Two-sided estimates on Dirichlet heat kernels for time-dependent parabolic operators with singular drifts in $C^{1,\alpha}$-domains

Sungwon Cho$^{a,1}$, Panki Kim$^{b,2}$, Hyein Park$^c$,*

$^a$ Department of Mathematics Education, Gwangju National University of Education, 93 Pilmunlo Bukgu, Gwangju 500-703, Republic of Korea
$^b$ Department of Mathematics and Research Institute of Mathematics, Seoul National University, Seoul 151-747, Republic of Korea
$^c$ Department of Mathematical Sciences, Korea Advanced Institute for Science and Technology, Daejeon 305-701, Republic of Korea

**Article info**

**Abstract**

In this paper, we establish sharp two-sided estimates for the Dirichlet heat kernels of a large class of time-dependent parabolic operators with singular drifts in $C^{1,\alpha}$-domain in $\mathbb{R}^d$, where $d \geq 1$ and $\alpha \in (0,1]$. Our operator is $L + \mu \cdot \nabla$, where $L$ is a time-dependent uniformly elliptic divergent operator with Dini continuous coefficients and $\mu$ is a signed measure on $(0, \infty) \times \mathbb{R}^d$ belonging to parabolic Kato class. Along the way, a gradient estimate is also established. Our method employs a combination of partial differential equations and perturbation techniques.

© 2011 Elsevier Inc. All rights reserved.

*Corresponding author.
E-mail addresses: scho@gnue.ac.kr (S. Cho), pkim@snu.ac.kr (P. Kim), hyein@kaist.ac.kr (H. Park).

$^1$ Sungwon Cho is supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0009685).

$^2$ Panki Kim is supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MEST) (2011-000251).

0022-0396/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
1. Introduction and setup

Two-sided estimates of heat kernel for parabolic operators in \( \mathbb{R}^d \) have a long history and many beautiful results have been established. In his classical results, Aronson [2] used the parabolic Harnack inequality, which was established by Moser [27], to obtain the heat kernel estimates. Later, Fabes and Stroock [15] used the heat kernel estimates to obtain parabolic Harnack inequality following Nash’s idea [28]. See [2,10,12,29] and the references therein for more details.

But, due to the complication near the boundary, two-sided estimates for (Dirichlet) heat kernel in \( C^{1,1} \)-domain have been established only recently. See [11–13] for the upper bound estimates and [34] for the lower bound estimates in \( C^{1,1} \)-domains for the Laplace operator. Regarding the strongly parabolic system, one may refer to [7,8] for the upper bounds. Recently in [6] the first named author proved the two-sided estimates for heat kernel of parabolic operators with Dini continuous diffusion coefficients and certain singular drifts, in every cylinder where the base domain \( D \) is a \( C^{1,\alpha} \)-domain lying in \( \mathbb{R}^d \) satisfying the connected line condition, which will be defined below.

Here and after, we assume that \( d \geq 1 \) and \( \alpha \in (0,1] \). Our operator \( L \) can be written

\[
L + \sum_{i=1}^{d} \mu^i \frac{\partial}{\partial x_i} , \tag{1.1}
\]

where

\[
Lu(t,x) = \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} (a_{ij}(t,x) \frac{\partial}{\partial x_j} u(t,x)) \tag{1.2}
\]

and the drift \( \mu^i \) is a signed measure in \((0,\infty) \times \mathbb{R}^d\), which belongs to the parabolic Kato class defined in Definition 3.1. Here, the operator \( \frac{d}{dt} - L \) is understood in a weak sense. Namely, \( (\frac{d}{dt} - L)u = f \) in \((t_1,t_2) \times D\) if and only if

\[
\int_{t_1}^{t_2} \int_{D} -u(t,x) \frac{\partial}{\partial t} \phi(t,x) + \sum_{i,j=1}^{d} a_{ij}(t,x) \frac{\partial}{\partial x_j} u(t,x) \frac{\partial}{\partial x_i} \phi(t,x) \, dt \, dx 
\]

\[
- \int_{(t_1,t_2) \times D} \sum_{i=1}^{d} \frac{\partial}{\partial x_i} u(t,x) \phi(t,x) \, d\mu^i(t,x) 
\]

\[
= \int_{D} \int_{t_1}^{t_2} f(t,x) \phi(t,x) \, dt \, dx
\]

for any \( \phi \in C^\infty_c((t_1,t_2) \times D) \). Here \( C^\infty_c((t_1,t_2) \times D) \) is the set of all smooth functions with compact support in \((t_1,t_2) \times D\), \( dt \, dx \) denotes the Lebesgue measure in \([0,\infty) \times \mathbb{R}^d\).

We assume that the operator \( L \) satisfies the uniform ellipticity condition, i.e., there exists a constant \( \lambda \in (0,1) \) such that for all \( (t,x) \in (0,\infty) \times \mathbb{R}^d \) and \( \xi \in \mathbb{R}^d \), we have

\[
\lambda |\xi|^2 \leq \sum_{i,j=1}^{d} a_{ij}(t,x) \xi_i \xi_j \quad \text{and} \quad \max_{|\xi| \leq 1} |a_{ij}(t,x)\xi| \leq \lambda^{-1}. \tag{1.3}
\]
We emphasize here that $a_{ij}(t,x)$ may not be symmetric. We further assume $(t,x) \rightarrow a_{ij}(t,x)$ is Dini continuous, if there exists a non-decreasing function $\psi : [0, \infty) \rightarrow [0, \infty)$ such that, for any $(t, x), (s, y) \in (0, \infty) \times \mathbb{R}^d$, 

$$\sum_{i,j=1}^{d} |a_{ij}(t,x) - a_{ij}(s,y)| \leq \psi(\sqrt{|t-s| + |x-y|}) \quad \text{and} \quad \int_{0^+} \frac{\psi(t)}{t} \, dt < \infty. \quad (1.4)$$

Here, $|x - y|$ means the Euclidean distance between $x$ and $y$, and we will use $B(x, r)$ to denote the open ball centered at $x \in \mathbb{R}^d$ with radius $r > 0$. A domain (connected and open set) $D$ in $\mathbb{R}^d$ (when $d \geq 2$) is said to be a $C^{1,\alpha}$-domain if there exist a localization radius $R_0 > 0$ and a constant $A_0 > 0$ such that for every $z \in \partial D$, there exist a $C^{1,\alpha}$-function $\phi = \phi_z : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$ satisfying $\phi(0) = 0$, $\nabla \phi(0) = (0,...,0)$, $\|\nabla \phi\|_{\infty} \leq A_0$, $|\nabla \phi(x_1) - \nabla \phi(x_2)| \leq A_0|x_1 - x_2|$ for $x_1, x_2 \in \mathbb{R}^{d-1}$, and an orthonormal coordinate system $CS_z$: $y = (y_1,...,y_{d-1},y_d) = : (\tilde{y}, y_d)$ with its origin at $z$ such that $B(z, R_0) \cap D = \{ y = (\tilde{y}, y_d) \in CS_z : |\tilde{y} | < R_0, y_d > \phi(\tilde{y}) \}$. The pair $(R_0, A_0)$ will be called the characteristics of $C^{1,\alpha}$-domain $D$. A $C^{1,\alpha}$-domain in $\mathbb{R}$ is simply an open interval or an open ray in $\mathbb{R}$. Note that a $C^{1,\alpha}$-domain can be unbounded.

We assume that our $C^{1,\alpha}$-domain $D$ in $\mathbb{R}^d$ with characteristics $(R_0, A_0)$ satisfies the following connected line condition with characteristics $(\chi_1, \chi_2)$: there exist $\chi_i > 0$, $i = 1, 2$, and $\epsilon_0 = \epsilon_0(R_0, A_0, d) > 0$ such that for any $\epsilon \leq \epsilon_0$ and $x, y \in D$ with $\text{dist}(x, \partial D) \geq \epsilon$, $\text{dist}(y, \partial D) \geq \epsilon$, one can find a length parameterized curve $l \subset D$ connecting $x$ and $y$ such that the length $|l|$ of $l$ is less than or equal to $\chi_1|y - x|$ and $\text{dist}(l(u), \partial D) \geq \chi_2 \epsilon$ for every $u \in [0, |l|]$. It is easy to check any bounded $C^{1,\alpha}$-domain and the domain above the graph of any $C^{1,\alpha}$-function satisfy connected line condition.

Here and in the sequel, we use “:=” as a way of definition. Throughout this paper, we fix a $C^{1,\alpha}$-domain $D$ satisfying the connected line condition with characteristics $(R_0, A_0)$ and $(\chi_1, \chi_2)$, respectively and let $\rho(x) := \text{dist}(x, \partial D)$ with the convention that $\rho(\cdot) \equiv \infty$ if $D = \mathbb{R}^d$.

A function $q_D(t;x,s,y)$ defined in $Q_T \times Q_T$, $Q_T := (0,T) \times D$, is called the heat kernel of $\partial_t - \mathcal{L}$ in $Q_T$ if $q_D(t;x,s,y)$ satisfies the following three conditions: For any fixed $(s, y) \in Q_T$,

$$\int_{D}^{T} q_D(t;x,s,y) \frac{\partial}{\partial t} \phi(t,x) + \sum_{i,j=1}^{d} a_{ij}(t,x) \frac{\partial}{\partial x_j} q_D(t;x,s,y) \frac{\partial}{\partial x_i} \phi(t,x) \, dt \, dx$$

$$- \sum_{i=1}^{d} \int_{(0,T) \times D} \frac{\partial}{\partial x_i} q_D(t;x,s,y) \phi(t,x) \, d\mu(t,x) = \phi(s,y) \quad \text{for any } \phi \in C_c^\infty((0, T) \times D); \quad (F1)$$

$$q_D(t;x,s,y) = 0 \quad \text{for any } x \in \partial D; \quad (F2)$$

$$\lim_{t \rightarrow s^+} \int_{D} q_D(t;x,s,y) \, dy = u_0(y) \quad \text{for any } x \in D, \, u_0 \in C_c^\infty(D). \quad (F3)$$

See also [19, Definition 3.15] and [25] for $D = \mathbb{R}^d$ for other definitions. When $D = \mathbb{R}^d$, we simply denote $q_{\mathbb{R}^d}$ by $q$.

It is easy to see that, for any fixed $(s, y) \in Q_T$ and $\phi \in C_c^\infty((s, T) \times D)$,

$$\int_{s}^{T} \int_{D} -q_D(t;x,s,y) \frac{\partial}{\partial t} \phi(t,x) + \sum_{i,j=1}^{d} a_{ij}(t,x) \frac{\partial}{\partial x_j} q_D(t;x,s,y) \frac{\partial}{\partial x_i} \phi(t,x) \, dt \, dx$$
\[-\sum_{i=1}^{d} \int_{(s,T) \times D} \frac{\partial}{\partial \chi_i} q_D(t, x; s, y) \phi(t, x) \, d\mu^i(t, x) = 0. \]  

(1.5)

Thus for any continuous function \( u_0 \) on \( D \) and fixed \( s \), the function defined by \( u(t, x) := \int_D q_D(t, x; s, y) u_0(y) \, dy \) solves the following initial value problem in a weak sense:

\[
\begin{cases}
(\partial_t - \mathcal{L})u(t, x) = 0 & \text{in } (s, T) \times D, \\
u(t, x) = 0 & \text{for } x \in \partial D, \\
\lim_{t \to s^+} u(t, x) = u_0(x). 
\end{cases}
\]  

(IVP)

For \( a, b \in \mathbb{R} \), let \( a \wedge b := \min\{a, b\} \) and \( a \vee b := \max\{a, b\} \). We denote

\[ \psi_0(t, x, y) := \left(1 \wedge \frac{\rho(x)}{\sqrt{t}}\right) \left(1 \wedge \frac{\rho(y)}{\sqrt{t}}\right) \quad \psi_1(t, y) := \left(1 \wedge \frac{\rho(y)}{\sqrt{t}}\right) \tau^{-\frac{1}{2}}, \]

when \( t > 0, x, y \in D \). If \( t \leq 0, \psi_0 \) and \( \psi_1 \) are defined to be zero. For \( c > 0, I_c \) is the Gaussian kernel:

\[ I_c(t, x; s, y) := (t - s)^{-d/2} e^{-|x-y|^2}, \]

for \( t > s \). If \( t \leq s, I_c \) is defined to be zero.

We frequently use the phrase “depending on \( \mu \) only via the rate at which \( \max_{1 \leq i \leq d} N_{\mu i}(r) \) goes to zero”, which means that the statement is true for any family of \( d \)-dimensional vector-valued signed measures \( \nu \)'s on \((0, \infty) \times \mathbb{R}^d \) with \( \max_{1 \leq i < d} N_{\mu i}(r) \leq \max_{1 \leq i \leq d} N_{\mu i}(r), r > 0 \). See Section 3.1 for the definition of the (parabolic) Kato (semi) norm \( N_{\mu i} \). The main results of this paper can be stated as follows:

**Theorem 1.1.** Let \( d \geq 1, \alpha \in (0, 1] \) and \( T > 0 \), and suppose \( D \) is a \( C^{1, \alpha} \)-domain satisfying the connected line condition with characteristics \((R_0, A_0)\) and \((\chi_1, \chi_2)\), respectively. Let \( L \) be an operator of the form

\[ Lu(t, x) = \sum_{i,j=1}^{d} \frac{\partial}{\partial \chi_i} \left( a_{ij}(t, x) \frac{\partial}{\partial \chi_j} u(t, x) \right) \quad \text{in } (0, \infty) \times D, \]

satisfying the uniform ellipticity condition with Dini continuous coefficients \( a_{ij}(t, x) \) and a vector-valued signed measure \( \mu = (\mu^1, \ldots, \mu^d) \) on \((0, \infty) \times \mathbb{R}^d \) belonging to the Kato class defined in Definition 3.1. Then the heat kernel \( q_D(t, x; s, y) \) of \[ \frac{\partial}{\partial t} - L - \mu \cdot \nabla x \] (in the initial value problem) and its gradient \( \nabla_x q_D(t, x; s, y) \) exist and are jointly continuous and the following Gaussian estimates hold: there exist positive constants \( c_i = c_i(d, \lambda, \chi_1, \chi_2, \psi, T, A_0, R_0, \alpha, \mu), i = 1, 2, \) depending on \( \mu \) only via the rate at which \( \max_{1 \leq i \leq d} N_{\mu i}(r) \) goes to zero, such that for \( x, y \in D \) and \( 0 < s < t < s + T \),

\[ c^{-1}_1 \psi_0(t-s, x, y) I^{-1}_{c^{-1}}(t, x; s, y) \leq q_D(t, x; s, y) \leq c_1 \psi_0(t-s, x, y) I^{-1}_{c_2}(t, x; s, y) \]

(1.6)

and

\[ |\nabla_x q_D(t, x; s, y)| \leq c_1 \psi_1(t-s, x, y) I^{-1}_{c_2}(t, x; s, y). \]

(1.7)

A heat kernel of \( \partial_t + L + \sum_{i=1}^{d} \mu^i \frac{\partial}{\partial \chi_i} \) of the final value problem in \((0, T) \times D \) is a Borel function \( \hat{q}_D(s, x; t, y) \) satisfying the following three conditions: For any fixed \((t, y) \in (0, T) \times D \) and any \( \phi \in C^\infty_c((0, t) \times D), \)
terplay can be extended to a large class of time-inhomogeneous parabolic operators. We show that Brownian motion. The heat kernel of such operator is the transition density function of infinitesimal generator of $X$. These require very careful and detailed estimates throughout our proofs. As in [22], the uniform convergence of heat kernels and their gradient show that the approximation scheme proposed above is also well-suited for the purpose of this paper. The uniform convergence of $q^0_D$ is essential for our approach to establish (1.6) and (1.7), and they can be regarded as stability results for the heat kernels under perturbations. The general strategy of proving the uniform convergence of heat kernels is similar to that of [22]. However, we have to overcome quite a few new difficulties due to the fact that our operators are time-dependent and the drifts are signed measures on $[0, \infty) \times \mathbb{R}^d$ (not time-dependent signed measures of the form $d\mu^1(t, x) = \mu^1(t)(dx) dt$). These require very careful and detailed estimates throughout our proofs.

Lastly, we give the probabilistic counterpart and some consequences of Theorem 1.1 in Section 4. It is well known that there are close relationships between second order elliptic differential operators and diffusion processes. For certain second order elliptic differential operator that satisfies the maximum principle, there is a diffusion process $X$ associated with it so that such operator is the infinitesimal generator of $X$. A prototype is the celebrated interplay between the Laplacian and the Brownian motion. The heat kernel of such operator is the transition density function of $X$. Such interplay can be extended to a large class of time-inhomogeneous parabolic operators. We show that...
\(\hat{q}_D\) is transition density of a time-inhomogeneous diffusion \(X\) which can be approximated by nice diffusions in the sense of weak convergence. From the time-homogeneous case, we obtain two-sided Green function estimates, Theorem 4.8, along with so-called 3G-Theorem, Corollary 4.9.

The remainder of this paper is organized as follows. In Section 2, we prove the estimate and joint continuity of the gradient of \(p^0_D\). In Section 3, we use the results on \(p^0_D\) in Section 2 and give the proof of Theorem 1.1. Section 4 discuss some probabilistic counterparts of Theorem 1.1.

Concluding the introductory section, we present some standard notations: We will follow the standard multi-index notation for derivatives: given a vector \(\mathbf{n} := (n_1, \ldots, n_d)\), where each \(n_k\) is a non-negative integer, let \(|\mathbf{n}| := n_1 + \cdots + n_d\) and denote \(D^\mathbf{n}_x u(t, x) := \frac{\partial^{n_1} u(t, x)}{\partial x_1^{n_1}} \cdots \frac{\partial^{n_d} u(t, x)}{\partial x_d^{n_d}}\) and \(|D^\mathbf{n}_x u(t, x)| := (\sum_{|\mathbf{n}|=k} |D^\mathbf{n}_x u(t, x)|^2)^{1/2}\). We define \(\nabla_x u := (\frac{\partial}{\partial x_1} u, \ldots, \frac{\partial}{\partial x_d} u)\). Note that \(|\nabla_x u(t, x)| = |D_1^0 u(t, x)|\). We use \(\partial_t, \partial_y, \partial_{y_i}, \partial_{y_j}\) instead of \(\frac{\partial}{\partial t}, \frac{\partial}{\partial y}, \frac{\partial}{\partial y_i} \\frac{\partial}{\partial y_j}\) respectively to simplify notations. We use \(c_1, c_2, \ldots\) to denote generic constants, whose exact values are not important and can be changed from one appearance to another. The labeling of the constants \(c_1, c_2, \ldots\) starts anew in the statement of each result. The values of the constants \(C_1, C_2, \ldots\) will remain the same, and the dependence of the constants \(c_1, c_2, \ldots\) on the universal constants, the dimension \(d\), and the constants \(C_1, C_2, \ldots\) will not be mentioned explicitly.

The set of strict positive natural numbers is denoted by \(\mathbb{N}\), whereas \(\mathbb{N}_0 := \mathbb{N} \cup \{0\}\). For a measurable set \(A \subset \mathbb{R}^d\), we use \(m_d(A)\) to denote the \(d\)-dimensional Lebesgue measure of \(A\). We sometimes use the convention of

\[
\mu((a, b) \times U) = \int_U \mu(U, x) = \int_{(a, b) \times U} d\mu(t, x)
\]

for every \((a, b) \times U \subseteq (-\infty, \infty) \times \mathbb{R}^d\). For the reader's convenience, we summarize in Table 1 heat kernel notations that appear frequently in this paper.

### 2. Gradient estimate

We recall that \(L\) is the operator defined in (1.2). In this section, we will prove a gradient estimate of the heat kernel \(p^0_D\) of \(\partial_t - L\) in \(Q_T = (0, T) \times D\), which will be used in Section 3.

The following two-sided bound on \(p^0_D(t, x; s, y)\) was established in [6]. We write it in a slightly different form for our purpose:

**Theorem 2.1.** Let \(T > 0\), \(\alpha \in (0, 1]\), and \(D\) be \(C^{1,\alpha}\)-domain satisfying the connected line condition with characteristics \((R_0, \Lambda_0)\) and \((\chi_1, \chi_2)\), respectively. Suppose that \(L\) is an operator of the form in (1.2) satisfying (1.3) in \(Q_T\) with Dini continuous coefficients \(a_{ij}\). Then, there exist positive constants \(C_1 = \ldots\)
\[C_1(d, \lambda, \chi_1, \chi_2, \psi, T, A_0, R_0, \alpha), \quad C_2 = C_2(d, \lambda, \chi_1, \chi_2, \psi, A_0, R_0, \alpha)\) such that for \(x, y \in D\) and \(0 \leq s < t \leq s + T,\)

\[
\frac{1}{C_1} \psi_0(t - s, x, y) \Gamma_1(t, x; s, y) \leq p^{B}_D(t, x; s, y) \leq C_1 \psi_0(t - s, x, y) \Gamma_2(t, x; s, y).
\]

The lower bound on \(p^{B}_D(t, x; s, y)\) is stated in [6] for bounded \(C^{1,\alpha}\)-domains (with connected ball condition). However, by the Harnack inequality using connected ball condition directly, the proofs there work without boundedness condition.

Some properties regarding Gaussian kernel are enlisted in the following lemma. Especially, all the necessary properties from Dini continuity are stated in (i) and (ii).

**Lemma 2.2.** For any fixed \(\beta > 0\) and \(\epsilon > 0\), there exists \(\delta > 0\) depending on \(\epsilon, \beta, \psi\) such that for every \(R \leq \delta\), and \(E_R := \{(s, y) : 0 < s < R^2, |y| < R\}\), we have

(i) \[\int_{E_R} \psi(\sqrt{s} + |y|)|\nabla_y \Gamma_{\beta}(s, y; 0, 0)| \, ds \, dy \leq \epsilon R,\]

(ii) \[\int_{E_R} \psi(\sqrt{s} + |y|)|D^2_y \Gamma_{\beta}(s, y; 0, 0)| \, ds \, dy \leq \epsilon.
\]

Moreover, for fixed \(\beta > 0\) there exists \(c > 0\) depending on \(\beta\) such that for every \(R < \infty\), and \(E'_R := \{(s, y) : 0 < s < R^2, \, R/2 < |y| < R\} \cup \{(s, y) : R^2/2 < s < R^2, \, |y| < R\}\),

we have

(iii) \[\int_{E'_R} |\nabla_y \Gamma_{\beta}(s, y; 0, 0)| \, ds \, dy \leq c R,\]

(iv) \[\int_{E'_R} |D^2_y \Gamma_{\beta}(s, y; 0, 0)| \, ds \, dy \leq c.
\]

**Proof.** The lemma can be proved by some elementary calculations using the definition of Dini continuity, the fact \(\sup_{\rho > 0} \rho^c e^{-\rho} < \infty\), for all \(c' > 0\) and the estimates \(|\nabla \Gamma_{\beta}(s, y; 0, 0)| \leq c s^{-d/2 - 1/2} e^{-\beta \frac{|y|^2}{s}}\) and \(|D^2 \Gamma_{\beta}(s, y; 0, 0)| \leq c s^{-d/2 - 1} e^{-\beta \frac{|y|^2}{s}}\). We skip the details. \(\Box\)

The following lemma, which generalizes the result of M. Gröter and K. Widman [20, Lemma 3.1] to the parabolic case, is the key step to obtain gradient estimates:

**Lemma 2.3.** Suppose that \(L\) is of the form (1.2) with smooth coefficients a\(_{ij}\). Let \(u \in C((s_a, t_a) \times B(x_0, 3r/2))\) such that \((\partial_t - L)u = 0\) in \((s_a, t_a) \times B(x_0, 3r/2)\). Then there exists \(C_0 = C_0(d, \lambda, \psi) > 0\) such that for any \((t, x) \in \bar{E} := (s_a, t_a) \times B(x_0, r),\)

\[|\nabla_x u(t, x)|_{d_E(t, x)} \leq C_0(\sqrt{t_a - s_a} + 1) \sup_{(s, y) \in \bar{E}} |u(s, y)|,
\]
where \( d_E(t, x) := \sqrt{t - s_x} \wedge \text{dist}(x, \partial B(x_s, r)) \). In particular,\[
|\nabla_x u(t_s, x_s)| \leq \frac{C_0}{1 \wedge r \wedge \sqrt{t_s - s_x}} \sup_{(s, y) \in E} |u(s, y)|.
\]

**Proof.** Without loss of generality, we assume \( s_x = 0 \) and \( x_s = 0 \). By [24, Theorem 6.6] and the remark after it, the solution \( u \) is smooth. Thus,\[
M_0 := \sup_{(t, x) \in E} |u(t, x)| < \infty, \quad M_1 := \sup_{(t, x) \in E} |\nabla_x u(t, x)| d_E(t, x) < \infty.
\]

Choose \((t_0, x_0) \in E\) such that \(|\nabla_x u(t_0, x_0)| d_E(t_0, x_0) > \frac{1}{2} M_1\). Let \( \eta \in C^\infty(\mathbb{R}^{d+1}), 0 \leq \eta \leq 1, \) be a cut-off function satisfying the following:\[
\eta(t, x) \equiv 1 \quad \text{in} \quad t_0 - t \leq \frac{r_0^2}{2} \quad \text{and} \quad |x - x_0| \leq \frac{r_0}{2},
\]
\[
\eta(t, x) \equiv 0 \quad \text{in} \quad t_0 - t \geq r_0^2 \quad \text{or} \quad |x - x_0| \geq r_0.
\]
\[
|\nabla_x \eta(t, x)| \leq \frac{4}{r_0}, \quad |D^2_x \eta(t, x)| \leq \frac{8}{r_0^2}, \quad |\partial_t \eta(t, x)| \leq \frac{4}{r_0^2},
\]

where \( r_0 \leq d_E(t_0, x_0)/2 \) will be chosen later. Let \( \Gamma_{t_0, x_0} \) be the heat kernel of the operator \( \partial_t - L_0 \) of the form in (1.2) with symmetric constant coefficients \( \sigma_i^j := \frac{1}{2} (a_{ij}(t_0, x_0) + a_{ji}(t_0, x_0)) \). Since \( u \) is smooth and \( \Gamma_{t_0, x_0}(t, x; \cdot, \cdot), \nabla \Gamma_{t_0, x_0}(t, x; \cdot, \cdot) \) belong to \( L^1(E) \), for \((t, x) \in E\), we may insert \( \eta(\cdot, \cdot) \Gamma_{t_0, x_0}(t, x; \cdot, \cdot) \) as a test function, and get
\[
0 = \int_E \left[ \sum_{i,j=1}^d \partial_y^i (a_{ij}(s, y) \partial_{y^i} u(s, y)) - \partial_s u(s, y) \right] \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy
\]
\[
= \int_E - \sum_{i,j=1}^d a_{ij}(s, y) \partial_{y^i} u(s, y) \partial_{y^j} (\eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y))
\]
\[
- \partial_t u(s, y) \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy
\]
\[
= \int_E \sum_{i,j=1}^d (a_{ij}(t_0, x_0) - a_{ij}(s, y)) \partial_{y^i} u(s, y) \partial_{y^j} (\eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y)) \, ds \, dy
\]
\[
- \lim_{\epsilon \to 0^+} \int_{B(0, r)} \int_0^{t-\epsilon} \sum_{i,j=1}^d a_{ij}(t_0, x_0) \partial_{y^i} u(s, y) \partial_{y^j} (\eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y))
\]
\[
+ \partial_t u(s, y) \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy.
\]

Using integration by parts, (2.3) and the fact \( \Gamma_{t_0, x_0}(t, x; s, y) = 0 \) for \( t \leq s \), we get
\[
0 = \int_E \sum_{i,j=1}^d (a_{ij}(t_0, x_0) - a_{ij}(s, y)) \partial_{y^i} u(s, y) \partial_{y^j} (\eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y)) \, ds \, dy
\]
\[ + \lim_{\epsilon \to 0^+} \int_{B(0,r)} \int_{0}^{t-\epsilon} \sum_{i,j=1}^{d} a_{ij}(t_0, x_0) u(s, y) \partial_y u(s, y) \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \\
- \int_{B(0,r)} u(s, y) \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) |_{s=0}^{t-\epsilon} \, dy \\
+ \int_{B(0,r)} \int_{0}^{t-\epsilon} u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \right]. \]

Hence we get

\[ u(t, x) \eta(t, x) = \lim_{\epsilon \to 0^+} \int_{B(0,r)} u(t - \epsilon, y) \eta(t - \epsilon, y) \Gamma_{t_0, x_0}(t, x; t - \epsilon, y) \, dy \]

\[ = \int_{E} \sum_{i,j=1}^{d} (a_{ij}(t_0, x_0) - a_{ij}(s, y)) \partial_y u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \]

\[ + \lim_{\epsilon \to 0^+} \int_{B(0,r)} \int_{0}^{t-\epsilon} \sum_{i,j=1}^{d} a_{ij}(t_0, x_0) u(s, y) \partial_y u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \\
+ u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy. \]

Switching the role of \( i \) and \( j \), and taking the average of the both sides, we obtain

\[ u(t, x) \eta(t, x) = \int_{E} \sum_{i,j=1}^{d} (a_{ij}(t_0, x_0) - a_{ij}(s, y)) \partial_y u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \]

\[ + \lim_{\epsilon \to 0^+} \int_{B(0,r)} \int_{0}^{t-\epsilon} \sum_{i,j=1}^{d} a^0_{ij} u(s, y) \partial_y u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \\
+ u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy. \]

First using the fact

\[ \sum_{i,j=1}^{d} a_{ij} \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) + \partial_y \Gamma_{t_0, x_0}(t, x; s, y) = 0 \quad \text{in} \ (0, t - \epsilon) \times B(0,r), \]

and then differentiating with respect to \( x_k, k = 1, \ldots, d \), and setting \((t, x) = (t_0, x_0)\), we obtain

\[ \nabla_x u(t_0, x_0) = \int_{E} \sum_{i,j=1}^{d} (a_{ij}(t_0, x_0) - a_{ij}(s, y)) \partial_y u(s, y) \partial_y \eta(s, y) \Gamma_{t_0, x_0}(t, x; s, y) \, ds \, dy \big|_{(t, x) = (t_0, x_0)} \]
\[
+ \lim_{\epsilon \to 0^+} \int_{B(0,r)} \int_{t-\epsilon}^{t} \sum_{i,j=1}^{d} (2a_{ij}^0 u \partial_{y_i} \eta \nabla_x \partial_{y_j} \Gamma_{t_0,x_0}(t,x,s,y) \\
+ a_{ij}^0 u(\partial_{y_i} \eta) \nabla_x \Gamma_{t_0,x_0}(t,x,s,y)) + u(\partial_s \eta) \nabla_x \Gamma_{t_0,x_0}(t,x,s,y) \, ds \, dy|_{(t,x)=(t_0,x_0)}.
\]

(2.5)

Here and below, we sometimes use \( u, \eta \) instead of \( u(s,y), \eta(s,y) \) to simplify notations. Note that \( \eta(s,y) \) and \( \Gamma_{t_0,x_0}(t,x,s,y) \) are non-zero only when \( 0 \leq t_0 - s < r_0^2 \), and \( |x_0 - y| \leq r_0 \). Let \( F := \{ (s,y) : 0 \leq t_0 - s < r_0^2, \ |x_0 - y| \leq r_0 \} \) and \( F' := \{ (s,y) : 0 \leq t_0 - s < r_0^2, \ r_0^2 \leq |x_0 - y| \leq r_0 \} \). Using \( r_0 \leq d_E(t_0,x_0)/2 \) and \( d_E(t_0,x_0) \leq 2d_E(s,y) \),

\[
\sup_{F} |\nabla_y u(s,y)| \leq \sup_{|x_0-y| \leq d_E(t_0,x_0)/2} |\nabla_y u(s,y)| \\
\leq \sup_{|x_0-y| \leq d_E(t_0,x_0)/2} d_E(s,y) |\nabla_y u(s,y)| \frac{1}{d_E(t_0,x_0)} \leq 2M_1 d_E(t_0,x_0),
\]

and we have

\[
|\nabla_x u(t_0,x_0)| \leq \frac{2dM_1}{d_E(t_0,x_0)} \int_{F} \psi(\sqrt{|t_0-s|} + |x_0-y|) |\nabla_y \eta| |\nabla_y \Gamma_{t_0,x_0}(t_0,x_0; s,y) |
\]

\[
+ \eta |D^2_y \Gamma_{t_0,x_0}(t_0,x_0; s,y)| ds \, dy
\]

\[
+ c_1(d, \lambda) M_0 \int_{F'} |D^2_y \eta| |\nabla_y \Gamma_{t_0,x_0}(t_0,x_0; s,y)| + |\nabla_y \eta| |D^2_y \Gamma_{t_0,x_0}(t_0,x_0; s,y)| ds \, dy
\]

\[
+ c_2(d) M_0 \int_{F'} |(\partial_s \eta) \nabla_y \Gamma_{t_0,x_0}(t_0,x_0; s,y)| ds \, dy
\]

=: \frac{2dM_1}{d_E(t_0,x_0)} f_1 + c_1 M_0 f_2 + c_2 M_0 f_3.
\]

By (2.4), and estimates (i) and (ii) of Lemma 2.2, there exists \( \delta > 0 \) depending on \( d, \lambda, \psi \) such that for \( r_0 := \min(\delta, d_E(t_0,x_0)/2) \), we get \( f_1 \leq (8d)^{-1} \). Also, by the estimates (iii) and (iv) of Lemma 2.2, \( f_2 \vee f_3 \leq c_3(\delta \wedge d_E(t_0,x_0))^{-1} \). Thus, by our choice of \( (t_0,x_0) \),

\[
\frac{1}{2} M_1 \leq |\nabla_x u(t_0,x_0)| d_E(t_0,x_0) \leq \frac{4}{4} M_1 + c_4 M_0 \frac{d_E(t_0,x_0)}{\delta \wedge d_E(t_0,x_0)}.
\]

Since \( \delta \) depends only on \( d, \lambda, \psi \), we conclude that \( M_1 \leq c_5(d, \lambda, \psi)((r + \sqrt{t_* - s_*}) \vee 1) M_0 \). \( \square \)

**Lemma 2.4.** Under the same condition as in Lemma 2.3, for any \( \epsilon_0 > 0 \), there exists \( \delta_0 = \delta_0(d, \lambda, \psi, r, t_* - s_*) \leq \sqrt{t_* - s_*} r/s \) such that for any \((\tilde{t}, \tilde{x}), (\hat{t}, \hat{x}) \in [(t_* + s_*)/2 - \delta_0^2, (t_* + s_*)/2] \times \tilde{B}(x_0, \delta_0)\),

\[
|\nabla_x u(\tilde{t}, \tilde{x}) - \nabla_x u(\hat{t}, \hat{x})| \leq \frac{\epsilon_0}{1 \wedge r \vee \sqrt{t_* - s_*}} \sup_{E} |u| \tag{2.6}
\]
Proof. Without loss of generality, we assume \( s_\ast = 0, x_\ast = 0 \). Let \( \eta(t, x) \in C^\infty(\mathbb{R}^{d+1}), 0 \leq \eta \leq 1 \), be a cut-off function such that \( \eta(t, x) \equiv 1 \) when \( t \leq r_0^2/2 \) and \( |x| \leq r_0/2 \), \( \eta(t, x) \equiv 0 \) when \( t > r_0^2 \) or \( |x| > r_0 \), and \( \|\nabla_x \eta\|_\infty \leq 4/r_0 \), \( \|D^2_x \eta\|_\infty \leq 8/r_0^2 \), \( \|\partial_x \eta\|_\infty \leq 4/r_0^2 \), where \( r_0 \leq \frac{\sqrt{r_0^2 + r}}{8} \) will be fixed later.

We denote \( \tilde{\eta}(t, x) := \eta(t - s, \tilde{x} - y) \) and \( \hat{\eta}(t, x) := \eta(t - s, \hat{x} - y) \). Let \( \Gamma_{\tilde{\eta}} \) be the heat kernel of \( \partial_t - \sum_{i,j=1}^d a_{ij}(\tilde{x}, \partial_{\tilde{x}})\hat{x}_{ij} \). Without loss of generality, we assume \( a_{ij} \) is symmetric, if not, we make it symmetric by replacing \( a_{ij} \) with \( (a_{ij} + a_{ji})/2 \), as in the proof of Lemma 2.3. By (2.5), the following identity holds for any \( (t, \hat{x}) \in E = (0, t_\ast] \times B(0, r) \):

\[
\nabla_x u(t, \hat{x}) = \int_E \sum_{i,j=1}^d (a_{ij}(\tilde{t}, \tilde{x}) - a_{ij}(s, y)) \partial_{yj} u \nabla_x \partial_{yj}(\tilde{\eta}(t, x; s, y)) ds dy |_{(t, x) = (\tilde{t}, \tilde{x})}
\]

\[
+ \sum_{i,j=1}^d (2a_{ij}(\tilde{t}, \tilde{x})u \partial_{yj}(\tilde{\eta}(s, y)) \nabla_x \partial_{yj} \Gamma_{\tilde{\eta}}(t, x; s, y)
\]

\[
+ a_{ij}(\tilde{t}, \tilde{x})u(\partial_{yj}(s, \tilde{y})) \nabla_x \Gamma_{\tilde{\eta}}(t, x; s, y)) + u(\partial_s \tilde{\eta}) \nabla_x \Gamma_{\tilde{\eta}}(t, x; s, y) ds dy |_{(t, x) = (\tilde{t}, \tilde{x})}.
\]

For \( (t, \hat{x}) \in E_{r_0} := (\frac{t_\ast}{2} - 2r_0^2, \frac{t_\ast}{2}] \times B(0, 2r_0) \), we define

\[
f(\tilde{t}, \tilde{x}, s, y) = \int_E \sum_{i,j=1}^d (2a_{ij}(\tilde{t}, \tilde{x})u \partial_{yj}(\tilde{\eta}(s, y)) \nabla_x \partial_{yj} \Gamma_{\tilde{\eta}}(t, x; s, y)
\]

\[
+ a_{ij}(\tilde{t}, \tilde{x}) \hat{\eta}(t, x; s, y)) \nabla_x \Gamma_{\tilde{\eta}}(t, x; s, y)) + u(\partial_s \tilde{\eta}) \nabla_x \Gamma_{\tilde{\eta}}(t, x; s, y) |_{(t, x) = (\tilde{t}, \tilde{x})}.
\]

Then, for every \( (\tilde{t}, \tilde{x}), (\hat{t}, \hat{x}) \in E_{r_0}, |\nabla_x u(\tilde{t}, \tilde{x}) - \nabla_x u(\hat{t}, \hat{x})| \leq I_1 + I_2 \), where

\[
I_1 := \int_E \sum_{i,j=1}^d \left| a_{ij}(\tilde{t}, \tilde{x}) - a_{ij}(s, y) \right| \| \partial_{yj} u \nabla_x \partial_{yj}(\tilde{\eta}(t, x; s, y)) \|_{(t, x) = (\tilde{t}, \tilde{x})}
\]

\[
+ \left| a_{ij}(\tilde{t}, \tilde{x}) - a_{ij}(s, y) \right| \| \partial_{yj} u \nabla_x \partial_{yj} \Gamma_{\tilde{\eta}}(t, x; s, y) \|_{(t, x) = (\tilde{t}, \tilde{x})} \, ds \, dy,
\]

\[
I_2 := \left| \int_E \left[ f(\tilde{t}, \tilde{x}, s, y) u(s, y) - f(\hat{t}, \hat{x}, s, y) u(s, y) \right] ds \, dy \right|.
\]

Note that \( \hat{\eta} \) and \( \Gamma_{\tilde{\eta}} \) are non-zero only when \( 0 \leq \tilde{t} - s < r_0^2 \), and \( |\tilde{x} - y| \leq r_0 \). Hence, our domain \( E \) can be restricted to \( E_{\tilde{t}, \tilde{x}} := \{(s, y): 0 \leq \tilde{t} - s < r_0^2, |\tilde{x} - y| \leq r_0 \} \). Then, by the definition of \( \tilde{\eta} \),

\[
I_1 \leq 4d \sup_{(\tilde{t}, \tilde{x}) \in E_{\tilde{t}, \tilde{x}}} \int_{(\tilde{t}, \tilde{x})} \psi(\sqrt{\tilde{t} - s + |\tilde{x} - y|}) |\nabla_y u| \times (|\nabla_x \hat{\eta} \nabla_y \Gamma_{\tilde{\eta}}(\tilde{t}, \tilde{x}; s, y) + \hat{\eta} |D_y^2 \Gamma_{\tilde{\eta}}(\tilde{t}, \tilde{x}; s, y)|) \, ds \, dy
\]

\[
\leq \frac{16d}{r_0} \sup_{(\tilde{t}, \tilde{x}) \in E_{\tilde{t}, \tilde{x}}} \int_{(\tilde{t}, \tilde{x})} \psi(\sqrt{\tilde{t} - s + |\tilde{x} - y|}) |\nabla_y u| \, |\nabla_y \Gamma_{\tilde{\eta}}(\tilde{t} - s, \tilde{x} - y; 0, 0) | \, ds \, dy
\]
\[ + 4d \sup_{(\tilde{t}, \tilde{x}) \in E_0} \int_{F_{\tilde{t}, \tilde{x}}} \psi \left( \sqrt{t - s} + |\tilde{x} - y| \right) |\nabla_y u| |D_y^2 \Gamma_{\tilde{t}, \tilde{x}}(\tilde{t} - s, \tilde{x} - y; 0, 0)| \, ds \, dy \]
\[ =: \frac{16d}{r_0} J_1 + 4d J_2. \]

Fix \( \epsilon_0 > 0 \). By Lemma 2.2(i), we can choose \( r_0 \leq \frac{\sqrt{r \wedge t}}{8} \) small enough such that

\[ J_1 = \int_{F_{\tilde{t}, \tilde{x}}} \psi \left( \sqrt{t - s} + |\tilde{x} - y| \right) |\nabla_y u| \left| \nabla_y \Gamma_{\tilde{t}, \tilde{x}}(\tilde{t} - s, \tilde{x} - y; 0, 0) \right| \, ds \, dy \]
\[ \leq \sup_E |\nabla_y u| \int_{F_{\tilde{t}, \tilde{x}}} \psi \left( \sqrt{t - s} + |\tilde{x} - y| \right) \left| \nabla_y \Gamma_{\tilde{t}, \tilde{x}}(\tilde{t} - s, \tilde{x} - y; 0, 0) \right| \, ds \, dy \]
\[ \leq \frac{\epsilon_0 r_0}{32C_0} \sup_E |\nabla_y u|. \]

Similarly, using Lemma 2.2(ii), we get \( J_2 \leq \frac{\epsilon_0}{8C_0} \sup_E |\nabla_y u| \). Hence, by Lemma 2.3,

\[ I_1 \leq \frac{\epsilon_0}{2C_0} \sup_E |\nabla_y u| \leq \frac{\epsilon_0}{2(1 \wedge r \wedge \sqrt{t})} \sup_{(s, y) \in E} |u(s, y)|. \]

Now we compute \( I_2 \). Note that \( \partial_y \tilde{y} \) are non-zero only if

\[ 0 \leq \tilde{t} - s < r_0^2, \quad r_0/2 \leq |\tilde{x} - y| \leq r_0 \quad \text{or} \quad r_0^2/2 < \tilde{t} - s < r_0^2, \quad |\tilde{x} - y| \leq r_0. \]

For every \( \delta_0 \leq r_0 \) and \((\tilde{t}, \tilde{x}), (\tilde{\tilde{t}}, \tilde{x}) \in E_{\delta_0} := [t_s/2 - \delta_0^2, t_s/2] \times \overline{B(0, \delta_0)} \), let

\[ H \:= \left\{ (s, y) \left| \frac{r_0^2}{8} \leq \tilde{t} - s < 8r_0^2, \ r_0/8 \leq |\tilde{x} - y| \leq 8r_0 \right. \right\}. \]

Note that \( H \subset E \). Moreover, \( \{0 \leq \tilde{t} - s < r_0^2, \ r_0/2 \leq |\tilde{x} - y| \leq r_0 \} \subset H, \{r_0^2/2 < \tilde{t} - s < r_0^2, \ |\tilde{x} - y| \leq r_0 \} \subset H, \ (0 \leq \tilde{t} - s < r_0^2, \ r_0/2 \leq |\tilde{x} - y| \leq r_0 \} \subset H \) and \( \{r_0^2/2 < \tilde{t} - s < r_0^2, \ |\tilde{x} - y| \leq r_0 \} \subset H \). Thus

\[ I_2 \leq \left( \sup_E |u| \right) \sup_{(\tilde{t}, \tilde{x}), (\tilde{\tilde{t}}, \tilde{x}) \in E_{\delta_0}} \int_H \left| f(\tilde{t}, \tilde{x}, s, y) - f(\tilde{\tilde{t}}, \tilde{x}, s, y) \right| \, ds \, dy. \]

Put the second supremum inside the integrals, and then the integrand is less than or equal to

\[ 2 \sup_{(\tilde{t}, \tilde{x}), (\tilde{\tilde{t}}, \tilde{x}) \in E_{\delta_0}} \left| f(\tilde{t}, \tilde{x}, s, y) \right| 1_H(s, y). \]

If \( \delta_0 \leq \frac{r_0}{8} \), we see that either \( |\tilde{x} - y| \geq 14\delta_0 \) or \( |\tilde{t} - s| \geq 31\delta_0^2 \), therefore the integrand is bounded uniformly. By continuity of \( f(\tilde{t}, \tilde{x}, s, y) \) with respect to \( (\tilde{t}, \tilde{x}) \) and using Lebesgue dominated convergence theorem, \( I_2 \) goes to zero as \( \delta_0 \) goes to zero. \( \square \)

Let \( \tilde{a}_{ij}(t, x) := a_{ij}(T - t, x) \) and \( \tilde{L}u := \partial_t \tilde{a}_{ij}(\partial_j \partial_i u) \). If \( a_{ij} \) is smooth and \( \tilde{p}_D^0 \) is the heat kernel of \( \partial_t - \tilde{L} \), then it is well known that
\[ p^0_D(t, x; s, y) = \hat{p}^0_D(T - s, y; T - t, x) \quad \text{for } 0 \leq s < t \leq T. \]  
(2.7)

(See [17, Theorem 15] and our Lemma 3.9 for the proof.)

**Lemma 2.5.** When \( a_{ij} \) is smooth, \( \nabla_s p^0_D(t, x; s, y) \) is jointly continuous on \( (0, \infty) \times D \times (0, t) \times D \).

**Proof.** Fix \( s_0 < t_0 \) and \( x_0, y_0 \in D \) and let \( T := 2t_0 \). Choose \( r_0 = \sqrt{s_0} \) such that \( F_1 := (t_0, t_0 + r_0^2] \times B(x_0, r_0) \), \( F_2 := (s_0, s_0 + r_0^2] \times B(y_0, r_0) \) are relatively compact open subsets of \( (0, 2t_0) \times D \) with \( F_1 \cap F_2 = \emptyset \). Using (2.7), for any \( (\hat{t}, \hat{x}), (\hat{t}, \hat{x}) \in F_1 \), \( (s, y), (s, y) \in F_2 \),

\[ |\nabla_s p^0_D(\hat{t}, \hat{x}; \hat{t}, \hat{x}, \hat{y}) - \nabla_s p^0_D(\hat{t}, \hat{x}; \hat{t}, \hat{x}, \hat{y})| \leq |\nabla_s p^0_D(\hat{t}, \hat{x}; \hat{t}, \hat{x}, \hat{y}) - \nabla_s p^0_D(\hat{t}, \hat{x}; \hat{t}, \hat{x}, \hat{y})| + |\nabla_s p^0_D(T - \hat{s}, \hat{y}; T - \hat{t}, \hat{x}) - \nabla_s p^0_D(T - \hat{s}, \hat{y}; T - \hat{t}, \hat{x})|. \]

Since \( F_1 \) and \( F_2 \) are disjoint, the heat kernel \( p^0_D(t, x; s, y) \) can be viewed as a solution of \((\partial_t - L)u = 0\). Thus by Lemma 2.4, given \( \epsilon > 0 \), we can choose \( \delta = \delta(\epsilon) \) such that for any \( (\hat{t}, \hat{x}), (\hat{t}, \hat{x}) \in [(2t_0 + r_0^2)/2 - \delta^2, (2t_0 + r_0^2)/2] \times B(x_0, \delta) \) and \( (s, y), (s, y) \in [(2s_0 + r_0^2)/2 - \delta^2, (2s_0 + r_0^2)/2] \times B(y_0, \delta) \),

\[ |\nabla_s p^0_D(\hat{t}, \hat{x}; \hat{t}, \hat{x}, \hat{y}) - \nabla_s p^0_D(\hat{t}, \hat{x}; \hat{t}, \hat{x}, \hat{y})| \leq \frac{2\epsilon}{1 \wedge t_0} \sup_{(t, x) \in F_1, (s, y) \in F_2} |p^0_D(t, x; s, y)| < \infty. \]  
(2.8)

Now the main result of this section follows:

**Theorem 2.6.** Let \( D \) be \( C^{1,\alpha} \)-domain satisfying the connected line condition with characteristics \( (R_0, \Lambda_0) \) and \( (\chi_1, \chi_2) \), respectively. Also, \( L \) is an operator of the form in (1.2) satisfying (1.3)–(1.4). Then, for some constant \( C_3 = C_3(d, \lambda, \chi_1, \chi_2, \psi, T, \Lambda_0, R_0) \), the heat kernel \( p^0_D \) of the operator \( \partial_t - L \) in \( (0, T) \times D \) satisfies the following gradient estimate: for every \( x, y \in D \) and \( 0 \leq s < t \leq s + T \),

\[ |\nabla_x p^0_D(t, x; s, y)| \leq C_3 \psi_1(t - s, y) \Gamma_{C_2/2}(t, x; s, y). \]  
(2.9)

\[ |\nabla_y p^0_D(t, x; s, y)| \leq C_3 \psi_1(t - s, x) \Gamma_{C_2/2}(t, x; s, y). \]  
(2.10)

Here \( C_2 \) is the constant in Theorem 2.1. Especially when \( D = \mathbb{R}^d \), these estimates also hold with \( \psi_1 = 1 \). Also, \( \nabla_x p^0_D(t, x; s, y) \) and \( \nabla_y p^0_D(t, x; s, y) \) are jointly continuous when \( s < t \).

**Proof.** Assume \( a_{ij} \) is smooth. Using (2.1) with \( E = (1 + s, t) \times B(x, \rho(x)/2 \wedge \sqrt{t - s}) \) and Theorem 2.1, for \( t > s \),

\[ |\nabla_x p^0_D(t, x; s, y)| \leq \frac{c_1}{1 \wedge \rho(x) \wedge \sqrt{t - s}} \sup_{(v, z) \in E} p^0_D(v, z; s, y) \]
\[ \leq \frac{c_2}{1 \wedge \rho(x) \wedge \sqrt{t - s}} \sup_{(v, z) \in E} \left( 1 \wedge \frac{\rho(z)}{\sqrt{v - s}} \right) \left( 1 \wedge \frac{\rho(y)}{\sqrt{v - s}} \right) \frac{1}{(v - s)^{d/2}} e^{-c_2 \frac{(z - y)^2}{v - s}} \]
\[ \leq \frac{c_3}{1 \wedge \rho(x) \wedge \sqrt{t - s}} \left( 1 \wedge \frac{\rho(x)}{\sqrt{t - s}} \right) \left( 1 \wedge \frac{\rho(y)}{\sqrt{t - s}} \right) \sup_{z \in B(x, \sqrt{t - s})} \frac{1}{(t - s)^{d/2}} e^{-c_2 \frac{(z - y)^2}{v - s}}.
\]

Considering the cases \( \rho(x) \leq \sqrt{t - s} \) and \( \rho(x) > \sqrt{t - s} \) separately, we see that

\[ \left( \frac{1}{1 \wedge \rho(x) \wedge \sqrt{t - s}} \right) \left( 1 \wedge \frac{\rho(x)}{\sqrt{t - s}} \right) \leq \sqrt{\frac{\rho(x)}{1 \wedge \sqrt{t - s}}}, \]
On the other hand, since \(|x - y|^2/(t - s) \leq 2(|x - z|^2 + |z - y|^2)/(t - s) \leq 2(1 + |z - y|^2/(t - s))\) for all \(z \in B(x, \sqrt{t - s})\), we see that
\[
\sup_{z \in B(x, \sqrt{t - s})} e^{-C_2 \frac{|z - y|^2}{t - s}} \leq e^{C_2 \frac{|x - y|^2}{2(t - s)}},
\]

For Dini continuous \(a_{ij}\), without loss of generality, assume \(a_{ij}\) is defined in \(\mathbb{R}^{d+1}\). Consider \(a_{ij}^m := a_{ij} \ast \phi_m\), where \(\phi_m(t, x) = 2^{-m(d+1)} \phi(t/2^m, x/2^m) \geq 0\), \(\phi \in C^\infty_0((-1, 1) \times B_1)\) and \(\int_{\mathbb{R}^{d+1}} \phi = 1\). Since \(a_{ij}\) is continuous, \(a_{ij}^m\) is uniformly convergent to \(a_{ij}\) in \(E\). Furthermore, note that \(\sum_{i,j=1}^d |a_{ij}^m(t, x) - a_{ij}^n(t, y)| \leq \psi(\sqrt{t - s} + |x - y|)\). Let \(p_D^{0,m}\) be the heat kernel of \(\partial_t - L_m\), where \(L_m\) is the same form of operator \(L\) replacing \(a_{ij}\) with \(a_{ij}^m\). It is known that \(p_D^{0,m}\) weakly converges to \(p_D^0\). (See the proof of \([2\text{, Theorem 9}]\).)

Since, by \((2.8)\) and \((2.9)\), \(\nabla_x p_D^{0,m}\) are equicontinuous and uniformly bounded, using Arzela–Ascoli theorem and standard diagonal argument, there exists a subsequence \(\nabla_x p_D^{0,m_k}\) converging uniformly on every compact subset of \(Q_T \times Q_T \cap \{s < t\}\). Since \(\nabla_x p_D^{0,m_k}\) is continuous, \(\nabla_x p_D^{0}\) is continuous on \(Q_T \times Q_T \cap \{s < t\}\).

For \((2.10)\), let \(x, y \in D\) and \(0 \leq s < t \leq s + T\). Note that by \((2.7)\), \(p_D^0(t, t; s, y) = \tilde{p}_D^0(2t - s, y; t, x)\), where \(\tilde{p}_D^0\) is the heat kernel of \(\partial_t - L\) replacing \(a_{ij}(t, x)\) by \(\tilde{a}_{ij}(t, x) := a_{ij}(2t - t, x)\). Thus, by \((2.9)\),
\[
|\nabla_x p_D^0(t, x; s, y)| = |\nabla_y \tilde{p}_D^0(2t - s, y; t, x)| \leq \frac{C_1}{\sqrt{t - s}} \left(1 + \frac{\rho(x)}{\sqrt{t - s}}\right) \Gamma_{C_2/2}(t, x; s, y).
\]

3. Stability of heat kernel under perturbation through parabolic Kato class

For a signed measure \(\nu\) on \([0, \infty) \times \mathbb{R}^d\), we use \(\nu^+\) and \(\nu^-\) to denote its positive and negative parts, and \(|\nu| = \nu^+ + \nu^-\) its total variation. We can extend \(\nu\) to \(\mathbb{R}^{d+1}\) by taking \(\nu \equiv 0\) on \((-\infty, 0) \times \mathbb{R}^d\).

For a signed measure \(\nu\) on \([0, \infty) \times \mathbb{R}^d\) and \(h > 0\), we define
\[
N_\nu(h) := \sup_{(t, x) \in (0, \infty) \times \mathbb{R}^d} \int_{(t-h, t) \times \mathbb{R}^d} (t - s)^{-\frac{1}{2}} \Gamma_{C_2/(32)}(t, x; s, y) d|\nu|(s, y)
+ \sup_{(s, y) \in (0, \infty) \times \mathbb{R}^d} \int_{(s, s+h) \times \mathbb{R}^d} (t - s)^{-\frac{1}{2}} \Gamma_{C_2/(32)}(t, x; s, y) d|\nu|(t, x).
\]

Here the constant \(C_2\) comes from the constant in Theorem 2.1.

**Definition 3.1.** We say that a signed measure \(\nu\) on \([0, \infty) \times \mathbb{R}^d\) belongs to the (parabolic) Kato class \(K_d\) if \(\lim_{h \uparrow 0} N_\nu(h) = 0\). A function \(f\) on \([0, \infty) \times \mathbb{R}^d\) is said to be in the (parabolic) Kato class \(K_d\) if \(f(t, x) dt dx \in K_d\).

More generally, we say that a \(d\)-dimensional vector-valued signed measure \(\mu := (\mu^1, \ldots, \mu^d)\) on \([0, \infty) \times \mathbb{R}^d\) belongs to the Kato class \(K_d\) if each \(\mu^i\) belongs to \(K_d\). Similarly, a \(d\)-dimensional vector-valued function \(V = (V^1, \ldots, V^d)\) on \([0, \infty) \times \mathbb{R}^d\) belongs to \(K_d\) if each \(V^i\) belongs to the Kato class \(K_d\).

We emphasize that \(\mu^i\) may not be absolutely continuous with respect to the Lebesgue measure on \([0, \infty) \times \mathbb{R}^d\) (see \([3, \text{p. 792}]\)). In Section 3.1, we give a couple of examples on signed measures belonging to Kato class. Regarding Example 3.2, one may refer to \([19\text{, Lemmas 3.7 and 3.8}]\).
3.1. Examples

Example 3.2. Let $\mu$ be a signed measure on $[0, \infty) \times \mathbb{R}^d$, and assume there exist a constant $\delta > 0$ and a positive function $\alpha$ on $[0, \infty)$ such that

$$|\mu|((t - h, t) \times B_{g(h)}(x)) \vee |\mu|((s, s + h) \times B_{g(h)}(x)) \leq \alpha(h)g(h)^{\delta}, \tag{3.1}$$

for any $t, s > 0, x \in \mathbb{R}^d$, and any positive function $g$ on $[0, \infty)$ and arbitrary small $h > 0$. Furthermore, we assume

$$\lim_{h \downarrow 0} h^{-d/2-1/2+\delta/2} \sum_{k=1}^{\infty} \alpha((1/k)^2(d+1)/h)k^{-\delta/(d+1)} = 0. \tag{3.2}$$

Then $\mu$ belongs to $K_d$. In fact, if we let $C_3 := C_2/32$ for simplicity,

$$\int \int_{\mathbb{R}^d} I_{C_3}(t, x; s, y)(t - s)^{-1/2} d|\mu|(s, y)$$

$$= \int_{0}^{t} \int_{h^{-d/2-1/2}} \int_{h^{-d/2-1/2}} |\mu|(\cdot) \, d\lambda$$

$$\leq \alpha(h) \int_{0}^{\infty} (\lambda^{-1} \ln(h^{d/2+1/2}/\lambda))^{\delta/2} d\lambda.$$ 

Using the fact $-x\ln x$ is increasing near zero and the assumption (3.1),

$$I_1(h) \leq \int_{0}^{t} |\mu|((s, y): s \in (t - h, t), \ y \in \mathbb{R}^d, |x - y| < \left\{ -C_3^{-1}h \ln(h^{d/2+1/2}/\lambda) \right\}^{1/2} \, d\lambda,$$

$$\leq \alpha(h) \int_{0}^{\infty} (\lambda^{-1} \ln(h^{d/2+1/2}/\lambda))^{\delta/2} d\lambda.$$ 

We know $h^{\delta/2-d/2-1/2}\alpha(h) \to 0$ as $h \to 0$ by (3.2), and therefore $I_1(h) \to 0$ as $h \to 0$. Also,

$$I_2(h) \leq \sum_{k=1}^{\infty} \int_{kh^{-d/2-1/2}} |\mu|((s, y): s \in (t - h, t), |x - y|$$

$$< \left\{ -C_3^{-1}(t - s) \ln(k^{-1}(h/(t - s))^{(d+1)/2}) \right\}^{1/2} \, d\lambda.$$ 

Note the logarithm function is positive when $t - s < (1/k)^2(d+1)/h$. Hence,
\[ I_2(h) \leq \sum_{k=1}^{\infty} h^{-d/2-1/2} |\mu| \left( (s, y): t - s < (1/k)^{2/(d+1)} h, |x - y| < c_3^{-1/2} h^{1/2} \right) \]
\[ \times \frac{1}{k^{1/(d+1)}} k^{1/(d+1)} \left( \frac{t - s}{h} \right)^{1/2} \left\{ \ln \left( k^{-1} (h/t - s)^{(d+1)/2} \right) \right\}^{1/2} d\lambda \]
\[ \leq c_1 M h^{-d/2-1/2+\delta/2} \sum_{k=1}^{\infty} \alpha \left( (1/k)^{2/(d+1)} h k^{-\delta/d+1} \right), \]

by (3.1), where \( M := \sup_{x<1} x^{\delta/2} (\ln 1/(x^{(d+1)/2}))^{\delta/2} \). We apply (3.2) again and conclude \( I_2(h) \to 0 \) as \( h \to 0 \). From the same argument,

\[ \sup_{(s, y) \in (0, \infty) \times \mathbb{R}^d} \int_{(s, s+h) \times \mathbb{R}^d} (t-s)^{-1/2} \Gamma_{C_3}(t; x, y) d|\mu|(t, x) \]
also goes to zero as \( h \to 0 \).

**Example 3.3.** For every \( s \in [0, \infty) \), let \( \mu_s \) be a measure on \( \mathbb{R}^d \), and assume there exist \( \delta > 0 \) and a positive function \( \alpha \) on \( [0, \infty) \) such that for arbitrary small \( r > 0 \),

\[ |\mu_s| (B_r(x)) \leq \alpha(s) r^\delta, \quad \text{for all } x \in \mathbb{R}^d, \quad (3.3) \]
\[ \sup_{t \in (0, \infty)} \int_{t-h}^{t} \alpha(s)(t-s)^{\delta/2-d/2-1/2} ds \to 0 \quad \text{as } h \to 0 \quad (3.4) \]

and

\[ \sup_{s \in (0, \infty)} \int_{s}^{s+h} \alpha(t)(t-s)^{\delta/2-d/2-1/2} dt \to 0 \quad \text{as } h \to 0. \quad (3.5) \]

Then \( d\mu_s(y) ds \in K_d \). In fact, if we let \( C_3 := C_2/32 \),

\[ \int_{\mathbb{R}^d} \Gamma_{C_3}(t; x, y)(t-s)^{-1/2} d|\mu_s|(y) \]
\[ = \int_{0}^{\infty} |\mu_s|(y: \Gamma_{C_3}(t; x, y)(t-s)^{-1/2} > \lambda) d\lambda \]
\[ \leq \int_{0}^{(t-s)^{-d/2-1/2}} |\mu_s|(y: |x - y| < \left\{ -c_3^{-1}(t-s) \ln((t-s)^{d/2+1/2}\lambda) \right\}^{1/2}) d\lambda, \]

and using the assumption (3.3),

\[ \int_{\mathbb{R}^d} \Gamma_{C_3}(t; x, y)(t-s)^{-1/2} d|\mu_s|(y) \leq \alpha(s) C_3^{-\delta/2} (t-s)^{\delta/2-d/2-1/2} \int_{0}^{1} (-\ln \eta)^{\delta/2} d\eta. \]
Then,

\[
\sup_{(t,x) \in (0,\infty) \times \mathbb{R}^d} \int_{t-h}^{t} \int_{\mathbb{R}^d} \Gamma_{C_3}(t,x,s,y)(t-s)^{-1/2} d|\mu_s|(y) ds \leq c_1 \sup_{t \in (0,\infty)} \int_{t-h}^{t} \alpha(s)(t-s)^{3/2-d/2-1/2} ds.
\]

By (3.4), the right-hand side goes to zero as \( h \to 0 \). Using (3.5) instead of (3.4),

\[
\sup_{(s,y) \in (0,\infty) \times \mathbb{R}^d} \int_{s}^{s+h} \int_{\mathbb{R}^d} \Gamma_{C_3}(t,x,s,y)(t-s)^{-1/2} d|\mu_{t-s}|(x) dt
\]

also goes to zero as \( h \to 0 \).

Recall that \( m_d \) is the \( d \)-dimensional Lebesgue measure for \( d \geq 1 \), and let \( \mu_s(dy) = f(s,y)ds \), where \( f \in L^{p_1}(m_1, L^{p_2}(m_d)) \) and \( p_1, p_2 \geq 1 \). Let \( q_i \) satisfy \( 1/p_1 + 1/q_i = 1, i = 1, 2 \). Then, one can check (3.4)-(3.5) are true with \( \delta = d/q_2 \) if \( (d/2q_2) - d/2 - 1/2 = q_1 + 1 > 0 \). Moreover, if \( |f(s,y)| \leq f_1(s)g_1(y) \), where \( f_1 \in L^{\infty}(\mathbb{R}^1) \) and \( g_1 \in L^{p_2}(\mathbb{R}^d) \), then (3.4)-(3.5) are true with \( \delta = d/q_2 \) if \( p_2 > d \).

**Example 3.4.** Let \( \mu_0 \) be the Cantor–Lebesgue measure on \([0,1]\) (see [4] for the construction), and extend it to \( \mathbb{R} \) by giving 0 outside of \([0,1]\). Also define \( d\mu_0 = d\mu_0(-s) \) for \( t-h < s < t \). Note that for every \( s \in [0,T] \), \( \mu_0([s, s+1]) = \mu_0(0,1] \). Now we define

\[
d\mu_s(a,b) = d\mu_0(a)db \quad \text{for } t-h < s < t.
\]

Then, for the \( d \)-dimensional ball \( B(x,r) \), \( \mu_s(B(x,r)) \leq c_1 r^{1+\log 2/\log 3} \) by [3, Example 2.2], and this satisfies the conditions (3.3) and (3.4).

### 3.2. Properties

The parabolic Kato class is a generalized version of elliptic Kato class in the following sense.

**Proposition 3.5.** (See Proposition 2.3 in [22].) Suppose that \( d \geq 3 \) and \( v \) is a signed measure on \( \mathbb{R}^d \). Then \( \nu(dx) dt \in K_d \) if and only if \( \lim_{r \to 0} \sup_{x \in \mathbb{R}^d} \int_{|x-y| \leq r} |x-y|^{-d+1} d|\nu|(y) = 0 \).

The next two lemmas follow easily from the well-known 3P-type inequalities (see [31, Lemmas 3.2 and 3.3]). Recall that we use the notational convention (1.8).

**Lemma 3.6.** There exists a positive constant \( M \) depending on \( d,C_2 \) such that for any signed measure \( \nu \) on \([0,\infty) \times \mathbb{R}^d \), for \( t > s \) and \( (x,y) \in D \times D \),

\[
\begin{align*}
(i) \quad & \int_{D} \int_{s}^{t} \psi_0(t-v,x,z)\Gamma_{C_2/4}(t,x,v,z)\psi_1(v-s,y)\Gamma_{C_2/2}(v,z,s,y) |\nu|(v,z) \\
& \leq M \psi_0(t-s,x,y)\Gamma_{C_2/4}(t,x,s,y)N_{\nu}(t-s), \\
(ii) \quad & \int_{D} \int_{s}^{t} \psi_1(t-v,z)\Gamma_{C_2/4}(t,x,v,z)\psi_1(v-s,y)\Gamma_{C_2/2}(v,z,s,y) |\nu|(v,z) \\
& \leq M \psi_1(t-s,y)\Gamma_{C_2/4}(t,x,s,y)N_{\nu}(t-s).
\end{align*}
\]
Lemma 3.7. There exists a positive constant $M$ depending on $d, C_2$ such that for any signed measure $v$ on $[0, \infty) \times \mathbb{R}^d$, for $t > s$ and $(x, y) \in D \times D$,

\[
\begin{align*}
(i) & \quad \int_D \int_s^t \psi_1(t - v, x) \Gamma_{C_2/4}(t, x; v, z) \psi_0(v - s, z, y) \Gamma_{C_2/2}(v, z; s, y) d|v|(v, z) \\
& \quad \leq M \psi_0(t - s, x, y) \Gamma_{C_2/4}(t, x; s, y) N_v(t - s), \\
(ii) & \quad \int_D \int_s^t \psi_1(t - v, x) \Gamma_{C_2/4}(t, x; v, z) \psi_1(v - s, z) \Gamma_{C_2/2}(v, z; s, y) d|v|(v, z) \\
& \quad \leq M \psi_1(t - s, x) \Gamma_{C_2/4}(t, x; s, y) N_v(t - s).
\end{align*}
\]

Proof. The claim (ii) follows from Lemma 3.6(ii). Moreover, (i) follows from the inequality

\[
\psi_1(t - v, x) \Gamma_{C_2/4}(t, x; v, z) \psi_0(v - s, z, y) \Gamma_{C_2/2}(v, z; s, y) \\
\leq M \psi_0(t - s, x, y) \Gamma_{C_2/4}(t, x; s, y) \\
\times \left( \psi_1(t - v, z) \Gamma_{C_2/8}(t, x; v, z) + \psi_1(v - s, z) \Gamma_{C_2/8}(v, z; s, y) \right),
\]

where $M = M(d, C_2)$ is a positive constant. This is a slightly modified version of the inequality in [31, Lemma 3.2], and can be proved by the same method. We skip the details. □

3.3. Perturbation for bounded functions

Suppose that $B(t, x) = (B^1(t, x), \ldots, B^d(t, x))$ is a bounded vector-valued function which clearly belongs to the Kato class $K_d$. Also, recall that $D$ is a $C^{1,\alpha}$-domain satisfying the connected line condition with characteristics $(R_0, \Lambda_0)$ and $(\chi_1, \chi_2)$, respectively and $p_D^0(t, x; s, y)$ is the heat kernel of $\partial_t - L$. By [2, Theorem 6], there exists a function $p_D(t, x; s, y)$ satisfying

\[
p_D(t, x; s, y) = p_D^0(t, x; s, y) + \int_D \int_s^t p_D(t, x; v, z) \nabla_z p_D^0(v, z; s, y) \cdot B(v, z) dv dz,
\]

which is a heat kernel of $\partial_t - L - B \cdot \nabla_x$. Moreover, the following Chapman–Kolmogorov equation holds:

\[
p_D(t, x; s, y) = \int_D p_D(t, x; v, z) p_D(v, z; s, y) dz, \quad \forall x, y \in D, \ 0 < s < v < t.
\]

Recall the constant $M$ in Lemma 3.6, and $C_3$ in Theorem 2.6. Choose $t_0$ small such that

\[
C_3 M \sum_{i=1}^d N_{B_i}(t_0) \leq \frac{1}{2}.
\]

We define $I^p_k(t, x; s, y)$ recursively for $k \in \mathbb{N}_0$ and $(t, x, y) \in (s, s + t_0] \times D \times D$:
\[ I^D_0(t, x; s, y) := p^0_D(t, x; s, y), \]
\[ I^B_{k+1}(t, x; s, y) := \int_D \int_s^t I^B_k(t, x; v, z) \nabla_z p^0_D(v, z; s, y) \cdot B(v, z) \, dv \, dz. \]

By induction, and using Lemma 3.6, Theorems 2.1 and 2.6,
\[ |I^B_k(t, x; s, y)| \leq C_1 \psi_0(t - s, x, y)(t - s)^{-\frac{d}{2}} e^{-\frac{C_2|x-y|^2}{4(t-s)}} \left( C_3 M \sum_{i=1}^d N_{B_i}(t - s) \right)^k. \] (3.9)

Here, \( C_1 \) and \( C_3 \) are the constants in Theorem 2.1 and Theorem 2.6 respectively. Thus, by iterating (3.6) and our choice of \( t_0 > 0 \), for \( (t, x, y) \in (s, s + t_0] \times D \times D \),
\[ p_D(t, x; s, y) = \sum_{k=0}^{\infty} I^B_k(t, x; s, y) \] (3.10)
and
\[ p_D(t, x; s, y) \leq \sum_{k=0}^{\infty} |I^B_k(t, x; s, y)| \leq 2C_1 \psi_0(t - s, x, y)(t - s)^{-\frac{d}{2}} e^{-\frac{C_2|x-y|^2}{4(t-s)}}. \] (3.11)

We also define \( J^B_k(t, x; s, y) \) for \( k \in \mathbb{N}_0 \) and \( (t, x, y) \in (s, s + t_0] \times D \times D \) by
\[ J^B_0(t, x; s, y) := \nabla_x p^0_D(t, x; s, y), \]
\[ J^B_{k+1}(t, x; s, y) := \int_D \int_s^t J^B_k(t, x; v, z) \nabla_z p^0_D(v, z; s, y) \cdot B(v, z) \, dv \, dz. \]

As in the inequality (3.9),
\[ |J^B_k(t, x; s, y)| \leq C_3 \left( 1 + \frac{\rho(y)}{\sqrt{t-s}} \right)(t - s)^{-\frac{d+1}{2}} e^{-\frac{C_2|x-y|^2}{4(t-s)}} \left( C_3 M \sum_{i=1}^d N_{B_i}(t - s) \right)^k. \] (3.12)

Note that (3.10) and (3.12) imply that for \( (t, x, y) \in (s, s + t_0] \times D \times D \),
\[ \nabla_x p_D(t, x; s, y) = \sum_{k=0}^{\infty} J^B_k(t, x; s, y) \] (3.13)
and
\[ |\nabla_x p_D(t, x; s, y)| \leq \sum_{k=0}^{\infty} |J^B_k(t, x; s, y)| \leq 2C_3 \left( 1 + \frac{\rho(y)}{\sqrt{t-s}} \right)(t - s)^{-\frac{d+1}{2}} e^{-\frac{C_2|x-y|^2}{4(t-s)}}. \] (3.14)

Now we are going to prove the lower estimate of \( p_D(t, x; s, y) \). Combining (3.8), (3.10), and (3.9), we have for every \( (t, x, y) \in (s, s + t_0] \times D \times D \),
\begin{align*}
|p_D^0(t, x; s, y) - p_D(t, x; s, y)| &\leq \sum_{k=1}^{\infty} |t_k^0(t, x; s, y)| \\
&\leq 2C_1C_3M \sum_{i=1}^{d} N_{B_i}(t_0) \psi_0(t - s, x, y)(t - s)^{-d/2} e^{-c_5|x-y|^2/(4(t-s))}.
\end{align*}

By Theorem 2.1,
\begin{equation}
p_D^0(t, x; s, y) \geq C_1^{-1} \psi_0(t - s, x, y)(t - s)^{-d/2} e^{-c_5|x-y|^2/(4(t-s))},
\end{equation}
so we have for $|x - y| \leq \sqrt{t-s}$ and $(t, x, y) \in (s, s + t_0) \times D \times D$,
\begin{equation}
p_D(t, x; s, y) \geq \left( C_1^{-1} e^{-C_2^{-1}} - 2C_1C_3M \sum_{i=1}^{d} N_{B_i}(t_0) \right) \psi_0(t - s, x, y)(t - s)^{-d/2}.
\end{equation}

Now we choose $t_1 \leq t_0$ small so that $C_1C_3M \sum_{i=1}^{d} N_{B_i}(t_1) < (2C_1)^{-1} e^{-C_2^{-1}}$. So for $(t, x, y) \in (s, s + t_1) \times D \times D$ and $|x - y| \leq \sqrt{t-s}$, we have
\begin{equation}
p_D(t, x; s, y) \geq (2C_1)^{-1} e^{-C_2^{-1}} \psi_0(t - s, x, y)(t - s)^{-d/2}.
\end{equation}

Now one can follow the proof in [34, Lemma 2.1] and conclude that there exist $C_4 \in (0, C_2/2)$ and $t_2 \leq t_1$ such that
\begin{equation}
p_D(t, x; s, y) \geq c_5 \psi_0(t - s, x, y)(t - s)^{-d/2} e^{-c_5|x-y|^2/(4(t-s))},
\end{equation}
for all $(t, x, y) \in (s, s + t_2) \times D \times D$.

Now using the Chapman–Kolmogorov equation, we can extend the two-sided estimates to any $T > 0$.

**Theorem 3.8.** Let $D$ be $C^{1,\alpha}$-domain satisfying the connected line condition with characteristics $(R_0, \Lambda_0)$ and $(\chi_1, \chi_2)$, respectively. Suppose that $B(t, x) = (B^1(t, x), \ldots, B^d(t, x))$ is a bounded vector-valued function. Then $p_D(t, x; s, y)$, which is the heat kernel $p_D(t, x; s, y)$ of $\partial_t - L - B \cdot \nabla_x$ in $D$, and $\nabla_x p_D$ are jointly continuous on $(t, x; s, y) \in (0, \infty) \times D \times (0, t) \times D$. Furthermore, for each $T > 0$, there exist positive constants $c_j = c_j(d, \lambda, \chi_1, \chi_2, \psi, T, \Lambda_0, R_0, \alpha, B)$, $1 \leq j \leq 3$, depending on $B$ only via the rate at which $\max_{1 \leq i \leq d} N_{B_i}(r)$ goes to zero, such that
\begin{align*}
c_1 \psi_0(t - s, x, y) \Gamma_{c_2}(t, x; s, y) &\leq p_D(t, x; s, y) \leq c_3 \psi_0(t - s, x, y) \Gamma_{c_2/4}(t, x; s, y) \\
\text{and} \\
|\nabla_x p_D(t, x; s, y)| &\leq c_3 \psi_1(t - s, y) \Gamma_{c_2/4}(t, x; s, y),
\end{align*}
for all $x, y \in D$ and $0 \leq s < t \leq s + T$.

**Proof.** Without loss of generality, we assume $s = 0$. Recall that $C_4 \geq C_2/2$ and put $\nu = \frac{C_2}{4C_4}$. By (3.7), (3.11) and (3.17), for $t \leq t_2$,
Lemma 3.9.

\[ p_D(2t, x; 0, y) \geq c_1 \int_D \left( 1 \wedge \rho(x) \right) \left( 1 \wedge \rho(z) \right) e^{-c_2 \frac{|x-z|^2}{4t}} \]
\[ \times \left( 1 \wedge \rho(z) \right) \left( 1 \wedge \rho(y) \right) e^{-c_2 \frac{|y-z|^2}{4t}} \, dz \]
\[ \geq c_2 \int_D p_D(2v, x; v, z) p_D(v, z; 0, y) \, dz \]
\[ = c_2 p_D(2v, x; 0, y) \geq c_3 \psi_0(2t, x, y) \Gamma_{4C^2_2}(2t, x; 0, y). \]

On the other hand, by (3.7) and (3.11), for \( t \leq t_2 \),

\[ p_D(2t, x; 0, y) \leq c_4 \left( 1 \wedge \frac{\rho(x)}{\sqrt{t}} \right) t^{-d/2} e^{-c_2 \frac{|x|^2}{4t}} t^{-d/2} e^{-c_2 \frac{|y|^2}{4t}} \, dz \]
\[ \leq c_5 \left( 1 \wedge \frac{\rho(x)}{\sqrt{2t}} \right) \left( 1 \wedge \frac{\rho(y)}{\sqrt{2t}} \right) (2t)^{-d/2} e^{-c_2 \frac{|x|^2}{4t}} \]

and, by (3.14),

\[ \left| \nabla_x p_D(2t, x; 0, y) \right| \leq \int_D \left| \nabla_x p_D(2t, x; t, z) \right| p_D(t, z; 0, y) \, dz \]
\[ \leq c_6 \left( 1 \wedge \frac{\rho(y)}{\sqrt{t}} \right) t^{-1/2} \int_D t^{-d/2} e^{-c_2 \frac{|x-z|^2}{4t}} t^{-d/2} e^{-c_2 \frac{|y-z|^2}{4t}} \, dz \]
\[ \leq c_7 \left( 1 \wedge \frac{\rho(y)}{\sqrt{2t}} \right) (2t)^{-d/2} e^{-c_2 \frac{|x|^2}{4t}}. \]

Repeating these procedures for finite number of times, we proved the theorem for fixed \( T > 0 \). Joint continuity of \( \nabla_x p_D \) follows from Theorem 2.6 and (3.13). \( \square \)

Let \( \tilde{p}_D \) be the heat kernel of \( \partial_t - \tilde{\mathcal{L}} + \nabla_x \cdot \tilde{B} \), \( \tilde{\mathcal{L}}u = \partial_t (\tilde{a}_{ij} \partial_{x_j} u) \), \( \tilde{a}_{ij}(v, z) := a_{ij}(T - v, z) \), \( \tilde{B}(v, z) := B(T - v, z) \), i.e.,

\[ \int_0^T \int_D \tilde{p}_D(t, x; s, y) \frac{\partial}{\partial t} \phi(t, x) + \sum_{i,j=1}^d \tilde{a}_{ij}(t, x) \frac{\partial}{\partial x_j} \tilde{p}_D(t, x; s, y) \frac{\partial}{\partial x_i} \phi(t, x) \, dt \, dx \]
\[ - \sum_{i=1}^d \int_0^T \int_D \tilde{B}(v, z) \tilde{p}_D(t, x; s, y) \frac{\partial}{\partial x_i} \phi(t, x) \, dx \, dt = \phi(s, y) \] (3.18)

for any \( \phi \in C_c^\infty((0, T) \times D) \), \( \lim_{t \to s^+} \int_D \tilde{p}_D(t, x; s, y) u_0(y) \, dy = u_0(x) \) for any \( x \in D \), \( u_0 \in C_c^\infty(D) \), and \( \tilde{p}_D(t, x; s, y) = 0 \) for any \( x \in \partial D \).

The next lemma will be used in the next subsection.

Lemma 3.9. Suppose \( \tilde{B} \) is smooth. Then \( p_D(t, x; s, y) = \tilde{p}_D(T - s, y; T - t, x) \).
Proof. Note that satisfies (1.3)–(1.4). By the argument of Theorem 2.6, we may assume the coefficient is smooth such that and are smooth unless \((t, x) \neq (s, y)\). For \(0 \leq s < t \leq T\),

\[
\int_{D} \int_{s+\epsilon}^{t-\epsilon} p_D(v, z; s, y) \partial_v \bar{p}_D(T - v, z; T - t, x) + \partial_v p_D(v, z; s, y) \bar{p}_D(T - v, z; T - t, x) \, dv \, dz
\]

\[
= \int_{D} p_D(v, z; s, y) \bar{p}_D(T - v, z; T - t, x) \big|_{s+\epsilon}^{t-\epsilon} \, dz.
\]

Since and are smooth, by (1.5) and (3.18),

\[
\partial_v p_D(v, z; s, y) = \sum_{i,j=1}^{d} \partial_{z_i} (a_{ij}(v, z) \partial_{z_j} p_D(v, z; s, y)) + \sum_{i=1}^{d} B^{i}(v, z) \partial_{z_i} p_D(v, z; s, y)
\]

and

\[
- \partial_v \bar{p}_D(T - v, z; T - t, x)
\]

\[
= \sum_{i,j=1}^{d} \partial_{z_i} (a_{ij}(v, z) \partial_{z_j} \bar{p}_D(T - v, z; T - t, x)) - \sum_{i=1}^{d} \partial_{z_i} (B^{i}(v, z) \bar{p}_D(T - v, z; T - t, x)).
\]

Using the fact that \(p_D(v, z; s, y) = \bar{p}_D(T - v, z; T - t, x) = 0\) for any \(z \in \partial D\), and integrating by parts,

\[
\int_{D} p_D(s + \epsilon, z; s, y) \bar{p}_D(T - s - \epsilon, z; T - t, x) \, dz = \int_{D} p_D(t - \epsilon, z; s, y) \bar{p}_D(T - t + \epsilon, z; T - t, x) \, dz.
\]

Taking \(\epsilon \to 0^+\), and by (F3),

\[
\bar{p}_D(T - s, y; T - t, x) = p_D(t, x; s, y). \quad \Box
\]

3.4. Perturbation for general case

In the remainder of this section, we assume \(\mu = (\mu^1, \ldots, \mu^d)\) belongs to \(K_d\) and fix it. The next lemma extends [22, Lemma 3.2] to the parabolic Kato class.

Lemma 3.10. For any bounded \((d - 1)\)-rectifiable subset \(A\) of \(\mathbb{R}^d\) and \(0 < T_1 \leq T_2 < \infty\), we have \(\sum_{i=1}^{d} |\mu^i|(\{T_1, T_2\} \times A) = 0\).

Proof. Since

\[
|\mu^i|(\{t - h^2, t\} \times B(x, R)) \leq 2^{(d+1)/2} h^{d+1} e^{C_2 R^2/h^2} \int_{[t-h^2, t+h^2] \times B(x, R)} (t+h^2-s)^{-1/2} \mathcal{I}_c(t, x; s-h^2, y) \, d|\mu^i|(s, y)
\]

\[
\leq 2^{(d+1)/2} h^{d+1} e^{C_2 R^2/h^2} N_{\mu^i}(2h^2),
\]

for any \(R, h > 0\).
\begin{align}
\sup_{(t,x) \in [t-h^2, t] \times B(x, R)} |\mu^1(t - h^2, t) \times B(x, R)| \leq 2^{(d+1)/2}h^{d+1}eC_2R^2h^2N_{\mu^1}(2h^2). \quad (3.19)
\end{align}

Let \(N(A, \varepsilon)\) be the smallest number of \(\varepsilon\)-balls needed to cover \(A\). So for each \(\varepsilon > 0\), there exists a sequence \(\{x_j\}_{1 \leq j \leq N(A, \varepsilon)}\) such that

\[
\sum_{i=1}^{d} |\mu^1([T_1, T_2] \times A)| \leq \sum_{j=1}^{N(A, \varepsilon)} \sum_{i=1}^{d} |\mu^1([T_1, T_2] \times B(x_j, \varepsilon)), \quad 1 \leq i \leq d.
\]

For each \(\varepsilon > 0\), we define \(t^\varepsilon_k := T_1 + (k-1)\varepsilon^2, k \geq 1\) and let \(N_\varepsilon\) be the smallest integer greater than or equal to \((T_2 - T_1)/\varepsilon^2\). Using (3.19), we get

\[
\sum_{i=1}^{d} |\mu^1([T_1, T_2] \times A)| \leq \sum_{j=1}^{N(A, \varepsilon)} \sum_{k=1}^{N_\varepsilon} \int d|\mu^1|(t, x) \\
\leq 2^{(d+1)/2}(T_2 - T_1 + \varepsilon^2) \varepsilon^{-1} eC_2 N(A, \varepsilon) N_{\mu^1}(2\varepsilon^2).
\]

Let \(A_\varepsilon := \{x \in \mathbb{R}^d: \text{dist}(x, A) \leq \varepsilon\}\). It is well known (see, for instance, [26, (5.4) and (5.6)]) that there exists a positive number \(c_1 = c_1(d)\) such that \(\varepsilon dN(A, \varepsilon) \leq c_1 m_\varepsilon(A_\varepsilon)\). Since \(A\) is \((d-1)\)-rectifiable, by [16, Theorem 3.2.39], there exists \(c_2 = c_2(A) > 0\) such that \(\lim_{\varepsilon \downarrow 0} \varepsilon^{-1} m_\varepsilon(A_\varepsilon) = c_2 < \infty\). Thus we have for any \(1 \leq i \leq d\),

\[
\sum_{i=1}^{d} |\mu^i([T_1, T_2] \times A)| \leq c_3 \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} m_\varepsilon(A_\varepsilon) N_{\mu^i}(2\varepsilon^2) = c_3 c_2 \lim_{\varepsilon \downarrow 0} N_{\mu^i}(2\varepsilon^2) = 0. \quad \square
\]

We fix a non-negative smooth function \(\varphi(t, x) = \Phi(t, |x|)\) in \(\mathbb{R}^{d+1}\) with \(\text{supp}[\varphi] \subset (0, 1) \times B(0, 1)\) and \(\iint \varphi(t, x) dt \, dx = 1\). For any positive integer \(n\), we put \(\varphi_n(t, x) = 2^{n(d+2)}\varphi(2^n t, 2^n x)\). For \(1 \leq i \leq d\), define

\[
U^1_n(t, x) := \int_{\mathbb{R}^{d+1}} \varphi_n(t - s, x - y) d\mu^1(s, y), \quad n \in \mathbb{N}. \quad (3.20)
\]

Put \(U_n(t, x) = (U^1_n(t, x), \ldots, U^d_n(t, x))\) and set

\[
d\mu^1_n(t, x) = U^1_n(t, x) dt \, dx \quad \text{and} \quad \mu_n := (\mu^1_n, \ldots, \mu^d_n). \quad (3.21)
\]

Note that \(U_n\) is bounded and smooth. The following lemma is simple, but essential in this paper. Recall that we fixed a signed measure \(\mu = (\mu^1, \ldots, \mu^d)\) belonging to \(K_d\).

**Lemma 3.11.** Each \(U^i_n\) belongs to the Kato class \(K_d\). Moreover \(N_{\mu_n}(h) \leq N_{\mu^i}(h)\) for every \(h > 0, 1 \leq i \leq d\).

**Proof.** By a change of variable,

\[
\int_{\mathbb{R}^d} \int_{t-h}^t (t - s)^{-1/2} \Gamma_3(t, x; s, y) |U^1_n| s \, ds \, dy
\]
\[
\begin{align*}
&\leq \int_{\mathbb{R}^d} \int_{t-h}^t \int_{\mathbb{R}^d} \int_{s-v}^s (t-s)^{-1/2} \Gamma_{C_3}(t, x; s, y) \varphi_n(s-v, y-z) \, ds \, dy \\
&\quad \; \, ds \\
&= \int_{\mathbb{R}^d} \int_{t-h-v}^t \int_{\mathbb{R}^d} \int_{r+w}^r (t-r-v)^{-1/2} \Gamma_{C_3}(t, x; r+v, z+w) \varphi_n(r, w) \, dr \, dw \\
&\quad \, ds \\
&= \int_{\mathbb{R}^d} \varphi_n(r, w) \int_{t-h-v}^t \int_{r+w}^r (t-r-v)^{-1/2} \Gamma_{C_3}(t, x; r+v, z+w) \, dr \, dw.
\end{align*}
\]

Similarly,

\[
\begin{align*}
&\int_{\mathbb{R}^d} \int_{s}^{s+h} (t-s)^{-1/2} \Gamma_{C_3}(t, x; s, y) \Big| U_n^i(t, x) \Big| dt \, dx \\
&\quad \, ds \\
&\leq \int_{\mathbb{R}^d} \varphi_n(r, w) \int_{s-r}^{s+h-r} \int_{s-v}^s (v+r-s)^{-1/2} \Gamma_{C_3}(v+r, z+w; s, y) \, dr \, dw.
\end{align*}
\]

Hence

\[
N_{\mu_n}(h) \leq \int_{\mathbb{R}^d} \varphi_n(r, w) N_{\mu_n}(h) \, dr \, dw = N_{\mu_n}(h).
\]

**Lemma 3.12.** Let \(0 < T_0 < T_1 < \infty\) and \(0 < \delta < T_0/2\). Suppose \(K\) is a compact subset of \(\mathbb{R}^d\) and \(U\) is a bounded domain with smooth boundary \(\partial U\). Then we have the following.

(i) For every \(s \geq 0\) and any uniformly continuous function \(f(t, v, x, y, z)\) on

\[E_s := \{(t, v, x, y, z): (t, x, y, z) \in [s + T_0, s + T_1] \times K \times K \times \overline{U}, \ v \in [s + \delta, t - \delta]\};\]

we have that for \(1 \leq i \leq d\),

\[
\lim_{n \to \infty} \sup_{(t, x, y) \in [s + T_0, s + T_1] \times K \times K} \left| \int_{s+\delta}^{t-\delta} f(t, v, x, y, z) (d\mu_n^i - d\mu_i^i)(v, z) \right| = 0.
\]

(ii) For every \(t \geq T_1\) and any uniformly continuous function \(g(s, v, x, y, z)\) on

\[F_t := \{(s, v, x, y, z): (s, x, y, z) \in [t - T_1, t - T_0] \times K \times K \times \overline{U}, \ v \in [s + \delta, t - \delta]\};\]

we have that for \(1 \leq i \leq d\),

\[
\lim_{n \to \infty} \sup_{(s, x, y) \in [t - T_1, t - T_0] \times K \times K} \left| \int_{s+\delta}^{t-\delta} g(s, v, x, y, z) (d\mu_n^i - d\mu_i^i)(v, z) \right| = 0.
\]
Proof. Proof of (ii) is similar to the one of (i). We give the proof of (i) only. Fix \( i, s \geq 0 \). Let \( E := E_s \) and extend \( f \) to be zero off \( E \). Let
\[
A_n := \{ w \in \mathbb{R}^d : \text{dist}(\partial U, w) \leq 2^{-n} \}.
\]

By Lemma 3.10, we have
\[
\lim_{n \to \infty} |\mu^i ([s + \delta - 2^{-2n}, s + t] \times A_n) = |\mu^i ([s + \delta, s + t] \times \partial U) = 0.
\]

Given \( \epsilon > 0 \), choose a large positive integer \( n_1 \) such that for every \( n \geq n_1 \),
\[
\left( \sup_{(t, v, x, y, z) \in E} |f(t, v, x, y, z)| \right) |\mu^i ([s + \delta - 2^{-2n}, s + t] \times A_n) < \frac{\epsilon}{4}. \tag{3.22}
\]

Let
\[
\psi_n^{t, x, y}(f)(a, w) := \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} \varphi_n(v - a, z - w) f(t, v, x, y, z) dv dz
\]
\[
= \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} \varphi(v, z) f(t, 2^{-2n} v + a, x, y, 2^{-n} z + w) dv dz. \tag{3.23}
\]

By definition of \( \mu_n^i \), we have for every \( (t, x, y) \in [s + T_0, s + T_1] \times K \times K \),
\[
\int_{U \times s + \delta} \int_{s + \delta}^{t - \delta} f(t, v, x, y, z) (d\mu_n^i - d\mu^i) (v, z)
\]
\[
= \int_{U \times s + \delta} \int_{s + \delta}^{t - \delta} \left[ \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} \varphi_n(v - a, z - w) d\mu^i(a, w) \right] f(t, v, x, y, z) dv dz
\]
\[
- \int_{U \times s + \delta} \int_{s + \delta}^{t - \delta} f(t, v, x, y, z) d\mu^i(v, z).
\]

Since \( f \) is zero outside \( E \), we extend the domain of \( z \) and \( v \), and then apply Fubini's theorem. Using the fact that \( \varphi_n(v - s, z - w) \) is non-zero only if \( s + \delta - 2^{-2n} \leq a \leq t - \delta \) and \( w \in U \cup A_n \), the above integral becomes
\[
\int_{U \cup A_n} \int_{s + \delta - 2^{-2n}}^{t - \delta} \left( \psi_n^{t, x, y}(f)(a, w) - f(t, a, x, y, w) \right) d\mu^i(a, w).
\]

By (3.22), for every large \( n \geq n_1 \) with \( 2^{-2n+1} < \delta \),
Using Lemma 3.11, we can choose $t_0 > 0$ small such that

$$C_3 M \sum_{i=1}^{d} N_{\mu_i}^i (t_0') \leq C_3 M \sum_{i=1}^{d} N_{\mu_i}^i (t_0') \leq \frac{1}{2}.$$  

(3.26)
As (3.10) and (3.13), we have for \((t, x, y) \in (s, s + t_0') \times D \times D,
\[
q_D^n(t, x; s, y) = \sum_{k=0}^{\infty} I_k^n(t, x; s, y), \quad \nabla_x q_D^n(t, x; s, y) = \sum_{k=0}^{\infty} J_k^n(t, x; s, y)
\]
and, using (3.26), by the same arguments as the ones in (3.9) and (3.12), we have, for \(k \in \mathbb{N}_0\) and \((t, x, y) \in (s, s + t_0') \times D \times D,
\[
\left| I_k^n(t, x; s, y) \right| + \left| J_k^n(t, x; s, y) \right| \leq C_1 2^{-k} \psi_0(t - s, x, y) (t - s)^{-\frac{d}{2}} e^{-\frac{C_t|x-y|^2}{4(t-s)}},
\]
(3.27)
\[
\left| I_k^n(t, x; s, y) \right| + \left| J_k^n(t, x; s, y) \right| \leq C_1 2^{-k} \psi_1(t - s, y) (t - s)^{-\frac{d}{2}} e^{-\frac{C_t|x-y|^2}{4(t-s)}}.
\]
(3.28)

With these inequalities, for \(t \in (s, s + t_0')\), we can define
\[
q_D(t, x; s, y) := \sum_{k=0}^{\infty} I_k^n(t, x; s, y), \quad r_D(t, x; s, y) := \sum_{k=0}^{\infty} J_k^n(t, x; s, y).
\]

Moreover, for \(t \in (s, s + t_0')\),
\[
q_D^n(t, x; s, y) + q_D(t, x; s, y) \leq 2C_1 \psi_0(t - s, x, y) (t - s)^{-\frac{d}{2}} e^{-\frac{C_t|x-y|^2}{4(t-s)}},
\]
\[
\left| \nabla_x q_D^n(t, x; s, y) \right| + \left| r_D(t, x; s, y) \right| \leq 2C_1 \psi_1(t - s, y) (t - s)^{-\frac{d}{2}} e^{-\frac{C_t|x-y|^2}{4(t-s)}}.
\]

In the remainder of this section, we will show that for each \(s \geq 0\), \(q_D^n(t, x; s, y)\) and \(\nabla_x q_D^n(t, x; s, y)\) converge uniformly on every compact subset of \(\{(t, x, y) \in (s, \infty) \times D \times D\}\), and show that for each \(t > 0\), \(q_D^n(t, x; s, y)\) and \(\nabla_x q_D^n(t, x; s, y)\) converge uniformly on every compact subset of \(\{(s, x, y) \in (0, t) \times D \times D\}\).

**Lemma 3.13.** Suppose that \(R\) is a positive number. Then, for any \(a > 0\), there exists positive constant \(c\) depending only on \(a\) and \(d\) such that for any measure \(\nu\) on \(\mathbb{R}^d\) and \(t > 0\),
\[
(i) \sup_{|x|, |y| < R} \int \int_{|x-z| \geq 4R} \frac{t}{s} \left( t - v \right)^{-d/2} e^{-\frac{a|x-z|^2}{t-s}} (v-s)^{-\frac{d+1}{2}} e^{-\frac{a|x-y|^2}{t-s}} \, dv(v, z)
\]
\[
\leq c(t-s)^{-d/2} e^{-\frac{a|x-y|^2}{t-s}}
\]
\[
\times \sup_{|w| < R} \int \int_{|w-z| \geq 2R} \left( t - v \right)^{-\frac{d+1}{2}} e^{-\frac{a|w-z|^2}{4(t-s)}} + (v-s)^{-\frac{d+1}{2}} e^{-\frac{a|w-y|^2}{4(t-s)}} \, dv(v, z),
\]
and
\[
(ii) \sup_{|x|, |y| < R} \int \int_{|x-z| \geq 4R} \frac{t}{s} \left( t - v \right)^{-\frac{d+1}{2}} e^{-\frac{a|x-z|^2}{t-s}} (v-s)^{-\frac{d+1}{2}} e^{-\frac{a|x-y|^2}{t-s}} \, dv(v, z)
\]
Therefore,

\[ \lim_{R \to \infty} \sup_{n: 2^{-2n+4} \leq \delta} \int_{|w-z| \geq 2R} \int_{s+\delta}^t (v-s)^{-(d+1)/2} e^{-\frac{\|w-z\|^2}{2(v-s)}} (d|\mu_n| + d|\mu^i|)(v, z) = 0, \]

\[ \lim_{R \to \infty} \sup_{n: 2^{-2n+4} \leq \delta} \int_{|w-z| \geq 2R} \int_s^{t-\delta} (t-v)^{-(d+1)/2} e^{-\frac{\|w-z\|^2}{2(v-s)}} (d|\mu_n| + d|\mu^i|)(v, z) = 0. \]

**Proof.** We will prove the first claim only. Since \( \varphi \) is a non-negative radial function supported by \((0, 1) \times B(0, 1)\), we have for any \( w \in \mathbb{R}^d \) and \( 2^{-2n+4} \leq \delta \),

\[ \int \int_{|w-z| \geq 2R} e^{-\frac{\|w-z\|^2}{2(v-s)}} d|\mu_n^i|(v, z) \]

For \( |x|, |y| < R \), if \( z \) satisfies \( |x-z| \geq 4R \), we have \( |y-z| \geq |x-z| - |x-y| \geq 4R - |x-y| > 2R \). Therefore,

\[ \int_{|x-z| \geq 4R} \int_{s+\delta}^t (v-s)^{-(d+1)/2} e^{-\frac{\|w-z\|^2}{2(v-s)}} d|\mu_n|(v, z) \leq \int_{|y-z| \geq 2R} \int_s^{t-\delta} (t-v)^{-(d+1)/2} e^{-\frac{\|w-z\|^2}{2(v-s)}} d|\mu_n|(v, z), \]

which ends the proof of (i). (ii) follows similarly. \( \square \)

**Lemma 3.14.** If \( \mu \in K_d \), we have that for any \( t, s, \delta > 0 \) with \( t-s > \delta \) and \( 1 \leq i \leq d \),
\[ \int_{|w-z| \geq 2R} \int_{s+\delta}^{t} \int \int \int e^{-\frac{C_2 |w-z|^2}{(w-v-y)^2}} \varphi_n(v-a, z-y) \, d\mu_i^j(a, y) \, dv \, dz. \]

Since \(|w-y| \leq |y-z| + |w-z| < 2^{-n} + |w-z| < (2^{-(n+1)} R^{-1} + 1)|w-z|\) and \(v-s \leq 2^{-2n} + (a-s) \leq \delta(2^{-2n})^{-1}(a-s) \leq 16 \delta\) in the integral above, for large \(R\), \(e^{-\frac{C_2 |w-z|^2}{(w-v-y)^2}} \leq e^{-\frac{C_2 |w-y|^2}{20|v-y|^2}}\). Thus for such large \(R\)'s, we get

\[ \int_{|w-z| \geq 2R} \int_{s+\delta}^{t} \int \int \int e^{-\frac{C_2 |w-y|^2}{20|v-y|^2}} \varphi_n(v-a, z-y) \, d\mu_i^j(a, y) \, dv \, dz \]

Thus

\[ \sup_{n: 2^{-2n+4} \leq \delta} \int_{|w-z| \geq 2R} \int_{s+\delta}^{t} (v-s)^{-\frac{(d+1)}{2}} e^{-\frac{C_2 |w-y|^2}{20|v-y|^2}} (d|\mu_i^j| + d|\mu_i^j|)(v, z) \]

\[ \leq 2\delta^{-\frac{d+1}{2}} \sup_{n: 2^{-2n+4} \leq \delta} \int_{|w-z| \geq 2R} \int_{s+\delta}^{t} e^{-\frac{C_2 |w-z|^2}{20|v-z|^2}} \, d|\mu_i^j| (v, z) \]

\[ \leq 2\delta^{-\frac{d+1}{2}} (t-s)^{\frac{d+1}{2}} \sup_{n: 2^{-2n+4} \leq \delta} \int_{|w-z| \geq 2R} \int_{s+\delta}^{t} (v-s)^{-\frac{(d+1)}{2}} e^{-\frac{C_2 |w-z|^2}{22|v-z|^2}} \, d|\mu_i^j| (v, z) \]

\[ \leq 2\delta^{-\frac{d+1}{2}} (t-s)^{\frac{d+1}{2}} e^{-\frac{3C_2 R^2}{220|v-z|^2}} N^{\mu_i}_{\mu_i}(t-s), \]

which goes to zero as \(R \to \infty\) since \(\mu_i^j\) belongs to the Kato class \(K_d\). \(\square\)

In the remainder of this section, \(t_0'\) stands for the constant \(t_0'\) defined in (3.26). Recall that we fixed \(\mu = (\mu_1, \ldots, \mu_d)\) belonging to \(K_d\).

**Lemma 3.15.** For any compact subsets \(K_1\) and \(K_2\) of \(D\), and \(T_0 \in (0, t_0')\), we have for all \(k \geq 0\),

\[ \lim_{n \to \infty} \sup_{(x,y) \in [s+t_0, s+t_0'] \times K_1 \times K_2} |l_i^{n+1}(t, x; s, y) - l_i^{n}(t, x; s, y)| = 0, \quad \forall s \geq 0, \]

\[ \lim_{n \to \infty} \sup_{(s,x,y) \in [t-t_0', t-T_0] \times K_1 \times K_2} |l_i^{n+1}(t, x; s, y) - l_i^{n}(t, x; s, y)| = 0, \quad \forall t \geq t_0'. \]
Proof. The proof is almost identical to the proof of Lemma 3.16 below. We skip the proof. □

Lemma 3.16. For any compact subsets $K_1$ and $K_2$ of $D$, and $T_0 \in (0, t'_0]$, we have for all $k \geq 0$,

$$\lim_{n \to \infty} \sup_{(t, x, y) \in [s + T_0, s + t'_0] \times K_1 \times K_2} |J^h_n(t, x; s, y) - J^h_k(t, x; s, y)| = 0, \quad \forall s \geq 0,$$

$$\lim_{n \to \infty} \sup_{(t, x, y) \in [t - t'_0, t - T_0] \times K_1 \times K_2} |J^h_n(t, x; s, y) - J^h_k(t, x; s, y)| = 0, \quad \forall t \geq t'_0.$$

Proof. Without loss of generality, we may assume that $T_0 < 2$ and we will prove the lemma when $K_1$ and $K_2$ are equal and are subsets of $B(0, \overline{c})$ for some $r > 0$. We let $K := K_1 = K_2$. The lemma is clearly valid for $k = 0$. We assume that the lemma is true for $k$, which implies that $J^h_k(t, x; s, y)$ is continuous for $t > s$.

Take $R > r$. Since $|x| < R$, we have $|x - z| > |z| - |x| > 4R$ for $|z| \geq 5R$. If $t \in [s + T_0, s + t'_0]$, by (2.9), (3.28), and Lemma 3.13(iii),

$$A(n, t, x; s, y) := \sum_{i=1}^{d} \left( \int_{|z| \geq 5R} \int_{s}^{t} |J^h_n(t, x; v, z)| \left| \partial_z\left( p_D^0(v, z, s, y) \right) \right| d|\mu_n^i|(v, z) \right.$$  

$$+ \int_{|z| \geq 5R} \int_{s}^{t} |J^h_k(t, x; v, z)| \left| \partial_z\left( p_D^0(v, z, s, y) \right) \right| d|\mu_k^i|(v, z) \right)$$  

$$\leq c_1 \sum_{i=1}^{d} \int_{|x - z| \geq 4R} \int_{s}^{t} (t - v)^{-\frac{d+1}{2}} e^{-\frac{c_2|x - y|^2}{4(v - s)}} (v - s)^{(d+1)/2}$$  

$$\times e^{-\frac{c_2|x - y|^2}{4(v - s)}} (d|\mu_n^i| + d|\mu_k^i|)(v, z)$$  

$$\leq c_2 T_0^{(d+1)/2} \sum_{i=1}^{d} \sup_{|w| < R} \int_{|w - z| \geq 2R} \int_{s}^{t} (t - v)^{-\frac{d+1}{2}} e^{-\frac{a|w - z|^2}{4(v - s)}}$$  

$$+ (v - s)^{(d+1)/2} e^{-\frac{a|w - z|^2}{4(v - s)}} (d|\mu_n^i| + d|\mu_k^i|)(v, z). \quad (3.29)$$

For any given $\varepsilon > 0$, by Lemma 3.11, we first choose $\delta < T_0/2$ such that

$$\sum_{i=1}^{d} N_{\mu_k^i}(\delta) + \sum_{i=1}^{d} \sup_{n \geq 1} N_{\mu_n^i}(\delta) < \frac{\varepsilon T_0^{(d+1)/2}}{2c_0}.$$

For this $\delta$, by Lemma 3.14, we can choose $R$ large enough so that

$$\sup_{n: 2^{-2n+4} \leq \delta} \int_{|w - z| \geq 2R} \int_{s + \delta}^{t} (t - v)^{-\frac{d+1}{2}} e^{-\frac{a|w - z|^2}{4(v - s)}} (d|\mu_n^i| + d|\mu_k^i|)(v, z)$$
\[
\int_{\mathbb{R}^d} (v - s)^{-(d+1)/2} e^{-\frac{|v-s|^2}{2(\sigma^2 + 1)}} (d|\mu^I_n| + d|\mu^I|)(v, z) + \int_{t-\delta}^{t} \int_{s+\delta}^{s+t_0-\delta} (t - v)^{-(d+1)/2} e^{-\frac{C_2|x-v|^2}{2(\sigma^2 + 1)}} (d|\mu^I_n| + d|\mu^I|)(v, z),
\]

which is finite. If \( \delta \) where

\[
\sup_{n: 2^{-2n+4} \leq \delta} \frac{\varepsilon T^{(d+1)/2}}{2C_0 d}.
\]

Therefore, by (3.29) and (3.30),

\[
\sup_{(t, x, y) \in [s + T_0, s + t_0] \times K \times K} A(n, t, x; s, y) < \varepsilon.
\]

Choose a smooth domain \( D_1 \subset \overline{D_1} \subset D \) such that for every \( z \in D \setminus D_1 \),

\[
\rho(z) < d_1 := \varepsilon \left( 1 + \frac{\varepsilon_0^{d/2}}{4C_1 L_1} \right),
\]

where \( \delta \) is to be chosen later, and

\[
L_1 = L_1(s, \delta) := \sup_{x \in K, n \geq 1} \sum_{i=1}^{d} \int_{D^i} \int_{s+\delta}^{s+t_0-\delta} (t - v)^{-(d+1)/2} e^{-\frac{C_2|x-v|^2}{2(\sigma^2 + 1)}} (d|\mu^I_n| + d|\mu^I|)(v, z),
\]

which is finite. If \( D = \mathbb{R}^d \), take \( D_1 = D = \mathbb{R}^d \).

Now we fix \( D' := D \cap B(0, 5R) \) and \( D'_1 := D_1 \cap B(0, 5R) \). We consider \( n \) with \( 2^{-2n+4} \leq \delta \) only. Let

\[
I(n, t, x; s, y) := \sum_{i=1}^{d} \left( \int_{D'_i \setminus D'_1} \int_{s+\delta}^{s+t_0-\delta} \int_{s}^{s+\delta} J_k^\mu(t, x; v, z) \left| \partial_z p_D^0(v, z; s, y) \right| d|\mu^I_n|(v, z)
\]

\[
+ \int_{D'_i \setminus D'_1} \int_{s+\delta}^{s+t_0-\delta} \int_{s}^{s+\delta} J_k^\mu(t, x; v, z) \left| \partial_z p_D^0(v, z; s, y) \right| d|\mu^I|(v, z),
\]

\[
II(n, t, x; s, y) := \sum_{i=1}^{d} \left( \int_{D'} \int_{s}^{s+\delta} \int_{s}^{s+\delta} J_k^\mu(t, x; v, z) \left| \partial_z p_D^0(v, z; s, y) \right| d|\mu^I_n|(v, z)
\]

\[
+ \int_{D'} \int_{s}^{s+\delta} \int_{s}^{s+\delta} J_k^\mu(t, x; v, z) \left| \partial_z p_D^0(v, z; s, y) \right| d|\mu^I|(v, z),
\]

\[
III(n, t, x; s, y) := \sum_{i=1}^{d} \left( \int_{D'} \int_{t-\delta}^{t} \int_{s+\delta}^{s+t_0-\delta} \int_{s}^{s+\delta} J_k^\mu(t, x; v, z) \left| \partial_z p_D^0(v, z; s, y) \right| d|\mu^I_n|(v, z)
\]

\[
+ \int_{D'} \int_{t-\delta}^{t} \int_{s+\delta}^{s+t_0-\delta} \int_{s}^{s+\delta} J_k^\mu(t, x; v, z) \left| \partial_z p_D^0(v, z; s, y) \right| d|\mu^I|(v, z),
\]

for every

\[ V(t, v, z) := \sum_{i=1}^{d} \int_{D_i}^{t-\delta} (t) \partial_z p_D(v, z; s, y) \, d\mu^i_n(v, z) \]

and

\[ V(n, t, x; s, y) := \sum_{i=1}^{d} \int_{D_i}^{t-\delta} \left| f_n(t, x; v, z) - f_n(t, x; v, z) \right| \partial_z p_D(v, z; s, y) \, d\mu^i_n(v, z) \]

Then we have

\[ |f_{n+1}(t, x; s, y) - f_n(t, x; s, y)| \leq C \sup_{w \in \mathbb{R}^d} \int_{s}^{s+\delta} (v - s)^{-1/2} e^{-\frac{C_1 |w - z|^2}{4(t - s)}} \left( d|\mu^i_n| + d|\mu^i| \right)(v, z) \]

Since \( \delta < T_0/2 \), by (3.27)-(3.28), we have

\[ II \leq C_2 2^{-2k} \left( \frac{2}{T_0} \right) \sum_{i=1}^{d} \int_{D_i}^{t-\delta} (t - v)^{-1/2} e^{-\frac{C_2 |w - z|^2}{4(t - s)}} \left( d|\mu^i_n| + d|\mu^i| \right)(v, z) \]

Here we used the inequality of \( T_0/2 = T_0 - T_0/2 \leq t - s - T_0/2 < t - s - \delta \leq t - v \). Similarly,

\[ III \leq C_2 2^{-2k} \left( \frac{2}{T_0} \right) \sum_{i=1}^{d} \int_{D_i}^{t-\delta} (v - s)^{-1/2} e^{-\frac{C_2 |w - z|^2}{4(t - s)}} \left( d|\mu^i_n| + d|\mu^i| \right)(v, z) \]

Given \( \epsilon > 0 \), using Lemma 3.11, we choose \( \delta < T_0/2 \) such that \( II \) and \( III \) are less than or equal to \( \epsilon/8 \) for every \( n \geq 1 \).

On the other hand, by (3.27), (3.28), and (3.31), we have

\[ \sup_{(x, y) \in K \times K, n \geq 1} I(n, t, x; s, y) \]

\[ \leq C_1 \sup_{x \in K, n \geq 1} \sum_{i=1}^{d} \int_{D_i}^{t-\delta} \left( 1 + \frac{\rho(z)}{\sqrt{t - v}} \right) (t - v)^{-1/2} (v - s)^{-1/2} \left( d|\mu^i_n| + d|\mu^i| \right)(v, z) \]

\[ \leq C_2 d_1 \delta^{-\frac{d+2}{2}} \sup_{x \in K, n \geq 1} \sum_{i=1}^{d} \int_{D_i}^{t-\delta} (t - v)^{-1/2} e^{-\frac{C_1 |w - z|^2}{4(t - s)}} \left( d|\mu^i_n| + d|\mu^i| \right)(v, z) \]
Then, by Lemma 3.15, for large $\frac{s+\delta}{\delta}$

\[\sup_{x \in K, n \geq 1} \sum_{i=1}^{d} \int_{\delta}^{s+\delta} \int_{s}^{t} (t - v)^{-(d+1)/2} e^{-\frac{C_2(t-v)}{4\|\mu^i_n\|}} (d|\mu^i_n| + d|\mu^j_n|)(v, z),\]

which is less than $\frac{1}{4}$ by definition of $d_1$.

Now we estimate IV. Let $f_i(t, v, x, y, z) := j_k^1(t, x; v, z) \partial_2 p^0(v, z; s, y)$. Since $j_k^1(t, x; v, z)$ and $\partial_2 p^0(v, z; s, y)$ are continuous by Theorem 2.6, $f_i(t, v, x, y, z)$ is uniformly continuous on $(t, x, y, z) \in [s + T_0, s + t'_0) \times K_1 \times K_1 \times \tilde{D}_{t_1}$. $v \in [s + \delta, t - \delta]$. Therefore, by Lemma 3.12,

\[\lim_{n \to \infty} \sup_{(t, x, y) \in [s + \delta, t - \delta]} IV(n, t, x; s, y) = 0.\]

Finally we estimate $V$. From (3.28), we easily see that

\[\sup_{(t, x, y) \in [s + T_0, s + t'_0] \times K \times K} V(n, t, x; s, y) \leq C_1 2^{-\frac{\delta}{2}} \sum_{i=1}^{d} \sup_{(t, x) \in [s + T_0, s + t'_0] \times K}_{s+\delta} \int_{s}^{t} j_k^1(t, x; v, z) - j_k^1(t, x; v, z) d|\mu^i_n|(v, z).\]

Thus, by the assumption on $j_k^1$, (3.28) and the dominated convergence theorem,

\[\lim_{n \to \infty} \sup_{(t, x, y) \in [s + T_0, s + t'_0] \times K \times K} V(n, t, x; s, y) = 0.\]

The proof for (ii) is similar. \(\square\)

**Theorem 3.17.**

(i) For every $s \geq 0$, $q_k^0(t, x; s, y)$ and $\nabla_x q_k^0(t, x; s, y)$ uniformly converge on any compact subset of $\{(t, x, y) : x, y \in D, t \in (s, \infty)\}$.

(ii) For every $t > 0$, $q_k^0(t, x; s, y)$ and $\nabla_x q_k^0(t, x; s, y)$ uniformly converge on any compact subset of $\{(x, s, y) : x, y \in D, s \in [0, t]\}$.

**Proof.** We fix $s \geq 0$ and prove the convergence of $q_k^0$ in (i) only. Let $K$ be any compact subset of $\{(t, x, y) : x, y \in D, t \in (s, s + t'_0]\}$. Given $\epsilon > 0$, using (3.27), choose $M$ such that

\[\sup_{n \geq 1} \left( \sum_{k=M}^{\infty} \left| I_k^i(t, x; s, y) \right| + \sum_{k=M}^{\infty} \left| I_k^j(t, x; s, y) \right| \right) < \epsilon/2.\]

Then, by Lemma 3.15, for large $n$

\[\sup_{(t, x, y) \in K} \left| q_k^0(t, x; s, y) - q_D(t, x; s, y) \right| \leq \sup_{(t, x, y) \in K} \sum_{k=1}^{M} \left| I_k^i(t, x; s, y) - I_k^j(t, x; s, y) \right| + \epsilon/2 < \epsilon.\]

Now we show the convergence for $t \in (s + t'_0, s + 2t'_0)$. Let $K_1, K_2$ be compact subsets of $D$ and $D_r := \{z \in D : \rho(z) < r\}$ for $r > 0$. The uniform upper bounds of $q_k^0$ and $q_D$ imply that
that the general case can be proved by repeating the previous argument. 

\[
\sup_{n \geq 1, t \in (s+t_{0}^{r}+\frac{2}{3}t_{0}^{r})} \int_{D_{r}} q_{D}^{n}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}(t - \frac{t_{0}^{r}}{2}, z; s, y) 
+ q_{D}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}(t - \frac{t_{0}^{r}}{2}, z; s, y) dz 
\leq c_{1}t_{0}^{r-d} \left( \int_{\{z \in D_{r} : |z| \leq R\}} \left( 1 + \frac{\rho(z)}{\sqrt{t_{0}^{r}}} \right) dz + \int_{\{z \in D_{r} : |z| \geq R\}} \exp \left( -c_{2} \frac{|z|^{2}}{t_{0}^{r}} \right) dz \right)
\]

for some positive constants \(c_{1}\) and \(c_{2}\), where \(R > 0\) is chosen so that for any given \(\varepsilon > 0\),

\[
c_{1}t_{0}^{r-d} \int_{\{z \in D_{r} : |z| \geq R\}} \exp \left( -c_{2} \frac{|z|^{2}}{t_{0}^{r}} \right) dz < \varepsilon/4.
\]

Then we choose \(0 < r < (t_{0}^{r})^{1/2}\) small such that \(c_{1}t_{0}^{r-d} |\{z \in D_{r} : |z| < R\}| < \frac{\varepsilon}{4}\). By (3.7), we have for \((t, x, y) \in (s+t_{0}^{r}, s+\frac{3}{2}t_{0}^{r}) \times K_{1} \times K_{2}\),

\[
\left| q_{D}^{n}(t, x; s, y) - \int_{D} q_{D}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}(t - \frac{t_{0}^{r}}{2}, z; s, y) dz \right|
= \left| \int_{D} q_{D}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}^{n}(t - \frac{t_{0}^{r}}{2}, z; s, y) - q_{D}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}(t - \frac{t_{0}^{r}}{2}, z; s, y) dz \right|
\leq c_{3} \sup_{t \in (s+t_{0}^{r}+\frac{2}{3}t_{0}^{r}), \rho(z) \geq r} \left| q_{D}^{n}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}(t - \frac{t_{0}^{r}}{2}, z; s, y) - q_{D}(t, x; t - \frac{t_{0}^{r}}{2}, z) q_{D}(t - \frac{t_{0}^{r}}{2}, z; s, y) \right| + \frac{\varepsilon}{2},
\]

for some positive constant \(c_{3}\). The first term in the last line above goes to zero as \(n \to \infty\) by the uniform convergence of \(q_{D}^{n}(t, x; s, y)\) on compact subsets of \(\{(t, x, y) : x, y \in D, t \in (s, s+t_{0}^{r})\}\). The general case can be proved by repeating the previous argument. \(\Box\)

We extend \(q_{D}(t, x; s, y)\) to every \(t \in (s, \infty)\) by

\[
q_{D}(t, x; s, y) = \lim_{n \to \infty} q_{D}^{n}(t, x; s, y).
\]

Since \(q_{D}(t, x; s, y)\) is the uniform limit of jointly continuous functions on any compact subset of \(\{(t, x, s, y) : x, y \in D, t \in (s, \infty)\}\), it is jointly continuous on \(\{(t, x, s, y) : x, y \in D, t \in (s, \infty)\}\). Using Theorem 3.17, we see that for any \((t, x) \in (s, \infty) \times D\) and any relatively compact open subset \(D_{0} \subseteq D\), the sequence of functions \(\{q_{D}^{n}(t, x; s, y) : n \geq 1\}\) is a Cauchy sequence in the Banach space \(C^{1}(D_{0})\), thus \(\nabla_{s} q_{D}(t, x; s, y)\) exists for every \((t, x, y) \in (s, \infty) \times D \times D\). By [32, Theorem 7.17], it follows that \(\nabla_{x} q_{D}(t, x; s, y)\) is equal to \(r_{D}(t, x; s, y) = \lim_{n \to \infty} \nabla_{y} q_{D}^{n}(t, x; s, y)\).

Moreover, by Theorem 3.17 and its proof, we have the Chapman–Kolmogorov equation for \(q_{D}\):

\[
q_{D}(t, x; s, y) = \int_{D} q_{D}(t, x; v, z) q_{D}(v, z; s, y) dz, \quad \forall x, y \in D, 0 < s < v < t.
\]
Theorem 3.18. Let $D$ be $C^{1,\alpha}$-domain satisfying the connected line condition with characteristics $(R_0, A_0)$ and $(\chi_1, \chi_2)$, respectively. Suppose that $\mu(t, x) = (\mu^1(t, x), \ldots, \mu^d(t, x))$ belongs to the Kato class $K_d$. Then, $q_D(t; x; s; y)$ defined in (3.32) is a heat kernel of $\partial_t - L - \mu \cdot \nabla_x$ in $D$. Furthermore, $q_D(t; x; s; y)$ and $\nabla_x q_D(t; x; s; y)$ are jointly continuous on $(t, x; s, y) \in (0, T) \times D \times (0, t) \times D$ for each $T > 0$, and there exist positive constants $c_j = c_j(d, \lambda, \chi_1, \chi_2, \psi, \nu, A_0, R_0, \alpha, \mu)$, $1 \leq j \leq 2$, depending on $\mu$ only via the rate at which $\max_{1 \leq i \leq d} N_{\mu^i}^0(t)$ goes to zero, such that

\[
 c_1^{-1} \psi_0(t - s, x, y) \Gamma_{c_2}(t; x, s, y) \leq q_D(t; x; s, y) \leq c_1 \psi_0(t - s, x, y) \Gamma_{c_2}(t; x, s, y) \tag{3.34}
\]

and

\[
 |\nabla_x q_D(t; x, s; y)| \leq c_1 \psi_1(t - s, y) \Gamma_{c_2}(t; x, s, y), \tag{3.35}
\]

for all $(t, x; s, y) \in (0, T] \times D \times (0, t) \times D$.

**Proof.** The inequalities (3.34) and (3.35) follow from Theorem 3.8, Lemma 3.11. Recall $I_k^\mu$ that we defined in (3.24), and $q_D$ defined using $I_k^\mu$. Also, note

\[
 I_{k+1}^{\mu}(t; x, s; y) = \int_{Q_T} \cdots \int_{Q_T} p_D^0(t; x, v_1, z_1) \nabla_{z_1} p_D^0(v_1, z_1; v_2, z_2) \cdot d\mu(v_1, z_1) \\
 \cdots \nabla_{z_{k+1}} p_D^0(v_{k+1}, z_{k+1}; s, y) \cdot d\mu(v_{k+1}, z_{k+1}).
\]

We will show that $q_D$ is indeed the heat kernel of $\partial_t - L - \mu \cdot \nabla_x$. For this, it is enough to show that $q_D$ satisfies (F1). Let $I_{1,\phi}^\mu(s, y) := \int_{Q_T} \phi(t; x) \nabla_x p_D^0(t; x, s; y) \cdot d\mu(t, x)$. Note that

\[
 \int_{Q_T} -p_D^0(t; x, s; y) \partial_t \phi(t; x) + \sum_{i,j=1}^d a_{ij}(t; x) \partial_{x_i} p_D^0(t; x, s; y) \partial_{x_j} \phi(t; x) \ dt \ dx \\
 - \int_{Q_T} \phi(t; x) \nabla_x p_D^0(t; x, s; y) \cdot d\mu(t, x) \\
 = \phi(s, y) - I_{1,\phi}^\mu(s, y).
\]

Moreover, with

\[
 I_{k+1}^{\mu,\phi}(s, y) := \int_{Q_T} \cdots \int_{Q_T} \phi(v_1, z_1) \nabla_{z_1} p_D^0(v_1, z_1; v_2, z_2) \cdot d\mu(v_1, z_1) \\
 \cdots \nabla_{z_{k+1}} p_D^0(v_{k+1}, z_{k+1}; s, y) \cdot d\mu(v_{k+1}, z_{k+1}),
\]

we also have

\[
 \int_{Q_T} -I_{k+1}^{\mu}(t; x, s; y) \partial_t \phi(t; x) + \sum_{i,j=1}^d a_{ij}(t; x) \partial_{x_i} I_{k+1}^{\mu}(t; x, s; y) \partial_{x_j} \phi(t; x) \ dt \ dx \\
 - \int_{Q_T} \phi(t; x) \nabla_x I_{k+1}^{\mu}(t; x, s; y) \cdot d\mu(t, x) \\
 = I_{k+1}^{\mu,\phi}(s, y) - I_{k+2}^{\mu,\phi}(s, y).
\]
Note that \(|I_{k+1}^{\mu,\phi}(s, y)| \leq \|\phi\|_\infty (N_\mu(T))^{k+1}\) and
\[
\int_{Q_T} -q_D(t, x; s, y) \partial_t \phi(t, x) + \sum_{i,j=1}^d a_{ij}(t, x) \partial_{x_j} q_D(t, x; s, y) \partial_{x_i} \phi(t, x) \, dt \, dx \\
- \int_{Q_T} \phi(t, x) \nabla_s q_D(t, x; s, y) \cdot d\mu(t, x) \\
= \phi(s, y) - I_{1}^{\mu,\phi}(s, y) + I_{2}^{\mu,\phi}(s, y) - I_{2}^{\mu,\phi}(s, y) + \cdots.
\]

Choosing small \(T\), we proved (F1). For large \(T\), we use (3.33) for \(q_D(t, x; s, y)\). Joint continuity of \(\nabla_s q_D\) follows from Theorem 3.17 and the continuity of \(\nabla_s q_D^0\).

**Remark 3.20.** The computations above show that our result is quite optimal. Namely, the condition of \(\mu \in K_d\) is necessary and sufficient for the convergence of \(I_k^{\mu}\). With Hölder continuous \(a_{ij}\) and in the whole domain, the results were obtained in [33]. Theorem 3.18 extends all results in [30, Chapter 3] to \(C^{1,\alpha}\)-domains.

**Remark 3.21.** Let \(\tilde{q}_D\) be the heat kernel of \(\partial_t - \tilde{L} + \nabla_x \cdot \tilde{\mu}\), \(\tilde{L} u = \partial_{x_j} (\tilde{a}_{ij} \partial_{x_j} u), \tilde{a}_{ij}(v, z) = a_{ij}(T - v, z), d\tilde{\mu}(v, z) = d\mu(T - v, z)\), i.e.,
\[
\int_{D}^T 0 \int \tilde{q}_D(t, x; s, y) \frac{\partial}{\partial t} \phi(t, x) + \sum_{i,j=1}^d \tilde{a}_{ij}(t, x) \frac{\partial}{\partial x_j} \tilde{q}_D(t, x; s, y) \frac{\partial}{\partial x_i} \phi(t, x) \, dt \, dx \\
- \sum_{i=1}^d \int_{Q_T} \tilde{q}_D(t, x; s, y) \frac{\partial}{\partial x_i} \phi(t, x) d\tilde{\mu}_i(t, x) = \phi(s, y)
\]
for any \(\phi \in C^\infty_c((0, T) \times D)\).

Note \(\tilde{a}\) satisfies (1.3)-(1.4). By Lemma 3.9, Theorems 3.17 and 3.18, it is easy to see that \(q_D(t, x; s, y) = \tilde{q}_D(T - s, y; T - t, x)\).

For applications to the next section, we prove the following:

**Theorem 3.22.** Let \(D\) be \(C^{1,\alpha}\)-domain satisfying the connected line condition with characteristics \((R_0, \Lambda_0)\) and \((\chi_1, \chi_2)\), respectively. Suppose that \(v(t, x) = (v^1(t, x), \ldots, v^d(t, x))\) belongs to the parabolic Kato class \(K_d\). Then there exists a heat kernel \(q_D(t, x; s, y)\) of \(\partial_t - L + \nabla_x \cdot v\) in \((0, T) \times D\). Furthermore, \(q_D(t, x; s, y)\) is jointly continuous on \((t, x; s, y) \in (0, T) \times D \times (0, t) \times D\) for each \(T > 0\), and there exist positive constants \(c_j = c_j(d, \lambda, \chi_1, \chi_2, \psi, T, A_0, R_0, \alpha, v), 1 \leq j \leq 2\), depending on given quantities and \(v\) only via the rate at which \(\max_{1 \leq i \leq d} N_{\nu^i}(r)\) goes to zero, such that
\[
c_1^{-1} \psi_0(t - s, x, y) \Gamma_{c_1}^{-1}(t, x; s, y) \leq q_D(t, x; s, y) \leq c_1 \psi_0(t - s, x, y) \Gamma_{c_2}^{-1}(t, x; s, y) \tag{3.36}
\]
for all \((t, x; s, y) \in (0, T) \times D \times (0, t) \times D\).
Specifically, $d\hat{v}(v, z) := d\nu(T-v, z)$, then $\hat{q}^v_0(T-s, y; T-t, x) = q^v_0(t, x; s, y)$. Now (3.36) is clear by Theorem 3.18. □

In fact, one can modify the proof of Theorem 3.18 and prove the gradient estimates also. We skip the details.

Remark 3.23. Let $\hat{L} u = \partial_t + \partial_x$ and $\hat{q}^v_0$ be the heat kernel of the adjoint operator of $\partial_x + \nu \cdot v$. Specifically, $(\partial_t + \partial_x + \nu \cdot v)\hat{w} := (-\partial_t + \partial_x + \nu \cdot \nabla_y)\hat{w}(s, y) = f$ in $Q_T$ if and only if

$$
\int_{Q_T} w(s, y)\partial_y \phi(s, y) + \sum_{i, j=1}^d a_{ij}(s, y)\partial_{y_j} w(s, y)\partial_{y_i} \phi(s, y) \, ds \, dy
$$

$$
- \int_{Q_T} \sum_{i=1}^d \partial_{y_i} w(s, y)\phi(s, y) \, d\nu^i(s, y)
$$

$$
= \int_{Q_T} f(s, y)\phi(s, y) \, ds \, dy
$$

for any $\phi \in C_c^\infty(Q_T)$. One can follow the idea of Theorem 3.18 to construct $\hat{q}^v_0$. Then $\hat{q}^v_0$ satisfies all the corresponding properties in Theorem 3.18.

Moreover, $q^v_0(t, x; s, y) = \hat{q}^v_0(s, y; t, x)$. This claim can be proved following the idea of proof for [2, Theorem 9(i)] or our argument in Remark 3.21. So we skip the proof.

Similar to (IVP), if we define $u$ to be $u(s, y) := \int_\mathbb{R} q^v_0(t, x; s, y) f(x) \, dx$, then $u$ solves the following in a weak sense:

$$
\begin{align*}
(\partial_t + \hat{L} + v \cdot \nabla_y)u &= 0 & \text{in } (0, t) \times D, \\
u(s, y) &= 0 & \text{for } y \in \partial D, \\
\lim_{s\to t^-} u(s, y) &= f(y) & \text{for } y \in D.
\end{align*}
$$

4. Time-inhomogeneous diffusions with measure-valued drifts

In this section, we discuss probabilistic counterpart of the main results of Section 3. Recall that $U_n(s, x) = (U^1_n(s, x), \ldots, U^n_n(s, x))$ is defined in (3.20), and consider the following final value problem:

$$
(\partial_t + L + U_n \cdot \nabla_x)u(s, x) = 0 & \text{in } [0, t] \times \mathbb{R}^d \quad \text{and} \quad u(t, \cdot) = f(\cdot). \quad (4.1)
$$

Since we will follow the standard notation to the probabilistic literature in this section (we mainly follow the notations in [19]), we will use the pair $(s, x)$, $(t, y)$ instead of $(t, x)$, $(s, y)$. For convenience, throughout this section, we fix $T > 0$ and consider the time interval $[0, T]$ only. We use $\mathcal{B}(\mathbb{R}^d)$ and $\mathcal{B}([0, T])$ to denote the Borel $\sigma$-algebra on $\mathbb{R}^d$ and $[0, T]$, respectively.

From (3.37), replacing $\nu$ by $U_n$, $\hat{L}$ by $L$,

$$
T^n_{s, t} f(x) = \begin{cases} 
\int_{\mathbb{R}^d} f(y)\hat{q}^n(s, x; t, y) \, dy & \text{if } 0 \leq s < t \leq T, \\
f(x) & \text{if } s = t \text{ and } 0 \leq s \leq T,
\end{cases}
$$

for every $f \in C_c(\mathbb{R}^d)$, is the solution of (4.1), where $\hat{q}^n(s, x; t, y)$ is the heat kernel of $\partial_t + L + U_n \cdot \nabla_x$ (in the final value problem).
Recall that a heat kernel of \( \partial_t + L + U_n \cdot \nabla_x \) in \((0, T) \times \mathbb{R}^d\) is the Borel function \(\tilde{q}^n(s, x; t, y)\) satisfying the following two conditions: For any fixed \((t, y) \in (0, T) \times \mathbb{R}^d\),
\[
\int_{(0, T) \times \mathbb{R}^d} -\hat{q}^n(s, x; t, y) \frac{\partial}{\partial s} \phi(s, x) - \sum_{i,j=1}^d a_{ij}(s, x) \frac{\partial}{\partial x_j} \hat{q}^n(s, x; t, y) \frac{\partial}{\partial x_i} \phi(s, x) \, ds \, dx \\
+ \int_{(0, T) \times \mathbb{R}^d} \sum_{i=1}^d \frac{\partial}{\partial x_i} \hat{q}^n(s, x; t, y) U^i_n(s, x) \phi(s, x) \, ds \, dx = \phi(t, y)
\]
for any \(\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)\), and
\[
\lim_{s \to t^-} \int_{\mathbb{R}^d} \hat{q}^n(s, x; t, y) u_0(y) \, dy = u_0(x) \quad \text{for any } x \in D, \ u_0 \in C_c^\infty(\mathbb{R}^d).
\]
On the other hand, from the definition and the existence and uniqueness of weak solutions to the final value problem on the space of the bounded continuous functions, \(\hat{q}^n(s, x; t, y)\) satisfies the Chapman–Kolmogorov equation and \(\int \hat{q}^n(s, x; t, y) \, dy = 1\). Thus, \(\hat{q}^n(s, x; t, y)\) is a transition density. See [19, Definition 1.2] and the remark after [19, Theorem 3.4]. Let \(\Omega := (\mathbb{R}^d)^{0,T}\). We define a process \(X^n\) and a probability measure \(P^n_{s,x}\) by \(X^n_t(w) = w(t)\) for \(w \in \Omega\), and
\[
P^n_{s,x}(X^n_{t_1} \in B_1, \ldots, X^n_{t_k} \in B_k)
= \int_{B_1} \hat{q}^n(s, x; t_1, x_1) \int_{B_2} \hat{q}^n(t_1, x_1; t_2, x_2) \cdots \int_{B_k} \hat{q}^n(t_{k-1}, x_{k-1}; t_k, x_k) \, dx_k \cdots dx_1.
\]
where \(s \leq t_1 < \cdots < t_k \leq T\), and \(B_j \in B(\mathbb{R}^d), 1 \leq j \leq k\). For the notational convenience, we use \(P^n_{s,x,X^n_t} \in \cdot\) instead of \(P^n_{s,x,X^n_t} \in \cdot\).

Let \(\mathcal{F}^n_{s,t}\), \(0 \leq s \leq t \leq T\), be the \(\sigma\)-algebra generated by \(X^n_t\) with \(s \leq v \leq t\) and let \(\mathcal{G}^n_{s,t} = \mathcal{F}^n_{s,t}\) be the completion of \(\mathcal{F}^n_{s,t}\) with respect to \(V_s = \{P^n_{v,x}: s \leq v \leq t, x \in \mathbb{R}^d\}\), which is right continuous, i.e., \(\mathcal{G}^n_{s,t} = \mathcal{G}^n_{v,t}\) (see [19, Lemma 2.18]).

The following proposition tells us that the process \(X^n\) can be considered as a standard process with continuous sample paths. We refer to [19] for the definition of strongly continuous backward Feller–Dynkin propagator and standard process.

**Proposition 4.1.** For every \(n \geq 1\), \(T^n_{s,t}\) is a strongly continuous backward Feller–Dynkin propagator associated with the transition density \(\hat{q}^n(s, x; t, y)\). Moreover, there exists a modification of \(X^n\) which is standard process whose sample paths are continuous.

**Proof.** Using Gaussian estimates of \(\hat{q}^n(s, x; t, y)\), one can check Kolmogorov’s criterion easily (see [19, Theorem 3.9]). Thus, by [19, Theorem 2.22], we only need to show that \(T^n_{s,t}\) is a strongly continuous backward Feller–Dynkin propagator in \(C_0(\mathbb{R}^d)\).

Clearly, using the Gaussian estimates of \(\hat{q}^n(s, x; t, y)\) (Theorem 3.22) and the dominated convergence theorem, we can show that \(T^n_{s,t}f(x) \to 0\) as \(|x| \to \infty\) for \(f \in C_0(\mathbb{R}^d)\). Moreover, since
\[
\left| T^n_{s,t}f(x_0) - T^n_{s,t}f(x) \right|
= \left| \int_{\mathbb{R}^d} f(x_0)\hat{q}^n(s, x_0; t, y) \, dy - \int_{\mathbb{R}^d} f(x)\hat{q}^n(s, x; t, y) \, dy \right|
\]
is continuous. Thus can apply Proposition 4.2.

Since our propagator is locally uniformly bounded, by [19, Theorem 2.1], to show $T^n_{s,t} f$ is strongly continuous, it is enough to show that it is strongly continuous in $t$ and $s$ separately. Moreover, since for $s \leq v < t$

$$\|T^n_{s,t} f - T^n_{v,t} f\|_{\infty} = \|T^n_{v,t} T^n_{s,v} f - T^n_{v,t} T^n_{v,v} f\|_{\infty} \leq \|T^n_{s,v} f - T^n_{v,v} f\|_{\infty} \leq \|T^n_{s,v} f - T^n_{v,v} f\|_{\infty},$$

we will show that for every fixed $s$ and $f \in C(\mathbb{R}^d)$, the function $t \to T^n_{s,t} f$ is continuous on $[s, T]$.

Fix $f \in C(\mathbb{R}^d)$ and $0 \leq s < T$. Recall that, by Theorem 3.22, $\hat{q}^n(s, x; t, y) \leq c_1 (t-s)^{-d/2} e^{-c_2 |x-y|^2} t^{s-t}$ for $0 < t - s < 1$ and let $c_3 := \int_{\mathbb{R}^d} e^{-c_2 |w|^2} dw$. Since $f$ is uniformly continuous, for any given $\varepsilon > 0$, there exists $\delta > 0$ such that if $|z| < \delta$, $\|f(\cdot + z) - f(\cdot)\|_{\infty} < \varepsilon/(2c_1c_3)$. There exists $t_\delta \in (0, 1)$ such that if $0 < v \leq t_\delta$, $\int_{|z|\geq \delta} e^{-c_2 |z|^2} d\varepsilon \leq \int_{|z|\geq \delta} e^{-c_2 |z|^2} d\varepsilon < \varepsilon/(4c_1\|f\|_{\infty})$.

Now using the fact that $T^n_{s,t} 1 = 1$, for every $s < t \leq t_\delta + s$,

$$\sup_{x \in \mathbb{R}^d} |T^n_{s,t} f(x) - T^n_{s,s} f(x)| \leq c_1 \sup_{x \in \mathbb{R}^d} \int_{|z| < \delta} |f(x + z) - f(x)| ((t-s)^{-d/2} e^{-c_2 |z|^2} t^{s-t}) d\varepsilon$$

$$+ c_1 \sup_{x \in \mathbb{R}^d} \int_{|z| \geq \delta} |f(x + z) - f(x)| (t-s)^{-d/2} e^{-c_2 |z|^2} t^{s-t}) d\varepsilon$$

$$\leq c_1 \frac{\varepsilon}{2c_1c_3} \int_{|z| \geq \delta} (t-s)^{-d/2} e^{-c_2 |z|^2} t^{s-t}) d\varepsilon + 2c_1\|f\|_{\infty} \int_{|z| \geq \delta} (t-s)^{-d/2} e^{-c_2 |z|^2} t^{s-t}) d\varepsilon < \varepsilon. \quad (4.3)$$

When $s \leq v < t \leq v + t_\delta$, $\|T^n_{s,t} f - T^n_{s,v} f\|_{\infty} = \|T^n_{v,t} T^n_{s,v} f - T^n_{v,v} T^n_{s,v} f\|_{\infty}$, and since $T^n_{s,v} f \in C(\mathbb{R}^d)$, we can apply $T^n_{s,v} f$ in (4.3) instead of $f$ and $v$ instead of $s$. 

Recall $\hat{q}(s, x; t, y) := \lim_{n \to \infty} \hat{q}^n(s, x; t, y)$ is the heat kernel of $\partial_s + L + \mu \cdot \nabla x$. Define

$$T_{s,t} f(x) = \begin{cases} \int_{\mathbb{R}^d} f(y) \hat{q}(s, x; t, y) dy & \text{if } 0 \leq s < t \leq T, \\ f(x) & \text{if } s = t \text{ and } 0 \leq s \leq T, \end{cases}$$

for every $f \in C(\mathbb{R}^d)$. Since $\hat{q}(s, x; t, y)$ also satisfies Chapman–Kolmogorov equation and $\int \hat{q}(s, x; t, y) dy = 1$, $\hat{q}(s, x; t, y)$ is also a transition density. We also define a process $X$ using $\hat{q}(s, x; t, y)$ and $P_{s,x}$ as in (4.2). Then we have a corresponding proposition to the previous one. The proof is almost identical, so we skip the proof.

**Proposition 4.2.** $T_{s,t}$ is a strongly continuous backward Feller–Dynkin propagator associated with the transition density $\hat{q}(s, x; t, y) dy$. Moreover, there exists a modification of $X$ which is standard process whose sample paths are continuous.
Due to Propositions 4.1 and 4.2, we will take the continuous versions of $X^n$ and $X$. Without loss of generality, we also assume $\Omega = C([0, T], \mathbb{R}^d)$. Under this setting we consider the weak convergence of the process $X^n$ to $X$. We first show the tightness of $X^n$. Without loss of generality, we consider the case $s = 0$ only. Recall that a sequence of $(X^n, P_{0,x})$ in $C([0, T], \mathbb{R}^d)$ with the uniform topology is tight if for every $\epsilon > 0$ there is a compact subset $K$ of $C([0, T], \mathbb{R}^d)$ such that $\sup_n P_{0,x}(X^n \notin K) \leq \epsilon$. The following lemma is sufficient to prove $X^n$ is tight due to [1].

**Proposition 4.3.** For any positive constants $T$, $\beta$ and $\epsilon$, there exists $\delta > 0$ independent of $x$ and $n$ such that

$$P_{0,x}\left( \sup_{t \leq \nu \leq t+\delta} |X^n_\nu - X^n_t| > \beta \right) \leq \epsilon, \quad \forall t \leq T - \delta.$$ 

**Proof.** Note that, by the Markov property of $X^n$ with respect to $\mathcal{F}_{t+n}^0$ (see [19, Theorem 2.13]),

$$P_{0,x}\left( \sup_{t \leq \nu \leq t+\delta} |X^n_\nu - X^n_t| > \beta \right) = E_{0,x}\left[ P_{t,\nu} \left( \sup_{t \leq \nu \leq t+\delta} |X^n_\nu - X^n_t| > \beta \right) \right]$$

$$\leq \sup_{y \in \mathbb{R}^d} P_{t,y}\left( \sup_{t \leq \nu \leq t+\delta} |X^n_\nu - X^n_t| > \beta \right).$$

Define $\mathcal{G}_{\nu}^{t,n}$-stopping time

$$S^n_t := \begin{cases} \inf\{v \geq t: |X^n_v - X^n_t| > \beta\} & \text{on } \cup_{t \leq \nu < T} \{|X^n_\nu - X^n_t| > \beta\}, \\ T & \text{elsewhere} \end{cases}$$

and

$$T^n_t := \begin{cases} \inf\{v \geq S^n_t: |X^n_v - X^n_{S^n_t}| > \beta/2\} & \text{on } \cup_{t \leq \nu < T} \{|X^n_\nu - X^n_{S^n_t}| > \beta/2\}, \\ T & \text{elsewhere} \end{cases}$$

Then, $\{T^n_t \leq \nu\}$ belongs to $\sigma$-algebra generated by $S^n \vee \rho, X^n_{S^n_t \vee \rho}, 0 \leq \rho \leq T$. Thus, by the strong Markov property with respect to $\{S^n_t, T^n\}$ and $\{P_{t,y}\}$ (cf. [19, Sections 2.7–2.10]),

$$P_{t,y}\left( |X^n_{t+\delta} - X^n_{S^n_t}| \geq \beta/2, \quad S^n_t \leq t + \delta \right)$$

$$= E_{t,y}\left[ P_{S^n_t, X^n_{S^n_t}} \left( \{X^n_{t+\delta} - X^n_{S^n_t}| \geq \beta/2: \quad S^n_t \leq t + \delta \right) \right]$$

and hence

$$P_{t,y}\left( \sup_{t \leq \nu \leq t+\delta} |X^n_\nu - X^n_t| > \beta \right)$$

$$\leq P_{t,y}(S^n_t \leq t + \delta)$$

$$\leq P_{t,y}(|X^n_{t+\delta} - X^n_t| \geq \beta/2) + P_{t,y}(|X^n_{t+\delta} - X^n_t| < \beta/2, \quad S^n_t \leq t + \delta)$$

$$\leq P_{t,y}(|X^n_{t+\delta} - X^n_t| \geq \beta/2) + P_{t,y}(|X^n_{t+\delta} - X^n_{S^n_t}| > \beta/2, \quad S^n_t \leq t + \delta)$$

$$\leq P_{t,y}(|X^n_{t+\delta} - X^n_t| \geq \beta/2) + E_{t,y}\left[ P_{S^n_t, X^n_{S^n_t}} \left( |X^n_{t+\delta} - X^n_{S^n_t}| \geq \beta/2: \quad S^n_t \leq t + \delta \right) \right]$$

$$\leq 2 \sup_{z \in \mathbb{R}^d, \xi \in [0, \delta]} P_{t+\delta, z}(|X^n_{t+\delta} - X^n_{t+\delta + \xi}| \geq \beta/2).$$
Therefore, by Theorem 3.22,

$$\sup_{n \geq 1} \mathbb{P}_{0,x} \left( \sup_{t \leq v \leq t + \delta} |X^n_v - X^n_t| > \beta \right) \leq 2 \sup_{n \geq 1, z \in \mathbb{R}^d, \xi \in [0, \delta]} \int_{B(z, \beta/2)^c} \hat{q}^n(t + \xi, z; t + \delta, w) \, dw$$

$$\leq c_1 \sup_{\xi \leq \delta} \int_{B(0, \beta/2 \sqrt{1 - \xi})} e^{-c_2 |y|^2} \, dy,$$

which is arbitrary small as $\delta$ goes to zero. □

For every open subset $U$ of $\mathbb{R}^d$ and $s \leq v \leq T$, let us define the first exit time of $U$ after time $v$ by

$$\tau^n_{U} = \left\{ \begin{array}{ll}
\inf\{t > v: X^n_t \notin U\} & \text{on } \bigcup_{s \leq t < T} \{t > v, X^n_t \notin U\}, \\
T & \text{elsewhere.}
\end{array} \right.$$  

We define for $0 \leq s < t \leq T$ and $P_{s,x}$,

$$X^{n,D}_t(\omega) = \left\{ \begin{array}{ll}
X^n_t(\omega) & \text{if } t < \tau^{n,s}_D(\omega), \\
\partial & \text{if } t \geq \tau^{n,s}_D(\omega),
\end{array} \right.$$  

where $\partial$ is a cemetery state. The process $(X^{n,D}, P_{s,x})$ is called the subprocess of $(X^n, P_{s,x})$ killed upon leaving $D$. For the remainder of this paper, we use the convention $f(\partial) = 0$.

For $0 \leq s \leq t \leq T$, $x, y \in \mathbb{R}^d$, let

$$\hat{q}^n_D(s, x; t, y) := \hat{q}^n(s, x; t, y) - \mathbb{E}_{s,x}[\hat{q}^n(\tau^{n,s}_D, t, y): \tau^{n,s}_D < t]. \quad (4.4)$$

Note that by the continuity of the sample paths of $X^n$, we have $\hat{q}^n_D(s, x; t, y) = 0$ for $x \in \mathbb{R}^d \setminus D$. Since

$$\mathbb{P}_{s,x}(X^n_t \in A, \tau^{n,s}_D < t) = \int_A \mathbb{E}_{s,x}[\hat{q}^n(\tau^{n,s}_D, t, y): \tau^{n,s}_D < t] \, dy,$$

and $\mathbb{P}_{s,x}(\tau^{n,s}_D = t) \leq \int_A \hat{q}^n(s, x; t, y) \, dy = 0$, for $0 \leq s \leq t \leq T$, we have $\mathbb{P}_{s,x}(X^{n,D}_t \in A) = \int_A \hat{q}^n_D(s, x; t, y) \, dy$. Hence, $\hat{q}^n_D(s, x; t, y)$ is the transition density of $X^{n,D}$. The following lemma is needed in Proposition 4.5.

**Lemma 4.4.** For every $x \in \partial D$, $s \geq 0$, $\mathbb{P}_{s,x}(\tau^{n,s}_D = s) = 1$.

**Proof.** Since $D$ is a $C^{1,0}$-domain, there exists a cone $A$ with vertex $x$ such that $A \cap B(x, r) \subset D^c$ for some $r > 0$. Without loss of generality, we assume $s = 0$, $x = 0$ and $r \leq 1$. Under $P_{0,x}$, we have

$$\mathbb{P}_{0,x}(\tau^{n,0}_D < t) \geq \mathbb{P}_{0,x}(X^n_t \in A \cap B(0, r)) \geq \mathbb{P}_{0,x}(X^n_t \in A) - \mathbb{P}_{0,x}(X^n_t \notin B(0, r))$$

$$= \int_A \hat{q}^n(0, 0; t, y) \, dy - \int_{B(0, r)^c} \hat{q}^n(0, 0; t, y) \, dy.$$

Using two-sided estimate of $\hat{q}^n$ and the change of variable $t^{-1/2}y = z$,
\[
\int_A \hat{q}^n(0, 0; t, y) dy \geq c_1 \int_A e^{-c_2 |z|^2} dz \quad \text{and} \\
\int_{B(0, r)^c} \hat{q}^n(0, 0; t, y) dy \leq c_3 \int_{B(0, r/\sqrt{t})^c} e^{-c_4 |z|^2} dz.
\]

Hence,
\[
P_{0,x}(\tau_D^{n,0} \leq t) \geq c_1 \int_A e^{-c_2 |z|^2} dz - c_3 \int_{B(0, r/\sqrt{t})^c} e^{-c_4 |z|^2} dz,
\]

which goes to \( c_1 \int_A e^{-c_2 |z|^2} dz > 0 \) as \( t \to 0 \). Now by zero-one law (for example, see [18, Theorem 2, p. 45]) and the fact \( \{\tau_D^{n,0} = 0\} \in \mathcal{G}_0 \), we conclude \( P_{0,x}(\tau_D^{n,0} = 0) = 1 \). \( \square \)

Using the fact that \( \hat{q}^n(s, x; t, y) \) is the heat kernel of \( \partial_s + L + \nabla \cdot \nabla_x \) in \( \mathbb{R}^d \) and (4.4) (see [19, Theorem 3.5]), we can prove the following.

**Proposition 4.5.** The function \( \hat{q}^n_{0,D} \), defined in (4.4) is the heat kernel of \( \partial_s + L + \nabla \cdot \nabla_x \) in \( D \).

**Proof.** Note that, by Remark 3.23, \( \hat{q}^n_{0,D}(s, x; t, y) \) is the heat kernel of \( \partial_s + L + \nabla \cdot \nabla_x \) in \( D \) if and only if it is the heat kernel of \( \partial_t - \hat{L} + \nabla_x \cdot U_n \) in \( D \) where \( \hat{L}u = \partial_x (a_{ij} \partial_x u) \). So we will prove the latter. Since, for any fixed \( (s, x) \in (0, T) \times D \) and any smooth function \( \phi \) vanishing near the parabolic boundary of \( (0, T) \times D \),

\[
\int_{(0, T) \times D} -\hat{q}^n(s, x; t, y) \frac{\partial}{\partial t} \phi(t, y) + \sum_{i,j=1}^d a_{ij}(t, y) \frac{\partial}{\partial y_j} \hat{q}^n(s, x; t, y) \frac{\partial}{\partial y_i} \phi(t, y) \, dt \, dy
\]

\[
- \int_{(0, T) \times D} \sum_{i=1}^d \hat{q}^n(s, x; t, y) U_n^i(t, y) \frac{\partial}{\partial y_i} \phi(t, y) \, dt \, dy
\]

\[
= \int_{(0, T) \times D} \hat{q}^n(s, x; t, y) \frac{\partial}{\partial t} \phi(t, y) + \sum_{i,j=1}^d a_{ij}(t, y) \frac{\partial}{\partial y_j} \hat{q}^n(s, x; t, y) \frac{\partial}{\partial y_i} \phi(t, y) \, dt \, dy
\]

\[
- \int_{(0, T) \times D} \sum_{i=1}^d \hat{q}^n(s, x; t, y) U_n^i(t, y) \frac{\partial}{\partial y_i} \phi(t, y) \, dt \, dy
\]

\[
+ \int_{(0, T) \times D} E_{s,x}[\hat{q}^n(\tau_{D}^{n,s}, x_{D}; t, y): \tau_{D}^{n,s} < t] \frac{\partial}{\partial t} \phi(t, y)
\]

\[
- \sum_{i,j=1}^d a_{ij}(t, y) \frac{\partial}{\partial y_j} E_{s,x}[\hat{q}^n(\tau_{D}^{n,s}, x_{D}; t, y): \tau_{D}^{n,s} < t] \frac{\partial}{\partial y_i} \phi(t, y)
\]

\[
+ \sum_{i=1}^d E_{s,x}[\hat{q}^n(\tau_{D}^{n,s}, x_{D}; t, y): \tau_{D}^{n,s} < t] U_n^i(t, y) \frac{\partial}{\partial y_i} \phi(t, y) \, dt \, dy.
\]
Using the Gaussian upper bound for \( \tilde{q}^n \) and \( \nabla \tilde{q}^n \) and the fact that \( \tilde{q}^n(s, x; t, y) \) is a heat kernel of \( \partial_t - L + \nabla_y \cdot U_n \) in \( \mathbb{R}^d \), the above is equal to

\[
\phi(s, x) + \mathbb{E}_{s,x} \left[ \int_{(0,T) \times D} \tilde{q}^n(\tau_D^{n,s}, \chi_{\tau_D^{n,s}}; t, y) \frac{\partial}{\partial t} \phi(t, y) \right] 
- \sum_{i,j=1}^d a_{ji}(t, y) \frac{\partial}{\partial y_j} \tilde{q}^n(\tau_D^{n,s}, \chi_{\tau_D^{n,s}}; t, y) \frac{\partial}{\partial y_i} \phi(t, y) 
+ \sum_{i=1}^d \tilde{q}^n(\tau_D^{n,s}, \chi_{\tau_D^{n,s}}; t, y) U_i(t, y) \frac{\partial}{\partial y_i} \phi(t, y) 
\]

thus, by (4.4),

\[
\lim_{t \to s+} \int_D \tilde{q}_D^n(s, x; t, y) u_0(y) \, dy 
= u_0(x) + \lim_{t \to s+} \int_D \mathbb{E}_{s,x} \left[ \tilde{q}^n(\tau_D^{n,s}, \chi_{\tau_D^{n,s}}; t, y) : \tau_D^{n,s} < t \right] u_0(y) \, dy.
\]

The second term is bounded above by \( \| u_0 \|_\infty P_{s,x}(\tau_D^{n,s} = s) \) which is zero by Lemma 4.4. \( \square \)

Recall that we have constructed a diffusion process \( X \) in Proposition 4.2. The following theorem is the main theorem in this section.

**Theorem 4.6.** Suppose that \( \mu \) belongs to the Kato class \( K_d \). Then, \( X^n \) converges weakly to \( X = \{X_t, P_{s,x} : s \leq t < T, \; x \in \mathbb{R}^d \} \). Moreover, \( \tilde{q}_D(s, x; t, y) = \lim_{n \to \infty} \tilde{q}_D^n(s, x; t, y) \) is transition density, which is jointly continuous, of the subprocess \( X^D \) of \( X \) killed upon leaving \( D \).

**Proof.** Without loss of generality, assume \( s = 0 \). By Proposition 4.3, the Markov property and [1], \( X^n \) is tight in \( C(0, T], \mathbb{R}^d) \). Since Theorem 3.17 tells us that every convergent sub-sequential limits of \( X^n \) are the same, \( X^n \) converges weakly to \( X \) in \( C(0, T], \mathbb{R}^d) \). Let \( \tau_D := \inf\{v > 0 : X_v \notin D\} \) be the first exit time of \( D \) for \( X \) under \( P_{0,x} \) and \( X^D \) be the subprocess of \( X \) killed upon leaving \( D \). Since \( X_t \) has continuous sample paths and \( D \) is a \( C^{1,\alpha} \)-domain, for every \( 0 < t < T \) and \( x \in \mathbb{R}^d \), \( P_{0,x}(\tau_D = t) \leq \int_D \tilde{q}_D(0, x; t, y) \, dy = 0 \). Thus \( \{t < \tau_D\} \) is open in \( C[0, T] \) whose boundary is \( P_{0,x} \)-null, and so for any bounded continuous function \( f \) in \( D \),

\[
\lim_{n \to \infty} \mathbb{E}_{0,x}[f(X^n_t) 1_{\{t < \inf\{v > 0 : X_v \notin D\}\}}] = \mathbb{E}_{0,x}[f(X_t) 1_{\{t < \tau_D\}}], \; \forall x \in D, \; t > 0
\]

(cf. Theorem 2.9.1(vi) in [14]). Therefore, by Theorem 3.17, \( \tilde{q}_D(0, x; t, y) := \lim_{n \to \infty} \tilde{q}_D^n(0, x; t, y) \) is the transition density for \( X^D \), which is jointly continuous. \( \square \)
For the remainder of this section, we assume $D$ is bounded $C^{1,\alpha}$-domain, and $a_{ij}$ and $\mu$ are time-independent. Define $q_D(t, x, y) := \hat{q}_D(0, x, t, y)$. In [23], the intrinsic ultracontractivity for $X^D$ is discussed when $a_{ij}$ is in $C^1$. Using the estimates in this paper, the intrinsic ultracontractivity for $X^D$ for the Dini continuous $a_{ij}$ can be proved exactly by the same way. We recall and use some results from [23]:

By [23, (5.26)], there exist $\mu_1 > 0$ and strictly positive continuous functions $\phi_1$ and $\psi_1$ such that $\|\phi_1\|_{L^2(D)} = \|\psi_1\|_{L^2(D)} = 1$ and, for every $(t, x, y) \in (0, \infty) \times D \times D$,

$$e^{\mu_1 t} \phi_1(x) = \int_D q_D(t, x, y) \phi_1(y) \, dy, \quad e^{\mu_1 t} \psi_1(x) = \int_D q_D(t, y, x) \psi_1(y) \, dy.$$  

Thus, by Theorem 3.22, we get

$$\phi_1(x) \asymp \rho(x) \quad \text{and} \quad \psi_1(x) \asymp \rho(x).$$  

(4.5)

Here and in the sequel, for two non-negative functions $f$ and $g$, the notation $f \asymp g$ means that there are positive constants $c_1$ and $c_2$ so that $c_1 g(x) \leq f(x) \leq c_2 g(x)$. Thus by combining (4.5) and [23, Corollary 5.10, Theorem 5.11], we have the heat kernel estimates for the large time as well.

**Theorem 4.7.** For any bounded $C^{1,\alpha}$-domain $D$ and $T > 0$, there exist positive constants $c_5, c_6 > 0$ such that

$$c_5 e^{-\mu_1 T} \rho(x) \rho(y) \leq q_D(t, x, y) \leq c_6 e^{-\mu_1 T} \rho(x) \rho(y), \quad \forall (t, x, y) \in [T, \infty) \times D \times D.$$  

Theorems 1.1 and 4.7 imply that the Green function $G_D(x, y) := \int_0^\infty q_D(t, x, y) \, dt$ is finite for $x \neq y$. We now extend [22, Theorem 6.2] to $C^{1,\alpha}$-domain and for every $d \geq 1$.

**Theorem 4.8.** The following Gaussian function estimate holds with implicit constants depending on $\mu$ only via the rate at which $\max_{1 \leq i \leq d} N_{\mu_i}(h)$ goes to zero:

$$G_D(x, y) \asymp \begin{cases} 
\frac{1}{|x-y|^{d-2}} (1 \wedge \frac{\rho(x) \rho(y)}{|x-y|^2}) & \text{when } d \geq 3, \\
\log(1 + \frac{\rho(x) \rho(y)}{|x-y|^2}) & \text{when } d = 2, \\
(\rho(x) \rho(y))^{1/2} \wedge \frac{\rho(x) \rho(y)}{|x-y|} & \text{when } d = 1.
\end{cases}$$

**Proof.** We skip the proof of the theorem since it is just an elementary direct integration (cf. [22, Theorem 6.2] and [5, Corollary 1.2]). □

The next corollary is an easy consequence of Theorem 4.8. See [9].

**Corollary 4.9 (3G-Theorem).** For any bounded $C^{1,\alpha}$-domain $D$ in $\mathbb{R}^d$, there exists a constant $c > 0$ depending on $\mu$ only via the rate at which $\max_{1 \leq i \leq d} N_{\mu_i}(r)$ goes to zero such that for $x, y \in D$,

$$\frac{G_D(x, y) G_D(y, z)}{G_D(x, z)} \leq c \begin{cases} 
\frac{|x-z|^{d-2}}{|x-y|^{d-2} |y-z|^{d-2}} & \text{when } d \geq 3, \\
(G_D(x, y) + G_D(y, z) + 1) & \text{when } d = 2, \\
1 & \text{when } d = 1.
\end{cases}$$
Acknowledgments

The main results of this paper were presented by the third named author at the First Joint Meeting of KMS and AMS held in Seoul, Republic of Korea from December 16 to 20, 2009, at the Sixth Cornell Probability Summer School held in Ithaca, USA from July 19 to 30, 2010 and at the Fourth International Conference on Stochastic Analysis and Its Applications held in Osaka, Japan from August 30 to September 3, 2010.

We thank Zhen-Qing Chen and Renming Song for their helpful comments. We thank Michael Röckner for informing us about the PhD thesis of Peng Jin [21] which discusses similar problems on $\mathbb{R}^d$. We also thank the referee for his (her) valuable comments.

References

[33] Q.S. Zhang, Gaussian bounds for the fundamental solutions of $\nabla (A\nabla u) + B\nabla u - u_t = 0$, Manuscripta Math. 93 (1997) 381–390.