Martingale Structure for General Thermodynamic Functionals of Diffusion Processes Under Second-Order Averaging

Hao Ge\textsuperscript{1,2} · Chen Jia\textsuperscript{3} · Xiao Jin\textsuperscript{4}

Received: 20 January 2021 / Accepted: 28 June 2021 / Published online: 28 July 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Novel hidden thermodynamic structures have recently been uncovered during the investigation of nonequilibrium thermodynamics for multiscale stochastic processes. Here we reveal the martingale structure for a general thermodynamic functional of inhomogeneous singularly perturbed diffusion processes under second-order averaging, where a general thermodynamic functional is defined as the logarithmic Radon–Nykodim derivative between the laws of the original process and a comparable process (forward case) or its time reversal (backward case). In the forward case, we prove that the regular and anomalous parts of a thermodynamic functional are orthogonal martingales. In the backward case, while the regular part may not be a martingale, we prove that the anomalous part is still a martingale. With the aid of the martingale structure, we prove the integral fluctuation theorem satisfied by the regular and anomalous parts of a general thermodynamic functional. Further extensions and applications to stochastic thermodynamics are also discussed, including the martingale structure and fluctuation theorems for the regular and anomalous parts of entropy production and housekeeping heat in the absence or presence of odd variables.

Keywords
Singularly perturbed diffusion · Model simplification · Multiscale · Stochastic thermodynamics · Entropy production · Housekeeping heat · Anomalous contribution · Fluctuation theorem

Communicated by Jorge Kurchan.

✉ Hao Ge
haoge@pku.edu.cn
✉ Chen Jia
chenjia@csrc.ac.cn

1 Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China
2 Biomedical Pioneering Innovation Center, Peking University, Beijing 100871, China
3 Applied and Computational Mathematics Division, Beijing Computational Science Research Center, Beijing 100193, China
4 School of Mathematical Sciences, Peking University, Beijing 100871, China
1 Introduction

Over the past two decades, significant progress has been made in stochastic thermodynamics [1–3], which has grown to become one of the most influential branches of statistical physics. In this field, an irreversible thermodynamic system is usually modeled by a nonstationary and nonsymmetric Markov process, where the breaking of stationarity describes Boltzmann’s irreversibility and the breaking of symmetry describes Prigogine’s irreversibility [4–6]. Along this line, an equilibrium state is defined as a stationary and symmetric Markov process and the deviation from equilibrium is usually quantified by the concept of entropy production [7,8]. Entropy production can be decomposed into two parts: the adiabatic part which is also referred to as housekeeping heat and the nonadiabatic part which is also known as free energy dissipation [4–6]. The former characterizes the deviation from symmetry and the latter characterizes the deviation from stationarity [9].

Entropy production, as well as its adiabatic and nonadiabatic parts, can also be defined along single stochastic trajectories. These thermodynamic quantities are all functionals of the stochastic trajectory of the underlying Markovian model. One of the major breakthroughs in stochastic thermodynamics is the finding that a broad class of thermodynamic functionals such as entropy production, housekeeping heat, free energy dissipation, and cycle fluxes (circulations) [10–12] satisfy various types of fluctuation theorems [13–23], which provide nontrivial generalizations of the second law of thermodynamics in terms of equalities rather than inequalities. These fluctuation theorems, which hold generally for states far from equilibrium, have garnered considerable attention from physicists, chemists, biologists, and mathematicians. Thus far, most of these fluctuation theorems have been numerically tested [24–28] or experimentally validated [29–33].

Mathematically, there is a universal approach for the investigation of fluctuation theorems. It has been shown that most thermodynamic functionals have a unified mathematical representation in terms of the logarithmic Radon–Nykodim derivative between the laws of the original process and some comparable process [34–36]:

\[ A_t = \log \frac{dP_t}{dQ_t}(X_s), \quad (1) \]

where \( A_t \) is a thermodynamic functional, \( P_t \) is the law of the original process \( X = (X_s)_{0 \leq s \leq t} \), and \( Q_t \) is the law of the comparable process \( Y = (Y_s)_{0 \leq s \leq t} \). If the thermodynamic functional of interest has the representation (1), then the following integral fluctuation theorem naturally follows:

\[ \mathbb{E} e^{-A_t} = 1. \]

It then follows from Jensen’s inequality that \( \mathbb{E} A_t \geq 0 \), which is just the conventional inequality in the second law of thermodynamics.

In recent years, another striking fact in stochastic thermodynamics beyond fluctuation theorems is that, under mild conditions, a wide class of thermodynamic functionals such as entropy production and housekeeping heat are actually martingales [37]. This not only reveals a deep connection between statistical physics and probability theory, but also enables us to apply the powerful tool of martingale theory to derive some novel statistical properties of such thermodynamic functionals, including universal equalities and inequalities for stopping times, passage statistics, and extrema [38–40]. An experimental validation of the infimum law of entropy production with a double electronic dot has been performed recently [41].

In stochastic thermodynamics, one of the most important mathematical models of a molecular system is the diffusion process, which generalizes the classical Langevin equation describing...
the stochastic movement of a mesoscopic particle in a fluid due to collisions with the fluid molecules. Diffusion processes have also been widely applied to model various stochastic systems in chemistry, biology, meteorology, and other disciplines [42]. In practice, a complex molecular system (such as biochemical reaction networks in living cells) often has multiple different time scales spanning many orders of magnitude [43]. For efficient modeling, one usually averages out rapidly varying components and only focuses on slowly varying components. This procedure is called averaging [44], coarse graining [45], or model simplification [46,47]. Mathematically, many previous papers focused on a two-timescale diffusion process whose generator has the singularly perturbed form of

\[ L^\epsilon = \epsilon^{-1}L_0 + L_1, \]

where \( \epsilon > 0 \) is a small parameter, \( L_0 \) is the generator governing fast components, and \( L_1 \) is the generator governing slow components [48–56]. The model simplification of such systems is called first-order averaging. In statistical physics, however, many important molecular systems such as the Langevin dynamics in the zero-mass limit [57] are actually singularly perturbed diffusion processes with a more complicated generator [57–59]

\[ L^\epsilon = \epsilon^{-2}L_0 + \epsilon^{-1}L_1 + L_2. \]

The model simplification of such systems is called second-order averaging which has also been studied extensively in probability theory [60–62].

Recently, it has been found that while model simplification such as first-order and second-order averaging retains the dynamic information of the original process, it loses some crucial thermodynamic information as a tradeoff [63–76]. In particular, under model simplification, it has been shown that entropy production and housekeeping heat usually decrease in the average sense [63,73]. The decreasing part emerging from the elimination of fast components is called the anomalous part of the corresponding thermodynamic functional. In addition, there have been some results showing that the anomalous part of entropy production and housekeeping heat satisfies fluctuation theorems [64,71]. However, it is unclear whether the anomalous parts of such thermodynamic functionals are martingales, and whether similar results also hold for a general thermodynamic functional defined by the representation (1).

In this paper, we investigate the martingale structure for a general thermodynamic functional with the representation (1) of inhomogeneous singularly perturbed diffusion processes under second-order averaging. Different thermodynamic functionals correspond to different comparable processes. For most thermodynamic functionals studied in the literature, the comparable process is either a diffusion process having the same diffusion coefficients as the original process or the time reversal of such a diffusion process under the reversed protocol [36]. To distinguish between these two scenarios, we refer to the former as the forward case and the latter as the backward case. Here we consider the martingale structure of a thermodynamic functional under the most general setup covering these two scenarios, leaving most known physical examples as special cases.

The main results of the present paper are stated as follows. Let \( F^\epsilon_t \) be a general forward or backward thermodynamic functional for an inhomogeneous singularly perturbed diffusion process and let \( F^{(1)}_t \) be the corresponding thermodynamic functional for the reduced process under second-order averaging. We prove that under mild conditions that are usually satisfied by real physical systems, the thermodynamic functional \( F^\epsilon_t \) has a weak limit \( F_t \) as \( \epsilon \to 0 \). The limit functional \( F_t \) is in general not equal to the thermodynamic functional \( F^{(1)}_t \) of the reduced process and thus can be decomposed into two parts:

\[ F_t = F^{(1)}_t + F^{(2)}_t. \]
where $F^{(1)}_t$ and $F^{(2)}_t$ are called the regular and anomalous parts of the limit functional, respectively.

In this paper, we prove that in the forward scenario, the thermodynamic functional $e^{-F^ε_t}$, the limit functional $e^{-F^ε_t}$, its regular part $e^{-F^{(1)}_t}$, and its anomalous part $e^{-F^{(2)}_t}$ are all martingales. Moreover, the regular and anomalous parts are actually orthogonal martingales. This provides an orthogonal martingale decomposition of the forward limit functional. In the backward scenario, the thermodynamic functional $e^{-F^ε_t}$, the limit functional $e^{-F^ε_t}$, and its regular part $e^{-F^{(1)}_t}$ may not be martingales due to a non-vanishing boundary term depending on the initial and final distributions. However, we prove that the anomalous part $e^{-F^{(2)}_t}$ is still a martingale. These results reveal a highly nontrivial relationship between model simplification and martingale. This martingale structure enables us to prove the following integral fluctuation theorem:

$$\mathbb{E} e^{-F^{(2)}_t} = 1,$$

where $τ$ is an arbitrary bounded stopping time. Therefore, as a byproduct, we prove that the anomalous part of a general thermodynamic functional satisfies the integral fluctuation theorem.

The structure of this paper is organized as follows. In Sect. 2, we introduce our model and the fundamental convergence theorem for singularly perturbed diffusion processes under second-order averaging. In Sects. 3 and 4, we state our main results about the martingale structure and fluctuation theorems for general forward and backward thermodynamic functionals. In Sects. 5 and 6, we apply our general theorems to stochastic thermodynamics. In particular, we establish the martingale structure for entropy production and housekeeping heat in the absence and presence of odd variables. Sections 7 and 8 are devoted to the proof of our main theorems.

2 Formulation

We consider a thermodynamic process modeled by a pair of inhomogeneous singularly perturbed diffusions $(X^ε, Y^ε) = (X^ε_t, Y^ε_t)_{0 ≤ t ≤ T}$ with a small parameter $ε > 0$ and a fixed ending time $T > 0$, where $X^ε$ is a diffusion with values in $\mathbb{R}^m$ and $Y^ε$ is a diffusion with values in $\mathbb{R}^n$. There are weak and strong interactions between these processes so that $X^ε$ is slowly varying and $Y^ε$ is fast varying. The system of diffusions $(X^ε, Y^ε)$ solves the stochastic differential equation (SDE)

$$\begin{align*}
&dX^ε_t = [b(X^ε_t, Y^ε_t, t) + \epsilon^{-1} f(X^ε_t, Y^ε_t, t)]dt + \sigma(X^ε_t, Y^ε_t, t)dW_t, \quad dY^ε_t = [\epsilon^{-1} g(X^ε_t, Y^ε_t, t) + \epsilon^{-2} c(X^ε_t, Y^ε_t, t)]dt + \epsilon^{-1} \eta(X^ε_t, Y^ε_t, t)dW_t,
&\text{where } W = (W_t)_{t ≥ 0} \text{ is a } p\text{-dimensional standard Brownian motion defined on a filtered probability space } (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}), b, f : \mathbb{R}^m × \mathbb{R}^n × \mathbb{R}^+ → \mathbb{R}^m \text{ and } c, g : \mathbb{R}^m × \mathbb{R}^n × \mathbb{R}^+ → \mathbb{R}^n \text{ are vector-valued functions, and } \sigma : \mathbb{R}^m × \mathbb{R}^n × \mathbb{R}^+ → \mathbb{R}^m × \mathbb{R}^p \text{ and } \eta : \mathbb{R}^m × \mathbb{R}^n × \mathbb{R}^+ → \mathbb{R}^n × \mathbb{R}^p \text{ are matrix-valued functions. Then } (X^ε, Y^ε) \text{ is an inhomogeneous diffusion process on } \mathbb{R}^{n+m} \text{ with drift}
&B(x, y, t) = \begin{pmatrix} b(x, y, t) + \epsilon^{-1} f(x, y, t) \\ \epsilon^{-1} g(x, y, t) + \epsilon^{-2} c(x, y, t) \end{pmatrix}.
\end{align*}$$
and diffusion matrix

\[
D(x, y, t) = \begin{pmatrix} 
\sigma(x, y, t) & -\sigma(x, y, t) \\
\epsilon^{-1} \eta(x, y, t) & \epsilon^{-1} \eta(x, y, t)
\end{pmatrix} \begin{pmatrix} 
\sigma(x, y, t) & -\sigma(x, y, t) \\
\epsilon^{-1} \eta(x, y, t) & \epsilon^{-1} \eta(x, y, t)
\end{pmatrix}^T = \begin{pmatrix} 
a(x, y, t) & -\epsilon^{-1} h(x, y, t) \\
\epsilon^{-1} h(x, y, t) & -\epsilon^{-2} \alpha(x, y, t)
\end{pmatrix},
\]

where

\[
a = \sigma \sigma', \quad h = \sigma \eta', \quad \alpha = \eta \eta'.
\]

Following previous papers [60–62], we assume that \((X^\epsilon, Y^\epsilon)\) starts from the same initial distribution for each \(\epsilon > 0\). Recall that the process is called homogenous if the drift and diffusion coefficients are independent of time \(t\). Otherwise, it is called inhomogeneous. Inhomogeneous diffusion processes are particularly important in statistical physics because they can be used to describe the evolution of a molecular system controlled by a time-varying protocol.

Throughout this paper, we use Einstein’s summation convention: if the same index appears twice in any term, once as an upper index and once as a lower index, that term is understood to be summed over all possible values of that index. Then the generator of \((X^\epsilon, Y^\epsilon)\) has the following second-order singularly perturbed form:

\[
L^\epsilon = \epsilon^{-2} L_0 + \epsilon^{-1} L_1 + L_2, \tag{3}
\]

where

\[
L_0 = \epsilon^i \partial y_i + \frac{1}{2} \alpha^{ij} \partial y_i y_j, \\
L_1 = f^i \partial x_i + g^i \partial y_i + h_{ij} \partial x_i y_j, \\
L_2 = b^i \partial x_i + \frac{1}{2} a_{ij} \partial x_i x_j.
\]

In the special case of \(f = 0, g = 0, h = 0\), the SDE (2) reduces to

\[
\begin{cases}
\quad dX^\epsilon_t = b(X^\epsilon_t, Y^\epsilon_t, t) dt + \sigma(X^\epsilon_t, Y^\epsilon_t, t) dW_t, \\
\quad dY^\epsilon_t = \epsilon^{-2} c(X^\epsilon_t, Y^\epsilon_t, t) dt + \epsilon^{-1} \eta(X^\epsilon_t, Y^\epsilon_t, t) dW_t.
\end{cases}
\]

In this case, the generator of \((X^\epsilon, Y^\epsilon)\) has the following first-order singularly perturbed form:

\[
L^\epsilon = \epsilon^{-2} L_0 + L_2.
\]

Compared with the first-order case, second-order singularly perturbed diffusions can deal with more complicated drift terms, as well as correlated noise terms between fast and slow components. In addition, most physics papers focus on Langevin equations with the diffusion matrix \(D\) being diagonal. Our setting can deal with the non-diagonal coupling case.

Treating \(x \in \mathbb{R}^m\) as a parameter, we next consider a family of \(\mathbb{R}^n\)-valued inhomogeneous diffusion processes \((Y^x)_x \in \mathbb{R}^m\) with generator \(L_0\), which solves the SDE

\[
dY^x_t = c(x, Y^x_t, t) dt + \eta(x, Y^x_t, t) dW_t.
\]

In this paper, we assume that the following regularity conditions are satisfied:

(a) The functions \(b, f, c, g, \sigma, \eta\) are bounded and smooth and all their partial derivatives are bounded.

(b) The diffusion matrix \(\alpha\) satisfies the uniformly elliptic condition, i.e. there exists \(\lambda > 0\) such that

\[
\xi^T \alpha(x, y, t) \xi \geq \lambda |\xi|^2, \quad \text{for any } \xi \in \mathbb{R}^n.
\]
(c) For any \( x \in \mathbb{R}^m \) and \( t \geq 0 \), the process \( Y^x \) has a pseudo-stationary density \( \rho = \rho(x, y, t) \) that is the solution of
\[
L_0^*\rho = 0, \quad \int_{\mathbb{R}^n} \rho(x, y, t) dy = 1,
\]
where \( L_0^* \) is the adjoint operator of \( L_0 \) with respect to the Lebesgue measure.

(d) For any \( x \in \mathbb{R}^m \) and \( t \geq 0 \), the function \( f \) satisfies the centering condition
\[
\int_{\mathbb{R}^n} f(x, y, t) \rho(x, y, t) dy = 0.
\]
In fact, the regularity condition (a) can be weakened by imposing several Lipschitz, Hölder, and growth conditions [60,61]. Since our main concern here is to understand the martingale structure of a general thermodynamic functional under model simplification, we do not pursue the weakest regularity conditions. Under the regularity conditions (a)–(c), it is a classical result that the pseudo-stationary density of \( Y^x \) must be unique, smooth, and positive for any \( x \in \mathbb{R}^m \) and \( t \geq 0 \) [7]. Moreover, under all the regularity conditions, it is also known that there exists a unique smooth solution \( \phi \) to the Poisson equation [60,61]
\[
\begin{aligned}
- L_0 \phi(x, y, t) &= f(x, y, t), \\
\int_{\mathbb{R}^n} \phi(x, y, t) \rho(x, y, t) dy &= 0.
\end{aligned}
\]

3 Martingale Structure for Forward Thermodynamic Functionals

In this section, we investigate the martingale structure for forward thermodynamic functionals of inhomogeneous singularly perturbed diffusion processes under second-order averaging. Before doing this, recall that for any \( \mathbb{R}^m \)-valued diffusion process \( Z = (Z_t)_{0 \leq t \leq T} \), the distribution (or the law) of \( Z \) is a probability measure \( P_Z \) on the path space \( C([0, T], \mathbb{R}^m) \) defined as
\[
P_Z(A) = \mathbb{P}(Z \in A),
\]
where \( A \subset C([0, T], \mathbb{R}^m) \) is an arbitrary Borel set.
In this paper, we define the forward thermodynamic functional for a diffusion process as the logarithmic Radon–Nikodym derivative between the laws of the original process and another one, which is called the comparable process. Specifically, let \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon) = (\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon)_{0 \leq t \leq T}\) be a comparable diffusion process solving the SDE

\[
\begin{align*}
\{ d\tilde{X}_t^\epsilon &= \tilde{b}(\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon, t) + \epsilon^{-1} f(\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon, t)dt + \sigma(\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon, t)dW_t, \\
\{ d\tilde{Y}_t^\epsilon &= \tilde{c}(\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon, t) + \epsilon^{-2} c(\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon, t)dt + \epsilon^{-1} \eta(\tilde{X}_t^\epsilon, \tilde{Y}_t^\epsilon, t)dW_t,
\end{align*}
\]

where \(\tilde{b} : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^m\) and \(\tilde{c} : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n\). The initial distribution of \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\) is chosen to be the same as that of \((X^\epsilon, Y^\epsilon)\). We also assume that the comparable process satisfies the regularity conditions (a)-(d). The forward thermodynamic functional for the original process is defined rigorously as follows.

**Definition 3.1** Let \(P_T^e\) be the law of the original process \((X^\epsilon, Y^\epsilon)\) and let \(\tilde{P}_T^e\) be the law of the comparable process \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\). Then the forward thermodynamic functional of \((X^\epsilon, Y^\epsilon)\) with respect to \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\) is defined as

\[
F_T^e = \log \frac{dP_T^e}{d\tilde{P}_T^e}(X^\epsilon, Y^\epsilon),
\]

if \(P_T^e\) and \(\tilde{P}_T^e\) are absolutely continuous with respect to each other.

The reason we use the same functions \(f, c, \sigma,\) and \(\eta\) in the comparable process instead of introducing new functions \(\hat{f}, \hat{c}, \hat{\sigma},\) and \(\hat{\eta}\) is to guarantee that (i) the Radon–Nikodym derivative defined in (8) exists and (ii) the forward thermodynamic functional \(F_T^e\) does not diverge as \(\epsilon \to 0\). Since the original and comparable processes are both singularly perturbed diffusions, it follows from Lemma 2.1 that \(X^\epsilon\) converges weakly in \(C([0, T], \mathbb{R}^m)\) to another diffusion \(X = (X_t)_{0 \leq t \leq T}\) with generator

\[
L = w^i(x, t)\partial_i + \frac{1}{2} A^{ij}(x, t)\partial_{ij},
\]

where \(w^i\) and \(A^{ij}\) are defined in Lemma 2.1, and \(\tilde{X}^\epsilon\) converges weakly in \(C([0, T], \mathbb{R}^m)\) to another diffusion \(\tilde{X} = (\tilde{X}_t)_{0 \leq t \leq T}\) with generator

\[
\tilde{L} = \tilde{w}^i(x, t)\partial_i + \frac{1}{2} \tilde{A}^{ij}(x, t)\partial_{ij},
\]

where

\[
\tilde{w}^i(x, t) = \int_{\mathbb{R}^n} (\tilde{b}^i + \partial_{x_j} \phi^i f^j + \partial_{y_j} \phi^i \tilde{c}^j + \partial_{x_i x_j} \phi^i \tilde{A}^{jk}) (x, y, t) \rho(x, y, t) dy.
\]

It is important to note that the two limiting processes, \(X\) and \(\tilde{X}\), have the same function \(\phi = (\phi^i)\) and the same diffusion matrix \(A = (A^{ij})\). This can be explained as follows. Since the original and comparable processes share the same functions \(f, c, \sigma,\) and \(\eta\), they also share the same generator \(L_0\). Hence the solution to the Poisson equation (6) for the comparable process is exactly the same as that for the original process. The forward thermodynamic function for the averaged process is defined rigorously as follows.

**Definition 3.2** Let \(P_T\) be the law of \(X\) and let \(\tilde{P}_T\) be the law of \(\tilde{X}\). Then the forward thermodynamic function of \(X\) with respect to \(\tilde{X}\) is defined as

\[
F_T^{(1)} = \log \frac{dP_T}{d\tilde{P}_T}(X),
\]
if $P_T$ and $\hat{P}_T$ are absolutely continuous with respect to each other.

**Remark 3.3** If we let the ending time $T$ change, then both $F^e_t$ and $F^{(1)}_t$ can be viewed as stochastic processes with $T$ being replaced by an arbitrary time $t$. Since $X^e \Rightarrow X$ in $C([0, T], \mathbb{R}^m)$, one may guess that the forward thermodynamic functional $F^e_t$ of the original process should converge weakly to the forward thermodynamic functional $F^{(1)}_t$ of the averaged process. However, recent studies in stochastic thermodynamics indicate that this is generally not true [5]. In fact, $F^e_t$ indeed converges weakly to a limit functional $F_t$, as stated in the following theorem. However, $F_t$ is in general not equal to $F^{(1)}_t$, which means that $F_t$ can be decomposed into two parts:

$$F_t = F^{(1)}_t + F^{(2)}_t.$$
In this section, we investigate the martingale structure for backward thermodynamic functionals of inhomogeneous singularly perturbed diffusion processes under second-order averaging. To avoid the initial layer effect, we assume that the original process \((X^ε, Y^ε)\) starts at \(t = -1\). While the original process starts at \(t = -1\), we will focus on the its dynamic behavior during \(t \in [0, T]\). The reason why we make this assumption is to guarantee the convergence of backward thermodynamic functionals under second-order averaging. Under this assumption, there is no need to impose additional requirements on the initial distribution of the original process at time 0.

In this paper, we define the backward thermodynamic functional for a diffusion process as the logarithmic Radon–Nikodym derivative between the laws of the original process and the comparable process. Specifically, let \((X_{R}^{R,e}, Y_{R}^{R,e}) = (\tilde{X}_{t}^{R,e}, \tilde{Y}_{t}^{R,e})_{0 \leq t \leq T}\) be the comparable diffusion process solving the SDE
\[
\begin{align*}
\frac{d\tilde{X}_{t}^{R,e}}{dt} &= (\tilde{b} + \epsilon^{-1} \tilde{f})(\tilde{X}_{t}^{R,e}, \tilde{Y}_{t}^{R,e}, T - t)dt + \sigma(\tilde{X}_{t}^{R,e}, \tilde{Y}_{t}^{R,e}, T - t)dW_{t}, \\
\frac{d\tilde{Y}_{t}^{R,e}}{dt} &= (\epsilon^{-1} \tilde{g} + \epsilon^{-2} \epsilon)(\tilde{X}_{t}^{R,e}, \tilde{Y}_{t}^{R,e}, T - t)dt + \epsilon^{-1} \eta(\tilde{X}_{t}^{R,e}, \tilde{Y}_{t}^{R,e}, T - t)dW_{t},
\end{align*}
\]
where \(\tilde{b}, \tilde{f} : \mathbb{R}^{m} \times \mathbb{R}^{n} \times \mathbb{R}^{+} \to \mathbb{R}^{m}\) and \(\tilde{g} : \mathbb{R}^{m} \times \mathbb{R}^{n} \times \mathbb{R}^{+} \to \mathbb{R}^{n}\). Unlike the forward case, the comparable process in the backward case is controlled by a reversed protocol since the temporal variable in the drift and diffusion coefficients is \(T - t\) rather than \(t\). The initial distribution of \((\tilde{X}_{R}^{R,e}, \tilde{Y}_{R}^{R,e})\) (at time 0) is chosen to be the same as the final distribution of \((X^{ε}, Y^{ε})\) (at time \(T\)). We also assume that the comparable process satisfies the regularity conditions (a)-(d). The backward thermodynamic functional for the original process is defined rigorously as follows.

**Definition 4.1** Let \(P_{T}^{ε}\) be the law of the original process \((X^{ε}, Y^{ε})\), let \(\tilde{P}_{T}^{R,e}\) be the law of the comparable process \((\tilde{X}_{R}^{R,e}, \tilde{Y}_{R}^{R,e})\), and let \(\tilde{Q}_{T}^{R,e}\) be the law of its time reversal \((\tilde{X}_{T-t}^{R,e}, \tilde{Y}_{T-t}^{R,e})_{0 \leq t \leq T}\). Then the backward thermodynamic functional of \((X^{ε}, Y^{ε})\) with respect to \((\tilde{X}_{R}^{R,e}, \tilde{Y}_{R}^{R,e})\) is defined as
\[
G_{T}^{ε} = \log \frac{dP_{T}^{ε}}{d\tilde{Q}_{T}^{R,e}}(X^{ε}, Y^{ε}),
\] if \(P_{T}^{ε}\) and \(\tilde{Q}_{T}^{R,e}\) are absolutely continuous with respect to each other.

Since the comparable process satisfies the regularity conditions (a)-(d), there exists a unique smooth solution \(\dot{\phi}\) to the Poisson equation
\[
\begin{align*}
- L_{0}\dot{\phi}(x, y, t) &= \tilde{f}(x, y, t), \\
\int_{\mathbb{R}^{n}} \dot{\phi}(x, y, t) \rho(x, y, t) dy &= 0.
\end{align*}
\]
It then follows from Lemma 2.1 that the slow component \(\tilde{X}_{R}^{R,e}\) of the comparable process converges weakly in \(C([0, T], \mathbb{R}^{m})\) to another diffusion \(\tilde{X}^{R} = (\tilde{X}_{t}^{R})_{0 \leq t \leq T}\) with generator
\[
\tilde{L} = \tilde{w}^{i}(x, T - t) \partial_{x_{i}} + \frac{1}{2} \tilde{A}^{ij}(x, T - t) \partial_{x_{i}x_{j}},
\]
where
\[
\tilde{w}^{i}(x, t) = \int_{\mathbb{R}^{n}} (\tilde{b}^{i} + \partial_{x_{j}} \tilde{b}^{i} \tilde{f}^{j} + \partial_{y_{j}} \tilde{f}^{i} \tilde{g}^{j} + \partial_{x_{j}x_{k}} \tilde{\phi}^{i} \tilde{h}^{jk})(x, y, t) \rho(x, y, t) dy.
\]
\[
\tilde{A}^{ij}(x, t) = \int_{\mathbb{R}^n} (\alpha^{ij} + \tilde{\phi}^{ij} \tilde{f}^{ij} + \tilde{\phi}^{ji} \tilde{f}^{ji} + \partial_x \tilde{\phi}^{ij} h^{ik} + \partial_x \tilde{\phi}^{ji} h^{ik})(x, y, t) \rho(x, y, t) dy.
\]

The backward thermodynamic functional for the averaged process is defined rigorously as follows.

**Definition 4.2** Let \( P_T \) be the law of \( X \), let \( \tilde{P}_T^R \) be the law of \( \tilde{X}_R \), and let \( \tilde{Q}_T^R \) be the law of its time reversal \( (\tilde{X}_R^T)_{0 \leq t \leq T} \). Then the backward thermodynamic function of \( X \) with respect to \( \tilde{X}_R \) is defined as

\[
G^{(1)}_T = \log \frac{dP_T}{d\tilde{Q}_T^R}(X.),
\]

if \( P_T \) and \( \tilde{Q}_T^R \) are absolutely continuous with respect to each other.

Similarly, if we let the ending time \( T \) change, then both \( G^\epsilon_T \) and \( G^{(1)}_t \) can be viewed as stochastic processes with \( T \) being replaced by an arbitrary time \( t \). In the backward case, we need the following two important assumptions.

**Assumption 4.3** The function \( \tilde{f} \) satisfies the compatible condition

\[
f + \tilde{f} = \nabla_x \cdot h + h \nabla \log \rho,
\]

where \( \nabla_x \cdot h \) is a vector-valued function whose \( i \)th component is given by \( (\nabla_x \cdot h)^i = \partial_{xj} h^{ij} \).

**Assumption 4.4** The pseudo-stationary density \( \rho \) satisfies the compatible condition

\[
2c = \nabla_x \cdot \alpha + \alpha \nabla \log \rho,
\]

where \( \nabla_x \cdot \alpha \) is a vector-valued function whose \( i \)th component is given by \( (\nabla_x \cdot \alpha)^i = \partial_{xj} \alpha^{ij} \).

Interestingly, the above two assumptions guarantee that the two limiting processes, \( X \) and \( \tilde{X}_R \), have the same functional forms of the diffusion coefficients (see Proposition 8.7 for details). Roughly speaking, Assumption 4.4 holds if and only if the diffusion \( Y^x \) is reversible during each infinitesimal time period. It is a classical result that this assumption has the following equivalent statement [7].

**Lemma 4.5** Assumption 4.4 holds if and only if the operator \( L_0 \) is symmetric with respect to the pseudo-stationary density \( \rho \), i.e.

\[
(L_0 f, g)_\rho = (f, L_0 g)_\rho, \quad f, g \in C_b^\infty(\mathbb{R}^n),
\]

where \( C_b^\infty(\mathbb{R}^n) \) is the space of all bounded smooth functions on \( \mathbb{R}^n \) whose all partial derivatives are bounded and

\[
(f, g)_\rho := \int_{\mathbb{R}^n} f(y)g(y)\rho(x, y, t) dy.
\]

**Remark 4.6** From the proof in Sect. 8, especially the expression (17), it is easy to see that the above two assumptions are the sufficient and necessary conditions that guarantee the backward thermodynamic function \( G^\epsilon_T \) not to diverge as \( \epsilon \to 0 \). If \( h = 0 \) (which means that the noise terms of fast and slow components are not correlated), then Assumption 4.3 reduces to \( \tilde{f} = -f \).

We emphasize that in this section, our theory is established when all slow and fast components of the original process are “even variables” (this concept will be explained in Sect. 6). However, in many physical systems, some variables of the original processes are actually “odd variables” and thus the function \( \tilde{f} \) should be replaced by another function \( \tilde{f}_\delta \) (see Sect. 6).
for details). For the classical model of Langevin dynamics, all slow components (position variables) are even variables and all fast components (velocity variables) are odd variables. In this case, we have $h = 0$ and $\tilde{f}_S = -f$, and thus Assumption 4.3 is automatically satisfied. This indicates that the above two assumptions are not excessive requirements for real physical systems.

For any $\epsilon > 0$ and any Borel set $K \subset \mathbb{R}^{n+m}$, let $\tau^\epsilon_K = \inf\{t \geq 0 : (X^\epsilon_t, Y^\epsilon_t) \notin K\}$ be the first exit time of $(X^\epsilon, Y^\epsilon)$ from $K$. We still need the following technical assumption.

**Assumption 4.7** For any $\delta > 0$, there exists a compact set $K \subset \mathbb{R}^{n+m}$ such that

\[
P\left(\inf_{\epsilon > 0} \tau^\epsilon_K \geq T\right) \geq 1 - \delta.
\]

**Remark 4.8** We emphasize that the above assumption is needed because we have used the matched asymptotic expansion theory of singularly perturbed diffusion processes [48,49,62] when proving the weak convergence of the backward thermodynamic functional in the path space and such theory is only developed for diffusion processes on compact manifolds. If the matched asymptotic expansion theory theory can be generalized to non-compact manifolds, then the above assumption may be relaxed; however, this is beyond the scope of the present paper. In Appendix A, we give an example showing that the above assumption is satisfied by a wide class of processes and thus is not too strong.

The martingale structure for backward thermodynamic functionals under second-order averaging is stated as follows.

**Theorem 4.9** Under Assumptions 4.3, 4.4, and 4.7, the following conclusions hold:

(a) For any $t \geq 0$, both $G^\epsilon_t$ and $G^{(1)}_t$ are well defined.
(b) The pair of processes $(X^\epsilon_t, G^\epsilon_t)_{0 \leq t \leq T}$ converges weakly to another pair of processes $(X_t, G_t)_{0 \leq t \leq T}$ in $C([0, T], \mathbb{R}^{m+1})$.
(c) If we decompose the limit functional into $G_t = G^{(1)}_t + G^{(2)}_t$, then $e^{-G^{(2)}_t}$ is a martingale.

**Proof** The proof of this theorem will be given in Sect. 8. \(\square\)

Unlike the forward case, the backward thermodynamic functional $G^\epsilon_t$, the limit functional $G_t$, and its regular part $G^{(1)}_t$ are generally not martingales. In both the forward and backward cases, a general thermodynamic functional has a weak limit under second-order averaging, which can be decomposed into the regular part and the anomalous part. Theorems 3.4 and 4.9 reveal a highly nontrivial relationship between model simplification and martingale. In the forward case, the limit functional and its two parts are all martingales. In the backward case, while the limit functional and its regular part may not be martingales, the anomalous part is still a martingale. Therefore, our theory indicates that the anomalous part of a general thermodynamic functional is always a martingale under model simplification.

**Remark 4.10** In this paper, our theorems are proved in a very general framework, which generalizes the setting in the physical literature in many ways. First, we reveal the martingale structure for a general thermodynamic functional defined as the logarithmic Radon–Nykodim derivative between the laws of the original process and a comparable process (or its time reversal) whose drift terms $\hat{b}$ and $\hat{g}$ (or $\tilde{b}$ and $\tilde{g}$) can be arbitrarily chosen, rather than focusing on specific thermodynamic functionals such as entropy production and housekeeping heat. It can be seen from the proof that the arbitrariness of the drift terms in the comparable process
greatly enhances the theoretical complexity. Second, we investigate model simplification under second-order averaging, which is physically more relevant and mathematically much more general and complicated than first-order averaging. In statistical physics, many important models such as the Langevin dynamics in the zero-mass limit are actually second-order singularly perturbed diffusions. Third, most previous papers focused on the case where the diffusion matrix $D$ is diagonal, while our theory can deal with the non-diagonal coupling case. In particular, our model covers the case of $h \neq 0$, which describes correlated noise terms between slow and fast components (such effect is particularly important in biochemical reaction networks).

**Corollary 4.11** (Integral fluctuation theorem) The backward functionals $G_t^r, G_t^{(1)}$, and $G_t^{(2)}$ satisfy the following integral fluctuation theorem:

(a) For any $t \geq 0$, we have

\[ \mathbb{E} e^{-G_t^r} = \mathbb{E} e^{-G_t^{(1)}} = 1. \]

(b) For any bounded stopping time $\tau$ with respect to $\{\mathcal{F}_t\}$, we have

\[ \mathbb{E} e^{-G_t^{(2)}} = 1. \]

**Proof** Since $e^{-G_t^{(2)}}$ is a martingale, part (b) follows from the optional sampling theorem. We next prove part (a). By Definition 4.1, we have

\[ \mathbb{E} e^{-G_t^r} = \mathbb{E} \frac{d \tilde{Q}_{t}^{R,R}}{d P_t^{\epsilon}}(X^\epsilon, Y^\epsilon) = \int_{C([0,t],\mathbb{R}^{m+n})} \frac{d \tilde{Q}_{t}^{R,R}}{d P_t^{\epsilon}}(w) d P_t^{\epsilon}(w) = 1. \]

Similarly, the equality $\mathbb{E} e^{-G_t^{(1)}} = 1$ can be proved by using Definition 4.2. \qed

Unlike the forward case, the limit functional $G_t$ in the backward case may not satisfy the integral fluctuation theorem because the family of random variables $(e^{-G_t^r})_{\epsilon>0}$ may not be uniformly integrable.

### 5 Physical Applications in the Absence of Odd Variables

In stochastic thermodynamics, most thermodynamic quantities can be represented as forward or backward thermodynamic functionals. Therefore, our general theory can be applied to understand the martingale structure and prove the integral fluctuation theorem for such thermodynamic quantities under second-order averaging. Two of the most important thermodynamic quantities among them are entropy production and housekeeping heat, which characterize the irreversibility of a molecular system. To define them rigorously, we need the following concepts.

#### 5.1 Some Concepts

In stochastic thermodynamics, there are many important processes associated with a given diffusion process $Z$. Specifically, let $Z = (Z_t)_{0 \leq t \leq T}$ be an $\mathbb{R}^m$-valued inhomogeneous diffusion process with generator

\[ A_t = b^i(x,t) \partial_i + \frac{1}{2} a^{ij}(x,t) \partial_{ij}. \]
For simplicity, we assume that the drift \( b = (b^i) \) and diffusion matrix \( a = (a^{ij}) \) are bounded and smooth, all their partial derivatives are bounded, and \( a \) is uniformly elliptic.

**Definition 5.1** The diffusion process with generator \( \mathcal{A}_{T-t} \) and with its initial distribution being chosen as the final distribution of \( Z \) (at time \( T \)) is called the original process \( Z \) under the reversed protocol.

For each \( t \), let \( p(x, t) \) be the probability density of \( Z_t \) and let \( \mathcal{B}_t \) be the adjoint operator of \( \mathcal{A}_t \) with respect to the probability density \( p(x, t) \) satisfying

\[
(\mathcal{A}_t f, g)_p = (f, \mathcal{B}_t g)_p, \quad f, g \in C^\infty_b(\mathbb{R}^m).
\]

It is easy to check that \( \mathcal{B}_t \) is given by

\[
\mathcal{B}_t = (-b^i + \partial_j a^{ij} + a^{ij} \log p) \partial_i + \frac{1}{2} a^{ij} \partial_{ij}.
\]

Clearly, \( \mathcal{B}_t \) is the generator of an inhomogeneous diffusion process with drift \(-b + \nabla \cdot a + a \nabla \log p\) and diffusion matrix \( a \). The following lemma is a classical result [7].

**Lemma 5.2** The time reversal \( (Z_{T-t})_{0 \leq t \leq T} \) of \( Z \) is a diffusion process with generator \( \mathcal{B}_{T-t} \).

For each \( t \), the pseudo-stationary density \( \mu(x, t) \) of \( Z \) is defined as the solution to the equation

\[
\mathcal{A}_t^* \mu = 0, \quad \int_{\mathbb{R}^m} \mu(x, t) dx = 1,
\]

where

\[
\mathcal{A}_t^* = (\partial_j a^{ij} - b^i) \partial_i + \frac{1}{2} a^{ij} \partial_{ij} + \frac{1}{2} \partial_i a^{ij} - \partial_i b^i
\]

is the adjoint operator of \( \mathcal{A}_t \) with respect to the Lebesgue measure. Suppose that the pseudo-stationary density exists and is unique for each \( t \). Moreover, let \( \mathcal{A}_t^\dagger \) be the adjoint operator of \( \mathcal{A}_t \) with respect to the pseudo-stationary density \( \mu(x, t) \) satisfying

\[
(\mathcal{A}_t f, g)_\mu = (f, \mathcal{A}_t^\dagger g)_\mu, \quad f, g \in C^\infty_b(\mathbb{R}^m).
\]

It is easy to check that \( \mathcal{A}_t^\dagger \) is given by

\[
\mathcal{A}_t^\dagger = (-b^i + \partial_j a^{ij} + a^{ij} \partial_j \log \mu) \partial_i + \frac{1}{2} a^{ij} \partial_{ij}.
\]

Clearly, \( \mathcal{A}_t^\dagger \) is the generator of an inhomogeneous diffusion process with drift \(-b + \nabla \cdot a + a \nabla \log \mu\) and diffusion matrix \( a \).

**Definition 5.3** The diffusion process with generator \( \mathcal{A}_t^\dagger \) and with its initial distribution being chosen as the initial distribution of \( Z \) is called the adjoint process of \( Z \).

It is easy to see that when \( Z \) is homogeneous and stationary, its time-reversed process and adjoint process are exactly the same.

### 5.2 Martingale Structure for Entropy Production

In statistical physics, the deviation of a molecular system from equilibrium is usually characterized by the concept of entropy production. From the trajectory perspective, the entropy production of the original and reduced processes is defined rigorously as follows [7,22].
**Definition 5.4** Let $P^\epsilon_T$ be the law of the original process $(X^\epsilon, Y^\epsilon)$ and let $Q^{R,\epsilon}_T$ be the law of the time reversal of $(X^\epsilon, Y^\epsilon)$ under the reversed protocol. Then the entropy production of the original process is defined as

$$S^\epsilon_{\text{tot}}(T) = \log \frac{dP^\epsilon_T}{dQ^{R,\epsilon}_T}(X^\epsilon, Y^\epsilon),$$

if $P^\epsilon_T$ and $Q^{R,\epsilon}_T$ are absolutely continuous with respect to each other. The entropy production $S^{(1)}_{\text{tot}}(T)$ of the averaged process $X$ can be defined similarly.

In fact, under the above definition, the entropy production of the original process has the following explicit expression (see Sect. 8 for the proof):

$$S^\epsilon_{\text{tot}}(T) = \log \frac{p^\epsilon(T_0, 0)}{p^\epsilon(T_0, T)} + 2 \int_0^T (B - \tilde{B})' D^{-1}(Z^\epsilon_s, s) dZ^\epsilon_s$$

$$- 2 \int_0^T (B - \tilde{B})' D^{-1} \tilde{B}(Z^\epsilon_s, s) ds + \int_0^T \nabla \cdot (B - \tilde{B})(Z^\epsilon_s, s) ds,$$

where $p^\epsilon(x, y, t)$ is the probability density of $(X^\epsilon_t, Y^\epsilon_t)$ for each $t$, $Z^\epsilon = (X^\epsilon, Y^\epsilon)'$, and

$$\tilde{B} = \frac{1}{2} \left( \frac{\nabla_y \cdot a + \epsilon^{-1} \nabla_x \cdot h}{\epsilon^{-1} \nabla_x \cdot h + \epsilon^{-2} \nabla_y \cdot a} \right).$$

In the special case that the diffusion matrix $D$ is a constant diagonal matrix, we have $\tilde{B} = 0$ and thus the entropy production reduces to

$$S^\epsilon_{\text{tot}}(T) = \log \frac{p^\epsilon(Z_0^\epsilon, 0)}{p^\epsilon(Z_0^\epsilon, T)} + 2 \int_0^T B' D^{-1}(Z^\epsilon_s, s) dZ^\epsilon_s + \int_0^T \nabla \cdot B(Z^\epsilon_s, s) ds$$

$$= \log \frac{p^\epsilon(Z_0^\epsilon, 0)}{p^\epsilon(Z_0^\epsilon, T)} + 2 \int_0^T B' D^{-1}(Z^\epsilon_s, s) dZ^\epsilon_s,$$

where the integral in the last equality is the Stratonovich integral. Note that this is exactly the entropy production defined in most physical literature [22], where the first part in the last equality is the increase in entropy of the system and the second part is the entropy production in the surrounding medium.

Under the reversed protocol, $(X^\epsilon, Y^\epsilon)$ becomes a diffusion process with drift $B(x, y, T - t)$ and diffusion matrix $D(x, y, T - t)$. In other words, it is a diffusion process solving the SDE

$$\begin{align*}
  dX^R_t &= (b + \epsilon^{-1} f)(X^R_t, Y^R_t, T - t) dt + \sigma(X^R_t, Y^R_t, T - t) dW_t, \\
  dY^R_t &= (\epsilon^{-1} g + \epsilon^{-2} c)(X^R_t, Y^R_t, T - t) dt + \epsilon^{-1} \eta(X^R_t, Y^R_t, T - t) dW_t.
\end{align*}$$

This is a special case of the comparable process (9) in the backward case with $\tilde{b} = b$, $\tilde{f} = f$, and $\tilde{g} = g$. Therefore, entropy production is a kind of backward thermodynamic functional. The next theorem then follows directly from Theorem 4.9.

**Theorem 5.5** Suppose that Assumptions 4.4 and 4.7 and the following compatible condition are satisfied:

$$2f = \nabla_y \cdot h + h \nabla_y \log \rho. \quad (10)$$

Then the following conclusions hold:

(a) For any $t \geq 0$, both $S^e_{\text{tot}}(t)$ and $S^{(1)}_{\text{tot}}(t)$ are well defined.
(b) The pair of processes \((X_t^\epsilon, S_{tot}^\epsilon(t))_{0 \leq t \leq T}\) converges weakly in \(C([0, T], \mathbb{R}^{m+1})\) to another pair of processes \((X_t, S_{tot}(t))_{0 \leq t \leq T}\).

(c) If we decompose the limit functional into \(S_{tot}(t) = S_{tot}^{(1)}(t) + S_{tot}^{(2)}(t)\), then \(e^{-S_{tot}^{(2)}(t)}\) is a martingale.

(d) For any \(t \geq 0\) and bounded stopping time \(\tau\) with respect to \(\mathcal{F}_t\), we have the following integral fluctuation theorem:

\[\mathbb{E}e^{-S_{tot}(\tau)} = \mathbb{E}e^{-S_{tot}^{(1)}(\tau)} = \mathbb{E}e^{-S_{tot}^{(2)}(\tau)} = 1.\]

**Remark 5.6** Recently, it has been reported that when the original process is homogeneous and stationary, its entropy production is a martingale [38]. Here we prove that under model simplification, the anomalous part of entropy production is always a martingale and hence satisfies the integral fluctuation theorem, whether the original process is homogeneous and stationary or not.

### 5.3 Martingale Structure for Housekeeping Heat

From the trajectory perspective, the housekeeping heat of the original and reduced processes is defined rigorously as follows.

**Definition 5.7** Let \(P^\epsilon_t\) be the law of the original process \((X^\epsilon, Y^\epsilon)\) and let \(\hat{P}^\epsilon_t\) be the law of its adjoint process. Then the housekeeping heat, also called adiabatic entropy production, of the original process is defined as

\[S^\epsilon_{hk}(T) = \log \frac{dP^\epsilon_T}{d\hat{P}^\epsilon_T}(X^\epsilon, Y^\epsilon),\]

if \(P^\epsilon_t\) and \(\hat{P}^\epsilon_t\) are absolutely continuous with respect to each other. The housekeeping heat \(S_{hk}^{(1)}(T)\) of the averaged process \(X\) can be defined similarly.

For each \(t\), let \(\mu^\epsilon(x, y, t)\) be the pseudo-stationary density of \((X^\epsilon, Y^\epsilon)\), which is assumed to exist and be unique. As discussed earlier, the adjoint process of \((X^\epsilon, Y^\epsilon)\) is a diffusion process with drift

\[\hat{B}(x, y, t) = (-B + \nabla \cdot D + D \nabla \log \mu^\epsilon)(x, y, t) = \begin{pmatrix} \hat{b}(x, y, t) + \epsilon^{-1} \hat{f}(x, y, t) \\ \epsilon^{-1} \hat{g}(x, y, t) + \epsilon^{-2} \hat{c}(x, y, t) \end{pmatrix},\]

and diffusion matrix \(D(x, y, t)\), where

\[
\begin{align*}
\hat{b} &= -b + \nabla_x \cdot \alpha + a \nabla_x \log \mu^\epsilon, \\
\hat{f} &= -f + \nabla_y \cdot h + h \nabla_y \log \mu^\epsilon, \\
\hat{g} &= -g + \nabla_x \cdot h' + h' \nabla_x \log \mu^\epsilon, \\
\hat{c} &= -c + \nabla_y \cdot \alpha + \alpha \nabla_y \log \mu^\epsilon.
\end{align*}
\]

In other words, it is a diffusion process solving the SDE

\[
\begin{align*}
&d\hat{X}^\epsilon_t = [\hat{b}(\hat{X}^\epsilon_t, \hat{Y}^\epsilon_t, t) + \epsilon^{-1} \hat{f}(\hat{X}^\epsilon_t, \hat{Y}^\epsilon_t, t)]dt + \sigma(\hat{X}^\epsilon_t, \hat{Y}^\epsilon_t, t)dW_t, \\
&d\hat{Y}^\epsilon_t = [\epsilon^{-1} \hat{g}(\hat{X}^\epsilon_t, \hat{Y}^\epsilon_t, t) + \epsilon^{-2} \hat{c}(\hat{X}^\epsilon_t, \hat{Y}^\epsilon_t, t)]dt + \epsilon^{-1} \eta(\hat{X}^\epsilon_t, \hat{Y}^\epsilon_t, t)dW_t.
\end{align*}
\]
In fact, under the above definition, the housekeeping heat of the original process has the following explicit expression (see Sect. 7 for the proof):

\[ S_{hk}^\epsilon(T) = \int_0^T (B - \hat{B})' D^{-1}(Z_s^\epsilon, s) dZ_s^\epsilon - \frac{1}{2} \int_0^T (B - \hat{B})' D^{-1}(B + \hat{B})(Z_s^\epsilon, s) ds \]

where \( Z_s^\epsilon = (X_s^\epsilon, Y_s^\epsilon)' \). In the special case that the diffusion matrix \( D \) is a constant diagonal matrix, we have \( \nabla \cdot D = 0 \) and thus the housekeeping heat reduces to

\[ S_{hk}^\epsilon(T) = \int_0^T (2B - D \nabla \log \mu^\epsilon)' D^{-1} (Z_s^\epsilon, s) dZ_s^\epsilon - \frac{1}{2} \int_0^T (2B - D \nabla \log \mu^\epsilon)' \nabla \log \mu^\epsilon(Z_s^\epsilon, s) ds \]

\[ = \int_0^T (2B - D \nabla \log \mu^\epsilon)' D^{-1} (Z_s^\epsilon, s) \circ dZ_s^\epsilon. \]

Note that this is exactly the housekeeping heat, i.e. the heat needed to maintain the violation of detailed balance, defined in most physical literature [78].

Under mild conditions, by using the matched asymptotic expansions of singularly perturbed diffusion processes [48,49,62], it can be proved that

\[ \mu^\epsilon(x, y, t) = \mu(x, t) \rho(x, y, t) + O(\epsilon), \]

\[ \nabla_x \log \mu^\epsilon(x, y, t) = \nabla_x \log \mu(x, t) + \nabla_x \log \rho(x, y, t) + O(\epsilon), \]

\[ \nabla_y \log \mu^\epsilon(x, y, t) = \nabla_y \log \rho(x, y, t) + O(\epsilon), \]

where \( \mu(x, t) \) is the pseudo-stationary density of the averaged process \( X \) which is also assumed to exist and be unique for each \( t \). Therefore, the drift coefficients have the following asymptotic expansions:

\[ \hat{b} = -b + \nabla_x \cdot a + a \nabla_x \log \mu + a \nabla_x \log \rho + O(1), \]

\[ \hat{f} = -f + \nabla_y \cdot h + h \nabla_y \log \rho, \]

\[ \hat{g} = -g + \nabla_x \cdot h' + h' \nabla_x \log \mu + h' \nabla_x \log \rho + O(1), \]

\[ \hat{c} = -c + \nabla_y \cdot \alpha + \alpha \nabla_y \log \rho, \]

where we have moved the \( O(\epsilon) \) terms of \( \hat{f} \) and \( \hat{c} \) into \( \hat{b} \) and \( \hat{g} \), respectively. If Equation (10) and Assumption 4.4 are satisfied, and if we ignore the \( O(1) \) terms of \( \hat{b} \) and \( \hat{g} \) (see Remark 3.5 for its justification), then we have \( \hat{f} = f \) and \( \hat{c} = c \), and thus the adjoint process of \((X^\epsilon, Y^\epsilon)\) is a special case of the comparable process (7) in the forward case. Therefore, housekeeping heat is a kind of forward thermodynamic functional. The next theorem then follows directly from Theorem 3.4.

**Theorem 5.8** Suppose that Assumption 4.4 and the following compatible condition are satisfied:

\[ 2f = \nabla_y \cdot h + h \nabla_y \log \rho. \]

Then the following conclusions hold:

(a) For any \( t \geq 0 \), both \( S_{hk}^\epsilon(t) \) and \( S_{hk}^{(1)}(t) \) are well defined.

(b) The pair of processes \((X_t^\epsilon, S_{hk}^\epsilon(t))_{0 \leq t \leq T}\) converges weakly in \( C([0, T], \mathbb{R}^{m+1}) \) to another pair of processes \((X_t, S_{hk}(t))_{0 \leq t \leq T}\).

(c) Both \( e^{-S_{hk}^\epsilon(t)} \) and \( e^{-S_{hk}(t)} \) are martingales.
(d) If we decompose the limit functional into \( S_{hk}(t) = S_{hk}^{(1)}(t) + S_{hk}^{(2)}(t) \), then \( e^{-S_{hk}^{(1)}(t)} \) and \( e^{-S_{hk}^{(2)}(t)} \) are orthogonal martingales.

(e) For any bounded stopping time \( \tau \) with respect to \( \{ \mathcal{F}_t \} \), we have the following integral fluctuation theorem:

\[
E e^{-S_{hk}(\tau)} = E e^{-S_{hk}^{(1)}(\tau)} = E e^{-S_{hk}^{(2)}(\tau)} = 1.
\]

Remark 5.9 Recently, it has been reported that the housekeeping heat of the original process is a martingale [40]. Here we prove that under model simplification, the regular and anomalous parts of housekeeping heat are always orthogonal martingales and hence satisfy the integral fluctuation theorem.

6 Physical Applications in the Presence of Odd Variables

6.1 General Theorem

In statistical physics, physical quantities are often divided into even and odd variables. For example, position, kinetic energy, and potential energy are even variables, while velocity and Lorentz force are odd variables. Intuitively, if a smooth function \( z(t) \) describes the trajectory of a classical particle, then its time reversal should be \( z(T - t) \). The velocity of the original trajectory is \( \dot{z}(t) \) and the velocity of the time-reversed trajectory is \( \dot{z}(T - t) \) rather than \( \dot{z}(T - t) \). Such kind of variables are called odd variables. Specifically, we give the following definition.

Definition 6.1 Let \( Z = (Z_t)_{0 \leq t \leq T} \) be a one-dimensional stochastic process.

(a) If the time reversal of \( Z \) is defined as \( (Z_{T-t})_{0 \leq t \leq T} \), then \( Z \) is called an even variable.

(b) If the time reversal of \( Z \) is defined as \( (-Z_{T-t})_{0 \leq t \leq T} \), then \( Z \) is called an odd variable.

Since the definition of forward thermodynamic functionals is irrelevant of time reversal, the results in the forward case will not change when odd variables are considered. Unlike the forward case, the definition of backward thermodynamic functionals depends on the time reversal of the comparable process. Therefore, all the results in the backward case discussed earlier only hold for systems whose components are all even variables and thus should be modified in the presence of odd variables.

From now on, we focus on the general case when each component of the original process \((X^\epsilon, Y^\epsilon)\) can be either an even or an odd variable. In the sense of Definition 6.1, the time reversal of the original process will be denoted by \((\delta X^\epsilon_{T-t}, \delta Y^\epsilon_{T-t})_{0 \leq t \leq T}\), where

\[
\delta = (\delta x_1, \ldots, \delta x_m, \delta y_1, \ldots, \delta y_n)
\]

with each component of \( \delta \) taking the value of 1 or \(-1\), corresponding to an even or an odd variable respectively.

In the backward case, let \((\tilde{X}^{R,\epsilon}, \tilde{Y}^{R,\epsilon})\) be the comparable process given in (9). Let \( P_T^\epsilon \) be the law of the original process \((X^\epsilon, Y^\epsilon)\) and let \( \tilde{Q}_T^{R,\epsilon} \) be the law of the time reversal of the comparable process in the sense of Definition 6.1. In other words, \( \tilde{Q}_T^{R,\epsilon} \) is the law of the time
reversal of $(\delta \tilde{X}^{R,e}, \delta \tilde{Y}^{R,e})$ in the usual sense. The backward thermodynamic functional of $(X^e, Y^e)$ with respect to $(\tilde{X}^{R,e}, \tilde{Y}^{R,e})$ is still defined as

$$G_T^e = \log \frac{dP_T^e}{d\tilde{Q}_T^{R,e}}(X^e, Y^e).$$

It then follows from Lemma 2.1 that $X^e \Rightarrow X$ and $\tilde{X}^{R,e} \Rightarrow \tilde{X}^R$ in $C([0, T], \mathbb{R}^m)$ as $\epsilon \to 0$. Similarly, let $P_T$ be the law of the averaged process $X$ and let $\tilde{Q}_T^{R,e}$ be the law of the time reversal of $\tilde{X}^R$ in the sense of Definition 6.1. The backward thermodynamic function of $X$ with respect to $\tilde{X}^R$ is still defined as

$$G_T^{(1)} = \log \frac{dP_T}{d\tilde{Q}_T}(X).$$

Clearly, $(\delta \tilde{X}^{R,e}, \delta \tilde{Y}^{R,e})$ is a diffusion process with drift

$$\tilde{B}_\delta = \left( \begin{array}{c} \delta_{xy}(x, y, T - t) + \epsilon^{-1} \delta_{xy}(x, y, T - t) \\ \epsilon^{-1} \tilde{c}_\delta(x, y, T - t) + \epsilon^{-2} \tilde{c}_\delta(x, y, T - t) \end{array} \right),$$

and diffusion matrix

$$D_\delta = \left( \begin{array}{cc} a_\delta(x, y, T - t) & \epsilon^{-1} h_\delta(x, y, T - t) \\ \epsilon^{-1} h_\delta'(x, y, T - t) & \epsilon^{-2} a_\delta(x, y, T - t) \end{array} \right),$$

where

$$\tilde{b}^i_\delta(x, y, t) = \delta_{xi} \tilde{b}^i(\delta x, \delta y, t), \quad \tilde{f}^i_\delta(x, y, t) = \delta_{xi} \tilde{f}^i(\delta x, \delta y, t),$$

$$\tilde{g}^i_\delta(x, y, t) = \delta_{yi} \tilde{g}^i(\delta x, \delta y, t), \quad \tilde{c}^i_\delta(x, y, t) = \delta_{yi} \tilde{c}^i(\delta x, \delta y, t),$$

$$a^{ij}_\delta(x, y, t) = \delta_{xi} \delta_{xj} a^{ij}(\delta x, \delta y, t),$$

$$h^{ij}_\delta(x, y, t) = \delta_{xi} \delta_{yj} h^{ij}(\delta x, \delta y, t),$$

$$a^{ij}_\delta(x, y, t) = \delta_{yi} \delta_{xj} \alpha^{ij}(\delta x, \delta y, t).$$

To make $(\delta \tilde{X}^{R,e}, \delta \tilde{Y}^{R,e})$ have the form of the comparable process (9), we need to impose the following compatible conditions.

**Assumption 6.2**

$$c_\delta = c, \quad a_\delta = a, \quad h_\delta = h, \quad \alpha_\delta = \alpha.$$

In fact, this assumption guarantees that the laws $P_T^e$ and $\tilde{Q}_T^{R,e}$ are absolutely continuous with respect to each other. In most cases of physical interests, we have $h = 0$ (which means that the noise terms of fast and slow components are not correlated), the slow components are all even variables, and the fast components are all odd variables [57–59]. In this case, we automatically have

$$a_\delta = a, \quad h_\delta = h, \quad \alpha_\delta = \alpha. \quad (11)$$

**Theorem 6.3** In the presence of odd variables, suppose that Assumptions 4.4, 4.7 and 6.2 and the following compatible condition are satisfied:

$$f + \tilde{f}_\delta = \nabla_y \cdot h + h \nabla_y \log \rho.$$

Then the conclusions in Theorem 4.9 hold.
6.2 Martingale Structure for Entropy Production

Let $P_T^e$ be the law of the original process $(X^e, Y^e)$ and let $Q_T^{R,e}$ be the law of the time reversal in the sense of Definition 6.1 of $(X^e, Y^e)$ under the reversed protocol. In other words, $Q_T^{R,e}$ is the law of the time reversal in the usual sense of $(\delta X^e, \delta Y^e)$ under the reversed protocol. In the presence of odd variables, the entropy production of the original process is still defined as

$$S_{\text{tot}}^e(T) = \log \frac{dP_T^e}{dQ_T^{R,e}}(X^e, Y^e).$$

The entropy production $S_{\text{tot}}^{(1)}(T)$ of the averaged process $X$ can be defined similarly. Under the reversed protocol, $(\delta X^e, \delta Y^e)$ becomes a diffusion process with drift

$$B_\delta = \begin{pmatrix} b_\delta(x, y, T - t) + \epsilon^{-1} f_\delta(x, y, T - t) \\ \epsilon^{-1} g_\delta(x, y, T - t) + \epsilon^{-2} c_\delta(x, y, T - t) \end{pmatrix}$$

and diffusion matrix $D_\delta$, where

$$b^i_\delta(x, y, t) = \delta x_i b^i(\delta x, \delta y, t), \quad f^i_\delta(x, y, t) = \delta x_i f^i(\delta x, \delta y, t),$$

$$g^i_\delta(x, y, t) = \delta y_i g^i(\delta x, \delta y, t), \quad c^i_\delta(x, y, t) = \delta y_i c^i(\delta x, \delta y, t).$$

If Assumption 6.2 is satisfied, then we have $c_\delta = c$ and $D_\delta = D$, and thus this is a special case of the comparable process (9) in the backward case with $\tilde{b} = b_\delta$, $\tilde{f} = f_\delta$, and $\tilde{g} = g_\delta$. The next theorem then follows directly from Theorem 6.3.

**Theorem 6.4** In the presence of odd variables, suppose that Assumptions 4.4, 4.7, and 6.2 and the following compatible condition are satisfied:

$$f + f_\delta = \nabla_y \cdot h + h \nabla_y \log \rho. \quad (12)$$

Then the conclusions in Theorem 5.5 hold.

**Example 6.5** In statistical physics, a classical model is the following Langevin equation under the zero-mass limit [57]:

$$\begin{cases} dX_t^e = \epsilon^{-1} Y_t^e, \\ dY_t^e = [\epsilon^{-1} g(X_t^e, V_t^e, t) - \epsilon^{-2} \gamma Y_t^e]dt + \epsilon^{-1} \eta dW_t, \end{cases}$$

where $X^e$ represents the position of a mesoscopic particle, $Y^e$ represents its velocity, $\gamma$ is the fraction coefficient of the particle with the fluid, and $\eta$ is a constant matrix. In this case, the slow components are all even variables and the fast components are all odd variables. Clearly, the relations (11) hold for this model with $h = 0$. Since $f(x, y, t) = y$ and $c(x, y, t) = -\gamma y$, we also have

$$f_\delta(x, y, t) = -y = -f(x, y, t),$$

$$c_\delta(x, y, t) = -\gamma y = c(x, y, t).$$

Therefore, for the Langevin equation, both Assumption 6.2 and the first compatible condition in the above theorem are automatically satisfied.
6.3 Martingale Structure for Housekeeping Heat

Since the adjoint process coincides with the time reversal when the original process is homogenous and stationary, we should also modify the definition of the adjoint process in the presence of odd variables.

**Definition 6.6** In the presence of odd variables, the adjoint process of \((X^\epsilon, Y^\epsilon)\) is defined as the adjoint process of \((\delta X^\epsilon, \delta Y^\epsilon)\) in the sense of Definition 5.3.

Let \(P_T^\epsilon\) be the law of the original process \((X^\epsilon, Y^\epsilon)\) and let \(\hat{P}_T^\epsilon\) be the law of its adjoint process in the sense of Definition 6.6. In other words, \(\hat{P}_T^\epsilon\) is the law of the adjoint process of \((\delta X^\epsilon, \delta Y^\epsilon)\) in the usual sense. In the presence of odd variables, the housekeeping heat of the original process is still defined as

\[
S_{hk}^{\epsilon}(T) = \log \frac{dP_T^\epsilon(X^\epsilon, Y^\epsilon)}{d\hat{P}_T^\epsilon(X^\epsilon, Y^\epsilon)}.
\]

The housekeeping heat \(S_{hk}^{(1)}(T)\) of the averaged process \(X\) can be defined similarly.

For each \(t\), let \(\mu^\epsilon(x, y, t)\) be the pseudo-stationary density of \((\delta X^\epsilon, \delta Y^\epsilon)\). As discussed earlier, the adjoint process of \((\delta X^\epsilon, \delta Y^\epsilon)\) in the usual sense is a diffusion process with drift

\[
\hat{B}_{\delta} = \begin{pmatrix}
\hat{b}_{\delta} + \epsilon^{-1} \hat{f}_{\delta} \\
\epsilon^{-1} \hat{g}_{\delta} + \epsilon^{-2} \hat{c}_{\delta}
\end{pmatrix}
\]

and diffusion matrix \(D_{\delta}\), where

\[
\begin{align*}
\hat{b}_{\delta} &= -b_{\delta} + \nabla_x \cdot a + a \nabla_x \log \mu^\epsilon, \\
\hat{f}_{\delta} &= -f_{\delta} + \nabla_y \cdot h + h \nabla_y \log \mu^\epsilon, \\
\hat{g}_{\delta} &= -g_{\delta} + \nabla_x \cdot h' + h' \nabla_x \log \mu^\epsilon, \\
\hat{c}_{\delta} &= -c_{\delta} + \nabla_y \cdot \alpha + \alpha \nabla_y \log \mu^\epsilon.
\end{align*}
\]

Under mild conditions, by using the matched asymptotic expansions of singularly perturbed diffusion processes \([48, 49, 62]\), it can be proved that

\[
\begin{align*}
\mu^\epsilon(x, y, t) &= \mu(x, t) \rho(x, y, t) + O(\epsilon), \\
\nabla_x \log \mu^\epsilon(x, y, t) &= \nabla_x \log \mu(x, t) + \nabla_x \log \rho(x, y, t) + O(\epsilon), \\
\nabla_y \log \mu^\epsilon(x, y, t) &= \nabla_y \log \rho(x, y, t) + O(\epsilon),
\end{align*}
\]

where \(\mu(x, t)\) is the pseudo-stationary density of the averaged process \(\delta X\). Therefore, the drift coefficients have the following asymptotic expansions:

\[
\begin{align*}
\hat{b}_{\delta} &= -b_{\delta} + \nabla_x \cdot a + a \nabla_x \log \mu + a \nabla_x \log \rho + O(1), \\
\hat{f}_{\delta} &= -f_{\delta} + \nabla_y \cdot h + h \nabla_y \log \rho, \\
\hat{g}_{\delta} &= -g_{\delta} + \nabla_x \cdot h' + h' \nabla_x \log \mu + h' \nabla_x \log \rho + O(1), \\
\hat{c}_{\delta} &= -c_{\delta} + \nabla_y \cdot \alpha + \alpha \nabla_y \log \rho,
\end{align*}
\]

where we have moved the \(O(\epsilon)\) terms of \(\hat{f}_{\delta}\) and \(\hat{c}_{\delta}\) into \(\hat{b}_{\delta}\) and \(\hat{g}_{\delta}\), respectively. If Equation (12) and Assumptions 4.4 and 6.2 are satisfied, then we have \(\hat{f}_{\delta} = f, \hat{c}_{\delta} = c, \) and \(D_{\delta} = D\), and thus the adjoint process of \((\delta X^\epsilon, \delta Y^\epsilon)\) in the usual sense is a special case of the comparable process (7) in the forward case. The next theorem then follows directly from Theorem 3.4.
Theorem 6.7 In the presence of odd variables, suppose that Assumptions 4.4 and 6.2 and the following compatible condition are satisfied:

\[ f + f_\delta = \nabla_y \cdot h + h \nabla_y \log \rho. \]

Then the conclusions in Theorem 5.8 hold.

7 Proof of Theorem 3.4

Recall that the original process \((X^\epsilon, Y^\epsilon)\) is a diffusion process with drift \(B(x, y, t)\) and diffusion matrix \(D(x, y, t)\). Moreover, the comparable process \((\hat{X}^\epsilon, \hat{Y}^\epsilon)\) is another diffusion process with drift

\[ \hat{B}(x, y, t) = \begin{pmatrix} \hat{b}(x, y, t) + \epsilon^{-1} f(x, y, t) \\ \epsilon^{-1} \hat{g}(x, y, t) + \epsilon^{-2} c(x, y, t) \end{pmatrix} \]

and the same diffusion matrix \(D(x, y, t)\). Without loss of generality, we assume that \(D\) is invertible.

Proof of Theorem 3.4 Since \((X^\epsilon, Y^\epsilon)\) and \((\hat{X}^\epsilon, \hat{Y}^\epsilon)\) have the same diffusion matrix \(D(x, y, t)\), it follows from Girsanov’s theorem that their laws are absolutely continuous with respect to each other and

\[ \frac{dP^\epsilon_t}{dP^\epsilon_t}(X^\epsilon_s, Y^\epsilon_s) = e^{\int_0^t (B-\hat{B})'D^{-1} \Sigma(X^\epsilon_s, Y^\epsilon_s, s)dW_s + \frac{1}{2} \int_0^t (B-\hat{B})'D^{-1}(B-\hat{B})(X^\epsilon_s, Y^\epsilon_s, s)ds}, \]

where

\[ \Sigma(x, y, t) = \begin{pmatrix} \sigma(x, y, t) \\ \epsilon^{-1} \eta(x, y, t) \end{pmatrix}. \]

Therefore, the thermodynamic functional \(F^\epsilon_t\) has the following explicit expression:

\[ F^\epsilon_t = \int_0^t (B-\hat{B})'D^{-1} \Sigma(X^\epsilon_s, Y^\epsilon_s, s)dW_s + \frac{1}{2} \int_0^t (B-\hat{B})'D^{-1}(B-\hat{B})(X^\epsilon_s, Y^\epsilon_s, s)ds, \]

and it is clear that \(e^{-F^\epsilon_t}\) is an exponential martingale. Moreover, using the inversion formula for block matrices, we have

\[ (B-\hat{B})'D^{-1} \Sigma = (b-\hat{b}, \epsilon^{-1}(g-\hat{g}))(a \epsilon^{-1}h \epsilon^{-1}h' \epsilon^{-2} \alpha \epsilon^{-1} \eta) = (b-\hat{b}, \epsilon^{-1}(g-\hat{g}))((a - h\alpha^{-1}h')^{-1} - \epsilon a^{-1}h(a - h'\alpha^{-1}h)^{-1} \epsilon^2 (a - h'a^{-1}h)^{-1}) (\sigma \epsilon^{-1} \eta), \]

from which it is easy to see that \((B-\hat{B})'D^{-1} \Sigma\) independent of \(\epsilon\). Thus

\[ (B-\hat{B})'D^{-1}(B-\hat{B}) = [(B-\hat{B})'D^{-1} \Sigma] \cdot [(B-\hat{B})'D^{-1} \Sigma]', \]
is also independent of $\epsilon$. As a result, the ordered triple $(X^\epsilon, F^\epsilon, Y^\epsilon)$ is a diffusion process solving the SDE

\[
\begin{align*}
\frac{dX^\epsilon_t}{dt} &= [b(X^\epsilon_t, Y^\epsilon_t, t) + \epsilon^{-1} f(X^\epsilon_t, Y^\epsilon_t, t)]dt + \sigma(X^\epsilon_t, Y^\epsilon_t, t)dW_t, \\
\frac{dF^\epsilon_t}{dt} &= \frac{1}{2} (B - \hat{B})' D^{-1} (B - \hat{B})(X^\epsilon_t, Y^\epsilon_t, t) dt + (B - \hat{B})' D^{-1} \Sigma(X^\epsilon_t, Y^\epsilon_t, t) dW_t, \\
\frac{dY^\epsilon_t}{dt} &= [\epsilon^{-1} g(X^\epsilon_t, Y^\epsilon_t, t) + \epsilon^{-2} c(X^\epsilon_t, Y^\epsilon_t, t)]dt + \epsilon^{-1} \eta(X^\epsilon_t, t) dW_t,
\end{align*}
\]

where $X^\epsilon$ and $F^\epsilon$ are slowly varying and $Y^\epsilon$ is rapidly varying. This SDE can be rewritten as

\[
\begin{align*}
\frac{d}{dt} \begin{pmatrix} X^\epsilon_t \\ F^\epsilon_t \\ Y^\epsilon_t \end{pmatrix} &= \begin{pmatrix} \tilde{b} \\ \tilde{f} \\ \tilde{\sigma} \end{pmatrix}(X^\epsilon_t, Y^\epsilon_t, t) dt + \begin{pmatrix} 0 \\ \Sigma(X^\epsilon_t, Y^\epsilon_t, t) \end{pmatrix} dW_t,
\end{align*}
\]

where

\[
\tilde{b} = \left( \frac{1}{2} (B - \hat{B})' D^{-1} (B - \hat{B}) \right), \quad \tilde{f} = \left( \begin{array}{c} f \\ 0 \end{array} \right), \quad \tilde{\sigma} = \left( \begin{array}{cc} \sigma & \sigma \\ \epsilon^{-1} \eta & \epsilon^{-1} \eta \end{array} \right).
\]

The diffusion matrix of $(X^\epsilon, F^\epsilon, Y^\epsilon)$ is given by

\[
\begin{pmatrix} \sigma & \sigma \\ \epsilon^{-1} \eta & \epsilon^{-1} \eta \end{pmatrix} = \begin{pmatrix} \tilde{a} & \epsilon^{-1} \tilde{h} \\ \epsilon^{-1} \tilde{h} & \epsilon^{-1} \tilde{h} \end{pmatrix},
\]

where

\[
\tilde{a} = \left( \begin{array}{cc} a & b - \hat{b} \\ (b - \hat{b})' (B - \hat{B})' D^{-1} (B - \hat{B}) \end{array} \right), \quad \tilde{h} = \left( \begin{array}{c} h \\ (g - \hat{g}) ' \end{array} \right).
\]

Since $\tilde{f} = 0$, the solution to the Poisson equation

\[
\begin{align*}
\begin{cases}
- \Lambda \phi(x, y, t) = \tilde{f}(x, y, t), \\
\int_{\mathbb{R}^n} \phi(x, y, t) \rho(x, y, t) dy = 0
\end{cases}
\end{align*}
\]

is given by $\phi = (\phi, 0)$. It thus follows from Lemma 2.1 that $(X^\epsilon, F^\epsilon) \Rightarrow (X, F)$ in $C([0, T], \mathbb{R}^{m+1})$, where $(X, F)$ is a diffusion process with generator

\[
\tilde{L} = \sum_{i=1}^{m+1} \tilde{w}^i(x, t) \partial_{x_i} + \frac{1}{2} \sum_{i,j=1}^{m+1} \tilde{A}^{ij}(x, t) \partial_{x_i x_j},
\]

where

\[
\begin{align*}
\tilde{w}^i(x, t) &= \int_{\mathbb{R}^n} (\tilde{b}^i + \partial_{x_j} \tilde{f}^j + \partial_{y_j} \tilde{f}^j g^j + \partial_{x_j x_k} \tilde{f}^j h^{jk})(x, y, t) \rho(x, y, t) dy, \\
\tilde{A}^{ij}(x, t) &= \int_{\mathbb{R}^n} (\tilde{a}^{ij} + \tilde{f}^j \tilde{f}^i + \tilde{f}^i \tilde{f}^i + \partial_{y_k} \tilde{f}^j \tilde{h}^{jk} + \partial_{y_k} \tilde{f}^j \tilde{h}^{ij})(x, y, t) \rho(x, y, t) dy.
\end{align*}
\]

Since $\tilde{f} = 0$, direct computations show that

\[
\begin{align*}
\tilde{w}^i(x, t) &= w^i(x, t), \quad \tilde{A}^{ij}(x, t) = A^{ij}(x, t), \quad 1 \leq i, j \leq m, \\
\tilde{w}^{m+1}(x, t) &= \frac{1}{2} \int_{\mathbb{R}^n} (B - \hat{B}) Y^{m+1}(B - \hat{B})(x, y, t) \rho(x, y, t) dy, \\
\tilde{A}^{m+1, j}(x, t) &= \int_{\mathbb{R}^n} (b^j - \hat{b}^j + \partial_{y_k} \phi^{j k} g^k - \partial_{y_k} \phi^{j k} \hat{g}^k)(x, y, t) \rho(x, y, t) dy, \quad 1 \leq j \leq m.
\end{align*}
\]
\[
A^{m+1,m+1}(x, t) = \int_{\mathbb{R}^n} (B - \hat{B})' D^{-1}(B - \hat{B})(x, y, t) \rho(x, y, t) dy.
\]

Therefore, \((X, F)\) can be viewed as the solution to the SDE

\[
\begin{align*}
\frac{dX_t}{dF_t} &= w^t(X_t, t)dt + (\bar{A}^{1/2})_{\cdot \cdot}(X_t, t) dB_t, \\
\frac{dF_t}{dF_t} &= \bar{w}^{m+1}(X_t, t)dt + (\bar{A}^{1/2})^{m+1, \cdot}(X_t, t) dB_t,
\end{align*}
\]

where \(\bar{A} = (\bar{A}^{ij})\) is an \((m + 1) \times (m + 1)\) matrix, \((\bar{A}^{1/2})_{\cdot \cdot}\) is the first row of \(\bar{A}^{1/2}\), \((\bar{A}^{1/2})^{m+1, \cdot}\) is the last row of \(\bar{A}^{1/2}\), and \(B = (B_t)_{t \geq 0}\) is an \((m + 1)\)-dimensional standard Brownian motion defined on some probability space. Since the diffusion matrix of \(X\) is \(A = (A^{ij})\), we have \(\bar{A}^{x, x} = A\), where \(\bar{A}^{x, x}\) is the matrix obtained from \(\bar{A}\) by retaining the first \(m\) rows and first \(m\) columns. Let \(\tilde{X}\) be the averaged process of the comparable process \((\hat{X}^x, \hat{Y}^x)\). Since \(X\) and \(\tilde{X}\) have the same diffusion matrix \(A(x, t)\), it follows from Girsanov’s theorem that their laws are absolutely continuous with respect to each other and

\[
\frac{dP_t}{dP_t}(X_s) = e^{\int_0^t (w - \hat{w})' A^{-1}(\bar{A}^{1/2})_{\cdot \cdot}(X_s, s) dB_s + \frac{1}{2} \int_0^t (w - \hat{w})' A^{-1}(w - \hat{w})(X_s, s) ds}.
\]

Therefore, the thermodynamic functional \(F_t^{(1)}\) has the following explicit expression:

\[
F_t^{(1)} = \int_0^t (w - \hat{w})' A^{-1}(\bar{A}^{1/2})_{\cdot \cdot}(X_s, s) dB_s + \frac{1}{2} \int_0^t (w - \hat{w})' A^{-1}(w - \hat{w})(X_s, s) ds,
\]

and it is clear that \(e^{-F_t^{(1)}}\) is an exponential martingale. Obviously, we have

\[
F_t = \int_0^t (\bar{A}^{1/2})^{m+1, \cdot}(X_s, s) dB_s + \int_0^t \bar{w}^{m+1}(X_s, s) ds.
\]

Therefore, we obtain

\[
F_t^{(2)} = F_t - F_t^{(1)} = \int_0^t \left[ (\bar{A}^{1/2})^{m+1, \cdot} - (w - \hat{w})' A^{-1}(\bar{A}^{1/2})_{\cdot \cdot} \right](X_s, s) dB_s \\
+ \int_0^t \left[ \bar{w}^{m+1} - \frac{1}{2} (w - \hat{w})' A^{-1}(w - \hat{w}) \right](X_s, s) ds.
\]

We next make a crucial observation that

\[
\bar{A}^{m+1,m+1} = 2\bar{w}^{m+1}, \quad \bar{A}^{m+1,x} = w - \hat{w},
\]

which shows that \(e^{-F_t}\) is an exponential martingale. Moreover, straightforward calculations show that

\[
\begin{align*}
&\left[ (\bar{A}^{1/2})^{m+1, \cdot} - (w - \hat{w})' A^{-1}(\bar{A}^{1/2})_{\cdot \cdot} \right](\bar{A}^{1/2})^{m+1, \cdot} - (w - \hat{w})' A^{-1}(\bar{A}^{1/2})_{\cdot \cdot} \\
&= \bar{A}^{m+1,m+1} - 2(\bar{A}^{1/2})^{m+1, \cdot}(\bar{A}^{1/2})_{\cdot \cdot} A^{-1}(w - \hat{w}) + (w - \hat{w})' A^{-1}(w - \hat{w}) \\
&= 2\bar{w}^{m+1} - (w - \hat{w})' A^{-1}(w - \hat{w}).
\end{align*}
\]

This shows that \(e^{-F_t^{(2)}}\) is also an exponential martingale. Since \(M_t = e^{-F_t^{(1)}}\) and \(N_t = e^{-F_t^{(2)}}\) are both martingales and \(M_t, N_t = e^{-F_t}\) is a martingale, we must have \([M, N]_t = 0\) and thus \(M\) and \(N\) are orthogonal martingales. 

\[\square\]
8 Proof of Theorem 4.9

The proof of Theorem 4.9 is much more complicated than that of Theorem 3.4. To do this, we consider an auxiliary process \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon) = (\tilde{X}^\epsilon_t, \tilde{Y}^\epsilon_t)_{0 \leq t \leq T}\) solving the SDE

\[
\begin{aligned}
    d\tilde{X}^\epsilon_t &= \frac{1}{2} (\partial_{x_j} a^{ij} + \epsilon^{-1} \partial_{y_j} h^{ij})(\tilde{X}^\epsilon_t, \tilde{Y}^\epsilon_t, t)dt + \sigma(\tilde{X}^\epsilon_t, \tilde{Y}^\epsilon_t, t)dW_t, \\
    d\tilde{Y}^\epsilon_t &= \frac{1}{2} (\epsilon^{-1} \partial_{x_j} h^{ij} + \frac{1}{2} \partial_{y_j} a^{ij})(\tilde{X}^\epsilon_t, \tilde{Y}^\epsilon_t, t)dt + \epsilon^{-1} \eta(\tilde{X}^\epsilon_t, \tilde{Y}^\epsilon_t, t)dW_t,
\end{aligned}
\]

with its initial distribution being chosen as that of \((X^\epsilon, Y^\epsilon)\). Then \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\) is a diffusion process with drift

\[
\tilde{B} = \frac{1}{2} \left( \nabla_x \cdot a + \epsilon^{-1} \nabla_y \cdot h \right),
\]

and diffusion matrix \(D\). Moreover, we consider another process \((\tilde{X}^{R, \epsilon}, \tilde{Y}^{R, \epsilon}) = (\tilde{X}^{R, \epsilon}_t, \tilde{Y}^{R, \epsilon}_t)_{0 \leq t \leq T}\) solving the SDE

\[
\begin{aligned}
    d\tilde{X}^{R, \epsilon}_t &= \frac{1}{2} (\partial_{x_j} a^{ij} + \epsilon^{-1} \partial_{y_j} h^{ij})(\tilde{X}^{R, \epsilon}_t, \tilde{Y}^{R, \epsilon}_t, T - t)dt + \sigma(\tilde{X}^{R, \epsilon}_t, \tilde{Y}^{R, \epsilon}_t, t)dW_t, \\
    d\tilde{Y}^{R, \epsilon}_t &= \frac{1}{2} (\epsilon^{-1} \partial_{x_j} h^{ij} + \frac{1}{2} \partial_{y_j} a^{ij})(\tilde{X}^{R, \epsilon}_t, \tilde{Y}^{R, \epsilon}_t, T - t)dt + \epsilon^{-1} \eta(\tilde{X}^{R, \epsilon}_t, \tilde{Y}^{R, \epsilon}_t, t)dW_t,
\end{aligned}
\]

with its initial distribution being chosen as that of \((\tilde{X}^{R, \epsilon}, \tilde{Y}^{R, \epsilon})\). Then \((\tilde{X}^{R, \epsilon}, \tilde{Y}^{R, \epsilon})\) is a diffusion process with drift \(\tilde{B}(x, y, T - t)\) and diffusion matrix \(D(x, y, T - t)\).

On the other hand, \((\tilde{X}^{R, \epsilon}, \tilde{Y}^{R, \epsilon})\) is a diffusion process with drift \(\tilde{B}(x, y, T - t)\) and the same diffusion matrix \(D(x, y, T - t)\), where

\[
\tilde{B} = \left( \begin{array}{c} \tilde{B} \\ \epsilon^{-1} f' \\ \epsilon^{-2} \tilde{c} \end{array} \right).
\]

Let \(\tilde{P}^e_T\) be the law of \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\), let \(\tilde{P}^{R, \epsilon}_T\) be the law of \((\tilde{X}^{R, \epsilon}, \tilde{Y}^{R, \epsilon})\), and let \(\tilde{Q}^{R, \epsilon}_T\) be the law of its time reversal \((\tilde{X}^{R, \epsilon}_{T-t}, \tilde{Y}^{R, \epsilon}_{T-t})_{0 \leq t \leq T}\). Then the backward thermodynamic functional \(G_T^\epsilon\) can be decomposed into three parts as

\[
G_T^\epsilon = \log \frac{dP_T^e(X^\epsilon, Y^\epsilon)}{dP_T^e(X^\epsilon, Y^\epsilon)} + \log \frac{d\tilde{P}_T^e(X^\epsilon, Y^\epsilon)}{d\tilde{P}_T^{R, \epsilon}(X^\epsilon, Y^\epsilon)} + \log \frac{d\tilde{Q}_T^{R, \epsilon}(X^\epsilon, Y^\epsilon)}{d\tilde{Q}_T^{R, \epsilon}(X^\epsilon, Y^\epsilon)} = I + II + III.
\]

We next calculate the explicit expressions of the three parts.

**Lemma 8.1**

\[
I = \int_0^T (B - \tilde{B})' D^{-1} \Sigma(X^\epsilon_s, Y^\epsilon_s, s) dW_s + \frac{1}{2} \int_0^T (B - \tilde{B})' D^{-1} (B - \tilde{B})(X^\epsilon_s, Y^\epsilon_s, s) ds.
\]

**Proof** Since \((X^\epsilon, Y^\epsilon)\) and \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\) have the same diffusion matrix, it follows from Girsanov’s theorem that their laws are absolutely continuous with respect to each other and

\[
\frac{dP_T^e(X^\epsilon, Y^\epsilon)}{dP_T^e(X^\epsilon, Y^\epsilon)} = e^{\int_0^T (B - \tilde{B})' D^{-1} \Sigma(X^\epsilon_s, Y^\epsilon_s, s) dW_s + \frac{1}{2} \int_0^T (B - \tilde{B})' D^{-1} (B - \tilde{B})(X^\epsilon_s, Y^\epsilon_s, s) ds},
\]

which completes the proof.

**Lemma 8.2** For each \(t\), let \(p^\epsilon(x, y, t)\) be the probability density of \((X^\epsilon_t, Y^\epsilon_t)\). Then

\[
II = \log \frac{p^\epsilon(X^\epsilon_0, Y^\epsilon_0, 0)}{p^\epsilon(X^\epsilon_T, Y^\epsilon_T, T)}.
\]

Springer
Proof Let $\tilde{p}(x, y, t|x', y', s)$ be the transition probability density of the auxiliary process $(\check{X}^e, \check{Y}^e)$ from time $s$ to time $t$. Then it satisfies the Kolmogorov forward equation
\[
\partial_t \tilde{p}(x, y, t|x', y', s) = (\tilde{L}_t)^* \tilde{p}(x, y, t|x', y', s),
\]
where $\tilde{L}_t$ is the generator of $(\check{X}^e, \check{Y}^e)$ and $(\tilde{L}_t)^*$ is the adjoint operator of $\tilde{L}_t$ with respect to the Lebesgue measure. By the definition of the auxiliary process, it is easy to check that $\tilde{L}_t$ is symmetric with respect to the Lebesgue measure. Therefore, we have
\[
\partial_t \tilde{p}(x, y, t|x', y', s) = \tilde{L}_t \tilde{p}(x, y, t|x', y', s).
\]

On the other hand, let $\bar{p}^R(x', y', t|x, y, s)$ be the transition probability density of the process $(\check{X}^R,e, \check{Y}^R,e)$ from time $s$ to time $t$. Then it satisfies the Kolmogorov backward equation
\[
\partial_t \bar{p}^R(x', y', t|x, y, s) = -\bar{L}^R_s \bar{p}^R(x', y', t|x, y, s),
\]
where $\bar{L}^R_s$ is the generator of $(\check{X}^R,e, \check{Y}^R,e)$. This shows that
\[
\partial_t \bar{p}^R(x', y', T - s|x, y, T - t) = \bar{L}^R_{T-t} \bar{p}^R(x', y', T - s|x, y, T - t)
= \tilde{L}_t \bar{p}^R(x', y', T - s|x, y, T - t),
\]
where have used the fact that $\bar{L}^R_{T-t} = \tilde{L}_t$. Therefore, both the functions $f(x, y, t) = \tilde{p}(x, y, t|x', y', s)$ and $g(x, y, t) = \bar{p}^R(x', y', T - s|x, y, T - t)$ satisfy the following Cauchy problem of the parabolic equation:
\[
\begin{align*}
\partial_t u &= \tilde{L}_t u, & s \leq t \leq T, \\
 u(s) &= \delta(x - x')\delta(y - y').
\end{align*}
\]

By the uniqueness of this parabolic equation, we have
\[
\tilde{p}(x, y, t|x', y', s) = \bar{p}^R(x', y', T - s|x, y, T - t).
\tag{13}
\]

To proceed, we consider the finite-dimensional measurable cylinder set
\[
E = \{ (x, y) \in C([0, T], \mathbb{R}^{m+n}) : x(t_0) \in A_0, y(t_0) \in B_0, x(t_1) \in A_1, y(t_1) \in B_1, \ldots, \\
x(t_N) \in A_N, \ldots, y(t_N) \in B_N, 0 = t_0 < t_1 < \cdots < t_N = T \},
\]
where $A_i$ are Borel sets in $\mathbb{R}^m$ and $B_i$ are Borel sets in $\mathbb{R}^m$. Then, we have
\[
\tilde{P}^e_T(E) = \int_{A_0 \times B_0 \times \cdots \times A_N \times B_N} \mathbb{P}(\check{X}^e_{t_0} \in dx_0, \tilde{Y}^e_{t_0} \in dy_0, \cdots, \check{X}^e_{t_N} \in dx_N, \tilde{Y}^e_{t_N} \in dy_N).
\]

For each $t$, let $\tilde{p}(x, y, t)$ be the probability density of $(\check{X}^e, \check{Y}^e)$ and let $\bar{p}^R(x, y, t)$ be the probability density of $(\check{X}^R,e, \check{Y}^R,e)$. It then follows from (13) that
\[
\mathbb{P}(\check{X}^e_{t_0} \in dx_0, \tilde{Y}^e_{t_0} \in dy_0, \cdots, \check{X}^e_{t_N} \in dx_N, \tilde{Y}^e_{t_N} \in dy_N)
= \tilde{p}(x_0, y_0, t_0) \tilde{p}(x_1, y_1, t_1|x_0, y_0, t_0) \cdots
\]
\[
= \tilde{p}(x_N, y_N, t_N|x_{N-1}, y_{N-1}, t_{N-1}) dx_0 dy_0 \cdots dx_N dy_N
\]
\[
= \tilde{p}(x_0, y_0, 0) \bar{p}^R(x_0, y_0, T - t_0|x_1, y_1, T - t_1) \cdots
\]
\[
= \bar{p}^R(x_{N-1}, y_{N-1}, T - t_{N-1}|x_N, y_N, T - t_N) dx_0 dy_0 \cdots dx_N dy_N
\]
\[
= \bar{p}^R(x_0, y_0, 0) \mathbb{P}(\check{X}^R,e_{T-t_N} \in dx_N, \tilde{Y}^R,e_{T-t_N} \in dy_N, \cdots, \check{X}^R,e_{T-t_0} \in dx_0, \tilde{Y}^R,e_{T-t_0} \in dy_0).
\]
Since \((\tilde{X}^\epsilon, \tilde{Y}^\epsilon)\) and \((X^\epsilon, Y^\epsilon)\) have the same initial distribution and \((\tilde{X}^{R,\epsilon}, \tilde{Y}^{R,\epsilon})\) and \((\tilde{X}^{R,\epsilon}, \tilde{Y}^{R,\epsilon})\) have the same initial distribution, which is the final distribution of \((X^\epsilon, Y^\epsilon)\), we have

\[
\tilde{p}(x_0, y_0, 0) = p^\epsilon(x_0, y_0, 0), \quad \tilde{p}^R(x_N, y_N, 0) = p^\epsilon(x_N, y_N, T).
\]

Therefore, we obtain

\[
\tilde{P}_T^\epsilon(E) = \int \mathbb{P}(\tilde{X}_{t_0}^e \in dx_0, \tilde{Y}_{t_0}^e \in dy_0, \ldots, \tilde{X}_{t_N}^e \in dx_N, \tilde{Y}_{t_N}^e \in dy_N)
\]

\[
= \int \frac{p^\epsilon(x_0, y_0, 0)}{p^\epsilon(x_N, y_N, T)} \mathbb{P}(\tilde{X}_{T-t_0}^{R,\epsilon} \in dx_0, \tilde{Y}_{T-t_0}^{R,\epsilon} \in dy_0, \ldots, \tilde{X}_{T-t_N}^{R,\epsilon} \in dx_N, \tilde{Y}_{T-t_N}^{R,\epsilon} \in dy_N)
\]

\[
= \int_E \frac{p^\epsilon(x(0), y(0), 0)}{\tilde{P}_T^R(x, y, T)} d\tilde{Q}_T^{R,\epsilon}(x, y),
\]

where the first two integrals are taken over \(A_0 \times B_0 \times \cdots A_N \times B_N\). Since the path space \(C([0, T], \mathbb{R}^{n+m})\) can be generated by finite-dimensional measurable cylinder sets, we finally obtain

\[
\frac{d\tilde{P}_T^\epsilon}{d\tilde{Q}_T^{R,\epsilon}}(X^\epsilon, Y^\epsilon) = \frac{p^\epsilon(X_0^\epsilon, Y_0^\epsilon, 0)}{p^\epsilon(X_T^\epsilon, Y_T^\epsilon, T)},
\]

which gives the desired result. \(\square\)

**Lemma 8.3**

\[
\text{III} = \int_0^T (\tilde{B} - \bar{B})'D^{-1} \Sigma(X_s^\epsilon, Y_s^\epsilon, s) dW_s + \frac{1}{2} \int_0^T (\tilde{B} - \bar{B})'D^{-1}(\tilde{B} - \bar{B})(X_s^\epsilon, Y_s^\epsilon, s) ds
\]

\[
+ \int_0^T (B - \bar{B})'D^{-1}(\tilde{B} - \bar{B})(X_s^\epsilon, Y_s^\epsilon, s) ds + \int_0^T \nabla \cdot (\tilde{B} - \bar{B})(X_s^\epsilon, Y_s^\epsilon, s) ds,
\]

**Proof** Since \((\tilde{X}^{R,\epsilon}, \tilde{Y}^{R,\epsilon})\) and \((\tilde{X}^{R,\epsilon}, \tilde{Y}^{R,\epsilon})\) have the same diffusion matrix, it follows from Girsanov’s theorem that their laws are absolutely continuous with respect to each other and

\[
\frac{d\tilde{P}_T^{R,\epsilon}}{d\tilde{P}_T^{R,\epsilon}}(\tilde{X}^\epsilon, \tilde{Y}^\epsilon) = e^{\int_0^T (\tilde{B} - \bar{B})'D^{-1}\Sigma(\tilde{X}_s^\epsilon, \tilde{Y}_s^\epsilon, s) ds + \frac{1}{2} \int_0^T (\tilde{B} - \bar{B})'D^{-1}(\tilde{B} - \bar{B})(\tilde{X}_s^\epsilon, \tilde{Y}_s^\epsilon, s) ds}.
\]

For convenience of notation, set \(\tilde{Z}^{R,\epsilon} = (\tilde{X}^{R,\epsilon}, \tilde{Y}^{R,\epsilon})\). Then we have

\[
d\tilde{Z}_t^{R,\epsilon} = \tilde{B}(\tilde{Z}_t^{R,\epsilon}, T - t) dt + \Sigma(\tilde{Z}_t^{R,\epsilon}, T - t) dW_t.
\]

This shows that

\[
\frac{d\tilde{P}_T^{R,\epsilon}}{d\tilde{P}_T^{R,\epsilon}}(\tilde{Z}^\epsilon) = e^{\int_0^T (\tilde{B} - \bar{B})'D^{-1}\Sigma(\tilde{Z}_s^{R,\epsilon}, T - s) ds + \frac{1}{2} \int_0^T (\tilde{B} - \bar{B})'D^{-1}(\tilde{B} + \bar{B})(\tilde{Z}_s^{R,\epsilon}, T - s) ds}.
\]

In other words, we have

\[
\log \frac{d\tilde{P}_T^{R,\epsilon}}{d\tilde{P}_T^{R,\epsilon}}(w) = \int_0^T (\tilde{B} - \bar{B})'D^{-1}(w(s), T - s)dw(s) - \int_0^T C(w(s), T - s) ds := M(w),
\]

where

\[
C = \frac{1}{2}(\tilde{B} - \bar{B})'D^{-1}(\tilde{B} + \bar{B}).
\]

Springer
Let \( \phi : C([0, T], \mathbb{R}^{m+n}) \to C([0, T], \mathbb{R}^{m+n}) \) be the time-reversed operator on the path space defined as

\[
\phi(w(\cdot)) = w(T - \cdot).
\]

Clearly, we have \( \phi_* \tilde{P}_T^{R,e} = \tilde{Q}_T^{R,e} \) and \( \phi_* \tilde{P}_T^{R,e} = \tilde{Q}_T^{R,e} \), where \( \phi_* \mu \) is the push-forward measure of \( \mu \) defined as \( \phi_* \mu(\cdot) = \mu(\phi^{-1}(\cdot)) \). It thus follows that for any Borel set \( E \subset C([0, T], \mathbb{R}^{m+n}) \),

\[
\tilde{Q}_T^{R,e}(E) = \tilde{P}_T^{R,e}(\phi^{-1}(E)) = \int_{\phi^{-1}(E)} e^{M(w)} d\tilde{P}_T^{R,e} = \int_E e^{M(\phi(w))} d\phi_* \tilde{P}_T^{R,e} = \int_E e^{M(\phi(w))} d\tilde{Q}_T^{R,e}.
\]

This shows that

\[
\frac{d \tilde{Q}_T^{R,e}}{d \tilde{Q}_T^{R,e}}(\cdot) = e^{M(\phi(\cdot))}.
\]

For convenience of notation, set \( Z^e = (X^e, Y^e)' \). Then we have

\[
\log \frac{d \tilde{Q}_T^{R,e}}{d \tilde{Q}_T^{R,e}}(Z^e) = \int_0^T (\bar{B} - \tilde{B})' D^{-1}(Z^e_{T-s}, T - s) dZ^e_{T-s} - \int_0^T C(Z^e_{T-s}, T - s) ds
\]

\[
= \int_0^T (\bar{B} - \tilde{B})' D^{-1}(Z^e_s, s) \ast dZ^e_s - \int_0^T C(Z^e_s, s) ds,
\]

where the first integral in the last equality is the right-endpoint stochastic integral (recall that the Itô integral is the left-endpoint stochastic integral and the Stratonovich integral is the middle-endpoint stochastic integral). Since the sum of the right-endpoint and left-endpoint stochastic integrals is twice the middle-endpoint stochastic integral, we have \( f * dZ = fdZ + dfdZ \), which implies that

\[
\log \frac{d \tilde{Q}_T^{R,e}}{d \tilde{Q}_T^{R,e}}(Z^e) = \int_0^T (\bar{B} - \tilde{B})' D^{-1}(Z^e_s, s) dZ^e_s + \int_0^T d(\bar{B} - \tilde{B})' D^{-1}(Z^e_s, s) dZ^e_s
\]

\[
- \int_0^T C(Z^e_s, s) ds
\]

\[
= \int_0^T (\bar{B} - \tilde{B})' D^{-1} \Sigma(Z^e_s, s) dW_s + \int_0^T (\bar{B} - \tilde{B})' D^{-1} B(Z^e_s, s) ds
\]

\[
+ \int_0^T \partial_j[(\bar{B} - \tilde{B})' D^{-1}]_i D^{ij}(Z^e_s, s) ds - \int_0^T C(Z^e_s, s) ds.
\]

By the definition of the auxiliary process, it is easy to see that \( \nabla \cdot D = 2\bar{B} \), which implies that

\[
\partial_j[(\bar{B} - \tilde{B})' D^{-1}]_i D^{ij} = \nabla \cdot (\bar{B} - \tilde{B}) - 2(\bar{B} - \tilde{B})' D^{-1} \bar{B}.
\]

Therefore, we obtain

\[
\log \frac{d \tilde{Q}_T^{R,e}}{d \tilde{Q}_T^{R,e}}(Z^e) = \int_0^T (\bar{B} - \tilde{B})' D^{-1} \Sigma(Z^e_s, s) dW_s + \frac{1}{2} \int_0^T (\bar{B} - \tilde{B})' D^{-1}(\bar{B} - \tilde{B})(Z^e_s, s) ds
\]

\[
+ \int_0^T (\bar{B} - \tilde{B})' D^{-1}(\bar{B} - \tilde{B})(Z^e_s, s) ds + \int_0^T \nabla \cdot (\bar{B} - \tilde{B})(Z^e_s, s) ds.
\]

which gives the desired result.
So far, we have obtained the explicit expressions of the three parts. Combining these three parts, we obtain the following corollary.

**Corollary 8.4**

\[ G_T^e = \log \frac{p^e(X_0^e, Y_0^e, 0)}{p^e(X_T^e, Y_T^e, T)} + \int_0^T (B + \tilde{B} - 2\tilde{B})' D^{-1} \Sigma(X_s^e, Y_s^e, s) dW_s \]

\[ + \frac{1}{2} \int_0^T (B + \tilde{B} - \tilde{B})' D^{-1}(B + \tilde{B} - \tilde{B})(X_s^e, Y_s^e, s) ds \]

\[ + \int_0^T \nabla \cdot (\tilde{B} - \tilde{B})(X_s^e, Y_s^e, s) ds. \]

To proceed, note that the second part can be rewritten as

\[ II = \log \frac{p^e(X_0^e, Y_0^e, 0)}{\rho(X_0^e, Y_0^e, 0)} - \log \frac{p^e(X_T^e, Y_T^e, T)}{\rho(X_T^e, Y_T^e, T)} + \log \frac{\rho(X_0^e, Y_0^e, 0)}{\rho(X_T^e, Y_T^e, T)}, \]

where \( \rho \) is the pseudo-stationary density of \( Y^x \) defined in (4). Therefore, the backward thermodynamic functional can be represented as

\[ G_T^e = I_T^e + H_T^e, \]

where

\[ I_T^e = \log \frac{p^e(X_0^e, Y_0^e, 0)}{\rho(X_0^e, Y_0^e, 0)} - \log \frac{p^e(X_T^e, Y_T^e, T)}{\rho(X_T^e, Y_T^e, T)}, \quad H_T^e = 1 + III + \log \frac{\rho(X_0^e, Y_0^e, 0)}{\rho(X_T^e, Y_T^e, T)}. \]

We shall next investigate the weak limits of \( H_T^e \) and \( I_T^e \) separately. The following lemma gives another explicit expression of \( H_T^e \), whose weak limit is much easier to deal with.

**Lemma 8.5**

\[ H_T^e = \int_0^T \sigma^{m+1}(X_s^e, Y_s^e, s) dW_s + \int_0^T (b^{m+1} + \epsilon^{-1} f^{m+1})(X_s^e, Y_s^e, s) ds, \]

where

\[ \sigma^{m+1} = (B + \tilde{B} - \tilde{B} - D \nabla \log \rho)' D^{-1} \Sigma, \]

\[ b^{m+1} = \frac{1}{2} \sigma^{m+1}(\sigma^{m+1})' + \frac{1}{\rho} \nabla \cdot [(\tilde{B} - \frac{1}{2} \nabla \cdot a - \frac{1}{2} a \nabla \log \rho) - \partial_s \log \rho], \]

\[ f^{m+1} = \frac{1}{\rho} \nabla \cdot [(\tilde{g} \rho) - \nabla \cdot (f \rho)]. \]

**Proof** By Ito’s formula, we have

\[ \log \frac{\rho(X_0^e, Y_0^e, 0)}{\rho(X_T^e, Y_T^e, T)} = - \int_0^T d \log \rho(X_s^e, Y_s^e, s) \]

\[ = - \int_0^T (\nabla \log \rho)' \Sigma(X_s^e, Y_s^e, s) dW_s \]

\[ - \int_0^T [(\nabla \log \rho)' B + \partial_s \log \rho + \frac{1}{2} \partial_{ij} \log \rho D^{ij}](X_s^e, Y_s^e, s) ds. \]

Since \( \nabla \cdot D = 2 \tilde{B} \), we have

\[ \frac{1}{2} \partial_{ij} \log \rho D^{ij} = \frac{1}{2} \nabla \cdot (D \nabla \log \rho) - \tilde{B}' \nabla \log \rho. \]
This shows that
\[
\log \frac{\rho(X^e_0, Y^e_0)}{\rho(X^e_T, Y^e_T)} = -\int_0^T (\nabla \log \rho)' \Sigma(X^e_s, Y^e_s, s) dW_s
- \int_0^T (B - \bar{B})' \nabla \log \rho(X^e_s, Y^e_s, s) ds
- \frac{1}{2} \int_0^T \nabla \cdot (D \nabla \log \rho)(X^e_s, Y^e_s, s) ds - \int_0^T \partial_s \log \rho(X^e_s, Y^e_s, s) ds.
\]

This equation, together with Corollary 8.4, shows that
\[
H^e_T = \int_0^T (B + \bar{B} - 2\bar{B} - D \nabla \log \rho)' D^{-1}(B + \bar{B} - 2\bar{B} - D \nabla \log \rho)(X^e_s, Y^e_s, s) ds
+ \frac{1}{2} \int_0^T (B + \bar{B} - 2\bar{B} - D \nabla \log \rho)' D^{-1}(B + \bar{B} - 2\bar{B} - D \nabla \log \rho)(X^e_s, Y^e_s, s) ds
+ \int_0^T (\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho)' \nabla \log \rho(X^e_s, Y^e_s, s) ds
+ \int_0^T \nabla \cdot ((\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho))(X^e_s, Y^e_s, s) ds - \int_0^T \partial_s \log \rho(X^e_s, Y^e_s, s) ds. \tag{14}
\]

It is easy to see that
\[
(\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho)' \nabla \log \rho + \nabla \cdot ((\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho) \rho) = \frac{1}{\rho} \nabla \cdot [(\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho) \rho].
\]

Inserting this equation into (14) yields
\[
H^e_T = \int_0^T (B + \bar{B} - 2\bar{B} - D \nabla \log \rho)' D^{-1}(B + \bar{B} - 2\bar{B} - D \nabla \log \rho)(X^e_s, Y^e_s, s) ds
+ \frac{1}{2} \int_0^T (B + \bar{B} - 2\bar{B} - D \nabla \log \rho)' D^{-1}(B + \bar{B} - 2\bar{B} - D \nabla \log \rho)(X^e_s, Y^e_s, s) ds
+ \int_0^T \frac{1}{\rho} \nabla \cdot [(\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho) \rho](X^e_s, Y^e_s, s) ds - \int_0^T \partial_s \log \rho(X^e_s, Y^e_s, s) ds. \tag{15}
\]

Moreover, it is easy to check that
\[
\bar{B} - \bar{B} - \frac{1}{2} D \nabla \log \rho
= \begin{pmatrix}
(b - \frac{1}{2} \nabla \alpha - \frac{1}{2} a \nabla x \log \rho) + e^{-1}(f - \frac{1}{2} \nabla y \cdot h - \frac{1}{2} h \nabla y \log \rho) \\
\epsilon^{-1}((\bar{g} - \frac{1}{2} \nabla x \cdot h' - \frac{1}{2} h' \nabla x \log \rho) + \epsilon^{-2}(c - \frac{1}{2} \nabla y \cdot \alpha - \frac{1}{2} \alpha \nabla y \log \rho)
\end{pmatrix}. \tag{16}
\]

By Assumption 4.4 and Lemma 4.5, the pseudo-stationary density $\rho$ satisfies
\[
2c = \nabla y \cdot \alpha + \alpha \nabla y \log \rho.
\]

Moreover, it follows from Assumption 4.3 that
\[
f - \frac{1}{2} \nabla y \cdot h - \frac{1}{2} h \nabla y \log \rho = \frac{1}{2} \nabla y \cdot h + \frac{1}{2} h \nabla y \log \rho - f.
\]
Therefore, the right side of (16) can be simplified as
\[
\tilde{B} - \tilde{b} - \frac{1}{2} D \nabla \log \rho = \left( \tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho + \epsilon^{-1} \left( \frac{1}{2} \nabla_y \cdot h + \frac{1}{2} h \nabla_y \log \rho - f \right) \right).
\]

This shows that
\[
\nabla \cdot \left[ (\tilde{B} - \tilde{b} - \frac{1}{2} D \nabla \log \rho) \rho \right] = \nabla_x \cdot \left[ (\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho) \rho \right] + \epsilon^{-1} \nabla_x \cdot \left[ \left( \frac{1}{2} \nabla_y \cdot h + \frac{1}{2} h \nabla_y \log \rho - f \right) \rho \right] + \epsilon^{-1} \nabla_y \cdot \left[ (\tilde{g} - \frac{1}{2} \nabla_x \cdot h' - \frac{1}{2} h' \nabla_x \log \rho) \rho \right] = \nabla_x \cdot \left[ (\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho) \rho \right] + \epsilon^{-1} \left[ \nabla_y \cdot (\tilde{g} \rho) - \nabla_x \cdot (f \rho) \right].
\]

Inserting this into (15) gives the desired result. \(\square\)

The following lemma gives the limit of \(I_T^\epsilon\).

**Lemma 8.6** For each \(t\), let \(p(x, t)\) be the probability density of \(X_t\). Then for any \(\eta > 0\), we have
\[
\lim_{\epsilon \to 0} \mathbb{P} \left( \sup_{t \leq T} \left| I_T^\epsilon - \log \frac{p(X_0^\epsilon, 0)}{p(X_0, 0)} \right| > \eta \right) = 0.
\]

**Proof** We first assume that the original process \((X^\epsilon, Y^\epsilon)\) starts at \(t = 0\). By the matched asymptotic expansions of singularly perturbed diffusion processes under second-order averaging [62], there exist two sequences of smooth functions \(u_i(x, y, t)\) and \(v_i(x, y, t)\) such that for any \(n \geq 0\) and any compact set \(K \subset \mathbb{R}^{m+n}\),
\[
\sup_{(x, y, t) \in K \times [0, T]} \left| \sum_{i=0}^{n} \epsilon^i u_i(x, y, t) + \sum_{i=0}^{n} \epsilon^i v_i(t/\epsilon, x, y) - p^\epsilon(x, y, t) \right| = O(\epsilon^{n+1}),
\]

where \(u_i(x, y, t)\) are called outer expansions and \(v_i(x, y, t)\) are called initial layer corrections. Moreover, the initial layer corrections \(v_i(x, y, t)\) decay exponentially fast:
\[
\sup_{(x, y) \in K} \left| v_i(x, y, t) \right| \leq c_1 e^{-c_2 t},
\]

where \(c_1\) and \(c_2\) are two constants independent of \(x\) and \(y\). Although the above two estimates are proved in [62] for singularly perturbed diffusions on compact Riemannian manifolds, the proof can be easily generalized to diffusions on the entire space [48, Sect. 6]. Combining the above two estimates, for any \(\delta > 0\), we have
\[
\sup_{(x, y, t) \in K \times [\delta, T]} \left| p^\epsilon(x, y, t) - u_0(x, y, t) \right| \leq \sup_{(x, y, t) \in K \times [0, T]} \left| p^\epsilon(x, y, t) - u_0(x, y, t) - v_0(t/\epsilon, x, y) \right| + \sup_{(x, y, t) \in K \times [\delta, T]} \left| v_0(t/\epsilon, x, y) \right| \leq O(\epsilon^2) + c_1 e^{-c_2 \delta/\epsilon}.
\]

Moreover, the outer expansion \(u_0(x, y, t)\) is given by [48]
\[
u_0(x, y, t) = p(x, t) \rho(x, y, t) > 0,
\]

\(\square\) Springer
where the positivity of $u_0$ follows from Harnack’s inequality [79]. It thus follows from the above two equations that
\[
\lim_{\epsilon \to 0} \sup_{(x,y,t) \in K \times [\delta, T]} \left| \frac{p^\epsilon(x, y, t)}{p(x, t)\rho(x, y, t)} - 1 \right| = 0.
\]
Since we have assumed that the original process $(X^\epsilon, Y^\epsilon)$ starts at $t = -1$, we can replace $\delta$ by $0$ in the above limit. This implies that
\[
\lim_{\epsilon \to 0} \sup_{(x,y,t) \in K \times [0, T]} \left| \log \frac{p^\epsilon(x, y, t)}{\rho(x, y, t)p(x, t)} \right| = 0.
\]
For any $\delta > 0$, it follows from Assumption 4.7 that there exists a compact set $K_\delta \subset \mathbb{R}^{m+n}$ such that
\[
\mathbb{P} \left( \inf_{\epsilon > 0} \tau^\epsilon_{K_\delta} < T \right) \leq \delta.
\]
Therefore, for any $\eta > 0$, when $\epsilon$ is sufficiently small, we have
\[
\mathbb{P} \left( \sup_{t \leq T} \left| \log \frac{p^\epsilon(X^\epsilon_t, Y^\epsilon_t)}{\rho(X^\epsilon_t, Y^\epsilon_t)p(X^\epsilon_t, t)} \right| > \eta \right) \leq \mathbb{P} \left( \sup_{t \leq T} \left| \log \frac{p^\epsilon(X^\epsilon_t, Y^\epsilon_t)}{\rho(X^\epsilon_t, Y^\epsilon_t)p(X^\epsilon_t, t)} \right| > \eta, \inf_{\epsilon > 0} \tau^\epsilon_{K_\delta} \geq T \right) + \mathbb{P} \left( \inf_{\epsilon > 0} \tau^\epsilon_{K_\delta} < T \right) \leq \delta.
\]
This finally shows that
\[
\lim_{\epsilon \to 0} \mathbb{P} \left( \sup_{t \leq T} \left| \log \frac{p^\epsilon(X^\epsilon_t, Y^\epsilon_t)}{\rho(X^\epsilon_t, Y^\epsilon_t)p(X^\epsilon_t, t)} \right| > \eta \right) = 0,
\]
which implies the desired result. \(\square\)

Interestingly, we have the following proposition, which shows that the two limiting processes, $X$ and $\tilde{X}^R$ have the same functional forms of the diffusion coefficients.

**Proposition 8.7** Under Assumptions 4.3 and 4.4, we have $A^{ij}(x, t) = \tilde{A}^{ij}(x, t)$, where $A^{ij}(x, t)$ is the diffusion coefficient of $X$ and $\tilde{A}^{ij}(x, T - t)$ is the diffusion coefficient of $\tilde{X}^R$.

**Proof** Using Assumption 4.3 and integration by parts, we obtain
\[
\int_{\mathbb{R}^n} (\tilde{\phi}^i \tilde{f}^j + \partial_y \tilde{\phi}^i h^{jk}) \rho dy = \int_{\mathbb{R}^n} \tilde{\phi}^i \tilde{f}^j \rho dy - \int_{\mathbb{R}^n} \tilde{\phi}^i \rho \partial_y (h^{jk} \rho) dy = \int_{\mathbb{R}^n} \tilde{\phi}^i \tilde{f}^j - \partial_y h^{jk} - h^{jk} \partial_y \log \rho \rho dy = -\int_{\mathbb{R}^n} \tilde{\phi}^i \tilde{f}^j \rho dy.
\]
By Assumption 4.4 and Lemma 4.5, $L^0$ is symmetric with respect to $\rho$ and thus
\[
\int_{\mathbb{R}^n} \tilde{\phi}^i \tilde{f}^j \rho dy = -(\tilde{\phi}^i, L_0 \tilde{\phi}^j)_\rho = -(L_0 \tilde{\phi}^i, \tilde{\phi}^j)_\rho = \int_{\mathbb{R}^n} \tilde{\phi}^i \tilde{f}^j \rho dy.
\]
Using Assumption 4.3 and integration by parts again yields
\[
\int_{\mathbb{R}^n} (\tilde{\phi}^i \tilde{f}^j + \partial_y \tilde{\phi}^i h^{jk}) \rho dy = \int_{\mathbb{R}^n} \phi^j (f^i - \partial_y h^{ik} - h^{ik} \partial_y \log \rho) \rho dy
\]
Combining the above two equations finally yields
\[
\int_{\mathbb{R}^n} (\tilde{\phi}^j f^i + \partial_{X_k} \tilde{\phi}^j h^{ik}) \rho dy = \int_{\mathbb{R}^n} (\phi^i f^i + \partial_{X_k} \phi^i h^{ik}) \rho dy.
\]
Similarly, we can prove that
\[
\int_{\mathbb{R}^n} (\tilde{\phi}^i \tilde{f}^j + \partial_{X_k} \tilde{\phi}^i h^{ik}) \rho dy = \int_{\mathbb{R}^n} (\phi^i f^j + \partial_{X_k} \phi^i h^{ik}) \rho dy.
\]
Combining the above two equations finally yields
\[
\tilde{A}^{ij}(x,t) = \int_{\mathbb{R}^n} (a^{ij} + \tilde{\phi}^j \tilde{f}^i + \partial_{X_k} \tilde{\phi}^j h^{ik} + \partial_{X_k} \tilde{\phi}^i h^{jk}) \rho dy = \int_{\mathbb{R}^n} (a^{ij} + \phi^j f^i + \partial_{X_k} \phi^j h^{ik} + \partial_{X_k} \phi^i h^{jk}) \rho dy = A^{ij}(x,t),
\]
which gives the desired result.

We next derive the explicit expressions of $G_T$ and $G_T^{(1)}$. Clearly, we have
\[
B + \tilde{B} - 2\tilde{B} - D \nabla \log \rho = \left( (b + \tilde{b} - \nabla_x \cdot a - a \nabla_x \log \rho) + e^{-1} (f + \tilde{f} - \nabla_y \cdot h - h \nabla_y \log \rho) \right).
\]
By Assumptions 4.3 and 4.4, the right side of the above equation can be simplified as
\[
B + \tilde{B} - 2\tilde{B} - D \nabla \log \rho = \left( b + \tilde{b} - \nabla_x \cdot a - a \nabla_x \log \rho \right).
\]
It is easy to check that the functions $\sigma^{m+1}$, $b^{m+1}$, and $f^{m+1}$ defined in Lemma 8.5 are all independent of $\epsilon$. It then follows from Lemma 8.5 that the ordered triple $(X^\epsilon, H^\epsilon, Y^\epsilon)$ is a diffusion process solving the SDE
\[
\begin{align*}
\left\{ \begin{array}{l}
dX^\epsilon_t = [b(X^\epsilon_t, Y^\epsilon_t, t) + e^{-1} f(X^\epsilon_t, Y^\epsilon_t, t)]dt + \sigma(X^\epsilon_t, Y^\epsilon_t, t) dW_t, \\
dH^\epsilon_t = [b^{m+1}(X^\epsilon_t, Y^\epsilon_t, t) + e^{-1} f^{m+1}(X^\epsilon_t, Y^\epsilon_t, t)]dt + \sigma^{m+1}(X^\epsilon_t, Y^\epsilon_t, t) dW_t, \\
dY^\epsilon_t = [e^{-1} g(X^\epsilon_t, Y^\epsilon_t, t)]dt + e^{-1} \eta(X^\epsilon_t, Y^\epsilon_t, t) dW_t,
\end{array} \right.
\end{align*}
\]
where $X^\epsilon$ and $H^\epsilon$ are slowly varying and $Y^\epsilon$ is rapidly varying. This SDE can be rewritten as
\[
\begin{align*}
\left\{ \begin{array}{l}
d\left( \begin{array}{c}
X^\epsilon_t \\
H^\epsilon_t
\end{array} \right) = [\tilde{b}(X^\epsilon_t, Y^\epsilon_t, t) + e^{-1} \tilde{f}(X^\epsilon_t, Y^\epsilon_t, t)]dt + \tilde{\sigma}(X^\epsilon_t, Y^\epsilon_t, t) dW_t, \\
dY^\epsilon_t = [e^{-1} g(X^\epsilon_t, Y^\epsilon_t, t)]dt + e^{-1} \eta(X^\epsilon_t, Y^\epsilon_t, t) dW_t,
\end{array} \right.
\end{align*}
\]
where
\[
\tilde{b} = \left( \begin{array}{c}
b \\
b^{m+1}
\end{array} \right), \quad \tilde{f} = \left( \begin{array}{c}
f \\
f^{m+1}
\end{array} \right), \quad \tilde{\sigma} = \left( \begin{array}{c}
\sigma \\
\sigma^{m+1}
\end{array} \right).
\]
The diffusion matrix of $(X^\epsilon, H^\epsilon, Y^\epsilon)$ is given by
\[
\left( \begin{array}{c}
\tilde{\sigma} \\
\tilde{\sigma} \eta
\end{array} \right) = \left( \begin{array}{c}
\tilde{a} \\
\tilde{a} \eta
\end{array} \right),
\]
where
\[
\tilde{a} = \left( \begin{array}{c}
a \\
(a + b - \nabla_x \cdot a - a \nabla_x \log \rho) \sigma^{m+1}(a^{m+1})
\end{array} \right).
\]
Obviously, we have
\[\hat{h} = \left( g + \bar{g} - \nabla_x \cdot h' - h' \nabla_x \log \rho \right).\]

Let \(\tilde{\phi} = (\phi, \phi^{m+1})'\) be solution to the Poisson equation
\[
\begin{cases}
- L_0 \tilde{\phi}(x, y, t) = f(x, y, t), \\
\int_{\mathbb{R}^n} \tilde{\phi}(x, y, t) \rho(x, y, t) dy = 0.
\end{cases}
\]

It thus follows from Lemma 2.1 that \((X^e, H^e) \Rightarrow (X, H)\) in \(C([0, T], \mathbb{R}^{m+1})\), where \((X, H)\) is a diffusion process with generator
\[
\tilde{L} = \sum_{i=1}^{m+1} \tilde{u}^i(x, t) \partial_{x_i} + \frac{1}{2} \sum_{i,j=1}^{m+1} \tilde{A}^{ij}(x, t) \partial_{x_i x_j},
\]

where
\[
\begin{align*}
\tilde{u}^i(x, t) &= \int_{\mathbb{R}^n} (\tilde{b}^i + \partial_{x_i} \tilde{\phi} f^i + \partial_{x_j} \tilde{\phi} g^i + \partial_{x_i x_k} \tilde{\phi} h^{ij}) (x, x, t) \rho(x, y, t) dy, \\
\tilde{A}^{ij}(x, t) &= \int_{\mathbb{R}^n} (\tilde{a}^{ij} + \phi^i \tilde{f}^j + \phi^j \tilde{f}^i + \partial_{x_i x_k} \phi^j h^{ik}) (x, y, t) \rho(x, y, t) dy.
\end{align*}
\]

It is easy to see that \(\tilde{u}^i(x, t) = u^i(x, t)\) and \(\tilde{A}^{ij}(x, t) = A^{ij}(x, t)\) for any \(1 \leq i, j \leq m\). Thus, \((X, H)\) can be viewed as the solution to the SDE
\[
\begin{cases}
\begin{align*}
&dX_i^t = u^i(X_t, t) dt + (\tilde{A}^{1/2})^{x \cdot} (X_t, t) dB_t, \\
&dH_t = \tilde{w}^{m+1}(X_t, t) dt + (\tilde{A}^{1/2})^{m+1 \cdot} (X_t, t) dB_t,
\end{align*}
\end{cases}
\]

where \(\tilde{A} = (\tilde{A}^{ij})\) is an \((m+1) \times (m+1)\) matrix, \((\tilde{A}^{1/2})^{x \cdot}\) is the first \(m\) rows of \(\tilde{A}^{1/2}\), \((\tilde{A}^{1/2})^{m+1 \cdot}\) is the last row of \(\tilde{A}^{1/2}\), and \(B = (B_t)_{t \geq 0}\) is an \((m+1)\)-dimensional standard Brownian motion defined on some probability space. Since the diffusion matrix of \(X\) is \(A(x, t) = (A^{ij}(x, t))\), we have \(\tilde{A}^{x \cdot} = A\), where \(\tilde{A}^{x \cdot}\) the matrix obtained from \(\tilde{A}\) by retaining the first \(m\) rows and first \(m\) columns. Let \(\tilde{X}^R\) be the averaged process of the comparable process \((\tilde{X}^{R,e}, \tilde{Y}^{R,e})\). By Proposition 8.7, the diffusion matrix of \(\tilde{X}^R\) is \(A(x, T - t)\) and thus \(\tilde{X}^R\) can be viewed as the solution to the SDE
\[
d\tilde{X}_t^R = \tilde{w}(\tilde{X}_t^R, T - t) dt + (\tilde{A}^{1/2})^{x \cdot} (\tilde{X}_t^R, T - t) dB_t.
\]

Imitating the proof of Corollary 8.4, we can also prove that the laws of \((X_t)_{0 \leq t \leq T}\) and \((\tilde{X}_t^R)_{0 \leq t \leq T}\) are absolutely continuous with respect to each other and the thermodynamic functional \(G_{T}^{(1)}\) has the following explicit expression:
\[
G_{T}^{(1)} = \log \frac{p(X_0, 0)}{p(X_T, T)} + \int_0^T (w + \tilde{w} - \nabla_x \cdot A)' A^{-1} (\tilde{A}^{1/2})^{x \cdot} (X_s, s) dB_s \\
+ \frac{1}{2} \int_0^T (w + \tilde{w} - \nabla_x \cdot A) A^{-1} (w + \tilde{w} - \nabla_x \cdot A)(X_s, s) ds \\
+ \int_0^T \nabla \cdot (\tilde{w} - \frac{1}{2} \nabla_x \cdot A)(X_s, s) ds. \tag{18}
\]

Obviously, we have
\[
H_T = \int_0^T (\tilde{A}^{1/2})^{m+1 \cdot} (X_s, s) dB_s + \int_0^T \tilde{w}^{m+1}(X_s, s) ds.
\]
Since \((X^\varepsilon, H^\varepsilon) \Rightarrow (X, H)\) in \(C([0, T], \mathbb{R}^{m+1})\), we have
\[
\left( X^\varepsilon_t, \log \frac{p(X^\varepsilon_0, 0)}{p(X^\varepsilon_t, t)} + H^\varepsilon_t \right) \Rightarrow \left( X_t, \log \frac{p(X_0, 0)}{p(X_t, t)} + H_t \right).
\]
in \(C([0, T], \mathbb{R}^{m+1})\) \cite[p. 350]{80}. Moreover, it follows from Lemma 8.6 that for any \(\eta > 0\),
\[
\lim_{\varepsilon \to 0} \mathbb{P} \left( \sup_{t \leq T} \left| I^\varepsilon_t - \log \frac{p(X^\varepsilon_0, 0)}{p(X^\varepsilon_t, t)} \right| > \eta \right) = 0.
\]

It then follows from the above two equations and \cite[p. 352, Lemma 3.31]{80} that
\[
(X^\varepsilon_t, G^\varepsilon_t) = (X^\varepsilon_t, I^\varepsilon_t + H^\varepsilon_t) \Rightarrow \left( X_t, \log \frac{p(X_0, 0)}{p(X_t, t)} + H_t \right)
\]
in \(C([0, T], \mathbb{R}^{m+1})\). Therefore, the limit functional \(G_t\) has the following explicit expression:
\[
G_t = \log \frac{p(X_0, 0)}{p(X_t, t)} + H_t.
\] \hspace{1cm} (19)

The following three lemmas play an important role in proving that the anomalous part of the limit functional \(G_t\) is a martingale.

**Lemma 8.8**
\[
\bar{w}^{m+1} = \frac{1}{2} \tilde{a}^{m+1,m+1} + \nabla_x \cdot \int_{\mathbb{R}^n} (\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho - \phi^{m+1} \tilde{f}) \rho dy.
\]

**Proof** On one hand, we have
\[
\int_{\mathbb{R}^n} \tilde{b}^{m+1} (x, y, t) \rho (x, y, t) dy
\]
\[
= \frac{1}{2} \int_{\mathbb{R}^n} \sigma^{m+1} (\sigma^{m+1})' \rho dy + \int_{\mathbb{R}^n} \nabla_x \cdot [(\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho) \rho] dy - \int_{\mathbb{R}^n} \partial_s \rho dy
\]
\[
= \frac{1}{2} \int_{\mathbb{R}^n} \tilde{a}^{m+1,m+1} \rho dy + \nabla_x \cdot \int_{\mathbb{R}^n} (\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho) \rho dy.
\]

On the other hand, it follows from Lemma 8.5, Assumption 4.3, and integration by parts that
\[
\int_{\mathbb{R}^n} \left[ \partial_{x_j} \phi^{m+1} f^j + \partial_{y_j} \phi^{m+1} g^j + \partial_{x_j y_k} \phi^{m+1} h^{jk} \right](x, y, t) \rho (x, y, t) dy
\]
\[
= \partial_{x_j} \int_{\mathbb{R}^n} \phi^{m+1} f^j \rho dy - \int_{\mathbb{R}^n} \phi^{m+1} [\nabla_x \cdot (f \rho) + \nabla_y \cdot (g \rho)] dy + \int_{\mathbb{R}^n} \partial_{x_j y_k} \phi^{m+1} h^{jk} \rho dy
\]
\[
= \partial_{x_j} \int_{\mathbb{R}^n} \phi^{m+1} f^j \rho dy + \int_{\mathbb{R}^n} \phi^{m+1} [\nabla_y \cdot (g \rho) - \nabla_x \cdot (f \rho)] dy
\]
\[
- \int_{\mathbb{R}^n} \phi^{m+1} \nabla_y \cdot [(g + \tilde{g}) \rho] dy
\]
\[
+ \partial_{x_j} \int_{\mathbb{R}^n} \partial_{x_k} \phi^{m+1} h^{jk} \rho dy - \int_{\mathbb{R}^n} \partial_{y_k} \phi^{m+1} \partial_{x_j} (h^{jk} \rho) dy
\]
\[
= \partial_{x_j} \int_{\mathbb{R}^n} \phi^{m+1} f^j \rho dy + \int_{\mathbb{R}^n} \phi^{m+1} f^{m+1} \rho dy + \int_{\mathbb{R}^n} \partial_{y_k} \phi^{m+1} (g^k + \tilde{g}^k) \rho dy
\]
\[
- \partial_{x_j} \int_{\mathbb{R}^n} \phi^{m+1} \partial_{x_k} (h^{jk} \rho) dy - \int_{\mathbb{R}^n} \partial_{y_k} \phi^{m+1} \partial_{x_j} (h^{jk} \rho) dy
\]
Lemma 8.9

This completes the proof. □

\[ \partial x_j \int_{\mathbb{R}^n} \phi^{m+1} (f^j - \partial_{x_k} h^{j k} - h^{j k} \partial_{y_k} \log \rho) \rho dy + \int_{\mathbb{R}^n} \phi^{m+1} f^{m+1} \rho dy \]

\[ + \int_{\mathbb{R}^n} \partial_{y_k} \phi^{m+1} (g^k + \tilde{g}^k - \partial_{x_j} h^{j k} - h^{j k} \partial_{x_j} \log \rho) \rho dy \]

\[ = - \nabla_x \cdot \int_{\mathbb{R}^n} \phi^{m+1} f^\rho \rho dy + \int_{\mathbb{R}^n} \phi^{m+1} f^{m+1} \rho dy + \int_{\mathbb{R}^n} \partial_{y_k} \phi^{m+1} \tilde{m}^{m+1,k} \rho dy \]

Combing the above two equations yields

\[ \tilde{w}^{m+1}(x, t) = \int_{\mathbb{R}^n} (\tilde{b}^{m+1} + \partial x_j \phi^{m+1} f^j + \partial y_j \phi^{m+1} g^j + \partial_{x_j y_k} \phi^{m+1} h^{j k}) \rho dy \]

\[ = \frac{1}{2} \int_{\mathbb{R}^n} (\bar{a}^{m+1,m+1} + 2\phi^{m+1} f^{m+1} + 2 \partial_{y_k} \phi^{m+1} \bar{m}^{m+1,k}) \rho dy \]

\[ + \nabla_x \cdot \int_{\mathbb{R}^n} (\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho - \phi^{m+1} \tilde{f}) \rho dy \]

\[ = \frac{1}{2} \bar{A}^{m+1,m+1} + \nabla_x \cdot \int_{\mathbb{R}^n} (\tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla_x \log \rho - \phi^{m+1} \tilde{f}) \rho dy. \]

This completes the proof.

Lemma 8.9

\[ \bar{A}^{m+1,x} = w + \tilde{w} - \nabla_x \cdot A. \]

**Proof** For any \( 1 \leq j \leq m \), we have

\[ \bar{A}^{m+1,j}(x, t) = \int_{\mathbb{R}^n} (\bar{a}^{m+1,j} + \phi^{m+1} f^j + \phi^j f^{m+1} + \partial_{x_k} \phi^{m+1} h^{j k} + \partial_{y_k} \phi^j \bar{m}^{m+1,k}) \rho dy. \]

It follows from Assumption 4.4 and Lemma 4.5 that \( L_0 \) is symmetric with respect to \( \rho \). Therefore, for any \( 1 \leq i, j \leq m + 1 \), we have

\[ \int_{\mathbb{R}^n} \phi^i f^j \rho dy = -(\phi^i, L_0 \phi^j)_\rho = -(L_0 \phi^i, \phi^j)_\rho = \int_{\mathbb{R}^n} \phi^j f^i \rho dy, \]

\[ \int_{\mathbb{R}^n} \tilde{\phi}^i f^j \rho dy = -(\tilde{\phi}^i, L_0 \phi^j)_\rho = -(L_0 \tilde{\phi}^i, \phi^j)_\rho = \int_{\mathbb{R}^n} \tilde{\phi}^j f^i \rho dy. \quad (20) \]

These two symmetric relations, together with Assumption 4.3 and integration by parts, show that

\[ \bar{A}^{m+1,j}(x, t) = \int_{\mathbb{R}^n} [\bar{a}^{m+1,j} + 2\phi^{m+1} (f^j - \frac{1}{2} \partial_{y_k} h^{j k} - h^{j k} \partial_{y_k} \log \rho) + \partial_{y_k} \phi^j \bar{m}^{m+1,k}] \rho dy \]

\[ = \int_{\mathbb{R}^n} [\bar{a}^{m+1,j} + \phi^{m+1} (f^j - \tilde{f}^j) + \partial_{y_k} \phi^j \bar{m}^{m+1,k}] \rho dy. \]

Moreover, for any \( 1 \leq i \leq m \), using Assumption 4.3, integration by parts, and (20) again yields

\[ w^i + \tilde{w}^i - (\nabla_x \cdot A)^i \]

\[ = \int_{\mathbb{R}^n} (b^i + \partial_{x_k} \phi^i f^j + \partial y_j \phi^i g^j + \partial_{x_j y_k} \phi^i h^{j k} \rho + \tilde{b}^i + \partial_{x_k} \tilde{\phi}^i f^j + \partial y_j \tilde{\phi}^i g^j + \partial_{x_j y_k} \tilde{\phi}^i h^{j k}) \rho dy \]

\[ - \partial_{x_j} \int_{\mathbb{R}^n} (a^{ij} + \phi^i f^j + \phi^j f^i + \partial_{y_k} \phi^i h^{j k} + \partial_{y_k} \phi^j h^{i k}) \rho dy \]
\[ = \int_{\mathbb{R}^n} (\tilde{b}^i + \tilde{b}^j - \partial_{x_j} a^{ij} - a^{ij} \partial_{x_j} \log \rho) \rho dy \]

\[ - \int_{\mathbb{R}^n} \phi^i \partial_{x_j} (f^j \rho) + \partial_{y_j} (g^j \rho) - \partial_{x_j \gamma} (h^{jk} \rho) dy \]

\[ + \int_{\mathbb{R}^n} \partial_{x_j} \tilde{b}^i (\tilde{f}^j - \partial_{x_k} h^{jk} - h^{jk} \partial_{x_k} \log \rho) \rho dy - \int_{\mathbb{R}^n} \tilde{b}^i \partial_{y_j} (\tilde{g}^j \rho) dy \]

\[ - \partial_{x_j} \int_{\mathbb{R}^n} \phi^j (f^i - \partial_{x_k} h^{ik} - h^{ik} \partial_{x_k} \log \rho) \rho dy \]

\[ = \int_{\mathbb{R}^n} \tilde{a}^{m+1,j} \rho dy + \int_{\mathbb{R}^n} \phi^i \partial_{y_j} (\tilde{g}^j \rho) dy \]

\[ - \int_{\mathbb{R}^n} \phi^i \partial_{x_j} (g^k + \tilde{g}^k - \partial_{x_j} h^{jk} - h^{jk} \partial_{x_j} \log \rho) \rho dy \]

\[ - \partial_{x_j} \int_{\mathbb{R}^n} \tilde{b}^i f^j \rho dy - \int_{\mathbb{R}^n} \tilde{b}^i \partial_{y_j} (\tilde{g}^j \rho) dy + \partial_{x_j} \int_{\mathbb{R}^n} \phi^j \tilde{f}^i \rho dy \]

\[ = \int_{\mathbb{R}^n} \tilde{a}^{m+1,i} \rho dy + \int_{\mathbb{R}^n} \phi^i f^{m+1} \rho dy + \int_{\mathbb{R}^n} \partial_{y_k} \phi^i \tilde{h}^{m+1,k} \rho dy - \int_{\mathbb{R}^n} \tilde{b}^i f^{m+1} \rho dy \]

\[ = \int_{\mathbb{R}^n} [\tilde{a}^{m+1,i} + \phi^{m+1} (f^i - \tilde{f}^i) + \partial_{y_k} \phi^i \tilde{h}^{m+1,k}] \rho dy = \tilde{A}^{m+1,i} (x, t). \]

This completes the proof. \(\square\)

**Lemma 8.10**

\[ \nabla_x \cdot \left( \tilde{w} - \frac{1}{2} \nabla_x \cdot A \right) = \nabla_x \cdot \int_{\mathbb{R}^n} \left( \tilde{b} - \frac{1}{2} \nabla_x \cdot a - \frac{1}{2} a \nabla \log \rho - \phi^{m+1} \tilde{f} \right) \rho dy. \]

**Proof** For any \(1 \leq i \leq m\), using Assumption 4.3, integration by parts, and (20) yields

\[ \tilde{w}^i - \frac{1}{2} (\nabla_x \cdot A)^i \]

\[ = \int_{\mathbb{R}^n} (\tilde{b}^i + \partial_{x_i} \phi^j \tilde{f}^j + \partial_{y_j} \phi^j \tilde{g}^j + \partial_{x_j \gamma} \phi^j h^{jk}) \rho dy \]

\[ - \frac{1}{2} \partial_{x_i} \int_{\mathbb{R}^n} (a^{ij} + \phi^i f^j + \phi^j f^i + \partial_{y_k} \phi^j h^{jk} + \partial_{y_k} \phi^j h^{ik}) \rho dy \]

\[ = \int_{\mathbb{R}^n} (\tilde{b}^i - \frac{1}{2} \partial_{x_j} a^{ij} - \frac{1}{2} a^{ij} \partial_{x_j} \log \rho) \rho dy \]

\[ + \int_{\mathbb{R}^n} \partial_{x_i} \phi^j (\tilde{f}^j - \partial_{x_k} h^{jk} - h^{jk} \partial_{x_k} \log \rho) \rho dy - \int_{\mathbb{R}^n} \tilde{b}^i \partial_{x_j} (\tilde{g}^j \rho) dy \]

\[ - \frac{1}{2} \partial_{x_i} \int_{\mathbb{R}^n} \phi^j (f^i - \partial_{x_k} h^{ik} - h^{ik} \partial_{x_k} \log \rho) \rho dy \]

\[ - \frac{1}{2} \partial_{x_i} \int_{\mathbb{R}^n} \phi^j (f^i - \partial_{x_k} h^{ik} - h^{ik} \partial_{x_k} \log \rho) \rho dy \]

\[ = \int_{\mathbb{R}^n} (\tilde{b}^i - \frac{1}{2} \partial_{x_j} a^{ij} - \frac{1}{2} a^{ij} \partial_{x_j} \log \rho) \rho dy - \partial_{x_i} \int_{\mathbb{R}^n} \tilde{b}^i f^j \rho dy \]

\[ - \int_{\mathbb{R}^n} \tilde{b}^i (\partial_{y_j} (\tilde{g}^j \rho) - \partial_{x_j} (f^j \rho)) dy \]

\[ + \frac{1}{2} \partial_{x_i} \int_{\mathbb{R}^n} \phi^j \tilde{f}^i \rho dy + \frac{1}{2} \partial_{x_i} \int_{\mathbb{R}^n} \phi^j \tilde{f}^i \rho dy \]
where we have used the fact that

Straightforward calculations show that

This shows that \( e^{-G_t^{(2)}} \) is an exponential martingale.  

**Acknowledgements** The authors acknowledge George G. Yin and Wenqing Hu for valuable comments and stimulating discussions. H. Ge is supported by NSFC (No. 11971037 and 11622101). C. Jia acknowledges support from the NSAF grant in National Natural Science Foundation of China (NSFC) with Grant No. U1930402.
Appendix A: An Class of Processes Satisfying Assumption 4.7

Clearly, Assumption 4.7 is automatically satisfied if the original process is defined on a compact set. However, in many cases, to guarantee this assumption to hold, we only need to require that the fast component is defined on a compact set. To give such an example, we consider the case where the fast process $Y^\epsilon$ is defined on a compact subset $K \subset \mathbb{R}^m$ and the slow process $X^\epsilon$ is a one-dimensional diffusion defined on the entire space. Moreover, we assume that $X^\epsilon_0 = 0$ and the drift and diffusion coefficients of the slow process satisfy $f = 0$, $\sigma = 1$, and

$$-\alpha x \leq b(x, y, t) \leq 0, \quad x \in \mathbb{R}, \quad y \in K, \quad t \in [0, T],$$

where $\alpha > 0$ is a constant. Under these conditions, it follows from the Ikeda–Watanabe comparison theorem [81, p. 437, Theorem 1.1] that for any $\epsilon > 0$, the slow process can be controlled by

$$U_t \leq X^\epsilon_t \leq W_t, \quad t \in [0, T], \quad (21)$$

where $U_t$ is an Ornstein-Uhlenbeck process satisfying

$$dU_t = -\alpha U_t + dW_t, \quad U_0 = 0.$$

Since $W_t$ is a continuous martingale, it follows from the Burkholder–Davis–Gundy inequality that for any $p > 0$, there exists a constant $D_p > 0$ independent of $T$ such that

$$\mathbb{E}[\sup_{t \leq T} |W_t|^p] \leq D_p T^{p/2}.$$

Recently, it has been proved that the Ornstein-Uhlenbeck process satisfies the following $L^p_{\max}$ inequality [82, 83]: for any $p > 0$, there exists a constant $C_p > 0$ independent of $\alpha$, $\beta$, and $T$ such that

$$\mathbb{E}[\sup_{t \leq T} |U_t|^p] \leq \frac{C_p}{\alpha^{p/2}} \log^{p/2}(1 + \alpha T).$$

For any $M > 0$, it then follows from Chebychev’s inequality that

$$\mathbb{P}(\sup_{t \leq T} |W_t| > M) \leq \frac{1}{M^p} \mathbb{E}[\sup_{t \leq T} |W_t|^p] \leq \frac{D_p}{M^p} T^{p/2},$$

$$\mathbb{P}(\sup_{t \leq T} |U_t| > M) \leq \frac{1}{M^p} \mathbb{E}[\sup_{t \leq T} |U_t|^p] \leq \frac{C_p}{M^p \alpha^{p/2}} \log^{p/2}(1 + \alpha T).$$

Thus for any $\delta > 0$, when $M$ is sufficiently large, it follows from (21) that

$$\mathbb{P}(\inf_{\epsilon > 0} \tau_{B_0(3M)}^\epsilon \geq T) \geq \mathbb{P}(\sup_{t \leq T} |X^\epsilon_t| \leq 2M \text{ for all } \epsilon > 0) \geq \mathbb{P}(\sup_{t \leq T} |W_t| + \sup_{t \leq T} |U_t| \leq 2M) \geq 1 - \mathbb{P}(\sup_{t \leq T} |W_t| > M) - \mathbb{P}(\sup_{t \leq T} |U_t| > M) \geq 1 - \delta,$$

where $B_0(3M) \subset \mathbb{R}^{m+1}$ is the closed ball centered at zero with radius $3M$. Therefore, Assumption 4.7 is satisfied.
References

1. Jarzynski, C.: Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. Annu. Rev. Condens. Matter Phys. 2, 329–351 (2011)
2. Seifert, U.: Stochastic thermodynamics, fluctuation theorems and molecular machines. Rep. Prog. Phys. 75, 126001 (2012)
3. Van den Broeck, C., Esposito, M.: Ensemble and trajectory thermodynamics: a brief introduction. Physica A 418, 6–16 (2015)
4. Esposito, M., Van den Broeck, C.: Three detailed fluctuation theorems. Phys. Rev. Lett. 104, 090601 (2010)
5. Esposito, M., Van den Broeck, C.: Three faces of the second law. I. Master equation formulation. Phys. Rev. E 82, 011143 (2010)
6. Ge, H., Qian, H.: Physical origins of entropy production, free energy dissipation, and their mathematical representations. Phys. Rev. E 81, 051133 (2010)
7. Jiang, D.-Q., Qian, M., Qian, M.-P.: Mathematical Theory of Nonequilibrium Steady States: On the Frontier of Probability and Dynamical Systems. Springer, Berlin (2004)
8. Zhang, X.-J., Qian, H., Qian, M.: Stochastic theory of nonequilibrium steady states and its applications. Part I. Phys. Rep. 510, 1–86 (2012)
9. Hong, L., Jia, C., Zhu, Y., Yong, W.-A.: Novel dissipative properties of the master equation. J. Math. Phys. 57, 103303 (2016)
10. Andrieux, D., Gaspard, P.: Fluctuation theorem for currents and Schnakenberg network theory. J. Stat. Phys. 127, 107–131 (2007)
11. Jia, C., Jiang, D.-Q., Qian, M.-P.: Cycle symmetries and circulation fluctuations for discrete-time and continuous-time Markov chains. Ann. Appl. Probab. 26, 2454–2493 (2016)
12. Ge, H., Jia, C., Jiang, D.-Q.: Cycle symmetry, limit theorems, and fluctuation theorems for diffusion processes on the circle. Stoch. Proc. Appl. 127, 1897–1925 (2017)
13. Evans, D.J., Cohen, E.G., Morriss, G.P.: Probability of second law violations in shearing steady states. Phys. Rev. Lett. 71, 2401 (1993)
14. Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in stationary states. J. Stat. Phys. 80, 931–970 (1995)
15. Jarzynski, C.: Nonequilibrium equality for free energy differences. Phys. Rev. Lett. 78, 2690 (1997)
16. Sekimoto, K.: Stochastic Energetics, vol. 799. Springer, Berlin (2010)
17. Kurchan, J.: Fluctuation theorem for stochastic dynamics. J. Phys. A: Math. Gen. 31, 3719 (1998)
18. Crooks, G.E.: Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. Phys. Rev. E 60, 2721 (1999)
19. Searles, D.J., Evans, D.J.: Fluctuation theorem for stochastic systems. Phys. Rev. E 60, 159 (1999)
20. Lebowitz, J.L., Spohn, H.: A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. J. Stat. Phys. 95, 333–365 (1999)
21. Hatano, T., Sasa, S.-I.: Steady-state thermodynamics of Langevin systems. Phys. Rev. Lett. 86, 3463 (2001)
22. Seifert, U.: Entropy production along a stochastic trajectory and an integral fluctuation theorem. Phys. Rev. Lett. 95, 040602 (2005)
23. Ritort, F.: Nonequilibrium fluctuations in small systems: From physics to biology. Adv. Chem. Phys. 137, 31 (2008)
24. Bonetto, F., Gallavotti, G., Garrido, P.: Chaotic principle: an experimental test. Physica D Nonlinear Phenom. 105, 226–252 (1997)
25. Hendrix, D., Jarzynski, C.: A fast growth method of computing free energy differences. J. Chem. Phys. 114, 5974–5981 (2001)
26. Van Zon, R., Cohen, E.: Extension of the fluctuation theorem. Phys. Rev. Lett. 91, 110601 (2003)
27. Giuliani, A., Zamponi, F., Gallavotti, G.: Fluctuation relation beyond linear response theory. J. Stat. Phys. 119, 909–944 (2005)
28. Williams, S.R., Searles, D.J., Evans, D.J.: Numerical study of the steady state fluctuation relations far from equilibrium. J. Chem. Phys. 124, 194102 (2006)
29. Ciliberto, S., Laroche, C.: An experimental test of the Gallavotti–Cohen fluctuation theorem. J. Phys. IV 8, P6–215 (1998)
30. Liphardt, J., Dumont, S., Smith, S.B., Tinoco, I., Bustamante, C.: Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality. Science 296, 1832–1835 (2002)
31. Garnier, N., Ciliberto, S.: Nonequilibrium fluctuations in a resistor. Phys. Rev. E 71, 060101 (2005)
32. Collin, D., et al.: Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. Nature 437, 231 (2005)
33. Imparato, A., Peliti, L., Pesce, G., Rusciano, G., Sasso, A.: Work and heat probability distribution of an optically driven Brownian particle: theory and experiments. Phys. Rev. E 76, 050101 (2007)
34. Maes, C.: On the origin and the use of fluctuation relations for the entropy. Sémin. Poincaré 2, 29–62 (2003)
35. Chernyak, V.Y., Chertkov, M., Jarzynski, C.: Path-integral analysis of fluctuation theorems for general Langevin processes. J. Stat. Mech. Theory Exp. 2006, P08001 (2006)
36. Chetrite, R., Gawedzki, K.: Fluctuation relations for diffusion processes. Commun. Math. Phys. 282, 469–518 (2008)
37. Chetrite, R., Gupta, S.: Two refreshing views of fluctuation theorems through kinematics elements and exponential martingale. J. Stat. Phys. 143, 543 (2011)
38. Neri, I., Roldán, É., Jülicher, F.: Statistics of infima and stopping times of entropy production and applications to active molecular processes. Phys. Rev. X 7, 011019 (2017)
39. Pigolotti, S., Neri, I., Roldán, É., Jülicher, F.: Generic properties of stochastic entropy production. Phys. Rev. Lett. 119, 140604 (2017)
40. Chetrite, R., Gupta, S., Neri, I., Roldán, É.: Martingale theory for housekeeping heat. Europhys. Lett. 124, 60006 (2019)
41. Singh, S., et al.: Records of entropy production in an electronic double dot. arXiv preprint (2017). arXiv:1712.01693
42. Pavliotis, G.: Stochastic Processes and Applications. Springer, Berlin (2014)
43. Jia, C., Grima, R.: Frequency domain analysis of fluctuations of mRNA and protein copy numbers within a cell lineage: theory and experimental validation. Phys. Rev. X 11, 021032 (2021)
44. Pavliotis, G., Stuart, A.: Multiscale Methods: Averaging and Homogenization. Springer, Berlin (2008)
45. Pigolotti, S., Vulpiani, A.: Coarse graining of master equations with fast and slow states. J. Chem. Phys. 128, 154114 (2008)
46. Jia, C.: Simplification of irreversible Markov chains by removal of states with fast leaving rates. J. Theor. Biol. 400, 129–137 (2016)
47. Jia, C.: Simplification of Markov chains with infinite state space and the mathematical theory of random gene expression bursts. Phys. Rev. E 96, 032402 (2017)
48. Khasminskii, R.Z., Yin, G.: On transition densities of singularly perturbed diffusions with fast and slow components. SIAM J. Appl. Math. 56, 1794–1819 (1996)
49. Khasminskii, R.Z., Yin, G.: On averaging principles: an asymptotic expansion approach. SIAM. J. Math. Anal. 35, 1534–1560 (2004)
50. Khasminskii, R.: Stochastic Stability of Differential Equations, vol. 66. Springer, Berlin (2011)
51. Jia, C.: Reduction of Markov chains with two-time-scale state transitions. Stochastics 88, 73–105 (2016)
52. Jia, C., Zhang, M.Q., Qian, H.: Emergent Lévy behavior in single-cell stochastic gene expression. Phys. Rev. E 96, 040402 (2017)
53. Jia, C., Yin, G.G., Zhang, M.Q., et al.: Single-cell stochastic gene expression kinetics with coupled positive-plus-negative feedback. Phys. Rev. E 100, 052406 (2019)
54. Jia, C., Grima, R.: Small protein number effects in stochastic models of autoregulated bursty gene expression. J. Chem. Phys. 152, 084115 (2020)
55. Jia, C., Grima, R.: Dynamical phase diagram of an auto-regulating gene in fast switching conditions. J. Chem. Phys. 152, 174110 (2020)
56. Jia, C.: Kinetic foundation of the zero-inflated negative binomial model for single-cell RNA sequencing data. SIAM J. Appl. Math. 80, 1336–1355 (2020)
57. Spinney, R.E., Ford, I.J.: Entropy production in full phase space for continuous stochastic dynamics. Phys. Rev. E 85, 051113 (2012)
58. Lee, H.K., Kwon, C., Park, H.: Fluctuation theorems and entropy production with odd-parity variables. Phys. Rev. Lett. 110, 050602 (2013)
59. Gè, H.: Time reversibility and nonequilibrium thermodynamics of second-order stochastic processes. Phys. Rev. E 89, 022127 (2014)
60. Pardoux, E., Veretennikov, Y., et al.: On the Poisson equation and diffusion approximation. I. Ann. Probab. 29, 1061–1085 (2001)
61. Pardoux, E., Veretennikov, A.Y., et al.: On Poisson equation and diffusion approximation 2. Ann. Probab. 31, 1166–1192 (2003)
62. Khasminskii, R., Yin, G.: Limit behavior of two-time-scale diffusions revisited. J. Differ. Equ. 212, 85–113 (2005)
63. Esposito, M.: Stochastic thermodynamics under coarse graining. Phys. Rev. E 85, 041125 (2012)
64. Kawaguchi, K., Nakayama, Y.: Fluctuation theorem for hidden entropy production. Phys. Rev. E 88, 022147 (2013)
65. Puglisi, A., Pigolotti, S., Rondoni, L., Vulpiani, A.: Entropy production and coarse graining in Markov processes. J. Stat. Mech. Theory Exp. 2010, P05015 (2010)
66. Rahav, S., Jarzynski, C.: Fluctuation relations and coarse-graining. J. Stat. Mech. Theory Exp. 2007, P09012 (2007)
67. Santillán, M., Qian, H.: Irreversible thermodynamics in multiscale stochastic dynamical systems. Phys. Rev. E 83, 041130 (2011)
68. Nakayama, Y., Kawaguchi, K.: Invariance of steady-state thermodynamics between different scales of description. Phys. Rev. E 91, 012115 (2015)
69. Ford, I.J., Laker, Z.P., Charlesworth, H.J.: Stochastic entropy production arising from nonstationary thermal transport. Phys. Rev. E 92, 042108 (2015)
70. Celani, A., Bo, S., Eichhorn, R., Aurell, E.: Anomalous thermodynamics at the microscale. Phys. Rev. Lett. 109, 260603 (2012)
71. Bo, S., Celani, A.: Entropy production in stochastic systems with fast and slow time-scales. J. Stat. Phys. 154, 1325–1351 (2014)
72. Lan, Y., Aurell, E.: The stochastic thermodynamics of a rotating Brownian particle in a gradient flow. Sci. Rep. 5, 12266 (2015)
73. Jia, C.: Model simplification and loss of irreversibility. Phys. Rev. E 93, 052149 (2016)
74. Marino, R., Eichhorn, R., Aurell, E.: Entropy production of a Brownian ellipsoid in the overdamped limit. Phys. Rev. E 93, 012132 (2016)
75. Bo, S., Celani, A.: Multiple-scale stochastic processes: decimation, averaging and beyond. Phys. Rep. 670, 1–59 (2017)
76. Birrell, J.: Entropy anomaly in Langevin–Kramers dynamics with a temperature gradient, matrix drag, and magnetic field. J. Stat. Phys. 173, 1549–1586 (2018)
77. Hu, W., Spiliopoulos, K., et al.: Hypoelliptic multiscale Langevin diffusions: large deviations, invariant measures and small mass asymptotics. Electron. J. Probab. 22, 1–38 (2017)
78. Speck, T., Seifert, U.: Integral fluctuation theorem for the housekeeping heat. J. Phys. A Math. Gen. 38, L581 (2005)
79. Herzog, D.P., Mattingly, J.C.: A practical criterion for positivity of transition densities. Nonlinearity 28, 2823 (2015)
80. Jacod, J., Shiryaev, A.: Limit Theorems for Stochastic Processes, 2nd edn. Springer, New York (2002)
81. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, 2nd edn. North-Holland, Amsterdam (1989)
82. Jia, C., Zhao, G.: Moderate maximal inequalities for the Ornstein–Uhlenbeck process. Proc. Am. Math. Soc. 148, 3607–3615 (2020)
83. Chen, X., Jia, C.: Identification of unstable fixed points for randomly perturbed dynamical systems with multistability. J. Math. Anal. Appl. 446, 521–545 (2017)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.