Potential Analysis (2024) 60:79–137
https://doi.org/10.1007/s11118-022-10039-4

Time Regularity for Local Weak Solutions of the Heat Equation on Local Dirichlet Spaces

Qi Hou[1] · Laurent Saloff-Coste[2]

Received: 12 October 2020 / Accepted: 10 August 2022 / Published online: 1 October 2022

Abstract

We study the time regularity of local weak solutions of the heat equation in the context of local regular symmetric Dirichlet spaces. Under two basic and rather minimal assumptions, namely, the existence of certain cut-off functions and a very weak L^2 Gaussian type upper bound for the heat semigroup, we prove that the time derivatives of a local weak solution of the heat equation are themselves local weak solutions. This applies, for instance, to local weak solutions of parabolic equations with uniformly elliptic symmetric divergence form second order operators with measurable coefficients. We describe some applications to the structure of ancient local weak solutions of such equations which generalize recent results of Colding and Minicozzi (Duke Math. J., **170**(18), 4171–4182 2021) and Zhang (Proc. Amer. Math. Soc., **148**(4), 1665–1670 2020).

Keywords Dirichlet form · Heat equation weak solution

Mathematics Subject Classification (2010) 31C25 · 60J46 · 58J35

1 Introduction

When $-P$ is the infinitesimal generator of a self-adjoint strongly continuous semigroup of operators $H_t = e^{-tP}$ acting on a Hilbert space $\mathbf{H}$, spectral theory implies the time regularity of any (global) solution $u(t) = H_t u_0$ of the equation $(\partial_t + P)u = 0$ with initial data $u_0 \in \mathbf{H}$. When $\mathbf{H} = L^2(X, m)$ for some nice measure space (X, m) and $-P$ is associated with a bilinear form $\mathcal{E}$ so that $\mathcal{E}(f, g) = \int_X f P g \, dm$ for enough functions f, g, it is often very useful to consider the concept of local weak solution of the equation $(\partial_t + P)u = 0$ in some

The authors were partially supported by NSF grant DMS 1404435 and DMS 1707589

✉ Qi Hou
qh55@cornell.edu; houqi@bimsa.cn

Laurent Saloff-Coste
lsc@math.cornell.edu

1 Yanqi Lake Beijing Institute of Mathematical Sciences and Applications, Huairou, China

2 Cornell University, Ithaca, NY, USA

open time-space cylinder $I \times \Omega \subset \mathbb{R} \times X$, in some appropriate sense. Such definition goes roughly as follows. A local weak solution u is a function defined on $I \times \Omega$ which *must* belong (locally) to a certain function space $\mathcal{F}$ (in the most classical case, $\mathcal{F}$ is related to the Sobolev space) and satisfies

$$-\int_{I\times\Omega} u\partial_t\phi\, dtdm + \int_I \mathcal{E}(u,\phi)dt = 0 \tag{1.1}$$

for all "test functions" ϕ compactly supported in $I \times \Omega$. The precise nature of the space $\mathcal{F}$ and of the space of test functions to be used here are an important part of such definition. When dealing with such a definition, the time regularity of a local weak solution is not automatic. Formally, one expects the time derivative of a local weak solution to be a local weak solution of (1.1), but the problem lies with the a priori requirement that $v = \partial_t u$ belongs locally to the space $\mathcal{F}$.

Consider the classical case when P is a symmetric locally uniformly elliptic second order operator with measurable coefficients $(a_{ij}(x))_{i,j=1}^n$ so that for any $f, g \in C_c^\infty(\mathbb{R}^n)$,

$$\mathcal{E}(f,g) = \int_{\mathbb{R}^n} \sum_{i,j=1}^n a_{ij}(x)\partial_i f(x)\partial_j g(x)\, dx.$$

The basic assumption, local uniform ellipticity, means that for any compact subset K there are $\epsilon_K > 0$ and $C_K < \infty$ such that

$$\max_{i,j} \sup_K\{|a_{ij}|\} \le C_K \text{ and } \sum_{i,j=1}^n a_{ij}\xi_i\xi_j \ge \epsilon_K\|\xi\|_2^2,\ \forall\xi = (\xi_i)_{i=1}^n.$$

For any open subset $\Omega \subset \mathbb{R}^n$ and open interval (a,b), $-\infty \le a < b \le \infty$, a local weak solution of $(\partial_t + P)u = 0$ in $(a,b)\times\Omega$ is a function u that locally belongs to $L^2((a,b) \to W^{1,2}(\Omega))$, such that

$$-\int_a^b\int_\Omega u(t,x)\partial_t\phi(t,x)dxdt + \int_a^b\int_\Omega \sum_{i,j=1}^n a_{ij}(x)\partial_i u(t,x)\partial_j\phi(t,x)\, dxdt = 0$$

for all functions $\phi \in C^\infty((a,b)\times\Omega)$ with compact support in $(a,b)\times\Omega$. Here, locally in $L^2((a,b) \to W^{1,2}(\Omega))$ means that u multiplied with any smooth function with compact support in $(a,b)\times\Omega$, is in $L^2((a,b) \to W^{1,2}(\Omega))$. It is clear from this definition that locally in $L^2((a,b) \to W^{1,2}(\Omega))$ and locally in $L^2((a,b) \to W_0^{1,2}(\Omega))$ are equivalent.

One consequence of the general results proved in this paper is that the iterated time derivatives $v_k(t,x) = \partial_t^k u(t,x)$ of any local weak solution u of the equation above are themselves locally in $L^2((a,b) \to W^{1,2}(\Omega))$ and are local weak solutions of the same equation in $(a,b)\times\Omega$. This follows from the following more general theorem. In this statement we assume that (X,m) is a locally compact separable metric measure space where m is a positive Radon measure with full support. In the following theorem, local weak solutions are in the sense of Definition 1.

Theorem 1 *Assume $(\mathcal{E},\mathcal{F})$ is a symmetric strongly local regular Dirichlet form on $L^2(X,m)$ whose intrinsic pseudo-metric is a continuous metric which induces the topology of X. For any local weak solution u of the associated heat equation in $(a,b)\times\Omega$, where $-\infty \le a < b \le \infty$ and $\Omega \subset X$ is an open set, the iterated time derivatives $v_k = \partial_t^k u$ are themselves local weak solutions of the same heat equation in $(a,b)\times\Omega$.*

Theorem 1 is an immediate consequence of our main results, Theorem 5 and Corollary 1. Theorem 5 is more general in several ways. First, it treats local Dirichlet forms, not just strongly local ones. Second, it replaces the existence of a continuous intrinsic pseudo-metric with weaker assumptions we now explain. One weakness of Theorem 1 is that it excludes fractal sets such as the Sierpinski gasket and the Sierpinski carpet (on such examples, the intrinsic pseudo-distance is identically equal to 0) as well as some infinite dimensional examples (e.g., on the infinite dimensional torus there are cases where the intrinsic pseudo-distance is infinite almost surely). These cases are in fact covered by Theorem 5 and Corollary 1. Indeed, Theorem 5 depends on the following two related types of assumptions which allow for spaces of the type just mentioned:

– the existence of good cut-off functions (in a sense that is somewhat weaker than most conditions of this type that exist in the literature);
– a very weak L^2 Gaussian bound, namely, the fact that for any $a > 0$ and any integer $k = 0, 1, 2, \dots$, for any disjoint compact sets V_1, V_2,

$$t^{-a} \sup_{\phi_1, \phi_2} \int_X \phi_2 \partial_t^k H_t \phi_1 dm \to 0 \ \ (\text{as } t \to 0)$$

where the sup is taken over all functions ϕ_1, ϕ_2 supported respectively in V_1, V_2 and with L^2-norm at most 1.

As an application of our results, we extend two recent structure theorems regarding ancient weak solutions, [9, 29, 38]. The first result of this type describes very general conditions under which any ancient (local) weak solution u of $(\partial_t + P)u = 0$ with "polynomial growth" must be of the form $u(t, x) = \sum_{k=1}^{d} t^k u_k(x)$ where all u_k's are of polynomial growth, u_d is a harmonic function, and other u_k's satisfy $-Pu_k = (k + 1)u_{k+1}$ in a weak sense. The integer d is related to the given growth degree of u. The second result describes very general conditions under which any ancient (local) weak solution of "exponential growth" is real analytic in time.

The general approach we take is to utilize the *heat semigroup* to study the time regularity properties of local weak solutions of the heat equation. The basic idea of deriving hypoelliticity type results from properties of the heat semigroup goes back to Kusuoka and Stroock's paper [24] which is written in the context of the heat equation associated with Hörmander sums of squares of vector fields in Euclidean spaces. It was also implemented in [7] to study distributional solutions of the Laplace equation on the infinite dimensional torus and other infinite dimensional compact groups.

This approach differs from the classical hypoellipticity viewpoint in the primary role it gives to the fundamental solution of the heat equation (here, in the very minimal form of the heat semigroup itself). On the contrary, traditional studies of hypoellipticity treat all solutions equally and are then used to deduce the basic regularity of the fundamental solution. In this paper we generalize the heat semigroup approach on hypoelliticity to the general setting of Dirichlet spaces on metric measure spaces. One natural goal is to cover rougher structures that make smoothness more elusive. Here, we treat a purely L^2-theory. In the companion paper [20], we further utilize this method to study the local boundedness and continuity properties of local weak solutions of the heat equation (the L^∞-type properties) under additional assumptions.

This work is organized as follows. Section 2 introduces the general Dirichlet space setup for this paper and defines the related notion of local weak solutions. Section 3 describes the two main hypotheses, the existence of certain cut-off functions and the notion of a very weak L^2 Gaussian bound. Section 4 states the main theorems proved in this paper, Theorem

5 and Corollary 1, and gives a sketch of the main idea of the proof while avoiding many long necessary computations and technical details. Section 5 gives a complete proof of Theorem 5 and Corollary 1. Section 6 is devoted to the results concerning the structure of ancient (local weak) solutions. Section 7 discusses briefly several typical examples that illustrate the results of this paper in a variety of different contexts. Lastly Section 8 verifies that the very weak L^2 Gaussian bound is satisfied under rather weak assumptions involving the existence of cut-off functions, and provides proofs of some lemmas regarding cut-off functions.

We remark that, in this paper, the Dirichlet forms we treat are symmetric, and are not time dependent. The independence on time is a crucial assumption for us, as we take advantage of the smoothness of the heat semigroup in time. The symmetry assumption can probably be replaced by some form of the sector condition but we leave this to a further study. For related but different results (under stronger assumptions) for nonsymmetric or time dependent Dirichlet spaces, we refer to [34, 35] and [26–28].

2 Dirichlet Spaces and Local Weak Solutions

2.1 Dirichlet Spaces

We briefly review some concepts and properties related to Dirichlet forms. A classical reference for (symmetric) Dirichlet forms is [16]. Let (X, d, m) be a metric measure space where X is locally compact separable, m is a Radon measure on X with full support, and d is some metric on X that we omit writing in the rest of the paper because we do not use it explicitly. For $p \in [1, \infty]$, we use $L^p(X, m)$ or $L^p(X)$ to denote the L^p-space on (X, m). Recall that the L^p-space is equipped with norm

$$\|f\|_{L^p(X)} = \left(\int_X |f|^p \, dm\right)^{1/p}$$

for $p \in [1, \infty)$, and

$$\|f\|_{L^\infty(X)} = \operatorname*{ess\,sup}_{x\in X} |f(x)|$$

where the essential supremum is with respect to the measure m. We use $\langle \cdot, \cdot \rangle_{L^2(X)}$ to denote the standard inner product on $L^2(X)$, i.e.,

$$\langle f, g\rangle_{L^2(X)} = \int_X fg \, dm.$$

Let $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form on $L^2(X, m)$, $\mathcal{F}$ denotes the domain of $\mathcal{E}$. By definition, a (symmetric) Dirichlet form is a closed symmetric form that further satisfies the Markov property. Here the term symmetric form refers to any symmetric, nonnegative definite, densely defined bilinear form. The domain $\mathcal{F}$ equipped with the $\mathcal{E}_1$ norm

$$\|f\|_{\mathcal{E}_1} := \left(\mathcal{E}(f, f) + \int_X f^2 \, dm\right)^{1/2}$$

is a Hilbert space.

Let $C_c(X)$ be the space of continuous functions in X with compact support. A Dirichlet form $(\mathcal{E}, \mathcal{F})$ is called regular, if $C_c(X)\cap\mathcal{F}$ is dense in $C_c(X)$ in the sup norm and dense in $\mathcal{F}$ in the $\mathcal{E}_1$ norm. Any subset $\mathcal{C} \subset C_c(X) \cap \mathcal{F}$ that is dense in these two senses is called a core of $\mathcal{E}$. Any u in the domain $\mathcal{F}$ of a regular Dirichlet form admits a quasi-continuous modification [16, Section 2.1]. In the following we do not specify quasi-continuous modifications of functions.

A Dirichlet form $(\mathcal{E}, \mathcal{F})$ is called local, if $\mathcal{E}(u, v) = 0$ for $u, v \in \mathcal{F}$ whenever $\operatorname{supp}\{u\}$ and $\operatorname{supp}\{v\}$ are disjoint and compact. Here $\operatorname{supp}\{f\}$ for any m-measurable function f denotes its (essential) support, i.e., the smallest closed subset F of X such that $f = 0$ m-a.e. outside F.

Regular Dirichlet forms satisfy the Beurling-Deny decomposition formula [16, Section 3.2]; as a corollary, any regular local Dirichlet form $(\mathcal{E}, \mathcal{F})$ can be written in the form

$$\mathcal{E}(u, v) = \int_X d\Gamma(u, v) + \int_X uv\, dk,\ \forall u, v \in \mathcal{F} \cap C_c(X).$$

This formula extends natually to all $u, v \in \mathcal{F}$ via quasi-continuous modification. Here dk is a positive Radon measure, called the killing measure, and $d\Gamma$ stands for the *energy measure*, which is a (Radon) measure-valued bilinear form first defined for any u in $\mathcal{F} \cap L^\infty(X)$ by

$$\int_X \phi\, d\Gamma(u, u) := \mathcal{E}(\phi u, u) - \frac{1}{2}\mathcal{E}(u^2, \phi)$$

for any $\phi \in \mathcal{F} \cap C_c(X)$, then extended by polarization for arbitrary pairs of $u, v \in \mathcal{F} \cap L^\infty(X)$. For $u \in \mathcal{F}$, the energy measure of u is the limit of the energy measures associated with the truncation functions $(u \wedge n) \vee (-n)$ as $n \to \infty$.

As a generalization of the classical energy integral $\int_{\mathbb{R}^n} \nabla u \cdot \nabla v\, dx$ in $\mathbb{R}^n$, that is, intuitively as a measure given by gradient square, the energy measure satisfies the following properties. As mentioned earlier, we do not specify quasi-continuous modifications of functions.

- (Leibniz rule [16, Lemma 3.2.5]) For any $u, v, w \in \mathcal{F}$ with $uv \in \mathcal{F}$ (e.g. $u, v \in \mathcal{F} \cap L^\infty(X)$),

$$d\Gamma(uv, w) = u\, d\Gamma(v, w) + v\, d\Gamma(u, w).$$

- (Chain rule [16, Theorem 3.2.2]) For any $u, v \in \mathcal{F}$, any $\Phi \in C^1(\mathbb{R})$ with bounded derivative and satisfies $\Phi(0) = 0$, then $\Phi(u) \in \mathcal{F}$, and

$$d\Gamma(\Phi(u), v) = \Phi'(u)\, d\Gamma(u, v).$$

- (Cauchy-Schwartz inequality [16, Lemma 5.6.1]) For any $f, g, u, v \in \mathcal{F} \cap L^\infty(X)$ (more generally, for any $u, v \in \mathcal{F}$, $f \in L^2(X, \Gamma(u, u))$, and $g \in L^2(X, \Gamma(v, v))$),

$$\begin{aligned}\int |fg|\, d|\Gamma(u, v)| &\leq \left(\int f^2\, d\Gamma(u, u)\right)^{1/2} \left(\int g^2 d\Gamma(v, v)\right)^{1/2} \\ &\leq \frac{C}{2}\int f^2\, d\Gamma(u, u) + \frac{1}{2C}\int g^2 d\Gamma(v, v). \qquad (2.1)\end{aligned}$$

The last inequality holds for any $C > 0$. The corresponding measure version is

$$|fg|\, d|\Gamma(u, v)| \leq \frac{C}{2} f^2\, d\Gamma(u, u) + \frac{1}{2C} g^2\, d\Gamma(v, v).$$

- (Strong locality [16, Corollary 3.2.1]) For any $u, v \in \mathcal{F}$, if on some precompact open set $U \Subset X$, $v \equiv C$ for some constant C, then

$$1_U\, d\Gamma(u, v) = 0.$$

Here the symbol $\Subset$ refers to precompact inclusion, i.e., $A \Subset B$ means that the closure $\overline{A}$ of A is a compact subset of B.

Any Dirichlet form $(\mathcal{E}, \mathcal{F})$ is associated with a corresponding Markov semigroup $(H_t)_{t>0}$, an (infinitesimal) generator $-P$ with dense domain $\mathcal{D}(P)$, and a Markov resolvent $(G_\alpha)_{\alpha>0}$ (in the sense of [16, Section 1.3]). The semigroup H_t and resolvent G_α have domain $L^2(X, m)$; the domain $\mathcal{D}(P)$ of $-P$ is dense in $\mathcal{F}$ with respect to the $\mathcal{E}_1$ norm. These are self-adjoint operators. By spectral theory, P has a spectral resolution $(E_\lambda)_{\lambda\geq 0}$ such that, for any $t > 0$,

$$PH_t = \int_0^\infty \lambda e^{-\lambda t}\, dE_\lambda.$$

As a consequence, for any $k \in \mathbb{N}$ where $\mathbb{N} = \{0, 1, 2, \ldots\}$,

$$\left\| \partial_t^k H_t \right\|_{L^2(X)\to L^2(X)} = \left\| P^k H_t \right\|_{L^2(X)\to L^2(X)} \leq (k/et)^k .$$

For any function $u_0 \in L^2(X, m)$, $u(t, x) := H_t u_0(x)$ is smooth in $t > 0$, and solves

$$\partial_t u = -Pu$$

in the strong sense. That is,

$$\lim_{h\to 0} \frac{u(t+h, \cdot) - u(t, \cdot)}{h} = -Pu(t, \cdot)$$

in $L^2(X, m)$.

Given the notations above, our main goal in this section is to define local weak solutions of the heat equation (with appropriate right-hand side f)

$$(\partial_t + P)u = f.$$

2.2 Function Spaces Associated with $(\mathcal{E}, \mathcal{F})$

To properly discuss candidate functions for local weak solutions, and later their properties, we first introduce some function spaces associated with $(\mathcal{E}, \mathcal{F})$. In choosing notations for these function spaces, we mostly follow [34], with a few exceptions that we will remark on later. Among these function spaces there are two prevalent types, one type consisting of functions that have compact support (all denoted with subscript "c"); the other type of functions that locally satisfy the required properties (all with subscript "loc").

Recall that $\mathcal{F} \subset L^2(X)$ and the inclusion is dense. Equating $L^2(X)$ with its dual with respect to the L^2 inner product, we get the Hilbert triple

$$\mathcal{F} \subset L^2(X) \subset \mathcal{F}'$$

in which the inclusions are dense and continuous. Intuitively, the "$\sim_c$" spaces are on the "$\mathcal{F}$" end, and the "$\sim_{\text{loc}}$" spaces are on the "$\mathcal{F}'$" (dual space) end. We consider the dual spaces of "$\sim_c$" spaces too.

We now give precise definitions of these spaces, organized in pairs, starting with the following two pairs:

- $\mathcal{F}_c(X) := \{f \in \mathcal{F} \mid f \text{ has compact (essential) support}\}$;
- $\mathcal{F}_{\text{loc}}(X) := \big\{ f \in L^2_{\text{loc}}(X, m) \mid$
  $\forall \text{compact } K \subset X\ \exists f^\sharp \in \mathcal{F} \text{ s.t. } f^\sharp = f\ m\text{-a.e. on } K \big\}$.

For any open subset $U \subset X$, define

- $\mathcal{F}_c(U) := \{f \in \mathcal{F} \mid f \text{ has compact (essential) support in } U\}$;

– $\mathcal{F}_{\text{loc}}(U) := \big\{ f \in L^2_{\text{loc}}(U, m) \mid$
$\forall$compact $K \subset U\ \exists f^\sharp \in \mathcal{F}$ s.t. $f^\sharp = f$ m-a.e. on $K \big\}$.

Remark 1 When $U \neq X$, by definition, there is an injection $i : \mathcal{F}_c(U) \hookrightarrow \mathcal{F}_c(X)$, and clearly $\mathcal{F}_{\text{loc}}(X) \hookrightarrow \mathcal{F}_{\text{loc}}(U)$ by restriction to U. Note, however, that $\mathcal{F}_{\text{loc}}(U)$ is not a subspace of $\mathcal{F}_{\text{loc}}(X)$.

Fix an arbitrary open interval $I = (a, b) \subset \mathbb{R}$, $-\infty \le a < b \le \infty$. Consider the following function spaces involving time and space. In defining these spaces, we switch freely between two viewpoints where elements in these spaces are viewed (1) as functions of time and space; (2) as functions on the time interval I with values in some (spatial) function space. The rigorous setup for the latter viewpoint is the theory of Bochner integrals, for which we refer to [37, Section 24].

In the sequel, when there is no ambiguity, we use the notation $u^t(\cdot)$ as an abbreviation for $u(t, \cdot)$. That is, for any fixed t, consider $u(t, y)$ as a function of y, denoted by u^t. Note that this is not any power of u or time derivative of u; the time derivative is denoted by $\partial_t u$.

First, we fix the notation for the "base space"

– $\mathcal{F}(I \times X) := L^2(I \to \mathcal{F})$, the L^2 space of functions on I with values in $\mathcal{F}$.

Remark 2 The space $L^2(I \to \mathcal{F})$ is the completion of the space of bounded continuous functions from I to $\mathcal{F}$, $C_b(I \to \mathcal{F})$, under the norm

$$\|u\|_{L^2(I \to \mathcal{F})} = \left(\int_I \|u^t\|^2_{\mathcal{E}_1} \, dt \right)^{1/2}.$$

The space $C_c^\infty(I \to \mathcal{F})$ of smooth compactly supported functions from I to $\mathcal{F}$ is also dense in $L^2(I \to \mathcal{F})$ with respect to the $\|\cdot\|_{L^2(I \to \mathcal{F})}$ norm. We use the notation $\mathcal{F}(I \times X)$ to clarify the use of notations $\mathcal{F}_c(I \times U)$ and $\mathcal{F}_{\text{loc}}(I \times U)$ for function spaces defined below. See also Remark 4.

Based on the "base space" $\mathcal{F}(I \times X)$, for any open subset $U \subset X$, define

– $\mathcal{F}_c(I \times U) := \{u \in \mathcal{F}(I \times X) \mid u$ is compactly supported in $I \times U\}$;
– $\mathcal{F}_{\text{loc}}(I \times U) := \big\{u \in L^2_{\text{loc}}(I \times U) \mid \forall$ open interval $I' \Subset I$, $\forall$ open subset $U' \Subset U$, $\exists u^\sharp \in \mathcal{F}(I \times X)$, s.t. $u^\sharp = u$ on $I' \times U'$ a.e.$\big\}$.

Here a.e. refers to $dt \times dm$-a.e.. The first two spaces $\mathcal{F}(I \times X)$ and $\mathcal{F}_c(I \times U)$ are subspaces of $L^2(I \times X)$ and $L^2(I \times U)$, respectively. We identify the L^2-spaces with their own duals (under the L^2 inner product), and denote the dual spaces of $\mathcal{F}(I \times X)$, $\mathcal{F}_c(I \times U)$ by $(\mathcal{F}(I \times X))'$, $(\mathcal{F}_c(I \times U))'$.

Remark 3 $(\mathcal{F}(I \times X))' = \big(L^2(I \to \mathcal{F})\big)' = L^2(I \to \mathcal{F}')$.

Remark 4 Here our notations are slightly different from the ones used in other places (e.g. [17, 34]). In the definition of $\mathcal{F}(I \times X)$, we do not require the functions to further be in $W^{1,2}(I \to \mathcal{F}')$, the space of functions in $L^2(I \to \mathcal{F}')$ with distributional time derivatives that belong to $L^2(I \to \mathcal{F}')$. The reason we consider the function spaces defined above instead of the ones obtained by taking the intersection with $W^{1,2}(I \to \mathcal{F}')$, is to put minimum assumptions in the definition of local weak solutions. Under our definition and hypotheses, such local weak solutions automatically satisfy better properties. In particular, we explain at the end of this section that under a very natural assumption on existence of cut-off functions, and when we require the right-hand side f to be locally in $L^2(I \to \mathcal{F}')$, our choice of definition of local weak solutions agrees with the definition used in other papers. This is verified by adapting the proof of Lemma 1 in [15].

To include more time derivatives we introduce the following notations for function spaces. For $k \in \mathbb{N}_+ := \{1, 2, \ldots\}$, define

- $\mathcal{F}^k(I \times X) := W^{k,2}(I \to \mathcal{F})$, the index-$k$ Sobolev space from I to $\mathcal{F}$;
- $\mathcal{F}^k_c(I \times U) := \{u \in \mathcal{F}^k(I \times X) \mid u$ is compactly supported in $I \times U\}$;
- $\mathcal{F}^k_{\text{loc}}(I \times U) := \{u \in L^2_{\text{loc}}(I \times U) \mid \forall$ open interval $I' \Subset I$, $\forall$ open subset $U' \Subset U$, $\exists u^\sharp \in \mathcal{F}^k(I \times X)$, s.t. $u^\sharp = u$ on $I' \times U'$ $a.e.\}$.

Here a.e. refers to $dt \times dm$-a.e.. When $k = 1$, these are the $\mathcal{F}$-spaces defined above. More details on the general theory of Sobolev spaces of functions with values in a Hilbert space can be found in for example [37, Section 25].

Remark 5 In general, we say that a function u is locally in some function space $\mathfrak{S}$, if for any compact subset of the underlying space with measure μ, there exists a function w in $\mathfrak{S}$ such that $w = u$ μ-a.e. on the compact set.

2.3 Notion of Local Weak Solutions

For any symmetric local regular Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(X, m)$, we define the following notion of local weak solutions of the associated heat equation (below $-P$ and $(H_t)_{t>0}$ are the corresponding generator and semigroup as before).

Definition 1 (local weak solution) Let $U \subset X$ be an open subset and $I \subset \mathbb{R}$ be an open interval. Let f be a function locally in $L^2(I \to \mathcal{F}')$. We say u is a *local weak solution* of the heat equation $(\partial_t + P)u = f$ on $I \times U$, if $u \in \mathcal{F}_{\text{loc}}(I \times U)$, and for any $\varphi \in \mathcal{F}_c(I \times U) \cap C^\infty_c(I \to \mathcal{F})$,

$$-\int_I \int_X u \cdot \partial_t \varphi \, dm dt + \int_I \mathcal{E}(u, \varphi) \, dt = \int_I \langle f, \varphi \rangle_{\mathcal{F}', \mathcal{F}} \, dt. \tag{2.2}$$

Here u in the integral is understood as $u^\sharp$ (relative to the support of φ) as in the definition for $\mathcal{F}_{\text{loc}}(I \times U)$. We take this convention throughout this paper. Note that $\mathcal{E}(u, \varphi)$ is well-defined (independent of the choice of $u^\sharp$) by the local property of $\mathcal{E}$. The symbol $\langle \cdot, \cdot \rangle_{\mathcal{F}', \mathcal{F}}$ stands for the pairing between elements in $\mathcal{F}'$ and $\mathcal{F}$.

We remark that we can define local weak solutions for more general right-hand side f, e.g., $f \in (\mathcal{F}_c(I \times U))'$. But in the propositions and theorems in this paper we always put more restrictions on f than f locally in $L^2(I \to \mathcal{F}')$; moreover, the results are interesting even for the case $f \equiv 0$, so here in the definition we do not aim to consider the most general right-hand side. With this choice, Definition 1 will be shown to be equivalent to the following variant, under a natural assumption on the existence of certain cut-off functions. As mentioned in Remark 4, the following definition is often adopted in the literature.

Definition 2 (local weak solution, variant) Let U, I, f be as in Definition 1. Let u be a function locally in $L^2(I \to \mathcal{F}) \cap W^{1,2}(I \to \mathcal{F}')$. u is called a local weak solution of the heat equation $(\partial_t + P)u = f$, if for any φ in $L^2(I \to \mathcal{F}) \cap W^{1,2}(I \to \mathcal{F}')$ with compact support in $I \times U$, for any subinterval $J \Subset I$,

$$\int_J \langle \partial_t u, \varphi \rangle_{\mathcal{F}',\mathcal{F}} \, dt + \int_J \mathcal{E}(u, \varphi) \, dt = \int_J \langle f, \varphi \rangle_{\mathcal{F}',\mathcal{F}} \, dt.$$

Note that in general,

$$\mathcal{F}_c(I \times U) \cdot \mathcal{F}_{\mathrm{loc}}(I \times U) \nsubseteq \mathcal{F}_c(I \times U),$$

roughly because $\mathcal{F}$ is not an algebra. Here

$$\mathcal{F}_c(I \times U) \cdot \mathcal{F}_{\mathrm{loc}}(I \times U) = \{gh \,|\, g \in \mathcal{F}_c(I \times U),\ h \in \mathcal{F}_{\mathrm{loc}}(I \times U)\}.$$

What we want to assume is that there is a subset of $\mathcal{F}_c(I \times U) \cap C(I \times U)$ that contains enough functions, each of which brings functions in $\mathcal{F}_{\mathrm{loc}}(I \times U)$ to $\mathcal{F}_c(I \times U)$ by multiplication (these can be thought of as cut-off functions with some nice properties). Here $C(I \times U)$ is the space of continuous functions in $I \times U$. We denote this subset of cut-off functions by $\mathfrak{C}(I \times U)$. Observe that we just need the existence of an analogous subset $\mathfrak{C}(U) \subset \mathcal{F}_c(U) \cap C(U)$, and then to construct $\mathfrak{C}(I \times U)$, take products of functions in $\mathfrak{C}(U)$ with standard cut-off functions in $C_c^\infty(I) \subset C_c^\infty(\mathbb{R})$. The following assumption makes precise what we require from the set $\mathfrak{C}(U) \subset \mathcal{F}_c(U) \cap C(U)$.

Assumption 2 There exists a subset $\mathfrak{C}(U) \subset \mathcal{F}_c(U) \cap C(U)$ such that

(i) for any pair of open sets $V \Subset U \Subset X$, there exists a function $\varphi \in \mathfrak{C}(U)$ such that $\varphi = 1$ on V, $\mathrm{supp}\{\varphi\} \subset U$;
(ii) for any $\varphi \in \mathfrak{C}(U)$, any $u \in \mathcal{F}_{\mathrm{loc}}(U)$, the product $\varphi u \in \mathcal{F}_c(U)$.

Remark 6 The requirement (i) in Assumption 2 is standard and easily fulfilled when the Dirichlet form is regular. The requirement (ii) is nontrivial. In general, only the products of functions in $\mathcal{F} \cap L^\infty(X)$ are guaranteed to belong to $\mathcal{F}$.

We now state the equivalence of the two definitions for local weak solutions.

Lemma 1 (**equivalence of definitions of local weak solutions**) *Under Assumption 2, Definition 1 is equivalent to Definition 2.*

Proof The proof follows essentially that of [15, Lemma 1]. □

3 Main Hypotheses

As summarized in the Introduction, two related types of assumptions play a key role in our analysis. We now introduce and elaborate on these assumptions. Let $(X, m, \mathcal{E}, \mathcal{F})$ be a symmetric local regular Dirichlet form as before, with the associated semigroup $(H_t)_{t>0}$.

3.1 Assumption on Existence of Cut-off Functions

For a pair of open sets $V \Subset U \Subset X$, by a cut-off function for the pair $V \subset U$ we mean a function $\eta \in \mathcal{F} \cap C(X)$ in between 0 and 1 such that $\eta = 1$ on V and $\mathrm{supp}\{\eta\} \subset U$. Such cut-off functions always exist for any pair of precompact open sets $V \subset U$ in a regular Dirichlet space, see [16, page 6 and Exercise 1.4.1]. For results in this paper we need the existence of cut-off functions that further have controlled energy, we explain what this means in the following assumption.

Assumption 3 (existence of nice cut-off functions) There exists some topological basis $\mathcal{TB}$ of X such that for any pair of open sets $V \Subset U$, $U, V \in \mathcal{TB}$, for any $0 < C_1 < 1$, there exist some constant $C_2(C_1, U, V) > 0$ and some cut-off function η for the pair $V \subset U$, such that for any $v \in \mathcal{F}$,

$$\int_X v^2 d\Gamma(\eta, \eta) \leq C_1 \int_X \eta^2 d\Gamma(v, v) + C_2 \int_{\mathrm{supp}\{\eta\}} v^2\, dm. \tag{3.1}$$

We call such η functions *nice cut-off functions* corresponding to C_1, C_2.

Remark 7 Later we show that in Assumption 3, the condition $U, V \in \mathcal{TB}$ for some topological basis $\mathcal{TB}$ is "redundant", in the sense that Assumption 3 implies automatically that nice cut-off functions in the sense of (3.1) exist for any pair of open sets $V \Subset U$. See Lemma 3. We also remark that Assumption 3 has a straightforward equivalent form that for any pair of precompact open sets U, V with disjoint closures, i.e., $\overline{U} \cap \overline{V} = \emptyset$, for any C_1 in $(0, 1)$, there exists a cut-off function η such that $\eta = 1$ on U, $\eta = 0$ on V, and there exists some constant $C_2(C_1, U, V) > 0$, such that for any $v \in \mathcal{F}$,

$$\int_X v^2 d\Gamma(\eta, \eta) \leq C_1 \int_X \eta^2 d\Gamma(v, v) + C_2(C_1, U, V) \int_{\mathrm{supp}\{\eta\}} v^2\, dm. \tag{3.2}$$

Let $\eta(x)$ be a nice cut-off function and $l(t)$ $(0 \leq l(t) \leq 1)$ be a smooth function on $\mathbb{R}$ with compact support, then the product $\eta(x)l(t)$ is a function in $\mathcal{F}_c(I \times X)$. We call such product functions *nice product cut-off functions*, and we denote such functions by adding an overline, i.e., $\overline{\eta}(t, x) := \eta(x)l(t)$.

Remark 8 If a cut-off function η for some pair $V \subset U$ satisfies that its corresponding energy measure is absolutely continuous with respect to m, and $d\Gamma(\eta, \eta)/dm$ is bounded, i.e.,

$$d\Gamma(\eta, \eta) \leq C\, dm \tag{3.3}$$

for some $C < \infty$, then η is a nice cut-off function and satisfies (3.1) with $C_1 = 0$ (hence any $0 < C_1 < 1$), $C_2 = C$. Here, C_2 is independent of C_1. We say in this special case that the cut-off function η has *bounded gradient*.

Conversely, if for some nice cut-off function η, (3.1) can be extended to hold true for $C_1 = 0$ and $C_2(0, U, V) < \infty$, then η has bounded gradient.

In particular, when the intrinsic pseudo-distance of the Dirichlet space,

$$\rho_X(x, y) = \sup\left\{\varphi(x) - \varphi(y) \,|\, \varphi \in \mathcal{F}_{\mathrm{loc}}(X) \cap C(X),\ d\Gamma(\varphi, \varphi) \leq dm\right\}, \tag{3.4}$$

is a continuous metric that induces the same topology of X, the Dirichlet space satisfies Assumption 3 with existence of cut-off functions with bounded gradient, and the cut-off functions can be explicitly constructed using the intrinsic distance. See Section 7.1 for more details.

Remark 9 Typical examples of Dirichlet spaces that satisfy Assumption 3 but do not possess cut-off functions with bounded gradient are some fractal spaces, including for example the Sierpinski gasket and the Sierpinski carpet. For fractal spaces, usually the existence of nice cut-off functions is guaranteed as a consequence of other properties like sub-Gaussian upper bounds satisfied by the Dirichlet space (heat kernel). In general, in such cases, there are no simple explicit constructions of cut-off functions satisfying (3.1). For references we mention [1] and [3].

Let $(X, m, \mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet space as before. We first verify that the cut-off functions in Assumption 3 indeed satisfy the conditions in Assumption 2.

Lemma 2 *Any nice cut-off function φ in the sense of (3.1) satisfies (ii) in Assumption 2. Namely, let $U \Subset X$ be some open set such that* $\operatorname{supp}\{\varphi\} \subset U$*, then for any $u \in \mathcal{F}_{\mathrm{loc}}(U)$, the product $\varphi u \in \mathcal{F}_c(U)$.*

Proof The support of the product function φu is clearly contained in U. To show $\varphi u \in \mathcal{F}$, recall that $u \in \mathcal{F}_{\mathrm{loc}}(U)$ means that u is in $L^2_{\mathrm{loc}}(U)$, and satisfies for any $V \Subset U$, there exists some $u^\sharp$ in $\mathcal{F}$ such that $u^\sharp = u$ m-a.e. on V. Pick some open set V such that $\operatorname{supp}\{\varphi\} \subset V \Subset U$, fix some $u^\sharp \in \mathcal{F}$ that agrees with u m-a.e. on V. Then

$$\begin{aligned}\|\varphi u^\sharp\|^2_{\mathcal{E}_1} &= \int_X (\varphi u^\sharp)^2\, dm + \int_X d\Gamma(\varphi u^\sharp, \varphi u^\sharp) + \int_X (\varphi u^\sharp)^2\, dk \\ &\leq \int_X (\varphi u^\sharp)^2\, dm + \int_X (\varphi u^\sharp)^2\, dk + 2\left[\int_X \varphi^2\, d\Gamma(u^\sharp, u^\sharp) + \int_X (u^\sharp)^2\, d\Gamma(\varphi, \varphi)\right].\end{aligned}$$

The first two terms are clearly finite, the third term is bounded above by $2(\mathcal{E}(u^\sharp, u^\sharp))^2$, and the last term is finite due to (3.1). Hence $\|\varphi u^\sharp\|_{\mathcal{E}_1} < +\infty$, and $\varphi u = \varphi u^\sharp \in \mathcal{F}_c(U)$. □

So far the examples we have described satisfy Assumption 3 for all pairs of open sets $V \Subset U$. The reason in Assumption 3 we only require nice cut-off functions to exist for pairs of open sets in some topological basis $\mathcal{TB}$ is to make the assumption easy to check for some infinite dimensional examples, like the infinite dimensional torus or the infinite product of Sierpinski gaskets.

In the next lemma we state the automatic extension of existence of nice cut-off functions for general pairs of open sets, given Assumption 3. We postpone the proof to Section 8.

Lemma 3 *Suppose Assumption 3 holds. Then for any open sets U, V with $V \Subset U$, for any constant $0 < C_1 < 1$, there exist some $C_2 = C_2(C_1, U, V) > 0$ and some nice cut-off function in the sense of (3.1) corresponding to C_1, C_2. In particular, U, V are not necessarily in $\mathcal{TB}$.*

Given any nice cut-off function and any function in the domain $\mathcal{F}$, by Lemma 2, their product belongs to $\mathcal{F}$. The energy of the product function satisfies the following estimate, which we later refer to as the gradient inequality.

Lemma 4 (gradient inequality) *Let η be a nice cut-off function associated with C_1, C_2 in the sense of (3.1), where $0 < C_1 < 1/4$, let $v \in \mathcal{F}$. Then*

$$\int_X d\Gamma(\eta v,\ \eta v) \leq \frac{1 - 2C_1}{1 - 4C_1} \int_X d\Gamma(\eta^2 v,\ v) + \frac{C_2}{1 - 4C_1} \int_{\operatorname{supp}\{\eta\}} v^2\, dm. \tag{3.5}$$

The point of the lemma is to bound the energy of the product function ηv on the left-hand side by L^2 integrals on the right-hand side, when $v \in \mathcal{D}(P)$. Indeed, the first integral then equals $\int_X \eta^2 v\, Pv\, dm$ (when $\mathcal{E}$ is strongly local).

It is easy to check the validity of this lemma in the special case when the cut-off function has bounded gradient. In this case, by expanding $\int_X d\Gamma(\eta v,\ \eta v)$ by the product rule and utilizing the upper bound $d\Gamma(\eta,\eta)/dm \le M$, we get

$$\int_X d\Gamma(\eta v,\ \eta v) \le \int_X d\Gamma(\eta^2 v,\ v) + M\int_{\mathrm{supp}\{\eta\}} v^2\,dm,$$

which is exactly (3.5) with $C_1 = 0$ and $C_2 = M$. In the general case, when the cut-off function does not have bounded gradient (thus C_1 in (3.1) must be taken as positive), (3.5) is less obvious, and we give the proof below.

Proof of Lemma 4 By the Cauchy-Schwartz inequality (2.1), and the assumption that η is a nice cut-off function associated with constants C_1, C_2,

$$\begin{aligned}
&\int_X d\Gamma(\eta v,\ \eta v) = \int_X \eta^2\,d\Gamma(v,v) + \int_X v^2\,d\Gamma(\eta,\eta) + 2\int_X \eta v\,d\Gamma(\eta,v)\\
&\ge \int_X \eta^2\,d\Gamma(v,v) + \int_X v^2\,d\Gamma(\eta,\eta) - \frac{1}{2}\int_X \eta^2\,d\Gamma(v,v) - 2\int_X v^2\,d\Gamma(\eta,\eta)\\
&= \frac{1}{2}\int_X \eta^2\,d\Gamma(v,v) - \int_X v^2\,d\Gamma(\eta,\eta)\\
&\ge \frac{1}{2}\int_X \eta^2\,d\Gamma(v,v) - \left[C_1\int_X \eta^2\,d\Gamma(v,v) + C_2\int_{\mathrm{supp}\{\eta\}} v^2\,dm\right]\\
&= \left(\frac{1}{2} - C_1\right)\int_X \eta^2\,d\Gamma(v,v) - C_2\int_{\mathrm{supp}\{\eta\}} v^2\,dm.
\end{aligned}$$

Hence as $C_1 < 1/4 < 1/2$,

$$\int_X \eta^2\,d\Gamma(v,v) \le \frac{1}{\frac{1}{2} - C_1}\int_X d\Gamma(\eta v,\ \eta v) + \frac{C_2}{\frac{1}{2} - C_1}\int_{\mathrm{supp}\{\eta\}} v^2\,dm. \tag{3.6}$$

On the other hand,

$$\begin{aligned}
\int_X d\Gamma(\eta v,\ \eta v) &= \int_X d\Gamma(\eta^2 v,\ v) + \int_X v^2\,d\Gamma(\eta,\eta)\\
&\le \int_X d\Gamma(\eta^2 v,\ v) + C_1\int_X \eta^2\,d\Gamma(v,v) + C_2\int_{\mathrm{supp}\{\eta\}} v^2\,dm.
\end{aligned}$$

Substituting the upper bound in (3.6) for $\int_X \eta^2\,d\Gamma(v,v)$ here, we get

$$\begin{aligned}
\int_X d\Gamma(\eta v,\ \eta v) &\le \int_X d\Gamma(\eta^2 v,\ v) + C_2\int_{\mathrm{supp}\{\eta\}} v^2\,dm\\
&+C_1\left(\frac{1}{\frac{1}{2} - C_1}\int_X d\Gamma(\eta v,\ \eta v) + \frac{C_2}{\frac{1}{2} - C_1}\int_{\mathrm{supp}\{\eta\}} v^2\,dm\right).
\end{aligned}$$

As $C_1 < 1/4$, this implies

$$\int_X d\Gamma(\eta v,\ \eta v) \le \frac{1-2C_1}{1-4C_1}\int_X d\Gamma(\eta^2 v,\ v) + \frac{C_2}{1-4C_1}\int_{\mathrm{supp}\{\eta\}} v^2\,dm.$$

□

In applications, we do not care about the exact constants. So in the following we consider $C_1 < 1/8$, and (3.5) implies

$$\int_X d\Gamma(\eta v,\, \eta v) \le 2\int_X d\Gamma(\eta^2 v,\, v) + 2C_2 \int_{\mathrm{supp}\{\eta\}} v^2\, dm. \tag{3.7}$$

Adding the positive terms $\int_X \eta^2 v^2\, dk$ to the left-hand side and $2\int_X \eta^2 v^2\, dk$ to the right-hand side of (3.7), we get

$$\mathcal{E}(\eta v,\, \eta v) \le 2\mathcal{E}(\eta^2 v,\, v) + 2C_2 \int_{\mathrm{supp}\{\eta\}} v^2\, dm. \tag{3.8}$$

3.2 L^2 Gaussian Type Upper Bound

In our treatment of the L^2 time regularity of local weak solutions, we rely much on the heat semigroup, which is smooth in time. Roughly speaking, we use the heat semigroup to construct an approximate sequence to a local weak solution u, and show that this approximate sequence (1) converges to u in some weak sense; (2) forms a Cauchy sequence in the space $\mathcal{F}^n(I \times X) = W^{n,2}(I \to \mathcal{F})$ for some integer $n \in [1, \infty]$. These two statements together then imply that u is (locally) in $\mathcal{F}^n(I \times X)$. To show the approximate sequence is Cauchy, we use the following (very weak) L^2 Gaussian type upper bound for the heat semigroup.

Assumption 4 (L^2 Gaussian type upper bound) For any two open sets $V_1, V_2 \Subset X$ with $\overline{V_1} \cap \overline{V_2} = \emptyset$, let

$$\mathcal{A}(V_1, V_2) := \left\{(g_1, g_2)\,\middle|\, \mathrm{supp}\{g_i\} \subset V_i,\ \|g_i\|_{L^2(V_i,m)} \le 1,\ i = 1, 2\right\}.$$

For any $a \ge 0$, any $n \in \mathbb{N}$, the semigroup H_t satisfies that

$$\lim_{t\to 0^+} \left(\sup_{(g_1,g_2)\in\mathcal{A}(V_1,V_2)} \left\{ \frac{1}{t^a} \left| \left\langle \partial_t^n H_t g_1,\, g_2 \right\rangle_{L^2(X)} \right| \right\} \right) = 0.$$

To simplify notation we write

$$G_{V_1,V_2}(a, n, t) := \max_{0\le k\le n} \sup \left\{ \frac{1}{t^a} \left| \left\langle \partial_t^k H_t g_1,\, g_2 \right\rangle_{L^2(X)} \right| \,\middle|\, (g_1, g_2) \in \mathcal{A}(V_1, V_2) \right\}.$$

In this notation, the condition above is

$$\lim_{t\to 0^+} G_{V_1,V_2}(a, n, t) = 0$$

for any $a \ge 0$, $n \in \mathbb{N}$. When this condition holds, we say that H_t satisfies the L^2 Gaussian type upper bound.

Remark 10 The L^2 Gaussian type bound above is a very weak upper bound. For example, from this bound itself we cannot tell if the heat semigroup admits a density, and even if we assume there is a density, neither can we say anything about the pointwise estimate of the density function. On the other hand, when there is some (global or local) pointwise Gaussian or sub-Gaussian upper bound, then the L^2 Gaussian bound is a very weak consequence. So we still name it "L^2 Gaussian type upper bound" after the name of the classical pointwise Gaussian or sub-Gaussian upper bound.

Typical examples where the L^2 Gaussian type upper bound for the heat semigroup holds are when there are enough cut-off functions with bounded gradient (see (3.3)), or when

Assumption 3 holds with $C_2(C_1, U, V) = C(U, V)C_1^{-\alpha}$ for some $C(U, V) > 0$ and $\alpha > 0$. More precisely, under Assumption 3 with cut-off functions with bounded gradient, one can define the distance between sets as follows (cf. [2, 19]). For any two precompact Borel sets U, V,

$$d(U, V) := \sup_{\substack{\phi \in \mathcal{F}_{\mathrm{loc}}(X) \cap L^\infty(X) \\ d\Gamma(\phi,\phi) \le dm}} \left\{ \operatorname*{ess\,inf}_{x \in U} \phi(x) - \operatorname*{ess\,sup}_{y \in V} \phi(y) \right\}, \tag{3.9}$$

where ess sup and ess inf are with respect to the measure m. The following more concrete L^2 Gaussian bound is a classical result, often referred to as the Takeda formula (cf. [36, Lemma 4.1]). See [12, Theorem 2] for a more related statement and proof. Let V_1, V_2 be two precompact measurable subsets of X with $\overline{V_1} \cap \overline{V_2} = \emptyset$. Then $0 < d(V_1, V_2) < \infty$. For any pair $(g_1, g_2) \in \mathcal{A}(V_1, V_2)$, any $t > 0$,

$$\left| \langle H_t g_1,\ g_2 \rangle_{L^2(X)} \right| \le \exp\left\{ -\frac{d(V_1, V_2)^2}{4t} \right\}. \tag{3.10}$$

Proofs for various kinds of Gaussian upper bounds usually use the so-called Davies' method, cf. e.g. [11]. To generalize the upper bound for terms like $\left| \left\langle \partial_t^n H_t g_1,\ g_2 \right\rangle_{L^2(X)} \right|$, one can use for example the complex analysis method from [10], or the method in [13].

However, when the existence of nice cut-off functions with bounded gradient is not guaranteed, there could be disjoint closed measurable sets U, V with distance $d(U, V) = 0$ (because roughly speaking, the only functions with bounded gradient are constant functions). Then this distance notion is not helpful in getting a Gaussian type upper bound.

Under Assumption 3 with cut-off functions satisfying the general inequality (3.1), or (3.2) as in the equivalent form of Assumption 3 (see Remark 7), when furthermore C_2 depends on C_1 in the specific form $C_2(C_1, U, V) = C(U, V)C_1^{-\alpha}$ for some $\alpha > 0$, $C(U, V) > 0$, by a modification of Davies' method, we can show that for any $t > 0$,

$$\left| \langle H_t g_1,\ g_2 \rangle_{L^2(X)} \right| \le \exp\left\{ -\left(\frac{1}{4^{\alpha+1} C(V_1, V_2) t} \right)^{\frac{1}{1+2\alpha}} \right\}. \tag{3.11}$$

Here again V_1, V_2 are two precompact Borel sets in X with $\overline{V_1} \cap \overline{V_2} = \emptyset$, $(g_1, g_2) \in \mathcal{A}(V_1, V_2)$. A relevant but different L^2 upper bound is given in [1, Proposition 2.3], under a different assumption concerning existence of cut-off functions.

Both bounds (3.10) and (3.11) imply that the semigroup H_t satisfies the L^2 Gaussian type upper bound in Assumption 4. Note that (formally) if we take $\alpha = 0$ and $C(V_1, V_2) = d(V_1, V_2)^{-2}$ in (3.11), then we recover (3.10). In Section 8 we give a proof of (3.11), as well as how this implies a similar bound for $\left| \left\langle \partial_t^n H_t g_1,\ g_2 \right\rangle_{L^2(X)} \right|$.

4 Statement of the Main Results and Overview of the Proof

4.1 Statement of the Main Results

In this section we state our results on the time regularity property of local weak solutions of the heat equation $(\partial_t + P)u = f$. Our main result is that the regularity in time of u is as good as that of the right-hand side f. Note that as a local weak solution on some time-space cylinder $I \times U \subset I \times X$, u satisfies the prerequisite $u \in \mathcal{F}_{\mathrm{loc}}(I \times U)$, so any of its "$\mathcal{F}(I \times X)$ representative" $u^\sharp$ automatically has distributional time derivatives of any order. The challenge hence lies in showing that these time derivatives belong to $\mathcal{F}(I \times X) = L^2(I \to \mathcal{F})$. Our main theorem is the following.

Theorem 5 *Let (X, m) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form satisfying Assumption 3 (existence of nice cut-off functions). Assume that the associated heat semigroup $(H_t)_{t>0}$ satisfies Assumption 4 (the L^2 Gaussian type upper bound). Let $I = (a, b)$, $-\infty \le a < b \le \infty$, be an open interval, $U \subset X$ be an open set, and f be a function locally in $W^{n,2}(I \to L^2(U))$ for some $n \in \mathbb{N}$. Let u be a local weak solution of $(\partial_t + P)u = f$ on $I \times U$. Then u is in $\mathcal{F}^n_{\text{loc}}(I \times U)$.*

In short, Theorem 5 claims that if the right-hand side f of the heat equation locally has L^2 time derivatives up to order n, then so does the local weak solution u, and its time derivatives up to order n locally belong to $L^2(I \to \mathcal{F})$. An important implication of Theorem 5 is that the time derivatives of u (up to order n) are local weak solutions of the corresponding heat equations.

Corollary 1 *Under the hypotheses in Theorem 5, if f is locally in the space $W^{n,2}(I \to L^2(U))$, then for any $1 \le k \le n$, $\partial_t^k u$ is a local weak solution of*

$$(\partial_t + P)\,\partial_t^k u = \partial_t^k f. \tag{4.1}$$

In particular, if u is a local weak solution of $(\partial_t + P)u = 0$ on $I \times U$, then all time derivatives of u are local weak solutions of the same heat equation on $I \times U$.

Remark 11 It will be evident after we present the proofs, that Theorem 5 and Corollary 1 are of a local nature. In fact, to obtain the conclusions of these results, we may ignore the Dirichlet form $(\mathcal{E}, \mathcal{F})$ and use instead the restricted Dirichlet form $\mathcal{E}^U_0$ on U, the domain of which is the completion of $\mathcal{F}_c(U)$ with respect to the $\mathcal{E}_1$ norm. The subscript 0 refers to Dirichlet boundary condition. It is enough to have the hypotheses in Theorem 5 hold for $\mathcal{E}^U_0$ and its corresponding semigroup $(H^U_t)_{t>0}$, to conclude that local weak solutions $u \in \mathcal{F}^n_{\text{loc}}(I \times U)$. See Sections 4.2 and 4.3 in the companion paper [20] for more details and for more examples illustrating this point.

4.2 Sketch of Proof for a Special Case of Theorem 5

In the next two sections we prove Theorem 5 and Corollary 1. In this subsection, we give an outline of proof for a simplified case to demonstrate some main ideas while avoiding certain technicalities. The sketched proof below is only for illustration and is not part of the rigorous proof in the next two sections. The simplified setting we consider here (Proposition 1 below) concerns a compact space X and local weak solutions u of the heat equation on $I \times X$. There, Assumption 3 and Assumption 4 are not needed. To treat the general context of Theorem 5, we need these further assumptions to conduct localization, which brings in more complications.

Recall the following convention: for any function $g(s, x)$, we write $g^s(x) := g(s, x)$. In the special case where X is compact and u is a local weak solution on the "full" time-space cylinder $I \times X$, since $\mathcal{F}_c(X) = \mathcal{F} = \mathcal{F}_{\text{loc}}(X)$, we know that u^t itself is in the domain of the Dirichlet form, and in particular, in $L^2(X)$. The spaces $\mathcal{F}_c(I \times X)$, $\mathcal{F}(I \times X)$, $\mathcal{F}_{\text{loc}}(I \times X)$ are different due to the inclusion of the open time interval $I = (a, b)$. We do need to multiply u with some smooth cut-off function in time, but in the outline proof below we ignore that technicality and pretend that the functions are globally good in time.

Proposition 1 (special case of Theorem 5) *Let (X, m) be a compact metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form. Given $I = (a, b) \subset \mathbb{R}$ and a function f that is locally in $W^{1,2}(I \to L^2(X))$, let u be a local weak solution of $(\partial_t + P)u = f$ on $I \times X$. Then u is locally in $\mathcal{F}^1(I \times X)$.*

Outline of Proof Let $\rho \in C_c^\infty((1, 2))$ be some smooth nonnegative cut-off function on $\mathbb{R}$ with $\int_{\mathbb{R}} \rho(t)\, dt = 1$. For any $\tau > 0$, define

$$\rho_\tau(t) := \frac{1}{\tau}\rho\left(\frac{t}{\tau}\right).$$

Then $\operatorname{supp}\{\rho_\tau\} \subset (\tau, 2\tau)$, and $\{\rho_\tau\}_{\tau>0}$ is an approximation to identity in $\mathbb{R}$. Note that $\partial_\tau \rho_\tau(t) = -\partial_t \bar{\rho}_\tau(t)$, where

$$\bar{\rho}_\tau(t) := \frac{t}{\tau^2}\rho\left(\frac{t}{\tau}\right).$$

Define an approximate sequence $\{u_\tau\}_{\tau>0}$ as follows. For any $\tau > 0$, let

$$u_\tau(s, x) := \int_I \rho_\tau(s - t) H_{s-t} u^t(x)\, dt,\ \ (s, x) \in I \times X.$$

Observe that because of the ρ_τ term, the integrand is nonzero only when $t \in I \cap (s - 2\tau,\ s - \tau)$. In particular, $t < s$ so that H_{s-t} is well-defined, and the integral makes sense as a Bochner integral. Note that when there is the notion of convolution and when H_t admits a density function (heat kernel), the approximate sequence above is exactly the convolution in time and space of u and the heat kernel (with a cut-off function ρ_τ in time).

Because H_t is smooth in time, it is easy to show that u_τ is smooth in time. More precisely, for any $\tau > 0$, $u_\tau \in C^\infty(I \to \mathcal{F})$. It is routine to verify that u_τ converges to u in $L^2(I \times X)$ as τ tends to 0. So to prove the proposition, it suffices to show that $\{u_\tau\}_{\tau>0}$ is Cauchy in $W^{1,2}(I \to \mathcal{F}) = \mathcal{F}^1(I \times X)$. Here by $\{u_\tau\}_{\tau>0}$ is Cauchy, we mean that for any subsequence τ_j that converges to 0 as j tends to infinity, the sequence $\{u_{\tau_j}\}_{j\in\mathbb{N}_+}$ is Cauchy.

To this end, we show that $\|\partial_\tau u_\tau\|_{W^{1,2}(I\to\mathcal{F})}$ is integrable in τ near 0, then for any $0 < \gamma' < \gamma$, by Minkowski's (integral) inequality,

$$\begin{aligned}\|u_\gamma - u_{\gamma'}\|_{W^{1,2}(I\to\mathcal{F})} &= \left\| \int_{\gamma'}^{\gamma} \partial_\tau u_\tau \, d\tau \right\|_{W^{1,2}(I\to\mathcal{F})} \\ &\le \int_0^\gamma \|\partial_\tau u_\tau\|_{W^{1,2}(I\to\mathcal{F})}\, d\tau \to 0 \ \text{ as } \gamma \to 0,\end{aligned}$$

thus $\{u_\tau\}_{\tau>0}$ is Cauchy in $W^{1,2}(I \to \mathcal{F})$. We first estimate $\|\partial_\tau \partial_s u_\tau\|_{L^2(I\times X)}$. By duality,

$$\|\partial_\tau \partial_s u_\tau\|_{L^2(I\times X)} = \sup_{\substack{\|\varphi\|_{L^2(I\times X)}\le 1 \\ \varphi\in C_c^\infty(I\to L^2(X))}} \langle \partial_\tau \partial_s u_\tau,\ \varphi\rangle_{L^2(I\times X)}\,.$$

Here $\langle\cdot,\cdot\rangle_{L^2(I\times X)}$ denotes the standard inner product on $L^2(I\times X)$. Using $\partial_\tau\rho_\tau(t)=-\partial_t\bar{\rho}_\tau(t)$ to express $\partial_\tau\partial_s u_\tau$, we have

$$\begin{aligned}
&\|\partial_\tau\partial_s u_\tau\|_{L^2(I\times X)}\\
&= \sup_{\substack{\|\varphi\|_{L^2(I\times X)}\le 1\\ \varphi\in C_c^\infty(I\to L^2(X))}} \left\{\left\langle\int_I \partial_s\left[\partial_t\bar{\rho}_\tau(s-t)H_{s-t}\right]u^t(x)\,dt,\ \varphi\right\rangle_{L^2(I\times X)}\right\}\\
&= \sup_{\substack{\|\varphi\|_{L^2(I\times X)}\le 1\\ \varphi\in C_c^\infty(I\to L^2(X))}} \left\{\int_I\int_I\int_X u^t(x)\,\partial_s\partial_t\left[\bar{\rho}_\tau(s-t)H_{s-t}\right]\varphi^s(x)\,dmdsdt\right.\\
&\qquad\left.-\int_I\int_I\int_X u^t(x)\,\partial_s\left[\bar{\rho}_\tau(s-t)\,\partial_t H_{s-t}\right]\varphi^s(x)\,dmdsdt\right\}.
\end{aligned}$$

From the second line to the third line we used the Fubini theorem and the self-adjointness of H_t to move $\partial_s[\partial_t\bar{\rho}_\tau(s-t)H_{s-t}]$ from the "u" side to the "φ" side, then used the product rule to redistribute ∂_t. Because $\partial_t H_{s-t}=PH_{s-t}$ and u is a local weak solution of $(\partial_t+P)u=f$ on $I\times X$, the above two terms in the curly brackets together, modulo a cut-off function in time that we omit in this proof (i.e., think of the function $(t,x)\mapsto\partial_s[\bar{\rho}_\tau(s-t)H_{s-t}]\varphi^s(x)$ as a test function), equals

$$-\int_I\int_I\int_X f(t,x)\,\partial_s\left[\bar{\rho}_\tau(s-t)H_{s-t}\right]\varphi^s(x)\,dmdtds.$$

By rewriting $\partial_s[\bar{\rho}_\tau(s-t)H_{s-t}]$ as $-\partial_t[\bar{\rho}_\tau(s-t)H_{s-t}]$ and using integration by parts, we get

$$\begin{aligned}
&\|\partial_\tau\partial_s u_\tau\|_{L^2(I\times X)}\\
&= \sup_{\substack{\|\varphi\|_{L^2(I\times X)}\le 1\\ \varphi\in C_c^\infty(I\to L^2(X))}} \left|\int_I\int_I\int_X \partial_t f(t,x)\,\bar{\rho}_\tau(s-t)H_{s-t}\varphi^s(x)\,dmdtds\right|.
\end{aligned}$$

Here we did not consider the boundary term, but that is not a problem once we add in the cut-off function in time in the rigorous proof in the next section. For the same reason, we think of $\|f\|_{W^{1,2}(I\to L^2(X))}$ as being finite, when more rigorously it should be f multiplied with some cut-off function in time. We now show that the above integral has an upper bound in terms of $\|f\|_{W^{1,2}(I\to L^2(X))}$, which in particular is independent of τ. First, by Hölder's inequality,

$$\begin{aligned}
&\left|\int_I\int_I\int_X \partial_t f(t,x)\,\bar{\rho}_\tau(s-t)H_{s-t}\varphi^s(x)\,dmdtds\right|\\
&= \left|\int_I\int_X \partial_t f(t,x)\left(\int_I \bar{\rho}_\tau(s-t)H_{s-t}\varphi^s(x)\,ds\right)dmdt\right|\\
&\le \|\partial_t f\|_{L^2(I\times X)}\left(\int_I\int_X\left(\int_I \bar{\rho}_\tau(s-t)H_{s-t}\varphi^s(x)\,ds\right)^2 dmdt\right)^{1/2}\\
&:= \|\partial_t f\|_{L^2(I\times X)}\left(C(\tau,\varphi)\right)^{1/2}.
\end{aligned}$$

To estimate $C(\tau, \varphi)$, first note that for any $t \in I$ (i.e., $a < t < b$),

$$\begin{aligned} \int_I \bar{\rho}_\tau(s-t)\,ds &= \int_a^b \frac{s-t}{\tau}\rho\left(\frac{s-t}{\tau}\right)\frac{1}{\tau}ds \\ &= \int_0^{\frac{b-t}{\tau}} \theta\rho(\theta)\,d\theta < 2\int_0^{\frac{b-t}{\tau}} \rho(\theta)\,d\theta \le 2, \end{aligned} \tag{4.2}$$

where in the last line we made a change of variable $\theta = (s-t)/\tau$, and used the fact that ρ is supported in $(1, 2)$. Let $r(t) := \int_I \bar{\rho}_\tau(s-t)\,ds$, then $0 \le r(t) < 2$. When $r(t) > 0$, by Jensen's inequality we have

$$\begin{aligned} &\left(\int_I \bar{\rho}_\tau(s-t)H_{s-t}\varphi^s(x)\,ds\right)^2 \\ &= (r(t))^2\left(\frac{1}{r(t)}\int_I \bar{\rho}_\tau(s-t)H_{s-t}\varphi^s(x)\,ds\right)^2 \\ &\le r(t)\int_I \bar{\rho}_\tau(s-t)\left(H_{s-t}\varphi^s(x)\right)^2 ds. \end{aligned}$$

The inequality holds for $r(t) = 0$ too. So the term $C(\tau, \varphi)$ satisfies that

$$\begin{aligned} C(\tau, \varphi) &\le 2\int_I\int_I \bar{\rho}_\tau(s-t)\int_X \left(H_{s-t}\varphi^s(x)\right)^2 dm\,dsdt \\ &\le 2\left(\sup_{s\in I}\int_I \bar{\rho}_\tau(s-t)\,dt\right)\int_I\int_X \varphi(s,x)^2\,dmds, \end{aligned}$$

where in the second inequality, we used the fact that the semigroup is a contraction semigroup, $\|H_{s-t}\varphi^s\|_{L^2(X)} \le \|\varphi^s\|_{L^2(X)}$. We can similarly check that $\sup_{s\in I}\int_I \bar{\rho}_\tau(s-t)\,dt < 2$. Hence

$$\begin{aligned} &\sup_{0<\tau<1} \|\partial_\tau\partial_s u_\tau\|_{L^2(I\times X)} \\ &\le \sup_{0<\tau<1} \sup_{\substack{\|\varphi\|_{L^2(I\times X)}\le 1 \\ \varphi\in C_c^\infty(I\to L^2(X))}} \|\partial_t f\|_{L^2(I\times X)}\,(C(\tau,\varphi))^{1/2} \le 2\|f\|_{W^{1,2}(I\to L^2(X))}. \end{aligned}$$

To estimate $\|\partial_\tau\partial_s u_\tau\|_{L^2(I\to\mathcal{F})}$, note that for any $\tau > 0$, $s \in I$, $\partial_\tau\partial_s u_\tau$ belongs to $\mathcal{D}(P)$. Thus

$$\left(\int_I \mathcal{E}(\partial_\tau\partial_s u_\tau,\ \partial_\tau\partial_s u_\tau)\,ds\right)^{1/2} \le \|\partial_\tau\partial_s u_\tau\|^{1/2}_{L^2(I\times X)}\,\|P(\partial_\tau\partial_s u_\tau)\|^{1/2}_{L^2(I\times X)}.$$

It is shown above that $\sup_{0<\tau<1}\|\partial_\tau\partial_s u_\tau\|_{L^2(I\times X)} < \infty$, so it suffices to show that for $0 < \tau < 1$,

$$\|P(\partial_\tau\partial_s u_\tau)\|_{L^2(I\times X)} \le \frac{C}{\tau}$$

for some constant C that depends only on f. Running the estimates above with H_{s-t} replaced by PH_{s-t}, we can get the desired estimate. See the rigorous proof in the next section for more details. □

5 Proof of the Main Results

5.1 Proof of Theorem 5 - General Strategy

In this section we prove Theorem 5. To verify $u \in \mathcal{F}^n_{\mathrm{loc}}(I \times U)$, we show that for any $J \times V \Subset I \times U$, there exists some function in $\mathcal{F}^n(I \times X)$ that equals $\overline{\psi} u$ a.e. over $J \times V$. Here $\overline{\psi}(s, x) := \psi(x)w(s)$ is some nice product cut-off function such that $\overline{\psi} \equiv 1$ on some $J_{\overline{\psi}} \times V_{\overline{\psi}}$ where $J \times V \Subset J_{\overline{\psi}} \times V_{\overline{\psi}}$; $\mathrm{supp}\{\overline{\psi}\} \subset I_{\overline{\psi}} \times U_{\overline{\psi}}$ for some $I_{\overline{\psi}} \times U_{\overline{\psi}} \Subset I \times U$. Our notational choice is that J, V are proper subsets of I, U, and subscripts mark which function these sets are "affiliated with".

More precisely, we first define an approximate sequence (now, with proper nice cut-off functions inserted) to the local weak solution u and show that the approximate sequence is Cauchy in $\mathcal{F}^n(I \times X)$. Next, we show that the sequence converges to $\overline{\psi} u$ in the L^2 sense (this step does not make use of the fact that u is a local weak solution). The limit of the approximate sequence then serves as the function in $\mathcal{F}^n(I \times X)$ that agrees with $\overline{\psi} u$ a.e. on $J \times V$.

The approximate sequence is defined as follows. Let ρ_τ be as in the last section, that is, $\rho(t) \in C_c^\infty((1, 2))$ is some nonnegative smooth function satisfying $\int_{\mathbb{R}} \rho(t)\, dt = 1$, and $\rho_\tau(t)$ is defined by $\rho_\tau(t) = (1/\tau)\rho(t/\tau)$ $(\tau > 0)$. Note that $\mathrm{supp}\{\rho_\tau\} \subset (\tau, 2\tau)$. Recall that $\partial_\tau \rho_\tau(t) = -\partial_t \bar{\rho}_\tau(t)$, where $\bar{\rho}_\tau(t) = (t/\tau^2)\rho(t/\tau)$. Let $\overline{\eta}(y, t) = \eta(y)l(t)$ be another nice product cut-off function which is 1 over some neighborhood of the support of $\overline{\psi}$. More precisely, $\overline{\eta} \equiv 1$ on some $J_{\overline{\eta}} \times V_{\overline{\eta}}$ where $J \times V \Subset I_{\overline{\psi}} \times U_{\overline{\psi}} \Subset J_{\overline{\eta}} \times V_{\overline{\eta}}$; $\mathrm{supp}\{\overline{\eta}\} \subset I_{\overline{\eta}} \times U_{\overline{\eta}}$ for some $I_{\overline{\eta}} \times U_{\overline{\eta}} \Subset I \times U$. Consider the sequence $\{\widetilde{u}_\tau\}_{\tau > 0}$ defined by

$$\widetilde{u}_\tau(s, x) := \int_I \rho_\tau(s - t) H_{s-t}\left(\overline{\eta}^t u^t\right)(x)\, dt, \ \ (s, x) \in I \times X.$$

Like in the definition of u_τ in Section 4.2, the integrand is nonzero only when $t \in I \cap (s - 2\tau, s - \tau)$. This guarantees the integral is well-defined.

As mentioned above, we claim that (1) the family $\left\{\overline{\psi}\widetilde{u}_\tau\right\}_{\tau>0}$ is Cauchy in $\mathcal{F}^n(I \times X)$ and hence has a limit in the same function space, here Cauchy means that any subsequence $\left\{\overline{\psi}\widetilde{u}_{\tau_k}\right\}_{k\in\mathbb{N}_+}$ with $\tau_k \to 0$ is a Cauchy sequence in $\mathcal{F}^n(I \times X)$; (2) $\overline{\psi}\widetilde{u}_\tau \to \overline{\psi}\,\overline{\eta} u = \overline{\psi} u$ in $L^2(I \times X)$ as $\tau \to 0$. So the two limit functions must equal a.e.; in particular, the "L^2-limit" $\overline{\psi} u$ in fact belongs to $\mathcal{F}^n(I \times X)$. Because $\overline{\psi} u = u$ a.e. on $J \times V$, and $J \times V$ is arbitrarily taken, the statement in Theorem 5 follows.

To prove $\left\{\overline{\psi}\widetilde{u}_\tau\right\}_{\tau>0}$ is Cauchy in $\mathcal{F}^n(I \times X) = W^{n,2}(I \to \mathcal{F})$, we first show that for each $\tau > 0$, $\overline{\psi}\widetilde{u}_\tau \in C^\infty(I \to \mathcal{F})$. It then suffices to prove the following two propositions.

Proposition 2 *Under the hypotheses in Theorem 5, for any nice product cut-off function $\overline{\psi}$ supported in $I \times U$,*

$$\max_{0\le k\le n} \sup_{0<\tau<1} \left\| \partial_\tau \partial_s^k \left(\overline{\psi}\widetilde{u}_\tau\right)\right\|_{L^2(I\times X)} < +\infty.$$

Proposition 3 *Under the hypotheses in Theorem 5, for any nice product cut-off function $\overline{\psi}$ supported in $I \times U$, for any* $0 < \tau < 1$,

$$\max_{0\le k\le n} \left(\int_I \mathcal{E}(\partial_\tau \partial_s^k(\overline{\psi}\widetilde{u}_\tau),\ \partial_\tau \partial_s^k(\overline{\psi}\widetilde{u}_\tau))\, ds \right)^{1/2} \lesssim \frac{1}{\sqrt{\tau}}.$$

Here $\lesssim$ means that the left-hand side is less than some finite positive constant C times the right-hand side, where C is independent of τ. These two propositions together imply that for any $0 < \gamma' < \gamma < 1$,

$$\begin{aligned}\|\overline{\psi}\widetilde{u}_\gamma - \overline{\psi}\widetilde{u}_{\gamma'}\|_{W^{n,2}(I\to\mathcal{F})} &= \left\| \int_{\gamma'}^{\gamma} \partial_\tau(\overline{\psi}\widetilde{u}_\tau)\,d\tau \right\|_{W^{n,2}(I\to\mathcal{F})} \\ &\le \int_0^\gamma \left\|\partial_\tau(\overline{\psi}\widetilde{u}_\tau)\right\|_{W^{n,2}(I\to\mathcal{F})} d\tau \lesssim \int_0^\gamma \frac{1}{\sqrt{\tau}}\,d\tau = 2\sqrt{\gamma},\end{aligned}$$

which tends to 0 as γ tends to 0. Here, the first inequality is by Minkowski's inequality and enlarging the domain of integration to $[0, \gamma]$. It thus follows that the family $\left\{\overline{\psi}\widetilde{u}_\tau\right\}_{\tau>0}$ is Cauchy in $W^{n,2}(I \to \mathcal{F})$.

We now verify that $\overline{\psi}\widetilde{u}_\tau \in C^\infty(I \to \mathcal{F})$ for every $\tau > 0$. Recall that the norm $\|\cdot\|_{C^n(\mathbb{R})}$ is the sum of L^∞ norms of the function and its derivatives up to order n. Note that for any fixed $\tau > 0$ and $m, n \in \mathbb{N}$,

$$\begin{aligned}&\int_I |\partial_s^n \rho_\tau(s-t)| \left\|P^m H_{s-t}(\overline{\eta}^t u^t)\right\|_{L^2(X)} dt \\ &\le \frac{1}{(e\tau)^m}\|\rho_\tau\|_{C^n(\mathbb{R})}|\mathrm{supp}\{l\}|^{1/2}\|\overline{\eta}u\|_{L^2(I\times X)},\end{aligned}$$

which is a finite upper bound and independent of $s \in I$. Here $|\mathrm{supp}\{l\}|$ is the one-dimensional Lebesgue measure of $\mathrm{supp}\{l\}$. It follows that all $\partial_s^k\widetilde{u}_\tau$, $k \in \mathbb{N}$, are well-defined as Bochner integrals and are in $L^\infty(I \to \mathcal{F})$. Hence $\widetilde{u}_\tau \in C^\infty(I \to \mathcal{F})$. More precisely, we have

$$\begin{aligned}\left\|\partial_s^k\widetilde{u}_\tau\right\|_{L^\infty(I\to\mathcal{F})} &= \sup_{s\in I}\left(\mathcal{E}_1(\partial_s^k\widetilde{u}_\tau,\ \partial_s^k\widetilde{u}_\tau)\right)^{1/2} \\ &\le \sup_{s\in I}\left(\left\|\partial_s^k\widetilde{u}_\tau\right\|_{L^2(X)}\left\|P\partial_s^k\widetilde{u}_\tau\right\|_{L^2(X)} + \left\|\partial_s^k\widetilde{u}_\tau\right\|_{L^2(X)}^2\right)^{1/2}.\end{aligned}$$

The estimate above implies that the right-hand side here is finite. The conclusion that $\overline{\psi}\widetilde{u}_\tau \in C^\infty(I \to \mathcal{F})$ then follows from applying the gradient inequality (3.8).

In the next two subsections we prove Proposition 2. We present the proof in two steps. In the first step we express and split $\left\|\partial_\tau\partial_s^k(\overline{\psi}\widetilde{u}_\tau)\right\|_{L^2(I\times X)}$ into three parts; in the second step we estimate each part and show that they are all bounded above independent of $0 < \tau < 1$ and $0 \le k \le n$.

5.2 Proof of Proposition 2 - Step 1

We first compute $\partial_\tau\widetilde{u}_\tau(s, x)$. For any $\tau > 0$, $(s, x) \in I \times X$,

$$\partial_\tau\widetilde{u}_\tau(s,x) = \int_I \partial_\tau\rho_\tau(s-t)H_{s-t}(\overline{\eta}^t u^t)(x)\,dt = \int_I \partial_t\bar{\rho}_\tau(s-t)H_{s-t}(\overline{\eta}^t u^t)(x)\,dt.$$

Recall that here

$$\bar{\rho}_\tau(s-t) = \frac{s-t}{\tau}\rho_\tau(s-t) = \frac{s-t}{\tau^2}\rho\left(\frac{s-t}{\tau}\right).$$

Let

$$\mathcal{T} := \{\varphi \mid \|\varphi\|_{L^2(I\times X)} \le 1,\ \varphi \in C_c^\infty(I \to L^2(X))\}.$$

Recall that $\overline{\psi}(s,x) = \psi(x)w(s)$. We have

$$\left\| \partial_\tau \partial_s^k (\overline{\psi}\widetilde{u}_\tau) \right\|_{L^2(I\times X)} = \sup_{\varphi\in\mathcal{T}} \left\langle \psi\,\partial_\tau \partial_s^k (w\widetilde{u}_\tau),\ \varphi \right\rangle_{L^2(I\times X)},$$

where

$$\begin{aligned}
&\left\langle \psi\,\partial_\tau \partial_s^k (w\widetilde{u}_\tau),\ \varphi \right\rangle_{L^2(I\times X)} \\
&= \int_I \int_X \left\{ \int_I \partial_s^k [w(s)(\partial_t \bar{\rho}_\tau(s-t))H_{s-t}](\overline{\eta}^t u^t)(x)\,dt \right\} \psi(x)\varphi(s,x)\,dmds \\
&= \int_I \int_I \int_X (\overline{\eta}^t u^t)(x)\,\partial_s^k [w(s)(\partial_t \bar{\rho}_\tau(s-t))H_{s-t}](\psi\varphi^s)(x)\,dmdtds.
\end{aligned}$$

The last line is by the Fubini Theorem (changing the integration order from $\int_I \int_X \int_I\ dtdmds$ to $\int_I \int_I \int_X\ dmdtds$) and by the self-adjointness of H_{s-t}. Using the product rule for ∂_t to rewrite $w(s)(\partial_t \bar{\rho}_\tau(s-t))H_{s-t}$ in the square bracket as $\partial_t(w(s)\bar{\rho}_\tau(s-t)H_{s-t}) - w(s)\bar{\rho}_\tau(s-t)\partial_t H_{s-t}$, altogether we get that

$$\begin{aligned}
&\left\| \partial_\tau \partial_s^k \left(\overline{\psi}\widetilde{u}_\tau\right) \right\|_{L^2(I\times X)} \\
&= \sup_{\varphi\in\mathcal{T}} \left\{ \int_I \int_I \int_X (\overline{\eta}^t u^t)(x)\,\partial_t \{\partial_s^k [w(s)\bar{\rho}_\tau(s-t)H_{s-t}](\psi\varphi^s)(x)\}\,dmdtds \right. \\
&\qquad \left. - \int_I \int_I \int_X (\overline{\eta}^t u^t)(x)\,\partial_s^k [w(s)\bar{\rho}_\tau(s-t)\partial_t H_{s-t}](\psi\varphi^s)(x)\,dmdtds \right\}.
\end{aligned}$$

In the last line, since $\partial_t H_{s-t} = PH_{s-t}$, the second term equals

$$\begin{aligned}
&\text{second term} \\
&= \int_I \int_I \int_X (\overline{\eta}^t u^t)(x)\ P[\partial_s^k (w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)(x)]\,dm\,dtds \\
&= \int_I \int_I \mathcal{E}(\overline{\eta}^t u^t,\ \partial_s^k (w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s))\,dtds.
\end{aligned}$$

To simplify notation, let

$$v_{k,\tau}(s,t,x) := \partial_s^k (w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)(x). \tag{5.1}$$

When $s,t\in I$ are fixed, we write $v_{k,\tau}^{s,t}(x) := v_{k,\tau}(s,t,x)$. It is clear that for any fixed $\tau > 0$ and $s,t \in I$, $v_{k,\tau}^{s,t} \in \mathcal{D}(P)$. Moreover, $v_{k,\tau} \in L^2(I^2 \to \mathcal{D}(P))$. Using $v_{k,\tau}$, we can rewrite the previous equality as

$$\begin{aligned}
&\left\| \partial_\tau \partial_s^k (\overline{\psi}\widetilde{u}_\tau) \right\|_{L^2(I\times X)} \\
&= \sup_{\varphi\in\mathcal{T}} \left\{ \int_I \int_I \int_X \overline{\eta}(t,x)u(t,x)\partial_t [v_{k,\tau}(s,t,x)]\,dmdtds \right. \\
&\qquad \left. - \int_I \int_I \mathcal{E}(\overline{\eta}(t,\cdot)u(t,\cdot),\ v_{k,\tau}(s,t,\cdot))\,dtds \right\}.
\end{aligned} \tag{5.2}$$

Recall that u is a local weak solution on $I \times U$. If in (5.2), $\overline{\eta}$ is not grouped with u but appears on the same side with $v_{k,\tau}$, then (5.2) is exactly

$$\sup_{\varphi\in\mathcal{T}} \int_I \langle f,\ \overline{\eta} v_{k,\tau}^s \rangle_{L^2(I\times X)}\ ds.$$

This observation inspires us to write (5.2) as this term plus the difference, and then estimate them each separately. More precisely, using (5.2), we have

$$\begin{aligned}&\left\|\partial_\tau \partial_s^k(\overline{\psi}\widetilde{u}_\tau)\right\|_{L^2(I\times X)}\\&\le \sup_{\varphi\in\mathcal{T}}|A_k(\tau,\varphi)|+\sup_{\varphi\in\mathcal{T}}|B_k(\tau,\varphi)|+\sup_{\varphi\in\mathcal{T}}|C_k(\tau,\varphi)|,\end{aligned}$$

where

$$\begin{aligned}A_k(\tau,\varphi)&=\int_I\int_I\int_X(\overline{\eta}^t u^t)\,\partial_t v_{k,\tau}^{s,t}-u^t\,\partial_t(\overline{\eta}^t v_{k,\tau}^{s,t})\,dmdtds\\&=-\int_I\int_I\int_X u(t,x)(\partial_t\overline{\eta}(t,x))v_{k,\tau}(s,t,x)\,dmdtds;\end{aligned}$$

$$\begin{aligned}B_k(\tau,\varphi)&=-\int_I\int_I\mathcal{E}(\overline{\eta}^t u^t,\,v_{k,\tau}^{s,t})\,dtds+\int_I\int_I\mathcal{E}(u^t,\,\overline{\eta}^t v_{k,\tau}^{s,t})\,dtds\\&=-\int_I\int_I\int_X d\Gamma(\overline{\eta}^t u^t,\,v_{k,\tau}^{s,t})\,dtds+\int_I\int_I\int_X d\Gamma(u^t,\,\overline{\eta}^t v_{k,\tau}^{s,t})\,dtds;\end{aligned}$$

$$C_k(\tau,\varphi)=\int_I\left\langle f,\,\overline{\eta}v_{k,\tau}^s\right\rangle_{L^2(I\times X)}ds=\int_I\int_I\int_X f(t,x)\overline{\eta}(t,x)v_{k,\tau}(s,t,x)\,dmdtds.$$

5.3 Proof of Proposition 2 - Step 2

Next we estimate $|A_k(\tau,\varphi)|$, $|B_k(\tau,\varphi)|$, and $|C_k(\tau,\varphi)|$ individually. We will see that the upper bounds we find for $|A_k|$, $|B_k|$, $|C_k|$ involve some L^2 or $\mathcal{E}_1$ norms of the local weak solution u on some precompact subsets of $I\times X$ (hence the norms are well-defined). To conveniently express these norms of u, we introduce a nice (product) cut-off function that lives in (i.e., has compact support in) $I\times U$ and is flat 1 on some open set that covers the supports of all other cut-off functions in the whole proof. We denote this cut-off function by $\overline{\Psi}(t,x)=\Psi(x)n(t)$. It can be determined after all other cut-off functions in the proof of Theorem 5 are introduced.

For $A_k(\tau,\varphi)$, note that $\partial_t\overline{\eta}(t,x)$ is only nonzero for $t\in\left(J_{\overline{\eta}}\right)^c$ (i.e., away from where $l(t)\equiv 1$), while $s\in I_{\overline{\psi}}\Subset J_{\overline{\eta}}$ because of $w(s)$. Hence for small τ, more precisely, for $\tau<\min\left\{d(I_{\overline{\psi}},\,\left(J_{\overline{\eta}}\right)^c)/2,\;1/2\right\}=:c_0$,

$$\partial_t\overline{\eta}(t,x)v_{k,\tau}(s,t,x)\equiv 0.$$

So $A_k(\tau,\varphi)=0$ for $0<\tau<c_0$. For $\tau\ge c_0$, first note that for any $s,t\in I$,

$$\begin{aligned}&\left\|\partial_s^k(w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)\right\|_{L^2(X)}\\&\le 3^k\|w\|_{C^k(\mathbb{R})}\max_{0\le a\le k}\sup_{\tau<s-t<2\tau}\left|\partial_s^a\left(\frac{s-t}{\tau^2}\rho\left(\frac{s-t}{\tau}\right)\right)\right|\times\\&\qquad\max_{0\le b\le k}\sup_{\tau<s-t<2\tau}\left\|\partial_s^b H_{s-t}\right\|_{L^2(X)\to L^2(X)}\left\|\psi\varphi^s\right\|_{L^2(X)}.\end{aligned}$$

Here $\|\partial_s^b H_{s-t}\|_{L^2(X)\to L^2(X)}=\|P^b H_{s-t}\|_{L^2(X)\to L^2(X)}\le(b/e(s-t))^b\le(b/e\tau)^b$ since $\tau<s-t<2\tau$. Direct computation shows that for any $0\le a\le k$,

$$\partial_s^a\left(\frac{s-t}{\tau^2}\rho\left(\frac{s-t}{\tau}\right)\right)=\frac{a}{\tau^{a+1}}\,\rho^{(a-1)}\left(\frac{s-t}{\tau}\right)+\frac{s-t}{\tau^{a+2}}\,\rho^{(a)}\left(\frac{s-t}{\tau}\right),$$

which is bounded above by $(a+2)\tau^{-(a+1)}\|\rho\|_{C^k(\mathbb{R})}$. So there exists some $C(k,\overline{\psi},\rho)>0$, such that

$$\left\|\partial_s^k(w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)\right\|_{L^2(X)} \leq \frac{C(k,\overline{\psi},\rho)}{\tau^{2k+1}}\left\|\varphi^s\right\|_{L^2(X)}.$$

It follows that

$$\begin{aligned}
|A_k(\tau,\varphi)| &= \left|\int_I\int_I\int_X u^t\,\partial_t\overline{\eta}^t\,\partial_s^k(w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)\,dmdtds\right| \\
&\leq \int_I\int_I \left\|u^t\partial_t\overline{\eta}^t\right\|_{L^2(X)}\left\|\partial_s^k(w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)\right\|_{L^2(X)}dtds \\
&\leq \frac{C(k,\overline{\psi},\rho)\,\|\partial_t\overline{\eta}\|_{L^\infty(I\times X)}}{\tau^{2k+1}}\int_{I_{\overline{\psi}}}\left\|\varphi^s\right\|_{L^2(X)}ds\int_{I_{\overline{\eta}}}\left\|u^t\right\|_{L^2(U_{\overline{\eta}})}dt \\
&\leq \widetilde{C}(k,\overline{\eta},\overline{\psi},\rho)\,\|\varphi\|_{L^2(I\times X)}\left\|\overline{\Psi}u\right\|_{L^2(I\times X)}.
\end{aligned}$$

Here the constant $\widetilde{C}(k,\overline{\eta},\overline{\psi},\rho) := |I_{\overline{\eta}}|^{1/2}|I_{\overline{\psi}}|^{1/2}C(k,\overline{\psi},\rho)\,\|\partial_t\overline{\eta}\|_{L^\infty(I\times X)}\,c_0^{-(2k+1)}$ depends only on the two cut-off functions $\overline{\eta}$, $\overline{\psi}$, the function ρ, and the sum of the trinomial coefficients that is bounded by 3^k. Note that c_0 is determined by the cut-off functions since $c_0=\min\left\{d(I_{\overline{\psi}},\,(J_{\overline{\eta}})^c)/2,\ 1/2\right\}$. The function $\overline{\Psi}$ is equal to 1 on the support of $\overline{\eta}$ as introduced at the beginning of this subsection. So

$$C_A(n,\overline{\eta},\overline{\psi},\rho) := \max_{0\leq k\leq n}\widetilde{C}(k,\overline{\eta},\overline{\psi},\rho)<\infty.$$

Recall that we take supremum over functions φ with $\|\varphi\|_{L^2(I\times X)}\leq 1$. Hence

$$\max_{0\leq k\leq n}\sup_{0<\tau<1}\sup_{\varphi\in\mathcal{T}}|A_k(\tau,\varphi)|\leq C_A(n,\overline{\eta},\overline{\psi},\rho)\left\|\overline{\Psi}u\right\|_{L^2(I\times X)}. \tag{5.3}$$

For $B_k(\tau,\varphi)$, because $\overline{\eta}(t,y)=l(t)\eta(y)$ and $\eta\equiv 1$ on $V_{\overline{\eta}}$, by the strong locality of the energy measure $d\Gamma$, the two terms in $B_k(\tau,\varphi)$ satisfy that

$$1_{V_{\overline{\eta}}}\,d\Gamma(\overline{\eta}^t u^t,\ v_{k,\tau}^{s,t}) = 1_{V_{\overline{\eta}}}\,d\Gamma(u^t,\ \overline{\eta}^t v_{k,\tau}^{s,t}).$$

In other words,

$$d\Gamma(\overline{\eta}^t u^t,\ v_{k,\tau}^{s,t}) - d\Gamma(u^t,\ \overline{\eta}^t v_{k,\tau}^{s,t}) = d\Gamma(\overline{\eta}^t u^t,\ \Phi v_{k,\tau}^{s,t}) - d\Gamma(u^t,\ \Phi\overline{\eta}^t v_{k,\tau}^{s,t}) \tag{5.4}$$

for any "bowl-shaped" Φ that equals 0 inside $V_{\overline{\eta}}$ and becomes 1 before it reaches the boundary of $V_{\overline{\eta}}$, provided that the products of the functions are still in the domain $\mathcal{F}$. To later utilize the L^2 Gaussian type upper bound to estimate, we take Φ to be a nice cut-off function "disjointly supported" from ψ. More precisely, recall that $V_{\overline{\psi}}\Subset U_{\overline{\psi}}\Subset V_{\overline{\eta}}\Subset U_{\overline{\eta}}$. Let V', U' be two open sets that sit in the middle of this chain, and let V'', U'' be two open sets at the right end of the chain, i.e.,

$$V_{\overline{\psi}}\Subset U_{\overline{\psi}}\Subset V'\Subset U'\Subset V_{\overline{\eta}}\Subset U_{\overline{\eta}}\Subset V''\Subset U''\Subset U.$$

Let $V_\Phi := V''\setminus\overline{U'}$ and $U_\Phi := U''\setminus\overline{V'}$. Then $V_\Phi\Subset U_\Phi$, and there exists a nice cut-off function that is 1 on V_Φ and supported in U_Φ. We fix such a function and denote it by Φ. The existence of Φ is guaranteed by Lemma 3, or we can take the difference of two nice cut-off functions (for pairs $V''\subset U''$ and $V'\subset U'$) and show that the difference still satisfies (3.1). The nice cut-off function Φ has the desired "bowl-shape", satisfies equation (5.4), and has disjoint support from ψ. In summary,

$$\begin{aligned}&|B_k(\tau,\varphi)|\\&=\left|-\int_I\int_I\int_X d\Gamma(\overline{\eta}^t u^t,\ \Phi v^{s,t}_{k,\tau})\,dtds+\int_I\int_I\int_X d\Gamma(u^t,\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\right|,\end{aligned}$$

where by the Cauchy-Schwartz inequality for the strongly local part of $\mathcal{E}$ and Hölder's inequality applied to the integrals on I, we have

$$\begin{aligned}&\left|\int_I\int_I\int_X d\Gamma(\overline{\eta}^t u^t,\ \Phi v^{s,t}_{k,\tau})\,dtds\right|\\&\quad\le\left(\int_{I_{\overline{\psi}}}\int_I\int_X d\Gamma(\overline{\eta}^t u^t,\ \overline{\eta}^t u^t)\,dtds\right)^{1/2}\left(\int_I\int_{I_{\overline{\eta}}}\int_X d\Gamma(\Phi v^{s,t}_{k,\tau},\ \Phi v^{s,t}_{k,\tau})\,dtds\right)^{1/2}\\&\quad\le\left(\left|I_{\overline{\psi}}\right|\int_I\mathcal{E}(\overline{\eta}^t u^t,\ \overline{\eta}^t u^t)\,dt\right)^{1/2}\left(\int_I\int_{I_{\overline{\eta}}}\mathcal{E}(\Phi v^{s,t}_{k,\tau},\ \Phi v^{s,t}_{k,\tau})\,dtds\right)^{1/2},\end{aligned}$$

and similarly (recall that $\overline{\Psi}$ equals 1 on the supports of all other nice cut-off functions),

$$\begin{aligned}&\left|\int_I\int_I\int_X d\Gamma(u^t,\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\right|=\left|\int_I\int_I\int_X d\Gamma(\overline{\Psi}^t u^t,\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\right|\\&\quad\le\left(\left|I_{\overline{\psi}}\right|\int_I\mathcal{E}(\overline{\Psi}^t u^t,\ \overline{\Psi}^t u^t)\,dt\right)^{1/2}\left(\int_I\int_I\mathcal{E}(\Phi\overline{\eta}^t v^{s,t}_{k,\tau},\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\right)^{1/2}.\end{aligned}$$

Hence

$$\begin{aligned}|B_k(\tau,\varphi)|\le{}&\left|I_{\overline{\psi}}\right|^{1/2}\left(\|\overline{\eta}u\|_{L^2(I\to\mathcal{F})}+\left\|\overline{\Psi}u\right\|_{L^2(I\to\mathcal{F})}\right)\times\\&\left[\left(\int_I\int_{I_{\overline{\eta}}}\mathcal{E}(\Phi v^{s,t}_{k,\tau},\ \Phi v^{s,t}_{k,\tau})\,dtds\right)^{1/2}+\left(\int_I\int_I\mathcal{E}(\Phi\overline{\eta}^t v^{s,t}_{k,\tau},\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\right)^{1/2}\right].\end{aligned}\tag{5.5}$$

It remains to estimate the two integrals in the square bracket. The estimates for the two terms are almost identical, so here we only do it for the second term, $\left(\int_I\int_I\mathcal{E}(\Phi\overline{\eta}^t v^{s,t}_{k,\tau},\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\right)^{1/2}$. Recall that $v^{s,t}_{k,\tau}\in L^2(I^2\to\mathcal{D}(P))$, we first want to move both $\Phi\overline{\eta}$ to one side in $\mathcal{E}(\cdot,\cdot)$, in order to rewrite the $\mathcal{E}$-integral as an L^2-integral of something times $Pv^{s,t}_{k,\tau}$. To this end we apply the gradient inequality. Using (3.8) applied to the nice cut-off function $\Phi\eta$, we get that

$$\begin{aligned}\int_I\int_I\mathcal{E}(\Phi\overline{\eta}^t v^{s,t}_{k,\tau},\ \Phi\overline{\eta}^t v^{s,t}_{k,\tau})\,dtds\ \le{}&2\int_I\int_I\left|\int_X(\Phi\overline{\eta}^t)^2 v^{s,t}_{k,\tau}\,Pv^{s,t}_{k,\tau}\,dm\right|dtds\\&+2C_2\int_I\int_{I_{\overline{\eta}}}\int_{\mathrm{supp}\{\Phi\eta\}}(v^{s,t}_{k,\tau})^2\,dmdtds.\end{aligned}$$

Here C_2 is associated with $\Phi\eta$. The first integral is (plugging in (5.1))

$$\begin{aligned}&2\int_I\int_I\left|\int_X(\Phi\overline{\eta}^t)^2 v^{s,t}_{k,\tau}\,Pv^{s,t}_{k,\tau}\,dm\right|dtds\\&\quad=2\int_I\int_I\left|\int_X(\Phi\overline{\eta}^t)^2 v^{s,t}_{k,\tau}\,\partial_s^k(w(s)\bar{\rho}_\tau(s-t)PH_{s-t})(\psi\varphi^s)(x)\,dm\right|dtds.\end{aligned}$$

Up to derivatives of $w(s)\bar{\rho}_\tau(s-t)$ which are bounded by some multiple of $1/(s-t)^{k+1}$, this integral is bounded above by a sum of integrals of the form

$$\int_I \int_I \left| \left\langle (\Phi \overline{\eta}^t)^2 v_{k,\tau}^{s,t},\ \partial_s^m H_{s-t}(\psi \varphi^s) \right\rangle_{L^2(X)} \right| dtds$$

for $1 \leq m \leq k+1$. Note that the functions $(\Phi \overline{\eta}^t)^2 v_{k,\tau}^{s,t}$ and $\psi \varphi^s$ have disjoint supports because Φ and ψ do. The second integral above equals

$$2C_2 \int_I \int_{I_{\overline{\eta}}} \int_X 1_{\mathrm{supp}\{\Phi\eta\}} v_{k,\tau}^{s,t} \times v_{k,\tau}^{s,k}\, dmdtds,$$

which is essentially the sum of

$$\int_I \int_{I_{\overline{\eta}}} \left| \left\langle 1_{\Phi\eta} v_{k,\tau}^{s,t},\ \partial_s^m H_{s-t}(\psi \varphi^s) \right\rangle_{L^2(X)} \right| dtds$$

for $0 \leq m \leq k$ (up to derivatives of $w(s)\bar{\rho}_\tau(s-t)$). Here, for simplicity, we write $1_{\Phi\eta} := 1_{\mathrm{supp}\{\Phi\eta\}}$. Again, by construction, Φ and ψ have disjoint supports, so the functions $1_{\Phi\eta} v_{k,\tau}^{s,t}$ and $\psi\varphi^s$ have disjoint supports. We can then use the L^2 Gaussian type upper bound to estimate each such term. Note also that as $\tau < s-t < 2\tau$, for any $0 \leq a \leq k$,

$$\left| \partial_s^a \left(\frac{s-t}{\tau^2} \rho \left(\frac{s-t}{\tau} \right) \right) \right| \leq \frac{a+2}{\tau^{a+1}} \|\rho\|_{C^k(\mathbb{R})} \leq \frac{2^{a+1}(a+2)}{(s-t)^{a+1}} \|\rho\|_{C^k(\mathbb{R})}.$$

In summary, we have

$$\begin{aligned}
&\int_I \int_I \mathcal{E}(\Phi \overline{\eta}^t v_{k,\tau}^{s,t},\ \Phi \overline{\eta}^t v_{k,\tau}^{s,t})\, dtds \\
&\leq 2(1+C_2) \times 3^k \|w\|_{C^k(\mathbb{R})} \times 2^{k+1}(k+2)\|\rho\|_{C^k(\mathbb{R})} \max_{\substack{1\leq a\leq k+1 \\ 0\leq m\leq k+1}} \int_I \int_{I_{\overline{\eta}}} \frac{1}{(s-t)^a} \times \\
&\quad \left(\left| \left\langle (\Phi \overline{\eta}^t)^2 v_{k,\tau}^{s,t},\ \partial_s^m H_{s-t}(\psi\varphi^s) \right\rangle_{L^2(X)} \right| + \left| \left\langle 1_{\Phi\eta} v_{k,\tau}^{s,t},\ \partial_s^m H_{s-t}(\psi\varphi^s) \right\rangle_{L^2(X)} \right| \right) dtds \\
&\leq 2^{k+2} 3^k (1+C_2)(k+2)\|w\|_{C^k(\mathbb{R})} \|\rho\|_{C^k(\mathbb{R})} \times \sup_{\tau<r<2\tau} G(k+1,k+1,r) \times \\
&\quad \int_I \int_{I_{\overline{\eta}}} \left(\left\| (\Phi\overline{\eta}^t)^2 v_{k,\tau}^{s,t} \right\|_{L^2(X)} + \left\| 1_{\Phi\eta} v_{k,\tau}^{s,t} \right\|_{L^2(X)} \right) \|\psi\varphi^s\|_{L^2(X)}\, dtds.
\end{aligned}$$

The last inequality is obtained by letting $r := s-t$ and applying the L^2 Gaussian type upper bound. Here $G(k+1,k+1,r) := G_{U_\Phi, U_\psi}(k+1,k+1,r)$ as defined in Assumption 4. So there exists some $C(k,w,\rho,C_2) > 0$, such that

$$\begin{aligned}
\int_I \int_I \mathcal{E}(\Phi \overline{\eta}^t v_{k,\tau}^{s,t},\ \Phi \overline{\eta}^t v_{k,\tau}^{s,t})\, dtds &\leq C(k,w,\rho,C_2) \sup_{\tau<r<2\tau} G(k+1,k+1,r) \times \\
&\left\{ \left\| (\Phi\overline{\eta})^2 v_{k,\tau} \right\|_{L^2(I\times I\times X)} + \left\| 1_{\Phi\eta} v_{k,\tau} \right\|_{L^2(I_{\overline{\eta}}\times I\times X)} \right\} \|\psi\varphi\|_{L^2(I_{\overline{\eta}}\times I\times X)}.
\end{aligned}$$

Next we estimate $\left\|1_{\Phi\eta}v_{k,\tau}\right\|_{L^2(I_{\overline{\eta}}\times I\times X)}$ (and $\left\|(\Phi\overline{\eta})^2 v_{k,\tau}\right\|_{L^2(I\times I\times X)}$). Note that

$$\begin{aligned}
\left\|1_{\Phi\eta}v_{k,\tau}\right\|^2_{L^2(I_{\overline{\eta}}\times I\times X)} &= \int_{I_{\overline{\eta}}}\int_I \left\langle 1_{\Phi\eta}v_{k,\tau},\ v_{k,\tau}\right\rangle_{L^2(X)}\,dsdt \\
&= \int_{I_{\overline{\eta}}}\int_I\int_X 1_{\Phi\eta}v_{k,\tau}(s,t,x)\times\partial_s^k(w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)(x)\,dmdsdt \\
&\le \widetilde{C}(k,w,\rho)\sup_{\tau<r<2\tau}G(k+1,k,r)\times\left\|1_{\Phi\eta}v_{k,\tau}\right\|_{L^2(I_{\overline{\eta}}\times I\times X)}\|\psi\varphi\|_{L^2(I\times X)}\left|I_{\overline{\eta}}\right|^{1/2}
\end{aligned}$$

for some $\widetilde{C}(k,w,\rho)>0$. The left-hand side $\left\|1_{\Phi\eta}v_{k,\tau}\right\|^2_{L^2(I_{\overline{\eta}}\times I\times X)}$ and the right-hand side of the inequality have a common factor $\left\|1_{\Phi\eta}v_{k,\tau}\right\|_{L^2(I_{\overline{\eta}}\times I\times X)}$ that cancel each other. The same argument works for $\left\|(\Phi\overline{\eta})^2 v_{k,\tau}\right\|_{L^2(I\times I\times X)}$ (as $0\le\Phi\overline{\eta}\le 1$). So

$$\begin{aligned}
&\left\|(\Phi\overline{\eta})^2 v_{k,\tau}\right\|_{L^2(I\times I\times X)}+\left\|1_{\Phi\eta}v_{k,\tau}\right\|_{L^2(I_{\overline{\eta}}\times I\times X)} \\
&\le 2\widetilde{C}(k,w,\rho)\sup_{\tau<r<2\tau}G(k+1,k,r)\times\|\psi\varphi\|_{L^2(I\times X)}\left|I_{\overline{\eta}}\right|^{1/2}.
\end{aligned}$$

Combining the two estimates above gives

$$\begin{aligned}
&\int_I\int_I\mathcal{E}(\Phi\overline{\eta}^t v_{k,\tau}^{s,t},\ \Phi\overline{\eta}^t v_{k,\tau}^{s,t})\,dtds\le C(k,w,\rho,C_2)\times \\
&2\widetilde{C}(k,w,\rho)\|\psi\|^2_{L^\infty(X)}\left|I_{\overline{\eta}}\right|\times\sup_{\tau<r<2\tau}G(k+1,k+1,r)^2\times\|\varphi\|^2_{L^2(I\times X)}.
\end{aligned}\tag{5.6}$$

By Assumption 4, $\sup_{0<r<2}G_{U_\Phi,U_\psi}(k+1,k+1,r)<\infty$.

Recall that C_2 here is the constant associated with $\Phi\eta$. To include the upper bound for the other term in (5.5), $\int_I\int_{I_{\overline{\eta}}}\mathcal{E}(\Phi v_{k,\tau}^{s,t},\ \Phi v_{k,\tau}^{s,t})\,dtds$, in the next few lines we write $C_2(\Phi\eta)$ and $C_2(\Phi)$ for the C_2 constants associated with the nice cut-off functions $\Phi\eta$ and Φ, respectively. Let

$$\begin{aligned}
C_B(n,\overline{\eta},\overline{\psi},\rho,\Phi) &:= \max_{0\le k\le n}\left[C(k,w,\rho,C_2(\Phi\eta))^{1/2}+C(k,w,\rho,C_2(\Phi))^{1/2}\right]\times \\
&(2\widetilde{C}(k,w,\rho))^{1/2}\|\psi\|_{L^\infty(X)}\left|I_{\overline{\eta}}\right|^{1/2}\times\sup_{0<r<2}G(k+1,k+1,r)\times\left|I_{\overline{\psi}}\right|^{1/2}.
\end{aligned}$$

Then C_B is finite, and we obtain the estimate for $B_k(\tau,\varphi)$ from (5.5) and (5.6)

$$\begin{aligned}
&\max_{0\le k\le n}\sup_{0<\tau<1}\sup_{\varphi\in\mathcal{T}}|B_k(\tau,\varphi)| \\
&\le C_B(n,\overline{\eta},\overline{\psi},\rho,\Phi)\left(\|\overline{\eta}u\|_{L^2(I\to\mathcal{F})}+\left\|\overline{\Psi}u\right\|_{L^2(I\to\mathcal{F})}\right).
\end{aligned}\tag{5.7}$$

Last, we estimate the term $C_k(\tau,\varphi)$. The idea is to first use the product rule for differentiation in time (∂_s) to expand and rewrite

$$\begin{aligned}
v_{k,\tau}^{s,t} &= \partial_s^k(w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s) \\
&= \sum_{a=0}^k\binom{k}{a}\partial_s^{k-a}w(s)\times\partial_s^a(\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s) \\
&= \sum_{a=0}^k\binom{k}{a}\partial_s^{k-a}w(s)\times(-1)^a\partial_t^a(\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s),
\end{aligned}$$

then move all the ∂_t^a on $\bar{\rho}_\tau(s-t)H_{s-t}$, $0 \le a \le k$, to $\overline{\eta} f$, using integration by parts. More precisely,

$$\begin{aligned}
&|C_k(\tau,\varphi)| \\
&= \left| \int_I \int_I \int_X f(t,x)\overline{\eta}(t,x)\, \partial_s^k (w(s)\bar{\rho}_\tau(s-t)H_{s-t})(\psi\varphi^s)(x)\, dmdtds \right| \\
&= \left| \sum_{a=0}^k \binom{k}{a} \int_I \partial_s^{k-a} w(s) \left\langle \partial_t^a(\overline{\eta}^t f^t),\ \bar{\rho}_\tau(s-t)H_{s-t}(\psi\varphi^s)\right\rangle_{L^2(I\times X)} ds \right| \\
&\le 2^k \|w\|_{C^k(\mathbb{R})} \max_{0\le a\le k} \left| \int_I \int_I \int_X \partial_t^a(\overline{\eta}^t f^t)\, \bar{\rho}_\tau(s-t)H_{s-t}(\psi\varphi^s)\, dmdtds \right|.
\end{aligned}$$

For any $0 \le a \le k$, note that by Hölder's inequality and Minkowski's inequality,

$$\begin{aligned}
&\left| \int_I \int_X \partial_t^a(\overline{\eta}^t f^t) \int_I \bar{\rho}_\tau(s-t)H_{s-t}(\psi\varphi^s)\, ds\, dmdt \right| \\
&\le \int_I \left\| \partial_t^a(\overline{\eta}^t f^t)\right\|_{L^2(X)} \left\| \int_I \bar{\rho}_\tau(s-t)H_{s-t}(\psi\varphi^s)\, ds \right\|_{L^2(X)} dt \\
&\le \int_I \left\| \partial_t^a(\overline{\eta}^t f^t)\right\|_{L^2(X)} \int_I \bar{\rho}_\tau(s-t) \left\| H_{s-t}(\psi\varphi^s)\right\|_{L^2(X)} ds\, dt.
\end{aligned}$$

To further estimate, we apply Hölder's inequality to the integral in t. First,

$$\max_{0\le a\le k} \left(\int_I \left\| \partial_t^a(\overline{\eta}^t f^t)\right\|_{L^2(X)}^2 dt \right)^{1/2} \le \|\overline{\eta} f\|_{W^{k,2}(I\to L^2(X))}.$$

Second, by Jensen's inequality and the fact that H_t is a contraction semigroup on $L^2(X)$,

$$\begin{aligned}
&\left(\int_I \left(\int_I \bar{\rho}_\tau(s-t) \left\| H_{s-t}(\psi\varphi^s)\right\|_{L^2(X)} ds \right)^2 dt \right)^{1/2} \\
&\le \left(\int_I r(t) \int_I \bar{\rho}_\tau(s-t) \left\| H_{s-t}(\psi\varphi^s)\right\|_{L^2(X)}^2 ds\, dt \right)^{1/2} \\
&\le \left(2 \sup_{s\in I} \int_I \bar{\rho}_\tau(s-t)\, dt \times \int_I \left\| \psi\varphi^s\right\|_{L^2(X)}^2 ds \right)^{1/2} \le 2\, \|\psi\varphi\|_{L^2(I\times X)},
\end{aligned}$$

where $r(t) = \int_I \bar{\rho}_\tau(s-t)\, ds \le 2$. See (4.2) and the paragraph below it for more detailed computations. Altogether we get that

$$\begin{aligned}
&|C_k(\tau,\varphi)| \\
&\le 2^k\, \|w\|_{C^k(\mathbb{R})} \max_{0\le a\le k} \left| \int_I \int_X \partial_t^a(\overline{\eta}^t f^t) \int_I \bar{\rho}_\tau(s-t)H_{s-t}(\psi\varphi^s)\, ds\, dmdt \right| \\
&\le 2^{k+1}\, \|w\|_{C^k(\mathbb{R})}\, \|\overline{\eta} f\|_{W^{k,2}(I\to L^2(X))}\, \|\psi\varphi\|_{L^2(I\times X)}.
\end{aligned}$$

Hence

$$\max_{0\le k\le n} \sup_{0<\tau<1} \sup_{\varphi\in\mathcal{T}} |C_k(\tau,\varphi)| \le C_C(n,\overline{\psi})\ \|\overline{\eta} f\|_{W^{n,2}(I\to L^2(X))}, \tag{5.8}$$

where $C_C(n,\overline{\psi}) := 2^{n+1}\, \|w\|_{C^k(\mathbb{R})}\, \|\psi\|_{L^\infty(X)} = 2^{n+1}\, \|w\|_{C^k(\mathbb{R})}$.

In the above estimates for A_k, B_k, C_k, we kept terms like $\|\overline{\eta} f\|_{W^{k,2}(I\to L^2(X))}$ and $\left\|\overline{\Psi} u\right\|_{L^2(I\to\mathcal{F})}$, since u, f are only assumed to be locally in those function spaces. If we take any proper representative $u^\sharp$, $f^\sharp$, we can bound those norms by the corresponding norms of $u^\sharp$ and $f^\sharp$.

Combining the estimates (5.3), (5.7), and (5.8) for $A_k(\tau,\varphi)$, $B_k(\tau,\varphi)$, and $C_k(\tau,\varphi)$, we complete the proof of Proposition 2. To finish with the proof of $\{\overline{\psi}\widetilde{u}_\tau\}_{\tau>0}$ being Cauchy in $W^{n,2}(I\to\mathcal{F})$, we still need to prove Proposition 3.

5.4 Proof of Proposition 3

Our aim is to show that for $0\le k\le n$, for any $0<\tau<1$,

$$\int_I \mathcal{E}(\partial_\tau\partial_s^k(\overline{\psi}\widetilde{u}_\tau),\ \partial_\tau\partial_s^k(\overline{\psi}\widetilde{u}_\tau))\,ds \lesssim \frac{1}{\tau}. \tag{5.9}$$

By the gradient inequality (3.8),

$$\begin{aligned}
&\int_I \mathcal{E}(\partial_\tau\partial_s^k(\overline{\psi}\widetilde{u}_\tau),\ \partial_\tau\partial_s^k(\overline{\psi}\widetilde{u}_\tau))\,ds\\
&\le 2\int_I \mathcal{E}(\psi^2\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s),\ \partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s))\,ds\\
&\quad +2C_2\int_I\int_{\mathrm{supp}\{\psi\}}(\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s))^2\,dmds.
\end{aligned} \tag{5.10}$$

Here C_2 is associated with ψ, we write it as $C_2(\psi)$ sometimes to specify its dependence. The proof of Proposition 2 implies that the second term is bounded, i.e.,

$$\begin{aligned}
&\max_{0\le k\le n}\sup_{0<\tau<1} C_2\int_I\int_{\mathrm{supp}\{\psi\}}(\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s))^2\,dmds\\
&\le \sup_{0<\tau<1} C_2\,\|\partial_\tau(w\widetilde{u}_\tau)\|^2_{W^{n,2}(I\to L^2(U_{\overline{\psi}}))}\le M_1
\end{aligned}$$

for some constant M_1. More precisely, we may fix some nice cut-off function $\mathfrak{p}$ in between ψ and η, i.e., $\mathfrak{p}\equiv 1$ on $V_{\mathfrak{p}}$, $\mathrm{supp}\{\mathfrak{p}\}\subset U_{\mathfrak{p}}$, where $U_{\overline{\psi}}\Subset V_{\mathfrak{p}}\Subset U_{\mathfrak{p}}\Subset V_{\overline{\eta}}$. Then

$$\int_I\int_{\mathrm{supp}\{\psi\}}(\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s))^2\,dmds\le\int_I\int_X \mathfrak{p}^2(\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s))^2\,dmds.$$

Take the product $\overline{\mathfrak{p}}(s,x):=\mathfrak{p}(x)w(s)$ to replace $\overline{\psi}$ and run the estimates in the proof of Proposition 2, we get that the right-hand side (taking supremum in $0<\tau<1$ and $0\le k\le n$) is bounded from above by some $\widetilde{M}_1=\widetilde{M}_1(n,\rho,\overline{\eta},\mathfrak{p},w,\overline{\Psi},\Phi,u,f)$. So the M_1 above can be taken as $M_1=C_2(\psi)\widetilde{M}_1$.

Next we estimate the first term in (5.10). Because $\widetilde{u}_\tau\in C^\infty(I\to\mathcal{D}(P))$,

$$\begin{aligned}
&\int_I \mathcal{E}(\psi^2\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s),\ \partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s))\,ds\\
&=\int_I\int_X \psi^2\partial_\tau\partial_s^k(w(s)\widetilde{u}_\tau^s)\times\partial_\tau\partial_s^k(w(s)P\widetilde{u}_\tau^s)\,ds\\
&\le \left\|\partial_\tau\partial_s^k(\overline{\psi}\widetilde{u}_\tau)\right\|_{L^2(I\times X)}\left\|\partial_\tau\partial_s^k(\overline{\psi}P\widetilde{u}_\tau)\right\|_{L^2(I\times X)}.
\end{aligned}$$

The first L^2 norm is exactly the quantity treated in Proposition 2, it is bounded above uniformly in $0<\tau<1$. Thus, (5.9) follows after we show that the second L^2 norm,

$$\left\|\partial_\tau\partial_s^k(\overline{\psi}P\widetilde{u}_\tau)\right\|_{L^2(I\times X)}\lesssim\frac{1}{\tau}.$$

Replace $\widetilde{u}_\tau$ by $P\widetilde{u}_\tau$ in the proof of Proposition 2. By the same arguments, $\left\|\partial_\tau \partial_s^k(\overline{\psi} P\widetilde{u}_\tau)\right\|_{L^2(I\times X)}$ breaks into three parts $A'_k(\tau,\varphi)$, $B'_k(\tau,\varphi)$, $C'_k(\tau,\varphi)$, and the estimates for A'_k and B'_k look almost identical to that for A_k and B_k. We write about the estimate for $C'_k(\tau,\varphi)$ here. The only difference is that instead of using $\|H_{s-t}\|_{L^2(X)\to L^2(X)} \le 1$ as in the estimate for C_k, here we use $\|PH_{s-t}\|_{L^2(X)\to L^2(X)} \le 1/e(s-t) \le 1/e\tau$ for $\tau < s-t < 2\tau$.

$$\begin{aligned}
&C'_k(\tau,\varphi)\\
&= \int_I\int_I\int_X f(t,x)\overline{\eta}(t,x)\, \partial_s^k(w(s)\bar{\rho}_\tau(s-t)PH_{s-t})(\psi\varphi^s)(x)\, dm(x)dtds.
\end{aligned}$$

As in the estimate for C_k, the estimate for C'_k comes down to estimate

$$\begin{aligned}
&\max_{0\le a\le k}\left|\int_I\int_X \partial_t^a(\overline{\eta}^t f^t)\int_I \bar{\rho}_\tau(s-t)PH_{s-t}(\psi\varphi^s)\, ds\, dmdt\right|\\
&\le \|\overline{\eta} f\|_{W^{k,2}(I\to L^2(X))}\left[2\int_I\int_I \bar{\rho}_\tau(s-t)\left\|PH_{s-t}(\psi\varphi^s)\right\|^2_{L^2(X)} dsdt\right]^{1/2}\\
&\le \|\overline{\eta} f\|_{W^{k,2}(I\to L^2(X))}\sup_{s\in I}\left\{2\int_I \bar{\rho}_\tau(s-t)\, dt\right\}^{1/2}\left(\int_I \frac{1}{(e\tau)^2}\left\|\psi\varphi^s\right\|^2_{L^2(X)} ds\right)^{1/2}\\
&\le \frac{2}{e\tau}\|\overline{\eta} f\|_{W^{k,2}(I\to L^2(X))}\|\psi\varphi\|_{L^2(I\times X)}.
\end{aligned}$$

Hence indeed for any $0\le k\le n$, $0<\tau<1$,

$$\begin{aligned}
\left\|\partial_\tau\partial_s^k(\overline{\psi} P\widetilde{u}_\tau)\right\|_{L^2(I\times X)} &= \sup_{\substack{\|\varphi\|_{L^2(I\times X)}\le 1\\ \varphi\in C_c^\infty(I\to L^2(X))}}\left\langle \psi\,\partial_\tau\partial_s^k(w(s)P\widetilde{u}_\tau),\ \varphi\right\rangle_{L^2(I\times X)}\\
&\le \sup_{\varphi\in\mathcal{T}}\left|A'_k(\tau,\varphi)\right| + \sup_{\varphi\in\mathcal{T}}\left|B'_k(\tau,\varphi)\right| + \sup_{\varphi\in\mathcal{T}}\left|C'_k(\tau,\varphi)\right| \le \frac{M_2}{\tau},
\end{aligned}$$

where $M_2 = M_2(n,\rho,\overline{\eta},\overline{\psi},\overline{\Psi},\Phi,u,f) < \infty$.

5.5 Convergence of the Approximate Sequence in L^2 Sense

Proposition 2 and Proposition 3 together imply that the approximate sequence $\{\overline{\psi}\widetilde{u}_\tau\}_{\tau>0}$ is Cauchy in $W^{n,2}(I\to\mathcal{F})$ (recall that by this we mean any subsequence $\{\overline{\psi}\widetilde{u}_{\tau_j}\}_{j\in\mathbb{N}_+}$ with $\tau_j\to 0$ is a Cauchy sequence). As we explained at the beginning of this section, to finish with the proof of Theorem 5, it remains to show that the approximate sequence converges to $\overline{\psi}u$ in some weak sense. We prove a slightly more general result (Proposition 4). In this proposition we treat the class of strongly continuous semigroups on $L^2(X)$, as roughly the same proof works. This is a larger class of semigroups as the semigroups not necessarily satisfy the Markov property or correspond to a Dirichlet form.

Let $(H_t)_{t>0}$ be a strongly continuous semigroup of bounded linear operators on $L^2(X)$. Then there exist some $M\ge 1$, $\omega\ge 0$, so that for any $t>0$,

$$\|H_t\|_{L^2(X)\to L^2(X)} \le Me^{\omega t}.$$

Let $I=(a,b)$ be an open interval, $-\infty\le a<b\le\infty$. For any function w in $L^2(I\times X)$, for any $s\in I$, $0<\tau<1$, define

$$A_\tau w(s,x) := \int_I \rho_\tau(s-t)H_{s-t}(w^t)(x)\, dt. \tag{5.11}$$

As before, the integrand is nonzero only when $t \in I \cap (s-2\tau, s-\tau)$ because of the ρ_τ term, and the integral is well-defined. To be more rigorous, A_τ is first defined on $C_c(I \to L^2(X))$, then extended to the whole $L^2(I \times X)$ under the fact that for any $w \in C_c(I \to L^2(X))$,

$$\sup_{0<\tau<1} \|A_\tau w\|_{L^2(I\times X)} \le C\|w\|_{L^2(I\times X)}$$

for some $C < \infty$. This operator norm bound can be verified using a combination of Minkowski's inequality and Jensen's inequality as we did in the estimates of $C_k(\tau, \varphi)$. When τ is not small enough, $A_\tau w$ is the zero function on $I \times X$. Similar to checking that $\widetilde{u}_\tau \in C^\infty(I \to \mathcal{F})$ for any $\tau > 0$, we can show that for any $\tau > 0$, $A_\tau w \in C^\infty(I \to \mathcal{F})$.

Proposition 4 *Let $(H_t)_{t>0}$ be a strongly continuous semigroup of bounded linear operators on $L^2(X)$. Then for any w in $L^2(I \times X)$, $A_\tau w$ defined as in (5.11) converges to w in $L^2(I \times X)$, as τ tends to 0.*

Proof We first show that for any w in $C_c(I \to L^2(X))$, $A_\tau w$ converges to w in $L^2(I \times X)$. Because $C_c(I \to L^2(X))$ is dense in $L^2(I \times X)$ and

$$\sup_{0<\tau<1} \|A_\tau\|_{L^2(I\times X)\to L^2(I\times X)} < +\infty,$$

the statement holds for all w in $L^2(I \times X)$. For $w \in C_c(I \to L^2(X))$,

$$\|A_\tau w - w\|_{L^2(I\times X)} \le \left\| \int_I \rho_\tau(\cdot - t)\left[H_{\cdot - t}(w^t) - w^{\cdot}\right] dt \right\|_{L^2(I\times X)} + \left\| \left(1 - \int_I \rho_\tau(\cdot - t)\, dt\right) w^{\cdot} \right\|_{L^2(I\times X)}.$$

When $a = -\infty$, or $a \neq -\infty$ and $s - a \ge 2\tau$, $(s-t)/\tau$ runs over the full $(1, 2)$ as t runs over I, so $1 - \int_I \rho_\tau(s-t)\, dt = 1 - 1 = 0$. So $1 - \int_I \rho_\tau(s-t)\, dt$ can only be nonzero when $a \neq -\infty$ and $a < s < a + 2\tau$, which is an interval of length 2τ. Because w is in $C_c(I \to L^2(X))$, we conclude that the second term

$$\left\| \left(1 - \int_I \rho_\tau(\cdot - t)\, dt\right) w^{\cdot} \right\|_{L^2(I\times X)} \to 0 \text{ as } \tau \to 0.$$

Next we estimate the first term. In the following, let $J \subset I$ be a bounded open interval such that $\operatorname{supp}\{w\} \subset J \times X$. Then for τ small enough, $w^{s-r} \equiv 0$ for all $s \in J^c$ and $0 < r < 2\tau$. That is, the interval where w is supported on is at distance larger than 2τ from J^c. The first term is the $L^2(I \times X)$ norm of

$$\begin{aligned}
&\int_I \rho_\tau(s-t)\left[H_{s-t}(w^t) - w^s\right] dt \\
&= \int_I \rho_\tau(s-t) H_{s-t}(w^t - w^s)\, dt + \int_I \rho_\tau(s-t)\left[H_{s-t}(w^s) - w^s\right] dt.
\end{aligned}$$

The L^2 norm of the first part, $\left\|\int_I \rho_\tau(s-t) H_{s-t}(w^t - w^s)\, dt\right\|_{L^2(I\times X)}$, is bounded from above by (let $r = s - t$ and apply Minkowski's inequality)

$$\begin{aligned}
&\int_\tau^{2\tau} \rho_\tau(r) \left\| H_r(w^{\cdot - r} - w^{\cdot})\right\|_{L^2(I\times X)} dr = \int_\tau^{2\tau} \rho_\tau \left\| \left\| H_r(w^{\cdot - r} - w^{\cdot})\right\|_{L^2(X)} \right\|_{L^2(I)} dr \\
&\le \int_\tau^{2\tau} \rho_\tau(r) \left\| M e^{\omega r} \left\| w^{\cdot - r} - w^{\cdot} \right\|_{L^2(X)} \right\|_{L^2(I)} dr \le C \sup_{\substack{s\in J \\ \tau<r<2\tau}} \left\| w^{s-r} - w^s \right\|_{L^2(X)},
\end{aligned}$$

which tends to 0 as $\tau \to 0$. The constant C can be taken as $Me^{2\omega}|J|^{1/2}$. By the same reasoning, the L^2 norm of the second part satisfies

$$\left\| \int_I \rho_\tau(s-t) \left[H_{s-t}(w^s) - w^s \right] dt \right\|_{L^2(I\times X)} \le C \sup_{\substack{s\in J \\ \tau<r<2\tau}} \left\| H_r(w^s) - w^s \right\|_{L^2(X)}.$$

The right-hand side tends to 0 as $\tau \to 0$ essentially because $\{s \mapsto H_r(w^s)\}_{r>0} \subset C_c(J \to L^2(X))$ is equicontinuous in s. The details are as follows. First note that for any fixed $r > 0$, any $s, t \in J$,

$$\begin{aligned} &\left\| H_r(w^s) - w^s \right\|_{L^2(X)} \\ &\le \left\| H_r(w^s - w^t) \right\|_{L^2(X)} + \left\| H_r(w^t) - w^t \right\|_{L^2(X)} + \left\| w^t - w^s \right\|_{L^2(X)} \\ &\le 2Me^{\omega r} \left\| w^t - w^s \right\|_{L^2(X)} + \left\| H_r(w^t) - w^t \right\|_{L^2(X)}. \end{aligned}$$

For any $\epsilon > 0$, any $s \in J$, there is some $\tau_0(s) > 0$ such that

(1) for any $r < \tau_0(s)$, $\|H_r(w^s) - w^s\|_{L^2(X)} < \epsilon$ (since $H_r(w^s) \to w^s$ in $L^2(X)$);
(2) $\left\| w^t - w^s \right\|_{L^2(X)} < \epsilon$, for any $|s - t| < \tau_0(s)$ (since $w \in C_c(J \to L^2(X))$).

Let $B(s; \tau_0(s)) := (s - \tau_0(s),\ s + \tau_0(s))$, $s \in J$. Because $\overline{J}$ is compact and $\{B(s; \tau_0(s))\}_{s\in J}$ covers $\overline{J}$, there exists some finite subcover $\{B(s_k; \tau_0(s_k))\}_{k=1}^N$. Hence there exists some fixed τ_0 ($\tau_0 = \min_{1\le k\le N}\{\tau_0(s_k)\}$) such that

(1) for any $r < \tau_0$, any s_k, $1 \le k \le N$, $\|H_r(w^{s_k}) - w^{s_k}\|_{L^2(X)} < \epsilon$;
(2) for any $s \in J$, there exists some s_k such that $s \in B(s_k; \tau_0(s_k))$. So $\|w^s - w^{s_k}\|_{L^2(X)} < \epsilon$.

Therefore,

$$\sup_{\substack{s\in J \\ \tau<r<2\tau}} \left\| H_r(w^s) - w^s \right\|_{L^2(X)} \to 0 \ \text{ as } \tau \to 0.$$

This completes the proof of Proposition 4. □

Note that for any local weak solution u in Theorem 5, the function $\widetilde{u}_\tau$ is exactly $A_\tau(\overline{\eta} u)$. So Proposition 4 applies to $\widetilde{u}_\tau$, and it follows that $\overline{\psi}\widetilde{u}_\tau \to \overline{\psi} u$ in $L^2(I \times X)$ as $\tau \to 0$. This completes the proof of Theorem 5.

5.6 Proof of Corollary 1

In this subsection we prove Corollary 1, which says essentially that time derivatives of local weak solutions of the heat equation are still local weak solutions.

Proof of Corollary 1 By Theorem 5, u belongs to $\mathcal{F}^n_{\text{loc}}(I \times U)$. For any test function φ in $\mathcal{F}_c(I \times U) \cap C_c^\infty(I \to \mathcal{F})$, $\partial_t^k \varphi$ for any $1 \le k \le n$ is also a test function. By definition of local weak solutions on $I \times U$, u satisfies

$$-\int_I \int_X u\, \partial_t^{k+1} \varphi \, dm dt + \int_I \mathcal{E}(u, \partial_t^k \varphi)\, dt = \int_I \int_X f \partial_t^k \varphi \, dm dt. \tag{5.12}$$

To show $\partial_t^k u$ is a local weak solution of the heat equation (4.1), intuitively it suffices to do integration by parts k times to move ∂_t^k to the u and f sides of the integrals. We now justify this procedure.

For the justification of integration by parts for the first and third integrals in (5.12), we only describe the first step and the remaining is clear by induction. By the Fubini-Tonelli theorem, suppose $\operatorname{supp}\{\varphi\} \subset J \times V \Subset I \times U$, since

$$\int_I \int_X \left| u\, \partial_t^{k+1} \varphi \right| dmdt \le \|u\|_{L^2(J\times V)} \|\varphi\|_{W^{k+1,2}(I\to L^2(U))} < \infty,$$

we can switch the order of integration and get

$$-\int_I \int_X u\, \partial_t^{k+1} \varphi\, dmdt = -\int_X \int_I u\, \partial_t^{k+1} \varphi\, dtdm = \int_X \int_I \partial_t u\, \partial_t^k \varphi\, dtdm,$$

where the second equality is by integration by parts and that φ is compactly supported in time. The same argument works for the integral

$$\int_I \int_X f \partial_t^k \varphi\, dmdt = -\int_X \int_I \partial_t f\, \partial_t^{k-1} \varphi\, dtdm.$$

For the second term in (5.12), let

$$\varphi_n(t, \cdot) := \frac{\varphi(t + 1/n, \cdot) - \varphi(t, \cdot)}{1/n},$$

then $\varphi_n \to \varphi$ in $C^\infty(I \to \mathcal{F})$. By the Cauchy-Schwartz inequality for $\mathcal{E}$,

$$\left| \int_I \mathcal{E}(u, \varphi_n - \partial_t \varphi)\, dt \right| \le \int_I (\mathcal{E}(u,u))^{1/2} (\mathcal{E}(\varphi_n - \partial_t\varphi,\ \varphi_n - \partial_t\varphi))^{1/2}\, dt \to 0$$

as $n \to \infty$. Here $\mathcal{E}(u,u)$ is understood as $\mathcal{E}(u^\sharp, u^\sharp)$ where $u^\sharp \in \mathcal{F}^1(I \times X) = W^{1,2}(I \to \mathcal{F})$ and agrees with u a.e. on a neighborhood of $\operatorname{supp}\{\varphi\}$. For n large enough,

$$\int_I \mathcal{E}(u, \varphi_n)\, dt = \int_I \mathcal{E}(u_n, \varphi)\, dt,$$

where $u_n(t, \cdot) := (u(t - 1/n, \cdot) - u(t, \cdot))\,/(1/n)$. More rigorously, u_n should be defined in terms of the $u^\sharp$ above. Then $u_n \to \partial_t u$ in $\mathcal{F}(I \times X) = L^2(I \to \mathcal{F})$ as $n \to \infty$. Passing to the limit then shows that

$$\int_I \mathcal{E}(u, \partial_t \varphi)\, dt = -\int_I \mathcal{E}(\partial_t u, \varphi)\, dt.$$

In summary, after k times of integration by parts, $1 \le k \le n$, (5.12) becomes

$$(-1)^{k+1} \int_I \int_X \partial_t^k u\, \partial_t \varphi\, dmdt + (-1)^k \int_I \mathcal{E}(\partial_t^k u, \varphi)\, dt = (-1)^k \int_I \int_X \partial_t^k f \varphi\, dmdt,$$

thus $\partial_t^k u$ is a local weak solution of (4.1) on $I \times U$. The statement in Corollary 1 for $f = 0$ then follows. □

6 Application to Ancient Solutions

Ancient solutions of a heat equation, also called ancient caloric functions, are the heat operator/equation version of "global harmonic functions". It is therefore interesting from a variety of points of view to study the structure of various linear spaces of such functions, especially spaces defined by particular growth conditions. In this section, we generalize the results in [9, 38] on the structures of ancient caloric functions with certain growth types. The results of [9, 38] are set in the setting of smooth Riemannian manifolds where solutions are smooth and time derivatives of solutions are automatically solutions themselves, and this fact is a key ingredient in the proof. Therefore, the extension to the setting of ancient local weak solutions in Dirichlet spaces (with appropriate properties) presented here provides a good illustration of the usefulness of the property that time derivatives of local weak solutions are local weak solutions themselves (Corollary 1). For instance, in $\mathbb{R}^n$, Corollary 1 allows us to show that polynomial growth ancient local weak solutions of the heat equation associated with divergence form uniformly elliptic operators with bounded measurable coefficients, admit structural properties similar to those proved in [9] for classical ancient caloric functions.

6.1 Statement of Results

As before we assume that $(\mathcal{E}, \mathcal{F})$ is a symmetric regular local Dirichlet form. We call a local weak solution u of $(\partial_t + P)u = 0$ on $(-\infty, b) \times X$ for some $b > 0$ an *ancient local weak solution*, or *ancient solution* for short. We assume $(X, m, \mathcal{E}, \mathcal{F})$ satisfies the assumption on existence of nice cut-off functions (Assumption 3), and the following further assumption. Recall that $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$.

Assumption 6 For any precompact open set $V \Subset X$, any $C_1 > 0$, any $n \in \mathbb{N}_+$, there exist

(1) an exhaustion of X, $\{W^n_{V,i}\}_{i\in\mathbb{N}_+}$, with each set $W^n_{V,i}$ covering V. That is, $\{W^n_{V,i}\}_{i\in\mathbb{N}_+}$ is a sequence of increasing open sets, satisfying

$$V \subset W^n_{V,1},\ W^n_{V,i} \Subset W^n_{V,i+1},\ \bigcup_{i=1}^{\infty} W^n_{V,i} = X.$$

(2) a sequence of cut-off functions $\{\varphi^n_{V,i}\}_{i\in\mathbb{N}_+}$, satisfying that each $\varphi^n_{V,i} =: \varphi_i$ is a cut-off function for the pair $W^n_{V,i} \subset W^n_{V,i+1}$, i.e., $\varphi_i \in \mathcal{F} \cap C_c(X)$ and is in between 0 and 1, $\varphi_i = 1$ on $W^n_{V,i}$, $\operatorname{supp}\{\varphi_i\} \subset W^n_{V,i+1}$; φ_i further satisfies that for any $v \in \mathcal{F}$,

$$\int_X v^2\, d\Gamma(\varphi_i, \varphi_i) \le C_1 \int_X \varphi_i^2\, d\Gamma(v, v) + \frac{1}{n} \int_{\operatorname{supp}\{\varphi_i\}} v^2\, dm. \tag{6.1}$$

When X is compact and 1 belongs to $\mathcal{F}$, Assumption 6 trivially holds because we can take all $W^n_{V,i}$ to be the whole space X, and take all φ_i to be the constant function 1. For noncompact spaces, in the most classical setting $(\mathbb{R}^d, dx)$ with the Dirichlet form of the d-dimensional Brownian motion, if $V \subset B(0; R)$ where $B(0; R)$ stands for the ball of radius R centered at the origin, we can take $W^n_{V,i} = B(0; R + cin^{1/2})$ for some $c \ge 1$. It is standard to construct nice cut-off functions φ_i for each pair $W^n_{V,i} \subset W^n_{V,i+1}$ such that

$$d\Gamma(\varphi_i, \varphi_i) \le \frac{2}{c^2 n}\, dx,$$

which implies (6.1) with $C_1 = 0$. See also the end of Sections 6.1 and 7.1.

In the following theorems we consider two types of ancient solutions, one with polynomial L^2 growth bound, and the other with exponential L^2 growth bound. We first remark that for any ancient local weak solution u, by Theorem 5, u is locally in $W^{\infty,2}((-\infty, b) \to \mathcal{F}) \subset C^\infty((-\infty, b) \to \mathcal{F})$. As generalizations of results in [9, 38], we state the following two theorems on the structure of ancient solutions in the Dirichlet space setting.

Theorem 7 *Let (X, m) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form on X. Assume that the Dirichlet space $(X, m, \mathcal{E}, \mathcal{F})$ satisfies Assumptions 3 and 4, and when X is not compact, further satisfies Assumption 6. Let $(H_t)_{t>0}$ and $-P$ be* *the corresponding semigroup and generator. Let u be a local weak solution of $(\partial_t + P)u = 0$ on $(-\infty, b) \times X$ for some $b > 0$, i.e., u is an ancient solution of the heat equation. Suppose u satisfies the L^2 polynomial growth condition, namely, for any open subset $V \Subset X$, any $i \in \mathbb{N}_+$, there exist positive constants $b_u, d_u, C_{u,V,i} > 0$ (b_u, d_u are independent of V, i), such that for any $T > 1$, $n \in \mathbb{N}_+$,*

$$\left(\int_{[-T,\,0]\times W^n_{V,i}} |u(t,x)|^2 \, dmdt \right)^{1/2} \leq C_{u,V,i} \max\left\{ T^{d_u},\, n^{b_u} \right\}. \tag{6.2}$$

Then there exists some $N > 0$ such that for any $k > N$,

$$\partial_t^k u = 0.$$

More precisely, u is a polynomial in time, with

$$u(t,x) = u(0,x) + \partial_t u(0,x)\, t + \partial_t^2 u(0,x)\, \frac{1}{2!} t^2 + \cdots + \partial_t^N u(0,x) \frac{1}{N!} t^N.$$

Here $N = \lfloor d_u - \frac{1}{2} \rfloor$, the largest integer not exceeding $d_u - 1/2$.

For ancient solutions of the exponential growth type, we only need one sequence of exhaustion to get sufficient estimates, so we fix $n = 1$ and some precompact open set V_0, and consider the sequence $W^1_{V_0,i} =: W_i$ only.

Theorem 8 *Let (X, m) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form on X. Assume that the Dirichlet space $(X, m, \mathcal{E}, \mathcal{F})$ satisfies Assumptions 3 and 4, and when X is not compact, further satisfies Assumption 6. Let $(H_t)_{t>0}$ and $-P$ be the corresponding semigroup and generator. Let u be a local weak solution of $(\partial_t + P)u = 0$ on $(-\infty, b) \times X$ for some $b > 0$, i.e., u is an ancient solution of the heat equation. Suppose u satisfies the L^2 exponential growth condition, namely, there exists some $c_u > 0$, such that for any $T > 1$, any $i \in \mathbb{N}_+$,*

$$\int_{[-T,\,0]\times W_i} |u(t,x)|^2 \, dmdt \leq e^{c_u (T+i)}. \tag{6.3}$$

Then u is analytic in $t \in (-\infty, 0]$, in the sense that for any precompact open set $V \subset X$,

$$\left\| u(t,\cdot) - \sum\nolimits_{i=0}^{k} \frac{\partial_t^i u(0,\cdot)}{i!} t^i \right\|_{L^2(V)} \to 0 \text{ as } k \to \infty, \tag{6.4}$$

and the convergence is uniform in $t \in [a, 0]$ for any $a < 0$.

We first make some remarks about the two theorems.

Remark 12 In Theorem 7, if we denote $\partial_t^k u(0,x)/k! =: u_k(x)$ and let $N = \lfloor d_u - 1/2 \rfloor$, then $\{u_k\}_{k=0}^N$ satisfies

$$\begin{aligned} -Pu_k(x) &= (k+1)\, u_{k+1}(x), \quad \text{for } 0 \leq k \leq N-1, \\ -Pu_N(x) &= 0, \end{aligned}$$

both in the sense that for any $\varphi \in \mathcal{F}_c(X)$,

$$-\mathcal{E}(u_k,\, \varphi) = (k+1) \int_X u_{k+1} \varphi \, dm,$$

for any $1 \le k \le N$ ($u_{N+1} = 0$). In other words, u_k is a local weak solutions of $-Pu_k = (k+1)u_{k+1}$ on X. In addition, by [9, Corollary 0.5], each u_k is a linear combination of $u(t_i, \cdot)$, $i = 0, 1, \ldots, N$, where $-1 < t_0 < t_1 < \cdots < t_N \le 0$ are arbitrarily fixed numbers. It follows that all u_k's satisfy the L^2 growth bound that for any precompact open set $V \Subset X$, any $i, n \in \mathbb{N}_+$, there exists some constant $\widetilde{C}_{u,V,i} > 0$, such that

$$\left(\int_{W^n_{V,i}} |u_k(x)|^2 \, dm\right)^{1/2} \le \widetilde{C}_{u,V,i} \, n^{b_u}.$$

Remark 13 In Theorem 8, if we write $u(t,x) = \sum_{k=0}^{\infty} a_k(x) t^k / k!$ where the two sides equal in the above L^2 sense, then the $a_k(x)$ functions are $a_k(x) = \partial_t^k u(0, x)$. A Caccioppoli type estimate for local weak solutions can be derived from the proof of Proposition 6, namely, for any ancient local weak solution v of the heat equation $(\partial_t + P)v = 0$, for any $T > 0$, there exists some $K > 0$ such that

$$\sup_{t \in [-T,0]} \int_{W_i} |v(t,x)|^2 \, dm \le K \int_{[-T-1,0] \times W_{i+1}} |v(t,x)|^2 \, dm dt,$$

where $i \in \mathbb{N}_+$ and W_i is defined as in Theorem 8. For any k, by taking $T = 1$ and $v = \partial_t^k u(t,x)$ which by Corollary 1 is a local weak solution, and by using the inequality in Proposition 5 given in the next section, we get that $a_k(x)$ satisfies the L^2 upper bound that for some constant $C_k > 0$,

$$\int_{W_i} |a_k(x)|^2 \, dm \le C_k \, e^{c_u(i+5k+3)}.$$

Remark 14 The conclusion in Theorem 8 is in the L^2 sense. By Proposition 6, for any ancient (local weak) solution u, it is also true that the partial sum $\sum_{i=0}^{k} \partial_t^i u(0,x) t^i / i!$ tends to u in the energy integral over any precompact set as k tends to infinity, uniformly in time on any finite interval. If the (essential) supremum of u over each time-space cylinder can be controlled by the L^2 integral of u over some cylinder, then we can make the conclusion in Theorem 8 an (m-a.e.) pointwise conclusion. For example, some ultracontractivity property of the heat semigroup is sufficient for this purpose. See the companion paper [20].

As a corollary of Theorem 7, we recover in the current setting the dimension result in [9, Corollary 0.5] under an additional condition on the polynomial volume growth of the sets, $W^n_{V,i}$. Here $V \Subset X$ is an arbitrarily fixed open set. We first define the appropriate function spaces. For each $d, b \in \mathbb{N}_+$, let $\mathcal{P}_{d,b}(X)$ denote the vector space of all ancient (local weak) solutions u of $(\partial_t + P)u = 0$ satisfying that for any $i \in \mathbb{N}_+$, there exists some constant $B_{u,V,i} > 0$, such that

$$\operatorname*{ess\,sup}_{[-T,0] \times W^n_{V,i}} |u(t,x)| \le B_{u,V,i} \max\left\{T^d, n^b\right\}. \tag{6.5}$$

Let $\mathcal{H}_b(X)$ denote the vector space of all local weak solutions v of $Pv = 0$ on X with polynomial growth bound with exponent b, that is, for any $i \in \mathbb{N}_+$, there exists some constant $D_{v,V,i} > 0$, such that

$$\operatorname*{ess\,sup}_{x \in W^n_{V,i}} |v(x)| \le D_{v,V,i} \, n^b.$$

Corollary 2 *Under the hypotheses of Theorem 7, and the additional assumption that for some precompact open set $V \Subset X$, for any $n, i \in \mathbb{N}_+$, the sets $W^n_{V,i}$ satisfy some polynomial volume growth bound*

$$m(W^n_{V,i}) \le E_{V,i}\, n^a \tag{6.6}$$

where $E_{V,i}, a > 0$ are constants. Then

$$\dim \mathcal{P}_{d,b}(X) \le (d+1) \dim \mathcal{H}_b(X).$$

Proof Take any $u \in \mathcal{P}_{d,b}(X)$. Note that (6.5) and (6.6) together imply the L^2 growth condition (6.2) for some d_u, b_u (e.g., $2d_u = 2d + 1$, $2b_u = 2b + a$). Hence by Theorem 7, u is a polynomial in time with $\partial_t^k u = 0$ for k large enough. The growth condition (6.5) then implies that $\partial_t^k u = 0$ for $k > d$. As in Remark 12, let $u_k = (\partial_t^k u(0, x))/k!$. As shown in the proof of [9, Corollary 0.5], for any fixed $t_0, t_1, \ldots, t_d \in (-1, 0]$ that are distinct, there exist numbers b_j^k such that for any $0 \le k \le d$,

$$u_k(x) = \sum_{j=0}^{d} b_j^k u(t_j, x).$$

Because all $|t_j| < 1$, and $u \in \mathcal{P}_{d,b}(X)$, for any $i, n \in \mathbb{N}_+$,

$$\operatorname*{ess\,sup}_{x \in W^n_{V,i}} |u_k(x)| \le (d+1) \max_{0 \le j \le d} \left| b_j^k \right| \times B_{u,V,i}\, n^b.$$

This implies that $u_k \in \mathcal{H}_b(X)$. By the same arguments as those in equation 1.21 through equation 1.23 in the proof of [9, Corollary 0.5], it follows that

$$\dim \mathcal{P}_{d,b}(X) \le (d+1) \dim \mathcal{H}_b(X).$$

$\square$

We make some final remarks about the two assumptions on existence of cut-off functions, Assumption 3 and Assumption 6.

First, Assumption 3 focuses on for any fixed pair of open sets $V \Subset U$, in particular they could be very close to each other, for any small $C_1 > 0$, the existence of a cut-off function for the pair $V \subset U$ that satisfies (3.1). There C_2 depends on C_1, U, V and is usually a large number when C_1 is small and U, V are close. Intuitively, the cut-off function is steep. In contrast, in Assumption 6, the focus is on for any fixed initial set $V \Subset X$ and fixed C_1, for small C_2 ($C_2 = 1/n$ for large n), the existence of an exhaustion and cut-off functions for each pair of adjacent open sets therein. Intuitively, for large n, the sets in the exhaustion are far apart, and the cut-off functions have flat shapes.

Regarding the validity of Assumption 6, we remark that in general Dirichlet spaces which have some notion of distance that interacts well with the energy measure, this assumption is satisfied. Roughly speaking, for large n, to find $W^n_{V,i}$'s and φ_i's, we just require $W^n_{V,i}$ and the complement of $W^n_{V,i+1}$ to be separated by a large enough distance. For example, consider a Dirichlet space $(X, m, \mathcal{E}, \mathcal{F})$ that admits "nice metric cut-off functions", namely, there exists some distance d that defines the same topology of X, such that for any pair of open sets $V \Subset U$, any $0 < C_1 < 1$, there exists some nice cut-off function φ satisfying that for any $v \in \mathcal{F}$,

$$\int_X v^2\, d\Gamma(\varphi, \varphi) \le C_1 \int_X \varphi^2\, d\Gamma(v, v) + C(C_1) d(V, U^c)^{-\beta} \int_{\mathrm{supp}\{\varphi\}} v^2\, dm,$$

where $\beta > 0$ and $C(C_1)$ is some positive function of C_1. Assume that $V \subset B(x_0; R)$ where the radius R is with respect to the distance d. Then we can take $W^n_{V,i} = B(x_0; R + ain^{1/\beta})$, for any a satisfying $a^\beta \geq C(C_1)$. For example,

(1) when the Dirichlet space admits a nice intrinsic distance, it is a special case of the discussion above with $\beta = 2$;
(2) when the Dirichlet space is the standard Dirichlet form on the Sierpinski gasket and d is the Euclidean metric, the discussion above applies with $\beta = \log 5/\log 2$, which is the walk dimension d_w of the Sierpinski gasket.

6.2 Proofs of Theorem 7 and Theorem 8

6.2.1 Overview and a Key Estimate

There are two difficulties in generalizing the structure results for ancient caloric functions to the current Dirichlet space setting. The first difficulty is in formulating proper assumptions on the existence of cut-off functions in order to adapt estimates of the form

$$\left| \int fg \nabla v \cdot \nabla w \, dx \right| \leq \left(\int (fg)^2 \, dx \right)^{1/2} \left(\int |\nabla v|^2 |\nabla w|^2 \, dx \right)^{1/2}$$

to estimates in terms of energy measures, especially when the energy measure is singular with respect to the measure m in the metric measure space (X, m). The second difficulty concerns whether time derivatives of an ancient (local weak) solution are still ancient solutions. This is answered positively by Corollary 1, which thus plays an essential role.

In this subsection we state the key estimate and use it to prove Theorem 7 and Theorem 8. The estimate is about bounding the L^2 integral of time derivatives of an ancient solution u over some time-space cylinder by the L^2 integral of u over some larger time-space cylinder, where the spatial sets are ones in an exhaustion of X. Let $C_1, C > 0$ be two fixed constants. Let $\{W_i\}_{i \in \mathbb{N}_+}$ be an exhaustion of X, let $\{\varphi_i\}_{i \in \mathbb{N}_+}$ be a sequence of cut-off functions where each φ_i is a cut-off function for the pair $W_i \subset W_{i+1}$, and satisfies that for any $v \in \mathcal{F}$,

$$\int_X v^2 \, d\Gamma(\varphi_i, \varphi_i) \leq C_1 \int_X \varphi_i^2 \, d\Gamma(v, v) + C \int_{\operatorname{supp}\{\varphi_i\}} v^2 \, dm.$$

We call such a pair $\left(\{W_i\}_{i \in \mathbb{N}_+}, \{\varphi_i\}_{i \in \mathbb{N}_+}\right)$ an exhaustion of X corresponding to C_1, C. The key estimate is as follows.

Proposition 5 *Let (X, m) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form on X. Assume that the Dirichlet space $(X, m, \mathcal{E}, \mathcal{F})$ satisfies Assumptions 3 and 4, and when X is not compact, further satisfies Assumption 6. Let $(H_t)_{t>0}$ and $-P$ be the corresponding semigroup and generator. Let u be an ancient (local weak) solution of $(\partial_t + P)u = 0$. Let $J := (c, 0]$ where $-\infty < c < 0$ is an arbitrarily fixed number, let $J_{-s} := (c - s, 0]$. Take $C_1 = 1/16$ and fix an arbitrary $C > 0$. Let $\left(\{W_i\}_{i \in \mathbb{N}_+}, \{\varphi_i\}_{i \in \mathbb{N}_+}\right)$ be an exhaustion of X corresponding to C_1, C, the existence of which is guaranteed by Assumption 6. Then for any $i, k \in \mathbb{N}_+$,*

$$\int_J \int_{W_i} \left(\partial_t^k u\right)^2 dm dt \leq \left(1200 \left(C + \frac{1}{r}\right)^2\right)^k \int_{J_{-2kr}} \int_{W_{i+3k}} u^2 \, dm dt.$$

In Lemma 5 below it can be seen that C_1 can take any value less than $1/9$, we take $C_1 = 1/16$ for convenience.

We now use Proposition 5 to prove Theorem 7 and Theorem 8. The proof of the proposition is given in Section 6.3. Note that the inequality in Proposition 5 remains true if we take J as a closed interval $J = [c, 0]$. We use $J = (c, 0]$ in the statement of the proposition for convenience in describing the cut-off function in time in its proof.

6.2.2 Proof of Theorem 7

To show $\partial_t^k u = 0$ for k large enough, we follow the idea in [9] and show that the L^2 integral of such $\partial_t^k u$ over any time-space cylinder is zero. Consider an arbitrary cylinder $[-T,\ 0] \times V$ where $T > 0$ and $V \subset X$ is a precompact open subset. For any $n \in \mathbb{N}_+$, Assumption 6 guarantees the existence of an exhaustion $\left(\{W_{V,i}^n\}_{i\in\mathbb{N}_+},\ \{\varphi_i\}_{i\in\mathbb{N}_+}\right)$ of X corresponding to $C_1 = 1/16$, $C = 1/n$. In particular, $[-T, 0] \times V \subset [-T, 0] \times W_{V,1}^n$, for any n. Taking $J = [-T, 0]$, $r = n$, $i = 1$ in Proposition 5 gives

$$\int_J \int_{W_{V,1}^n} \left(\partial_t^k u\right)^2 dmdt \le \left(\frac{5000}{n^2}\right)^k \int_{J_{-2kn}} \int_{W_{V,1+3k}^n} u^2\, dmdt.$$

Then by the growth condition of u, we have

$$\int_{[-T,\ 0]\times W_{V,1}^n} \left(\partial_t^k u\right)^2 dmdt \le \frac{5000^k \left(C_{u,V,1+3k} \max\left\{(\lceil T\rceil + 2kn)^{d_u},\ n^{b_u}\right\}\right)^2}{n^{2k}}.$$

Because for any fixed k with $2k > 2(d_u + b_u)$, the right-hand side tends to 0 as n tends to infinity, by the discussion above, we conclude that for $k > d_u + b_u$,

$$\partial_t^k u = 0.$$

Hence u is a polynomial in t. Applying the growth bound (6.2) to u in the explicit polynomial form

$$u(t, x) = \sum_{k=0}^{N} \partial_t^k u(0, x) \frac{1}{k!} t^k,$$

we have for any $W_{V,i}^n$, for any $T > 0$,

$$\begin{aligned}
&\int_{[-T,0]\times W_{V,i}^n} |u(t, x)|^2\, dmdt \\
&= c_{2N+1} T^{2N+1} + c_{2N} T^{2N} + c_{2N-1} T^{2N-1} + \cdots + c_1 T \\
&\le \left(C_{u,V,i} \max\{T^{d_u},\ n^{b_u}\}\right)^2,
\end{aligned}$$

where

$$c_{2N+1} = \frac{1}{2N+1} \int_{W_{V,i}^n} \left(\frac{1}{N!} \partial_t^N u(0, x)\right)^2 dm.$$

Let T tend to infinity, it follows that $2N + 1 \le 2d_u$. We conclude that $\partial_t^k u = 0$ for $k > d_u - 1/2$.

6.2.3 Proof of Theorem 8

This proof follows the idea in [38]. By the Taylor expansion formula (expansion in t), for any $t < 0$,

$$u(t,x) = \sum_{i=0}^{k} \frac{\partial_t^i u(0,x)}{i!} t^i + \int_0^t \partial_s^{k+1} u(s,x) \frac{(t-s)^k}{k!}\, ds$$

as L^2 functions in x. So to prove the statement in Theorem 8, we prove that for any precompact open set V, for any $a < 0$,

$$\sup_{a \le t \le 0} \int_V \left(\int_0^t \partial_s^{k+1} u(s,x) \frac{(t-s)^k}{k!}\, ds \right)^2 dm(x) \to 0 \tag{6.7}$$

as $k \to \infty$. We first bound the integral by Jensen's inequality,

$$\begin{aligned} &\left| \int_V \left(\int_0^t \partial_s^{k+1} u(s,x) \frac{(t-s)^k}{k!}\, ds \right)^2 dm(x) \right| \\ &\le \frac{|t|}{(k!)^2} \int_V \int_t^0 \left(\partial_s^{k+1} u(s,x) \times (t-s)^k \right)^2 dsdm \\ &\le \frac{|t|^{2k+1}}{(k!)^2} \int_V \int_t^0 \left(\partial_s^{k+1} u(s,x) \right)^2 dsdm. \end{aligned} \tag{6.8}$$

Recall the notation introduced in the statement of Theorem 8, i.e., $W_j := W^1_{V_0,j}$ for some fixed $V_0 \Subset X$. Intuitively, by fixing $n = 1$ (or any fixed integer), we are looking at open sets whose sizes grow linearly. Because $V \Subset X$ and $\{W_j\}_{j \in \mathbb{N}_+}$ is an exhaustion of X, there exists some j_0 such that for all $j \ge j_0$, $V \subset W_j$. By Proposition 5, for any $r > 0$,

$$\begin{aligned} &\int_t^0 \int_{W_{j_0}} \left(\partial_s^{k+1} u(s,x) \right)^2 dmdt \\ &\le \left(1200 \left(1 + \frac{1}{r} \right)^2 \right)^{k+1} \int_{t-2(k+1)r}^0 \int_{W_{j_0+3(k+1)}} u(s,x)^2\, dmdt. \end{aligned}$$

By the exponential growth assumption (6.3) on u, we conclude that (take for example $r = 1$) for any $t \in [a, 0]$,

$$\int_t^0 \int_{W_{j_0}} \left(\partial_s^{k+1} u(s,x) \right)^2 dmdt \le (5000)^{k+1} e^{c_u(|a| + j_0 + 5(k+1))}.$$

Substituting this bound back to (6.8), noting that $V \subset W_{j_0}$, we get

$$\begin{aligned} &\left| \int_V \left(\int_0^t \partial_s^{k+1} u(s,x) \frac{(t-s)^k}{k!}\, ds \right)^2 dm(x) \right| \\ &\le \frac{|a|^{2k+1}}{(k!)^2} (5000)^{k+1} e^{c_u(|a| + j_0 + 5(k+1))} \to 0 \;\; (k \to \infty). \end{aligned}$$

This completes the proof of Theorem 8.

6.3 Proof of the Key Estimate

In this subsection we give the proof of Proposition 5, which is an iteration of the following proposition.

Proposition 6 *Under the same hypotheses as Proposition 5, for $C_1 = 1/16$ and for any $C > 0$, let $\left(\{W_i\}_{i\in\mathbb{N}_+}, \{\varphi_i\}_{i\in\mathbb{N}_+}\right)$ be an exhaustion of X corresponding to C_1, C. Then for any $r > 0$, there exist constants K_1, K_2 (dependent on C and r) such that for any $i \in \mathbb{N}_+$,*

$$\begin{aligned}\int_J \int_{W_i} (\partial_t u)^2\, dmdt \;&\le\; K_1(C,r) \int_{J_{-r}} \left(\int_{W_{i+2}} d\Gamma(u,u) + \int_{W_{i+2}} u^2\, dk \right) dt \\ &\le\; K_2(C,r) \int_{J_{-2r}} \int_{W_{i+3}} u^2\, dmdt.\end{aligned}$$

Here

$$K_1(C,r) = 200\left(C + \frac{1}{r}\right),\; K_2(C,r) = 1200\left(C + \frac{1}{r}\right)^2.$$

6.3.1 Proof of Proposition 6

We present the proof in three steps.

Step 1. We first prove two Caccioppoli type inequalities, one of which gives the second inequality in the proposition. Recall that $J = (c, 0]$ for some $c < 0$, and $J_{-s} = (c-s, 0]$ for any $s > 0$. Let l be a smooth cut-off function in time (on $(-\infty, 0]$) that equals 1 on $[c, 0]$, has compact support in $(c - r, 0]$, and satisfies $0 \le l \le 1$, $\|l'\|_{L^\infty} := \sup_{t\in(-\infty,0]} |l'(t)| \le 2/r$. It can be easily extended into a function in $C_c^\infty(\mathbb{R})$, in the following we only use its part on $(-\infty, 0]$. By construction, each φ_i is a nice cut-off function for the pair $W_i \subset W_{i+1}$, and satisfies that for any $v \in \mathcal{F}$,

$$\int_X v^2\, d\Gamma(\varphi_i,\varphi_i) \le C_1 \int_X \varphi_i^2\, d\Gamma(v,v) + C \int_{\mathrm{supp}\{\varphi_i\}} v^2\, dm. \tag{6.9}$$

The product $\varphi_i(x)l(t)$ is then a nice product cut-off function for the pair $J \times W_i \subset J_{-r} \times W_{i+1}$. First we have

$$\begin{aligned}&\int_{J_{-r}} \int_{W_{i+1}} 2u(\partial_t u)\varphi_i^2 l^2\, dmdt + \int_{J_{-r}} \int_{W_{i+1}} u^2 \varphi_i^2 \left(l^2\right)'\, dmdt \\ &= \int_{J_{-r}} \partial_t \left(\int_{W_{i+1}} u^2\varphi_i^2 l^2\, dm \right) dt = \left(\int_{W_{i+1}} u^2 \varphi_i^2 l^2\, dm \right)\Bigg|_{t=0} \ge 0.\end{aligned}$$

On the other hand, since u is an ancient local weak solution of the heat equation $(\partial_t + P)u = 0$, φ_i is supported in W_{i+1}, the first term above is

$$\begin{aligned}&\int_{J_{-r}} \int_{W_{i+1}} 2u(\partial_t u)\varphi_i^2 l^2\, dmdt = -2 \int_{J_{-r}} l^2 \mathcal{E}(u,\, u\varphi_i^2)\, dt \\ &= \; -2 \int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2\, d\Gamma(u,u)\, dt - 4 \int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i u\, d\Gamma(\varphi_i, u)\, dt \\ &\quad -2 \int_{J_{-r}} l^2 \int_X u^2 \varphi_i^2\, dkdt \\ &:= I + II + III.\end{aligned}$$

By the Cauchy-Schwartz inequality (2.1), the middle term

$$II = -4\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i u \, d\Gamma(\varphi_i, u)\, dt$$
$$\leq 4\left(\frac{1}{4}\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt + 1 \times \int_{J_{-r}} l^2 \int_{W_{i+1}} u^2 \, d\Gamma(\varphi_i, \varphi_i)\, dt\right),$$

where by (6.9), the part $\int_{J_{-r}} l^2 \int_{W_{i+1}} u^2 \, d\Gamma(\varphi_i, \varphi_i)\, dt$ is further bounded by

$$\int_{J_{-r}} l^2 \int_{W_{i+1}} u^2 \, d\Gamma(\varphi_i, \varphi_i)\, dt$$
$$\leq C_1 \int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt + C\int_{J_{-r}} l^2 \int_{W_{i+1}} u^2 \, dmdt.$$

So $I + II$ is bounded by

$$I + II = -2\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt - 4\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i u \, d\Gamma(\varphi_i, u)\, dt$$
$$\leq -\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt + 4\int_{J_{-r}} l^2 \int_{W_{i+1}} u^2 \, d\Gamma(\varphi_i, \varphi_i)\, dt$$
$$\leq -(1-4C_1)\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt + 4C\int_{J_{-r}} l^2 \int_{W_{i+1}} u^2 \, dmdt.$$

Combining the estimates so far and apply $\left\|(l^2)'\right\|_{L^\infty} = \left\|2ll'\right\|_{L^\infty} \leq 4/r$, we get

$$(1-4C_1)\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt + 2\int_{J_{-r}} l^2 \int_X u^2\varphi_i^2 \, dkdt$$
$$\leq \left(4C + \frac{4}{r}\right)\int_{J_{-r}} \int_{W_{i+1}} u^2 \, dmdt,$$

which then implies

$$\int_J \left(\int_{W_i} d\Gamma(u,u) + \int_{W_i} u^2 \, dk\right) dt$$
$$\leq (1-4C_1)^{-1}\left(4C + \frac{4}{r}\right)\int_{J_{-r}} \int_{W_{i+1}} u^2 \, dmdt. \tag{6.10}$$

This proves the second inequality in Proposition 6. It is a Caccioppoli type inequality that we will use later in the proof. We also need a version of the inequality with $\varphi_i^2 u^2$ instead of u^2 on the right-hand side of (6.10). To get that we repeat the above computations, replacing φ_i with φ_i^2, l with l^2. First,

$$-\int_{J_{-r}} \int_{W_{i+1}} u^2 \varphi_i^4 \left(l^4\right)' \, dmdt$$
$$\leq \int_{J_{-r}} \int_{W_{i+1}} 2u(\partial_t u)\varphi_i^4 l^4 \, dmdt = -2\int_{J_{-r}} l^4 \mathcal{E}(u,\, u\varphi_i^4)\, dt$$
$$= -2\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^4 \, d\Gamma(u,u)\, dt - 8\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^3 u \, d\Gamma(\varphi_i, u)\, dt$$
$$-2\int_{J_{-r}} l^4 \int_X u^2 \varphi_i^4 \, dkdt,$$

where

$$-2\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^4 \, d\Gamma(u,u)\, dt - 8\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^3 u \, d\Gamma(\varphi_i, u)\, dt$$
$$\leq -\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^4 \, d\Gamma(u,u)\, dt + 16\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^2 u^2 \, d\Gamma(\varphi_i, \varphi_i)\, dt.$$

Instead of applying (6.9) as before, we now use Lemma 5 below to get that the term

$$\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^2 u^2 \, d\Gamma(\varphi_i, \varphi_i)\, dt \leq C_0 \int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 u^2 \, dmdt.$$

Here $C_0 = C_0(C_1, C, l)$ is independent of u. Combining the estimates as before (note that $|(l^4)'| = 4l^3|l'| \leq (8/r)l^2$), we thus have

$$\begin{aligned} &\int_{J_{-r}} l^4 \int_{W_{i+1}} \varphi_i^4 \, d\Gamma(u,u)\, dt + \int_{J_{-r}} l^4 \int_X u^2 \varphi_i^4 \, dkdt \\ &\leq 4\left(4C_0 + \frac{2}{r}\right) \int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 u^2 \, dmdt. \end{aligned} \tag{6.11}$$

Here both sides of the inequality are integrals over the same set $J_{-r} \times W_{i+1}$.

Step 2. Next we estimate the L^2 norm of $\partial_t u$, which by Corollary 1 is also an ancient solution. In this step we do some preparatory work. Because $\partial_t u$ is a local weak solution of $(\partial_t + P)\partial_t u = 0$ on $(-\infty, b) \times X$ for some $b > 0$, we have

$$\begin{aligned} \int_{J_{-r}} \int_{W_{i+1}} (\partial_t u \, \varphi_i l)^2 \, dmdt &= -\int_{J_{-r}} l^2 \, \mathcal{E}(u,\ \partial_t u \, \varphi_i^2)\, dt \\ &= -\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u, \partial_t u)\, dt - \int_{J_{-r}} l^2 \int_{W_{i+1}} 2\varphi_i \, \partial_t u \, d\Gamma(u, \varphi_i)\, dt \\ &\quad - \int_{J_{-r}} l^2 \int_X u \, \partial_t u \, \varphi_i^2 \, dkdt. \end{aligned} \tag{6.12}$$

To estimate the right-hand side, we first show that

$$\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(\partial_t u, u) = \frac{1}{2}\partial_t \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u) \right)$$

to replace the first term in (6.12). By Theorem 5, u locally belongs to the space $C^\infty((-\infty, b) \to \mathcal{F})$. Fix any $t \in (-\infty, 0]$, let

$$v_n^t(x) := \frac{u(t + 1/n,\ x) - u(t, x)}{1/n}.$$

Then $v_n^t \to \partial_t u$ locally in $\mathcal{F}$. In particular,

$$\lim_{n\to\infty} \int_{W_{i+1}} d\Gamma(v_n^t - \partial_t u,\ v_n^t - \partial_t u) = 0.$$

By definition,

$$
\begin{aligned}
&\partial_t \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u, u) \right) \\
&= \lim_{n\to\infty} \frac{\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u(t+1/n, \cdot), \, u(t+1/n, \cdot)) - \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u(t, \cdot), \, u(t, \cdot))}{1/n} \\
&= \lim_{n\to\infty} \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^{t+\frac{1}{n}}, \, \frac{u^{t+\frac{1}{n}} - u^t}{1/n} \right) + \int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^t, \, \frac{u^{t+\frac{1}{n}} - u^t}{1/n} \right) \right) \\
&= \lim_{n\to\infty} \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^{t+\frac{1}{n}}, \, v_n^t \right) + \int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^t, \, v_n^t \right) \right).
\end{aligned}
$$

More accurately, u should be some $u^\sharp \in C^\infty((-\infty, b) \to \mathcal{F})$ that agrees with u on some time-space cylinder covering $\{t\} \times W_{i+1}$, and v_n^t is given using $u^\sharp$ in its definition. Then using the Cauchy-Schwartz inequality (2.1) and noting that $u^{t+1/n} \to u^t$ locally in $\mathcal{F}$; $v_n^t \to \partial_t u$ locally in $\mathcal{F}$, we can check that the limit above is

$$
2 \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u^t, \, \partial_t u).
$$

For example,

$$
\begin{aligned}
&\int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^{t+\frac{1}{n}}, \, v_n^t \right) - \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u^t, \, \partial_t u) \\
&= \int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^{t+\frac{1}{n}} - u^t, \, v_n^t \right) + \int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^t, \, v_n^t - \partial_t u \right) \\
&\leq \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(u^{t+\frac{1}{n}} - u^t, \, u^{t+\frac{1}{n}} - u^t \right) \right)^{1/2} \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(v_n^t, \, v_n^t) \right)^{1/2} \\
&\quad + \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u^t, \, u^t) \right)^{1/2} \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma \left(v_n^t - \partial_t u, \, v_n^t - \partial_t u \right) \right)^{1/2},
\end{aligned}
$$

where each summand is a product of one term uniformly bounded in n, and one term that tends to 0 as n tends to infinity. Here to show the uniform boundedness it is useful to use the estimate

$$
\left| \left(\int_X f \, d\Gamma(v, v) \right)^{1/2} - \left(\int_X f \, d\Gamma(w, w) \right)^{1/2} \right| \leq \left(\int_X f \, d\Gamma(v - w, v - w) \right)^{1/2}
$$

where f is any nonnegative bounded Borel function, and $v, w \in \mathcal{F} \cap L^\infty(X)$ (cf. Chapter 3 in [16]). So we conclude that

$$
\frac{1}{2} \partial_t \left(\int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u, u) \right) = \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(\partial_t u, u).
$$

Step 3. We are now ready to estimate (6.12). The first term is

$$
\begin{aligned}
-\int_{J_{-r}} l^2 \int_{W_{i+1}} \varphi_i^2\, d\Gamma(u, \partial_t u)\, dt &= -\frac{1}{2} \int_{J_{-r}} l^2\, \partial_t \left(\int_{W_{i+1}} \varphi_i^2\, d\Gamma(u,u) \right) dt \\
&= -\frac{1}{2} \int_{J_{-r}} \partial_t \left(l^2 \int_{W_{i+1}} \varphi_i^2\, d\Gamma(u,u) \right) dt + \frac{1}{2} \int_{J_{-r}} \left(l^2\right)' \int_{W_{i+1}} \varphi_i^2\, d\Gamma(u,u)\, dt \\
&\leq \frac{1}{2} \int_{J_{-r}} \left(l^2\right)' \int_{W_{i+1}} \varphi_i^2\, d\Gamma(u,u)\, dt. \qquad (6.13)
\end{aligned}
$$

The second term in (6.12) satisfies

$$
\begin{aligned}
&\left| -\int_{J_{-r}} l^2 \int_{W_{i+1}} 2\varphi_i\, \partial_t u\, d\Gamma(u, \varphi_i)\, dt \right| \\
&= \left| -\int_{J_{-r}} l^2 \int_{W_{i+1}} 2\varphi_i \varphi_{i+1} \partial_t u\, d\Gamma(u, \varphi_i)\, dt \right| \\
&\leq \epsilon \int_{J_{-r}} \int_{W_{i+1}} l^4 \varphi_i^2 \left(\partial_t u\right)^2\, d\Gamma(\varphi_i, \varphi_i)\, dt + \frac{1}{\epsilon} \int_{J_{-r}} \int_X \varphi_{i+1}^2\, d\Gamma(u,u)\, dt
\end{aligned}
$$

for any $\epsilon > 0$. Here φ_{i+1} is the nice cut-off function for the pair $W_{i+1} \subset W_{i+2}$. In particular, $\varphi_{i+1} \equiv 1$ on $\operatorname{supp}\{\varphi_i\}$. By Lemma 5,

$$
\int_{J_{-r}} \int_{W_{i+1}} l^4 \varphi_i^2 \left(\partial_t u\right)^2\, d\Gamma(\varphi_i, \varphi_i) dt \;\leq C_0 \int_{J_{-r}} \int_{W_{i+1}} l^2 \varphi_i^2 \left(\partial_t u\right)^2\, dm dt,
$$

where $C_0 \leq 3C + 1/r$. Thus

$$
\begin{aligned}
&\left| -\int_{J_{-r}} l^2 \int_{W_{i+1}} 2\varphi_i\, \partial_t u\, d\Gamma(u, \varphi_i)\, dt \right| \\
&\leq \epsilon C_0 \int_{J_{-r}} \int_{W_{i+1}} \left(\partial_t u\, \varphi_i l\right)^2\, dm dt + \frac{1}{\epsilon} \int_{J_{-r}} \int_X \varphi_{i+1}^2\, d\Gamma(u,u)\, dt. \qquad (6.14)
\end{aligned}
$$

The last term in (6.12) satisfies by the Cauchy-Schwartz inequality

$$
\begin{aligned}
-\int_{J_{-r}} l^2 \int_X u \partial_t u\, \varphi_i^2\, dk dt &= -\int_{J_{-r}} \int_X \varphi_{i+1} l^2 u \partial_t u\, \varphi_i^2\, dk dt \\
&\leq \frac{1}{2c} \int_{J_{-r}} \int_X \varphi_{i+1}^2 u^2\, dk dt + \frac{c}{2} \int_{J_{-r}} \int_X l^4 (\partial_t u)^2 \varphi_i^4\, dk dt \qquad (6.15)
\end{aligned}
$$

for any $c > 0$. Now we plug in $C_1 = 1/16$ and take $\epsilon = (2C_0)^{-1}$, then by (6.11), (6.12), (6.13), (6.14), and (6.15),

$$
\begin{aligned}
&\int_{J_{-r}} \int_{W_{i+1}} \left(\partial_t u\, \varphi_i l\right)^2\, dm dt \\
&\leq \frac{1}{2} \int_{J_{-r}} \left(l^2\right)' \int_{W_{i+1}} \varphi_i^2\, d\Gamma(u,u)\, dt + \frac{1}{2} \int_{J_{-r}} \int_{W_{i+1}} \left(\partial_t u\, \varphi_i l\right)^2\, dm dt \\
&\quad + 2C_0 \int_{J_{-r}} \int_X \varphi_{i+1}^2\, d\Gamma(u,u)\, dt + \frac{1}{2c} \int_{J_{-r}} \int_X \varphi_{i+1}^2 u^2\, dk dt \\
&\quad + c\left(8C_0 + \frac{4}{r}\right) \int_{J_{-r}} \int_X l^2 (\partial_t u)^2 \varphi_i^2\, dm dt.
\end{aligned}
$$

Letting $c = 4^{-1}\left(8C_0 + 4r^{-1}\right)^{-1}$ and combining coefficients for the term on the left-hand side, we get

$$\begin{aligned}
&\int_{J_{-r}} \int_{W_{i+1}} (\partial_t u\, \varphi_i l)^2 \; dmdt \\
&\leq 2\int_{J_{-r}} \left(l^2\right)' \int_{W_{i+1}} \varphi_i^2 \, d\Gamma(u,u)\, dt + 8C_0 \int_{J_{-r}} \int_X \varphi_{i+1}^2 \, d\Gamma(u,u)\, dt \\
&\quad 8\left(8C_0 + \frac{4}{r}\right) \int_{J_{-r}} \int_X u^2 \varphi_{i+1}^2 \, dkdt \\
&\leq 8\left(8C_0 + \frac{4}{r}\right) \int_{J_{-r}} \left(\int_{W_{i+2}} d\Gamma(u,u) + \int_{W_{i+2}} u^2 \, dk \right) dt. \qquad (6.16)
\end{aligned}$$

Take J_{-r} and W_{i+2} on the left in (6.10) with $C_1 = 1/16$, combine (6.10) and (6.16), and recall that $C_0 \leq 3C + 1/r$, we obtain that

$$\begin{aligned}
&\int_J \int_{W_i} (\partial_t u)^2 \, dmdt \\
&\leq 200\left(C + \frac{1}{r}\right) \int_{J_{-r}} \left(\int_{W_{i+2}} d\Gamma(u,u) + \int_{W_{i+2}} u^2 \, dk \right) dt \\
&\leq 1200\left(C + \frac{1}{r}\right)^2 \int_{J_{-2r}} \int_{W_{i+3}} u^2 \, dmdt.
\end{aligned}$$

Let $K_1(C,r) := 200\,(C + 1/r)$ and $K_2(C,r) := 1200\,(C + 1/r)^2$. This completes the proof of Proposition 6. Note that by taking C small and r large enough, we can make the coefficients $K_1(C,r)$ and $K_2(C,r)$ as small as needed.

Straightforward iterations lead to Proposition 5.

6.3.2 A Technical Lemma

Last we state and prove the technical lemma used in the proof of Proposition 6.

Lemma 5 *Let (X, m) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form on X. Assume that the Dirichlet space $(X, m, \mathcal{E}, \mathcal{F})$ satisfies Assumption 3. Let $(H_t)_{t>0}$ and $-P$ be the corresponding semigroup and generator. Let I, I' be two intervals where $I = (a,b)$ or $(a,b]$, $-\infty \leq a < b \leq \infty$, $I \subset I'$. Let u be a local weak solution of the heat equation $(\partial_t + P)u = 0$ on $I' \times X$. Let $\overline{\varphi}(t,x) := \varphi(x)l(t)$ be a nice product cut-off function, where φ corresponds to coefficients C_1, C_2 with $C_1 < 1/9$, and $l \in C_c^\infty(I)$ is a smooth cut-off function on $\mathbb{R}$. Then there exists some C_0 that depends on $\overline{\varphi}$, or equivalently, on C_1, C_2, l, such that*

$$\int_I \int_X \overline{\varphi}^2 u^2 \, d\Gamma(\overline{\varphi}, \overline{\varphi})\, dt \leq C_0 \int_I \int_X \overline{\varphi}^2 u^2 \, dmdt.$$

The last inequality says when the same cut-off function with bounded energy is both in the integrand and in the energy measure, the net effect is the same as having a cut-off function with bounded gradient in the energy measure. Observe that for $\int_X \varphi^2\, d\Gamma(\varphi,\varphi)$, it is easy to check that

$$\int_X \varphi^2\, d\Gamma(\varphi,\varphi) \le (1-C_1)^{-1} C_2 \int_{\operatorname{supp}\{\varphi\}} \varphi^2\, dm.$$

In the lemma we generalize this observation to bound $\int_I \int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi},\overline{\varphi})$, for local weak solutions u.

Proof Because $\overline{\varphi}$ is a nice product cut-off function associated with C_1, C_2,

$$\int_I \int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi},\overline{\varphi})\, dt \le C_1 \int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt + C_2 \int_I \int_X l^2 \overline{\varphi}^2 u^2\, dm dt.$$

To estimate $\int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt$, we make the following two observations

(i) $\int_I \int_X \overline{\varphi}^2\, d\Gamma(u,\overline{\varphi}^2 u)\, dt = \int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt - \int_I \int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi},\overline{\varphi})\, dt;$

(ii) $\int_I \int_X \overline{\varphi}^2\, d\Gamma(u,\overline{\varphi}^2 u)\, dt$
$= \int_I \int_X d\Gamma(u,\overline{\varphi}^4 u)\, dt - 2\int_I \int_X \overline{\varphi}^2 u\, d\Gamma(\overline{\varphi}u,\overline{\varphi})\, dt + 2\int_I \int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi},\overline{\varphi})\, dt.$

The middle term in (the right-hand side of) (ii) satisfies (by (2.1))

$$\begin{aligned}&\left|2\int_I \int_X \overline{\varphi}^2 u\, d\Gamma(\overline{\varphi}u,\overline{\varphi})\, dt\right| \\ &\quad \le \epsilon \int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt + \frac{1}{\epsilon}\int_I \int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi},\overline{\varphi})\, dt,\end{aligned}$$

for any $\epsilon > 0$. To estimate the first term in (ii), note that u being a local weak solution implies that (using Definition 2)

$$\begin{aligned}&\int_I \int_X d\Gamma(u,\overline{\varphi}^4 u)\, dt = -\int_I \int_X \partial_t u\, \overline{\varphi}^4 u\, dm dt - \int_I \int_X \overline{\varphi}^4 u^2\, dk dt \\ &\le -\frac{1}{2}\int_I \int_X (\partial_t(l^4 u^2 \varphi^4) - \partial_t(l^4) u^2 \varphi^4)\, dm dt \le 2\|l'\|_{L^\infty(I)} \int_I \int_X l^3 u^2 \varphi^4\, dm dt.\end{aligned}$$

Combining (i)(ii) and the estimates above, we get that

$$\begin{aligned}&\int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt \le 2\|l'\|_{L^\infty(I)} \int_I \int_X l^3 \varphi^4 u^2\, dm dt \\ &\qquad + \epsilon \int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt + \left(\frac{1}{\epsilon}+3\right)\int_I \int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi},\overline{\varphi})\, dt \\ &\le 2\|l'\|_{L^\infty(I)} \int_I \int_X l^3 \varphi^4 u^2\, dm dt + \left[\epsilon + C_1\left(\frac{1}{\epsilon}+3\right)\right]\int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt \\ &\qquad + \left(\frac{1}{\epsilon}+3\right) C_2 \int_I \int_X l^2 \overline{\varphi}^2 u^2\, dm dt.\end{aligned}$$

When $C_1 < 1/9$, we can pick ϵ small so that $\epsilon + C_1\left(\epsilon^{-1}+3\right) < 1$. Let $\alpha = 1 - \left[\epsilon + C_1\left(\epsilon^{-1}+3\right)\right] > 0$, the above estimate is equivalent to

$$\begin{aligned}&\int_I \int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u,\overline{\varphi}u)\, dt \\ &\quad \le \frac{1}{\alpha}\left\{2\|l'\|_{L^\infty(I)} \int_I \int_X l^3 \varphi^4 u^2\, dm dt + \left(\frac{1}{\epsilon}+3\right) C_2 \int_I \int_X l^2 \overline{\varphi}^2 u^2\, dm dt\right\}.\end{aligned}$$

Letting $K := \alpha^{-1}\left[2\|l'\|_{L^\infty(I)} + \left(\epsilon^{-1}+3\right)C_2\right]$ and noting that $l\varphi^2 \le 1$, $l^2 \le 1$, we get

$$\int_I\int_X \overline{\varphi}^2\, d\Gamma(\overline{\varphi}u, \overline{\varphi}u)\, dt \le K \int_I\int_X \overline{\varphi}^2 u^2\, dmdt.$$

Combining this with the very first inequality, we get

$$\int_I\int_X \overline{\varphi}^2 u^2\, d\Gamma(\overline{\varphi}, \overline{\varphi})\, dt \le C_1 K \int_I\int_X \overline{\varphi}^2 u^2\, dmdt + C_2 \int_I\int_X \overline{\varphi}^2 u^2\, dmdt.$$

Letting $C_0 = C_1 K + C_2$ gives the inequality in the lemma. To apply this lemma to the proof of Proposition 6, let $C_1 = 1/16$ and $C_2 = C$, plug in $\|l'\|_{L^\infty(I)} \le 2/r$ (here $I = (c-r, 0]$), and take for example $\alpha = 5/16$, we get that $\epsilon = 1/4$, $C_0 \le 3C + 1/r$. ☐

7 Examples

In this section we list some examples to which our theorems apply. We group them according to the types of nice cut-off functions they admit. Note that the properties we require on the nice cut-off functions involve only the energy measure associated with the Dirichlet form, so in the following we describe examples of strongly local Dirichlet forms; our theorems apply to local Dirichlet forms whose strongly local parts belong to the following examples as well.

7.1 Dirichlet Spaces with Good Intrinsic Distance

In [34], Sturm showed that for a symmetric strongly local regular Dirichlet space, when the topology induced by the intrinsic distance (3.4), that is,

$$\rho_X(x, y) = \sup\left\{\varphi(x) - \varphi(y) \mid \varphi \in \mathcal{F}_{\text{loc}}(X) \cap C(X),\ d\Gamma(\varphi, \varphi) \le dm\right\},$$

is equivalent to the original topology on X, one can use the intrinsic distance to construct nice cut-off functions with bounded gradient. More precisely, for $V \Subset U \Subset X$, define

$$\eta(x) := \frac{\left(\frac{1}{\sqrt{2}}\rho_X(V, U^c) - \rho_X(x, V)\right)_+}{\frac{1}{\sqrt{2}}\rho_X(V, U^c)}.$$

Here $(\cdot)_+$ denotes the positive part. Clearly $\eta = 1$ on V and $\operatorname{supp}\{\eta\} \subset U$. Further, η is in $\mathcal{F}_{\text{loc}}(X) \cap C(X)$, and

$$d\Gamma(\eta, \eta) \le \frac{2}{\rho_X(V, U^c)^2}\, dm. \tag{7.1}$$

See [34, Lemma 1.9]. It clearly follows that such Dirichlet spaces satisfy Assumption 3 and Assumption 6 (pick the exhaustion $\{W^n_{V,i}\}_{i\in\mathbb{N}_+}$ to be balls with radii r_i's that increase fast enough). By Lemma 6, these Dirichlet spaces satisfy the L^2 Gaussian type upper bound. Thus all results in this paper apply to this type of examples which includes:

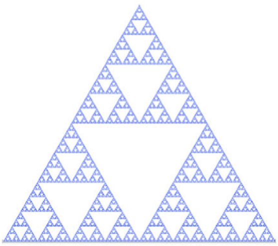
(1) Weighted Riemannian manifolds with Dirichlet forms associated with any locally uniformly elliptic second order divergence form operator with locally bounded measurable coefficients. See, for example, [33]. This includes the example we described in the Introduction.
(2) Riemannian polyhedra under minimal local assumptions (cf. [14, 32] and [8]).

(3) Alexandrov spaces and their Dirichlet space structures as considered for instance in [25, 31].

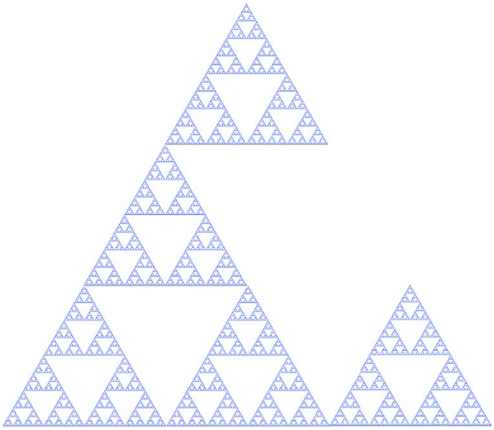
7.2 Fractal Type Dirichlet Spaces

For fractal spaces, Assumption 3 is a nontrivial hypothesis to check. It is well known that in many fractal spaces the only functions in $\mathcal{F}_{\mathrm{loc}}(X) \cap C(X)$ with bounded gradient are constant functions (cf. e.g. [18]). More generally, in a recent paper [23], it was shown that for a very general class of Dirichlet spaces, two-sided off-diagonal heat kernel estimate with walk-dimension strictly larger than two implies the singularity of the energy measure with respect to the symmetric measure.

On the other hand, many fractal spaces admit cut-off functions satisfying the inequality (3.1) in Assumption 3. We remark that the existence of cut-off functions satisfying (3.1) on such examples is highly nontrivial, and although their existence is known, there is in general no direct geometric construction of such cut-off functions. Some first examples satisfying Assumption 3 include the Sierpinsket gasket and its non-compact extension as in the following pictures. (The picture on left ($\mathcal{SG}$) is from Wikipedia, the picture on right ($\mathcal{ISG}$) is created by shifting copies of $\mathcal{SG}$.)



(a) the Sierpinski gasket $\mathcal{SG}$



(b) the infinite Sierpinski gasket $\mathcal{ISG}$

One example of obtaining the validity of Assumption 3 in an indirect way is [1, Theorem 1.12]. In [1], Andres and Barlow showed that certain Dirichlet spaces satisfying some pointwise heat kernel upper bound (called HKU(Ψ) in [1]), must satisfy a condition called CSA(Ψ). The CSA(Ψ) condition guarantees that the Dirichlet space admits cut-off functions that satisfy some more specific version of the inequality (3.1). CSA(Ψ) also implies Assumption 6. Therefore, under the HKU(Ψ) condition, all our results apply.

Consider the following refined version of Assumption 3: there is some distance d_X that defines the same topology of X, such that Assumption 3 holds with C_2 in (3.1) satisfying $C_2 \le CC_1^{-\alpha} d_X(V, U^c)^{-\beta}$ for some constants $C, \alpha, \beta > 0$. While this condition by itself does not imply any pointwise heat kernel upper bound like HKU(Ψ), it does imply the L^2 Gaussian type upper bound by Lemma 6. It is obvious that this condition also implies Assumption 6 (taking balls given by d_X, see the final remark before Section 6.2 for more details). So this refined version of Assumption 3 is a sufficient condition for all our results. In [3], Barlow and Murugan proved that this condition is quite typical.

7.3 Infinite Products of Dirichlet Spaces of the First Two Types

The first examples of this type are the infinite dimensional torus $\mathbb{T}^\infty$ and the infinite product of Sierpinski gaskets $\mathcal{SG}^\infty$, the first one being a special case of the class of compact (more generally, locally compact) connected metrizable (infinite dimensional) groups, cf. [5], and the second one the simplest of the infinite product of compact fractal spaces. To have some noncompact examples, consider $\mathbb{R} \times \mathbb{T}^\infty$, or the Iwasawa's example (cf. [5, 22]), or replace one piece of Sierpinski gasket in the product $\mathcal{SG}^\infty$ with the infinite Sierpinski gasket $\mathcal{ISG}$. This type of examples does not satisfy a property often satisfied in the previous two types of examples, namely, for these infinite dimensional spaces, the volume doubling property (local or global) cannot hold.

A general treatment of the elliptic diffusion on infinite product spaces like $\mathbb{T}^\infty$ is given in [4], and their results apply more generally to anomalous diffusion on infinite products of fractal spaces too. On a locally compact connected metrizable group G that is unimodular, one usually starts with a heat (convolution) semigroup, or a (left-invariant) Laplacian of the form $L = -\sum_{i,j=1}^\infty a_{ij} X_i X_j$, where $(a_{ij})_{i,j=1}^\infty$ is symmetric positive definite and $\{X_i\}_{i\in\mathbb{N}_+}$ is a projective basis of the left-invariant vector fields on G (i.e., a basis of the projective Lie algebra of G), and then consider the associated (left-invariant) Dirichlet form. Depending on the coefficients, the Dirichlet form may or may not have non-degenerate intrinsic distance. See [5].

For general product spaces X that have rougher differential structures, like $\mathcal{SG}^\infty$, it is easier and more convenient to consider only the "diagonal Dirichlet form", namely, for any diagonal matrix $(a_{ii})_{i=1}^\infty$ with all $a_{ii} > 0$, consider

$$\mathcal{E}(f, g) = \sum_{i=1}^\infty a_{ii} \int \mathcal{E}_i(f, g)\, d\left(\bigotimes_{j\neq i} m_j\right). \tag{7.2}$$

Here $\mathcal{E}_i$ stands for the standard Dirichlet form on the i-th factor of X, m_j stands for the normalized Hausdorff measure on the j-th factor of X, and f, g are proper functions.

Some infinite product examples do possess non-degenerate intrinsic distances that define the same topology (e.g. when the coefficient matrix for the Laplacian on $\mathbb{T}^\infty$ is diagonal and satisfies $\sum_{i=1}^\infty a_{ii}^{-1} < \infty$), in which case Assumption 3 and Assumption 6 follow. More generally, one can show that the cut-off function assumptions (Assumption 3 and Assumption 6) are satisfied using the fact that each factor in the infinite product possesses nice cut-off functions in the senses required.

More precisely, since the product topology is generated by cylindric sets (sets that are direct products of open sets of the first few factors, and the whole space for all remaining factors), for pairs of cylindric sets it is easy to construct a nice cut-off function being a product of nice cut-off functions for pairs of open sets on the first few factors, namely,

$$\varphi(\mathbf{x}) := \prod_{i=1}^N \varphi_i(x_i) \tag{7.3}$$

for some $N \in \mathbb{N}_+$. We verify this for the simpler case when the Dirichlet form is defined as in (7.2) (for the group case this is when the coefficient matrix is diagonal).

Suppose $\varphi_i(x_i)$ is a nice cut-off function on the i-th factor X_i of the infinite product space $X = \prod_{i=1}^{\infty} X_i$, satisfying that for any $v \in \mathcal{D}(\mathcal{E}_i)$,

$$\int_{X_i} v^2\, d\Gamma_i(\varphi_i, \varphi_i) \le C_1 \int_{X_i} \varphi_i^2\, d\Gamma_i(v, v) + C_2 \int_{\mathrm{supp}\{\varphi_i\}} v^2\, dm_i. \tag{7.4}$$

Here Γ_i represents the energy measure on X_i, and C_1, C_2 are the same for all factors X_i. Then for any $f \in \mathcal{D}(\mathcal{E})$, for the function φ defined as in (7.3),

$$\begin{aligned}
&\int_X f^2\, d\Gamma(\varphi, \varphi)\\
&= \sum_{i=1}^{N} a_{ii} \int_{\prod_{k\neq i} X_k} \left(\int_{X_i} f^2\, d\Gamma_i(\varphi_i, \varphi_i) \right) \left(\prod_{\substack{j=1\\ j\neq i}}^{N} \varphi_j(x_j) \right)^2 d\left(\bigotimes_{k\neq i} m_k \right)\\
&\le \sum_{i=1}^{N} a_{ii} \Bigg[C_1 \int_{\prod_{k\neq i} X_k} \left(\int_{X_i} (\varphi_i)^2\, d\Gamma_i(f, f) \right) \left(\prod_{\substack{j=1\\ j\neq i}}^{N} \varphi_j(x_j) \right)^2 d\left(\bigotimes_{k\neq i} m_k \right)\\
&\quad + C_2 \int_{\mathrm{supp}\{\varphi\}} f^2\, dm \Bigg].
\end{aligned}$$

In the last line we bounded the product of φ_i's by 1. Then since

$$\int_X \varphi^2\, d\Gamma(f, f) = \sum_{i=1}^{\infty} a_{ii} \int_{\prod_{k\neq i} X_k} \int_{X_i} \varphi^2\, d\Gamma_i(f, f)\, d\left(\bigotimes_{k\neq i} m_k \right),$$

we conclude that

$$\int_X f^2\, d\Gamma(\varphi, \varphi) \le C_1 \int_X \varphi^2\, d\Gamma(f, f) + \left(\sum_{i=1}^{N} a_{ii} \right) C_2 \int_{\mathrm{supp}\{\varphi\}} f^2\, dm.$$

Thus these infinite product spaces satisfy Assumption 3. Using cylindric open sets to build an exhaustion of X, we can also easily check that these infinite product spaces satisfy Assumption 6. By Lemma 6, these spaces satisfy the L^2 Gaussian type upper bound, given that each factor X_i satisfies the hypotheses of the lemma. We remark that here we do not have additional requirements on the coefficient matrix $(a_{ii})_{i=1}^{\infty}$ except that all $a_{ii} > 0$.

Remark 15 On infinite dimensional compact groups, when the Laplacian L is bi-invariant, one can define more function spaces associated with L that capture the smoothness of functions and define corresponding distributional solutions of the heat equation $(\partial_t + L)u = 0$. These are broader classes of solutions than the local weak solutions we consider in this paper. In the new settings one can consider the time regularity and other spatial regularity properties of the distributional solutions of the heat equation, under more assumptions on the associated heat (convolution) semigroup, cf. [6, 7]. In [21], the authors show that for these bi-invariant Laplacians L and other left-invariant Laplacians $\widetilde{L}$ that have comparable Dirichlet forms with L, the distributional solutions are smooth, with repeated time and spatial derivatives belonging to the function spaces associated with L and $\widetilde{L}$. These results provide generalizations of the results in [7] and describe hypoellipticiy type properties of $\partial_t + L$.

8 The Weak Gaussian Bound and other Lemmas

8.1 The L^2 Type Gaussian Bound

In this subsection we prove an L^2 Gaussian type upper bound assuming the existence of cut-off functions satisfying (3.2) with $C_2(C_1, U, V) = C_1^{-\alpha} C(U, V)$ for some $\alpha > 0$ and $C(U, V) > 0$. Our proof is a modification of the classical proof of L^2 Gaussian bound when there are enough cut-off functions with bounded gradient. For references that discuss about stronger (sub)-Gaussian estimates under stronger assumptions, we mention [11] and [30]. The last part in this subsection about transitioning to estimates on derivatives of the heat semigroup is a straightforward modification of the methods in [10].

The following is the main lemma for L^2 Gaussian type upper bound. Its proof is close to for example the beginning part of the proof in [30].

Lemma 6 *Let (X, m) be a metric measure space and $(\mathcal{E}, \mathcal{F})$ be a symmetric regular local Dirichlet form on X. Assume that the Dirichlet space satisfies Assumption 3 and that for any precompact open sets U, V with disjoint closures, C_2 in (3.2) is of the form $C_2 = C_1^{-\alpha} C(U, V)$ for some $\alpha > 0$ and $C(U, V) > 0$. Then for any such open sets U, V, for any $f, g \in L^2(X)$ with supp$\{f\} \subset U$, supp$\{g\} \subset V$, for any $t > 0$,*

$$\left|\langle H_t f,\ g\rangle_{L^2(X)}\right| \le \exp\left\{-\frac{1}{2}\left(\frac{1}{2C(U, V)t}\right)^{\frac{1}{1+2\alpha}}\right\} \|f\|_{L^2(X)} \|g\|_{L^2(X)} . \tag{8.1}$$

When there exist enough nice cut-off functions with bounded gradient, Lemma 6 is a classical result obtained from the so-called Davies' Method. We adapt it to include the case when there only exist nice cut-off functions with bounded energy (as specified in the statement above). In the proof below we refer to the cut-off functions (that equal to 1 on U and 0 on V) corresponding to some C_1, C_2 with $C_2 = C_1^{-\alpha} C(U, V)$ in short as nice cut-off functions.

Proof For any fixed $\lambda > 0$, any nice cut-off function ϕ, define the perturbed semigroup $\left(H_t^{\lambda\phi}\right)_{t>0}$ by

$$H_t^{\lambda\phi} f := e^{-\lambda\phi} H_t \left(e^{\lambda\phi} f\right).$$

For any $f, g \in L^2(X)$ with supp$\{f\} \subset U$, supp$\{g\} \subset V$, first observe that

$$\left|\left\langle H_t^{\lambda\phi} f,\ g\right\rangle_{L^2(X)}\right| = e^{\lambda} \left|\langle H_t f,\ g\rangle_{L^2(X)}\right| . \tag{8.2}$$

On the other hand,

$$\left|\left\langle H_t^{\lambda\phi} f,\ g\right\rangle_{L^2(X)}\right| \le \left\|H_t^{\lambda\phi} f\right\|_{L^2(X)} \|g\|_{L^2(X)} .$$

We estimate $\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}$ by looking at its (square's) time derivative first.

$$\begin{aligned}
\frac{d}{dt}\left(\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}^2\right) &= \int_X 2\left(H_t^{\lambda\phi} f\right)\frac{d}{dt} H_t^{\lambda\phi} f\, dm \\
&= \int_X 2\left(H_t^{\lambda\phi} f\right) e^{-\lambda\phi}\frac{d}{dt} H_t\left(e^{\lambda\phi} f\right) dm = -2\mathcal{E}\left(e^{-\lambda\phi} H_t^{\lambda\phi} f,\ e^{\lambda\phi} H_t^{\lambda\phi} f\right) \\
&= -2\mathcal{E}\left(H_t^{\lambda\phi} f,\ H_t^{\lambda\phi} f\right) + 2\lambda^2 \int_X \left(H_t^{\lambda\phi} f\right)^2 d\Gamma(\phi,\phi). \qquad (8.3)
\end{aligned}$$

Because ϕ is a nice cut-off function associated with C_1, C_2, we have

$$\begin{aligned}
&\int_X \left(H_t^{\lambda\phi} f\right)^2 d\Gamma(\phi,\phi) \\
&\le C_1 \int_X \phi^2\, d\Gamma\left(H_t^{\lambda\phi} f,\ H_t^{\lambda\phi} f\right) + C_2 \int_{\mathrm{supp}\{\phi\}} \left(H_t^{\lambda\phi} f\right)^2 dm \\
&\le C_1 \mathcal{E}\left(H_t^{\lambda\phi} f,\ H_t^{\lambda\phi} f\right) + C_2 \int_{\mathrm{supp}\{\phi\}} \left(H_t^{\lambda\phi} f\right)^2 dm.
\end{aligned}$$

Substituting this bound back to (8.3) gives

$$\begin{aligned}
\frac{d}{dt}\left(\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}^2\right) &= -2\mathcal{E}\left(H_t^{\lambda\phi} f,\ H_t^{\lambda\phi} f\right) + 2\lambda^2 \int_X \left(H_t^{\lambda\phi} f\right)^2 d\Gamma(\phi,\phi) \\
&\le \left(-2 + 2\lambda^2 C_1\right)\mathcal{E}\left(H_t^{\lambda\phi} f,\ H_t^{\lambda\phi} f\right) + 2\lambda^2 C_2 \int_{\mathrm{supp}\{\phi\}} \left(H_t^{\lambda\phi} f\right)^2 dm.
\end{aligned}$$

When $-2 + 2\lambda^2 C_1 \le 0$ ($C_1 \le 1/\lambda^2$), we can drop the first term and get

$$\frac{d}{dt}\left(\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}^2\right) \le 2\lambda^2 C_2 \left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}^2.$$

Observe that at $t = 0$, $\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}^2 = \|f\|_{L^2(X)}^2$, so Gronwall's inequality gives

$$\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)}^2 \le \|f\|_{L^2(X)}^2 \exp\left(2\lambda^2 C_2 t\right).$$

Combining this with (8.2), we have

$$\left|\langle H_t f,\ g\rangle_{L^2(X)}\right| \le e^{-\lambda}\left\| H_t^{\lambda\phi} f \right\|_{L^2(X)} \|g\|_{L^2(X)} \le \|f\|_{L^2(X)}\|g\|_{L^2(X)} \exp\left(-\lambda + \lambda^2 C_2 t\right).$$

Take ϕ corresponding to $C_1 = 1/\lambda^2$ and let

$$\lambda = \left(\frac{1}{2C(U,V)t}\right)^{\frac{1}{1+2\alpha}}.$$

As $C_2 = C_1^{-\alpha} C(U,V)$, we have $\lambda = 2\lambda^2 C_2 t$, and

$$\left|\langle H_t f,\ g\rangle_{L^2(X)}\right| \le \|f\|_{L^2(X)}\|g\|_{L^2(X)} \exp\left\{-\frac{1}{2}\left(\frac{1}{2C(U,V)t}\right)^{\frac{1}{1+2\alpha}}\right\}.$$

$\square$

Remark 16 When C_2 satisfies the more explicit dependence $C_2(C_1, U, V) = CC_1^{-\alpha} d_X(U,V)^{-\beta}$ for some $C, \alpha, \beta > 0$ and some distance d_X on X that defines the same topology, substituting $C(U,V) = d_X(U,V)^{-\beta}$ in the above L^2 Gaussian type bound gives the L^2 version of the sub-Gaussian upper bound. For example, for fractals with walk dimension d_w, $C_2 = C_1^{1-d_w/2} d_X(U, V)^{-d_w}$ up to a multiplicative constant (see, for example, [30, Lemma 2.1]), then in our expression, $\alpha = (d_w/2) - 1$, $\beta = d_w$, and the exponential term in the upper bound for $\left|\langle H_t f, g\rangle_{L^2(X)}\right|$ is $\exp\left\{-c\left(\frac{d_X(U,V)^{d_w}}{t}\right)^{1/(d_w-1)}\right\}$ for some constant $c > 0$.

Next we estimate $\left|\left\langle \partial_t^k H_t f, g\right\rangle_{L^2(X)}\right|$ where $k \in \mathbb{N}_+$. The estimate essentially follows from a straightforward adaptation of Proposition 2.2 in [10]. For another approach on obtaining estimates on time derivatives of $\langle H_t f, g\rangle_{L^2(X)}$, see [13]. We first record a lemma. In the following, $\mathbb{C}_+$ denotes the right half plane $\mathbb{C}_+ := \{z \in \mathbb{C} \,|\, \mathrm{Re}(z) > 0\}$; $\mathbb{R}_+ := (0, \infty)$.

Lemma 7 *Suppose that F is an analytic function on $\mathbb{C}_+$. Assume that, for given numbers $A, B, \gamma > 0$,*

$$|F(z)| \le B, \ \forall z \in \mathbb{C}_+,$$

and for some $0 < a \le 1$,

$$|F(t)| \le Ae^{at} e^{-\left(\frac{\gamma}{t}\right)^a}, \ \forall t \in \mathbb{R}_+.$$

Then

$$|F(z)| \le B \exp\left(-\mathrm{Re}\left[\left(\frac{\gamma}{z}\right)^a\right]\right), \ \forall z \in \mathbb{C}_+. \tag{8.4}$$

When $a = 1$, this is exactly Proposition 2.2 in [10]. The proof of Lemma 7 is close to that of the proposition in [10] (essentially, replace ζ with ζ^a in that proof), and we omit it here.

Lemma 8 *Under the hypotheses in Lemma 6, for any $f, g \in L^2(X)$ with supp$\{f\} \subset U$, supp$\{g\} \subset V$, where U, V are precompact open sets with disjoint closures, for any $n \in \mathbb{N}_+$, $t > 0$,*

$$\begin{aligned}&\left|\left\langle \partial_t^n H_t f, \ g\right\rangle_{L^2(X)}\right| \\ &\quad \le n! \frac{2^n}{t^n} \|f\|_{L^2(X)} \|g\|_{L^2(X)} \exp\left\{-\frac{1}{4}\left(\frac{1}{3C(U,V)t}\right)^{\frac{1}{1+2\alpha}}\right\}.\end{aligned} \tag{8.5}$$

Proof For $t > 0$, set $F(t) := \langle H_t f, g\rangle_{L^2(X)}$. By spectral calculus, for any $z \in \mathbb{C}$ with $\mathrm{Re}(z) > 0$,

$$H_z v = \int_0^{+\infty} e^{-z\lambda} dE_\lambda v$$

is well-defined for all $v \in L^2(X)$, hence $F(z)$ can be analytically extended to $z \in \mathbb{C}_+$. Moreover,

$$\|H_z f\|^2_{L^2(X)} = \int_0^\infty e^{-2\,\mathrm{Re}(z)\lambda} \, d\langle E_\lambda f, f\rangle_{L^2(X)} \le \|f\|^2_{L^2(X)},$$

so $F(z)$ satisfies $|F(z)| \le \|f\|_{L^2(X)} \|g\|_{L^2(X)}$. By Lemma 6, for $t > 0$,

$$|F(t)| \le \exp\left\{-\frac{1}{2}\left(\frac{1}{2C(U,V)t}\right)^{\frac{1}{1+2\alpha}}\right\} \|f\|_{L^2(X)} \|g\|_{L^2(X)}.$$

So by Lemma 7,

$$|F(z)| \le \|f\|_{L^2(X)} \|g\|_{L^2(X)} \exp\left(-\mathrm{Re}\left[\left(\frac{\gamma}{z}\right)^{\frac{1}{1+2\alpha}}\right]\right), \tag{8.6}$$

where $\gamma = (4^{1+\alpha}C(U,V))^{-1}$.

Recall that we have by Cauchy's integral formula the expression of the n-th derivative of $F(z)$ using the integral over some circle $\mathcal{C}$ with radius r around z,

$$F^{(n)}(z) = \frac{n!}{2\pi i}\oint_{\mathcal{C}} \frac{F(\xi)}{(\xi - z)^{n+1}}\, d\xi = \frac{n!}{2\pi}\int_0^{2\pi} \frac{F(z + re^{i\theta})}{r^n e^{in\theta}}\, d\theta. \tag{8.7}$$

Consider $z = t \in \mathbb{R}_+$. Take for example $r = t/2$. Then (8.6) gives the bound

$$\begin{aligned}\left|F\left(t + \frac{t}{2}e^{i\theta}\right)\right| &\le \|f\|_{L^2(X)} \|g\|_{L^2(X)} \exp\left(-\mathrm{Re}\left[\left(\frac{\gamma}{t + \frac{t}{2}e^{i\theta}}\right)^{\frac{1}{1+2\alpha}}\right]\right)\\ &\le \|f\|_{L^2(X)} \|g\|_{L^2(X)} \exp\left\{-\frac{1}{2}\left(\frac{2\gamma}{3t}\right)^{\frac{1}{1+2\alpha}}\right\}.\end{aligned}$$

Indeed, suppose

$$\frac{1}{1 + \frac{1}{2}e^{i\theta}} = \frac{1 + \frac{1}{2}e^{-i\theta}}{1 + \frac{1}{4} + \cos\theta} = re^{i\varphi}.$$

Then $|\varphi| \le \pi/6$ and $r = \sqrt{(\frac{5}{4} + \cos\theta)^{-1}} \ge 2/3$. Then

$$\mathrm{Re}\left[\left(\frac{1}{1 + \frac{1}{2}e^{i\theta}}\right)^{\frac{1}{1+2\alpha}}\right] = r^{\frac{1}{1+2\alpha}} \cos\left(\frac{\varphi}{1+2\alpha}\right) \ge r^{\frac{1}{1+2\alpha}} \cos\varphi > \frac{1}{2}\left(\frac{2}{3}\right)^{\frac{1}{1+2\alpha}}.$$

Substituting the above bound of $\left|F(t + \frac{t}{2}e^{i\theta})\right|$ in (8.7), we get

$$\left|F^{(n)}(t)\right| = \left|\left\langle \partial_t^n H_t f,\ g\right\rangle_{L^2(X)}\right| \le n!\frac{2^n}{t^n} \|f\|_{L^2(X)} \|g\|_{L^2(X)} \exp\left\{-\frac{1}{2}\left(\frac{2\gamma}{3t}\right)^{\frac{1}{1+2\alpha}}\right\}.$$

Plugging in the expression of γ gives (8.5). □

In the proofs in previous sections, the exact form of the upper bounds is not important, we only need the property that the upper bound, divided by any positive power of t, tends to 0 as t tends to 0. So we use Assumption 4.

8.2 Other Lemmas

In this subsection we prove Lemma 3 on existence of nice cut-off functions for general pairs of open sets. Starting with the existence of nice cut-off functions for pairs in a topological basis $\mathcal{TB}$ in the sense of Assumption 3, we now construct nice cut-off functions for any pair of open sets $V \Subset U$ (Lemma 3).

In the next two lemmas we first discuss properties of the sum and product of two nice cut-off functions. By taking maximum if necessary, we assume that all cut-off functions correspond to the same C_1, C_2.

Lemma 9 (**sum of nice cut-off functions**) *Let η_1, η_2 be two nice cut-off functions for some $V_1 \Subset U_1$, $V_2 \Subset U_2$, respectively, where V_1, U_1, V_2, U_2 are open subsets of X. Suppose η_1, η_2 both correspond to C_1, C_2. Then their sum $\eta := \eta_1 + \eta_2$ satisfies that for any $v \in \mathcal{F}$,*

$$\begin{aligned}&\int_X v^2\, d\Gamma(\eta_1+\eta_2,\ \eta_1+\eta_2)\\ &\le\ 2C_1\int_X (\eta_1+\eta_2)^2\, d\Gamma(v,v) + 4C_2\int_{\operatorname{supp}\{\eta_1+\eta_2\}} v^2\, dm.\end{aligned}$$

Proof The energy measure $d\Gamma(\eta_1+\eta_2,\ \eta_1+\eta_2)$ equals

$$d\Gamma(\eta_1+\eta_2,\ \eta_1+\eta_2) = d\Gamma(\eta_1,\eta_1) + 2d\Gamma(\eta_1,\eta_2) + d\Gamma(\eta_2,\eta_2).$$

By applying the Cauchy-Schwartz inequality (2.1), we get that for any $v \in \mathcal{F}$,

$$\begin{aligned}&\int_X v^2\, d\Gamma(\eta_1+\eta_2,\ \eta_1+\eta_2) \le 2\int_X v^2\, d\Gamma(\eta_1,\eta_1) + 2\int_X v^2\, d\Gamma(\eta_2,\eta_2)\\ &\le\ 2C_1\int_X (\eta_1+\eta_2)^2\, d\Gamma(v,v) + 4C_2\int_{\operatorname{supp}\{\eta_1+\eta_2\}} v^2\, dm.\end{aligned}$$

The last line follows from that η_1, η_2 are nice cut-off functions corresponding to C_1, C_2; $\eta_1, \eta_2 \ge 0$; $\operatorname{supp}\{\eta_1\},\ \operatorname{supp}\{\eta_2\} \subset \operatorname{supp}\{\eta_1+\eta_2\}$. $\square$

In general, by induction, given k nice cut-off functions $\eta_1, \ldots, \eta_k$ corresponding to C_1, C_2, their sum satisfies that for any $v \in \mathcal{F}$,

$$\begin{aligned}&\int_X v^2\, d\Gamma(\eta_1+\cdots+\eta_k,\ \eta_1+\cdots+\eta_k)\\ &\le\ kC_1\int_X (\eta_1+\cdots+\eta_k)^2\, d\Gamma(v,v) + k^2C_2\int_{\operatorname{supp}\{\eta_1+\cdots+\eta_k\}} v^2\, dm. \qquad (8.8)\end{aligned}$$

We can then normalize the sum by dividing by k to get a nice cut-off function for the pair $\bigcap_{i=1}^k V_i \Subset \bigcup_{i=1}^k U_i$.

Lemma 10 (product of nice cut-off functions) *Let η_1, η_2 be two nice cut-off functions for some $V_1 \Subset U_1$, $V_2 \Subset U_2$, respectively, where V_1, U_1, V_2, U_2 are open subsets of X. Suppose η_1, η_2 both correspond to C_1, C_2, and $0 < C_1 < 1/4$. Then the product function $\eta := \eta_1\eta_2$ is still a nice cut-off function satisfying*

$$\begin{aligned}&\int_X v^2\, d\Gamma(\eta_1\eta_2,\ \eta_1\eta_2)\\ &\le\ 16C_1\int_X (\eta_1\eta_2)^2\, d\Gamma(v,v) + 8C_2\int_{\operatorname{supp}\{\eta_1\eta_2\}} v^2\, dm. \qquad (8.9)\end{aligned}$$

Proof By the product rule for the energy measure,

$$d\Gamma(\eta_1\eta_2,\ \eta_1\eta_2) = \eta_1^2\, d\Gamma(\eta_2,\eta_2) + 2\eta_1\eta_2\, d\Gamma(\eta_1,\eta_2) + \eta_2^2\, d\Gamma(\eta_1,\eta_1).$$

Then by the Cauchy-Schwartz inequality (2.1), for any $v \in \mathcal{F}$,

$$\int_X v^2\, d\Gamma(\eta_1\eta_2,\ \eta_1\eta_2) \le 2\int_X v^2\eta_1^2\, d\Gamma(\eta_2,\eta_2) + 2\int_X v^2\eta_2^2\, d\Gamma(\eta_1,\eta_1). \tag{8.10}$$

Because η_1, η_2 are associated with C_1, C_2, for any $\beta > 0$,

$$\begin{aligned}
&\int_X v^2\eta_1^2\, d\Gamma(\eta_2,\eta_2) + \int_X v^2\eta_2^2\, d\Gamma(\eta_1,\eta_1) \\
&\le C_1\left[\int_X \eta_2^2\, d\Gamma(\eta_1 v,\ \eta_1 v) + \int_X \eta_1^2\, d\Gamma(\eta_2 v,\ \eta_2 v)\right] + 2C_2\int_{\mathrm{supp}\{\eta_1\eta_2\}} v^2\, dm \\
&\le C_1\left[2(1+\beta)\int_X \eta_1^2\eta_2^2\, d\Gamma(v,v) + \left(1+\frac{1}{\beta}\right)\int_X \eta_1^2 v^2\, d\Gamma(\eta_2,\eta_2)\right. \\
&\quad \left. + \left(1+\frac{1}{\beta}\right)\int_X \eta_2^2 v^2\, d\Gamma(\eta_1,\eta_1)\right] + 2C_2\int_{\mathrm{supp}\{\eta_1\eta_2\}} v^2\, dm.
\end{aligned}$$

The second inequality is obtained by expanding $d\Gamma(\eta_1 v,\ \eta_1 v)$ and $d\Gamma(\eta_2 v,\ \eta_2 v)$ using the product rule, then applying the Cauchy-Schwartz inequality (2.1). So

$$\begin{aligned}
&\left(1 - C_1\left(1+\frac{1}{\beta}\right)\right)\left[\int_X v^2\eta_1^2\, d\Gamma(\eta_2,\eta_2) + \int_X v^2\eta_2^2\, d\Gamma(\eta_1,\eta_1)\right] \\
&\le 2C_1\,(1+\beta)\int_X \eta_1^2\eta_2^2\, d\Gamma(v,v) + 2C_2\int_{\mathrm{supp}\{\eta_1\eta_2\}} v^2\, dm.
\end{aligned}$$

For $C_1 < 1/4$, we can take $\beta = 1$, then

$$\frac{2C_1(1+\beta)}{1 - C_1\left(1+\frac{1}{\beta}\right)} = \frac{4C_1}{1-2C_1} < 8C_1,$$

and the above inequality becomes

$$\begin{aligned}
&\int_X v^2\eta_1^2\, d\Gamma(\eta_2,\eta_2) + \int_X v^2\eta_2^2\, d\Gamma(\eta_1,\eta_1) \\
&\le 8C_1\int_X \eta_1^2\eta_2^2\, d\Gamma(v,v) + 4C_2\int_{\mathrm{supp}\{\eta_1\eta_2\}} v^2\, dm.
\end{aligned} \tag{8.11}$$

Combining (8.10) and (8.11), we get (8.9). □

To extend Assumption 3, we use a construction similar to the standard construction of partitions of unity to obtain cut-off functions for general pairs of open sets and then check that the so-obtained functions satisfy (3.1). We first state the following lemma on using open sets in the basis $\mathcal{TB}$ to cover any compact set.

Lemma 11 *For any compact set $K \subset X$ and any open neighborhood U of K ($K \subset U \Subset X$), there exist two finite open covers $\mathcal{C}_1 = \{U_1, U_2, \ldots, U_n\}$ and $\mathcal{C}_2 = \{V_1, V_2, \ldots, V_m\}$, such that all U_j, V_i are elements in $\mathcal{TB}$; $K \subset \bigcup_{i=1}^m V_i \subset \bigcup_{j=1}^n U_j \subset U$; $\mathcal{C}_2$ is subordinate to $\mathcal{C}_1$, i.e., for any $V_i \in \mathcal{C}_2$, there exists some $U_j \in \mathcal{C}_1$ such that $V_i \Subset U_j$.*

Proof For any point $p \in K$, there exists an open neighborhood $U_p \in \mathcal{TB}$ such that $p \in U_p \Subset U$ since $\mathcal{TB}$ is a topology basis and X is regular (to ensure there is some U_p that is precompact in U). Then $\{U_p \mid p \in K\}$ is an open cover of K, which has a finite sub-cover $\mathcal{C}_1 = \{U_{p_1}, U_{p_2}, \ldots, U_{p_n}\}$. We rename U_{p_j} as U_j.

Now we construct $\mathcal{C}_2$ from $\mathcal{C}_1$. For any point $p \in K$, there exists some U_j, $j = 1, 2, ..., n$, such that $p \in U_j$. Then there exists some smaller open neighborhood $V_p \in \mathcal{TB}$ such that $p \in V_p \Subset U_j$, $\left\{V_p \,\middle|\, p \in K\right\}$ is an open cover of K. Let $\left\{V_{p_1}, V_{p_2}, \ldots, V_{p_m}\right\}$ be a finite sub-cover, then this gives the $\mathcal{C}_2$ open cover we wanted, after renaming V_{p_i} as V_i. □

Next we proceed to prove the lemma on the automatic extension of the applicability of Assumption 3 from pairs of open sets in a topological basis to all open sets.

Proof of Lemma 3 For any pair of open sets $V \Subset U$, for any $0 < C_1 < 1$, we want to construct a nice cut-off function ψ for the pair $V \subset U$ corresponding to C_1 in (3.1). Pick another open set V' such that $V \Subset V' \Subset U \Subset X$. Applying Lemma 11 to the compact set $K = \overline{V'}$ with open neighborhood U, we get two finite open covers $\mathcal{C}_1 = \{O_1, \ldots, O_n\}$ and $\mathcal{C}_2 = \{\Omega_1, \ldots \Omega_m\}$ such that $\mathcal{C}_2$ is subordinate to $\mathcal{C}_1$, and that both cover $\overline{V'}$ and are contained in U. Applying Lemma 11 to the compact set $\overline{U} \setminus V'$ with open neighborhood $X \setminus \overline{V}$, we get two more finite open covers $\mathcal{C}_1' = \left\{O_1', \ldots, O_{n'}'\right\}$ and $\mathcal{C}_2' = \left\{\Omega_1', \ldots \Omega_{m'}'\right\}$, such that $\mathcal{C}_2'$ is subordinate to $\mathcal{C}_1'$, both $\mathcal{C}_1', \mathcal{C}_2'$ cover $\overline{U} \setminus V'$, and are contained in $X \setminus \overline{V}$.

For any $0 < C < 1$, apply Assumption 3 to each pair $\Omega_i \Subset O_j$ and $\Omega_i' \Subset O_j'$. Because all $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_1', \mathcal{C}_2'$ are finite covers, there are finitely many nice cut-off functions $\{\eta_1, \ldots, \eta_r\}$ and $\{\varphi_1, \ldots, \varphi_k\}$ for pairs $\Omega_i \Subset O_j$ and $\Omega_i' \Subset O_j'$, respectively, where all cut-off functions correspond to $C_1 =: C$ in (3.1). Let

$$\eta := \eta_1 + \cdots + \eta_r, \quad \varphi := \sum_{i=1}^{k} \varphi_i + \sum_{j=1}^{r} \eta_j.$$

Then $1 \le \varphi \le k + r$ on U, and $\varphi = \eta$ on $\overline{V}$, since all φ_i's vanish on $\overline{V}$. Hence η/φ is well-defined on U and becomes 0 before it reaches the boundary of U since η is supported in U. By extending the quotient by 0 outside U, we obtain a function ψ satisfying

$$\psi(x) = \begin{cases} \eta/\varphi, & x \in U, \\ \\ 0, & x \in U^c \end{cases} = \begin{cases} 1, & x \in \overline{V}, \\ \text{between 0 and 1}, & x \in U \setminus \overline{V}, \\ 0, & x \in U^c. \end{cases}$$

Hence it remains to show that ψ satisfies (3.1). By the lemmas on the sum and product of nice cut-off functions (Lemma 9 and Lemma 10), we only need to show $1/\varphi$ satisfies (3.1) for $u \in \mathcal{F}$ with support in U (since ψ is supported in U). For any $u \in \mathcal{F}$ with support in U,

$$\begin{aligned} \int_X u^2 \, d\Gamma(1/\varphi, 1/\varphi) &= \int_X u^2 \left(-\frac{1}{\varphi^2}\right)^2 d\Gamma(\varphi, \varphi) \\ &\le \int_X u^2 \, d\Gamma(\varphi, \varphi) \le C' \int_X \varphi^2 \, d\Gamma(u, u) + C_2 \int_{\mathrm{supp}\{\varphi\}} u^2 \, dm, \end{aligned}$$

where $C' = (k + r)C$ is by (8.8) and our definition of φ; C_2 can be computed correspondingly. Moreover, since $1 \le \varphi \le k + r$, $1 \le \varphi^2 \le (k + r)^2$, we get that $\varphi \le (k + r)^2/\varphi$ on U, hence

$$\int_X u^2 \, d\Gamma(1/\varphi, 1/\varphi) \le C' \int_X \frac{(k + r)^4}{\varphi^2} \, d\Gamma(u, u) + C_2 \int_{\mathrm{supp}\{\varphi\}} u^2 \, dm,$$

which is indeed of the form (3.1). By picking a proper C, ψ would correspond to the given C_1 in (3.1). □

Data Availability Data sharing not applicable to this paper as no data sets were generated or analysed in the current study.

References

1. Andres, S., Barlow, M.T.: Energy inequalities for cutoff functions and some applications. J. Reine Angew. Math. **699**, 183–215 (2015). https://doi.org/10.1515/crelle-2013-0009
2. Ariyoshi, T., Hino, M.: Small-time asymptotic estimates in local Dirichlet spaces. Electron. J. Probab. **10**(37), 1236–1259 (2005). https://doi.org/10.1214/EJP.v10-286
3. Barlow, M.T., Murugan, M.: Stability of the elliptic Harnack inequality. Ann. of Math. (2) **187**(3), 777–823 (2018). https://doi.org/10.4007/annals.2018.187.3.4
4. Bendikov, A., Saloff-Coste, L.: Elliptic diffusions on infinite products. J. Reine Angew. Math. **493**, 171–220 (1997). https://doi.org/10.1515/crll.1997.493.171
5. Bendikov, A., Saloff-Coste, L.: Invariant local Dirichlet forms on locally compact groups. Ann. Fac. Sci. Toulouse Math. (6) **11**(3), 303–349 (2002)
6. Bendikov, A., Saloff-Coste, L.: Spaces of smooth functions and distributions on infinite-dimensional compact groups. J. Funct. Anal. **218**(1), 168–218 (2005). https://doi.org/10.1016/j.jfa.2004.06.006
7. Bendikov, A., Saloff-Coste, L.: Hypoelliptic bi-invariant Laplacians on infinite dimensional compact groups. Canad. J. Math. **58**(4), 691–725 (2006). https://doi.org/10.4153/CJM-2006-029-9
8. Bendikov, A., Saloff-Coste, L., Salvatori, M., Woess, W.: The heat semigroup and Brownian motion on strip complexes. Adv. Math. **226**(1), 992–1055 (2011). https://doi.org/10.1016/j.aim.2010.07.014
9. Colding, T.H., Minicozzi, W.P. II.: Optimal bounds for ancient caloric functions. Duke Math. J. **170**(18), 4171–4182 (2021). https://doi.org/10.1215/00127094-2021-0015
10. Coulhon, T., Sikora, A.: Gaussian heat kernel upper bounds via the Phragmén-Lindelöf theorem. Proc. Lond. Math. Soc. (3) **96**(2), 507–544 (2008). https://doi.org/10.1112/plms/pdm050
11. Davies, E.B.: Heat Kernels and Spectral Theory. Cambridge Tracts in Mathematics, vol. 92. Cambridge University Press, Cambridge (1989). https://doi.org/10.1017/CBO9780511566158
12. Davies, E.B.: Heat kernel bounds, conservation of probability and the Feller property, pp. 99–119. https://doi.org/10.1007/BF02790359. Festschrift on the occasion of the 70th birthday of Shmuel Agmon (1992)
13. Davies, E.B.: Non-Gaussian aspects of heat kernel behaviour. J. London Math. Soc. (2) **55**(1), 105–125 (1997). https://doi.org/10.1112/S0024610796004607
14. Eells, J., Fuglede, B.: Harmonic Maps between Riemannian Polyhedra. Cambridge Tracts in Mathematics, vol. 142. Cambridge University Press, Cambridge (2001). With a preface by M. Gromov
15. Eldredge, N., Saloff-Coste, L.: Widder's representation theorem for symmetric local Dirichlet spaces. J. Theoret. Probab. **27**(4), 1178–1212 (2014). https://doi.org/10.1007/s10959-013-0484-1
16. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet forms and symmetric Markov processes, De Gruyter Studies in Mathematics, vol. 19, extended edn. Walter de Gruyter & Co., Berlin (2011)
17. Gyrya, P., Saloff-Coste, L.: Neumann and Dirichlet heat kernels in inner uniform domains. Astérisque (336) viii+144 (2011)
18. Hino, M.: On singularity of energy measures on self-similar sets. Probab. Theory Related Fields **132**(2), 265–290 (2005). https://doi.org/10.1007/s00440-004-0396-1
19. Hino, M., Ramírez, J.A.: Small-time Gaussian behavior of symmetric diffusion semigroups. Ann. Probab. **31**(3), 1254–1295 (2003). https://doi.org/10.1214/aop/1055425779
20. Hou, Q., Saloff-Coste, L.: Rough hypoellipticity for the heat equation in Dirichlet spaces (2020)
21. Hou, Q., Saloff-Coste, L.: Perturbation results concerning gaussian estimates and hypoellipticity for left-invariant laplacians on compact groups (2021)
22. Iwasawa, K.: On some types of topological groups. Ann. of Math. (2) **50**, 507–558 (1949). https://doi.org/10.2307/1969548
23. Kajino, N., Murugan, M.: On singularity of energy measures for symmetric diffusions with full off-diagonal heat kernel estimates. Ann. Probab. **48**(6), 2920–2951 (2020). https://doi.org/10.1214/20-AOP1440
24. Kusuoka, S., Stroock, D.: Applications of the Malliavin calculus. II. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32**(1), 1–76 (1985)
25. Kuwae, K., Machigashira, Y., Shioya, T.: Sobolev spaces, Laplacian, and heat kernel on Alexandrov spaces. Math. Z. **238**(2), 269–316 (2001). https://doi.org/10.1007/s002090100252
26. Lierl, J.: Local behavior of solutions of quasilinear parabolic equations on metric spaces. arXiv:1708.06329 (2017)

27. Lierl, J.: Parabolic Harnack inequality on fractal-type metric measure Dirichlet spaces. Rev. Mat. Iberoam. **34**(2), 687–738 (2018). https://doi.org/10.4171/RMI/1001
28. Lierl, J.: Parabolic Harnack inequality for time-dependent non-symmetric Dirichlet forms. J. Math. Pures Appl. (9) **140**, 1–66 (2020). https://doi.org/10.1016/j.matpur.2020.01.001
29. Lin, F., Zhang, Q.S.: On ancient solutions of the heat equation. Comm. Pure Appl. Math. **72**(9), 2006–2028 (2019). https://doi.org/10.1002/cpa.21820
30. Murugan, M., Saloff-Coste, L.: Davies' method for anomalous diffusions. Proc. Amer. Math. Soc. **145**(4), 1793–1804 (2017). https://doi.org/10.1090/proc/13324
31. Otsu, Y., Shioya, T.: The Riemannian structure of Alexandrov spaces. J. Differential Geom. **39**(3), 629–658 (1994). http://projecteuclid.org/euclid.jdg/1214455075
32. Pivarski, M., Saloff-Coste, L.: Small time heat kernel behavior on Riemannian complexes. New York J. Math. **14**, 459–494 (2008)
33. Saloff-Coste, L.: Uniformly elliptic operators on Riemannian manifolds. J. Differential Geom. **36**(2), 417–450 (1992). http://projecteuclid.org/euclid.jdg/1214448748
34. Sturm, K.T.: Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations. Osaka J. Math. **32**(2), 275–312 (1995). http://projecteuclid.org/euclid.ojm/1200786053
35. Sturm, K.T.: Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. J. Math. Pures Appl. (9) **75**(3), 273–297 (1996)
36. Takeda, M.: On a martingale method for symmetric diffusion processes and its applications. Osaka J. Math. **26**(3), 605–623 (1989). http://projecteuclid.org/euclid.ojm/1200781700
37. Wloka, J.: Partial Differential Equations. Cambridge University Press, Cambridge (1987). https://doi.org/10.1017/CBO9781139171755, Translated from the German by C. B. Thomas and M. J. Thomas
38. Zhang, Q.S.: A note on time analyticity for ancient solutions of the heat equation. Proc. Amer. Math. Soc. **148**(4), 1665–1670 (2020). https://doi.org/10.1090/proc/14830

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