Adjoint solution method for inverse and optimization problems.

Marc Bonnet

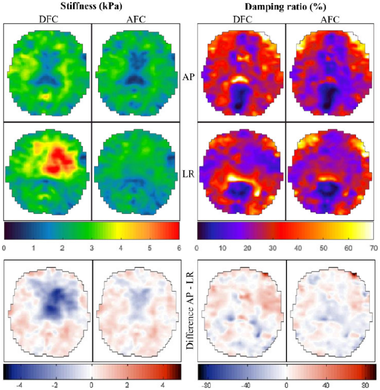
Propagation des Ondes: Etudes Mathématiques et Simulation (POEMS)
UMR 7231 CNRS-INRIA-ENSTA
Unité de Mathématiques Appliquées
ENSTA Paris

mbonnet@ensta.fr

Contents

Optimization and inversion based on physical ODE/PDE models

- Optimization (inverse, optimal control...) problems based on PDE/ODE models arise in many areas of science and engineering
- Often, find some parameter, control variable...given data or target on state variable(s)
- Widespread use of PDE-constrained optimization. Possibly most well-known use: full-waveform inversion (FWI) in geophysics
- Includes optimization problem yielding MAP (Bayesian) estimates
- Includes treatment of constitutive identification using error in constitutive relation (my last two talks)



S. Kurtz, PhD 2023

Bayesian solution approach (Angèle's talks), short reminder

- Definition of a conditional probability $P(A|B)$:

  $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

  $P(A \cap B) = P(B \cap A)$ (symmetry) yields Bayes' formula:

  $$P(A|B)P(B) = P(B|A)P(A)$$

- **Idea:** use Bayes to "invert" the parameter-data relationship between $\boldsymbol{p}$ and $\boldsymbol{d}$:

  $$f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}}) \propto f_{\mathcal{D}|\mathcal{P}}(\boldsymbol{d}_{\text{obs}}|\boldsymbol{p}) f_{\mathcal{P}}(\boldsymbol{p})$$

  $f_{\mathcal{P}}(\boldsymbol{p})$: probability density describing prior information on $\boldsymbol{p}$

  $f_{\mathcal{D}|\mathcal{P}}(\boldsymbol{d}|\boldsymbol{p})$: probability density describing the effect of the forward problem (probability density describing modeling and measurement uncertainties)

  $f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}})$: probability density on $\boldsymbol{p}$ given $\boldsymbol{d}$ (defines a solution of the inverse problem)

- Estimators on $\boldsymbol{p}$ extracted by post-processing $f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}})$. In particular:

  $$\boldsymbol{p}_{\text{MAP}} = \arg\max_{p} f_{\mathcal{P}|\mathcal{D}}(\boldsymbol{p}|\boldsymbol{d}_{\text{obs}}) \qquad \textit{Maximum a posteriori estimate}$$

  Main focus of e.g. Tarantola's first book on inverse problems (1987)

PDE-constrained optimization, general considerations

- Forward solution u and parameter p linked by ODE/PDE model $E(p,u)=0$:

  seek p solving $\quad \min_{p,u} J(p,u)$ s.t. $E(p,u)=0$ (and possibly $u \in \mathcal{U}_{\text{adm}},\ p \in \mathcal{P}_{\text{adm}} \ldots$)

- Assume (usually) $E(p,u)=0$ solvable for u given p (can be subject to admissibility constraints on p):
  u implicit function of p through $E(p,u)=0$: (often-nonlinear) solution mapping $p \mapsto u(p)$

  Reduced objective and minimization: $\quad \min_{p \in \mathcal{P}_{\text{adm}}} \widehat{J}(p) := J\big(p, u(p)\big)$

- Iterative optimization algorithms, compute minimizing sequence $p_n \to p$

  Need (at least) evaluation of each $\widehat{J}(p_n)$; requires $u_n = u(p_n)$ via PDE solve.

- Optimization algorithms using evaluations of $\boldsymbol{\nabla}\widehat{J}$ (sometimes $\boldsymbol{\nabla}\boldsymbol{\nabla}\widehat{J}$) avoid excessive number of (costly) PDE solves. Descent directions usually defined using full gradient $\boldsymbol{\nabla}\widehat{J}$.
- (KKT) optimality conditions (basis of some algorithms) expressed with derivatives of J, E.
- Numerical derivatives of $\widehat{J}$ for P-dim. parameter approximation $p \in \text{span}(\pi_1, \ldots, \pi_P)$, e.g.:

$$\boldsymbol{\nabla}\widehat{J}(p) \approx \Big(\frac{\widehat{J}(p+h\pi_1) - \widehat{J}(p)}{h}, \ldots, \frac{\widehat{J}(p+h\pi_P) - \widehat{J}(p)}{h} \Big) \qquad h\text{: small step}$$

  Total evaluation cost for $\widehat{J}(p)$, $\boldsymbol{\nabla}\widehat{J}(p)$: $P+1$ PDE solves (at least once per optim. iteration).

- Can be greatly improved: analytical sensitivity or (better still) adjoint solution methods

A few words about derivatives

Derivative

Let $f:\mathcal{X}\to\mathcal{Y}$ ($\mathcal{X},\mathcal{Y}$ normed vector spaces). If $\boxed{f(a+h)=f(a)+f'(a)h+o(\|h\|_{\mathcal{X}})}$ with linear continuous operator $f'(a)\in\mathcal{L}(\mathcal{X},\mathcal{Y})$, we say that $f'(a)$ is the (Fréchet) derivative of f at a.

- Finite-dim. case: $f'(a)$ Jacobian matrix of f at a;
- $\mathcal{Y}=\mathbb{R}$: $f'(a)$ is a continuous linear functional, $f'(a)h=\langle f'(a),h\rangle$ ($\langle\cdot,\cdot\rangle$ duality bracket);
- $\mathcal{Y}=\mathbb{R}$ and $\mathcal{X}$ Hilbert: there is $g(a)\in\mathcal{X}$ (gradient of f at a) such that $f'(a)h=\big(g(a),h\big)_{\mathcal{X}}$
- If f multivariate, e.g. $(x,y)\mapsto f(x,y)$, we write

$$f(a+h,b+k)=f(a,b)+\partial_x f(a)h+\partial_y f(b)k+o(\|(h,k)\|)$$

First-order KKT conditions (equality constraint)

Let (p,u) verify $E(p,u)=0$ and $\partial_u E(p,u)$ linear continuous with bounded inverse. Assume $\widehat{J}$ has a local extremum at p. Then, there exists $\lambda=\lambda(p,u)$ such that

$$\partial_\lambda\mathcal{L}(p,u,\lambda)=0,\quad \partial_u\mathcal{L}(p,u,\lambda)=0,\quad \partial_p\mathcal{L}(p,u,\lambda)=0,$$

where $\mathcal{L}(p,u,\lambda):=J(p,u)+\langle E(p,u),\lambda\rangle$ is the Lagrangian associated with the constrained optimization problem.

Reduced objective derivative using state sensitivities

$$\min_{p\in\mathcal{P},u\in\mathcal{U}} J(p,u) \quad \text{subject to} \quad E(p,u)=0$$

- Let p,u verify $E(p,u)=0$ (i.e. $u=u(p)$). Assume $\partial_u E(p,u(p))\in\mathcal{L}(\mathcal{U},\mathcal{U}')$ invertible with bounded inverse. By implicit function thm., $u(\cdot)$ differentiable at p and

$$\partial_u E(p,u)\, u'(p) + \partial_p E(p,u) = 0$$

To compute $u'(p)q$ for given $q\in\mathcal{P}$: solve $\partial_u E(p,u)[u'(p)q] = -\partial_p E(p,u)q$.

- Then, differentiation of reduced objective $\widehat{J} = J(u(\cdot),\cdot)$ at p yields

$$\widehat{J}'(p)q = \partial_u J(p,u)\, u'(p)q + \partial_p J(p,u)q$$

- For P-dim. parameter approximation $p\in\text{span}(\pi_1,\ldots,\pi_P)$:

$$\nabla\widehat{J}(p) = \big(\widehat{J}'(p)\pi_1,\ldots,\widehat{J}'(p)\pi_P\big)^{\mathsf{T}}$$

A priori entails solving P sensitivity problems:

$$\widehat{J}'(p)\pi_k = \partial_u J(p,u)\, u'(p)\pi_k + \partial_p J(p,u)\pi_k,$$

$$\text{with} \quad \partial_u E(p,u)u'(p)\pi_k = -\partial_p E(p,u)\pi_k \quad (1\le k\le P)$$

- Descent directions require the full gradient, e.g. $\boldsymbol{d} = -\mathbf{H}\cdot\nabla\widehat{J}(p)$ using quasi-Newton
- **Adjoint solution method:** provides $\nabla\widehat{J}(p)$ without actually solving sensitivity problems

Adjoint solution method: basic mechanism

Adjoint solution method: provides $\nabla\widehat{J}(p)$ without actually solving sensitivity problems

$$\widehat{J}'(p)\pi_k = \partial_u J(p,u)\, u'(p)\pi_k + \partial_p J(p,u)\pi_k,$$
$$\text{with} \quad \partial_u E(p,u)u'(p)\pi_k = -\partial_p E(p,u)\pi_k \quad (1 \leq k \leq P)$$

i.e. $$\widehat{J}'(p)\pi_k = g\, v_k + \partial_p J(p,u)\pi_k, \qquad \text{with} \quad Av_k = f_k \quad (1 \leq k \leq P)$$

- Basic task: to evaluate one linear functional $\langle g, v_k \rangle$ on (possibly infinitely) many v_k solving $Av_k = f_k$.
- Associate adjoint solution λ to g by $A^\star\lambda = g$, use transposition ($\langle x, Ay \rangle = \langle A^\star x, y \rangle$):

$$\langle g, v_k \rangle = \langle A^\star\lambda, v_k \rangle = \langle \lambda, Av_k \rangle = \langle \lambda, f_k \rangle \qquad \boxed{\langle g, v_k \rangle = \langle \lambda, f_k \rangle \quad \text{for all } k}$$

Reduced objective derivative using adjoint solution

- In present context (with $q = \pi_1, \dots, \pi_P$ for the discrete gradient):
  $g = -\partial_u J(p,u)$, $v_k = u'(p)q$, $f_k = -\partial_p E(p,u)q$, $q \in \mathcal{P}$ arbitrary.
- Revisit reduced objective derivative evaluation using adjoint solution:

$$\widehat{J}'(p)q = -\big\langle \partial_u J(p,u)\,,\, \big[\partial_u E(p,u)\big]^{-1}\,\partial_p E(p,u)q\,\big\rangle_{\mathcal{U}',\mathcal{U}} + \partial_p J(p,u)q$$
$$= \big\langle \underbrace{-\big[\partial_u E(p,u)\big]^{-\star}\,\partial_u J(p,u)}_{\lambda}\,,\, \partial_p E(p,u)q\,\big\rangle_{\mathcal{U}',\mathcal{U}} + \partial_p J(p,u)q$$

The **adjoint solution** $\lambda \in \mathcal{U}$ solves $\big[\partial_u E(p,u)\big]^{\star}\lambda = -\partial_u J(p,u)$. Then, for any $q \in \mathcal{P}$:
$\widehat{J}'(p)q = \big\langle \lambda, \, \partial_p E(p,u)q\,\big\rangle_{\mathcal{U}',\mathcal{U}} + \partial_p J(p,u)q$ i.e.

- Same result on $\widehat{J}'(p)$ found using Lagrangian $\mathcal{L}(p,u,\lambda) := J(p,u) + \big\langle E(p,u), \lambda \big\rangle$:

  $$\widehat{J}'(p)q = \partial_p \mathcal{L}(p,u(p),\lambda(p))q$$

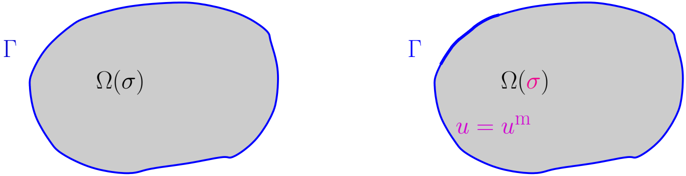
  with forward and adjoint problems coinciding with first two KKT conditions

  $$\partial_\lambda \mathcal{L}(p,u,\lambda) = 0, \quad \partial_u \mathcal{L}(p,u,\lambda) = 0.$$
- Third(remaining) KKT condition in the form $\partial_p \mathcal{L}(p,u(p),\lambda(p)) = 0$: 1st order stationarity condition for unconstrained minimization of $\widehat{J}(p)$.

Example: inverse conductivity (EIT) problem

- Forward problem: find electrostatic potential u in Ω given source f and conductivity σ

$$-\text{div}\,(\sigma\boldsymbol{\nabla} u) = f \quad \text{in } \Omega, \qquad \partial_n u = 0 \quad \text{on } \partial\Omega \qquad \big(E(\sigma,u)=0\big)$$



Variational formulation, $\mathcal{U} = \big\{ w \in H^1(\Omega), \langle w \rangle_\Omega = 0 \big\}$:

$$\text{Find } u\in\mathcal{U}, \quad A(\sigma,u,w) - F(w) = 0 \quad \text{for all } w\in\mathcal{U}$$

$$\text{i.e. } E(\sigma,u)=0, \; E : \mathcal{P}\times\mathcal{U}\to\mathcal{U}'$$

$$\begin{cases} A(\sigma,u,w) := \displaystyle\int_\Omega \sigma\boldsymbol{\nabla} u\cdot\boldsymbol{\nabla} w \,\text{d}V \\ F(w) := \displaystyle\int_\Omega f\,w,\,\text{d}S \end{cases}$$

- Inverse problem: estimate σ from measurements u^{m} of u on Γ.

  Optimization approach: regularized output least-squares with PDE constraint linking u to σ

$$\min_{\sigma,u} J(\sigma,u) \qquad \text{subject to } E(\sigma,u)=0$$

$$\text{e.g.} \quad J(\sigma,u) = \frac{1}{2}\int_\Gamma |u-u^{\text{m}}|^2 \,\text{d}S + \alpha R(\sigma)$$

  Often additional inequality constraints (mostly left out in this talk), e.g. $\sigma\geq\sigma_0$, $\sigma\in\mathcal{K}\ldots$

Linear PDE constraint in weak form

$$\min_{p\in\mathcal{P},u\in\mathcal{U}} J(p,u) \qquad \text{subject to} \qquad A(p,u,w)-F(w)=0 \quad \text{for all } w\in\mathcal{U}$$

where $A(p,\cdot,\cdot)$ bilinear, continuous form, and thus $\partial_u A(p,u,w)z = A(p,z,w)$.

- Variational formulations for state sensitivity and adjoint problems:

$$A(p,u'(p)q,w) = -\partial_p A(p,u,w)q \qquad \forall w\in\mathcal{U} \qquad \text{(sensitivity)}$$
$$A(p,w,\lambda) = -\partial_u J(p,u)w \qquad \forall w\in\mathcal{U} \qquad \text{(adjoint)}$$

- Combine with $w=-\lambda$ and $w=u'(p)q$:

$$\left.\begin{aligned} -A(p,u'(p)q,\lambda) &= \partial_p A(p,u,\lambda)q \\ A(p,u'(p)q,\lambda) &= -\partial_u J(p,u)u'(p)q \end{aligned}\right\} \implies \partial_u J(p,u)u'(p)q = \partial_p A(p,u,\lambda)q$$

The **adjoint solution** $\lambda\in\mathcal{U}$ solves $A(p,w,\lambda) = -\partial_u J(p,u)w \quad \forall w\in\mathcal{U}$. Then, for any $q\in\mathcal{P}$:

$$\widehat{J}'(p)q = \partial_p A(p,u,\lambda)q + \partial_p J(p,u)q \quad \text{i.e.} \quad \boxed{\widehat{J}'(p) = \partial_p A(p,u,\lambda) + \partial_p J(p,u)}$$

Inverse conductivity pb.: $p=\sigma$, $A(\sigma,u,w)$ trilinear and $J'(u)w = \int_{\Gamma_N}(u-u^m)w\,dS$, hence

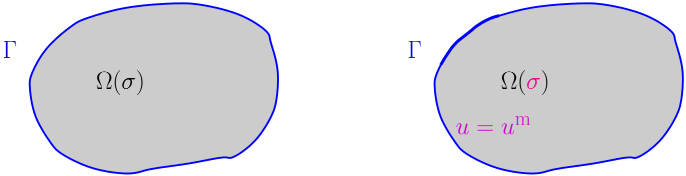
$$A(\sigma,u'(\sigma)s,w) = -A(s,u,w) \qquad \forall w\in\mathcal{U} \qquad \text{(sensitivity)}$$
$$A(\sigma,w,\lambda) = -\big((u-u^m),w\big)_{\Gamma_N} \qquad \forall w\in\mathcal{U} \qquad \text{(adjoint)}$$

$$\boxed{\widehat{J}'(\sigma)s = \int_\Omega s\boldsymbol{\nabla}u(\sigma)\cdot\boldsymbol{\nabla}\lambda(\sigma)\,dV + \alpha R'(\sigma)s}$$

Example: diffusivity identification for the heat equation

- Forward problem: find temperature u in Ω given source f and diffusivity σ

$$\partial_t u - \text{div}\,(\sigma \nabla u) = f \;\text{ in } \Omega, \quad \partial_n u = 0 \;\text{ on } \partial\Omega, \quad u(\cdot,0) = 0 \;\text{ in } \Omega \qquad \big(E(\sigma,u)=0\big)$$



Variational formulation, $\mathcal{U} = H^1([0,T]\times\Omega)\,\}$:

$$\text{Find } u \in \mathcal{U}, \quad \begin{cases} (\partial_t u, w)_\Omega + A(\sigma,u,w) - F(w) = 0 \quad t\in[0,T] \quad \text{for all } w \in H^1(\Omega) \\ u(\cdot,0) = 0 \;\text{ in } \Omega \end{cases}$$

- Inverse problem: estimate σ from measurements u^{m} of u on $\Gamma\times[0,T]$.

$$\boxed{\min_{\sigma,u} J(\sigma,u) \qquad \text{s.t. } E(\sigma,u)=0} \qquad \text{e.g. } J(\sigma,u) = \frac{1}{2}\int_0^T\!\!\int_\Gamma |u-u^{\mathrm{m}}|^2\,\mathrm{d}S\,\mathrm{d}t + \alpha R(\sigma).$$

Same general approach for full-waveform inversion (FWI, geophysics) and many other cases.

Adjoint solution, heat equation example

- Variational formulations for the (forward) sensitivity and (backward) adjoint problems:

Find $u' = u'(\sigma)s \in \mathcal{U}$, $\begin{cases} \big(\partial_t u', w\big)_\Omega + A(\sigma, u', w) = -A(s, u, w) & t \in [0, T] \quad \forall w \in H^1(\Omega) \\ u'(\cdot, 0) = 0 & \text{in } \Omega \end{cases}$

Find $\lambda \in \mathcal{U}$, $\begin{cases} -\big(w, \partial_t \lambda\big)_\Omega + A(\sigma, w, \lambda) = -\big(u - u^{\text{m}}, w\big)_\Gamma & t \in [0, T] \quad \forall w \in H^1(\Omega) \\ \lambda(\cdot, T) = 0 & \text{in } \Omega \end{cases}$

- Set $w = \lambda$, $w = -u'$, integrate over $[0, T]$, combine, use initial/final conditions:

$$\left.\begin{aligned} &\big(\partial_t u', \lambda\big)_\Omega + A(\sigma, u', \lambda) = -A(s, u, \lambda) \\ &\big(u', \partial_t \lambda\big)_\Omega - A(\sigma, u', \lambda) = \big(u - u^{\text{m}}, u'\big)_\Gamma \end{aligned}\right\}$$

$$\Longrightarrow \int_0^T \big(u - u^{\text{m}}, u'\big)_\Gamma \,\text{d}t - \int_0^T A(s, u, \lambda)\,\text{d}t = \int_0^T \partial_t \big(u', \lambda\big)_\Omega \,\text{d}t = 0$$

i.e. $$\partial_u J(\sigma, u)u' = \int_0^T A(s, u, \lambda)\,\text{d}t$$

$$\widehat{J}'(\sigma)s = \partial_u J(\sigma, u)u'(\sigma)s + \partial_\sigma J(\sigma, u) = \int_0^T A(s, u, \lambda)\,\text{d}t + \alpha R'(\sigma)s$$

Additional considerations

Find $\lambda \in \mathcal{U}$, $\begin{cases} -\big(w, \partial_t \lambda\big)_\Omega + A(\sigma, w, \lambda) = -\big(u - u^{\text{m}}, w\big)_\Gamma & t \in [0, T] \quad \forall w \in H^1(\Omega) \\ \lambda(\cdot, T) = 0 & \text{in } \Omega \end{cases}$

- Time-backward adjoint solution with final condition.

  Solving adjoint problem requires forward solution over whole duration $[0, T]$ (here),
  at final time (occasionally).

- In general, gradient evaluation needs forward and adjoint solutions over whole duration $[0, T]$.

- Can cause significant memory problems in large-scale applications. Mitigation includes
  - ▷ treating $[0, T]$ piecewise and recomputing parts of forward history.
  - ▷ "parareal" method [P.L. Lions, Y. Maday, G. Turini 01]

- Same result on $\widehat{J}'(y)$ by 1st order stationarity conditions $\partial_\lambda \mathcal{L} = 0$, $\partial_u \mathcal{L} = 0$ of Lagrangian

$$\mathcal{L}(\sigma, u, \lambda) := J(\sigma, u) + \int_0^T \big\{ \big(\partial_t u, \lambda\big)_\Omega + A(\sigma, u, \lambda) - F(\lambda) \big\} \, \text{d}t$$

- Time reversal $t = T - \tau$ yields adjoint problem as IVP for the heat eq.

Find $\lambda \in \mathcal{U}$, $\begin{cases} \big(w, \partial_t \lambda\big)_\Omega + A(\sigma, w, \lambda) = -\big((u - u^{\text{m}})(T - \bullet), w\big)_\Gamma & \tau \in [0, T] \quad \forall w \in H^1(\Omega) \\ \lambda(\cdot, 0) = 0 & \text{in } \Omega \end{cases}$

and time convolution form of the objective function derivatives

Adjoint-then-discretize, or vice versa

Discretized setting of constrained optimization problem, time-dependent example:

- Objective function:

$$\widehat{J}(\sigma) = J(\sigma, u_1, \ldots, u_N)$$

$$\widehat{J}'(\sigma)s = \partial_\sigma J(\sigma, u_1, \ldots, u_N)s + \sum_{n=1}^{N} \partial_{u_n} J(\sigma, u_1, \ldots, u_N) u_n'(\sigma)s$$

- Forward problem with backward Euler (implicit) time stepping, $h := \Delta t$:

$$Mu_0 = 0, \quad \big(M + hK(\sigma)\big) u_{n+1} = Mu_n + F_{n+1} \qquad n = 0, 1, \ldots, N-1$$

- Forward sensitivity problem:

$$Mu_0' = 0, \quad \big(M + hK(\sigma)\big) u_{n+1}' = Mu_n' - hK(s)u_{n+1} \qquad n = 0, 1, \ldots, N-1 \qquad (\mathcal{S}_n)$$

- Adjoint problem (can be found from relevant Lagrangian):

$$M\lambda_N = 0, \quad \big(M + hK(\sigma)\big)\lambda_n = M\lambda_{n+1} - \partial_{u_{n+1}} J(\sigma, \ldots)s \qquad n = N-1, \ldots, 0 \qquad (\mathcal{A}_n)$$

- Combine, use initial and final conditions:

$$\sum_{n=0}^{N-1} \Big\{ \lambda_n^{\mathsf{T}}(\mathcal{S}_n) - u_{n+1}'^{\mathsf{T}}(\mathcal{A}_n) \Big\} = \sum_{n=1}^{N} \Big\{ \partial_{u_n} J(\sigma, u_1, \ldots, u_N) u_n'(\sigma)s - h\lambda_{n-1}^{\mathsf{T}} K(s) u_n \Big\}$$

Objective function derivatives using discrete adjoint method:

$$\widehat{J}'(\sigma)s = \partial_\sigma J(\sigma, u_1, \ldots, u_N)s + h \sum_{n=1}^{N} \lambda_{n-1}^{\mathsf{T}} K(s) u_n$$

Second-order derivatives

Minimization under linear PDE constraint:

$$\min_{p,u} J(p,u) \quad \text{s.t.} \quad A(p,u,w) - F(w) = 0 \text{ for all } w \in \mathcal{U}$$

- Second-order derivative $\widehat{J}''(p)(q,r)$ of $\widehat{J}(p) := J(u(p))$? Useful e.g. to
  - ▷ Find q solving $\widehat{J}''(p)(q,r) + \widehat{J}'(p)r = 0$ for all r
    (Newton step / descent dir. based on $\partial_p \widehat{J}(p) = 0$, minimize quadratic approx. of $\widehat{J}$...)
  - ▷ Evaluate / check second-order optimality conditions
- Main idea: differentiate first-order derivative. Adjoint-solution approach (for $J = J(u)$) gives

  $$\widehat{J}'(p)r = -\partial_p A\big(p, u(p), \lambda(p)\big)r$$

  Then

  $$\widehat{J}''(p)(q,r) = -\partial_p A\big(p, u'(p)q, \lambda(p)\big)r - \partial_p A\big(p, u(p), \lambda'(p)q\big)r - \partial_{pp} A\big(p, u(p), \lambda(p)\big)(q,r)$$

  with the forward and adjoint solution derivatives u' and λ' satisfying

  $$A\big(u'(p)q, w, p\big) = -\partial_p A\big(u(p), w, p\big)q \qquad \text{for all } w \in \mathcal{U}$$

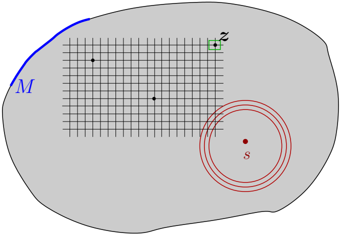
  $$A\big(w, \lambda'(p)q, p\big) = -\partial_p A\big(u(p), w, p\big)q \qquad \text{for all } w \in \mathcal{U}$$

For P-dim. parameter approximation setting:

- ▷ Evaluation of $\widehat{J}''(p)(q,\cdot)$: needs $u(p), \lambda(p), u'(p)q, \lambda'(p)q$ (4 solutions)
- ▷ Evaluation of $\widehat{J}''(p)$: needs $u(p), \lambda(p)$, and $u'(p)r, \lambda'(p)r$ for all r ($2+2P$ solutions)

Method used e.g. in full-waveform inversion [Metivier et al. 2012].

Wave-based identification



- Standard approach: PDE-constrained minimization of (e.g. output least-squares) cost functional:

$$\min_B \mathcal{J}(B), \quad \text{e.g. } \mathcal{J}(B) := J(u_\mathrm{B}) = \frac{1}{2}\int_M \left| u_\mathrm{B} - u_\mathrm{obs} \right|^2 \mathrm{d}M$$

  Entails repeated evaluations of forward (and adjoint if gradient-based) solutions u_B

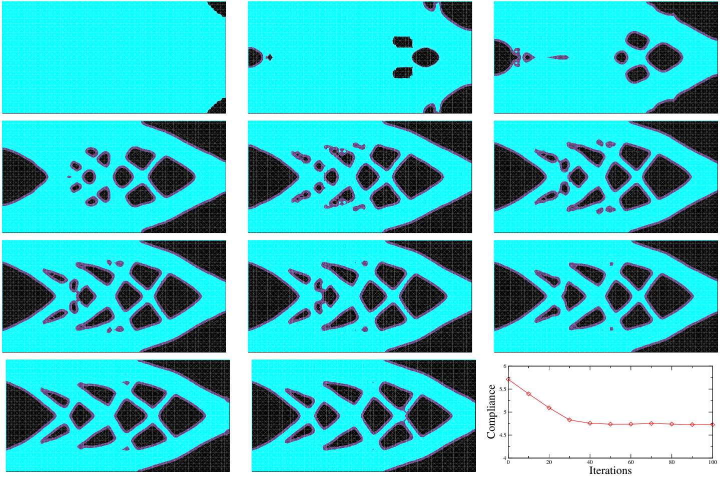
- Impetus for development of non-iterative, qualitative, sampling-based identification. General idea: focus on finding the support of B, define a function ϕ to determine whether $\boldsymbol{z} \notin B$ or $\boldsymbol{z} \in B$ based on value $\phi(\boldsymbol{z})$ at *sampling points* $\boldsymbol{z}$.
  - Linear sampling method (Colton, Kirsch '96), factorization method (Kirsch'98)
    Mathematically well-justified; require abundant data
  - Topological derivative (this talk):
    Conditional partial mathematical justification so far; any overdetermined data

A Qualitative Approach to Inverse Scattering Theory (F. Cakoni, D. Colton, 2014)

Topological derivative concept (topology optimization)

TD: sensitivity analysis tool [Eschenauer et al 94; Sokolowski, Zochowski 99; Garreau et al 01...]

- Initially introduced and applied for topology optimization

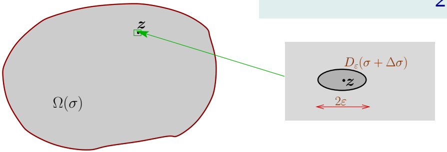


- Later proved useful also for qualitative flaw identification

PhD G. Delgado (2014), adv. G. Allaire — Topology optimization combining topological and shape derivatives

Topological derivative concept (conducting inclusion identification)

- Objective functional (as example): $\widehat{J}(D) = J(u_D) = \dfrac{1}{2}\displaystyle\int_{\partial\Omega} |u_D - u^{\mathrm{m}}|^2 \, \mathrm{d}S$



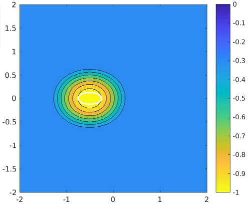
- Consider small trial penetrable inclusion $D_\varepsilon = \boldsymbol{z} + \varepsilon\mathcal{D} \Subset \Omega$ ($\mathcal{D}, \Delta\sigma$ prescribed)

u: background, u_D: (true or trial) object D, u_ε: small trial object D_ε

- J has expansion $J(u_\varepsilon) = J(u) + \varepsilon^d \mathcal{T}(\boldsymbol{z}) + o(\varepsilon^d)$.
- $\mathcal{T}(\boldsymbol{z}) = \mathcal{T}(\boldsymbol{z}; \mathcal{D}, \sigma, \Delta\sigma)$: **topological derivative** (TD) of J at $\boldsymbol{z} \in \Omega$.

Heuristic idea for flaw imaging: find D by seeking regions in Ω where $\mathcal{T}(\boldsymbol{z}; \mathcal{D})$ is **most negative**.

- Proposed defect indicator function: $\boldsymbol{z} \mapsto \mathcal{T}(\boldsymbol{z}; \mathcal{D})$
- Qualitative estimation (location, size, number) of buried objects.
- Heuristic involves both magnitude and sign of $\mathcal{T}(\boldsymbol{z}; \mathcal{D})$



TD as imaging functional for qualitative flaw identification

- Connects qualitative inverse scattering to PDE-constrained approaches
  - ▷ Framework allowing to use any available data
  - ▷ Most developments to date in connection with (acoustic, elastodynamic, electromagnetic) linear wave models
  - ▷ Bounded or unbounded propagation domains
- Many empirical validations of this heuristic *even for macroscopic defects*
  - ▷ using synthetic data (many references)
  - ▷ using experimental data [Tokmashev, Tixier, Guzina 2013].
  - ▷ simplified variants, experimental data [Dominguez et al. 2005, Rodriguez et al. 2014].
- Heuristic proved in some cases:
  - ▷ far-field data, medium perturbation in zeroth-order PDE term, “moderate” scatterers [Bellis, B, Cakoni 2013],
  - ▷ near-field data, medium perturbation in leading-order PDE term, “moderate” scatterers [B, Cakoni 2019; B. 2022],
  - ▷ small obstacles [Ammari et al. 2012, Wahab 15],
- High-frequency behavior [Guzina, Pourahmadian 2015] (TD emphasizes *boundaries*).

- Usually want to compute the field $\boldsymbol{z} \mapsto \mathcal{T}(\boldsymbol{z}) \implies$ Purely numerical evaluation **impractical**
- Analytical TD formulas rely on asymptotic approximation of u_ε. Several approaches, e.g.:
  - ▷ Isolate finite region $C \subset \Omega$ around $D_\varepsilon(\boldsymbol{z})$, asymptotic form of DtN on ∂C
  - ▷ Find asymptotic form of shape derivative (homothetic dilatation of D_ε) as $\varepsilon \to 0$;
  - ▷ Find asymptotic form of (volume or boundary) integral equation (this lecture)

Topological derivative: evaluation using adjoint solution

Focus here on medium perturbations in leading-order PDO term

- $\mathcal{T}(\boldsymbol{z})$ found as leading-order contribution in $J'(u)(u_\varepsilon - u)$ as $\varepsilon \to 0$
- Variational formulations for background, perturbed and adjoint solutions:

$$A(u,w) = F(w) \qquad \forall w \in \mathcal{U} \qquad \text{(background)}$$

$$\big(\Delta\sigma \boldsymbol{\nabla} u_\varepsilon, \boldsymbol{\nabla} w\big)_{D_{\varepsilon,z}} + A(u_\varepsilon, w) = F(w) \qquad \forall w \in \mathcal{U} \qquad \text{(perturbed)}$$

$$A(w,\lambda) = J'(u)w \qquad \forall w \in \mathcal{U} \qquad \text{(adjoint)}$$

- Combine with $w = \lambda,\ -\lambda,\ u_\varepsilon - u$:

$$\left.\begin{aligned} A(u,\lambda) &= F(\lambda) \\ -\big(\Delta\sigma \boldsymbol{\nabla} u_\varepsilon, \boldsymbol{\nabla} \lambda\big)_{D_{\varepsilon,z}} - A(u_\varepsilon,\lambda) &= -F(\lambda) \\ A(u_\varepsilon - u,\lambda) &= J'(u)(u_\varepsilon - u) \end{aligned}\right\} \quad \partial_u J(u)(u_\varepsilon - u) = -\big(\Delta\sigma \boldsymbol{\nabla} u_\varepsilon, \boldsymbol{\nabla} \lambda\big)_{D_{\varepsilon,z}}$$

- Use $\boldsymbol{\nabla}\lambda = \boldsymbol{\nabla}\lambda(\boldsymbol{z}) + o(1)$ and (known) asymptotic approximation in $D_{\varepsilon,z}$ of the form

$$\begin{cases} u_\varepsilon(\boldsymbol{x}) = u(\boldsymbol{z}) + \varepsilon\, \boldsymbol{U}(\boldsymbol{x}/\varepsilon)\cdot\boldsymbol{\nabla} u(\boldsymbol{z}) + o(\varepsilon) \\ \boldsymbol{\nabla} u_\varepsilon(\boldsymbol{x}) = \qquad \boldsymbol{\nabla}\boldsymbol{U}(\boldsymbol{x}/\varepsilon)\cdot\boldsymbol{\nabla} u(\boldsymbol{z}) + o(1) \end{cases} \text{in } D_{\varepsilon,z} \qquad \big(\boldsymbol{U} = \boldsymbol{U}(\bullet;\mathcal{D},\sigma,\Delta\sigma)\big)$$

- Topological derivative:

$$J(u_\varepsilon) = J(u) + \varepsilon^d \mathcal{T}(\boldsymbol{z}) + o(\varepsilon^d) \quad \text{with} \quad \boxed{\mathcal{T}(\boldsymbol{z}) = -\boldsymbol{\nabla} u(\boldsymbol{z})\cdot\mathbf{A}\cdot\boldsymbol{\nabla}\lambda(\boldsymbol{z})}$$

$\mathbf{A} = \mathbf{A}(\mathcal{D},\sigma,\Delta\sigma) = \int_{\mathcal{D}} \Delta\sigma \boldsymbol{\nabla}\boldsymbol{U}\, \mathrm{d}V$: polarization tensor (known analytically for simple $\mathcal{D}$)

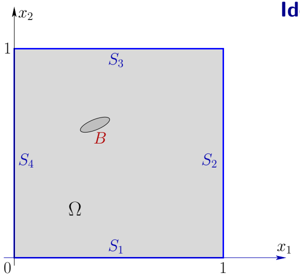
Topological derivative: evaluation using adjoint solution

$J(u_\varepsilon) = J(u) + \varepsilon^d \mathcal{T}(\boldsymbol{z}) + o(\varepsilon^d)$ with $\boxed{\mathcal{T}(\boldsymbol{z}) = -\boldsymbol{\nabla} u(\boldsymbol{z}) \cdot \mathbf{A} \cdot \boldsymbol{\nabla} \lambda(\boldsymbol{z})}$

- In practice: (i) compute background and adjoint solutions, (ii) evaluate $\boldsymbol{z} \mapsto \mathcal{T}(\boldsymbol{z})$
- Similar TD formulas in other cases (Maxwell, elasticity, time-harmonic or transient waves...)

- **Computation of the TD field:**
  - Background and adjoint solutions $u, \hat{u}$ defined on **same** (reference) configuration
    $\Longrightarrow$ Evaluation of $\hat{\hat{u}}$ and $\mathcal{T}$ at moderate extra cost
  - Computation of $\mathcal{T}(\boldsymbol{z})$ straightforward with **standard** methods (FEM, BEM...)
  - Experimental information exploited via the adjoint solution $\hat{u}$
- **Corresponding results available in many other cases, e.g.**
  - Potential problems, elasticity: B, Delgado (2014); Delgado, B (2015); Garreau, Guillaume, Masmoudi (2001); Novotny et al. (2003); Sokolowski, Zochowski (2001); Vogelius, Volkov (2000); Schneider, Andrä (2013);
  - Acoustics: Feijoo (2004); Guzina, B (2006); Nemitz, MB (2008)
  - Electromagnetism: Masmoudi, Pommier, Samet (2005);
  - Elastodynamics: Guzina, B (2004); Guzina, Chikichev (2007); Bellis, Impériale (2013)
  - Time domain: Dominguez, Gibiat, Esquerré (2005); B (2006); Amstutz, Takahashi, Wexler (2008); Tokmashev, Tixier, Guzina (2013)
  - Cracks: Amstutz, Horchani, Masmoudi (2005); Bellis, B (2012); B (2011)
  - Image processing: Auroux, Jaafar Belaid, Rjaibi (2010); Larnier, Masmoudi (2013)
  - Related imaging functionals: Rodriguez, Sahuguet, Gibiat, Jacob (2012)

Example: FEM-based computation of $\mathcal{T}$ (2D time-domain wave eqn.)



Identification of impenetrable scatterer(s) in 2-D acoustic medium.

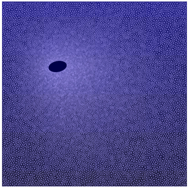
- Normalized scalar wave equation

$$\Delta u_{\rm B}^{(k)} - \partial_{tt} u_{\rm B}^{(k)} = 0$$

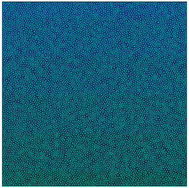
$$\partial_n u_{\rm B}^{(k)} = \begin{cases} 1 & (\text{on } S_k) \\ 0 & (\text{on } S_j,\ j \neq k) \end{cases} \qquad \partial_n u_{\rm B}^{(k)} = 0 \quad (\text{on } \partial B)$$

- Matlab (very simple) implementation, T3 elements.
- Newmark time-marching ($\beta = \frac{1}{4}, \gamma = \frac{1}{2}$, inconditionally stable)
- $M = S_1 \cup S_2 \cup S_3 \cup S_4$ (simulated measurements for $0 \leq t \leq T$)

$$\mathcal{J}^{(k)}(B) = \frac{1}{2} \int_0^T \int_{S_1+S_2+S_3+S_4} |u_{\rm B}^{(k)} - u_{\rm obs}^{(k)}|^2 \, \mathrm{d}s \, \mathrm{d}t$$



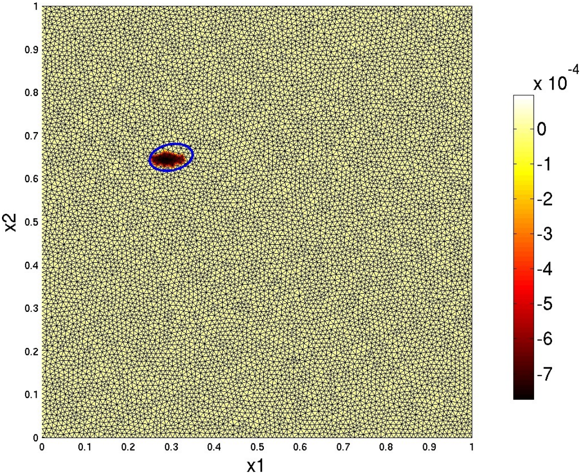
Computation of synthetic data $u_{obs}^{(k)}$



Computation of $u^{(k)}$ and $\hat{u}^{(k)}$

Bellis, B, *Int. J. Solids Struct.* (2010)

FEM-based computation of $\mathcal{T}$: identification of a single scatterer



Thresholded TS field:

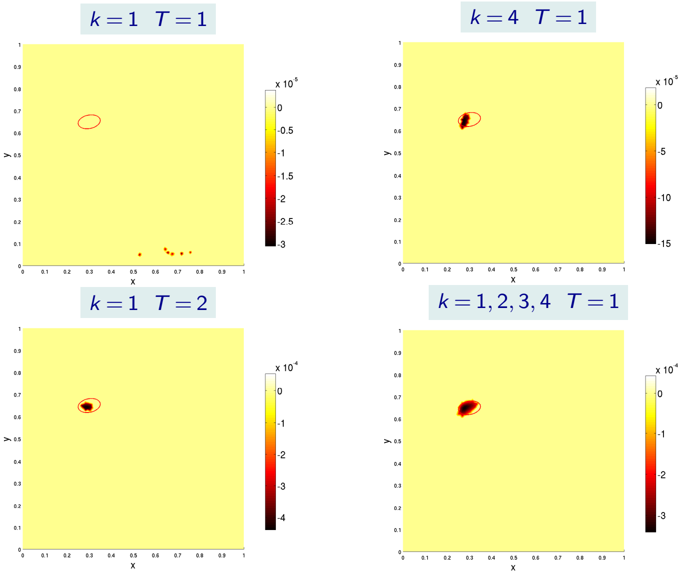
$$\hat{\mathcal{T}}_3 = \mathcal{T}_3 \quad (\mathcal{T}_3 \leq \alpha \mathcal{T}_3^{\text{Min}})$$
$$= 0 \quad (\mathcal{T}_3 > \alpha \mathcal{T}_3^{\text{Min}})$$

$$k = 1, \; T = 2$$
$$\alpha = .75$$

$$\mathcal{T}(\boldsymbol{z}) = \Big(2\pi \boldsymbol{\nabla} u^{(k)} \star \boldsymbol{\nabla} \hat{u}^{(k)} + \frac{4\pi}{3} \frac{1}{c^2} u^{(k)} \star \hat{u}^{(k)} \Big)(\boldsymbol{z}, T)$$

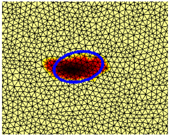
Bellis, B, *Int. J. Solids Struct.* (2010)

FEM-based computation of $\mathcal{T}$: identification of a single scatterer

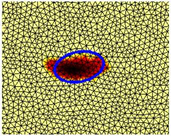


Bellis, B, *Int. J. Solids Struct.* (2010)

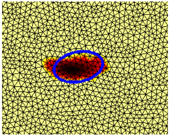
FEM-based computation of $\mathcal{T}$: identification of a single scatterer



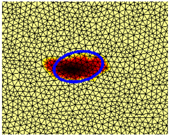
(a) $\alpha = 0.1$



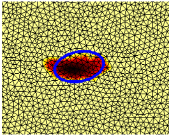
(b) $\alpha = 0.2$



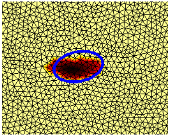
(c) $\alpha = 0.3$



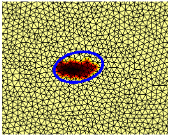
(d) $\alpha = 0.4$



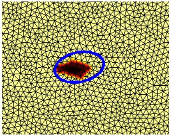
(e) $\alpha = 0.5$



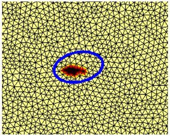
(f) $\alpha = 0.6$



(g) $\alpha = 0.7$



(h) $\alpha = 0.8$

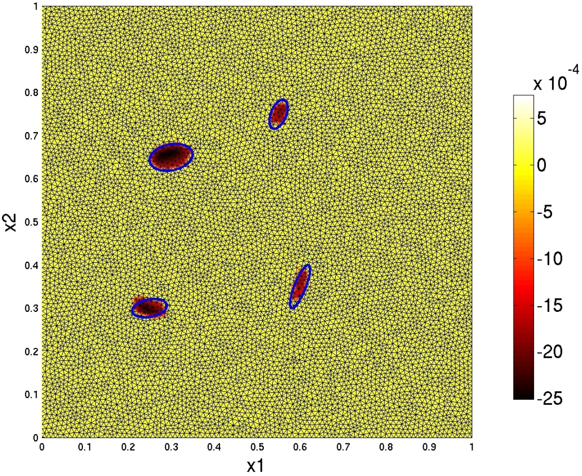


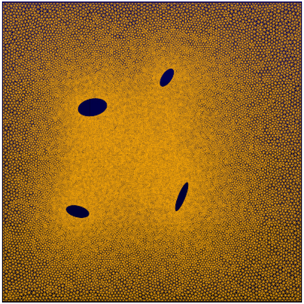
(i) $\alpha = 0.9$

Influence of cut-off parameter α.

Bellis, B., *Int. J. Solids Struct.* (2010)

FEM-based computation of $\mathcal{T}$: simultaneous identification of a multiple scatterer





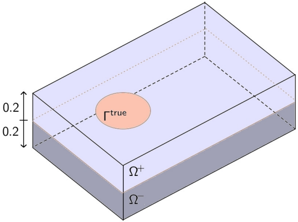
Computation of synthetic data $u^{(k)}$

$k = 1, 2, 3, 4, \;\; T = 2, \;\; \alpha = 0.5$

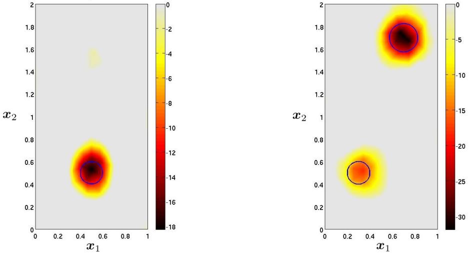
Bellis, B, *Int. J. Solids Struct.* (2010)

Example: Imaging of interface cracks (3-D transient elasticity, FEM)

- FEM-based time domain 3D simulations in *stiff/soft* bi-material domain
- Gaussian time distribution of compressional loading on top face
- Adimensionalization w.r.t. longitudinal wave velocity, $T = 1$
- Observation on top face



- Topological derivative $\mathbb{T}(\cdot, T) \leqslant 0$ at interface



- Extension to multilayered domains
- Study of other type of interface, e.g. fiber reinforced composites

Bellis (2011)

Experimental study (Tokmashev, Tixier, Guzina 2013)

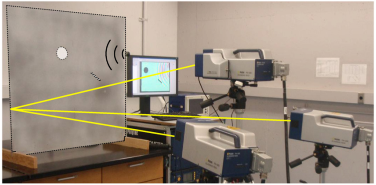


Figure 2. Three-dimensional motion sensing via laser Doppler vibrometer (LDV) system.

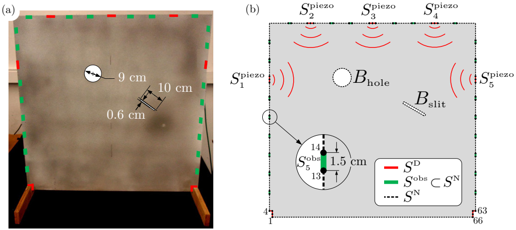
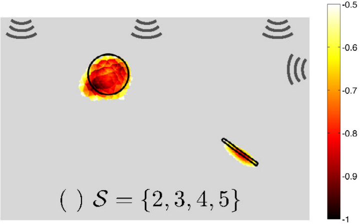


Figure 4. Testing configuration: (a) photograph of the damaged plate, and (b) boundary conditions and spatial arrangement of the LDV scan points for five individual source locations ($S_k^{\text{piezo}}, k = \overline{1,5}$).



Topological derivative $\boldsymbol{z} \mapsto \mathcal{T}(\boldsymbol{z})$ for

$$J(\boldsymbol{u}_D) = \frac{1}{2}\int_0^T \int_S |\boldsymbol{u}_D - \boldsymbol{u}^{\text{m}}|^2 \, \mathrm{d}S \, \mathrm{d}t$$

Closing remark: different derivatives but same adjoint solution

- Assume fixed basic setting (physical configuration, objective function).

  e.g. conductivity problem with objective function $J(u) = \frac{1}{2}\int_{\partial\Omega} |u - u^{\mathrm{m}}|^2 \,\mathrm{d}S$

  Adjoint solution λ only depends on (choice of) J. Here:
  - ▷ Material parameter perturbation ($\sigma \to \sigma + s$):
  $$\widehat{J}(\sigma + s) - \widehat{J}(\sigma) = \int_{\Omega} s \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} \lambda \,\mathrm{d}V + o(\|s\|)$$
  - ▷ Inclusion shape perturbation ($D \to D + \boldsymbol{\theta}(D)$):
  $$\widehat{J}(D + \boldsymbol{\theta}(D)) - \widehat{J}(D) = \int_{\partial D} \Delta\sigma (\boldsymbol{\nabla}_S u \cdot \boldsymbol{\nabla}_S \lambda) \boldsymbol{\theta} \cdot \boldsymbol{n} \,\mathrm{d}V + o(\|\boldsymbol{\theta}\|)$$
  - ▷ Topology change via small-inclusion nucleation ($\emptyset \to D_{\varepsilon,z}(\Delta\sigma)$):
  $$\widehat{J}(D_{\varepsilon,z}) - \widehat{J}(\emptyset) = \varepsilon^d \, \boldsymbol{\nabla} u(\boldsymbol{z}) \cdot \mathbf{A}(\mathcal{D}, \sigma, \Delta\sigma^\star) \cdot \boldsymbol{\nabla} \lambda(\boldsymbol{z}) + o(\|\varepsilon^d\|)$$
- Formulas are bilinear in same (forward and adjoint) solutions, differ in details.
- One adjoint solution per objective function, applies (even simultaneously) to all types of sensitivity

See e.g. Céa, Garreau, Guillaume, Masmoudi 2000

Thank you for listening!

Any questions?