Electronic Supplementary Material (ESI) for Molecular BioSystems.

Appendix 1

Below is the derivation that the standard form of the marginal likelihood expressed as a normal distribution

$$p_{y|m}\left(y \mid m\right) = N\left(y, A(m)\cdot\bar{\theta}(m)+b(m), V_y + A(m)\cdot V_\theta(m)\cdot A(m)^T\right) \tag{10}$$

can be expressed as a product of the likelihood evaluated at the best-fit parameter set, the prior evaluated at the best-fit parameter set, and an expression containing the determinant of the posterior variance

$$p_{y|m}\left(y \mid m\right) = p_{y|\theta,m}\left(y \mid \tilde{\theta}(y,m), m\right)\cdot p_{\theta|m}\left(\tilde{\theta}(y,m) \mid m\right)\cdot\left\|2\cdot\pi\cdot V_{\tilde{\theta}}(y,m)\right\|^{\frac{1}{2}} \tag{11}$$

For brevity, the arguments for $A(m)$, $\bar{\theta}(m)$, $\tilde{\theta}(y,m)$ $V_{\bar{\theta}}(m)$, and $V_{\tilde{\theta}}(m)$ will be elided and $2\cdot\pi$ will be replaced by τ. Furthermore, $b(m)$ will be elided because y can always be redefined as the difference between the data and the intercept.

Expanding Equation 11 with the equivalent expressions for the likelihood (Equation 7) and prior (Equation 8) gives:

$$p_{y|m}\left(y \mid m\right) = \left\|\tau\cdot V_y\right\|^{-\frac{1}{2}}\cdot\exp\left(-\frac{1}{2}\cdot\left(y-A\cdot\tilde{\theta}\right)^T\cdot V_y^{-1}\cdot\left(y-A\cdot\tilde{\theta}\right)\right)\cdot\left\|\tau\cdot V_\theta\right\|^{-\frac{1}{2}}\cdot\exp\left(-\frac{1}{2}\cdot\left(\tilde{\theta}-\bar{\theta}\right)^T\cdot V_\theta^{-1}\cdot\left(\tilde{\theta}-\bar{\theta}\right)\right)\cdot\left\|\tau\cdot V_{\tilde{\theta}}\right\|^{\frac{1}{2}} \tag{A1}$$

Combining the exponential terms under one exponential gives:

$$\begin{aligned} p_{y|m}\left(y \mid m\right) = {} & \left\|\tau\cdot V_y\right\|^{-\frac{1}{2}}\cdot\left\|\tau\cdot V_\theta\right\|^{-\frac{1}{2}}\cdot\left\|\tau\cdot V_{\tilde{\theta}}\right\|^{\frac{1}{2}} \\ & \cdot\exp\left(-\frac{1}{2}\cdot\left(\left(y-A\cdot\tilde{\theta}\right)^T\cdot V_y^{-1}\cdot\left(y-A\cdot\tilde{\theta}\right)+\left(\tilde{\theta}-\bar{\theta}\right)^T\cdot V_\theta^{-1}\cdot\left(\tilde{\theta}-\bar{\theta}\right)\right)\right) \end{aligned} \tag{A2}$$

This equation will be considered in two parts: the product of determinants in the front:

$$\left\|\tau\cdot V_y\right\|^{-\frac{1}{2}}\cdot\left\|\tau\cdot V_\theta\right\|^{-\frac{1}{2}}\cdot\left\|\tau\cdot V_{\tilde{\theta}}\right\|^{\frac{1}{2}} \tag{A3}$$

and the sum in the exponential:

$$\left(y-A\cdot\tilde{\theta}\right)^T\cdot V_y^{-1}\cdot\left(y-A\cdot\tilde{\theta}\right)+\left(\tilde{\theta}-\overline{\theta}\right)^T\cdot V_\theta^{-1}\cdot\left(\tilde{\theta}-\overline{\theta}\right) \tag{A4}$$

The maximum a posteriori parameter set has the following definition for linear Gaussian models:

$$\tilde{\theta}\left(y,m\right)=V_{\tilde{\theta}}\cdot\left(A^T\cdot V_y^{-1}\cdot\left(y-b\right)+V_\theta^{-1}\cdot\overline{\theta}\right) \tag{A5}$$

Replacing $\tilde{\theta}$ in Equation A4 with its definition (Equation A5) gives:

$$\begin{aligned}&\left(y-A\cdot V_{\tilde{\theta}}\cdot\left(A^T\cdot V_y^{-1}\cdot\left(y-b\right)+V_\theta^{-1}\cdot\overline{\theta}\right)\right)^T\cdot V_y^{-1}\cdot\left(y-A\cdot V_{\tilde{\theta}}\cdot\left(A^T\cdot V_y^{-1}\cdot\left(y-b\right)+V_\theta^{-1}\cdot\overline{\theta}\right)\right)\\&+\left(V_{\tilde{\theta}}\cdot\left(A^T\cdot V_y^{-1}\cdot\left(y-b\right)+V_\theta^{-1}\cdot\overline{\theta}\right)-\overline{\theta}\right)^T\cdot V_\theta^{-1}\cdot\left(V_{\tilde{\theta}}\cdot\left(A^T\cdot V_y^{-1}\cdot\left(y-b\right)+V_\theta^{-1}\cdot\overline{\theta}\right)-\overline{\theta}\right)\end{aligned} \tag{A6}$$

Expanding those terms gives:

$$\begin{aligned}&y^T\cdot V_y^{-1}\cdot y\\&-2\cdot y^T\cdot V_y^{-1}\cdot A\cdot\left(A^T\cdot V_y^{-1}\cdot A+V_\theta^{-1}\right)^{-1}\cdot\left(A^T\cdot V_y^{-1}\cdot y+V_\theta^{-1}\cdot\overline{\theta}\right)\\&+\left(A^T\cdot V_y^{-1}\cdot y+V_\theta^{-1}\cdot\overline{\theta}\right)^T\cdot\left(A^T\cdot V_y^{-1}\cdot A+V_\theta^{-1}\right)^{-1}\cdot A^T\cdot V_y^{-1}\cdot A\cdot\left(A^T\cdot V_y^{-1}\cdot A+V_\theta^{-1}\right)^{-1}\cdot\left(A^T\cdot V_y^{-1}\cdot y+V_\theta^{-1}\cdot\overline{\theta}\right)\\&+\overline{\theta}^T\cdot V_\theta^{-1}\cdot\overline{\theta}\\&-2\cdot\overline{\theta}^T\cdot V_\theta^{-1}\cdot\left(A^T\cdot V_y^{-1}\cdot A+V_\theta^{-1}\right)^{-1}\cdot\left(A^T\cdot V_y^{-1}\cdot y+V_\theta^{-1}\cdot\overline{\theta}\right)\\&+\left(A^T\cdot V_y^{-1}\cdot y+V_\theta^{-1}\cdot\overline{\theta}\right)^T\cdot\left(A^T\cdot V_y^{-1}\cdot A+V_\theta^{-1}\right)^{-1}\cdot V_\theta^{-1}\cdot\left(A^T\cdot V_y^{-1}\cdot A+V_\theta^{-1}\right)^{-1}\cdot\left(A^T\cdot V_y^{-1}\cdot y+V_\theta^{-1}\cdot\overline{\theta}\right)\end{aligned} \tag{A7}$$

Expanding those terms further gives:

$$
\begin{aligned}
& y^T \cdot V_y^{-1} \cdot y \\
& -2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot y \\
& -2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta} \\
& + y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot y \\
& +2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta} \\
& + \bar{\theta}^T \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta} \\
& + \bar{\theta}^T \cdot V_\theta^{-1} \cdot \bar{\theta} \\
& -2 \cdot \bar{\theta}^T \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot y \\
& -2 \cdot \bar{\theta}^T \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta} \\
& + y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot y \\
& +2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta} \\
& + \bar{\theta}^T \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta}
\end{aligned}
\tag{A8}
$$

We will group each term according to whether it is quadratic, linear, or constant with respect to y.

Quadratic terms

Collecting the four terms from Equation A8 that are quadratic with respect to y and factoring y from the front and back gives:

$$
y \cdot \begin{pmatrix}
V_y^{-1} \\
-2 \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \\
+ V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \\
+ V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1}
\end{pmatrix} \cdot y
\tag{A9}
$$

Factoring the $-A^T \cdot V_y^{-1}$ term from the front and $A^T \cdot V_y^{-1}$ from the back of the second, third, and fourth terms of the sum gives:

$$y \cdot \begin{pmatrix} V_y^{-1} \\ -V_y^{-1} \cdot A \cdot \begin{pmatrix} 2 \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot A^T \cdot V_y^{-1} \end{pmatrix} \cdot y \qquad \text{(A10)}$$

Factoring $\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1}$ from the front and back of the third and fourth terms to collapse the third and fourth terms together gives:

$$y \cdot \begin{pmatrix} V_y^{-1} \\ -V_y^{-1} \cdot A \cdot \begin{pmatrix} 2 \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right) \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot A^T \cdot V_y^{-1} \end{pmatrix} \cdot y \qquad \text{(A11)}$$

Cancelling $\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)$ in the third term gives:

$$y \cdot \begin{pmatrix} V_y^{-1} \\ -V_y^{-1} \cdot A \cdot \begin{pmatrix} 2 \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot A^T \cdot V_y^{-1} \end{pmatrix} \cdot y \qquad \text{(A12)}$$

Collapsing the second and third terms which now have the same form gives:

$$y \cdot \left(V_y^{-1} - V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot A^T \cdot V_y^{-1}\right) \cdot y \qquad \text{(A13)}$$

Finally, applying the Woodbury identity gives:

$$y \cdot \left(V_y + A \cdot V_\theta \cdot A^T\right)^{-1} \cdot y \qquad \text{(A14)}$$

Linear terms

Collecting from Equation A8 the terms that are linear with respect to y and factoring $-2 \cdot y^T \cdot V_y^{-1} \cdot A$ from the front and $V_\theta^{-1} \cdot \bar{\theta}$ from the back gives:

$$-2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \begin{pmatrix} \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ +\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \cdot \bar{\theta} \tag{A15}$$

Collapsing the first and third terms which have the same form gives:

$$-2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \begin{pmatrix} 2\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \cdot \bar{\theta} \tag{A16}$$

Factoring $\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1}$ from the second and third terms gives:

$$-2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \begin{pmatrix} 2\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right) \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \cdot \bar{\theta} \tag{A17}$$

Cancelling $\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)$ in the second term gives:

$$-2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \begin{pmatrix} 2\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \cdot \bar{\theta} \tag{A18}$$

Collapsing the first and second terms which now have the same form gives:

$$-2 \cdot y^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot V_\theta^{-1} \cdot \bar{\theta} \tag{A19}$$

Transforming $V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1}$ with the positive definite variant of the Woodbury identity gives:

$$-2 \cdot y^T \cdot \left(V_y + A \cdot V_\theta \cdot A^T \right)^{-1} \cdot A \cdot V_\theta \cdot V_\theta^{-1} \cdot \overline{\theta} \tag{A20}$$

Cancelling $V_\theta \cdot V_\theta^{-1}$ gives:

$$-2 \cdot y^T \cdot \left(V_y + A \cdot V_\theta \cdot A^T \right)^{-1} \cdot A \cdot \overline{\theta} \tag{A21}$$

Constant terms

Collecting from Equation A8 the terms that are not functions of $\hat{y}$ and factoring $\overline{\theta}$ from the front and back gives:

$$\overline{\theta}^T \cdot \begin{pmatrix} V_\theta^{-1} \\ -2 \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \\ +V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \\ +V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \end{pmatrix} \cdot \overline{\theta} \tag{A22}$$

Factoring $-V_\theta^{-1}$ from the front and V_θ^{-1} from the back of the second, third, and fourth terms gives:

$$\overline{\theta}^T \cdot \begin{pmatrix} V_\theta^{-1} \\ -V_\theta^{-1} \cdot \begin{pmatrix} 2 \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot A^T \cdot V_y^{-1} \cdot A \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \cdot V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1} \right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \end{pmatrix} \cdot \overline{\theta} \tag{A23}$$

Factoring $\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1}$ from the front and back of the third and fourth terms gives:

$$\bar{\theta}^T \cdot \begin{pmatrix} V_\theta^{-1} \\ -V_\theta^{-1} \cdot \begin{pmatrix} 2 \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right) \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \end{pmatrix} \cdot \bar{\theta} \qquad \text{(A24)}$$

Canceling $\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)$ in the third term gives:

$$\bar{\theta}^T \cdot \begin{pmatrix} V_\theta^{-1} \\ -V_\theta^{-1} \cdot \begin{pmatrix} 2 \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \\ -\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \end{pmatrix} \cdot V_\theta^{-1} \end{pmatrix} \cdot \bar{\theta} \qquad \text{(A25)}$$

Collapsing the second and third terms which now have the same form gives:

$$\bar{\theta}^T \cdot \left(V_\theta^{-1} - V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot V_\theta^{-1}\right) \cdot \bar{\theta} \qquad \text{(A26)}$$

Applying the Woodbury identity to $V_\theta^{-1} - V_\theta^{-1} \cdot \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \cdot V_\theta^{-1}$ gives:

$$\bar{\theta}^T \cdot A^T \cdot \left(V_y + A V_\theta A^T\right)^{-1} \cdot A \cdot \bar{\theta} \qquad \text{(A27)}$$

Determinant terms

Combining all determinant terms from Equation A3 under a single exponent gives:

$$\left(\left\|\tau \cdot V_y\right\| \cdot \left\|\tau \cdot V_\theta\right\| \cdot \left\|\tau \cdot V_{\tilde{\theta}}\right\|^{-1}\right)^{-\frac{1}{2}} \qquad \text{(A28)}$$

Pulling out τ from each determinant gives:

$$\left(\tau^{n_y} \cdot \left\|V_y\right\| \cdot \tau^{n_\theta} \cdot \left\|V_\theta\right\| \cdot \tau^{-n_\theta} \cdot \left\|V_{\tilde{\theta}}\right\|^{-1}\right)^{-\frac{1}{2}} \qquad \text{(A29)}$$

Cancelling $\tau^{n_\theta} \cdot \tau^{-n_\theta}$ gives:

$$\left(\tau^{n_y} \cdot \left\|V_y\right\| \cdot \left\|V_\theta\right\| \cdot \left\|V_{\tilde{\theta}}\right\|^{-1}\right)^{-\frac{1}{2}} \quad \text{(A30)}$$

The posterior variance for a linear Gaussian model given by:

$$V_{\tilde{\theta}} = \left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1} \quad \text{(A31)}$$

Replacing $V_{\tilde{\theta}}$ in Equation A30 with its linear definition (Equation A31) gives:

$$\left(\tau^{n_y} \cdot \left\|V_y\right\| \cdot \left\|V_\theta\right\| \cdot \left\|\left(A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right)^{-1}\right\|^{-1}\right)^{-\frac{1}{2}} \quad \text{(A32)}$$

Cancelling the stacked negative exponents gives:

$$\left(\tau^{n_y} \cdot \left\|V_y\right\| \cdot \left\|V_\theta\right\| \cdot \left\|A^T \cdot V_y^{-1} \cdot A + V_\theta^{-1}\right\|\right)^{-\frac{1}{2}} \quad \text{(A33)}$$

Applying a determinant identity gives:

$$\left(\tau^{n_y} \cdot \left\|V_y + A \cdot V_\theta \cdot A^T\right\|\right)^{-\frac{1}{2}} \quad \text{(A34)}$$

Finally, bringing the τ back into the determinant gives:

$$\left\|\tau\left(V_y + A \cdot V_\theta \cdot A^T\right)\right\|^{-\frac{1}{2}} \quad \text{(A35)}$$

Recombination

Recombining the simplified quadratic (Equation A14), linear (Equation A21), and constant (Equation A27) expressions, which together equal Equation A8, gives:

$$y \cdot \left(V_y + A \cdot V_\theta \cdot A^T\right)^{-1} \cdot y - 2 \cdot y^T \cdot \left(V_y + A \cdot V_\theta \cdot A^T\right)^{-1} \cdot A \cdot \bar{\theta} + \bar{\theta}^T \cdot A^T \cdot \left(V_y + A \cdot V_\theta A^T\right)^{-1} \cdot A \cdot \bar{\theta} \quad \text{(A36)}$$

Factoring the quadratic equation gives:

$$\left(y - A\cdot\overline{\theta}\right)^{T}\cdot\left(V_y + A\cdot V_\theta\cdot A^{T}\right)^{-1}\cdot\left(y - A\cdot\overline{\theta}\right) \tag{A37}$$

Replacing the simplified determinant (Equation A35) and the simplified quadratic equation (Equation A37) in Equation A2 gives:

$$p_{y|m}\left(y\mid m\right) = \left\|\tau\cdot\left(V_y + A\cdot V_\theta\cdot A^{T}\right)\right\|^{-\frac{1}{2}}\cdot\exp\left(-\frac{1}{2}\cdot\left(y - A\cdot\overline{\theta}\right)^{T}\cdot\left(V_y + A\cdot V_\theta\cdot A^{T}\right)^{-1}\cdot\left(y - A\cdot\overline{\theta}\right)\right) \tag{A38}$$

Given that the probability density function of a normal distribution is:

$$N\left(x,\mu,V\right) = \left\|\tau\cdot V\right\|^{-\frac{1}{2}}\cdot\exp\left(-\frac{1}{2}\cdot\left(x-\mu\right)^{T}\cdot V^{-1}\cdot\left(x-\mu\right)\right) \tag{A39}$$

it is clear that the linearization formula (Equation 11) is equivalent to a normal distribution of variable y with a mean of $A\cdot\overline{\theta}$ and a variance of $V_y + A\cdot V_\theta\cdot A^{T}$, which is the standard marginal likelihood given in Equation 10.