

Yes! And it has a beautiful closed form. Let me work through it.

By Bayes:

$$P(X = \pm 1 | Y_T = y) = \frac{P(X = \pm 1) \cdot \mathcal{N}(y; \pm 1, T\sigma^2)}{f_{Y_T}(y)}$$

So:

$$\mathbb{E}[X | Y_T = y] = P(X=1|y) - P(X=-1|y) = \frac{(1-a)e^{-(y-1)^2/(2T\sigma^2)} - ae^{-(y+1)^2/(2T\sigma^2)}}{ae^{-(y+1)^2/(2T\sigma^2)} + (1-a)e^{-(y-1)^2/(2T\sigma^2)}}$$

This looks messy, but the difference of exponents simplifies nicely:

$$\frac{(y+1)^2 - (y-1)^2}{2T\sigma^2} = \frac{4y}{2T\sigma^2} = \frac{2y}{T\sigma^2}$$

Factoring $e^{-(y-1)^2/(2T\sigma^2)}$ out of numerator and denominator and letting $r = a/(1-a)$ (the prior odds):

$$\mathbb{E}[X | Y_T = y] = \frac{1 - r e^{-2y/(T\sigma^2)}}{1 + r e^{-2y/(T\sigma^2)}}$$

This is exactly the form of \tanh with a shifted argument. Using $\tanh(z) = (1 - e^{-2z})/(1 + e^{-2z})$:

$$\boxed{\mathbb{E}[X | Y_T = y] = \tanh\left(\frac{y}{T\sigma^2} - \frac{1}{2} \log \frac{a}{1-a}\right)}$$

A few things to notice about this formula:

The denominator $T\sigma^2$ is the cumulative variance. The argument is essentially signal-to-noise: as T grows, the SNR shrinks, the \tanh argument flattens, and the function gets less responsive to y .

For uniform prior $a = 1/2$: the log term vanishes and you get the clean result $\tanh(y/T\sigma^2)$ — exactly the formula from our earlier two-point widget, but with the noise variance σ^2 replaced by the cumulative variance $T\sigma^2$.

Prior bias. When $a \neq 1/2$, the \tanh gets horizontally shifted by $\frac{T\sigma^2}{2} \log \frac{a}{1-a}$. So the “decision point” where $\mathbb{E}[X|y] = 0$ moves off zero — biased toward the cluster with smaller prior weight (because you need more y -evidence to override a strong prior).

Asymptotic behavior as $T \rightarrow \infty$: the $y/(T\sigma^2)$ term vanishes and you're left with $\tanh(-\frac{1}{2} \log(a/(1-a)))$. Using the identity $\tanh(\frac{1}{2} \log(p/q)) = \frac{p-q}{p+q}$:

$$\mathbb{E}[X | Y_\infty = y] \rightarrow 1 - 2a = \mathbb{E}[X]$$

This is exactly the prior mean — the unconditional expectation. When you've added so much noise that Y tells you nothing about X , the best guess collapses to the prior mean. Which is exactly right.

Want to plot this tanh function as a widget alongside the forward distribution?