Coefficient Recovery Using Compressive Sensing on Lotka-Volterra Models of Ecosystem Dynamics

BIOL490, Fall 2020

Author: Amanda Gu[1]

Supervisors: Eric Pedersen[2] and Simone Brugiapaglia[3]

Committee Member: Pedro Peres-Neto[4]

[1] McGill University

[2] Concordia University

[3] Concordia University

[4] Concordia University

Received: 18 Dec. 2020

Abstract The ability to create quantitative ecological models which closely follow empirical data is powerful. Models inform many management and conservation policies, and also provide a theoretical framework in which simplified cause-and-effect relationships illuminate key characteristics and interactions[1]. Accordingly, ecologists have devoted significant time to crafting models that give realistic dynamics, and these models require the input of adjustable parameters set using empirical observations. The challenge lies in how these parameters can be estimated. Unfortunately, barriers unique to marine ecosystems result in observational undersampling[2], and traditional methods for solving multivariate auto-regressive models struggle when faced with short replicates of noisy data[3]. The method proposed herein uses compressive sensing, a ℓ_1 minimization technique for recovering sparse linear relationships, to estimate the coefficients of a Lotka-Volterra system[4]. This method is then tested by generating realistic Lotka-Volterra systems, producing short replicates, and using these replicates to estimate the coefficients of the initial system.

1 Introduction

A foundational concern of community ecology is predicting how population abundances within a community will change over time. Within the context of fisheries, there is often a subset of focal or managed species of particular interest to us and we wish to understand how the population will behave in given conditions[2]. Early fisheries models

operated under the assumption that the dynamics of focal species could be predicted based on historical abundances and removal rates. Unfortunately, the task is not so simple. It is now commonly accepted that accurate predictions also require, at minimum, an understanding of the network of underlying trophic relationships, which we refer to as a "food web"[5][6][7].

Causality is a difficult relation to sift out of data, especially in the case of multivariate auto-regressive models. Even in coupled equation systems with chaotic behaviour, false or "mirage" correlations are common[8]. Variables that are correlated across the data series might not be in actuality, and this problem is exacerbated by the nature of data that is available: often only short replicates[1]. Moreover, traditional methods for solving for the coefficients of such systems require large amounts of data; like the ten or more sample points recommended for each parameter estimated for Least Squares[9]. As mathematicians further refine the techniques and lower bounds on the required amount of data to reconstruct systems, the question for ecologists becomes how much data is necessary to estimate dynamics, and whether time-series data is now, by itself, sufficient to estimate trophic links and coefficients of interactions. The central goal of this project is to investigate how variations in other populations affect the growth rate of our focal species and the amount of data required to come to these conclusions.

The magnitude and nature of these effects can be alternatively conceptualized as appropriate coefficients for the Lotka-Volterra equations. This model is one of the most widely used and earliest explored dynamical systems modelling the size of predator and prey biomass and it can be viewed as an approximation of the true dynamics near equilibrium. Developed independently by Lotka[10] and Volterra[11], the model is relatively straight-forward and can be expressed as a system of first-order quadratic ordinary differential equations (ODEs). Its coefficients must be set based on data that is often noisy and underdetermined. However, it is reasonable to assume that the coefficients are sparse (ie. they contain many zeros), as a given species likely does not have trophic links to the majority of other species present in the system. On the contrary, by dynamical systems theory, a necessary condition to having any sort of stable system is to have a certain degree of sparseness. Systems that are too interconnected are far less stable than their sparse counterparts[12].

The recovery of sparse systems is a widely studied topic in information theory, with a number of methods available to choose from[13]. Compressive sensing is but one potential method for fitting such models; it is used to reconstruct sparse signals from a limited number of linear measurements by means of convex optimization techniques[4][14].

2 Methodology

2.1 Generating Realistic Lotka-Volterra Time Series

Lotka-Volterra equations can take the form

$$\dot{\mathbf{x}} = \text{diag}(\mathbf{x})(\mathbf{r} + A\mathbf{x}),$$

or

$$\frac{dx_i}{dt} = r_i x_i + \sum_{j=1}^{n} a_{ij} x_i x_j, \qquad i = 1, \ldots, n \tag{1}$$

where $\mathbf{r}$ is a vector of length n representing intrinsic birth/death rates, $\mathbf{x}$ is the state of the system and A is the $n \times n$ community matrix. The element a_{ij} of A is the effect that species j has on species i. $diag(\mathbf{x})$ is the $n \times n$ matrix with elements of the vector $\mathbf{x}$ along its diagonal and all other elements being 0. Equation (1) is the ODE that defines the growth rate of one particular species, x_i, where n is the total number of species in the system.

Within this framework, a narrower objective statement arises: given a time series of $\mathbf{x}$ and $\dot{\mathbf{x}}$ values, can we find A and $\mathbf{r}$ that accurately predicts the values of $\mathbf{x}$ and $\dot{\mathbf{x}}$ in the future? How much data do we need to achieve this?

Before testing the efficacy of compressive sensing at recovering realistic Lotka-Volterra dynamics, we must first have realistic dynamics to recover. The Lotka-Volterra model does not include any guidance on the choosing of trophic links, and so the question we are faced with is given the number of species, n, and the desired connectance level or sparsity, S, what mechanisms can be used to create dynamics that would resemble realistic food web dynamics? We define S to be the proportion of existing links out of the total possible links in the ecosystem; that is $S = \frac{\text{non-zero links}}{n^2}$.

The Lotka-Volterra model also does not specify what the value of those coefficients should be. We used metabolic-based scaling factors from Yodzi and Innes'[15] to set intrinsic growth rates and interaction coefficients, and generated predator-prey interaction patterns via the Niche Model of food web interactions [16]. The model is based on the very simple principle that larger organisms eat smaller organisms and was chosen for its ability to mimic key structural features of food webs. For example, the percentage of species with no predators and food chain length match empirical observations much better than randomly assembled unstructured food webs. The niche model begins by assigning each

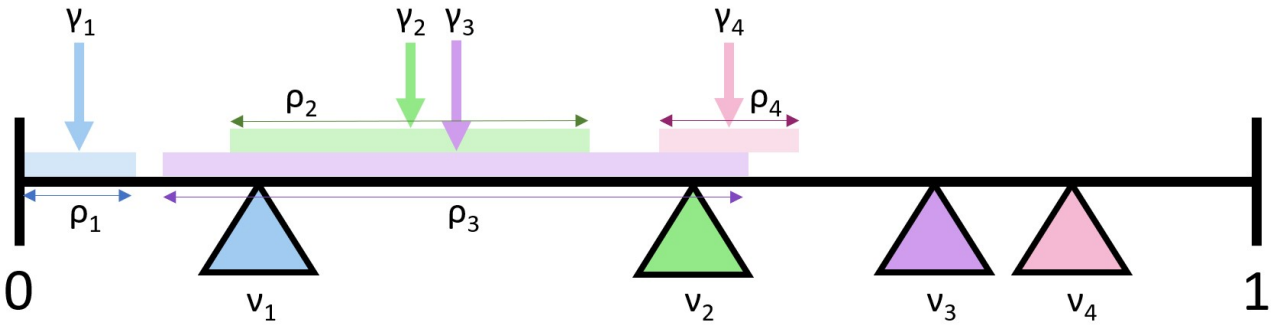

Figure 1. Adapted Williams and Martinez, 2000 [16], a visual representation of how the niche model generates trophic links. The triangles represent the position of each species with respect to their ν_i, and the coloured bars, ie. the ρ_i-length interval centred at γ_i, represent the niche ranges of the prey.

species a position in a hierarchy with a niche value, ν_i, drawn from a uniform distribution from 0 to 1. This niche value is a representation of the log-body size and ties nicely into the allometric coefficients that are used to determine Lotka-Volterra coefficients. Following this, the model assigns a range in $[0, 1]$ to each species and the species will

consume all species with niche values within this range. The width of the range, ρ_i, is drawn from a beta-distribution with the parameters $\alpha = 1$, $\beta = \frac{1}{2S} - 1$. The centre, γ_i, is then uniformly drawn from the interval $[\rho_i/2, \nu_i]$[16].

Species i then consumes all species with a niche value in $[\gamma_i - \rho_i/2, \gamma_i + \rho_i/2]$[16].

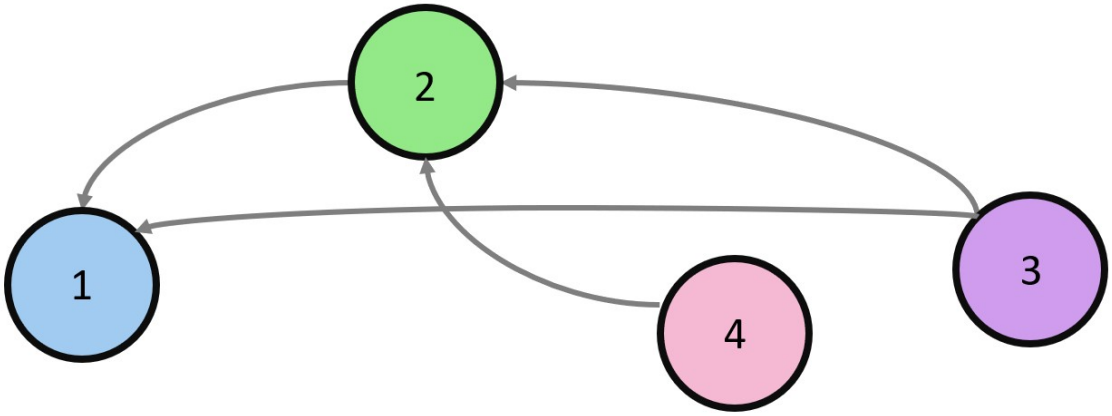

Figure 2. The food web generated based on the niche values, radii, and centre points in Figure 1
.

From the niche model, we obtain a food web with the appropriate size and connectance in the form of a matrix, where $a_{ij} = 0$ if j does not eat i and and $a_{ij} = 1$ if j eats i. This represents the trophic links but does not capture the magnitude of these relationships. In order to convert our food web into a system of ODEs, we look to a Type I, Taylor series approximation of the Holling Type II response Lotka-Volterra system defined by Yodzis and Innes[15], outlined in Equation (2). The basic premise of the coefficients which realistically mimic predator-prey dynamics are parameters which represent growth rate, mortality, predation rate, and conversion efficiencies. These parameters, are in turn determined by metabolic type (eg. herbivorous vertebrate ectotherm) and scale allometrically. That is to say, these rates scale logarithmically with body size.

$$\begin{aligned} \frac{dR}{dt} &= R\left(1 - \frac{R}{K}\right) - \left(\frac{vw}{1-\delta} f_2\right) P \frac{R}{R_0} \\ \frac{dP}{dt} &= Pv\left(-1 + w\frac{R}{R_0}\right) \end{aligned} \tag{2}$$

where

$$v = \frac{a_T}{f_r a_r}\left(\frac{m_P}{m_C}\right)^{0.25}, \qquad w = \frac{f_J a_J}{a_T}$$

P is the biomass of the predator species, and R is the biomass of the resources species.

Using the empirically determined rates and coefficients in Table 1, plug in the values into Equation (2) to obtain dynamics.

Coefficient	Subset of Species	Description	Value
δ	Carnivore	% energy lost during ingestion and metabolization	0.15
δ	Herbivore	% energy lost during ingestion and metabolization	0.55
a_T	Vertebrate ectotherm	Baseline respiration rate (Energy expenditure in free existence)	2.3
a_J	Vertebrate ectotherm	Maximal rate of energy ingestion	8.9
a_r	Vertebrate ectotherm	Physiological capacity for production	6.6
a_r	Phytoplankton (producer)	Physiological capacity for production	0.4
f_r	Vertebrate ectotherm	Production biomass ratio	1
f_r	Phytoplankton (producer)	Production biomass ratio	1
f_e	Prey	Fraction eaten	1
R_0	Prey	Half-saturation density of prey	0.5
m_R	Prey	Mass of prey	ν_R
m_P	Consumer	Mass of predator	ν_P

Table 1: Empirical rates and universal coefficients which represent the efficacy and probability of natural processes for various metabolic organism types.

Before constructing our ecosystem, we must first extend the two species dynamics to a multi-species model. This can be accomplished rather intuitively using the following equations. Let P be an arbitrary non-resource, R be an arbitrary resource, and a subscript of 0 indicate the half-saturation density of that species. E.g. P_0 is the half-saturation density of the species P.

$$\begin{aligned} \frac{dR}{dt} &= R\left(1-\frac{R}{K}\right) - \sum_{Q\in\text{Predators}} \frac{QRvwf_e}{(1-\delta)R_0 n_Q} \\ \frac{dP}{dt} &= -Pv + Pvw \sum_{R'\in\text{Prey}} R' \frac{w}{R'_0 n_P} - \sum_{Q\in\text{Predators}} \frac{QPvwf_e}{(1-\delta)P_0 n_Q} \end{aligned} \tag{3}$$

Examining Equations (3) and (1), we can directly populate the entries of A and r using the parameters in Equation (3).

We assume that the biomasses of the resource species grow logistically at equal rates in the absence of predators and there is no interspecific competition between resource species. Therefore, if species i is a resource, j is its consumer, x_i and x_j are the biomass of species i and j respectively, we see can see by plugging the ODEs into our matrix form

that

$$\begin{aligned} r_i &= 1, \\ a_{ii} &= \frac{1}{K}, \\ a_{ij} &= -\frac{vw}{1-\delta} f_e \frac{1}{x_{i0}\phi_j}, \end{aligned}$$

where ϕ_j is the number of prey that species j has.

Conversely, let species i be a consumer, j its predator, and k its prey. Starvation causes the biomass of i to decrease exponentially at rate v in the absence of prey. As a function of some of the rates in table 1, the biomass increases by $\frac{v}{x_{k0}\phi_i} x_i x_k$ for each prey species it has and decreases by a factor of $vwf_e \frac{1}{(1-\delta)x_{i0}\phi_j}$. Therefore, we have

$$\begin{aligned} r_i &= -v \\ a_{ij} &= -\frac{vwf_e}{(1-\delta)x_{i0}\phi_j} \\ a_{ik} &= \frac{w}{x_{k0}\phi_i}. \end{aligned}$$

We then use the generated systems to test the process outlined in the following Section 2.2.

2.2 Recovering Hidden Dynamics With Compressive Sensing

Compressive sensing is a signal-recovery method, which means that we are attempting to recover some information, called a signal, from some data that we have. This data is obtained by somehow transforming the initial signal, in this case by multiplying the signal by a measurement matrix. Let us define the signal and matrix with respect to the Lotka-Volterra system; in this case, the signal we are trying to recover is the set of coefficients (A, r). The method used in this experiment closely follows that which was outlined in Schaeffer and Tran, 2018[17].

2.2.1 The measurement model: From unknown coefficients to time series data

In our case, the data we have are *bursts* of time series. A burst is a time series of some fixed length, and each burst has different initial conditions (generated from the exponential distribution). In ecology, they are often referred to as replicates or short community time series. We are assuming that we have access to both growth rate and biomass of our focal set of species at each time, though this is a non-trivial assumption as rates of change cannot be easily measured empirically and are not typically available for real data. Nevertheless, it is possible to, from state measurements, approximate them with smoothing functions[18]. Writing all of this information in matrix form, we can generalize a vector of growth rates, $\dot{\mathbf{x}}$ in our initial system, to the following matrix of velocity for the kth burst out of K total

bursts.

$$V^{(k)} = \begin{pmatrix} \dot{x_1}(k,t_0) & \dot{x_2}(k,t_0) & \cdots & \dot{x_n}(k,t_0) \\ \dot{x_1}(k,t_1) & \dot{x_2}(k,t_1) & \cdots & \dot{x_n}(k,t_1) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{x_1}(k,t_m) & \dot{x_2}(k,t_m) & \cdots & \dot{x_n}(k,t_m) \end{pmatrix} \tag{4}$$

Here, $\dot{x_i}(k,t_j)$ is the rate of change of species x_i at time t_j. Moreover, let $x_i(k,t_j)$ be the biomass of species x_i at time t_j.

We can see in Equation (1) that each $\dot{x_i}(k,t_j)$ term can be written out as a function of x_i, and x_ix_j for all j from 1 to n (where n is number of species). x_i and x_ix_j are then what we refer to as the standard polynomial basis. Taking this basis, consider that the community matrix can be decomposed into the following matrix product of a coefficient matrix and a time-dependent basis matrix. In this case, we may apply Equation (1) to the entirety of the velocity matrix, $V^{(k)}$, note that each column, which represents the velocity of one species over time, is a function of x_i and x_ix_j at the given time value.

Writing this relationship between $V^{(k)}$ and the state measurements as a matrix-vector product, we have

$$V^{(k)} = \begin{pmatrix} x_1(k,t_0) & x_2(k,t_0) & \cdots & x_n(k,t_0) & x_1(k,t_0)^2 & x_1(k,t_0)x_2(k,t_0) & \cdots & x_n(k,t_0)^2 \\ x_1(k,t_1) & x_2(k,t_1) & \cdots & x_n(k,t_1) & x_1(k,t_1)^2 & x_1(k,t_1)x_2(k,t_1) & \cdots & x_n(k,t_1)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1(k,t_m) & x_2(k,t_m) & \cdots & x_n(k,t_m) & x_1(k,t_m)^2 & x_1(k,t_m)x_2(k,t_m) & \cdots & x_n(k,t_m)^2 \end{pmatrix} C,$$

where C is a matrix of coefficients of dimension $(n^2+n) \times n$. Each column of C corresponds to the coefficients of the ODE for the ith species.

This can be streamlined even further by noting that for each $\frac{dx_i}{dt}$, there are certain basis vectors that must have a coefficient of 0. By Equation (1), we can define a basis matrix for each species. Define

$$B_i^{(k)} = \begin{pmatrix} x_i(k,t_0) & x_1(k,t_0)x_i(k,t_0) & x_2(k,t_0)x_i(k,t_0) & \cdots & x_n(k,t_0)x_i(k,t_0) \\ x_i(k,t_1) & x_1(k,t_1)x_i(k,t_1) & x_2(k,t_1)x_i(k,t_1) & \cdots & x_n(k,t_1)x_i(k,t_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_i(k,t_m) & x_1(k,t_m)x_i(k,t_m) & x_2(k,t_m)x_i(k,t_m) & \cdots & x_n(k,t_m)x_i(k,t_m) \end{pmatrix}.$$

We now have that the ith column of $V^{(k)}$, which represents the rates of change of one species over the whole burst, is a product of this new basis matrix and the ith column of a new coefficient matrix, C', which has the terms that must be zero removed. This new sensing matrix also makes the relationship between A, r, $V^{(k)}$, and C' more

straightforward and explicit.

$$V^{(k)}(\ ,i) = B_i^{(k)} C'(\ ,i) \tag{5}$$

$$C' = \begin{pmatrix} r_1 & r_2 & \cdots & r_n \\ a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} \tag{6}$$

Here, $(\ ,i)$ represents the ith column of a matrix. Notice that because of how we ordered the basis elements in $B_i^{(k)}$, we have that $C'(1,i)$ is the value of r_i and $C'(j+1,i)$ is the value of a_{ij}.

By the nature of matrix multiplication, if there are K bursts, we can "stack" the matrices in Equation (5) as follows.

$$\begin{pmatrix} V^{(1)}(\ ,i) \\ V^{(2)}(\ ,i) \\ \vdots \\ V^{(K)}(\ ,i) \end{pmatrix} = \begin{pmatrix} B_i^{(1)} \\ B_i^{(2)} \\ \vdots \\ B_i^{(K)} \end{pmatrix} C'(\ ,i). \tag{7}$$

Again, to make this equation easier to denote, let us define the matrices in the above equation as

$$V(\ ,i) = B_i C'(\ ,i) \tag{8}$$

Recall that the values of V and the basis matrices, $B_i^{(k)}$, are known and we are now attempting to recover the sparse signal $C'(\ ,i)$, or the columns of the coefficient matrix C', one by one.

2.2.2 Recovery via ℓ_1 minimization: From time series data to coefficients

Using the l_1 measure, the convex optimization problem

$$\hat{C}(,i) := \underset{C'(\ ,i):\ V(\ ,i)=B_i C'(\ ,i)}{\operatorname{argmin}} (||C'(\ ,i)||_1) \tag{9}$$

can be solved to obtain C' and therefore $\mathbf{r}$ and A with a high-degree of accuracy. Here, the subscript $C'(\ ,i):\ V(\ ,i) = B_i C'(\ ,i)$ denotes the support of the possible solutions: our final guess for $C'(\ ,i)$, must satisfy this condition that, when multiplied by B_i, we obtain $V(\ ,i)$. This formula assumes no noise as this is a strict equality; further discussion on noisy data is included in Section 3. $argmin(||C'(\ ,i)||_1)$ denotes the minimum argument, or the $C'(\ ,i)$ in the support that has the minimum ℓ_1 norm, where $||\mathbf{x}||_1 = \sum_1^n |x_i|$. In other words, this equation sets our guess, $\hat{C}(\ ,i)$, to be the C_i' with the smallest ℓ_1 norm that satisfies the condition $V(\ ,i) = B_i C'(\ ,i)$.

Norms are generalizations of the notion of length, where ℓ_2 is the common Euclidean norm. The ℓ_p norm is defined and denoted as follows: for a vector $\mathbf{x} = (x_1, x_2 \cdots, x_n)$, the ℓ_p norm of $\mathbf{x}$ is $||\mathbf{x}||_p = \sqrt[p]{\sum_1^n (x_i)^p}$. Recall that the ℓ_1 norm is the one used in compressive sensing, i.e. the absolute value of the magnitude of the vector defines the penalty term. We chose the ℓ_1 norm (the sum of the absolute values of all elements) because of its geometry, and specifically its convexity, and therefore it maximizes sparsity. This is not a new idea, methods such have LASSO also use ℓ_1 minimization in conjunction with the optimization problem of minimizing squared-error lost to estimate the linear relationship between predictor and response variables[19]. To understand why this works, it is first important to understand some basic properties of norms.

Loosely speaking, the convexity of a norm measures the "pointy-ness" of the unit ball along the basis vectors. The unit ball is the subset, U of all vectors that have a norm equal to one, ie. $\forall x \in U, ||x||_1 = 1$. Graphically, it can be visualized as some shape around the origin such that all points on the perimeter of the shape are "equidistant" from the origin within the norm. Figure 1 displays the shape of unit balls for various p norms. For the familiar ℓ_2 (Euclidean norm) in $\mathbb{R}^2$, the unit ball takes on the form of a circle. In this case, $U := \{(y_1, y_2) \in \mathbb{R}_2 \text{ s.t. } ||(y_1, y_2)||_2 = \sqrt{y_1^2 + y_2^2} = 1\}$, or the unit ball U is the set of all points (y_1, y_2) such that the ℓ_2 norm of (y_1, y_2), $\sqrt{y_1^2 + y_2^2}$ is equal to 1. Clearly, this is a circle with radius 1. This result can also be generalized to obtain that for any circle centred at the origin in $\mathbb{R}^2$, all points on its perimeter will have the same ℓ_2 norm.

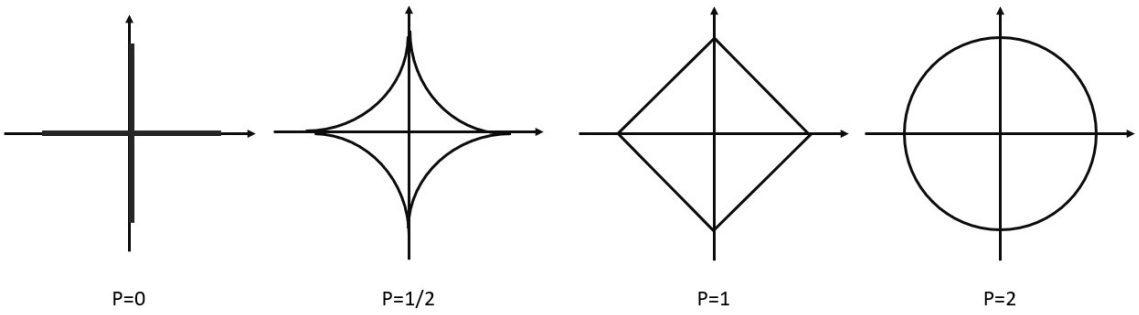

Figure 3. Unit balls in various ℓ_p norms.

As shown in Figure 3 , for the ℓ_1 norm on $\mathbb{R}^2$, the unit "ball" is actually shaped like a diamond with its points at $(0, 1)$, $(1, 0)$, $(-1, 0)$, and $(0, -1)$. Just as in the case of the Euclidean norm, where for all points x and y on a circle centred at the origin's perimeter, $||x||_2 = ||y||_2$, we have that for all points x and y on the perimeter of a diamond centred at $(0, 0)$, $||x||_1 = ||y||_1$. Recall that we are solving an underdetermined matrix problem. As such, there is a family of solutions that satisfies the condition $V(\ , i) = B_i C'(\ , i)$ in the optimization problem in Equation (9). Returning to an $\mathbb{R}^2$ case, we might obtain a line of possible solutions. When searching for the vector with the smallest norm, we obtain the vector that intersects the smallest diamond (recall that all points on a given diamond have the same norm). As the unit ball in ℓ_1 is "pointy", it is almost certain that this solution intersects the diamond at one of the points along a basis vector, meaning that the other basis vectors have a value of 0 at this particular solution.

Conversely, for a non-convex norm, like the Euclidean norm, the element with the smallest norm (i.e. the element that intersects the smallest circle) is almost certainly not sparse. However, when determining which ℓ_p norm to use, this

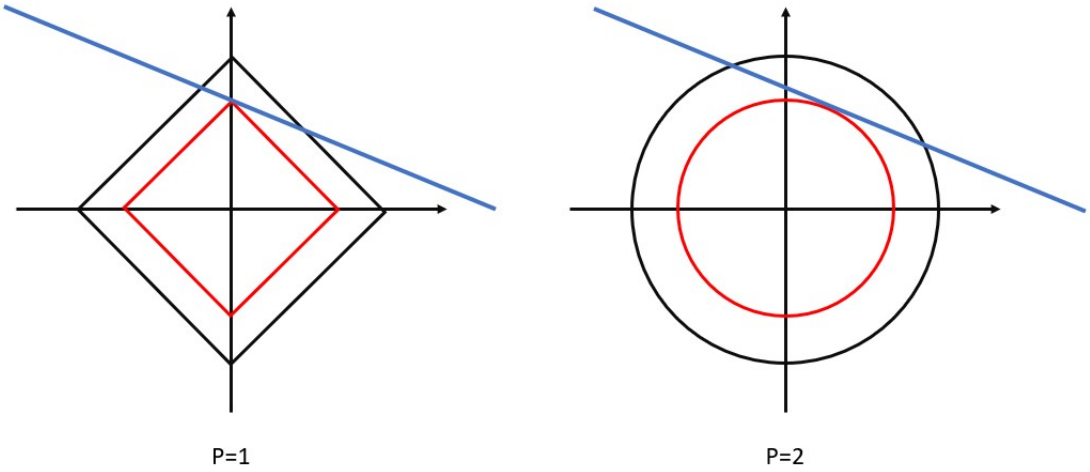

Figure 4. The same solution hyperplane (blue line) and the solution with the smallest ℓ_1 norm (point where it intersects red diamond) is sparse, while the same is not true for the solution with the smallest ℓ_2 norm.

argument does not restrict our p options to $p = 1$. Examining Figure 3, it is immediately clear that ℓ_1 is not the only convex ℓ_p norm: in fact, for all $p < 1$, the norms grow only more and more convex as $p \longrightarrow 0$, and at $p = 0$, it is certain that the *argmin* found is sparse. Unfortunately, the ℓ_0 minimization is NP hard to compute and as such, is not a very useful tool for large datasets[20].

In the case of noisy data, the conditions of the convex optimization would need to be relaxed in order to prevent overfitting and to generate the best possible solution. Instead of the strict equality presented in Equation (9), the problem

$$\hat{C}(,i) := \underset{C'(\ ,i):||V(\ ,i)-B_iC'(\ ,i)||_2\leq\sigma}{\operatorname{argmin}} (||C'(\ ,i)||_1) \tag{10}$$

enforces the condition that the Euclidean norm of the difference between $V(\ ,i)$ and $B_iC'(\ ,i)$is less than some error term σ [17].

2.3 Toy Example Parameters & Metrics

In this experiment, we use the burst length $M = 8$, representing times 0-0.7 for a food web with $n = 10$ species. The burst length was specifically chosen to be small, fewer than the number of species in the system, while being long enough to capture meaningful interactions between species. The connectance $S = 0.3$ was picked as it is on the higher end of connectances empirically observed in food webs[21]. If the method works when $S = 0.3$, it will work for food webs with a connectance of less than 0.3.

At time 0, the biomasses were picked using an exponential distribution (so that all biomasses were above 0). M initial condition vectors are picked, one for each burst. After generating the Lotka-Volterra system as described in Section 2.1, the R package deSolve[22], an ODE solver, can calculate the biomass of all species at each of the times for each burst using the system and initial conditions. We now have the state data which we attempt to recover the system with. Noise can be added at this point by generating a matrix of the same dimensions as the measurements with elements pulled from a Gaussian. Regardless of if noise is added, we construct the monomial matrices $B_i^{(k)}$, and use the R package CVXR[23], which solves the convex optimization problem presented through an ADMM-based first-order method[24]. Following this, we can clean up and de-noise the matrix by setting coefficients with an order of 10^{-5} or less to 0.

There are two main objectives we are aiming to estimate are the trophic links (ie. which coefficients are non-zero) and the magnitude of interactions (ie. the value of the coefficients themselves). The success with which compressive sensing attains these objectives can be measured in many different ways, but we have chosen to consider only the following metrics, which track the approximation quality of the recovered coefficients and dynamics:

1) Percentage of coefficients that were recovered

In Sections 3 and 4, the topic of unrecovered coefficients is covered in greater detail, but in essence, when there is not enough information,compressive sensing is unable to return any estimate. The percentage of coefficients unrecovered is

$$\frac{\text{number of coefficients unrecovered}}{n^2+n}.$$

The denominator is the total number of coefficients we are trying to recover.

2) Euclidean distance between A and $\hat{A}$

Euclidean distance is perhaps the most straightforward, and answers the question "How close is $\hat{A}$ to A?". First, bind the r and A together to form a single $n+1 \times n$ matrix. Take the Euclidean distance as follows.

$$d(A,\hat{A}) = \left(\sum_{i=1}^{n+1}\sum_{j=1}^{n} |a_{ij}-\hat{a}_{ij}|^2\right)^{\frac{1}{2}}$$

3) Difference in leading Eigenvalue of A and $\hat{A}$

Eigenvalues describe the magnitudes by which a space is stretched by a transformation represented by a matrix. The leading Eigenvalue is important, as it is the dominant stretching action performed. Understanding the system in this way allows us to understand the stability of the system.

4) Type I and II error of tropic links recovered

Type I and type II error are both measures of trophic link recovery. Type I error refers to false negatives, and type II error refers to false positives. Both types of error were calculated as a percentage of the total amount of the possible error of that type. I.e. in the case of type I error, the total number of false negatives is the number of trophic links in the initial system and the total amount of possible type II error is the number of non-active links in the initial system.

5) Average absolute error using new initial conditions

With A, r, $\hat{A}$, and $\hat{r}$, a natural metric to determine the accuracy of the recovered system is to generate new initial conditions, and solve the initial and estimated systems. Subtracting the biomass of each species in the two sequences then taking the absolute value and median of all values for a given time yields the median absolute error.

3 Results

With an empirical, a recovered systems, and a new set of initial conditions, the most direct results we obtain are two new time series by solving for both systems, thereby obtaining predicted states and the actual states of the ecosystem. One such example is shown in Figure 5, where the predicted and actual dynamics are plotted for each of the ten species. We used two bursts of data, i.e. $K = 2$ to construct our estimate.

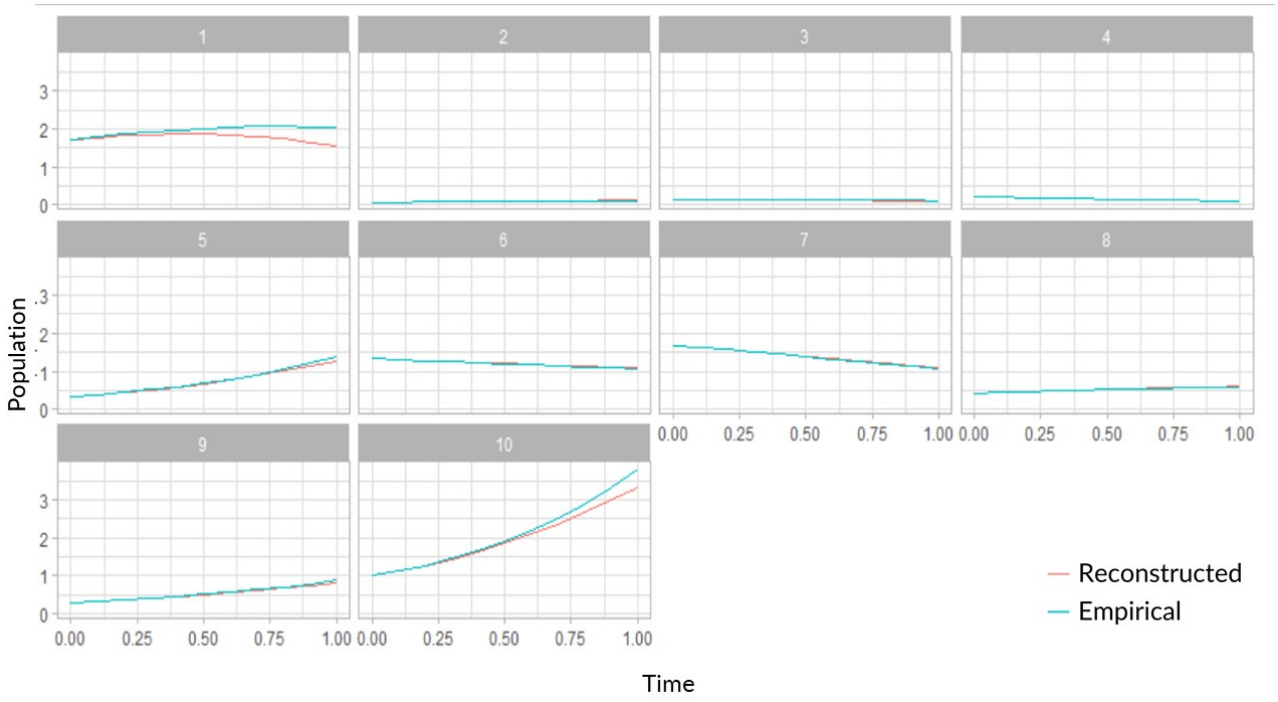

Figure 5. The biomasses of species 1 through 10 in the empirical system and reconstructed system, which was recovered with 2 bursts of data.

As can be seen in Figure 6, the phase transition between 30% of coefficients being unrecoverable and all of the coefficients being recoverable occurs rather abruptly between two and four bursts. With five bursts, compressive sensing is consistently returning an estimate for all coefficients. Recall that two bursts is only sixteen measurements and four bursts is thirty-two state measurements. This is a full order of magnitude smaller than the 110 coefficients we are attempting to recover.

Note that in Figure 6 and all following, the points have been *jittered*, that is, a small amount of random noise has been added. This prevents the points from being clustered too closely, and that lateral spread allows the individual

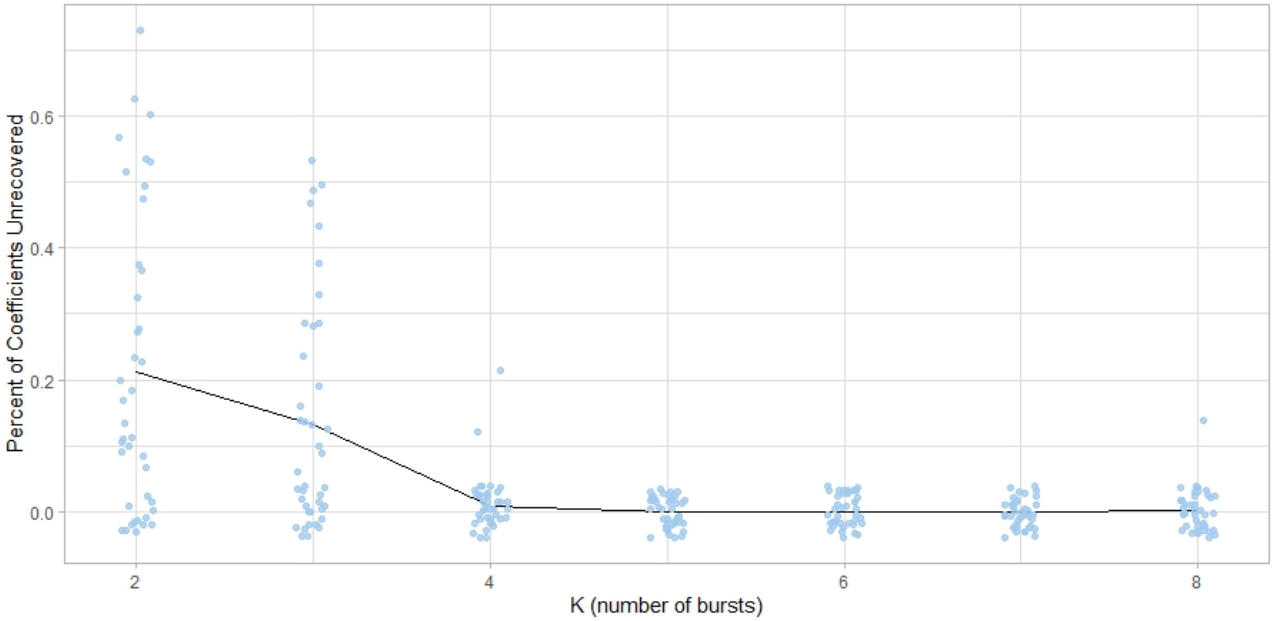

Figure 6. Jittered plot showing percentage of unrecovered coefficients as a function of burst length

points to be more visible. Additionally, the following metrics were measured only from systems which were entirely recovered, that is, all metrics are conditional upon successful recovery. For example, the trophic links in a coefficient matrix that was 80% recovered would not be included in the count of those that were recovered, nor would the trophic links in the initial matrix be included in the trophic links compressive sensing was attempting to recover.

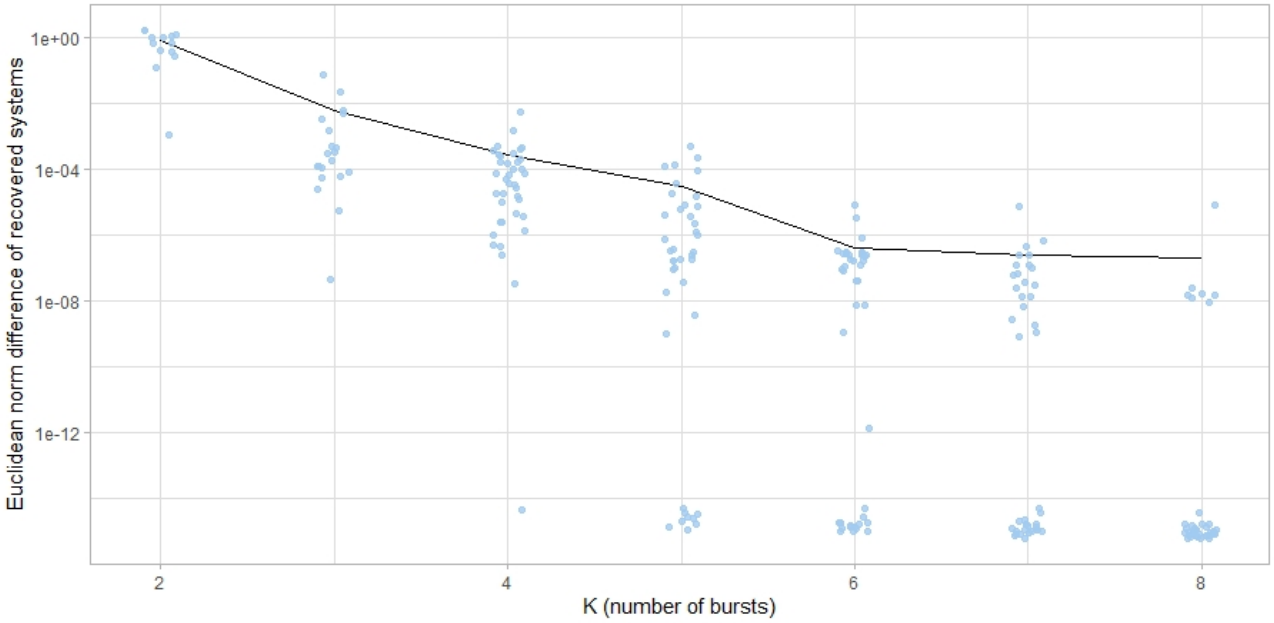

Figure 7. Jittered plot showing the Euclidean norm of the difference between A and $\hat{A}$

In Figure 7, as expected, the average Euclidean distance shrinks as the number of bursts increases. The y-axis is logarithmically scaled and the decrease is actually quite a bit more than what a quick glance reveals. At six bursts, the

error is, on average below 10^{-6}, and in any application context, is functionally zero. Bifurcation can also be observed at that point where the median drops significantly, and exact recovery to machine precision can be observed in an increasing percentage of systems as the number of bursts increases.

Unsurprisingly, Figure 8 very closely mirrors Figure 7. Since the order of magnitudes are so small and compressive sensing is essentially exactly reconstructing the matrix after five bursts, so here, a smaller norm difference is equivalent to closer leading Eigenvalues. This represents more similar behaviour by the systems around equilibrium and stability properties. In both graphs we see a significant amount of bifurcation, and so the performance of compressive sensing is actually better than what the mean might suggest. With seven bursts, this method is recovering the exact system to mach precision more than half of the time.

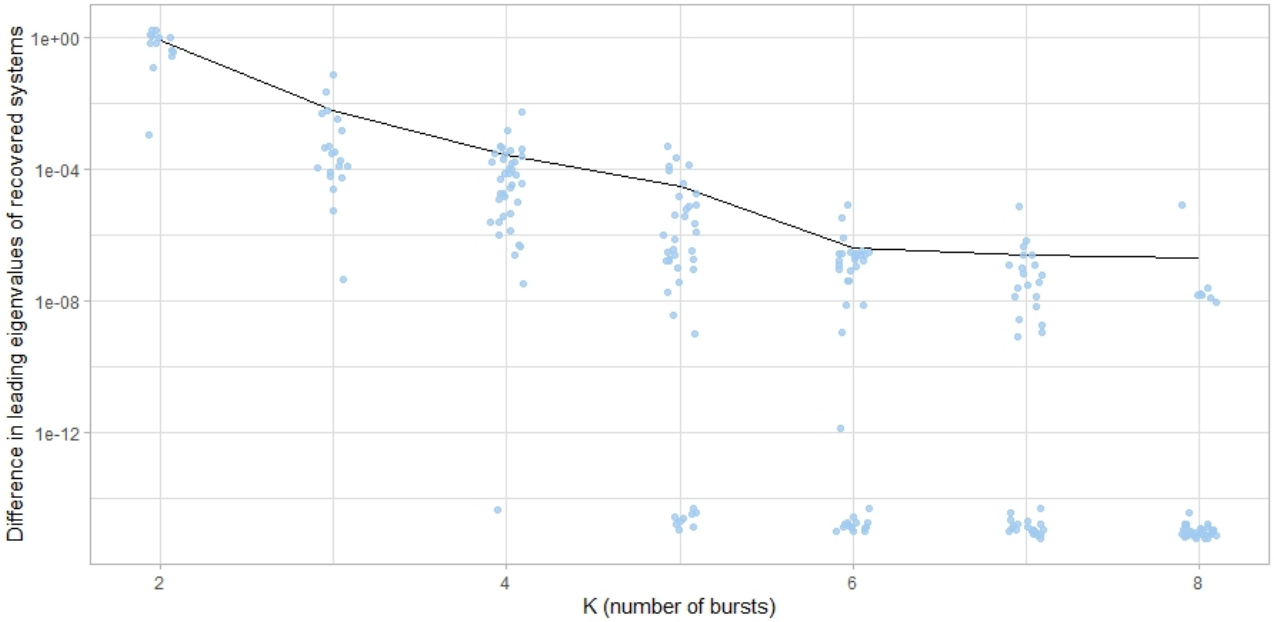

Figure 8. Jittered plot showing the difference between the dominant Eigenvalue of A and $\hat{A}$

Figure 9 shows both the Type I and Type II Error for trophic link recovery. For two bursts, there is a fairly significant amount of error; more false negatives than false positives. However, by three bursts, for all matrices that are recovered in their entirety, every single trophic link is correctly identified, and that holds for all bursts greater than three as well.

Figure 10 follows similar trends to the previous metrics in that more bursts leads to better results. Note that the y-axis is logarithmically scaled, and so while there is significant improvement between all burst lengths, we can conclude that there is a certain point at which more bursts are not necessary. When using empirical data, the difference between an error of 10^{-12} and 10^{-16} would be negligible.

Having covered one of the major issues with data extensively (not enough), we now turn to the other: what if the data is noisy? This is a critical consideration because empirical noiseless data is impossible. While stability to noise

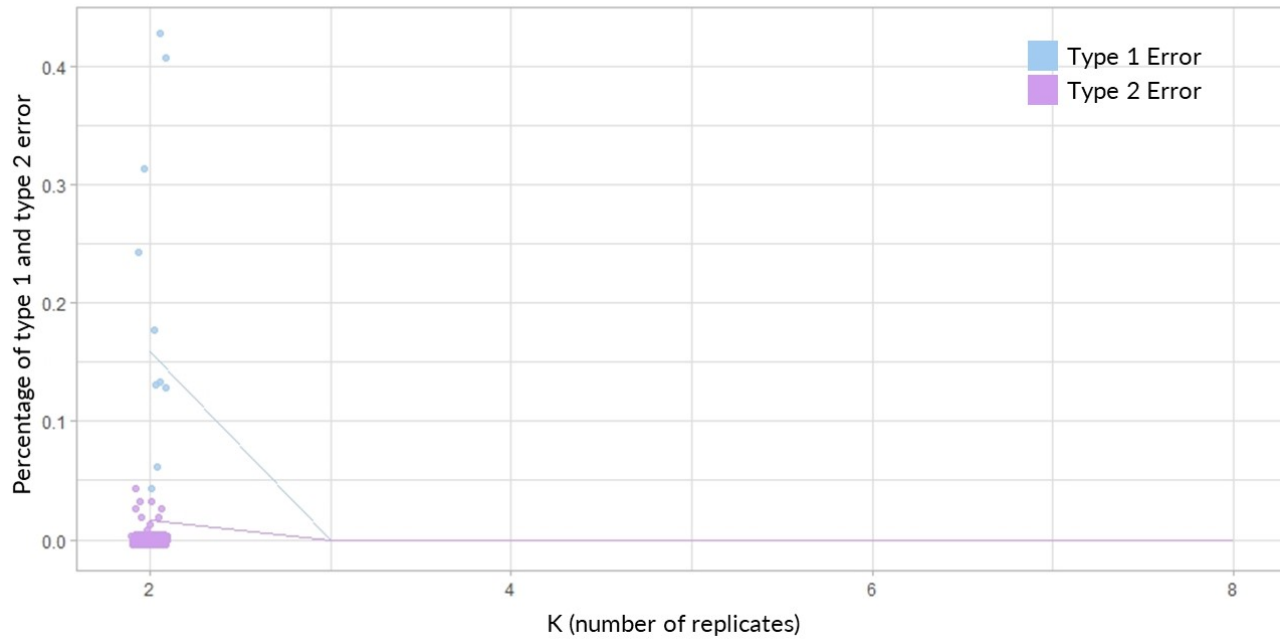

Figure 9. Jittered plot showing the Type I and Type II error of $\hat{A}$ with respect to A as a function of number of bursts

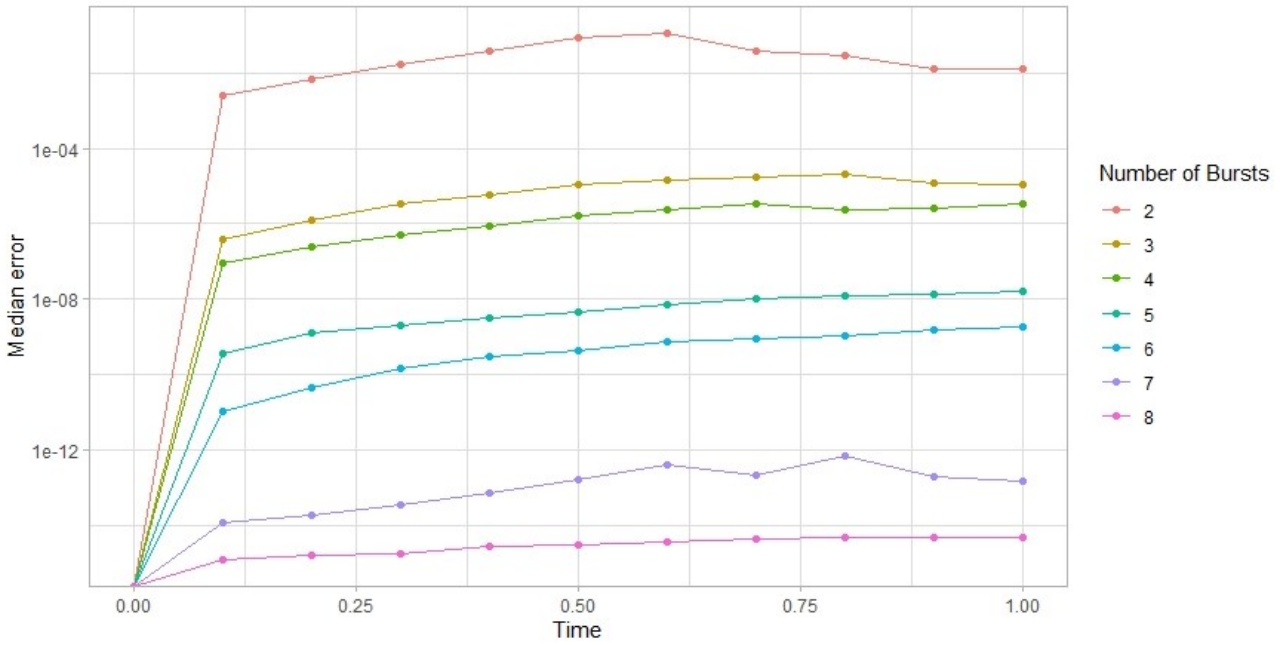

Figure 10. Median error of reconstructed biomass of each species as a function of time

was not a parameter tested in this paper, it has been shown that this method of compressive sensing is noise-resilient. We measure noise as follows: if ϵ is a matrix of the same dimension as the measurement matrix, X, generated through a Gaussian distribution, let us define the noise ratio as

$$\text{Noise ratio} = \frac{||\epsilon||_2}{||X||_2} \times 100\% \tag{11}$$

Gaussian noise of up to 5% allows for stable reconstruction, after 5%, we see a marked increase in ℓ_2 error, with noise over 6% yielding inaccurate and unstable results[17].

4 Discussion

Preliminary results are extremely promising. Compressive sensing was able to estimate, with a great degree of accuracy, the coefficients of the underlying Lotka-Volterra system generated via the niche model and allometric coefficients.

Compressive sensing is by no means the only way to estimate how dependent variables drive independent variables. In fact, the entire field of regression analysis is centred on this question. There are, however, a few characteristics of compressive sensing which make it particularly suited to ecological contexts over other common methods such as ordinary least squares regression (a form of linear regression).

The number of possible connections grow exponentially with the number of species. As such, linear regression, which solves overdetermined systems, becomes very data hungry very quickly. A general rule of thumb is that for each variable we aim to solve for, ten to twenty data points are required[9]. This is less of an issue if the trophic links are known; in that case, some of the coefficients can be set as zero, and so we reduce the data required whereby reducing coefficients to solve for. However, even when trophic links are "known" it is important to consider that they are not a certainty and new links are still being uncovered[1]. Taking into account non-trophic relationships becomes even more difficult, as we run into the same issue of the data required growing exponentially while ethological links and their comparative strengths are still relatively unknown. Conversely, using a monomial decomposition then compressive sensing has a theoretical lower bound on data required that scales logarithmically with the number of species[17]. This affords us the luxury of not having to choose which links are "active" in the estimated final system, and instead allows the algorithm to pick the best possible answer from a larger solution space, and ensures preconceived notions and biases play less of a role. While feasibility of compressive sensing was explored within a dynamical system of purely trophic links, it does not require all links to be trophic. In fact, the reverse is true: compressive sensing makes no distinction between trophic, behavioural, or any other type of link, potentially allowing for a deeper understanding of ethological links between species.

Linear regression also makes some key assumptions about the inputs which are not guaranteed to be satisfied. Firstly, observations must be made independently of each other. Given that time series data represents the biomass of a species over time, it is trivial to see that data taken at one time is not independent of the data taken at lag one[9]. Multicollinearity, or the strong correlation of explanatory variables is common, especially among short time bursts, and even in chaotic systems[3]. This can obfuscate the true nature of relationships and complicates statistical and inferential interpretation[8][25].

This is of course not to say the compressive sensing is a flawless solver. The major challenge in reconstruction for compressive sensing are phase transitions, which is a general term for a sizable threshold which marks a sudden shift in the likelihood of a given property being observed. That is, there is a threshold such that for any amount of data

below this amount, compressive sensing is unable to return a result at all; there is no gradual breakdown[26]. The tests we ran used two to eight bursts because all simulations with one burst, failed to yield any estimate at all. This is a property inherent to the solving method itself, and a potential future avenue for research would be to explore how phase transitions are affected by having access to data on only certain focal species within the system. Would it still be possible to estimate the dynamics accurately[1]? Are there certain ecosystems that are more resistant and likely to be recovered by fewer bursts, and what properties might they share?

In summary, compressive sensing has proven capable of parameter estimation in an ecological context, and unsurprisingly, the potential it has demonstrated simply opens up more avenues for future research. Compressive sensing is exceptional at recovering trophic links without any inputs detailing interactions and it is reasonable to assume that it would perform equally well in recovering other types of interactions in other models. This fits nicely with the demands of fishery modeling, where undersampled data is common and current models often ignore linkages in lower trophic levels[1]. Whether the trophic links can be recovered using compressive sensing, the value of the coefficients determined using least squares or LASSO, and whether this would even be efficient remains to be seen. Further study on the recovery of focal species dynamics and whether the best reconstructed set of interactions would still be sparse would also be necessary to better understand the full extent of compressive sensing's capabilities in reconstructing population dynamics.

References

1. W. Fennel and T. Neumann. Introduction. In *Introduction to the Modelling of Marine Ecosystems*, pages 1–12. Elsevier, 2015.
2. C. J. Walters and S. J. D. Martell. *Fisheries ecology and management*. Princeton University Press, Princeton, N.J, 2004. OCLC: ocm53483591.
3. M. H. Graham. CONFRONTING MULTICOLLINEARITY IN ECOLOGICAL MULTIPLE REGRESSION. *Ecology*, 84(11):2809–2815, November 2003.
4. D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, April 2006.
5. M. Pascual and J. A. Dunne, editors. *Ecological networks: linking structure to dynamics in food webs*. Santa Fe Institute studies in the sciences of complexity. Oxford University Press, Oxford ; New York, 2006. OCLC: ocm60312026.
6. R. G. Wetzel and G. E. Likens. Predator-Prey Interactions. In *Limnological Analyses*, pages 257–262. Springer New York, New York, NY, 2000.
7. P. Karlsson T. Jonsson and A. Jonsson. Trophic interactions affect the population dynamics and risk of extinction of basal species in food webs. *Ecological Complexity*, 7(1):60–68, March 2010.
8. J. H. Kim. Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, 72(6):558–569, December 2019.
9. F. E. Harrell. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. Springer Series in Statistics. Springer International Publishing, Cham, 2015.

10. A. J. Lotka. Elements of Physical Biology. 116(2917).
11. V. Volterra. Fluctuations in the Abundance of a Species considered Mathematically. *Nature*, 118(2972):558–560, October 1926.
12. R. M. May. Will a Large Complex System be Stable? *Nature*, 238(5364):413–414, August 1972.
13. S. L. Brunton and J. N. Kutz. *Data-driven science and engineering: machine learning, dynamical systems, and control*. Cambridge University Press, Cambridge, 2019.
14. E. J. Candes, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, February 2006.
15. P. Yodzis and S. Innes. Body Size and Consumer-Resource Dynamics. *The American Naturalist*, 139(6):1151–1175, June 1992.
16. R. J. Williams and N. D. Martinez. Simple rules yield complex food webs. *Nature*, 404(6774):180–183, March 2000.
17. G. Tran H. Schaeffer and R. Ward. Extracting Sparse High-Dimensional Dynamics from Limited Data. *SIAM Journal on Applied Mathematics*, 78(6):3279–3295, January 2018.
18. E. J. Pedersen, M. Koen-Alonso, and T. D. Tunney. Detecting regime shifts in communities using estimated rates of change. *ICES Journal of Marine Science*, 77(4):1546–1555, July 2020.
19. R. Tibshirani T. Hastie and M. Wainwright. *Statistical Learning with Sparsity: The Lasso and Generalizations*.
20. M. Fornasier and H. Rauhut. Compressive Sensing. In Otmar Scherzer, editor, *Handbook of Mathematical Methods in Imaging*, pages 205–256. Springer New York, New York, NY, 2015.
21. R. J. Williams J. A. Dunne and N. D. Martinez. Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences*, 99(20):12917–12922, October 2002.
22. T. Petzoldt K. Soetaert and R. W. Setzer. Solving differential equations in R: Package deSolve. *Journal of Statistical Software*, 33(9):1–25, 2010.
23. B. Narasimhan A. Fu and S. Boyd. CVXR : An R Package for Disciplined Convex Optimization. *Journal of Statistical Software*, 94(14), 2020.
24. S. Boyd. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2010.
25. Hal Caswell. Matrix Methods for Population Analysis. In S. Tuljapurkar and H. Caswell, editors, *Structured-Population Models in Marine, Terrestrial, and Freshwater Systems*, pages 19–58. Springer US, Boston, MA, 1997.
26. C. Zhang Z. Yang and L. Xie. On Phase Transition of Compressed Sensing in the Complex Domain. *IEEE Signal Processing Letters*, 19(1):47–50, January 2012.

5 Appendix

5.1 Variables & Definitions

Note that not all variables are included in this list. Allometric coefficients in Table **??** are omitted as are the variable briefly used to describe norms, as they are not used elsewhere in the paper.

Variable	**Definition**
n	Number of species
S	Connectance or Sparsity
$\mathbf{x}$	Vector of biomasses of all species at a given time
x_i	Biomass of i^{th} species, i^{th} value of $\mathbf{x}$
$\dot{\mathbf{x}}$	Rate of change vector for biomasses of all species at a given time
t	Time
$\mathbf{r}$	n-length vector of species growth rates
r_i	Growth rate of species i, i^{th} value of $\mathbf{r}$
A	$n \times n$ interaction matrix, Lotka-Volterra coefficients
a_{ij}	Effect of species j on species i, Lotka-Volterra coefficient
ν_i	Niche value of i^{th} species
ρ_i	Niche range diameter of i^{th} species
γ_i	Niche range centre of i^{th} species
R	Biomass of arbitrary resource species, Yodzi & Innes
P	Biomass of arbitrary resource species, Yodzi & Innes
v, w	Allometric rates, Yodzi & Innes
ϕ_i	Number of prey of species i
k	Index of current burst
K	Total number of bursts
$V^{(k)}$	Velocity matrix for k^{th} burst
$\dot{x_i}(k, t_m)$	Velocity of the biomass of species x_i in k^{th} burst at time m
M	Burst length, given as $M = 8$
C	Full coefficient matrix for all monomial bases, dimension $n \times (n^2 + n)$
$B_i^{(k)}$	Monomial basis for i^{th} species in k^{th} burst
C'	Reduced coefficient matrix, all guaranteed zero terms removed
V	"Stacked" list of velocity matrices for all k in K
$B_i^{(k)}$	"Stacked" list of monomial for all k in K
$\hat{C}$	Estimate of C'
$\hat{A}$	Estimate of A
$\hat{\mathbf{r}}$	Estimate of $\mathbf{r}$