Sequential design of experiments on a stochastic multi-fidelity simulator

STROH Rémi[a,b]
BECT Julien[a], VAZQUEZ Emmanuel[a]
DEMEYER Séverine[b], FISCHER Nicolas[b]
(a) Laboratoire des Signaux et Systèmes (L2S)
(b) Laboratoire National de métrologie et d'Essais (LNE)

MEASUREMENT AND STANDARDS

COMPETITIVENESS AND SAFETY VECTOR

03/22/2017
GdR MASCOT-NUM 2017

Outline

1. Introduction
2. Sequential design of experiments
3. Academic example

Outline

1. Introduction

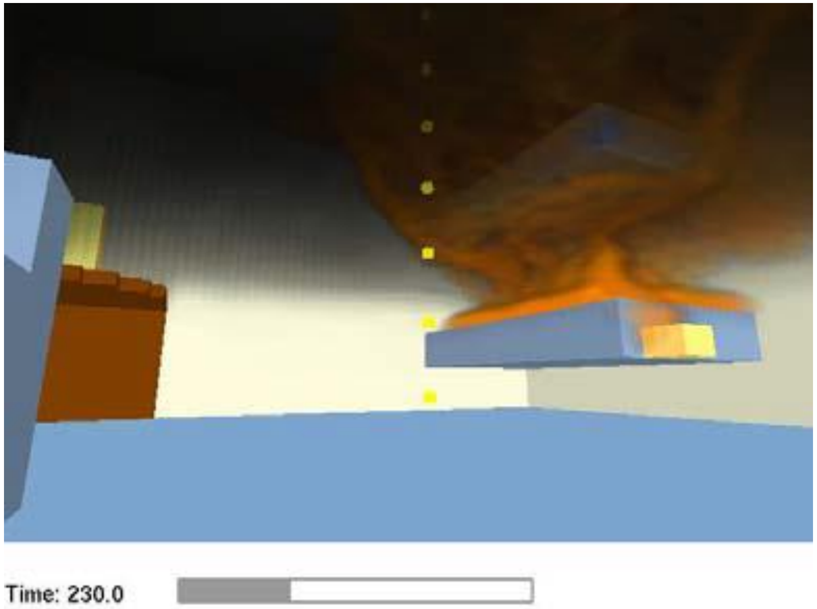
2. Sequential design of experiments

3. Academic example

Context

- Fire safety: conformity of a smoke extraction system
  - Expensive experiments → use of numerical models

Real Experiment



Fire Dynamics Simulator (FDS)

- Images from [Kerber, 2005]

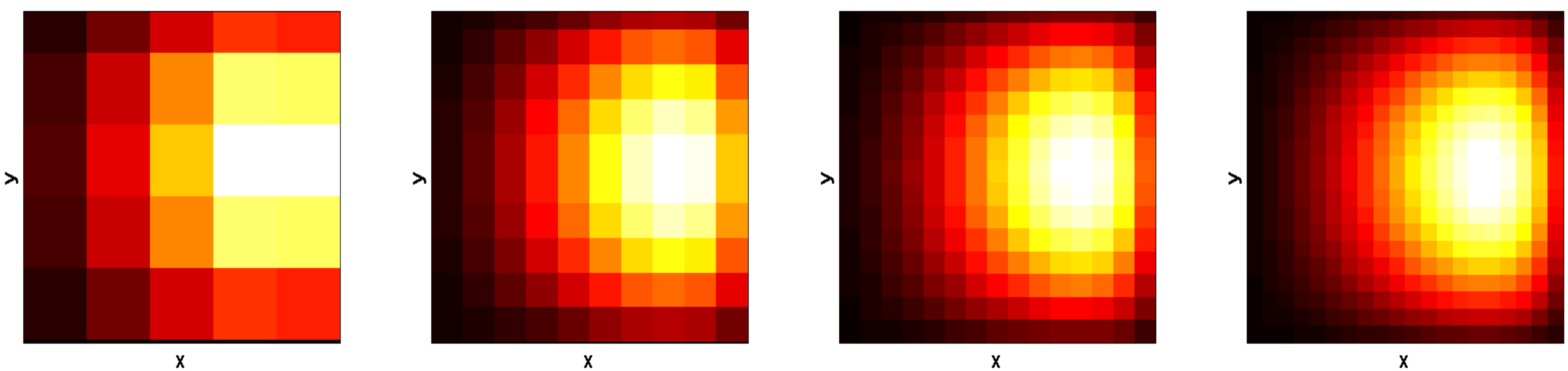
Properties of numerical models

- Main properties of the considered simulators:
  1. Multi-fidelity
  2. Tunable cost
  3. Stochastic outputs

Multi-fidelity

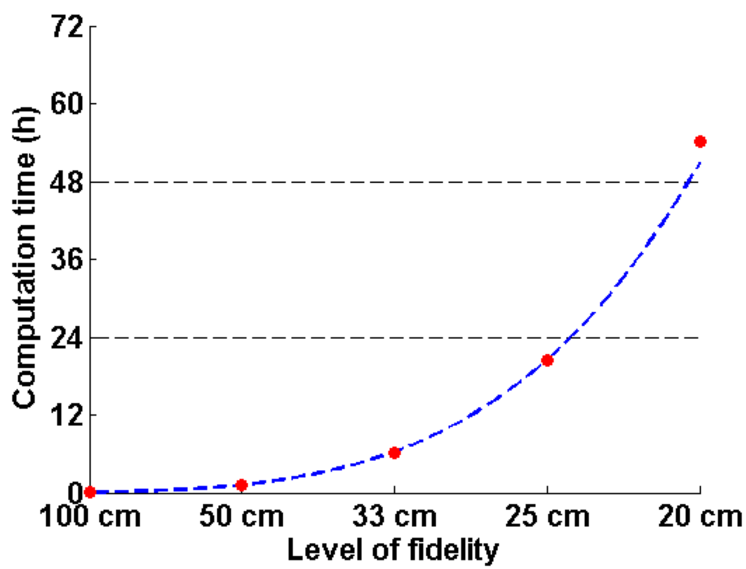
- Multi-fidelity: same physical phenomenon
  ➔ several models with various accuracy

Low fidelity → High fidelity



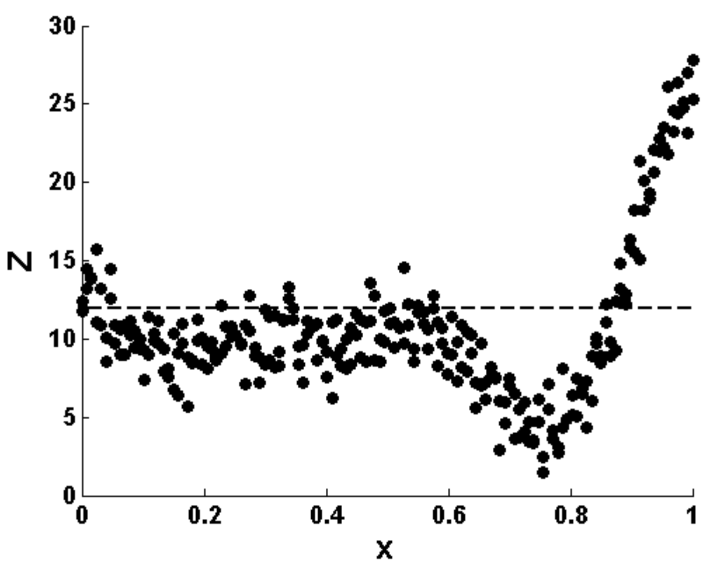
Tunable cost

- Cost of observation: function of the fidelity $C(t)$
  - Cheap simulation, but low fidelity
  - High fidelity simulation, but time-consuming



Stochastic outputs

- Stochastic: same input ➔ different outputs



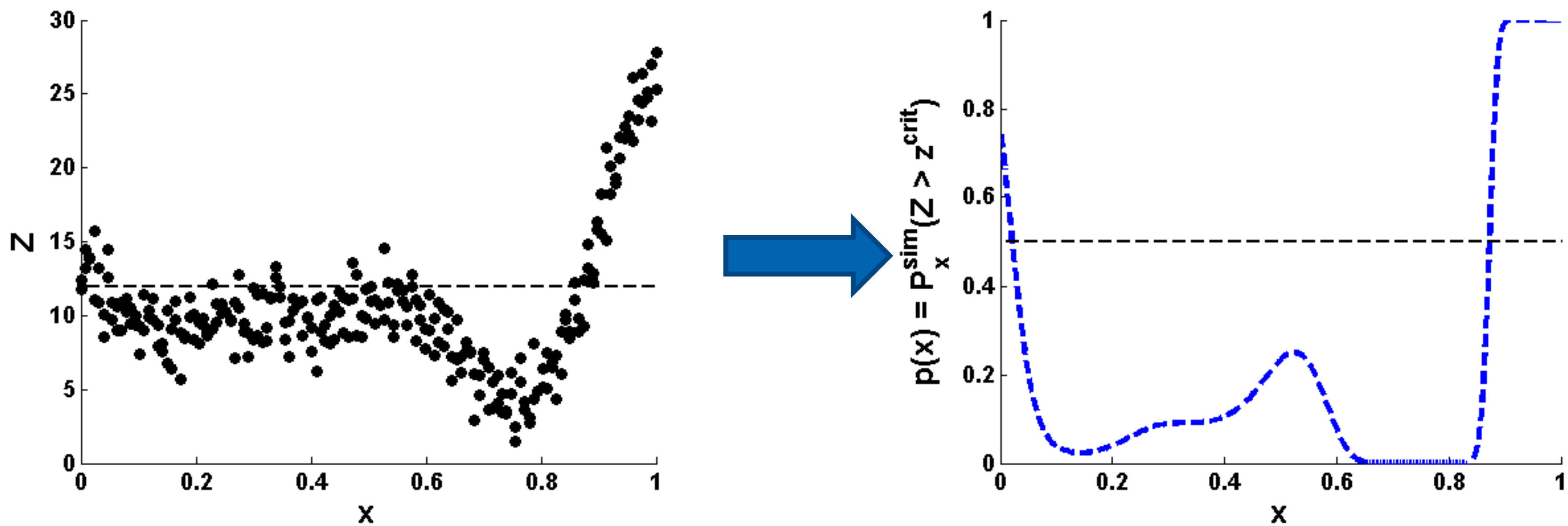
Quantity of Interest

$(x,t) \longrightarrow$ **Simulator** $\longrightarrow Z \longrightarrow \mathbb{P}^{\text{sim}}_{x,t^{HF}}\left(Z > z^{crit}\right)$

- Probability of exceeding a critical threshold z^{crit}

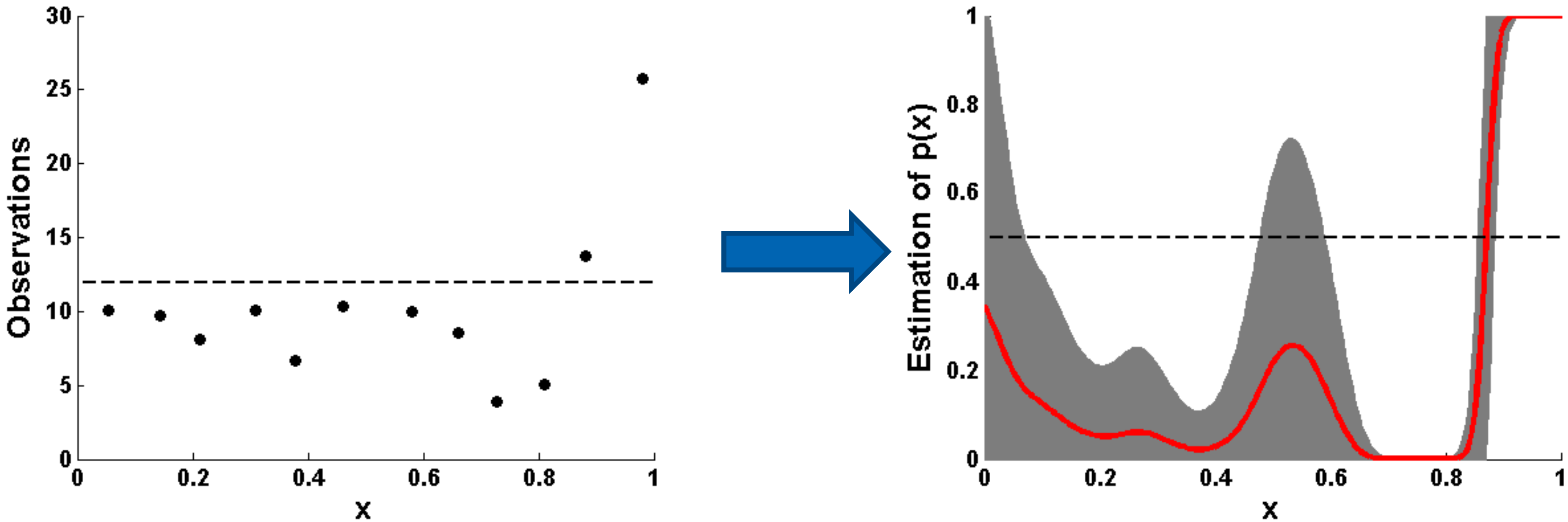
$$p(x, t^{HF}) = \mathbb{P}^{sim}_{x,t^{HF}}\left(Z > z^{crit}\right)$$

  - t^{HF}: the highest-fidelity level



Design of experiments

- Goal: selecting $(x_1, t_1), \ldots, (x_n, t_n)$ to estimate the function p with a minimal cost $C(t_1) + \cdots + C(t_n)$
  - Observations $(x_i, t_i; z_i)_{1 \le i \le n}$ ➔ Estimation $\hat{p}_n$ of p



Sequential design of experiments

- Goal: selecting $(x_1, t_1), \dots, (x_n, t_n)$ to estimate the function p with a minimal cost $C(t_1) + \cdots + C(t_n)$

- Sequential design
  - use the n first observations to select the $(n+1)^{th}$ observation

Outline

1. Introduction
2. Sequential design of experiments
3. Academic example

Model of the simulator

- Prior distribution:
  - Output Z at x, t follows a normal distribution
    $$Z|\xi \sim \mathcal{N}\big(\xi(x,t), \lambda(x,t)\big)$$
  - Mean function ξ: Gaussian process
    $$\xi \sim \mathcal{GP}(m, k)$$
- Posterior distribution: kriging
  - Mean function $\xi|\chi_n$
    $$\xi|\chi_n \sim \mathcal{GP}(m_n, k_n)$$

Auto-regressive multi-fidelity covariance

- Mean function

$$\xi(x,t) = \begin{cases} \xi_{LF}(x) & \text{if } t = 1 \\ \rho\xi_{LF}(x) + \delta(x) & \text{if } t = 2 \end{cases}$$

  - ξ_{LF}: low-fidelity simulator
  - $\xi_{HF} = \rho\xi_{LF} + \delta$: high-fidelity simulator, linked to the low-fidelity by a linear relationship

➔ Covariance function

$$k\big((x,t),(x',t')\big) = \begin{cases} k_{LF}(x,x') & if\ t = t' = 1 \\ \rho k_{LF}(x,x') & if\ t \neq t' \\ \rho^2 k_{LF}(x,x') + k_\delta(x,x') & if\ t = t' = 2 \end{cases}$$

  - k_{LF}: covariance of the low-fidelity simulator
  - k_δ: covariance of the difference between high- and low-fidelity levels
  - ρ: correlation between the low- and high-fidelity levels

[Kennedy and O'Hagan, 2000], [Le Gratiet and Cannamela, 2015]

Non-stationary multi-fidelity covariance

- Mean function

$$\xi(x,t) = \xi_0(x) + \epsilon(x,t)$$

  - ξ_0: ideal simulator (Ex: mesh size = 0)
  - ϵ: system error between ideal and real simulators at t

➔ Covariance function

$$k\big((x,t),(x',t')\big) = k_0(x,x') + r(t,t') \cdot k_\epsilon(x,x')$$

  - k_0: covariance of ξ_0
  - k_ϵ: covariance of ϵ according to x
  - r: rules the decrease of the error
- [Picheny and Ginsbourger, 2013], [Tuo et al., 2014]

Probability of exceeding the critical threshold

- Probability of exceeding the critical threshold:

$$p(x,t) = \mathbb{P}^{sim}_{x,t}\left(Z > z^{\text{crit}} \middle| \chi_n\right) = \Phi\left(\frac{\xi(x,t) - z^{crit}}{\sqrt{\lambda(x,t)}}\right)$$

- First and second moments
  - Expectation: $\mathbb{E}_n[p(x,t)] = \Phi(u_n(x,t)) = \hat{p}_n(x,t)$
    Variance: $\mathbb{V}ar_n[p(x,t)] = \Phi_2(u_n(x,t), u_n(x,t); r_n(x,t)) - \Phi^2(u_n(x,t))$
  - $u_n(x,t) = \frac{m_n(x,t) - z^{crit}}{\sqrt{\sigma_n^2(x,t) + \lambda(x,t)}}$ $\quad r_n(x,t) = \frac{\sigma_n^2(x,t)}{\sigma_n^2(x,t) + \lambda(x,t)}$
  - $\sigma_n^2(x,t) = k_n\big((x,t),(x,t)\big)$
  - Φ: cumulative distribution function (cdf) of the normal distribution
    Φ_2: cdf of the bivariate normal distribution

Measure of uncertainty

- Measure of uncertainty

$$H_n = \mathbb{E}_n[\|\hat{p}_n(\cdot, t^{HF}) - p(\cdot, t^{HF})\|^2] = \int_{\mathbb{X}} \mathbb{V}ar_n[p(x, t^{HF})]dx$$

  - $\mathbb{L}^2$-norm of the error of the estimator at the highest level of fidelity

Stepwise Uncertainty Reduction

- Measure of uncertainty

$$H_n = \mathbb{E}_n[\|\hat{p}_n(\cdot, t^{HF}) - p(\cdot, t^{HF})\|^2] = \int_{\mathbb{X}} \mathbb{V}ar_n[p(x, t^{HF})]dx$$

- Stepwise uncertainty reduction algorithm

$$(x_{n+1}, t_{n+1}) = \underset{x,t}{\operatorname{argmin}}\{\mathbb{E}_n[H_{n+1} | X_{n+1} = x, T_{n+1} = t]\}$$

- [Vazquez and Bect, 2009]

Stepwise Uncertainty Reduction

- Measure of uncertainty

$$H_n = \mathbb{E}_n[\|\hat{p}_n(\cdot, t^{HF}) - p(\cdot, t^{HF})\|^2] = \int_{\mathbb{X}} \mathbb{V}ar_n[p(x, t^{HF})]dx$$

- Stepwise uncertainty reduction algorithm

$$(x_{n+1}, t_{n+1}) = \underset{x,t}{\operatorname{argmin}}\{\mathbb{E}_n[H_{n+1} | X_{n+1} = x, T_{n+1} = t]\}$$

- Analytical expression

$$\mathbb{E}_n[H_{n+1} | X_{n+1} = x, T_{n+1} = t]$$
$$= \int_{\mathbb{X}} \left[\Phi_2(u_n(y, t^{HF}), u_n(y, t^{HF}); r_n(y, t^{HF})) - \Phi_2\left(u_n(y, t^{HF}), u_n(y, t^{HF}); \tilde{r}_n\left((\boldsymbol{x}, \boldsymbol{t}), (y, t^{HF})\right)\right)\right] dy$$

- $$\tilde{r}_n\left((x, t), (y, t^{HF})\right) = \frac{k\left((x,t),(y,t^{HF})\right)^2}{\left(\sigma_n^2(x,t)+\lambda(x,t)\right)\cdot\left(\sigma_n^2(y,t^{HF})+\lambda(y,t^{HF})\right)}$$

Cost of observation

- Different costs $C(x,t)$ of observations
  ➔ Trade-off between H_n reduction and cost $C(x,t)$

- [Huang et al. 2006], [Le Gratiet and Cannamela, 2015]: comparison between benefit and cost

Maximum Speed of Uncertainty Reduction

- Maximum Speed of Uncertainty Reduction (MSUR)

$$(x_{n+1}, t_{n+1}) = \underset{x,t}{\operatorname{argmax}} \left\{ \frac{H_n - \mathbb{E}_n[H_{n+1} | X_{n+1} = x, T_{n+1} = t]}{C(x,t)} \right\}$$

  - MSUR = Benefit/Cost
  - Adaptable for any measure of uncertainty H_n
  - If C is constant ➔ equivalent to SUR algorithm

Simplification

- If the cost depends only on the level $C(x,t) = C(t)$

➔ Algorithm: separate optimization of the point x and the level t

1. $x^\star(t) = \underset{x}{\mathrm{argmin}}\{\mathbb{E}_n[H_{n+1}|X_{n+1} = x, T_{n+1} = t]\}$
2. $t_{n+1} = \underset{t}{\mathrm{argmax}}\left\{\frac{H_n - \mathbb{E}_n[H_{n+1}|X_{n+1}=x^\star(t), T_{n+1}=t]}{C(t)}\right\}$
3. $x_{n+1} = x^\star(t_{n+1})$

Outline

1. Introduction
2. Sequential design of experiments
3. Academic example

Outline

1. Introduction

2. Sequential design of experiments

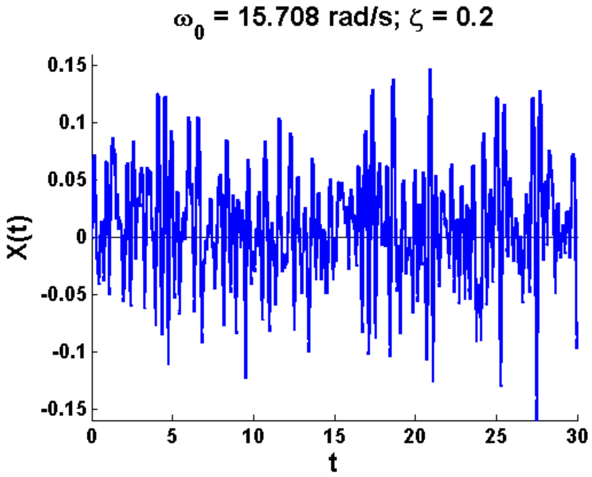
3. Academic example
   a. Presentation of the example
   b. Comparison of sequential designs of experiments

Stochastic damped harmonic oscillator

- Consider a damped harmonic simulator with random drive

$$\ddot{X}(t) + 2\zeta\omega_0\dot{X}(t) + \omega_0^2 X(t) = W(t)$$

  - ω_0: the undamped angular frequency
  - ζ: the damping ratio
  - W: a Brownian motion, with spectral density $S = 1$
  - Initial conditions: $X(t=0) = 0$, $\dot{X}(t=0) = 0$



- [Au and Beck, 2001]

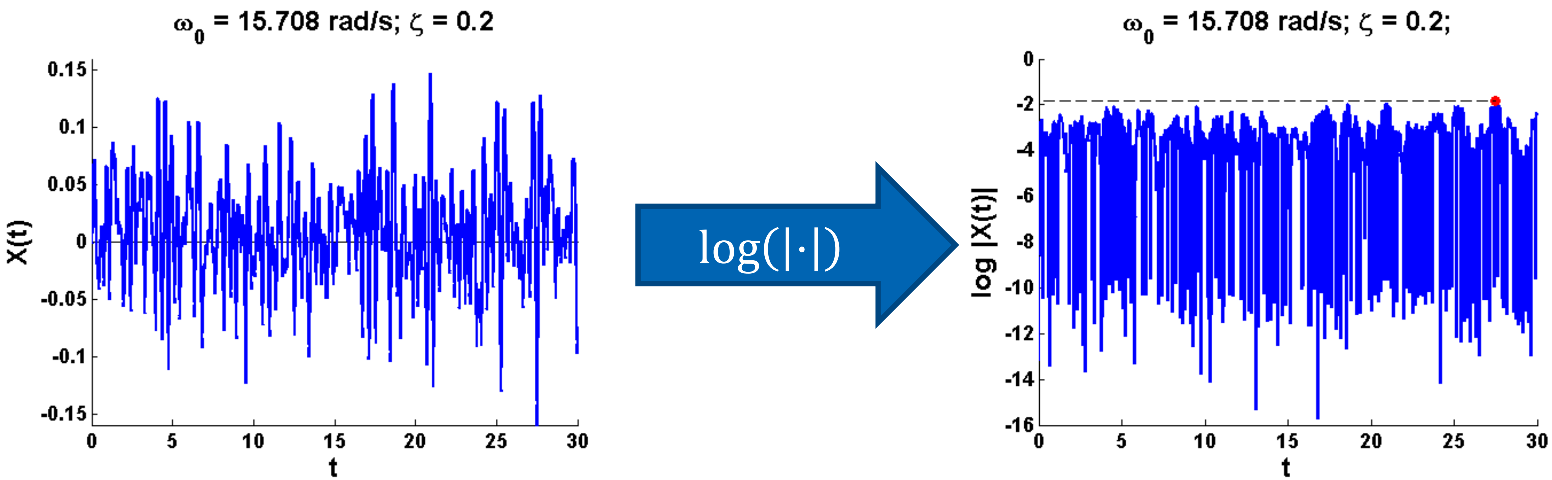
Simulator

- Consider a damped harmonic simulator with random drive

$$\ddot{X}(t) + 2\zeta\omega_0\dot{X}(t) + \omega_0^2 X(t) = W(t)$$

- Ideal simulator

$$F: (\omega_0, \zeta) \mapsto \max_{0 \leq t \leq t^{\mathrm{end}} = 30} \{\log|X(t)|\}$$



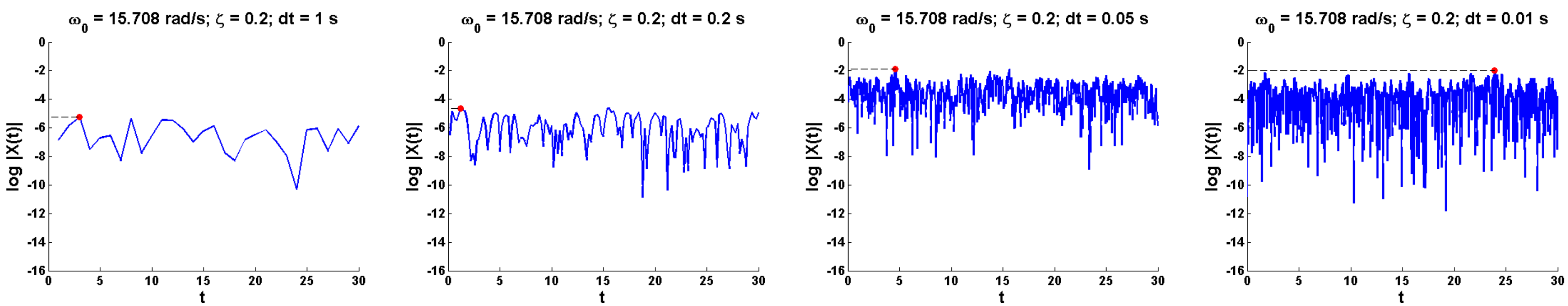
Academic multi-fidelity simulator

- Approximation by an *explicit Exponential Euler Scheme*

$$X(n \cdot dt) \approx \tilde{X}_n$$

- Multi-fidelity simulator

$$f : (\omega_0, \zeta, dt) \mapsto \max_{0 \leq n \leq \left\lceil \frac{t^{\mathrm{end}}}{dt} \right\rceil} \left\{ \log \left| \tilde{X}_n \right| \right\}$$

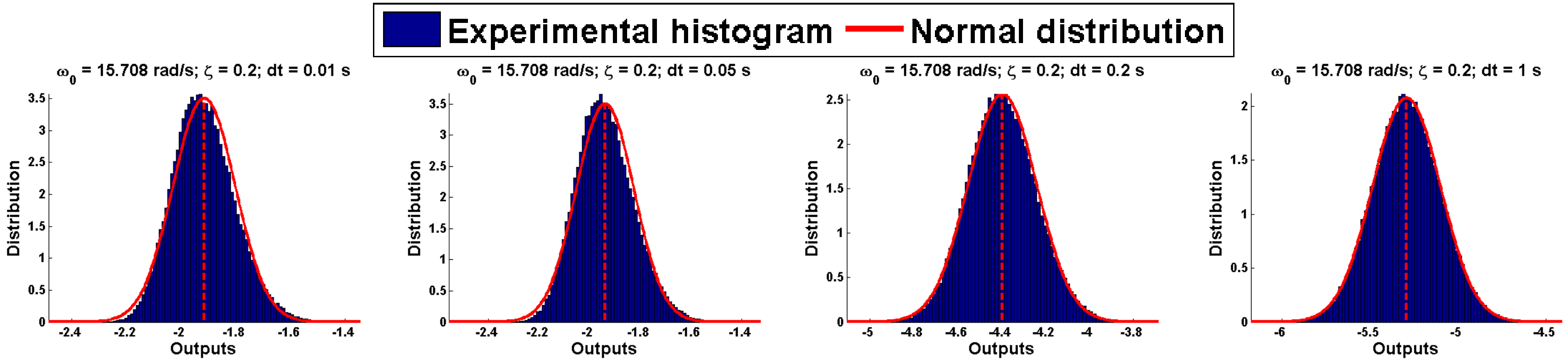


$dt = 1$ s $\quad dt = 0.2$ s $\quad dt = 0.05$ s $\quad dt = 0.01$ s

- [Jentzen and Kloeden, 2009]

Output distributions at a fixed input

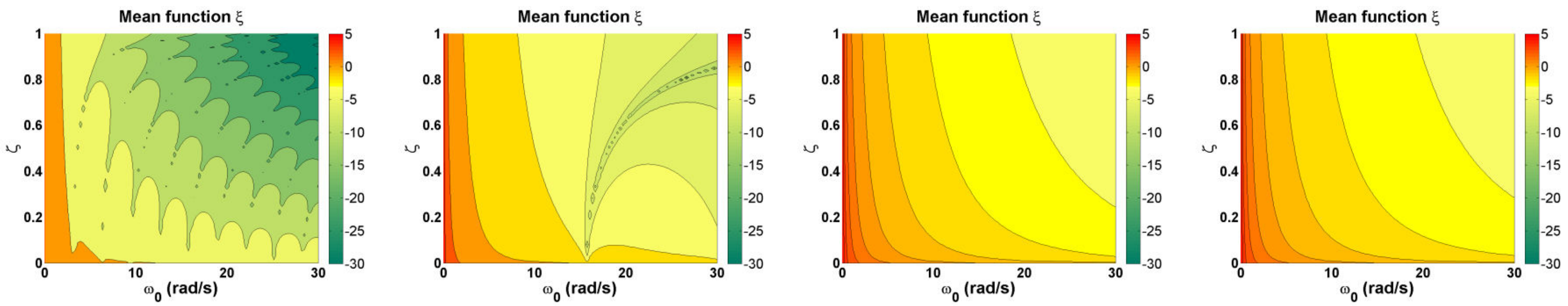
- The output distribution at (ω_0, ζ, dt) can be approximated by a normal distribution
  - 10^5 simulations at $\omega_0 = 15.708\ \mathrm{rad}/s$ and $\zeta = 0.2$



$dt = 1$ s | $dt = 0.2$ s | $dt = 0.05$ s | $dt = 0.01$ s

Mean function

- Mean function ξ
  - 10^5 simulations
  - $0 \leq \omega_0 \leq 30\ \mathrm{rad/s},\ 0 \leq \zeta \leq 1$ - Grid: 100×100



$dt = 1$ s

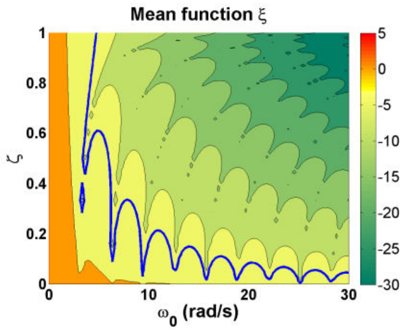
$dt = 0.2$ s

$dt = 0.05$ s

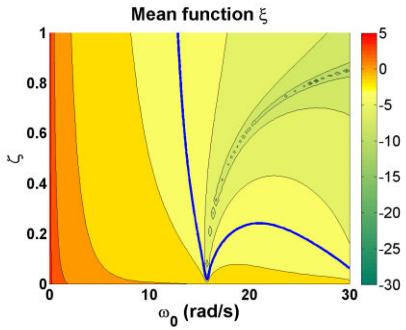
$dt = 0.01$ s

Threshold

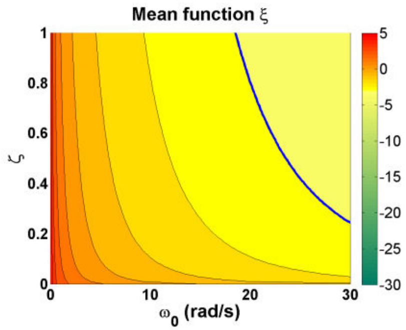
- Critical threshold $z^{crit} = -3$



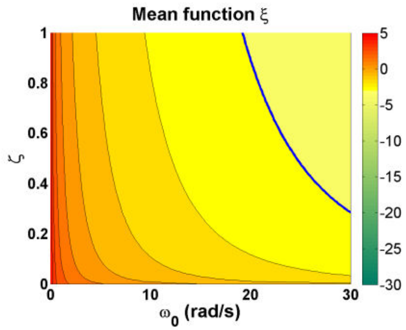
$dt = 1$ s



$dt = 0.2$ s



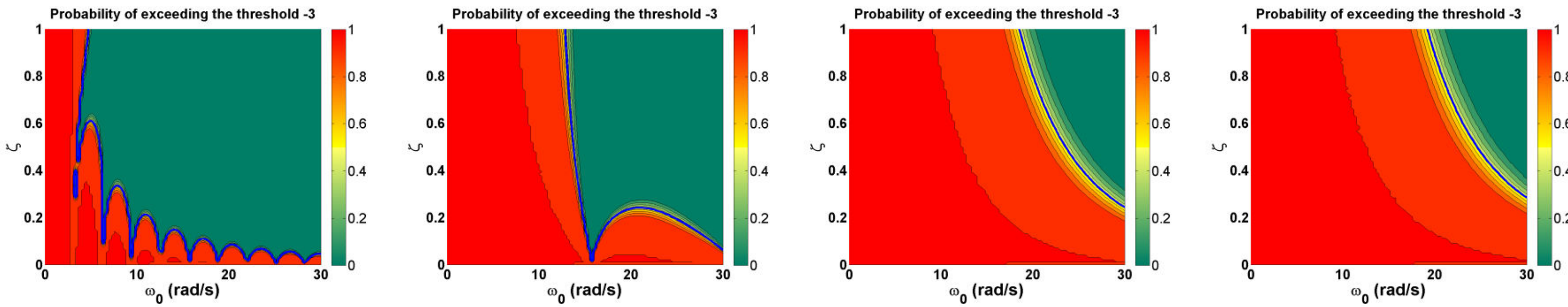
$dt = 0.05$ s



$dt = 0.01$ s

Probability of exceeding the threshold

- True probability of exceeding the threshold $p(\omega_0, \zeta; dt)$

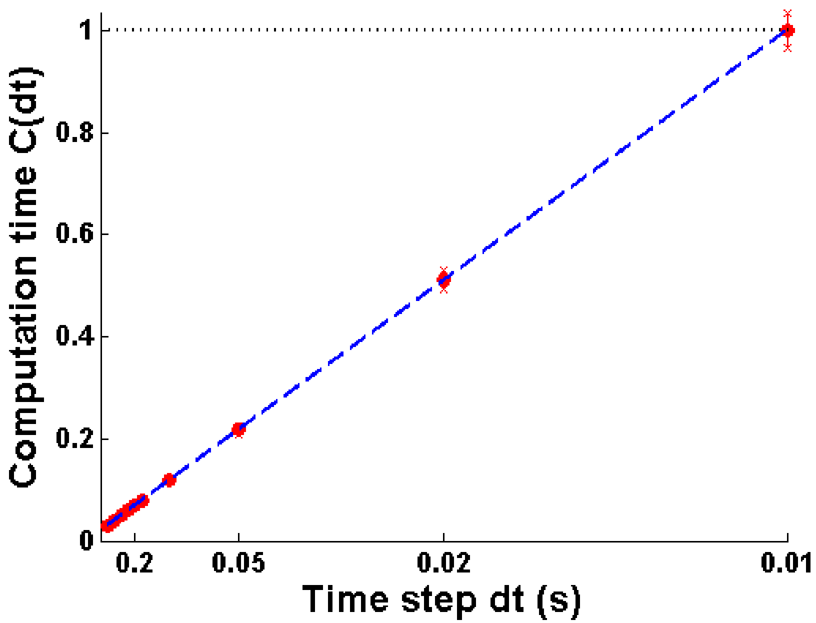


$dt = 1$ s $dt = 0.2$ s $dt = 0.05$ s $dt = 0.01$ s

Computation time

- Computation time $C(dt)$: linear in $1/dt$.

Time step dt	1 s	0.2 s	0.05 s	0.01 s
CPU Time (ms)	0.799 ms	1.85 ms	5.78 ms	26.7 ms
Cost function $C(dt)$	0.030 ¤	0.069 ¤	0.217 ¤	1.00 ¤



Outline

1. Introduction
2. Sequential design of experiments
3. Academic example
   a. Presentation of the example
   b. Comparison of sequential designs of experiments

Definition of the problem

- Target: probability of exceeding the threshold at the highest level of fidelity $dt = 0.01$ s

- Initial design: Nested LHS on 5 levels

dt (s)	1.00	0.50	0.33	0.25	0.20	0.17	0.10	0.05	0.02	0.01
Nb. points	180	60	20	10	5	0	0	0	0	0

- [Qian, 2009]

- Initial budget: 9.87 ¤
  1 ¤ = cost for 1 observation at the level $dt = 0.01$ s
  = cost for 1.96 observations at the level $dt = 0.02$ s
  = cost for 33.4 observations at the level $dt = 1$ s
  = …

Comparison between designs of experiments

- Initial budget: 9.87 ¤
  Supplementary budget: 10 ¤
- 6 designs of experiments (DoE)
  - 5 Single level DoE
  - Multi-level DoE

Sequential design	Criterion	Nb. Points Initial design	Nb. Points Final design
Single level ($dt = 0.17\ s$)	SUR	275	275+145 = 420
Single level ($dt = 0.10\ s$)	SUR	275	275 + 85 = 360
Single level ($dt = 0.05\ s$)	SUR	275	275 + 46 = 321
Single level ($dt = 0.02\ s$)	SUR	275	275 + 19 = 294
Single level ($dt = 0.01\ s$)	SUR	275	275 + 10 = 285
Multi-level	MSUR	275	275 + ? = ?

Comparison between designs of experiments

- Initial budget: 9.87 ¤
  Supplementary budget: 10 ¤
- 6 designs of experiments (DoE)
  - 5 Single level DoE
  - Multi-level DoE

- Same model:
  - Same covariance function
  - Hyper-parameters estimated on a large design
  - Fixed hyper-parameters during the sequential designs
- Each DoE: 12 repetitions

Criteria of comparison

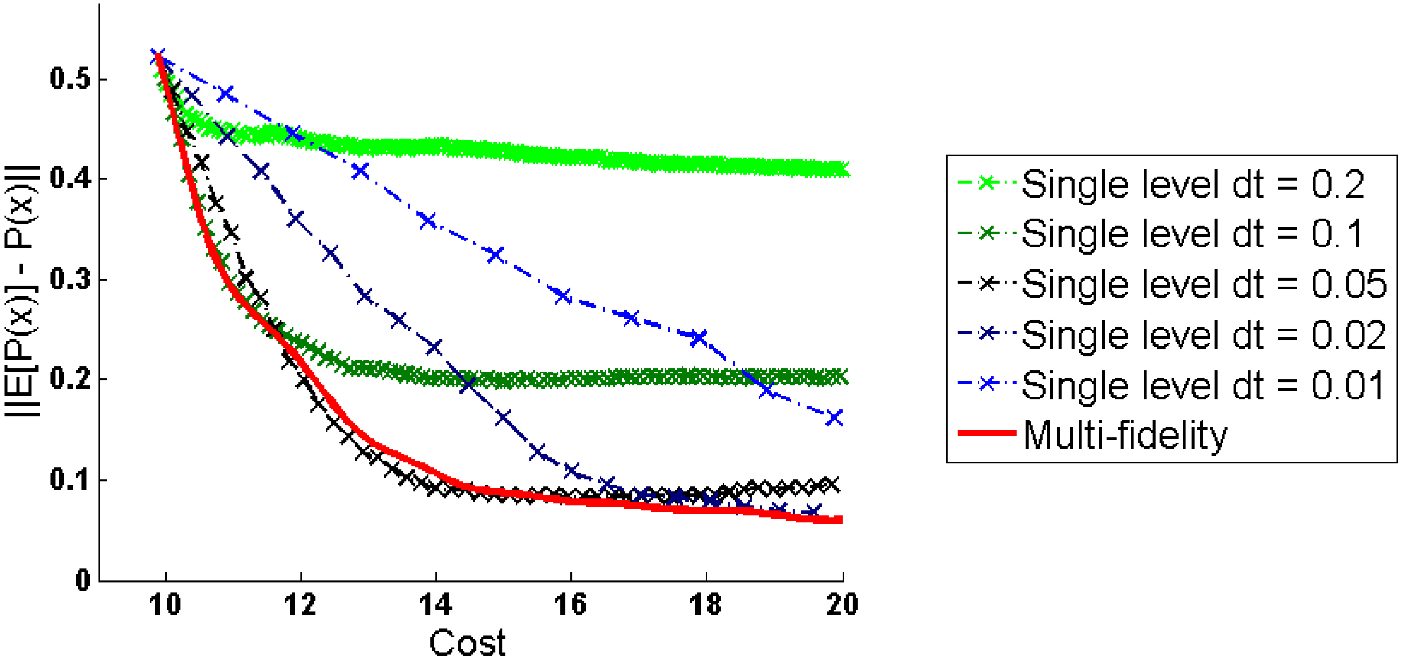
$\mathbb{L}^2$-error on the probability function

$$\sqrt{\int_{[0;30]\times[0;1]} (\hat{p}_n(x,t^{HF}) - p(x,t^{HF}))^2 dx}$$

$$t^{HF} = 0.01\ s$$

Multi-fidelity is better

$\mathbb{L}^2$-error on the probability function



- Low-fidelity levels are biased
  High-fidelity levels are slow
- In this example, multi-fidelity finds the best trade-off

Conclusion

- Goal: sequential design of experiments to estimate probability on stochastic multi-fidelity numerical models

- New SUR criteria to estimate probability of exceeding a threshold on stochastic simulator

- Adaptation to multi-fidelity model ➔ Maximum Speed of Uncertainty Reduction (MSUR)
  - MSUR = (Uncertainty Reduction)/Cost

- Results on an academic example ➔ automatic trade-off between cost and fidelity

Do you have any questions?

Thank you for your attention!

References (1/5): SUR-algorithm

1. Stephen I. N. Kerber. Evaluation of the Ability of Fire Dynamic Simulator to Simulate Positive Pressure Ventilation in the Laboratory and Practical Scenarios. *PhD thesis, University of Maryland*, December 2005. http://hdl.handle.net/1903/3243

2. Emmanuel Vazquez and Julien Bect. A sequential Bayesian algorithm to estimate a probability of failure. *IFAC Proceedings* Volumes, 42(10):546–550, 2009.

3. Clément Chevalier, Julien Bect, David Ginsbourger, Emmanuel Vazquez, Victor Picheny, and Yann Richet. Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set. *Technometrics*, 56(4):455–465, 2014. http://dx.doi.org/10.1080/00401706.2013.860918

References (2/5): auto-regressive multi-fidelity

4. Marc C. Kennedy and Antony O'Hagan. Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87(1):1–13, 2000. http://biomet.oxfordjournals.org/content/87/1/1.abstract

5. Loïc Le Gratiet and Claire Cannamela. Cokriging-based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes. *Technometrics*, 57(3):418–427, 2015. http://dx.doi.org/10.1080/00401706.2014.928233

References (3/5): non-stationary multi-fidelity

6. Victor Picheny and David Ginsbourger. A nonstationary space-time Gaussian process model for partially converged simulations. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):57–78, 2013. http://dx.doi.org/10.1137/120882834

7. Rui Tuo, C. F. Jeff Wu, and Dan Yu. Surrogate modeling of computer experiments with different mesh densities. *Technometrics*, 56(3):372–380, 2014. http://dx.doi.org/10.1080/00401706.2013.842935

References (4/5): multi-fidelity sequential design

8. Deng Huang, Theodore T. Allen, William I. Notz, and R. A. Miller. Sequential kriging optimization using multiple-fidelity evaluations. *Structural and Multidisciplinary Optimization*, 32(5):369–382, 2006. http://dx.doi.org/10.1007/s00158-005-0587-0

9. Loïc Le Gratiet and Claire Cannamela. Cokriging-based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes. *Technometrics*, 57(3):418–427, 2015. http://dx.doi.org/10.1080/00401706.2014.928233

10. Lawrence D. Phillips and Carlos A. Bana e Costa. Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Annals of Operations Research*, 54(1):51-68, 2007. http://dx.doi.org/10.1007/s10479-007-0183-3

References (5/5): academic example

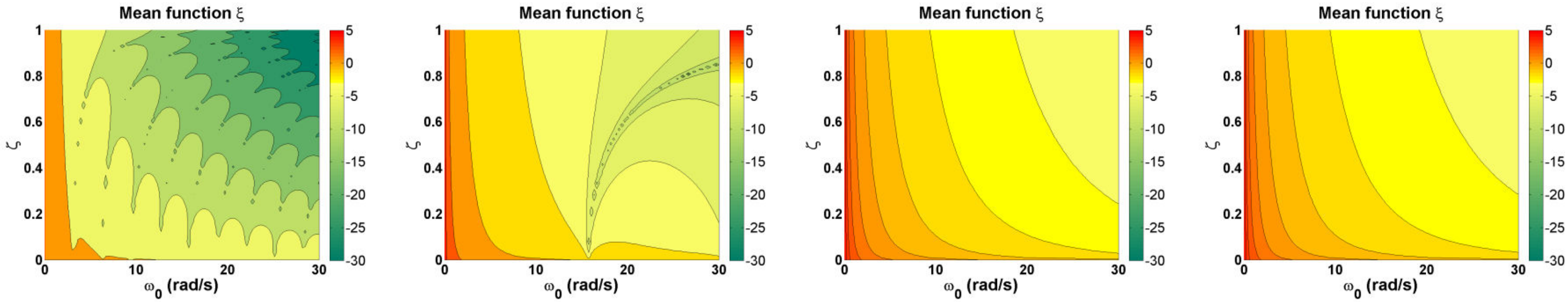
11. Siu-Kui Au and James L. Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277, 2001. http://www.sciencedirect.com/science/article/pii/S0266892001000194

12. Arnulf Jentzen and Peter E. Kloeden. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space–time noise. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 465(2102): 649–667, 2009. http://rspa.royalsocietypublishing.org/content/465/2102/649

13. Peter Z. G. Qian. Nested latin hypercube designs. *Biometrika*, 96(4):957–970, 2009. http://biomet.oxfordjournals.org/content/early/2009/09/22/biomet.asp045.abstract

Mean and variance functions

- Mean function ξ and variance function λ



$dt = 1$ s $dt = 0.2$ s $dt = 0.05$ s $dt = 0.01$ s