

**Towards optimization-based
error bounds for
uncertain PDEs**

**Arnold Neumaier
University of Vienna
Vienna, Austria**

Overview

1. The problem posed

– motivation and context

2. The error in linear sensitivity analysis

– beyond the dual weighted residual (DWR) method

3. An optimization approach to uncertainty

– not PDE specific

4. Linear elliptic equations

– functional analytic norm bounds

Summary

This is a report on work in progress.

Currently we have complete theoretical results for systems of uniformly elliptic PDEs, and partial results for uniformly parabolic and symmetric hyperbolic problems.

The implementation of the elliptic case is part of the PhD work of a student who just began.

1. The problem posed

Most partial differential equations solved in practice represent only an approximate model of the corresponding real life situation.

Modeling errors can be accounted for by

- adding parameters (constants or functions) to the model, and
- specifying the uncertainty in these parameters.

Each parameter θ selects a particular **scenario** from the set of possibilities.

One solves the equations for a particular scenario, or for just a few scenarios.

But one is interested in how the solution varies over the full (infinite) set of allowed scenarios.

Usually the quality is assessed by means of an overall merit function or a few key quantities computed from the solution.

We call these quantities **response functionals**, and assume that they are smooth.

Traditionally, the response functionals are evaluated by a **local sensitivity analysis**:

- Taylor expand around the nominal solution
- ignore higher order terms
- neglect discretization errors

However, these approximations may affect the errors significantly, especially for problems with low smoothness or highly nonlinear dependence on the parameters.

(This is frequent even for linear equations.)

Recent results by MUHANNA & MULLEN and NEUMAIER & POWNUK for large-scale truss structures (perhaps the simplest finite element problems) show that there are now methods to get

- **rigorous and realistic worst case error bounds**
- for uncertain linear systems of equations
- with thousands of variables and
- realistic parameter uncertainties (up to 15% rel. error).

The true error is overestimated by a computable margin, which typically is at most a few percent.

These methods are much faster than Monte Carlo methods (which always underestimate the worst case).

To extend this to partial differential equations,
we must cope with

- uncertainty due to parameters in the equations
 - is tractable as in finite dimensions
- errors introduced by discretization and numerical integration
 - needs additional tools from functional analysis

General procedure, adapted from experience with finite-dimensional problems:

- 1. Look at existence proofs for solutions.
 - These use some functional analysis existence argument.
- 2. Replace it by a constructive semilocal version.
 - This gives a verifiable existence condition for a solution close to a given approximation.
- 3. Find explicit values for the constant involved.
 - This boils down to finding computable bounds for inverse operators.
 - Requires special techniques for each problem type.

- 4. Combine with sensitivity analysis to ensure the **quadratic approximation property**:
 - For uncertainties in the problem formulation of order $O(\varepsilon)$, the overestimation of the bounds is of the order $O(\varepsilon^2)$.
- 5. If necessary, take additional measures to reduce the **wrapping effect**,
 - the exponential loss of accuracy when iterating estimates.

2. The error in linear sensitivity analysis

We consider the solution of a functional equation

$$F(u) = 0.$$

Given an approximate solution u_0 of $F(u) = 0$, standard sensitivity analysis can be used to compute an approximation

$$\gamma \approx g(u) - g(u_0). \quad (1)$$

We assume that bounds derivable from functional analysis are available which bound the error in this approximation rigorously,

$$|g(u) - g(u_0) - \gamma| \leq \delta. \quad (2)$$

Note that if $\delta < |\gamma|$ then the quotient

$$\bar{q} := (|\gamma| + \delta)/(|\gamma| - \delta) \quad (3)$$

is a computable bound for the **efficiency index**

$$q := (|\gamma| + \delta)/|g(u) - g(u_0)|,$$

the overestimation factor of the bound $|\gamma| + \delta$ for the error $|g(u) - g(u_0)|$ of the response functional evaluated at u_0 .

Therefore, knowing a bound $\delta \ll |\gamma|$ allows a check on the accuracy (rather than only the size) of this bound.

To find out what is needed to obtain a suitable δ , we use the **slope representation**

$$F(u) = F(u_0) + F[u, u_0](u - u_0), \quad g(u) = g(u_0) + g[u, u_0](u - u_0),$$

with a linear operator $F[u, u_0]$ and a linear functional $g[u, u_0]$.

F and g can be written down explicitly; in some cases (such as when F and g is linear) by an algebraic formula, in other cases by means of an integral derived from the mean value theorem. In any case, in a suitable norm,

$$F[u, u_0] = \partial_u F(u_0) + O(\|u - u_0\|), \quad g[u, u_0] = \partial_u g(u_0) + O(\|u - u_0\|).$$

Since $F(u) = 0$, we find that $u - u_0 = -F[u, u_0]^{-1}F(u_0)$ if the inverse exists, so that

$$g(u) - g(u_0) = g[u, u_0](u - u_0) = -g[u, u_0]F[u, u_0]^{-1}F(u_0).$$

Introducing a linear functional w_0 to be chosen later, we can rewrite the **sensitivity analysis error**

$$\Delta := g(u) - g(u_0) - w_0 F(u_0) = -g[u, u_0] F[u, u_0]^{-1} F(u_0) - w_0 F(u_0)$$

in the form

$$\Delta = -(g[u, u_0] + w_0 F[u, u_0]) F[u, u_0]^{-1} F(u_0). \quad (4)$$

The factor in parentheses deviates from $\partial_u g(u_0) + w_0 \partial_u F(u_0)$ by $O(\|u - u_0\|)$; hence it is natural to choose w_0 such that $\partial_u g(u_0) + w \partial_u F(u_0) \approx 0$ by approximately solving the so-called **adjoint equation**

$$\partial_u g(u_0) + w \partial_u F(u_0) = 0.$$

Solving the adjoint equation is the heart of the **dual weighted residual (DWR)** method of BECKER & RANNACHER for approximating the variation of the response functional, used to guide the adaptive refinement process of finite element methods for solving PDEs.

The above exact expression (4) for the sensitivity analysis error Δ is the key for getting reliable error bounds (rather than the uncontrolled approximations of the DWR) having the quadratic approximation property.

One only needs to find ways to bound Δ in a way which preserves its obvious second order accuracy.

For the case of systems of linear uniformly elliptic equations and linear g , we derive computable bounds of this type in Section 4 below.

For nonlinear elliptic equations there is the additional difficulty that the operator to be inverted in the exact error expression still contains the unknown u .

Thus additional analysis is needed to establish computable error bounds, using appropriate fixed point theorem to establish the existence of a solution near u_0 .

Therefore, bounds can be obtained only if u_0 is a sufficiently good approximation to a solution.

For uniformly parabolic equations and
for symmetric hyperbolic equations,
a similar approach is possible
but has only been partially worked out.

3. An optimization approach to uncertainty

We collect all uncertain parameters into a (possibly infinite-dimensional) vector θ .

Each admissible choice of θ defines a possible **scenario**; the set Θ of all admissible scenario defines the region of uncertainty.

We shall employ global optimization techniques to obtain bounds that cover both

- the errors in solving the differential equations, and
- the errors caused by the uncertainty in the parameters.

Assumptions

The problem to be solved when uncertainty is modeled explicitly takes the form

$$F(u, \theta) = 0 \quad \text{for some } \theta \in \Theta, \quad (5)$$

parameterized by θ .

We assume that that for fixed θ , the solution is uniquely determined and depends smoothly on θ .

We write the response functional of interest as $g(u, \theta)$, and assume for simplicity that it is real-valued.

The discussion of Section 2 applies for any particular scenario θ , with $F(u) = F(u, \theta)$ and $g(u) = g(u, \theta)$.

Finding approximations

Since the uncertainty region contains infinitely many scenarios θ , we cannot compute independent approximations for each θ .

To reduce the work to a finite and tractable amount, we compute an initial approximation $u_0 = u_0(\theta)$ for $u = u(\theta)$ by interpolation from approximate solutions u_l of (1) for a small number of scenarios $\theta_l \in \Theta$.

In the simplest case, a single scenario $\theta_0 \in \Theta$ suffices, leading to a θ -independent approximation u_0 .

However, better results are obtained if sufficiently many scenarios are used to ensure that $u_0(\theta)$ captures the essential variation in θ , since the errors will then be smaller.

In principle, one can increase the number of scenarios used adaptively until the error bounds obtained are satisfying.

The sensitivity analysis approximation γ also depends on the scenario θ .

It is usually obtained with the help of an approximate solution of a θ -dependent dual problem.

However, since we do not expect to need a highly accurate dual solution to get reasonably accurate bounds, the dual problem must usually be solved only once, at a nominal scenario θ_0 in the center of the uncertainty set,

But when uncertainties are large, a θ -dependent dual approximation may be important.

Finally, the bound δ for the error in γ will also depend on θ . The resulting error bound

$$|g(u(\theta), \theta) - g(u_0(\theta), \theta) - \gamma(\theta)| \leq \delta(\theta). \quad (6)$$

can be rewritten in the form of an enclosure

$$g(u(\theta), \theta) - g(u_0(\theta), \theta) \in [\gamma(\theta) - \delta(\theta), \gamma(\theta) + \delta(\theta)]. \quad (7)$$

This leads to the desired rigorous error bounds

$$g(u(\theta), \theta) - g(u_0(\theta), \theta) \in [\underline{e}, \bar{e}] \quad \text{for all } \theta \in \Theta, \quad (8)$$

where

$$\begin{aligned} \underline{e} &:= \min\{\gamma(\theta) - \delta(\theta) \mid \theta \in \Theta\}, \\ \bar{e} &:= \max\{\gamma(\theta) + \delta(\theta) \mid \theta \in \Theta\}. \end{aligned}$$

The computation of

$$\underline{e} := \min\{\gamma(\theta) - \delta(\theta) \mid \theta \in \Theta\},$$

$$\bar{e} := \max\{\gamma(\theta) + \delta(\theta) \mid \theta \in \Theta\}.$$

is a *global* optimization problem, since we really need the worst case to have a valid bound.

For safety tests, usually only an inequality for $g^*u(\theta)$ must be verified, and it suffices to compute one of \underline{e}, \bar{e} .

Solving the global optimization problems

If θ is not too high-dimensional, these global optimization problems are usually computational tractable.

For example, with the formulas derived later for elliptic PDEs with a polynomial dependence of the problem data on θ , the resulting optimization problems are purely algebraic if the number of uncertain variable is finite.

Their coefficients can be determined analytically or by numerical integration, using the approximate solution of the primal and dual equation.

Thus the error bounding process is completely decoupled from the PDE solver, and the work for the bounds is independent of the work spent in computing the approximations.

The resulting version of the optimization problem can be specified in an algebraic modeling language such as AMPL or GAMS, and solved with a variety of modern solvers.

In particular, the prize winning global solver BARON (available online via the NEOS interface) handles successfully many global optimization problems of substantial size.

Thus even large uncertainties and strongly nonlinear dependence of the solution on the parameters can be handled reliably.

The global solver BARON solves global optimization problems on the basis of the **branch-and-bound** principle.

An initial box bounding the set of admissible scenarios is recursively subdivided.

On each subbox, analytic estimates are used to check

- whether the subbox contains no scenario better than the best scenario already found (in which case the box is discarded),
- whether one can tighten the bounds without losing any better scenario.

The analytic estimates are obtained automatically by means of

- **interval analysis**, used to control the nonlinearities, which must be bounded over a whole subbox, and
- **linear relaxations**, which transform the problem into a sequence of relaxed linear programs.

Curse of dimensionality

Because of the need to search vast regions of space to reliably find the global optimum, the number of variables a global solver can handle is limited for problems where the branch tree created grows exponentially fast.

However, if the uncertainties are small, the problem is likely to have only one locally optimal point, in which case a local minimization is sufficient to locate it.

We sacrifice the guarantee of knowing whether we indeed have the global minimum, but have the advantage that a much larger suite of solvers is available.

To increase the confidence in the bounds we may optimize from several different starting points, resulting in the same optimum if the assumption is valid.

Rounding errors

If the PDE data depend continuously differentiably on θ then the bounds derived below also depend continuously differentiably on the data.

Using **automatic differentiation** techniques, it is not difficult to transform a program calculating the desired bounds into a program that provides together with the bounds also the gradient with respect to the uncertain parameters.

In the COCONUT optimization environment, this can even be done with intervals, and with slopes in place of the gradient.

This makes it possible to get bounds which remain rigorous even in the presence of rounding errors.

Probabilistic uncertainty

Often the uncertainties in the problem description can be specified only approximately, and a worst case analysis is inappropriate since the worst case has an exceedingly small probability.

Rather, one wants to get some information on the variation of the solution over the likely scenarios.

To describe these scenarios, we shall use the concept of **clouds**. Clouds describe the rough shapes of typical samples of various size, without fixing the details of the distribution.

They allow the representation of incomplete stochastic information in a clearly understandable and computationally attractive way.

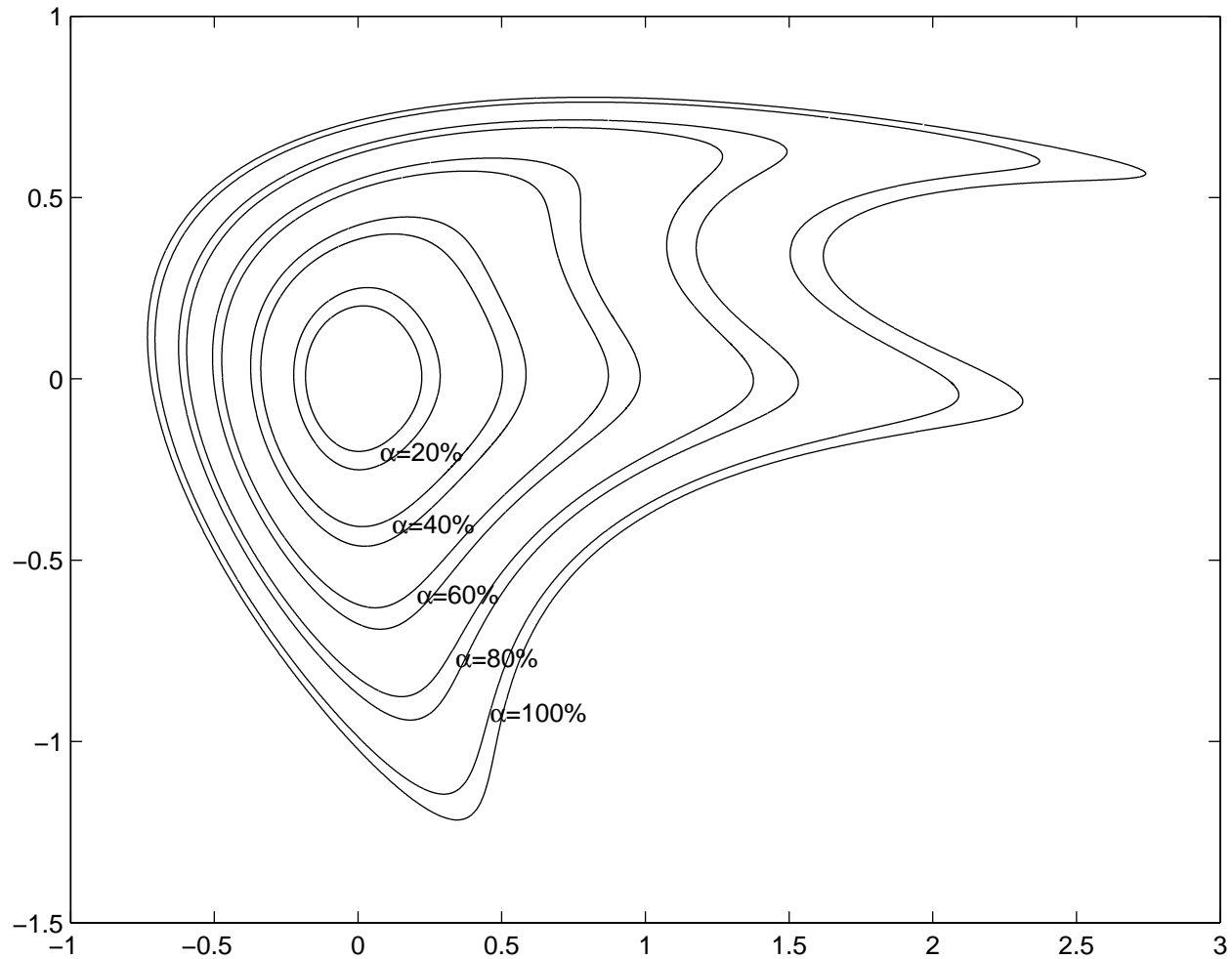


Figure 1: A 2-dimensional cloud with inner and outer boundaries for the confidence levels

Clouds have a well-defined operational semantics within probability theory.

Therefore computations with clouds give reliable probabilistic information.

Clouds use only probabilistic information readily available in practice.

This contrasts with traditional probabilistic reliability analysis, which requires the exact knowledge of the multivariate probability distribution of the uncertainty, a requirement nearly impossible to validate in practice.

The use of clouds permits a worst case analysis without losing track of important probabilistic information.

All computed probabilities are safeguarded against uncertainty in a way consistent with the limited information about probability distributions typically available in the multiparameter case.

The cloud approach has been applied successfully to the robust optimization of spacecraft design problems of the European Space Agency ESA.

The special case of interest for large-scale models is a **confocal cloud**.

It is defined by a continuous **potential** V which assigns to each scenario θ a potential function $V(\theta)$ defining the shape of the cloud, and a **lower probability** $\underline{\alpha}(s)$ and an **upper probability** $\bar{\alpha}(s)$ defining the fuzzy boundary of the cloud.

Writing $\Pr(A)$ for the probability of a statement A involving θ (which is now a random vector), the defining property of a cloud is that, for all s ,

$$\underline{\alpha}(s) \leq \Pr(V(\theta) < s) \leq \Pr(V(\theta) \leq s) \leq \bar{\alpha}(s), \quad (9)$$

where $\underline{\alpha}$ and $\bar{\alpha}$ are strictly increasing continuous functions of s mapping the range of V to $[0, 1]$.

For a given failure probability ε and $\alpha = 1 - \varepsilon$, the so-called **α -cut** describes

– an inner region \underline{C}_α of **α -relevant** scenarios with $V(\theta) < \underline{s}_\varepsilon$, and

– a (generally larger) region \overline{C}_α of **α -reasonable** scenarios with $V(\theta) < \overline{s}_\varepsilon$, where

$$\overline{\alpha}(\underline{s}_\varepsilon) = 1 - \varepsilon, \quad \underline{\alpha}(\overline{s}_\varepsilon) = 1 - \varepsilon$$

define $\underline{s}_\varepsilon$ and \overline{s}_ε with $\underline{s}_\varepsilon \leq \overline{s}_\varepsilon$.

For continuous random variables, the conditions defining the cloud guarantee that for every α there is a region C_α with $\underline{C}_\alpha \subseteq C \subseteq \overline{C}_\alpha$ whose measure is a fraction α of the measure of the set of all scenarios considered possible.

To apply the techniques of the previous section, we define regions of uncertainty labelled by the potential value s representing the boundary of a level set in the cloud,

$$\Theta_s := \{\theta \mid V(\theta) \leq s\}.$$

Then the uncertainty is specified by a family of deterministic conditions $\theta \in \Theta_s$.

The probabilistic information is restricted to the interpretation of s and is given in the form

$$\Pr(\theta \in \Theta(s)) \in [\underline{\alpha}(s), \bar{\alpha}(s)].$$

The potential determines the shape of the cloud.

Of primary interest for practical work in higher dimensions are **elliptic clouds**, a sort of fuzzy ellipsoids defined by potentials of the form

$$V(\theta) = \|A\theta - b\|^2,$$

and **polyhedral clouds** defined by potentials of the form

$$V(\theta) = \max_k |(A\theta - b)_k|.$$

The shape of these clouds can be easily adapted to uncertain probabilistic information.

The values of the coefficients are chosen to match the overall shape of the data set, based on readily available partial information about 1- and 2-dimensional marginal distributions (FUCHS & NEUMAIER).

The distribution of the resulting 1-dimensional random number $V(\theta)$ can be estimated reliably, even for high-dimensional θ .

Thus all information needed to define a cloud can be determined from a reasonable amount of data.

4. Linear elliptic equations

To complete the general scheme outlined above, one needs explicit rules for computing with sufficiently little overestimation the required bounds on the sensitivity analysis error Δ .

We have completed the theoretical analysis of this step for the case of linear, uniformly elliptic partial differential equations and linear response functionals.

We write the linear PDE (including the boundary conditions) in the operator form

$$Au = f, \quad (10)$$

and write the linear response functional as $g(u) = g^*u$.

Assuming the availability of approximations u_0 to the primal equation (10) and v_0 to the dual equation

$$A^*v = g, \quad (11)$$

we define the **residuals**

$$r := f - Au_0, \quad s := g - A^*v.$$

In terms of our results from Section 2, we may take

$$F(u) = Au - f.$$

The slopes take the simple form

$$F[u, u_0] = A, \quad g[u, u_0] = g^*,$$

and with $w_0 := -v_0^*$, the sensitivity analysis error Δ from (4) becomes

$$\Delta = -(g^* - v_0^* A) A^{-1} (A u_0 - f) = s^* A^{-1} r,$$

independent of the unknown u .

Note that the uniform ellipticity of A implies that A^{-1} exists.

Thus the error bounding problem is reduced to the problem of bounding the bilinear form $s^* A^{-1} r$ in terms of suitable norms in which r and s are small.

Note that the computed approximations usually have low differentiability properties, so that the computed residuals are singular.

Thus one must take care in defining and computing these norms.

A Hilbert space setting

The natural abstract setting for 'elliptic' error bounds is given by two real Hilbert spaces \mathbb{U} and \mathbb{H} of square integrable functions with respect to suitable measures.

We write the inner product of $u, v \in \mathbb{U}$ as u^*v , and that of $h, k \in \mathbb{H}$ as h^*k .

We assume that for some dense subspace U (of smooth functions) of \mathbb{U} , we have linear operators

$$L, N : U \rightarrow \mathbb{H};$$

L is assumed to be injective. Typically, L is a differential operator, while N contains no derivatives.

We assume that the linear operator A figuring in the equation of interest is expressible as

$$A = L^* M L + N^* L - L^* N \quad (12)$$

with a bounded, symmetric positive definite operator

$$M : \mathbb{H} \rightarrow \mathbb{H}$$

with bounded inverse, expressing ellipticity.

$N^* L - L^* N$ is the antisymmetric part of A , accounting for convection.

$$S := L^*ML$$

is the symmetric, positive definite part of A , accounting for diffusion. Under our assumptions,

$$\|u\|_\theta := \sqrt{u^*Su} = \sqrt{(Lu)^*M(Lu)} \quad (13)$$

defines a norm on U , called the **energy norm**.

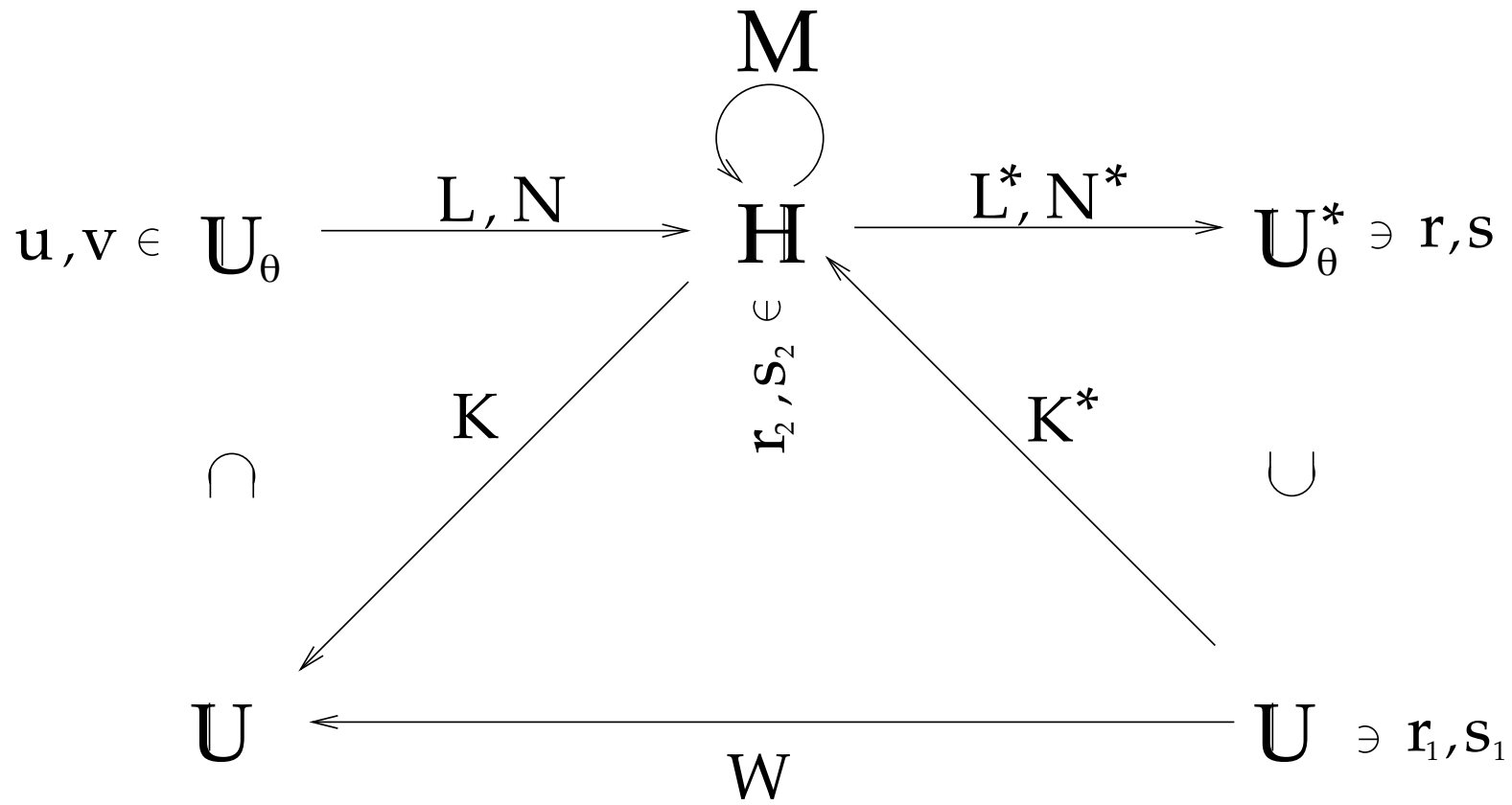
We define the Hilbert space \mathbb{U}_θ as the closure of U with respect to the energy norm.

\mathbb{U}_θ is embedded as vector space into \mathbb{U} by inclusion, but has the topology induced by the Sobolev type energy norm.

Note that in general, L, M, N , and hence the energy norm and \mathbb{U}_θ may depend on the scenario θ .

We write \mathbb{U}_θ^* for the dual of \mathbb{U}_θ with respect to the inner product of \mathbb{U} , and f^*u for the image of $u \in \mathbb{U}_\theta$ under $f \in \mathbb{U}_\theta^*$.

L extends by continuity to a linear mapping from \mathbb{U}_θ to \mathbb{H} , and the dual L^* is a linear mapping from \mathbb{H} to \mathbb{U}_θ^* .



The representation theorem of Riesz implies the invertibility of $S = L^*ML$, leading to the formula

$$\|r\|_{\theta}^* = \|S^{-1}r\|_{\theta} = \sqrt{r^*S^{-1}r} \quad (14)$$

for the dual norm of $r \in \mathbb{U}_{\theta}^*$, and the formula

$$\|L^*h\|_{\theta}^* \leq \|h\|_M := \sqrt{h^*M^{-1}h} \quad \text{for } h \in \mathbb{H}. \quad (15)$$

Both the operator $A = L^*ML + N^*L - L^*N$ and its dual $A^* = L^*ML - N^*L + L^*N$ map \mathbb{U}_θ into \mathbb{U}_θ^* . Thus the equations $Au = f$ and $A^*v = g$ make sense for all $f, g \in \mathbb{U}_\theta^*$.

The Lax-Milgram theorem implies an existence proof for the solutions, and the desired bounds for expressions of the form $s^*A^{-1}r$:

Theorem

The operator $A = L^*ML + N^*L - L^*N$ is invertible, and

$$|s^*A^{-1}r| \leq \|s\|_\theta^* \|r\|_\theta^* \quad \text{for } r, s \in \mathbb{U}_\theta^*. \quad (16)$$

This reduces our estimation problem to that of finding reasonable bounds for the dual norms of residuals.

But general right hand sides are singular functions.

Even when the right hand side is smooth, approximations u_0 or v_0 of low smoothness lead to singular residuals.

To avoid numerical computations with such singular functions, we write these as the sum of a square integrable term and the derivative of another square integrable term.

More specifically, we consider the primal equation

$$Au = f = f_1 + L^* f_2, \quad f_1 \in \mathbb{U}, \quad f_2 \in \mathbb{H},$$

and the associated dual equation

$$A^* v = g = g_1 + L^* g_2, \quad g_1 \in \mathbb{U}, \quad g_2 \in \mathbb{H}.$$

Clearly, f_2, g_2 are determined by f, g only up to the addition of a function from the space

$$\mathbb{S} := \{h \in \mathbb{H} \mid L^* h \in \mathbb{U}\} \tag{17}$$

of smooth functions in \mathbb{H} .

Given an approximate primal solution $u_0 \in \mathbb{U}_\theta$, the primal residual $r = f_1 + L^*f_2 - Au_0$ can be written with an arbitrary $b \in \mathbb{S}$ in the form

$$r = (f_1 - N^*Lu_0 + L^*h) + L^*(f_2 - MLu_0 + Nu_0 - h),$$

with square integrable terms in both brackets. Thus

$$\|r\|_\theta^* \leq \|f_1 - N^*Lu_0 + L^*h\|_\theta^* + \|f_2 - MLu_0 + Nu_0 - h\|_M.$$

$h \in \mathbb{S}$ is a smoothness requirement amounting to $L^*h \in \mathbb{U}$.

A good choice for b is therefore obtained by smoothing the expression $f_2 - MLu_0 + Nu_0 \approx h \in \mathbb{S}$.

Similarly, given an approximate dual solution $v_0 \in \mathbb{U}_\theta$, the dual residual $s = g_1 + L^*g_2 - A^*v_0$ can be written with an arbitrary $c \in \mathbb{S}$ in the form

$$s = (g_1 + N^*Lv_0 + L^*k) + L^*(g_2 - MLv_0 - Nv_0 - k),$$

with square integrable terms in both brackets. Thus

$$\|s\|_\theta^* \leq \|g_1 + N^*Lv_0 + L^*k\|_\theta^* + \|g_2 - MLv_0 - Nv_0 - k\|_M.$$

Now $k \in \mathbb{S}$ should be chosen by smoothing the expression $g_2 - MLv_0 - Nv_0 \approx k \in \mathbb{S}$.

It remains to discuss how the dual norm

$$\|w\|_{\theta}^* = \sqrt{F^* S^{-1} w}$$

of a square integrable function $w \in \mathbb{U}$ can be obtained without using computations involving S^{-1} .

Theorem

If, for some $\beta > 0$, some linear operator $K : \mathbb{H} \rightarrow \mathbb{U}$, and some positive definite symmetric operator $W : \mathbb{U} \rightarrow \mathbb{U}$,

$$E := LK + K^* L^* - K^* W^{-1} K - \beta^2 M^{-1}$$

is positive semidefinite on $\mathbb{H}_0 = \{h \in \mathbb{H} \mid Kh \in \mathbb{U}_\theta\}$ then

$$\beta \|w\|_{\theta}^* \leq \|w\|_W := \sqrt{w^* W w} \quad \text{for } w \in \mathbb{U}. \quad (18)$$

Proof. S^{-1} maps \mathbb{U} to \mathbb{U} since $\mathbb{U}_\theta \subseteq \mathbb{U} \subseteq \mathbb{U}_\theta^*$.

Thus $Z := MLS^{-1}$ maps \mathbb{U} into \mathbb{H} , and

$$W - \beta^2 Z^* M^{-1} Z = (W - KZ)^* W^{-1} (W - KZ) + Z^* E Z.$$

Now $Z^* L = 1$ implies $Z^* M^{-1} Z = Z^* L S^{-1} = S^{-1}$.

Thus $w^* W w - \beta^2 w^* S^{-1} w = w^* (W - \beta^2 Z^* M^{-1} Z) w \geq 0$
for all $w \in \mathbb{U}$ with $Zw \in \mathbb{H}_0$.

By continuity, this extends to all $w \in \mathbb{U}$,
and proves the assertion. □

In the applications to systems of elliptic PDEs, W, M, K are multiplication operators over a domain Ω , and L is a first order differential operator.

In this case, E is also a multiplication operator; the derivatives cancel by integration by parts.

Thus checking positive semidefiniteness reduces to a purely algebraic problem tractable by means of **semidefinite programming, interval analysis, or global optimization.**

For example, in case of the Poisson equation in a d -dimensional domain Ω with homogeneous Dirichlet boundary conditions,

$$\mathbb{U} = H_0^2(\Omega), \quad \mathbb{H} = L^2(\Omega)^d, \quad L = \nabla, \quad W = M = 1.$$

Choose K as multiplication by $se^T x$, where e is the all-one vector, and s an adjustable constant.

The semidefiniteness of E is equivalent to the algebraic condition $s - s^2(e^T x)^2 - \beta^2 \geq 0$ for all $x \in \Omega$. Writing $\omega := \max_{x \in \Omega} |e^T x|$, this is equivalent to $\beta^2 \leq s - s^2\omega^2$. Choosing $s = 1/2\omega^2$, we see that we can take $\beta = 1/2\omega$.

A more versatile ansatz for K , and solving a more complex optimization problem for the free parameters in the ansatz, gives a constant β close to best possible.

In general, the complete bound calculations are reduced to computing

- an approximate (fine grid) solution of the primal problem,
- the low resolution (coarse grid) solution of a dual problem,
- four residuals (on the fine grid),
- two applications of smoothing fine grid data,
- an algebraic check for semidefiniteness, and
- the integrations necessary to find the norms.

Typically, therefore, the work should not be much higher than that for the computation of the approximation alone.

In contrast to standard treatments which involve constants that are difficult to obtain for a general domain, all our assumptions can be verified rigorously by computer.

Of course, the integration errors and rounding errors made in evaluating the bounds themselves remain unaccounted for.

For a fully rigorous implementation in the mathematical sense, one would also need to bound these errors.

Because of the a posteriori nature of our bounds, an error control is needed *only* for the evaluation of the residual bounds.

The approximations u_0, v_0, b and c may be arbitrary and hence can be computed without error control.

Thus the rounding error control can be restricted to fairly simple arithmetic operations and to numerical integration.

Summary

By combining functional analytic tools with global optimization techniques, worst case bounds can be found (at least for uniformly elliptic PDEs) that cover

- the errors made in solving the differential equations:
 - discretization errors
 - numerical integration errors
 - boundary approximation errors
- the errors caused by parameter uncertainties.

This provides reliable tools for the assessment of uncertainty in the solution of (at least elliptic) partial differential equations.

The bounds are independent of the way the approximations are obtained.

This allows an independent verification of the quality of approximate solvers.

A traditional probabilistic reliability analysis does not account for discretization errors.

Moreover it would require the exact knowledge of the multivariate probability distribution of the uncertainty, a requirement nearly impossible to validate in case of many uncertain parameters.

Using clouds, the above optimization techniques can also be applied to probabilistic uncertainty. Clouds use only probabilistic information readily available in practice from 1- and 2-dimensional marginal distributions.

Web site with background material

Uncertainty modeling in high dimensions

<http://www.mat.univie.ac.at/~neum/clouds.html>

Details for linear elliptic PDEs

Certified error bounds for uncertain elliptic equations,
J. Comput. Appl. Math. 218 (2008), 125–136.

<http://www.mat.univie.ac.at/~neum/papers.html#pdebounds>