A DUAL PROCESS FOR THE COUPLED WRIGHT-FISHER DIFFUSION

MARTINA FAVERO†, HENRIK HULT∗, AND TIMO KOSKI

ABSTRACT. The coupled Wright-Fisher diffusion is a multi-dimensional Wright-Fisher diffusion for multi-locus and multi-allelic genetic frequencies, expressed as the strong solution to a system of stochastic differential equations that are coupled in the drift, where the pairwise interaction among loci is modelled by an inter-locus selection.

In this paper, a dual process to the coupled Wright-Fisher diffusion is derived, which contains transition rates corresponding to coalescence and mutation as well as single-locus selection and double-locus selection. The coalescence and mutation rates correspond to the typical transition rates of Kingman’s coalescent process. The single-locus selection rate not only contains the single-locus selection parameters in a form that generalises the rates for an ancestral selection graph, but it also contains the double-selection parameters to include the effect of the pairwise interaction on the single locus. The double-locus selection rate reflects the particular structure of pairwise interactions of the coupled Wright-Fisher diffusion.

Moreover, in the special case of two loci, two alleles, with selection and parent independent mutation, the stationary density for the coupled Wright-Fisher diffusion and the transition rates of the dual process are obtained in an explicit form.

1. INTRODUCTION

The coupled Wright-Fisher diffusion was introduced by Aurell, Ekeberg and Koski [2] with the purpose of analysing networks of loci in recombining populations of bacteria, or more precisely, detecting couples of loci co-evolving under strong selective pressure when the linkage disequilibrium is low across the genome. The model includes parent dependent mutation, interlocus selection and free recombination. Mutation is assumed to occur independently at each locus, while selection consists of first and second order selective interaction among loci. The biological significance of this type of interaction is found in [9].

The model considers $L$ different loci where, at each locus, a number of variants (alleles) is possible. The allele types at locus $l$ are labelled by $1, \ldots, M_l$, thus assuming that the type space at each locus is finite.

The coupled Wright-Fisher diffusion is obtained as the weak limit of a sequence of discrete Wright-Fisher models characterised by the assumption that the evolution of the population at one locus is conditionally independent on the other loci given that the previous generation at each locus is known. Here we state the definition of the diffusion as solution of a system of stochastic differential equations, without reference to the underlying discrete model.

MSC2010 subject classifications. Primary 60J70, 92D25 ; secondary 60J60, 92D10.

Key words and phrases. Wright-Fisher diffusion, Markov processes, duality, population genetics, ancestral graph.

† Corresponding author (mfavero@kth.se)

∗ Hult’s research is supported by the Swedish Research Council
The coupled Wright-Fisher diffusion, \( \mathbf{X} = \{ \mathbf{X}(t), t \geq 0 \} \), represents the evolution of the vector of all frequencies of allele types at each locus. Let
\[
\mathbf{X}^{(l)}(t) = (X_1^{(l)}(t), \ldots, X_M^{(l)}(t))^T
\]
represent the vector of frequencies at locus \( l \), with \( X_i^{(l)}(t) \) being the frequency of allele type \( i \) at locus \( l \), then
\[
\mathbf{X}(t) = (\mathbf{X}^{(1)}(t), \ldots, \mathbf{X}^{(L)}(t))^T.
\]

The process \( \mathbf{X} \) is the strong solution to the system of stochastic differential equations
\[
d\mathbf{X}(t) = \mu(\mathbf{X}(t))dt + D(\mathbf{X}(t))\nabla V(\mathbf{X}(t))dt + D^{1/2}(\mathbf{X}(t))d\mathbf{W}(t), \tag{1.1}
\]
where \( V \) is a specific quadratic function encoding the structure of the interactions, while the mutation vector \( \mu \) and the diffusion matrix \( D \) have the following block structure
\[
\mu(\mathbf{x}) = \begin{pmatrix}
\mu^{(1)}(\mathbf{x}^{(1)}) \\
\vdots \\
\mu^{(L)}(\mathbf{x}^{(L)})
\end{pmatrix}, \quad D(\mathbf{x}) = \begin{pmatrix}
D^{(1)}(\mathbf{x}^{(1)}) & & \\
& \ddots & \\
& & D^{(L)}(\mathbf{x}^{(L)})
\end{pmatrix},
\]
with \( \mu^{(l)} : \mathbb{R}^{M_l} \to \mathbb{R}^{M_l} \) and \( D^{(l)} : \mathbb{R}^{M_l} \to \mathbb{R}^{M_l \times M_l} \). The functions \( V, \mu \) and \( D \) are described in detail in the next section. The process \( \mathbf{W} = (\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(L)}) \) is a multidimensional Brownian motion with \( \mathbf{W}^{(l)} \) having the same dimension of \( \mathbf{X}^{(l)} \).

The system of SDEs (1.1) consists of \( L \) systems of equations for \( \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(L)} \), coupled by the drift term \( D \nabla V \). Note that, if \( \nabla V = 0 \), there is no interaction among the loci and the coupled Wright-Fisher diffusion consists of \( L \) independent Wright-Fisher diffusions, that is, each \( \mathbf{X}^{(l)} \) solves
\[
d\mathbf{X}^{(l)}(t) = \mu^{(l)}(\mathbf{X}^{(l)}(t))dt + D^{(l)}(\mathbf{X}^{(l)}(t))d\mathbf{W}^{(l)}(t),
\]
which is the SDE for a single-locus, multi-type Wright-Fisher diffusion with mutations. In fact, the selective interaction among different loci is completely described by the drift term \( D \nabla V \).

An interesting feature of the coupled Wright-Fisher diffusion, addressed by Aurell et al. \cite{2} as one of the main motivations for its introduction, is its stationary density which appeared, in a more general form, as a conjecture by Kimura over half a century ago. In \cite{6}, Kimura suggests a Wright-Fisher model for multi-locus and multi-allelic genetic frequencies and conjectures that the stationary density is of the form \( \pi(\bar{x})e^{m(\bar{x})} \), where \( \pi \) is the product of Dirichlet densities and \( m \) is a generic mean fitness term. The coupled Wright-Fisher diffusion is constructed so that the quadratic function \( V \) could replace the generic \( m \). Indeed, under the assumption of parent independent mutations, the stationary density of the coupled Wright-Fisher diffusion is known up to a normalising constant \( Z \), and corresponds to the one conjectured by Kimura with \( m = 2V \),
\[
p(\bar{x}) = \frac{1}{Z} \pi(\bar{x})e^{2V(\bar{x})}, \tag{1.2}
\]
see Section \cite{2} for the definition of \( \pi \) and \( V \). Note that, since the sum of the frequencies at each locus is equal to one, some components can be written as function of the others, which makes the distribution degenerate. Therefore, in order to deal with densities, a reduced process must be considered, that is, the last component at each locus is omitted and \( \bar{x} \) will be used, in place of \( x \), to denote the states of the reduced process.
In this paper the main result concerns the derivation of a dual process for the coupled Wright-Fisher diffusion. See [5] for a complete survey on duality for Markov processes. Markov duality has proven to be a useful technique in population genetics, several duality relationships have been established between different types of Wright-Fisher diffusions, modelling the evolution of allele frequencies forward in time, and reverse-time genealogical processes. The simplest and best known duality relationship in this framework is the one between the classic Wright-Fisher diffusion and the block counting process of the Kingman’s coalescent, which is a moment duality between a diffusion process and a jump process. Generalisations of this fundamental dual relation lead to the construction of many genealogical process, e.g. the ancestral selection graph [7, 8]. For an overview on the role of duality in population genetics see [4]. As pointed out in [4], the great value of the dual process is that it describes the history of a sample, more precisely, the posterior distribution given an allelic configuration at the present time which is needed for statistical inference. For example, knowing the backward dynamics, or a good approximation of them, turns out to be useful to construct good proposal distributions in importance sampling algorithms, see [10].

The main result in this paper is Theorem 5.1, which provides a description of the transition rates of a pure jump Markov process, \( \mathcal{N} = \{ \mathcal{N}(t), t \geq 0 \} \), that is dual to the coupled Wright-Fisher diffusion, \( \mathbf{X} \), through the duality relationship

\[
\mathbb{E}[F(\mathbf{X}(t), \mathbf{n})|\mathbf{X}(0) = \mathbf{x}] = \mathbb{E}[F(\mathbf{x}, \mathcal{N}(t))|\mathcal{N}(0) = \mathbf{n}],
\]

(1.3)

where \( F \) is a moment duality function which will be determined. The derivation use a generator approach as in [4] and [3]. It is based on the duality relationship of the infinitesimal generators

\[
L F(\cdot, \mathbf{n})(\mathbf{x}) = L^D F(\mathbf{x}, \cdot)(\mathbf{n}),
\]

(1.4)

where \( L \) is the generator of the coupled Wright-Fisher diffusion and \( L^D \) the unknown generator of the dual process. By proposing an appropriate duality function \( F \), the generator \( L^D \) of the dual process can be identified, from which transition rates of the dual process are obtained. It is known that the duality relationship (1.4) implies (1.3) under certain conditions on \( F \), see [5] for details.

The paper is outlined as follows. In Section 2 a background on the coupled Wright-Fisher diffusions is provided. Section 3 outlines the general generator approach to derive the dual process. In Section 4 the case of one locus, two allele types and parent independent mutations is considered. In this case the dual process is related to the ancestral selection graph, moreover, explicit formulas for the stationary density of the diffusion and the transition rates of the dual process are obtained. The main result is provided in Section 5 where the dual process is derived in the general multi-locus setting. The final Section 6 provides additional details in the case of two loci, two alleles, selection and parent independent mutations, more precisely, the transition rates of the dual process are expressed in terms of beta and confluent hypergeometric functions.

2. Preliminaries on the coupled Wright-Fisher diffusion

In this section the coupled Wright-Fisher diffusion is introduced and the explicit expression for its infinitesimal generator is provided. The notation in this section differs slightly from that in [2]. In [2], the frequency of the last allele type at each locus is omitted, being function of the other frequencies, whereas in this paper an expanded version of the diffusion is considered,
which includes all the frequencies. Since the frequencies sum up to one the descriptions are equivalent. For our purpose we find the expanded version more convenient to work with.

For a given integer \( L \geq 1 \), the number of loci, let \( M_1, \ldots, M_L \) be positive integers representing the number of alleles at each locus. Put \( M = \sum_{l=1}^{L} M_l \). A vector \( \mathbf{x} \in \mathbb{R}^M \) is interpreted as the concatenation of \( L \) vectors with lengths \( M_1, \ldots, M_L \), i.e. \( \mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)})^T \) with \( \mathbf{x}^{(l)} \in \mathbb{R}^{M_l} \) \( l = 1, \ldots, L \), and the coordinate \( i \) in vector \( \mathbf{x}^{(l)} \) is denoted by \( x^{(l)}_i \). Similarly, a matrix \( \mathbf{A} \in \mathbb{R}^{M \times M} \) consists of \( L^2 \) blocks with dimensions \( (M_l \times M_r)_{l,r=1,\ldots,L} \). The block at position \((l,r)\) is denoted by \( \mathbf{A}^{(lr)} \) and its component at position \((i,j)\) is denoted by \( A^{(lr)}_{ij} \).

Furthermore by \( \mathbf{e}_i^{(l)} \) it is denoted the unit vector in \( \mathbb{R}^{M_l} \) with the \( i \)th component of its \( l \)th building vector being equal to 1.

In the following, each of the terms appearing in (1.1) will be described, starting from the interaction drift term. The quadratic function \( V : \mathbb{R}^M \rightarrow \mathbb{R} \) is given by

\[
V(\mathbf{x}) = \mathbf{x}^T \mathbf{h} + \frac{1}{2} \mathbf{x}^T \mathbf{J} \mathbf{x},
\]

where \( \mathbf{h} \in \mathbb{R}^M_+ \) and \( \mathbf{J} \in \mathbb{R}^{M \times M}_+ \) is a symmetric block matrix with the blocks on the diagonal equal to zero matrices, i.e. \( J^{(ll)} = 0 \in \mathbb{R}^{M_l \times M_l} \) and \( J^{(lr)} = (J^{(r)}(l))^T \) for all \( l, r = 1, \ldots, L \). The vector \( \mathbf{h} \) and matrix \( \mathbf{J} \) contain the selection parameters, expressing, respectively, the single locus selection and the selective interaction among pairs of loci. Note that \( \nabla V(\mathbf{x}) = \mathbf{h} + \mathbf{J} \mathbf{x} \), since the matrix \( \mathbf{J} \) is symmetric.

Let \( g(\mathbf{x}) = D(\mathbf{x}) \nabla V(\mathbf{x}) \). Then, the components of \( g(\mathbf{x}) \) are

\[
g_i^{(l)}(\mathbf{x}) = \sum_{k=1}^{M_l} d_{ik}^{(l)}(\mathbf{x}) \tilde{h}_k^{(l)}(\mathbf{x}), \quad \text{with} \quad \tilde{h}_k^{(l)}(\mathbf{x}) = h_k^{(l)} + \sum_{r=1}^{L} \sum_{m=1}^{M_r} J_{km}^{(lr)} x^{(r)}_m. \tag{2.1}
\]

The drift function \( \mu \) models the mutations. It is assumed that mutations occur independently at each locus, in particular, at the \( l \)th locus the mutation rate is \( \theta_l^l / 2 \) and the probability matrix of mutations is \( P^{(l)} = (P_{ij}^{(l)})_{i,j=1,\ldots,M_l} \). The transition rates of mutation from type \( i \) to type \( j \) at locus \( l \) are thus \( u_{ij}^{(l)} = \theta_l^l P_{ij}^{(l)} \). As in the standard Wright-Fisher model with parent dependent mutations, the components of the drift function are defined by

\[
\mu_i^{(l)}(\mathbf{x})^{(l)} = \sum_{j=1}^{M_l} [u_{ji}^{(l)} x_j^{(l)} - u_{ij}^{(l)} x_i^{(l)}]. \tag{2.2}
\]

Finally, the components of the diagonal block \( D^{(l)}(\mathbf{x})^{(l)} \) of the diffusion matrix \( D(\mathbf{x}) \) are defined by

\[
d_{ij}^{(l)}(\mathbf{x})^{(l)} = x_i^{(l)} (\delta_{ij} - x_j^{(l)}) \quad \text{with} \quad \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \tag{2.3}
\]

which is characteristic for Wright-Fisher processes.

Having defined \( \mu, D, \) and \( V \), a compact definition of the coupled Wright-Fisher diffusion can be given, in terms of its infinitesimal generator. The coupled Wright-Fisher diffusion \( \{X(t)\}_{t \geq 0} \)
is a $M$-dimensional diffusion process on the state space

$$S = \left\{ x \in [0,1]^M \text{ s.t. } \sum_{i=1}^{M} x_i^{(l)} = 1 \quad \forall l = 1, \ldots, L \right\},$$

with generator

$$Lf(x) = \sum_{l=1}^{L} \sum_{i=1}^{M_l} \left( \mu_i^{(l)}(x) + g_i^{(l)}(x) \right) \frac{\partial f}{\partial x_i^{(l)}}(x) + \frac{1}{2} \sum_{i,j=1}^{M_l} d_{ij}^{(l)}(x) \frac{\partial^2 f}{\partial x_i^{(l)} \partial x_j^{(l)}}(x),$$

(2.4)

where $\mu$, $g$ and $d$ are given by (2.2), (2.1) and (2.3), respectively. The generator $L$ is defined on the domain $C^2(S)$.

Before proceeding with the derivation of the dual process, the stationary density (1.2) is considered. For existence of a density, the coupled Wright-Fisher diffusion must be restricted to the state space

$$\bar{S} = \left\{ \bar{x} \in [0,1]^{M-L} \text{ s.t. } \sum_{i=1}^{M_l-1} \bar{x}_i^{(l)} \leq 1 \quad \forall l = 1, \ldots, L \right\},$$

where

$$\bar{x} = (\bar{x}_1^{(1)}, \ldots, \bar{x}_{M_1-1}^{(1)}, \ldots, \bar{x}_1^{(L)}, \ldots, \bar{x}_{M_L-1}^{(L)})^T \in \bar{S}$$

corresponds to

$$x = \left( \bar{x}_1^{(1)}, \ldots, \bar{x}_{M_1-1}^{(1)}, \ldots, \bar{x}_1^{(L)}, \ldots, \bar{x}_{M_L-1}^{(L)} \right)^T \in S.$$ 

If there are no interactions among loci, the coupled Wright-Fisher diffusion consists of $L$ independent Wright-Fisher diffusions and the stationary density is well known when the mutations are parent independent. Wright himself proved that the stationary distribution of a single-locus, multi-type Wright-Fisher diffusion with parent independent mutations is Dirichlet [11]. Therefore, the stationary density of independent Wright-Fisher diffusions is the product of Dirichlet densities. More precisely, let

$$\pi(\bar{x}) = \prod_{l=1}^{L} \pi_l(\bar{x}^{(l)}), \quad \text{with} \quad \pi_l(\bar{x}) = \prod_{i=1}^{M_l-1} (\bar{x}_i^{(l)})^{2u_i^{(l)}-1} \left(1 - \sum_{i=1}^{M_l-1} \bar{x}_i^{(l)}\right)^{2u_{M_l}^{(l)}-1}.$$ 

$\pi(\bar{x})$ is the non-normalized stationary density of a coupled Wright-Fisher diffusion with no interaction among loci. In the presence of interaction, Aurell et al. [2] prove that there is an additional exponential factor in the stationary density, that is

$$p(\bar{x}) = \frac{1}{Z} \pi(\bar{x}) e^{2V(\bar{x})}, \quad (2.5)$$

with $V$ defined on the restricted space by naturally defining the missing frequencies as one minus the sum of the other frequencies at the same locus. It is assumed that the mutations are parent independent, i.e. $u_{ij}^{(l)} = u_{ij}^{(l)} \forall l$. The form of the stationary density is explicit up to a normalising constant. In general, it is difficult to compute the normalising constant $Z$.
explicitly, but under additional assumptions it can be computed numerically, as demonstrated in Section 4 and 6.

3. Outline of the derivation of a dual process

To derive a process that is dual to the coupled Wright-Fisher diffusion, a generator approach will be used as in [4], where the authors find a dual process for a multi-locus Wright-Fisher diffusion with recombination. In this section the method will be explained, in general terms.

Let \( L \) be the generator of the diffusion process (2.4) and \( L^{D} \) be the unknown generator of a dual process. Suppose that the following relationship holds

\[
L F(\cdot, n)(x) = L^{D} F(x, \cdot)(n) \quad x \in S, \quad n \in \mathbb{N}^M, \tag{3.1}
\]

for some duality function \( F \) that needs to be determined. By using the relationship above the transition rates of a dual process can be identified from its generator. To pursue this approach, it is necessary to compute the left hand side of (3.1) by applying the generator \( L \) to the dual function \( F \), considered as a function of \( x \), and rewrite it on form

\[
L F(\cdot, n)(x) = \sum_{\hat{n}} q(n, \hat{n}) \left[ F(x, \hat{n}) - F(x, n) \right], \tag{3.2}
\]

for some non-negative coefficients \( q(n, \hat{n}) \), \( \hat{n} \in \mathbb{N}^M \), \( \hat{n} \neq n \). In the light of the duality relationship, expression (3.2) can be interpreted as the generator \( L^{D} \) applied to the dual function \( F \), considered as a function of \( n \). Consequently, the dual process obtained this way is a pure jump process on the discrete space \( \mathbb{N}^M \) with transition rate matrix \( Q = (q(\cdot, \cdot)) \), the off-diagonal elements being the non-negative coefficients in (3.2) and the diagonal elements being chosen so that the rows sums are equal to 0. The alleged duality relationship is validated once the transition rates and the proper duality function are determined.

Consider the following proposal for the duality function, \( F \). The inspiration for the proposal comes from the duality function for the single locus Wright-Fisher diffusion with mutations. It can be generalised to the multi-loci setting by taking

\[
F(x, n) = \frac{1}{k(n)} \prod_{l=1}^{L} \prod_{i=1}^{M_l} (x_l^{(i)})^{n_l^{(i)}}, \tag{3.3}
\]

for some function \( k : \mathbb{N}^M \to \mathbb{R} \) that is determined as follows.

Let \( \tilde{X} \) be distributed according to the stationary distribution of the diffusion process \( \{X(t)\}_{t \geq 0} \), when such a density exists. Then \( \mathbb{E} \left[ L F(\tilde{X}, n) \right] = 0 \). Therefore, by taking expectation under the stationary distribution in (3.2), it follows that

\[
\sum_{\hat{n}} q(n, \hat{n}) \mathbb{E} \left[ F(\tilde{X}, \hat{n}) - F(\tilde{X}, n) \right] = 0,
\]

which implies that \( \mathbb{E} \left[ F(\tilde{X}, \cdot) \right] \) must be constant, the constant can be taken to be equal to 1, and consequently,

\[
k(n) = \mathbb{E} \left[ \prod_{l=1}^{L} \prod_{i=1}^{M_l} (\tilde{X}_l^{(i)})^{n_l^{(i)}} \right]. \tag{3.4}
\]
To find the transition rates of the dual process, it remains to obtain an expression of the form (3.2). In fact, it is sufficient to obtain an expression of the form
\[ \mathcal{L}F(\cdot, n)(x) = \sum_{\hat{n} \neq n} q(n, \hat{n}) F(x, \hat{n}) + q(n, n) F(x, n), \] (3.5)
with the requirement that \( q(n, \hat{n}) \) is positive for \( \hat{n} \neq n \) (it will be soon clear that \( q(n, n) \) is thus negative). In fact, once (3.5) is obtained, it is possible to derive expression (3.2) from it as follows. Rewriting (3.5) yields
\[ \mathcal{L}F(\cdot, n)(x) = \sum_{\hat{n} \neq n} q(n, \hat{n}) [F(x, \hat{n}) - F(x, n)] + \left( \sum_{\hat{n} \neq n} q(n, \hat{n}) + q(n, n) \right) F(x, n). \] (3.6)
Keeping in mind that \( \mathbb{E} \left[ \mathcal{L}F(\tilde{X}, n) \right] = 0 \) and that \( k \) is chosen such that \( \mathbb{E} \left[ F(\tilde{X}, n) \right] \) is constant, one can apply expectation with respect to the stationary distribution to get
\[ \sum_{\hat{n} \neq n} q(n, \hat{n}) + q(n, n) = 0. \] (3.7)
Therefore (3.6) implies (3.2) and it remains to write \( \mathcal{L}F \) as in (3.5) by finding the positive coefficients \( q(n, \hat{n}) \). Furthermore, (3.7) can be used to find a recursion formula for the function \( k \), as demonstrated in the following section.

Throughout the rest of the paper, the emphasis will be on obtaining an expression of the type (3.5). This approach is first illustrated in a simpler case (single locus), in order to lighten the formulas and highlight the ideas, and is subsequently used in the general case of coupled Wright-Fisher diffusions. The simpler case turns out to be closely related to a well known model: the ancestral selection graph.

### 4. The ancestral selection graph

When only one locus is considered, the coupled Wright-Fisher diffusion is simply a Wright-Fisher diffusion with selection. Let \( L = 1, \ M_1 = 2 \) and assume that mutations are parent independent, i.e. \( u_{ij} = u_j \) for \( i = 1, 2 \). Then the matrix of pairwise selection parameters is the zero matrix and the quadratic function \( V \) becomes linear
\[ V(x) = h_1 x_1 + h_2 x_2. \]
Let \( j(i) \) be the index opposite to \( i \),
\[ j(i) = \begin{cases} 2 & \text{if } i = 1 \\ 1 & \text{if } i = 2 \end{cases}. \]
Then, the drift terms can be written as follows
\[ \mu_i(x) = u_i x_{j(i)} - u_{j(i)} x_i, \]
\[ g_i(x) = h_i x_i (1 - x_i) - h_{j(i)} x_i x_{j(i)}, \quad i = 1, 2. \]
The diffusion process solving (1.1) under the assumptions above is a two-types Wright-Fisher diffusion with selection and parent independent mutations. It is known that the genealogical process corresponding to this type of Wright-Fisher diffusion is embedded in a graph with coalescing and branching structure, the ancestral selection graph, studied by Krone and Neuhauser [7, 8]. A dual process for the same type of diffusion is derived using a generator approach also
by Etheridge and Griffiths in [3], where the authors derive a dual process for the finite population size Moran model and use it to find the limiting transition rates of the dual process for the diffusion. Note that here is assumed, as in [3], that the types of individuals in the sample \( n \) are known, thus mutations are included in the dual process rather than superimposed afterwards on it, as in [7].

Following the outline in Section 3, a dual process is derived as follows. By applying the generator \( \mathcal{L} \) to the duality function \( F \) in (3.3), rewriting the derivatives of \( F \), and rearranging the terms yields

\[
\mathcal{L}F(\cdot, n)(x) = \sum_{i=1,2} (u_i x_{j(i)} - u_{j(i)} x_i) \frac{n_i}{x_i} F(x, n) + \sum_{i=1,2} x_i(h_i - h_i x_i - h_{j(i)} x_{j(i)}) \frac{n_i}{x_i} F(x, n)
\]

\[
+ \frac{1}{2} \sum_{i=1,2} x_i(1 - x_i) \frac{n_i(n_i - 1)}{(x_i)^2} F(x, n) - x_1 x_2 \frac{n_1 n_2}{x_1 x_2} F(x, n)
\]

\[
= \sum_{i=1,2} \frac{n_i(n_i - 1)}{2} \frac{1}{x_i} F(x, n) + \sum_{i=1,2} u_i n_i \frac{x_{j(i)}}{x_i} F(x, n) - \sum_{i=1,2} h_i(n_i + n_{j(i)}) x_i F(x, n)
\]

\[
- \left\{ \frac{n}{2}(n - 1) + \sum_{i=1,2} n_i u_{j(i)} - \sum_{i=1,2} n_i h_i \right\} F(x, n),
\]

where \( n = n_1 + n_2 \). To obtain an expression of the form (3.5) the expression in the last display can be rewritten as follows. First replace \( x_i = 1 - x_{j(i)} \) to obtain positive coefficients for the selection terms, then use the identities, for \( i = 1, 2 \),

\[
\frac{1}{x_i} F(x, n) = \frac{k(n - e_i)}{k(n)} F(x, n - e_i), \tag{4.1}
\]

\[
\frac{x_{j(i)}}{x_i} F(x, n) = \frac{k(n + e_{j(i)} - e_i)}{k(n)} F(x, n + e_{j(i)} - e_i), \tag{4.2}
\]

\[
x_i F(x, n) = \frac{k(n + e_i)}{k(n)} F(x, n + e_i), \tag{4.3}
\]

where \( e_i, i = 1, 2 \), are the unit vectors in \( \mathbb{N}^2 \). Finally it yields,

\[
\mathcal{L}F(\cdot, n)(x) = \sum_{i=1,2} \frac{n_i(n_i - 1)}{2} \frac{k(n - e_i)}{k(n)} F(x, n - e_i)
\]

\[
+ \sum_{i=1,2} u_i n_i \frac{k(n + e_{j(i)} - e_i)}{k(n)} F(x, n + e_{j(i)} - e_i)
\]

\[
+ \sum_{i=1,2} h_{j(i)} n \frac{k(n + e_i)}{k(n)} F(x, n + e_i)
\]

\[
- \left\{ \frac{n}{2}(n - 1) + \sum_{i=1,2} n_i u_{j(i)} + \sum_{i=1,2} n_{j(i)} h_i \right\} F(x, n),
\]
which is the desired expression. As demonstrated in Section 3, the transition rates of a dual process can be identified directly from this expression. Therefore a dual process for the Wright-Fisher diffusion considered in this section is a pure jump process on the state space $\mathbb{N}^2$, with transition rates as follows. The dual process, in state $n$,

- jumps to state $n - e_i$ at rate
  $$q(n, n - e_i) = \frac{n_i(n_i - 1) k(n - e_i)}{2};$$
  \[\text{coalescence}\]

- jumps to state $n + e_{j(i)} - e_i$ at rate
  $$q(n, n + e_{j(i)} - e_i) = u_i n_i \frac{k(n + e_{j(i)} - e_i)}{k(n)};$$
  \[\text{mutation}\]

- jumps to state $n + e_i$ at rate
  $$q(n, n + e_i) = h_{j(i)} n \frac{k(n + e_i)}{k(n)}.$$
  \[\text{selection}\]

As anticipated, the dual process just described corresponds to the limiting process in [3] which mirrors the block counting process of the ancestral selection graph. From the transition rates $q$, it is observed that three types of events are possible for the dual process: mutation, coalescence and branching. The first two appear also in the Kingman’s coalescent, while the latter is a virtual addition to the true genealogical process which is characteristic of the ancestral selection graph. Seen forward in time, it corresponds to the event that two potential parents are chosen and only the one carrying the advantageous allele reproduces. Backward in time, when a branching happens, the individual splits into two individuals: its true parent and its virtual (potential) parent.

To complete the identification of the transition rates, $q(n, n)$ is defined as the coefficient of $F(x, n)$ in (4.4),

$$q(n, n) = -\frac{n}{2} (n - 1) - \sum_{i=1}^{2} n_i u_{j(i)} - \sum_{i=1}^{2} n_{j(i)} h_i.$$  

Furthermore, (3.7) ensures that the rows sums of the transition matrix equal zero and provides a recursion formula for $k$

$$k(n) = \left\{ \frac{n}{2} (n - 1) + \sum_{i=1}^{2} n_i u_{j(i)} + \sum_{i=1}^{2} n_{j(i)} h_i \right\} k(n)$$

$$= \sum_{i=1}^{2} \frac{n_i(n_i - 1)}{2} k(n - e_i) + \sum_{i=1}^{2} u_i n_i k(n + e_{j(i)} - e_i) \sum_{i=1}^{2} h_{j(i)} n k(n + e_i).$$

with boundary conditions $k(e_1) = \rho_1$, $k(e_2) = 1 - \rho_1$, where $\rho_1$ is the probability that the ultimate ancestor is of type 1, i.e. $\rho_1 = \int_0^1 x p(x) dx$. This type of recursion formula is usually too large to be solved directly, importance sampling techniques have been developed to estimate $k$, see [10].
In general it is not possible to find a closed-form expression for $k$ and thus for the transition rates. However, when the mutations are parent independent, as in this example, the stationary density is explicitly known up to a normalizing constant $Z$ and thus $k$ can be written as an integral with respect to the stationary density

$$k(n) = \frac{1}{Z} \int_0^1 x^{n_1+2u_1-1} (1-x)^{n_2+2u_2-1} e^{2[h_1 x + h_2 (1-x)]} \, dx.$$ 

The integral above cannot be computed analytically but it is related to the confluent hypergeometric function of the first kind, the Kummer’s function, which can be efficiently computed numerically. The idea of using Kummer’s functions is taken from [2] and [7], where this type of functions have been used to find respectively a series representation for the normalising constant in an example and a representation for the probability $p_1$. Let $\text{1F1}$ be the confluent hypergeometric function, then, using its integral representation, it yields

$$k(n) = \frac{1}{Z} e^{2h_2} \frac{\Gamma(n_1 + 2u_1 + n_2 + 2u_2)}{\Gamma(n_1 + 2u_1) \Gamma(n_2 + 2u_2)} \text{1F1}(n_1 + 2u_1, n_1 + 2u_1 + n_2 + 2u_2, 2(h_1 - h_2)),$$

$$Z = \frac{\Gamma(2u_1 + 2u_2)}{\Gamma(2u_1) \Gamma(2u_2)} \text{1F1}(2u_1, 2u_1 + 2u_2, 2(h_1 - h_2)),$$

see [1] for a complete collection of definitions and properties of confluent hypergeometric functions.

5. A multi-locus dual process

In this section the derivation in the previous section is extended to the general multi-locus setting, $L \geq 1$ and $M_l \geq 2$, $l = 1, \ldots, L$. The diffusion process is given by (1.1), with generator (2.4), where $\mu$, $g$ and $d$ are given by (2.2), (2.1) and (2.3), respectively. Let $k$ be defined by (3.4).

**Theorem 5.1.** The dual process for the coupled Wright-Fisher diffusion is a pure jump process on the state space $\mathbb{N}^M$ with the following transition rates. From the current state, $n$, the dual process jumps to

- $n - e^{(l)}_i$, $i = 1, \ldots, M_l, l = 1, \ldots, L$, at rate
  $$q(n, n - e^{(l)}_i) = n^{(l)}_i (n^{(l)}_i - 1) k(n - e^{(l)}_i) / k(n);$$  
  [coalescence]

- $n - e^{(l)}_i + e^{(l)}_j$, $i, j = 1, \ldots, M_l, l = 1, \ldots, L$, at rate
  $$q(n, n - e^{(l)}_i + e^{(l)}_j) = n^{(l)}_i u^{(l)}_{ij} k(n - e^{(l)}_i + e^{(l)}_j) / k(n);$$  
  [mutation]
A DUAL PROCESS FOR THE COUPLED WRIGHT-FISHER DIFFUSION

\[ n + e_j^{(l)}, j = 1, \ldots, M_l, l = 1, \ldots, L, \text{ at rate} \]

\[
q(n, n + e_j^{(l)}) = \left( n^{(l)} \sum_{k=1 \atop k \neq j}^{M_l} k_k^{(l)} + \sum_{r=1 \atop r \neq l}^{L} \sum_{i=1}^{M_l} n_i^{(r)} J_{ji}^{(l)} \right) \frac{k(n + e_j^{(l)})}{k(n)};
\]

[single-locus selection]

\[ n + e_j^{(l)} + e_h^{(r)}, j = 1, \ldots, M_l, h = 1, \ldots, M_r, l, r = 1, \ldots, L, l \neq r, \text{ at rate} \]

\[
q(n, n + e_j^{(l)} + e_h^{(r)}) = \left( (n^{(l)} + n^{(r)}) \sum_{k=1 \atop k \neq j}^{M_l} \sum_{m=1 \atop m \neq h}^{M_r} J_{km}^{(l)} \right) \frac{k(n + e_j^{(l)} + e_h^{(r)})}{k(n)}.
\]

[double-locus selection]

where \( n^{(l)} = \sum_{i=1}^{M_l} n_i^{(l)} \).

Note that the mutation and coalescence jumps involve one locus at the time. The coalescence and mutation rates are similar to the typical transition rates of the Kingman’s coalescent process with mutations, the only difference being the function \( k \) which, despite having the same structure, is based on a different stationary density and depends on all the loci, not only on the one where the jump takes place. The single-locus selection rate not only contains the single-locus selection parameters in a form that generalises the rates in Section 4, but it also contains the double-selection parameters to include the effect of the pairwise interaction on the single locus. The single-locus selection, as the coalescence and mutation, jumps involve one locus at the time. Finally, the double-locus selection rate reflects the particular structure of pairwise interactions of the coupled Wright-Fisher diffusion and it is, to the best of our knowledge, a novel type of transition rate appearing in genealogical processes related to Wright-Fisher diffusions. The double-locus selection jumps correspond to simultaneous branchings at two different loci. The explicit parts of the transition rates (not depending on the function \( k \)) have a very natural interpretation. As in the simpler case studied in Section 4, the basic principle is that weak types branch at a higher rate. The difference is that, while in the simpler case there are only two types, a viable type and a weaker type, here there are many types and many loci all influencing each others branching rates. To understand this behaviour in greater detail, some terms will be investigated more thoroughly. The term

\[
n^{(l)}(l) \sum_{k=1 \atop k \neq j}^{M_l} h_k^{(l)},
\]

arises purely from the single-locus selection and contributes to the rate of adding a gene of type \( j \) at locus \( l \). It depends on the single-locus viability of the other allele types (all except type \( j \)) at locus \( l \), the higher their viability, the higher the rate of adding type \( j \), and of course it is also directly proportional to the number of genes at locus \( l \), \( n^{(l)} \).
The rate of adding a couple of genes of type \( j \) at locus \( l \) and of type \( h \) at locus \( r \) is related to the term

\[
(n^{(l)} + n^{(r)}) \sum_{k=1}^{M_l} \sum_{m=1}^{M_r} J_{km}^{(lr)}.
\]

It depends on the viability of the other couples of allele types (all except couple \( j, h \)), the higher their viability, the higher the rate of adding type \( j \) and \( h \) at locus \( l \) and \( r \), respectively. Again the rate is directly proportional to the number of genes at loci \( l \) and \( r \).

Although the interpretation of some parts of the transition rates is straightforward, the function \( k \) remains implicit, even for the simpler Kingman’s coalescent process with parent dependent mutations. When the mutations are parent independent, the stationary density is known up to a normalising constant and \( k \) can be expressed as an integral that sometimes can be computed numerically, as shown in Section 4 and as it will be shown in Section 6, where a series representation of \( k \) involving Kummer’s and Beta functions will be given. However, in the general case, we have to settle for a recursion formula for \( k \).

**Proof of Theorem 5.1** Following the outline in Section 3, a dual process is derived as follows.

By applying the generator \( \mathcal{L} \) to the duality function \( F \) in (3.3) each term in the expression for \( \mathcal{L}F \) is treated separately. As in the previous section, the terms corresponding to mutation and diffusion can be easily rewritten in the required form. Summing the mutation terms over allele types at locus \( l \) yields

\[
\sum_{i=1}^{M_l} \mu_i^{(l)}(x^{(l)}) \frac{\partial F}{\partial x_i^{(l)}}(x, n) = \sum_{i,j=1}^{M_l} n_i^{(l)} n_j^{(l)} \frac{u_{ij}}{x_i^{(l)}} F(x, n) - \sum_{i,j=1}^{M_l} n_i^{(l)} u_{ij} F(x, n).
\]

By using identity (4.2) at locus \( l \) the mutation terms can be rewritten in the desired form

\[
\sum_{i=1}^{M_l} \mu_i^{(l)}(x^{(l)}) \frac{\partial F}{\partial x_i^{(l)}}(x, n) = \sum_{i,j=1}^{M_l} n_i^{(l)} n_j^{(l)} \frac{k(n - e_i^{(l)} + e_j^{(l)})}{k(n)} F(x, n - e_i^{(l)} + e_j^{(l)}) - \sum_{i,j=1}^{M_l} n_i^{(l)} u_{ij} F(x, n).
\]

For the diffusion terms, the diagonal and off-diagonal terms are written separately as

\[
d_{ii}^{(l)}(x^{(l)}) \frac{\partial^2 F}{\partial x_i^{(l)} \partial x_i^{(l)}}(x, n) = n_i^{(l)}(n_i^{(l)} - 1) \frac{1}{x_i^{(l)}} F(x, n) - n_i^{(l)}(n_i^{(l)} - 1) F(x, n),
\]

\[
d_{ij}^{(l)}(x^{(l)}) \frac{\partial^2 F}{\partial x_i^{(l)} \partial x_j^{(l)}}(x, n) = -n_i^{(l)} n_j^{(l)} F(x, n) \quad i \neq j.
\]

Summing the diffusion terms at locus \( l \) and rearranging yields

\[
\frac{1}{2} \sum_{i,j=1}^{M_l} d_{ij}^{(l)}(x^{(l)}) \frac{\partial^2 F}{\partial x_i^{(l)} \partial x_j^{(l)}}(x, n) = \sum_{i=1}^{M_l} n_i^{(l)}(n_i^{(l)} - 1) \frac{1}{2} \frac{1}{x_i^{(l)}} F(x, n) - \frac{1}{2} n_i^{(l)}(n_i^{(l)} - 1) F(x, n).
\]
Now use identity (4.1) at locus $l$ to obtain
\[
\frac{1}{2} \sum_{i,j=1}^{M_l} a_{ij}^{(l)}(x^{(l)}) \frac{\partial^2 F}{\partial x_i^{(l)} \partial x_j^{(l)}}(x, n) = \sum_{i=1}^{M_l} n_i^{(l)} (n_i^{(l)} - 1) \frac{k(n - e_i^{(l)})}{k(n)} F(x, n - e_i^{(l)}) - \frac{1}{2} n_i^{(l)} (n_i^{(l)} - 1) F(x, n),
\] (5.2)

Next, consider the interaction terms. Using the definition of $g^{(l)}$ and rewriting the derivatives of $F$ yields,
\[
\sum_{i=1}^{M_l} g_i^{(l)}(x) \frac{\partial F}{\partial x_i^{(l)}}(x, n) = \sum_{i=1}^{M_l} \sum_{k=1}^{M_r} n_i^{(l)} h_k^{(l)} (\delta_{ik} - x_k^{(l)}) F(x, n)
\]
\[
+ \sum_{r=1}^{L} \sum_{i=1}^{M_l} \sum_{k=1}^{M_r} \sum_{m=1}^{M_r} n_i^{(l)} j_{km}^{(lr)} (\delta_{ik} - x_k^{(l)}) x_m^{(r)} F(x, n).
\] (5.3)

Note that the first group of sums, $S_1$, contains the single-locus selection parameters while the second, $S_2$, contains the pairwise selection parameters. Each of them will be treated separately. The single-locus selection term can be rearranged into
\[
S_1 = \sum_{i=1}^{M_l} n_i^{(l)} h_i^{(l)} F(x, n) - \sum_{i=1}^{M_l} \sum_{k=1}^{M_r} n_i^{(l)} h_k^{(l)} x_k^{(l)} F(x, n).
\]

As in Section 4, the fact that the sum of the frequencies at each locus equals one is used. Since $x_k^{(l)} = 1 - \sum_{j \neq k}^{M_l} x_j^{(l)}$, the terms can be rearranged to obtain
\[
S_1 = - \sum_{i=1}^{M_l} \sum_{k=1}^{M_r} n_i^{(l)} h_k^{(l)} F(x, n) + \sum_{j=1}^{M_l} \left( n_i^{(l)} \sum_{k=1}^{M_r} h_k^{(l)} \right) x_j^{(l)} F(x, n).
\]

The second part of (5.3) can be expressed as
\[
S_2 = \sum_{r=1}^{L} \sum_{i=1}^{M_l} \sum_{m=1}^{M_r} n_i^{(l)} j_{im}^{(lr)} x_m^{(r)} F(x, n) - \sum_{r=1}^{L} \sum_{k=1}^{M_r} \sum_{m=1}^{M_r} n_i^{(l)} j_{km}^{(lr)} x_k^{(l)} x_m^{(r)} F(x, n).
\]

This time the equality
\[
-x_k^{(l)} x_m^{(r)} = -1 + \sum_{j=1}^{M_l} \sum_{h=1}^{M_r} (1 - \delta_{hm} \delta_{jk}) x_j^{(l)} x_h^{(r)} ,
\] (5.4)
will be used. To see that (5.4) holds, the fact that the frequencies sum up to one at each locus is used multiple times, as follows,

\[-x_k^{(l)} x_m^{(r)} = - x_k^{(l)} \left( 1 - \sum_{h=1 \atop h \neq m}^{M_r} x_h^{(r)} \right)\]

\[= - x_k^{(l)} + \sum_{h=1}^{M_r} (1 - \delta_{hm}) x_k^{(l)} x_h^{(r)}\]

\[= - 1 + \sum_{j=1 \atop j \neq k}^{M_l} x_j^{(l)} \sum_{h=1}^{M_r} x_h^{(r)} + \sum_{h=1}^{M_r} (1 - \delta_{hm}) x_k^{(l)} x_h^{(r)}\]

\[= - 1 + \sum_{j=1}^{M_l} \sum_{h=1}^{M_r} (1 - \delta_{jk}) x_k^{(l)} x_h^{(r)} + \sum_{j=1}^{M_l} \sum_{h=1}^{M_r} \delta_{jk} (1 - \delta_{hm}) x_k^{(l)} x_h^{(r)}\]

\[= - 1 + \sum_{j=1}^{M_l} \sum_{h=1}^{M_r} (1 - \delta_{hm} \delta_{jk}) x_j^{(l)} x_h^{(r)}\]

Applying (5.4) in the expression for \(S_2\) and rearranging the terms, leads to

\[S_2 = - \sum_{r=1}^{L} \sum_{k=1}^{M_l} \sum_{m=1}^{M_r} n_i^{(l)} J_{km}^{(lr)} F(x, n) + \sum_{r=1}^{L} \sum_{m=1}^{M_r} \left( \sum_{i=1}^{M_l} n_i^{(l)} J_{im}^{(lr)} \right) x_m^{(r)} F(x, n)\]

\[+ \sum_{r=1}^{L} \sum_{k=1}^{M_l} \sum_{m=1}^{M_r} \sum_{j=1 \atop j \neq k}^{M_l} n_i^{(l)} \sum_{m=1}^{M_r} \sum_{h=1 \atop h \neq m}^{M_r} J_{km}^{(lr)} x_j^{(l)} x_h^{(r)} F(x, n)\]

Summing over \(l\) and putting similar terms together yields

\[\sum_{l=1}^{L} \sum_{i=1}^{M_l} g_i^{(l)}(x) \frac{\partial F}{\partial x_i}(x, n) = - \left( \sum_{l=1}^{L} \sum_{i=1}^{M_l} \sum_{k=1 \atop k \neq i}^{M_l} n_i^{(l)} h_{k}^{(l)} + \sum_{l=1}^{L} \sum_{r=1}^{L} \sum_{m=1}^{M_r} \sum_{i=1}^{M_l} n_i^{(l)} J_{km}^{(lr)} \right) F(x, n)\]

\[+ \sum_{l=1}^{L} \sum_{j=1}^{M_l} \left( n_i^{(l)} \sum_{k=1 \atop k \neq j}^{M_l} h_{k}^{(l)} + \sum_{r=1}^{L} \sum_{i=1}^{M_l} n_i^{(r)} J_{ji}^{(lr)} \right) x_j^{(l)} F(x, n)\]

\[+ \sum_{l=1}^{L} \sum_{r=1}^{L} \sum_{m=1}^{M_r} \sum_{j=1}^{M_l} \sum_{h=1 \atop h \neq m}^{M_r} \left( n_i^{(l)} + n_i^{(r)} \right) J_{km}^{(lr)} x_j^{(l)} x_h^{(r)} F(x, n)\].
Use the identities (4.3) at locus \( l \) and
\[
x_j^{(l)} x_h^{(r)} F(x, n) = \frac{k(n + e_j^{(l)} + e_h^{(r)})}{k(n)} F(x, n + e_j^{(l)} + e_h^{(r)})
\]
for the mixed terms involving loci \( l \) and \( r \), in order to rewrite the selection terms in the desired form
\[
\sum_{l=1}^{L} \sum_{i=1}^{M_l} g_i^{(l)}(x) \frac{\partial F}{\partial x_i^{(l)}}(x, n) =
\]
\[
- \left( \sum_{l=1}^{L} \sum_{i=1}^{M_l} \sum_{k=1}^{M_l} n_i^{(l)} k_k^{(l)} + \sum_{l=1}^{L} \sum_{r \neq l}^{L} \sum_{i=1}^{M_l} \sum_{m=1}^{M_r} n_i^{(l)} j_{k_m}^{lr} \right) \frac{k(n + e_j^{(l)})}{k(n)} F(x, n + e_j^{(l)})
\]
\[
+ \sum_{l=1}^{L} \sum_{j=1}^{M_l} \left( n_i^{(l)} k_k^{(l)} + \sum_{r \neq l}^{L} \sum_{i=1}^{M_l} n_i^{(l)} j_{k_m}^{lr} \right) \frac{k(n + e_j^{(l)} + e_h^{(l)})}{k(n)} F(x, n + e_j^{(l)} + e_h^{(l)}).
\]

The terms corresponding to mutation (5.1), diffusion (5.2) and selection (5.5) are now written in form (3.5). It is finally possible to identify the transition rates of the dual process and this completes the proof.

6. Two loci, two alleles, with selection and parent independent mutation

In this section a particular example will be considered, where there are two loci, \( L = 2 \), and two allele types at each locus, \( M_1 = M_2 = 2 \). The pairwise interactions are represented by the matrix
\[
J = \begin{pmatrix}
0 & 0 & J_1 & 0 \\
0 & 0 & 0 & J_2 \\
J_1 & 0 & 0 & 0 \\
0 & J_2 & 0 & 0
\end{pmatrix},
\]
and there is no single-locus selection, \( h = 0 \). Furthermore, parent independent mutations are assumed.

In this special case, the function \( k \), in (3.4), and consequently the transition rates of the dual process can be computed rather efficiently. The main difficulty in the computation is that the normalising constant of the stationary density (2.5) is unknown. It may be noted that computing the normalising constant and the function \( k \) are closely related problems. In fact, by defining
\[
I(a_1, a_2, b_1, b_2) = \int_0^1 \int_0^1 x^{a_1-1} (1 - x)^{a_2-1} y^{b_1-1} (1 - y)^{b_2-1} e^{2[J_1xy + J_2(1-x)(1-y)]} dx dy,
\]
the normalising constant can be written as
\[ Z = I(2u_1^{(1)}, 2u_2^{(1)}, 2u_1^{(2)}, 2u_2^{(2)}), \]
and the function \( k \) as
\[ k(n) = \frac{1}{Z} I(n_1^{(1)} + 2u_1^{(1)}, n_2^{(1)} + 2u_1^{(2)}, n_1^{(2)}, n_2^{(2)} + 2u_2^{(2)}). \]

The integral \( I \) cannot be computed analytically, but it is possible to find a series representation of it in terms of Beta and Kummer functions, which can be truncated to numerically evaluate the function \( k \). The following formula is derived by a straightforward, albeit cumbersome, application of definitions and properties of Kummer functions
\[
I(a_1, a_2, b_1, b_2) = e^{2J_2} B(a_1, a_2) \sum_{n=0}^{\infty} \frac{[a_1]_n}{[a_1 + a_2]_n} \frac{(-2J_2)^n}{n!} \times \sum_{k=0}^{n} \binom{n}{k} \left(-\frac{J_1 + J_2}{J_2}\right)^k B(k + b_1, b_2) _1F_1(k + b_1, k + b_1 + b_2, -2J_2),
\]
where \( B \) is the Beta function, \(_1F_1\) is the Kummer function and \([a]_n = a(a+1) \cdots (a+n-1)\), for \( n > 0, [a]_0 = 1 \).

As an illustration, the stationary density of independent Wright-Fisher diffusions, with \( J_1 = J_2 = 0 \), is compared to the stationary density of the coupled Wright-Fisher diffusion, with \( J_1 = J_2 = 2 \), in Figure 6.1. Both distributions have mutation rates \( u_1^{(1)} = u_2^{(1)} = u_1^{(2)} = u_2^{(2)} = 0.8 \).

**Figure 6.1.** Stationary density of a coupled Wright-Fisher diffusion for two loci, two alleles, with no interaction (left) and nonzero interaction (right). Mutation parameters: \( u_1^{(1)} = u_2^{(1)} = u_1^{(2)} = u_2^{(2)} = 0.8. \) Double-locus selection parameters: \( J_1 = J_2 = 0 \) (left), \( J_1 = J_2 = 2 \) (right).

**References**

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Dover publications Inc., 1970.
[2] E. Aurell, M. Ekeberg, and T. Koski. Networks of loci by a multilocus Wright-Fisher model with selection and mutation, and a conjecture by Motoo Kimura. *ArXiv e-prints*, art. arXiv:1906.00716, 2019.

[3] A. M. Etheridge and R. C. Griffiths. A coalescent dual process in a Moran model with genic selection. *Theoretical Population Biology*, 75:320–330, 2009.

[4] R. C. Griffiths, P. A. Jenkins, and S. Lessard. A coalescent dual process for a Wright-Fisher diffusion with recombination and its application to haplotype partitioning. *Theoretical Population Biology*, 112:126–138, 2016.

[5] S. Jensen and N. Kurt. On the notion(s) of duality for Markov processes. *Probability Surveys*, 11:59–120, 2014.

[6] M. Kimura. Stochastic processes and distribution of gene frequencies under natural selection. *Cold Spring Harbor Symposia on Quantitative Biology*, 20:33–53, 1955.

[7] S. M. Krone and C. Neuhauser. Ancestral processes with selection. *Theoretical Population Biology*, 51:210–237, 1997.

[8] C. Neuhauser and S. M. Krone. The genealogy of samples in models with selection. *Genetics*, 154:519–534, 1997.

[9] M. J. Skwark, N.J.Croucher, S.Puranen, C. Chewapreecha, and M.Pesonen et al. Interacting networks of resistance, virulence and core machinery genes identifies by genome-wide epistasis analysis. *PLoS Genetics*, 13(2), 2017.

[10] M. Stephens. Inference under the coalescent. In D. Balding, M. Bishop, and C. Cannings, editors, *Handbook of Statistical Genetics*, chapter 26, pages 878–908. Whiley, Chichester, UK, 2007.

[11] S. Wright. Adaptation and selection. In G. L. Jepson, E. Mayr, and G.G. Simpson, editors, *Genetics, Palaeontology, and Evolution*, pages 365–389. Princeton Univ. Press, 1949.