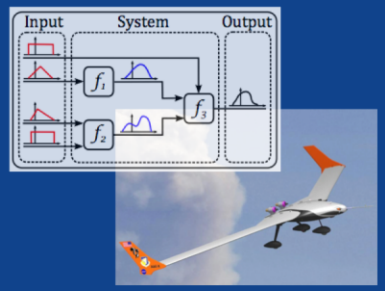


MURI-14:

A unified mathematical and algorithmic framework for managing multiple information sources of multi-physics systems

MURI: Managing multiple information sources of multi-physics systems

AFOSR Computational Mathematics Program Review
August 3, 2015

Cornell University

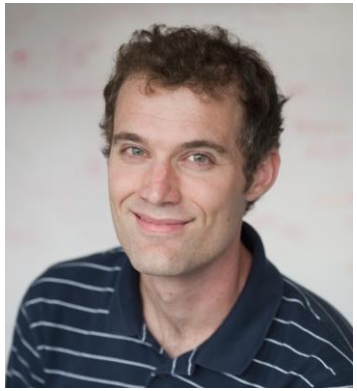
SANTA FE INSTITUTE

Outline

- MURI project overview
- Multi-information source example:
  Estimating QoI statistics
- Multi-information source optimization (Frazier)
- Reduced-order modeling (Mignolet)
- Multi-physics coupling (Martins)

MURI PIs

Douglas Allaire
Texas A&M U.



Peter Frazier
Cornell U.

Joaquim Martins
U. Michigan

Youssef Marzouk
MIT

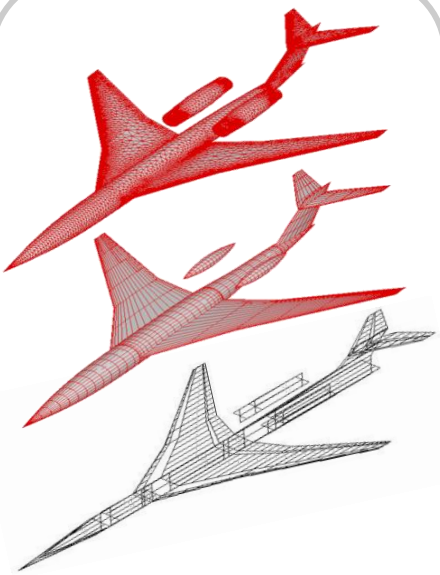
Marc Mignolet
Arizona State U.

Karen Willcox
MIT (lead PI)

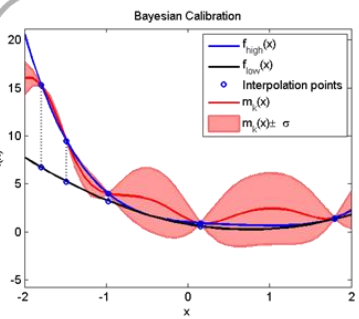
David Wolpert
SFI/MIT

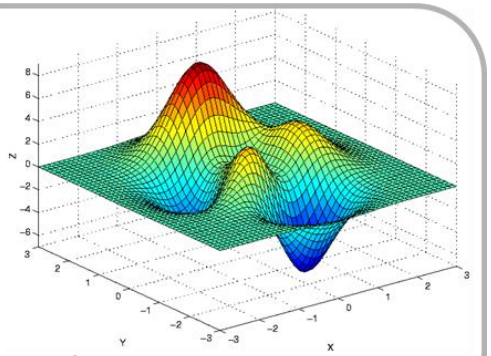
Project motivation

- Analysis and design typically begin with low-fidelity models and progressively incorporate higher fidelity tools
- Many information sources available: multifidelity models, historical data, operational data, experimental data, expert opinions
  - Telling us different things about the system: the collective information they provide is greater than the individual parts



Physics-based models





Data-fit/surrogate/reduced models

Source: wikipedia.org

Experts

Source: www.directindustry.com

Experimental data

Challenges and research gap

- Existing multifidelity optimization approaches calibrate low-fidelity models or replace low-fidelity results with higher fidelity data
- Major challenges not addressed by current multifidelity approaches:
  - Managing a broad range of information sources
  - Certifying overall analysis and design results
  - Adapting to new decision goals
- Further challenges for **multi-physics systems**, where different information sources relate to different parts of the system and where there are choices for how to couple (or decouple) the various parts

Research objectives

1. Develop statistical approaches for defining and quantifying fidelity
2. Establish decision-theoretic methods for optimally managing sources of uncertain multi-physics information
3. Create reduced-order models with goal-driven adaptation to multi-physics interactions and with quantified uncertainty
4. Formulate an information-theoretic approach for handling multi-physics coupling
5. Create a scalable framework for solving multi-physics analysis and design problems under uncertainty

Research thrusts

1. Develop statistical approaches for defining and quantifying fidelity
2. Establish decision-theoretic methods for optimally managing sources of uncertain multi-physics information

RT1: Optimal information-source management

3. Create reduced-order models with goal-driven adaptation to multi-physics interactions and with quantified uncertainty

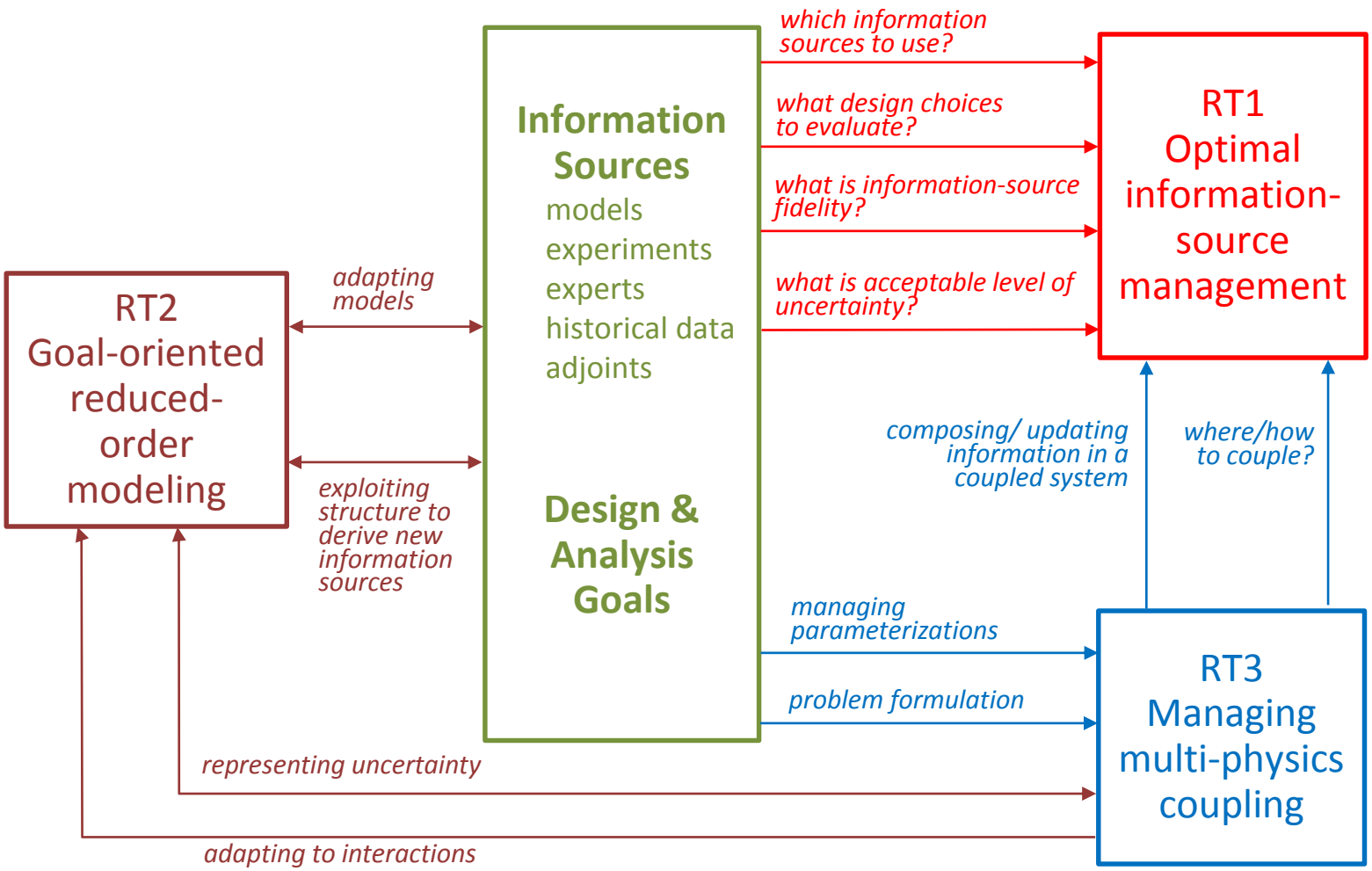
RT2: Goal-oriented ROMs for the multi-source multi-physics setting

4. Formulate an information-theoretic approach for handling multi-physics coupling
5. Create a scalable framework for solving multi-physics analysis and design problems under uncertainty

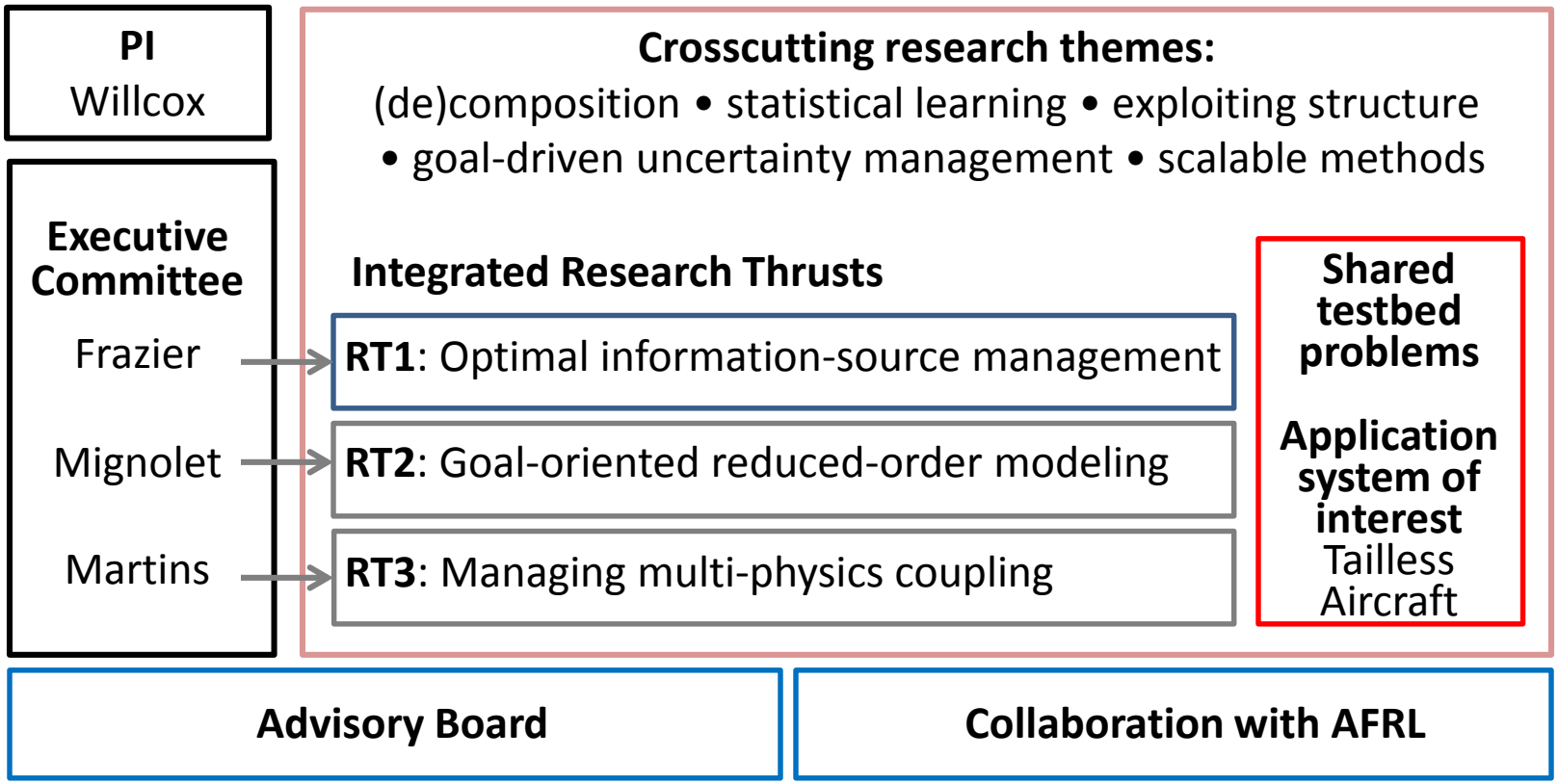
RT3: Managing coupling in multi-physics systems

Research thrust integration

A scalable multi-source framework for certifiable solution of multi-physics analysis and design problems under uncertainty



Overall project structure



MURI PIs: Allaire, Frazier, Martins, Marzouk, Mignolet, Willcox, Wolpert

Advisory Board: P. Beran (AFRL, Chair), R. Canfield (VA Tech), R. Liebeck (Boeing), S. Mahadevan (Vanderbilt) J. Scarcello (Lockheed), B. Smarslok (AFRL), B. Stanford (NASA)

Multi-IS example: Estimating QoI statistics

B. Peherstorfer
postdoc MIT

- input $z \in D$
- quantity of interest (QoI) $y \in Y$
- high-fidelity model $f^{(1)}: D \to Y$ with cost $w_1 > 0$ ("truth")

- Goal: given random input variable *Z*, estimate statistics *s* of $f^{(1)}(Z)$, e.g., mean:

$$s = E\left[f^{(1)}(Z)\right]$$

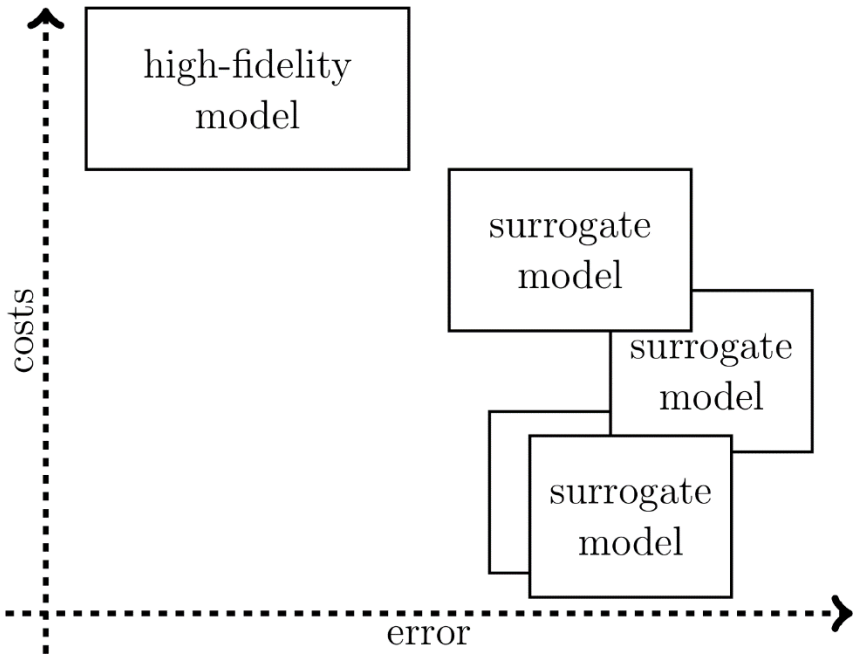
- Monte Carlo estimator for s using n realizations $z_1, \dots, z_n$ of *Z:*

$$\bar{y}_n^{(1)} = \frac{1}{n}\sum_{i=1}^{n} f^{(1)}(z_i)$$

has costs nw_1

Surrogate models as information sources

- Often many surrogates of different type available, with different costs
  - Interpolation, regression models
  - Projection-based reduced models
  - Simplified-physics models
  - Coarse-grid approximations
  - Machine-learning models (SVMs)



- Models do not necessarily form a hierarchy (cf. MLMC)

- View each model as an information source (IS), where the collective information is greater than the information provided by a single model
  - How to combine these models?
  - How to balance model evaluations among them?

Multifidelity Monte Carlo (MFMC)

- Our previous work considered two models, combined in a control variate framework *(Ng & Willcox 2012, 2014)*

Here:

- high-fidelity IS $f^{(1)}: D \rightarrow Y$ ("truth")
- *k – 1* surrogate IS's $f^{(2)}, \ldots, f^{(k)}: D \rightarrow Y$
- IS $f^{(i)}$ has cost w_i
- m_i evaluations for IS i, with

$$m_1 \leq m_2 \leq \ldots \leq m_k$$

Multifidelity Monte Carlo (MFMC)

- Draw m_k realizations $z_1, \dots, z_{m_k}$ of Z and evaluate $f^{(i)}$

$$f^{(i)}(z_1), \dots, f^{(i)}(z_{m_i})$$

- MFMC estimator uses the MC estimates $\bar{y}_{m_1}^{(1)}, \dots, \bar{y}_{m_k}^{(k)}$ and $\bar{y}_{m_1}^{(2)}, \dots, \bar{y}_{m_{k-1}}^{(k)}$ as control variates

$$\hat{s} = \bar{y}_{m_1}^{(1)} + \sum_{i=2}^{k} \alpha_i \left(\bar{y}_{m_i}^{(i)} - \bar{y}_{m_{i-1}}^{(i)} \right)$$

- MFMC estimator is unbiased, even with no error bounds for surrogates

Multifidelity Monte Carlo (MFMC)

- MFMC estimator

$$\hat{s} = \bar{y}_{m_1}^{(1)} + \sum_{i=2}^{k} \alpha_i \left(\bar{y}_{m_i}^{(i)} - \bar{y}_{m_{i-1}}^{(i)} \right)$$

- The costs of the MFMC estimator are

$$c(\hat{s}) = \sum_{i=1}^{k} w_i m_i$$

- Distinguishing feature of MFMC method is the optimal selection of the number of IS evaluations $m_1 \leq m_2 \leq \ldots \leq m_k$ and of coefficients $\alpha_2, \ldots, \alpha_k$ that is applicable to general IS's (e.g., any type of surrogate model, database curve fits, etc.)

Balancing the number of IS evaluations

- Minimize the MSE of the MFMC estimator for a given computational budget p

- MFMC estimate $\hat{s}$ is unbiased; MSE is given by $\mathrm{Var}[\hat{s}]$

$$\mathrm{Var}[\hat{s}] = \frac{\sigma_1^2}{m_1} + \sum_{i=2}^{k} \left(\frac{1}{m_{i-1}} - \frac{1}{m_i} \right) \left(\alpha_i^2 \sigma_i^2 - 2\alpha_i \rho_i \sigma_i \sigma_1 \right)$$

  - σ_i^2 is variance of $f^{(i)}(Z)$
  - ρ_i is correlation coefficient between $f^{(1)}(Z)$ and $f^{(i)}(Z)$

- Leads to optimization problem

$$\min_{m \in \mathbb{R}^k, \alpha_2, \ldots, \alpha_k \in \mathbb{R}} \quad \mathrm{Var}[\hat{s}]$$

$$\text{such that} \quad m_{i-1} \leq m_i, i = 2, \ldots, k$$

$$0 \leq m_1$$

$$c(\hat{s}) = w^T m = p$$

Balancing the number of IS evaluations

- Optimization problem has unique (analytic) solution if

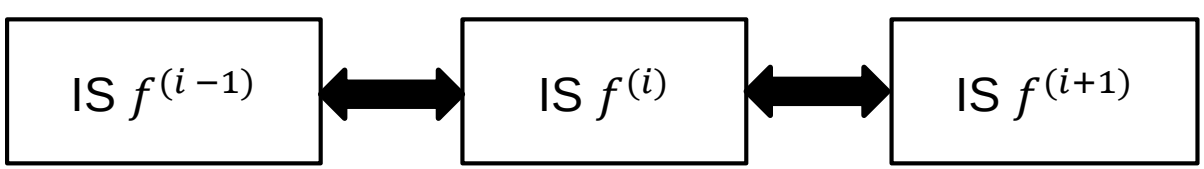
$$\rho_1^2 > \rho_2^2 > \ \dots > \rho_k^2$$

and

$$\frac{w_{i-1}}{w_i} > \frac{\rho_{i-1}^2 - \rho_i^2}{\rho_i^2 - \rho_{i+1}^2}, \qquad i = 2, \dots, k$$

- The costs/correlation ratio establishes a relationship between
  - preceding IS $f^{(i-1)}$,
  - current IS $f^{(i)}$, and
  - succeeding IS $f^{(i+1)}$

- If $f^{(1)}, \dots, f^{(k)}$ violate conditions, can construct a subset of IS's that satisfy conditions



The interactions between the IS's impact the behavior of the MFMC estimator; not the properties of the IS's alone

MFMC variance reduction

- Let $\bar{y}_n^{(1)}$ be the (benchmark) Monte Carlo estimator with computational budget p
- Ratio of MSE of MFMC estimator $\hat{s}$ and MSE of $\bar{y}_n^{(1)}$ is

$$\frac{e(\hat{s})}{e(\bar{y}_n^{(1)})} = \left(\sum_{i=1}^{k} \sqrt{\frac{w_i}{w_1}\left(\rho_i^2 - \rho_{i+1}^2\right)} \right)^2$$

- The MFMC estimator has lower MSE than the Monte Carlo estimator if

$$\sum_{i=1}^{k} \sqrt{\frac{w_i}{w_1}\left(\rho_i^2 - \rho_{i+1}^2\right)} < 1$$

- Condition on the collective whole of the models (sum), not on properties of each model separately
- The interaction between the models is what drives the MFMC estimator, not model properties alone

Example: three IS's ($k = 3$)

- truth model $f^{(1)}$, cost w_1
- surrogate model $f^{(2)}$, cost w_2, correlation with truth ρ_2
- surrogate model $f^{(3)}$, cost w_3, correlation with truth ρ_3

- feasibility conditions

$$1 > \rho_2^2 > \rho_3^2, \qquad \frac{w_1}{w_2} > \frac{\rho_1^2 - \rho_2^2}{\rho_2^2 - \rho_3^2}, \qquad \frac{w_2}{w_3} > \frac{\rho_2^2 - \rho_3^2}{\rho_3^2}$$

- variance reduction (low S $\longrightarrow$ low variance)

$$\mathrm{S}(w_1, w_2, w_3, \rho_2, \rho_3) = \sum_{i=1}^{k} \sqrt{\frac{w_i}{w_1}\left(\rho_i^2 - \rho_{i+1}^2\right)}$$

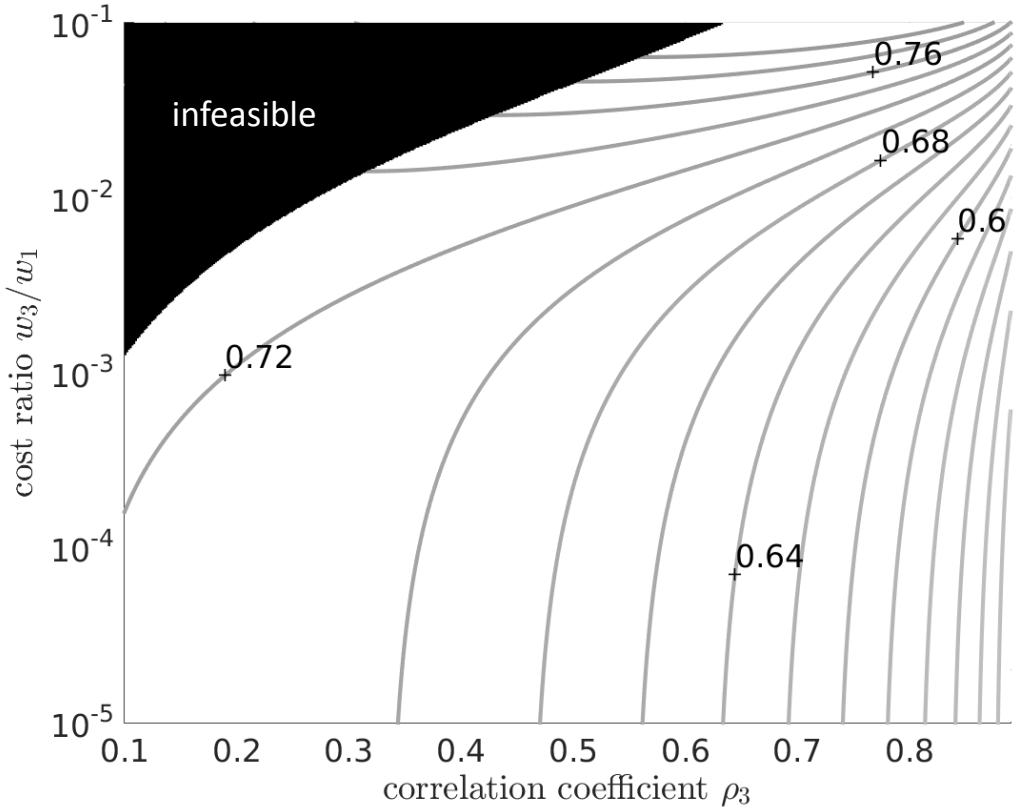
$$= \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}\left(\rho_2^2 - \rho_3^2\right)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

< 1 for the MFMC estimator to be efficient

Example: three IS's ($k = 3$)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

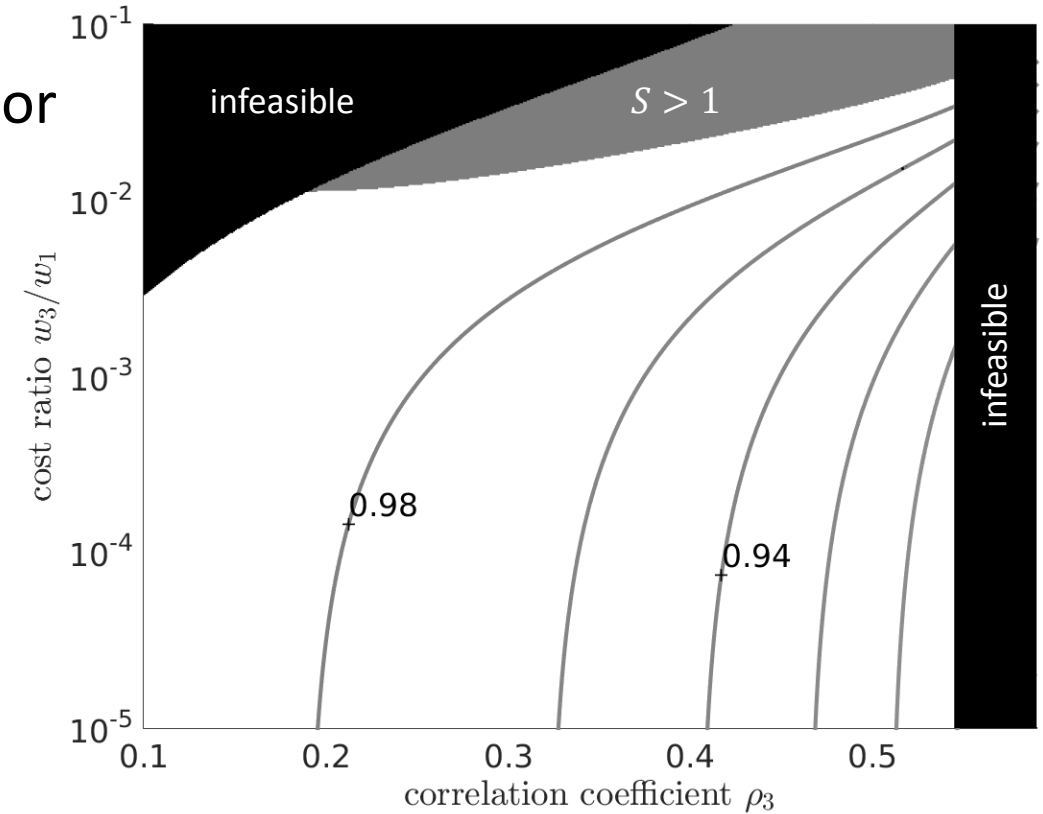
- Set $w_2/w_1 = 0.1, \quad \rho_2 = 0.9$
- Using $f^{(1)}, f^{(2)}$ only $\longrightarrow$ larger variance than MC estimator (S > 1)
- Vary w_3/w_1 and ρ_3; plot contours of S
- If costs w_3 high (> $0.01w_1$): third term in S can dominate, increasing correlation ρ_3 can lead to larger S
- If w_3 low, then increase of ρ_3 always reduces S



Example: three IS's ($k = 3$)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

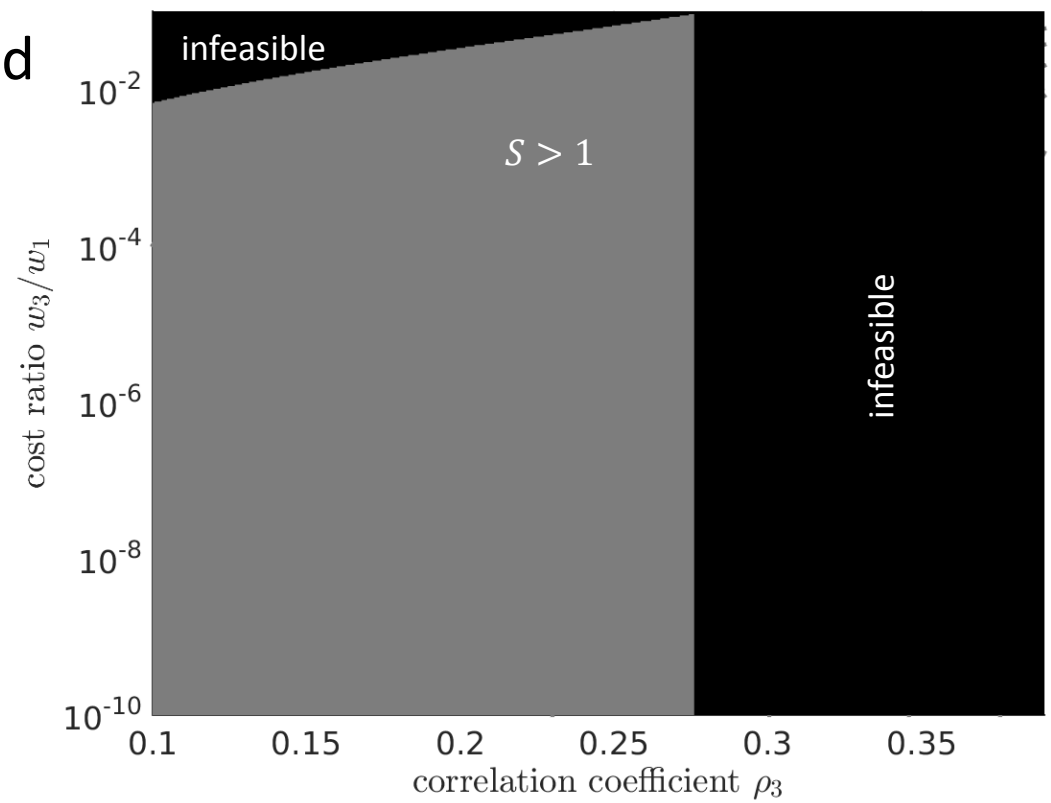
- Set $w_2/w_1 = 0.1$, $\rho_2 = 0.6$
- Increasing the correlation can violate feasibility condition: if $\rho_3 \approx \rho_2$ then denominator $\rho_2^2 - \rho_3^2$ in feasibility condition becomes small and condition is violated



Example: three IS's ($k = 3$)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

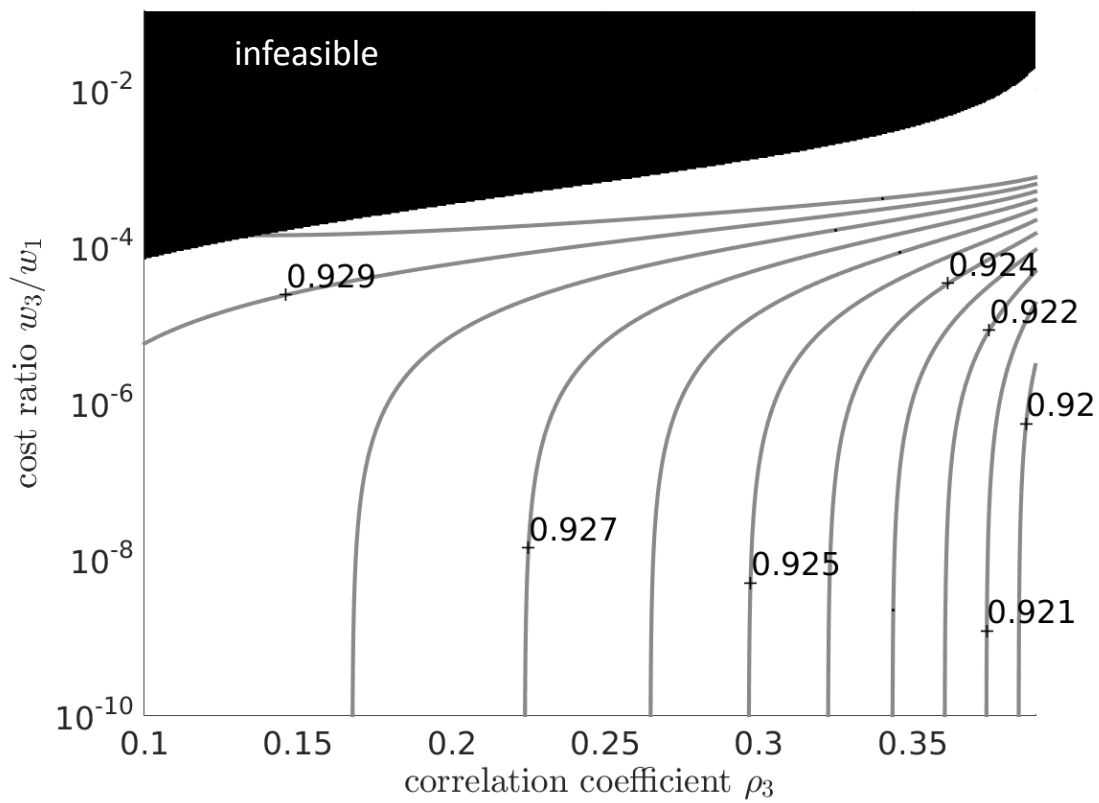
- Set $w_2/w_1 = 0.1$, $\rho_2 = 0.4$
- IS $f^{(2)}$ has high costs and low correlation
- Variance cannot be improved by adding a third IS
- Any third IS will lead either to a violation of the feasibility condition or to a higher variance than the MC estimator ($S > 1$)
- Have to remove/change IS $f^{(2)}$ to reduce variance



Example: three IS's ($k = 3$)

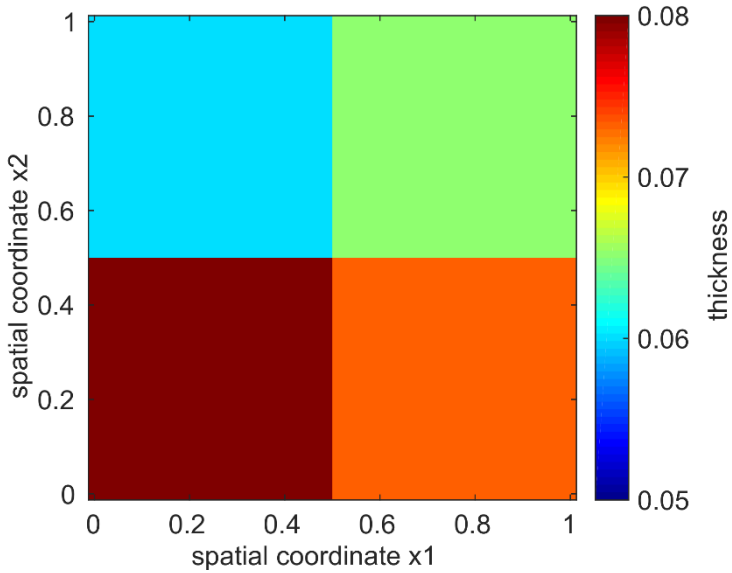
$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

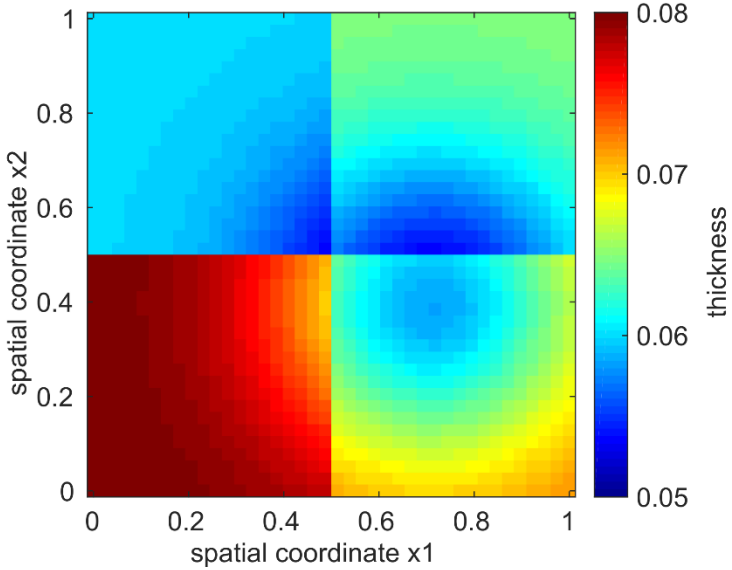
- Set $w_2/w_1 = 10^{-4}$, $\rho_2 = 0.4$
- Decreasing the costs of $f^{(2)}$ releases deadlock; adding a third model can improve the variance



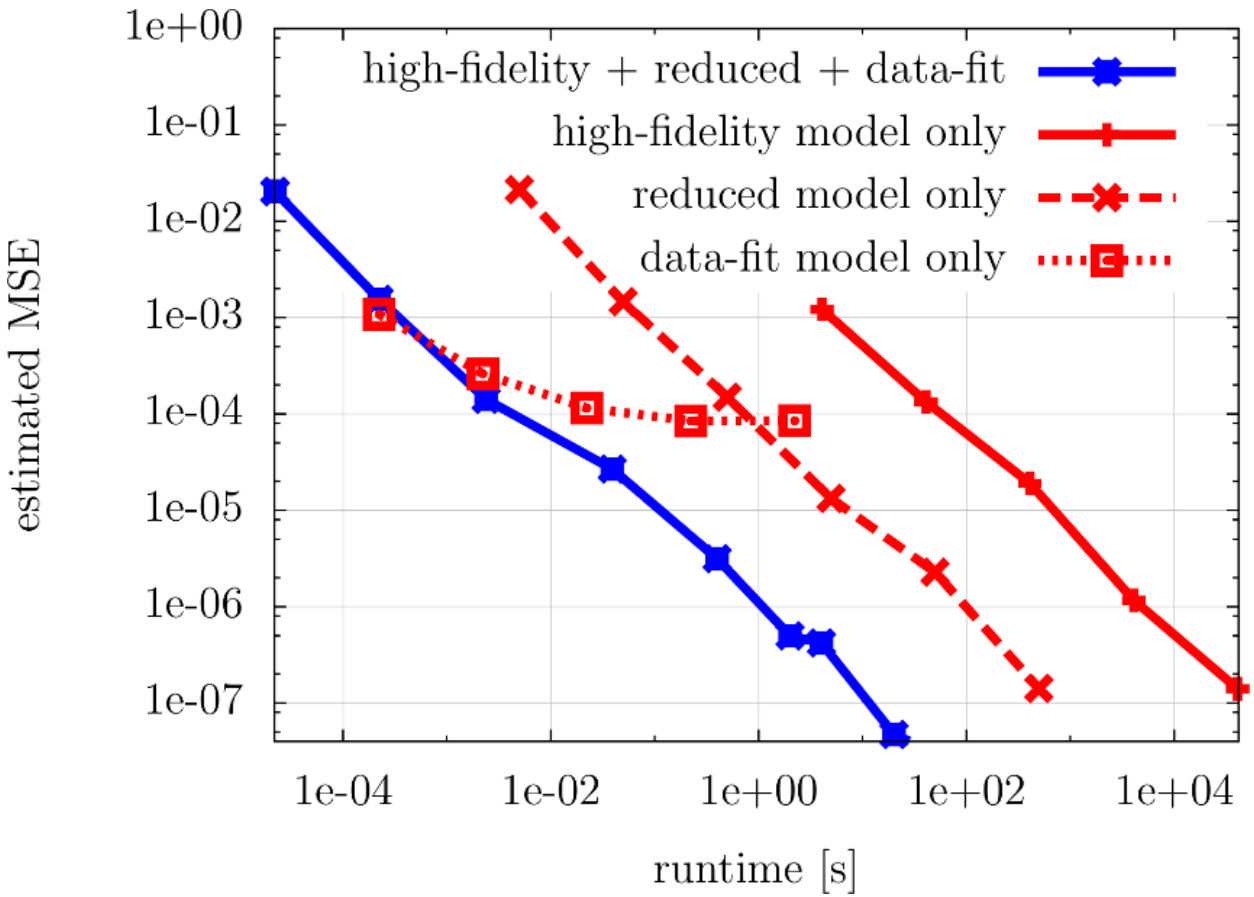
Locally damaged plate

- Locally damaged plate
- Inputs: thickness, load, two damage parameters
- Inputs uniformly distributed in [0.05, 0.1] x [1, 100] x [0, 0.2] x (0, 0.05]
- QoI: maximum deflection of plate
- Six models available
  - High-fidelity model: FEM, 300 dof
  - Reduced model: POD, 10 dof
  - Reduced model: POD, 2 dof
  - Reduced model: POD, 5 dof
  - Data-fit model: linear interpolation, 256 pts
  - Support vector machine: 256 pts
- Variance, correlation, runtime estimated from 100 samples



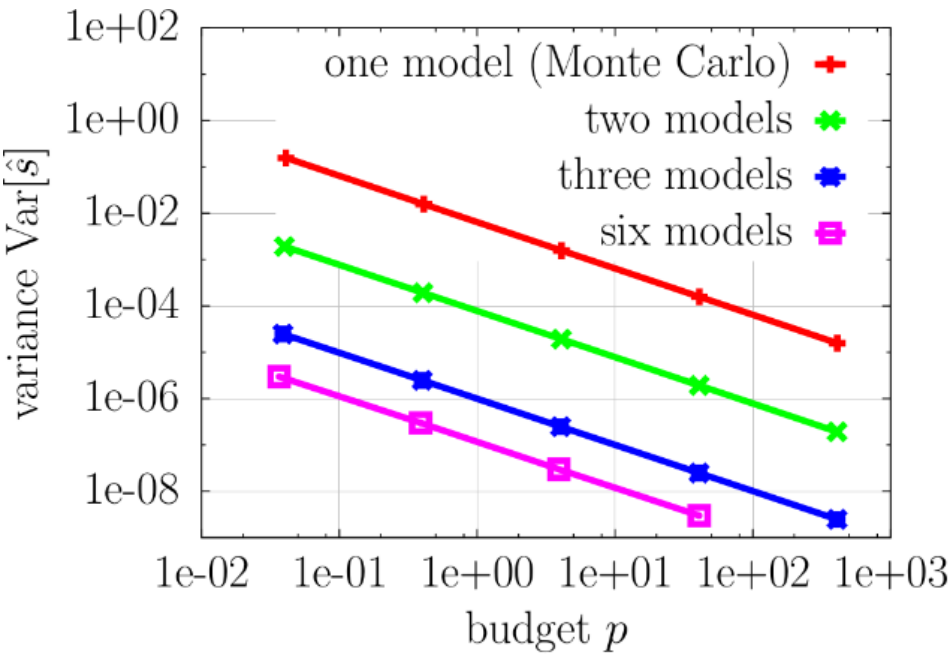


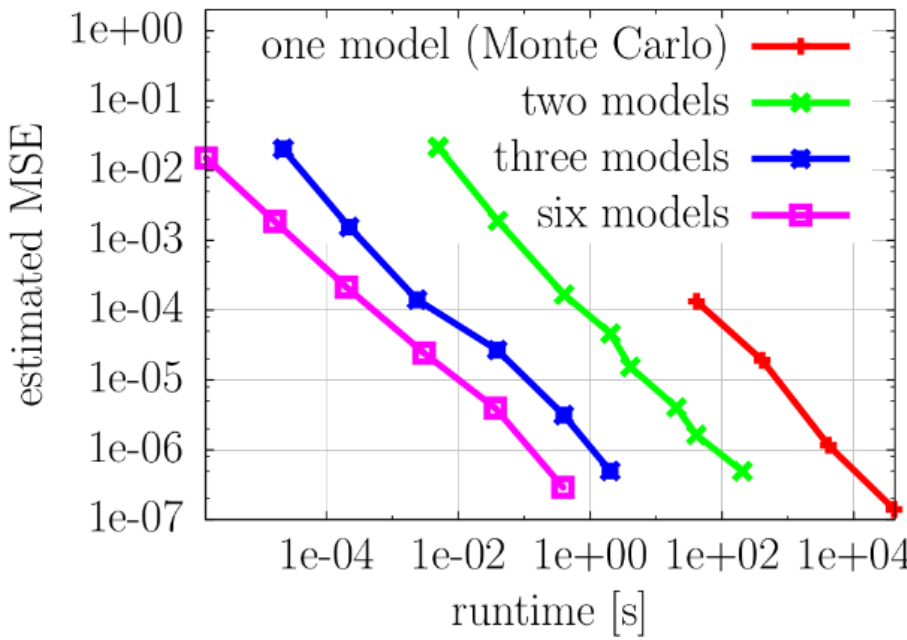
Locally damaged plate



- Combine high-fidelity + reduced (POD, 10) + data-fit (linear interp, 256)
- MFMC achieves almost 4 orders of magnitude improvement over MC with high-fidelity, and two orders over MC with reduced model
- Reduced and data-fit model lead to biased estimator, MFMC is unbiased

Locally damaged plate





- Successively add reduced (POD, 10), data-fit (linear interp, 256), and then all others
- Adding data-fit model reduces variance, even though data-fit model is poor approximation of high-fidelity model (left)
- Adding additional reduced and SVM models reduces variance only slightly
- Predicted variance reduction (left) fits well to numerically obtained MSE (right)

Multi-IS example: Conclusions

- MFMC combines general IS's (e.g., surrogate models of any type) to accelerate the estimation of statistics
- Optimal allocation of number of evaluations across IS's
- MFMC estimator is unbiased
- Properties of the collective whole of the IS's drive the efficiency of the MFMC estimator
- Conditions on IS costs vs. correlations could be used constructively (i.e., to inform reduced model creation, perhaps to inform "useful" experiments)
- Can we extend this kind of analysis to inform IS selection for design decisions?

Doug Allaire, Peter Frazier, David Wolpert

RT1: OPTIMAL INFORMATION-SOURCE MANAGEMENT

Research Thrust 1's objectives

- Develop statistical methods for **predicting** system performance using information from multiple sources.
  - Example: Using results from a computational fluid dynamics simulation at 2 different mesh sizes, predict performance at a new point.
- Develop methods for **deciding** how to optimally allocate computational / experimental effort across information sources, to achieve a desired goal.
  - Example: Decide at which mesh size and system design to run the simulation next, to best support finding the best design.

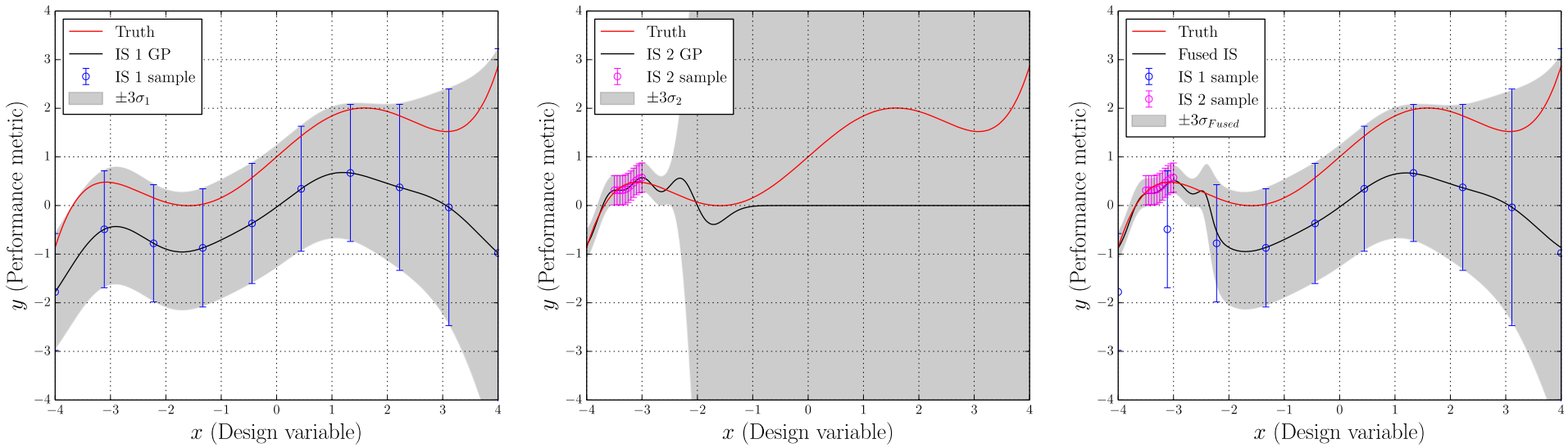
We are working on these MISO problems: "traditional" multi-fidelity optimization

- In multi-fidelity optimization, we have a range of computational models, with increasing accuracy and increasing cost.

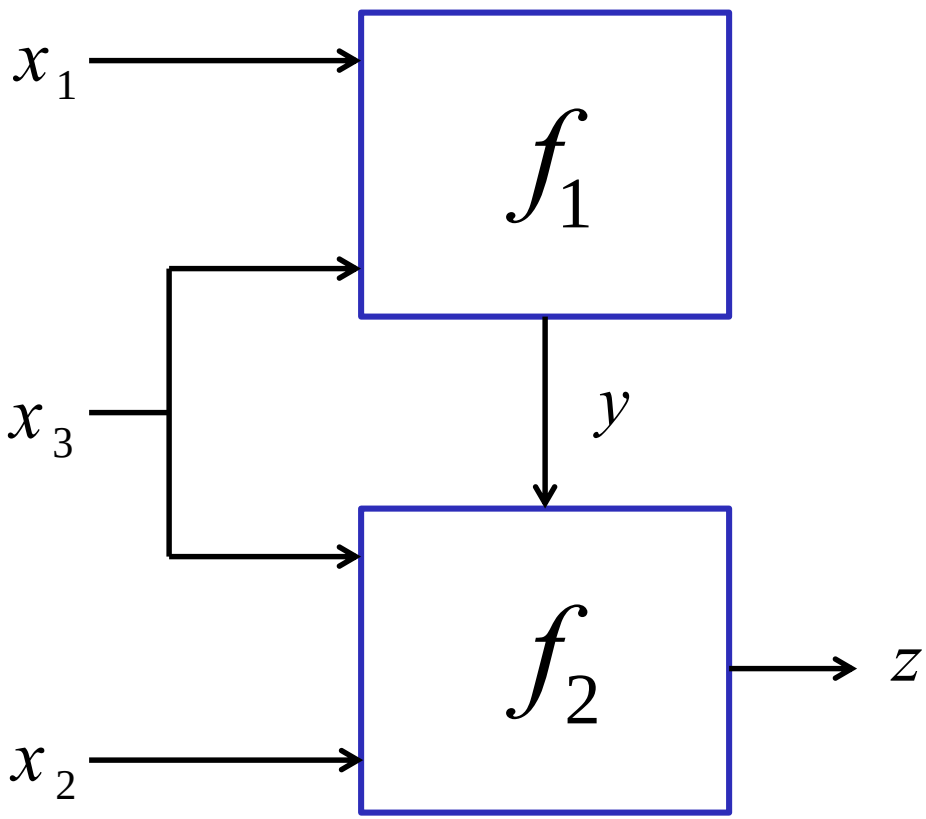
- Examples:
  - a PDE,
    where we can choose the fineness of the mesh.
  - a steady state simulation,
    where we can choose how long to run it.
  - evaluation of a system on historical data,
    where we can choose how much data to use.

We are working on these MISO problems: ”non-traditional” multi-fidelity optimization

- Information sources do not have to be ordered by accuracy.
- We are working on a problem with two information sources:
  - Information source 1 is noise-free, but biased.
  - Information source 2 gives unbiased observations of the objective with stochastic noise.



We are working on these MISO problems: two information sources with coupling



- Our design variable is $[x_1, x_2, x_3]$
- Information source 1 takes input x_1, x_3, and outputs y.
- Information source 2 takes inputs x_2, x_3 and outputs our objective z.
- Information source 1 could be an aerodynamic simulation of a wing
- Information source 2 could be a structural simulation of a wing, that includes aerodynamic forces.

Here's how we solve this multi-information source optimization problem

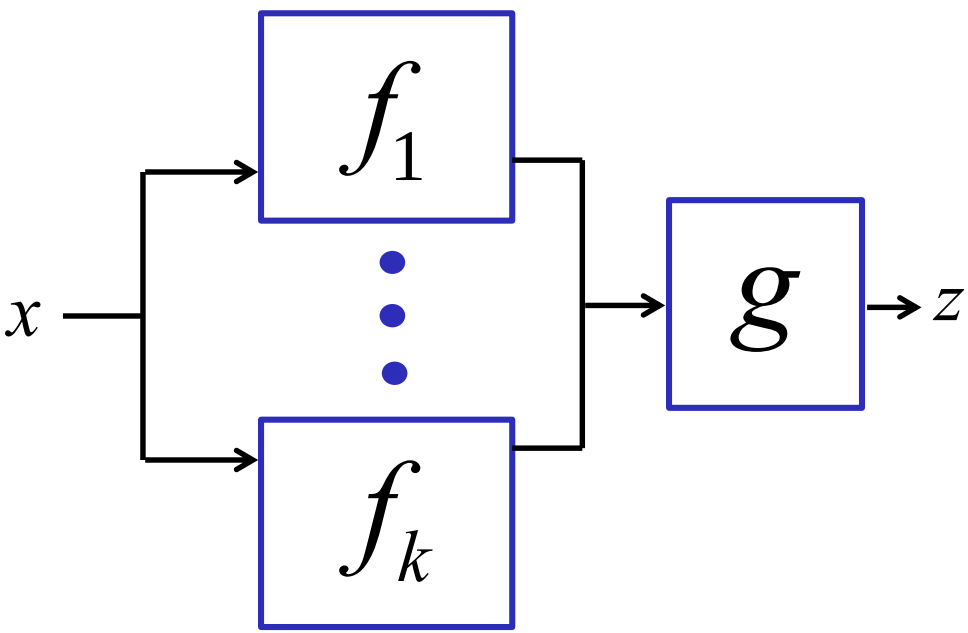
1. Build a statistical model of each information source, and the relationship between them, based on some initial data.

2. Until we decide to stop:
   - Use the statistical model and the end-goal to decide at which design to evaluate, and which information source to use.

     We use value of information analysis.
   - Update the statistical model based on the new information.

3. Choose a final solution based on all the data collected.

We are working on these MISO problems: information sources with combined outputs



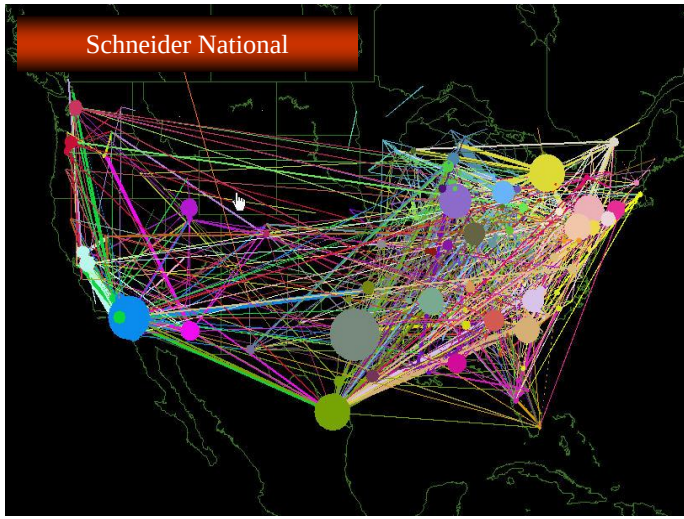
- The outputs from the k information sources are combined by a common function g.
- Information sources are uncoupled.
- Information sources take common input x.
- The function g is known. Our algorithm and examples suppose it is linear.
- We want to maximize $g(x) := g(f_1(x),\ldots,f_k(x))$.

Application: Simulation Calibration

- Our simulator takes parameter vector x.
- We want to calibrate x to this observed data:
  - observed environmental condition e_s
  - real-world output y_s
- The simulator's prediction of y_s is $SIM(x, e_s)$
- Information source s is the squared error under environmental condition e_s

$$f_s(x) = (SIM(x, e_s) - y_s)^2$$

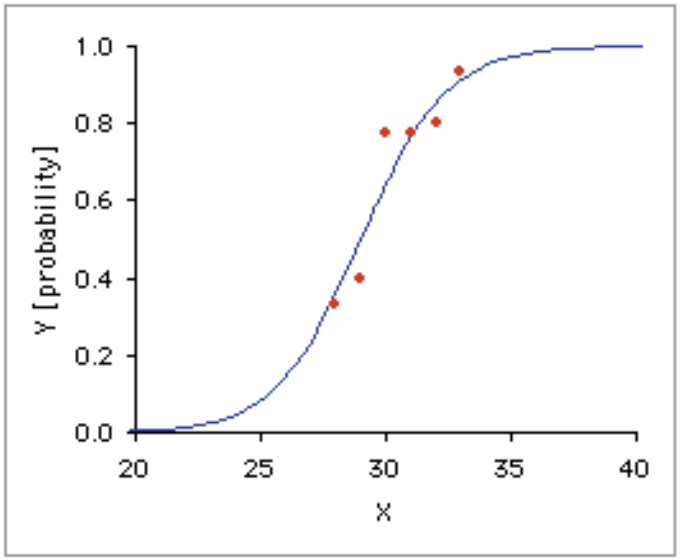
- We want to choose x to minimize $g(x) = f_1(x) + \ldots + f_k(x)$



Application: Training Machine Learning Models

- We have a machine learning model, e.g., logistic regression.
- It takes a vector of parameters x.
- We have many terabytes of training data ($INPUT_j$, $OUTPUT_j$).
- We want to choose x to maximize the log-likelihood $\Sigma_j \log p(OUTPUT_j \mid INPUT_j, x)$.
- Our data is spread across k disks.
  Information source s
  is the log-likelihood of the data on disk s.

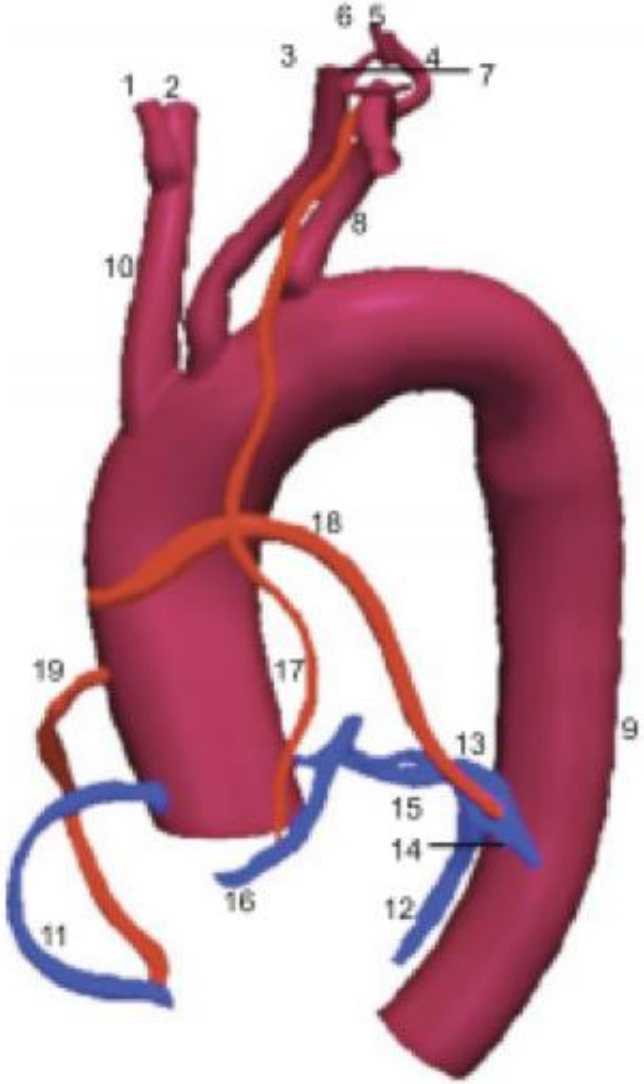
  $f_s(x) = \Sigma_{j \text{ is on disk } s} \log p(OUTPUT_j \mid INPUT_j, x)$.
- We want to choose x to maximize
  $g(x) = f_1(x) + \ldots + f_k(x)$



Application: Simulation Optimization

- We have a stochastic simulator $f(x,\omega)$
  - x is something we control
  - ω is some external source of randomness, with a known probability distribution p.
- We want to solve $\min_x E[f(x,\omega)]$

- Suppose ω takes finitely many values, $\omega_1,\ldots,\omega_k$. [Later, we relax this.]
- Let $f_s(x) := f(x,\omega_s)$
- Let $g(x) = g(f_1(x),\ldots,f_k(x)) = \Sigma_s\, p(\omega_s)\, f_s(x)$
- Our goal is $\min_x g(x)$.



Here's how we solve this multi-information source optimization problem

1. Build a statistical model of each information source, and the relationship between them, based on some initial data.

2. Until we decide to stop:
   - Use the statistical model and the end-goal to decide at which design to evaluate, and which information source to use.

     We use value of information analysis.
   - Update the statistical model based on the new information.

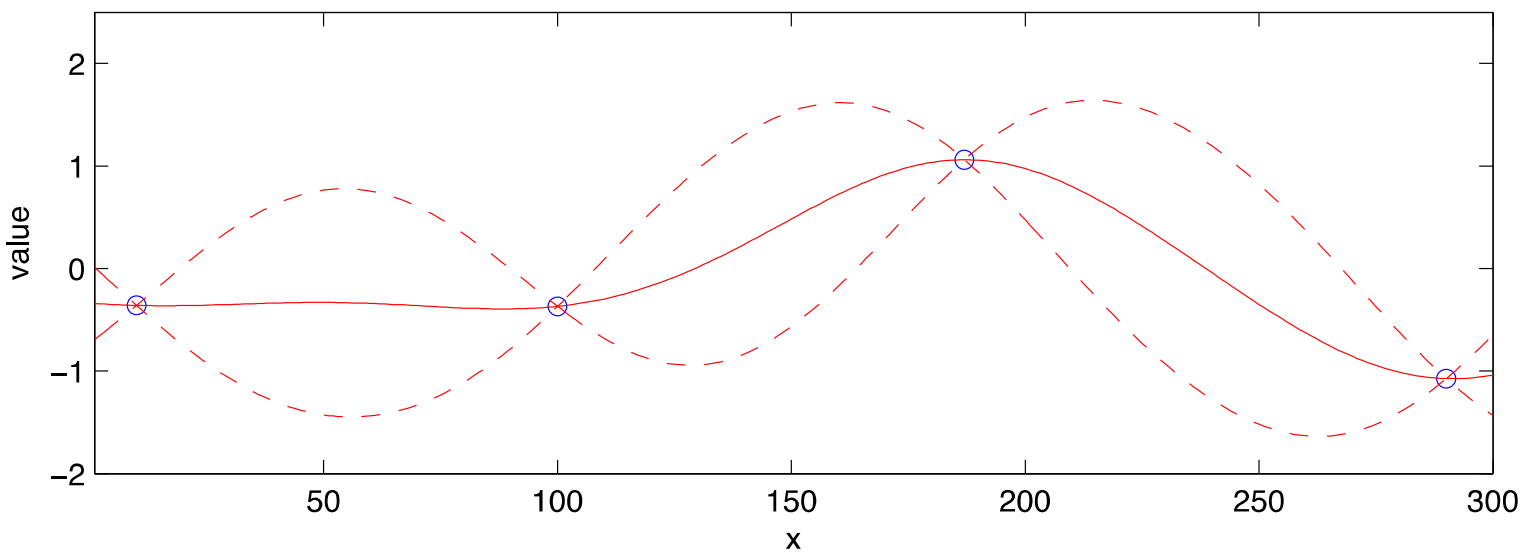
3. Choose a final solution based on all the data collected.

Here's our statistical model

- We model each information source as a Gaussian process

$$f_s(\bullet) \sim GP(\mu_s(\bullet), \Sigma_s(\bullet,\bullet)).$$

- If we observe $f_s(x)$ for some s and x, then our posterior on each $f_s(\bullet)$ will still be a Gaussian process, $GP(\mu_{ns}(\bullet), \Sigma_{ns}(\bullet,\bullet))$.



- Our Gaussian process priors can be independent across s. Or, if we wish to model related information sources, they can be correlated [more later].

Here's our value of information analysis (Part 1)

- Suppose we've seen n points, at information sources of our choosing, $y_j = f_{s(j)}(x_j)$, j=1,..,n.

- Under our posterior,
  $E_n[f_s(x)] = \mu_{ns}(x)$.
  $m_n(x) = E_n[g(x)] = \Sigma_s \mu_{ns}(x)$.

- If we were to stop now,
  we would choose $\text{argmax}_x\, m_n(x)$,
  with value $m_n^* = \max_x m_n(x)$.

Here's our value of information analysis (Part 2)

- Suppose instead we took one additional sample, from information source s_{n+1} at point x_{n+1}, observing y_{n+1}.

- Our posterior on that information source would change (the posterior on the others would change if we used a correlated prior)
  $E_{n+1}[f_s(x)] = \mu_{n+1,s}(x)$.
  $m_{n+1}(x) = E_{n+1}[g(x)] = \Sigma_s \mu_{n+1,s}(x)$.

- If we then stopped sampling, we would choose $\text{argmax}_x\, m_{n+1}(x)$, with value $m_{n+1}^* = \max_x m_{n+1}(x)$.

Here's our value of information analysis (Part 3)

- After n samples, our solution value in $m_n^* = \max_x m_n(x)$.
- After n+1 samples, our solution value in $m_{n+1}^* = \max_x m_{n+1}(x)$.
- The difference, $m_{n+1}^* - m_n^*$ is the improvement in solution quality due to sample n+1.

Here's our value of information analysis (Part 3)

- After n samples, our solution value in $m_n^* = \max_x m_n(x)$.
- After n+1 samples, our solution value in $m_{n+1}^* = \max_x m_{n+1}(x)$.
- The difference, $m_{n+1}^* - m_n^*$ is the improvement in solution quality due to sample n+1.

- When we choose s_{n+1} and x_{n+1}, we don't know y_{n+1} or m_{n+1}^*, but we know the **distribution** of y_{$_{n+1}$}.
- We compute the value of information for sampling x with information source s:

$$VOI_n(x,s) = E_n[m_{n+1}^* - m_n^* \mid x_{n+1}=x, s_{n+1}=s].$$

- We sample the point and information source with the largest value of information.

Here's our algorithm

1. Build a Gaussian process prior on the information sources using some initial data

2. Until we decide to stop:
   - Solve $\text{argmax}_{x,s}\, \text{VOI}_n(x,s)$.
   - Sample at x, with information source s.
   - Update the statistical model based on the new information.

3. Choose a final solution based on all the data collected.

We can do simulation optimization over continuous random variables too

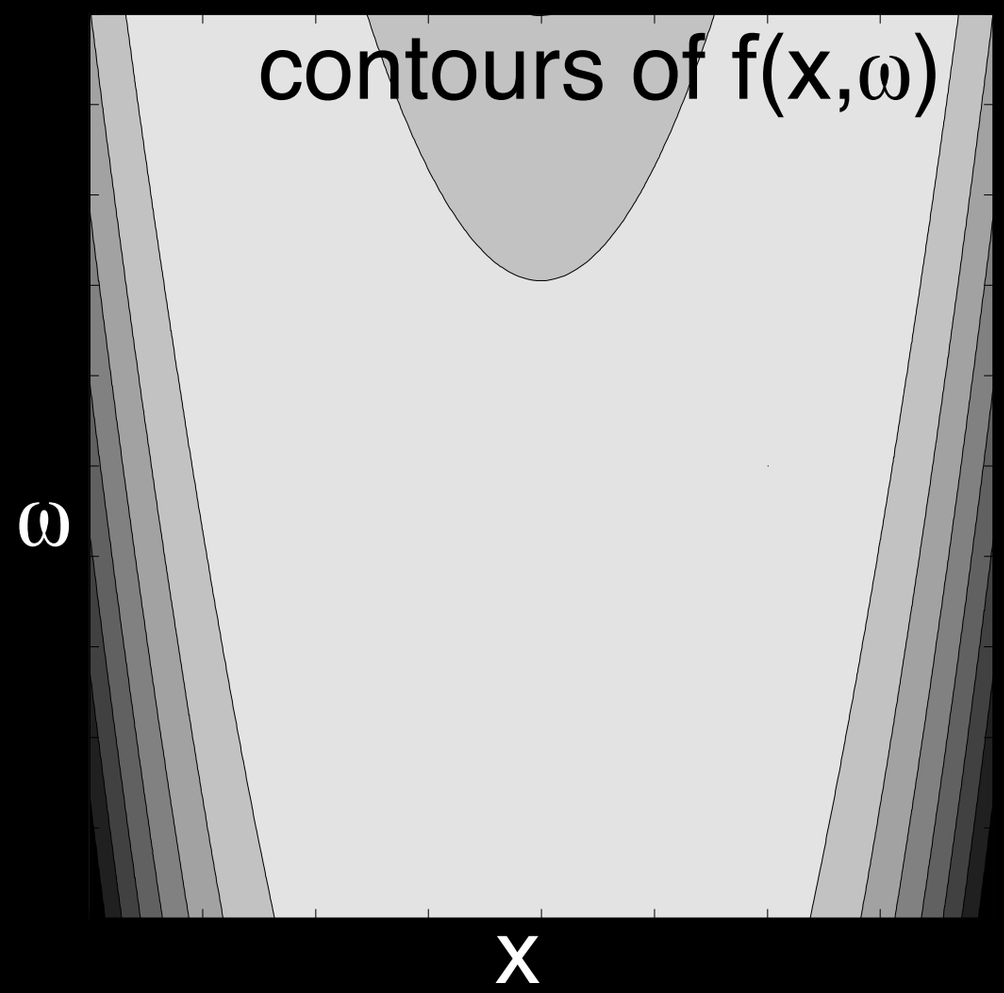
- The sth information source is $f_s(x) = f(x, \omega_s)$, the output of an expensive simulator with random noise ω_s.
  - $g(x) = \Sigma_{s=1,\ldots,k}\, p(\omega_s)\, f(x, \omega_s)$
  - This is not that different from:

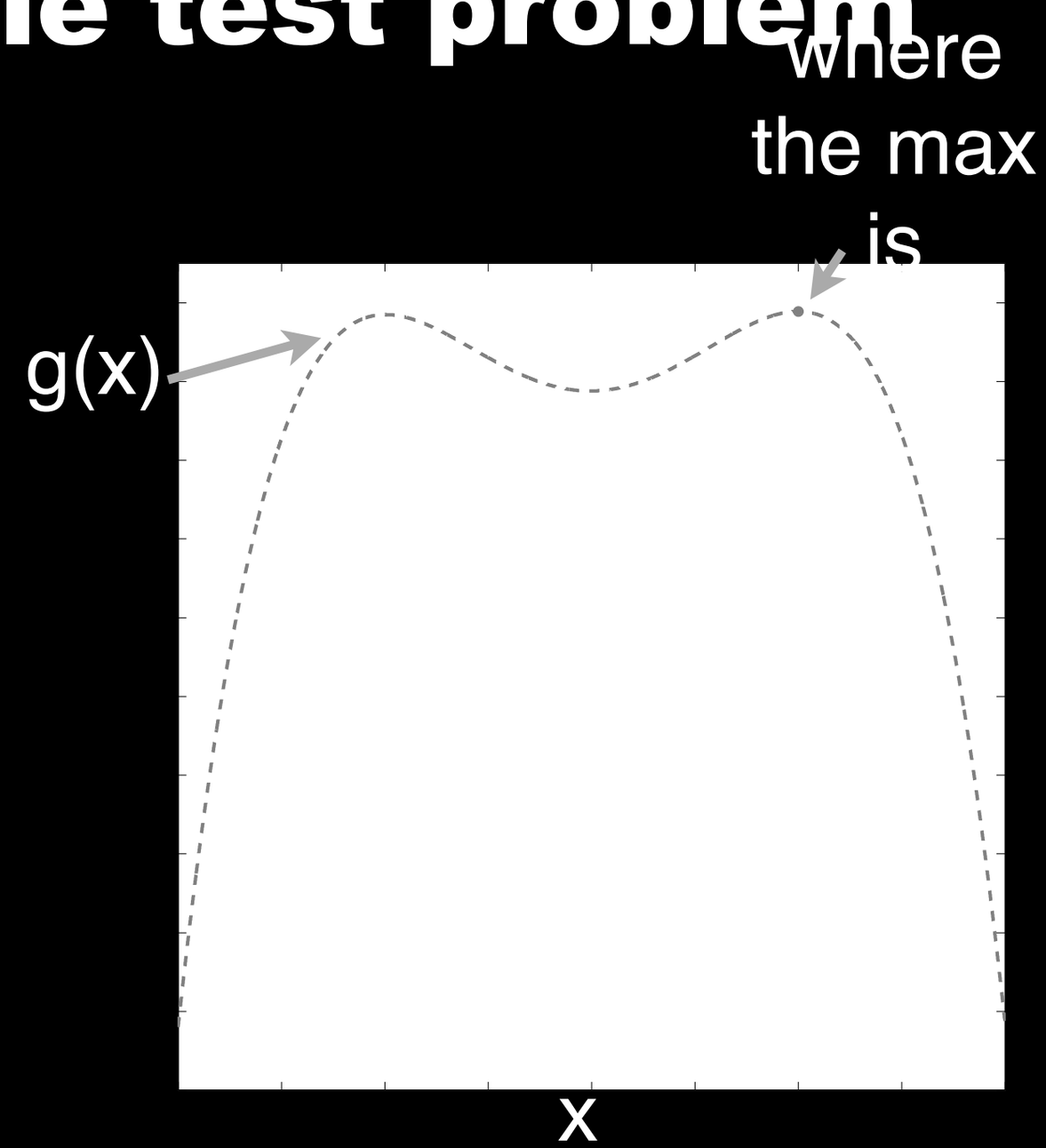
$$g(x) = \int p(\omega) f(x, \omega)\, d\omega$$

s

- To model the correlation between the information sources we put a Gaussian process on f, as a function of both x and ω.
- Then f(•, ω) is also a GP for each ω, with correlation across ω.

Let's show how this algorithm solves a simple test problem



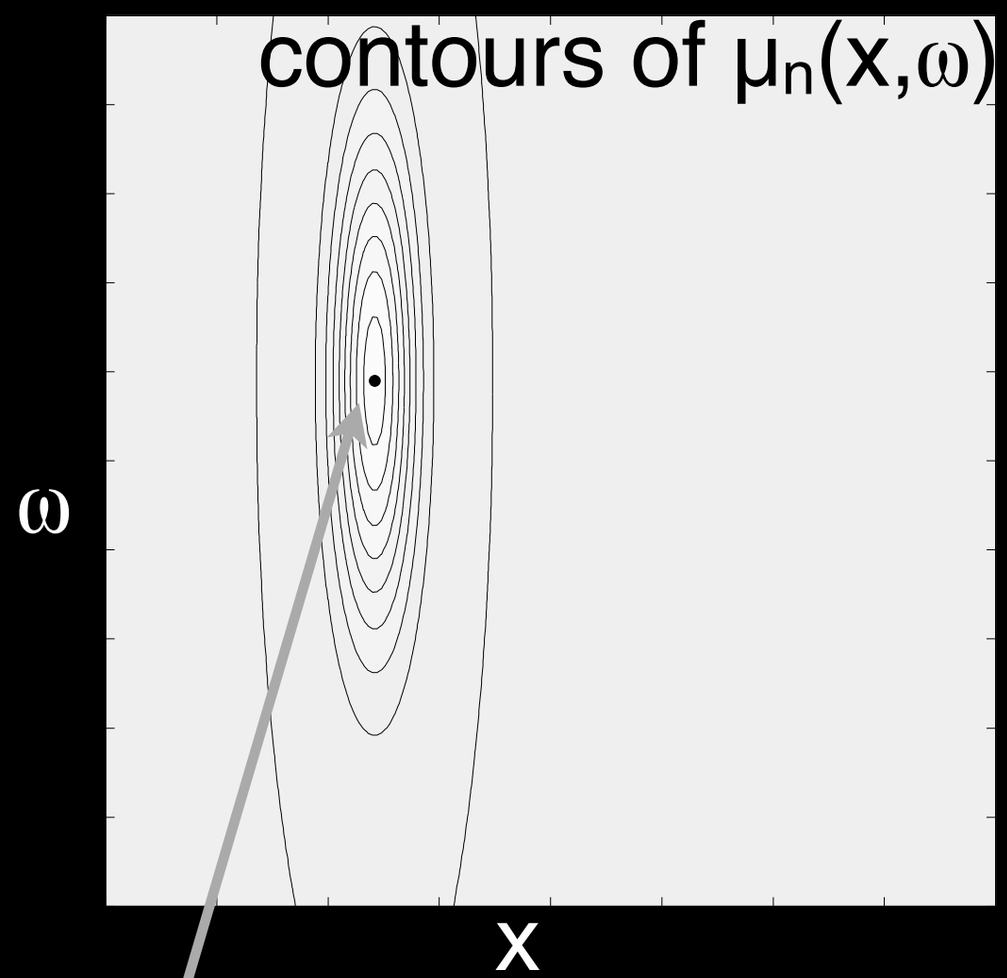


$\omega \sim \text{Normal}(1,1/9)$.

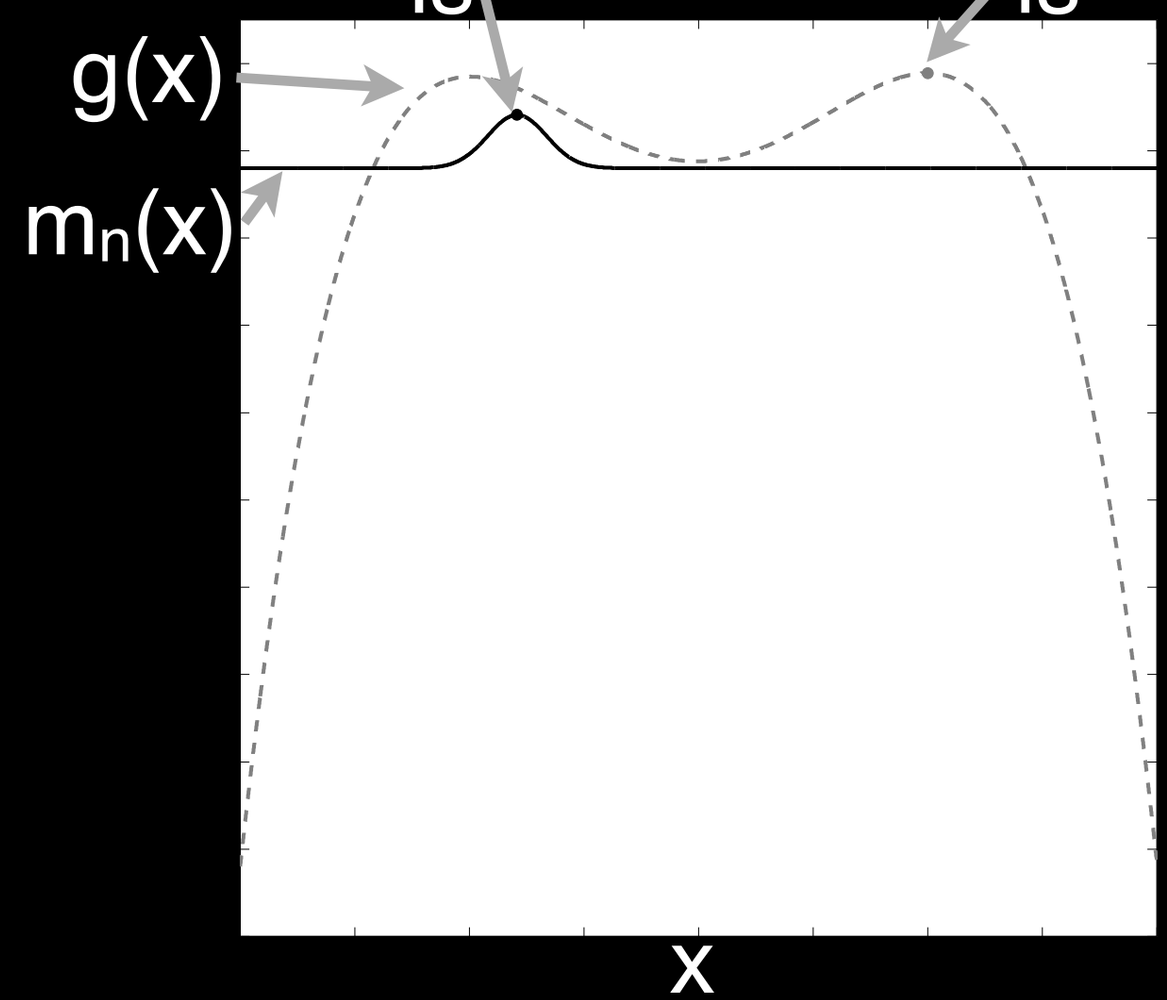
$f(x,\omega)=-100(\omega-x^2)^2-(1-x)^2$.

$g(x)=(1-x)^2 + 100[(1-x^2)^2+(1/9)]$

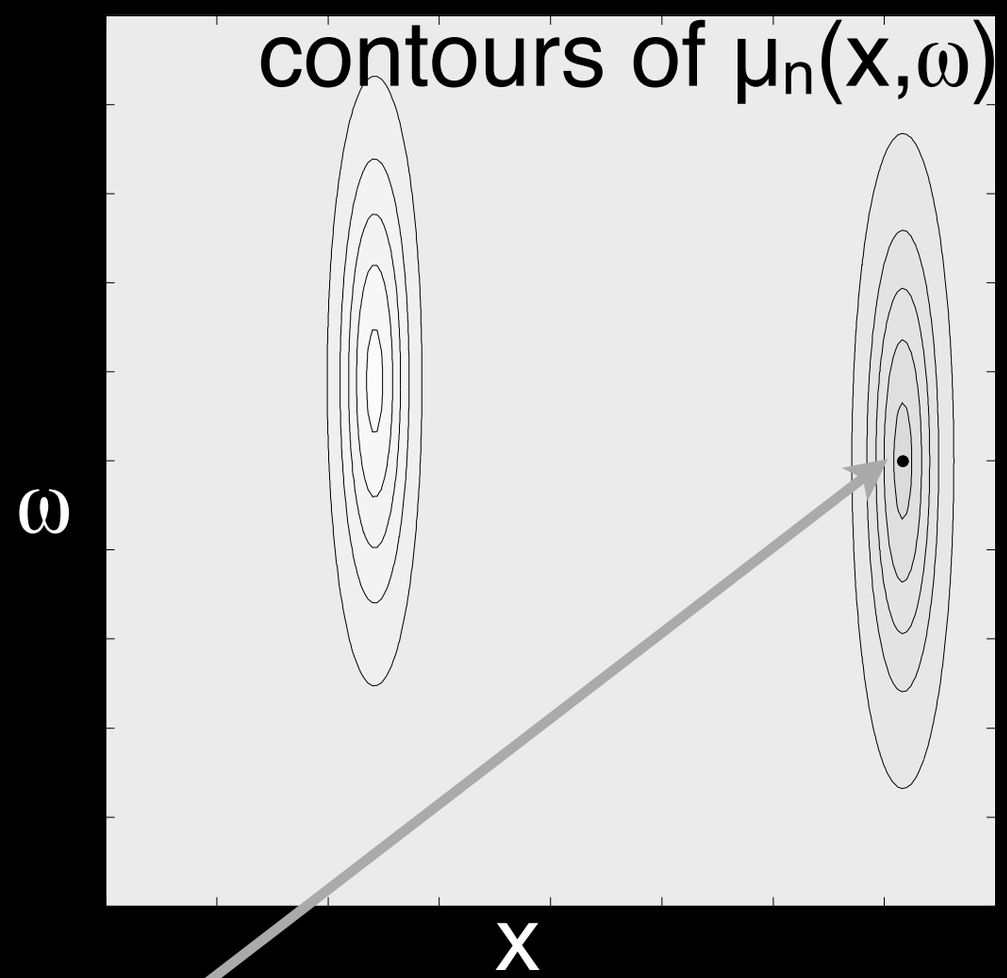
Here's the algorithm solving a test problem



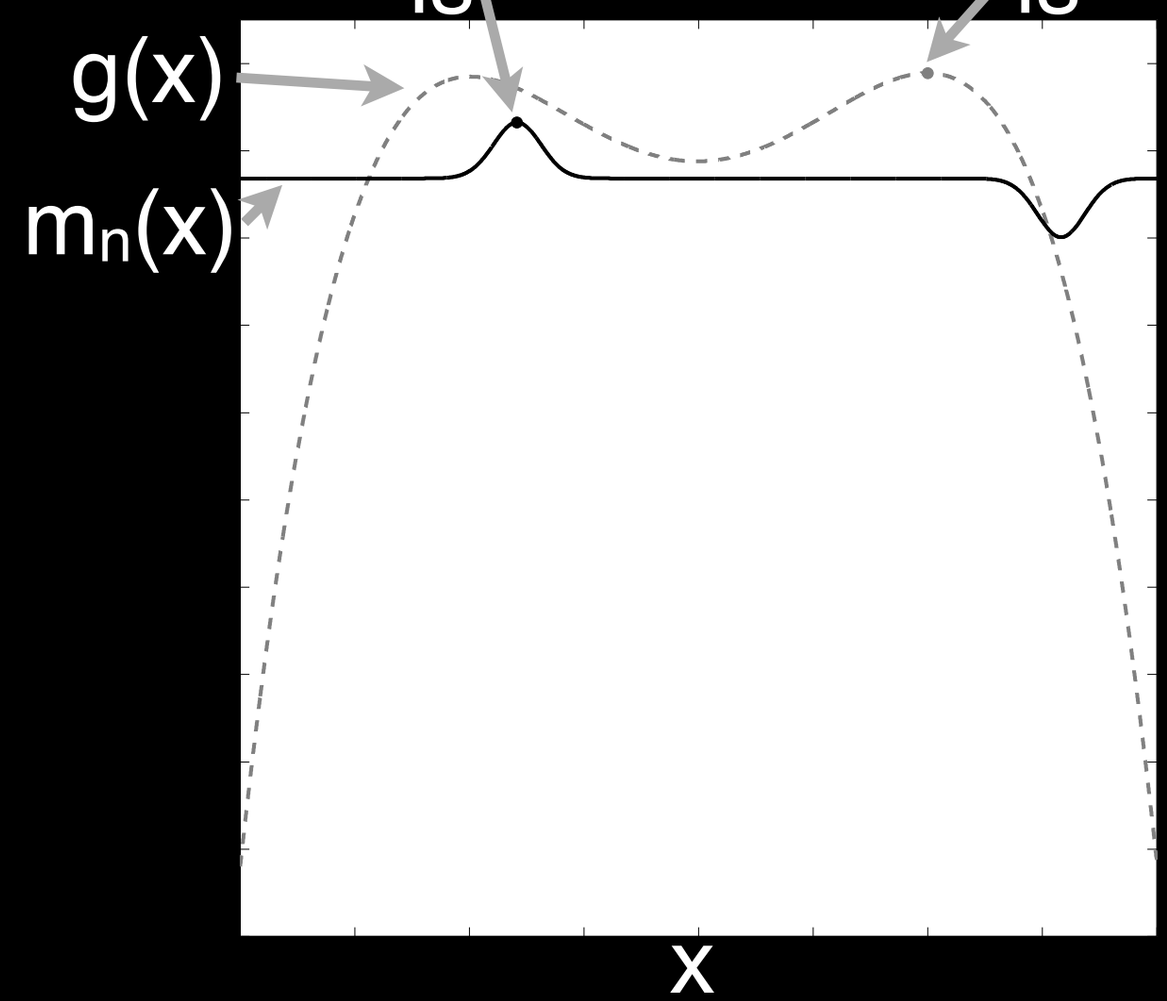
where we last sampled



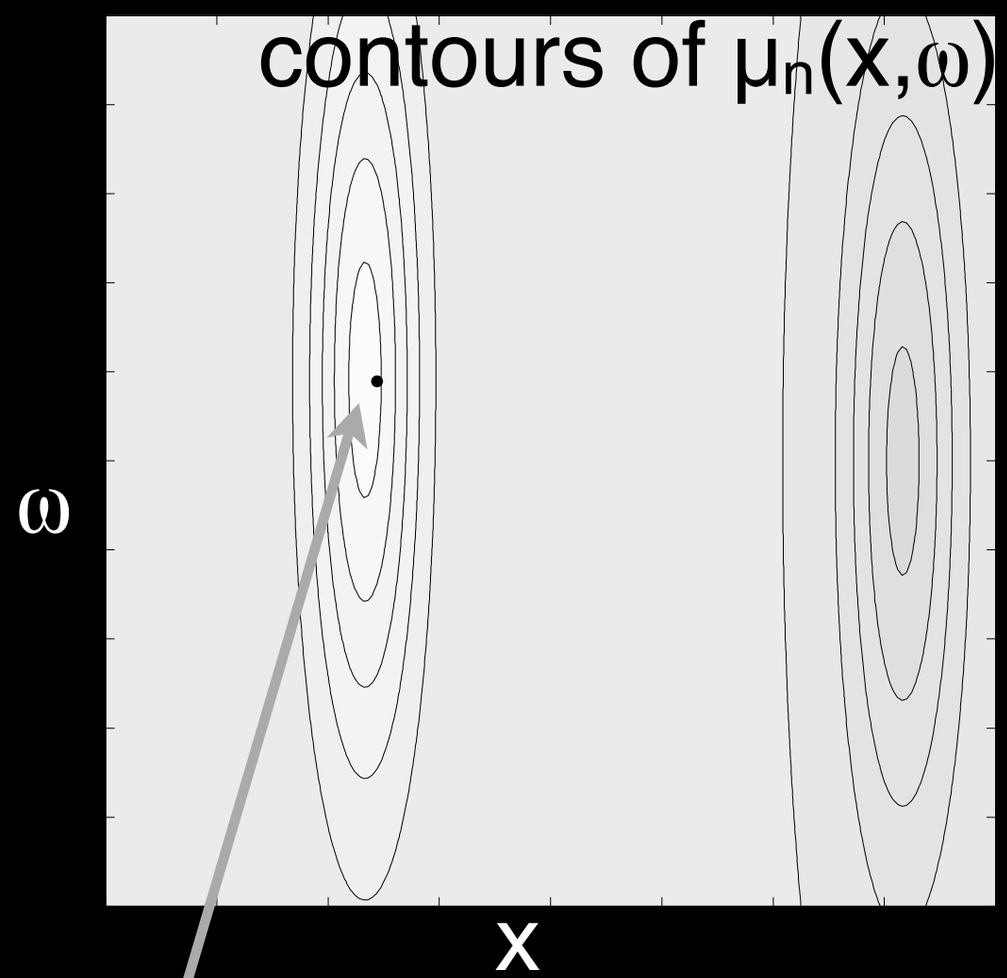
Here's the algorithm solving a test problem



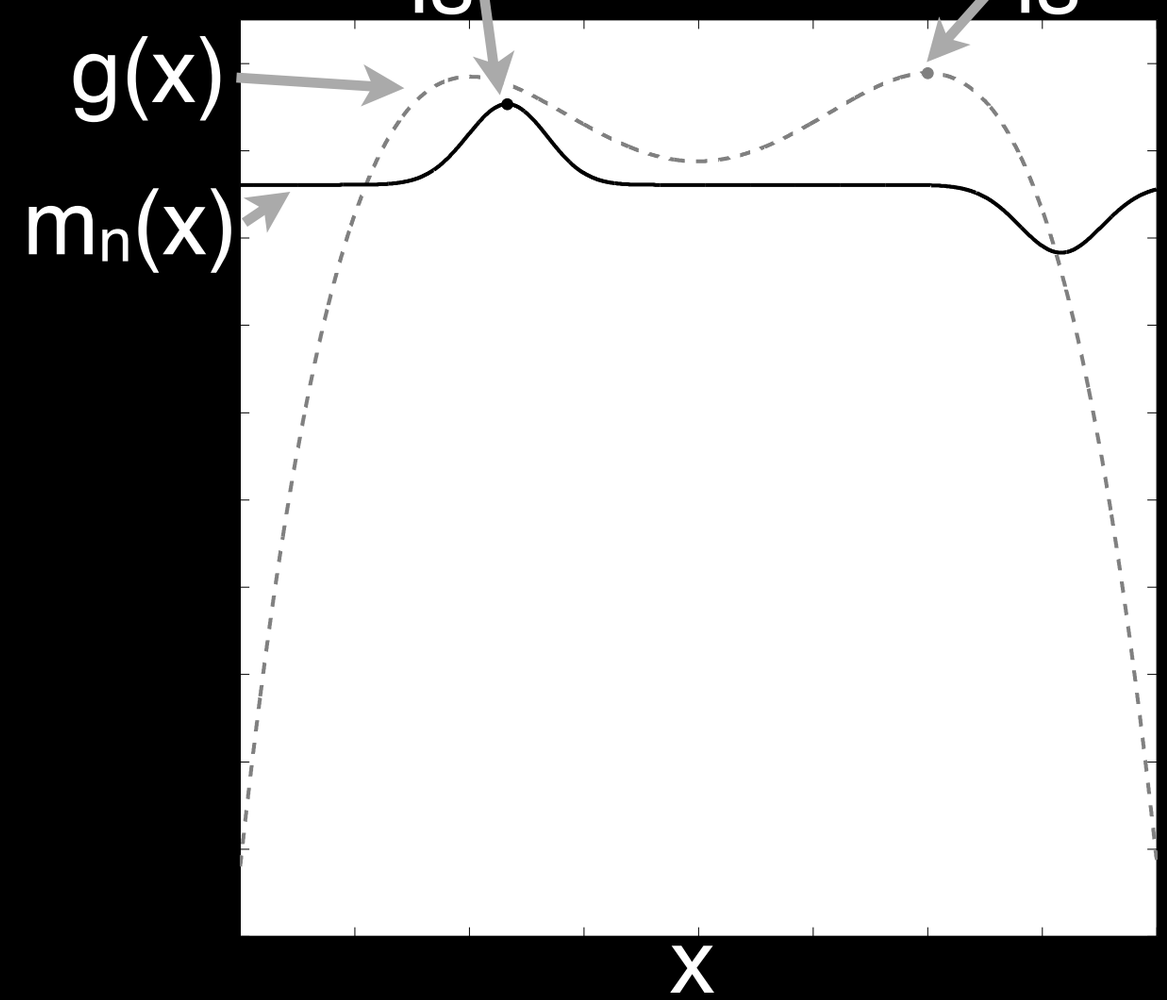
where we last sampled



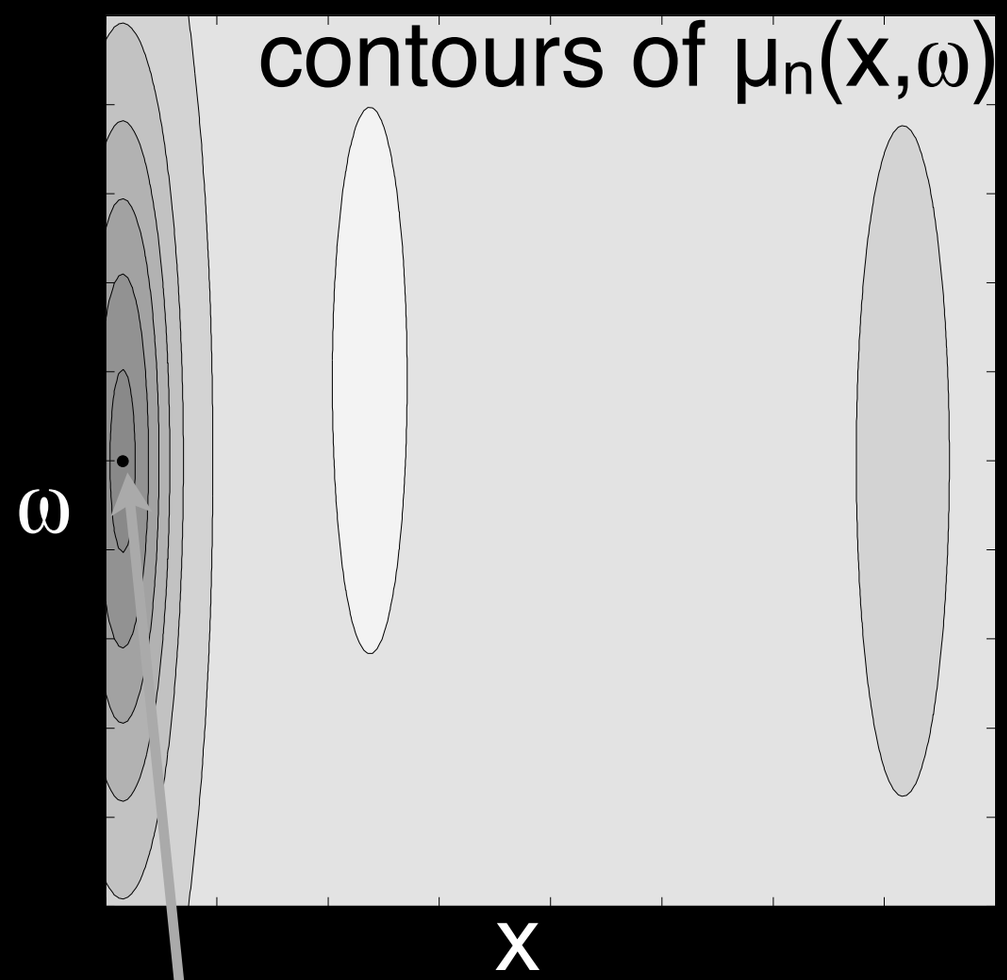
Here's the algorithm solving a test problem



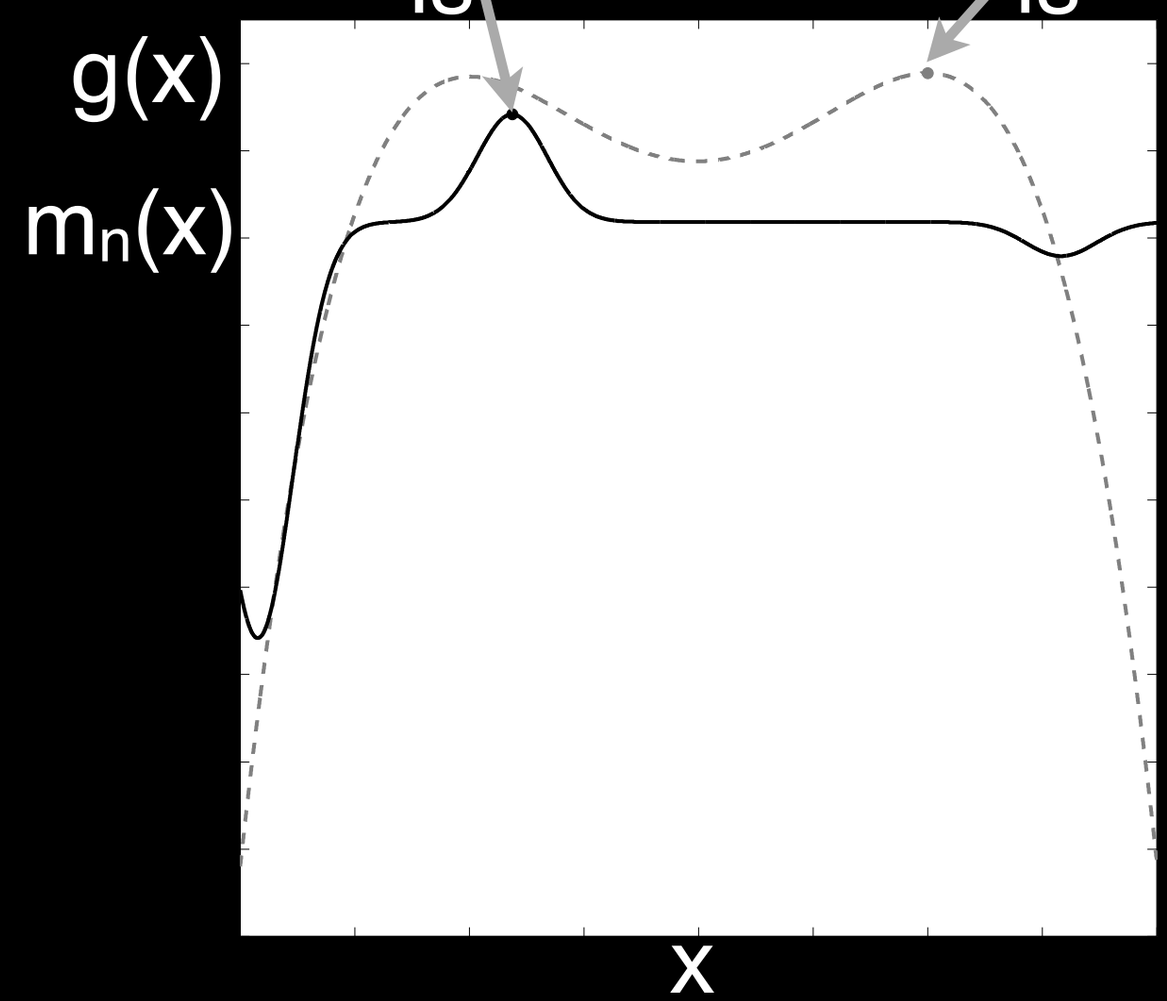
where we last sampled



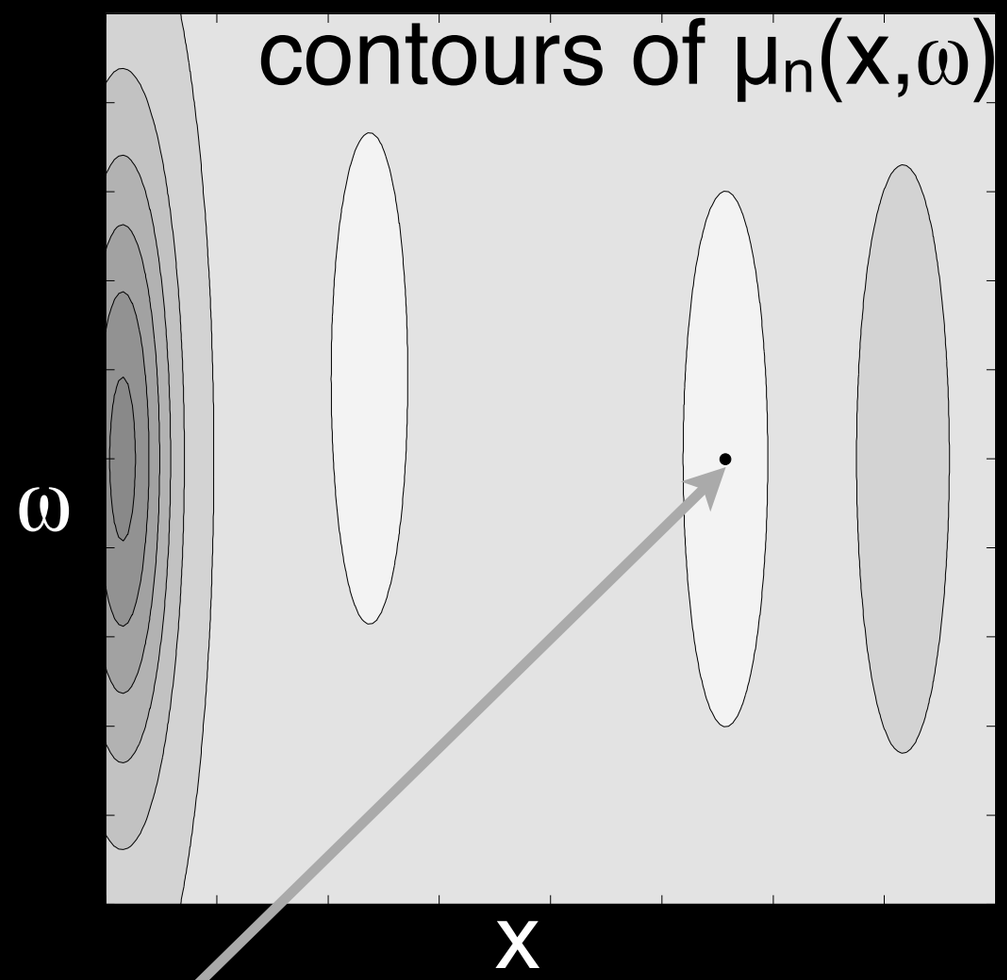
Here's the algorithm solving a test problem



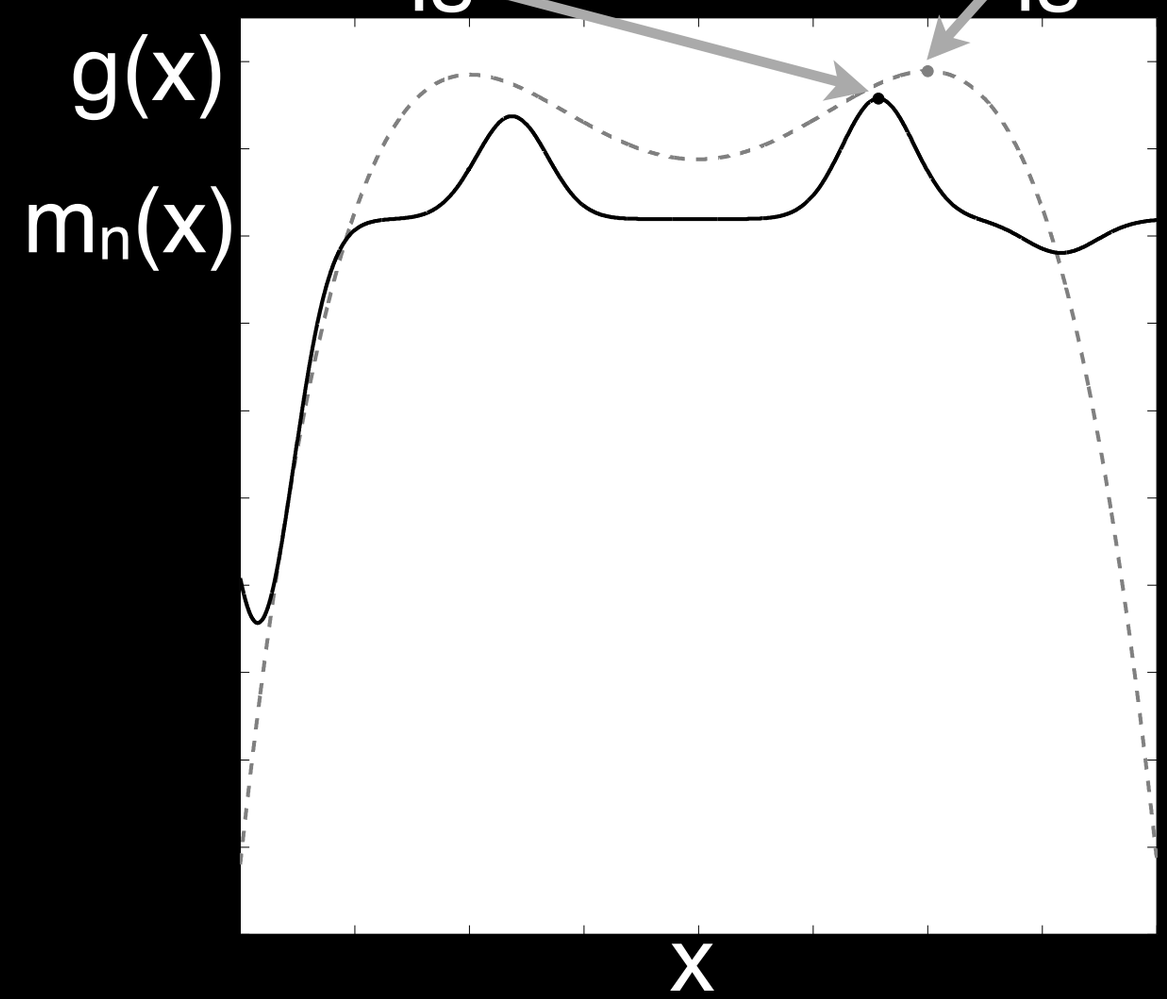
where we last sampled



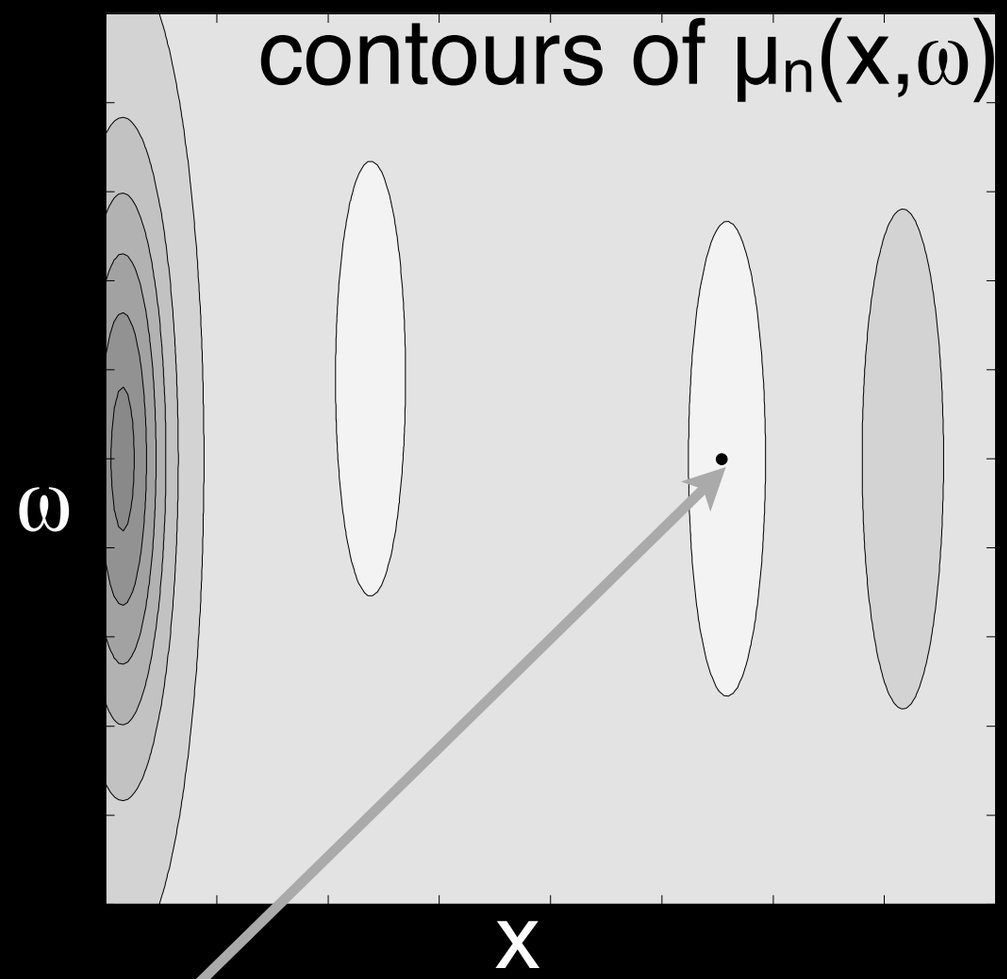
Here's the algorithm solving a test problem



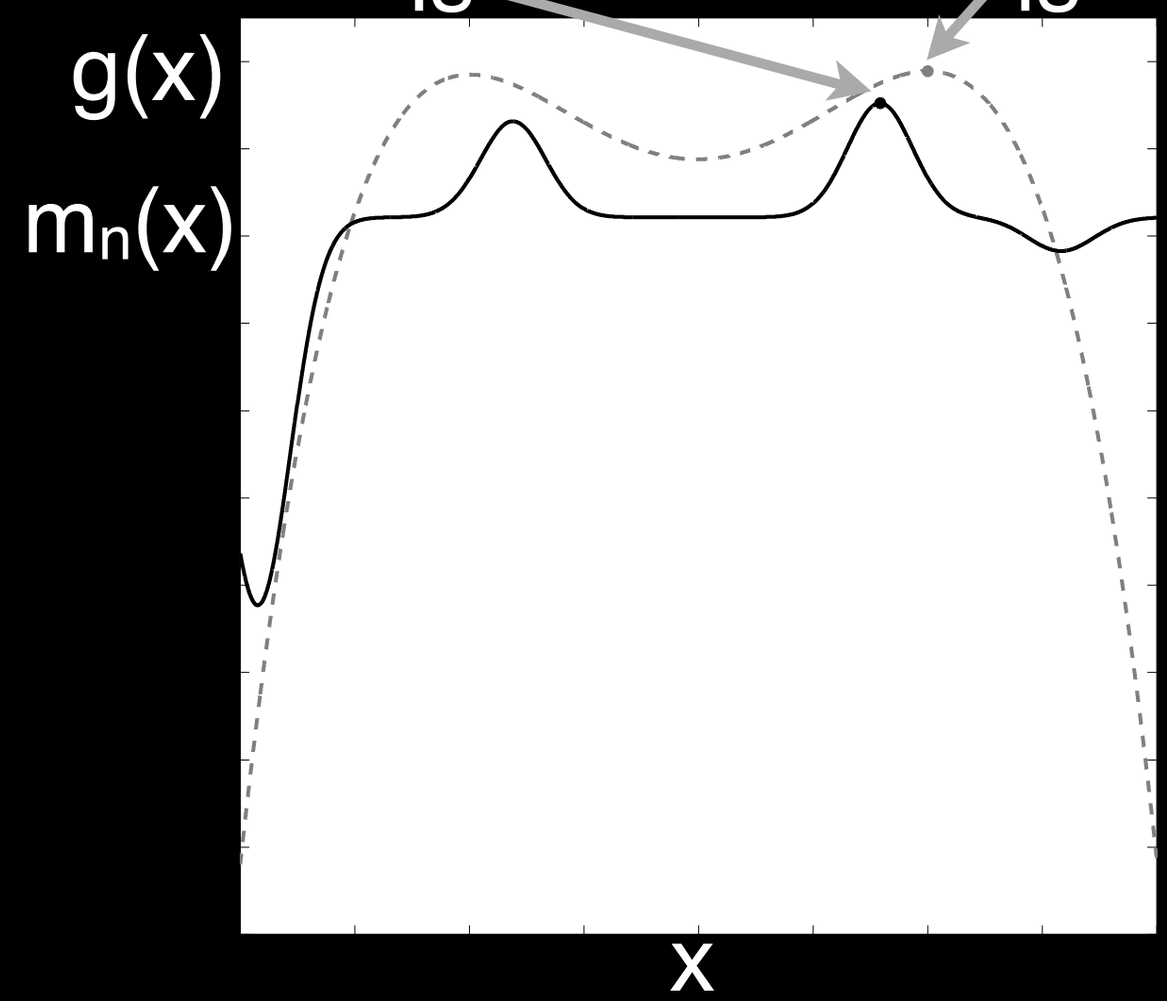
where we
last sampled



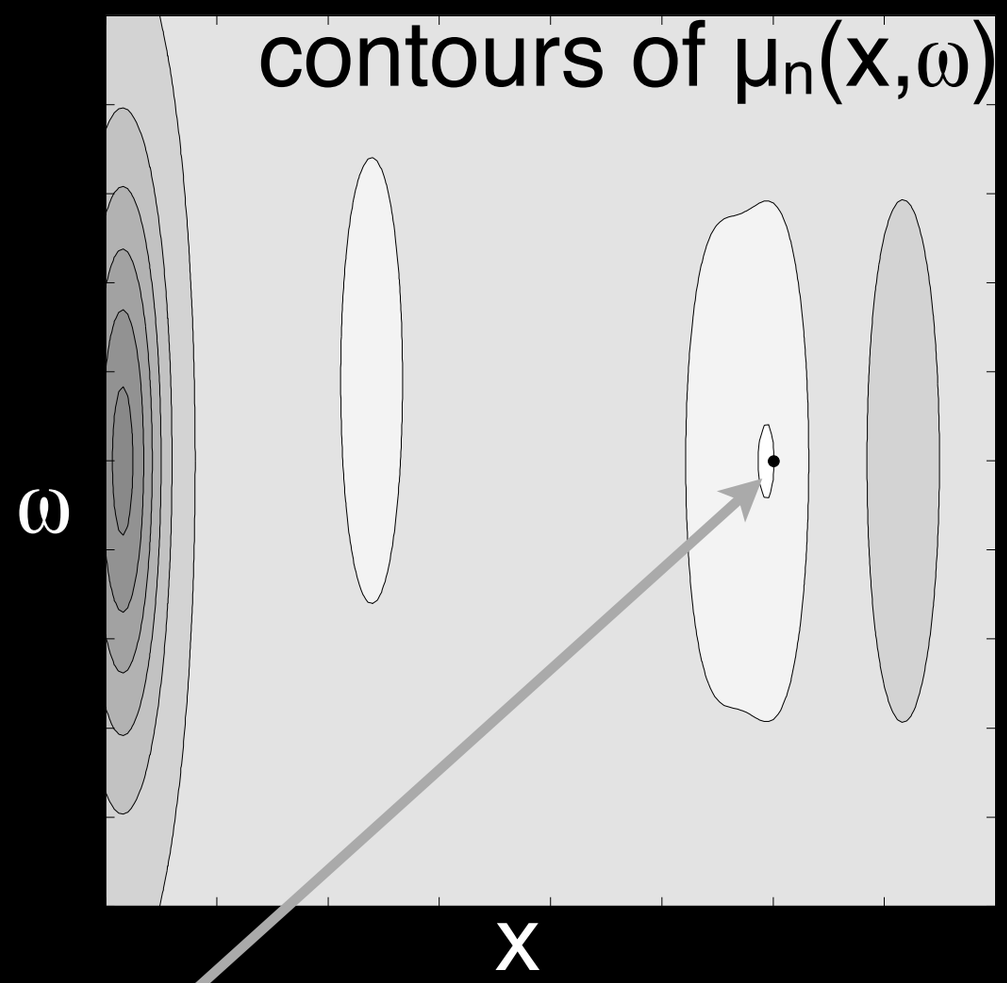
Here's the algorithm solving a test problem



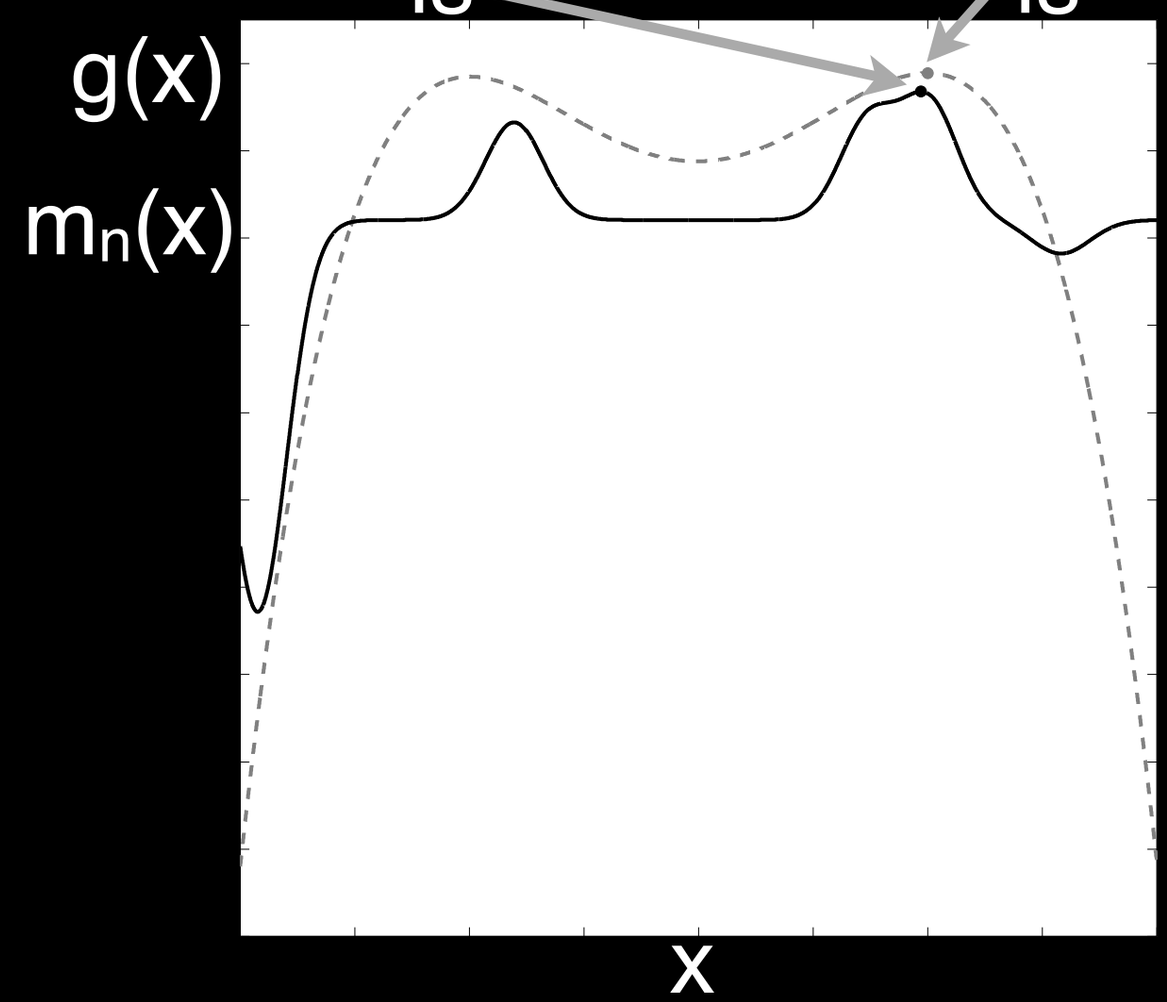
where we last sampled



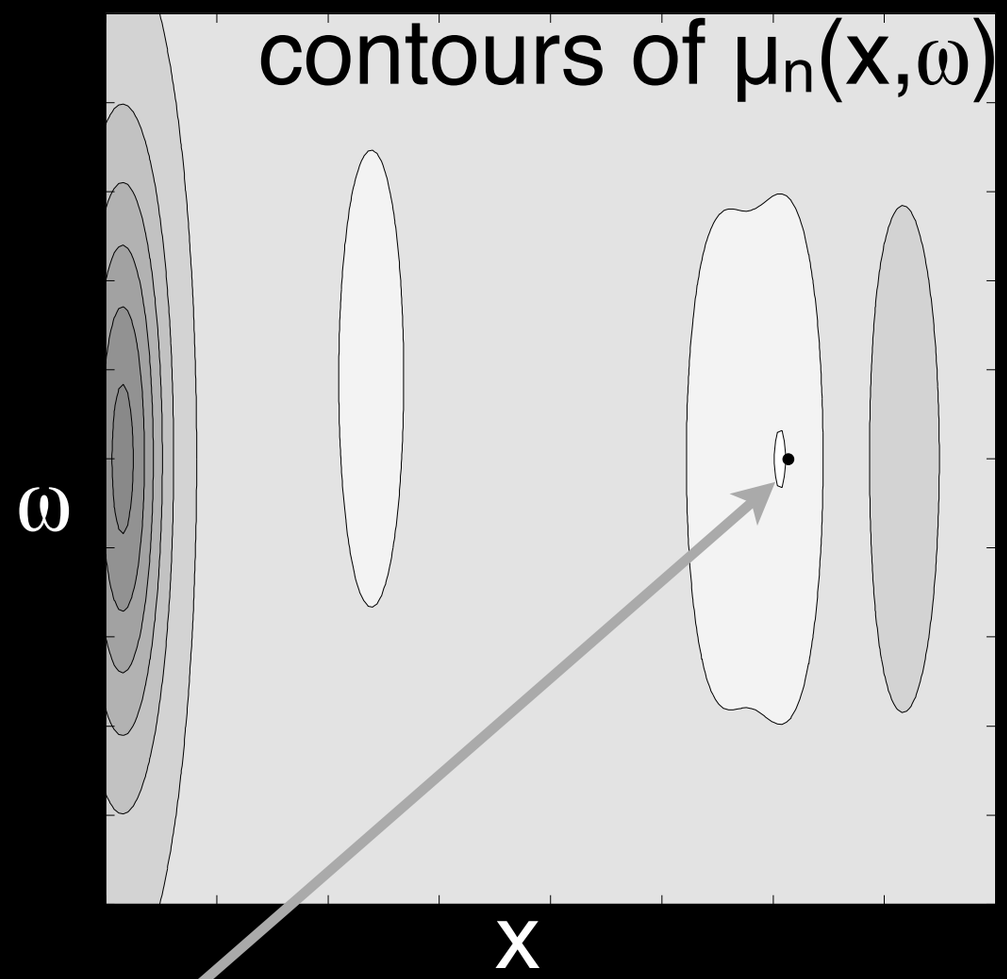
Here's the algorithm solving a test problem



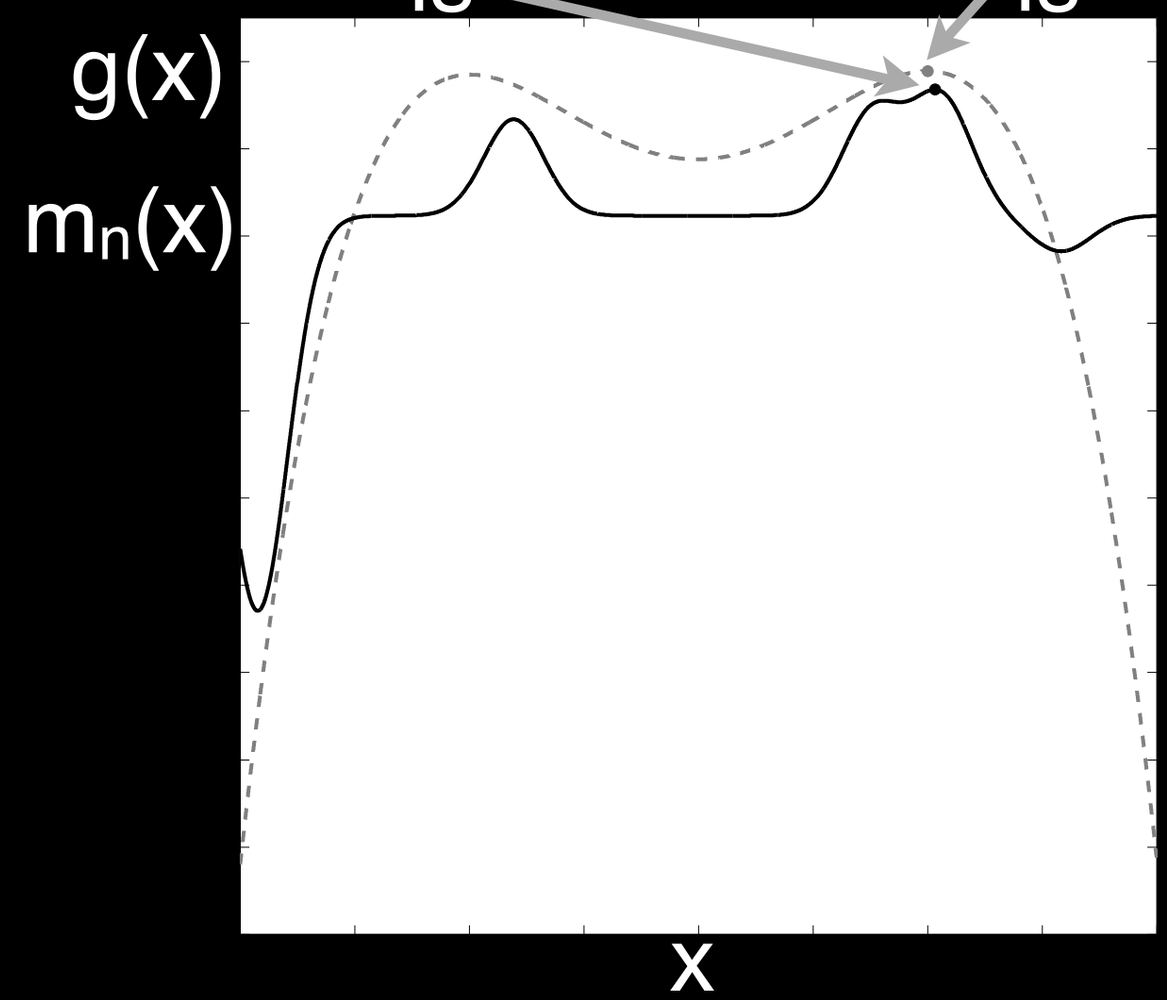
where we last sampled



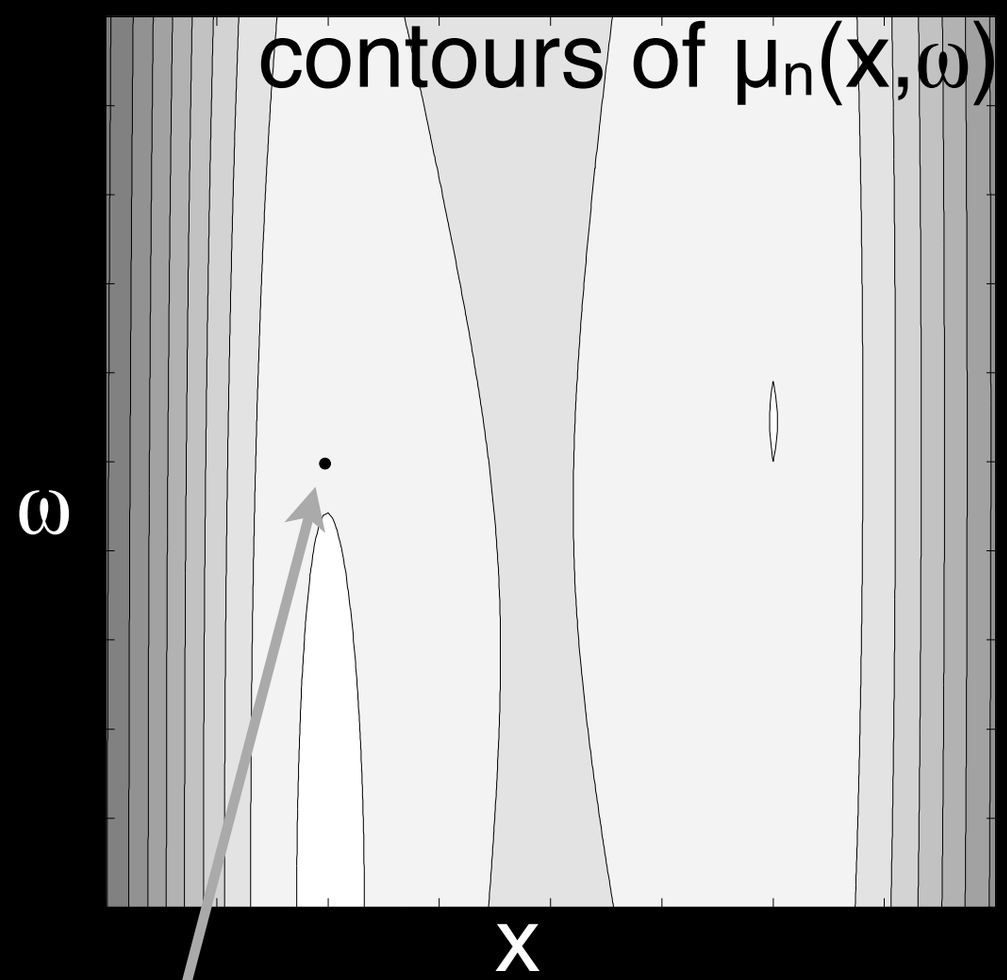
Here's the algorithm solving a test problem



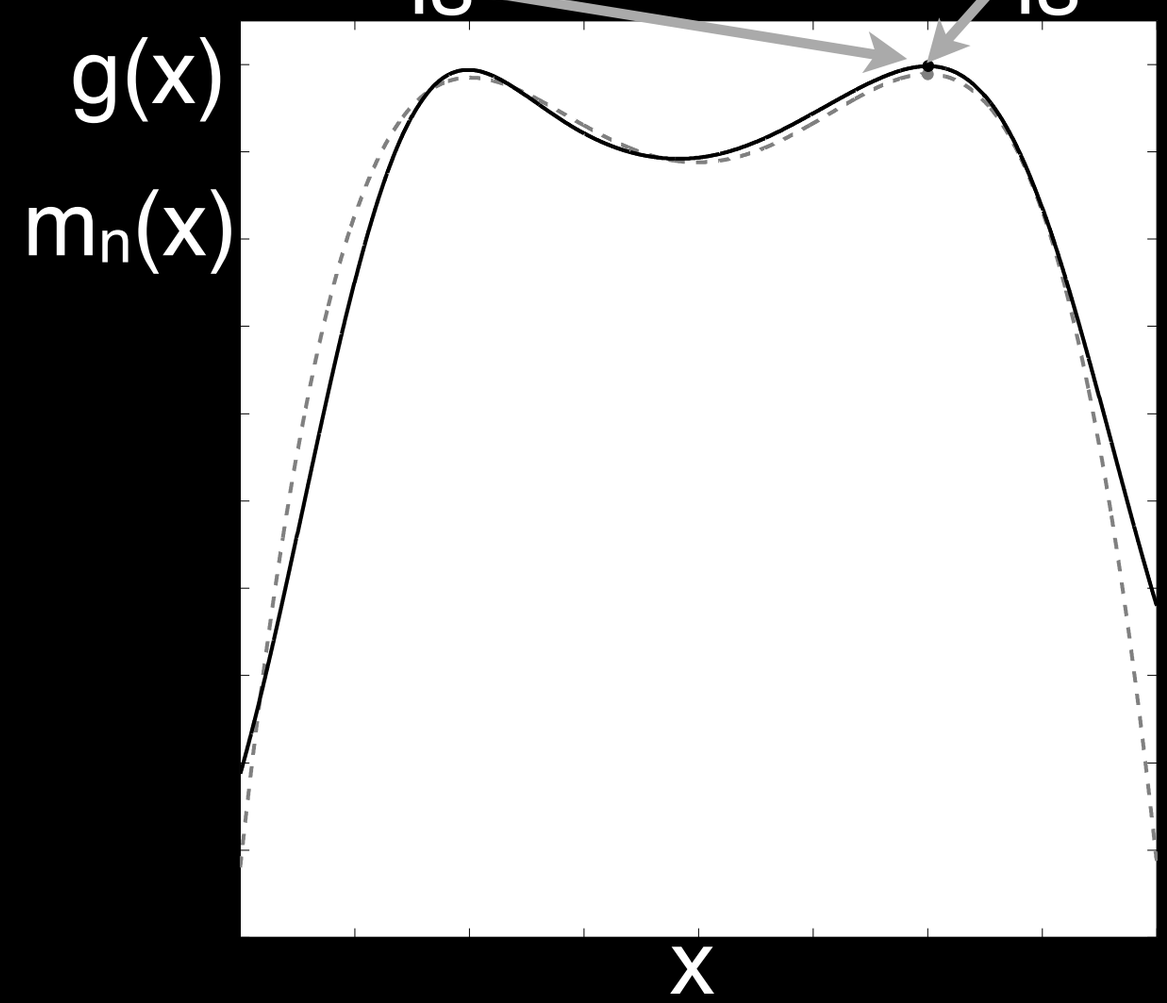
where we last sampled



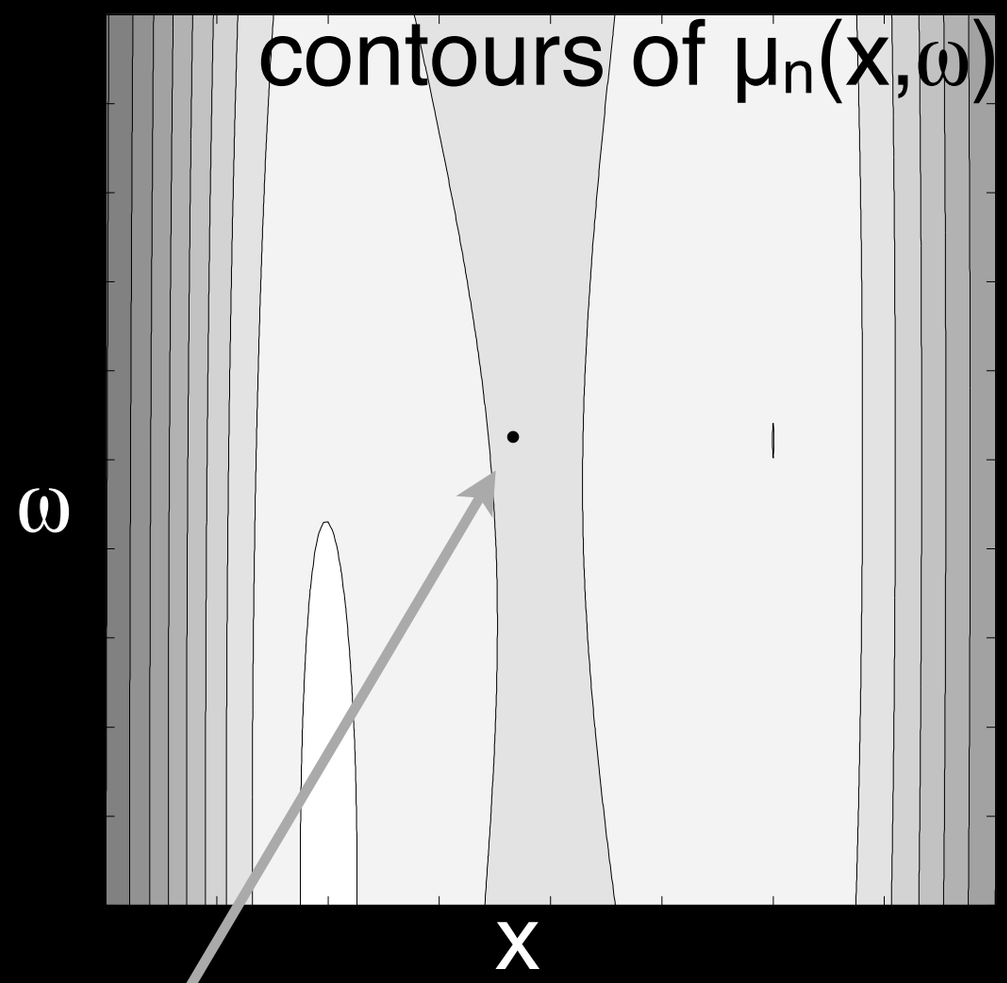
Here's the algorithm solving a test problem



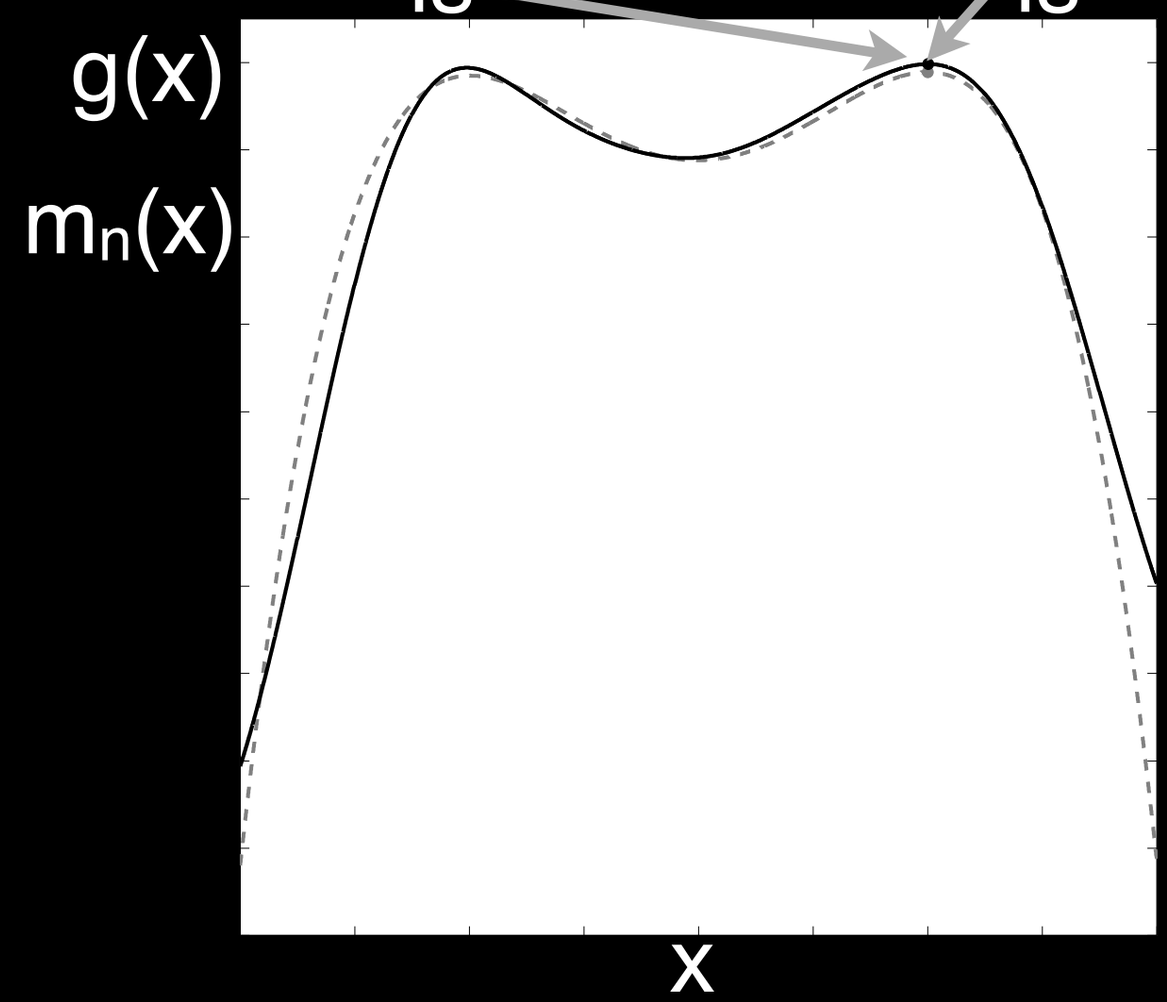
where we last sampled



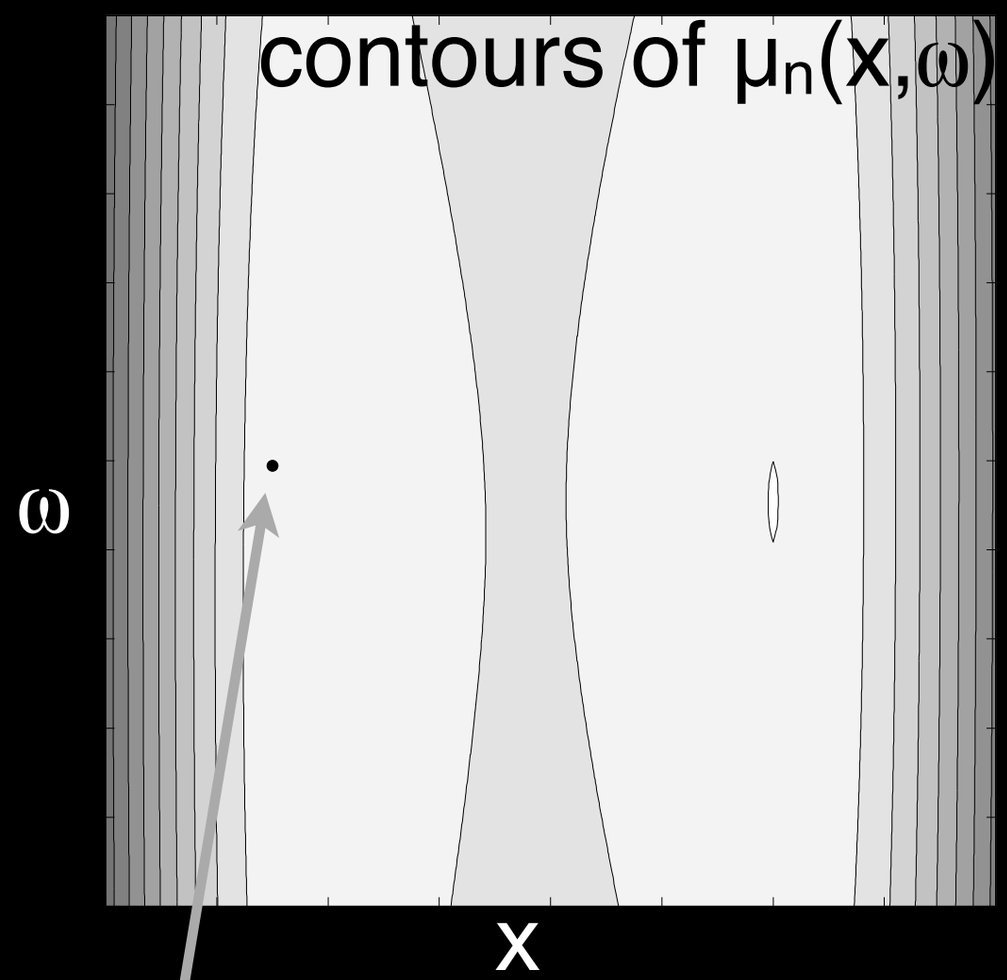
Here's the algorithm solving a test problem



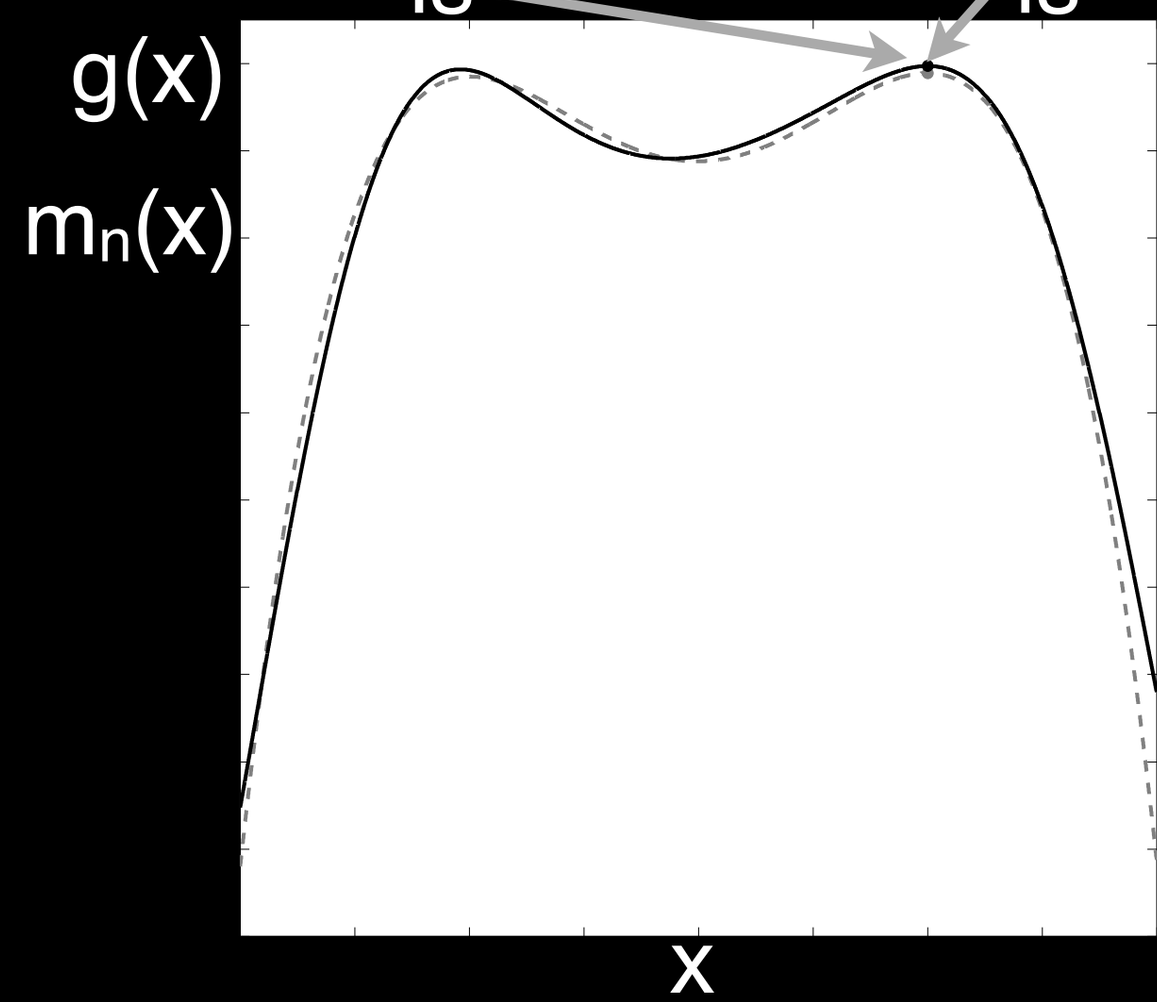
where we last sampled



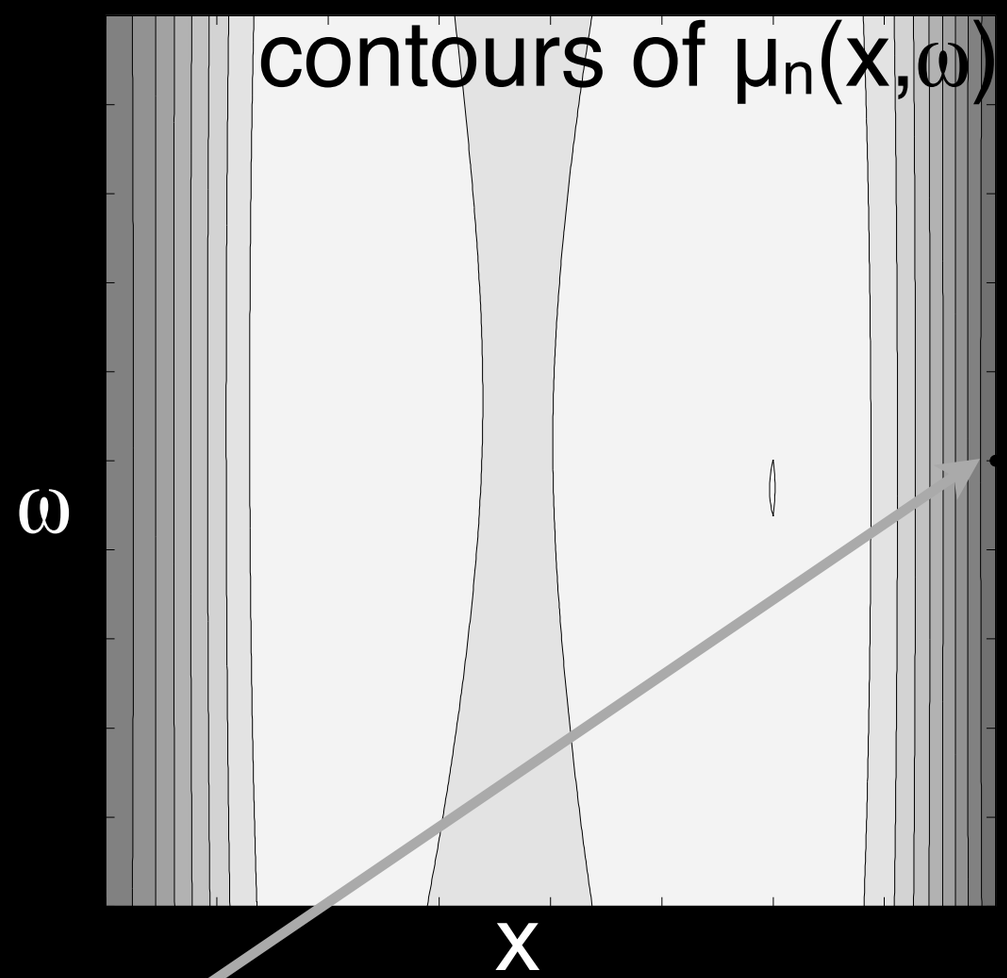
Here's the algorithm solving a test problem



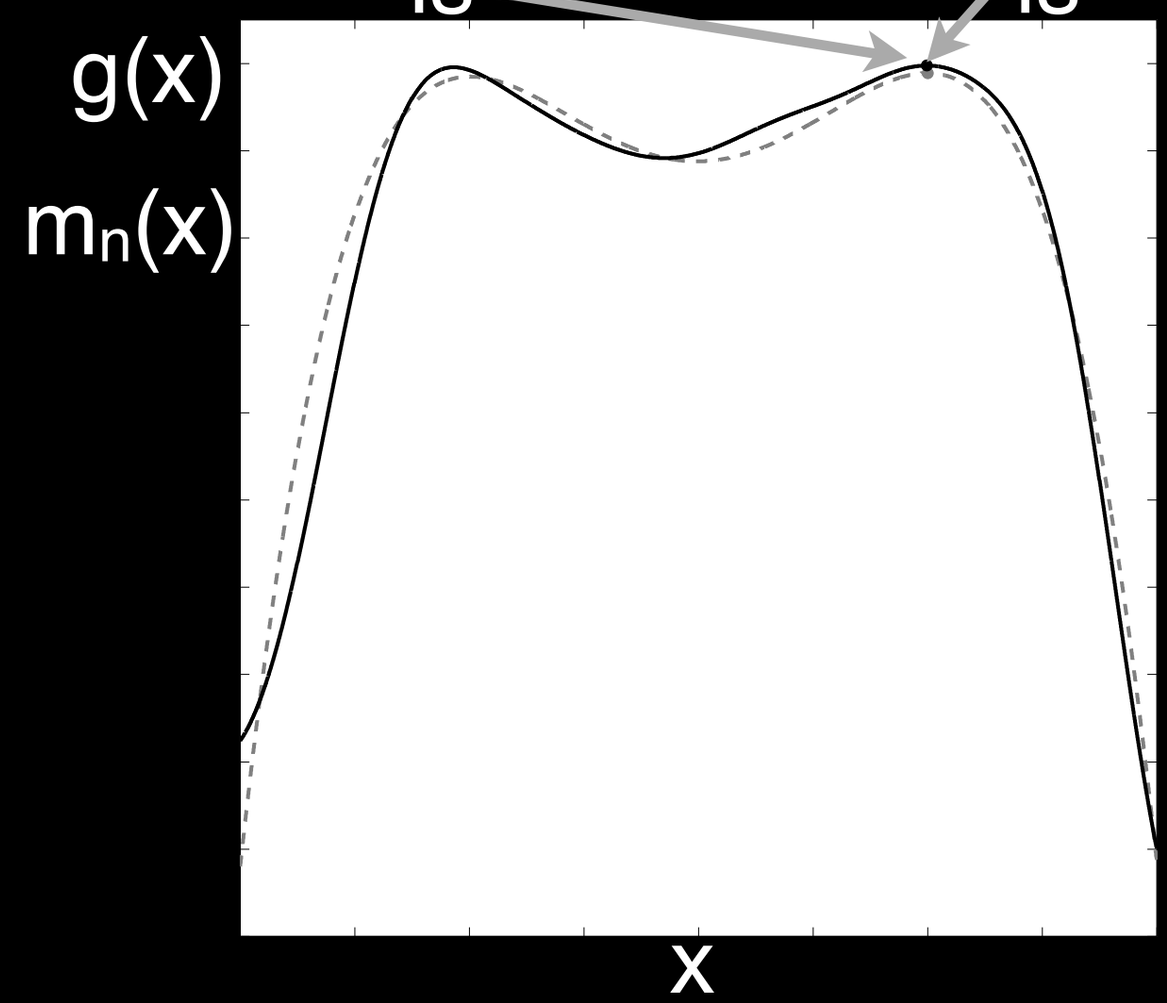
where we
last sampled



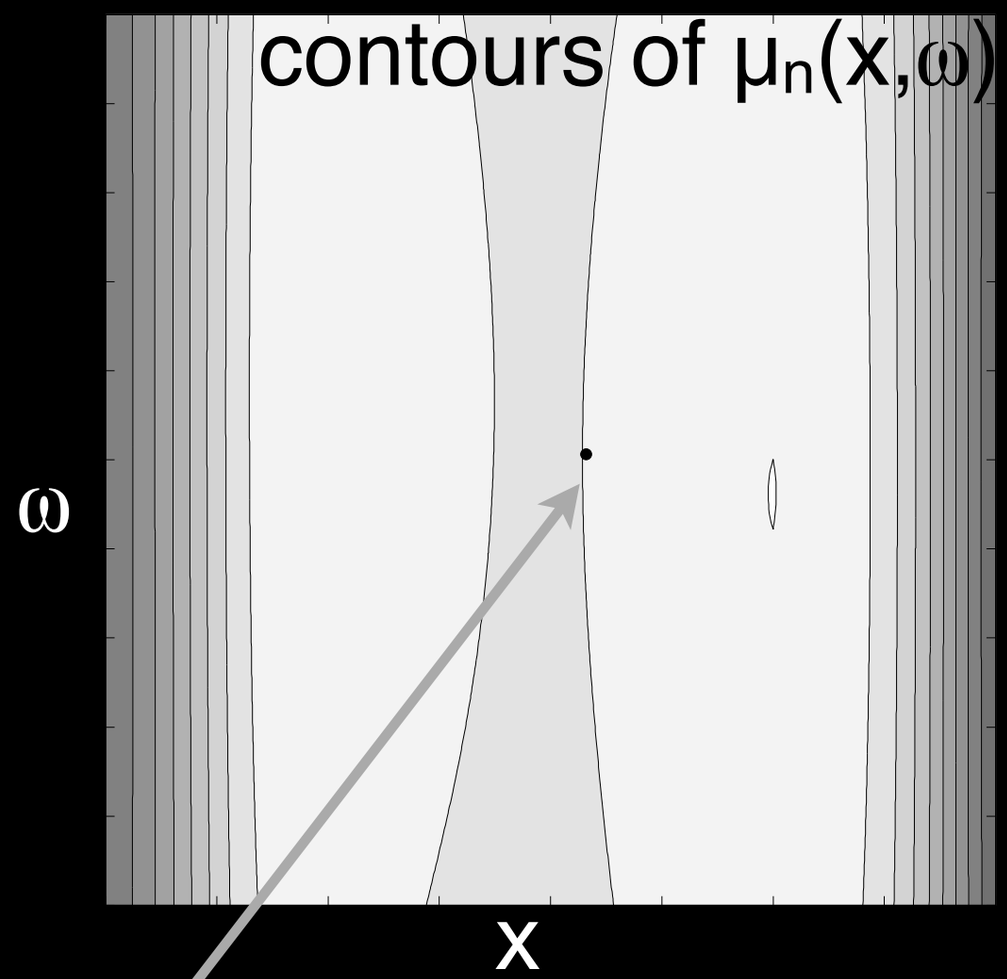
Here's the algorithm solving a test problem



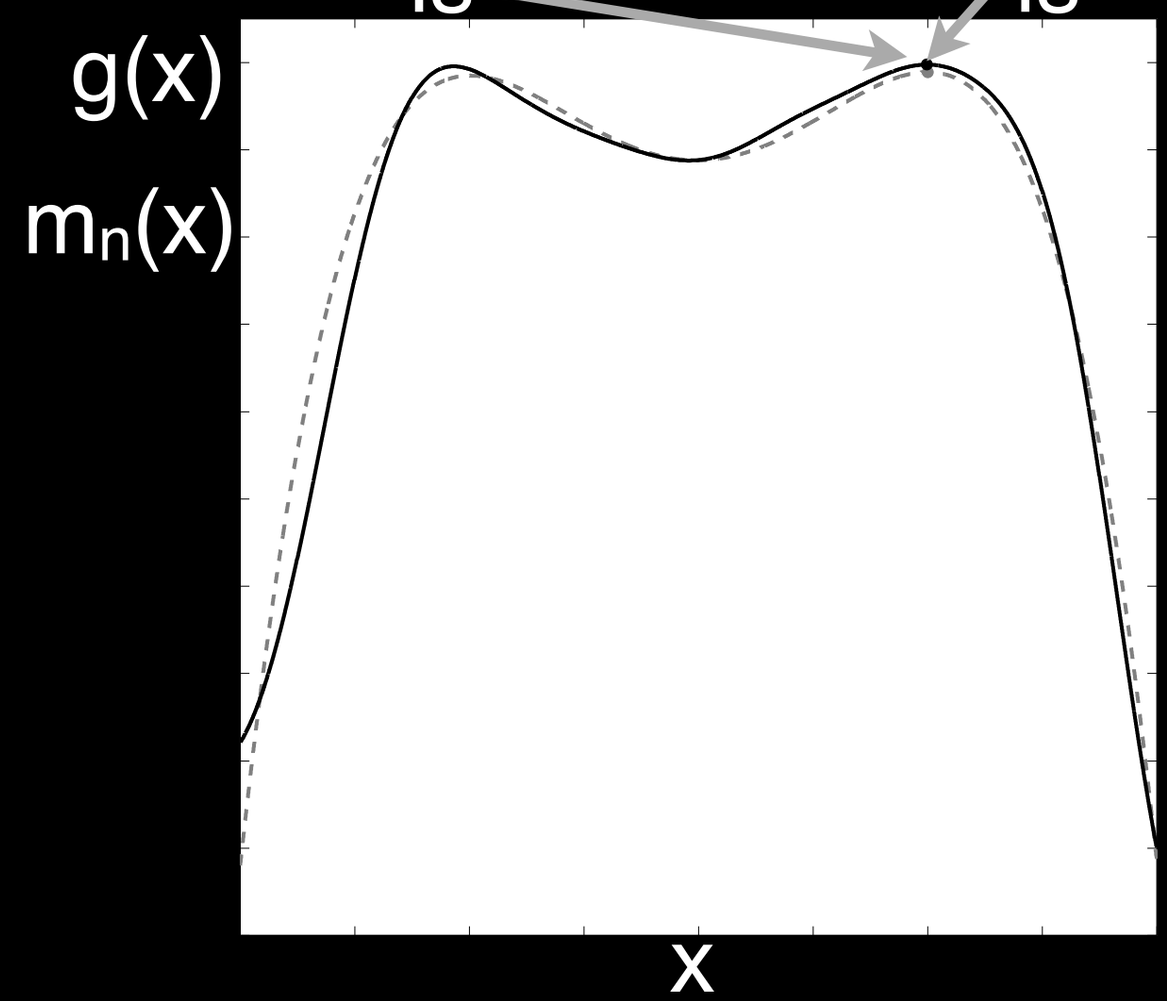
where we last sampled



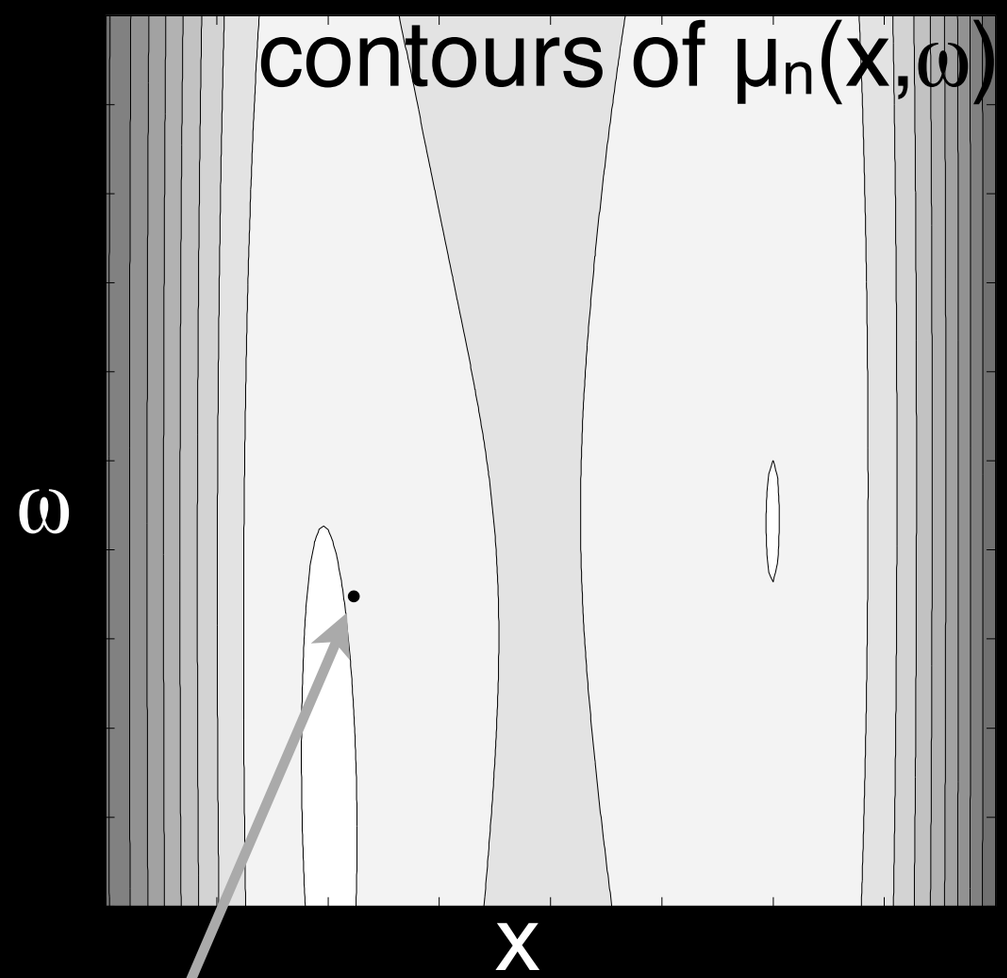
Here's the algorithm solving a test problem



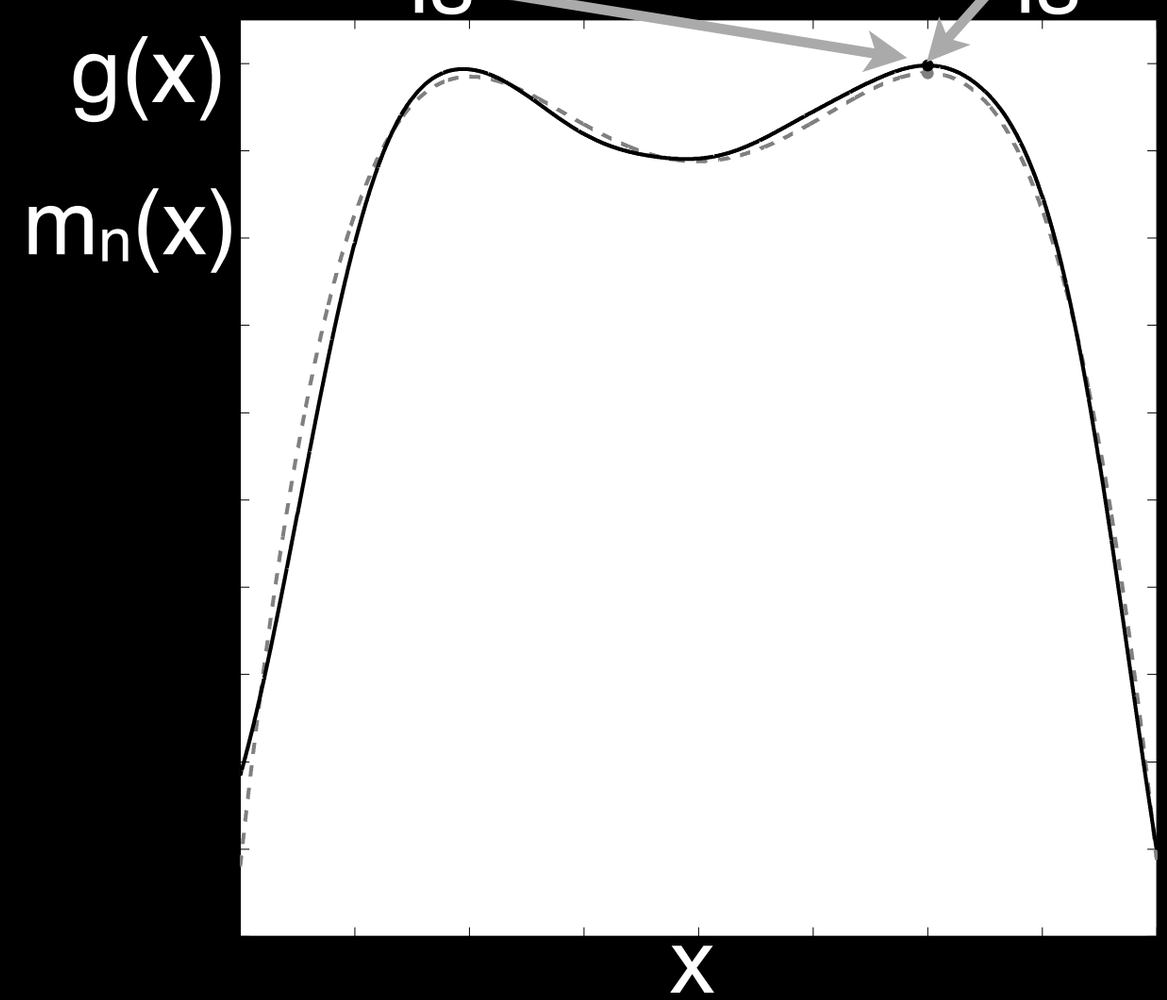
where we last sampled



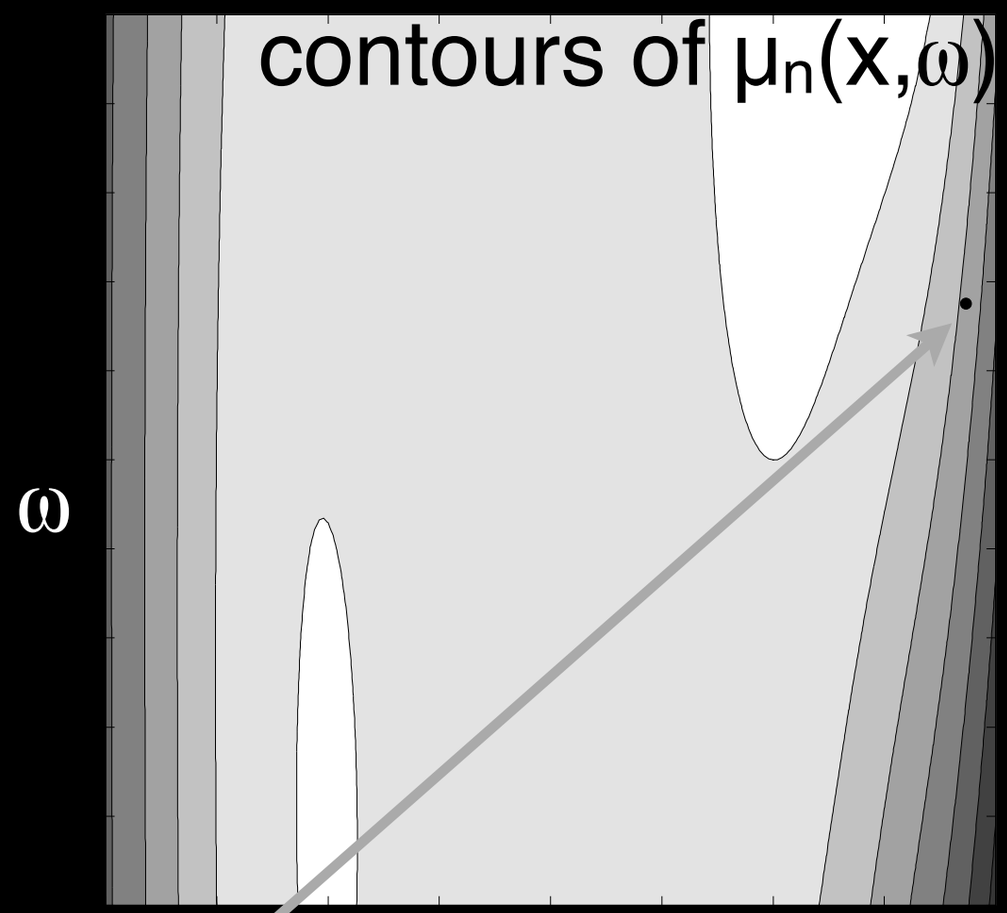
Here's the algorithm solving a test problem



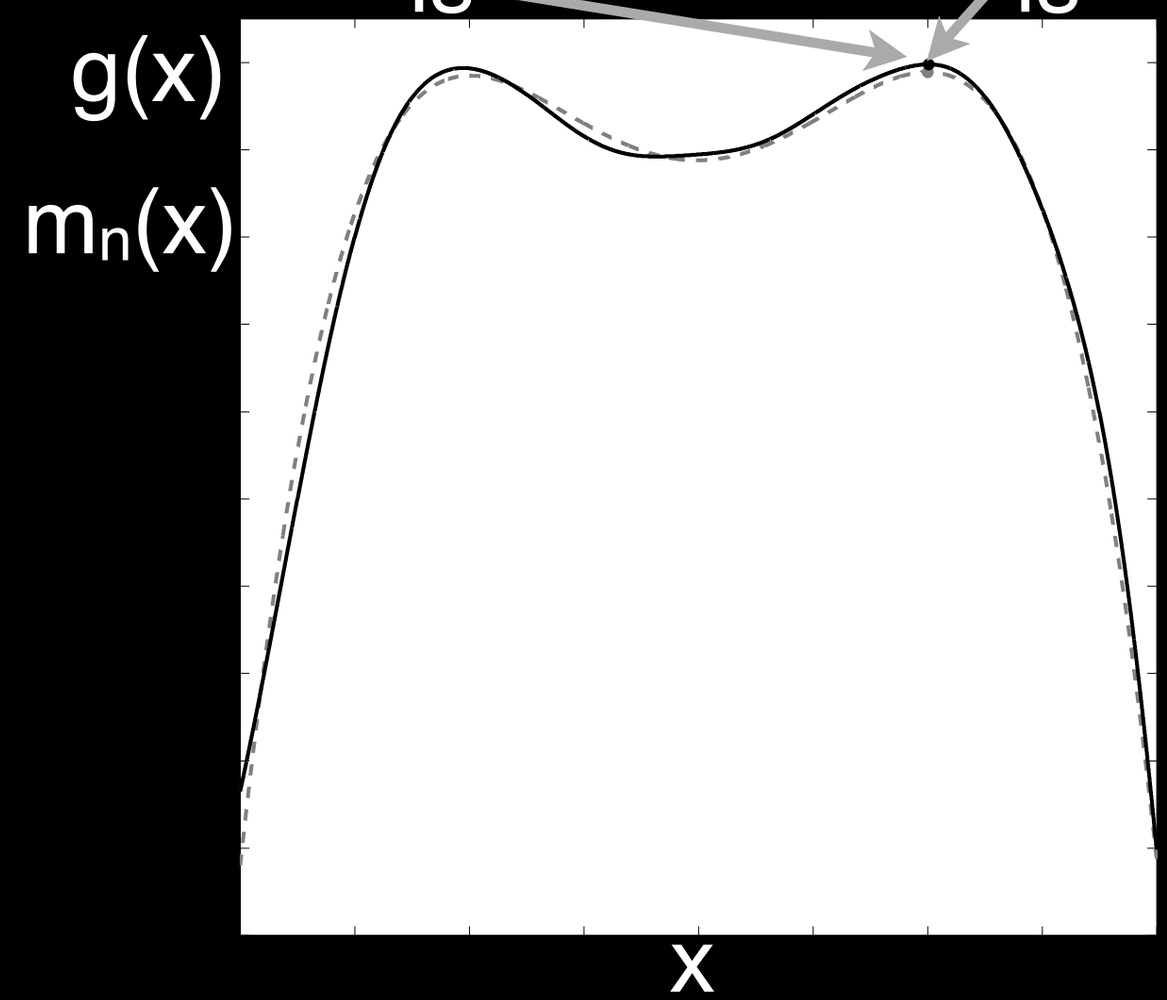
where we last sampled



Here's the algorithm solving a test problem



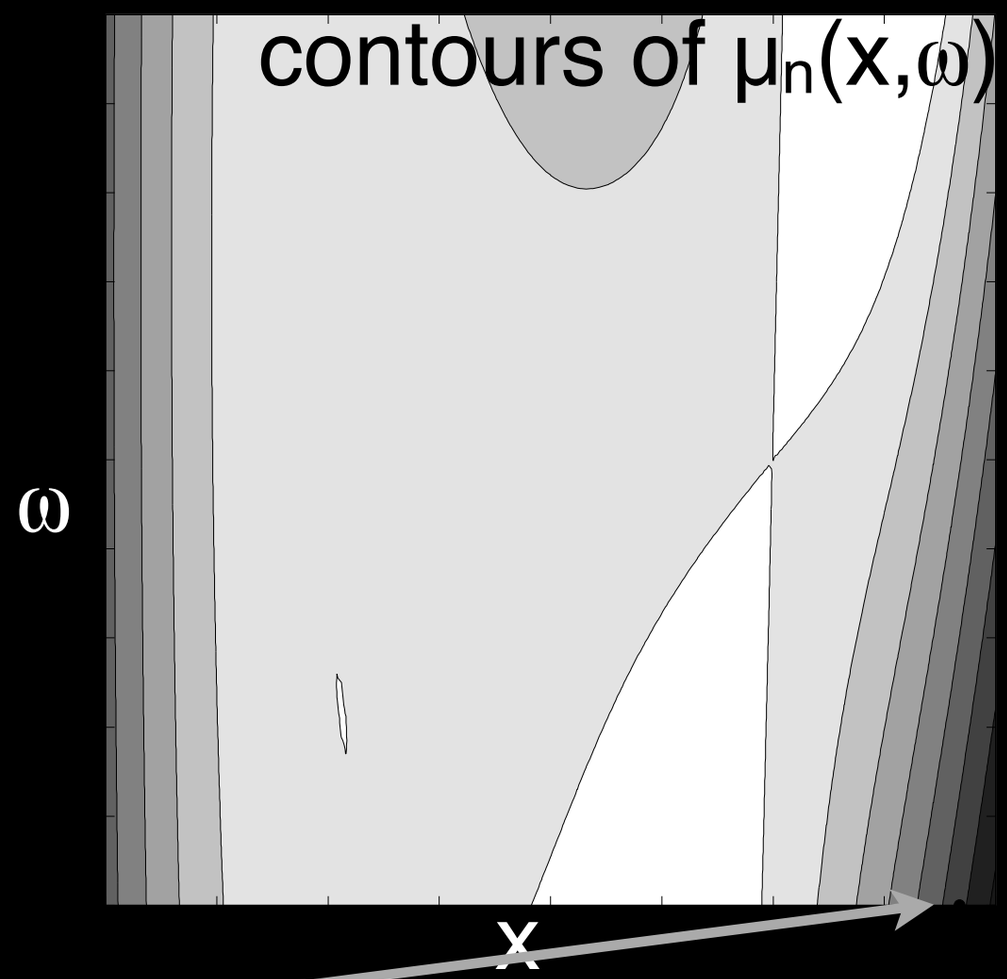
where we last sampled



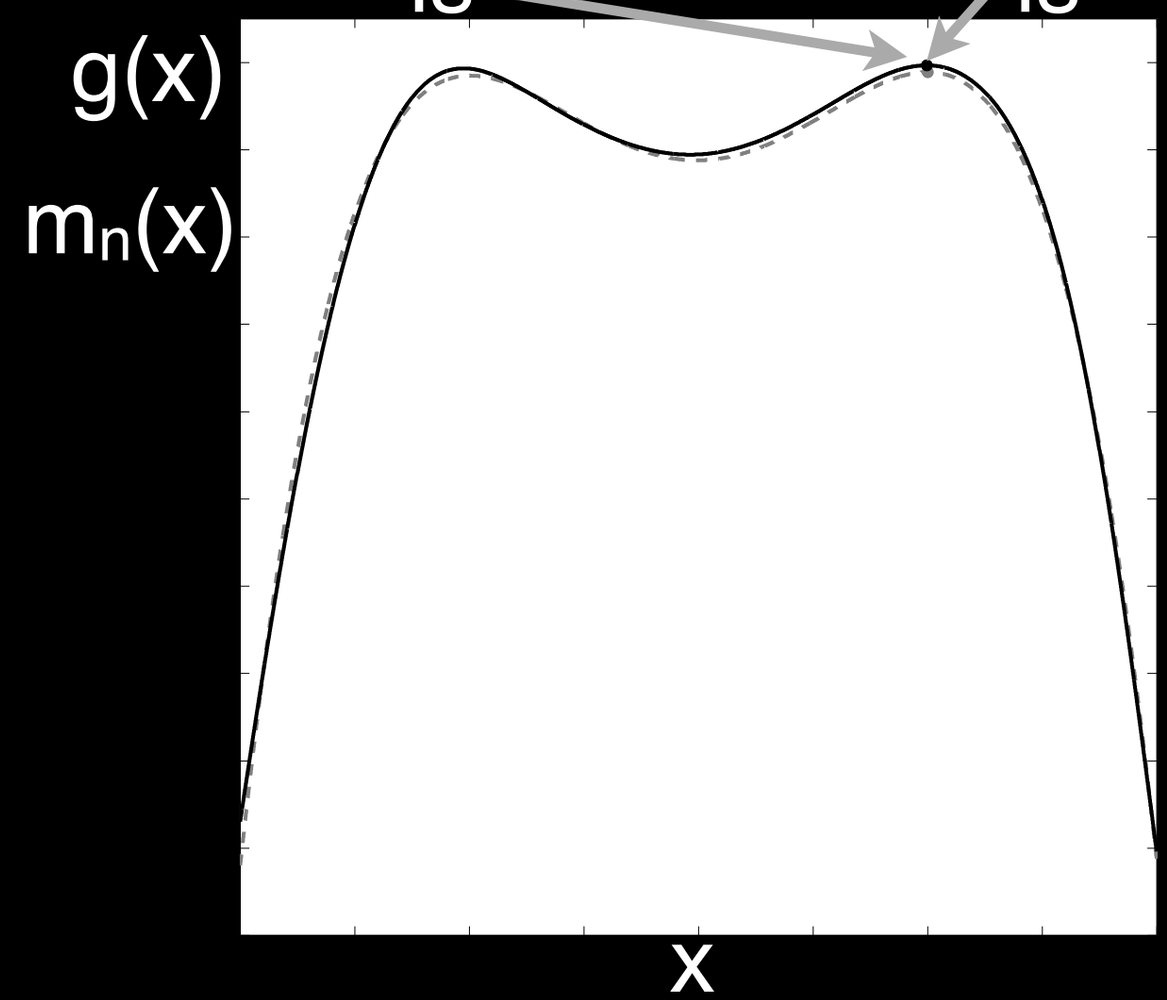
where we think the max is

where the max is

Here's the algorithm solving a test problem



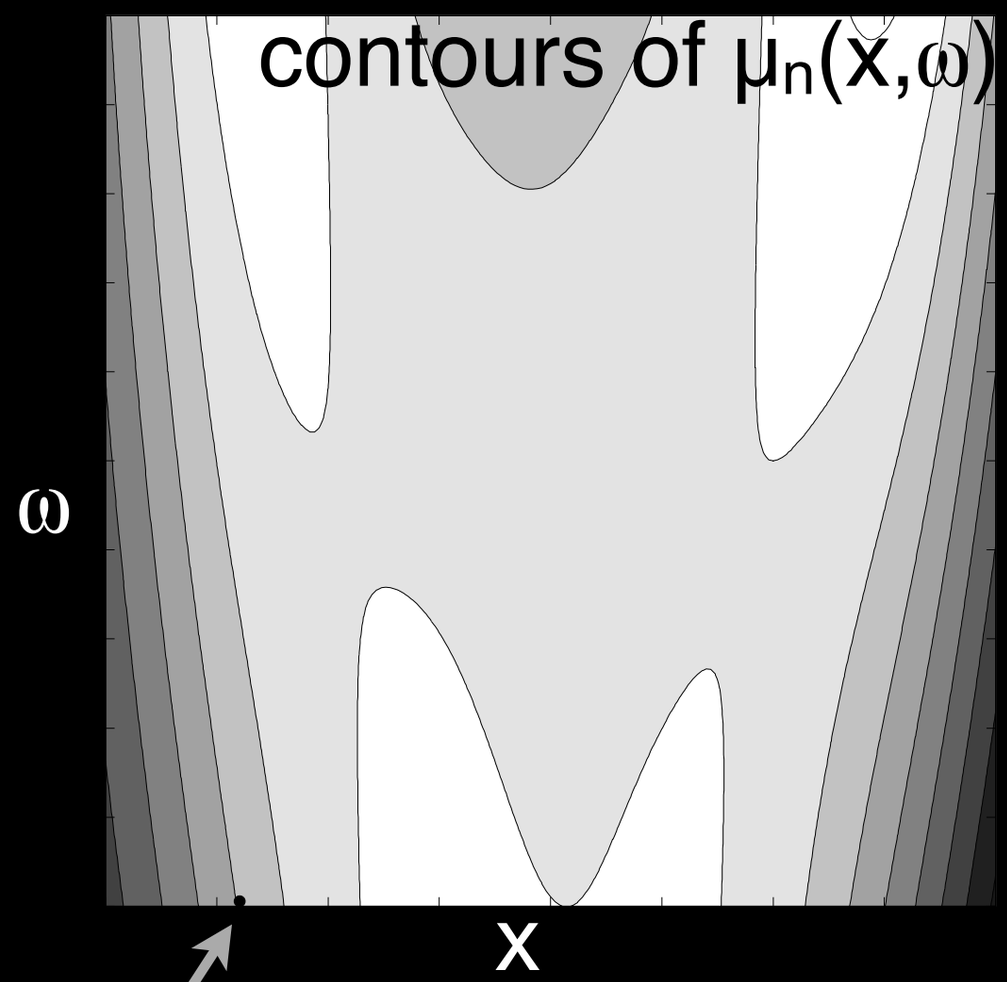
where we last sampled



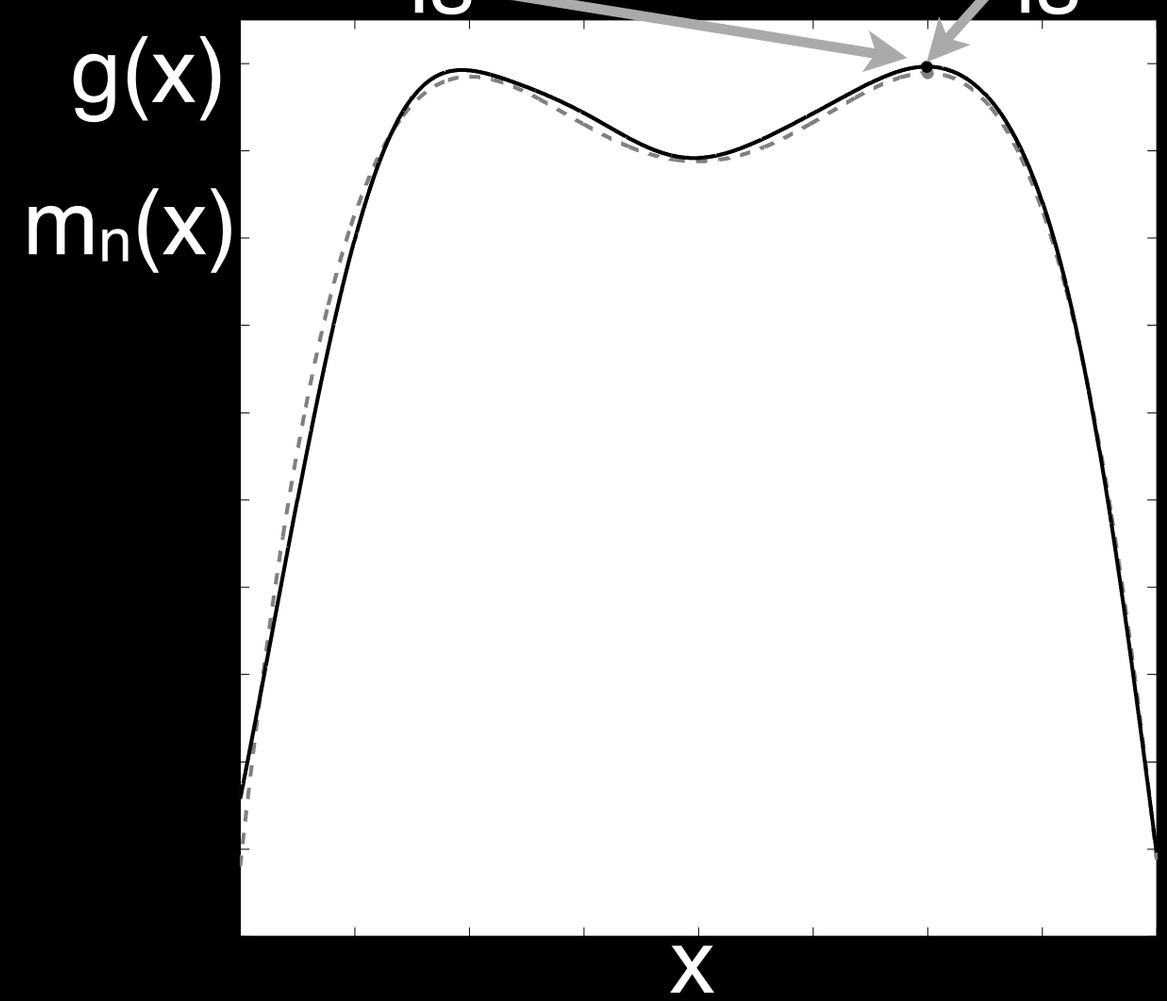
where we think the max is

where the max is

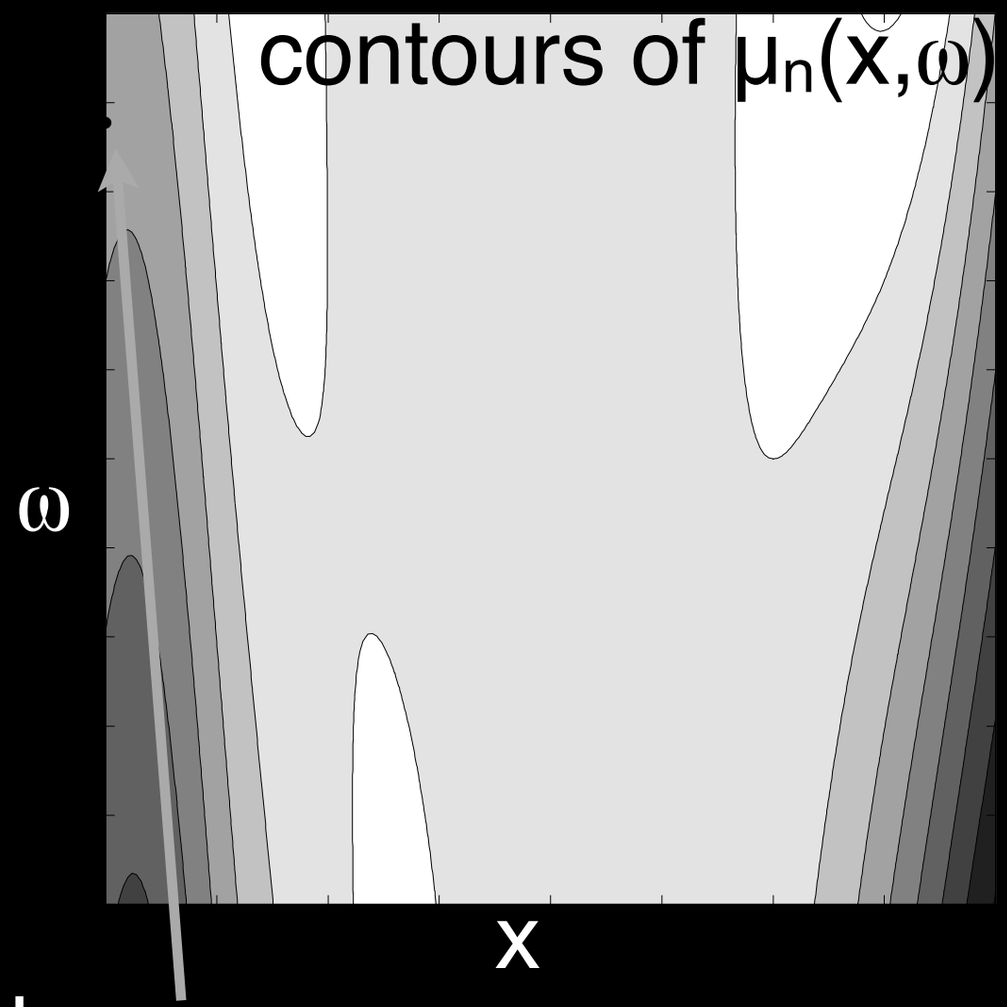
Here's the algorithm solving a test problem



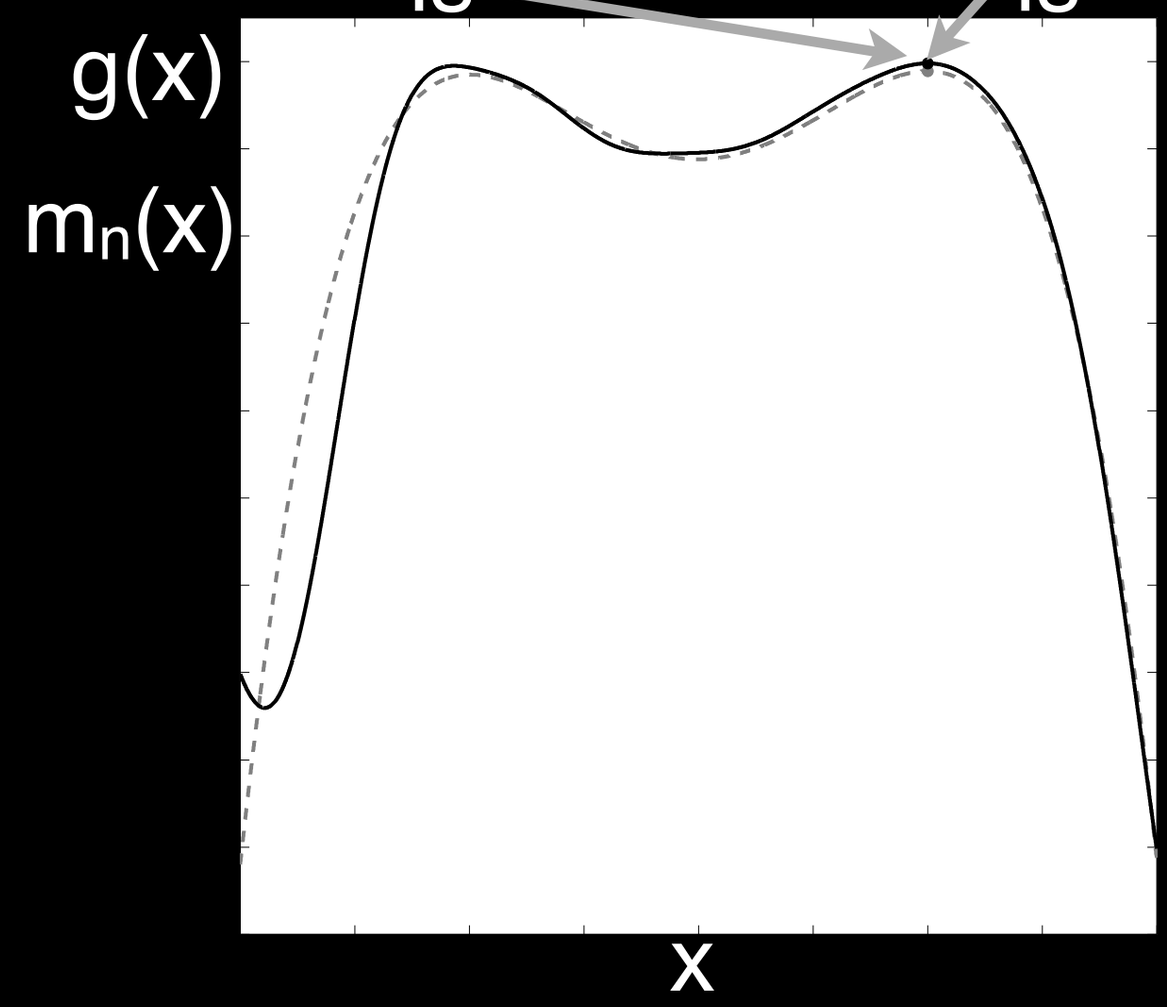
where we last sampled



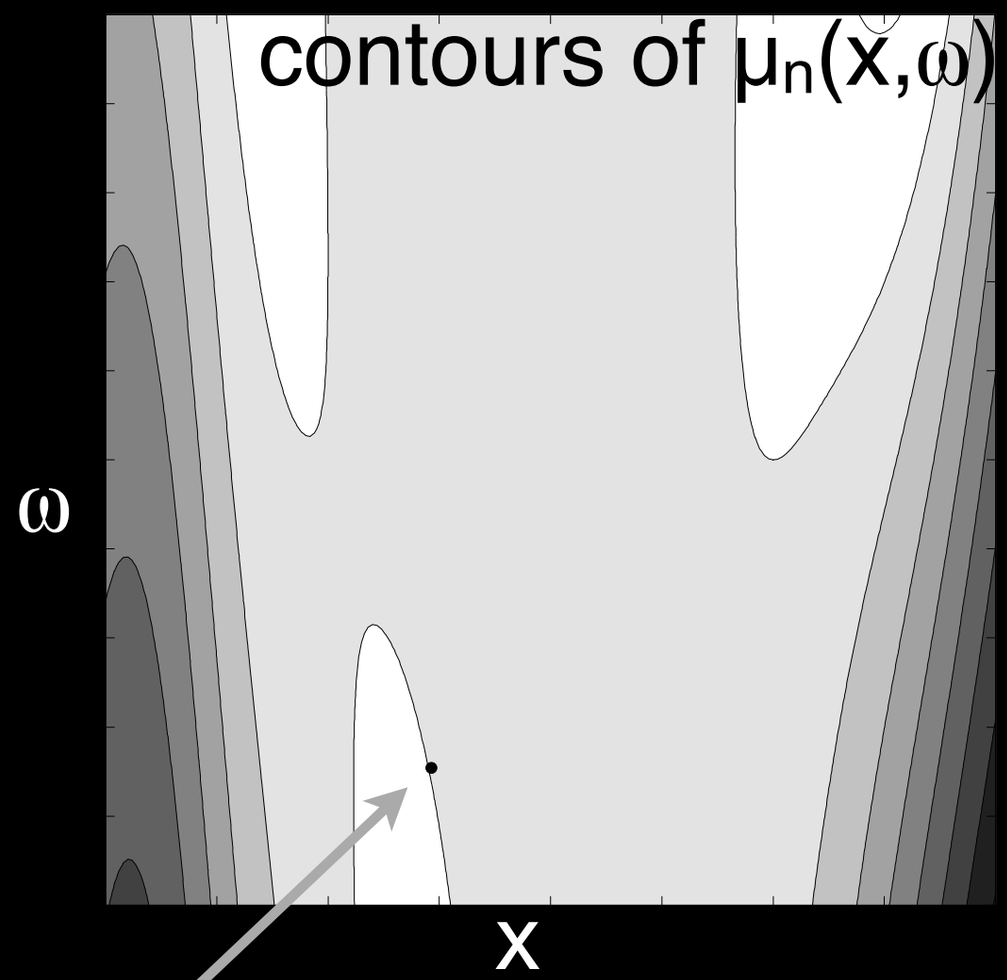
Here's the algorithm solving a test problem



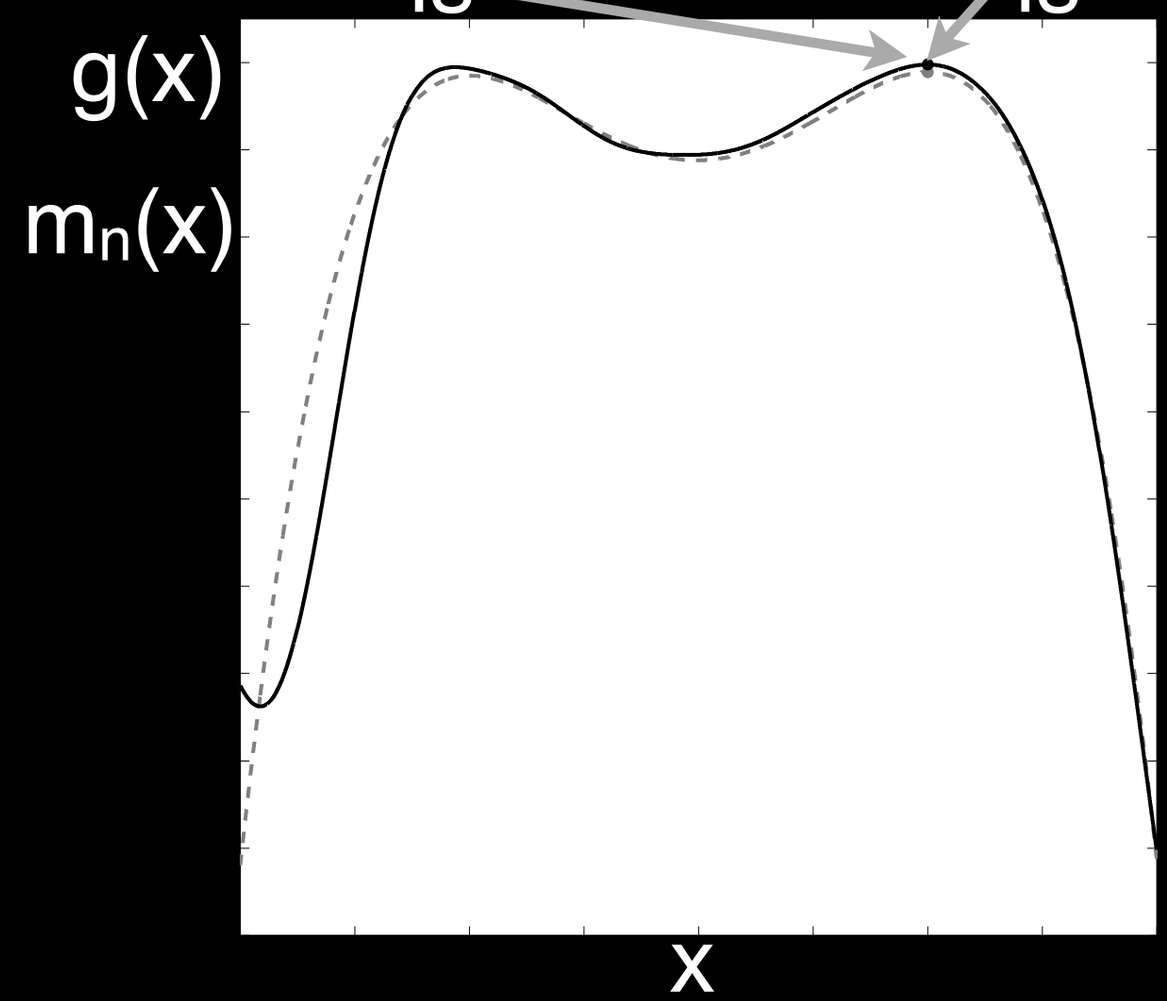
where we last sampled



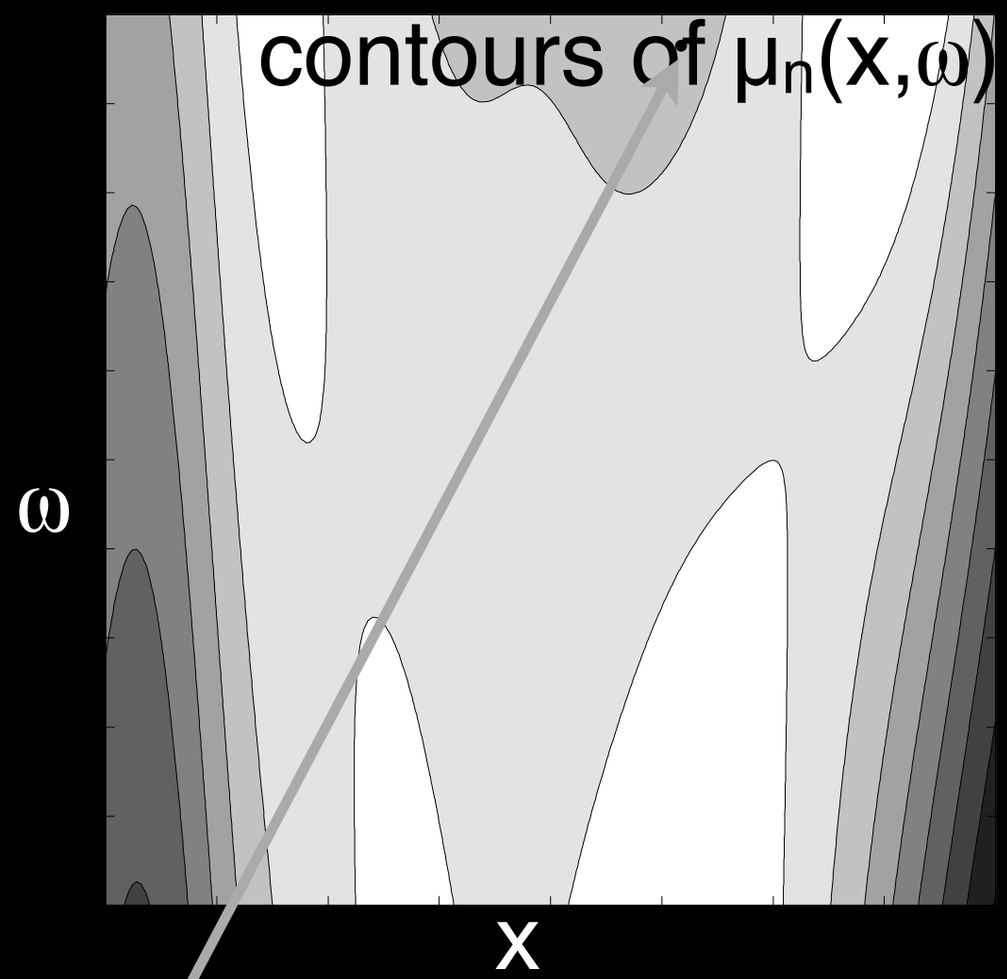
Here's the algorithm solving a test problem



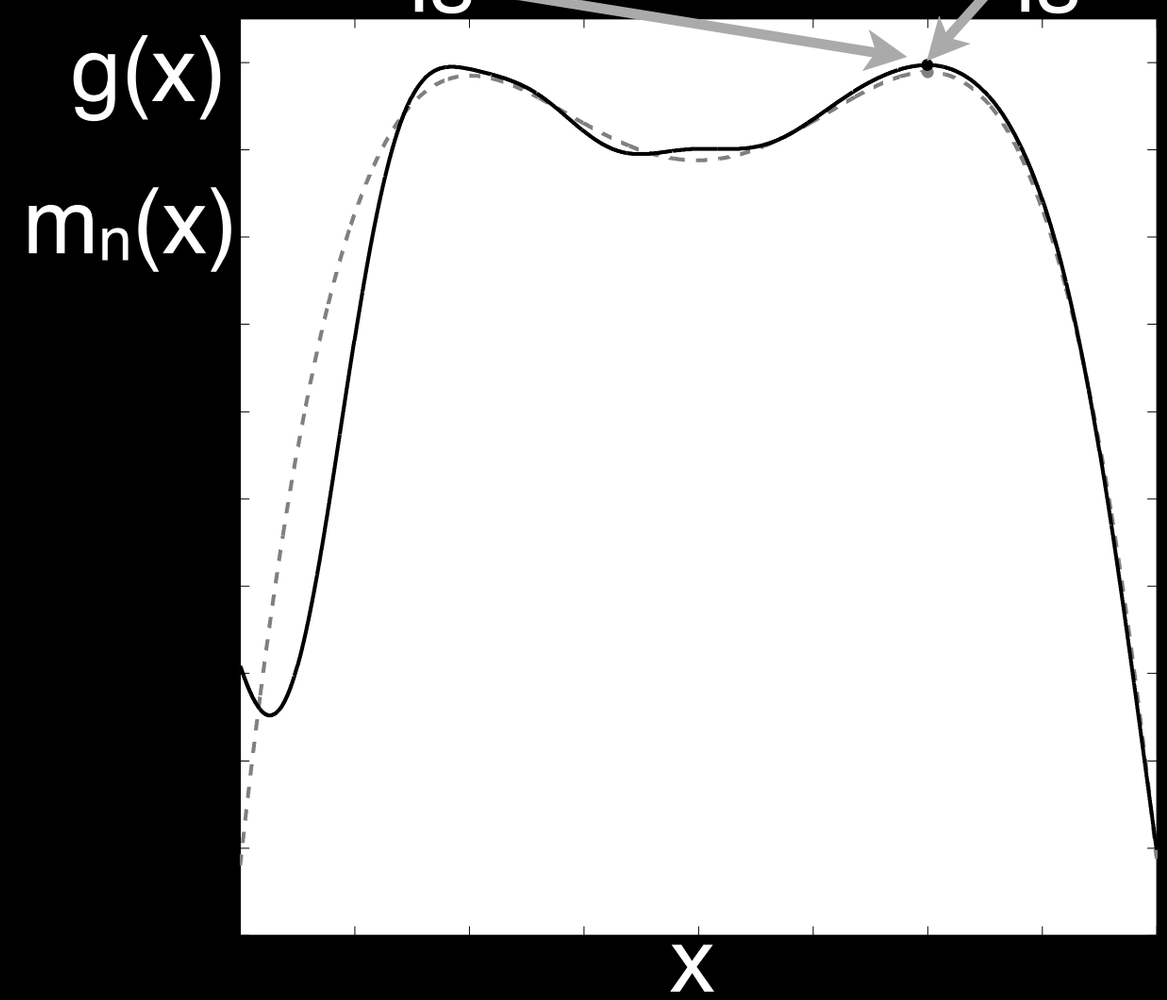
where we last sampled



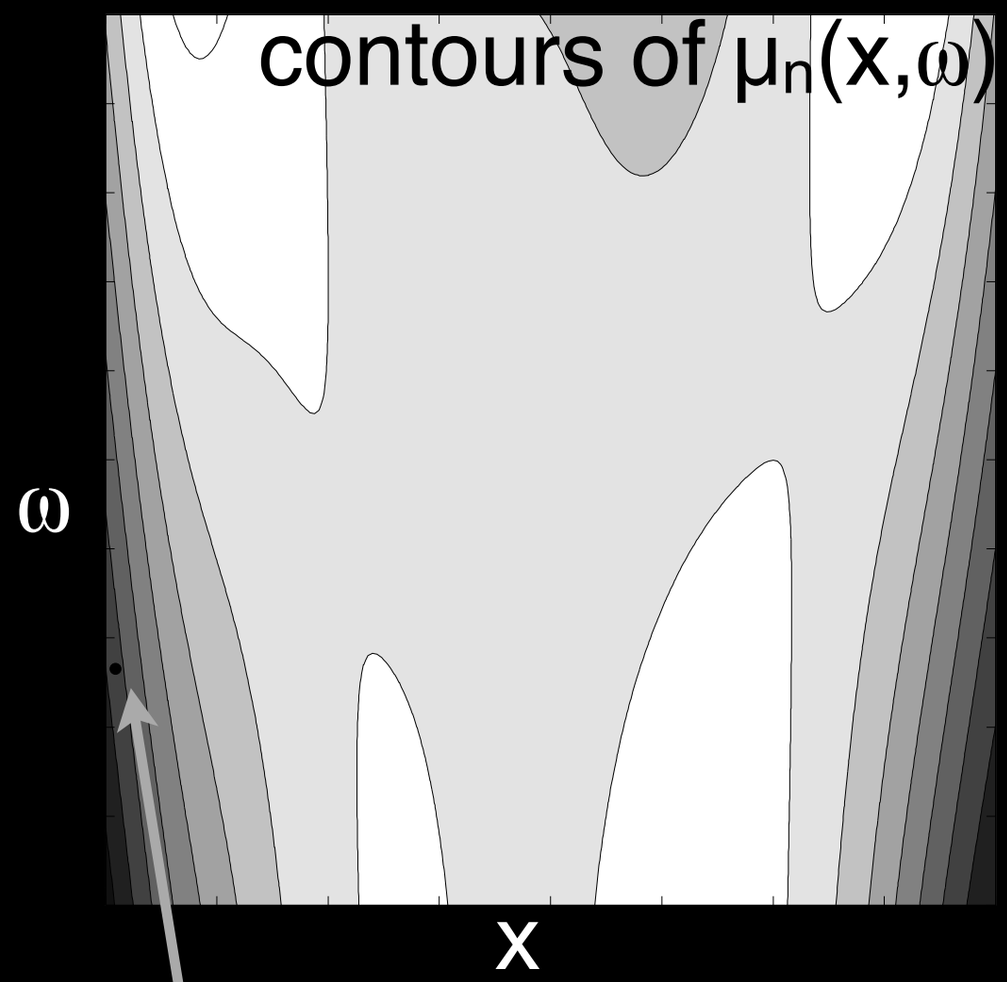
Here's the algorithm solving a test problem



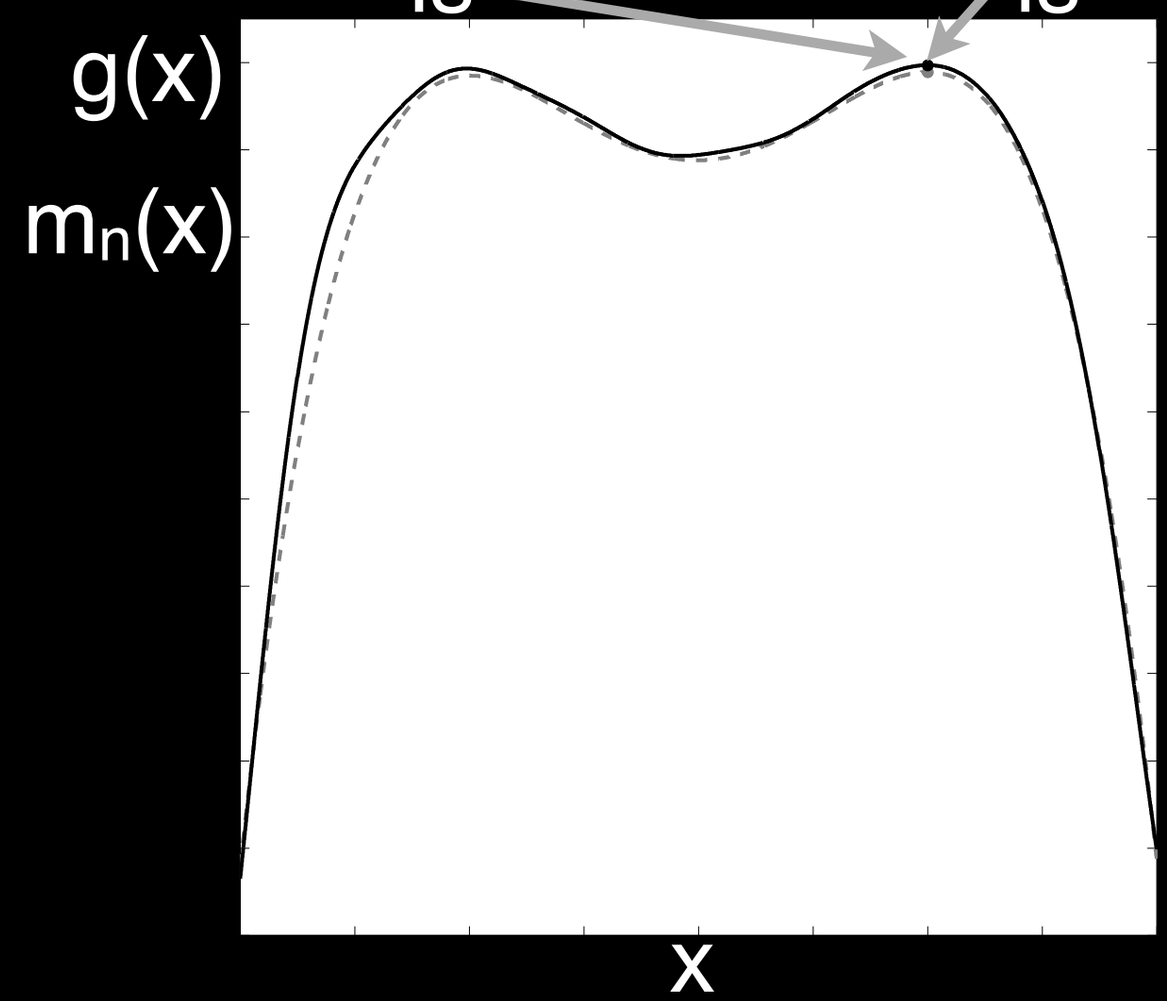
where we last sampled



Here's the algorithm solving a test problem



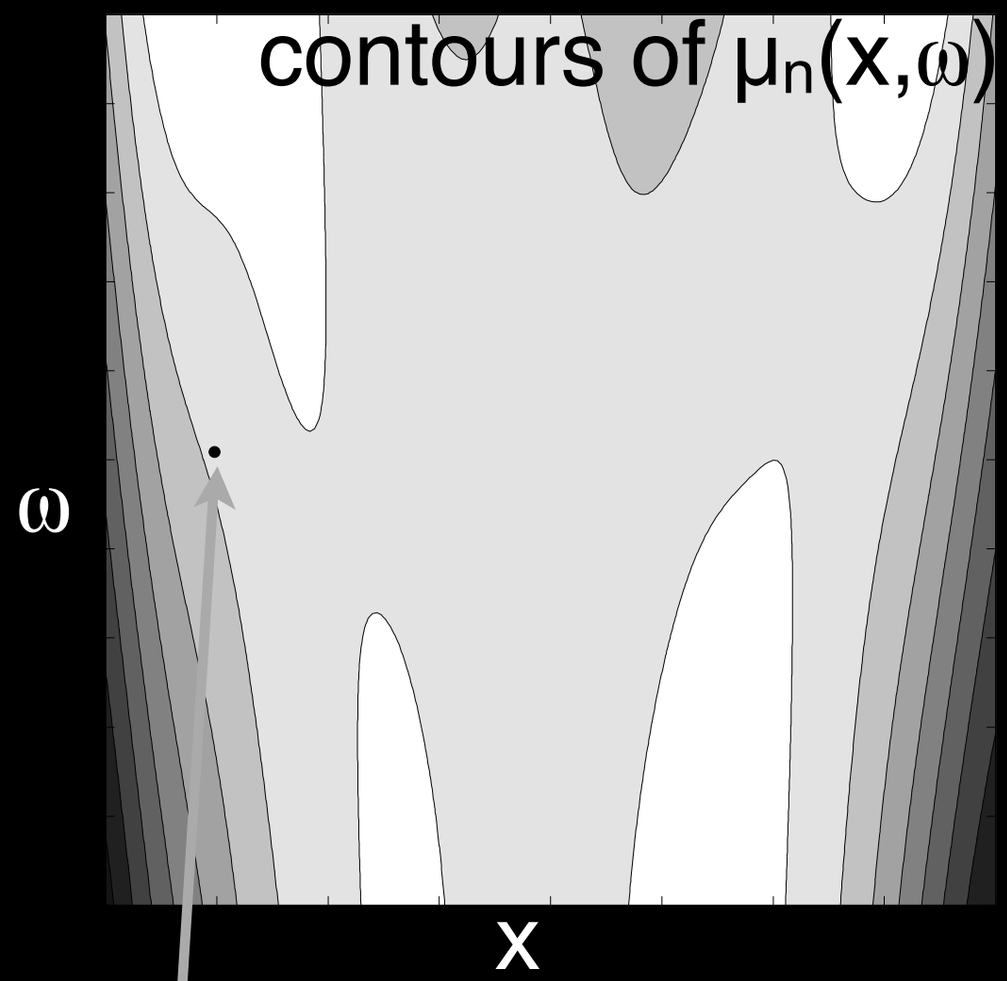
where we last sampled



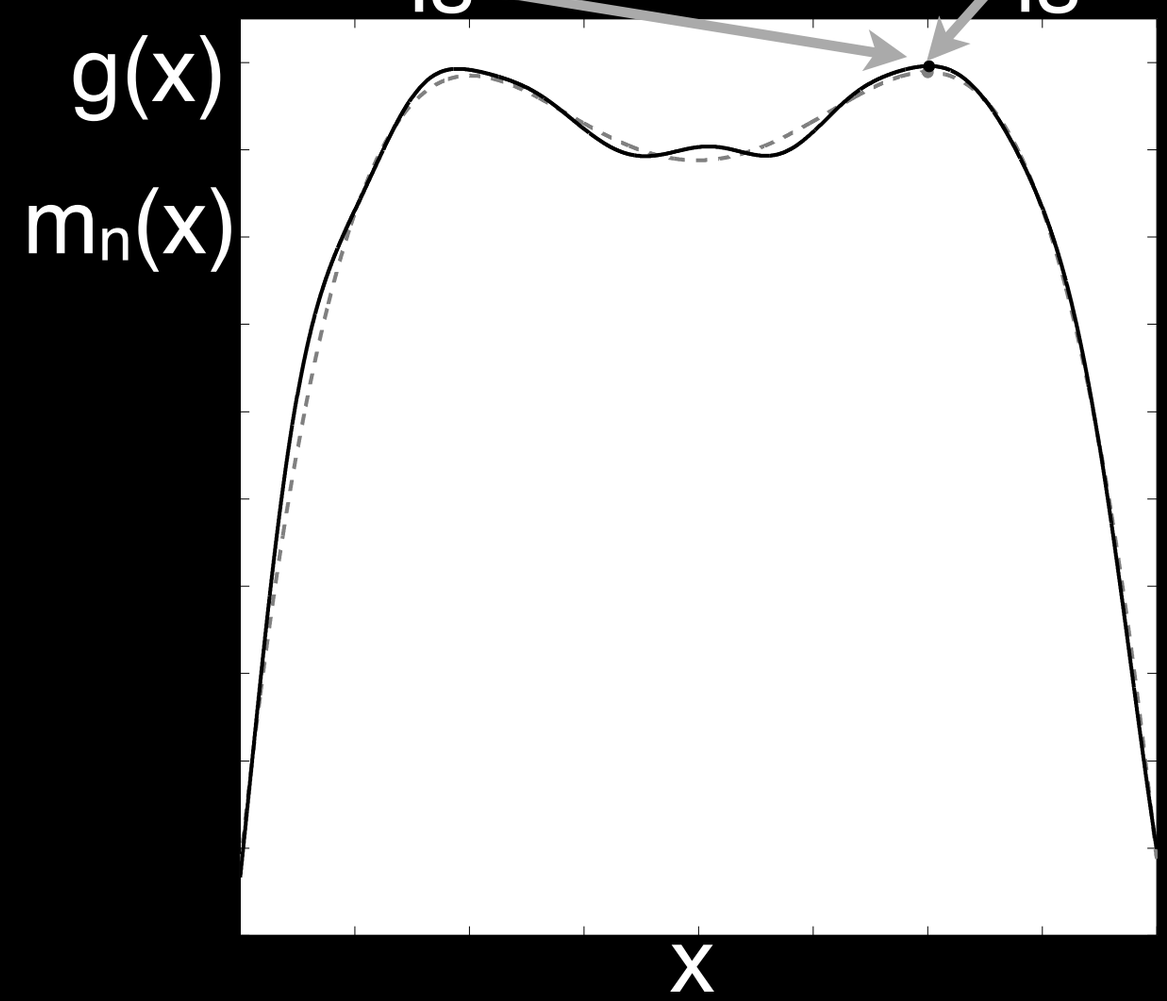
where we think the max is

where the max is

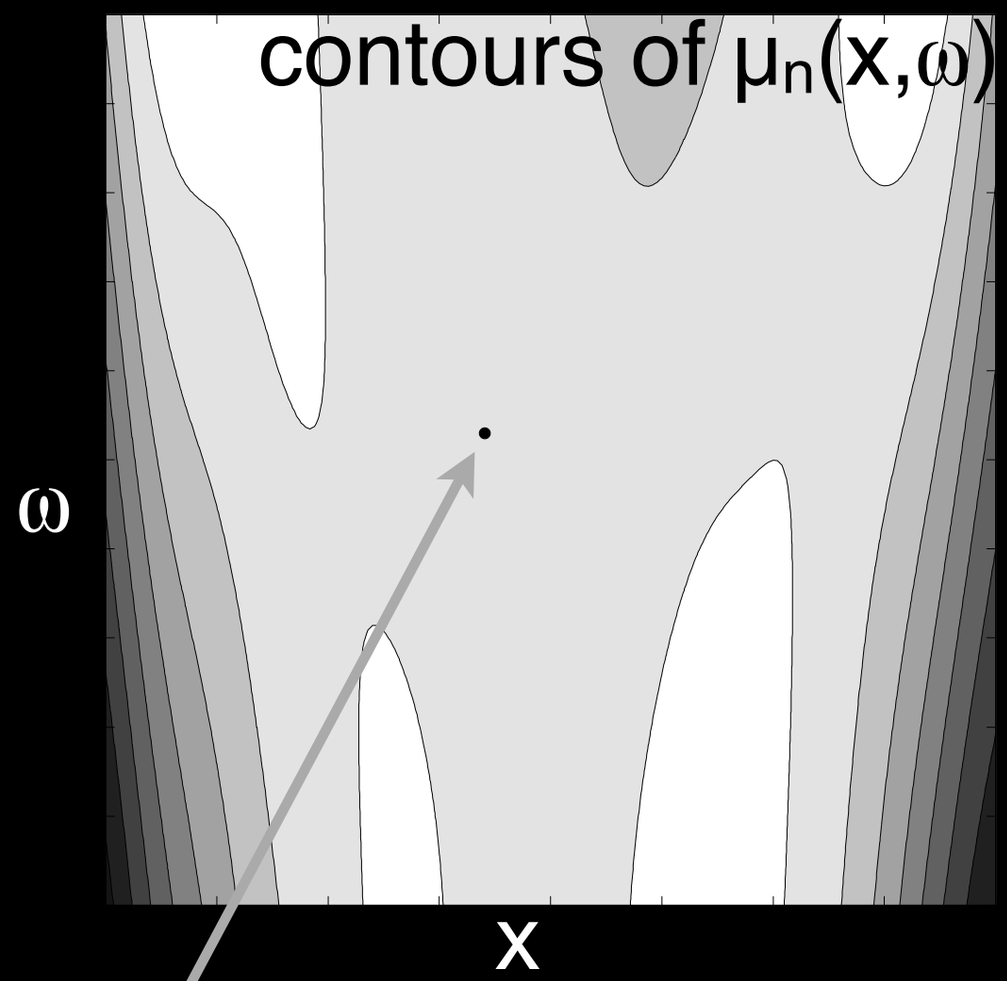
Here's the algorithm solving a test problem



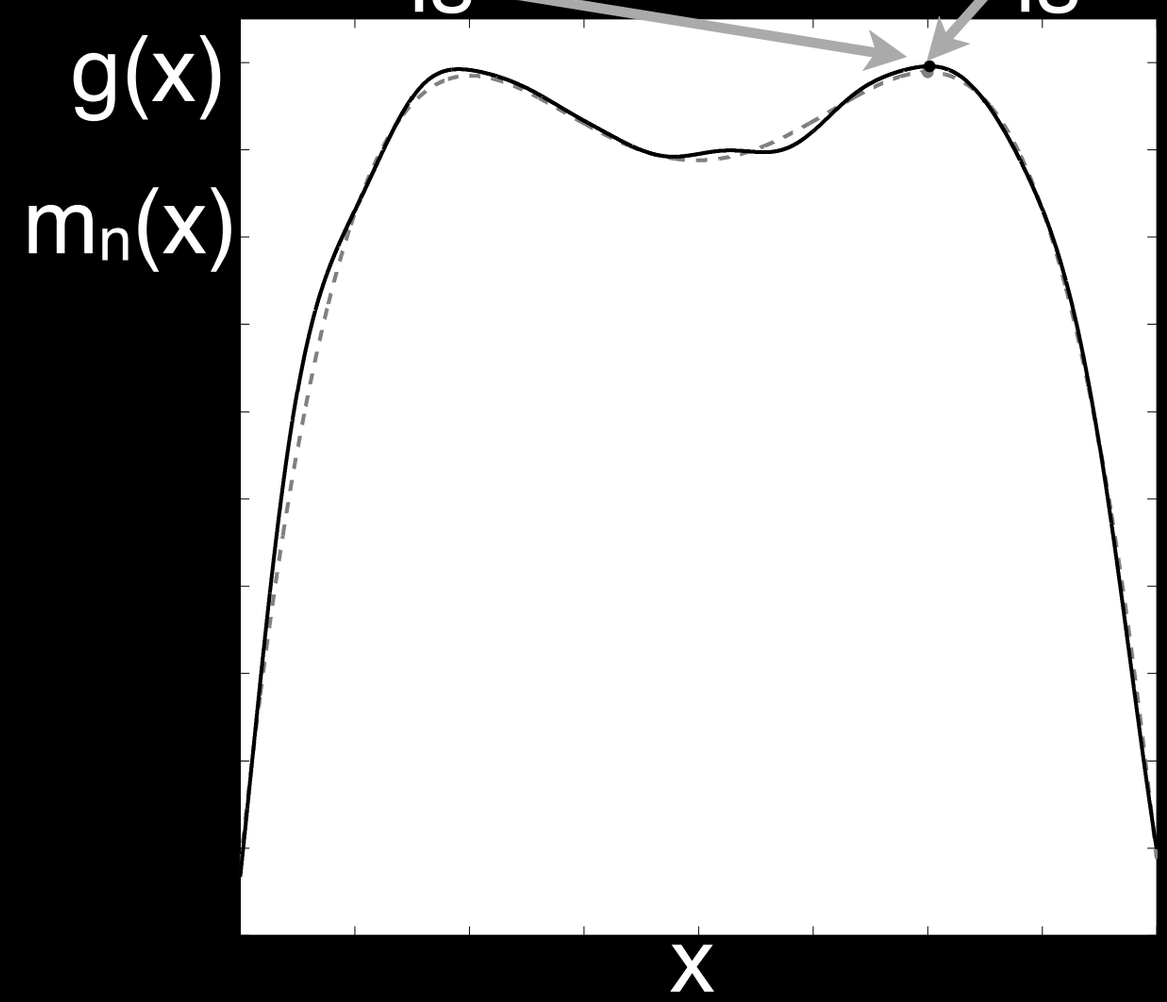
where we last sampled



Here's the algorithm solving a test problem



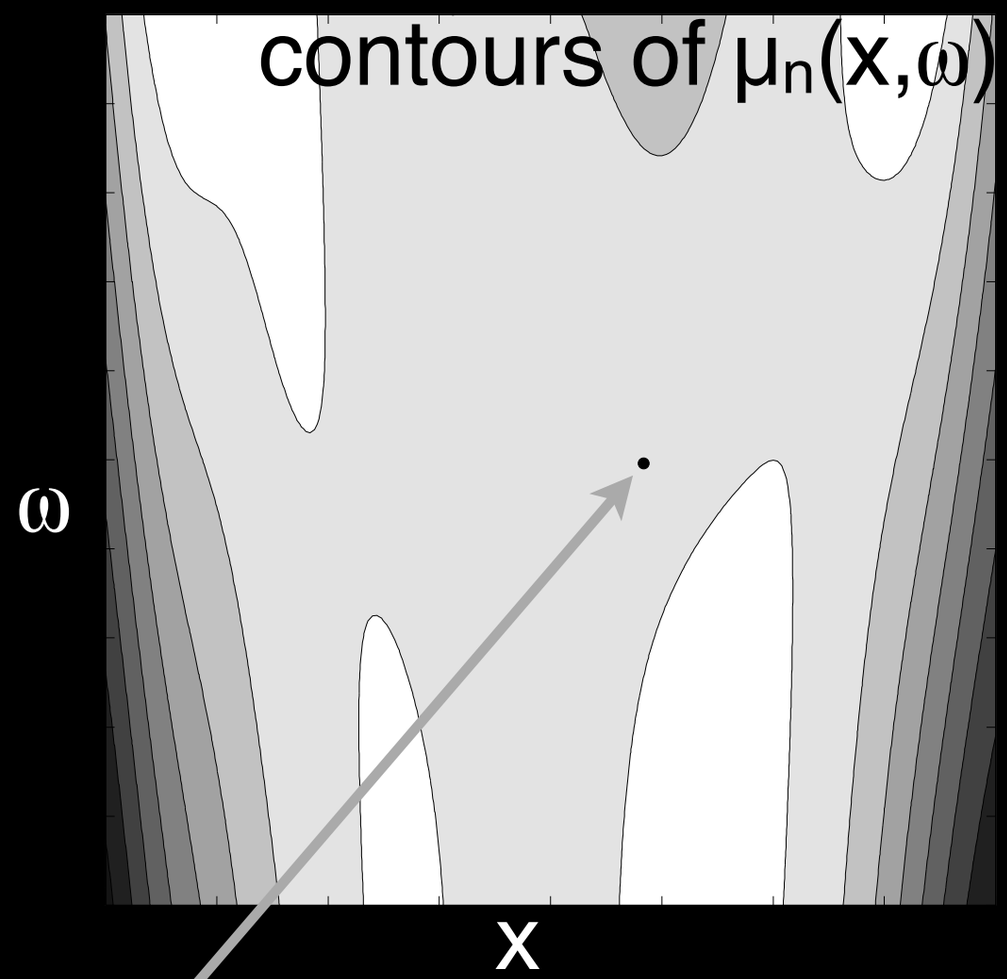
where we last sampled



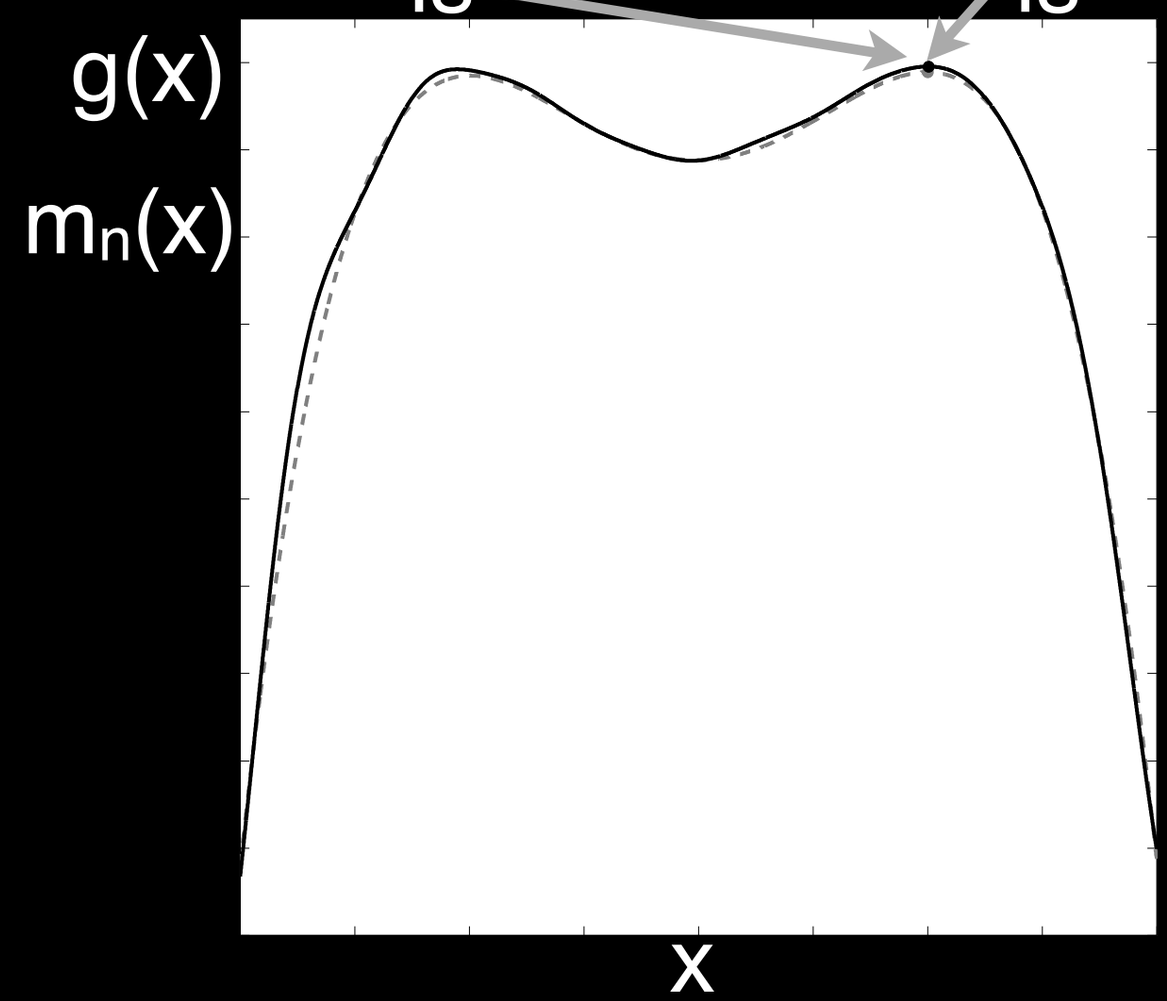
where we think the max is

where the max is

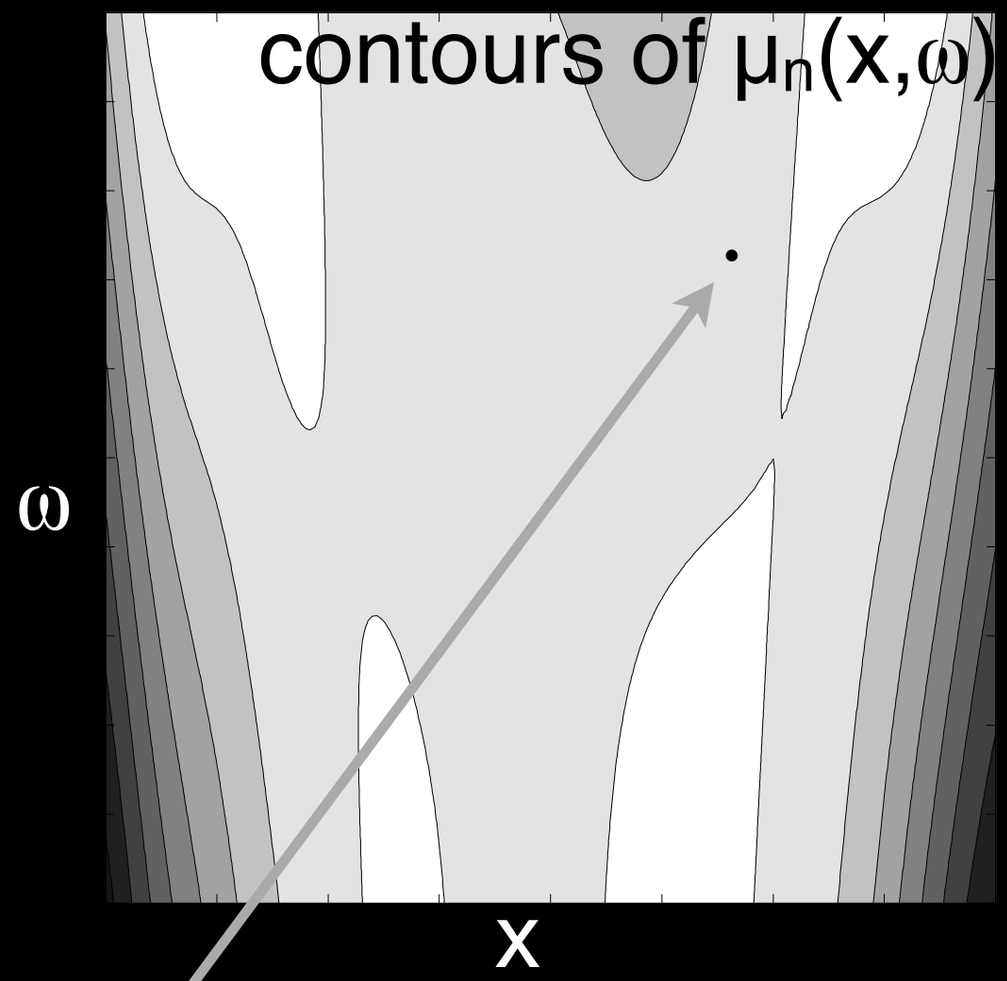
Here's the algorithm solving a test problem



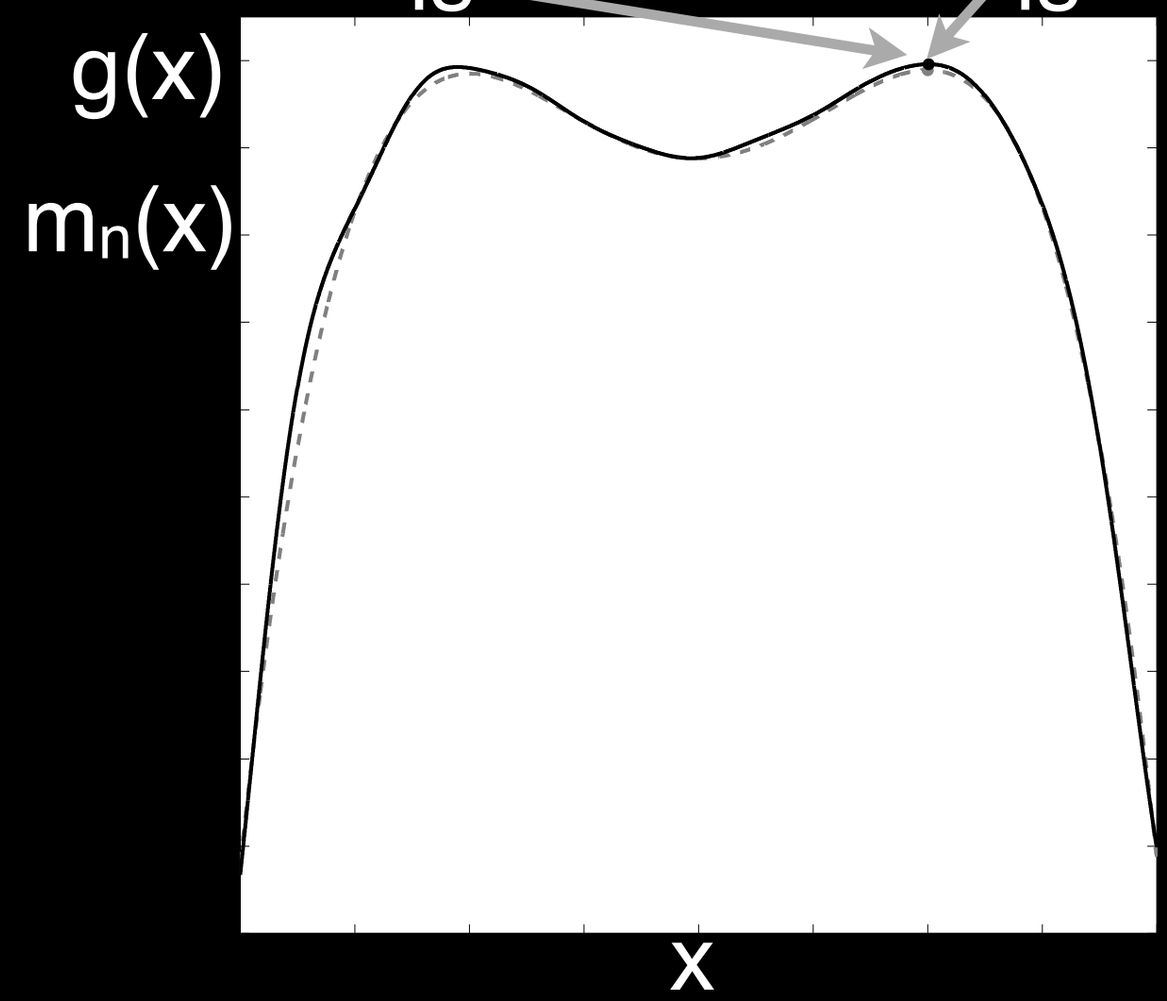
where we last sampled



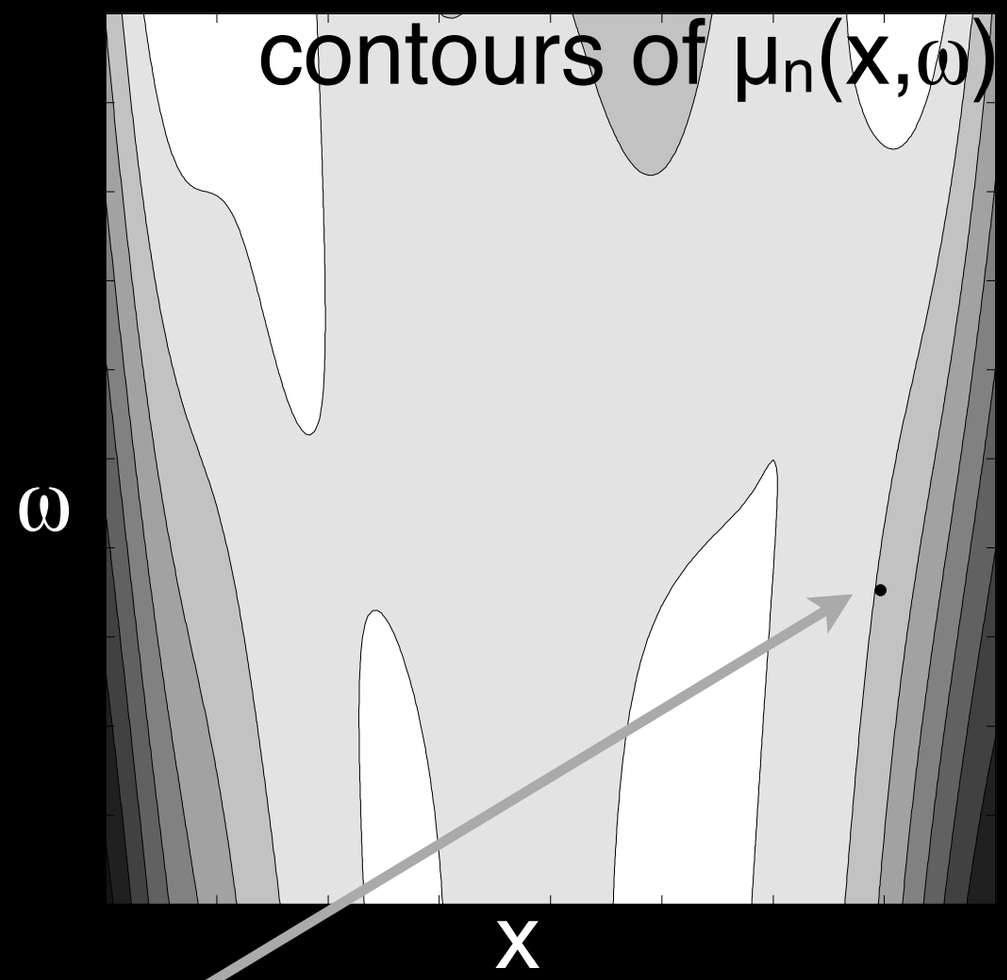
Here's the algorithm solving a test problem



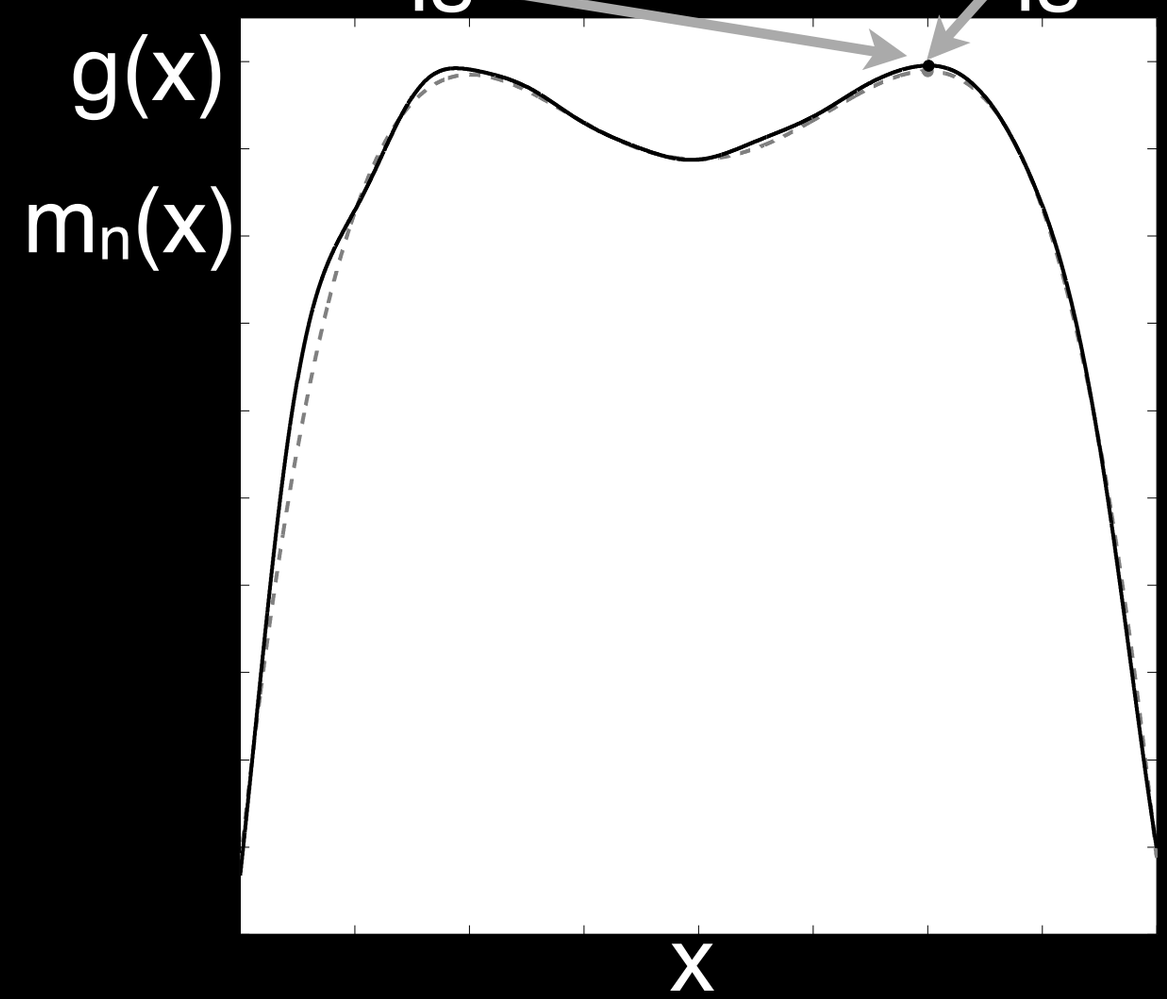
where we last sampled



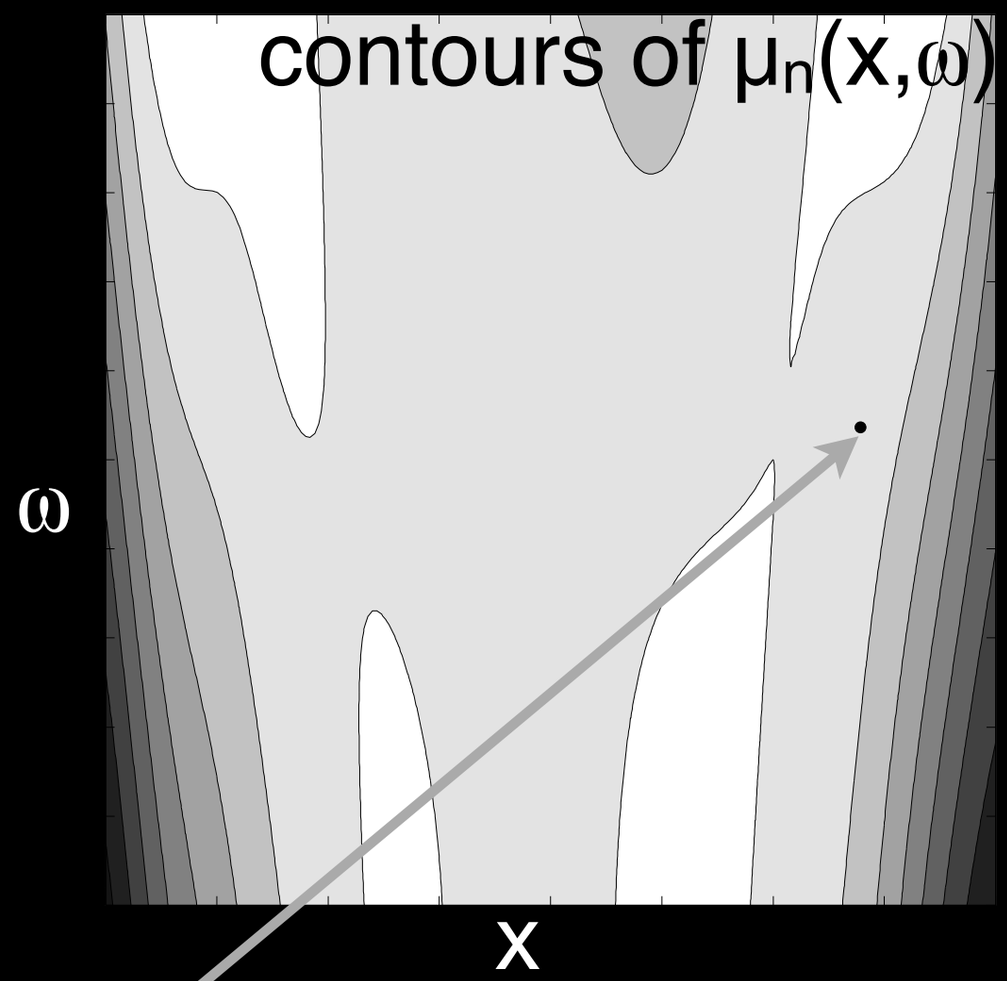
Here's the algorithm solving a test problem



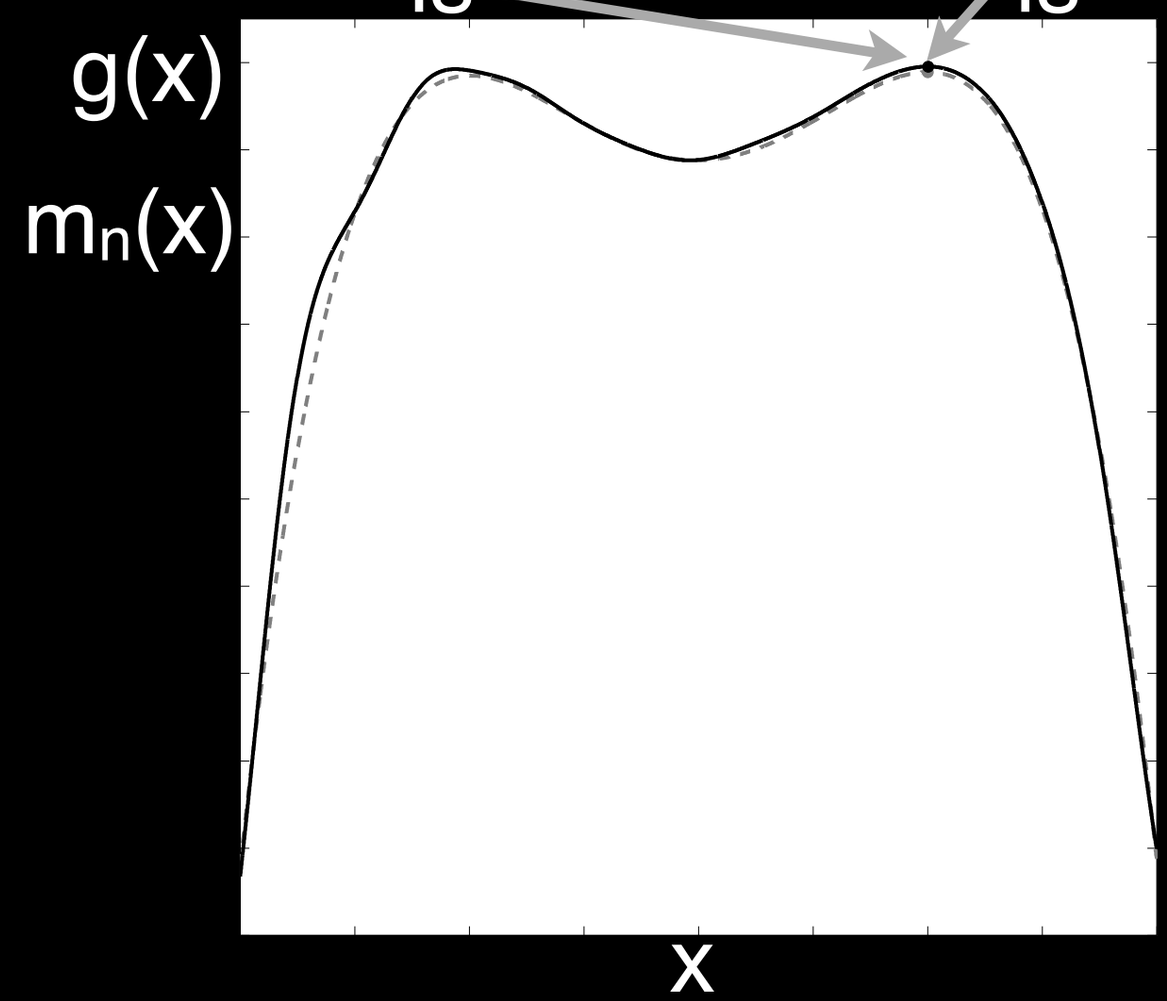
where we last sampled



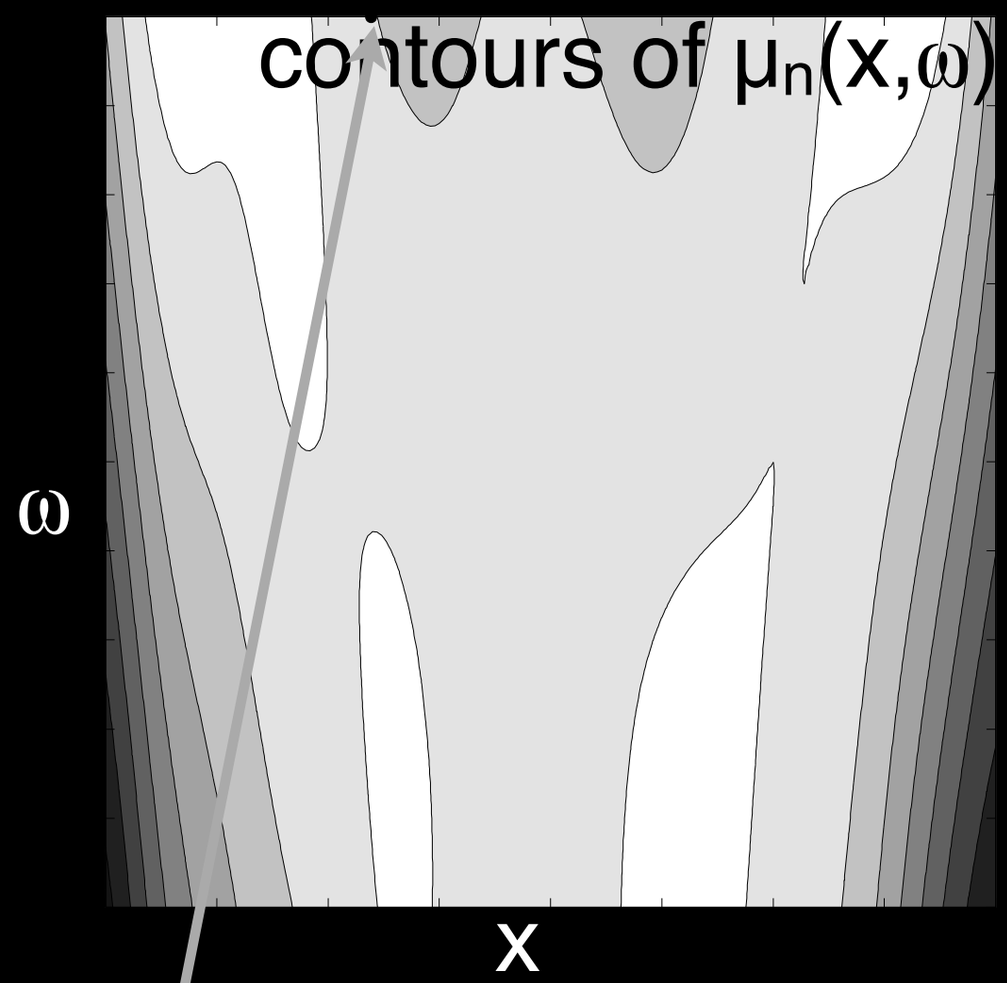
Here's the algorithm solving a test problem



where we last sampled



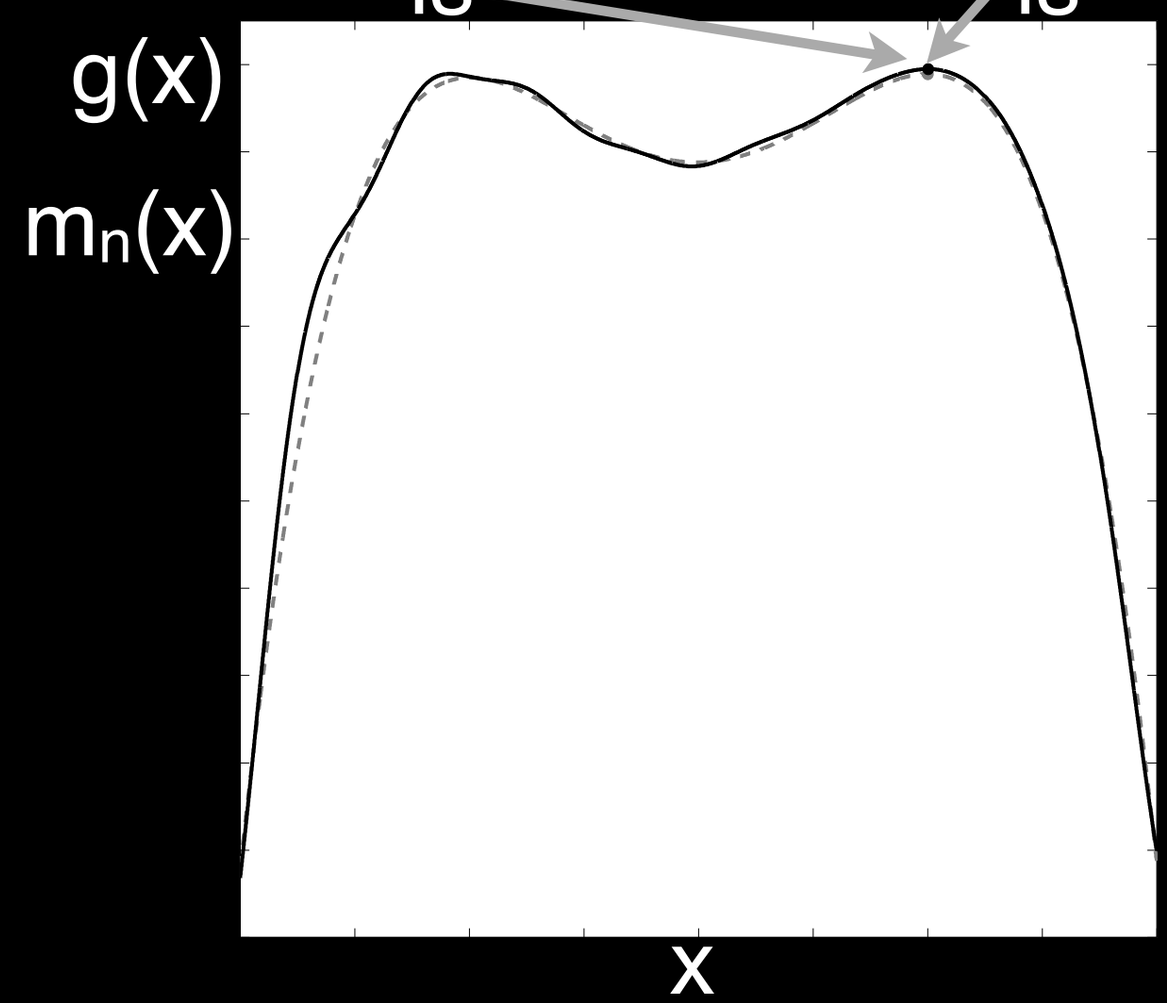
Here's the algorithm solving a test problem



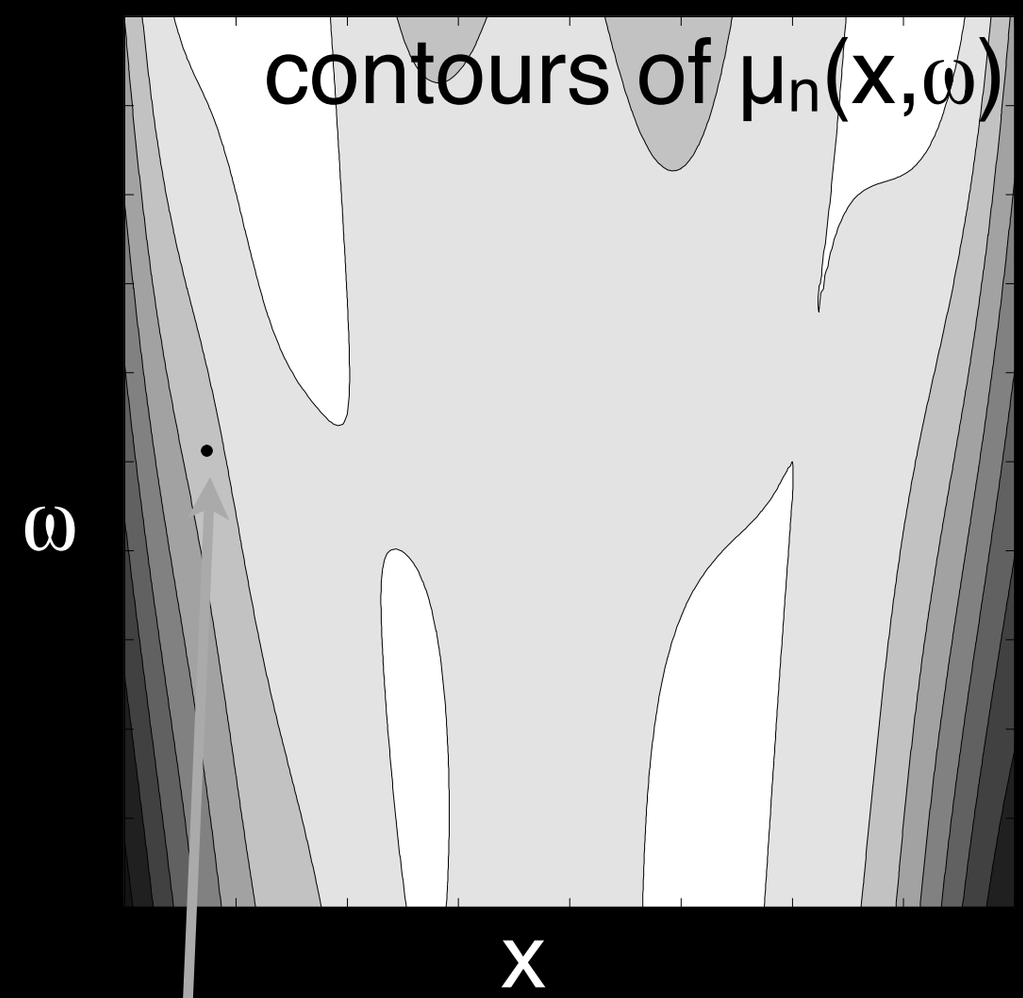
where we last sampled

where we think the max is

where the max is



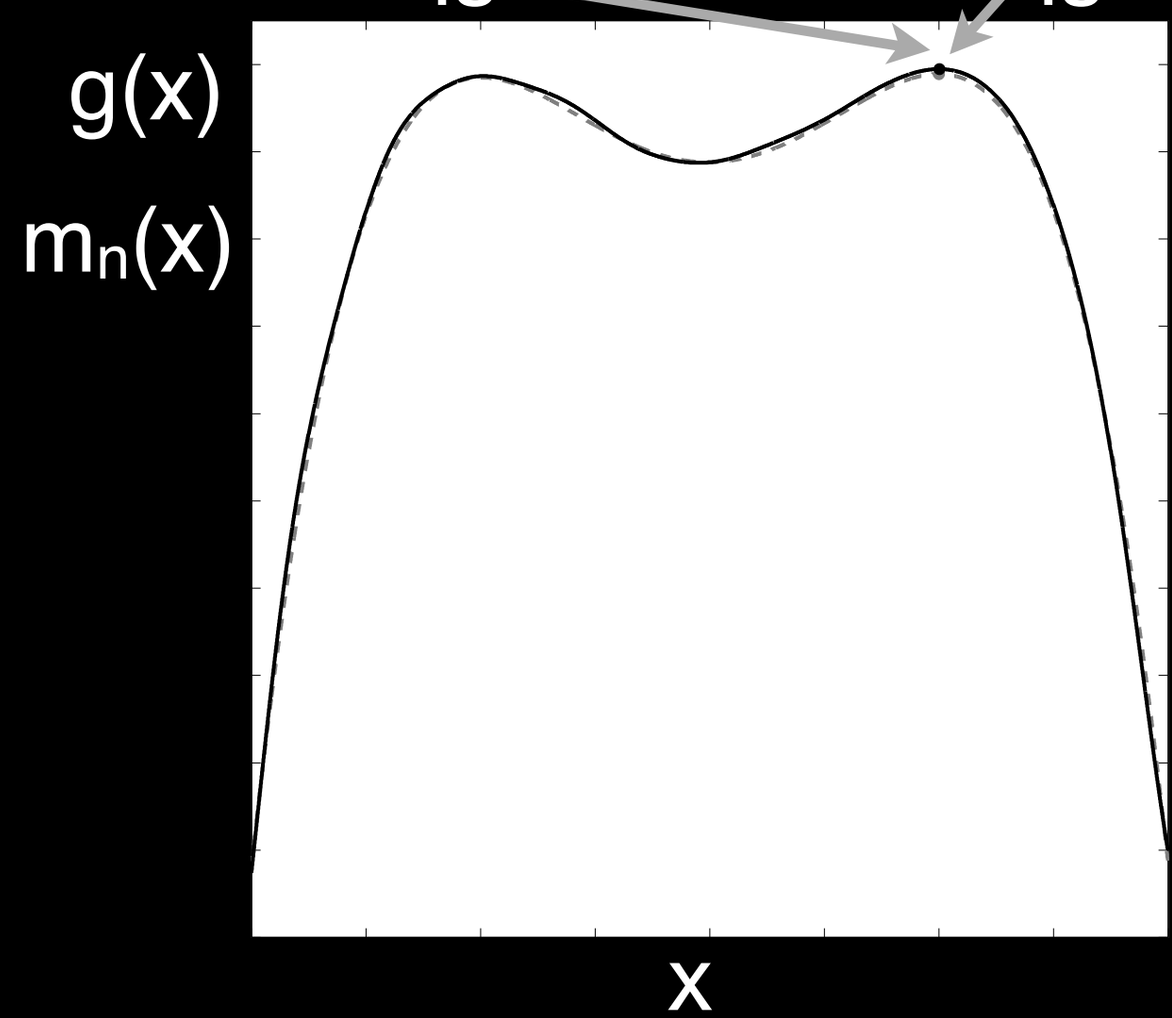
Here's the algorithm solving a test problem



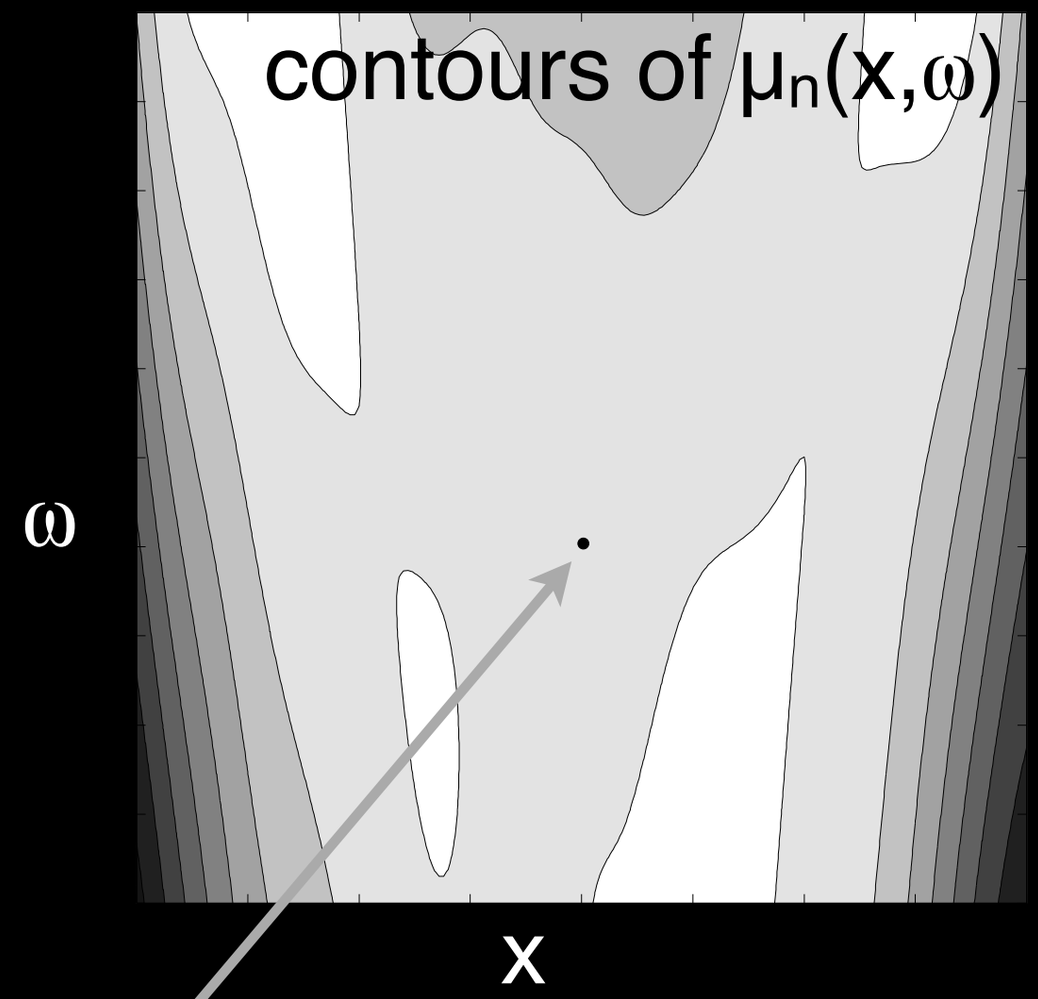
where we last sampled

where we think the max is

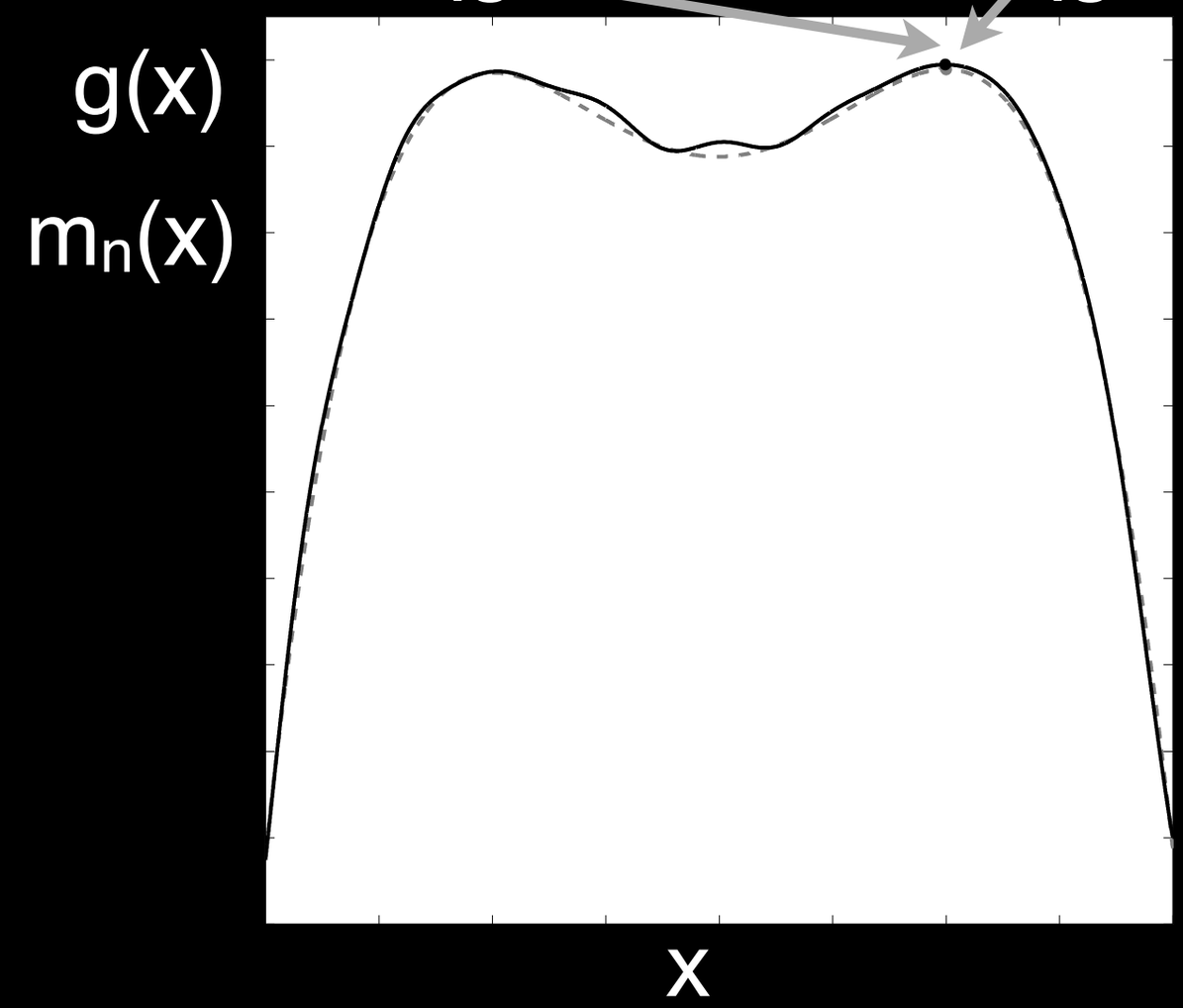
where the max is



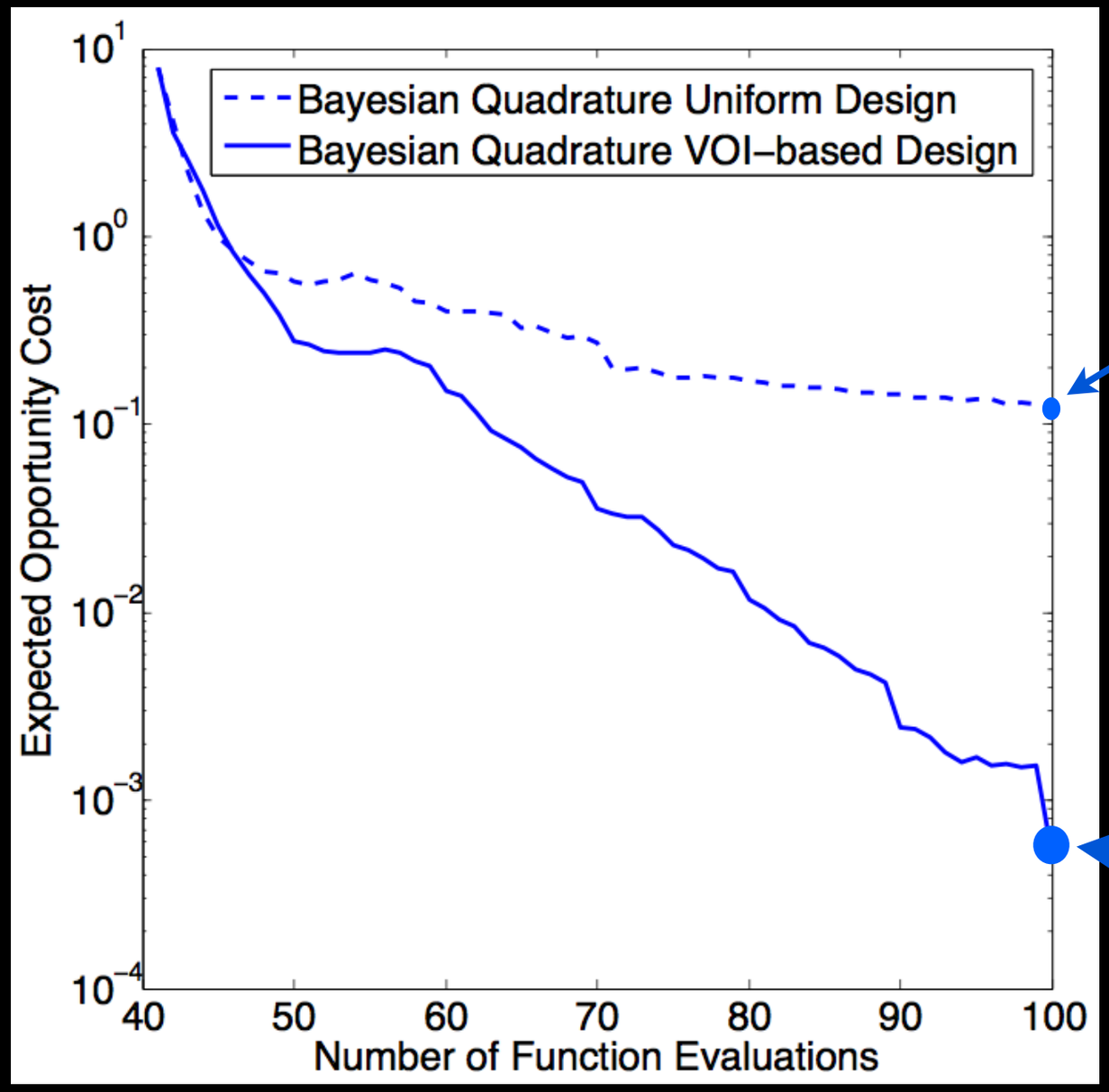
Here's the algorithm solving a test problem



where we
last sampled



The algorithm works well on a test problem

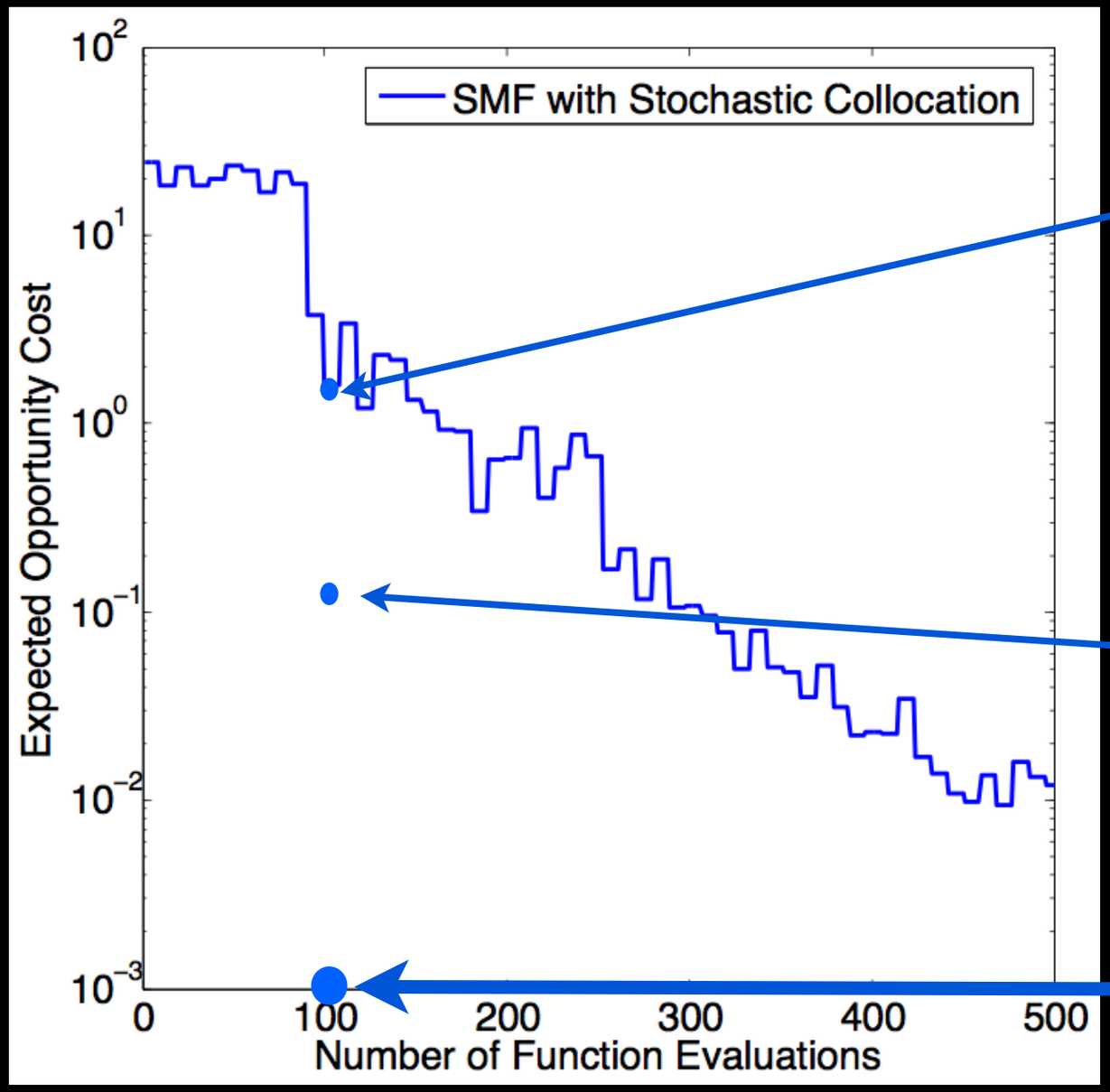


In 100 function evaluations:

A simple benchmark (sampling at random, and using Bayesian quadrature) achieves accuracy of 10^{-1}.

Our algorithm achieves accuracy better than 10^{-3}.

The algorithm works well on a test problem



In 100 function evaluations:

A benchmark from the literature (Surrogate management framework with stochastic collocation [Sankaran and Marsden, 2011]) achieves accuracy of 10^{0}.

The simple benchmark (sampling at random, and using Bayesian quadrature) achieves accuracy of 10^{-1}.

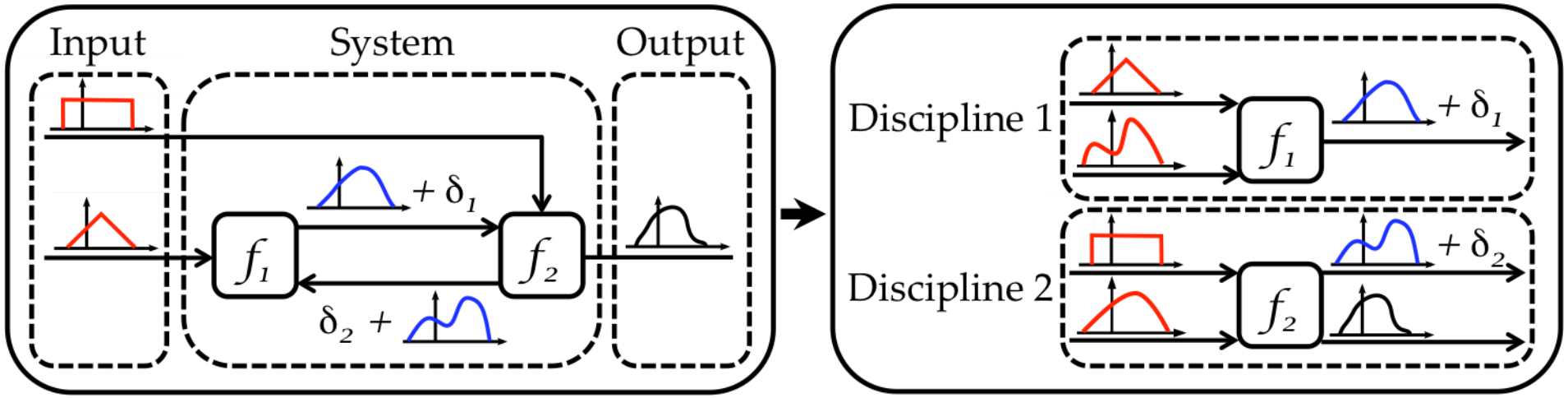
Our algorithm achieves accuracy better than 10^{-3}.

Some other recent progress

- Compositional Uncertainty Quantification
- Stack MC

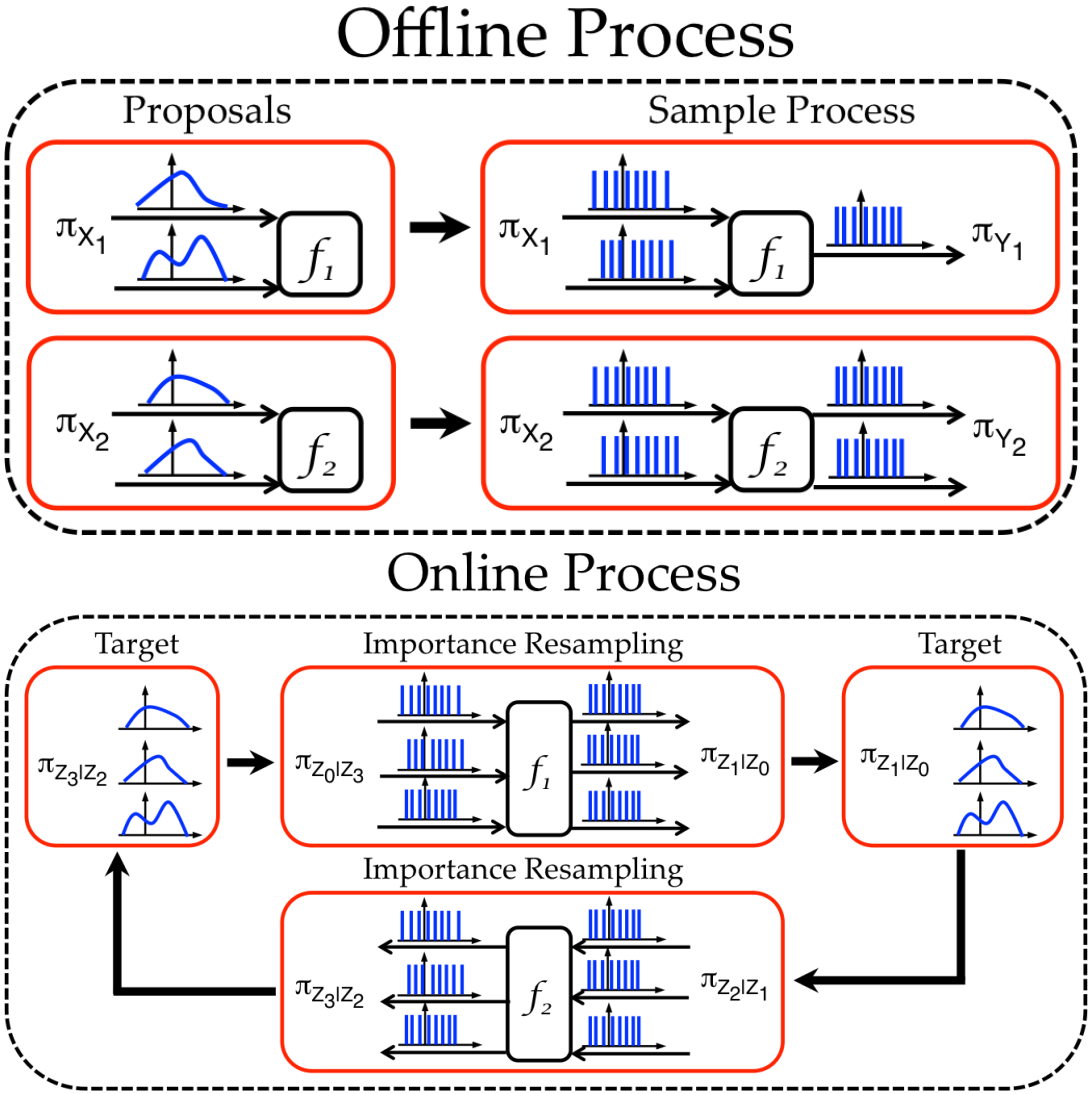
Compositional Uncertainty Quantification

- Developing a compositional uncertainty analysis methodology for coupled, multidisciplinary systems
  - This capability enables the efficient assessment of the expected value of information of introducing a new information source or sampling in a different location in the design space
  - This capability enables rapid uncertainty quantification updates when a new information source is incorporated
  - We are currently considering model discrepancy and parametric uncertainty
  - Builds off the work of Amaral, Allaire, and Willcox, 2014.



Compositional Uncertainty Quantification

- Offline
  - Sample information sources *independently* (this can also be previously generated data sets from other uses of a given information source)
- Online
  - Synthesize information source input and output samples so as to ensure correct dependency among the variables is captured
  - Key ingredients: Gibbs sampling, sequential importance resampling, density estimation
  - Developing adaptive sampling algorithm to combat potential sample impoverishment during the Gibbs sampling iterations



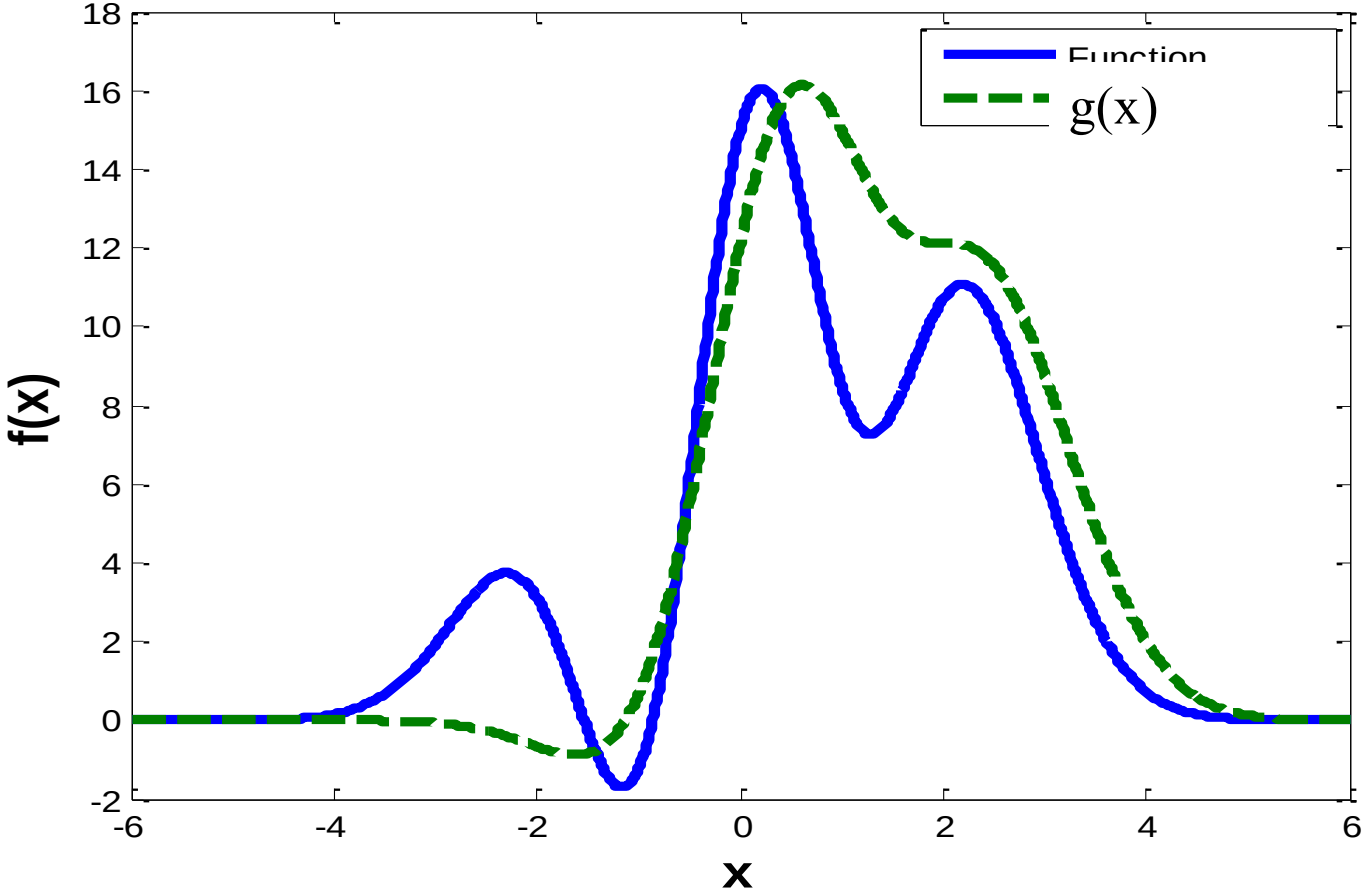
(Ghoreishi and Allaire, 2015)

Key aspects: Full system level evaluations are not necessary
Few online single information source evaluations may be required

Stack MC

1) We want to estimate the integral E[f] via MC.

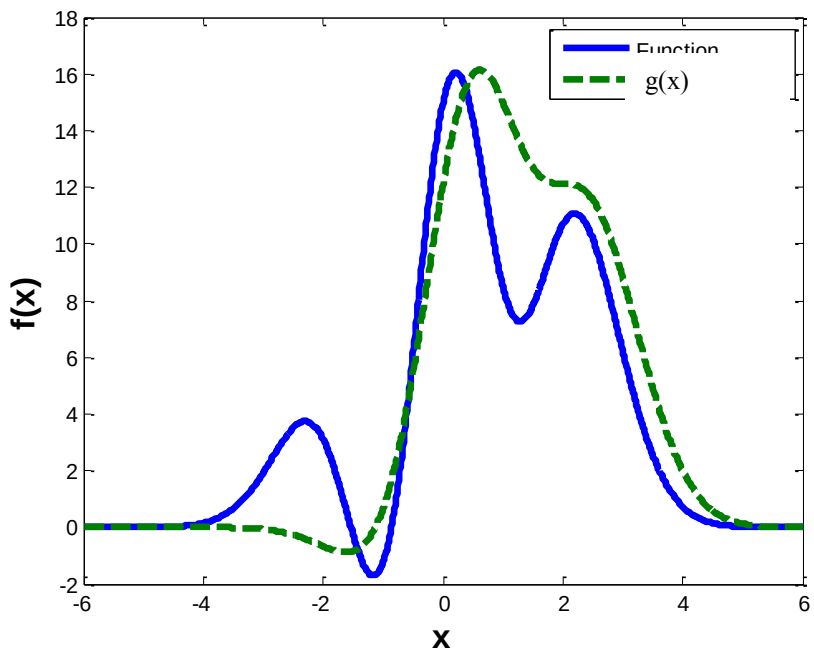
2) Suppose we had some g(x) that we can integrate in closed form, and that closely follows f(x). [choosing g(x) is discussed later]

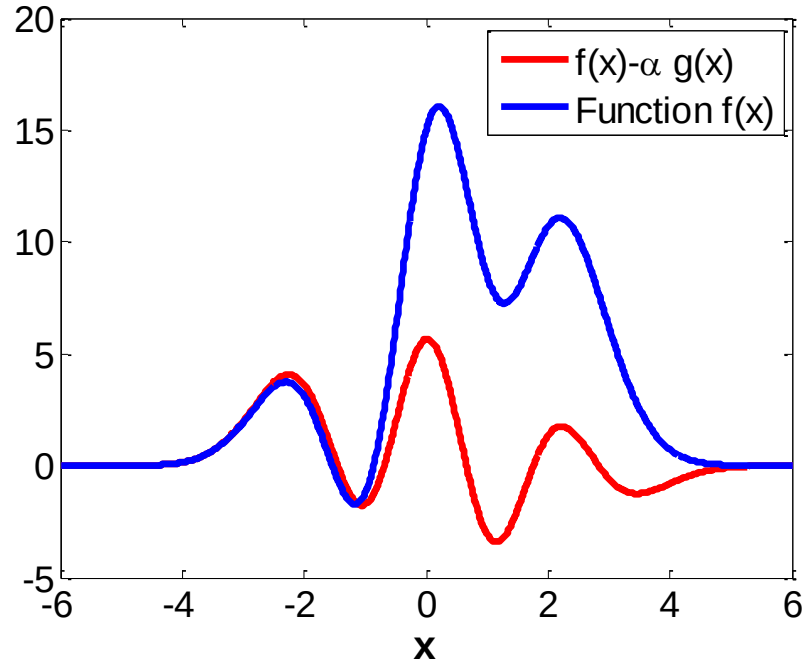


Stack MC

1) f(x) – g(x) varies less than f(x).

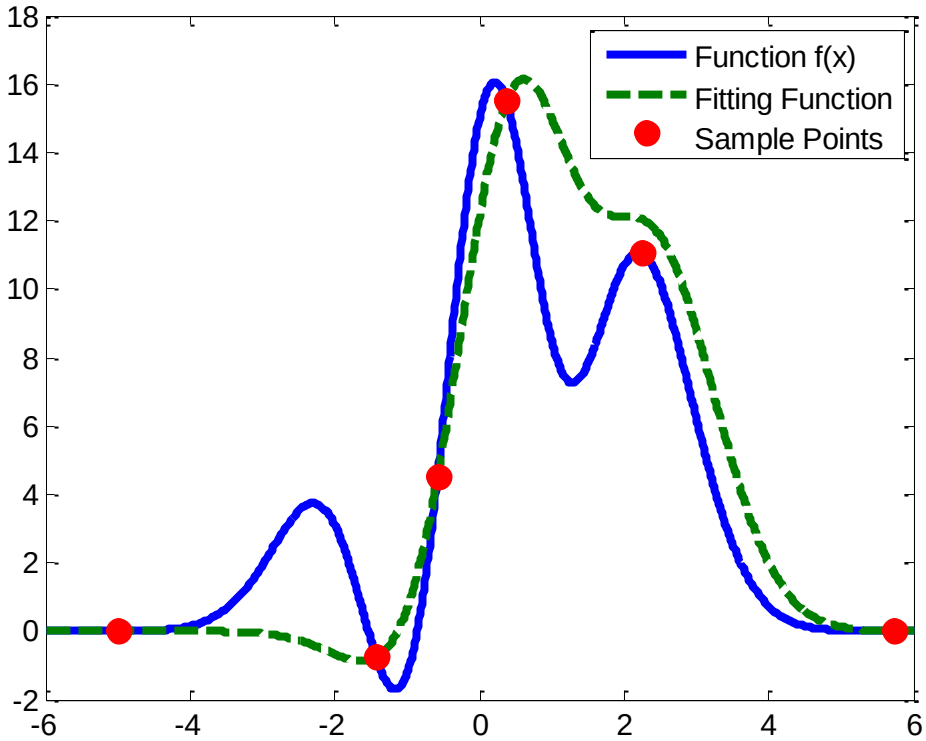
2) So use MC to estimate E[f – g], and add that estimate to the (known) E(g) to get the final estimate.



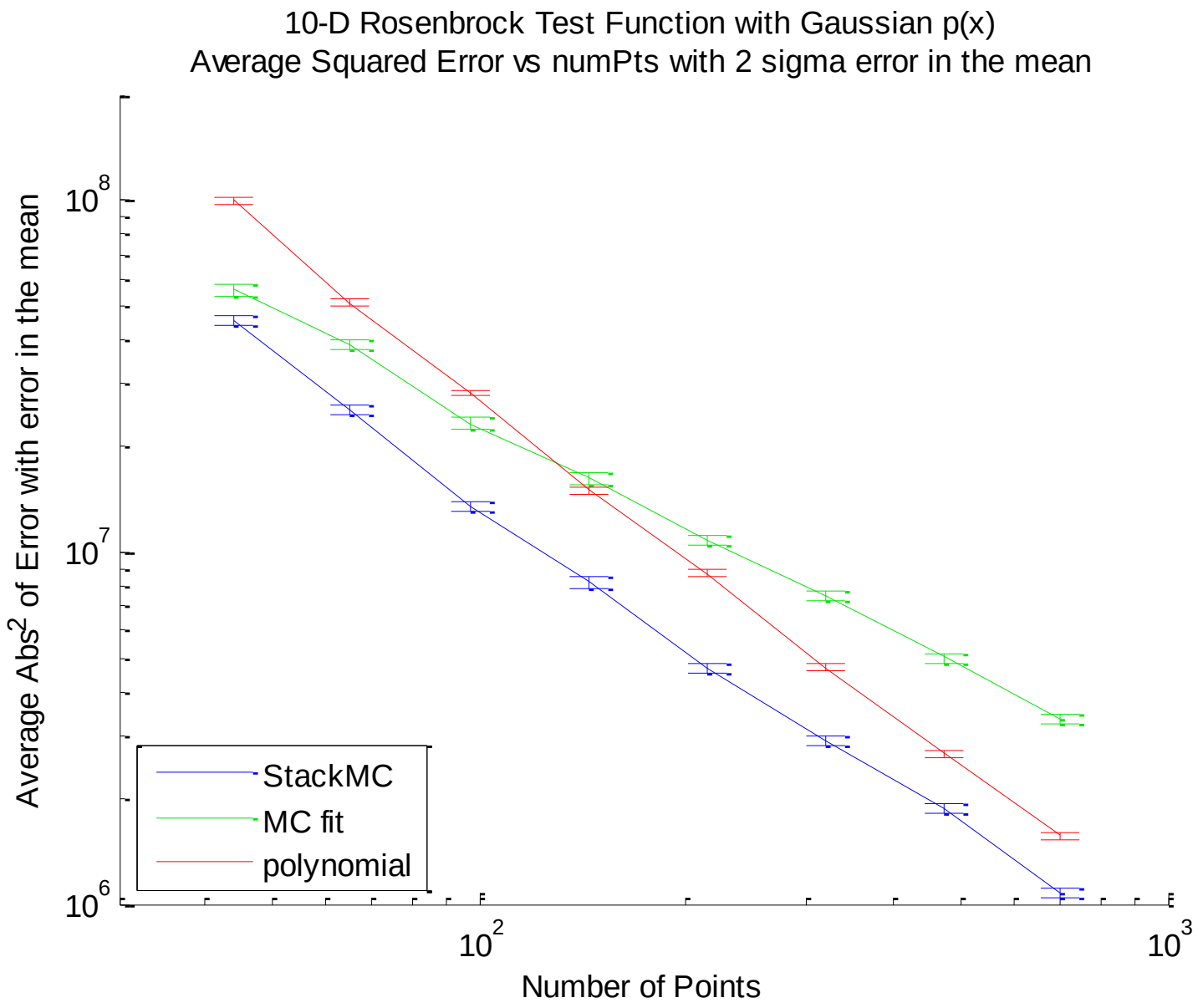


Stack MC

1) Here is how we choose g:
   - Fit g(x) to f(x) from a subset of the sample points among a class of functions we can integrate, using any machine learning technique;
2) Estimate E(f – g) using the remaining points.



Stack MC works pretty well



RT2: GOAL-ORIENTED ROMS FOR THE MULTI-SOURCE MULTI-PHYSICS SETTING

Research Thrust 2 (RT2) : Reduced Order Modeling

Goal-oriented ROMs for the multi-source multi-physics setting

Marc Mignolet, Karen Willcox

Overview

RT2.1: Uncertainty Modeling and Management in ROMs

RT2.2: ROMs Adaptation to Multidisciplinary Interactions

RT2.3: Compositional ROM Strategies

RT2.4: Localized and Adaptive ROMs

Research Thrust 2 (RT2) : Reduced Order Modeling

RT2.1: Uncertainty Modeling and Management in ROMs

Modeling:
- in FOM vs. directly in ROM?
- for different physics (structure, thermal,…)
- for interaction between different physics

Management: Define and use overall ROM uncertainty to assess adequacy of ROM, need for bases update, etc.

RT2.2: ROMs Adaptation to Multidisciplinary Interactions

- Detect strong/changing multidisciplinary interactions
- Develop methodology for adaptation of ROMs to unresolved/poorly resolved multi-physics interactions
- Estimate multidisciplinary uncertainty level from single-discipline counterparts

Research Thrust 2 (RT2) : Reduced Order Modeling

RT2.3: Compositional ROM Strategies

- For design/analysis of β within $\alpha+\beta$ with uncertainty on α, on β, on α and β
- For different physics (structure, thermal,…) and their interactions

RT2.4: Localized and Adaptive ROMs

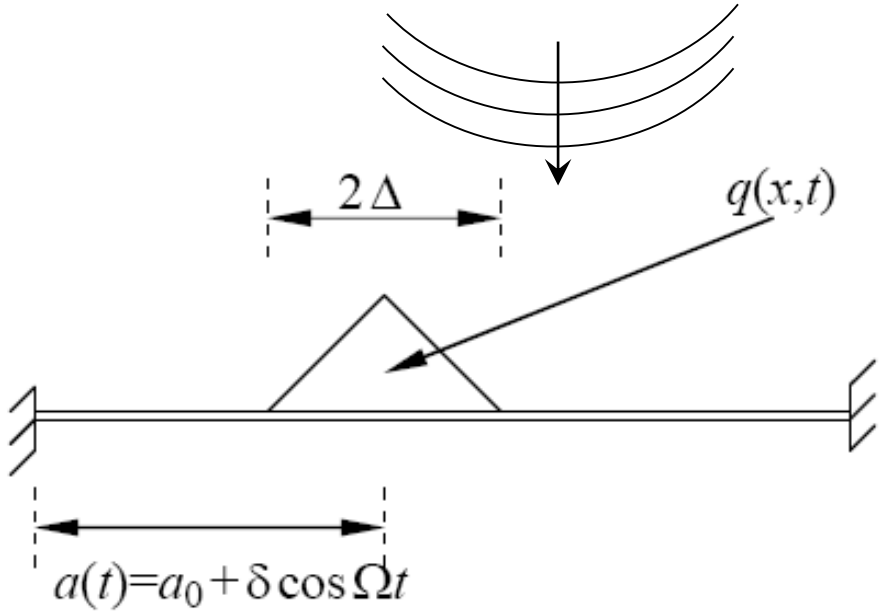
Need for (online) adaptation of models:

- Adapt to changing multi-physics coupling requirements as decision process proceeds
- Feedback of goal-driven uncertainty to adapt models
- Adapt model structure, order, fidelity, etc. to a local region (in state space or in parameter space)

Research Thrust 2 (RT2) : Uncertainty Modeling

A Testbed Problem: oscillating heat flux acting on beam

Goal: Analyze propagation of uncertainty in multiphysics system



Beam Length (L)	0.2286 m
Cross-section Width (w)	0.0127 m
Cross-section Thickness (h)	7.88 10^{-4} m
Density	2700 kg/m^3
Young's Modulus	73,000 MPa
Shear Modulus	27,730 MPa
Coeff. Thermal Expansion	2.5 10^{-5} /°C
Mesh (CBEAM)	40
Δ	.2 L
a_0	.5 L
Maximum Temperature	10 °C
δ	.075 L

Structural (17 modes)

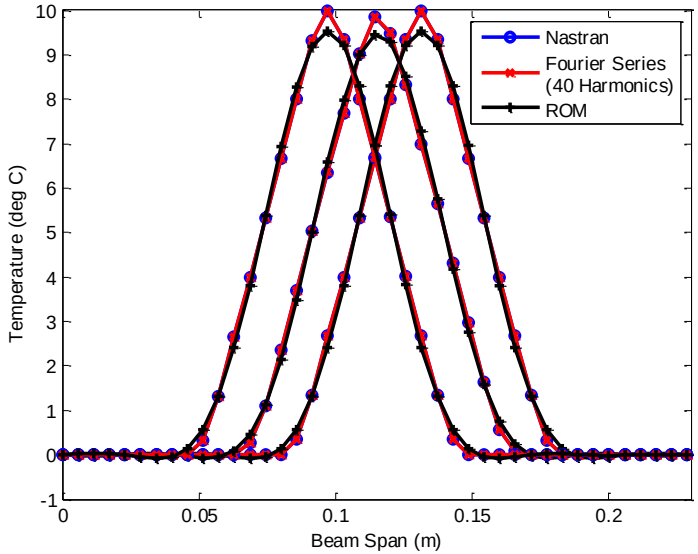
$$M_{ij}\,\ddot{q}_j + D_{ij}\,\dot{q}_j + K_{ij}^{(1)}\,q_j - K_{ijl}^{(th)}\,q_j\,\tau_l + K_{ijl}^{(2)}\,q_j\,q_l + K_{ijlp}^{(3)}\,q_j\,q_l\,q_p = F_i + F_{il}^{(th)}\tau_l$$

Thermal (12 modes)

$$B_{ij}\,\frac{d\left(\tau_j\right)}{dt} + \tilde{K}_{ij}\tau_j = P_i$$

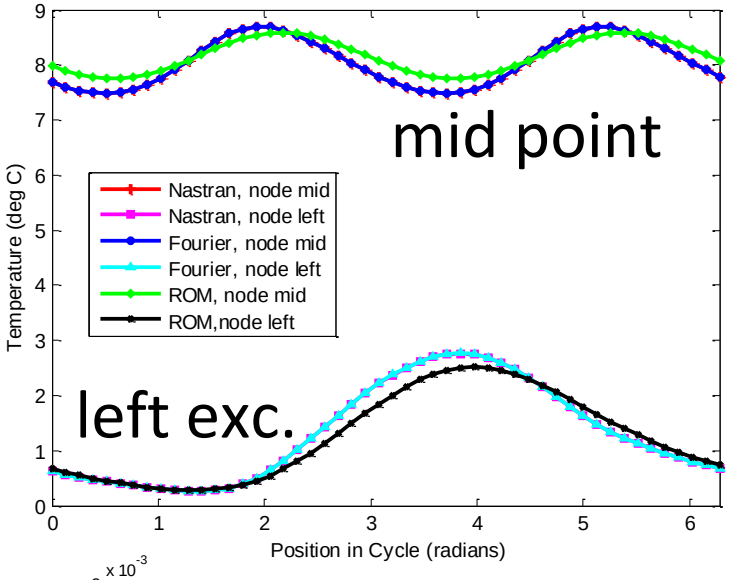
Research Thrust 2 (RT2) : Testbed Mean Model

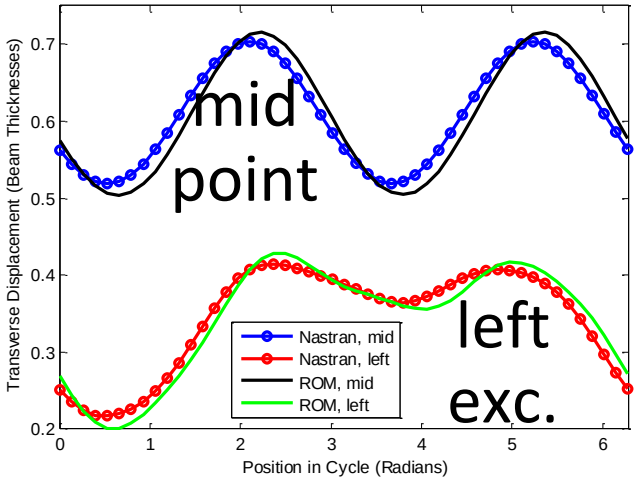
A Testbed Problem: oscillating heat flux acting on beam



Temperature at 0, 1/4, 1/2 period

vs. time

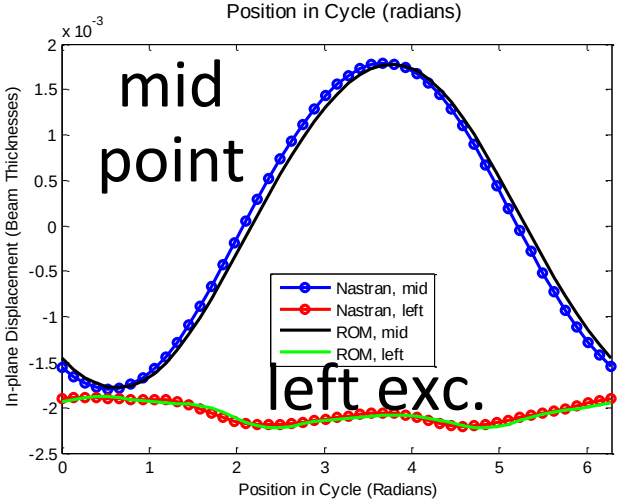




Structure

Transverse

In-Plane



Research Thrust 2 (RT2) : Uncertainty Modeling

How do we introduce uncertainty on thermal-structural (coeff. thermal expansion), thermal (conductance, capacitance), and structural (stiffness, mass, damping) properties?

- in FOM, then project on ROM: expensive in nonlinear ROMs
- directly on ROM:
  maximum entropy nonparametric approach (Soize, 2000)

- joint distribution of random parameters not specified a priori but derived to yield maximum of entropy
- physical requirements (mean specified, positive definiteness, …) added as constraints
- solution is computationally elegant

Research Thrust 2 (RT2) : Uncertainty Modeling

<u>Example</u>: Application to random symmetric, positive definite, non-singular matrices $\underline{\underline{A}}$ of given mean $\overline{\underline{\underline{A}}}$.

Maximize

$$S = -\int_{\Omega} p_{\underline{\underline{A}}}\left(\underline{\underline{a}}\right) \ln\, p_{\underline{\underline{A}}}\left(\underline{\underline{a}}\right) d\underline{\underline{a}}$$

with the constraints:

positive definiteness: $\Omega = \left\{\underline{\underline{a}} = \underline{\underline{\tilde{L}}}\,\underline{\underline{\tilde{L}}}^{T}; \tilde{L}_{ij}, i, j = 1,\ldots,n : \left[\tilde{L}_{ij} \in \left(-\infty,+\infty\right), i > j\right] \cap \left[\tilde{L}_{ii} \in \left[0,+\infty\right)\right]\right\}$

total probability: $\int_{\Omega} p_{\underline{\underline{A}}}\left(\underline{\underline{a}}\right) d\underline{\underline{a}} = 1$

mean model specified: $E\left[\underline{\underline{A}}\right] = \int_{\Omega} \underline{\underline{a}}\; p_{\underline{\underline{A}}}\left(\underline{\underline{a}}\right) d\underline{\underline{a}} = \overline{\underline{\underline{A}}}$

non-singularity: $\int_{\Omega} \ln\left[\det\left(\underline{\underline{a}}\right)\right] p_{\underline{\underline{A}}}\left(\underline{\underline{a}}\right) d\underline{\underline{a}} = \nu \;\; \text{finite}$

Research Thrust 2 (RT2) : Uncertainty Modeling

Solution:

Cholesky decompose $\underline{\underline{\bar{A}}}$ and $\underline{\underline{A}}$ as

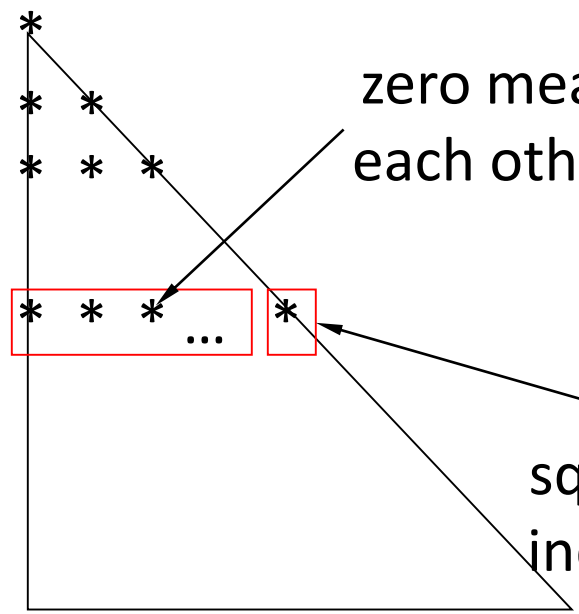
$$\underline{\underline{\bar{A}}} = \underline{\underline{\bar{L}}}\,\underline{\underline{\bar{L}}}^T \qquad \underline{\underline{A}} = \underline{\underline{\bar{L}}}\,\underline{\underline{G}}\,\underline{\underline{\bar{L}}}^T \qquad \underline{\underline{G}} = \underline{\underline{H}}\,\underline{\underline{H}}^T$$

it is found that

$$p_{\underline{\underline{G}}}\left(\underline{\underline{g}}\right) = \bar{C}\left[\det\left(\underline{\underline{g}}\right)\right]^{\lambda-1} \exp\left[-\mathrm{tr}\left(\underline{\underline{\mu}}^T\,\underline{\underline{g}}\right)\right]$$

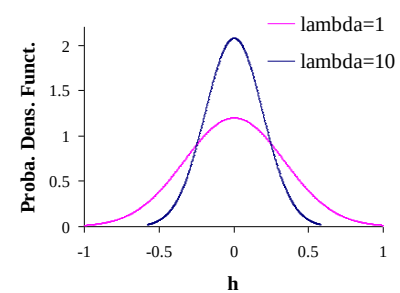
or equivalently

$\underline{\underline{H}}$ =



zero mean Gaussian, independent of each other with identical standard dev.

square root of Gamma, independent of all others

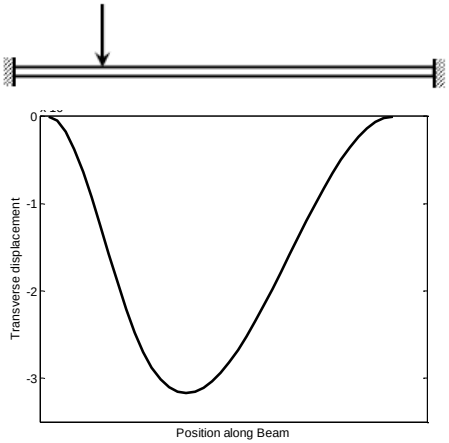


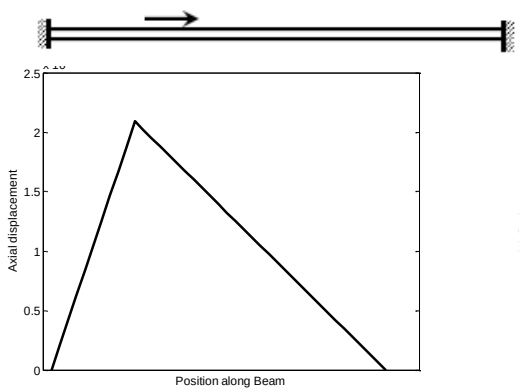
Research Thrust 2 (RT2) : Uncertainty Modeling

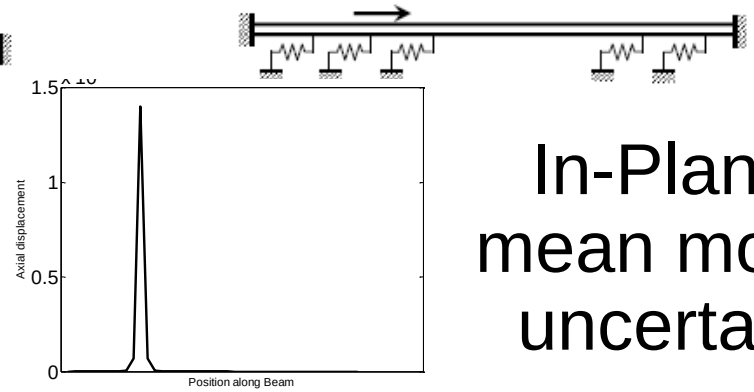
No topology involved (done by Soize, 2008)… Appropriate if:

_ ROM order is much smaller than FOM's

_ response is global expressed in global modes. OK for structural modeling (for which developed/implemented).

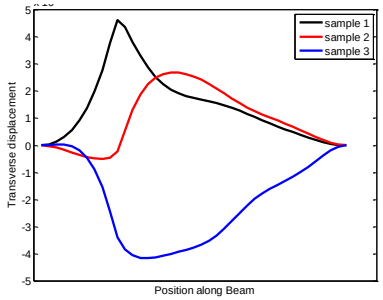






In-Plane mean mod./ uncertain

uncertainty in beam curvature introduces excitation in lower order (transverse) modes



Transverse uncertain

Research Thrust 2 (RT2) : Uncertainty Modeling

What about the thermal problem? Depends on boundary cond.
Analyze temperature distribution induced by a point heat flux

beam with fixed temperature on bottom - local response

standard nonparametric not desirable, revised approach being formulated

FOM to ROM implemented

beam with adiabatic on bottom – global response

standard nonparametric method applicable

Research Thrust 2 (RT2) : Uncertainty Modeling

Temperature affecting about 1/2 beam and global structural effects $\Longrightarrow$ max entropy nonparametric tentatively applicable for thermal-structural interaction terms $K_{ijl}^{(th)}$ and $F_{il}^{(th)}$.

<u>Procedure:</u>

(1) third order tensor $K_{ijl}^{(th)}$ of mean model is reshaped into a rectangular matrix

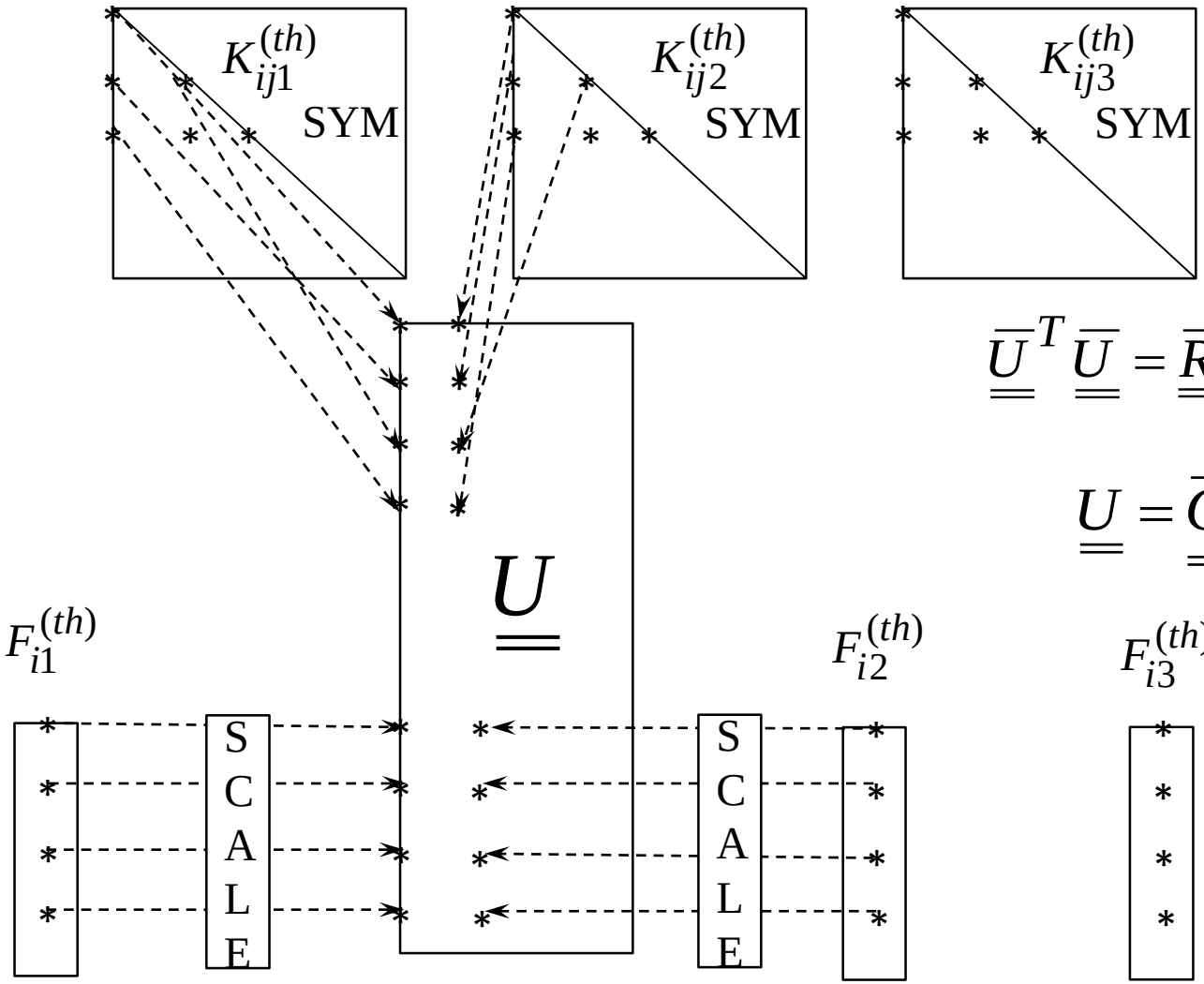
(2) mean model thermal forces $F_{il}^{(th)}$ included in rectangular matrix

(3) QR decomposition applied to reduce the uncertainty to its kernel (*R*)

(4) *R* matrix randomized in nonparametric form, mean *Q* retained

(5) Coefficient of variation of buckling temperature = 0.05

Research Thrust 2 (RT2) : Uncertainty Modeling



$$\underline{\underline{\overline{U}}} = \underline{\underline{\overline{Q}}}\,\underline{\underline{\overline{R}}}$$

$$\underline{\underline{\overline{U}}}^T \underline{\underline{\overline{U}}} = \underline{\underline{\overline{R}}}^T \underline{\underline{\overline{R}}} = \underline{\underline{\overline{R}}}^2 = \left(\underline{\underline{\overline{L}}}\,\underline{\underline{\overline{L}}}^T\right)^2$$

$$\underline{\underline{U}} = \underline{\underline{\overline{Q}}}\,\underline{\underline{R}} = \underline{\underline{\overline{Q}}}\,\underline{\underline{\overline{L}}}\,\underline{\underline{H}}\,\underline{\underline{H}}^T\,\underline{\underline{\overline{L}}}^T$$

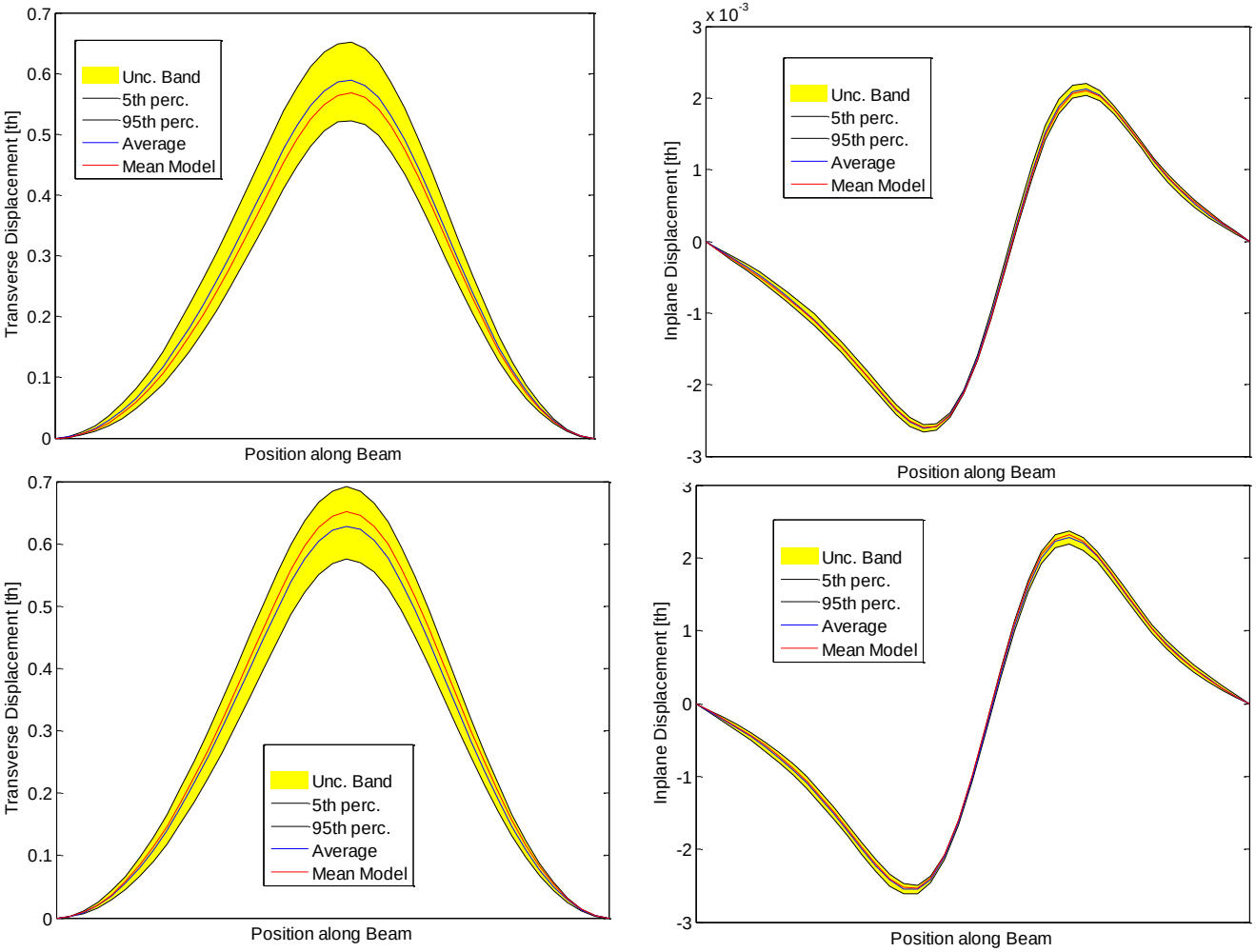
$$\underline{\underline{\overline{Q}}} = \underline{\underline{\overline{U}}}\,\underline{\underline{\overline{R}}}^{-1}$$

Testbed Results: uncertain K_{ijl} only, flux only

Transverse

In-Plane

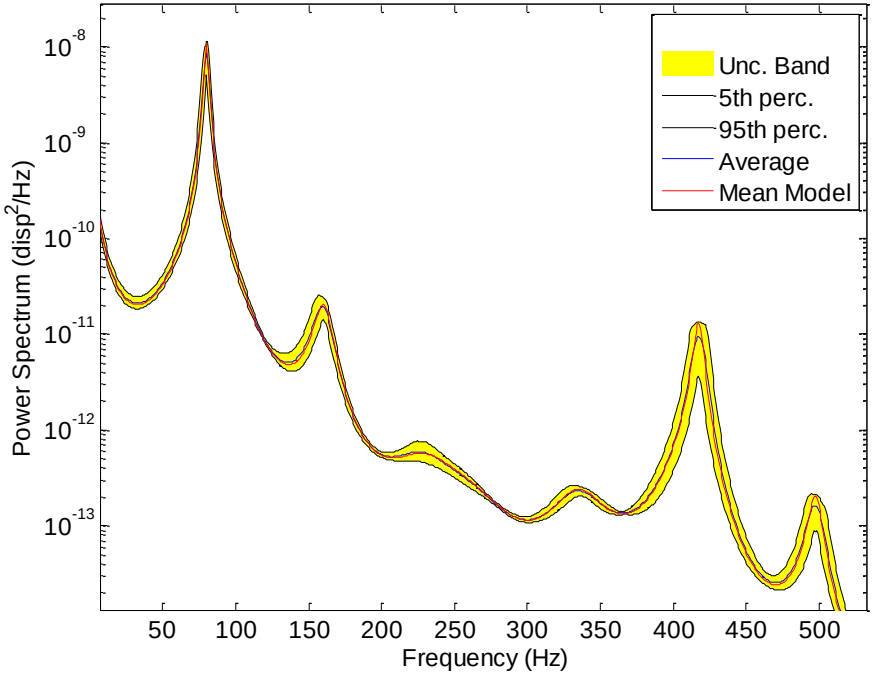
Snapshots of Response



Peak flux at right

Peak flux at middle

Testbed Results: uncertain K_{ijl} only, flux + acoustic

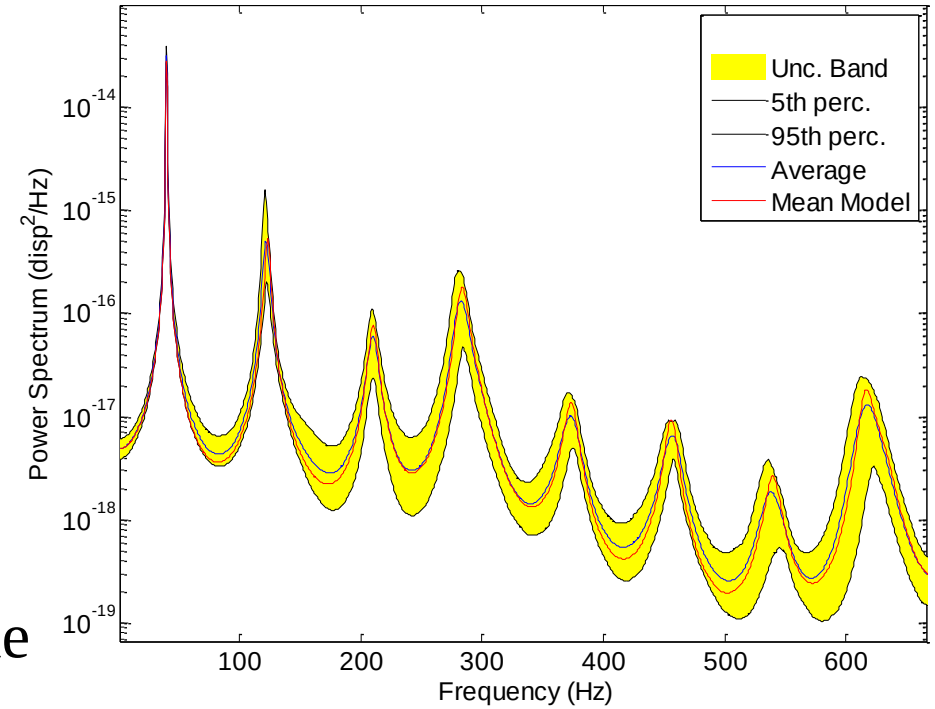


White Noise Excitation [0,1024] Hz 130dB OASPL. Transverse response at middle: mean 0.45th, std. 0.43 th

Transverse

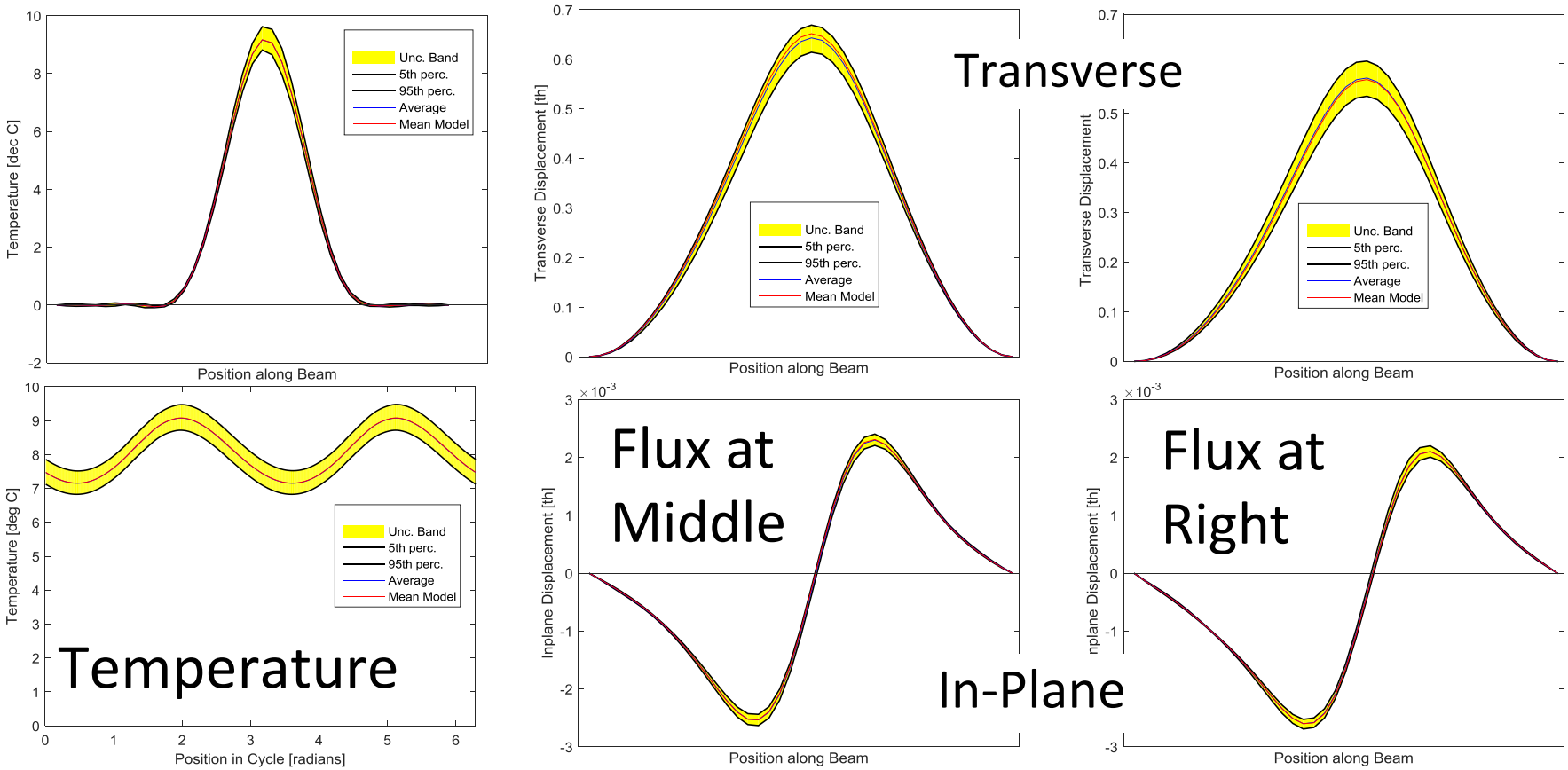
Beam Middle Point

In-Plane

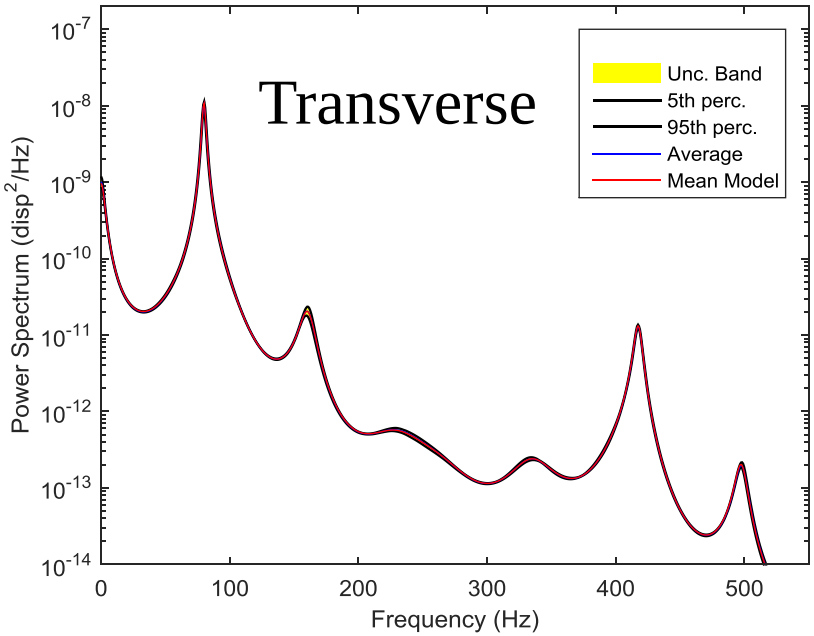


Research Thrust 2 (RT2) : Uncertain Conductance

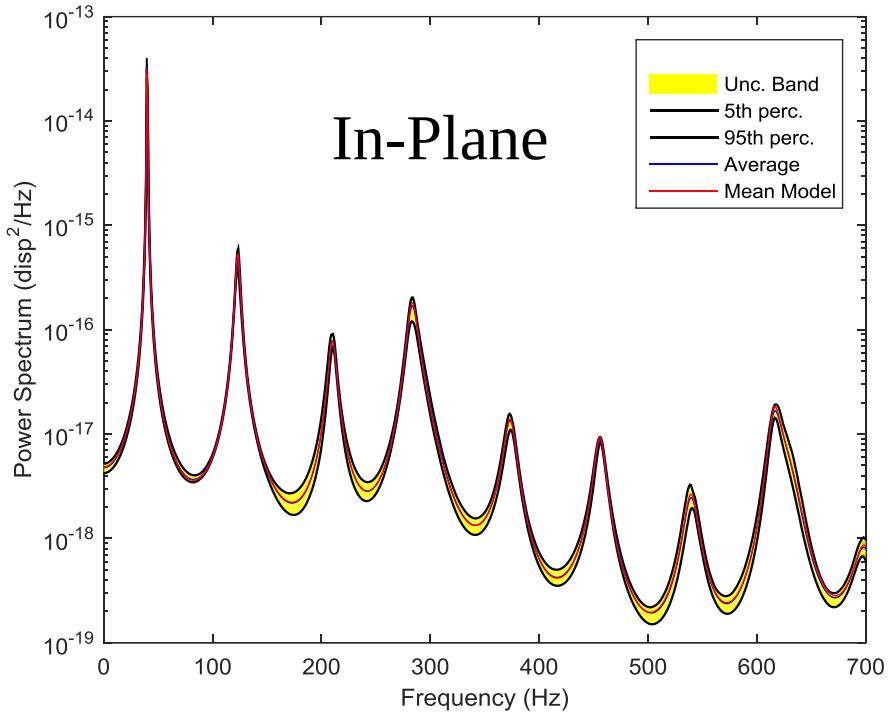
Conductance simulated in finite element model as homogenous random field with given mean, variance, and correlation length, projected on thermal ROM basis. Correlation through thickness.



Research Thrust 2 (RT2) : Uncertain Conductance



White Noise Excitation [0,1024] Hz 130dB OASPL. Transverse response at middle: mean 0.45th, std. 0.43 th



Beam Middle Point

Research Thrust 2 (RT2) : Status Summary

* Work on all 4 objectives of Research Thrust 2 initiated

* Work on uncertainty modeling focused on interaction between different physics and possibility of carrying out the modeling directly in ROM vs. in the FOM then projected on ROM

* Work on composition ROM strategies currently focused on a multi-bay panel and the sequential design of its panel with uncertainty on others (e.g., reflecting lack of design)

Joaquim Martins, Youssef Marzouk

RT3: MANAGING COUPLING IN MULTI-PHYSICS SYSTEMS

Here's the Research Thrust 3 (RT3) team

Joaquim Martins

UofM

Youssef Marzouk

MIT

Karen Willcox

MIT

Doug Allaire

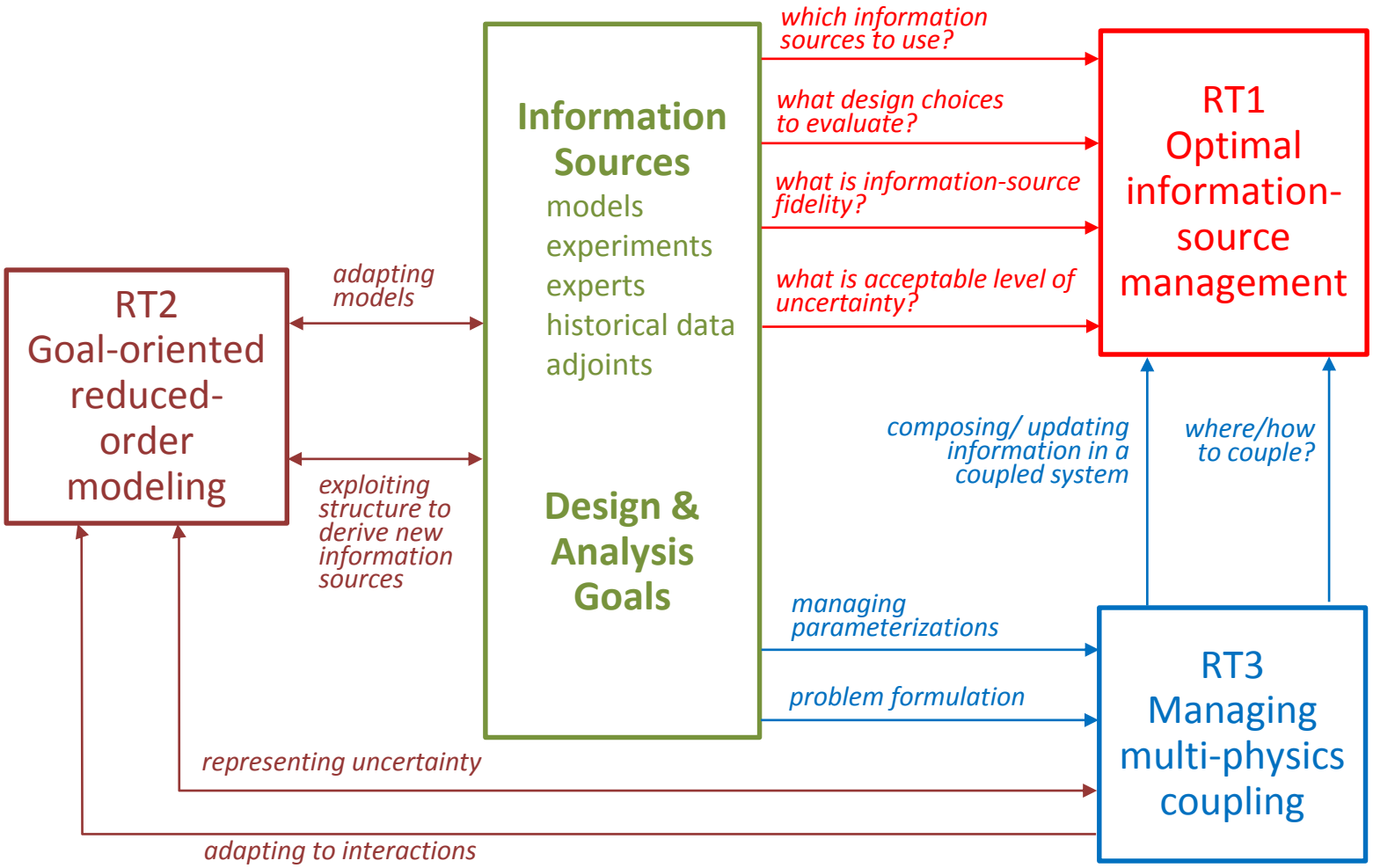
Texas A&M

Marc Mignolet

ASU

Here's how RT3 interfaces with the other thrusts

A scalable multi-source framework for certifiable solution of multi-physics analysis and design problems under uncertainty



Research Thrust 3's objectives

- Develop a scalable multi-source approach for formulating and solving multi-physics analysis and design problems
  - Mathematical formulation
  - Algorithmic framework
  - Software implementation
- Develop new information-theoretic and statistical approaches for characterizing, modeling and handling multi-physics coupling
- Apply the developed methods to an application of interest

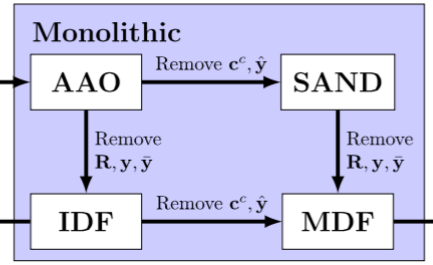
Background: multi-physics/multidisciplinary coupling

- Multi-physics coupling studies in the context of specific disciplines for both analysis and design optimization
  - Fluid-structure interaction [Kenway and Martins 2014]
  - MOOSE framework [Gaston et al. 2009]

- General architectures for solving multidisciplinary design optimization problems
  - Monolithic
  - Distributed individual discipline feasible
  - Distributed multidisciplinary feasible

Background: MDO architectures

$$
\begin{aligned}
\text{minimize} \quad & f_0(\mathbf{x},\mathbf{y}) + \sum_{i=1}^{N} f_i(\mathbf{x}_0,\mathbf{x}_i,\mathbf{y}_i) \\
\text{with respect to} \quad & \mathbf{x},\hat{\mathbf{y}},\mathbf{y},\bar{\mathbf{y}} \\
\text{subject to} \quad & \mathbf{c}_0(\mathbf{x},\mathbf{y}) \geq 0 \\
& \mathbf{c}_i(\mathbf{x}_0,\mathbf{x}_i,\mathbf{y}_i) \geq 0 \qquad \text{for} \quad i = 1,\ldots,N \\
& \mathbf{c}_i^c = \hat{\mathbf{y}}_i - \mathbf{y}_i = 0 \qquad \text{for} \quad i = 1,\ldots,N \\
& \mathbf{R}_i(\mathbf{x}_0,\mathbf{x}_i,\hat{\mathbf{y}}_{j\neq i},\bar{\mathbf{y}}_i,\mathbf{y}_i) = 0 \qquad \text{for} \quad i = 1,\ldots,N
\end{aligned}
$$



Nomenclature

f	Objective	$\mathbf{c}$	Design constraints
$\mathbf{c}^c$	Consistency constraints	$\mathbf{R}$	Analysis constraints
$\mathbf{x}$	Design variables	$\mathbf{y}$	Coupling variables
$\bar{\mathbf{y}}$	State variables	$\hat{\mathbf{y}}$	Variable copies
$_0$	Shared data	$_i$	Discipline i data

Distributed IDF

Multilevel

CO: Copies of the shared variables are created for each discipline, together with corresponding consistency constraints. Discipline subproblems minimize difference between the copies of shared and local variables subject to local constraints. System subproblem minimizes objective subject to shared constraints subject to consistency constraints.

BLISS-2000: Discipline subproblems minimize the objective with respect to local variables subject to local constraints. A surrogate model of the local optima with respect to the shared variables is maintained. Then, system subproblem minimizes objective with respect to shared design and coupling variables subject to shared design and consistency constraints, considering the disciplinary preferences.

QSD: Each discipline is assigned a "budget" for a local objective and the discipline problems maximize the margin in their local constraints and the budgeted objective. System subproblem minimizes a shared objective and the budgets of each discipline subject to shared design constraints and positivity of the margin in each discipline.

Penalty

ATC: Copies of the shared variables are used in discipline subproblems together with the corresponding consistency constraints. These consistency constraints are relaxed using a penalty function. System and discipline subproblems solve their respective relaxed problem independently. Penalty weights are increased until the desired consistency is achieved.

IPD/EPD: Applicable to MDO problems with no shared objectives or constraints. Like ATC, copies of shared variables are used for every discipline subproblem and the consistency constraints are relaxed with a penalty function. Unlike ATC, the simple structure of the disciplinary subproblems is exploited to compute post-optimality sensitivities to guide the system subproblem.

ECO: As in CO, copies of the shared design variables are used. Disciplinary subproblems minimize quadratic approximations of the objective subject to local constraints and linear models of nonlocal constraints. Shared variables are determined by the system subproblem, which minimizes the total violation of all consistency constraints.

Distributed MDF

CSSO: In system subproblem, disciplinary analyses are replaced by surrogate models. Discipline subproblems are solved using surrogates for the other disciplines, and the solutions from these discipline subproblems are used to update the surrogate models.

BLISS: Coupled derivatives of the multidisciplinary analysis are used to construct linear subproblems for each discipline with respect to local design variables. Post-optimality derivatives from the solutions of these subproblems are computed to form the system linear subproblem, which is solved with respect to shared design variables.

MDOIS: Applicable to MDO problems with no shared objectives, constraints, or design variables. Discipline subproblems are solved independently assuming fixed coupling variables, and then a multidisciplinary analysis is performed to update the coupling.

ASO: System subproblem is like that of MDF, but some disciplines solve a discipline optimization subproblem within the multidisciplinary analysis with respect to local variables subject to local constraints. Coupled post-optimality derivatives from the discipline subproblems are computed to guide the system subproblem.

[Martins and Lambe, 2013]

Background: limitations of current methods

- Decomposition pre-determined along disciplinary boundaries
  - Need: Methods for automatic discovery of new decomposition
- Existing coupling approaches consider only single fidelity multi-physics models
  - Need: Methods that handle multifidelity coupling and multi-source information
- Multi-physics/multidisciplinary coupling is considered deterministic
  - Need: Methods with approximate coupling considering uncertainty

We propose several new ideas to address these needs

1. A new mathematical view of the multifidelity multidisciplinary analysis and design optimization problem
2. A new information-theoretic approach to learning and approximating multi-physics coupling
3. A heterogeneous approach for solving a multi-physics system
4. Automatic coupled sensitivity analysis via a new coupled adjoint method that considers multifidelity models
5. A compositional goal-driven approach to quantification of uncertainty

A new mathematical view of multidisciplinary problems

Any computational model can be decomposed into

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_q \end{pmatrix}$$

Input variables State variables Output variables

The state variables and output variables are computed by the functions:

$$\mathcal{Y} = \begin{pmatrix} \mathcal{Y}_1 \\ \vdots \\ \mathcal{Y}_p \end{pmatrix}, \quad \mathcal{F} = \begin{pmatrix} \mathcal{F}_1 \\ \vdots \\ \mathcal{F}_q \end{pmatrix}$$

We encapsulate the whole system as residuals and variables

Then, the full system has the following form:

$$u = \begin{pmatrix} x \\ y \\ f \end{pmatrix} \quad \text{and} \quad R(u) = \begin{pmatrix} x - x^* \\ -\mathcal{R}(x, y) \\ f - \mathcal{F}(x, y) \end{pmatrix}$$

The numerical model can be encapsulated as:

$$\mathcal{G} : x \mapsto \mathcal{F}(x, \mathcal{Y}(x))$$

We want the total derivative of the outputs with respect to the inputs, i.e., $\partial \mathcal{G} / \partial x$.

Now we derive a unifying formula for computing the desired total derivatives

We start with the following identity:

$$\frac{\partial R}{\partial u} \qquad \frac{\partial R}{\partial u}^{-1} \qquad = \qquad \mathcal{I}$$

Which expands to :

$$\begin{bmatrix} \mathcal{I} & 0 & 0 \\ -\frac{\partial \mathcal{R}}{\partial x} & -\frac{\partial \mathcal{R}}{\partial x} & 0 \\ -\frac{\partial \mathcal{F}}{\partial x} & -\frac{\partial \mathcal{F}}{\partial x} & \mathcal{I} \end{bmatrix} \begin{bmatrix} A^{xx} & A^{xy} & A^{xf} \\ A^{yx} & A^{yy} & A^{yf} \\ A^{fx} & A^{fy} & A^{ff} \end{bmatrix} = \begin{bmatrix} \mathcal{I} & 0 & 0 \\ 0 & \mathcal{I} & 0 \\ 0 & 0 & \mathcal{I} \end{bmatrix}$$

We can express both adjoint and direct methods in a single equation

Block-forward substitution for the first block-column yields:

$$\begin{bmatrix} A^{xx} \\ A^{yx} \\ A^{fx} \end{bmatrix} = \begin{bmatrix} \mathcal{I} \\ -\dfrac{\partial \mathcal{R}}{\partial y}^{-1} \dfrac{\partial \mathcal{R}}{\partial x} \\ \dfrac{\partial \mathcal{F}}{\partial x} - \dfrac{\partial \mathcal{F}}{\partial y} \dfrac{\partial \mathcal{R}}{\partial y}^{-1} \dfrac{\partial \mathcal{R}}{\partial x} \end{bmatrix}$$

Given the previous definition,

$$\frac{\partial \mathcal{G}}{\partial x} = \begin{bmatrix} \dfrac{\partial \mathcal{F}}{\partial x} & \dfrac{\partial \mathcal{F}}{\partial y} \end{bmatrix} \begin{bmatrix} \mathcal{I} \\ \dfrac{\partial \mathcal{Y}}{\partial x} \end{bmatrix} = \frac{\partial \mathcal{F}}{\partial x} - \frac{\partial \mathcal{F}}{\partial y} \frac{\partial \mathcal{R}}{\partial y}^{-1} \frac{\partial \mathcal{R}}{\partial x}$$

We need to define an inverse function that computes the states with respect to the residuals

By the inverse function theorem, we have an inverse function

$$R^{-1} : r \mapsto u | R(u) = r$$

That satisfies:

$$\frac{\partial(R^{-1})}{\partial r} = \left[\frac{\partial R}{\partial u}\right]^{-1}$$

We now define this as the total derivatives of the states with respect to the residuals:

$$\frac{\mathrm{d}u}{\mathrm{d}r} = \frac{\partial(R^{-1})}{\partial r}$$

We now have all the terms in the unifying formula for computing the total derivatives

We can then derive the unifying equation for system derivatives:

$$\frac{\partial R}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}r} = \mathcal{I} = \frac{\partial R}{\partial u}^T \frac{\mathrm{d}u}{\mathrm{d}r}^T$$

$$\begin{bmatrix} \mathcal{I} & 0 & 0 \\ -\frac{\partial R}{\partial x} & -\frac{\partial R}{\partial y} & 0 \\ -\frac{\partial F}{\partial x} & -\frac{\partial F}{\partial y} & \mathcal{I} \end{bmatrix} \begin{bmatrix} \mathcal{I} & 0 & 0 \\ \frac{\mathrm{d}y}{\mathrm{d}x} & \frac{\mathrm{d}y}{\mathrm{d}r} & 0 \\ \frac{\mathrm{d}f}{\mathrm{d}x} & \frac{\mathrm{d}f}{\mathrm{d}r} & \mathcal{I} \end{bmatrix} = \begin{bmatrix} \mathcal{I} & 0 & 0 \\ 0 & \mathcal{I} & 0 \\ 0 & 0 & \mathcal{I} \end{bmatrix}$$

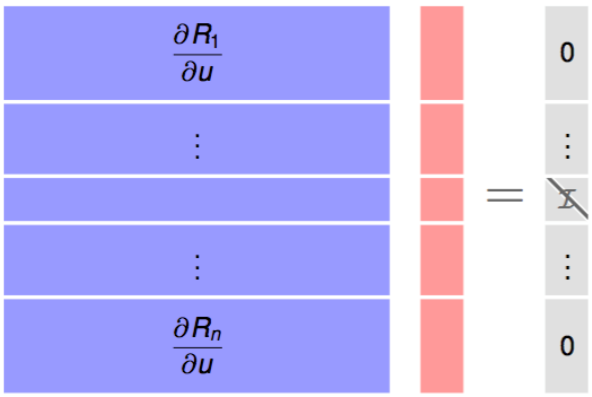
The various methods for computing derivatives can be unified

$$\frac{\partial R}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}r} = \mathcal{I} = \frac{\partial R}{\partial u}^T \frac{\mathrm{d}u}{\mathrm{d}r}^T$$

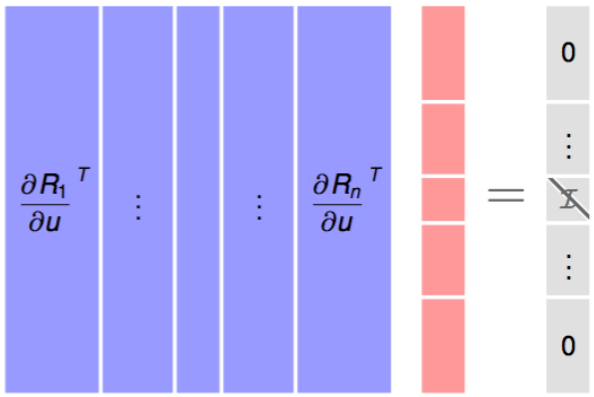
- Finite differences
- Chain rule
- Direct method and adjoint method
- Algorithmic differentiation

[Martins and Hwang, 2013]

Implementation leads to efficient derivative computation



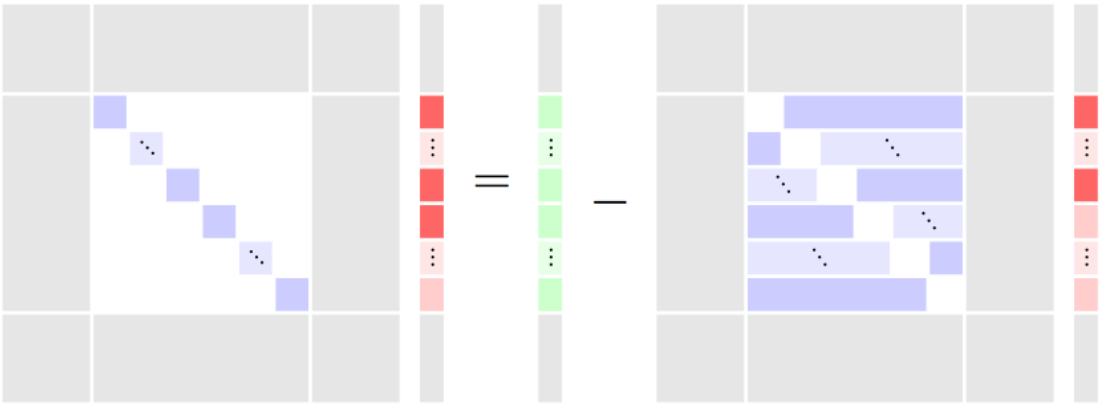
$$\frac{\partial R}{\partial u}\frac{du}{dr} = \mathcal{I}$$



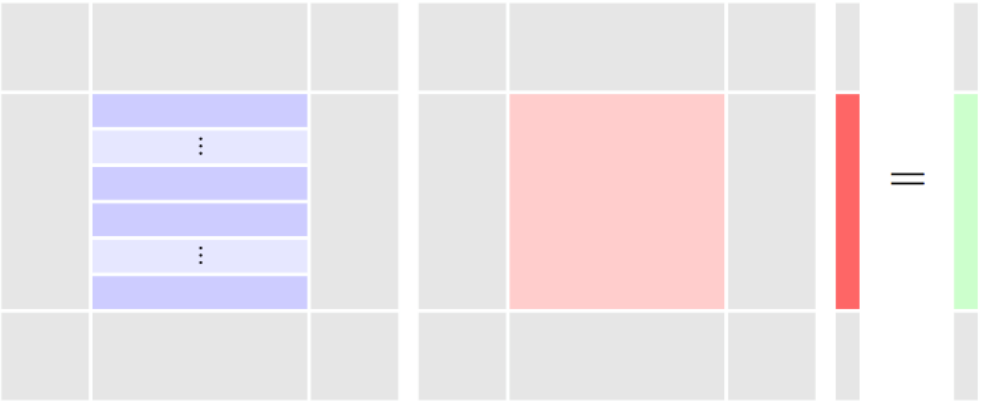
$$\frac{\partial R}{\partial u}^T\frac{du}{dr}^T = \mathcal{I}$$

The system can be decomposed and parallelized

Block Gauss-Seidel



Preconditioned
Krylov subspace methods

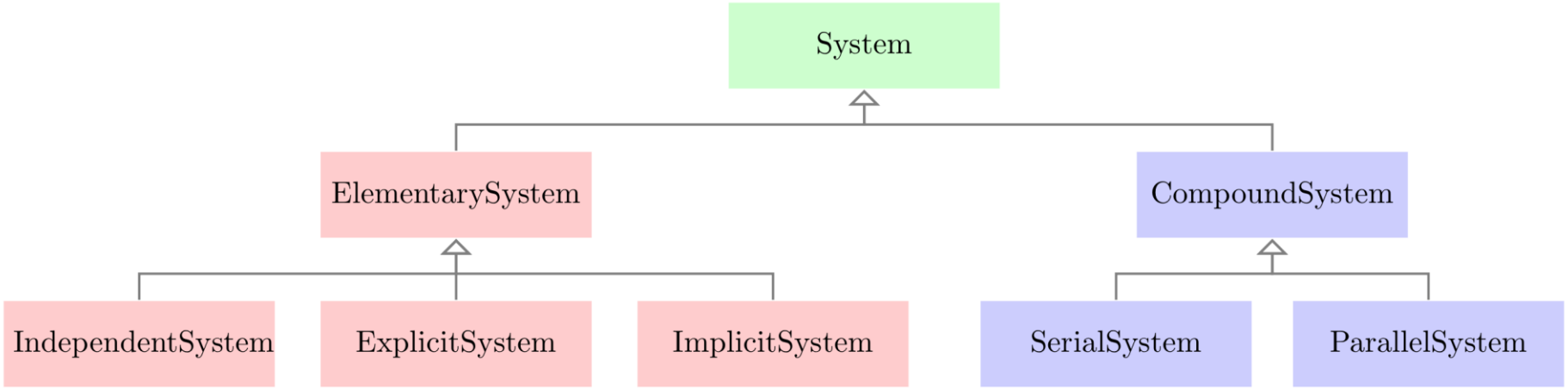


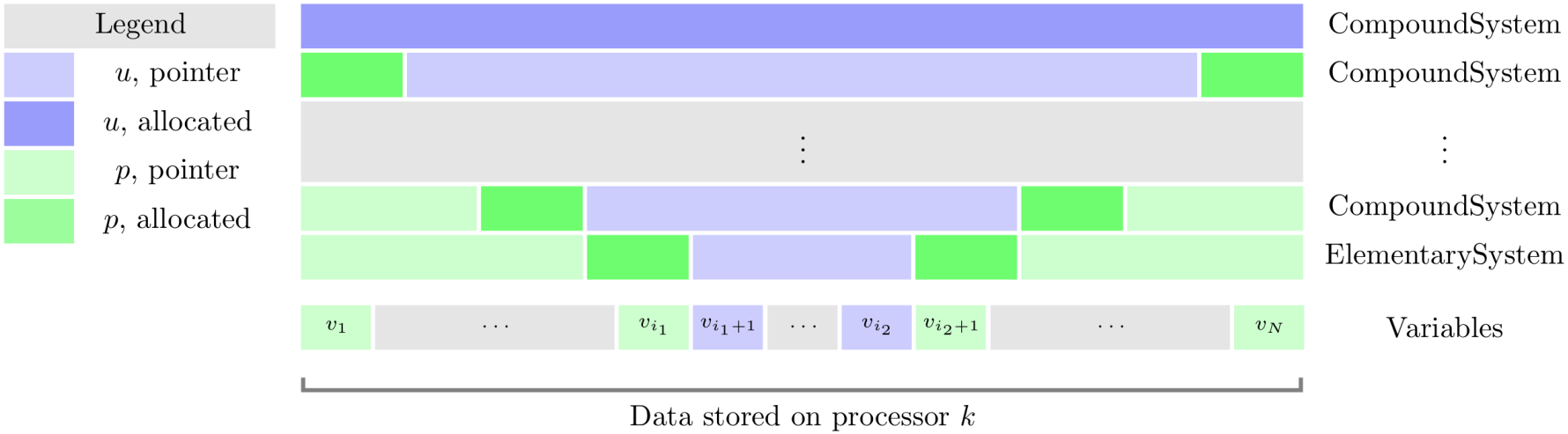
Implementation uses a compact API

		System classes		
		System	*CompoundSystem*	*ElementarySystem*
Virtual methods	*apply_nonlinear*	—	Recursive	User-implemented
	apply_linear		Recursive	User-implemented or FD*
	solve_nonlinear	Newton with line search	Nonlinear block Gauss–Seidel/Jacobi	Optional
	solve_linear	Krylov-type with preconditioning	Linear block Gauss–Seidel/Jacobi	Optional

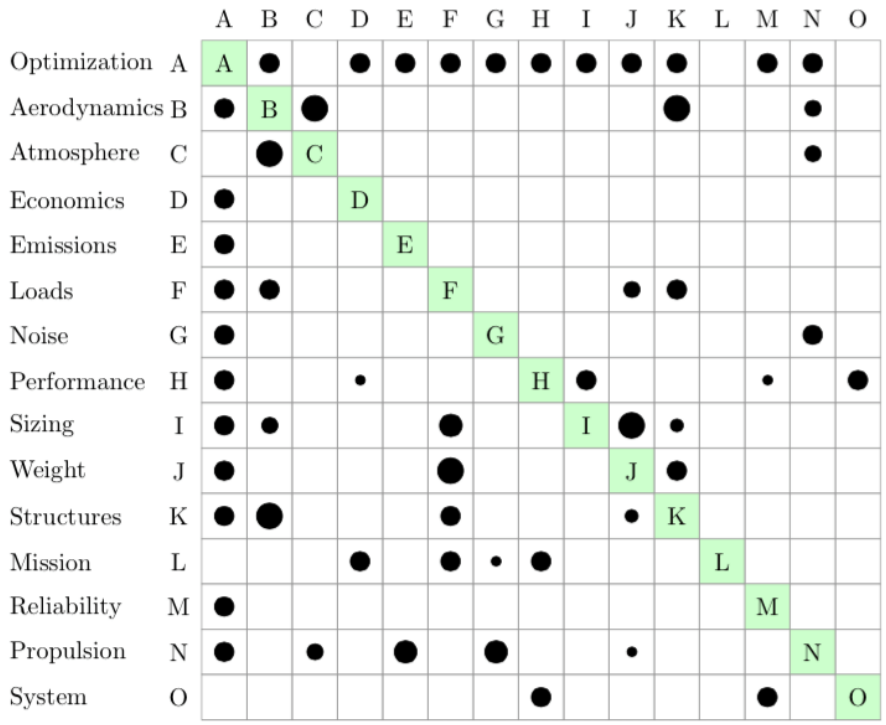
*FD: finite-difference approximation of the Jacobian.

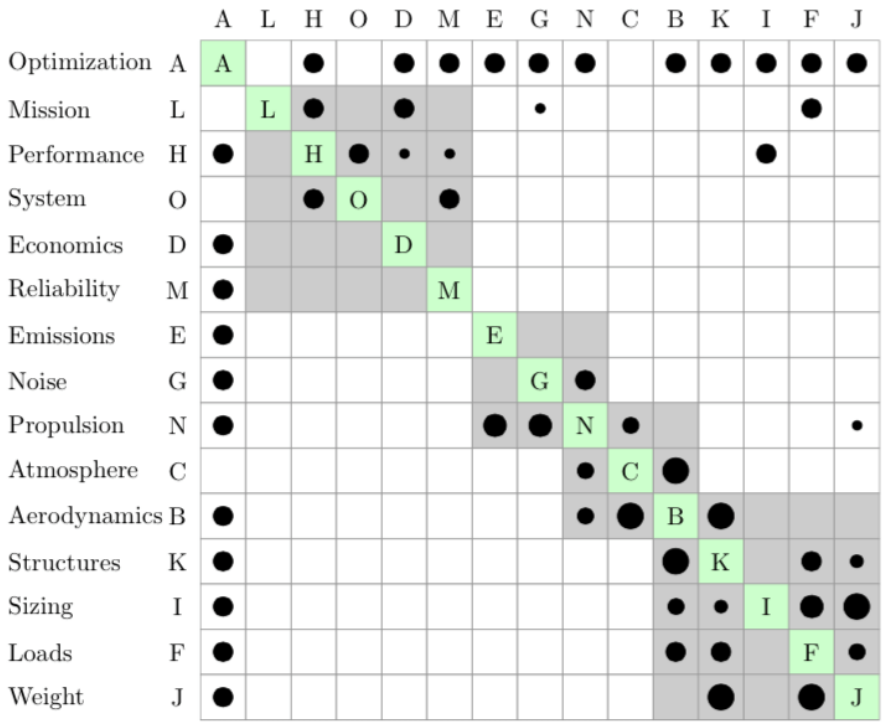
Implementation uses a hierachical approach





Example of automatic decomposition





[Lu and Martins, 2012]

Research questions

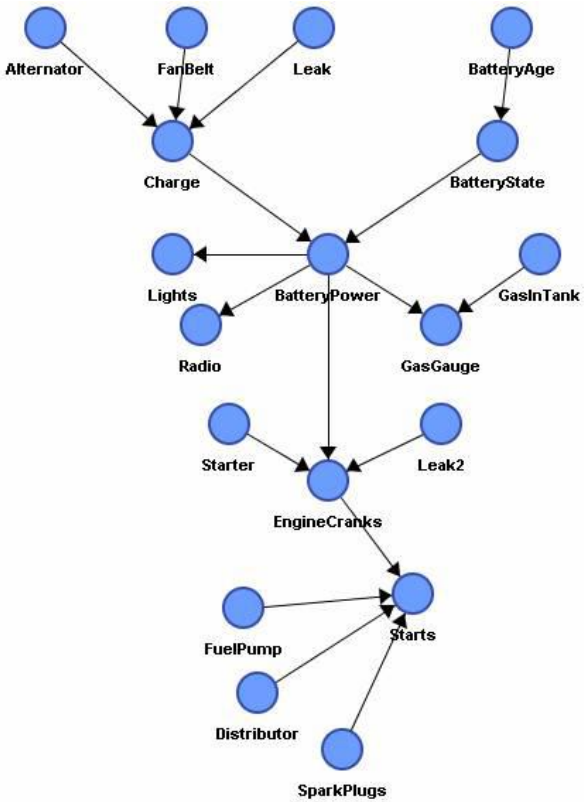
- How can we extend this theory and algorithmic framework for uncertainty quantification?
- How can we integrate multiple versions/fidelities of a given into the system?
- How can we exploit the weak coupling in an automated and mathematically sound way?
- Are there reordering and nesting approaches that can improve the computational performance?

Variational approaches to multi-physics coupling

- We've talked about efficient computation with given couplings, and even DSM re-ordering.
- But some fundamental questions remain:
  - *Which* disciplines should be coupled? How *tightly* do they need to interact?
  - Given *constraints* on computational resources, specific *prediction goals,* or specific *design optimization objectives*, which couplings can be neglected?
  - How should uncertainties affect the coupling strategy?
  - Can we answer these questions optimally?

Factorizing the joint probability distribution

- Let ***u*** represent the complete state (inputs, component states, outputs) of the multi-physics system
- We cannot easily work with the joint probability distribution *p(**u**)*, but what about its approximations?
  - **Key idea: *decomposition = conditional independence***
  - Removing interactions between disciplines is equivalent to **factorizing** the joint probability distribution *p*
  - Example: fluid + non-deforming solid
- This idea is central to probabilistic graphical models (PGMs)



Research questions

- A particular quantity of interest (QoI) u_i is represented by a node of the graphical model
  - QoI could be something we want to predict, or it could be a design optimization objective
- Variational approach: which decoupling will yield a *best* approximation of the marginal probability distribution of this QoI?

$$\min_{q \in \mathcal{Q}_\ell} D_{\mathrm{KL}}\left(\varphi \middle\| \psi\right) \quad \text{where} \quad \begin{aligned} \psi(u_i) &= \int p(u_i, \boldsymbol{u}_{\sim i})\, d\boldsymbol{u}_{\sim i} \\ \varphi(u_i) &= \int q(u_i, \boldsymbol{u}_{\sim i})\, d\boldsymbol{u}_{\sim i} \end{aligned}$$

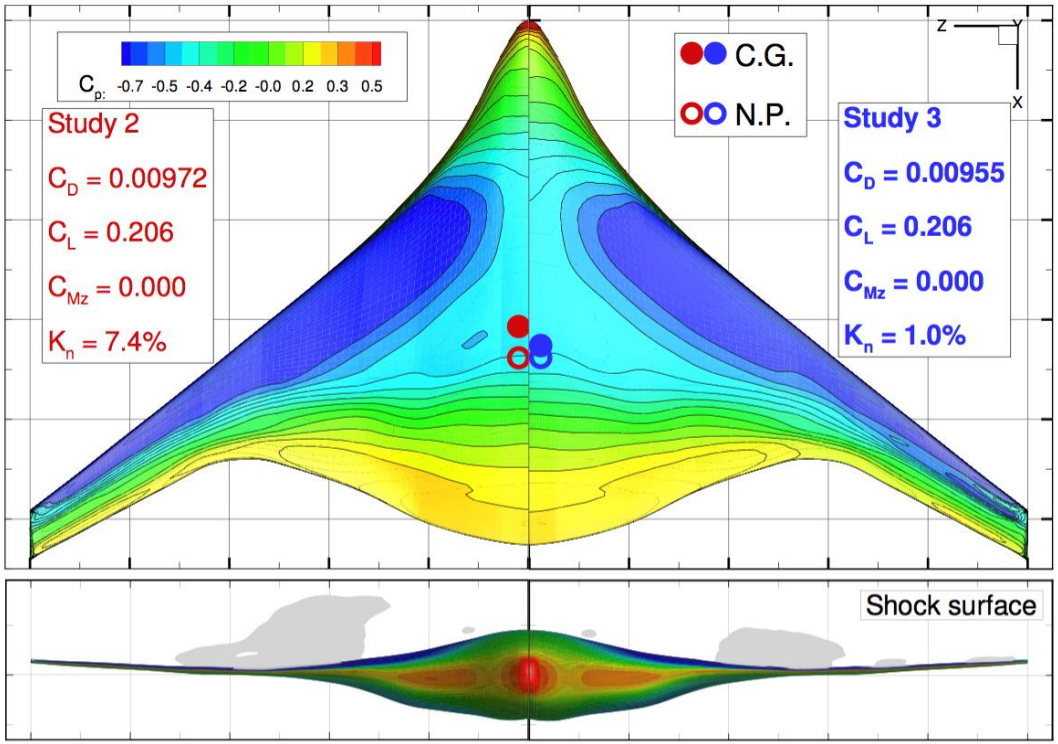
  - Possible choices of approximating class $\mathcal{Q}_\ell$: *Couple at most ℓ variables? Separate particular disciplines? Impose one-way versus two-way couplings?*

Research questions

- Challenges we are considering:
  - Useful representations of the joint distribution (beyond discrete or Gaussian probability distributions)
  - Efficient algorithms for extracting marginal distributions, given some coupling structure
  - Use in design optimization: acceptable coupling may depend on the current value of the design variables

APPLICATION SYSTEM OF INTEREST

Application: Multidisciplinary analysis and design of a blended-wing body aircraft

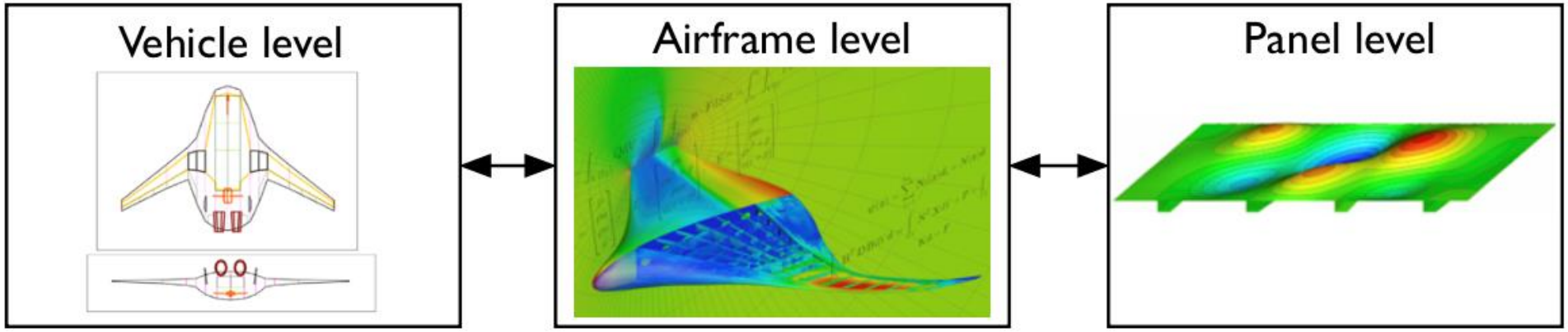


[Lyu and Martins, 2014]

- Aerodynamic design inherently closely coupled with trim and stability
- Wing flexibility (through structures) also affects trim and stability
- Thermal interaction due to electric systems and propulsion integrations

Design with multiple levels of detail and multiple information sources

- Multiple system parametric representations
- Multiple resolution of the physics
- Multiple numerical models and experimental data



Conceptual sizing tools including all relevant disciplines with low-fidelity

Preliminary design of aerodynamic shape and structural sizing with trim, stability and strength constraints

Detailed design with full aero-thermo-structural analysis

Summary

http://mcubed.mit.edu/

A scalable multi-source framework for certifiable solution of multi-physics analysis and design problems under uncertainty