

# Diffusion approximations via Stein's method and time changes

Mikołaj J. Kasprzak,

*Mikołaj Kasprzak  
Department of Statistics  
University of Oxford  
24-29 St Giles'  
Oxford, OX1 3LB  
United Kingdom*

*e-mail:* [kasprzak@stats.ox.ac.uk](mailto:kasprzak@stats.ox.ac.uk)

**Abstract:** We extend the ideas of [Bar90] and use Stein's method to obtain a bound on the distance between a scaled time-changed random walk and a time-changed Brownian Motion. We then apply this result to bound the distance between a time-changed compensated scaled Poisson process and a time-changed Brownian Motion.

This allows us to bound the distance between a process whose dynamics resemble those of the Moran model with mutation and a process whose dynamics resemble those of the Wright-Fisher diffusion with mutation upon noting that the former may be expressed as a difference of two time-changed Poisson processes and the diffusive part of the latter may be expressed as a time-changed Brownian Motion.

The method is applicable to a much wider class of examples satisfying the Stroock-Varadhan theory of diffusion approximation ([SV79]).

**MSC 2010 subject classifications:** Primary 60B10, 60F17; secondary 60J70, 60J65, 60E05, 60E15.

**Keywords and phrases:** Stein's method, functional convergence, time-changed Brownian Motion, Moran model, Wright-Fisher diffusion.

## 1. Introduction

In his seminal paper [Ste72], Charles Stein introduced a method for proving normal approximations and obtained a bound on the speed of convergence to the standard normal distribution. He observed that a random variable  $Z$  has the standard normal law if and only if  $\mathbb{E}Zf(Z) = \mathbb{E}f'(Z)$  for all smooth functions  $f$ . Therefore, if, for a random variable  $W$  with mean 0 and variance 1,  $\mathbb{E}f'(W) - \mathbb{E}Wf(W)$  is close to zero for a large class of functions  $f$ , then the law of  $W$  should be approximately Gaussian. He then proposed that, instead of evaluating  $|\mathbb{E}h(W) - \mathbb{E}h(Z)|$  directly for a given function  $h$ , one can first find an  $f = f_h$  solving the following *Stein equation*:

$$f'(w) - wf(w) = h(w) - \mathbb{E}h(Z)$$

and then find a bound on  $|\mathbb{E}f'(W) - \mathbb{E}Wf(W)|$ . This approach often turns out to be much easier, due to some bounds on the solutions  $f_h$ , which can be

derived in terms of the derivatives of  $h$ . Since then, Stein's method has been significantly developed and extended to approximations by distributions other than normal.

The aim of Stein's method is to find a bound on the quantity  $|\mathbb{E}_{\nu_n} h - \mathbb{E}_\mu h|$ , where  $\mu$  is the target (known) distribution,  $\nu_n$  is the approximating law and  $h$  is chosen from a suitable class of real-valued test functions  $\mathcal{H}$ . The idea is to find an operator  $\mathcal{A}$  acting on a class of real-valued functions such that

$$(\forall f \in \text{Domain}(\mathcal{A}) \quad \mathbb{E}_\nu \mathcal{A}f = 0) \iff \nu = \mu,$$

where  $\mu$  is our target distribution. In the next step, for a given function  $h \in \mathcal{H}$ , a solution  $f = f_h$  to the following Stein equation:

$$\mathcal{A}f = h - \mathbb{E}_\mu h$$

is sought and its properties studied. Finally, using various mathematical tools (among which the most popular are Taylor's expansions in the continuous case, Malliavin calculus, as described in [NP12], and coupling methods), a bound is sought for the quantity  $|\mathbb{E}_{\nu_n} \mathcal{A}f_h|$ .

An accessible account of the method can be found, for example, in the surveys [LRS17] and [Ros11] as well as the books [BHJ92] and [CGS11], which treat the cases of Poisson and normal approximation, respectively, in detail. [Swa16] is a database of information and publications connected to Stein's method.

Approximations by infinite-dimensional laws have not been covered in the Stein's method literature very widely, with the notable exceptions of [Bar90], [BJ09] and recently [CD13]. We will focus on the ideas taken from [Bar90], which provides bounds on the Brownian Motion approximation of a one-dimensional scaled random walk and some other one-dimensional processes including scaled sums of locally dependent random variables and examples from combinatorics. In the sequel, we show that the approach presented in [Bar90] can be extended to time-changes of Brownian Motion, including diffusions in the natural scale.

The most important example we apply our theory to is the approximation of the Moran model with mutation by the Wright-Fisher diffusion with mutation. The former, first introduced in [Mor58] as an alternative for the Wright-Fisher model (first formally described in [Wri31]) is one of the simplest and most important models of the genetic drift, i.e. the change in the frequencies of alleles in a population. It assumes that the population is divided into two allelic types (A and a) and the frequency of each of the alleles is governed by a birth-death process and an independent mutation process. Specifically, in a population of size  $n$ , at exponential rate  $\binom{n}{2}$ , a pair of genes is sampled uniformly at random. Then one of them is selected at random to die and the other one gives birth to another gene of the same allelic type. In addition, every gene of type a changes its type independently at rate  $\nu_2$  and every gene of type A changes its type independently at rate  $\nu_1$ . The model then looks at the proportion of a-genes in the population.

On the other hand, the Wright-Fisher model is a discrete Markov chain and does not allow for overlapping generations. Specifically, each step represents a

generation. In generation  $k$  each of the  $n$  individuals chooses its parent independently, uniformly at random from the individuals present in generation  $k-1$  and inherits its genetic type. This model also then looks at the proportion of  $a$ -genes in the population.

The Moran model turns out to be easier to study mathematically. It may be proved, for instance using the Stroock-Varadhan theory of diffusion approximation (see [SV79]), that it converges weakly to the Wright-Fisher diffusion, which is also a scaling limit of the Wright-Fisher model. We show how to put a bound on the speed of this convergence. The Wright-Fisher diffusion is often used in practice in genetics for inference concerning large populations. It is given by the equation:

$$dM(t) = \gamma(M(t))dt + \sqrt{M(t)(1-M(t))}dB_t,$$

where  $\gamma : [0, 1] \rightarrow \mathbb{R}$  encompasses mutation. For a discussion of probabilistic models in genetics see [Eth12].

In Section 2 we introduce the space of test functions we will find the bounds for. In Section 3 we present our main results. Theorem 3.1 shows how the approach in [Bar90] can be extended to the approximation of a scaled, time-changed random walk by a time-changed Brownian Motion. In Theorem 3.2 we apply Theorem 3.1 to look at the distance between a time-changed Poisson Process and a time-changed Brownian Motion. Theorem 3.4 shows how this can be extended to find the speed of convergence of a process whose dynamics resemble those of the Moran model with mutation to one whose dynamics resemble those of the Wright-Fisher diffusion with mutation. The bound obtained therein makes it possible to analyse the impact the mutation rates and the number of individuals have on the quality of the approximation and the interplay between those parameters. Section 5.2 provides proofs of the results presented in Section 2 and comments on how the proof of Theorem 3.4 may be adapted to find the speed of convergence in other examples satisfying the Stroock-Varadhan theory of diffusion approximation (see [SV79]).

In what follows,  $\|\cdot\|$  will always denote the sup norm and  $D = D[0, 1] = D([0, 1], \mathbb{R})$  will be the Skorokhod space of càdlàg real-valued functions on  $[0, 1]$ .

## 2. Space M

Let us define:

$$\|f\|_L := \sup_{w \in D[0,1]} \frac{|f(w)|}{1 + \|w\|^3},$$

and let  $L$  be the Banach space of the continuous functions  $f : D[0, 1] \rightarrow \mathbb{R}$  such that  $\|f\|_L < \infty$ . We now let  $M \subset L$  consist of the twice Fréchet differentiable functions  $f$ , such that:

$$\|D^2 f(w + h) - D^2 f(w)\| \leq k_f \|h\|, \quad (2.1)$$

for some constant  $k_f$ , uniformly in  $w, h \in D[0, 1]$ . By  $D^k f$  we mean the  $k$ -th Fréchet derivative of  $f$  and the  $k$ -linear norm  $B$  on  $L$  is defined to be  $\|B\| = \sup_{\{h: \|h\|=1\}} |B[h, \dots, h]|$ . Note the following:

**Lemma 2.1.** *For every  $f \in M$ ,  $\|f\|_M < \infty$ , where:*

$$\begin{aligned} \|f\|_M := & \sup_{w \in D[0,1]} \frac{|f(w)|}{1 + \|w\|^3} + \sup_{w \in D[0,1]} \frac{\|Df(w)\|}{1 + \|w\|^2} + \sup_{w \in D[0,1]} \frac{\|D^2f(w)\|}{1 + \|w\|} \\ & + \sup_{w, h \in D[0,1]} \frac{\|D^2f(w+h) - D^2f(w)\|}{\|h\|}. \end{aligned}$$

*Proof.* Note that for  $f \in M$  it is possible to find a constant  $K_f$  satisfying:

$$\begin{aligned} \text{A)} \quad & \|Df(w)\| \leq \|Df(w) - Df(0)\| + \|Df(0)\| \\ & \stackrel{\text{MVT}}{\leq} \|w\| \sup_{\theta \in [0,1]} \|D^2f(\theta w)\| + \|Df(0)\| \\ & \leq \|w\| \left[ \sup_{\theta \in [0,1]} (\|D^2f(\theta w) - D^2f(0)\| + \|D^2f(0)\|) \right] + \|Df(0)\| \\ \text{(2.1)} \quad & \leq \|w\| \left[ k_f \sup_{\theta \in [0,1]} \|\theta w\| + \|D^2f(0)\| \right] + \|Df(0)\| \\ & \leq k_f \|w\|^2 + \|D^2f(0)\| (1 \vee \|w\|^2) + \|Df(0)\| < K_f (1 + \|w\|^2); \\ \text{B)} \quad & \|D^2f(w)\| \leq \|D^2f(w) - D^2f(0)\| + \|D^2f(0)\| \\ \text{(2.1)} \quad & \leq k_f \|w\| + \|D^2f(0)\| < K_f (1 + \|w\|); \\ \text{C)} \quad & \left| f(w+h) - f(w) - Df(w)[h] - \frac{1}{2} D^2f(w)[h, h] \right| \leq K_f \|h\|^3, \end{aligned} \quad (2.2)$$

uniformly in  $w, h \in D$ , where the last inequality follows by Taylor's theorem and (2.1). Therefore:

$$\begin{aligned} \|f\|_M = & \sup_{w \in D} \frac{|f(w)|}{1 + \|w\|^3} + \sup_{w \in D} \frac{\|Df(w)\|}{1 + \|w\|^2} + \sup_{w \in D} \frac{\|D^2f(w)\|}{1 + \|w\|} \\ & + \sup_{w, h \in D} \frac{\|D^2f(w+h) - D^2f(w)\|}{\|h\|} < \infty. \end{aligned}$$

□

We now let  $M^0 \subset M$  be the class of functionals  $g \in M$  such that:

$$\|g\|_{M^0} := \|g\|_M + \sup_{w \in D} |g(w)| + \sup_{w \in D} \|Dg(w)\| + \sup_{w \in D} \|D^2g(w)\| < \infty.$$

This is Proposition 3.1 of [BJ09]:

**Proposition 2.2.** *Suppose that, for each  $n \geq 1$ , the random element  $Y_n$  of  $D[0, 1]$  is piecewise constant with intervals of constancy of length at least  $r_n$ . Let  $(Z_n)_{n \geq 1}$  be random elements of  $D[0, 1]$  converging weakly in  $D[0, 1]$ , with respect to the Skorokhod topology, to a random element  $Z \in C([0, 1], \mathbb{R})$ . If:*

$$|\mathbb{E}g(Y_n) - \mathbb{E}g(Z_n)| \leq C\tau_n \|g\|_{M^0} \quad (2.3)$$

for each  $g \in M^0$  and if  $\tau_n \log^2(1/r_n) \xrightarrow{n \rightarrow \infty} 0$ , then  $Y_n \Rightarrow Z$  in  $D[0, 1]$  (weakly in the Skorokhod topology).

A similar result holds when  $Y_n$  is a continuous-time Markov chain:

**Proposition 2.3.** *Suppose that, for each  $n \geq 1$ , the random element  $Y_n$  of  $D[0, 1]$  is a continuous-time Markov chain with mean holding time  $\frac{1}{\lambda_n} \rightarrow 0$ , identically distributed for each state. Let  $(Z_n)_{n \geq 1}$  be random elements of  $D[0, 1]$  converging weakly in  $D[0, 1]$ , with respect to the Skorokhod topology, to a random element  $Z \in C([0, 1], \mathbb{R})$ . Suppose further that:*

$$|\mathbb{E}g(Y_n) - \mathbb{E}g(Z_n)| \leq C\tau_n \|g\|_{M^0} \quad (2.4)$$

for each  $g \in M^0$  and that  $\tau_n \log^2((\lambda_n)^3) \xrightarrow{n \rightarrow \infty} 0$ . Then  $Y_n \Rightarrow Z$  in  $D[0, 1]$  (weakly in the Skorokhod topology).

We provide a proof of Proposition 2.3 in the Appendix.

### 3. Main results

Theorem 3.1 below is an extension of Theorem 1 in [Bar90] to the case of a time-changed scaled random walk:

**Theorem 3.1.** *Let  $X_1, X_2, \dots$  be i.i.d. with mean 0, variance 1 and finite third moment. Let  $s : [0, 1] \rightarrow [0, \infty)$  be a strictly increasing, continuous function with  $s(0) = 0$ . Define:*

$$Y_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor ns(t) \rfloor} X_i, \quad t \in [0, 1]$$

and let  $(Z(t), t \in [0, 1]) = (B(s(t)), t \in [0, 1])$ , where  $B$  is a standard Brownian Motion. Suppose that  $g \in M$ . Then:

$$\begin{aligned} |\mathbb{E}g(Y_n) - \mathbb{E}g(Z)| &\leq \|g\|_M \frac{30 + 54 \cdot 5^{1/3} s(1)}{\sqrt{\pi \log 2}} n^{-1/2} \sqrt{\log(2s(1)n)} \\ &\quad + \|g\|_M s(1) \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} s(1)^{3/2} \right) \mathbb{E}|X_1|^3 n^{-1/2} \\ &\quad + \|g\|_M \frac{2160}{\sqrt{\pi} (\log 2)^{3/2}} n^{-3/2} (\log(2s(1)n))^{3/2}. \end{aligned}$$

In Theorem 3.1 we do not claim that our bounds are sharp. Our bound in Theorem 3.1 is of the same order as the one obtained in the original case in [Bar90]. This result can also be extended in a straightforward way to instances in which the time change is random and independent of the step sizes of the random walk. We can obtain this by conditioning on the time change.

Theorem 3.2 below treats a time-changed Poisson process and can also be extended to random time changes, independent of the Poisson process of interest, by conditioning.

**Theorem 3.2.** Suppose that  $P$  is a Poisson process with rate 1 and  $S^{(n)} : [0, 1] \rightarrow [0, \infty)$  is a sequence of increasing continuous functions, such that  $S^{(n)}(0) = 0$ . Let  $S : [0, 1] \rightarrow [0, \infty)$  be also increasing and continuous. Let  $Z(t) = B(S(t)), t \in [0, 1]$  where  $B$  is a standard Brownian Motion and

$$\tilde{Y}_n(t) = \frac{P(nS^{(n)}(t)) - nS^{(n)}(t)}{\sqrt{n}}, \quad t \in [0, 1].$$

Then, for all  $g \in M$ :

$$\begin{aligned} |\mathbb{E}g(\tilde{Y}_n) - \mathbb{E}g(Z)| &\leq \|g\|_M \left\{ \left( 2 + \frac{27\sqrt{2}}{2\sqrt{\pi}} S(1) \right) \sqrt{\|S - S^{(n)}\|} + \frac{27\sqrt{2}}{2\sqrt{\pi}} \|S - S^{(n)}\|^{3/2} \right. \\ &\quad + n^{-1/2} \left[ \frac{30 + 54 \cdot 5^{1/3} S(1)}{\sqrt{\pi \log 2}} \sqrt{\log(2S(1)n)} \right. \\ &\quad + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} S^{(n)}(1)^{3/2} \right) S^{(n)}(1)(1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 2 \log n}{\log \log(n + 2)} \Big] \\ &\quad + n^{-1} \frac{9\sqrt{S^{(n)}(1)}}{2} \left( 1 + 3nS^{(n)}(1) \right)^{1/2} \left[ 4 + \frac{16701 + 128(\log n)^3}{(\log \log(n + 3))^3} \right]^{1/3} \\ &\quad \left. + n^{-3/2} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(2S(1)n))^{3/2} + 8 + \frac{33402 + 256(\log n)^3}{(\log \log(n + 3))^3} \right] \right\}. \end{aligned}$$

**Remark 3.3.** The bound in Theorem 3.2 goes to 0 as long as the time changes  $S^n \rightarrow S$  uniformly.

Theorem 3.4 below gives a bound on the speed of convergence of a process whose dynamics are similar to those of the Moran model with mutation to a process whose dynamics are similar to those of the Wright-Fisher diffusion with mutation. In the Moran model with mutation, in a population of size  $n$ , each individual carries a particular gene of one of the two forms:  $A$  and  $a$ . Each individual has exactly one parent and offspring inherit the genetic type of their parent. Now, at exponential rate  $\binom{n}{2}$  a pair of genes is sampled uniformly at random from the population. One of the pair is selected at random to die and the other one splits in two. In addition, every individual of type  $A$  changes its type independently at rate  $\nu_2$  and every individual of type  $a$  changes its type independently at rate  $\nu_1$ .

**Theorem 3.4.** Let  $X_n(t)$  be the proportion of type  $a$  genes in the population at time  $t \in [0, 1]$  under the Moran model with mutation rates  $\nu_1, \nu_2$ , as described above. Let  $(X(t), t \in [0, 1])$  denote the Wright-Fisher diffusion given by:

$$dX(t) = (\nu_2 - (\nu_1 + \nu_2)X(t))dt + \sqrt{X(t)(1 - X(t))}dB_t.$$

Suppose that:

$$\begin{cases} \mathbb{P}[X_n(t + \Delta t) - X_n(t) = \frac{1}{n}] = n^2 R_1^{(n)}(t) \Delta t \\ \mathbb{P}[X_n(t + \Delta t) - X_n(t) = -\frac{1}{n}] = n^2 R_{-1}^{(n)}(t) \Delta t \end{cases} \quad (3.1)$$

and:

$$M_n(t) = \frac{1}{n}P_1 \left( n^2 R_1^{(n)}(t) \right) - \frac{1}{n}P_{-1} \left( n^2 R_{-1}^{(n)}(t) \right), \quad (3.2)$$

for  $t \in [0, 1]$ , where  $P_1$  and  $P_{-1}$  are i.i.d. rate 1 Poisson processes, independent of  $X_n$ .

Suppose further that

$$M(t) = W \left( \int_0^t X(s)(1 - X(s))ds \right) + \int_0^t (\nu_2 - (\nu_1 + \nu_2)X(s))ds, \quad (3.3)$$

where  $W$  is a standard Brownian Motion, independent of  $X$ . Then, for any  $g \in M$ :

$$\begin{aligned} & |\mathbb{E}g(M_n) - \mathbb{E}g(M)| \\ & \leq \|g\|_M \left\{ \left( 18 + \nu_1^{1/2} + 47\nu_1^{3/4} + 31\nu_1^{3/2} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right) \right. \\ & \quad \cdot \left( 1.02 \cdot 10^6 + 425\nu_2^{1/2} + 623\nu_2 + 39\nu_2^{3/2} + 7\nu_2^{5/2} \right) \\ & \quad + (12 + 3\nu_2 + 3\nu_2^2 + 9\nu_2^3) \left( 1.02 \cdot 10^6 + 425\nu_1^{1/2} + 623\nu_1 + 39\nu_1^{3/2} + 7\nu_1^{5/2} \right) \\ & \quad \left. + 7 \left( \frac{1}{2}(1 + 2\nu_2)(\nu_1 + \nu_2) + 31(\nu_1 + \nu_2)^3 \right) \right\} n^{-1/4} \\ & \quad + \|g\|_M 2112 \left[ \left( 18 + \nu_1^{1/2} + 47\nu_1^{3/4} + 31\nu_1^{3/2} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right) (\log(n^2/4 + \nu_2 n))^{3/2} \right. \\ & \quad \left. + (12 + 3\nu_2 + 3\nu_2^2 + 9\nu_2^3) (\log(n^2/4 + \nu_1 n))^{3/2} \right] n^{-3}. \end{aligned}$$

**Remark 3.5.** If  $\nu_1 \geq 1$  and  $\nu_2 \geq 1$  then we can write:

$$\begin{aligned} & |\mathbb{E}g(M_n) - \mathbb{E}g(M)| \\ & \leq \|g\|_M \left[ (18 + 79\nu_1^{3/2} + 13\nu_2^3)(1.02 \cdot 10^6 + 1094\nu_2^{5/2}) \right. \\ & \quad + (12 + 15\nu_2^3)(1.02 \cdot 10^6 + 1094\nu_2^{5/2}) \\ & \quad + 7(31.5\nu_1^3 + 32.5\nu_2^3 + \nu_1\nu_2(1 + 93\nu_1 + 93\nu_2)) \left. \right] n^{-1/4} \\ & \quad + \|g\|_M 2112(18 + 79\nu_1^{3/2} + 13\nu_2^3)n^{-3} (\log(n^2/4 + \nu_1 n))^{3/2} \\ & \quad + \|g\|_M 2112(12 + 15\nu_2^3)n^{-3} (\log(n^2/4 + \nu_1 n))^{3/2}. \end{aligned}$$

The approximation gets worse as the mutation rates increase and the number of individuals decreases. Should we want to make the mutation rates depend on  $n$  and be of the same order, we will require them to be  $o(n^{1/22})$  in order for the bound to converge to 0 as  $n \rightarrow \infty$ .

**Remark 3.6.** Note that the Moran model  $X_n$  jumps up by  $\frac{1}{n}$  with intensity  $n^2 R_1^{(n)}(t)$  and down by  $\frac{1}{n}$  with intensity  $n^2 R_{-1}^{(n)}(t)$ , as defined in (3.1). Using the ideas from [Kur12] we can write  $X_n$  in the following form:

$$X_n(t) = \frac{1}{n}\tilde{P}_1 \left( n^2 R_1^{(n)}(t) \right) - \frac{1}{n}\tilde{P}_{-1} \left( n^2 R_{-1}^{(n)}(t) \right)$$

for some independent rate 1 Poisson processes  $\tilde{P}_1$  and  $\tilde{P}_{-1}$  which are, however, not independent of  $R_1^{(n)}$  and  $R_{-1}^{(n)}$ . In Theorem 3.4 we consider a similar process, given by (3.2), whose definition uses Poisson processes  $P_1$  and  $P_{-1}$  independent of  $R_1^{(n)}$  and  $R_{-1}^{(n)}$ .

Similarly, the diffusive part of the Wright-Fisher diffusion may be expressed as a time-changed Brownian Motion  $\tilde{W} \left( \int_0^t X(s)(1-X(s))ds \right)$ . However, we consider a process given by (3.3) with the assumption that  $W$  is independent of  $X$ .

We will appeal to Theorem 3.2 to obtain the bounds. The time changes we apply in this case are random and therefore we will first condition on them.

**Remark 3.7.** The Moran model  $M_n$  converges weakly to the Wright-Fisher diffusion  $M$ , which can be proved using, for instance, the Stroock-Varadhan theory of diffusion approximation [SV79]. However, our paper does not provide the tools necessary for treating this convergence with Stein's method and obtaining bounds on its rate.

A key idea used in the proof will be the Donnelly-Kurtz look-down construction coming from [DK96] and described in Chapter 2.10 of [Eth12]. The  $n$ -particle look-down process is denoted by a vector  $(\psi_1(t), \dots, \psi_n(t))$  with each index representing a "level" and each of  $\psi_i$ 's representing the type of the individual at level  $i$  at time  $t$ . Individual at level  $k$  is equipped with an exponential clock with rate  $k-1$ , independent of other individuals, and at the times the clock rings it "looks down" at a level chosen uniformly at random from  $\{1, \dots, k-1\}$  and adopts the type of the individual at that level. In addition, the type of each individual evolves according to the mutation process. A comparison of the generators of the Moran model and the look-down process shows that, as long as the two are started from the same initial exchangeable condition, they produce the same distribution of types in the population. In addition, it may be shown that the Wright-Fisher diffusion may be represented as the proportion of type  $a$  individuals in the population, in which the types are distributed according to the infinite-particle look-down process. The corresponding Moran model is then the proportion of type  $a$  individuals among the ones located on the first  $n$  levels. Due to exchangeability, we may then describe the Moran model  $X_n$  at a fixed time, as depending on the Wright-Fisher diffusion  $X$  in the following way:  $nX_n(t) \sim \text{Binomial}(n, X(s))$ .

**Remark 3.8.** Our bound in Theorem 3.4 is sufficient to conclude that process  $M_n$  converges to process  $M$  weakly on compact intervals. This follows from Proposition 2.3. Using the notation therein, in this case,  $\tau_n = n^{-1/4}$  and  $\lambda_n = \frac{n(n-1)}{2}$ .



#### 4. Setting up Stein's method

Let us first define:

$$A_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor ns(1) \rfloor} Z_i \mathbb{1}_{[i/n, s(1)]}(s(t)) = n^{-1/2} \sum_{i=1}^{\lfloor ns(1) \rfloor} Z_i \mathbb{1}_{[s^{-1}(i/n), 1]}(t), \quad (4.1)$$

for  $Z_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$ . In a preparation for the proof of Theorem 3.1, we will apply Stein's method to find the distance between  $A_n$  and  $Y_n$ .

##### 4.1. The Stein equation

We first note that if  $U_1, U_2, \dots$  are i.i.d. Ornstein-Uhlenbeck processes with stationary law  $\mathcal{N}(0, 1)$ , then defining:

$$W_n(t, u) = n^{-1/2} \sum_{i=1}^{\lfloor ns(t) \rfloor} U_i(u), \quad u \geq 0, t \in [0, 1],$$

we obtain that the law of  $A_n$  is stationary for  $(W_n(\cdot, u))_{u \geq 0}$ . Denote the generator of  $(W_n(\cdot, u))_{u \geq 0}$  by  $\mathcal{A}_n$ . By properties of stationary distributions,  $\mathbb{E}_\mu \mathcal{A}_n f = 0$  for all  $f \in \text{Domain}(\mathcal{A}_n)$  if and only if  $\mu = \mathcal{L}(A_n)$ . Therefore, we can treat

$$\mathcal{A}_n f = g - \mathbb{E}g(A_n) \quad (4.2)$$

as our Stein equation.

In the next subsection, for any  $g$  from a suitable class of functions, we will find an  $f$  satisfying equation (4.2). Then, in the sequel, we will find a bound on  $|\mathbb{E} \mathcal{A}_n f(Y_n)|$ , which will readily give us a bound on  $|\mathbb{E}g(Y_n) - \mathbb{E}g(A_n)|$ .

**Proposition 4.1.** *The generator  $\mathcal{A}_n$  of the process  $(W_n(\cdot, u))_{u \geq 0}$  acts on any  $f \in M$  in the following way:*

$$(\mathcal{A}_n f)(w) := -Df(w)[w] + \mathbb{E} D^2 f(w) \left[ A_n^{(2)} \right].$$

Before we prove this result, we need a lemma:

**Lemma 4.2.** *Letting  $\mathcal{F}_{n,u} = \sigma(W_n(\cdot, v), v \leq u)$ , we have:*

$$W_n(\cdot, u + v) - e^{-v} W_n(\cdot, u) \stackrel{\mathcal{D}}{=} \sigma(v) A_n(\cdot).$$

*Proof.* We first note that for each  $i \geq 1$  we can construct independent standard Brownian Motions  $B_i$  such that  $(X_i(u), u \geq 0) = (e^{-u} B_i(e^{2u}), u \geq 0)$ . Then:

$$\begin{aligned}
W_n(\cdot, u+v) - e^{-v}W(\cdot, u) &= n^{-1/2} \sum_{k=1}^{\lfloor ns(\cdot) \rfloor} U_k(u+v) - e^{-v}n^{-1/2} \sum_{k=1}^{\lfloor ns(\cdot) \rfloor} U_k(u) \\
&\stackrel{\mathcal{D}}{=} n^{-1/2} e^{-(u+v)} \sum_{k=1}^{\lfloor ns(\cdot) \rfloor} \left[ B_k(e^{2(u+v)}) - B_k(e^{2u}) \right] \\
&\stackrel{\mathcal{D}}{=} n^{-1/2} \sigma(v) \sum_{k=1}^{\lfloor ns(\cdot) \rfloor} Z_k = \sigma(v) A_n(\cdot).
\end{aligned}$$

□

*Proof of Proposition 4.1.* Note that the semigroup of  $(W_n(\cdot, u))_{u \geq 0}$ , acting on  $L$ , is defined by:

$$(T_{n,u}f)(w) := \mathbb{E}[f(W_n(\cdot, u)) | W_n(\cdot, 0) = w]$$

and by Lemma 4.2 we readily obtain that:

$$(T_{n,u}f)(w) = \mathbb{E}[f(we^{-u} + \sigma(u)A_n(\cdot))]. \quad (4.3)$$

We can define the generator by:  $\mathcal{A}_n = \lim_{u \searrow 0} \frac{T_{n,u} - I}{u}$ . We also have that for  $f \in M$ :

$$\begin{aligned}
&\left| (T_{n,u}f)(w) - f(w) - \mathbb{E}Df(w)[\sigma(u)A_n - w(1 - e^{-u})] \right. \\
&\quad \left. - \frac{1}{2}\mathbb{E}D^2f(w)[\{\sigma(u)A_n - w(1 - e^{-u})\}^{(2)}] \right| \\
&\stackrel{(2.2)}{\leq} K_f \mathbb{E}\|\sigma(u)A_n - w(1 - e^{-u})\|^3 \\
&\leq 4K_f [\sigma^3(u)\mathbb{E}\|A_n\|^3 + (1 - e^{-u})^3\|w\|^3] \leq K_3(1 + \|w\|^3)u^{3/2}
\end{aligned}$$

for a constant  $K_3$  depending only on  $f$ , where the last inequality follows from the fact that for  $u \geq 0$ ,  $\sigma^3(u) \leq 3u^{3/2}$  and  $(1 - e^{-u})^3 \leq u^{3/2}$ . So:

$$\begin{aligned}
&\left| (T_{n,u}f - f)(w) + uDf(w)[w] - u\mathbb{E}D^2f(w)[A_n^{(2)}] \right| \\
&\leq \left| (T_{n,u}f)(w) - f(w) - \mathbb{E}Df(w)[\sigma(u)A_n - w(1 - e^{-u})] \right. \\
&\quad \left. - \frac{1}{2}\mathbb{E}D^2f(w)[\{\sigma(u)A_n - w(1 - e^{-u})\}^{(2)}] \right| + |\sigma(u)\mathbb{E}Df(w)[A_n]| \\
&\quad + |(u - 1 + e^{-u})Df(w)[w]| + \left| \left( \frac{\sigma^2(u)}{2} - u \right) \mathbb{E}D^2f(w)[A_n^{(2)}] \right| \\
&\quad + \left| \frac{(1 - e^{-u})^2}{2} D^2f(w)[w^{(2)}] \right| + |\sigma(u)(1 - e^{-u})\mathbb{E}D^2f(w)[A_n, w]|
\end{aligned}$$

$$\begin{aligned}
&\leq K_1 (1 + \|w\|^3) u^{3/2} + |\sigma(u) \mathbb{E} Df(w)[A_n]| + |u - 1 + e^{-u}| \|Df(w)\| \|w\| \\
&\quad + \left| \frac{\sigma^2(u)}{2} - u \right| \|D^2 f(w)\| \mathbb{E} \|A_n\|^2 \\
&\quad + \frac{(1 - e^{-u})^2}{2} \|D^2 f(w)\| \|w\|^2 + \sigma(u)(1 - e^{-u}) \|D^2 f(w)\| \|w\| \mathbb{E} \|A_n\| \\
(2.2) \quad &\leq K_1 (1 + \|w\|^3) u^{3/2} + |\sigma(u) \mathbb{E} Df(w)[A_n]| + |u - 1 + e^{-u}| K_f (1 + \|w\|^2) \|w\| \\
&\quad + \left| \frac{\sigma^2(u)}{2} - u \right| K_f (1 + \|w\|) \mathbb{E} \|A_n\|^2 + \frac{(1 - e^{-u})^2}{2} K_f (1 + \|w\|) \|w\|^2 \\
&\quad + \sigma(u)(1 - e^{-u})(1 + \|w\|) \|w\| \mathbb{E} \|A_n\| \\
&\leq 3u^{3/2} (K_1 (1 + \|w\|^3) + K_f (1 + \|w\|^2) \|w\| + K_f (1 + \|w\|) \mathbb{E} \|A_n\|^2 \\
&\quad + K_f (1 + \|w\|) \|w\|^2 + (1 + \|w\|) \|w\| \mathbb{E} \|A_n\|) + |\sigma(u) \mathbb{E} Df(w)[A_n]| \\
&\leq K_4 (1 + \|w\|^3) u^{3/2}, \tag{4.4}
\end{aligned}$$

for some constant  $K_4$  depending only on  $f$ . The last inequality follows from the fact that:

$$\begin{aligned}
\mathbb{E} Df(w)[A_n] &= \mathbb{E} Df(w) \left[ n^{-1/2} \sum_{k=1}^{\lfloor ns(1) \rfloor} Z_k \mathbb{1}_{[s^{-1}(k/n), 1]} \right] \\
&= n^{-1/2} \sum_{k=1}^{\lfloor ns(1) \rfloor} Df(w) \left[ \mathbb{1}_{[s^{-1}(k/n), 1]} \right] \mathbb{E}[Z_k] \\
&= 0.
\end{aligned}$$

It follows that for any  $f \in M$ :

$$\mathcal{A}_n f(w) = \lim_{u \searrow 0} \frac{T_{n,u} f(w) - f(w)}{u} = -Df(w)[w] + \mathbb{E} D^2 f(w) \left[ A_n^{(2)} \right].$$

□

#### 4.2. Solving Stein's equation

**Proposition 4.3.** *Suppose that  $g \in M$  satisfies  $\mathbb{E} g(A_n) = 0$ . Then the equation:  $\mathcal{A}_n f = g$  is solved by:*

$$f = \phi_n(g) = - \int_0^\infty T_{n,u} g du \tag{4.5}$$

for  $T_{n,u}$  defined by (4.3). Furthermore,  $\phi_n(g) \in M$  and the following inequalities hold:

$$\begin{aligned} A) \quad & \|D\phi_n(g)(w)\| \leq \left(1 + \frac{4}{3}\mathbb{E}\|A_n\|^2\right) \|g\|_M(1 + \|w\|^2), \\ B) \quad & \|D^2\phi_n(g)(w)\| \leq \left(\frac{1}{2} + \frac{\mathbb{E}\|A_n\|}{3}\right) \|g\|_M(1 + \|w\|), \\ C) \quad & \|D^2\phi_n(g)(w+h) - D^2\phi(g)(w)\| \leq \frac{\|g\|_M}{3}\|h\|. \end{aligned} \quad (4.6)$$

*Proof.* The proof will follow the procedure used to prove [KDV17, Proposition 1].

**Step 1:** First, we show that  $\phi_n(g) \in M$  and that (4.6) holds. Assume  $g \in M$  and  $\mathbb{E}[g(A_n)] = 0$ . Note that if, for instance:

$$|g(w) - g(x)| \leq C_g(1 + \|w\|^2 + \|x\|^2)\|w - x\| \quad (4.7)$$

uniformly in  $w, x \in D[0, 1]$  then:

$$\begin{aligned} & \lim_{t \rightarrow \infty} \int_0^t |T_{n,u}g(w)| du = \lim_{t \rightarrow \infty} \int_0^t |\mathbb{E}g(we^{-u} + \sigma(u)A_n)| du \\ & \leq \lim_{t \rightarrow \infty} \left[ \int_0^t |\mathbb{E}[g(we^{-u} + \sigma(u)A_n) - g(\sigma(u)A_n)]| du + \int_0^t |\mathbb{E}[g(\sigma(u)A_n) - g(A_n)]| du \right] \\ & \leq C_g \lim_{t \rightarrow \infty} \left[ \int_0^t \mathbb{E}[(1 + \|e^{-u}w + \sigma(u)A_n\|^2 + \sigma^2(u)\|A_n\|^2)e^{-u}\|w\|] du \right. \\ & \quad \left. + \int_0^t \mathbb{E}[(1 + (\sigma^2(u) + 1)\|A_n\|^2)|(\sigma(u) - 1)A_n|] du \right] \\ & \leq C_g \lim_{t \rightarrow \infty} \left[ \int_0^t [e^{-u}\|w\| + 2e^{-3u}\|w\|^3 + 3\sigma^2(u)e^{-u}\|w\|\mathbb{E}\|A_n\|^2] du \right. \\ & \quad \left. + \int_0^t (\sigma(u) - 1)\mathbb{E}[(1 + (\sigma^2(u) + 1)\|Z\|^2)|\|A_n\|] du \right] \\ & \leq C(1 + \|w\|^3), \end{aligned} \quad (4.8)$$

for some constant  $C$ . For such  $g$ ,  $f = \phi_n(g) = -\int_0^\infty T_{n,u}g du$  exists. Note that every  $g \in M$  satisfies condition (4.7) because for such  $g$ :

$$\begin{aligned} |g(w) - g(x)| & \stackrel{(2.2)}{\leq} K_g\|w - x\|^3 + \left| Dg(x)[w - x] + \frac{1}{2}D^2g(x)[w - x, w - x] \right| \\ & \leq K_g\|w - x\|^3 + \|Dg(x)\|\|w - x\| + \frac{1}{2}\|D^2g(x)\|\|w - x\|^2 \\ & \stackrel{(2.2)}{\leq} K_g\|w - x\| \left( \|w - x\|^2 + 1 + \|x\|^2 + \frac{1}{2}\|w - x\|(1 + \|x\|) \right) \\ & \leq K_g\|w - x\| \left( 2\|w\|^2 + 2\|x\|^2 + 1 + \|x\|^2 + \frac{1}{2}(\|w\| + \|x\| + \|w\|\|x\| + \|x\|^2) \right) \\ & \leq C_g(1 + \|w\|^2 + \|x\|^2)\|w - x\| \end{aligned}$$

uniformly in  $w, x$  because  $\|w\| \leq 1 + \|w\|^2$ ,  $\|x\| \leq 1 + \|x\|^2$  and  $\|w\|\|x\| \leq \|w\|^2 + \|x\|^2$ .

Now take  $g \in M$ , such that  $\mathbb{E}g(A_n) = 0$  and note that for  $\phi_n$  defined in (4.5) we get:

$$\phi_n(g)(w+h) - \phi_n(g)(w) \stackrel{(4.3)}{=} -\mathbb{E} \int_0^\infty [g((w+h)e^{-u} + \sigma(u)A_n) - g(we^{-u} + \sigma(u)A_n)] du$$

and so dominated convergence (which can be applied because of (4.8)) gives:

$$D^k \phi_n(g)(w) = -\mathbb{E} \int_0^\infty e^{-ku} D^k g(we^{-u} + \sigma(u)A_n) du, \quad k = 1, 2. \quad (4.9)$$

Also, observe that:

$$\begin{aligned} \text{A)} \quad & \|D\phi_n(g)(w)\| \\ & \leq \int_0^\infty e^{-u} \mathbb{E} \|Dg(we^{-u} + \sigma(u)A_n)\| du \\ & \stackrel{(2.2)}{\leq} \|g\|_M \int_0^\infty e^{-u} (1 + \mathbb{E} \|we^{-u} + \sigma(u)A_n\|^2) du \\ & \leq \|g\|_M \int_0^\infty (e^{-u} + 2\|w\|^2 e^{-3u} + 2\mathbb{E} \|A_n\|^2 (e^{-u} - e^{-3u})) du \\ & \leq \left(1 + \frac{4}{3} \mathbb{E} \|A_n\|^2\right) \|g\|_M (1 + \|w\|^2), \\ \text{B)} \quad & \|D^2\phi_n(g)(w)\| \\ & \leq \int_0^\infty e^{-2u} \mathbb{E} \|D^2g(we^{-u} + \sigma(u)A_n)\| du \\ & \stackrel{(2.2)}{\leq} \|g\|_M \int_0^\infty e^{-2u} (1 + \mathbb{E} \|we^{-u} + \sigma(u)A_n\|) du \\ & \leq \|g\|_M \int_0^\infty e^{-2u} (1 + e^{-u} \|w\| + \sigma(u) \mathbb{E} \|A_n\|) du \\ & \leq \|g\|_M \left(\frac{1}{2} + \frac{\|w\|}{3} + \frac{\mathbb{E} \|A_n\|}{3}\right) \\ & \leq \left(\frac{1}{2} + \frac{\mathbb{E} \|A_n\|}{3}\right) \|g\|_M (1 + \|w\|), \\ \text{C)} \quad & \|D^2\phi_n(g)(w+h) - D^2\phi_n(g)(w)\| \\ & \leq \left\| \mathbb{E} \int_0^\infty e^{-2u} D^2g((w+h)e^{-u} + \sigma(u)A_n) - e^{-2u} D^2g(we^{-u} + \sigma(u)A_n) du \right\| \\ & \stackrel{g \in M, (2.1)}{\leq} \|g\|_M \int_0^\infty e^{-2u} e^{-u} \|h\| du = \frac{\|g\|_M}{3} \|h\|, \end{aligned} \quad (4.10)$$

uniformly in  $g \in M$ , which proves (4.6). It follows by (4.8) and (4.10) that  $\phi_n(g) \in M$ . Thus,  $\phi_n(g)$  is in the domain of the semigroup and of  $\mathcal{A}_n$ . Also, similarly, for any  $t > 0$ ,  $\int_0^t T_{n,u} g du \in M$ .

**Step 2:** We now show that for all  $t > 0$  and for all  $g \in M$ :

$$T_{n,t}g - g = \mathcal{A}_n \left( \int_0^t T_{n,u}g du \right). \quad (4.11)$$

We will prove it by following the steps of the proof of Proposition 1.5 on p. 9 of [EK86]. Observe that for all  $w \in D[0, 1]$  and  $h > 0$ :

$$\begin{aligned} & \frac{1}{h} [T_{n,h} - I] \int_0^t T_{n,u}g(w) du = \frac{1}{h} \int_0^t [T_{n,u+h}g(w) - T_{n,u}g(w)] du \\ &= \frac{1}{h} \int_t^{t+h} T_{n,u}g(w) du - \frac{1}{h} \int_0^h T_{n,u}g(w) du \\ &\stackrel{(4.3)}{=} \frac{1}{h} \int_t^{t+h} \mathbb{E}[g(we^{-u} + \sigma(u)A_n)] du - \frac{1}{h} \int_0^h \mathbb{E}[g(we^{-u} + \sigma(u)A_n)] du. \end{aligned} \quad (4.12)$$

Taking  $h \rightarrow 0$  on the left-hand side gives  $\mathcal{A}_n \left( \int_0^t T_{n,u}g(w) du \right)$  because  $\int_0^t T_{n,u}g(w) du$  is in the domain of  $\mathcal{A}_n$ , as shown in **Step 1**. In order to analyse the right-hand side note that:

$$\begin{aligned} & \left| \frac{1}{h} \int_0^h \mathbb{E}[g(we^{-u} + \sigma(u)A_n)] - g(w) du \right| \\ &\stackrel{\text{MVT}}{\leq} \frac{1}{h} \int_0^h \mathbb{E} \left[ \|w(e^{-u} - 1) + \sigma(u)A_n\| \sup_{c \in [0,1]} \|Dg(cw + (1-c)(we^{-u} + \sigma(u)A_n))\| \right] du \\ &\leq \frac{\|g\|_M}{h} \int_0^h \mathbb{E} \left[ (\|w\|(1 - e^{-u}) + \sigma(u)\|A_n\|) (1 + 3\|w\|^2 + 3\|w\|^2 e^{-2u} + 3\sigma^2(u)\|A_n\|^2) \right] du \\ &= \frac{\|g\|_M}{h} \mathbb{E} \left\{ (1 + 3\|w\|^2 + 3\|A_n\|^2) (\|w\|(-1 + h + \cosh(h) - \sinh(h)) \right. \\ &\quad \left. + \|A_n\|e^{-h}(-\sqrt{e^{2h} - 1} + e^h(h + \log(1 + e^{-h}\sqrt{-1 + e^{2h}}))) \right. \\ &\quad \left. + 3\|w\|(\|w\|^2 - \|A_n\|^2) \left( \frac{e^{-3h}}{6}(e^h - 1)^2(e^h + 2) \right) \right. \\ &\quad \left. + 3(\|w\|^2\|A_n\| - \|A_n\|^3) \frac{1}{3} \left( \sqrt{1 - e^{-2h}} - \sqrt{e^{-6h}(e^{2h} - 1)} \right) \right\} \xrightarrow{h \rightarrow 0} 0. \end{aligned} \quad (4.13)$$

Similarly:

$$\left| \frac{1}{h} \int_t^{t+h} \mathbb{E}[g(we^{-u} + \sigma(u)A_n)] du - \mathbb{E}[g(we^{-t} + \sigma(t)A_n)] \right| \xrightarrow{h \rightarrow 0} 0$$

Therefore, as  $h \rightarrow 0$ , the right-hand side of (4.12) converges to  $T_{n,t}g - g$ , which proves (4.11).

**Step 3:** We note that for any  $h > 0$  and for any  $f \in M$ :

$$\frac{1}{h} [T_{n,s+h}f - T_{n,s}f] = T_{n,s} \left[ \frac{T_{n,h} - I}{h} f \right]$$

and therefore for any  $w \in D[0, 1]$ , by dominated convergence (which can be applied because of (4.4)):

$$\begin{aligned} \left( \frac{d}{ds} \right)^+ T_{n,s}f(w) &= \lim_{h \searrow 0} T_{n,s} \left[ \frac{T_{n,h} - I}{h} f(w) \right] = \lim_{h \searrow 0} \mathbb{E} \left[ \frac{T_{n,h} - I}{h} f(we^{-s} + \sigma(s)A_n) \right] \\ &= \mathbb{E} \left[ \lim_{h \searrow 0} \frac{T_{n,h} - I}{h} f(we^{-s} + \sigma(s)A_n) \right] = T_{n,s} \mathcal{A}_n f(w) \end{aligned}$$

and, similarly, for  $s > 0$ ,  $\left( \frac{d}{ds} \right)^- T_{n,s}f = T_{n,s} \mathcal{A}_n f$  because:

$$\begin{aligned} &\lim_{h \searrow 0} \frac{1}{-h} [T_{n,s-h}f - T_{n,s}f](w) - T_{n,s} \mathcal{A}_n f(w) \\ &= \lim_{h \searrow 0} T_{n,s-h} \left[ \left( \frac{T_{n,h} - I}{h} - \mathcal{A}_n \right) f \right](w) + \lim_{h \searrow 0} (T_{n,s-h} - T_{n,s}) \mathcal{A}_n f(w) \\ &= \lim_{h \searrow 0} \mathbb{E} \left[ \left( \frac{T_{n,h} - I}{h} - \mathcal{A}_n \right) f(we^{-s+h} + \sigma(s-h)A_n) \right] \\ &\quad + \lim_{h \searrow 0} \mathbb{E} [\mathcal{A}_n f(we^{-s+h} + \sigma(s-h)A_n) - \mathcal{A}_n f(we^{-s} + \sigma(s)A_n)] \\ &\stackrel{(4.4)}{=} 0 \end{aligned}$$

again, by dominated convergence. It can be applied because of (4.4) and the observation that:

$$\begin{aligned} &|\mathcal{A}_n f(we^{-s+h} + \sigma(s-h)A_n) - \mathcal{A}_n f(we^{-s} + \sigma(s)A_n)| \\ &= |-Df(we^{-s+h} + \sigma(s-h)A_n)[we^{-s+h} + \sigma(s-h)A_n] \\ &\quad + \sum_{i=1}^n D^2 f(we^{-s+h} + \sigma(s-h)A_n)[\mathbb{1}_{[i/n, 1]}^{(2)}] \\ &\quad - Df(we^{-s} + \sigma(s)A_n)[we^{-s} + \sigma(s)A_n] + \sum_{i=1}^n D^2 f(we^{-s} + \sigma(s)A_n)[\mathbb{1}_{[i/n, 1]}^{(2)}] \Big| \\ &\leq \|f\|_M (1 + \|we^{-s+h} + \sigma(s-h)A_n\|^2) \|we^{-s+h} + \sigma(s-h)A_n\| \\ &\quad + n\|f\|_M (1 + \|we^{-s+h} + \sigma(s-h)A_n\|) \\ &\quad + \|f\|_M (1 + \|we^{-s} + \sigma(s)A_n\|) \|we^{-s} + \sigma(s)A_n\| + \|f\|_M n (1 + \|we^{-s} + \sigma(s)A_n\|) \\ &\leq \|f\|_M (1 + 2\|w\|^2 e^{-2s+2} + 2\sigma^2(s-1)\|A_n\|^2) (\|w\|e^{-s+1} + \sigma(s-1)\|A_n\|) \\ &\quad + n\|f\|_M (1 + \|we^{-s+1} + \sigma(s-1)A_n\|) \\ &\quad + \|f\|_M (1 + \|we^{-s} + \sigma(s)A_n\|) \|we^{-s} + \sigma(s)A_n\| + n\|f\|_M (1 + \|we^{-s} + \sigma(s)A_n\|) \end{aligned}$$

for all  $h \in [0, 1]$ .

Thus, for all  $w \in D[0, 1]$  and  $s > 0$ :  $\frac{d}{ds}T_{n,s}f(w) = T_{n,s}\mathcal{A}_nf(w)$  and so, by the Fundamental Theorem of Calculus:

$$T_{n,r}f(w) - f(w) = \int_0^r T_{n,s}\mathcal{A}_nf(w)ds. \quad (4.14)$$

Applying (4.14) to  $f = \int_0^t T_{n,u}gdu$  we obtain:

$$T_{n,r} \int_0^t T_{n,u}g(w)du - \int_0^t T_{n,u}g(w)du = \int_0^r T_{n,s}\mathcal{A}_n \left( \int_0^t T_{n,u}g(w)du \right) ds$$

Now, we take  $t \rightarrow \infty$ . We apply dominated convergence, which is allowed because of (4.8) and the following bound for  $h_{t,u}(w) = \int_0^t T_{n,u}g(w)du$ :

$$\begin{aligned} & |\mathcal{A}_nh_{t,u}(w)| \\ & \leq \int_0^t \mathbb{E} \left| e^{-u} Dg(we^{-u} + \sigma(u)A_n)[w] \right| du \\ & \quad + \sum_{i=1}^n \int_0^t \mathbb{E} \left| e^{-2u} D^2g(we^{-u} + \sigma(u)A_n) \left[ \mathbb{1}_{[i/n, 1]}^{(2)} \right] \right| du \\ & \leq \int_0^\infty \mathbb{E} \left| e^{-u} Dg(we^{-u} + \sigma(u)A_n)[w] \right| du \\ & \quad + \sum_{i=1}^n \int_0^\infty \mathbb{E} \left| e^{-2u} D^2g(we^{-u} + \sigma(u)A_n) \left[ \mathbb{1}_{[i/n, 1]}^{(2)} \right] \right| du \\ & \stackrel{(4.10)^{A,B}}{\leq} \left( 1 + \frac{4}{3} \mathbb{E} \|A_n\|^2 \right) \|g\|_M (1 + \|w\|^2) \|w\| + n \left( \frac{1}{2} + \frac{\mathbb{E} \|A_n\|}{3} \right) \|g\|_M (1 + \|w\|), \end{aligned}$$

where the first inequality follows again by dominated convergence applied because of (4.8) in order to exchange integration and differentiation in a way similar to (4.9). As a result, we obtain:

$$\begin{aligned} T_{n,r} \int_0^\infty T_{n,u}g(w)du - \int_0^\infty T_{n,u}g(w)du &= \int_0^r T_{n,s} \lim_{t \rightarrow \infty} \mathcal{A}_n \left( \int_0^t T_{n,u}g(w)du \right) ds \\ &\stackrel{(4.11)}{=} - \int_0^r T_{n,s}g(w)ds. \end{aligned}$$

Now, dividing both sides by  $r$  and taking  $r \rightarrow 0$ , we obtain:

$$\begin{aligned} \mathcal{A}_n \left( \int_0^\infty T_{n,u}g(w)du \right) &= - \lim_{r \rightarrow 0} \frac{1}{r} \int_0^r T_{n,s}g(w)ds \\ &= - \lim_{r \rightarrow 0} \left[ \frac{1}{r} \int_0^r \mathbb{E} g(we^{-s} + \sigma(s)A_n)ds \right] \\ &\stackrel{(4.13)}{=} -g(w), \end{aligned}$$

which finishes the proof.  $\square$



**Remark 4.4.** It is an easy consequence of Propositions 4.1 and 4.3 that for  $g \in M$ :

$$\mathcal{A}_n \phi_n(g)(w) = -D\phi_n(g)(w)[w] + \mathbb{E} D^2 \phi_n(g)(w) \left[ A_n^{(2)} \right].$$

## 5. Proofs of Theorems 3.1, 3.2 and 3.4

### 5.1. Proof of Theorem 3.1

#### 5.1.1. Discretisation of Brownian Motion

Let  $A_n$  be as in (4.1). Now, note that we can first realise  $B$  and then set  $A_n(t) = B\left(\frac{\lfloor ns(t) \rfloor}{n}\right)$  for  $t \in [0, 1]$  so that:

$$\sup_{t \in [0, 1]} |A_n(t) - Z(t)| = \sup_{t \in [0, 1]} \left| B\left(\frac{\lfloor ns(t) \rfloor}{n}\right) - B(s(t)) \right| = \sup_{t \in [0, s(1)]} \left| B(t) - B\left(\frac{\lfloor nt \rfloor}{n}\right) \right|.$$

By Lemma 3 of [FN10] we get:

$$\begin{aligned} \text{A) } \mathbb{E} \|A_n - Z\| &\leq \mathbb{E} \left[ \sup_{t, s \in [0, s(1)], |t-s| \leq \frac{1}{n}} |B(t) - B(s)| \right] \\ &\leq \frac{5}{\sqrt{\pi}} \cdot \frac{6}{\sqrt{\log 2}} n^{-1/2} \sqrt{\log(2s(1)n)} \\ \text{B) } \mathbb{E} \|A_n - Z\|^2 &\leq \mathbb{E} \left[ \left( \sup_{t, s \in [0, s(1)], |t-s| \leq \frac{1}{n}} |B(t) - B(s)| \right)^2 \right] \\ &\leq \frac{5}{2} \cdot \left( \frac{6}{\sqrt{\log 2}} \right)^2 n^{-1} \log(2s(1)n) \\ \text{C) } \mathbb{E} \|A_n - Z\|^3 &\leq \mathbb{E} \left[ \left( \sup_{t, s \in [0, s(1)], |t-s| \leq \frac{1}{n}} |B(t) - B(s)| \right)^3 \right] \\ &\leq \frac{5}{\sqrt{\pi}} \cdot \left( \frac{6}{\sqrt{\log 2}} \right)^3 n^{-3/2} (\log(2s(1)n))^{3/2} \end{aligned} \quad (5.1)$$

and therefore we obtain, for any  $g \in M$ :

$$\begin{aligned} &|\mathbb{E}g(A_n) - \mathbb{E}g(Z)| \\ &\stackrel{\text{MVT}}{\leq} \mathbb{E} \left[ \sup_{c \in [0, 1]} \|Dg((1-c)Z + cA_n)\| \|Z - A_n\| \right] \\ &\leq \|g\|_M \mathbb{E} \left[ \sup_{c \in [0, 1]} (1 + \|Z + c(A_n - Z)\|^2) \|Z - A_n\| \right] \\ &\leq \|g\|_M \mathbb{E} \left[ (1 + 2\|Z\|^2 + 2\|Z - A_n\|^2) \|Z - A_n\| \right] \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Hölder}}{\leq} \|g\|_M \left\{ \mathbb{E}\|Z - A_n\| + 2\mathbb{E}\|Z - A_n\|^3 + 2 \left( \mathbb{E}\|Z\|^3 \right)^{2/3} \left( \mathbb{E}\|A_n - Z\|^3 \right)^{1/3} \right\} \\
& \stackrel{(5.1)}{\leq} \|g\|_M \left\{ \frac{5}{\sqrt{\pi}} \cdot \frac{6}{\sqrt{\log 2}} n^{-1/2} \sqrt{\log(2s(1)n)} \right. \\
& \quad + \frac{10}{\sqrt{\pi}} \cdot \left( \frac{6}{\sqrt{\log 2}} \right)^3 n^{-3/2} (\log(2s(1)n))^{3/2} \\
& \quad \left. + 2 \left( \mathbb{E}\|Z\|^3 \right)^{2/3} \left( \frac{5}{\sqrt{\pi}} \right)^{1/3} \cdot \frac{6}{\sqrt{\log 2}} n^{-1/2} \sqrt{\log(2s(1)n)} \right\} \\
& \stackrel{\text{Doob's } L^3}{\leq} \|g\|_M \left( \frac{30}{\sqrt{\pi} \log 2} + \frac{2 \cdot 5^{1/3} \cdot 6}{\pi^{1/6} \sqrt{\log 2}} \left( \left( \frac{3}{2} \right)^3 \cdot 2 \sqrt{\frac{2}{\pi}} \right)^{2/3} s(1) \right) n^{-1/2} \sqrt{\log(2s(1)n)} \\
& \quad + \|g\|_M \frac{10}{\sqrt{\pi}} \cdot \left( \frac{6}{\sqrt{\log 2}} \right)^3 n^{-3/2} (\log(2s(1)n))^{3/2} \\
& = \|g\|_M \frac{30 + 54 \cdot 5^{1/3} s(1)}{\sqrt{\pi} \log 2} n^{-1/2} \sqrt{\log(2s(1)n)} \\
& \quad + \|g\|_M \frac{2160}{\sqrt{\pi} (\log 2)^{3/2}} n^{-3/2} (\log(2s(1)n))^{3/2}. \tag{5.2}
\end{aligned}$$

### 5.1.2. Applying Stein's method

Let  $g \in M$  and  $g_n = g - \mathbb{E}[g(A_n)]$ . Let  $f_n = \phi_n(g_n)$ , as in (4.5). First, note that:

$$\mathbb{E} Df_n(Y_n)[Y_n] = n^{-1/2} \sum_{j=1}^{\lfloor ns(1) \rfloor} \mathbb{E} Df_n(Y_n)[X_j \mathbb{1}_{[s^{-1}(j/n), 1]}].$$

We now let  $Y_n^j = n^{-1/2} \sum_{k \neq j} X_k \mathbb{1}_{[s^{-1}(k/n), 1]} = Y_n - n^{-1/2} X_j \mathbb{1}_{[s^{-1}(j/n), 1]}$  and observe that, by Taylor's theorem:

$$\begin{aligned}
& \left| n^{-1/2} \mathbb{E} X_j Df_n(Y_n) [\mathbb{1}_{[s^{-1}(j/n), 1]}] - \mathbb{E} \left\{ n^{-1/2} X_j Df_n(Y_n^j) [\mathbb{1}_{[s^{-1}(j/n), 1]}] \right. \right. \\
& \quad \left. \left. + n^{-1} (X_j)^2 D^2 f_n(Y_n^j) \left[ (\mathbb{1}_{[s^{-1}(j/n), 1]} e_i)^{(2)} \right] \right\} \right| \\
& = \left| \mathbb{E} \left[ n^{-1/2} X_j Df_n \left( Y_n^j + n^{-1/2} X_j \mathbb{1}_{[j/n, 1]} \right) [\mathbb{1}_{[s^{-1}(j/n), 1]}] \right. \right. \\
& \quad \left. \left. - n^{-1/2} X_j Df_n(Y_n^j) [\mathbb{1}_{[s^{-1}(j/n), 1]}] \right. \right. \\
& \quad \left. \left. - n^{-1} (X_j)^2 D^2 f_n(Y_n^j) \left[ (\mathbb{1}_{[s^{-1}(j/n), 1]})^{(2)} \right] \right] \right| \\
& \stackrel{(4.6)C)}{\leq} \frac{n^{-3/2}}{6} \|g_n\|_M \mathbb{E}|X_j|^3 \tag{5.3}
\end{aligned}$$

because, clearly,  $\|\mathbb{1}_{[s^{-1}(j/n), 1]}\| = 1$ . Also, in the last inequality we have used the fact that  $X_j$  is independent of  $Y_n^j$ . We can now sum (5.3) over  $j = 1, 2, \dots, \lfloor ns(1) \rfloor$

and use the fact that  $X_j$ 's are independent of  $Y_n^j$ 's and that  $X_j$ 's have mean 0 and variance 1 to obtain:

$$\left| \mathbb{E} Df_n(Y_n)[Y_n] - n^{-1} \sum_{j=1}^{\lfloor ns(1) \rfloor} D^2 f_n(Y_n^j) \left[ \left( \mathbb{1}_{[s^{-1}(j/n), 1]} \right)^{(2)} \right] \right| \leq \frac{n^{-1/2}}{6} s(1) \|g_n\|_M \mathbb{E}|X_1|^3.$$

We notice that for  $\mathcal{A}_n$  defined in Proposition 4.1, using Remark 4.4, we obtain:

$$\begin{aligned} |\mathbb{E} \mathcal{A}_n f_n(Y_n)| &= \left| \mathbb{E} Df_n(Y_n)[Y_n] - \mathbb{E} D^2 f_n(Y_n) \left[ A_n^{(2)} \right] \right| \\ &\leq \left| \mathbb{E} Df_n(Y_n)[Y_n] - n^{-1} \sum_{j=1}^{\lfloor ns(1) \rfloor} \mathbb{E} D^2 f_n(Y_n^j) \left[ \left( \mathbb{1}_{[s^{-1}(j/n), 1]} \right)^{(2)} \right] \right| \\ &\quad + n^{-1} \left| \sum_{j=1}^{\lfloor ns(1) \rfloor} \mathbb{E} \left\{ D^2 f_n(Y_n) \left[ \left( \mathbb{1}_{[s^{-1}(j/n), 1]} \right)^{(2)} \right] - D^2 f_n(Y_n^j) \left[ \left( \mathbb{1}_{[s^{-1}(j/n), 1]} \right)^{(2)} \right] \right\} \right| \\ &\leq \frac{n^{-1/2}}{6} \|g_n\|_M \mathbb{E}|X_1|^3 \\ &\quad + n^{-1} \left| \sum_{j=1}^{\lfloor ns(1) \rfloor} \mathbb{E} \left\{ D^2 f_n \left( Y_n^j + n^{-1/2} X_j \mathbb{1}_{[s^{-1}(j/n), 1]} \right) \left[ \left( \mathbb{1}_{[s^{-1}(j/n), 1]} \right)^{(2)} \right] \right. \right. \\ &\quad \left. \left. - D^2 f_n(Y_n^j) \left[ \left( \mathbb{1}_{[s^{-1}(j/n), 1]} \right)^{(2)} \right] \right\} \right| \\ &\stackrel{(4.6)^C}{\leq} \frac{n^{-1/2}}{6} \|g_n\|_M \mathbb{E}|X_1|^3 + n^{-1} \frac{\|g_n\|_M}{3} \sum_{j=1}^{\lfloor ns(1) \rfloor} n^{-1/2} \mathbb{E} \|X_j \mathbb{1}_{[s^{-1}(j/n), 1]}\| \\ &\leq \frac{n^{-1/2}}{6} s(1) \|g_n\|_M \mathbb{E}|X_1|^3 + \frac{n^{-1/2}}{3} s(1) \|g_n\|_M \mathbb{E} \|X_1 \mathbb{1}_{[s^{-1}(j/n), 1]}\| \\ &\leq \frac{n^{-1/2} s(1) \|g_n\|_M}{2} \mathbb{E}|X_1|^3. \end{aligned}$$

The last inequality follows by Jensen's inequality:

$$\mathbb{E}|X_1| \leq \sqrt{\mathbb{E}|X_1|^2} = 1 = (\mathbb{E}|X_1|^2)^{3/2} \leq \mathbb{E}|X_1|^3. \quad (5.4)$$

Now, note that this gives:

$$\begin{aligned} |\mathbb{E} g(Y_n) - \mathbb{E} g(A_n)| &= |\mathbb{E} g_n(Y_n)| = |\mathbb{E} \mathcal{A}_n f_n(Y_n)| \\ &\leq \frac{n^{-1/2}}{2} s(1) \|g_n\|_M \mathbb{E}|X_1|^3 \\ &\leq \frac{n^{-1/2}}{2} s(1) (\|g\|_M + \mathbb{E} g(A_n)) \mathbb{E}|X_1|^3 \\ &\leq \frac{n^{-1/2}}{2} s(1) (2 + \mathbb{E} \|A_n\|^3) \|g\|_M \mathbb{E}|X_1|^3. \end{aligned} \quad (5.5)$$

Also, recall that, by Doob's  $L^p$  inequality (see, e.g. Theorem 1.7 of Chapter II in [RY91]), if  $M$  is a right-continuous martingale then for every  $t > 0$  and  $p > 1$ :

$$\mathbb{E} \left[ \left( \sup_{s \in [0, t]} |M_s| \right)^p \right] \leq \left( \frac{p}{p-1} \right)^p \mathbb{E} |M_t|^p.$$

Note that  $A_n$  is an integrable process, adapted to its natural filtration and for any  $t \in [s^{-1}(m/n), s^{-1}((m+1)/n))$  and  $r \in [s^{-1}(l/n), s^{-1}((l+1)/n))$  and  $r < t$ :

$$\begin{aligned} \mathbb{E} [A_n(t) | \{A_n(u) : u \leq r\}] &= n^{-1/2} \mathbb{E} \left[ \sum_{k=1}^m Z_k \middle| Z_1, \dots, Z_l \right] \\ &= n^{-1/2} \sum_{k=1}^l Z_k + n^{-1/2} \mathbb{E} \left[ \sum_{k=l+1}^m Z_k \right] \\ &= A_n(r) \end{aligned}$$

and so  $A_n$  is a right-continuous martingale. Applying Doob's  $L^3$  inequality to it yields:

$$\mathbb{E} \|A_n\|^3 \leq \left( \frac{3}{2} \right)^3 \mathbb{E} |A_n(1)|^3 = \left( \frac{3}{2} \right)^3 \cdot 2 \sqrt{\frac{2}{\pi}} s(1)^{3/2},$$

because  $A_n(1) \sim \mathcal{N}(0, \frac{\lfloor ns(1) \rfloor}{n})$ . Therefore (5.5) gives:

$$|\mathbb{E}g(Y_n) - \mathbb{E}g(A_n)| \leq n^{-1/2} \left( 1 + \left( \frac{3}{2} \right)^3 \cdot \sqrt{\frac{2}{\pi}} s(1)^{3/2} \right) \|g\|_{Ms(1)} \mathbb{E}|X_1|^3. \quad (5.6)$$

Combining this with (5.2):

$$\begin{aligned} |\mathbb{E}g(Y_n) - \mathbb{E}g(Z)| &\leq \|g\|_M \frac{30 + 54 \cdot 5^{1/3} s(1)}{\sqrt{\pi} \log 2} n^{-1/2} \sqrt{\log(2s(1)n)} \\ &\quad + \|g\|_{Ms(1)} \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} s(1)^{3/2} \right) \mathbb{E}|X_1|^3 n^{-1/2} \\ &\quad + \|g\|_M \frac{2160}{\sqrt{\pi} (\log 2)^{3/2}} n^{-3/2} (\log(2s(1)n))^{3/2}, \end{aligned} \quad (5.7)$$

which proves Theorem 3.1.  $\square$

## 5.2. Proof of Theorem 3.2

Note that  $(P(\lfloor nS^{(n)}(t) \rfloor), t \in [0, 1])$  can be expressed in the following way:

$$P(\lfloor nS^{(n)}(t) \rfloor) - \lfloor nS^{(n)}(t) \rfloor = \sum_{i=1}^{\lfloor nS^{(n)}(t) \rfloor} X_i,$$

where  $(X_i + 1)$ 's are i.i.d.  $\text{Poisson}(1)$ . Therefore, we can express  $(\tilde{Y}_n(t), t \in [0, 1])$  in the following way:

$$\tilde{Y}_n(t) = n^{-1/2} \left\{ \sum_{i=1}^{\lfloor nS^{(n)}(t) \rfloor} X_i + P\left(nS^{(n)}(t)\right) - P\left(\lfloor nS^{(n)}(t) \rfloor\right) - \left(nS^{(n)}(t) - \lfloor nS^{(n)}(t) \rfloor\right) \right\}.$$

We also define:

$$Y_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nS^{(n)}(t) \rfloor} X_i.$$

Note that  $|nS^{(n)}(t) - \lfloor nS^{(n)}(t) \rfloor| \leq 1$  for all  $t \geq 0$ . Also, observe that for all  $t \geq 0$ :

$$\left| P\left(nS^{(n)}(t)\right) - P\left(\lfloor nS^{(n)}(t) \rfloor\right) \right| \leq P\left(\lfloor nS^{(n)}(t) \rfloor + 1\right) - P\left(\lfloor nS^{(n)}(t) \rfloor\right).$$

By the independence of increments of a Poisson process:

$$\begin{aligned} \text{A) } \mathbb{E}\|\tilde{Y}_n - Y_n\| &\leq n^{-1/2} \left[ 1 + \mathbb{E} \left[ \max_{1 \leq i \leq n} \bar{P}_i \right] \right] \\ \text{B) } \mathbb{E}\|\tilde{Y}_n - Y_n\|^3 &\leq n^{-3/2} \left[ 4 + 4\mathbb{E} \left[ \max_{1 \leq i \leq n} \bar{P}_i^3 \right] \right], \end{aligned} \quad (5.8)$$

where  $\bar{P}_1, \dots, \bar{P}_n \stackrel{\text{i.i.d.}}{\sim} \text{Poisson}(1)$ . Using the trick from [Das11], we note that, by Jensen's inequality applied to function  $\exp(x \log \log(n+2))$ :

$$\begin{aligned} \exp \left( \log \log(n+2) \cdot \mathbb{E} \left[ \max_{1 \leq i \leq n} \bar{P}_i \right] \right) &\leq \mathbb{E} \left[ \exp \left( \log \log(n+2) \cdot \max_{1 \leq i \leq n} \bar{P}_i \right) \right] \\ &= \mathbb{E} \left[ \max_{1 \leq i \leq n} \exp \left( \log \log(n+2) \cdot \bar{P}_i \right) \right] \\ &\leq n \mathbb{E} \left[ \exp \left( \log \log n \cdot \bar{P}_1 \right) \right] \\ &= n \exp \left( \log(n+2) - 1 \right) \\ &\leq (1 + 2e^{-1}) n^2 \end{aligned} \quad (5.9)$$

and by Jensen's inequality applied to function  $\exp(x^{1/3} \log \log(n+3))$ , which is convex for  $x \geq \frac{8}{(\log \log(n+3))^3}$ :

$$\begin{aligned} &\exp \left( \log \log(n+3) \cdot \left\{ \mathbb{E} \left[ \max_{1 \leq i \leq n} \bar{P}_i^3 \right] \left( \max_{1 \leq i \leq n} \bar{P}_i^3 \right) \geq \frac{8}{(\log \log(n+3))^3} \right\}^{1/3} \right) \\ &\leq \mathbb{E} \left[ \exp \left( \log \log(n+3) \cdot \max_{1 \leq i \leq n} \bar{P}_i \right) \left( \max_{1 \leq i \leq n} \bar{P}_i^3 \right) \geq \frac{8}{(\log \log(n+3))^3} \right] \\ &\leq n \mathbb{E} \left[ \exp \left( \log \log(n+3) \cdot \bar{P}_1 \right) \left( \max_{1 \leq i \leq n} \bar{P}_i^3 \right) \geq \frac{8}{(\log \log(n+3))^3} \right] \end{aligned}$$

$$\begin{aligned}
&\leq \frac{n \exp(\log(n+3) - 1)}{\mathbb{P} \left[ (\max_{1 \leq i \leq n} \bar{P}_i) \geq \frac{2}{(\log \log(n+3))} \right]} \\
&\leq \frac{n \exp(\log(n+3) - 1)}{\mathbb{P} \left[ \bar{P}_1 \geq \frac{2}{(\log \log 4)} \right]} \\
&\leq \frac{(1 + 3e^{-1})n^2}{1 - \frac{1957}{720e}}. \tag{5.10}
\end{aligned}$$

Now, combining (5.8), (5.9) and (5.10), we obtain:

$$\begin{aligned}
\text{A) } \mathbb{E} \|\tilde{Y}_n - Y_n\| &\leq n^{-1/2} \left[ 1 + \frac{\log(1 + 2e^{-1}) + 2 \log n}{\log \log(n+2)} \right] \\
\text{B) } \mathbb{E} \|\tilde{Y}_n - Y_n\|^3 &\leq n^{-3/2} \left[ 4 + 4 \left( \frac{\log \left( \frac{1+3e^{-1}}{1-\frac{1957}{720e}} \right) + 2 \log n}{\log \log(n+3)} \right)^3 + \frac{32}{(\log \log(n+3))^3} \right] \\
&\leq n^{-3/2} \left[ 4 + \frac{16701 + 128(\log n)^3}{(\log \log(n+3))^3} \right]. \tag{5.11}
\end{aligned}$$

We also note that:

$$\begin{aligned}
\left[ \mathbb{E} \|\tilde{Y}_n\|^3 \right]^{2/3} &\stackrel{\text{Doob}}{\leq} n^{-1} \left( \frac{3}{2} \right)^2 \left[ \mathbb{E} \left| P(nS^{(n)}(1)) - nS^{(n)}(1) \right|^3 \right]^{2/3} \\
&\leq \left( \frac{3}{2} \right)^2 n^{-1} \left[ \mathbb{E} \left| P(nS^{(n)}(1)) - nS^{(n)}(1) \right|^4 \right]^{1/2} \\
&= \left( \frac{3}{2} \right)^2 n^{-1/2} \sqrt{S^{(n)}(1)} \left( 1 + 3nS^{(n)}(1) \right)^{1/2}. \tag{5.12}
\end{aligned}$$

Then, for every  $g \in M$ :

$$\begin{aligned}
|\mathbb{E}g(Y_n) - \mathbb{E}g(\tilde{Y}_n)| &\stackrel{\text{MVT}}{\leq} \mathbb{E} \left[ \sup_{c \in [0,1]} \|Dg((1-c)\tilde{Y}_n + cY_n)\| \|Y_n - \tilde{Y}_n\| \right] \\
&\leq \|g\|_M \mathbb{E} \left[ \sup_{c \in [0,1]} (1 + \|\tilde{Y}_n + c(Y_n - \tilde{Y}_n)\|^2) \|Y_n - \tilde{Y}_n\| \right] \\
&\leq \|g\|_M \mathbb{E} \left[ (1 + 2\|\tilde{Y}_n\|^2 + 2\|Y_n - \tilde{Y}_n\|^2) \|Y_n - \tilde{Y}_n\| \right] \\
&\stackrel{\text{Hölder}}{\leq} \|g\|_M \left\{ \mathbb{E}\|Y_n - \tilde{Y}_n\| + 2\mathbb{E}\|Y_n - \tilde{Y}_n\|^3 + 2(\mathbb{E}\|\tilde{Y}_n\|^3)^{2/3}(\mathbb{E}\|Y_n - \tilde{Y}_n\|^3)^{1/3} \right\} \\
&\stackrel{(5.11), (5.12)}{\leq} \|g\|_M \left\{ n^{-1/2} \left[ 1 + \frac{\log(1 + 2e^{-1}) + 2 \log n}{\log \log(n+2)} \right] + n^{-3/2} \left[ 8 + \frac{33402 + 256(\log n)^3}{(\log \log(n+3))^3} \right] \right. \\
&\quad \left. + \frac{9}{2} n^{-1} \sqrt{S^{(n)}(1)} \left( 1 + 3nS^{(n)}(1) \right)^{1/2} \left[ 4 + \frac{16701 + 128(\log n)^3}{(\log \log(n+3))^3} \right]^{1/3} \right\}. \tag{5.13}
\end{aligned}$$

Let  $A_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nS^{(n)}(t) \rfloor} Z_i, t \in [0, 1]$  for  $Z_i \stackrel{i.i.d}{\sim} \mathcal{N}(0, 1)$ . By (5.6):

$$\begin{aligned} |\mathbb{E}g(Y_n) - \mathbb{E}g(A_n)| &\leq n^{-1/2} \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} S^{(n)}(1)^{3/2} \right) S^{(n)}(1) \|g\|_M \mathbb{E}|X_1|^3 \\ &\leq n^{-1/2} \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} S^{(n)}(1)^{3/2} \right) S^{(n)}(1) \|g\|_M (1 + 2e^{-1}) \end{aligned} \quad (5.14)$$

because  $X_1 \stackrel{\mathcal{D}}{=} P(1) - 1$ . Now let  $\tilde{A}_n(t) = n^{-1/2} \sum_{i=1}^{\lfloor nS(t) \rfloor} Z_i, t \in [0, 1]$ . Then:

$$\begin{aligned} \text{A) } \mathbb{E}\|A_n - \tilde{A}_n\| &= n^{-1/2} \mathbb{E} \left[ \sup_{t \in [0, 1]} \left| \sum_{i=\lfloor nS(t) \wedge S^{(n)}(t) \rfloor + 1}^{\lfloor nS(t) \vee S^{(n)}(t) \rfloor} Z_i \right| \right] \\ &= n^{-1/2} \mathbb{E} \left[ \sup_{t \in [0, 1]} \left| \sum_{i=1}^{\lfloor nS(t) \vee S^{(n)}(t) \rfloor - (\lfloor nS(t) \wedge S^{(n)}(t) \rfloor + 1)} Z_i \right| \right] \\ &\stackrel{Doob, Jensen}{\leq} 2n^{-1/2} \sqrt{\mathbb{E} \left| \sum_{i=1}^{\sup_{t \in [0, 1]} (\lfloor nS(t) \vee S^{(n)}(t) \rfloor - (\lfloor nS(t) \wedge S^{(n)}(t) \rfloor + 1))} Z_i \right|^2} \leq 2\sqrt{\|S - S^{(n)}\|} \\ \text{B) } \mathbb{E}\|A_n - \tilde{A}_n\|^3 &= n^{-3/2} \mathbb{E} \left[ \sup_{t \in [0, 1]} \left| \sum_{i=\lfloor nS(t) \wedge S^{(n)}(t) \rfloor + 1}^{\lfloor nS(t) \vee S^{(n)}(t) \rfloor} Z_i \right|^3 \right] \\ &= n^{-3/2} \mathbb{E} \left[ \sup_{t \in [0, 1]} \left| \sum_{i=1}^{\lfloor nS(t) \vee S^{(n)}(t) \rfloor - (\lfloor nS(t) \wedge S^{(n)}(t) \rfloor + 1)} Z_i \right|^3 \right] \\ &\stackrel{Doob}{\leq} \left( \frac{3}{2} \right)^3 n^{-3/2} \mathbb{E} \left[ \left| \sum_{i=1}^{\sup_{t \in [0, 1]} (\lfloor nS(t) \vee S^{(n)}(t) \rfloor - (\lfloor nS(t) \wedge S^{(n)}(t) \rfloor + 1))} Z_i \right|^3 \right] \\ &\leq 2\sqrt{\frac{2}{\pi}} \left( \frac{3}{2} \right)^3 \|S - S^{(n)}\|^{3/2} \\ \text{C) } (\mathbb{E}\|\tilde{A}_n\|^3)^{2/3} &\stackrel{Doob}{\leq} 2n^{-1} \left( \frac{3}{2} \right)^2 \pi^{-1/3} (S(1)^{3/2})^{2/3} = \frac{9}{2\pi^{1/3}} n^{-1} S(1). \end{aligned} \quad (5.15)$$

Therefore:

$$\begin{aligned} |\mathbb{E}g(A_n) - \mathbb{E}g(\tilde{A}_n)| &\stackrel{\text{MVT}}{\leq} \mathbb{E} \left[ \sup_{c \in [0, 1]} \|Dg((1-c)\tilde{A}_n + cA_n)\| \|A_n - \tilde{A}_n\| \right] \\ &\leq \|g\|_M \mathbb{E} \left[ \sup_{c \in [0, 1]} (1 + \|\tilde{A}_n + c(A_n - \tilde{A}_n)\|^2) \|A_n - \tilde{A}_n\| \right] \end{aligned}$$

$$\begin{aligned}
&\leq \|g\|_M \mathbb{E} \left[ (1 + 2\|\tilde{A}_n\|^2 + 2\|A_n - \tilde{A}_n\|^2) \|A_n - \tilde{A}_n\| \right] \\
&\stackrel{\text{Hölder}}{\leq} \|g\|_M \left\{ \mathbb{E}\|A_n - \tilde{A}_n\| + 2\mathbb{E}\|A_n - \tilde{A}_n\|^3 + 2(\mathbb{E}\|\tilde{A}_n\|^3)^{2/3} \left( \mathbb{E}\|A_n - \tilde{A}_n\| \right)^{1/3} \right\} \\
&\stackrel{(5.15)}{\leq} \|g\|_M \left\{ 2\sqrt{\|S - S^{(n)}\|} + \frac{27\sqrt{2}}{2\sqrt{\pi}} \|S - S^{(n)}\|^{3/2} + \frac{27\sqrt{2}}{2\sqrt{\pi}} S(1) \sqrt{\|S - S^{(n)}\|} \right\}.
\end{aligned} \tag{5.16}$$

By (5.2) we get for  $Z = B \circ S$ :

$$\begin{aligned}
|\mathbb{E}g(\tilde{A}_n) - \mathbb{E}g(Z)| &\leq \|g\|_M \frac{30 + 54 \cdot 5^{1/3} S(1)}{\sqrt{\pi} \log 2} n^{-1/2} \sqrt{\log(2S(1)n)} \\
&\quad + \|g\|_M \frac{2160}{\sqrt{\pi} (\log 2)^{3/2}} n^{-3/2} (\log(2S(1)n))^{3/2}.
\end{aligned} \tag{5.17}$$

Theorem 3.2 now follows from (5.13), (5.14), (5.16), (5.17).  $\square$

### 5.3. Proof of Theorem 3.4

Note that  $X_n$  jumps up by  $\frac{1}{n}$  with intensity  $\frac{1}{2}n^2X_n(t)(1-X_n(t)) + n\nu_2(1-X_n(t))$  and down by  $\frac{1}{n}$  with intensity  $\frac{1}{2}n^2X_n(t)(1-X_n(t)) + n\nu_1X_n(t)$ . To see this observe that a jump occurs with intensity  $\binom{n}{2}$  and it is an up-jump if the first gene chosen was of type a, the second of type A and the one with type A died (which happens with probability  $\frac{1}{2}X_n(t)\frac{n(1-X_n(t))}{n-1}$ ) or if the first one chosen was of type A, the second of type a and the type A gene died (which happens with probability  $\frac{1}{2}(1-X_n(t))\frac{nX_n(t)}{n-1}$ ). In addition, there are  $n(1-X_n(t))$  genes of type A and each of them mutates into type a at rate  $\nu_2$ . Hence:

$$\begin{cases} \mathbb{P} [X_n(t + \Delta t) - X_n(t) = \frac{1}{n}] = \frac{1}{2}n^2X_n(t)(1-X_n(t))\Delta t + n\nu_2(1-X_n(t))\Delta t \\ \mathbb{P} [X_n(t + \Delta t) - X_n(t) = -\frac{1}{n}] = \frac{1}{2}n^2X_n(t)(1-X_n(t))\Delta t + n\nu_1X_n(t)\Delta t. \end{cases}$$

Therefore:

$$\begin{aligned}
M_n(t) &= \frac{P_1 \left( n^2 R_1^{(n)}(t) \right) - n^2 R_1^{(n)}(t)}{n} - \frac{P_{-1} \left( n^2 R_{-1}^{(n)}(t) \right) - n^2 R_{-1}^{(n)}(t)}{n} \\
&\quad + \int_0^t (\nu_2 - (\nu_1 + \nu_2) M_n(s)) ds,
\end{aligned}$$

where  $P_1, P_{-1}$  are i.i.d. Poisson processes with rate 1, independent of  $X_n$ , and

$$\begin{cases} R_1^{(n)}(t) := \int_0^t \left( \frac{1}{2}X_n(s) + \frac{\nu_2}{n} \right) (1 - X_n(s)) ds \\ R_{-1}^{(n)}(t) := \int_0^t \left( \frac{1}{2}(1 - X_n(s)) + \frac{\nu_1}{n} \right) X_n(s) ds \end{cases} \quad \text{for } t \in [0, 1].$$

Also let:

$$\begin{cases} R_1(t) = R_{-1}(t) := \int_0^t \frac{1}{2}X(s)(1-X(s))ds \\ I_n(t) := \int_0^t (\nu_2 - (\nu_1 + \nu_2)X_n(s)) ds \\ I(t) := \int_0^t (\nu_2 - (\nu_1 + \nu_2)X(s)) ds \end{cases} \quad \text{for } t \in [0, 1].$$



Let us denote  $Z_1 = B_1 \circ R_1$ ,  $Z_{-1} = B_{-1} \circ R_{-1}$ , where  $B_1$  and  $B_{-1}$  are i.i.d. standard Brownian Motions, independent of  $X$  and:

$$Y_n^1(\cdot) := \frac{P_1\left(n^2 R_1^{(n)}(\cdot)\right) - n^2 R_1^{(n)}(\cdot)}{n}, \quad Y_n^{-1}(\cdot) := \frac{P_{-1}\left(n^2 R_{-1}^{(n)}(\cdot)\right) - n^2 R_{-1}^{(n)}(\cdot)}{n}.$$

Now, for any  $g \in M$ :

$$\begin{aligned} & |\mathbb{E}g(M_n) - \mathbb{E}g(M)| \\ & \leq \left| \mathbb{E} \left[ \mathbb{E} \left[ g(Y_n^1 - Y_n^{-1} + I_n) - g(Z_1 - Y_n^{-1} + I_n) | Y_n^{-1}, I_n, R_1^{(n)}, R_{-1}^{(n)}, R_1 \right] \right. \right. \\ & \quad + \mathbb{E} \left[ \mathbb{E} \left[ g(Z_1 - Y_n^{-1} + I_n) - g(Z_1 - Z_{-1} + I_n) | Z_1, I_n, R_1^{(n)}, R_{-1}^{(n)}, R_1 \right] \right. \\ & \quad \left. \left. + \mathbb{E} \left[ \mathbb{E} \left[ g(Z_1 - Z_{-1} + I_n) - g(Z_1 - Z_{-1} + I) | Z_1, Z_{-1}, R_1 \right] \right] \right] \right| \\ & = \left| \mathbb{E} \left[ \mathbb{E} \left[ g^{(1)}(Y_n^1) - g^{(1)}(Z_1) | Y_n^{-1}, I_n, R_1^{(n)}, R_1 \right] \right. \right. \\ & \quad + \mathbb{E} \left[ \mathbb{E} \left[ g^{(-1)}(Y_n^{-1}) - g^{(-1)}(Z_{-1}) | Z_1, I_n, R_{-1}^{(n)}, R_{-1} \right] \right. \\ & \quad \left. \left. + \mathbb{E} \left[ \mathbb{E} \left[ g^{(0)}(I_n) - g^{(0)}(I) | Z_1, Z_{-1}, R_1, R_{-1} \right] \right] \right] \right|, \end{aligned} \quad (5.18)$$

where  $g^{(1)}(x) = g(x - Y_n^{-1} + I_n)$ ,  $g^{(-1)}(x) = g(Z_1 - x + I_n)$ ,  $g^{(0)}(x) = g(Z_1 - Z_{-1} + x)$ . Note that, given  $R_1$  and  $R_1^{(n)}$ ,  $Y_n^1$  and  $Z_1$  are independent of  $Y_n^{-1}$  and  $I_n$ . Similarly, given  $R_{-1}$  and  $R_{-1}^{(n)}$ ,  $Y_n^{-1}$  and  $Z_{-1}$  are independent of  $Z_1$  and  $I_n$ . Also, given  $R_1$  and  $R_{-1}$ ,  $I_n$  and  $I$  are independent of  $Z_1$  and  $Z_{-1}$ .

We now apply Theorem 3.2 to obtain:

$$\begin{aligned} & \text{A) } \mathbb{E} \left[ \mathbb{E} \left[ g^{(1)}(Y_n^1) - g^{(1)}(Z_1) | Y_n^{-1}, I_n, R_1^{(n)}, R_1 \right] \right] \\ & \leq \mathbb{E} \left[ \mathbb{E} \left[ \|g^{(1)}\|_M | R_1, R_1^{(n)} \right] \right. \\ & \quad \cdot \left\{ \left( 2 + \frac{27\sqrt{2}}{2\sqrt{\pi}} R_1(1) \right) \sqrt{\|R_1 - R_1^{(n)}\|} + \frac{27\sqrt{2}}{2\sqrt{\pi}} \|R_1 - R_1^{(n)}\|^{3/2} \right. \\ & \quad + n^{-1} \left[ \frac{30 + 54 \cdot 5^{1/3} R_1(1)}{\sqrt{\pi \log 2}} \sqrt{\log(2R_1(1)n^2)} + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} R_1^{(n)}(1)^{3/2} \right) \right. \\ & \quad \left. \left. \cdot R_1^{(n)}(1)(1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 4 \log n}{\log \log(n^2 + 2)} \right] \right. \\ & \quad + n^{-2} \frac{9\sqrt{R_1^{(n)}(1)}}{2} \left( 1 + 3n^2 R_1^{(n)}(1) \right)^{1/2} \left[ 4 + \frac{16701 + 1024(\log n)^3}{(\log \log(n^2 + 3))^3} \right]^{1/3} \\ & \quad \left. \left. + n^{-3} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(2R_1(1)n^2))^{3/2} + 8 + \frac{33402 + 2048(\log n)^3}{(\log \log(n^2 + 3))^3} \right] \right\} \right] \\ & \leq \mathbb{E} \left[ \mathbb{E} \left[ \|g^{(1)}\|_M | R_1, R_1^{(n)} \right] \left\{ \left( 2 + \frac{27\sqrt{2}}{16\sqrt{\pi}} \right) \sqrt{\|R_1 - R_1^{(n)}\|} + \frac{27\sqrt{2}}{2\sqrt{\pi}} \|R_1 - R_1^{(n)}\|^{3/2} \right. \right. \right. \end{aligned}$$

$$\begin{aligned}
& + n^{-1} \left[ \frac{30 + \frac{27}{4} \cdot 5^{1/3}}{\sqrt{\pi \log 2}} \sqrt{\log(n^2/4)} + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} \left( \frac{1}{8} + \frac{\nu_2}{n} \right)^{3/2} \right) \right. \\
& \cdot \left. \left( \frac{1}{8} + \frac{\nu_2}{n} \right) (1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 4 \log n}{\log \log(n^2 + 2)} \right] \\
& + n^{-2} \frac{9 \sqrt{\frac{1}{8} + \frac{\nu_2}{n}}}{2} \left( 1 + \frac{3n^2}{8} + 3n\nu_2 \right)^{1/2} \left[ 4 + \frac{16701 + 1024(\log n)^3}{(\log \log(n^2 + 3))^3} \right]^{1/3} \\
& + n^{-3} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(n^2/4 + \nu_2 n))^{3/2} + 8 + \frac{33402 + 2048(\log n)^3}{(\log \log(n^2 + 3))^3} \right] \Big] \\
\text{B)} \quad & \left| \mathbb{E} \left( \mathbb{E} \left\{ g^{(-1)}(Y_n^{-1}) - g^{(-1)}(Z_{-1}) \mid Z_1, I_n \right\} \mid R_{-1}, R_{-1}^{(n)} \right) \right| \\
\leq & \mathbb{E} \left[ \mathbb{E} \left[ \|g^{(-1)}\|_M \mid R_{-1}, R_{-1}^{(n)} \right] \right. \\
& \cdot \left\{ \left( 2 + \frac{27\sqrt{2}}{2\sqrt{\pi}} R_{-1}(1) \right) \sqrt{\|R_{-1} - R_{-1}^{(n)}\|} + \frac{27\sqrt{2}}{2\sqrt{\pi}} \|R_{-1} - R_{-1}^{(n)}\|^{3/2} \right. \\
& + n^{-1} \left[ \frac{30 + 54 \cdot 5^{1/3} R_{-1}(1)}{\sqrt{\pi \log 2}} \sqrt{\log(2R_{-1}(1)n^2)} + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} R_{-1}^{(n)}(1)^{3/2} \right) \right. \\
& \cdot \left. R_{-1}^{(n)}(1) (1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 4 \log n}{\log \log(n^2 + 2)} \right] \\
& + n^{-2} \frac{9 \sqrt{R_{-1}^{(n)}(1)}}{2} \left( 1 + 3n^2 R_{-1}^{(n)}(1) \right)^{1/2} \left[ 4 + \frac{16701 + 1024(\log n)^3}{(\log \log(n^2 + 3))^3} \right]^{1/3} \\
& + n^{-3} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(2R_{-1}(1)n^2))^{3/2} + 8 + \frac{33402 + 2048(\log n)^3}{(\log \log(n^2 + 3))^3} \right] \Big\} \\
& \leq \mathbb{E} \left[ \mathbb{E} \left[ \|g^{(1)}\|_M \mid R_{-1}, R_{-1}^{(n)} \right] \left\{ \left( 2 + \frac{27\sqrt{2}}{16\sqrt{\pi}} \right) \sqrt{\|R_{-1} - R_{-1}^{(n)}\|} + \frac{27\sqrt{2}}{2\sqrt{\pi}} \|R_{-1} - R_{-1}^{(n)}\|^{3/2} \right. \right. \\
& + n^{-1} \left[ \frac{30 + \frac{27}{4} \cdot 5^{1/3}}{\sqrt{\pi \log 2}} \sqrt{\log(n^2/4)} + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} \left( \frac{1}{8} + \frac{\nu_1}{n} \right)^{3/2} \right) \right. \\
& \cdot \left. \left( \frac{1}{8} + \frac{\nu_1}{n} \right) (1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 4 \log n}{\log \log(n^2 + 2)} \right] \\
& + n^{-2} \frac{9 \sqrt{\frac{1}{8} + \frac{\nu_1}{n}}}{2} \left( 1 + \frac{3n^2}{8} + 3n\nu_1 \right)^{1/2} \left[ 4 + \frac{16701 + 1024(\log n)^3}{(\log \log(n^2 + 3))^3} \right]^{1/3} \\
& + n^{-3} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(n^2/4 + \nu_1 n))^{3/2} + 8 + \frac{33402 + 2048(\log n)^3}{(\log \log(n^2 + 3))^3} \right] \Big\} \Big], \\
& \tag{5.19}
\end{aligned}$$

where we have used the fact that  $R_1(1), R_{-1}(1) \leq \frac{1}{8}$  and  $R_1^{(n)} \leq \frac{1}{8} + \frac{\nu_2}{n}$ ,  $R_{-1}^{(n)} \leq$

$\frac{1}{8} + \frac{\nu_1}{n}$ . We also note that:

$$\begin{aligned} & \left| \mathbb{E} \left[ \mathbb{E} \left[ g^{(0)}(I_n) - g^{(0)}(I) | Z_1, Z_{-1}, R_1, R_{-1} \right] \right] \right| \\ & \leq \mathbb{E} \left[ \mathbb{E} \left[ \|g^{(0)}\|_M | R_1, R_{-1} \right] \left[ \|I_n - I\| + 2\|I_n - I\|^3 + 2\|I\| \|I_n - I\| \right] \right] \\ & \leq \mathbb{E} \left[ \mathbb{E} \left[ \|g^{(0)}\|_M | R_1, R_{-1} \right] \left[ \|I_n - I\| + 2\|I_n - I\|^3 + 2\nu_2 \|I_n - I\| \right] \right]. \quad (5.20) \end{aligned}$$

Now note that:

$$\begin{aligned} \text{A)} \quad & \mathbb{E} \left[ \sup_{w, h \in D} \frac{\|D^2 g(w + h - Y_n^{-1} + I_n) - D^2 g(w - Y_n^{-1} + I_n)\|}{\|h\|} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\ & \leq \sup_{x, y \in D} \frac{\|D^2 g(x + y) - D^2 g(x)\|}{\|y\|} \\ \text{B)} \quad & \mathbb{E} \left[ \sup_{w \in D} \frac{\|D^2 g(w - Y_n^{-1} + I_n)\|}{1 + \|w\|} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\ & \leq \mathbb{E} \left[ \sup_{w \in D} \frac{\|D^2 g(w - Y_n^{-1} + I_n)\|}{1 + \|w - Y_n^{-1} + I_n\|} \cdot \frac{1 + \|w - Y_n^{-1} + I_n\|}{1 + \|w\|} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\ & \stackrel{Doob}{\leq} \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \left[ 1 + 2\sqrt{\mathbb{E} \left[ (Y_n^{-1}(1))^2 \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right]} + \nu_2 \right] \\ & = \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \left[ 1 + 2\sqrt{R_{-1}^{(n)}(1)} + \nu_2 \right] \\ & \leq \left( 1 + 2\sqrt{\frac{1}{8} + \frac{\nu_1}{n}} + \nu_2 \right) \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \\ \text{C)} \quad & \mathbb{E} \left[ \sup_{w \in D} \frac{\|Dg(w - Y_n^{-1} + I_n)\|}{1 + \|w\|^2} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\ & \leq \mathbb{E} \left[ \sup_{w \in D} \frac{\|Dg(w - Y_n^{-1} + I_n)\|}{1 + \|w - Y_n^{-1} + I_n\|^2} \cdot \frac{1 + \|w - Y_n^{-1} + I_n\|^2}{1 + \|w\|^2} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\ & \stackrel{Doob}{\leq} 3 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \left[ 1 + 4\mathbb{E} \left[ (Y_n^{-1}(1))^2 \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] + \nu_2^2 \right] \\ & = 3 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \left[ 1 + 4R_{-1}^{(n)}(1) + \nu_2^2 \right] \\ & \leq \left( \frac{9}{2} + \frac{12\nu_1}{n} + 3\nu_2^2 \right) \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \\ \text{D)} \quad & \mathbb{E} \left[ \sup_{w \in D} \frac{|g(w - Y_n^{-1})|}{1 + \|w\|^3} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\ & \leq \mathbb{E} \left[ \sup_{w \in D} \frac{|g(w - Y_n^{-1} + I_n)|}{1 + \|w - Y_n^{-1} + I_n\|^3} \cdot \frac{1 + \|w - Y_n^{-1} + I_n\|^3}{1 + \|w\|^3} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \end{aligned}$$

$$\begin{aligned}
&\stackrel{Doob}{\leq} 9 \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right) \left[ 1 + \frac{27}{8} \mathbb{E} \left[ |Y_n^{-1}(1)|^3 \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] + \nu_2^3 \right] \\
&\leq 9 \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right) \left[ 1 + \frac{27}{8} \left( 3 \left( R_{-1}^{(n)}(1) \right)^2 + n^{-1} R_{-1}^{(n)}(1) \right)^{3/4} + \nu_2^3 \right] \\
&\leq \left( 9 + \frac{243}{8} \cdot \left( 3 \left( \frac{1}{8} + \frac{\nu_1}{n} \right)^2 + \frac{1}{8n} + \frac{\nu_1}{n^2} \right)^{3/4} + 9\nu_2^3 \right) \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right).
\end{aligned}$$

Therefore:

$$\begin{aligned}
&\mathbb{E} \left[ \|g^{(1)}\|_M \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\leq \left[ 9 + 2\sqrt{\frac{1}{8} + \frac{\nu_1}{n}} + \frac{12\nu_1}{n} + \frac{243}{8} \cdot \left( \frac{3}{64} + \frac{6\nu_1 + 1}{8n} + \frac{\nu_1 + \nu_1^2}{n^2} \right)^{3/4} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right] \|g\|_M
\end{aligned} \tag{5.21}$$

Similarly:

$$\begin{aligned}
&\text{A) } \mathbb{E} \left[ \sup_{w, h \in D} \frac{\|D^2 g(Z_1 - (w + h) + I_n) - D^2 g(Z_1 - w + I_n)\|}{\|h\|} \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\leq \sup_{x, y \in D} \frac{\|D^2 g(x + y) - D^2 g(x)\|}{\|y\|} \\
&\text{B) } \mathbb{E} \left[ \sup_{w \in D} \frac{\|D^2 g(Z_1 - w + I_n)\|}{1 + \|w\|} \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\leq \mathbb{E} \left[ \sup_{w \in D} \frac{\|D^2 g(Z_1 - w + I_n)\|}{1 + \|Z_1 - w + I_n\|} \cdot \frac{1 + \|Z_1 - w + I_n\|}{1 + \|w\|} \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\stackrel{Doob}{\leq} \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \left[ 1 + 2\sqrt{\mathbb{E} \left[ (Z_1(1))^2 \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right]} + \nu_2 \right] \\
&= \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \left[ 1 + 2\sqrt{R_1(1)} + \nu_2 \right] \\
&\leq \left( 1 + \frac{\sqrt{2}}{2} + \nu_2 \right) \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \\
&\text{C) } \mathbb{E} \left[ \sup_{w \in D} \frac{\|Dg(Z_1 - w + I_n)\|}{1 + \|w\|^2} \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\leq \mathbb{E} \left[ \sup_{w \in D} \frac{\|Dg(Z_1 - w + I_n)\|}{1 + \|Z_1 - w + I_n\|^2} \cdot \frac{1 + \|Z_1 - w + I_n\|^2}{1 + \|w\|^2} \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\stackrel{Doob}{\leq} 3 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \left[ 1 + 4\mathbb{E} \left[ (Z_1(1))^2 \mid R_1, R_1^{(n)}, R_{-1}^{(n)} \right] + \nu_2^2 \right] \\
&= 3 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \left[ 1 + 4R_1(1) + \nu_2^2 \right] \leq \left( \frac{9}{2} + 3\nu_2^2 \right) \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right)
\end{aligned}$$

$$\begin{aligned}
\text{D)} \quad & \mathbb{E} \left[ \sup_{w \in D} \frac{|g(Z_1 - w + I_n)|}{1 + \|w\|^3} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \leq \mathbb{E} \left[ \sup_{w \in D} \frac{|g(Z_1 - w + I_n)|}{1 + \|Z_1 - w + I_n\|^3} \cdot \frac{1 + \|Z_1 - w + I_n\|^3}{1 + \|w\|^3} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \stackrel{Doob}{\leq} 9 \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right) \left[ 1 + \frac{27}{8} \mathbb{E} \left[ |Z_1(1)|^3 \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] + \nu_2^3 \right] \\
& \leq 9 \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right) \left[ 1 + \frac{27\sqrt{2}}{4\sqrt{\pi}} R_1(1)^{3/2} + \nu_2^3 \right] \\
& \leq \left( 9 + \frac{243}{64\sqrt{\pi}} + 9\nu_2^3 \right) \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right).
\end{aligned}$$

Therefore:

$$\mathbb{E} \left[ \|g^{(-1)}\|_M | R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \leq \left( 9 + \frac{243}{64\sqrt{\pi}} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right) \|g\|_M. \quad (5.22)$$

Also:

$$\begin{aligned}
\text{A)} \quad & \mathbb{E} \left[ \sup_{w, h \in D} \frac{\|D^2 g(Z_1 - Z_{-1} + (w + h)) - D^2 g(Z_1 - Z_{-1} + w)\|}{\|h\|} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \leq \sup_{x, y \in D} \frac{\|D^2 g(x + y) - D^2 g(x)\|}{\|y\|} \\
\text{B)} \quad & \mathbb{E} \left[ \sup_{w \in D} \frac{\|D^2 g(Z_1 - Z_{-1} + w)\|}{1 + \|w\|} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \leq \mathbb{E} \left[ \sup_{w \in D} \frac{\|D^2 g(Z_1 - Z_{-1} + w)\|}{1 + \|Z_1 - Z_{-1} + w\|} \cdot \frac{1 + \|Z_1 - Z_{-1} + w\|}{1 + \|w\|} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \stackrel{Doob}{\leq} \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \left[ 1 + 2\sqrt{\mathbb{E} \left[ (Z_1(1) - Z_{-1}(1))^2 \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right]} \right] \\
& = \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \left[ 1 + 2\sqrt{2R_1(1)} \right] \\
& \leq 2 \left( \sup_{w \in D} \frac{\|D^2 g(w)\|}{1 + \|w\|} \right) \\
\text{C)} \quad & \mathbb{E} \left[ \sup_{w \in D} \frac{\|Dg(Z_1 - Z_{-1} + w)\|}{1 + \|w\|^2} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \leq \mathbb{E} \left[ \sup_{w \in D} \frac{\|Dg(Z_1 - Z_{-1} + w)\|}{1 + \|Z_1 - Z_{-1} + w\|^2} \cdot \frac{1 + \|Z_1 - Z_{-1} + w\|^2}{1 + \|w\|^2} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
& \stackrel{Doob}{\leq} 2 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \left[ 1 + 4\mathbb{E} \left[ (Z_1(1) - Z_{-1}(1))^2 \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \right] \\
& = 2 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) [1 + 8R_1(1)]
\end{aligned}$$

$$\begin{aligned}
&\leq 4 \left( \sup_{w \in D} \frac{\|Dg(w)\|}{1 + \|w\|^2} \right) \\
&\quad \text{D) } \mathbb{E} \left[ \sup_{w \in D} \frac{|g(Z_1 - Z_{-1} + w)|}{1 + \|w\|^3} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\leq \mathbb{E} \left[ \sup_{w \in D} \frac{|g(Z_1 - Z_{-1} + w)|}{1 + \|Z_1 - Z_{-1} + w\|^3} \cdot \frac{1 + \|Z_1 - Z_{-1} + w\|^3}{1 + \|w\|^3} \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \\
&\stackrel{Doob}{\leq} 4 \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right) \left[ 1 + \frac{27}{8} \mathbb{E} \left[ |Z_1(1) - Z_{-1}(1)|^3 \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \right] \\
&\leq 4 \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right) \left[ 1 + \frac{27\sqrt{2}}{4\sqrt{\pi}} (2R_1(1))^{3/2} \right] \\
&\leq \left( 4 + \frac{27\sqrt{2}}{8\sqrt{\pi}} \right) \left( \sup_{w \in D} \frac{|g(w)|}{1 + \|w\|^3} \right).
\end{aligned}$$

So:

$$\mathbb{E} \left[ \|g^{(0)}\|_M \middle| R_1, R_1^{(n)}, R_{-1}^{(n)} \right] \leq \left( 4 + \frac{27\sqrt{2}}{8\sqrt{\pi}} \right) \|g\|_M. \quad (5.23)$$

Now, the Moran model and the Wright-Fisher diffusion can be coupled using the Donnelly-Kurtz look-down construction (see the discussion below Theorem 3.4). In this construction first the Wright-Fisher diffusion  $M$  is realised and then the Moran model  $M_n$  is constructed by describing  $nM_n(s)$  as a **Binomial**( $n, M(s)$ ) random variable. Note that:

$$\begin{aligned}
&\text{A) } \mathbb{E} \sqrt{\|R_1 - R_1^{(n)}\|} \\
&\leq \mathbb{E} \sqrt{\int_0^1 \left| \frac{1}{2} X(s)(1 - X(s)) - \left( \frac{1}{2} X_n(s) + \frac{\nu_2}{n} \right) (1 - X_n(s)) \right| ds} \\
&\stackrel{Jensen}{\leq} \sqrt{\int_0^1 \mathbb{E} \left| \frac{1}{2} X(s)(1 - X(s)) - \left( \frac{1}{2} X_n(s) + \frac{\nu_2}{n} \right) (1 - X_n(s)) \right| ds} \\
&\leq \left\{ \int_0^1 \left( \frac{1}{2} \mathbb{E} [\mathbb{E}[|X_n(s) - X(s)| | X(s)]] + \frac{1}{2} \mathbb{E} [\mathbb{E}[|X_n^2(s) - X^2(s)| | X(s)]] \right. \right. \\
&\quad \left. \left. + \frac{\nu_2}{n} \mathbb{E} [\mathbb{E}[|1 - X_n(s)| | X(s)]] \right) ds \right\}^{1/2} \\
&\leq \left\{ \int_0^1 \left( \frac{1}{2} \mathbb{E} [\sqrt{\text{Var}[X_n(s) | X(s)]}] + \frac{1}{2} \mathbb{E} [\sqrt{\text{Var}[X_n^2(s) | X(s)]}] \right. \right. \\
&\quad \left. \left. + \frac{1}{2} \mathbb{E} \left[ \frac{X(s)(1 - X(s))}{n} \right] + \frac{\nu_2}{n} \mathbb{E}[1 - X(s)] \right) ds \right\}^{1/2} \\
&= \left\{ \int_0^1 \left( \frac{1}{2} \mathbb{E} \left[ \sqrt{\frac{X(s)(1 - X(s))}{n}} \right] + \frac{1}{2} \mathbb{E} \left[ \left( \frac{X(s)(1 - 7X(s) + 6nX(s))}{n^3} \right. \right. \right. \right. \\
&\quad \left. \left. \left. + \frac{X(s)(12X^2(s) - 20nX^2(s) + 4n^2X^2(s) - 6X^3(s) + 10nX^3(s))}{n^3} \right) \right] \right) ds \right\}^{1/2}
\end{aligned}$$

$$\begin{aligned}
& \left. - \frac{8X^4(s)}{n} \right)^{1/2} \Big] + \mathbb{E} \left[ \left( \frac{(X(s) + 2\nu_2)(1 - X(s))}{2n} \right) \right] ds \Big\}^{1/2} \\
& \leq \sqrt{\frac{1}{4}n^{-1/2} + \frac{1}{2}\sqrt{13n^{-3} + 16n^{-2} + 4n^{-1}} + \frac{1 + 2\nu_2}{2}n^{-1}} \\
& \text{B) } \mathbb{E} \|R_1 - R_1^{(n)}\|^{3/2} \\
& \leq \mathbb{E} \left[ \left( \int_0^1 \left| \frac{1}{2}X(s)(1 - X(s)) - \left( \frac{1}{2}X_n(s) + \frac{\nu_2}{n} \right) (1 - X_n(s)) \right| ds \right)^{3/2} \right] \\
& \stackrel{\text{Hölder}}{\leq} \mathbb{E} \left[ \int_0^1 \left| \frac{1}{2}X(s)(1 - X(s)) - \left( \frac{1}{2}X_n(s) + \frac{\nu_2}{n} \right) (1 - X_n(s)) \right|^{3/2} ds \right] \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ \mathbb{E} |X_n(s) - X(s)|^{3/2} \middle| X(s) \right] \right. \\
& \quad + \frac{\sqrt{2}}{4} \mathbb{E} \left[ \mathbb{E} |X_n^2(s) - X^2(s)|^{3/2} \middle| X(s) \right] \\
& \quad \left. + \frac{\nu_2^{3/2}}{n^{3/2}} \mathbb{E} \left[ \mathbb{E} [|1 - X_n(s)|^{3/2} | X(s)] \right] \right\} ds \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ (\text{Var} [X_n(s) | X(s)])^{3/4} \right] \right. \\
& \quad + \frac{1}{2} \mathbb{E} \left[ \mathbb{E} \left| X_n^2(s) - \left( \frac{X(s)(1 - X(s))}{n} + X^2(s) \right) \right|^{3/2} \middle| X(s) \right] \\
& \quad + \frac{1}{2} \mathbb{E} \left[ \left( \frac{X(s)(1 - X(s))}{n} \right)^{3/2} \right] \\
& \quad \left. + \frac{\nu_2^{3/2}}{n^{3/2}} \mathbb{E} \left[ \left( \frac{(1 - X(s))X(s)}{n} + (1 - X(s))^2 \right)^{3/4} \right] \right\} ds \\
& \leq \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ (\text{Var} [X_n(s) | X(s)])^{3/4} \right] + \left( \frac{1}{16} + \nu_2^{3/2} \right) n^{-3/2} \right. \\
& \quad \left. + \frac{1}{2} \mathbb{E} \left[ \mathbb{E} \left| X_n^2(s) - \left( \frac{X(s)(1 - X(s))}{n} + X^2(s) \right) \right|^{3/2} \middle| X(s) \right] \right\} ds \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ (\text{Var} [X_n(s) | X(s)])^{3/4} \right] \right. \\
& \quad \left. + \frac{1}{2} \mathbb{E} \left[ (\text{Var} [X_n^2(s) | X(s)])^{3/4} \right] + \left( \frac{1}{16} + \nu_2^{3/2} \right) n^{-3/2} \right\} ds \\
& = \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ \left( \frac{X(s)(1 - X(s))}{n} \right)^{3/4} \right] + \frac{1}{2} \mathbb{E} \left[ \left( \frac{X(s)(1 - 7X(s))}{n^3} \right) \right] \right\} ds
\end{aligned}$$

$$\begin{aligned}
& + \frac{X(s)(6nX(s) + 12X^2(s) - 20nX^2(s) + 4n^2X^2(s) - 6X^3(s))}{n^3} \\
& + \frac{X(s)(10nX^3(s) - 8n^2X^3(s))}{n^3} \Big)^{3/4} \Big] + \left( \frac{1}{16} + \nu_2^{3/2} \right) n^{-3/2} \Big\} ds \\
& \leq \sqrt{3} \left( \frac{\sqrt{2}}{32} n^{-3/4} + \frac{1}{2} (13n^{-3} + 16n^{-2} + 4n^{-1})^{3/4} + \left( \frac{1}{16} + \nu_2^{3/2} \right) n^{-3/2} \right) \\
\text{C) } & \mathbb{E} \sqrt{\|R_{-1} - R_{-1}^{(n)}\|} \\
& \leq \mathbb{E} \sqrt{\int_0^1 \left| \frac{1}{2} X(s)(1 - X(s)) - \left( \frac{1}{2}(1 - X_n(s)) + \frac{\nu_1}{n} \right) X_n(s) \right| ds} \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{\int_0^1 \mathbb{E} \left| \frac{1}{2} X(s)(1 - X(s)) - \left( \frac{1}{2}(1 - X_n(s)) + \frac{\nu_1}{n} \right) X_n(s) \right| ds} \\
& \leq \left\{ \int_0^1 \left( \frac{1}{2} \mathbb{E} [\mathbb{E}[|X_n(s) - X(s)| | X(s)]] + \frac{1}{2} \mathbb{E} [\mathbb{E}[|X_n^2(s) - X^2(s)| | X(s)]] \right. \right. \\
& \quad \left. \left. + \frac{\nu_1}{n} \mathbb{E} [\mathbb{E}[X_n(s) | X(s)]] \right) ds \right\}^{1/2} \\
& \leq \left\{ \int_0^1 \left( \frac{1}{2} \mathbb{E} [\sqrt{\text{Var}[X_n(s) | X(s)]}] + \frac{1}{2} \mathbb{E} [\sqrt{\text{Var}[X_n^2(s) | X(s)]}] \right. \right. \\
& \quad \left. \left. + \frac{1}{2} \mathbb{E} \left[ \frac{X(s)(1 - X(s))}{n} \right] + \frac{\nu_1}{n} \mathbb{E}[X(s)] \right) ds \right\}^{1/2} \\
& = \left\{ \int_0^1 \left( \frac{1}{2} \mathbb{E} \left[ \sqrt{\frac{X(s)(1 - X(s))}{n}} \right] + \frac{1}{2} \mathbb{E} \left[ \left( \frac{X(s)(1 - 7X(s) + 6nX(s))}{n^3} \right. \right. \right. \right. \\
& \quad \left. \left. \left. + \frac{X(s)(12X^2(s) - 20nX^2(s) + 4n^2X^2(s) - 6X^3(s) + 10nX^3(s))}{n^3} \right. \right. \right. \\
& \quad \left. \left. \left. - \frac{8X^4(s)}{n} \right)^{1/2} \right] + \mathbb{E} \left[ \frac{X(s)(1 - X(s) + 2\nu_1)}{2n} \right] \right) ds \right\}^{1/2} \\
& \leq \sqrt{\frac{1}{4} n^{-1/2} + \frac{1}{2} \sqrt{13n^{-3} + 16n^{-2} + 4n^{-1}} + \frac{1 + 2\nu_1}{2} n^{-1}} \\
\text{D) } & \mathbb{E} \|R_{-1} - R_{-1}^{(n)}\|^{3/2} \\
& \leq \mathbb{E} \left[ \left( \int_0^1 \left| \frac{1}{2} X(s)(1 - X(s)) - \left( \frac{1}{2}(1 - X_n(s)) + \frac{\nu_1}{n} \right) X_n(s) \right| ds \right)^{3/2} \right] \\
& \stackrel{\text{Hölder}}{\leq} \mathbb{E} \left[ \int_0^1 \left| \frac{1}{2} X(s)(1 - X(s)) - \left( \frac{1}{2}(1 - X_n(s)) + \frac{\nu_1}{n} \right) X_n(s) \right|^{3/2} ds \right] \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} [\mathbb{E}[|X_n(s) - X(s)|^{3/2} | X(s)]] \right.
\end{aligned}$$



$$\begin{aligned}
& + \frac{\sqrt{2}}{4} \mathbb{E} \left[ \mathbb{E} |X_n^2(s) - X^2(s)|^{3/2} \middle| X(s) \right] \\
& + \frac{\nu_1^{3/2}}{n^{3/2}} \mathbb{E} \left[ \mathbb{E} [|X_n(s)|^{3/2} | X(s)] \right] \Bigg\} ds \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ (\text{Var} [X_n(s) | X(s)])^{3/4} \right] \right. \\
& + \frac{1}{2} \mathbb{E} \left[ \mathbb{E} \left| X_n^2(s) - \left( \frac{X(s)(1-X(s))}{n} + X^2(s) \right) \right|^{3/2} \middle| X(s) \right] \\
& + \frac{1}{2} \mathbb{E} \left[ \left( \frac{X(s)(1-X(s))}{n} \right)^{3/2} \right] \\
& + \left. \frac{\nu_1^{3/2}}{n^{3/2}} \mathbb{E} \left[ \left( \frac{(1-X(s))X(s)}{n} + X^2(s) \right)^{3/4} \right] \right\} ds \\
& \leq \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ (\text{Var} [X_n(s) | X(s)])^{3/4} \right] + \left( \frac{1}{16} + \nu_1^{3/2} \right) n^{-3/2} \right. \\
& + \frac{1}{2} \mathbb{E} \left[ \mathbb{E} \left| X_n^2(s) - \left( \frac{X(s)(1-X(s))}{n} + X^2(s) \right) \right|^{3/2} \middle| X(s) \right] \Bigg\} ds \\
& \stackrel{\text{Jensen}}{\leq} \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ (\text{Var} [X_n(s) | X(s)])^{3/4} \right] + \frac{1}{2} \mathbb{E} \left[ (\text{Var} [X_n^2(s) | X(s)])^{3/4} \right] \right. \\
& + \left. \left( \frac{1}{16} + \nu_1^{3/2} \right) n^{-3/2} \right\} ds \\
& = \sqrt{3} \int_0^1 \left\{ \frac{\sqrt{2}}{4} \mathbb{E} \left[ \left( \frac{X(s)(1-X(s))}{n} \right)^{3/4} \right] + \frac{1}{2} \mathbb{E} \left[ \left( \frac{X(s)(1-7X(s))}{n^3} \right. \right. \right. \\
& + \frac{X(s)(6nX(s) + 12X^2(s) - 20nX^2(s) + 4n^2X^2(s) - 6X^3(s))}{n^3} \\
& + \left. \left. \left. \frac{X(s)(10nX^3(s) - 8n^2X^3(s))}{n^3} \right)^{3/4} \right] + \left( \frac{1}{16} + \nu_1^{3/2} \right) n^{-3/2} \right\} ds \\
& \leq \sqrt{3} \left( \frac{\sqrt{2}}{32} n^{-3/4} + \frac{1}{2} (13n^{-3} + 16n^{-2} + 4n^{-1})^{3/4} + \left( \frac{1}{16} + \nu_1^{3/2} \right) n^{-3/2} \right). \tag{5.24}
\end{aligned}$$

Also:

$$\begin{aligned}
\text{A) } \mathbb{E} \|I - I_n\| & \leq \int_0^1 (\nu_1 + \nu_2) \mathbb{E} [\mathbb{E} [|X_n(s) - X(s)| | X(s)]] ds \\
& \leq \int_0^1 (\nu_1 + \nu_2) \mathbb{E} \left[ \sqrt{\text{Var} [X_n(s) | X(s)]} \right] ds
\end{aligned}$$

$$\begin{aligned}
&\leq \int_0^1 (\nu_1 + \nu_2) \mathbb{E} \left[ \sqrt{\frac{X(s)(1-X(s))}{n}} \right] ds \leq \frac{1}{2}(\nu_1 + \nu_2)n^{-1/2} \\
&\quad \text{B) } \mathbb{E} \|I - I_n\|^3 \stackrel{\text{Hölder}}{\leq} \int_0^1 \mathbb{E} |(\nu_1 + \nu_2)(X_n(s) - X(s))|^3 ds \\
&\leq (\nu_1 + \nu_2)^3 \int_0^1 \left[ \mathbb{E} \left[ (X_n(s) - X(s))^4 \right] \right]^{3/4} ds \\
&\leq (\nu_1 + \nu_2)^3 \int_0^1 \mathbb{E} \left[ \left( \frac{X(s)(1-7X(s) + 7nX(s) + 12X^2(s) - 18nX^2(s))}{n^4} \right. \right. \\
&\quad \left. \left. + \frac{X(s)(6n^2X^2(s) - 6X^3(s) + 11nX^3(s) - 6n^2X^3(s) + n^3X^3(s))}{n^4} \right)^{3/4} \right] ds \\
&\leq (\nu_1 + \nu_2)^3 (13n^{-4} + 18n^{-3} + 6n^{-2} + n^{-1})^{3/4}. \tag{5.25}
\end{aligned}$$

We now combine (5.18), (5.19), (5.20), (5.21), (5.22), (5.23), (5.24), (5.25) to obtain:

$$\begin{aligned}
&|\mathbb{E}g(M) - \mathbb{E}g(M_n)| \\
&\leq \|g\|_M \left[ 9 + 2\sqrt{\frac{1}{8} + \frac{\nu_1}{n}} + \frac{12\nu_1}{n} + \frac{243}{8} \cdot \left( \frac{3}{64} + \frac{6\nu_1 + 1}{8n} + \frac{\nu_1 + \nu_1^2}{n^2} \right)^{3/4} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right] \\
&\quad \cdot \left\{ \left( 2 + \frac{27\sqrt{2}}{16\sqrt{\pi}} \right) \sqrt{\frac{1}{4}n^{-1/2} + \frac{1}{2}\sqrt{13n^{-3} + 16n^{-2} + 4n^{-1}} + \frac{1+2\nu_2}{2}n^{-1}} \right. \\
&\quad \left. + \frac{27\sqrt{2}}{2\sqrt{\pi}}\sqrt{3} \left( \frac{\sqrt{2}}{32}n^{-3/4} + \frac{1}{2}(13n^{-3} + 16n^{-2} + 4n^{-1})^{3/4} + \left( \frac{1}{16} + \nu_2^{3/2} \right) n^{-3/2} \right) \right. \\
&\quad \left. + n^{-1} \left[ \frac{30 + \frac{27}{4} \cdot 5^{1/3}}{\sqrt{\pi} \log 2} \sqrt{\log(n^2/4)} + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} \left( \frac{1}{8} + \frac{\nu_2}{n} \right)^{3/2} \right) \right. \right. \\
&\quad \left. \cdot \left( \frac{1}{8} + \frac{\nu_2}{n} \right) (1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 4 \log n}{\log \log(n^2 + 2)} \right] \\
&\quad \left. + n^{-2} \frac{9\sqrt{\frac{1}{8} + \frac{\nu_2}{n}}}{2} \left( 1 + \frac{3}{8}n^2 + 3\nu_2n \right)^{1/2} \left[ 4 + \frac{16701 + 1024(\log n)^3}{(\log \log(n^2 + 3))^3} \right]^{1/3} \right. \\
&\quad \left. + n^{-3} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(n^2/4 + \nu_2n))^{3/2} + 8 + \frac{33402 + 2048(\log n)^3}{(\log \log(n^2 + 3))^3} \right] \right\} \\
&\quad + \|g\|_M \left( 9 + \frac{243}{64\sqrt{\pi}} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right) \\
&\quad \cdot \left\{ \left( 2 + \frac{27\sqrt{2}}{16\sqrt{\pi}} \right) \sqrt{\frac{1}{4}n^{-1/2} + \frac{1}{2}\sqrt{13n^{-3} + 16n^{-2} + 4n^{-1}} + \frac{1+2\nu_1}{2}n^{-1}} \right. \\
&\quad \left. + \frac{27\sqrt{2}}{2\sqrt{\pi}}\sqrt{3} \left( \frac{\sqrt{2}}{32}n^{-3/4} + \frac{1}{2}(13n^{-3} + 16n^{-2} + 4n^{-1})^{3/4} + \left( \frac{1}{16} + \nu_1^{3/2} \right) n^{-3/2} \right) \right.
\end{aligned}$$

$$\begin{aligned}
& + n^{-1} \left[ \frac{30 + \frac{27}{4} \cdot 5^{1/3}}{\sqrt{\pi \log 2}} \sqrt{\log(n^2/4)} + \left( 1 + \left( \frac{3}{2} \right)^3 \sqrt{\frac{2}{\pi}} \left( \frac{1}{8} + \frac{\nu_1}{n} \right)^{3/2} \right) \right. \\
& \cdot \left( \frac{1}{8} + \frac{\nu_1}{n} \right) (1 + 2e^{-1}) + 1 + \frac{\log(1 + 2e^{-1}) + 4 \log n}{\log \log(n^2 + 2)} \Big] \\
& + n^{-2} \frac{9 \sqrt{\frac{1}{8} + \frac{\nu_1}{n}}}{2} \left( 1 + \frac{3}{8} n^2 + 3\nu_1 n \right)^{1/2} \left[ 4 + \frac{16701 + 1024(\log n)^3}{(\log \log(n^2 + 3))^3} \right]^{1/3} \\
& + n^{-3} \left[ \frac{2160}{\sqrt{\pi}(\log 2)^{3/2}} (\log(n^2/4 + \nu_1 n))^{3/2} + 8 + \frac{33402 + 2048(\log n)^3}{(\log \log(n^2 + 3))^3} \right] \Big\} \\
& + \|g\|_M \left( 4 + \frac{27\sqrt{2}}{8\sqrt{\pi}} \right) \\
& \cdot \left[ \frac{1}{2} (1 + 2\nu_2)(\nu_1 + \nu_2) n^{-1/2} + 2(\nu_1 + \nu_2)^3 (13n^{-4} + 18n^{-3} + 6n^{-2} + n^{-1})^{3/4} \right] \\
& \leq \|g\|_M \left\{ \left( 18 + \nu_1^{1/2} + 47\nu_1^{3/4} + 31\nu_1^{3/2} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right) \right. \\
& \cdot \left( 1.02 \cdot 10^6 + 425\nu_2^{1/2} + 623\nu_2 + 39\nu_2^{3/2} + 7\nu_2^{5/2} \right) \\
& + (12 + 3\nu_2 + 3\nu_2^2 + 9\nu_2^3) \left( 1.02 \cdot 10^6 + 425\nu_1^{1/2} + 623\nu_1 + 39\nu_1^{3/2} + 7\nu_1^{5/2} \right) \\
& + 7 \left( \frac{1}{2} (1 + 2\nu_2)(\nu_1 + \nu_2) + 31(\nu_1 + \nu_2)^3 \right) \Big\} n^{-1/4} \\
& + \|g\|_M 2112 \left[ \left( 18 + \nu_1^{1/2} + 47\nu_1^{3/4} + 31\nu_1^{3/2} + \nu_2 + 3\nu_2^2 + 9\nu_2^3 \right) (\log(n^2/4 + \nu_2 n))^{3/2} \right. \\
& + (12 + 3\nu_2 + 3\nu_2^2 + 9\nu_2^3) (\log(n^2/4 + \nu_1 n))^{3/2} \Big] n^{-3}. \quad \square
\end{aligned}$$

**Remark 5.1.** The term of order  $n^{-1/4}$  appearing in the bound obtained in Theorem 3.4 is unexpected. It comes from the comparison of the two time changes  $R_1^{(n)}$  and  $R_1$  applied at certain points in the proof to the Poisson process and to Brownian Motion respectively.

One would expect  $n^{-1} \sqrt{\log n^2}$  to be closer to the true speed of convergence. This is because our process  $M_n$  can be expressed as a difference of two scaled Poisson processes with parameters of order  $n^2$ , which resemble scaled random walks stopped at the  $n^2$ -th step. Order  $n^{-1} \sqrt{\log n^2}$  would therefore be in line with the results presented in [Bar90] in the context of scaled random walks and indeed with Theorem 3.1.

We can therefore guess that our bound is not sharp, yet a method avoiding the comparison of the aforementioned time changes is needed to improve it.

**Remark 5.2.** A strategy similar to the one used in the proof of Theorem 3.4 may be used to obtain bounds on the distance between other continuous-time Markov chains and diffusions. A key ingredient in the proof is, however, a way of coupling the two for any fixed time.

### Appendix: Proof of Proposition 2.3

Note that the proof of Proposition 3.1 of [BJ09] readily applies in this case up to and excluding (3.4) and it suffices to prove that  $\liminf_{n \rightarrow \infty} \mathbb{P}[Y_n \in B] \geq \mathbb{P}[Z \in B]$  for all sets  $B$  of the form  $B = \bigcap_{1 \leq l \leq L} B_l$ , where  $B_l = \{w \in D : \|w - s_l\| < \gamma_l\}$ ,  $s_l \in C([0, 1], \mathbb{R})$  and  $\gamma_l$  is such that  $\mathbb{P}[Z \in \partial B_l] = 0$ .

We will condition on the fact that the minimum holding time (interval of constancy of  $Y_n$ ) is of length greater than  $r_n = \lambda_n^{-3}$ . It follows from Theorems 2.1 and 2.2 of Chapter 5 in [Dev86] that if we condition on the number of holding times being equal to  $i$ , their lengths are distributed uniformly over the simplex  $A_i = \{(x_1, \dots, x_i) : x_j \geq 0, \sum_{j=1}^i x_j \leq 1\}$ . Note that the probability of the minimum of them being greater or equal to  $r_n$  is  $(1 - ir_n)^i$  if  $i \leq r_n$  and 0 otherwise. This is because  $\text{Vol}(A_i) = \frac{1}{i!}$  and  $\text{Vol}\left(\{(x_1, \dots, x_i) : x_j \geq r_n, \sum_{j=1}^i x_j \leq 1\}\right) = \frac{(1 - ir_n)^i}{i!}$ . Therefore:

$$\begin{aligned} & \mathbb{P}[\text{minimal waiting time} \geq r_n] \\ &= \sum_{i=1}^{\infty} \mathbb{P}[\text{minimal waiting time} \geq r_n | \#\text{waiting times} = i] \mathbb{P}[\#\text{waiting times} = i] \\ &= \sum_{i=1}^{\lfloor \lambda_n^3 \rfloor} (1 - i\lambda_n^{-3})^i e^{-\lambda_n} \frac{(\lambda_n)^{i-1}}{(i-1)!} \xrightarrow{n \rightarrow \infty} 1. \end{aligned} \quad (5.26)$$

To see this note the following:

$$\begin{aligned} \text{A) } \sum_{i=\lceil \lambda_n^{5/4} \rceil}^{\lfloor \lambda_n^3 \rfloor} (1 - i\lambda_n^{-3})^i e^{-\lambda_n} \frac{(\lambda_n)^{i-1}}{(i-1)!} &\leq e^{-\lambda_n} \left( \lambda_n^3 - \lambda_n^{5/4} \right) \left( 1 - \lambda_n^{-7/4} \right)^{\lambda_n^{5/4}} \frac{\lambda_n^{\lambda_n^{5/4}-1}}{\left( \lceil \lambda_n^{5/4} \rceil - 1 \right)!} \\ &\leq \frac{\lambda_n^2 \lceil \lambda_n^{5/4} \rceil}{e^{\lambda_n}} \cdot \lambda_n^{-\frac{1}{8} \lceil \lambda_n^{5/4} \rceil + \frac{9}{8} \lceil \lambda_n^{9/8} \rceil} \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

$$\text{B) } \sum_{i=1}^{\lceil \lambda_n^{5/4} \rceil - 1} (1 - i\lambda_n^{-3})^i e^{-\lambda_n} \frac{(\lambda_n)^{i-1}}{(i-1)!} \geq (1 - \lambda_n^{-7/4})^{\lceil \lambda_n^{5/4} \rceil} e^{-\lambda_n} \sum_{i=1}^{\lceil \lambda_n^{5/4} \rceil - 1} \frac{(\lambda_n)^{i-1}}{(i-1)!} \xrightarrow{n \rightarrow \infty} 1,$$

where the convergence in B) holds since  $(1 - \lambda_n^{-7/4})^{\lceil \lambda_n^{5/4} \rceil} \rightarrow 1$ ,  $e^{-\lambda_n} \sum_{i=1}^{\infty} \frac{(\lambda_n)^{i-1}}{(i-1)!} = 1$  and:

$$e^{-\lambda_n} \sum_{i=\lceil \lambda_n^{5/4} \rceil}^{\infty} \frac{(\lambda_n)^{i-1}}{(i-1)!} \leq e^{-\lambda_n} \frac{\lambda_n^{\lceil \lambda_n^{5/4} \rceil}}{\lceil \lambda_n^{5/4} \rceil!} \cdot \frac{\lceil \lambda_n^{5/4} \rceil + 1}{\lceil \lambda_n^{5/4} \rceil + 1 - \lambda_n} \xrightarrow{n \rightarrow \infty} 0$$

for instance, by Proposition A.2.3(ii) of [BHJ92]. Furthermore, note that for  $g_{l,n}^*$  defined by (3.6) in [BJ09]:

$$\begin{aligned}
& \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Y_n) \right] \\
&= \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Y_n) \middle| \text{minimal waiting time} \geq r_n \right] \mathbb{P} [\text{minimal waiting time} \geq r_n] \\
&\quad + \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Y_n) \middle| \text{minimal waiting time} < r_n \right] \mathbb{P} [\text{minimal waiting time} < r_n] \\
&= \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Y_n) \middle| \text{minimal waiting time} \geq r_n \right] \mathbb{P} [\text{minimal waiting time} \geq r_n] \\
&\hspace{15em} (5.27)
\end{aligned}$$

because:

$$\begin{aligned}
0 &\leq \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Y_n) \middle| \text{minimal waiting time} < r_n \right] \mathbb{P} [\text{minimal waiting time} < r_n] \\
&\leq \liminf_{n \rightarrow \infty} \mathbb{P} [\text{minimal waiting time} < r_n] \stackrel{(5.26)}{=} 0.
\end{aligned}$$

Following the same steps as in [BJ09], we obtain:

$$\begin{aligned}
& \liminf_{n \rightarrow \infty} \mathbb{P} [Y_n \in B] \geq \liminf_{n \rightarrow \infty} \mathbb{P} [Y_n \in B \text{ and minimal waiting time} \geq r_n] \\
&\geq \liminf_{n \rightarrow \infty} \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Y_n) \middle| \text{minimal waiting time} \geq r_n \right] \mathbb{P} [\text{minimal waiting time} \geq r_n] \\
&\stackrel{(2.3), (5.27)}{\geq} \liminf_{n \rightarrow \infty} \left\{ \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Z_n) \right] - C\tau_n \left\| \prod_{l=1}^L g_{l,n}^* \right\|_{M^0} \right\} \\
&\geq \liminf_{n \rightarrow \infty} \left\{ \mathbb{E} \left[ \prod_{l=1}^L g_{l,n}^*(Z_n) \right] - C''\tau_n p_n^2 (\epsilon\gamma)^{-2} \eta_n^{-3} \right\} \\
&\stackrel{\text{Fatou}}{\geq} \mathbb{E} \left[ \liminf_{n \rightarrow \infty} \prod_{l=1}^L g_{l,n}^*(Z_n) \right] \geq \mathbb{P} \left[ \bigcap_{1 \leq l \leq L} (\|Z - s_l\| < \gamma_l(1 - \theta)) \right].
\end{aligned}$$

□

## Acknowledgements

The author would like to thank Professor Gesine Reinert, Professor Alison Etheridge and Professor Andrew Barbour for many helpful discussions and Professor Gesine Reinert for constructive comments on the early versions of the paper. The author is also grateful to Dr Sebastian Vollmer for pointing out a mistake in an early version of the proof of Proposition 4.3.

## References

- [Bar90] A.D. Barbour. Stein's Method for Diffusion Approximations. *Probability Theory and Related Fields*, 84:297–322, 1990.
- [BHJ92] A.D. Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*. Oxford Studies in Probability. Clarendon Press, 1992.
- [BJ09] A.D. Barbour and S. Janson. A functional combinatorial central limit theorem. *Electronic Journal of Probability*, 14(81):2352–2370, 2009.
- [CD13] L. Coutin and L. Decreusefond. Stein's method for Brownian Approximations. *Communications on Stochastic Analysis*, 7(3):349–372, 2013.
- [CGS11] L.H.Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal Approximation by Steins Method*. Probability and Its Applications. Springer Verlag, 2011.
- [Das11] G. Dasarathy. A Simple Probability Trick for Bounding the Expected Maximum of  $n$  Random Variables. [www.cs.cmu.edu/~gautamd/Files/maxGaussians.pdf](http://www.cs.cmu.edu/~gautamd/Files/maxGaussians.pdf), 2011. Accessed on 30/11/2016.
- [Dev86] L. Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, 1986.
- [DK96] P. Donnelly and T.G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *The Annals of Probability*, 24(2):698–742, 1996.
- [EK86] S.N. Ethier and T.G. Kurtz. *Markov processes: characterization and convergence*. Wiley, New York, 1986.
- [Eth12] A. Etheridge. *Some Mathematical Models from Population Genetics*. Lecture Notes in Mathematics. Springer-Verlag, 2012.
- [FN10] M. Fischer and G. Nappo. On the Moments of the Modulus of Continuity of Ito Processes. *Stochastic Analysis and Applications*, 28(1):103–122, 2010.
- [KDV17] M.J. Kasprzak, A. B. Duncan, and S.J. Vollmer. Note on A. Barbour's paper on Steins method for diffusion approximations. *Electron. Commun. Probab.*, 22(23):1–8, 2017.
- [Kur12] T. Kurtz. Time change equations for diffusion processes. [www.math.wisc.edu/~kurtz/Lectures/Chenai12.pdf](http://www.math.wisc.edu/~kurtz/Lectures/Chenai12.pdf), 2012. Accessed on 19/05/2016.
- [LRS17] C. Ley, G. Reinert, and Y. Swan. Stein's method for comparison of univariate distributions. *Probability Surveys*, 14:1–52, 2017.
- [Mor58] P.A.P. Moran. Random Processes in Genetics. *Proc. Camb. Phil. Soc.*, (54):60–71, 1958.
- [NP12] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus*. Cambridge tracts in Mathematics. Cambridge University Press, 2012.
- [Ros11] N. Ross. Fundamentals of Stein's Method. *Probability Surveys*, 8:210–293, 2011.
- [RY91] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*.

- A Series of Comprehensive Studies in Mathematics. 1991.
- [Ste72] Ch. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.*, 2:583–602, 1972.
- [SV79] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer, 1979.
- [Swa16] Y. Swan. A gateway to Stein's Method. <https://sites.google.com/site/steinsmethod/home>, 2016. Accessed on 19/05/2016.
- [Wri31] S. Wright. Evolution in Mendelian populations. *Genetics*, (16):97–159, 1931.