A Bayesian Approach to Optimal Sequential Experimental Design using Approximate Dynamic Programming

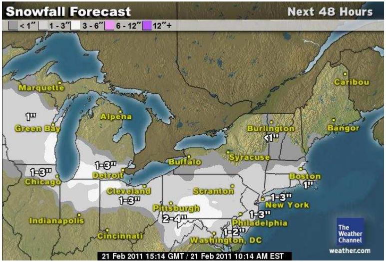
Xun Huan

Massachusetts Institute of Technology
Department of Aeronautics and Astronautics

February 13, 2013

Motivation

- Experimental data are crucial for developing and refining models:
  - parameter inference
  - prediction
  - model selection
- “Optimally”-chosen experiments lead to substantial savings



(Sources: left—Argonne National Labs; right—www.weather.com)

Challenges

Optimal experimental design (OED):

- open-loop design:
  - theory for linear models well developed [Atkinson 92]
  - analytical results not available for nonlinear designs, and numerical approaches often rely on linearization, Gaussian approximation, and "best guess" parameters [Box 59, Ford 89, Chaloner 95, Chu 08]
  - general design framework free of these assumptions [Müller 98] difficult to solve numerically [Ryan 03, van den Berg 03, Terejanu 12]
  - open-loop is sub-optimal for multiple experiments!
- closed-loop design:
  - mostly greedy approach (sub-optimal) [Cavagnaro 12, Solonen 12]
  - dynamic programming truly optimal (POMDP formulation [Chong 09]) but computationally feasible for only "simple" applications [Brockwell 03, Christen 03, Wathen 06, Müller 07]

Scope and Objective

Scope:

Optimal closed-loop design via dynamic programming for

- nonlinear and computationally intensive (PDE-based) models
- continuous design and data spaces of multiple dimensions
- the purpose of parameter inference, using an information measure in the objective

Objective:

develop numerical tools that find the optimal closed-loop design via dynamic programming in a computationally feasible manner

Outline

1. Formulations and Numerical Methods
2. Proof-of-Concept Example
3. Future Work

Outline

1. Formulations and Numerical Methods
2. Proof-of-Concept Example
3. Future Work

Experimental Goal

Interested in experiments whose data are valuable for parameter inference, taking a Bayesian design approach

$$\overbrace{f(\theta|y,d)}^{\textit{posterior}} = \frac{\overbrace{f(y|\theta,d)}^{\textit{likelihood}}\overbrace{f(\theta|d)}^{\textit{prior}}}{\underbrace{f(y|d)}_{\textit{evidence}}}$$

θ — parameters of interest
y — noisy measurements or data
d — design variables or controls

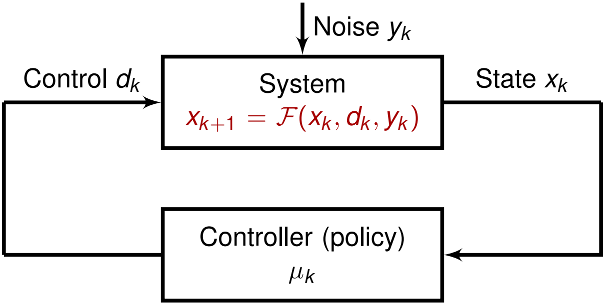
Closed-Loop Dynamic Programming Formulation

State: posterior PDFs $x_k = f(\theta|I_k)$ where $I_k = \{d_0, y_0, \ldots, d_{k-1}, y_{k-1}\}$

Control: $d_k = \mu_k(x_k) \in \mathcal{U} \subseteq \mathbb{R}^{n_u}$; $\pi = \{\mu_0, \mu_1, \ldots, \mu_{N-1}\}$ is a policy

Noise: $y_k \in \mathbb{R}^{n_w}$ distributed according to likelihood $f(y_k|\theta, d_k)$

System: Bayes' Theorem $\quad x_{k+1} = \mathcal{F}(x_k, d_k, y_k) = \dfrac{f(y_k|\theta, d_k)x_k}{f(y_k|d_k, I_k)}$



Closed-Loop Dynamic Programming Formulation

Finite-horizon, discrete-time, perfect state information

Value functions (Bellman equation):

$$J_k(x_k) = \max_{d_k} \mathbb{E}\left[g_k(x_k, d_k, y_k) + J_{k+1}\left(\mathcal{F}(x_k, d_k, y_k)\right)\right]$$

$$J_N(x_N) = \int_{\mathcal{H}} x_N \ln\left[\frac{x_N}{x_0}\right] d\theta$$

for $k = 0, \ldots, N-1$; policy implicitly in arg-max $d_k^* = \mu_k(x_k)$

Objective: expected total reward

$$\mathbb{E}_{y_0,\ldots,y_{N-1}}\left[J_N\left(x_N\right) + \sum_{k=0}^{N-1} g_k\left(x_k, \mu_k(x_k), y_k\right)\right]$$

Numerical Tools

Evaluating terminal reward (information gain):

- need to evaluate the expectation of Kullback-Leibler divergence
- possible numerical approaches: Laplace approximation, binning, quadrature, kernel density estimation—can have large errors or poor scaling with dimension [Long 12, Guest 09, Sebastiani 97, Khan 07]
- we use a doubly-nested Monte Carlo estimator [Ryan 03]

Numerical Tools

Stochastic optimization:

- need to optimize value of noisy Monte Carlo estimator
- Stochastic approximation (e.g. Robbins-Monro [Robbins 51]): steepest-descent-like using an unbiased gradient estimator, difficult to select stepsize
- Sample average approximation [Shapiro 91, Kleywegt 02]: fix random variables at a seed, optimize resulting deterministic instance

■ **Work-to-date:** developed gradient expressions for stochastic approximation and sample average approximation, conducted empirical performance studies on open-loop design problems [Huan 13b, Huan 13a]

Numerical Tools

Polynomial chaos (surrogate model): [Wiener 38, Ghanem 91, Le Maître 10]
Replace forward model with polynomial expansions:

$$G(\xi) \approx \sum_{|\mathbf{i}|_1=0}^{p} G_{\mathbf{i}} \Psi_{\mathbf{i}}(\xi_1, \xi_2, \ldots, \xi_n)$$

- coefs $G_{\mathbf{i}}$, basis random variables ξ_j, orthogonal polynomials $\Psi_{\mathbf{i}}$
- non-intrusive approach to compute expansion coefficients via sparse pseudo-spectral approximation [Conrad 13]

■ **Work-to-date:** open-loop design using polynomial approximation of the forward model, over the product space of the uncertain parameters and the design variables [Huan 13b, Huan 13a]

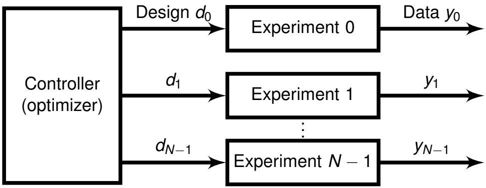
Two remaining issues (future work):

1. how to numerically represent the states $x_k = f(\theta|I_k)$?
2. how to deal with the "combinatorial explosion" and lack of analytical solutions in the dynamic programming framework?

Open-Loop and Greedy Designs

For comparisons, we will also consider other design approaches

Open-loop design:
No feedback of data, clump all experiments in a batch and perform one-stage closed-loop design



Greedy policy:
Update after each experiment and then perform open-loop design for the next experiment only

Outline

Linear-Gaussian Model

$$y_k = d_k\theta + \epsilon_k$$

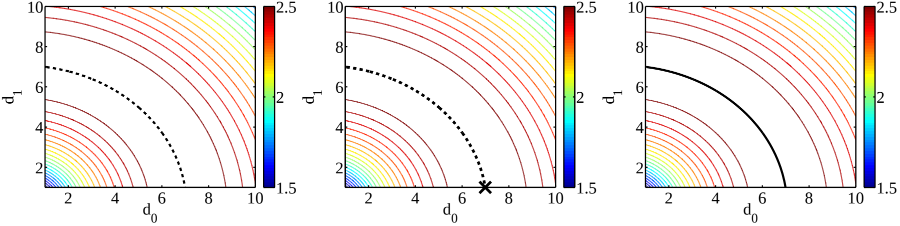
- prior: $\theta \sim \mathcal{N}(s_0, \sigma_0^2) = \mathcal{N}(7, 3^2)$
- noise: $\epsilon_k \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$
- linear-Gaussian problem: conjugate family, posteriors (i.e., all states) will be Gaussian:

$$x_{k+1} = \left(s_{k+1}, \sigma_{k+1}^2\right) = \left(\frac{\frac{y_k/d_k}{\sigma_\epsilon^2/d_k^2} + \frac{s_k}{\sigma_k^2}}{\frac{1}{\sigma_\epsilon^2/d_k^2} + \frac{1}{\sigma_k^2}}, \frac{1}{\frac{1}{\sigma_\epsilon^2/d_k^2} + \frac{1}{\sigma_k^2}}\right)$$

- control: $d_k \in [1, 10]$
- stage cost quadratic in control: $g_k = -0.01 d_k^2$
- two experiments ($N = 2$)

Linear-Gaussian Example: $\sigma^2_{\epsilon,k} = 1$

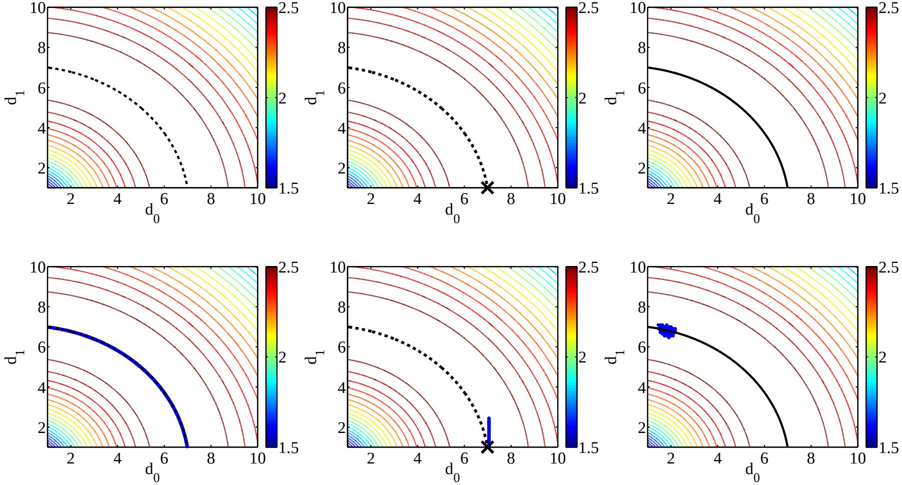
$\sigma^2_{\epsilon,k} = 1$ (constant)



open-loop greedy dynamic programming

Linear-Gaussian Example: $\sigma^2_{\epsilon,k} = 1$

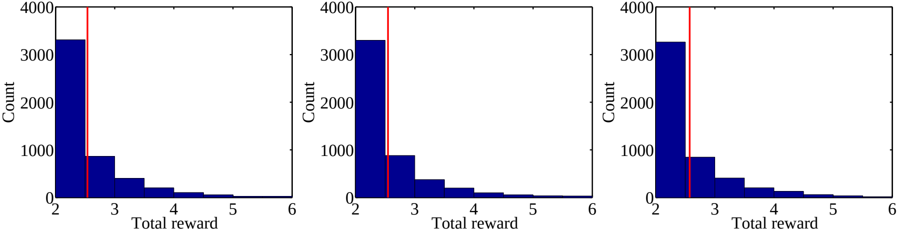
$\sigma^2_{\epsilon,k} = 1$ (constant)



open-loop greedy dynamic programming

Linear-Gaussian Example: $\sigma^2_{\epsilon,k} = 1$

$\sigma^2_{\epsilon,k} = 1$ (constant)



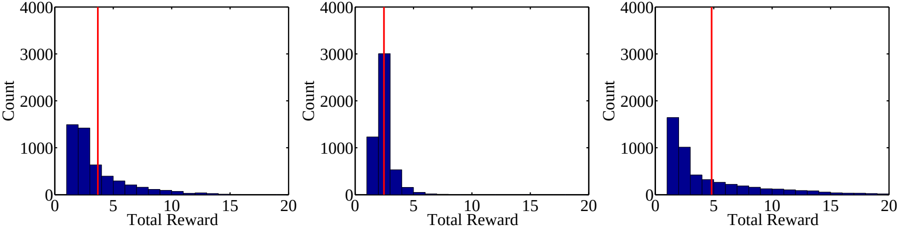
open-loop (mean 2.54)

greedy (mean 2.55)

dynamic programming (mean 2.57)

Linear-Gaussian Example: $\sigma^2_{\epsilon,k} = (\frac{\sqrt{7}}{5} d_k^{s_k - 6.5})^2$

$\sigma^2_{\epsilon,k} = (\frac{\sqrt{7}}{5} d_k^{s_k - 6.5})^2$ (non-constant)



open-loop (mean 3.68) | greedy (mean 2.78) | dynamic programming (mean 4.94)

Outline

State Representation

Two remaining issues (future work):

1. how to numerically represent the states $x_k = f(\theta|I_k)$?
2. how to deal with the "combinatorial explosion" and lack of analytical solutions in the dynamic programming framework?

State representation:
state variables x_k are general (non-Gaussian) PDFs

- Gaussian mixture model
- sequential Monte Carlo (particle filtering) [Ristic 04]
- exponential family principle component analysis [Roy 05]
- random variable mapping $\mathcal{G}$ where [Moselhy 12]

$$\theta|y_{1:k}, d_{1:k} = \mathcal{G}(\theta|y_{1:k-1}, d_{1:k-1})$$

- will select an existing method that works well

Approximate Dynamic Programming

Approximate Dynamic Programming [Bertsekas 96, Sutton 98, Powell 07]

Problem simplifications:
open-loop, greedy, myopic, open-loop feedback control, rolling-horizon, discretization and aggregation [Bertsekas 00]

Value function approximation:
parameterized linear architecture

$$\tilde{J}_k(x_k) = \sum_i r_{k,i} z_{k,i}(x_k)$$

- $z_{k,i}$ features, selection crucial and heavily relies on heuristics
- $r_{k,i}$ weights, trained via regression, sampling, quadrature, etc
- one-step lookahead policy: perform a step of dynamic programming before using the approximation function
- rollout algorithm: one-step lookahead with approximation being from a heuristic policy, equivalent to one step of policy iteration

Approximate Dynamic Programming

One-step lookahead backward induction:

$$\begin{aligned} \tilde{J}_N(x_N) &= \Pi J_N(x_N) \\ \tilde{J}_k(x_k) &= \Pi \max_{d_k} \mathbb{E}[g_k(x_k, d_k, y_k) + \tilde{J}_{k+1}(\mathcal{F}(x_k, d_k, y_k))] \end{aligned}$$

- potential exponential error buildup
- Π can use e.g. regression, but what state measure to use?

■ **Work-to-date**: one-step lookahead backward induction used for linear-Gaussian model

Forward trajectory simulation: [Powell 07]

- simulates trajectories from current approx functions (exploitation)
- update approximation from trajectories (e.g., temporal differencing [Sutton 88])
- flexibility in stopping or continuing refinement
- issue: exploration vs. exploitation
  - heuristic exploration techniques (e.g., ϵ-greedy [Singh 00])
  - reflect uncertainty of states (e.g., assign density on their values)

Approximate Dynamic Programming

Iterative forward-backward sweeps:

- obtain sample paths, construct approximation functions using backward induction based on these state measure approximations
- iterative “batching” allows the use of more efficient techniques such as quadrature
- can use prior knowledge on state space to form the initial set of value function approximations

Sequential Bayesian inference structure:

- reacheable state space can be narrowed down based on the problem structure
- depends on a good choice of state representation
- example: linear-Gaussian variance state component follows

$$\sigma_{k+1}^2 = \frac{1}{\frac{1}{\sigma_\epsilon^2/d_k^2} + \frac{1}{\sigma_k^2}}$$

Approximate Dynamic Programming

Q-factors: [Watkins 89, Watkins 92]

$$Q_k(x_k, d_k) \equiv \mathbb{E}\left[g_k(x_k, d_k, y_k) + J_{k+1}\left(\mathcal{F}(x_k, d_k, y_k)\right)\right]$$

Bellman equation:

$$Q_k(x_k, d_k) = \mathbb{E}\left[g_k(x_k, d_k, y_k) + \max_{d_{k+1}} Q_{k+1}\left(\mathcal{F}(x_k, d_k, y_k), d_{k+1}\right)\right]$$

- model-free operation once Q_k are available

$$\mu_k^*(x_k) = \arg\max_{d_k} Q_k(x_k, d_k)$$

- Q-factor approximation with $\hat{Q}_k$, may optimize analytically
- sparse quadrature may now be used for $\mathbb{E}$, jointly over θ and an design-independent version of y_k
- $\Pi\hat{\mathbb{E}}$ now unbiased (c.f. $\Pi \max \hat{\mathbb{E}}$ is biased)
- drawback: input of Q_k has dimension $\dim(x_k) + \dim(d_k)$

Horizon Changes

Intended stopping rule: (upon satisfactory information gain)

$$J_k(x_k) = \begin{cases} \max\left\{\max_{d_k} \mathbb{E}\left[g_k(x_k, d_k, y_k) + J_{k+1}(x_{k+1})\right], \int_{\mathcal{H}} \ln\left[\frac{x_k}{x_0}\right] x_k \, d\theta\right\} & \text{if } x_k \neq T \\ 0 & \text{if } x_k = T \end{cases}$$

$$J_N(x_N) = \begin{cases} \int_{\mathcal{H}} \ln\left[\frac{x_N}{x_0}\right] x_N \, d\theta & \text{if } x_N \neq T \\ 0 & \text{if } x_N = T \end{cases}$$

T is the absorbing terminal state

Unexpected changes to the number of experiments:

- redo dynamic programming for new horizon from current state
- formulations robust to horizon change
  - form stopping problem if probabilities of horizon changes are known
  - “greedification”
  - incremental information gain

$$J_k(x_k) = \max_{d_k}\left\{\mathbb{E}\left[\int_{\mathcal{H}} \ln\left[\frac{x_{k+1}}{x_k}\right] x_{k+1} \, d\theta + J_{k+1}(x_{k+1})\right]\right\}, \qquad J_N = 0$$

Applications

Combustion kinetics:

- choose initial temperature and concentrations to infer reaction kinetic parameters from ignition delay time measurements
- experiments with fixed form and stationary parameters

Diffusion-convection source inversion:

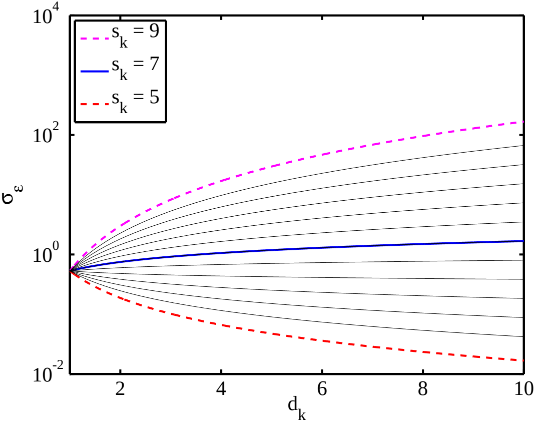
- choose concentration measurement locations and *times* under distance penalties to infer source location and other parameters
- source and parameters may be time-dependent
- experiments with varying form and stationary or non-stationary parameters

Proposed Schedule

Completed	
09/08–08/11	• open-loop design for multiple experiments • gradient-free stochastic optimization methods (SPSA and NM) • combustion application [Huan 13b]
06/11–05/12	• gradient-based stochastic optimization methods (RM and SAA-BFGS) • diffusion source-inversion application [Huan 13a]
11/11–01/13	• closed-loop DP design formulation • analytical solutions and numerical ADP results for linear-Gaussian cases
Future	
02/13–06/13	*ADP*: • additional literature review • implement existing techniques (e.g., Q-factors) • create new efficient and accurate techniques for experimental design
07/13–09/13	*state representation*: • compare and choose methods to represent the states (PDFs) • possible candidates: GMM, SMC, random variable mapping
10/13–01/14	• combine tools together, run cases on the application problems • explore horizon-change and stopping problem
02/14–06/14	• write and defend thesis

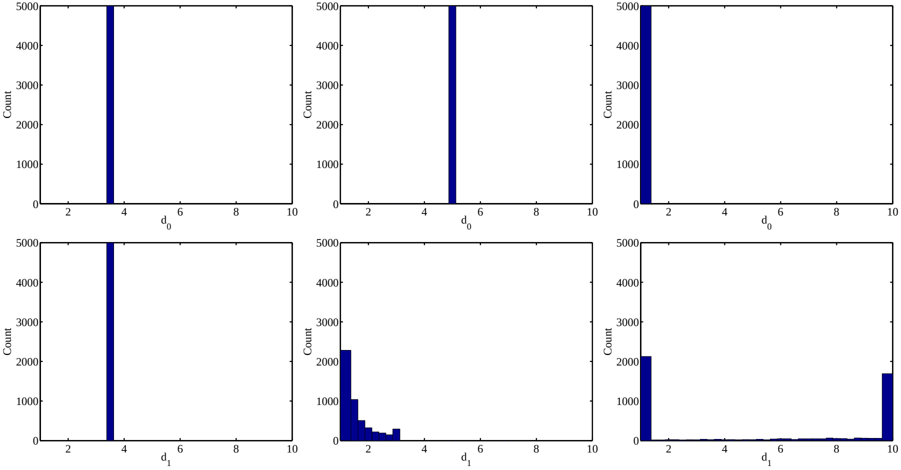
Linear-Gaussian Example: $\sigma^2_{\epsilon,k} = (\frac{\sqrt{7}}{5} d_k^{s_k - 6.5})^2$

$\sigma^2_{\epsilon,k} = (\frac{\sqrt{7}}{5} d_k^{s_k - 6.5})^2$ (non-constant)



Linear-Gaussian Example: $\sigma^2_{\epsilon,k} = (\frac{\sqrt{7}}{5} d_k^{s_k - 6.5})^2$

$\sigma^2_{\epsilon,k} = (\frac{\sqrt{7}}{5} d_k^{s_k - 6.5})^2$ (non-constant)



open-loop greedy dynamic programming

References I

A. C. Atkinson & A. N. Donev.
Optimum experimental designs.
Oxford Statistical Science Series. Oxford University Press, 1992.

Dimitri P. Bertsekas & John N. Tsitsiklis.
Neuro-dynamic programming.
Athena Scientific, 1996.

Dimitri P. Bertsekas.
Dynamic programming and optimal control.
Athena Scientific, 2nd edition, 2000.

G. E. P. Box & H. L. Lucas.
Design of Experiments in Non-Linear Situations.
Biometrika, vol. 46, no. 1/2, pages 77–90, 1959.

Anthony E. Brockwell & Joseph B. Kadane.
A Gridding Method for Bayesian Sequential Decision Problems.
Journal of Computational and Graphical Statistics, vol. 12, no. 3, pages 566–584, 2003.

Daniel R. Cavagnaro, Jay I. Myung, Mark A. Pitt & Janne V. Kujala.
Adaptive Design Optimization: A Mutual Information-Based Approach to Model Discrimination in Cognitive Science.
Neural Computation, vol. 22, no. 4, pages 887–905, 2012.

Kathryn Chaloner & Isabella Verdinelli.
Bayesian Experimental Design: A Review.
Statistical Science, vol. 10, no. 3, pages 273–304, 1995.

References II

Edwin K. Chong, Christopher M. Kreucher & Alfred O. Hero Iii.
Partially Observable Markov Decision Process Approximations for Adaptive Sensing.
Discrete Event Dynamic Systems, vol. 19, no. 3, pages 377–422, 2009.

J.Andrés Christen & Miguel Nakamura.
Sequential stopping rules for species accumulation.
Journal of Agricultural, Biological, and Environmental Statistics, vol. 8, pages 184–195, 2003.

Yunfei Chu & Juergen Hahn.
Integrating Parameter Selection with Experimental Design Under Uncertainty for Nonlinear Dynamic Systems.
AIChE Journal, vol. 54, no. 9, pages 2310–2320, 2008.

P. Conrad & Y. Marzouk.
Adaptive Smolyak Pseudospectral Approximation.
SIAM Journal on Scientific Computing, 2013.
submitted.

Ian Ford, D. M. Titterington & Kitsos Christos.
Recent Advances in Nonlinear Experimental Design.
Technometrics, vol. 31, no. 1, pages 49–60, 1989.

Roger Ghanem & Pol Spanos.
Stochastic finite elements: A spectral approach.
Springer, 1991.

T. Guest & A. Curtis.
Iteratively Constructive Sequential Design of Experiments and Surveys with Nonlinear Parameter-Data Relationships.
Journal of Geophysical Research, vol. 114, no. B04307, pages 1–14, 2009.

References III

Xun Huan & Youssef M. Marzouk.
Gradient-Based Stochastic Optimization Methods in Bayesian Experimental Design.
2013.

Xun Huan & Youssef M. Marzouk.
Simulation-Based Optimal Bayesian Experimental Design for Nonlinear Systems.
Journal of Computational Physics, vol. 232, no. 1, pages 288–317, 2013.

Shiraj Khan, Sharba Bandyopadhyay, Auroop R. Ganguly, Sunil Saigal, David J. Erickson, Vladimir Protopopescu & George Ostrouchov.
Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data.
Phys. Rev. E, vol. 76, page 026209, 2007.

Anton J. Kleywegt, Alexander Shapiro & Tito Homem-de-Mello.
The Sample Average Approximation Method for Stochastic Discrete Optimization.
SIAM Journal on Optimization, vol. 12, no. 2, pages 479–502, 2002.

O. P. Le Maître & O. M. Knio.
Spectral methods for uncertainty quantification: With applications to computational fluid dynamics.
Springer, 2010.

Q. Long, M. Scavino, R. Tempone & S. Wang.
Fast estimation of expected information gains for Bayesian experimental designs based on Laplace approximations.
Computer Methods in Applied Mechanics and Engineering, 2012.

References IV

Tarek A. El Moselhy & Youssef M. Marzouk.
Bayesian inference with optimal maps.
Journal of Computational Physics, vol. 231, no. 23, pages 7815 – 7850, 2012.

Peter Müller.
Simulation Based Optimal Design.
In Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting, pages 459–474. Oxford University Press, 1998.

Peter Müller, Don A. Berry, Andy P. Grieve, Michael Smith & Michael Krams.
Simulation-based sequential Bayesian design.
Journal of Statistical Planning and Inference, vol. 137, no. 10, pages 3140 – 3150, 2007.

Warren B. Powell.
Approximate dynamic programming: Solving the curses of dimensionality.
Wiley-Interscience, 2007.

Branko Ristic, Sanjeev Arulampalam & Neil Gordon.
Beyond the kalman filter: Particle filters for tracking applications.
Artech House Publishers, 2004.

Herbert Robbins & Sutton Monro.
A Stochastic Approximation Method.
The Annals of Mathematical Statistics, vol. 22, no. 3, pages 400–407, 1951.

Nicholas Roy, Geoffrey J. Gordon & Sebastian Thrun.
Finding approximate POMDP solutions through belief compression.
J. Artif. Intell. Res. (JAIR), vol. 23, pages 1–40, 2005.

References V

Kenneth J. Ryan.
Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model.
Journal of Computational and Graphical Statistics, vol. 12, no. 3, pages 585–603, September 2003.

Paola Sebastiani & Henry P. Wynn.
Bayesian Experimental Design and Shannon Information.
In In 1997 Proceedings of the Section on Bayesian Statistical Science, pages 176–181, 1997.

Alexander Shapiro.
Asymptotic Analysis of Stochastic Programs.
Annals of Operations Research, vol. 30, no. 1, pages 169–186, 1991.

Satinder Singh, Tommi Jaakkola, MichaelL. Littman & Csaba Szepesvári.
Convergence Results for Single-Step On-Policy Reinforcement-Learning Algorithms.
Machine Learning, vol. 38, pages 287–308, 2000.

Antti Solonen, Heikki Haario & Marko Laine.
Simulation-Based Optimal Design Using a Response Variance Criterion.
Journal of Computational and Graphical Statistics, vol. 21, no. 1, pages 234–252, 2012.

Richard S. Sutton.
Learning to predict by the methods of temporal differences.
Machine Learning, vol. 3, pages 9–44, 1988.

R.S. Sutton & A.G. Barto.
Reinforcement learning: An introduction, volume 1.
Cambridge Univ Press, 1998.

References VI

Gabriel Terejanu, Rochan R. Upadhyay & Kenji Miki.
Bayesian Experimental Design for the Active Nitridation of Graphite by Atomic Nitrogen.
Experimental Thermal and Fluid Science, vol. 36, pages 178–193, 2012.

Jojanneke van den Berg, Andrew Curtis & Jeannot Trampert.
Optimal Nonlinear Bayesian Experimental Design: an Application to Amplitude Versus Offset Experiments.
Geophysical Journal International, vol. 155, no. 2, pages 411–421, November 2003.

J. Kyle Wathen & J. Andrés Christen.
Implementation of Backward Induction for Sequentially Adaptive Clinical Trials.
Journal of Computational and Graphical Statistics, vol. 15, no. 2, pages 398–413, 2006.

C.J.C.H. Watkins.
Learning from Delayed Rewards.
PhD thesis, King's College, 1989.

C.J.C.H. Watkins & P. Dayan.
Q-learning.
Machine learning, vol. 8, no. 3, pages 279–292, 1992.

Norbert Wiener.
The Homogeneous Chaos.
American Journal of Mathematics, vol. 60, no. 4, pages 897–936, 1938.