

# $\Sigma$ GS node: Derivations

Bart van Erp

May 25, 2021

## 1 Overview

The node factor is given by

$$p(x \mid s, n) = \mathcal{N}_C(x \mid 0, e^s + e^n, 0), \quad (1)$$

where

$$\mathcal{N}_C(x \mid \mu, \sigma^2, 0) = \frac{1}{\pi \sigma^2} e^{-\frac{1}{\sigma^2} |x - \mu|^2}, \quad (2)$$

so

$$p(x \mid s, n) = \frac{1}{\pi(e^s + e^n)} e^{-\frac{1}{e^s + e^n} |x|^2} \quad (3)$$

The approximate posterior is given by

$$q(x, s, n) = q(x)q(s)q(n), \quad (4)$$

where

$$q(x) = \mathcal{N}_C(x \mid m_x, v_x, 0) \quad (5)$$

$$q(s) = \mathcal{N}(s \mid m_s, v_s) \quad (6)$$

$$q(n) = \mathcal{N}(n \mid m_n, v_n) \quad (7)$$

The log-pdf can be found as

$$\ln p(x \mid s, n) = -\ln(e^s + e^n) - \frac{1}{e^s + e^n} |x|^2 + \text{const} \quad (8)$$

### 1.1 Approximations

VMP will result in intractable computations because of the non-linear term in the node factor. Therefore we will approximate this non-linear function by a first-order (vector) Taylor expansion. The individual derivatives are given as

$$\frac{\partial}{\partial s} \ln(e^s + e^n) = \frac{e^s}{e^s + e^n} = \frac{1}{1 + e^{n-s}} = \sigma(s - n), \quad (9)$$

$$\frac{\partial}{\partial n} \ln(e^s + e^n) = \frac{e^n}{e^s + e^n} = \frac{1}{1 + e^{s-n}} = \sigma(n - s), \quad (10)$$

$$\frac{\partial}{\partial s} \frac{1}{e^s + e^n} = \frac{-e^s}{(e^s + e^n)^2} = -e^{-s} \sigma(s - n)^2 \quad (11)$$

and

$$\frac{\partial}{\partial n} \frac{1}{e^s + e^n} = \frac{-e^n}{(e^s + e^n)^2} = -e^{-n} \sigma(n - s)^2 \quad (12)$$

where we specify the sigmoid function  $\sigma(x)$  as

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (13)$$

## 2 Message $\nu(x)$

The message  $\nu(x)$  can be calculated as

$$\ln \nu(x) = \mathbb{E}_{q(s)q(n)} [\ln p(x | s, n)] \quad (14)$$

The exponential terms cause issues. Let us focus on the first term  $\log(e^s + e^n)$ . The expectation with respect to this term cannot be calculated exactly. Instead we need to approximate it (see Zotero). Here we approximate it with the first-order Taylor expansion around its mean as

$$\ln(e^s + e^n) \approx \ln(e^{m_s} + e^{m_n}) + \sigma(m_s - m_n)(s - m_s) + \sigma(m_n - m_s)(n - m_n). \quad (15)$$

Now the expectation simply becomes

$$\mathbb{E}_{q(s)q(n)} [\ln p(x | s, n)] \approx \ln(e^{m_s} + e^{m_n}). \quad (16)$$

Similarly the second term in the log-pdf can be approximated as

$$\mathbb{E}_{q(s)q(n)} \left[ \frac{1}{e^s + e^n} \right] \approx \frac{1}{e^{m_s} + e^{m_n}}. \quad (17)$$

Using the above approximation we can approximate the log-message as

$$\begin{aligned} \ln \nu(x) &= \mathbb{E}_{q(s)q(n)} [\ln p(x | s, n)] \\ &\approx -\frac{1}{e^{m_s} + e^{m_n}} |x|^2 + \text{const} \end{aligned} \quad (18)$$

and identify it as

$$\boxed{\nu(x) \propto \mathcal{N}_{\mathcal{C}}(x | 0, e^{m_s} + e^{m_n}, 0)} \quad (19)$$

## 3 Message $\nu(s)$

For the message  $\nu(s)$  we again need to perform some approximations. Below you can find the approximations and their pros and cons. The expectation with respect to  $x$  is well-defined and simplifies the problem to

$$\begin{aligned} \ln \nu(s) &= \mathbb{E}_{q(n)q(x)} [\ln p(x | s, n)] + \text{const} \\ &= \mathbb{E}_{q(n)} \left[ -\ln(e^s + e^n) - \frac{1}{e^s + e^n} (v_X + |m_x|^2) \right] + \text{const} \end{aligned} \quad (20)$$

### 3.1 Solution 1: Full Taylor expansion + LaPlace marginal

If we perform a full Taylor expansion (just like by  $\nu(x)$ ) the message can be found as

$$\ln \nu(s) \approx -s \left( \sigma(m_s - m_n) - e^{-m_s} \sigma(m_s - m_n)^2 \right) \quad (21)$$

This message has a linear log-pdf and is improper. An LaPlace approximation for this message is not useful, however, the resulting marginal  $q(s)$  can be approximated using LaPlace.

### 3.2 Solution 2: Partial Taylor expansion + LaPlace message

Alternatively, we can linearize the log-message only with respect to  $n$ , yielding

$$\ln \nu(s) = -\ln(e^s + e^{m_n}) - \frac{1}{e^s + e^{m_n}}(v_y + |m_x|^2) + \text{const} \quad (22)$$

This message has a sigmoidal shape. The log-pdf is shaped horizontally and bends of linearly. This shape is not perse appropriate for approximating it with a Gaussian distribution, but if we want, we could try. The parial derivative can be found as

$$\frac{\partial}{\partial s} \ln \nu(s) = \frac{e^s}{e^s + e^{m_n}} \left( \frac{v_x + |m_x|^2}{e^s + e^{m_n}} - 1 \right) \quad (23)$$

The mean can be found by calculating the corresponding zeros, which is found as

$$s_0 = \ln(v_x + |m_x|^2 - e^{m_n}), \quad (24)$$

from which it can be seen that this mode/mean is not always defined.

### 3.3 Solution 3: Partial Taylor expansion + LaPlace marginal

To make sure we can calculate the resulting marginal, the message

$$\ln \nu(s) = -\ln(e^s + e^{m_n}) - \frac{1}{e^s + e^{m_n}}(v_x + |m_x|^2) + \text{const} \quad (25)$$

can be propagated along the edge. Using the incoming message, the marginal distribution can be expressed as a function of  $s$ . Using numerical optimization (as no closed-form expression is present) the mode and variance around the mode can be determined, allowing for a Gaussian LaPlace approximation of the marginal.

### 3.4 Alternative solutions

Besides the LaPlace approximation, we can solve the problem by moment matching (using for example quadrature integration or importance sampling).

## 4 Message $\nu(n)$

By argument of symmetry, above strategies can be followed for the message  $\nu(n)$ .

## 5 Free energy

The local free energy can be calculated (by a Taylor expansion) as

$$\begin{aligned} U &= -\mathbb{E}_{q(x)q(s)q(n)} [\ln p(x \mid s, n)] \\ &\approx \ln(\pi) + \ln(e^{m_s} + e^{m_n}) + \frac{v_x + |m_x|^2}{e^{m_s} + e^{m_n}} \end{aligned} \tag{26}$$