RESEARCH PAPER

JOURNAL OF Evolutionary Biology .ૡ૯seb WILEY

Extinction in complex communities as driven by adaptive dynamics

Vu A. T. Nguyen [ID] | Dervis Can Vural [ID]

University of Notre Dame, South Bend, IN, USA

Correspondence
Dervis Can Vural, University of Notre Dame, South Bend, IN, USA.
Email: dvural@nd.edu

Abstract

In a complex community, species continuously adapt to each other. On rare occasions, the adaptation of a species can lead to the extinction of others, and even its own. 'Adaptive dynamics' is the standard mathematical framework to describe evolutionary changes in community interactions, and in particular, predict adaptation driven extinction. Unfortunately, most authors implement the equations of adaptive dynamics through computer simulations that require assuming a large number of questionable parameters and fitness functions. In this study, we present analytical solutions to adaptive dynamics equations, thereby clarifying how outcomes depend on any computational input. We develop general formulas that predict equilibrium abundances over evolutionary time scales. Additionally, we predict which species will go extinct next, and when this will happen.

KEYWORDS
adaptation, adaptive dynamics, adaptive extinction, evolution of ecological communities, evolutionary suicide, simulation, theory

1 | INTRODUCTION

For many ecological communities, the changes in population abundances due to interspecies interactions are much faster than the evolutionary changes in the traits that mediate these interactions. This affords us the theoretical convenience of dealing with population dynamics equations with unchanging interaction parameters (Turchin, 2003). However, to describe the population dynamics of interacting communities over long periods, or of those that evolve rapidly, one must also take into account how species coadapt and interact differently over time (Dawkins & Krebs, 1979; Dietl & Kelley, 2002; Maor & Shirasu, 2005; Stahl & Bishop, 2000; Vural et al., 2015; Weitz et al., 2005). To illustrate, when a pathogen crosses a species barrier, a previously zero interaction parameter between it and its host becomes finite. Conversely, when a pathogen (perhaps the same pathogen) kills all but resistant individuals, a previously finite interaction parameter drops to zero.

Amending population dynamics equations to include such adaptive changes to interspecies interactions is referred as 'adaptive dynamics' (AD) (Abrams, 1999; Abrams, 2000; Cortez et al., 2020; Dieckmann & Law, 1996; Dieckmann et al., 1995; Gandon et al., 2008; Gavrilets, 1997; Hochberg & Holt, 1995; Jian et al., 2016; Lehtinen & Geritz, 2019; Lion, 2018; Loeuille & Hauzy, 2018; Marrow et al., 1996; Rosenzweig et al., 1987; Rosenzweig & Schaffer, 1978; Schaffer & Rosenzweig, 1978). We should caution that AD is applicable only to asexual, well-mixed communities, in the low mutation rate limit.

As with much of classical population dynamics, AD typically focuses on demonstrating the stability of communities—even in the face of perpetual evolutionary arms races (Cortez & Patel, 2017; Cortez et al., 2020; Jian et al., 2016; Lehtinen & Geritz, 2019; Weitz et al., 2005). However, adaptive changes in interspecies interactions can occasionally lead to catastrophic displacements in equilibrium abundances, and even extinction, as was suggested theoretically

(Boldin & Kisdi, 2016; Gyllenberg & Parvinen, 2001; Leimar, 2002; Marrow et al., 1996; Matsuda & Abrams, 1994; Matsuda & Abrams, 1994; Parvinen, 2005; Parvinen, 2010; Parvinen & Dieckmann, 2013) and empirically (Conover & Munch, 2002; Fiegna & Velicer, 2003; Howard et al., 2004; Muir & Howard, 1999; Olsen et al., 2004; Rankin & López-Sepulcre, 2005). This phenomenon is referred as 'evolutionary suicide'. Analytical studies of adaptive extinction have so far been limited to only single species mutating (Boldin & Kisdi, 2016; Gyllenberg & Parvinen, 2001; Matsuda & Abrams, 1994; Matsuda & Abrams, 1994; Parvinen, 2005; Parvinen & Dieckmann, 2013) or coevolving two-species communities (Leimar, 2002; Marrow et al., 1996; Parvinen, 2010).

Prey-predator, host-parasite, plant-herbivore coevolution, as well as suicidal over-hunting, are standard textbook material. However, in a complex community, arms races need not be an exclusively two-species phenomenon. The equilibrium abundances of two species may change only marginally as they coevolve, while still causing a chain of events that lead to the extinction of other, possibly even far removed, species. Furthermore, in a complex community, evolutionary arms races need not take place between two species privately and end when one goes extinct. Arms races could involve so many species and span such long times that speciation and extinction events can be viewed as ordinary background events that merely decorate the Hutchinsonian evolutionary stage.

In this paper, we are primarily concerned with calculating the rate of extinction in a complex community, as driven by the parallel co-evolution of its members. We will be interested in predicting which species will next go extinct, when this will happen, and how fast a community will lose its members to complex evolutionary arms races. While doing so, we will also obtain how equilibrium abundances and interaction mediating traits evolve.

As general, widely applicable and conceptually insightful AD models are, in practice, questions of this kind can presently be addressed only by elaborate computer simulations that require a specific choice of initial conditions for the species abundances and interaction structure. Furthermore, these initial conditions must be accompanied by an army of functions that relate evolving traits to evolving interspecies interaction parameters. These functions are parameterized by an army of constants, typically randomly chosen according to specific probability distributions, which themselves are parameterized by arbitrarily chosen constants.

For example, (Ito & Dieckmann, 2007) discovered that if there is frequency-dependent disruptive selection in one trait and weak directional selection in another, then the community can enter a cycle of recurrent adaptive radiations and extinctions. Although this provides a mechanism for continuous introduction and removal of species it requires very particular interaction functions that allow trait branching. Likewise, (Johansson & Dieckmann, 2009) investigated the consequences of extinctions in ecological communities and found evolutionary keystone species that can switch the community from one evolutionary domain of attraction to another. Although their work carefully describes the convergence landscape, due to the combinatorially large possible number

of trajectories it is extremely difficult to gain general insights into how fast a large, complex community would decline under adaptive dynamics. Other theoretical studies that examine extinction-speciation dynamics suffer from the similar difficulties, of having to assume highly specific parameter sets (Dercole, 2003; Ito & Dieckmann, 2014; Kisdi et al., 2002).

Given the high dimensionality of initial conditions and model parameters, and given the sensitivity of the outcomes to the precise values of some of these parameters, numerical simulations provide little general insight into the macroscopic consequences of adaptive dynamics for short, medium and long time scales. For short time scales, we should be able to predict how species abundances change upon evolving interspecies interactions. For medium time scales, we should be able to predict how often a species will go extinct upon an accumulation of such adaptations; and for long time scales, we should be able to predict if and when an entire community would accumulate enough adaptations to annihilate itself. Furthermore, we should be able to make these predictions for any community, without having to assume functions, parameters or probability distributions that are not only highly specific and questionable, but also highly consequential. Lastly, since it is practically impossible to measure an interaction matrix in entirety, we would like to know which interaction parameters suffice to make 'reasonably accurate' predictions, and empirically measure only those.

Here, we take a step towards addressing the above questions by deriving accurate and general analytical solutions to adaptive dynamics equations. Since our results are analytical, they allow us to see clearly how population trajectories and extinction times depend on model inputs, without necessarily assuming what they are.

Our work consists of two parts. In part one, we assume that the interspecies interactions A_{ij} change in time according to $\mathbf{A}(t) = \mathbf{H}_0 + \mathbf{H}_1 t + \mathbf{H}_2 t^2 + \ldots$ and calculate the equilibrium abundances and extinction times of species in terms of $\mathbf{H}_0, \mathbf{H}_1, \mathbf{H}_2, \ldots$ This way, if $A_{ij}(t)$ is measured empirically at multiple time points, one can fit a line or polynomial and from there extrapolate equilibrium abundances and extinction times.

Despite being a useful tool for extrapolation, the result of part one is really a stepping stone. Its true strength will become apparent in part two of our paper, where we derive what the $\mathbf{H}$'s are, starting with the canonical equations for adaptive dynamics, which take into account the coupling between evolving traits and interaction matrix elements that depend on these traits.

Equipped with the results of both part one and two, one could now measure how the interactions $\mathbf{A}(\vec{s})$ depend on traits $\vec{s} = \{s_1, s_2, \ldots\}$, fit a line to each, and then feed these slopes into our analytical formulas to get future abundances $\vec{x}(t)$, trait evolution $\vec{s}(t)$ and extinction times τ. After we analytically obtain $\vec{x}(t)$, $\vec{s}(t)$ and τ, we apply them to specific AD models which we gather from the literature. We compare our general analytical formulas with the numerical solutions of three specific models (which, as is customary, we run with a large number of made-up parameter values and ad-hoc functional forms): One where $\mathbf{A}(t)$ is a random matrix, where each matrix element changes at its own random constant rate, one where

—WILEY—

$\mathbf{A}(\vec{s})$ depends on the difference between two evolving traits, as in (Rosenzweig et al., 1987; Rosenzweig & Schaffer, 1978; Schaffer & Rosenzweig, 1978), and one where $\mathbf{A}(\vec{s})$ has a Gaussian dependence on evolving traits, as in (Marrow et al., 1996; Marrow et al., 1992).

Throughout, we verify our analytical results with deterministic (mean field) simulations, as well as stochastic simulations (cf. Numerical Methods), and find very good agreement.

We conclude our introduction with an exposition of canonical adaptive dynamics, as much as it relates to the present work. In classical population dynamics, the growth rate (fitness) $f_i(\vec{x})$ of the abundance x_i of species i depends on the abundances of other species in the community, $\vec{x} = \{x_1, x_2, \ldots\}$. That is, $dx_i/dt = f_i(\vec{x}) x_i$.

There are many population dynamics models with varying degrees of complexity and realism (Turchin, 2003), and adaptive dynamics can be imposed on any of them. Here, we will follow the original AD formalism which assumes the simple Lotka-Volterra (LV) fitness $f_i = r_i + \sum_j A_{ij} x_j$, where A_{ij} quantifies how the jth species affects the ith species and r_i the growth rate of the species in isolation. The LV model is a rather crude representation of reality, however, under certain biotic and abiotic conditions, LV does overlap with more realistic models. These conditions are well established and well understood (cf. O'Dwyer (2018) and references therein). Furthermore, due to its analytical simplicity, LV also forms the basis of most empiric measurements of interspecies interactions (Delmas et al., 2019).

While standard LV equations keep A_{ij} unchanging, AD considers parameters $A_{ij}(\vec{s})$ that depend on some traits $\vec{s} = \{s_1, s_2, \ldots\}$. The specific functional form of $A_{ij}(\vec{s})$ depends on the ecological details of the system of interest (Abrams, 1999; Abrams, 2000; Cortez et al., 2020; Dieckmann & Law, 1996; Dieckmann et al., 1995; Gandon et al., 2008; Gavrilets, 1997; Hochberg & Holt, 1995; Jian et al., 2016; Lehtinen & Geritz, 2019; Lion, 2018; Loeuille & Hauzy, 2018; Marrow et al., 1996; Rosenzweig et al., 1987; Rosenzweig & Schaffer, 1978; Schaffer & Rosenzweig, 1978). However, the defining aspect of all AD models is that novel strains with differing traits (and therefore different interaction matrix elements) will be introduced, and subsequently, LV equations will govern the fate of these strains, thereby playing the role of natural selection.

It was shown in (Dieckmann & Law, 1996) that in the rare mutation / fast selection limit, the traits governing the interspecies interactions evolve, on average, according to the so called canonical equation,

$$\frac{ds_{k\alpha}}{dt} = \frac{1}{2} \mu_{k\alpha} \sigma_{k\alpha}^2 \frac{x_k}{\lambda} \frac{\partial f_k}{\partial s_{k\alpha}} \tag{1}$$

where $s_{k\alpha}$ is the αth trait of species k, $\mu_{k\alpha}$, and $\sigma_{k\alpha}^2$ are the mutation rate and variance of mutation strength for $s_{k\alpha}$, λ is the abundance corresponding to a single individual so that x_k/λ represents the number of individuals, and f_k is the LV fitness defined above. This equation makes intuitive sense: traits change at a rate proportional to how much they increase fitness and also proportional to the rate at which a population can generate variation. This equation constitutes the starting point of the present work (Figure 1).

2 | RESULTS

Setting $dx_i/dt \to 0$ with $x_i \neq 0$ in the LV equations, gives the coexistent equilibrium condition:

$$\vec{x} = -\mathbf{A}^{-1} \vec{r}. \tag{2}$$

from which we qualitatively see that as small changes accumulate in $\mathbf{A}$, the equilibrium abundances $\vec{x}$ will move along some trajectory, which occasionally might cross zero. Our goal is to determine $\vec{x}(t)$, the components i of $\vec{x}$ that crosses zero, and the times τ_i for which $x_i(\tau_i) = 0$.

2.1 | Implicit adaptive dynamics

According to AD, as traits $\vec{s}$ climb up a fitness landscape, the interactions $A_{ij}(\vec{s})$ that depend on them also change, leading to a difference in equilibrium abundances. Suppose we somehow know $A_{ij}(t)$, in the form.

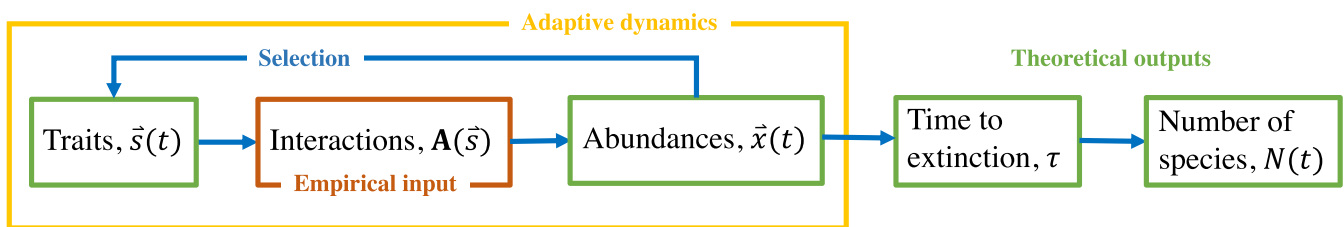$$\mathbf{A}(t) = \mathbf{H}_0 + \mathbf{H}_1 t + \mathbf{H}_2 t^2 + \ldots \tag{3}$$

FIGURE 1 Schematics of our approach. Here, we analytically solve canonical adaptive dynamics equations (yellow), where interaction mediating traits evolve as to climb up a fitness landscape, as defined according to Lotka-Volterra-type equations. We take the trait dependence of interspecies interactions $\mathbf{A}(\vec{s})$ as given (red), and from here, calculate (green), (1) how equilibrium abundances shift, $\vec{x}(t)$; (2) the time τ it takes for this shift to lead to a single extinction and (3) the species that gets extinct the soonest; (4) the number of species as a function of time $N(t)$, as many such extinctions accumulate; and (5) the evolution of the traits $\vec{s}(t)$ while all the above is happening. The blue arrows indicate causal links, and also show which quantity is used to calculate which other

where the **H** matrices could have been obtained experimentally, by measuring **A**(t) at few time points and fitting a line or polynomial; or theoretically, by modelling the evolution of $\bar{s}(t)$ and knowing **A**($\bar{s}$).

Of course, in reality, the interaction matrix elements will not change all together continuously (mean-field dynamics). Instead, small discrete changes to randomly selected matrix elements will accumulate (stochastic dynamics). Of course, when mutation strength is small, over long times, the two approaches give identical results (as shown in Figures 2 and 3). Thus, we proceed with analysis based off of a mean-field description of Equation 3.

We find how abundances evolve by expanding $\bar{x}(t) = \bar{\varepsilon}_0 + \bar{\varepsilon}_1 t + \bar{\varepsilon}_2 t^2 + \ldots$, inserting this and Equation 3 into Equation 2, and match the coefficients of every power of t (for details, see Appendix A),

$$\bar{\varepsilon}_m = -\mathbf{A}_0^{-1} \sum_{k=1}^{m} \mathbf{H}_k \bar{\varepsilon}_{m-k}. \tag{4}$$

Lastly, we evaluate all $\bar{\varepsilon}$'s and plug them into the definition of $\bar{x}(t)$.

$$\bar{x}(t) = \bar{x}_0 + \left[\mathbf{U}_1 t + \left(\mathbf{U}_1^2 + \mathbf{U}_2 \right) t^2 + \left(\mathbf{U}_1^3 + \mathbf{U}_1 \mathbf{U}_2 + \mathbf{U}_2 \mathbf{U}_1 + \mathbf{U}_3 \right) t^3 + \ldots \right] \bar{x}_0 \tag{5}$$

where $\mathbf{U}_k \equiv -\mathbf{A}_0^{-1} \mathbf{H}_k$. The pattern here is that the coefficient of the mth power of time consists of all **U**'s that sum up to m, in every possible order. For example, the coefficient of t^4 is a sum of $\mathbf{U}_1^4$, $\mathbf{U}_1^2 \mathbf{U}_2$, $\mathbf{U}_2 \mathbf{U}_1^2$, $\mathbf{U}_1 \mathbf{U}_2 \mathbf{U}_1$, $\mathbf{U}_2^2$, $\mathbf{U}_1 \mathbf{U}_3$, $\mathbf{U}_3 \mathbf{U}_1$ and $\mathbf{U}_4$.

Note that Equation 5 is an infinite series and depending on the nature of the original interaction matrix $\mathbf{A}_0$ might have a finite radius of convergence, which would prohibit us from knowing $\bar{x}(t)$ beyond a certain time. However, there is a quick workaround this: Suppose we already evaluated $\bar{x}(t)$ all the way till $t = T$ and would like to see beyond the radius of convergence. All we have to do is to reset the time to zero: we substitute $\mathbf{A}(T)$ and $\bar{x}(T)$ as initial conditions $\mathbf{A}_0$ and $\bar{x}_0$ and just use the same formula, now with a radius of convergence pushed forward in time. A principled way to determine the T at which to do this replacement is given in the final Appendix.

Using Equation 5, we can obtain the time to first extinction. We set $x_i(\tau) = 0$, solve for τ, and pick the i that has the smallest τ value (for details, see Appendix B),

$$\tau_m = \min_i \frac{-[\bar{x}_0]_i}{[\bar{\varepsilon}_1]_i + \sum_{j=2}^{m} [\bar{\varepsilon}_j]_i \tau_{m-1}^{j-1}}. \tag{6}$$

where $[\bar{\varepsilon}_j]_i$ denotes the ith component of the jth $\bar{\varepsilon}$ vector, and $\min_i()$ indicates to chose the value of i that yields the smallest positive value of (). This is an iterated solution. We start by plugging in $\tau_0 = 0$ on the right hand side to get τ_1. Then, we plug in τ_1 on the right hand side to get τ_2, and so on, such that τ_m rapidly approaches to the true value of τ with increasing m.

Evaluating the first approximation $m = 1$ for the time to extinction will turn out to be insightful,

$$\tau_1 = \min_i \left([\bar{x}_0]_i / \left[\mathbf{A}_0^{-1} \mathbf{H}_1 \bar{x}_0 \right]_i \right). \tag{7}$$

This result informs us of two risk factors for adaptation driven extinction: a small numerator and a large denominator. The first confirms what we might expect intuitively, that the species with the smallest abundances to begin with, are more likely to go extinct. The second risk factor is less intuitive. Species i with large $[\bar{\varepsilon}_1]_i = [\mathbf{A}_0^{-1} \mathbf{H}_1 \bar{x}_0]_i$ values are also more likely to go extinct. We tested the convergence of this scheme in Appendix B which shows rapid convergence to the true extinction time when an extinction event occurs within the domain of convergence for Equation 5.

Equation 7 also suggests that larger communities will lose species faster, since the minimum of a larger list of numbers will be smaller. We confirm this trend in our numerical simulations (Figure 5, first row).

We should emphasize that Equations 5 and 6 (and the approximate Equation 7) are very general: As long as we are provided with **H**'s we can track $\bar{x}(t)$ and determine extinction times regardless of the community structure, and regardless of the evolutionary mechanism that gives rise to the **H**'s. In fact, the next section will be devoted to obtaining **H** matrices starting with the canonical equation of adaptive dynamics, which will be simply plugged into the above formulas. However, before doing so, let us illustrate a more practical use of the results we obtained so far.

Suppose that **A** is empirically measured at two times and then a line is fit, $\mathbf{A} = \mathbf{A}_0 + \mathbf{H}_1 t$. In this special case, Equation 5 reduces to a geometric sum and can be evaluated in closed form,

$$\bar{x}(t) = \bar{x}_0 + \mathbf{U}_1 t (\mathbf{I} - \mathbf{U}_1 t)^{-1} \bar{x}_0. \tag{8}$$

where $\mathbf{U}_1 = -\mathbf{A}_0^{-1} \mathbf{H}_1$. This formula is exact, and luckily, even has an infinite radius of convergence (Appendix C). We compare it against numerical solutions and stochastic simulations and find excellent agreement (Figure 2).

To obtain the time to first extinction τ, we again set the left side of Equation 8 to zero, and identify which component i crosses zero first (see Appendix D).

$$\tau_m = \min_i \frac{-[\bar{x}_0]_i}{[\mathbf{U}_1 (\mathbf{I} - \mathbf{U}_1 \tau_{m-1})^{-1} \bar{x}_0]_i} \tag{9}$$

where again, we start the iteration with $\tau_0 = 0$ on the right side, and $\min_i$ indicates that the smallest positive value should be selected.

The first iteration for this special case gives us a τ_1 that is identical to Equation 7, which of course is expected, since the general case depends only on $\mathbf{H}_1$ to begin with.

Comparing Equation 9 to numerical simulations of a $N = 100$ species community (see Numerical Methods), we find high accuracy and rapid convergence. The simple expression $\tau_1 = \min_i \left(-[\bar{x}_0]_i / [\mathbf{U}_1 \bar{x}_0]_i \right)$ gives us a reasonably good estimation to begin with, off only by 13%. The next iteration $\tau_2 = \min_i \left(-[\bar{x}_0]_i / [\mathbf{U}_1 (\mathbf{I} - \mathbf{U}_1 \tau_1)^{-1} \bar{x}_0]_i \right)$ brings this error down to 3%.

To better understand the qualitative behaviour of Equation 8, we write $\bar{x}_0$ as a linear combination of the eigenvectors of $\mathbf{U}_1$ (see Appendix D).
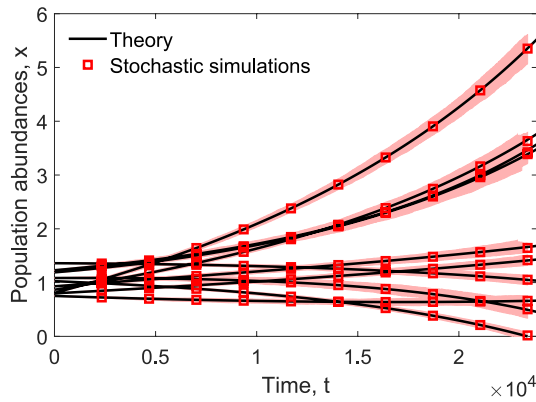
FIGURE 2 Equilibrium abundances $\bar{x}(t)$ in a community whose interactions are changing at a fixed rate. Our analytic formula (Equation 8) correctly tracks each species abundance over time and can also be used to predict an extinction. We sampled 1,000 stochastic trajectories to find the min and max bounds shown by the red shaded region using $\zeta = 0.1$ (cf. Numerical Methods) for a community with $N = 10$ species
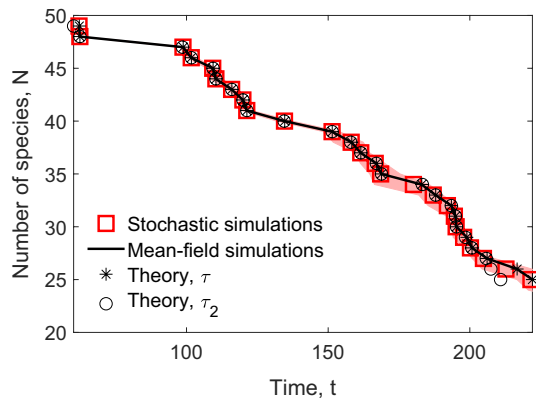

FIGURE 3 Decline of a community as driven by species coadapting. Here, we compare the extant number of species in a coevolving community where random matrix elements of **A** are mutated incrementally (stochastic simulations) and where all matrix elements change all together, continuously (mean-field simulations), to analytical theory Equation 8 and the second order approximation Equation 8. The shaded red regions contain all (5,000) simulated trajectories using the minimum and maximum extinction times at each community size. The initial $N = 50$ community has an interaction density $\rho = 1.0$ and randomly selects 50% ($\zeta = 0.5$) of interaction elements to change at each time-step (cf. Numerical Methods)

$$\bar{x}(t) = x_0 + \sum_{j,k=1}^{N} \frac{(\mathbf{V}^{-1})_{jk}[\bar{x}_0]_k \lambda_j t}{1 - \lambda_j t} \bar{v}_j \qquad (10)$$

where λ_j and $\bar{v}_j$ are the jth eigenvalue and eigenvector of $\mathbf{U}_1 = -\mathbf{A}_0^{-1}\mathbf{H}_1$; and **V** is a matrix formed by writing the jth eigenvector $\bar{v}_j$ into the jth column. We observe in Equation 10, that the denominator approaches zero, as time approaches one over the largest eigenvalue. This means that the abundances of many species will either boom or crash simultaneously depending on the sign of the

corresponding component of $\bar{v}_j$. Interestingly, we observed a similar singularity in a previous work that explored ways to edit community composition (Nguyen & Vural, 2020).

The first order formula for extinction time τ_1 is reasonably accurate, even though it can identify who goes extinct only 54% of the time (although still better than random guessing, 1%). This is because many species start crashing around the same time. Fortunately, the second iteration ($m = 2$) increases the success rate to 93%. Since the first order extinction time only takes into account the initial downward slope $\bar{\varepsilon}_1$, this observation suggests that about half the time, the species initially in rapid decline are not the ones that goes extinct first. Half of the time, second order effects (i.e. evolution of neighbours of neighbours in the interaction network) will accelerate an otherwise slow decline, a scenario we mentioned in the introduction.

Incidentally, if we are allowed to make two best guesses instead of one, extinctions can be predicted 71% ($m = 1$) and 99% ($m = 2$) of the time.

Conveniently, our formulas for extinction time Equations 6 and 9 need not be used only for the soonest extinction. To obtain the times for latter extinctions, one simply removes the row and column from the interaction matrix corresponding to the species that went extinct, and substitute this reduced interaction matrix for $\mathbf{A}_0$ in Equations 6 and 9. This can be repeated to get a series of extinction times. Then, these extinction times can be put together to obtain the number of extant species as a function of time $N(t)$ (Figure 3).

In Figures 2 and 3 and first column of Figure 5, we plot Equations 8 and 9 and the repeated use of the latter to get $N(t)$. We compare these results with mean-field simulations and stochastic simulations and find excellent agreement.

We should caution that the iterated use of Equations 6 and 9 to obtain $N(t)$ neglects the possibility that extinct species can return back later. In other words, we remove species from the interaction matrix permanently, thereby neglecting the (seemingly rare) scenario where a community becomes uninhabitable for a species, after some time becomes habitable again, and the originally extinct species migrates or mutates back and successfully fixes.

2.2 | Canonical adaptive dynamics

In the previous section, we obtained how equilibrium abundances change for any community, given $\mathbf{H}_1$, $\mathbf{H}_2$,... In this section, we will obtain these **H** matrices assuming the canonical equation of adaptive dynamics, which relates interaction values with the evolution of traits they depend on.

Specifically, we assume that $A_{ij}(s_{11}, s_{21}, ...)$ depends on some traits $\bar{s} = \{s_{11}, s_{21}, ...\}$. If at present, $t = 0$ these traits are $\bar{s}_0$, for small changes in trait values we can Taylor expand.

$$A_{ij}(\bar{s}) \approx A_{ij}(\bar{s}_0) + \sum_{k\alpha} \frac{\partial A_{ij}}{\partial s_{k\alpha}}\bigg|_{\bar{s}_0} \delta s_{k\alpha} \qquad (11)$$

where $\delta s_{k\alpha} = (ds_{k\alpha}/dt)\,\delta t$ is the amount that the αth trait of the kth species can change within some time interval δt. The time interval for which this approximation is valid depends on the functional form of $A_{ij}(\bar{s})$.

Then, we write the chain rule

$$\frac{dA_{ij}}{dt} = \sum_{k\alpha} \frac{\partial A_{ij}}{\partial s_{k\alpha}} \frac{ds_{k\alpha}}{dt}, \tag{12}$$

plug in Equation 1 for $ds_{k\alpha}/dt$, with $f_k = r_k + \sum_l A_{kl} x_l$, expand $\bar{x}(t) = \bar{\varepsilon}_1 + \bar{\varepsilon}_1 t + \bar{\varepsilon}_2 t^2 + \dots$ and $A(t) = H_0 + H_1 t + H_2 t^2 + \dots$ and match polynomial coefficients to get (Appendix E)

$$[H_k]_{ij} = \sum_{nq\alpha} \frac{\mu_{n\alpha}\sigma_{n\alpha}^2}{2k\lambda} C_{ijn\alpha} \sum_{s=0}^{k-1} [\bar{\varepsilon}_s]_n C_{nqn\alpha} [\bar{\varepsilon}_{k-s-1}]_q \tag{13}$$

where $C_{ijk\alpha} = \partial A_{ij}/\partial s_{k\alpha}$ depends on the specifics of the biology, that is how interactions depend on traits. The indices n and q run over all the species, whereas α runs over all the traits. Note that in this formula, each order of H depends on previous orders of ε's, which themselves depend on previous orders of H, via Equation 4. Thus, Equation 13 can be evaluated order by order, starting with $H_0 = A_0$ and $\bar{\varepsilon}_0 = \bar{x}_0$.

This is our central result. With H's in hand, one evaluates Equations 5 and 6 to get abundances $\bar{x}(t)$, extinction times τ and number of extant species $N(t)$. Furthermore, now that we know $\bar{x}(t)$, we can plug it in Equation 1 and obtain the evolution of traits,

$$s_{k\beta}(t) = s_{k\beta}(0) + \frac{\mu_{k\beta}\sigma_{k\beta}^2}{2\lambda} \sum_{\alpha,q} \left[[\bar{x}_0]_k C_{kqk\alpha}[\bar{x}_0]_q t + ([\bar{x}_0]_k C_{kqk\alpha}[\bar{\varepsilon}_1]_q + [\bar{\varepsilon}_1]_k C_{kqk\alpha}[\bar{x}_0]_q) \frac{t^2}{2} + \dots \right] \tag{14}$$

where $s_{k\beta}$ is the β-th trait for species k. In Figure 4, we compare numerical simulations (cf. Applications and Numerical Methods sections) to these analytical formulae and find excellent agreement.
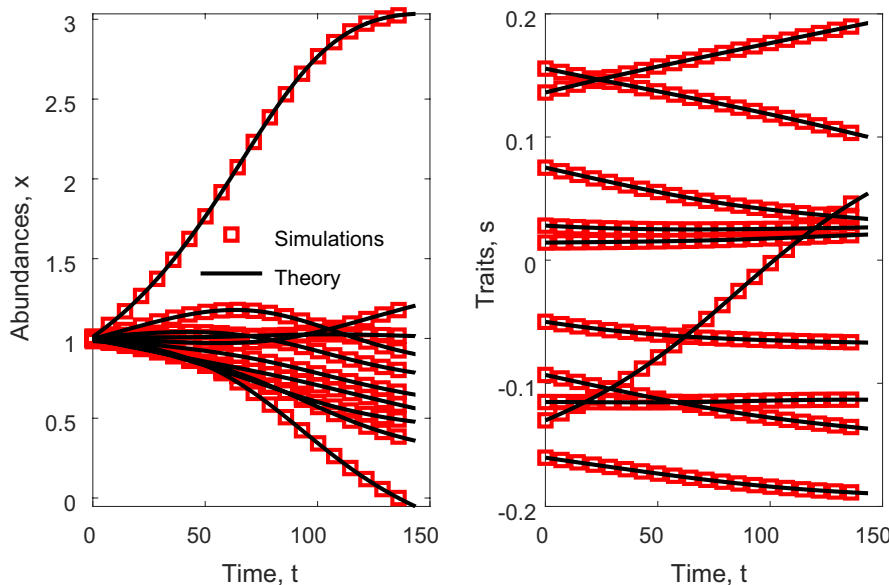
2.3 | Applications

We now use our central result, Equation 13 (as plugged into Equations 5 and 6) to solve some classical models gathered from the AD literature. All we need to do is to extract C_{ijkl} from a specific model, and plug it into Equation 13. We will illustrate how to do this for two models. In both models, species happen to have only have one evolving trait, so α can only be 1 in $C_{ijk\alpha}$, and $s_{i\alpha}$. For brevity, we will omit writing 1 each time, $C_{ijkl} = C_{ijk}$ and $s_{i1} = s_i$.

2.3.1 | Escalating traits (Linear A(s))

We start with the model in (Rosenzweig et al., 1987; Rosenzweig & Schaffer, 1978; Schaffer & Rosenzweig, 1978), which assumes an arms race scenario between every predator-prey pair. Over evolutionary time scales, each species develops their trait to counter the progress of their opponent. This has been empirically shown by Brodie and Brodie through the correlation between the toxin resistance of garter snakes and toxicity of newts in various geographical areas (Brodie & Brodie, 1990; Brodie & Brodie, 1991; Brodie & Brodie, 1999a; Brodie & Brodie, 1999b). Their work suggests snakes have greater toxin resistance in order to consume newts with higher toxicity.

The interaction between i and j depends on the difference between their two traits $A_{ij}(s_i, s_j) = (s_i - s_j)M_{ij}$ and $A_{ji}(s_i, s_j) = (s_i - s_j)M_{ji}$. Note that reciprocal interactions use the same trait difference but can have different signs (due to M). The first term scales the strength of the interaction with the difference in traits and allows for the scenario where this difference changes in sign. For example, if we were

FIGURE 4 Equilibrium abundances $\bar{x}(t)$ in an arms race adaptive dynamics community. Our analytic formula (Equation 5) correctly tracks each species abundance over time and can also be used to predict the traits over time. The $N = 50$ community was generated using the parameters presented in Methods using up to $O(t^4)$ for Equations 5 and 14. For visual clarity, we only show the first 10 species (unsorted)

modelling two types of carnivorous fish, if a prey species evolves to be larger than its predator then the role of predator and prey can be flipped.

Next, we evaluate $C_{ijk} = \partial A_{ij}/\partial s_k$. For all $i < j$, $C_{ijk} = M_{ij}\left(\delta_{ik} - \delta_{jk}\right)$, $C_{jik} = M_{ji}\left(\delta_{ik} - \delta_{jk}\right)$, $C_{iik} = 0$ where δ_{ik} is the Kronecker delta. We plug these into Equation 13 and compare the analytical model against numerical simulations of the model, and find excellent agreement (Figure 5, middle column).

2.3.2 | Bounded traits (Gaussian A(s))

Next, we consider the model in (Marrow et al., 1996; Marrow et al., 1992) where interactions are bounded, such that $A_{ij}(\vec{s})$ has a Gaussian peak. Specifically,

$$A_{ij}(s_i, s_j) = M_{ij}\exp\left[-\beta_{ij}^2 - \gamma_{ij}^2 + 2[\mathbf{P_5}]_{ij}\beta_{ij}\gamma_{ij}\right]$$
$$\beta_{ij} = (s_i - [\mathbf{P_1}]_{ij})/[\mathbf{P_2}]_{ij} \quad \gamma_{ij} = (s_j - [\mathbf{P_3}]_{ij})/[\mathbf{P_4}]_{ij} \tag{15}$$

where all model parameters are contained inside the matrices $\mathbf{P_1}$, $\mathbf{P_2}$, ..., $\mathbf{P_5}$. Similar to the linear $\mathbf{A}(\vec{s})$ model, we use $\mathbf{M}$ to incorporate the signs, trade-off and density of interactions. In this model, the interactions cannot change signs so one can incorporate a more realistic predator-prey type of interactions by restricting the positive interactions to be weaker than the negative interactions (cf. Numerical Methods).

Next, we evaluate $C_{ijk} = \partial A_{ij}/\partial s_k$. For all $i < j$

$$C_{ijk} = \delta_{ik}\frac{2A_{ij}}{[\mathbf{P_2}]_{ij}}\left([\mathbf{P_5}]_{ij}\frac{s_j - [\mathbf{P_3}]_{ij}}{[\mathbf{P_4}]_{ij}} - \frac{s_i - [\mathbf{P_1}]_{ij}}{[\mathbf{P_2}]_{ij}}\right)$$
$$+ \delta_{jk}\frac{2A_{ij}}{[\mathbf{P_4}]_{ij}}\left([\mathbf{P_5}]_{ij}\frac{s_i - [\mathbf{P_1}]_{ij}}{[\mathbf{P_2}]_{ij}} - \frac{s_j - [\mathbf{P_3}]_{ij}}{[\mathbf{P_4}]_{ij}}\right),$$
$$C_{jik} = \delta_{jk}\frac{2A_{ji}}{[\mathbf{P_4}]_{ji}}\left([\mathbf{P_5}]_{ji}\frac{s_i - [\mathbf{P_1}]_{ji}}{[\mathbf{P_2}]_{ji}} - \frac{s_j - [\mathbf{P_3}]_{ji}}{[\mathbf{P_4}]_{ji}}\right) \tag{16}$$
$$+ \delta_{ik}\frac{2A_{ji}}{[\mathbf{P_2}]_{ji}}\left([\mathbf{P_5}]_{ji}\frac{s_j - [\mathbf{P_3}]_{ji}}{[\mathbf{P_4}]_{ji}} - \frac{s_i - [\mathbf{P_1}]_{ji}}{[\mathbf{P_2}]_{ji}}\right),$$
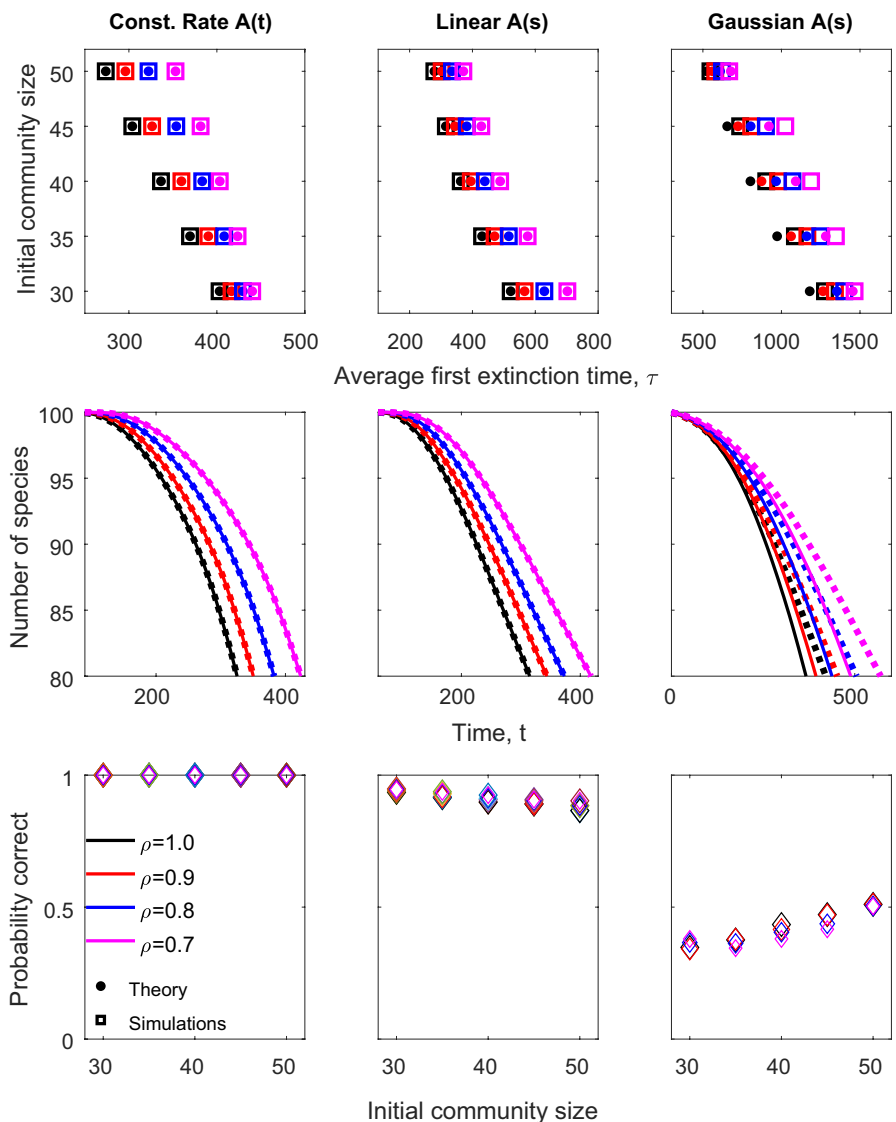$$C_{iik} = 0.$$

FIGURE 5 Adaptive Extinction. We gathered classical AD models from the literature (Marrow et al., 1996; Marrow et al., 1992; Rosenzweig et al., 1987; Rosenzweig & Schaffer, 1978; Schaffer & Rosenzweig, 1978), and compared their specific numerical solutions to our general formulae. To do so, we sampled 5,000 communities starting with varying initial number of species and interaction densities $\rho = (0.7, 0.8, 0.9, 1.0)$ as (black, red, blue, magenta), respectively. The results from adaptive dynamics (AD) are shown as squares and iterating the theory (middle and right columns) gives the results shown as solid/filled circles and the constant rate $\mathbf{A}(t)$ case (left column) uses the exact solution shown in Equation 10. Each simulation started with a community of size between $N = 30$ and $N = 100$ for interaction densities $\rho = [0.7, 0.8, 0.9, 1.0]$ Top: average first extinction time. Middle: number of surviving species over time. Bottom: probability of correctly predicting the first extinct species

As each species changes their trait monomorphically, we update the values of A according to each interaction model and then determine the population abundances using the Equation 2. These numerical simulations are compared against our analytical formula Equation 13 in the third column of Figure 4.

As expected, analytical theory agrees with simulations during some initial time, and then gradually departs (Figure 5, third column). This is because of our linearity assumption Equation 11. The Gaussian $A(\bar{s})$ can be approximated by a line only for small trait changes. Of course, one could remedy this by occasionally correcting the slope of $A(\bar{s})$. However, even without doing so, our formula retains its accuracy over a timescale that spans many extinctions (Figure 5, third row, second column). Note, however, that we can accurately pinpoint which species will go extinct only about half the time (Figure 5, third row, third column).

3 | NUMERICAL METHODS

We have applied our analytical formulae to solve some models gathered from the literature, which so far had only specific numerical solutions. In this section, we outline the initial conditions and parameter values we chose for reproducing these numerical solutions, for the purpose of comparing them against our analytical formulae (Figure 6).

All initial communities were generated using normal distribution for the population abundances and trait parameters for varying community sizes and interaction types. Across the three sampled models, we use the same average population abundance $\langle x \rangle = 1$ and standard deviation $\langle \sigma_x \rangle = 0.1$. For the two models which simulates AD traits, we initialized the traits $\langle s \rangle = 0$ with standard deviation $\langle \sigma_s \rangle = 0.1$. The classification matrix M randomly assigns predator-prey interactions for each pair (i, j). Additionally, we modify the interaction density, $0 \le \rho \le 1$, which is the probability that we do not assign $M_{ij} = M_{ji} = 0$ to a matrix element.

The initial communities for the constant rate $A(t)$ use the interactions generated by the linear $A(\bar{s})$ but does not simulate the traits. Instead, we randomly sampled the values of H from a normal distribution $H_{ij} = \mu \times N[1, 1]$ for $\mu = 10^{-4}$ and kept self-competition and initially nonexistent interactions from changing. The linear $A(\bar{s})$ model uses $\mu_k = 10^{-3}$, the Gaussian $A(\bar{s})$ model uses $\mu_k = 10^{-2}$, and both of these model use $\sigma_k = 10^{-3}$ and $\lambda_k = 10^{-3}$ for all species k. The stochastic trajectories were generated for Figures 2 and 3 by assigning each off-diagonal element a probability of change, $0 < \zeta \le 1$ per unit time. In order to keep the average change over time the same, the rate of change was modified to $H_{ij} \rightarrow H_{ij}/\zeta$ so as ζ decreased from unity each element experiences rare but stronger changes than the mean field.

The Gaussian $A(\bar{s})$ model requires additional parameters for its peaks (P_1, P_3) and widths (P_2, P_4). These parameters are symmetric for each interaction pair ($P_{ij} = P_{ji}$) with an average peak magnitude $\langle |P_1| \rangle = \langle |P_3| \rangle = 5$ and widths $\langle P_3 \rangle = \langle P_4 \rangle = 10$. The peak values have a standard deviation of 0.1 and are randomly chosen as either
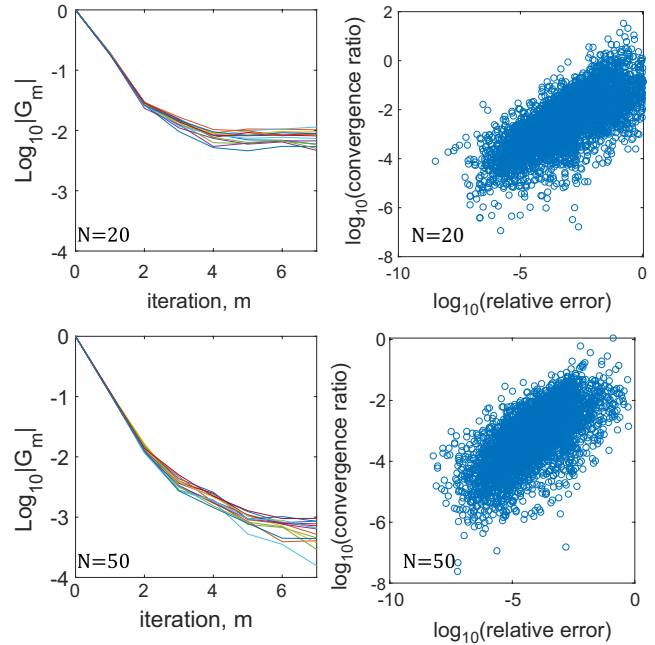

FIGURE 6 Convergence to true extinction time. Left: Average relative error, G_m, for initial ecosystem sizes of $N = 20$ (top) and $N = 50$ (bottom) with interaction density $\rho = [0.3, 0.5, 1.0]$ and scaling parameter $\mu = [10^{-4}, 10^{-3}, ..., 1]$ for $A(t) = A(0) + H_1 t$. Each line represents one parameter set (ρ, μ) averaged across 5,000 random ecosystems. The zeroth order starts the iteration with $\tau_0 = 0$ so it always starts with 100% error. Right: Convergence ratio estimated using Equation B6 for the $m = 10$ iteration compared to G_{10}. Generally, higher relative errors in τ positively correlate with high values of the convergence ratio

positive or negative with equal probability. The width values have unit standard deviation and are strictly positive. The coupling coefficients (P_5) for each interaction pair are uniformly random between zero and unity while keeping it symmetric. The sign of each Gaussian interaction cannot change as the community evolves its trait so positive interaction values $A_{ij} > 0$ are given an inefficient mass-transfer penalty of 10% in comparison with the negative interaction values ($A_{ij} = -0.1A_{ji}$). The overall magnitude of the Gaussian interactions is scaled up by a factor of 50 in order to accelerate changes.

Similar to (Marrow et al., 1996), we assume that the intrinsic birth/death vector does not change with mutations for both models. However, we also assume that the self-interaction for each species is also kept constant while (Marrow et al., 1996) allows the self-interaction of the prey to change with its trait and there is no self-limiting interaction for the predator. To keep the model simple, the self-interaction for all species is kept at −10.

In our models, we assume that a random walk in interaction space replicates a random walk in trait space of AD in a low mutation rate and high selection limit. Comparing the stochastic effect of this random walk to our deterministic mean-field model (Figures 2 and 3) shows that our deterministic (mean-field) approximations agree with the average stochastic behaviour. Therefore, we can average over the stochastic effects of randomly walking in interaction space by just using our mean-field theory.

4 | CONCLUSION

We obtained general analytical formulas for species abundances (Equations 5 and 8), evolution of interaction mediating traits (Equation 14), and a series of extinction times (Equations 6 and 9) for complex communities involved in multi-party arms races. We then showed how to evaluate these formulas when the species interactions are governed by the standard adaptive dynamics equations (Equation 13), applied these analytical formulas to solve AD models gathered from the literature (Marrow et al., 1996; Marrow et al., 1992; Rosenzweig et al., 1987; Rosenzweig & Schaffer, 1978; Schaffer & Rosenzweig, 1978) and compared our analytical solutions to specific numerical solutions of these models.

Having analytical formulas, as opposed to numerical solutions, show us the effect of every single parameter on the abundances and series of extinctions. For example, we find that a crucial quantity governing the extinction is $[\mathbf{A}_0^{-1}\mathbf{H}_1\bar{\mathbf{x}}_0]_i$ (for species i), which is the most prominent factor in almost all of our results. We can also see the precise dependence of relative initial populations on extinction time, as well as the influence of community size. Another interesting observation is, if all initial populations are modified by a constant factor, then Equation 6 shows us that $\bar{x}(t)$ scales by the same amount, but the extinction times (which after all, are obtained by setting the left hand zero), remain invariant. In other words, communities with small abundances exhibit similar adaptive extinction times with those with large abundances.

In closing, we mention two important problems that still remain open. In the present work, we analytically obtained a string of extinction times, which we stitched together to obtain the number of extant species $N(t)$ as a function of time. However, we were not able to obtain a closed form formula for $N(t)$ that would make apparent the conditions for which a community would approach to a fixed number of stably coexisting species (in the red queen sense), versus vanish entirely. For communities that do approach a stable number of coexisting coevolving species, we would like to know what this number is, as a function of initial community size and interaction parameters.

Such a result might also provide some insight into May's paradox (May, 1972), which, as many authors before us have pointed out, can be resolved by structured (as opposed to random) interaction matrices.

The evolutionary process that structure interaction matrices is adaptive dynamics. Thus, having a closed form formula for $\lim_{t\to\infty} N(t)$ would allow us to catalogue the interaction structures that allow nonvanishing communities, and those that lead to maximal (possibly, even divergent) number of species. It would also allow us to see what, if any, is the upper bound to the number of species that can coexist while also coevolving.

The second open problem is the following. Presently, our adaptive dynamics solutions accept as inputs, the slope of the function $\mathbf{A}(\bar{s})$. In other words, one must empirically measure how interactions change with respect to some traits, and fit lines. It seems to us that it should be feasible to generalize the calculations here to any arbitrary function of $\mathbf{A}(\bar{s})$, which then would allow evolutionary predictions that hold for a much longer time.

The main predictive challenge in testing the feedback between evolutionary and population dynamical timescales lies in determining the mapping from traits to interactions. It is currently not feasible to fully determine $N \times N$ functions, $\mathbf{A}(\bar{s})$, empirically. However, it *may* be feasible to measure the local rates of change in interactions $d\mathbf{A}/dt$. As such, our implicit AD approach, Equation 5 could allow for an easier (albeit less mechanistic) integration of observation and theory. When our analytical model is used with the local rates of change in the slopes, we can accurately track the abundance levels over long periods of time, albeit not indefinitely. However, coupling analytical theory with periodic measurements of the local values of $d\mathbf{A}/dt$ one can avoid mapping traits to interactions and still retain high accuracy in predicting abundance trajectories and series of extinctions.

CONFLICT OF INTEREST
The authors declare no conflict of interests.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/jeb.13796.

OPEN RESEARCH BADGES

This article has earned an Open Materials Badge for making publicly available the components of the research methodology needed to reproduce the reported procedure and analysis.

DATA AVAILABILITY STATEMENT
All code used in this analysis are available and uploaded as supplementary information. The extinction-package.zip is the said code.

ORCID
Vu A. T. Nguyen https://orcid.org/0000-0002-4712-5578
Dervis Can Vural https://orcid.org/0000-0002-0495-8086

REFERENCES
Abrams, P. A. (1999). The adaptive dynamics of consumer choice. *The American Naturalist*, *153*(1), 83–97.
Abrams, P. A. (2000). The evolution of predator-prey interactions: Theory and evidence. *Annual Review of Ecology and Systematics*, *31*(1), 79–105.
Boldin, B., & Kisdi, E. (2016). Evolutionary suicide through a non-catastrophic bifurcation: Adaptive dynamics of pathogens with frequency-dependent transmission. *Journal of Mathematical Biology*, *72*(4), 1101–1124.
Brodie, E. D. III, & Brodie, E. D. Jr. (1990). Tetrodotoxin resistance in garter snakes: an evolutionary response of predators to dangerous prey. *Evolution*, *44*(3), 651–659.
Brodie, E. D., & Brodie, E. D. (1991). Evolutionary response of predators to dangerous prey: reduction of toxicity of newts and resistance of garter snakes in island populations. *Evolution*, *45*(1), 221–224.
Brodie, E. D. III, & Brodie, E. D. Jr. (1999a). Costs of exploiting poisonous prey: evolutionary trade-offs in a predator-prey arms race. *Evolution*, *53*(2), 626–631.

Brodie, E. D. III, & Brodie, E. D. Jr. (1999b). Predator-prey arms races: asymmetrical selection on predators and prey may be reduced when prey are dangerous. *BioScience*, 49(7), 557–568.

Conover, D. O., & Munch, S. B. (2002). Sustaining fisheries yields over evolutionary time scales. *Science*, 297(5578), 94–96.

Cortez, M. H., & Patel, S. (2017). The effects of predator evolution and genetic variation on predator–prey population-level dynamics. *Bulletin of Mathematical Biology*, 79(7), 1510–1538.

Cortez, M. H., Patel, S., & Schreiber, S. J. (2020). Destabilizing evolutionary and eco-evolutionary feedbacks drive empirical eco-evolutionary cycles. *Proceedings of the Royal Society B: Biological Sciences*, 287(1919), 20192298.

Dawkins, R., & Krebs, J. R. (1979). Arms races between and within species. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161), 489–511.

Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Dalla Riva, G. V., Fortin, M.-J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., & Poisot, T. (2019). Analysing ecological networks of species interactions. *Biological Reviews*, 94(1), 16–36.

Dercole, F. (2003). Remarks on branching-extinction evolutionary cycles. *Journal of Mathematical Biology*, 47(6), 569–580.

Dieckmann, U., & Law, R. (1996). The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34(5–6), 579–612.

Dieckmann, U., Marrow, P., & Law, R. (1995). Evolutionary cycling in predator-prey interactions: Population dynamics and the red queen. *Journal of Theoretical Biology*, 176(1), 91–102.

Dietl, G. P., & Kelley, P. H. (2002). The fossil record of predator-prey arms races coevolution and escalation hypotheses. *The Paleontological Society Papers*, 8, 353–374.

Fiegna, F., & Velicer, G. J. (2003). Competitive fates of bacterial social parasites: persistence and self–induced extinction of myxococcus xanthus cheaters. *Proceedings of the Royal Society of London. Series B, Biological sciences*, 270(1523), 1527–1534.

Gandon, S., Buckling, A., Decaestecker, E., & Day, T. (2008). Host–parasite coevolution and patterns of adaptation across time and space. *Journal of Evolutionary Biology*, 21(6), 1861–1866.

Gavrilets, S. (1997). Coevolutionary chase in exploiter–victim systems with polygenic characters. *Journal of Theoretical Biology*, 186(4), 527–534.

Gyllenberg, M., & Parvinen, K. (2001). Necessary and sufficient conditions for evolutionary suicide. *Bulletin of Mathematical Biology*, 63(5), 981–993.

Hochberg, M. E., & Holt, R. D. (1995). Refuge evolution and the population dynamics of coupled host|parasitoid associations. *Evolutionary Ecology*, 9(6), 633–661.

Howard, R. D., Andrew DeWoody, J., & Muir, W. M. (2004). Transgenic male mating advantage provides opportunity for Trojan gene effect in a fish. *Proceedings of the National Academy of Sciences*, 101(9), 2934–2938.

Ito, H. C., & Dieckmann, U. (2007). A new mechanism for recurrent adaptive radiations. *The American Naturalist*, 170(4), E96–E111.

Ito, H. C., & Dieckmann, U. (2014). Evolutionary branching under slow directional evolution. *Journal of Theoretical Biology*, 360, 290–314.

Jian, Z., Wang, J., & Huang, G. (2016). Evolutionary diversification of prey and predator species facilitated by asymmetric interactions. *PLoS One*, 11(9), e0163753.

Johansson, J., & Dieckmann, U. (2009). Evolutionary responses of communities to extinctions. *Evolutionary Ecology Research*, 11, 561–588.

Kisdi, E., Jacobs, F. J. A., & Geritz, S. A. H. (2002). Red queen evolution by cycles of evolutionary branching and extinction. *Selection*, 2(1–2), 161–176.

Lehtinen, S. O., & Geritz, S. A. H. (2019). Coevolution of cannibalistic predators and timid prey: Evolutionary cycling and branching. *Journal of Theoretical Biology*, 483, 110001.

Leimar, O. (2002). Evolutionary change and darwinian demons. *Selection*, 2(1–2), 65–72.

Lion, S. (2018). Theoretical approaches in evolutionary ecology: Environmental feedback as a unifying perspective. *The American Naturalist*, 191(1), 21–44.

Loeuille, N., & Hauzy, C. (2018). Multidimensionality of plant defenses and herbivore niches: Implications for eco-evolutionary dynamics. *Journal of Theoretical Biology*, 445, 110–119.

Maor, R., & Shirasu, K. (2005). The arms race continues: battle strategies between plants and fungal pathogens. *Current Opinion in Microbiology*, 8(4), 399–404.

Marrow, P., Dieckmann, U., & Law, R. (1996). Evolutionary dynamics of predator-prey systems: An ecological perspective. *Journal of Mathematical Biology*, 34(5–6), 556–578.

Marrow, P., Law, R., & Cannings, C. (1992). The coevolution of predator|prey interactions: ESSs and Red Queen dynamics. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 250(1328), 133–141.

Matsuda, H., & Abrams, P. A. (1994). Timid consumers: Self-extinction due to adaptive change in foraging and anti-predator effort. *Theoretical Population Biology*, 45(1), 76–91.

Matsuda, H., & Abrams, P. A. (1994). Runaway evolution to self-extinction under asymmetrical competition. *Evolution*, 48(6), 1764–1772.

May, R. M. (1972). Will a large complex system be stable? *Nature*, 238(5364), 413–414.

Muir, W. M., & Howard, R. D. (1999). Possible ecological risks of transgenic organism release when transgenes affect mating success: Sexual selection and the Trojan gene hypothesis. *Proceedings of the National Academy of Sciences*, 96(24), 13853–13856.

Nguyen, V. A T, & Vural, D. C. (2020). Theoretical guidelines for editing ecological communities. *arXiv*, preprint:.

O'Dwyer, J. P. (2018). Whence lotka-volterra? *Theoretical Ecology*, 11(4), 441–452.

Olsen, E. M., Heino, M., Lilly, G. R., Morgan, M. J., Brattey, J., Ernande, B., & Dieckmann, U. (2004). Maturation trends indicative of rapid evolution preceded the collapse of northern cod. *Nature*, 428(6986), 932–935.

Parvinen, K. (2005). Evolutionary suicide. *Acta Biotheoretica*, 53(3), 241–264.

Parvinen, K. (2010). Adaptive dynamics of cooperation may prevent the coexistence of defectors and cooperators and even cause extinction. *Proceedings of the Royal Society B: Biological Sciences*, 277(1693), 2493–2501.

Parvinen, K., & Dieckmann, U. (2013). Self-extinction through optimizing selection. *Journal of Theoretical Biology*, 333, 1–9.

Rankin, D. J., & López-Sepulcre, A. (2005). Can adaptation lead to extinction? *Oikos*, 111(3), 616–619.

Rosenzweig, M. L., Brown, J. S., & Vincent, T. L. (1987). Red Queens and ESS: The coevolution of evolutionary rates. *Evolutionary Ecology*, 1(1), 59–94.

Rosenzweig, M. L., & Schaffer, W. M. (1978). Homage to the red queen. II. Coevolutionary response to enrichment of exploitation ecosystems. *Theoretical Population Biology*, 14(1), 158–163.

Schaffer, W. M., & Rosenzweig, M. L. (1978). Homage to the red queen. I. Coevolution of predators and their victims. *Theoretical Population Biology*, 14(1), 135–157.

Stahl, E. A., & Bishop, J. G. (2000). Plant–pathogen arms races at the molecular level. *Current Opinion in Plant Biology*, 3(4), 299–304.

Turchin, P. (2003). *Complex population dynamics: a theoretical/empirical synthesis* (Vol. 35). Princeton University Press.

Vural, D. C., Isakov, A., & Mahadevan, L. (2015). The organization and control of an evolving interdependent population. *Journal of the Royal Society Interface*, 12(108), 20150044.

Weitz, J. S., Hartman, H., & Levin, S. A. (2005). Coevolutionary arms races between bacteria and bacteriophage. *Proceedings of the National*

Academy of Sciences, 102(27), 9535–9540. https://doi.org/10.1073/pnas.0504062102

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

APPENDIX A

IMPLICIT ADAPTIVE DYNAMICS−ABUNDANCES

For a time-dependent interaction matrix, $\mathbf{A}(t)$, we wish to find a time-dependent form for the equilibrium population abundances, $\bar{x}(t)$, through the fixed-point constraint $\mathbf{A}\bar{x} = -\vec{r}$ given the intrinsic birth/death rates $\vec{r}$. Then by expanding the time-dependencies as a polynomial with respect to time, t, we can then write

$$\bar{x}(t) = \bar{\varepsilon}_0 + \bar{\varepsilon}_1 t + \bar{\varepsilon}_2 t^2 + \ldots$$

and assume that the form of $\mathbf{A}(t)$ is also given as

$$\mathbf{A}(t) = \mathbf{H}_0 + \mathbf{H}_1 t + \mathbf{H}_2 t^2 + \ldots \quad \text{(A1)}$$

where all coefficients $(\mathbf{H}, \bar{\varepsilon})$ are constant with respect to time. This then gives

$$\left(\mathbf{H}_0 + \mathbf{H}_1 t + \mathbf{H}_2 t^2 + \ldots\right)\left(\bar{\varepsilon}_0 + \bar{\varepsilon}_1 t + \bar{\varepsilon}_2 t^2 + \ldots\right) = -\vec{r}$$

equating each power of t gives

$$\mathbf{H}_0 \bar{\varepsilon}_0 = -\vec{r}$$
$$\mathbf{H}_0 \bar{\varepsilon}_1 + \mathbf{H}_1 \bar{\varepsilon}_0 = 0$$
$$\mathbf{H}_0 \bar{\varepsilon}_2 + \mathbf{H}_1 \bar{\varepsilon}_1 + \mathbf{H}_2 \bar{\varepsilon}_0 = 0$$
$$\vdots$$
$$\sum_{k=0}^{m} \mathbf{H}_k \bar{\varepsilon}_{m-k} = 0$$

where $m > 0$. The zeroth order gives our starting point by definition $\mathbf{H}_0 = \mathbf{A}(t = 0)$ and $\bar{\varepsilon}_0 = \bar{x}(t = 0)$. We then sequentially calculate the higher order coefficients for the population abundances as

$$\bar{\varepsilon}_1 = -\mathbf{H}_0^{-1}\mathbf{H}_1 \bar{\varepsilon}_0$$
$$\bar{\varepsilon}_2 = -\mathbf{H}_0^{-1}\left(\mathbf{H}_1 \bar{\varepsilon}_1 + \mathbf{H}_2 \bar{\varepsilon}_0\right)$$
$$\vdots$$
$$\bar{\varepsilon}_m = -\mathbf{H}_0^{-1}\sum_{k=1}^{m} \mathbf{H}_k \bar{\varepsilon}_{m-k}$$

where each coefficient only depends on terms of a lower order. This will also work for population dependent mutation rates as long as we can write $\mathbf{H}_m$ using terms up to order $m - 1$.

For example, if the interactions are only linearly changing with time then all higher orders are removed, $\mathbf{H}_2 = \mathbf{H}_3 = \cdots = 0$, and the coefficients can be simplified to

$$\bar{\varepsilon}_1 = -\mathbf{H}_0^{-1}\mathbf{H}_1 \bar{\varepsilon}_0$$
$$\bar{\varepsilon}_2 = -\mathbf{H}_0^{-1}\mathbf{H}_1 \bar{\varepsilon}_1 = (-\mathbf{H}_0^{-1}\mathbf{H}_1)^2 \bar{\varepsilon}_0$$
$$\vdots$$
$$\bar{\varepsilon}_m = (-\mathbf{H}_0^{-1}\mathbf{H}_1)^m \bar{\varepsilon}_0. \quad \text{(A2)}$$

APPENDIX B

IMPLICIT ADAPTIVE DYNAMICS−TIME TO EXTINCTION

In this section, we will develop a scheme to easily estimate the extinction time for each species as we iteratively increase the accuracy of the polynomial expansion of $\bar{x}(t)$. Given that $\mathbf{A}(t) = \mathbf{A}_0 + \mathbf{H}_1 t + \mathbf{H}_2 t^2 + \ldots$ we can track the population abundances over time using the coefficients given by Equation A2. This gives

$$\bar{x}(t) = \bar{x}_0 + \bar{\varepsilon}_1 t + \bar{\varepsilon}_2 t^2 + \ldots$$

which can be solved to find the extinction time for each species $x_i(\tau) = 0$ and pick the earliest extinction time. However, this requires that the series be truncated at some order t^m and the calculated extinction time will vary slightly as the order is changed. We wish to quickly estimate the extinction time using a *mth* order polynomial by iteratively calculating the extinction time $\{\tau_1, \tau_2, \ldots, \tau_m\}$ as we increase the time order m. As we increase the order we should quickly converge to the true extinction time by relying on our previous estimates without needing to solve for the roots of a high order polynomial. The basic idea is to convert the equations from $0 = [\bar{x}]_i(\tau)$ to

$$0 = [\bar{x}_0]_i + [\bar{\varepsilon}_1]_i \tau_m + [\bar{\varepsilon}_2]_i \tau_m \tau_{m-1} + \ldots + [\bar{\varepsilon}_m]_i \tau_m \tau_{m-1}^{m-1} \quad \text{(B1)}$$

by replacing $\tau^k = \tau_m \tau_{m-1}^{k-1}$ such that the highest order extinction time τ_m appears in every power of time and can be easily solved for by factoring it out for each species i to solve for the earliest positive extinction time. This approximation should then converge as $\tau_m \to \tau$.

For a community of N species, this system of equations has N possible extinction times. There will be an extinction time for each species, however, there cannot be multiple extinction times for the community. Instead, the species with the earliest extinction event will die and the community structure will rearrange after its death. So as we are iterating this method we must pick for this specific extinction time. In other words, given the N possible extinction times returned by the iteration, we will pick the earliest *positive* extinction time.

For $m = 1$, we have for the *i*th species

$$0 = [\bar{x}_0]_i + [\bar{\varepsilon}_1]_i \tau_{1i}$$

to get

$$\tau_{1i} = \frac{-[\overline{x}_0]_i}{[\overline{\varepsilon}_1]_i}. \tag{B2}$$

where we discard any negative extinction times and pick the earliest positive time, $\tau_1 = min_i \tau_{1i}$, as our first order approximation to the extinction time. Then for $m = 2$

$$0 = [\overline{x}_0]_i + [\overline{\varepsilon}_1]_i \tau_{2i} + [\overline{\varepsilon}_2]_i \tau_{2i} \tau_1$$

to get

$$\tau_{2i} = \frac{-[\overline{x}_0]_i}{[\overline{\varepsilon}_1]_i + [\overline{\varepsilon}_2]_i \tau_1} \tag{B3}$$

where we plug-in the result for τ_4, and then again discard the negative extinction times and pick the earliest positive, $\tau_2 = min_i \tau_{2i}$. This procedure can be iterated m times to get

$$\tau_{mi} = \frac{-[\overline{x}_0]_i}{[\overline{\varepsilon}_1]_i + \sum_{j=2}^{m} [\overline{\varepsilon}_j]_i \tau_{m-1}^{j-1}}. \tag{B4}$$

and applying our positive minimization at each step.

Assuming that the true extinction time is τ then each iteration of this method has an error δ_m such that $\tau_m = \tau + \delta_m$. Assuming that the errors are small relative to the true extinction time, $\delta_m/\tau = 1$ we can then estimate the error ratio δ_m/δ_{m-1} by using the approximation

$$\tau_m^k = (\tau + \delta_m)^k \approx \tau^k \left(1 + \frac{k\delta_m}{\tau}\right).$$

Then for τ_m and the extinction of the ith species we have

$$\tau + \delta_m = \frac{-[\overline{x}_0]_i}{[\overline{\varepsilon}_1]_i + \sum_{j=2}^{m} [\overline{\varepsilon}_j]_i (\tau + \delta_{m-1})^{j-1}} \approx \frac{-[\overline{x}_0]_i}{[\overline{\varepsilon}_1]_i + \sum_{j=2}^{m} [\overline{\varepsilon}_j]_i \left(\tau^{j-1} + (j-1)\delta_{m-1}\tau^{j-2}\right)} \tag{B5}$$

Now we assume that the expansion is sufficiently high enough order to closely approximate the true extinction time such that

$$0 \approx [\overline{x}_0]_i + [\overline{\varepsilon}_1]_i \tau + [\overline{\varepsilon}_2]_i \tau^2 + \ldots + [\overline{\varepsilon}_m]_i \tau^m$$

We can then multiply Equation B5 on the top and bottom with τ to simplify to

$$\tau + \delta_m \approx \frac{-[\overline{x}_0]_i \tau}{-[\overline{x}_0]_i + \sum_{j=2}^{m} (j-1)[\overline{\varepsilon}_j]_i \tau^{j-1} \delta_{m-1}}.$$

Now assuming the residue terms are small relative to the initial population sizes we can then again apply the binomial approximation to get

$$\tau + \delta_m \approx \tau \left(1 + \frac{1}{[\overline{x}_0]_i} \sum_{j=2}^{m} (j-1)[\overline{\varepsilon}_j]_i \tau^{j-1} \delta_{m-1}\right)$$

so the convergence ratio is given by

$$\left|\frac{\delta_m}{\delta_{m-1}}\right| \approx \frac{1}{[\overline{x}_0]_i} \left|\sum_{j=2}^{m} (j-1)[\overline{\varepsilon}_j]_i \tau^j\right|. \tag{B6}$$

This can be evaluated before hand for each species to determine the convergence of their extinction times. For our system, the coefficients ε are very small relative to the initial population abundances which allow this method to rapidly converge.

We tested the convergence of this scheme for $N = 20$ and $N = 50$ by varying the parameters used to generate an adaptive dynamics given by $\mathbf{A}(t) = \mathbf{A}_{(0)} + \mathbf{H}_1 t$ for interaction densities $\rho = [0.3, 0.5, 1.0]$ and a wide range of random constant $\mathbf{H}_1$ with different rate scales $\mu = [10^{-4}, 10^{-3},\ldots,1]$. Figure 6 shows how well this scheme convergences to the true extinction time, τ, given by numerically evaluating Equation 10 based on starting with the zeroth order or the second order estimate of the earliest extinction time. We then compare the relative error, $G_m = (\tau - \tau_m)/\tau$, for each iteration m. This result shows good convergence even when we start with $\tau_0 = 0$.

However, there are several limitations involved in this calculation. First, there is no guarantee that any species will go extinct. For small ecosystems, it is possible that all returned coefficients are positive such that all species are growing to infinity. Second, because we are dealing with only a single expansion around the initial starting time this requires that extinction must occur within the domain of convergence of our polynomial series given in Equation 5. This is, in general, not true, especially for more complex dynamics involving higher orders of H_m. Overall, the best method to determine the extinct time involves iterating our full scheme as shown in Figure 5 when neither of the above conditions are reliably satisfied.

Unfortunately, there is no way to mathematically isolate this condition. However, we have shown that when these conditions are meant, our scheme provides a good estimate of the extinction time. For example, truncating the polynomial series at second order can reasonably estimate the extinction time within 1%–10% error on average.

APPENDIX C

Species abundances for the special case of $\mathbf{A} = \mathbf{A}_0 + \mathbf{H}_1 t$

For an initial n-species community described by $(\mathbf{A}, \vec{r})$ and the mean-field approximation to the changes of the interactions $\mathbf{H}_t$, we have that the fixed point of the system at some time t can be calculated as $\overline{x}(t) = -(\mathbf{A}_0 + \mathbf{H}t)^{-1}\vec{r} = \left(\mathbf{I} + \mathbf{A}_0^{-1}\mathbf{H}t\right)\overline{x}_0$ where $\overline{x}_0$ is the initial fixed point of the system. Using the implicit form $\left(\mathbf{I} + \mathbf{A}_0^{-1}\mathbf{H}t\right)\overline{x}(t) = \overline{x}_0$ and if the initial and time-dependent fixed points can be written as a linear combination of the projections on the eigenvectors of the operator $\mathbf{U} = -\mathbf{A}_0^{-1}\mathbf{H}$ we then have that

$$\sum_{j=1}^{n} c_j(t)\bar{v}_j - \lambda_j t c_j(t)\bar{v}_j = \sum_{j=1}^{n} c_j(0)\bar{v}_j$$

where $\bar{x}_0 = \sum_{j=1}^{N} c_j(0)\bar{v}_j$, $\bar{x}(t) = \sum_{j=1}^{N} c_j(t)\bar{v}_j$ and $\mathbf{U}\bar{v}_j = \lambda_j \bar{v}_j$. If the eigenvectors are linearly independent and normalized, we can solve for $c_j(t)$ to get

$$c_j(t) = \frac{c_j(0)}{1 - \lambda_j t}.$$

Finally, this allows us to write down the fixed point at some time t to be

$$\bar{x}(t) = \sum_{j=1}^{N} \frac{c_j(0)}{1 - \lambda_j t} \bar{v}_j.$$

Note that $c_j(0)$ is the length of $\bar{x}_0$ along the jth eigenvector. That is, $\bar{c}(0) = \mathbf{V}^{-1}\bar{x}_0$, where $\mathbf{V}$ is a matrix whose jth column is the jth eigenvector $\bar{v}_j$.

The possible extinction time, τ_i, for the i-th species is the earliest real, positive root of the n order polynomial, $x_i(\tau_i) = 0$, and if this root does not exist then the i-th species does not go extinct in the current community. The *actual* extinction of the current community, $\tau = min_i \tau_i$, is the earliest extinction time of all the species and all later possible extinction times are invalid. This is because as soon as a species dies the community reduces in size and the operator $\mathbf{U}$ also changes. The additional constraints on the applicability of this result can be found in Appendix F.

APPENDIX D

Time to extinction for the special case of $\mathbf{A} = \mathbf{A}_0 + \mathbf{H}_1 t$

In this section, we will approximate the extinction time for communities which have an interaction matrix changing as $\mathbf{A}(t) = \mathbf{A}_0 + \mathbf{H}t$. Although the exact solution is given by Equation 10, it is nontrivial to calculate the extinction time. For a community composed of N species, we must solve a system of N equations where each equation is the root of a N order polynomial with respect to time t. Then, filter the resulting roots for a realistic extinction time. Additionally, we hope that writing down the expression for the extinction time will provide some insights in how community parameters such as interaction density, interaction strength, etc. can alter the rate of collapse for the community.

We will apply a perturbation scheme on the exact solution given by Equation 8 to approximate the extinction time of the community. First setting Equation 8 to zero to find the extinction time τ

$$0 = \bar{x}_0 + \mathbf{U}_1 \tau (\mathbf{I} - \mathbf{U}_1 \tau)^{-1} \bar{x}_0$$

gives us a system of equations used to determine the extinction time for each species. However, only the first species which goes extinct will realistically do so. The other species will rearrange themselves into a smaller community which we must solve separately to determine the next extinction time.

We apply a perturbation scheme in order to estimate the extinction time of the community. First, we substitute τ_m and τ_{m-1} for the two τ's,

$$0 = \bar{x}_0 + \mathbf{U}_1 \tau_m (\mathbf{I} - \mathbf{U}_1 \tau_{m-1})^{-1} \bar{x}_0. \tag{D1}$$

Then by repeatedly substituting in our guess of the extinction time into τ_{m-1} we can solve for a more accurate extinction time τ_m. Using the first guess $\tau_0 = 0$, we can find the extinction time ($m = 1$) for the ith species

$$\tau_{1i} = \frac{-[\bar{x}_0]_i}{[\mathbf{U}_1 \bar{x}_0]_i}. \tag{D2}$$

This term approximates the behaviour of $\bar{x}(t)$ near $t = 0$. If the species has a positive slope then the extinction time returned will be *negative* and if it appears that the population abundance is declining then the extinction time will be *positive*. However, the species with the earliest positive extinction time will die first which will cause the community to rearrange in its absence. Therefore, we must examine each extinction time, discarding the negative extinction times, and select only the earliest extinction time. We perform this positive minimization and define it as the first order approximation, $\tau_1 = min_i \tau_{1i}$. This minimization also specifies which species is going extinct in the approximation.

We can now plug this estimate into Equation D1 and solve for the next estimate ($m = 2$)

$$\tau_{2i} = \frac{-[\bar{x}_0]_i}{[\mathbf{U}_1(\mathbf{I} - \mathbf{U}_1 \tau_1)^{-1} \bar{x}_0]_i} \tag{D3}$$

This will again give N possible extinction times from which we discard the negative times and pick the earliest positive time, $\tau_2 = min_i \tau_{2i}$. This process can be repeated m times to find

$$\tau_{mi} = \frac{-[\bar{x}_0]_i}{[\mathbf{U}_1(\mathbf{I} - \mathbf{U}_1 \tau_{m-1})^{-1} \bar{x}_0]_i} \tag{D4}$$

where a positive minimization is required after each step and tells us which species is likely to die.

If the scheme converges to the true extinction time $[\bar{x}]_i(\tau) = 0$ for the extinction of species i then we would expect the error between each iteration to decrease in magnitude. We now ask how does the error, δ_m, in the mth estimate change as m increases.

First, we shift to the basis which diagonalizes $\mathbf{U}_1$. By defining the ith eigenvalues λ_i and normalized eigenvector $\bar{v}_i$ of $\mathbf{U}_1$ as

$$\mathbf{U}_1 \bar{v}_i = \lambda_i \bar{v}_i$$

we can find the projection of $\bar{x}_0$ on this basis as

$$\bar{x}_0 = \sum_{k=1}^{N} c_k \bar{v}_k$$

where the projection coefficients are given by

$$\vec{c} = \mathbf{V}^{-1}\vec{x}_0$$

and $\mathbf{V}$ is the matrix of eigenvectors with $\vec{v}_i$ in its ith column.

Then, substituting the error terms into Equation D4 in the new basis we have

$$\tau + \delta_m = \frac{-[\vec{x}_0]_i}{\sum_{k=1}^{N} \frac{\lambda_k}{1 - \lambda_k(\tau + \delta_{m-1})} c_k[\vec{v}_k]_i} \tag{D5}$$

where the term in the denominator is approximated as

$$\frac{\lambda_k}{1 - \lambda_k(\tau + \delta_{m-1})} = \frac{\lambda_k}{(1 - \lambda_k \tau)\left(1 - \frac{\lambda_k \delta_{m-1}}{1 - \lambda_k \tau}\right)} \approx \frac{\lambda_k}{1 - \lambda_k \tau}\left(1 + \frac{\lambda_k \delta_{m-1}}{1 - \lambda_k \tau}\right) \tag{D6}$$

assuming that $|\lambda_k \delta_{m-1}| = |1 - \lambda_k \tau|$. If we apply this assumption to all eigenvalues in the summation then we have must have $\delta_{m-1} = |1 - \lambda_{max}\tau|/|\lambda_{max}|$ where λ_{max} is the eigenvalue with the largest magnitude. If our iteration has a small enough error to the true extinction time we can then find

$$\tau + \delta_m \approx \frac{-[\vec{x}_0]_i}{\sum_{k=1}^{N} \frac{\lambda_k}{1 - \lambda_k \tau}\left(1 + \frac{\lambda_k \delta_{m-1}}{1 - \lambda_k \tau}\right) c_k[\vec{v}_k]_i}. \tag{D7}$$

Using the definition for the extinction time

$$0 = [\vec{x}_0]_i + \tau \sum_{k=1}^{N} \frac{\lambda_k}{1 - \lambda_k \tau} c_k[\vec{v}_k]_i$$

we can simplify the denominator to

$$\tau + \delta_m = \frac{-[\vec{x}_0]_i}{\frac{-[\vec{x}_0]_i}{\tau} + \delta_{m-1} \sum_{k=1}^{N} \frac{\lambda_k}{1 - \lambda_k \tau}\left(\frac{\lambda_k}{1 - \lambda_k \tau}\right) c_k[\vec{v}_k]_i} = \frac{\tau}{1 - \delta_{m-1} z} \tag{D8}$$

where $z = \frac{\tau}{[\vec{x}_0]_i} \sum_{k=1}^{N} \frac{\lambda_k}{1 - \lambda_k \tau}\left(\frac{\lambda_k}{1 - \lambda_k \tau}\right) c_k[\vec{v}_k]_i$ depends upon the true extinction time of the ith species. Assuming that $|\delta_{m-1}| = |z|$, we finally arrive at

$$\delta_m \approx \delta_{m-1} \tau z \tag{D9}$$

so the ratio between the new and old errors

$$\left|\frac{\delta_m}{\delta_{m-1}}\right| = \left|\frac{\tau^2}{[\vec{x}_0]_i} \sum_{k=1}^{N} \frac{\lambda_k}{1 - \lambda_k \tau}\left(\frac{\lambda_k}{1 - \lambda_k \tau}\right) c_k[\vec{v}_k]_i\right|$$

will decrease as long as this ratio is less than unity. Although there are N possible extinction times to examine in each iteration, we always pick the earliest positive extinction time so there should only be one associated time 'error' for the convergence.

APPENDIX E

CANONICAL ADAPTIVE DYNAMICS

Here, we derive the $\mathbf{H}$ matrices for the special case of mutation driven AD using the canonical form of AD. We assume that an interaction function has been chosen and that the derivatives of each

interaction with respect to their dependent traits are known. We then use the framework of AD to construct the each order of $\mathbf{H}$ shown in Equation A1.

In standard adaptive dynamics for each species, we can assign a number of traits S_k for the kth species and calculate the total rate of change for the interactions based on all possible changes in the traits. Let $s_{k\alpha}$ denote the αth trait of species k and sum over all possible changes for the interaction A_{ij} to get

$$\frac{dA_{ij}}{dt} = \sum_{k=1}^{N} \sum_{\alpha=1}^{S_k} \frac{\partial A_{ij}}{\partial s_{k\alpha}} \frac{ds_{k\alpha}}{dt}$$

where the rate of change for each traits, $s_{k\alpha}$, is given by (Dieckmann & Law, 1996) as

$$\frac{ds_{k\alpha}}{dt} = \frac{1}{2} \mu_{k\alpha} \sigma_{k\alpha}^2 \frac{x_k}{\lambda} \frac{\partial f_k}{\partial s_{k\alpha}}$$

where $s_{k\alpha}$ is the αth trait of specie k, $\mu_{k\alpha}$ and $\sigma_{k\alpha}^2$ are the mutation rate and variance of mutation strength for $s_{k\alpha}$, λ is the abundance corresponding to a single individual so that x_k/λ represents the number of individuals, and f_k is the LV fitness given by $f_k = r_k + \sum_l A_{kl} x_l$. The interspecies interactions $A_{ij}(\vec{s}_i, \vec{s}_j, \ldots)$ are functions of some traits of species i, j, or more.

We now evaluate the slope of the fitness hill

$$\frac{\partial f_k}{\partial s_{k\alpha}} = \sum_{q=1}^{N} \left[\frac{\partial A_{kq}}{\partial s_{k\alpha}} x_q + A_{kq} \frac{\partial x_q}{\partial s_{k\alpha}}\right]. \tag{E1}$$

Note that f is the relative fitness of a new mutant as it first emerges in a wild-type population where every other individual is s. The first term quantifies the relative fitness of a mutant individual that interacts differently with others, while the second term quantifies the relative fitness of the newly emerged mutant due to a change it causes to the other species' abundances. Since the new mutant is very few in abundance compared to the wild type, the abundances of others will not depend on its trait value, and thus $dx_q/ds_{k\alpha} = 0$. To understand this point better with specific examples, the reader could refer to (Dieckmann & Law, 1996; Marrow et al., 1996) (in their notation $dx/ds' = 0$). Therefore, we have

$$\frac{ds_{k\alpha}}{dt} = \frac{1}{2} \mu_{k\alpha} \sigma_{k\alpha}^2 \frac{x_k}{\lambda} \sum_{q=1}^{N} \frac{\partial A_{kq}}{\partial s_{k\alpha}} x_q.$$

Finally, we can write down the rate of change for each pair of interactions as

$$\frac{dA_{ij}}{dt} = \frac{1}{2\lambda} \sum_{k=1}^{N} x_k \sum_{\alpha=1}^{S_k} \mu_{k\alpha} \sigma_{k\alpha}^2 \frac{\partial A_{ij}}{\partial s_{k\alpha}} \sum_{q=1}^{N} \frac{\partial A_{kq}}{\partial s_{k\alpha}} x_q. \tag{E2}$$

The interaction derivatives can be any general function of time, t. In order write $d\mathbf{A}/dt$ as shown in Equation A1 we must now assume that each derivative is approximately constant

—WILEY

$$\frac{\partial A_{ij}}{\partial s_{k\alpha}} = C_{ijk\alpha}.$$

Although we take this simplifying assumption, it is not completely necessary. As long as $\frac{\partial A_{ij}}{\partial s_j}$ can be written as a power series of t then this method will still work but will require updating to fully match the constraints of Equation A1. A fully general prescription for this derivation is outside the scope of this Appendix.

We now seek to break down, term by term, the time dependencies of each expression shown in Equation E2. First, using $[\bar{x}]_q = [\bar{\varepsilon}_0]_q + [\bar{\varepsilon}_1]_q t + [\bar{\varepsilon}_2]_q t^2 + \dots$ we have

$$\sum_{q=1}^{N} C_{kqk\alpha}[\bar{x}]_q = \sum_{q=1}^{N} C_{kqk\alpha}\left([\bar{\varepsilon}_0]_q + [\bar{\varepsilon}_1]_q t + [\bar{\varepsilon}_2]_q t^2 + \dots\right).$$

Then

$$x_k \sum_{q=1}^{N} C_{kqk\alpha} x_q$$

$$= \sum_{q=1}^{N} C_{kqk\alpha}\left([\bar{\varepsilon}_0]_k + [\bar{\varepsilon}_1]_k t + \dots\right)\left([\bar{\varepsilon}_0]_q + [\bar{\varepsilon}_1]_q t + \dots\right)$$

$$= \sum_{m=0}^{\infty}\left(\sum_{s=0}^{m}[\bar{\varepsilon}_s]_k \sum_{q=1}^{N} C_{kqk\alpha}[\bar{\varepsilon}_{m-s}]_q\right) t^m$$

We then arrive at

$$\frac{dA_{ij}}{dt} = \sum_{m=0}^{\infty}\sum_{k=1}^{N}\sum_{\alpha=1}^{S_k}\frac{\mu_{k\alpha}\sigma_{k\alpha}^2}{2\lambda}C_{ijk\alpha} \times \left(\sum_{s=0}^{m}[\bar{\varepsilon}_s]_k \sum_{q=1}^{N} C_{kqk\alpha}[\bar{\varepsilon}_{m-s}]_q\right) t^m.$$

This arranges the summation with respect to their power of t for dA/dt. To find the expression for $\mathbf{A}(t)$ we finally *integrate* over time to find the $m > 0$ order of $\mathbf{H_m}$ in Equation A1 as

$$[\mathbf{H_m}]_{ij} = \sum_{q,k=1}^{N}\sum_{\alpha=1}^{S_k}\frac{\mu_{k\alpha}\sigma_{k\alpha}^2}{2m\lambda}C_{ijk\alpha}\sum_{s=0}^{m-1}[\bar{\varepsilon}_s]_k C_{kqk\alpha}[\bar{\varepsilon}_{m-s-1}]_q. \quad (E3)$$

Then using $\bar{x}(t)$ generated we can then predict the behaviour of the traits by integrating ds/dt over time. Given the initial trait values as $\bar{s}_0 = \bar{s}(t=0)$ we then have

$$s_{k\beta}(t) = s_{k\beta}(0) + \frac{\mu_{k\beta}\sigma_{k\beta}^2}{2\lambda}\sum_{\alpha=1}^{S_k}\left[[\bar{x}_0]_k \sum_{q=1}^{N} C_{kqk\alpha}[\bar{x}_0]_q t + \left([\bar{x}_0]_k \sum_{q=1}^{N} C_{kqk\alpha}[\bar{\varepsilon}_1]_q + [\bar{\varepsilon}_1]_k \sum_{q=1}^{N} C_{kqk\alpha}[\bar{x}_0]_q\right)\frac{t^2}{2} + \dots\right] \quad (E4)$$

where $s_{k\beta}$ is the β-th trait for species k.

APPENDIX F

TECHNICAL DISCUSSION ON THE CONVERGENCE AND RATE OF CONVERGENCE OF SERIES SOLUTIONS

We now demonstrate why repeated iterations of perturbation theory are required in order to examine the abundances for longer times. Consider the constant rate model $\mathbf{A}(t) = \mathbf{A}_{(0)} + \mathbf{H}_1 t$ for which we have both the analytical solution using the diagonal basis of the operator $\mathbf{U} = -\mathbf{A}_{(0)}^{-1}\mathbf{H}$ and the perturbation scheme described in the previous section. We now use the results from the perturbation theory to derive the analytic solution. First, note that $\mathbf{U}$ is also used to get the coefficients in the perturbation theory. Then by projecting the initial population abundance vector into the eigenvectors of $\mathbf{U}$, we have that $\bar{\varepsilon}_0 = \sum_j c_j(0)\bar{v}_j$ where $\mathbf{U}\bar{v}_j = \lambda_j\bar{v}_j$. Therefore, the polynomial expansion from the previous section can be written as

$$\bar{x}(t) = \sum_j\left(1 + \lambda_j t + \lambda_j^2 t^2 + \dots\right)c_j(0)\bar{v}_j = \sum_j c_j(0)\bar{v}_j \sum_{m=0}^{\infty}(\lambda_j t)^m = \sum_j \frac{c_j(0)\bar{v}_j}{1 - \lambda_j t}$$

which is the same result as the analytic method. However, there are two restrictions for the valid time domain of this solution. First, the solution can be divergent for certain values of λ. Second, when condensing the geometric series (polynomial expansion) from perturbation theory to the analytic expression we have implicitly assumed that we are within the convergence region of the geometric series. Therefore the spectral radius, $|\lambda|_{max}$, of the operator $\mathbf{U}$ restricts the convergent domain of $t < \frac{1}{|\lambda|_{max}}$.

The polynomial series from perturbation theory is more restrictive than the analytic result from using eigenvalues due to the convergence condition. Even if there are no singularities for positive (real) times the result from perturbation theory only converges for a time interval restricted by the magnitude of the largest eigenvalue. This restriction is always applied because we are approximating the analytic result with a series which in application is truncated. This restriction applies to any form of dA/dt because the perturbation theory approximates the operation of applying a matrix inversion continuously over time.

Typically, the expansion will blow up to infinity for long times and can alternate between positive and negative infinity as the truncated order is increased. Therefore, we can use a heuristic estimate to determine a good time step as

$$T = \left(\frac{\chi}{max_i\left|\varepsilon_i^{(m)}\right|}\right)^{-m} \quad (F1)$$

where χ is the allowed error in the approximation for $\bar{x}(t)$ and assuming that a sufficiently accurate order m has been chosen for

$\bar{x}(t) = \bar{x}_0 + \bar{\varepsilon}_1 t + \dots + \bar{\varepsilon}_m t^m$. This is because the highest order term serves as an estimation for the error in our approximation of $\bar{x}(t)$ when you compare two power series for $\bar{x}(t)$ of different truncation orders $m-1$ and m.