Check for
updates

Study the genetic variation using Eta functions

S. Mashayekhi[1] · **S. Sedaghat**[2]

Abstract

This paper studies the genetic variation within species using the Eta base functions. We consider the House of Cards Kingman's model to study genetic variation. This model analyzes the balance between mutation and selection within species, while other forces that produce and maintain genetic variation cause only perturbations. Since this model is a nonlinear integral equation, we introduce a new numerical method for solving the nonlinear mixed Volterra–Fredholm integral equations. We find the best approximation of unknown functions to solve the integral equation using the Eta functions and collocation method. We derive the error bounds of the numerical method and solve some numerical examples to show the high accuracy of the new numerical technique. Using this numerical method, we study the behavior of the probability density of genes within species by considering three different cases of fitnesses for individuals.

Keywords Genetic variation · House of Cards Kingman's model · Eta functions · Best approximation · Mixed Volterra–Fredholm integral equations

Mathematics Subject Classification 9208

1 Introduction

Genetics study genetic variation and heredity in population (Griffiths et al. 2005). Genes consist of deoxyribonucleic acid (DNA) that contains the code, or blueprint, used to synthesize a protein. Population genetics aims to understand the forces that produce and maintain genetic variation within species. These forces include mutation, recombination, natural selection, population structure, and the random transmission of genetic material from parents to offspring (Wakeley 2009). A Mutation, changing in DNA sequence, may be transmitted to descendants by DNA replication, resulting in a sector or patch of cells having an abnormal

✉ S. Mashayekhi
   smashayekhi@fsu.edu; smashay1@kennesaw.edu

1   Department of Mathematics, Kennesaw State University, Marietta, GA 30060, Georgia

2   Imam Khomeini International University-Buein Zahra Higher Education Center of Engineering and Technology, Qazvin, Iran

∅ Springer ⨌BMAC

function, such as cancer. Based on Natural selection, organisms more adapted to their environment are more likely to survive. This process causes species to change and diverge over time. (Kingman 1978) considered the equilibrium of a population as existing because of a balance between two factors, mutation, and selection, while other phenomena cause only perturbations. In this model, a nonlinear integral equation, an individual with a larger type value is more productive. Kingman's model have been studied in different literatures, see for example Bürger (Bürger 1986, 1989, 1998, 2000; Steinsaltz et al. 2005; Evans et al. 2013; Yuan 2017, 2020). We study the properties of this model using a new numerical model developed for Volterra-Fredholm integral equations.

There are different types of integral equations, Volterra integral equations, Fredholm integral equations and Volterra–Fredholm integral equations (Liu et al. 2020; Amin et al. 2020).

Since nonlinear integral equations are usually difficult to get their exact solution, many authors have worked on analytical methods, and numerical methods for the solution of this kind of equation (Binh and Ninh 2019). Some of these methods are the successive approximations method (Chen et al. 1997; Tricomi 1985; Wazwaz 2011), Newton-Kantorovich method (Chen et al. 1997), Adomian decomposition method (Wazwaz 2011), Homotopy analysis method (Liao 2003), iterative numerical method (Ziari and Bica 2019).

In this paper, to study the genetic variation using the Kingman model, we introduce a new numerical method for the nonlinear Volterra–Fredholm integral equations. The technique relies on using the Eta functions (Mashayekhi and Ixaru 2020).

Eta functions have been introduced by (Ixaru 1984) and have been used in developing the new numerical methods in several literatures. For example, see (Ixaru 1997; Liviu Gr and Guido Vanden 2004; Cardone et al. 2010; Coleman and Ixaru 2006; Kim et al. 2002, 2003; Ixaru 2002; Conte et al. 2010) and references therein.

Recently, Mashayekhi and Ixaru (2020) have used the Eta functions to find the best approximation of a function. The paper results show the advantages of using the Eta function compared to sets of orthogonal or nonorthogonal functions for finding the best approximation of a high oscillatory function.

The present paper introduces a new direct computational method for solving the nonlinear Volterra–Fredholm integral equations. This method uses the Eta-based functions as base functions to find the best approximation of the solution of the integral equations. An essential property of the Eta-based functions is that they tend to the polynomial when the involved frequencies tend to zero. Thus, the Eta-based functions are suitable for attaining a good approximation of high oscillatory functions and polynomials. This property brings the excellent opportunity to approximate the solution of the dynamical systems when we do not know the behavior of the exact solution. In this paper, we introduce the best approximation of the unknown function using Eta-based functions and convert the solution of the nonlinear Volterra–Fredholm integral equations to the solution of a set of nonlinear equations using the collocation method. We derive the error bounds of the numerical method and show the high accuracy of the new numerical technique by solving some numerical examples.

Using the new numerical method, we study the properties of the House of Cards Kingman's model. We consider three different cases for the fitness of individuals with different values for the mutation rate to study the behavior of the probability density of genes within species. Since the exact solution, in this case, is unknown, using the Eta-based functions allow us to consider all possible situation for the exact solution, which includes trigonometric function, hyperbolic functions, or polynomials.

The paper is organized as follows: Sect. 2 introduces Kingman's model. Section 3 presents the new numerical method for the nonlinear Volterra–Fredholm integral equations. In Sect. 4,

we study the genetic variation based on Kingman's model using the new numerical method. In Sect. 5, we report the conclusion.

2 The balance between selection and mutation using Kingman's model

In this section, we describe the House of Cards Kingman's model, which is a nonlinear integral equation (Kingman 1980). To study the properties of a gene, we use two terms, allele and locus. While locus is the position of an allele in the chromosome, an allele is a version of a gene. Let consider a single locus and suppose that the possible alleles at this locus are listed as $A_1, A_2, \ldots, A_k$. Suppose that w_{ij} are the fitnesses, which are the probability of an individual with genotype $A_i A_j$ to survive and reproduce, and have the following form

$$w_{ij} = F(\xi_i, \xi_j, \varepsilon_{ij}), \tag{1}$$

for some function $F : (0, 1)^3 \to [0, \infty)$ which is symmetric in its first two arguments. Suppose that in a particular generation the ξ_i values of the genes in the gamete pool have an empirical distribution approximated by a probability density $\rho(t)$, $\quad 0 < t < 1$. The individuals in the next generation inherit a ξ_i value from each gamete, and so may be describe by a pair of ξ_i values (ξ_i, ξ_j) whose joint distribution is $\rho(t)\rho(s)$. Individuals with this description have on average a fitness $K(\xi_i, \xi_j)$ which by Eq. (1) and based on the definition of marginal probability is given by

$$K(t, s) = \int_0^1 F(t, s, z)dz, \tag{2}$$

so that the joint density in the mature population is

$$\frac{\rho(t)\rho(s)K(t, s)}{W}, \tag{3}$$

where

$$W = \int_0^1 \int_0^1 \rho(t)\rho(s)K(t, s)dtds. \tag{4}$$

Thus, without mutation, the ξ_i values in the next gamete pool would have density

$$\rho(t) = \int_0^1 W^{-1}\rho(t)\rho(s)K(t, s)ds. \tag{5}$$

If a mutation happens at rate μ then mutation replaces this by $(1 - \mu)\rho(t) + \mu$ so the ξ_i values in the next gamete pool would have density

$$\rho(t) = (1 - \mu) \int_0^1 W^{-1}\rho(t)\rho(s)K(t, s)ds + \mu. \tag{6}$$

Let assume $K^*(t, s, \rho(s)) = W^{-1}\rho(t)\rho(s)K(t, s)$ then we can rewrite Eq. (6) as

$$\rho(t) = (1 - \mu) \int_0^1 K^*(t, s, \rho(s))ds + \mu.$$

Since $K^*(t, s, \rho(s))$ depends on W^{-1}, Eq. (6) is a nonlinear Fredholm integral equation determining $\rho(t)$ in terms of K and μ. In the next section, we develop the numerical method to study the properties of Eq. (6). Although Eq. (6) is a Fredholm integral equation, we solve more general cases, which include Fredholm and Volterra integral equations, in the next section to show the advantages of the new numerical method for solving more general cases.

3 Integral equation

In this section, we develop a new numerical method based on the Eta functions for solving the system of Fredholm and Volterra integral equations

$$P(t) = G(t) + \lambda_1 \int_0^t K_1(t, s, P(s))ds + \lambda_2 \int_0^1 K_2(t, s, P(s))ds, \quad 0 \le t, s \le 1, \quad (7)$$

where

$$
\begin{aligned}
P(t) &= [\rho_1(t), \ldots, \rho_n(t)]^T, \\
G(t) &= [g_1(t), \ldots, g_n(t)]^T, \\
K_1(t, s, P(s)) &= [\kappa_1^1(t, s, P(s)), \ldots, \kappa_n^1(t, s, P(s))]^T, \\
K_2(t, s, P(s)) &= [\kappa_1^2(t, s, P(s)), \ldots, \kappa_n^2(t, s, P(s))]^T
\end{aligned}
$$

and λ_1 and λ_2 are constant vectors. It needs to mention Eq. (7) is a general form of Eq. (6), and we will study the properties of Eq. (6) using the new numerical method developed in this section. For more details please see Sect. 4.

3.1 Best approximation

Suppose $f(t) \in L^2[0, 1]$ and

$$f_{M+1}^F(t) = F^T D = \sum_{i=1}^{M+1} d_i f_i(t), \quad (8)$$

is a best approximation to f out of F where

$$F(t) = [f_1(t), f_2(t), \ldots, f_{M+1}(t)]^T, \qquad D = [d_1, d_2, \ldots, d_{M+1}]^T,$$

are the base functions and coefficients vector, respectively. The details of the best approximation have been presented on Mashayekhi and Ixaru (2020). The base functions have a crucial role in deriving the best approximation. This paper aims to show the advantages of using the Eta function as a based function for finding the best approximation to solve integral equations.

Eta functions $\eta_m(Z)$ with $m > 0$ and $Z \ne 0$ are generated by recurrence (Ixaru 1984, 1997):

$$\eta_m(Z) = \frac{\eta_{m-2}(Z) - (2m - 1)\eta_{m-1}(Z)}{Z}, \quad m = 1, 2, 3, \ldots \quad (9)$$

where

$$
\eta_{-1}(Z) = \begin{cases} \cos(|Z|^{\frac{1}{2}}) & Z \le 0 \\ \cosh(Z^{\frac{1}{2}}) & Z > 0, \end{cases}
\qquad
\eta_0(Z) = \begin{cases} \frac{\sin(|Z|^{\frac{1}{2}})}{|Z|^{\frac{1}{2}}} & Z < 0 \\ 1 & Z = 0 \\ \frac{\sinh(Z^{\frac{1}{2}})}{Z^{\frac{1}{2}}} & Z > 0, \end{cases}
$$

For $Z = 0$ these functions have the following values

$$\eta_m(0) = \frac{1}{(2m + 1)!!}, \quad m = 1, 2, \ldots$$

where !! is a double factorial (Ixaru 1984). These functions are used to introduce the set of Eta-based functions in the following way:

$$\phi_m(t) = t^{m-1}\eta_{\lfloor \frac{m}{2} \rfloor - 1}(Z(t)), \ m = 1, 2, \cdots \tag{10}$$

where $\lfloor \frac{m}{2} \rfloor$ is the integer part of $\frac{m}{2}$, and $Z(t) = -\omega^2 t^2$ in the trigonometric case and $Z(t) = \omega^2 t^2$ in the hyperbolic case. An important property of the Eta-based functions is that they tend to power functions when $\omega = 0$ (For more details please see (Ixaru 1997)).

In this paper, to show the advantages of Eta based functions, we consider three different choices for the set of base functions $F(t)$ in Eq. (8):

- We choose Eta-based functions as a base. In this case $F(t) = [\phi_1(t), \phi_2(t), \ldots, \phi_{M+1}(t)]$ is defined on $t \in [0, 1]$ where $\phi_i(t)$ has been introduced in Eq. (10).
- We choose Legendre polynomials as a base. These polynomials are orthogonal on the interval $[-1, 1]$ and have the recurrence relation (Kirkwood 2018)

$$(m + 1)L_{m+1}(t) = (2m + 1)t L_m(t) - m L_{m-1}(t),$$

where $L_0(t) = 1$ and $L_1(t) = t$. In this case

$$F(t) = [L_0(2t - 1), L_1(2t - 1), \ldots, L_M(2t - 1)]$$

is defined on $t \in [0, 1]$.
- We choose $F(t) = [\psi_0(t), \psi_1(t), \ldots, \psi_M(t)]$ as a base where

$$\psi_i(t) = \begin{cases} cos(i \times t), & \text{if i is even} \\ sin(i \times t), & \text{if i is odd} \end{cases}$$

and $t \in [0, 1]$. In some specific cases, we could consider only $sin(i \times t)$ or $cos(i \times t)$ as the base.

It has been shown that the Eta-based functions are the best candidate as a base to find the approximation of trigonometric and hyperbolic functions (Mashayekhi and Ixaru 2020).

3.2 Numerical technique

In this section, we develop a new numerical technique for solving the system of Fredholm and Volterra integral equations in Eq. (7).

Using Eq. (8) the best approximation of $\rho_i(t)$ in Eq. (7) is

$$\rho_i(t) = F^T(t)D_i, \tag{11}$$

and

$$P(t) = \hat{F}(t)\hat{D}, \tag{12}$$

where $\hat{D}$ is a $n(M + 1) \times 1$ vector given by

$$\hat{D} = [D_1, D_2, \ldots, D_n]^T, \tag{13}$$

and

$$\hat{F}(t) = I_n \otimes F^T(t), \tag{14}$$

where I_n is the n dimensional identity matrix. Also, $\hat{F}(t)$ is $n \times n(M + 1)$ matrix and $\otimes$ shows Kronecker product (Mashayekhi and Sedaghat 2021). Replacing Eqs. (12) in (7), we

have

$$\hat{F}(t)\hat{D} = G(t) + \lambda_1 \int_0^t K_1(t, s, \hat{F}(s)\hat{D})ds + \lambda_2 \int_0^1 K_2(t, s, \hat{F}(s)\hat{D})ds, \quad 0 \le t, s \le 1,$$
(15)

Using the Gauss-Legendre numerical integration for evaluating the integral in Eq. (15), we get

$$\hat{F}(t)\hat{D} = G(t) + \lambda_1 \sum_{i=0}^p \omega_i K_1\left(t, \frac{t}{2} + \frac{t}{2}\gamma_i, \hat{F}\left(\frac{t}{2} + \frac{t}{2}\gamma_i\right)\hat{D}\right)$$

$$+\lambda_2 \sum_{i=0}^p \omega_i K_2\left(t, \frac{1}{2} + \frac{1}{2}\gamma_i, \hat{F}\left(\frac{1}{2} + \frac{1}{2}\gamma_i\right)\hat{D}\right), \quad 0 \le t \le 1, \quad (16)$$

where ω_i and γ_i are weights and nods of Gauss–Legendre given in Smith (1965). Using Eq. (16), we introduce the residual of the problem as

$$R(t, \hat{D}) = \hat{F}(t)\hat{D} - G(t) - \lambda_1 \sum_{i=0}^p \omega_i K_1\left(t, \frac{t}{2} + \frac{t}{2}\gamma_i, \hat{F}\left(\frac{t}{2} + \frac{t}{2}\gamma_i\right)\hat{D}\right)$$

$$-\lambda_2 \sum_{i=0}^p \omega_i K_2\left(t, \frac{1}{2} + \frac{1}{2}\gamma_i, \hat{F}\left(\frac{1}{2} + \frac{1}{2}\gamma_i\right)\hat{D}\right), \quad (17)$$

and collocate Eq. (17) at the extreme points of the Chebyshev polynomial (Mashayekhi and Sedaghat 2021) to get $n(M + 1)$ nonlinear equations which can be solved for the elements of $\hat{D}$. We use Newton's iterative method to solve the nonlinear equations for the elements of $\hat{D}$. Finally, we calculate $P(t)$ given in Eq. (12).

Remark To apply Newton's iterative method, in the first stage, we set $M = 1$ and then use Newton's iterative process with a random initial condition. For $M = 1$, the equation has few terms, and the random initial condition gives us some results. Next, we increase the value of M until a satisfactory convergence is achieved. We use the approximate solution in the first stage as our initial guess of this stage. We continue this approach until the results are similar to a required number of decimal places for two consecutive M values.

3.3 Best approximation errors

In this section, we derive the error bound of the numerical method presented in Sect. 3.2. For the sake of simplicity, but without any loss of generality, we describe convergence analysis for $n = 1$ and $\rho_1 = \rho$. First, review some relevant properties of hypergeometric functions, Sobolev space, and best approximation using Legendre polynomials.

Generalized hypergeometric functions. For any real a_i, b_j with $b_j \ne 0, -1, -2, \dots$ $(i = 1, \dots, p; \ j = 1, \dots, q)$, the generalized hypergeometric function is a power series defined by Kilbas (2006)

$$_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; Z) = \sum_{k=0}^\infty \frac{(a_1)_k \dots (a_p)_k}{(b_1)_k \dots (b_q)_k} \frac{Z^k}{k!},$$
(18)

where the Pochhammer or ascending factorial symbol for $a \in \mathbb{R} - \{0\}$ is defined as

$$(a)_0 = 1, \quad (a)_k = a(a+1)(a+2)\dots(a+k-1) = \Gamma(a+k)/\Gamma(a), \quad k \in \mathbb{N}.$$

Here, $\Gamma(.)$ is the usual Gamma function. It is known that this series is absolutely convergent for all values of Z if $p \leq q$, for $p = q + 1$, the series converges for $|Z| < 1$ and for $|Z| = 1$ under some additional conditions. If $p > q + 1$, the series is divergent.

Products of two hypergeometric functions $_0F_1(\gamma; Z)_0F_1(\delta; Z)$ is obtained as Bateman (1953),

$$_0F_1(\gamma; Z)_0F_1(\delta; Z) = {}_2F_3\left(\frac{1}{2}\gamma + \frac{1}{2}\delta, \frac{1}{2}\gamma + \frac{1}{2}\delta - \frac{1}{2}; \gamma, \delta, \gamma + \delta - 1; 4Z\right). \quad (19)$$

Now, we recall the following lemma to obtain the presentation of Eta-based functions in terms of the hypergeometric functions.

Lemma 1 *Legendre duplication formula is defined as Kilbas (2006)*

$$\Gamma(2m) = \frac{2^{2m-1}}{\sqrt{\pi}}\Gamma(m)\Gamma\left(m + \frac{1}{2}\right), \quad Re\, m > 0.$$

Theorem 1 *The Eta-based functions $\phi_m(t)$ can be represented in terms of the hypergeometric function*

$$\phi_m(t) = \frac{2^{-\lfloor\frac{m}{2}\rfloor}\sqrt{\pi}}{\Gamma(\lfloor\frac{m}{2}\rfloor + \frac{1}{2})}t^{m-1}{}_0F_1\left(\lfloor\frac{m}{2}\rfloor + \frac{1}{2}; \frac{Z}{4}\right), \quad (20)$$

where $Z = -\omega^2 t^2$ in the trigonometric case and $Z = \omega^2 t^2$ in the hyperbolic case.

Proof Series expansion of $\eta_m(Z)$ satisfy the following relation Ixaru (1997):

$$\eta_m(Z) = 2^m \sum_{q=0}^{\infty} \frac{(q+m)!}{q!(2q + 2m + 1)!}Z^q, \quad m = 0, 1, \ldots \quad (21)$$

from Eqs. (10) and (21) and by using Legendre duplication formula we obtain

$$\phi_m(t) = 2^{\lfloor\frac{m}{2}\rfloor - 1}t^{m-1}\sum_{k=0}^{\infty}\frac{\Gamma(k+\lfloor\frac{m}{2}\rfloor)}{\Gamma(2k+2\lfloor\frac{m}{2}\rfloor)}\frac{Z^k}{k!} = 2^{\lfloor\frac{m}{2}\rfloor - 1}t^{m-1}\sum_{k=0}^{\infty}\frac{\sqrt{\pi}2^{1-2k-2\lfloor\frac{m}{2}\rfloor}}{\Gamma(k+\lfloor\frac{m}{2}\rfloor+\frac{1}{2})}\frac{Z^k}{k!}$$

$$= \frac{2^{-\lfloor\frac{m}{2}\rfloor}\sqrt{\pi}}{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})}t^{m-1}\sum_{k=0}^{\infty}\frac{1}{(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})_k}\frac{(\frac{Z}{4})^k}{k!} = \frac{2^{-\lfloor\frac{m}{2}\rfloor}\sqrt{\pi}}{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})}t^{m-1}{}_0F_1(\lfloor\frac{m}{2}\rfloor + \frac{1}{2}; \frac{Z}{4}).$$

Sobolev Norm. The Sobolev norm of integer order $r \geq 0$ in the interval $(0, 1)$, is given by

$$\| \rho \|_{H^r(0,1)} = \left(\sum_{j=0}^{r}\int_0^1 |\rho^{(j)}(t)|^2 w(t)dt\right)^{\frac{1}{2}} = \left(\sum_{j=0}^{r}\| \rho^{(j)} \|_{L^2(0,1)}^2\right)^{\frac{1}{2}}, \quad (22)$$

where $\rho^{(j)}$ denotes the derivative of ρ of order j and $H^r(0, 1)$ is a Sobolev space. □

Theorem 2 *Suppose that $\rho \in H^r(0, 1)$ with $r \geq 0$, and $L_m(2t - 1)$, are the well-known shifted Legendre polynomials defined on the interval $[0, 1]$. Assume that*

$$\sum_{m=0}^{M} d_m L_m(2t - 1) \in \Pi_M$$

denotes the best approximation of ρ out of shifted Legendre polynomials, where Π_M is the space of all polynomials of degree less than or equal to M. Then we have Mashayekhi and Sedaghat (2021)

$$\left\| \rho - \sum_{m=0}^{M} d_m L_m(2t-1) \right\|_{L^2(0,1)} \leq cM^{-r} |\rho|_{H^{r;M}(0,1)}, \tag{23}$$

where

$$|\rho|_{H^{r;M}(0,1)} = \left(\sum_{i=\min\{r,M+1\}}^{r} \left\| \rho^{(i)} \right\|_2^2 \right)^{\frac{1}{2}},$$

and $c > 0$ is a constant which is independent of M.

In the following theorem, we derive the error bound of the approximation using the Eta-based functions as a base.

Theorem 3 Suppose that $\rho \in H^r(0,1)$ with $r \geq 0$, and ϕ_m ($m = 1, \ldots, M+1$) are the Eta-based functions defined on the interval $[0,1]$. Assume that $\rho_M(t) = \sum_{m=1}^{M+1} d_m \phi_m(t)$ denotes the best approximation of ρ out of Eta-based functions. Then we have

$$\left\| \rho - \sum_{m=1}^{M+1} d_m \phi_m(t) \right\|_{L^2(0,1)} \leq cM^{-r} |\rho|_{H^{r;M}(0,1)} + \sum_{m=1}^{M+1} \frac{2^{-\lfloor \frac{m}{2} \rfloor} \sqrt{\pi}}{\Gamma(\lfloor \frac{m}{2} \rfloor + \frac{1}{2})} \varepsilon_m |d_m|. \tag{24}$$

Proof Using Theorem 1, we have

$$\| \rho(t) - \rho_{M+1}(t) \|_{L^2(0,1)} = \left\| \rho - \sum_{m=1}^{M+1} d_m \phi_m(t) \right\|_{L^2(0,1)} = \left\| \rho(t) - \sum_{m=1}^{M+1} \varsigma_m \sum_{k=0}^{\infty} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} \right\|_{L^2(0,1)}$$

$$= \left\| \rho(t) - \sum_{m=1}^{M+1} \varsigma_m \left(\sum_{k=0}^{N} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} + \sum_{k=N+1}^{\infty} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} \right) \right\|_{L^2(0,1)}$$

$$\leq \left\| \rho(t) - \sum_{m=1}^{M+1} \varsigma_m \sum_{k=0}^{N} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} \right\|_{L^2(0,1)} + \left\| \sum_{m=1}^{M+1} \varsigma_m \sum_{k=N+1}^{\infty} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} \right\|_{L^2(0,1)},$$

where

$$\varsigma_m = d_m \frac{2^{-\lfloor \frac{m}{2} \rfloor} \sqrt{\pi}}{\Gamma(\lfloor \frac{m}{2} \rfloor + \frac{1}{2})}.$$

Since the best approximation of a given function $\rho \in H^r(0,1)$ is unique, using Theorem 2 we get

$$\left\| \rho(t) - \sum_{m=1}^{M+1} \sum_{k=0}^{N} \frac{(\mp \omega^2)^k \varsigma_m}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} \right\|_{L^2(0,1)} = \left\| \rho(t) - \sum_{m=0}^{M} d_m L_m(2t-1) \right\|_{L^2(0,1)}$$

$$\leq cM^{-r} |\rho|_{H^{r;M}(0,1)}. \tag{25}$$

Since the series $\sum_{k=0}^{\infty} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!}$ is convergent, for each m we have ε_m as follows

$$\left| \sum_{k=N+1}^{\infty} \frac{(\mp \omega^2)^k}{\left(\lfloor \frac{m}{2} \rfloor + \frac{1}{2} \right)_k} \frac{t^{2k+m-1}}{4^k k!} \right| < \varepsilon_m, \tag{26}$$

using Eq. (26) for all values of $|t| < 1$, we have

$$\left\| \sum_{m=1}^{M+1} \varsigma_m \sum_{k=N+1}^{\infty} \frac{(\mp\omega^2)^k}{\left(\lfloor\frac{m}{2}\rfloor+\frac{1}{2}\right)_k} \frac{t^{2k+m-1}}{4^k k!} \right\|_{L^2(0,1)} \leq \sum_{m=1}^{M+1} \left(|\varsigma_m| \left\| \sum_{k=N+1}^{\infty} \frac{(\mp\omega^2)^k}{\left(\lfloor\frac{m}{2}\rfloor+\frac{1}{2}\right)_k} \frac{t^{2k+m-1}}{4^k k!} \right\|_{L^2(0,1)} \right)$$

$$\leq \sum_{m=1}^{M+1} \varepsilon_m |\varsigma_m| \leq \sum_{m=1}^{M+1} \frac{2^{-\lfloor\frac{m}{2}\rfloor}\sqrt{\pi}}{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})} \varepsilon |d_m|,$$

this completes the proof. $\qquad\square$

Now, Using Theorem 3, we derive the error bound for the numerical method presented in Sect. 3.2.

Theorem 4 *Suppose that $\rho \in H^r(0,1)$ with $r \geq 0$ is the exact solution of Eq. (7) and $\bar{\rho}_{M+1} = F^T \bar{D} = \sum_{m=1}^{M+1} \bar{d}_m \phi_m(t)$ is its approximation given by the method proposed in Sect. 3.2, then,*

$$\|\rho - \bar{\rho}_{M+1}\|_{L^2(0,1)} \leq cM^{-r}|\rho|_{H^{r;M}(0,1)} + \sum_{m=1}^{M+1} \frac{2^{-\lfloor\frac{m}{2}\rfloor}\sqrt{\pi}}{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})} \varepsilon_m |d_m|$$

$$+ \|D - \bar{D}\|_{L^2(0,1)} \left(\sum_{m=1}^{M+1} \frac{\pi 2^{-2\lfloor\frac{m}{2}\rfloor}}{(-1+2m)\{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})\}^2} {}_2F_3 \left(m-\frac{1}{2}, \lfloor\frac{m}{2}\rfloor; m+\frac{1}{2}, 2\lfloor\frac{m}{2}\rfloor, \lfloor\frac{m}{2}\rfloor+\frac{1}{2}; \mp w^2\right) \right)^{\frac{1}{2}},$$

where $\sum_{m=1}^{M+1} d_m \phi_m(t) = F^T D$ denotes the approximation of ρ using the set of Eta-based functions.

Proof Let assume $\rho_{M+1}(t) = F^T D = \sum_{m=1}^{M+1} d_m \phi_m(t)$, we have

$$\|\rho - \bar{\rho}_{M+1}\|_{L^2(0,1)} \leq \|\rho - \rho_{M+1}\|_{L^2(0,1)} + \|\rho_{M+1} - \bar{\rho}_{M+1}\|_{L^2(0,1)}. \tag{27}$$

Using Theorem 3, we have the upper bound for $\|\rho - \rho_{M+1}\|_{L^2(0,1)}$ so we need to find an upper bound for $\|\rho_{M+1} - \bar{\rho}_{M+1}\|_{L^2(0,1)}$. To do this, using the Schwarz'z inequality, Theorem 1 and Eq. (19) we obtain

$$\|\rho_{M+1} - \bar{\rho}_{M+1}\|_{L^2(0,1)}^2 = \left\| \sum_{m=1}^{M+1} d_m \phi_m(t) - \sum_{m=1}^{M+1} \bar{d}_m \phi_m(t) \right\|_{L^2(0,1)}^2 = \left\| \sum_{m=1}^{M+1} (d_m - \bar{d}_m)\phi_m(t) \right\|_{L^2(0,1)}^2$$

$$= \int_0^1 \left| \sum_{m=1}^{M+1} (d_m - \bar{d}_m)\phi_m(t) \right|^2 dt \leq \left(\sum_{m=1}^{M+1} |d_m - \bar{d}_m|^2 \right) \int_0^1 \sum_{m=1}^{M+1} |\phi_m(t)|^2 \, dt$$

$$= \left(\sum_{m=1}^{M+1} |d_m - \bar{d}_m|^2 \right) \int_0^1 \sum_{m=1}^{M+1} \frac{\pi 2^{-2\lfloor\frac{m}{2}\rfloor} t^{2m-2}}{\{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})\}^2} \left\{ {}_0F_1 \left(\lfloor\frac{m}{2}\rfloor+\frac{1}{2}; \frac{\mp w^2 t^2}{4}\right) \right\}^2 dt$$

$$= \left(\sum_{m=1}^{M+1} |d_m - \bar{d}_m|^2 \right) \int_0^1 \sum_{m=1}^{M+1} \frac{\pi 2^{-2\lfloor\frac{m}{2}\rfloor} t^{2m-2}}{\{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})\}^2} {}_2F_3 \left(\lfloor\frac{m}{2}\rfloor, \lfloor\frac{m}{2}\rfloor+\frac{1}{2}; \lfloor\frac{m}{2}\rfloor+\frac{1}{2}, 2\lfloor\frac{m}{2}\rfloor; \mp w^2 t^2\right) dt$$

$$= \left(\sum_{m=1}^{M+1} |d_m - \bar{d}_m|^2 \right) \sum_{m=1}^{M+1} \frac{\pi 2^{-2\lfloor\frac{m}{2}\rfloor}}{(-1+2m)\{\Gamma(\lfloor\frac{m}{2}\rfloor+\frac{1}{2})\}^2} {}_2F_3 \left(m-\frac{1}{2}, \lfloor\frac{m}{2}\rfloor; m+\frac{1}{2}, 2\lfloor\frac{m}{2}\rfloor, \lfloor\frac{m}{2}\rfloor+\frac{1}{2}; \mp w^2\right),$$

this completes the proof. $\qquad\square$

Remark In this section, to derive the error bound of the Eta-based function, we have used the Legendre polynomials, so the error bound has more terms than the error bound of using the Legendre polynomials as a base. We plan to derive the error bound of the Eta-based function directly in future work.

3.4 Numerical test

In this section, we use the method presented in Sect. 3.2 to solve seven examples to show the accuracy of the present method. In all examples, we use the extreme points of the Chebyshev polynomial as the collocation points (Mashayekhi and Sedaghat 2021) while we consider three different choices for the set of base functions: Eta-based function, Legendre polynomials, and trigonometric functions. In different examples, we consider a different form of non-linearity such as $\rho^2(s)$, $cos(\rho(s))$ and $e^{\rho(s)}$. The error presented for all examples is an absolute error. For all examples, we consider two different cases $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_1 = 1$, $\lambda_2 = 0$ (Fredholm and Volterra cases). Since the results for both cases are similar, we present one of them in the Tables. At the end of the examples, we analyze the results of these examples.

- **Example 1:** In this example, we assume $n = 1$, $K_2(t, s, \rho(s)) = k(t, s)\rho^2(s)$ where $k(t, s) = ts$ and the exact solution is $\rho(t) = cos(t)$. The absolute error is presented in Table 1. In this table, we choose the three first terms of the base for all three different choices of base functions.
- **Example 2:** In this example, we assume $n = 1$, $K_2(t, s, \rho(s)) = k(t, s)\rho^2(s)$ where $k(t, s) = tcos(s)$ or $k(t, s) = scos(t)$ and the exact solution is $\rho(t) = cos(t)$. The absolute error is presented in Table 2. In this table, we choose four first terms of the base for all three different choices of base functions.
- **Example 3:** In this example, we assume $n = 1$, $K_2(t, s, \rho(s)) = k(t, s)cos(\rho(s))$ where $k(t, s) = ts$ and the exact solution is $\rho(t) = cos(t)$. The absolute error is presented in Table 3. In this table, we choose four first terms of the base for all three different choices of base functions.

Table 1 Absolute error (Example 1)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	3.9×10^{-12}	5.5×10^{-4}	3.8×10^{-12}
0.4	8.1×10^{-12}	1.0×10^{-3}	7.7×10^{-12}
0.6	1.2×10^{-11}	4.5×10^{-4}	1.2×10^{-11}
0.8	1.6×10^{-11}	6.2×10^{-5}	1.6×10^{-11}

Table 2 Absolute error (Example 2)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	1.5×10^{-12}	1.0×10^{-4}	1.5×10^{-12}
0.4	2.9×10^{-12}	1.0×10^{-4}	3.3×10^{-12}
0.6	4.4×10^{-12}	1.4×10^{-4}	5.0×10^{-12}
0.8	5.8×10^{-12}	2.6×10^{-5}	6.4×10^{-12}

Table 3 Absolute error (Example 3)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	4.7×10^{-12}	1.4×10^{-4}	4.1×10^{-12}
0.4	9.1×10^{-12}	2.2×10^{-5}	9.2×10^{-12}
0.6	1.3×10^{-11}	1.9×10^{-5}	1.3×10^{-11}
0.8	1.8×10^{-11}	1.4×10^{-4}	1.7×10^{-11}

Table 4 Absolute error (Example 4)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	3.8×10^{-11}	1.1×10^{-4}	3.5×10^{-11}
0.4	7.5×10^{-11}	8.6×10^{-5}	7.9×10^{-11}
0.6	1.1×10^{-10}	1.1×10^{-4}	1.1×10^{-10}
0.8	1.5×10^{-10}	1.2×10^{-5}	1.5×10^{-10}

Table 5 Absolute error (Example 5)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	1.4×10^{-11}	5.6×10^{-4}	2.4×10^{-5}
0.4	2.9×10^{-11}	4.7×10^{-4}	7.2×10^{-5}
0.6	4.4×10^{-11}	6.4×10^{-4}	1.1×10^{-4}
0.8	5.9×10^{-11}	1.2×10^{-4}	3.8×10^{-5}

Table 6 Absolute error for $\rho_1(t)$ (Example 6)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	1.4×10^{-11}	5.6×10^{-4}	2.4×10^{-5}
0.4	2.9×10^{-11}	4.7×10^{-4}	7.2×10^{-5}
0.6	4.4×10^{-11}	6.4×10^{-4}	1.1×10^{-4}
0.8	5.9×10^{-11}	1.2×10^{-4}	3.8×10^{-5}

- **Example 4:** In this example, we assume $n = 1$, $K_2(t, s, \rho(s)) = k(t, s)e^{\rho(s)}$ where $k(t, s) = ts$ and the exact solution is $\rho(t) = cos(t)$. The absolute error is presented in Table 4. In this table, we choose four first terms of the base for all three different choices of base functions.
- **Example 5:** In this example, we assume $n = 1$, $K_2(t, s, \rho(s)) = k(t, s)\rho^2(s)$ where $k(t, s) = ts$ and the exact solution is $\rho(t) = tsinh(t)$. The absolute error is presented in Table 5. In this table, we choose four first terms of the base for all three different choices of base functions.
- **Example 6:** In this example, we assume $n = 2$, $\kappa_1^2(t, s, \rho_1(s), \rho_2(s)) = \rho_1(s) + \rho_2(s)$, and $\kappa_2^2(t, s, \rho_1(s), \rho_2(s)) = \rho_1(s) - \rho_2(s)$ with the exact solution $\rho_1(t) = tsinh(t)$, $\rho_2(t) = cosh(t)$. Tables 6 and 7 show the absolute error for this example.

Table 7 Absolute error for $\rho_2(t)$ (Example 6)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	1.0×10^{-11}	3.3×10^{-4}	1.6×10^{-6}
0.4	1.7×10^{-11}	9.4×10^{-5}	1.8×10^{-6}
0.6	2.6×10^{-11}	1.6×10^{-4}	6.2×10^{-7}
0.8	3.7×10^{-11}	3.6×10^{-4}	2.3×10^{-5}

Table 8 Absolute error (Example 7)

t	Eta-based functions	Legendre polynomials	Trigonometric functions
0.2	2.9×10^{-12}	4.0×10^{-5}	1.5×10^{-2}
0.4	5.9×10^{-12}	1.0×10^{-5}	3.4×10^{-3}
0.6	8.9×10^{-12}	2.3×10^{-5}	1.9×10^{-2}
0.8	1.1×10^{-11}	2.3×10^{-5}	2.7×10^{-2}

In these tables, we choose the four first terms of the base for all three different choices of base functions.

- **Example 7:** In this example, we assume $n = 1$, $K_2(t, s, \rho(s)) = k(t, s)\rho^2(s)$ where $k(t, s) = ts$ and the exact solution is $\rho(t) = t\cos(t)$. Table 8 shows the absolute error for this example. In this table, we choose the five first terms of the base for all three different choices of base functions.

These numerical examples show the Eta-based functions are much better than the Legendre polynomials when the exact solution of the integral equation is trigonometric functions. Also, in this case, trigonometric and Eta-based functions have the same accuracy. The Eta-based functions are much better than the Legendre polynomials and trigonometric functions when the exact solution of the integral equation is an exponential function or has one of the following forms

$$\Delta(t) = \delta_1(t)\sin(\omega t) + \delta_2(t)\cos(\omega t) \text{ or } \Delta(t) = \delta_1(t)\sinh(\omega t) + \delta_2(t)\cosh(\omega t),$$

where ω is a constant and $\Delta_1(t)$, $\Delta_2(t)$ are continuous functions. These results are consistent with the reported results in Mashayekhi and Ixaru (2020). The results of this section convince us to use Eta-based functions to study the House of Cards Kingman's model in Eq. (6). Since the exact solution of House of Cards Kingman's model is unknown, using the Eta-based functions allowed us to consider all possible situations for the exact solution, including a trigonometric function, hyperbolic functions, or polynomials.

4 Genetic variation using Kingman's model

In this section, we study the House of Cards Kingman's model in Eq. (6) using the numerical method introduced in Sect. 3.2. We consider three different cases for the fitness of individuals, $K(t, s)$, in Eq. (6) with different values for the mutation rate. Since we do not have an exact solution for Eq. (6), we have assumed a different format of the involved frequencies, $Z(t)$ in Eq. (10), to consider all possible cases including trigonometric function, hyperbolic
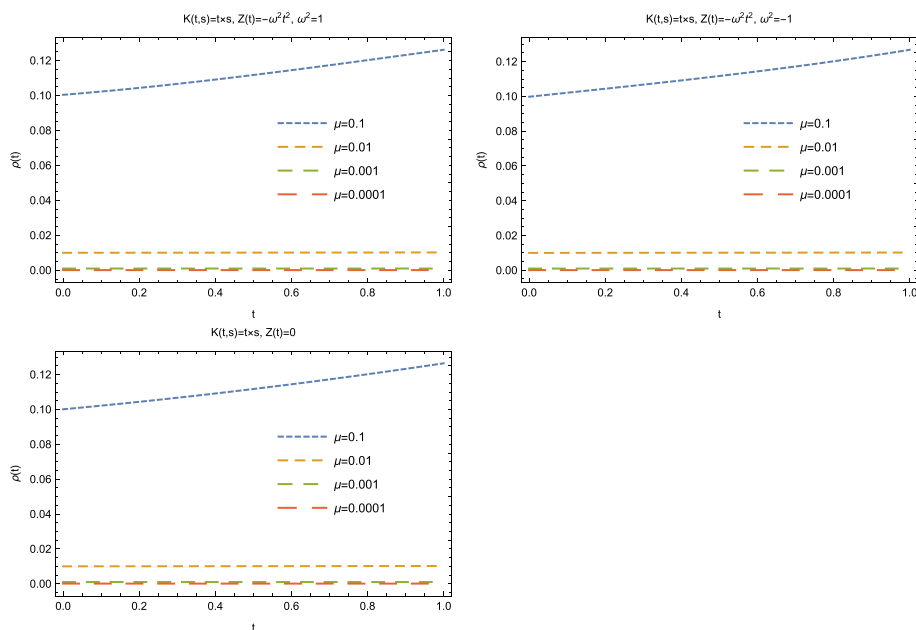
Fig. 1 The probability density of genes, $\rho(t)$, for $\mu < 1$ with $K(t,s) = t \times s$

functions, or polynomials. These three different cases include $Z(t) = t^2$, $Z(t) = -t^2$, and $Z(t) = 0$, which introduce hyperbolic, trigonometric, and polynomial functions, respectively. Figures 1, 2, and 3 show the probability density of genes within species, $\rho(t)$, for different values of $Z(t)$ and different values of fitnesses, $K(t,s)$, which includes $K(t,s) = t \times s$, $K(t,s) = cos(t) \times cos(s)$ and $K(t,s) = cosh(t) \times cosh(s)$ with the mutation rate less than one ($\mu < 1$). These figures show despite different choices for $K(t,s)$, the value of the probability density, $\rho(t)$, does not change dramatically over time. Figures 4, 5, and 6 show the probability density of genes within species, $\rho(t)$, for different values of $Z(t)$ and different values of fitnesses, $K(t,s)$, which includes $K(t,s) = t \times s$, $K(t,s) = cos(t) \times cos(s)$ and $K(t,s) = cosh(t) \times cosh(s)$ with the mutation rate greater than one ($\mu > 1$). These figures show the behavior of the probability density, $\rho(t)$, depends on the value of $K(t,s)$. These figures show for $K(t,s) = t \times s$, $K(t,s) = cosh(t) \times cosh(s)$ the value of the density, $\rho(t)$, decreases through time while the value of the density, $\rho(t)$, increases through time by choosing $K(t,s) = cos(t) \times cos(s)$. These results show the probability density of individuals within species does not depend on the fitness of individuals when the mutation rate is less than one. At the same time, this is not the case when the mutation rate is greater than one.
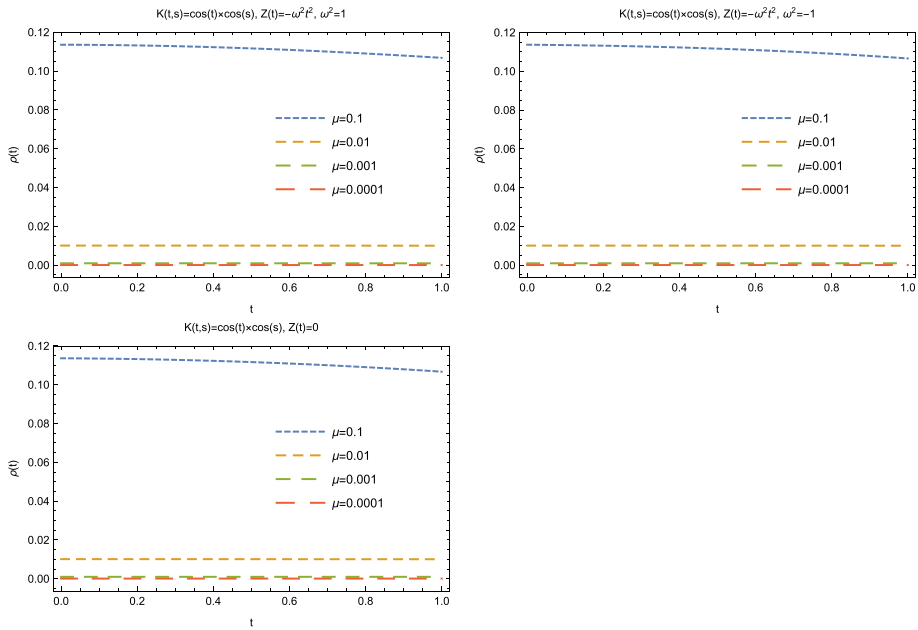
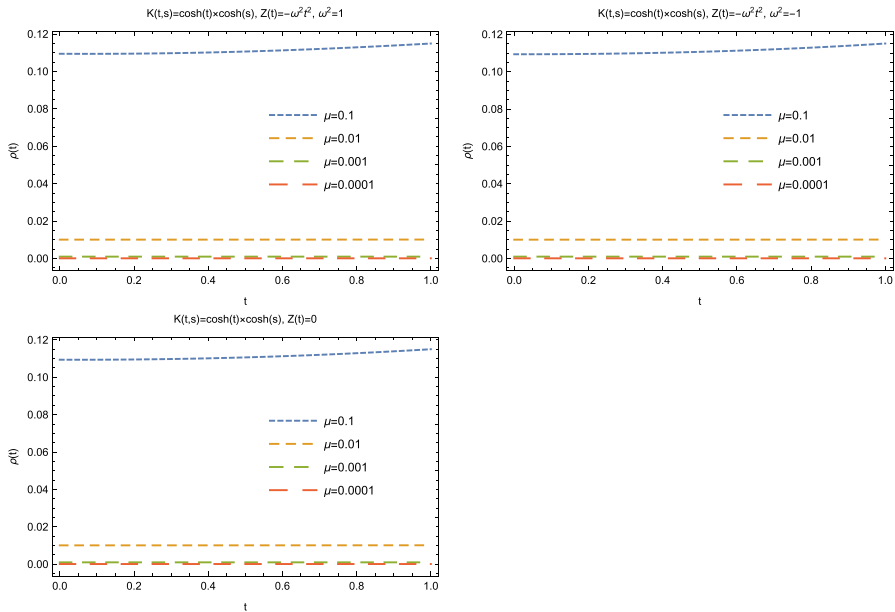**Fig. 2** The probability density of genes, $\rho(t)$, for $\mu < 1$ with $K(t,s) = cos(t) \times cos(s)$

Fig. 3 The probability density of genes, $\rho(t)$, for $\mu < 1$ with $K(t,s) = cosh(t) \times cosh(s)$
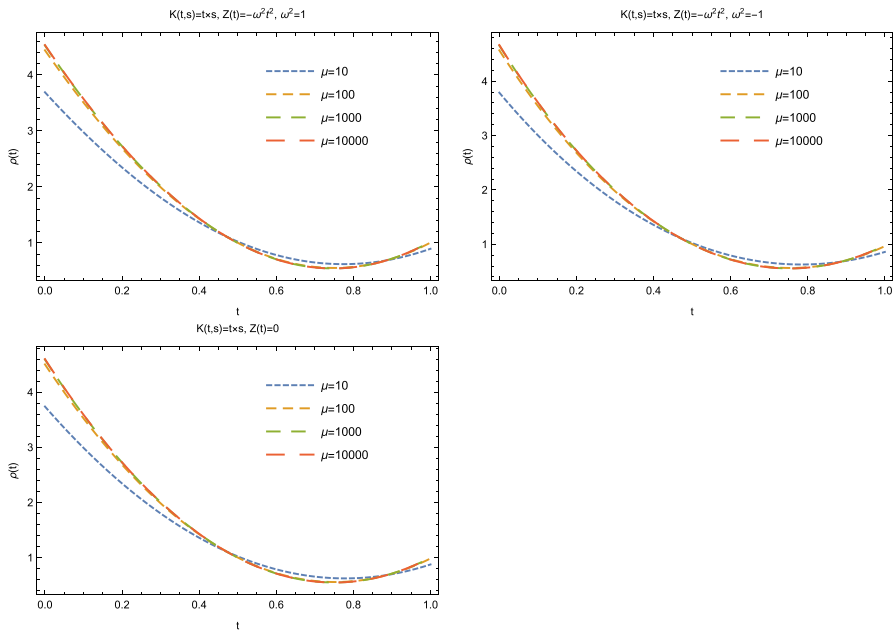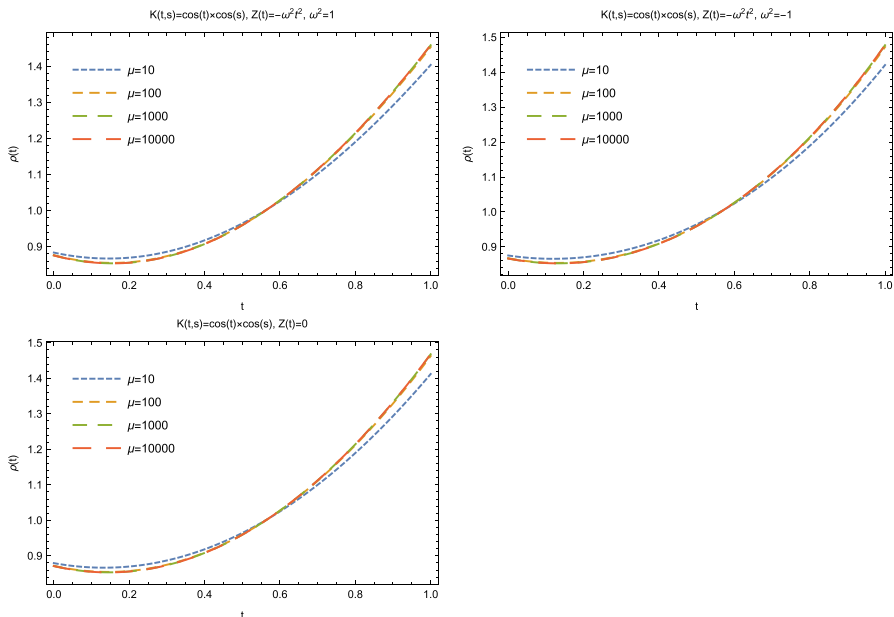
Fig. 4 The probability density of genes, $\rho(t)$, for $\mu > 1$ with $K(t, s) = t \times s$

Fig. 5 The probability density of genes, $\rho(t)$, for $\mu > 1$ with $K(t, s) = cos(t) \times cos(s)$
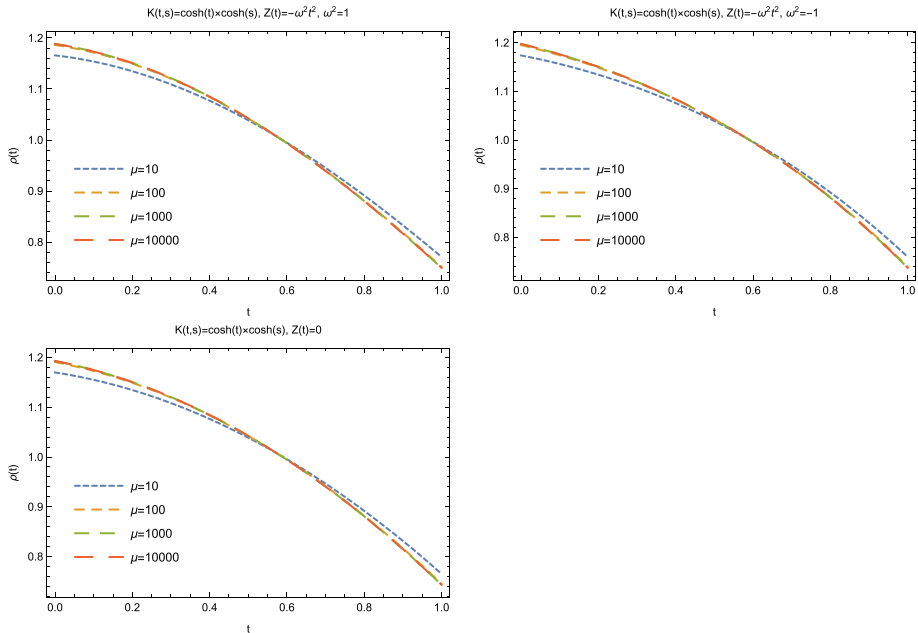
Fig. 6 The probability density of genes, $\rho(t)$, for $\mu > 1$ with $K(t, s) = cos(t) \times cos(s)$

5 Conclusion

We review the House of Cards Kingman's model using the new numerical method to study the genetic variation, which solves the nonlinear Volterra–Fredholm integral equations. The process is based on using the Eta-based functions. We used the best approximation and the extreme points of the Chebyshev polynomial as the collocation points. We derive the numerical method's error bounds and demonstrate the numerical technique's accuracy by solving some numerical tests. Since the exact solution for the House of Cards Kingman's model is unknown, using the Eta-based functions allowed us to consider all possible situations for the exact solution, including trigonometric function, hyperbolic functions, or polynomials. Using the numerical method, we show that the probability density of individuals within species depends on the mutation rate.

References

Abdul-Majid W (2011) Linear and nonlinear integral equations, vol 639. Springer
Amin R, Shah K, Asif M, Khan I (2020) Efficient numerical technique for solution of delay volterra–fredholm integral equations using haar wavelet. Heliyon 6(10):e05108
Anatoliĭ K (2006) Theory and applications of fractional differential equations
Bateman Harry (1953) Higher transcendental functions [volumes i–iii], vol 1. McGraw-Hill Book Company
Binh NT, Van Ninh K (2019) Parameter continuation method for solving nonlinear fredholm integral equations of the second kind. Bull Malaysian Math Sci Soc 42(6):3379–3407

Bürger Reinhard (2000) The mathematical theory of selection, recombination, and mutation. John Wiley & Sons

Bürger R (1986) On the maintenance of genetic variation: global analysis of kimura's continuum-of-alleles model. J Math Biol 24(3):341–351

Bürger R (1989) Mutation-selection balance and continuum-of-alleles models. Math Comput Modell 12(9):1181

Bürger R (1998) Mathematical properties of mutation-selection models. Genetica 102:279–298

Cardone A, Ixaru LG, Paternoster B (2010) Exponential fitting direct quadrature methods for volterra integral equations. Numer Algorithms 55(4):467–480

Coleman JP, Ixaru LG (2006) Truncation errors in exponential fitting for oscillatory problems. SIAM J Numer Anal 44(4):1441–1465

Conte D, Esposito E, Paternoster B, Ixaru LG (2010) Some new uses of the ηm (z) functions. Comput Phys Commun 181(1):128–137

Frank Charles KJ (1980) Mathematics of genetic diversity. SIAM

Giacomo TF (1985) Integral equations, vol 5. Courier corporation

Gr IL (1984) Numerical methods for differential equations and applications

Griffiths JF, Griffiths AJF, Wessler SR, Lewontin RC, Gelbart WM, Suzuki DT, Miller JH et al (2005) An introduction to genetic analysis. Macmillan

Guanrong C, Mingjun, C Zhongying C (1997) Approximate solutions of operator equations, vol 9. World Sci

Ixaru LG, Vanden Berghe G (2004) Exponential fitting, vol 568. Springer Science & Business Media

Ixaru L Gr (1997) Operations on oscillatory functions. Comput Phys Commun 105(1):1–19

Ixaru LG (2002) Lilix-a package for the solution of the coupled channel schrödinger equation. Comput Phys Commun 147(3):834–852

James K (2018) Mathematical physics with partial differential equations. Academic Press

John W (2009) Coalescent theory: an introduction. Number 575: 519.2 WAK

Kim KJ, Cools R, Ixaru LG (2002) Quadrature rules using first derivatives for oscillatory integrands. J Comput Appl Math 140(1–2):479–497

Kim KJ, Cools R, Ixaru LG (2003) Extended quadrature rules for oscillatory integrands. Appl Numer Math 46(1):59–73

Kingman JFC (1978) A simple model for the balance between selection and mutation. J Appl Prob 15(1):1–12

Mashayekhi S, Ixaru LG (2020) The least-squares fit of highly oscillatory functions using eta-based functions. J Comput Appl Math 376:112839

Mashayekhi S, Sedaghat S (2021) Fractional model of stem cell population dynamics. Chaos Solitons Fractals 146:110919

Neil ES, David S, Kenneth WW (2013) A mutation-selection model with recombination for general genotypes, vol. 222. Am Math Soc

Shijun L (2003) Beyond perturbation: introduction to the homotopy analysis method. CRC Press

Smith Francis J (1965) Quadrature methods based on the euler-maclaurin formula and on the clenshaw-curtis method of integration. Numer Math 7(5):406–411

Steinsaltz D, Evans SN, Wachter KW (2005) A generalized model of mutation-selection balance with applications to aging. Adv Appl Math 35(1):16–33

Xiaoyan L, Jin X, Zhi L, Jiahuan H (2020) The cardinal spline methods for the numerical solution of nonlinear integral equations. J Chem 2020

Yuan Linglong (2017) A generalization of kingman's model of selection and mutation and the lenski experiment. Math Biosci 285:61–67

Yuan L (2020) Kingman's model with random mutation probabilities: convergence and condensation ii. J Stat Phys 181(3):870–896

Ziari Shokrollah, Bica Alexandru Mihai (2019) An iterative numerical method to solve nonlinear fuzzy volterra-hammerstein integral equations. J Intell Fuzzy Syst 37(5):6717–6729