# Mutual Information-based Optimal Design for estimation in 1-D SDEs - Full Observation Case

Alexandre Iolov, Susanne Ditlevsen, André Longtin
<aiolo040 at uottawa dot ca>, alongtin at uottawa dot ca

January 25, 2015

**Abstract**

Here we try to summarize the pieces deemed 'salvagable' from our work on Optimal Design for SDEs based on maximizing the Mutual Information between the process realizations and some prior belief distribution over the parameters.

# Contents

# 1 SDE Parameter Estimation and Optimal Design

The basic goal of 'Optimal Design' is to perturb a dynamical system in an 'optimal' way such as to 'best' estimate its structural parameters.

Consider the generic parametrized and controlled SDE system

$$dX = U(x, \alpha(x,t); \theta)\,\mathrm{d}t + \sqrt{2D}dW \tag{1}$$

and, for illustration case, specialize eq. (1) to the OU system:

$$dX = \underbrace{(\alpha + \beta(\mu - X)}_{U(x,\alpha;\theta)}\,\mathrm{d}t + \underbrace{\sigma}_{\sqrt{2D}} dW \tag{2}$$

Whose parameter set is:

$$\theta = \{\mu, \beta, \sigma\}$$

or a subset thereof.

We want to estimate $\theta$ from observations $\{X_t\}$, given that we have latitude in choosing a perturbation (control) policy $\alpha(\cdot)$. We will assume that we have cts. observations (i.e. very-high-frequency observations).

## 1.1 Some basic concepts/notation

We will call $f$ the probability transition density associated with $X_t$,

$$f(x, t|x_0; \theta, \alpha(\cdot))dx = \mathbb{P}[X_t \in x + dx|x_0, \alpha, \theta] \tag{3}$$

For a fixed parameter set, $\theta$, and control policy, $\alpha(\cdot)$, $f$ is the solution to a Fokker-Planck equation:

$$\partial_t f = \mathcal{L}_{\theta,\alpha(\cdot)}[f] \tag{4}$$

where the differential operator $\mathcal{L}$ is given by:

$$\mathcal{L}_{\theta,\alpha(\cdot)}[f] = D \cdot \partial_x^2[f] - \partial_x[U(x, \alpha(x,t); \theta) \cdot f]$$

We will call $\mathcal{L}$ the forward operator and its adjoint, $\mathcal{L}^*$, the backward operator.

# 2 Optimal Design using Mutual Information

Our main point-of-departure in obtaining $\alpha(\cdot)$ is the concept of Mutual Information (here are two books on information theory - [1, 4] :), see appendix A for definitions).

In particular, two references suggest that one should use the Mutual Information [5, 2] as the criterion for designing a parameter estimation experiment. (Designing the experiment is synonymous to choosing the control policy $\alpha(\cdot)$.)

In order to apply the concept of mutual information, we need to have a prior on the parameter set, $\theta$. We will call this prior:

$$\rho(\theta).$$

Naturally a optimally-designed experiment is one which reveals most the actual value of $\theta$. That is we want to choose $\alpha(\cdot)$ such as to maximize the Information from the random variable $\theta$, given observations of the random variable $X_t$.

First we need to consider the posterior of $\theta$, $p(\theta|X_t)$ given an observation $X_t$ for $t$ fixed

$$p(\theta|X_t;\alpha) = \frac{f(X_t,t|\theta;\alpha)\cdot\rho(\theta)}{\int_\Theta f(X_t,t|\theta;\alpha)\cdot\rho(\theta)\,\mathrm{d}\theta} \tag{5a}$$

That is the transition density is the likelihood of the parameters.

$$L(X_t|\theta) = f(X_t,t|x_0,0;\theta,\alpha(\cdot)) \tag{5b}$$

However, if we have many observations $\{X_{t_n}\}$, the likelihood is, of course, more complicated. Then

$$L(\{X_n\}|\theta) = \prod_n (f(X_{t_{n+1}},t_{n+1}|X_{t_n},t_n;\theta) \tag{6a}$$

and the parameter posterior is then

$$p(\theta|\{X_{t_n}\};\alpha) = \frac{L(\{X_{t_n}\},t|\theta;\alpha)\cdot\rho(\theta)}{\int_\Theta L(\{X_{t_n}\},t|\theta;\alpha)\cdot\rho(\theta)\,\mathrm{d}\theta} \tag{6b}$$

Once we have the concepts of the prior, posterior and likelihood, $\rho, p, L$, then the mutual information, $I$, between the two random variables, $\theta$, $X_t$ is:

$$I(\alpha) = \int_\Theta\int_X \log\left[\frac{p(\theta|x;\alpha)}{\rho(\theta)}\right]\cdot L(x|\theta;\alpha)\cdot\rho(\theta)\,\mathrm{d}\theta\,\mathrm{d}x \tag{7}$$

This is straight from equations 6,8,9 of [5] (That paper is on the OptEstimate Mendeley group). Also see appendix A.

Replacing the posterior in eq. (7), with the bayesian formula from eq. (5a) or 6b, we get:

$$I(\alpha) = \int_\Theta\int_X \log\left[\frac{L(x|\theta;\alpha)}{\int_\Theta L(x|\theta;\alpha)\cdot\rho(\theta)\,\mathrm{d}\theta}\right]\cdot L(x|\theta;\alpha)\cdot\rho(\theta)\,\mathrm{d}\theta\,\mathrm{d}x \tag{8}$$

If $I$ is deemed to encode the information in $\theta$ given $X_t$, it is natural to seek the policy $\alpha(\cdot)$ that maximizes $I$.

However eq. (8) can appear deceptively simple. The integral over $X$, $\int_X \mathrm{d}x$, is really a multi-dimensional integral over all observations. Thus just the evaluation of $I$ is only possible through Monte Carlo methods, unless we are dealing with very few, say $N < 5$ observations. However if you have only one observation, as with the posterior given eq. (5a), then we have a very simple single integral over $x$, which can be easily evaluated approximately using any quadrature rule.

A crucial assumption in the sequel is that we will NOT be updating the prior on the fly, that is we will always have the same $\rho(\theta)$ in eq. (8), even though new observations can obviously be used to update the prior, $\rho$, through Bayes' formula.

# 3 Simplest Example without any reference to Optimal Control Concepts

Reconsider the system in eq. (2)

$$dX = \underbrace{(\alpha + \beta(\mu - X)}_{U(x,\alpha)}\,\mathrm{d}t + \underbrace{\sigma}_{\sqrt{2D}}\,dW$$

We will rederive the standard Maximum Likelihood formulas for $\theta = \{\mu, \beta, \sigma\}$ given the non-zero, but known evolution of $\alpha(\cdot)$.

$X_t$ can be solved explicitly for as:

$$\begin{aligned}
dX =& (\alpha + \beta(\mu - X)dt + \sigma dW \\
(e^{\beta t} dX + e^{\beta t}\beta X_t)dt =& e^{\beta t}(\alpha + \beta\mu)dt + \sigma e^{\beta t} dW \\
X e^{\beta t} - X_0 =& \int e^{\beta t}(\alpha + \beta\mu)dt + \int \sigma e^{\beta t} dW \\
\implies X_t =& e^{-\beta t} X_0 + \frac{(\alpha + \beta\mu)}{\beta} \cdot (1 - e^{-\beta t}) + \sigma \cdot \sqrt{\frac{1 - e^{-2\beta t}}{2\beta}} \cdot \xi
\end{aligned}$$

where $\xi$ is a standard normal RV. Crucially, we have assumed that $\alpha$ is constant, o/w the integral $\int e^{\beta t}\alpha(t)dt$ cannot be solved explicitly.

Then if we have known initial conditions (ICs), i.e. if $X_0$ is constant

$$X_t \sim N\left( e^{-\beta t} x_0 + \frac{(\alpha + \beta\mu)}{\beta} \cdot (1 - e^{-\beta t}), \quad \sigma \cdot \sqrt{\frac{1 - e^{-2\beta t}}{2\beta}} \right)$$

Consider discrete observations $X_n, t_n$ obtained at uniform $t_n$, such that $\alpha$ is constant in between observations. Then the transition probabilities $p_n(X_n|X_{n-1})$ are given by:

$$\begin{aligned}
p_n(X_n|X_{n-1}; \mu, \beta, \sigma; \Delta_n) \propto & \frac{\sqrt{\beta}}{\sigma\sqrt{1 - e^{-2\beta\Delta_n}}} \\
& \cdot \exp\left( \frac{\left(X_n - (\frac{\alpha}{\beta} + \mu) - (X_{n-1} - \frac{\alpha}{\beta} - \mu) \cdot e^{-\beta\Delta_n}\right)^2 \cdot \beta}{\sigma^2(1 - e^{-2\beta\Delta_n})} \right)
\end{aligned}$$

TODO: Change symbol $p \rightarrow f$ to keep consistent with FP notation.

The likelihood is simply the product:

$$L(\{X_n\}|;\mu,\beta,\sigma;\Delta_n;\alpha) = \prod_n p_n(X_n|X_{n-1};\mu,\beta,\sigma;\Delta_n) \tag{9}$$

And the log-likelihood of $X_n, t_n$ is

$$\begin{aligned} l(\beta,\mu,\sigma|) =& \sum \log p_n(X_n|X_{n-1}) \\ =& \frac{N}{2}\log\frac{\beta}{\sigma^2(1-e^{-2\beta\Delta_n})} \\ & - \sum_n \left(X_n - (\frac{\alpha}{\beta}+\mu) - (X_{n-1} - \frac{\alpha}{\beta} - \mu)\cdot e^{-\beta\Delta_n}\right)^2 \cdot \frac{\beta}{\sigma^2(1-e^{-2\beta\Delta_n})} \end{aligned}$$

ML estimators for $\mu,\beta,\sigma$ are obtained via setting $\partial_\theta l$ to zero for each parameter $\theta$. (ignore for now that $\Delta_n$ is not the same throughout). However, it turns out that it is easier to first compute the ML estimate for $\sigma$ and then plug it back into the likelihood, $l$, to simplify things:

$$\begin{aligned} \partial_\sigma l() =& -\frac{N}{\sigma} + 2\sum_n \frac{\left(X_n - e^{-\beta\Delta_n}X_{n-1} - (\frac{\alpha}{\beta}+\mu)\cdot(1-e^{-\beta\Delta_n})\right)^2\beta}{\sigma^3\cdot(1-e^{-2\beta\Delta_n})} \\ \implies \hat{\sigma}^2 =& 2\sum_n \frac{\left(X_n - e^{-\hat\beta\Delta_n}X_{n-1} - (\frac{\alpha}{\hat\beta}+\mu)\cdot(1-e^{-\hat\beta\Delta_n})\right)^2\hat\beta}{N\cdot(1-e^{-2\hat\beta\Delta_n})} \end{aligned} \tag{10}$$

With that the likelihood becomes:

$$l(\beta,\mu|\,X_n) = -N\log\left(\frac{\sum_n\left(X_n - e^{-\beta\Delta_n}X_{n-1} - (\frac{\alpha}{\beta}+\mu)\cdot(1-e^{-\beta\Delta_n})\right)^2}{N}\right) - 1$$

Now maximizing the negative of a log plus a const is the same as minimizing the argument of the log. So we need to *minimize*

$$l(\beta,\mu|\,X_n) \equiv \sum_n\left(X_n - e^{-\beta\Delta_n}X_{n-1} - (\frac{\alpha}{\beta}+\mu)\cdot(1-e^{-\beta\Delta_n})\right)^2$$

Let's differentiate

$$\partial_\mu l() = 2\sum_n\left(X_n - e^{-\beta\Delta_n}X_{n-1} - (\frac{\alpha}{\beta}+\mu)\cdot(1-e^{-\beta\Delta_n})\right)\cdot(1-e^{-\beta\Delta_n}) \tag{11a}$$

$$\begin{aligned} \partial_\beta l() =& 2\sum_n\left(X_n - e^{-\beta\Delta_n}X_{n-1} - (\frac{\alpha}{\beta}+\mu)\cdot(1-e^{-\beta\Delta_n})\right) \\ & \times\left(\Delta_n e^{-\beta\Delta_n}X_{n-1} + \frac{\alpha}{\beta^2}(1-e^{-\beta\Delta_n}) - (\frac{\alpha}{\beta}+\mu)(\Delta_n e^{-\beta\Delta_n})\right) \end{aligned} \tag{11b}$$

The $\mu$ equation is straight-forward to solve

$$\hat{\mu} = \frac{\sum_n \left(X_n - e^{-\beta\Delta_n} X_{n-1} - \frac{\alpha}{\beta}(1 - e^{-\beta\Delta_n})\right)}{N(1 - e^{-\beta\Delta_n})} \tag{12}$$

This means that we only need to solve one equation in one unknown

$$\begin{aligned} 0 = \sum_n & \left(X_n - e^{-\beta\Delta_n} X_{n-1} - (\frac{\alpha}{\beta} + \mu) \cdot (1 - e^{-\beta\Delta_n})\right) \\ & \times \left(\Delta_n e^{-\beta\Delta_n} X_{n-1} + \frac{\alpha}{\beta^2}(1 - e^{-\beta\Delta_n}) - (\frac{\alpha}{\beta} + \mu)(\Delta_n e^{-\beta\Delta_n})\right) \end{aligned} \tag{13}$$

for $\beta$ and then plug into eqs. (10) and (12)

(I have checked that in the case of $\alpha = 0$ this reduces to well-known expressions!)

Note: we've been quite cavalier about the constancy of $\Delta_n$, mostly treating it as a constant, in order to focus on the impact of $\alpha$. Later we can go back and be a little more rigorous, treating $\Delta_n$ as a function of $n$.

## 3.1 Optimal Design

Now we proceed to apply the Mutual Information criterion in order to choose the controls $\alpha(t) = \alpha(t_{n-1})$, assumed piecewise constant over $\Delta_n$, such as to facilitate the estimation of the parameters $\mu, \beta, \sigma$?

Let us call the prior over the parameters $\theta = \{\mu, \beta, \sigma\}$:

$$\rho(\theta)$$

and the posterior over $\theta$ given $X$:

$$p(\theta|X;\alpha) = \frac{L(x|\theta;\alpha) \cdot \rho(\theta)}{\int_\Theta L(x|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta} \tag{14}$$

Where $L(X|\theta;\alpha)$ is the likelihood of $X$ given in eq. (9). $X$ could represent only one observation, $X_n$ or a set of observations $\{X_k\}_n^{n+K}$.

Then we seek to find the $\alpha$ that maximizes the mutual information:

$$I(\alpha) = \int_\Theta \int_X \log\left[\frac{p(\theta|x;\alpha)}{\rho(\theta)}\right] \cdot L(x|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta\,\mathrm{d}x \tag{15}$$

This is straight from equations 6,8,9 of [5] (That paper is on Mendeley)

Replacing the posterior in eq. (15), with the Bayesian formula from eq. (14), we get:

$$I(\alpha) = \int_\Theta \int_X \log\left[\frac{L(x|\theta;\alpha)}{\int_\Theta L(x|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta}\right] \cdot L(x|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta\,\mathrm{d}x \tag{16}$$

Again, the whole magic is to find the $\alpha$ that gives the highest value of $I$.

Now we need to consider how are we going to represent/choose the parameter prior distribution $\rho(\theta) = \rho(\mu, \beta, \sigma)$?

One very simplistic way to proceed is to let the system roll on unperturbed, then to obtain ML estimates for $\theta$ using a small initial segment of

$X_t$ and then to take a Gaussian approximations centred at the so-obtained estimates.

So

$$\rho(\theta) \propto \exp\Big((\theta - \hat{\theta}) \cdot \hat{\Xi}^{-1} \cdot (\theta - \hat{\theta})\Big)) \tag{17}$$

It is not entirely clear how to compute the covariance matrix, $\Xi$, perhaps something related to the Fisher Information of the ML estimates might be viable.

### 3.1.1 Curse of Dimensionality given mulit-observations

As we already mentioned in eq. (15) we have integration wrt. the RV $x$. If, this is only one observation of the process, then $x$ is just $X_n$, but if we take this to be $K$ observations: then 'x' is $\{X_k\}_n^{n+K}$ and we have a $K$ dimensional integral... However it does factor as we can integrate backwards, first wrt $X_{n+K}$ then $x_{n+K-1}$ and so on all the way to $X_n$.

One thing to do is for fixed $\theta$ to sample a few, say $M$, $X$ paths. Then the objective will look like:

$$I(\alpha) = \int_\Theta \sum_{X_i|\theta} \log\left[\frac{L(X_i|\theta;\alpha)}{\int_\Theta L(X_i|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta}\right] \cdot \rho(\theta)\,\mathrm{d}\theta \tag{18}$$

Now suppose that $\theta$ is chosen using some kind of a Gauss-Hermite quadrature scheme in 3-d. Say that requires 125 points (that is only five points in each direction, so quite conservative, but also should give reasonable accuracy). That means that to evaluate $J$ we need to sample $M \times 125$ paths. And then do the summation. But wait, there's more. For each $x, \theta$ pair we also need to do the integral for the normalizing constant... so now we have $M \times 125^2$ process samples in order to naively evaluate $I$. Thus we might first consider, what if we only considered 1-slice ($X_n$) observations.

## 3.2 1-Slice Illustration

Let us make a simple proof-of-concept.

We will take for the true parameters

$$\beta = .05; \mu = -60; \sigma = .1;$$

Now since the values of $\mu, \sigma$ are determined by $\beta$, let us reduce the prior to a one-dimensional Gaussian $\rho(\beta) \propto \exp(-(\beta - \hat{\beta})^2/{\sigma_\beta}^2)$ and then for a given $\beta$ we will obtain $\mu, \sigma$ from the formulas in eqs. (10) and (12).

We will only focus on the next observation, so the likelihood is also a Gaussian distribution (conditional on $\mu, \beta, \sigma$) in 1-d.

Let us write it out using $x_0$ as the current value and $x$ as the future value and $\Delta$ as the time interval until the next observation: then the terms $L, \rho$ in the mutual information

$$I(\alpha) = \int_\Theta \int_X \log\left[\frac{L(x|\theta;\alpha)}{\int_\Theta L(x|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta}\right] \cdot L(x|\theta;\alpha) \cdot \rho(\theta)\,\mathrm{d}\theta\,\mathrm{d}x$$

are given by:

$$L(x|\theta;\alpha) = \frac{\beta}{\sigma\sqrt{2\pi(1-e^{-2\beta\Delta})}} \cdot \exp\left(\frac{\left(x-(\frac{\alpha}{\beta}+\mu)-(x_0-\frac{\alpha}{\beta}-\mu)\cdot e^{-\beta\Delta}\right)^2\cdot\beta}{\sigma^2(1-e^{-2\beta\Delta})}\right)$$

$$\rho(\theta) = \frac{1}{\sigma_\beta\sqrt{2\pi}}\exp(-\frac{(\beta-\hat{\beta})^2}{\sigma_\beta^2})$$

Thus for each $x$ we need to form a Gaussian integral for the normalizing constant. And on top of that we need to make a two dimensional independent gaussian integral for the $x, \beta$.

Let us illustrate this. Start with a sample of duration 50 ms sampled at .1 ms (500 observed pts.). Then if we take a coarser sampling of 1s to create our boot-strap of estimates for $\beta$ (50 observations in each sub-sample) we get:

$$\hat{\hat{\beta}} \quad \sigma_{\hat{\beta}}$$
$$0.0881, 0.0182$$
$$0.0974, 0.0514$$
$$\boxed{0.0902, 0.0562}\text{//used in subsequent simulations}$$
$$0.3060, 0.2968$$
$$0.0696, 0.0386$$
$$0.1947, 0.0790$$
$$0.1702, 0.0624$$
$$0.1706, 0.0481\text{//way off including too small std}$$
$$0.1332, 0.0879$$

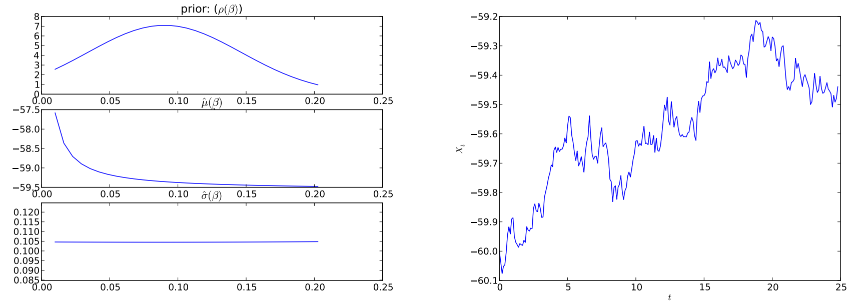
That is actually ok. In 1 case, we couldn't actually estimate $\beta$ (the data implies there is no solution to eq. (13)). In the other 9 cases, 8 times the true value (.05) is within 2 standard deviations of the mean estimate and in 5 cases it is within 1 standard deviation of the mean.

Note that calculating $\sigma_\beta$ as

$$\sigma_\beta^2 = \sum(\hat{\hat{\beta}} - \hat{\beta})^2$$

where $\hat{\hat{\beta}}$ are bootstrap estimates from a reduced data set helps increase the value of $\sigma_\beta^2$. If instead of $\hat{\beta}$, we use the mean of the bootstrap values themselves to calculate $\sigma_\beta$, we will get smaller values for $\sigma_\beta$, which will usually put the true value more than two standard deviations from the short-time estimate.

Let's use the values of $\hat{\beta}, \sigma_{\hat{\beta}} = 0.0902, 0.0562$ and visualize the resulting distributions for $\beta$ and the functions $\mu(\beta), \sigma(\beta, \mu)$, see fig. 1. Essentially what we see is that $\mu$ depends on $\beta$ inversely while $\sigma$ does not really depend on $\beta$ given the data and a $\mu$ estimate. These two facts seem reasonable given eqs. (10) and (12), where $\mu$ is inversely proportional to $\beta$ and $\sigma$ is proportional to the product of $\mu\beta$ once the term $\exp(-\beta\Delta)$ is ignored, and ignoring it is likely justified as long as $\Delta$ is small enough...




(a) prior

(b) underlying $X$ path

Figure 1: An example of a path, a prior over $\beta$ built based on the path and the resulting functional relations between $\beta$, $\mu$ and $\sigma$, which are from eqs. (10) and (12). The prior for $\beta$ and the resulting $\mu, \sigma$ are obtained using the path in (b)

At this point we might start to think a normal prior for an a priori positive random variable, like $\beta$, is a bad idea, especially if we are near zero and the std. dev. of the $\beta$ belief distribution is non-negligible. Perhaps, we should use a log-normal prior or a gamma distribution ...

Now let's calculate the integral of the likelihood wrt. the prior for a fixed $x$, ie. the marginal forward distribution of $x$.

$$p(x) = \int_\Theta L(x|\theta; \alpha) \cdot \rho(\theta) \, \mathrm{d}\theta$$

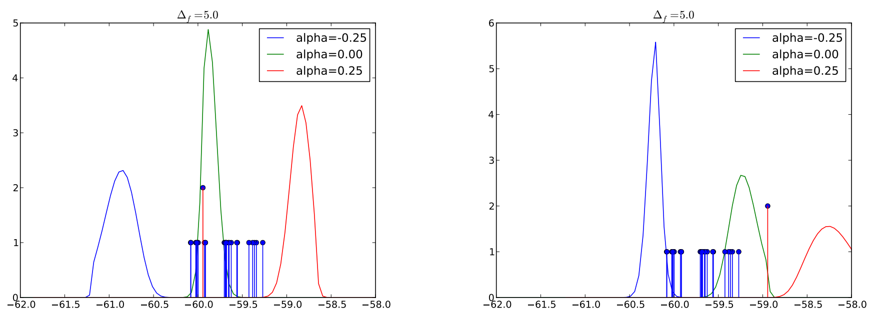
This is also the normalizing constant in bayes rule. We will use a forward horizon of $\Delta_f = 5$ (ms) (Recall the data used to generate the prior distributions had $\Delta = 0.1$ and has length $T = 50$ ms). $p(x)$ is shown in fig. 2a. Essentially, fig. 2a is telling us that the current value of $x$ is lower than the estimate for the long-term mean, all things considered. Our current estimates for $\beta, \mu, \sigma$ are letting us believe that it should move up towards approximately 59.5 provided that we continue with the current value of $\alpha = 0$.

The two opposing proposed values for $\alpha = \pm 0.25$ result in much more spread distributions for $X_{t+\Delta_f}$. In particular, one might argue that $\alpha = -0.25$ is most informative, since the forward distribution has the most spread. (if we intuitively equate spread with information). This is consistent with the notion that the most informative experiments are the ones that lead $X$ furthest from its current equilibrium (given $\alpha = 0$). In this case the negative is better than positive, since the current value of $X_t$ is already negative in relation to the current equilibrium. Basically this is consistent with the following selection mechanism of $\alpha$: If $X_t > \mu$ choose $\alpha_{\max}$ otherwise if $X_t < \mu$ chose $\alpha_{\min}$. This is illustrated in fig. 2b, where we artificially move the value of $X_t$ to the right of the (current)

| $\alpha$ | $I(\alpha)$ |
|---|---|
| -0.25 | 0.688 |
| 0.00 | 0.100 |
| 0.25 | 0.419 |

Table 1: values for the mutual information, $I(\alpha)$, for various values of $\alpha$, starting from the value of $X_t$, (the red stem in fig. 2a)

equilibrium, which makes the forward distribution arising from $\alpha = 0.25$ more informative (higher spread), then the one with $\alpha = -0.25$.




(a) Left of equilibrium

(b) Right of equilibrium

Figure 2: Marginal of $x$ (normalizing constant), the tall red stem is the current value of $X_t$. This is the last observed value based on which we compute transition densities. The lower, blue stems are all the other data previously observed. The blue and red curves are forward densities given different applied forward values for $\alpha$. The green curve is the forward density given the hitherto used value of $\alpha = 0$. It mostly coincides with the so-far observed data(the blue stems). In float (b) we artificially move the starting point to the right. Note that the green curves in a), b) are not the same, since they depend on the value of $X_t$ as well as the observed data ($X_t$ is different in the two panels, but all other data points, the blue stems, are the same)

Let us verify this intuition formally, by calculating $I(\alpha)$, for $\alpha = [-0.25, 0, 0.25]$.

Aside: Unfortunately, naively calculating the double integral (in Python using 'quad' or 'romberg') is not numerically efficient. Calculating $I(\alpha)$ for a single $\alpha$ takes on the order of 10 secs, once you relax the quadrature tolerances without incurring any significant error... Since we are dealing with Gaussian-type integrals, a Gauss-Hermite Integration scheme might be very effective, but we will leave that for now.

Crushing through the integration with brute force we get the result in table 1. We have used the same starting (current) value of $X_t$ to form the forward likelihoods as in fig. 2a and indeed we get the result.

$$I(-0.25) > I(.25) > I(.0)$$

which is consistent with our expectations after calculating the corresponding $p(x)|\alpha$ in fig. 2a. This basically says that it should be most informative to stimulate down ($\alpha < 0$), and it should be least informative to do nothing $\alpha = 0$.

### 3.3 Follow-up Estimation

Let's now see what that means in practice. Pretend that we have done the analysis instantaneously and let the process unroll further from $X_t$ for some time. Then we will check if there is any advantage to using the mutually most informative $\alpha$, i.e. $\alpha = \arg\max I(\alpha)$. First we generate 10 forward trajectories of duration $2\Delta_f = 10$ (ms). They are shown in fig. 3. Now what we would like to see is that the estimates corresponding

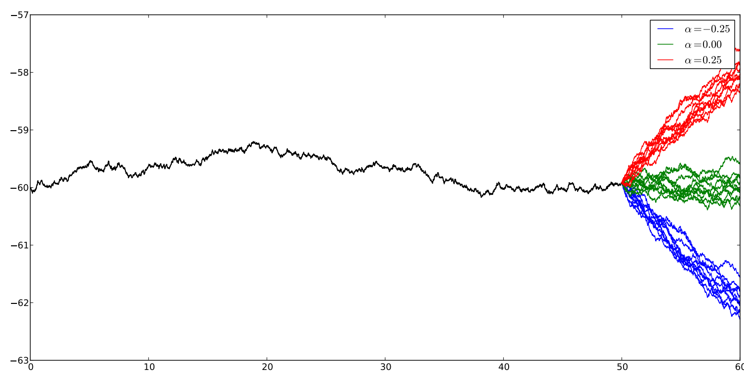



Figure 3: Different Trajectories perturbed by different values of $\alpha$ after the MI calculation

to $\alpha = -.25$ are 'better' than the ones corresponding to $\alpha = .25$ and that they are much better than the ones corresponding to $\alpha = .0$.

Let's see: The resulting estimates for the 10 trajectories are shown in fig. 4. Well, what do you know, visually, it is clear that indeed $\alpha = -.25$ is 'better' than $\alpha = .25$, which in turn is better than $\alpha = .0$.

### 3.4 Bang-Bang?

We expect that it is actually best to apply maximum inhibition or maximum excitation. That is we expect that $I(\alpha_2) > I(\alpha_1)$ for $|\alpha_2| > |\alpha_1|$. We verify this in table 2, where we see the general tendency that bigger is more informative than smaller and negative is more informative than positive. Table 2 and our intuition suggest then that the choice for which

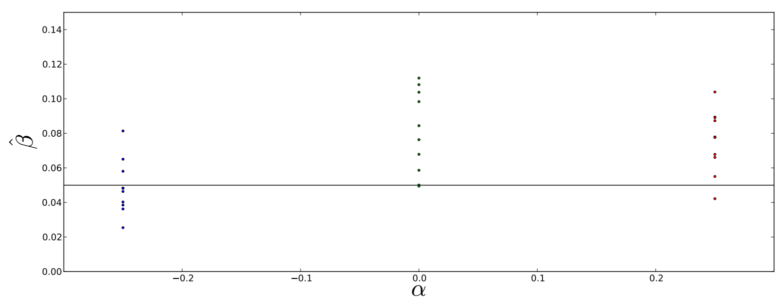



Figure 4: The estimates for $\beta$ given the three perturbed trajectories (one is actually un-perturbed ($\alpha = .0$)). The solid black line indicates the true value of $\beta$

| $\alpha$ | $I(\alpha)$ |
|---|---|
| -2 | 2.31 |
| -1.00 | 1.714 |
| -0.50 | 1.182 |
| 0.50 | 0.885 |
| 1.00 | 1.563 |
| 2.00 | 2.20 |

Table 2: values of MI

is always between the two extremes st. $\alpha_{\text{most informative}} \in \{\alpha_{\min}, \alpha_{\max}\}$.

Now we wonder what happens to the calculated value of $I$ as we move $\Delta_f$, see fig. 5. Now this is interesting. As $\Delta_f$ increases, we have a raise in the mutual information. Intuitively this is obvious, since bigger $\Delta_f$ means more data. However, recall that we are only considering the information contained in the final value (at $t = \Delta_f$). That is also why $I$ levels off eventually, if we were considering the full path and not just the final value, then it should continue increasing monotonically, although perhaps with a decreasing slope. However! It is also clear that while in the current context and for the current example at this time, inhibition is best $I(-) > I(+)$. At some point in the future excitation will be better! That is as the $X$ variable settles into its new, lower, equilibrium, $I(+) > I(-)$.

The question becomes how to formulate this problem as to decide on when to switch. Let me explain why, potentially, this is an interesting

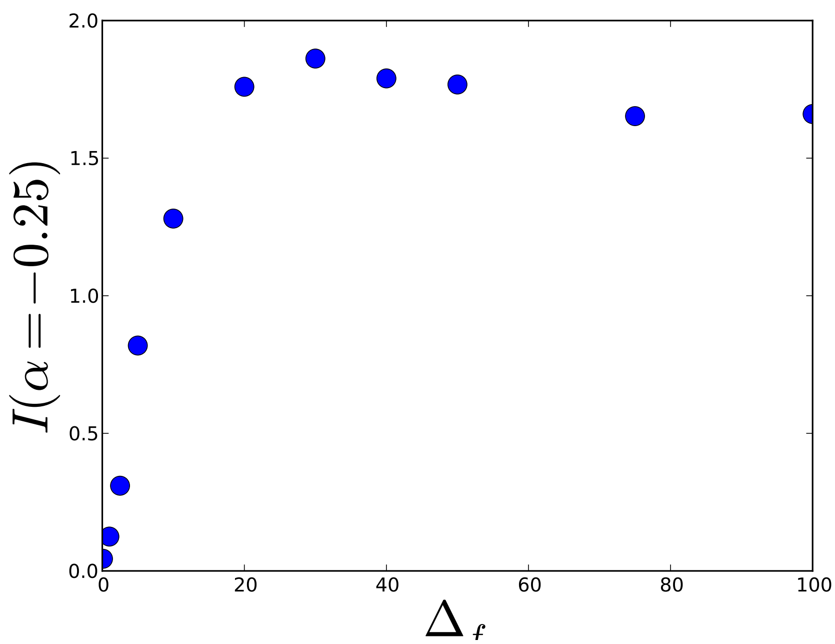



Figure 5: values of MI while varying $\Delta_f$. Note that the right-most value is a little iffy, as the integration routines warn about possible problems with the various integrals

problem, what is a possible solution and why that solution is possibly very difficult to enact:

## 3.5 Optimal Switching for Optimal Design

Let us recap where (we think) we are: There is an observed OU process $X_t$. We estimate it on the fly and have estimates for $\beta, \mu, \sigma$ or more precisely we have a distribution for $\beta$ and a one-to-one relation between $\beta$ and $\mu, \sigma$.

We have a control of $\alpha$ with which we can stimulate $X$. We want to use $\alpha$ to improve the observations of $\beta, \mu, \sigma$. We use the Mutual Information criterion, $I(\alpha)$ to select $\alpha$. From the structure of the OU process, it is conjectured and empirically observed that the bigger in magnitude $\alpha$ the more informative it will be. Thus, in the absence of an energy cost, we are left only to select between the two extreme values of $\alpha$, $[\alpha_{\min}, \alpha_{\max}]$, which we can assume to be just $[\pm\alpha_{\max}]$ for some $\alpha_{\max} > 0$. Thus at any time, $t$, we can calculate $I(\alpha_{\min}), I(\alpha_{\max})$ and choose the $\alpha$ associated with the larger $I$.

However, in a sense this is greedy and thus not necessarily optimal.

Here is a simple way to think about it:

Imagine that $\alpha_{\min}, \alpha_{\max}$ correspond to two equilibria $x_-, x_+$. Our intuition is that there is a mid-point $x_{mid}$ between $x_-, x_+$ st. if $X_t < x_{mid}$, we should apply $\alpha_{\max}$ and conversely. It is now clear that this can easily result in chattering - being below $x_{mid}$ we stimulate which sends us above $x_{mid}$ and then we inhibit, since above $x_{mid}$ it is most informative to inhibit and so on.

When you add the noise, it is hard to see if that is even a better idea than doing nothing.

Of course, things are not so simple as the distribution on the parameters, $\rho$ may make $x_{mid}$ itself move as $\rho$ shifts and shrinks, but let's ignore that for now.

At this point it becomes clear that we need a way to choose the switching time to switch between $\alpha_{\max}, \alpha_{\min}$ using something more sophisticated than just the instantaneous value of $I(\alpha)$. This is related to the problem of how to choose the observation time $\Delta_f$ in the formulation of $I(\alpha)$.

Basically we need to select what we are going to do now in part based on what we can do later. And that is Dynamic Optimization!

The main difference here from standard Dynamic Optimization is that our state is not so much the value of $X_t$ but the value of $\hat{\beta}_t, \hat{\mu}_t, \hat{\sigma}_t$, the estimates at time $t$. Once we realize this, we also realize our main challenge - the updates for $\beta, \mu, \sigma$ using the ML formulas are non-Markovian! That is we look back on all the old data when taking the new data into account to form $\hat{\beta}_t, \hat{\mu}_t, \hat{\sigma}_t$.

We could consider further one of two things

1. come up with a heuristic way of choosing $\Delta_f$ before which to consider switching (if $\Delta_f$ is large enough, we will always switch (I think))
2. Come up with an incremental form for updating $\beta$

In the next section we do NEITHER:) Instead we consider the Optimal Design problem, ie. the selection of $\alpha(\cdot)$ from the point-of-view of Stochastic Optimal Control Theory.

# 4 Finding the optimal design $\alpha(\cdot)$ using Stochastic Optimal Control

In principle now, we have a criterion, the mutual information in eq. (8), using which to select the most informative perturbation $\alpha(\cdot)$. Intutitively, the 'optimal' $\alpha(t)$ depends on the up-to $t$ realization of $X_s; s \leq t$. This is clearly a problem from optimal stochastic control. However, the objective in eq. (8) is not in the form needed to apply dynamic programing (or the maximum principle for that matter). The multi-dimensional integral in $X$ makes things 'too' complicated.

On the other hand, a possible simplification is to take the mutual information criterion in eq. (8) using the single observation likelihood in eq. (5b) and just integrate it over time.

$$J(\alpha) = \int_\theta \int_0^T \int_{X_t} \log\left(\frac{f(x,t|\theta;\alpha(\cdot))}{\int_\theta f(x,t|\theta;\alpha(\cdot))\rho(\theta)\,\mathrm{d}\theta}\right) f(x,t|\theta;\alpha(\cdot))\cdot\rho(\theta)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\theta \tag{19}$$

where the $dx$ integral has the same dimension as the dimension of the SDE not the dimension of the SDE times the number of observations. I should repeat that $J$ is NOT the mutual information between a realization of the process, $\{X_t\}_0^T$ and the parameter set, $\theta$. It is the time-integral of the individual mutual informations between each $X_t$ at time $t$ and the parameter set $\theta$.

The reason we would like to consider this simplification is that now $J$ can be written as

$$J(\alpha) = \mathbb{E}_\theta\left[\mathbb{E}_{X_t|\theta;\alpha(\cdot)}\left[\int_0^T \log\left(\frac{f(X_t,t|\theta;\alpha(\cdot))}{\int_\theta f(X_t,t|\theta;\alpha(\cdot))\rho(\theta)\,\mathrm{d}\theta}\right)\mathrm{d}t\right]\right] \tag{20}$$

which almost has the classical form of a stochastic optimal control problem, except for two non-standard features:

1. We have an extra outer expectation, this is the integration wrt. to the parameter prior.
2. There is a forward-backward coupling in the determinations of the optimal control, meaning we can't just back out, the optimal control with a backwards solution to an HJB equation

Of these two features, we discuss pt. 1 first.

## 4.1 Stochastic Optimal Control with a Prior

We would like to apply Dynamic Programing to the problem of maximizing eq. (20). However section 4.1.1 shows why this is impossible (or at least why the first thing one thinks of does not work).

Instead, in section 4.1.2, we use the Maximum Principle in order to find the optimal $\alpha(\cdot)$.

### 4.1.1 Dynamic Programing with a Prior

WARNING: THIS SECTION ULTIMATLY EXPLAINS WHY YOU CANNOT! USE eq. (26). I.E. WHY DYNAMIC PROGRAMING CANNOT! BE USED TO FIND THE CONTROL, $\alpha(\cdot)$.

In order to apply Dynamic Programing, i.e. in order to set up an HJB PDE, let us write our objective as:

$$J(x_0, 0; \alpha) = \mathbb{E}_\theta \left[ \mathbb{E}_{X_0^T|\theta,\alpha} \left[ \int_0^T r(X_t|\theta)\,\mathrm{d}t \right] \right] \tag{21}$$

where, in our case, the reward $r$ is the mutual information between $X_t$ and $\theta$

$$r(x|\theta) = \log \left( \frac{f(x,t|\theta;\alpha(\cdot))}{\int_\theta f(x,t|\theta;\alpha(\cdot))\rho(\theta)} \right) \tag{22}$$

But we can consider $r(\cdot)$ as a generic function of $X_t$.

Now we try to follow the standard dynamic programing approach to obtain an HJB-type equation.

Call $w$ the value function, i.e. the optimal reward-to-go.

$$w(x_t, t) = \sup_{\alpha(\cdot)} \mathbb{E}_\theta \left[ \mathbb{E}_{X_t^T|\theta,\alpha} \left[ \int_t^T r(X_t|\theta)\,\mathrm{d}t \right] \right] \tag{23}$$

Or in particular, starting from $x_0$ at $t = 0$

$$w(x_0, 0) = \sup_{\alpha(\cdot)} \mathbb{E}_\theta \left[ \mathbb{E}_{X_0^T|\theta,\alpha} \left[ \int_0^T r(X_t|\theta)\,\mathrm{d}t \right] \right] \tag{24}$$

$$\mathbb{E}_{X_0^T|\theta,\alpha}[\cdot]$$

means the expectation over the full trajectory of $X$ from 0 to $T^*$, $X_0$ held fixed, while evolving under the parameter set $\theta$ and the control policy $\alpha$.

$$\mathbb{E}_{\Delta X|\theta,\alpha}[\cdot]$$

Is the expectation over the single point realization of $\Delta X = X(\Delta t) = X_{\Delta t}$, again under a given parameter set and policy, $\theta, \alpha$.

The Markovian nature of $X_t|\theta$ implies that for $x_0$ fixed.

$$\mathbb{E}_{X_0^T|\theta,\alpha}[\cdot] = \mathbb{E}_{\Delta X|\theta} \left[ \mathbb{E}_{X_{\Delta t}^T|\theta,\alpha}[\cdot|\Delta X] \right]$$

With that we can start deriving an equation for $w$, starting from

$$\begin{aligned} w(x_0, 0) &= \sup_{\alpha(\cdot)} \mathbb{E}_\theta \left[ \mathbb{E}_{X_0^T|\theta,\alpha} \left[ \int_0^T r(X_t|\theta)\,\mathrm{d}t \right] \right] \\ &= \sup_{\alpha(\cdot)} \mathbb{E}_\theta \left[ r(x_0)\Delta t \right] + \mathbb{E}_\theta \left[ \mathbb{E}_{X_0^T|\theta,\alpha} \left[ \int_{\Delta t}^T r(X_t|\theta)\,\mathrm{d}t \right] \right] \end{aligned}$$

All we've done so far is split the time integral into an incremental initial part which is approximately equal to $r(x_0)\Delta t$ and the rest $\int_{\Delta t}^T$. Now let's focus on the second term, $\mathbb{E}_\theta \left[ \mathbb{E}_{X_0^T|\theta,\alpha} \left[ \int_{\Delta t}^T r(X_t|\theta)\, \mathrm{d}t \right] \right]$ and condition on $x_0 + \Delta X$:

$$\mathbb{E}_\theta \left[ \mathbb{E}_{X_0^T|\theta,\alpha} \left[ \int_{\Delta t}^T r(X_t|\theta)\, \mathrm{d}t \right] \right] = \mathbb{E}_\theta \left[ \mathbb{E}_{\Delta X|\theta,\alpha} \left[ \mathbb{E}_{X_{\Delta t}^T|\theta,\alpha} \int_{\Delta t}^T r(X_t|\theta)\, \mathrm{d}t | X_{\Delta t} \right] \right]$$

Now here is the main problem! We WOULD LIKE to say that

$$\mathbb{E}_\theta \left[ \mathbb{E}_{\Delta X|\theta,\alpha} \left[ \mathbb{E}_{X_{\Delta t}^T|\theta,\alpha} \int_{\Delta t}^T r(X_t|\theta)\, \mathrm{d}t | X_{\Delta t} \right] \right] = \\ \mathbb{E}_\theta \left[ \mathbb{E}_{\Delta X|\theta,\alpha} \left[ \underbrace{\mathbb{E}_{\boldsymbol{\theta}}}_{\uparrow\texttt{add this?}\uparrow} \left[ \mathbb{E}_{X_{\Delta t}^T|\theta,\alpha} \int_{\Delta t}^T r(X_t|\theta)\, \mathrm{d}t | X_{\Delta t} \right] \quad \right] \right] \quad (25)$$

Because then we could plug in $w(x_0 + \Delta X, \Delta t)$ in:

$$\mathbb{E}_\theta \left[ \mathbb{E}_{X_{\Delta t}^T|\theta,\alpha} \int_{\Delta t}^T r(X_t|\theta)\, \mathrm{d}t | X_{\Delta t} \right] = w(x_0 + \Delta X, \Delta t)$$

And then the rest rolls off easily to get the PDE:

$$\partial_t w(x,t) + \sup_{\alpha(x,t)} \left\{ \mathbb{E}_\theta \left[ \mathcal{L}^*{}_\theta[w] + r(x|\theta) \right] \right\} = 0 \quad (26)$$

Where $\mathcal{L}^*{}_\theta$ is the generator (backward Kolmogorov operator) corresponding to the SDE eq. (1) for fixed parameters, $\theta$,

$$\mathcal{L}^*{}_\theta[\cdot] = U(x,\alpha;\theta)\partial_x[\cdot] + D\partial_x^2[\cdot]$$

HOWEVER! Can we just put the extra $\mathbb{E}_\theta$ in eq. (25)?

It comes down to what exactly is the meaning of the prior on SDE parameters. Is it that:

1. You choose $\theta$ at each time-step (infinitesimally going to 0) let $X$ evolve accordingly for an increment $dt$ and then choose $\theta$ again.
   or
2. You choose $\theta$ at time 0 and then let $X$ evolve accordingly to this once-and-for-all fixed $\theta$

If it is the former, then we can indeed add the extra expectation wrt. $\theta$ in eq. (25) and then use eq. (26) to compute $w$. If it is the latter, then we cannot.

HOWEVER! Assuming pt.1, i.e. that we re-choose $\theta$ at each incremenent, fundamentally violates the basic point of the parameter estimation.

Let me explain.

We suppose that the underlying process $X$ is governed by a single value of $\theta$, we just don't know which. We would like to observe the trajectory of $X$ so as to determine which is the underlying value of $\theta$, but if we re-choose $\theta$ at each time-increment of $X$'s evolution, then there is NO single $\theta$ and indeed the whole estimation problem is moot.

So adding the inner expectation wrt. $\theta$ in eq. (25) is wrong! And solving eq. (26) will not at all help in finding the maximally informative stimulus $\alpha(\cdot)$.

More mundanely, I actually, computed the solution to eq. (26) for the Double-Well Potential and it gives non-sense results (for $\alpha$ AND $w\ldots$)

We must try something else!

#### 4.1.2 Maximum Principle with a Prior

Let us go back to the original objective, eq. (19) or equivalently eq. (20)

$$J(\alpha) = \int_\theta \int_0^T \int_X \log\left(\frac{f(x,t|\theta;\alpha(\cdot))}{\int_\theta f(x,t|\theta;\alpha(\cdot))\rho(\theta)\,\mathrm{d}\theta}\right) f(x,t|\theta;\alpha(\cdot))\cdot\rho(\theta)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}\theta$$

$J$ then is a functional of a family of distributions $f(|\theta)$ parametrized by $\theta$.

We will write this as:

$$J(\alpha) = \mathbb{E}_\theta\left[\int_0^T \int_{\Omega_X} r(f(x,t|\theta;\alpha(\cdot)))\cdot f(x,t|\theta;\alpha(\cdot))\,\mathrm{d}x\,\mathrm{d}t\right]$$

Then optimizing $J$ looks a lot like the a generic problem in optimizing over PDEs with the added complexity of the outer expectation (the one wrt. $\theta$).

We now attempt to set up a Pontryagin-Type equation for the optimal value of $\alpha$: We start by augmenting the objective with the dynamics:

$$J = \mathbb{E}_\theta\left[\int_0^T \int_{\Omega_X} r(f)\cdot f - p\cdot(\partial_t f - \mathcal{L}_{\theta;\alpha}[f])\,\mathrm{d}x\,\mathrm{d}t\right] \tag{27}$$

where $p = p(x,t|\theta;\alpha(\cdot))$ is the adjoint co-state and $\partial_t f - \mathcal{L}_{\theta;\alpha}[f]$ is the Fokker-Planck equation, eq. (4).

What we are going to do is calculate the differential of $J$ wrt. $\alpha(\cdot)$ and then use this in a gradient ascent procedure. First what we would like to do is transfer all the differentials from $f$ to $p$. That is we will integrate $p\cdot(\partial_t f - \mathcal{L}[f])$ by parts so that only $f$ appears in the expression without any of its derivatives. This is a standard exercise, we show it in detail:

$$\begin{aligned}
&-\int_0^T \int_{\Omega_X} p\cdot(\partial_t f - \mathcal{L}[f])\,\mathrm{d}x\,\mathrm{d}s = \\
&= -\int_0^T \int_{\Omega_X} p\cdot(\partial_t f_0 - D\cdot\partial_x^2 f + \partial_x[U\cdot f]\,\mathrm{d}x\,\mathrm{d}t \quad \text{// what is } \mathcal{L} \\
&= \int_0^T \int_{\Omega_X} \partial_t p f\,\mathrm{d}x\,\mathrm{d}t + \int_{\Omega_X} pf\,\mathrm{d}x\Big|_{t=0}^T \quad \text{// the time-derivative pieces} \\
&\quad + \int_0^T \int_{\Omega_X} (D\partial_x^2 p + U\partial_x p)\cdot f\,\mathrm{d}t\,\mathrm{d}x \text{// the space-derivative pieces} \\
&\quad + \int_0^T \Big(pUf - pD\partial_x f + \partial_x pDf\Big|_{x=x_-}^{x_+}\,\mathrm{d}t \text{// the BC terms in 1-d}
\end{aligned}$$

Thus the terminal and boundary conditions of $p$ are chosen given those of $f$ and the objective $J$. Since there are no boundary or terminal terms contributing to our $J$, only the boundary conditions for $f$ impact the choice of boundary conditions for $p$. In particular the following has to be true:

$$
\begin{aligned}
pf\Big|_{t=T} &= 0 && \text{Null TCs}\\
pUf - pD\partial_x f + \partial_x pDf\Big|_{x=x_-,x_+} &= 0 && \text{Null BCs}
\end{aligned}
$$

This usually implies that $p(T) \equiv 0$, since there are usually no a priori restrictions on $f$ at the terminal time. For the boundary terms, if the forward density has reflecting boundaries such that:

$$
Uf - D\partial_x f\Big|_{x=x_-,x_+} = 0
$$

then the BC terms for the adjoint are just the simple Neumann BCs:

$$
\partial_x p\Big|_{x=x_-,x_+} = 0
$$

Once this integration-by-parts is done and the appropriate BCs applied, we can return to the augmented objective, eq. (27) which now looks like:

$$
J = \mathbb{E}_\theta\left[\int_0^T \int_{\Omega_X} \log\left(\frac{f_\theta}{\int_\theta f_\theta \cdot \rho(\theta)\,\mathrm{d}\theta}\right) \cdot f_\theta + (\partial_t p_\theta + \mathcal{L}^*{}_{\theta;\alpha}[p_\theta]) \cdot f_\theta \,\mathrm{d}x\,\mathrm{d}t\right] \tag{28}
$$

with $\mathcal{L}^*$ the adjoint operator to $\mathcal{L}$.

The next step is to take the differential of $J$ wrt. the control $\alpha(x,t)$

In order to make things simpler, we will write the integral wrt. $\theta$ as a sum, i.e.:

$$
\int_\Theta f(\theta)\rho(\theta)\,\mathrm{d}\theta = \mathbb{E}_\theta[f(\theta)] = \sum_\theta w_\theta(f(\theta))
$$

where the weights $w_\theta$ approximate the density $\rho(\theta)$. If one assumes a discrete prior this is just a different way of writing the integral, if the prior is assumed continuous, then this is an approximation. Then $J$ reads like:

$$
J = \mathbb{E}_\theta\left[\int_0^T \int_{\Omega_X} \log(f_\theta) \cdot f_\theta - \log(\sum_\theta w_\theta f_\theta) \cdot f_\theta + (\partial_t p_\theta + \mathcal{L}^*{}_{\theta;\alpha}[p_\theta]) \cdot f_\theta \,\mathrm{d}x\,\mathrm{d}t\right]
$$

and its differential wrt. $\alpha$ for given $x,t$ is:

$$
\begin{aligned}
\delta J|_{x,t} = &\sum_\theta \Big[\frac{\delta f_\theta}{f_\theta} \cdot f_\theta - \frac{w_\theta \delta f_\theta}{\sum_\theta w_\theta f_\theta} \cdot f_\theta + \log\left(\frac{f_\theta}{\sum_\theta w_\theta f_\theta}\right) \cdot \delta f_\theta \\
&+ (\partial_t p_\theta + \mathcal{L}^*{}_{\theta;\alpha} p_\theta) \cdot \delta f_\theta + \delta\alpha \cdot (\partial_x p_\theta \cdot f_\theta)\Big] \\
= &\sum_\theta \left[\left(1 - \frac{w_\theta f_\theta}{\sum_\theta w_\theta f_\theta} + \log\left(\frac{f_\theta}{\sum_\theta w_\theta f_\theta}\right) + (\partial_t p_\theta + \mathcal{L}^*{}_{\theta;\alpha} p_\theta)\right) \cdot \delta f_\theta \right.\\
&\left. + \delta\alpha \cdot (\partial_x p_\theta \cdot f_\theta)\right]
\end{aligned}
$$

Now we need to knock out the $\delta f_\theta$ terms so that we are left with only $\delta\alpha$ terms. Thus we set the coeffiecent of $\delta f_\theta$ to zero, which completes the evolution equation for a given $p_\theta$

$$-\partial_t p_\theta = \mathcal{L}^*{}_{\theta;\alpha}[p_\theta] + 1 - \frac{w_\theta f_\theta}{\sum_\theta w_\theta f_\theta} + \log\left(\frac{f_\theta}{\sum_\theta w_\theta f_\theta}\right) \tag{29}$$

And once $p_\theta, f_\theta$ are solved for, the differential wrt. $\alpha$ comes out to:

$$\frac{\delta J}{\delta\alpha}\Big|_{x,t} = \sum_\theta (\partial_x p_\theta \cdot f_\theta) \tag{30}$$

Equation (30) forms the key ingredient in our gradient search for the most informative perturbation $\alpha^*(x,t) = \arg\max J(\alpha)$.

### 4.1.3 Stationary Maximum Principle with a Prior

Let us now consider a variation on the Maximum Principle approach, when we just maximize the Mutual Information between the stationary distribution and the prior.

This amounts to changing the objective in eq. (19) to

$$J(\alpha) = \int_\theta \int_X \log\left(\frac{f(x|\theta;\alpha(\cdot))}{\int_\theta f(x|\theta;\alpha(\cdot))\rho(\theta)\,\mathrm{d}\theta}\right) f(x|\theta;\alpha(\cdot)) \cdot \rho(\theta)\,\mathrm{d}x\,\mathrm{d}\theta$$

In effect this is a simplified version of section 4.1.2, where we ignore the time evolution of $f$ and focus on its long-term equilibrium.

The calculations are very similar with the exception of $\partial_t f = 0$. We start by augmenting the objective with the dynamics:

$$J = \mathbb{E}_\theta\left[\int_{\Omega_X} r(f)\cdot f + p\cdot \mathcal{L}_{\theta;\alpha}[f]\,\mathrm{d}x\right] \tag{31}$$

where $p = p(x|\theta;\alpha(\cdot))$ is the statinoary adjoint co-state and $\mathcal{L}_{\theta;\alpha}[f]$ is the right hand of the Fokker-Planck equation, eq. (4).

What we are going to do is calculate the differential of $J$ wrt. $\alpha(\cdot)$ and then use this in a gradient ascent procedure. First what we would like to do is transfer all the differentials from $f$ to $p$. That is we will integrate $p\cdot(\mathcal{L}[f])$ by parts so that only $f$ appears in the expression without any of its derivatives. This is done exactly as before:

$$\begin{aligned}
&\int_{\Omega_X} p\cdot(\mathcal{L}[f])\,\mathrm{d}x = \\
&= \int_{\Omega_X} (\mathcal{L}^*[p])\cdot f\,\mathrm{d}x // \text{ the space-derivative pieces} \\
&\quad + \int_0^T \Big(pUf - pD\partial_x f + \partial_x pDf\Big|_{x=x_-}^{x_+} // \text{ the BC terms in 1-d}
\end{aligned}$$

Thus we keep the BCs as in the time-dependent section section 4.1.2, and ditch the TCs:

$$pUf - pD\partial_x f + \partial_x pDf\Big|_{x=x_-,x_+} = 0 \qquad \text{Null BCs}$$

Once this integration-by-parts is done and the appropriate BCs applied, we can return to the augmented objective, eq. (27) which now looks like:

$$J = \mathbb{E}_\theta \left[ \int_{\Omega_X} \log \left( \frac{f_\theta}{\int_\theta f_\theta \cdot \rho(\theta)\,\mathrm{d}\theta} \right) \cdot f_\theta + ({\mathcal{L}^*}_{\theta;\alpha}[p_\theta]) \cdot f_\theta \,\mathrm{d}x \right] \tag{32}$$

Taking the differential of $J$ wrt. the control $\alpha(x,t)$ gives

$$\begin{aligned} \delta J|_x = \sum_\theta \Bigg[ & \left( 1 - \frac{w_\theta f_\theta}{\sum_\theta w_\theta f_\theta} + \log \left( \frac{f_\theta}{\sum_\theta w_\theta f_\theta} \right) + {\mathcal{L}^*}_{\theta;\alpha} p_\theta \right) \cdot \delta f_\theta \\ & + \delta\alpha \cdot (\partial_x p_\theta \cdot f_\theta) \Bigg] \end{aligned}$$

This gives us the differential equation for the adjoint $p$

$$0 = {\mathcal{L}^*}_{\theta;\alpha}[p_\theta] + 1 - \frac{w_\theta f_\theta}{\sum_\theta w_\theta f_\theta} + \log \left( \frac{f_\theta}{\sum_\theta w_\theta f_\theta} \right) \tag{33}$$

And once $p_\theta, f_\theta$ are solved for, the differential wrt. $\alpha$ comes out to:

$$\left.\frac{\delta J}{\delta \alpha}\right|_x = \sum_\theta (\partial_x p_\theta \cdot f_\theta) \tag{34}$$

Equation (34) forms the key ingredient in our gradient search for the most informative *stationary* perturbation $\alpha^*(x) = \arg\max J(\alpha)$. It is exactly the same as eq. (30) and it is imagine that Equation (34) can be derived from eq. (30) without going to all the trouble above. . .

# 5 Illustrative Example - Double Well Potential

## 5.1 Maximum Principle Approach - Time-Dependent Case

We now follow up the theoretical developments from section 4.1.2 with a concrete example.

Our first test problem will be the problem on estimating the double-well potential barrier height as in Sec. 4 of the latest draft of Hooper et al. [3] on arXiv. (From June 7th, 2013). We shall use exactly the same parameter values etc. as in sec. 4 in [3].

We would now like to compute the forward density and the adjoint functions $\{f_\theta, p_\theta\}$ for the double-well problem.

Let's explicitly state the evolution equations for $f_\theta, p_\theta$

$$\begin{gathered} \partial_t f_\theta(x,t;\theta,\alpha(\cdot)) = -\partial_x[U(x;A,\alpha)\cdot f_\theta(x,t)] + D\partial_x^2 f_\theta(x,t) \\ \begin{cases} f_\theta(x,0) & = & \delta(x-x_0) & \text{delta function at some } x_0 \\ U f_\theta - D\partial_x f_\theta \big|_{x=x_-,x_+} & \equiv & 0 & \text{reflecting BCs at some } x_-, x_+ \end{cases} \end{gathered} \tag{35}$$

$$
-\partial_t p_\theta(x,t) = D\partial_x^2 p_\theta(x,t) + U(x;A,\alpha(x,t)) \cdot \partial_x p_\theta(x,t) \\
+1 - \frac{w_\theta f_\theta}{\sum_\theta p_{\theta\theta} f_\theta} + \log\left(\frac{f_\theta}{\sum_\theta w_\theta f_\theta}\right) \tag{36}
$$

$$
\begin{cases} \partial_x p_\theta(x,t)|_{x=x_-,x_+} & = & 0 & \text{BCs} \\ p_\theta(x,T^*) & = & 0 & \text{TCs} \end{cases}
$$

where,

$$
\begin{aligned}
U(x;A,\alpha) \quad &= -\left(4x^3 - 4x - A\tfrac{x}{c}e^{-(x/c)^2/2}\right) + \alpha(x,t) \\
&= -\nabla_x\left(x^4 - 2x^2 + Ae^{-(x/c)^2/2}\right) + \alpha(x,t) \\
&= -\nabla_x\left(\mathcal{V}(x) + \mathcal{A}(x)\right)
\end{aligned}
$$

Having computed $f_\theta, p_\theta$, we compute the gradient $\delta J/\delta\alpha$ as

$$
\frac{\delta J}{\delta \alpha}|_{x,t} = \sum_\theta w_\theta(\partial_x p_\theta \cdot f_\theta)
$$

This just restates eq. (30).

Following [3] (with some deviation in the exact values) we set the parameters as $\sigma = 1. \implies D = 0.5$, (actually it is a little smaller in [3], but this eases the numerics) and $c = 0.3$.

We will use $A = 4$ as the value of $A$ under which the actual process evolves, but that is not important in the computation of the optimal control, $\alpha^*$. For the prior on $A$ we use a uniform over $[2,5]$, which we represent with only $N_\theta = 2$ uniform points - $[2.0, 5.0]$. In general it is not clear how to start the forward density (what its ICs should be). My first attempt - to use a delta mass at $x_0 = 0$, i.e. to assume the process starts at the crest of the barrier, resulted in poor convergence properties for the gradient ascent procedure (by delta mass, we mean a very narrow Gaussian, of course). Instead, using a broad Gaussian distribution centred at the crest gave better results.

The control is constrained to lie in the set $[-10, 10]$, i.e. $\alpha_{\max} = 10$. The space is constrained to $x \in [-2., 2.]$, i.e. $x_-, x_+ = -2., 2.$, using It is further discretized using $\Delta x = .1$, i.e. with 101 uniform points, $[-5, -4.9 \ldots 5.]$.

Let's first see what happens when we run the solver until $T = 5$ (Similar to the $T$ value in [3].)

The experiment proceeds as follows: For our initial guess we take

$$
\alpha_0(x,t) \equiv 0
$$

Then we would like to see that

$$
\operatorname{sgn}\left(\frac{\delta J}{\delta \alpha}\right)\Big|_{x,t} = -\operatorname{sgn}(x)
$$

That is that for negative $x$ we want to drive to the right, $\alpha > 0$ and for positive $x$ we want to drive to the right $\alpha < 0$. Let's see.

The results are visualized in fig. 6. Let's discuss fig. 6. The most important plots are on the right, $\delta J$ which indicate how we are supposed to be changing $\alpha$. It is clear that except for the very beginning $t \approx 0$, and the end $t \approx T$, $\delta J$ is essentially constant in time.

Let's focus then on what happens in the bulk of time in the middle. Indeed we have that

$$\operatorname{sgn}\left(\frac{\delta J}{\delta \alpha}\right)\Big|_{x,t} = -\operatorname{sgn}(x)$$

which implies that we should push the particle to the right (resp. left) depending on whether we are to the left (resp. right) of the barrier at $x = 0$. Everything looks good, except for two points.

1. The magnitude of $\delta J$ is very small. If we were to take steps of size $s = 1.$, it would take us thousands of iteratons to get to what we expect to be the right solution, i.e. bang-bang at $\alpha_{\max} = 10$.
2. The behaviour of $\delta J$ is 'wrong' or 'surprising near the end, $t \approx T$, which appear around $t > 4.75$ and there is also some discrepancies near the beginning, the negative wiggles for $t \approx 0$, which vanish by the time $t > 0.5$

Both issues are minor, but we shall keep them in mind when we go through a full iteration of the gradient descent.

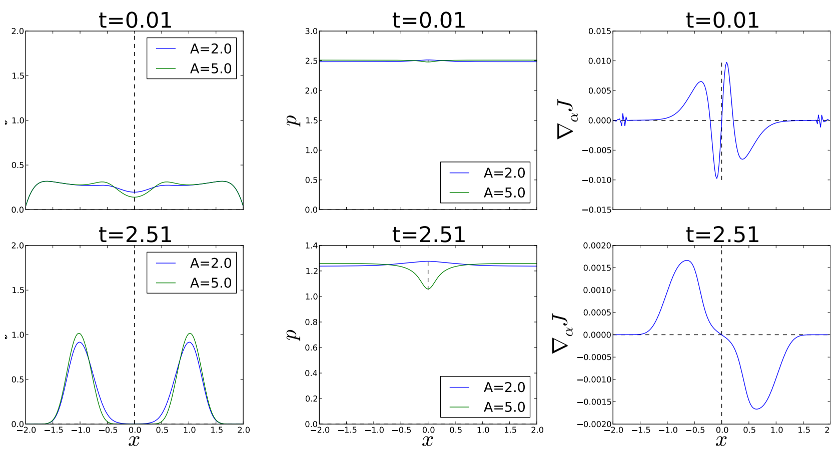



Figure 6: Solution to the test Double Well potential problem using $\alpha \equiv 0$.

### 5.1.1 Going through a full gradient descent iteration

See figs. 7 and 8, basically, we are (almost) able to converge to the bang-bang control and from fig. 8 we are led to believe that the bang-bang control is indeed optimal.

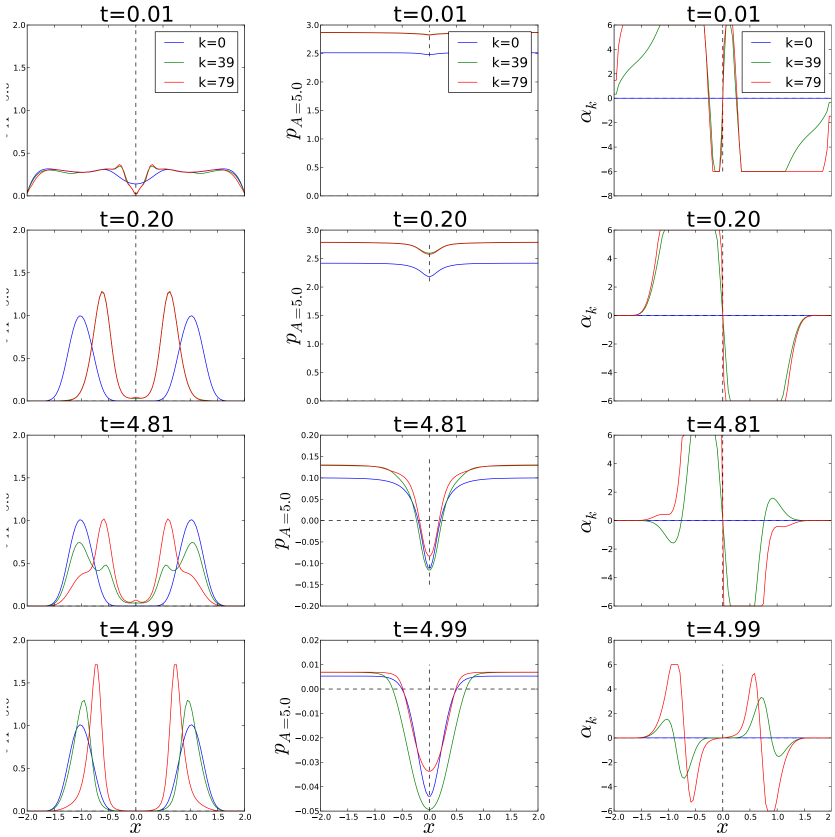



Figure 7: Solution to the test Double Well potential problem starting from $\alpha \equiv 0$ until convergence. Note that except at the very beginning, $t \approx 0$ and very end of the interval $t \approx T$, we have converged to the bang-bang solution.

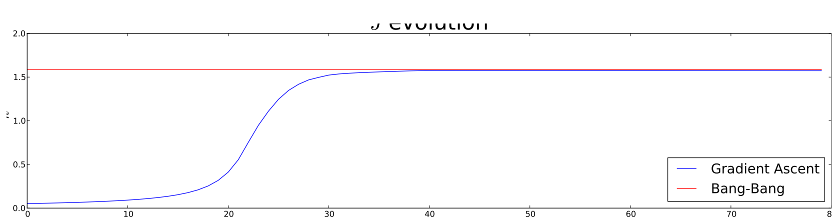



Figure 8: The evolution of $J$ during the Gradient Ascent, we see it goes up and eventually asymptotes with the value of $J$ corresponding to the bang-bang control.

# 6 Second Illustrative Example - the time-constant for the OU model

In section 5.1, we got decent results for the Double-Well potential problem. Let's now apply the same techniques, (from section 4.1.2) on the OU process.

We have seen this a few times, now (eq. (2))

$$dX = \underbrace{(\alpha + \beta(\mu - X)}_{U(x,\alpha;\theta)} \mathrm{d}t + \underbrace{\sigma}_{\sqrt{2D}} dW$$

Since, we've seen the details in section 5.1, we'll just skip to the chase:

(Known) Parameter values are given by:

For the fixed parameters we assume $\sigma = 1. \implies D = 0.5$, $\mu = 0$.

For the prior on $\tau_c = 1/\beta$ we use a uniform over $[.5, 2]$, which we represent with only $N_\theta = 2$ uniform points located at $[.5, 2]$.

For ICs for the forward density, we take a very broad Gaussian centred at 0 (the equilibrium).

The control is constrained to lie in the set $[-1, 1]$, i.e. $\alpha_{\max} = 1$. The space is constrained to $x \in [-2., 2.]$, i.e. $x_-, x_+ = -2., 2.$.

The $\alpha_k$ iterates and the gradient-ascent progress of $J$ are shown in figs. 9 and 10. We see that again, we converge close to the bang-bang control (which in this case does the exact opposite than it did in the Double-Well test case). We also see that the bang-bang control is likely MI-optimal (the red curve in fig. 9).

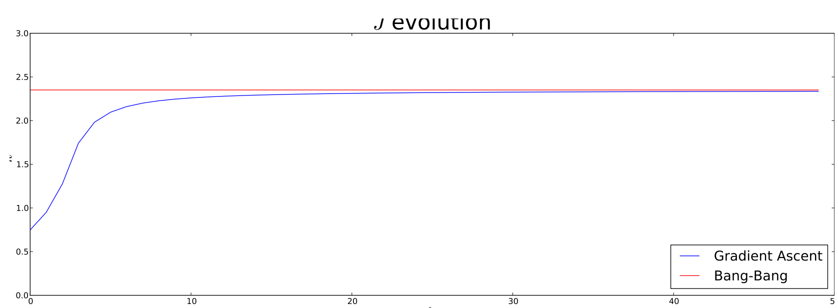



Figure 9: The evolution of $J$ during the Gradient Ascent Ornstein-Uhlenbeck, we see it goes up and eventually asymptotes to the value of $J$ corresponding to the bang-bang control.

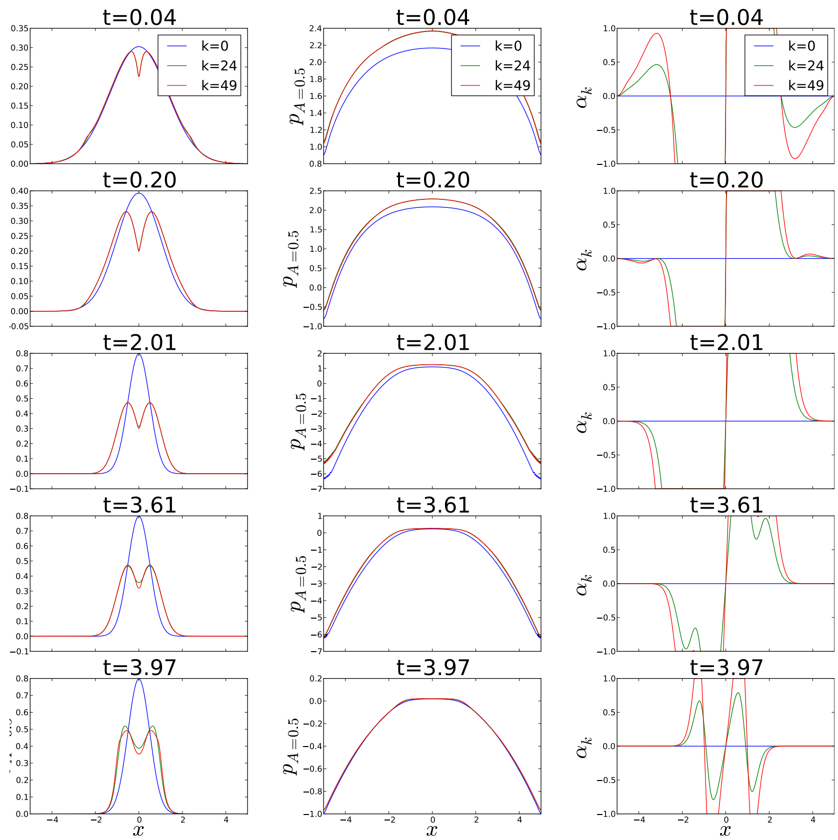



Figure 10: Solution to the test Ornstein-Uhlenbeck problem starting from $\alpha \equiv 0$ until convergence. Note that except at the very beginning, $t \approx 0$ and very end of the interval $t \approx T$, we have converged to the bang-bang solution.

## 6.1 Is bang-bang (in space, i.e. feedback) optimal?

There is some speculation (Susanne!) whether a bang-bang in space control is optimal. Perhaps a bang-bang control in time is better? In particular a bang-bang control in space is a feedback control which pulls away from the equilibrium ($\mu = 0$, known) always in the direction depending on the current position of $X_t$. On the other hand, it might be better to alternatively pull to the left and then to the right. ('Better' in the sense that one gets better estimates). Let's then run a simulation comparison.

On one hand we will take the feedback, MI-optimal bang-bang solution which we obtained in the beginning of this section (section 6). On the other hand we will divide the time interval in $N$ equal length segments and alternate Up/Down on each adjacent segment, for illustration see fig. 11. When we apply the controls (using identical streams of random numbers) we get the trajectories shown in fig. 12. We also show the reference trajectory, using no control at all (labeled 'placebo'). There is obviously quite a big similarity between all trajectories, since we are using a fairly small value of $\alpha_{\max} = 1$.

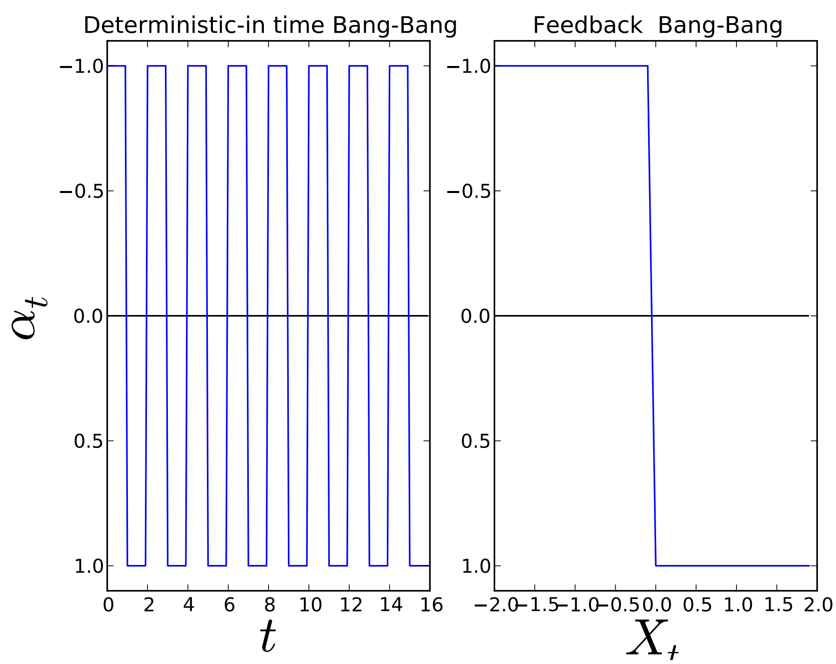



Figure 11: Illustration of the two competing bang-bang controls, space and time

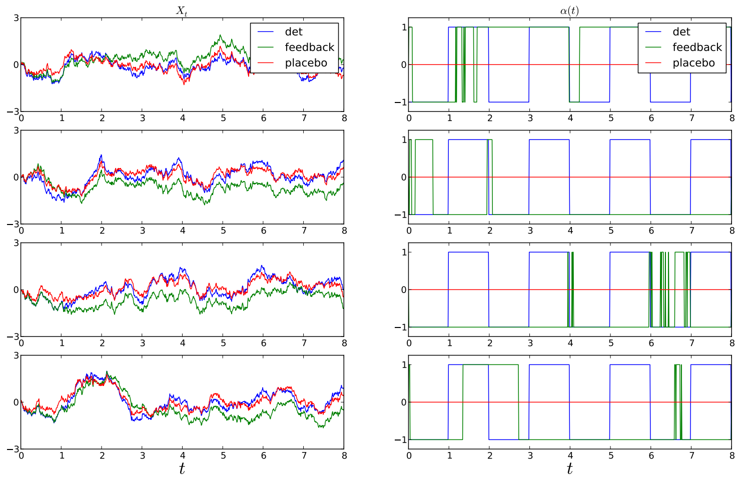



Figure 12: Illustration of the trajectories resulting from the two competing controls, bang-bang in space and time

### 6.1.1 Estimating $\beta$

We now need a MaxLikelihood formula for $\beta$ given the known $\mu = 0$ and $\sigma = 1.$. Consider discrete observations $X_n, t_n$ obtained at uniform $t_n$. Then the transition probabilities $p_n(X_n|X_{n-1})$ are given by:

$$p_n(X_n|X_{n-1};\beta;\alpha_n,\Delta_n) \propto \sqrt{\frac{\beta}{1-e^{-2\beta\Delta_n}}} \cdot \exp\left(-\frac{\left(X_n - (\frac{\alpha_n}{\beta}) - (X_{n-1} - \frac{\alpha_n}{\beta}) \cdot e^{-\beta\Delta_n}\right)^2 \cdot \beta}{(1-e^{-2\beta\Delta_n})}\right)$$

which makes the log-likelihood look like:

$$\begin{aligned} l(\beta|\ldots) =& \sum \log p_n(X_n|X_{n-1}) \\ =& \frac{N}{2}(\log\beta - \log(1-e^{-2\beta\Delta_n}) \\ & - \sum_n \left(\frac{\left(X_n - (\frac{\alpha_n}{\beta}) - (X_{n-1} - \frac{\alpha_n}{\beta}) \cdot e^{-\beta\Delta_n}\right)^2 \cdot \beta}{(1-e^{-2\beta\Delta_n})}\right) + \text{const} \end{aligned}$$

ML estimators for $\beta$ are obtained via setting $\partial_\beta l$ to zero.

| $T_f$: | 8 | 16 | 32 |
|---|---|---|---|
| placebo : | (1.41, 0.66) | (1.19, 0.41) | (1.09, 0.27) |
| det : | (1.37, 0.60) | (1.16, 0.37) | (1.08, 0.24) |
| feedback : | (1.29, 0.39) | (1.14, 0.24) | (1.06, 0.15) |

Table 3: (Mean/st.deviation) of the $\beta$- ML Estimates obtained using the three controls, given different value for $T_f$. Although not substantially, the bang-bang feedback control is clearly superior as it has a lower bias and a lower variance. We have used $N_{traj} = 1000$ to form the statistics for each $T_f$ and control. The 'true' value is $\beta = 1$.

i.e. we must solve for

$$\begin{aligned}\partial_\beta l(\beta|\ldots) =& \frac{N}{2}\left(\frac{1}{\beta} - \frac{2\Delta_n e^{-2\beta\Delta_n}}{(1-e^{-2\beta\Delta_n})}\right) \\ & - \sum_n \text{a horrendous mess} \\ =& 0 \end{aligned} \tag{37}$$

where 'a horrendous mess' is actually (using SAGE - python's CAS):

$$- \frac{2\left(X_{n-1}e^{(-\Delta\beta)} - \frac{\left(e^{(-\Delta\beta)}-1\right)\alpha_n}{\beta} - X_n\right)^2 \Delta\beta e^{(-2\,\Delta\beta)}}{\left(e^{(-2\,\Delta\beta)}-1\right)^2} + \frac{2\left(X_{n-1}e^{(-\Delta\beta)} - \frac{\left(e^{(-\Delta\beta)}-1\right)\alpha_n}{\beta} - X_n\right)\left(\Delta X_{n-1}e^{(-\Delta\beta)} - \frac{\Delta\alpha_n e^{(-\Delta\beta)}}{\beta} - \frac{\left(e^{(-\Delta\beta)}-1\right)\alpha_n}{\beta^2}\right)\beta}{e^{(-2\,\Delta\beta)}-1} - \frac{\left(X_{n-1}e^{(-\Delta\beta)} - \frac{\left(e^{(-\Delta\beta)}-1\right)\alpha_n}{\beta} - X_n\right)^2}{e^{(-2\,\Delta\beta)}-1} \tag{38}$$

!!! Once that is all done, we can compute some estimates for the trajectories from fig. 12 and tabulate those in table 3. Although not substantially, the feedback bang-bang control is clearly superior as it has a lower bias and a lower variance than the estimates obtained when using the deterministic, bang-bang in time, control. I've tried with other values for the switching frequency of the deterministic control (switch every .5, 1, 2 time units) and the results do NOT change. I've also tried with tweaking the value of $\alpha_{max}$, again, same results.

In table 3, we also show the results for the base-case (called 'placebo') where $\alpha(x,t) \equiv 0$.

Finally, in fig. 13, we show something that might be of interest. It is the graph of the score function of the $\beta$ ML estimates, $\partial_\beta l$, eq. (37) as a function of $\beta$ for several different trajectories $X_n$. Recall that an estimate is obtained by finding the solution to $\partial_\beta l(\beta) = 0$. What we see is that the root function corresponding to the feedback control, tends to be much steeper. Intuitively, I associate this with a more robust estimation process, since small vertical perturbations will have small effects on the root $\beta$, while for the flatter curves (the ones corresponding to the placebo and the deterministic control), small vertical perturbations will much more drastically change the $\beta$ estimate.

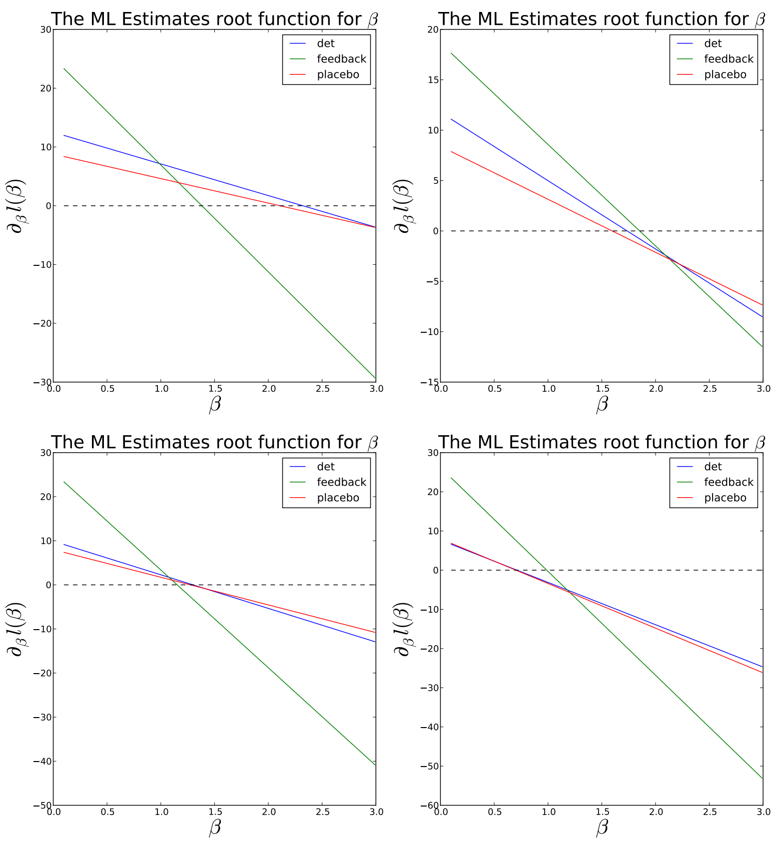



Figure 13: THe root function eq. (37) for a few trajectories

### 6.1.2 Sweep through $\beta_{true}, \sigma$

For completeness we repeat the exercise above, i.e. we recalculate the results from table 3 for several different values of the 'true' value of $\beta$ and of $\sigma$. See table 4

| $(\beta_{true}, \sigma)$: | (5.00,0.25) | (5.00,4.00) | (1.00,0.25) | (1.00,4.00) | (0.20,0.25) | (0.20,4.00) |
|---|---|---|---|---|---|---|
| placebo: | (11.57, 0.79 ) | (0.27, 0.06 ) | (2.51, 0.43 ) | (0.08, 0.05 ) | (0.56, 0.18 ) | (0.06, 0.06 ) |
| det: | (9.09, 0.52 ) | (0.32, 0.11 ) | (2.33, 0.31 ) | (0.10, 0.07 ) | (0.73, 0.23 ) | (0.07, 0.07 ) |
| feedback: | (7.78, 0.42 ) | (0.84, 0.09 ) | (1.26, 0.08 ) | (0.34, 0.07 ) | (0.30, 0.02 ) | (0.18, 0.07 ) |

Table 4: Sweep through the $\beta_{true}, \sigma$ parameters and the effect on the estimates. $Tf = 16$. We've used $N = 100$ to form the statistics. In brackets are displayed the mean estimate (out $N$) and the estimates standard deviation. We have set the initial value to $X_0 = 2.$, if we set it to $X_0 = .0$ the feedback-based estimator will be much more dominant, especially for low noise, presumably because for low noise the other systems spend too much time near the equilibrium if they already start there, and being near equilibrium there is no restoring force and thus $\beta$ is harder to estimate...

## 6.2 Multi-Parameter Case

So far, we have worked in the simplest possible context - with only one unknown parameter. Let's try to ramp it up a bit to TWO unknown parameters. In the OU case this means, $\{\beta, \mu\}$.

WAIT! First let's consider if we only had $\mu$ uncertainty? Suppose we knew $\beta = 1$ and only cared about $\mu$, which could be one of $-1, 1$. Then, after some crunching it turns out (Results NOT shown) that the best thing to do is do nothing, $\alpha \equiv 0$. I don't know why, but that is what the numerics suggest.... Again, by 'best thing' we mean the control, $\alpha(x, t)$, which maximizes $J[\alpha]$ in eq. (19).

So as a first guess, the optimal control for when both $\beta, \mu$ are uncertain should be some combination of 'bang-bang' and 'null'. Bang-bang is best for $\beta$ and 'null' (do nothing/ $\alpha = 0$) is (we think) best for $\mu$. We let the prior be equally weighted between the following (cartesian -product ) possibilities:

$$(\beta, \mu) = \{.5, 2\} \times \{-1, 1\} = (.5, -1), \ldots (2, 1)$$

i.e. in the prior, each of the 4 possibilities has probability 0.25. The calculations after that are exactly as in the earlier section when we only had an uncertainty over $\beta$. We show the distributions and optimal control in fig. 15 and the objective value in fig. 14.

Strictly speaking, what we see is quite novel and interesting! The optimal control is a strange combination of 'bang-bang' outside for $x \notin [-1, 1]$. But null ($\alpha = 0.$), for $xin[-1, 1]$. This is slightly cheating, as we take that that to be the case as initial guess (for $\alpha_k$) and if we started from different initial guess for $\alpha_k$ we will not be able to converge to this, although, that is likely more a fault of the gradient algorithm...

However, we see that the difference is really very small. WHereas in the $\beta$ uncertainty only, fig. 9, we could triple the value of the objective, $J$ by going from the initial guess ($\alpha = 0$) to the optimal control, here the difference between bang-bang, do-nothing, and optimal control is very, very slight!

Uf!

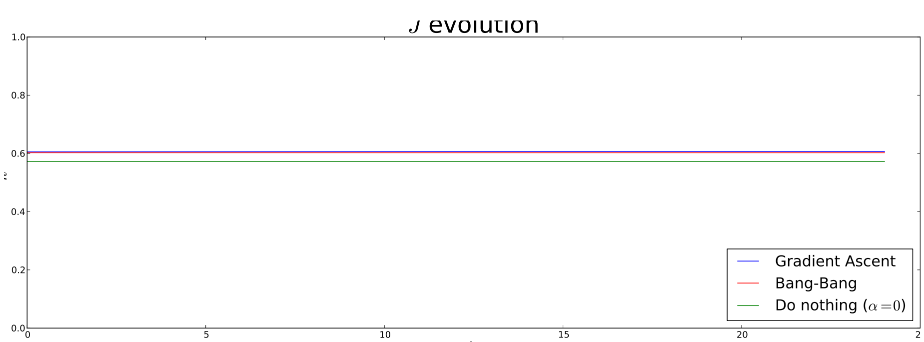



Figure 14: The iterations for $J$ when both $\tau_c, \mu$ are uncertain.

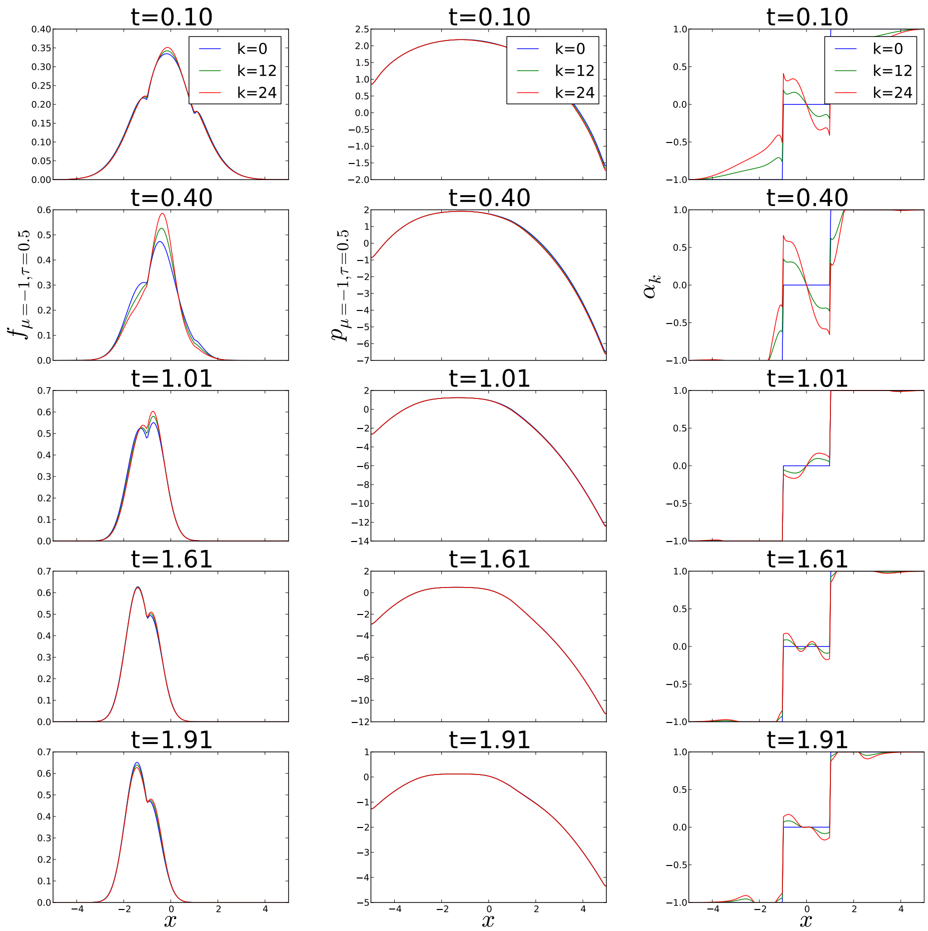



Figure 15: The gradient ascent iterations for $f, p, \alpha$ when both $\tau_c, \mu$ are uncertain. The initial guess for $\alpha_0$ is in 'blue' on the right-most

# A Mutual Info calculation

Here we show why eq. (7) for the Mutual Information agrees with the usual definition of the Mutual Information, which for the random variables, $X, \theta$ is

$$I(X,\theta) = \int_\Theta \int_X p(x,\theta) \cdot \log\left(\frac{p(x,\theta)}{p(x)p(\theta)}\right) \mathrm{d}x \, \mathrm{d}\theta \tag{39}$$

First of all, $p(\theta)$ is just the prior of $\theta$,

$$p(\theta) = \rho(\theta)$$

The joint distribution is

$$p(x,y) = L(x|\theta)\rho(\theta)$$

while the $x$ marginal is

$$p(x) = \int_\Theta L(x|\theta)\rho(\theta) \, \mathrm{d}\theta$$

Plugging the three expressions into the definition in eq. (39) gives:

$$I = \int_\Theta \int_X L(x|\theta)\rho(\theta) \cdot \log\left(\frac{L(x|\theta)\rho(\theta)}{\int_\Theta L(x|\theta)\rho(\theta) \, \mathrm{d}\theta \cdot \rho(\theta)}\right) \mathrm{d}x \, \mathrm{d}\theta. \tag{40}$$

And after canceling $\rho(\theta)$ inside the log, we get eq. (8) which is equivalent to eq. (7).

# References


[1] Thomas Cover and Joy Thomas. Elements of Information Theory. Wiley-Interscience; 2 edition, 2006 edition, 2006.

[2] Jeremy Lewi, Robert Butera, and Liam Paninski. Sequential optimal design of neurophysiology experiments. Neural computation, 21(3):619–87, March 2009.

[3] Kevin K Lin, Giles Hooker, and Bruce Rogers. Control Theory and Experimental Design in Diffusion Processes. pages 1–23.

[4] David J. C. MacKay. Information Theory, Inference and Learning Algorithms. Cambridge University Press, 2003.

[5] Jay I Myung, Daniel R Cavagnaro, and Mark a Pitt. A Tutorial on Adaptive Design Optimization. Journal of mathematical psychology, 57(3-4):53–67, January 2013.