A Appendix: Main ideas of proofs

The main idea of the proof will be outlined and the relevant theorems will be presented. Further detailed definitions will be provided later.

Ω is a probability space and $b(t,\mathbf{x}), \sigma_B(t,\mathbf{x}), \sigma_L(t,\mathbf{x})$ are a scalar function from $\mathbb{R}^d$ to $\mathbb{R}$ a under some smooth condition. If a $\mathbb{R}^d$-valued stochastic process $(\overrightarrow{X}_t)_{t\in[0,T]}$ is a solution of a Stochastic Differential Equations (SDE) driven by Lévy process, $d\overrightarrow{X}_t = b(t,\overrightarrow{X}_{t-})dt + \sigma_L(t,\overrightarrow{X}_{t-})dL_t^\alpha$, the generator $\mathcal{L}_t$ satisfies

$$\mathcal{L}_t u(\mathbf{x}) = b(t,\mathbf{x})\nabla u(\mathbf{x}) + \int \left[u(\mathbf{x}+\sigma_L(t,\mathbf{x})\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot\sigma_L(t,\mathbf{x})\mathbf{y}\right]\nu(d\mathbf{y}). \tag{14}$$

where ν is a symmetric Lévy measure of L_t^α. If for all $(t,\mathbf{x})$, $\sigma_L(t,x) > 0$, then

$$\mathcal{L}_t u(\mathbf{x}) = b(t,\mathbf{x})\nabla u(\mathbf{x}) + \int \left[u(\mathbf{x}+\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot\mathbf{y}\right]\frac{1}{\sigma_L^d(t,\mathbf{x})}\tilde{\nu}(d\mathbf{y}). \tag{15}$$

where $\tilde{\nu}(A) = \nu(\phi^{-1}(A))$ such that A is a borel measurable sets and ϕ is a function, $\phi(\mathbf{x}) = \sigma_L(t,\mathbf{x})\cdot\mathbf{x}$.

We know the form of generator $\mathcal{L}_t$ of the given a weak solution of the SDE. Therefore we can get the time-reversal formula of the operator $\mathcal{L}_t$ [7] such that

$$\overleftarrow{\mathcal{L}}_t u(\mathbf{x}) = \overleftarrow{b}(t,\mathbf{x})\cdot\nabla u(\mathbf{x}) + \int_{\mathbb{R}^n}\int \left[u(\mathbf{x}+\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot\mathbf{y}\right]\frac{1}{\sigma_L^d(t)}\frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})}\tilde{\nu}(d\mathbf{y}). \tag{16}$$

where $\mathbf{p}_t(\mathbf{x})$ is a marginal density function of the solution $(\overrightarrow{X}_t)_{t\in[0,T]}$ and the drift $\overleftarrow{b}(t,\mathbf{x})$ is given by

$$b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \int_{\mathbb{R}^n}\mathbf{y}\cdot\left(1+\frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})}\right)\frac{1}{\sigma_L^d(t)}\tilde{\nu}(d\mathbf{y}) \quad \mathbf{p_t} - \text{a.e.} \tag{17}$$

Time-reversal of SDEs driven by Lévy process takes the form of Lévy-type stochastic integral. This means that the equation (16) can be seen as the generator for a solution of some Lévy-type stochastic integral. It is uncertain whether the SDE exists in its precise form and whether $\overleftarrow{b}(t,\mathbf{x})$ can be expressed in a simple manner. To answer these questions, the proof is divided into two parts. The first part is to determine the SDE representation of a generator $\mathcal{L}_t$ of the form (16), and the second part is to find the exact form of $\overleftarrow{b}(t,\mathbf{x})$.

B Time-reversal of SDE

In this chapter, we present proof that, under certain conditions, the time-reversal formula can be transformed into an exact formula based on the generator of a general Markov process with a jump kernel. First, we will briefly review some essential lemmas. Lemma B.1 states that there always exists a homogeneous Markov process that corresponds to an inhomogeneous Markov process. Lemma B.1 explains that there exists an SDE representation of a homogeneous Markov process with a particular generator. Lemma B.5 introduces the general time-reversal formula. By transforming time-inhomogeneous Markov processes and finding the SDE representation for a specific generator, we can determine the SDE representation for the generator of the reverse-time process. From these lemmas, we can derive the reverse-time SDE and obtain stochastic samplings.

B.1 Time-Reversal of General Markov process with jump kernel

Let $\overrightarrow{X}_t$ be an $\mathbb{R}^d$-valued continuous time inhomogeneous Markov process on an probability space $(\Omega,\mathcal{A},\mathbb{P})$ where Ω is a set, $\mathcal{A}$ is a σ-algebra, and $\mathbb{P}$ is a probability measure. The evolution system is defined as

$$\mathcal{T}(s,t)u(\mathbf{x}) = \mathbb{E}(u(\overrightarrow{X}_t)|\overrightarrow{X}_s = \mathbf{x}) \text{ for } s \le t, s,t \in [0,1]. \tag{18}$$

and this operator is well-defined on the set of Borel measurable function u on $\mathbb{R}^d$, denoted by $B(\mathbb{R}^d)$. The operator is linear and positive preserving with $\mathcal{T}(s,t)\mathbf{I} = \mathbf{I}$ and $\mathcal{T}(s,t) = \mathcal{T}(s,r)\mathcal{T}(r,t)$ for $s \leq r \leq t$ where $\mathbf{I}$ is a identity operator. This operator is also strongly continuous such that for each $v, w \in \mathbb{R}, v \leq w$ and $s \leq t$ $\lim_{(s,t)\to(v,w)} ||U(s,t)u - U(v,w)u||_\infty = 0$ where $||\cdot||_\infty$ is the supreme norm. For all $u \in C_\infty(\mathbb{R}^d)$, the set of a continuous function with vanishing at ∞, the generators of the evolution system is given by

$$\mathcal{L}_s u = \lim_{h\to 0} \frac{\mathcal{T}(s,s+h)u - u}{h} \text{ for each } s \in \mathbb{R}. \tag{19}$$

A family of linear operators $\mathcal{T}(s,t)$ on C_∞ is a Feller evolution system if it is a strongly continuous, positive, contraction semigroup on C_∞.

Definition B.1 (Space-time process)**.** Let $\mathcal{B}$ be a Borel algebra in $\mathbb{R}^d$ and an a state space $(\mathbb{R}_+ \times \mathbb{R}^d, \tilde{\mathcal{B}})$ with $\tilde{\mathbf{x}} \in \mathbb{R}_+ \times \mathbb{R}^d$ and σ-algebra $\tilde{\mathcal{B}} = \{B \in \mathbb{R}_+ \times \mathbb{R}^d | B_s \in \mathcal{B}\}$ where the cuts $B_s := \{x : (s,x) \in B\}$ are elements of the Borel σ-algebra on $\mathbb{R}^d$, and a new sample space $(\tilde{\Omega}, \tilde{\mathcal{A}})$ with $\tilde{w} = (s,w) \in \mathbb{R}_+ \times \Omega = \tilde{\Omega}$ and $\tilde{\mathcal{A}} = \{A \subset \mathbb{R}_+ \times \Omega | A_s \in \mathcal{A}, \forall s \in \mathbb{R}_+\}$. A space-time process $(\tilde{X}_t)$ is defined by

$$\tilde{X}_t(\tilde{w}) = (s+t, \overrightarrow{X}_{s+t}(w)). \tag{20}$$

with the probability measure for $A \in \tilde{\mathcal{A}}$ and $\tilde{\mathbf{x}} \in \mathbb{R}_+ \times \mathbb{R}^d$ such that $\tilde{P}_{\tilde{\mathbf{x}}}(A) = \tilde{P}(A|\tilde{X}_0 = (s,\mathbf{x})) \doteq P(A_s|\overrightarrow{X}_s = \mathbf{x})$ and the transition probabilities are given by $\tilde{P}(\tilde{X}_t \in B|\tilde{X}_0 = \tilde{\mathbf{x}}) = \tilde{P}(\tilde{X}_t \in B|\tilde{X}_0 = (s,\mathbf{x})) = P(\overrightarrow{X}_{s+t} \in B_{s+t}|\overrightarrow{X}_s = \mathbf{x})$ where $B \in \tilde{B}, \tilde{\mathbf{x}} \in \mathbb{R}_+ \times \mathbb{R}^d$. The transition function is defined by $\tilde{P}(t, \tilde{\mathbf{x}}, B) = P(s, \mathbf{x}; s+t, B_{s+t})$.

Lemma B.1. *Given a inhomogeneous Markov process (X_t), the space-time process $(\tilde{X}_t)$ on $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ is a homogeneous Markov process.*

Proof. See Transformation 3.1 in [4]. □

Lemma B.2. *Let $(\overrightarrow{X}_t)$ be the stochastic process with a Feller evolution system $U(s,t)$ and the generator of $(\overrightarrow{X}_t)$ be $\mathcal{L}_t$. Let $\widetilde{X}_t$ be its space-time process with associated the semigroup $\mathcal{T}(t)$ by $\mathcal{T}_t u(\tilde{\mathbf{x}}) = \tilde{\mathbb{E}}(u(\tilde{X}_t)|\tilde{X}_0 = \tilde{\mathbf{x}})$ for $\tilde{\mathbf{x}} \in \mathbb{R}_+ \times \mathbb{R}^d$ and $u \in B_b(\mathbb{R}_+ \times \mathbb{R}^d)$, the set of bounded Borel measurable functions. Then the extended generator $\tilde{\mathcal{L}}$ of $\mathcal{T}_t$ is given for all $u \in C_\infty\left([0,1] \times \mathbb{R}^d\right)$ satisfying some conditions,*

$$\tilde{\mathcal{L}}u(\widetilde{\mathbf{x}}) = \frac{\partial}{\partial s}u(s,\mathbf{x}) + \mathcal{L}_s u_s(\mathbf{x}) \quad \text{where } \widetilde{\mathbf{x}} = (s,\mathbf{x}) \text{ and } u_s(\mathbf{x}) = u(s,\mathbf{x}). \tag{21}$$

Proof. See Theorem 3.2 in [4]. □

A Markov process typically has a generator that takes the form

$$\begin{aligned} \mathcal{L}u(\mathbf{x}) &= \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i \partial x_j}u(\mathbf{x}) + b(\mathbf{x}) \cdot \nabla u(\mathbf{x}) \\ &+ \int_{\mathbb{R}^d} \left(u(\mathbf{x}+\mathbf{y}) - u(\mathbf{x}) - \mathbf{1}_{B_1}(\mathbf{y})\mathbf{y} \cdot \nabla u(\mathbf{x})\right) \eta(\mathbf{x}, d\mathbf{y}). \end{aligned} \tag{22}$$

where $b(\mathbf{x})$ is a locally bounded $\mathbb{R}^d$-valued function and (a_{ij}) is a locally bounded and $d \times m$ matrix-valued function, B_1 is the ball with a radius of one and a center of zero and η satisfies

$$\int_{\mathbb{R}^d} 1 \wedge \left|y^2\right| \eta(\mathbf{x}, d\mathbf{y}) < \infty. \tag{23}$$

Suppose there exist $\lambda : \mathbb{R}^d \times S \to [0,1], \hat{\gamma} : \mathbb{R}^d \times S \to \mathbb{R}^d$, and a σ-finite measure v on a measurable space $(S, \mathcal{S})$, where S is a set satisfying $\mathcal{S} \subset \text{dom}(\lambda)$ and $\mathcal{S}$ is a σ-algebra defined on S such that

$$\eta(\mathbf{x}, \Gamma) = \int_S \lambda(\mathbf{x}, \mathbf{y})\mathbf{1}_\Gamma(\hat{\gamma}(\mathbf{x}, \mathbf{y}))\nu(d\mathbf{y}). \tag{24}$$

We decompose S into $S_1 \cup S_2$ such that $1_{S_1} = 1_{B_1}(\hat{\gamma}((s,\mathbf{x}),\mathbf{y}))$ and $1_{S_2} = 1_{B_1^c}(\hat{\gamma}((s,\mathbf{x}),\mathbf{y}))$. We can rewrite the form of the generator is

$$\begin{aligned}\mathcal{L}u(\mathbf{x}) &= \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\frac{\partial^2}{\partial x_i \partial x_j}u(\mathbf{x}) + b(\mathbf{x})\cdot\nabla u(\mathbf{x}) \\ &\quad + \int_S \lambda(\mathbf{x},\mathbf{y})u(\mathbf{x},\hat{\gamma}(\mathbf{x},\mathbf{y})) - u(\mathbf{x}) - 1_{S_1}(\mathbf{y})\hat{\gamma}(\mathbf{x},\mathbf{y})\cdot\nabla u(\mathbf{x}))\nu(d\mathbf{y}).\end{aligned}$$

Lemma B.3. *Let the generator $\mathcal{L}$ be the form of (22). Let ξ be a Poisson random measure on $[0,1]\times S\times[0,\infty)$ with mean measure $m\times\nu\times m$. We define ξ as*

$$\widetilde{\xi}(A) = \xi(A) - m\times\nu\times m(A). \tag{25}$$

and $(S,\mathcal{S})$ be a measurable space, μ be a σ-finite measure on $(S,\mathcal{S})$. Assume that for each compact $K\subset\mathbb{R}^d$,

$$\sup_{\mathbf{x}\in K}\Big(|b(\mathbf{x}) + \int_{S_1}\lambda(\mathbf{x},\mathbf{u})|\hat{\gamma}(\mathbf{x},\mathbf{u})|^2\nu(d\mathbf{u})\int_{S_2}\lambda(\mathbf{x},\mathbf{u})|\hat{\gamma}(\mathbf{x},\mathbf{u})|\wedge 1\nu(d\mathbf{u})\Big) < \infty. \tag{26}$$

Then $\overrightarrow{X}_t$ satisfies a stochastic differential equation of the form

$$\begin{aligned}\overrightarrow{X}_t &= \overrightarrow{X}_0 + \int_0^t b(\overrightarrow{X}_{s-})ds \\ &\quad + \int_{s=0}^{s=t}\int_{S_1}\int_{v=0}^{v=\lambda(\overrightarrow{X}_{s-},\mathbf{u})}\hat{\gamma}(\overrightarrow{X}_{s-},\mathbf{u})\widetilde{\xi}(dv\times d\mathbf{u}\times ds) \\ &\quad + \int_{s=0}^{s=t}\int_{S_2}\int_{v=0}^{v=\lambda(\overrightarrow{X}_{s-},\mathbf{u})}\hat{\gamma}(\overrightarrow{X}_{s-},\mathbf{u})\xi(dv\times d\mathbf{u}\times ds).\end{aligned} \tag{27}$$

Proof. See Theorem 2.3 in [19] □

Lemma B.4. *Let $\lambda((s,\mathbf{x}),\mathbf{y}) = \frac{p_s(\mathbf{x}+\mathbf{y})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)$ for $\sigma_L(s)\geq 0$ and $\hat{\gamma}((s,\mathbf{x}),\mathbf{y})$ be $(0,\mathbf{y})$ and $\nu(d\mathbf{y})$ be a Lévy measure such that it is a Borel measure on $\mathbb{R}^d$ and $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(||\mathbf{x}||^2\wedge 1)\nu(d\mathbf{x}) < \infty$ with $S_1 = B_1(\mathbf{y})$. If $(\overrightarrow{X}_t)$ has the corresponding generator $\mathcal{L}_t$*

$$\mathcal{L}_t u(\mathbf{x}) = b(\mathbf{x})\cdot\nabla u(\mathbf{x}) + \int_{\mathbb{R}^d}[u(\mathbf{x}+\mathbf{y}) - u(\mathbf{x}) - \mathbf{y}\cdot\nabla u(\mathbf{x})1_{S_1}(\mathbf{y})]\frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})}\sigma_L^\alpha(t)\nu(d\mathbf{y}). \tag{28}$$

where $u\in B_b(\mathbb{R}^d)$. Then the corresponding generator $\tilde{\mathcal{L}}$ of the space-time process $\tilde{X}_t$ is

$$\begin{aligned}\tilde{\mathcal{L}}u(s,\mathbf{x}) &= (1,b(\mathbf{x}))\cdot\nabla u(s,\mathbf{x}) + \int_{\mathbb{R}^d}[u((s,\mathbf{x}) \\ &\quad + \hat{\gamma}((s,\mathbf{x}),\mathbf{y})) - u(s,\mathbf{x}) - \gamma((s,\mathbf{x}),\mathbf{y})\cdot\nabla u(s,\mathbf{x})1_{S_1}(\mathbf{y})]\lambda((s,\mathbf{x}),\mathbf{y})\nu(d\mathbf{y}).\end{aligned} \tag{29}$$

where $u\in C_\infty([0,1]\times\mathbb{R}^d)$.

Proof.

$$\tilde{\mathcal{L}}u(s,\mathbf{x}) = \frac{\partial}{\partial s}u(s,\mathbf{x}) + \mathcal{L}_s u_s(\mathbf{x}) \quad \text{for } u_s(\mathbf{x}) = u(s,\mathbf{x}) \tag{30}$$

$$\begin{aligned} &= \frac{\partial}{\partial s}u(s,\mathbf{x}) + b(\mathbf{x})\cdot\nabla u_s(\mathbf{x}) \\ &\quad + \int [u_s(\mathbf{x}+\mathbf{y}) - u_s(\mathbf{x}) - \mathbf{y}\cdot\nabla u_s(\mathbf{x})1_{S_1}(\mathbf{y})]\,\frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})}\sigma_L^\alpha(t)\nu(d\mathbf{y}) \end{aligned} \tag{31}$$

$$\begin{aligned} &= (1, b(\mathbf{x}))\cdot\nabla u(s,\mathbf{x}) \\ &\quad + \int [u(s,\mathbf{x}+\mathbf{y}) - u(s,\mathbf{x}) - (0,\mathbf{y})\cdot\nabla u(\mathbf{x})1_{S_1}(\mathbf{y})]\,\frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})}\sigma_L^\alpha(t)\nu(d\mathbf{y}) \end{aligned} \tag{32}$$

$$\begin{aligned} &= (1, b(\mathbf{x}))\cdot\nabla u(s,\mathbf{x}) \\ &\quad + \int [u((s,\mathbf{x})+(0,\mathbf{y})) - u(s,\mathbf{x}) - (0,\mathbf{y})\cdot\nabla u(\mathbf{x})1_{S_1}(y)]\,\frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})}\sigma_L^\alpha(t)\nu(d\mathbf{y}). \end{aligned} \tag{33}$$

□

Theorem B.1. *A generator $\mathcal{L}_t$ has a jump kernel driven by the isotropic α-stable Lévy process with Lévy measure ν represented by (28). ξ be a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d \times [0,\infty)$ with mean measure $m\times\nu\times m$ such that $\mathbb{E}[\xi(dv\times d\mathbf{y}\times ds)] = dv\times\nu(d\mathbf{y})\times ds$ and $\tilde{\xi}(A) = \xi(A) - m\times\nu\times m(A)$. Then the SDE representation of the generator $\tilde{\mathcal{L}}$ satisfies*

$$\begin{aligned} \overrightarrow{X}_t &= \overrightarrow{X}_0 + \int_0^t b(s,\overrightarrow{X}_{s-})ds + \int_{s=0}^{s=t}\int_{||\mathbf{y}||<1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)} \mathbf{y}\cdot\tilde{\xi}(dv\times d\mathbf{y}\times ds) \\ &\quad + \int_{s=0}^{s=t}\int_{||\mathbf{y}||\geq 1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)} \mathbf{y}\cdot\xi(dv\times d\mathbf{y}\times ds) \\ &= \overrightarrow{X}_0 + \int_0^t b(s,\overrightarrow{X}_{s-})ds + Y_t. \end{aligned} \tag{34}$$

where

$$Y_t = \int_{s=0}^{s=t}\int_{||\mathbf{y}||<1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)} \mathbf{y}\cdot\tilde{\xi}(dv\times d\mathbf{y}\times ds) \tag{35}$$

$$+ \int_{s=0}^{s=t}\int_{||\mathbf{y}|>1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)} \mathbf{y}\cdot\xi(dv\times d\mathbf{y}\times ds) \tag{36}$$

such that the characterstic function of Y_t follows:

$$\exp\Big(\int_0^t \sigma_L^\alpha(s)\Big[\int_{\mathbb{R}^d}\Big(e^{i\langle\mathbf{u},\mathbf{y}\rangle} - 1 - i\langle\mathbf{u},\mathbf{y}\rangle\cdot 1_{||\mathbf{y}||<1}(\mathbf{y})\Big)\frac{p_s(\mathbf{y}+\mathbf{x})}{p_s(\mathbf{x})}d\nu(\mathbf{y})\Big]ds\Big). \tag{37}$$

We also decompose Y_t in such a way that

$$Y_t = \int_0^t \sigma_L(s)dL_s^\alpha + Z_t \tag{38}$$

where the Levy symbol for Z_t is expressed as

$$\int_0^t \sigma_L^\alpha(s)\Big[\int_0^1\int_{\mathbb{R}^d}\Big(e^{i\langle\mathbf{u},\mathbf{y}\rangle} - 1 - i\langle\mathbf{u},\mathbf{y}\rangle\cdot 1_{||\mathbf{y}||<1}(\mathbf{y})\Big)\frac{\langle\mathbf{y},\nabla p_s(\mathbf{x}+u\mathbf{y})\rangle}{p_s(\mathbf{x})}\nu(d\mathbf{y})du\Big]ds. \tag{39}$$

Furthermore, dZ_t is characterized by the intensity measure $\tilde{\nu}(\mathbf{x},d\mathbf{y})$ following

$$\tilde{\nu}(\mathbf{x},d\mathbf{y}) = \frac{1}{p_s(\mathbf{x})}\Big\langle\mathbf{y},\int_0^1\nabla p_s(\mathbf{x}+u\cdot\mathbf{y})du\Big\rangle\nu(d\mathbf{y}). \tag{40}$$

This ensures that $\int_{||\mathbf{y}||<1}||\mathbf{y}||\tilde{\nu}(\mathbf{x},d\mathbf{y}) < \infty$, providing the guarantee of finite variation for Z_t in accordance with Remark 7.12 in [28].

Proof. $\lambda((s,\mathbf{x}),\mathbf{y})$ is $\frac{p_s(\mathbf{x}+\mathbf{y})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)$ for $\sigma_L(s)\geq 0$ and $\hat{\gamma}((s,\mathbf{x}),\mathbf{y})$ is $(0,\mathbf{y})$ with $S_1=\{||\mathbf{y}||<1\}$ and $S_2=\{||\mathbf{y}||\geq 1\}$. We know λ satisfies

$$\int_{\mathbb{R}}\lambda((s,\mathbf{x}),\mathbf{y})1_{S_1}(\mathbf{y})||r((s,\mathbf{x}),\mathbf{y})||^2+1_{S_2}(\mathbf{y})\nu(d\mathbf{y}) \tag{41}$$

$$=\int_{||\mathbf{y}||<1}\left[\frac{p_s(\mathbf{x}+\mathbf{y})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)||\mathbf{y}||^2\nu(d\mathbf{y})\right]dy+\int_{||\mathbf{y}||\geq 1}\frac{p_s(\mathbf{x}+\mathbf{y})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)\nu(d\mathbf{y})<\infty. \tag{42}$$

Since $\int_{S_1}\lambda((s,\mathbf{x}),\mathbf{y})||\hat{\gamma}((s,\mathbf{x}),\mathbf{y})||^2\nu(d\mathbf{u})+\int_{S_2}\int\lambda((s,\mathbf{x}),\mathbf{y})|\hat{\gamma}((s,\mathbf{x}),\mathbf{y})|\wedge 1\nu(d\mathbf{u}))$ is well-defined and continuous with respect to $(s,\mathbf{x})$ and $b(s,\mathbf{x})$ is locally bounded $\mathbb{R}$-valued function, we can apply Lemma $B.2$ to the transformed homogeneous generator $\tilde{\mathcal{L}}$ of the inhomogeneous generator $\mathcal{L}_t$ from Lemma $B.4$. Now, we find the corresponding stochastic process Y_t by using Lemma B.3 such that

$$\begin{aligned}Y_t=&\int_{s=0}^{s=t}\int_{||\mathbf{y}||<1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)}\mathbf{y}\cdot\tilde{\xi}(dv\times d\mathbf{y}\times ds)\\&+\int_{s=0}^{s=t}\int_{||\mathbf{y}||>1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)}\mathbf{y}\cdot\xi(dv\times d\mathbf{y}\times ds).\end{aligned} \tag{43}$$

We know that

$$\begin{aligned}&\mathbb{E}\left[\exp\langle i\langle\mathbf{u},Y_t\rangle\right]=\mathbb{E}\Big[\exp\langle i\mathbf{u},\int_{s=0}^{s=t}\int_{||\mathbf{y}||<1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(s)}\mathbf{y}\cdot\tilde{\xi}(dv\times d\mathbf{y}\times ds)\rangle\\&+\int_{s=0}^{s=t}\int_{||\mathbf{y}||>1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\overrightarrow{X}_{s-})}{p_s(\overrightarrow{X}_{s-})}\sigma_L^\alpha(t)}\mathbf{y}\cdot\xi(dv\times d\mathbf{y}\times ds)\rangle\Big].\end{aligned} \tag{44}$$

Thus,

$$\mathbb{E}[\exp(i(\mathbf{u},Y_t))]=\mathbb{E}\Big[\exp(i\langle\mathbf{u},\int_{s=0}^{s=t}\int_{||\mathbf{y}||<1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\mathbf{x})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)}\mathbf{y}\cdot\tilde{\xi}(dv\times d\mathbf{y}\times ds) \tag{45}$$

$$+\int_{s=0}^{s=t}\int_{||\mathbf{y}||\geq 1}\int_{v=0}^{v=\frac{p_s(\mathbf{y}+\mathbf{x})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)}\mathbf{y}\cdot\xi(dv\times d\mathbf{y}\times ds)\rangle)\Big] \tag{46}$$

$$=\exp\Big(\int_0^t\int_{\mathbb{R}^d}\int_0^{\frac{p_s(\mathbf{y}+\mathbf{x})}{p_s(\mathbf{x})}\sigma_L^\alpha(s)}(e^{i\langle\mathbf{u},\mathbf{y}\rangle}-1-i\langle\mathbf{u},\mathbf{y}\rangle\cdot 1_{||\mathbf{y}||<1}(\mathbf{y}))dv\times\nu(d\mathbf{y})\times ds)\Big) \tag{47}$$

$$=\exp\Big(\int_0^t\sigma_L^\alpha(s)\left[\int_{\mathbb{R}^d}\left(e^{i\langle\mathbf{u},\mathbf{y}\rangle}-1-i\langle\mathbf{u},\mathbf{y}\rangle\cdot 1_{||\mathbf{y}||<1}(\mathbf{y})\right)\frac{p_s(\mathbf{y}+\mathbf{x})}{p_s(\mathbf{x})}\nu(d\mathbf{y})\right]ds\Big). \tag{48}$$

Given that s is within the interval $[0,1]$ and $\mathbf{x},\mathbf{y}$ belong to $\mathbb{R}^d$, we can obtain an useful a representation for $\frac{p_s(\mathbf{x}+\mathbf{y})}{p_s(\mathbf{x})}$ by applying Fundamental theorem of calculus [12],

$$\frac{p_s(\mathbf{x}+\mathbf{y})}{p_s(\mathbf{x})}=1+\int_0^1\frac{\langle\mathbf{y},\nabla p_s(\mathbf{x}+u\mathbf{y})\rangle}{p_s(\mathbf{x})}du. \tag{49}$$

Hence, we can break down the characteristic function of Y_t in a manner such that

$$\begin{aligned}&\mathbb{E}\left[\exp\left(\langle i\langle\mathbf{u},Y_t\rangle\right)\right]=\exp\Big(\int_0^t\sigma_L^\alpha(s)\left[\int_{\mathbb{R}^d}\left(e^{i\langle\mathbf{u},\mathbf{y}\rangle}-1-i\langle\mathbf{u},\mathbf{y}\rangle\cdot 1_{||\mathbf{y}||<1}(\mathbf{y})\right)\nu(d\mathbf{y})\right]ds\Big)+\\&\exp\Big(\int_0^t\sigma_L^\alpha(s)\left[\int_{\mathbb{R}^d}\left(e^{i\langle\mathbf{u},\mathbf{y}\rangle}-1-i\langle\mathbf{u},\mathbf{y}\rangle\cdot 1_{||\mathbf{y}||<1}(\mathbf{y})\right)\int_0^1\frac{\langle\mathbf{y},\nabla p_s(\mathbf{x}+u\mathbf{y})\rangle}{p_s(\mathbf{x})}\nu(d\mathbf{y})du\right]ds\Big)\end{aligned} \tag{50}$$

$$=\mathbb{E}\left[\exp\left(i\langle\mathbf{u},\int_0^t\sigma_L(s)dL_s^\alpha\right)\right]\cdot\mathbb{E}[\exp\left(i\langle\mathbf{u},Z_t\right)] \tag{51}$$

in which the Levy symbol for Z_t is expressed as

$$\int_0^t \sigma_L^\alpha(s)\left[\int_0^1 \int_{\mathbb{R}^d}\left(e^{i\langle \mathbf{u},\mathbf{y}\rangle} - 1 - i\langle \mathbf{u},\mathbf{y}\rangle \cdot 1_{||\mathbf{y}||<1}(\mathbf{y})\right)\frac{\langle \mathbf{y}, \nabla p_s(\mathbf{x}+u\mathbf{y})\rangle}{p_s(\mathbf{x})}\nu(d\mathbf{y})du\right]ds. \quad (52)$$

so that $Y_t = \int_0^t \sigma_L(s)dL_s^\alpha + Z_t$. □

Up until now, we have established that inhomogeneous Markov processes which fulfill specific conditions have SDE representations. Later, we will look into how the time-reversal of the generator appears in the case of a homogeneous Markov process. The Lemma B.1 will be utilized to derive the time-reversal formula for SDE driven by Lévy process. We will then use the time-reversal formula from Theorem 5.7 in [7] to introduce a new type of generative model called LIM.

Lemma B.5. *Consider a Markov process $(\overrightarrow{X}_t)$ with a generator $\mathcal{L}_t$ that is defined on the set of continuous functions with compact support, $C_c^1(\mathbb{R}^d)$ such that $\mathcal{L}_t u(x) = b(t,x)\cdot\nabla u(\mathbf{x}) + \int_{\mathbb{R}^n}[u(\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot[\mathbf{y}-\mathbf{x}]^\delta]\overrightarrow{J}_{t,\mathbf{x}}(d\mathbf{y})$, $(t,\mathbf{x}) \in [0,T]\times\mathbb{R}^n$ for some $\delta > 0$, where $b(t,\mathbf{x})$ is a vector field, and the jump kernel is $\overrightarrow{J}_{t,\mathbf{x}}(d\mathbf{y})$. Let $[\mathbf{x}]^\delta \doteq 1_{||\mathbf{x}||\le\delta}x$. Under certain conditions, the generator of the reverse-time process, $\overleftarrow{\mathcal{L}}_t$, is given by.*

$$\overleftarrow{\mathcal{L}}_t u(\mathbf{x}) = \overleftarrow{b}(t,\mathbf{x})\cdot\nabla u(\mathbf{x}) + \int_{\mathbb{R}^n}\int[u(\mathbf{y}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot[\mathbf{y}-\mathbf{x}]^\delta\overleftarrow{J}_{t,\mathbf{x}}(d\mathbf{y}). \quad (53)$$

where $\mathbf{p}_t(d\mathbf{y})$ is a marginal distribution of $(\overrightarrow{X}_t)$ such that it satisfies $\mathbf{p}_t(d\mathbf{y})\overleftarrow{J}_{t,\mathbf{x}}(d\mathbf{x}) = \mathbf{p}_t(d\mathbf{x})\overrightarrow{J}_{t,\mathbf{x}}(d\mathbf{y})$ for almost every t and the backward drift $\overleftarrow{b}(t,\mathbf{x})$ is given by

$$b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \int_{\mathbb{R}^n}[\mathbf{y}-\mathbf{x}]^\delta(\overrightarrow{J}_{t,\mathbf{x}} + \overleftarrow{J}_{t,\mathbf{x}})(d\mathbf{y}) \quad \mathbf{p_t} - (a.e). \quad (54)$$

Proof. See Theorem 5.7 in [7]. □

Assuming the marginal distribution has a density function $p_t(\mathbf{x})$ such that $\mathbf{p}_t(d\mathbf{x}) = p_t(\mathbf{x})d\mathbf{x}$ and $\overrightarrow{J}_{t,\mathbf{x}}(d\mathbf{y})$ is a symmetric kernel with $\overrightarrow{J}_{t,\mathbf{x}}(d\mathbf{y}) = v_t(\mathbf{y}-\mathbf{x})d\mathbf{y}$ for some isotropic Lévy measure v_t that is a Borel measure such that $v_t(\{0\}) = 0$ and $\int 1\wedge||\mathbf{y}||^2 v_t(d\mathbf{y}) < \infty$ for each $t\in[0,T]$. Then $\overleftarrow{J}_{t,\mathbf{x}}(d\mathbf{y}) = \frac{p_t(\mathbf{y})}{p_t(\mathbf{x})}v_t(\mathbf{y}-\mathbf{x})d\mathbf{y}$. It satisfies $b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \int_{||y||\le\delta}\mathbf{y}\cdot\frac{p_t(\mathbf{y}+\mathbf{x})}{p_t(\mathbf{x})}v_t(\mathbf{y})d\mathbf{y}$. Since v_t is symmetric, δ can be ∞ such that

$$\begin{aligned}\overleftarrow{\mathcal{L}}_t u(\mathbf{x}) &= \overleftarrow{b}(t,\mathbf{x})\cdot\nabla u(\mathbf{x}) + \int_{\mathbb{R}^n}\left[u(\mathbf{y}+\mathbf{x}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot[\mathbf{y}]^\delta\right]\nu_t(d\mathbf{y}) \\ &= \overleftarrow{b}(t,\mathbf{x})\cdot\nabla u(\mathbf{x}) + \int_{\mathbb{R}^n}\left[u(\mathbf{y}+\mathbf{x}) - u(\mathbf{x}) - \nabla u(\mathbf{x})\cdot\mathbf{y}\right]\nu_t(d\mathbf{y}).\end{aligned}$$

Thus,

$$b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \int_{\mathbb{R}^n}\mathbf{y}\cdot\frac{p_t(\mathbf{y}+\mathbf{x})}{p_t(\mathbf{x})}\nu_t(\mathbf{y})d\mathbf{y} \quad \mathbf{p_t} - \text{(a.e)}. \quad (55)$$

B.2 Fractional Calculus

Fractional calculus is a concept that extends differentiation. To begin with, we will define the Fractional Laplacian, which is utilized to depict the drift term $\overleftarrow{b}(t,\mathbf{x})$ of the time-reverse formula for SDEs driven by isotropic α-stable Lévy process.

Definition B.2 (Fractional Laplacian). Let the fourier transformation of f be $\mathcal{F}\{f\}(\mathbf{u}) = \int_{\mathbf{x}\in\mathbb{R}^d} e^{i\langle\mathbf{x},\mathbf{u}\rangle}f(\mathbf{x})d\mathbf{x}$. The fractional Laplacian $\Delta^{\frac{\alpha}{2}}$ for $\alpha > -1$ follows,

$$\Delta^{\frac{\alpha}{2}}f(\mathbf{x}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}||\mathbf{u}||^\alpha\mathcal{F}\{f\}(\mathbf{u})e^{-i\langle\mathbf{u},\mathbf{x}\rangle}d\mathbf{u}. \quad (56)$$

The fractional Laplacian is a linear operator that is a more general version of the original Laplacian, represented as $\Delta^{\frac{\alpha}{2}}$. The minus sign is omitted in the fractional Laplacian for ease of use, as stated in [27].

B.3 1-dimensional isotropic alpha-stable Lévy processs

We will focus on the case of one-dimensional isotropic α-stable Lévy process. This type of process has a isotropic Lévy measure ν that follows $\nu(dy) = \frac{C}{|y|^{1+\alpha}}dy$, where $C = \frac{\Gamma(\alpha+1)\sin(\alpha\pi/2)}{\pi}$. Using equation (55), we can estimate the drift term $\overleftarrow{b}(t, x)$.

Lemma B.6. *Given a $\mathbb{R}$-valued stochastic process $(\overrightarrow{X}_t)$ that solves the equation $d\overrightarrow{X}_t = -\frac{\beta(t)}{\alpha}\overrightarrow{X}_t + (\beta(t))^{1/\alpha}dL_t^\alpha$, the jump kernel of $\overrightarrow{X}_t$ can be represented as follows.*

$$\overrightarrow{J}_t(x, dy) = \frac{\Gamma(\alpha+1)\sin(\alpha\pi/2)}{\pi}\frac{\sigma_L^\alpha(t)dy}{|y-x|^{\alpha+1}}. \tag{57}$$

Proof. See Lemma 4.6 in [32]. □

By Lemma B.6, the drift term of the $\mathbb{R}$-valued solution $(\overrightarrow{X}_t)$ to $d\overrightarrow{X}_t = -\frac{\beta(t)}{\alpha}\overrightarrow{X}_{t-} + (\beta(t))^{1/\alpha}dL_t^\alpha$ satisfies

$$b(t, x) + \overleftarrow{b}(t, x) = \frac{\Gamma(\alpha+1)\sin(\alpha\pi/2)}{\pi}\sigma_L^\alpha(t)\int_{\mathbb{R}^n} y \cdot \frac{p_t(x+y)}{p_t(y)}\frac{1}{|y|^{1+\alpha}}dy \quad \mathbf{p_t} - \text{(a.e)}. \tag{58}$$

Therefore, the Markov generator $\overleftarrow{\mathcal{L}}_t$ of $(\overleftarrow{X}_t)$ is the form of (28). So, we can use Theorem B.1 to $\overleftarrow{\mathcal{L}}_t$ such that the reverse-time SDE of $\overleftarrow{X}_t$ is $d\overleftarrow{X}_t = -\overleftarrow{b}(t, \overleftarrow{X}_{t+})d\bar{t} + \sigma_L^\alpha(t)d\bar{L}_t^\alpha + d\bar{Z}_t$ where $\bar{d}t$ is an infinitesimal negative timestep, $\bar{L}_t^\alpha$ is the backward version of the isotropic α-stable Lévy process, and $\bar{Z}_t$ is the backward version of Z_t where Z_t is a Lévy-type stochastic integral [3] such that $\mathbb{E}[Z_t] = 0$ with finite variation. See more detail in Theorem B.1 for the definition of Z_t.

We shall now figure out the exact form of the integral representation of $\overleftarrow{b}(t, x)$. We arrive at an useful equation for it.

Lemma B.7. $\int_0^\infty \frac{\sin x}{x^\alpha}dx = \cos(\frac{\pi\alpha}{2}) \cdot \Gamma(1-\alpha)$.

Lemma B.8. *For $1 < \alpha < 2$, $\int_{-\infty}^\infty \frac{y}{|y|^{\alpha+1}}e^{-i(u,y)}dy = -2 \cdot iu|u|^{\alpha-2}\cos(\frac{\pi\alpha}{2})\Gamma(1-\alpha)$.*

Proof. Let $uy = k$. If $u > 0$,

$$\int_{-\infty}^\infty \frac{y}{|y|^{\alpha+1}}e^{-i(u,y)}dy = |u|^{\alpha-1}\int_{-\infty}^\infty \frac{k}{|k|^{\alpha+1}}e^{ik}dk.$$

If $u < 0$,

$$\int_{-\infty}^\infty \frac{y}{|y|^{\alpha+1}}e^{-i(u,y)}dy = -|u|^{\alpha-1}\int_{-\infty}^\infty \frac{k}{|k|^{\alpha+1}}e^{-ik}dk.$$

Therefore,

$$\begin{aligned}&\int \frac{y}{|y|^{\alpha+1}}e^{-i(u,y)}dy = -\text{sgn}(u)|u|^{\alpha-1}\int_{-\infty}^\infty \frac{k}{|k|^{\alpha+1}}e^{ik}dk\\ &= -2iu|u|^{\alpha-2}\int_0^\infty \frac{\sin k}{k^\alpha}dk = -2 \cdot iu|u|^{\alpha-2}\cos(\frac{\pi\alpha}{2})\Gamma(1-\alpha).\end{aligned}$$

□

Theorem B.2. *If $d\overrightarrow{X}_t = b(t, \overrightarrow{X}_{t-})dt + \sigma_L(t)dL_t^\alpha$ then the reverse-time SDE with respect to backward integral is $d\overleftarrow{X}_t = -\overleftarrow{b}(t, \overleftarrow{X}_{t+})\bar{d}t + \sigma(t)d\bar{L}_t^\alpha + d\bar{Z}_t$ with $\overleftarrow{b}(t, x)$ satisfying*

$$b(t, x) + \overleftarrow{b}(t, x) = \sigma_L^\alpha(t) \cdot \alpha \cdot \frac{\Delta^{\frac{\alpha-2}{2}}\nabla_x p_t(x)}{p_t(x)}. \tag{59}$$

Proof.

$$\begin{aligned}
\Delta^{\frac{\alpha-2}{2}}\nabla p_t(x) &= -\frac{1}{(2\pi)^d}\int iu|u|^{(\alpha-2)}e^{-i(u,x)}\hat{p}_t(u)du \\
&= \frac{1}{(2\pi)^d}\cdot\frac{1}{2\cdot\cos(\pi\alpha/2)\Gamma(1-\alpha)}\int\int\frac{y}{|y|^{\alpha+1}}e^{-i(u,y+x)}\hat{p}_t(u)dudy \\
&= \frac{1}{2\cdot\cos(\pi\alpha/2)\Gamma(1-\alpha)}\int p_t(x+y)\frac{y}{|y|^{\alpha+1}}dy \\
&= \frac{1}{\alpha}\int C\cdot p_t(x+y)\frac{y}{|y|^{\alpha+1}}dy \text{ for } C = \frac{\sin(\pi\alpha/2)\Gamma(\alpha+1)}{\pi}.
\end{aligned}$$

since $\Gamma(1-\alpha)\Gamma(\alpha) = \frac{\pi}{\sin\pi\alpha}$ and $\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha$. Thus, $b(t,x)+\overleftarrow{b}(t,x) = \sigma_L^\alpha(t)\cdot\alpha\cdot\frac{\Delta^{\frac{\alpha-2}{2}}\nabla_x p_t(x)}{p_t(x)}$. □

For one dimension, we have so far identified the exact form of the time-reversal formula. Let we derive the exact form for the d-dimensional isotropic α-stable Lévy process based on this.

B.4 d-dimensional isotropic alpha-stable Lévy process

Lemma B.9. *Let the constant C satisfies*

$$\frac{1}{C} = 2\int_0^\infty \frac{1-\cos k}{k^{\alpha+1}}dk\cdot\int_{|\theta_{d-1}|<\frac{\pi}{2}}(\cos\theta_{d-1})^\alpha d\sigma(\theta_1,\ldots,\theta_{d-1}) \tag{60}$$

$$= \frac{\pi}{\Gamma(\alpha+1)\sin(\alpha\pi/2)}\cdot\int_{\theta_{d-1}<\frac{\pi}{2}}(\cos\theta_{d-1})^\alpha d\sigma(\theta_1,\ldots,\theta_{d-1}). \tag{61}$$

For $1<\alpha<2$, the integral $\int_{\mathbf{u}\in\mathbb{R}^d}\frac{e^{-i\langle\mathbf{u},\mathbf{y}\rangle}-1}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}$ follows

$$\int_{\mathbf{u}\in\mathbb{R}^d}\frac{e^{-i\langle\mathbf{u},\mathbf{y}\rangle}-1}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y} = -\frac{1}{C}||\mathbf{u}||^\alpha. \tag{62}$$

Proof. Let $I = \int_{\mathbf{u}\in\mathbb{R}^d}\frac{e^{i\langle\mathbf{u},\mathbf{y}\rangle}-1}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}$. The integral I converges since $\frac{1}{||\mathbf{y}||^{d+\alpha}}$ is a Lévy measure. For given $\mathbf{u}$, we fix an axis which is parallel to the direction of $\mathbf{u}$ and take $\hat{\mathbf{u}} = \frac{\mathbf{u}}{||\mathbf{u}||}$. As the dimension of $\mathbb{R}^d$ is d, we can find an orthogonal basis $\tilde{B}$ such that

$$\tilde{B} = \{\tilde{e}_1,\ldots,\tilde{e}_{d-1},\hat{\mathbf{u}}\}. \tag{63}$$

The standard basis B is denoted by $B = \{\hat{e}_1,\ldots,\hat{e}_d\}$. For any $\mathbf{y}\in\mathbb{R}^d$, we can represent $\mathbf{y}$ as $\mathbf{y} = \sum_{j=1}^d y_j\hat{e}_j = \sum_{i=1}^d \tilde{u}_i\tilde{e}_i$. Since $\tilde{B}$ is orthogonal, the measure $d\mathbf{y}$ follows the equation,

$$d\mathbf{y} = dy_1\cdots dy_n = d\tilde{\mathbf{u}}_1\cdots d\tilde{\mathbf{u}}_d. \tag{64}$$

From the observation, we apply a spherical coordinate to the basis $\tilde{B}$ following,

$$\tilde{u}_1 = r\sin\theta_1\prod_{i=2}^{d-1}\sin\theta_i \tag{65}$$

and

$$\tilde{u}_m = r\cos\theta_{m-1}\prod_{k=m}^{d-1}\sin\theta_k \tag{66}$$

for $m\in\{2,\ldots,d-1\}$, and $\tilde{u}_d = r\cos\theta_{d-1}$. Since I converges in the sense of improper integral, we can use the change of variable for the integral with the spherical measure $d\sigma$ such that

$$d\sigma = d\sigma(\theta_1\ldots,\theta_{d-1}) = \prod_{k=2}^{d-1}\sin^{k-1}\theta_k d\theta_1,\ldots d\theta_{d-1}. \tag{67}$$

From the coordinate transformation, we obtain a polar-coordinate-based representation of the above integral.

$$\int_{\mathbf{u}\in\mathbb{R}^d}\frac{e^{-i\langle\mathbf{u},\mathbf{y}\rangle}-1}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}=\int_{\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{e^{-i||\mathbf{u}||r\cos\theta_{d-1}}-1}{|r|^{\alpha+1}}drd\sigma \tag{68}$$

$$\begin{aligned}&=\int_{\sigma\in\mathcal{S}^{d-1},0<\theta_{d-1}<\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{e^{-i||\mathbf{u}||r\cos\theta_{d-1}}-1}{|r|^{\alpha+1}}drd\sigma\\&+\int_{\sigma\in\mathcal{S}^{d-1},\pi>\theta_{d-1}>\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{e^{-i||\mathbf{u}||r\cos\theta_{d-1}}-1}{|r|^{\alpha+1}}drd\sigma\end{aligned} \tag{69}$$

$$\begin{aligned}&=\int_{\sigma\in\mathcal{S}^{d-1},0<\theta_{d-1}<\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{e^{-i||\mathbf{u}||r\cos\theta_{d-1}}-1}{|r|^{\alpha+1}}drd\sigma\\&+\int_{\sigma\in\mathcal{S}^{d-1},0<\theta_{d-1}<\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{e^{i||\mathbf{u}||r\cos\theta_{d-1}}-1}{|r|^{\alpha+1}}drd\sigma\end{aligned} \tag{70}$$

$$=2\int_{\sigma\in\mathcal{S}^{d-1},0<\theta_{d-1}<\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{\cos(||\mathbf{u}||r\cos\theta_{d-1})-1}{|r|^{\alpha+1}}drd\sigma. \tag{71}$$

If $k=||\mathbf{u}||r\cos\theta_{d-1}$, $r=\frac{k}{||\mathbf{u}||\cos\theta_{d=1}}$, $dr=\frac{dk}{||\mathbf{u}||\cos\theta_{d-1}}$,

$$\int_{\sigma\in\mathcal{S}^{d-1},0<\theta_{d-1}<\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{\cos(||\mathbf{u}||r\cos\theta_{d-1})-1}{|r|^{\alpha+1}}drd\sigma \tag{72}$$

$$=\int_{\sigma\in\mathcal{S}^{d-1},0<\theta_{d-1}<\frac{\pi}{2}}\int_{r=0}^{\infty}\frac{\cos(k)-1}{|k|^{\alpha+1}}||\mathbf{u}||^{\alpha}(\cos\theta_{d-1})^{\alpha}dkd\sigma \tag{73}$$

$$=\Big(-2\int_0^{\infty}\frac{1-\cos k}{k^{\alpha+1}}\int_{\sigma\in\mathcal{S}^{d-1},|\theta_{d-1}|<\frac{\pi}{2}}(\cos\theta_{d-1})^{\alpha}d\sigma\Big)||\mathbf{u}||^{\alpha} \tag{74}$$

$$=\frac{\pi}{\Gamma(\alpha+1)\sin(\alpha\pi/2)}\cdot\int_{\sigma\in\mathcal{S}^{d-1},\theta_{d-1}<\frac{\pi}{2}}(\cos\theta_{d-1})^{\alpha}d\sigma(\theta_1,\ldots,\theta_{d-1}) \tag{75}$$

$$\doteq\frac{1}{C} \tag{76}$$

Thus, if a Lévy process X_t has the Lévy measure $\nu(d\mathbf{y})=\frac{C}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}$, then the characteristic function $\psi(\mathbf{u})=\mathbb{E}[e^{i\langle\mathbf{u},\mathbf{x}\rangle}]$ follows $\psi(\mathbf{u})=e^{-||\mathbf{u}||^{\alpha}}$. □

Lemma B.10. *Let $1<\alpha<2$. Then the integral $\int_{\mathbb{R}^d}\frac{C\cdot e^{-i\langle\mathbf{u},\mathbf{y}\rangle}\mathbf{y}}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}$ can be represented as:*

$$\int_{\mathbb{R}^d}\frac{C\cdot e^{-i\langle\mathbf{u},\mathbf{y}\rangle}\mathbf{y}}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}=-\alpha i\mathbf{u}||\mathbf{u}||^{\alpha-2}. \tag{77}$$

Proof. Because α is greater than 1 and less than 2, the mean is finite, so $\int_{u\in\mathbb{R}^d}\frac{C\cdot e^{-i\langle\mathbf{u},\mathbf{y}\rangle}\mathbf{y}}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}$ is defined in the sense of an improper integral, and convergence is guaranteed. Let $I=\int_{\mathbf{u}\in\mathbb{R}^d}\frac{C\cdot e^{-i\langle\mathbf{u},\mathbf{y}\rangle}\mathbf{y}}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y}$. Similar to B.9, transforming to spherical coordinates and expressing I is as follows.

$$\int_{\mathbf{u}\in\mathbb{R}^d} \frac{C\cdot e^{-i\langle \mathbf{u},\mathbf{y}\rangle}\mathbf{y}}{||\mathbf{y}||^{d+\alpha}}d\mathbf{y} \tag{78}$$

$$= \int_{\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty} \frac{Ce^{-i||\mathbf{u}||r\cos\theta}}{r^\alpha}(\sin\theta_1\prod_{k=2}^{d-1}\sin\theta_k^{d-1}\tilde{u}_1\tilde{e}_1 + \sum_{m=2}^{d-2}\cos\theta_{m-1}\prod_{k=m}^{d-1}\sin\theta_k\tilde{u}_m\tilde{e}_m + \cos\theta_{d-1}\hat{\mathbf{u}})drd\sigma \tag{79}$$

$$= I_1\tilde{e}_1 + \cdots + I_m\tilde{e}_m + \cdots + I_d\hat{\mathbf{u}} = I_d\hat{\mathbf{u}}. \tag{80}$$

$$\tag{81}$$

where

$$I_1 = \int_{\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{Ce^{-i||\mathbf{u}||r\cos\theta_{d-1}}}{r^\alpha}(\sin\theta_1\prod_{k=2}^{d-1}\sin\theta_k^{d-1}\tilde{u}_1)drd\sigma = 0 \tag{82}$$

since $\int_{0<\theta_1<2\pi}\sin\theta_1 d\theta_1 = 0$. For $m\in\{2,\ldots,d-1\}$,

$$I_m = \int_{\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{Ce^{-i||\mathbf{u}||r\cos\theta_{d-1}}}{r^\alpha}(\cos\theta_{m-1}\prod_{k=m}^{d-1}\sin\theta_k\tilde{u}_m)drd\sigma = 0. \tag{83}$$

since

$$\int_{0<\theta_{m-1}<2\pi}\cos\theta_{m-1}\sin^{m-2}\theta_{m-1}\sin\theta_{m-1}d\theta_{m-1} = 0 \quad \text{for } m=2 \tag{84}$$

and

$$\int_{0<\theta_{m-1}<\pi}\cos\theta_{m-1}\sin^{m-2}\theta_{m-1}\sin\theta_{m-1}d\theta_{m-1} = 0 \quad \text{for } m\in\{3,\ldots,d-1\} \tag{85}$$

and

$$I_d = \int_{0<\theta_{d-1}<\frac{\pi}{2},\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{Ce^{-i||\mathbf{u}||r\cos\theta}}{r^\alpha}(\cos\theta_{d-1})drd\sigma \tag{86}$$

$$= \int_{0<\theta_{d-1}<\frac{\pi}{2},\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{Ce^{-i||\mathbf{u}||r\cos\theta}}{r^\alpha}(\cos\theta_{d-1})drd\sigma \tag{87}$$

$$+ \int_{\frac{\pi}{2}<\theta_{d-1}<\pi,\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{Ce^{-i||\mathbf{u}||r\cos\theta_{d-1}}}{r^\alpha}(\cos\theta_{d-1})drd\sigma \tag{88}$$

$$= -2i\int_{0<\theta_{d-1}<\frac{\pi}{2},\sigma\in\mathcal{S}^{d-1}}\int_{r=0}^{\infty}\frac{C\sin(||\mathbf{u}||r\cos\theta_{d-1})}{r^\alpha}(\cos\theta_{d-1})drd\sigma. \tag{89}$$

If $k = ||\mathbf{u}||r\cos\theta_{d-1}$ then $r = \frac{k}{||\mathbf{u}||\cos\theta_{d-1}}$, $dr = \frac{dk}{||\mathbf{u}||\cos\theta_{d-1}}$ and we can get the equation:

$$I_d\hat{\mathbf{u}} = \Big(-2i \int_{0<\theta_{d-1}<\frac{\pi}{2}, \sigma\in\mathcal{S}^{d-1}} \int_{r=0}^{\infty} \frac{C\sin(||\mathbf{u}||r\cos\theta_{d-1})}{r^{\alpha}} (\cos\theta_{d-1}) dr d\sigma \Big)\hat{\mathbf{u}} \tag{90}$$

$$= \Big(-2i \int_{0<\theta_{d-1}<\frac{\pi}{2}, \sigma\in\mathcal{S}^{d-1}} \int_{k=0}^{\infty} \frac{C\sin(k)}{k^{\alpha}} (||\mathbf{u}||^{\alpha-1} (\cos\theta_{d-1})^{\alpha} \cos\theta_{d-1}) dk d\sigma \Big)\hat{\mathbf{u}} \tag{91}$$

$$= \Big(-2i \int_{k=0}^{\infty} \frac{\sin(k)}{k^{\alpha}} dk \int_{0<\theta_{d-1}<\frac{\pi}{2}, \sigma\in\mathcal{S}^{d-1}} ((\cos\theta_{d-1})^{\alpha} \cos\theta_{d-1}) d\sigma \cdot C \cdot ||\mathbf{u}||^{\alpha-1} \Big)\hat{\mathbf{u}} \tag{92}$$

$$\begin{aligned} &= -2i\cos(\frac{\pi\alpha}{2})\Gamma(1-\alpha) \\ &\quad \cdot \frac{\Gamma(\alpha+1)\sin(\frac{\pi\alpha}{2})}{\pi} \frac{\int_{0<\theta_{d-1}<\frac{\pi}{2}, \sigma\in\mathcal{S}^{d-1}} ((\cos\theta_{d-1})^{\alpha} \cos\theta_{d-1}) d\sigma}{\int_{0<\theta_{d-1}<\frac{\pi}{2}, \sigma\in\mathcal{S}^{d-1}} ((\cos\theta_{d-1})^{\alpha} \cos\theta_{d-1}) d\sigma} ||\mathbf{u}||^{\alpha-1}\hat{\mathbf{u}} \end{aligned} \tag{93}$$

$$= -2i \frac{\cos(\frac{\pi\alpha}{2})\sin(\frac{\pi\alpha}{2})}{\pi} \Gamma(\alpha+1)\Gamma(1-\alpha)\hat{\mathbf{u}} \tag{94}$$

$$= -i \frac{\sin(\alpha\pi)}{\pi} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \frac{\pi}{\sin(\alpha\pi)} \hat{\mathbf{u}} \tag{95}$$

$$= -i\alpha||\mathbf{u}||^{\alpha-1}\hat{\mathbf{u}} \tag{96}$$

$$= -i\alpha\mathbf{u}||\mathbf{u}||^{\alpha-2}. \tag{97}$$

□

Theorem B.3. $b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \frac{\int_{\mathbf{u}\in\mathbb{R}^d} -i\mathbf{u}||\mathbf{u}||^{\alpha-2} e^{-i\alpha\langle\mathbf{u},\mathbf{x}\rangle} \hat{p}_t(\mathbf{u}) d\mathbf{u}}{p_t(\mathbf{x})}$ *for* $\sigma_L = 1$*. If* $\sigma_L \neq 1$ *then*

$$b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \sigma_L^{\alpha}(t) \cdot \frac{\int_{\mathbf{u}\in\mathbb{R}^d} -i\mathbf{u}||\mathbf{u}||^{\alpha-2} e^{-i\alpha\langle\mathbf{u},\mathbf{x}\rangle} \hat{p}_t(\mathbf{u}) d\mathbf{u}}{p_t(\mathbf{x})}$$

Proof. When $\sigma_L(t) = 1$,

$$b(t,\mathbf{x}) + \overleftarrow{b}(t,\mathbf{x}) = \int_{\mathbf{y}\in\mathbb{R}^d} \frac{p_t(\mathbf{x}+\mathbf{y})}{p_t(\mathbf{x})} \frac{C}{||\mathbf{y}||^{d+\alpha}} \mathbf{y} d\mathbf{y} \tag{98}$$

$$= \frac{\int_{\mathbf{y}\in\mathbb{R}^d} \int_{\mathbf{u}\in\mathbb{R}^d} \hat{p}_t(\mathbf{u}) e^{-i\langle\mathbf{u},\mathbf{x}+\mathbf{y}\rangle} \frac{C}{||\mathbf{y}||^{d+\alpha}} \mathbf{y} d\mathbf{u} d\mathbf{y}}{p_t(\mathbf{x})} \tag{99}$$

$$= \frac{\int_{\mathbf{u}\in\mathbb{R}^d} \hat{p}_t(\mathbf{u}) e^{-i\langle\mathbf{u},\mathbf{x}\rangle} [\int_{\mathbf{y}\in\mathbb{R}^d} e^{-i\langle\mathbf{u},\mathbf{y}\rangle} \frac{C}{||\mathbf{y}||^{d+\alpha}} \mathbf{y} d\mathbf{y} d\mathbf{u}}{p_t(\mathbf{x})} \tag{100}$$

$$= \frac{\int_{\mathbf{u}\in\mathbb{R}^d} -i\alpha\mathbf{u}||\mathbf{u}||^{\alpha-2} e^{-i\langle\mathbf{u},\mathbf{x}\rangle} \hat{p}_t(\mathbf{u})] d\mathbf{u}}{p_t(\mathbf{x})} \tag{101}$$

$$= \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})}. \tag{102}$$

□

Corollary B.1. *If* $d\overrightarrow{X}_t = b(t,\overrightarrow{X}_t)dt + \sigma_L(t)dL_t^{\alpha}$ *is given, then the time-reversal of SDE follows*

$$d\overleftarrow{X}_t = \Big(b(t,\overleftarrow{X}_{t+}) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\overleftarrow{X}_{t+})}{p_t(\overleftarrow{X}_{t+})} \Big) \bar{d}t + \sigma_L(t)\bar{d}L_t^{\alpha} + \bar{d}Z_t. \tag{103}$$

where $\bar{d}t$ *is an infinitesimal negative timestep,* $\bar{L}_t^{\alpha}$ *is the backward version of the isotropic* α*-stable Lévy process, and* $\bar{Z}_t$ *is the backward version of* Z_t *where* Z_t *is a Lévy-type stochastic integral [3] such that* $\mathbb{E}[Z_t] = 0$ *with finite variation.*

B.5 Combined models

The forward carré du champ is a process that is defined as $\overrightarrow{\Gamma}_t(u,v) = \mathcal{L}_t(uv) - u\mathcal{L}_t v - v\mathcal{L}_t u$, where $dom\overrightarrow{\Gamma}_t = (u,v); u, v, uv \in$ dom $\mathcal{L}_t$. The IbP of the Carré du champ is as follows: if $u \in$ dom$\overleftarrow{L}$ and $\overleftarrow{L}u \in L^1(q)$, then for almose every t

$$\int_{\mathbb{R}^n} \Big\{(\mathcal{L}_t u + \overleftarrow{\mathcal{L}}_t u)v + \overleftarrow{\Gamma}_t(u,v)\Big\} dq_t = 0. \tag{104}$$

By (104), the proof of the time-reversal formula is based on the integration by parts formula for the carré du champ. As a result, the reverse-formula is dependent on the form of the Carré du champ.

If the forward generator $\mathcal{L}_t$ can be decomposed into $\mathcal{L}_t = \mathcal{L}_t^1 + \mathcal{L}_t^2$, then its Carré de champ also can be decomposed into $\overrightarrow{\Gamma}_t(u,v) = \overrightarrow{\Gamma}^1_t(u,v) + \overrightarrow{\Gamma}^2_t(u,v)$ such that $\overrightarrow{\Gamma}^1_t(u,v)$ is the Carré du champ of $\mathcal{L}_t^1$ and $\overrightarrow{\Gamma}^2_t(u,v)$ is the Carré du champ of $\mathcal{L}_t^2$ [9]. Since Carreé du champ $\overrightarrow{\Gamma}_t$ is only determined by operator $\mathcal{L}_t$, and if it satisfies $\overrightarrow{\Gamma}_t(u,v) = \overrightarrow{\Gamma}^1_t(u,v) + \overrightarrow{\Gamma}^2_t(u,v)$ then

$$\int_{\mathbb{R}^n} (\overleftarrow{\mathcal{L}}_t u)v = \int_{\mathbb{R}^n} (\mathcal{L}_t u)v + \overleftarrow{\Gamma}_t(u,v) dq_t \tag{105}$$

$$= \int_{\mathbb{R}^n} (\mathcal{L}_t u)v + \int_{\mathbb{R}^n} \overrightarrow{\Gamma}^1_t(u,v) + \int_{\mathbb{R}^n} \overrightarrow{\Gamma}^2_t(u,v) dq_t. \tag{106}$$

The reverse-formula is derived from the decomposition of $\int_{\mathbb{R}^n} \overrightarrow{\Gamma}^1_t(u,v)dq_t$ and $\int_{\mathbb{R}^n} \overrightarrow{\Gamma}^2_t(u,v)dq_t$. Knowing the terms of each integral allows one to find the time-reversal formula for $\mathcal{L}_t$. From this, we can extend the result to the time-reversal of jump-diffusion processes. The general form of the reverse SDE is given by:

$$\begin{aligned} d\overleftarrow{X}_t = &\Big(b(t, \overleftarrow{X}_{t+}) - \sigma_B^2(t)\nabla \log p_t(\overleftarrow{X}_{t+}) - \sigma_L^\alpha(t) \cdot \alpha \cdot S_t^{(\alpha)}(\overleftarrow{X}_{t+})\Big)\bar{d}t \\ &+ \sigma_B(t)d\bar{B}_t + \sigma_L(t)d\bar{L}_t^\alpha + d\bar{Z}_t. \end{aligned} \tag{107}$$

where $\bar{d}t$ is an infinitesimal negative timestep, $\bar{L}_t^\alpha$ is the backward version of the isotropic α-stable Lévy process, and $\bar{Z}_t$ is the backward version of Z_t where Z_t is a Lévy-type stochastic integral [3] such that $\mathbb{E}[Z_t] = 0$ with finite variation.

C Probability ODE

C.1 Probability ODE for isotropic alpha-stable Lévy process

In this chapter, we discuss the fractional Fokker-Planck equation, which is a extended version of the Fokker-Planck equation that considers fractional derivatives. The goal of this equation is to determine the existence of probability fractional ODEs. Before proving the existence of probability fractional ODEs, some useful lemmas will be presented.

Lemma C.1. *Let $1 < \alpha < 2$. One can divide fractional Laplacian into* $\Delta^{\frac{\alpha}{2}} f(\mathbf{x}) = \sum_{i=1}^d \Big(-\partial_{x_i}^2 \Delta^{\frac{\alpha-2}{2}} f(\mathbf{x})\Big) = \sum_{i=1}^d \Big(-\partial_{x_i}\Delta^{\frac{\alpha-2}{2}}\partial_{x_i} f(\mathbf{x})\Big)$

Proof.

$$\Delta^{\frac{\alpha}{2}} f(x) = \frac{1}{(2\pi)^d} \int ||\mathbf{u}||^{\alpha} \hat{f}(\mathbf{u}) e^{-i\langle \mathbf{u},\mathbf{x}\rangle} d\mathbf{u} \tag{108}$$

$$= \frac{1}{(2\pi)^d} \int ||\mathbf{u}||^{2} \cdot ||\mathbf{u}||^{\alpha-2} \hat{f}(\mathbf{u}) e^{-i\langle \mathbf{u},\mathbf{x}\rangle} d\mathbf{u} \tag{109}$$

$$= \frac{1}{(2\pi)^d} \int \sum_{i=1}^{d} |u_i|^2 ||\mathbf{u}||^{\alpha-2} \hat{f}(\mathbf{u}) e^{-i\langle \mathbf{u},\mathbf{x}\rangle} d\mathbf{u} \tag{110}$$

$$= \frac{1}{(2\pi)^d} \sum_{i=1}^{d} \int |u_i|^2 ||\mathbf{u}||^{\alpha-2} \hat{f}(\mathbf{u}) e^{-i\langle \mathbf{u},\mathbf{x}\rangle} d\mathbf{u} \tag{111}$$

$$= \frac{1}{(2\pi)^d} \sum_{i=1}^{d} \int u_i ||\mathbf{u}||^{\alpha-2} u_i \hat{f}(\mathbf{u}) e^{-i\langle \mathbf{u},\mathbf{x}\rangle} d\mathbf{u} \tag{112}$$

$$= \sum_{i=1}^{d} \Big(- \partial_{x_i}^2 \Delta^{\frac{\alpha-2}{2}} f(\mathbf{x}) \Big) \tag{113}$$

$$= \sum_{i=1}^{d} \Big(- \partial_{x_i} \Delta^{\frac{\alpha-2}{2}} \partial_{x_i} f(\mathbf{x}) \Big). \tag{114}$$

□

Lemma C.2 (Fractional Fokker-Planck equation for isotropic α-stable Lévy process)**.** *Given the SDE $d\overrightarrow{X}_t = b(t, \overrightarrow{X}_{t-})dt + \sigma_L(t) dL_t^{\alpha}$, where $b(t,\mathbf{x})$ and $\sigma_L(t)$ are measurable and satisfy a Lipschitz condition, then the marginal density function $p_t(\mathbf{x})$ follows the fractional Fokker-Planck equation for isotropic α-stable Lévy process as*

$$\frac{\partial p_t(\mathbf{x})}{\partial t} = -\nabla \cdot [b(t,\mathbf{x}) p_t(\mathbf{x})] - \sigma_L^{\alpha}(t) \Delta^{\frac{\alpha}{2}} p_t(\mathbf{x}). \tag{115}$$

The proof of this result can be found in [35].

Theorem C.1 (Existence of Probability ODE)**.** *If the marginal density function $p_t(\mathbf{x})$ follows the fractional Fokker-Planck equation, it can be expressed as:*

$$\frac{\partial p_t(\mathbf{x})}{\partial t} = -\nabla \cdot \Big([b(t,\mathbf{x}) p_t(\mathbf{x})] - \sigma_L^{\alpha}(t) \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \Big) p_t(\mathbf{x}) \tag{116}$$

And the SDE for $\overrightarrow{X}_t$ becomes:

$$d\overrightarrow{X}_t = \Big(b(t, \overrightarrow{X}_t) - \sigma_L^{\alpha}(t) \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\overrightarrow{X}_t)}{p_t(\overrightarrow{X}_t)} \Big) dt \tag{117}$$

Proof.

$$\frac{\partial p(\mathbf{x})}{\partial t} = -\nabla \cdot [b(t,\mathbf{x}) p_t(\mathbf{x})] - \sigma_L(t) \Delta^{\frac{\alpha}{2}} p_t(\mathbf{x}) \tag{118}$$

$$= -\sum_{i=1}^{d} \partial_{x_i} (b_i(t,\mathbf{x}) p_t(\mathbf{x})) + \sum_{i=1}^{d} \sigma_L^{\alpha}(t) \partial_{x_i} \Delta^{\frac{\alpha-2}{2}} \partial_{x_i} p_t(\mathbf{x}) \tag{119}$$

$$= -\sum_{i=1}^{d} \Big[\partial_{x_i} b_i(t,\mathbf{x}) p_t(\mathbf{x}) - \sigma_L^{\alpha}(t) \partial_{x_i} \Delta^{\frac{\alpha-2}{2}} \partial_{x_i} p_t(\mathbf{x}) \Big] \tag{120}$$

$$= -\sum_{i=1}^{n} \partial_{x_i} \left(\left[b_i(t, x_t) - \sigma_L^{\alpha}(t) \frac{\Delta^{\frac{\alpha-2}{2}} \partial_{x_i} p_t(\mathbf{x})}{p_t(\mathbf{x})} \right] p_t(\mathbf{x}) \right) \tag{121}$$

$$= -\nabla \cdot \Big[\Big((b(t,\mathbf{x}) - \sigma_L^{\alpha}(t) \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \Big) p_t(\mathbf{x}) \Big]. \tag{122}$$

□

Lemma C.3 (General fractional Fokker-Planck equation)**.** *If the Lévy-driven stochastic SDE is given as*

$$d\overrightarrow{X}_t = b(t, \overrightarrow{X}_{t-})dt + \sigma_B(t)dB_t + \sigma_L(t)dL_t^\alpha. \tag{123}$$

then the marginal density function $p_t(\mathbf{x})$ satisfies the General fractional-Fokker-Planck equation,

$$\frac{\partial p_t(\mathbf{x})}{\partial t} = -\nabla \cdot [b(t, \mathbf{x})p_t(\mathbf{x})] + \frac{\sigma_B^2(t)}{2}\nabla p_t(\mathbf{x}) - \sigma_L^\alpha(t)\Delta^{\frac{\alpha}{2}} p_t(\mathbf{x}). \tag{124}$$

Corollary C.1 (The general Probability ODE)**.** *If $p_t(\mathbf{x})$ follows the fractional Fokker-Planck equation, then the marginal density function $p_t(\mathbf{x})$ satisfies the following expression:*

$$\frac{\partial p_t(\mathbf{x})}{\partial t} = -\nabla \cdot \left[b(t, \mathbf{x}) - \frac{\sigma_B^2(t)}{2}\nabla \log p_t(\mathbf{x}) - \sigma_L^\alpha(t) \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \right]. \tag{125}$$

The process $\overrightarrow{X}_t$ also satisfies the following differential equation:

$$d\overrightarrow{X}_t \overset{d}{=} \left[b(t, \overrightarrow{X}_{t-}) - \frac{\sigma_B^2(t)}{2}\nabla \log p_t(\overrightarrow{X}_{t-}) - \sigma_L^\alpha(t) \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\overrightarrow{X}_{t-})}{p_t(\overrightarrow{X}_{t-})} \right] dt. \tag{126}$$

Proof. We apply the same method in Lemma C.3. □

D General OU process

The weak solution to the SDE, $\overrightarrow{X}_t = \mathbf{x}_0 e^{-\beta t} + (\alpha \cdot \beta)^{1/\alpha} \int_0^t e^{-\beta(t-s)} dL_s^\alpha$ in (127) can be expressed as

$$d\overrightarrow{X}_t = -\beta \overrightarrow{X}_t dt + (\alpha \cdot \beta)^{1/\alpha} dL_t^\alpha. \tag{127}$$

the solution of the SDE for $\overrightarrow{X}_0 = \mathbf{x}_0$ is

$$\overrightarrow{X}_t \overset{d}{=} \mathbf{x}_0 e^{-\beta t} + (\alpha \cdot \beta)^{1/\alpha} \int_0^t e^{-\beta(t-s)} dL_s^\alpha. \tag{128}$$

In the following section, we will investigate how to derive the solution of equation (127).

Lemma D.1. *Given α with $1 < \alpha < 2$ and f is a measurable function such that $f : [0, T] \to \mathbb{R}$ with $\int_0^T |f(s)|^\alpha ds < \infty$. Let $\mathbb{R}^d$-valued $\overrightarrow{X}_t = \int_0^t f(s)dL_s^\alpha$ then*

$$\overrightarrow{X}_t \sim \mathcal{S}\alpha\mathcal{S}\left(\int_0^t |f(s)|^\alpha ds)^{1/\alpha} \right). \tag{129}$$

Proof. If $f(t) = \sum_{i=1}^N a_i \chi_{(t_{i-1}, t_i]}$ with $t_0 = 0, t_N = t$,

$$\overrightarrow{X}_t = \int_0^t \sum_{i=1}^N a_i 1_{(t_{i-1}, t_i]}(s) dL_s^\alpha = \sum_{i=1}^n a_i [L_{t_i}^\alpha - L_{t_{i-1}}^\alpha] \overset{d}{=} \sum_{i=1}^N a_i L_{\Delta t_i}^\alpha, \ \Delta t_i = t_i - t_{i-1}. \tag{130}$$

Using the above equation,

$$\mathbb{E}\left[e^{i\langle \mathbf{u}, \overrightarrow{X}_t \rangle} \right] = \mathbb{E}\left[e^{i\langle \mathbf{u}, \sum_{i=1}^N a_i L_{\Delta t_i}^\alpha \rangle} \right] = \prod_{i=1}^N \mathbb{E}\left[e^{i\langle \mathbf{u}, a_i L_{\Delta i_i}^\alpha \rangle} \right] \tag{131}$$

$$= \prod_{i=1}^N e^{-||\mathbf{u}||^\alpha |a_i|^\alpha \Delta t_i} = e^{-\sum_{i=1}^N |a_i|^\alpha \Delta t_i ||\mathbf{u}||^\alpha} = e^{-(\int_0^t |f(s)|^\alpha ds)||\mathbf{u}||^\alpha}. \tag{132}$$

Thus, we conclude that $\overrightarrow{X}_t \sim \mathcal{S}\alpha\mathcal{S}\left(\int_0^t |f(s)|^\alpha ds)^{1/\alpha} \right)$.

The proof for the lemma is demonstrated for a case where the function f is not a simple function.

For simplicity, it is assumed that $f(t)$ is non-negative, but if it is not, it can be decomposed into two non-negative functions $f^+(t)$ and $f^-(t)$. Then, a sequence of simple functions f_n that approaches $f(t)$ can be constructed. The vector-valued process X_t^n is defined as the integral of $f_n(s)dL_s^\alpha$ from 0 to t. As the integral of $|f(s)|^\alpha$ is finite, the dominated convergence theorem [12] can be applied to show that the limit of X_t^n as n approaches infinity is equal to X_t for all values of t in the interval $[0, T]$.

$$\mathbb{E}[e^{i\langle \mathbf{u}, \overrightarrow{X}_t \rangle}] = \lim_{n\to\infty} \mathbb{E}[e^{i\langle \mathbf{u}, \overrightarrow{X}_t^n \rangle}] = \lim_{n\to\infty} e^{-(\int_0^t |f_n(s)|^\alpha ds)||\mathbf{u}||^\alpha} = e^{-(\int_0^t |f(s)|^\alpha ds)||\mathbf{u}||^\alpha}. \tag{133}$$

Thus, we obtain $\overrightarrow{X}_t \sim \mathcal{S}\alpha\mathcal{S}\left(\int_0^t |f(s)|^\alpha ds)^{1/\alpha}\right)$ when f is a measurable function. □

Theorem D.1. *If $a(t)$ is equal to $e^{-\beta t}$, $\gamma(t)$ is equal to $(1-e^{-\alpha\beta t})^{1/\alpha} = (1-(a(t))^\alpha)^{1/\alpha}$, and $\overrightarrow{X}_t$ is equal to $a(t)\mathbf{x}_0 + \gamma(t)\boldsymbol{\epsilon}$ for some random variable $\boldsymbol{\epsilon}$ with the α-stable distribution $\mathcal{S}\alpha\mathcal{S}$, then $\overrightarrow{X}_t$ is a weak solution to the stochastic differential equation (SDE), $d\overrightarrow{X}_t = -\beta\overrightarrow{X}_t dt + (\alpha\cdot\beta)^{1/\alpha} dL_t^\alpha$. Furthermore, any weak solution of this SDE can be represented as $\overrightarrow{X}_t \overset{d}{=} \mathbf{x}_0 e^{-\beta t} + (\alpha\cdot\beta)^{1/\alpha}\int_0^t e^{-\beta(t-s)} dL_s^\alpha$.*

Proof. Use Lemma D.1. □

Lemma D.2. *If $\overrightarrow{X}_t$ is a weak solution to the SDE $d\overrightarrow{X}_t = -\frac{\beta(t)}{\alpha}\overrightarrow{X}_t dt + \beta(t)^{1/\alpha} dL_t^\alpha$, then it can be represented by*

$$\overrightarrow{X}_t \overset{d}{=} e^{-\int_0^t \frac{\beta(s)}{\alpha} ds}\overrightarrow{X}_0 + \int_0^t e^{-\int_u^t \frac{\beta(s)}{\alpha} ds}\beta(u)^{1/\alpha} dL_u^\alpha. \tag{134}$$

The function $a(t) = e^{-\int_0^t \frac{\beta(s)}{\alpha} ds}$ is defined and it is stated that the scale parameter $\gamma(t)$ of $\int_0^t e^{-\int_u^t \frac{\beta(s)}{\alpha} ds}(\beta(u)^{1/\alpha} dL_t^\alpha$ satisfies $\gamma^\alpha(t) = (1-a^\alpha(t))$.

Proof.

$$\begin{aligned} d(e^{\int_0^t \frac{\beta(s)}{\alpha} ds}) &= e^{\int_0^t \frac{\beta(s)}{\alpha} ds}\cdot\frac{\beta(t)}{\alpha} dt + e^{\int_0^t \frac{\beta(s)}{\alpha} ds}\Big(-\frac{\beta(t)}{\alpha}\overrightarrow{X}_t dt + (\beta(t))^{1/\alpha} dL_t^\alpha\Big) \\ &= e^{\int_0^t \frac{\beta(s)}{\alpha} ds}(\beta(t))^{1/\alpha} dL_t^\alpha. \end{aligned}$$

$\overrightarrow{X}_t = e^{-\int_0^t \frac{\beta(s)}{\alpha} ds}X_0 + \int_0^t e^{-\int_u^t \frac{\beta(s)}{\alpha} ds}\beta(u)^{1/\alpha} dL_t^\alpha$. If we set $a(t) = e^{-\int_0^t \frac{\beta(s)}{\alpha} ds}$ then $\frac{d}{dt}\log a(t) = -\frac{\beta(s)}{\alpha}$. And the scale parameter $\gamma(t)$ satisfies

$$\begin{aligned} \gamma^\alpha(t) &= \int_0^t \frac{a(t)^\alpha}{a(u)^\alpha}(\beta(u))du = \int_0^t \frac{a^\alpha(t)}{a^\alpha(u)}(-\alpha)\frac{d}{dt}\log a(u)du = a^\alpha(t)\int_0^t \frac{-\alpha}{a^\alpha(u)}\frac{a'(u)}{a(u)}du. \\ &= a^\alpha(t)\int_0^t (-\alpha)\frac{a'(u)}{a^{\alpha+1}(u)}du = a^\alpha(t)\int_0^t \frac{d}{du}(a^{-\alpha}(u))du = a^\alpha(t)[a^{-\alpha}(t) - a^{-\alpha}(0)]. \\ &= (1-a^\alpha(t)). \end{aligned}$$

□

E Numerical methods and Convergences

The practical way to solve SDEs is to use a numerical method, and the Euler-Maruyama method is a popular choice for this. The approximation of the solution obtained from the Euler-Maruyama method denoted as $(\tilde{X}_t)_{t\in[0,T]}$, and the actual solution of the SDE, $(X_t)_{t\in[0,T]}$, both have their own measures, referred to as μ_η and μ respectively. For the Euler-Maruyama method to converge, it's important that these two measures are similarly distributed. To assess the accuracy of the approximation, it's necessary to measure the difference between the measures, which can be done using either the Wasserstein-1 distance $\mathcal{W}_1$ or the bounded Lipschitz distance $\mathcal{W}_{bL}$.

Theorem E.1. *A function $b(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is considered twice continuously differentiable for $\mathbf{x}$ and $\sigma_L(t) : \mathbb{R}^d \to \mathbb{R}$ is a bounded function. Given a SDE $d\overrightarrow{X}_t = b(t, \overrightarrow{X}_t)dt + \sigma_L(t)dL_t^\alpha$, there exists constants $\theta_1, \theta_2 > 0$ and $\theta_3, K \geq 0$ such that $\langle b(t, \mathbf{x}) - b(t, \mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \leq -\theta_1 ||\mathbf{x} - \mathbf{y}||^2 + K$ for $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ and $||\nabla b(t, \mathbf{x}) \cdot \nu|| \leq \theta_2 ||\nu||,\ ||\nu_1 \cdot \nabla^2 \nu_2|| \leq \theta_3 ||\nu_1|| ||\nu_2||$ for $\forall \nu_1, \nu_2 \in \mathbb{R}^d$. Then, there exists a constant C such that for every step size $\eta <$ min$\{1, \theta_1/(\theta_2^2), 1/\theta_1\}$, the Wasserstein-1 distance between two measures μ and μ_η satisfies $\mathcal{W}_1(\mu, \mu_\theta) \leq C\eta^{\frac{2}{\alpha}-1}$.*

Proof. See more detail in Theorem 1.2 in [6]. □

Theorem E.2. *Let a function $b(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ be twice continuously differentiable for $\mathbf{x} \in \mathbb{R}^d$ and satisfies **Assumption A** [6] for uniformly in t, and $\sigma_L(t) : \mathbb{R}^d \to \mathbb{R}$ are bounded. Given a SDE $d\overrightarrow{X}_t = b(t, \overrightarrow{X}_t)dt + \sigma_L(t)dL_t^\alpha$, there exists a constant C such that for small step size $\eta << 1$, the Wasserstein-1 distance between two measures μ and μ_η satisfies $\mathcal{W}_1(\mu, \mu_\theta) \leq C\eta^{\frac{2}{\alpha}-1}$.*

Proof. We can deduce Theorem [6] by following exactly the proof of Theorem 1.2 in [6]. If we use the quadratic schedule to the stochastic sampling of LIM, the bound of the Wasserstein-1 distance between the invariant measure of the solution and of an approximation satisfies $\mathcal{W}_1(\mu, \mu_{\frac{1}{N^2}}) \leq C\eta^{\frac{3}{\alpha}-1}$ where $\mu_{\frac{1}{N^2}}$ is the invariant measure of the approximation following the Euler-Maruyama scheme.

Let $\widetilde{Z}_1, \widetilde{Z}_2, \cdots$ be an iid sequence of d-dimensional random vectors, which are Pareto distributed, i.e.

$$\widetilde{Z}_1 \sim p(\mathbf{z}) = \frac{\alpha}{\sigma_{d-1} ||\mathbf{z}||^{\alpha+d}} \chi_{(1,\infty)}(||\mathbf{z}||)$$

We denote by $\sigma_{d-1} = 2\pi^{\frac{d}{2}} / \Gamma\left(\frac{d}{2}\right)$ the surface area of the unit sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$. We will approximate the SDE (1.1) by the following Euler-Maruyama scheme:

$$Y_0 = x, \quad Y_{k+1} = Y_k + \eta_{k+1} b\left(\eta_1 + \cdots + \eta_k, Y_k\right) + \frac{\eta_{k+1}^{1/\alpha}}{\sigma} \sigma_L(\eta_1 + \ldots + \eta_k) \widetilde{Z}_{k+1}, \quad k = 0, 1, 2, \ldots \tag{135}$$

where $\sigma^\alpha = \alpha / \left(\sigma_{d-1} C_{d,\alpha}\right)$. We denote the initial point $X_0 = \mathbf{x}$ for a given $\mathbf{x} \in \mathbb{R}^d$. we use this also for Y_k^y for a given $\mathbf{y} \in \mathbb{R}^d$. By P_t, Q_k and Q_k we denote the Markov semigroups of X_t, Y_k and Y_k, respectively, i.e.

$$P_t f(\mathbf{x}) = \mathbb{E}[f\left(X_t^{\mathbf{x}}\right)], \quad Q_{\eta_1 + \cdots + \eta_k} f(\mathbf{x}) = \mathbb{E}[f\left(Y_k^{\mathbf{x}}\right)]. \tag{136}$$

for a bounded measurable function $f : \mathbb{R}^d \to \mathbb{R}, \mathbf{x} \in \mathbb{R}^d, t \geq 0$ and $k = 0, 1, 2, \ldots$ For $i \in \{1, \ldots, N\}$, i-th step size η_i follows $\eta_i = \frac{2(N-i)+1}{N^2}$ for the quadratic schedule. The key idea for deriving the bounds of the Wasserstein-1 distance is to use the Duhamel principle:

$$P_{\eta_1 + \cdots + \eta_N} h(\mathbf{x}) - Q_{\eta_1 + \cdots + \eta_N} h(\mathbf{x}) = \sum_{i=1}^{N} Q_{\eta_1 + \cdots + \eta_{i-1}} \left(P_{\eta_{N-i+1}} - Q_{\eta_i}\right) P_{\eta_1 + \cdots + \eta_{(N-i)}} h(\mathbf{x}). \tag{137}$$

Through the Duhamel principle,

$$\mathcal{W}_1\left(\text{law}\left(X_{\eta_1 + \cdots + \eta_N}\right), \text{law}\left(Y_{\eta_1 + \cdots + \eta_N}\right)\right) = \sup_{h \in \text{Lip}(1)} \left|P_{\eta_1 + \cdots + \eta_N} h(x) - Q_{\eta_1 + \cdots + \eta_N} h(x)\right| \tag{138}$$

$$\leq \sum_{i=1}^{N} \sup_{h \in \text{Lip}(1)} \left|Q_{\eta_1 + \cdots + \eta_{i-1}} \left(P_{\eta_{N-i+1}} - Q_{\eta_i}\right) P_{\eta_1 + \cdots + \eta_{(N-i)}} h(\mathbf{x})\right| \tag{139}$$

We can find the upper bound of the Wasserstein-1 distance. The difference from [6] is that the paper assumes a constant step size, but in the case of a quadratic schedule, it depends on i. When these differences are taken into account, the following holds.

$$\left|P_{\eta_{N-i+1}}f(\mathbf{x})-Q_{\eta_i}f(\mathbf{x})\right|\le C(1+|x|)\left(\|\nabla f\|_\infty+\left\|\nabla^2 f\right\|_{\mathrm{HS},\infty}\right)\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i). \quad (140)$$

Using the Lemma 3.1 [6],

$$\left|\left(P_{\eta_{N-i+1}}-Q_{\eta_i}\right)P_{\eta_1+\cdots+\eta_{N-i}}h(\mathbf{x})\right| \quad (141)$$

$$\le C(1+||\mathbf{x}||)\left(\left\|\nabla P_{\eta_1+\cdots+\eta_{N-i}}h\right\|_\infty+\left\|\nabla^2 P_{\eta_1+\cdots+\eta_{N-i}}h\right\|_{\mathrm{HS},\infty}\right)\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i) \quad (142)$$

$$\le C(1+||\mathbf{x}||)[\eta_1+\cdots+\eta_{N-i}]^{-1/\alpha}\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i) \quad (143)$$

$$\le C(1+||\mathbf{x}||)\frac{1}{N^{-2/\alpha}}\Big[\frac{(N-i)\cdot(N+i)}{N^2}\Big]^{-1/\alpha}\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i) \quad (144)$$

$$\le \tilde{C}(1+||\mathbf{x}||)\frac{1}{N^{-2/\alpha}}(N-i)^{-\frac{1}{\alpha}}(N+i)^{-\frac{1}{\alpha}}\max{}^{\frac{2}{\alpha}}(\eta_{N-i+1},\eta_i). \quad (145)$$

Combining this with the equation (2.4) in [6], then

$$\sup_{h\in\mathrm{Lip}(1)}\left|Q_{\eta_1+\cdots+\eta_{i-1}}\left(P_{\eta_{N-i+1}}-Q_{\eta_i}\right)P_{\eta_1+\cdots+\eta_{(N-i)}}h(x)\right| \quad (146)$$

$$\le C\left(1+\mathbb{E}[\left|Y_{i-1}^{\mathbf{x}}\right|]\right)\frac{1}{N^{-2/\alpha}}(N-i)^{-\frac{1}{\alpha}}(N+i)^{-\frac{1}{\alpha}}\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i) \quad (147)$$

$$\le C((1+||\mathbf{x}||)\frac{1}{N^{-2/\alpha}}(N-i)^{-\frac{1}{\alpha}}(N+i)^{-\frac{1}{\alpha}}\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i) \quad (148)$$

When $i\le\frac{N+1}{2}$, $\max^{2/\alpha}(\eta_{N-i+1},\eta_i)=\eta_i^{2/\alpha}$, $i>\frac{N+1}{2}$, $\max^{2/\alpha}(\eta_{N-i+1},\eta_i)=\eta_{N-i+1}^{2/\alpha}$. Thus,

$$\sum_{i=1}^{N}\frac{1}{N^{-2/\alpha}}(N-i)^{-\frac{1}{\alpha}}(N+i)^{-\frac{1}{\alpha}}\max{}^{2/\alpha}(\eta_{N-i+1},\eta_i) \quad (149)$$

$$\begin{aligned}&=\sum_{i=1}^{\frac{N+1}{2}}\frac{1}{N^{2/\alpha}}(N-i)^{-\frac{1}{\alpha}}(N+i)^{-\frac{1}{\alpha}}\Big(2i-1\Big)^{2/\alpha}\\&+\sum_{i=\frac{N+3}{2}}^{N}\frac{1}{N^{2/\alpha}}(N-i)^{-\frac{1}{\alpha}}(N+i)^{-\frac{1}{\alpha}}\Big(2(N-i)+1\Big)^{2/\alpha}\end{aligned} \quad (150)$$

$$\le\frac{1}{N^{2/\alpha}}\int_0^{\frac{N+1}{2}}(N-y)^{-\frac{1}{\alpha}}(N+y)^{-\frac{1}{\alpha}}(2y-1)^{2/\alpha}dy \quad (151)$$

$$+\frac{1}{N^{2/\alpha}}\int_{\frac{N+1}{2}}^{N}(N-y)^{-\frac{1}{\alpha}}(N+y)^{-\frac{1}{\alpha}}(2(N-y)+1)^{2/\alpha}dy. \quad (152)$$

Since

$$\int_0^{\frac{N+1}{2}}(N-y)^{-\frac{1}{\alpha}}(N+y)^{-\frac{1}{\alpha}}(2y-1)^{2/\alpha}dy \quad (153)$$

$$\le\Big(\int_0^{\frac{N+1}{2}}(N^2-y^2)^{-\frac{2}{\alpha}}dy\Big)^{\frac{1}{2}}\Big(\int_0^{\frac{N+1}{2}}(2y-1)^{\frac{4}{\alpha}}dy\Big)^{\frac{1}{2}} \quad (154)$$

$$\le C_1(\frac{1}{N})^{-1} \quad (155)$$

for some constant C_1 and

$$\int_{\frac{N+1}{2}}^{N}(N-y)^{-\frac{1}{\alpha}}(N+y)^{-\frac{1}{\alpha}}(2(N-y)+1)^{\frac{2}{\alpha}}dy \quad (156)$$

$$\le\Big(\int_{\frac{N+1}{2}}^{N}(N^2-y^2)^{-\frac{2}{\alpha}}dy\Big)^{\frac{1}{2}}\Big(\int_{\frac{N+1}{2}}^{N}(2(N-y)+1)^{\frac{4}{\alpha}}dy\Big)^{\frac{1}{2}} \quad (157)$$

$$\le C_2(\frac{1}{N})^{-1} \quad (158)$$

for some constants C_2.

This gives the upper bound

$$\sum_{i=1}^{N-1} \sup_{h \in \mathrm{Lip}(1)} \left| Q_{\eta_1+\cdots+\eta_{i-1}} \left(P_{\eta_{N-i+1}} - Q_{\eta_i} \right) P_{\eta_1+\cdots+\eta_{(N-i)}} h(x) \right| \quad \leq C^*(1+||\mathbf{x}||) \Big(\frac{1}{N} \Big)^{\frac{2}{\alpha}-1}. \tag{159}$$

From the same technique in [6], we can get the conclusion, $\mathcal{W}_1(\mu, \mu_{\frac{1}{N^2}}) \leq C^* \left(\frac{1}{N}\right)^{\frac{2}{\alpha}-1}$. It is possible to obtain the bound of the Wasserstein-1 distance $\mathcal{W}_1(\mu, \mu_{\frac{1}{N}^2})$ when applying the quadratic schedule to the fast stochastic sampling, replacing the inequality $\max^{2/\alpha}(\eta_{N-i+1}, \eta_i)$ with $\max^{1+\frac{1}{\alpha}}(\eta_{N-i+1}, \eta_i)$ and following the same proof above. □

Given the specific form of the drift term $\overleftarrow{b}(t, \mathbf{x})$ in the reverse-time SDE (107), It can be confirmed that the fractional score function meets the requirements outlined in Theorem E.2 if p_{data} is distributed according to a Gaussian distribution. Applying the Euler-Maruyama method to the diffusion process, it's been shown that the bounded Lipschitz distance between the invariant measures of the solution and its approximation using the Euler-Maruyama scheme with step size η is bounded by $O(\eta^{\frac{1}{2}})$ [36]. Additionally, since $\mathcal{W}_{bL} \leq \mathcal{W}_1$, it follows that $\mathcal{W}_{bL}(\mu, \mu_\eta) = O(\eta^{\frac{2}{\alpha}-1})$.

Corollary E.1 (Euler-Maruyama)**.** *Suppose the fractional score function in the SDE satisfies the conditions stated in Theorem E.2, and $a(t), \gamma(t)$ are bounded. Then, there exists a Markov chain $(\mathbf{x}_t)$ that follows the Euler-Maruyama scheme:*

$$\mathbf{x}_t = \left(1 + \frac{\beta(s)}{\alpha} \cdot \Delta t\right) \mathbf{x}_s + \alpha \cdot (\beta(s) \cdot \Delta t) S_t^{(\alpha)}(\mathbf{x}_s) + (\beta(s)\Delta t)^{1/\alpha} \boldsymbol{\epsilon}. \tag{160}$$

Here, $\boldsymbol{\epsilon} \sim \mathcal{S}\alpha\mathcal{S}^d(1)$ for $s > t$, and $\Delta t = s - t \ll 1$ such that the time step Δt is small. As a result of the conditions being satisfied, the Wasserstein-1 distance between the invariant measures of the solution and $(\mathbf{x}_t)$ is bounded by $(\Delta t)^{\frac{2}{\alpha}-1}$ [6].

For $t < s$, the solution of equation (5) can be represented as an integral, utilizing the semi-linear structure of the reverse SDE,

$$\begin{aligned}
\mathbf{x}_t &= \frac{a(t)}{a(s)} \mathbf{x}_s - \alpha \cdot a(t) \int_s^t \frac{\beta(u)}{a(u)} S_u^{(\alpha)}(\mathbf{x}_u) du + \int_s^t (\beta(u))^{\frac{1}{\alpha}} \frac{a(t)}{a(u)} d\bar{L}_t^\alpha && (161) \\
&= \frac{a(t)}{a(s)} \mathbf{x}_s - \alpha^2 \cdot a(t) \int_s^t \frac{d}{du} \Big(\frac{1}{a(u)} \Big) S_u^{(\alpha)}(\mathbf{x}_u) du + \int_s^t (\beta(u))^{\frac{1}{\alpha}} \frac{a(t)}{a(u)} d\bar{L}_t^\alpha. && (162)
\end{aligned}$$

From (161), We can get more faster sampling method by using an approximation for the second term $\int_s^t \frac{\beta(u)}{a(u)} S_u^{(\alpha)}(\mathbf{x}_u) du$. This term can be approximated as $(\int_s^t \frac{\beta(u)}{a(u)} du) \cdot S_s^{(\alpha)}(\mathbf{x}_s)$ and it is possible to calculate $\int_s^t \beta(u) \frac{a(t)}{a(u)}$, the scale parameter γ of which follows $\gamma^\alpha = \left| \int_s^t \frac{d}{du} \left(e^{-\int_u^t \beta(k)} \right) \right|$.

Theorem E.3 (Variant of Euler-Maruyama with dynamic time increment)**.** *Suppose the fractional score function in the SDE satisfies the conditions stated in Theorem E.2, and $a(t), \gamma(t)$ are bounded. Then, there exists a Markov chain $(\mathbf{x}_t)$ that follows:*

$$\mathbf{x}_t = \frac{a(t)}{a(s)} \mathbf{x}_s + \alpha^2 \Big(\frac{a(t)}{a(s)} - 1 \Big) S_s^{(\alpha)}(\mathbf{x}_s) + \left(\Big(\frac{a(t)}{a(s)} \Big)^\alpha - 1 \right)^{\frac{1}{\alpha}} \boldsymbol{\epsilon} \tag{163}$$

Here, $\boldsymbol{\epsilon} \sim \mathcal{S}\alpha\mathcal{S}^d(1)$ for $s > t$, and $\Delta t = s - t \ll 1$. As a result of the conditions being satisfied, the Wasserstein-1 distance between the invariant measures of the solution and $(\mathbf{x}_t)$ is bounded by $(\Delta t)^{\frac{1}{\alpha}}$ [6]. A modified version of the variant-Euler-Maruyama(v-Euler-Maruyama) exists to improve sample quality.

The paper [6] outlines a way to find bounds on the Wasserstein-1 distance by determining bounds for the approximation of the drift term, J_1, and the approximation of the stochastic term, J_2. In the equation, only the bound for the approximation of the drift term, J_1, is utilized. Furthermore, since the law of a weak solution is same to the law of the strong solution, the the Wasserstein-1 distance between the invariant measures for the strong solution and the approximation can be used for the that of the weak solution and the approximation as well.

Theorem E.4 (Probability fractional ODE)**.** *Let $b(t,\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ and $\sigma_L(t) : \mathbb{R} \to \mathbb{R}$ be functions that satisfy the Lipschitz condition as stated in [31]. For the SDE $d\overrightarrow{X}_t = b(t, \overrightarrow{X}_{t-})dt + \sigma(t)dL_t^\alpha$, the solution $(X_t)_{t\in[0,T]}$ of the SDE satisfies the following ODE:*

$$d\overrightarrow{X}_t \overset{d}{=} \left(b(t, \overrightarrow{X}_{t-}) - \sigma_L^\alpha(t) S_t^{(\alpha)}(\overrightarrow{X}_{t-}) \right) dt. \tag{164}$$

Due to its semilinear structure, the solution to (164) can be represented as an integral. This is shown in Lemma D.2. The Euler method can then be used to find the solution.

Corollary E.2 (Deterministic ODE sampling)**.** *If the drift term in (164) is Lipschitz continuous and the solution $\mathbf{x}_t$ has a bounded second derivative, then a sequence $(\mathbf{x}_t)$ can be obtained using the Euler-scheme:*

$$\mathbf{x}_t = \frac{a(t)}{a(s)} \mathbf{x}_s + \alpha \cdot \left(\frac{a(t)}{a(s)} - 1 \right) \cdot S_s^{(\alpha)}(\mathbf{x}_s) \tag{165}$$

where $s > t$. When the step size is Δt, the global truncation error is bounded by $O(\Delta t)$ [5].

F Fractional score function for Lévy-Itō Models

Lemma F.1. *The density function of $\mathcal{S}\alpha\mathcal{S}^d(1)$ for $\mathbf{x} \in \mathbb{R}^d$ is represented by $q_\alpha(\mathbf{x})$. Given an initial value $\mathbf{x}_0$ that follows the distribution p_{data} and a random variable $\boldsymbol{\epsilon}$ that follows the distribution $\mathcal{S}\alpha\mathcal{S}^d(1)$, the value of X_t can be represented as $\mathbf{x}_t = a(t)\mathbf{x}_0 + \gamma(t)\boldsymbol{\epsilon}$. The transition density function $p_t(\mathbf{x}_t|\mathbf{x}_0)$ given $\mathbf{x}_0$, can be expressed as $p_t(\mathbf{x}_t|\mathbf{x}_0) = \frac{q_\alpha(\boldsymbol{\epsilon})}{\gamma^d(t)}$.*

Proof. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^s$, we denote $\mathbf{x} \le \mathbf{y}$ if $[\mathbf{x}]_i \le [\mathbf{y}]_i$ for all $i \in \{1, \ldots, d\}$. Let $\overrightarrow{X}_t$ and Y be defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ where the transition density function of $\overrightarrow{X}_t$ is $p_t(\mathbf{x}_t|\mathbf{x}_0) = \frac{d\mathbb{P}(\overrightarrow{X}_t \le \mathbf{x}_t | \overrightarrow{X}_t = \mathbf{x}_0)}{d\mathbf{x}_t}$ and the density function of Y is $q_\alpha(y) = \frac{d\mathbb{P}(Y \le y)}{dy}$. Let $\overrightarrow{X}_t = a(t)\mathbf{x}_0 + \gamma(t)\boldsymbol{\epsilon}$. Then

$$\mathbb{P}(\overrightarrow{X}_t \le \mathbf{x}_t | \overrightarrow{X}_0 = \mathbf{x}_0) = \mathbb{P}(a(t)\mathbf{x}_0 + \gamma(t)Y \le \mathbf{x}_t) \text{ since } \overrightarrow{X}_t = a(t)\mathbf{x}_0 + \gamma(t)Y \tag{166}$$

$$= \mathbb{P}(Y \le \frac{\mathbf{x}_t - a(t)\mathbf{x}_0}{\gamma(t)}) \tag{167}$$

$$= \mathbb{P}(Y \le \boldsymbol{\epsilon}). \tag{168}$$

Since the probability density function $q_\alpha(\boldsymbol{\epsilon})$ satisfies the relation $q_\alpha(\boldsymbol{\epsilon}) = \frac{\partial \cdots \partial \mathbb{P}(Y \le \boldsymbol{\epsilon})}{\partial_{\epsilon_1} \cdots \partial_{\epsilon_d}}$, we obtain $p_t(\mathbf{x}_t|\mathbf{x}_0) = \frac{q_\alpha(\boldsymbol{\epsilon})}{\gamma^d(t)}$. ☐

Theorem F.1 (Fractional Denoising Score Matching (fDSM))**.** *For parameter θ, we define two losses $L_1(\theta, t)$ and $L_2(\theta, t)$ for $t \in [0, 1]$ such that*

$$L_1(\theta, t) = \mathbb{E}_{\mathbf{x}_t \sim p_t(\mathbf{x}_t)} \left[\left\| S_\theta(\mathbf{x}_t, t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}_t)}{p_t(\mathbf{x}_t)} \right\|_2^2 \right]. \tag{169}$$

and

$$L_2(\theta, t) = \mathbb{E}_{\mathbf{x}_0 \sim p_{data}(\mathbf{x}_0), \mathbf{x}_t \sim p_t(\mathbf{x}_t|\mathbf{x}_0)} \left[\left\| S_\theta(\mathbf{x}_t, t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}_t|\mathbf{x}_0)}{p_t(\mathbf{x}_t|\mathbf{x}_0)} \right\|_2^2 \right]. \tag{170}$$

Then two losses are equivalent, meaning that there exists a constant C satisfying $L_1(\theta, t) = L_2(\theta, t) + C$.

Proof. For $t \in [0,1]$, due to the monotone convergence theorem, it holds that

$$L_1(\theta,t) = \mathbb{E}_{\mathbf{x}_t \sim p_t(\mathbf{x}_t)} \left[\frac{1}{2} \left\| S_\theta(\mathbf{x}_t,t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}_t)}{p_t(\mathbf{x}_t)} \right\|_2^2 \right] \tag{171}$$

$$= \lim_{r\to\infty} \int_{|\mathbf{x}|<r} p_t(\mathbf{x}) \frac{1}{2} \left\| S_\theta(\mathbf{x},t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \right\|_2^2 d\mathbf{x} \tag{172}$$

where the last integral can be decomposed into two terms $L_1(\theta,t;r)$ and $C_1(r)$ as follows:

$$\int_{|\mathbf{x}|<r} p_t(\mathbf{x}) \frac{1}{2} \left\| S_\theta(\mathbf{x},t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \right\|_2^2 d\mathbf{x} \tag{173}$$

$$= \int_{|\mathbf{x}|<r} p_t(\mathbf{x}) \left[\frac{1}{2} \|S_\theta(\mathbf{x},t)\|_2^2 - \left\langle S_\theta(\mathbf{x},t), \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \right\rangle \right] d\mathbf{x} \quad (=: L_1(\theta,t;r)) \tag{174}$$

$$+ \int_{|\mathbf{x}|<r} \frac{1}{2} p_t(\mathbf{x}) \left\| \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \right\|_2^2 d\mathbf{x} \quad (=: C_1(r)) \tag{175}$$

$$= L_1(\theta,t;r) + C_1(r). \tag{176}$$

Note that $C_1(r)$ is independent of θ and well-defined for any $r > 0$ provided that p_{data} has the compact support or $p_{\text{data}}(\mathbf{x}) \sim e^{-|\mathbf{x}|^2}$. Similarly, due to the monotone convergence theorem,

$$L_2(\theta,t) = \mathbb{E}_{\mathbf{x}_0 \sim p_{\text{data}}(\mathbf{x}_0), \mathbf{x}_t \sim p_t(\mathbf{x}_t|\mathbf{x}_0)} \left[\frac{1}{2} \left\| S_\theta(\mathbf{x}_t,t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}_t|\mathbf{x}_0)}{p_t(\mathbf{x}_t|\mathbf{x}_0)} \right|_2^2 \right] \tag{177}$$

$$= \lim_{r\to\infty} \int_{|\mathbf{x}|<r} \int_{\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) p_t(\mathbf{x}|\mathbf{x}_0) \frac{1}{2} \left\| S_\theta(\mathbf{x},t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} \right\|_2^2 d\mathbf{x}_0 d\mathbf{x} \tag{178}$$

where the last integral term can be decomposed into two terms $L_2(\theta,t;r)$ and $C_2(r)$ as follows:

$$\int_{|\mathbf{x}|<r} \int_{\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) p_t(\mathbf{x}|\mathbf{x}_0) \frac{1}{2} \left\| S_\theta(\mathbf{x},t) - \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} \right\|_2^2 d\mathbf{x}_0 d\mathbf{x} \tag{179}$$

$$= \int_{|\mathbf{x}|<r} \int_{\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) p_t(\mathbf{x}|\mathbf{x}_0) \left[\frac{1}{2} \|S_\theta(\mathbf{x},t)\|_2^2 - \left\langle S_\theta(\mathbf{x},t), \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} \right\rangle \right] d\mathbf{x}_0 d\mathbf{x} \tag{180}$$

$$+ \int_{|\mathbf{x}|<r} \int_{\mathbb{R}^d} \frac{1}{2} p_{\text{data}}(\mathbf{x}_0) p_t(\mathbf{x}|\mathbf{x}_0) \left\| \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} \right\|_2^2 d\mathbf{x}_0 d\mathbf{x} = L_2(\theta,t;r) + C_2(r). \tag{181}$$

Also, note that $C_2(r)$ is independent of θ and well-defined for any $r > 0$ with the same condition as we already mentioned. However, $C_1(r)$ and $C_2(r)$ may diverge as $r \to \infty$. Thus, we control $C_3(r) := C_1(r) - C_2(r)$ instead of controlling C_1 and C_2 individually. Observe that for any $r > 0$,

$$\int_{|\mathbf{x}|<r} p_t(\mathbf{x}) \left\langle S_\theta(\mathbf{x},t), \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} \right\rangle d\mathbf{x} = \int_{|\mathbf{x}|<r} \left\langle S_\theta(\mathbf{x},t), \Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}) \right\rangle d\mathbf{x} \tag{182}$$

$$= \int_{|\mathbf{x}|<r} \left\langle S_\theta(\mathbf{x},t), \int_{\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) \Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0) d\mathbf{x}_0 \right\rangle d\mathbf{x} \tag{183}$$

$$= \int_{|\mathbf{x}|<r} \left\langle S_\theta(\mathbf{x},t), \int_{\mathbf{x}_0} p_{\text{data}}(\mathbf{x}_0) p_t(\mathbf{x}|\mathbf{x}_0) \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} \right\rangle d\mathbf{x}_0 d\mathbf{x} \tag{184}$$

$$= \int_{|\mathbf{x}|<r} \int_{\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) p_t(\mathbf{x}|\mathbf{x}_0) \left\langle S_\theta(\mathbf{x}), \frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} \right\rangle d\mathbf{x}_0 d\mathbf{x}. \tag{185}$$

Thus, we can conclude that $L_1(\theta,t;r) = L_2(\theta,t;r)$ for any $r>0$ so that

$$L_1(\theta,t;r) + C_1(r) = L_2(\theta,t;r) + C_1(r) = L_2(\theta,t;r) + C_2(r) + C_3(r). \tag{186}$$

Recall that $L_1(\theta,t;r)+C_1(r) \uparrow L_1(\theta,t)$ and $L_2(\theta,t;r)+C_2(r)\uparrow L_2(\theta,t)$ as $r\to\infty$. Thus, the limit of $C_3(r)$ exists as $r\to\infty$ and we write it as C_3. Therefore,

$$\begin{aligned} L_1(\theta,t) &= \lim_{r\to\infty}\left(L_1(\theta,t;r)+C_1(r)\right) && (187)\\ &= \lim_{r\to\infty}\left(L_2(\theta,t;r)+C_2(r)+C_3(r)\right) = L_2(\theta,t)+C_3. && (188)\end{aligned}$$

Consequently, we show that $L_1(\theta,t)$ and $L_2(\theta,t)$ are equivalent. ◻

According to Theorem 4.3, the label of the fractional SDM is

$$\begin{aligned} \frac{\Delta^{\frac{\alpha-2}{2}}\nabla p_t(\mathbf{x}|\mathbf{x}_0)}{p_t(\mathbf{x}|\mathbf{x}_0)} &= \frac{1}{\sigma_L^{\alpha-1}(t)}\frac{\int_{\mathbf{u}\in\mathbb{R}^d}\left(-i\mathbf{u}||\mathbf{u}||^{\alpha-2}e^{-i\langle\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\sigma_L(t)},u\rangle}e^{-||\mathbf{u}||^{\alpha}}\right)d\mathbf{u}}{\int_{\mathbf{u}\in\mathbb{R}^d}\left(e^{-i\langle\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\sigma_L(t)},\mathbf{u}\rangle}e^{-||\mathbf{u}||^{\alpha}}\right)d\mathbf{u}} && (189)\\ &= \frac{1}{\sigma_L^{\alpha-1}(t)}\frac{\Delta^{\frac{\alpha-2}{2}}\nabla q_\alpha(\mathbf{x})}{q_\alpha(\mathbf{x})}. && (190)\end{aligned}$$

where $\mathbf{x} = \frac{\mathbf{x}-a(t)\mathbf{x}_0}{\sigma_L(t)}$.

Even though it may seem complex, the integral representation of the label of fDSM can be simplified to a straightforward linear function.

Lemma F.2. $\frac{\Delta^{\frac{\alpha-2}{2}}\nabla q_\alpha(\mathbf{x})}{q_\alpha(\mathbf{x})}$ *has a 2-dimensional integral representation such as*

$$\frac{-i\int_0^\pi\int_0^\infty e^{-ir||\mathbf{x}||\cos\theta}e^{-r^\alpha}r^{d+\alpha-2}\cos\theta\sin^{d-2}\theta drd\theta}{\int_0^\pi\int_0^\infty e^{-ir||\mathbf{x}||\cos\theta}e^{-r^\alpha}r^{d-1}\sin^{d-2}\theta drd\theta}\frac{\mathbf{x}}{||\mathbf{x}||}. \tag{191}$$

Proof. With a fixed $\mathbf{x}$ in $\mathbb{R}^d$, we select an axis that is aligned with the direction of $\mathbf{x}$. Then, we define $\hat{x}$ as $\frac{x}{||x||}$ and include it as the last component of an orthogonal basis $\tilde{B}$ of $\mathbb{R}^d$, where $\tilde{B} = \tilde{e}_1,\dots,\tilde{e}_{d-1},\hat{x}(\doteq\tilde{e}_d)$. The choice of points does not affect the independence of the two bases. $\tilde{B}$ is formed by rotating the basis B so that the absolute value of the determinant of the jacobian $\frac{\partial(u_1,\dots,u_d)}{\partial(\tilde{u}_1,\dots,\tilde{u}_d)}$ is 1. If $\mathbf{f} = f_1\hat{e}_1+\cdots f_d\hat{e}_d$, we denote $[\mathbf{f}]_B$ as $[\mathbf{f}]_B = (f_1,\dots,f_d)$ for basis B, $[\mathbf{f}]_{\tilde{B}} = (\tilde{f}_1,\dots,\tilde{f}_d)$ for basis $\tilde{B}$. Given $\mathbf{u}\in\mathbb{R}^d$, $\mathbf{u}$ can represented as $\mathbf{u} = u_1\hat{e}_1+\ldots+u_m\hat{e}_m+\ldots+u_d\hat{e}_d = \tilde{u}_1\tilde{e}_1+\cdots+\tilde{u}_m\tilde{e}_m+\cdots+\tilde{u}_d\hat{x}$ and $\mathbf{x} = x_1\hat{e}_1+\cdots x_d\hat{e}_d = ||\mathbf{x}||\hat{x}$ with $[\mathbf{u}]_B = u = (u_1,\dots,u_d)$ and $[\mathbf{u}]_{\tilde{B}} = \tilde{u} = (\tilde{u}_1,\dots\tilde{u}_d)$, $[\mathbf{x}]_B = (x_1,\dots,x_d)\in\mathbb{R}^d$, $\tilde{x} = (0,\dots,0,||\mathbf{x}||)$.

Then $u_m = \sum_{j=1}^d\tilde{u}_j\langle\hat{e}_m,\tilde{e}_j\rangle = \sum_{j=1}^d\tilde{u}_j(u)\langle\hat{e}_m,\tilde{e}_j\rangle$ for each $m\in\{1,\dots,d\}$ by orthogonality of the basis B.

Lemma F.3 (Change of variable)**.** *T is a 1-1 $\mathcal{C}'$-mapping of an open set $E\subset R^d$ into R^d such that $J_T(\mathbf{x})\neq 0$ for all $\mathbf{x}\in E$. Iff is a continuous function on R^k whose support is compact and lies in $T(E)$, then*

$$\int_{R^d} f(\mathbf{y})d\mathbf{y} = \int_{R^d} f(T(\mathbf{x}))\,|J_T(\mathbf{x})|\,d\mathbf{x}. \tag{192}$$

We recall that J_T is the Jacobian of T.

For a given $\mathbf{x}\in\mathbb{R}^d$, let us define a rotation operator $T_{\mathbf{x}}:\mathbb{R}^d\to\mathbb{R}^d$ such that $T_{\mathbf{x}}(\mathbf{u}) = \tilde{\mathbf{u}}$ where

$$u_m = \sum_{j=1}^d\tilde{u}_j\langle\tilde{e}_j,\hat{e}_m\rangle = \sum_{j=1}^d\mathrm{Proj}_j(T_{\mathbf{x}}(\mathbf{u}))\langle\hat{e}_m,\tilde{e}_j\rangle,\quad m\in\{1,\dots,d\} \tag{193}$$

where $\text{Proj}_i(\mathbf{x}) = x_i$ is the projection map onto the corresponding basis of $\tilde{B}$. Also, note that $\|T_{\mathbf{x}}(\mathbf{u})\| = \|\mathbf{u}\|$ and $\langle \mathbf{x}, \mathbf{u}\rangle = \|\mathbf{x}\| \cdot \text{Proj}_d(T_{\mathbf{x}}(\mathbf{u}))$ holds for any $\mathbf{u} \in \mathbb{R}^d$. Then for each m,

$$\int_{u\in\mathbb{R}^d} (-iu_m||\mathbf{u}||^{\alpha-2} e^{-i\langle \mathbf{x},\mathbf{u}\rangle} e^{-||\mathbf{u}||^\alpha}) d\mathbf{u} \tag{194}$$

$$\overset{(193)}{=} -i \int_{\mathbb{R}^d} \sum_{j=1}^{d} \text{Proj}_j(T_{\mathbf{x}}(\mathbf{u})) \langle \hat{e}_m, \tilde{e}_j\rangle ||\mathbf{u}||^{\alpha-2} e^{-i\langle \mathbf{x},\mathbf{u}\rangle} e^{-||\mathbf{u}||^\alpha} d\mathbf{u} \tag{195}$$

$$= -i \sum_{j=1}^{d} \langle \hat{e}_m, \tilde{e}_j\rangle \int_{\mathbb{R}^d} \text{Proj}_j(T_{\mathbf{x}}(\mathbf{u})) \|T_{\mathbf{x}}(\mathbf{u})\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \text{Proj}_d(T_{\mathbf{x}}(\mathbf{u}))} e^{-\|T_{\mathbf{x}}(\mathbf{u})\|^\alpha} d\mathbf{u} \tag{196}$$

$$\overset{(192)}{=} -i \sum_{j=1}^{d} \langle \hat{e}_m, \tilde{e}_j\rangle \int_{\mathbb{R}^d} \text{Proj}_j(\tilde{\mathbf{u}}) \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \text{Proj}_d(\tilde{\mathbf{u}})} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}} \tag{197}$$

$$= -i \sum_{j=1}^{d} \langle \hat{e}_m, \tilde{e}_j\rangle \int_{\mathbb{R}^d} \tilde{u}_j \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \tilde{u}_d} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}}. \tag{198}$$

Therefore, we get

$$(\Delta)^{\frac{\alpha-2}{2}} \nabla q_\alpha(\mathbf{x}) = \int_{\mathbb{R}^d} -i\mathbf{u}\|\mathbf{u}\|^{\alpha-2} e^{-i\langle \mathbf{x},\mathbf{u}\rangle} e^{-\|\mathbf{u}\|^\alpha} d\mathbf{u} \tag{199}$$

$$\overset{\text{Def.}}{=} \sum_{m=1}^{d} \left(\int_{\mathbb{R}^d} -Qu_m \|\mathbf{u}\|^{\alpha-2} e^{-i\langle \mathbf{x},\mathbf{u}\rangle} e^{-\|\mathbf{u}\|^\alpha} d\mathbf{u} \right) \hat{e}_m \tag{200}$$

$$\overset{(198)}{=} \sum_{m=1}^{d} \left(-i \sum_{j=1}^{d} \langle \hat{e}_m, \tilde{e}_j\rangle \int_{\mathbb{R}^d} \tilde{u}_j \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \tilde{u}_d} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}} \right) \hat{e}_m \tag{201}$$

$$= \sum_{j=1}^{d} \left(-i \int_{\mathbb{R}^d} \tilde{u}_j \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \tilde{u}_d} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}} \right) \tilde{e}_j \tag{202}$$

where the last equality holds since $\tilde{e}_j = \sum_{m=1}^{d} \langle \hat{e}_m, \tilde{e}_j\rangle \hat{e}_m$ for each $j \in \{1, \ldots, d\}$.

The Cartesian coordinates are converted to spherical coordinates to perform the above integration. For $j = 1$,

$$I_1 := \int_{\mathbb{R}^d} \tilde{u}_1 \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \tilde{u}_d} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}} \tag{203}$$

$$= \int_{\mathbb{S}^{d-1}} \int_0^\infty r^{\alpha-1} \sin\theta_1 e^{-ir\|\mathbf{x}\| \cos\theta_{d-1}} e^{-r^\alpha} \prod_{k=2}^{d-1} \sin\theta_k dr d\sigma_{d-1}, \tag{204}$$

The same calculation is performed for $j \in \{2, \ldots, d-1\}$, and the results are represented as

$$I_j := \int_{\mathbb{R}^d} \tilde{u}_j \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \tilde{u}_d} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}} \tag{205}$$

$$= \int_{\mathbb{S}^{d-1}} \int_0^\infty r^{\alpha-1} \cos\theta_{j-1} e^{-ir\|\mathbf{x}\| \cos\theta_{d-1}} e^{-r^\alpha} \prod_{k=j}^{d-1} \sin\theta_k dr d\sigma_{d-1} \tag{206}$$

and for $j = d$,

$$I_d := \int_{\mathbb{R}^d} \tilde{u}_d \|\tilde{\mathbf{u}}\|^{\alpha-2} e^{-i\|\mathbf{x}\|\cdot \tilde{u}_d} e^{-\|\tilde{\mathbf{u}}\|^\alpha} d\tilde{\mathbf{u}} \tag{207}$$

$$= \int_{\mathbb{S}^{d-1}} \int_0^\infty r^{\alpha-1} \cos\theta_{d-1} e^{-ir\|\mathbf{x}\| \cos\theta_{d-1}} e^{-r^\alpha} dr d\sigma_{d-1} \tag{208}$$

$$= \int_0^{2\pi} \int_0^{\pi} \cdots \left[\int_0^{\pi} \int_0^{\infty} r^{\alpha-1} \cos\theta_{d-1} e^{-ir\|\mathbf{x}\| \cos\theta_{d-1}} e^{-r^\alpha} J_d dr d\theta_{d-1} \right] \cdots d\theta_1. \tag{209}$$

Here J_d denotes the Jacobian

$$J_d = (-1)^{d-1} r^{d-1} \prod_{k=2}^{d-1} \sin^{k-1}\theta_k, \tag{210}$$

where $\theta_1 \in [0, 2\pi)$ and $\theta_j \in [0, \pi)$ for $j \in \{2, \ldots, d\}$. It is shown that $I_1 = I_2 = \cdots = I_{d-1} = 0$ holds. while $q_\alpha(\mathbf{x})$ can be calculated as an integral over $\mathbb{R}^d$.

$$\begin{aligned}
q_\alpha(\mathbf{x}) &= \int_{\mathbb{R}^d} e^{-i\langle \mathbf{x}, \mathbf{u}\rangle} e^{-\|\mathbf{u}\|^\alpha} d\mathbf{u} && (211)\\
&= \int_{\mathbb{S}^{d-1}} \int_0^\infty e^{-ir\|\mathbf{x}\|\cos\theta_{d-1}} e^{-r^\alpha} dr d\sigma_{d-1} && (212)\\
&= \int_0^{2\pi} \int_0^\pi \cdots \left[\int_0^\pi \int_0^\infty e^{-ir\|\mathbf{x}\|\cos\theta_{d-1}} e^{-r^\alpha} J_d dr d\theta_{d-1}\right] \cdots d\theta_1 =: I_0. && (213)
\end{aligned}$$

Then we get the desired result

$$\frac{(\Delta)^{\frac{\alpha-2}{2}} \nabla q_\alpha(\mathbf{x})}{q_\alpha(\mathbf{x})} = \sum_{j=1}^d \frac{I_j}{I_0} \tilde{e}_j = \frac{I_d}{I_0} \tilde{e}_d = \frac{I_d}{I_0} \frac{\mathbf{x}}{\|\mathbf{x}\|}. \tag{214}$$

I_d and I_0 have the same term, $\int_0^{2\pi} \int_0^\pi \cdots \int_0^\pi \prod_{k=2}^{d-1} \sin^{k-1}\theta_k d\theta_{d-2} \cdots d\theta_1$, $\frac{I_d}{I_0}$ can be represented as the 2-dimensional integral,

$$\frac{(\Delta)^{\frac{\alpha-2}{2}} \nabla q_\alpha(\mathbf{x})}{q_\alpha(\mathbf{x})} = \frac{(-i\int_0^\pi \int_0^\infty e^{-ir|||\mathbf{x}||\cos\theta_{d-1}} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta_{d-1} \sin^{d-2}\theta_{d-1} dr d\theta_{d-1})}{(\int_0^\pi \int_0^\infty e^{-ir|||\mathbf{x}||\cos\theta_{d-1}} e^{-r^\alpha} r^{d-1} \sin^{d-2}\theta_{d-1} dr d\theta_{d-1})} \frac{\mathbf{x}}{|||\mathbf{x}||}. \tag{215}$$

$\square$

Lemma F.4. $\frac{-i\int_0^\pi \int_0^\infty e^{-ir|||\mathbf{x}||\cos\theta} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin\theta^{d-2} dr d\theta}{\int_0^\pi \int_0^\infty e^{-ir|||\mathbf{x}||\cos\theta_{d-1}} e^{-r^\alpha} r^{d-1} \sin\theta^{d-2} dr d\theta} \frac{\mathbf{x}}{|||\mathbf{x}||} = -\frac{\mathbf{x}}{\alpha}.$

Proof. The integral is split into two parts and then combined to form a final result. We first estimate the numerator of the integral.

$$\begin{aligned}
I_d :&= \int_0^\pi \int_0^\infty e^{-ir||\mathbf{x}||\cos\theta} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta dr d\theta && (216)\\
&= \int_0^{\pi/2} \int_0^\infty e^{-ir||\mathbf{x}||\cos\theta} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta dr d\theta && (217)\\
&+ \int_{\pi/2}^\pi \int_0^\infty e^{-ir||\mathbf{x}||\cos\theta} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta dr d\theta && (218)\\
&= \int_0^{\pi/2} \int_0^\infty e^{-ir||\mathbf{x}||\cos\theta} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta dr d\theta && (219)\\
&- \int_0^{\pi/2} \int_0^\infty e^{ir|||\mathbf{x}||\cos\theta} e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta dr d\theta && (220)\\
&= -2i \cdot \int_0^{\pi/2} \int_0^\infty \sin(r|||\mathbf{x}||\cos\theta) e^{-r^\alpha} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta dr d\theta. && (221)
\end{aligned}$$

The expression is then changed to a new form by using $y = r\cos\theta$ and the relationship between y, r, and θ. The integral is simplified by using the derivative of $(r^2 - y^2)^{(d-1)/2}$ with respect to y and then integrated to arrive at the final result.

$$I_d = -2i \cdot \int_0^{\infty} \int_0^{\pi/2} \sin(r||\mathbf{x}|| \cos\theta) e^{-r^{\alpha}} r^{d+\alpha-2} \cos\theta \sin^{d-2}\theta d\theta dr \tag{222}$$

$$= -2i \cdot \int_0^{\infty} e^{-r^{\alpha}} r^{d+\alpha-2} dr \Big[\int_0^r \sin(||\mathbf{x}||y) \cdot \frac{y}{r} \cdot \sin^{d-2}\theta \frac{1}{r\sin\theta} dy \Big] \tag{223}$$

$$= -2i \cdot \int_0^{\infty} e^{-r^{\alpha}} r^{d+\alpha-2} r^{-d+1} dr \Big[\int_0^r \sin(||\mathbf{x}||y) \cdot y \cdot (r^2-y^2)^{\frac{d-3}{2}} dy \Big] \tag{224}$$

$$= -2i \cdot \int_0^{\infty} e^{-r^{\alpha}} r^{\alpha-1} dr \Big[\int_0^r \sin(||\mathbf{x}||y) \cdot y \cdot (r^2-y^2)^{\frac{d-3}{2}} dy \Big]. \tag{225}$$

The integral $\int_0^r \sin(||\mathbf{x}||y) \cdot y \cdot (r^2-y^2)^{\frac{d-3}{2}} dy$ is simplified as follows. Using the property $y(r^2 - y^2)^{(d-3)/2} = -\frac{1}{d-1}\frac{d}{dy}(r^2-y^2)^{(d-1)/2}$, we can rewrite the integral as

$$\int_0^r \sin(||\mathbf{x}||y) y (r^2-y^2)^{(d-3)/2} dy = -\frac{1}{d-1} \int_0^r \sin(||\mathbf{x}||y) \frac{d}{dy}(r^2-y^2)^{(d-1)/2} dy \tag{226}$$

$$= -\frac{1}{d-1}(r^2-y^2)^{(d-1)/2} \sin(||\mathbf{x}||y)\big|_0^r + \frac{||\mathbf{x}||}{d-1} \int_0^r (r^2-y^2)^{(d-1)/2} \cos(||\mathbf{x}||y) dy \tag{227}$$

$$= \frac{||\mathbf{x}||}{d-1} \int_0^r (r^2-y^2)^{(d-1)/2} \cos(||\mathbf{x}||y) dy. \tag{228}$$

We then focus on estimating $\frac{||\mathbf{x}||}{d-1} \int_0^r (r^2-y^2)^{(d-1)/2} \cos(||\mathbf{x}||y) dy$.

$$\int_0^r (r^2-y^2)^{(d-1)/2} \cos(||\mathbf{x}||y) dy = r^{(d-1)} \int_0^r (1-(\frac{y}{r})^2)^{\frac{d-1}{2}} \cos(||\mathbf{x}||y) dy \tag{229}$$

$$= \frac{r^{d-1}}{||\mathbf{x}||} \int_0^{||\mathbf{x}||r} (1-(\frac{y}{||\mathbf{x}||r})^2)^{\frac{d-1}{\cdot}2} \cos(y) dy. \tag{230}$$

Setting $y = ||\mathbf{x}||r\cos k$, yields $dy = ||\mathbf{x}||r(-\sin k)dk$ and results in $y = 0 \Rightarrow k = \frac{\pi}{2}$, and $y = r||\mathbf{x}|| \Rightarrow k = 0$. Hence,

$$\int_0^r (r^2-y^2)^{(d-1)/2} \cos(||\mathbf{x}||y) dy = r^{(d-1)} \int_0^r (1-(\frac{y}{r})^2)^{\frac{d-1}{2}} \cos(||\mathbf{x}||y) dy \tag{231}$$

$$= r^d \int_0^{\frac{\pi}{2}} \sin^{d-1}(k) \cos(||\mathbf{x}||r\cos(k)) dk. \tag{232}$$

Finally, using the integral representation of the Bessel function of the first kind (as given in [1] .360, 9.1.20),

$$J_{\nu}(z) = \frac{\left(\frac{1}{2}z\right)^{\nu}}{\pi^{\frac{1}{3}}\Gamma\left(\nu+\frac{1}{2}\right)} \int_0^{\pi} \cos(z\cos\theta) \sin^{2\nu}\theta d\theta \tag{233}$$

$$\Rightarrow r^d \int_0^{\frac{\pi}{2}} \sin^{d-1}(k) \cos(||\mathbf{x}||r\cos(k)) dk = r^{d/2} 2^{\frac{d}{2}-1} \sqrt{\pi} \frac{\Gamma(\frac{d}{2}+\frac{1}{2})}{||\mathbf{x}||^{d/2}} J_{\frac{d}{2}}(r||\mathbf{x}||). \tag{234}$$

Thus $\frac{||\mathbf{x}||}{d-1} \int_0^r (r^2-y^2)^{(d-3)/2} \cos(||\mathbf{x}||y) dy = r^{d/2} 2^{\frac{d}{2}-1} \sqrt{\pi} \frac{\Gamma(\frac{d}{2}+\frac{1}{2})}{||\mathbf{x}||^{d/2-1}(d-1)} J_{\frac{d}{2}}(r||\mathbf{x}||)$.

The denominator of the integral is estimated as follows:

$$I_0 := \int_0^{\pi} \int_0^{\infty} e^{-ir|||\mathbf{x}|| \cos\theta_{d-1}} e^{-r^{\alpha}} r^{d-1} \sin\theta_{d-1}^{d-2} dr d\theta_{d-1} \tag{235}$$

$$= \int_0^{\pi/2} \int_0^{\infty} e^{-ir|||\mathbf{x}|| \cos\theta} e^{-r^{\alpha}} r^{d-1} \sin_{d-2}\theta dr d\theta \tag{236}$$

$$+ \int_{\pi/2}^{\pi} \int_0^{\infty} e^{-ir||\mathbf{x}|| \cos\theta} e^{-r^{\alpha}} r^{d-1} \sin_{d-2}\theta dr d\theta \tag{237}$$

$$= 2 \cdot \int_0^{\infty} e^{-r^{\alpha}} r dr \int_0^{\pi/2} \cos(r|||\mathbf{x}|| \cos\theta) \sin^{d-2}\theta d\theta. \tag{238}$$

To estimate the denominator, a change of variables is made: $y = r\cos\theta$. Then $\cos\theta = \frac{y}{r}$, $\sin\theta = \frac{(r^2-y^2)^{1/2}}{r}$, and $dy = -r\sin\theta d\theta$ with $d\theta = -\frac{dy}{r\sin\theta}$. The integral becomes:

$$I_0 = 2\cdot\int_0^\infty e^{-r^\alpha} r^{d-2} dr\Big[\int_0^r \cos(||\mathbf{x}||y)\sin^{d-3}\theta dy\Big] \tag{239}$$

$$= 2\cdot\int_0^\infty e^{-r^\alpha} r^{d-2} dr\Big[\int_0^r \cos(||\mathbf{x}||y) r^{-d+3}(r^2-y^2)^{\frac{d-3}{2}} dy\Big] \tag{240}$$

$$= 2\cdot\int_0^\infty e^{-r^\alpha} r dr\Big[\int_0^r \cos(||\mathbf{x}||y)(r^2-y^2)^{\frac{d-3}{2}} dy\Big]. \tag{241}$$

In a similar manner, the integral $\int_0^r (r^2-y^2)^{(d-3)/2}\cos(||\mathbf{x}||y)dy$ can be estimated.

$$\int_0^r (r^2-y^2)^{(d-3)/2}\cos(||\mathbf{x}||y)dy = r^{\frac{d}{2}-1}2^{\frac{d}{2}-2}\sqrt{\pi}\frac{\Gamma(\frac{d}{2}-\frac{1}{2})}{||\mathbf{x}||^{\frac{d}{2}-1}} J_{\frac{d}{2}-1}(r||\mathbf{x}||). \tag{242}$$

The formula for I is rearranged as follows:

$$I = -\frac{\int_0^\infty e^{-r^\alpha} r^{\frac{d}{2}} 2^{\frac{d}{2}-1} r^{\alpha-1}\sqrt{\pi}\Gamma(\frac{d}{2}+\frac{1}{2}) J_{\frac{d}{2}}(r||\mathbf{x}||)dr\frac{1}{||\mathbf{x}||^{\frac{d}{2}(d-1)}}}{\int_0^\infty e^{-r^\alpha} r^{\frac{d}{2}-1} r 2^{\frac{d}{2}-2}\sqrt{\pi}\Gamma(\frac{d}{2}-\frac{1}{2}) J_{\frac{d}{2}-1}(r||\mathbf{x}||)dr\frac{1}{||\mathbf{x}||^{\frac{d}{2}-1}}} \tag{243}$$

$$= -\frac{\int_0^\infty r^{\alpha-1+d/2} e^{-r^\alpha} J_{d/2}(r||\mathbf{x}||)dr}{\int_0^\infty r^{d/2} e^{-r^\alpha} J_{d/2-1}(r||\mathbf{x}||)dr}. \tag{244}$$

$$\tag{245}$$

The integral of I is represented as a series of r using the series representation of the Bessel function of the first kind: ([1], p.360, 9.1.10),

$$J_\alpha(x) = \sum_{m=0}^\infty \frac{(-1)^m}{m!\Gamma(m+\alpha+1)}\left(\frac{x}{2}\right)^{2m+\alpha}. \tag{246}$$

The result is:

$$\int_0^\infty r^{\alpha-1+d/2} e^{-r^\alpha} J_{d/2}(r||\mathbf{x}||)dr = \int_0^\infty r^{2m+\alpha-1+d} e^{-r^\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!\Gamma(m+\frac{d}{2}+1)}(\frac{||\mathbf{x}||}{2})^{2m+\frac{d}{2}} dr. \tag{247}$$

The Dominanted Convergence Theorem [13], Theorem 2.25 on p. 55) can be used to exchange the series and the integral, but first, we must check the convergence of the series. For this, we need to perform the test on the absolute series. Let $f_m = \frac{(-1)^m}{m!\Gamma(m+\frac{d}{2}+1)} r^{2m+\alpha+d-1} e^{-r^\alpha}(\frac{||\mathbf{x}||}{2})^{2m+d/2}$. The absolute value of this can be expressed as $|f_m| = \frac{1}{m!\Gamma(m+\frac{d}{2}+1)} r^{2m+\alpha+d-1} e^{-r^\alpha}(\frac{||\mathbf{x}||}{2})^{2m+d/2}$, and the integral of this from 0 to infinity is $\int_0^\infty |f_m|dr = \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{m!\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m+d/2}$. As a result, the right-hand side of (247) becomes $\sum_0^\infty\int_0^\infty |f_m|dr = \sum_0^\infty \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{m!\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m+d/2}$. We can use the Dominanted Convergence Theorem if we can show that the series converges. Let $a_m = \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{m!\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m+d/2} = |a_m|$. Using the asymptotic approximation of the Gamma function $\Gamma(x)$, $\Gamma(x+\alpha)\sim x^\alpha\Gamma(x)$ when $x\to\infty$ ([1], p.259, 6.1.39),

$$\lim_{m\to\infty}\frac{|a_{m+1}|}{|a_m|} = \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha}+\frac{2}{\alpha})}{(m+1)(m+d/2)\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}\cdot(\frac{||\mathbf{x}||}{2})^2 \tag{248}$$

$$= \lim_{m\to\infty}\frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})\cdot(\frac{d}{\alpha}+\frac{2m}{\alpha})^{2/\alpha}}{(m+1)(m+d/2)\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}\cdot(\frac{||\mathbf{x}||}{2})^2 \tag{249}$$

$$= \lim_{m\to\infty}\frac{(\frac{d}{\alpha}+\frac{2m}{\alpha})^{2/\alpha}}{(m+1)(m+d/2)}\cdot(\frac{||\mathbf{x}||}{2})^2 \text{ since } \frac{2}{\alpha}<2 \tag{250}$$

$$= 0 < 1. \tag{251}$$

By applying the Ratio test, it has been determined that the series converges. As a result, the Dominanted Convergence Theorem can be used on the numerator of I.

$$\int_0^\infty r^{\alpha-1+d/2}e^{-r^\alpha}J_{d/2}(r||\mathbf{x}||)dr = \sum_{m=0}^\infty \frac{(-1)^m}{m!\Gamma(m+\frac{d}{2}+1)}\int_0^\infty r^{2m+\alpha+d-1}e^{-r^\alpha}dr \cdot (\frac{||\mathbf{x}||}{2})^{2m+d/2} \tag{252}$$

$$= \frac{1}{\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!\Gamma(m+\frac{d}{2}+1)} \cdot \Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha}+1)(\frac{||\mathbf{x}||}{2})^{2m+d/2} \tag{253}$$

$$= \frac{1}{\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!} \cdot \frac{(\frac{d}{\alpha}+\frac{2m}{\alpha})}{(m+\frac{d}{2})}\frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m+d/2} \tag{254}$$

$$= \frac{1}{\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!} \cdot \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m+d/2} \cdot \frac{2}{\alpha}. \tag{255}$$

By changing the variable from r to $k = r^\alpha$, the integral of $r^{2m+\alpha+d-1}e^{-r^\alpha}$ over the interval $(0, \infty)$ can be estimated as $\frac{1}{\alpha}\int_0^\infty e^{-k}k^{\frac{2m+d+1}{\alpha}-1}dk = \frac{\Gamma(\frac{d+2m+\alpha}{\alpha})}{\alpha}$. The calculation of the denominator of I follows the same method.

$$\int_0^\infty r^{d/2}e^{-r^\alpha}J_{d/2-1}(r||\mathbf{x}||)dr = \sum_{m=0}^\infty \frac{(-1)^m}{m!\Gamma(m+\frac{d}{2})}\int_0^\infty r^{2m+d-1}e^{-r^\alpha}dr \cdot (\frac{||\mathbf{x}||}{2})^{2m+d/2-1} \tag{256}$$

$$= \frac{1}{\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!}\frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})} \cdot (\frac{||\mathbf{x}||}{2})^{2m+d/2-1}. \tag{257}$$

We have determined the expressions for both the numerator and denominator of the integral I. Now, we will find the linear formula for I.

$$I = -\frac{\frac{1}{\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!} \cdot \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m+d/2} \cdot \frac{2}{\alpha}}{\frac{1}{\alpha}\sum_{m=0}^\infty \frac{(-1)^m}{m!}\frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})} \cdot (\frac{||\mathbf{x}||}{2})^{2m+d/2-1}} \tag{258}$$

$$= -\frac{\sum_{m=0}^\infty \frac{(-1)^m}{m!} \cdot \frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m}}{\sum_{m=0}^\infty \frac{(-1)^m}{m!}\frac{\Gamma(\frac{d}{\alpha}+\frac{2m}{\alpha})}{\Gamma(m+\frac{d}{2})}(\frac{||\mathbf{x}||}{2})^{2m}}\frac{||\mathbf{x}||}{2}\frac{2}{\alpha} \tag{259}$$

$$= -\frac{||\mathbf{x}||}{\alpha}. \tag{260}$$

Therefore,

$$\frac{(-i\int_0^\pi\int_0^\infty e^{-ir||\mathbf{x}||\cos\theta}e^{-r^\alpha}r^{d+\alpha-2}\cos\theta(\sin\theta)^{d-2}drd\theta)}{(\int_0^\pi\int_0^\infty e^{-ir||\mathbf{x}||\cos\theta}e^{-r^\alpha}r^{d-1}(\sin\theta)^{d-2}drd\theta)}\frac{\mathbf{x}}{||\mathbf{x}||} = -\frac{||\mathbf{x}||}{\alpha} \cdot \frac{\mathbf{x}}{||\mathbf{x}||} = -\frac{\mathbf{x}}{\alpha}. \tag{261}$$

□

Let's proceed to examine the fractional score function of the marginal density function of the solution (127). According to Lemma F.1 $p_t(\mathbf{x}_t|\mathbf{x}_0) = \frac{q_\alpha(\frac{\mathbf{x}_t - a(t)\mathbf{x}_0}{\gamma(t)})}{\gamma^d(t)}$ is established. In that case, the Foruier transformation of the transition density function is $\mathcal{F}\{p\}(\mathbf{u}) = e^{-||\gamma(t)\mathbf{u}||^\alpha}e^{ia(t)\mathbf{x}_0}$. Therefore,

$$p_t(\mathbf{x}) = \int_{\mathbf{x}_t\in\mathbb{R}^d} p(\mathbf{x}|\mathbf{x}_0)p_{\text{data}}(\mathbf{x}_0)d\mathbf{x}_0 \tag{262}$$

$$= \frac{1}{(2\pi)^d}\int_{\mathbf{x}\in\mathbb{R}^d}\Big(\int_{\mathbf{u}\in\mathbb{R}^d} e^{-||\gamma(t)\mathbf{u}||^\alpha}e^{ia(t)\langle\mathbf{x}_0,\mathbf{u}\rangle}e^{-i\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{u}\Big)p_{\text{data}}(\mathbf{x}_0)d\mathbf{x}_0 \tag{263}$$

$$= \frac{1}{(2\pi)^d}\int_{\mathbf{u}\in\mathbb{R}^d} e^{-||\gamma(t)\mathbf{u}||^\alpha}e^{-i\langle\mathbf{x},\mathbf{u}\rangle}\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})d\mathbf{u}. \tag{264}$$

and

$$\Delta^{\frac{\alpha-2}{2}}\nabla p_t(\mathbf{x}) = \frac{-i}{(2\pi)^d}\int_{\mathbf{u}\in\mathbb{R}^d}\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})e^{-||\gamma(t)\mathbf{u}||^\alpha}||\mathbf{u}||^{\alpha-2}\mathbf{u}e^{-i\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{u}. \tag{265}$$

Since $-\frac{1}{\alpha(\gamma(t))^\alpha}\frac{\partial}{\partial u_i}e^{-||\gamma(t)\mathbf{u}||^\alpha} = u_i||\mathbf{u}||^{\alpha-2}e^{-||\gamma(t)\mathbf{u}||^\alpha}$ for each $i \in \{1,\ldots,d\}$,

$$[\Delta^{\frac{\alpha-2}{2}}\nabla p_t(\mathbf{x}_t)]_j = \frac{i}{(2\pi)^d\cdot\alpha(\gamma(t))^{\alpha-1}}\int_{\mathbf{u}\in\mathbb{R}^d}\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})\frac{\partial}{\partial u_j}[e^{-||\gamma(t)\mathbf{u}||^\alpha}]e^{-i\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{u} \tag{266}$$

$$= \frac{-i}{(2\pi)^d\alpha(\gamma(t))^\alpha}\int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\frac{\partial}{\partial u_j}[\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})e^{-i\langle\mathbf{x},\mathbf{u}\rangle}]d\mathbf{u} \tag{267}$$

$$= -\frac{x_j p_t(\mathbf{x})}{\alpha\gamma^{\alpha-1}(t)} - \frac{i}{(2\pi)^d\alpha(\gamma(t))^\alpha}\int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\frac{\partial}{\partial u_j}[\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})]e^{-i\langle\mathbf{x},\mathbf{u}\rangle}]d\mathbf{u}. \tag{268}$$

Thus,

$$\begin{aligned}\frac{\Delta^{\frac{\alpha-2}{2}}\nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} &= -\frac{\mathbf{x}}{\alpha(\gamma(t))^\alpha}\\ &- \frac{i}{(2\pi)^d\alpha(\gamma(t))^\alpha}\frac{\int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\nabla[\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})]e^{-i\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{u}}{p_t(\mathbf{x})}\end{aligned} \tag{269}$$

$$= -\frac{\mathbf{x}}{\alpha(\gamma(t))^\alpha} - \frac{1}{(2\pi)^d}\frac{i}{\alpha(\gamma(t))^\alpha}\int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\nabla[\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})]\frac{e^{-i\langle\mathbf{x},\mathbf{u}\rangle}}{p_t(\mathbf{x})}d\mathbf{u}. \tag{270}$$

This representation (269) also can be applied to any integrable function $p_t(\mathbf{x})$ since the normalization term cancels out in both the numerator and denoimiator. If $p_t(\mathbf{x}) = \frac{q_t(\mathbf{x})}{N}$ where $N = \int_{\mathbf{x}\in\mathbb{R}^d}q_t(\mathbf{x})d\mathbf{x}$, then $\frac{\Delta^{\frac{\alpha-2}{2}}\nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} = \frac{\Delta^{\frac{\alpha-2}{2}}\nabla q_t(\mathbf{x})}{q_t(\mathbf{x})}$.

Let's consider the function $\phi(\mathbf{x}, t)$ defined as follows:

$$\phi(\mathbf{x}, t) = -\frac{i}{(2\pi)^d}\int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\nabla[\mathcal{F}\{p_{\text{data}}\}](a(t)\mathbf{u})\frac{e^{-i\langle\mathbf{x},\mathbf{u}\rangle}}{p_t(\mathbf{x})}d\mathbf{u}. \tag{271}$$

Then the fractional score function is represeted as:

$$\frac{\Delta^{\frac{\alpha-2}{2}}\nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} = \frac{1}{\alpha\gamma^\alpha(t)}\left(a(t)\phi(\mathbf{x}, t) - \mathbf{x}\right) \tag{272}$$

Since $p_{\text{data}}(\mathbf{x})$ is an integrable function, it can be expressed as a simple function of the form $p_{\text{data}}(\mathbf{x}) = \sum_{i=1}^n a_i\chi_{E_i}(\mathbf{x})$, then the Fourier transform of p_{data}, denoted by $\mathcal{F}p_{\text{data}}$, follows the following relationship:

$$\mathcal{F}\{p_{\text{data}}\}(\mathbf{u}) = \sum_{i=1}^n a_i\int_{\mathbf{y}\in E_i}e^{i\langle\mathbf{u},\mathbf{y}\rangle}d\mathbf{y}. \tag{273}$$

By taking the gradient of equation (273),

$$\nabla[\mathcal{F}\{p_{\text{data}}\}](\mathbf{u}) = \sum_{i=1}^n a_i\int_{\mathbf{y}\in E_i}e^{i\langle\mathbf{u},\mathbf{y}\rangle}\cdot i\cdot\mathbf{y}d\mathbf{y}. \tag{274}$$

Thus,

$$\phi(\mathbf{x},t) = -\frac{i}{(2\pi)^d}\int_{\mathbf{u}\in\mathbb{R}^d} e^{-||\gamma(t)\mathbf{u}||^\alpha}\nabla[\mathcal{F}\{p_{\text{data}}\}](a(t)\mathbf{u})\frac{e^{-i\langle\mathbf{x},\mathbf{u}\rangle}}{p_t(\mathbf{x})}d\mathbf{u} \tag{275}$$

$$= \sum_{i=1}^n a_i\int_{\mathbf{y}\in E_i}\mathbf{y}\cdot\left[\frac{1}{(2\pi)^d}\int_{\mathbf{u}\in\mathbb{R}^d} e^{-||\gamma(t)\mathbf{u}||^\alpha}e^{-i\langle\mathbf{u},\mathbf{x}-a(t)\mathbf{y}\rangle}d\mathbf{u}\right]d\mathbf{y}/p_t(\mathbf{x}) \tag{276}$$

$$= \frac{1}{\gamma^d(t)}\sum_{i=1}^n a_i\int_{\mathbf{y}\in E_i}\mathbf{y}\cdot q_\alpha\left(\frac{\mathbf{x}-a(t)\mathbf{y}}{\gamma(t)}\right)d\mathbf{y}/p_t(\mathbf{x}) \tag{277}$$

$$= \frac{1}{\gamma^d(t)}\int_{y\in\mathbb{R}^d}\sum_{i=1}^n a_i\chi_{E_i}\cdot\mathbf{y}\cdot q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{y}}{\gamma(t)})d\mathbf{y}/p_t(\mathbf{x}) \tag{278}$$

$$= \frac{1}{\gamma^d(t)}\int_{y\in\mathbb{R}^d}\mathbf{y}\cdot p_{\text{data}}(\mathbf{y})q_\alpha\left(\frac{\mathbf{x}-a(t)\mathbf{y}}{\gamma(t)}\right)d\mathbf{y}/p_t(\mathbf{x}) \tag{279}$$

$$= \frac{\int_{\mathbf{x}_0\in\mathbb{R}^d}\mathbf{x}_0\cdot p_{\text{data}}(\mathbf{x}_0)q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d}p_{\text{data}}(\mathbf{x_0})q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}. \tag{280}$$

Now, let's consider the general cases where p_{data} is not a simple function but an integrable function. In this case, since the integrable function is measurable, we can find a sequence of simple functions, denoted by $\{p_n\}_{n=1}^\infty$, such that as n approaches infinity, $p_n \to p_{\text{data}}$ almost everywhere and under the L_1 norm with $p_n \leq p_{n+1}$ for all $n \in \mathbb{N}$.

For each p_n, the corresponding $\phi_n(\mathbf{x},t)$ satisfies

$$\phi_n(\mathbf{x},t) = \frac{\int_{\mathbf{x}_0\in\mathbb{R}^d}\mathbf{x}_0\cdot p_n(\mathbf{x}_0)q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d}p_n(\mathbf{x_0})q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}. \tag{281}$$

By the Dominant Convergence Theorem,

$$\lim_{n\to\infty}\phi_n(\mathbf{x},t) = \lim_{n\to\infty}\left(\frac{\int_{\mathbf{x}_0\in\mathbb{R}^d}\mathbf{x}_0\cdot p_n(\mathbf{x}_0)q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d}p_n(\mathbf{x_0})q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}\right) \tag{282}$$

$$= \frac{\lim_{n\to\infty}\int_{\mathbf{x}_0\in\mathbb{R}^d}\mathbf{x}_0\cdot p_n(\mathbf{x}_0)q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}{\lim_{n\to\infty}\int_{\mathbf{x}_0\in\mathbb{R}^d}p_n(\mathbf{x_0})q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0} \tag{283}$$

$$= \frac{\int_{\mathbf{x}_0\in\mathbb{R}^d}\mathbf{x}_0\cdot\lim_{n\to\infty}p_n(\mathbf{x}_0)q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d}\lim_{n\to\infty}p_n(\mathbf{x_0})q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0} \tag{284}$$

$$= \frac{\int_{\mathbf{x}_0\in\mathbb{R}^d}\mathbf{x}_0\cdot p_{\text{data}}(\mathbf{x}_0)q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d}p_{\text{data}}(\mathbf{x_0})q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)})d\mathbf{x}_0}. \tag{285}$$

We need to show that $\lim_{n\to\infty}\phi_n(\mathbf{x},t) = \phi(\mathbf{x},t)$. Since p_n converges to p_{data} in the L_1 norm, it follows that $\mathcal{F}\{p_n\}$ converges to $\mathcal{F}\{p_{\text{data}}\}$ almost everywhere. By the Dominant Convergence Theorem, we get

$$\lim_{n\to\infty}\int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\mathcal{F}\{p_n\}(a(t)\mathbf{u})e^{-i\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{u} = \int_{\mathbf{u}\in\mathbb{R}^d}e^{-||\gamma(t)\mathbf{u}||^\alpha}\mathcal{F}\{p_{\text{data}}\}(a(t)\mathbf{u})e^{-i\langle\mathbf{x},\mathbf{u}\rangle}d\mathbf{u}. \tag{286}$$

Since $|[\mathbf{y}\cdot p_n(\mathbf{y})e^{i\langle\mathbf{y},\mathbf{u}\rangle}]_i| \leq \mathbf{y}_i\cdot p_n(\mathbf{y})$ and $\lim_{n\to\infty}\int_{\mathbf{y}\in\mathbb{R}^d}\mathbf{y}\cdot p_n(\mathbf{y})d\mathbf{y} = \int_{\mathbf{y}\in\mathbb{R}^d}\mathbf{y}\cdot p_{\text{data}}(\mathbf{y})d\mathbf{y}$ with using the Monotone Convergence Theorem for each component $i \in \{1,\ldots,n\}$, we induce

$$\lim_{n\to\infty}\nabla[\mathcal{F}\{p_n\}](\mathbf{u}) = \lim_{n\to\infty} i\int_{\mathbf{y}\in\mathbb{R}^d}\mathbf{y}\cdot p_n(\mathbf{y})e^{i\langle\mathbf{y},\mathbf{u}\rangle}d\mathbf{y} \tag{287}$$

$$= i\int_{\mathbf{y}\in\mathbb{R}^d}\mathbf{y}\cdot p_{\text{data}}(\mathbf{y})e^{i\langle\mathbf{y},\mathbf{u}\rangle}d\mathbf{y} \tag{288}$$

$$= \nabla[\mathcal{F}\{p_{\text{data}}\}](\mathbf{u}). \tag{289}$$

By the generalized Dominant Convergence Theorem [12][Excercise 20 on p 59],

$$\lim_{n\to\infty} i \int_{\mathbf{u}\in\mathbb{R}^d} e^{-||\gamma(t)\mathbf{u}||^\alpha} \nabla[\mathcal{F}\{p_n\}](a(t)\mathbf{u}) e^{-i\langle \mathbf{x},\mathbf{u}\rangle} d\mathbf{u} \tag{290}$$

$$= \int_{\mathbf{u}\in\mathbb{R}^d} e^{-||\gamma(t)\mathbf{u}||^\alpha} \nabla[\mathcal{F}\{p_{\text{data}}\}](a(t)\mathbf{u}) e^{-i\langle \mathbf{x},\mathbf{u}\rangle} d\mathbf{u}. \tag{291}$$

From this result, we can get $\lim_{n\to\infty} \phi_n(\mathbf{x},t) = \phi(\mathbf{x},t)$. Thus,

$$\phi(\mathbf{x},t) = \frac{\int_{\mathbf{x}_0\in\mathbb{R}^d} \mathbf{x}_0 \cdot p_{\text{data}}(\mathbf{x}_0) q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)}) d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)}) d\mathbf{x}_0}. \tag{292}$$

Combining (292) and (272) , we conclude

$$\frac{\Delta^{\frac{\alpha-2}{2}} \nabla p_t(\mathbf{x})}{p_t(\mathbf{x})} = \frac{1}{\alpha\gamma^\alpha(t)} \left(a(t) \frac{\int_{\mathbf{x}_0\in\mathbb{R}^d} \mathbf{x}_0 \cdot p_{\text{data}}(\mathbf{x}_0) q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)}) d\mathbf{x}_0}{\int_{\mathbf{x}_0\in\mathbb{R}^d} p_{\text{data}}(\mathbf{x}_0) q_\alpha(\frac{\mathbf{x}-a(t)\mathbf{x}_0}{\gamma(t)}) d\mathbf{x}_0} - \mathbf{x} \right). \tag{293}$$

Therefore, we obtain the integral representation of the fractional score function from the perspective of the data distribution. This representation allows us to precisely understand the direction indicated by the fractional score function.

G Experiment Detail

G.1 Model architecture

Our model uses U-Net [31] following DDPM [15], which replaces weight normalization [30] with group normalization [37] for simple implementation. We set the hidden layer dimension of our model suitable for the dataset, such that CIFAR10 (32×32) is [128, 256, 256, 256], CelebA (64×64) is [128, 256, 256, 256, 1024], and both CelebA-HQ and LSUN (256×256) are [128, 256, 256, 256, 1024, 1024], but fix the number of residual blocks with 2 in each resolution level, and add self-attention block only in 16×16 resolution level. These setups are also following DDPM([15], [39]). We also use extended model architecture, NCSN++ (deep) following [39] to improve sample quality on CIFAR10. It has 8 residual blocks in each resolution level.

The exponential moving average(EMA) rate set to be 0.9999 for CIFAR10 and CelebA, 0.999 for CelebA-HQ and LSUN, because it works better for models trained with VP perturbations.

Continuous diffusion time $t \in [0, 1)$ is injected into the model through Transformer sinusoidal position embedding [34] after adding with 0.00001, and we use swish function for all datasets as the activation function.

We train our CIFAR10 model for 200k steps with batch size 256, CelebA model for 600k steps with batch size 256, and CelebA-HQ, LSUN model for 1.3M steps with batch size 16. All training and experiments are conducted on NVIDIA A100 GPU, NVIDIA V100 GPU and NVIDIA GeForce RTX 3090, and we tune the batch size for sampling adjusted for computation resources.

G.2 Noise scheduling

We use cosine noise schedule as [25], and limit the maximum time $T = 0.9946$ for all datasets. We consider the variance preserving type with modification which is suitable for Lévy noise, so it can have different value according to α.

We can get $\beta(t)$ and marginal $\log \alpha_t$ from

$$\beta(t) = \frac{\alpha}{1+s}\frac{\pi}{2} \tan\left(\frac{t+s}{1+s}\frac{\pi}{2}\right) \tag{294}$$

$$\log a(t) = \log\left(\cos\left(\frac{\pi}{2}\cdot\frac{t+s}{1+s}\right)\right) - \log\left(\cos\left(\frac{\pi}{2}\cdot\frac{s}{1+s}\right)\right) \tag{295}$$

where $s = 0.008$, following [25]. The scale parameter can be computed by the relation $\gamma(t) = (1 - a^\alpha(t))^{\frac{1}{\alpha}}$.

G.3 Smooth L1 Loss

L_2 loss has been widely used for stable training because it is differentiable in entire range, but it is sensitive to outliers due to its square term. Therefore, we choose smooth L1 Loss for stable training instead of L2 Loss, since Lévy noise has much more outliers than Gaussian noise. For smooth L_1 Loss, as β varies, the L_1 segment of the loss has a constant slope of 1. As the β approaches to zero, smooth L_1 Loss converges to L_1 Loss, but as the β approaches to ∞, it converges to a constant 0 loss. We evaluate sample quality according to β in the CIFAR10 dataset with FID, and Precision-Recall at NFE 500.

β	FID	Precision	Recall
0.5	3.90	0.752	**0.688**
1.0	**3.37**	0.752	**0.688**
2.0	4.02	0.754	0.686
3.0	4.02	**0.756**	0.687
4.0	3.90	0.753	**0.688**

Table 7: β selection table.

G.4 Clamping / Threshold

Different from Gaussian distribution, α-stable distribution can have large-scale noise at lower α values, which leads to sample quality degradation. To prevent this problem, we use 2 kinds of heuristics in the training and sampling phase.

α-stable distributions can be simulated by generating samples of 1-dimensional α-stable distribution and gaussian distribution. If $0 < \alpha < 2$, $A \sim \mathbf{S}\left(\alpha/2, 1, 2(\cos\pi\alpha/4)^{2/\alpha}, 0\right)$ and $\mathbf{G} \sim N(0, Q)$, then α-stable distribution follows this relation [25]:

$$\mathbf{X} = A^{1/2}\mathbf{G} \tag{296}$$

A will actually increase up to 10000 or more if no action is taken. However, excessive noise can make sampling poor since the pixel values of image data range from 0 to 1. To address this issue, we introduce the clamping method, $A \leftarrow \text{clamp}(A, -20, 20)$. Thus, we generate samples of α-stable distribution as follow:

$$\mathbf{X} = \text{clamp}(A, -20, 20)^{1/2}\mathbf{G} \tag{297}$$

Another heuristic is the clamp threshold, which is re-normalization in the sampling phase. The clamp threshold is a heuristic that aids in the convergence of a sequence during sampling. At each reverse sampling step, the normalization of the norm is done using that threshold if the norm of the sample exceeds the threshold. It can be observed that the generated images vary slightly and there are differences in quality depending on the clamp threshold.

Figure 9: The effect of the clamp threshold on CelebA-HQ.

G.5 Evaluation metrics

G.5.1 FID (Fréchet Inception Distance) score and Recall

To evaluate generated sample quality, we choose traditional metric, FID score [14] where a lower score means better sample quality. Moreover, we use Recall [20] , where a higher score is better, as metric to evaluate sample diversity, since Recall measures the fraction of the training data manifold covered by the generator.

After computing both mean/variance of distributions in the training dataset and generating the same number of samples as training dataset (CIFAR10: 50k, CelebA-HQ: 30k). In CelebA and LSUN, which have large training dataset, we randomly choose 50k samples from training dataset 5 iteration. Then we calculate the distance between two distributions as FID score and the probability that generated samples falls within the support of distribution of training dataset by using the pre-trained Inception-V3 model.

G.5.2 Likelihood computation

We can compute the exact likelihood on any input data in the same way as [39]. By replacing the score $\nabla_x \log p_t(X_t)$ with score model $S_\theta(X_t, t)$, we can rewrite (164) as

$$d\overrightarrow{X}_t = \underbrace{\left(b(t, \overrightarrow{X}_t) - S_\theta(\overrightarrow{X}_t, t)\sigma_L^\alpha(t)\right)}_{=:\tilde{f}_\theta(\overrightarrow{X}_t, t)} dt. \tag{298}$$

Then we can compute the log-likelihood of $p_0(X_0)$ such that

$$\log p_0(\overrightarrow{X}_0) = \log p_T(\overrightarrow{X}_T) + \int_0^T \nabla \cdot \tilde{f}_\theta(\overrightarrow{X}_t, t)dt. \tag{299}$$

where $\overleftarrow{X}_T$ is noise mapping to $\overleftarrow{X}_0$ which can be obtained by solving the probability ODE in (298) with ODE solver. Because of the expensive computation of $\nabla \cdot \tilde{f}_\theta(\overrightarrow{X}_t, t)$, we estimate it by using the Skilling-Hutchinson trace estimator [38][16].

To solve the integral term, we choose the RK45 ODE solver [10] which can be used as `solve-ivp` function in `scipy.integrate` library. We also set parameters `atol=1e-5`, `rtol=1e-5`. We use a test dataset applied uniform dequantization, and take the average of the bits/dim values over 5 repeats for exact likelihood computation. By changing initial time t_0 of integral $\int_{t_0}^T \nabla \cdot \tilde{f}_\theta(\overrightarrow{X}_t, t)dt$ after adding 0.001, we compute bits/dim with varied number of function evaluations(NFE).

G.6 Wall clock time

Previously, Lévy sampling supported by `scipy.stats` library was much slower than pytorch based Gaussian noise sampling since it is a numpy based method. Therefore, we migrate it to pytorch version to raise the speed of Lévy sampling upto Gaussian noise. As you can see in below table, our wall clock time per NFE become as fast as Gaussian, and we show that the halved NFE leads to a reduction in the actual sampling speed.

Model	wall clock time/NFE (sec)	NFE	total wall clock time (sec)	FID↓
CIFAR10 (32×32)				
DDPM cont. (VP) [35]	0.0261	1000	26.0675	3.24
LIM-DDPM cont. (Ours)	0.0285	500	14.2709	3.37
CelebA (64×64)				
DDPM cont. (VP) [35]	0.0800	1000	79.9665	3.21
LIM-DDPM cont. (Ours)	0.0763	500	38.1757	1.57

Table 8: Wall clock time on CIFAR10 (32×32) and CelebA (64×64).

G.7 Imputation: Additional samples

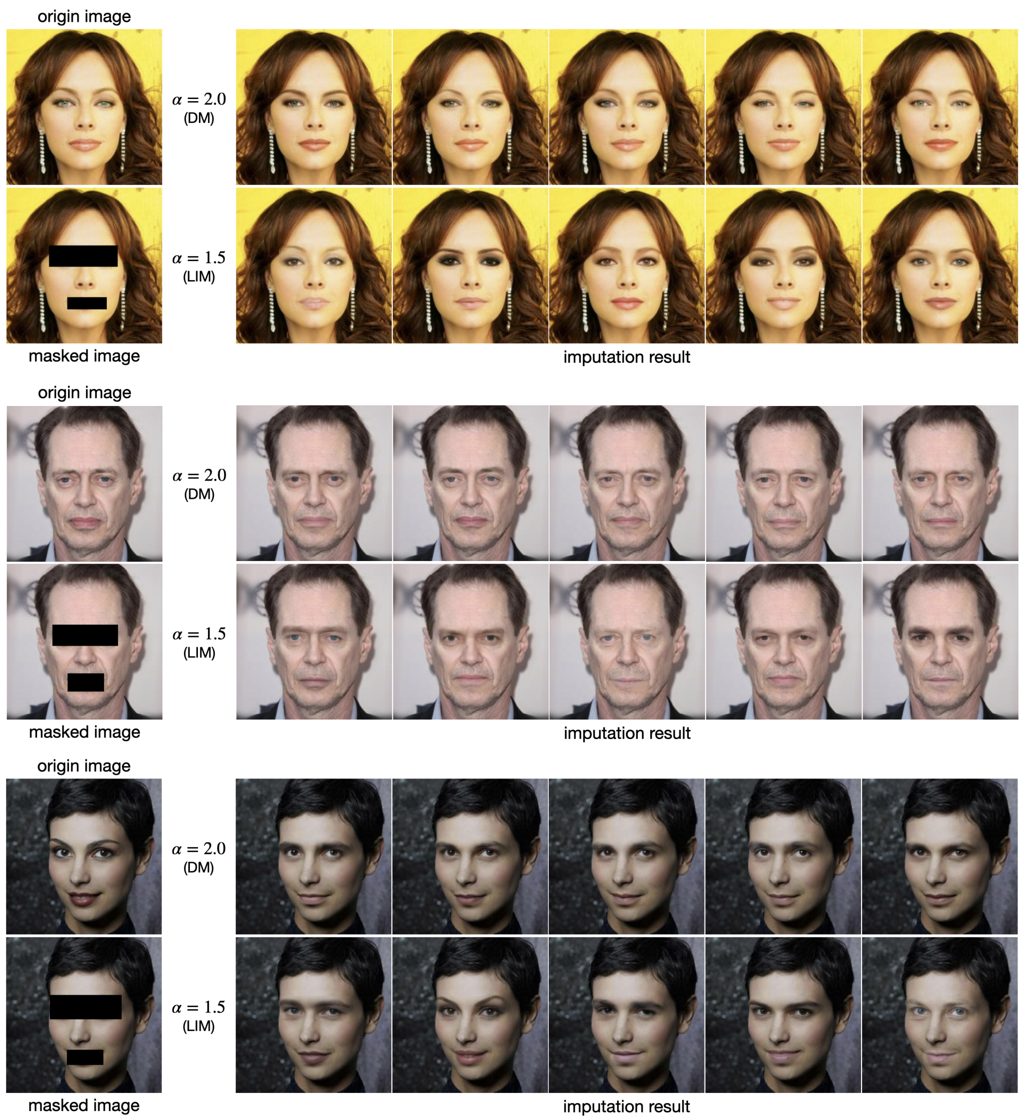

Figure 10: Additional imputation results using CelebA-HQ dataset. Diffusion model(DM) generates only similar eyes and lips, while LIM generates much more diverse range of shape and color of eyes and lips.

G.8 Additional samples on high dimensional datasets

We also evaluate our model on high dimensional datasets, such as CelebA-HQ, LSUN-Bedroom, and LSUN-Church (256×256), which are reported in below figures.

Figure 11: Additional samples for CelebA-HQ (256×256) at NFE 200.

Figure 12: Additional samples for LSUN-Bedroom (256×256) at NFE 200.

Figure 13: Additional samples for LSUN-Church (256×256) at NFE 200.