A. Proofs

Theorem 2. *Let $p \in \mathcal{L}$ and $q \in \mathcal{H}$ such that $q = T_{\#}p$, where T is a diffeomorphism. Then, for all $M > 0$ and all $z_0 > 0$ there is $z > z_0$, such that $T'(z) > M$. Conversely, if T is a Lipschitz-continuous map & $p \in \mathcal{L}$, then, $T_{\#}p \in \mathcal{L}$.*

Proof. We prove this by contradiction. Assume on the contrary that there exists a diffeomorphism $T : \mathbb{R} \to \mathbb{R}$, such that $q = T_{\#}p$ and $\exists\, M > 0,\ z_0 > 0$, such that $\forall z > z_0,\ T'(z) \leq M$. Because T is a univariate diffeomorphism, it is a strictly monotonic function. Without loss of generality, consider a strictly increasing function T, such that $0 < T'(z) \leq M$ for all $z > z_0$. Since, $p \in \mathcal{L}$, we have

$$\int_{\mathsf{Z}} e^{\lambda_1 z} p(z)\,\mathrm{d}z < \infty, \qquad \text{for some } \lambda_1 > 0 \tag{4}$$

Furthermore, since $q \in \mathcal{H}$, we have

$$\int_{\mathsf{X}} e^{\lambda x} q(x)\,\mathrm{d}x = \infty, \qquad \forall\, \lambda > 0 \tag{5}$$

$$\implies \quad \int_{\mathsf{Z}} e^{\lambda T(z)} p(z)\,\mathrm{d}z = \infty, \qquad \forall \lambda > 0, \qquad [\because \text{change of variables}] \tag{6}$$

Split the domain Z into : $\mathsf{Z}_+ = \mathsf{Z} \cap \{z \geq 0\}$ and $\mathsf{Z}_- = \mathsf{Z} \cap \{z < 0\}$. The integral over the negative part trivially converges since:

$$\int_{\mathsf{Z}_-} e^{\lambda T(z)} p(z)\,\mathrm{d}z \leq \int_{\mathsf{Z}_-} e^{\lambda T(0)} p(z)\,\mathrm{d}z \leq e^{\lambda T(0)},$$

where we used that T is increasing. Next, we split the integral over Z_+ into two parts: integral from 0 to z_0 and from z_0 to ∞. The first integral is clearly finite since it is an integral of a continuous function over a compact set in $\mathbb{R}$. Thereafter, integrating the inequality on a slope, we get $\forall\, z > z_0$: $T(z) \leq Mz + T(z_0)$. Then:

$$\int_{z_0}^{\infty} e^{\lambda T(z)} p(z)\,\mathrm{d}z \leq e^{T(z_0)} \int_{z_0}^{\infty} e^{\lambda M z} p(z)\,\mathrm{d}z. \qquad \forall\, \lambda > 0 \tag{7}$$

$$\leq e^{T(z_0)} \int_{\mathsf{Z}} e^{\lambda M z} p(z)\,\mathrm{d}z, \qquad \forall\, \lambda > 0 \tag{8}$$

Choose λ such that $\lambda M = \lambda_1$. Then, the integral must be finite because p is light-tailed leading to the desired contradiction. □

Proposition 1. *Let p be a density with $fQ_p \sim (1-u)^{\alpha}$ as $u \to 1^-$. Then, $0 < \alpha < 1$ iff $supp(p) = [a, b]$ where $b < \infty$ i.e. p has a support bounded from above.*

Proof. Let $0 < \alpha < 1$.

$$\begin{aligned} fQ_p(u) \sim (1-u)^{\alpha} &\iff Q(u) \sim (1-u)^{\delta} + c, \qquad 0 < \delta < 1,\ c \text{ is a finite constant} \\ &\iff \lim_{u \to 1^-} Q(u) \to c \\ &\iff F_p^{-1}(1) = c \iff p \text{ has support bounded from above.} \end{aligned}$$

A similar argument proves the reverse direction. □

Proposition 3. *Let p be a distribution with $Q_p(u) \sim (1-u)^{-\gamma}$ as $u \to 1^-$. Then, $\int_{z_0}^{\infty} z^{\omega} p(z) dz$ exists and is finite for some z_0 iff $\omega < \frac{1}{\gamma}$.*

Proof.

$$\int_{z_0}^{\infty} z^{\omega} p(z) dz \text{ exists} \iff \int_{u_0}^{1} Q_p^{\omega}(u)\,\mathrm{d}u \text{ exists for some } u_0 > 0 \tag{9}$$

$$\iff \int_{u_0}^{1-\epsilon} Q_p^{\omega}(u)\,\mathrm{d}u \text{ exists} \quad \& \quad \int_{1-\epsilon}^{1} Q_p^{\omega}(u)\,\mathrm{d}u \text{ exists} \tag{10}$$

The first integral is finite because the integrand is non-singular. For the second integrand, we can use the asymptotic behaviour of the quantile function by choosing ϵ very close to 1. Subsequently, the integral exists and converges if and only if $1 - \omega\gamma > 0 \iff \omega < \frac{1}{\gamma}$. □

Proposition 4. *Let p be a $\omega_p^{-1}-$heavy distribution, q be a $\omega_q^{-1}-$heavy distribution and T be a diffeomorphism such that $q := T_{\#}p$. Then for small $\epsilon > 0$, $T(z) = o(|z|^{\omega_p/\omega_q - \epsilon})$.*

Proof. The integral

$$\mathbb{E}_q[|x|^{\omega_q - \epsilon}] = \int_{\mathbb{R}} |x|^{\omega_q - \epsilon} q(x)\, \mathrm{d}x \tag{11}$$

$$= \int_{\mathbb{R}} |T(z)|^{\omega_q - \epsilon} p(z)\, \mathrm{d}z \tag{12}$$

converges for $0 < \epsilon < \omega_q$, because q is ω_q^{-1}-heavy. Because T is a univariate diffeomorphism, it is a strictly monotone function. Without loss of generality, let us consider T to be positive increasing function and investigate the right asymptotic. Consider the function $T(z)^{\omega_q-\epsilon}/z^{\omega_p}$ for big positive z. Assume there is a sequence $\{z_i\}_{i=1}^{\infty}$, such that $\lim_i z_i = +\infty$ and the sequence $T(z_i)^{\omega_q-\epsilon}/z_i^{\omega_p}$ does not converge to zero. In other words, there exists $a > 0$, such that for any $N > 0$ there exists $z_j > N$, such that $T(z_j)^{\omega_q-\epsilon}/z_j^{\omega_p} > a$. Let us work with this infinite sub-sequence $\{z_j\}$. Because $T(z)$ is increasing function, we can estimate its integral from the left by its left Riemannian sum with respect to the sequence of points $\{z_j\}$:

$$\int_N^{\infty} T(z)^{\omega_q-\epsilon} p(z)\, \mathrm{d}z \geq \sum_j T(z_j)^{\omega_q-\epsilon} p(\Delta z_j) > a \sum_j z_j^{\omega_p} p(\Delta z_j).$$

Since, p is $\omega_p^{-1}-$heavy, the series on the right hand side diverges as a left Riemannian sum of a divergent integral. But this contradicts to the convergence of the integral on the left hand side. Hence, our assumption was wrong and for all sequences $\{z_i\}$ we have: $\lim T(z_i)^{\omega_q-\epsilon}/z_i^{\omega_p} = 0$. Hence, $|T(z)|^{\omega_q-\epsilon} = o(|z|^{\omega_p})$ which leads to the desired result that $|T(z)| = o(|z|^{\omega_p/\omega_q-\epsilon})$. □

Proposition 5. *Under the same assumptions as in Lemma 2 (App.B), if $\mathsf{X} \sim \varepsilon_d(0, \mathbf{I}, F_R)$ is ω^{-1}-heavy, then the conditional distribution of $\mathsf{X}_2|(\mathsf{X}_1 = \mathbf{x}_1)$ is $(\omega + d_1)^{-1}$-heavy where $\mathsf{X}_1 \subseteq \mathbb{R}^{d_1}$.*

Proof. The density function of the conditional $p(x|X_1 = \mathbf{x}_1)$ is proportional to $g_R((x-\mu^*)^T\Sigma^{*-1}(x-\mu^*))$, where $x \in \mathbb{R}^{d_2}$ and g_R is the same function as for the distribution of X (see (Cambanis et al., 1981)). Then, because it is a d_2-dimensional elliptical distribution, it is α-heavy iff $\mu_l = \int_0^{\infty} r^{l+d_2-1} g_R(r^2) dr < \infty$ for all $0 < l < \alpha$. It is given that X is ω^{-1}-heavy, which is equivalent to $\int_0^{\infty} r^{l+d-1} g_R(r^2) dr < \infty$, $\forall 0 < l < \omega$. Because $d = d_1 + d_2$, one gets that $\int_0^{\infty} r^{\tilde{l}+d_2-1} g_R(r^2) dr < \infty$, $\forall\ 0 < \tilde{l} < \omega + d_1$, hence $\mathsf{X}_2|\mathsf{X}_1 = \mathbf{x}_1$ is $(\omega + d_1)^{-1}$-heavy. □

Theorem 3. *Let $\mathsf{Z} \subseteq \mathbb{R}^d$ be a random variable with density function p that is light-tailed and $\mathsf{X} \subseteq \mathbb{R}^d$ be a target random variable with density function q that is heavy-tailed. Let $\mathbf{T} : \mathsf{Z} \to \mathsf{X}$ be such that $q = \mathbf{T}_{\#}p$, then $\mathbf{T}$ cannot be a Lipschitz function.*

Proof. On the contrary, assume that $\mathbf{T}$ is $M-$Lipschitz. Since $q(\mathbf{x})$ is heavy tailed we have that $\forall\ \lambda > 0$

$$\int_{\mathbf{x}} e^{\lambda\|\mathbf{x}\|} q(\mathbf{x})\, \mathrm{d}\mathbf{x} = \infty \tag{13}$$

$$\implies \int_{\mathbf{z}} e^{\lambda\|T(\mathbf{z})\|} p(\mathbf{z})\, \mathrm{d}\mathbf{z} = \infty \tag{14}$$

$$\int_{\mathbf{z}} e^{\lambda\|T(\mathbf{z})\|} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \leq \int_{\mathbf{z}} e^{\lambda M\|\mathbf{z}\|} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \tag{15}$$

Since $p(\mathbf{z})$ is light-tailed there exists a $\lambda > 0$ such that the right hand side of the equation above is finite. This gives us the required contradiction. □

Corollary 3. *Under the same set-up as in Theorem 3, there exists an index $i \in [d]$ such that $\|\nabla_{\mathbf{z}} T_i\|$ is unbounded.*

Proof. We will prove this using contradiction; assume that $\forall (i,j) \in [d]^2,\ \frac{\partial T_i}{\partial z_j} \leq M < \infty$. Assume for simplicity that $\mathbf{T}(0) = c < \infty$. Therefore, we have

$$T_i(\mathbf{z}) - T_i(0) = \int_{\mathbf{r}(0\to\mathbf{z}):0}^{\mathbf{z}} \nabla T_i \cdot \mathrm{d}\vec{r} \tag{16}$$

$$\implies \qquad |T_i(\mathbf{z}) - T_i(0)| \leq M \sum_{i=1}^{d} |z_i| \tag{17}$$

Since, $q(\mathbf{z})$ is heavy tailed, $\exists\ \mathbf{u} \in \mathcal{B}_1$ such that $\forall\ \kappa > 0$

$$\int_{\mathbb{R}^d} e^{\kappa \mathbf{u}^T \mathbf{x}} q(\mathbf{x})\, \mathrm{d}\mathbf{x} = \infty \tag{18}$$

$$i.e. \qquad \int_{\mathbb{R}^d} e^{\kappa \mathbf{u}^T \mathbf{T}(\mathbf{z})} p(\mathbf{z})\, \mathrm{d}\mathbf{z} = \infty \qquad [\text{ change of variables }] \tag{19}$$

We have

$$\int_{\mathbb{R}^d} e^{\kappa \mathbf{u}^T \mathbf{T}(\mathbf{z})} p(\mathbf{z})\, \mathrm{d}\mathbf{z} = \int_{\mathbb{R}^d} \prod_{i=1}^{d} e^{\kappa u_i T_i(\mathbf{z})} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \tag{20}$$

$$\leq C \int_{\mathbb{R}^d} \prod_{i=1}^{d} e^{\kappa |u_i||T_i(\mathbf{z})|} p(\mathbf{z})\, \mathrm{d}\mathbf{z}, \qquad [\, C = \text{ finite constant} \,] \tag{21}$$

$$\leq C \int_{\mathbb{R}^d} \prod_{i=1}^{d} e^{\kappa M \sum_{i=1}^{d} |u_i||z_i|} p(\mathbf{z})\, \mathrm{d}\mathbf{z}, \qquad [\, u = \max |u_i| \,] \tag{22}$$

$$\leq \tilde{C} \int_{\mathbb{R}^d} e^{\kappa M \sum_{i=1}^{d} |u_i||z_i|} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \tag{23}$$

$$= \tilde{C} \int_{\mathbb{R}^d} e^{\kappa M \sum_{i=1}^{d} \text{sign}(z_i)|u_i| z_i} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \tag{24}$$

Partition $\mathbb{R}^d$ into 2^d sets $U_k,\ k \in [2^d]$, i.e. $\mathbb{R}_d = \cup_{k=1}^{2^d} U_k$ such that if $\mathbf{a} = (a_1, a_2, \cdots, a_d) \in U_i$, and $\mathbf{b} = (b_1, b_2, \cdots, b_d) \in U_j,\ i \neq j$, then there exists at least one index $m \in [d]$ such that $\text{sign}(a_m) \neq \text{sign}(b_m)$. Subsequently, we can rewrite the integral above as

$$\tilde{C} \int_{\mathbb{R}^d} e^{\kappa M \sum_{i=1}^{d} \text{sign}(z_i)|u_i| z_i} p(\mathbf{z})\, \mathrm{d}\mathbf{z} = \tilde{C} \sum_{k=1}^{2^d} \int_{U_k} e^{\kappa M \sum_{i=1}^{d} \text{sign}(z_i)|u_i| z_i} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \tag{25}$$

$$= \tilde{C} \sum_{k=1}^{2^d} \int_{U_k} e^{\kappa M \mathbf{w}^T \mathbf{z}} p(\mathbf{z})\, \mathrm{d}\mathbf{z}, \qquad w_i = \text{sign}(z_i) \cdot |u_i| \tag{26}$$

$$\tag{27}$$

We will prove that each integral over the set U_k is finite.

$$\int_{U_k} e^{\kappa M \mathbf{w}^T \mathbf{z}} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \ \leq \int_{\mathbb{R}^d} e^{\kappa M \mathbf{w}^T \mathbf{z}} p(\mathbf{z})\, \mathrm{d}\mathbf{z} \tag{28}$$

Since $p(\mathbf{z})$ is light-tailed, we know that for any $\mathbf{u} \in \mathcal{B}_1$, there exists a $\lambda > 0$ such that $\int_{\mathbb{R}^d} e^{\lambda \mathbf{u}^T \mathbf{z}} p(\mathbf{z})\, \mathrm{d}\mathbf{z} < \infty$. Choose any $\mathbf{u} \in \mathcal{B}_1$, then for $\lambda = \kappa M / \|\mathbf{w}\|$ we have that the above integral is finite. This directly implies that

$$\sum_{k=1}^{2^d} \int_{U_k} e^{\kappa M \mathbf{w}^T \mathbf{z}} p(\mathbf{z})\, \mathrm{d}\mathbf{z} < \infty \tag{29}$$

Hence, we have our contradiction. □

Proposition 6. *Let $\mathsf{Z} \sim \varepsilon_d(0, \mathbf{I}, F_S)$ and $\mathsf{X} \sim \varepsilon_d(0, \mathbf{I}, F_R)$ have densities p and q respectively where F_R is heavier tailed than F_S. If $\mathbf{T} : \mathsf{Z} \to \mathsf{X}$ is an increasing triangular map such that $q := \mathbf{T}_{\#}p$, then all diagonal entries of $\nabla \mathbf{T}$ and $\mathsf{det}|\nabla \mathbf{T}|$ are unbounded.*

Proof. We need to show that

$$\lim_{z_j \to \infty} \frac{\partial T_{jj}}{\partial z_j} = \lim_{z_j \to \infty} \frac{fQ_{p,j|<j}}{fQ_{q,j|<j}} \to \infty, \qquad \forall\, j \in [d] \tag{30}$$

Thus, all we need to show is that the generating variate R^* of the conditional distribution for the target is heavier than the generating variate S^* of the conditional distribution of the source. From §3, we know that the tail exponent in the asymptotics of the density quantile function characterize the degree of heaviness. Furthermore, we also know that asymptotical behaviour of the density quantile function is directly related to the asymptotical behaviour of the density function since if f is a density function, the cdf is given by $F(x) = \int f(x)\,\mathrm{d}x$, the quantile function therefore is $Q = F^{-1}$ and the density quantile function is the reciprocal of the derivative of the quantile function i.e. $fQ = 1/Q'$. Hence, we need to ensure that asymtotically, the density of R^* is heavier than the density of S^*. Using the result of the cdf of a conditional distribution as given by Eq.(15) in (Cambanis et al., 1981) we have that asymptotically

$$f_{R^*}(x) = Cx^{d_1 - d} f_R(x) \tag{31}$$

where d_1 is the dimension of the partition that is being conditioned upon. Since, R is heavier tailed than S, we have that R^* is heavier tailed than S^* for all the conditional distributions. □

Theorem 4. *Let p be a light-tailed density and $\mathbf{T}$ be a triangular transformation such that $T_j(z_j;\ z_{<j}) = \sigma_j \cdot z_j + \mu_j$. If, $\sigma_j(z_{<j})$ is bounded above and $\mu_j(z_{<j})$ is Lipschitz then the target density $q := \mathbf{T}_{\#}p$ is light-tailed.*

Proof. Here, we will prove the result in two-dimensions and the higher-dimensional proof will follow directly. Following the definition of class $\mathcal{H}$ and $\mathcal{L}$ as given in the beginning of Section 3, we will show that for all direction vectors $\mathbf{v} \in \mathcal{B}$ where $\mathcal{B} := \{\mathbf{v} : \|\mathbf{v}\| = 1\}$, the univariate random variable $\mathbf{v}^T\mathbf{x} \in \mathcal{L}$ i.e. there is no direction on the hyper-sphere where the marginal distribution of the push-forward random variable is heavy-tailed.

$$\begin{aligned}
\int_{\mathbf{x}} \mathsf{exp}(\lambda \cdot \mathbf{v}^T\mathbf{x}) q(\mathbf{x})\,\mathrm{d}\mathbf{x} &= \int_{\mathbf{z}} \mathsf{exp}(\lambda \cdot \mathbf{v}^T\mathbf{T}(\mathbf{z})) p(\mathbf{z})\,\mathrm{d}\mathbf{z} \\
&= \int_{\mathbf{z}} \mathsf{exp}(\lambda v_1 z_1 + \lambda v_2 \cdot \sigma \cdot z_2 + \lambda v_2 \cdot \mu) p(\mathbf{z})\,\mathrm{d}\mathbf{z} \\
&\leq \int_{\mathbf{z}} \mathsf{exp}(\lambda v_1 z_1 + \lambda v_2 \cdot B \cdot z_2 + \lambda v_2 \cdot M \cdot z_1) p(\mathbf{z})\,\mathrm{d}\mathbf{z} \\
&= \int_{\mathbf{z}} \mathsf{exp}(\tilde{\lambda} \cdot \mathbf{u}^T\mathbf{z}) p(\mathbf{z})\,\mathrm{d}\mathbf{z} < \infty, \qquad \forall\, \tilde{\lambda} > 0, \forall \mathbf{u} \in \mathcal{B}
\end{aligned}$$

where B is the upper bound of $\sigma(\cdot)$, M is the Lipschitz constant of $\mu(\cdot)$ and the final inequality follows from the fact that $p(\mathbf{z})$ is a light-tailed distribution. □

B. Useful Results, Figures, and Examples

Example 2. *Let $p \sim \mathcal{N}(0, 1)$ and $q \sim t_1(0, 1)$. Then, T such that $q := T_{\#}p$ is given by:*

$$\begin{aligned}
T(z) &= G^{-1} \circ F = \tan\Big(\frac{\pi}{2}\,\mathrm{erf}(\frac{z}{\sqrt{2}})\Big) \\
\&, \quad T'(z) &= \sqrt{\pi} e^{-\frac{z^2}{2}} \sec^2\Big(\frac{\pi}{2}\,\mathrm{erf}(\frac{z}{\sqrt{2}})\Big)
\end{aligned}$$

where $\mathrm{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2}\,\mathrm{d}s$ *is the error function. Furthermore,* $fQ_p(u) \sim (1-u)\big(-2\log(1-u)\big)^{1/2}$ *and* $fQ_q(u) \sim (1-u)^2$ *and hence,* $\lim_{z\to\infty} T'(z) = \lim_{u \to 1^-} (1-u)^{-1}\big(-2\log(1-u)\big)^{1/2} \to \infty$.

Similarly, for $p \sim \text{uniform}[0,1]$*:*

$$T(z) = G^{-1} \circ F = \tan\Big(\pi(z - \frac{1}{2})\Big)$$
$$\&, \quad T'(z) = \pi \sec^2\Big(\pi(z - \frac{1}{2})\Big)$$

and $fQ_p(u) = 1$*. Thus,* $\lim_{z\to\infty} T'(z) = \lim_{u\to 1^-}(1-u)^{-2} \to \infty$*.*

Example 3 (Pushing uniform to normal)**.** *Let* p *be uniform over* $[0,1]$ *and* $q \sim \mathcal{N}(\mu, \sigma^2)$ *be normal distributed. The unique increasing transformation*

$$\begin{aligned} T(z) &= G^{-1} \circ F = \mu + \sqrt{2}\sigma \cdot \text{erf}^{-1}(2z-1) \\ &= \mu + \sqrt{2}\sigma \cdot \sum_{k=0}^{\infty} \frac{\pi^{k+1/2} c_k}{2k+1} (z - \tfrac{1}{2})^{2k+1}, \end{aligned}$$

where $\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2} \, \mathrm{d}s$ *is the error function, which was Taylor expanded in the last equality. The coefficients* $c_0 = 1$ *and* $c_k = \sum_{m=0}^{k-1} \frac{c_m c_{k-1-m}}{(m+1)(2m+1)}$*. We observe that the derivative of* T *is an infinite sum of squares of polynomials. Both uniform and normal distributions are considered "light-tailed" (all their higher moments exist and are finite). However, an increasing transformation from uniform to normal distribution has unbounded slope. Density quantile functions help us to reveal this precisely:* $fQ_p(u) = 1$ *and* $fQ_q(u) \sim (1-u)\big(-2\log(1-u)\big)^{1/2}$ *i.e. Normal distribution is "relatively" heavier tailed than uniform distribution explaining the asymptotic divergence of this transformation. However, note that this characterization does not follow immediately from Theorem 2. Indeed, density quantiles provide a more granular definition of heavy-tailedness based on the tail-exponent* α *and shape exponent* β*.*

Lemma 1 (Marginal distributions of an elliptical distribution are elliptical, (Frahm, 2004))**.** *Let* $\mathsf{X} = (\mathsf{X}_1, \mathsf{X}_2) \sim \varepsilon_d(\boldsymbol{\mu}, \Sigma, F_R)$ *where* $\mathsf{X}_1 \subseteq \mathbb{R}^{d_1}$ *and* $\mathsf{X}_2 \subseteq \mathbb{R}^{d_2}$ *partition* X *such that* $d_1 + d_2 = d$*. Let* $\boldsymbol{\mu}_1 \in \mathbb{R}^{d_1}, \boldsymbol{\mu}_2 \in \mathbb{R}^{d_1}$ *and* $\Sigma_{11} \in \mathbb{R}^{d_1 \times d_1}, \Sigma_{12} \in \mathbb{R}^{d_1 \times d_2}, \Sigma_{22} \in \mathbb{R}^{d_2 \times d_2}$ *be the corresponding partitions of* $\boldsymbol{\mu}$ *and* Σ *respectively. Then,* $\mathsf{X}_i \sim \varepsilon_{d_i}(\boldsymbol{\mu}_i, \Sigma_{ii}, F_R),\ i \in \{1,2\}$*.*

Lemma 2 (Conditional distributions of an elliptical distribution are elliptical, (Cambanis et al., 1981; Frahm, 2004))**.** *Let* $\mathsf{X} \sim \varepsilon_d(\boldsymbol{\mu}, \Sigma, F_R)$ *where* $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) \in \mathbb{R}^d$ *and* $\Sigma \in \mathbb{R}^{d \times d}$ *is p.s.d with* $\text{rank}(\Sigma) = r$ *and* $\Sigma = \mathbf{A}\mathbf{A}^T$ *where* $\mathsf{X} \stackrel{d}{=} \boldsymbol{\mu} + R\mathbf{A}U^{(r)}$*. Further, let* $\mathsf{X}_1 \subseteq \mathbb{R}^{d_1}$ *and* $\mathsf{X}_2 \subseteq \mathbb{R}^{d_2}$ *partition* X *such that* $d_1 + d_2 = d$*. Let* $\boldsymbol{\mu}_1 \in \mathbb{R}^{d_1}, \boldsymbol{\mu}_2 \in \mathbb{R}^{d_2}$ *and* $\Sigma_{11} \in \mathbb{R}^{d_1 \times d_1}, \Sigma_{12} \in \mathbb{R}^{d_1 \times d_2}, \Sigma_{22} \in \mathbb{R}^{d_2 \times d_2}$ *be the corresponding partitions of* $\boldsymbol{\mu}$ *and* Σ *respectively. If the conditional random vector* $\mathsf{X}_2|(\mathsf{X}_1 = \mathbf{x}_1)$ *exists then*

$$\mathsf{X}_2|(\mathsf{X}_1 = \mathbf{x}_1) \stackrel{d}{=} \boldsymbol{\mu}^* + R^* \Sigma^* U^{(d_2)}$$

where $\boldsymbol{\mu}^* = \boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1),\ \Sigma^* = \Sigma_{22} + \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12},\ R^* = \Big(\big(R^2 - h(\mathbf{x}_1)\big)^{1/2} | \mathsf{X}_1 = \mathbf{x}_1\Big)$ *where* $h(\mathbf{x}_1) = (\mathbf{x}_1 - \boldsymbol{\mu}_1)\Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)^T$*.*

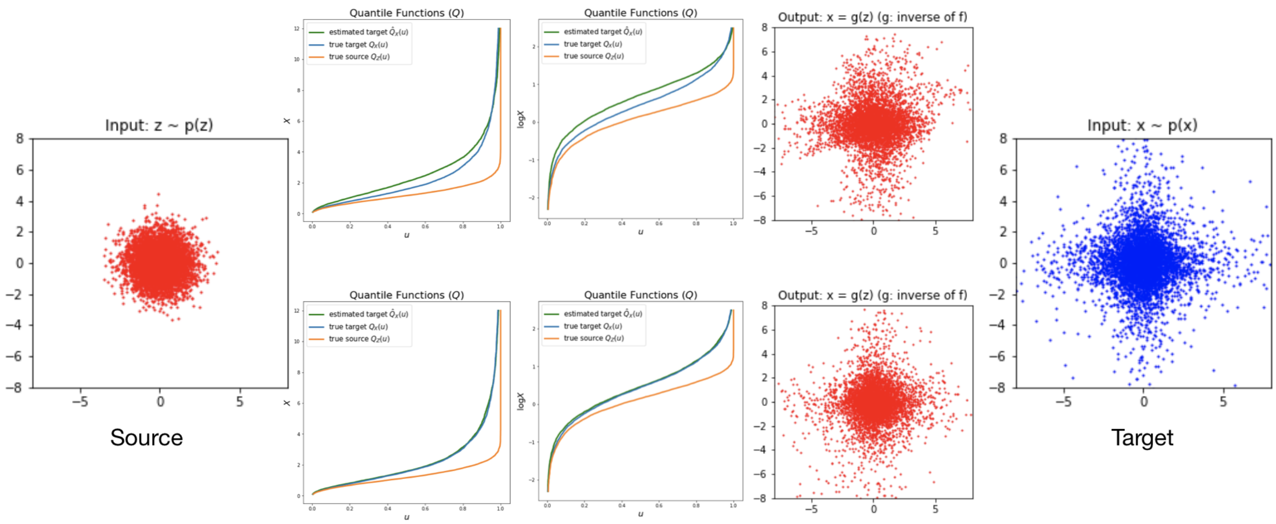

Figure 6. Results for SOS-Flows with degree of polynomial $r = 2$ for two and three blocks. The first and last column plots the samples from the source (Gaussian) and target (student-t) distribution respectively. The two rows from top to bottom in second-fourth columns correspond to results from transformations learned using two, and three compositions (or blocks). The second and third column depict the quantile and log-quantile (for clearer illustration of differences) functions of the source (orange), target (blue), and estimated target (green) and the fourth column plots the samples drawn from the estimated target density. The estimated target quantile function matches exactly with the quantile function of the target distribution illustrating that the higher-order polynomial flows like SOS flows can capture heavier tails of a target. This is further reinforced by their respective tail-coefficients which were estimated to be $\gamma_{\text{source}} = 0.15$, $\gamma_{\text{target}} = 0.81$, $\gamma_{\text{estimated}-\text{target},2} = 0.76$, $\gamma_{\text{estimated}-\text{target},3} = 0.81$. Best viewed in color.