DISCRETE APPROXIMATION OF NON-COMPACT OPERATORS DESCRIBING CONTINUUM-OF-ALLELES MODELS

Oliver Redner

Institut für Mathematik und Informatik, Universität Greifswald, Jahnstr. 15a, 17487 Greifswald, Germany (redner@uni-greifswald.de)

Abstract We consider the eigenvalue equation for the largest eigenvalue of certain kinds of non-compact linear operators given as the sum of a multiplication and a kernel operator. It is shown that, under moderate conditions, such operators can be approximated arbitrarily well by operators of finite rank, which constitutes a discretization procedure. For this purpose, two standard methods of approximation theory, the Nyström and the Galerkin method, are generalized. The operators considered describe models for mutation and selection of an infinitely large population of individuals that are labeled by real numbers, commonly called continuum-of-alleles (COA) models.

Keywords: continuum-of-alleles models; mutation–selection models; discretization; kernel operators; multiplication operators; non-compact operators; Nyström method; Galerkin method

AMS 2000 *Mathematics subject classification:* Primary 47A58; 45C05
Secondary 47B34

1. Introduction

This article is concerned with eigenvalue equations on $\mathrm{L}^1(I)$ of the form

$$r(x)\,p(x) + \int_I \big[u(x,y)\,p(y) - u(y,x)\,p(x)\big]\,\mathrm{d}y = \lambda\,p(x) \qquad \text{for all } x \in I. \tag{1.1}$$

Here, p is a probability density on the set I, which is either taken to be a compact interval $[a,b]$ or the real line $\mathbb{R}$, i.e., $p \in \mathrm{L}^1(I)$ with $p \geq 0$ and $\int_I p(x)\,\mathrm{d}x = 1$. Sufficient conditions for the existence and uniqueness of solutions of (1.1) were given by Bürger, see [**4**, Ch. IV.3], in which case λ is the largest eigenvalue.

If one is interested in a discrete approximation of (1.1), one faces the problem that the operator acting on p is the sum of a multiplication operator and a kernel operator; and the former is never compact (apart from trivial cases). Therefore, a direct application of most standard methods of approximation theory fails because, for these, compactness is a prerequisite. In this article, it will be shown that, under some moderate extra conditions, these methods can nevertheless be applied.

One motivation to study equations of the form (1.1) is their occurrence in population genetics, which is concerned with the (micro)evolution of the genetic composition of populations. For many situations, individuals are adequately described by a continuous scalar variable, representing, for example, a quantitative character under selection. This leads to the definition of so-called continuum-of-alleles (COA) models, in which individuals are identified with this variable, referred to as their type. Usually, selection is then modeled by type-dependent fitness values, whereas mutation is described, for every source type, by a probability distribution for the mutant types. For a recent review of and relevant literature on COA models, see [**4**].

In population genetics, evolution may quite generally be assumed to proceed in continuous time, with overlapping generations, or in discrete generations. For the COA model, in both cases, equilibrium is described by an equation of the form (1.1), compare [**12**]. Here, I is the set of possible types. Assuming the population to be effectively infinite, we represent it by the probability density p.

The notation chosen here best fits the case of continuous time, where $r(x)$ describes the effective reproduction rate of type x (i.e., the difference of its birth and death rate), the so-called Malthusian fitness, and $u(x, y)$ is the mutation rate $u_1(y)$ of type y times the density $m(x, y)$ of mutant types x, conditioned on a mutation to occur for y. With discrete generations, $r(x)$ has the interpretation of the expected number of offspring of an individual of type x, i.e., its Wrightian fitness, and mutation is assumed to occur during reproduction with some probability $\mu(y)$ for type y. The distribution of mutant types is again given by $m(x, y)$, hence $u(x, y) = m(x, y)\,\mu(y)\,r(y)$. In both cases, λ equals the equilibrium mean fitness $\int_I r(x)\,p(x)\,\mathrm{d}x$.

There are several reasons why it is desirable to approximate a COA model by a model with discrete types. One reason is the need for numerical investigations of COA models, since most of them are not tractable analytically. These inevitably require a discrete formulation of the model. Another reason is that recently a simple characterization of the equilibrium of discrete mutation–selection models has been found [**8**] (see also [**7**, **6**, **2**]); this takes the form of a scalar maximum principle in a limit of infinitely many types that densely fill a compact interval. Gaining a better understanding of the relation between models with discrete and continuous types is therefore promising to enable a transfer of some of these results.

This article starts with a summary of Bürger's results on (1.1) in Section 2, since these form the basis for our treatment. We will then consider two methods to approximate compact kernel operators and extend them to our case. One, the Nyström method, is applicable to continuous functions r and u on compact intervals I and involves sampling* of these functions. This is presented in Section 3. The other one, the Galerkin method, is based on projections to finite-dimensional subspaces and works—in principle—for a broad class of compact operators. In our case, however, one has to make relatively strong assumptions, e.g., that the functions r and u are, in some sense, uniformly continuous. Then, it turns out that the local averaging in the projection process can be replaced by sampling again (if an additional condition is satisfied). This is discussed in Section 4. A comparison of both methods in Section 5 and an outlook in Section 6 complete this article.

* The term *sampling* is used in the meaning also used in signal processing: Instead of a continuous function one considers its values at a (properly chosen) finite set of points.

2. General properties

Let us first put the equilibrium condition (1.1) in operator notation. Since we are interested in probability densities, we will consider $\mathrm{L}^1(I)$, or a subspace thereof, as the underlying function space. We define the total mutation rate of type x as

$$u_1(x) = \int_I u(y,x)\,\mathrm{d}y \tag{2.1}$$

and, for notational brevity,

$$w = u_1 - r\,.$$

Then, (1.1) is equivalent to the eigenvalue equation

$$(A+\lambda)p = 0\,, \tag{2.2}$$

where, for elements f of the function space and all $x \in I$,

$$(Tf)(x) = w(x)f(x)\,, \tag{2.3}$$

$$(Uf)(x) = \int_I u(x,y)f(y)\,\mathrm{d}y\,, \tag{2.4}$$

$$A = T - U\,. \tag{2.5}$$

As mentioned above, being a (non-zero) multiplication operator, T cannot be compact (compare [**16**, Thm. 2.1]). Strong results like analogs to the Perron–Frobenius theorem, however, are only available for compact, or at least power compact*, operators, see Schaefer [**20**, Ch. V]. Therefore one considers the following family of kernel operators:

$$(K_\alpha f)(x) = \int_I k_\alpha(x,y)f(y)\,\mathrm{d}y\,,$$

where

$$k_\alpha(x,y) = \frac{u(x,y)}{w(y)+\alpha}\,.$$

These are, under conditions that will be given shortly, power compact or even compact. Their connection to the operator A from (2.5) is stated in the following

Lemma 2.1. [**3**, Prop. 2.1(i)] *Let T, U be operators in a Banach space X, with U being bounded, T densely defined, i.e., $\overline{D(T)} = X$, and $T+\alpha$ invertible. Then f is an eigenvector of $A = T - U$ with eigenvalue $-\alpha$, i.e., $0 \neq f \in D(A) = D(T)$ and*

$$(A+\alpha)f = 0\,,$$

if and only if $g = (T+\alpha)f$ is an eigenvector of $K_\alpha = U(T+\alpha)^{-1}$ with eigenvalue 1,

$$(K_\alpha - 1)g = 0\,.$$

* An operator is said to be *power compact* if one of its powers is compact.

So, explicitly in our case, the eigenvalue equation (2.2) is equivalent to

$$(K_\lambda - 1)q = 0 \tag{2.6}$$

with $q = (T + \lambda)p$. This equation can now be used to find sufficient conditions for the existence and uniqueness of a solution of (2.2).

An important class of bounded kernel operators from $\mathrm{L}^q(I)$ into $\mathrm{L}^p(I)$ $(1 \le p,\, q \le \infty)$ are the Hille–Tamarkin operators, see [**11**, Sec. 11.3]. Their kernels need to satisfy

$$|\!|\!|K|\!|\!|_{pq} := \|k_1\|_p < \infty \qquad \text{with} \qquad k_1(x) = \|k(x,.)\|_{q'}\,, \tag{2.7}$$

where $(Kf)(x) = \int_I k(x,y)\, f(y)\,\mathrm{d}y$, $k(x,.)$ denotes the function $y \mapsto k(x,y)$, and q' is the conjugate exponent to q satisfying $\frac{1}{q} + \frac{1}{q'} = 1$, $1 \le q' \le \infty$. The Hille–Tamarkin norm $|\!|\!|.|\!|\!|_{pq}$ turns the set $\mathcal{H}_{pq}(I)$ of all Hille–Tamarkin operators into a Banach space [**11**, Thm. 11.5]. Here, we are interested in $p = q = 1$, in which case (2.7) yields

$$|\!|\!|K|\!|\!|_{11} = \int_I \operatorname*{ess\,sup}_{y\in I} |k(x,y)|\,\mathrm{d}x < \infty$$

and K^2 is compact for every $K \in \mathcal{H}_{11}(I)$ [**11**, Thm. 11.9].

Let us now turn to kernel operators that are power compact, positive, and irreducible. An operator is called *positive* if it maps the set of non-negative functions into itself, for which, in the case of kernel operators, non-negativity of the kernel is necessary and sufficient [**11**, p. 122]. A kernel operator is *irreducible* if its kernel satisfies [**20**, Exm. 4 in Sec. V.6]

$$\int_{I\setminus J}\int_J k(x,y)\,\mathrm{d}x\,\mathrm{d}y > 0 \qquad \text{for all measurable } J \subset I \text{ with } |J|,\ |I\setminus J| > 0.$$

Here, $|J|$ denotes the Lebesgue measure of a measurable set J. Then, the theorem of Jentzsch [**20**, Thm. V.6.6], which parallels the Perron–Frobenius theorem for matrices, states that the spectral radius is an algebraically simple eigenvalue with an (up to normalization) unique positive eigenfunction (i.e., strictly positive a.e.*) and the only eigenvalue with a positive eigenfunction.

In our case, the following requirements are sufficient for the K_α to be Hille–Tamarkin operators [**3**, Sec. 3].

(U1) u is non-negative and measurable.

(U2) $u_1(x)$ from (2.1) exists for a.e. $x \in \mathbb{R}$ and $u_1 \in \mathrm{L}^\infty(I)$, i.e., u_1 is essentially bounded. (By Hölder's inequality, this implies that U is bounded, cf. [**3**, Prop. 3.1(ii)].)

(T1) $w = u_1 - r$ is measurable and satisfies $\operatorname{ess\,inf}_{x\in I} w(x) = 0$. (The latter can be achieved, without loss of generality, by adding a suitable constant to r.)

* The abbreviation 'a.e.' stands for 'almost every' or 'almost everywhere' and means that the set at which the condition it refers to is not fulfilled has zero (Lebesgue) measure.

(T2) $(w+1)^{-1} \in \mathrm{L}^\infty(I)$ is then already a consequence of (T1).

(U4) $\int_I \operatorname{ess\,sup}_{y\in I} u(x,y)/(w(y)+\alpha)\,\mathrm{d}x < \infty$ for one (and then for all) $\alpha > 0$.

For $\alpha > 0$, K_α is irreducible if U is [**3**, proof of Thm. 2.2(c)], i.e.,

$$\int_{I\setminus J}\int_J u(x,y)\,\mathrm{d}x\,\mathrm{d}y > 0 \qquad \text{for all measurable } J \subset I \text{ with } |J|,\ |I\setminus J| > 0. \tag{2.8}$$

To keep the equilibrium distribution from having atoms, we assume that there is a set $J \subset I$ with positive measure for which $\operatorname{ess\,inf}_{x\in J} w(x) = 0$ such that

$$\operatorname*{ess\,inf}_{x,y\in J} u(x,y) \int_J (w(x))^{-1}\,\mathrm{d}x > 1 \tag{2.9}$$

or the integral diverges [**4**, cond. 3” in Sec. IV.3].

Putting everything together, we have the following

Theorem 2.2 (Bürger). *Under the above conditions,* (1.1) *has a unique positive solution* $p \in \mathrm{L}^1(I)$ *with* $\|p\|_1 = 1$, *for which* $\lambda > 0$ *is the largest spectral value of* $-A$ *from* (2.5).

Proof. See the above, [**3**, Thm. 3.5], and [**4**, Sec. IV.3]. □

Note that, due to (T1), p is positive if and only if $q = (w+\alpha)p$ is, for $\alpha > 0$.

Another result that will be needed in the sequel is

Lemma 2.3. [**3**, Lemmas 1–3 and Thm. 2.2(ii)] *Under the above conditions, the spectral radius* $\rho(K_\alpha)$ *is, as a function of* α, *strictly decreasing and satisfies* $\rho(K_\lambda) = 1$ *as well as* $\lim_{\alpha\to\infty} \rho(K_\alpha) = 0$. *Thus,* $\rho(K_\alpha) < 1$ *implies* $\alpha > \lambda$ *and* $\rho(K_\alpha) > 1$ *implies* $\alpha < \lambda$.

Throughout the rest of this article, all the above criteria are assumed to be satisfied, namely (U1), (U2), (U4), (T1), (T2), (2.8), and (2.9).

3. Discretization—compact interval

Let the interval I be compact and $\mathrm{C}(I)$ denote the Banach space of bounded, continuous functions equipped with the supremum norm $\|f\|_\infty = \sup_{x\in I} |f(x)|$. We consider operators K of the form

$$(Kf)(x) = \int_I k(x,y) f(y)\,\mathrm{d}y \qquad \text{for all } x \in I \tag{3.1}$$

with a continuous kernel $k\colon I \times I \to \mathbb{R}$. First note these two basic results:

Proposition 3.1. *Any* K *of the form* (3.1) *maps* $\mathrm{L}^1(I)$ *into* $\mathrm{C}(I) \subset \mathrm{L}^1(I)$.

Proof. We follow the proof of [**5**, Thm. 2.1], where this is shown for $\mathrm{L}^2(I)$, which, since I is compact, is a subspace of $\mathrm{L}^1(I)$. Let $f \in \mathrm{L}^1(I)$ and $x, \xi \in I$ be given. Then,

$$|(Kf)(x) - (Kf)(\xi)| \leq \int_I |k(x,y) - k(\xi,y)|\,|f(y)|\,\mathrm{d}y \leq \sup_{y \in I} |k(x,y) - k(\xi,y)|\,\|f\|_1 \,.$$

Due to the uniform continuity of k in $I \times I$, we have

$$\lim_{\xi \to x} \sup_{y \in I} |k(x,y) - k(\xi,y)| = 0\,,$$

from which the continuity of Kf follows. □

Proposition 3.2. *An operator K of the form* (3.1) *is compact from* $\mathrm{C}(I)$ *or* $\mathrm{L}^1(I)$ *to either of the two spaces.*

Proof. Follow the proof of [**5**, Thm. 2.10] (or [**15**, XVII.4]), where this is shown for $\mathrm{L}^2(I) \subset \mathrm{L}^1(I)$, and use Hölder's inequality whenever the Cauchy–Schwarz inequality is used. Alternatively, see [**20**, Exm. 3 in Sec. IV.10]. □

Thus, if in our case the functions r and u are continuous, also the kernel k_α is, for every $\alpha > 0$. It then follows from Proposition 3.1 that the equilibrium density p is continuous as well. Therefore we can restrict our attention to $\mathrm{C}(I)$ in our quest for a solution of the eigenvalue equation (2.2). This makes the Nyström method applicable as a discretization procedure, which will be presented now.

3.1. The Nyström method

The Nyström method is based on quadratures, which are used for numerical integration, cf., e.g., Kress [**14**, Ch. 12]. We will use this (slightly restricted)

Definition 3.3. A *quadrature rule* Q_n is a mapping of the form

$$Q_n \colon \mathrm{C}(I) \to \mathbb{R}\,, \qquad f \mapsto Q_n f = \sum_{k=1}^{N_n} \alpha_{n,k} f(t_{n,k})\,,$$

with $n \in \mathbb{N}$, $N_n \in \mathbb{N}$, *quadrature points* $t_{n,k} \in I$, and *quadrature weights* $\alpha_{n,k} > 0$, for $k \in \mathcal{N}_n := \{1, \ldots, N_n\}$. A sequence of quadrature rules, or simply a *quadrature*, (Q_n) is said to be *convergent* if

$$Q_n f \to Qf \qquad \text{for all } f \in \mathrm{C}(I), \tag{3.2}$$

where $Q \colon \mathrm{C}(I) \to \mathbb{R}$ is the linear functional that assigns to each f its integral, i.e., $Qf = \int_I f(x)\,\mathrm{d}x$.

Another notion that is important for the Nyström method is the collectively compact convergence of operators. The standard reference for this matter is [**1**].

Definition 3.4. A sequence (K_n) of (compact) operators in a Banach space X is *collectively compact* if the set $\{K_n B : n \in \mathbb{N}\}$ is relatively compact (i.e., its closure is compact) for every bounded set $B \subset X$. If furthermore the sequence converges pointwise to an operator K one speaks of *collectively compact convergence*, in symbols $K_n \xrightarrow{\mathrm{cc}} K$.

As a direct consequence of this definition, K is compact (as well as all K_n). The central result for the Nyström method is

Theorem 3.5. *Let K be a compact kernel operator of the form* (3.1) *whose eigenvalue equation*

$$(K-\nu)g=0 \tag{3.3}$$

is to be approximated. To this end, let $(Q_n)_{n\in\mathbb{N}}$ be a convergent quadrature with the notation as in Definition 3.3. *A* complete discretization *is given by the $N_n\times N_n$ matrices $\boldsymbol{K}_n$ with entries*

$$K_{n,k\ell}=\alpha_{n,\ell}\,k(t_{n,k},t_{n,\ell})\,,$$

a partial discretization *by means of the operators K_n on* $\mathrm{C}(I)$ *with*

$$(K_nf)(x)=\sum_{k=1}^{N_n}\alpha_{n,k}\,k(x,t_{n,k})f(t_{n,k})=Q_n(k(x,.)f)\,. \tag{3.4}$$

Consider the corresponding eigenvalue equations

$$(\boldsymbol{K}_n-\nu_n)\boldsymbol{g}_n=0 \qquad \text{and} \qquad (K_n-\nu_n)g_n=0\,, \tag{3.5}$$

where $\boldsymbol{g}_n$ is an N_n-dimensional vector with components $g_{n,k}$, and $g_n\in\mathrm{C}(I)$. Then, under the above conditions the following statements are true:

(a) *Both eigenvalue equations in* (3.5) *are equivalent and connected via*

$$g_n(x)=\sum_{k=1}^{N_n}\alpha_{n,k}\,k(x,t_{n,k})g_{n,k}\,. \tag{3.6}$$

(b) *For every $\nu\neq 0$ from* (3.3) *there is a sequence (ν_n) of eigenvalues of* (3.5) *such that $\nu_n\to\nu$ as $n\to\infty$. Conversely, every non-zero limit point of any sequence (ν_n) of eigenvalues of* (3.5) *is an eigenvalue of* (3.3).

(c) *Every bounded sequence (g_n) of eigenfunctions of* (3.5) *associated with eigenvalues $\nu_n\to\nu\neq 0$ contains a convergent subsequence; the limit of any convergent subsequence $(g_{n_i})_i$ is an eigenfunction of* (3.3) *associated with the eigenvalue ν (unless the limit is zero).*

Proof. (a) is the statement of [**14**, Thm. 12.7] or [**5**, Lemma 3.15]. (b) and (c) rely on $K_n\xrightarrow{\mathrm{cc}}K$, which is shown, e.g., in [**1**, Props. 2.1, 2.2], [**14**, Thm. 12.8], or [**5**, Thm. 3.22]. The statements then follow from [**1**, Thms. 4.11, 4.17]. $\square$

We will restrict ourselves to quadratures that allow for disjoint partitions of I with intervals $I_{n,k}$, i.e., $I_{n,k}\cap I_{n,\ell}\neq\emptyset$ and $\bigcup_{k=0}^{N_n}I_{n,k}=I$, such that $t_{n,k}\in I_{n,k}$ and $|I_{n,k}|=\alpha_{n,k}$ (with $k\in\mathcal{N}_n$). For such quadratures it is easy to see that*

$$\|Q_n\|=\sum_{k=1}^{N_n}\alpha_{n,k}=|I| \tag{3.7}$$

* If not noted otherwise, the following convention for operator norms is used. If an operator maps a space X into itself, we denote its norm by the same symbol as the norm of X, e.g., $\|.\|_X$, or $\|.\|_1$ for L^1; in all other cases the unornamented symbol $\|.\|$ is used.

and that the partitions are unique (up to the boundary points of the intervals). Furthermore we have

Lemma 3.6. *Let (Q_n) be a convergent quadrature that allows for partitions of I as described above. Then* $\lim_{n\to\infty} \max_{k\in\mathcal{N}_n} |I_{n,k}| = 0$.

Proof. Assume the contrary. Then there are an $\varepsilon > 0$ and sequences $(n_i)_i$ and $(k_i)_i$ with $\lim_{i\to\infty} n_i = \infty$ such that $|I_{n_i,k_i}| \geq \varepsilon$. Due to the compactness of I, these can be chosen in a way that $\lim_{i\to\infty} t_{n_i,k_i} =: t$ exists. Now consider the continuous function $f(x) = \max\{1 - 2|x-t|/\varepsilon, 0\}$. For this we have $Qf \leq \varepsilon/2$, but $\lim_{i\to\infty} Q_{n_i} f \geq \varepsilon \lim_{i\to\infty} f(t_{n_i,k_i}) = \varepsilon$, which contradicts the convergence of the quadrature (3.2). □

3.2. Application to the COA model

In our case of the COA model with a compact interval I and continuous functions r and u, the complete discretization is given by the following $N_n \times N_n$ matrices:

$$T_{n,k\ell} = \delta_{k\ell}\, w(t_{n,k}) \geq 0\,, \tag{3.8}$$

$$U_{n,k\ell} = \alpha_{n,\ell}\, u(t_{n,k}, t_{n,\ell}) \geq 0\,, \tag{3.9}$$

$$\boldsymbol{A}_n = \boldsymbol{T}_n - \boldsymbol{U}_n, \qquad \boldsymbol{K}_{\alpha,n} = \boldsymbol{U}_n(\boldsymbol{T}_n + \alpha)^{-1} \quad \text{for } \alpha > -\min_{k\in\mathcal{N}_n} w(t_{n,k}).$$

The eigenvalue equations to be solved are

$$(\boldsymbol{A}_n + \lambda_n)\boldsymbol{p}_n = 0 \qquad \text{with } \boldsymbol{p}_n > 0.$$

Here, $-\boldsymbol{A}_n + c$ is positive with a suitable constant c. We further have to assume that the $\boldsymbol{A}_n$ are irreducible (which might not be the case for special choices of the $t_{n,k}$, e.g., if $u_1(t_{n,k}) = 0$ for some k). Then, due to the Perron–Frobenius theorem, there exist (up to normalization) unique positive $\boldsymbol{p}_n$ belonging to the eigenvalues $-\lambda_n = -\rho(-\boldsymbol{A}_n + c) + c$, where $\rho(\boldsymbol{M})$ denotes the spectral radius of a matrix $\boldsymbol{M}$. With $\boldsymbol{q}_n = (\boldsymbol{T}_n + \lambda_n)\boldsymbol{p}_n$ also the eigenvalue equations

$$(\boldsymbol{K}_{\lambda_n,n} - 1)\boldsymbol{q}_n = 0 \tag{3.10}$$

are solved (and vice versa), cf. Lemma 2.1.

Both $\boldsymbol{K}_{\lambda_n,n}$ and $\boldsymbol{q}_n$ can be embedded into $\mathrm{C}(I)$ as described by (3.4) and (3.6). Then, with Theorem 3.5, one might conclude the convergence $\|q_n - q\|_\infty \to 0$. In the end, however, we are interested in the population vectors $\boldsymbol{p}_n$ and their convergence to the density p. It might be easiest to interpret the vectors $\boldsymbol{p}_n$ as point measures on I. But then the best one can hope for is weak convergence since the set of point measures is closed under the total variation norm. It will turn out that we can indeed achieve norm convergence if we embed the $\boldsymbol{p}_n$ into $\mathrm{L}^1(I)$ the following way. We choose a disjoint partition of I as above and let

$$p_n = \sum_{k=1}^{N_n} p_{n,k} 1_{I_{n,k}}\,,$$

where 1_J denotes the characteristic function of a set J. (Note that $p_{n,k}$ denotes the k-th component of $\boldsymbol{p}_n \in \mathbb{R}^{N_n}$, whereas p_n is an L^1 function.) Thus the $\boldsymbol{p}_n$ can be interpreted as probability densities on I, if we normalize them such that $\|p_n\|_1 = 1$. This is most easily expressed using the induced norm $\|\boldsymbol{f}\|_{(n)} := \sum_{k=1}^{N_n} \alpha_{n,k}|f_k|$ on $\mathbb{R}^{N_n}$. Convergence in total variation then corresponds to $\|p_n - p\|_1 \to 0$ [**18**, Thm. 6.13].*

3.3. Convergence of eigenvalues and eigenvectors

We now come to prove the main approximation result:

Theorem 3.7. *With the notation and assumptions from Sections* 2 *and* 3.2*,*

(a) $\lim_{n\to\infty} \lambda_n = \lambda > 0$ *and*

(b) $\lim_{n\to\infty} \|p_n - p\|_1 = 0$*, i.e., the probability measures corresponding to these densities converge in total variation.*

The idea of the proof is as follows. In the following two lemmas, we first determine an upper and a lower bound for the λ_n and conclude that there is a convergent subsequence. Then we show that every convergent subsequence converges to λ and hence the sequence itself. By Theorem 3.5, this implies the convergence of a subsequence of $(q_n/\|q_n\|_\infty)$ to a (non-negative) limit function. Since, due to Theorem 2.2, the latter is unique, we conclude that it is $q/\|q\|_\infty$. With this, part (b) can be shown.

Lemma 3.8. *There is a constant $M > 0$ such that $|\lambda_n| \le M$ for all $n \in \mathbb{N}$.*

Proof. Using (3.8) and (3.9), one checks

$$\begin{aligned}|\lambda_n| = \frac{\|\lambda_n \boldsymbol{p}_n\|_{(n)}}{\|\boldsymbol{p}_n\|_{(n)}} = \frac{\|\boldsymbol{A}_n \boldsymbol{p}_n\|_{(n)}}{\|\boldsymbol{p}_n\|_{(n)}} &\le \sup_{\|\boldsymbol{f}\|_{(n)}=1} \sum_{k=1}^{N_n} \alpha_{n,k} \left| \sum_{\ell=1}^{N_n} (T_{n,k\ell} - U_{n,kl}) f_\ell \right| \\ &\le \max_k w(t_{n,k}) + \max_{k,\ell} u(t_{n,k}, t_{n,\ell}) \sum_{k=1}^{N_n} \alpha_{n,k} \\ &\le \|w\|_\infty + \|u\|_{\mathrm{C}(I\times I)} \sup_m \|Q_m\| =: M > 0\,.\end{aligned}$$

Here, $\|Q_m\| = |I|$ due to (3.7). More generally, $\sup_m \|Q_m\| < \infty$ holds for any convergent quadrature according to the theorem of Banach–Steinhaus, compare [**19**, Thm. 2.5]. □

Lemma 3.9. $\liminf_{n\to\infty} \lambda_n > 0$.

Proof. We start by following Bürger [**4**, p. 134] and show that the spectral radius $\rho(K_\alpha)$ is larger than 1 for sufficiently small $\alpha > 0$, from which then $\lambda > \alpha > 0$ follows by Lemma 2.3. Let J be the interval from (2.9). Then we have

$$(K_\alpha 1_J)(x) = \int_J \frac{u(x,y)}{w(y)+\alpha}\,\mathrm{d}y \ge 1_J(x) \operatorname*{ess\,inf}_{x',y'\in J} u(x',y') \int_J (w(y)+\alpha)^{-1}\,\mathrm{d}y$$

* One may also define operator analogs of the $\boldsymbol{A}_n$, see [**17**, Sec. II.2.1.2].

and thus

$$\|K_\alpha{}^m\|_1{}^{1/m} \geq \operatorname*{ess\,inf}_{x,y\in J} u(x,y) \int_J (w(y)+\alpha)^{-1}\,\mathrm{d}y \qquad \text{for all } m\in\mathbb{N},$$

which implies for the spectral radius

$$\rho(K_\alpha) \geq \operatorname*{ess\,inf}_{x,y\in J} u(x,y) \int_J (w(y)+\alpha)^{-1}\,\mathrm{d}y. \tag{3.11}$$

The RHS is, as a function of α, strictly decreasing. Thus, as a consequence of B. Levi's monotone convergence theorem [**9**, Thm. III.12.22], also

$$\lim_{\alpha\searrow 0} \rho(K_\alpha) \geq \operatorname*{ess\,inf}_{x,y\in J} u(x,y) \int_J (w(y))^{-1}\,\mathrm{d}y > 1$$

according to (2.9) (including divergence of both sides).

Now we choose $\alpha > 0$ such that the RHS of (3.11) is larger than or equal to $1+\varepsilon$, with a sufficiently small $\varepsilon > 0$. Furthermore, we pick, according to the convergence of the quadrature, an n_0 with $\operatorname{ess\,inf}_{x,y\in J} u(x,y)|Q_n(w+\alpha)^{-1} - Q(w+\alpha)^{-1}| < \varepsilon/2$ for all $n \geq n_0$. This way

$$\begin{aligned}(K_{\alpha,n}1_J)(x) &= Q_n\big(u(x,.)(w+\alpha)^{-1}1_J\big) \geq 1_J(x) \operatorname*{ess\,inf}_{x',y'\in J} u(x',y')Q_n(w+\alpha)^{-1} \\ &\geq 1_J(x)\left(\operatorname*{ess\,inf}_{x',y'\in J} u(x',y')Q(w+\alpha)^{-1} - \frac{\varepsilon}{2}\right) \geq 1_J(x)\left(1+\frac{\varepsilon}{2}\right).\end{aligned}$$

Hence, by Lemma 2.3, $\lambda_n > \alpha > 0$ for all $n \geq n_0$, from which the claim follows. □

Proof of Theorem 3.7. By Lemmas 3.8 and 3.9, the sequence $(\lambda_n)_n$ has a convergent subsequence $(\lambda_{n_i})_i$ with limit $\lambda' \in\,]0,M]$. Consider $(K_{\lambda'}f)(x) = Q(k_{\lambda'}(x,.)f)$ as well as $(K_nf)(x) := (K_{\lambda_n,n}f)(x) = Q_n(T+\lambda_n)^{-1}(T+\lambda')(k_{\lambda'}(x,.)f)$. We first show that the 'distorted' quadrature $\tilde{Q}_{n_i} = Q_{n_i}(T+\lambda_{n_i})^{-1}(T+\lambda')$ is convergent. Note that, for i_0 large enough, such that $\inf_{j\geq i_0}\lambda_{n_j} > 0$, and $i \geq i_0$,

$$\begin{aligned}\|(T+\lambda_{n_i})^{-1}(T+\lambda)-1\|_\infty &= \sup_{\|f\|_\infty\leq 1}\left\|\frac{w+\lambda}{w+\lambda_{n_i}}f - f\right\|_\infty \\ &\leq \|(w+\inf_{j\geq i_0}\lambda_{n_j})^{-1}\|_\infty\,|\lambda-\lambda_{n_i}|\,\|f\|_\infty \to 0\,.\end{aligned} \tag{3.12}$$

Then, since (Q_n) is convergent by assumption, we have, for all $f\in \mathrm{C}(I)$,

$$\|(\tilde{Q}_{n_i}-Q)f\|_\infty \leq \|Q_{n_i}((T+\lambda_{n_i})^{-1}(T+\lambda')-1)f\|_\infty + \|(Q_{n_i}-Q)f\|_\infty \to 0,$$

where the first term vanishes in the limit due to $\sup_m \|Q_m\| < \infty$ and (3.12).

With this it follows from Theorem 3.5 that $\rho(K_n) = 1$ is also an eigenvalue of $K_{\lambda'}$ going with a non-negative eigenfunction. The latter is even a.e. positive since, due to the irreducibility (2.8) of $K_{\lambda'}$, there cannot be a set with positive measure on which a non-negative eigenfunction vanishes.* But since, according to Theorem 2.2, there is, up to normalization, only one positive eigenfunction, we have $\lambda' = \lambda$. Therefore every convergent subsequence of $(\lambda_n)_n$ converges to λ, and thus, due to the boundedness, also the sequence itself. This proves part (a).

Along the same line of reasoning, $(a_n q_n)$, with $a_n = 1/\|q_n\|_\infty$, has a convergent subsequence with an a.e. positive limit function, which equals aq with $a = 1/\|q\|_\infty$. Therefore, $\|a_n q_n - aq\|_\infty \to 0$ for $n \to \infty$. Now consider

$$\begin{aligned}
\|a_n p_n - ap\|_\infty &= \|a_n p_n - (T+\lambda)^{-1} aq\|_\infty \\
&= \max_{k\in\mathcal{N}_n} \sup_{x\in I_{n,k}} \left|(w(t_{n,k})+\lambda_n)^{-1} a_n q_{n,k} - (w(x)+\lambda)^{-1} aq(x)\right| \\
&\le \max_{k\in\mathcal{N}_n} \left|(w(t_{n,k})+\lambda_n)^{-1} - (w(t_{n,k})+\lambda)^{-1}\right| a_n q_n(t_{n,k}) + \\
&\quad \max_{k\in\mathcal{N}_n} (w(t_{n,k})+\lambda)^{-1} \left|a_n q_n(t_{n,k}) - aq(t_{n,k})\right| + \\
&\quad \max_{k\in\mathcal{N}_n} \sup_{x\in I_{n,k}} \left|(w(t_{n,k})+\lambda)^{-1} aq(t_{n,k}) - (w(x)+\lambda)^{-1} aq(x)\right| .
\end{aligned}$$

The first term is bounded from above by

$$|\lambda - \lambda_n| \, \|(w + \inf_{m\ge n_0} \lambda_m)^{-1} (w+\lambda)^{-1}\|_\infty ,$$

for $n \ge n_0$ with sufficiently large n_0, and vanishes for $n \to \infty$ due to $\lambda_n \to \lambda$. The second term vanishes due to the uniform convergence of the $a_n q_n$ towards aq, and the third due to the uniform continuity of $(w+\lambda)^{-1} q$ and Lemma 3.6. With this, $a_n p_n \to ap$ in $\mathrm{L}^\infty(I)$ and thus in $\mathrm{L}^1(I)$. Hence, $a_n \to a$ and $p_n \to p$ in $\mathrm{L}^1(I)$, which proves part (b). ☐

4. Discretization—unbounded interval

Now we assume the types to be taken from $I = \mathbb{R}$ and the functions r and u to be continuous. It will be one aim of this section to analyze what further conditions have to be imposed in order to allow for a discretization procedure similar to the one in the previous section. In order to do so, we start by a summary of the relevant theory.

4.1. The Galerkin method

In the Galerkin method, an approximation of compact operators is achieved using projections to finite-dimensional subspaces. This method has been reviewed, e.g., by Krasnosel'skii et al. [**13**, Sec. 18]. The results needed in the sequel are collected in

Theorem 4.1. *Let K be a compact linear operator on the Banach space Y. Consider the eigenvalue equation*

$$(K - \nu) g = 0 , \tag{4.1}$$

* Let $\tilde{q}$ be the eigenfunction and $J = \{x : \tilde{q}(x) > 0\}$ with $0 < |J|$. Assume $|J| < |I|$. Then, for $x \in I\backslash J$, we have $0 = \tilde{q}(x) = \int_J k_{\lambda'}(x,y)\tilde{q}(y)\,\mathrm{d}y$, which implies, for a.e. $y \in J$, that $k_{\lambda'}(x,y)\tilde{q}(y) = 0$ and thus $u(x,y) = 0$, contradicting (2.8).

which is to be approximated. To this end, let (Y_n) be a sequence of closed subspaces of Y with bounded projections P_n onto them. On these subspaces, let the compact linear operators K_n be defined, together with the eigenvalue equations

$$(K_n - \nu_n)g_n = 0\,. \tag{4.2}$$

Assume that

$$\|K_n - P_nK\|_{Y_n} \to 0\,, \quad \|K - P_nK\|_Y \to 0 \qquad \text{as } n \to \infty. \tag{4.3}$$

Then the following statements are true:

(a) *For every $\nu \neq 0$ from (4.1) there is a sequence (ν_n) of eigenvalues of (4.2) such that $\nu_n \to \nu$ as $n \to \infty$. Conversely, every non-zero limit point of any sequence (ν_n) of eigenvalues of (4.2) is an eigenvalue of (4.1).*

(b) *Every bounded sequence (g_n) of eigenvectors of (4.2) associated with eigenvalues $\nu_n \to \nu \neq 0$ contains a convergent subsequence; the limit of any convergent subsequence $(g_{n_i})_i$ is an eigenvector of (4.1) associated with the eigenvalue ν (unless the limit is zero).*

Proof. See [**13**, Thms. 18.1, 18.2]. □

A sufficient condition for the validity of the second assumption in (4.3) is given by

Proposition 4.2. *Let X be a normed space, Y a Banach space, and $K\colon X \to Y$ a compact linear operator. For bounded linear operators $P_n\colon Y \to Y$ $(n \in \mathbb{N})$ with $P_n \to \mathbb{1}$ pointwise for $n \to \infty$, the operators P_nK approximate K, i.e., $\|P_nK - K\| \to 0$.*

Proof. Follow the proof of [**21**, Thm. II.3.5], where the additional assumptions on X and (P_n) are not used. □

4.2. Application to kernel operators

In our case of the COA model we have $X = Y = \mathrm{L}^1(\mathbb{R})$ and K is of the form

$$(Kf)(x) = \int_{\mathbb{R}} k(x,y)f(y)\,\mathrm{d}y \qquad \text{for all } x \in \mathbb{R} \tag{4.4}$$

with a measurable kernel $k\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Therefore, for the Galerkin method to work, it is necessary that, for $\mathrm{L}^1(\mathbb{R})$, operators P_n as in Proposition 4.2 exist. We will explicitly construct such operators using a sequence $(\{I_{n,k} : 1 \le k \le N_n\})_n$ of families of disjoint intervals that get finer and finer and also ultimately cover every bounded interval.*

Proposition 4.3. *Let Y be the Banach space $\mathrm{L}^1(\mathbb{R})$ and finite-dimensional subspaces Y_n of Y chosen to consist of all step functions with prescribed (bounded) intervals $I_{n,k}$ $(k \in \mathcal{N}_n := \{1, \dots, N_n\})$ with the following properties:*

* Both properties are formally captured by (I1) in Proposition 4.3.

(I1) *For every bounded interval $I \subset \mathbb{R}$ and every $\varepsilon > 0$ there is an n_0 such that, for all $n \geq n_0$, a set $L \subset \mathcal{N}_n$ exists for which $I_{n,L} := \bigcup_{\ell \in L} I_{n,\ell}$ satisfies $|I \backslash I_{n,L}| = 0$ and $|I_{n,L} \backslash I| < \varepsilon$. (We then say that I is ε-*optimally covered.)

(I2) $|I_{n,k} \cap I_{n,\ell}| = 0$ *for all* $n \in \mathbb{N}$ *and* $1 \leq k < \ell \leq N_n$.

Then, with the characteristic functions $\varphi_{n,k} = 1_{I_{n,k}}$, the projections P_n onto the subspaces Y_n spanned by $\{\varphi_{n,k} : k \in \mathcal{N}_n\}$ are given by

$$P_n f = \sum_{k=1}^{N_n} \varphi_{n,k} \frac{1}{|I_{n,k}|} \int_{I_{n,k}} f(x)\, \mathrm{d}x \qquad \text{for } f \in \mathrm{L}^1(\mathbb{R}),$$

where $\int_{I_{n,k}} f(x)\, \mathrm{d}x$ are conditional expectations (*compare* [**20**, Thm. IV.2.4]*). *The projections satisfy $\|P_n\|_1 = 1$ and $P_n \to \mathbb{1}$ pointwise.*

Proof. Obviously, the subspaces Y_n are closed, finite-dimensional, and the P_n are, due to (I2), projections onto them. Since

$$\|P_n f\| = \sum_{k=1}^{N_n} \int_{I_{n,k}} |f(x)|\, \mathrm{d}x \leq \int_{\mathbb{R}} |f(x)|\, \mathrm{d}x = \|f\|_1 \qquad \text{for every } f \in \mathrm{L}^1(\mathbb{R})$$

and $\|P_n \varphi_{n,k}\| = \|\varphi_{n,k}\|$ for every $k \in \mathcal{N}_n$, we have $\|P_n\| = 1$.

We now show that $P_n \to \mathbb{1}$ pointwise. To this end, let $f \in \mathrm{L}^1(\mathbb{R})$ and $\varepsilon > 0$ be given. Remember that the set of all step functions is, by definition, dense in $\mathrm{L}^1(\mathbb{R})$, compare [**15**, Sec. VI.3]. Therefore, we can find a step function $\psi = \sum_{k=1}^{m} \psi_k 1_{J_k}$ (with bounded intervals J_k) that satisfies $\|f - \psi\|_1 < \varepsilon/3$. Due to (I1) we can now choose an n_0 such that $|\bigcup_{k=1}^{m} J_k \backslash \bigcup_{k=1}^{N_n} I_{n,k}| = 0$ for all $n \geq n_0$. Then, the only contributions to $\|P_n \psi - \psi\|_1$ are due to mismatches at the boundaries of the J_k. Therefore, let J_k^+ and J_k^- ($k \in \mathcal{N}_n$) be open intervals of measure $\eta = \varepsilon/(12m \max_k |\psi_k|)$ that contain the right and left boundary points of J_k, respectively. Choosing $n_1 \geq n_0$ according to (I1) large enough such that every $J_k^\pm$ is η-optimally covered for $n \geq n_1$, we have $\|P_n \psi - \psi\|_1 < 2 \sum_{k=1}^{m} 2\eta\psi_k \leq \varepsilon/3$ for $n \geq n_1$. Putting everything together yields, for $n \geq n_1$,

$$\|P_n f - f\|_1 \leq \|P_n(f - \psi)\|_1 + \|P_n \psi - \psi\|_1 + \|\psi - f\|_1 < \varepsilon\,,$$

which proves $\|P_n f - f\| \to 0$ for $n \to \infty$ and thus the approximation property. □

With respect to a kernel operator K of the form (4.4) and some $f = \sum_{k=1}^{N_n} \varphi_{n,k} f_k$ in Y_n, the above procedure amounts to the discretization

$$\begin{aligned}(P_n K f) &= \sum_{k=1}^{N_n} \varphi_{n,k} \frac{1}{|I_{n,k}|} \left(\int_{I_{n,k}} \int_{\mathbb{R}} k(x,y) \sum_{\ell=1}^{N_n} f_\ell \varphi_{n,\ell}(y)\, \mathrm{d}y\, \mathrm{d}x \right) \\ &= \sum_{k=1}^{N_n} \varphi_{n,k} \sum_{\ell=1}^{N_n} \frac{1}{|I_{n,k}|} \left(\int_{I_{n,k}} \int_{I_{n,\ell}} k(x,y)\, \mathrm{d}y\, \mathrm{d}x \right) f_\ell =: \sum_{k=1}^{N_n} \varphi_{n,k} \sum_{\ell=1}^{N_n} M_{n,k\ell} f_\ell\end{aligned}$$

* See also [**17**] for a discussion of the connection to the approximation property of Banach spaces.

with an $N_n \times N_n$ matrix $\boldsymbol{M}_n = (M_{n,k\ell})$. The corresponding eigenvalue equation is

$$\boldsymbol{M}_n \boldsymbol{g}_n = \nu_n \boldsymbol{g}_n\,, \qquad \text{or equivalently} \quad P_n K g_n = \nu_n g_n\,,$$

where $g_n \in Y_n$ is granted due to the projection property. An example of intervals $I_{n,k}$ satisfying (I1) and (I2) is $I_{n,k} = [-n+2^{-n}(k-1), -n+2^{-n}k]$ with $k \in \mathcal{N}_n = \{1, \dots, 2^{n+1}n\}$.

With respect to compactness of K, following Jörgens [**11**, Secs. 11, 12], we extract

Proposition 4.4. *A kernel operator K on $\mathrm{L}^1(\mathbb{R})$ of the form* (4.4) *is compact if it satisfies the following conditions:*

(C1) *The function $x \to k(x, .)$ from $\mathbb{R}$ to $\mathrm{L}^1(\mathbb{R})$ is continuous and bounded.*

(C2) *For every $\varepsilon > 0$ there exists a finite open covering $(V_1, \dots, V_n)$ of $\mathbb{R}$ and points $x_j \in V_j$ such that $\|k(x, .) - k(x_j, .)\|_1 < \varepsilon$ for all $x \in V_j$ and all j.*

(C3) *The function $y \to k(., y)$ from $\mathbb{R}$ to $\mathrm{L}^1(\mathbb{R})$ is continuous and bounded.*

(C4) *For every $\varepsilon > 0$ there exists a finite open covering $(W_1, \dots, W_n)$ of $\mathbb{R}$ and points $y_j \in W_j$ such that $\|k(., y) - k(., y_j)\|_1 < \varepsilon$ for all $y \in W_j$ and all j.*

Proof. First, as in [**11**, Sec. 12.4], we consider the dual system $\langle \mathrm{C}(\mathbb{R}), \mathrm{C}_1(\mathbb{R})\rangle$ with the bilinear form $\langle f, g\rangle = \int_{\mathbb{R}} f(x) g(x)\, \mathrm{d}x$. Here, $\mathrm{C}(\mathbb{R})$ is equipped with the supremum norm $\|.\|_\infty$ and $\mathrm{C}_1(\mathbb{R}) := \mathrm{C}(\mathbb{R}) \cap \mathrm{L}^1(\mathbb{R})$ with the norm $|\!|\!|.|\!|\!| := \max\{\|.\|_\infty, \|.\|_1\}$. With this, we define the transposed K^{T} of K via $(K^{\mathrm{T}} g)(y) = \int_{\mathbb{R}} g(x) k(x, y)\, \mathrm{d}x$, for all $y \in \mathbb{R}$. Then, by (C1)–(C4) and [**11**, Thms. 12.2, 12.3], the compactness of K and K^{T} on $\mathrm{C}(\mathbb{R})$ follows.

As both K^{T} and K are bounded as operators on $\mathrm{C}(\mathbb{R})$, they are, at the same time, Hille–Tamarkin operators in $\mathcal{H}_{\infty\infty}(\mathbb{R})$ since the respective norm, $|.|_{\infty\infty}$ in (2.7), is just given by $\sup_{x\in\mathbb{R}} \int_{\mathbb{R}} k(x, y)\, \mathrm{d}y$ and $\sup_{y\in\mathbb{R}} \int_{\mathbb{R}} k(x, y)\, \mathrm{d}x$, respectively [**11**, Thms. 12.2, 12.3]. Then, according to [**11**, Thm. 11.5], K and K^{T} can also be regarded as bounded operators on $\mathrm{L}^1(\mathbb{R})$; thus, both map $\mathrm{C}_1(\mathbb{R})$ into itself. Due to [**11**, Thm. 12.6] there is, for every $\varepsilon > 0$, an operator of finite rank, K_ε, with $|\!|\!|K_\varepsilon - K|\!|\!| < \varepsilon$, where $|\!|\!|A|\!|\!| := \max\{\|A\|_\infty, \|A^{\mathrm{T}}\|_\infty\}$ is a norm for the Banach algebra of all operators on $\mathrm{C}(\mathbb{R})$ that map $\mathrm{C}_1(\mathbb{R})$ into itself and have a transposed of the same kind. We have $|\!|\!|Af|\!|\!| \le |\!|\!|A|\!|\!|\,|\!|\!|f|\!|\!|$ for $f \in \mathrm{C}_1(\mathbb{R})$, see [**11**, Sec. 12.4]. Thus, $|\!|\!|A|\!|\!|$ can serve as an upper bound for the operator norm of A on $\mathrm{C}_1(\mathbb{R})$. Therefore, K is compact as an operator on $\mathrm{C}_1(\mathbb{R})$ and can be approximated by K_ε. Furthermore, according to [**11**, Thm. 11.5], $\|K_\varepsilon - K\|_1 \le |\!|\!|(K_\varepsilon - K)^{\mathrm{T}}|\!|\!| < \varepsilon$ holds. Hence, K is compact as an operator on $\mathrm{L}^1(\mathbb{R})$ as well. □

4.3. Application to the COA model

Checking the compactness of K_α by conditions (C1)–(C4) of Proposition 4.4, we would be able to apply Theorem 4.1 and approximate K_α by operators of finite rank. However, the original system is described by the (non-compact) operator $A = T - U$, not by some K_α. It will be shown that it is indeed possible to discretize the operators T and U directly by applying the projections P_n from Proposition 4.3, if further restrictions apply. Then, even more generally, the approximation can be done by choosing arbitrary points in the intervals $I_{n,k}$ at which the functions w and u are sampled. Both procedures will now be described in detail.

In the first setting, K_λ is approximated by $K_n := P_n U(P_n T + \lambda_n)^{-1}$. Explicitly, for $f \in Y_n$ with $f = \sum_{k=1}^{N_n} f_k \varphi_{n,k}$, it reads

$$\begin{aligned} P_n T f &= \sum_{k=1}^{N_n} \varphi_{n,k} \frac{1}{|I_{n,k}|} \int_{I_{n,k}} w(x)\,\mathrm{d}x\, f_k &&= \sum_{k=1}^{N_n} \varphi_{n,k} w(t^w_{n,k}) f_k\,, \\ P_n U f &= \sum_{k,\ell=1}^{N_n} \varphi_{n,k} \frac{1}{|I_{n,k}|} \int_{I_{n,k}} \int_{I_{n,\ell}} u(x,y)\,\mathrm{d}y\,\mathrm{d}x\, f_\ell &&= \sum_{k,\ell=1}^{N_n} \varphi_{n,k} |I_{n,\ell}| u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell}) f_\ell\,, \end{aligned}$$

with appropriate points $t^w_{n,k}$, $t^{ux}_{n,k\ell} \in I_{n,k}$ and $t^{uy}_{n,k\ell} \in I_{n,\ell}$ that satisfy

$$\frac{1}{|I_{n,k}|} \int_{I_{n,k}} u(x, t^{uy}_{n,k\ell})\,\mathrm{d}x = u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})\,. \tag{4.5}$$

These exist due to the continuity of w and u. But more generally, we may pick the points arbitrarily from the respective intervals.

In either case, we define the $N_n \times N_n$ matrices $\boldsymbol{T}_n$, $\boldsymbol{U}_n$, and $\boldsymbol{A}_n := \boldsymbol{T}_n - \boldsymbol{U}_n$ via

$$T_{n,kk} := w(t^w_{n,k})\,, \qquad U_{n,k\ell} := |I_{n,\ell}|\, u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})\,. \tag{4.6}$$

The corresponding operators in Y_n are given by

$$T_n f = \sum_{k=1}^{N_n} \varphi_{n,k} T_{n,kk} f_k\,, \qquad U_n f = \sum_{k,\ell=1}^{N_n} \varphi_{n,k} U_{n,k\ell} f_\ell\,, \qquad A_n = T_n - U_n\,,$$

again with $f = \sum_{k=1}^{N_n} f_k \varphi_{n,k}$. For notational convenience, we also define the matrices $\boldsymbol{P}_{\alpha,n}$ by

$$P_n K_\alpha f = \sum_{k,\ell=1}^{N_n} \varphi_{n,k} \frac{1}{|I_{n,k}|} \int_{I_{n,k}} \int_{I_{n,\ell}} \frac{u(x,y)}{w(y)+\alpha}\,\mathrm{d}y\,\mathrm{d}x f_\ell =: \sum_{k,\ell=1}^{N_n} \varphi_{n,k} P_{\alpha,n,k\ell} f_\ell\,. \tag{4.7}$$

The eigenvalue equation to be solved is

$$(\boldsymbol{A}_n + \lambda_n)\boldsymbol{p}_n = 0\,,$$

which is equivalent to

$$(A_n + \lambda_n) p_n = 0\,, \tag{4.8}$$

where $p_n = \sum_{k=1}^{N_n} p_{n,k} \varphi_{n,k} \in Y_n$. With $K_{\alpha,n} = U_n(T_n + \alpha)^{-1}$, $\alpha > -\min_{k \in \mathcal{N}_n} w(t_{n,k})$, and $q_n = (T_n + \lambda_n) p_n$ also the eigenvalue equation

$$(K_{\lambda_n,n} - 1) q_n = 0$$

is solved (and vice versa), cf. Lemma 2.1. (The inequality $\lambda_n > -\min_{k\in\mathcal{N}_n} w(t_{n,k})$ follows from Theorem 2.2.)

For these procedures to be valid approximations, the first condition in (4.3), that is, $\|K_n - P_n K\|_{Y_n} \to 0$, has to be true for $K = K_\lambda$ and $K_n = U_n(T_n + \lambda_n)^{-1}$. This, however, is not given automatically. Problems arise from the fact that in K_n the averaging defined by P_n (or, more generally, the sampling) is applied to the enumerator and denominator of k_{λ_n} separately, whereas in $P_n K$ the quotient k_λ is averaged as such. It turns out that some additional requirements of uniform continuity are sufficient for the convergence. This is made precise in the following two propositions.

Proposition 4.5. *Suppose that the following conditions are true:*

(S1) $u(x,.)$ *is uniformly continuous for all* $x \in \mathbb{R}$.

(S2) k_α *is uniformly continuous on* $I \times \mathbb{R}$ *for all* $\alpha > 0$ *and all bounded* $I \subset \mathbb{R}$.

(S3) *There is a function* $w_{\min}\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$, *satisfying*

$$\int_{\mathbb{R}} \sup_{y\in\mathbb{R}} \frac{u(x,y)}{w_{\min}(y) + \alpha}\,\mathrm{d}x < \infty \qquad \text{for all } \alpha > 0,$$

and an $n_0 \in \mathbb{N}$ *such that* $w(y) \geq w_{\min}(y')$ *for all* $n \geq n_0$, $\ell \in \mathcal{N}_n$, *and* y, $y' \in I_{n,\ell}$.

Then, for $K = K_\alpha$ *and* $K_n = P_n U(P_n T + \alpha)^{-1}$ *with any* $\alpha > 0$ *and the projections* P_n *from Proposition 4.3, the first condition in* (4.3) *is fulfilled, i.e.,* $\|K_n - P_n K\|_{Y_n} \to 0$ *as* $n \to \infty$. *The same is true for* $K_n = K_{\alpha,n} = U_n(T_n + \alpha)^{-1}$ *with the more general discretization from above if in addition to* (S1)–(S3) *the following condition is satisfied:*

(S4) *There is a function* $u_{\max}\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$, *satisfying*

$$\int_{\mathbb{R}} \sup_{y\in\mathbb{R}} \frac{u_{\max}(x,y)}{w_{\min}(y) + \alpha}\,\mathrm{d}x < \infty \qquad \text{for all } \alpha > 0,$$

and an $n_1 \geq n_0$ *such that* $u(x,y) \leq u_{\max}(x',y)$ *for all* $n \geq n_1$, $k \in \mathcal{N}_n$, $y \in \mathbb{R}$, *and* x, $x' \in I_{n,k}$.

Let us split the rather technical proof into a couple of digestible lemmas.

Lemma 4.6. *If conditions* (S1) *and* (S2) *are true, then for every* $\varepsilon > 0$ *and every compact interval* $I \subset \mathbb{R}$ *there is an* n_2 *such that for all* $n \geq n_2$ *and all* k, $\ell \in \mathcal{N}_n$ *with* $I_{n,k} \cap I \neq \emptyset$ *we have*

$$\frac{1}{|I_{n,\ell}|}\left|P_{\alpha,n,k\ell} - \frac{U_{n,k\ell}}{T_{n,\ell\ell} + \alpha}\right| < \frac{\varepsilon}{|I|}\,.$$

Proof. Let ε and I be given as above and $I_0 = \bigcup_{n\in\mathbb{N}} \bigcup_{k:I_{n,k}\cap I\neq\emptyset} I_{n,k}$, which is a bounded interval due to (I1) from Proposition 4.3. By assumptions (S1) and (S2), u and

k_α are uniformly continuous on $\overline{I_0} \times \mathbb{R}$. Further, $(w+\alpha)^{-1}$ is bounded by α^{-1}. Thus, there is an n_2 such that, for every $n \geq n_2$ and $k, \ell \in \mathcal{N}_n$ with $I_{n,k} \cap I \neq \emptyset$,

$$\begin{aligned}
\left| \frac{1}{|I_{n,k}|} \frac{1}{|I_{n,\ell}|} \int_{I_{n,k}} \int_{I_{n,\ell}} \frac{u(x,y)}{w(y)+\alpha} \,\mathrm{d}y \,\mathrm{d}x - \frac{u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})}{w(t^{w}_{n,\ell})+\alpha} \right| \\
= \left| \frac{u(t^{kx}_{n,k\ell}, t^{ky}_{n,k\ell})}{w(t^{ky}_{n,k\ell})+\alpha} - \frac{u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})}{w(t^{w}_{n,\ell})+\alpha} \right| \\
\leq \left| \frac{u(t^{kx}_{n,k\ell}, t^{ky}_{n,k\ell})}{w(t^{ky}_{n,k\ell})+\alpha} - \frac{u(t^{ux}_{n,k\ell}, t^{w}_{n,\ell})}{w(t^{w}_{n,\ell})+\alpha} \right| + \frac{|u(t^{ux}_{n,k\ell}, t^{w}_{n,\ell}) - u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})|}{w(t^{w}_{n,\ell})+\alpha} < \frac{\varepsilon}{|I|} \,.
\end{aligned}$$

Here, the points $t^{kx}_{n,k\ell} \in I_{n,k}$ and $t^{ky}_{n,k\ell} \in I_{n,\ell}$ are chosen such that the first equality holds, which is possible due to the continuity of k_α. From this the claim follows easily with (4.6) and (4.7). □

Lemma 4.7. *For every $\varepsilon > 0$ there is a compact interval I_1 such that, for all intervals $I \supset I_1$ and all $n \in \mathbb{N}$,*

$$\sum_{\substack{k \\ I_{n,k} \cap I = \emptyset}} |I_{n,k}| \max_{\ell \in \mathcal{N}_n} \frac{P_{\alpha,n,k\ell}}{|I_{n,\ell}|} < \varepsilon \,.$$

Proof. Due to (U4) there is a compact interval I_1 such that, for all $I \supset I_1$,

$$\begin{aligned}
\sum_{\substack{k \\ I_{n,k} \cap I = \emptyset}} |I_{n,k}| \max_{\ell \in \mathcal{N}_n} \frac{P_{\alpha,n,k\ell}}{|I_{n,\ell}|} &\leq \sum_{\substack{k \\ I_{n,k} \cap I = \emptyset}} |I_{n,k}| \frac{1}{|I_{n,k}|} \int_{I_{n,k}} \max_{y \in \mathbb{R}} \frac{u(x,y)}{w(y)+\alpha} \,\mathrm{d}x \\
&\leq \int_{\mathbb{R} \setminus I_1} \max_{y \in \mathbb{R}} \frac{u(x,y)}{w(y)+\alpha} \,\mathrm{d}x < \varepsilon \,,
\end{aligned}$$

which proves the claim. □

Lemma 4.8. *If condition* (S3) *is true, and if*

(i) $U_{n,k\ell} = \frac{|I_{n,\ell}|}{|I_{n,k}|} \int_{I_{n,k}} u(x, t^{uy}_{n,k\ell}) \,\mathrm{d}x$ *for all* $k, \ell \in \mathcal{N}_n$ *or*

(ii) *condition* (S4) *is fulfilled,*

then for every $\varepsilon > 0$ there is a compact interval I_2 such that, for all intervals $I \supset I_2$ and all $n \in \mathbb{N}$,

$$\sum_{\substack{k \\ I_{n,k} \cap I = \emptyset}} |I_{n,k}| \max_{\ell \in \mathcal{N}_n} \frac{U_{n,k\ell}}{|I_{n,\ell}|(T_{n,\ell\ell}+\alpha)} < \varepsilon \,.$$

Proof. In case (i) we have, using (4.5),

$$\begin{aligned}\sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} |I_{n,k}| \max_{\ell\in\mathcal{N}_n} \frac{u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})}{w(t^w_{n,\ell})+\alpha} &= \sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} \max_{\ell\in\mathcal{N}_n} \frac{\int_{I_{n,k}} u(x, t^{uy}_{n,k\ell})\,\mathrm{d}x}{w(t^w_{n,\ell})+\alpha} \\ &\le \sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} \int_{I_{n,k}} \max_{y\in\mathbb{R}} \frac{u(x,y)}{w_{\min}(y)+\alpha}\,\mathrm{d}x \le \int_{\mathbb{R}\setminus I_2} \max_{y\in\mathbb{R}} \frac{u(x,y)}{w_{\min}(y)+\alpha}\,\mathrm{d}x < \varepsilon\end{aligned}$$

for some compact interval I_2, due to (S3), and all intervals $I \supset I_2$. In case (ii) we can find, due to (S4), a compact interval I_2 such that, for all intervals $I \supset I_2$,

$$\begin{aligned}\sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} |I_{n,k}| \max_{\ell\in\mathcal{N}_n} \frac{u(t^{ux}_{n,k\ell}, t^{uy}_{n,k\ell})}{w(t^w_{n,\ell})+\alpha} &\le \sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} |I_{n,k}| \max_{y\in\mathbb{R}} \frac{u(t^{ux}_{n,k\ell}, y)}{w_{\min}(y)+\alpha} \\ &\le \sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} \max_{y\in\mathbb{R}} \frac{\int_{I_{n,k}} u_{\max}(x,y)\,\mathrm{d}x}{w_{\min}(y)+\alpha} \le \int_{\mathbb{R}\setminus I_2} \max_{y\in\mathbb{R}} \frac{u_{\max}(x,y)}{w_{\min}(y)+\alpha}\,\mathrm{d}x < \varepsilon\,.\end{aligned}$$

This completes the proof. □

Proof of Proposition 4.5. Let $\varepsilon > 0$ be given. Choose a compact interval I such that $I \supset I_1 \cup I_2$ with I_1 and I_2 from Lemmas 4.7 and 4.8. Let $I_3 = \overline{\bigcup_{n,k: I_{n,k}\cap I\neq\emptyset} I_{n,k}}$. Further, let n_0 be as in (S3), n_1 as in (S4) (or $n_0 = n_1$ if not applicable), n_2 as in Lemma 4.6, and $n \ge \max\{n_0, n_1, n_2\}$. Then

$$\begin{aligned}\|P_n K_\alpha - K_{\alpha,n}\|_{Y_n} &= \sup_{\substack{f\in Y_n\\ \|f\|_{Y_n}\le 1}} \sum_{k=1}^{N_n} |I_{n,k}| \left| \sum_{\ell=1}^{N_n} \left(P_{\alpha,n,k\ell} - \frac{U_{n,k\ell}}{T_{n,\ell\ell}+\alpha}\right) f_\ell \right| \\ &\le \Bigg(\sum_{\substack{k\\ I_{n,k}\cap I\neq\emptyset}} + \sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} \Bigg) |I_{n,k}| \max_{\ell\in\mathcal{N}_n} \frac{1}{|I_{n,\ell}|} \left| P_{\alpha,n,k\ell} - \frac{U_{n,k\ell}}{T_{n,\ell\ell}+\alpha} \right| \\ &\le \sum_{\substack{k\\ I_{n,k}\cap I\neq\emptyset}} |I_{n,k}| \frac{\varepsilon}{|I_3|} + \sum_{\substack{k\\ I_{n,k}\cap I=\emptyset}} |I_{n,k}| \left(\max_{\ell\in\mathcal{N}_n} \frac{P_{\alpha,n,k\ell}}{|I_{n,\ell}|} + \max_{\ell\in\mathcal{N}_n} \frac{U_{n,k\ell}}{|I_{n,\ell}|(T_{n,\ell\ell}+\alpha)} \right) < 3\varepsilon\end{aligned}$$

according to Lemmas 4.6–4.8. From this the claim follows. □

Proposition 4.9. *Let $\alpha_n > -\min_{k\in\mathcal{N}_n} w(t_{n,k})$ with $\alpha_n \to \alpha > 0$ as $n \to \infty$ and the hypotheses of Proposition* 4.5 *be satisfied. Then $\|K_{\alpha_n,n} - P_n K_\alpha\|_{Y_n} \to 0$.*

Proof. Consider

$$\begin{aligned}&\|P_n K_\alpha - U_n(T_n+\alpha_n)^{-1}\|_{Y_n} \\ &\qquad \le \|P_n K_\alpha - U_n(T_n+\alpha)^{-1}\|_{Y_n} + \|U_n[(T_n+\alpha_n)^{-1} - (T_n+\alpha)^{-1}]\|_{Y_n}\,.\end{aligned}$$

The first term tends to zero as $n \to \infty$ according to Proposition 4.5. For the second, choose n_0 such that $\inf_{n \geq n_0} \alpha_n > 0$. Then, for $n \geq n_0$,

$$\begin{aligned}\|U_n[(T_n+\alpha_n)^{-1} - (T_n+\alpha)^{-1}]\|_{Y_n} \\ = |\alpha - \alpha_n| \, \|U_n(T_n+\alpha_n)^{-1}(T_n+\alpha)^{-1}\|_{Y_n} \leq |\alpha-\alpha_n| \, \|U\|_Y (\inf_{n\geq n_0} \alpha_n)^{-1}\alpha^{-1} .\end{aligned}$$

This vanishes as $n \to \infty$ since all constants that occur are finite, from which the claim follows. □

4.4. Convergence of eigenvalues and eigenvectors

Let us now show

Theorem 4.10. *With the notation and assumptions from Section* 2 *and* λ, p, λ_n, p_n *as in* (2.2) *and* (4.8), *we have*

(a) $\lim_{n\to\infty} \lambda_n = \lambda > 0$ *and*

(b) $\lim_{n\to\infty} \|p_n - p\|_1 = 0$, *i.e., the probability measures corresponding to these densities converge in total variation.*

The plan is the same as described in Section 3.3. The proofs, however, are quite different due to the more general setup.

Lemma 4.11. *There is a constant* $M > 0$ *such that* $\limsup_{n\to\infty} \lambda_n \leq M$.

Proof. Choose an $\alpha > 0$ such that $\|K_\alpha\|_Y \leq 1 - \varepsilon$ for some $0 < \varepsilon < 1$, which is possible since $\|K_\alpha\|_Y \to 0$ for $\alpha \to \infty$. Then, for all $n \geq n_0$ with some n_0, due to Propositions 4.2 and 4.5, $|\|P_n K_\alpha\|_Y - \|K_\alpha\|_Y| \leq \varepsilon/3$ and $|\|K_{\alpha,n}\|_{Y_n} - \|P_n K_\alpha\|_{Y_n}| \leq \varepsilon/3$. For these n, we have

$$\begin{aligned}\rho(K_{\alpha,n}) \leq \|K_{\alpha,n}\|_{Y_n} \leq \|P_n K_\alpha\|_{Y_n} + \varepsilon/3 &\leq \|P_n K_\alpha\|_Y + \varepsilon/3 \\ &\leq \|K_\alpha\|_Y + 2\varepsilon/3 \leq 1 - \varepsilon/3 < 1\end{aligned}$$

and thus $\lambda_n < \alpha$ by Lemma 2.3. Then, with $M = \alpha$, the claim follows. □

Lemma 4.12. $\liminf_{n\to\infty} \lambda_n > 0$.

Proof. In a modification of the proof of Lemma 3.9, we choose $\alpha > 0$ such that $\rho(K_\alpha) \geq 1 + \varepsilon$ with a sufficiently small $\varepsilon > 0$. We know from the theorem of Jentzsch [**20**, Thm. V.6.6] that $\rho(K_\alpha)$ is a simple eigenvalue of K_α and the only one with a positive eigenfunction. The same is true for $\rho(K_{\alpha,n})$ with respect to $K_{\alpha,n}$ (as an operator in Y_n). Theorem 4.1 together with Proposition 4.5 implies that there is a sequence of eigenvalues ν_n of $K_{\alpha,n}$ with limit $\rho(K_\alpha)$. Therefore, $\liminf_{n\to\infty} \rho(K_{\alpha,n}) \geq \rho(K_\alpha) \geq 1 + \varepsilon$ and thus $\lambda_n > \alpha > 0$ for sufficiently large n. From this the claim follows. □

Proof of Theorem 4.10. From Lemma 4.11 and 4.12 we conclude that there is a convergent subsequence $(\lambda_{n_i})_i$ with limit $\lambda' \in \,]0, M]$. Then, due to Proposition 4.9, $K_{\lambda_{n_i},n_i}$ converges to $P_n K_{\lambda'}$ in norm. Hence, $\lim_{i\to\infty} \rho(K_{\lambda_{n_i},n_i}) = \rho(K_{\lambda_{n_i},n_i}) = 1$ is an eigenvalue of $K_{\lambda'}$ by Theorem 4.1. Furthermore, a subsequence of $(a_{n_i} q_{n_i})$, where $a_n = 1/\|q_n\|_1$, converges to an eigenfunction $\tilde{q}$ of $K_{\lambda'}$, and $\tilde{q} \geq 0$ (but $\tilde{q} \neq 0$). As there is only one non-negative eigenfunction by Theorem 2.2, we conclude $\lambda' = \lambda$ and $\tilde{q} = aq$ with $a = 1/\|q\|_1$. Since this is true for every convergent subsequence of (λ_n), the claim of part (a) and the convergence $a_n q_n \to aq$ follow.

Now, let n_0 be sufficiently large such that $\alpha := \inf_{n\geq n_0} \lambda_n > 0$. Then, for $n \geq n_0$,

$$\begin{aligned}
\|a_n p_n - (T+\lambda)^{-1} aq\|_1 &= \|(P_n T + \lambda_n)^{-1} a_n q_n - (T+\lambda)^{-1} aq\|_1 \\
&\leq \|[(P_n T + \lambda_n)^{-1} - (P_n T + \lambda)^{-1}] a_n q_n\|_1 + \\
&\quad\ \|(P_n T + \lambda)^{-1}(a_n q_n - aq)\|_1 + \\
&\quad\ \|[(P_n T + \lambda)^{-1} - (T+\lambda)^{-1}] aq\|_1 \\
&\leq \tfrac{1}{\alpha\lambda}|\lambda - \lambda_n| + \tfrac{1}{\lambda}\|a_n q_n - aq\|_1 + \tfrac{1}{\lambda^2}\|(\mathbb{1} - P_n) T aq\|_1 \to 0\,.
\end{aligned}$$

With this, $a_n p_n \to ap$ in $\mathrm{L}^1(I)$, hence $a_n \to a$ and $p_n \to p$, which proves part (b). □

5. Comparison of both methods

Both approaches, the application of the Nyström method in the case of a compact interval and of the Galerkin method in the case of an unbounded interval, effectively lead to the same approximation procedure in our case of the COA model. First, one chooses appropriate intervals $I_{n,k}$ and points $t_{n,k} \in I_{n,k}$ (also for an unbounded interval the use of identical points $t^w_{n,k} = t^{ux}_{n,k\ell} = t^{uy}_{n,\ell k} = t_{n,k}$ seems reasonable in many cases). Then, the operators T and U from (2.3) and (2.4), respectively, are approximated by matrices $\boldsymbol{T}_n$ and $\boldsymbol{U}_n$, cf. (3.8), (3.9), and (4.6). For these, the (finite-dimensional) eigenvalue problem $(\boldsymbol{T}_n - \boldsymbol{U}_n + \lambda_n)\boldsymbol{p}_n = 0$ is solved. Here, the eigenvectors $\boldsymbol{p}_n$ are considered as probability *densities* on I. Then, under the conditions described above, the eigenvalues λ_n converge to λ and the measures corresponding to the $\boldsymbol{p}_n$ converge in total variation to the equilibrium distribution described by the solution p of the original problem (1.1).

The differences between the two approaches lie on the intermediate technical level of the compact operators K_α and $K_{\alpha,n}$ and the solutions q and q_n of the equivalent eigenvalue problems (2.6), (3.5), and (4.2). Here, in the first case we have collectively compact convergence $K_{\lambda_n,n} \xrightarrow{\text{cc}} K_\lambda$ going together with $\|q_n - q\|_\infty \to 0$, whereas in the second case $\|P_n K_\lambda - K_\lambda\|_Y \to 0$ in $Y = \mathrm{L}^1(\mathbb{R})$ and $\|K_{\lambda_n,n} - P_n K_\lambda\|_{Y_n} \to 0$ in the subspaces Y_n going together with $\|q_n - q\|_1 \to 0$. On this level, neither does $\|K_{\lambda_n,n} - K_\lambda\|_\infty \to 0$ hold in the first case, compare [**14**, Thm. 12.8], nor any kind of collectively compact convergence in the second.

Both methods may, strictly speaking, only be applied to continuous mutation kernels u. This excludes, for example, Γ-distributions (reflected at the source type), where $u(x,y) \propto |x-y|^{\Theta-1}\exp(-d\,|x-y|)$, which have poles for $x = y$ if $\Theta \in \,]0,1[$ and $d > 0$. These distributions incorporate biologically desirable properties, such as strong leptokurticity, and have been used, e.g., in [**10**]. However, kernels as the above may be approximated arbitrarily well by continuous ones in the sense that the norm of the difference operator—and thus the difference of the largest eigenvalues—gets arbitrarily small. Then, the procedures described here may be applied to these continuous kernels.

6. Outlook

This article shows that most reasonable COA models can be approximated arbitrarily well by models with discrete types. Therefore, one can expect both model classes to behave quite similarly. For certain mutation–selection models with discrete types, a simple maximum principle for the equilibrium mean fitness λ was recently found [**8**] (see also [**7**, **6**, **2**]). It takes the form

$$\lambda \simeq \sup_{x \in I} \big(r(x) - g(x) \big)$$

and holds as an exact identity in a limit of infinitely many types that densely fill a compact interval I. In the simplest case, a linear ordering of types is assumed and mutation is taken to only connect every type x with its two neighbors at rates $u^{\pm}(x)$. Then, the function g is given as $g(x) = u^{+}(x) + u^{-}(x) - 2\sqrt{u^{+}(x)\, u^{-}(x)}$. In a subsequent analysis [**6**], models with three types of mutation—and hence six neighbors of every type—were considered. For these, g is given as the sum of three terms of the above pattern (and x has three components), one for each type of mutation.

In the light of the findings presented here, one may conjecture that also for certain COA models the above characterization is valid with an appropriate function g. First steps in [**17**], both analytical and numerical, corroborate this conjecture with

$$g(x) = \int_I \Big(u(x,y) - \sqrt{u(x,y)\, u(y,x)} \Big)\, \mathrm{d}y \,,$$

which generalizes the additive structure of g found in [**6**] with respect to a continuum of possible mutations. The important prerequisite seems to be the possibility to approximate every local subsystem, corresponding to a small interval $J \subset I$, by a COA model whose mutation kernel is of the form $u(x,y) = \exp(\gamma\,(x-y))\, h(|x-y|)$. Then, in a limit $\nu \to \infty$, where h is replaced by $h_\nu(|x-y|) = \nu\, h(\nu\, |x-y|)$, the above expression seems to become exact. A rigorous proof for this statement seems feasible in the near future.

Acknowledgements. I thank Michael Baake, Ellen Baake, Tini Garske, Joachim Hermisson, and Reinhard Bürger for discussions and helpful comments on the manuscript. Support by a PhD scholarship of the Studienstiftung des deutschen Volkes is gratefully acknowledged. Furthermore, I express my gratitude to the Erwin Schrödinger International Institute for Mathematical Physics in Vienna for support during a stay in December 2002, where the manuscript for this article was completed. Many thanks also go to an anonymous referee for hints at how to simplify the main proofs and improve the presentation.

References

1. P. M. Anselone, *Collectively Compact Operator Approximation Theory and Applications to Integral Equations* (Prentice-Hall, Englewood Cliffs, 1971).
2. E. Baake, M. Baake, A. Bovier and M. Klein, An asymptotic maximum principle for essentially linear evolution models, Preprint q-bio.PE/0311020.
3. R. Bürger, Perturbations of positive semigroups and applications to population genetics, *Math. Z.* **197** (1988), 259–272.
4. R. Bürger, *The Mathematical Theory of Selection, Recombination, and Mutation* (Wiley, Chichester, 2000).
5. H. W. Engl, *Integralgleichungen* (Springer, Wien, 1997).
6. T. Garske and U. Grimm, A maximum principle for the mutation–selection equilibrium of nucleotide sequences, to appear in *Bull. Math. Bio.*
7. H.-O. Georgii and E. Baake, Supercritical multitype branching processes: the ancestral types of typical individuals, *Adv. Appl. Prob.* **35** (2003), 1090–1110.
8. J. Hermisson, O. Redner, H. Wagner and E. Baake, Mutation–selection balance: Ancestry, load, and maximum principle, *Theor. Pop. Biol.* **62** (2002), 9–46.
9. E. Hewitt and K. Stromberg, *Real and Abstract Analysis* (Springer, Berlin, 1969).
10. W. G. Hill, Predictions of response to artificial selection from new mutations, *Genet. Res.* **40** (1982), 255–278.
11. K. Jörgens, *Lineare Integraloperatoren* (Teubner, Stuttgart, 1970).
12. M. Kimura, A stochastic model concerning the maintenance of genetic variability in quantitative characters, *Proc. Natl. Acad. Sci. U.S.A.* **54** (1965), 731–736.
13. M. A. Krasnosel'skii, G. M. Vainikko, P. P. Zabreiko, Y. B. Rutitskii and V. Y. Stetsenko, *Approximate Solution of Operator Equations* (Wolters-Noordhoff, Groningen, 1972).
14. R. Kress, *Linear Integral Equations*, 2nd ed. (Springer, Heidelberg, 1999).
15. S. Lang, *Real and Functional Analysis*, 3rd ed. (Springer, New York, 1993).
16. K. E. Morrison, Spectral approximation of multiplication operators, *New York J. Math.* **1** (1995), 75–96.
17. O. Redner, *Models for Mutation, Selection, and Recombination in Infinite Populations*, PhD thesis, Univ. Greifswald (Shaker, Aachen, 2003).
18. W. Rudin, *Real and Complex Analysis*, 3rd ed. (McGraw-Hill, New York, 1986).
19. W. Rudin, *Functional Analysis*, 2nd ed. (McGraw-Hill, New York, 1991).
20. H. H. Schaefer, *Banach Lattices and Positive Operators* (Springer, Berlin, 1974).
21. D. Werner, *Funktionalanalysis*, 3rd ed. (Springer, Berlin, 2000).