Merging two Gaussian distributions

We want to merge two Gaussians over x and v into one over v

$$p(x \mid v, z)p(v \mid g) = \mathcal{N}(x; zAv, \sigma_x I)\mathcal{N}(v; 0, C_v) \tag{1}$$

The Gaussian over x spelled out is

$$\mathcal{N}(x; zAv, \sigma_x I) = \sqrt{\frac{1}{(2\pi)^{Dx}\sigma_x^{Dx}}} e^{-\frac{1}{2\sigma_x}(x-zAv)^T(x-zAv)} \tag{2}$$

$$-\frac{1}{2}(x - zAv)^T(x - zAv) = -\frac{1}{2}(x^T x - zv^T A^T x - zx^T Av + z^2 v^T A^T Av) \tag{3}$$

as $v^T A^T x = (x^T Av)^T$, and both are scalars, thus equal to their transposes, it's also true that $v^T A^T x = x^T Av$

$$\begin{aligned} -\frac{1}{2}(x - zAv)^T(x - zAv) = -\frac{1}{2}(x^T x - 2zx^T Av + z^2 v^T A^T Av) = \\ = -\frac{x^T x}{2} + zx^T Av - \frac{z^2}{2} v^T A^T Av \end{aligned} \tag{4}$$

we have the identity for any symmetric matrix M and vector b that

$$-\frac{1}{2}v^T Mv + b^T v = -\frac{1}{2}(v - M^{-1}b)^T M(v - M^{-1}b) + \frac{1}{2}b^T M^{-1}b \tag{5}$$

making the substitution $M = z^2 A^T A$ and $b = (zx^T A)^T = zA^T x$, yielding $M^{-1} = \frac{1}{z^2}(A^T A)^{-1}$ and $M^{-1}b = \frac{1}{z}(A^T A)^{-1}A^T x = \frac{1}{z}A^+ x$, where A^+ is the Moore-Penrose pseudoinverse of A. Thus we get

$$\begin{aligned} -\frac{1}{2}(x - zAv)^T(x - zAv) = \\ = -\frac{x^T x}{2} - \frac{1}{2}(v - \frac{1}{z}A^+ x)^T z^2 A^T A(v - \frac{1}{z}A^+ x) + \frac{1}{2}(A^T x)^T (A^T A)^{-1} A^T x = \\ = -\frac{x^T x}{2} - \frac{1}{2}(v - \frac{1}{z}A^+ x)^T z^2 A^T A(v - \frac{1}{z}A^+ x) + \frac{1}{2}x^T AA^{-1}A^{-T}A^T x = \\ = -\frac{x^T x}{2} - \frac{1}{2}(v - \frac{1}{z}A^+ x)^T z^2 A^T A(v - \frac{1}{z}A^+ x) + \frac{x^T x}{2} = \\ = -\frac{1}{2}(v - \frac{1}{z}A^+ x)^T z^2 A^T A(v - \frac{1}{z}A^+ x) \end{aligned} \tag{6}$$

which implies

$$e^{-\frac{1}{2\sigma_x}(x-zAv)^T(x-zAv)} = e^{-\frac{1}{2}(v-\frac{1}{z}A^+x^T)^T\frac{z^2}{\sigma_x}A^TA(v-\frac{1}{z}A^+x)} \tag{7}$$

meaning that

$$\mathcal{N}(x; zAv, \sigma_x I) = \alpha\mathcal{N}(v; \frac{1}{z}A^+x, \frac{\sigma_x}{z^2}(A^TA)^{-1}) \tag{8}$$

and as the formulas in the exponents are equal, the constant α is given by the ratio of the normalisation terms

$$\sqrt{\frac{1}{(2\pi)^{Dx}\sigma_x^{Dx}}} = \alpha\sqrt{\frac{1}{(2\pi)^{Dv}\det(\frac{\sigma_x}{z^2}(A^TA)^{-1})}} \tag{9}$$

$$\alpha = \sqrt{\frac{(2\pi)^{Dv}\frac{\sigma_x^{D_v}}{z^{2D_v}}\det((A^TA)^{-1})}{(2\pi)^{Dx}\sigma_x^{Dx}}} \tag{10}$$

$$\alpha = \sqrt{\frac{(2\pi)^{Dv}\sigma_x^{D_v}}{(2\pi)^{Dx}\sigma_x^{Dx}z^{2D_v}\det(A^TA)}} \tag{11}$$

making the simplifying assumption $D_x = D_v$ we arrive to

$$\mathcal{N}(x; zAv, \sigma_x I) = \frac{1}{\sqrt{\det(A^TA)}}\frac{1}{z^{D_v}}\mathcal{N}(v; \frac{1}{z}A^+x, \frac{\sigma_x}{z^2}(A^TA)^{-1}) \tag{12}$$

we can merge two Gaussian distributions over v into one by using the following formula

$$\mathcal{N}(v; \mu_1, C_1)\mathcal{N}(v; \mu_2, C_2) = \mathcal{N}(\mu_1; \mu_2, C_1 + C_2)\mathcal{N}(v; \mu_c, C_c) \tag{13}$$

where $C_c = (C_1^{-1} + C_2^{-1})^{-1}$ and $\mu_c = C_c(C_1^{-1}\mu_1 + C_2^{-1}\mu_2)$. Substitution to these formulas yields

$$\begin{gathered}\frac{1}{\sqrt{\det(A^TA)}}\frac{1}{z^{D_v}}\mathcal{N}(v; \frac{1}{z}A^+x, \frac{\sigma_x}{z^2}(A^TA)^{-1})\mathcal{N}(v; 0, C_v) = \\ \frac{1}{\sqrt{\det(A^TA)}}\frac{1}{z^{D_v}}\mathcal{N}(\frac{1}{z}A^+x; 0, \frac{\sigma_x}{z^2}(A^TA)^{-1} + C_v)\mathcal{N}(v; \mu_c, C_c)\end{gathered} \tag{14}$$

$$C_c = (\frac{z^2}{\sigma_x}(A^TA) + C_v^{-1})^{-1} \tag{15}$$

$$\mu_c = C_c\frac{z}{\sigma_x}(A^TA)A^+x = \frac{z}{\sigma_x}C_cA^Tx \tag{16}$$