Stochastic Differential Equations
Assignment 2

Raja Shan Zaker Mahmood, s144102

December 21, 2018

Preliminary Remarks

The Matlab code for all the simulations is enclosed in an .m file called $MAIN\_assignment2.m$. When references are given to definitions without source, the course notes are meant.

Question 1. A stochastic Predator-Prey model

We consider two coupled SDE's, governing stochastic processes N_t and P_t. Unless otherwise stated, parameters are $r = 1, K = 1, \beta = 10, \epsilon = 0.1, \mu = 0.05, \sigma_N = \sigma_P = 0.1$.

$$
\begin{aligned}
dN_t &= \left(rN_t(1 - N_t/K) + \beta N_t P_t\right)dt + \sigma_N N_t dB_t^{(1)} \\
dP_t &= \epsilon\beta N_t P_t dt - \mu P_t dt + \sigma_P P_t dB_t^{(2)}
\end{aligned}
$$

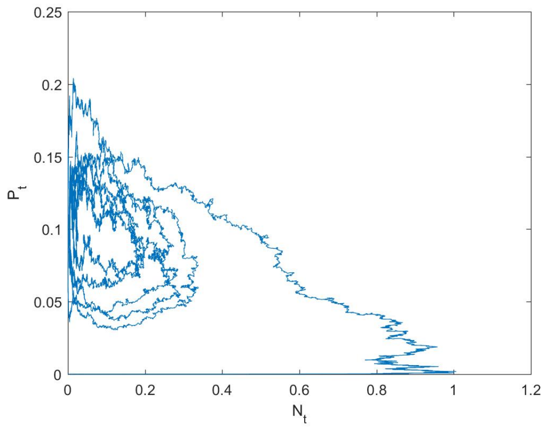
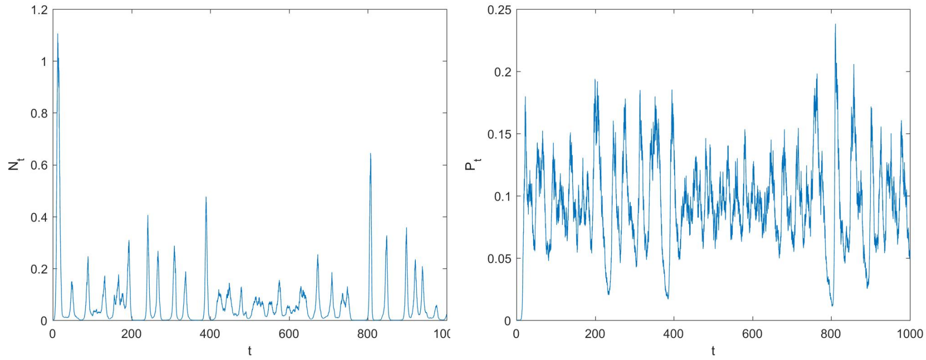


Figure 1: An illustration of the phase-plane dynamics of the stochastic predator-prey model. We see the initial transient and the trajectories forming limit cycles. $T = [0, 200]$, $dt = 0.002$.

(a)

A run of simulation up to $t = 1000$, is shown in figure 2. To get a better view of the transient initial behaviour for these initial conditions, a simulation over a shorter interval is included in figure 3. To estimate the mean, variance and covariance, we have a range of choices. We could perform the simulation for many different seeds, constructing a histogram of the statistics. We have a choice of whether or not to include the initial transient. If we keep it, we can modify the time of simulation, to mitigate the influence of it and approach the statistics of the stationary distribution. Not being entirely scientific, we here choose a longer simulation time for a single realization, instead of many shorter ones. Trial and error deemed that the statistics do not differ significantly once time goes to several thousands.



(a) (b)

Figure 2: Solution of the system, using a step length $dt = 0.01$, on a time interval $T = [0, 1000]$. The explicit Euler-Maruyama method is used.

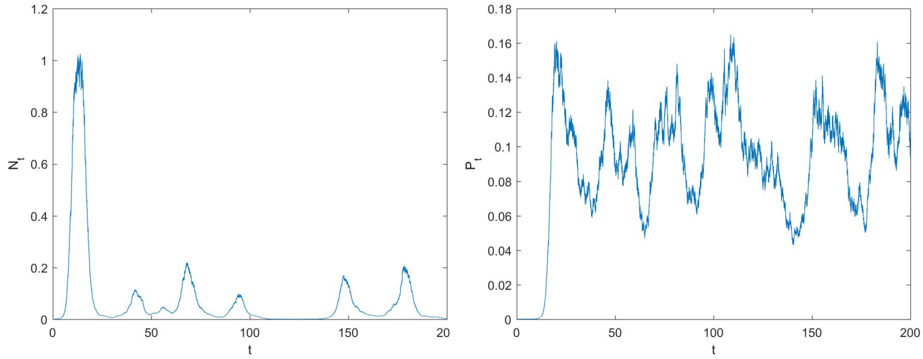
With the tools at our disposal, we cannot say whether the long term behaviour is the same for all realizations in the probability space of the system (cue a diversion into ergodicity theory). To be safely away from the transient, we could take the values of the statistics from $t = 100$ and onward, until a final time of $t = 40000$, which is about as far as numerical stability will take us. The contribution from the transient is entirely trivial by this point, so we just start at $t = 0$.

mean_N	mean_P	var_N	var_P	cov_{NP}
0.057	0.090	0.0082	0.0015	$\begin{pmatrix} 0.0082 & -0.00061 \\ -0.0061 & 0.0015 \end{pmatrix}$

The numbers vary, even with the same seed, again due to numerical instability.

(b)

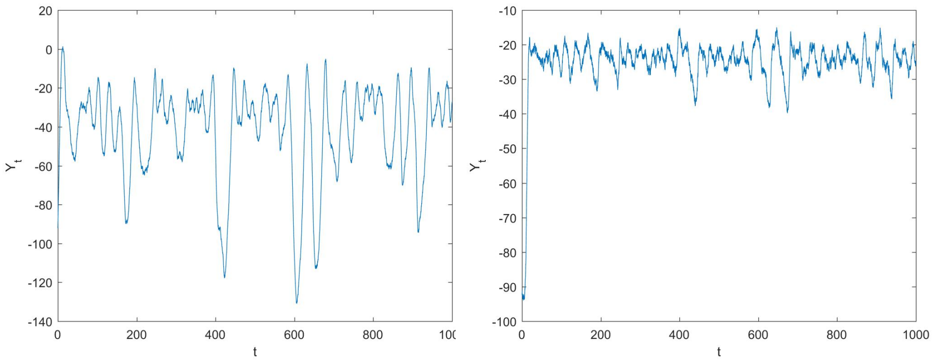
We use the Lamperti transform, but will not go into explicit detail on how the system looks using these coordinates. The implementation in the code provides a reference and we just follow the standard transform introduced in section 7.5.1. There is a slight wrinkle in nomenclature, $X_t = (N_t, P_t)^T$, Y_t are the Lamperti transformed coordinates and $Z_t = (N_t, P_t)$ are the back-transformed coordinates. The solution in the transformed coordinates is displayed in figure 4. The back-transformed system



(a) (b)

Figure 3: A better look at the initial transient, for the same parameters as for the previous figure, but on an interval $T = [0, 200]$ and $dt = 0.002$.

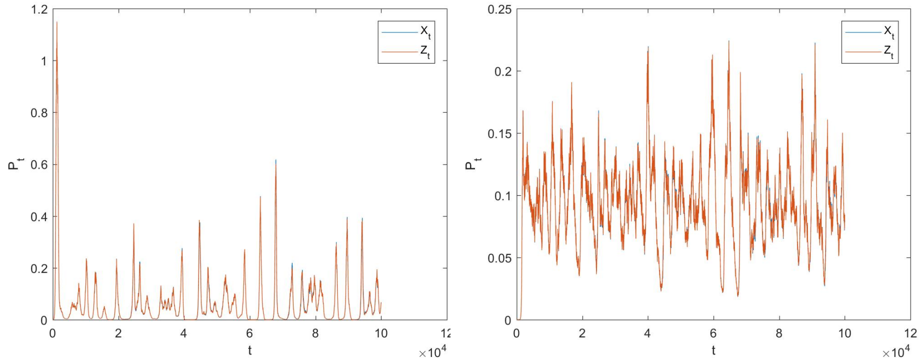
and the original Euler-Maruyama simulation of the system are displayed in figure 5.



(a) (b)

Figure 4: A simulation of the solution to the predator-prey system, transformed by the Lamperti transform. The transformed system is solved using the Stochastic Heun method. Same parameters as in figure 2

As the methods seem to result in very close results, an obvious performance measure would be the time to execute the simulation of a system of a given size. The results are shown in figure 6. It is seen from the experiment that the Heun method is slower but that it seems to have the same asymptotic scale factor (not sure about the terminology here, it seems to be linear in grid size). About 30% slowdown is not a big concession to make for increased stability. It should be mentioned that the stability with respect to noise is greatly improved using the Lamperti transform, as can be tried in the code. It was attempted for $\sigma = 0.3$, step size 0.01, time $t = 10000$ and at 10 I have chosen not to include the plots, as they just show complete numerical instability for the untransformed system and fair reproduction in the transformed and back-transformed one.
The validity of this method follows from the result in exercise 7.79. The Lamperti



(a) (b)

Figure 5: The back-transformed solution from the stochastic Heun method, overlaid over a solution using the Euler-Maruyama method. Both solutions were calculated using the same parameters and the same seed. As can be seen, they are nigh indistinguishable.

transforms of the Ito and the Stratonovich sde's coincide and so, when we transform to these coordinates, it is legitimate to use a numerical method based on the Stratonovich formulation of the stochastic integral. I have elected not to include the derivation, as it is also included in the notes.

(c)

We identify the functions $f(n,p), g(n,p)$, were we have substituted the stochastic processes with coordinates to ease up on notation.

$$f(n,p) = \begin{pmatrix} (rn(1-n/K)-\beta np) \\ \epsilon\beta np - \mu p \end{pmatrix}$$
$$g(n,p) = \begin{pmatrix} \sigma_N n \\ \sigma_P p \end{pmatrix}$$

Now we find the diffusion and the advection. We use the formulation on page 207 for finding the gradient of the diffusion matrix.

$$D = \frac{1}{2}\begin{pmatrix} \sigma_N n \\ \sigma_P p \end{pmatrix}\begin{pmatrix} \sigma_N n \\ \sigma_P p \end{pmatrix}^T = \begin{pmatrix} \sigma_N^2 n^2 & \sigma_N n \sigma_P p \\ \sigma_N n \sigma_P p & \sigma_P^2 p^2 \end{pmatrix}$$

$$\begin{aligned} u = f - \nabla D &= \begin{pmatrix} (rn(1-n/K)-\beta np) \\ \epsilon\beta np - \mu p \end{pmatrix} - \begin{pmatrix} 2\sigma_N^2 + \sigma_N \sigma_P p \\ 2\sigma_p \end{pmatrix} \\ &= \begin{pmatrix} rn(1-n/K)-\beta np - 2\sigma_N^2 + \sigma_N\sigma_P p \\ \epsilon\beta np - \mu p - 2\sigma_p \end{pmatrix} \end{aligned}$$

From this we can write the Forward Kolmogorov equation for the unknown probability density ϕ.

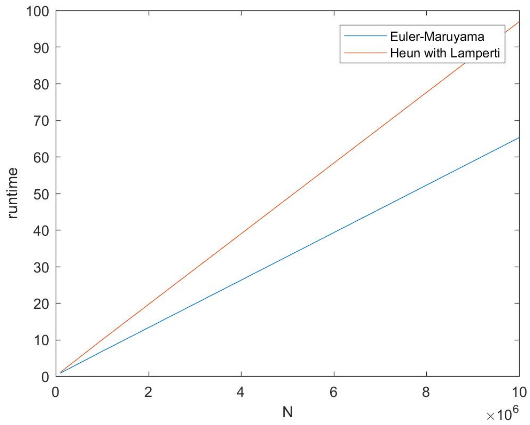


Figure 6: The execution time as a function of grid size.

$$
\begin{aligned}
\dot{\phi} = -\nabla\cdot\left(\begin{pmatrix} rn(1-n/K)-\beta np-2\sigma_N^2+\sigma_N\sigma_P p \\ \epsilon\beta np-\mu p-2\sigma_p \end{pmatrix}\phi - \frac{1}{2}\begin{pmatrix} \sigma_N^2 n^2 & \sigma_N n\sigma_P p \\ \sigma_N n\sigma_P p & \sigma_P^2 p^2 \end{pmatrix}\nabla\phi\right) \\
= \\
-\nabla\cdot\left(\left(\begin{pmatrix} (rn(1-n/K)-\beta np-2\sigma_N^2+\sigma_N\sigma_P p)\,\phi \\ (\epsilon\beta np-\mu p-2\sigma_p)\,\phi \end{pmatrix} - \frac{1}{2}\begin{pmatrix} (\sigma_N^2 n^2\phi_N+\sigma_N n\sigma_P p\phi_P) \\ (\sigma_N n\sigma_P p\phi_N+\sigma_P^2 p^2\phi_p) \end{pmatrix}\right)\right) \\
= \\
-\nabla\cdot\begin{pmatrix} (rn(1-n/K)+\beta np-2\sigma_N^2+\sigma_N\sigma_P p)\,\phi-\frac{1}{2}(\sigma_N^2 n^2\phi_N+\sigma_N n\sigma_P p\phi_P) \\ (\epsilon\beta np-\mu p-2\sigma_p)\,\phi-(\sigma_N n\sigma_P p\phi_N+\sigma_P^2 p^2\phi_p) \end{pmatrix} \\
= \\
-(r\phi+rn\phi_N-\frac{2rn}{K}\phi-\frac{rn^2}{K}\phi_N-\beta p\phi-\beta np\phi_N+2\sigma_N^2\phi_N+\sigma_N\sigma_P p\phi_N-\sigma_N^2 n\phi_N-\frac{1}{2}\sigma_N^2 n^2\phi_{NN} \\
+\sigma_N\sigma_P p\phi_P+\sigma_N\sigma_P np\phi_{PN}+\epsilon\beta n\phi+\epsilon\beta np\phi_P-\mu\phi-\mu p\phi_P-2\sigma_P\phi_P \\
-\sigma_N n\sigma_P\phi_N-\sigma_N n\sigma_P p\phi_{NP}+2\sigma_P^2 p\phi_P+\sigma_P^2 p^2\phi_{PP})
\end{aligned}
$$

Where ϕ_N, ϕ_P denotes ϕ differentiated with respect to N and P respectively. I can spot at least one term (the cross term), that can be eliminated from that humongous expression.

(d)

We write up the sensitivity equation governing the sensitivity matrix S_t. We define the initial condition $x = [N_0, P_0]^T$. For the choice $N_0 = 0, P_0 = 0$ and

$$dS_t(x) = \begin{pmatrix} r & 0 \\ 0 & -\mu \end{pmatrix} S_t(x)dt + \begin{pmatrix} \sigma_N & 0 \\ 0 & 0 \end{pmatrix} S_t(x)dB^{(1)} + \begin{pmatrix} 0 & 0 \\ 0 & \sigma_P \end{pmatrix} S_t(x)dB^{(2)}$$

The equations thus decouple.

$$S_t(x) = \begin{pmatrix} \exp\left((r - \frac{1}{2}\sigma_N)t + \sigma_N B_t\right) & 0 \\ 0 & \exp\left((-\mu - \frac{1}{2}\sigma_P)t + \sigma_P B_t\right) \end{pmatrix}$$

To yield the Lyapunov exponent, we need the find the largest singular value of the sensitivity matrix. As the matrix is diagonal, this reduces to chooses the largest of the diagonal values. This will, in general depend on the parameters chosen for the problem, so we calculate the Lyapunov exponent for both cases. For the parameters in (a), it is the first case.

$$\begin{aligned} \bar{\lambda}_1 = \lim_{t\to\infty} \sup \frac{1}{t} \log\left(\exp\left((r - \frac{1}{2}\sigma_N)t + \sigma_N B_t^{(1)}\right)\right) = \lim_{t\to\infty} \sup \left((r - \frac{1}{2}\sigma_N) + \frac{\sigma_N B_t^{(2)}}{t}\right) \\ = r - \frac{1}{2}\sigma_N \end{aligned}$$

Where we have used that bronian motion scales as $\sqrt{t}$. By a similar argument, we get that if the second singular value is largest, the Lyapunov exponent is of the form. The stability of the solution at this IC is thus determined by the sign of the Lyapunov exponent. Consider it as an analogue to the real part of the eigenvalues in the jacobian, when we linearize nonlinear dynamical systems in deterministic dynamics.

$$\bar{\lambda}_2 = -\mu - \frac{1}{2}\sigma_P$$

Now we assume that parameters are chosen such that the zero solution is unstable. That is, so that the above Lyapunov exponent is positive. We then consider an initial condition $N_0 = n > 0$, $P_0 = 0$. In this case, the stability equation has a more complicated form. Calling $x = (n, 0)^T$.

$$dS_t(x) = \begin{pmatrix} r - \frac{2rn}{K} & -\beta n \\ 0 & \epsilon\beta n - \mu \end{pmatrix} S_t(x)dt + \begin{pmatrix} \sigma_N & 0 \\ 0 & 0 \end{pmatrix} S_t(x)dB_t^{(1)} + \begin{pmatrix} 0 & 0 \\ 0 & \sigma_P \end{pmatrix} S_t(x)dB_t^{(2)}$$

This equation does not decouple, as it did for the trivial case. Instead, we have 4, pairwise coupled sde's. We write the sensitivity matrix in term of components, where we suppress the dependence on x and t.

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

This allows us to write the sensitivity equation for each component.

$$\begin{aligned} dS_{11} &= \left(\left(r - \frac{2rn}{K}\right) S_{11} - \beta n S_{21}\right) dt + S_{11}\sigma_N dB_t^{(1)} \\ dS_{12} &= \left(\left(r - \frac{2rn}{K}\right) S_{12} - \beta n S_{22}\right) dt + S_{12}\sigma_N dB_t^{(1)} \\ dS_{21} &= \left(\epsilon\beta n - \mu\right) S_{21} dt + S_{21}\sigma_P dB_t^{(2)} \\ dS_{22} &= \left(\epsilon\beta n - \mu\right) S_{22} dt + S_{22}\sigma_P dB_t^{(2)} \end{aligned}$$

What follows is probably not legitimate. The last two we can solve analytically.

$$S_{22} = \exp\left((\epsilon\beta n - \mu - \frac{1}{2}\sigma_P^2)t + \sigma_P B_t^{(2)}\right)$$

Now we can zero out the off-diagonal terms because of the IC (?). This gives Lyapunov exponents $\epsilon\beta n - \mu - \frac{1}{2}\sigma_P^2$ and for S_{11} a similar argument we get $r - \frac{2rn}{K} - \frac{1}{2}\sigma_N^2$. We see that the lyapunov exponents are the same as for the zero solution, except for an extra term. So if we assume instability, we have. $r - \frac{1}{2}\sigma_N^2 > 0$ or $-\mu - \frac{1}{2}\sigma_P^2 > 0$. And it did not work.

Question 2. A scalar mean-reverting process with state-dependent noise intensity

We consider the scalar Ito SDE.

$$dX_t = -\lambda X_t dt + \sigma\sqrt{1 + X_t^2}dB_t$$

(a)

To determine the stationary distribution, we set the temporal derivative equal to zero in the forward Kolmogorov equation. As we are dealing with a scalar sde, we can follow the steps from the course notes, to arrive at the following form of the stationary distribution.

$$\phi(x) = \frac{1}{Z}\exp\left(\int_{x_0}^{x}\frac{u(y)}{D(y)}dy\right)$$

Z and x_0 are set arbitrarily, but for simplicity, we set $x_0 = 0$ and later find a Z that normalizes the distribution. We find the advection and the diffusion, u and D respectively.

$$\begin{aligned} D(x) &= \frac{1}{2}\sigma^2\left(1 + x^2\right) \\ u(x) &= -\lambda x - \sigma^2 x \end{aligned}$$

So calculating, we find.

$$\begin{aligned} \phi(x) &= \frac{1}{Z}\exp\left(\int_0^x \frac{-\lambda y - \sigma^2 y}{\frac{1}{2}\sigma^2\left(1 + y^2\right)}dy\right) \\ &= \frac{1}{Z}\left(x^2 + 1\right)^{-\frac{\sigma^2 + \lambda}{\sigma^2}} \end{aligned}$$

The normalization constant Z can be found by straightforward means.

$$Z = \int_{-\infty}^{\infty} \left(x^2+1\right)^{-\frac{\sigma^2+\lambda}{\sigma^2}} dx$$
$$= \frac{\sqrt{\pi}\Gamma\left(\frac{\sigma^2+2\lambda}{2\sigma^2}\right)}{\Gamma\left(\frac{\sigma^2+\lambda}{\sigma^2}\right)}$$

Where Γ is the gamma function. This gives the stationary distribution for the process.

$$\phi(x) = \frac{\left(x^2+1\right)^{-\frac{\sigma^2+\lambda}{\sigma^2}} \Gamma\left(\frac{\sigma^2+\lambda}{\sigma^2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\sigma^2+2\lambda}{2\sigma^2}\right)}$$

We calculate the expectation of x.

$$\mathbb{E}(x) = \int_{-\infty}^{\infty} |x|\phi(x)\, dx = \frac{\Gamma\left(\frac{\sigma^2+\lambda}{\sigma^2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\sigma^2+2\lambda}{2\sigma^2}\right)} \int_0^{\infty} |x| \left(x^2+1\right)^{-\frac{\sigma^2+\lambda}{\sigma^2}} dx$$
$$= \lim_{x\to\infty} \left[\frac{\left(\left(x^2+1\right)^{\frac{-\lambda}{\sigma^2}} - 1\right)\Gamma\left(\frac{\lambda}{\sigma^2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\sigma^2+2\lambda}{2\sigma^2}\right)} \right]$$

We see that the condition for this term to be bounded for any x, is that λ is positive. We find the second mode of x.

$$\mathbb{E}(x^2) = \frac{\Gamma\left(\frac{\sigma^2+\lambda}{\sigma^2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\sigma^2+2\lambda}{2\sigma^2}\right)} \int_{-\infty}^{\infty} x^2 \left(x^2+1\right)^{-\frac{\sigma^2+\lambda}{\sigma^2}} dx = \frac{\sigma^2}{-\sigma^2+2\lambda}$$

This should be bonded for finite parameter values. In general, for the stationary distribution to exist, we expect it to satisfy zero-flux of probability.

(b)

For an initial condition $X_0 = x$, we can readily find the expecion of X_t, by operating with the expectation on both sides of the SDE.

$$d\mathbb{E}^x(X_t) = -\lambda\mathbb{E}^x(X_t)\, dt + \mathbb{E}^x\left(\sigma\sqrt{1+X_t^2}dB_t\right)$$
$$= -\lambda\mathbb{E}^x(X_t)\, dt$$
$$\to$$
$$\mathbb{E}^x(X_t) = xe^{-\lambda t}$$

Where we used the linearity of expectation in the first step and the martingale property of the stochastic integral in the second. To find the expectation of X_t^2, we use Ito's lemma. For the process $Y_t = X_t^2$.

$$dY_t = -2\lambda X_t^2 dt + \sigma^2(1 + X_t^2)dt + 2X_t\sigma\sqrt{1 + X_t^2}dB_t$$

We then take the expectation of both sides in this equation as well, to find the differential equation governing it.

$$d\mathbb{E}\left(X_t^2\right) = -2\lambda\mathbb{E}(X_t^2)dt + \sigma^2\mathbb{E}(X_t^2)dt + \sigma^2 dt$$

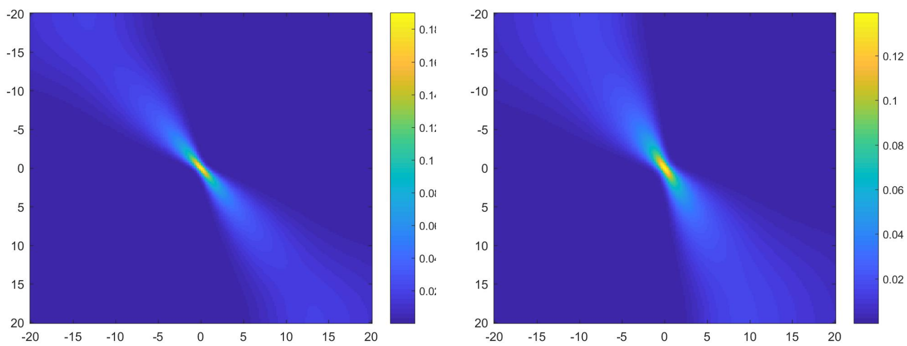
$$\rightarrow$$

$$\mathbb{E}\left(X_t^2\right) = \frac{\sigma^2}{-\sigma^2 + 2\lambda} + xe^{-(-\sigma^2+2\lambda)t}$$

So the condition for the expression to remain bounded for all t, is that $2\lambda > \sigma^2$. The first of the conditions found in this part is the same as for the stationary distribution, but the boundedness of the variance of the stationary distribution did have the same dependence as the expectation of the second mode.

(c)

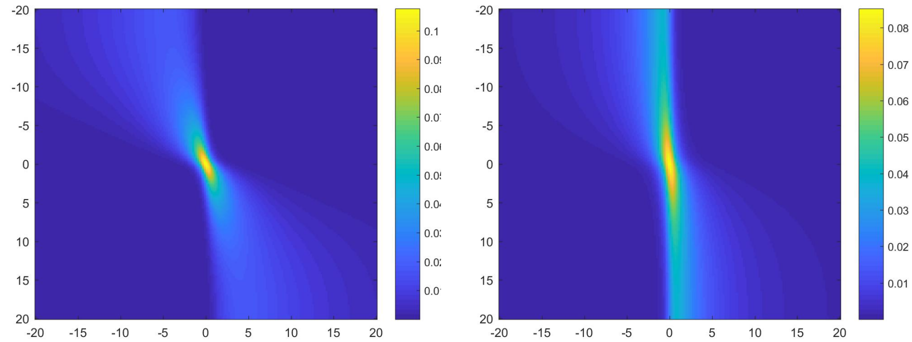
The transition probabilities are plotted for different time steps in figures 7 and 8. The probabilities were calculated with the program $fvade.m$, which we encountered in the exercises. It calculates the generator and the probabilities are calculated with a matrix exponential over a time period.



(a) (b)

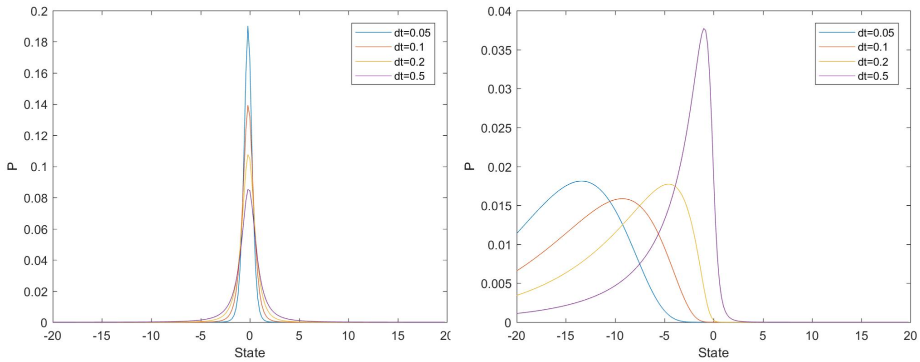
Figure 7: (a) $dt = 0.05$ (b) $dt = 0.1$

Our interpretation follows that from Markov Chain analysis (as that is what this is). The probability in each field the the probability of going from the cartesian x-coordinate to the cartesian y-coordinate. We see that small state values have a very narrow distribution that spread out with time. We see an illustration of this in figure 9. It is also seen that outherlying states will be more probable to go towards the center (state 0) as we look at a longer time interval. This is in line with what we would expect by looking at the Ito sde. It is mean reverting after all and the drift term drives it back towards the center, while thhe stochastic term pushes out.



(a) (b)

Figure 8: (a) $dt = 0.2$ (b) $dt = 0.5$



(a) (b)

Figure 9: (a) The probability distribution of the next step, for a process in state 0. (b) for a process in state -19

(d)

As a first sanity check on the implementation, we plot the likelihood function and see whether it matches with the one specified. It does and the result is illustrated in figure 10, for different values of the process. We expect to see a gaussian centered around the process and an increasing variance.
For the implementation of the filter, we consider the estimated and predicted distributions presented in 10.2. The implemented filters are plotted against the data in figure 11.
For measure of accuracy, we have a host of different options. An argument can be made for removing the four or so extreme outliers, but without knowing any more about the nature of the nature, that cannot be recommended. We can quantify absolute error via the L^1 norm, which basically takes the sum of the absolute distance between data and the distributions from the filter. For the specific choice of $dt = 0.1$, we get a L^1 error of 260.7 for ϕ and 300.71 for ψ. It makes sense that the estimated distribution is more accurate than the predicted one, as it incorporates more infor-

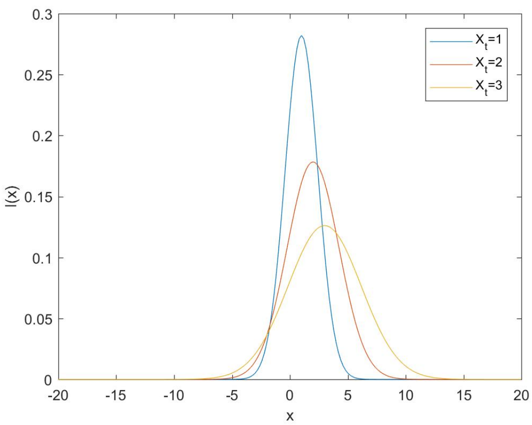


Figure 10: The likelihood functions for different choices of the stochastic process X_t

mation.

(e)

We now introduce a control, modifying the process as follows.

$$dX_t = -\lambda X_t dt + U_t dt + \sigma\sqrt{1+X_t^2}dB_t$$

Given is a performance objective.

$$\mathbb{E}\left[\int_0^T \frac{1}{2}qX_t^2 + \frac{1}{2}U_t^2 dt + \frac{1}{2}\alpha X_T\right]^2$$

We write up the Hamilton-Jacobi-Bellman equation for the problem. For the value function $V(x,t)$, using the coordinate x for the process X_t and the coordinate u for U_t, and suppressing the dependence on the independent variables in V.

$$\frac{\partial V}{\partial t} + \sup_u \left[\frac{\partial V}{\partial x}(-\lambda x + u) + \frac{1}{2}\sigma^2\left(1+x^2\right)\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}qx^2 + \frac{1}{2}u^2\right] = 0$$

We take the supremum of the expression in the brackets, by taking the derivative with respect to u and setting it equal to 0.

$$\frac{\partial V}{\partial x} + u = 0$$

$$\rightarrow$$

$$u = -\frac{\partial V}{\partial x}$$

We insert this into the HJB equation, yielding a non-linear pde.

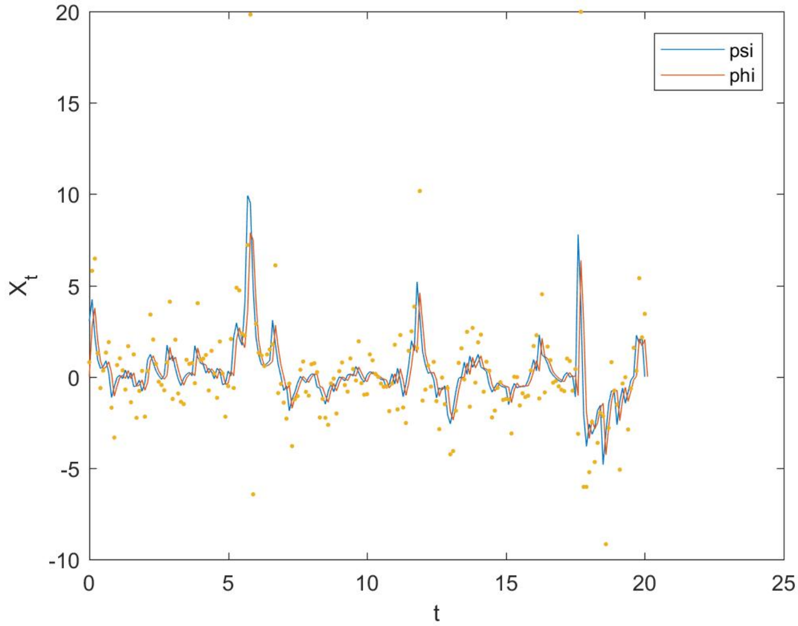


Figure 11: The bayesian filter applied to the data. The implementation can be found in the code. We see that both distributions follow each other closely and seem to provide a good fit for the data.

$$\frac{\partial V}{\partial t} - \lambda x \frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2\left(1+x^2\right)\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}qx^2 - \frac{1}{2}\left(\frac{\partial V}{\partial x}\right)^2 = 0$$

Now using the hint, we use that $V(x,t) = V(0,t) + \frac{1}{2}rx^2$ for some r. Inserting and using that $V(x,t)$ is thus a sum of a time dependent and a space dependent part.

$$\dot{V}(0,t) - \lambda x(rx) + \frac{1}{2}\sigma^2\left(1-x^2\right)r + \frac{1}{2}qx^2 - \frac{1}{2}(rx)^2 = 0$$

To solve this, we gather powers of x as this must hold for any x.

$$\dot{V}(0,t) + \frac{1}{2}\sigma^2 r = 0$$
$$\left(-\lambda r + \frac{1}{2}\sigma^2 + \frac{1}{2}q - \frac{1}{2}r^2\right)x^2$$

From the first equation, we get.

$$V(0,t) = -\frac{1}{2}\sigma^2 rt + C$$

Where C is an arbitrary integration constant. For the second equation to hold for all x, we must have that factor consisting of a polynomial in r is zero. This allows us to find r.

$$\begin{aligned} -\lambda r + \frac{1}{2}\sigma^2 + \frac{1}{2}q - \frac{1}{2}r^2 = 0 \\ \rightarrow \\ r = -\lambda \pm \sqrt{\lambda^2 + \sigma^2 + q} \end{aligned}$$

What remain now is to fix the integration constant C.

$$V(x, T) = -\frac{1}{2}\sigma^2 r T + C + \frac{1}{2}r x^2 = \frac{1}{2}\alpha x^2$$

For this to be the case for any x, we see that we must choose C to be $\sigma^2 r T$ and $\alpha = r$. So that.

$$V(x, t) = -\frac{1}{2}\sigma^2 r (t - T) + \frac{1}{2}r x^2$$

Where $\alpha = r$. And also.

$$U(x) = -r x$$