Supplementary Material for: Variational Inference Optimized Using the Curved Geometry of Coupled Free Energy

Amenah Al-Najafi, Kenric P. Nelson, Igor Oliveira

Appendix

1 Proof of Lemma 1

Proof. Applying the coupled logarithm and expanding the expression using Eq. 3 in paper.

$$\ell_\kappa(\boldsymbol{\theta}) \equiv \frac{1}{\alpha} \ln_\kappa \left(\exp_\kappa^{-\frac{1+d\kappa}{\alpha}} \left(\eta(\boldsymbol{\theta}) \cdot T(\boldsymbol{x}) \oplus_\kappa \ln_\kappa \left(\frac{h(\boldsymbol{x})}{Z_\kappa(\eta(\boldsymbol{\theta}))} \right)^{-\frac{\alpha}{1+d\kappa}} \right)^{-\frac{\alpha}{1+d\kappa}} \right) \tag{1}$$

$$\begin{aligned}
&= \frac{1}{\alpha} \left(\eta(\boldsymbol{\theta}) \cdot T(\boldsymbol{x}) \oplus_\kappa \ln_\kappa \left(\frac{h(\boldsymbol{x})}{Z_\kappa(\eta(\boldsymbol{\theta}))} \right)^{-\frac{\alpha}{1+d\kappa}} \right) \\
&= \frac{1}{\alpha} \left(\eta(\boldsymbol{\theta}) \cdot T(\boldsymbol{x}) \oplus_\kappa \frac{1}{\kappa} \left(\left(\frac{h(\boldsymbol{x})}{Z_\kappa(\eta(\boldsymbol{\theta}))} \right)^{-\frac{\alpha\kappa}{1+d\kappa}} - 1 \right) \right) && (2) \\
&= \frac{1}{\alpha} \left(\eta(\boldsymbol{\theta}) \cdot T(\boldsymbol{x}) \oplus_\kappa \frac{1}{\kappa} \left(R(\boldsymbol{\zeta}) - 1 \right) \right). && (3)
\end{aligned}$$

Let $\boldsymbol{\zeta} := (\boldsymbol{\theta}, \boldsymbol{x}, d, \alpha)$, and define

$$R(\boldsymbol{\zeta}) = \left(\frac{h(\boldsymbol{x})}{Z_\kappa(\eta(\boldsymbol{\theta}))} \right)^{-\frac{\alpha\kappa}{1+d\kappa}}. \tag{4}$$

Then, the expression becomes

$$\begin{aligned}\ell_\kappa(\boldsymbol{\theta}) =& \frac{1}{\alpha}\left[\eta(\boldsymbol{\theta})T(\boldsymbol{x}) + \frac{1}{\kappa}\left(R(\boldsymbol{\zeta})-1\right) + \kappa\cdot\eta(\boldsymbol{\theta})T(\boldsymbol{x})\cdot\frac{1}{\kappa}(R(\boldsymbol{\zeta})-1)\right]\\ =& \frac{1}{\alpha}\left[\eta(\boldsymbol{\theta})T(\boldsymbol{x})R(\boldsymbol{\zeta}) + \frac{1}{\kappa}(R(\boldsymbol{\zeta})-1)\right].\end{aligned}$$

Now, we obtain

$$\frac{\partial \ell_\kappa(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i} = \frac{1}{\alpha}\left[\frac{\partial}{\partial \boldsymbol{\theta}_i}\left(\eta(\boldsymbol{\theta})T(\boldsymbol{x})R(\boldsymbol{\zeta})\right) + \frac{1}{\kappa}\cdot\frac{\partial(R(\boldsymbol{\zeta})-1)}{\partial \boldsymbol{\theta}_i}\right],$$

where the derivative of $\eta(\boldsymbol{\theta})T(\boldsymbol{x})R(\boldsymbol{\zeta})$ is

$$\frac{\partial}{\partial \boldsymbol{\theta}_i}\left(\eta(\boldsymbol{\theta})T(\boldsymbol{x})R(\boldsymbol{\zeta})\right) = \left(\frac{\partial \eta(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i}\right)\cdot T(\boldsymbol{x})\cdot R(\boldsymbol{\zeta}) + \eta(\boldsymbol{\theta})\cdot T(\boldsymbol{x})\cdot\frac{\partial R(\boldsymbol{\zeta})}{\partial \boldsymbol{\theta}_i} \tag{5}$$

and the derivative of $(R(\boldsymbol{\zeta})-1)$ is

$$\frac{1}{\kappa}\cdot\frac{\partial}{\partial \boldsymbol{\theta}_i}(R(\boldsymbol{\zeta})-1) = \frac{1}{\kappa}\cdot\frac{\partial R(\boldsymbol{\zeta})}{\partial \boldsymbol{\theta}_i}.$$

Since

$$R(\boldsymbol{\zeta}) = \left(\frac{h(\boldsymbol{x})}{Z_\kappa(\eta(\boldsymbol{\theta}))}\right)^{-r} = h(\boldsymbol{x})^{-r}\cdot Z_\kappa(\eta(\boldsymbol{\theta}))^r, \quad \text{where } r := \frac{\alpha\kappa}{1+d\kappa},$$

hence,

$$\begin{aligned}\frac{\partial R(\boldsymbol{\zeta})}{\partial \boldsymbol{\theta}} &= h(\boldsymbol{x})^{-r}\cdot\frac{\partial}{\partial \boldsymbol{\theta}}\left(Z_\kappa(\eta(\boldsymbol{\theta}))^r\right)\\ &= h(\boldsymbol{x})^{-r}\cdot r\cdot Z_\kappa(\eta(\boldsymbol{\theta}))^{r-1}\cdot\frac{\partial Z_\kappa(\eta(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}}.\end{aligned}$$

By applying the chain rule on $(Z_\kappa(\eta(\boldsymbol{\theta})))$, we can obtain

$$\begin{aligned}\frac{\partial Z_\kappa(\eta(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}_i} &= \frac{\partial Z_\kappa}{\partial \eta}\cdot\frac{\partial \eta(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i}\\ \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i} &= r\cdot h(\boldsymbol{x})^{-r}\cdot Z_\kappa(\eta(\boldsymbol{\theta}))^{r-1}\cdot\frac{\partial Z_\kappa(\eta)}{\partial \eta}\cdot\frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_i}.\end{aligned} \tag{6}$$

From Eqs. (5) and (6), we have

$$\frac{\partial \ell_\kappa(\boldsymbol{\theta})}{\partial \theta_i} = \frac{1}{\alpha}\left[T(\boldsymbol{x})\cdot R(\boldsymbol{\zeta})\cdot\frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_i} + \left(\frac{1}{\kappa} + T(\boldsymbol{x})\cdot\eta(\boldsymbol{\theta})\right)\cdot\frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i}\right], \tag{7}$$

where $R(\boldsymbol{\zeta}) = h(\boldsymbol{x})^{-r} \cdot Z_\kappa(\eta(\boldsymbol{\theta}))^r$, $r = \frac{\alpha\kappa}{1+d\kappa}$, and

$$\frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i} = r \cdot h(\boldsymbol{x})^{-r} \cdot Z_\kappa(\eta(\boldsymbol{\theta}))^{r-1} \cdot \frac{\partial Z_\kappa(\eta)}{\partial \eta} \cdot \frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_i}.$$

For the second derivative, using the first derivative in Eq. (7), we have

$$\frac{\partial^2 \ell_\kappa(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i} = \frac{1}{\alpha}\left(\frac{\partial A}{\partial \theta_j} + \frac{\partial B}{\partial \theta_j}\right), \tag{8}$$

where

$$\begin{aligned} A :=& T(\boldsymbol{x}) \cdot R(\boldsymbol{\zeta}) \cdot \frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_i}, \\ B :=& \left(\frac{1}{\kappa} + T(\boldsymbol{x}) \cdot \eta(\boldsymbol{\theta})\right) \cdot \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i}. \end{aligned}$$

Here, the derivative of (A) can be obtained as

$$\frac{\partial A}{\partial \theta_j} = T(\boldsymbol{x}) \cdot \left[\frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_j} \cdot \frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_i} + R(\boldsymbol{\zeta}) \cdot \frac{\partial^2 \eta(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right]$$

and the derivative of (B) can be obtained as

$$\begin{aligned} \frac{\partial B}{\partial \theta_j} &= \frac{\partial}{\partial \theta_j}\left(\left(\frac{1}{\kappa} + T(\boldsymbol{x}) \cdot \eta(\boldsymbol{\theta})\right) \cdot \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i}\right) \\ &= T(\boldsymbol{x}) \cdot \frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_j} \cdot \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i} + \left(\frac{1}{\kappa} + T(\boldsymbol{x}) \cdot \eta(\boldsymbol{\theta})\right) \cdot \frac{\partial^2 R(\boldsymbol{\zeta})}{\partial \theta_j \partial \theta_i}. \end{aligned}$$

Substituting everything into the full Hessian expression, we can obtain

$$\begin{aligned} \frac{\partial^2 \ell_\kappa(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i} = & \frac{1}{\alpha}\Bigg[T(\boldsymbol{x}) \cdot \left(\frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_j} \cdot \frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_i} + R(\boldsymbol{\zeta}) \cdot \frac{\partial^2 \eta(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right) \\ & + T(\boldsymbol{x}) \cdot \frac{\partial \eta(\boldsymbol{\theta})}{\partial \theta_j} \cdot \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i} + \left(\frac{1}{\kappa} + T(\boldsymbol{x}) \cdot \eta(\boldsymbol{\theta})\right) \cdot \frac{\partial^2 R(\boldsymbol{\zeta})}{\partial \theta_j \partial \theta_i}\Bigg]. \end{aligned}$$

The results can be further simplified by considering the coupled exponential family (i.e., $\eta(\boldsymbol{\theta}) = \boldsymbol{\theta}$) instead of the curved coupled exponential family as follows:

$$\frac{\partial \ell_\kappa(\boldsymbol{\theta})}{\partial \theta_i} = \frac{1}{\alpha}\left[T(\boldsymbol{x}) \cdot R(\boldsymbol{\zeta}) + \left(\frac{1}{\kappa} + T(\boldsymbol{x}) \cdot \boldsymbol{\theta}\right) \cdot \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i}\right] \tag{9}$$

$$\frac{\partial^2 \ell_\kappa(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} = \frac{1}{\alpha}\left[T(\boldsymbol{x})\left(\frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_j} + \frac{\partial R(\boldsymbol{\zeta})}{\partial \theta_i}\right) + \left(\frac{1}{\kappa} + T(\boldsymbol{x}) \cdot \boldsymbol{\theta}\right) \cdot \frac{\partial^2 R(\boldsymbol{\zeta})}{\partial \theta_j \partial \theta_i}\right]. \tag{10}$$

The geometrical quantities on the manifold $\mathcal{S}$ can be derived from Eqs. (9) and (10), i.e., the Fisher metric tensors and the affine connection can be given by

$$
\begin{aligned}
g_{ij} &= -\mathbb{E}_{\boldsymbol{X}}\left[\frac{\partial^2 \ell_\kappa(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right] \\
&= \frac{1}{\alpha}\left[\mathbb{E}_{\boldsymbol{X}}\left[T(\boldsymbol{x})\right]\left(\frac{\partial R(\boldsymbol{\zeta})}{\partial\theta_j}+\frac{\partial R(\boldsymbol{\zeta})}{\partial\theta_i}\right)+\left(\frac{1}{\kappa}+\mathbb{E}_{\boldsymbol{X}}\left[T(\boldsymbol{x})\right]\cdot\boldsymbol{\theta}\right)\cdot\frac{\partial^2 R(\boldsymbol{\zeta})}{\partial\theta_j\partial\theta_i}\right].
\end{aligned}
$$

By substituting the first derivative in Eq. (9) and the second derivative in Eq. (10), we obtain the full expression for Γ_{ijk} as

$$
\begin{aligned}
\Gamma_{ijk} &= \mathbb{E}_{\boldsymbol{X}}\left[\frac{\partial^2 \ell_\kappa(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\frac{\partial \ell_\kappa(\boldsymbol{\theta})}{\partial\theta_k}\right] \\
&= \frac{1}{\alpha^2}\cdot\mathbb{E}_{\boldsymbol{X}}\left[(A_1+A_2)\cdot(B_1+B_2)\right].
\end{aligned}
$$

$\square$

---

2 Coupled Free Energy derivation

Proof of Theorem 1

Theorem 2.1 (Coupled Free Energy for Multivariate Coupled Gaussian)**.** *Let $\boldsymbol{x}\in\mathbb{R}^d$ and $\mathbf{z}\in\mathbb{R}^n$ be random vectors, and suppose that the posterior $q(\mathbf{z}\mid\boldsymbol{x})$ and the prior $p(\mathbf{z})$ are both multivariate coupled Gaussian with their joint PDFs defined as:*

$$
f\left(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma},\kappa\right)\equiv\begin{cases}\frac{1}{Z(\boldsymbol{\Sigma},\kappa)}\left(1+\kappa\left|(\boldsymbol{x}-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right|\right)_+^{-\frac{1+d\kappa}{2\kappa}}, & \kappa\neq0,\kappa>-1/d;\\ \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right), & \kappa=0.\end{cases}
$$

Given a Coupled Variational Autoencoder (CVAE) with $\kappa\neq0$, the Coupled Free Energy (CFE) takes the following form with a coupled divergence term plus a reconstruction loss term:

$$
\begin{aligned}
\mathcal{F}_{\theta,\phi,\kappa}(\boldsymbol{x}) &= \frac{1}{2}\mathbb{E}_{\mathbf{z}\sim Q_\phi^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z}|\boldsymbol{x})}\left[\ln_\kappa\left(p_\theta(\mathbf{z})^{\frac{-2}{1+d\kappa}}\right)-\ln_\kappa\left(q_\phi(\mathbf{z}\mid\boldsymbol{x})^{\frac{-2}{1+d\kappa}}\right)\right]d\mathbf{z} \\
&+\frac{1}{2}\mathbb{E}_{\mathbf{z}\sim Q_\phi^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z}|\boldsymbol{x})}\left[\ln_\kappa\left(\exp_\kappa^{\frac{-(1+d\kappa)}{2}}\left((\boldsymbol{x}-\bar{\boldsymbol{x}}_{x|z})^\top\boldsymbol{\Sigma}_{x|z}^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}_{x|z})\oplus_\kappa\ln_\kappa\left(\frac{1}{Z_q}\right)^{\left(\frac{-2}{1+d\kappa}\right)}\right)\right)^{\frac{-2}{1+d\kappa}}\right]
\end{aligned}
$$

The CFE simplifies to:

$$
\begin{aligned}
\mathcal{F}_{\theta,\phi,\kappa}(\boldsymbol{x}) &= -\frac{d\left(1+\kappa A_q\right)}{2}+\frac{1+\kappa A_p}{2}\left((\boldsymbol{\mu}_p-\boldsymbol{\mu}_q)^T\boldsymbol{\Sigma}_p^{-1}(\boldsymbol{\mu}_p-\boldsymbol{\mu}_q)+\operatorname{tr}\left(\boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_q\right)\right) \\
&-\frac{A_q}{2}+\frac{A_p}{2}-\mathbb{E}_{\mathbf{z}\sim Q_\phi^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z}|\boldsymbol{x})}\left[\frac{1}{2}\left((\boldsymbol{x}-\bar{\boldsymbol{x}}_{x|z})^\top\boldsymbol{\Sigma}_{x|z}^{-1}(\boldsymbol{x}-\bar{\boldsymbol{x}}_{x|z})\oplus_\kappa A_{x|z}\right)\right]
\end{aligned}
$$

Proof. 1. Calculating the Kullback-Leibler divergence

$$
q(z\mid x)=\left(1+\kappa\left(\left((z-\mu_q)^\intercal\Sigma_q^{-1}(z-\mu_q)\right)\oplus_\kappa\left(\ln_\kappa\left(\frac{1}{Z}\right)^{-\frac{2}{1+d\kappa}}\right)\right)\right)^{-\frac{1+d\kappa}{2\kappa}}
$$

$$p(z)=\left(1+\kappa\left(\left((z-\mu_p)^\intercal\,\Sigma_p^{-1}(z-\mu_p)\right)\oplus_\kappa\left(\ln_\kappa\left(\frac{1}{Z}\right)^{-\frac{2}{1+d\kappa}}\right)\right)\right)^{-\frac{1+d\kappa}{2\kappa}}$$

For simplicity, we set $A_q=\ln_\kappa\left(\frac{1}{Z_q}\right)^{-\frac{2}{1+d\kappa}}$

$$\begin{aligned}\frac{1}{2}\ln_\kappa\left(q(\mathbf{z}\mid\mathbf{x})\right)^{\frac{-2}{1+d\kappa}} &=\frac{1}{2\kappa}\left(\left(q(\mathbf{z}\mid\mathbf{x})\right)^{\frac{-2\kappa}{1+d\kappa}}-1\right)\\ &=\frac{1}{2\kappa}\left(1+\kappa\left(\left((z-\mu_q)^\intercal\,\Sigma_q^{-1}\,(z-\mu_q)\right)\oplus_\kappa A_q\right)-1\right)\\ &=\frac{1}{2}\left((z-\mu_q)^\intercal\,\Sigma_q^{-1}\,(z-\mu_q)\right)+\frac{A_q}{2}+\frac{\kappa}{2}\left((z-\mu_q)^\intercal\,\Sigma_q^{-1}\,(z-\mu_q)\right)A_q\\ &=\frac{1}{2}\left((z-\mu_q)^\intercal\,\Sigma_q^{-1}\,(z-\mu_q)\right)\oplus_{2\kappa}\frac{A_q}{2}\end{aligned}$$

The same thing with $\frac{1}{2}\ln_\kappa\left(p(\mathbf{z})\right)^{\frac{-2}{1+d\kappa}}=\frac{1}{2}\left((z-\mu_p)^\intercal\,\Sigma_p^{-1}\,(z-\mu_p)\right)\oplus_{2\kappa}\frac{A_p}{2}$

$$D_\kappa\left[q(z\mid x)\|p(z)\right]=\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}\left(\frac{1}{2}\ln_\kappa p(z\mid x)^{\frac{-2}{1+d\kappa}}-\frac{1}{2}\ln_\kappa q(z)^{\frac{-2}{1+d\kappa}}\right)$$

$$\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}\left(\frac{1}{2}\left((z-\mu_p)^\intercal\,\Sigma_p^{-1}\,(z-\mu_p)\right)\oplus_{2\kappa}\frac{A_p}{2}-\frac{1}{2}\left((z-\mu_q)^\intercal\,\Sigma_q^{-1}\,(z-\mu_q)\right)\oplus_{2\kappa}\frac{A_q}{2}\right) \tag{11}$$

Take the first item from eq.11:

$$\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}\left(\frac{1}{2}\left((z-\mu_p)^\intercal\,\Sigma_p^{-1}\,(z-\mu_p)\right)\oplus_{2\kappa}\frac{A_p}{2}\right)=$$

$$\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}\left(\frac{1}{2}\left((z-\mu_p)^\intercal\Sigma_p^{-1}\,(z-\mu_p)\right)+\frac{A_p}{2}+\frac{\kappa}{2}\left((z-\mu_p)^\intercal\Sigma_p^{-1}\,(z-\mu_p)\right)^2A_p\right)$$

$$\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}\left(\frac{1}{2}\,(z-\mu_p)^\intercal\,\Sigma_p^{-1}\,(z-\mu_p)\right)(1+\kappa A_p)+\frac{A_p}{2}$$

Let's define the centered variable:

$$y_p=z-\mu_p$$

The quantity we want to compute is:

$$\mathbb{E}\left[y_p^\intercal\Sigma_p^{-1}y_p\right]$$

Express the quadratic form using the trace

The identity for the trace of a scalar:

$$a^\intercal M a = \mathrm{Tr}\left(M a a^\intercal\right)$$

where a is a vector and M is a matrix. Applying this here:

$$y_p^\intercal \Sigma_p^{-1} y_p = \mathrm{Tr}\left(\Sigma_p^{-1} y_p y_p^\intercal\right)$$

Taking the expectation:

$$\mathbb{E}\left[\frac{1}{2}(y_p^\intercal \Sigma_p^{-1} y_p)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\frac{1}{2}(\Sigma_p^{-1} y_p y_p^\intercal)\right)\right]$$

Since the trace of an expectation is the expectation of the trace:

$$= \mathrm{Tr}\left(\frac{1}{2}\Sigma_p^{-1}\mathbb{E}\left[y_p y_p^\intercal\right]\right)$$

$$Q^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z} \mid \mathbf{x}) \sim \frac{\Gamma(\frac{1+\kappa(d+2)}{2\kappa})}{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right)\sqrt{(\frac{\pi}{\kappa})^d|\Sigma_q|}}\left(1+\kappa\left|(z-\mu_q)^\intercal {\Sigma_q}^{-1}(z-\mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} \tag{12}$$

The first term

Let an eigenvalue decomposition $\mathbf{\Sigma}_q^{-1} = U_q \lambda_q^{-1} U_q$ Applying this transformation:

$$\mathbb{E}\left[y_q y_q^\intercal\right] = \frac{\Gamma(\frac{1+\kappa(d+2)}{2\kappa})\frac{1}{2}\mathbf{\Sigma}_q^{-1}}{\Gamma\left(\frac{2\kappa+1}{2\kappa}\right)\sqrt{(\frac{\pi}{\kappa})^d|\Sigma_q|}}\int\left(y_q y_q^\intercal\right)\left(1+\kappa\left|y_q^\intercal U_q \lambda_q^{-1} U_q^\intercal y_q\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} \tag{13}$$

Let $s_q = U_q^\intercal y_q, \Rightarrow y_q = U_q s_q$ and $ds_q = |U_q^\intercal| dy_q \Rightarrow ds_q = dy_q$ Applying in the 13

$$\begin{aligned}
&= \frac{\Gamma(\frac{1+\kappa(d+2)}{2\kappa})\frac{1}{2}\mathbf{\Sigma}_q^{-1}}{\Gamma\left(\frac{2\kappa+1}{2\kappa}\right)\sqrt{(\frac{\pi}{\kappa})^d|\Sigma_q|}}\int\left(U_q s_q s_q^\intercal U_q^\intercal\right)\left(1+\kappa\left|s_q^\intercal \lambda_q^{-1} s_q\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds_q \\
&= \frac{\Gamma(\frac{1+\kappa(d+2)}{2\kappa})\frac{1}{2}\mathbf{\Sigma}_q^{-1}}{\Gamma\left(\frac{2\kappa+1}{2\kappa}\right)\sqrt{(\frac{\pi}{\kappa})^d|\Sigma_q|}} U_q\left(\int\left(s_q s_q^\intercal\right)\left(1+\kappa\left|s_q^\intercal \lambda_q^{-1} s_q\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds_q\right) U_q^\intercal
\end{aligned}$$

λ_q is Positive Definite Matrix,

$$\begin{aligned}
I &= \int\left(s_q s_q^\intercal\right)\left(1+\kappa\left(s_q^\intercal \lambda_q^{-1} s_q\right)\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds_q \\
&= \int\left(s_q s_q^\intercal\right)\left(1+\kappa\sum_{r=1}^{d}\frac{s_{q_r}^2}{\lambda_{q_r}}\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds_q \\
&= \int\left(s_{q_i} s_{q_j}\right)\left(1+\kappa\sum_{r=1}^{d}\frac{s_{q_r}^2}{\lambda_{q_r}}\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds_q
\end{aligned}$$

We consider the integral:

$$I = \int \left(s_{q_i} s_{q_j}\right) \left(1 + \kappa \sum_{r=1}^{d} \frac{s_{q_r}^2}{\lambda_{q_r}}\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds_q$$

Applying the Gamma Function Representation, We use the identity:

$$(1 + \kappa \sum_{r=1}^{d} \frac{s_{q_r}^2}{\lambda_{q_r}})^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t(1+\kappa \sum_{r=1}^{d} \frac{s_{q_r}^2}{\lambda_{q_r}})} dt$$

where:

$$\alpha = \frac{1 + \kappa(d+2)}{2\kappa}$$

Substituting this into the given integral:

$$I = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t} \left(\int s_{q_i} s_{q_j} e^{-t\kappa \sum_{r=1}^{d} \frac{s_{q_r}^2}{\lambda_{q_r}}} ds_q \right) dt$$

Evaluating the Inner Integral. The inner integral is:

$$\begin{aligned} J &= \int s_{q_i} s_{q_j} e^{-t\kappa \sum_{r=1}^{d} \frac{s_{q_r}^2}{\lambda_{q_r}}} ds \\ &= \int s_{q_i} s_{q_j} \prod_{r=1}^{d} \int_{-\infty}^{\infty} e^{-t\kappa \frac{s_{q_r}^2}{\lambda_{q_r}}} ds_r \end{aligned}$$

Since the matrix (λ_q) is diagonal, the integral factorizes into a product of independent integrals.

Thus, if $(i \neq j)$, we obtain:

$$\left(\prod_{r \neq i,j}^{d} \int_{-\infty}^{\infty} e^{-t\kappa \frac{s_r^2}{\lambda_r}} ds_r \right) \cdot \left(\int s_i e^{-t\kappa \frac{s_{q_i}^2}{\lambda_{q_i}}} ds_{q_i} \right) \cdot \left(\int s_{q_j} e^{-t\kappa \frac{s_{q_j}^2}{\lambda_{q_i}}} ds_{q_j} \right) = 0$$

The second and third term = 0, due to symmetry (since the integral of an odd function over a symmetric domain is zero).

If $i = j$, then:

$$\begin{aligned} &= \prod_{r \neq i} \int_{-\infty}^{\infty} e^{-t\kappa \frac{s_{q_r}^2}{\lambda_{q_r}}} ds_r . \int s_{q_i}^2 e^{-t\kappa \frac{s_{q_i}^2}{\lambda_{q_i}}} ds_{q_i} \\ &= \prod_{r \neq i} \frac{\sqrt{\pi}}{\sqrt{\frac{t\kappa}{\lambda_{q_r}}}} . \frac{\sqrt{\pi}}{2\sqrt{\frac{(t\kappa)^3}{\lambda_{q_i}^3}}} \end{aligned}$$

$$=\prod_{r\neq i}\frac{\sqrt{\pi}.\sqrt{\lambda_{q_r}}}{\sqrt{t\kappa}}.\frac{\sqrt{\pi}\sqrt{\lambda_{q_i}}}{2\sqrt{(t\kappa)^3}}.\lambda_{q_i}$$

$$=\frac{\sqrt{(\pi^d).\prod_{r=1}^{d}\lambda_{q_r}}}{2\sqrt{(t\kappa)^{d+2}}}.\lambda_{q_i}$$

$$=\frac{\sqrt{(\pi)^d|\mathbf{\Sigma}_q|}}{2\sqrt{(t\kappa)^{d+2}}}.\mathbf{\Sigma}_q$$

Substituting Back into the Main Integral For $i=j$, the integral simplifies to:

$$I=\frac{1}{\Gamma(\alpha)}\int_0^\infty t^{\alpha-1}e^{-t}\left(\frac{\sqrt{(\pi)^d|\mathbf{\Sigma}_q|}}{2\sqrt{(t\kappa)^{d+2}}}.\mathbf{\Sigma}_q\right)dt$$

Rearranging the constants:

$$I=\frac{\sqrt{(\pi)^d|\mathbf{\Sigma}_q|}}{2\sqrt{(\kappa)^{d+2}}\Gamma(\alpha)}.\mathbf{\Sigma}_q\int_0^\infty t^{\alpha-1}(t)^{-\frac{d+2}{2}}e^{-t}dt.$$

$$I=\frac{\sqrt{\pi^d|\mathbf{\Sigma}_q|}.\mathbf{\Sigma}_q}{2\sqrt{(\kappa)^{d+2}}\Gamma(\alpha)}\int_0^\infty t^{\alpha-1-(d+2)/2}e^{-t}dt.$$

The integral now takes the form of a Gamma function:

$$\int_0^\infty t^{\beta-1}e^{-t}dt=\Gamma(\beta),$$

where:

$$\beta=\alpha-\frac{d+2}{2}.$$

Thus, the integral evaluates to:

$$I=\frac{\sqrt{\pi^d|\mathbf{\Sigma}_q|}.\mathbf{\Sigma}_q}{2\sqrt{(\kappa)^{d+2}}\Gamma(\alpha)}\Gamma\left(-1-\frac{d}{2}+\alpha\right).$$

$$\left(\frac{\sqrt{\pi^d|\mathbf{\Sigma}_q|}.\mathbf{\Sigma}_q}{2\sqrt{(\kappa)^{d+2}}\Gamma(\frac{1+\kappa(d+2)}{2\kappa})}\Gamma\left(\frac{1}{2\kappa}\right)\right)(1+\kappa A_q)+A_q. \tag{14}$$

The second term

$$J=\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}\left((z-\mu_p)^\intercal\,\mathbf{\Sigma}_p^{-1}\,(z-\mu_p)\right)$$

Define the Expectation Since $\mathbf{z}$ is distributed according to Q with mean μ_q and covariance $\mathbf{\Sigma}_q$, we need to compute:

$$\mathbb{E}_Q\left[(\mathbf{z}-\mu_p)^T\mathbf{\Sigma}_p^{-1}(\mathbf{z}-\mu_p)\right].$$

Rewriting $\mathbf{z}$ in terms of its mean μ_q:

$$\mathbf{z} - \mu_p = (\mathbf{z} - \mu_q) + (\mu_q - \mu_p).$$

Thus, expanding the quadratic term:

$$(\mathbf{z} - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_p) = (\mathbf{z} - \mu_q + \mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_q + \mu_q - \mu_p).$$

Expand the Quadratic Form, expanding the expression:

$$\begin{aligned}(\mathbf{z} - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_p) = \\ & (\mathbf{z} - \mu_q)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_q) + 2(\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_q) + \\ & (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p).\end{aligned}$$

Taking the expectation on both sides:

$$\begin{aligned}\mathbb{E}_Q \left[(\mathbf{z} - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_p) \right] = \\ & \mathbb{E}_Q \left[(\mathbf{z} - \mu_q)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_q) \right] + 2(\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} \mathbb{E}_Q [\mathbf{z} - \mu_q] \\ & + (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p).\end{aligned}$$

Since $\mathbb{E}_Q[\mathbf{z} - \mu_q] = 0$, the middle term vanishes, reducing the equation to:

$$\begin{aligned}\mathbb{E}_Q \left[(\mathbf{z} - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_p) \right] =& \mathbb{E}_Q \left[(\mathbf{z} - \mu_q)^T \boldsymbol{\Sigma}_p^{-1} (\mathbf{z} - \mu_q) \right] + (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p) \\ =& \boldsymbol{\Sigma}_p^{-1} \mathbb{E}_Q \left[(\mathbf{z} - \mu_q)^T (\mathbf{z} - \mu_q) \right] + (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p) \\ =& \boldsymbol{\Sigma}_p^{-1} I + (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p) + (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p) \\ =& \left(\boldsymbol{\Sigma}_p^{-1} \frac{\sqrt{\pi^d |\boldsymbol{\Sigma}_q|}.\boldsymbol{\Sigma}_q \Gamma\left(\frac{1}{2\kappa}\right)}{2\sqrt{(\kappa)^{d+2}} \Gamma(\frac{1+\kappa(d+2)}{2\kappa})} + (\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p) \right) (1 + \kappa A_p) + A_p\end{aligned}$$

Where I eq.14 is the previous result.

$$\begin{aligned}&\frac{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)}{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right) \sqrt{\left(\frac{\pi}{\kappa}\right)^n |\Sigma_q|}} \\ &\times \left[\left(\boldsymbol{\Sigma}_q^{-1} \cdot \frac{\sqrt{\pi^d |\boldsymbol{\Sigma}_q|}\, \boldsymbol{\Sigma}_q \, \Gamma\left(\frac{1}{2\kappa}\right)}{2\sqrt{\kappa^{d+2}}\, \Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \right) (1 + \kappa A_q) + A_q \right. \\ &\quad - \left(\boldsymbol{\Sigma}_p^{-1} \cdot \frac{\sqrt{\pi^n |\boldsymbol{\Sigma}_q|}\, \boldsymbol{\Sigma}_q \, \Gamma\left(\frac{1}{2\kappa}\right)}{2\sqrt{\kappa^{n+2}}\, \Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \right) (1 + \kappa A_p) + A_p \,] \\ &\quad - \left((\mu_q - \mu_p)^T \boldsymbol{\Sigma}_p^{-1} (\mu_q - \mu_p) \right) (1 + \kappa A_p) + A_p\end{aligned}$$

$$-\frac{d\left(1+\kappa A_q\right)}{2}+\frac{1+\kappa A_p}{2}\left((\mu_p-\mu_q)^T\boldsymbol{\Sigma}_p^{-1}(\mu_p-\mu_q)+\operatorname{tr}\left(\boldsymbol{\Sigma}_{\mathbf{p}}^{-\mathbf{1}}\boldsymbol{\Sigma}_{\mathbf{q}}\right)\right)-\frac{A_q}{2}+\frac{A_p}{2}$$

---

2. Reconstruction

$$p_\theta(x|z)=\left(1+\kappa\left(\left(\left(x-\mu_{x|z}\right)^{\intercal}\boldsymbol{\Sigma}_{x|z}^{-1}(x-\mu_{x|z})\right)\oplus_\kappa\left(\ln_\kappa\left(\frac{1}{Z_{x|z}}\right)^{-\frac{2}{1+d\kappa}}\right)\right)\right)^{-\frac{1+d\kappa}{2\kappa}}$$

$$\mathbb{E}_{Q\left(\frac{2\kappa}{1+d\kappa}\right)(\mathbf{z}|\mathbf{x})}[\frac{1}{2}\ln_\kappa p_\theta(x|z)^{\frac{-2}{1+d\kappa}}],$$

For simplicity, we set $A_{x|z}=\ln_\kappa\left(\frac{1}{Z_{x|z}}\right)^{-\frac{2}{1+d\kappa}}$

$$\begin{aligned}\frac{1}{2}\ln_\kappa\left(q(\mathbf{z}\mid\mathbf{x})\right)^{\frac{-2}{1+d\kappa}}=&\frac{1}{2\kappa}\left(\left(q(\mathbf{z}\mid\mathbf{x})\right)^{\frac{-2\kappa}{1+d\kappa}}-1\right)\\=&\frac{1}{2\kappa}\left(1+\kappa\left(\left(\left(x-\mu_{x|z}\right)^{\intercal}\boldsymbol{\Sigma}_{x|z}^{-1}\left(x-\mu_{x|z}\right)\right)\oplus_\kappa A_{x|z}\right)-1\right)\\=&\left(\frac{1}{2}\left(x-\mu_{x|z}\right)^{\intercal}\boldsymbol{\Sigma}_{x|z}^{-1}\left(x-\mu_{x|z}\right)\right)\oplus_{2\kappa}\frac{A_{x|z}}{2}\end{aligned}$$

We assume that

$$p_\theta(x|z)=\mathcal{N}_\kappa(\hat{x},1),$$

where $\bar{x}$ represents the reconstructed image from the model.

Taking the expectation over $\left(z_Q\sim Q^{\left(\frac{2\kappa}{1+d\kappa}\right)}(z|x)\right)$, we get:

$$\begin{aligned}&\mathbb{E}_{z_Q\sim Q\left(\frac{2\kappa}{1+d\kappa}\right)(z|x)}\left[\frac{1}{2}\ln_\kappa p(x|z)^{\frac{-2}{1+d\kappa}}\right]\\&=\mathbb{E}_{z_Q\sim Q\left(\frac{2\kappa}{1+d\kappa}\right)(z|x)}\left[\frac{1}{2}\ln_\kappa p(x|z)^{\frac{-2}{1+d\kappa}}\right]\\&=\mathbb{E}_{z_Q\sim Q\left(\frac{2\kappa}{1+d\kappa}\right)(z|x)}\left[\frac{1}{2}\left(x-\bar{x}_{x|z}\right)^{\intercal}\left(x-\bar{x}_{x|z}\right)\oplus_{2\kappa}\frac{A_{x|z}}{2}\right]\\&=\frac{1}{2}\mathbb{E}_{z_Q\sim Q\left(\frac{2\kappa}{1+d\kappa}\right)(z|x)}\left[\left(x-\bar{x}_{x|z}\right)^2\oplus_\kappa A_{x|z}\right]\end{aligned}$$

□

2.1 Derivation of the Normalization Constant

Proof.

$$q(z \mid x) \equiv \frac{1}{Z(\mathbf{\Sigma}, \kappa)} \left(1 + \kappa \left|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+d\kappa}{2\kappa}}$$

$$\begin{aligned}
Q^{\frac{2\kappa}{1+d\kappa}}(z \mid x) =& \frac{\frac{1}{Z(\mathbf{\Sigma},\kappa,\alpha)} \left(1 + \kappa \left|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}}}{\int \left(\frac{1}{Z(\mathbf{\Sigma},\kappa,\alpha)} \left(1 + \kappa \left|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}}\right)} \\
=& \frac{\left(1 + \kappa \left|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}}}{\int \left(\left(1 + \kappa \left|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}}\right)}
\end{aligned}$$

The normalization constant Z is given by the integral in the denominator

$$Z_Q = \int \left(\left(1 + \kappa \left|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}}\right) dz$$

To simplify the integral, we define:

$$y_q = z - \mu_q$$

Since ($\mathbf{\Sigma}$) is a covariance matrix, we can perform eigenvalue decomposition:

$$\mathbf{\Sigma}_q^{-1} = U_q \lambda_q^{-1} U_q^T$$

where: (U_q) is an orthogonal matrix (coordinate transformation). (λ_q) is a diagonal matrix containing eigenvalues.

Substituting this into the $|(z - \mu_q)^\intercal \mathbf{\Sigma}_q^{-1}(z - \mu_q)|$ form:

$$y_q^T \mathbf{\Sigma}_q^{-1} y_q = y_q^T U_q \lambda_q^{-1} U_q^T y_q \tag{15}$$

Using the transformation:

$$s_q = U_q^T y_q$$

substitute in (4)

$$s_q^T \lambda_q^{-1} s_q$$

The integral becomes:

$$Z_Q = \int \left(1 + \kappa \left(s_q^T \lambda_q^{-1} s_q\right)\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds$$

Since (λ_q^{-1}) is diagonal, we can express $(s_q^T \lambda_q^{-1} s_q)$ as a sum of squared terms:

$$s_q^T \lambda_q^{-1} s_q = \begin{bmatrix} s_{q_1} & s_{q_2} & \dots & s_{q_d} \end{bmatrix} \begin{bmatrix} \Lambda_{q_1}^{-1} & 0 & 0 & \dots & 0 \\ 0 & \Lambda_{q_2}^{-1} & 0 & \dots & 0 \\ 0 & 0 & \Lambda_{q_3}^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \Lambda_{q_d}^{-1} \end{bmatrix} \begin{bmatrix} s_{q_1} \\ s_{q_2} \\ \vdots \\ s_{q_d} \end{bmatrix} = \sum_{i=1}^{d} \lambda_{q_i}^{-1} s_{q_i}^2$$

We started with:

$$(1 + \kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}})^{-\frac{1+\kappa(d+2)}{2\kappa}}$$

Applying the Gamma function integral representation:

$$(1+u)^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t(1+u)} dt$$

where $(\alpha = \frac{1+\kappa(d+2)}{2\kappa})$ and $(u = \kappa \sum_{i=1}^{d} \frac{s_i^2}{\lambda_i})$, we obtain:

$$(1 + \kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}})^{-\frac{1+\kappa(d+2)}{2\kappa}} = \frac{1}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \int_0^\infty t^{\frac{1+\kappa(d+2)}{2\kappa}-1} e^{-t(1+\kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}})} dt$$

Thus, our integral transforms into:

$$Z_Q = \int \left(1 + \kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}}\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} ds$$

Substituting the Gamma representation:

$$Z_Q = \int \frac{1}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \int_0^\infty t^{\frac{1+\kappa(d+2)}{2\kappa}-1} e^{-t(1+\kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}})} dt\, ds$$

$$Z_Q = \frac{1}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \int_0^\infty t^{\frac{1+\kappa(d+2)}{2\kappa}-1} e^{-t} \left[\int e^{-t\kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}}} ds\right] dt$$

Solving the Inner Integral

$$I = \int e^{-t\kappa \sum_{i=1}^{d} \frac{s_{q_i}^2}{\lambda_{q_i}}} ds_q$$

$$I = \prod_{i=1}^{d} \int_{-\infty}^{\infty} e^{-t\kappa \frac{s_{q_i}^2}{\lambda_{q_i}}} ds_{q_i}.$$

Using the Gaussian integral formula:

$$\int_{-\infty}^{\infty} e^{-t\kappa \frac{s^2}{\lambda}} ds = \sqrt{\frac{\pi\lambda}{t\kappa}}.$$

Thus, the product over all i gives:

$$I = \prod_{i=1}^{d} \sqrt{\frac{\pi\lambda_{q_i}}{t\kappa}}.$$
$$I = \left(\frac{\pi}{t\kappa}\right)^{\frac{d}{2}} \sqrt{|\mathbf{\Sigma}_q|}.$$

Now substituting (I) into (Z):

$$Z_Q = \frac{1}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \sqrt{|\mathbf{\Sigma}_q|} \int_0^{\infty} t^{\frac{1+\kappa(d+2)}{2\kappa}-1} e^{-t} \left(\frac{\pi}{t\kappa}\right)^{\frac{d}{2}} dt.$$

Rewriting:

$$Z_Q = \frac{1}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \sqrt{|\mathbf{\Sigma}_q|} \left(\frac{\pi}{\kappa}\right)^{\frac{d}{2}} \int_0^{\infty} t^{\frac{1+\kappa(d+2)}{2\kappa}-1-\frac{d}{2}} e^{-t} dt.$$

The remaining integral as the Gamma function:

$$\int_0^{\infty} t^{p-1} e^{-t} dt = \Gamma(p), \quad \text{for } p > 0.$$

(p) is:

$$\begin{aligned} p =& \frac{1+\kappa(d+2)}{2\kappa} - \frac{d}{2}. \\ =& \frac{1+2\kappa}{2\kappa} \end{aligned}$$

So:

$$\int_0^{\infty} t^{\frac{1+\kappa(d+2)}{2\kappa}-1-\frac{d}{2}} e^{-t} dt = \Gamma\left(\frac{1+2\kappa}{2\kappa}\right).$$

Final Expression Thus, we obtain:

$$Z_Q = \frac{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right)}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \sqrt{(\frac{\pi}{\kappa})^d |\mathbf{\Sigma}_q|}. \tag{16}$$

$$Q^{\frac{2\kappa}{1+d\kappa}}(z \mid x) = \frac{1}{Z_Q(\mathbf{\Sigma}, \kappa)} \left(1 + \kappa_Q \left|(z-\mu_q)^{\intercal} \mathbf{\Sigma}_Q^{-1} (z-\mu_q)\right|\right)^{-\frac{1+d\kappa_Q}{2\kappa_Q}}$$

$$\frac{1+d\kappa_Q}{2\kappa_Q} = (\frac{1+d\kappa_q}{2\kappa_q})(1+\frac{2\kappa_q}{1+d\kappa_q})$$
$$\kappa_Q = \frac{\kappa_q}{1+2\kappa_q} \tag{17}$$

To find $(\mathbf{\Sigma}_Q)$ from the given equation:

$$\kappa_q|(z-\mu_q)^T\mathbf{\Sigma}_q^{-1}(z-\mu_q)| = \kappa_Q|(z-\mu_q)^T\mathbf{\Sigma}_Q^{-1}(z-\mu_q)|$$

Since determinants operate on the eigenvalues of the matrices, we can rewrite the equation as:

$$\kappa_q|\mathbf{\Sigma}_q^{-1}||(z-\mu_q)^T(z-\mu_q)| = \kappa_Q|\mathbf{\Sigma}_Q^{-1}||(z-\mu_q)^T(z-\mu_q)|$$

$$\kappa|\mathbf{\Sigma}_q^{-1}| = \kappa_Q|\mathbf{\Sigma}_Q^{-1}|$$

$$|\mathbf{\Sigma}_Q^{-1}| = \frac{\kappa}{\kappa_Q}|\mathbf{\Sigma}_q^{-1}|$$

Taking the determinant on both sides:

$$|\mathbf{\Sigma}_Q|^{-1} = \frac{\kappa_q}{\kappa_Q}|\mathbf{\Sigma}_q|^{-1}$$

$$|\mathbf{\Sigma}_Q|^{-1} = \frac{\kappa_q}{\frac{\kappa_q}{1+2\kappa_q}}|\mathbf{\Sigma}_q|^{-1} = (1+2\kappa_q)|\mathbf{\Sigma}_q|^{-1} \tag{18}$$

$$Q^{\frac{2\kappa}{1+d\kappa}}(z \mid x) \equiv \frac{\Gamma(\frac{1+\kappa(d+2)}{2\kappa})}{\Gamma(\frac{1+2\kappa}{2\kappa}\sqrt{\frac{\pi}{\kappa}|\mathbf{\Sigma}_q|}}\left(1+\frac{\kappa}{1+2\kappa}\left|(z-\mu_q)^\intercal(1+2\kappa)\mathbf{\Sigma}_q{}^{-1}(z-\mu_q)\right|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} \tag{19}$$

$$Q^{\frac{2\kappa}{1+d\kappa}}(z \mid x) \equiv \frac{\Gamma(\frac{1+\kappa(d+2)}{2\kappa})}{\Gamma(\frac{1+2\kappa}{2\kappa})\sqrt{\frac{\pi}{\kappa}|\mathbf{\Sigma}_q|}}\left(1+\kappa\left|(z-\mu_q)^\intercal\,\mathbf{\Sigma}_q{}^{-1}(z-\mu_q)\right|\right)^{-\frac{1+2\kappa(d+2)}{2\kappa}} \tag{20}$$

□