Optimization in the presence of uncertainty: applications in designing random heterogeneous media

P.S. Koutsourelakis

Continuum Mechanics Group
Technical University of Munich, Germany

Boston, February 26 2013

Deterministic topology optimization

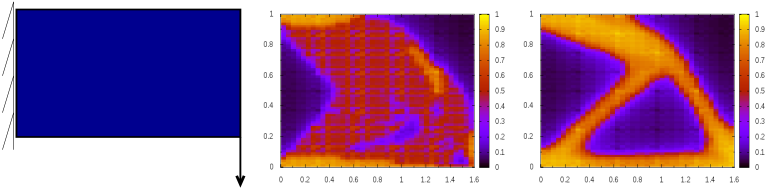
Shape/topology optimization:

$$\min_{\boldsymbol{d}} \quad compliance(\boldsymbol{d}) = \boldsymbol{f}^T \boldsymbol{u}(\boldsymbol{d})$$

such that:

$$\boldsymbol{K}(\boldsymbol{d})\boldsymbol{u}(\boldsymbol{d}) = \boldsymbol{f} \quad \text{(governing equation)}$$
$$\int d(\boldsymbol{x})\, d\boldsymbol{x} \leq V_0, \quad \text{(volume fraction)}$$
$$d(\boldsymbol{x}) \in [0,1]$$

$$d(\boldsymbol{x}) = \begin{cases} 1, & material \\ 0, & void \end{cases}$$



(a) domain (b) $compliance(\boldsymbol{d}) \approx 50$ (c) $compliance(\boldsymbol{d}) \approx 55$

Figure: Adjoint-based gradient optimization - $O(1000)$ forward runs

Stochastic topology optimization

Shape/topology optimization:

$compliance(\boldsymbol{d}, \boldsymbol{\theta}) = \boldsymbol{f}^T \boldsymbol{u}(\boldsymbol{d}, \boldsymbol{\theta})$

$\boldsymbol{K}(\boldsymbol{d}, \boldsymbol{\theta})\boldsymbol{u}(\boldsymbol{d}, \boldsymbol{\theta}) = \boldsymbol{f}$ (governing equation)

$\int d(\boldsymbol{x})\, d\boldsymbol{x} \leq V_0$, (volume fraction)

$d(\boldsymbol{x}) \in [0, 1]$

$$d(\boldsymbol{x}) = \begin{cases} 1, & material \\ 0, & void \end{cases}$$

$\boldsymbol{\theta} \sim \pi(\boldsymbol{\theta})$, (random material properties)

Stochastic topology optimization

Average design: $\min_{\boldsymbol{d}} E[compliance(\boldsymbol{d}, \boldsymbol{\theta})] = \int \boldsymbol{f}^T \boldsymbol{u}(\boldsymbol{d}, \boldsymbol{\theta})\, \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$

or:

Robust design: $\min_{\boldsymbol{d}} Pr[compliance(\boldsymbol{d}, \boldsymbol{\theta}) \geq c_0] = \int 1_{\boldsymbol{f}^T \boldsymbol{u}(\boldsymbol{d}, \boldsymbol{\theta}) \geq c_0}\, \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$

<u>such that:</u>

$\boldsymbol{K}(\boldsymbol{d})\boldsymbol{u}(\boldsymbol{d}) = \boldsymbol{f}$ (governing equation)

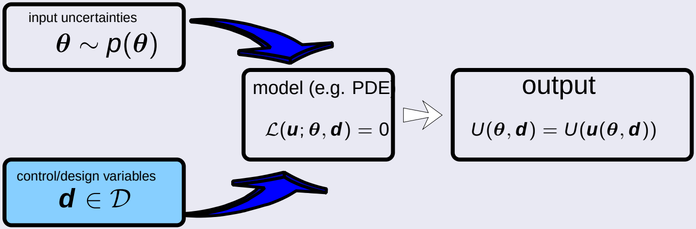
$\int d(\boldsymbol{x})\, d\boldsymbol{x} \leq V_0$, (volume fraction)

$d(\boldsymbol{x}) \in [0, 1]$

$\boldsymbol{\theta} \sim \pi(\boldsymbol{\theta})$

Motivation

Uncertainty quantification



- uncertainties $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$, $n_\theta >> 1$
- design/control variables $\boldsymbol{d} \in \mathcal{D} \subset \mathbb{R}^{n_d}$, $n_d >> 1$
- Goal - Design/Optimization: Can we *efficiently* optimize w.r.t $\boldsymbol{d}$ and some output utility $U(\boldsymbol{\theta}, \boldsymbol{d})$

$$V(\boldsymbol{d}) = \int U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

Motivation

Optimize the *expected* utility $V(\boldsymbol{d})$:

$$V(\boldsymbol{d}) = \int U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

Solution Strategies:

- Extend deterministic optimization tools:
  - Draw N samples from a distribution $q(\theta)$ (usually $q(\theta) = p(\theta)$)
  - Optimize $\log V_N(\boldsymbol{d}) = \frac{1}{N} \sum_{i=1}^{N} \log \frac{U(\theta_i, \boldsymbol{d}) p(\theta_i)}{q(\theta_i)}$
  - After some algebra:

$$\log V(\boldsymbol{d}) = \log V_N(\boldsymbol{d}) + KL(q(\theta) | p(\theta | \boldsymbol{d}), \quad p(\theta | \boldsymbol{d})) = \frac{U(\theta, \boldsymbol{d}) p(\theta)}{V(\boldsymbol{d})}$$

- Surrogate (reduced-order) models

Motivation

Optimize the *expected* utility $V(\boldsymbol{d})$:

$$V(\boldsymbol{d}) = \int U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

Solution Strategies:

- Extend deterministic optimization tools:
  - Draw N samples from a distribution $q(\boldsymbol{\theta})$ (usually $q(\boldsymbol{\theta}) = p(\boldsymbol{\theta})$)
  - Optimize $\log V_N(\boldsymbol{d}) = \frac{1}{N} \sum_{i=1}^{N} \log \frac{U(\boldsymbol{\theta}_i, \boldsymbol{d}) p(\boldsymbol{\theta}_i)}{q(\boldsymbol{\theta}_i)}$
  - After some algebra:

$$\log V(\boldsymbol{d}) = \log V_N(\boldsymbol{d}) + KL(q(\boldsymbol{\theta})|p(\boldsymbol{\theta}|\boldsymbol{d}), \quad p(\boldsymbol{\theta}|\boldsymbol{d})) = \frac{U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})}{V(\boldsymbol{d})}$$

- Surrogate (reduced-order) models

Motivation

Optimize the *expected* utility $V(\boldsymbol{d})$:

$$V(\boldsymbol{d}) = \int U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

Solution Strategies:

- Extend deterministic optimization tools:
  - Draw N samples from a distribution $q(\boldsymbol{\theta})$ (usually $q(\boldsymbol{\theta}) = p(\boldsymbol{\theta})$)
  - Optimize $\log V_N(\boldsymbol{d}) = \frac{1}{N} \sum_{i=1}^{N} \log \frac{U(\boldsymbol{\theta}_i, \boldsymbol{d}) p(\boldsymbol{\theta}_i)}{q(\boldsymbol{\theta}_i)}$
  - After some algebra:

$$\log V(\boldsymbol{d}) = \log V_N(\boldsymbol{d}) + KL(q(\boldsymbol{\theta}) | p(\boldsymbol{\theta}|\boldsymbol{d}), \quad p(\boldsymbol{\theta}|\boldsymbol{d})) = \frac{U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})}{V(\boldsymbol{d})}$$

- Surrogate (reduced-order) models

Motivation

Optimize the *expected* utility $V(\boldsymbol{d})$:

$$V(\boldsymbol{d}) = \int U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})\, d\boldsymbol{\theta}, \quad \boldsymbol{J} = \frac{\partial V(\boldsymbol{d})}{\partial \boldsymbol{d}} = \int \frac{\partial U(\boldsymbol{\theta}, \boldsymbol{d})}{\partial \boldsymbol{d}} p(\boldsymbol{\theta})\, d\boldsymbol{\theta}$$

Solution Strategies:

- Stochastic Approximation (Robbins & Monro 1951)
  - Perform gradient ascent/descent i.e.:

$$\boldsymbol{d}^{(k+1)} = \boldsymbol{d}^{(k)} \pm \alpha_k \hat{\boldsymbol{J}}(\boldsymbol{d}^{(k)})$$

  where:
    - $\alpha_k > 0$, $\alpha_k \to 0$, $\sum_{k=0}^{\infty} \alpha_k = +\infty$ and $\sum_{k=0}^{\infty} \alpha_k^2 < +\infty$.
    - $\hat{\boldsymbol{J}}(\boldsymbol{d}^{(k)}) = \text{unbiased estimator} \left(\frac{\partial V}{\partial \boldsymbol{d}} = \int \frac{\partial U(\boldsymbol{\theta}, \boldsymbol{d})}{\partial \boldsymbol{d}} p(\boldsymbol{\theta})\, d\boldsymbol{\theta} \right)$ (i.e. with Monte Carlo and a single $\boldsymbol{\theta}-$sample

Approach

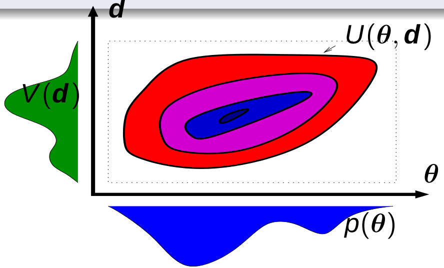
Optimize the *expected* utility $V(\boldsymbol{d})$:

$$V(\boldsymbol{d}) = \int U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}) \, d\boldsymbol{\theta}$$

We adopt a *sampling* approach (*Müller 1999*) in the joint $\boldsymbol{\theta} \times \boldsymbol{d}$ space [a]:

$$\pi(\boldsymbol{\theta}, \boldsymbol{d}) \propto U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})$$

Note that the $\boldsymbol{d}$-coordinates of $(\boldsymbol{\theta}, \boldsymbol{d})$ samples from $p(\boldsymbol{\theta}, \boldsymbol{d})$ will concentrate on the maxima of V.



[a] $U(\boldsymbol{\theta}, \boldsymbol{d})$ is assumed positive or in general bounded from below

Approach

the good:

- dimensionality becomes less of an issue (Monte Carlo samplers are the best tools we have to deal with high-dimensional problems)
- *sensitivity* w.r.t $\boldsymbol{d}$ can be assessed.

the bad:

- we have to work on the joint space $\theta \times \boldsymbol{d}$
- Monte Carlo can be very demanding in terms of forward runs.
- multiple local maxima/minima of $V(\boldsymbol{d})$

Approach

the good:

- dimensionality becomes less of an issue (Monte Carlo samplers are the best tools we have to deal with high-dimensional problems)
- *sensitivity* w.r.t $\boldsymbol{d}$ can be assessed.

the bad:

- we have to work on the joint space $\boldsymbol{\theta} \times \boldsymbol{d}$
- Monte Carlo can be very demanding in terms of forward runs.
- multiple local maxima/minima of $V(\boldsymbol{d})$

Approach

Sequential Monte Carlo:

- A combination of Importance sampling and MCMC that provides a particulate approximation $\{(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)}), \boldsymbol{W}^{(i)}\}_{i=1}^{N}$ (e.g. Doucet et al. 2001):

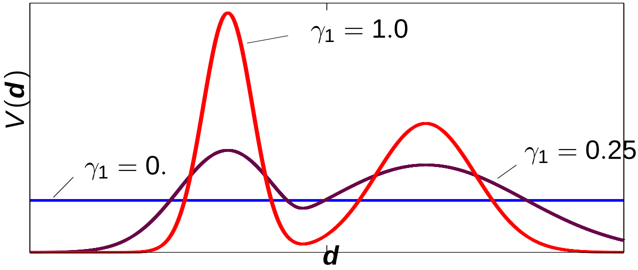
$$\pi(\boldsymbol{\theta}, \boldsymbol{d}) \propto U(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}) \approx \sum_{i=1}^{N} \boldsymbol{W}^{(i)} \delta_{\boldsymbol{\theta}^{(i)}}(\boldsymbol{\theta}) \delta_{\boldsymbol{d}^{(i)}}(\boldsymbol{d})$$

- *almost sure* convergence of expectations of π-measurable functions (Del Moral 2004)
- Operates on a sequence of distributions.
- In contrast to MCMC, it can overcome trapping in modes of the distributions

Approach

We operate on a *sequence* of distributions (from simple to complicated) (Amzal et al 2003, Johansen et al 2006, Kück et al. 2006):

$$\pi_\gamma(\boldsymbol{\theta}, \boldsymbol{d}) \propto U^\gamma(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}), \quad \gamma \in [0, 1]$$



Approach

We operate on a *sequence* of distributions (from simple to complicated):

$$\pi_\gamma(\boldsymbol{\theta}, \boldsymbol{d}) \propto U^\gamma(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta}), \quad \gamma \in [0, 1]$$

***Adaptive* SMC (PSK, *J. Comp. Phys.* 2009, Sternfels, PSK, *Int. J. Mult. Comp. Eng* 2010):**

- If γ increases slowly, we do too many forward runs (cost)
- If γ increases too fast we loose accuracy (accuracy)

Adaptive SMC

- Generate initial particle population $\{(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)}), W^{(i)}\}_{i=1}^{N}$ from $\pi_{\gamma=0} \equiv p(\boldsymbol{\theta})$. Set $\gamma_{current} = 0$.
- Iterate until $\gamma_{current} = 1$.
  - **Reweight**: Find γ_{next} based on the relative reduction in the Effective Sample Size ESS :

$$w^{(i)} = W^{(i)} \frac{\pi_{\gamma_{next}}(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)})}{p_{\gamma_{current}}(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)})}, \quad ESS = \frac{(\sum_{i=1}^{N} w^{(i)})^2}{\sum_{i=1}^{N} (w^{(i)})^2}$$

  - **Resample:** If ESS drops below a specified threshold (typically $N/2$) , then resample.
  - **Rejuvenate:** Move particles using a $p_{\gamma_{next}}$-invariant MCMC kernel:
    - We employed a Metropolis-adjusted Langevin (MALA) sampler which implies calculation of U as well as derivatives $\frac{\partial U}{\partial \theta}$
    - These were calculated using *adjoint formulations*
  - Set $\gamma_{current} = \gamma_{next}$

Adaptive SMC

- Generate initial particle population $\{(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)}), W^{(i)}\}_{i=1}^{N}$ from $\pi_{\gamma=0} \equiv p(\boldsymbol{\theta})$. Set $\gamma_{current} = 0$.
- Iterate until $\gamma_{current} = 1$.
  - **Reweight**: Find γ_{next} based on the relative reduction in the Effective Sample Size ESS :

$$w^{(i)} = W^{(i)} \frac{\pi_{\gamma_{next}}(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)})}{p_{\gamma_{current}}(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)})}, \quad ESS = \frac{(\sum_{i=1}^{N} w^{(i)})^2}{\sum_{i=1}^{N} (w^{(i)})^2}$$

  - **Resample:** If ESS drops below a specified threshold (typically $N/2$) , then resample.
  - **Rejuvenate:** Move particles using a $p_{\gamma_{next}}$-invariant MCMC kernel:
    - We employed a Metropolis-adjusted Langevin (MALA) sampler which implies calculation of U as well as derivatives $\frac{\partial U}{\partial \theta}$
    - These were calculated using *adjoint formulations*
  - Set $\gamma_{current} = \gamma_{next}$

Adaptive SMC

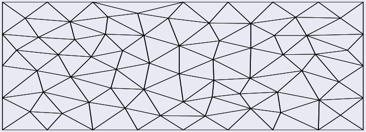
- Generate initial particle population $\{(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)}), W^{(i)}\}_{i=1}^{N}$ from $\pi_{\gamma=0} \equiv p(\boldsymbol{\theta})$. Set $\gamma_{current} = 0$.
- Iterate until $\gamma_{current} = 1$.
  - **Reweight**: Find γ_{next} based on the relative reduction in the Effective Sample Size ESS :

$$w^{(i)} = W^{(i)} \frac{\pi_{\gamma_{next}}(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)})}{p_{\gamma_{current}}(\boldsymbol{\theta}^{(i)}, \boldsymbol{d}^{(i)})}, \quad ESS = \frac{(\sum_{i=1}^{N} w^{(i)})^2}{\sum_{i=1}^{N} (w^{(i)})^2}$$

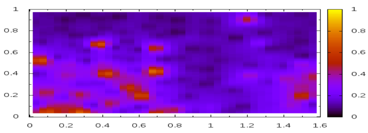
  - **Resample:** If ESS drops below a specified threshold (typically $N/2$) , then resample.
  - **Rejuvenate:** Move particles using a $p_{\gamma_{next}}$-invariant MCMC kernel:
    - We employed a Metropolis-adjusted Langevin (MALA) sampler which implies calculation of U as well as derivatives $\frac{\partial U}{\partial \boldsymbol{\theta}}$
    - These were calculated using *adjoint formulations*
  - Set $\gamma_{current} = \gamma_{next}$

Multi-resolution Adaptive SMC

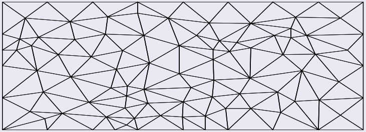
Adaptive refinement driven by inferential uncertainty:



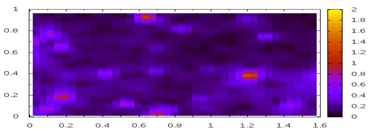
(a) resolution 1



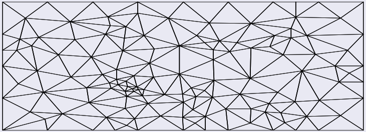
(b) variance of ***d***



(c) resolution 2



(d) variance of ***d***



(e) resolution 3

Multi-resolution Adaptive SMC

- Coarser: $\pi_c(\boldsymbol{d}, \boldsymbol{\theta}) \propto U_c(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})$
- Finer: $\pi_f(\boldsymbol{d}, \boldsymbol{\theta}) \propto U_f(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})$
- Bridge: $\pi_{cf,\gamma}(\boldsymbol{d}, \boldsymbol{\theta}) \propto U_c^{1-\gamma}(\boldsymbol{\theta}, \boldsymbol{d}) U_f^{\gamma}(\boldsymbol{\theta}, \boldsymbol{d}) p(\boldsymbol{\theta})$

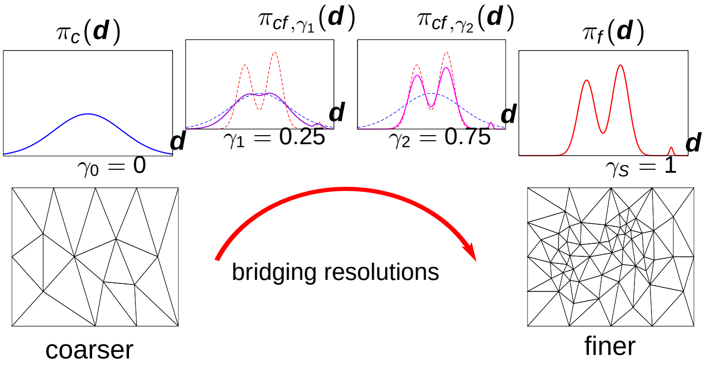


Figure: Resolution-bridging densities (PSK, *J. Comp. Phys.* 2009)

SMC/MCMC on implicitly defined manifold

- MALA:

$$\text{Step 1:} \quad \boldsymbol{d}_2 = \boldsymbol{d}_1 + \frac{\sigma^2}{2}\nabla \log \pi(\boldsymbol{d}_1) + \sigma \boldsymbol{z}_1, \quad \boldsymbol{z}_1 \sim N(0, \boldsymbol{I})$$

$$\left(\boldsymbol{d}_1 = \boldsymbol{d}_2 + \frac{\sigma^2}{2}\nabla \log \pi(\boldsymbol{d}_2) + \sigma \boldsymbol{z}_2, \quad \boldsymbol{z}_2 \sim N(0, \boldsymbol{I})\right)$$

$$\text{Step 2:} \quad p(accept) = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{d}_2 \to \boldsymbol{d}_1)}{\pi(\boldsymbol{d}_1)p(\boldsymbol{d}_1 \to \boldsymbol{d}_2)} = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{z}_2)}{\pi(\boldsymbol{d}_1)p(\boldsymbol{z}_1)}$$

- MALA on manifold $V(\boldsymbol{d}) = V_0$:

$$\text{Step 1:} \quad \boldsymbol{d}_1' = \boldsymbol{d}_1 + \frac{\sigma^2}{2}\nabla \log \pi(\boldsymbol{d}_1) + \sigma \boldsymbol{z}_1, \quad \boldsymbol{z}_1 \sim N(0, \boldsymbol{I}) 1_{\boldsymbol{d}_1' \perp \nabla V(\boldsymbol{d}_1)}$$

$$\text{Step 2:} \quad \boldsymbol{d}_2 = arg\min_{\boldsymbol{d} \in V(\boldsymbol{d}) = V_0} |\boldsymbol{d} - \boldsymbol{d}_1'|^2$$

$$\text{Step 3:} \quad p(accept) = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{d}_2 \to \boldsymbol{d}_2')}{\pi(\boldsymbol{d}_1)p(\boldsymbol{d}_1 \to \boldsymbol{d}_1')} = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{z}_2)}{\pi(\boldsymbol{d}_1)p(\boldsymbol{z}_1)}$$

SMC/MCMC on implicitly defined manifold

- MALA:

Step 1: $\boldsymbol{d}_2 = \boldsymbol{d}_1 + \frac{\sigma^2}{2}\nabla \log \pi(\boldsymbol{d}_1) + \sigma \boldsymbol{z}_1, \quad \boldsymbol{z}_1 \sim N(0, \boldsymbol{I})$

$\left(\boldsymbol{d}_1 = \boldsymbol{d}_2 + \frac{\sigma^2}{2}\nabla \log \pi(\boldsymbol{d}_2) + \sigma \boldsymbol{z}_2, \quad \boldsymbol{z}_2 \sim N(0, \boldsymbol{I})\right)$

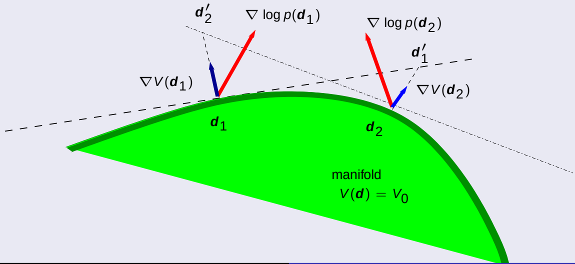
Step 2: $p(accept) = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{d}_2 \to \boldsymbol{d}_1)}{\pi(\boldsymbol{d}_1)p(\boldsymbol{d}_1 \to \boldsymbol{d}_2)} = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{z}_2)}{\pi(\boldsymbol{d}_1)p(\boldsymbol{z}_1)}$

- MALA on manifold $V(\boldsymbol{d}) = V_0$:

Step 1: $\boldsymbol{d}_1' = \boldsymbol{d}_1 + \frac{\sigma^2}{2}\nabla \log \pi(\boldsymbol{d}_1) + \sigma \boldsymbol{z}_1, \quad \boldsymbol{z}_1 \sim N(0, \boldsymbol{I}) 1_{\boldsymbol{d}_1' \perp \nabla V(\boldsymbol{d}_1)}$

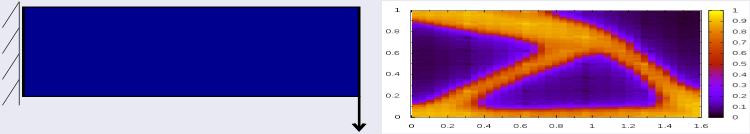
Step 2: $\boldsymbol{d}_2 = arg\min_{\boldsymbol{d} \in V(\boldsymbol{d}) = V_0} |\boldsymbol{d} - \boldsymbol{d}_1'|^2$

Step 3: $p(accept) = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{d}_2 \to \boldsymbol{d}_2')}{\pi(\boldsymbol{d}_1)p(\boldsymbol{d}_1 \to \boldsymbol{d}_1')} = \frac{\pi(\boldsymbol{d}_2)p(\boldsymbol{z}_2)}{\pi(\boldsymbol{d}_1)p(\boldsymbol{z}_1)}$



Numerical Illustration

Deterministic topology optimization



(a) domain (b) $compliance(\boldsymbol{d}) \approx 55$

Figure: Deterministic topology optimization - $O(1000)$ forward runs

Stochastic topology optimization

- $dim(\boldsymbol{d}) = 5120$ (design variables), $dim(\boldsymbol{\theta}) = 5120$ (random variables)
- $\log \boldsymbol{\theta} \sim N(\boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$
  - $E[\theta_i] = 1$
  - $\boldsymbol{\Sigma}_\theta Cov[\log \theta(\boldsymbol{x}_i), \log \theta(\boldsymbol{x}_j)] = e^{-|\boldsymbol{x}_i - \boldsymbol{x}_j|/l_0}$
  - $l_0 = 0.1$ (correlation length)
- Objective: $\boxed{max_{\boldsymbol{d}} Pr[compliance \leq 100]}$

Numerical Illustration

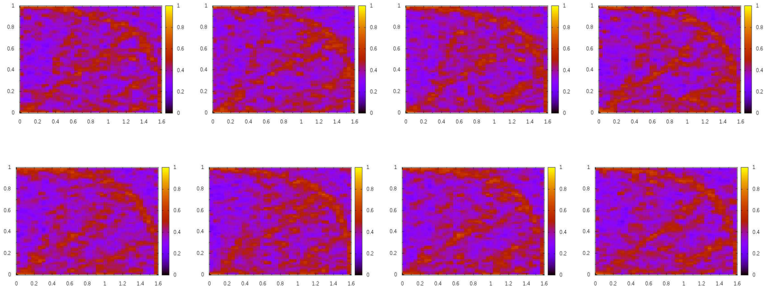
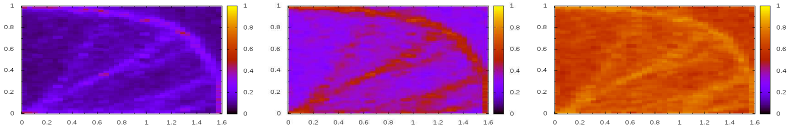


Figure: Particle realizations

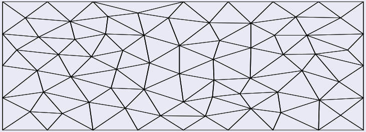


(a) mean-variance (b) mean (c) mean+variance

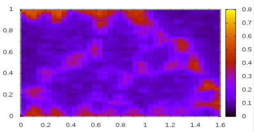
Figure: Stochastic topoloy optimization - $180,000$ forward runs

Numerical Illustration

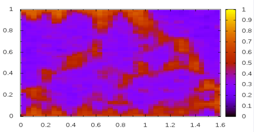
Multi-resolution inference



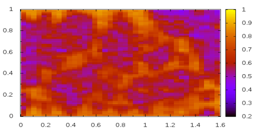
(a)



(b) mean-variance



(c) mean

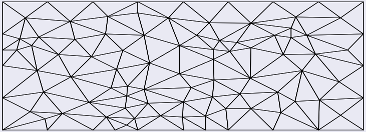


(d) mean+variance

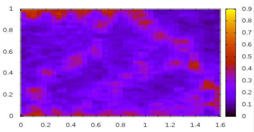
Figure: Resolution 1 - Cost 3 runs of reference(finest) model

Numerical Illustration

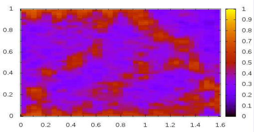
Multi-resolution inference



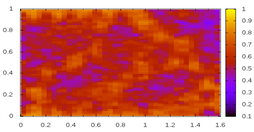
(a)



(b) mean-variance



(c) mean

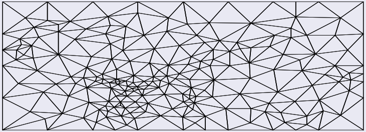


(d) mean+variance

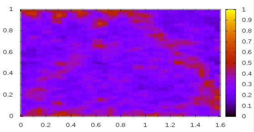
Figure: Resolution 2 - Cost 7 runs of reference(finest) model

Numerical Illustration

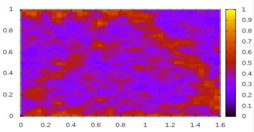
Multi-resolution inference



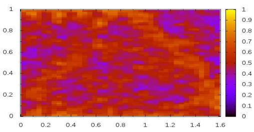
(a)



(b) mean-variance



(c) mean

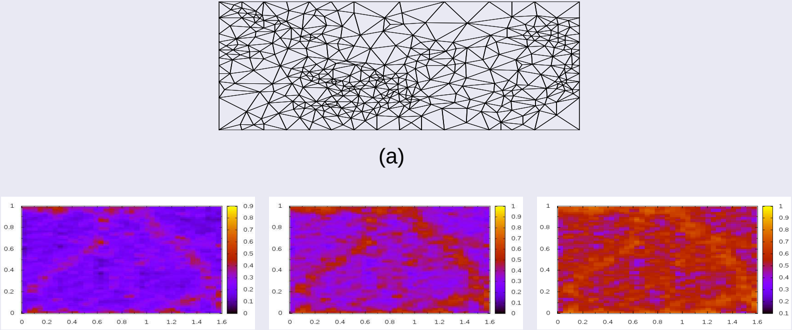


(d) mean+variance

Figure: Resolution 10 - Cost 80 runs of reference(finest) model

Numerical Illustration

Multi-resolution inference

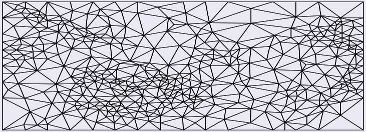


(a)

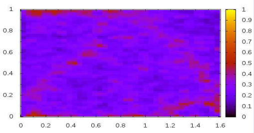
(b) mean-variance (c) mean (d) mean+variance

Figure: Resolution 20 - Cost 506 runs of reference(finest) model

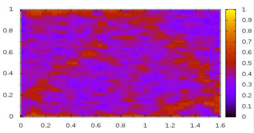
Multi-resolution inference



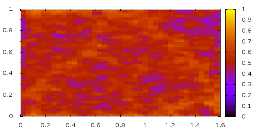
(a)



(b) mean-variance



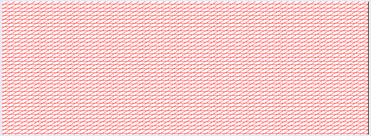
(c) mean



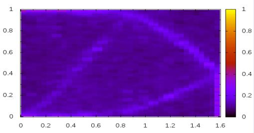
(d) mean+variance

Figure: Resolution 20 - Cost 886 runs of reference(finest) model

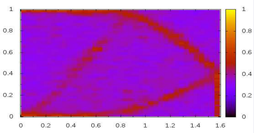
Multi-resolution inference



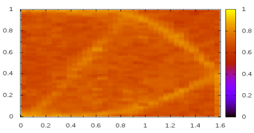
(a)



(b) mean-variance



(c) mean



(d) mean+variance

Figure: Resolution Final - Cost 56, 641 runs of reference (finest) model

Conclusions

- A general, non-intrusive approach that converts the optimization problem to a sampling one.
- Suitable for high-dimensional problems in terms of random and design variables.
- Can make use of derivatives (available through adjoint formulations) to expedite inference
- Can exploit various sequence of distributions (defined through less-expensive, approximate models) to expedite computations
- Open questions:
  - Can we identify all optima?
  - Can we actually estimate $V(\boldsymbol{d})$?