We are interested in $p(\theta|Y)$. Normal Bayes rule gives

$$p(\theta|Y) = p(Y|\theta)p(\theta)/p(Y).$$

The trouble is that computing $p(Y|\theta)$ is really hard.

Let z represent some outcome from the model such that $p(Y|z,\theta) = p(Y|z)$.

Ozge is currently filtering using some rule like $\rho(y,z) \leq \delta$. This can be considered R^2, but really anything.

We need to compute

$$p(Y|\theta) = \int_{\rho(Y,z)\leq\delta} p(Y|z,\theta)p(z|\theta,)\mathrm{d}z + \int_{\rho(Y,z)>\delta} p(Y|z,\theta)p(z|\theta)\mathrm{d}z$$

If $p(Y|z,\theta) \approx 0$ when $\rho(Y,z) > \delta$, meaning our rule is sufficiently loose that it only eliminates bad stuff, then

$$p(Y|\theta) \approx \int_{\rho(Y,z)\leq\delta} p(Y|z,\theta)p(z|\theta)\mathrm{d}z$$

Moreover, $p(Y|z,\theta) = p(Y|z)$.

We find it much easier to understand $p(z|\theta, \rho(y,z) \leq \delta)$ than $p(z|\theta)$. When we build our emulator, we only have data from $z|\theta, \rho(y,z) \leq \delta$.

This gives us

$$p(Y|\theta) \approx \frac{p(\rho(Y,Z) \leq \delta|\theta)}{p(\rho(Y,Z) \leq \delta|\theta)} \int p(Y|z)p(z|\theta)\mathrm{d}z$$

$$p(Y|\theta) \approx p(\rho(Y,Z) \leq \delta|\theta) \int p(Y|z) \frac{p(z|\theta)}{p(\rho(Y,Z) \leq \delta|\theta)}\mathrm{d}z$$

or

$$p(Y|\theta) \approx p(\rho(Y,Z) \leq \delta|\theta) \int p(Y|z)p(z|\theta, \rho(y,z) \leq \delta)\mathrm{d}z$$

Note that we had been using

$$p(Y|\theta) \approx \int_{\rho(Y,z)\leq\delta} p(Y|z)p(z|\theta, \rho(y,z) \leq \delta)\mathrm{d}z$$

This is what is in the code now.

The front term matters! A good way to approximate it is

$$p(\rho(Y,Z) \leq \delta|\theta) = \frac{1}{n}\sum_{i=1}^{n} I(\rho(Y,Z) \leq \delta|\theta)$$

Basically, carry the empirical probability of eliminating solutions.