

May 9, 2025

1. Deriving the normalization

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

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

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

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 $|(z - \mu_q)^\top \mathbf{\Sigma}_q^{-1} (z - \mu_q)|$ form:

$$y_q^T \mathbf{\Sigma}_q^{-1} y_q = y_q^T U_q \lambda_q^{-1} U_q^T y_q \quad (1)$$

Using the transformation:

$$s_q = U_q^T y_q$$

substitute in (4)

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

The integral becomes:

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

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 $(u = \kappa \sum_{i=1}^d \frac{s_{q_i}^2}{\lambda_{q_i}})$, 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:

$$I = \prod_{i=1}^d \sqrt{\frac{\pi\lambda_{q_i}}{t\kappa}}.$$

$$I = \left(\frac{\pi}{t\kappa}\right)^{\frac{d}{2}} \sqrt{|\Sigma_q|}.$$

Now substituting (I) into (Z) :

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

Rewriting:

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

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:

$$Z_Q = \frac{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right)}{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \sqrt{\left(\frac{\pi}{\kappa}\right)^d |\Sigma_q|}. \quad (2)$$

$$\begin{aligned} Q^{\frac{2\kappa}{1+d\kappa}}(z | x) &= \frac{1}{Z_Q(\Sigma, \kappa)} \left(1 + \kappa_Q \left|(z - \mu_q)^\top \Sigma_Q^{-1}(z - \mu_q)\right|\right)^{-\frac{1+d\kappa_Q}{2\kappa_Q}} \\ \frac{1+d\kappa_Q}{2\kappa_Q} &= \left(\frac{1+d\kappa_q}{2\kappa_q}\right) \left(1 + \frac{2\kappa_q}{1+d\kappa_q}\right) \\ \kappa_Q &= \frac{\kappa_q}{1+2\kappa_q} \end{aligned} \quad (3)$$

To find (Σ_Q) from the given equation:

$$\kappa_q |(z - \mu_q)^T \Sigma_q^{-1}(z - \mu_q)| = \kappa_Q |(z - \mu_q)^T \Sigma_Q^{-1}(z - \mu_q)|$$

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

$$\kappa_q |\Sigma_q^{-1}| |(z - \mu_q)^T (z - \mu_q)| = \kappa_Q |\Sigma_Q^{-1}| |(z - \mu_q)^T (z - \mu_q)|$$

$$\kappa |\Sigma_q^{-1}| = \kappa_Q |\Sigma_Q^{-1}|$$

$$|\Sigma_Q^{-1}| = \frac{\kappa}{\kappa_Q} |\Sigma_q^{-1}|$$

Taking the determinant on both sides:

$$|\Sigma_Q|^{-1} = \frac{\kappa_q}{\kappa_Q} |\Sigma_q|^{-1}$$

$$|\Sigma_Q|^{-1} = \frac{\kappa_q}{\frac{\kappa_q}{1+2\kappa_q}} |\Sigma_q|^{-1} = (1+2\kappa_q) |\Sigma_q|^{-1} \quad (4)$$

$$Q^{\frac{2\kappa}{1+d\kappa}}(z | x) \equiv \frac{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)}{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right) \sqrt{\frac{\pi}{\kappa} |\Sigma_q|}} \left(1 + \frac{\kappa}{1+2\kappa} |(z - \mu_q)^\top (1+2\kappa) \Sigma_q^{-1}(z - \mu_q)|\right)^{-\frac{1+\kappa(d+2)}{2\kappa}} \quad (5)$$

$$= \frac{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)}{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right) \sqrt{\frac{\pi}{\kappa} |\Sigma_q|}} \left(1 + \kappa |(z - \mu_q)^\top \Sigma_q^{-1}(z - \mu_q)|\right)^{-\frac{1+2\kappa(d+2)}{2\kappa}} \quad (6)$$

2. Calculating the Kullback-Leibler divergence

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

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

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

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

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

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

$$\mathbb{E}_{Q(\frac{2\kappa}{1+d\kappa})(\mathbf{z} \mid \mathbf{x})} \left(\frac{1}{2} \left(((z - \mu_p)^\top \Sigma_p^{-1} (z - \mu_p)) \oplus_{2\kappa} \frac{A_p}{2} - \frac{1}{2} \left(((z - \mu_q)^\top \Sigma_q^{-1} (z - \mu_q)) \oplus_{2\kappa} \frac{A_q}{2} \right) \right) \right) \quad (7)$$

Take the first item from eq.7:

$$\begin{aligned} &\mathbb{E}_{Q(\frac{2\kappa}{1+d\kappa})(\mathbf{z} \mid \mathbf{x})} \left(\frac{1}{2} \left(((z - \mu_p)^\top \Sigma_p^{-1} (z - \mu_p)) \oplus_{2\kappa} \frac{A_p}{2} \right) \right) = \\ &\mathbb{E}_{Q(\frac{2\kappa}{1+d\kappa})(\mathbf{z} \mid \mathbf{x})} \left(\frac{1}{2} \left(((z - \mu_p)^\top \Sigma_p^{-1} (z - \mu_p)) + \frac{A_p}{2} + \frac{\kappa}{2} \left(((z - \mu_p)^\top \Sigma_p^{-1} (z - \mu_p))^2 A_p \right) \right) \right) \\ &\mathbb{E}_{Q(\frac{2\kappa}{1+d\kappa})(\mathbf{z} \mid \mathbf{x})} \left(\frac{1}{2} (z - \mu_p)^\top \Sigma_p^{-1} (z - \mu_p) \right) (1 + \kappa A_p) + \frac{A_p}{2} \end{aligned}$$

Let's define the centered variable:

$$y_p = z - \mu_p$$

The quantity we want to compute is:

$$\mathbb{E} [y_p^\top \Sigma_p^{-1} y_p]$$

Express the quadratic form using the trace

The identity for the trace of a scalar:

$$a^\top M a = \text{Tr} (M a a^\top)$$

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

$$y_p^\top \Sigma_p^{-1} y_p = \text{Tr} (\Sigma_p^{-1} y_p y_p^\top)$$

Taking the expectation:

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

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

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

$$Q^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z} \mid \mathbf{x}) \sim \frac{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)}{\Gamma\left(\frac{1+2\kappa}{2\kappa}\right) \sqrt{\left(\frac{\pi}{\kappa}\right)^d |\Sigma_q|}} (1 + \kappa |(z - \mu_q)^\top \Sigma_q^{-1} (z - \mu_q)|)^{-\frac{1+\kappa(d+2)}{2\kappa}} \quad (8)$$

The first term

Let an eigenvalue decomposition $\Sigma_q^{-1} = U_q \lambda_q^{-1} U_q^\top$ Applying this transformation:

$$\mathbb{E} [y_q y_q^\top] = \frac{\Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right) \frac{1}{2} \Sigma_q^{-1}}{\Gamma\left(\frac{2\kappa+1}{2\kappa}\right) \sqrt{\left(\frac{\pi}{\kappa}\right)^d |\Sigma_q|}} \int (y_q y_q^\top) (1 + \kappa |y_q^\top U_q \lambda_q^{-1} U_q^\top y_q|)^{-\frac{1+\kappa(d+2)}{2\kappa}} \quad (9)$$

Let $s_q = U_q^\top y_q, \Rightarrow y_q = U_q s_q$ and $ds_q = |U_q^\top| dy_q \Rightarrow ds_q = dy_q$ Applying in the 9

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

λ_q is Positive Definite Matrix,

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

We consider the integral:

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

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:

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

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:

$$\begin{aligned}
&= \prod_{r \neq i} \int_{-\infty}^{\infty} e^{-t\kappa \frac{s_{qr}^2}{\lambda_{qr}}} ds_r \cdot \int s_{qi}^2 e^{-t\kappa \frac{s_{qi}^2}{\lambda_{qi}}} ds_{qi} \\
&= \prod_{r \neq i} \frac{\sqrt{\pi}}{\sqrt{\frac{t\kappa}{\lambda_{qr}}}} \cdot \frac{\sqrt{\pi}}{2\sqrt{\frac{(t\kappa)^3}{\lambda_{qi}^3}}} \\
&= \prod_{r \neq i} \frac{\sqrt{\pi} \cdot \sqrt{\lambda_{qr}}}{\sqrt{t\kappa}} \cdot \frac{\sqrt{\pi} \sqrt{\lambda_{qi}}}{2\sqrt{(t\kappa)^3}} \cdot \lambda_{qi} \\
&= \frac{\sqrt{(\pi^d) \cdot \prod_{r=1}^d \lambda_{qr}}}{2\sqrt{(t\kappa)^{d+2}}} \cdot \lambda_{qi} \\
&= \frac{\sqrt{(\pi^d) |\Sigma_q|}}{2\sqrt{(t\kappa)^{d+2}}} \cdot \Sigma_q
\end{aligned}$$

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

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

Rearranging the constants:

$$\begin{aligned}
I &= \frac{\sqrt{(\pi^d) |\Sigma_q|}}{2\sqrt{(\kappa)^{d+2}} \Gamma(\alpha)} \cdot \Sigma_q \int_0^{\infty} t^{\alpha-1} (t)^{-\frac{d+2}{2}} e^{-t} dt. \\
I &= \frac{\sqrt{\pi^d |\Sigma_q|} \cdot \Sigma_q}{2\sqrt{(\kappa)^{d+2}} \Gamma(\alpha)} \int_0^{\infty} t^{\alpha-1-(d+2)/2} e^{-t} dt.
\end{aligned}$$

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:

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

$$\left(\frac{\sqrt{\pi^d |\Sigma_q|} \cdot \Sigma_q}{2\sqrt{(\kappa)^{d+2}} \Gamma\left(\frac{1+\kappa(d+2)}{2\kappa}\right)} \Gamma\left(\frac{1}{2\kappa}\right) \right) (1 + \kappa A_q) + A_q. \quad (10)$$

The second term

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

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

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

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:

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

Expand the Quadratic Form, expanding the expression:

$$(\mathbf{z} - \mu_p)^T \Sigma_p^{-1} (\mathbf{z} - \mu_p) = (\mathbf{z} - \mu_q)^T \Sigma_p^{-1} (\mathbf{z} - \mu_q) + 2(\mu_q - \mu_p)^T \Sigma_p^{-1} (\mathbf{z} - \mu_q) + (\mu_q - \mu_p)^T \Sigma_p^{-1} (\mu_q - \mu_p).$$

Taking the expectation on both sides:

$$\mathbb{E}_Q \left[(\mathbf{z} - \mu_p)^T \Sigma_p^{-1} (\mathbf{z} - \mu_p) \right] = \mathbb{E}_Q \left[(\mathbf{z} - \mu_q)^T \Sigma_p^{-1} (\mathbf{z} - \mu_q) \right] + 2(\mu_q - \mu_p)^T \Sigma_p^{-1} \mathbb{E}_Q [\mathbf{z} - \mu_q] + (\mu_q - \mu_p)^T \Sigma_p^{-1} (\mu_q - \mu_p).$$

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

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

Where I eq.10 is the previous result.

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

$$-\frac{d(1+\kappa A_q)}{2} + \frac{1+\kappa A_p}{2} ((\mu_p - \mu_q)^T \Sigma_p^{-1} (\mu_p - \mu_q) + \text{tr}(\Sigma_p^{-1} \Sigma_q)) - \frac{A_q}{2} + \frac{A_p}{2}$$

5. Reconstruction

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

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

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

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

We use the reparametrization trick: For $z \sim q(z|x)$, we can write

$$z = \mu_z + \Sigma_z^{\frac{1}{2}} \epsilon, \quad \epsilon \sim \mathcal{N}_\kappa(0, I_n).$$

$$\begin{aligned} \mu_{x|z} &= Dz + b_D \\ &= D(\mu_z + \Sigma_z^{\frac{1}{2}} \epsilon) + b_D \end{aligned}$$

$$(x - \mu_{x|z})^\top \Sigma_{x|z}^{-1} (x - \mu_{x|z})$$

$$\mu_{x|z} = Dz + b_D = D(\mu_z + \Sigma_z^{1/2} \epsilon) + b_D = D\mu_z + D\Sigma_z^{1/2} \epsilon + b_D$$

For simplicity, assume $(\Sigma_{x|z} = I)$ Expand the expression $(x - \mu_{x|z})$

$$x - \mu_{x|z} = x - (D\mu_z + D\Sigma_z^{1/2} \epsilon + b_D) = (x - D\mu_z - b_D) - D\Sigma_z^{1/2} \epsilon$$

Plug into the full quadratic form

$$(x - \mu_{x|z})^\top (x - \mu_{x|z}) = \left((x - D\mu_z - b_D) - D\Sigma_z^{1/2} \epsilon \right)^\top \left((x - D\mu_z - b_D) - D\Sigma_z^{1/2} \epsilon \right)$$

Use the identity:

$$(a - b)^\top(a - b) = a^\top a - 2a^\top b + b^\top b$$

Let: $a = x - D\mu_z - b_D$, $b = D\Sigma_z^{1/2}\epsilon$

Then:

$$= (x - D\mu_z - b_D)^\top(x - D\mu_z - b_D) - 2(x - D\mu_z - b_D)^\top D\Sigma_z^{1/2}\epsilon + \epsilon^\top (\Sigma_z^{1/2})^\top D^\top D\Sigma_z^{1/2}\epsilon$$

Take expectation over $\epsilon \sim \mathcal{N}_\kappa(0, I)$

We now compute:

$$\mathbb{E}_\epsilon [(x - \mu_{x|z})^\top(x - \mu_{x|z})] =$$

$$(x - D\mu_z - b_D)^\top(x - D\mu_z - b_D) - 2(x - D\mu_z - b_D)^\top D\Sigma_z^{1/2} \underbrace{\mathbb{E}[\epsilon]}_{=0} + \mathbb{E}[\epsilon^\top (\Sigma_z^{1/2})^\top D^\top D\Sigma_z^{1/2}\epsilon]$$

The second term vanishes (mean is zero). For the third term, we use:

$$\mathbb{E}_\epsilon[\epsilon^\top A\epsilon] = \frac{\text{Tr}(A)}{1 + (-n + d)\kappa}, \quad \text{if } \epsilon \sim \mathcal{N}_\kappa(0, I_n)$$

Here, $(A = (\Sigma_z^{1/2})^\top D^\top D\Sigma_z^{1/2})$, so:

$$\mathbb{E}_\epsilon[\epsilon^\top A\epsilon] = \frac{\text{Tr}(D\Sigma_z D^\top)}{1 + (-n + d)\kappa}$$

Final Result: (needs to be checked, since expectation is over \mathbf{Q} not \mathbf{q})

$$\mathbb{E}_{z_Q \sim Q^{(\frac{2\kappa}{1+d\kappa})}(z|x)} \left[\frac{1}{2} \ln_\kappa(p_\theta(x|z))^{\frac{-2}{1+d\kappa}} \right] = \frac{1}{2} ((x - D\mu_z - b_D)^\top(x - D\mu_z - b_D) + \text{Tr}(D\Sigma_{z_Q} D^\top)) \oplus_{2\kappa} \frac{A_{x|z_Q}}{2}$$

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

$z_Q \sim Q^{(\frac{2\kappa}{1+d\kappa})}(z|x) = \mathcal{N}_\kappa(\mu_{z_Q}, \Sigma_{z_Q})$, We assume that

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

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

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

$$\begin{aligned} \mathbb{E}_{z_Q \sim Q^{(\frac{2\kappa}{1+d\kappa})}(z|x)} \left[\frac{1}{2} \ln_{\kappa} p(x|z)^{\frac{-2}{1+d\kappa}} \right] \\ = \mathbb{E}_{z_Q \sim Q^{(\frac{2\kappa}{1+d\kappa})}(z|x)} \left[\frac{1}{2} \ln_{\kappa} p(x|z)^{\frac{-2}{1+d\kappa}} \right] \left[\frac{1}{2} (x - \bar{x}_{x|z})^{\top} (x - \bar{x}_{x|z}) \oplus_{2\kappa} \frac{A_{x|z}}{2} \right] \\ = \frac{1}{4} \mathbb{E}_{z_Q \sim Q^{(\frac{2\kappa}{1+d\kappa})}(z|x)} \left[\ln_{\kappa} p(x|z)^{\frac{-2}{1+d\kappa}} \right] \left[(x - \bar{x}_{x|z})^2 \oplus_{\kappa} A_{x|z} \right] \end{aligned}$$

Correction:

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

This research introduces the concept of coupled probability as a alternative to standard probability through the application of nonlinear statistical coupling theory. The distribution represented by (Q) represents a distribution that accounts for this coupling and is used as a weighting tool across probabilities. Our approach produces more stable expectation values because are applying coupling not to the probability density function (PDF) directly, but to the cumulative distribution function (CDF)

The coupled probability is defined as:

$$P_i^{(r)} \equiv \frac{p_i^{1+r}}{\sum_{j=1}^N p_j^{1+r}}, \quad \text{where } r = \frac{\alpha\kappa}{1+d\kappa}$$

The reweighted distribution acts as a relative risk bias measure (r) which models the nonlinear state interactions (coupling) from the original distribution. The primary advantage of employing (Q) lies in its ability to produce more stable estimation especially when compared to traditional methods that suffer from instability or divergence under high variance conditions.

One particularly elegant result is that when the coupling parameter $(\kappa) \rightarrow \infty$, the system naturally converges to a value where $(\kappa = \frac{1}{2})$, which guarantees a measurable mean and a finite variance across all (κ) values. The property enables machine learning algorithms to achieve robust and efficient learning because traditional methods often faced stability challenges.

In the following, we extend the VAE ELBO to the case where both the prior and posterior are modeled as coupled Gaussian distributions within the CVAE framework.

Theorem 0.1 (Coupled Free Energy for Multivariate Coupled Gaussian). *Let $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{z} \in \mathbb{R}^n$ be random vectors, and suppose that the posterior $q(\mathbf{z} | \mathbf{x})$ and the prior $p(\mathbf{z})$ are both multivariate coupled Gaussian, with their probability densities defined as:*

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \kappa) \equiv \left\{ \frac{1}{Z(\boldsymbol{\Sigma}, \kappa)} (1 + \kappa |(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})|)_+^{-\frac{1+d\kappa}{2\kappa}}, \quad \kappa \neq 0. \right.$$

Given a Coupled Variational Autoencoder (CVAE), the Coupled Free Energy (CFE) takes the following form:

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

such that:

- (a) *The coupled divergence term between the multivariate coupled Gaussians in the CFE, given by:*

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

simplifies to the following closed-form expression:

$$-\frac{d(1 + \kappa A_q)}{2} + \frac{1 + \kappa A_p}{2} ((\boldsymbol{\mu}_p - \boldsymbol{\mu}_q)^\top \boldsymbol{\Sigma}_p^{-1}(\boldsymbol{\mu}_p - \boldsymbol{\mu}_q) + \text{tr}(\boldsymbol{\Sigma}_p^{-1} \boldsymbol{\Sigma}_q)) - \frac{A_q}{2} + \frac{A_p}{2}.$$

where A_q and A_p denote the normalization terms of the coupled posterior and prior distributions, respectively, and $\text{tr}(\cdot)$ denotes the trace operator.

- (b) *The coupled reconstruction term is given by:*

$$-\frac{1}{2} \mathbb{E}_{\mathbf{z} \sim Q_{\phi}^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z}|\mathbf{x})} \left[\ln_{\kappa} \left(\exp_{\kappa}^{\frac{-(1+d\kappa)}{2}} \left((\mathbf{x} - \bar{\mathbf{x}}_{x|z})^\top \boldsymbol{\Sigma}_{x|z}^{-1}(\mathbf{x} - \bar{\mathbf{x}}_{x|z}) \right) \oplus_{\kappa} \left(\ln_{\kappa} \left(\frac{1}{Z_q} \right) \right)^{\left(\frac{-2}{1+d\kappa}\right)} \right) \right]$$

and is defined based on the coupled likelihood model $p(\mathbf{x} | \mathbf{z})$. The term simplifies to:

$$-\mathbb{E}_{\mathbf{z} \sim Q_{\phi}^{\left(\frac{2\kappa}{1+d\kappa}\right)}(\mathbf{z}|\mathbf{x})} \left[\frac{1}{2} \left((\mathbf{x} - \bar{\mathbf{x}}_{x|z})^\top \boldsymbol{\Sigma}_{x|z}^{-1}(\mathbf{x} - \bar{\mathbf{x}}_{x|z}) \oplus_{\kappa} A_{x|z} \right) \right]$$

This result provides a closed-form expression for the coupled Free Energy in terms of the coupled divergence and reconstruction cost. It recovers the standard VAE formulation in the limit as $(\kappa \rightarrow 0)$.

The derivation will be included in a follow-up manuscript currently in preparation.

A_q and A_p denote, respectively, $\left[\ln_{\kappa} \left(\frac{1}{Z_q} \right) \right]^{-\frac{2}{1+d\kappa}}$ and $\left[\ln_{\kappa} \left(\frac{1}{Z_p} \right) \right]^{-\frac{2}{1+d\kappa}}$.