S.1 Proof of (14)

Recall that the model parameters are \( \theta = (\theta_l, \theta_h) \). We express \( p(y_h(x_0)|y_l, y_h) \) as

\[
p(y_h(x_0)|M_k, y_l, y_h) = \int_{\theta_l} \int_{\theta_h} p(y_h(x_0)|M_k, y_h, \theta_l, \theta_h) p(\theta_l, \theta_h|M_k, y_l, y_h) d\theta_h d\theta_l
\]

\[
= \int_{\theta_l} \int_{\theta_h} p(y_h(x_0)|M_k, y_h, \eta_l, \theta_h) p(\theta_h|M_k, \eta_l, y_h) d\theta_h p(\eta_l|M_k, y_l, \theta_l) p(\theta_l|y_l) d\theta_l
\]

\[
= \int_{\theta_l} p(y_h(x_0)|M_k, y_h, \eta_l) p(\eta_l|M_k, y_l, \theta_l) p(\theta_l|y_l) d\theta_l
\]

where

\[
p(y_h(x_0)|M_k, \eta_l, y_h) = \int_{\theta_h} p(y_h(x_0)|M_k, y_h, \eta_l, \theta_h) p(\theta_h|M_k, \eta_l, y_h) d\theta_h
\]

\[
= \int_{\alpha, \sigma_e^2} p(y_h(x_0)|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) p(\alpha, \sigma_e^2, \lambda|M_k, \eta_l, y_h) \ d\alpha \ d\sigma_e^2
\]

\[
= \int_{\lambda} p(y_h(x_0)|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) p(\alpha, \sigma_e^2|M_k, \lambda, \eta_l, y_h) \ d\alpha \ d\sigma_e^2
\]

In order to get the expression of \( p(y_h(x_0)|M_k, \eta_l, y_h) \), perform the integration in (24) in the following two steps:

(i) Integrate out \( \alpha \) and \( \sigma_e^2 \);

(ii) Integrate out \( \lambda \).

Step (i) integrate out \( \alpha \) and \( \sigma_e^2 \). We denote the inner integration in (24) by \( p(y_h(x_0)|M_k, y_h, \eta_l, \lambda) \), that is,

\[
p(y_h(x_0)|M_k, y_h, \eta_l, \lambda) = \int_{\sigma_e^2, \alpha} p(y_h(x_0)|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) p(\alpha, \sigma_e^2|M_k, \lambda, y_h, \eta_l) \ d\alpha \ d\sigma_e^2
\]

\[
\propto \int_{\sigma_e^2, \alpha} p(y_h(x_0)|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) p(y_h|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) \ d\alpha \ d\sigma_e^2
\]

Given the kernel width \( \lambda \), the linkage model can be considered as a linear regression model \( y_h = F_\lambda \alpha + \epsilon_h \). Recall that \( \epsilon_h \sim N(0, \sigma_e^2 I) \). Therefore,

\[
(y_h|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) \sim N(F_\lambda \alpha, \sigma_e^2 I)
\]

\[
(y_h(x_0)|M_k, \eta_l, \alpha, \sigma_e^2, \lambda) \sim N(F_\lambda(x_0)\alpha, \sigma_e^2)
\]

These are the same results as in (10) and (11). Given that the prior distribution of \( \alpha \) and \( \sigma_e^2 \) is \( p(\alpha, \sigma_e^2) \propto \sigma_e^{-2} \), Gelman et al. (2003, Page 359) stated that under this priors, the posterior predictive distribution of \( y_h(x_0) \), conditioned on the data and kernel width \( \lambda \), is

\[
(y_h(x_0)|M_k, y_h, \eta_l(x_0), \lambda) \sim t_{m_e - 2}(F_\lambda(x_0)\hat{\alpha}, s^2(1 + F_\lambda(x_0)(F_\lambda^T F_\lambda)^{-1} F_\lambda(x_0)^T))
\]

where \( \hat{\alpha} = (F_\lambda^T F_\lambda)^{-1} F_\lambda^T y_h \) and \( s^2 = \frac{1}{m_e - 2}(y_h - F_\lambda \hat{\alpha})^T(y_h - F_\lambda \hat{\alpha}) \). This is how (14) is obtained. Consequently, after \( \alpha \) and \( \sigma_e^2 \) are integrated out, (24) becomes

\[
p(y_h(x_0)|M_k, \eta_l, y_h) = \int_{\lambda} p(y_h(x_0)|M_k, y_h, \eta_l(x_0), \lambda) p(\lambda|M_k, \eta_l, y_h) d\lambda
\]
Step (ii), integrate out $\lambda$. Recall that $\lambda$ has a discrete distribution. Thus, the integration in (25) can be written as a summation (and (15) is obtained):

$$p(y_h(x_0) | M_k, \eta_l, y_h) = \sum_{\lambda_1=1,2,...,\lambda_0} \sum_{\lambda_d=1,2,...,\lambda_0} p(y_h(x_0) | M_k, y_h, \eta_l, \lambda) p(\lambda | M_k, y_h, \eta_l),$$

where

$$p(\lambda | M_k, y_h, \eta_l) \propto p(\lambda) p(y_h | M_k, \eta_l, \lambda). \tag{26}$$

The marginal distribution of the high-resolution data given the inputs $\eta_l$ and the kernel width $\lambda$ is as follows

$$p(y_h | M_k, \eta_l, \lambda) = \int_{\sigma_e^2, \alpha} p(y_h, \alpha, \sigma_e^2 | M_k, \eta_l, \lambda) \, d\alpha \, d\sigma_e^2 \tag{27}$$

$$= \int_{\sigma_e^2, \alpha} p(y_h | M_k, \eta_l, \alpha, \sigma_e^2, \lambda) \, d\alpha \, d\sigma_e^2$$

$$= \int_{\sigma_e^2} \int_{\alpha} (2\pi)^{-m_h/2} (\sigma_e^2)^{-m_h/2} \exp \left\{ -\frac{1}{2\sigma_e^2} (y_h - F_\lambda^T y_h - F_\lambda \alpha)^2 \right\} \, d\alpha \, d\sigma_e^2$$

$$= \int_{\sigma_e^2} \int_{\alpha} \exp \left\{ -\frac{1}{2\sigma_e^2} \left[ (y_h - F_\lambda^T y_h - F_\lambda \alpha) + (\alpha - \hat{\alpha})^T F_\lambda^T F_\lambda (\alpha - \hat{\alpha}) \right] \right\} \, d\alpha \, d\sigma_e^2$$

$$= \int_{\sigma_e^2} \int_{\alpha} \exp \left\{ -\frac{1}{2\sigma_e^2} (\alpha - \hat{\alpha})^T F_\lambda^T F_\lambda (\alpha - \hat{\alpha}) \right\} \, d\alpha \, d\sigma_e^2$$

$$= \int_{2\pi} \int_{\sigma_e^2, \alpha} \exp \left\{ -\frac{1}{2\sigma_e^2} (y_h - F_\lambda^T y_h - F_\lambda \alpha)^2 \right\} \, d\alpha \, d\sigma_e^2$$

$$= (2\pi)^{-m_h+1} |F_\lambda^T F_\lambda|^{-1/2} \Gamma \left( \frac{m_h}{2} - 1 \right) \left[ \frac{(y_h - F_\lambda \hat{\alpha})^T (y_h - F_\lambda \hat{\alpha})}{2\sigma_e^2} \right]^{-m_h+1}.$$

Note that in the second line of the above derivation, we utilize that (11) which specifies the distribution of $p(y_h | \eta_l, \alpha, \sigma_e^2, \lambda)$. Given the above, (26) can now be written as

$$p(\lambda | y_h, \eta_l) \propto p(\lambda) |F_\lambda^T F_\lambda|^{-1/2} \left[ \frac{(y_h - F_\lambda \hat{\alpha})^T (y_h - F_\lambda \hat{\alpha})}{2} \right]^{-m_h+1}.$$

This shows how (14) is obtained.
S.2 Proof of (20)

To prove (20), we just integrate (27) over $\lambda$. As $\lambda$ takes discrete values, we end up with a summation over possible values for $\lambda$ as in (20):

$$p(y_h|M_k, \eta_l) = \sum_{\lambda_1=1,2,\ldots,\lambda_0} \sum_{\lambda_2=1,2,\ldots,\lambda_0} \cdots \sum_{\lambda_d=1,2,\ldots,\lambda_0} (2\pi)^{-\frac{m_h}{2}+1}|F_{\lambda}^T F_{\lambda}|^{-\frac{1}{2}} \Gamma(\frac{m_h}{2} - 1) \left[ (y_h - F_{\lambda} \hat{\alpha})^T (y_h - F_{\lambda} \hat{\alpha}) \right].$$