**Dynamic modeling of spike count data with Conway-Maxwell Poisson variability**

Ganchao Wei1\* and Ian H. Stevenson2-4

*1 Department of Statistics, University of Connecticut*

*2 Department of Psychological Sciences, University of Connecticut*

*3 Department of Biomedical Engineering, University of Connecticut*

*4 Connecticut Institute for Brain and Cognitive Science, University of Connecticut*

*\* Corresponding Author:* ganchao.wei@uconn.edu

# Abstract

In many areas of the brain, neural spiking activity covaries with features of the external world, such as sensory stimuli or an animal’s movement. Experimental findings suggest that the variability of neural activity changes over time and may provide information about the external world beyond the information provided by the average neural activity. To flexibly track time-varying neural response properties, here we developed a dynamic model with Conway-Maxwell-Poisson (CMP) observations. The CMP distribution can flexibly describe firing patterns that are both under- and over-dispersed relative to the Poisson distribution. Here we track parameters of the CMP distribution as they vary over time. Using simulations, we show that a normal approximation can accurately track dynamics in state vectors for both the mean and dispersion parameters. We then fit our model to neural data from neurons in V1 area and “place cells” in the hippocampus. We find that this method out-performs previous dynamic model based on the Poisson distribution. This model provides a flexible framework for tracking time-varying non-Poisson count data and may also have applications beyond neuroscience.

# Introduction

The neural spiking activities will change across the time, even in the controlled settings with constant stimuli. There is growing evidence that not only mean of neural response is non-stationary, but the variability also varies. The drift in both mean and variance provide information about the external world. In most previous research, the neural variability is assumed to be Poisson distributed. However, many experiments suggest that the neurons can be substantially more or less variable than Poisson.

To flexibly track time-varying neural response, in terms of both mean and variance, we developed a dynamic GLM with Conway-Maxwell-Poisson (CMP) observations. The CMP distribution can flexibly account for both over- and under-dispersion in spike count data. To get the closed-form posterior for state vectors with CMP likelihood, we fit the model using a global Gaussian approximation (the Laplace approximation). The procedure can be done efficiently because of the Markovian assumption in state space. We further provide an efficient way to estimate the process noise. After illustrating the proposed method by simulations, we applied it to neurons in V1 area and “place cells” in the hippocampus. The new method can not only capture more information, i.e., dynamics in mean and variance, but also outperforms previous methods based on Poisson assumptions.

Drift…

(Rokni et al. 2007)

(Chestek et al. 2007)

(Stevenson et al. 2011)

(Tomko and Crapper 1974)

Stability…

(Dickey et al. 2009)

(Steinmetz et al. 2021) [shows that correcting for electrode drift reduces some instability]

[some amount of drift is functional] adaptation/plasticity…

(Lesica et al. 2007)

Variability…

(Ghanbari et al. 2019)

(Stevenson 2016)

(Fenton and Muller 1998)

(Barbieri et al. 2001)

(Maimon and Assad 2009)

(Churchland et al. 2010)

(Churchland et al. 2011)

(Eden et al. 2004)

(Brown et al. 2001)

(DeWeese and Zador 1998)

# Methods

Here we consider a dynamic GLM with Conway-Maxwell Poisson (CMP) observations to describe time-varying spike counts. We first introduce the model. Although the CMP distribution allows us to flexibly model non-Poisson variability, one major challenge with using this model is that there are no closed-form posteriors for the CMP likelihood. Here, we fit the model using a global Gaussian approximation, and we discuss several additional technical challenges that arise when using the CMP distribution with a dynamic GLM.

## Dynamic Conway-Maxwell Poisson Model

A count observation , such as the spike count for a neuron, is assumed to follow the CMP distribution, with parameters and . The probability mass function (pmf) of CMP is

, where is the normalizing constant. The parameter controls different dispersion patterns, i.e. equi- (), over- () or under-dispersion (). Three common distributions occur as special cases: 1) the Poisson (), 2) geometric (), and 3) Bernoulli ().

For multiple observations up to steps, such as simultaneous spike counts from neurons, denote the counts at time bin as , for The corresponding CMP parameters at are and . Previous work has examined the CMP-GLM (Chatla and Shmueli 2018; Sellers and Shmueli 2010), and here we focus on the dynamic version of this GLM. The CMP parameters at are modeled by two log-linear models, and , with and , and and denote known predictors. Under the CMP-GLM, the parameters are static. Here, we assume that they progress linearly with a Gaussian noise.

The observations follow conditionally independent CMP distributions, given the state vector .

While the state vector evolves linearly with Gaussian noise:

Given the initial state mean , covariance , linear dynamics and process covariance .

## Inference by Gaussian approximation

To fit the model to data we need to estimate the time-varying state vector In this section, we first assume **F** and **Q** are known. Since the observations are CMP distributed, we cannot estimate in closed form. Instead, here we approximate it by a multivariate Gaussian distribution, , with . The parameters of this Gaussian are found by a global Laplace approximation, i.e., and . The log-posterior is given by:

, where is the log-likelihood. The log-posterior is concave (Gupta et al. 2014), and the Markovian structure of the state vector dynamics makes it possible to optimize by Newton-Raphson (NR) in time (Paninski et al. 2010). After the Newton update, we can further quantify the uncertainty for the CMP parameters and the underlying rates (see Appendix).

There are several technical challenges involved with performing the Newton update with CMP observations. Firstly, in order to find the gradient and Hessian we need to calculate moments of and , which have no closed forms (Shmueli et al. 2005). We can calculate these moments by truncated summation. However, when and , truncated summation is computationally costly since we need many steps for accurate approximation. In this case, we approximate the moments using previous (Chatla and Shmueli 2018; Gaunt et al. 2019) asymptotic results (see Appendix). A second challenge is that the Hessian is not robust to outliers. Outliers often result in the Hessian being close to singular or even positive-definite (see details in Appendix). To ensure robustness, we use Fisher scoring where the observed information is replaced by the expected information. Finally, a third challenge is that the Newton updates take a long time to converge if the initial state estimate is far from the maximum of the posterior, especially when is large. To resolve this issue, we use a smoothing estimate with local Gaussian approximation as a “warm start”. Forward filtering for a dynamic Poisson model has been previously described (Eden et al. 2004), and, here we implement CMP filtering following the same rationale. Let and be the mean and variance for the one-step prediction density and and be mean and variance for the posterior density, then the filtering update for step is given by

Here, to again ensure robustness, we use Fisher scoring when updating the state covariance. We then find smoothed estimates using a backward pass ((Rauch et al. 1965)). Although doing smoothing is fast, the estimates can be inaccurate, especially when there are large changes in the state vectors. In the forward filtering stage, the Gaussian approximation at each step is conducted locally at the recursive prior . This will be statistically inefficient when the recursive prior is too far away from the posterior mode, or when there is a large change in the state vector. Moreover, Fisher scoring reduces the efficiency of the smoother even further. The smoother provides reasonable initial estimates, but estimation accuracy is substantially improved by using Newton’s method to find the global Laplace approximation for the posterior.

## Estimating process noise

For the applications to neural data examined here, we assume that . However, we still need to estimate the process noise . When is small, especially when , different values will have a substantial influence on estimation. One possible way to estimate is to use an Expectation Maximization (EM) algorithm (Macke et al. 2011). However, using the Laplace approximation for during the E-step breaks the usual guarantee of non-decreasing likelihoods in EM, and, hence, may lead to divergence. To avoid that, we could sample the posterior directly by MCMC. However, the lack of closed-form moments for the CMP distribution makes sampling computationally intensive. Here, to estimate robustly and quickly, we instead assume is diagonal and estimate it by maximizing the prediction likelihood in the filtering stage, as in (Wei and Stevenson 2021).

# Results

## Tracking the mean and dispersion of spike counts over time

To illustrate how the dynamic CMP model can track both time-varying mean and dispersion, we simulated a neuron with a time-varying tuning curve, where the response to hypothetical visual stimuli shifts over 100 trials. Here, the neuron’s tuning curve is determined by a linear combination of cubic B-spline basis functions with equally-spaced knots. The stimulus that evokes the highest average response – the “preferred orientation” – is initially ~80 deg, but shifts over the course of the experiment, and the response amplitude also increases over time (Fig. 1A). Meanwhile, the dispersion pattern also changes: the responses are initially over-dispersed relative to a Poisson distribution and then become under-dispersed (Fig. 1B). Noisy observations are sampled from the Conway-Maxwell Poisson distribution at each time (Fig. 1C), mimicking the types of experimental observations collected during adaptation experiments in primary visual cortex (Dragoi et al. 2000) . We then fitted the simulated spike observations using the same predictor variables as the generative model: the covariates for capture the tuning curve with as a 10-knot cubic spline basis expansion of the orientation, and the covariate for the dispersion parameter does not depend on the stimulus orientation . The fitted results match the ground truth well, for both the mean (Fig. 1C) and Fano factors (variance-to-mean ratio, Fig. 1D).

|  |
| --- |
| Chart, histogram  Description automatically generated with medium confidence |
| **Figure 1. A simulated neuron with a shifting firing and dispersion pattern.** (A) The tuning curve of the neuron shifts over time, with the preferred stimulus orientation changing and the response amplitude increasing. (B) At the same time, the variability in spiking changes from being over-dispersed relative to a Poisson distribution to under-dispersed. This leads to a decreasing Fano factor (variance-to-mean ratio) from 1.9 to 0.4 overall. (C) To illustrate the shifts, we show the tuning curve at two time points: Trial 20 (blue) and Trial 80 (red). Dots denote observed spike counts. The solid lines are the ground truth in mean firing rate, while the corresponding dashed lines are the fitted values. (D) To illustrate the shift in dispersion over time we show the true (solid) and estimated (dashed) Fano Factor for two specific stimuli as a function of time. The dispersion for the early preferred orientation is shown in cyan, while the dispersion for the late preferred orientation is shown in yellow. |

This model-based approach provides estimates of tuning curves and dispersion at each time point. In cases where the tuning curve and variability change simultaneously, this approach can efficiently track both. By using the model with CMP observations, rather than Poisson or negative-binomial observations, the Fano Factor can be both <1 (under-dispersed) and >1 (over-dispersed).

Changes in tuning have been widely documented in systems neuroscience both due to changing environment and spontaneous nonstationarity. Changes in variability also occur, but have been less well studied. With the CMP model, the mean and dispersion are both tracked and, thus, changes in variability can occur even when the mean is stable. To illustrate this potential, we simulated a neuron whose mean firing rate is controlled to be constant, but whose Fano factor varies over time. Here is a 5-knot cubic B-spline basis expansion of the orientation and . The model recovers the true mean firing rate (Fig. 2A) and capture the fluctuations in variance (Fig. 2B) at the same time. However, the estimated Fano factor is somewhat oversmoothed when the process noise is optimized by maximizing the predictive likelihood (see Methods).

|  |
| --- |
|  |
| **Figure 2. Constant mean with fluctuations in dispersion.** (A) The first two panels show the true mean firing rate and the simulated observations. The last panel show the fitted mean response. (B) Although the mean response is constant, the Fano factor varies across the trial (black line). The colored line show the fitted result. |

Although dynamic Poisson models have been applied in some neuroscientific settings, when spike counts are not Poisson distributed the model estimates can be biased. Since the dispersion influences estimates of the process noise , estimates of the mean in the dynamic Poisson model can be affected by over- or under-dispersion. To illustrate this interaction here we simulate a place cell from the hippocampus whose “place field” drifts over time. The true mean is determined by a Gaussian function where the preferred position varies over time. The spike counts are then generated by CMP distributions, here over-dispersed with constant dispersion parameter (). We fit 1000 observations randomly sampled from 100 “runs” of a linear-track. We find that, in this data-limited regime, the dynamic Poisson model and the dynamic CMP model give substantially different estimates of the time-varying place field (Fig. 3A). The dynamic Poisson model, in this case, under-estimates the firing rate at the true preferred position and under-estimates the uncertainty (Fig. 3B).

|  |
| --- |
|  |
| **Figure 3. Dynamic CMP and dynamic Poisson estimates differ.** Simulated over-dispersed place cell spiking is generated by the dynamic CMP model with . (A) We then fit dynamic CMP and dynamic Poisson models with a 2-knot B-spline expansion for position ( for the CMP). (B) When evaluating the response at the true preferred position for each run, the dynamic Poisson estimates are biased (under-estimated) and the uncertainty is also underestimated. The solid line gives the MAP estimates of mean firing rate, and the dashed lines show one S.D. credible intervals. The standard deviations of dynamic CMP estimates are calculated using the truncated summations (see details in Appendix), while the standard deviations for the dynamic Poisson model are from a log-normal distribution. |

## Application to Experimental Data

We next applied our method to two publicly available datasets of extracellular spike recordings: 1) Utah array recordings of visually-evoked activity from anesthetized macaque primary visual cortex (“V1 data”), and 2) multi-shank silicon probe recordings from hippocampus of a rat running back-and-forth on a linear maze (“HC data”).

### **V1 data**

In the V1 dataset (Kohn and Smith 2016), anesthetized macaque monkeys viewed full-field sinusoidal grating movies while neural activity was recorded by a 96-channel “Utah” array. The waveform segments were sorted by hand with modified competitive mixture decomposition methods (Shoham et al. 2003). Monkeys viewed a movie of sinusoidal gratings with ~100 different drift directions presented in random order (300ms each, 30s movie in total), and the movie was repeated 120 times. For further details on how the data were obtained, see (Kelly et al. 2010; Smith and Kohn 2008).

In this analysis, we set bin size to be 0.3s, but ignoring the spike counts for the first 0.05s for each bin to avoid the marginal effect (?).

* Monkey #1
* Mention CRCNS pvc-11 somewhere
* Here we consider spike counts 50-350ms after stimulus onset.
* https://crcns.org/data-sets/vc/pvc-11/about

The spike counts show that there are two increasing place fields (Fig. 4A), the major one is around 4 degree while the minor one is around 1 degree.

Considering the sequential measurement effect, we analyze the data with single neuron input. The predictors and are circular cubic spline basis expansion, with 5 and 3 knots, of the grating directions. Fitting the model with full and half data (in a speckled pattern) can both recover the pattern successfully (Fig. 4B). When evaluating the Fano factors using 5-step sliding windows for these two place fields (Fig. 4C), we will see that the Fano factors decrease in two place fields. The fitted Fano factors also follow the observed trend and are mostly within the 1-standard error range, which is obtained by Bayes Bootstrapping. We then compare the performance for 7 models, when fitting the data from 74 neurons in V1 area (Fig. 4D). 4 models are dynamic models: (1) dynamic CMP, with 5 knots for and 3 knots for , denoted as dCMP-(5,3); (2) dynamic CMP with , dCMP-(5,1); (3) dynamic CMP with constant , dCMP-(5)- (fit by coordinate descent) and (4) dynamic Poisson, dPoi-(5). The remaining 3 models are static: (1) static CMP, sCMP-(5,3); (2) static CMP, sCMP-(5,1) and (3) static Poisson, sPoi-(5). The held-out log-likelihoods (bits/ spike) relative to homogeneous static Poisson show that the CMP-based models, both dynamic and static perform similarly (dynamic ones are a bit better). These CMP-based models significantly outperform the Poisson-based models.

|  |
| --- |
|  |
| **Figure 4. Mouse V1 data.** Fit the data with 5 circular basis for and 3 circular basis for . (A) The spiking counts of one neuron from V1 area. Two place fields (found by max mean of CMP firing rate, using mean of spikes is gives weird orientation) were marked by dashed lines. The fitted results of using all and half data (hold out in the speckled pattern). (C) The observed Fano factors, averaging by a 15-step sliding window, of two place fields, overlaid by fitted results using all and half data. The dashed lines are one standard deviation from the window estimates, obtained by Bayes boosting. (D) Compare the performance of different (7) models for 74 neurons from V1 area. In these models, 5 are dynamical and the remaining 3 are static. The training and test log-likelihood (bits/ spike), with respect to the homogeneous static Poisson model for all neurons are shown in grey lines. The solid orange lines show the medians, and the dashed lines show the first and third quartiles. |

### Hippocampus data

Citation: (Mizuseki et al. 2013)

In the HC dataset, the rat was running back and forth along the 250cm linear track. The recording holds up to ~66 min. The recorded spikes are further sorted by program KlustaKwick (~~<https://github.com/klustateam/klustakwik>~~) for automatic spike sorting, then by Klusters, <http://neurosuite.sourceforge.net/> for manual adjustment.

* Bin size
* Exact recording from file name - ec014\_468?
* CRCNS hc-3
* For more details see (REF)
* https://crcns.org/data-sets/hc/hc-3/about-hc-3

The spike counts (Fig. 5A) show that there are two place fields. Unlike data in V1 area, the neural response are more position-specific and sparser. Moreover, the spikes are orientation-specific: the spiking activities are different for the same position but different orientations. We then fit the data using 12-knots for , and let . The spiking pattern can be successfully recovered in terms of mean response (Fig. 5B). We then evaluate the fitted Fano factors for two place fields (Fig. 5C). Compared to the neural activities in V1 area, the spikes in Hippocampus area fluctuate more. The dispersion patterns are also specific in both position and orientation. Compared to models based on Poisson distribution or static models, the dynamic CMP also outperforms others. Figure 5D shows the training and held-out log-likelihood (bits/spike) relative to static homogeneous Poisson for 78 neurons, after extracting 5% data as test in the speckled pattern. Unlike V1 data, all dynamic models are generally better than static models. Within groups of dynamic or static models, CMP-based models are consistently better than the Poisson-based models.

Comment: (1) I drop dCMP-(12)-, since unlike V1, it’s very painful for even 1 neuron. (2) Since the signal is too sparse, I just extract 5% data. (3) There are 97 neurons in HC totally, but 19 of them cannot fit with 12 knots. Maybe they are too sparse. I tried using 6 knots, some resolved but some neurons still fail to fit. Fitting with 6 knots performs similarly.

|  |
| --- |
|  |
| **Figure 5. Hippocampus data.** (A) the spikes counts of the neuron in hippocampus area. The blue lines show the linear tracks. The black circles are spiking counts, with radiuses correspond to number of spikes. (B) the heatmap shows fitted mean firing rate. The colored lines show two putative place fields, which are founded by fitted mean response. (C) The log10-Fano factor for two place fields and two directions. (D) The training and test log-likelihoods (bits/spike) with respect to homogeneous static Poisson for 4 models in 78 neurons, when extracting 5% data as test. 4 models are combinations of dynamic/static and CMP/Poisson. The red solid lines are median values, while the dashed lines showing the first and third quartiles. |

# Discussion

Omitted variables can increase the apparent variability of observations via the law of total variance. For example, in the hippocampus, place cell firing is highly variable on different passes through the field (cite Fenton). This may be partially due to joint selectivity to position, speed, and head direction, as well as the influence of theta phase. Here, rather than model these distinct covariates assuming Poisson observations, we allow the variability to be non-Poisson and introduce a dynamic GLM model with CMP observations. The proposed model can track the fluctuations in mean and dispersion for neural spiking activities simultaneously. The global Laplace approximation with smoothing estimates initialization provide an accurate and computational efficient model estimates. After validating our method by simulations, we further applied it to neurons in V1 area and “place cells” in the hippocampus. All these examples show that our model can successfully capture dynamics in response mean and dispersion, and the proposed method outperforms models based on Poisson observations or static parameters.

Although the current model works well for fitting the neural spikes, there are some potential improvements. First, the state transition matrix is currently assumed to be known and fixed to be identity for convenience. This doesn’t allow for interactions between state vectors and may limit the usage in some situation. Because the state vectors progress linearly with a Gaussian noise, the estimation of is straightforward. The EM algorithm for Poisson linear dynamics system (PLDS), a modified model for Poisson factor analysis model can be found in (Macke et al. 2011). Secondly, although the CMP distribution can flexibly model over- and under-dispersed data, the assumed mean-variance relationship may not be appropriate in some cases. To give a more flexible modeling of the dispersion, we can instead consider the generalized count (GC) distribution(del Castillo and Pérez-Casany 2005), which is a further generalization of CMP distribution. Combining the idea of dynamical factor analysis, (Gao et al. 2015) proposed GC linear dynamical system (GCLDS) model. However, in their GCLDS, the dispersion function is static. To track the fluctuations in dispersion, we can further model dynamically as for in the proposed dynamic CMP model.

# Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1931249.

[Some references are not correct, but they are fixed in LaTeX]

# References

**Barbieri R**, **Quirk MC**, **Frank LM**, **Wilson MA**, **Brown EN**. Construction and analysis of non-Poisson stimulus-response models of neural spiking activity. *J Neurosci Methods* 105: 25–37, 2001.

**Brown EN**, **Nguyen DP**, **Frank LM**, **Wilson MA**, **Solo V**. An analysis of neural receptive field plasticity by point process adaptive filtering. *Proc Natl Acad Sci* 98: 12261–12266, 2001.

**del Castillo J**, **Pérez-Casany M**. Overdispersed and underdispersed Poisson generalizations. *J Stat Plan Inference* 134: 486–500, 2005.

**Chatla SB**, **Shmueli G**. Efficient estimation of COM–Poisson regression and a generalized additive model. *Comput Stat Data Anal* 121: 71–88, 2018.

**Chestek CA**, **Batista AP**, **Santhanam G**, **Yu BM**, **Afshar A**, **Cunningham JP**, **Gilja V**, **Ryu SI**, **Churchland MM**, **Shenoy K V**. Single-neuron stability during repeated reaching in macaque premotor cortex. *J Neurosci* 27: 10742–10750, 2007.

**Churchland AK**, **Kiani R**, **Chaudhuri R**, **Wang XJ**, **Pouget A**, **Shadlen MN**. Variance as a Signature of Neural Computations during Decision Making. *Neuron* 69: 818–831, 2011.

**Churchland MM**, **Yu BM**, **Cunningham JP**, **Sugrue LP**, **Cohen MR**, **Corrado GS**, **Newsome WT**, **Clark AM**, **Hosseini P**, **Scott BB**, **Bradley DC**, **Smith M a**, **Kohn A**, **Movshon JA**, **Armstrong KM**, **Moore T**, **Chang SW**, **Snyder LH**, **Lisberger SG**, **Priebe NJ**, **Finn IM**, **Ferster D**, **Ryu SI**, **Santhanam G**, **Sahani M**, **Shenoy K V**. Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nat Neurosci* 13: 369–378, 2010.

**DeWeese M**, **Zador A**. Asymmetric Dynamics in Optimal Variance Adaptation. *Neural Comput* 10: 1179–1202, 1998.

**Dickey AS**, **Suminski A**, **Amit Y**, **Hatsopoulos NG**. Single-unit stability using chronically implanted multielectrode arrays. *J Neurophysiol* 102: 1331–1339, 2009.

**Dragoi V**, **Sharma J**, **Sur M**. Adaptation-Induced Plasticity of Orientation Tuning in Adult Visual Cortex. *Neuron* 28: 287–298, 2000.

**Eden UT**, **Frank LM**, **Barbieri R**, **Solo V**, **Brown EN**. Dynamic Analysis of Neural Encoding by Point Process Adaptive Filtering. *Neural Comput* 16: 971–998, 2004.

**Fenton AA**, **Muller RU**. Place cell discharge is extremely variable during individual passes of the rat through the firing field. *Proc Natl Acad Sci U S A* 95: 3182–3187, 1998.

**Gao Y**, **Buesing L**, **Shenoy K V**, **Cunningham JP**. *High-dimensional neural spike train analysis with generalized count linear dynamical systems*. 2015.

**Gaunt RE**, **Iyengar S**, **Olde Daalhuis AB**, **Simsek B**, **Robert Gaunt BE**. An asymptotic expansion for the normalizing constant of the Conway-Maxwell-Poisson distribution. *Ann Inst Stat Math* 71: 163–180, 2019.

**Ghanbari A**, **Lee CM**, **Read HL**, **Stevenson IH**. Modeling stimulus-dependent variability improves decoding of population neural responses. *J Neural Eng* 16, 2019.

**Gupta RC**, **Sim SZ**, **Ong SH**. Analysis of discrete data by Conway–Maxwell Poisson distribution. *AStA Adv Stat Anal* 98: 327–343, 2014.

**Kelly RC**, **Smith MA**, **Kass RE**, **Lee TS**. Local field potentials indicate network state and account for neuronal response variability. *J Comput Neurosci* 29: 567–579, 2010.

**Kohn A**, **Smith MA**. Utah array extracellular recordings of spontaneous and visually evoked activity from anesthetized macaque primary visual cortex (V1) [Online]. *CRCNS.org* 2016.http://dx.doi.org/10.6080/K0NC5Z4X.

**Lesica NA**, **Jin J**, **Weng C**, **Yeh CI**, **Butts DA**, **Stanley GB**, **Alonso JM**. Adaptation to Stimulus Contrast and Correlations during Natural Visual Stimulation. *Neuron* 55: 479–491, 2007.

**Macke JH**, **Buesing L**, **Cunningham JP**, **Yu BM**, **Shenoy K V.**, **Sahani M**. Empirical models of spiking in neural populations. *Adv Neural Inf Process Syst* 24, 2011.

**Maimon G**, **Assad J a.** Beyond Poisson: Increased Spike-Time Regularity across Primate Parietal Cortex. *Neuron* 62: 426–440, 2009.

**Mizuseki K**, **Sirota A**, **Pastalkova E**, **Diba K**, **Buzsáki G**. Multiple single unit recordings from different rat hippocampal and entorhinal regions while the animals were performing multiple behavioral tasks [Online]. *CRCNS.org* , 2013http://dx.doi.org/10.6080/K09G5JRZ.

**Paninski L**, **Ahmadian Y**, **Ferreira DG**, **Koyama S**, **Rahnama Rad K**, **Vidne M**, **Vogelstein J**, **Wu W**. A new look at state-space models for neural data. *J. Comput. Neurosci.* 29Springer: 107–126, 2010.

**Rauch HE**, **Tung F**, **Striebel CT**. Maximum likelihood estimates of linear dynamic systems. *AIAA J* 3: 1445–1450, 1965.

**Rokni U**, **Richardson AG**, **Bizzi E**, **Seung HS**. Motor learning with unstable neural representations. *Neuron* 54: 653–666, 2007.

**Sellers KF**, **Shmueli G**. A flexible regression model for count data. *https://doi.org/101214/09-AOAS306* 4: 943–961, 2010.

**Shmueli G**, **Minka TP**, **Kadane JB**, **Borle S**, **Boatwright P**. A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *J R Stat Soc Ser C (Applied Stat* 54: 127–142, 2005.

**Shoham S**, **Fellows MR**, **Normann RA**. Robust, automatic spike sorting using mixtures of multivariate t-distributions. *J Neurosci Methods* 127: 111–122, 2003.

**Smith MA**, **Kohn A**. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J Neurosci* 28: 12591–12603, 2008.

**Steinmetz NA**, **Aydin C**, **Lebedeva A**, **Okun M**, **Pachitariu M**, **Bauza M**, **Beau M**, **Bhagat J**, **Böhm C**, **Broux M**, **Chen S**, **Colonell J**, **Gardner RJ**, **Karsh B**, **Kloosterman F**, **Kostadinov D**, **Mora-Lopez C**, **O’Callaghan J**, **Park J**, **Putzeys J**, **Sauerbrei B**, **van Daal RJJ**, **Vollan AZ**, **Wang S**, **Welkenhuysen M**, **Ye Z**, **Dudman JT**, **Dutta B**, **Hantman AW**, **Harris KD**, **Lee AK**, **Moser EI**, **O’Keefe J**, **Renart A**, **Svoboda K**, **Häusser M**, **Haesler S**, **Carandini M**, **Harris TD**. Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science (80- )* 372, 2021.

**Stevenson IH**. Flexible models for spike count data with both over- and under- dispersion. *J Comput Neurosci* 41: 29–43, 2016.

**Stevenson IH**, **Cherian A**, **London BM**, **Sachs NA**, **Lindberg E**, **Reimer J**, **Slutzky MW**, **Hatsopoulos NG**, **Miller LE**, **Kording KP**. Statistical assessment of the stability of neural movement representations. *J Neurophysiol* 106: 764–774, 2011.

**Tomko GJ**, **Crapper DR**. Neuronal Variability - Nonstationary Responses to Identical Visual-Stimuli. *Brain Res* 79: 405–418, 1974.

**Wei G**, **Stevenson IH**. Tracking Fast and Slow Changes in Synaptic Weights From Simultaneously Observed Pre- and Postsynaptic Spiking. *Neural Comput* 33: 2682–2709, 2021.

# Appendix

## Quantifying Uncertainties

After convergence, we have an approximation of the log-posterior , and we can use this approximation to quantify he uncertainty about the CMP parameters, as well as about the mean rate at each time.

The CMP parameters are log-normal distributed. Let , then . Denote the variance of CMP parameters as . The variance can be easily found in Wiki. If I write it out, I need to define some useless notations. . and .

The conditional mean firing rate is , whose variance can be calculated by the Delta method:

We can calculate the moments as in the first section of the appendix, or we can use a simpler approximation when or . Then and .

## Moments approximation for Conway-Maxwell Poisson distribution

To estimate the state-vector for the dynamic CMP model, we need to find first and second moments for and . For ,

Generally, these moments can be calculated by truncated summation.

However, when and , we need many steps for accurate approximation. In this case, we make use of a previous asymptotic results for efficient calculation. Let , and

Then the moments are (The highlights can be found in the reference, the covariance is calculated by myself):

## Gradient and Hessian of the log-posterior

We estimate the state vector by maximizing the log-posterior with Newton-Raphson updates. Denote , the -th update of NR algorithm is .

The gradient is:

The Hessian:

, where

When , the Hessian may be ill-conditioned or even positive-definite. To ensure the robustness, do Fisher scoring, i.e. replace the observed information by the expected information , so that .