

## TABLE OF CONTENTS

	Page
<b>1 Inferring species interactions</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Methodology . . . . .	2
1.2.1 Population models . . . . .	2
1.2.2 Simulation procedure . . . . .	4
1.2.3 Interaction strength . . . . .	4
1.2.4 Inference method . . . . .	4
1.2.5 Examples . . . . .	4
1.3 Results . . . . .	7
1.3.1 Lotka-Volterra system . . . . .	7
1.3.2 Holling system . . . . .	10
1.3.3 Range sampling . . . . .	10
1.4 TEMP : other results . . . . .	10
1.5 Application to IBM (optional) . . . . .	11
1.5.1 Testing functionl response . . . . .	11
1.5.2 Extend methodology (3 and 4 species) . . . . .	11
1.5.3 Results . . . . .	11
1.6 Discussion . . . . .	11
<b>Bibliography</b>	<b>15</b>



## INFERRING SPECIES INTERACTIONS

## 1.1 Motivation

This section focuses on the interactions between species. The importance of this topic was discussed at length in the introduction (section ??). However we re-iterate here that inter-specific interactions are one of the major mechanisms that drive ecological processes. For example trophic interactions (e.g. predator-prey) define the pathways along which energy flows through an ecological community [REF], and are responsible for both top-down and bottom-up regulation of species abundances [REF]. Similar mutualistic and competitive interactions play an important role in shaping communities [REF]. All inter-specific interactions play a role in generating the temporal pattern of species abundances that we call *population dynamics*. A well known example of this is the phenomenon of predator-prey oscillations, which are thought to have been exhibited by hare and lynx populations in Hudson Bay, Canada [REF]. Predator-prey oscillations will play an important role in this chapter.

The modelling approach in the previous chapters was *individual based*. Therefore in our simulations the interactions occurred between individuals of a given species, as they do in nature. However the possibility for two individuals to interact was determined by the species interaction network - if a link between two species exists in the network then individuals belonging to those species may interact.

.... Lacking suitable data - for which time-series and interaction (strengths) known..

Focus in predator-prey interactions..can be extended..

In this chapter we investigate one novel method for inferring species interaction strengths...

Outline exactly what we will do e.g. two species focus, prey-dependence

## 1.2 Methodology

The general goal is to quantify inter-specific interaction strengths from observed population dynamics. As discussed, we currently lack suitable empirical data to do this (section 1.1). Therefore we develop a methodology using *simulated* population dynamics. In the future this method could be applied to empirical abundance time-series (see discussion in section 1.6). The population models used for simulation are, for the most part, standard ordinary differential equation (ODE) models. These are outlined in section 1.2.1. However we also present a preliminary application of our methodology to dynamics generated using the individual-based model from the previous chapters (section 1.5). To quantify the interaction strengths between species from their simulated dynamics we adapt a method from [1]. In short the method is used to fit a generalised Lotka-Volterra (GLV) model to the dynamics, giving numerical estimates of the interaction terms. The method is detailed in section 1.2.4 and the use of the GLV model to quantify interaction strengths is justified in section 1.2.3.

### 1.2.1 Population models

To simulate population dynamics we use coupled ordinary differential equation (ODE) models. Although we focus on dynamics generated by predator-prey interactions, the ODE modelling framework may be adapted to model other types of interaction (e.g. mutualism, competition). ODE models have been used extensively to simulate predator-prey dynamics at the species level [REFS]. They have several limitations in their usefulness, which are discussed in 1.6. However they are a suitable choice for us because the equations explicitly contain a term that defines the interactions between species. Therefore we are able to simulate populations dynamics for which we know the analytic form of all inter-specific interactions involved. This is the key to being able to test our method for quantifying interaction strengths [SECTION?].

We focus on two-species systems. However the modelling framework, and the methodology in general may be extended to larger systems (see section 1.5). The growth rate for each species is defined by an ODE, which takes the general form:

$$(1.1) \quad \frac{dx_i}{dt} = G_i(x_i) + \sum_{j=1}^N F_{ij}(x_i, x_j) + \xi_i,$$

where  $x_i$  is the population density of species  $i$ ;  $N$  is the number of species;  $G_i$  is the intrinsic growth function of species  $i$ ;  $F_{ij}$  is a functional coupling between species  $i$  and  $j$ ; and  $\xi_i$  is an additive noise term (see below).

We now restrict the discussion to two species systems. However all that follows may be extended to larger systems with little modification. This possibility is discussed in section 1.6<sup>1</sup>.

---

<sup>1</sup>Or results presented in section...

Writing equation 1.1 for both species, and expanding the coupling function  $F_{ij}$  in terms of the *functional response*, we can express the entire system as:

$$(1.2) \quad \begin{aligned} \frac{dx_0}{dt} &= G_0(x_0) + a_{01}x_1H(x_0, x_1) + \xi_0, \\ \frac{dx_1}{dt} &= G_1(x_1) + a_{10}x_1H(x_0, x_1) + \xi_1 \end{aligned}$$

where species 0 is the prey; species 1 is the predator; the  $a_{ij}$  are constant coefficients and  $H(x_0, x_1)$  is the functional response (FR) of the predator. The other symbols are the same as in equation 1.1. The FR gives the rate at which a single predator consumes prey i.e. the per-capita predation rate. Numerous functional forms have been proposed for the FR, and as we discussed the correct form is the subject of a long running debate [REFS]. However there are some widely used forms for both the FR and the intrinsic growth function  $G_i$ , which we adopt. We use the following intrinsic growth functions:

$$(1.3) \quad G_0(x_0) = r_0x_0 \left(1 - \frac{x_0}{K_c}\right)$$

$$(1.4) \quad G_1(x_1) = -r_1x_1,$$

where  $r_i \in \mathbf{R}^+$  is the intrinsic growth rate of species  $i$  and  $K_c$  is the carrying capacity of the prey species. Therefore the predator species has an exponential intrinsic mortality, as in the Lotka-Volterra equations [REFS], whereas the prey species has a saturating growth function, as in the Rosenzweig-MacArthur model [REF]. For the FR we focus mainly on those proposed by Holling in the 1950s [REFs], which have been most widely used [REFS]. However other proposed forms of the FR are important in the debate and these are discussed in section 1.6. There are three Holling FRs, types I, II, and III which can be expressed as:

$$(1.5) \quad H_I(x_0, x_1) = x_0,$$

$$(1.6) \quad H_{II}(x_0, x_1) = \frac{x_0}{x_0 + K_s},$$

$$(1.7) \quad H_{III}(x_0, x_1) = \frac{x_0^2}{x_0^2 + K_s^2},$$

where  $K_s$  is the saturation constant for the predator, giving the prey density at which the per-predator consumption rate reaches half-maximum<sup>2</sup>. MORE ON FRS: PCITURES (OR IN SECTIONS ON IS??)

---

<sup>2</sup>This quantity can be related to other biologically meaningful values: search time and handling time.

**Additive noise term** Discuss what it is, why and deterministic limit. Process error [REF].

We include additive noise  $\xi_i$  in some simulations (see section ??). Therefore the (deterministic) population dynamics of a species is entirely determined by its intrinsic growth and interaction functions. SORT THIS OUT.

## 1.2.2 Simulation procedure

## 1.2.3 Interaction strength

There are numerous metrics available that can be used to quantify the strength of interactions between species. These other metrics are discussed in chapter ?? and in more detail in [REFS]. Of these metrics there is one that is a natural choice because we are able to make a direct comparison between the interaction strengths estimated from the population dynamics (using the method in section 1.2.4), and those evaluated from the simulation model (equations 1.1). This metric is called the *interaction matrix* (IM). The elements of the IM,  $\alpha_{ij}$ , quantify the effect of a small change in the population density of species  $j$  on the per capita growth rate of species  $i$ . Therefore the IM elements are given by:

$$(1.8) \quad \alpha_{ij} = \frac{\partial}{\partial x_j} \left( \frac{1}{x_i} \frac{dx_i}{dt} \right),$$

where  $x_i$  and  $x_j$  are the population densities of species  $i$  and  $j$  respectively. In the case of our two species systems the IM is a  $2 \times 2$  matrix, but trivially extends to quantify all pair-wise interactions between species in a  $N$ -species system. Since our simulation model has an explicit form for  $\frac{dx_0}{dt}$  and  $\frac{dx_1}{dt}$  we can evaluate the partial derivative in equation 1.8 to obtain analytic forms for all the IM elements ( $\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11}$ ). That is, we can calculate the interaction strengths exactly from our simulation model.

However, as stated, our goal is quantify the interaction strengths using only the population dynamics. In order to do this we fit an ODE model to the

## 1.2.4 Inference method

## 1.2.5 Examples

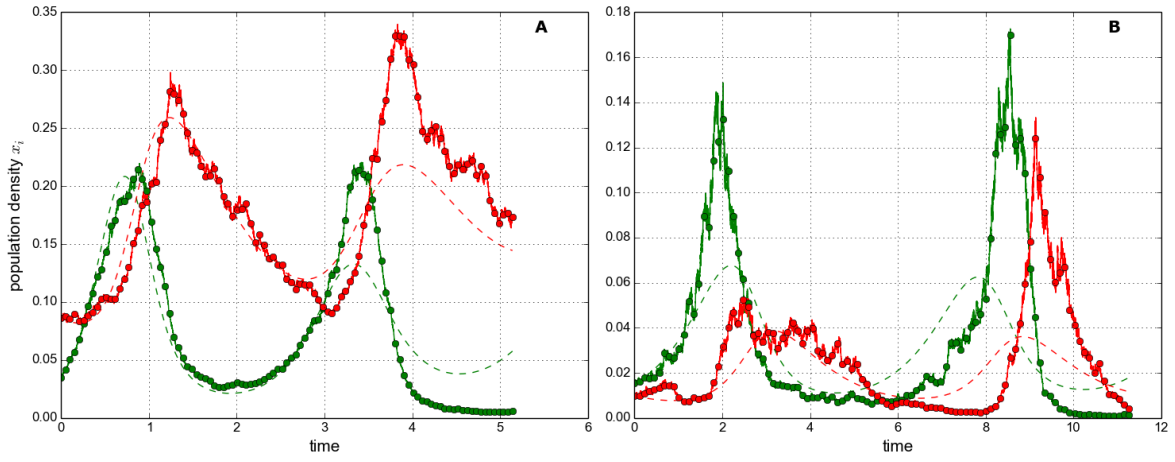


Figure 1.1: Example linear dynamics. 100 sampels. Two different parameter sets. A: noise=20. B:noise=50.

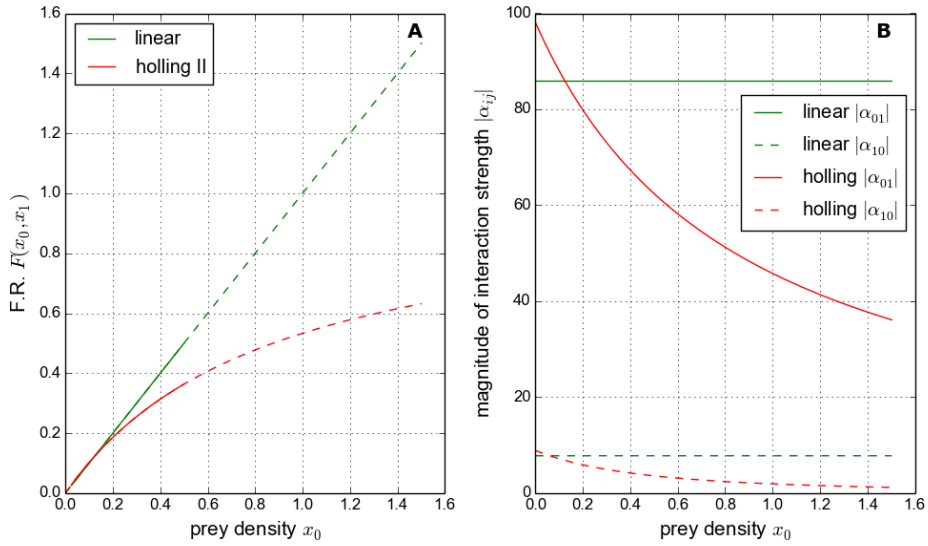


Figure 1.2: Example FR. Parameters?

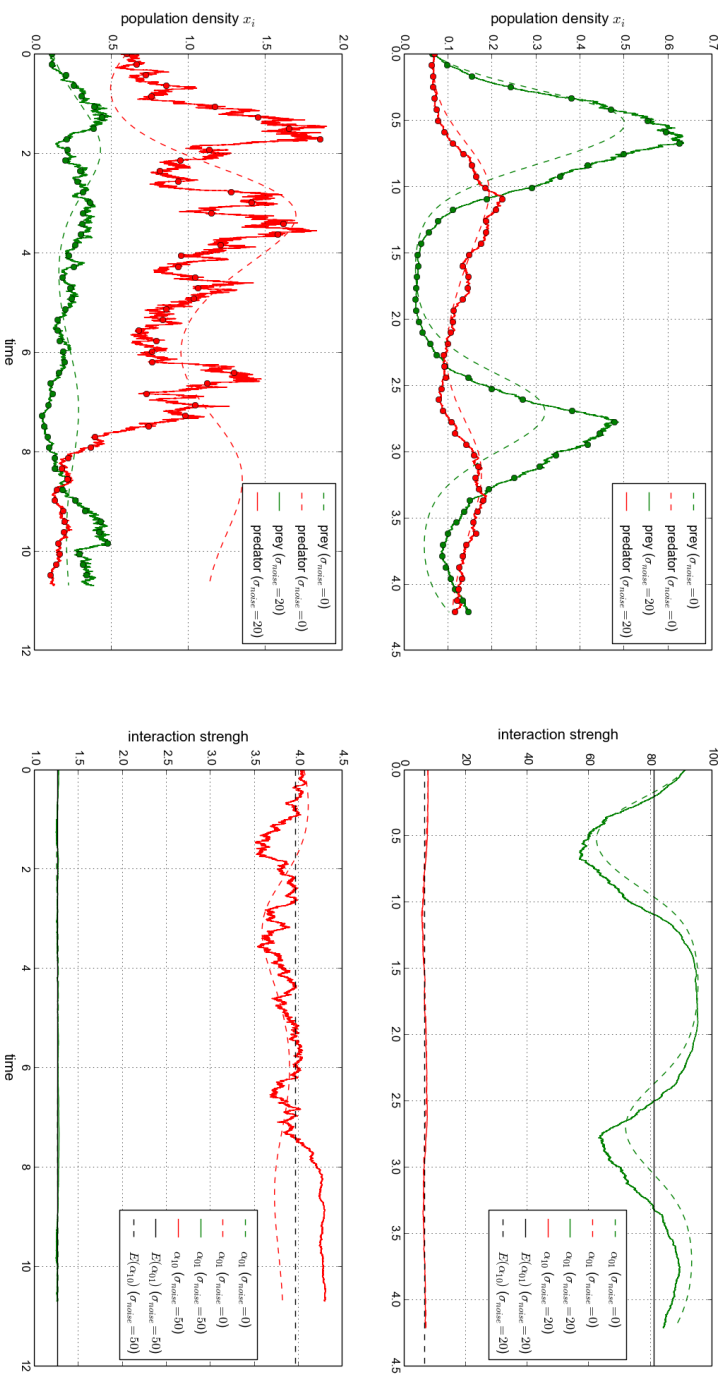


Figure 1.3: Example HOLLING II dynamics.



## 1.3 Results

In this sections we characterise the numerical performance of the method, described in section 1.2.4, for estimating the strength of interactions between species. The method is tested on the dynamics of two different ODE systems: a Lotka-Volterra (LV) and a Holling type II (HII) system. In the first case it is simply a test of a model fitting procedure. This is because the method works by fitting a generalised Lotka-Volterra (GLV) model to the dynamics, and he LV systems can be expressed as a GLV systems. Therefore we are simply simulating using one model, and then testing a method of estimating the model parameters form the simulated dynamics. We test the effects of noise and sampling frequency. In the second case, the HII system cannot be expressed as a GLV system. Therefore the GLV model that we fit can only approximate the dynamics and we cannot make a direct comparison between the parameters of the simulation model and the GLV model used for estimation. In this case we comapre to the mean interaction strengths (see section 1.3.2).

### 1.3.1 Lotka-Volterra system

Initially we run repeated simulations of the LV model using a single parameter set. We investigate how the numerical estimates of the model parameters respond to two variables: the level of noise in the simulations; and the number of samples used for estimation. Other variables are held constant using the simulation procedure described in section 1.2.2. We then generalise these results by looking at the relative error in the estimates, for repeated simulations using an ensemble of 100 selected parameter sets (as described in section 1.2.2).

**Single parameter set.** Here we can make direct comparison between model parameters. The GLV model for two species has six parameters:  $r_0, r_1, J_{00}, J_{01}, J_{10}, J_{11}$ . These correspond respectively to the following constant values of the LV system used for simulations:  $A, -1, -A, -B, C, 0$  (see equation ??). In general we find that the numerical estimates perform well at low noise intensities and poorly at high noise intensities. This is illustrated in figures 1.4 and 1.5. We also find that the estimates improve with the number of samples used, up to a point. Beyond this point the use of more samples does little to improve to estimates, and in some cases makes them worse. This behaviour is illustrated in figures 1.6 and 1.7. These patterns were found to hold across all parameter sets investiagted, but are only shown using a single parameter set here for clarity.

In panel **A** of figures 1.4 and 1.5 we see that the mean value of the estimates approaches the true value for low noise and, in panel **B** that the variance in the estimates approaches zero. This tells us that the method consistently gives a good fit of the GLV to the dynamics of the LV system, even when only 100 sample points are used (figure 1.4). As the noise intesitiy is increased the mean values of the estimates deviate from the true values, and the standard deviation in the estimates increases. Comparing the two figures we see that the response to noise is very similar

whether 100 or 10,000 samples are used. A notable exception to this is a spike in the variance in panel **B** of figure. However this appears to be a single statistically anomolous result and not part of the trend. Panel **C** of both figures shows that the error function, which is minimised by the estimation method, increases with noise for both species. This cannot be directly compared between the two plots because of the different number of samples used. However it indicates that in both cases (100 and 10,000 samples) the quality of the fit is high in the deterministic case, and decreases with noise.

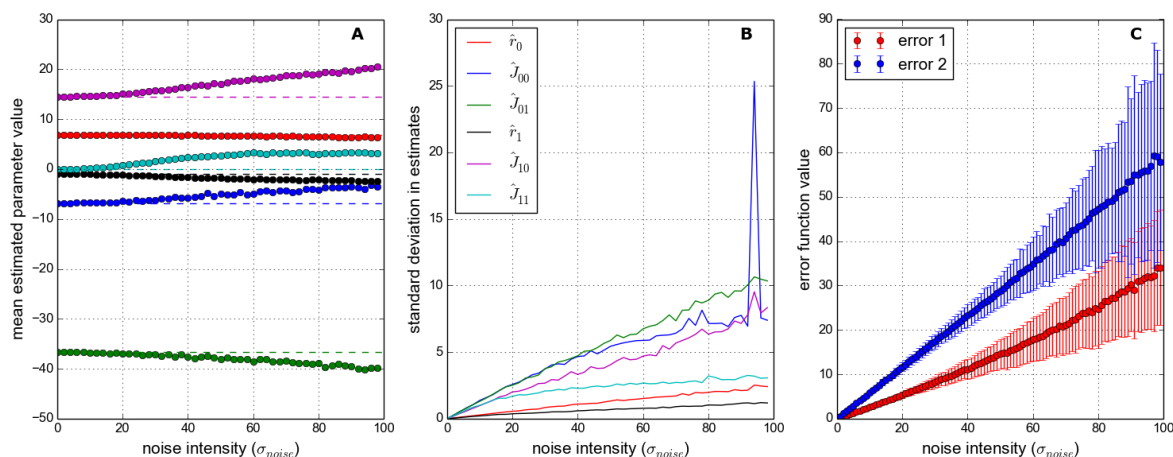


Figure 1.4: Effect of noise on numerical estimates. Here the method uses 100 samples from simulated dynamics. All simulations run using the LV model with a single parameter set. The noise intensity varies between 0 (deterministic) and 100. See section 1.2.5 for an intuition of how noisy this is. 1000 repeat simulations run at each noise level. **Panel A:** Mean estimated parameter values (each dot representing mean over 1000 repeats). The ‘true’ parameter values of the simulation model are shown by dashed lines. **Panel B:** Standard deviation in estimates. **Panel C:** Value of the error functions used in the estimation method, one for each species. The dots show the mean error, and the bars show  $\pm$  one standard deviation.

We now look at how the estimates respond to the number of samples used, in the cases of low and high noise intensity. Figure 1.6 shows the low noise case, with  $\sigma_{noise} = 10$ . In panel **A** we see that the mean value of the estimates quickly converges to close to the true parameter values, as the number of samples increases. Panel **B** shows that the standard deviation in the estimates quickly becomes small, but non-zero. Above about 32 samples there is no visible improvement in the estimates, as measured by the mean or standard deviation. In figure 1.7 we see the effect of a higher noise intensity. Here we have  $\sigma_{noise} = 50$ . Panel **A** shows that the estimates do not converge on the true parameter values, even for large numbers of samples. Also the standard deviation in the estimates, shown in panel **B**, is higher than in the low noise case. Again we find that there is little, if any, improvement in the estimates beyond about 32 samples.

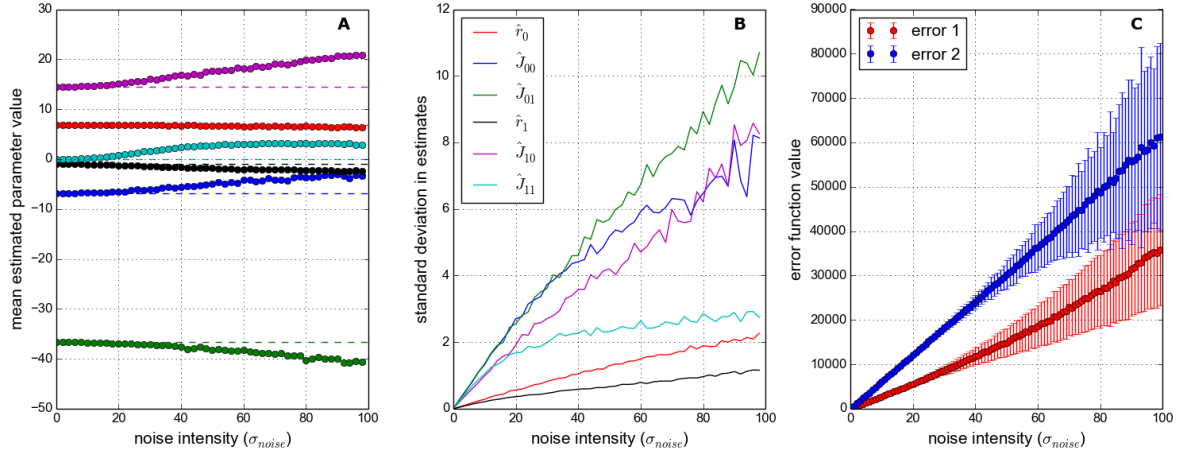


Figure 1.5: Exactly as in figure 1.4 but using 10,000 samples from the simulated dynamics.

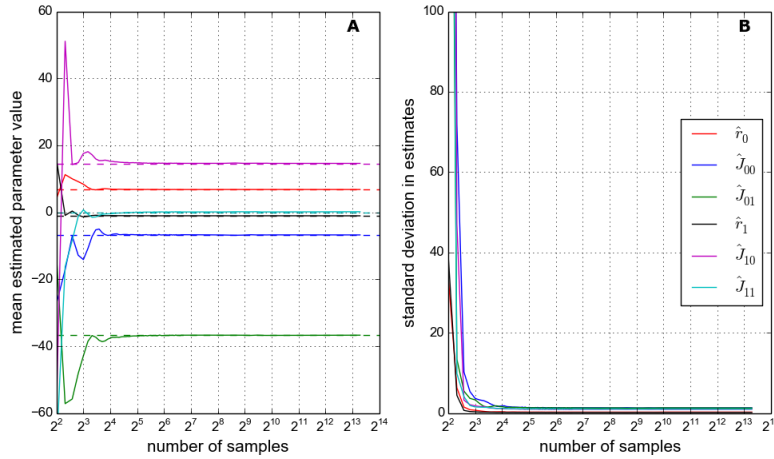


Figure 1.6: Effect of the number of samples on numerical estimates. All simulations run using the LV model with a single parameter set. The noise intensity  $\sigma_{noise} = 10$ . Number of samples ranges from 4 to 10,000. Samples drawn from simulated dynamics at equal intervals. 1000 repeat simulations for each number of samples. **Panel A:** Solid lines show mean estimated parameter values. Dashed lines show the ‘true’ parameter values of the simulation model. **Panel B:** Standard deviation in estimates.

**Ensemble of parameter sets.** Run 10 repeats for each of 100 parameter sets. In general the trends described above hold across the ensemble..

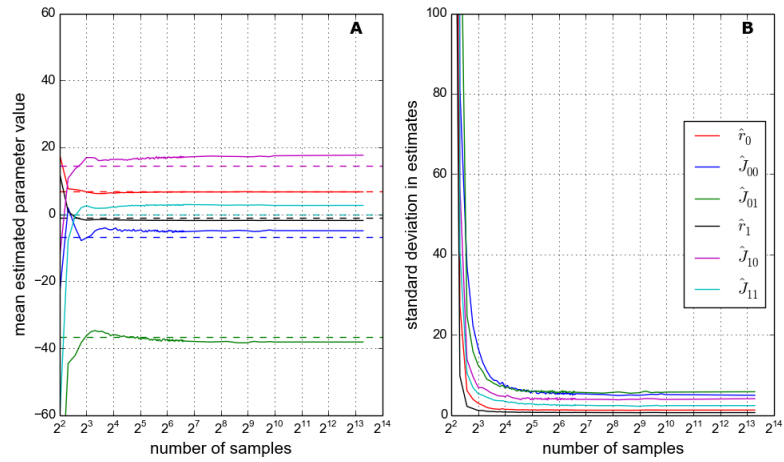


Figure 1.7: Exactly as in figure 1.6 but with noise intensity  $\sigma_{noise} = 50$ .

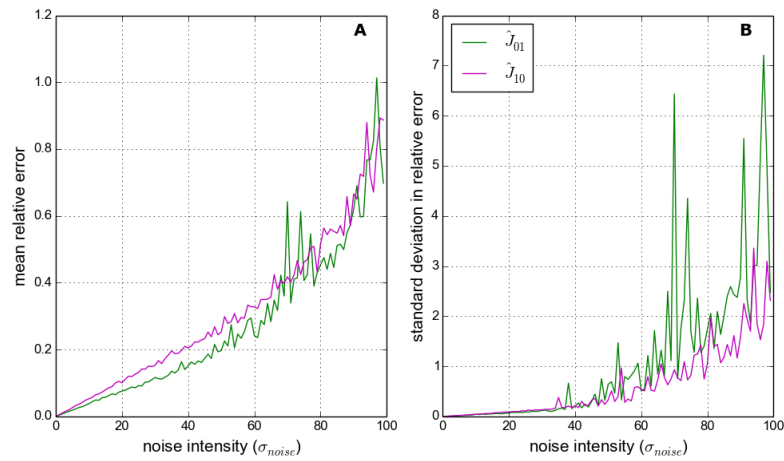


Figure 1.8: Nonsense. 1000 samples used.

### 1.3.2 Holling system

### 1.3.3 Range sampling

## 1.4 TEMP : other results

This section shows some plots which I was not planning to put into the thesis but are worth discussing..

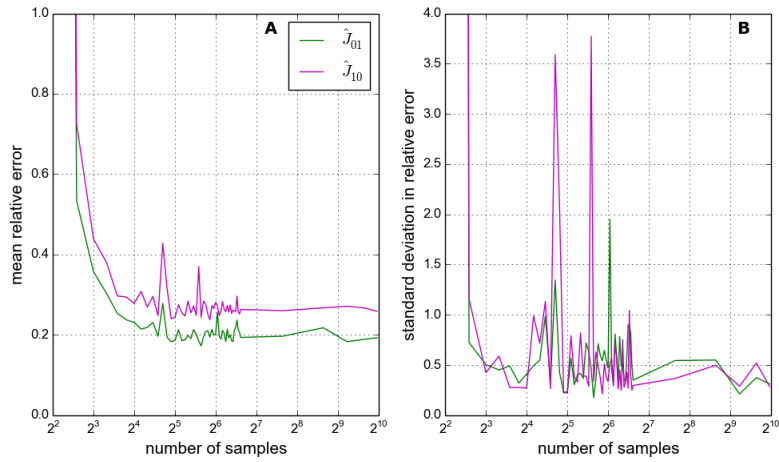


Figure 1.9: Nonsense. Noise is 50.

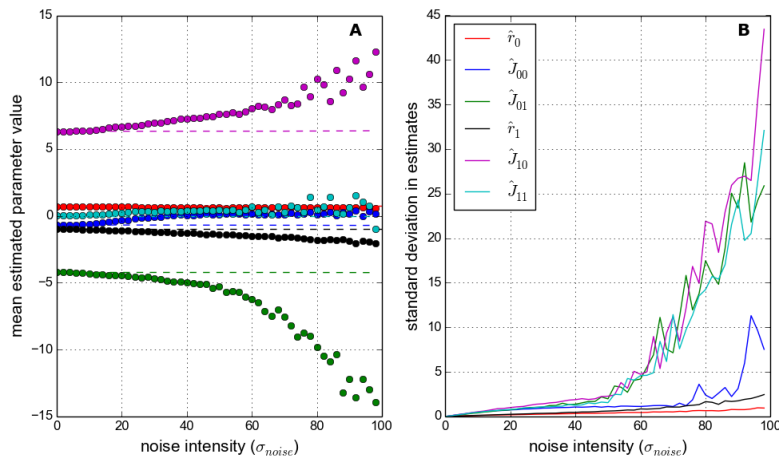


Figure 1.10: Nsamples is 10,000.

## 1.5 Application to IBM (optional)

### 1.5.1 Testing functionl response

### 1.5.2 Extend methodology (3 and 4 species)

### 1.5.3 Results

## 1.6 Discussion

Points referenced in text above, make sure to discuss them!..

- Discuss how this methodology could be used on empircal data...

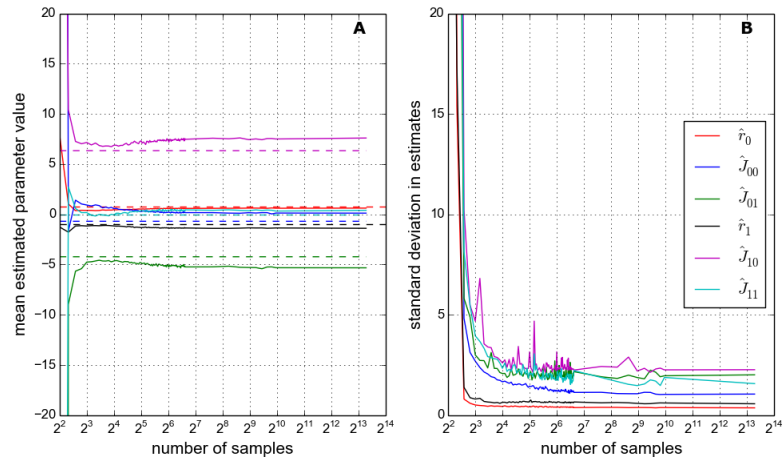


Figure 1.11: Noise is 50.

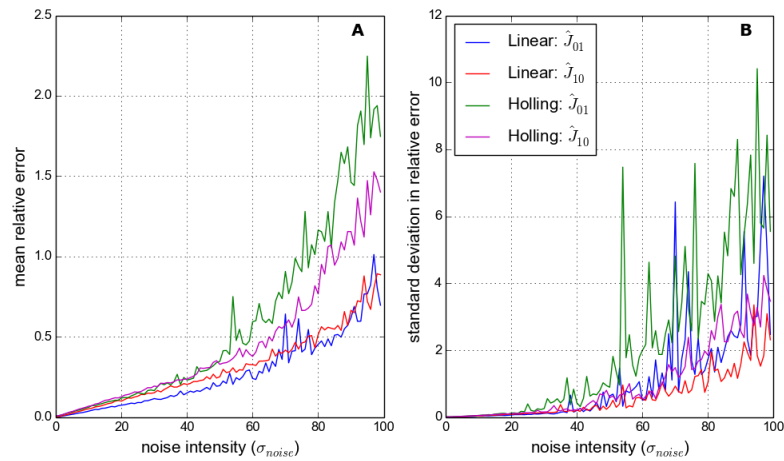


Figure 1.12: Noise is 50.

- Limitations of ODE models (non-spatial, response to debate on FR)
- Possibility of extending to more than two species systems (if this is actually done then change ref in text above)
- discussion of other forms of FR (not H) - or discussed already in intro?
- good GLV fit to LV even with 100 sample points - realistic?

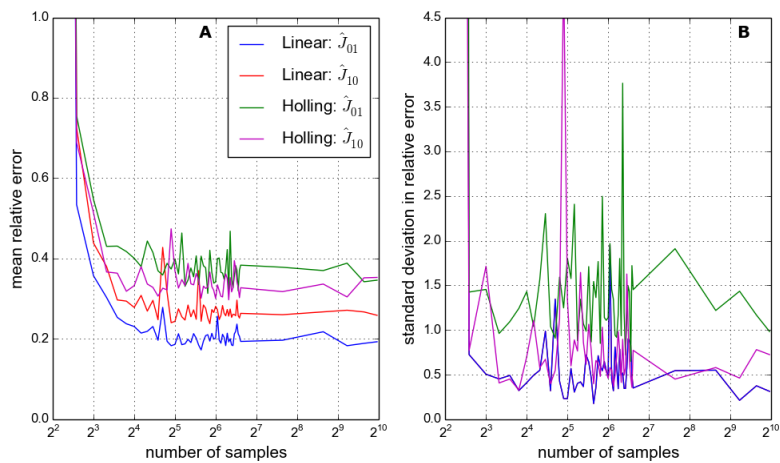


Figure 1.13: Noise is 50.

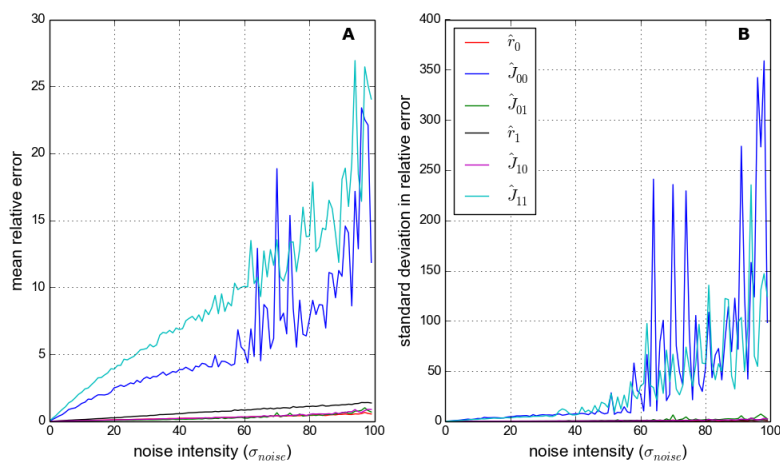


Figure 1.14: Nonsense. 1000 samples used.

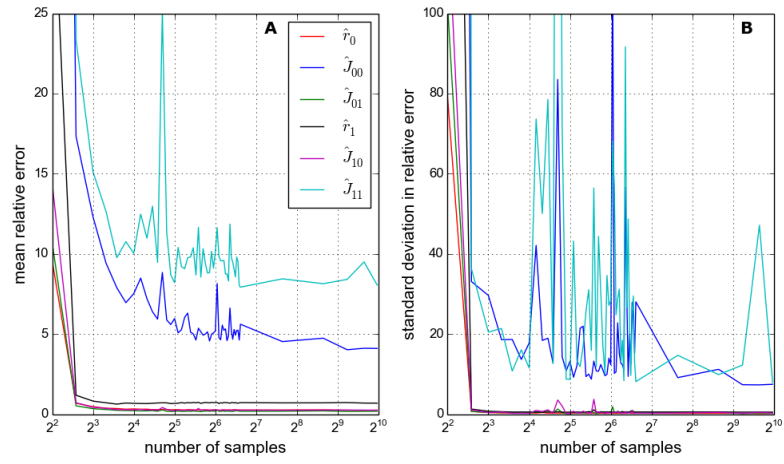


Figure 1.15: Nonsense. Noise is 50.



## BIBLIOGRAPHY

- [1] S. G. SHANDILYA AND M. TIMME, *Inferring network topology from complex dynamics*, New Journal of Physics, 13 (2011), p. 013004.

