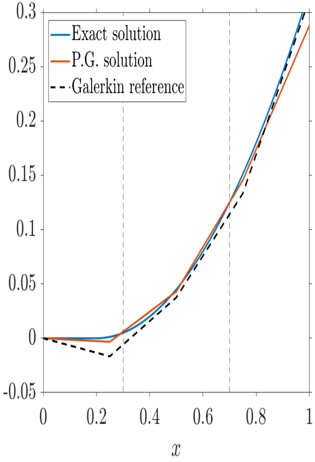


How to obtain the red solution?

IDEAS:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.
- Define the discrete solution $u_{h,\xi} \in \mathbb{U}_h$ obtained using this *neurally-controlled* weak-form.

Motivating example 1: incorporation of data.



How to obtain the red solution?

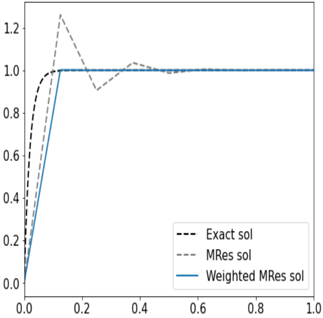
IDEAS:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.
- Define the discrete solution $u_{h,\xi} \in \mathbb{U}_h$ obtained using this *neurally-controlled* weak-form.
- Assuming you have a set of reliable data $\{\overline{q_i}\}_{i=1}^{N_d} \subset \mathbb{R}$ (e.g., known quantities of interest of the exact solution), then try to minimize the following cost functional:

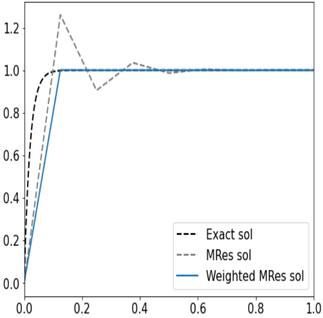
$$J(u_{h,\xi}) := \sum_{i=1}^{N_d} \frac{1}{2} |q_i(u_{h,\xi}) - \overline{q_i}|^2 \quad \longrightarrow \min,$$

where $q_i : \mathbb{U} \to \mathbb{R}$ are QoI functionals. (supervised learning)

Motivating example 2: incorporation of qualitative attributes.



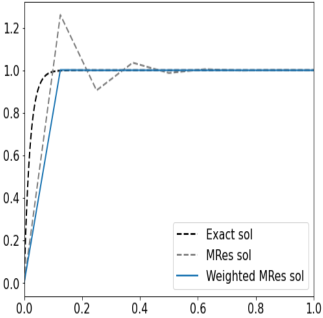
Motivating example 2: incorporation of qualitative attributes.



Same idea:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.

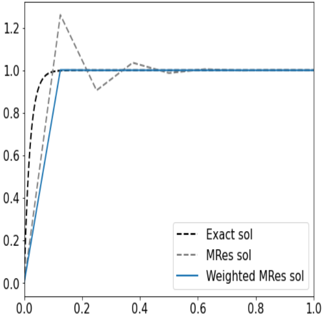
Motivating example 2: incorporation of qualitative attributes.



Same idea:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.
- Define the discrete solution $u_{h,\xi} \in \mathbb{U}_h$ obtained using this *neurally-controlled* weak-form.

Motivating example 2: incorporation of qualitative attributes.



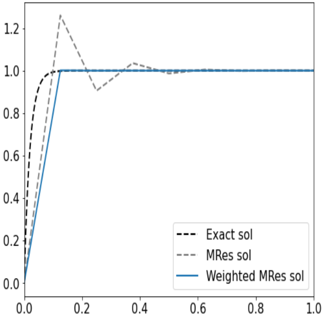
Same idea:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.
- Define the discrete solution $u_{h,\xi} \in \mathbb{U}_h$ obtained using this *neurally-controlled* weak-form.
- Try to minimize the total variation:

$$J(u_{h,\xi}) := \|u'_{h,\xi}\|_{L^1} \quad \longrightarrow \min$$

(unsupervised learning)

Motivating example 2: incorporation of qualitative attributes.



Same idea:

- Incorporate a neural-network ξ as a control variable into the discrete weak-formulation.
- Define the discrete solution $u_{h,\xi} \in \mathbb{U}_h$ obtained using this *neurally-controlled* weak-form.
- Try to minimize the total variation:

$$J(u_{h,\xi}) := \|u'_{h,\xi}\|_{L^1} \quad \longrightarrow \text{min}$$

(unsupervised learning)

Q: How to intervene a discrete formulation in order to incorporate a control parameter?

Discrete formulations

- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not, with inner-product $(\cdot,\cdot)_{\mathbb{V}}$)
- $b(\cdot,\cdot) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ (continuous bilinear form of the PDE)
- $f \in \mathbb{V}^*$ RHS of the PDE $b(u,v) = f(v) \quad \forall v \in \mathbb{V}$
- $B : \mathbb{U} \to \mathbb{V}^*$ (induced operator) $Bw = b(w,\cdot) \in \mathbb{V}^*, \;\; w \in \mathbb{U}$

The parent: (Residual minimization form)

$$u_h = \arg\min_{w_h \in \mathbb{U}_h} \|f - Bw_h\|_{\hat{\mathbb{V}}^*} = \arg\min_{w_h \in \mathbb{U}_h} \left(\sup_{v \in \hat{\mathbb{V}}} \frac{|f(v) - b(w_h, v)|}{\|v\|_{\mathbb{V}}} \right)$$

The offspring: (Mixed/saddle-point form)

$$\begin{aligned} (r, v)_{\mathbb{V}} + b(u_h, v) &= f(v), \quad && \forall v \in \hat{\mathbb{V}} \\ b(w_h, r) &= 0, && \forall w_h \in \mathbb{U}_h \end{aligned}$$

The holy spirit: (Petrov-Galerkin w/optimal test functions form)

$$b(u_h, v_h) = f(v_h)\,, \qquad \forall v_h \in V_h := R_{\hat{\mathbb{V}}}^{-1} B \mathbb{U}_h$$

Theoretical formalism

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not)

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not)
- $a(\xi; \cdot, \cdot) : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ (controlled equivalent inner-product)

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not)
- $a(\xi; \cdot, \cdot) : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ (controlled equivalent inner-product)
- $b(\cdot, \cdot) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ (variational form of the PDE)

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not)
- $a(\xi; \cdot, \cdot) : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ (controlled equivalent inner-product)
- $b(\cdot, \cdot) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ (variational form of the PDE)
- $\mathbb{Z}$ (space for observations) and $Q : \mathbb{U} \to \mathbb{Z}$ (observation operator)

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not)
- $a(\xi; \cdot, \cdot) : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ (controlled equivalent inner-product)
- $b(\cdot, \cdot) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ (variational form of the PDE)
- $\mathbb{Z}$ (space for observations) and $Q : \mathbb{U} \to \mathbb{Z}$ (observation operator)

State problem: Given $\xi \in \mathbb{X}$ & $f \in \mathbb{V}^*$, find $u_{h,\xi} \in \mathbb{U}_h$ (and $r_\xi \in \hat{\mathbb{V}}$) s.t.

$$\boxed{\begin{aligned} a(\xi; r_\xi, v) + b(u_{h,\xi}, v) &= f(v), && \forall v \in \hat{\mathbb{V}}, \\ b(w_h, r_\xi) &= 0, && \forall w_h \in \mathbb{U}_h\,, \end{aligned}}$$

Discrete state problem and associated cost functional

Functional spaces and operators:

- $\xi \in \mathbb{X}$ (Hilbert space for the control variable)
- $\mathbb{U}_h \subset \mathbb{U}$ (Hilbert trial space for discrete solutions)
- $\hat{\mathbb{V}} \subseteq \mathbb{V}$ (Hilbert test space, discrete or not)
- $a(\xi;\cdot,\cdot) : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ (controlled equivalent inner-product)
- $b(\cdot,\cdot) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ (variational form of the PDE)
- $\mathbb{Z}$ (space for observations) and $Q : \mathbb{U} \to \mathbb{Z}$ (observation operator)

State problem: Given $\xi \in \mathbb{X}$ & $f \in \mathbb{V}^*$, find $u_{h,\xi} \in \mathbb{U}_h$ (and $r_\xi \in \hat{\mathbb{V}}$) s.t.

$$\begin{aligned} a(\xi; r_\xi, v) + b(u_{h,\xi}, v) &= f(v), \quad \forall v \in \hat{\mathbb{V}}, \\ b(w_h, r_\xi) &= 0, \quad \forall w_h \in \mathbb{U}_h\,, \end{aligned}$$

Reduced cost functional: Given desirable observations or attributes $z_o \in \mathbb{Z}$,

$$j(\xi) := j_1(\xi) + \alpha j_2(\xi) := \frac{1}{2}\|Q(u_{h,\xi}) - z_o\|_{\mathbb{Z}}^2 + \frac{\alpha}{2}\|\xi\|_{\mathbb{X}}^2$$

Neural quasi-minimization

- The idea is to minimize $j(\cdot)$ over a *computable* subset $\mathcal{M}_n \subset \mathbb{X}$

Neural quasi-minimization

- The idea is to minimize $j(\cdot)$ over a *computable* subset $\mathcal{M}_n \subset \mathbb{X}$
- $\mathcal{M}_n \subset \mathbb{X}$ will be the manifold of all the realizations functions for a fixed neural-network architecture parametrized by n

Neural quasi-minimization

- The idea is to minimize $j(\cdot)$ over a *computable* subset $\mathcal{M}_n \subset \mathbb{X}$
- $\mathcal{M}_n \subset \mathbb{X}$ will be the manifold of all the realizations functions for a fixed neural-network architecture parametrized by n
- Warning: $\mathcal{M}_n \subset \mathbb{X}$ may not be convex and closed in the topology of $\mathbb{X}$!!

Neural quasi-minimization

- The idea is to minimize $j(\cdot)$ over a *computable* subset $\mathcal{M}_n \subset \mathbb{X}$
- $\mathcal{M}_n \subset \mathbb{X}$ will be the manifold of all the realizations functions for a fixed neural-network architecture parametrized by n
- Warning: $\mathcal{M}_n \subset \mathbb{X}$ may not be convex and closed in the topology of $\mathbb{X}$!!
- The cost functional $j(\cdot)$ may have an infimum when restricted to $\mathcal{M}_n$, but not a minimizer in $\mathcal{M}_n$

Neural quasi-minimization

- The idea is to minimize $j(\cdot)$ over a *computable* subset $\mathcal{M}_n \subset \mathbb{X}$
- $\mathcal{M}_n \subset \mathbb{X}$ will be the manifold of all the realizations functions for a fixed neural-network architecture parametrized by n
- Warning: $\mathcal{M}_n \subset \mathbb{X}$ may not be convex and closed in the topology of $\mathbb{X}$!!
- The cost functional $j(\cdot)$ may have an infimum when restricted to $\mathcal{M}_n$, but not a minimizer in $\mathcal{M}_n$

Definition (Quasi-minimization concepts; Shin, Zhang & Karniadakis)

Let $\{\mathcal{M}_n\}$ be a sequence of subsets of $\mathbb{X}$ and $\{\delta_n\} \to 0^+$. A quasi-minimizing sequence $\{\bar{\xi}_n\} \subset \mathbb{X}$ consists of quasi-minimizers $\bar{\xi}_n \in \mathcal{M}_n$ satisfying:

$$\boxed{j(\bar{\xi}_n) \leq \inf_{\xi_n \in \mathcal{M}_n} j(\xi_n) + \frac{\delta_n}{2}} \tag{1}$$

The problem to find $\bar{\xi}_n \in \mathcal{M}_n$ satisfying (1) will be our *quasi-minimizing control problem*.

Main result 1 (independent interest)

Theorem

Assume that $j : \mathbb{X} \to \mathbb{R}$ *is Gâteaux differentiable, with derivative* $j' : \mathbb{X} \to \mathbb{X}^*$ *satisfying for all* $\xi, \eta \in \mathbb{X}$*:*

- $\langle j'(\xi) - j'(\eta), \xi - \eta \rangle_{\mathbb{X}^*,\mathbb{X}} \geq \gamma \|\xi - \eta\|_{\mathbb{X}}^2$ *(strong convexity of* j*)*
- $\|j'(\xi) - j'(\eta)\|_{\mathbb{X}^*} \leq L \|\xi - \eta\|_{\mathbb{X}}$ *(Lipschitz continuity of* j'*)*

Then the following hold true

1. $j(\cdot)$ *has a unique minimizer* $\bar{\xi} \in \mathbb{X}$*, which satisfies* $j'(\bar{\xi}) = 0$ *in* $\mathbb{X}^*$
2. *For any* $\mathcal{M}_n \subset \mathbb{X}$ *and* $\delta_n > 0$*,* $j(\cdot)$ *has a quasi-minimizer* $\bar{\xi}_n \in \mathcal{M}_n$*.*
3. *Any quasi-minimizer satisfies the following quasi-optimal error estimate:*

$$\boxed{\|\bar{\xi} - \bar{\xi}_n\|_{\mathbb{X}}^2 \leq \frac{L}{\gamma} \inf_{\xi_n \in \mathcal{M}_n} \|\bar{\xi} - \xi_n\|_{\mathbb{X}}^2 + \frac{\delta_n}{\gamma}}$$

Main result 1 (independent interest)

Theorem

Assume that $j : \mathbb{X} \to \mathbb{R}$ *is Gâteaux differentiable, with derivative* $j' : \mathbb{X} \to \mathbb{X}^*$ *satisfying for all* $\xi, \eta \in \mathbb{X}$*:*

- $\langle j'(\xi) - j'(\eta), \xi - \eta \rangle_{\mathbb{X}^*,\mathbb{X}} \geq \gamma \|\xi - \eta\|_{\mathbb{X}}^2$ *(strong convexity of* j*)*
- $\|j'(\xi) - j'(\eta)\|_{\mathbb{X}^*} \leq L \|\xi - \eta\|_{\mathbb{X}}$ *(Lipschitz continuity of* j'*)*

Then the following hold true

1. $j(\cdot)$ *has a unique minimizer* $\bar{\xi} \in \mathbb{X}$*, which satisfies* $j'(\bar{\xi}) = 0$ *in* $\mathbb{X}^*$
2. *For any* $\mathcal{M}_n \subset \mathbb{X}$ *and* $\delta_n > 0$*,* $j(\cdot)$ *has a quasi-minimizer* $\bar{\xi}_n \in \mathcal{M}_n$*.*
3. *Any quasi-minimizer satisfies the following quasi-optimal error estimate:*

$$\boxed{\|\bar{\xi} - \bar{\xi}_n\|_{\mathbb{X}}^2 \leq \frac{L}{\gamma} \inf_{\xi_n \in \mathcal{M}_n} \|\bar{\xi} - \xi_n\|_{\mathbb{X}}^2 + \frac{\delta_n}{\gamma}}$$

Example 1 (PINNs): $j(\xi) = \frac{1}{2}\|f - B\xi\|_{\mathbb{L}}^2$

Main result 1 (independent interest)

Theorem

Assume that $j : \mathbb{X} \to \mathbb{R}$ *is Gâteaux differentiable, with derivative* $j' : \mathbb{X} \to \mathbb{X}^*$ *satisfying for all* $\xi, \eta \in \mathbb{X}$*:*

- $\langle j'(\xi) - j'(\eta), \xi - \eta \rangle_{\mathbb{X}^*,\mathbb{X}} \geq \gamma \|\xi - \eta\|_{\mathbb{X}}^2$ *(strong convexity of* j*)*
- $\|j'(\xi) - j'(\eta)\|_{\mathbb{X}^*} \leq L \|\xi - \eta\|_{\mathbb{X}}$ *(Lipschitz continuity of* j'*)*

Then the following hold true

1. $j(\cdot)$ *has a unique minimizer* $\bar{\xi} \in \mathbb{X}$*, which satisfies* $j'(\bar{\xi}) = 0$ *in* $\mathbb{X}^*$
2. *For any* $\mathcal{M}_n \subset \mathbb{X}$ *and* $\delta_n > 0$*,* $j(\cdot)$ *has a quasi-minimizer* $\bar{\xi}_n \in \mathcal{M}_n$*.*
3. *Any quasi-minimizer satisfies the following quasi-optimal error estimate:*

$$\boxed{\|\bar{\xi} - \bar{\xi}_n\|_{\mathbb{X}}^2 \leq \frac{L}{\gamma} \inf_{\xi_n \in \mathcal{M}_n} \|\bar{\xi} - \xi_n\|_{\mathbb{X}}^2 + \frac{\delta_n}{\gamma}}$$

Example 1 (PINNs): $j(\xi) = \frac{1}{2}\|f - B\xi\|_{\mathbb{L}}^2$

Example 2 (Deep Ritz method): $j(\xi) = \frac{1}{2} b(\xi, \xi) - f(\xi)$

Main result 2

Going back to our original problem ...

- Let $j(\xi) := j_1(\xi) + \alpha j_2(\xi) := \dfrac{1}{2}\|Q(u_{h,\xi}) - z_o\|_{\mathbb{Z}}^2 + \dfrac{\alpha}{2}\|\xi\|_{\mathbb{X}}^2$
- Let $S_h : \mathbb{X} \to \mathbb{U}_h$ be the control-to-state operator, i.e., $S_h(\xi) = u_{h,\xi}$

Main result 2

Going back to our original problem ...

- Let $j(\xi) := j_1(\xi) + \alpha j_2(\xi) := \frac{1}{2}\|Q(u_{h,\xi}) - z_o\|_{\mathbb{Z}}^2 + \frac{\alpha}{2}\|\xi\|_{\mathbb{X}}^2$
- Let $S_h : \mathbb{X} \to \mathbb{U}_h$ be the control-to-state operator, i.e., $S_h(\xi) = u_{h,\xi}$

Theorem

Assume $S_h(\cdot)$ differentiable, $S_h(\cdot)$ and $S_h'(\cdot)$ uniformly bounded on $\mathbb{X}$, and $S_h'(\cdot)$ Lipschitz continuous. Then

1. $j_1, j_2, j : \mathbb{X} \to \mathbb{R}$ *are Gâteaux differentiable with* $j_1', j_2', j' : \mathbb{X} \to \mathbb{X}^*$ *Lipschitz continuous.*
2. *For* $\alpha > 0$ *large enough (or* j_1 *convex),* $j(\cdot)$ *is strongly convex.*

Main result 2

Going back to our original problem ...

- Let $j(\xi) := j_1(\xi) + \alpha j_2(\xi) := \frac{1}{2}\|Q(u_{h,\xi}) - z_o\|_{\mathbb{Z}}^2 + \frac{\alpha}{2}\|\xi\|_{\mathbb{X}}^2$
- Let $S_h : \mathbb{X} \to \mathbb{U}_h$ be the control-to-state operator, i.e., $S_h(\xi) = u_{h,\xi}$

Theorem

Assume $S_h(\cdot)$ differentiable, $S_h(\cdot)$ and $S_h'(\cdot)$ uniformly bounded on $\mathbb{X}$, and $S_h'(\cdot)$ Lipschitz continuous. Then

1. *$j_1, j_2, j : \mathbb{X} \to \mathbb{R}$ are Gâteaux differentiable with $j_1', j_2', j' : \mathbb{X} \to \mathbb{X}^*$ Lipschitz continuous.*
2. *For $\alpha > 0$ large enough (or j_1 convex), $j(\cdot)$ is strongly convex.*

Notice that in our case ...

$$\begin{aligned} A(\xi)r + BS_h(\xi) &= f \\ B^* r &= 0 \end{aligned}$$

$$\begin{aligned} A(\xi)r'(\xi)\eta + BS_h'(\xi)\eta &= -[A'(\xi)\eta]r(\xi) \\ B^* r'(\xi)\eta &= 0 \end{aligned}$$

Hence, suitable conditions on $A(\cdot)$ will imply desired conditions on $S_h(\cdot)$, viz., $A(\cdot)$ Gateaux differentiable and uniformly bounded from above and below; $A'(\cdot)$ Lipschitz continuous and uniformly bounded.

Numerical Experiments

Point value control for weighted least-squares

Problem: $\begin{cases} u' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = 0 \end{cases}$ with $\lambda >> 1$.

Point value control for weighted least-squares

Problem: $\begin{cases} u' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} = L^2(0,1)$
- $(Bu, v)_{L^2} := \int_0^1 (u' + \lambda u)v = \lambda \int_0^1 v =: (f, v)_{L^2}$

Point value control for weighted least-squares

Problem: $\begin{cases} u' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} = L^2(0,1)$
- $(Bu, v)_{L^2} := \int_0^1 (u' + \lambda u) v = \lambda \int_0^1 v =: (f, v)_{L^2}$

Neurally-controlled discrete formulation:

$$\underbrace{\begin{array}{|llll|} \left(\dfrac{r}{\omega(\xi)}, v\right)_{L^2} + (Bu_{h,\xi}, v)_{L^2} & = (f, v)_{L^2} & \forall v \in \hat{\mathbb{V}} \\ (Bw_h, r)_{L^2} & = 0 & \forall w_h \in \mathbb{U}_h \end{array}}_{\text{mixed form}}$$

Point value control for weighted least-squares

Problem: $\begin{cases} u' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} = L^2(0,1)$
- $(Bu, v)_{L^2} := \int_0^1 (u' + \lambda u) v = \lambda \int_0^1 v =: (f, v)_{L^2}$

Neurally-controlled discrete formulation:

$$\underbrace{\begin{aligned} \left(\frac{r}{\omega(\xi)}, v\right)_{L^2} + (Bu_{h,\xi}, v)_{L^2} &= (f, v)_{L^2} & \forall v \in \hat{\mathbb{V}} \\ (Bw_h, r)_{L^2} &= 0 & \forall w_h \in \mathbb{U}_h \end{aligned}}_{\text{mixed form}} \qquad \underbrace{\begin{aligned} \left(\omega(\xi) Bw_h, f - Bu_{h,\xi}\right)_{L^2} = 0 \\ \forall w_h \in \mathbb{U}_h \end{aligned}}_{\text{weighted LS form}}$$

Point value control for weighted least-squares

Problem: $\begin{cases} u' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} = L^2(0,1)$
- $(Bu, v)_{L^2} := \int_0^1 (u' + \lambda u)v = \lambda \int_0^1 v =: (f, v)_{L^2}$

Neurally-controlled discrete formulation:

$$\underbrace{\begin{array}{|lll|} \left(\frac{r}{\omega(\xi)}, v\right)_{L^2} + (Bu_{h,\xi}, v)_{L^2} & = (f, v)_{L^2} & \forall v \in \hat{\mathbb{V}} \\ (Bw_h, r)_{L^2} & = 0 & \forall w_h \in \mathbb{U}_h \end{array}}_{\text{mixed form}} \quad \underbrace{\begin{array}{|c|} \Big(\omega(\xi) Bw_h, f - Bu_{h,\xi}\Big)_{L^2} = 0 \\ \forall w_h \in \mathbb{U}_h \end{array}}_{\text{weighted LS form}}$$

Cost functional:

$$j(\xi) := \frac{1}{2}|u(h) - u_{h,\xi}(h)|^2 \qquad \text{(assume that we know the value of } u(h))$$

Point value control for weighted least-squares

Weight: $\omega(\xi(x)) = 1 + \dfrac{M}{1+\exp(-\xi(x))}$ $\qquad (1 \le \omega \le 1 + M)$

ANN: $\mathcal{M}_8 := \left\{ \eta_8(x) = \sum_{j=1}^{8} c_j \,\mathrm{ReLU}(W_j x + b_j) \;\middle|\; c_j, W_j, b_j \in \mathbb{R} \right\}$

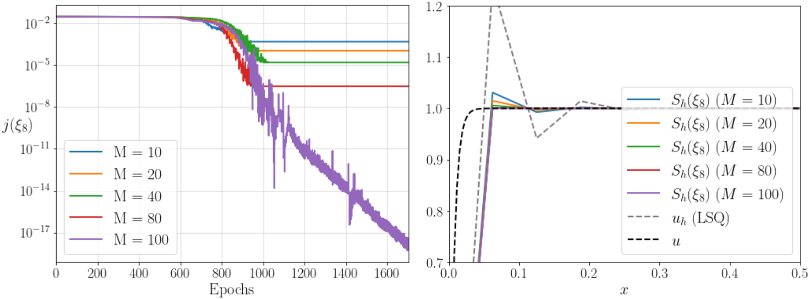


Figure: Point value control for weighted least-squares. Minimization of the cost functional for several values of M (left). Overshoot control of the discrete solutions (right).

Total variation control for weighted discrete-dual residual minimization

Problem: $\begin{cases} -u'' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \end{cases}$ with $\lambda >> 1$.

Total variation control for weighted discrete-dual residual minimization

Problem: $\begin{cases} -u'' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} \subset H^1_{(0}(0,1)$ conforming piecewise quadratics on same mesh
- $b(u,v) := \int_0^1 (u'v' + \lambda uv) = \lambda \int_0^1 v =: f(v)$

Total variation control for weighted discrete-dual residual minimization

Problem: $\begin{cases} -u'' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} \subset H^1_{(0}(0,1)$ conforming piecewise quadratics on same mesh
- $b(u,v) := \int_0^1 (u'v' + \lambda uv) = \lambda \int_0^1 v =: f(v)$

Neurally-controlled discrete formulation:

$$\begin{array}{lll} \left(\omega(\xi) r', v'\right)_{L^2} + b(u_{h,\xi}, v) & = f(v) & \forall v \in \hat{\mathbb{V}} \\ b(w_h, r) & = 0 & \forall w_h \in \mathbb{U}_h \end{array}$$

Total variation control for weighted discrete-dual residual minimization

Problem: $\begin{cases} -u'' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \end{cases}$ with $\lambda >> 1$.

Variational formulation:

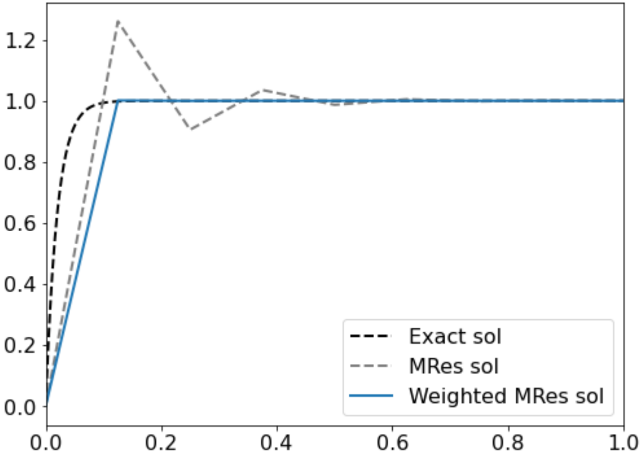
- Trial: $\mathbb{U}_h \subset H^1_{(0}(0,1)$ conforming piecewise linear on uniform mesh (size h)
- Test: $\hat{\mathbb{V}} \subset H^1_{(0}(0,1)$ conforming piecewise quadratics on same mesh
- $b(u,v) := \int_0^1 (u'v' + \lambda uv) = \lambda \int_0^1 v =: f(v)$

Neurally-controlled discrete formulation:

$$\begin{array}{|lll|} \hline \left(\omega(\xi)r', v'\right)_{L^2} + b(u_{h,\xi}, v) & = f(v) & \forall v \in \hat{\mathbb{V}} \\ b(w_h, r) & = 0 & \forall w_h \in \mathbb{U}_h \\ \hline \end{array}$$

Cost functional: $j(\xi) := \|u'_{h,\xi}\|_{L^1}$ (total variation)

Total variation control for weighted discrete-dual residual minimization



L^1-based control for overconstrained advection-reaction (Guermond)

Problem: $\begin{cases} \vec{\beta}\cdot\nabla u + u = 1 & \text{in } (0,1)^2 \\ u(0,x_2) = 0 & \end{cases}$ with $\vec{\beta} = (1,0)$.

L^1-based control for overconstrained advection-reaction (Guermond)

Problem: $\begin{cases} \vec{\beta} \cdot \nabla u + u = 1 & \text{in } (0,1)^2 \\ u(0, x_2) = 0 \end{cases}$ with $\vec{\beta} = (1, 0)$.

Overconstrained least-squares formulation:

FEM space: $\mathbb{U}_h \subset \{w \in H^1 : w(0, x_2) = w(1, x_2) = 0\}$ overconstrained piecewise linear on uniform mesh (size h)

$$u_h \in \mathbb{U}_h \quad \text{s.t.} \int_\Omega (1 - u_h - \vec{\beta} \cdot \nabla u_h)(\vec{\beta} \cdot \nabla w_h + w_h) = 0, \quad \forall w_h \in \mathbb{U}_h$$

L^1-based control for overconstrained advection-reaction (Guermond)

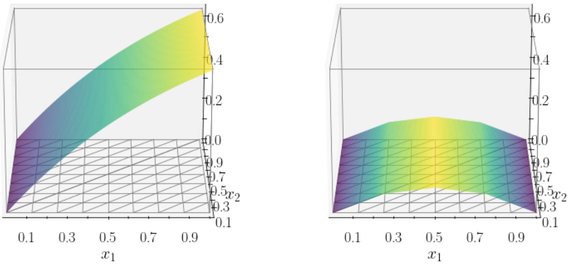
Problem: $\begin{cases} \vec{\beta}\cdot\nabla u + u = 1 & \text{in } (0,1)^2 \\ u(0,x_2) = 0 \end{cases}$ with $\vec{\beta} = (1,0)$.

Overconstrained least-squares formulation:

FEM space: $\mathbb{U}_h \subset \{w \in H^1 : w(0,x_2) = w(1,x_2) = 0\}$ overconstrained piecewise linear on uniform mesh (size h)

$$u_h \in \mathbb{U}_h \quad \text{s.t.} \int_\Omega (1 - u_h - \vec{\beta}\cdot\nabla u_h)(\vec{\beta}\cdot\nabla w_h + w_h) = 0, \quad \forall w_h \in \mathbb{U}_h$$

Exact v/s discrete solution:



L^1-based control for overconstrained advection-reaction (Guermond)

Overconstrained weighted least-squares formulation:

FEM space: $\mathbb{U}_h \subset \{w \in H^1 : w(0, x_2) = w(1, x_2) = 0\}$ overconstrained piecewise linear on uniform mesh (size h)

$$u_{h,\xi} \in \mathbb{U}_h \quad \text{s.t.} \int_\Omega \omega(\xi)(1 - u_{h,\xi} - \vec{\beta} \cdot \nabla u_{h,\xi})(\vec{\beta} \cdot \nabla w_h + w_h) = 0, \quad \forall w_h \in \mathbb{U}_h$$

L^1-based control for overconstrained advection-reaction (Guermond)

Overconstrained weighted least-squares formulation:
FEM space: $\mathbb{U}_h \subset \{w \in H^1 : w(0,x_2) = w(1,x_2) = 0\}$ overconstrained piecewise linear on uniform mesh (size h)

$$u_{h,\xi} \in \mathbb{U}_h \quad \text{s.t.} \int_\Omega \omega(\xi)(1 - u_{h,\xi} - \vec{\beta}\cdot\nabla u_{h,\xi})(\vec{\beta}\cdot\nabla w_h + w_h) = 0, \quad \forall w_h \in \mathbb{U}_h$$

Cost functional: $j(\xi) := \|1 - u_{h,\xi} - \vec{\beta}\cdot\nabla u_{h,\xi}\|_{L^1}$ (mimicking L^1 MinRes)

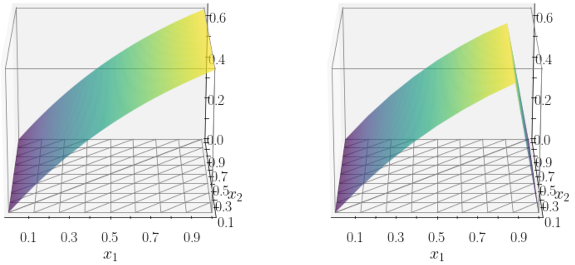
L^1-based control for overconstrained advection-reaction (Guermond)

Overconstrained weighted least-squares formulation:
FEM space: $\mathbb{U}_h \subset \{w \in H^1 : w(0,x_2) = w(1,x_2) = 0\}$ overconstrained piecewise linear on uniform mesh (size h)

$$u_{h,\xi} \in \mathbb{U}_h \quad \text{s.t.} \int_\Omega \omega(\xi)(1 - u_{h,\xi} - \vec{\beta}\cdot\nabla u_{h,\xi})(\vec{\beta}\cdot\nabla w_h + w_h) = 0, \quad \forall w_h \in \mathbb{U}_h$$

Cost functional: $j(\xi) := \|1 - u_{h,\xi} - \vec{\beta}\cdot\nabla u_{h,\xi}\|_{L^1}$ (mimicking L^1 MinRes)

Exact v/s discrete solution:



Bibliography

I. Brevis, I. M. & K.G. Van der Zee,
A machine-learning minimal-residual (ML-MRes) framework for goal-oriented finite element discretizations
Comput. Math. Appl., 95 (2021), pp.186–199.

I. Brevis, I. M. & K.G. Van der Zee,
Neural Control of Discrete Weak Formulations: Galerkin, Least-Squares & Minimal-Residual Methods with Quasi-Optimal Weights.
arXiv:2206.07475 (2022)

THX !!!

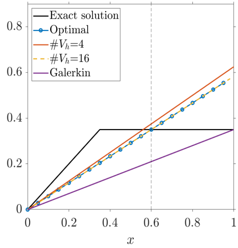
- European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 777778 (MATHROCKS).

MATH ROCKS

(bcam) basque center for applied mathematics

ignacio.muga@pucv.cl

Numerical experiments: 1D diffusion with one QoI

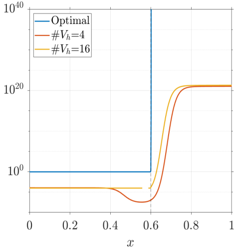


$$-u'' = \delta_\lambda \text{ in } (0,1), \qquad u(0) = u'(1) = 0$$

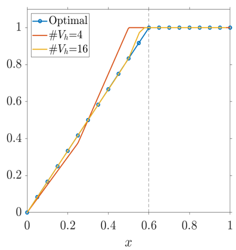
$$\mathbb{U}_h = \text{span}\{x\} \qquad \text{QoI} = 0.6$$

$$\text{ANN}(x;\theta) = \sum_{j=1}^{5} \theta_{j3}\sigma(\theta_{j1}x + \theta_{j2})$$

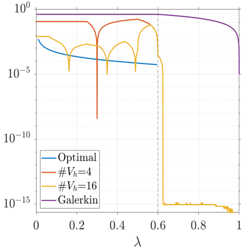
σ is the logistic sigmoid function.



(a) Trained weight



(b) Optimal test function



(c) Relative errors in QoI

Numerical experiments: 1D advection with one QoI

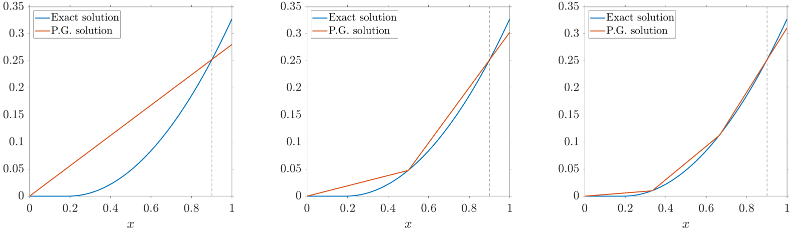
$$\begin{cases} u' = (x-\lambda)\mathbb{1}_{[\lambda,1]}(x) \\ u(0) = 0 \end{cases}$$

QoI$=0.9$

$$\text{ANN}(x;\theta) = \sum_{j=1}^{5} \theta_{j3}\sigma(\theta_{j1}x + \theta_{j2})$$

$\dim \mathbb{V}_h = 128$

Exact v/s Discrete solutions ($\lambda = 0.19$)



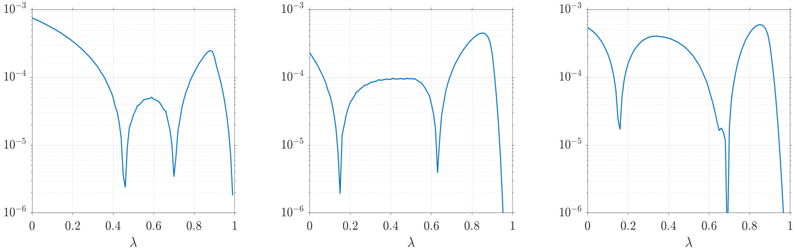
(a) One element (b) Two elements (c) Three elements

Numerical experiments: 1D advection with one QoI

$$\begin{cases} u' = (x-\lambda)\mathbb{1}_{[\lambda,1]}(x) \\ u(0) = 0 \end{cases} \qquad \text{QoI}=0.9 \qquad \lambda \in [0,1]$$

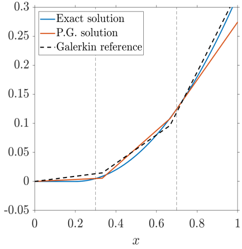
$$\mathrm{ANN}(x;\theta) = \sum_{j=1}^{5} \theta_{j3}\sigma(\theta_{j1}x + \theta_{j2}) \qquad \dim \mathbb{V}_h = 128$$

Absolute error in QoI

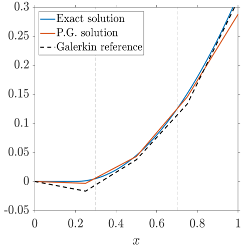


(a) One DoF (b) Two DoF (c) Three DoF

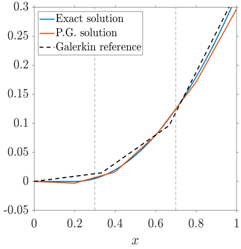
Numerical experiments: 1D advection with multiple QoIs



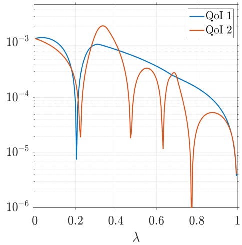
(a) Three elements



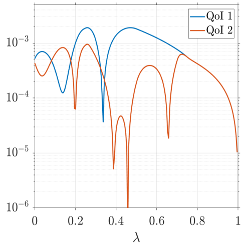
(b) Four elements



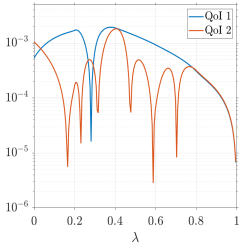
(c) Five elements



(a) Three DoF



(b) Four DoF



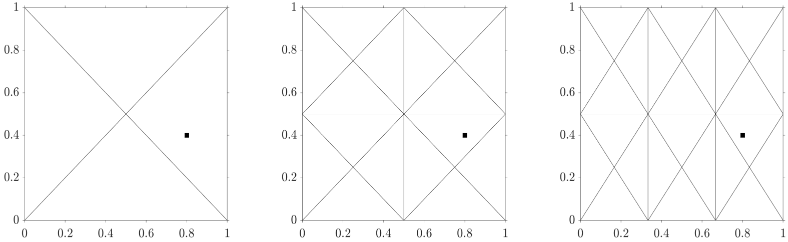
(c) Five DoF

Numerical experiments: 2D diffusion with one QoI

$$\begin{cases} -\Delta u_\lambda = f_\lambda & \text{in } \Omega = [0,1]^2 \\ u_\lambda = 0 & \text{over } \partial\Omega \end{cases}$$

f_λ is chosen such that the exact solution is:
$u_\lambda(x) = \sin(\pi x_1)\sin(\lambda\pi x_1)\sin(\pi x_2)\sin(\lambda\pi x_2)$

QoI: $$q(u) = \frac{1}{|\Omega_0|}\int_{\Omega_0} u\, dx$$

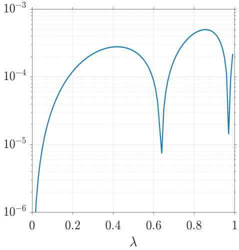


(a) One DoF (b) Five DoF (c) Eight DoF

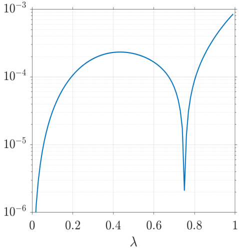
Numerical experiments: 2D diffusion with one QoI

$$\mathrm{ANN}(x_1, x_2; \theta) = \sum_{j=1}^{5} \theta_{j4}\sigma(\theta_{j1}x_1 + \theta_{j2}x_2 + \theta_{j3}) \qquad \dim \mathbb{V}_h = 1024$$

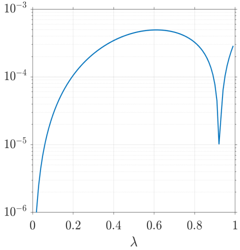
Absolute error in QoI



(a) One DoF



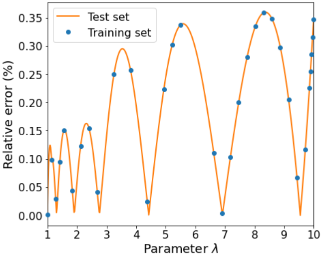
(b) Five DoF



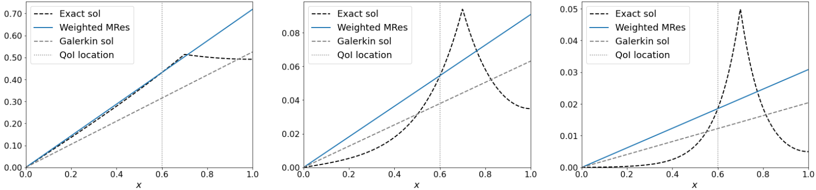
(c) Eight DoF

Numerical experiments: parameter λ on the left hand side

$$\begin{cases} -u'' + \lambda u = \delta_{x_0} & \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{QoI}=0.6 \\ x_0 = 0.7 \end{cases}$$



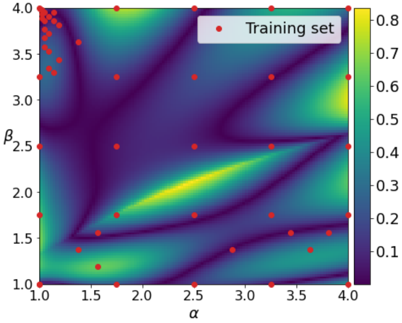
Exact v/s Discrete solutions



(a) $\lambda = 1$; rel. err. $= 0.00\%$ (b) $\lambda = 5.5$; rel. err. $= 0.34\%$ (c) $\lambda = 10$; rel. err. $= 0.35\%$

Numerical experiments: two parameters on the left hand side

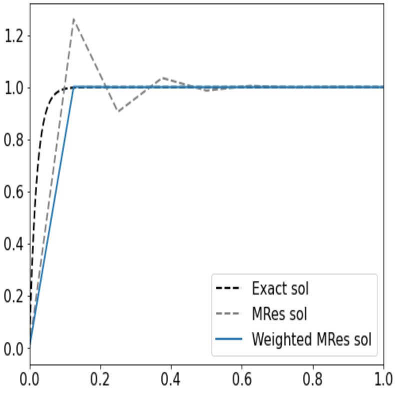
$$\begin{cases} -\alpha^2 u'' + \beta^2 u = \delta_{x_0} \quad \text{in } (0,1) \\ u(0) = u'(1) = 0 \\ \text{QoI}=0.6 \\ x_0 = 0.7 \end{cases}$$



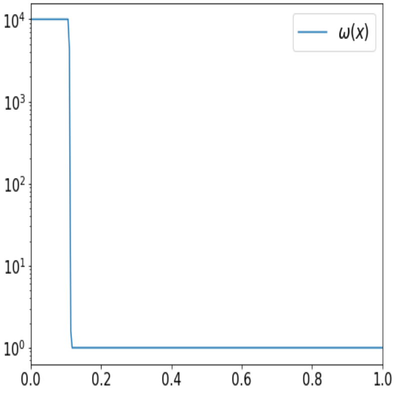
Numerical experiments: Diffusion–Reaction unsupervised training

$$\begin{cases} -u'' + \lambda u = \lambda & \text{in } (0,1) \\ u(0) = u'(1) = 0 \end{cases}$$

$$J(\omega) = \|u'_{h,\lambda,\omega}\|_{L^1}$$



(a) Exact v/s discrete solutions



(b) Trained weight