

# 1 Objectives

Calculate mutual information using numerical integration for a complex, linearized signal model  $\vec{z}(\mu, \vec{k})$ .

$$\vec{z}(\mu, \vec{k}) = \int_{\Omega} M(\vec{x}) e^{-s(\mu, \vec{x})} e^{-2\pi i \vec{k} \cdot \vec{x}} d\vec{x} + \nu = \mathcal{G}(\mu, \vec{k}) + \nu \quad \nu \sim \mathcal{N}(\vec{0}, \Sigma_{\nu}) \quad \Sigma_{\nu} = \begin{bmatrix} \sigma_{\nu,r}^2 & 0 \\ 0 & \sigma_{\nu,i}^2 \end{bmatrix} \quad (1)$$

Assume normal distribution for the model parameter, optical attenuation coefficient  $\mu$ .

$$\mu \sim \mathcal{N}(\vec{\mu}_{\mu}, \Sigma_{\mu}) \quad \vec{\mu}_{\mu} = \begin{bmatrix} \mu_{\mu,r} \\ 0 \end{bmatrix} \quad \Sigma_{\mu} = \begin{bmatrix} \sigma_{\mu,r}^2 & 0 \\ 0 & 0 \end{bmatrix} \quad (2)$$

# 2 Work

Mutual information can be calculated from the difference of entropies. In the case of normal distributions, it is necessary to know the covariance matrices of the distributions  $p(\mathbf{z})$  and  $p(\mathbf{z}|\mu)$ .

$$I(\mu; z) = H(z) - H(z|\mu) = \frac{1}{2} \ln \left( (2\pi e)^2 \cdot |\Sigma_z| \right) - \frac{1}{2} \ln \left( (2\pi e)^2 \cdot |\Sigma_{\nu}| \right) \quad (3)$$

The real and imaginary parts of the signal vary independently of one another, so the covariance matrix  $\Sigma_{\nu}$  is diagonal.

$$\Sigma_{\nu} = \begin{bmatrix} \sigma_{\nu,r}^2 & 0 \\ 0 & \sigma_{\nu,i}^2 \end{bmatrix} \quad (4)$$

The quantities  $\sigma_{\nu,r}^2$  and  $\sigma_{\nu,i}^2$  are measured, so the determinant  $|\Sigma_{\nu}|$  is known, and one term of the mutual information expression can be immediately calculated.

$$|\Sigma_{\nu}| = \sigma_{\nu,r}^2 \sigma_{\nu,i}^2 \quad (5)$$

Obtaining  $\Sigma_z$  requires the bulk of the calculations. From the definition of covariance matrix,

$$\Sigma_z = \mathbb{E} \left[ (\vec{z} - \mathbb{E}[\vec{z}]) (\vec{z} - \mathbb{E}[\vec{z}])^T \right] = \mathbb{E} \left[ \begin{pmatrix} z_r - \bar{z}_r \\ z_i - \bar{z}_i \end{pmatrix} \begin{pmatrix} z_r - \bar{z}_r & z_i - \bar{z}_i \end{pmatrix} \right] \quad (6)$$

Written explicitly in matrix form, the covariance matrix for  $\Sigma_z$  is

$$\Sigma_z = \begin{bmatrix} \mathbb{E} \left[ (z_r - \bar{z}_r)^2 \right] & \mathbb{E} \left[ (z_r - \bar{z}_r) (z_i - \bar{z}_i) \right] \\ \mathbb{E} \left[ (z_i - \bar{z}_i) (z_r - \bar{z}_r) \right] & \mathbb{E} \left[ (z_i - \bar{z}_i)^2 \right] \end{bmatrix} \quad (7)$$

The determinant of the covariance matrix  $\Sigma_z$  can be calculated with knowledge of  $\mathbb{E} \left[ (z_r - \bar{z}_r)^2 \right]$ ,  $\mathbb{E} \left[ (z_i - \bar{z}_i)^2 \right]$ , and  $\mathbb{E} \left[ (z_r - \bar{z}_r) (z_i - \bar{z}_i) \right]$ .

$$|\Sigma_z| = \mathbb{E} \left[ (z_r - \bar{z}_r)^2 \right] \mathbb{E} \left[ (z_i - \bar{z}_i)^2 \right] - (\mathbb{E} \left[ (z_r - \bar{z}_r) (z_i - \bar{z}_i) \right])^2 \quad (8)$$

$\mathbb{E} \left[ (z_r - \bar{z}_r)^2 \right]$ ,  $\mathbb{E} \left[ (z_i - \bar{z}_i)^2 \right]$ , and  $\mathbb{E} \left[ (z_r - \bar{z}_r) (z_i - \bar{z}_i) \right]$  are difficult to obtain analytically and must be calculated using Gauss-Hermite quadrature. The general form for Gauss-Hermite quadrature is

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{i=1}^n w_i f(x_i) \quad (9)$$

It is necessary to evaluate the distribution  $p(\mathbf{z})$  before Gauss-Hermite quadrature can be performed to obtain the covariance matrix of  $z$ . The distribution  $p(\mathbf{z})$  can be shown to be equal to the product of three one-dimensional normal distributions.

$$p(\mathbf{z}) = \int p(\mathbf{z}|\mu) p(\mu) d\mu = \int S \frac{1}{\sqrt{2\pi\sigma_{tot}^2}} \exp \left( -\frac{(\mu - \mu_{tot})^2}{2\sigma_{tot}^2} \right) d\mu \quad (10)$$

The statistics for normal distribution which results from the product of the three normal distributions are given by the following equations. It should be noted that the normal distribution resulting from the product of multiple normal distributions is not normalized, which can be seen from the factor  $S$ .

$$\sigma_{tot}^2 = \left( \frac{a_r^2}{\sigma_{\nu,r}^2} + \frac{a_i^2}{\sigma_{\nu,i}^2} + \frac{1}{\sigma_\mu^2} \right)^{-1} \quad (11)$$

$$\mu_{tot} = \left( \frac{\mu_\mu}{\sigma_\mu^2} + \frac{(z_r - b_r)/a_r}{\sigma_{\nu,r}^2/a_r} + \frac{(z_i - b_i)/a_i}{\sigma_{\nu,i}^2/a_i} \right) \sigma_{tot} \quad (12)$$

$$S = \left[ 2\pi \sqrt{\frac{(\sigma_{\nu,r}/a_r)^2 (\sigma_{\nu,i}/a_i)^2 \sigma_\mu^2}{\sigma_{tot}^2}} \right]^{-1} \exp \left[ -\frac{1}{2} \left( \frac{(z_r - b_r)^2}{\sigma_{\nu,r}^2} + \frac{(z_i - b_i)^2}{\sigma_{\nu,i}^2} + \frac{\mu_\mu^2}{\sigma_\mu^2} - \frac{\mu_{tot}^2}{\sigma_{tot}^2} \right) \right] \quad (13)$$

Change of variables for Gauss-Hermite quadrature is done by the following equations:

$$x = \frac{\mu - \mu_{tot}}{\sqrt{2}\sigma_{tot}} \quad (14)$$

$$\mu = \sqrt{2}\sigma_{tot}x + \mu_{tot} \quad (15)$$

$$d\mu = \sqrt{2}\sigma_{tot}dx \quad (16)$$

The distribution  $p(\mathbf{z})$  can be evaluated in this manner, but the result is not physical and may not be useful. At the very least it is very difficult to obtain further statistics about the distribution  $p(\mathbf{z})$ .

$$p(\mathbf{z}) = S \frac{\sqrt{2}\sigma_{tot}}{\sqrt{2\pi}\sigma_{tot}^2} \int_{-\infty}^{\infty} e^{-x^2} dx = \frac{S}{\sqrt{\pi}} (\sqrt{\pi}) = S \quad (17)$$

$$p(\mathbf{z}) = \left[ 2\pi \sqrt{\frac{(\sigma_{\nu,r}/a_r)^2 (\sigma_{\nu,i}/a_i)^2 \sigma_\mu^2}{\sigma_{tot}^2}} \right]^{-1} \exp \left[ -\frac{1}{2} \left( \frac{(z_r - b_r)^2}{\sigma_{\nu,r}^2} + \frac{(z_i - b_i)^2}{\sigma_{\nu,i}^2} + \frac{\mu_\mu^2}{\sigma_\mu^2} - \frac{\mu_{tot}^2}{\sigma_{tot}^2} \right) \right] \quad (18)$$

### 3 Directions

The fact that  $p(\mathbf{z})$  depends on values of  $\mu < 0$  causes problems. This results from  $p(\mu) \sim \mathcal{N}$ . Gauss-Hermite quadrature might not be an option for finite integration bounds. May need to redo the calculation using a uniform distribution for  $p(\mu)$  and Gauss-Legendre quadrature.

An analytical evaluation of mutual information  $I(\mu; z)$  might be possible if it is acceptable to represent the optical attenuation coefficient as a complex number, i.e.  $\vec{\mu} = \begin{bmatrix} \mu_r \\ 0 \end{bmatrix}$ . This still leaves the issue of  $\mu$  being allowed to be negative, which is unphysical.

The change of variables from  $p(\mathbf{z}|\mu) = \mathcal{N}_{\vec{z}}(\vec{a}(\mu - \hat{\mu}) + \vec{b}, \Sigma_\nu)$  to  $p(\mathbf{z}|\mu) = \mathcal{N}_\mu$  is ambiguous. Is it then a one-dimensional normal distribution? Does this treat the complex element correctly? It may be better to investigate the complex normal distribution from the start.

Also, need to ensure that  $p(\mathbf{z})$  is normalized. Entropy of a normal distribution is likely different for an unnormalized distribution.