Probabilistic Analysis of Replicator-Mutator Equations

Lijun Bo∗ Huafu Liao†

March 16, 2020

Abstract

This paper introduces a general class of Replicator-Mutator equations on a multi-
dimensional fitness space. We establish a novel probabilistic representation of weak
solutions of the equation by using the theory of Fockker-Planck-Kolmogorov (FPK)
equations and a martingale extraction approach. The examples with closed-form prob-
babilistic solutions for different fitness functions considered in the existing literature are
provided. We also construct a particle system and prove a general convergence result
to any solution to the FPK equation associated with the extended Replicator-Mutator
equation with respect to a Wasserstein-like metric adapted to our probabilistic frame-
work.

AMS 2000 subject classifications: 92B05, 35K15, 60H10, 60G46.

Keywords: Replicator-Mutator equations; Probabilistic representation; FPK equa-
tions; martingale extraction; particle system.

1 Introduction

In this paper, we introduce a general class of Replicator-Mutator equations on a fitness
space specified as a domain of \( \mathbb{R}^n \) with \( n \geq 1 \). The classical replicator-mutator equation
in evolutionary genetics described as an integro-differential Cauchy problem on the one-
dimensional fitness space \( \mathbb{R} \) is proposed by Kimura (1965). This type of equations with the
linear fitness function is used to model the evolution of RNA virus populations by Tsimring
(1996) based on a mean-field approach. For the case where some phenotypes are infinitely
well-adapted (this typically corresponds to the fitness functions which are unbounded from
above), a so-called replicator-mutator model with cut-off at large phenotype is studied
by Rouzine et al. (2003) and Sniegowski and Gerrish (2010), while Rouzine et al. (2008)
propose a proper stochastic treatment for large phenotypic trait region for the linear fitness
case.

In recent years, rigorous mathematical treatments on existence and behaviours of
solutions of Replicator-Mutator equations describing as a Cauchy problem on the one-
dimensional fitness space for different types of fitness functions are developed. Alfaro and Carles
(2014) reduce the equation with the linear fitness function to a standard heat equation
by applying a tricky transform of the solution based on the Avron-Herbst formula, and
therefore, one can compute its solution explicitly. For the quadratic fitness case, a earlier
work Fleming (1979) studies a model for inheritance of continuous polygenic traits and
the equilibrium density of gametic types is found approximately. Bürger (1998) provides
an overview of the mathematical properties of various deterministic mutation-selection

∗Email: lijunbo@ustc.edu.cn, School of Mathematical Sciences, University of Science and Technology of
China, Hefei, Anhui Province, 230026, China, and Wu Wen Tsun Key Laboratory of Mathematics, Chinese
Academy of Science, Hefei, Anhui Province 230026, China.
†E-mail: lhflhf@mail.ustc.edu.cn, School of Mathematical Sciences, University of Science and Technol-
ogy of China, Hefei, Anhui Province, 230026, China.
models. Recently, Alfaro and Carles (2017) can also transform the equation to a heat equation by using a generalized lens transform of the Schrödinger equation. They also prove that, for any initial density, there is always extinction of the equation with the positive quadratic fitness function at a finite time. Alfaro and Veruet (2019) establish an explicit solution representation of the equation as a non-local Cauchy problem via the underlying Schrödinger spectral elements when the fitness function is confining, and further a result on the long time behaviour related to that in Bürger (1988) in terms of the principal eigenfunction is proved therein.

The methods used in the papers reviewed above rely on the solution-based transform or the eigenfunction expansion of operators. Our approach used in this paper to study the Replicator-Mutator equations is from a completely probabilistic perspective. To be more precise, the first goal of this paper is to provide a novel probabilistic representation of the (weak) solution of a general class of Replicator-Mutator equations on a multi-dimensional fitness space. We prove that the solution of the extended Replicator-Mutator equation can be in fact expressed in terms of probability transition density functions for some Itô diffusion processes. The method used in this paper is jointly based on the probability representation of solutions to a class of Fockker-Planck-Kolmogorov (FPK) equations and a martingale extraction approach (see Qin and Linetsky (2016) where this approach is applied to the pricing of derivatives). Motivated by the probabilistic solution representation of the equation, we design a particle system in which the state process of homogeneous particles follows a system of stochastic differential equations (SDEs). The second goal of this paper is to establish a general convergence result (with respect to a Wasserstein-like distance adapted to our probabilistic framework) to any solution to our FPK equation associated with the extended Replicator-Mutator equation. In particular, easy byproducts of this result are uniqueness of solutions to the extended Replicator-Mutator equation and an explicit convergence rate to the solution of the equation. The strategy used in this paper is based upon a propagator method (see Del Moral and Miclo (2000), Sznitman (1991) and Xu (2018)). We also provide examples which admit the closed-form probabilistic solution of the extended Replicator-Mutator equation for different fitness functions considered in the aforementioned works.

The paper is organized as follows: we introduce in Section 2 the extended Replicator-Mutator equation under a probabilistic framework. Section 3 establishes a novel probabilistic representation of the (weak) solution of the extended Replicator-Mutator equation and provide examples with different fitness functions considered in the existing literature. Section 4 constructs a particle system and prove a general convergence result to any solution to the FPK equation associated with the extended Replicator-Mutator equation. Thus result is usually called propagation of chaos which was first formulated by Kac (1956).

2 Extended Replicator-Mutator equations

In this section, we introduce a class of extended Replicator-Mutator (RM) equations as a non-local Cauchy problem on a domain in \( \mathbb{R}^n \) for \( n \geq 1 \) using a probabilistic framework.

2.1 Itô diffusion process

This section presents a class of SDEs which is related to the representation of our extended RM equation on a domain \( D \) in \( \mathbb{R}^n \). Let \( T > 0 \) be an arbitrary fixed time horizon and \( D \) a domain in \( \mathbb{R}^n \), i.e., an open connected subset of \( \mathbb{R}^n \). Consider the continuous functions \( b : D \to \mathbb{R}^{n \times 1} \) and \( \sigma : D \to \mathbb{R}^{n \times m} \), where \( m \geq 1 \). We then introduce the following second-order differential operator acted on \( C^2(D) \), which is given by, for \( f \in C^2(D) \),

\[
Af(x) := b(x)^T \nabla_x f(x) + \frac{1}{2} \text{tr}[\sigma \sigma^T (x) \nabla_x^2 f(x)], \quad x \in D.
\]
where $\nabla_x = (\partial_{x_1}, \ldots, \partial_{x_n})^\top$ denotes the gradient operator, $\nabla_x^2$ is the corresponding Hessian matrix and $\text{tr}$ denotes the trace operator. The operator $\mathcal{A}$ acted on $C^2(D)$ is in fact the infinitesimal generator of the (Markovian) solution of SDE given by, for $(t, x) \in [0, T] \times D$,
\begin{align}
    dX^{t,x}_s &= b(X^{t,x}_s)ds + \sigma(X^{t,x}_s)dW_s, \quad s \in [t, T] \\
    X^{t,x}_t &= x \in D, \quad (2)
\end{align}
where $W = (W_t)_{t \in [0,T]}$ is an $m$-dimensional Brownian motion on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ satisfying the usual conditions.

We impose the following assumptions on SDE (2) so to guarantee the existence a unique strong solution of the equation:

$(\mathbf{A}_{b,\sigma})$ (i) $b : D \rightarrow \mathbb{R}^{n \times 1}$ and $\sigma : D \rightarrow \mathbb{R}^{n \times m}$ are locally Lipschitz continuous; or (ii) for $m = n = 1$, $b : D \rightarrow \mathbb{R}$ is locally Lipschitz continuous and $\sigma : D \rightarrow \mathbb{R}$ is Hölder continuous with exponent $\gamma \in (0, 1]$.

$(\mathbf{A}_X)$ For all $(t, x) \in [0, T] \times D$, the solution $X^{t,x} = (X^{t,x}_s)_{s \in [t,T]}$ of (2) neither explodes nor leaves $D$ before $T$, i.e., $\mathbb{P}(\sup_{s \in [0,T]} |X^{t,x}_s|) = 1$ and $\mathbb{P}(X^{t,x}_s \in D, \forall s \in [t,T]) = 1$.

By Theorem 5.2.5 of Karatzas and Shreve (1991), the assumption $(\mathbf{A}_{b,\sigma})$-(i) implies that SDE (2) admits a unique strong solution $X^{t,x}$ up to a possibly finite random explosion time. Then, the assumption $(\mathbf{A}_X)$ yields that this explosion time should be greater than $T$, $\mathbb{P}$-a.s., and hence $X^{t,x}$ is well-defined on $[t, T]$. For the case with $m = n = 1$ (i.e., SDE (2) is a one-dimensional equation), by Proposition 5.2.13 of Karatzas and Shreve (1991), the assumption $(\mathbf{A}_{b,\sigma})$-(ii) gives that SDE (2) admits a unique strong solution $X^{t,x}$ up to a possibly finite random explosion time. Therefore, the assumption $(\mathbf{A}_X)$ yields that this explosion time is greater than $T$, a.s., and hence $X^{t,x}$ is well-defined on $[t, T]$.

### 2.2 Extended RM equations on $[0, T] \times D$

This section introduces a class of extended RM equations as a non-local Cauchy problem on $[0, T] \times D$. More precisely, let $\mathcal{A}^*$ be the adjoint operator of $\mathcal{A}$ defined by (1)\(^1\). Then, the extended RM equation considered in this paper, is given by, for $(t, x) \in (0, T] \times D$,
\begin{align}
    \begin{cases}
        \partial_t u(t, x) = \mathcal{A}^* u(t, x) + \left( g(x) - \int_D g(y) u(t, y) dy \right) u(t, x); \\
        u(0, x) = u_0(x),
    \end{cases} \quad (3)
\end{align}
where $u_0 : D \rightarrow \mathbb{R}_+$ is a probability density function, i.e., $\int_D u_0(x) dx = 1$, and $g : D \rightarrow \mathbb{R}$ is referred to as the fitness function (see, e.g. Kimura (1965) and Tsimring (1996)). The condition satisfied by the fitness function $g$ will be imposed later.

We next give the definition of a weak solution of the above extended RM equation (3). To this purpose, let $C_0^\infty(D)$ be the space of infinitely differentiable functions with compact support. Then, a function $u : [0, T] \times D \rightarrow \mathbb{R}$ is called a weak solution of the RM equation (3) if it satisfies the following variational form: for all test functions $f \in C_0^\infty(D)$,
\begin{align}
    \langle u(t), f \rangle = \langle u_0, f \rangle + \int_0^t \langle u(s), (\mathcal{A} + g)f \rangle ds - \int_0^t \langle u(s), g \rangle \langle u(s), f \rangle ds, \quad (4)
\end{align}
where the integral $\langle u(t), f \rangle := \int_D u(t, x) f(x) dx$. One of objectives of the paper is to establish a closed-form probabilistic representation of the weak solution to the RM equation (3)\(^1\).

\(^1\)It can be seen that if $(b, \sigma)$ are sufficiently smooth, then $\mathcal{A}^* f(x) = \frac{1}{2} \sum_{i,j=1}^n \partial_{x_i}^2 \left( (\sigma \sigma^\top)_{ij}(x) f(x) \right) - \sum_{i=1}^n \partial_{x_i} (b_i(x) f(x))$ for $x \in D$. 

3
in the above distributional sense. We provide sufficient conditions satisfied by the fitness function \( g \) under which the solution of (4) admits an explicit form which can be expressed in terms of the transition density function of some Itô diffusion processes. The main strategy for achieving this aim is to provide a probabilistic representation of the solution of (3) by using the theory of FPK equations and a martingale extraction approach.

The second objective of this paper is to establish the propagation of chaos of our RM equation (3) for a relatively large class of fitness functions \( g \). It is in fact equivalent to the convergence of the empirical measure of a (homogeneous) particle system to an arbitrary solution to the FPK equation related to (4) with respect to a Wasserstein-like distance adapted to our probabilistic framework. We also establish a rate of convergence in the propagation of chaos. This in particular implies uniqueness in the class of all weak solutions of our RM equation (4).

### 3 FPK equation and examples

This section introduces a class of FPK equations associated with the RM equation (4). We establish a probabilistic representation of the solution of the FPK equation and then the probabilistic solution of (4) follows by assuming the absolute continuity of the initial date of the FPK equation w.r.t. Lebesgue measure. Finally, we provide examples with closed-form probabilistic solutions for different fitness functions considered in the existing literature.

#### 3.1 Probabilistic solution of FPK equation

Denote by \( \mathcal{P}(D) \) the set of probability measures on \( \mathcal{B}_D \) (i.e., the \( \sigma \)-algebra generated by the open subsets of \( D \)). For \( p \geq 1 \), let \( \mathcal{P}_p(D) \subset \mathcal{P}(D) \) be the set of probability measures on \( \mathcal{B}_D \) with finite \( p \)-order moment. We introduce the FPK equation associated with (4), which is given by, for \( f \in C_0^\infty(D) \),

\[
\langle \mu_t, f \rangle = \langle \rho_0, f \rangle + \int_0^t \langle \mu_s, (A + g)f \rangle \, ds - \int_0^t \langle \mu_s, f \rangle \, \langle \mu_s, g \rangle \, ds, \quad t \in [0, T],
\]

where the initial datum \( \rho_0 \in \mathcal{P}(D) \) and the integral \( \langle \mu_t, f \rangle := \int_D f \, d\mu_t \) for \( t \in [0, T] \).

**Remark 3.1.** If the initial datum \( \rho_0 \) of the FPK equation (5) admits a density function given by \( u_0(x) \) for \( x \in D \), namely, \( \rho_0(dx) = u_0(x) \, dx \), then, for any \( t \in [0, T] \), the solution \( \mu_t(dx) = u(t, x) \, dx \), where \( u(t, x) \) for \( (t, x) \in [0, T] \times D \) is the (weak) solution of the RM equation (4).

The following lemma provides a probabilistic representation of the \( \mathcal{P}(D) \)-valued solution \( \mu = (\mu_t)_{t \in [0, T]} \) for the FPK equation.

**Lemma 3.2.** Let assumptions \( (A_{b,\sigma}) \) and \( (A_X) \) hold. Assume that the fitness function \( g : D \to \mathbb{R} \) satisfies that

\[(A_g) \quad \text{For any} \ (t, x) \in [0, T] \times D, \ \mathbb{E}[\exp(\int_0^t g(X_s^x) \, ds)] < +\infty. \ \text{Moreover, for} \ t \in [0, T], \ x \to \mathbb{E}[\exp(\int_0^t g(X_s^x) \, ds)] \ \text{is in} \ L^1(D; \rho_0), \ \text{where} \ X_t^x := X_t^{0,x} \ \text{for} \ t \in [0, T].
\]

Let us define that, for \( t \in [0, T] \),

\[
\mu_t = \frac{\int_D \mathbb{E}\left[ \delta_{X_t^x} \exp\left(\int_0^t g(X_s^x) \, ds\right)\right] \rho_0(dy)}{\int_D \mathbb{E}\left[ \exp\left(\int_0^t g(X_s^x) \, ds\right)\right] \rho_0(dy)}, \quad \text{on} \ \mathcal{B}_D,
\]

where \( \delta \) denotes Dirac-delta measure. Then, \( \mu = (\mu_t)_{t \in [0, T]} \) defined by (6) is a \( \mathcal{P}(D) \)-valued solution of the FPK equation (5).
Proof. It is obvious to see that \( \mu_t \) defined by (6) is a probability measure on \( \mathcal{B}_D \) for any \( t \in [0, T] \), where \( \mu_0 = \rho_0 \). We define that

\[
    h_t := \int_D \mathbb{E} \left[ \exp \left( \int_0^t g(X^y_s) \, ds \right) \right] \rho_0(dy), \quad t \in [0, T],
\]

which is well-defined due to the assumption \((A_y)\). By virtue of the representation (6), we have that, for \( s \in [0, T] \),

\[
    h_s(\mu_s, g) = \int_D \mathbb{E} \left[ g(X^y_s) \exp \left( \int_0^s g(X^y_r) \, dr \right) \right] \rho_0(dy).
\]

Integrate on both sides of the above equality w.r.t. \( s \) from 0 to \( t \in (0, T) \), it follows from the Fubini’s theorem that

\[
    \int_0^t h_s(\mu_s, g) \, ds = \int_D \mathbb{E} \left[ \int_0^t g(X^y_s) \exp \left( \int_0^s g(X^y_r) \, dr \right) \, ds \right] \rho_0(dy)
    = \int_D \mathbb{E} \left[ \exp \left( \int_0^t g(X^y_s) \, ds \right) \right] \rho_0(dy) - 1 = h_t - 1.
\]

This yields the ODE given by \( h'_t = h_t(\mu_t, g) \) with \( h_0 = 1 \). Then, it holds that, for \( t \in [0, T] \),

\[
    \exp \left( \int_0^t \langle \mu_s, g \rangle \, ds \right) = \int_D \mathbb{E} \left[ \exp \left( \int_0^t g(X^y_s) \, ds \right) \right] \rho_0(dy). \tag{7}
\]

It follows from (6) and (7) that

\[
    \mu_t = \int_{\mathbb{R}} \mathbb{E} \left[ \delta_{X^y_t} \exp \left( \int_0^t (g(X^y_s) - \langle \mu_s, g \rangle) \, ds \right) \right] \rho_0(dy), \quad \text{on } \mathcal{B}_D. \tag{8}
\]

Then, for any test function \( f \in C^\infty_c(D) \), by applying Itô formula to \( f(X^y_t) \), which yields that

\[
    \mathbb{E} \left[ f(X^y_t) \exp \left( \int_0^t (g(X^y_s) - \langle \mu_s, g \rangle) \, ds \right) \right] = f(y)
    \quad + \int_0^t \mathbb{E} \left[ \exp \left( \int_0^s (g(X^y_r) - \langle \mu_r, g \rangle) \, dr \right) A f(X^y_t) \right] \, ds
    \quad + \int_0^t \mathbb{E} \left[ \exp \left( \int_0^s (g(X^y_r) - \langle \mu_r, g \rangle) \, dr \right) f(X^y_t)(g(X^y_t) - \langle \mu_s, g \rangle) \right] \, ds.
\]

Integrate on both sides of the above equality w.r.t. \( \rho_0(dy) \). Then, by virtue of (8), we arrive at

\[
    \langle \mu_t, f \rangle = \langle \rho_0, f \rangle + \int_0^t \langle \mu_s, Af \rangle \, ds + \int_0^t \langle \mu_s, gf - \langle \mu_s, g \rangle f \rangle \, ds.
\]

This yields the FPK equation (5). \( \square \)

### 3.2 Closed-form representation of \( \mu \) given by (6)

We study a closed-form representation of the \( \mathcal{P}(D) \)-valued function \( \mu \) given by (6) in Lemma 3.2 by applying the martingale approach. We find that the obtained explicit form of \( \mu \) can be in fact represented in terms of the transition density function of a class of Itô diffusion processes. Moreover, we provide examples in which the expressions of \( \mu \) and its density admit complete closed-form.
By (6), the representation of \( \mu \) can be reduced to identify the following function: for \( f \in C_0(D) \), i.e., the space of bounded continuous functions on \( D \), define

\[
\Phi^f(t, x) := \mathbb{E} \left[ f(X^t_x) \exp \left( \int_0^t g(X^s_x) \, ds \right) \right], \quad (t, x) \in [0, T] \times D, \tag{9}
\]

where the process \( X^x = (X^x_t)_{t \in [0, T]} \) is the unique strong solution of SDE (2). We next apply a martingale approach to give a closed-form representation of \( \mu \) when the coefficients of equation (b, \( \sigma, g \)) satisfy an additional constraint.

**Lemma 3.3.** Let assumptions \((A_{b, \sigma}), (A_X)\) and \((A_g)\) hold. Assume additionally that:

(i) \( g \in C^2(D) \) and satisfies that, for all \( x \in D, \)

\[
A_g(x) = C_1 \in \mathbb{R}, \quad \text{and} \quad \nabla_x g(x)^\top \sigma(x) = C_2 \in \mathbb{R}^{1 \times m}, \tag{10}
\]

where the operator \( A \) acted on \( C^2(D) \) is given by (1).

(ii) for \( y \in D \), let \( \bar{X}^y_t = (\bar{X}^y_t)_{t \in [0, T]} \) be the strong solution of the following SDE:

\[
\bar{X}^y_t = y + \int_0^t \bar{b}(s, \bar{X}^y_s) \, ds + \int_0^t \sigma(\bar{X}^y_s) \, dW_s, \tag{11}
\]

where the time-dependent function \( \bar{b}(t, x) := b(x) - t\sigma(x)C_2^\top \) for \( (t, x) \in [0, T] \times D \).

Then, for any \( T > 0 \), the solution \( \mu \) given by (6) of the FPK equation (5) admits the following form:

\[
\mu_t(dx) = \frac{e^{tg(x)}}{\int_D e^{tg(z)} \mathbb{P}(\bar{X}^y_t \in dz)} \rho_0(dy), \quad t \in [0, T]. \tag{12}
\]

Moreover, if \( \rho_0(dx) = u_0(x)dx \) and the density function \( \bar{p}(t, y; x) \) of \( \bar{X}^y_t \) for \( t \in [0, T] \)

exists, namely

\[
\mathbb{P}(\bar{X}^y_t \in dx) = \bar{p}(t, y; x)dx,
\]

then, for all \( t \in [0, T] \), the probability measure \( \mu_t(dx) = u(t, x)dx \). Here \( u(t, x) \) satisfies the extended RM equation (4), and it has the following representation given by

\[
u(t, x) = \frac{e^{tg(x)}}{\int_D e^{tg(z)} (\int_D \mathbb{P}(t, y; z)u_0(y)dy) \, dz}, \quad (t, x) \in [0, T] \times D. \tag{13}
\]

**Proof.** Integration by parts yields that, for \( (t, y) \in [0, T] \times D, \)

\[
\begin{aligned}
\int_0^t g(X^y_s) \, ds &= t g(X^y_t) - \int_0^t s A_g(X^y_s) \, ds - \int_0^t s \nabla_x g(X^y_s)^\top \sigma(X^y_s) \, dW_s \\
&= t g(X^y_t) + \int_0^t \left\{ \frac{1}{2} s^2 \nabla_x g(X^y_s)^\top (\sigma \sigma^\top) (X^y_s) \nabla_x g(X^y_s) - s A_g(X^y_s) \right\} \, ds \\
&\quad - \int_0^t s \nabla_x g(X^y_s)^\top \sigma(X^y_s) \, dW_s - \frac{1}{2} \int_0^t s^2 \nabla_x g(X^y_s)^\top (\sigma \sigma^\top) (X^y_s) \nabla_x g(X^y_s) \, ds.
\end{aligned}
\]

Note that, by (10) in the condition (i), we have that \( A_g(x) \equiv C_1 \in \mathbb{R} \) and \( \nabla_x g(x)^\top \sigma(x) = C_2 \in \mathbb{R}^{1 \times m} \). Then, by verifying the Novikov’s condition, we can define a \((\mathbb{P}, \mathbb{F})\)-martingale as follows:

\[
\mathcal{N}_t := \exp \left( -\int_0^t sC_2 \, dW_s - \frac{C_2 C_2^\top}{6} t^3 \right), \quad t \in [0, T].
\]


Therefore, by virtue of (9), for \((t, y) \in [0, T] \times D,\)
\[
\Phi^f(t, y) = \mathbb{E} \left[ N_t f(X^y_t) \exp \left( t g(X^y_t) + \frac{C_2 C_2^\top}{6} t^3 - \frac{C_1}{2} t^2 \right) \right].
\]
(14)

Let us define \(\frac{dQ}{dP}\big|_{F_t} = N_t\), for \(t \in [0, T]\). Then \(B_t := W_t + C_2^\top t\) for \(t \in [0, T]\) is an \(m\)-dimensional Brownian motion under the probability measure \(Q\). Moreover, under the new probability measure \(Q\), the dynamics of the process \(X^y(t) = (X^y_t)_{t \in [0, T]}\) is given by, for \(y \in D,\)
\[
X^y_t = y + \int_0^t (b(X^y_s) - s \sigma(X^y_s) C_2^\top) ds + \int_0^t \sigma(X^y_s) dB_s.
\]
By applying (14), we arrive at
\[
\Phi^f(t, y) = \exp \left( \frac{C_2 C_2^\top}{6} t^3 - \frac{C_1}{2} t^2 \right) \mathbb{E}^Q \left[ f(X^y_t) e^{t g(X^y_t)} \right]
= \exp \left( \frac{C_2 C_2^\top}{6} t^3 - \frac{C_1}{2} t^2 \right) \int_D f(x) e^{tg(x)} Q(X^y_t \in dx).
\]
Then, the solution representation (6) yields that, for \(t \in [0, T],\)
\[
\int_D f(x) \mu_t(dx) = \frac{\int_D f(x) e^{tg(x)} \int_D Q(X^y_t \in dx) \rho_0(dy)}{\int_D e^{tg(x)} \int_D Q(X^y_t \in dx) \rho_0(dy)}. \tag{15}
\]
Note that \(Q(X^y_t \in dx) = \mathbb{P}(\overline{X^y_t} \in dx)\). Therefore, the desired results (12) and (13) follows from the equality (15).

Lemma 3.3 works well for the case where the drift and volatility functions \((b, \sigma)\) are constant matrix and the fitness function \(g\) is a linear mapping. This is documented in the following example:

Examples 3.4. Consider \(m = n, b(x) \equiv b \in \mathbb{R}^n\) and \(\sigma(x) \equiv \sigma \in \mathbb{R}^{n \times n}\). Let \(\sigma \in \mathbb{R}^{n \times n}\) be invertible and its invertible matrix is given by \(\sigma^{-1}\). In this example, SDE (2) is reduced to a drift-Brownian motion described as:
\[
X^x_t = x + bt + \sigma W_t, \quad t \in [0, T].
\]
Therefore, the domain \(D = \mathbb{R}^n\). For any \(C_2 \in \mathbb{R}^{1 \times n}\), we consider the following fitness function given by
\[
g(x) = C_2 \sigma^{-1} x, \quad x \in \mathbb{R}^{n \times 1}.
\]
(16)
Then, a direct calculation yields that, for all \(t \in [0, T],\)
\[
\mathbb{E} \left[ \exp \left( \int_0^t g(X^x_s) ds \right) \right] = e^{tc_2 \sigma^{-1} x + \frac{t^2}{2} C_2 \sigma^{-1} b} \mathbb{E} \left[ \exp \left( t C_2 W_1 - C_2 \int_0^t s dW_s \right) \right]
= \exp \left( tc_2 \sigma^{-1} x + \frac{t^2}{2} C_2 \sigma^{-1} b + \frac{\sqrt{3}}{6} t^3 \sqrt{C_2 C_2^\top} \right).
\]
If the initial datum \(\rho_0\) of the FPK equation (5) satisfies \(\int_D e^{tc_2 \sigma^{-1} x} \rho_0(dx) < \infty\), then the assumption \((A_g)\) holds. Thus, we have from (16) that
\[
A_g \equiv C_1 := C_2 \sigma^{-1} b \in \mathbb{R}.
\]
Using the condition (ii) of Lemma 3.3, the process $\mathbf{X}^y_t = (X^y_t)_{t \in [0,T]}$ is given by

$$
\mathbf{X}^y_t = y + bt - \frac{\sigma C_2^\top}{2} t^2 + \sigma W_t, \quad (t, y) \in [0, T] \times \mathbb{R}^n.
$$

Hence, under the original probability measure $\mathbb{P}$, we have that

$$
\mathbf{X}^y_t \sim N \left( y + bt - \frac{\sigma C_2^\top}{2} t^2, \sigma \sigma^\top t \right), \quad t \in (0, T].
$$

This implies that the density function $\overline{p}(t, y; x)$ of $\mathbf{X}^y_t$ has the following closed-form representation given by $\overline{p}(0, y; x) = \delta_{x-y}$, and for $(t, x, y) \in (0, T] \times \mathbb{R}^2$,

$$
\overline{p}(t, y; x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\sigma \sigma^\top) t}} \exp \left\{ - \frac{1}{2 t} \left[ x - y - bt + \frac{\sigma C_2^\top}{2} t^2 \right]^\top \left( \sigma \sigma^\top \right)^{-1} \left[ x - y - bt + \frac{\sigma C_2^\top}{2} t^2 \right] \right\}. \tag{17}
$$

It follows from (13) in Lemma 3.3 that the density function $\mu_t$ is given by

$$
u(t, x) = \frac{e^{C_2 \sigma^{-1} x t} \int_{\mathbb{R}^n} \overline{p}(t, y; x) u_0(y) dy}{\int_{\mathbb{R}^n} e^{C_2 \sigma^{-1} x t} \left( \int_{\mathbb{R}^n} \overline{p}(t, y; x) u_0(y) dy \right) dz}, \quad (t, x) \in [0, T] \times \mathbb{R}^n. \tag{18}
$$

This in fact establishes the probabilistic representation of the weak solution to the following RM equation with the fitness space $D = \mathbb{R}^n$:

$$
\begin{align*}
\partial_t u(t, x) &= \frac{\sigma \sigma^\top}{2} \Delta u(t, x) - b^\top \nabla_x u(t, x) + \left( g(x) - \int_{\mathbb{R}^n} g(y) u(t, y) dy \right) u(t, x); \\
u(0, x) &= u_0(x), \quad x \in \mathbb{R}^n.
\end{align*} \tag{19}
$$

A special case of Example 3.4 above is the one-dimensional RM equation with $b \equiv 0$, $\sigma = \sqrt{2}$ and the fitness function $g(x) = x$ for $x \in \mathbb{R}$. Then, Eq. (19) becomes that

$$
\begin{align*}
\partial_t u(t, x) &= \Delta u(t, x) + \left( x - \int_{\mathbb{R}} y u(t, y) dy \right) u(t, x); \\
u(0, x) &= u_0(x), \quad x \in \mathbb{R}.
\end{align*} \tag{20}
$$

The explicit solution to the RM equation (20) has been studied by Alfaro and Carles (2014). We next verify that the probabilistic solution (18) of the RM equation (20) coincides with the solution form given by (2.3) of Theorem 2.3 in Alfaro and Carles (2014) via a tricky transform of the solution based on the Avron-Herbst formula. In fact, in view of Lemma 3.3, from (16) and (17), it results in $C_2 = \sqrt{2}$, and hence

$$
\overline{p}(t, y; x) = \frac{1}{2\sqrt{\pi} t} \exp \left( - \frac{(x - y + t^2)^2}{4 t} \right), \quad (t, x, y) \in [0, T] \times \mathbb{R}^2. \tag{21}
$$

This yields that the probability solution $u(t)$ of the RM equation (20) is given by, for $(t, x, y) \in [0, T] \times \mathbb{R}^2$,

$$
\begin{align*}
u(t, x) &= \frac{e^{tx} \int_{\mathbb{R}} \exp \left( \frac{(x - y + t^2)^2}{4 t} \right) u_0(y) dy}{\int_{\mathbb{R}} e^{tz} \left( \int_{\mathbb{R}} \exp \left( \frac{(z - y + t^2)^2}{4 t} \right) u_0(y) dy \right) dz}.
\end{align*} \tag{22}
$$
By (6) and the assumption \((A_g)\), it is not difficult to verify that, for \(t \in [0, T)\),

\[
\int_{\mathbb{R}} xu(t, x)dx = \int_{\mathbb{R}} x\mu_t(dx) = \frac{\int_D \mathbb{E} \left[ X^y_t \exp \left( \int_0^t g(X^y_s)ds \right) \right] \rho_0(dy)}{\int_D \mathbb{E} \left[ \exp \left( \int_0^t g(X^y_s)ds \right) \right] \rho_0(dy)} < \infty.
\]

A direct calculation gives that

\[
\int_{\mathbb{R}} e^{tz} \left( \int_{\mathbb{R}} \exp \left( -\frac{(z - y + t^2)^2}{4t} \right) u_0(y)dy \right) dz = \int_{\mathbb{R}^2} \exp \left( -\frac{(z - y + t^2)^2}{4t} + ty \right) u_0(y)dydz = \sqrt{4\pi t} \int_{\mathbb{R}} e^{ty} u_0(y)dy.
\]

Therefore, for \((t, x) \in [0, T) \times \mathbb{R}\),

\[
u(t, x) = \frac{e^{tx} \int_{\mathbb{R}} \exp \left( -\frac{(x-y+t^2)^2}{4t} \right) u_0(y)dy}{\int_{\mathbb{R}} e^{tz} \left( \int_{\mathbb{R}} \exp \left( -\frac{(z - y + t^2)^2}{4t} \right) u_0(y)dy \right) dz} = \frac{e^{tx}}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{ty} u_0(y)dy \tag{23}
\]

this is the solution representation given by (2.3) of Theorem 2.3 in Alfaro and Carles (2014).

However, the condition (10) in Lemma 3.3 on the coefficients \((b, \sigma, g)\) is very restrictive. In order to bypass this constraint on \((b, \sigma, g)\), we next apply a martingale extraction approach to identify the function \(\Phi^f\) defined by (9). More precisely, recall the operator \(A\) acted on \(C^2(D)\) which is defined by (1). Let \((\lambda, \phi)\) be an eigenpair (if exists) of the following characteristic equation given by

\[
(A + g(x))\phi(x) = -\lambda\phi(x), \quad x \in D. \tag{24}
\]

Then, it follows from (24) and Itô formula that, for \(y \in D\),

\[
M^y_t := \frac{1}{\phi(y)} \exp \left( \lambda t + \int_0^t g(X^y_s)ds \right) \phi(X^y_t), \quad t \in [0, T] \tag{25}
\]

is a local \((\mathbb{P}, \mathbb{F})\)-martingale with \(M^y_0 = 1\). Then, by (9), it holds that

\[
\Phi^f(t, y) = e^{-\lambda t} \phi(y) \mathbb{E} \left[ M^y_t \frac{f(X^y_t)}{\phi(X^y_t)} \right], \quad (t, y) \in [0, T] \times D. \tag{26}
\]

A key observation is that the expectation in (26) does not depend on the path of the process \(X^y\). Then, this gives that

**Theorem 3.5.** Let assumptions \((A_{b,\sigma})\), \((A_X)\) and \((A_g)\) hold. Assume additionally that, for all \(y \in D\),

(i) there exists an eigenpair \((\lambda, \phi)\) of the characteristic equation (24) such that the process \(M^y = (M^y_t)_{t \in [0, T]}\) defined by (25) is a \((\mathbb{P}, \mathbb{F})\)-martingale.

(ii) let \(X^y = (X^y_t)_{t \in [0, T]}\) be the strong solution of the following SDE: for \(t \in [0, T]\),

\[
X^y_t = y + \int_0^t (b(X^y_s) + \sigma(X^y_s)\sigma(x)^\top)ds + \int_0^t \sigma(X^y_s)dW_s, \tag{27}
\]

where the function \(\sigma(x) := \frac{\nabla \phi(x)^\top}{\phi(x)}\sigma(x)\) for \(x \in D\).
Then, for any $T > 0$, the solution $\mu$ given by (6) of the FPK equation (5) admits the following form:

$$
\mu_t(dx) = \frac{1}{\phi(x)} \int_D \phi(y) \mathbb{P}(\mathbb{X}^y_t \in dx) \rho_0(dy), \quad t \in (0, T].
$$

(28)

Moreover, if $\rho_0(dx) = u_0(x)dx$ and the density function $\mathbb{P}(t, y; x)$ of $\mathbb{X}^y_t$ for $t \in (0, T]$ exists, in other words,

$$
\mathbb{P}(\mathbb{X}^y_t \in dx) = \mathbb{P}(t, y; x)dx,
$$

then, for all $t \in [0, T]$, the probability measure $\mu_t(dx) = u(t, x)dx$. Here, $u(t, x)$ satisfies the extended RM equation (4), and it admits the following representation given by

$$
u(t, x) = \frac{1}{\phi(x)} \int_D \phi(y) \mathbb{P}(t, y; x) u_0(y)dy, \quad (t, x) \in [0, T] \times D.
$$

(29)

Proof. By using (25) and (24), the dynamics of the martingale $M^y$ defined by (25) is given by, for $y \in D$,

$$
dM^y_t = M^y_t \nabla \phi(X^y)\sigma(X^y) + \frac{\nabla \phi(X^y)}{\phi(X^y)}dW_t = M^y_t \sigma(X^y) dW_t, \quad M^y_0 = 1.
$$

(30)

Define $\frac{dM_t}{d\mathbb{P}} |_{\mathbb{F}_t} = M^y_t$ for $t \in [0, T]$. Then, from Girsanov’s theorem, it follows that $B_t := W_t - \int_0^t \sigma(X^y) dW_t$ is an $m$-dimensional Brownian motion under the new probability measure $\mathbb{Q}$. Then, under $\mathbb{Q}$, the dynamics of the process $X^y = (X^y_t)_{t \in [0, T]}$ is given by, for $(t, y) \in [0, T] \times D$,

$$
X^y_t = y + \int_0^t (b(X^y_s) + \sigma(X^y_s)\sigma(X^y_s)) ds + \int_0^t \sigma(X^y_s) dB_s.
$$

By virtue of (26), we have that, for $(t, y) \in [0, T] \times D$,

$$
\Phi^J(t, y) = e^{-\lambda t} \mathbb{E}^\mathbb{Q} \left[ \frac{f(X^y_T)}{\phi(X^y_T)} \right] = e^{-\lambda t} \mathbb{E} \left[ \frac{f(x)}{\phi(x)} \right] \mathbb{Q}(X^y_T \in dx).
$$

Apply Lemma 3.2, we obtain that, for all test functions $f \in C^\infty_0(D)$,

$$
\int_D f(x) \mu(dx) = \frac{\int_D f(x) \mathbb{P}(X^y_T \in dx) \rho_0(dy)}{\int_D \mathbb{P}(X^y_T \in dx) \rho_0(dy)}. \quad (31)
$$

Note that $\mathbb{Q}(X^y_T \in dx) = \mathbb{P}(\mathbb{X}^y_T \in dx)$. Then, the desired result follows from (31). $
$

In order to apply Theorem 3.5, the key point is to identify an eigenpair $(\lambda, \phi)$ of the characteristic equation (24) so to satisfy the condition (i) of Theorem 3.5. We next provide examples in which we illustrate how to verify this condition (i) and derive the probability solution (28)-(29) explicitly. We first consider the one-dimensional case, i.e., $n = m = 1$. In this case, let $D = (\alpha, \beta)$ with $-\infty \leq \alpha < \beta \leq +\infty$. For this purpose, for a function $V : D \to \mathbb{R}$ which is locally Hölder continuous, we define the following operator acted on $C^2(D)$ as:

$$
\mathcal{H}_V := \mathcal{A} + V.
$$

(32)

We call a positive function $\phi \in C^2(D)$ a positive harmonic function of $\mathcal{H}_V$ if $\mathcal{H}_V \phi = 0$. A function $\phi \in C^2(D)$ is said to be an invariant function of the semigroup generated by $\mathcal{H}_V$ if $\mathbb{E}[e^{\int_0^t \mathcal{H}_V(X^y_s) ds} \phi(X^y_T)] = \phi(x)$ for $(t, x) \in [0, T] \times D$. The following result provides the necessary and sufficient condition under which a positive harmonic function of $\mathcal{H}_V$ on $D$ is also an invariant function of the semigroup generated by this operator.
Lemma 3.6 (Theorem 5.18. in Pinsky (1995)). Let the assumption \((A_{b,\sigma})\) and \((A_X)\) hold and \(x_0 \in D\). Then, a positive harmonic function of \(H_V\) is also an invariant function of the semigroup generated by \(H_V\) if and only if the following conditions hold:

\[
\int_{x_0}^{x_0} \left[ \frac{1}{\phi^2(x)} \exp \left( -\int_{x_0}^{x_0} \frac{2b(z)}{\sigma^2(z)} dz \right) \int_{x_0}^{x_0} \frac{\phi^2(y)}{\sigma^2(y)} \exp \left( \int_{x_0}^{y} \frac{2b(z)}{\sigma^2(z)} dz \right) dy \right] dx = +\infty,
\]

\[
\int_{x_0}^{\beta} \left[ \frac{1}{\phi^2(x)} \exp \left( -\int_{x_0}^{\beta} \frac{2b(z)}{\sigma^2(z)} dz \right) \int_{x_0}^{\beta} \frac{\phi^2(y)}{\sigma^2(y)} \exp \left( \int_{x_0}^{y} \frac{2b(z)}{\sigma^2(z)} dz \right) dy \right] dx = +\infty. \tag{33}
\]

Lemma 3.6 can be used to verify the martingale property of \(M^y = (M^y_t)_{t \in [0,T]}\) defined by (25) in some one-dimensional cases. We next provide examples in the one-dimensional case by applying Lemma 3.6.

Examples 3.7. Consider \(m = n = 1\), \(b(x) = \kappa(\theta - x)\), and \(\sigma(x) = \sigma\) for \(x \in D := \mathbb{R}\), where \(\theta, \kappa \in \mathbb{R}\) with \(\kappa \neq 0\) and \(\sigma > 0\). Then, assumptions \((A_{b,\sigma})\) and \((A_X)\) are satisfied. In this case, the process \(X^{t,x} = (X^{t,x}_s)_{s \in [t,T]}\) given by (2) satisfies that, for all \((t,x) \in [0,T] \times \mathbb{R}\),

\[
dX^{t,x}_s = \kappa(\theta - X^{t,x}_s)ds + \sigma dW_s, \quad X^{t,x}_t = x \in \mathbb{R}. \tag{34}
\]

The fitness function is given by \(g(x) = -x\) for \(x \in \mathbb{R}\). We next compute \(\mathbb{E}[\exp(-\int_0^t X_s^0 ds)]\) for \((t,x) \in [0,T] \times \mathbb{R}\). In fact, we have that

\[
\int_0^t X_s^0 ds = \left(1 - e^{-\kappa t}\right)(x - \kappa^{-1}\theta) + \kappa \theta t + \sigma W_t - \sigma e^{-\kappa t} \int_0^t e^{\kappa s} dW_s.
\]

This yields that \(\int_0^t X_s^0 ds \sim N(\mu(t,x),\sigma^2(t))\) for \(t \in (0,T]\), where \(\mu(t,x) := (1 - e^{-\kappa t})(x - \kappa^{-1}\theta) + \kappa \theta t + \sigma^2(t) := \sigma^2 t + \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t}) - \frac{2\sigma^2}{\kappa}(1 - e^{-\kappa t})\). Therefore, for \(t \in [0,T]\),

\[
\mathbb{E}\left[\exp\left(-\int_0^t X_s^0 ds\right)\right] = e^{-\mu(t,x) + \frac{1}{2}\sigma^2(t)}. \tag{35}
\]

Moreover, it results in

\[
\int_{\mathbb{R}} \mathbb{E}\left[\exp\left(-\int_0^t X_s^0 ds\right)\right] \rho_0(dx) = e^{\frac{\sigma^2(t)}{2} + (1 - e^{-\kappa t})\kappa^{-1}\theta - \kappa \theta t} \int_{\mathbb{R}} e^{-(1 - e^{-\kappa t})x} \rho_0(dx).
\]

This implies that if the initial data \(\rho_0\) of the FPK equation (5) satisfies that

\[
\int_{\mathbb{R}} e^{-(1 - e^{-\kappa t})x} \rho_0(dx) < \infty, \tag{36}
\]

then the assumption \((A_g)\) is satisfied.

Define \(V(x) := g(x) + \lambda = \lambda - x\) for \(x \in \mathbb{R}\). Thus, the characterize equation (24) becomes \(H_V \phi = 0\). For \(\lambda \in \mathbb{R}\), let us define

\[
\mu(\lambda) := \begin{cases} 
\frac{1}{\kappa} \left(\lambda - \theta + \frac{\sigma^2}{2\kappa^2}\right), & \kappa > 0; \\
\frac{1}{|\kappa|} \left(\lambda - \theta + \frac{\sigma^2}{2\kappa^2} + \kappa\right), & \kappa < 0.
\end{cases} \tag{37}
\]

We take \(\lambda\) so that \(\mu(\lambda) \leq 0\). Then, together with Lemma 3.6, by applying Proposition 6.2 and Theorem 6.2 in Qin and Linetsky (2016), the associated semigroup generated by \(H_V\) (depending on \(\lambda\)) has invariant functions given by:
(i) for $\mu(\lambda) < 0$, the invariant functions are of the form:

\[
\phi_\lambda(x) = C_1 e^{\frac{z(x)^2}{4}} E_\mu(z(x) - x) + C_2 e^{\frac{z(x)^2}{4}} E_\mu(\alpha - x(x)), \quad C_1, C_2 > 0,
\]

where $\alpha := \sigma \sqrt{\frac{2}{|\kappa|^2}}, z(x) := \sqrt{\frac{2|\kappa|}{\sigma}}(\theta - x)$, and $\epsilon := \text{sign}(\kappa)$.

(ii) for $\mu(\lambda) = 0$, the invariant functions are of the form:

\[
\phi_\lambda(x) = C_1 e^{\frac{z(x)^2}{4}} E_\mu(z(x) - x), \quad C_1 > 0.
\]

Here, $E_\mu(\cdot)$ denotes the Weber parabolic cylinder function. Noting that the following recurrence relations on the Weber parabolic cylinder function hold:

\[
E_{\mu+1}(z) - zE_\mu(z) + \mu E_{\mu-1}(z) = 0, \quad E'_\mu(z) + \frac{1}{2}zE_\mu(z) - \mu E_{\mu-1}(z) = 0.
\]

Then, we arrive at the following relations by

\[
\begin{align*}
\frac{d}{dz} \left(e^{\frac{z^2}{4}} E'_\mu(z - \alpha)\right) &= \frac{\alpha}{2} e^{\frac{z^2}{4}} E_\mu(z - \alpha) + \mu e^{\frac{z^2}{4}} E_{\mu-1}(z - \alpha); \\
\frac{d}{dz} \left(e^{\frac{z^2}{4}} E_\mu(\alpha - z)\right) &= \frac{\alpha}{2} e^{\frac{z^2}{4}} E_\mu(\alpha - z) - \mu e^{\frac{z^2}{4}} E_{\mu-1}(z - \alpha); \\
\frac{d}{dz} \left(e^{-\frac{z^2}{4}} E_\mu(z - \alpha)\right) &= -\frac{\alpha}{2} e^{-\frac{z^2}{4}} E_\mu(z - \alpha) - e^{\frac{z^2}{4}} E_{\mu+1}(z - \alpha); \\
\frac{d}{dz} \left(e^{-\frac{z^2}{4}} E_\mu(\alpha - z)\right) &= -\frac{\alpha}{2} e^{-\frac{z^2}{4}} E_\mu(\alpha - z) + e^{\frac{z^2}{4}} E_{\mu+1}(\alpha - z).
\end{align*}
\]

Therefore, the dynamic of $X_t^y$ given in the condition (ii) of Theorem 3.5 are respectively as follows in terms of the sign of $\kappa$:

\[
\begin{align*}
dX_t^y &= b^+(X_t^y) + \sigma dW_t, \quad \kappa > 0; \\
&= b^+(X_t^y) + \sigma dW_t, \quad \kappa < 0.
\end{align*}
\]

The drift coefficient functions in the above SDEs are given by:

\[
\begin{align*}
b^+(x) &= \kappa \theta - \frac{\sigma^2}{\kappa} - \kappa x + \sqrt{2|\kappa|\sigma} \frac{C_2 E_{\mu-1}(z(x) - x)}{C_2 E_\mu(\alpha - z(x)) + C_1 E_\mu(\alpha - x(x))}; \\
b^-(x) &= \kappa \theta + \frac{\sigma^2}{\kappa} - \kappa x + \sqrt{2|\kappa|\sigma} \frac{C_1 E_{\mu+1}(z(x) - x)}{C_2 E_\mu(\alpha - z(x)) + C_1 E_\mu(\alpha - x(x))}.
\end{align*}
\]

Define $W_t^y := \sigma^{-1}y + W_t$, $N_t^{y,\pm} := \exp(\int_0^t \sigma^{-1}b^+\left(\sigma W_s^y\right) dW_s^y - \frac{1}{2\sigma^2} \int_0^t b^+ \left(\sigma W_s^y\right)^2 ds)$ and $Y_t^y := \sigma^{-1}X_t^y$ for $t \in [0, T]$. Using Exercise 5.5.38 in Karatzas and Shreve (1991), it follows that, for all $t \in (0, T]$ and $z \in \mathbb{R}$,

\[
P(Y_t^y \leq z) = \mathbb{E} \left[N_t^{y,\text{sgn}(k)} \mathbb{1}_{W_t^y \leq z}\right] = \int_{-\infty}^{z} \mathbb{E} \left[N_t^{y,\text{sgn}(k)} \middle| W_t^y = r\right] p_0 \left(t, \frac{y}{\sigma} ; r\right) dr,
\]

where $p_0(t, y; x) := \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right)$ for $(t, x, y) \in (0, T] \times \mathbb{R}^2$. Therefore, for all $(t, x, y) \in (0, T] \times \mathbb{R}^2$,

\[
P(X_t^y \leq x) = P \left(Y_t^y \leq \frac{x}{\sigma}\right) = \int_{-\infty}^{\frac{x}{\sigma}} \Phi_{\text{sgn}(k)}(t, y) p_0 \left(t, \frac{y}{\sigma} ; r\right) dr.
\]
Here $\Phi^\pm(t,r,y) := \mathbb{E}[N_{y,t}^{b,\pm}|W_t = r]$ for all $(t,r,y) \in (0,T) \times \mathbb{R}^2$. Then, the density function $p(t,y;x)$ of $X_t^y$ for $t \in (0,T)$ is given by, for $(t,x,y) \in (0,T) \times \mathbb{R}^2$,

$$
p(t,y;x) = \frac{1}{\sigma} q_{\text{exp}}(\alpha, \beta) \frac{x}{\sqrt{y}} \mathbf{1}_{[0,T]}(t) p_0 \left( t, \frac{y}{\sigma} \right).
$$

(41)

Then, by the solution representation (29) in Theorem 3.5, and (36), for any probability density function $u_0$ satisfying $\int_{\mathbb{R}} e^{-\sigma - u_0(x)} dx < \infty$, we have the solution of the extended RM equation (4) given by

$$
u(t,x) = \frac{1}{\sigma} \int_{\mathbb{R}} \phi(\sigma, y) \mathbb{E}(t, y) u_0(y) dy\int_{\mathbb{R}} \mathbb{E}(t, y) dy dy,
$$

(42)

Examples 3.8. Consider $m = n = 1$, $b(x) = a + bx$, and $\sigma(x) = \sigma \sqrt{x}$ for $x \in D := (0, \infty)$, where $a, \sigma > 0$ and $b \in \mathbb{R}$. Let the Feller condition hold, i.e., $2a \geq \sigma^2$. Then, assumptions $(A_{b,\sigma})$ and $(A_X)$ are satisfied. In this case, the process $X^{t,x} = (X^{t,x}_s)_{s \in [t,T]}$ given by (2) satisfies that, for all $(t,x) \in (0,T) \times D$,

$$
dX^{t,x}_s = (a + bX^{t,x}_s) ds + \sigma \sqrt{X^{t,x}_s} dW_s,
$$

(43)

and for any probability density function $u_0(x)$ on $x \in (0, \infty)$, we have the solution of the extended RM equation (4) given by

$$
u(t,x) = \frac{1}{\sigma} \int_{\mathbb{R}} \phi(\sigma, y) \mathbb{E}(t, y) u_0(y) dy\int_{\mathbb{R}} \mathbb{E}(t, y) dy dy,
$$

(44)

Examples 3.9. Consider $m = n = 1$, $b(x) = 0$ and $\sigma(x) = \sqrt{2} \sigma$ for $x \in D = \mathbb{R}$, where $\sigma > 0$. Then, assumptions $(A_{b,\sigma})$ and $(A_X)$ are satisfied. In this case, the process $X^{t,x} = (X^{t,x}_s)_{s \in [t,T]}$ given by (2) satisfies that, for all $(t,x) \in (0,T) \times \mathbb{R}$,

$$
X^{t,x}_s = y + \sqrt{2} \sigma (W_s - W_t),
$$

(45)

Here, we consider the polynomial confining fitness function of degree $2q$ which is introduced in Alfaro and Verdua (2019), i.e., for a given $q \in \mathbb{N},$

$$
g(x) = -x^{2q} + \sum_{i=0}^{2q-1} a_i x^i, \quad x \in \mathbb{R},
$$

(46)
where the sequence of constants $\alpha_l$ with $l = 0, \ldots, 2q - 1$. Then, the resulting equation (3) becomes the integro-differential Cauchy problem studied in Alfaro and Veruete (2019). Note that the operator $A$ acted on $C^2(\mathbb{R})$ is given by

$$Af(x) = \sigma^2 f''(x), \quad x \in \mathbb{R}.$$  

Then, the eigenpair $(\lambda, \phi)$ of the characteristic equation (24) is related to the spectral basis of the Schrödinger operator $-\mathcal{A} - g$ with confining potential $-g.$ By virtue of (46), it can be seen that $g : \mathbb{R} \to \mathbb{R}$ is continuous and $\lim_{|x| \to \infty} g(x) = -\infty.$ Then, $-\mathcal{A} - g$ has discrete spectrum, i.e., there is an orthonormal basis $(\phi_k)_{k \geq 1} \subset L^2(\mathbb{R})$ such that $(\mathcal{A} + g)\phi_k = -\lambda_k \phi_k$ and $\|\phi_k\|_{L^2(\mathbb{R})} = 1,$ with corresponding eigenvalues

$$\lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k \to \infty$$

of finite multiplicity. Moreover, Proposition 2.6 in Alfaro and Veruete (2019) gives that

$$\|\phi_k\|_{L^\infty(\mathbb{R})} \leq C k^{-\frac{q}{2q+1}}, \quad k \geq 1.$$  

Up to subtracting a constant to $g,$ we can assume without loss of generality that $g \leq 0.$ This yields the assumption $(\mathbf{A}_g).$ Recall the local martingale $M^y$ defined by (25) with the eigenpair $(\lambda_k, \phi_k)$ for $k \geq 1.$ Then, it holds that, $M^0_t = 1$ and for $t \in (0, T],

$$dM_{t}^{k,y} = \frac{\sqrt{2}\sigma}{\phi_k(y)} \exp \left( \lambda t + \int_0^t g(X^y_s)ds \right) \phi'_k(X^y_t) dW_t.$$  

Note that, the normalized eigenfunction $\phi_k$ for $k \geq 1$ satisfies that

$$\sigma^2 \int_{\mathbb{R}} |\phi'_k(x)|^2 dx - \int_{\mathbb{R}} g(x)|\phi_k(x)|^2 dx = \lambda_k,$$

this yields that $\phi'_k(\pm \infty) = 0.$ Therefore $\phi'_k$ for $k \geq 1$ is bounded on $\mathbb{R}.$ This implies that the local martingale $M^{k,y}$ defined by (25) with the eigenpair $(\lambda_k, \phi_k)$ is a $(\mathcal{F}_t)$-martingale. In other words, the condition (i) of Theorem 3.5 holds.

For $k \geq 1$ and $y \in \mathbb{R},$ consider the following SDE given by

$$\overline{Y}^k_t, y = \frac{y}{\sqrt{2}\sigma} + \int_0^t \overline{\sigma}_k(\sqrt{2}\sigma \overline{Y}^k_s, y) ds + \overline{W}_t, \quad t \in [0, T],$$

where $\overline{\sigma}_k(x) := \sqrt{2}\sigma \phi'_k(x)/\phi_k(x)$ for $x \in \mathbb{R}.$ Then, the process $\overline{Y}^{k,y} = (\overline{Y}^k_t, y)_{t \in [0, T]}$ given in Theorem 3.5-(ii) is given by

$$\overline{Y}^{k,y}_t = \sqrt{2}\sigma Y^k_t, \quad t \in [0, T].$$

By applying Exercise 5.5.38 in Karatzas and Shreve (1991), we arrive at, for all $t \in (0, T]$ and $z \in \mathbb{R},$

$$\mathbb{P}(\overline{Y}^k_t, y \leq z) = \mathbb{E} \left[ N^k_t, y \mathbb{1}_{W^y_t \leq z} \right] = \int_{-\infty}^z \mathbb{E} \left[ N^k_t, y | W^y_t = r \right] p_0 \left( t, \frac{y}{\sqrt{2}\sigma}; r \right) dr,$$

where $W^y_t = \frac{y}{\sqrt{2}\sigma} + \overline{W}_t$ and $N^k_t, y := \exp\left( \int_0^t \overline{\sigma}_k(\sqrt{2}\sigma W^y_s) dW^y_s - \frac{1}{2} \int_0^t \overline{\sigma}_k^2(\sqrt{2}\sigma W^y_s) ds \right)$ for $(t, y) \in [0, T] \times \mathbb{R}.$ Therefore, for all $(t, x, y) \in (0, T] \times \mathbb{R}^2,$

$$\mathbb{P}(\overline{Y}^{k,y}_t \leq x) = \mathbb{P} \left( \overline{Y}^{k,y}_t \leq \frac{x}{\sqrt{2}\sigma} \right) = \int_{-\infty}^{x/\sqrt{2}\sigma} \Phi_k(t, r, y) p_0 \left( t, \frac{y}{\sqrt{2}\sigma}; r \right) dr.$$
Here \( \Phi_k(t, r, y) := \mathbb{E}[N^k_{t, y} | W_t^y = r] \) for all \((t, r, y) \in (0, T] \times \mathbb{R}^2 \). Then, the density function \( \overline{p}_k(t, y; x) \) of \( X_t^y \) for \( t \in (0, T] \) exists, and it is given by, for \((t, x, y) \in (0, T] \times \mathbb{R}^2 \),

\[
\overline{p}_k(t, y; x) = \frac{1}{\sqrt{2\sigma}} \Phi_k \left( t, \frac{x}{\sqrt{2\sigma}}, y \right) p_0 \left( t, \frac{y}{\sqrt{2\sigma}} \right),
\]

(49)

Then, for \( k \geq 1 \), the function defined below

\[
u_k(t, x) = \frac{1}{\phi_k(y)} \int_{\mathbb{R}} \phi_k(y) \overline{p}_k(t, y; x) u_0(y) dy, \quad (t, x) \in [0, T] \times \mathbb{R}
\]

(50)
is a (weak) solution of of the extended RM equation (4). If the uniqueness of solutions of the extended RM equation (4) holds, then \( \nu_k = u_l \) for \( k \neq l \). The uniqueness of solutions of the extended RM equation (4) will be implied by the propagation of chaos which will be established in Section 4.

The example with multi-dimensional fitness space and quadratic fitness function is provided below:

**Examples 3.10.** Consider \( m = n \geq 1 \), \( b(x) = b + Bx \), and \( \sigma(x) \equiv \sigma \) for \( x \in D := \mathbb{R}^n \),

where \( b \in \mathbb{R}^n, B \in \mathbb{R}^{n \times n} \) and \( \sigma \in \mathbb{R}^{n \times n} \) which is an invertible matrix, so that \( a := \sigma \sigma^\top \) is strictly positive definite. The fitness function is given by \( g(x) = -r(x) \) for \( x \in \mathbb{R}^n \), and

\[
r(x) = \alpha + \delta^\top x + x^\top Gx, \quad x \in \mathbb{R}^n,
\]

(51)

where \( \alpha \in \mathbb{R}, \delta \in \mathbb{R}^n \), and \( G \in \mathbb{R}^{n \times n} \) is a symmetric positive semi-definite matrix. They are taken to be such that \( r(x) \) is non-negative for all \( x \in \mathbb{R}^n \). Then, the assumption (A_2) holds. Let \( V = g(x) + \lambda = \lambda - r(x) \) for \( x \in \mathbb{R}^n \), then the characteristic equation (24) becomes \( \mathcal{H}_V \phi = 0 \). Let \( H \in \mathbb{R}^{n \times n} \) satisfies the following continuous-time algebraic Riccati equation:

\[
2HaH - B^\top H - HB - G = 0,
\]

(52)

and \( v \in \mathbb{R}^n \) satisfies the following linear equation:

\[
2HaH - B^\top v - 2Hb - \delta = 0.
\]

(53)

By Section 6.2.2 of Qin and Linetsky (2016), the eigenpair \((\lambda, \phi)\) of the characteristic equation (24) which satisfies the condition (i) of Theorem 3.5 is given by

\[
(\lambda, \phi(x)) = \left( \alpha - \frac{1}{2} v^\top av + \text{tr}(aV) + v^\top b, \exp \left( -v^\top x - x^\top Hx \right) \right), \quad x \in \mathbb{R}^n,
\]

(54)

where \( \text{tr} \) denotes the trance operator. Define \( \overline{\sigma}(x) = -(v + 2Hx)^\top \sigma \) for \( x \in \mathbb{R}^n \). Therefore, the process \( X^y = (X^y_s)_{t \in [0, T]} \) given in the condition (ii) of Theorem 3.5 satisfies that, for \( y \in \mathbb{R}^n \),

\[
X^y_t = y + \int_0^t (b + BX^y_s + \sigma \sigma(X^y_s)^\top) ds + \sigma W_t.
\]

(55)

Note that the process \( X^y \) is an n-dimensional affine process and hence the transition density function \( \overline{p}(t, y; x) \) of \( X^y \) exists and the characteristic function of \( \overline{p}(t, y; \cdot) \) has exponential-affine dependence on \( y \in \mathbb{R}^n \), see Theorem 2.7 of Duffie, et al. (2003). In fact, we can rewrite the process \( X^y \) described as (55) as in the form given by, for \( y \in \mathbb{R}^n \),

\[
X^y_t = e^{\Gamma y} + \int_0^t e^{\Gamma(t-s)} \beta ds + \int_0^t e^{\Gamma(t-s)} \sigma W_s, \quad t \in [0, T],
\]
where $\beta := b - \sigma \sigma^\top v$ and $\Gamma = B - 2\sigma \sigma^\top H$. This yields that, for $t \in (0, T]$, 
\[ \overline{X}_t^y \sim N(\mu_t, \Sigma_t), \]
where $\mu_t := e^{t \Gamma} y + \int_0^t e^{(t-s) \beta} ds$ and $\Sigma_t := \int_0^t e^{(t-s) \beta} \sigma \sigma^\top e^{(t-s) \beta} ds$. Therefore, the transition density function of $\overline{X}_t^y$ is given by: for $(t, x, y) \in (0, T] \times \mathbb{R}^{2n},$
\[
\mathbb{P}(t, y; x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma_t)}} \exp \left( -\frac{1}{2} (x - \mu_t)^\top \Sigma_t^{-1} (x - \mu_t) \right). \tag{56}
\]
Then, by (29) of Theorem 3.5, for any probability density function $u_0(x)$ on $x \in \mathbb{R}^n$, we have the solution of the extended RM equation (4) given by
\[
u(t, x) = \frac{1}{\phi(x)} \frac{\int_{\mathbb{R}^n} \phi(y) \mathbb{P}(t, y; x) u_0(y) dy}{\int_{\mathbb{R}^n} \phi(y) \mathbb{P}(t, y; x) u_0(y) dy}, \quad (t, x) \in [0, T] \times \mathbb{R}^n.
\]

4 Propagation of chaos

The aim of this section is to establish the propagation of chaos of the FPK equation (5). It is equivalent to the convergence of the empirical measure of a particle system to an arbitrary solution to the FPK equation (5) with respect to a suitable distance adapted to our probabilistic framework.

4.1 Particle system

Building upon the probabilistic solution form of the extended RM equation discussed in Section 3, we construct a suitable coupling between a particle system and the realization of the solution to the FPK equation (5).

Before introducing the state dynamics of the particle system proposed in this section, we first impose the following assumption on the fitness function used in this section:

(A$_{gl}$) (i) the fitness function $g : D \to \mathbb{R}$ is continuous and bounded from above; (ii) there exists a polynomial $Q_g : \mathbb{R}^+ \to \mathbb{R}^+$ such that
\[ |g(x) - g(y)| \leq Q_g(|x| + |y||x - y|), \quad \text{for all } x, y \in D. \tag{57} \]

Note that the condition (i) on the boundedness from above of $g$ in assumption (A$_{gl}$) implies the assumption (A$_g$). It can be seen that the fitness functions in the examples presented in Section 3 satisfy the condition (57) in assumption (A$_{gl}$). We next describe the particle system. More precisely, applying the solution representation (6) of the FPK equation given in Lemma 3.2, we introduce the following particle system as follows: let the number of particles be given by $N \geq 1$. For $i = 1, \ldots, N$, the dynamics of the state process of the $i$-th particle is given by, for $t \in [0, T]$,
\[ dX_t^i = b(X_t^i) dt + \sigma(X_t^i) dW_t^i, \quad X_0^i \in D, \tag{58} \]
where $W^i = (W_t^i)_{t \in [0, T]}, \ i = 1, \ldots, N$ and $W = (W_t)_{t \in [0, T]}$ are independent ($m$-dimensional) Brownian motions under the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Here, the initial values $(X_0^i)_{i \geq 1}$ is assumed to satisfy that

(A$_{X_0}$) For $q \geq 2$, the sequence of r.v.s $(X_0^i)_{i \geq 1}$ is i.i.d. according to the probability distribution $\rho_0 \in \mathcal{P}_{\deg(Q_g)+1}^{\deg(Q_g)}(D)$, where $\deg(Q_g)$ denotes the degree of the polynomial $Q_g$ which is given in the assumption (A$_{gl}$).
We also strengthen the assumption on the coefficients \((b, \sigma)\) of SDE (2) imposed in \((A_{b,\sigma})\) used in the previous section as:

\[(A'_{b,\sigma})\ b: D \to \mathbb{R}^{n \times 1} \text{ and } \sigma: D \to \mathbb{R}^{n \times m} \text{ are Lipschitz continuous with linear growth.}\]

The current assumption \((A'_{b,\sigma})\) can yield the assumption \((A_{X})\), and moreover, under the assumption \((A_{X_{t}})\), for \(q \geq 2\), there exists a constant \(C = C_{g,q,b,\sigma,T} > 0\) such that

\[
E \left[ \sup_{t \in [0,T]} |X_{t}^{i}(\deg(Q_{q})+1)q| \right] \leq C \left\{ 1 + E \left[ |X_{0}^{i}(\deg(Q_{q})+1)q| \right] \right\} < \infty, \tag{59}\]

see also Chapter 5 in Karatzas and Shreve (1991).

Motivated from the solution representation (6) of the FPK equation (5) given in Lemma 3.2, we introduce the following sequence of \(\mathcal{P}(D)\)-valued processes defined by, for \(N \geq 1\), and \(t \in [0,T]\),

\[
\mu_{t}^{N} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_{t}^{i}} \exp \left( \int_{0}^{t} g(X_{s}^{i})ds \right), \quad \text{on } \mathcal{B}_{D}. \tag{60}\]

Similarly to (8), for \(N \geq 1\), the \(\mathcal{P}(D)\)-valued process \(\mu^{N} = (\mu_{t}^{N})_{t \in [0,T]}\) also satisfies that, for \(t \in [0,T]\),

\[
\mu_{t}^{N} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_{t}^{i}} \exp \left( \int_{0}^{t} (g(X_{s}^{i}) - \langle \mu_{s}^{N}, g \rangle)ds \right), \quad \text{on } \mathcal{B}_{D}. \tag{61}\]

From the representation (61), \((\mu_{t}^{N})_{N \geq 1}\) is in fact a sequence of empirical measure-valued processes with an exponential weight.

### 4.2 Metric between \(\mu^{N}\) and \(\mu\)

This section introduces an appropriate metric between \(\mu^{N}\) and \(\mu\) for establishing the propagation of chaos. We first give the following spaces. Let \(\mathcal{P}^{*}(D)\) be the set of sub-probability measures on \(\mathcal{B}_{D}\), i.e., for any \(\mu \in \mathcal{P}^{*}(D)\), \(\mu\) is a finite measure on \(\mathcal{B}_{D}\) such that \(\mu(D) \leq 1\). We next introduce the (Alexandroff) one-point compactification. Add one point which is outside of \(D\) to \(D\) called "\(\star\)" and define \(D_{\star} := D \cup \{\star\}\). Let \(D\) be topologized by a topology \(\mathcal{T}\), and we then can define a topology \(\mathcal{T}^{\star}\) for \(D_{\star}\) as follows: (i) each open subset of \(D\) is also in \(\mathcal{T}^{\star}\), i.e., \(\mathcal{T} \subset \mathcal{T}^{\star}\); (ii) for each compact set \(C \subset D\), define an element \(U_{C} \in \mathcal{T}^{\star}\) by \(U_{C} := (D \setminus C) \cup \{\star\}\). Let us define a bijection \(\iota: \mathcal{P}(D) \to \mathcal{P}(D_{\star})\) as: for any \(A \in \mathcal{B}(D_{\star})\),

\[
(\iota \mu)(A) := \mu(A \cap D) + (1 - \mu(D)) \delta_{\star}(A), \quad \mu \in \mathcal{P}(D). \tag{62}\]

Then, the integral of \(\mu \in \mathcal{P}(D)\) w.r.t. a measurable function \(f: D_{\star} \to \mathbb{R}\) is defined as (if it is well-defined):

\[
\int_{D_{\star}} f(x)(\iota \mu)(dx) = \int_{D} f(x)(\mu(dx) + f(\star)(1 - \mu(D))) = \langle \mu, f \rangle + f(\star)(1 - \mu(D)). \tag{63}\]

Consider \(\mu = (\mu_{t})_{t \in [0,T]}\) as an arbitrary \(\mathcal{P}(D)\)-valued solution of the FPK equation (5), and we then define that

\[
\mu_{t}^{g} := \exp \left( \int_{0}^{t} (\iota \mu_{s}, g)ds \right) \mu_{t}, \quad t \in [0,T]. \tag{64}\]
This implies that, for any test function \( f \in C_0^\infty(D) \),
\[
\langle \mu_t^q, f \rangle = \langle \rho_0, f \rangle + \int_0^t \langle \mu_s^q, (A + g) f \rangle ds, \quad t \in [0, T].
\] (65)
For the sequence of empirical measure-valued processes \( (\mu^N)_{N \geq 1} \) defined by (60), similarly to (64), we can also define
\[
\mu_t^{g,N} := \exp \left( \int_0^t \langle \mu_s^q, g \rangle ds \right) \mu_t^N, \quad t \in [0, T].
\] (66)
Hereafter, without loss of generality, we assume that \( g(x) \leq 0 \) for all \( x \in D \) by the assumption \((A,tl)\). Then, for any \( t \in (0, T] \), it follows from (64) and (66) that \( \mu_t^q \) and \( \mu_t^{g,N} \) are sub-probability measures on \( B_D \), i.e., \( \mu_t^q, \mu_t^{g,N} \in \mathcal{P}^q(D) \) (note that \( \mu_0^{g,N} = \mu^N \) and \( \mu_0^q = \mu_0 \) and hence \( \mu_0^{g,N}, \mu_0^q \in \mathcal{P}(D) \)).

For the parameter \( q \geq 2 \) given in the assumption \((A,tl)\), we next establish a metric \( d_{q,T} \) between \( \mu = (\mu_t)_{t \in [0,T]} \) and \( \mu^N = (\mu_t^{N})_{t \in [0,T]} \) given as follows:
\[
d_{q,T}(\mu, \mu^N) := \left\{ \mathbb{E} \left[ \sup_{t \in [0,T]} d_{BL}(\mu_t^{g,N}, \mu_t^q)^q \right] \right\}^{1/q},
\] (67)
where \( d_{BL} \) is Fortet-Mourier distance (see, e.g. Section 11.2 of Dudley (2002), page 390), which is in fact defined as:
\[
d_{BL}(\mu_t^{g,N}, \mu_t^q) = \sup_{x \in D_\ast} \left( \int_{D_\ast} \psi(x) (\mu_t^{g,N} - \mu_t^q)(dx) \right).
\] (68)
Here, \( D_\ast \) is the set of (bounded) Lipschitzian continuous functions \( \psi : D_\ast \to \mathbb{R} \) satisfying \( \| \psi \|_\infty + \| \psi \|_{\text{Lip}} \leq 1 \) (where \( \| \psi \|_\infty := \sup_{x \in D_\ast} |\psi(x)| \) and \( \| \psi \|_{\text{Lip}} \) denotes the Lipschitzian coefficient of \( \psi \)). Let \( | \cdot | \) be the Euclidean norm. Then, we can define a metric \( d_* \) on \( D_\ast \) as in Mandelkern (1989): fix \( x_0 \in D \) and define \( l(x) := \frac{1}{1 + |x - x_0|} \) for \( x \in D \). For \( x_1, x_2 \in D \), define \( d_*(x_1, x_2) := |x_1 - x_2| \wedge (l(x_1) + l(x_2)), d_*(x, x) := l(x) \) for \( x \in D \), and \( d_*(x, x) = 0 \). Then, the Lipschitzian coefficient of \( \psi \) (as a seminorm) is given by
\[
\| \psi \|_{\text{Lip}} = \sup_{x_1 \neq x_2, x_1, x_2 \in D} \frac{|\psi(x_1) - \psi(x_2)|}{d_*(x_1, x_2)}.
\]
This implies that, for any \( \psi \in R_\ast \), and \( x_1, x_2 \in D \),
\[
|\psi(x_1) - \psi(x_2)| \leq |x_1 - x_2| \wedge (l(x_1) + l(x_2)) \leq |x_1 - x_2|.
\] (69)
In other words, \( \psi \in R_\ast \) is also a (bounded) Lipschitzian continuous function on \( D \) with the Lipschitzian coefficient being less than one.

### 4.3 Propagator of FPK equation

This section establishes the propagator corresponding to the transformed FPK equation (65), which is defined as: for \( (t, x) \in [0, T] \times D \),
\[
P_t^q f(x) := \mathbb{E} \left[ f(X_t^{L,x}) \exp \left( \int_t^T g(X_s^{L,x}) ds \right) \right], \quad f \in C_b(D),
\] (70)
where the process \( X_t^{L,x} = (X_s^{L,x})_{s \in [t, T]} \) is the unique strong solution of SDE (2), i.e.,
\[
X_s^{L,x} = x + \int_t^s b(X_r^{L,x}) dr + \int_t^s \sigma(X_r^{L,x}) dW_r, \quad s \in [t, T].
\]
The following lemma can be proved by verifying the conditions of Theorem 1 and Lemma 2 in Health and Schweizer (2000).
Lemma 4.1. Let assumptions \((A_{b,\sigma}), (A_X)\) and \((A_{\eta})\)-(i) hold. Suppose also that \((A_{D,\sigma})\) there exists a sequence \((D_k)_{k\in\mathbb{N}}\) of bounded domains with \(\bigcup_{k=1}^{\infty} D_k = D\) such that \(\bigcup_{k=1}^{\infty} D_k = D\), each \(D_k\) has a \(C^2\)-boundary; and for each \(k \geq 1\), \(\sigma\sigma^\top(x)\) is uniform elliptic on \(\mathbb{R}^n\) for \((t,x) \in [0,T) \times D_k\).

Then, the propagator \(P_{t,T}^g f\) defined by (70) satisfies that, for \((t,x) \in (0,T) \times D\),

\[
\partial_t P_{t,T}^g f(x) + (A + g)P_{t,T}^g f(x) = 0, \quad P_{0,T}^g f(x) = f(x). \tag{71}
\]

The operator \(A\) is defined by (1). Moreover, \(P_{t,T}^g f \in C^{1,2}((0,T] \times D) \cap C([0,T] \times D)\) and there exists a unique classical solution of the Cauchy problem (71).

The propagator defined by (70) can be used to establish the following relation satisfied by \(\mu_{t,T}^g - \mu_{0,T}^g\) for any fixed \(T > 0\):

Lemma 4.2. Let the conditions of Lemma 4.1 hold. Then, for any fixed \(T > 0\), it holds that

\[
\langle \mu_{t,T}^g - \mu_{0,T}^g, f \rangle = \langle \mu_{0,T}^g - \mu_{0,0}^g, P_{0,T}^g f \rangle + \frac{1}{N} \sum_{i=1}^{N} \int_0^T \exp \left( \int_0^s g(X_i^s) \, dr \right) \nabla_x P_{s,T}^g f(X_i^s) \sigma(X_i^s) dW_i^s,
\]

where \(P_{t,T}^g f\) for \(t \in [0,T]\) is the propagator defined by (70).

Proof. Recall the state process of the particle system \(X^i = (X_i^t)_{t \in [0,T]}\) defined by (58) for \(i \geq 1\). Lemma 4.1 allows us to apply Itô’s formula to \(P_{t,T}^g f(X_i^t)\) with \(t \in [0,T]\), and we have that, for \(t \in [0,T]\),

\[
P_{t,T}^g f(X_i^t) = P_{0,T}^g f(X_i^0) + \int_0^t (\partial_s + A) P_{s,T}^g f(X_i^s) \, ds + \int_0^t \nabla_x P_{s,T}^g f(X_i^s) \sigma(X_i^s) \, dW_i^s.
\]

Thus, by Eq. (71) in Lemma 4.1, it yields that

\[
P_{t,T}^g f(X_i^t) \exp \left( \int_0^t g(X_i^s) \, ds \right) = P_{0,T}^g f(X_i^0) + \int_0^t \exp \left( \int_0^s g(X_i^s) \, dr \right) \nabla_x P_{s,T}^g f(X_i^s) \sigma(X_i^s) \, dW_i^s.
\]

Recall (66). Then, using (61), it follows that

\[
\langle \mu_{t,T}^g, P_{t,T}^g f \rangle = \exp \left( \int_0^t \langle \mu_s^g, g \rangle \, ds \right) \langle \mu_{t,T}^g, P_{t,T}^g f \rangle = \frac{1}{N} \sum_{i=1}^{N} P_{t,T}^g f(X_i^t) \exp \left( \int_0^t g(X_i^s) \, ds \right)
\]

\[= \langle \mu_{0,T}^g, P_{0,T}^g f \rangle + \frac{1}{N} \sum_{i=1}^{N} \int_0^t \exp \left( \int_0^s g(X_i^s) \, dr \right) \nabla_x P_{s,T}^g f(X_i^s) \sigma(X_i^s) \, dW_i^s.
\]

By equalities (65) and (71) in Lemma 4.1, it holds that

\[
\partial_t \langle \mu_t^g, P_{t,T}^g f \rangle = 0, \quad \text{for all } t \in [0,T]. \tag{73}
\]

Therefore \(\langle \mu_t^g, P_{t,T}^g f \rangle = \langle \mu_0^g, P_{0,T}^g f \rangle\) for all \(t \in [0,T]\). This gives that

\[
\langle \mu_{t,T}^g - \mu_{0,T}^g, f \rangle = \langle \mu_{0,T}^g - \mu_{0,0}^g, P_{0,T}^g f \rangle + \frac{1}{N} \sum_{i=1}^{N} \int_0^t \exp \left( \int_0^s g(X_i^s) \, dr \right) \nabla_x P_{s,T}^g f(X_i^s) \sigma(X_i^s) \, dW_i^s.
\]

In the equality, taking \(t = T\) and using \(P_{T,T}^g f = f\) given in Lemma 4.1, we arrive at (72). Thus, we complete the proof of the lemma.
By (63), we have that, for all $\psi \in \mathcal{R}_1$,
\[
\int_{D^*} \psi(x)(\mu^0_{pq} - \mu^0_q)(dx) = \int_D \psi_s(x)(\mu^0_{pq} - \mu^0_q)(dx) = \langle \mu^0_{pq} - \mu^0_q, \psi_s \rangle,
\]
where $\psi_s(x) := \psi(x) - \psi(\star)$ for $x \in D$. It follows from (69) that $\psi_s$ is a (bounded) Lipschitzian continuous function on $D$ with the Lipschitzian coefficient being less than one. Then, we have the following estimate of the gradient $\nabla_x P_{t,T} \psi_s$ which is given by

Lemma 4.3. Let assumptions $\mathbf{(A'_{t,b})}$, $\mathbf{(A_{D,a})}$ and $\mathbf{(A_{g'})}$ hold. Then, there exists a polynomial $\hat{Q}_g = \hat{Q}_{g,b,s,T} : \mathbb{R}_+ \to \mathbb{R}_+$ (depending only on $g, b, \sigma, T$) with $\deg(\hat{Q}_g) = \deg(Q_g)$, such that, for all $\psi \in \mathcal{R}_1$,
\[
|\nabla_x P_{t,T} \psi_s(x)| \leq \hat{Q}_g(|x|), \quad (t, x) \in [0, T] \times D.
\]

Proof. Note that $\|\psi_s\|_\infty \leq 2$ for all $\psi \in \mathcal{R}_1$. Then, in view of (70), for all $x_1, x_2 \in D$,
\[
\left| P^g_{t,T} \psi_s(x_1) - P^g_{t,T} \psi_s(x_2) \right|^2 \leq 2E \left[ \left| \psi_s(X^{t,x_1}_T) \right| \psi_s(X^{t,x_2}_T) \right] \left[ \exp \left( 2 \int_t^T g(X^{t,x}_s) ds \right) \right] + 2E \left[ \left| \psi_s(X^{t,x_2}_T) \right| \left| \psi_s(X^{t,x_2}_T) \right| \right] \left[ \exp \left( \int_t^T g(X^{t,x_2}_s) ds \right) \right] \leq 2E \left[ \left| X^{t,x_1}_T - X^{t,x_2}_T \right| \right]^2 + 8E \left[ \left| \int_t^T g(X^{t,x_1}_s) ds - \int_t^T g(X^{t,x_2}_s) ds \right| \right] \leq 2E \left[ \left| X^{t,x_1}_T - X^{t,x_2}_T \right| \right]^2 + 8T \int_t^T E \left[ |g(X^{t,x_1}_s) - g(X^{t,x_2}_s)|^2 \right] ds.
\]

We first note that, it follows from the assumption $\mathbf{(A_{g'})}$-(ii) that, for some polynomial $Q_g : \mathbb{R}_+ \to \mathbb{R}_+$ and for all $s \in [t, T]$,
\[
E \left[ \left| g(X^{t,x_1}_s) - g(X^{t,x_2}_s) \right| \right] \leq E \left[ \left| Q_g \left( \left| X^{t,x_1}_s \right| + \left| X^{t,x_2}_s \right| \right) \right| \right] \left| X^{t,x_1}_s - X^{t,x_2}_s \right|^2 \leq \left\{ \frac{E \left[ \left| X^{t,x_1}_s \right| + \left| X^{t,x_2}_s \right| \right]}{2} \right\} \left\{ \frac{E \left[ \left| X^{t,x_1}_s \right| - X^{t,x_2}_s \right|^4}{4} \right\} \frac{1}{2}.
\]

On the other hand, by the assumption $\mathbf{(A'_{b,s})}$, we have that, for $k \geq 1$,
\[
E \left[ \sup_{s \in [t,T]} \left| X^{t,x_1}_s - X^{t,x_2}_s \right|^{2k} \right] \leq |x_1 - x_2|^{2k} + C_k \int_t^T E \left[ \left| X^{t,x_1}_s - X^{t,x_2}_s \right|^{2k} \right] ds,
\]
where $C_k = C_{b,s,T,k}$ is a positive constant which depends on $b, \sigma, T, k$. Then, the Gronwall’s lemma yields that
\[
E \left[ \sup_{s \in [t,T]} \left| X^{t,x_1}_s - X^{t,x_2}_s \right|^{2k} \right] \leq |x_1 - x_2|^{2k} e^{(T-t)C_k}.
\]

Then, it follows from (76), (77) and (79) that, for all $(t, x_1, x_2) \in [0, T] \times D^2$,
\[
\left| P^g_{t,T} \psi_s(x_1) - P^g_{t,T} \psi_s(x_2) \right|^2 \leq 2E \left[ \left| X^{t,x_1}_T - X^{t,x_2}_T \right| \right]^2 + 8T \int_t^T E \left[ |g(X^{t,x_1}_s) - g(X^{t,x_2}_s)|^2 \right] ds \leq 2|x_1 - x_2|^2 e^{(T-t)C_1} + 8T^2 \int_t^T \left\{ Q_g \left( \left| X^{t,x_1}_s \right| + \left| X^{t,x_2}_s \right| \right) \right\} \left\{ \frac{E \left[ \left| X^{t,x_1}_s \right| - X^{t,x_2}_s \right|^4}{4} \right\} \frac{1}{2} ds.
\]
\[2e^{(T-t)C_1}|x_1 - x_2|^2 + 8T^2e^{\frac{(T-t)C_2}{2}}|x_1 - x_2|^2 \int_t^T \left\{ \mathbb{E} \left[ Q^g_s \left( |X^s_{t,x_1}| + |X^s_{t,x_2}| \right) \right] \right\} \frac{1}{2} ds.\]

It is not difficult to show from (59) that there exists a polynomial \( \hat{Q}_g := \hat{Q}_{g,b,\sigma,T} : \mathbb{R}_+ \to \mathbb{R}_+ \) (which depends on \( g, b, \sigma, T \) only) such that \( \deg(\hat{Q}_g) = 2 \deg(Q_g) \), and for all \((t, x_1, x_2) \in [0, T] \times D^2,\)
\[2e^{(T-t)C_1} + 8T^2e^{\frac{(T-t)C_2}{2}} \int_t^T \left\{ \mathbb{E} \left[ Q^g_s \left( |X^s_{t,x_1}| + |X^s_{t,x_2}| \right) \right] \right\} \frac{1}{2} ds \leq \hat{Q}_g(|x_1| + |x_2|).\]

Therefore, we arrive at, for all \((t, x_1, x_2) \in [0, T] \times D^2,\)
\[|P^g_{t,T}\psi_s(x_1) - P^g_{t,T}\psi_s(x_2)|^2 \leq \hat{Q}_g(|x_1| + |x_2|)|x_1 - x_2|^2. \quad (80)\]

Thanks to Lemma 4.1, we have that \( P^g_{t,T}\psi_s \in C^{1,2}((0, T] \times D) \cap C([0, T] \times D) \) by the fact of \( \psi_s \in C_b(D) \). By letting \( x_2 \to x_1 \) in (80), it follows that, for all \((t, x) \in [0, T] \times D,\)
\[|\nabla_x P^g_{t,T}\psi_s(x)| \leq \hat{Q}_g^\frac{1}{2}(2|x|). \quad (81)\]

Then, the estimate (75) follows from (81) and the fact that \( \hat{Q}_g^\frac{1}{2}(2|x|) \leq \hat{Q}_g(|x|) \) for some polynomial \( \hat{Q}_g := \hat{Q}_{g,b,\sigma,T} : \mathbb{R}_+ \to \mathbb{R}_+ \) with \( \deg(\hat{Q}_g) = \frac{1}{2} \deg(Q_g) = \frac{1}{2} \deg(\hat{Q}_g) \). Thus, we complete the proof of the lemma.

We next discuss the estimate (75) in Lemma 4.3 for the one-dimensional case (i.e., \( m = n = 1 \)), but the volatility function \( \sigma : D \to \mathbb{R} \) of SDE (2) is only Hölder continuous. To this purpose, we impose the following assumption:

\textbf{(A}_{1d}\textbf{)} \ for \( m = n = 1, \ b : D \to \mathbb{R} \) is Lipschitz continuous and \( \sigma : D \to \mathbb{R} \) is Hölder continuous with exponent \( \gamma \in \left[ \frac{1}{2}, 1 \right) \).

\textbf{Lemma 4.4.} Let assumptions \textbf{(A}_{1d}\textbf{),} \ (\textbf{A}_{D,\sigma}\textbf{)} and \textbf{(A}_{gl}\textbf{)} with \( \deg(Q_g) = 0 \) hold. Then, there exists a constant \( K = K_{g,b,\sigma,T} > 0 \) such that, for all \( \psi \in R_1, \)
\[|\nabla_x P^g_{t,T}\psi_s(x)| \leq K, \quad (t, x) \in [0, T] \times D. \quad (82)\]

\textbf{Proof.} Note that, under the assumption \textbf{(A}_{1d}\textbf{),} the estimate (78) in the proof of Lemma 4.3 does not hold. However, the assumption \textbf{(A}_{1d}\textbf{)} implies that SDE (2) satisfies the Yamada-Watanabe condition given in Proposition 5.2.13 of Karatzas and Shreve (1991). This yields from this proposition that, for all \( s \in [t, T], \) there exists a constant \( C = C_{b,\sigma,T} > 0 \) such that
\[\mathbb{E} \left[ |X^s_{t,x_1} - X^s_{t,x_2}| \right] \leq |x_1 - x_2| + C \int_t^T \mathbb{E} \left[ |X^s_{t,x_1} - X^s_{t,x_2}| \right] ds.\]

Then, the Gronwall’s lemma yields that, for all \( s \in [t, T], \)
\[\mathbb{E} \left[ |X^s_{t,x_1} - X^s_{t,x_2}| \right] \leq |x_1 - x_2|e^{(T-t)C}. \quad (83)\]

Note that, under the assumption \textbf{(A}_{gl}\textbf{)} with \( \deg(Q_g) = 0 \), the fitness function \( g : D \to \mathbb{R} \) is hence Lipschitz continuous. Similarly to (76), we obtain from (83) that
\[|P^g_{t,T}\psi_s(x_1) - P^g_{t,T}\psi_s(x_2)| \leq \mathbb{E} \left[ |X^s_{T,x_1} - X^s_{T,x_2}| \right] + 2 \int_t^T \mathbb{E} \left[ |g(X^s_{t,x_1}) - g(X^s_{t,x_2})| \right] ds \leq (1 + 2T\|g\|_{\text{Lip}})e^{TC}|x_1 - x_2|.

This proves the estimate (82). \hfill \square
Building upon the auxiliary results established by the above lemmas, we next prove the main result of this section on the propagation of chaos of the FPK equation (5) with respect to the metric defined by (67):

**Theorem 4.5.** Let assumptions \( (\mathbf{A}_X, \mathbf{A}^\prime) \), \( (\mathbf{A}^\prime, \mathbf{A}) \), \( (\mathbf{A}_D, \mathbf{A}) \) and \( (\mathbf{A}^\prime, \mathbf{A}) \) with \( \deg(Q_g) = 0 \) hold. Let \( \mu = (\mu_t)_{t \in [0,T]} \) be an arbitrary \( \mathcal{P}(\mathcal{D}) \)-valued solution of the FPK equation (5) and \( \mu^N = (\mu^N_t)_{t \in [0,T]} \) be the \( \mathcal{P}(\mathcal{D}) \)-valued process given by (60) in Section 4.1. Then, for any \( T > 0 \) and \( N \geq 1 \), there exists a constant \( C > 0 \) which is independent of \( N \) such that, for any \( p \geq 1 \),

\[
d_{q,T}(\mu, \mu^N) \leq C \left( \alpha(p, q, n, N) + \frac{1}{N^{q-1}} \right), \quad q \geq 2,
\]

where the metric \( d_{q,T}(\cdot, \cdot) \) is defined by (67), and the first rate of convergence rate \( \alpha(p, q, n, N) \) in (84) is given by

\[
\alpha(p, q, n, N) := \begin{cases} 
N^{-\frac{1}{2}} + N^{-\frac{p-q}{p}}, & q > \frac{n}{2}, \ p \neq 2q; \\
N^{-\frac{1}{4}} \ln(1 + N) + N^{-\frac{p-q}{p}}, & q = \frac{n}{2}, \ p \neq 2q; \\
N^{-\frac{n}{2}} + N^{-\frac{p-q}{p}}, & q < \frac{n}{2}, \ p \neq \frac{n}{q},
\end{cases}
\]

**Proof.** Using (72) in Lemma 4.2, it results in, for all \( \psi \in \mathcal{R}_1 \),

\[
\langle \mu^N_T - \mu_T, \psi \rangle = \langle \mu^N_0 - \mu_0, P^N_{0,T} \psi \rangle + \frac{1}{N} \sum_{i=1}^N \int_0^T \exp \left( \int_0^s g(X^i_s) \right) \nabla_x P^N_{s,T} \psi(X^i_s)^\top \sigma(X^i_s) dW^i_s.
\]

We next consider the estimate of the first term of the r.h.s. of the equality (86). Note that

\[
\mu^N_0 = \frac{1}{N} \sum_{i=1}^N \delta_{X^i_0}, \quad \mu_0 = \rho_0.
\]

Let \( \rho_0 := P_{0,T} \psi \) with \( \psi \in \mathcal{R}_1 \). Then, it follows from the assumption \( (\mathbf{A}_X) \) that \( (\phi_0(X^i_0))_{i \geq 1} \) are i.i.d. r.v.s. Let \( \rho_0 := \mathcal{L}(\phi_0(X^i_0)), \) i.e., the law of the r.v. \( \phi_0(X^i_0) \), and define \( \tilde{\mu}^N := \frac{1}{N} \sum_{i=1}^N \delta_{\phi_0(X^i_0)}. \) Then, for \( I(x) = x \) for \( x \in D \), we obtain from (89) and the Kantorovich-Rubinstein dual formula (see Villani (2003)) that, for \( q \geq 2 \),

\[
\mathbb{E} \left[ \left| \langle \mu^N_0 - \mu_0, P^N_{0,T} \psi \rangle \right|^q \right] \leq \mathbb{E} \left[ \| \mu^N_0 - \mu_0, \phi_0 \|^q \right] \leq \mathbb{E} \left[ \| \mu^N_0 - \rho_0 \|^q \right] \leq \mathbb{E} \left[ \| \tilde{\mu}^N - \rho_0 \|^q \right].
\]

By (70) and the fact that \( \| \psi \|_{\infty} \leq 2 \), we have from the assumption (\( \mathbf{A}_g \))-(i) (without loss of generality, we have assumed that \( g \leq 0 \)) that \( \| \phi_0 \|_{\infty} \leq 2 \). Therefore \( \int_D |x|^p \rho_0(dx) = \int_D |\phi_0(x)|^p \rho_0(dx) \leq 2^p \) for any \( p \geq 1 \). Note that, under the assumption \( (\mathbf{A}_X) \), by Glivenko-Cantelli’s theorem, the empirical measure \( \tilde{\mu}^N \) tends weakly to \( \rho_0 \) as \( N \to \infty \). Moreover, Theorem 1 of Fournier and Guillin (2015) yields that, there is a constant \( C \) depending only on \( n, p, q \) such that

\[
\mathbb{E} \left[ \| \tilde{\mu}^N - \rho_0 \|^q \right] \leq C \left( \int_D |x|^p \rho_0(dx) \right)^{\frac{q}{p}} \alpha(p, q, n, N) \leq 2^p C \alpha(p, q, n, N),
\]
where \( \mathcal{W}_q \) denotes the Wasserstein metric with order \( q \geq 2 \) on the \( q \)-th order Wasserstein space \( \mathcal{P}_q(D) \) (see Villani (2003)), and the rate of convergence of rate is given by (85). Then, in view of (88) and (89), there is a constant \( C \) depending only on \( p, q, n \) such that

\[
E \left[ \left\| \langle \mu_0 - \mu_{0,T}^g \rangle \right\|_q^q \right] \leq C \alpha(p, q, n, N).
\]

(90)

On the other hand, apply the estimate (75) in Lemma 4.3 or the estimate (82) in Lemma 4.4. Then, for the estimate of the second of the r.h.s. of the equality (86), under the assumption \((\mathbf{A}_{X_0})\) or the H"{o}lder continuity given in the assumption \((\mathbf{A}_{1q})\), there exists a constant \( C = C_{g,b,\sigma,q,T} > 0 \) which may differ from line to line such that, \( \mathbb{P}\)-a.s.

\[
\hat{Q}_g(|X^1_s|) |\sigma(X^1_s)|^2 \leq P_1 \left( \sup_{t \in [0,T]} |X^1_t| \right), \quad s \in [0,T],
\]

(92)

where \( P_1 : \mathbb{R}_+ \to \mathbb{R}_+ \) is a polynomial with \( \text{deg}(P_1) = \text{deg}(Q_g) + 1 \). By (59) under the assumption \((\mathbf{A}_{X_0})\), it follows that

\[
E \left[ \sup_{t \in [0,T]} |X^1_t|^{\text{deg}(P_1)q} \right] \leq C \left\{ E \left[ |X^1_0|^{\text{deg}(P_1)q} \right] + 1 \right\} < \infty,
\]

(93)

for some constant \( C = C_{g,b,\sigma,T} > 0 \). Then, by applying (91), (92) and (93), there exists a constant \( C = C_{g,b,\sigma,T,q} > 0 \) which may differ from line to line such that

\[
E \left[ \sup_{t \in [0,T]} \left\| \frac{1}{N} \sum_{i=1}^N \int_0^t \exp \left( \int_0^s g(X^i_s)ds \right) \nabla_x P^g_{s,T} \psi_*(X^i_s) \sigma(X^i_s) dW^i_s \right\|^q \right] \\
\leq \frac{C}{N^{q-1}} E \left[ \left( \int_0^T \check{Q}_g^2(|X^1_s|) |\sigma(X^1_s)|^2 ds \right)^{q/2} \right] \\
\leq \frac{C}{N^{q-1}} \left\{ E \left[ |X^1_0|^{\text{deg}(P_1)q} \right] + 1 \right\}.
\]

(94)

Then, the estimate (84) can be deduced from (94) and (90) jointly. Thus, we complete the proof of the theorem.
Recall that Example 3.8 relates to the one-dimensional RM equation with $b(x) = a + bx$, $\sigma(x) = \sigma \sqrt{x}$ and $g(x) = -x$ for $x \in D = (0, \infty)$; Example 3.9 relates to the one-dimensional RM equation with $b(x) \equiv 0$, $\sigma(x) = \sigma \sqrt{2}$ and $g(x) = -x^{2q} + \sum_{l=0}^{2q-1} \alpha_l x^l$ given by (46) for $x \in D = (-\infty, \infty)$; Example 3.10 relates to the $n$-dimensional RM equation with $b(x) = b + Bx$, $\sigma(x) \equiv \sigma$, and $g(x) = -r(x)$ given by (51) for $x \in D = \mathbb{R}^n$. It is not difficult to verify the validity of assumptions $(A_{1d})$, $(A_{D,\sigma})$ and $(A_{gl})$ with deg$(Q_g) = 0$ for Example 3.8, and the validity of assumptions $(A_{b,\sigma})$, $(A_{D,\sigma})$ and $(A_{gl})$ for Example 3.9 and Example 3.10. Therefore, the propagation of chaos with the rate of convergence (84) in Theorem 4.5 under the assumption $(A_{X_0})$ on the initial data of the particle system holds for these examples.

References

Alfaro, M., and R. Carles (2014): Explicit solutions for replicator-mutator equations: extinction versus acceleration. *SIAM J. Appl. Math.* 74, 1919-1934.

Alfaro, M., and R. Carles (2017): Replicator-mutator equations with quadratic fitness. *Proc. Amer. Math. Soc.* 145, 5315-5327.

Alfaro, M., and M. Veruete (2019): Evolutionary branching via replicator-mutator equations. *J. Dyn. Diff. Eqn.* 31, 2029-2052.

Bürger, R. (1988): Perturbations of positive semigroups and applications to population genetics. *Math. Z.* 197, 259-272.

Bürger, R. (1998): Mathematical properties of mutation-selection models. *Genetica* 102, Article number: 279.

Del Moral, P., and L. Miclo (2000): A Moran particle system approximation of Feynman-Kac formulae. *Stoch. Process. Appl.* 86, 193-216.

Dudley, R.M. (2002): *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics Book 74, Cambridge University Press, Cambridge.

Duffie, D., D. Filipović, and W. Schachermayer (2003): Affine processes and applications in finance. *Ann. Appl. Probab.* 13, 984-1053.

Fleming, W.H. (1979): Equilibrium distributions of continuous polygenic traits. *SIAM J. Appl. Math.* 36, 148-168.

Fournier, N., and A. Guillin (2015): On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory & Rel. Fields* 162, 707-738.

Health, D., and M. Schweizer (2000): Martingale versus PDEs in finance: An equivalence result with examples. *J. Appl. Probab.* 37, 947-957.

Kac, M. (1956): *Foundations of Kinetic Theory*. In Proceedings of the Third Berkeley Symposium on Math. Stats. & Probab. 1954-1955, vol. III, 171-197. University of California Press, Berkeley and Los Angeles.

Karatzas, I., and S.E. Shreve (1991): *Borwnian motion and Stochastic Calculus*. Springer-Verlag, New York.

Kimura, M. (1965): A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Nat. Acad. Sci.* 54, 731-736.
Mandelkern, M. (1989): Metrization of the one-point compactification. *Proceedings of AMS* 107, 1111-1115.

Pinsky, R.G. (1995): *Positive Harmonic Function and Diffusion*. Cambridge University Press, Cambridge.

Qin, L., and V. Linetsky (2016): Positive eigenfunctions of Markovian pricing operators: Hansen-Scheinkman factorization and Ross recovery. *Opers. Res.* 64, 99-117.

Rouzine, I., E. Brunet, and C. Wilke (2008): The traveling-wave approach to asexual evolution: Muller’s ratchet and speed of adaptation. *Theor. Popul. Biol.* 73, 24-46.

Rouzine, I., J. Wakekey, and J. Coffin (2003): The solitary wave of asexual evolution. *Proc. Natl. Acad. Sci.* 100, 587-592.

Sniegowski, P., and P. Gerrish (2010): Beneficial mutations and the dynamics of adaptation in asexual populations. *Phil. Trans. Royal Soc. B* 365, 1255-1263.

Sznitman, A.S. (1991): *Topics in Propagation of Chaos*. In École D’Été de Probabilités de Saint-Flour XIX-1989. Lecture Notes in Math. 1464, 165-251. Springer, Berlin.

Tsimring, L., H. Levine, and D. Kessler (1996): RNA virus evolution via a fitness-space model. *Phys. Rev. Lett.* 76, 4440-4443.

Villani, C. (2003): *Topics in Optimal Transportation*. Graduate Studies in Mathematics, Volume: 58, AMS.

Xu, L.P. (2018): Uniqueness and propagation of chaos for the Boltzmann equation with moderately soft potentials. *Ann. Appl. Probab.* 28, 1136-1189.