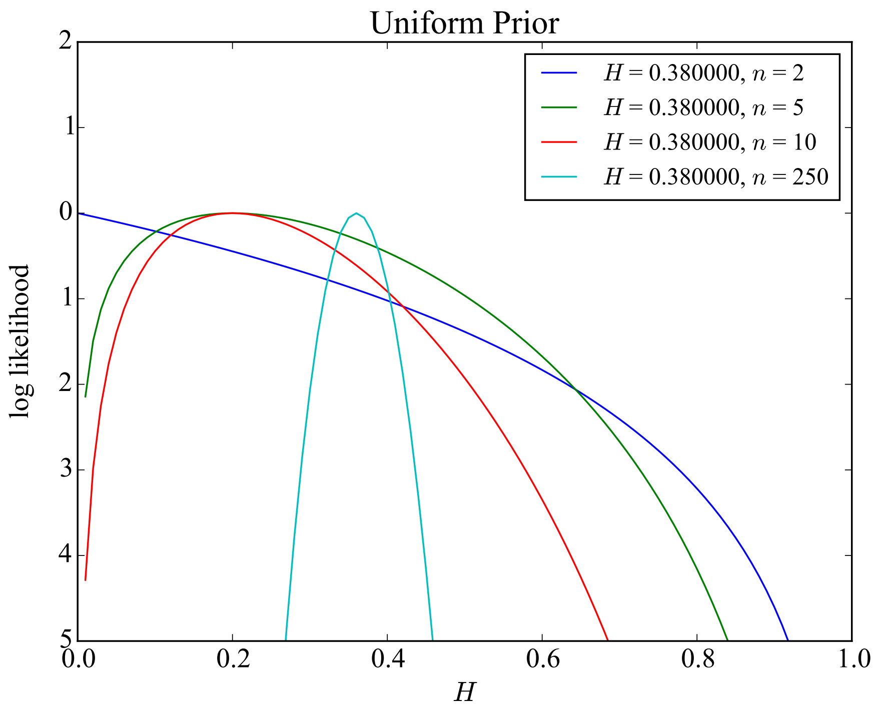

Figure 1: Likelihoods for multiple n with an uniform prior.

Theorem (Bayes' Theorem)**.** The probability that a system is well-described a model with parameter set $\mathbf{X}$, given a dataset $\mathbf{D}$ and initial knowledge I, is:

$$\Pr(\mathbf{X}|\mathbf{D}, I) = \frac{\Pr(\mathbf{D}|X, I) \cdot \Pr(\mathbf{X}|I)}{\Pr(\mathbf{D}|I)}.$$

$\Pr(\mathbf{X}|I)$ is called the *prior*, $\Pr(\mathbf{X}|\mathbf{D}, I)$ the *posterior*, and $\Pr(\mathbf{D}|\mathbf{X}, I)$ the *likelihood*.

1 Coin tossing

We consider a sequence of tosses of a biased coin. The coin can be modeled with a single parameter $H \in [0, 1]$, the probability of heads on a single toss. Thus our parameter set is $\mathbf{X} = \{H\}$.

The dataset $\mathbf{D}$ can be reduced to a tuple (n, h) where h is the number of heads in n tosses. There is no other prior information: $I = \emptyset$. The probability of obtaining some data given the parameter H is:

$$\Pr(\mathbf{D}|\mathbf{X}, I) = \Pr\left((n, h)|H\right) = \binom{n}{h} H^h (1 - H)^{n-h},$$

since there are $\binom{n}{h}$ ways of getting h heads (each with probability H) and $n - h$ tails (each with probability $1 - H$).

In this simulation, we choose $H = 0.38$ and three prior distributions: an uniform one, a Gaussian with $\sigma = 0.1$ centered around $\mu = 0.5$, representing a belief that the coin is fair, and a Gaussian with $\sigma = 0.05$ centered around $\mu = 0.5$, representing a stronger belief that the coin is fair. Results for each prior are shown in Figures 1-3.

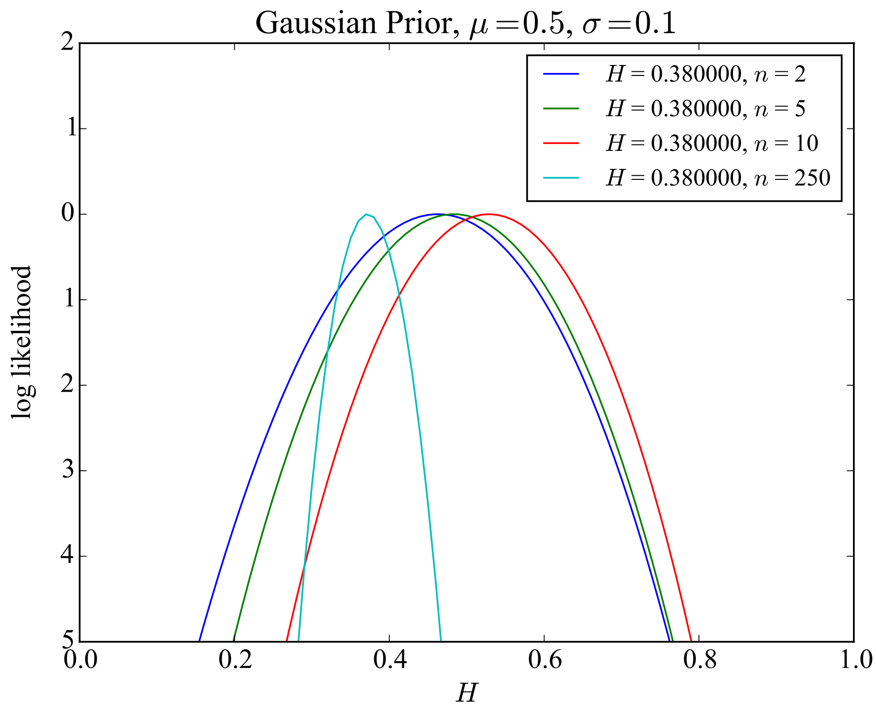

Figure 2: Likelihoods for multiple n with an wide Gaussian prior.

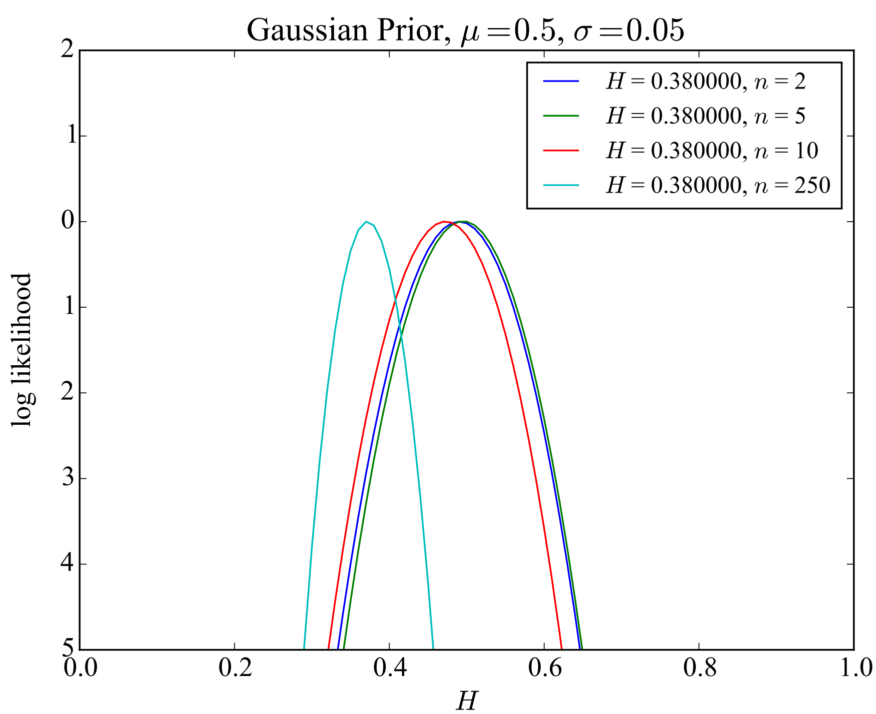

Figure 3: Likelihoods for multiple n with a narrow Gaussian prior.

2 Lighthouse problem

We have a lighthouse at a distance β from the shore and at a location α along the shore. The lighthouse emits flashes at random angles θ_k which arrive at the shore at locations x_k. Let:

$$\begin{aligned}
f_\theta(\theta) &\equiv \Pr(\theta_k = \theta) \text{ be the pdf of } \theta_k, \\
F_\theta(\theta) &\equiv \Pr(\theta_k \leq \theta) \text{ be the cdf of } \theta_k, \\
f_x(x) &\equiv \Pr(x_k = x) \text{ be the pdf of } x_k, \\
F_x(x) &\equiv \Pr(x_k \leq x) \text{ be the cdf of } x_k.
\end{aligned}$$

Since the lighthouse emits flashes at random angles, we assume θ_k is uniformly distributed:

$$f_\theta(\theta) = \frac{1}{2\pi}, \qquad F_\theta(\theta) = \int_0^\theta f_\theta(\tilde{\theta})\, d\tilde{\theta} = \frac{\theta}{2\pi}.$$

To derive f_x, we first derive F_x:

$$F_x(x) = \Pr(x_k \leq x) = \Pr\Big(\theta_k \leq \hat{\theta}(x)\Big) = F_\theta\Big(\hat{\theta}(x)\Big),$$

where $\hat{\theta}(x)$ is the angle that creates a flash at x. To find it, we use the geometry of the problem:

$$\hat{\theta}(x) = \arctan\left(\frac{x-\alpha}{\beta}\right).$$

Thus:

$$\begin{aligned}
F_x(x) &= \frac{1}{2\pi} \arctan\left(\frac{x-\alpha}{\beta}\right). \\
f_x(x) &= \frac{d}{dx} F_x(x) = \frac{\beta/2\pi}{(x-\alpha)^2 + \beta^2}.
\end{aligned}$$

This makes intuitive sense: the distribution is a Lorentzian centered around α, the location of the the lighthouse along the shore. The further the lighthouse is, the wider the distribution is.

For our study we choose $\alpha = 1.0$ and $\beta = 1.5$. We again use three prior distributions: an uniform one, a wide Gaussian centered at $\mu = 1.5$ with $\sigma = 0.3$, and a narrow Gaussian with $\mu = 1.5$ with $\sigma = 0.1$.

Results for the one-dimensional case, where β is known but α is not, are shown in Figures 4-6.

In the two-dimensional case, we use identical priors for each parameter. Results are shown in Figures 7-9.

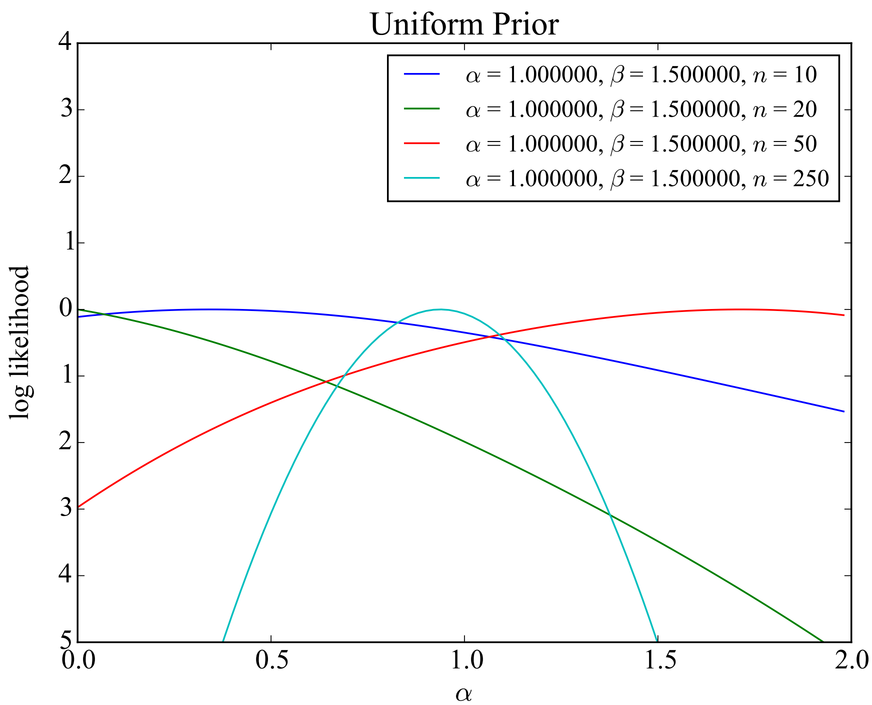

Figure 4: Likelihoods for multiple n with an uniform prior.

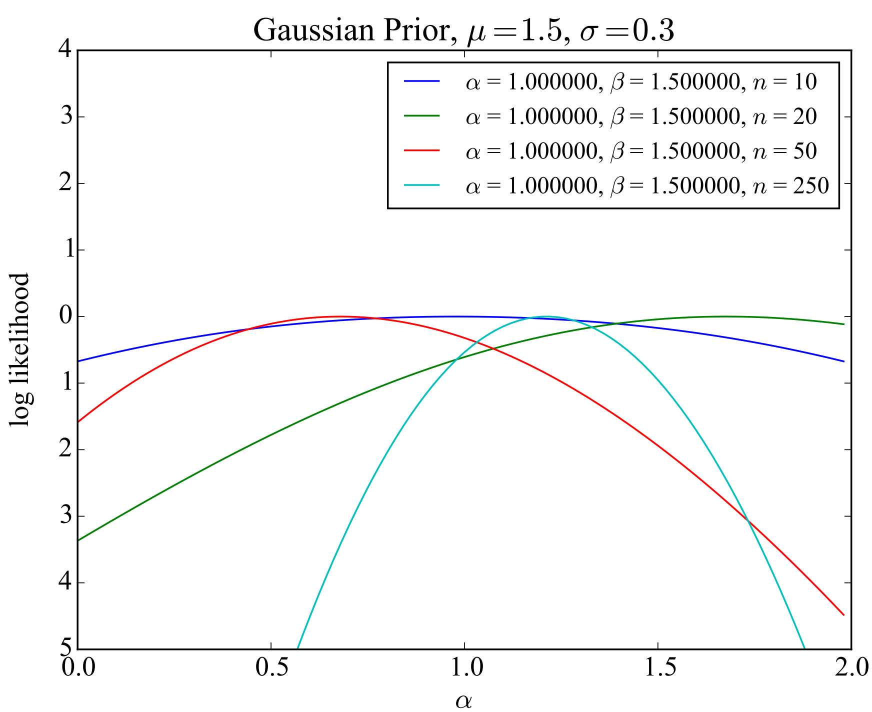

Figure 5: Likelihoods for multiple n with a wide Gaussian prior.

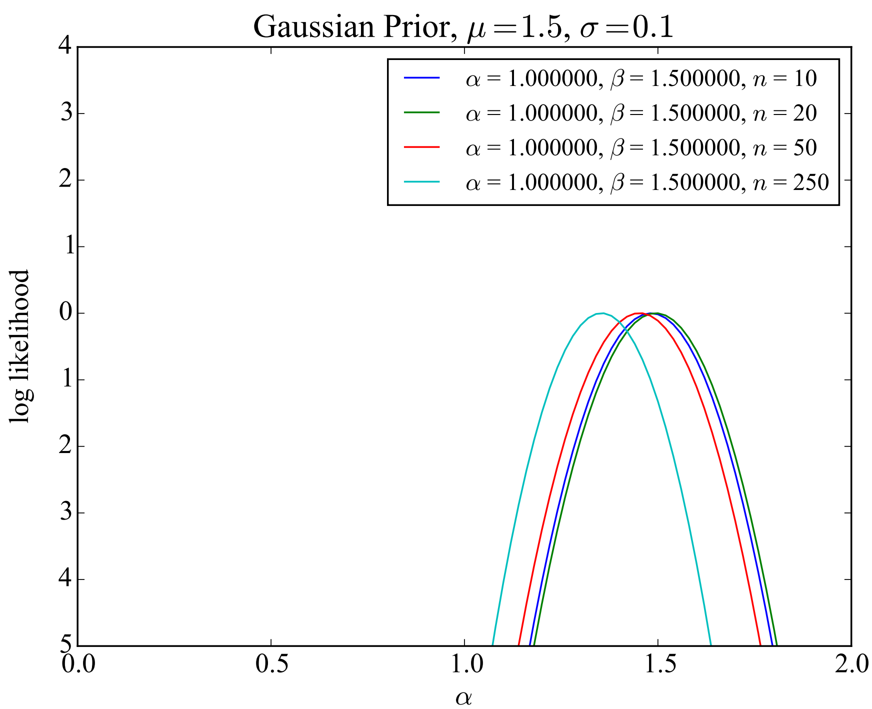

Figure 6: Likelihoods for multiple n with a narrow Gaussian prior.

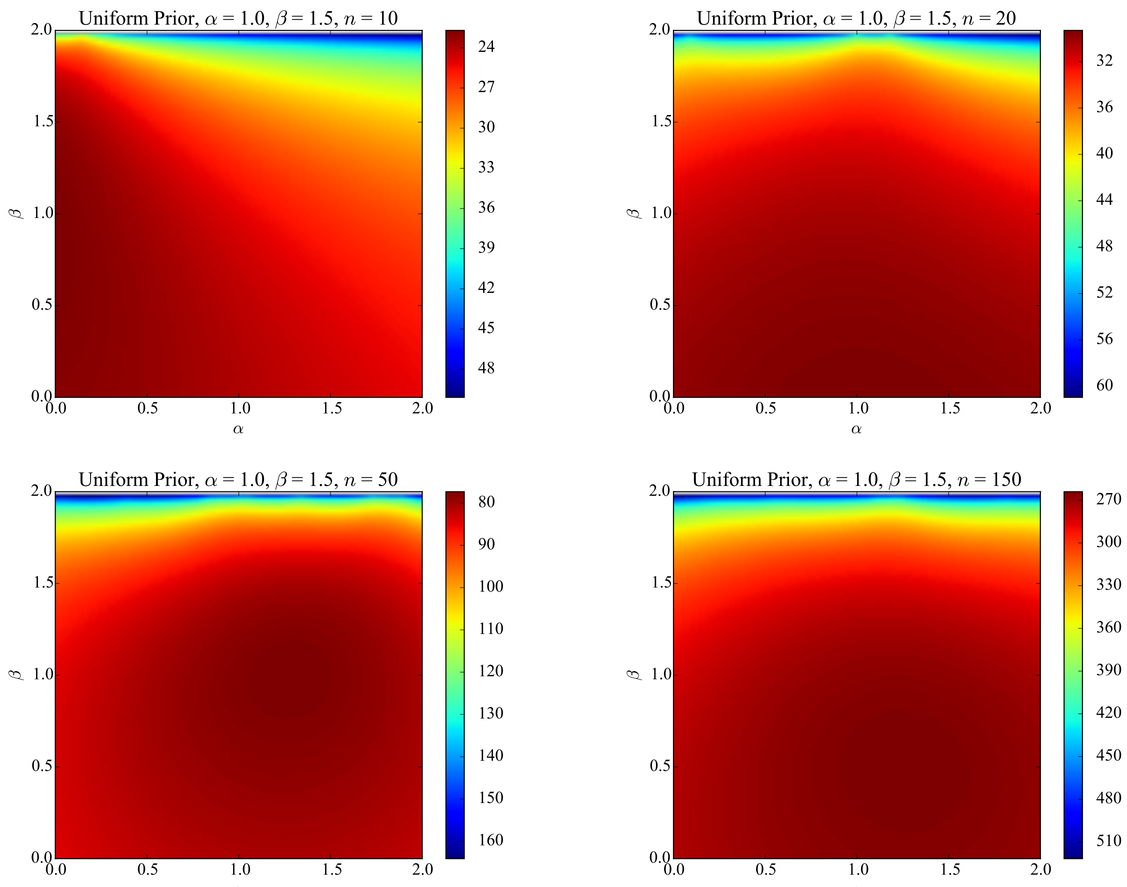

Figure 7: Likelihoods for multiple n with an uniform prior.

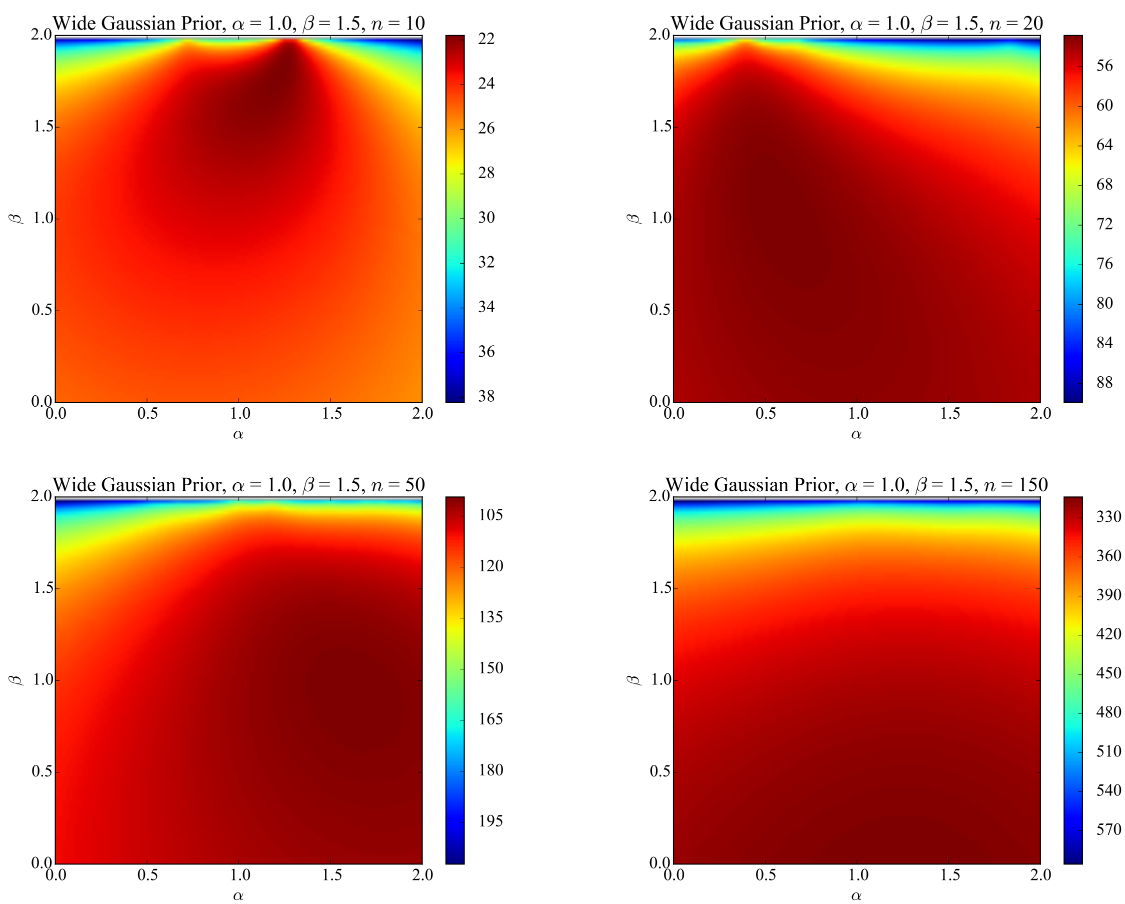

Figure 8: Likelihoods for multiple n with a wide Gaussian prior.

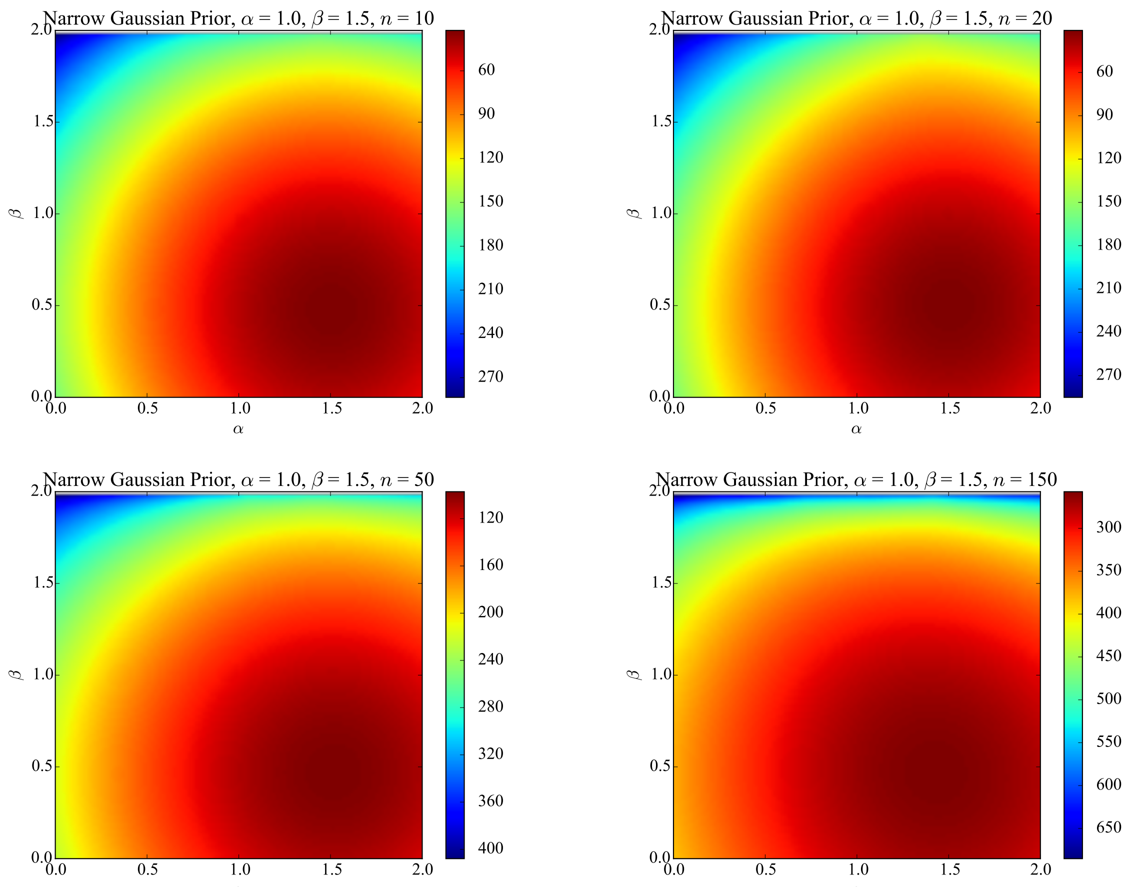

Figure 9: Likelihoods for multiple n with a narrow Gaussian prior.