A priori and a posteriori analyses of the DPG method

Jay Gopalakrishnan

Portland State University

ICERM Workshop on
Robust Discretization and Fast Solvers for Computable Multi-Physics Models

Brown University, May 2013

Thanks: AFOSR, NSF
Principal Collaborator in DPG research: Leszek Demkowicz.

- Three avenues to DPG methods
  - 
  - 
  - 

- \textit{A priori} error analysis
  - 
  - 
  - 

- \textit{A posteriori} error analysis
- Fast solvers
- Examples
Three avenues to DPG methods

- Least-squares Galerkin method
- Petrov-Galerkin with optimal test space
- Mixed Galerkin method
"Petrov-Galerkin" schemes (PG)

PG schemes are distinguished by different trial and test (Hilbert) spaces.

The problem: \[
\text{P.D.E.} + \text{boundary conditions}.
\]

Variational form:

\[
\begin{aligned}
\text{Find } x \text{ in a trial space } X \text{ satisfying } \\
b(x, y) &= \ell(y) \\
\text{for all } y \text{ in a test space } Y.
\end{aligned}
\]

Discretization:

\[
\begin{aligned}
\text{Find } x_h \text{ in a discrete trial space } X_h \subset X \text{ satisfying } \\
b(x_h, y_h) &= \ell(y_h) \\
\text{for all } y_h \text{ in a discrete test space } Y_h \subset Y.
\end{aligned}
\]

For PG schemes, \( X_h \neq Y_h \) in general.
Elements of theory

- **Variational formulation:**

  \[
  C\|x\|_X \leq \sup_{y \in Y} \frac{|b(x, y)|}{\|y\|_Y} + \text{a uniqueness condition} \quad \Rightarrow \quad \text{wellposedness}
  \]

- **Babuška-Brezzi theory:**

  \[
  C\|x_h\|_X \leq \sup_{y_h \in Y_h} \frac{|b(x_h, y_h)|}{\|y_h\|_Y} \quad \Rightarrow \quad \|x - x_h\|_X \leq C \inf_{w_h \in X_h} \|x - w_h\|_X.
  \]

- **Difficulty:** Exact inf-sup condition \( \not\Rightarrow \) Discrete inf-sup condition
Elements of theory

- **Variational formulation:**

\[
\begin{align*}
C\|x\|_X &\leq \sup_{y \in Y} \frac{|b(x, y)|}{\|y\|_Y} \\
\text{Exact inf-sup condition} &+ \text{a uniqueness condition} \implies \text{wellposedness}
\end{align*}
\]

- **Babuška-Brezzi theory:**

\[
\begin{align*}
C\|x_h\|_X &\leq \sup_{y_h \in Y_h} \frac{|b(x_h, y_h)|}{\|y_h\|_Y} \\
\text{Discrete inf-sup condition} &\implies \|x - x_h\|_X \leq C \inf_{w_h \in X_h} \|x - w_h\|_X.
\end{align*}
\]

- **Difficulty:** Exact inf-sup condition $\not\Rightarrow$ Discrete inf-sup condition

Is there a way to find a stable **test** space for **any** given **trial** space (thus giving a stable method automatically)?
The ideal method

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$b(x_h, y) = \ell(y), \quad \forall y \in Y_h^{\text{opt}} \overset{\text{def}}{=} T(X_h),$$

where $T : X \mapsto Y$ is defined by

$$(Tw, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y.$$  

Rationale:

[Demkowicz+G 2011]
The ideal method

Pick any \( X_h \subseteq X \). The ideal DPG method finds \( x_h \in X_h \) such that

\[
  b(x_h, y) = \ell(y), \quad \forall y \in Y_{h_{\text{opt}}} \overset{\text{def}}{=} T(X_h),
\]

where \( T : X \mapsto Y \) is defined by

\[
  (T_w, y)_Y = b(w, y), \quad \forall w \in X, \; y \in Y.
\]

[Rationale:]

- **Q:** Which function \( y \) maximizes \( \frac{|b(x, y)|}{\|y\|_Y} \) for any given \( x \)?

- **A:** \( y = Tx \) is the maximizer. ← Optimal test function.

**DPG Idea:** If the discrete test space contains the optimal test functions,

exact inf-sup condition \( \implies \) discrete inf-sup condition.
The ideal method

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$b(x_h, y) = \ell(y), \quad \forall y \in Y_h^{\text{opt}} \overset{\text{def}}{=} T(X_h),$$

where $T : X \mapsto Y$ is defined by

$$(Tw, y)_Y = b(w, y), \quad \forall w \in X, y \in Y.$$

[A.1] $\{w \in X : b(w, y) = 0 \quad \forall y \in Y\} = \{0\}$.

[A.2] $\exists C_1, C_2 > 0$ such that $C_1\|y\|_Y \leq \sup_{w \in X} \frac{|b(w, y)|}{\|w\|_X} \leq C_2\|y\|_Y$.

Theorem (DPG Quasioptimality)

[A.1–A.2] $\implies \|x - x_h\|_X \leq \frac{C_2}{C_1} \inf_{w_h \in X_h} \|x - w_h\|_X$. 
The ideal method

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$b(x_h, y) = \ell(y), \quad \forall y \in Y_{h}^{\text{opt}} \overset{\text{def}}{=} T(X_h),$$

where $T : X \mapsto Y$ is defined by

$$(Tw, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y.$$

But ... can we really compute $Tx$?

- For a few problems, $Tx$ can be calculated in closed form.
- When $Tx$ cannot be hand calculated, we overcome two difficulties:
  - Redesign formulation so that $T$ is local (by hybridization).
  - Approximate $T$ by a computable (finite-rank) $T^r$. 
Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$b(x_h, y) = \ell(y), \quad \forall y \in Y_h^{\text{opt}} \overset{\text{def}}{=} T(X_h),$$

where $T : X \mapsto Y$ is defined by

$$(Tw, y)_Y = b(w, y), \quad \forall w \in X, \; y \in Y.$$
Trivial Example 1

**Standard FEM is an iDPG method**

Problem

Given \( F \in H^{-1}(\Omega) \), find \( u \in H^1_0(\Omega) \) solving:

\[
\int_{\Omega} \nabla u \cdot \nabla v = F(v), \quad \forall v \in H^1_0(\Omega).
\]

Recall

Pick any \( X_h \subseteq X \). The ideal DPG method finds \( x_h \in X_h \) such that

\[
b(x_h, y) = \ell(y), \quad \forall y \in Y_h^{\text{opt}} \overset{\text{def}}{=} T(X_h),
\]

where \( T : X \mapsto Y \) is defined by

\[
(T w, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y.
\]
Standard FEM is an iDPG method

Problem

Given $F \in H^{-1}(\Omega)$, find $u \in H^1_0(\Omega)$ solving:

$$\int_{\Omega} \nabla u \cdot \nabla v = F(v), \quad \forall v \in H^1_0(\Omega).$$

Set $X = Y = H^1_0(\Omega)$ and

$$(v, y)_Y = \int_{\Omega} \nabla v \cdot \nabla y.$$

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$b(x_h, y) = \ell(y), \quad \forall y \in Y_{h opt} \overset{\text{def}}{=} T(X_h),$$

where $T : X \mapsto Y$ is defined by

$$(Tw, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y.$$
Standard FEM is an iDPG method

Problem

Given \( F \in H^{-1}(\Omega) \),

find \( u \in H^1_0(\Omega) \) solving:

\[
\int_{\Omega} \nabla u \cdot \nabla v = F(v), \quad \forall v \in H^1_0(\Omega).
\]

\[\text{Set } X = Y = H^1_0(\Omega) \text{ and} \]

\[ (v, y)_Y = \int_{\Omega} \nabla v \cdot \nabla y. \]

Then \((\cdot, \cdot)_Y = b(\cdot, \cdot) \implies T = \text{identity, so} \]

\[ Y_{opt}^h = X_h. \]

Pick any \( X_h \subseteq X \). The ideal DPG method finds \( x_h \in X_h \) such that

\[ b(x_h, y) = \ell(y), \quad \forall y \in Y_{opt}^h \overset{\text{def}}{=} T(X_h), \]

where \( T : X \mapsto Y \) is defined by

\[ (Tw, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y. \]
Three avenues to DPG methods

▶ Petrov-Galerkin with optimal test functions
▶ Least-squares Galerkin method

A priori error analysis

▶ Ideal DPG method

A posteriori error analysis

Fast solvers

Examples

▶ Example 1 (Standard FEM)
**Trivial Example 2**

**L^2-based least squares method is an ideal DPG method**

Problem

Given an \( f \in L^2(\Omega) \) and a linear continuous bijective \( A : X \to L^2(\Omega) \), find \( u \in X \) satisfying \( Au = f \).

Recall

Pick any \( X_h \subseteq X \). The ideal DPG method finds \( x_h \in X_h \) such that

\[
b(x_h, y) = \ell(y), \quad \forall y \in Y_h^{\text{opt}} \overset{\text{def}}{=} T(X_h),
\]

where \( T : X \mapsto Y \) is defined by

\[
(Tw, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y.
\]
**L²-based least squares method is an ideal DPG method**

**Problem**

Given an \( f \in L^2(\Omega) \) and a linear continuous bijective \( A : X \to L^2(\Omega) \), find \( u \in X \) satisfying \( Au = f \).

- Set \( Y = L^2(\Omega) \), \( b(x, y) = (Ax, y)_Y \), \( \ell(y) = (f, y)_Y \).

Recall

Pick any \( X_h \subseteq X \). The ideal DPG method finds \( x_h \in X_h \) such that

\[
 b(x_h, y) = \ell(y), \quad \forall y \in Y_h^{\text{opt}} \overset{\text{def}}{=} T(X_h),
\]

where \( T : X \leftrightarrow Y \) is defined by

\[
 (Tw, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y.
\]
**Trivial Example 2**

**$L^2$-based least squares method is an ideal DPG method**

**Problem**

Given an $f \in L^2(\Omega)$ and a linear continuous bijective $A : X \rightarrow L^2(\Omega)$, find $u \in X$ satisfying $Au = f$.

- Set $Y = L^2(\Omega)$, $b(x, y) = (Ax, y)_Y$, $\ell(y) = (f, y)_Y$.
- Then $(Tw, y)_Y = (Aw, y) \implies T = A \implies Y_{h}^{\text{opt}} = AX_h \implies$ iDPG equations become Normal equations:

\[
(Ax_h, Aw_h)_Y = (f, Aw_h)_Y \quad \forall w_h \in X_h.
\]

**Recall**

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

\[
b(x_h, y) = \ell(y), \quad \forall y \in Y_{h}^{\text{opt}} \overset{\text{def}}{=} T(X_h),
\]

where $T : X \mapsto Y$ is defined by

\[
(Tw, y)_Y = b(w, y), \quad \forall w \in X, y \in Y.
\]
The least-squares avenue

- Least-squares Galerkin method
- DPG methods
- Mixed Galerkin method
- Petrov-Galerkin with optimal test space
Definitions

- **Riesz map:**

  \[ R_Y : Y \rightarrow Y^* : \quad (R_Y y)(v) = (y, v)_Y, \quad \forall y, v \in Y. \]

- **Operator generated by the form:**

  \[ B : X \rightarrow Y^* : \quad Bx(y) = b(x, y), \quad \forall x \in X, y \in Y. \]

- **Trial-to-Test operator** \( T : X \leftrightarrow Y \) **was defined by**

  \[ (Tw, y)_Y = b(w, y), \quad \forall w \in X, y \in Y. \]

  \[ \Longrightarrow \quad T = R_Y^{-1} \circ B. \]

- **Energy norm on** \( X \):

  \[ \|z\|_X \overset{\text{def}}{=} \|Tz\|_Y. \]
Theorem (DPG methods are least-squares methods)

The following are equivalent statements:

i) \( x_h \in X_h \) is the unique solution of the ideal DPG method.

ii) \( x_h \) is the best approximation to \( x \) from \( X_h \) in the energy norm:

\[
\| x - x_h \|_X = \inf_{z_h \in X_h} \| x - z_h \|_X
\]

iii) \( x_h \) minimizes residual in the following sense:

\[
x_h = \arg \min_{z_h \in X_h} \| \ell - Bz_h \|_{Y^*}.
\]

Proof of \((i) \iff (ii)\):

\[
b(x - x_h, y_h) = 0 \quad \forall y_h \in Y_h^{\text{opt}} \iff b(x - x_h, Tz_h) = 0 \quad \forall z_h \in X_h
\]

\[
\iff (T(x - x_h), Tz_h)_Y = 0 \quad \forall z_h \in X_h.
\]
Theorem (DPG methods are least-squares methods)

The following are equivalent statements:

i) \( x_h \in X_h \) is the unique solution of the ideal DPG method.

ii) \( x_h \) is the best approximation to \( x \) from \( X_h \) in the energy norm:

\[
\| x - x_h \|_X = \inf_{z_h \in X_h} \| x - z_h \|_X
\]

iii) \( x_h \) minimizes residual in the following sense:

\[
x_h = \arg \min_{z_h \in X_h} \| \ell - Bz_h \|_{Y^*}.
\]

Proof of (ii) \( \iff \) (iii).

\[
\| x - z_h \|_X = \| T(x - z_h) \|_Y = \| R_Y^{-1}B(x - z_h) \|_Y
\]

\[
= \| B(x - z_h) \|_{Y^*} = \| \ell - Bz_h \|_{Y^*}.
\]
Example 3: An ODE

Pavlovian integration by parts, or not?

1D transport eq.: \[ \begin{cases} u' = f & \text{in } (0, 1), \\ u(0) = u_0 & \text{(inflow b.c.)} \end{cases} \]

Variational form: Find \( u \) in \( H^1 \), satisfying \( u(0) = u_0 \), &

\[
\begin{aligned}
&\int_0^1 u' v = \int_0^1 f v, \\
&\quad b(u,v) - \int_0^1 l(v), \\
&\forall v \in L^2.
\end{aligned}
\]
Example 3: An ODE

Pavlovian integration by parts, or not?

1D transport eq. \[ u' = f \quad \text{in} \ (0, 1), \]
\[ u(0) = u_0 \quad \text{(inflow b.c.)} \]

Variational form: \[
\begin{array}{l}
\text{Find } u \text{ in } H^1, \text{ satisfying } u(0) = u_0, \& \\
\int_0^1 u' v = \int_0^1 f v, \quad \forall v \text{ in } L^2.
\end{array}
\]

Ultra-weak form: \[
\begin{array}{l}
\text{Find } u \in L^2, \text{ and a number } \hat{u}_1 \in \mathbb{R}, \text{ satisfying} \\
- \int_0^1 u v' + \hat{u}_1 v(1) = \int_0^1 f v + u_0 v(0), \quad \forall v \in H^1.
\end{array}
\]
Example 3: An ODE

**Pavlovian integration by parts, or not?**

1D transport eq.

\[
\begin{aligned}
\quad & u' = f \quad \text{in } (0, 1), \\
\quad & u(0) = u_0 \quad \text{(inflow b.c.)}
\end{aligned}
\]

**Variational form:**

Find \( u \) in \( H^1 \), satisfying \( u(0) = u_0, \& \)

\[
\int_0^1 u' v = \int_0^1 f v, \quad \forall v \in L^2.
\]

**Ultra-weak form:**

Find \( u \in L^2 \), and a number \( \hat{u}_1 \in \mathbb{R} \), satisfying

\[
- \int_0^1 u v' + \hat{u}_1 v(1) = \int_0^1 f v + u_0 v(0), \quad \forall v \in H^1.
\]
One-dimensional results using spectral trial space

Spectral Petrov-Galerkin solutions

- \( p=1 \)
- \( p=5 \)
- \( p=9 \)
- Exact solution

\( L^2 \) Least-Squares solutions

- \( p=1 \)
- \( p=5 \)
- \( p=9 \)
- Exact solution

[Click here to download FEniCS code.]
Three avenues to DPG methods
- Petrov-Galerkin with optimal test functions
- Least-squares Galerkin method
- Ideal DPG method
- Practical DPG method

A priori error analysis
- Ideal DPG method
- Practical DPG method

A posteriori error analysis

Fast solvers

Examples
- Example 1 (Standard FEM)
- Example 2 ($L^2$-based least-squares)
- Example 3 (An ODE)
### The practical method

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$b(x_h, y) = \ell(y), \quad \forall y \in Y_{h}^{\text{opt}} \overset{\text{def}}{=} T(X_h),$$

where $T : X \mapsto Y$ is defined by

$$(T_w, y)_Y = b(w, y), \quad \forall w \in X, y \in Y.$$

Pick any $X_h \subseteq X$. The practical DPG method finds $x_h^r \in X_h$, using a (finite-dimensional) $Y_r \subseteq Y$, such that

$$b(x_h^r, y) = \ell(y), \quad \forall y \in Y_{h}^r \overset{\text{def}}{=} T^r(X_h),$$

where $T^r : X \mapsto Y_r$ is defined by

$$(T^r_w, y)_Y = b(w, y), \quad \forall w \in X, y \in Y_r.$$
The practical method

Pick any $X_h \subseteq X$. The ideal DPG method finds $x_h \in X_h$ such that

$$x_h = \arg \min_{z_h \in X_h} \| \ell - Bz_h \|_{Y^*}.$$ 

Pick any $X_h \subseteq X$. The practical DPG method finds $x_h^r \in X_h$, using a (finite-dimensional) $Y^r \subseteq Y$, such that

$$x_h^r = \arg \min_{z_h \in X_h} \| \ell - Bz_h \|_{(Y^r)^*}.$$
Error analysis of the practical DPG method

[A.1] \( \{ w \in X : b(w, y) = 0 \ \forall y \in Y \} = \{0\} \).

[A.2] \( \exists C_1, C_2 > 0 \text{ such that } C_1 \|y\|_Y \leq \sup_{w \in X} \frac{|b(w, y)|}{\|w\|_X} \leq C_2 \|y\|_Y \).

[A.3] \( \exists \Pi : Y \mapsto Y^r \) and \( C_\Pi > 0 \) such that for all \( w_h \in X_h \) and \( y \in Y \),

\[ b(w_h, y - \Pi y) = 0, \quad \|\Pi y\|_Y \leq C_\Pi \|y\|_Y. \]

Theorem (A priori estimates for practical DPG method [G+Qiu 2013])

[A.1–A.3] \( \Rightarrow \|x - x_h^r\|_x \leq \frac{C_2 C_\Pi}{C_1} \inf_{w_h \in X_h} \|x - w_h\|_x. \)
The ‘D’ in ‘DPG’

For the residual minimization in

$$x_h = \arg \min_{z_h \in X_h} \| \ell - Bz_h \|_{Y^*}$$

to be feasible, the dual norm $\| \cdot \|_{Y^*}$ must be easily computable!

- “Negative-norm least-squares” uses multigrid or operators spectrally equivalent to the dual norm. [Bramble+Pasciak+Lazarov 1997]

- DPG methods reformulate problems to localize the dual norm computation (to parallel element-by-element computations). DPG methods have discontinuous test function space

$$Y = \prod_{K \in \text{mesh}} Y(K),$$

which have locally invertible Riesz maps.
Example 4: The Dirichlet problem

A new weak form for the old Laplacian

Find \( u \):
\[
-\Delta u = f, \quad \text{on } \Omega, \\
u = 0, \quad \text{on } \partial \Omega.
\]

Let \( \Omega_h \) be a mesh of \( \Omega \) and \( K \in \Omega_h \) be a mesh element. Then:
\[
\int_K \vec{\nabla} u \cdot \vec{\nabla} v - \int_{\partial K} (n \cdot \vec{\nabla} u) v = \int_K f v.
\]

This allows test function \( v \in Y \) to be in a “broken” Sobolev space
\[
Y = H^1(\Omega_h) := \prod_{K \in \Omega_h} H^1(K).
\]
Example 4: The Dirichlet problem

A new weak form for the old Laplacian

Find $u$: \[ -\Delta u = f, \quad \text{on } \Omega, \]
\[ u = 0, \quad \text{on } \partial \Omega. \]

Let $\Omega_h$ be a mesh of $\Omega$ and $K \in \Omega_h$ be a mesh element. Then:

\[ \int_K \nabla \cdot \nabla v - \int_{\partial K} (n \cdot \nabla u) v = \int_K f v. \]

\[ \sum_{K \in \Omega_h} \left[ \int_K \nabla \cdot \nabla v - \int_{\partial K} \hat{q}_n v \right] = \int_{\Omega} f v. \]

This allows test function $v \in Y$ to be in a “broken” Sobolev space

\[ Y = H^1(\Omega_h) := \prod_{K \in \Omega_h} H^1(K). \]
Functional setting for the Laplacian

Want $X$ and $Y$ to make $B : X \to Y^*$ a continuous bijection, i.e., the form

$$b(x, y) = (Bx)(y) \quad \text{on} \quad X \times Y$$

must satisfy a uniqueness and inf-sup condition.

- Set $b((u, \hat{q}_n), v) = \sum_{K \in \Omega_h} \left[ \int_K \vec{\nabla} u \cdot \vec{\nabla} v - \int_{\partial K} \hat{q}_n v \right]$.

- We seek $u$ in $H^1_0(\Omega)$ and $\hat{q}_n$ in $H^{-1/2}(\partial \Omega_h)$.
Functional setting for the Laplacian

Want $X$ and $Y$ to make $B : X \rightarrow Y^*$ a continuous bijection, i.e., the form

$$b(x, y) = (Bx)(y) \quad \text{on} \quad X \times Y$$

must satisfy a uniqueness and inf-sup condition.

- Set
  $$b\left((u, \hat{q}_n), v\right) = \sum_{K \in \Omega_h} \left[ \int_K \vec{\nabla} u \cdot \vec{\nabla} v - \int_{\partial K} \hat{q}_n v \right].$$

- We seek $u$ in $H^1_0(\Omega)$ and $\hat{q}_n$ in $H^{-1/2}(\partial \Omega_h)$.

**Definition (of $H^{-1/2}(\partial \Omega_h)$, the space of numerical fluxes)**

Define the element-by-element trace operator $\text{tr}_n$ by

$$\text{tr}_n : H(\text{div}, \Omega) \rightarrow \prod_{K \in \Omega_h} H^{-1/2}(\partial K), \quad \text{tr}_n r|_{\partial K} = r \cdot n|_{\partial K}.$$ 

and set $H^{-1/2}(\partial \Omega_h) = \text{ran}(\text{tr}_n)$. 
Functional setting for the Laplacian

Want $X$ and $Y$ to make $B : X \rightarrow Y^*$ a continuous bijection, i.e., the form

$$b(x, y) = (Bx)(y) \text{ on } X \times Y$$

must satisfy a uniqueness and inf-sup condition.

- Set $b((u, \hat{q}_n), v) = \sum_{K \in \Omega_h} \left[ \int_K \vec{\nabla} u \cdot \vec{\nabla} v - \int_{\partial K} \hat{q}_n v \right]$.
- We seek $u$ in $H^1_0(\Omega)$ and $\hat{q}_n$ in $H^{-1/2}(\partial \Omega_h)$.

Theorem

With $X = H^1_0(\Omega) \times H^{-1/2}(\partial \Omega_h)$ and $Y = H^1(\Omega_h)$, the operator $B$ is a continuous bijection and has a continuous inverse.

[Demkowicz+G 2013]
Discrete spaces for the Laplacian

Trial subspace \( X_h \subseteq X \equiv H_0^1(\Omega) \times H^{-1/2}(\partial \Omega_h) \): Approximate

\[ u \text{ by Lagrange FE of degree } \leq p + 1, \quad \forall \ K \in \Omega_h, \]
\[ \hat{q}_n \text{ by polynomials of degree } \leq p, \quad \forall \text{ mesh edges}. \]

Test subspace \( Y^r \subseteq H^1(\Omega_h) \): Set, for some \( r \geq 0 \),

\[ Y^r = \{ v : v|_K \in P_r(K), \quad \forall K \in \Omega_h \}. \]
Discrete spaces for the Laplacian

Trial subspace \( X_h \subseteq X \equiv H^1_0(\Omega) \times H^{-1/2}(\partial \Omega_h) \): Approximate

\[
\begin{align*}
    u & \quad \text{by Lagrange FE of degree } \leq p + 1, \quad \forall K \in \Omega_h, \\
    \hat{q}_n & \quad \text{by polynomials of degree } \leq p, \quad \forall \text{ mesh edges}.
\end{align*}
\]

Test subspace \( Y^r \subseteq H^1(\Omega_h) \): Set, for some \( r \geq 0 \),

\[
Y^r = \{ v : v|_K \in P_r(K), \quad \forall K \in \Omega_h \}.
\]

Pick any \( X_h \subseteq X \). The practical DPG method finds \( x^r_h \in X_h \), using a (finite-dimensional) \( Y^r \subseteq Y \), such that

\[
b(x^r_h, y) = \ell(y), \quad \forall y \in Y^r_h \overset{\text{def}}{=} T^r(X_h),
\]

where \( T^r : X \mapsto Y^r \) is defined by

\[
(T^r w, y)_Y = b(w, y), \quad \forall w \in X, \ y \in Y^r.
\]
Discrete spaces for the Laplacian

Trial subspace $X_h \subseteq X \equiv H^1_0(\Omega) \times H^{-1/2}(\partial \Omega_h)$: Approximate

\[ u \text{ by Lagrange FE of degree } \leq p + 1, \quad \forall K \in \Omega_h, \]
\[ \hat{q}_n \text{ by polynomials of degree } \leq p, \quad \forall \text{ mesh edges}. \]

Test subspace $Y^r \subseteq H^1(\Omega_h)$: Set, for some $r \geq 0$,

\[ Y^r = \{ \nu : \nu|_K \in P_r(K), \quad \forall K \in \Omega_h \}. \]

Computation of $T^r$ is local:

Apply: \((T^r w, y)_Y = b(w, y)\)

\[ \Rightarrow (T^r(u, \hat{q}_n), y)_{H^1(\Omega_h)} = b((u, \hat{q}_n), y), \quad \forall y \in Y^r. \]

\[ \Rightarrow (T^r(u, \hat{q}_n), y)_{H^1(K)} = \int_{K} \tilde{\nabla} u \cdot \tilde{\nabla} y - \int_{\partial K} \hat{q}_n y, \quad \forall K \in \Omega_h. \]
Discrete spaces for the Laplacian

Trial subspace $X_h \subseteq X \equiv H^1_0(\Omega) \times H^{-1/2}(\partial \Omega_h)$: Approximate

- $u$ by Lagrange FE of degree $\leq p + 1$, $\forall K \in \Omega_h$,
- $\hat{q}_n$ by polynomials of degree $\leq p$, $\forall$ mesh edges.

Test subspace $Y^r \subseteq H^1(\Omega_h)$: Set, for some $r \geq 0$,

$$Y^r = \{ v : v|_K \in P_r(K), \forall K \in \Omega_h \}.$$ 

To prove optimal convergence, we must choose $r$ so that [A.3] holds.

[A.3] $\exists \Pi : Y \mapsto Y^r$ and $C_\Pi > 0$ such that for all $w_h \in X_h$ and $y \in Y$,

$$b(w_h, y - \Pi y) = 0, \quad \|\Pi y\|_Y \leq C_\Pi \|y\|_Y.$$
Discrete spaces for the Laplacian

Trial subspace $X_h \subseteq X \equiv H^1_0(\Omega) \times H^{-1/2}(\partial \Omega_h)$: Approximate

- $u$ by Lagrange FE of degree $\leq p + 1$, $\forall K \in \Omega_h$,
- $\hat{q}_n$ by polynomials of degree $\leq p$, $\forall$ mesh edges.

Test subspace $Y^r \subseteq H^1(\Omega_h)$: Set, for some $r \geq 0$,

$$Y^r = \{ v : v|_K \in P_r(K), \forall K \in \Omega_h \}.$$

**Theorem (Verification of [A.3])**

Let $\Omega_h$ be a simplicial shape-regular finite element mesh in $N$-space dimensions. For any $p \geq 0$, whenever $r \geq p + N$, there exists a continuous $\Pi : Y \to Y^r$ such that for all $(w_h, \hat{s}_{n,h}) \in X_h$,

$$\int_K \nabla w_h \cdot \nabla (v - \Pi v) - \int_{\partial K} \hat{s}_{n,h} (v - \Pi v) = 0, \quad \forall K \in \Omega_h.$$
Three avenues to DPG methods

- Petrov-Galerkin with optimal test functions
- Least-squares Galerkin method

A priori error analysis

- Ideal DPG method
- Practical DPG method

A posteriori error analysis

Fast solvers

Examples

- Example 1 (Standard FEM)
- Example 2 ($L^2$-based least-squares)
- Example 3 (An ODE)
- Example 4 (Diffusion)
Preconditioning

Abstractly, \( b(x_h^r, y) = \ell(y) \quad \forall y \in Y_h^r = T^r(X_h), \)

\[
\Rightarrow \quad b(x_h^r, T^r z_h) = \ell(T^r z_h) \quad \forall z_h \in X_h
\]

\[
\Rightarrow \quad (T^r x_h^r, T^r z_h)_Y = \ell(T^r z_h) \quad \forall z_h \in X_h.
\]

**Lemma**

\[ [A.1–A.3] \quad \Rightarrow \quad \frac{C_1}{C_{II}} \|x\|_X \leq \|T^r x\|_Y \leq C_2 \|x\|_X \]

for all \( x \in X_h. \)

This implies that any preconditioner spectrally equivalent to the \((\cdot, \cdot)_X\)-inner product is also a preconditioner for the practical DPG method.
Example: A BDDC preconditioner

\[ b( (u, \hat{q}_n), v ) = \sum_{K \in \Omega_h} \left[ \int_K \vec{\nabla} u \cdot \vec{\nabla} v - \int_{\partial K} \hat{q}_n v \right] \]

\[ X = H^1_0(\Omega) \times H^{-1/2}(\partial \Omega_h), \]

Implementation in NGSolve with Lukas Kogler & Joachim Schöberl

1. **Statically condense** the stiffness matrix to \( u|_{\partial \Omega_h} \) and \( \hat{q}_n \).
2. **Apply a BDDC preconditioner** as follows:
   1. Do a wire basket coarse solve.
   2. Add inverses of small blocks of \( u|_{\partial \Omega_h} \)-unknowns on each interface.
   3. Add inverses of small blocks of \( \hat{q}_n \)-unknowns on each interface.
Example: A BDDC preconditioner

\[ b( (u, \hat{q}_n), v) = \sum_{K \in \Omega} \left[ \int_K \vec{\nabla} u \cdot \vec{\nabla} v - \int_{\partial K} \hat{q}_n v \right] \]

\[ X = H_0^1(\Omega) \times H^{-1/2}(\partial \Omega_h), \]

Implementation in NGSolve with Lukas Kogler & Joachim Schöberl

<table>
<thead>
<tr>
<th>( p + 1 )</th>
<th>diagonal</th>
<th>BDDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>142</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>159</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>180</td>
<td>77</td>
</tr>
<tr>
<td>7</td>
<td>202</td>
<td>78</td>
</tr>
<tr>
<td>8</td>
<td>209</td>
<td>88</td>
</tr>
<tr>
<td>9</td>
<td>243</td>
<td>90</td>
</tr>
</tbody>
</table>

- Used a small fixed 8 x 8 mesh
- Number of preconditioned conjugate gradient iterations are reported.
Three avenues to DPG methods
  ▶ Petrov-Galerkin with optimal test functions
  ▶ Least-squares Galerkin method
  ▶ Mixed Galerkin method

A priori error analysis
  ▶ Ideal DPG method
  ▶ Practical DPG method

A posteriori error analysis

Fast solvers

Examples
  ▶ Example 1 (Standard FEM)
  ▶ Example 2 ($L^2$-based least-squares)
  ▶ Example 3 (An ODE)
  ▶ Example 4 (Diffusion)
Built-in error estimator in DPG methods

Results for Carter’s flat plate problem: (courtesy of Jesse Chan)

Adaptivity shows no preasymptotics.

Iteration 0

Supersonic flow impinging over a flat plate (Ma = 3, Re = 1000). Used Petrov-Galerkin implementation in Camillia package with $h$-adaptivity, $p = 2$, starting with a mesh of just two elements.
Results for Carter’s flat plate problem: (courtesy of Jesse Chan)

Adaptivity shows no preasymptotics.

Supersonic flow impinging over a flat plate (Ma = 3, Re = 1000). Used Petrov-Galerkin implementation in Camillia package with $h$-adaptivity, $p = 2$, starting with a mesh of just two elements.
Results for Carter’s flat plate problem: (courtesy of Jesse Chan)

Adaptivity shows no preasymptotics.

Supersonic flow impinging over a flat plate (Ma = 3, Re = 1000). Used Petrov-Galerkin implementation in Camillia package with $h$-adaptivity, $p = 2$, starting with a mesh of just two elements.
The mixed method approach

- Least-squares Galerkin method
- Petrov-Galerkin with optimal test space
- Mixed Galerkin method

DPG methods
Error representation function

**Residual:** \( \rho = \ell - Bx_h. \)

**Error representation function:** \( \varepsilon^r = R_{Yr}^{-1}(\ell - Bx_h). \)

It can be practically computed by
\[
(\varepsilon^r, y)_{Y} = \ell(y) - b(x_h, y), \quad \forall y \in Y^r.
\]

**Error estimator:** \( \eta = \|\varepsilon^r\|_Y. \)  

---

Petrov-Galerkin solve \( \rightarrow \) \( \varepsilon^r \) by local postprocessing

Least-squares \( \rightarrow \) \( \varepsilon^r \) is Riesz inverse of residual

Mixed method \( \rightarrow \) \( \varepsilon^r \) is one of the variables

[Demkowicz+G+Niemi 2012]
Theorem (Reinterpretation of DPG as a mixed method)

The following are equivalent statements:

i) \( x_h \in X_h \) solves the practical DPG method.

ii) \( x_h \in X_h \) and \( \varepsilon^r \in Y^r \) solve the mixed formulation

\[
(\varepsilon^r, y)_Y + b(x_h, y) = \ell(y) \quad \forall y \in Y^r, \quad (1a)
\]

\[
b(z_h, \varepsilon^r) = 0 \quad \forall z_h \in X_h. \quad (1b)
\]

Proof.

\((i) \implies (ii)\): Eq. (1a) is just the definition of \( \varepsilon^r \). For (1b),

\[
b(z_h, \varepsilon^r) = (T^r z_h, \varepsilon^r)_Y = (T^r z_h, R_{Y^r}^{-1}(\ell - Bx_h))_Y = (T^r z_h, T^r(x - x_h))_Y
\]

\[
= b(x - x_h, T^r z_h) = 0.
\]

\((ii) \implies (i)\): Similar.
Theorem (Reinterpretation of DPG as a mixed method)

The following are equivalent statements:

i) \( x_h \in X_h \) solves the practical DPG method.

ii) \( x_h \in X_h \) and \( \varepsilon^r \in Y^r \) solve the mixed formulation

\[
(\varepsilon^r, y)_{Y} + b(x_h, y) = \ell(y) \quad \forall y \in Y^r, \tag{1a}
\]

\[
b(z_h, \varepsilon^r) = 0 \quad \forall z_h \in X_h. \tag{1b}
\]

[Dahmen+Huang+Schwab+Welper 2012] studied similar mixed formulations and found techniques other than localization by discontinuous spaces to make the method practical.
Recall the previous assumptions

[A.1] \( \{ w \in X : b(w, y) = 0 \ \forall y \in Y \} = \{0\} \).

[A.2] \( \exists C_1, C_2 > 0 \) such that \( C_1 \|y\|_Y \leq \sup_{w \in X} \left| \frac{b(w, y)}{\|w\|_X} \right| \leq C_2 \|y\|_Y \).

[A.3] \( \exists \Pi : Y \mapsto Y' \) and \( C_\Pi > 0 \) such that for all \( w_h \in X_h \) and \( y \in Y \),

\[
b(w_h, y - \Pi y) = 0, \quad \|\Pi y\|_Y \leq C_\Pi \|y\|_Y.
\]

Optimal a priori estimates followed from these assumptions.

*We now show that a posteriori error estimators also follow from the same assumptions [A.1–A.3].*
A posteriori error estimates

**Theorem (Reliability & Efficiency of DPG error estimator)**

Suppose [A.1–A.3] hold. Let \( F \in Y^* \),

- \( x = B^{-1}F \),
- \( x_h \in X_h \) be the DPG solution,
- \( \eta = \| F - Bx_h \| (Y^r)^* = \| \varepsilon^r \|_Y \) be the error estimator,
- \( \text{osc}(F) \overset{\text{def}}{=} \| F \circ (1 - II) \|_{Y^*} \).

Then

\[
C_1^2 \| x - x_h \|_X^2 \leq \eta^2 + (C_{II} \eta + \text{osc}(F))^2 , \quad \leftarrow \text{Reliability}
\]

\[
\eta^2 \leq C_2^2 \| x - x_h \|_X^2 . \quad \leftarrow \text{Efficiency}
\]

- “Efficiency” is trivial in least-square methods.
- Proof of “Reliability” uses \( II \) critically.
A posteriori error estimates

Theorem (Reliability & Efficiency of DPG error estimator)

Suppose [A.1–A.3] hold. Let $F \in Y^*$,

- $x = B^{-1}F$,
- $\tilde{x}_h \in X_h$ be the DPG solution,
- $\tilde{\eta} = \|F - B\tilde{x}_h\| (Y^r)^* = \|\varepsilon^r\|_Y$ be the error estimator,
- $\text{osc}(F) \overset{\text{def}}{=} \|F \circ (1 - \Pi)\|_Y^*$.

Then

$$C_1^2 \|x - \tilde{x}_h\|_X^2 \leq \tilde{\eta}^2 + (C_\Pi \tilde{\eta} + \text{osc}(F))^2,$$

$\leftarrow$ Reliability

$$\tilde{\eta}^2 \leq C_2^2 \|x - \tilde{x}_h\|_X^2.$$

$\leftarrow$ Efficiency

- “Efficiency” is trivial in least-square methods.
- Proof of “Reliability” uses $\Pi$ critically.
Results for Dirichlet problem with \( f(x, y) = e^{-100(x^2+y^2)} \) aside.

No need to code an error estimator for driving adaptivity in DPG methods.

The mixed formulation is standard Galerkin, so it is easily implementable in codes without support for Petrov-Galerkin forms.

[Click here to download FEniCS code for this experiment.]
Error estimator in the Laplace example

- Results for Dirichlet problem with $f(x, y) = e^{-100(x^2+y^2)}$ aside.

- No need to code an error estimator for driving adaptivity in DPG methods.

- The mixed formulation is standard Galerkin, so it is easily implementable in codes without support for Petrov-Galerkin forms.

[Click here to download FEniCS code for this experiment.]
Error estimator in the Laplace example

Results for Dirichlet problem with $f(x, y) = e^{-100(x^2+y^2)}$ aside.

No need to code an error estimator for driving adaptivity in DPG methods.

The mixed formulation is standard Galerkin, so it is easily implementable in codes without support for Petrov-Galerkin forms.

[Click here to download FEniCS code for this experiment.]
Example 5: Stresses in Stokes flow

Second order system:

\[
\frac{1}{2} \Delta \vec{u} - \vec{\nabla} p = \vec{f} \quad \text{in } \Omega, \\
\nabla \cdot \vec{u} = 0 \quad \text{in } \Omega.
\]

No slip B.C.: \( \vec{u} = \vec{0} \) on \( \partial \Omega \).

For uniqueness: \((p, 1)_\Omega = 0\).

Convert to first order system:

\[
\sigma + p \delta - \varepsilon(\vec{u}) = 0, \quad \text{(definition of true fluid stress } \sigma) \\
\nabla \cdot \sigma = \vec{f}. \quad \text{(since } \nabla \cdot \sigma = \frac{1}{2} \Delta \vec{u} - \vec{\nabla} p)\]
Example 5: Stresses in Stokes flow

Second order system:

\[
\frac{1}{2} \Delta \vec{u} - \nabla p = \vec{f} \quad \text{in } \Omega, \\
\nabla \cdot \vec{u} = 0 \quad \text{in } \Omega.
\]

No slip B.C.: \( \vec{u} = \vec{0} \) on \( \partial \Omega \).
For uniqueness: \( (p, 1)_{\Omega} = 0 \).

Convert to first order system:

Apply deviatoric:

\[
D\tau = \tau - \frac{\text{tr } \tau}{N} \delta
\]

\[
\sigma + p \delta - \varepsilon(\vec{u}) = 0, \\
\nabla \cdot \sigma = \vec{f}.
\]
Example 5: Stresses in Stokes flow

Second order system:

\[ \frac{1}{2} \Delta \vec{u} - \nabla p = \vec{f} \quad \text{in } \Omega, \]
\[ \nabla \cdot \vec{u} = 0 \quad \text{in } \Omega. \]

No slip B.C.: \( \vec{u} = \vec{0} \) on \( \partial \Omega \).
For uniqueness: \( (p, 1)_\Omega = 0 \).

Convert to first order system:

Apply deviatoric:

\[ D\tau = \tau - \frac{\text{tr } \tau}{N} \delta \]
\[ D\sigma - \varepsilon(\vec{u}) = 0, \]
\[ \nabla \cdot \sigma = \vec{f}. \quad \text{And} \quad (\text{tr } \sigma, 1)_\Omega = 0. \]
Example 5: Stresses in Stokes flow

Second order system:

\[ \frac{1}{2} \Delta \mathbf{u} - \nabla p = \mathbf{f} \quad \text{in } \Omega, \]
\[ \nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega. \]

No slip B.C.: \( \mathbf{u} = \mathbf{0} \) on \( \partial \Omega \).
For uniqueness: \( (p, 1)_\Omega = 0 \).

Convert to first order system:

\[ D\sigma - \varepsilon(\mathbf{u}) = 0, \]
\[ \nabla \cdot \sigma = \mathbf{f}. \]

And \( (\text{tr } \sigma, 1)_\Omega = 0 \).

DPG form with \( x = (\sigma, \mathbf{u}, \mathbf{h}, \hat{\sigma}_n, \alpha) \) and \( y = (\tau, \overline{\mathbf{v}}, \omega) \):

\[ b(x, y) = (D\sigma, \tau)_\Omega + (\mathbf{u}, \nabla \cdot \tau)_\Omega + \langle \mathbf{h}, \tau n \rangle_{\partial \Omega} + (\alpha, \text{tr } \tau)_\Omega \]
\[ + (\sigma, \varepsilon(\overline{\mathbf{v}}))_{\Omega_h} - \langle \hat{\sigma}_n, \overline{\mathbf{v}} \rangle_{\partial \Omega_h} + (\text{tr } \sigma, \omega)_\Omega. \]
**Spaces for Stokes example**

**DPG form with** $x = (\sigma, \vec{u}, \hat{u}, \hat{\sigma}_n, \alpha)$ and $y = (\tau, \vec{v}, \omega)$:

$$b(x, y) = (D\sigma, \tau)_\Omega + (\vec{u}, \nabla \cdot \tau)_{\Omega_h} - \langle \hat{u}, \tau n \rangle_{\partial \Omega_h} + (\alpha, \text{tr} \ \tau)_{\Omega}$$

$$+ (\sigma, \varepsilon(\vec{v}))_{\Omega_h} - \langle \hat{\sigma}_n, \vec{v} \rangle_{\partial \Omega_h} + (\text{tr} \ \sigma, \omega)_{\Omega}.$$  

**Trial and test spaces:**

$$X = L^2(\Omega; S) \times L^2(\Omega)^N \times H^{1/2}(\partial \Omega_h)^N \times H^{-1/2}(\partial \Omega_h)^N \times \mathbb{R},$$

$$Y = H(\text{div}, \Omega_h; S) \times H^1(\Omega_h)^N \times \mathbb{R}.$$  

**Discrete spaces:**

$$X_h = \{ (\sigma, \vec{u}, \hat{u}, \hat{\sigma}_n, \alpha) \in X : \sigma|_K \in P_p(K; S), \vec{u}|_K \in P_p(K)^N, \forall \text{elements } K,$$

$$\hat{u}|_F \in P_{p+1}(F)^N, \hat{\sigma}_n|_F \in P_p(F)^N, \forall \text{interfaces } F, \alpha \in \mathbb{R} \},$$

$$Y^r = \{ (\tau, \vec{v}, \omega) \in Y : \omega \in \mathbb{R},$$

$$\tau|_K \in P_{p+2}(K; S), \vec{v}|_K \in P_{p+N}(K)^N, \forall \text{elements } K \}.$$
Suppose $\Omega_h$ is a shape-regular simplicial mesh of $\Omega$ and $p \geq 0$. Then [A.1–A.3] holds for the Stokes example. Consequently, $\exists$ mesh-independent constants $c_1, \ldots, c_4 > 0$ such that

$$
\| x - x_h \|_X \leq c_1 \min_{\xi_h \in X_h} \| x - \xi_h \|_X,
$$

$$
c_4 \| x - x_h \|_X^2 - c_2 \text{osc}(F)^2 \leq \eta^2 \leq c_3 \| x - x_h \|_X^2.
$$

- Verification of [A.3] uses degrees of freedom of symmetric matrix polynomials in [G+Guzmán 2011].
- Proof proceeds by taking the incompressible limit of a similar elasticity discretization.
Osborn’s singular solution:

\[ u = \text{curl} \left( a_+ s_+ + a_- s_- + c_+ - c_- \right), \]

where

\[ s_\pm = r^{1+z} \sin((z \pm 1)\theta), \quad c_\pm = r^{1+z} \cos((z \pm 1)\theta), \]
\[ a_\pm = -z \cot(3z\pi/2)/(z \pm 1), \]
\[ z^2 = \sin^2 (3z\pi/2) \quad [z = \text{root with smallest real part}]. \]

Results from an h-adaptive algorithm with \( \eta \) as estimator and \( p = 2 \):
Stokes solution on L-shaped domain

Stokes example: h–adaptivity on L–shaped domain

![Graph](image)

Error $\|x - x_h\|_X$ and estimator $\eta$

$\sigma_{xy}$

$\sigma_{xx}$

$\sigma_{yy}$

Jay Gopalakrishnan
Stokes solution on L-shaped domain

Effectivity during the adaptive process

Effectivity index $\rho$. 

$$\rho = \frac{\eta}{\|x - x_h\|_X}$$
Stokes solution on L-shaped domain

Effectivity for perturbed adaptive iterates

Effectivity $\tilde{\rho}$ vs. # Degrees of freedom

After $x_h$ randomly perturbed by 5%.

$\tilde{\rho} = \frac{\eta}{\|x - \tilde{x}_h\|_X}$

Jay Gopalakrishnan
Conclusion

- Three avenues to DPG methods
  - Petrov-Galerkin with optimal test functions
  - Least-squares Galerkin method
  - Mixed Galerkin method
- A priori error analysis
  - Ideal DPG method
  - Practical DPG method
- A posteriori error analysis
- Fast solvers
- Examples
  - Example 1 (Standard FEM)
  - Example 2 ($L^2$-based least-squares)
  - Example 3 (An ODE)
  - Example 4 (Diffusion)
  - Example 5 (Stokes)