

# **Diffusing Ideas**

Software, noise and building mathematical toys

**Robert J. Hardwick**

February 23, 2023



# Introduction

*Diffusing Ideas* is a book of research exploration and software development which I have written for the interest of mathematically-inclined programmers and computational scientists. It's the output of many interrelated projects over several years which have sought to generalise the computational mathematics of simulating, statistically inferring, manipulating and automatically controlling stochastic phenomena as far as possible.

The book accompanies a lot of new open-source scientific software written predominantly in Go [1], with a little Python [2] to top it off. A major motivation for creating these new tools is to prepare a foundation of code from which to develop new and more complex applications. I also hope that the resulting framework will enable anyone to explore and study new phenomena effectively, regardless of their scientific background.

The need to properly test all this software has also provided a wonderful excuse to study and play with an extensive range of mathematical toy models. I've chosen these models based on a fairly broad background of interests, but also to illustrate the remarkable cross-disciplinary applicability of stochastic processes. However, I've often found that mathematical formalities can obscure the computations that a programmer must implement. So, while I've tried to be as ambitious as possible with the level of technical sophistication in these models, I've also tried to write the expressions in a computer-friendly way where feasible<sup>1</sup> and added code snippets throughout.

A quick note on the code: any software that I describe in this book (including the software which compiles the book itself [3]) will always be shared under a MIT License [4] in a public Git repository.<sup>2</sup> Forking these repositories and submitting pull requests for new features or applications is strongly encouraged too, though I apologise in advance if I don't follow these up very quickly as all of this work has to be conducted independently in free time, outside of work hours.

No quest would be complete without a guide, so I think this introduction should end with a list of the key milestones in the book; comprising its four major parts. These parts each correspond to answering one of the following interdependent research questions:

**Part 1.** How do we simulate a general set of stochastic phenomena?

**Part 2.** How do we then learn/identify the answer to **Part 1** from real-world data?

**Part 3.** How do we simulate a general set of control policies to interact with the answer to **Part 1**?

**Part 4.** How do we then optimise the answer to **Part 3** to achieve a specified control objective?

---

<sup>1</sup>For example, I'll typically be thinking more in terms of 'matrices' and less about 'operators'.

<sup>2</sup>The repositories will always be somewhere on this list: <https://github.com/umbralcalc?tab=repositories>.



# Table of contents

<b>1</b>	<b>Building a generalised simulator</b>	<b>3</b>
1.1	Computational formalism . . . . .	3
1.2	Useful probabilistic concepts . . . . .	10
1.3	Software design . . . . .	14
1.4	Implementation details . . . . .	16
<b>2</b>	<b>Simulating a financial market</b>	<b>19</b>
2.1	Introducing Q-Hawkes processes . . . . .	19
<b>3</b>	<b>Quantum jumps on generic networks</b>	<b>21</b>
3.1	The Lindblad equation . . . . .	21
<b>4</b>	<b>Empirical dynamical emulators</b>	<b>25</b>
<b>5</b>	<b>Inferring dynamical 2D maps</b>	<b>27</b>
5.1	Adapting the stochadex formalism . . . . .	27
<b>6</b>	<b>Learning from ants on curved surfaces</b>	<b>29</b>
6.1	Diffusive limits for ant interactions . . . . .	29
<b>7</b>	<b>Hydrodynamic ensembles from input data</b>	<b>31</b>
7.1	The Boltzmann/Navier-Stokes equations . . . . .	31
<b>8</b>	<b>Generalised statistical inference tools</b>	<b>33</b>
8.1	Likelihood-free methods . . . . .	33
<b>9</b>	<b>Interacting with systems in general</b>	<b>37</b>
9.1	Parameterising general interactions . . . . .	37
<b>10</b>	<b>Angling for freshwater fish</b>	<b>39</b>
10.1	A large-scale Lotka-Volterra model . . . . .	39
<b>11</b>	<b>Managing a Rugby match</b>	<b>41</b>
11.1	Introduction . . . . .	41
11.2	Designing the event simulation engine . . . . .	41

11.3 Linking to player attributes . . . . .	42
11.4 Deciding on gameplay actions . . . . .	42
11.5 Writing the game itself . . . . .	42
<b>12 Influencing house prices</b>	<b>43</b>
<b>13 Optimising actions for control objectives</b>	<b>47</b>
<b>14 Resource allocation for epidemics</b>	<b>49</b>
<b>15 Quantum system control</b>	<b>51</b>
<b>16 Other models</b>	<b>53</b>

# **Part 1. How do we simulate a general set of stochastic phenomena?**





# Building a generalised simulator

**Concept.** To design and build a generalised simulation engine that is able to generate samples from a ‘Pokédex’ of possible stochastic processes that a researcher might encounter. A ‘Pokédex’ here is just my fanciful description for a very general class of multidimensional stochastic processes that pop up everywhere in taming the mathematical wilds of real-world phenomena, and which also leads to a name for the software itself: the ‘stochadex’. With such a thing pre-built and self-contained, it can become the basis upon which to build generalised software solutions for a lot of different problems. For the mathematically-inclined, this chapter will require the introduction of a new formalism which we shall refer back to throughout the book. For the programmers, the public Git repository for the code that is described in this chapter can be found here: <https://github.com/umbralcalc/stochadex>.

## 1.1 Computational formalism

Before we dive into software design of the stochadex, we need to mathematically define the general computational approach that we’re going to take. Because the language of stochastic processes is primarily mathematics, I’d argue this step is essential in enabling a really general description. From experience, it seems reasonable to start by writing down the following formula which describes iterating some arbitrary process forward in time (by one finite step) and adding a new row each to some matrix  $X' \rightarrow X$

$$X_{t+1}^i = F_{t+1}^i(X', t), \quad (1.1)$$

where:  $i$  is an index for the dimensions of the ‘state’ space;  $t$  is the current time index for either a discrete-time process or some discrete approximation to a continuous-time process;  $X$  is the next version of  $X'$  after one timestep (and hence one new row has been added); and  $F_{t+1}^i(X', t)$  as the latest element of an arbitrary matrix-valued function. As we shall discuss shortly,  $F_{t+1}^i(X', t)$  may represent not just operations on deterministic variables, but also on stochastic ones. There is also no requirement for the function to be continuous.

The basic computational idea here is illustrated in Fig. 1.1; we iterate the matrix  $X$  forward in time by a row, and use its previous version  $X'$  as an entire matrix input into a function which

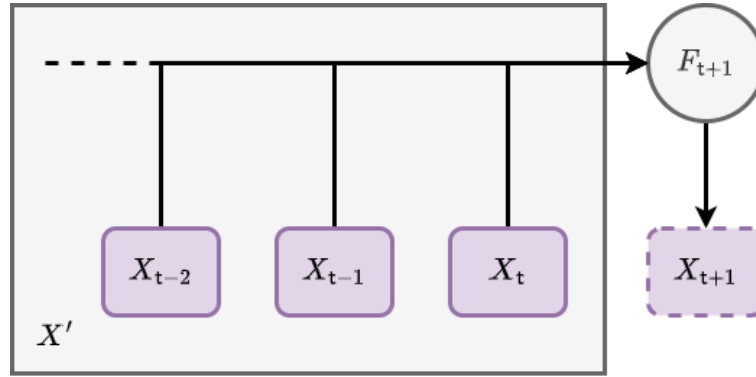


Figure 1.1: Graph representation of Eq. (1.1).

populates the elements of its latest rows. In Go you could easily write something with the same idea in it, and the code would probably look like this.

```

1 type StateElement float64
2 type StateVector []StateElement
3 type StateHistory []StateVector
4
5 // iterate the state history forward in time by one step
6 func IterationFormula(
7     stateHistory StateHistory,
8     timeStepNumber TimeStepNumber,
9 ) StateVector {
10     for t, stateVector := range stateHistory {
11         for i, stateElement := range stateVector {
12             // do something
13         }
14     }
15     return newestStateVector
16 }
```

Pretty simple! But why go to all this trouble of storing matrix inputs for previous values of the same process? It's true that this is mostly redundant for *Markovian* phenomena, i.e., processes where their only memory of their history is the most recent value they took. However, for a large class of stochastic processes a full memory<sup>1</sup> of past values is essential to consistently construct the sample paths moving forward. This is true in particular for *non-Markovian* phenomena, where the latest values don't just depend on the immediately previous ones but can depend on values which occurred much earlier in the process as well.

For more complex physical models and integrators, the distinct notions of 'numerical timestep' and 'total elapsed continuous time' will crop up quite frequently. Hence, before moving on further details, it will be important to define the total elapsed time variable  $t(t)$  for processes which are defined in continuous time. Assuming that we have already defined some function  $\delta t(t)$  which returns the specific change in continuous time that corresponds to the step  $t-1 \rightarrow t$ , we will always

<sup>1</sup>Or memory at least within some window.

be able to compute the total elapsed time through the relation

$$t(t) = \sum_{t'=0}^t \delta t(t'). \quad (1.2)$$

This seems a lot of effort, no? Well it's important to remember that our steps in continuous time may not be constant, so by defining the  $\delta t(t)$  function and summing over it we can enable this flexibility in the computation. In case the summation notation is no fun for programmers; in Go we're implicitly doing this.

```

1 type TimeStepNumber int
2
3 // both below are in some unit, e.g., seconds
4 type TimeIncrement    float64
5 type TotalElapsedTime float64
6
7 // get the next increment from this step number forward
8 func TimeIncrementFunction(timeStepNumber TimeStepNumber) TimeIncrement {
9     // compute the next increment
10    return nextTimeIncrement
11 }
12
13 // compute the total time elapsed up to the input step number
14 func TotalTimeElapsedFunction(timeStepNumber TimeStepNumber) TotalElapsedTime {
15     totalElapsedTime := 0.0
16     for t := 0; t < timeStepNumber; t++ {
17         totalElapsedTime += TimeIncrementFunction(t)
18     }
19     return totalElapsedTime
20 }
```

So, now that we've mathematically defined a really general notion of iterating the stochastic process forward in time, it makes sense to discuss some simple examples. For instance, it is frequently possible to split  $F$  up into deterministic (denoted  $D$ ) and stochastic (denoted  $S$ ) matrix-valued functions like so

$$F_{t+1}^i(X', t) = D_{t+1}^i(X', t) + S_{t+1}^i(X', t). \quad (1.3)$$

In the case of stochastic processes with continuous sample paths, it's also nearly always the case with mathematical models of real-world systems that the deterministic part will at least contain the term  $D_{t+1}^i(X', t) = X_t^i$  because the overall system is described by some stochastic differential equation. This is not a really requirement in our general formalism, however.

What about the stochastic term? For example, if we wanted to consider a *Wiener process noise*, we can define  $W_t^i$  is a sample from a Wiener process for each of the state dimensions indexed by  $i$  and our formalism becomes

$$S_{t+1}^i(X', t) = W_{t+1}^i - W_t^i. \quad (1.4)$$

One draws the increments  $W_{t+1}^i - W_t^i$  from a normal distribution with a mean of 0 and a variance equal to the length of continuous time that the step corresponded to  $\delta t(t+1)$ , i.e., the probability density  $P_{t+1}^i(x)$  of the increments  $x^i = W_{t+1}^i - W_t^i$  is

$$P_{t+1}^i(x) = \text{NormalPDF}[x^i; 0, \delta t(t+1)]. \quad (1.5)$$

Note that for state spaces with dimensions  $> 1$ , we could also allow for non-trivial cross-correlations between the noises in each dimension. In Go, the Wiener process can be represented within our formalism like this.

```

1 import (
2     "math"
3     "math/rand"
4 )
5
6 // generate a new Wiener process increment for a state element
7 func NewWienerProcessIncrement(timeStepNumber TimeStepNumber) StateElement {
8     // use the time increment function we defined earlier
9     timeIncrement := TimeIncrementFunction(timeStepNumber)
10    // multiply by the square-root here as
11    // it is proportional to the standard deviation
12    value := math.Sqrt(timeIncrement) * rand.NormFloat64()
13    return StateElement(value)
14 }
15
16 // returns the state vector from the S(X',t) function we defined
17 // in the main text above
18 func SFunction(
19     stateHistory StateHistory,
20     timeStepNumber TimeStepNumber,
21 ) StateVector {
22     stateVector := stateHistory[0]
23     sFunctionValue := make(StateVector, 0)
24     // note that we don't care about the history in this case = Markovian
25     for _ = range stateVector {
26         increment := NewWienerProcessIncrement(timeStepNumber)
27         sFunctionValue = append(sFunctionValue, increment)
28     }
29     return sFunctionValue
30 }

```

In another example, to model *geometric Brownian motion noise* we would simply have to multiply  $X_t^i$  to the Wiener process like so

$$S_{t+1}^i(X', t) = X_t^i(W_{t+1}^i - W_t^i). \quad (1.6)$$

Here we have implicitly adopted the Itô interpretation to describe this stochastic integration. Given a carefully-defined integration scheme other interpretations of the noise would also be possible with our formalism too, e.g., Stratonovich<sup>2</sup> or others within the more general ‘ $\alpha$ -family’ [5, 6, 7]. The Go code for any of these should hopefully be fairly straightforward to deduce based on the lines I’ve already written above.

We can imagine even more general processes that are still Markovian. One example of these in a single-dimension state space would be to define the noise through some general function of the Wiener process like so

$$S_{t+1}^0(X', t) = g[W_{t+1}^0, t(t+1)] - g[W_t^0, t(t)] \quad (1.7)$$

$$= \left[ \frac{\partial g}{\partial t} + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \right] \delta t(t+1) + \frac{\partial g}{\partial x} (W_{t+1}^0 - W_t^0), \quad (1.8)$$

---

<sup>2</sup>Which would implicitly give  $S_{t+1}^i(X', t) = (X_{t+1}^i + X_t^i)(W_{t+1}^i - W_t^i)/2$  for Eq. (1.6).

where  $g(x, t)$  is some continuous function of its arguments which has been expanded out with Itô's Lemma on the second line. Note also that the computations in Eq. (1.8) could be performed with numerical derivatives in principle, even if the function were extremely complicated. This is unlikely to be the best way to describe the process of interest, however, the mathematical expressions above can still be made a bit more meaningful to the programmer in this way. The code would probably look something like this.

```

1 // some function
2 func G(wienerProcessSample float64, timeStepNumber TimeStepNumber) float64 {
3     // return something
4 }
5
6 // discretely represents the dg/dt expression in the equation above
7 func DgDt(
8     newWienerProcessSample float64,
9     previousWienerProcessSample float64,
10    timeStepNumber TimeStepNumber,
11 ) float64 {
12     return (G(newWienerProcessSample, timeStepNumber) -
13             G(previousWienerProcessSample, timeStepNumber)) /
14             float64(TimeIncrementFunction(timeStepNumber))
15 }
16
17 // discretely represents the dg/dx expression in the equation above
18 func DgDx(
19     newWienerProcessSample float64,
20     previousWienerProcessSample float64,
21     timeStepNumber TimeStepNumber,
22 ) float64 {
23     return (G(newWienerProcessSample, timeStepNumber) -
24             G(previousWienerProcessSample, timeStepNumber)) /
25             (newWienerProcessSample - previousWienerProcessSample)
26 }
27
28 // discretely represents the d^2g/dx^2 expression in the equation above
29 func D2gDx2(
30     // newDgDx and previousDgDx could be passed in here
31     newWienerProcessSample float64,
32     previousWienerProcessSample float64,
33     timeStepNumber TimeStepNumber,
34 ) float64 {
35     // newDgDx and previousDgDx are the result of applying the function
36     // for dg/dx defined above on two different timesteps
37     return (newDgDx - previousDgDx) /
38             (newWienerProcessSample - previousWienerProcessSample)
39 }

```

Let's now look at a more complicated type of noise. For example, *fractional Brownian motion*  $[B_H]_t$  with Hurst exponent  $H$ . Following Ref. [8], we can simulate this process in one of our state space dimensions by modifying a standard Wiener process like so

$$S_{t+1}^0(X', t) = \frac{(W_{t+1}^0 - W_t^0)}{\delta t(t)} \int_{t(t)}^{t(t+1)} dt' \frac{(t' - t)^{H-\frac{1}{2}}}{\Gamma(H + \frac{1}{2})} {}_2F_1\left(H - \frac{1}{2}; \frac{1}{2} - H; H + \frac{1}{2}; 1 - \frac{t'}{t}\right), \quad (1.9)$$

where  $S_{t+1}^0(X', t) = [B_H]_{t+1} - [B_H]_t$ . The integral in Eq. (1.9) can be approximated using an appropriate numerical procedure (like the trapezium rule). In the expression above  ${}_2F_1$  and  $\Gamma$  are

the ordinary hypergeometric and gamma functions, respectively. A discretised form of this integral is written below in Go to try and disentangle some of the complexity in the expression.

```

1 import "scientificgo.org/special"
2
3 type HurstExponent float64
4
5 // computes the integral term in the fractional Brownian motion process
6 // defined above using some scientific libraries
7 func ComputeTheIntegral(
8     currentTime TotalElapsedTime,
9     nextTime TotalElapsedTime,
10    hurstExponent HurstExponent,
11    numberOfIntegrationSteps int,
12 ) float64 {
13     currentT := float64(currentTime)
14     nextT := float64(nextTime)
15     integralStepSize := nextT - currentT/float64(numberOfIntegrationSteps)
16     h := float64(hurstExponent)
17     a := []float64{h - 0.5, 0.5 - h}
18     b := []float64{h + 0.5}
19     integralValue := 0.0
20     // implements the trapezium rule in a loop over the steps
21     // between the current and the next point in time
22     for t := 0; t < numberOfIntegrationSteps; t++ {
23         t1 := currentT + float64(t)*integralStepSize
24         t2 := t1 + integralStepSize
25         functionValue1 := (math.Pow(t1-currentT, h-0.5) / math.Gamma(h+0.5)) *
26             special.HypPFQ(a, b, 1.0-t1/currentT)
27         functionValue2 := (math.Pow(t2-currentT, h-0.5) / math.Gamma(h+0.5)) *
28             special.HypPFQ(a, b, 1.0-t2/currentT)
29         integralValue += 0.5 * (functionValue1 + functionValue2) * integralStepSize
30     }
31     return integralValue
32 }

```

So far we have mostly been discussing noises with continuous sample paths, but we can easily adapt our computation to discontinuous sample paths as well. For instance, *Poisson process noises* would generally take the form

$$S_{t+1}^i(X', t) = [N_\lambda]_{t+1}^i - [N_\lambda]_t^i, \quad (1.10)$$

where  $[N_\lambda]_t^i$  is a sample from a Poisson process with rate  $\lambda$ . One can think of this process as counting the number of events which have occurred up to the given interval of time, where the intervals between each successive event are exponentially distributed with mean  $1/\lambda$ . Such a simple counting process could be simulated exactly by explicitly setting a newly-drawn exponential variate to the next continuous time jump  $\delta t(t+1)$  and iterating the counter. Other exact methods exist to handle more complicated processes involving more than one type of ‘event’, such as the Gillespie algorithm [9] — though these techniques are not always applicable in every situation.

Is using step size variation always possible? If we consider a *time-inhomogeneous Poisson process noise*, which would generally take the form

$$S_{t+1}^i(X', t) = [N_{\lambda(t+1)}]_{t+1}^i - [N_{\lambda(t)}]_t^i, \quad (1.11)$$

the rate  $\lambda(t)$  has become a deterministically-varying function in time. In this instance, it likely not be accurate to simulate this process by drawing exponential intervals with a mean of  $1/\lambda(t)$

because this mean could have changed by the end of the interval which was drawn. An alternative approach (which is more generally capable of simulating jump processes but is an approximation) first uses a small time interval  $\tau$  such that the most likely thing to happen in this period is nothing, and then the probability of the event occuring is simply given by

$$p(\text{event}) = \frac{\lambda(t)}{\lambda(t) + \frac{1}{\tau}}. \quad (1.12)$$

This idea can be applied to phenomena with an arbitrary number of events and works well as a generalised approach to event-based simulation, though its main limitation is worth remembering; in order to make the approximation good,  $\tau$  often must be quite small and hence our simulator must churn through a lot of steps. From now on we'll refer to this well-known technique as the *rejection method*.

There are a few extensions to the simple Poisson process that introduce additional stochastic processes. *Cox (doubly-stochastic) processes*, for instance, are basically where we replace the time-dependent rate  $\lambda(t)$  with independent samples from some other stochastic process  $\Lambda(t)$ . For example, a Neyman-Scott process [10] can be mapped as a special case of this because it uses a Poisson process on top of another Poisson process to create maps of spatially-distributed points. In our formalism, a two-state implementation of the Cox process noise would look like

$$S_{t+1}^0(X', t) = \Lambda(t+1) \quad (1.13)$$

$$S_{t+1}^1(X', t) = [N_{S_{t+1}^0}]_{t+1}^i - [N_{S_t^0}]_t^i. \quad (1.14)$$

Another extension is *compound Poisson process noise*, where it's the count values  $[N_\lambda]_t^i$  which are replaced by independent samples  $[J_\lambda]_t^i$  from another probability distribution, i.e.,

$$S_{t+1}^i(X', t) = [J_\lambda]_{t+1}^i - [J_\lambda]_t^i. \quad (1.15)$$

Note that the rejection method of Eq. (1.12) can be employed effectively to simulate any of these extensions as long as a sufficiently small  $\tau$  is chosen.

All of the examples we have discussed so far are Markovian. Given that we have explicitly constructed the formalism to handle non-Markovian phenomena as well, it would be worthwhile going some examples of this kind of process too. *Self-exciting process noises* would generally take the form

$$S_{t+1}^0(X', t) = \mathcal{I}_{t+1}(X', t) \quad (1.16)$$

$$S_{t+1}^1(X', t) = [N_{S_{t+1}^0}]_{t+1}^i - [N_{S_t^0}]_t^i, \quad (1.17)$$

where the stochastic rate  $\mathcal{I}_{t+1}(X', t)$  now depends on the history explicitly. Amongst other potential inputs we can see, e.g., Hawkes processes [11] as an example of above by substituting

$$\mathcal{I}_{t+1}(X', t) = \mu + \sum_{t'=0}^t \gamma[t(t) - t(t')](S_{t'}^1 - S_{t'-1}^1), \quad (1.18)$$

where  $\gamma$  is the 'exciting kernel' and  $\mu$  is some constant background rate.

Note that this idea of integration kernels could also be applied back to our Wiener process. For example, another type of non-Markovian phenomenon that frequently arises across physical and

life systems integrates the Wiener process history like so

$$S_{t+1}^0(X', t) = W_{t+1}^0 - W_t^0 \quad (1.19)$$

$$S_{t+1}^1(X', t) = \frac{1}{T} \sum_{t'=0}^t e^{-\frac{t(t)-t(t')}{T}} (S_{t'}^0 - S_{t'-1}^0), \quad (1.20)$$

where  $T$  is some decay coefficient which quantifies the length of memory in continuous time.

So we've introduced the basic elements of our computational formalism and demonstrated how flexible the approach can be in simulating just about any stochastic phenomenon imaginable. Before progressing to algorithm design, it will be helpful to discuss some useful concepts that should enable us analyse the system later on in the book.

## 1.2 Useful probabilistic concepts

The general stochastic process that we defined with Eq. (1.1) also has an implicit *master equation* associated to it which fully describes the time evolution of the *probability density function*  $P_{t+1}(x)$  of the most recent matrix row  $x = X_{t+1}$  at time  $t$ . This can be written as

$$P_{t+1}(x) = \frac{1}{t} \sum_{t'=0}^t \int_{\omega_{t'}} dx' P_{t'}(x') P_{(t+1)t'}(x|x'), \quad (1.21)$$

where at the moment we are assuming the state space is continuous in each dimension and  $P_{(t+1)t'}(x|x')$  is the conditional probability that the matrix row at time  $(t+1)$  will be  $x = X_{t+1}$  given that the row at time  $t'$  was  $x' = X_{t'}$ . This is a very general equation which should almost always apply to any continuous stochastic phenomenon we want to study in due course. To try and understand what this equation is saying, I find it's helpful to think of an iterative relationship between probabilities, which are connected by their relative conditional probabilities. This kind of thinking is also illustrated in Fig. 1.2.

The factor of  $1/t$  in Eq. (1.21) is a normalisation factor — this just normalises the sum of all probabilities to 1 given that there is a sum over  $t'$ . Note that, if the process is defined over continuous time, we would need to replace

$$\frac{1}{t} \sum_{t'=0}^t \rightarrow \frac{1}{t(t)} \sum_{t'=0}^t \delta t(t'). \quad (1.22)$$

But what is  $\omega_t$ ? You can think of this as just the domain of possible  $x'$  inputs into the integral which will depend on the specific stochastic process we are looking at.

What if we wanted the joint distribution of both rows  $P_{(t+1)t'}(x, x')$ ? One way to obtain this would be to extend Eq. (1.21) such that both matrix rows are marginalised over separately like so

$$P_{(t+1)t'}(x, x') = \frac{1}{(t'-1)t} \sum_{t''=0}^t \sum_{t'''=0}^{t'-1} \int_{\omega_{t''}} dx'' \int_{\omega_{t'''}} dx''' P_{t''t'''}(x'', x''') P_{(t+1)t''}(x|x'') P_{t't'''}(x'|x'''). \quad (1.23)$$



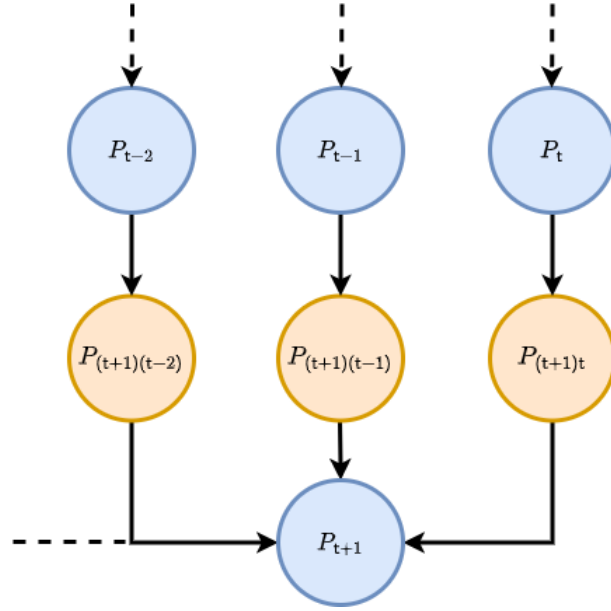


Figure 1.2: Graph representation of Eq. (1.21).

Given Eqs. (1.21) and (1.23) it's also possible to work out what the conditional probabilities would look like using the simple relation

$$P_{(t+1)t'}(x|x') = \frac{P_{(t+1)t'}(x, x')}{P_{t'}(x')}. \quad (1.24)$$

The implicit notation in Eq. (1.21) can hide some staggering complexity. To analyse the system in more detail, we can also do a kind of Kramers-Moyal expansion [12, 13] for each point in time to approximate the overall equation like this

$$\begin{aligned} P_{t+1}(x) = & \frac{1}{t} \sum_{t'=0}^t P_{t'}(x) - \frac{1}{t} \sum_{t'=0}^t \sum_{i=1}^d \frac{\partial}{\partial x^i} [\alpha_{(t+1)t'}^i(x) P_{t'}(x)] \\ & + \frac{1}{2t} \sum_{t'=0}^t \sum_{i=1}^d \sum_{j=1}^d \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} [\beta_{(t+1)t'}^{ij}(x) P_{t'}(x)] + \dots, \end{aligned} \quad (1.25)$$

in which we have assumed that the state space is  $d$ -dimensional. In this expansion, we also needed to define these new integrals

$$\alpha_{(t+1)t'}^i(x) = \int_{\omega_{t'}} dx' (x' - x)^i P_{(t+1)t'}(x'|x) \quad (1.26)$$

$$\beta_{(t+1)t'}^{ij}(x) = \int_{\omega_{t'}} dx' (x' - x)^i (x' - x)^j P_{(t+1)t'}(x'|x). \quad (1.27)$$

So the matrix notation of Eq. (1.21) can indeed hide a very complicated calculation. Truncating the expansion at second-order, Eq. (1.25) tells us that there can be first and second derivatives contributing to the flow of probability to each element of the row  $x = X_{t+1}$  which depend on every element of the matrix  $X'$ . The probability does indeed *flow*, in fact. We can define a quantity known as the ‘probability current’  $J_{(t+1)t'}(x)$  from  $t'$  to  $(t+1)$  which illustrates this through the following continuity relation

$$P_{t+1}(x) - \frac{1}{t} \sum_{t'=0}^t P_{t'}(x) = \frac{1}{t} \sum_{t'=0}^t [P_{t+1}(x) - P_{t'}(x)] = -\frac{1}{t} \sum_{t'=0}^t J_{(t+1)t'}(x). \quad (1.28)$$

By inspection of Eq. (1.25) we can therefore also deduce that

$$J_{(t+1)t'}^i(x) = \alpha_{(t+1)t'}^i(x) P_{t'}(x) - \frac{1}{2} \sum_{j=1}^d \frac{\partial}{\partial x^j} [\beta_{(t+1)t'}^{ij}(x) P_{t'}(x)] + \dots \quad (1.29)$$

What would happen if we assumed that  $\alpha$  and  $\beta$  were just arbitrary time-dependent functions? For example, let’s make the following assumptions

$$\alpha_{(t+1)t'}^i(x) = \mu^i(t') - x^i \quad (1.30)$$

$$\beta_{(t+1)t'}^{ij}(x) = 2\Sigma^{ij}(\theta, t'), \quad (1.31)$$

where  $\mu(t')$  is an arbitrary vector-valued function of the timestep and  $\Sigma(\theta, t')$  is an arbitrary matrix (often known as the ‘diffusion tensor’) which depends on both the timestep and a set of hyperparameters  $\theta$ . If we now also assume stationarity of  $P_{t'}(x) = P_{t''}(x)$  for any  $t'$  and  $t''$  such that

$$P_{t+1}(x) = \frac{1}{t} \sum_{t'=0}^t P_{t'}(x), \quad (1.32)$$

we can solve Eq. (1.25) to obtain the following stationary solution

$$P_{t'}(x) = \text{MultivariateNormalPDF}[x; \mu(t'), \Sigma(\theta, t')]. \quad (1.33)$$

I actually hid a little bit of the detail in that last step; the solution also required the identification that the flow of probability between timesteps vanishes uniquely for each and every  $t'$  such that  $J_{(t+1)t'}(x) = 0$ .

It’s possible to take this derivation a bit further by expanding Eq. (1.23) in a similar fashion, truncating it to second-order, assuming only time-dependent terms and then solving it in the stationary limit. By plugging this solution (and its corresponding marginal distribution equivalent) into Eq. (1.24), it’s possible to get something that looks like this conditional distribution

$$P_{(t+1)t'}(x|x') \propto \exp \left\{ -\frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d [x - f(t+1)]^i [K^{-1}(\theta, t+1, t+1)]^{ij} [x - f(t+1)]^j \right. \\ \left. + \sum_{i=1}^d \sum_{j=1}^d [x - f(t+1)]^i [K^{-1}(\theta, t+1, t')]^{ij} [x' - f(t')]^j \right\}, \quad (1.34)$$

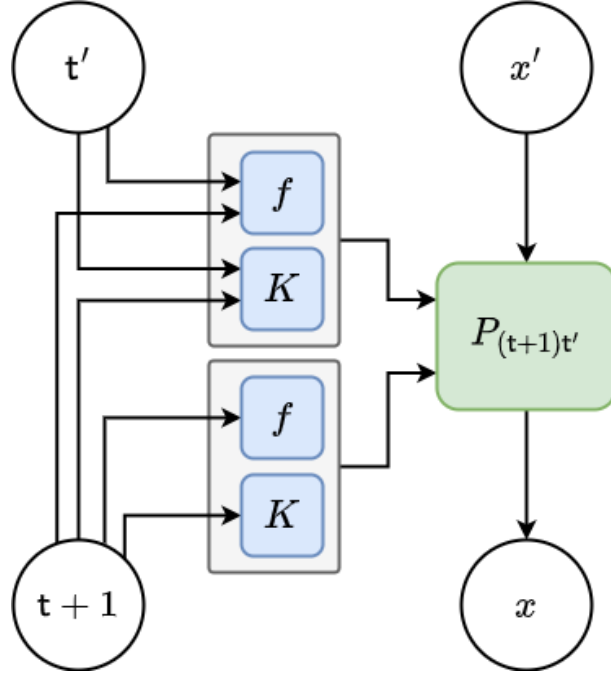


Figure 1.3: Graph representation of the generative process implied by Eq. (1.34).

where  $K(\theta, t+1, t')$  is some arbitrary covariance matrix that encodes how the correlation structure varies with the between compared states at two different timesteps and  $K^{-1}(\theta, t+1, t')$  denotes taking its inverse. Eq. (1.34) may look a bit familiar to some readers who like using Gaussian processes from the machine learning literature [14] — this version implies a *generative* model for a future  $x$  value (which I’ve illustrated in Fig 1.3), in contrast to the more standard equation used to *infer* values of  $f$ . These are two sides of the same coin though.

What other processes can be described by Eq. (1.21)? For Markovian phenomena, the equation no longer depends on timesteps older than the immediately previous one, hence the expression reduces to just

$$P_{t+1}(x) = \int_{\omega_t} dx' P_t(x') P_{(t+1)t}(x|x'). \quad (1.35)$$

It’s also easy to show that Eq. (1.25) naturally simplifies into the more usually applied Kramers-Moyal expansion when considering a Markovian process — you just remove the sum over  $t'$  and the  $1/t$  normalisation factor.

Note that an analog of Eq. (1.21) exists for discrete state spaces as well. We just need to replace the integral with a sum and the schematic would look something like this

$$P_{t+1}(x) = \frac{1}{t} \sum_{t'=0}^t \sum_{\omega_{t'}} P_{t'}(x') P_{(t+1)t'}(x|x'), \quad (1.36)$$

where we note that the  $P$ 's in the expression above all now refer to *probability mass functions*. Because the state space is now discrete, we cannot immediately intuit an approximative expansion from this expression.

Ok, as a brief aside; if one is really determined to use a similar approach to the one we derived above, it's quite straightforward to rewrite it in terms of continuous-valued characteristic functions like so

$$\varphi_{t+1}(s) = \frac{1}{t} \sum_{t'=0}^t \int_{\ell_{t'}} ds' \mathcal{C}(s') \varphi_{t'}(s') \varphi_{(t+1)t'}(s|s') \quad (1.37)$$

$$\mathcal{C}(s') = \frac{1}{(2\pi)^d} \sum_{\omega_{t'}} e^{-i(s' \cdot x')}, \quad (1.38)$$

where  $\ell_{t'}$  defines all the continuous values that the vector  $s'$  can possibly have at time  $t'$ . In the expression above,  $\mathcal{C}(s')$  acts like is a kind of comb<sup>3</sup> to map the continuous frequency domain of  $s'$  onto the discrete state space of  $x'$ . Note also that  $\mathcal{C}(s')$  uses the imaginary number  $i$  and, to be visually tidier, the dot product notation  $a \cdot b$  just means the sum of vector elements:  $a \cdot b = \sum_{\forall k} a^k b^k$ . In principle, one can perform an approximative expansion on Eq. (1.37) like we did for continuous state spaces. This isn't always the most practical way of analysing the system though.

We have one more important example to discuss and then we can cap off this analysis subsection. In the even-simpler case where  $x$  is just a vector of binary 'on' or 'off' states, Eq. (1.36) reduces to

$$P_{t+1}^i = \frac{1}{t} \sum_{t'=0}^t \sum_{j=1}^d P_{t'}^j P_{(t+1)t'}^{ij} = \frac{1}{t} \sum_{t'=0}^t \sum_{j=1}^d [P_{t'}^j A_{(t+1)t'}^{ij} + (1 - P_{t'}^j) B_{(t+1)t'}^{ij}], \quad (1.39)$$

where  $P_{t'}^i$  now represents the probability that element  $x^i = 1$  (is 'on') at time  $t'$ . The matrices  $A$  and  $B$  are defined as conditional probabilities where the previous state in time  $P_{t'}^j$  was either 'on' or 'off', respectively.

### 1.3 Software design

So I've proposed a computational formalism and done a bit of analysis on it to demonstrate that it can cope with a variety of different stochastic phenomena. Now I think we're ready to summarise what we want the stochadex software package to be able to do. But what's so complicated about Eq. (1.1)? Can't we just implement an iterative algorithm with a single function? It's true that the fundamental concept is very straightforward, but as I'll discuss in due course; the stochadex needs to have a lot of configurable features so that it's applicable in different situations. Ideally, the stochadex sampler should be designed to try and maintain a balance between performance and flexibility of utilisation.

If we begin with the obvious first set of criteria; we want to be able to freely configure the iteration function  $F$  of Eq. (1.1) and the timestep function  $t$  of Eq. (1.2) so that any process we want can be described. The point at which a simulation stops can also depend on some algorithm termination condition which the user should be able to specify up-front.

Once someone has written the code to create these functions for the stochadex, I want to then be able to recall them in future only with configuration files while maintaining the possibility

---

<sup>3</sup>This is very similar to how a 'Dirac comb' works in signal processing [15].

of changing their simulation run parameters. This flexibility should facilitate our uses for the simulation later in the book, and from this perspective it also makes sense that the parameters should include the random seed and initial state value.

The state history matrix  $X$  should be configurable in terms of its number of rows — what we’ll call the ‘state width’ — and its number of columns — what we’ll call the ‘state history depth’. If we were to keep increasing the state width up to millions of elements or more, it’s likely that on most machines the algorithm performance would grind to a halt when trying to iterate over the resulting  $X$  within a single thread. Hence, before the algorithm or its performance in any more detail, we can pre-empt the requirement that  $X$  should be represented in computer memory by a set of partitioned matrices which are all capable of communicating to one-another downstream. In this paradigm, I’d like the user to be able to configure which state partitions are able to communicate with each other without having to write any new code.

For convenience, it seems sensible to also make the outputs from stochadex runs configurable. A user should be able to change the form of output that they want through, e.g., some specified function of  $X$  at the time of outputting data. The times that the stochadex should output this data can also be decided by some user-specified condition so that the frequency of output is fully configurable as well.

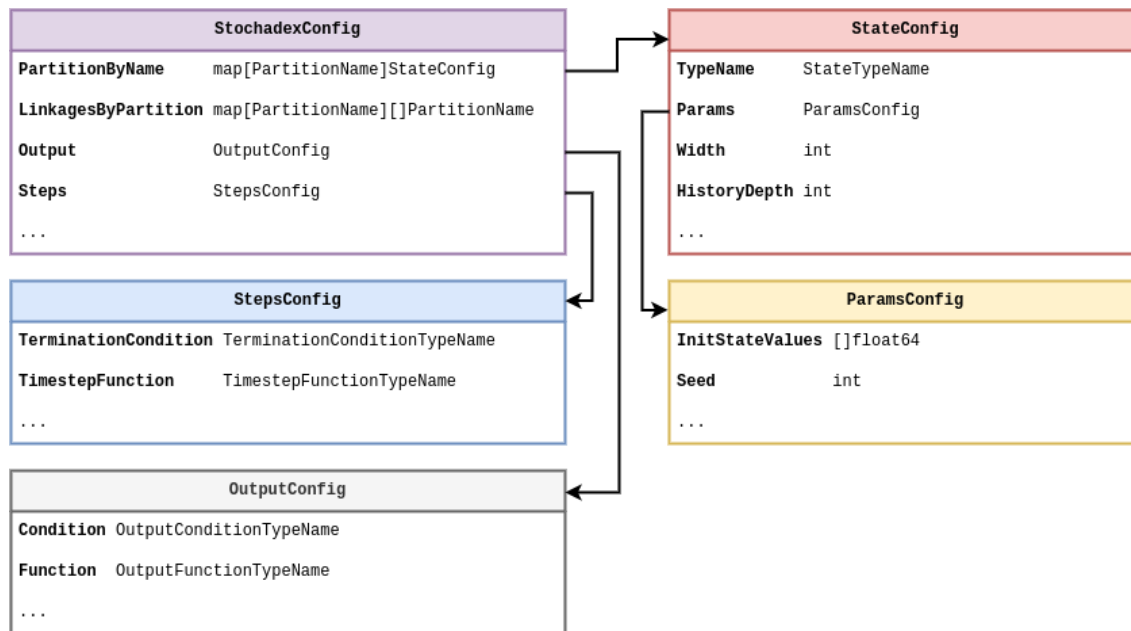


Figure 1.4: A relational summary of the core data types in the stochadex.

In summary, I’ve put together a schematic of data types and their relationships in Fig. 1.4. In this diagram there is some indication of the data type that I propose to store each piece information in (in Go syntax), and the diagram as a whole should serve as a useful guide to the basic structure of configuration files for the stochadex.

It’s clear that in order to simulate Eq. (1.1), we need an iterative algorithm which reapplies a user-specified function to the continually-updated history. But let’s now return to the point I

made earlier about how the performance of such an algorithm will depend on the size of the state history matrix  $X$ . The key bit of the algorithm design that isn't so straightforward is: how do we successfully split this state history up into separate partitions in memory while still enabling them to communicate effectively with each other? Other generalised simulation frameworks — such as SimPy [16], StoSpa [17] and FLAME GPU [18] — have all approached this problem in different ways, and with different software architectures.

In Fig. 1.5 I've illustrated what a loop involving two separate state partitions looks like in the stochadex simulator. Each partition is handled by concurrently running execution threads of the same process, while a separate process is used to handle the outputs from the algorithm. While this diagram illustrates only a single use for multiple processes, it's obviously true that we may run many of these whole diagrams at once on a multicore machine to generate a batch of independent ensembles if necessary.

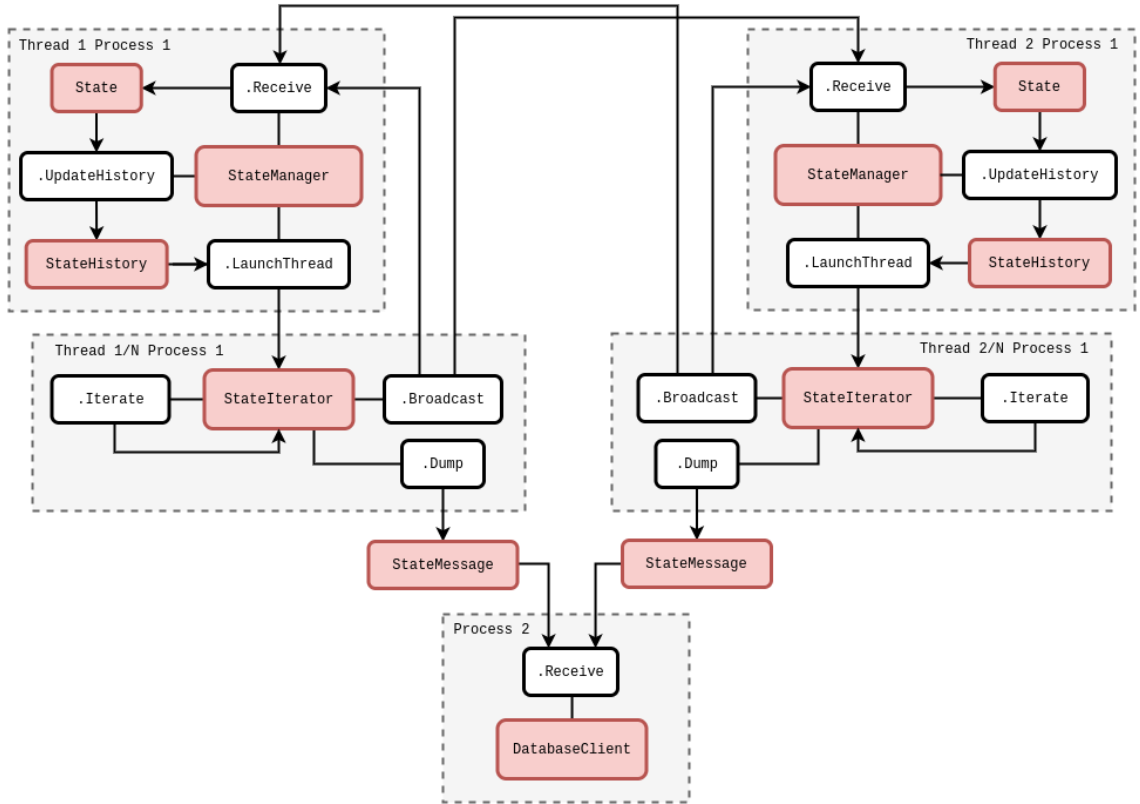


Figure 1.5: A loop of the stochadex simulation algorithm with two state partitions.

## 1.4 Implementation details

Now that the design of the algorithm and its basic data types have been outlined, we can finally turn to the implementation details. As I mentioned in the introduction, most of the core software in this

book has been written in Go — the main reason for this is mostly because it is a performant language that I enjoy developing with. More objectively, the key feature of Go that made it attractive for building the stochadex are its lightweight concurrency primitives called ‘goroutines’ [\[19\]](#) which made it easy and quick to piece together a scalable pipeline without excessive head-scratching. Ideal for projects in your limited free time!





# Simulating a financial market

**Concept.** The idea here is to use the Q-Hawkes processes and the Bouchaud work to come up with some interesting simulations of financial markets.

## 2.1 Introducing Q-Hawkes processes



# Quantum jumps on generic networks

**Concept.** The idea is to follow this sort of thing [here](#) to simulate the Lindblad equation over an arbitrary network of entangled states.

## 3.1 The Lindblad equation



**Part 2. How do we then  
learn/identify the answer to Part 1  
from real-world data?**



# Empirical dynamical emulators

**Concept.** The idea here is to leverage our probabilistic formalism to be able to empirically emulate a wide variety stochastic phenomena within a generalised framework.





# Inferring dynamical 2D maps

**Concept.** The idea here is

## 5.1 Adapting the stochadex formalism



# Learning from ants on curved surfaces

**Concept.** The idea here is

## 6.1 Diffusive limits for ant interactions



# Hydrodynamic ensembles from input data

**Concept.** The idea here is

## 7.1 The Boltzmann/Navier-Stokes equations



# Generalised statistical inference tools

**Concept.** The idea here is to extend the stochadex with tools for very generalised statistical inference (ABC algorithms and the like) that will work in nearly every situation. Probably need to exploit the phase space analogy of the formalism.

## 8.1 Likelihood-free methods





**Part 3. How do we simulate a  
general set of control policies to  
interact with the answer to Part 1?**



# Interacting with systems in general

**Concept.** The idea here is

## 9.1 Parameterising general interactions



# Angling for freshwater fish

**Concept.** The idea here is

## 10.1 A large-scale Lotka-Volterra model

Inspired by the empirical dynamical modelling approach to sockeye salmon in Ref. [20], but also desiring a generative model which has some link to the classic causal models promoted by mathematical ecology; the goal here is to create and calibrate a stochastic model which predicts the fish counts, weights, lengths and ages for each species in each area based on the past system states. To do this, we will combine some well-known models from mathematical ecology with supervised learning.

The one-step master equation for the proposed stochastic simulation is given implicitly by

$$\frac{d}{dt}P(\dots, n_i, \dots, t) = \sum_{\forall i} \mathcal{T}_i^+(\dots, n_i - 1, \dots, \mathbf{f}, t)P(\dots, n_i - 1, \dots, t) \quad (10.1)$$

$$+ \sum_{\forall i} \mathcal{T}_i^-(\dots, n_i + 1, \dots, \mathbf{f}, t)P(\dots, n_i + 1, \dots, t) \quad (10.2)$$

$$- \sum_{\forall i} \left[ \mathcal{T}_i^+(\dots, n_i, \dots, \mathbf{f}, t) + \mathcal{T}_i^-(\dots, n_i, \dots, \mathbf{f}, t) \right] P(\dots, n_i, \dots, t), \quad (10.3)$$

where the time  $t$  is defined in units of years and  $\mathcal{T}_i^+$  and  $\mathcal{T}_i^-$  are the transition coefficients for the  $i$ -th species, which depend not only on the counts for all species  $n_1, n_2, \dots$ , but also (in principle) on a larger feature space  $\mathbf{f}$  generated by the available data up to time  $t$ .

The famous Lotka-Volterra system, with some modifications for fishing and a larger set of species, would suggest transition coefficients of the form

$$\mathcal{T}_i^+(\dots, n_i, \dots, \mathbf{f}, t) = \mathcal{T}_i^+(\dots, n_i, \dots) = \Lambda_i(n_i) + n_i \alpha_i \sum_{\forall i' \text{ prey}} n_{i'} \quad (10.4)$$

$$\mathcal{T}_i^-(\dots, n_i, \dots, \mathbf{f}, t) = \mathcal{T}_i^-(\dots, n_i, \dots) = n_i \mu_i + n_i \gamma_i + n_i \beta_i \sum_{\forall i' \text{ pred}} n_{i'} , \quad (10.5)$$

where:  $\Lambda_i(n_i) = \tilde{\Lambda}_i n_i e^{-\lambda_i(n_i-1)}$  is the density-dependent birth rate;  $\mu_i$  is the species death rate;  $\alpha_i$  is the increase in the baseline birth rate per fish caused by the increase in prey population;  $\beta_i$  is the rate per fish of predation of the species; and  $\gamma_i$  accounts for the rate of recreational fishing per fish of the species. To approach the present data-driven simulation problem, we're going to generalise this model by training  $\mathcal{T}_i^+(\dots, n_i, \dots, \mathbf{f}, t)$  and  $\mathcal{T}_i^-(\dots, n_i, \dots, \mathbf{f}, t)$  directly from the data and generated features.

Look into the likelihood from, e.g., an electrofishing survey such as in Ref. [21]...

$$\text{Likelihood} = \sum_{\text{data}} \text{NB}[\text{data}; w_{i,\text{survey}} \langle n_i(t_{\text{data}}) \rangle, k_{i,\text{survey}}] , \quad (10.6)$$

# Managing a Rugby match

**Concept.** The idea here is

## 11.1 Introduction

Since the basic game engine will run using the [stochadex](#) sampler, the novelties in this project are all in the design of the rugby match model itself. And, in this instance, I'm not especially keen on spending a lot of time doing detailed data analysis to come up with the most realistic values for the parameters that are dreamed up here. Even though this would also be interesting.

One could do this data analysis, for instance, by scraping player-level performance data from one of the excellent websites that collect live commentary data such as [rugbypass.com](#) or [espn.co.uk/rugby](#).

This game is primarily a way of testing out the interface of the stochadex for other users to build projects with. This should help to both iron out some of the kinks in the design, as well as prioritise adding some more convenience methods for event-based modelling into its code base.

## 11.2 Designing the event simulation engine

We need to begin by specifying an appropriate event space to live in when simulating a rugby match. It is important at this level that events are defined in quite broadly applicable terms, as it will define the state space available to our stochastic sampler and hence the simulated game will never be allowed to exist outside of it. So, in order to capture the fully detailed range of events that are possible in a real-world match, we will need to be a little imaginative in how we define certain gameplay elements when we move through the space.

The diagrams below sum up what should hopefully work as a decent initial approximation while providing a little context with specific examples of play action.

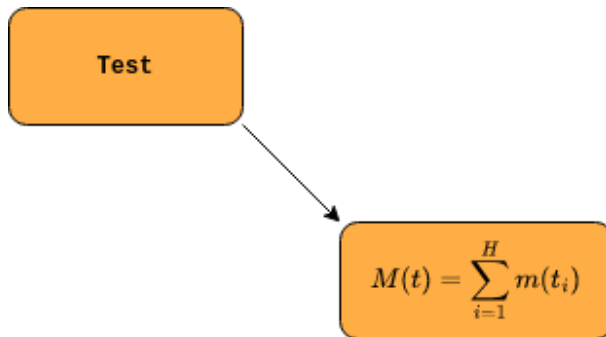


Figure 11.1: Simplified event graph of a rugby union match - replace with drawio.

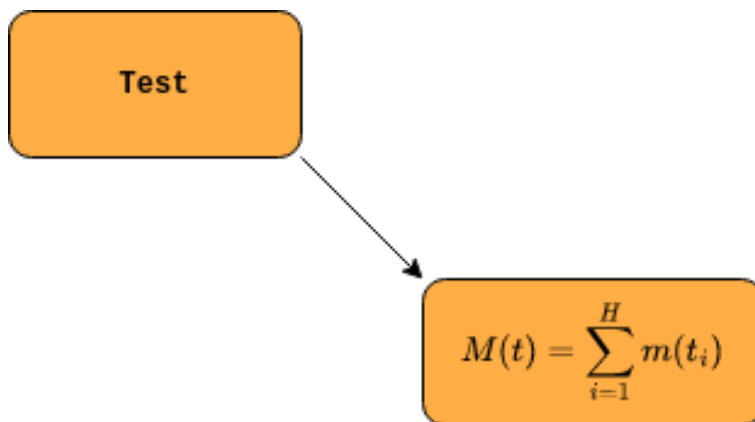


Figure 11.2: Optional model ideas - replace with drawio.

### 11.3 Linking to player attributes

### 11.4 Deciding on gameplay actions

### 11.5 Writing the game itself



# Influencing house prices

**Concept.** The idea here is



**Part 4. How do we then optimise the answer to Part 3 to achieve a specified control objective?**



# Optimising actions for control objectives

**Concept.** The idea controlhere is



## Resource allocation for epidemics

**Concept.** The idea here is to limit the spread of some abstract epidemic through the correct time-dependent resource allocation.





# Quantum system control

**Concept.** The idea here is to follow stuff along these lines [here](#).



## Other models

**Concept.** The idea here is



# Bibliography

- [1] “The Go Programming Language.” <https://go.dev/>.
- [2] “The Python Programming Language.” <https://www.python.org/>.
- [3] “The Diffusing Ideas GitHub Repository.” <https://github.com/umbralcalc/diffusing-ideas>.
- [4] “Open Source Initiative: MIT License.” <https://opensource.org/licenses/MIT>.
- [5] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. Elsevier, 1992.
- [6] H. Risken, *Fokker-planck equation*, in *The Fokker-Planck Equation*, pp. 63–95. Springer, 1996.
- [7] L. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales 2: Ito Calculus*, vol. 1, pp. xiv+480. Cambridge University Press, 04, 2000. 10.1017/CBO9781107590120.
- [8] L. Decreusefond et al., *Stochastic analysis of the fractional brownian motion, Potential analysis* **10** (1999) 177–214.
- [9] D. T. Gillespie, *Exact stochastic simulation of coupled chemical reactions, The journal of physical chemistry* **81** (1977) 2340–2361.
- [10] J. Neyman and E. L. Scott, *Statistical approach to problems of cosmology, Journal of the Royal Statistical Society: Series B (Methodological)* **20** (1958) 1–29.
- [11] A. G. Hawkes, *Spectra of some self-exciting and mutually exciting point processes, Biometrika* **58** (1971) 83–90.
- [12] H. A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions, Physica* **7** (1940) 284–304.
- [13] J. Moyal, *Stochastic processes and statistical physics, Journal of the Royal Statistical Society. Series B (Methodological)* **11** (1949) 150–210.
- [14] K. P. Murphy, *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [15] D. Brandwood, *Fourier transforms in radar and signal processing*. Artech House, 2012.
- [16] “SimPy: a process-based discrete-event simulation framework.” <https://gitlab.com/team-simpy/simpy/>.

- [17] “StoSpa: A C++ package for running stochastic simulations to generate sample paths for reaction-diffusion master equation.” <https://github.com/BartoszBartmanski/StoSpa>.
- [18] “FLAME GPU: A GPU accelerated agent-based simulation library for domain independent complex systems simulation.” <https://github.com/FLAMEGPU/FLAMEGPU2/>.
- [19] “A Tour of Go: Goroutines.” <https://go.dev/tour/concurrency/1>.
- [20] H. Ye, R. J. Beamish, S. M. Glaser, S. C. Grant, C.-h. Hsieh, L. J. Richards et al., *Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling*, *Proceedings of the National Academy of Sciences* **112** (2015) E1569–E1576.
- [21] “Electrofishing to assess a river’s health.” <https://environmentagency.blog.gov.uk/2015/10/29/electrofishing-to-assess-a-rivers-health/>.