CONTINUUM APPROXIMATION OF INVASION PROBABILITIES

REBECCA K. BORCHERING† AND SCOTT A. MCKINLEY‡

Abstract. In the last decade there has been growing criticism of the use of Stochastic Differential Equations (SDEs) to approximate discrete state-space, continuous-time Markov chain population models. In particular, several authors have demonstrated the failure of Diffusion Approximation, as it is often called, to approximate expected extinction times for populations that start in a quasi-stationary state.

In this work we investigate a related, but distinct, population dynamics property for which Diffusion Approximation fails: invasion probabilities. We consider the situation in which a few individual are introduced into a population and ask whether their collective lineage can successfully invade. Because the population count is so small during the critical period of success or failure, the process is intrinsically stochastic and discrete. In addition to demonstrating how and why the Diffusion Approximation fails in the large population limit, we contrast this analysis with that of a sometimes more successful alternative WKB-like approach. Through numerical investigations, we also study how these approximations perform in an important intermediate regime. In a surprise, we find that there are times when the Diffusion Approximation performs well: particularly when parameters are near-critical and the population size is small to intermediate.

Key words. invasion probability, stochastic population model, diffusion approximation

1. Introduction. Invasion events are fundamental in population biology. In the study of disease there are interesting examples both at the cellular and whole-organism level: while studying the onset or recovery of an infection, one might look at the probability that a single virion can infect a target cell and proliferate [9]; in epidemiology, the goal might be to estimate the probability that a newly introduced pathogen will become endemic in a naive host population [4]. The same multi-scale interest in invasions appears in the study of population genetics: at the multi-organism scale one studies the probability that a novel allele can fix in a population, possibly affecting the population’s overall fitness [17, 21, 25]; at the cellular level, the invasion of a mutation in a stem cell population has been studied as an important first step in certain cancers [27]. What distinguishes the dynamics at different scales are the state-dependent rates of events. While multi-organism models tend to feature movement and direct interaction among individuals [7], cellular dynamics may be more spatially static with indirect interactions (competition for resources, or signaling at a distance, for example) [29].

It is natural to model these population dynamics using continuous-time, discrete-state-space Markov chains. It is easy to encode nonlinear interactions through state-dependent transition rates and there are numerous straightforward simulation techniques available that can be exact, but costly [13], or inexact (with respect to boundary interactions), but efficient [3]. It is also straightforward to write down difference equations whose solutions represent the probability of invasion from a small number of individuals, or the mean extinction time starting from a large population size. Often
it is even possible to find exact solutions for these systems of difference equations, but, importantly, it is difficult to interpret how these exact solutions depend on the parameters of the original system.

To overcome both computational and analytical challenges, it has become common to use Stochastic Differential Equations (SDEs) as a continuous-state-space approximation for the discrete-state-space Markov chains. Using Kurtz’s Diffusion Approximation Theorem as a guide [19, 2], there is a natural way to translate Markov chain transition rates into the drift and diffusion terms of an SDE. (See Section 3.1 for discussion.) This technique has been used prominently to model epidemics [1], neuronal activity [14, 10] and branching processes with logistic growth limits [20]. In the SDE setting, solutions to invasion probability and mean extinction time questions can be obtained by solving certain ODEs (or PDEs, depending on the dimension). These ODEs can also be obtained without appealing to an SDE model by applying a naïve Taylor series approximation of the Kolmogorov Equations of the Markov chain model. (See Section 3.1 again for discussion.) However, multiple authors have shown that there are important differences between Diffusion Approximations and their corresponding CTMCs. For example, Doering and co-authors [11, 12] have shown that the mean extinction time for SDE models can differ substantially from that of the corresponding Markov chain model. Newby and Bressloff also documented the issue for neuronal models [16, 22, 8].

While this issue of the mean extinction time has received considerable recent attention, the probability of invasion question was studied, and in many ways resolved, much earlier. Frank Ball and colleagues [4, 5] showed that the initial phase of a certain class of epidemic models can be well-approximated by branching processes for which an exact probability of invasion can be calculated. For example, in the large population limit, the probability that a single infectious individual will cause an epidemic in the stochastic Susceptible-Infectious-Susceptible (SIS) and stochastic Susceptible-Infectious-Recovered (SIR) models is $1 - d/b$, where $b$ is the rate at which an infectious individual infects susceptible members of the population and $d$ is the rate at which the pathogen is cleared from an infectious individual. (Throughout this work, we will use the letters $b$ and $d$ to evoke the words “birth” and “death” to correspond to population increases and decreases, regardless of the particular application.) This formula is appealing because it is trivial to understand the impact of model parameters, but we note that because this method relies on an infinite population size limit, it accordingly gives no guidance on how invasion probabilities scale with finite population sizes. Moreover, the Branching Process Approximation is only valid when the dynamics are “asymptotically linear” in a sense that will described after Assumption 1.1 below. Heuristically, this means that a “birth event” only depends on one member of the population of interest. This is often not the case in chemical reactions (see Doering, 2007 [12] for an example) and a recent investigation into the impact of sensing and decision-making on animal movement has revealed non-linear encounter rates in predator-prey models [15].

The above considerations motivate the present work, in which we analyze the performance of standard continuum approximations for the probability of invasion in Markov chain models. The mathematical framework allows for nonlinear transition rate functions and provides a natural way to estimate probability of invasion when the population size $N$ is finite. We perform asymptotic analysis ($N \to \infty$) on the Diffusion Approximation and a WKB-like Exponential Approximation, showing that the former yields an incorrect value in essentially all cases. When valid, the Exponential Approximation appears to yield the correct large $N$ limit. Remarkably though,
when the population size is small or intermediate, and when the parameters of the problem form a system that is near-critical, we show numerically that the Diffusion Approximation captures invasion probabilities much better than the Exponential Approximation. Rather than declare one approximation or another the “winner” we think it would be most useful for practitioners to keep in mind the tradeoffs involved in choosing which method to implement in their own work.

1.1. Mathematical Framework. For a given population scale \( N \), consider a continuous-time Markov chain \( X_N(t) \) that takes its values in the set of integers \( n \in \mathbb{N} \). We consider the class of models with transitions of size one and the transition rates from the state \( n \) are given by

\[
\begin{align*}
n \to n + 1 & \text{ at rate } \lambda_N(n) \\
n \to n - 1 & \text{ at rate } \mu_N(n)
\end{align*}
\]

with \( \lambda_N(0) = \mu_N(0) = 0 \). While \( \lambda_N(n) = 0 \) is allowable for some \( n > 0 \), we require that the death rate satisfies \( \mu_N(n) > 0 \) for all \( n \in \mathbb{N} \). It follows that 0 is the only absorbing state.

**Assumption 1.1** (Transition rate shape functions). There exist functions \( \lambda : \mathbb{R}_+ \to \mathbb{R}_+ \) and \( \mu : \mathbb{R}_+ \to \mathbb{R}_+ \) such that

\[
\lambda_N(n) = N\lambda\left(\frac{n}{N}\right) \quad \text{and} \quad \mu_N(n) = N\mu\left(\frac{n}{N}\right).
\]

Furthermore,

i) there exists an \( A \in (0, \infty] \) such that \( \lambda \) and \( \mu \) are differentiable for all \( x \in (0, A) \);

ii) \( \lambda(0) = \mu(0) = 0; \lambda(x) \geq 0 \) for all \( x \in (0, \infty) \); \( \mu(x) > 0 \) for all \( x \in (0, \infty) \); and

iii) there exists a value \( x_* > 0 \) such that \( \lambda(x_*) = \mu(x_*) \) and for all \( x \in (0, x_*) \), either \( \lambda(x) > \mu(x) \) or \( \mu(x) < \lambda(x) \). In the former case, the system is called supercritical; in the latter, subcritical.

We will refer to \( x_* \) as the minimal rate-balanced point. When the system is discrete, rate-balanced points are not always precisely achieved. For a given system size \( N \), we define the discrete analogue \( N_* \) as follows.

**Definition 1.2.** For a given \( N \in \mathbb{N} \), \( N_* := \min\{n \in \mathbb{N} : \lambda_N(n) \leq \mu_N(n)\} \) for supercritical systems and \( N_* := \min\{n \in \mathbb{N} : \lambda_N(n) \geq \mu_N(n)\} \) for subcritical systems.

If the transition rate shape functions are differentiable at zero with \( \lambda'(0) > 0 \) and \( \mu'(0) > 0 \), then we call the system asymptotically linear.

**Example 1** (Stochastic epidemics). For a population size \( N \), and constants \( b > 0 \) and \( d > 0 \), the size of the infectious population in a stochastic (non-density dependent) Susceptible-Infectious-Susceptible system is defined by the transition rates

\[
\begin{align*}
n \to n + 1 & \text{ at rate } bn\frac{(N - n)}{N} \\
n \to n - 1 & \text{ at rate } dn
\end{align*}
\]

These dynamics are characterized by the rate shape functions

\[
\lambda(x) = \begin{cases} 
bn(1 - x), & x \in [0, 1] \\
0, & x > 1
\end{cases}
\quad \text{and} \quad 
\mu(x) = dx, \text{ for all } x \in [0, \infty).
\]
The minimal rate-balanced point $x_* = 1 - d/b$, and for the process $X_N$, $N_* = \lceil N(1 - d/b) \rceil$.

**Example 2** (Density-dependent mortality). In this case, the population scale $N$ plays the role of the carrying capacity of an environment. Again, let $d > 0$ and let $b > 0$ be the growth rate of the population when scarcity of resources is not a factor. Then we consider the dynamics set by the rate shape functions

$$
\lambda(x) = bx \quad \text{and} \quad \mu(x) = dx^2
$$

for $x \in [0, \infty)$. We note that this is the standard logistic growth model with the intrinsic growth rate parameter $b = r$ and death rate $d = r/K$ for a carrying capacity $K$. The resulting transition rates are similar to the Logistic Branching Process presented by Amaury Lambert (2005) [20]. The rate-balanced value for this system is $x_* = b/d$, and for the process $X_N$, we have $N_* = \lceil bN/d \rceil$.

**Example 3** (Resource-constrained birth). Let $b > 0$, $d > 0$ and $a > 0$. We define our dynamics according to the rate shape functions

$$
\lambda(x) = bxe^{-x/a} \quad \text{and} \quad \mu(x) = dx
$$

for $x \in [0, \infty)$. Note that $x_* = a \ln \left( \frac{b}{\beta} \right)$, while $N_* = \lceil Na \ln \left( \frac{b}{\beta} \right) \rceil$.

**Example 4** (General nonlinear single term model). Let $b, \beta, d, \delta > 0$ with $\beta \neq \delta$. We define our dynamics according to the rate shape functions

$$
\lambda(x) = bx^\beta \quad \text{and} \quad \mu(x) = dx^\delta
$$

for $x \in [0, \infty)$. If $\beta < \delta$ we say that $\lambda$ has the leading order. If $\beta > \delta$ we say that $\mu$ has the leading order. In this case, $x_* = (b/d)^{\beta-\delta}$, and for the process $X_N$, $N_* = \lceil (b/d)^{\beta-\delta} N \rceil$.

While conducting our asymptotic analysis, we will often prefer an alternative assumption for the transition rate functions that features explicit series expansions for the birth and the death rates.
CONTINUUM APPROXIMATION OF INVASION PROBABILITIES

ASSUMPTION 1.3 (Series representation for the rate shape functions). Let $\lambda$ and $\mu$ be as in Assumption 1.1. Additionally, we assume that there exist constants $b, \beta, d, \delta, > 0$ with $\{b_n\}_{n \in \mathbb{N}}, \{d_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}$, and an integer $m \in \mathbb{N}$, such that

$$
\lambda(x) = bx^{\beta} \left(1 + \sum_{n=1}^{\infty} b_n x^{n/m}\right) \text{ and } \mu(x) = dx^{\delta} \left(1 + \sum_{n=1}^{\infty} d_n x^{n/m}\right),
$$

for all $x \in [0, x_\ast]$. For simplicity in the analysis, we also assume that $\beta$ and $\delta$ are rational numbers and $m$ is chosen such that $\beta m$ and $\delta m$ are integer values.

1.2. Approximations for Invasion Probabilities. The conditions of Assumption 1.1 ensure that as $N \to \infty$, the stochastic processes $X_N(t)$ converge pathwise to an associated ODE. Kurtz [19] defines the sense of this convergence rigorously as follows: Let $\lambda$ and $\mu$ be locally Lipschitz functions. Define $Y_N(t) = \frac{1}{N} X_N(t)$, and suppose $Y_N(0) \to x_0$. Define $x(t)$ to be the solution to

$$
\dot{x} = \lambda(x) - \mu(x), \text{ with } x(0) = x_0.
$$

Then

$$
\lim_{N \to \infty} \sup_{s \in [0, t]} |Y_N(s) - x(s)| = 0, \text{ almost surely.}
$$

A visual demonstration of this result is shown in Figure 1 for the stochastic Susceptible-Infectious-Susceptible epidemic model (Example 1 above). Conditioned on not hitting zero, the solutions converge to the heteroclinic connection between zero and the minimal rate-balanced point $x_\ast$. In our case, zero is an unstable fixed point of the limiting ODE. Thus, there is an apparent contradiction since, although a solution with initial condition zero should be zero for all $t \geq 0$, the fraction of solutions that avoid extinction does not go to zero as $N \to \infty$. This conflict is resolved by noting that the theorem above only applies for a fixed time window. It is indeed true that for any fixed time $t$, $X_N(t) \to 0$ as $N \to \infty$. However, it is also true that, conditioned on successful invasion, the hitting time for $N_\ast$ (Definition 1.2) tends to infinity as well.

We call the probability of hitting $N_\ast$ before 0 the invasion probability and adopt the following notation.

DEFINITION 1.4. For the process $X_N(t)$, we define the hitting time $\tau_N$ as follows:

$$
\tau_N := \inf \{t > 0 : X_N(t) \in \{0, N_\ast\}\}.
$$

We define the associated extinction probabilities $q_N(k), k \in \{0, 1, \ldots, N_\ast\}$ by

$$
q_N(k) := P \{X_N(\tau_N) = 0 | X_N(0) = k\}.
$$

From this we define the asymptotic probability of invasion $p_{\text{invasion}}(k)$ (conditioned on $k$ individuals at time zero) to be

$$
p_{\text{invasion}}(k) = \lim_{N \to \infty} 1 - q_N(k).
$$

It is possible to obtain an exact solution for these invasion probabilities, and we provide a formula in Section 2. In 1983, Frank Ball [4] introduced a slightly different definition of an invasion probability (which he referred to as the probability of a true epidemic) and, in the asymptotically linear case, proved that the probability of a true
epidemic converges to the survival probability of an associated branching process. To define this associated branching process, consider the number of offspring that any given individual gives birth to before expiring. In the asymptotically linear case, the leading order terms of the birth and death rates are \( b \) and \( d \), respectively. Therefore, the number of births before the individual dies has a Geometric distribution with “success probability” \( d/(b+d) \).

**Definition 1.5 (Branching Process Approximation).** Suppose that the transition rate shape functions \( \lambda \) and \( \mu \) satisfy Assumptions 1.1 and 1.3 and that the leading order exponents satisfy \( \beta = \delta = 1 \). Define the discrete time stochastic process \( \{Z_n\}_{n \geq 0} \) to be \( Z_0 = k \) and for \( N \in \mathbb{N} \),

\[
Z_{n+1} = \sum_{j=1}^{Z_n} \xi_{n+1,j}
\]

where \( \xi_{i,j} \sim \text{Geom}(d/(b+d)) \), \((i,j) \in \mathbb{N}\) are independent and identically distributed.

Then, we define

\[
p_{\text{branch}}(k) := \mathbb{P} \left\{ \lim_{n \to \infty} Z_n = \infty \mid Z_0 = k \right\} = \begin{cases} 1 - \left( \frac{d}{b} \right)^k, & b \geq d \\ 0, & b < d \end{cases}
\]

Ball [4] proved that \( p_{\text{invasion}}(k) = p_{\text{branch}}(k) \) in this setting. What the Branching Process Approximation lacks is a clear interpretation in asymptotically nonlinear cases (e.g. Example 4) and a means to consider what the invasion probability is for finite \( N \). For these cases we consider the Diffusion Approximation and the Exponential Approximation. The derivations of these approximations are given in Sections 3 and 4, respectively. In what follows we introduce the small parameter \( \epsilon \), which corresponds to the inverse of the population scale parameter \( N \).

**Table 1**

Summary of asymptotic invasion probabilities (see Definition 1.4). *These results are only proved for the asymptotically linear case (i.e. when \( \beta = \delta = 1 \)). †We define the Exponential Approximation to be zero for subcritical regimes because the asymptotic limit tends to zero, but the approach to zero is through negative values.

| \( \beta \) \( \delta \) | \( \beta = \delta \) \( b \geq d \) | \( \beta = \delta \) \( b < d \) | \( \beta > \delta \) |
|---|---|---|---|
| Exact Solution* | \( 1 - \left( \frac{d}{b} \right)^k \) | 0 | |
| Branching Process* | \( 1 - \left( \frac{d}{b} \right)^k \) | 0 | |
| Diffusion Approx. | \( 1 - e^{-2k} \) | \( 1 - e^{-2k \left( \frac{b-d}{b+d} \right)} \) | 0 | 0 |
| Exponential Approx. † | 1 | \( 1 - \left( \frac{d}{b} \right)^k \) | 0 | 0 |

| Supercritical | Subcritical |
DEFINITION 1.6 (Diffusion Approximation). Let $\epsilon > 0$ be given. Then for $x \in (0,x_*)$, let $u_\epsilon(x)$ be the solution to the boundary value problem

$$
\mathcal{E}(\epsilon) \frac{1}{2} (\lambda(x) + \mu(x))u''_\epsilon(x) + (\lambda(x) - \mu(x))u'_\epsilon(x) = 0
$$

with $u_\epsilon(0) = 1$ and $u_\epsilon(x_*) = 0$.

Then, setting $\epsilon = 1/N$, we define the Diffusion Approximation for invasion probabilities to be the relationship

$$
p_{\text{diffusion}}(k) := 1 - u_\epsilon(k\epsilon) \approx 1 - q_N(k).
$$

DEFINITION 1.7 (Exponential Approximation). Let $\epsilon > 0$ be given. Then for $x \in (0,x_*)$, define $w_\epsilon(x)$ to be

$$
w_\epsilon(x) = \exp \left( -\frac{1}{\epsilon} \int_0^x \ln r_\epsilon(y) \, dy \right),
$$

where $r_\epsilon(x)$ satisfies

$$
\epsilon r'_\epsilon(x) \{ r_\epsilon(x) - (1 + \rho(x)) \} = r_\epsilon(x)(1 + \rho(x)) - \rho(x),
$$

with $\rho(x) = \lambda(x)/\mu(x)$. Then the Exponential Approximation for invasion probabilities is defined by the relationship

$$
p_{\text{exp}}(k) := 1 - w_\epsilon(k\epsilon) \approx 1 - q_N(k).
$$

In Table 1 we summarize the results in Sections 3 and 4. Consistent with asymptotic analyses of the Diffusion Approximation for corresponding extinction probabilities, we find that the Diffusion Approximation disagrees with the Branching Process Approximation in the large $N$ limit. Remarkably, when the leading orders of the transition rate shape functions $\lambda$ and $\mu$ do not match, the Diffusion Approximation reduces to two cases, ignoring all detailed information contained in the rate functions. In contrast, the Exponential Approximation does provide the right large $N$ limit. We note however, that the Exponential Approximation is only well-defined in the supercritical case.

In Section 5 we investigate the problem numerically and find mixed results. In the large $N$ limit, the Exponential Approximation agrees with the Markov chain model, even in the asymptotically nonlinear cases we consider. However, when the population size parameter $N$ is of small or intermediate size, say $N = 50$, it is common that the Diffusion Approximation more faithfully represents the Markov chain model, especially when the parameter set is near-critical.

2. Exact solution for invasion probabilities.

PROPOSITION 2.1. For fixed $N \in \mathbb{N}$, let $X_N(t)$ be a CTMC with transition rate shape functions $\lambda$ and $\mu$ as defined in Assumption 1.1. Let $\{q_N(n)\}_{n=0}^{N_*}$ be the extinction probabilities as defined in Definition 1.4. Then the collection $\{q_N(n)\}_{n=0}^{N_*}$ satisfies the system of difference equations

$$
\lambda_N(n)q_N(n+1) + \mu_N(n)q_N(n-1) - (\lambda_N(n) + \mu_N(n))q_N(n) = 0,
$$

with $q_N(0) = 1$ and $q_N(N_*) = 0$.  

\textit{Proof.} Though the proof appears in standard texts, we include it for completeness. Fix $N \in \mathbb{N}$ and let an initial condition $n \in \{1, \ldots, N_\ast - 1\}$ be given. Define $T := \inf\{t > 0 : X_N(t) \neq n\}$, the first time the process changes states. Then, applying the Strong Markov property, we have
\[
\mathbb{P}\{X_N(\tau_N) = 0 \mid X_N(0) = n\} = \mathbb{P}\{X_N(\tau_N) = 0 \mid X_N(0) = n\} \mathbb{P}\{X_N(T) = n + 1 \mid X_N(0) = n\} + \mathbb{P}\{X_N(\tau_N) = 0 \mid X_N(T) = n - 1 \} \mathbb{P}\{X_N(T) = n - 1 \mid X_N(0) = n\}
\]
\[
= \mathbb{P}\{X_N(\tau_N) = 0 \mid X_N(0) = n\} \frac{\lambda_N(n)}{\lambda_N(n) + \mu_N(n)} + \mathbb{P}\{X_N(\tau_N) = 0 \mid X_N(0) = n - 1\} \frac{\mu_N(n)}{\lambda_N(n) + \mu_N(n)}.
\]
Applying the $q_N(n)$ notation from Definition 1.4 and multiplying through by $\lambda_N(n) + \mu_N(n)$, we have (8), where $q_N(0) = \mathbb{P}\{X(\tau_N) = 0 \mid X_N(0) = 0\} = 1$ and $q_N(N_\ast) = 0$. \hfill \Box

The following argument for finding the exact solution is similar to one presented on birth-death processes in Norris [23], for example, and has a similar form to the exact solutions for mean extinction times presented by Doering et al [11, 12].

\textbf{Proposition 2.2 (Exact solution for extinction probability).} Let $X_N(t)$ be defined as in Proposition 2.1, with associated extinction probabilities $\{q_N(n)\}_{n=0}^{N_\ast}$. Then for $k \in \{1, 2, \ldots, N_\ast - 1\}$,

\[
q_N(k) = \prod_{m=1}^{k+1} \left( 1 - \left( 1 + \sum_{j=m}^{N_\ast} \frac{1}{\rho_N(j)} \right)^{-1} \right)
\]

where we define $\rho_N(0) = 1$ and for $j \geq 1$, $\rho_N(j) = \frac{\lambda_N(j)}{\mu_N(j)}$.

\textit{Proof.} Suppressing the dependence on $N$, we write

\[
q_n := q_N(n), \lambda_n := \lambda_N(n), \mu_n := \mu_N(n), \rho_n := \frac{\lambda_n}{\mu_n}.
\]

Rearranging (8), we have

\[
\lambda_n(q_{n+1} - q_n) + \mu_n(q_{n-1} - q_n) = 0.
\]

Introducing $\delta_n := q_{n-1} - q_n$, we can solve for $\delta_{n+1}$ and extrapolate to a dependence on $\delta_k$:

\[
\delta_{n+1} = \frac{\mu_n}{\lambda_n} \delta_n = \cdots = \frac{\mu_n \mu_{n-1} \cdots \mu_k}{\lambda_n \lambda_{n-1} \cdots \lambda_k} \delta_k.
\]

Recalling Definition 1.4, we have the boundary conditions $q_0 = 1$ and $q_{N_\ast} = 0$. It follows that

\[
1 = q_0 - q_{N_\ast} = \sum_{n=1}^{N_\ast} \delta_n = \sum_{n=1}^{k-1} \delta_n + \delta_k + \sum_{n=k+1}^{N_\ast} \left( \prod_{j=k}^{n-1} \frac{1}{\rho_j} \right) \delta_k.
\]

Noticing that $\sum_{n=1}^{k-1} \delta_n = 1 - q_{k-1}$, we have

\[
q_{k-1} = \delta_k \left( 1 + \sum_{n=k+1}^{N_\ast} \left( \prod_{j=k}^{n-1} \frac{1}{\rho_j} \right) \right)
\]
CONTINUUM APPROXIMATION OF INVASION PROBABILITIES

and solving for $\delta_k$

$$\delta_k = q_{k-1} \left( 1 + \sum_{n=k+1}^{N} \left( \frac{1}{\rho_j} \right)^{n-1} \right)$$

$$q_{k-1} - q_k = q_{k-1} \left( 1 + \sum_{n=k+1}^{N} \left( \frac{1}{\rho_j} \right)^{n-1} \right)$$

$$q_k = q_{k-1} \left( 1 - \left( 1 + \sum_{n=k+1}^{N} \left( \frac{1}{\rho_j} \right)^{n-1} \right) \right).$$

From this recursion we obtain

$$q_k = q_0 \left( 1 - \left( 1 + \sum_{n=1}^{N} \left( \frac{1}{\rho_j} \right)^{n-1} \right) \right) \cdots \left( 1 - \left( 1 + \sum_{n=k+1}^{N} \left( \frac{1}{\rho_j} \right)^{n-1} \right) \right)$$

and since $q_0 = 1$, the result follows.

3. Diffusion Approximation.

3.1. Motivations for the Diffusion Approximation. There are two points of view on the derivation of the Diffusion Approximation for invasion probabilities. They yield the same result. The first point of view (Motivation 1, below) is common in the physics literature [6]. It involves substituting a smooth function $u_\epsilon$ into the difference equation (8) and then converting it to a differential equation by writing out Taylor expansions and matching terms. In a second point of view (Motivation 2, below), we define a Stochastic Differential Equation whose infinitesimal first and second moments are defined to match that of the original birth-death chain in a sense discussed in [26] and [2]. Then the same Diffusion Approximation results from computing the probability that this SDE hits the rate-balanced state before going extinct, starting from the value $k\epsilon = k/N$, where $k$ is the number of initially introduced individuals.

Motivation 1: Taylor series approximation

To begin this approach, we assume that $u(x)$ is a smooth function and appeal to the following formal argument. We start with the difference equation

$$\lambda_n q_{n+1} - (\lambda_n + \mu_n) q_n + \mu_n q_{n-1} = 0$$

and rewrite it using the rate shape functions according to Assumption 1.1, $\lambda_n = N\lambda(n/N)$ and $\mu_n = N\mu(n/N)$, along with $q_n = u(n/N)$:

$$N\lambda \left( \frac{n}{N} \right) u \left( \frac{n+1}{N} \right) - N \left[ \lambda \left( \frac{n}{N} \right) + \mu \left( \frac{n}{N} \right) \right] u \left( \frac{n}{N} \right) + N\mu \left( \frac{n}{N} \right) u \left( \frac{n-1}{N} \right) = 0.$$  

Writing $x = n/N$ and $\epsilon = 1/N$,

$$\epsilon^{-1} \lambda(x) u(x + \epsilon) - \epsilon^{-1} [\lambda(x) + \mu(x)] u(x) + \epsilon^{-1} \mu(x) u(x - \epsilon) = 0.$$  

We then apply a Taylor series approximation to the $u(x + \epsilon)$ and $u(x - \epsilon)$ terms:

$$u(x + \epsilon) = u(x) + \epsilon u'(x) + \frac{\epsilon^2}{2} u''(x) + O(\epsilon^3)$$

$$u(x - \epsilon) = u(x) - \epsilon u'(x) + \frac{\epsilon^2}{2} u''(x) + O(\epsilon^3).$$
Substituting these into (9) and neglecting the higher order terms, we have

$$ (\lambda(x) - \mu(x))u'(x) + \frac{\epsilon}{2} (\lambda(x) + \mu(x))u''(x) = 0 $$

which is the form seen in Definition 1.6.

**Motivation 2: Hitting probabilities for an SDE approximation**

Let \( \epsilon = 1/N \) and \( y = n/N \), and consider the following calculation for the infinitesimal mean:

\[
\alpha_{\epsilon}(y) = \lim_{h \to 0} \frac{1}{h} \left( \mathbb{E}(X_N(t + h) - X_N(t) \mid X_N(t) = n) \right) \\
= \lim_{h \to 0} \frac{1}{h} \left( \frac{1}{N} \mathbb{P}\left\{ X(t + h) - X(t) = \frac{1}{N} \mid X_N(t) = n \right\} \\
- \frac{1}{N} \mathbb{P}\left\{ X(t + h) - X(t) = -\frac{1}{N} \mid X_N(t) = n \right\} + o(h) \right) \\
= \lim_{h \to 0} \frac{1}{h} \left( \frac{h}{N} \lambda_N(n) - \frac{h}{N} \mu_N(n) + o(h) \right) = \lambda(y) - \mu(y).
\]

In the last equality we used Assumption 1.1: \( \lambda_N(n) = N\lambda(n/N) \) and \( \mu_N(n) = N\mu(n/N) \). Note that the infinitesimal mean does not, in fact, depend on \( N \). On the other hand, a factor of \( 1/\sqrt{N} \) does appear in the calculation of the infinitesimal second moment:

\[
\sigma^2_{\epsilon}(y) = \lim_{h \to 0} \frac{1}{h} \left( \mathbb{E}\left( (X_N(t + h) - X_N(t))^2 \mid X_N(t) = n \right) \right) \\
= \lim_{h \to 0} \frac{1}{h} \left( \frac{h}{N^2} (\lambda_N(n) + \mu_N(n)) + o(h) \right) \\
= \frac{1}{N} (\lambda(y) + \mu(y)).
\]

Using these values, we define the SDE approximation as follows.

**Definition 3.1 (SDE Approximation).** Let \( \epsilon > 0 \). We say that a process \( Y_\epsilon(t) \) is an SDE approximation for the family of CTMCs defined with rate shape functions \( \lambda \) and \( \mu \) if it solves the Itô Stochastic Differential Equation:

\[
dY_\epsilon(t) = a(Y_\epsilon(t)) dt + \sqrt{\sigma(Y_\epsilon(t))} dW(t)
\]

where

\[
a(y) = \lambda(y) - \mu(y) \text{ and } \sigma^2(y) = \lambda(y) + \mu(y).
\]

Specifically, if \( \epsilon = 1/N \), then the SDE associated with the CTMC \( X_N(t) \) is \( Y_\epsilon(t) \).

**Proposition 3.2.** Let \( Y_\epsilon(t) \) be defined as in (10) and (11). Let \( T_\epsilon := \inf\{ t > 0 : Y_\epsilon(t) \in \{0, y_\epsilon\} \} \). Then \( u_\epsilon(x) := \mathbb{P}\{ Y_\epsilon(T_\epsilon) = 0 | Y_\epsilon(0) = x \} \) satisfies the boundary value problem (3).

**Proof.** The proof follows quickly from stochastic calculus. For a general continuous sample-path SDE with infinitesimal generator \( L \), the probability \( u(y) \) that the process hits a value \( a \) before \( b \) starting from \( y \in [a, b] \) satisfies the boundary value problem

\[
L_\epsilon u_\epsilon(y) = 0, \text{ for } y \in [a, b],
\]
with \( u(a) = 1 \) and \( u(b) = 0 \) [18].

It remains to note that the generator of \( Y_\epsilon(t) \) is

\[
L_\epsilon := a(y) \frac{d}{dy} + \epsilon \frac{\sigma^2(y)}{2} \frac{d^2}{dy^2}.
\]

### 3.2. Analysis of \( p_{\text{diffusion}}(k) \)

It is useful to introduce the functions

\[
\psi(x) := \frac{2(\lambda(x) - \mu(x))}{\lambda(x) + \mu(x)} \quad \text{and} \quad \Psi(x) := \int_0^x \psi(y) dy.
\]

We note that \( \psi(x) \) in (12) can be expressed exclusively in terms of the ratio \( \rho(x) := \lambda(x)/\mu(x) \) as

\[
\psi(x) := \frac{2(\rho(x) - 1)}{\rho(x) + 1}.
\]

However, in order to retain the ability to directly interpret results in terms of explicit birth and death rate parameters, we generally perform our analyses and present our results using the rate shape functions \( \lambda \) and \( \mu \) directly, as they are characterized in Assumption 1.1. An exception arises in Section 4, where we find that it is most natural to analyze the Exponential Approximation in terms of the ratio of rate functions \( \rho(x) \).

**Proposition 3.3.** Let \( u_\epsilon(x) \) be the Diffusion Approximation for the extinction probabilities, satisfying (3) of Definition 1.6. Then

\[
(13) \quad u_\epsilon(x) = 1 - \frac{\int_0^x e^{-\Psi(y)/\epsilon} dy}{\int_0^{x_*} e^{-\Psi(y)/\epsilon} dy},
\]

where \( x_* = \min\{x : \lambda(x) = \mu(x)\} \).

**Proof.** It we let \( h(x) = u_\epsilon'(x) \) then the ODE (3) becomes

\[
h'(x) + \frac{1}{\epsilon} \psi(x) h(x) = 0,
\]

which has general solution \( h(x) = C_1 e^{-\Psi(x)/\epsilon} \). Integrating to get \( u_\epsilon(x) = \int_0^x h(y) dy \) and enforcing the boundary conditions \( u_\epsilon(0) = 1 \) and \( u_\epsilon(x_*) = 1 \) yields the solution (13).

We note that a similar formula can be found in Pakmadan et al. [24].

We summarize the asymptotic behavior of the Diffusion Approximation as follows.

**Theorem 3.4.** Suppose that the transition rate shape functions \( \lambda \) and \( \mu \) each admit a series expansion as defined by Assumption 1.3. If the leading order exponents \( \beta \) and \( \delta \) are equal, then for \( k \in \mathbb{N} \),

\[
p_{\text{diffusion}}(k) = \begin{cases} 
1 - \exp \left( -\frac{2k(b-d)}{b+d} \right), & b > d; \\
0, & b \leq d.
\end{cases}
\]

Otherwise, if \( \beta \neq \delta \), then

\[
p_{\text{diffusion}}(k) = \begin{cases} 
1 - e^{-2k}, & \beta < \delta; \\
0, & \beta > \delta.
\end{cases}
\]
Remarkably, the Diffusion Approximation essentially ignores the detailed structure of the rate functions when they have different leading orders. The integrals that appear in (13) are in a form compatible with applying Watson’s Lemma [28]. In fact, Watson’s Lemma can be applied directly to the denominator. The subtlety in the present analysis is that, in the numerator, the upper limit of integration depends on $\epsilon$. To proceed, we follow Laplace’s method and introduce a substitution. Suppose first that we are in the supercritical case, as defined by Assumption 1.1, so that for all $x \in (0, x_*)$ we have that $\psi(x) > 0$. It follows that $\Psi(x)$ is increasing on that interval and therefore invertible. Letting $t = \Psi(y)$ we have

\begin{equation}
\int_0^x e^{-\Psi(y)/\epsilon} \, dy = \int_0^{\Psi(x)} f(t)e^{-t/\epsilon} \, dt
\end{equation}

where

\begin{equation}
f(t) := \frac{1}{\psi(\Psi^{-1}(t))} = \frac{\lambda(\Psi^{-1}(t)) + \mu(\Psi^{-1}(t))}{2(\lambda(\Psi^{-1}(t)) - \mu(\Psi^{-1}(t)))}.
\end{equation}

On the other hand, in the subcritical case, $\psi(x) < 0$ when $x \in (0, x_*)$. Introducing the notation $\psi_r(x) = -\psi(x)$ and $\Psi_r(x) = -\Psi(x)$, we substitute $t = \Psi_r(y)$ to attain

\begin{equation}
\int_0^x e^{-\Psi(y)/\epsilon} \, dy = \int_0^{\Psi_r(x)} f_r(t)e^{t/\epsilon} \, dt,
\end{equation}

noting that the exponential in the second integral has a positive exponent, where

\begin{equation}
f_r(t) := \frac{1}{\psi_r(\Psi_r^{-1}(t))} = \frac{\mu(\Psi_r^{-1}(t)) + \lambda(\Psi_r^{-1}(t))}{2(\mu(\Psi_r^{-1}(t)) - \lambda(\Psi_r^{-1}(t)))}.
\end{equation}

The proof follows from a combination of three lemmas. In Lemma 3.7 we assume that $f(t)$ admits a series expansion and then perform an asymptotic analysis on the ratios that appear in integrals of the form (14) and (16). In Lemmas 3.5 and 3.6 we establish that $\psi(x)$ and $f(t)$ admit series expansions when $\lambda(x)$ and $\mu(x)$ do, and provide the coefficients and powers of the first few terms. After presenting these lemmas and their proofs, we complete the proof of Theorem 3.4.

**Lemma 3.5.** Suppose that the transition rate shape functions $\lambda$ and $\mu$ satisfy Assumption 1.3. Then there exists an expansion for $\psi(x)$ of the form

\begin{equation}
\psi(x) = cx^{\kappa/m} \left(1 + \sum_{n=1}^{\infty} c_n x^{n/m} \right)
\end{equation}

where

\begin{equation}
c = \begin{cases} 
\min\{ n \geq 1 : b_n \neq d_n \} & \text{if } \beta = \delta \text{ and } b = d; \text{ and } \\
0, & \text{otherwise},
\end{cases}
\end{equation}

and the value of $c$ is given in the following table.

| $c$ | $\beta < \delta$ | $\beta = \delta$ | $\beta > \delta$ |
|-----|------------------|------------------|------------------|
| $b \neq d$ | 2 | $2(b - d)$ | $\frac{2(b - d)}{b + d}$ |
| $b = d$ | 2 | $(b_\kappa - d_\kappa)$ | $-2$ |
CONTINUUM APPROXIMATION OF INVASION PROBABILITIES

Proof. Substituting the expansions (1) in for \( \lambda \) and \( \mu \), we have

\[
\psi(x) = \frac{2(\lambda(x) - \mu(x))}{\lambda(x) + \mu(x)} \\
= \frac{2(bx^\beta(1 + \sum_{n=1}^{\infty} b_n x^{n/m}) - dx^\delta(1 + \sum_{n=1}^{\infty} d_n x^{n/m}))}{bx^\beta(1 + \sum_{n=1}^{\infty} b_n x^{n/m}) + dx^\delta(1 + \sum_{n=1}^{\infty} d_n x^{n/m})}.
\]

(20)

Case 1: \( \beta = \delta, b \neq d \). We can factor out \( x^\beta = x^\delta \) from the numerator and denominator. Then (20) simplifies to

\[
\psi(x) = \frac{2(b - d)}{b + d} \left( 1 + \sum_{n=1}^{\infty} \zeta_n^- x^{n/m} \right) \left( 1 + \sum_{n=1}^{\infty} \zeta_n^+ x^{n/m} \right)
\]

where \( \zeta_n^\pm := (bb_n \pm dd_n)/(b \pm d) \). Following the notation introduced in Proposition A.2 of the Appendix, we have

\[
\psi(x) = \frac{2(b - d)}{b + d} \left( 1 + \sum_{n=1}^{\infty} c_n x^{n/m} \right)
\]

so that in the notation of (18), we have \( c = 2(b - d)/(b + d) \) and \( \kappa = 0 \). The remaining coefficients can be expressed in terms of \( \{\zeta_n^+\} \) and \( \{\zeta_n^-\} \):

\[
c_n = \zeta_n^- + \left( \sum_{j=1}^{n-1} \zeta_{n-j}^- \hat{\zeta}_j^+ \right) + \hat{\zeta}_n^+.
\]

Case 2: \( \beta = \delta, b = d \).

In this case we can factor \( bx^\beta = dx^\delta \) from the numerator and denominator of (20), which leaves

\[
\psi(x) = \frac{2(\sum_{n=1}^{\infty} (b_n - d_n)x^{n/m})}{2 + \sum_{n=1}^{\infty} (b_n + d_n)x^{n/m}} = \left( \sum_{n=\kappa}^{\infty} \xi_n^- x^{n/m} \right) \left( 1 + \sum_{n=1}^{\infty} \hat{\xi}_n^+ x^{n/m} \right)
\]

where \( \xi_n^+ = (b_n + d_n)/2 \) for all \( n \geq 1 \), and \( \xi_n^- = b_n - d_n \) for \( n \geq \kappa \). The form of \( \hat{\xi}_n^+ \) follows from Proposition A.2 in the Appendix. It follows that \( \psi(x) \) can be written in the form of (18) with the leading order exponent being \( \kappa/m \) and

\[
c = b_\kappa - d_\kappa \text{, and } c_n = \xi_n^- + \sum_{j=\kappa}^{n-1} \xi_{n-j}^- \hat{\xi}_j^+.
\]

Case 3: \( \beta < \delta \)

In this case we begin by factoring out \( bx^\beta \) from the numerator and denominator of equation (20). After cancellation we have

\[
\psi(x) = \frac{2 \left( 1 + \sum_{n=1}^{\infty} b_n x^{n/m} - \frac{b}{d} x^{\delta-\beta} \left( 1 + \sum_{n=1}^{\infty} d_n x^{n/m} \right) \right)}{1 + \sum_{n=1}^{\infty} b_n x^{n/m} + \frac{d}{b} x^{\delta-\beta} \left( 1 + \sum_{n=1}^{\infty} d_n x^{n/m} \right)}.
\]
From this form and Proposition A.2, we see that $\kappa = 0$ and $c = 2$. Furthermore, the existence of the sequence $c_n$ follows from our assumption that $(\beta - \delta)m$ is an integer.

**Case 4: $\beta > \delta$**

In this case we begin by factoring out $dx^\delta$ from the numerator and denominator of equation (20). After cancellation and some rearrangement we have

$$
\psi(x) = \frac{-2\left(1 + \sum_{n=1}^{\infty} d_n x^{n/m} - b \frac{1}{x^{\beta - \delta}} \left(1 + \sum_{n=1}^{\infty} b_n x^{n/m}\right)\right)}{1 + \sum_{n=1}^{\infty} d_n x^{n/m} + b \frac{1}{x^{\beta - \delta}} \left(1 + \sum_{n=1}^{\infty} b_n x^{n/m}\right)}
$$

Similar to Case 3, we see that $\kappa = 0$ and $c = -2$.

From here, the proof proceeds in two steps. First we need to rewrite the integral $\Psi(x)$ presented in (12) in a form that is compatible with Lemma 3.7.

**Lemma 3.6.** Suppose that the shape functions $\lambda$ and $\mu$ satisfy Assumption 1.3 and we write $\psi$ as described in Lemma 3.5. Furthermore, let $\kappa \in \mathbb{Z}^+$ be as defined in Lemma 3.5. Then $f(t)$ admits a series expansion of the form

$$
(21) \quad f(t) = at^\alpha \left(1 + \sum_{n=1}^{\infty} a_n t^{n/(m+\kappa)}\right),
$$

where $m$ is the positive integer defined in Assumption 1.3 and

$$
a = \frac{1}{c} \left(\frac{cm}{\kappa + m}\right)^{\kappa/(\kappa + m)}, \quad \alpha = -\frac{\kappa}{\kappa + m}, \quad \text{and} \quad a_1 = c_1 \left(\frac{cm}{\kappa + m}\right)^{-1/(\kappa + m)} \left(\frac{\kappa}{\kappa + m + 1} - 1\right).
$$

**Proof.** It will be convenient to write the series expansion for $\psi$ in the form

$$
\psi(x) = c \sum_{n=0}^{\infty} c_n x^{\frac{1}{m}(n+\kappa)}
$$

with the convention that $c_0 = 1$. For the main part of the proof, we will take the $c > 0$ case. At the end, we will explain how the argument changes when $c < 0$. Integrating term-by-term, we have

$$
\Psi(x) = c \sum_{n=0}^{\infty} \frac{c_n m}{n + \kappa + m} x^{\frac{1}{m}(n+\kappa+m)}.
$$

It will be useful to define $h(x) := \Psi(x^m)$ with a series expansion written as

$$
h(x) = \sum_{n=0}^{\infty} h_n x^{n+\kappa+m}, \quad \text{where} \quad h_n = \frac{c c_n m}{n + \kappa + m}.
$$

Since $c > 0$, $\Psi(x)$ is increasing on $[0, x_\ast)$ and therefore $h(x)$ is increasing on $[0, x^{\frac{1}{m}})$. Setting $t = h(x)$, by Proposition A.3, we can express the inverse of $h$ in terms of the series

$$
h^{-1}(t) = \sum_{n=1}^{\infty} \tilde{h}_n t^{n/(\kappa+m)}
$$

where

$$
\tilde{h}_1 = \frac{1}{h_0^{1/(\kappa+m)}}, \quad \tilde{h}_2 = -\frac{1}{(\kappa + m)h_0^{1+2/(\kappa+m)}}.
$$
Recalling that \( \psi(z) = y^{1/m} \), we have
\[
I_t(x) := \int_0^x e^{-\Psi(y)/\epsilon} \, dy = \int_0^{x^{1/m}} mz^{m-1} e^{-h(z)} \, dz.
\]
Then, taking \( t = h(z) \), and observing that \( \frac{d}{dz} h(z) = \frac{d}{dz} \Psi(z^m) = \psi(z^m) mz^{m-1} \), we have
\[
I_t(x) = \int_0^{h(x^{1/m})} m(h^{-1}(t))^{m-1} \frac{1}{h'(h^{-1}(t))} e^{-t/\epsilon} \, dt = \int_0^{h(x^{1/m})} \frac{1}{\psi(h^{-1}(t))^m} e^{-t/\epsilon} \, dt.
\]
To determine the series expansion for \( f(t) := 1/\psi(h^{-1}(t)^m) \), we need to analyze the expansion for the denominator:
\[
\psi(h^{-1}(t)^m) = c \sum_{n=0}^{\infty} \epsilon_n (h^{-1}(t))^{n+\kappa}
\]
(22)
\[
=c \sum_{n=0}^{\infty} \epsilon_n \left( \sum_{j=1}^{\infty} \frac{\tilde{h}_j t^{j/(\kappa+m)}}{m} \right)^{n+\kappa}.
\]
It is helpful to rewrite \( h^{-1}(t) \) as follows
\[
h^{-1}(t) = \tilde{h}_1 t^{1/(\kappa+m)} \left( 1 + \frac{\tilde{h}_2}{\tilde{h}_1} t^{1/(\kappa+m)} + O(t^{2/(\kappa+m)}) \right).
\]
Recalling that \( \psi(x) = c x^{\kappa/m} + cc_1 x^{(\kappa+1)/m} + O(x^{(\kappa+2)/m}) \), we observe
\[
\psi(h^{-1}(t)^m) = c(h^{-1}(t))^{\kappa} + cc_1 h^{-1}(t)^{\kappa+1} + O(h^{-1}(t)^{\kappa+2})
\]
\[
= c \tilde{h}_1^{\kappa} t^{\kappa/(\kappa+m)} \left( 1 + \frac{\tilde{h}_2}{\tilde{h}_1} t^{1/(\kappa+m)} + O(t^{2/(\kappa+m)}) \right) \kappa
\]
\[
+ cc_1 \tilde{h}_1^{\kappa+1} t^{(\kappa+1)/(\kappa+m)} \left( 1 + \frac{\tilde{h}_2}{\tilde{h}_1} t^{1/(\kappa+m)} + O(t^{2/(\kappa+m)}) \right) \kappa+1
\]
\[
+ O(t^{(\kappa+2)/(\kappa+m)}).
\]
Then by applying Proposition A.1 twice, we obtain
\[
\psi(h^{-1}(t)^m) = c \tilde{h}_1^{\kappa} t^{\kappa/(\kappa+m)} \left( 1 + \left( \frac{k \tilde{h}_2}{\tilde{h}_1} + c_1 \tilde{h}_1 \right) t^{1/(\kappa+m)} \right) + O(t^{(\kappa+2)/(\kappa+m)}).
\]
Proposition A.2 then yields
\[
f(t) = \frac{1}{\psi(h^{-1}(t)^m)} = \frac{1}{c \tilde{h}_1^{\kappa} t^{\kappa/(\kappa+m)}} \left( 1 - \left( \frac{k \tilde{h}_2}{\tilde{h}_1} + c_1 \tilde{h}_1 \right) t^{1/(\kappa+m)} \right) + O(t^{(\kappa+2)/(\kappa+m)}).
\]
After simplifying, the stated results for \( a, \alpha \) and \( a_1 \) hold.

When \( c < 0 \), we replace \( \psi \) and \( \Psi \) with \( \psi_r \) and \( \Psi_r \), respectively. The leading term of \( \psi_r \) is then \(-c\), which is positive and exact same procedure holds.

**Lemma 3.7.** Suppose that \( f : \mathbb{R} \rightarrow \mathbb{R} \) admits a series expansion of the form (21) with \( a > 0 \), \( \alpha > -1 \), \( \{a_n\}_{n=1}^{\infty} \subset \mathbb{R} \), \( \kappa \in \mathbb{Z}_+ \) and \( m \in \mathbb{N} \). Furthermore, let \( g : \mathbb{R} \rightarrow \mathbb{R} \) be a continuously differentiable function that is monotonically increasing or decreasing for all \( t \in (0,t_*) \), with \( g(0) = 0 \). Define

\[
R_k^\pm [f; g] := \lim_{\epsilon \to 0} \frac{\int_0^{g(k\epsilon)} f(t)e^{\pm t/\epsilon} dt}{\int_0^{g(t_*)} f(t)e^{\pm t/\epsilon} dt}.
\]

Then

\[
R_k^+ [f; g] = 0;
\]

\[
R_k^- [f; g] = 1 - e^{-kg'(0)}.
\]

**Proof.** Our method is to replace \( f(t) \) with its series expansion and consider the integration term-by-term. Our results will be expressed in terms of the incomplete gamma function

\[
\gamma(s, x) := \int_0^x t^{s-1}e^{-t} dt.
\]

Near zero, the incomplete gamma function has the behavior that

\[
\lim_{x \to 0} \frac{\gamma(s, x)}{x^s} = \frac{1}{s}.
\]

As usual, we define \( \Gamma(s) = \lim_{x \to \infty} \gamma(s, x) \). Note that for any \( s > -1 \) and \( A > 0 \), under the substitution \( y = t/\epsilon \), we have

\[
\int_0^A t^{s-1}e^{-t/\epsilon} dt = \epsilon^s \int_0^{A/\epsilon} y^{s-1}e^{-y} dy = \epsilon^s \gamma(s, A/\epsilon).
\]

Therefore,

\[
\lim_{\epsilon \to 0} \frac{\int_0^A t^{s-1}e^{-t/\epsilon} dt}{\epsilon^s} = \Gamma(s).
\]

Writing \( G \) for \( g(k\epsilon) \) and \( g(t_*) \) respectively, it follows that the numerator and the denominator appearing in \( R_k^- [f; g] \) have the form

\[
\int_0^G f(t)e^{-t/\epsilon} dt = a \int_0^G (t^\alpha + a_1 t^\alpha t_*/(n+m) + \sum_{n=2}^{\infty} a_n t^\alpha t^{n/(n+m)}) e^{-t/\epsilon} dt
\]

\[
= a e^{\alpha+1} \gamma(\alpha + 1, G/\epsilon) + O(\epsilon^{\alpha+1/(n+m)}).
\]

In taking the ratio, the coefficients \( ae^{\alpha+1} \) cancel, and neglecting higher order terms, we have

\[
R_k^- [f; g] = \lim_{\epsilon \to 0} \frac{\gamma(\alpha + 1, g(k\epsilon)/\epsilon)}{\gamma(\alpha + 1, g(t_*)/\epsilon)} = \frac{\gamma(\alpha + 1, kg'(0))}{\Gamma(\alpha + 1)}.
\]

When \( \alpha = 0 \), this takes the form

\[
R_k^- [f; g] = 1 - e^{-kg'(0)}.
\]
For the plus-sign case, we have
\[ R^+_k [f; g] = \lim_{\epsilon \to 0} \frac{\int_0^{g(k\epsilon)} t^\alpha e^{t/\epsilon} dt}{\int_0^{g(t^*_\epsilon)} t^\alpha e^{t/\epsilon} dt}, \]

Since \( g' \) is continuous, let \( M := \max_{t \in [0, t_*/2]} g'(t) \). Then by the Taylor Remainder Theorem, since \( g(0) = 0 \) we have
\[ g(k\epsilon) = k\epsilon g'(\xi), \]
for some \( \xi \in (0, t_*/2) \). It follows that \( g(k\epsilon) \leq kM\epsilon \) for all \( t \in [0, t_*/2] \). So for the numerator we have
\[ \lim_{\epsilon \to 0} \int_0^{g(k\epsilon)} t^\alpha e^{t/\epsilon} dt \leq \lim_{\epsilon \to 0} \int_0^{kM\epsilon} t^\alpha e^{kM} dt = 0. \]

To study the denominator, we again make the substitution \( y = t/\epsilon \) and obtain
\[ \lim_{\epsilon \to 0} \int_0^{g(t^*_\epsilon)} t^\alpha e^{t/\epsilon} dt = \lim_{\epsilon \to 0} \epsilon \int_0^{g(t^*_\epsilon)/\epsilon} (\epsilon y)^\alpha e^{\epsilon y} dy \]
\[ = \lim_{\epsilon \to 0} \epsilon^{\alpha+1} \int_0^{g(t^*_\epsilon)/\epsilon} y^\alpha e^{\epsilon y} dy. \]

Noticing that the integral on the right hand side diverges to infinity, we rewrite the expression and apply l'Hôpital's rule to find that the denominator diverges to infinity as \( \epsilon \to 0 \). It follows that the ratio tends to \( 0 \) as \( \epsilon \to 0 \). \( \square \)

We introduced the notation \( f \) and \( g \) to emphasize that the contribution to the final value comes from the limit of integration, not the function \( f \) in the integrand.

**Proof of Theorem 3.4.** First, recalling Definition 1.6 and the result from Proposition 13, we find that
\[ p_{\text{diffusion}}(k) = \frac{\int_0^x e^{-\Psi(y)/\epsilon} dy}{\int_0^x e^{-\Psi(y)/\epsilon} dy}. \]

Noticing the familiar form of this expression, we apply Lemma 3.6 with \( g \) set to \( \Psi \) to rewrite \( p_{\text{diffusion}}(k) \) in a form compatible with Lemma 3.7.

When \( \beta > \delta \) or \( \beta = \delta \) with \( b \leq d \) Lemma 3.7 implies
\[ p_{\text{diffusion}}(k) = R^+_k [f; \Psi] = 0. \]
Alternatively, if \( \beta < \delta \) or \( \beta = \delta \) with \( b > d \), Lemma 3.7 shows that
\[ p_{\text{diffusion}}(k) = R^-_k [f; \Psi] = 1 - e^{-k\Psi'(0)}. \]

Since \( \Psi'(0) = \psi'(0) \), by Definition 12, the final result is attained by evaluating \( \psi(0) \) using Lemma 3.5.

**4. Exponential approximation.** Doering et al [11, 12] demonstrated that making a WKB-type ansatz of the form \( q_n \approx \sigma_n e^{-V_n} \) for some functions \( \sigma \) and \( V \), can be an accurate method for constructing a continuum approximation for solving Kolmogorov equations. In the motivation that follows we provide a different analytical justification for this observation than has been presented elsewhere. We do this by transforming the system of equations defined by (8) into an equation for ratios instead of differences, then applying a Taylor series expansion technique similar to what is presented as Motivation 2 for the Diffusion Approximation.
4.1. Motivation for the Exponential Approximation. For \( n = 1, 2, \ldots, N_\ast \), define

\[ a_N(n) := q_N(n - 1)/q_N(n). \]

Since \( q_N(0) = 1 \), we can write

\[ q_N(n) = \prod_{k=1}^{n} \frac{1}{a_N(k)} = \prod_{k=1}^{n} \exp(- \ln a_N(k)), \]

where

\[ \frac{\lambda_N(n)}{a_N(n + 1)} - (\lambda_N(n) + \mu_N(n)) + \mu_N(n)a_N(n) = 0. \]

We can rewrite the exponent on the right-hand side to look like a Riemann sum:

\[ q_N(n) = \exp \left( -N \sum_{k=1}^{n} \ln \left( a_N(k) \right) \frac{1}{N} \right). \]

We note that the above form of the extinction probability is identical to that of the mean extinction time found in Doering et al. (2007) \[12\], where in our case \( \Delta V_n = - \ln(a_N(k)) \). Mimicking the continuum approximations and Taylor expansions presented as motivation for the Diffusion Approximation, we introduce the function \( r_\epsilon(x) \) defined below. For \( x = n/N \) and \( \epsilon = 1/N \), we will write

\[ a_N(n) \approx r_\epsilon \left( \frac{n}{N} \right) = r_\epsilon(x). \]

Then,

\[ \sum_{i=1}^{n} \ln \left( a_N(i) \right) \frac{1}{N} \approx \int_{0}^{x} \ln r_\epsilon(y) \, dy \]

and the heuristic assertion is that (26) becomes

\[ q_N(n) \approx \exp \left( - \frac{1}{\epsilon} \int_{0}^{x} \ln r_\epsilon(y) \, dy \right). \]

We note that this is the same form as presented in Definition 1.7 but it remains to characterize the function \( r_\epsilon(x) \). Indeed, following a Taylor series expansion technique, we will show that \( r_\epsilon(x) \) should have the form

\[ r_\epsilon(x) = \rho(x) + \epsilon \frac{\rho'(x)}{\rho(x) - 1} + \epsilon^2 \frac{\rho''(x)\rho^2(x) - 2(\rho'(x))^2}{2(\rho(x) - 1)^3} + O(\epsilon^3). \]

Dividing (25) through by \( \mu_N(n) \) and multiplying by \( a_N(n + 1) \) we have

\[ \frac{\lambda_N(n)}{\mu_N(n)} + a_N(n)a_N(n + 1) = \left( \frac{\lambda_N(n)}{\mu_N(n)} + 1 \right) a_N(n + 1). \]

Let

\[ \rho(x) := \lambda(x)/\mu(x), \]
then using the definition of the transition rate shape functions $\lambda$ and $\mu$, we have
\[
\frac{\lambda_N(n)}{\mu_N(n)} = \frac{N\lambda\left(\frac{n}{N}\right)}{N\mu\left(\frac{n}{N}\right)} = \rho\left(\frac{n}{N}\right).
\]
Thus,
\[
\rho\left(\frac{n}{N}\right) + r\left(\frac{n}{N}\right) r\left(\frac{n+1}{N}\right) = \left(\rho\left(\frac{n}{N}\right) + 1\right) r\left(\frac{n+1}{N}\right),
\]
and, substituting $x$ and $\epsilon$ for $n/N$ and $1/N$, we arrive at
\[
\rho(x) + r(x) r(x + \epsilon) - (1 + \rho(x)) r(x + \epsilon) = 0.
\]
Neglecting higher order terms, we make the substitution
\[r_\epsilon(x) \approx r_0(x) + \epsilon r_1(x) + \epsilon^2 r_2(x),\]
and obtain
\[
\rho(x) + (r_0(x) + \epsilon r_1(x) + \epsilon^2 r_2(x))(r_0(x + \epsilon) + \epsilon r_1(x + \epsilon) + \epsilon^2 r_2(x + \epsilon))
\]
\[- (1 + \rho(x))(r_0(x + \epsilon) + \epsilon r_1(x + \epsilon) + \epsilon^2 r_2(x + \epsilon)).
\]
Now we perform a Taylor expansion on the $r_0(x + \epsilon)$, $r_1(x + \epsilon)$, and $r_2(x + \epsilon)$ terms. Organizing the terms by powers of $\epsilon$, we find the following system of equations:
\[
\epsilon^0 : \rho(x) + r_0'(x) - (1 + \rho(x)) r_0(x) = 0,
\]
\[
\epsilon^1 : r_0(x) r_0'(x) + 2 r_0(x) r_1(x) - (1 + \rho(x))(r_0'(x) + r_1(x)) = 0,
\]
\[
\epsilon^2 : \frac{1}{2} r_0(x) r_0''(x) + r_0(x) r_1'(x) + 2 r_0(x) r_2(x) + r_1(x) r_0'(x) + r_1(x) r_0''(x) + r_1(x) r_1'(x) + r_2(x)
\]
\[- \frac{1}{2}(1 + \rho(x)) r_0''(x) - (1 + \rho(x)) r_1'(x) - (1 + \rho(x)) r_2(x) = 0.
\]
The first of these equations yields two solutions for $r_0(x)$: $r_0(x) = \rho(x)$ and $r_0(x) = 1$. To show that the former must be true, consider that if $r_0(x) \equiv 1$, then $r_0'(x) \equiv 0$. Substituting into the $\epsilon^1$ equation, it follows that $r_1(x) = 0$. Continuing in this way, we find $r_i(x) = 0$ for all $i \geq 1$. This corresponds to the solution $q_N(n) = 1$ for all $n \in \{0, \ldots, N_\epsilon\}$, which does not satisfy the boundary condition $q_N(N_\epsilon) = 0$. Therefore, we adopt the solution $r_0(x) = \rho(x)$. Feeding this into the order $\epsilon^1$ equation allows us to solve for $r_1(x)$. Similarly, substituting these solutions into the $\epsilon^2$ equation yields a solution for $r_2(x)$ and so on:
\[
r_0(x) = \rho(x),
\]
\[
r_1(x) = \frac{\rho'(x)}{\rho(x) - 1},
\]
\[
r_2(x) = \frac{\rho''(x)\rho^2(x) - 2(\rho'(x))^2}{2(\rho(x) - 1)^3}.
\]

4.2. Analysis of $p_{\text{exp}}(k)$. For the analysis that follows, we define what we call an nth order exponential approximation. Let
\[r_{\epsilon,n}(x) = r_0(x) + \epsilon r_1(x) + \cdots + \epsilon^n r_n(x)\]
where the functions $r_j(x)$ are defined according to the procedure given in the preceding section. Then, we define
\[
p_{\text{exp},n}(k) := 1 - \lim_{\epsilon \to 0} w_{\epsilon,n}(k\epsilon) = \lim_{\epsilon \to 0} 1 - \exp\left(-\frac{1}{\epsilon} \int_0^{k\epsilon} \ln r_{\epsilon,n}(y) \, dy\right).
\]
THEOREM 4.1 (Leading Order Exponential Approximation for Supercritical Systems). Suppose that the transition rate shape functions $\lambda$ and $\mu$ form a supercritical system, as defined in Assumption 1.1. Further suppose that $\lambda$ and $\mu$ each admit a series expansion, as defined by Assumption 1.3.

If the leading order exponents $\beta$ and $\delta$ are equal, then for any $n \in \mathbb{N},$

$$p_{\exp,0}(k) = 1 - \left(\frac{d}{b}\right)^k \text{ when } b \geq d.$$  

Otherwise, when $\beta < \delta,$ $p_{\exp,0}(k) = 1.$

Remark 4.2. When $\lambda$ and $\mu$ form a subcritical system this method of approximation is not well defined. In this regime, the function $w_{\epsilon,n}(k\epsilon)$ is often greater than one, meaning that $1 - w_{\epsilon,n}(k\epsilon)$ is negative and not a reasonable estimate for a probability. For example, note that this is what happens to equation (31) when $\beta = \delta$ and $b < d.$ The proposed value for $p_{\text{invasion}}(k)$ is negative and cannot be an invasion probability. A practitioner could take the Exponential Approximation to be zero whenever the system of study is subcritical.

The proof of Theorem 4.1 is an immediate results of the following lemma.

LEMMA 4.3. Suppose that the shape functions $\lambda$ and $\mu$ each admit a series expansion as defined by Assumption 1.3, then $w_{\epsilon,0}(k\epsilon)$ has the form:

$$\exp\left(-\frac{1}{\epsilon} \int_{0}^{k\epsilon} \ln \rho(y) dy\right) = \left(\frac{d}{b}\right)^k \left(\frac{\epsilon}{k\epsilon}\right)^{(\beta-\delta)k} \exp\left(\frac{d_1 - b_1}{1+1/m} k(\epsilon)^{1/m} + O((\epsilon)^{2/m})\right).$$

Proof. First we find a series expansion for $\rho(x).$ From the definition of $\rho$ in equation (29) and the series expansions for $\lambda$ and $\mu$ assumed in Assumption 1.3, we have

$$\rho(x) = \frac{b}{d} x^{\beta-\delta} \left(1 + \sum_{n=1}^{\infty} b_n x^{n/m}\right) \left(1 + \sum_{n=1}^{\infty} d_n x^{n/m}\right)^{-1}.$$

By applying Proposition A.2 and separating out the leading order terms, we obtain

$$\rho(x) \approx \frac{b}{d} x^{\beta-\delta} \left(1 + b_1 x^{1/m} + \sum_{n=2}^{\infty} b_n x^{n/m}\right) \left(1 - \left(d_1 x^{1/m} + \sum_{n=2}^{\infty} d_n x^{n/m}\right) + \left(\sum_{n=1}^{\infty} d_n x^{n/m}\right)^2\right).$$

Collecting ordered terms yields the expansion

$$\rho(x) = \frac{b}{d} x^{\beta-\delta} \left(1 + (b_1 - d_1) x^{1/m} + O(x^{2/m})\right).$$

To consider the leading order approximation we use one term of $r_{\epsilon}(x),$ i.e. $r_{\epsilon}(x) = \rho(x).$ Then by plugging equation (33) into Definition 1.7, we have:

$$w_{\epsilon,0}(x) = \exp\left(-\frac{1}{\epsilon} \int_{0}^{x} \ln \left(\frac{b}{d} y^{\beta-\delta} \left(1 + (b_1 - d_1) y^{1/m} + O(y^{2/m})\right)\right) dy\right),$$
which can then be expressed in terms of having \( k \) individuals introduced
\[
\begin{align*}
    w_{x,0}(k\epsilon) &= \exp \left( -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(b/d) dy \right) \exp \left( -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(y^{\beta - \delta}) dy \right) \\
    &\times \exp \left( -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(1 + (b_1 - d_1)y^{1/m} + O(y^{2/m})) dy \right).
\end{align*}
\] 

Considering the first term in the product of equation (34),
\[
\begin{align*}
    \exp \left( -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(b/d) dy \right) &= \exp \left( -\frac{1}{\epsilon} k\epsilon \ln(b/d) \right), \\
    &= \left( \frac{d}{b} \right)^k.
\end{align*}
\]

Then using an expansion for \( \ln(1 + x) \) and integrating the result, the third term in the product becomes
\[
\begin{align*}
    \int_0^{k\epsilon} \ln(1 + (b_1 - d_1)y^{1/m} + O(y^{2/m})) dy &= \int_0^{k\epsilon} (b_1 - d_1)y^{1/m} + O(y^{2/m}) dy, \\
    &= \frac{b_1 - d_1}{1 + 1/m} (k\epsilon)^{1+1/m} + O((k\epsilon)^{1+2/m}).
\end{align*}
\]

Thus the third term in the product of equation (34) can be rewritten as
\[
\begin{align*}
    \exp \left( -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(1 + (b_1 - d_1)y^{1/m} + O(y^{2/m})) dy \right) &= \exp \left( \frac{d_1 - b_1}{1 + 1/m} k(k\epsilon)^{1/m} + O((k\epsilon)^{2/m}) \right).
\end{align*}
\]

Rearranging and integrating the exponent from the second term in equation (34), we have
\[
\begin{align*}
    -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(y^{\beta - \delta}) dy &= -\frac{\beta - \delta}{\epsilon} \int_0^{k\epsilon} \ln y \ dy, \\
    &= -\frac{\beta - \delta}{\epsilon} \left[ \ln y - y \right]_0^{k\epsilon}, \\
    &= -(\beta - \delta)(k \ln(k\epsilon) - k).
\end{align*}
\]

Therefore, the second term becomes
\[
\begin{align*}
    \exp \left( -\frac{1}{\epsilon} \int_0^{k\epsilon} \ln(y^{\beta - \delta}) dy \right) &= \left( \frac{e}{k\epsilon} \right)^{(\beta - \delta)k}.
\end{align*}
\]

The result follows by rewriting equation (34) using equations (35), (36), and (37). 

In Lemma 4.3, we show that the Exponential Approximation exhibits a non-trivial dependence on \( k \), the number of individuals introduced in the population. This dependence results from the relationship between the number of introduced individuals and the population scale \( N \). We explicitly included the number of individuals introduced as a fraction of the population \( (k\epsilon) \) in the error term to record that error terms do not necessarily approach zero with \( \epsilon \) if the number of individuals introduced is also allowed to vary.
To conclude our analysis, we return our attention to the functions \( r_n(x) \) contained within equation (30). While we do not currently have a complete result for analyzing higher-order exponential approximations \( p_{\exp,n}(k) \) with \( n > 0 \), we would like to report the following recursion formula that can be used to generate these terms.

**Proposition 4.4.** For \( n \geq 2 \), \( r_n(x) \) has the form

\[
\begin{align*}
  r_n(x) &= \frac{1}{r_0(x) - 1} \left( \sum_{i=0}^{n-1} r_i^{(n-i)}(x) - \sum_{i=1}^{n-1} r_i(x) \left( \sum_{j=0}^{n-i} r_j^{(n-i-j)}(x) (n-i-j)! \right) \right). 
\end{align*}
\]

**Proof.** We begin by recalling the equation

\[
\rho(x) + r_1(x) + (1 + \rho(x)) r_0(x) = 0.
\]

We then rewrite \( r_1(x) \) as a series in powers of \( \epsilon \)

\[
\rho(x) + \left( \sum_{i=0}^{n} \epsilon^i r_i(x) \right) \left( \sum_{i=0}^{n} \epsilon^i r_i(x) + \epsilon \right) - (1 + \rho(x)) \left( \sum_{i=0}^{n} \epsilon^i r_i(x) + \epsilon \right) = 0.
\]

Substituting in Taylor expansions for the \( r_i(x + \epsilon) \) terms we obtain

\[
\rho(x) + \left( \sum_{i=0}^{n} \epsilon^i r_i(x) \right) \left( \sum_{i=0}^{n} \epsilon^i \sum_{j=0}^{n} \frac{\epsilon^j r_i^{(j)}(x)}{j!} \right) - (1 + \rho(x)) \left( \sum_{i=0}^{n} \epsilon^i \sum_{j=0}^{n} \frac{\epsilon^j r_i^{(j)}(x)}{j!} \right) = 0.
\]

Since \( r_0(x) = \rho(x) \) this reduces to

\[
\rho(x) + \left( \sum_{i=1}^{n} \epsilon^i r_i(x) \right) \left( \sum_{i=0}^{n} \epsilon^i \sum_{j=0}^{n} \frac{\epsilon^j r_i^{(j)}(x)}{j!} \right) - \left( \sum_{i=0}^{n} \epsilon^i \sum_{j=0}^{n} \frac{\epsilon^j r_i^{(j)}(x)}{j!} \right) = 0.
\]

To solve for \( r_n(x) \) we will consider the \( \epsilon^n \) terms. First we consider the term

\[
\sum_{i=0}^{n} \sum_{j=0}^{n} \frac{\epsilon^{i+j} r_i^{(j)}(x)}{j!}
\]

and notice that the \( \epsilon^n \) terms are contained in the series

\[
\sum_{i=0}^{n} \epsilon^i \frac{\epsilon^{n-i} r_i^{(n-i)}(x)}{(n-i)!} = \epsilon^n \sum_{i=0}^{n} \frac{r_i^{(n-i)}(x)}{(n-i)!}.
\]

Now considering the second term in (39), we proceed by isolating the \( \epsilon^{n-i} \) terms from the double summation in that product

\[
\sum_{j=0}^{n-i} \epsilon^j \frac{\epsilon^{n-i-j} r_j^{(n-i-j)}(x)}{(n-i-j)!} = \epsilon^{n-i} \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}(x)}{(n-i-j)!}.
\]

Thus, the \( \epsilon^n \) terms in (39) can be expressed as

\[
\left( \sum_{i=1}^{n} \epsilon^i r_i(x) \right) \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}(x)}{(n-i-j)!} - \epsilon^n \sum_{i=0}^{n} \frac{r_i^{(n-i)}(x)}{(n-i)!}.
\]
CONTINUUM APPROXIMATION OF INVASION PROBABILITIES

It follows from (39) that

\[ \sum_{i=1}^{n} r_i(x) \left( \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}}{(n-i-j)!} \right) - \sum_{i=0}^{n} r_i^{(n-i)}(x) = 0. \]

Then for the first term we have

\[ \sum_{i=1}^{n} r_i(x) \left( \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}}{(n-i-j)!} \right) = r_n(x) \sum_{j=0}^{n} \frac{r_j^{(n-n-j)}}{(n-n-j)!} + \sum_{i=1}^{n-1} r_i(x) \left( \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}}{(n-i-j)!} \right) \]

\[ = r_n(x) r_0(x) + \sum_{i=1}^{n-1} r_i(x) \left( \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}}{(n-i-j)!} \right). \]

After also separating out the \( r_n(x) \) term from the second series we have

\[ r_n(x) r_0(x) + \sum_{i=1}^{n-1} r_i(x) \left( \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}}{(n-i-j)!} \right) - r_n(x) - \sum_{i=0}^{n-1} \frac{r_i^{(n-i)}}{(n-i)!} = 0 \]

and

\[ r_n(x) (r_0(x) - 1) = \sum_{i=0}^{n-1} \frac{r_i^{(n-i)}}{(n-i)!} - \sum_{i=1}^{n-1} r_i(x) \left( \sum_{j=0}^{n-i} \frac{r_j^{(n-i-j)}}{(n-i-j)!} \right). \]

The final result is obtained by solving for \( r_n(x) \).

5. Numerical observations for invasion probability approximations. In the preceding sections, our main analytical results for the Branching Process, Diffusion, and Exponential Approximations focus on the large population \((N \to \infty)\) limit. In this section, we validate our asymptotic results numerically and then shift our attention to consider the behavior of the approximations when the population size is small or intermediate. We challenge the methods in their ability to best approximate the exact solution as defined and calculated in Section 2. We focus our investigation on the specific examples presented in the introduction, and obtain exact solutions for the approximations whenever possible. We use numerical integration when it is not possible to obtain an exact solution for an approximation. In this way, we verify our analytical results and explore how the conditions of a system (e.g. population size and subcritical vs. supercritical dynamics) have an impact on whether a particular approximation method should be deemed fit for use. We set the death rate parameter \((d)\) to be one by default.

5.1. Diffusion approximation methods fail for large populations when dynamics are supercritical. To complement our analysis in the previous section, we used exact solutions when possible and numerical integration when necessary (Simpson’s method coded in R) to evaluate the Diffusion Approximation and Exponential Approximation for invasion probabilities. In Figure 2, the results are shown for Examples 1, 2, and 3 for a range of population sizes and parameter value choices. We chose the exponents and coefficients so that the dynamics are supercritical and the Exponential Approximation is well defined. In each panel, we highlight that as the total population size becomes large, the values calculated numerically approach their corresponding analytically determined limit (indicated by \(\ast\)).
For the supercritical systems displayed in Figure 2 the Diffusion Approximation yields a different answer than the exact solution in the large population size limit. This discrepancy between the Diffusion Approximation and the exact solution confirms our asymptotic analysis and is most apparent when the leading coefficients of the birth and death rates are dissimilar. As displayed in the last column of Figure 2, the Diffusion Approximation fails to match the exact solution when the dynamics are far from critical. For population sizes greater than 100 in Figure 3, the Diffusion Approximation fails to match the exact solution for both near critical and far from critical dynamics. The inability of the Diffusion Approximation to approach one is predicted by our asymptotic result that $p_{\text{diffusion}}(k) = 1 - e^{-2k}$ when $\beta < \delta$. This can be seen in the middle row of Figure 2 and the top row of Figure 3. These plots also highlight the phenomenon that, if the leading order exponents satisfy $\beta < \delta$, then in the large population limit, the Diffusion Approximation completely ignores the parameters of the rate functions.
By contrast, the Diffusion Approximation succeeds in characterizing the large population size behavior when the dynamics are dominated by the death rate ($\mu$ is leading order). This is seen in the bottom row of Figure 3, as predicted by our result that $p_{\text{diffusion}}(k) = 0$ when $\beta > \delta$. (See Theorem 3.4.)

5.2. Diffusion approximation methods can work well for small populations that exhibit near critical dynamics. In the examples we have studied, the Diffusion Approximation consistently outperforms the Exponential Approximation when the population is small and the parameter set is near critical. This result is first demonstrated in Figure 2. Moving from left to right in this multi-panel figure, the parameter characterizing the birth rate ($b$) moves farther away from the death rate parameter ($d$) which is set to one by default. As such, the first column shows that the Diffusion Approximation closely approximates the exact solution for the invasion probability for small population sizes.

The Diffusion Approximation’s transition from exceptional to poor performance is even more clearly demonstrated when studying Example 4. In the top row of Figure 3 we see that the Diffusion Approximation captures non-monotonic features of the exact solution that the Exponential Approximation misses entirely. In this example, the dynamics are dictated by the birth rate since $\lambda$ features the lower leading order term. When the dynamics are near critical $b = 0.7$ and $b = 1.1$, there is a range of small population sizes where the Diffusion Approximation tracks the exact solution.

5.3. When leading order terms match, higher order terms matter: for small, but not large population sizes. When the leading order of the “birth” and “death” rates are the same and their leading coefficients are equal, subtleties in the outcomes are determined by the first pair of mismatched coefficients. This result
is displayed prominently in Figure 4 for which we chose both rate functions to be asymptotically linear ($\beta = \delta = 1$), to share the same leading coefficient ($b = d = 1$), but to have different values for the coefficients $b_1$ and $d_1$ (see Assumption 1.3).

**Fig. 4. Identical Leading Order Terms.** Comparison between the diffusion and exponential approximations when both rate equations are asymptotically linear ($\beta = \delta = 1$) and the leading coefficients match ($b = d = 1$). In both panels, the invasion probability tends to zero as the population size ($N$) becomes large. Differences between the left (supercritical) and right (subcritical) panels are driven by the leading pair of terms with mismatched coefficients ($b_1 \neq d_1$).

In the left panel of Figure 4, $b_1 > d_1$ and the system is supercritical (as defined in Assumption 1.1). In this case, both the Diffusion Approximation and Exponential Approximation are well defined. As the population size becomes large, the approximations correctly predict that the invasion probability approaches zero. The corresponding analytical results are presented in Theorems 3.4 and 4.1, respectively, with their asymptotic limits indicated in the plot as black and gray “*”s. Limitations of the numerical integration procedure prevent us from displaying the Diffusion Approximation for the full range of population sizes (i.e. up to 10,000). The main difference between the two approximations in this regime is the rate of convergence to zero, with respect to increases in population size. Consistent with the Diffusion Approximation’s success for relatively small population sizes, it initially tracks the exact solution.

In the right panel of Figure 4, $b_1 < d_1$ and the dynamics are subcritical. In Theorem 4.1, we noted that the Exponential Approximation does not hold when $b < d$. As a direct consequence of equation (32), we further observe that the Exponential Approximation will be invalid for sufficiently small population sizes. In particular, when $d = b$ and $\beta = \delta$, we have

$$w_{e,0}(k\epsilon) = \exp\left(\frac{d_1 - b_1}{1 + 1/m}(k\epsilon)^{1/m} + O((k\epsilon)^{2/m})\right) > 1.$$  

Since the Exponential Approximation is $1 - w_e(k\epsilon)$, this returns an invalid value less than zero. From a practical point-of-view, one could simply define the Exponential Approximation to be zero in such a circumstance.

**5.4. Approximation success depends on the initial number of individuals introduced in the population.** When more than one individual is initially
introduced in a population, the probability of invasion increases. In Figure 5, we display results for each approximation method along with the exact solution for the probability of extinction when \(1 \leq k < Nx_*\) individuals are initially introduced in the population. By definition, for all larger values of \(k\), the invasion probability is one.

As shown in the top row of Figure 5, when the leading order terms match, the Diffusion Approximation performs well and tracks the exact solution over the full range of initial numbers of individuals introduced. Typically the Exponential Approximation with two terms (dark gray dashed line) better approximates the exact solution than the Exponential Approximation (light gray dashed line). However, when \(k\) is close to \(Nx_*\) the Exponential Approximation with two terms sharply turns up, away from the exact solution. This numerical result is in line with expectations from our analytical results in equation (28) since the additional higher order term is undefined for \(\rho(x) = 1\), i.e. when \(\lambda(x) = \mu(x)\).

In special cases, it is possible to compute the Exponential Approximation by hand using Definition 1.7. We validated the Exponential Approximation for the epidemic extinction probability (Example 1) by comparing the exact result for the leading order approximation,

\[
w_{r,0}(k\epsilon) = \left(\frac{d}{b(1-k\epsilon)}\right)^{k},
\]

with the numerical integration result plotted in Figure 5. There was not a visible difference between the value found by hand and the numerically obtained value.

We also compared the exact and numerical values for the Exponential Approximation with the analytical approximation found in Lemma 4.3. During this investigation, we found that since the range of potential number of introduced individuals scales with the total population size, it is important to keep track of the \(k\) parameter in the error term in equation (32). As \(k\) approaches \(Nx_*\), the error term becomes
significant (even for large population sizes). For fixed \( k \), the population size can be chosen large enough to yield an approximation with the desired level of accuracy.

6. Conclusion. Applied mathematicians, physicists, and biologists use a variety of techniques to approximate Markov chain models for population processes. The widespread use of SDEs, in particular, raises the question of whether these approximations faithfully reproduce results for basic probability questions, especially, hitting times and splitting probabilities. In this work we focused on studying the probability that the lineage of a few newly introduced individuals will invade a larger population. We analyzed the Diffusion Approximation and a simple Exponential Approximation rigorously in the large population limit and numerically for finite size populations. We have were able to show that both population size and the state of sub- versus super-criticality play an important role in determining which approximation method performs better.

Similar to recent analogous results for mean hitting times, we found that, in the large population limit, the Diffusion Approximation does not agree with the asymptotic invasion probability of the original Markov chain system (Theorem 3.4). In fact, when the leading order terms are mismatched (see the case when \( \beta \neq \delta \) in Theorem 3.4 and Example 2 in the bottom row of Figure 5), the Diffusion Approximation takes on a value that does not depend on the parameters of the Markov chain’s rate functions at all. Interestingly though, when the dynamics are near critical and the population of interest is small, we found that the Diffusion Approximation often performs quite well. This can be seen throughout the figures generated by our numerical investigation, which is described in Section 5. By contrast, for supercritical systems that are far from critical, the Exponential Approximation nearly matches the exact solution for large populations, while the Diffusion Approximation visibly misses the mark. The rigorous expression of this observation can be found in Theorem 4.1. There, we show that for asymptotically linear supercritical systems the Exponential Approximation provides the correct limiting result. This is displayed for a stochastic epidemic model (Example 1) and a resource-constrained population model (Example 3) in Figures 2 and 5.

Our work highlights that invasion probabilities are an important testing ground for evaluating approximation methods. Unlike mean hitting time calculations, the simplicity of the fundamental difference equations determining invasion probabilities permits the explicit evaluation of limits. At this stage, it remains an open question to completely understand which approximation is best for a given set of circumstances. It is possible, but not at all clear, that taking higher order approximations of the Exponential Approximation or a fully formulated WKB-type approximation would succeed under all conditions. In any case, it remains true that careful consideration should be taken when choosing an approximation method for evaluating properties of continuous-time, discrete state-space Markov chains.

Acknowledgements. The authors would like to thank José M. Ponciano and Sergei S. Pilyugin for thoughtful conversations in the development of this work.

Appendix A. Series Formulas.

Proposition A.1 (Series Expansion I). For a given power series with \( \gamma \geq 0 \) and \( \nu > 0 \), there exist coefficients \( \bar{a}_n \) such that

\[
\left( \sum_{n=0}^{\infty} a_n x^{n/m+\gamma} \right)^\nu = \sum_{n=0}^{\infty} \bar{a}_n x^{n/m+\gamma\nu},
\]
where the first few coefficients are
\[ \hat{a}_0 = a_0^\nu, \quad \hat{a}_1 = \nu a_0^{\nu-1} a_1, \quad \text{and} \quad \hat{a}_2 = \nu a_0^{\nu-1} \left( a_2 + \frac{(\nu - 1)a_1^2}{2a_0} \right). \]

**Proof.** The proof begins by factoring out the leading term in the given series
\[ \left( \sum_{n=0}^{\infty} a_n x^{n/m + \gamma} \right)^\nu = (a_0 x^\gamma)^\nu \left( 1 + \sum_{n=1}^{\infty} \frac{a_n}{a_0} x^{n/m} \right)^\nu. \]

Then make the substitution \( y = \sum_{n=1}^{\infty} \frac{a_n}{a_0} x^{n/m} \) and find the Taylor expansion of \((1 + y)^\nu\) at 1,
\[ (1 + y)^\nu = 1 + \nu y + \frac{\nu(\nu - 1)}{2} y^2 + O(y^3) \]
They by plugging in the series for \( y \) and collecting terms of the same power, we have
\[ \left( 1 + \sum_{n=1}^{\infty} \frac{a_n}{a_0} x^{n/m} \right)^\nu = 1 + \frac{\nu a_1}{a_0} x^{1/m} + \frac{\nu}{a_0} \left( a_2 + \frac{(\nu - 1)a_1^2}{2a_0^2} \right) x^{2/m} + O(x^{3/m}). \]

The form of the new series is found by multiplying each term by the original leading order term \((a_0 x^\gamma)^\nu\).

**Proposition A.2 (Series Expansion II).** Suppose a power series is given in the form, \( \sum_{n=0}^{\infty} a_n x^{n/m} \), with \( a_0 \neq 0, a_n \in \mathbb{R} \) for \( n \geq 1 \), and \( m > 0 \). Whenever \( x \) is such that \( \left| \sum_{n=1}^{\infty} a_n x^{n/m} \right| < |a_0| \), then
\[ \frac{1}{\sum_{n=0}^{\infty} a_n x^{n/m}} = \sum_{n=0}^{\infty} \hat{a}_n x^{n/m}, \]

where
\[ \hat{a}_0 = \frac{1}{a_0}, \quad \hat{a}_1 = \frac{-a_1}{a_0} \quad \text{and} \quad \hat{a}_2 = \frac{a_1^2 - a_2}{a_0}. \]

**Proof.** Suppose there is a given power series that satisfies the conditions of the proposition. Further suppose that the inequality \( \left| \sum_{n=1}^{\infty} a_n x^{n/m} \right| < |a_0| \) holds. By rearranging our original expression, we find a familiar form
\[ \frac{1}{\sum_{n=0}^{\infty} a_n x^{n/m}} = \frac{1}{a_0} \left( 1 + \frac{1}{a_0} \sum_{n=1}^{\infty} a_n x^{n/m} \right). \]

Since we have assumed that \( \left| \sum_{n=1}^{\infty} a_n x^{n/m} \right| < |a_0| \),
\[ \frac{1}{\sum_{n=0}^{\infty} a_n x^{n/m}} = \frac{1}{a_0} \left( \sum_{j=0}^{\infty} (-1)^j \left( \sum_{n=1}^{\infty} a_n x^{n/m} \right)^j \right) \]
\[ = \frac{1}{a_0} \left( 1 - (a_1 x^{1/m} + a_2 x^{2/m} + \ldots) + (a_1^2 x^{2/m} + \ldots) + \ldots \right) \]
\[ = \frac{1}{a_0} \left( 1 - a_1 x^{1/m} + (a_1^2 - a_2) x^{2/m} + \ldots \right) \]

The first few coefficients of the resulting series are obtained by distributing the \((a_0)^{-1}\) in the last line.

\[ \square \]
Proposition A.3 (Series Inversion [28]). Suppose $f(x)$ is increasing on $[0, x_*)$. Further assume that $f(0) = 0$ and $f(x)$ can be written as a formal power series:

$$f(x) = \sum_{n=0}^{\infty} a_n x^{n+\nu}.$$  

Let $t = f(x)$. Then the inverse of the series $f$ can be expressed as

$$x = \sum_{n=1}^{\infty} \tilde{a}_n t^{n/\nu}, \text{ as } t \to 0^+,$$

where

$$\tilde{a}_1 = \frac{1}{a_0^{1/\nu}} \text{ and } \tilde{a}_2 = -\frac{a_1}{\nu a_0^{1+2/\nu}}.$$  

Proof. We first rewrite $f(x)$ as follows

$$f(x) = x^{\nu} \sum_{n=0}^{\infty} a_n x^n.$$  

Setting $t = f(x)$, we have

$$t^{1/\nu} = xg(x), \text{ where } g(x) := \left( \sum_{n=0}^{\infty} a_n x^n \right)^{1/\nu}.$$  

Then, by Proposition A.1, we have

$$g(x) = \sum_{n=0}^{\infty} \tilde{a}_n x^n,$$

where

$$\tilde{a}_0 = a_0^{1/\nu}, \quad \tilde{a}_1 = \frac{a_0^{1/\nu-1} a_1}{\nu}, \quad \tilde{a}_2 = a_0^{1/\nu-1} \left( \frac{1}{\nu} a_1^2 \right) \frac{1}{2a_0} + a_2.$$  

Then,

$$t^{1/\nu} = a_0^{1/\nu} x \left( 1 + \frac{a_1}{\nu a_0} x + \frac{1}{\nu a_0} \left( \frac{1}{\nu} a_1^2 - \frac{1}{2a_0} + a_2 \right) x^2 + O(x^3) \right),$$

and

$$t^{2/\nu} = a_0^{2/\nu} x^2 \left( 1 + \frac{2a_1}{\nu a_0} x + \left( \frac{2}{\nu a_0} \left( \frac{1}{\nu} a_1^2 - \frac{1}{2a_0} + a_2 \right) + \frac{a_1}{\nu a_0} \right) x^2 + O(x^3) \right).$$

Choose $\tilde{a}_1 = a_0^{-1/\nu}$ so that it cancels the leading coefficient of $t^{1/\nu}$. From here we use powers of $t^{1/\nu}$ to find an expansion of the form

$$x \sim \sum_{n=1}^{\infty} \tilde{a}_n t^{n/\nu}.$$  

By collecting ordered terms, we find the corresponding two term approximation,

$$\tilde{a}_1 t^{1/\nu} + \tilde{a}_2 t^{2/\nu} = x + \left( \frac{a_1}{\nu a_0} + \tilde{a}_2 a_0^{2/\nu} \right) x^2 + O(x^3).$$
We then choose
\[ \tilde{a}_2 = -\frac{a_1}{\nu a_1^{1+2/\nu}} \]
so that the coefficient for the \( x^2 \) term is zero. Increasingly precise approximations of \( x \) can be found by keeping track of the higher order \( t \) terms and choosing each \( \tilde{a}_n \) to appropriately cancel out these terms.

REFERENCES
[1] L. J. Allen, An introduction to stochastic epidemic models, in Mathematical Epidemiology, Springer, 2008, pp. 81–130.
[2] L. J. Allen and E. J. Allen, A comparison of three different stochastic population models with regard to persistence time, Theor. Popul. Biol., 64 (2003), pp. 439–449.
[3] D. F. Anderson and D. J. Higham, Multilevel monte carlo for continuous time markov chains, with applications in biochemical kinetics, Multiscale Model. Simul., 10 (2012), pp. 146–179.
[4] F. Ball, The threshold behaviour of epidemic models, J. Appl. Probab., 20 (1983), pp. 227–241.
[5] F. Ball and P. Donnelly, Strong approximations for epidemic models, Stochastic Process. Appl., 55 (1995), pp. 1–21.
[6] H. C. Berg, Random walks in biology, Princeton University Press, 1993.
[7] R. K. Borcherding, S. E. Bellan, J. M. Flynn, J. R. C. Pulliam, and S. A. McKinley, Resource-driven encounters among consumers and implications for the spread of infectious disease, Journal of The Royal Society Interface, 14 (2017).
[8] P. C. Bressloff and J. M. Newby, Stochastic hybrid model of spontaneous dendritic nmda spikes, Phys. Biol., 11 (2014), p. 016006.
[9] J. M. Conway, B. P. Konrad, and D. Coombs, Stochastic analysis of pre-and postexposure prophylaxis against hiv infection, SIAM Journal on Applied Mathematics, 73 (2013), pp. 904–928.
[10] S. Ditlevsen and P. Greenwood, The morris-lear neuron model embeds a leaky integrate-and-fire model, J. Math. Biol., 67 (2013), pp. 239–259, https://doi.org/10.1007/s00285-012-0552-7.
[11] C. R. Doering, K. V. Sargsyan, and L. M. Sander, Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the fokker-planck approximation, Multiscale Model. Simul., 3 (2005), pp. 283–299.
[12] C. R. Doering, K. V. Sargsyan, L. M. Sander, and E. Vandenberg, Asymptotics of rare events in birth-death processes bypassing the exact solutions, J. Phys. Condens. Matter, 19 (2007), p. 065145.
[13] D. T. Gillespie et al., Exact stochastic simulation of coupled chemical reactions, J. Phys. Chem., 81 (1977), pp. 2340–2361.
[14] V. Giorno, P. Lansky, A. G. Noble, and L. M. Ricciardi, Diffusion approximation and first-passage-time problem for a model neuron, Biol. Cybern., 58 (1988), pp. 387–404, https://doi.org/10.1007/BF00361346.
[15] A. M. Hein and S. A. McKinley, Sensory information and encounter rates of interacting species, PLoS Comput. Biol., 9 (2013), p. e1003178.
[16] J. P. Keener and J. M. Newby, Perturbation analysis of spontaneous action potential initiation by stochastic ion channels, Phys. Rev. E, 84 (2011), p. 011918.
[17] M. Kimura, Diffusion models in population genetics, Journal of Applied Probability, 1 (1964), pp. 177–232.
[18] F. C. Klebaner, Introduction to stochastic calculus with applications, vol. 57, World Scientific, 2005.
[19] T. G. Kurtz, Approximation of population processes, vol. 36, SIAM, 1981.
[20] A. Lambert, The branching process with logistic growth, Ann. Appl. Probab., 15 (2005), pp. 1506–1535.
[21] C. Lopez-Garcia, A. M. Klein, B. D. Simon, and D. J. Winton, Intestinal stem cell replacement follows a pattern of neutral drift, Science, 330 (2010), pp. 822–825.
[22] J. M. Newby, P. C. Bressloff, and J. P. Keener, Breakdown of fast-slow analysis in an excitable system with channel noise, Phys. Rev. Lett., 111 (2013), p. 128101.
[23] J. R. Norris, Markov Chains, no. 2, Cambridge University Press, 1998.
[24] K. Pakdaman, M. Thieullen, and G. Wainrib, Diffusion approximation of birth-death processes: Comparison in terms of large deviations and exit points, Stat. Probab. Lett., 80 (2010), pp. 1121–1127.
[25] H. J. Snipperit, L. G. Van Der Flier, T. Sato, J. H. Van Es, M. Van Den Born, C. Kroon-Veenboer, N. Barker, A. M. Klein, J. Van Rheenen, B. D. Simons, et al., *Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells*, Cell, 143 (2010), pp. 134–144.

[26] C. Tier and F. B. Hanson, *Persistence in density dependent stochastic populations*, Mathematical Biosciences, 53 (1981), pp. 89–117.

[27] D. Wodarz and N. Komarova, *Computational biology of cancer: lecture notes and mathematical modeling*, World Scientific, 2005.

[28] R. Wong, *Asymptotic Approximations of Integrals*, SIAM, 2001.

[29] H. Youssefpour, X. Li, A. Lander, and J. Lowengrub, *Multispecies model of cell lineages and feedback control in solid tumors*, Journal of theoretical biology, 304 (2012), pp. 39–59.