Poisson-Gamma Neural Variability in the Visual Cortex

Yilun Kuang

Dec 20, 2020

1 Introduction

In neural coding, a population of sensory neurons responds to the same repeated stimuli with differing spiking variability. The stochasticity in the neural responses is characterized as "trial-to-trial variability". A classical model in neural encoding captures the response variability using the poisson rate model:

$$N \sim \text{Poisson}(\lambda), \quad \lambda \geq 0,$$

where $\lambda = \mu \Delta t$ is the rate parameter, μ denotes the mean spike rate, Δt denotes the duration of the counting window, and N represents the spike count (Goris et al., 2014) [4] . So we have the spike count distribution:

$$P(N = n|\lambda) = \frac{\lambda^n}{n!}\exp(-\lambda)$$

with equal mean and variance, i. e. $E[N] = \text{Var}[N] = \lambda$. In sensory system and particularly in visual cortex, there is usually higher variance than the mean in the population responses, a phenomenon called as "overdispersion" (Taouali et al., 2016) [6]. Overdispersion in neural data represents modulatory influences (attention, adapation etc) coming from non-sensory sources, and the response variability increases along the visual pathways from LGN, V1, V2, to MT (Goris et al., 2014) [4].

2 Poisson-Gamma Model

To capture the overdispersion, this project paper implements a doubly-stochastic poisson-gamma model by Goris et al (2014) [4]. We define the modulatory influence, or gain signal, as G. Given the poisson-gamma framework, the mean spike rate μ is the product of a stimuli-independent gain signal G and a function of the stimuli $f(S)$:

$$\mu = f(S)G \implies \lambda = f(S)G\Delta t.$$

The gain signal follows a gamma distribution:

$$G \sim \text{gamma}(r, s),$$

where the gamma distribution is parameterized by the shape parameter r and the scale parameter s. Here we set $E[G] = rs = 1$, $\text{Var}[G] = rs^2 = s$, where $s = \sigma_G^2$, $r = \frac{1}{\sigma_G^2}$ (later in the numerical simulation we will explore the decoding accuracy of both $[r, s] = [\sigma_G^2, \sigma_G^2]$ and $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$). Here σ_G^2 denotes the variances of the gain signal G. Now since $\lambda = f(S)G\Delta t$ where $f(S)\Delta t$ is a constant, λ is now a random variable Λ with a gamma distribution:

$$\Lambda \sim \text{gamma}(r, sf(S)\Delta t)$$

$$P(\Lambda = \lambda) = \frac{\lambda^{r-1}\exp(-\lambda/[sf(S)\Delta t])}{\Gamma(r)[sf(S)\Delta t]^r}.$$

So for $P(N|\Lambda)$ and $P(\Lambda)$ as a poisson-gamma mixture, the marginal distribution of N should be negative-binomial (Gelman et al., 2013, p. 44) [3].

$$\text{Neg-Bin}(n|r,s) = \int \text{Poisson}(n|\lambda)\text{Gamma}(\lambda|r,s)d\lambda.$$

Now we can compute the marginal distribution of N:

$$\begin{aligned}P(N=n) &= \int_0^\infty P(N=n|\Lambda \in \lambda)P(\Lambda \in \lambda)d\lambda \\ &= \int_0^\infty \frac{\lambda^n}{n!}\exp(-\lambda)\frac{\lambda^{\ r-1}\exp\left(-\lambda/[s(f(S)\Delta t)]\right)}{\Gamma(r)[s(f(S)\Delta t)]^r}d\lambda.\end{aligned}$$

Let $\alpha = s(f(S)\Delta t)$, we have.

$$\begin{aligned}&\int_0^\infty \frac{\lambda^n}{n!}\exp(-\lambda)\frac{\lambda^{\ r-1}\exp\left(-\lambda/\alpha\right)}{\Gamma(r)\alpha^r}d\lambda \\ &= \int_0^\infty \frac{\alpha^{-r}}{n!\ \Gamma(r)}\ \lambda^{n+r-1}e^{-(1+\frac{1}{\alpha})\lambda}\ d\lambda \\ &= \frac{\alpha^{-r}}{n!\ \Gamma(r)}\ \Gamma(n+r)\left(\frac{\alpha}{\alpha+1}\right)^{n+r}\int_0^\infty \frac{1}{\Gamma(n+r)\left(\frac{\alpha}{\alpha+1}\right)^{n+r}}\ \lambda^{n+r-1}e^{-(1+\frac{1}{\alpha})\lambda}\ d\lambda \\ &= \frac{\alpha^{-r}}{n!\ \Gamma(r)}\ \Gamma(n+r)\left(\frac{\alpha}{\alpha+1}\right)^{n+r} = \frac{\Gamma(n+r)}{\Gamma(N+1)\ \Gamma(r)}\ \alpha^{-r}\left(\frac{\alpha}{\alpha+1}\right)^{n+r} \\ &= \frac{\Gamma(n+r)}{\Gamma(N+1)\ \Gamma(r)}\ \alpha^{-r}\left(\frac{\alpha}{\alpha+1}\right)^{n+r} = \frac{\Gamma(n+r)}{\Gamma(N+1)\ \Gamma(r)}\ \left(\frac{1}{\alpha+1}\right)^{r}\left(\frac{\alpha}{\alpha+1}\right)^{n}.\end{aligned}$$

So we have the spike count N following a negative-binomial, or poisson-gamma mixture, distribution:

$$N \sim \text{Neg-Bin}(r,\alpha),\ \text{where}$$

$$P(N=n) = \frac{\Gamma(n+r)}{\Gamma(N+1)\ \Gamma(r)}\ \left(\frac{1}{\alpha+1}\right)^{r}\left(\frac{\alpha}{\alpha+1}\right)^{n},$$

where

$$\begin{aligned}E[N] &= r\alpha = rs(f(S)\Delta t) = f(S)\Delta t \\ \text{Var}[N] &= r\alpha + r\alpha^2 = f(S)\Delta t + (\sigma_G^2)(f(S)\Delta t)^2.\end{aligned}$$

To quantify the extent of overdispersion, we use Fano Factor F, defined as the variance over the mean. For poisson rate model, $F_{\text{Poi}} = 1$. For the negative binomial model, we have

$$F_{\text{Neg-Bin}} = \frac{\text{Var}[N]}{E[N]} = 1 + \alpha = 1 + \sigma_G^2(f(S)\Delta t).$$

(Notice that $F_{\text{Neg-Bin}}$ is not changed for either $[r,s] = [\sigma_G^2, \sigma_G^2]$ or $[r,s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$). So the overdispersion in the negative-binomial neural population relative to the poisson population is

$$F' = F_{\text{Neg-Bin}} - F_{\text{Poi}} = \sigma_G^2(f(S)\Delta t),$$

where F' is determined by the product of the gamma parameter σ_G^2 and the stimulus tuning.

3 Encoding

A numerical simulation is performed to determine the fluctuations in the Fano Factor with varying stimulus values and gamma parameters. For simplicity, we assume Δt to be 1. A gaussian tuning curve with $g = 15, b = 0.1, \sigma = 5$ is chosen with the preferred orientations of 20 to 40 and the true stimuli values ranged from 0 to 60 for a population of 50 neurons. The gamma gain signal G is parametrized by the scale parameter s and the shape parameter r from 1 to 81 in 50 steps. The gamma values are generated by 100 trials for every different scale parameter s and put through a poisson random number generator to create a poisson-gamma distribution for every different gamma parameters and stimulus values. Then fano factors are plotted for varying stimulus values and gamma parameters.

3.1 $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$

If we set $r = s^{-1}$, we will have the following figure

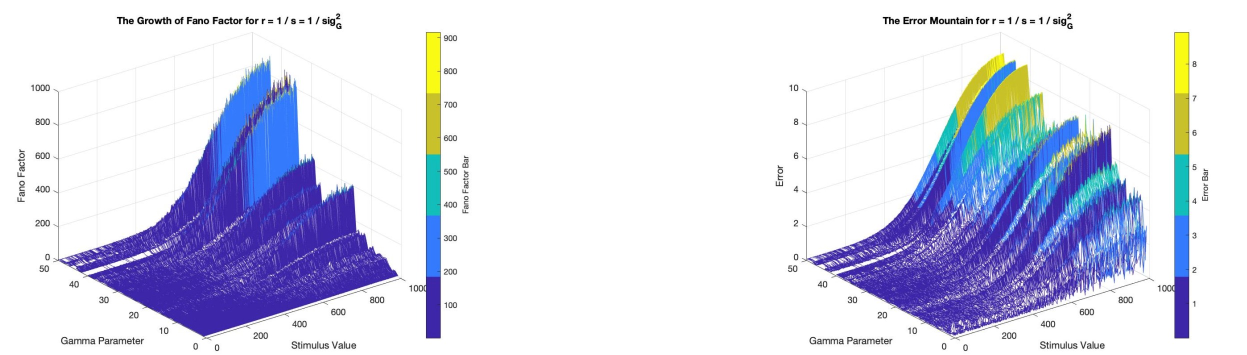

Figure 1: Encoding Performance for $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$.

It can be seen that fano factor generally increase for increasing gamma parameters and stimulus values, but there are some fluctuations. This fluctuability could be due to matlab underflow since there might be a parameter combination such as $[r, s] = [\frac{1}{81}, 81]$ that is not numerically stable. To determine the extent to which the growth in fano factor is not coming from the product of σ_G^2 and $f(S)\Delta t$, we define the error term

$$\mathcal{E}_F = |f(S)\Delta t - \hat{f(S)}\Delta t| \text{ ,where } \hat{f(S)}\Delta t = \frac{F_{\text{Neg-Bin}} - 1}{\sigma_G^2},$$

$$\text{for } F_{\text{Neg-Bin}} = 1 + \sigma_G^2(f(S)\Delta t).$$

Plot $\mathcal{E}_F$ and we get the Error Mountain above. Since $\mathcal{E}_F$ fluctuates on a small scale in general, it can be concluded that the excess in fano factors comes from the product between σ_G^2 and $f(S)\Delta t$, as our derivation predicted.

3.2 $[r, s] = [\sigma_G^2, \sigma_G^2]$

Set $r = s = \sigma_G^2$ and plot the fano graph. From the figure, it is obvious that the fano factors increase as gamma parameters and stimulus values increase. The error is also small and it corresponds to our initial hypothesis.

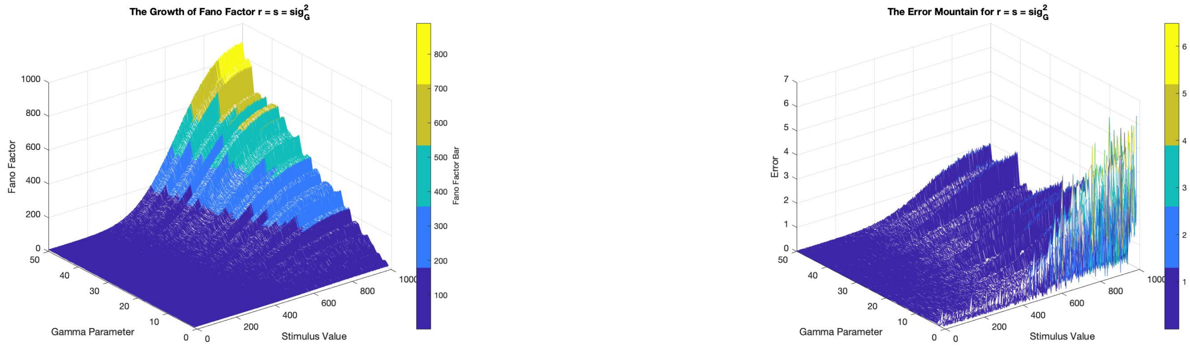

Figure 2: Encoding Performance for $[r, s] = [\sigma_G^2, \sigma_G^2]$.

4 Decoding

During the encoding process, the parameters of the gain signal G is parametrized as $r = \sigma_G^2, s = \sigma_G^2$. For decoding, both $[r, s] = [\sigma_G^2, \sigma_G^2]$ and $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$ is explored. Here we will do $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$ first (different from the encoding part above, where $[r, s] = [\sigma_G^2, \sigma_G^2]$).

4.1 $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$

Rewrite r as s^{-1}, α as $s(f(S)\Delta t)$ we get:

$$P(N = n) = \frac{\Gamma(n + s^{-1})}{\Gamma(N + 1)\ \Gamma(s^{-1})} \left(1 + \alpha\right)^{-s^{-1}} \left(\frac{\alpha}{\alpha + 1}\right)^{n}$$

$$= \frac{\Gamma(n + s^{-1})}{\Gamma(N + 1)\ \Gamma(s^{-1})} \left(1 + s(f(S)\Delta t)\right)^{-s^{-1}} \left(\frac{s(f(S)\Delta t)}{s(f(S)\Delta t) + 1}\right)^{n},$$

where

$$N \sim \text{Neg-Bin}(s, f(S)\Delta t).$$

The log likelihood is then

$$l(s, f(S)\Delta t) = \log\left(\prod_{i=1}^{N} \frac{\Gamma(n_i + s^{-1})}{n_i!\ \Gamma(s^{-1})} \left(1 + s(f(S)\Delta t)\right)^{-s^{-1}} \left(\frac{s(f(S)\Delta t)}{s(f(S)\Delta t) + 1}\right)^{n_i}\right)$$

Simplify then we get the log likelihood equation (Piegorsch, 1990) [5]:

$$l(s, f(S)\Delta t) \propto \frac{1}{m} \sum_{i=1}^{m} \sum_{v=0}^{n_i - 1} \log\{1 + sv\} + \bar{n} \log\{f(S)\Delta t\} - (\bar{n} + s^{-1}) \log\{1 + sf(S)\Delta t\}.$$

Take the partial derivative of l with respect to $f(S)\Delta t$ we get

$$\nabla_{f(S)\Delta t}\, l = \frac{\bar{n}}{f(S)\Delta t} - \frac{(\bar{n} + s^{-1}) \times s}{1 + sf(S)\Delta t}$$

$$= \frac{\bar{n}}{f(S)\Delta t} - \frac{1 + s\bar{n}}{1 + sf(S)\Delta t}$$

Now set $\nabla_{f(S)\Delta t}\, l = 0$, then we have $f(\hat{S})\Delta t = \bar{n}$ (Piegorsch, 1990) [5]. The ML estimator for the parameter $f(S)\Delta t$ is just the sample mean $\bar{n}$, i. e. the mean spike count (Anscombe, 1950) [1]. This result from statistical literature implies that the stimulus value $f(S)$ can be decoded independently regardless of the values of the gain signal. It is consistent with the claim that "the precision of decoding of the PM and the NBM under the nontuned dispersion hypothesis are very similar" (Taouali et al., 2016, p. 441) [6].

Population decoding is performed in nbm_decoding.m. To prevent unexpected fluctuations in the population responses, the Matlab command nbinrnd is used to encode a negative binomial population instead of nesting gamrnd inside poissrnd. Maximum likelihood estimation is performed on the encoding population. There are trials where the sample variance exceeds the sample mean such that the parameter estimation is not possible for the maximum likelihood estimation of the negative binomial parameters. Given that this project wants to evaluate the decoding accuracy of a negative binomial population, it is legitimate to throw away trials where certain parameter combinations can't generate an adequate amount of sample variance over the sample mean for the population response to be "negative binomial". The missing data is then imputed using the KNN imputation package for the purpose of plotting the decoding bias colormap without missing values. After the data imputation, the decoding bias is plotted (decoding variance is not included in here as all variances are close to 0).

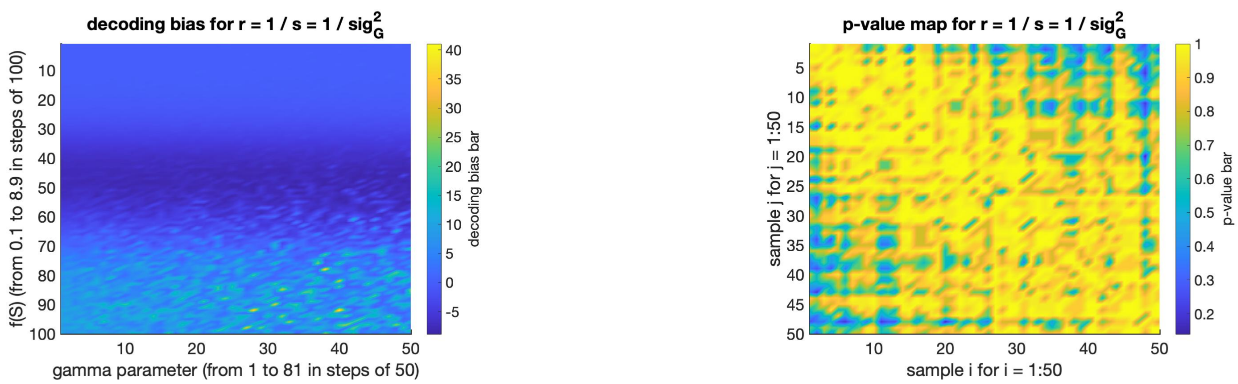

Figure 3: Decoding Metrics for $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$.

A visual inspection tells us that there might be some but not significant increases in the decoding bias as we increase the gamma parameters. A two-sample Kolmogorov-Smirnov test is performed over 50 samples of decoding bias, each corresponding to a given gamma parameter, for every two columns. There are no p-value smaller than 0.05 and we fail to reject the null hypothesis that 50 samples of decoding bias are coming from the same distribution. In other words, the decoding bias is not changed with respect to the variations in the gamma parameters, confirming the theoretical derivation that the decoding accuracy is not affected by the gain signal G.

4.2 $[r, s] = [\sigma_G^2, \sigma_G^2]$

Rewrite r as β^{-1}, α as $s(f(S)\Delta t)$ we get:

$$
\begin{aligned}
P(N = n) &= \frac{\Gamma(n + \beta^{-1})}{\Gamma(N+1)\,\Gamma(\beta^{-1})} \Big(1 + \alpha\Big)^{-\beta^{-1}} \left(\frac{\alpha}{\alpha + 1}\right)^{n} \\
&= \frac{\Gamma(n + \beta^{-1})}{\Gamma(N+1)\,\Gamma(\beta^{-1})} \Big(1 + s(f(S)\Delta t)\Big)^{-\beta^{-1}} \left(\frac{s(f(S)\Delta t)}{s(f(S)\Delta t) + 1}\right)^{n},
\end{aligned}
$$

where

$$N \sim \text{Neg-Bin}(\beta, sf(S)\Delta t).$$

Then by Piegorsch (1990) we have $(sf(\hat{S})\Delta t) = \bar{n}$ for $\nabla_{f(S)\Delta t}\ l = 0$ (Piegorsch, 1990) [5]. The decoding accuracy is affected by the scale parameter s.

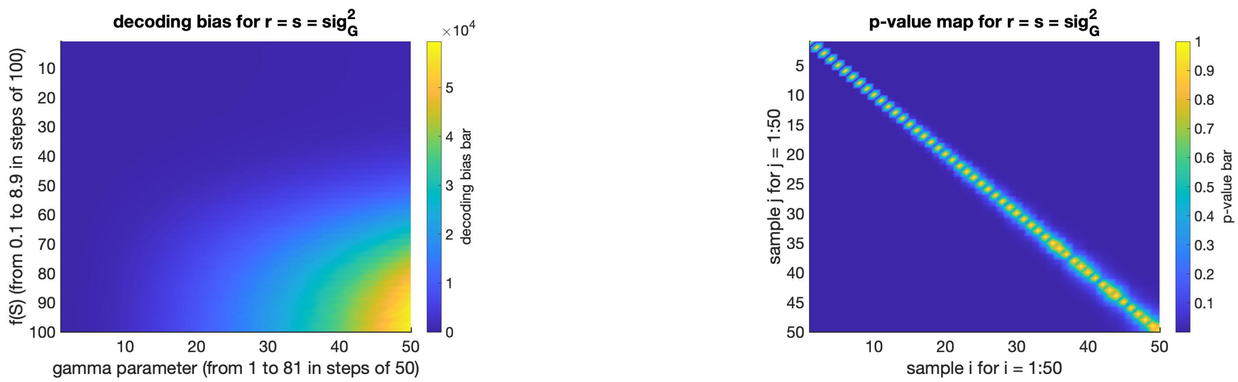

Figure 4: Decoding Metrics for $[r, s] = [\sigma_G^2, \sigma_G^2]$.

With the same procedure and the Kolmogorov-Smirnov test described above, we can conclude that we reject the null hypothesis that 50 samples of decoding bias comes from the same distribution, i. e. the gain signal does affect the decoding accuracy for the $f(S)\Delta t$ given the assumption that $[r, s] = [\sigma_G^2, \sigma_G^2]$, as opposed to $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$ where the decoding accuracy is not affected by the gamma parameters.

5 Conclusion

This project has shown that a poisson-gamma encoding model from Goris et al (2014) creates excess variability in the population responses, which is reflected in the growth of Fano Factor with varying gamma values and stimulus values [4]. A further decoding of population response through both mathematical derivations and numerical simulations has shown that the decoding accuracy of population responses is not affected by the gain signal represented as convergent inputs to the visual cortex along with the sensory inputs, given the assumptions that $[r, s] = [\frac{1}{\sigma_G^2}, \sigma_G^2]$. The decoding accuracy is affected when we assume that $[r, s] = [\sigma_G^2, \sigma_G^2]$. The stability or instability of the decoding process given the gain signal might convey important functional roles of the convergent inputs during visual processings. From a Bayesian perspective, the conjugate prior of the poisson distribution (poisson population) is the gamma distribution (gain signal), highlighting the possibility for bayesian inference (Fink, 1997) [2]. Further research should evaluate the optimality or suboptimality of the gain signal as it is transmitted to the visual pathway along with the sensory stimuli.

References

[1] Francis J Anscombe. "Sampling theory of the negative binomial and logarithmic series distributions". In: *Biometrika* 37.3/4 (1950), pp. 358–382.

[2] Daniel Fink. "A compendium of conjugate priors". In: 46 (1997).

[3] Andrew Gelman et al. *Bayesian data analysis*. CRC press, 2013.

[4] Robbe LT Goris, J Anthony Movshon, and Eero P Simoncelli. "Partitioning neuronal variability". In: *Nature neuroscience* 17.6 (2014), pp. 858–865.

[5] Walter W Piegorsch. "Maximum likelihood estimation for the negative binomial dispersion parameter". In: *Biometrics* (1990), pp. 863–867.

[6] Wahiba Taouali et al. "Testing the odds of inherent vs. observed overdispersion in neural spike counts". In: *Journal of Neurophysiology* 115.1 (2016), pp. 434–444.