# Coupled Chemical Reactions in Stochastic Pi-calculus

Andrew Phillips

December 2, 2007

## 1 Introduction

This document describes how a number of spatially homogeneous chemical systems can be modeled in the stochastic pi-calculus and simulated using the Stochastic Pi Machine[1]. Each of the systems presented here was previously defined as a set of reaction equations, which were simulated using the Gillespie algorithm. For further details on the models and the references to the original literature, see: Gillespie, D. (1977) *Exact stochastic simulation of coupled chemical reactions*. J. Chem. Phys. 81 2340-2361.

## 2 Radioactive Decay Reactions

One of the simplest systems that can be simulated is the irreversible isomerization reaction, commonly referred to as radioactive decay. In this system, a species of molecule $X$ decays with rate $c$ to a species $Z$:

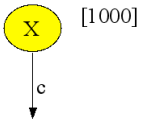
$$X \xrightarrow{c} Z \tag{1}$$

This can be modeled in the stochastic pi-calculus as a process $X()$, which performs a stochastic delay $\tau_c$ with rate $c$ and then executes the process $Z()$:

$$X() \;=\; \tau_c.Z() \tag{2}$$

---

```
directive sample 10.0 1000
directive plot X()
val c = 0.5

let X() = delay@c
run 1000 of X()
```



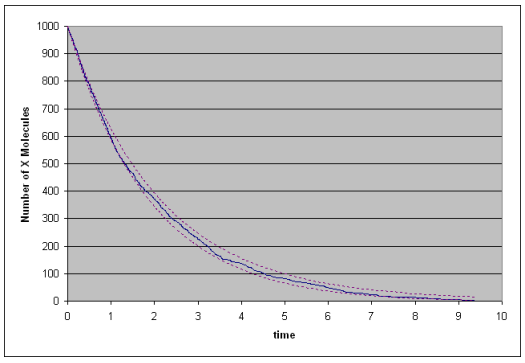




Figure 1: SPiM code and simulation results for reaction (2) with $c = 0.5$ and $X_0 = 1000$. The two-standard deviation envelope (dotted) has been calculated from the stochastic formulation of (1) and superimposed on the results

---

This system was simulated up to time $t = 10$, with $c = 0.5$ and an initial number of $X$ molecules $X_0 = 1000$. The number of $X$ molecules was then plotted versus time. The SPiM code for this simulation is given in Figure 1, together with the corresponding simulation results. A graphical representation of the code is also shown, in

[1] simulator and documentation available from http://research.microsoft.com/~aphillip/spim/

which an oval represents a choice of zero or more actions, and a box represents a parallel composition of one or more processes. Note that since the process $Z()$ does not participate in any reactions it can be omitted from the pi-calculus model.

In this simple example, it is possible to analytically solve the stochastic formulation of (1) and calculate the mean and rms deviation. It turns out that the stochastic mean $X^{(1)}(t) = X_0 e^{-ct}$ and the deviation $\Delta(t) = (X_0 e^{-ct}(1 - e^{-ct}))^{1/2}$. The two-standard deviation envelope, defined as $X^{(1)}(t) \pm \Delta(t)$, was superimposed on the simulation results for Figure 1 in order to compare them with the predictions of the stochastic formulation. One can observe that the stochastic fluctuations of a given simulation generally lie within the boundaries of this envelope.

# 3 The Malek-Mansour and Nicolis Reaction

The following system of reactions was once proposed as a refutation of the basic stochastic hypothesis:

$$\begin{array}{c} X + Y \xrightarrow{c_1} 2Y + X \\ 2Y \xrightarrow{c_2} Z \end{array} \tag{3}$$

In particular, Malek-Mansour and Nicolis showed that the stochastic formulation of this system based on a Master equation has only a single steady-state solution at $Y = 0$, while the deterministic formulation has *two* steady-state solutions, one at $Y = 0$ that is mathematically unstable and another at $Y = c_1 X/c_2$ that is mathematically stable. As a result, they concluded that the stochastic formulation destroys the stable solution of the deterministic formulation, and preserves only the trivial unstable solution. They hypothesised that, according to the stochastic formulation, even if the system starts with a large number of $Y$ molecules it will eventually reach a steady state at $Y = 0$, in apparent contradiction with the deterministic formulation.

In order to check this hypothesis, the system of equations (3) can be modeled as a pi-calculus process and simulated in SPiM. Each $X$ molecule is modeled as a process $X()$, which can perform an input on channel $c_1$ and remain as $X()$. Each $Y$ molecule is modeled as a process $Y()$, which can either perform an output on $c_1$ and evolve to two parallel copies of $Y()$, or perform an input on $c_2$ and evolve to $Z()$, or perform an output on $c_2$.

$$\begin{array}{rl} X() & = ?c_1.X() \\ Y() & = !c_1.(Y() \mid Y()) \quad + \quad ?c_2.Z() \quad + \quad !c_2 \end{array} \tag{4}$$

The input and output on $c_2$ are used to model the fact that two $Y$ molecules can interact with each other to produce a $Z$ molecule. In this model, a given pair of $Y$ molecules can interact in two possible ways: either the first $Y$ molecule can perform an input on $c_2$ and the second molecule can perform an output on $c_2$, or vice-versa. As a result, the rate of channel $c_2$ needs to be adjusted so that $rate(c_2)$ in the pi-calculus model (4) is equal to $c_2/2$ in the reaction model (3).

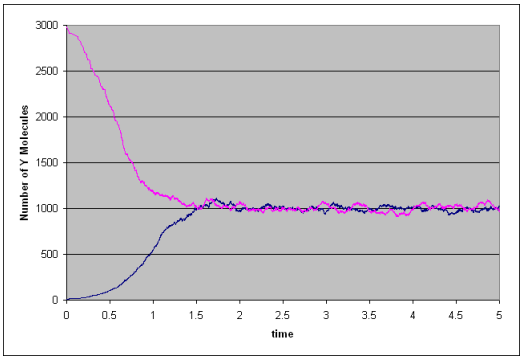
The system was simulated up to time $t = 5$, with $rate(c_1) = 5.0$, $rate(c_2) = 0.0025$ and an initial number of $Y$ molecules $Y_0 = 10$. The number of $Y$ molecules was plotted over time and the simulation was then repeated with $Y_0 = 3000$. The SPiM code for the first simulation is given in Figure 2, together with the results for both simulations. As with the previous system, the process $Z()$ does not participate in any reactions and can be omitted from the pi-calculus model.

The simulation results show that different initial conditions of $Y_0 = 10$ and $Y_0 = 3000$ lead to a situation in which the number of $Y$ molecules fluctuates in an apparently stable manner around the steady state value of $c_1 X/c_2 = 1000$, as predicted by the deterministic formulation of (3). Although in theory the number of $Y$ molecules will eventually reach 0 as $t \longrightarrow \infty$, in practice the system will continue to oscillate indefinitely around the steady state value of $c_1 X/c_2$, with a very low probability of randomly fluctuating from this steady state value to $Y = 0$. In fact, analytical calculations have shown that the variance about the steady-state mean $Y_s^{(1)}$ is given by $\Delta_s^2 = (3/2)Y_s^{(1)}$ which gives a standard deviation of about 39 for a steady state value of 1000. This comparison between analytical calculation and simulation results illustrates how stochastic simulations can help clarify the subtle differences between deterministic and stochastic formulations of chemical systems.

# 4 The Lotka Reactions

The Lotka reactions can be used to model a simple predator-prey ecosystem, in which a prey species $Y_1$ feeds on an inexhaustible food source $X$ to reproduce, a predator species $Y_2$ feeds on $Y_1$ to reproduce and the predator species

```
directive sample 5.0 1000
directive plot Y()
new c1@5.0:chan
new c2@0.0025:chan

let X() = ?c1; X()
let Y() =
  do !c1; (Y() | Y())
  or !c2
  or ?c2
run (X() | 10 of Y())
```



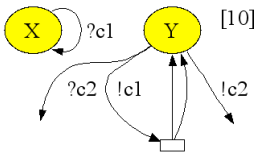




Figure 2: SPiM code and simulation results for reaction (4) with $rate(c_1) = 5$, $rate(c_2) = 0.005$ and initial values $Y_0 = 10$. The simulation results for $Y_0 = 3000$ are also given.

$Y_2$ can die of natural causes:

$$\begin{aligned} X + Y_1 &\xrightarrow{c_1} 2Y_1 + X \\ Y_1 + Y_2 &\xrightarrow{c_2} 2Y_2 \\ Y_2 &\xrightarrow{c_3} Z \end{aligned} \tag{5}$$

This system can be given a deterministic formulation using differential equations, which can be shown to have a steady state of $Y_1 = Y_{1s} = c_3/c_2$ and $Y_2 = Y_{2s} = c_1 X/c_2$. Therefore, if the system has initial populations $Y_1 = Y_{1s}$ and $Y_2 = Y_{2s}$ at time $t = 0$, the deterministic formulation predicts that this situation should persist indefinitely.

In order to check this hypothesis, the system of equations (5) can be modeled as a pi-calculus process and simulated in SPiM. The inexhaustible food source $X$ is modeled as a process $X()$, which can be eaten by performing an input on channel $c_1$ and then remain as $X()$. The prey $Y_1$ is modeled as a process $Y_1()$, which can eat by performing an output on $c_1$ and then reproduce as two $Y_1()$ processes in parallel, or be killed by performing an input on $c_2$ and then disappear. The predator $Y_2$ is modeled as a process $Y_2()$, which can eat by performing an output on $c_2$ and then reproduce as two $Y_2()$ processes, or die of natural causes by performing a stochastic delay $\tau_{c_3}$ and then disappear.

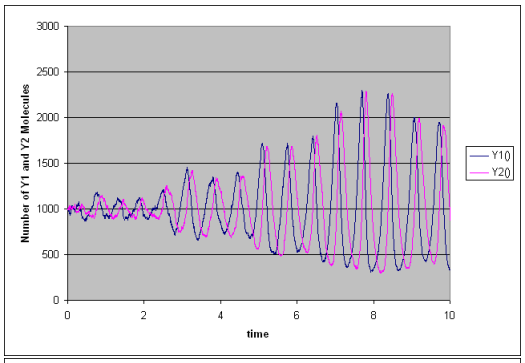
$$\begin{aligned} X() &= ?c_1.X() \\ Y_1() &= !c_1.(Y_1() \mid Y_1()) \quad + \quad ?c_2 \\ Y_2() &= !c_2.(Y_2() \mid Y_2()) \quad + \quad \tau_{c_3} \end{aligned} \tag{6}$$

This system was simulated up to time $t = 30$, with $rate(c_1) = 10.0$, $rate(c_2) = 0.01$, $c_3 = 10.0$, initial populations $Y_1 = Y_2 = 1000$ and an inexhaustible species $X$. The SPiM code for the simulation is given in Figure 3, together with the corresponding simulation results. The results show that, instead of remaining at a constant value of 1000, the number of $Y_1$ and $Y_2$ species oscillates with a fairly stable frequency and phase, but markedly unstable amplitude. Figure 3(a) shows how the predator population lags behind that of the prey, Figure 3(b) shows the stability of the frequency and instability of the amplitude of the oscillations in the prey population and Figure 3(c) shows the counter clockwise orbits traced out in the $Y_1Y_2$ plane.

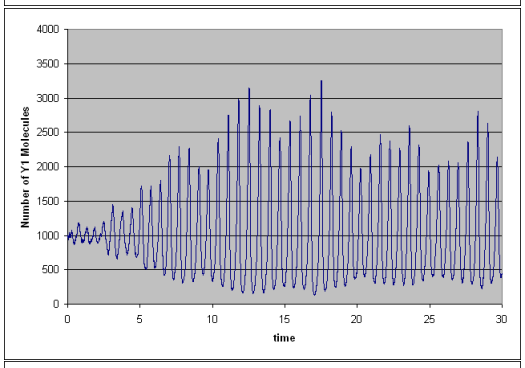
The simulation results can be logically explained by the fact that a rise in the prey population provides additional food for the reproduction of the predators, resulting in a rise in predator population shortly afterwards. This is turn leads to an increase in consumption of prey species, resulting in a decline in the prey population, followed closely by a decline in predator population, and so on. The results can also be explained by analysing the stability of the solutions of the deterministic formulation. Such analysis shows that the orbits in the $Y_1Y_2$ plane are *neutrally stable*, i.e. when perturbed slightly to a point $(Y_{11}, Y_{21})$ off the orbit, the system will begin orbiting on the solution

orbit that passes through the new point $(Y_{11}, Y_{21})$. Therefore, any random fluctuations in $Y_1$ and $Y_2$ will result in the system wandering between neutrally stable orbits. Furthermore, the wide fluctuations in amplitude indicate that it is only a matter of time before the orbits intersect with either the $Y_1$ or $Y_2$ axis. Therefore, as $t \longrightarrow \infty$ either the $Y_1$ prey species becomes extinct and the $Y_2$ predator species dies out soon afterwards, or the $Y_2$ predator species becomes extinct and the $Y_1$ species tends to infinity. This contrasts with the predictions of the deterministic formulation, which suggest that the populations of predator and prey will remain constant over time. These results indicate the importance of taking into account stochastic fluctuations when trying to predict the behaviour of a system.

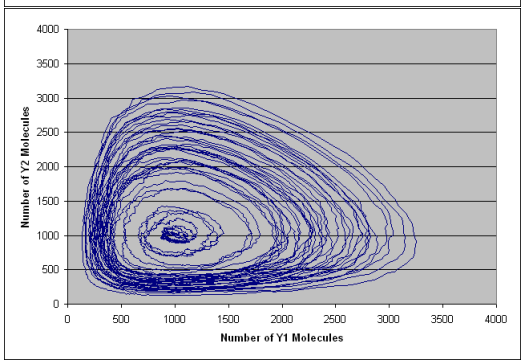
```
directive sample 30.0 1000
directive plot Y1(); Y2()
new c1@10.0:chan
new c2@0.01:chan
val c3 = 10.0

let X() = ?c1; X()
let Y1() =
  do !c1; (Y1() | Y1())
  or ?c2
let Y2() =
  do !c2; (Y2() | Y2())
  or delay@c3
run (X() | 1000 of Y1() | 1000 of Y2())
```



a



b



c

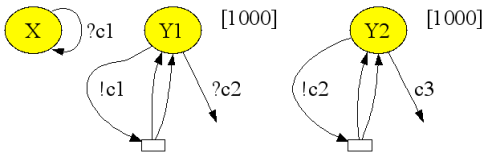



Figure 3: SPiM code and simulation results for the Lotka reactions (6) with $rate(c_1) = 10.0$, $rate(c_2) = 0.01$, $c_3 = 10.0$ and initial values $Y_1 = Y_2 = 1000$. Results for (a) $Y_1, Y_2$ vs. $t$ with $0 < t \leq 10$, (b) $Y_1$ vs. $t$ with $0 < t \leq 30$ and (c) $Y_2$ vs. $Y_1$.

A number of variations of the Lotka reactions can also be simulated. In particular, the food source $X$ can be made finite by changing the definition of reaction $c_1$:

$$X + Y_1 \stackrel{c_1}{\rightarrow} 2Y_1$$

This can be modeled in the pi-calculus by changing the corresponding definition of process $X()$:

$$X() \ = ?c_1$$

The resulting system can be simulated in SPiM by starting with a large quantity of food source $X$, as shown in Figure 4. The simulation results indicate that the depletion of the prey food source $X$ is more detrimental to the predator than to the prey. In this simulation the predators become extinct at $t \simeq 21$, after which the remaining food source $X$ is consumed by the prey for reproduction.

A more realistic system can be defined by adding a reaction to allow the prey to die of natural causes:

$$Y_1 \xrightarrow{c_4} Z$$

This can be modeled in the pi-calculus by changing the definition of the corresponding process $Y_1()$:

$$Y_1() = !c_1.(Y_1() \mid Y_1()) \quad + \quad ?c_2 \quad + \quad \tau_{c_4}$$

The resulting system can be simulated in SPiM by taking $c_4 = c_3$, as shown in Figure 4. As expected, both the predator $Y_1$ and the prey $Y_2$ eventually become extinct. However, it is interesting to note that the predator species becomes extinct significantly before the prey, even though they have the same life expectancy ($1/c_3 = 1/c_4$). More surprisingly, over 40% of the initial food source remains after both the predator and prey have become extinct. These results indicate how useful and sometimes unexpected insight can be gained through stochastic simulation of systems.

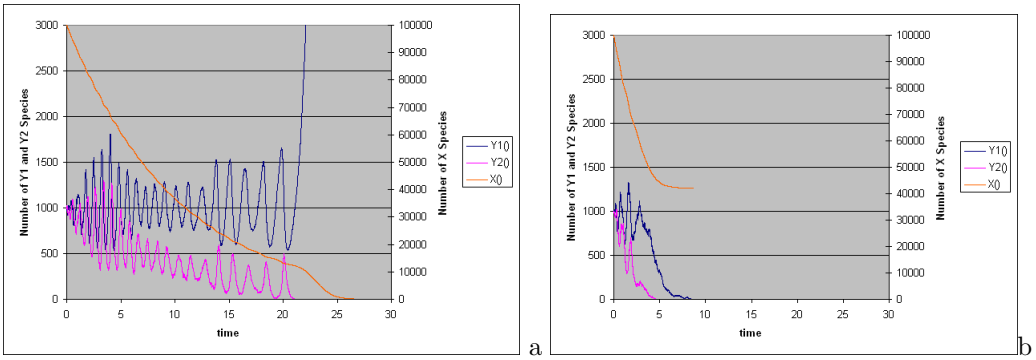
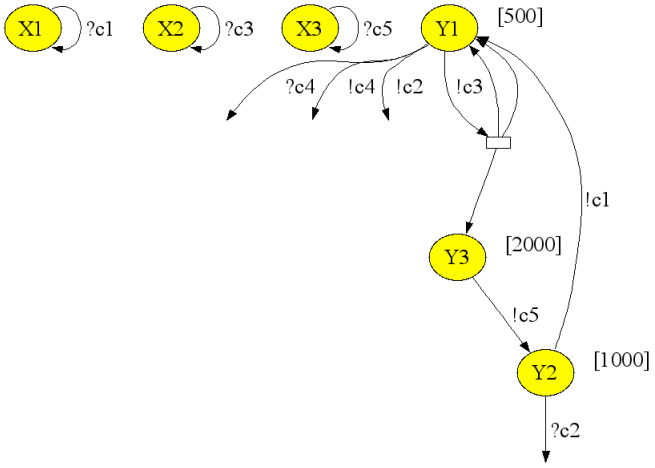
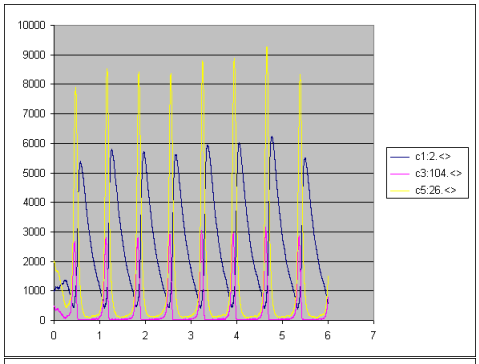



Figure 4: Simulation results for the Lotka reactions (6) but with a limited number of $X$ species. (a) Simulation up to time $t = 30$, with $rate(c_1) = 0.0001$, $rate(c_2) = 0.01$, $c_3 = 10.0$. Initial values $Y_1 = Y_2 = 1000$, $X = 10^5$. (b) Simulation with an additional reaction $Y_1 \xrightarrow{c_4} Z$ that allows the prey to die of natural causes. Simulation up to time $t = 30$, with $rate(c_1) = 0.0002$, $rate(c_2) = 0.01$, $c_3 = 10.0$, $c_4 = 10.0$. Initial values $Y_1 = Y_2 = 1000$, $X = 10^5$.

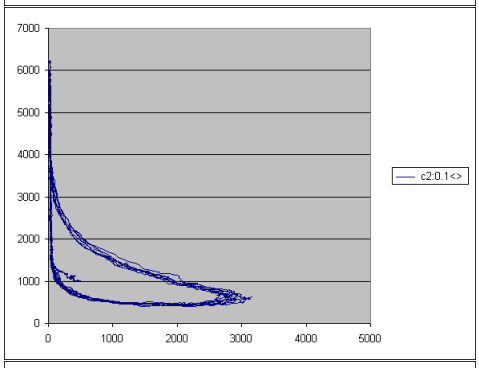
# A The Oregonator Reactions

```
directive sample 6.0 10000
directive plot Y1(); Y2(); Y3()
new c1@2.0:chan
new c2@0.1:chan
new c3@104.0:chan
new c4@0.008:chan (* 0.016 / 2 *)
new c5@26.0:chan

let X1() = ?c1; X1()
let X2() = ?c3; X2()
let X3() = ?c5; X3()
let Y1() =
  do !c2
  or !c3; (Y3() | Y1() | Y1())
  or !c4
  or ?c4
and Y2() =
  do !c1; Y1()
  or ?c2
and Y3() = !c5; Y2()
run (X1() | X2() | X3() | 500 of Y1())
run ( 1000 of Y2() | 2000 of Y3())
```

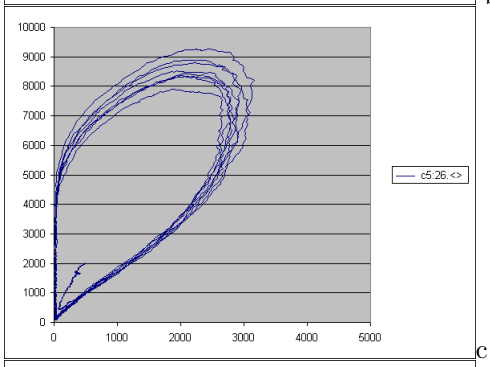


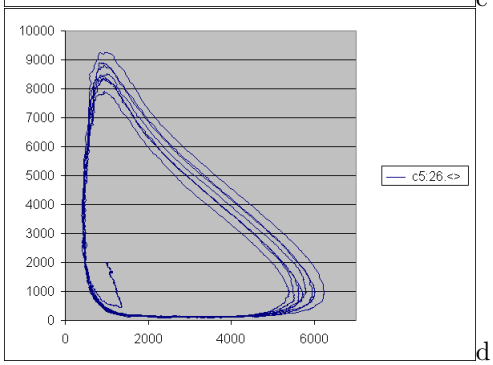

a




b




c

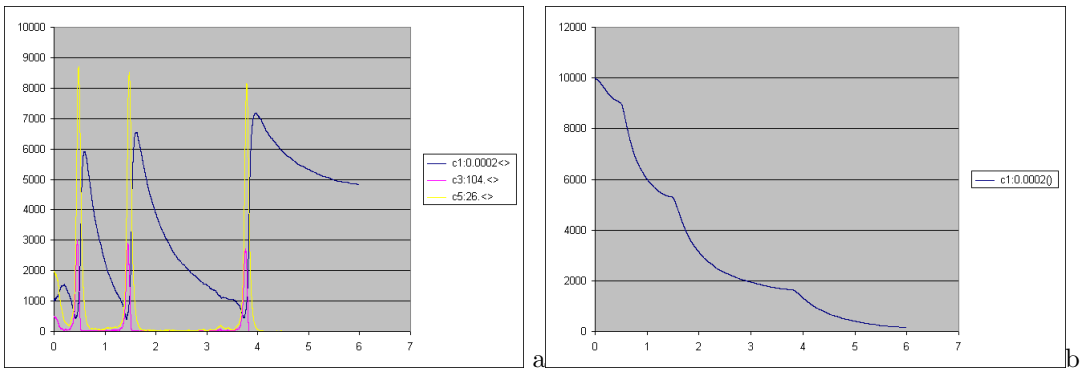



d

Figure 5: SPiM code and simulation results for the Oregonator reactions (7) with $c_1 = 2$, $c_2 = 0.1$, $c_3 = 104$, $c_4 = 0.016$ and $c_5 = 26$. Initial values $Y_1 = 500$, $Y_2 = 1000$, $Y_3 = 2000$. (a) $Y_1, Y_2, Y_3$ vs. $t$ for $0 \leq t \leq 6$. (b) $Y_2$ vs. $Y_1$ for $0 \leq t \leq 6$ (c) $Y_3$ vs. $Y_1$ (d) $Y_3$ vs. $Y_2$

$$
\begin{array}{rcl}
X_1 + Y_2 & \xrightarrow{c_1} & X_1 + Y_1 \\
Y_1 + Y_2 & \xrightarrow{c_2} & Z_1 \\
X_2 + Y_1 & \xrightarrow{c_3} & X_2 + 2Y_1 + Y_3 \\
2Y_1 & \xrightarrow{c_4} & Z_2 \\
X_3 + Y_3 & \xrightarrow{c_5} & X_3 + Y_2
\end{array}
\tag{7}
$$

**Definition A.1.** Oregonator Reactions




a b

Figure 6: Simulation of Oregonator reactions (7) as in Figure 24, but with a limited number of $X_1$ species. Rate $c_1 = 0.0002$. Initial values $X_1 = 10^4$. Results for $Y_1, Y_2, Y_3$ vs. $t$ and $X_1$ vs. $t$, respectively.

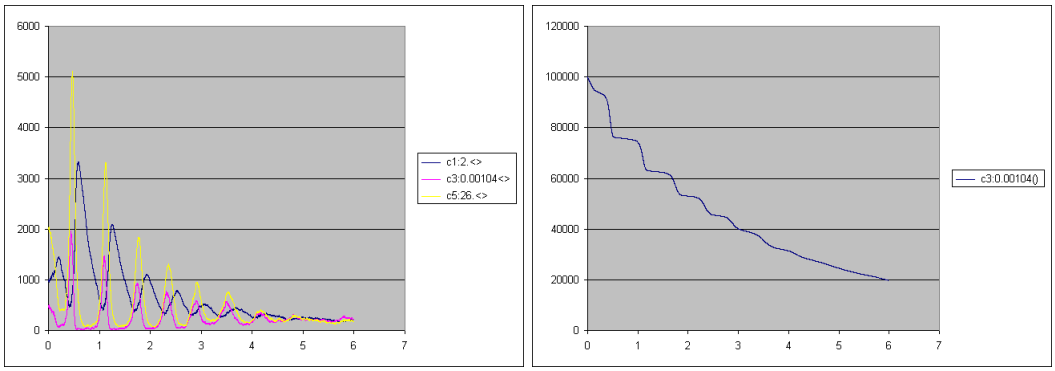



Figure 7: Simulation of Oregonator reactions (7) as in Figure 24, but with a limited number of $X_2$ species. Rate $c_3 = 0.00104$. Initial values $X_2 = 10^5$. Results for $Y_1, Y_2, Y_3$ vs. $t$ and $X_2$ vs. $t$, respectively.