Mathematical Biology

Evolutionary dynamics of complex traits in sexual populations in a heterogeneous environment: how normal?

Léonard Dekens[1,2]

Abstract

When studying the dynamics of trait distribution of populations in a heterogeneous environment, classical models from quantitative genetics choose to look at its system of moments, specifically the first two ones. Additionally, in order to close the resulting system of equations, they often assume the local trait distributions are Gaussian [see for instance Ronce and Kirkpatrick (Evolution 55(8):1520–1531, 2001. https://doi.org/10.1111/j.0014-3820.2001.tb00672.x.37)]. The aim of this paper is to introduce a mathematical framework that follows the whole trait distribution (without prior assumption) to study evolutionary dynamics of sexually reproducing populations. Specifically, it focuses on complex traits, whose inheritance can be encoded by the infinitesimal model of segregation (Fisher in Trans R Soc Edinb 52(2):399–433, 1919. https://doi.org/10.1017/S0080456800012163). We show that it allows us to derive a regime in which our model gives the same dynamics as when assuming Gaussian local trait distributions. To support that, we compare the stationary problems of the system of moments derived from our model with the one given in Ronce and Kirkpatrick (Evolution 55(8):1520–1531, 2001. https://doi.org/10.1111/j.0014-3820.2001.tb00672.x.37) and show that they are equivalent under this regime and do not need to be otherwise. Moreover, under this regime of equivalence, we show that a separation bewteen ecological and evolutionary time scales arises. A fast relaxation toward monomorphism allows us to reduce the complexity of the system of moments, using a slow-fast analysis. This reduction leads us to complete, still in this regime, the analytical description of the bistable asymmetrical equilibria numerically found in Ronce and Kirkpatrick (Evolution 55(8):1520–1531, 2001. https://doi.org/10.1111/j.0014-

---

The codes to reproduce the figures of this article are available at https://github.com/ldekens/Evolutionary-dynamics-of-complex-traits-in-sexual-populationsin-a-heterogeneous-environment.

✉ Léonard Dekens
dekens@math.univ-lyon1.fr

[1] Institut Camille Jordan, UMR 5208 UCBL/CNRS, Université de Lyon, Villeurbanne, France

[2] DRACULA Project Team, INRIA Grenoble - Rhône-Alpes, Institut Camille Jordan, Villeurbanne, France

3820.2001.tb00672.x.37). More globally, we provide explicit modelling hypotheses that allow for such local adaptation patterns to occur.

Keywords Quantitative genetic model · Heterogeneous environment · Asymptotic analysis · Infinitesimal model · Sexual reproduction

Mathematics Subject Classification 35Q92 · 92D15 · 35R09 · 34E13

Contents

Introduction

Most species occupy heterogeneous environments, in which the spatial structure is expected to play a significant role in the evolution of the diversity of a species. As a result of the balance between the mixing effect of migration connecting the different habitats of a species and the selective pressure reducing diversity within each habitat, several equilibrium states encoding the local adaptation of a species can be reached. Will the species succeed to persist in a wide range of habitat available and thus thrive as a generalist species? Will it become adapted to specific sets of conditions as what we
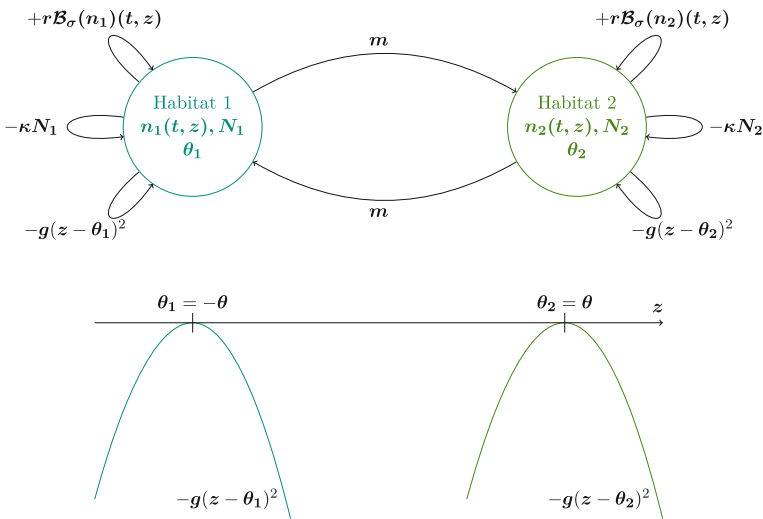
Fig. 1 Heterogeneous symmetrical environment framework for a quantitative trait z. The upper part of the figure illustrates the different biological forces acting in each habitat (reproduction, competition for resources, selection) and between them (migration). The lower part of the figure draws the local quadratic selection functions considered, where θ_1 and θ_2 are the local optimal traits. The parameters are the same in both habitats, except for the local optimal traits

call a specialist species? Evolutionary biology fields have taken a sustained interest in these questions, in population genetics (Lythgoe 1997; Nagylaki and Lou 2001; Bürger and Akerman 2011; Akerman and Bürger 2014), adaptive dynamics (Meszéna et al. 1997; Day 2000) or quantitative genetics (Tufto 2000; Ronce and Kirkpatrick 2001; Hendry et al. 2001; Yeaman and Guillaume 2009; Débarre et al. 2013, 2015; Mirrahimi 2017; Lavigne 2019; Mirrahimi and Gandon 2020). Here we adopt the framework of quantitative genetics, which models the adaptation of a continuous trait without giving explicitly its underlying genetic architecture. Additionally, we specifically choose to analyse the influence of sexual reproduction as mating system.

Model We build our model within a biological framework shared with classical studies (Ronce and Kirkpatrick 2001; Hendry et al. 2001; Débarre et al. 2013). We consider a sexual population whose individuals are characterized by a quantitative phenotypic trait $z \in \mathbb{R}$ and evolving in a heterogeneous environment constituted by two habitats that we will assume to be symmetric (i.e, sharing the same ecological parameters except for their optimal traits), as illustrated in Fig. 1.

The density of population at a given time t with respect to a phenotype z in habitat $i \in \{1, 2\}$ is denoted $n_i(t, z) \in L^1(\mathbb{R}_+ \times \mathbb{R})$, for which we further assume that $z^k n_i(t, z) \in L^1(\mathbb{R}_+ \times \mathbb{R})$ for $k < 4$.

Local maladaptation is the source of mortality in our model: stabilizing selection acts quadratically in each patch toward an optimal phenotype $\theta_i \in \mathbb{R}$ with an intensity $g > 0$. Define θ as half the distance between the two local optima: $\theta := \frac{|\theta_2 - \theta_1|}{2}$. Up to a translation in the phenotypic space, we can consider without loss of generality that $0 < \theta_2 = -\theta_1 = \theta$. Additionally, competition for resources regulates the total size

of the subpopulation $N_i(t) = \int_{\mathbb{R}} n_i(t, z') \, dz'$ in each patch with an intensity $\kappa > 0$. The mortality rate of an individual with phenotypic trait $z \in \mathbb{R}$ is thus given by:

$$M[n_i(t, z)] = -g(z - \theta_i)^2 - \kappa N_i.$$

Migration between the two patches occurs symmetrically at a rate $m > 0$. The exchange of individuals from patch i to patch j of a given phenotype $z \in \mathbb{R}$ at time $t \geq 0$ is thereby:

$$m \left(n_j(t, z) - n_i(t, z) \right).$$

Finally, we denote by $\mathcal{B}_\sigma(n_i)(t, z)$ the number of new individuals that are born at time $t \geq 0$ in patch i with a phenotype $z \in \mathbb{R}$ due to sexual reproduction. That phenomenon is occurring at a rate $r > 0$, and the parameter σ is a measure of the segregational variance linked to the trait inheritance process. The sexual reproduction operator is at this point still unspecified and will be defined below. However, we will consider that it respects the following conservative properties:

$$\forall t \in \mathbb{R}_+, \int_{\mathbb{R}} \mathcal{B}_\sigma(n_i)(t, z) \, dz = \int_{\mathbb{R}} n_i(t, z) \, dz, \quad \int_{\mathbb{R}} z \mathcal{B}_\sigma(n_i)(t, z) \, dz$$
$$= \int_{\mathbb{R}} z \, n_i(t, z) \, dz.$$

The dynamics of the local trait distributions are therefore given by:

$$\begin{cases} \frac{\partial n_1}{\partial t}(t, z) = r\mathcal{B}_\sigma(n_1)(t, z) - g(z - \theta_1)^2 n_1(t, z) - \kappa N_1(t)n_1(t, z) + m\left(n_2(t, z) - n_1(t, z)\right), \\ \\ \frac{\partial n_2}{\partial t}(t, z) = r\mathcal{B}_\sigma(n_2)(t, z) - g(z - \theta_2)^2 n_2(t, z) - \kappa N_2(t)n_2(t, z) + m\left(n_1(t, z) - n_2(t, z)\right). \end{cases}$$
$$(1)$$

System of moments and Gaussian assumption Quantitative genetics studies often model the dynamics of the sizes of the subpopulations $N_1 > 0$ and $N_2 > 0$ and their mean traits $\overline{z}_1$ and $\overline{z}_2$ (where $N_i := \int_{\mathbb{R}} n_i(t, z) \, dz$ and $\overline{z}_i := \frac{1}{N_i} \int_{\mathbb{R}} z \, n_i(t, z) \, dz$). Although we intend to follow the dynamics of the whole trait distributions, for the sake of comparison, we derive ordinary differential equations for the first moments of the trait distributions by integrating (1) with regard to z:

$$\begin{cases} \frac{dN_i}{dt} = \left[r - \kappa N_i(t) - g(\overline{z}_i(t) - \theta_i)^2 - g\sigma_i^2\right] N_i(t) + m\left(N_j(t) - N_i(t)\right), \\ \\ \frac{d\overline{z}_i}{dt} = 2\sigma_i^2 g(\theta_i - \overline{z}_i(t)) - g\psi_i^3 + m\frac{N_j(t)}{N_i(t)}(\overline{z}_j(t) - \overline{z}_i(t)). \end{cases}$$
$$(2)$$

where $\sigma_i^2 := \frac{1}{N_i} \int_{\mathbb{R}} (z - \overline{z}_i)^2 \, n_i(t, z) \, dz$ and $\psi_i^3 := \int_{\mathbb{R}} \frac{1}{N_i} \int (z - \overline{z}_i)^3 \, n_i(t, z) \, dz$ are respectively the variance and the third central moment of the trait distribution of each subpopulation (see "Appendix A" for details about the derivation). At this point, a common key assumption used to close the system that arises in quantitative genetics

models is the normality of such a trait distribution, with a constant variance (Hendry et al. 2001; Ronce and Kirkpatrick 2001). In Ronce and Kirkpatrick (2001), such an assumption results in the following system (with their original notations for the parameters):

$$
\begin{cases}
\frac{dN_i}{dt} = \left[r_0(1 - \frac{N_i}{K}) - \frac{\gamma}{2}\sigma_p{}^2 - \frac{\gamma}{2}(\overline{z_i} - \theta_i)^2 \right] N_i + m(N_j - N_i), \\[2ex]
\frac{d\overline{z_i}}{dt} = \sigma_g{}^2 \gamma(\theta_i - \overline{z_i}) + m\frac{N_j}{N_i}(\overline{z_j} - \overline{z_i}).
\end{cases}
$$

where $\sigma_p{}^2$ and $\sigma_g{}^2$ are respectively the constant phenotypic and genotypic variance, differing additively by a constant variance due to environmental effects $\sigma_e{}^2$ ($\sigma_p{}^2 = \sigma_g{}^2 + \sigma_e{}^2$). With this method, the authors of Ronce and Kirkpatrick (2001) analyse the equilibria of the system above, by distinguishing two types of equilibrium:

- Symmetrical equilibrium, where both local populations have equal size and are equally maladapted to their local habitat. The species survives in both habitats, and is therefore characterized as a generalist species. The authors derived this equilibrium analytically.
- Asymmetrical equilibria, where the species mainly inhabits one habitat to which it is adapted. It acts as a source for the other habitat that is almost deserted, if it were not for a few unsuccessful migrants, sent from the first habitat, and therefore poorly adapted to the second one (the sink). This type of equilibrium characterizes a specialist species, that can only live in a restricted set of environments . The authors numerically explored this type of equilibrium and derived approximations for low migration rates.

However, this approach disregards the effect of higher moments of the trait distribution (like the skewness), that may become significant due to the presence of gene flow, as pointed out in Yeaman and Guillaume (2009) and Débarre et al. (2015).

The infinitesimal model of sexual reproduction To account for the influence of higher moments calls for models bypassing any prior assumption on the trait distribution, both to assess the validity of the Gaussian approximation or examine the departure from it. Therefore, it is necessary to make explicit the interplay between sexual reproduction and phenotypic inheritance. The infinitesimal model of sexual reproduction, first introduced by Fisher (1919) offers a simple way to tackle this issue for complex traits. Consequently, it has been used both in several biological studies [under truncation selection in Turelli and Barton (1994), or in a continent-island model in Tufto (2000)] and mathematical ones (Mirrahimi and Raoul 2013; Bourgeron et al. 2017; Raoul 2017). Aligning with these, we choose it in our study to model trait inheritance due to sexual reproduction. The classical version of this model translates the stochasticity of the segregation process by the fact that the offspring trait variable $\mathcal{Z}$ (conditioned to the parental traits $\mathcal{Z}_1 = z_1$ and $\mathcal{Z}_2 = z_2$) follows a Gaussian law centered in the mean parental trait and with a segregational variance of $\frac{\sigma^2}{2}$:

$$
\mathcal{Z}|\{\mathcal{Z}_1 = z_1, \mathcal{Z}_2 = z_2\} \sim \frac{z_1 + z_2}{2} + \mathcal{N}\left(0, \frac{\sigma^2}{2}\right). \tag{3}
$$

Consequently, this model makes a normal assumption, not on the distribution of trait in the population, but on the distribution of offspring within each family, with a fixed and constant segregational variance (Turelli 2017). A common Mendelian interpretation of this mixing model is that the trait results from the expression of a large number of alleles with small additive effects (Fisher 1919; Bulmer 1971; Lange 1978). Recently, a rigorous framework of the use of that model in various biological contexts has been derived in Barton et al. (2017).

The regime of small variance $\sigma^2 \ll \theta^2$ There also has been increasing mathematical interest in developing integro-differential equations for the whole trait distribution to study qualitatively quantitative genetics models (Magal and Webb 2000; Diekmann et al. 2005; Desvillettes et al. 2008). A framework introduced by Diekmann et al. (2005) to study asexual models in the regime of small mutations led to first rigorous results in Perthame and Barles (2008) in the context of homogeneous environment. Next, it has been extended to study spatially heterogeneous environment where asexual species evolve, like in Mirrahimi (2017) that successfully characterizes the equilibrium states by using a Hamilton–Jacobi approach in the limit of small mutations. For sexually reproducing populations, using the infinitesimal model in an asymptotic regime allowed Mirrahimi and Raoul (2013) to study invasions by phenotypically structured populations. More recently, using the infinitesimal model in a small variance regime led Bouin et al. (2018) to formally derive features of the underlying trait distribution of a population under a changing environment. Their formal derivations have next been justified in a homogeneous space framework in Calvez et al. (2019). Our work aligns with these studies: our main analysis lies in the small variance regime: $\sigma^2 \ll \theta^2$, namely when the diversity introduced by sexual reproduction is small compared to the heterogeneity of the environment (recall that $\theta = \frac{|\theta_2 - \theta_1|}{2}$).

Contributions We use the infinitesimal model operator and the formalism of small segregational variance to study evolutionary dynamics of a sexually reproducing population under stabilizing selection in a heterogeneous and symmetrical environment in an integrated model (Sect. 1). From the PDE system on the local trait distributions, we derive a system of ODE on their moments. In the particular asymptotic regime considered: $\sigma^2 \ll \theta^2$, our ODE system approximates the one of Ronce and Kirkpatrick (2001) (Sect. 1):

$$\begin{cases} \frac{dN_i}{dt} = \left[r - \kappa N_i(t) - g(\overline{z}_i(t) - \theta_i)^2 - g\sigma^2 \right] N_i(t) + m\left(N_j(t) - N_i(t)\right) + \mathcal{O}\left(\frac{\sigma^4}{\theta^4}\right), \\[2mm] \frac{d\overline{z}_i}{dt} = 2\sigma^2 g(\theta_i - \overline{z}_i(t)) + m \frac{N_j(t)}{N_i(t)} (\overline{z}_j(t) - \overline{z}_i(t)) + \mathcal{O}\left(\frac{\sigma^4}{\theta^4}\right). \end{cases} \tag{4}$$

To support that, we provide a numerical comparison between the two models, showing their equivalence in the small variance regime, and their discrepancy when this variance becomes large (Sect. 2). By doing so, we are justifying the validity of the Gaussian assumption on local trait distributions in this small variance regime. Next, we show that, in the regime of small variance, our system of moments can be reformulated as a slow-fast system (Sect. 3), which highlights the blending force of our sexual reproduction operator that strains monomorphism to quickly emerge at the metapopulation level.

The study of the corresponding unperturbed problem, with a reduced complexity, leads to the complete analytical description of the equilibria in the asymptotic regime of small variance. In particular, it gives the conditions of existence of bistable asymmetrical equilibria numerically observed by Ronce and Kirkpatrick (2001) (Sect. 4).

To replace this study in a broader context, let us first recall some findings of Ronce and Kirkpatrick (2001), our reference moment-based model in the quantitative genetics field. It makes a Gaussian assumption on the local trait distributions, without specifying any particular mode of reproduction. The authors numerically found that bistable mirrored asymmetrical equilibria can exist, allowing source-sink dynamics to completely reverse after a demographical loss event. Based on their study, however, it remains unclear which hypotheses on the inheritance process allow for such dynamics to arise. More recently, two studies interested in the equilibria states of asexual populations highlight the need for precise hypotheses with regard to such conclusions. If the authors of Débarre et al. (2013) indicate that asymmetrical equilibria can be locally stable in a restrained range of mutational parameters, Mirrahimi (2017) and Mirrahimi and Gandon (2020) show through using a continuum-of-alleles model that, under broader mutational parameters, only a single stable symmetrical equilibrium can arise in a symmetrical setting. Here, we claim that we can explain the dynamics of the analysis done in Ronce and Kirkpatrick (2001) via a model on phenotypic densities dynamics, analogous to Mirrahimi (2017) and Mirrahimi and Gandon (2020) but with a sexual reproduction operator derived from the infinitesimal model and in a small segregational variance regime. We thereby make explicit the details of another mechanism that can provide with those locally bistable asymmetrical equilibria, which relies on the blending effect of the infinitesimal model in a regime of small segregational variance.

1 The infinitesimal model and the regime of small variance

In this section, we present the specific framework in which we choose to perform our analysis. We first present some properties of the infinitesimal model operator in general, then its relationship with the specific regime of small variance. Then, we will show that the asymptotic approximation allows us to formally derive a closed system for the dynamics of the moments.

Let us define the following rescaled variables and parameters to get a dimensionless system:

$$z := \frac{z}{\theta}, \quad g := \frac{g\theta^2}{r}, \quad m := \frac{m}{r}, \quad \varepsilon := \frac{\sigma}{\theta}, \quad t := \varepsilon^2 rt,$$

$$n_{\varepsilon,i}(t, z) := \frac{\kappa}{r} n_i(t, z), \quad N_{\varepsilon,i}(t) = \frac{\kappa}{r} N_i(t),$$

and the reproduction operator $\mathcal{B}_\varepsilon(n_{\varepsilon,i})(t, z) = \mathcal{B}_{\sigma(n_i)(t,z)}$. Then, (1) gives the rescaled system:

$$\begin{cases} \varepsilon^2 \frac{\partial n_{\varepsilon,1}}{\partial t}(t,z) = \mathcal{B}_\varepsilon(n_{\varepsilon,1})(t,z) - g(z+1)^2 n_{\varepsilon,1}(t,z) - N_{\varepsilon,1}(t)n_{\varepsilon,1}(t,z) + m\left(n_{\varepsilon,2}(t,z) - n_{\varepsilon_1}(t,z)\right), \\[2mm] \varepsilon^2 \frac{\partial n_{\varepsilon,2}}{\partial t}(t,z) = \mathcal{B}_\varepsilon(n_{\varepsilon,2})(t,z) - g(z-1)^2 n_{\varepsilon,2}(t,z) - N_{\varepsilon,2}(t)n_{\varepsilon,2}(t,z) + m\left(n_{\varepsilon,1}(t,z) - n_{\varepsilon,2}(t,z)\right). \end{cases}$$
(5)

From the remaining of this section and unless specified otherwise, we will refer to that system for all analysis purposes.

1.1 The sexual reproduction operator

Presentation For modelling the segregation process resulting from sexual reproduction, we use the infinitesimal model, first introduced in Fisher (1919). It is inspired originally from the observation that the phenotypic variance among families does not seem to depend on their breeding values (Galton 1877). Although this can be formulated solely from a phenotypic perspective, Fisher (1919) gives a Mendelian interpretation by proposing to consider that the quantitative trait z results from the infinitesimally small additive effects of a large number of alleles. That interpretation, in the spirit of a central limit theorem, has been followed on (Bulmer 1971, 1980; Lange 1978; Barton et al. 2017). It leads to (3). With our notations, we can express the number of individuals born at time t with trait z in habitat i by:

$$\mathcal{B}_\varepsilon(n_\varepsilon)(t,z) = \frac{1}{\sqrt{\pi}\varepsilon} \int_{\mathbb{R}^2} \exp\left[\frac{-(z - \frac{z_1+z_2}{2})^2}{\varepsilon^2}\right] n_\varepsilon(t,z_1)\frac{n_\varepsilon(t,z_2)}{N_\varepsilon(t)} dz_1 dz_2. \quad (6)$$

The scaled segregational variance $\frac{\varepsilon^2}{2}$ is assumed to be constant with regard to time and independent of the parental traits. These are strong biological assumptions. Their relevance in the context of a spatially structured population will be the subject of a forthcoming work.

Equilibria under random mating only To study the behaviour of the reproduction operator (6), it is informative to consider the conservative case where a sexually reproducing population only experiences random mating, without any structure due to space or mating preferences:

$$\varepsilon^2 \frac{\partial n_\varepsilon}{\partial t}(t,z) = \frac{1}{\sqrt{\pi}\varepsilon} \int_{\mathbb{R}^2} \exp\left[\frac{-(z - \frac{z_1+z_2}{2})^2}{\varepsilon^2}\right] n_\varepsilon(t,z_1)\frac{n_\varepsilon(t,z_2)}{N_\varepsilon(t)} dz_1 dz_2 \\ - n_\varepsilon(t,z), \quad (7)$$

(the term $-n_\varepsilon(t,z)$ is meant to keep the size of the population constant by balancing birth and death). Then, every Gaussian distribution of variance ε^2 (arbitrarily centered) is a stable distribution under (7) (see "Appendix B"). Furthermore, it is shown in Raoul (2017) that there are no other equilibrium and that the convergence toward such a Gaussian distribution is exponential in quadratic Wasserstein distance. Therefore, with this operator of sexual reproduction, a fixed and finite variance in trait at equilibrium arises under random mating only and without selection.

1.2 The regime of small variance: $\varepsilon^2 \ll 1$

The framework presented in this section is inspired by a methodology developed in Diekmann et al. (2005) and Perthame and Barles (2008) that uses asymptotic regime in partial differential equations in order to derive analytical features of quantitative genetics models. In a regime where few diversity is introduced by reproduction at each generation, the continuous trait distributions are expected to converge toward Dirac masses concentrated on some specific traits. Performing a suitable transformation on the trait distribution allows to unfold the singularities of these Dirac masses and define more regular objects to study and calculate, in order to follow trait densities. That methodology has already been successfully applied for asexual populations, in homogeneous (Perthame and Barles 2008) and heterogeneous space (Mirrahimi 2017), then in other frameworks such as the study of adaptation to a changing environment (Bouin et al. 2018), and lately for sexual populations in homogeneous space (Calvez et al. 2019). Applying a similar approach as described above, we will show that, within a regime of small variance yet to be defined, we can reduce the complexity of the system while rigorously justify that reduction.

In our context, a relative measure of diversity introduced by reproduction comes from comparing the variance of the segregation process to a measure of habitats' difference (recall that $\theta = \frac{|\theta_2 - \theta_1|}{2}$):

$$\frac{\sigma^2}{\theta^2} = \varepsilon^2.$$

One can thus define the small variance regime by $\sigma^2 \ll \theta^2$, or equivalently $\varepsilon^2 \ll 1$. Moreover, we perform the unfolding of singularities by shaping the traits distributions according to:

$$n_{\varepsilon,i} = \frac{1}{\sqrt{2\pi}\varepsilon} e^{-\frac{U_{\varepsilon,i}}{\varepsilon^2}}. \tag{8}$$

The exponential form, known as the Hopf–Cole transform in scalar conservation laws, presumes that $U_{\varepsilon,i}$ will be a more regular object to analyze when $\varepsilon^2 \ll 1$ than $n_{\varepsilon,i}$, which we expect to converge toward a sum of Dirac distributions centered at the minima of $U_{\varepsilon,i}$. In fact, Bouin et al. (2018) performed a formal analysis on the behaviour of the reproduction term in the regime of small variance under such a formalism. They found that, for the various contributions to be well-balanced in the equation (reproduction and mortality) when $\varepsilon^2 \ll 1$, $U_{\varepsilon,i}$ is formally constrained to have the following expansion with regard to successive powers of ε^2 (see "Appendix C"):

$$U_{\varepsilon,i}(z) = \frac{(z - z_i^*)^2}{2} + \varepsilon^2 u_{\varepsilon,i}, \tag{9}$$

where z_i^* is a byproduct of the formal analysis and $u_{\varepsilon,i}$ is the following order term in the expansion. It leads to:

$$n_{\varepsilon,i} = \frac{1}{\sqrt{2\pi}\,\varepsilon} e^{-\frac{(z-z_i^*)^2}{2\varepsilon^2}} e^{-u_{\varepsilon,i}(z)}. \tag{10}$$

Let us interpret this formalism. For $\varepsilon^2 \ll 1$, the leading term in the expansion (10) is precisely the Gaussian distribution of (yet unknown) mean z_i^* and variance ε^2, namely a distribution we know to be at equilibrium under random mating only. Only considering this term would be to assume that the trait distribution is Gaussian. As we want to capture the departure from normality, we introduce the term $u_{\varepsilon,i}$, which we can see as the next order term in the expansion of $\log(n_{\varepsilon,i})$ with regard to successive powers of ε. It embodies the correction to the Gaussian distribution due to the effect of selection, competition and migration. The study of its analytical properties is beyond the scope of this paper and will be the project of a forthcoming paper. For now, we will assume that such a limit exist and we will use it in our analysis without rigorously justifying it.

1.3 Derivation of the dynamics of the moments in the regime of small variance

Although our method describes directly the trait distribution, we propose to formally derive the equations describing the dynamics of the first three moments of the trait distribution from its dynamics under the small variance of segregation ($\varepsilon^2 \ll 1$) to compare our framework to other quantitative genetic studies. Toward that purpose, we define (assuming persistence of each subpopulation):

$$N_{\varepsilon,i}(t) = \int_{\mathbb{R}} n_{\varepsilon,i}(t,z)\,dz, \quad \bar{z}_{\varepsilon,i}(t) = \frac{1}{N_{\varepsilon,i}} \int_{\mathbb{R}} z\,n_{\varepsilon,i}(t,z)\,dz,$$

$$\sigma_{\varepsilon,i}^2(t) = \frac{1}{N_{\varepsilon,i}} \int_{\mathbb{R}} \left(\bar{z}_{\varepsilon,i} - z\right)^2 n_{\varepsilon,i}(t,z)\,dz, \quad \psi_\varepsilon^3 = \frac{1}{N_\varepsilon} \int_{\mathbb{R}} (z - \bar{z}_{\varepsilon,i})^3 n_\varepsilon(z)\,dz. \tag{11}$$

Let us omit for a moment the time dependency. Using the expression (10) and under the formal assumption that $u := \lim_{\varepsilon \to 0} u_\varepsilon$ is sufficiently regular, we get the following expansions (where $v_{i,\varepsilon}$ is the expansion term of order ε^4 of $U_{\varepsilon,i}$—see "Appendix D"):

$$\begin{cases} N_{\varepsilon,i} = e^{-u_i(z_i^*)}\left[1 + \varepsilon^2\left(\frac{(\partial_z u_i(z_i^*))^2}{2} - \frac{\partial_{zz} u_i(z_i^*)}{2} - v_{i,\varepsilon}(z_i^*)\right)\right] + \mathcal{O}(\varepsilon^4), \\ \bar{z}_{\varepsilon,i} = z_i^* - \varepsilon^2 \partial_z u_i(z_i^*) + \mathcal{O}(\varepsilon^4), \\ \sigma_{\varepsilon,i}^2 = \varepsilon^2 + \mathcal{O}(\varepsilon^4), \\ \psi_{\varepsilon,i}^3 = \mathcal{O}(\varepsilon^4). \end{cases} \tag{12}$$

These expansions are informative, particularly the one describing the rescaled variances of the trait distributions $\sigma_{\varepsilon,1}^2$ and $\sigma_{\varepsilon,2}^2$ (third line of (12)). We can observe that they are both equivalent to twice the rescaled segregational variance $\frac{\varepsilon^2}{2}$ (which is given as a parameter of the model—see (6)) when the latter is small. The local rescaled variances in trait $\sigma_{\varepsilon,1}^2$ and $\sigma_{\varepsilon,2}^2$ are thereby asymptotically constant and independent of the local environment.

Now, from scaling (2), we obtain:

$$
\begin{cases}
\varepsilon^2 \frac{dN_{\varepsilon,i}}{dt} = \left[1 - N_{\varepsilon,i}(t) - g(\overline{z}_{\varepsilon,i}(t) - (-1)^i)^2 - g\,\sigma_i(t)^2\right] N_{\varepsilon,i}(t) + m\big(N_{\varepsilon,j}(t) - N_{\varepsilon,i}(t)\big), \\[2mm]
\varepsilon^2 \frac{d\overline{z}_{\varepsilon,i}}{dt} = 2g\,\sigma_i(t)^2((-1)^i - \overline{z}_{\varepsilon,i}(t)) - g\,\psi_i^3(t) + m\,\frac{N_{\varepsilon,j}(t)}{N_{\varepsilon,i}(t)}(\overline{z}_{\varepsilon,j}(t) - \overline{z}_{\varepsilon,i}(t)).
\end{cases}
$$

Next, using the formal expansions of the variances and skews given by (12) when $\varepsilon^2 \ll 1$ yields:

$$
\begin{cases}
\varepsilon^2 \frac{dN_{\varepsilon,i}}{dt} = \left[1 - N_{\varepsilon,i}(t) - g(\overline{z}_{\varepsilon,i}(t) - (-1)^i)^2 - g\varepsilon^2\right] N_{\varepsilon,i}(t) + m\big(N_{\varepsilon,j}(t) - N_{\varepsilon,i}(t)\big) + \mathcal{O}(\varepsilon^4), \\[2mm]
\varepsilon^2 \frac{d\overline{z}_{\varepsilon,i}}{dt} = 2\varepsilon^2 g((-1)^i - \overline{z}_{\varepsilon,i}(t)) + m\,\frac{N_{\varepsilon,j}(t)}{N_{\varepsilon,i}(t)}(\overline{z}_{\varepsilon,j}(t) - \overline{z}_{\varepsilon,i}(t)) + \mathcal{O}(\varepsilon^4),
\end{cases}
\tag{13}
$$

which is equivalent to (4).

Remark 1.1 (Relationship between the rescaling of time and small variance regime $\varepsilon^2 \ll 1$) The small variance regime $\sigma^2 \ll \theta^2$ (or equivalently $\varepsilon^2 \ll 1$) considers the case where the variance introduced by reproduction is very small compared to the phenotypic gap between the two habitats (recall that $\theta = \frac{|\theta_2 - \theta_2|}{2}$). Therefore, it takes a very long ecological time to bridge the gap. An interpretation of that intuition can be seen in the rescaled system (13). The effects of the ecology (migration, population growth, death by competition and selection) are of order 1. The evolutionary effects (how selection shifts the mean traits of both subpopulations toward the local optima) are represented by the terms $2\varepsilon^2 g((-1)^i - \overline{z}_{\varepsilon,i}(t))$, and are therefore comparatively very small (of order ε^2). This discrepancy is the motivation of the change in time scales $t = \varepsilon^2 T$ to capture the slow dynamics of the local mean traits. It is also behind the motivation for the slow-fast analysis (see Sect. 3).

Remark 1.2 (Relationship between the small variance regime and the weak selection approximation) A widespread regime studied in quantitative genetics models using the Gaussian assumption of trait distributions is the weak selection approximation (Turelli and Barton 1994; Tufto 2000; Turelli 2017). As we showed formally that the local trait distributions are well approximated by Gaussian distributions in the small variance regime (see (10)), it is natural to examine if the regime of small variance $\sigma^2 \ll \theta^2$ and the weak selection approximation are equivalent.

However, the small variance regime $\sigma^2 \ll \theta^2$ presents an alternative that seems to differ from the weak selection approximation:

1. Either the segregational variance σ^2 is of order 1, and therefore θ^2 must be large, ie. the local optimal traits are far apart. However, this has an indirect consequence on the strength of selection g, which must be small, since $g = \frac{g\theta^2}{r}$ must be of order 1 to be relevant in the rescaled system (13). Nevertheless, this framework is distinct from the weak selection approximation, in the sense that the effective selection felt by an individual adapted to one patch and migrating to the other is of order $g\theta^2$, hence of order 1.

2. Either the segregational variance σ^2 is small compared to θ^2, the latter being of order 1, as well as the other parameters of the system. Therefore, in that case, the selection does not need to be weak. A way to get such a small segregational variance can be illustrated by the following with haploid individuals: suppose that we consider L loci that contribute to the focal quantitative trait additively, and that, at each locus, two alleles segregate, having opposite effects of $\pm\frac{a}{2\sqrt{L}}$, where a is a parameter that scales the magnitude of the effect. An estimation of the variance in the offsprings of two mates is $\hat{\sigma}^2 = \frac{a^2 D}{L}$, where $D < L$ is the number of differences between their respective genetic backgrounds. So $\hat{\sigma}^2 = \mathcal{O}(a^2)$ can be uniformly small provided that the allelic effect size parameter a is small. This calculus is similar as in the numerical simulations performed in Tufto (2000) (equation 10), which also considers small segregational variances with the infinitesimal model (see Figure 1 of Tufto 2000).

2 Equivalence with a moment based model

2.1 Presentation of the moment based model

In Ronce and Kirkpatrick (2001), the authors present a quantitative genetic model to tackle the same problem: the evolutionary dynamics of a species under the effects of stabilizing selection and migration between two symmetric patches. Let us first recall the model and indicate the parameters. Stabilizing selection toward a local phenotypic optima $\theta_i \in \mathbb{R}$ is added to competition for resources within each patch to build the fitness of an individual of phenotype z in patch i:

$$r_i(z) = r_0 \left(1 - \frac{N_i}{K}\right) - \frac{\gamma}{2}(z - \theta_i)^2,$$

where $r_0 > 0$ is the maximal fitness at low density, $K > 0$ the carrying capacity of each environment (assumed to be the same in both of them), and $\gamma > 0$ the intensity of the selection. Migration occurs symmetrically between the two patches at a rate $m > 0$. The mode of reproduction is left unspecified (it is however noteworthy that reproduction in most quantitative genetics models is implicitly sexual), but phenotypes and breeding values are assumed to follow a Gaussian distribution within each population, of constant genetic ($\sigma_g^2 > 0$) and phenotypic ($\sigma_p^2 > 0$) variances, independent of the patch with:

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2,$$

where $\sigma_e^2 > 0$ is the environmental variance. The analysis is focused on the ordinary differential equation system of the first two moments of the local trait distributions (assuming persistence of each subpopulation). Namely, the sizes of the subpopulations

(N_1, N_2) and the mean phenotypic traits $(\overline{z_1}, \overline{z_2})$:

$$
\begin{cases}
\frac{dN_i}{dt} = \left[r_0(1 - \frac{N_i}{K}) - \frac{\gamma}{2}\sigma_p{}^2 - \frac{\gamma}{2}(\overline{z_i} - \theta_i)^2 \right] N_i + m(N_j - N_i), \\[2mm]
\frac{d\overline{z_i}}{dt} = \sigma_g{}^2 \gamma (\theta_i - \overline{z_i}) + m\frac{N_j}{N_i}(\overline{z_j} - \overline{z_i}).
\end{cases}
\tag{14}
$$

2.2 Formal comparison

Let us consider (14) in the case where we neglect the additional variance due to the environment, so that all the variation in trait results from the genetic variance. We will denote this variance by ς^2, so that: $\sigma_p{}^2 = \sigma_g{}^2 := \varsigma^2$. Then, let us also consider the equations of the trait distribution moments derived from our model (4), when disregarding the errors of $\mathcal{O}\left(\frac{\sigma^4}{\theta^4}\right)$. Then, the dynamics of the moments and their stationary states are equivalent under the change of parameters:

$$
r = r_0, \quad g = \frac{\gamma}{2}, \quad \kappa = \frac{r_0}{K}, \quad \sigma^2 = \varsigma^2, \quad \sigma_e = 0.
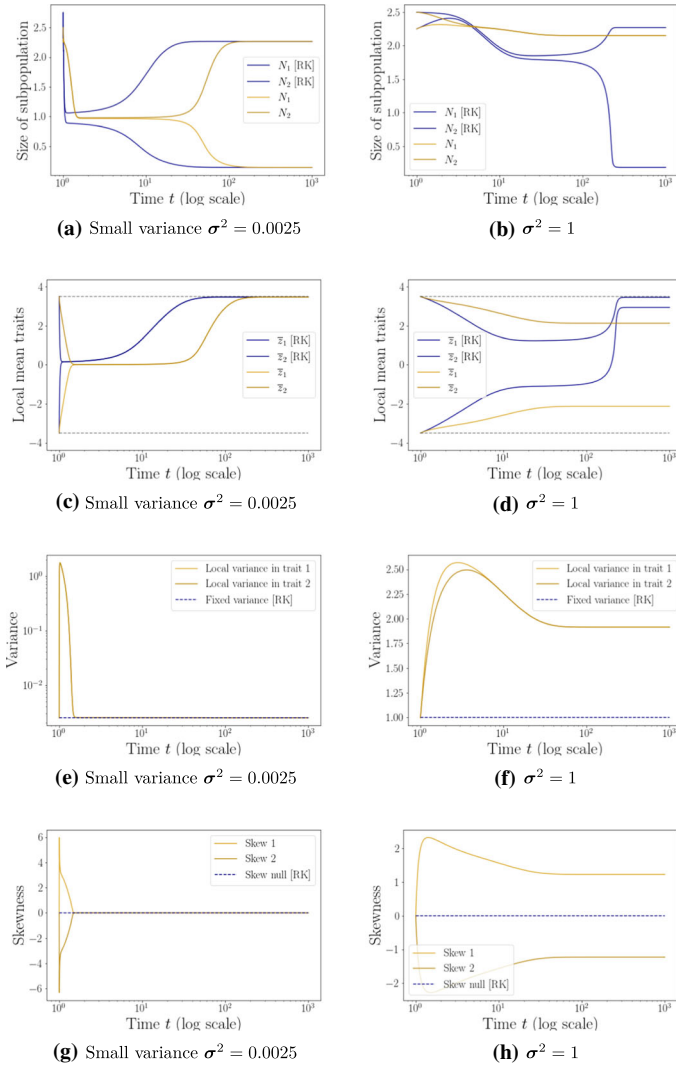\tag{15}
$$

This change of parameters is only possible because, in both models, the variance in trait in the subpopulations is derived from a single parameter encoding the genetic stochasticity ($\sigma_g{}^2$ in Ronce and Kirkpatrick (2001) and σ^2 in our model). Particularly, the variance is independent from the other biological parameters, which is a structural difference with asexual models (see Mirrahimi 2017).

2.3 Numerical comparison

In this subsection, we provide results from numerical simulations performed to confirm this formal equivalence between the stationary states of the two models under the regime of small variance in which we expect this link to hold. In these simulations, we follow two systems:

- The first one is a discretization of (1) under the infinitesimal model assumption for segregational variance, where we follow the evolution of the local trait distributions $n_i(t, \cdot)$. We then compute at each time the sizes, mean traits and variances in trait of the subpopulations $N_i(t)$, $\overline{z}_i(t)$ and σ_i. We emphasize the fact that we do not deduce $N_i(t)$ and $\overline{z}_i(t)$ from the system of moments (4).
- The second one is the system of moments (14) provided in the article Ronce and Kirkpatrick (2001), initialized by integration of $n_i(0, \cdot)$. We denote the respective quantitites $N_{i,RK}(t)$ and $\overline{z}_{i,RK}(t)$.

We then compare the evolution of the sizes and the mean traits of the subpopulations given by both systems. We also provide the evolution of the variance and the skewness in trait in both subpopulations compared to the value of the fixed and constant variance σ_g and the skew null of the Gaussian approximation, for it can shed some lights on the divergence of the two systems. The results are displayed in Fig. 2. Details about numerical domains and schemes can be consulted in "Appendix H".

(a) Small variance $\sigma^2 = 0.0025$

(b) $\sigma^2 = 1$

(c) Small variance $\sigma^2 = 0.0025$

(d) $\sigma^2 = 1$

(e) Small variance $\sigma^2 = 0.0025$

(f) $\sigma^2 = 1$

(g) Small variance $\sigma^2 = 0.0025$

(h) $\sigma^2 = 1$

Fig. 2 Numerial comparison of the model described (1) under the infinitesimal model assumption for segregational variance (yellow lines) with Ronce and Kirkpatrick (2001)'s model (blue line) in small (left panel) and large (right panel) variance regime. All parameters are the same or given by (15) and initial conditions are the same for both models. The left panel shows the results in the small variance regime ($\sigma^2 = 2.5 \times 10^{-3}$). Both models converge quickly to an asymmetrical equilibrium where both subpopulations are adapted to the second habitat ($\overline{z_1} = \overline{z_2} \approx \theta_2$). The right panel shows the results when not in the small variance regime ($\sigma^2 = 1$): the same link of equivalence does not hold. The discrepancy can be explained by looking at the local variances and skews in trait. They are asymptotically supporting a Gaussian assumption with fixed variance on the trait distributions in the small variance regime (note the logarithmic scale for the y-axis of **e**), but not when the segregational variance is larger. Note the logarithmic time scale for the sake of clarity (color figure online)

Parameters of the simulations The value of the parameters were taken from Ronce and Kirkpatrick (2001) (the optimal phenotypes are translated without loss of generality to reduce the numbers of parameters):

$$m = 0.1, \quad \gamma = 0.1, \quad r_0 = 1 + \frac{\gamma}{2}\sigma_p{}^2, \quad K = 2.5\,r_0, \quad \theta = |\frac{\theta_2 - \theta_1}{2}| = 3.5,$$

where the value of $\sigma_g{}^2 = \sigma_p{}^2 = \sigma^2$ determines completely the parameters. Two values are chosen for $\sigma^2 = \sigma_g{}^2 = \sigma_p{}^2$: the first, $\sigma^2 = 0.0025$, is set to assess the regime of small variance ($\sigma^2 \ll \theta^2$) in which our formal link of equivalence should hold. The second, $\sigma^2 = 1$, comes from the value set in Ronce and Kirkpatrick (2001) and illustrates the discrepancy between the two models when not in the small variance regime.

Initial conditions In both simulations, the initial conditions are the same, conditioned to the value of σ, for we want to be close to the equilibrium when under random mating only and selection only, as if the two habitats were disconnected at first. We consider two populations locally adapted to their habitats, but one is a little smaller in size than the other. To do so, we set:

$$\begin{cases} n_1(0, z) = \frac{9}{10\kappa} \dfrac{e^{-\frac{(z+\theta)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \\[3mm] n_2(0, z) = \frac{1}{\kappa} \dfrac{e^{-\frac{(z-\theta)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma} \end{cases}$$

Results of the numerical comparison As Fig. 2a and c display the dynamics of the mean traits and population size in both subpopulations in the regime of small variance ($\sigma^2 = 0.0025$), it confirms numerically that both the model used in Ronce and Kirkpatrick (2001) and the model described by (1) under the infinitesimal model assumption for segregational variance share similar dynamics (except maybe at initial times when the migratory fluxes are transiently high). When not in this regime ($\sigma^2 = 1$), Fig. 2b and d show that it does not need to be the case : the model used in Ronce and Kirkpatrick (2001) converges toward a monomorphic asymmetrical equilibrium whereas the model described by (1) under the infinitesimal model assumption for segregational variance converges toward a dimorphic symmetrical equilibrium. The four bottom plots give an intuition of the source of this discrepancy. In the regime of small variance, we can see with Fig. 2e the variances in trait of the subpopulations in our model match the fixed genetical variance assumed by the gaussian approximation made in Ronce and Kirkpatrick (2001) (note the logarithmic scale for the y-axis on this figure). Moreover, Fig. 2g shows that the skew in both distributions are very small, as expected by our formal expansions, which makes the Gaussian approximation consistent. On the contrary, when not in the regime of small variance, Fig. 2f shows that the stationary variances in trait in both subpopulations derived from the model described by (1) under the infinitesimal model assumption for segregational variance are significantly greater than the prescribed fixed variance $\sigma_g{}^2$ of Ronce and

Kirkpatrick (2001). It is also important to note that with the former, even if the variance of segregation within families is held constant, the local variances in trait (byproducts of our numerical analysis) vary over time. The presence of respectively negative and positive skews (Fig. 2h) for the subpopulations confirms that the gaussian approximation breaks down in this regime in the model described by (1) under the infinitesimal model assumption for segregational variance, hence the discrepancy in the outcomes with Ronce and Kirkpatrick (2001).

The two models have their own limit. Ronce and Kirkpatrick (2001) assumes that the variance in traits is the same in both subpopulations and constant through time and disregards any skewness in the local trait distributions. The assumption on the model described by (1) with the infinitesimal model acts on the segregation : variance in each family is constant and independent of parental traits or habitat. As a result of that discrepancy between the models, their results differ on some ranges of parameters, as the previous figures show (Fig. 2b, d), while they match on others (Fig. 2a, c). To determine the range of parameters on which each model is closer to an explicit genetic model that includes drift, individual-based simulations are to be carried. That is the prospect of future work.

For now, since we have shown that the model described by (1) under the infinitesimal model assumption for segregational variance is equivalent to Ronce and Kirkpatrick (2001)'s one in the regime of small variance, we will next develop a slow-fast analysis that will reduce the complexity of the system (Sect. 3) in the limit of vanishing variance in order to complete the equilbrium analysis done in Ronce and Kirkpatrick (2001) (Sect. 4).

3 Slow–fast system in small variance regime

In this section, we will see that the small variance regime allows for a separation of time scales to arise, as (13) can be seen as a slow-fast system when $\varepsilon^2 \ll 1$. Using a singular perturbation approach similar to the one described in Levin and Levinson (1954), we will show that it converges in the limit of small variance to the following system, constrained in having $N_1^* > 0$, $N_2^* > 0$:

$$
\begin{cases}
\left[1 - N_1^* - g(z^* + 1)^2 - m\right] N_1^* + m N_2^* = 0, \\
\left[1 - N_2^* - g(z^* - 1)^2 - m\right] N_2^* + m N_1^* = 0, \\
\frac{dz^*}{dt} = 2g \left(\frac{\frac{N_2^*}{N_1^*} - \frac{N_1^*}{N_2^*}}{\frac{N_2^*}{N_1^*} + \frac{N_1^*}{N_2^*}} - z^* \right).
\end{cases}
\tag{16}
$$

Until further notice, let us consider ourselves in the regime of small variance: $\varepsilon^2 \ll 1$.

Monomorphism in the regime of small variance The slow-fast system reduces the complexity of the system (13) from four equations to three (see (16)), as the local mean traits $\bar{z}_{\varepsilon,1}$ and $\bar{z}_{\varepsilon,2}$ both relax rapidly toward the same value $z^*(t)$. Since asymptotically, the mean traits in both subpopulations are the same and the local variances in trait are

infinitesimally small, the metapopulation can be considered as *monomorphic* in $z^*(t)$, which we call the dominant trait.

Biological interpretation of the slow-fast analysis in terms of separation between ecological and evolutionary time scales The limit system (16) highlights the separation of ecological and evolutionary time scales in the limit of small variance, seen from the evolutionary perspective. Indeed, the two first equations of (16) are algebraic and therefore describe an instantaneous equilibrium reached by the local population sizes N_1^* and N_2^*. This equilibrium can be seen as an ecological one, as it results from the balanced actions of birth, death and migration. It depends on the value of the trait z^*, which changes according to the last differential equation. As explained in the previous paragraph, this differential equation results from the changes in local mean traits driven by local selection (attested here by the prefactor g), weighted by the discrepancy between local population sizes. Consequently, the dynamics of z^* can be seen as evolutionary dynamics, constrained to occur on the manifold of ecological equilibrium defined by the first two equations (considered as instantaneously reached on the evolutionary time scale considered).

3.1 Slow–fast system formulation

As we expect monomorphism to occur rapidly in the regime of small variance, let us operate the following change in variables:

$$\delta_\varepsilon = \frac{\bar{z}_{\varepsilon,2} - \bar{z}_{\varepsilon,1}}{2\varepsilon^2}, \quad z_\varepsilon^* = \frac{\bar{z}_{\varepsilon,2} + \bar{z}_{\varepsilon,1}}{2}.$$

Then (13) is equivalent to:

$$\begin{cases} \varepsilon^2 \frac{dN_{\varepsilon,1}}{dt} = \left[1 - N_{\varepsilon,1}(t) - g(z_\varepsilon^*(t) + 1 - \varepsilon^2\delta_\varepsilon(t))^2 - g\varepsilon^2\right]N_{\varepsilon,1}(t) + m\big(N_{\varepsilon,2}(t) - N_{\varepsilon,1}(t)\big) + \mathcal{O}(\varepsilon^4), \\[2mm] \varepsilon^2 \frac{dN_{\varepsilon,2}}{dt} = \left[1 - N_{\varepsilon,2}(t) - g(z_\varepsilon^*(t) - 1 + \varepsilon^2\delta_\varepsilon(t))^2 - g\varepsilon^2\right]N_{\varepsilon,2}(t) + m\big(N_{\varepsilon,1}(t) - N_{\varepsilon,2}(t)\big) + \mathcal{O}(\varepsilon^4), \\[2mm] \varepsilon^2 \frac{d\delta_\varepsilon(t)}{dt} = 2g - m\left(\frac{N_{\varepsilon,2}(t)}{N_{\varepsilon,1}(t)} + \frac{N_{\varepsilon,1}(t)}{N_{\varepsilon,2}(t)}\right)\delta_\varepsilon(t) + \mathcal{O}(\varepsilon^2), \\[2mm] \frac{dz_\varepsilon^*}{dt} = -2gz_\varepsilon^*(t) + m\left(\frac{N_{\varepsilon,2}(t)}{N_{\varepsilon,1}(t)} - \frac{N_{\varepsilon,1}(t)}{N_{\varepsilon,2}(t)}\right)\delta_\varepsilon(t) + \mathcal{O}(\varepsilon^2). \end{cases}$$

$$(17)$$

Let us denote $\Omega = (\mathbb{R}_+^*)^2 \times \mathbb{R}$ and $\bar{Y} = (N_1, N_2, \delta)$ the elements of Ω. Let us define $F : \Omega \to \mathbb{R}$ and $G : \mathbb{R} \times \Omega \to \mathbb{R}^3$ by :

$$\forall (z, (N_1, N_2, \delta)) \in \mathbb{R} \times \Omega,$$

$$F(N_1, N_2, \delta) = m\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right)\delta,$$

$$G(z, N_1, N_2, \delta) = \begin{pmatrix} \left[1 - N_1 - g(z+1)^2 - m\right] N_1 + m N_2 \\ \left[1 - N_2 - g(z-1)^2 - m\right] N_2 + m N_1 \\ 2g - m \left(\frac{N_2}{N_1} + \frac{N_1}{N_2}\right) \delta \end{pmatrix}, \qquad (18)$$

where F and G are respectively in $C^\infty(\Omega, \mathbb{R})$ and $C^\infty(\mathbb{R} \times \Omega, \mathbb{R}^3)$.

Let the following be called the perturbed system (P_ε), where $\varepsilon > 0$ is a vanishing parameter and $\nu_{N,\varepsilon}$ and $\nu_{z,\varepsilon}$ are uniformly bounded as $\varepsilon \to 0$:

$$(P_\varepsilon) \quad \begin{cases} \varepsilon^2 \frac{d\bar{Y}_\varepsilon}{dt} = G(z_\varepsilon, \bar{Y}_\varepsilon) + \varepsilon^2 \nu_{N,\varepsilon}(t), \\ \frac{dz_\varepsilon}{dt} = -2g z_\varepsilon + F(\bar{Y}_\varepsilon) + \varepsilon^2 \nu_{z,\varepsilon}(t), \\ (z_\varepsilon(0), \bar{Y}_\varepsilon(0)) = (z_0^\varepsilon, \bar{Y}_0^\varepsilon). \end{cases} \qquad (19)$$

One can verify that any solution of (17) also solves (P_ε). The framework is concordant with fast/slow system studies, like in Levin and Levinson (1954). We seek to establish the convergence over a finite time interval of the solutions of (P_ε) towards the solution of the unperturbed system (P_0), when $(z_0^\varepsilon, \bar{Y}_0^\varepsilon)$ is close enough to $(z_0^*, \bar{Y}_0^*)$ which verifies $G(z_0^*, \bar{Y}_0^*) = 0$:

$$(P_0) \quad \begin{cases} G(z^*(t), \bar{Y}^*(t)) = 0, \\ \frac{dz^*}{dt} = -2g z^* + F(\bar{Y}^*) \\ (z^*(0), \bar{Y}^*(0)) = (z_0^*, \bar{Y}_0^*), \end{cases} \qquad (20)$$

The first line $G(z^*(t), \bar{Y}^*(t)) = 0$ in (20) defines the slow manifold, parametrized by the slow variable $z^*(t)$, whereas the equation $\frac{dz^*}{dt} = -2g z^* + F(\bar{Y}^*)$ (second line) encodes the slow dynamic on that manifold. The slow manifold can be interpreted as the set of fast equilibria $\bar{Y}^*(t)$ corresponding to the levels given by slow variables $z^*(t)$. We will first assess the number of coexisting fast equilibria for any given parameter set $(g, m) \in \mathbb{R}_+^{*2}$ and value of the slow variable z^*. We will show that there exists either one or none of those, which constrains our proof of convergence to apply when $(z_0^\varepsilon, \bar{Y}_0^\varepsilon)$ is close enough to $(z_0^*, \bar{Y}_0^*)$ (the latter being on the slow manifold). Then, we will show that those fast equilibria are locally stable in Lemma 6. This lemma represents the essential condition for the convergence to apply on the finite time interval $[0, t^*]$, where t^* will be subsequently defined (see Levin and Levinson (1954) and "Appendix E" for the detailed proof). We state the following theorem:

Theorem 3.1 *Let $(\bar{Y}^*, z^*)$ be solution of (20) on $[0, t^*]$ with initial conditions $(z_0^*, \bar{Y}_0^*)$, located on the slow manifold (ie. such that $G\left(z^*(t), \bar{Y}^*(t)\right) = 0$ for $t \in [0, t^*]$). For $0 < \varepsilon < 1$, let $(\bar{Y}_\varepsilon, z_\varepsilon)$ be solution of (19) on $[0, t^*]$ with initial conditions $(z_0^\varepsilon, \bar{Y}_0^\varepsilon)$. Then, as $\max(\varepsilon, |z_0^\varepsilon - z_0^*|, |\bar{Y}_0^\varepsilon - \bar{Y}_0^*|) \to 0$, $(\bar{Y}_\varepsilon, z_\varepsilon)$ converges toward $(\bar{Y}^*, z^*)$ uniformly on $[0, t^*]$.*

3.2 Number of coexisting fast equilibria

Let us explicit that fast equilibria corresponding to $z^* \in \mathbb{R}$ are $\bar{Y}^* = (N_1^*, N_2^*, \delta^*) \in \Omega = (\mathbb{R}_+^*)^2 \times \mathbb{R}$ verifying: $G(z^*, \bar{Y}^*) = 0$, ie. the system:

$$
\begin{cases}
\left[1 - N_1^* - g(z^* + 1)^2 - m\right] N_1^* + m N_2^* = 0, \\
\left[1 - N_2^* - g(z^* - 1)^2 - m\right] N_2^* + m N_1^* = 0, \\
2g - m\left(\frac{N_2^*}{N_1^*} + \frac{N_1^*}{N_2^*}\right) \delta^* = 0.
\end{cases}
\tag{21}
$$

We stress that this definition of fast equilibria requires both sizes of the subpopulations to be positive (we can notice that the two first equations of (21) do not allow for one population to go extinct while the other one persists). The objective is to identify how many coexisting fast equilibria there are for each set of parameter $(g, m, z^*) \in (\mathbb{R}_+^*)^2 \times \mathbb{R}$. To that purpose, let us first notice that the fast equilibria can be defined only using their demographic ratio $\frac{N_2^*}{N_1^*}$.

Lemma 1 *For $z^* \in \mathbb{R}$, let us define:*

$$
P_{z^*}(X) = X^3 - f_1(z^*)X^2 + f_2(z^*)X - 1,
$$

where

$$
f_1(z^*) = 1 + \frac{g}{m}(z^* + 1)^2 - \frac{1}{m}, \quad f_2(z^*) = 1 + \frac{g}{m}(z^* - 1)^2 - \frac{1}{m}.
$$

If (N_1^*, N_2^*, δ^*) is a fast equilibrium, then: $\rho^* = \frac{N_2^*}{N_1^*}$ is a positive root of P_{z^*} greater than $f_1(z^*)$. Conversely, if ρ^* is a positive root of P_{z^*} greater than $f_1(z^*)$, then:

$$
(N_1^*, N_2^*, \delta^*) = \left(m[\rho^* - f_1(z^*)], \, m\,\rho^*\,[\rho^* - f_1(z^*)], \, \frac{2g}{m\left(\rho^* + \frac{1}{\rho^*}\right)} \right) \in \Omega,
$$

is a fast equilibrium corresponding to z^* and $\rho^* = \frac{N_2^*}{N_1^*}$.

Consequently, the number of fast equilibria corresponding to z^* is the number of positive roots of $P_{z^*}(X)$ greater than $f_1(z^*)$.

Proof (Proof of Lemma 1) For $z^* \in \mathbb{R}$, since $\bar{Y}^* \in \Omega = \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}$, one can notice that (21) is equivalent to:

$$
\begin{cases}
\frac{N_2^*}{N_1^*} = \frac{g(z^*-1)^2 + m - 1 - m\frac{N_1^*}{N_2^*}}{g(z^*+1)^2 + m - 1 - m\frac{N_2^*}{N_1^*}}, \\
N_1^* = m\frac{N_2^*}{N_1^*} + 1 - g(z^* + 1)^2 - m, \\
\delta^* = \frac{2g}{m\left(\frac{N_2^*}{N_1^*} + \frac{N_1^*}{N_2^*}\right)}.
\end{cases}
$$

$$\iff \begin{cases} \left[\frac{N_2^*}{N_1^*}\right]^3 - \left[\frac{N_2^*}{N_1^*}\right]^2 \left[1 + \frac{g}{m}(z^*+1)^2 - \frac{1}{m}\right] + \left[\frac{N_2^*}{N_1^*}\right] \left[1 + \frac{g}{m}(z^*-1)^2 - \frac{1}{m}\right] - 1 = 0, \\ N_1^* = m \left[\frac{N_2^*}{N_1^*} - \left(1 + \frac{g}{m}(z^*+1)^2 - \frac{1}{m}\right)\right], \\ \delta^* = \frac{2g}{m\left(\frac{N_2^*}{N_1^*} + \frac{N_1^*}{N_2^*}\right)}. \end{cases}$$

Hence the result. □

Remark 3.1 Thanks to the symmetrical setting of the habitats, one can notice that, for all $z^* \in \mathbb{R}$, $P_{-z^*}(X) = X^3 P_{z^*}(1/X)$ and $f_1(-z^*) = f_2(z^*)$. Hence, the number of positive roots of P_{z^*} that are greater than $f_1(z^*)$ is the number of positive roots of P_{-z^*} that are greater than $f_2(z^*)$. Therefore, from now on, we will consider that $z^* \geq 0$ without loss of generality.

The Lemma 2 shows that multiple fast equilibria cannot coexist and fast equilibria do not need to exist for any given set of parameters $(g, m, z^*) \in \mathbb{R}_+^{*\,2} \times \mathbb{R}_+$.

Lemma 2 *Let $z^* \geq 0$. Then:*

(i) *If P_{z^*} has more than a single positive root, then they are all lower than $f_1(z^*)$. Hence, no fast equilibrium can exist in this configuration.*
(ii) *If P_{z^*} has a single positive root ρ^*, then:*

$$\left[\rho^* > f_1(z^*)\right] \iff \left[f_1(z^*) \leq 0\right] \vee \left[P_{z^*}(f_1(z^*)) < 0\right].$$

Proof (Proof of Lemma 2) Let $z^* \geq 0$. As $P_{z^*}(0) = -1$, and the leading coefficient is 1, P_{z^*} has at least one positive root and has either 1 or 3 positive roots.

(i) Let us assume that P_{z^*} has three positive roots x_1, x_2, x_3. Then $f_1(z^*) = x_1 + x_2 + x_3 > \max\{x_1, x_2, x_3\}$, since the three roots are positive.
(ii) Let us assume now that P_{z^*} has a single positive root ρ^*. As $P_{z^*}(0) = -1 < 0$ and the leading coefficient of P_{z^*} is 1, we deduce that, for $y > 0$: $y < \rho^* \iff P_{z^*}(y) < 0$. Hence the result. □
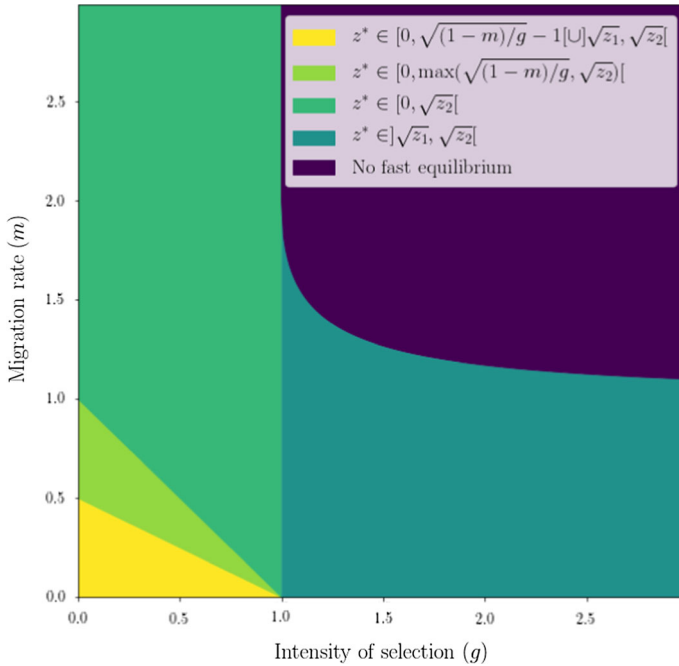
The second point of the Lemma 2 allows us to precise in the next proposition the conditions on z^* such that a fast equilibrium exists, depending on $(g, m) \in \mathbb{R}_+^{*\,2}$ (see also Fig. 3):

Proposition 3.1 *For $(g, m, z^*) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+$ such that P_{z^*} has a single positive root, let us define:*

$$\Delta = \frac{4}{g^2}\left[m^2 - 4g(m-1)\right], \quad z_1 = \frac{1}{2}\left[\frac{2(g+1-m)}{g} - \sqrt{\Delta}\right],$$

$$z_2 = \frac{1}{2}\left[\frac{2(g+1-m)}{g} + \sqrt{\Delta}\right].$$

The following holds:

∗ *If $g \geq 1$ and:*

Fig. 3 Description of the conditions imposed on $z^* \geq 0$ to get a fast equilibrium depending on the pair (g, m) under the preliminary assumption that P_{z^*} has a single positive root, according to the results of Proposition 3.1. When selection is smaller than 1, symmetrical fast equilibria exist ($z^* = 0$), and do not when selection is larger than 1. When both migration and selection are both too strong, no fast equilibrium can exist

⋄ $m < 2g\left(1 - \sqrt{1 - \frac{1}{g}}\right)$, then for all $z^* \in]\sqrt{z_1}, \sqrt{z_2}[$, there exists a single fast equilibrium, and none otherwise.

⋄ $m \geq 2g\left(1 - \sqrt{1 - \frac{1}{g}}\right)$ (ie. $\Delta \leq 0$), then for all $z^* \geq 0$, there exists no fast equilibria.

∗ If $g < 1$, then :

⋄ If $m \leq \frac{1-g}{2}$, then, for $z^* \in [0, \sqrt{\frac{1-m}{g}} - 1[\cup]\sqrt{z_1}, \sqrt{z_2}[$, there exists a single fast equilibrium associated to z^*, and none otherwise.

⋄ If $\frac{1-g}{2} < m < 1 - g$, then, for $z^* \in [0, \max\left(\sqrt{\frac{1-m}{g}} - 1, \sqrt{z_2}\right)[$, there exists a single fast equilibrium associated to z^*, and none otherwise.

⋄ If $1 - g \leq m$, then, for $0 \leq z^* < \sqrt{z_2}$, there exists a single fast equilibrium associated to z^*, and none otherwise.

The proof of Proposition 3.1 is located in "Appendix F".

Finally, we examine the conditions upon which P_{z^*} has three positive roots. Due to the high degrees of the polynomials involved, an analytical condition on $(g, m) \in \mathbb{R}_+^{*2}$ has only been found when $z^* \in [-1, 1]$:

Proposition 3.2 *If $1 + 2m \geq g$, for all $z^* \in [-1, 1]$, P_{z^*} has a single positive root.*
If $1 + 2m < g$, there exists an interval $I \neq \emptyset$ centered in 0 such that for all $z^* \in I$, P_{z^*} has three distinct positive roots.

Proof The proof will require three lemma. The first one states conditions upon which P_{z^*} has three distinct positive roots for $z^* \in \mathbb{R}$. The second one gives an explicit condition determining if $P_0 = P_{z^*=0}$ has one ($1 + 2m \geq g$) or three distinct positive roots ($1 + 2m < g$). The third one shows that if there exists a $z^* \in [-1, 1]\backslash\{0\}$ such that P_{z^*} has three distinct positive roots, then P_0 also has three distinct positive roots. □

Lemma 3 *Let $z^* \in \mathbb{R}$. $P_{z^*}(X) = X^3 - f_1(z^*)X^2 + f_2(z^*)X - 1$ has three distinct positive roots if and only if the three following conditions hold simultaneously:*

  (i) $f_1(z^*) > 0$,
  (ii) $f_2(z^*) > 0$,
  (iii) $\Delta(z^*) := f_1(z^*)^2 f_2(z^*)^2 - 4(f_1(z^*)^3 + f_2(z^*)^3) + 18 f_1(z^*) f_2(z^*) - 27 > 0.$

Proof (Proof of Lemma 3) Let $(x_1, x_2, x_3) \in \mathbb{C}^{*3}$ be the roots of P_{z^*}. Since $x_1 x_2 x_3 = 1$, we have:

$$f_1(z^*) = x_1 + x_2 + x_3, \quad f_2(z^*) = \frac{x_1 x_2 + x_2 x_3 + x_3 x_1}{x_1 x_2 x_3} = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3}.$$

Let us assume first that x_1, x_2, x_3 are positive and distinct. Then they are real and from the latter, $f_1(z^*) > 0$ and $f_2(z^*) > 0$. Moreover, they are real and distinct if and only if the discriminant of P_{z^*} is positive, hence condition (iii).

Conversely, let us assume (i), (ii) and (iii). Then x_1, x_2, x_3 are real and distinct. Since $P_{z^*}(0) < 0$, two of them (for example x_2 and x_3) share the same sign. Suppose that they are negative (they cannot be 0 since $P_{z^*}(0) = -1$). Then (i) yields:

$$x_1 > |x_2| + |x_3|.$$

Hence :

$$f_2(z^*) = \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} = \frac{1}{x_1} - \frac{1}{|x_2|} - \frac{1}{|x_3|} < \frac{1}{|x_2| + |x_3|} - \frac{1}{|x_2|} - \frac{1}{|x_3|} < 0,$$

which contradicts (ii). Hence x_1, x_2, x_3 are positive and distinct. □

Lemma 4 *$P_0 = P_{z^*=0}$ has three distinct positive roots if and only if $g > 1 + 2m$ and one positive root otherwise.*

Proof (Proof of Lemma 4) One can notice that $f_1(0) = f_2(0) = 1 + \frac{g}{m} - \frac{1}{m}$ and:

$$\Delta(0) = f_1(0)^4 - 8 f_1(0)^3 + 18 f_1(0)^2 - 27 = (f_1(0) + 1)(f_1(0) - 3)^3.$$

Hence, the precedent lemma ensures that P_0 has three distinct positives roots only in the region where $f_1(0) = f_2(0) = 1 + \frac{g}{m} - \frac{1}{m} > 0$ and $\Delta(0) > 0$. That occurs if and only if $f_1(0) > 3$, which yields $g > 1 + 2m$. □

Lemma 5 *If there exists $z^* \in [-1, 1]$ such that P_{z^*} has three distinct positive roots, then P_0 has three distinct positive roots.*

Proof (Proof of Lemma 5) We recall that we study the case $z^* > 0$ without loss of generality. Let us consider $z^* \in]0, 1]$ such that P_{z^*} has three distinct positive roots. From Remark 3.1, P_{-z^*} has also three distinct positive roots. Thereby, Lemma 3 implies that $f_i(\pm z^*) > 0$, $i = 1, 2$ and $\delta(\pm z^*) > 0$.

It is clear that f_1 is strictly increasing on $] - 1, 1[$ and f_2 is strictly decreasing on $] - 1, 1[$. As $f_i(\pm z^*) > 0$, $i = 1, 2$, we get that $f_1 > 0$ and $f_2 > 0$ on $[-z^*, z^*]$, in particular $f_1(0) > 0$ and $f_2(0) > 0$.

Moreover, let us introduce the function $g : z \mapsto f_1(z)^2 - 3 f_2(z)$. For $z \in] - 1$, $1[$, $g'(z) = 2 f_1'(z) f_1(z) - 3 f_2'(z) > 0$, because $f_1(z) > 0$, $f_1'(z) > 0$ and $f_2'(z) < 0$. Therefore, g is increasing on $] - 1, 1[$. One can also notice that $g(z)$ is the quarter of the discriminant of $P_z'(X)$. As P_{z^*} and P_{-z^*} have three distinct positive roots, by Rolle's theorem, P_{z^*}' and P_{-z^*}' have two distinct positive roots. Therefore, $g(-z^*)$ and $g(z^*)$ are positive. As g is increasing on $[-z^*, z^*]$, we get: $0 < g(0) = f_1(0)(f_1(0) - 3)$. Since $f_1(0) > 0$ and $g(0) > 0$, we have $3 < f_1(0) = 1 + \frac{g}{m} - \frac{1}{m}$. By the Lemma 4, P_0 has then three distinct positive roots. $\square$

The successive applications of Lemmas 4 and 5 are sufficient to conclude.

3.3 Fast relaxation towards the slow manifold

We hereby prove the following lemma on the stability of the slow manifold:

Lemma 6 *For $(z, \bar{Y}) \in \mathbb{R} \times \Omega$ such that $G(z, \bar{Y}, 0) = 0$, $J_G(z, \bar{Y}) := \partial_{\bar{Y}} G(z, \bar{Y}, 0)$ is invertible. Furthermore, its eigenvalues are real and negative.*

Proof For $(z, \bar{Y}) \in \mathbb{R} \times \Omega$ such that $G(z, \bar{Y}, 0) = 0$, we have:

$$J_G(z, \bar{Y}) = \begin{pmatrix} -2N_1 + [1 - g(z+1)^2 - m] & m & 0 \\ m & -2N_2 + [1 - g(z-1)^2 - m] & 0 \\ \frac{m\,\delta}{N_1}\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right) & -\frac{m\,\delta}{N_2}\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right) & -m\left(\frac{N_2}{N_1} + \frac{N_1}{N_2}\right) \end{pmatrix}.$$

Since $G(z, \bar{Y}, 0) = 0$, (18) leads to:

$$J_G(z, \bar{Y}) = \begin{pmatrix} -m\frac{N_2}{N_1} - N_1 & m & 0 \\ m & -m\frac{N_1}{N_2} - N_2 & 0 \\ \frac{m\,\delta}{N_1}\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right) & -\frac{m\,\delta}{N_2}\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right) & -m\left(\frac{N_2}{N_1} + \frac{N_1}{N_2}\right) \end{pmatrix}$$

$$= \begin{pmatrix} & & 0 \\ & J & 0 \\ \frac{m\,\delta}{N_1}\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right) & -\frac{m\,\delta}{N_2}\left(\frac{N_2}{N_1} - \frac{N_1}{N_2}\right) & -m\left(\frac{N_2}{N_1} + \frac{N_1}{N_2}\right) \end{pmatrix}$$

so that we can compute :

$$\det J_G(z, \bar{Y}) = -m \left(\frac{N_2}{N_1} + \frac{N_1}{N_2} \right) \left[m \frac{N_2^2}{N_1} + m \frac{N_1^2}{N_2} + N_1 N_2 \right] < 0.$$

Hence $J_G(z, \bar{Y})$ is invertible. A first eigenvalue is $-m \left(\frac{N_2}{N_1} + \frac{N_2}{N_1} \right) < -2m$. The last two eigenvalues are those of the upper left block J. We have:

$$\text{tr}(J) < -2m < 0, \quad \det(J) = m \frac{N_1^2}{N_2} + m \frac{N_1^2}{N_2} + N_1 N_2 > 0,$$

and:

$$\text{tr}(J)^2 - 4 \det(J) = 4m^2 + \left(m \frac{N_2}{N_1} - m \frac{N_1}{N_2} + N_1 - N_2 \right)^2 > 4m^2 > 0.$$

Hence J has two real negative eigenvalues and consequently, $J_G(z, \bar{Y})$ has three real negative eigenvalues. □

4 Analytical description of the equilibria in the limit of vanishing variance

In this section, we will perform an equilibrium analysis for the stationary problem in the limit of vanishing variance. As numerically illustrated in Sect. 2, under this regime, our model (1) leads to the same dynamics of the moments as in Ronce and Kirkpatrick (2001). Consequently, this equilibrium analysis corresponds to the one made in Ronce and Kirkpatrick (2001) (in the limit of vanishing variance where their system of four equations converges to the system (16)). Recall from the introduction that the study done in Ronce and Kirkpatrick (2001) reveals two types of equilibrium:

- Symmetrical equilibrium, where both populations are of the same size and equally maladapted to their local habitat (corresponding to a generalist species). Such an equilibrium is derived analytically by the authors. It is worthy to note that in the small variance regime, this equilibrium becomes monomorphic.
- Asymmetrical equilibria, where one larger population of locally adapted individuals acts as a source for the other more poorly adapted smaller population (corresponding to a specialist species). The authors numerically explored this type of equilibrium and derived approximations for low migration rates. One aim of this section is to characterize such equilibria analytically.

The fast/slow analysis done in Sect. 3 gives us the opportunity to go further in the equilibrium analysis in the small variance regime, as the asymptotic system (16) presents a reduced complexity (three equations instead of four). Moreover, adopting

the notation $\rho^* = \frac{N_2^*}{N_1^*} > 0$ and using the polynomial previously defined:

$$P_{z^*}(X) = X^3 - X^2\left[1 + \frac{g}{m}(z^* + 1)^2 - \frac{1}{m}\right] + X\left[1 + \frac{g}{m}(z^* - 1)^2 - \frac{1}{m}\right] - 1,$$

the Lemma 1 implies that (16) is equivalent to:

$$\begin{cases} P_{z^*}(\rho^*) = 0, \\ \frac{dz^*}{dt} = 2g\left(\frac{\rho^{*2}-1}{\rho^{*2}+1} - z^*\right), \end{cases} \tag{22}$$

with the constraint $\rho^* > \max\left(1 + \frac{g}{m}(z^* + 1)^2 - \frac{1}{m}, 0\right)$ (ie. $N_1^* > 0$). This reduction in the regime of small variance allows us in a second time to derive analytical expressions of every possible equilibrium $(z^*, N_1^*, N_2^*) \in \mathbb{R} \times \mathbb{R}_+^{*\,2}$ from solving:

$$\left[P_{\frac{\rho^{*2}-1}{\rho^{*2}+1}}(\rho^*) = 0\right] \wedge \left[\rho^* > \max\left(1 + \frac{g}{m}\frac{4\rho^{*4}}{(\rho^* + 1)^2} - \frac{1}{m}, 0\right)\right], \tag{23}$$

and next setting:

$$\begin{cases} z^* = \frac{\rho^{*2}-1}{\rho^{*2}+1}, \\ N_1^* = m\left[\rho^* - \left[1 + \frac{g}{m}(z^* + 1)^2 - \frac{1}{m}\right]\right], \\ N_2^* = m\left[\frac{1}{\rho^*} - \left[1 + \frac{g}{m}(z^* - 1)^2 - \frac{1}{m}\right]\right]. \end{cases} \tag{24}$$

We will show that there exists a unique symmetrical equilibrium, which correspond to the monomorphic one analytically found by Ronce and Kirkpatrick (2001) (in the regime of small variance). We will then show that there can additionally exist a mirrored pair of asymmetrical equilibria uniquely defined, corresponding to the ones found numerically by Ronce and Kirkpatrick (2001).

4.1 Equilibrium analysis

The objective of this section is to find the steady states (z^*, N_1^*, N_2^*) of the system (16) that lie in $\mathbb{R} \times \mathbb{R}_+^{*\,2}$ (or equivalently, solve (23) and set (24)). Henceforth, we will call these (z^*, N_1^*, N_2^*) *equilibria*. The systems (23) and (24) imply that

$(z^*, N_1^*, N_2^*) \in \mathbb{R} \times \mathbb{R}_+^{*\,2}$ is an equilibrium if and only if $\bar{Y}^* = \left(N_1^*, N_2^*, \frac{2g}{m\left[\rho^* + \frac{1}{\rho^*}\right]}\right)$

is a fast equilibrium corresponding to $z^* = \frac{\rho^{*2}-1}{\rho^{*2}+1}$. As a corollary of the Proposition 3.1, we get that the following region of parameters does not allow for any equilibria to exist:

Corollary 1 *If* $\left[[g \geq 1] \wedge \left[m \geq 2g \left(1 - \sqrt{1 - \frac{1}{g}} \right) \right] \right]$, *then there can exist no equilibria as defined by* (23) *and* (24), *i.e. that leads to* $N_1^* > 0$ *and* $N_2^* > 0$.

Remark 4.1 Although our analysis is not meant to describe extinction, we observe numerically that the system goes to extinction in the region defined in the previous corollary (see Fig. 6).

From now on and until further notice, we will thus consider $(m, g) \in \mathbb{R}_+^{*2}$ such that:

$$[g < 1] \vee \left[m < 2g \left(1 - \sqrt{1 - \frac{1}{g}} \right) \right].$$

4.1.1 Symmetric equilibrium: fixation of a generalist species

Definition 1 We call symmetric equilibrium the $(z^*, N_1^*, N_2^*) \in \mathbb{R} \times \mathbb{R}_+^{*2}$ solutions of (23) and (24) where both subpopulations have the same size: $N_1^* = N_2^* = N^* > 0$.

We first state that there can only exist one viable symmetrical equilibrium:

Proposition 4.1 *There exists a single symmetric equilibrium when* $g < 1$, *given by* $(0, 1 - g, 1 - g)$ *and none when* $g \geq 1$.

Proof Regarding (23): we have $\rho^* = 1$ is a positive root of:

$$P_{z^*=0}(X) = X^3 - \left(1 + \frac{g-1}{m} \right) X^2 + \left(1 + \frac{g-1}{m} \right) X - 1,$$

that additionally satisfies:

$$\rho^* > 1 + \frac{g-1}{m} \iff 1 > g.$$

Hence the symmetrical equilibrium is uniquely defined by $(0, 1 - g, 1 - g)$ (from considering (24)). $\qquad \square$

In this case, as 0 is the middle point between the local optimal phenotypes -1 in habitat 1 and 1 in habitat 2, each subpopulation is equally maladapted.

Remark 4.2 The existence of this equilibrium (or the associated extinction when it is not viable) was expected, for we consider symmetrical habitats and thus symmetrical dynamics. Therefore, under symmetrical initial conditions, the outcome is necessarily symmetrical.

4.1.2 Asymmetric equilibrium: specialist species

We define as asymmetric equilibrium any solution of (24) in $\mathbb{R} \times \mathbb{R}_+^{*2}$ that is not a symmetric equilibrium.

Remark 4.3 One can notice that the system (23) is invariant under the transformation $\rho^* \mapsto \frac{1}{\rho^*}$ or equivalently (24) under $(z^*, N_1^*, N_2^*) \mapsto (-z^*, N_2^*, N_1^*)$. Thus, we do not lose in generality if we look for equilibria with $N_1^* < N_2^*$ instead of $N_1^* \neq N_2^*$: to each asymmetrical equilibrium with $N_1^* < N_2^*$, we can associate its mirrored version.

This section is dedicated to confirm the numerical intuition of Ronce and Kirkpatrick (2001) and show that there exists a range of parameters such that a unique mirrored couple of asymmetrical equilibria exists.

Proposition 4.2 *Let* $(m, g) \in \mathbb{R}_+^{*\,2}$ *be such that:*

$$[1 + 2m < 5g] \wedge \left[m^2 > 4g\,(m-1)\right]. \tag{25}$$

Then there exists a single asymmetrical equilibrium (z^*, N_1^*, N_2^*) *with* $N_1^* < N_2^*$, *given by:*

$$\begin{cases} N_1^* = (1-m) + m\rho - 4g\frac{\rho^{*4}}{(\rho^{*2}+1)^2}, \\ N_2^* = (1-m) + \frac{m}{\rho^*} - 4g\frac{1}{(\rho^{*2}+1)^2}, \\ z^* = \frac{\rho^{*2}-1}{\rho^{*2}+1} \neq 0, \end{cases} \tag{26}$$

where $\rho^* = \frac{y^* + \sqrt{y^{*2}-4}}{2}$ *and* $y^*\left(= \rho^* + \frac{1}{\rho^*}\right)$ *is the only root greater than 2 of the polynomial:*

$$S(Y) = Y^3 + \frac{(1-4g)}{m}Y^2 - \frac{4g}{m}Y + \frac{4g}{m}.$$

*Conversely, if the condition (25) is not verified, there can be no asymmetrical equilibria.*

Remark 4.4 For $g > 1$, $m > 0$, we have the equivalence:

$$[1 + 2m < 5g] \wedge \left[m^2 > 4g\,(m-1)\right] \iff \left[m < 2g\left(1 - \sqrt{1 - \frac{1}{g}}\right)\right].$$

Figure 4 summarizes the conditions obtained with Propositions 4.1 and 4.2. It illustrates the analytical range of parameters where the different types of equilibrium exist when the strength of selection g and the migration rate m vary. In the region where none of the conditions are met, we observe numerically that the system leads to extinction (upper right region). In the intermediate green triangle, the two asymmetrical equilibria coexist with the symmetrical equilibrium.

Proof (Proof of Proposition 4.2)
The first part of the proof is directed to solve the equation given in (23) and consists in two lemmas. The second part of the proof examines the conditions under which
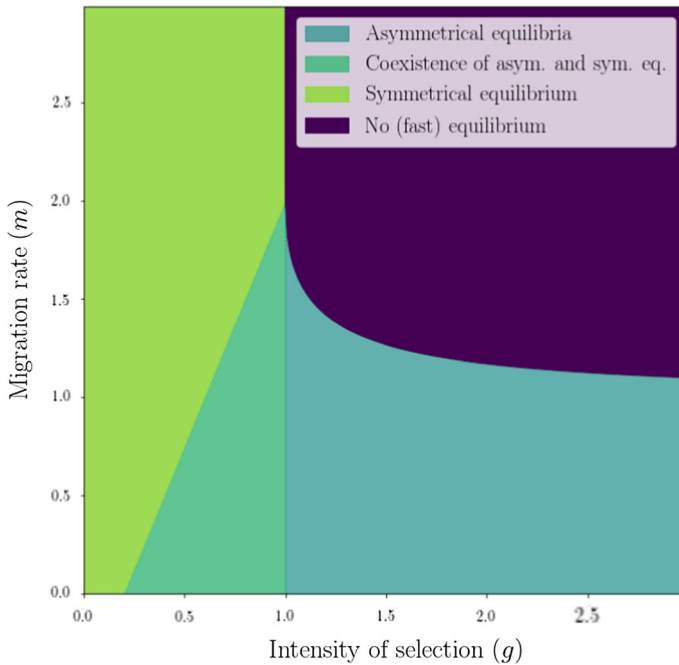
Fig. 4 Regions of existence of the equilibria, according to Propositions 4.1 and 4.2. The symmetrical equilibrium is only determined by the intensity of selection, regardless of the migration rate. The asymmetrical equilibria cannot exist for large migration rate ($m > 2$) or small intensity of selection. The limit of the blue region is given by $m = 1$ when g goes to ∞. Interestingly enough, at intermediate migration: $m \in [1, 2]$, asymmetrical equilibria only exist for a bounded range of positive g: selection cannot be too strong nor too weak. Moreover, see Sect. 4.2 for stability results about these equilibria to determine which equilibria prevail when both symmetrical and asymmetrical coexist (turquoise triangular region) (color figure online)

such solutions verify the inequality constraint given by (23). It consists in a lemma that involves tedious computations. Consequently, the second part of the proof is left to be consulted in "Appendix G".   □

First part of the proof (23) provides us with a close equation: $P_{\frac{\rho^{*2}-1}{\rho^{*2}+1}}(\rho^*) = 0$. Solving it seems necessary, however, the direct search for solutions of this equation leads to consider a seventh degree polynomial. The first part of the proof consists in two lemmas. We first rely on the symmetry of the system noticed by Remark 4.3 ((z^*, ρ^*) is solution if and only if $\left(-z^*, \frac{1}{\rho^*}\right)$ is too) to reduce the complexity from a seventh degree polynomial to a third degree polynomial S:

Lemma 7 *Let us define:*

$$S(Y) = Y^3 + \frac{(1 - 4g)}{m}Y^2 - \frac{4g}{m}Y + \frac{4g}{m}.$$

Then, we have the following relation for $\rho^* \in \mathbb{R}_+^* \setminus \{1\}$:

$$S\left(\rho^* + \frac{1}{\rho^*}\right) = \frac{(1+\rho^{*2})^2}{(\rho^* - 1)\rho^{*3}} P_{\frac{\rho^{*2}-1}{\rho^{*2}+1}}(\rho^*).$$

As for $\rho^* \in \mathbb{R}_+^* \setminus \{1\}$, $\rho^* + \frac{1}{\rho^*} > 2$, we next look for the number of roots of S greater than 2:

Lemma 8 *Let $a > 0, b \in \mathbb{R}$. Let us define $b(a) := \frac{5a}{4} - 2$. Then: if $b \geq b(a)$, $S(Y) = Y^3 + (b-a)Y^2 - aY + a$, has no root greater than 2. If $b < b(a)$, S has a single root greater than 2.*

The successive application of the Lemma 7 and Lemma 8 with:

$$\begin{cases} b = \frac{1}{m}, \\ a = \frac{4g}{m} > 0, \end{cases}$$

yields that there exists a unique solution to (23) if and only if $1 + 2m < 5g$, and therefore to (24) in $\mathbb{R} \times \mathbb{R}_+^{*2}$ which is exactly (26). Proving the two lemmas concludes the first part of the proof.

Proof (Proof of Lemma 7) Let us consider $\rho^* \in \mathbb{R}^* \setminus \{1\}$. Then we have:

$$\frac{(1+\rho^{*2})^2}{(\rho^* - 1)\rho^{*3}} P_{\frac{\rho^{*2}-1}{\rho^{*2}+1}}(\rho^*)$$

$$= \frac{2 - 4g}{m} + \frac{(3m - 4g)}{m}\left(\rho^* + \frac{1}{\rho^*}\right) + \frac{(1 - 4g)}{m}\left(\rho^{*2} + \frac{1}{\rho^{*2}}\right) + \rho^{*3} + \frac{1}{\rho^{*3}}$$

$$= \frac{2 - 4g}{m} + \frac{(3m - 4g)}{m}\left(\rho^* + \frac{1}{\rho^*}\right) + \frac{(1 - 4g)}{m}\left(\rho^{*2} + \frac{1}{\rho^{*2}}\right) + \rho^{*3} + \frac{1}{\rho^{*3}}.$$

Since:

$$\rho^{*2} + \frac{1}{\rho^{*2}} = \left(\rho^* + \frac{1}{\rho^*}\right)^2 - 2,$$

$$\rho^{*3} + \frac{1}{\rho^{*3}} = \left(\rho^* + \frac{1}{\rho^*}\right)^3 - 3\left(\rho^* + \frac{1}{\rho^*}\right),$$

we have:

$$\frac{(1+\rho^{*2})^2}{(\rho^* - 1)\rho^{*3}} P_{\frac{\rho^{*2}-1}{\rho^{*2}+1}}(\rho^*) = \left(\rho^* + \frac{1}{\rho^*}\right)^3 + \frac{1 - 4g}{m}\left(\rho^* + \frac{1}{\rho^*}\right)^2$$

$$- \frac{4g}{m}\left(\rho^* + \frac{1}{\rho^*}\right) + \frac{4g}{m} = S\left(\rho^* + \frac{1}{\rho^*}\right).$$

$\square$

Proof (Proof of Lemma 8)

As $S(0) = a > 0$ and since S goes to $-\infty$ in $-\infty$, S has always a negative root.

Thereby, the case that we take interest in is included within the case where all three roots Z_1, Z_2, Z_3 of S are real. Furthermore, we have the following relations:

$$\begin{cases} Z_1 Z_2 Z_3 = -a < 0, \\ Z_1 Z_2 + Z_2 Z_3 + Z_3 Z_1 = -a < 0. \end{cases}$$

From the first relation, we deduce that S has an even number of positive roots, so either 0 or 2. The second relation leads to a contradiction if all roots are negative. Thus S has necessarily two positive roots and one negative.

Moreover, we have:

$$\frac{1}{Z_1} + \frac{1}{Z_2} + \frac{1}{Z_3} = \frac{Z_1 Z_2 + Z_2 Z_3 + Z_3 Z_1}{Z_1 Z_2 Z_3} = 1.$$

Without loss of generality, let us assume that $Z_3 < 0$. If the remaining two positive roots were greater than 2, then we would get:

$$1 < \frac{1}{Z_1} + \frac{1}{Z_2} \leq \frac{1}{2} + \frac{1}{2} = 1$$

which is a contradiction. Hence at most one is greater than or equal to 2.

The only fact that is left to prove is that such a root exists. Let $S_a(X) = X^3 + (b(a) - a)X^2 - aX + a$. Under the choice of $b(a)$, we can verify that $S_a(2) = 0$. Consequently, the following holds:

$$b < b(a) \iff S(2) < S_a(2) = 0.$$

Therefore, because S goes to $+\infty$ in $+\infty$, if $b > b(a)$, S has an even number of roots greater than 2. Thereby, from the previous part of the proof, in that case, S do not have any roots greater than 2. If $b = b(a)$, 2 is the only root of S greater than or equal to 2. If $b < b(a)$, S has at least one root strictly greater than 2. This root is unique by the argument above (which was independent of b). □

Second part of the proof The second part of the proof is dedicated to show that for all $(m, g) \in \mathbb{R}_+^{*\,2}$ verifying (25), the solution $\rho^* > 0$ that we found in the first part of the proof verifies the constraint given in (23). It consists in the following lemma, that is obtained after tedious calculations done in part with the help of the software Mathematica, so the proof is left to be consulted in "Appendix G".

Lemma 9 *Let $(m, g) \in \mathbb{R}_+^{*\,2}$ verifying (25), and $\rho^* > 0$ be the unique solution of the equation $\left[P_{\frac{\rho^{*2}-1}{\rho^{*2}+1}}(\rho^*) = 0 \right]$. Then:*

$$\rho^* > 1 + \frac{g}{m} \frac{4\rho^{*4}}{(\rho^* + 1)^2} - \frac{1}{m}.$$

Consequently: for $(g, m) \in \mathbb{R}_+^{*\,2}$ such that $1 + 2m < 5g$ and $m^2 > 4g\,(m - 1)$, ρ^* defined in Proposition 4.2 defines an equilibrium with positive subpopulation sizes.

Conversely: if (25) is not met, either $1 + 2m > 5g$, in which case no asymmetrical equilibrium can exist from Lemmas 7 and 8, or $m^2 < 4g\,(m - 1)$ (which implies that $g > 1$), in which case Remark 4.4 and Corollary 1 implies that no equilibrium can exist. $\qquad \square$

4.2 Stability analysis

In this subsection, we examine the stability of the equilibria of the system (22) that we described previously.

Proposition 4.3 *Let $(z^*, N_1^*, N_2^*) \in \mathbb{R} \times \mathbb{R}_+^{*\,2}$ be an equilibrium and $\rho^* = \frac{N_2^*}{N_1^*}$. Then the equilibrium is locally stable (respectively unstable) if:*

$$\frac{4\rho^*}{(\rho^{*2} + 1)^2} \times \frac{1}{P'_{z^*}(\rho^*)} \times \frac{2g}{m} \left[z^* \left(\rho^* - \rho^{*2} \right) - \left(\rho^* + \rho^{*2} \right) \right] + 1 > 0 \quad (resp. < 0).$$

Proof If $(z^*, N_1^*, N_2^*) \in \mathbb{R} \times \mathbb{R}_+^{*\,2}$ is an equilibrium and $\rho^* = \frac{N_2^*}{N_1^*}$, then (N_1^*, N_2^*) is a fast equilibrium associated to z^* (Lemma 1), which implies that P_{z^*} has a single positive root (without multiplicity) that is ρ^* (Lemma 2). Hence ρ^* cannot be a double root of P_{z^*}, which yields: $P'_{z^*}(\rho^*) \neq 0$.

(22) implies that the local stability of the equilibria can be examined by the following system:

$$\begin{cases} \mathcal{G}(z^*, \rho^*) := P_{z^*}(\rho^*) = 0, \\ \rho^* > \left[1 + \frac{g}{m}(z^* + 1)^2 - \frac{1}{m} \right], \\ \frac{dz^*}{dt} = \mathcal{F}(z^*, \rho^*) := 2g \left(\frac{\rho^{*2}-1}{\rho^{*2}+1} - z^* \right). \end{cases}$$

As $\partial_\rho \mathcal{G}(z^*, \rho^*) = P'_{z^*}(\rho^*) \neq 0$, we apply the implicit function theorem to get U a open neighbourhood of z^* and a smooth function $\rho : U \to \mathbb{R}_+^*$ such that:

$$\forall z \in U, \mathcal{G}(z, \rho(z)) = 0.$$

For $z \in U$, we define $f : U \to \mathbb{R}, z \mapsto \mathcal{F}(z, \rho(z))$. Hence, (z^*, N_1^*, N_2^*) is locally stable (resp. unstable) if :

$$f'(z^*) = \nabla \mathcal{F}(z^*, \rho^*) \cdot \left(\frac{1}{\frac{d\rho}{dz}(z^*)} \right)$$

$$= \partial_\rho \mathcal{F}(z^*, \rho^*) \left[-\left(\partial_\rho \mathcal{G}(z^*, \rho^*) \right)^{-1} \partial_z \mathcal{G}(z^*, \rho^*) \right] - 2g < 0 \quad (\text{resp.} > 0).$$

Since we have:

$$\partial_\rho \mathcal{F}(z^*, \rho^*) = 2g \frac{4\rho^{*2}}{(\rho^* + 1)^2}, \quad \left(\partial_\rho \mathcal{G}(z^*, \rho^*) \right)^{-1} = \frac{1}{P'_{z^*}(\rho^*)},$$

and

$$\partial_z \mathcal{G}(z^*, \rho^*) = -2 \frac{g}{m}(z^* + 1)\rho^{*2} + 2 \frac{g}{m}(z^* - 1)\rho^*,$$

the considered equilibrium is locally stable (reps. unstable) if:

$$\frac{4\rho^*}{(\rho^{*2} + 1)^2} \times \frac{1}{P'_{z^*}(\rho^*)} \times \frac{2g}{m} \left[z^* \left(\rho^* - \rho^{*2} \right) - \left(\rho^* + \rho^{*2} \right) \right] + 1 > 0 \quad (\text{resp.} < 0).$$

$\square$

Corollary 2 *The symmetrical equilibrium* $z^* = 0, \rho^* = 1$ *is locally stable (resp. unstable) if* $5g < 1 + 2m$ *(resp.* $5g > 1 + 2m$*) (ie, when it is alone).*

Proof If $z^* = 0$ and $\rho^* = 1$, we have:

$$\frac{4\rho^*}{(\rho^{*2} + 1)^2} \times \frac{1}{P'_{z^*}(\rho^*)} \times \frac{2g}{m} \left[z^* \left(\rho^* - \rho^{*2} \right) - \left(\rho^* + \rho^{*2} \right) \right] + 1$$

$$= \frac{1}{3 - 2 \left(1 + \frac{g}{m} - \frac{1}{m} \right) + \left(1 + \frac{g}{m} - \frac{1}{m} \right)} \times \frac{-4g}{m} + 1 = \frac{1 + 2m - 5g}{1 + 2m - g}.$$

We recall that for the symmetrical equilibrium to exist, we need: $g < 1$, which imply: $g < 1 + 2m$. Hence the result. $\square$

Analytical derivations are more tedious for asymmetrical equilibria. However, when $1 + 2m > g$, we showed that P_{z^*} has a single (without multiplicity) positive root $\rho(z^*)$ for all $z^* \in [-1, 1]$ (Proposition 3.2). The function $\rho : [-1, 1] \to \mathbb{R}_+^*, z \mapsto \rho(z)$ is therefore smooth (where $\rho(z)$ designates the single positive root of P_z). Thus, we can globally define the smooth function f similarly as in Proposition 4.3 on $]-1, 1[$:

$$f : \begin{cases}]-1, 1[\to \mathbb{R} \\ z \mapsto 2g \left(\frac{\rho(z)^2 - 1}{\rho(z)^2 + 1} - z \right), \end{cases} \quad .$$

That leads to the following result:

Corollary 3 *Let $5g > 1+2m > g$. Then the asymmetrical equilibria are locally stable.*

Proof Let (z^*, N_1^*, N_2^*) be an asymmetrical equilibrium. We recall that $z^* = \frac{\rho^{*2}-1}{\rho^{*2}+1} \in$ $]-1, 1[$. From the previous corollary, the symmetric equilibrium is locally unstable, i.e.:

$$f'(0) > 0.$$

Moreover, from Proposition 3.2, $P_{z^*=1}$ has a single positive root, and we can extend f in 1 by continuity and calculate :

$$f(1) = 2g \left(\frac{\rho^2(1) - 1}{\rho^2(1) + 1} - 1 \right) = -\frac{4g}{\rho^2(1) + 1} < 0.$$

Since 0 and z^* are the only zeros of f on $[0, 1]$ (from the uniqueness of the mirrored couple of asymmetric equilibria) and $f'(0) > 0$, f is positive on $]0, z^*[$ ane negative on $]z^*, 1]$. Hence, the asymmetrical equilibria are locally stable. □

To illustrate the diversity of cases in both the number of equilibria and their stability, we display in Fig. 5 the graph of the function f defined above as a function of the dominant trait z when $g = 1.5$ and m takes the following values :

1. $m = 0.02$. There are multiple branches near the origin (yellow curve), as the function f is multi-valued. Indeed, we are in the case where: $1 + 2m < g$. Therefore, for z^* near 0, there is three distinct positive roots for P_{z^*} (from Proposition 3.2), which leads to non-viable fast equilibria (from Lemma 2). Therefore, if the initial dominant trait is near 0, the system will go to extinction.
2. $m = 0.25$, so that the equality $1 + 2m = g$ holds, which is the limit case of the folding near the origin.
3. $m = 1$. For each value of the dominant trait z^*, there is only one root to P_{z^*}. There are three equilibria, an unstable symmetric and two stable asymmetric equilibria $(1 + 2m < 5g)$.
4. $m = 3.25$, so that the equality $1 + 2m = 5g$ holds. This displays the limit of existence of the asymmetrical equilibria (see Proposition 4.2). The three equilibria are merging and exchanging stability.
5. $m = 5$. As m grows further, the asymmetric equilibria do not exist anymore. Therefore, only the symmetric one is left and is stable.

5 Discussion

Contributions In this paper, we have studied the evolutionary dynamics of a complex trait under stabilizing selection in a heterogeneous environment in a sexually reproducing population. To model the process of inheritance of this trait, we have used a mixing
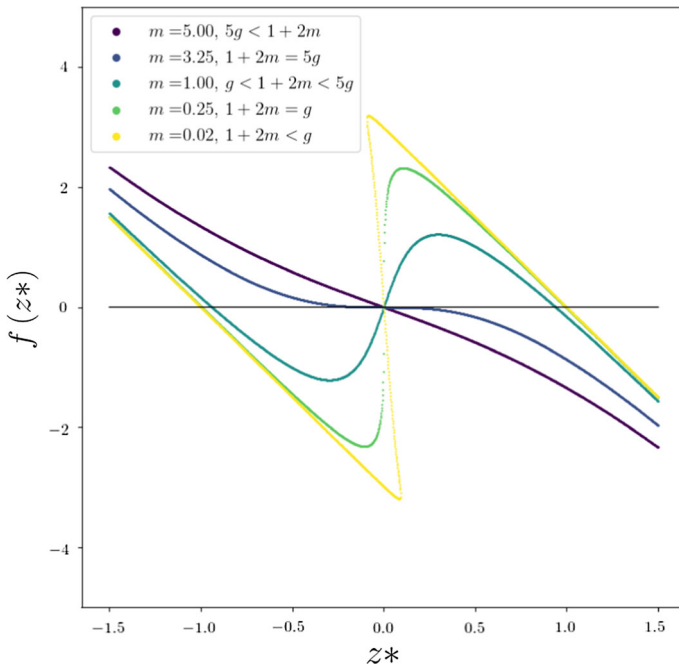
Fig. 5 Graph of the function f for $g = 1.5$ and $m \in \{0.02, 0.25, 1, 3.25, 5\}$. The dominant traits z^* of the different equilibria are located where the curve crosses the horizontal line ($f(z^*) = 0$). An equilibrium is locally stable if the slope of f at the point of equilibrium is negative. For decreasing values of m (dark to light colors), at first, only the symmetric equilibrium exists and is stable (see also Corollary 2). Then, the asymmetrical equilibria emerge (in the parameter region indicated in Proposition 4.2) and are bistable (see also Corollary 3), while the symmetric equilibrium becomes unstable. For small values of m, the curve folds near the origin, as for z^* near 0, P_{z^*} has three distinct positive roots (Proposition 3.2). For those z^*, the fast equilibria are all non viable (Lemma 2): numerically, the system goes to extinction if the initial dominant trait is near 0 (color figure online)

sexual reproduction operator according to the infinitesimal model (Fisher 1919; Bulmer 1971; Barton et al. 2017), assuming that the segregational variance is constant and independent of the families. We have set our analysis in a regime of small variance of segregation, aligning with a framework developed by Diekmann et al. (2005), Perthame and Barles (2008) and recurrently used with the infinitesimal model (Bouin et al. 2018; Calvez et al. 2019). By doing so, we showed two types of result. First, we compared the system of moments derived from our model in the limit of small variance with a seminal work in quantitative genetics (Ronce and Kirkpatrick 2001), showing their equivalence in that limit, while bypassing any prior normality assumption on the trait distributions. Next, we showed that this small variance regime discriminates two time scales, allowing to perform a slow-fast analysis, which reduces the complexity of our system in the asymptotic limit. Thus, we were able to fully derive analytically its equilibria thanks to algebraic arguments of symmetry reflecting the symmetrical habitats. The theoretical outcomes of our model are shown in the upper panel of Fig. 6. They are to be compared to numerical outcomes shown in the lower panel, where the same colours indicate the same types of equilibria. For the numerical analysis, for

each couple of parameters (m, g), the initial state is the same: both local distributions are normal of same mean (0.2) and same variance $\varepsilon^2 = 2.5 \times 10^{-3}$. The initial state is taken as monomorphic so that it falls within the scope of the slow-fast analysis. Moreover, the color yellow is attributed to simulations whose final state does not meet the small segregational variance regime analysis prediction, which in particular states that the distribution of trait in the metapopulation has a variance of order ε^2 (see (12) and recall that the population is monomorphic (Sect. 3)). In the two simulations that present the color yellow, the variance in trait in the metapopulation is of approximately $3\varepsilon^2$, which exceeds the chosen threshold ($2\varepsilon^2$). The detailed setting and scoring of the simulations involved in the lower panel of Fig. 6 are available in "Appendix I".

One can notice that the justification of the validity of the Gaussian approximation of local trait distributions in the regime of small variance [see Sect. 1 and Bouin et al. (2018)] and most of the slow-fast analysis (Sect. 3) are robust when introducing asymmetries in our model, or changing the selection functions. However, we stress that our analytical derivation of the equilibria in the asymptotic limit uses specific arguments that rely crucially on the symmetries between habitats in our model (see Remark 4.3 and Proposition 4.2).

Robustness with regard to dimorphic initial state The theoretical outcomes given in Fig. 6 are in particular a consequence of the reduction of system due to the slow-fast theorem, which applies provided that the initial state is close enough from a fast equilibrium from the slow manifold (see Theorem 3.1). Those fast equilibria are monomorphic. A natural question would be to ask to what extent those results apply for an initial state that is dimorphic. This would model for example two initially isolated subpopulations, locally adapted, that are suddenly being connected. Here we give a numerical taste of what a more complete answer could look like. We display Fig. 7 using the same methodology and scoring than for the lower panel of Fig. 6, the only difference being the initial state, now constituted by two locally adapted subpopulations, slightly asymmetrical in size (see "Appendix I" for details). To get a sense of what could occur in the regime of vanishing variance, we choose to display the results for two values of ε^2: $\varepsilon^2 = 2.5 \times 10^{-3}$ (upper panel) and $\varepsilon^2 = 6.25 \times 10^{-4}$ (lower panel). Both panels of Fig. 7 and the lower panel of Fig. 6 are globally quite similar, except for the yellow region that is much wider in both panels of Fig. 7. Particularly, there is a net trend for strong selection and small migration. That is expected, because the initial state of the simulations involved in both panels of Fig. 7 is presumably far from the conditions asked by Theorem 3.1. These simulations suggest that, in this particular range of parameters, the fast relaxation to a monomorphic state, that is central in Theorem 3.1, breaks down and dimorphism is maintained. However, we can note that this yellow region decreases for decreasing values of ε^2 (difference between upper and lower panel of Fig. 7). That suggests that our analysis remains quite robust to dimorphic initial states in the limit of vanishing variance.

Comparison with asexual studies In Sect. 4, we found that bistable asymmetrical equilibria can exist in our system (Proposition 4.2, Corollary 3). That is a strong difference with the findings of Mirrahimi (2017) and Mirrahimi and Gandon (2020): with a similar mesoscopic model but using an asexual reproduction operator with frequent mutations of small effects, they find that symmetrical habitats lead to a single
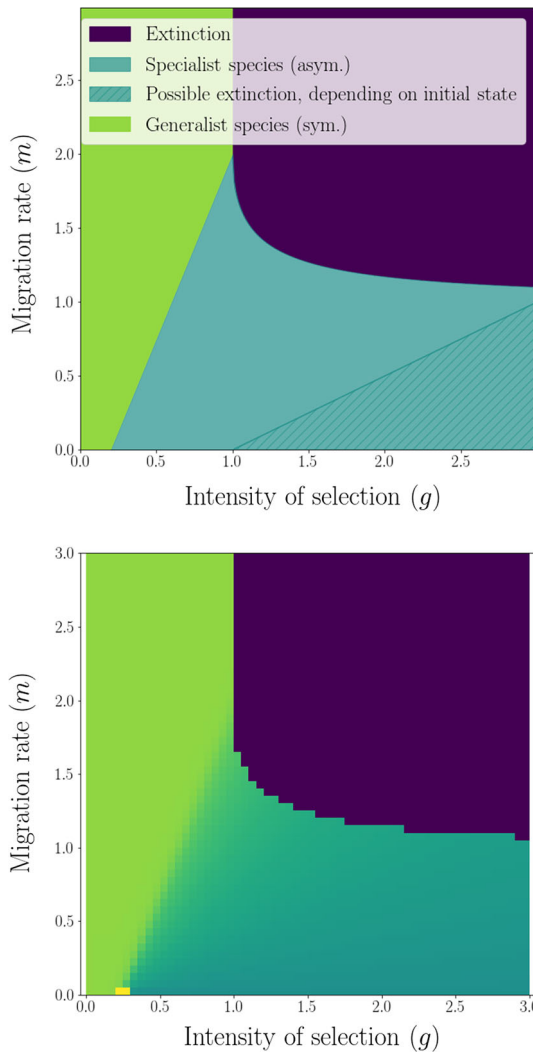
Fig. 6 Summary of the different theoretical (upper panel, in the limit of vanishing segregational variance) and numerical (lower panel, $\varepsilon^2 = 2.5 \times 10^{-3}$) outcomes of our model when selection (g) and migration (m) are varying. The same colors represents the same outcomes in both figures. Figure 4 complemented by the stability analysis (see Sect. 4.2) gives the upper figure. In the dashed region, the system goes to one of the asymmetrical equilibrium, except if the initial conditions are too symmetrical (the system goes then numerically to extinction, typically due to the folding near $z^* = 0$ of the yellow curve in Fig. 5). For the lower figure, all simulations share the same initial state: the metapopulation is monomorphic and asymmetrical as local distributions are both normal with same mean (0.2) and same variance ($2\varepsilon^2$). Hence, the potential extinction in the dashed region does not occur and the numerical analyis falls within the scope of the slow-fast analysis (Sect. 3). The color yellow is attributed to simulations whose final state does not meet the small segregational variance regime analysis prediction, which, in particular, states that the distribution of trait in the metapopulation has a variance of order ε^2 (see (12) and recall that the population is monomorphic (Sect. 3)). For more details on the simulations and their scoring resulting in the lower panel figure, see "Appendix I" (color figure online)
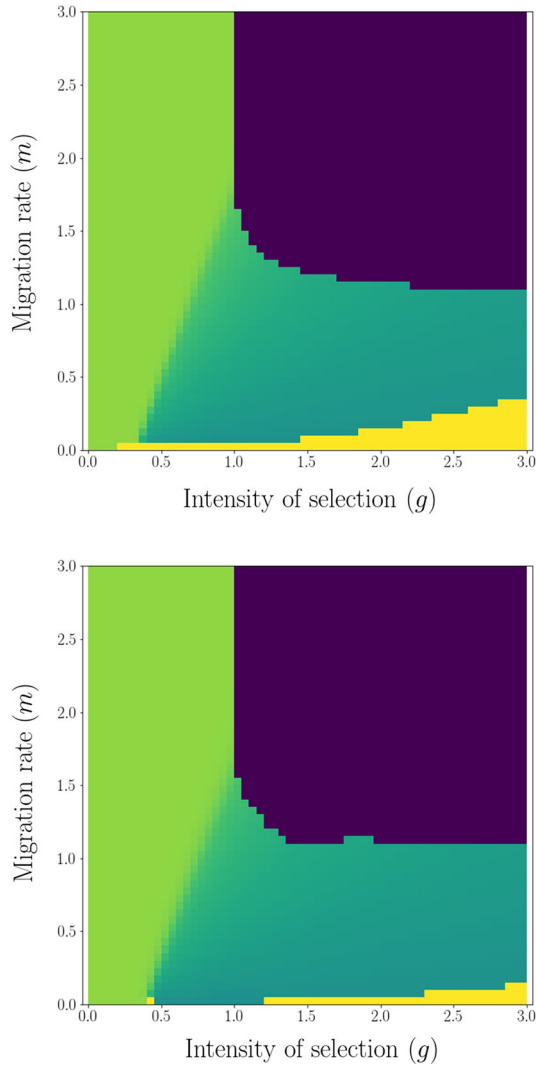
Fig. 7 Numerical outcomes with initial dimorphic state (locally adapted subpopulations), for $\varepsilon^2 = 2.5 \times 10^{-3}$ (upper panel) and $\varepsilon^2 = 6.25 \times 10^{-4}$ (lower panel). The colors for both figures results from the same scoring scheme as for Fig. 6 (see "Appendix I" for details). The results are quite similar as Fig. 6, except for the yellow region. In the upper panel ($\varepsilon^2 = 2.5 \times 10^{-3}$), the yellow region is wider than when the initial state is monomorphic (Fig. 6), increasingly so for stronger selection. That highlights the numerical cases where the population ends up dimorphic as the species adapts locally to each deme's optimum, for strong selection and small migration. This is expected as the fast convergence toward a monomorphism state induced by Theorem 3.1 in the limit of vanishing variance of segregation is likeky to break down, as the initial state is far from the slow manifold and the segregational variance is small but not zero. However, this yellow region decreases as the value of ε^2 decreases, as indicated by the lower panel ($\varepsilon^2 = 6.25 \times 10^{-4}$). That suggests that our analysis remains quite robust to dimorphic initial states in the limit of vanishing variance (color figure online)

stable symmetrical equilibrium, either monomorphic or dimorphic. In particular, if migration is small enough compared to selection, each subpopulation adapts to their habitats and dimorphism occurs at the metapopulation scale. In our case, the mixing effect of the infinitesimal operator of sexual reproduction does not allow for such a local adaptation to occur in the limit of small variance. In Sect. 3, we showed that it forces monomorphism quickly and the only option to adapt to strong forces of selection is an asymmetrical equilibrium (Proposition 4.2, Fig. 6) that describes a source sink scenario. One population is adapted to its habitat, and the other is essentially composed by poorly adapted migrants ; the choice of which depends on the initial conditions.

Our findings share notable similarities with some in Débarre et al. (2013), which conducts a hybrid analysis on asexual populations with tools of adaptive dynamics applied to quantitative genetics equations. Particularly, under gradual evolution (when mutations are rare and of small effects), they state that asymmetrical equilibria can be reached if the population is initially monomorphic, under a similar range of migration and selection parameters as indicated by our analysis. To solve for them, they assume that the distributions of traits around each peak found using adaptive dynamics is Gaussian, of constant variance related to the mutational variance (which is small by hypotheses). That is similar to the framework that naturally arise from the hypotheses of our model, should the mutational variance be replaced by the segregational variance. Consequently, we suggest that the asymmetrical equilibria found in Débarre et al. (2013) should have the same coordinates as the ones found in our analysis. However, there is a substantial difference in the dynamics leading to those equilibria. Even with an initially dimorphic metapopulation, our hypotheses on sexual reproduction typically strains toward monomorphism. With the same initial state, Débarre et al. (2013) indicate that dimorphism is typically maintained in the range of parameters where asymmetrical equilibria exist.

Gaussian assumption In our study, we consider a regime where the segregational variance is small compared to how far apart the local optimal traits are. While this small variance regime is more general than the standard weak selection approximation widely used in quantitative genetics model using the Gaussian assumption (see Remark 1.2), we formally show that the local trait distributions can still be well approximated by normal distributions within this regime (Sect. 1). Hence, asymptotically, in the regime of small variance, the findings of our model are equivalent to Ronce and Kirkpatrick (2001), which relies on a Gaussian assumption of local trait distributions. This link of equivalence relies on the hypothesis that the genetic (and phenotypic) variance is constant, which we interpreted in our model to be twice the segregational variance in the limit of vanishing variance. Furthermore, together with the last paragraph, our study gives some elements of explanation to why the findings of Ronce and Kirkpatrick (2001) are structurally different from Mirrahimi (2017) and Mirrahimi and Gandon (2020), and closer to Débarre et al. (2013).

Constant segregational variance in a heterogeneous environment Our model relies on using the infinitesimal model with a constant segregational variance, independent of the mates deme. That is a strong assumption. However, in the perspective of linking the present study to population genetics approaches, one can question the limits of such

a modelling assumption with regard to a Mendelian interpretation of this model. A future work is planned to examine it through conducting individual based simulations.

Appendices

A System of moments derived from our model

Here, we derive the system of moments (2) from (1). In the preliminary computations, we will omit the time and deme dependency for the sake of clarity. We will then denote n the trait distribution density, N the size of the population, $\overline{z}$ the mean trait, σ^2 the mean variance, ψ^3 the third central moment and θ the optimal phenotype.

Preliminary integration of the selection term We have:

$$\int_{\mathbb{R}} (z - \theta)^2 n(z) dz = \int_{\mathbb{R}} \left[(z - \overline{z})^2 + (\overline{z} - \theta)^2 + 2(z - \overline{z})(\overline{z} - \theta) \right] n(z)\, dz$$
$$= \sigma^2 N + (\overline{z} - \theta)^2 N,$$

and:

$$\int_{\mathbb{R}} (z - \overline{z})(z - \theta)^2 n dz$$
$$= \int_{\mathbb{R}} \left[(z - \overline{z})^3 + (z - \overline{z})(\overline{z} - \theta)^2 + 2(z - \overline{z})^2(\overline{z} - \theta) \right] n(z)\, dz$$
$$= 2\sigma^2(\overline{z} - \theta)N + \psi N.$$

Size of the subpopulations Recalling that $N_i(t) = \displaystyle\int_{\mathbb{R}} n_i(t, z)\, dz$, we get from the preliminary computations by integrating (1):

$$\frac{dN_i}{dt} = \int_{\mathbb{R}} \frac{\partial n_i}{\partial t}(t, z) dz$$
$$= \int_{\mathbb{R}} r\mathcal{B}_\sigma(n_i)(t, z) - g(z - \theta_i)^2 n_i(t, z)$$
$$\quad - \kappa N_i(t) n_i(t, z) + m\, (n(t, z) - n(t, z))\ dz$$
$$= \left[r - \kappa N_i(t) - g(\overline{z}_i(t) - \theta_i)^2 - g\sigma_i^2 \right] N_i(t) + m\big(N_j(t) - N_i(t)\big).$$

Local mean trait Recalling that $\overline{z}_i(t) = \frac{1}{N_i(t)} \int_{\mathbb{R}} z\, n_i(t, z)\, dz$, we have, thanks to the preliminary computations:

$$
\begin{aligned}
\frac{dz_i}{dt} &= \frac{1}{N_i} \int_{\mathbb{R}} z \frac{\partial n_i}{\partial t}(t, z) dz - \frac{1}{N_i{}^2} \frac{dN_i}{dt} \int_{\mathbb{R}} z n_i(t, z) dz \\
&= \frac{1}{N_i} \int_{\mathbb{R}} (z - \overline{z}_i) \frac{\partial n_i}{\partial t}(t, z) dz \\
&= \frac{1}{N_i} \int_{\mathbb{R}} (z - \overline{z}_i) \Big[-g(z - \theta_i)^2 n_i(t, z) + m \left(n_j(t, z) - n_i(t, z) \right) \Big] dz \\
&= 2g\sigma_i{}^2(\theta_i - \overline{z}_i) - g\psi_i{}^3 + m \frac{N_j}{N_i}(\overline{z}_j - \overline{z}_i).
\end{aligned}
$$

B Equilibria of a dynamical system under the infinitesimal model of reproduction with random mating only

In this subsection, we show that (7) admits any Gaussian of variance ε^2 as equilibrium. That is equivalent to state that:

Proposition B.1 *For $\mu \in \mathbb{R}$, the Gaussian distribution G_{μ, ε^2} of mean μ and variance ε^2 is a fixed point of the operator $\mathcal{B}_\varepsilon$, namely:*

$$
\mathcal{B}_\varepsilon(G_{\mu, \varepsilon^2}) = G_{\mu, \varepsilon^2}.
$$

Proof We can first notice that $\mathcal{B}_\varepsilon$ can be written using a double convolution product: $\qquad\square$

Lemma 10 *For $f \in \mathcal{L}_1(\mathbb{R})$, $\int_{\mathbb{R}} f \neq 0$, we have:*

$$
\mathcal{B}_\varepsilon(f) = \frac{4}{\displaystyle\int_{\mathbb{R}} f(z')\, dz'} G_{0, \frac{\varepsilon^2}{2}} * F * F,
$$

where $F : z \mapsto f(2z)$.

Proof (Proof of Lemma 10) For $f \in \mathcal{L}_1(\mathbb{R})$, $\int_{\mathbb{R}} f \neq 0$, a straight-forward computation yields:

$$
\begin{aligned}
\mathcal{B}_\varepsilon(f)(z) &= \frac{1}{\sqrt{\pi}\varepsilon} \iint_{\mathbb{R}^2} \exp\left[\frac{-(z - \frac{z_1 + z_2}{2})^2}{\varepsilon^2} \right] \frac{f(z_1) f(z_2)}{\displaystyle\int_{\mathbb{R}} f(z')\, dz'} dz_1 dz_2 \\
&= \frac{1}{\displaystyle\int_{\mathbb{R}} f(z')\, dz'} \int_{\mathbb{R}} \int_{\mathbb{R}} G_{0, \frac{\varepsilon^2}{2}} \left((z - \frac{z_1}{2}) - \frac{z_2}{2} \right) F(\frac{z_2}{2}) dz_2\, F(\frac{z_1}{2})\, dz_1
\end{aligned}
$$

$$= \frac{2}{\int_{\mathbb{R}} f(z')\, dz'} \int_{\mathbb{R}} G_{0,\frac{\varepsilon^2}{2}} * F(z - \frac{z_1}{2}) F(\frac{z_1}{2})\, dz_1$$

$$= \frac{4}{\int_{\mathbb{R}} f(z')\, dz'}\, G_{0,\frac{\varepsilon^2}{2}} * F * F(z).$$

□

If $f = G_{\mu,\varepsilon^2}$, then we find $F = \frac{1}{2} \times G_{\frac{\mu}{2},\frac{\varepsilon^2}{4}}$. Besides, as the convolution product of two Gaussian kernels G_{μ_1,σ_1^2} and G_{μ_2,σ_2^2} is the Gaussian kernel $G_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}$, Proposition B.1 is a corollary of the previous lemma.

C Formal expansion within the exponential formalism for n_ε

In this subsection, we will remove the deme dependency for the sake of clarity. To formally derive (9), let us consider the following formal expansion of U_ε with regard to successive orders of ε^2:

$$U_\varepsilon = u_0 + \varepsilon^2 u_\varepsilon.$$

The aim is to characterize u_0 thanks to the behaviour of the reproduction term when $\varepsilon \ll 1$, which we expect neither to diverge nor to vanish:

$$\frac{\mathcal{B}_\varepsilon(n_\varepsilon)}{n_\varepsilon}(z)$$
$$= \frac{1}{\sqrt{\pi}\varepsilon} \iint_{\mathbb{R}^2} \frac{\exp\left[\frac{1}{\varepsilon^2}\left[-\left[z - \frac{z_1+z_2}{2}\right]^2 + u_0(z) - u_0(z_1) - u_0(z_2)\right]\right] \exp\left[u_\varepsilon(z) - u_\varepsilon(z_1) - u_\varepsilon(z_2)\right] dz_1 dz_2}{\int_{\mathbb{R}} \exp\left[-\frac{u_0(z')}{\varepsilon^2} - u(z')\right] dz'}$$

Then, we have several considerations to make. First, if we assume that u_0 reaches its minimum at a non degenerate point z^*, then the following modified expression of the denominator:

$$\frac{1}{\sqrt{\pi}\varepsilon} \int_{\mathbb{R}} \exp\left[-\frac{1}{\varepsilon^2}\left[u_0(z') - \min u_0\right] - u(z')\right] dz',$$

will have its integrand concentrate around the minimum of u_0 and will converge as $\varepsilon \ll 1$. Therefore it is relevant to introduce this minimum both at the numerator and the denominator.

Then, since we expect the numerator not to diverge nor to vanish uniformly as $\varepsilon \ll 1$, we need that:

$$\forall z \in \mathbb{R}, \max_{(z_1,z_2)} \left[-\left(z - \frac{z_1+z_2}{2}\right)^2 + u_0(z) - u_0(z_1) - u_0(z_2) + \min u_0\right] = 0.$$

(27)

As shown in Bouin et al. (2018), thanks to some convexity arguments, this leads necessarily to choose u_0 as a quadratic function in z, hence its decomposition:

$$u_0(z) = u(z^*) + \frac{(z - z^*)^2}{2}, \tag{28}$$

where z^* is realizing the minimum of u_0. Note that $u(z^*) = 0$, due to the Laplace method of integration, since:

$$N_\varepsilon = \frac{1}{\sqrt{2\pi}\,\varepsilon} \int_{\mathbb{R}} \exp\left[-\frac{U_\varepsilon(z)}{\varepsilon^2}\right] dz \underset{\varepsilon \to 0}{\approx} \frac{\exp\left[-\frac{u(z^*)}{\varepsilon^2}\right]}{\sqrt{U_\varepsilon''(z^*)}}.$$

So either $u(z^*) = 0$, either there is extinction or explosion of the population size. That yields (9).

Convexity arguments from Bouin et al. (2018). Let us recall the arguments of convexity involved in Bouin et al. (2018) to show that functional constraint (27) leads in our case to u_0 being quadratic:

1. First, they show that u_0 has some regularities (continuous and has left and right derivative everywhere), for (27) implies that $z \mapsto u_0(z) - z^2$ is concave as minimum of affine functions:

$$\forall z \in \mathbb{R}, \quad u_0(z) - z^2 = \min_{(z_1, z_2)} \left[-z(z_1 + z_2) + \frac{(z_1 + z_2)^2}{4} + u_0(z_1) + u_0(z_2) \right].$$

2. Next, they introduce the Legendre convex conjugate

$$\hat{u}_0 : p \mapsto \sup_{z \in \mathbb{R}} \left[(z - z^*)p - u(z) \right],$$

and show that it satisfies the following functional equality, by commuting the different sup operators while computing $\hat{u}_0(p)$ using (27):

$$\forall p \in \mathbb{R}, \quad \hat{u}_0(p) = \frac{p^2}{4} + 2\,\hat{u}_0\left(\frac{p}{2}\right). \tag{29}$$

3. As $\hat{u}_0$ is convex by definition, it is continuous and admits left and right derivative everywhere. Moreover, $\hat{u}_0$ has a minimum in 0 and $\hat{u}_0(0) = -u(z^*) = 0$. Therefore (29) implies by recursion:

$$\forall p > 0 \ \text{(resp.} < 0), \quad \hat{u}_0(p) = \frac{p^2}{2} + \hat{u}_0{}'(0^+)\,p \ \text{(resp. } \hat{u}_0{}'(0^-)\,p). \tag{30}$$

Note that 0 being a minimum of $\hat{u}_0$ implies that: $\hat{u}_0{}'(0^-) \leq 0 \leq \hat{u}_0{}'(0^+)$.

4. The next step aims at showing that u_0 is equal to its convex bi-conjugate

$$\hat{\hat{u}}_0 : z \mapsto \sup_{p \in \mathbb{R}} \left[p \, (z - z^*) - \hat{u}_0(p) \right],$$

which is computable from (30):

$$\hat{\hat{u}}_0 : z \mapsto \begin{cases} \frac{(z - z^* - \hat{u}_0(0^-))^2}{2} & \text{if } z < z^* + \hat{u}_0(0^-) \\ 0 & \text{if } z^* + \hat{u}_0(0^-) \leq z \leq z^* + \hat{u}_0(0^+) \\ \frac{(z - z^* - \hat{u}_0(0^+))^2}{2} & \text{if } z > z^* + \hat{u}_0(0^+). \end{cases} \tag{31}$$

Standard convexity analysis shows also that $\hat{\hat{u}}_0$ is the lower convex envelope of u_0.

The first implication is that u_0 and $\hat{\hat{u}}_0$ coincide on $\mathbb{R} \backslash [z^* + \hat{u}_0(0^-), z^* + \hat{u}_0(0^+)]$, because $\hat{\hat{u}}_0$ is strictly convex there.

The second implication is that $u_0 \left(z^* + \hat{u}_0(0^+) \right) = \hat{\hat{u}}_0 \left(z^* + \hat{u}_0(0^+) \right) = 0$ (resp. $z^* + \hat{u}_0(0^-)$), since $z^* + \hat{u}_0(0^+)$ (resp. $z^* + \hat{u}_0(0^-)$) is an extremal point of the graph of $\hat{\hat{u}}_0$. One can show using (27) that the midpoint between any zeros of u_0 is still a zero of u_0 (recall that $u_0 \geq 0$). Hence, by density and continuity of u_0, u_0 vanishes on $[z^* + \hat{u}_0(0^-), z^* + \hat{u}_0(0^+)]$.

5. Finally, since u_0 satisfies (31) and we need N_ε not to explode when ε vanishes, we necessarily obtain that $\hat{u}_0(0^-) = \hat{u}_0(0^+)$. Hence u_0 quadratic.

D Formal approximations of the trait distributions moments in the regime of small variance $\varepsilon^2 \ll 1$

This appendix is dedicated to formally explain (12). We remove the time and the deme dependency for the sake of clarity. We denote n_ε the trait distribution density, N_ε the size of the population, $\overline{z}_\varepsilon$ the mean trait, σ_ε^2 the variance and ψ_ε the third central moment. Let us also recall that the computations are performed using the exponential formalism introduced in (10) while considering the following formal expansion of u_ε in the regime of small variance:

$$u_\varepsilon = u + \varepsilon^2 \, v + \mathcal{O}(\varepsilon^4).$$

Size of population We have:

$$\begin{aligned} N_\varepsilon &= \int_{\mathbb{R}} n_\varepsilon(z) \, dz \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\varepsilon} e^{-\frac{(z - z^*)^2}{2\varepsilon^2}} e^{-u(z) - \varepsilon^2 \, v(z) + \mathcal{O}(\varepsilon^4)} dz \\ &= \int_{\mathbb{R}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{-u(z^* + \varepsilon y) - \varepsilon^2 v(z^* + \varepsilon y) + \mathcal{O}(\varepsilon^4)} dy \quad \left(y := \frac{z - z^*}{\varepsilon} \right) \end{aligned}$$

$$= \int_{\mathbb{R}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{-[u(z^*)+\varepsilon y u'(z^*)+\frac{\varepsilon^2 y^2}{2}u''(z^*)+\frac{\varepsilon^3 y^3}{6}u'''(z^*)+\mathcal{O}(\varepsilon^4)]-\varepsilon^2 v(z^*)-\varepsilon^3 y v'(z^*)+\mathcal{O}(\varepsilon^4)} dy$$

$$= \int_{\mathbb{R}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{-u(z^*)} e^{-\left[\varepsilon y u'(z^*)+\varepsilon^2\left[\frac{y^2 u''(z^*)}{2}+v(z^*)\right]+\varepsilon^3\left[\frac{y^3}{6}u'''(z^*)+y v'\right]+\mathcal{O}(\varepsilon^4)\right]} dy$$

$$= \int_{\mathbb{R}} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{-u(z^*)} \left[1 - \varepsilon y u'(z^*) - \varepsilon^2 \left[\frac{y^2 u''(z^*)}{2} + v(z^*) \right] - \varepsilon^3 \left[\frac{y^3}{6} u'''(z^*) - y v'(z^*) \right] \right.$$
$$\left. + \frac{1}{2} \left[\varepsilon^2 y^2 u'(z^*)^2 + \varepsilon^3 \left[y^3 u'(z^*) u''(z^*) + 2 y u'(z^*) v(z^*) \right] \right] - \frac{\varepsilon^3 y^3 u'(z^*)^3}{6} + \mathcal{O}(\varepsilon^4) \right]$$

$$= e^{-u(z^*)} \left[1 + \varepsilon^2 \left[\frac{u'^2(z^*)}{2} - \frac{u''(z^*)}{2} - v(z^*) \right] \right] + \mathcal{O}(\varepsilon^4),$$

from the computations of the moments of a Gaussian. *Mean trait* Similarly as above, we have:

$$\bar{z}_\varepsilon = \int_{\mathbb{R}} z \frac{n_\varepsilon}{N_\varepsilon} dz$$

$$= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} z \frac{1}{\sqrt{2\pi}\varepsilon} e^{-\frac{(z-z^*)^2}{2\varepsilon^2}} e^{-u(z)-\varepsilon^2 v(z)+\mathcal{O}(\varepsilon^4)} dz$$

$$= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} (z^* + \varepsilon y) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{-u(z^*+\varepsilon y)-\varepsilon^2 v(z^*+\varepsilon y)+\mathcal{O}(\varepsilon^4)} dy, \qquad \left(y := \frac{z-z^*}{\varepsilon} \right)$$

$$= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} (z^* + \varepsilon y) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} e^{-u(z^*)} \left[1 - \varepsilon y u'(z^*) + \varepsilon^2 \left[\frac{y^2 u'(z^*)^2}{2} - \frac{y^2 u''(z^*)}{2} - v(z^*) \right] \right.$$
$$\left. + \varepsilon^3 \left[-\frac{y^3}{6} u'''(z^*) - y v'(z^*) + \frac{y^3 u'(z^*) u''(z^*)}{2} + y u'(z^*) v(z^*) - \frac{3 y^3 u'(z^*)^3}{6} \right] + \mathcal{O}(\varepsilon^4) \right]$$

$$= \frac{e^{-u(z^*)} \left[z^* \left(1 + \varepsilon^2 \left[\frac{u'^2(z^*)}{2} - \frac{u''(z^*)}{2} - v(z^*) \right] \right) - \varepsilon^2 u'(z^*) \right] + \mathcal{O}(\varepsilon^4)}{e^{-u(z^*)} \left(1 + \varepsilon^2 \left[\frac{u'^2(z^*)}{2} - \frac{u''(z^*)}{2} - v(z^*) \right] \right) + \mathcal{O}(\varepsilon^4)}$$

$$= z^* - \varepsilon^2 u'(z^*) + \mathcal{O}(\varepsilon^4).$$

Variance Using the previous formal computations and methodology, we get:

$$\sigma_\varepsilon^2 = \frac{1}{N_\varepsilon} \int_{\mathbb{R}} (z - \bar{z}_\varepsilon)^2 n_\varepsilon(z) dz$$

$$= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} \left[(z - z^*)^2 + (z^* - \bar{z}_\varepsilon)^2 + 2(z - z^*)(z^* - \bar{z}_\varepsilon) \right] n_\varepsilon(z) dz$$

$$= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} \left[\varepsilon^2 y^2 + 2\varepsilon^3 y u'(z^*) + \mathcal{O}(\varepsilon^4) \right] \left[1 - \varepsilon y u' + \mathcal{O}(\varepsilon^2) \right] e^{-u(z^*)} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$$

$$= \frac{\varepsilon^2 e^{-u(z^*)}}{e^{-u(z^*)} \left[1 + \mathcal{O}(\varepsilon^2) \right]}$$

$$= \varepsilon^2 + \mathcal{O}(\varepsilon^4).$$

Third central moment We compute, using the same change in variable $y := \frac{z - z^*}{\varepsilon}$:

$$
\begin{aligned}
\psi_\varepsilon^3 &= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} (z - \bar{z}_\varepsilon)^3 n_\varepsilon(z) dz \\
&= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} \Big[(z - z^*)^3 + (z^* - \bar{z}_\varepsilon)^3 + 3(z - z^*)^2 (z^* - \bar{z}_\varepsilon) + 3(z - z^*)(z^* - \bar{z}_\varepsilon)^2 \Big] \\
&\quad n_\varepsilon(z) dz \\
&= \frac{1}{N_\varepsilon} \int_{\mathbb{R}} \Big[\varepsilon^3 y^3 + \mathcal{O}(\varepsilon^4) \Big] \Big[e^{-u(z^*)} + \mathcal{O}(\varepsilon) \Big] dz \\
&= \mathcal{O}(\varepsilon^4).
\end{aligned}
$$

E Fast/slow system: proof of Theorem 3.1

This appendix is dedicated to prove Theorem 3.1.

Let $(z_0^*, \bar{Y}_0^*) \in \mathbb{R} \times \Omega$ (we recall that $\Omega = (\mathbb{R}_+^*)^2 \times \mathbb{R}$) be on the slow manifold, ie. such that $G(z_0^*, \bar{Y}_0^*) = 0$. From Lemma 6 of fast relaxation towards the slow manifold, the jacobian matrix $J_G(z_0^*, \bar{Y}_0^*)$ is invertible. Consequently, the implicit function theorem gives us U open neighbourhood of z_0^* in $\mathbb{R}$, V open neighbourhood of $(z_0^*, \bar{Y}_0^*)$ in $\mathbb{R} \times \Omega$ and $\phi \in C^\infty(U, V)$ such that :

$$
\forall (z^*, \bar{Y}^*) \in V, \ G(z^*, \bar{Y}^*, 0) = 0 \implies \bar{Y}^* = \phi(z^*).
$$

Hence, we can define a notation that we shall use henceforth:

$$
\forall z \in U, \ J_z := J_G(z, \phi(z)).
$$

If K is a compact subset of U such that $z_0^* \in \mathring{K}$, we can define the Cauchy problem (E_0) by the following :

$$
(E_0) \quad \begin{cases} \frac{dz^*}{dt} = -2g z^*(t) + F\left(\phi(z^*(t)) \right), \\ z^*(0) = z_0^*, \end{cases} \tag{32}
$$

for $t \leq t^*$, that we define as the following:

$$
t^* := \inf\{t > 0, z^*(t) \notin K\}.
$$

It is similar to (20) with the initial conditions $(z^*(0), \bar{Y}_0^*) = (z_0^*, \phi(z_0^*))$. A essential part of the proof relies in the fact that we can define the following uniform positive constant, thanks to Lemma 6 of fast relaxation:

$$
\lambda_K = -\frac{1}{2} \max_{z \in K} \{\lambda \in \mathbf{Sp}(J_z)\} > 0.
$$

As the first step, we state the following lemma whose proof will be provided at the end of this appendix. It defines a uniform control constant $\gamma > 0$:

Lemma 11 *There exists $\gamma > 0$ such that:*

$$\max_{z \in K,\, s \geq 0} \left\| \left\| e^{\lambda_K s} e^{J_z s} \right\| \right\| \leq \gamma.$$

($\| \cdot \|_{\mathcal{M}_3(\mathbb{R})}$ *is noted $\| \cdot \|$*).

The next step is to show the convergence of solutions of (P_ε) (19) towards those of (P_0) (20) on a time interval, yet to be defined, that will be shown to be uniform with regard to ε and the initial conditions, provided that they are small enough. For that purpose, it is more convenient to consider the system (R_ε) verified by the residuals $r_z^\varepsilon(t) = z_\varepsilon(t) - z^*(t)$ and $r_Y^\varepsilon(t) = \bar{Y}_\varepsilon(t) - \bar{Y}^*(t)$:

$$(R_\varepsilon) \quad \begin{cases} \varepsilon^2 \frac{dr_Y^\varepsilon}{dt} = G(z^*(t) + r_z^\varepsilon(t), \bar{Y}^*(t) + r_Y^\varepsilon(t)) - G(z^*(t), \bar{Y}^*(t)) - \varepsilon^2 \frac{d\bar{Y}^*}{dt} + \varepsilon^2 v_{N,\varepsilon}(t), \\[2mm] \frac{dr_z^\varepsilon}{dt} = -2g r_z^\varepsilon(t) + F(\bar{Y}^*(t) + r_Y^\varepsilon) - F(\bar{Y}^*(t)) + \varepsilon^2 v_{z,\varepsilon}(t), \\[2mm] (r_z^\varepsilon(0), r_Y^\varepsilon(0)) = (z_0^\varepsilon - z_0^*, \bar{Y}_0^\varepsilon - \bar{Y}_0^*), \end{cases} \quad (33)$$

and introduce some further definitions.

Because K is a compact set, there exists $\delta_K > 0$ such that the following set is a compact subset of V:

$$\bar{K}_{\delta_K} = \{(z, \bar{Y}) \in \mathbb{R} \times \Omega | \exists z^* \in K, \ |(z, \bar{Y}) - (z^*, \phi(z^*))| \leq \delta_K \} \subset V.$$

Let us consider from now $(z_0^\varepsilon, N_0^\varepsilon) \in \bar{K}_{\delta_K}$. Then we define $\Delta = \min \left(\frac{\lambda_K}{4C\gamma}, \delta_K \right)$ and $T = \min(t^*, \frac{\lambda_K}{4C'\gamma})$, where:

$$C = \max \left(\|\partial_{\bar{Y}}^2 G\|_{\infty, \bar{K}_{\delta_K}}, \|\partial_z G\|_{\infty, \bar{K}_{\delta_K}}, \|\partial_{\bar{Y}} F\|_{\infty, \Pi_\Omega(\bar{K}_{\delta_K})} \right)$$

and :

$$C' = \max_{t \leq t^*} \left\| \left\| \partial_t J_{z^*(t)} \right\| \right\|,$$

where Π_Ω is the projection from $\mathbb{R} \times \Omega$ on Ω. One can notice from these definitions and from Lemma 11, that $\gamma, \Delta, T, \lambda_K, C, C'$ do not depend on ε and are uniform on $[0, t^*]$. Specifically taking $\Delta \leq \frac{\lambda_K}{4C\gamma}$ and $T \leq \frac{\lambda_K}{4C'\gamma}$ will turn out to be important in the proof.

On the time region $[0, T]$, we will show that we can control explicitly the various perturbed terms that arise. We can now state the following proposition, whose proof constitutes the core of the resolution of the problem:

Proposition E.1 *As* $\max(\varepsilon, |r_z^\varepsilon(0)|, |r_Y^\varepsilon(0)|) \to 0$, $(\bar{Y}_\varepsilon, z_\varepsilon)$ *converges toward* $(\bar{Y}^*, z^*)$ *uniformly on* $[0, T]$.

For the final step, we will show that we can reiterate the process on each interval of time $[jT, \min\{(j+1)T, t^*\}]$ with $\forall j \leq \lfloor \frac{t^*}{T} \rfloor$, $jT \leq t_\varepsilon^*$. Thus, for sufficiently small ε and initial conditions, the control remains valid until t^*, hence Theorem 3.1.

For convenience, we will denote by $f * g\,(t)$ the convolution product of f and g at time $t > 0$:

$$f * g\,(t) = \int_0^t f(\tau) g(t-\tau) d\tau.$$

Proof (Proof of Proposition E.1)

Let $\varepsilon \in]0, 1]$. Let us define an auxiliary time t_ε^*:

$$t_\varepsilon^* = \min\left(t^*, \inf\{t > 0, |r_z^\varepsilon| + |r_Y^\varepsilon| > \Delta\}\right).$$

It ensures that the perturbed trajectory stays inside of $\bar{K}_{\delta_K}$ when $t \leq t_\varepsilon^*$.

Let us highlight the main steps of the proof:

1. preliminary controls on r_Y^ε by $|r_Y^\varepsilon(0)|$ and $\frac{1}{\varepsilon^2}|r_z^\varepsilon| * e^{-\frac{\lambda_K}{2\varepsilon^2}}$ thanks to the regularity of G, the fast relaxation properties (Lemma 6 and Lemma 11) and Gronwall's lemma.
2. control $|r_z^\varepsilon|$ by $|r_z^\varepsilon(0)|$ and $|r_Y^\varepsilon|$.
3. finish the control on r_Y^ε by using the latter and Gronwall's lemma.

1. For $t \leq \min(T, t_\varepsilon^*)$, we can introduce new terms in the equation from (33) on r_Y^ε:

$$\frac{dr_Y^\varepsilon}{dt} = \frac{J_{z^*(0)}}{\varepsilon^2} r_Y^\varepsilon + \frac{1}{\varepsilon^2}\left[G(z^*(t), \bar{Y}^*(t) + r_Y^\varepsilon(t)) - G(z^*(t), \bar{Y}^*(t)) - J_{z^*(0)} r_Y^\varepsilon\right]$$

$$+ \frac{1}{\varepsilon^2}\left[G(z^*(t) + r_z^\varepsilon(t), \bar{Y}^*(t) + r_Y^\varepsilon(t)) - G(z^*(t), \bar{Y}^*(t) + r_Y^\varepsilon(t))\right]$$

$$- \phi'(z^*(t))(-2gz^*(t) + F(\phi(z^*(t)))) + \nu_{N,\varepsilon}(t)$$

$$= \frac{J_{z^*(0)}}{\varepsilon^2} r_Y^\varepsilon + A_1(t) + A_2(t) + A_3(t).$$

Since $t \leq \min(T, t_\varepsilon^*)$ and G is C^∞ on $\bar{K}_{\delta_K} \times [0, 1]$, we can control A_1:

$$|A_1(t)| \leq \frac{1}{\varepsilon^2}\left[|G(z^*(t), \bar{Y}^*(t) + r_Y^\varepsilon(t)) - G(z^*(t), \bar{Y}^*(t) - J_{z^*(t)} r_Y^\varepsilon|\right]$$

$$+ \frac{1}{\varepsilon^2}\left[\|\|J_{z^*(t)} - J_{z^*(0)}\|\|\,|r_Y^\varepsilon(t)|\right]$$

$$\leq \frac{1}{\varepsilon^2}\left[\|\partial_{\bar{Y}}^2 G\|_{\infty, \bar{K}_{\delta_K}}\,|r_Y^\varepsilon(t)|^2 + T \max_{t \leq t^*}\|\|\partial_t J_{z^*(t)}\|\|\,|r_Y^\varepsilon(t)|\right]$$

$$\leq \frac{1}{\varepsilon^2}(C\Delta + C'T)|r_Y^\varepsilon(t)|,$$

and A_2:

$$|A_2(t)| = \frac{1}{\varepsilon^2}|G(z^*(t) + r_z^\varepsilon(t), \bar{Y}^*(t) + r_Y^\varepsilon(t)) - G(z^*(t), \bar{Y}^*(t) + r_Y^\varepsilon(t))|$$

$$\leq \frac{1}{\varepsilon^2}\|\partial_z G\|_{\infty, \bar{K}_{\delta_K}}|r_z^\varepsilon(t)| \leq \frac{C}{\varepsilon^2}|r_z^\varepsilon(t)|,$$

and A_3:

$$|A_3(t)| = |-\phi'(z^*(t))(-2gz^*(t) + F(\phi(z^*(t)))) + \nu_{N,\varepsilon}(t)| \leq C'',$$

for some constant C'' independent of ε and $z^*(0) \in K$. Using Duhamel formulas, we get, for $t \leq \min(T, t_\varepsilon^*)$:

$$r_Y^\varepsilon(t) = e^{\frac{J_{z^*(0)}t}{\varepsilon^2}}r_Y^\varepsilon(0) + \left[e^{\frac{J_{z^*(0)} \cdot}{\varepsilon^2}} * (A_1 + A_2 + A_3)\right](t). \tag{34}$$

Hence, applying Lemma 11 yields:

$$|r_Y^\varepsilon(t)| \leq \gamma|r_Y^\varepsilon(0)|e^{-\frac{\lambda_K t}{\varepsilon^2}} + \frac{\gamma}{\varepsilon^2}\left[\left(C|r_z^\varepsilon| + (C\Delta + C'T)|r_Y^\varepsilon|\right) * e^{-\frac{\lambda_K \cdot}{\varepsilon^2}}\right](t)$$

$$+ \varepsilon^2\gamma\frac{C''}{\lambda_K}$$

$$\leq A^{r_z^\varepsilon}(t) + \frac{\gamma(C\Delta + C'T)}{\varepsilon^2}\int_0^t |r_Y^\varepsilon(\tau)| e^{\frac{\lambda_K}{\varepsilon^2}(\tau - t)}d\tau,$$

where $A^{r_z^\varepsilon}(t) := \gamma|r_Y^\varepsilon(0)|e^{-\frac{\lambda_K t}{\varepsilon^2}} + \frac{\gamma C}{\varepsilon^2}\left(|r_z^\varepsilon| * e^{-\frac{\lambda_K \cdot}{\varepsilon^2}}\right)(t) + \varepsilon^2\gamma\frac{C''}{\lambda_K}$.

Applying Gronwall inequality to $r_Y^\varepsilon(t)e^{\frac{\lambda_K t}{\varepsilon^2}}$ yields:

$$|r_Y^\varepsilon(t)| \leq A^{r_z^\varepsilon}(t) + \frac{\gamma(C\Delta + C'T)}{\varepsilon^2}\left[A^{r_z^\varepsilon} * e^{\left(\frac{-\lambda_K}{\varepsilon^2} + \frac{\gamma(C\Delta + C'T)}{\varepsilon^2}\right)\cdot}\right](t). \tag{35}$$

Having fixed $\Delta \leq \frac{\lambda_K}{4C\gamma}$ and $T \leq \frac{\lambda_K}{4C'\gamma}$ in the preliminaries ensures that $e^{\left(\frac{-\lambda_K}{\varepsilon^2} + \frac{\gamma(C\Delta + C'T)}{\varepsilon^2}\right)\cdot}$ defines a negative exponential term, that we can dominate by $e^{-\frac{\lambda_K}{2\varepsilon^2}\cdot}$. Hence:

$$|r_Y^\varepsilon(t)| \leq A^{r_z^\varepsilon}(t) + \left[A^{r_z^\varepsilon} * \frac{\lambda_K}{2\varepsilon^2}e^{-\frac{\lambda_K}{2\varepsilon^2}\cdot}\right](t). \tag{36}$$

Making $A^{r_z^\varepsilon}$ explicit gives:

$$|r_Y^\varepsilon(t)| \leq \gamma|r_Y^\varepsilon(0)|e^{-\frac{\lambda_K t}{\frac{\lambda_K}{\Delta}}} + \frac{\gamma C}{\varepsilon^2}|r_z^\varepsilon| * e^{-\frac{\lambda_K}{\varepsilon^2}\cdot}(t) + \varepsilon^2\gamma\frac{C''}{\lambda_K}$$

$$+ \left[\left(\gamma |r_Y^\varepsilon(0)| e^{-\frac{\lambda_K \cdot}{\varepsilon^2}} + \frac{\gamma C}{\varepsilon^2} \left[|r_z^\varepsilon| * e^{-\frac{\lambda_K \cdot}{\varepsilon^2}} \right] + \varepsilon^2 \gamma \frac{C''}{\lambda_K} \right) * \left(\frac{\lambda_K}{2\varepsilon^2} e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}} \right) \right] (t)$$

$$\leq \gamma |r_Y^\varepsilon(0)| \left[e^{-\frac{\lambda_K t}{\varepsilon^2}} + e^{-\frac{\lambda_K \cdot}{\varepsilon^2}} * \left(\frac{\lambda_K}{2\varepsilon^2} e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}} \right)(t) \right] + \varepsilon^2 \gamma \frac{C''}{\lambda_K} \left(\left(1 + \int_0^t \frac{\lambda_K}{2\varepsilon^2} e^{-\frac{\lambda_K}{2\varepsilon^2}(\tau - t)} dt \right) \right)$$

$$+ \frac{\gamma C}{\varepsilon^2} |r_z^\varepsilon| * \left(e^{-\frac{\lambda_K \cdot}{\varepsilon^2}} + e^{-\frac{\lambda_K \cdot}{\varepsilon^2}} * \frac{\lambda_K}{2\varepsilon^2} e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}} \right)(t), \tag{37}$$

thanks to the associativity of the convolution product. One can compute that, for $t \geq 0$:

$$e^{-\frac{\lambda_K t}{\varepsilon^2}} + e^{-\frac{\lambda_K \cdot}{\varepsilon^2}} * \left(\frac{\lambda_K}{2\varepsilon^2} e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}} \right)(t) = e^{-\frac{\lambda_K t}{\varepsilon^2}} + \frac{\lambda_K}{2\varepsilon^2} \int_0^t e^{-\frac{\lambda_K}{\varepsilon^2}\tau} e^{-\frac{\lambda_K}{2\varepsilon^2}(t - \tau)} d\tau$$

$$= e^{-\frac{\lambda_K t}{\varepsilon^2}} + \frac{\lambda_K}{2\varepsilon^2} \int_0^t e^{-\frac{\lambda_K}{2\varepsilon^2}(t + \tau)} d\tau = e^{-\frac{\lambda_K}{2\varepsilon^2}t}.$$

Hence, replacing those terms in (37) yields:

$$|r_Y^\varepsilon(t)| \leq \gamma |r_Y^\varepsilon(0)| e^{-\frac{\lambda_K}{2\varepsilon^2}t} + 2\varepsilon^2 \gamma \frac{C''}{\lambda_K} + \frac{C\gamma}{\varepsilon^2} |r_z^\varepsilon| * e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}}(t). \tag{38}$$

2. The next step is to gain similarly some control on $|r_z^\varepsilon|$. Using Duhamel formula on the equation from (33) on r_z^ε gives, for $t \leq \min(T, t_\varepsilon^*)$:

$$r_z^\varepsilon(t) = r_z^\varepsilon(0) e^{-2gt} + \left(\left[F(N^* + r_Y^\varepsilon) - F(N^*) + \varepsilon^2 v_{z,\varepsilon} \right] * e^{-2g \cdot} \right)(t),$$

which yields:

$$|r_z^\varepsilon(t)| \leq |r_z^\varepsilon(0)| e^{-2gt} + \varepsilon^2 \frac{\|v_{z,\varepsilon}\|_\infty}{2g} + \|\partial_{\bar{Y}} F\|_{\infty, \Pi_\Omega(\bar{K}_{\delta_K})} \left(|r_Y^\varepsilon| * e^{-2g \cdot} \right)(t).$$

Hence:

$$|r_z^\varepsilon(t)| \leq |r_z^\varepsilon(0)| e^{-2gt} + \varepsilon^2 \frac{\|v_{z,\varepsilon}\|_\infty}{2g} + C \left(|r_Y^\varepsilon| * e^{-2g \cdot} \right)(t). \tag{39}$$

At that point, it is clear that it is sufficient to control $|r_Y^\varepsilon|$ and $|r_z^\varepsilon(0)|$ in order to control $|r_z^\varepsilon(t)|$ for sufficiently small ε.

3. Plugging the latter in (38) gives:

$$|r_Y^\varepsilon(t)| \leq \gamma |r_Y^\varepsilon(0)| e^{-\frac{\lambda_K}{2\varepsilon^2}t} + \frac{C\gamma}{\varepsilon^2} |r_z^\varepsilon(0)| \left(e^{-2g \cdot} * e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}} \right)(t) + \varepsilon^2 \frac{C\gamma \|v_{z,\varepsilon}\|_\infty}{\lambda_K g}$$

$$+ 2\varepsilon^2 \gamma \frac{C''}{\lambda_K} + \frac{\gamma C^2}{\varepsilon^2} \left[|r_Y^\varepsilon| * \left(e^{-2g \cdot} * e^{-\frac{\lambda_K \cdot}{2\varepsilon^2}} \right) \right](t). \tag{40}$$

Similarly as the computation above, we have, for $\varepsilon^2 < \min(\frac{\lambda_K}{8g}, 1)$ and $t \geq 0$:

$$e^{-2g\cdot} * e^{-\frac{\lambda_K}{2\varepsilon^2}\cdot}(t) = \frac{1}{\frac{\lambda_K}{2\varepsilon^2} - 2g}\left(e^{-2gt} - e^{-\frac{\lambda_K}{2\varepsilon^2}t}\right) \leq \frac{4\varepsilon^2}{\lambda_K}e^{-2gt}.$$

Hence, for $\varepsilon^2 < \min(\frac{\lambda_K}{8g}, 1)$, we get from (40):

$$
\begin{aligned}
|r_Y^\varepsilon(t)| &\leq \gamma|r_Y^\varepsilon(0)|e^{-\frac{\lambda_K}{2\varepsilon^2}t} + \frac{2\gamma C}{\lambda_K}|r_z^\varepsilon(0)|e^{-2gt} + \varepsilon^2\frac{C\gamma\|v_{z,\varepsilon}\|_\infty}{\lambda_K g} + 2\varepsilon^2\gamma\frac{C''}{\lambda_K} \\
&\quad + \frac{2\gamma C^2}{\lambda_K}\left(|r_Y^\varepsilon| * e^{-2g\cdot}\right)(t) \\
&\leq C_0^\varepsilon(t) + \frac{2\gamma C^2}{\lambda_K}\left(|r_Y^\varepsilon| * e^{-2g\cdot}\right)(t),
\end{aligned}
$$

where we define: $C_0^\varepsilon(t) := \gamma|r_Y^\varepsilon(0)|e^{-\frac{\lambda_K}{2\varepsilon^2}t} + \frac{2\gamma C}{\lambda_K}|r_z^\varepsilon(0)|e^{-2gt} + \varepsilon^2\frac{C\gamma\|v_{z,\varepsilon}\|_\infty}{\lambda_K g} + 2\varepsilon^2\gamma\frac{C''}{\lambda_K}$.

Using once again Gronwall inequality on $|r_Y^\varepsilon|e^{2g\cdot}$ yields:

$$|r_Y^\varepsilon(t)| \leq C_0^\varepsilon(t) + \frac{2\gamma C^2}{\lambda_K}\left(C_0^\varepsilon * e^{\left(-2g + \frac{2\gamma C^2}{\lambda_K}\right)\cdot}\right)(t). \tag{41}$$

Recalling that:

$$C_0^\varepsilon(t) = \gamma|r_Y^\varepsilon(0)|e^{-\frac{\lambda_K}{2\varepsilon^2}t} + \frac{2\gamma C}{\lambda_K}|r_z^\varepsilon(0)|e^{-2gt} + \varepsilon^2\frac{C\gamma\|v_{z,\varepsilon}\|_\infty}{\lambda_K g} + 2\varepsilon^2\gamma\frac{C''}{\lambda_K},$$

we get that, thanks to (41) and (39), for a given $0 < \delta < \Delta$, there exists $\eta_\delta > 0$ depending only on δ, g, m, K, t^*, F, G, $\|v_{z,\varepsilon}\|_\infty$ such that :

$$\forall(\varepsilon, |r_Y^\varepsilon(0)|, |r_z^\varepsilon(0)|) \in [0, \eta_\delta]^3, \max_{t \leq \min(T, t_\varepsilon^*)}|r_Y^\varepsilon(t)| + |r_z^\varepsilon(t)| \leq \delta.$$

Recalling that $t_\varepsilon^* = \min\left(t^*, \inf\{t > 0, |r_z^\varepsilon| + |r_Y^\varepsilon| > \Delta\}\right)$, we get that $T \leq t_\varepsilon^*$, for $\delta < \Delta$ and $(\varepsilon, |r_Y^\varepsilon(0)|, |r_z^\varepsilon(0)|) \in [0, \eta_\delta]^3$. Consequently, the convergence is uniform on $[0, T]$. $\square$

Proof (Proof of Theorem 3.1) One can notice that the control obtained in the proof of Proposition 1 can be applied on any time interval $[a, a + T]$ with $a \in [0, t^* - T]$, provided that $(\varepsilon, |r_Y^\varepsilon(a)|, |r_z^\varepsilon(a)|)$ are small enough. Therefore, we can reiterate the control a finite number of times on the intervals $[jT, \min\{(j + 1)T, t^*\}]$ with $\forall j \leq \lfloor\frac{t^*}{T}\rfloor$. Hence, the uniform convergence on $[0, t^*]$. $\square$

Proof (Proof of Lemma 11) Recall that for all $z \in K$, J_z has real negative eigenvalues, uniformly bounded over K by $-2\lambda_K < -\lambda_K$. Let us define, for $z \in K$:

$$f_{\lambda_K, z} : \mathbb{R}_+ \to \mathbb{R}_+, \quad s \mapsto \left\| \left| e^{J_z s} e^{\lambda_K s} \right| \right\|.$$

For all $z \in \mathbb{K}$, $f_{\lambda_K, z}$ is continuous. Moreover, Theorem 2.34 of Chicone (1999) ensures that $f_{l,z}$ is bounded for all $l < 2\lambda_K$.

We can thus define :

$$\Gamma_{\lambda_K} : K \to \mathbb{R}_+^*, \quad z \mapsto \max_{s \geq 0} f_{\lambda_K, z}(s).$$

Let us show that Γ_{λ_K} is a continuous function. Let $z_0 \in K$ and $\varepsilon > 0$.

One can first notice that, for $s \geq 0$:

$$f_{\lambda_K, z}(s) = f_{\frac{3\lambda_K}{2}, z}(s) e^{-\frac{\lambda_K}{2} s} < \Gamma_{\frac{3\lambda_K}{2}, z} e^{-\frac{\lambda_K}{2} s}.$$

Thus, $f_{\lambda_K, z}$ vanishes when s goes to infinity. As a consequence, there exists $s_0 \geq 0$ such that:

$$\Gamma_{\lambda_K}(z_0) = \left\| \left| e^{J_{z_0} s_0} e^{\lambda_K s_0} \right| \right\|.$$

Furthermore, for $l \in]\lambda_K, 2\lambda_K[$, we have:

$$\Gamma_l(z_0) = \left\| \left| e^{J_{z_0} s_0} e^{l s_0} \right| \right\| = \Gamma_{\lambda_K}(z_0) e^{(l - \lambda_K) s_0}.$$

We can therefore choose $l \in]\lambda_K, 2\lambda_K[$ such that $\Gamma_{\lambda_K}(z_0) \leq \Gamma_l(z_0) \leq \Gamma_{\lambda_K}(z_0) + \varepsilon$.

As $z \mapsto J_z$ is a continuous function, there exists $\delta > 0$ that ensures that for if $z \in K$ and $|z - z_0| \leq \delta$, then:

$$\left\| \left| J_z - J_{z_0} \right| \right\| < \frac{l - \lambda_K}{2 \Gamma_l(z_0)}.$$

Let us consider such a z.

As $e^{J_z s}$ is solution of the ODE : $y' = J_{z_0} y + (J_z - J_{z_0}) y$, we obtain, for $s \geq 0$:

$$e^{J_z s} = e^{J_{z_0} s} + e^{J_{z_0} \cdot} * (J_z - J_{z_0}) e^{J_z \cdot}(s).$$

Hence :

$$\left\| \left| e^{J_z t} \right| \right\| \leq \Gamma_l(z_0) e^{-ls} + \frac{l - \lambda_K}{2} \left\| \left| e^{J_z \cdot} \right| \right\| * e^{-l \cdot}$$

From applying Gronwall's inequality to $t \mapsto \left\| \left| e^{J_z s} \right| \right\| e^{ls}$, it comes that, for $s \geq 0$:

$$\left\| \left| e^{J_z s} \right| \right\| \leq \Gamma_l(z_0) e^{-\left(l - \frac{l - \lambda_K}{2} \right) t} \leq \Gamma_l(z_0) e^{-\left(\frac{l + \lambda_K}{2} \right) s}$$

$$\leq \left[\Gamma_{\lambda_K}(z_0) + \varepsilon \right] e^{-\lambda_K s}.$$

Hence:

$$\Gamma_{\lambda_K}(z) \leq \Gamma_{\lambda_K}(z_0) + \varepsilon.$$

Moreover, recall that t_0 was defined so that :

$$\Gamma_{\lambda_K}(z_0) = \left\| \left| e^{J_{z_0} s_0} e^{\lambda_K s_0} \right| \right\|.$$

Then, by continuity of $z \mapsto e^{J_z s_0}$, there exists $\delta' > 0$ that ensures that for $|z - z_0| \leq \delta'$, we have:

$$\left\| \left| e^{J_z s_0} e^{\lambda_K s_0} \right| \right\| \geq \left\| \left| e^{J_{z_0} s_0} e^{\lambda_K s_0} \right| \right\| - \varepsilon.$$

Hence:

$$\Gamma_{\lambda_K}(z) \geq \Gamma_{\lambda_K}(z_0) - \varepsilon.$$

In conclusion, if $|z - z_0| \leq \min(\delta, \delta')$, then $|\Gamma_{\lambda_K}(z) - \Gamma_{\lambda_K}(z_0)| \leq \varepsilon$. Hence Γ_{λ_K} is continuous over K. Furthermore, as K is a compact set, Γ_{λ_K} is bounded, by γ. □

F Proof of Proposition 3.1

This appendix is dedicated to the proof of Proposition 3.1.

Proof Let $(g, m, z^*) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+$ be such that P_{z^*} has a single positive root. From Lemma 1, this root defines a fast equilibrium if it is greater than $f_1(z^*)$. From Lemma 2, that is the case if and only if $f_1(z^*)$ is negative or $P_{z^*}(f_1(z^*))$ is negative.

First, regarding the sign of $f_1(z^*)$, we have:

$$f_1(z^*) < 0 \iff (z^* + 1)^2 < \frac{1 - m}{g},$$

which requires that $m < 1$. If $m < 1$ then:

$$f_1(z^*) < 0 \iff 0 \leq z^* < \sqrt{\frac{1 - m}{g}} - 1,$$

which requires that $m + g < 1$. Hence:

$$f_1(z^*) < 0 \iff [m + g < 1] \wedge [z^* < \sqrt{\frac{1 - m}{g}} - 1].$$

Next, regarding the sign of $P_{z^*}(f_1(z^*))$, we compute:

$$P_{z^*}(f_1(z^*)) = f_1(z^*) f_2(z^*) - 1$$

$$= \left(1 + \frac{g}{m}(z^* + 1)^2 - \frac{1}{m}\right)\left(1 + \frac{g}{m}(z^* - 1)^2 - \frac{1}{m}\right) - 1$$

$$= \frac{g^2}{m^2}\left[z^{*4} + z^{*2}\frac{2(m - g - 1)}{g} + \frac{(g - 1)(2m + g - 1)}{g^2}\right]$$

Let us define:

$$Q(X) = X^2 + X\frac{2(m - g - 1)}{g} + \frac{(g - 1)(2m + g - 1)}{g^2},$$

z_1, z_2 its two roots and $\Delta = \frac{4}{g^2}\left[m^2 - 4g(m - 1)\right]$ its discriminant. From the computation above,

$$P_{z^*}(f_1(z^*)) < 0 \iff [\Delta > 0] \wedge \left[z^{*2} \in]z_1, z_2[\right].$$

We have:

$$\Delta > 0 \iff m^2 - 4gm + 4g > 0$$

$$\iff [g < 1] \vee \left[[g \geq 1] \wedge \left[\left[0 < m < 2g\left(1 - \sqrt{1 - \frac{1}{g}}\right)\right]\right.\right.$$

$$\left.\left.\vee \left[m > 2g\left(1 + \sqrt{1 - \frac{1}{g}}\right)\right]\right]\right]$$

and:

$$z_1 z_2 = \frac{(g - 1)(2m + g - 1)}{g^2}, \quad z_1 + z_2 = \frac{2(g + 1 - m)}{g}.$$

Consequently:

◇ if $g \geq 1$, then $2m + g - 1 > 0$ and then $z_1 z_2 \geq 0$. If, additionally, $m < 2g\left(1 - \sqrt{1 - \frac{1}{g}}\right)$, then $m < 2 \leq g + 1$ ($g \mapsto 2g - 2\sqrt{g^2 - g}$ is decreasing on $[1, +\infty[)$. Therefore, we get: $z_1 + z_2 > 0$ and thus, $z_2 > 0$ and $z_1 \geq 0$. At last, if $m > 2g\left(1 + \sqrt{1 - \frac{1}{g}}\right)$, then $m > 2g \geq g + 1$, which implies $z_1 + z_2 < 0$ and thus $z_1 < 0$, $z_2 \leq 0$.

◇ if $g < 1$, then $z_1 + z_2 \geq 0$ if and only if $m \leq g + 1$ and $z_1 z_2 \geq 0$ if and only if $m \leq \frac{1-g}{2}$ (which is lower than $g + 1$).

Hence the result. □

G Proof of Lemma 9

This section is dedicated to proving Lemma 9, which concludes the proof of Proposition 4.2.

Proof (Proof of Lemma 9)

Let $(m, g) \in \mathbb{R}_+^{*\,2}$ verify (25). Then, from the first part of the proof of Proposition 4.2, there exists a unique $\rho^* > 0$ that is solution of the equation in (23). Let us define N_1^* and N_2^* such as in (26). Then we have: $0 < \rho^* = \frac{N_2^*}{N_1^*}$. Thus:

$$N_1^* > 0 \iff N_2^* > 0 \iff \frac{1}{m}(N_1^* + N_2^*) > 0.$$

Borrowing once again the notations: $a = \frac{4g}{m}$, $b = \frac{1}{m}$ and $y^* = \rho^* + \frac{1}{\rho^*}$ (unique root of S larger than 2), (26) leads to:

$$
\begin{aligned}
\frac{1}{m}(N_1^* + N_2^*) &= 2\left(\frac{1}{m} - 1\right) + y^* - \frac{4g}{m}\frac{y^{*2} - 2}{y^{*2}} \\
&= \frac{1}{y^{*2}}\left[y^{*3} + \left[\frac{1 - 2m}{m} + \frac{1}{m} - \frac{4g}{m}\right]y^{*2} + 2 \times \frac{4g}{m}\right] \\
&= \frac{1}{y^{*2}}\left[S(y^*) + (\frac{1 - 2m}{m})y^{*2} + \frac{4g}{m}y^* + \frac{4g}{m}\right].
\end{aligned}
$$

As $S(y^*) = 0$, we get:

$$N_1^* > 0 \iff N_2^* > 0 \iff (1 - 2m)y^{*2} + 4gy^* + 4g > 0.$$

This is always true whenever $m \leq \frac{1}{2}$. Otherwise, let us suppose henceforth that $2m > 1$. The condition above is equivalent to:

$$y^* < c + \sqrt{c^2 + 2c}, \quad \text{where: } c = \frac{2g}{2m - 1} > 0.$$

Let us show that: $c + \sqrt{c^2 + 2c} \geq 2$. It is sufficient to show that: $c \geq \frac{2}{3}$, which is equivalent to having: $3g + 1 \geq 2m$. In this proof, we are considering $(m, g) \in \mathbb{R}_+^{*\,2}$ such that $1 + 2m < 5g$ and $4g(m - 1) < m^2$. Let us show that such pairs verify $3g + 1 \geq 2m$:

◇ if $g \leq 1$, then $m < \frac{5g-1}{2} \leq \frac{3g+1}{2}$.
◇ if $g \geq 1$, then $m < 2g - 2\sqrt{g^2 - g}$ which is a decreasing function on $[1, +\infty[$, which takes the value 2 when $g = 1$. Hence it is always dominated by $g \mapsto \frac{3g+1}{2}$ on this interval.

Hence $c + \sqrt{c^2 + 2c} \geq \frac{2}{3} + \sqrt{\frac{4}{9} + \frac{4}{3}} = 2$. Therefore, as y^* is the only root of S greater than 2, we get the following equivalence:

$$y^* < c + \sqrt{c^2 + c} \iff S\left(c + \sqrt{c^2 + 2c}\right) > 0.$$

The rest of the proof is dedicated to examine the conditions on (m, g) under which:

$$S\left(c + \sqrt{c^2 + 2c}\right) > 0.$$

Let us set $Q := \sqrt{c^2 + 2c} = \sqrt{4g \frac{g+2m-1}{(2m-1)^2}}$. Tedious computations done with the help of Mathematica show that: $S(c) = Q^2 \left[\frac{g(4-6m)+(2m-1)^2}{m\,(2m-1)} \right]$, and we next compute:

$$
\begin{aligned}
S(c + Q) &= S(c) + Q^2 \left[3c + \frac{1-4g}{m} \right] + Q \left[Q^2 + 3c^2 + 2c \frac{(1-4g)}{m} - \frac{4g}{m} \right] \\
&= Q^2 \left[\frac{g(4-6m)+(2m-1)^2}{m\,(2m-1)} + \frac{6g}{2m-1} + \frac{1-4g}{m} \right] \\
&\quad + Q \left[4c^2 + 2c \frac{(m+1-4g)}{m} - \frac{4g}{m} \right] \\
&= Q \left[2Q \frac{\left(2m^2 - m - 4g\,(m-1)\right)}{m(2m-1)} - \frac{4g\left(4g\,(m-1)+2m^2-5m+2\right)}{m(2m-1)^2} \right].
\end{aligned}
$$

Hence:

$$
\begin{aligned}
&S(c + Q) > 0 \\
&\iff Q\left(2m^2 - m - 4g\,(m-1)\right) > 2g \frac{\left(4g\,(m-1)+2m^2-5m+2\right)}{(2m-1)} \\
&\iff \sqrt{g+2m-1}\left(2m^2 - m - 4g\,(m-1)\right) \\
&\quad > 2\sqrt{g}\left(4g\,(m-1)+2m^2-5m+2\right).
\end{aligned}
$$

Let us study different cases corresponding to different ranges of value of $m > \frac{1}{2}$.

If $m = 1$, then the last line is equivalent to :

$$\sqrt{1+g} > -2\sqrt{g},$$

which is true for all $g > 0$.

If $\frac{1}{2} < m < 1$, then:

$$4g\,(m-1) + 2m^2 - 5m + 2 = 4g\,(m-1) + 2(m-2)\left(m - \frac{1}{2}\right) < 0,$$

and:

$$2m^2 - m - 4g(m-1) = 2m\left(m - \frac{1}{2}\right) + 4g(1-m) > 0.$$

Hence, for all g such that $1 + 2m < 5g$ and $m^2 > 4g(m-1)$:

$$\sqrt{g + 2m - 1}\left(2m^2 - m - 4g(m-1)\right) > 2\sqrt{g}\left(4g(m-1) + 2m^2 - 5m + 2\right).$$

If $\underline{m > 1}$, then:

$$2m^2 - m > m^2 > 4g(m-1).$$

Hence, if: $4g(m-1) + 2m^2 - 5m + 2 < 0$, then, for all g such that $1 + 2m < 5g$ and $m^2 > 4g(m-1)$:

$$\sqrt{g + 2m - 1}\left(2m^2 - m - 4g(m-1)\right) > 2\sqrt{g}\left(4g(m-1) + 2m^2 - 5m + 2\right).$$

Otherwise, if $4g(m-1) + 2m^2 - 5m + 2 \geq 0$, then:

$$
\begin{aligned}
&S(c+Q) > 0 \\
&\iff \sqrt{g + 2m - 1}\left(2m^2 - m - 4g(m-1)\right) > 2\sqrt{g}\left(4g(m-1) + 2m^2 - 5m + 2\right) \\
&\iff \left(1 + \frac{2m-1}{g}\right)\left(2m^2 - 2 - 4g(m-1)\right)^2 > 4\left(4g(m-1) + 2m^2 - 5m + 2\right)^2.
\end{aligned}
$$

Let us note $x := \frac{2m-1}{g}$. Then, the latter is equivalent to:

$$
\begin{aligned}
&(1+x)\left[(m-1)x + (x - 4(m-1))\right]^2 - \left[(m-1)x - (x - 4(m-1))\right]^2 > 0 \\
&\iff 4(m-1)x(x - 4(m-1)) + x(mx - 4(m-1))^2 > 0 \\
&\iff 4(m-1)x - 16(m-1)^2 + m^2x^2 - 8mx(m-1) + 16(m-1)^2 > 0 \\
&\iff m^2x^2 + 4x(m-1)(1-2m) > 0 \\
&\iff m^2x^2 - 4x^2g(m-1) > 0 \\
&\iff m^2 > 4g(m-1).
\end{aligned}
$$

$\square$

H Details of the numerical analysis carried out in Sect. 2

Domains We consider a bounded trait domain $[-z_{\mathbf{max}}, z_{\mathbf{max}}]$, discretised in a mesh $(z_k)_{0 \leq k < K}$ (K odd) with regard to the step length $\delta z > 0$, and a time domain $[0, T_{\mathbf{max}}]$,

discretised in a mesh $\left(t^l\right)_{0 \leq l < L}$ with regard to the step length $\delta t > 0$. In the the simulations involved in Fig. 2, we use the following values for the parameters:

$$z_{\mathbf{max}} = 7, \quad T_{\mathbf{max}} = 1000, \quad \delta z = 1.6 \times 10^{-2}, \quad \delta t = 5 \times 10^{-3}.$$

Scheme For $i \in \{1, 2\}, 0 \leq l < L$, we approximate the trait distributions $n_i(t^l, \cdot)$ by $\left(n_{i,k}^l\right)_{0 \leq k < K}$ with the following semi-implicit scheme:

$$\frac{\sigma^2}{\delta t} \left(n_{i,k}^{l+1} - n_{i,k}^l\right) = r \, B_{i,k}^l - \left(g \, (z_k - \theta_i)^2 + \kappa \, N_i^l + m\right) n_{i,k}^{l+1} + m \, n_{j,k}^{l+1},$$

where $N_i^l = \sum_{k=0}^{K-2} n_{i,k}^l \, \delta z$ and $B_{i,k}^l$ is a discretisation of the reproduction operator $\mathcal{B}_\sigma \left(n_i(t^l, z_k)\right)$. In the next paragraph, we detail how we compute $\left(B_{i,k}^l\right)_{0 \leq k < K}$.

We approximate the system of moments of Ronce and Kirkpatrick (2001) following a similar semi-implicit scheme.

Discretization of the reproduction operator The discretization of the reproduction operator is in accordance with the double convolution form shown in Lemma 10, as it increases greatly the computational speed in comparison to a double loop. However, the half-arguments involved in Lemma 10 calls for a special attention to the meshes involved.

Let us define two auxiliary trait meshes

1. $\left(\tilde{z}_{k'}\right)_{0 \leq k' < 2K-1}$ on $[-z_{\mathbf{max}}, z_{\mathbf{max}}]$, with step length $\frac{\delta z}{2}$,
2. $\left(\hat{z}_{k''}\right)_{0 \leq k'' < 4K-3}$ on $[-2z_{\mathbf{max}}, 2z_{\mathbf{max}}]$, with step length $\frac{\delta z}{2}$.

We define the vector $(G_{k'})_{0 \leq k' < 2K-1}$ discretising the Gaussian kernel involved in our reproduction operator on the trait grid $\left(\tilde{z}_{k'}\right)_{0 \leq k' < 2K-1}$:

$$G_{k'} = \frac{1}{\sqrt{\pi}\sigma} \exp\left[-\frac{\tilde{z}_{k'}^2}{\sigma^2}\right].$$

We next define the vector $\left(\hat{B}_{i,k''}^l\right)_{0 \leq k'' < 4K-3}$ resulting from the following double discrete convolution (denoted $*$):

$$\left(\hat{B}_{i,k''}^l\right)_{0 \leq k'' < 4K-3} = \frac{1}{N_i^l} \left(n_{i,k}^l\right)_{0 \leq k < K} * \left(n_{i,k}^l\right)_{0 \leq k < K} * (G_{k'})_{0 \leq k' < 2K-1}$$

We use a convolution algorithm with default settings: the size of the output is the sum of entry vector sizes minus one, and out of bounds index entries are extrapolated as 0. A straight-forward computation shows that $\left(\hat{B}_{i,k''}^l\right)_{0 \leq k'' < 4K-3}$ is the approximation

of the reproduction operator on the mesh $\left(\hat{z}_{k''}\right)_{0 \le k'' < 4K-3}$:

$$
\begin{aligned}
\hat{B}^l_{i,k''} &= \frac{\delta z^2}{N^l_i} \sum_{k_1=0}^{4K-4} n^l_{i,k_1} \sum_{k_2=0}^{3K-2} n^l_{i,k_2}\, G_{k''-k_1-k_2} \\
&= \frac{\delta z^2}{N^l_i} \sum_{k_1=0}^{4K-4} n^l_{i,k_1} \sum_{k_2=0}^{3K-2} n^l_{i,k_2}\, \frac{1}{\sqrt{\pi}\sigma} \exp\left[-\frac{\tilde{z}^2_{k''-k_1-k_2}}{\sigma^2} \right] \\
&= \frac{\delta z^2}{N^l_i} \sum_{k_1=0}^{4K-4} n^l_{i,k_1} \sum_{k_2=0}^{3K-2} n^l_{i,k_2}\, \frac{1}{\sqrt{\pi}\sigma} \exp\left[-\frac{\left(-z_{\max} + \frac{\delta z}{2}(k''-k_1-k_2)\right)^2}{\sigma^2} \right] \\
&= \frac{\delta z^2}{N^l_i} \sum_{k_1=0}^{4K-4} n^l_{i,k_1} \sum_{k_2=0}^{3K-2} n^l_{i,k_2}\, \frac{1}{\sqrt{\pi}\sigma} \\
&\qquad \exp\left[-\frac{\left(-2z_{\max} + \frac{k''\delta z}{2} - \frac{(-z_{\max}+k_1\delta z)+(-z_{\max}+k_2\delta z)}{2}\right)^2}{\sigma^2} \right] \\
&= \frac{\delta z^2}{N^l_i} \sum_{k_1=0}^{4K-4} n^l_{i,k_1} \sum_{k_2=0}^{3K-2} n^l_{i,k_2}\, \frac{1}{\sqrt{\pi}\sigma} \exp\left[-\frac{\left(\hat{z}_{k''} - \frac{z_{k_1}+z_{k_2}}{2}\right)^2}{\sigma^2} \right].
\end{aligned}
$$

Thus, we interpolate $\left(\hat{B}^l_{i,k''}\right)_{0 \le k'' < 4K-3}$ at the entries corresponding to $(z_k)_{0 \le k < K}$ to obtain $\left(B^l_{i,k}\right)_{0 \le k < K}$.

I Numerical outcomes details: Figs. 6 and 7

Numerical setting The lower panel of Fig. 6 has been produced by running 3600 simulations, one for each couple of migration rate $m \in [0.01, 3]$ and intensity of selection $g \in [0.01, 3]$, for $t \le T_{max} \in \left[\frac{300}{\varepsilon^2}, \frac{600}{\varepsilon^2}\right]$, with a criteria to cut the simulation short at a time greater than $\frac{300}{\varepsilon^2}$ if the difference between two consecutive steps is small enough. The value of the other parameters are the same for each simulation: $r = 1, \theta = 1, \kappa = 1, \varepsilon = 0.05$, as well as the initial state:

$$
\begin{cases}
n^0_1(z) = 0.99 \times \dfrac{e^{-\frac{(z-0.2)^2}{2\varepsilon^2}}}{\sqrt{2\pi}\varepsilon}, \\[3mm]
n^0_2(z) = \dfrac{e^{-\frac{(z-0.2)^2}{2\varepsilon^2}}}{\sqrt{2\pi}\varepsilon}.
\end{cases}
$$

The initial state is taken as monomorphic, as the aim of this figure is to be compared to the theoretical outcomes that are predicted within the scope of the slow-fast analysis

as stated in Theorem 3.1 (so when the initial state is close enough from the slow manifold).

Scoring Each simulation final state (n_1^f, n_2^f) is attributed a score between 0 and 1 according to the following scheme:

1. If $\max\left(N_1^f, N_2^f\right) < 0.01$, then the score is 0 (for extinction) and is corresponding to the deep purple color. Else, the score is a positive number (lower than 1) according to what follows.
2. If the variance in trait of the metapopulation is greater than $2\varepsilon^2$, the score is 1 (corresponding to the color yellow). This would be the case if the final state is dimorphic, but more generally, this is to highlight the simulations whose final state does not fall in the small segregational variance regime analysis prediction (which in particular predicts that the distribution of trait in the metapopulation is monomorphic (see Sect. 3), with a variance of order ε^2 (see (12)).
3. If both conditions above are not met, then the score S is given according to the following formula:

$$S = \frac{5}{6} - \frac{1}{3} \frac{\left|N_2^f - N_1^f\right|}{N_1^f + N_2^f}.$$

This formula discriminates between symmetrical equilibria (which are characterized by equal population sizes, see Proposition 4.1), which typically have a score of $\frac{5}{6}$ (corresponding to the color light green), and asymmetrical equilibria, which have a discrepancy in local population sizes and therefore have a typically much lower score (in the blue tones).

Adjustments for Fig. 7 The methodology is the same for the lower panel of Fig. 6 and both panels of Fig. 7, at the exception of the initial state, set as:

$$\begin{cases} n_1^0(z) = 0.9 \times \dfrac{e^{-\frac{(z+1)^2}{2\varepsilon^2}}}{\sqrt{2\pi}\varepsilon}, \\ n_2^0(z) = \dfrac{e^{-\frac{(z-1)^2}{2\varepsilon^2}}}{\sqrt{2\pi}\varepsilon}. \end{cases}$$

and of the time step for the lower panel of Fig. 7, which is refined to keep up with the smaller value of ε^2.

References

Akerman A, Bürger R (2014) The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. J Math Biol 68(5):1135–1198. https://doi.org/10.1007/s00285-013-0660-z

Barton NH, Etheridge AM, Véber A (2017) The infinitesimal model: definition, derivation, and implications. Theor Popul Biol 118:50–73. https://doi.org/10.1016/j.tpb.2017.06.001

Bouin E et al (2018) Equilibria of quantitative genetics models beyond the Gaussian approximation I: maladaptation to a changing environment. (In preparation)

Bourgeron T et al (2017) Existence of recombination–selection equilibria for sexual populations. arXiv:1703.09078 [math, q-bio]

Bulmer MG (1971) The effect of selection on genetic variability. Am Nat 105(943):201–211

Bulmer MG (1980) The mathematical theory of quantitative genetics. Oxford University Press

Bürger R, Akerman A (2011) The effects of linkage and gene flow on local adaptation: a two-locus continent-island model. Theor Popul Biol 80(4):272–288. https://doi.org/10.1016/j.tpb.2011.07.002

Calvez V, Garnier J, Patout F (2019) Asymptotic analysis of a quantitative genetics model with nonlinear integral operator. J École Polytech 6:537–579. https://doi.org/10.5802/jep.100

Chicone C (1999) Ordinary differential equations with applications. Springer, Berlin

Day T (2000) Competition and the effect of spatial resource heterogeneity on evolutionary diversification. Am Nat 155(6):790–803

Débarre F, Ronce O, Gandon S (2013) Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. J Evol Biol 26(6):1185–1202. https://doi.org/10.1111/jeb.12132

Débarre F, Yeaman S, Guillaume F (2015) Evolution of quantitative traits under a migration–selection balance: when does skew matter? Am Nat 186(S1):S37–S47. https://doi.org/10.1086/681717

Desvillettes L et al (2008) On selection dynamics for continuous structured populations. Commun Math Sci 6(3):729–747. https://doi.org/10.4310/CMS.2008.v6.n3.a10

Diekmann O et al (2005) The dynamics of adaptation: an illuminating example and a Hamilton–Jacobi approach. Theor Popul Biol 67(4):257–271. https://doi.org/10.1016/j.tpb.2004.12.003

Fisher RA (1919) The correlation between relatives on the supposition of mendelian inheritance. Trans R Soc Edinb 52(2):399–433. https://doi.org/10.1017/S0080456800012163

Galton F (1877) Typical laws of heredity 1. Nature 15:492–495

Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. Evolution 55(3):459–466. https://doi.org/10.1111/j.0014-3820.2001.tb00780.x

Lange K (1978) Central limit theorems of pedigrees. J Math Biol 6(1):59–66. https://doi.org/10.1007/BF02478517

Lavigne F et al (2019) When sinks become sources: adaptive colonization in asexuals. bioRxiv https://doi.org/10.1101/433235. https://www.biorxiv.org/content/early/2019/05/03/433235

Levin JJ, Levinson N (1954) Singular perturbations of non-linear systems of differential equations and an associated boundary layer equation. J Ration Mech Anal 3:247–270

Lythgoe KA (1997) Consequences of gene flow in spatially structured populations. Genet Res 69(1):49–60. https://doi.org/10.1017/S0016672397002644

Magal P, Webb GF (2000) Mutation, selection, and recombination in a model of phenotype evolution. Discrete Contin Dyn Syst A 6(1):221–236. https://doi.org/10.3934/dcds.2000.6.221

Meszéna G, Czibula I, Geritz S (1997) Adaptive dynamics in a 2-patch environment: a toy model for allopatric and parapatric speciation. J Biol Syst 05(02):265–284. https://doi.org/10.1142/S0218339097000175

Mirrahimi S (2017) A Hamilton–Jacobi approach to characterize the evolutionary equilibria in heterogeneous environments. Math Models Methods Appl Sci 27(13):2425–2460. https://doi.org/10.1142/s0218202517500488

Mirrahimi S, Gandon S (2020) Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. Genetics 214(2):479–491. https://doi.org/10.1534/genetics.119.302868

Mirrahimi S, Raoul G (2013) Dynamics of sexual populations structured by a space variable and a phenotypical trait. Theor Popul Biol 84:87–103. https://doi.org/10.1016/j.tpb.2012.12.003

Nagylaki T, Lou Y (2001) Patterns of multiallelic polymorphism maintained by migration and selection. Theor Popul Biol 59(4):297–313. https://doi.org/10.1006/tpbi.2001.1526

Perthame B, Barles G (2008) Dirac concentrations in Lotka–Volterra parabolic PDEs. Indiana Univ Math J 57(7):3275–3302. https://doi.org/10.1512/iumj.2008.57.3398

Raoul G (2017) Macroscopic limit from a structured population model to the Kirkpatrick–Barton model. arXiv:1706.04094 [math]

Ronce O, Kirkpatrick M (2001) When sources become sinks: migrational meltdown in heterogeneous habitats. Evolution 55(8):1520–1531. https://doi.org/10.1111/j.0014-3820.2001.tb00672.x.37

Tufto J (2000) Quantitative genetic models for the balance between migration and stabilizing selection. Genet Res 76(3):285–293. https://doi.org/10.1017/S0016672300004742

Turelli M (2017) Commentary: Fisher's infinitesimal model: a story for the ages. Theor Popul Biol 118:46–49. https://doi.org/10.1016/j.tpb.2017.09.003

Turelli M, Barton NH (1994) Genetic and statistical analyses of strong selection on PolygenicTraits: what, me normal? Genetics 138:913–941. https://doi.org/10.1093/genetics/138.3.913

Yeaman S, Guillaume F (2009) Predicting adaptation under migration load: the role of genetic skew. Evolution 63(11):2926–2938. https://doi.org/10.1111/j.1558-5646.2009.00773.x