

A New Kalman Filter for Sigmoidal Nonlinearities

Dean R. Freestone, Parham Aram, David B. Grayden, Dragan Nesic,

2012-10-09

Abstract

1 Introduction

This paper describes a new Kalman filter that is constructed for systems with sigmoidal nonlinearities. In this paper we will provide a derivation for the mean of a Gaussian random variable that is transformed by an

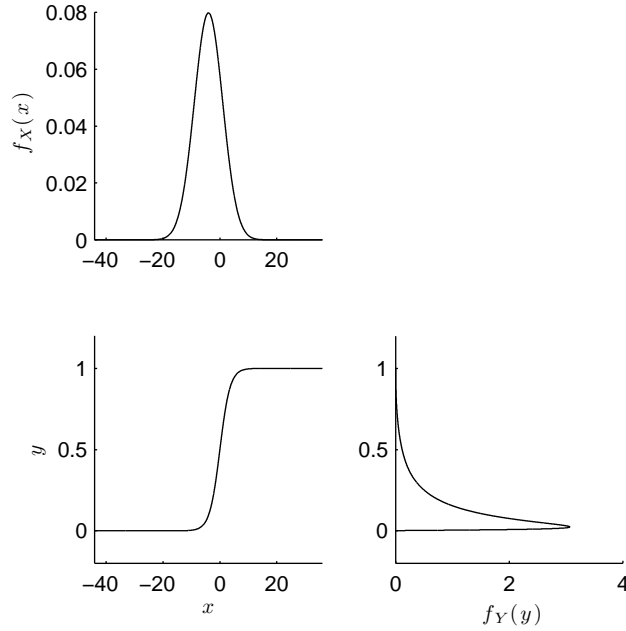


Figure 1: Transformation of a Gaussian through a sigmoid.

1.1 Expected Value of a Gaussian Transformed by an Error Function Sigmoid

Let's look at the transformation of a Gaussian random variable through a typical sigmoidal function. Let our Gaussian RV be described by the probability density function

$$f_X(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(x - \bar{x})^2}{2\sigma_x^2}\right). \quad (1)$$

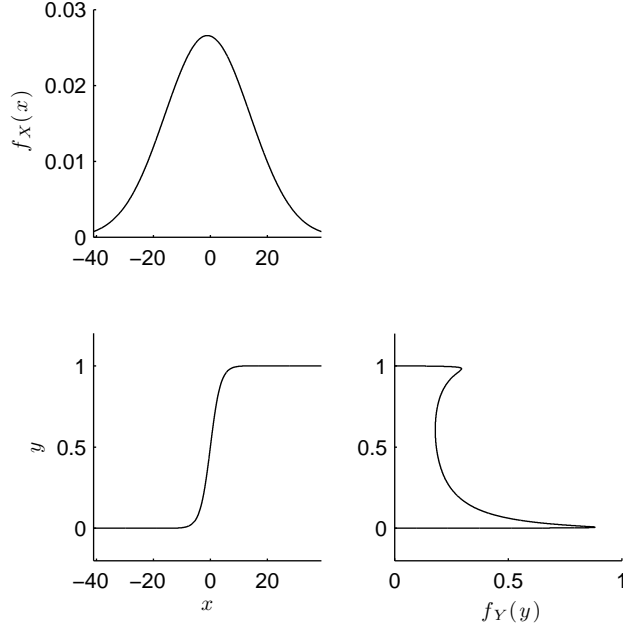


Figure 2: **Transformation of a Gaussian through a sigmoid.**

An alternate form of the sigmoid function than the one that is typically used is

$$g(v) = \frac{1}{\sqrt{2\pi}\varsigma} \int_{-\infty}^v \exp\left(-\frac{(z-v_0)^2}{2\varsigma^2}\right) dz \quad (2)$$

$$= \frac{1}{2} \left(\operatorname{erf}\left(\frac{v-v_0}{\varsigma\sqrt{2}}\right) + 1 \right). \quad (3)$$

Now we consider the mapping of the GRV (defined in equation ??) through this new sigmoid. In a similar vein as before, we will first consider a simple case where $v_0 = 0$ and $\varsigma = 1$. So we have

$$Y = g(X) \quad (4)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{z^2}{2}\right) dz \quad (5)$$

$$= \frac{1}{2} \left(\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) + 1 \right) \quad (6)$$

and

$$f_Y(y) = |h'(y)| f_X[h(y)] \quad (7)$$

$$= \frac{h'(y)}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(h(y) - \bar{x})^2}{2\sigma_x^2}\right) \quad (8)$$

where

$$h(y) = g^{-1}(Y). \quad (9)$$

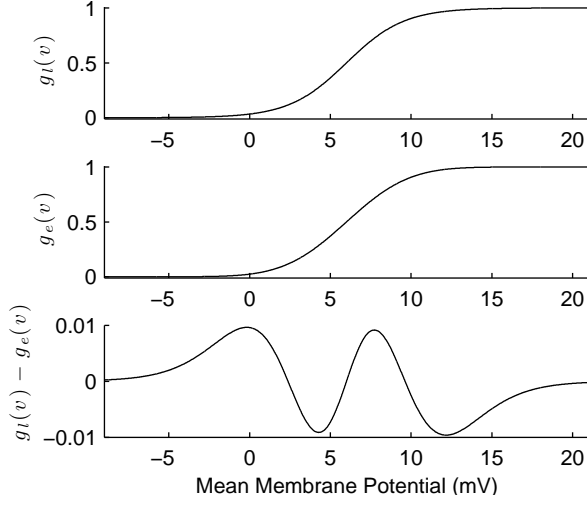


Figure 3: caption

As before, we have the expected value of the transformed Gaussian as

$$\mathbb{E}(y) = \int_{-\infty}^{\infty} y f_Y(y) dy \quad (10)$$

$$= \int_{-\infty}^{\infty} \frac{y h'(y)}{\sigma_x \sqrt{2\pi}} \exp\left(-\frac{(h(y) - \bar{x})^2}{2\sigma_x^2}\right) dy. \quad (11)$$

Now the papers by Brenden J. Frey (Variational inference for continuous sigmoidal Bayesian networks, 1999 and Variational learning in nonlinear Gaussian belief networks, Neural Computation 1999) provide a closed-form solution to the expectation. This solution is

$$\mathbb{E}(y) = g\left(\frac{\bar{x}}{\sqrt{1 + \sigma^2}}\right) \quad (12)$$

$$= \frac{1}{2} \left(\operatorname{erf}\left(\frac{\bar{x}}{\sqrt{2}\sqrt{1 + \sigma^2}}\right) + 1 \right). \quad (13)$$

This is shown in Figure ??, where the a Monte-Carlo simulation was used to test the analytic result. The predicted expected value and the actual numerical expected value correspond nicely.

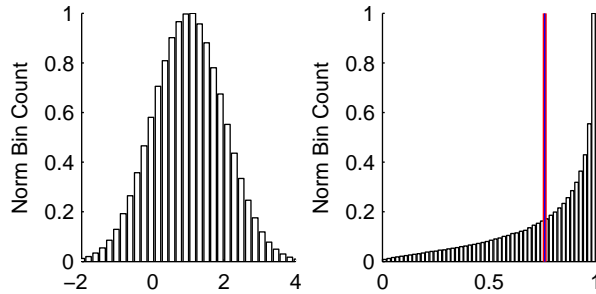


Figure 4: Monte Carlo simulation result confirming the analytic solution.

The derivation proceeds as follows. The expected value is defined by

$$\mathbb{E}[y] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (14)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^x \exp\left(-\frac{z^2}{2}\right) f_X(x) dz dx. \quad (15)$$

Now we can make the substitution $z = w - x$ to get x out of the integral terminal giving

$$\mathbb{E}[y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \int_{-\infty}^0 \exp\left(-\frac{(w-x)^2}{2}\right) f_X(x) dw dx. \quad (16)$$

The order of integration can be changed without changing the limits of integration giving

$$\mathbb{E}[y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp\left(-\frac{(w-x)^2}{2}\right) f_X(x) dx dw \quad (17)$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp\left(-\frac{(w-x)^2}{2}\right) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx dw \quad (18)$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp\left(-\frac{(w-x)^2}{2} - \frac{(x-\mu)^2}{2\sigma^2}\right) dx dw \quad (19)$$

From Wolfram Alpha, our target solution is

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(w-x)^2}{2} - \frac{(x-\mu)^2}{2\sigma^2}\right) dx = \frac{\sqrt{2\pi}\sigma}{\sqrt{\sigma^2+1}} \exp\left(-\frac{(w-\mu)^2}{2(\sigma^2+1)}\right), \quad (20)$$

which leads to

$$\mathbb{E}[y] = \frac{1}{2\pi} \frac{\sqrt{2\pi}}{\sqrt{\sigma^2+1}} \int_{-\infty}^0 \exp\left(-\frac{(w-\mu)^2}{2(\sigma^2+1)}\right) dw \quad (21)$$

Now we need to integrate out x , so we collect all the x -related terms

$$\mathbb{E}[y] = \frac{1}{2\pi\sigma} \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(w^2 - 2wx + x^2) - \frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2)\right) dx dw \quad (22)$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}(\sigma^2 w^2 - \sigma^2 2wx + \sigma^2 x^2 + x^2 - 2\mu x + \mu^2)\right) dx dw \quad (23)$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^0 \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}((\sigma^2+1)x^2 - 2\mu x - \sigma^2 2wx + \sigma^2 w^2 + \mu^2)\right) dx dw \quad (24)$$

$$= \frac{1}{2\pi\sigma} \int_{-\infty}^0 \exp\left(-\frac{1}{2\sigma^2}(\sigma^2 w^2 + \mu^2)\right) \int_{-\infty}^{\infty} \exp\left(-\frac{\sigma^2+1}{2\sigma^2}x^2 + \frac{\sigma^2 w + \mu}{\sigma^2}x\right) dx dw \quad (25)$$

$$(26)$$

Now we integrate out x . Let

$$\int_{-\infty}^{\infty} \exp \left(-\frac{\sigma^2 + 1}{2\sigma^2} x^2 + \frac{\sigma^2 w + \mu}{\sigma^2} x \right) dx = \int_{-\infty}^{\infty} \exp (ax^2 + bx) dx \quad (27)$$

$$= \int_{-\infty}^{\infty} \exp \left(a \left(x + \frac{b}{2a} \right)^2 + \frac{b^2}{4a} \right) dx \quad (28)$$

$$= \exp \left(-\frac{b^2}{4a} \right) \int_{-\infty}^{\infty} \exp \left(\frac{\left(x + \frac{b}{2a} \right)^2}{a^{-1}} \right) dx \quad (29)$$

$$= \exp \left(-\frac{b^2}{4a} \right) \frac{1}{\sqrt{-a}} \int_{-\infty}^{\infty} \exp (-c^2) dc \quad (30)$$

$$= \frac{\sqrt{\pi}}{\sqrt{-a}} \exp \left(-\frac{b^2}{4a} \right), \quad (31)$$

where

$$a = -\frac{\sigma^2 + 1}{2\sigma^2} \quad (32)$$

$$b = \frac{\sigma^2 w + \mu}{\sigma^2} \quad (33)$$

$$c = \frac{x + \frac{b}{2a}}{\sqrt{-a^{-1}}} \quad (34)$$

$$\frac{dc}{dx} = \frac{1}{\sqrt{-a^{-1}}} \quad (35)$$

$$dx = dc \sqrt{-a^{-1}}. \quad (36)$$

Now substituting back to replace a and b we get

$$\frac{\sqrt{\pi}}{\sqrt{-a}} \exp \left(-\frac{b^2}{4a} \right) = \frac{\sqrt{\pi}}{\sqrt{\frac{\sigma^2 + 1}{2\sigma^2}}} \exp \left(\frac{\left(\frac{\sigma^2 w + \mu}{\sigma^2} \right)^2}{\frac{4(\sigma^2 + 1)}{2\sigma^2}} \right) \quad (37)$$

$$= \frac{\sqrt{2\pi}\sigma}{\sqrt{\sigma^2 + 1}} \exp \left(\frac{\sigma^2}{2(\sigma^2 + 1)} \left(\frac{\sigma^2 w + \mu}{\sigma^2} \right)^2 \right) \quad (38)$$

$$= \frac{\sqrt{2\pi}\sigma}{\sqrt{\sigma^2 + 1}} \exp \left(\frac{(\sigma^2 w + \mu)^2}{2\sigma^2(\sigma^2 + 1)} \right) \quad (39)$$

$$(40)$$

Now putting it back together again (this agrees with Wolfram Alpha) we get

$$\mathbb{E}[y] = \frac{1}{2\pi\sigma} \int_{-\infty}^0 \exp \left(-\frac{1}{2\sigma^2} (\sigma^2 w^2 + \mu^2) \right) \frac{\sqrt{2\pi}\sigma}{\sqrt{\sigma^2 + 1}} \exp \left(\frac{(\mu + \sigma^2 w)^2}{2\sigma^2(\sigma^2 + 1)} \right) dw \quad (41)$$

$$= \frac{1}{2\pi\sigma} \frac{\sqrt{2\pi}\sigma}{\sqrt{\sigma^2 + 1}} \int_{-\infty}^0 \exp \left(-\frac{1}{2\sigma^2} (\sigma^2 w^2 + \mu^2) \right) \exp \left(\frac{(\mu + \sigma^2 w)^2}{2\sigma^2(\sigma^2 + 1)} \right) dw \quad (42)$$

$$= \frac{1}{2\pi} \frac{\sqrt{2\pi}}{\sqrt{\sigma^2 + 1}} \int_{-\infty}^0 \exp \left(-\frac{(\sigma^2 + 1)(\sigma^2 w^2 + \mu^2)}{2\sigma^2(\sigma^2 + 1)} + \frac{(\mu + \sigma^2 w)^2}{2\sigma^2(\sigma^2 + 1)} \right) dw \quad (43)$$

Now we will attempt to expand and simplify the exponent

$$-\frac{(\sigma^2 + 1)(\sigma^2 w^2 + \mu^2)}{2\sigma^2(\sigma^2 + 1)} + \frac{(\mu - \sigma^2 w)^2}{2\sigma^2(\sigma^2 + 1)} = -\frac{1}{2\sigma^2(\sigma^2 + 1)} \left((\sigma^2 + 1)(\sigma^2 w^2 + \mu^2) - (\mu + \sigma^2 w)^2 \right) \quad (44)$$

$$= -\frac{1}{2\sigma^2(\sigma^2 + 1)} (\sigma^4 w^2 + \sigma^2 \mu^2 + \sigma^2 w^2 + \mu^2 - (\mu^2 + 2\mu\sigma^2 w + \sigma^4 w^2)) \quad (45)$$

$$= -\frac{1}{2\sigma^2(\sigma^2 + 1)} (\sigma^2 \mu^2 + \sigma^2 w^2 - 2\mu\sigma^2 w) \quad (46)$$

$$= -\frac{1}{2(\sigma^2 + 1)} (\mu^2 + w^2 - 2\mu w) \quad (47)$$

$$= -\frac{(w - \mu)^2}{2(\sigma^2 + 1)} \quad (48)$$

Now putting it back together we get (this agrees with Wolfram too!)

$$\mathbb{E}[y] = \frac{1}{2\pi} \frac{\sqrt{2\pi}}{\sqrt{\sigma^2 + 1}} \int_{-\infty}^0 \exp\left(-\frac{(w - \mu)^2}{2(\sigma^2 + 1)}\right) dw \quad (49)$$

Now let

$$z = \frac{w - \mu}{\sqrt{\sigma^2 + 1}} \quad (50)$$

$$\frac{dz}{dw} = \frac{1}{\sqrt{\sigma^2 + 1}} \quad (51)$$

$$dw = \sqrt{\sigma^2 + 1} dz, \quad (52)$$

giving

$$\mathbb{E}[y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\mu}{\sqrt{\sigma^2 + 1}}} \exp\left(-\frac{z^2}{2}\right) dz \quad (53)$$

$$= g\left(\frac{\mu}{\sqrt{\sigma^2 + 1}}\right), \quad (54)$$

as required.

As mentioned above, the more general form of the error function sigmoid has slope (variance) and threshold (mean) parameters. This is

$$g(v) = \frac{1}{2} \left(\operatorname{erf}\left(\frac{v - v_0}{\varsigma\sqrt{2}}\right) + 1 \right). \quad (55)$$

It can easily be shown that the expected value of a Gaussian (the membrane voltage v) transformed by this guy is

$$\mathbb{E}[g(v)] = \frac{1}{2} \left(\operatorname{erf}\left(\frac{\mu - v_0}{\sqrt{2(\varsigma^2 + \sigma^2)}}\right) + 1 \right). \quad (56)$$

1.2 The Variance

1.2.1 Approximating the Error Function

Recall that

$$g(v) = \frac{1}{\sqrt{2\pi}\varsigma} \int_{-\infty}^v \exp\left(-\frac{(z-v_0)^2}{2\varsigma^2}\right) dz \quad (57)$$

$$= \frac{1}{2} \left(\operatorname{erf}\left(\frac{v-v_0}{\varsigma\sqrt{2}}\right) + 1 \right). \quad (58)$$

A recent paper has provided Chernoff-type upper and lower bounds on the Gaussian error function (Seok-Ho Chang et al. 2011, IEEE Trans. on Comms., Chernoff-Type Bounds for the Gaussian Error Function)

Now

$$g(v) \leq a_1 \exp(-b_1 z^2), \quad z - v_0 < 0 \quad (59)$$

$$\leq 1 - a_2 \exp(-b_2 z^2), \quad z - v_0 > 0, \quad (60)$$

where

$$z = \frac{v - v_0}{\varsigma} \quad (61)$$

$$a_1 = 0.5 \quad (62)$$

$$b_1 = 0.5 \quad (63)$$

$$a_2 = \sqrt{\frac{2e}{\pi}} \sqrt{\frac{b_2 - 1}{b_2}} \quad (64)$$

$$b_2 = 1.08 \quad (65)$$

$$\mathbb{E}[y^2] \leq \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{v_0} \left(a_1 \exp\left(-b_1 \frac{(x-v_0)^2}{\varsigma^2}\right) \right)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \quad (66)$$

$$+ \frac{1}{\sqrt{2\pi}\sigma} \int_{v_0}^{\infty} \left(1 - a_2 \exp\left(-b_2 \frac{(x-v_0)^2}{\varsigma^2}\right) \right)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx. \quad (67)$$

Now we need to integrate out x to give our upper bound on the variance. Let's do the first term now

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{v_0} \left(a_1 \exp\left(-b_1 \frac{(x-v_0)^2}{\varsigma^2}\right) \right)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \frac{a_1^2}{\sqrt{2\pi}\sigma} \int_{-\infty}^{v_0} \exp\left(-2b_1 \frac{(x-v_0)^2}{\varsigma^2}\right) \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \quad (68)$$

$$= \frac{1}{4\sqrt{2\pi}\sigma} \int_{-\infty}^{v_0} \exp\left(-\frac{(x-v_0)^2}{\varsigma^2} - \frac{(x-\mu)^2}{2\sigma^2}\right) dx \quad (69)$$

expand the brackets in the exponent giving

$$-\frac{(x-v_0)^2}{\varsigma^2} - \frac{(x-\mu)^2}{2\sigma^2} = -\frac{2\sigma^2(x-v_0)^2 + \varsigma^2(x-\mu)^2}{2\sigma^2\varsigma^2} \quad (70)$$

$$= -\frac{2\sigma^2(x^2 - 2xv_0 + v_0^2) + \varsigma^2(x^2 - 2\mu x + \mu^2)}{2\sigma^2\varsigma^2} \quad (71)$$

$$= -\frac{(2\sigma^2 + \varsigma^2)x^2 - (2\sigma^2 2v_0 + \varsigma^2 2\mu)x + 2\sigma^2 v_0^2 + \varsigma^2 \mu^2}{2\sigma^2\varsigma^2} \quad (72)$$

$$(73)$$

2 A Simple Neural Mass

We begin by defining the post-synaptic potential of population n as a result of an input firing rate from population m .

$$v_n(t) = v_{r,n} + \int_{-\infty}^t \frac{\alpha_{mn}}{\tau_{mn}} h_{mn}(t-t') g_m(v_m(t')) dt' \quad (74)$$

$$v_n(t) - v_{r,n} = \int_{-\infty}^t \frac{\alpha_{mn}}{\tau_{mn}} h_{mn}(t-t') g_m(v_m(t')) dt' \quad (75)$$

$$\tilde{v}_n(t) = \int_{-\infty}^t \frac{\alpha_{mn}}{\tau_{mn}} h_{mn}(t-t') g_m(v_m(t')) dt', \quad (76)$$

where α_{mn} is the gain for the post-synaptic response kernel, $h_{mn}(t)$ from neural population m to n and τ_{mn} is the time constant. Also, $g_m(v_m(t'))$ describes the input firing rate as a function of the pre-synaptic membrane potential. The resting membrane potential of the post-synaptic population is denoted by $v_{r,n}$, $v_n(t)$ is the post-synaptic membrane potential and $\tilde{v}_n(t)$ is the deviation of the membrane from the resting potential. The index n (post-synaptic) may represent either the pyramidal (p), excitatory interneuron (spiny stellate) (e), inhibitory interneuron (i), fast inhibitory ($i1$) or slow inhibitory ($i2$) populations.

The form of $h_{mn}(t)$ is

$$h_{mn}(t) = \eta(t) t \exp\left(-\frac{t}{\tau_{mn}}\right) \quad (77)$$

where $\eta(t)$ is the Heaviside step function.

This convolution can conveniently be written as

$$D\tilde{v}_n(t) = \frac{\alpha_{mn}}{\tau_{mn}} g_m(v_m(t)), \quad (78)$$

where the linear differential operator, D , is

$$D = \frac{d^2}{dt^2} + \frac{2}{\tau_{mn}} \frac{d}{dt} + \frac{1}{\tau_{mn}^2} \quad (79)$$

This allows the dynamics of the neural mass to be described by the differential equation

$$\frac{d^2 \tilde{v}_n(t)}{dt^2} + \frac{2}{\tau_{mn}} \frac{d\tilde{v}_n(t)}{dt} + \frac{1}{\tau_{mn}^2} \tilde{v}_n(t) = \frac{\alpha_{mn}}{\tau_{mn}} g_m(v_m(t')). \quad (80)$$

This second-order ODE can be written as two coupled first-order ODEs by defining

$$z_n(t) = \frac{d\tilde{v}_n(t)}{dt}. \quad (81)$$

This gives the system

$$\frac{d\tilde{v}_n(t)}{dt} = z_n(t) \quad (82)$$

$$\frac{dz_n(t)}{dt} = \frac{\alpha_{mn}}{\tau_{mn}} g_m(v_m(t')) - \frac{2}{\tau_{mn}} z_n(t) - \frac{1}{\tau_{mn}^2} \tilde{v}_n(t). \quad (83)$$

There is a sigmoidal relationship between the mean membrane potential and firing rate of each of the populations that is described by

$$g(\tilde{v}_n(t)) = \frac{1}{1 + \exp(\varsigma_n(v_{0n} - \tilde{v}_n(t)))} \quad (84)$$

$$g(\tilde{v}_n(t)) = \frac{1}{1 + \exp(\varsigma_n(v_{0n} + v_{rn} - v_n(t)))} \quad (85)$$

$$g(v_n(t)) = \frac{1}{1 + \exp(\varsigma_n(\tilde{v}_{0n} - v_n(t)))} \quad (86)$$

where $\tilde{v}_{0n} = v_{0n} + v_{rn}$. Note that in this formulation we are absorbing the maximal firing rate, which is typically a linear coefficient on the sigmoid, into the PSP gain (α_{mn}). This removes redundant parameters that can not be recovered by estimation methods. The parameters ς_n and v_{0n} describe the slope of the sigmoid (variance of firing thresholds within the populations) and the mean firing threshold, respectively. The parameter \tilde{v}_{0n} describes the deviation from the resting membrane potential, which becomes our lumped threshold parameter. For ease of notation we can drop the *tilde* remembering the resting membrane potential resides within this parameter / variable. Figure ?? depicts a standard neural mass.

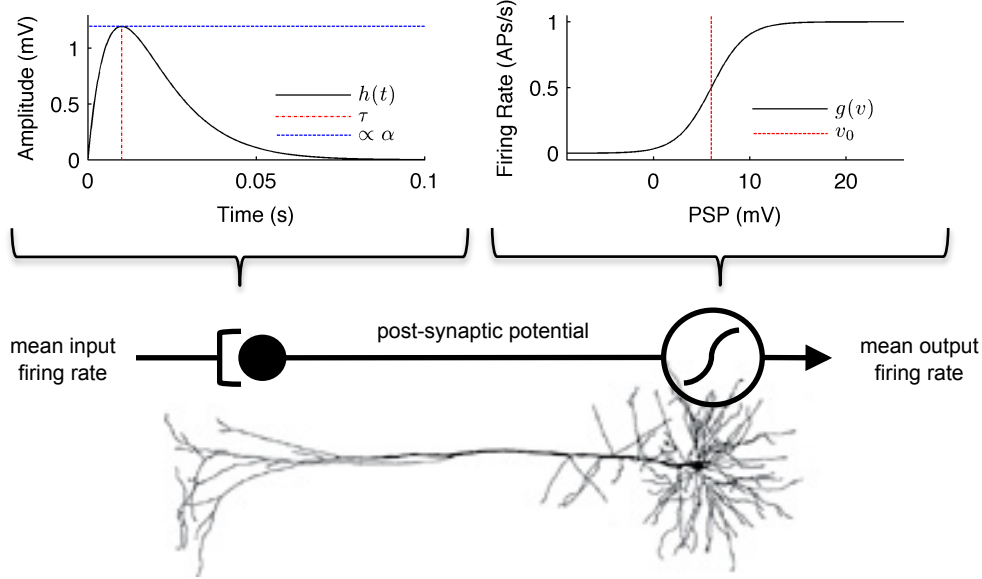


Figure 5: **Graphical Representation of the Neural Mass Model.** The neural mass model converts an input firing rate to a mean post-synaptic potential by a convolution with the post-synaptic response kernel. The membrane potential is converted to output firing rate by the sigmoidal activation function.

Since this model can be considered a statistical approximation to the actual brain dynamics we allow for model error and unmodelled inputs to be incorporated through the disturbance term $e_n(t)$ giving

$$dv_n(t) = z_n(t) dt \quad (87)$$

$$dz_n(t) = \alpha_{mn} \varsigma_{mn} g_m(v_m(t)) dt - 2\varsigma_{mn} z_n(t) dt - \varsigma_{mn}^2 v_n(t) dt + dw_n(t), \quad (88)$$

where ς_{mn} is the reciprocal of the synaptic time constant and $w_n(t)$ is a Wiener process that is independent of $v_n(t)$ and $w_n(t) - w_n(0) \sim \mathcal{N}(0, t)$. Note that we have switched notation using Ito's notation.

3 Discretisation

Next we need a discrete time version of the neural mass model in order to be able to apply the estimation methods. Let the time step be denoted by δt . The discretisation is performed using the Euler-Maruyama method by

$$v_n(t + \delta t) = z_n(t) \delta t + v_n(t) \quad (89)$$

$$z_n(t + \delta t) = (\alpha_{mn} \varsigma_{mn} g_m(v_m(t)) - 2\varsigma_{mn} z_n(t) - \varsigma_{mn}^2 v_n(t)) \delta t + z_n(t) + e_n(t), \quad (90)$$

where $e_n(t) = w_n(t + \delta t) - w_n(t) \sim \mathcal{N}(0, \delta t \sigma_n^2)$.

4 Jansen and Rit

An example of the model of a cortical column is shown in Figure ??.

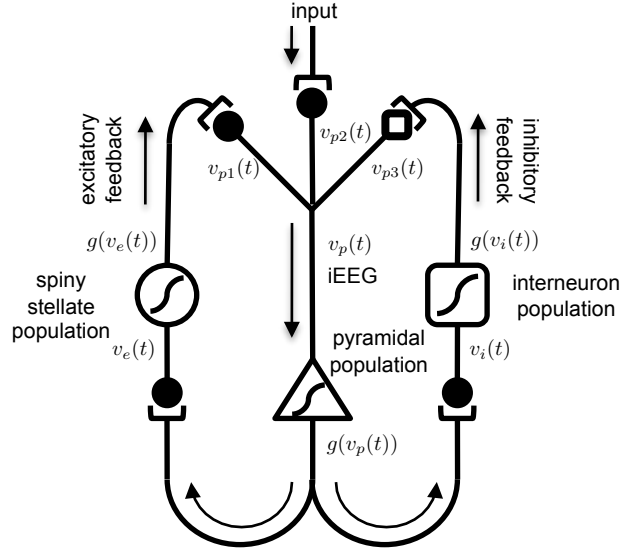


Figure 6: **Model of a Cortical Column.** The model shows three interconnected neural masses, which are pyramidal neurons, excitatory spiny stellate cells, and inhibitory interneurons. The specific subtype of neural population is defined by the parameters that describe the post-synaptic response kernels.

$$d v_e(t) = z_e(t) dt \quad (91)$$

$$d z_e(t) = (\alpha_{pe} \zeta_{pe} g(v_p(t)) - 2 \zeta_{pe} z_e(t) - \zeta_{pe}^2 v_e(t)) dt + d w_e(t) \quad (92)$$

$$d v_i(t) = z_i(t) dt \quad (93)$$

$$d z_i(t) = (\alpha_{pi} \zeta_{pi} g(v_p(t)) - 2 \zeta_e z_i(t) - \zeta_e^2 v_i(t)) dt + d w_i(t) \quad (94)$$

$$d v_{p1}(t) = z_{p1}(t) dt \quad (95)$$

$$d z_{p1}(t) = (\alpha_{ep} \zeta_{ep} g(v_e(t)) - 2 \zeta_e z_{p1}(t) - \zeta_e^2 v_{p1}(t)) dt + d w_{p1}(t) \quad (96)$$

$$d v_{p2}(t) = z_{p2}(t) dt \quad (97)$$

$$d z_{p2}(t) = (\alpha_{ip} \zeta_{ip} g(v_i(t)) - 2 \zeta_i z_{p2}(t) - \zeta_i^2 v_{p2}(t)) dt + d w_{p2}(t) \quad (98)$$

$$d v_{p3}(t) = z_{p3}(t) dt \quad (99)$$

$$d z_{p3}(t) = (\alpha_{xp} \zeta_{xp} \mu_x(t) - 2 \zeta_e z_{p3}(t) - \zeta_e^2 v_{p3}(t)) dt + d \alpha_{xp} w_x(t). \quad (100)$$

$$v_e^{t+1} = z_e^t \delta t + v_e^t \quad (101)$$

$$z_e^{t+1} = (\alpha_{pe} g(v_p^t) - 2\zeta_e z_e^t - \zeta_e^2 v_e^t) \delta t + z_e^t + \varepsilon_e^t \quad (102)$$

$$v_i^{t+1} = z_i^t \delta t + v_i^t \quad (103)$$

$$z_i^{t+1} = (\alpha_{pi} g(v_p^t) - 2\zeta_e z_i^t - \zeta_e^2 v_i^t) \delta t + z_i^t + \varepsilon_i^t \quad (104)$$

$$v_{p1}^{t+1} = z_{p1}^t \delta t + v_{p1}^t \quad (105)$$

$$z_{p1}^{t+1} = (\alpha_{ep} g(v_e^t) - 2\zeta_e z_{p1}^t - \zeta_e^2 v_{p1}^t) \delta t + z_{p1}^t + \varepsilon_{p1}^t \quad (106)$$

$$v_{p2}^{t+1} = z_{p2}^t \delta t + v_{p2}^t \quad (107)$$

$$z_{p2}^{t+1} = (\alpha_{ip} g(v_i^t) - 2\zeta_i z_{p2}^t - \zeta_i^2 v_{p2}^t) \delta t + z_{p2}^t + \varepsilon_{p2}^t \quad (108)$$

$$v_{p3}^{t+1} = z_{p3}^t \delta t + v_{p3}^t \quad (109)$$

$$z_{p3}^{t+1} = (\alpha_{xp} \mu_x(t) - 2\zeta_e z_{p3}^t - \zeta_e^2 v_{p3}^t) \delta t + z_{p3}^t + \varepsilon_x^t. \quad (110)$$

The measured output (iEEG) is given by

$$y^t = v_p^t + \epsilon^t \quad (111)$$

where

$$v_p^t = v_{p1}^t + v_{p2}^t - v_{p3}^t \quad (112)$$

and ϵ^t is the observation noise. The model can be written in a compact form as

$$\mathbf{x}^{t+1} = \mathbf{A}\mathbf{x}^t + \mathbf{f}(\mathbf{x}^t) + \mathbf{B}\mu_x^t + \boldsymbol{\varepsilon}^t \quad (113)$$

$$y^t = \mathbf{C}\mathbf{x}^t + \epsilon^t \quad (114)$$

where $A = \bigoplus_{j=1}^5 A_j$ and

$$A_j = \begin{cases} \begin{bmatrix} 1 & \delta t \\ -\zeta_i^2 \delta t & 1 - 2\zeta_i \delta t \end{bmatrix}, & j = 4 \\ \begin{bmatrix} 1 & \delta t \\ -\zeta_e^2 \delta t & 1 - 2\zeta_e \delta t \end{bmatrix}, & \text{otherwise} \end{cases} \quad (115)$$

and

$$\mathbf{B} = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1]^\top, \quad (116)$$

where \top denotes transpose operator. The non-linearity $\mathbf{f}(\cdot)$ is defined as

$$\mathbf{f}(\mathbf{x}^t) = \begin{bmatrix} 0 \\ \alpha_{pe} \delta t g(x_5^t + x_7^t - x_9^t) \\ 0 \\ \alpha_{pi} \delta t g(x_5^t + x_7^t - x_9^t) \\ 0 \\ \alpha_{ep} \delta t g(x_1^t) \\ 0 \\ \alpha_{ip} \delta t g(x_3^t) \\ 0 \\ 0 \end{bmatrix} \quad (117)$$

The observation matrix \mathbf{C} , is given by

$$\mathbf{C} = [0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0 \ -1 \ 0] \quad (118)$$

Parameter	Description	Value	Relation to JR
α_{pe}	Average excitatory synaptic gain	2193.75 mV	$2e_0AC_1$
α_{pi}	Average inhibitory synaptic gain	548.4375 mV	$2e_0AC_3$
α_{ep}	Average excitatory synaptic gain	1755.0 mV	$2e_0AC_2$
α_{ip}	Average inhibitory synaptic gain	3712.5 mV	$2e_0BC_4$
α_{xp}	Average synaptic input gain	16.25 mV	$2e_0A$
ζ_e	Average excitatory time constant	100 s^{-1}	a
ζ_i	Average inhibitory time constant	50 s^{-1}	b
v_0	Mean firing threshold	6 mV	v_0
ς	Slope of the sigmoid	$0.56 \text{ mV}^{-1}\text{spike s}^{-1}$	r
$\mu_x(t)$	Input	0	$p(t)$

Table 1: **Parameters of the model.**

Need to find

$$[\mathbf{Ax}_t + f(\mathbf{x}_t) + \varepsilon_t - \hat{\mathbf{x}}_{t+1}] [\mathbf{Ax}_t + f(\mathbf{x}_t) + \varepsilon_t - \hat{\mathbf{x}}_{t+1}]^\top \quad (119)$$

$$= (\mathbf{Ax}_t) (\mathbf{Ax}_t)^\top + (\mathbf{Ax}_t) f^\top(\mathbf{x}_t) + (\mathbf{Ax}_t) \varepsilon_t^\top - (\mathbf{Ax}_t) \hat{\mathbf{x}}_{t+1}^\top \quad (120)$$

$$+ f(\mathbf{x}_t) (\mathbf{Ax}_t)^\top + f(\mathbf{x}_t) f^\top(\mathbf{x}_t) + f(\mathbf{x}_t) \varepsilon_t^\top - f(\mathbf{x}_t) \hat{\mathbf{x}}_{t+1}^\top \quad (121)$$

$$+ \varepsilon_t (\mathbf{Ax}_t)^\top + \varepsilon_t f^\top(\mathbf{x}_t) + \varepsilon_t \varepsilon_t^\top - \varepsilon_t \hat{\mathbf{x}}_{t+1}^\top \quad (122)$$

$$- \hat{\mathbf{x}}_{t+1} (\mathbf{Ax}_t)^\top - \hat{\mathbf{x}}_{t+1} f^\top(\mathbf{x}_t) - \hat{\mathbf{x}}_{t+1} \varepsilon_t^\top + \hat{\mathbf{x}}_{t+1} \hat{\mathbf{x}}_{t+1}^\top \quad (123)$$