The Annals of Probability
2003, Vol. 31, No. 3, 1254–1295

SMALL-TIME GAUSSIAN BEHAVIOR OF SYMMETRIC DIFFUSION SEMIGROUPS

BY MASANORI HINO[1] AND JOSÉ A. RAMÍREZ[2]

Kyoto University and Cornell University

This work is involved with the short-time asymptotics of diffusion semigroups in a general setting. A generalization of Fang's version of Varadhan's formula is proven for general Dirichlet spaces that are local and conservative. The intrinsic metric appearing in the formula is characterized by pointwise distance for canonical Dirichlet spaces on loop groups.

1. Introduction. Recall Varadhan's formula for the heat kernel density $p_t(x, y)$ on a Riemannian manifold:

$$\lim_{t\downarrow 0} t \log p_t(x, y) = -\frac{d(x, y)^2}{2},$$

where $d(x, y)$ is the Riemannian distance. One can integrate the kernel over two positive measure sets in the form

$$P_t(A, B) = \int_A \int_B p_t(x, y)\, d\mu(y)\, d\mu(x),$$

where μ is the volume measure on the manifold. Then, under mild assumptions on A and B, a Laplace-type estimation gives that

$$\lim_{t\downarrow 0} t \log P_t(A, B) = -\frac{d(A, B)^2}{2},$$

where the distance between two sets is defined in the natural way by means of infimums. The proof of this last formula in a very general setting is the main concern of this article. We start by describing the framework.

Let $(\Omega, \mathscr{B}, \mu)$ be a probability space and $L^p = L^p(\Omega, \mu)$, $p \in [1, \infty]$, the corresponding L^p-space with norm $\|\cdot\|_{L^p}$. The inner product on L^2 will be denoted by $(\cdot,\cdot)_{L^2}$. We consider a Dirichlet form $\mathscr{E}$ with domain $\mathbb{D} \subset L^2$. That is, $(\mathscr{E}, \mathbb{D})$ is a densely defined, positive semidefinite and symmetric bilinear closed form satisfying that if $f \in \mathbb{D}$ then $f \wedge 1 \in \mathbb{D}$ and $\mathscr{E}(f \wedge 1, f \wedge 1) \le \mathscr{E}(f, f)$. We further assume that this Dirichlet form is conservative and local. Namely,

Received February 2002; revised July 2002.

[1]Research supported in part by Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Encouragement of Young Scientists, 13740095, 2001.

[2]Research supported in part by Liftoff employment from the Clay Mathematical Institute.

AMS 2000 subject classifications. Primary 31C25, 60J60, 47D07; secondary 58J65.

Key words and phrases. Dirichlet spaces, heat kernel, short-time asymptotics, intrinsic metric, loop group.

- $1 \in \mathbb{D}$ and $\mathcal{E}(1, 1) = 0$;
- for C^∞-functions F, G on $\mathbb{R}$ with compact support and $\operatorname{supp} F \cap \operatorname{supp} G = \varnothing$, $\mathcal{E}(F(f), G(f)) = 0$ for any $f \in \mathbb{D}$.

The latter condition has several equivalent expressions. For example, it can be replaced by

- if $f, g \in \mathbb{D}$ satisfies $fg = 0$ a.e., then $\mathcal{E}(f, g) = 0$;
- $\mathcal{E}(|f|, |f|) = \mathcal{E}(f, f)$ for any $f \in \mathbb{D}$.

See [7] and [20] for excellent references in these matters. Typically, the Dirichlet form given as the energy integral on Ω satisfies the above conditions. We note that when $(\mathcal{E}, \mathbb{D})$ is a quasiregular Dirichlet form the corresponding Markov process is a conservative diffusion process. We do not assume, however, the existence of such a probabilistic counterpart throughout this paper.

The Markovian semigroup and the nonpositive generator associated with $(\mathcal{E}, \mathbb{D})$ will be denoted by $\{T_t\}$ and $\mathcal{L}$, respectively. Let $\mathbb{D}_b = \mathbb{D} \cap L^\infty$. The space $\mathbb{D}_b$ is an algebra (see, e.g., Corollary 3.3.2 in [7]). The functional $I : \mathbb{D}_b \times \mathbb{D}_b \times \mathbb{D}_b \ni (f, g, h) \mapsto I_{f,g}(h) \in \mathbb{R}$ given by

$$I_{f,g}(h) = I(f, g; h) = \mathcal{E}(fh, g) + \mathcal{E}(gh, f) - \mathcal{E}(fg, h)$$

will be of extreme importance in what follows. We sometimes write $I_f(h) = I(f; h) = I_{f,f}(h)$. The main properties of these functionals that are used in this work are collected in Section 2.2. Define a subset $\mathbb{D}_0$ of $\mathbb{D}_b$ as

$$\mathbb{D}_0 = \{f \in \mathbb{D}_b \mid I_f(h) \le \|h\|_{L^1} \text{ holds for any } h \in \mathbb{D}_b\}. \tag{1.1}$$

With this, we define the intrinsic metric d as given by

$$\mathsf{d}(A, B) = \sup_{f \in \mathbb{D}_0} \left\{ \operatorname*{essinf}_{x \in B} f(x) - \operatorname*{esssup}_{y \in A} f(y) \right\} \tag{1.2}$$

for any two measurable sets A and B. Here we take a natural convention that $\sup \varnothing = -\infty$ and $\inf \varnothing = \infty$. Define

$$P_t(A, B) = \int_A T_t \mathbb{1}_B \, d\mu.$$

We want to study the short-time behavior of $P_t(A, B)$ in a logarithmic scale. The present work is intended to prove the following theorem.

THEOREM 1.1. *Under the above conditions, it is true that*

$$\lim_{t \downarrow 0} t \log P_t(A, B) = -\frac{\mathsf{d}(A, B)^2}{2} \tag{1.3}$$

for all measurable sets A and B.

The upper bound in Theorem 1.1 is virtually well known and follows from an argument introduced originally by Gaffney in [21]. This will be done in Section 2.4. The main trouble resides, then, in showing that the converse inequality also holds. That is, we will show that

$$\liminf_{t\downarrow 0} t\log P_t(A,B) \ge -\frac{\mathsf{d}(A,B)^2}{2} \tag{1.4}$$

holds. Sections 2.5 and 2.6 form the core of the proof of this fact.

This problem has been investigated thoroughly in many particular cases. As mentioned in the beginning, Theorem 1.1 follows from a stronger (pointwise) formula first established by Varadhan [44] in the finite-dimensional case (for more on this subject see Section 3.1). It was first proven to hold in an infinite-dimensional example by Fang [17], who considered the Ornstein–Uhlenbeck process on Wiener space. Recent work includes [3, 4, 18, 19, 45]. The previous work of the authors ([25] and [36]) dealt with the same problem in this general setting but under certain restrictions. This article intends to prove the estimate in the general case.

We recall the main ideas in the argument of [36]. Suppose that Ω has a differential structure and a gradient operator ∇ taking values in a Hilbert space with inner product $\langle\cdot,\cdot\rangle$ like finite-dimensional Riemannian manifolds and that $\mathcal{E}$ is expressed as $\mathcal{E}(f,g)=\frac{1}{2}\int_\Omega\langle\nabla f,\nabla g\rangle\,d\mu$. Then a simple calculation shows that $I(f,g;h)=\int_\Omega h\langle\nabla f,\nabla g\rangle\,d\mu$ and $\mathbb{D}_0=\{f\in\mathbb{D}_b \mid |\nabla f|\le 1 \text{ a.e.}\}$. Also, the function $u_t=-t\log T_t\mathbb{1}_A$ satisfies the following partial differential equation:

$$t(\partial_t u_t-\mathcal{L}u_t)=u_t-\tfrac{1}{2}|\nabla u_t|^2. \tag{1.5}$$

Thus, when we let $t\downarrow 0$, we expect that $|\nabla u_0|^2=2u_0$ or, what we can actually prove, $|\nabla u_0|^2\le 2u_0$. This implies that $|\nabla\sqrt{2u_0}|^2\le 1$. In other words, since it seems that $u_0=0$ on A, we find that

$$\lim_{t\downarrow 0}\sqrt{-2t\log T_t\mathbb{1}_A(x)}\le \mathsf{d}_A(x)$$

(with an appropriate definition of d_A), which is an instance of the result we want. Sadly, this simple argument will be very much obscured with technicalities in the main body of this paper.

In [36], L^2-methods are used to take limits in (1.5). However, in the absence of a spectral gap, that argument breaks down because we cannot even prove that $u_t\in L^2$. The present work uses a simple idea to overcome such difficulty. Roughly speaking, what we do is replace u_t by $u_t\wedge M$ ($M>0$) so that being in L^2 is no longer an issue. Section 2.1 deals with some basic definitions and properties regarding the cutoff functions used in the proof.

In Section 2.3, the distance function d_A referred to in the previous heuristic argument will be shown to exist and satisfy some properties. The following result will be proven.

THEOREM 1.2. *Let A be a positive measure set. There exists an (a.e.) unique $[0,\infty]$-valued measurable function d_A such that*

- $\mathsf{d}_A \wedge N \in \mathbb{D}_0$ *for any* $N \geq 0$;
- $\mathsf{d}_A = 0$ *a.e. on* A;
- d_A *is the (a.e.) largest function that satisfies the two previous requirements.*

Moreover, if B is another measurable set, then

$$\mathsf{d}(A,B) = \operatorname*{essinf}_{x\in B} \mathsf{d}_A(x).$$

Based on the described heuristic argument, it seems reasonable that we are able to prove a somewhat stronger result than Theorem 1.1. The best we can do is the following.

THEOREM 1.3. *The sequence of functions $u_t = -t\log T_t \mathbb{1}_A$ converges to $\mathsf{d}_A^2/2$ as $t \downarrow 0$ in the following senses:*

(i) $u_t \cdot \mathbb{1}_{\{u_t<\infty\}}$ *converges to* $\mathsf{d}_A^2/2 \cdot \mathbb{1}_{\{\mathsf{d}_A<\infty\}}$ *in probability.*

(ii) *If F is a bounded function on $[0,\infty]$ that is continuous on $[0,\infty)$, then $F(u_t)$ converges to $F(\mathsf{d}_A^2/2)$ in L^2.*

The final blow in the proof of this theorem is presented in Section 2.7. However it is very much based on the work of Sections 2.5 and 2.6.

Section 3 consists of some remarks related to the known results about asymptotic estimates on finite-dimensional spaces.

Our original motivation about Theorem 1.1 is to study the behavior of canonical diffusions on infinite-dimensional spaces with differential structure, such as path or loop spaces on Riemannian manifolds. Since the transition kernels are singular with respect to the underlying measure in general, the formulation in Theorem 1.1 is quite natural. Moreover, we expect that $\mathsf{d}(A,B)$ (or d_A) is derived from a pointwise distance. Such "distance," however, should not be compatible with the original topology of the underlying space. In order to look more closely at this aspect, we briefly review the result on path spaces over compact Lie groups due to Aida and Zhang [4], leaving some precise definitions in Section 4.

Let G be a connected and simply connected compact Lie group and e its unit element. Since G can be embedded in some general linear group, we may and will assume that G is a matrix group. The Lie algebra $\mathfrak{g} \equiv T_eG$ is identified with the totality of left-invariant vector fields on G. We fix an Ad_G-invariant inner product $\langle\cdot,\cdot\rangle_{\mathfrak{g}}$ and a positive number T. The based path space $\mathcal{P}G$ is given by

$$\mathcal{P}G = \{g \in C([0,T]\to G) \mid g(0) = e\},$$

which has a group structure by pointwise multiplication. Define

$$\mathcal{P}\mathfrak{g} = \{h \in C([0,T]\to \mathfrak{g}) \mid h(0) = 0\}$$

and its subspace

$$H = \left\{ h \in \mathcal{P}\mathfrak{g} \,\middle|\, h \text{ is absolutely continuous and } \int_0^T |\dot{h}(s)|_{\mathfrak{g}}^2 \, ds < \infty \right\}.$$

For $h \in \mathcal{P}\mathfrak{g}$, we define $e^h \in \mathcal{P}G$ by $e^h(t) = e^{h(t)}$, $t \in [0, T]$. Here H is a separable Hilbert space under the inner product $(h_1|h_2) = \int_0^T \langle \dot{h}_1(t), \dot{h}_2(t) \rangle_{\mathfrak{g}} \, dt$, which is regarded as a tangent space on each point of $\mathcal{P}G$.

Let $\mathfrak{C}$ be the set of smooth cylindrical functions on $\mathcal{P}G$ defined in (4.1). For $f \in \mathfrak{C}$ and $h \in H$, we set $\partial_h f(g) = \frac{d}{ds} f(e^{sh} g)|_{s=0}$. To each $f \in \mathfrak{C}$, we associate a unique H-valued function $\nabla^{\mathcal{P}} f$ on $\mathcal{P}G$ such that $(\nabla^{\mathcal{P}} f(g)|h) = \partial_h f(g)$ for every $g \in \mathcal{P}G$ and $h \in H$.

A pre-Dirichlet form

$$\mathcal{E}^{\mathcal{P}}(f_1, f_2) = \frac{1}{2} \int_{\mathcal{P}G} (\nabla^{\mathcal{P}} f_1 | \nabla^{\mathcal{P}} f_2) \, d\mu, \qquad f_1, f_2 \in \mathfrak{C},$$

where μ is the Brownian motion measure, is known to be closable. Denote its closure by $(\mathcal{E}^{\mathcal{P}}, \mathcal{F}^{\mathcal{P}})$, which is a conservative and local Dirichlet form. In order to describe the intrinsic metric $\mathsf{d}^{\mathcal{P}}(A, B)$ and the intrinsic distance function $\mathsf{d}_A^{\mathcal{P}}$ associated with $(\mathcal{E}^{\mathcal{P}}, \mathcal{F}^{\mathcal{P}})$ in terms of the geometry of the underlying space, define, for $g_1, g_2 \in \mathcal{P}G$,

$$d^{\mathcal{P}}(g_1, g_2) = \begin{cases} \left(\int_0^T |v(t)^{-1} \dot{v}(t)|_{\mathfrak{g}}^2 \, dt \right)^{1/2}, & \\ & \text{if } v := g_1 g_2^{-1} \text{ is absolutely continuous,} \\ \infty, & \text{otherwise.} \end{cases}$$

This is regarded as the energy of the path $g_1 g_2^{-1}$. Note that $d^{\mathcal{P}}(g_1, g_2) = d^{\mathcal{P}}(g_2, g_1) = d^{\mathcal{P}}(g_1 g_2^{-1}, \mathbf{e})$ from Ad_G-invariance of the inner product of $\mathfrak{g}$, where $\mathbf{e}$ is a constant path taking the value e. A subset A of $\mathcal{P}G$ is called $d^{\mathcal{P}}$-open if each $g \in A$ has a constant $r > 0$ such that $g' \in A$ for every $g' \in \mathcal{P}G$ with $d^{\mathcal{P}}(g', g) < r$.

For a subset A of $\mathcal{P}G$, we define

$$\bar{\mathsf{d}}_A^{\mathcal{P}}(g) = \inf_{g' \in A} d^{\mathcal{P}}(g, g'), \qquad g \in \mathcal{P}G.$$

Here we set $\inf \varnothing = \infty$ as usual.

THEOREM 1.4 ([4]; see also [25]). *Let A be a Borel set of $\mathcal{P}G$. Then $\bar{\mathsf{d}}_A^{\mathcal{P}}$ is universally measurable and $\mathsf{d}_A^{\mathcal{P}} \geq \bar{\mathsf{d}}_A^{\mathcal{P}}$ μ-a.e. Moreover, $\mathsf{d}_A^{\mathcal{P}} = \bar{\mathsf{d}}_A^{\mathcal{P}}$ μ-a.e. if A is $d^{\mathcal{P}}$-open and $\mu(A) > 0$. In particular, for Borel sets A, B of $\mathcal{P}G$,*

$$\mathsf{d}^{\mathcal{P}}(A, B) = \max \left\{ \operatorname*{essinf}_{g \in B} \bar{\mathsf{d}}_A^{\mathcal{P}}(g), \operatorname*{essinf}_{g \in A} \bar{\mathsf{d}}_B^{\mathcal{P}}(g) \right\}$$

if A or B is $d^{\mathcal{P}}$-open.

Note that A is $d^{\mathcal{P}}$-open if A is open with respect to the uniform topology on $\mathcal{P}G$. Theorem 1.4 indicates that the intrinsic metric is governed by the energy of the path, which is not compatible with the original topology. This is not surprising since the gradient $\nabla^{\mathcal{P}}$ is considered only in the directions along H. The assumptions of A (or B) in the theorem cannot be removed completely, because $\mathsf{d}_A^{\mathcal{P}}$ does not change when A is replaced by A' which is equal to A, μ-a.e., while $\bar{\mathsf{d}}_A^{\mathcal{P}}$ may do.

Unlike finite-dimensional Riemannian manifolds, such identification is not confirmed straightforwardly in general. For example, when G is replaced by general Riemannian manifolds, similar claims have not been proven yet. This is partly because we do not have enough information of the domain of the (canonical) Dirichlet form, which is related to the fact that we do not have appropriate mollifiers to smooth the measurable functions. (Path spaces over Lie groups are exceptional cases; the natural semigroup plays the role of a satisfactory mollifier.) The situation seems worse for the loop space case.

In Section 4, we identify the intrinsic distance in the case of loop spaces over compact Lie groups, solving the difficulty mentioned above. Let G be the same as above. The based loop group $\mathcal{L}G$ is a subgroup of $\mathcal{P}G$ given by

$$\mathcal{L}G = \{g \in \mathcal{P}G \mid g(T) = e\}.$$

Let

$$\mathcal{L}\mathfrak{g} = \{h \in \mathcal{P}\mathfrak{g} \mid h(T) = 0\}, \qquad H_0 = H \cap \mathcal{L}\mathfrak{g},$$

where H_0 is regarded as a tangent space of $\mathcal{L}G$. When $\mathcal{C}$ is considered as a function space on $\mathcal{L}G$, each $f \in \mathcal{C}$ has a unique H_0-valued function $\nabla^{\mathcal{L}} f$ on $\mathcal{L}G$ such that $(\nabla^{\mathcal{L}} f(g)|h) = \partial_h f(g)$ for every $g \in \mathcal{L}G$ and $h \in H_0$. Let ν be a Borel probability measure on $\mathcal{L}G$ satisfying conditions (M0), (M1) and (M2) given in Section 4.1. Two important probabilities on $\mathcal{L}G$, the pinned Brownian motion measure and the heat kernel measure, satisfy these conditions (Proposition 4.2). Define a pre-Dirichlet form $(\mathcal{E}^{\mathcal{L}}, \mathcal{C})$ by

$$\mathcal{E}^{\mathcal{L}}(f_1, f_2) = \frac{1}{2} \int_{\mathcal{L}G} (\nabla^{\mathcal{L}} f_1 | \nabla^{\mathcal{L}} f_2)\, d\nu, \qquad f_1, f_2 \in \mathcal{C}, \tag{1.6}$$

and denote its closure by $(\mathcal{E}^{\mathcal{L}}, \mathcal{F}^{\mathcal{L}})$. The corresponding intrinsic metric or distance function will be denoted by $\mathsf{d}^{\mathcal{L}}(A, B)$ or $\mathsf{d}_A^{\mathcal{L}}$ in the same way as in $\mathcal{P}G$.

We define by (4.4) the shortest path metric $d^{\mathcal{L}}(\cdot, \cdot)$ on $\mathcal{L}G$ induced by $d^{\mathcal{P}}$. A subset A of $\mathcal{L}G$ is called $d^{\mathcal{L}}$-open if every $g \in A$ has a constant $r > 0$ such that $g' \in A$ for all $g' \in \mathcal{L}G$ with $d^{\mathcal{L}}(g', g) < r$. Open sets in the uniform topology are $d^{\mathcal{L}}$-open (see Remark 4.9).

For a subset A of $\mathcal{L}G$, we set

$$\bar{\mathsf{d}}_A^{\mathcal{L}}(g) = \inf_{g' \in A} d^{\mathcal{L}}(g, g'), \qquad g \in \mathcal{L}G.$$

Let us recall that a Suslin space is a metrizable space that is a continuous image of a certain Polish space, and a subset A of a metrizable space is called a Suslin set if the subspace A endowed with a relative topology is a Suslin space. All Borel sets are Suslin sets and all Suslin sets are universally measurable. (See, e.g., [8] for the proofs.)

Now, our claim is as follows.

THEOREM 1.5. *Assume that* ν *satisfies conditions* (M0), (M1) *and* (M2) *in Section* 4.1. *Let* A *be a Suslin set of* $\mathcal{L}G$. *Then* $\bar{\mathsf{d}}_A^{\mathcal{L}}$ *is universally measurable and* $\mathsf{d}_A^{\mathcal{L}} \ge \bar{\mathsf{d}}_A^{\mathcal{L}}$, ν-*a.e. If* A *is* $d^{\mathcal{L}}$-*open in addition, then* $\mathsf{d}_A^{\mathcal{L}} = \bar{\mathsf{d}}_A^{\mathcal{L}}$, ν-*a.e. In particular, for Suslin sets* A, B *of* $\mathcal{L}G$,

$$\mathsf{d}^{\mathcal{L}}(A, B) = \max\left\{\operatorname*{essinf}_{g\in B} \bar{\mathsf{d}}_A^{\mathcal{L}}(g), \operatorname*{essinf}_{g\in A} \bar{\mathsf{d}}_B^{\mathcal{L}}(g)\right\}$$

if A *or* B *is* $d^{\mathcal{L}}$-*open. In other words, if* A *is Suslin and* $d^{\mathcal{L}}$-*open and* B *is measurable, then*

$$\mathsf{d}^{\mathcal{L}}(A, B) = \operatorname*{essinf}_{g\in B}\left[\inf_{g'\in A} d^{\mathcal{L}}(g, g')\right].$$

Theorems 1.1 and 1.5 express that the Varadhan estimates capture the natural metric on $\mathcal{L}G$ as a submanifold of $\mathcal{P}G$. For the proof of Theorem 1.5, much effort is devoted to prove the Rademacher theorem of the following version (Theorem 4.18): if a bounded measurable function of f on $\mathcal{L}G$ is $d^{\mathcal{L}}$-Lipschitz with Lipschitz constant 1, then $f \in \mathcal{F}^{\mathcal{L}}$ and $\|\nabla^{\mathcal{L}} f\| \le 1$, ν-a.e. Because we do not have good mollifiers on $\mathcal{L}G$, unlike on $\mathcal{P}G$, we follow the idea of Gross [22] to extend the domain of the function f to $\mathcal{P}G$ and reduce the problem to the analysis on $\mathcal{P}G$, which is easier to handle. The localization argument needs the quasi-sure analysis in our proof, so the case when ν is a pinned Brownian motion μ_e is the most suitable from the viewpoint of our proof. When ν is a heat kernel measure, we utilize the fact that it is absolutely continuous with respect to μ_e, which was proven by Driver and Srimurthy [14].

2. Proofs of Theorems 1.1–1.3.

2.1. *Definitions and cutoffs.* For technical reasons, we do not work directly with $u_t \wedge M$. Instead, consider a concave function $g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying the following properties:

- $g(x)$ is bounded and three times continuously differentiable;
- $g(x) = x$ for $x \le 1$ and $0 < g'(x) \le 1$ for any $x \in [0, \infty)$;
- there is a positive constant C such that $0 \le -g''(x) \le Cg'(x)$ for all $x \ge 0$.

Notice that these conditions imply that $\lim_{x\uparrow\infty} g(x) = L$ exists and the convergence is monotone. For example, a smooth function g such that $g(0) = 0$, g' is nonincreasing and

$$g'(x) = \begin{cases} 1, & \text{if } 0 \le x \le 1, \\ e^{-x}, & \text{if } x \ge 2, \end{cases}$$

will satisfy the requirements.

Define our main cutoff function at level K by $\phi^K(x) = Kg(x/K)$. To simplify the notation, we do not show the dependence on K explicitly for most of the present work. That is, we use ϕ for ϕ^K whenever the value of K is clear from the context. The following functions will also play a prominent role, so we give them names. Let

$$\Phi(x) = \int_0^x \phi'(s)^2\,ds, \qquad \Psi(x) = x\phi'(x)^2, \qquad \Xi(x) = \sqrt{\Phi(x)+1}.$$

From the conditions above, we have the following estimates:

$$\begin{aligned} &0 < \phi'(x) \le 1, \qquad\qquad 0 \le -\phi''(x) \le \frac{C}{K}\phi'(x), \\ &0 \le \Psi(x) \le \Phi(x) \le \int_0^x \phi'(s)\,ds = \phi(x) \le LK, \\ &\Psi(x) = \Phi(x) = x \qquad \text{on } [0, K]. \end{aligned} \tag{2.1}$$

All these functions extend continuously on $[0, \infty]$. This is clearly true for ϕ, Φ and Ξ. Regarding Ψ, notice that

$$\Psi(x) \le \frac{1}{x}\left(\int_0^x \phi'(s)\,ds\right)^2 = \frac{\phi(x)^2}{x} \to 0 \qquad \text{as } x \to \infty.$$

Notice also that

$$\Phi^K(x) - \Phi^K\big(\Phi^M(x)\big) \le \Phi^K(\infty) - \Phi^K(M), \tag{2.2}$$

since the inequality is trivial when $0 \le x \le M$ and it is implied by $\Phi^K(x) \le \Phi^K(\infty)$ and $\Phi^K(\Phi^M(x)) \ge \Phi^K(M)$ when $x > M$.

As mentioned earlier, instead of working directly with $u_t (= -t\log T_t\mathbb{1}_A)$, we are going to study the behavior of the function $\phi_t = \phi(u_t)$ as $t \downarrow 0$. The advantage in considering this function is that, trivially, $0 \le \phi_t \le LK$ a.e. and therefore the collection $\{\phi_t\}_{t>0}$ is bounded in L^2.

The following result will be needed later.

LEMMA 2.1. *The functions $F(x) = \Phi(-t\log x)$ and $G(x) = \Xi(-t\log x)$ are convex for $x \in [0, 1]$ if $t > 0$ is sufficiently small.*

PROOF. Compute

$$F'(x) = -\frac{t}{x}\Phi'(-t\log x),$$

$$F''(x) = \frac{t}{x^2}\Phi'(-t\log x) + \frac{t^2}{x^2}\Phi''(-t\log x)$$
$$= \frac{t}{x^2}\big((\phi')^2 + 2t\phi'\phi''\big)\big|_{-t\log x}.$$

Thus, we can choose t small enough to have a positive second derivative because of (2.1).

The convexity of G is proven in the same way. We have

$$G''(x) = \frac{t}{4x^2}\left[\frac{1}{\Xi}\left(2(\phi')^2 + t\left(4\phi'\phi'' - \frac{(\phi')^4}{\Phi+1}\right)\right)\right]_{-t\log x}.$$

Again, the properties stated in (2.1) imply that the right-hand side is positive if $t > 0$ is small. □

Finally, for lack of a better place, we include here a little lemma that will be very useful later on.

LEMMA 2.2. *Suppose that F is a concave upper semicontinuous function defined on $\mathbb{R}$. Then, seen as a map from L^2 to L^2, it defines a weakly semicontinuous function. That is, if $f_n \rightharpoonup f$ weakly in L^2 and $F(f_n) \in L^2$ for each n, then $F^* \le F(f)$ a.e. for any weak limit F^* of $F(f_n)$.*

PROOF. Notice that $f \mapsto -(F(f), g)_{L^2}$ is a (strongly) lower semicontinuous convex map if $g \ge 0$. Therefore, Corollary 3.8 in [9] applies and completes the proof. □

2.2. *Properties of $I_{f,g}(h)$.* First of all, notice that $I(\cdot,\cdot;\cdot)$ is a linear functional of each argument. The following lemma is a consequence of the representation given as in Proposition 1.2.3.3 of [7].

LEMMA 2.3. *Let f, g, h, h_1 and h_2 be in $\mathbb{D}_b$.*

(i) *If $h_1 \le h_2$ a.e., then $I_f(h_1) \le I_f(h_2)$, in particular, $I_f(h) \ge 0$ if $h \ge 0$ a.e.*;
(ii) $I_{f,g}(h)^2 \le I_f(|h|)I_g(|h|)$;
(iii) $I_{f,g}(h_1h_2)^2 \le I_f(h_1^2)I_g(h_2^2)$;
(iv) $\sqrt{I_{f+g}(h)} \le \sqrt{I_f(h)} + \sqrt{I_g(h)}$ *if $h \ge 0$ a.e.*

PROOF. For $f, g, h \in L^\infty$ and $t > 0$, let

$$I_{f,g}^{(t)}(h) = \mathcal{E}^{(t)}(fh, g) + \mathcal{E}^{(t)}(gh, f) - \mathcal{E}^{(t)}(fg, h),$$

where $\mathcal{E}^{(t)}(f,g) = t^{-1}(f - T_t f, g)_{L^2}$. When $f = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$, $g = \sum_{i=1}^n \beta_i \mathbb{1}_{A_i}$ and $h = \sum_{i=1}^n \gamma_i \mathbb{1}_{A_i}$ with $\alpha_i, \beta_i, \gamma_i \in \mathbb{R}$ and disjoint sets $A_i \in \mathcal{B}$ with $\bigcup_{i=1}^n A_i = \Omega$, we have, as in [7], Proposition 1.2.3.3,

$$I_{f,g}^{(t)}(h) = \frac{1}{t} \sum_{i,j=1}^{n} (\alpha_i - \alpha_j)(\beta_i - \beta_j)\gamma_i P_t(A_i, A_j). \tag{2.3}$$

From this expression and the limiting argument, we have the claims (i)–(iv) with I being replaced by $I^{(t)}$. Since $I_{f,g}(h) = \lim_{t\downarrow 0} I_{f,g}^{(t)}(h)$ when $f, g, h \in \mathbb{D}_b$, we reach the conclusion. $\square$

By this lemma, in order to check whether a function f belongs to $\mathbb{D}_0$, it is enough to consider only nonnegative functions h in (1.1). Another important consequence of Lemma 2.3 is that $I_f(h) \le 2C\mathcal{E}(f,f)$ if $h \le C$ a.e. Also, $\langle f, g\rangle_h = I_{f,g}(h)$ is a pre-inner product when $h \ge 0$ a.e. Moreover, we have, when $|h| \le C$ a.e.,

$$|I_{f,g}(h)| \le 2C\sqrt{\mathcal{E}(f,f)}\sqrt{\mathcal{E}(g,g)}.$$

In particular, we can extend $I(\cdot,\cdot;h)$ continuously on $\mathbb{D}\times\mathbb{D}$ and $\langle f,\cdot\rangle_h$ is a bounded linear functional in $\mathbb{D}$. Hence, we have the following result.

LEMMA 2.4. *Let $\{f_n\}$ be a sequence of functions in $\mathbb{D}$. Then*:

(i) *If $f_n \rightharpoonup f$ weakly in $\mathbb{D}$, then*

$$\liminf_{n\uparrow\infty} I_{f_n}(h) \ge I_f(h)$$

for any nonnegative $h \in \mathbb{D}_b$.

(ii) *If $f_n \to f$ strongly in $\mathbb{D}$, then*

$$\lim_{n\uparrow\infty} I_{f_n}(h) = I_f(h)$$

for any $h \in \mathbb{D}_b$.

LEMMA 2.5. *Suppose that $f(t,x) = f_t(x)$ is a bounded jointly measurable function for $(t,x) \in (0,T]\times\Omega$. Also suppose that $f_t \in \mathbb{D}$ for each $t \in (0,T]$ and that*

$$\int_0^T \mathcal{E}(f_t, f_t)\,dt < \infty.$$

If we denote

$$\bar{f}_T = \frac{1}{T}\int_0^T f_t\,dt,$$

then $\bar{f}_T \in \mathbb{D}$ *and the following is true for any nonnegative* $h \in \mathbb{D}_b$:

$$I_{\bar{f}_T}(h) \le \frac{1}{T}\int_0^T I_{f_t}(h)\,dt. \tag{2.4}$$

PROOF. The hypotheses imply that, by Theorem 3.6.20 in [15], $\bar{f}_T$ is in $\mathbb{D}$, the domain of the closed linear operator $\sqrt{-\mathcal{L}}$. Now let g and h be in $\mathbb{D}_b$ with $h \ge 0$ a.e. Then

$$0 \le I_{f_t - g}(h) = I_{f_t}(h) - 2I_{f_t,g}(h) + I_g(h).$$

We can integrate over $(0, T]$ and, by the same theorem in [15], obtain

$$2I_{\bar{f}_T,g}(h) - I_g(h) \le \frac{1}{T}\int_0^T I_{f_t}(h)\,dt.$$

Letting $g = \bar{f}_T$, we obtain (2.4). $\square$

Next, we write I in a convenient form by using Theorem 1.5.2.1 from [7]. The way this representation characterizes locality is crucial for the proof of Theorem 1.1 to be presented below.

LEMMA 2.6. *Let* f, g, h *be functions in* $\mathbb{D}$. *There exists a* (*signed*) *Radon measure* σ *on* $\mathbb{R}^3$ *such that, for all* C^1*-functions with compact support* F *and* G *on* $\mathbb{R}$ *and* H *on* $\mathbb{R}^3$,

$$I_{F(f),G(g)}\big(H(f,g,h)\big) = \int_{\mathbb{R}^3} F'(x)G'(y)H(x,y,z)\,d\sigma(x,y,z).$$

PROOF. Theorem 1.5.2.1 in [7] provides us with a family of signed Radon measures $\{\sigma_{i,j}\}_{i,j=1}^3$ on $\mathbb{R}^3$ such that $\sigma_{i,j} = \sigma_{j,i}$ for all i, j, $\sum_{i,j=1}^3 \rho_i\rho_j\sigma_{i,j}$ is positive for all $(\rho_1, \rho_2, \rho_3) \in \mathbb{R}^3$, and, for any C^1-functions F_1, F_2 on $\mathbb{R}^3$ with compact supports,

$$\mathcal{E}\big(F_1(f,g,h), F_2(f,g,h)\big) = \sum_{i,j=1}^3 \int_{\mathbb{R}^3} \frac{\partial F_1}{\partial x_i}\frac{\partial F_2}{\partial x_j}\,d\sigma_{i,j}.$$

Then a simple calculation shows that

$$I_{F(f),G(g)}\big(H(f,g,h)\big) = 2\int_{\mathbb{R}^3} F'(x)G'(y)H(x,y,z)\,d\sigma_{1,2}(x,y,z). \qquad \square$$

Hereafter, f, g and h will be assumed to be in $\mathbb{D}_b$. From Lemma 2.6, we obtain that, for C^2-functions F and G on $\mathbb{R}$,

$$I_{F(f),G(g)}(h) = I_{f,g}\big(F'(f)G'(g)h\big), \tag{2.5}$$

by modifying F, G and identity functions appropriately to have compact supports. The identities $2\mathcal{E}(h, F(f)) = I_{h,F(f)}(1) = I_{f,h}(F'(f))$ and $I_f(F''(f)h) = I_{F'(f),f}(h)$, and another simple computation also give us the fundamental identity

$$\mathcal{E}\big(h, F(f)\big) = \mathcal{E}\big(F'(f)h, f\big) - \tfrac{1}{2}I_f\big(F''(f)h\big), \qquad F \in C^3(\mathbb{R}). \tag{2.6}$$

From (2.5) and (2.6), we obtain that, for $f \in \mathbb{D}_b$ uniformly bounded away from zero,

$$\mathcal{E}(h, \log f) = \mathcal{E}\left(\frac{h}{f}, f\right) + \frac{1}{2}I_f\left(\frac{h}{f^2}\right) = \mathcal{E}\left(\frac{h}{f}, f\right) + \frac{1}{2}I_{\log f}(h), \tag{2.7}$$

since log can be thought of as a smooth function when evaluated in arguments that are bounded away from zero. We also have, for the cutoffs of the last section,

$$\begin{aligned}\mathcal{E}\big(h, \Phi(f)\big) &= \mathcal{E}\big(h\phi'(f)^2, f\big) - I_f\big(h\phi'(f)\phi''(f)\big)\\ &= \mathcal{E}\big(h\phi'(f)^2, f\big) - I_{\phi(f),\phi'(f)}(h).\end{aligned} \tag{2.8}$$

The following lemma is crucial for the next section.

LEMMA 2.7. *The set* $\mathbb{D}_0$ *is convex. If a sequence* $\{f_n\}$ *in* $\mathbb{D}_0$ *converges weakly to* $f \in \mathbb{D}_b$ *in* $\mathbb{D}$, *then* $f \in \mathbb{D}_0$. *Moreover*, $\mathbb{D}_0$, *defined in* (1.1), *is closed under the operations* $\wedge$ *and* $\vee$.

PROOF. The first two assertions follow from Lemmas 2.5 and 2.4. For the proof of the last one, let

$$\mathbb{D}_0^{\#} = \big\{f \in \mathbb{D}_b \mid \text{there exists some } \alpha \geq 0 \text{ such that } I_f(h) \leq \alpha\|h\|_{L^1} \text{ for all } h \in \mathbb{D}_b\big\}.$$

Since $\mathbb{D}_b$ is dense in L^1, each $f \in \mathbb{D}_0^{\#}$ associates a unique element $\Gamma(f)$ in L^∞ such that

$$I_f(h) = \int_\Omega h\Gamma(f)\,d\mu, \qquad h \in \mathbb{D}_b.$$

Note that $f \in \mathbb{D}_0$ if and only if $\|\Gamma(f)\|_{L^\infty} \leq 1$. First, we will prove that $\mathbb{D}_0^{\#}$ is a vector lattice. Since $\mathbb{D}_0^{\#}$ is a vector space by Lemma 2.3(iv), we need only to prove that $f \in \mathbb{D}_0^{\#}$ implies $|f| \in \mathbb{D}_0^{\#}$. Take a sequence of C^1-functions $\{F_n\}$ on $\mathbb{R}$ such that $|F_n'| \leq 1$ everywhere and $F_n(f)$ converges to $|f|$ weakly in $\mathbb{D}$. The lemmas above imply that each $F_n(f)$, hence $|f|$, belongs to $\mathbb{D}_0^{\#}$.

The map $\Gamma : \mathbb{D}_0^{\#} \times \mathbb{D}_0^{\#} \to L^1$ can be defined by polarization; namely, $\Gamma(f, g) = (\Gamma(f+g) - \Gamma(f-g))/4$. This is symmetric and bilinear and, from Lemma 2.3(i), satisfies the positivity: $\Gamma(f, f) = \Gamma(f) \geq 0$ a.e. if $f \geq 0$ a.e. Then, by a standard argument, the Schwarz inequality

$$\Gamma(f, g) \leq \Gamma(f)^{1/2}\Gamma(g)^{1/2} \qquad \text{a.e., } f, g \in \mathbb{D}_0^{\#},$$

holds, and therefore, when $f, g \in \mathbb{D}_0^{\#}$,

$$|\Gamma(f) - \Gamma(g)| \le \Gamma(f-g)^{1/2}\Gamma(f+g)^{1/2} \qquad \text{a.e.} \tag{2.9}$$

As is proven in the same way in [7], Theorem 1.7.1.1, if $f \in \mathbb{D}_0^{\#}$ satisfies $f = 0$ on Ω_0 a.e. for some measurable set Ω_0 of Ω, then $\Gamma(f) = 0$ on Ω_0 a.e. In particular, by combining (2.9), we obtain that $\Gamma(f_1) = \Gamma(f_2)$ on Ω_0 a.e. when $f_1 = f_2$ on Ω_0 a.e. for $f_1, f_2 \in \mathbb{D}_0^{\#}$.

Now let f and g belong to $\mathbb{D}_0$. Set $\Omega_0 = \{f \le g\}$. Then $f \wedge g \in \mathbb{D}_0^{\#}$ and $\Gamma(f \wedge g) = \Gamma(f) \cdot \mathbb{1}_{\Omega_0} + \Gamma(g) \cdot \mathbb{1}_{\Omega \setminus \Omega_0}$ a.e., since $f \wedge g = f$ on Ω_0 and $f \wedge g = g$ on $\Omega \setminus \Omega_0$. Therefore, $\|\Gamma(f \wedge g)\|_{L^\infty} \le \|\Gamma(f)\|_{L^\infty} \vee \|\Gamma(g)\|_{L^\infty} \le 1$, which means that $f \wedge g \in \mathbb{D}_0$. The same argument applies to $f \vee g$. $\square$

2.3. *On the intrinsic distance.* This section contains the proof of Theorem 1.2. We define a distance function to a measurable set in the same spirit as was done in [25] and [36] (a separability condition was required in [36], which is not needed).

Given any measurable set A and positive number $N > 0$, consider the set

$$V_A^N = \{f \in \mathbb{D}_0 \mid f = 0 \text{ on } A \text{ and } 0 \le f \le N \text{ a.e.}\}.$$

The quantity

$$M = \sup\{\|f\|_{L^1} \mid f \in V_A^N\}$$

is clearly bounded by N. Take a sequence of functions $f_n \in V_A^N$ such that $\|f_n\|_{L^1} \to M$. Since $f_n \in \mathbb{D}_0$ and f_n is bounded, this sequence is bounded in $\mathbb{D}$. Therefore, it has a weakly convergent subsequence. Call d_A^N its limit and notice that $\mathsf{d}_A^N \in \mathbb{D}_0$ by Lemma 2.7 and $\|\mathsf{d}_A^N\|_{L^1} = (\mathsf{d}_A^N, 1)_{L^2} = \lim_{n\to\infty}(f_n, 1)_{L^2} = M$.

We now check that d_A^N is the (a.e.) largest element of V_A^N. Suppose there was a $g \in V_A^N$ with $\{g > d_A^N\}$ having positive measure. Then $g \vee \mathsf{d}_A^N \in V_A^N$ by Lemma 2.7 and

$$\|g \vee \mathsf{d}_A^N\|_{L^1} > \|\mathsf{d}_A^N\|_{L^1} = M,$$

which creates a contradiction.

By noting that $\mathsf{d}_A^N = \mathsf{d}_A^{N'} \wedge N$ for $N < N'$, $\mathsf{d}_A = \lim_{N\uparrow\infty} \mathsf{d}_A^N$ is well defined as a measurable function. The first three assertions of Theorem 1.2 are easy consequences of all this. For the last assertion, just notice that it is enough to consider in (1.2) only functions f in $\mathbb{D}_0$ such that $f = 0$ on A and $f \ge 0$ a.e.

2.4. *Proof of the upper bound.* The upper bound is proven in the standard way, namely, Gaffney's method [21]. We just have to hide the fact that we no longer have a "gradient." The actual statement is stronger because it gives a bound on $P_t(A, B)$ for each $t > 0$.

THEOREM 2.8. *With the definitions given in the Introduction, the following bound holds*:

$$P_t(A,B) \le \sqrt{\mu(A)\mu(B)}\, e^{-\mathsf{d}(A,B)^2/2t}, \qquad t>0.$$

PROOF. Let $v_t = T_t\mathbb{1}_A$ and $w \in \mathbb{D}_0$. Fix $\alpha \in \mathbb{R}$ and consider

$$f(t) = \int_\Omega (e^{\alpha w} v_t)^2\, d\mu.$$

Differentiating and using the relation (2.5) and Lemma 2.3(iii), we get

$$\begin{aligned} f'(t) &= -2\mathcal{E}(e^{2\alpha w} v_t, v_t) \\ &= -I_{v_t}(e^{2\alpha w}) - 2\alpha I_{v_t,w}(v_t e^{2\alpha w}) \\ &\le \alpha^2 I_w(v_t^2 e^{2\alpha w}) \le \alpha^2 f(t). \end{aligned}$$

Solving this differential inequality, we have

$$f(t) \le f(0) e^{\alpha^2 t}. \tag{2.10}$$

Suppose that $\mathsf{d}(A,B) < \infty$. By setting $w = \mathsf{d}_A \wedge \mathsf{d}(A,B)$, (2.10) implies that

$$\|e^{\alpha w} T_t \mathbb{1}_A\|_{L^2} \le \sqrt{\mu(A)}\, e^{\alpha^2 t/2}.$$

A similar calculation by letting $v_t = T_t\mathbb{1}_B$ gives that

$$\|e^{-\alpha w} T_t \mathbb{1}_B\|_{L^2} \le \sqrt{\mu(B)} e^{-\alpha \mathsf{d}(A,B) + \alpha^2 t/2}.$$

Finally, use the Schwarz inequality to show that

$$P_t(A,B) \le \|e^{\alpha w} T_{t/2}\mathbb{1}_A\|_{L^2} \|e^{-\alpha w} T_{t/2}\mathbb{1}_B\|_{L^2} \le \sqrt{\mu(A)\mu(B)}\, e^{-\alpha \mathsf{d}(A,B)+\alpha^2 t/2}$$

and optimize in α to obtain

$$P_t(A,B) \le \sqrt{\mu(A)\mu(B)}\, e^{-\mathsf{d}(A,B)^2/2t}.$$

This gives the result in the case of finite distance. If $\mathsf{d}(A,B) = \infty$, then the above procedure can be slightly modified to give $P_t(A,B) = 0$ for all $t \ge 0$. For this, use $w = \mathsf{d}_A \wedge M$ to obtain

$$P_t(A,B) \le \sqrt{\mu(A)\mu(B)}\, e^{-M^2/2t}.$$

Let $M \uparrow \infty$ to conclude. $\square$

Notice that, if we denote $u_t = -t \log T_t \mathbb{1}_A$, $\Phi_t = \Phi(u_t)$ and $\Xi_t = \Xi(u_t)$, then the families $\{\Phi_t\}_{t>0}$ and $\{\Xi_t\}_{t>0}$ are bounded and therefore weakly relatively compact in L^2. We will need the following lemma when proving the lower bound estimate in Theorem 1.2 and the strong convergence stated in Theorem 1.3.

LEMMA 2.9. *It holds that*

$$\Phi_0(x) \geq \Phi\big(\mathsf{d}_A(x)^2/2\big) \qquad a.e.$$

for any L^2 weak limit Φ_0 of $\{\Phi_t\}_{t>0}$. Moreover,

$$\Xi_0(x) \geq \Xi\big(\mathsf{d}_A(x)^2/2\big) \qquad a.e.$$

is true for any weak limit point Ξ_0 of $\{\Xi_t\}_{t>0}$.

PROOF. Let C be a measurable set with $\mu(C) > 0$ Then, when limits are taken along the appropriate subsequence $\{t_k\} \downarrow 0$,

$$\begin{aligned}
\Phi\left(\frac{\mathsf{d}(A,C)^2}{2}\right) &\leq \Phi\left(\liminf_{k\uparrow\infty} -t_k \log P_{t_k}(A,C)\right) \\
&= \Phi\left(\liminf_{k\uparrow 0} -t_k \log\left(\frac{1}{\mu(C)}\int_C T_{t_k}\mathbb{1}_A\,d\mu\right)\right) \\
&= \liminf_{k\uparrow\infty} \Phi\left(-t_k \log\left(\frac{1}{\mu(C)}\int_C T_{t_k}\mathbb{1}_A\,d\mu\right)\right) \\
&\leq \liminf_{k\uparrow\infty} \frac{1}{\mu(C)}\int_C \Phi(-t_k \log T_{t_k}\mathbb{1}_A)\,d\mu \\
&= \liminf_{k\uparrow\infty} \frac{1}{\mu(C)}\int_C \Phi_{t_k}\,d\mu = \frac{1}{\mu(C)}\int_C \Phi_0\,d\mu.
\end{aligned}$$

Here Lemma 2.1 was used in the fourth line. Given $\varepsilon > 0$, suppose that $D_\varepsilon = \{\Phi_0 \leq \Phi(\mathsf{d}_A^2/2) - \varepsilon\}$ has positive measure. Then $C_\varepsilon = \{x \in D_\varepsilon \mid \Phi(\mathsf{d}_A^2(x)/2) \leq \Phi(\mathsf{d}(A,D_\varepsilon)^2/2) + \varepsilon/2\}$ also has positive measure. Hence, we can write

$$\begin{aligned}
\frac{1}{\mu(C_\varepsilon)}\int_{C_\varepsilon} \Phi\left(\frac{\mathsf{d}_A^2}{2}\right)d\mu &\leq \Phi\left(\frac{\mathsf{d}(A,C_\varepsilon)^2}{2}\right) + \frac{\varepsilon}{2} \\
&\leq \frac{1}{\mu(C_\varepsilon)}\int_{C_\varepsilon} \Phi_0\,d\mu + \frac{\varepsilon}{2} \\
&\leq \frac{1}{\mu(C_\varepsilon)}\int_{C_\varepsilon} \Phi\left(\frac{\mathsf{d}_A^2}{2}\right)d\mu - \frac{\varepsilon}{2},
\end{aligned}$$

which is a contradiction. Since $\varepsilon > 0$ is arbitrary, we obtain the first conclusion of the lemma.

The second statement is proven by the same procedure. $\square$

2.5. *First step in the proof of the lower bound.* We start by proving Lemma 2.10. It gives us an equation satisfied by $u_t = -t \log T_t \mathbb{1}_A$ (or functions of it) that is our main workhorse for the rest of the section. For a given Borel-measurable function ψ on $[0,\infty]$, we write ψ_t instead of $\psi(u_t)$.

LEMMA 2.10. *Suppose $f \in L^2$ with values in $[\varepsilon, 1]$ a.e. for some $\varepsilon > 0$. Let $u_t(x) = -t \log T_t f(x)$ and ϕ, Φ and Ψ as in Section 2.1. Let $\rho_t(x) \in H^1([t_0, t_1]; L^2) \cap L^2([t_0, t_1]; \mathbb{D}_b)$ with $0 < t_0 < t_1$. Then the function $t \mapsto (\Phi_t, \rho_t)_{L^2}$ is absolutely continuous and satisfies*

$$\begin{aligned}\partial_t(\Phi_t, \rho_t)_{L^2} &= (\Phi_t, \partial_t \rho_t)_{L^2} - \mathcal{E}(\Phi_t, \rho_t) - I(\phi_t, \phi_t'; \rho_t) \\ &\quad + \frac{1}{t}\left[(\Psi_t, \rho_t)_{L^2} - \frac{1}{2} I(\phi_t; \rho_t)\right]\end{aligned} \tag{2.11}$$

for almost every $t \in [t_0, t_1]$.

PROOF. Since log can be modified on $[0, \varepsilon)$ and $(1, \infty)$ to make it a smooth function of compact support, we can apply the formulas from the previous section. By (2.7), we have

$$\begin{aligned}(\rho_t, \partial_t u_t)_{L^2} &= -t\left(\frac{\rho_t}{T_t f}, \partial_t T_t f\right)_{L^2} + \frac{1}{t}(\rho_t, u_t)_{L^2} \\ &= t\mathcal{E}\left(\frac{\rho_t}{T_t f}, T_t f\right) + \frac{1}{t}(\rho_t, u_t)_{L^2} \\ &= -\mathcal{E}(\rho_t, u_t) - \frac{1}{2t} I_{u_t}(\rho_t) + \frac{1}{t}(\rho_t, u_t)_{L^2}.\end{aligned}$$

Using this and the definitions for ϕ, Φ and Ψ, we can compute

$$\begin{aligned}\big(\rho_t, \partial_t \Phi(u_t)\big)_{L^2} &= \big(\rho_t (\phi_t')^2, \partial_t u_t\big)_{L^2} \\ &= -\mathcal{E}\big(\rho_t(\phi_t')^2, u_t\big) + \frac{1}{t}\left[-\frac{1}{2} I_{u_t}\big(\rho_t(\phi_t')^2\big) + (\rho_t, \Psi_t)_{L^2}\right].\end{aligned}$$

By recalling (2.8), this is equal to

$$-\mathcal{E}(\rho_t, \Phi_t) - I_{\phi_t, \phi_t'}(\rho_t) - \frac{1}{2t} I_{\phi_t}(\rho_t) + \frac{1}{t}(\rho_t, \Psi_t)_{L^2}.$$

Therefore, (2.11) follows. $\square$

For a function $f : [0, \infty) \times \Omega \to \mathbb{R}$, we define its time average $\bar{f}$ by

$$\bar{f}_t(x) = \frac{1}{t} \int_0^t f_s(x)\, ds.$$

THEOREM 2.11. *The families $\{\bar{\phi}_t\}_{0<t<\tau}$ and $\{\bar{\Phi}_t\}_{0<t<\tau}$ are uniformly bounded in $\mathbb{D}$ for τ small enough.*

PROOF. Call

$$u_t^\delta = -t \log((1 - \delta) T_t \mathbb{1}_A + \delta),$$

where $\delta \in (0,1)$. We will use the notation ψ_t^δ for $\psi(u_t^\delta)$, where ψ is any Borel-measurable function on $[0,\infty)$. Taking $\rho_t \equiv 1$ in Lemma 2.10, we have, for $0 < \varepsilon < T$,

$$\begin{aligned}\int_\varepsilon^T \mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt \\ = -\int_\varepsilon^T \int_\Omega t\,\partial_t \Phi_t^\delta\,d\mu\,dt - 2\int_\varepsilon^T t\mathcal{E}\big((\phi')_t^\delta, \phi_t^\delta\big)\,dt + \int_\varepsilon^T \int_\Omega \Psi_t^\delta\,d\mu\,dt.\end{aligned} \tag{2.12}$$

The last term on the right-hand side of (2.12) is bounded by LKT. On the first term, we integrate by parts in the t variable. This gives

$$\begin{aligned}\int_\varepsilon^T \int_\Omega t\,\partial_t \Phi(u_t^\delta)\,d\mu\,dt &= t\|\Phi_t^\delta\|_{L^1}\big|_\varepsilon^T - \int_\varepsilon^T \|\Phi_t^\delta\|_{L^1}\,dt \\ &= T\|\Phi_T^\delta\|_{L^1} - \varepsilon\|\Phi_\varepsilon^\delta\|_{L^1} - \int_\varepsilon^T \|\Phi_t^\delta\|_{L^1}\,dt\end{aligned}$$

and a bound of LKT on this term.

Finally, for the second term in (2.12), we first notice that

$$\big|\mathcal{E}\big((\phi')_t^\delta, \phi_t^\delta\big)\big| = \left|\frac{1}{2} I_{u_t^\delta}\big((\phi'')_t^\delta (\phi')_t^\delta\big)\right| \le \frac{C}{2K} I_{u_t^\delta}\big(((\phi')_t^\delta)^2\big) = \frac{C}{K}\mathcal{E}(\phi_t^\delta, \phi_t^\delta),$$

because of (2.1) and Lemma 2.3. Now we can write

$$\begin{aligned}\left|\int_\varepsilon^T t\mathcal{E}\big((\phi')_t^\delta, \phi_t^\delta\big)\,dt\right| &\le \frac{C}{K}\int_\varepsilon^T t\mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt \\ &\le \frac{CT}{K}\int_\varepsilon^T \mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt.\end{aligned} \tag{2.13}$$

Putting things together, we obtain

$$\int_\varepsilon^T \mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt \le 2LKT + \frac{2CT}{K}\int_\varepsilon^T \mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt$$

or, rearranging terms and letting $\varepsilon \downarrow 0$,

$$\frac{1}{T}\int_0^T \mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt \le \frac{2LK}{1 - 2CT/K} \tag{2.14}$$

if $T < K/(2C)$. By Lemma 2.5,

$$\mathcal{E}(\bar\phi_T^\delta, \bar\phi_T^\delta) \le \frac{2LK}{1 - 2CT/K}. \tag{2.15}$$

Letting $\delta \downarrow 0$, we obtain that $\bar\phi_T \in \mathbb{D}$ and $\mathcal{E}(\bar\phi_T, \bar\phi_T)$ has the same bound. This finishes the proof for $\{\bar\phi_t\}_{0<t<\tau}$.

The case of $\{\bar\Phi_t\}_{0<t<\tau}$ is easily deduced from here. For this, notice that

$$\mathcal{E}(\Phi_t^\delta, \Phi_t^\delta) = \tfrac{1}{2} I_{u_t}\big(\phi'(u_t^\delta)^4\big) \le \tfrac{1}{2} I_{u_t}\big(\phi'(u_t^\delta)^2\big) = \mathcal{E}(\phi_t^\delta, \phi_t^\delta)$$

(use Lemma 2.3) or, after averaging in time and using convexity,

$$\mathcal{E}(\bar{\Phi}_t^\delta, \bar{\Phi}_t^\delta) \le \frac{1}{T}\int_0^T \mathcal{E}(\Phi_t^\delta, \Phi_t^\delta)\,dt \le \frac{1}{T}\int_0^T \mathcal{E}(\phi_t^\delta, \phi_t^\delta)\,dt \le \frac{2LK}{1-2CT/K}. \tag{2.16}$$

The last inequality above is just (2.14). Taking limits as $\delta \downarrow 0$, we obtain that $\mathcal{E}(\bar{\Phi}_t, \bar{\Phi}_t)$ is bounded in t. $\square$

At this point, we take a suitable sequence $\{t_k\}$ for which it holds that:

- $\phi_{t_k} \rightharpoonup \phi_0$ and $\Psi_{t_k} \rightharpoonup \Psi_0$ weakly in L^2,
- $\bar{\phi}_{t_k} \rightharpoonup \bar{\phi}_0$ and $\bar{\Phi}_{t_k} \rightharpoonup \bar{\Phi}_0$ weakly in $\mathbb{D}$,

for some functions $\phi_0, \Psi_0 \in L^2$ and $\bar{\phi}_0, \bar{\Phi}_0 \in \mathbb{D}_b$.

Let $\rho \in \mathbb{D}_b$ be a function that is nonnegative. Use (2.11) with this ρ independent of time and integration by parts to get

$$\begin{aligned}
&\tfrac{1}{2}\int_\varepsilon^T I_{\phi_t^\delta}(\rho)\,dt \\
&\quad = -t(\rho, \Phi_t^\delta)_{L^2}\Big|_\varepsilon^T + \int_\varepsilon^T (\rho, \Phi_t^\delta)_{L^2}\,dt \\
&\qquad - \int_\varepsilon^T t\mathcal{E}(\rho, \Phi_t^\delta)\,dt - \int_\varepsilon^T tI_{\phi_t^\delta,(\phi')_t^\delta}(\rho)\,dt + \int_\varepsilon^T (\rho, \Psi_t^\delta)_{L^2}\,dt.
\end{aligned} \tag{2.17}$$

We want to make estimates on the "gradient squared," that is, on $I(\phi_t; \rho)$. We start with the fourth term on the right-hand side. As in (2.13), we have

$$\left|\int_\varepsilon^T tI\big(\phi_t^\delta, (\phi')_t^\delta; \rho\big)\,dt\right| \le \frac{CT}{K}\int_\varepsilon^T I_{\phi_t^\delta}(\rho)\,dt. \tag{2.18}$$

We rearrange the terms in (2.17), let $\varepsilon \downarrow 0$, divide by T and use Lemma 2.5 to obtain

$$\begin{aligned}
&\left(1 - \frac{2CT}{K}\right)\frac{1}{2}I(\bar{\phi}_T^\delta; \rho) \\
&\qquad \le -(\rho, \Phi_T^\delta)_{L^2} + (\rho, \bar{\Phi}_T^\delta)_{L^2} + (\rho, \bar{\Psi}_T^\delta)_{L^2} - \frac{1}{T}\int_0^T t\mathcal{E}(\rho, \Phi_t^\delta)\,dt
\end{aligned}$$

if $T < K/(2C)$. Here we used (2.16) to assure the convergence in the last term on the right-hand side.

At this point, we take limits as $\delta \downarrow 0$. Each of the functions $\bar{\phi}^\delta$, Φ_t^δ, Ψ_t^δ and Ψ_t^δ is uniformly bounded and converges pointwise. On the other hand, $\Phi_t^\delta \rightharpoonup \Phi_t$ weakly in $L^2((0,T]; \mathbb{D})$ and $\bar{\phi}_t^\delta \rightharpoonup \bar{\phi}_t$ weakly in $\mathbb{D}$ by (2.15) and (2.16). Lower semicontinuity of the energy term on the left-hand side is enough to get

$$\begin{aligned}
&\left(1 - \frac{2CT}{K}\right)\frac{1}{2}I_{\bar{\phi}_T}(\rho) \\
&\qquad \le -(\rho, \Phi_T)_{L^2} + (\rho, \bar{\Phi}_T)_{L^2} + (\rho, \bar{\Psi}_T)_{L^2} - \frac{1}{T}\int_0^T t\mathcal{E}(\rho, \Phi_t)\,dt.
\end{aligned} \tag{2.19}$$

The first three terms on the right-hand side are fine for our purposes. In the last term, we use integration by parts to obtain

$$\begin{aligned}\frac{1}{T}\int_0^T t\mathcal{E}(\rho,\Phi_t)\,dt &= \frac{1}{T}t\int_0^t \mathcal{E}(\rho,\Phi_s)\,ds\Big|_0^T - \frac{1}{T}\int_0^T\int_0^t \mathcal{E}(\rho,\Phi_s)\,ds\,dt \\ &= T\mathcal{E}(\rho,\bar{\Phi}_T) - \frac{1}{T}\int_0^T t\mathcal{E}(\rho,\bar{\Phi}_t)\,dt \to 0\end{aligned}$$

as $T\downarrow 0$. Hence, we have an inequality of the form, by taking limits along $\{t_k\}$ in (2.19) and using Lemma 2.4,

$$\tfrac{1}{2}I_{\bar{\phi}_0}(\rho) \le -(\rho,\Phi_0)_{L^2} + (\rho,\bar{\Phi}_0)_{L^2} + (\rho,\bar{\Psi}_0)_{L^2}, \tag{2.20}$$

which is going to be our main tool in the following argument. In our first approximation, we drop the first term to obtain

$$\tfrac{1}{2}I_{\bar{\phi}_0}(\rho) \le (\rho,\bar{\Phi}_0)_{L^2} + (\rho,\bar{\Psi}_0)_{L^2} \le 2(\rho,\bar{\phi}_0)_{L^2}.$$

Then, for each $\varepsilon>0$,

$$I_{\sqrt{\bar{\phi}_0+\varepsilon}}(h) = \frac{1}{4}I_{\bar{\phi}_0}\left(\frac{h}{\bar{\phi}_0+\varepsilon}\right) \le \left(\bar{\phi}_0, \frac{h}{\bar{\phi}_0+\varepsilon}\right)_{L^2} \le \|h\|_{L^1}$$

for any nonnegative $h\in\mathbb{D}_b$. This means $\sqrt{\bar{\phi}_0+\varepsilon}\in\mathbb{D}_0$. Letting $\varepsilon\downarrow 0$, we obtain $\sqrt{\bar{\phi}_0}\in\mathbb{D}_0$. Therefore, we will know that $\bar{\phi}_0\le \mathsf{d}_A^2$ a.e. as long as we have that $\bar{\phi}_0=0$ on A. But this is easy (see, e.g., Lemma 3.7 in [36]).

Notice that the estimate is not sharp; we want $\bar{\phi}_0\le \mathsf{d}_A^2/2$. In order to improve it, we iterate the inequality that we have in order to obtain a sharp estimate.

LEMMA 2.12. *If the inequality*

$$\bar{\phi}_0^K(x) \le c\,\frac{\mathsf{d}_A(x)^2}{2}$$

holds true a.e. for some $c>1$ for every K and every weak limit in L^2, then the inequality

$$\bar{\phi}_0^K(x) \le \left(2-\frac{1}{c}\right)\frac{\mathsf{d}_A(x)^2}{2}$$

also holds true almost everywhere.

PROOF. Let us make explicit the dependence of Φ on the cutoff K for the following argument. Given K, we can choose $M<\infty$ such that $\Phi^K(M)\ge$

$\sup_x \Psi^K(x)$. Then $\Phi^K(\phi_t^M) \geq \Psi_t^K$ holds a.e. We use the convexity of $\Phi(-t\log(\cdot))$ (Lemma 2.1) to see that, for any nonnegative $\rho \in \mathbb{D}_b$ that is not identically 0,

$$\begin{aligned}(\Phi_t^K, \rho)_{L^2} &= \int_\Omega \Phi^K(-t\log(T_t \mathbb{1}_A))\rho\, d\mu \\ &\geq \|\rho\|_{L^1} \Phi^K\biggl(-t\log\biggl(\frac{1}{\|\rho\|_{L^1}}(T_t\mathbb{1}_A, \rho)_{L^2}\biggr)\biggr).\end{aligned}$$

Let $S_\rho = \{\rho > 0\}$. We can estimate the limit of the expression on the right-hand side because we have the upper bound available. We do it as follows:

$$\liminf_{t\downarrow 0} -t\log(T_t\mathbb{1}_A, \rho)_{L^2} \geq \liminf_{t\downarrow 0} -t\log P_t(A, S_\rho) \geq \frac{\mathsf{d}(A, S_\rho)^2}{2}.$$

We conclude that, in the limit,

$$(\Phi_0^K, \rho)_{L^2} \geq \|\rho\|_{L^1} \Phi^K\biggl(\frac{\mathsf{d}(A, S_\rho)^2}{2}\biggr) \geq \frac{\|\rho\|_{L^1}}{c} \operatorname*{essinf}_{x\in S_\rho} \Phi^K(\bar{\phi}_0^M).$$

Since Φ^K is concave,

$$\Phi^K(\bar{\phi}_t^M) = \Phi^K\biggl(\frac{1}{t}\int_0^t \phi_s^M\, ds\biggr) \geq \frac{1}{t}\int_0^t \Psi_s^K\, ds = \bar{\Psi}_t^K \qquad \text{a.e.}$$

Therefore, we get that $(\Phi_0^K, \rho)_{L^2} \geq c^{-1}\|\rho\|_{L^1} \operatorname{essinf}_{x\in S_\rho} \bar{\Psi}_0^K$. Combining this relation with (2.20), we obtain

$$\frac{1}{2} I_{\bar{\phi}_0^K}(\rho) \leq -\frac{\|\rho\|_{L^1}}{c} \operatorname*{essinf}_{x\in S_\rho} \bar{\Psi}_0^K(x) + (\rho, \bar{\Phi}_0^K)_{L^2} + (\rho, \bar{\Psi}_0^K)_{L^2}.$$

A minimal adaptation of Lemma 3.9 from [36] now implies that

$$I_{\bar{\phi}_0^K}(\rho) \leq \biggl(1 - \frac{1}{c}\biggr)(\rho, \bar{\Psi}_0^K)_{L^2} + (\rho, \bar{\Phi}_0^K)_{L^2} \leq \biggl(2 - \frac{1}{c}\biggr)(\rho, \bar{\phi}_0^K)_{L^2}.$$

The result follows from here. $\square$

After iterating this procedure, we conclude that $\bar{\phi}_0 \leq \mathsf{d}_A^2/2$ a.e. and, therefore, that $\bar{\Phi}_0 \leq \mathsf{d}_A^2/2$ a.e.

On the other hand, we have a partial converse.

LEMMA 2.13. *It holds that $\bar{\Phi}_0 \geq \Phi(\mathsf{d}_A^2/2)$ a.e. for every weak limit $\bar{\Phi}_0$.*

PROOF. By Lemma 2.9, for every nonnegative $\rho \in L^2$,

$$\biggl(\Phi\biggl(\frac{\mathsf{d}_A^2}{2}\biggr), \rho\biggr)_{L^2} \leq \liminf_{t\downarrow 0} (\Phi_t, \rho)_{L^2}.$$

Then

$$\left(\Phi\left(\frac{\mathsf{d}_A^2}{2}\right), \rho\right)_{L^2} \le \liminf_{t\downarrow 0} \frac{1}{t}\int_0^t (\Phi_s, \rho)_{L^2}\, ds \le \liminf_{k\to\infty} \big(\bar{\Phi}_{t_k}, \rho\big)_{L^2} = (\bar{\Phi}_0, \rho)_{L^2}.$$

This means that $\Phi(\mathsf{d}_A^2/2) \le \bar{\Phi}_0$ a.e. $\square$

Hence, $\bar{\Phi}_0 = \mathsf{d}_A^2/2$ on $\{\mathsf{d}_A^2/2 \le K\}$ a.e. independent of the choice of subsequences. Namely, $\bar{\Phi}_t \cdot \mathbb{1}_{D_K}$ converges weakly to $\mathsf{d}_A^2/2 \cdot \mathbb{1}_{D_K}$ in L^2 as $t \downarrow 0$, where $D_K = \{\mathsf{d}_A^2/2 \le K\}$. Moreover, the following lemma tells us that this extends to the whole space.

LEMMA 2.14. *The function $\bar{\Phi}_t$ has a unique weak L^2-limit $\bar{\Phi}_0$ as $t \downarrow 0$ and $\bar{\Phi}_0 = \Phi(\mathsf{d}_A^2/2)$.*

PROOF. As in the proof of Lemma 2.12, we will use the notation Φ^K to make explicit the dependence of Φ. Initially, we only have that $\bar{\Phi}_t^M \cdot \mathbb{1}_{D_M}$ converges weakly, and its limit satisfies $\bar{\Phi}_0^M = \mathsf{d}_A^2/2$ on D_M for any $M > 0$. But notice that, by using this, inequality (2.2) and the concavity of Φ^K, it is true that

$$\begin{aligned}
\bar{\Phi}_T^K &= \frac{1}{T}\int_0^T \Phi^K(u_t)\, dt \\
&\le \frac{1}{T}\int_0^T \Phi^K\big(\Phi^M(u_t)\big)\, dt + \big(\Phi^K(\infty) - \Phi^K(M)\big) \\
&\le \Phi^K\left(\frac{1}{T}\int_0^T \Phi^M(u_t)\, dt\right) + \Delta_{K,M} \\
&= \Phi^K\big(\bar{\Phi}_T^M\big) + \Delta_{K,M},
\end{aligned} \tag{2.21}$$

where

$$\Delta_{K,M} = \Phi^K(\infty) - \Phi^K(M) \to 0 \qquad \text{as } M \uparrow \infty.$$

We can now take limits in (2.21) as $T \downarrow 0$ along any subsequence. Using Lemma 2.2 (since Φ^K is concave), we deduce that $\bar{\Phi}_0^K \le \Phi^K(\mathsf{d}_A^2/2) + \Delta_{K,M}$ on D_M for any limit point $\bar{\Phi}_0^K$. But now M can be taken arbitrarily large, so we get that $\bar{\Phi}_0^K \le \Phi^K(\mathsf{d}_A^2/2)$ a.e. for any limit function. This, together with Lemma 2.13, proves that $\Phi^K(\mathsf{d}_A^2/2)$ is the limit of any converging subsequence and hence also the limit of $\bar{\Phi}_t^K$ as $t \downarrow 0$. $\square$

2.6. *Final step in the proof of the lower bound.* At this point, we have proven Lemma 2.14 which has a similar statement to that of Theorem 1.3 but just for $F = \Phi$ and only after averaging in time. We turn to the same Tauberian theorem as in [36] (Lemma 3.11) in order to get rid of the averages. The main step is given by the following.

LEMMA 2.15. *Let $\tau > 0$. Then, for $t \le \tau$,*

$$\lim_{t\downarrow 0}(T_{\tau-t}\mathbb{1}_B, \Phi_t)_{L^2} = (T_\tau \mathbb{1}_B, \bar{\Phi}_0)_{L^2} = \big(T_\tau \mathbb{1}_B, \Phi(\mathsf{d}_A^2/2)\big)_{L^2}.$$

PROOF. If $f(t) = (T_{\tau-t}\mathbb{1}_B, \Phi_t)_{L^2}$, in order to apply Lemma 3.11 from [36], we have to check two conditions, namely:

(i) $\bar{f}(t) \to (T_\tau \mathbb{1}_B, \bar{\Phi}_0)_{L^2}$ as $t \downarrow 0$;
(ii) $f(t) - f(s) \le M(t-s)/s$ for some constant $M > 0$.

For (i), write

$$\begin{aligned}
&\left|\frac{1}{T}\int_0^T f(t)\,dt - (T_\tau \mathbb{1}_B, \bar{\Phi}_0)_{L^2}\right| \\
&\qquad \le \frac{LK}{T}\int_0^T \|T_{\tau-t}\mathbb{1}_B - T_\tau \mathbb{1}_B\|_{L^2}\,dt + (T_\tau \mathbb{1}_B, \bar{\Phi}_T - \bar{\Phi}_0)_{L^2}
\end{aligned}$$

and notice that the two terms on the right-hand side converge to 0.

Regarding (ii), we start by using Lemma 2.10 to obtain, for $\delta > 0$,

$$\begin{aligned}
(T_{\tau-r}\mathbb{1}_B, \Phi_r^\delta)_{L^2}\big|_s^t = \int_s^t \bigg\{&-I\big(\phi_r^\delta, (\phi')_r^\delta; T_{\tau-r}\mathbb{1}_B\big) \\
&+ \frac{1}{r}\bigg[(T_{\tau-r}\mathbb{1}_B, \Psi_r^\delta)_{L^2} - \frac{1}{2}I(\phi_r^\delta; T_{\tau-r}\mathbb{1}_B)\bigg]\bigg\}\,dr.
\end{aligned}$$

Estimating in a similar way as was done in the previous subsection, we get

$$\begin{aligned}
(T_{\tau-r}\mathbb{1}_B, \Phi_r^\delta)_{L^2}\big|_s^t &\le \int_s^t \left\{\frac{C}{K}I(\phi_r^\delta; T_{\tau-r}\mathbb{1}_B) + \frac{LK}{r} - \frac{1}{2r}I(\phi_r^\delta; T_{\tau-r}\mathbb{1}_B)\right\}dr \\
&\le \int_s^t \frac{LK}{r}\,dr \le \frac{LK}{s}(t-s)
\end{aligned}$$

for $t \le K/(2C)$. Since the bound is uniform in $\delta > 0$, we obtain

$$f(t) - f(s) = (T_{\tau-r}\mathbb{1}_B, \Phi_r)_{L^2}\big|_s^t \le \frac{LK}{s}(t-s)$$

after taking limits as $\delta \downarrow 0$. $\square$

The previous result is easily seen to imply that, for every measurable set B,

$$\lim_{t\downarrow 0}\int_B \Phi_t\,d\mu = \int_B \Phi\left(\frac{\mathsf{d}_A^2}{2}\right)d\mu. \tag{2.22}$$

We see this by first writing

$$\int_B \Phi_t\,d\mu = (\Phi_t, T_{\tau-t}\mathbb{1}_B)_{L^2} + (\Phi_t, \mathbb{1}_B - T_{\tau-t}\mathbb{1}_B)_{L^2}$$

and then noticing that the modulus of the second term on the right-hand side is dominated by $LK\|\mathbb{1}_B - T_{\tau-t}\mathbb{1}_B\|_{L^2}$, which tends to 0 when one takes limits as $t \downarrow 0$ first and then as $\tau \downarrow 0$. In particular,

$$\Phi_t \rightharpoonup \Phi(\mathsf{d}_A^2/2) \qquad \text{weakly in } L^2 \text{ as } t \downarrow 0, \tag{2.23}$$

since any weak limit of $\{\Phi_t\}$ should coincide with $\Phi(\mathsf{d}_A^2/2)$ by (2.22).

Finally, we can give a proof of formula (1.4). We may assume that $\mathsf{d}(A, B) < \infty$. Let $\varepsilon > 0$ and $C_\varepsilon = B \cap \{\mathsf{d}_A \le \mathsf{d}(A, B) + \varepsilon\}$. From Lemma 2.1, Jensen's inequality and (2.22),

$$\begin{aligned}
&\limsup_{t\downarrow 0} \Phi\bigl(-t \log P_t(A, B)\bigr) \\
&\quad\le \limsup_{t\downarrow 0} \Phi\bigl(-t \log P_t(A, C_\varepsilon)\bigr) \\
&\quad= \limsup_{t\downarrow 0} \Phi\biggl(-t \log \frac{1}{\mu(C_\varepsilon)} \int_{C_\varepsilon} T_t\mathbb{1}_A \, d\mu\biggr) \\
&\quad\le \limsup_{t\downarrow 0} \frac{1}{\mu(C_\varepsilon)} \int_{C_\varepsilon} \Phi(-t \log T_t\mathbb{1}_A) \, d\mu \\
&\quad= \limsup_{t\downarrow 0} \frac{1}{\mu(C_\varepsilon)} \int_{C_\varepsilon} \Phi_t \, d\mu \\
&\quad= \frac{1}{\mu(C_\varepsilon)} \int_{C_\varepsilon} \Phi\biggl(\frac{\mathsf{d}_A^2}{2}\biggr) d\mu \\
&\quad\le \frac{1}{\mu(C_\varepsilon)} \int_{C_\varepsilon} \frac{\mathsf{d}_A^2}{2} \, d\mu \le \frac{(\mathsf{d}(A, B) + \varepsilon)^2}{2}.
\end{aligned}$$

Since ε is arbitrary and $\Phi^K(x) \uparrow x$ as $K \to \infty$ for each $x \ge 0$, we have found that

$$\limsup_{t\downarrow 0} -t \log P_t(A, B) \le \frac{\mathsf{d}(A, B)^2}{2}.$$

2.7. *Proof of strong convergence.* First, we note the following fact.

LEMMA 2.16. *For every $t > 0$, $\{T_t\mathbb{1}_A = 0\} = \{\mathsf{d}_A = \infty\}$ a.e.*

PROOF. By Theorem 2.8, it holds that $\{T_t\mathbb{1}_A = 0\} \supset \{\mathsf{d}_A = \infty\}$ a.e. On the other hand, from the argument originally due to Simon [41], $\{T_t\mathbb{1}_A = 0\} \subset \{T_s\mathbb{1}_A = 0\}$ if $s < t$. Indeed, it holds for any measurable set B that

$$\begin{aligned}
P_t(A, B) &= (\mathbb{1}_A, T_{t-s}T_s\mathbb{1}_B)_{L^2} \ge \bigl(\mathbb{1}_A \cdot T_s\mathbb{1}_B, T_{t-s}(\mathbb{1}_A \cdot T_s\mathbb{1}_B)\bigr)_{L^2} \\
&= \|T_{(t-s)/2}(\mathbb{1}_A \cdot T_s\mathbb{1}_B)\|_{L^2}^2 \ge \|T_{(t-s)/2}(\mathbb{1}_A \cdot T_s\mathbb{1}_B)\|_{L^1}^2 \\
&= P_s(A, B)^2.
\end{aligned}$$

Therefore, we have $\{T_t \mathbb{1}_A = 0\} \subset \{\mathsf{d}_A = \infty\}$ a.e. by Theorem 1.1. $\square$

Now we prove Theorem 1.3. We start by recalling that $\Xi_t = \sqrt{\Phi_t + 1}$ is a bounded sequence in L^2, so it is weakly relatively compact. Now $x \mapsto \sqrt{x+1}$ is a concave function so we apply Lemma 2.2 and (2.23) to deduce that

$$\Xi_0 \le \sqrt{\Phi\left(\frac{\mathsf{d}_A^2}{2}\right) + 1} = \Xi\left(\frac{\mathsf{d}_A^2}{2}\right) \qquad \text{a.e.}$$

for any weak limit Ξ_0 of $\{\Xi_t\}$. On the other hand, Lemma 2.9 states that $\Xi_0 \ge \Xi(\mathsf{d}_A^2/2)$ a.e. Therefore, $\Xi_t \rightharpoonup \Xi(\mathsf{d}_A^2/2)$ weakly in L^2. But, since we already had that $\Phi_t \rightharpoonup \Phi(\mathsf{d}_A^2/2)$ weakly,

$$\|\Xi_t\|_{L^2}^2 = (\Phi_t + 1, 1)_{L^2} \to \big(\Phi(\mathsf{d}_A^2/2) + 1, 1\big)_{L^2} = \|\Xi(\mathsf{d}_A^2/2)\|_{L^2}^2 \qquad \text{as } t \downarrow 0.$$

This implies that Ξ_t converges strongly. Therefore, when $F = \Xi$ the second assertion of Theorem 1.3 follows from here. Because of Lemma 2.16, we can restrict ourselves to $\{\mathsf{d}_A < \infty\}$. Then the convergence in probability of u_t will follow from the strong convergence of Ξ_t since the function Ξ is one to one and has a continuous inverse. This same fact implies that any other function of u_t can be expressed as the composition of a continuous function and Ξ_t on $\{\mathsf{d}_A < \infty\}$. The complete statement of the theorem easily follows from all this.

3. Remarks on some examples of finite dimensional spaces.

3.1. *Diffusions with degenerated coefficients.* Consider, on $\mathbb{R}^d$, a Dirichlet form given by

$$\mathcal{E}(f, g) = \frac{1}{2} \int_{\mathbb{R}^d} \langle \nabla f, a \nabla g \rangle \, d\mu,$$

where the matrix function $a : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is nonnegative definite, but might have degeneracies. Notice that by requiring $\mathcal{E}$ to be a Dirichlet form, that is, a closed form, we put restrictions on the degeneracies of a. For a complete characterization of allowed a's in one dimension and some partial results in $d > 1$, see [20], Theorem 3.1.6.

Under the additional condition of uniform strong ellipticity, it was proven by Davies [10] and Norris and Stroock [34] that Varadhan's formula holds, that is,

$$\lim_{t \downarrow 0} 2t \log p_t(x, y) = -d(x, y)^2, \tag{3.1}$$

where $p_t(x, y)$ is the heat kernel associated with the Dirichlet form. This easily implies our result Theorem 1.1 in this particular case as long as the sets are good enough. Hence, our work generalizes the formula to the degenerate case. Of course, it will be most interesting to prove Varadhan's formula (3.1) in this degenerate case and not just an integrated version of it. This problem has also been studied in the case of elliptic operators in Hörmander's sum of squares form ([33] contains references on this).

3.2. *Diffusions on fractals.* At a first glance, Theorem 1.1 suggests that the (rough) Gaussian estimate is universal for symmetric diffusions. It seems contradictory to the fact that many diffusions on fractal sets have different asymptotics from the Gaussian type. It is not, of course. Let us consider the Brownian motion on the Sierpinski gasket as an example. The associated Dirichlet form $(\mathcal{E}, \mathbb{D})$ with the underlying Hausdorff measure μ lies in the setting of Theorem 1.1. It is known that for all functions in the domain of the associated Dirichlet form except constant functions, their energy measures are singular with respect to μ (see [29]). Therefore, in this case, $\mathbb{D}_0$ consists of only constant functions. Then $\mathsf{d}(A, B) = 0$ for every A and B with positive measure. On the other hand, the following detailed estimate of the transition density p_t is known:

$$
\begin{aligned}
& c_1 t^{-d_s/2} \exp\left(-c_2\left(\frac{|x-y|^{d_w}}{t}\right)^{1/(d_w-1)}\right) \\
& \quad \le p_t(x,y) \le c_3 t^{-d_s/2} \exp\left(-c_4\left(\frac{|x-y|^{d_w}}{t}\right)^{1/(d_w-1)}\right).
\end{aligned}
\tag{3.2}
$$

Here $d_s = 2\log 3/\log 5$ (spectral dimension), $d_w = \log 5/\log 2$ (walk dimension) and c_i $(i = 1, 2, 3, 4)$ is a constant independent of t, x and y. Since $d_w > 2$, $1/(d_w - 1) < 1$ and there is no inconsistency; Theorem 1.1 treats just a degenerate asymptotics.

Let us now consider this situation from the reverse side. Suppose that $(\mathcal{E}, \mathbb{D})$ is irreducible. Then, noticing the fact that $\mathsf{d}(A, B) < \infty$ if A and B have positive measures, we conclude that d_w has to be greater than or equal to 2 if the transition density satisfies an (upper side) estimate in (3.2). The restriction of d_w of this type has been discussed in various frameworks (e.g., [6], [25] and [28]).

4. Identification of the intrinsic metric on loop groups.

4.1. *Definitions.* In this section, we give some auxiliary definitions in addition to what was stated about path and loop groups in the Introduction. We keep the notation introduced there.

We equip $\mathcal{P}G$ and $\mathcal{L}G$ with distance $\rho(g_1, g_2) = \sup_{0\le t\le T} \rho_G(g_1(t), g_2(t))$, where ρ_G is a left-invariant distance on G. Then they become separable and complete metric spaces. For a subset U of G, we introduce a subset $\mathcal{P}_U G$ of $\mathcal{P}G$ by

$$
\mathcal{P}_U G = \{g \in \mathcal{P}G \mid g(T) \in U\}.
$$

Both $\mathcal{P}\mathfrak{g}$ and $\mathcal{L}\mathfrak{g}$ are separable Banach spaces under the supremum norm. We denote the norm of H by $\|h\| = (h|h)^{1/2}$. The following elementary fact is proven simply by the Schwarz inequality.

LEMMA 4.1. *For each $h \in H$,*

$$\sup_{0 \le s \le T} |h(s)|_{\mathfrak{g}} \le \sqrt{T}\|h\|.$$

The space $\mathcal{C}$ (resp. $\mathcal{F}C_{\mathrm{b}}^{\infty}$) of smooth cylindrical functions on $\mathcal{P}G$ (resp. $\mathcal{P}\mathfrak{g}$) is defined by

$$\begin{aligned} \mathcal{C} &= \{f(g) = F(g(t_1), \dots, g(t_n)) \mid n \in \mathbb{N}, 0 < t_1 < \cdots < t_n \le T, \\ &\qquad F \in C^{\infty}(G^n)\}, \qquad (4.1) \\ \mathcal{F}C_{\mathrm{b}}^{\infty} &= \{f(w) = F(l_1(w), \dots, l_n(w)) \mid n \in \mathbb{N},\ l_i \in (\mathcal{P}\mathfrak{g})^*,\ F \in C_{\mathrm{b}}^{\infty}(\mathbb{R}^n)\}, \end{aligned}$$

where $(\mathcal{P}\mathfrak{g})^*$ represents the topological dual space of $\mathcal{P}\mathfrak{g}$ and $C_{\mathrm{b}}^{\infty}(\mathbb{R}^n)$ denotes the totality of bounded C^{∞}-functions on $\mathbb{R}^n$, the derivatives of which are all bounded.

We will define the Brownian motion measure μ on $\mathcal{P}G$. Consider the following SDE of Fisk–Stratonovich type:

$$du(t) = u(t) \circ dw(t), \qquad u(0) = e,$$

where $\{w(t)\}$ is a $\mathfrak{g}$-valued Brownian motion starting at 0. This SDE has a strong solution; namely, there exists an Itô map $I : \mathcal{P}\mathfrak{g} \to \mathcal{P}G$ such that $I(w) = u$. The Brownian motion measure μ is the induced measure of the Wiener measure λ on $\mathcal{P}\mathfrak{g}$ by I.

We will write $|||g|||_{\mathcal{P}G} = d^{\mathcal{P}}(g, \mathbf{e})$ for $g \in \mathcal{P}G$. We introduce a subset $\mathcal{F}\mathcal{P}G$ of $\mathcal{P}G$ by

$$\mathcal{F}\mathcal{P}G = \{g \in \mathcal{P}G \mid |||g|||_{\mathcal{P}G} < \infty\}.$$

For $g \in \mathcal{P}G$ and $h \in H$, we define $h + g \in \mathcal{P}G$ by $(h + g)(t) = v(t)g(t)$, where $v \in \mathcal{F}\mathcal{P}G$ is a unique solution to

$$v(t)^{-1}\dot{v}(t) = (\mathrm{Ad}\, g(t))\dot{h}(t), \qquad v(0) = e. \tag{4.2}$$

Conversely, given $g \in \mathcal{P}G$ and $v \in \mathcal{F}\mathcal{P}G$, there is a unique $h \in H$ such that (4.2) holds and $|||v|||_{\mathcal{P}G} = \|h\|$ from the Ad_G-invariance of the inner product of $\mathfrak{g}$. Namely, $d^{\mathcal{P}}(g_1, g_2) = \|h\|$ if $h + g_2 = g_1$. From Theorem 2.4 in [23], we also have $0 + g = g$ and $(h_1 + h_2) + g = h_1 + (h_2 + g)$ for $h_1, h_2 \in H$, $g \in \mathcal{P}G$. Therefore, by considering the map $I_0 : H \ni h \mapsto h + \mathbf{e} \in \mathcal{F}\mathcal{P}G$, we see that $\mathcal{F}\mathcal{P}G$ is a metric space under $d^{\mathcal{P}}$ and it is homeomorphic to H.

A function f on (a subset of) $\mathcal{P}G$ is called $d^{\mathcal{P}}$-Lipschitz if there exists a constant C such that $|f(g_1) - f(g_2)| \le C d^{\mathcal{P}}(g_1, g_2)$ for all g_1 and g_2. The best constant C is called the $d^{\mathcal{P}}$-Lipschitz constant of f.

We will define a distance-like function on $\mathcal{L}G$. First, for a continuous curve γ on $\mathcal{L}G$, namely, for $\gamma \in C([0, 1] \to \mathcal{L}G)$, we define the "$d^{\mathcal{P}}$-length" $\ell(\gamma)$ of γ by

$$\ell(\gamma) = \sup_{\substack{\Delta = \{0 = s_0 < s_1 < \cdots < s_n = 1\}: \\ \text{finite partition of } [0,1]}} \sum_i d^{\mathcal{P}}(\gamma(s_{i-1}), \gamma(s_i)) \in [0, \infty]. \tag{4.3}$$

Then we set, for $g_1, g_2 \in \mathcal{L}G$,

$$d^{\mathcal{L}}(g_1, g_2) = \inf_{\substack{\gamma \in C([0,1] \to \mathcal{L}G) \\ \gamma(0) = g_1,\ \gamma(1) = g_2}} \ell(\gamma) \in [0, \infty]. \tag{4.4}$$

Note that $d^{\mathcal{L}}(g_1, g_2) = d^{\mathcal{L}}(g_2, g_1) = d^{\mathcal{L}}(g_1 g_2^{-1}, \mathbf{e})$. We write $|||g|||_{\mathcal{L}G} = d^{\mathcal{L}}(g, \mathbf{e})$ for $g \in \mathcal{L}G$ and set $\mathcal{F}\mathcal{L}G = \{g \in \mathcal{L}G \mid |||g|||_{\mathcal{L}G} < \infty\}$. Clearly, $\mathcal{F}\mathcal{L}G \subset \mathcal{F}\mathcal{P}G \cap \mathcal{L}G$ and $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$ is a metric space. We define the notions of $d^{\mathcal{L}}$-Lipschitz functions and $d^{\mathcal{L}}$-Lipschitz constants in the natural way.

A Borel probability measure ν on $\mathcal{L}G$ is assumed to satisfy the following:

(M0) The pre-Dirichlet form $(\mathcal{E}^{\mathcal{L}}, \mathcal{C})$ defined in (1.6) is well defined and closable on $L^2(\mathcal{L}G, \nu)$.

Denote its closure by $(\mathcal{E}^{\mathcal{L}}, \mathcal{F}^{\mathcal{L}})$. The operator $\nabla^{\mathcal{L}}$ extends continuously to $\mathcal{F}^{\mathcal{L}}$. By a simple calculation, we have $I_f(h) = \int_{\mathcal{L}G} h \|\nabla^{\mathcal{L}} f\|^2 \, d\nu$ for f, $h \in \mathcal{F}^{\mathcal{L}} \cap L^\infty(\nu) = \mathbb{D}_b$, and $\mathbb{D}_0$, which is defined in Section 1, is expressed as

$$\mathbb{D}_0 = \{f \in \mathbb{D}_b \mid \|\nabla^{\mathcal{L}} f\| \le 1 \ \nu\text{-a.e.}\}.$$

We also extend $\nabla^{\mathcal{P}}$ continuously to $\mathcal{F}^{\mathcal{P}}$ for later use.

For $h \in H_0$, the shift operator θ_h on $\mathcal{L}G$ will be defined by $\theta_h(g) = e^h g$.

We further introduce the following conditions:

(M1) (Quasi-invariance of ν.) There is a dense subspace H_0' of H_0 and a constant $\delta_0 > 0$ such that, for every $h \in H_0'$ with $\|h\| \le \delta_0$, the induced measure of ν by θ_h is absolutely continuous with respect to ν (i.e., $\nu \circ \theta_h^{-1} \ll \nu$) and it holds that

$$\int_0^1 \left\| \frac{d\nu \circ \theta_{sh}^{-1}}{d\nu} \right\|_{L^2(\nu)} ds < \infty. \tag{4.5}$$

(M2) The measure ν is absolutely continuous with respect to μ_e, and the Radon–Nikodym derivative $d\nu/d\mu_e$ belongs to $\bigcup_{p>1} L^p(\mu_e)$. Here μ_e is the pinned Brownian motion measure, which is a conditional probability measure of μ given by $g(T) = e$.

In order to avoid the problem of measurability, we mean by (4.5) the following: there exists a Lebesgue integrable function ψ on $[0, 1]$ such that $\|d\nu \circ \theta_{sh}^{-1}/d\nu\|_{L^2(\nu)} \le \psi(s)$ for every $s \in [0, 1]$. Note that, under (M1), $\nu \circ \theta_h^{-1} \ll \nu$ for every $h \in H_0'$. This is proven by showing inductively that $\nu \circ \theta_{(j/N)h}^{-1} \ll \nu \circ \theta_{((j-1)/N)h}^{-1}$ for $j = 1, \dots, N$, $N \ge \delta_0/\|h\|$.

The typical examples of ν are given as follows.

PROPOSITION 4.2. *The following measures on $\mathcal{L}G$ satisfy* (M0), (M1) *and* (M2).

(i) *The pinned Brownian motion measure* μ_e.

(ii) *The heat kernel measure* μ_{heat} *with parameter* 1. *That is*, $T = 1$ *and* μ_{heat} *is the law at time* 1 *of the Brownian motion on* $\mathcal{L}G$ *starting at* $\mathbf{e}$. (*See* [13] *for a more precise definition*.)

4.2. *Preliminaries for the proofs of Theorem* 1.5 *and Proposition* 4.2. We split Theorem 1.5 into the following three claims. In the following, we always assume that ν satisfies (M0), (M1) and (M2), and A is a Suslin set of $\mathcal{L}G$.

PROPOSITION 4.3. *The function* $\bar{\mathsf{d}}_A^{\mathcal{L}}$ *is universally measurable*.

PROPOSITION 4.4. $\mathsf{d}_A^{\mathcal{L}} \le \bar{\mathsf{d}}_A^{\mathcal{L}}$, ν*-a.e if* A *is also* $d^{\mathcal{L}}$*-open*.

PROPOSITION 4.5. $\mathsf{d}_A^{\mathcal{L}} \ge \bar{\mathsf{d}}_A^{\mathcal{L}}$, ν*-a.e*.

We need some preparation for the proof of the propositions above. Let us first briefly review the quasi-sure analysis. (See, e.g., [27] and [32] for references.) Let L be the Ornstein–Uhlenbeck operator on $\mathcal{P}\mathfrak{g}$. For $p > 1$ and $r \in \mathbb{Z}_+$, the Sobolev space $\mathbb{D}^{r,p}(\mathcal{P}\mathfrak{g})$ on $\mathcal{P}\mathfrak{g}$ is a range of the operator $(1-L)^{-r/2}$ on $L^p(\mathcal{P}\mathfrak{g}, \lambda)$ with norm

$$\|f\|_{r,p} = \big\|(1-L)^{r/2} f\big\|_{L^p(\lambda)}.$$

The associated capacity $\mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}}$ is defined by

$$\mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}}(O) = \inf\big\{\|f\|_{r,p} \mid f \in \mathbb{D}^{r,p}(\mathcal{P}\mathfrak{g}),\ f \ge 0\ \lambda\text{-a.e. and } f \ge 1 \text{ on } O\ \lambda\text{-a.e.}\big\}$$

for open sets O in $\mathcal{P}\mathfrak{g}$, and

$$\mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}}(A) = \inf\big\{\mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}}(O) \mid O\colon \text{open}, O \supset A\big\}$$

for general sets A.

A function f on $\mathcal{P}\mathfrak{g}$ is called (r, p)-quasicontinuous if there exists a sequence of closed sets $\{A_k\}_{k=1}^{\infty}$ of $\mathcal{P}\mathfrak{g}$ such that $\mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}}(\mathcal{P}\mathfrak{g} \setminus A_k)$ converges to 0 as $k \to \infty$ and f is continuous on each A_k. A function f is called ∞-quasicontinuous if f is (r, p)-quasicontinuous for all $p > 1$ and $r \in \mathbb{Z}_+$. We say that a certain assertion holds ∞-quasi everywhere (∞-q.e. for short) if it holds except a set of zero (r, p)-capacity for every $p > 1$ and $r \in \mathbb{Z}_+$. Let $\mathbb{D}^{\infty}(\mathcal{P}\mathfrak{g}) = \bigcap_{p>1,\, r\in\mathbb{Z}_+} \mathbb{D}^{r,p}(\mathcal{P}\mathfrak{g})$. Each function f in $\mathbb{D}^{\infty}(\mathcal{P}\mathfrak{g})$ has an ∞-quasicontinuous modification, denoted by $\tilde{f}$ from now on. It is unique in the sense that if $\tilde{f}_1$ and $\tilde{f}_2$ are both ∞-quasicontinuous modifications, then $\tilde{f}_1 = \tilde{f}_2$, ∞-q.e. When $\{f_n\}$ converges to f in $\mathbb{D}^{\infty}(\mathcal{P}\mathfrak{g})$, we can take a subsequence $\{f_{n_k}\}$ such that its ∞-quasicontinuous modification $\{\tilde{f}_{n_k}\}$ converges to $\tilde{f}$, ∞-q.e.

The Itô map I induces a measure-theoretical isometry between $(\mathcal{P}\mathfrak{g}, \lambda)$ and $(\mathcal{P}G, \mu)$. Accordingly, it induces Sobolev spaces $\mathbb{D}^{r,p}(\mathcal{P}G)$ on $\mathcal{P}G$. Then, capacities $\mathrm{Cap}_{r,p}^{\mathcal{P}G}$ and quasi notions associated with them are also defined on $\mathcal{P}G$ in the same way as in $\mathcal{P}\mathfrak{g}$, and the same properties hold. Indeed, due to Shigekawa [40], I has an ∞-quasicontinuous modification (which will be fixed hereafter) inducing a quasihomeomorphism between $(\mathcal{P}\mathfrak{g}, \mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}})$ and $(\mathcal{P}G, \mathrm{Cap}_{r,p}^{\mathcal{P}G})$, and $\mathfrak{C}$ is dense in $\mathbb{D}^{r,p}(\mathcal{P}G)$ for every $p > 1$ and $r \in \mathbb{Z}_+$. (For the precise definition of quasihomeomorphism, see [40].) Furthermore, $\mathcal{F}^{\mathcal{P}} = \mathbb{D}^{1,2}(\mathcal{P}G)$ and it holds for $f \in \mathbb{D}^{1,2}(\mathcal{P}\mathfrak{g})$ that $\|D(f \circ I)\| = \|\nabla^{\mathcal{P}} f\|$, λ-a.e., where D represents the H-derivative in the sense of Malliavin calculus (see also Gross [22]).

For each $x \in G$, there exists some $p > 1$ and $r \in \mathbb{Z}_+$ such that the conditional measure λ_x of λ given by $I(w)(T) = x$ can be realized by an element of the topological dual space of $\mathbb{D}^{r,p}(\mathcal{P}\mathfrak{g})$, and $\mathrm{Cap}_{r,p}^{\mathcal{P}\mathfrak{g}}(A) = 0$ implies $\lambda_x(A) = 0$. Accordingly, the conditional measure μ_x of μ given by $g(T) = x$ is identified with an element of the dual space of $\mathbb{D}^{r,p}(\mathcal{P}G)$, and $\mathrm{Cap}_{r,p}^{\mathcal{P}G}(A) = 0$ implies $\mu_x(A) = 0$. The relation $\lambda_x \circ I^{-1} = \mu_x$ holds. Also, the following disintegration formula holds:

$$\int_G \left[\int_{\mathcal{P}G} f(g)\mu_x(dg) \right] p_T(x)m(dx) = \int_{\mathcal{P}G} f(g)\mu(dg) \tag{4.6}$$

for any (bounded or positive) Borel function f on $\mathcal{P}G$, where m is the Haar measure on G and p_T is a density of the law of $I(\cdot)(T)$ with respect to m, which is a strictly positive function.

For each $v \in \mathcal{P}G$, l_v denotes the multiplication of v from the left-hand side: $\mathcal{P}G \ni g \mapsto l_v g = vg \in \mathcal{P}G$. The following theorem is due to Malliavin and Malliavin [31]. See also [38], Lemma 3.1, for the proof.

THEOREM 4.6. *For each $v \in \mathcal{F}\mathcal{P}G$, the measure $\mu_e \circ l_v^{-1}$ is mutually absolutely continuous with respect to $\mu_{v(T)}$ and its Radon–Nikodym derivative $J_v := \frac{d(\mu_e \circ l_v^{-1})}{d\mu_{v(T)}}$ is given by*

$$J_v(u) = \frac{p_T(v(T))}{p_T(e)} \times j_v \circ I^{-1}(u),$$

where

$$j_v(w) = \exp\left(\int_0^T \big(\mathrm{Ad}\, I(w)(s)\big)\big(v(s)^{-1}\dot{v}(s)\big)\, dw(s) - \frac{\|\!|v|\!\|_{\mathcal{P}G}^2}{2} \right).$$

Here the ∞-quasicontinuous modification should be taken.

LEMMA 4.7. *For each $x \in G$, there exists a nondecreasing function ψ_x on $[0, \infty)$ such that $\|J_v\|_{L^2(\mu_x)} \le \psi_x(\|\!|v|\!\|_{\mathcal{P}G})$ for every $v \in \mathcal{F}\mathcal{P}G \cap \mathcal{P}_{\{x\}}G$.*

PROOF. Let $M_v = \int_0^T (\mathrm{Ad}\, I(w)(s))(v(s)^{-1}\dot{v}(s))\, dw(s)$. From a standard argument in the proof of the differentiability of the solution of SDE such as in [27], Chapter 5, Proposition 10.1, or [35], Theorem 2.2.2, $M_v \in \mathbb{D}^\infty(\mathcal{P}\mathfrak{g})$ and $\|M_v\|_{r,p} \le c_{r,p}|\!|\!|v|\!|\!|_{\mathcal{P}G}$ for some constant $c_{r,p}$ independent of v, for each $r \in \mathbb{Z}_+$ and $p > 1$. Then, for p and r large enough,

$$\|J_v\|_{L^2(\mu_x)} = \|j_v\|_{L^2(\lambda_x)} \le \mathrm{const} \cdot \|j_v^2\|_{r,p}^{1/2} \le \mathrm{const} \cdot \|\exp(2M_v - |\!|\!|v|\!|\!|^2_{\mathcal{P}G})\|_{r,p}^{1/2}.$$

We have

$$\begin{aligned}
&\|\exp(2M_v - |\!|\!|v|\!|\!|^2_{\mathcal{P}G})\|_{L^p(\lambda)} \\
&\quad = \left(\int_{\mathbb{R}} \exp(2ps - p|\!|\!|v|\!|\!|^2_{\mathcal{P}G}) \frac{1}{\sqrt{2\pi |\!|\!|v|\!|\!|^2_{\mathcal{P}G}}} \exp\left(\frac{-s^2}{2|\!|\!|v|\!|\!|_{\mathcal{P}G}}\right) ds\right)^{1/p} \\
&\quad = \exp\big((2p-1)|\!|\!|v|\!|\!|^2_{\mathcal{P}G}\big)
\end{aligned}$$

and

$$\begin{aligned}
&\|D\exp(2M_v - |\!|\!|v|\!|\!|^2_{\mathcal{P}G})\|_{L^p(\lambda)} \\
&\quad = \|2(DM_v)\exp(2M_v - |\!|\!|v|\!|\!|^2_{\mathcal{P}G})\|_{L^p(\lambda)} \\
&\quad \le 2\|M_v\|_{1,2p}^{1/2} \|\exp(2M_v - |\!|\!|v|\!|\!|^2_{\mathcal{P}G})\|_{L^{2p}(\lambda)}^{1/2} \\
&\quad \le 2c_{1,2p}^{1/2} |\!|\!|v|\!|\!|^{1/2}_{\mathcal{P}G} \exp\big((2p - 1/2)|\!|\!|v|\!|\!|^2_{\mathcal{P}G}\big).
\end{aligned}$$

Inductively, we get a similar estimate of $\|\exp(2M_v - |\!|\!|v|\!|\!|^2_{\mathcal{P}G})\|_{r,p}$ and reach the conclusion. □

Take an open neighborhood U of e in G so that U is diffeomorphic to an open ball centered at 0 in $\mathfrak{g}$ via exp, the exponential map. The inverse map of exp will be denoted by Log. Recall that the map $H \ni h \mapsto h + \mathbf{e} \in \mathcal{P}G$ is denoted by I_0. Then the following result holds (see [23], Lemma 2.1, and its proof).

LEMMA 4.8. (i) *For each $\varepsilon > 0$, there exists a neighborhood U_ε of e which is an image of an open ball centered at* 0 *in $\mathfrak{g}$ by* exp *and contained in U, enjoying the following property: if $v \in \mathcal{F}\mathcal{P}G$ satisfies $v(t) \in U_\varepsilon$, $t \in [0,T]$, then* $\operatorname{Log} v \in H$ *and* $\|\operatorname{Log} v\| \le (1+\varepsilon)|\!|\!|v|\!|\!|_{\mathcal{P}G}$.

(ii) *There is an open ball S in H centered at* 0 *such that $I_0(S) \subset \{v \in \mathcal{P}G \mid v(t) \in U,\ t \in [0,T]\}$ and $\chi := \operatorname{Log} \circ I_0$ is C^∞-diffeomorphic from S into H. Moreover, the Fréchet derivative χ' satisfies $\chi'(0) = \mathrm{Id}_H$.*

In particular, we can take $\varepsilon_0 > 0$ such that $\{h \in H \mid \|h\| \le \varepsilon_0\} \subset S \cap \chi(S)$ and

$$\sup_{\|h\| \le \varepsilon_0,\, \|k\| \le \varepsilon_0,\, h \ne k} \left(\frac{\|\chi(h) - \chi(k)\|}{\|h-k\|} \vee \frac{\|\chi^{-1}(h) - \chi^{-1}(k)\|}{\|h-k\|}\right) < \infty.$$

REMARK 4.9. From this lemma, any open set in $\mathcal{P}G$ is $d^{\mathcal{P}}$-open. In addition, since $|||v|||_{\mathcal{P}G} \le |||v|||_{\mathcal{L}G}$ for $v \in \mathcal{F}\mathcal{L}G$, any open set in $\mathcal{L}G$ is $d^{\mathcal{L}}$-open.

4.3. *Proofs of Propositions* 4.2–4.4.

PROOF OF PROPOSITION 4.2. (i) The condition (M0) follows from, for example, [24], Theorem 10.4. Let $h \in H_0$ with $\|h\| \le \varepsilon_0$. Then, from Lemma 4.8, $c := \sup_{0\le s\le 1} |||e^{sh}|||_{\mathcal{P}G} = \sup_{0\le s\le 1} \|\chi^{-1}(sh)\| < \infty$. Thus, by Lemma 4.7,

$$\left\| \frac{d\mu_e \circ \theta_{sh}^{-1}}{d\mu_e} \right\|_{L^2(\mu_e)} = \|J_{e^{sh}}\|_{L^2(\mu_e)} \le \psi_e(c).$$

Therefore, μ_e satisfies (M1). It is clear that (M2) holds.

(ii) The condition (M0) is proven in Theorem 4.14 of [13]. From Corollary 7.10, together with Theorem 7.4, Lemma 7.6 and Proposition 7.9 in [13], μ_{heat} satisfies (M1). The property (M2) is due to [14]. □

LEMMA 4.10. *For each* $\varepsilon \in (0, \varepsilon_0]$, *there exists some constant* $C_0 = C_0(\varepsilon) > 0$ *such that* $|||v|||_{\mathcal{L}G} \le C_0 |||v|||_{\mathcal{P}G}$ *for all* $v \in \mathcal{L}G$ *with* $|||v|||_{\mathcal{P}G} \le \varepsilon$. *Moreover*, $C_0(\varepsilon)$ *can be taken so that* $\lim_{\varepsilon\to 0} C_0(\varepsilon) = 1$.

PROOF. Let $v \in \mathcal{L}G$ satisfy $|||v|||_{\mathcal{P}G} \le \varepsilon$. Then $v(t) \in U$ for all $t \in [0, T]$ and $k = \operatorname{Log} v \in H_0$ is well defined. Define $\gamma \in C([0, 1] \to \mathcal{L}G)$ connecting $\mathbf{e}$ and v in $\mathcal{L}G$ by

$$\gamma(s)(t) = \exp(sk(t)), \qquad s \in [0, 1],\ t \in [0, T].$$

Then

$$|||v|||_{\mathcal{L}G} = d^{\mathcal{L}}(\mathbf{e}, v) \le \ell(\gamma) = \sup_{\Delta} \sum_i d^{\mathcal{P}}\big(\gamma(s_{i-1}), \gamma(s_i)\big),$$

where sup is taken for all finite partitions $\Delta = \{0 = s_0 < s_1 < \cdots < s_n = 1\}$.

Let $C(\varepsilon) = \sup_{h\in H,\ 0<\|h\|\le\varepsilon} (\|\chi(h)\|/\|h\|) \vee (\|\chi^{-1}(h)\|/\|h\|) < \infty$. Then

$$\begin{aligned} d^{\mathcal{P}}\big(\gamma(s_{i-1}), \gamma(s_i)\big) &= \big|\big|\big| \exp\big(s_{i-1}k(\cdot)\big) \exp\big(s_i k(\cdot)\big)^{-1} \big|\big|\big|_{\mathcal{P}G} \\ &= \big|\big|\big| \exp\big((s_{i-1} - s_i)k(\cdot)\big) \big|\big|\big|_{\mathcal{P}G} \\ &= \big|\big|\big| (I_0 \circ \chi^{-1})\big((s_{i-1} - s_i)k\big) \big|\big|\big|_{\mathcal{P}G} \\ &= \big\| \chi^{-1}\big((s_{i-1} - s_i)k\big) \big\| \\ &\le C(\varepsilon) \|(s_{i-1} - s_i)k\| \\ &= C(\varepsilon)(s_i - s_{i-1}) \|\chi \circ I_0^{-1}(v)\| \\ &\le C(\varepsilon)^2 (s_i - s_{i-1}) |||v|||_{\mathcal{P}G}. \end{aligned}$$

Therefore, we have

$$\|\!|v\|\!|_{\mathcal{L}G} \le \sup_{\Delta} \sum_i C(\varepsilon)^2 (s_i - s_{i-1}) \|\!|v\|\!|_{\mathcal{P}G} = C(\varepsilon)^2 \|\!|v\|\!|_{\mathcal{P}G}.$$

Moreover, $C(\varepsilon) \to 1$ as $\varepsilon \to 0$ since $\chi'(0) = \mathrm{Id}_H$. $\square$

LEMMA 4.11. *Let $h \in H_0$. Then $e^h \in \mathcal{F}\mathcal{L}G$ and $\|\!|e^{sh}\|\!|_{\mathcal{L}G}/|s| \to \|h\|$ as $s \to 0$.*

PROOF. When N is a sufficiently large integer, $e^{h/N} \in \mathcal{F}\mathcal{L}G$ by Lemmas 4.8 and 4.10. Then $\|\!|e^h\|\!|_{\mathcal{L}G} \le \sum_{i=1}^N d^{\mathcal{L}}(e^{(i-1)h/N}, e^{ih/N}) \le N\|\!|e^{h/N}\|\!|_{\mathcal{L}G} < \infty$. By virtue of Lemma 4.8,

$$\lim_{s\to 0} \frac{\|\!|e^{sh}\|\!|_{\mathcal{P}G}}{|s|} = \lim_{s\to 0} \frac{\|\chi^{-1}(sh)\|}{|s|} = \|h\|.$$

Since $\|\!|e^{sh}\|\!|_{\mathcal{P}G} \le \|\!|e^{sh}\|\!|_{\mathcal{L}G} \le C_0(\|\!|e^{sh}\|\!|_{\mathcal{P}G})\|\!|e^{sh}\|\!|_{\mathcal{P}G}$ and $C_0(\|\!|e^{sh}\|\!|_{\mathcal{P}G}) \to 1$ as $s \to 0$ by Lemma 4.10, we obtain the second assertion. $\square$

LEMMA 4.12. *Both $(\mathcal{F}\mathcal{P}G, d^{\mathcal{P}})$ and $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$ are Polish spaces. Moreover, $\mathcal{F}\mathcal{L}G$ is closed in $(\mathcal{F}\mathcal{P}G, d^{\mathcal{P}})$, and $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$ is homeomorphic to $(\mathcal{F}\mathcal{L}G, d^{\mathcal{P}}|_{\mathcal{F}\mathcal{L}G})$.*

PROOF. Since $(\mathcal{F}\mathcal{P}G, d^{\mathcal{P}})$ is homeomorphic to H, it is a Polish space. Suppose that a sequence $\{v_n\}$ in $\mathcal{F}\mathcal{L}G$ converges to v in $(\mathcal{F}\mathcal{P}G, d^{\mathcal{P}})$. Then $v \in \mathcal{L}G$. For large n, $d^{\mathcal{P}}(v_n, v) \le \varepsilon_0$. Then $d^{\mathcal{L}}(v_n, v) \le C_0 d^{\mathcal{P}}(v_n, v)$ by Lemma 4.10. Therefore, $\|\!|v\|\!|_{\mathcal{L}G} \le \|\!|v_n\|\!|_{\mathcal{L}G} + C_0\varepsilon_0 < \infty$, which means $v \in \mathcal{F}\mathcal{L}G$, and $v_n \to v$ in $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$. Hence, $\mathcal{F}\mathcal{L}G$ is closed in $(\mathcal{F}\mathcal{P}G, d^{\mathcal{P}})$ and $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$ is homeomorphic to $(\mathcal{F}\mathcal{L}G, d^{\mathcal{P}}|_{\mathcal{F}\mathcal{L}G})$. In particular, $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$ is a Polish space. $\square$

PROOF OF PROPOSITION 4.3. Let $p : \mathcal{F}\mathcal{L}G \times \mathcal{L}G \to \mathcal{L}G$ be defined by $p(v, g) = vg$. This is a continuous map when $\mathcal{F}\mathcal{L}G$ and $\mathcal{L}G$ are equipped with distance $d^{\mathcal{L}}$ and ρ, respectively. For each $r > 0$, $\{\bar{\mathsf{d}}_A^{\mathcal{L}} \le r\} = \bigcap_{n\in\mathbb{N}} p(B_n \times A)$, where

$$B_n = \{v \in \mathcal{F}\mathcal{L}G \mid \|\!|v\|\!|_{\mathcal{L}G} \le r + 1/n\}.$$

Since B_n is closed, it is a Polish space with relative topology by Lemma 4.12. Therefore, $\{\bar{\mathsf{d}}_A^{\mathcal{L}} \le r\}$ is a countable intersection of Suslin sets and, in particular, universally measurable. This completes the proof. $\square$

LEMMA 4.13. *Let $\{u_n\}_{n\in\mathbb{N}}$ and u belong to $\mathcal{F}\mathcal{P}G$ and let $\{g_n\}_{n\in\mathbb{N}}$ and g belong to $\mathcal{P}G$. If $d^{\mathcal{P}}(u_n, u) \to 0$ and $d^{\mathcal{P}}(g_n, g) \to 0$ as $n \to \infty$, then $d^{\mathcal{P}}(u_n g_n, ug) \to 0$ as $n \to \infty$.*

PROOF. Let $w = u_n u^{-1}$ and $v = g_n g^{-1}$. Then, noting that $w^{-1}\dot{w} = (\mathrm{Ad}\, u) \times (u_n^{-1}\dot{u}_n - u^{-1}\dot{u})$, we have

$$\begin{aligned} d^{\mathcal{P}}(u_n g_n, ug)^2 &= |||u_n v u^{-1}|||_{\mathcal{P}G}^2 \\ &= \int_0^T \left|(u_n v u^{-1})^{-1}(\dot{u}_n v u^{-1} + u_n \dot{v} u^{-1} - u_n v u^{-1}\dot{u}u^{-1})\right|_{\mathfrak{g}}^2 dt \\ &= \int_0^T \left|(\mathrm{Ad}\, u)\{(\mathrm{Ad}\, v^{-1})(u_n^{-1}\dot{u}_n) + v^{-1}\dot{v} - u^{-1}\dot{u}\}\right|_{\mathfrak{g}}^2 dt \\ &= \int_0^T \left|(\mathrm{Ad}\, v^{-1} - \mathrm{Id}_{\mathfrak{g}})(u_n^{-1}\dot{u}_n) + v^{-1}\dot{v} + (\mathrm{Ad}\, u^{-1})(w^{-1}\dot{w})\right|_{\mathfrak{g}}^2 dt. \end{aligned}$$

Therefore,

$$\begin{aligned} d^{\mathcal{P}}(u_n g_n, ug) \le \sup_{0\le t\le T} &\left\| \mathrm{Ad}\left(g(t)g_n(t)^{-1}\right) - \mathrm{Id}_{\mathfrak{g}} \right\|_{\mathfrak{g}\to\mathfrak{g}} |||u_n|||_{\mathcal{P}G} \\ &+ d^{\mathcal{P}}(g_n, g) + d^{\mathcal{P}}(u_n, u), \end{aligned}$$

where $\|\cdot\|_{\mathfrak{g}\to\mathfrak{g}}$ means the operator norm from $\mathfrak{g}$ to $\mathfrak{g}$. The right-hand side converges to 0 since $\rho(gg_n^{-1}, \mathbf{e}) \to 0$ as $n \to \infty$ and $\mathrm{Ad}\, e = \mathrm{Id}_{\mathfrak{g}}$. $\square$

LEMMA 4.14. *For every $f \in \mathcal{F}^{\mathcal{L}}$, $0 \le a < b \le 1$ and $h \in H_0'$, it holds that*

$$f(e^{bh}g) - f(e^{ah}g) = \int_a^b \left(\nabla^{\mathcal{L}} f(e^{sh}g)|h\right) ds, \qquad \nu\text{-a.e.} \tag{4.7}$$

REMARK 4.15. Since $\nu \circ \theta_h^{-1} \ll \nu$ for every $h \in H_0'$, $f(e^h \cdot)$ is a well-defined measurable function and independent of the choice of the ν-version of f.

PROOF. It is enough to prove (4.7) when $\|h\| \le \delta_0$. Indeed, for general $h \in H_0'$, take an integer N so that $\|h/N\| \le \delta_0$. Then

$$\begin{aligned} &f(e^{bh}g) - f(e^{ah}g) \\ &\quad = \sum_{j=1}^N \left\{ f\left(e^{(b-a)h/N} e^{(b-a)(j-1)h/N+ah} g\right) - f\left(e^{(b-a)(j-1)h/N+ah} g\right)\right\} \\ &\quad = \sum_{j=1}^N \int_0^{b-a} \left(\nabla^{\mathcal{L}} f\left(e^{sh/N} e^{(b-a)(j-1)h/N+ah} g\right) \Big| \frac{h}{N}\right) ds \\ &\quad = \sum_{j=1}^N \int_{(b-a)(j-1)/N+a}^{(b-a)j/N+a} \left(\nabla^{\mathcal{L}} f(e^{sh}g)|h\right) ds \\ &\quad = \int_a^b \left(\nabla^{\mathcal{L}} f(e^{sh}g)|h\right) ds. \end{aligned}$$

When $f \in \mathcal{C}$, it is easy to see that (4.7) is true. For general f, fix $h \in H_0'$ with $\|h\| \le \delta_0$ and take a sequence $\{f_n\}$ in $\mathcal{C}$ converging to f in $\mathcal{F}^{\mathcal{L}}$ and ν-a.e. From the quasi-invariance of ν, $f_n(e^{sh}g)$ converges to $f(e^{sh}g)$, ν-a.e. for each $s \in [0,1]$. Furthermore,

$$\begin{aligned}
&\left\| \int_a^b \{(\nabla^{\mathcal{L}} f_n(e^{sh})|h) - (\nabla^{\mathcal{L}} f(e^{sh})|h)\}\, ds \right\|_{L^1(\nu)} \\
&\quad \le \|h\| \int_a^b \left\| \|\nabla^{\mathcal{L}} f_n - \nabla^{\mathcal{L}} f\| \frac{d\nu \circ \theta_{sh}^{-1}}{d\nu} \right\|_{L^1(\nu)} ds \\
&\quad \le \|h\| \mathcal{E}^{\mathcal{L}}(f_n - f, f_n - f)^{1/2} \int_a^b \left\| \frac{d\nu \circ \theta_{sh}^{-1}}{d\nu} \right\|_{L^2(\nu)} ds \\
&\quad \to 0 \qquad \text{as } n \to \infty \qquad \text{[by (M1)]}.
\end{aligned}$$

Therefore, by letting $n \to \infty$ along an appropriate sequence in (4.7) with f being replaced by f_n, we have (4.7) for $f \in \mathcal{F}^{\mathcal{L}}$. $\square$

LEMMA 4.16. *Let $f \in \mathcal{F}^{\mathcal{L}}$.*

(i) *If a ν-version of f is $d^{\mathcal{L}}$-Lipschitz continuous with $d^{\mathcal{L}}$-Lipschitz constant K, then $\|\nabla^{\mathcal{L}} f\| \le K$, ν-a.e.*

(ii) *If $\|\nabla^{\mathcal{L}} f\| \le K$, ν-a.e., then, for each $v \in \mathcal{F}\mathcal{L}G$ and $\varepsilon > 0$, there is a $v' \in \mathcal{F}\mathcal{L}G$ such that $d^{\mathcal{L}}(v, v') < \varepsilon$, $\nu \circ l_{v'}^{-1} \ll \nu$ and*

$$|f(v'g) - f(g)| \le K(1+\varepsilon)(|||v|||_{\mathcal{L}G} + \varepsilon), \qquad \nu\text{-a.e. } g.$$

PROOF. (i) Fix a Borel-measurable version of $\nabla^{\mathcal{L}} f$. For each $h \in H_0'$ with $\|h\| \le \delta_0$, define $\phi : (0,1) \times \mathcal{L}G \to \mathbb{R}$ by $\phi(s,g) = (\nabla^{\mathcal{L}} f(e^{sh}g)|h)$. By Lemma 4.14, for each $0 < a < b < 1$, for ν-a.e. g,

$$\begin{aligned}
&\left| \frac{1}{b-a} \int_a^b \phi(s,g)\, ds \right| \\
&\qquad \le \frac{1}{b-a} |f(e^{bh}g) - f(e^{ah}g)| \le \frac{K}{b-a} |||e^{(b-a)h}|||_{\mathcal{L}G},
\end{aligned} \tag{4.8}$$

and the last term converges to $K\|h\|$ as $b - a \to 0$ by Lemma 4.11. Take $X_1 \subset \mathcal{L}G$ such that $\nu(X_1) = 1$ and (4.8) holds for all rational numbers a and b in $(0,1)$ for $g \in X_1$. Since $\| \int_0^1 |\phi(s,\cdot)|\, ds\|_{L^1(\nu)} < \infty$, we can take $X_2 \subset X_1$ with full ν-measure such that $\phi(\cdot, g) \in L^1((0,1))$ for any $g \in X_2$. By Lebesgue's density theorem, for each $g \in X_2$, $|\phi(s,g)| \le K\|h\|$ for every Lebesgue point s of $\phi(\cdot, g)$, that is, for a.e. s in $(0,1)$ with respect to the Lebesgue measure. From Fubini's theorem, for a.e. s in $(0,1)$, for ν-a.e. g, $|\phi(s,g)| \le K\|h\|$. Fix such s. Then $|(\nabla^{\mathcal{L}} f(g)|h)| = |\phi(0,g)| = |\phi(s, e^{-sh}g)| \le K\|h\|$, ν-a.e. g. Since H_0' is dense in H_0, we get the conclusion.

(ii) Let $v \in \mathcal{F}\mathcal{L}G$ and $\varepsilon > 0$. There exists a $\gamma \in C([0,1] \to \mathcal{L}G)$ connecting $\mathbf{e}$ and v such that $\ell(\gamma) \le |||v|||_{\mathcal{L}G} + \varepsilon/2$. Let $\Delta = \{0 = s_0 < s_1 < \cdots < s_N = 1\}$ be a finite partition of $[0,1]$. Define $v_i \in \mathcal{F}\mathcal{P}G$, $i = 1, \dots, N$, by $v_i = \gamma(s_i)\gamma(s_{i-1})^{-1}$. If $\max_{1\le i\le N}(s_i - s_{i-1})$ is sufficiently small, then $v_i(t) \in U_\varepsilon$ for every i and $t \in [0,T]$, where U_ε is given in Lemma 4.8(i). Let $h_i = \operatorname{Log} v_i \in H_0$. Define $\tilde\gamma \in C([0,1] \to \mathcal{L}G)$ by $\tilde\gamma(0) = \mathbf{e}$ and

$$\tilde\gamma(s) = e^{(s-s_{i-1})h_i/(s_i-s_{i-1})}\tilde\gamma(s_{i-1}) \qquad \text{if } s \in [s_{i-1}, s_i],\ i = 1, \dots, N.$$

Clearly, $\tilde\gamma(s_i) = \gamma(s_i)$, $i = 0, 1, \dots, N$. Therefore, we can take $h_i' \in H_0'$ near to h_i so that $e^{h_i'}(t) \in U_\varepsilon$ for $t \in [0,T]$ and $\gamma' \in C([0,1] \to \mathcal{L}G)$ defined by $\gamma'(0) = \mathbf{e}$ and

$$\gamma'(s) = e^{(s-s_{i-1})h_i'/(s_i-s_{i-1})}\gamma'(s_{i-1}) \qquad \text{if } s \in [s_{i-1}, s_i],\ i = 1, \dots, N,$$

satisfies that $d^{\mathcal{P}}(\gamma(s_i), \gamma'(s_i)) \le \varepsilon/(4N)$, $i = 1, \dots, N$. Indeed, we can take h_i' such that $\gamma'(s_i) = (\prod_{j=1}^{n_i} e^{h_j'/n_j})\gamma'(s_{i-1})$ is sufficiently near to $\tilde\gamma(s_i) = (\prod_{j=1}^{n_i} e^{h_j/n_j})\tilde\gamma(s_{i-1})$ in $\mathcal{F}\mathcal{P}G$ in view of Lemmas 4.8(ii) and 4.13, where $n_i \in \mathbb{N}$ is taken so that $\|h_i\| < \varepsilon_0 n_i$. Let $v' = \gamma'(1)$. Then, from Lemmas 4.14 and 4.8(i),

$$\begin{aligned}
|f(v'g) - f(g)| &\le \sum_{i=1}^N \bigl|f\bigl(\gamma'(s_i)g\bigr) - f\bigl(\gamma'(s_{i-1})g\bigr)\bigr| \\
&= \sum_{i=1}^N \left|\int_0^1 \bigl(\nabla^{\mathcal{L}} f\bigl(e^{h_i'}\gamma'(s_{i-1})g\bigr)\big|h_i'\bigr)\,ds\right| \\
&\le \sum_{i=1}^N K\|h_i'\| \\
&\le \sum_{i=1}^N K(1+\varepsilon) d^{\mathcal{P}}\bigl(\gamma'(s_i), \gamma'(s_{i-1})\bigr) \\
&\le K(1+\varepsilon)\left(\varepsilon/2 + \sum_{i=1}^N d^{\mathcal{P}}\bigl(\gamma(s_i), \gamma(s_{i-1})\bigr)\right) \\
&\le K(1+\varepsilon)(\varepsilon/2 + \ell(\gamma)) \\
&\le K(1+\varepsilon)(|||v|||_{\mathcal{L}G} + \varepsilon), \qquad \nu\text{-a.e. } g. \qquad \square
\end{aligned}$$

REMARK 4.17. As is seen from the proof above, assertion (ii) is strengthened to the following when we can take $H_0' = H_0$: if $\|\nabla^{\mathcal{L}} f\| \le K$, ν-a.e., then, for every $v \in \mathcal{F}\mathcal{L}G$, it holds that $\nu \circ l_v^{-1} \ll \nu$ and $|f(vg) - f(g)| \le K\|v\|_{\mathcal{L}G}$, ν-a.e. g.

PROOF OF PROPOSITION 4.4. Let $N > 0$. For each $v \in \mathcal{F}\mathcal{L}G$, let

$$\rho_v(g) = \begin{cases} \mathsf{d}_A^{\mathcal{L}}(g) \wedge N, & \text{if } vg \in A, \\ \infty, & \text{otherwise,} \end{cases} \qquad g \in \mathcal{L}G,$$

and

$$\bar{\rho}_v(g) = \begin{cases} |||v|||_{\mathcal{L}G}, & \text{if } vg \in A, \\ \infty, & \text{otherwise}, \end{cases} \qquad g \in \mathcal{L}G.$$

Both are measurable functions. By Lemma 4.16(ii), each $v \in \mathcal{F}\mathcal{L}G$ has a sequence $\{v_n\}$ in $\mathcal{F}\mathcal{L}G$ such that $v_n \to v$ in $\mathcal{F}\mathcal{L}G$, $\nu \circ l_{v_n}^{-1} \ll \nu$ and

$$\text{(4.9)} \quad \left|\mathsf{d}_A^{\mathcal{L}}(v_n g) \wedge N - \mathsf{d}_A^{\mathcal{L}}(g) \wedge N\right| \le (1+n^{-1})(|||v|||_{\mathcal{L}G} + n^{-1}), \qquad \nu\text{-a.e. } g.$$

Take a ν-version of $\mathsf{d}_A^{\mathcal{L}}$ so that $\mathsf{d}_A^{\mathcal{L}} = 0$ on A. Since A is $d^{\mathcal{L}}$-open, $\mathsf{d}_A^{\mathcal{L}}(v_n g) = 0$ for large enough n if $vg \in A$. Therefore, letting $n \to \infty$ in (4.9), we have $\rho_v(g) \le \bar{\rho}_v(g)$ ν-a.e. Take a countable dense set $\mathcal{L}_0$ of $(\mathcal{F}\mathcal{L}G, d^{\mathcal{L}})$. From the $d^{\mathcal{L}}$-openness of A again, we have

$$\bar{\mathsf{d}}_A^{\mathcal{L}}(g) = \inf_{v\in\mathcal{L}_0} \bar{\rho}_v(g) \ge \inf_{v\in\mathcal{L}_0} \rho_v(g) \ge \mathsf{d}_A^{\mathcal{L}}(g) \wedge N, \qquad \nu\text{-a.e.}$$

Letting $N \to \infty$, we get $\bar{\mathsf{d}}_A^{\mathcal{L}}(g) \ge \mathsf{d}_A^{\mathcal{L}}(g)$, ν-a.e. $\square$

4.4. *Proof of Proposition* 4.5. From the property of $\mathsf{d}_A^{\mathcal{L}}$, it is enough to prove that, for each $N > 0$, $\bar{\mathsf{d}}_A^{\mathcal{L}} \wedge N$ belongs to $\mathcal{F}^{\mathcal{L}}$ and $\|\nabla^{\mathcal{L}}(\bar{\mathsf{d}}_A^{\mathcal{L}} \wedge N)\| \le 1$, ν-a.e.

We note that $\bar{\mathsf{d}}_A^{\mathcal{L}}$ is $d^{\mathcal{L}}$-Lipschitz on $\mathcal{L}G$ with $d^{\mathcal{L}}$-Lipschitz constant (at most) 1; therefore, so is $\bar{\mathsf{d}}_A^{\mathcal{L}} \wedge N$. This is easily proven from the definition of $\bar{\mathsf{d}}_A^{\mathcal{L}}$. The rest to be proven is the following theorem of Rademacher type.

THEOREM 4.18. *Let f be a bounded measurable function on $\mathcal{L}G$ and $d^{\mathcal{L}}$-Lipschitz continuous. Then $f \in \mathcal{F}^{\mathcal{L}}$ and $\|\nabla^{\mathcal{L}} f\|$ is a.e. dominated by the $d^{\mathcal{L}}$-Lipschitz constant of f.*

By virtue of Lemma 4.16, we need only to prove that $f \in \mathcal{F}^{\mathcal{L}}$. Let $|f| \le M$ everywhere. For the proof, we will extend f to a neighborhood of $\mathcal{L}G$ in $\mathcal{P}G$, following an idea by Gross [22]. Since $|\mathrm{Log}((\exp b_1)(\exp(b_2 - b_1)))|_{\mathfrak{g}} \ge \text{const} \times |b_2|_{\mathfrak{g}}$ for b_1 and b_2 sufficiently near to 0 in $\mathfrak{g}$, we can take $\varepsilon' > 0$ such that $V_0 := \{\exp a \mid |a|_{\mathfrak{g}} < \varepsilon'\}$ satisfies $V_0 V_0^{-1} \subset U$ and

$$\sup_{|b|_{\mathfrak{g}}<\varepsilon',\, |b'|_{\mathfrak{g}}<\varepsilon',\, b\neq b'} \frac{|b-b'|_{\mathfrak{g}}}{|\mathrm{Log}((\exp b)(\exp(-b')))|_{\mathfrak{g}}} < \infty.$$

For each $x \in V_0$, set $w_x \in \mathcal{P}G$ by $w_x(t) = \exp((t/T)\,\mathrm{Log}\, x)$, $t \in [0,T]$. Define a function f_1 on $\mathcal{P}_{V_0}G$ by $f_1(g) = f(w_{g(T)}^{-1} g)$. Since the map $g \mapsto w_{g(T)}^{-1} g$ is continuous from $(\mathcal{P}_{V_0}G, \rho)$ to $(\mathcal{L}G, \rho)$, f_1 is a universally measurable function.

LEMMA 4.19. *The function f_1 is $d^{\mathcal{P}}$-Lipschitz on $\mathcal{P}_{V_0}G$.*

PROOF. Suppose that $g_1, g_2 \in \mathcal{P}_{V_0}G$ satisfy $d^{\mathcal{P}}(g_1, g_2) \le \varepsilon_0$. In the following, c_i denotes a constant independent of g_1 and g_2. Set $v = g_1 g_2^{-1} \in \mathcal{P}G$, $\xi_i = w_{g_i(T)}^{-1} \in \mathcal{P}G$, $a_i = T^{-1}\operatorname{Log} g_i(T) \in \mathfrak{g}$, $i = 1, 2$. Then, utilizing the calculation in the proof of Lemma 4.13, we have $|||\xi_1|||_{\mathcal{P}G} = \sqrt{T}|a_1|_{\mathfrak{g}}$, $d^{\mathcal{P}}(\xi_1, \xi_2) = \sqrt{T}|a_1 - a_2|_{\mathfrak{g}}$ and

$$
\begin{aligned}
d^{\mathcal{P}}(\xi_1 g_1, \xi_2 g_2) \le{}& \sup_{t\in[0,T]} \|\operatorname{Ad} v(t)^{-1} - \mathrm{Id}_{\mathfrak{g}}\|_{\mathfrak{g}\to\mathfrak{g}} \sqrt{T}|a_1|_{\mathfrak{g}} \\
&+ d^{\mathcal{P}}(g_1, g_2) + \sqrt{T}|a_1 - a_2|_{\mathfrak{g}}.
\end{aligned}
$$

Since Ad is differentiable, the first term is dominated by $c_1 \sup_{t\in[0,T]} |\operatorname{Log} v(t)|_{\mathfrak{g}} \times |a_1|_{\mathfrak{g}}$. By Lemmas 4.1 and 4.8, this is further dominated by $c_2 d^{\mathcal{P}}(g_1, g_2)$. We also have

$$
\begin{aligned}
|a_1 - a_2|_{\mathfrak{g}} &\le c_3 \big|\operatorname{Log}\big(\exp(Ta_1)\exp(-Ta_2)\big)\big|_{\mathfrak{g}} = c_3 \big|\operatorname{Log}\big(g_1(T) g_2(T)^{-1}\big)\big|_{\mathfrak{g}} \\
&\le c_4 \|\operatorname{Log}(g_1 g_2^{-1})\| = c_4 \|\chi \circ I_0^{-1}(g_1 g_2^{-1})\| \le c_5 d^{\mathcal{P}}(g_1, g_2).
\end{aligned}
$$

Therefore, $d^{\mathcal{P}}(\xi_1 g_1, \xi_2 g_2) \le c_6 d^{\mathcal{P}}(g_1, g_2)$ for some $c_6 \ge 1$.

Now, when $d^{\mathcal{P}}(g_1, g_2) \le \varepsilon_0/c_6$, we have, by Lemma 4.10,

$$
\begin{aligned}
|f_1(g_1) - f_1(g_2)| &= |f(\xi_1 g_1) - f(\xi_2 g_2)| \le d^{\mathcal{L}}(\xi_1 g_1, \xi_2 g_2) \\
&\le C_0 d^{\mathcal{P}}(\xi_1 g_1, \xi_2 g_2) \le C_0 c_6 d^{\mathcal{P}}(g_1, g_2).
\end{aligned}
$$

When $d^{\mathcal{P}}(g_1, g_2) > \varepsilon_0/c_6$, we have $|f_1(g_1) - f_1(g_2)| \le 2M \le (2Mc_6/\varepsilon_0) \times d^{\mathcal{P}}(g_1, g_2)$. $\square$

Denote by C_2 the $d^{\mathcal{P}}$-Lipschitz constant of $f_1|_{\mathcal{P}_{V_0}G}$. Take neighborhoods V_1 and V_2 of e in G such that $\overline{V_2} \subset V_1$ and $\overline{V_1} \subset V_0$. Fix a C^∞-function Ψ on G so that $0 \le \Psi \le 1$ on G, $\Psi = 1$ on V_2 and $\Psi = 0$ on $G \setminus V_1$. A function F on $\mathcal{P}G$ will be defined by

$$
F(g) = \begin{cases} f_1(g)\Psi(g(T)), & \text{if } g \in \mathcal{P}_{V_0}G, \\ 0, & \text{otherwise.} \end{cases}
$$

Clearly, F is universally measurable. From Lemmas 4.1 and 4.8, $d := \inf\{d^{\mathcal{P}}(g, g') \mid g \in \mathcal{P}_{V_1}G,\ g' \in \mathcal{P}_{G\setminus V_0}G\} > 0$ and the function $\mathcal{P}G \ni g \mapsto \Psi(g(T)) \in \mathbb{R}$ is $d^{\mathcal{P}}$-Lipschitz continuous. Let C_3 be its $d^{\mathcal{P}}$-Lipschitz constant.

LEMMA 4.20. *$F \in \mathcal{F}^{\mathcal{P}}$ and $\|\nabla^{\mathcal{P}} F\|$ is μ-essentially bounded.*

PROOF. We first prove that F is $d^{\mathcal{P}}$-Lipschitz continuous. When $g, g' \in \mathcal{P}_{V_0}G$,

$$
\begin{aligned}
|F(g) - F(g')| &\le \big|\big(f_1(g) - f_1(g')\big)\Psi(g(T))\big| + \big|f_1(g')\big(\Psi(g(T)) - \Psi(g'(T))\big)\big| \\
&\le |f_1(g) - f_1(g')| + M\big|\Psi(g(T)) - \Psi(g'(T))\big| \\
&\le (C_2 + MC_3) d^{\mathcal{P}}(g, g').
\end{aligned}
$$

When $g \in \mathcal{P}_{V_1}G$ and $g' \in \mathcal{P}_{G\setminus V_0}G$,

$$|F(g) - F(g')| = |F(g)| \le M \le Md^{-1}d^{\mathcal{P}}(g, g').$$

When $g, g' \in \mathcal{P}_{G\setminus V_1}G$, $|F(g) - F(g')| = 0$.

Now the assertions of the lemma are proven in the same way as in [4], Lemma 3.1. We will give it for completeness. Let C_4 be the $d^{\mathcal{P}}$-Lipschitz constant of F.

For each $h \in H$, it holds that

$$I(w+h) = h + I(w), \qquad \lambda\text{-a.e. } w.$$

Indeed, by Itô's formula, $v(t) := (I(w+h)I(w)^{-1})(t)$ satisfies $v(0) = 0$ and $v^{-1}\dot{v} = (\operatorname{Ad} I(w))\dot{h}$. Then

$$\big|(F \circ I)(w+h) - (F \circ I)(w)\big| = \big|F\big(h + I(w)\big) - F(I(w))\big| \le C_4\|h\|, \qquad \lambda\text{-a.e.}$$

From [5], Theorem 2.4, or [16], Theorem, $F \circ I \in \mathbb{D}^{1,2}(\mathcal{P}\mathfrak{g})$ and $\|D(F \circ I)\| \le C_4$, λ-a.e. This implies that $F \in \mathcal{F}^{\mathcal{P}}$ and $\|\nabla^{\mathcal{P}} F\| \le C_4$, μ-a.e. $\square$

We note that F is everywhere defined, not only almost everywhere.

Let $T_t^{\mathcal{P}}$ and $\mathcal{L}^{\mathcal{P}}$ be the semigroup and its generator associated with $(\mathcal{E}^{\mathcal{P}}, \mathcal{F}^{\mathcal{P}})$. As in the case of Wiener spaces, for $t > 0$, both $T_t^{\mathcal{P}} F$ and $\|\nabla^{\mathcal{P}} T_t^{\mathcal{P}} F\|^2$ belong to $\mathbb{D}^{\infty}(\mathcal{P}G)$, and $\|\nabla^{\mathcal{P}} T_t^{\mathcal{P}} F\| \le e^{-t/2} T_t^{\mathcal{P}}(\|\nabla^{\mathcal{P}} F\|) \le C_4$, μ-a.e. Then $\|\nabla^{\mathcal{P}} T_t^{\mathcal{P}} F\|$ has an ∞-quasicontinuous modification and it is dominated by C_4, ∞-q.e. From now on, we always take such a modification when we can. We also fix an ∞-quasicontinuous modification $\widetilde{T_t^{\mathcal{P}} F}$ for each $t > 0$. Take a subsequence $\{t_n\}$ decreasing to 0 so that $\widetilde{T_{t_n}^{\mathcal{P}} F} \to F$, μ-a.e. Let $F_n = \widetilde{T_{t_n}^{\mathcal{P}} F}$. Then, by taking account of the disintegration (4.6), there is an m-null set G_0 of G such that $F_n \to F$, μ_x-a.e. for every $x \in G \setminus G_0$, where m is the Haar measure on G. Take and fix $x \in V_2 \setminus G_0$. Let $a = \operatorname{Log} x$. For a function ϕ on $\mathcal{P}G$, define $\phi^x(g) = \phi(w_x g)$.

LEMMA 4.21. *There exists a constant $C_5 \ge 1$ such that, for every $\phi \in \mathcal{C}$,*

$$\|\nabla^{\mathcal{P}} \phi^x(g)\| \le C_5 \|\nabla^{\mathcal{P}} \phi(w_x g)\|, \qquad g \in \mathcal{P}G.$$

PROOF. Clearly, $\phi^x \in \mathcal{C}$ if $\phi \in \mathcal{C}$. Let $h \in H$ and $\varepsilon > 0$ small. Then $w_x e^{\varepsilon h} = e^{v} w_x$, where $v = \operatorname{Log}(w_x e^{\varepsilon h} w_x^{-1}) = (\operatorname{Ad} w_x)(\varepsilon h) = \varepsilon(\operatorname{Ad} w_x)(h)$. Since $(\operatorname{Ad} w_x)(h)(t) = e^{\operatorname{ad}(ta/T)} h(t) = e^{(t/T)\operatorname{ad} a} h(t)$,

$$\frac{d}{dt}(\operatorname{Ad} w_x)(h)(t) = \big(\operatorname{Ad} w_x(t)\big)\left\{\frac{(\operatorname{ad} a)h(t)}{T} + \dot{h}(t)\right\}.$$

Hence, there exists some constant C_5 independent of h such that $\|(\operatorname{Ad} w_x)h\| \le C_5\|h\|$. Since

$$\begin{aligned}
\varepsilon^{-1}\big(\phi^x(e^{\varepsilon h} g) - \phi^x(g)\big) &= \varepsilon^{-1}\big(\phi(w_x e^{\varepsilon h} g) - \phi(w_x g)\big) \\
&= \varepsilon^{-1}\big(\phi(e^{\varepsilon(\operatorname{Ad} w_x)h} w_x g) - \phi(w_x g)\big) \\
&\to \big(\nabla^{\mathcal{P}} \phi(w_x g) | (\operatorname{Ad} w_x)h\big) \qquad \text{as } \varepsilon \to 0,
\end{aligned}$$

we get the conclusion. $\square$

For each n, take a sequence of functions $\{F_{n,l}\}_{l\in\mathbb{N}}$ in $\mathcal{C}$ converging to F_n in $\mathbb{D}^\infty(\mathcal{P}G)$. Since $\|\nabla^{\mathcal{P}}(F_{n,l}-F_n)\|^2 \to 0$ in $\mathbb{D}^\infty(\mathcal{P}G)$, we have

$$\int (\|\nabla^{\mathcal{P}} F_{n,l} - \nabla^{\mathcal{P}} F_{n,m}\|^2 + |F_{n,l}-F_{n,m}|^2)^p \, d\mu_x \to 0 \qquad \text{as } l \ge m \to \infty$$

for all $p>1$. Take $p'>1$ so that $d\nu/d\mu_e \in L^{p'}(\mu_e)$. Let $q' = p'/(p'-1)$. Then

$$\begin{aligned}
&\int (\|\nabla^{\mathcal{L}} F^x_{n,l} - \nabla^{\mathcal{L}} F^x_{n,m}\|^2 + |F^x_{n,l} - F^x_{n,m}|^2)\, d\nu \\
&\quad \le \int (\|\nabla^{\mathcal{P}} F^x_{n,l} - \nabla^{\mathcal{P}} F^x_{n,m}\|^2 + |F^x_{n,l} - F^x_{n,m}|^2)\, d\nu \\
&\quad \le \int \Big(C_5^2 \|\nabla^{\mathcal{P}} F_{n,l}(w_x g) - \nabla^{\mathcal{P}} F_{n,m}(w_x g)\|^2 \\
&\qquad\qquad\qquad + |F_{n,l}(w_x g) - F_{n,m}(w_x g)|^2 \Big) \frac{d\nu}{d\mu_e}(g)\, \mu_e(dg) \\
&\quad \le C_5^2 \bigg\{ \int \Big(\|(\nabla^{\mathcal{P}} F_{n,l} - \nabla^{\mathcal{P}} F_{n,m})(w_x g)\| \\
&\qquad\qquad\qquad + |(F_{n,l} - F_{n,m})(w_x g)| \Big)^{2q'} \mu_e(dg) \bigg\}^{1/q'} \\
&\qquad \times \left\| \frac{d\nu}{d\mu_e} \right\|_{L^{p'}(\mu_e)} \\
&\quad \le C_5^2 \bigg\{ \int (\|\nabla^{\mathcal{P}} F_{n,l} - \nabla^{\mathcal{P}} F_{n,m}\| + |F_{n,l} - F_{n,m}|)^{2q'} \, d\mu_e \circ l_{w_x}^{-1} \bigg\}^{1/q'} \\
&\qquad \times \left\| \frac{d\nu}{d\mu_e} \right\|_{L^{p'}(\mu_e)} \\
&\quad \le C_5^2 \bigg\{ \int (\|\nabla^{\mathcal{P}} F_{n,l} - \nabla^{\mathcal{P}} F_{n,m}\| + |F_{n,l} - F_{n,m}|)^{4q'} \, d\mu_x \bigg\}^{1/2q'} \\
&\qquad \times \|J_{w_x}\|_{L^2(\mu_x)}^{1/q'} \left\| \frac{d\nu}{d\mu_e} \right\|_{L^{p'}(\mu_e)} \\
&\quad \to 0 \qquad \text{as } l \ge m \to \infty.
\end{aligned}$$

Therefore, $\{F^x_{n,l}\}_{l\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{F}^{\mathcal{L}}$ and the limit should be F^x_n. By taking a suitable subsequence (and ∞-quasicontinuous modification), we may assume that $\nabla^{\mathcal{L}} F^x_{n,l} \to \nabla^{\mathcal{L}} F^x_n$, ν-a.e. and $\|\nabla^{\mathcal{P}} F_{n,l}\| \to \|\nabla^{\mathcal{P}} F_n\|$, ∞-q.e., in particular, $\mu_e \circ l_{w_x}^{-1}$-a.e. Then

$$\begin{aligned}\|\nabla^{\mathcal{L}} F_n^x(g)\| &= \lim_{l\to\infty} \|\nabla^{\mathcal{L}} F_{n,l}^x(g)\| \le \liminf_{l\to\infty} \|\nabla^{\mathcal{P}} F_{n,l}^x(g)\| \\ &\le C_5 \liminf_{l\to\infty} \|\nabla^{\mathcal{P}} F_{n,l}(w_x g)\| \\ &= C_5 \|\nabla^{\mathcal{P}} F_n(w_x g)\| \le C_4 C_5, \qquad \nu\text{-a.e.}\end{aligned}$$

In particular, $\{F_n^x\}_{n\in\mathbb{N}}$ is bounded in $\mathcal{F}^{\mathcal{L}}$. Noticing that $F_n^x \to F^x$ $\mu_x \circ l_{w_x}^{-1}$-a.e., hence μ_e-a.e., that is, ν-a.e., we have $F^x \in \mathcal{F}^{\mathcal{L}}$ from the Banach–Saks theorem. But, since $F^x = f$ on $\mathcal{L}G$ from the way of constructing of F^x, we conclude $f \in \mathcal{F}^{\mathcal{L}}$. This finishes the proof of Theorem 4.18 and hence Proposition 4.5.

REMARK 4.22. (i) We will temporarily write $\mathcal{L}G^{(T)}$ for the loop group to emphasize the parameter T. For each $T > 0$, we can consider the heat kernel measure $\mu_{\mathrm{heat},T}$, which is the law at time T of the Brownian motion on $\mathcal{L}G^{(1)}$. The induced measure of $\mu_{\mathrm{heat},T}$ by the map

$$\mathcal{L}G^{(1)} \ni g \mapsto \big(t \mapsto g(t/T)\big) \in \mathcal{L}G^{(T)}$$

is proven to be absolutely continuous with respect to the pinned Wiener measure on $\mathcal{L}G^{(T)}$ by Driver and Srimurthy [14]. Since conditions (M0) and (M1) hold for such measures (see the same reference in the proof of Proposition 4.2), we can show that the claim in Theorem 1.5 is true also for $\mu_{\mathrm{heat},T}$.

(ii) Fang and Zhang [19] gave a large deviation estimate for the Brownian motion on $\mathcal{L}G$ with μ_{heat}. We should note that the rate function there is highly relevant to the $d^{\mathcal{P}}$-length defined by (4.3).

(iii) Aida [1] proved the essential self-adjointness of the generator of $(\mathcal{E}^{\mathcal{L}}, \mathcal{F}^{\mathcal{L}})$ when ν is the pinned Wiener measure. Utilizing this fact might simplify the proof of Theorem 1.5.

Acknowledgments. The main part of this paper was done while the first author visited Cornell University from September 2000 to March 2001. He thanks Professor L. Gross for his hospitality and fruitful discussion. He also thanks Professor S. Watanabe for his valuable comments. The second author would also like to thank Professor L. Saloff-Coste for very helpful discussion that led to improvement of the results in this paper. He also owes an apology to W. Stannat for not acknowledging in [36] that his comments were important in the development of the method that is shared by that and this paper.

REFERENCES

[1] AIDA, S. (1995). Sobolev spaces over loop groups. *J. Funct. Anal.* **127** 155–172.

[2] AIDA, S. and DRIVER, B. K. (2000). Equivalence of heat kernel measure and pinned Wiener measure on loop groups. *C. R. Acad. Sci. Paris Sér. I Math.* **331** 709–712.

[3] AIDA, S. and KAWABI, H. (1998). Short time asymptotics of certain infinite dimensional diffusion process. In *Proceedings of the Seventh Workshop on Stochastic Analysis and Related Topics* 77–124. Birkhäuser, Boston.

[4] AIDA, S. and ZHANG, T. S. (2002). On the small time asymptotics of diffusion processes on path groups. *Potential Anal.* **16** 67–78.

[5] AIDA, S. MASUDA, T. and SHIGEKAWA, I. (1994). Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* **126** 83–101.

[6] BARLOW, M. T. (1998). Diffusions on fractals. *Lectures on Probability Theory and Statistics*. *Lecture Notes in Math.* **1690** 1–121.

[7] BOULEAU, N. and HIRSH, F. (1991). *Dirichlet Forms and Analysis on Wiener Space*. de Gruyter, Berlin.

[8] BOURBAKI, N. (1989). *Elements of Mathematics*: *General Topology* Chapters 5–10. Springer, Berlin.

[9] BREZIS, H. (1983). *Analyse fonctionelle*: *Theorie et applications*. Masson, Paris.

[10] DAVIES, E. B. (1987). Explicit constants for Gaussian upper bounds on heat kernels. *Amer. J. Math.* **109** 319–334.

[11] DAVIES, E. B. (1989). *Heat Kernels and Spectral Theory*. Cambridge Univ. Press.

[12] DRIVER, B. K. (1994). A Cameron–Martin type quasi-invariance theorem for pinned Brownian motion on a compact Riemannian manifold. *Trans. Amer. Math. Soc.* **342** 375–395.

[13] DRIVER, B. K. (1997). Integration by parts and quasi-invariance for heat kernel measures on loop groups. *J. Funct. Anal.* **149** 470–547. [Correction (1998) *J. Funct. Anal.* **155** 297–301.]

[14] DRIVER, B. K. and SRIMURTHY, V. K. (2001). Absolute continuity of heat kernel measure with pinned Wiener measure on loop groups. *Ann. Prob.* **29** 691–723.

[15] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators I.* Interscience, New York.

[16] ENCHEV, O. and STROOCK, D. W. (1993). Rademacher's theorem for Wiener functionals. *Ann. Probab.* **21** 25–33.

[17] FANG, S. (1994). On the Ornstein–Uhlenbeck process. *Stochastics Stochastics Rep.* **46** 141–159.

[18] FANG, S. and ZHANG, T. S. (1999). On the small time behavior of Ornstein–Uhlenbeck processes with unbounded linear drift. *Probab. Theory Related Fields* **114** 487–504.

[19] FANG, S. and ZHANG, T. S. (2001). Large deviations for the Brownian motion on loop groups. *J. Theoret. Probab.* **14** 463–483.

[20] FUKUSHIMA, M., OSHIMA, Y. and TAKEDA, M. (1994). *Dirichlet Forms and Symmetric Markov Processes*. de Gruyter, Berlin.

[21] GAFFNEY, M. P. (1959). The conservation property of the heat equation on Riemannian manifolds. *Comm. Pure Appl. Math.* **12** 1–11.

[22] GROSS, L. (1991). Logarithmic Sobolev inequalities on loop groups. *J. Funct. Anal.* **102** 268–313.

[23] GROSS, L. (1992). Logarithmic Sobolev inequalities on Lie groups. *Illinois J. Math.* **36** 447–490.

[24] GROSS, L. (1993). Uniqueness of ground states for Schrödinger operators over loop groups. *J. Funct. Anal.* **112** 373–441.

[25] HINO, M. (2002). On short time asymptotic behavior of some symmetric diffusions on general state spaces. *Potential Anal.* **16** 249–264.

[26] HSU, E. P. (1999). Analysis on path and loop spaces. In *Probability Theory and Applications* (E. P. Hsu and S. R. S. Varadhan, eds.) 277–347. Amer. Math. Soc., Providence, RI.

[27] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. North-Holland, Amsterdam.

[28] KUSUOKA, S. (1987). A diffusion process on a fractal. In *Probabilistic Methods in Mathematical Physics. Proceedings of the Taniguchi International Symposium* (K. Itô and N. Ikeda, eds.) 251–274. Kinokuniya, Tokyo.

[29] KUSUOKA, S. (1989). Dirichlet forms on fractals and products of random matrices. *Publ. Res. Inst. Math. Sci.* **25** 659–680.

[30] MA, Z. M. and RÖCKNER, M. (1992). *Introduction to the Theory of* (*Non-Symmetric*) *Dirichlet Forms*. Springer, Berlin.

[31] MALLIAVIN, M.-P. and MALLIAVIN, P. (1990). Integration on loop groups. I. Quasi invariant measures. *J. Funct. Anal.* **93** 207–237.

[32] MALLIAVIN, P. (1997). *Stochastic Analysis*. Springer, Berlin.

[33] NORRIS, J. R. (1997). Heat kernel asymptotics and the distance function in Lipschitz Riemannian manifolds. *Acta Math.* **179** 79–103.

[34] NORRIS, J. R. and STROOCK, D. (1991). Estimates on the fundamental solution to heat flows with uniformly elliptic coefficients. *Proc. London Math. Soc. (3)* **62** 373–402.

[35] NUALART, D. (1995). *The Malliavin Calculus and Related Topics.* Springer, Berlin.

[36] RAMÍREZ, J. A. (2001). Short time asymptotics in Dirichlet spaces. *Comm. Pure Appl. Math.* **54** 259–293.

[37] RUDIN, W. (1991). *Functional Analysis*. McGraw-Hill, New York.

[38] SADASUE, G. (1995). Equivalence–singularity dichotomy for the Wiener measures on path groups and loop groups. *J. Math. Kyoto Univ.* **35** 653–662.

[39] SHIGEKAWA, I. (1994). Sobolev spaces of Banach-valued functions associated with a Markov process. *Probab. Theory Related Fields* **99** 425–441.

[40] SHIGEKAWA, I. (1995). A quasihomeomorphism on the Wiener space. In *Stochastic Analysis* (M. C. Cranston and M. A. Pinsk, eds.) 473–486. Amer. Math. Soc., Providence, RI.

[41] SIMON, B. (1973). Ergodic semigroups of positivity preserving self-adjoint operators. *J. Funct. Anal.* **12** 335–339.

[42] SRIMURTHY, V. K. (2000). On the equivalence of measures on loop space. *Probab. Theory Related Fields* **118** 522–546.

[43] STROOCK, D. W. (1999). An introduction to analysis on path space. In *Probability Theory and Applications* (E. P. Hsu and S. R. S. Varadhan, eds.) 227–276. Amer. Math. Soc., Providence, RI.

[44] VARADHAN, S. R. S. (1967). On the behaviour of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.* **20** 431–455.

[45] ZHANG, T. S. (2000). On the small time asymptotics of diffusion processes on Hilbert spaces. *Ann. Probab.* **28** 537–557.

GRADUATE SCHOOL OF INFORMATICS
KYOTO UNIVERSITY
KYOTO 606-8501
JAPAN
E-MAIL: hino@i.kyoto-u.ac.jp

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853
E-MAIL: ramirez@polygon.math.cornell.edu