

OPEN ACCESS

A Frank mixture copula family for modeling higher-order correlations of neural spike counts

To cite this article: Arno Onken and Klaus Obermayer 2009 *J. Phys.: Conf. Ser.* **197** 012019

View the [article online](#) for updates and enhancements.

You may also like

- [Local fluctuations of the signed traded volumes and the dependencies of demands: a copula analysis](#)
Shanshan Wang and Thomas Guhr
- [Determining the Core Radio Luminosity Function of Radio AGNs via Copula](#)
Zunli Yuan, Jiancheng Wang, D. M. Worrall et al.
- [Beyond linear neural envelope tracking: a mutual information approach](#)
Pieter De Clercq, Jonas Vanthornhout, Maaïke Vandermosten et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

A Frank mixture copula family for modeling higher-order correlations of neural spike counts

Arno Onken^{1,2} and Klaus Obermayer^{1,2}

¹ School of Electrical Engineering and Computer Science, Technische Universität Berlin, Berlin, Germany

² Bernstein Center for Computational Neuroscience Berlin, Berlin, Germany

E-mail: aonken@cs.tu-berlin.de

Abstract. In order to evaluate the importance of higher-order correlations in neural spike count codes, flexible statistical models of dependent multivariate spike counts are required. Copula families, parametric multivariate distributions that represent dependencies, can be applied to construct such models. We introduce the Frank mixture family as a new copula family that has separate parameters for all pairwise and higher-order correlations. In contrast to the Farlie-Gumbel-Morgenstern copula family that shares this property, the Frank mixture copula can model strong correlations. We apply spike count models based on the Frank mixture copula to data generated by a network of leaky integrate-and-fire neurons and compare the goodness of fit to distributions based on the Farlie-Gumbel-Morgenstern family. Finally, we evaluate the importance of using proper single neuron spike count distributions on the Shannon information. We find notable deviations in the entropy that increase with decreasing firing rates. Moreover, we find that the Frank mixture family increases the log likelihood of the fit significantly compared to the Farlie-Gumbel-Morgenstern family. This shows that the Frank mixture copula is a useful tool to assess the importance of higher-order correlations in spike count codes.

1. Introduction

How do populations of neurons in the brain communicate with each other? Single neurons transmit information via all-or-none action potentials, called spikes. Investigating spike counts of populations of neurons within short periods of time is one possibility to gain insight into the neural code of populations. Spike counts between different neurons are typically correlated [1, 2]. It is not clear, however, how important pairwise and higher-order correlations between spike counts of different neurons are in terms of the information that is conveyed. On the one hand, it is debated whether correlations matter at all [3, 4, 5, 6]. On the other hand, it is debated whether pairwise correlations versus higher-order correlations are sufficient for obtaining good estimates of the information [7, 8, 9]. New data acquisition techniques make it possible to record from many neurons simultaneously and to investigate the population coding questions by applying information theoretic methods. In order to apply measures from information theory statistical models of the neural responses are required. Recently, the Ising model was proposed as a model of parallel spike trains [8]. This model parametrizes only pairwise interactions between the responses of different neurons and assumes a very strict distribution for the single neuron spike counts.

Flexible parametric distributions with separate parameters for higher-order correlations are required in order to investigate the impact of higher-order correlations on information measures. The copula approach provides a promising method to construct such distributions [10, 11]. Copulas are multivariate distributions that can be used to generate spike count distributions with arbitrary marginal distributions and different dependence structures. One example of such a copula is the so-called Farlie-Gumbel-Morgenstern (FGM) copula family which has separate parameters for all pairwise and higher-order correlations [10]. However, there are tight constraints on the parameters and only weak correlations can be modeled using the FGM copula.

In this paper we introduce a new copula family and compare it to the FGM copula and to the Ising model. The family is constructed as a finite mixture of copula families of different orders. It has separate parameters for all pairwise and higher-order correlations and can model dependence strengths between independence and perfect deterministic dependence. We show that this model provides a much better fit to data generated by a network of leaky integrate-and-fire neurons. Moreover, we compare the copula-based models to the Ising model. The Ising model has an implicit probability distribution of spike counts, which has smaller variance than mean. We explore the relation to Poisson-like spike count distributions and overdispersed distributions that are more frequently observed experimentally [10, 12, 2]. We show that the deviation to these distributions is considerable in terms of the Shannon information.

In the next section we describe several models of neural responses, including the Ising model and copula-based models of dependent multivariate spike counts. Moreover, we introduce a new copula family, the Frank mixture family, that is capable of modeling strong higher-order correlations. We construct a network of leaky integrate-and-fire neurons that can be used to generate spike trains with detailed higher-order correlations. In section 3 we apply the statistical models to data generated by the leaky integrate-and-fire network and demonstrate the superior flexibility of the Frank mixture model compared to the Ising model and the FGM copula family. The paper concludes with a discussion of the advantages and limitations of the presented models.

2. Methods

Synthetic data are generated by a network of leaky integrate-and-fire neurons. We will first describe the network and then introduce several statistical models of correlated neural responses.

2.1. Network models

To generate data we applied the leaky integrate-and-fire neuron as a very simple yet successful neuron model that is based on membrane potentials. The membrane potential V is determined by the differential equation

$$\tau_m \frac{dV}{dt} = E_L - V + R_m I_s,$$

where τ_m denotes the time constant, E_L denotes the resting membrane potential, R_m is the total membrane resistance, and I_s is the synaptic input current. V is initialized by a value V_{init} . When V reaches a threshold V_{th} , the neuron emits a spike and V is reset to V_{reset} [13]. We parametrized the model with $\tau_m = 20$ ms, $R_m = 20$ M Ω , $V_{th} = -50$ mV, and $E_L = V_{reset} = V_{init} = -65$ mV, which are typical values found in [13].

We applied the α -function to model synaptic input [13]: $I_s = I_{max} \frac{t}{\tau_s} \exp(1 - t/\tau_s)$, where the presynaptic release occurs at time $t = 0$. The function has its peak I_{max} at time $t = \tau_s$. There is a decay which can be varied by the time constant τ_s . We used $I_{max} = 1$ nA for excitatory synapses, $I_{max} = -1$ nA for inhibitory synapses, and $\tau_s = 5$ ms.

We constructed a small network which allows us to vary pairwise correlations and third order correlations of the spike counts generated by the leaky integrate-and-fire neurons. The network is depicted in figure 1. Poisson spike generators project to the leaky integrate-and-fire neurons

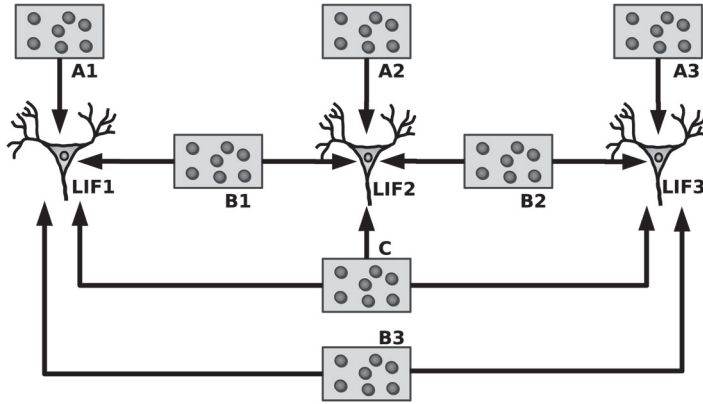


Figure 1. A network of leaky integrate-and-fire neurons to generate spike count data with higher-order correlations. Each rectangle with circles (A1-A3, B1-B3, C) represents two Poisson spike generators: one for the excitatory and one for the inhibitory input. The Poisson spike generators represent populations of neurons that project to three leaky integrate-and-fire neurons (LIF1-LIF3). Connections are illustrated as arrows.

LIF1-LIF3 via inhibitory and excitatory Poisson spike generators. The generators projecting to two neurons (B1-B3) vary the pairwise interactions, whereas the generator projecting to three neurons (C) varies the third order interaction. The rate of excitatory input is always four times the rate of the inhibitory input. We use the generators projecting to only one neuron (A1-A3) to keep the total input rates of the leaky integrate-and-fire neurons constant in spite of varying correlations.

2.2. Ising model

The Ising model is a very popular model in statistical mechanics. It is a maximum entropy model of binary variables called spins that have only pairwise interactions [14]. Recently, it was introduced to neuroscience as a model of spike responses of retinal ganglion cells [8, 9]. In this setting, spike trains are binned into short intervals. A value of either 1 or 0 is assigned to each bin depending on whether at least one spike falls into the bin or not. The bin size should be short enough such that no more than one spike falls into each bin. Successive bins are assumed to be independent, while correlations of parallel spike trains of other neurons are taken into account. The probability of these multivariate binary spike variables $\mathbf{r}$ can be modeled by the Ising distribution [8, 9]

$$P^{\text{Ising}}(\mathbf{r}) = \frac{1}{Z} \exp \left(\sum_i h_i r_i + \sum_{i < j} J_{i,j} r_i r_j \right),$$

where Z is the partition function $Z = \sum_{\mathbf{r}} P_I(\mathbf{r})$ and $r_i \in \{0, 1\}$.

Based on the Ising distribution we can calculate the implicit spike count distribution that arises because of the independence assumption within single neuron spike trains. For spike count bins of length $T \times$ (bin size of binary variables) we get:

$$\begin{aligned} P_T^{\text{Ising}}(\mathbf{r}) &= \sum_{\{\mathbf{Y} | \sum_t Y_{ti} = r_i\}} \prod_{t=1}^T P^{\text{Ising}}(\mathbf{Y}_t) \\ &= \frac{1}{Z^T} \exp \left(\sum_i h_i r_i \right) \sum_{\{\mathbf{Y} | \sum_t Y_{ti} = r_i\}} \exp \left(\sum_{i < j} J_{i,j} Y_{ti} Y_{tj} \right), \end{aligned}$$

where $r_i \in \{0, \dots, T\}$. Due to the independence assumption of successive bins the marginals

of P_T^{Ising} (i.e. the spike count distributions of single neurons) take the form of a binomial distribution $P^{\text{Bin}}(x) = \binom{T}{x} p^x (1-p)^{T-x}$, where p is the probability of a spike in the bin.

2.3. Copulas

A d -copula C is defined as a cumulative distribution function (CDF) on the d -dimensional unit hypercube $[0, 1]^d$ that has uniform marginals [15]. There is a theorem by Sklar which states that copulas can be used to construct multivariate distributions with arbitrary marginal distributions. If C is a copula and $F_1, \dots, F_d$ are marginal CDF's then the function $F_{\mathbf{X}}(r_1, \dots, r_d) \equiv C(F_1(r_1), \dots, F_d(r_d))$ is again a cumulative distribution function and has the marginal distributions $F_1, \dots, F_d$ [15].

Different dependence structures can be modeled by different copulas. Many parametric copula families can be found in the literature. One example is the Frank copula family [16]. Its CDF takes the form

$$C_{\theta}^{\text{Frank}}(\mathbf{u}) = \begin{cases} -\frac{1}{\theta} \ln \left(1 + \left(\prod_{i=1}^d (e^{-\theta u_i} - 1) \right) (e^{-\theta} - 1)^{1-d} \right), & \text{if } \theta \neq 0, \\ \prod_{i=1}^d u_i, & \text{if } \theta = 0. \end{cases} \quad (1)$$

The bivariate Frank family is commutative and radial symmetric: its probability density c_{θ} abides by $\forall (u, v) \in [0, 1]^2 : c_{\theta}(u, v) = c_{\theta}(1-u, 1-v)$ [16]. The scalar parameter θ controls the strength of dependence. As $\theta \rightarrow \pm\infty$ the copula approaches deterministic positive/negative dependence: knowledge of one variable implies knowledge of the other (so-called Fréchet-Hoeffding bounds [15]). The linear correlation coefficient is capable of measuring this dependence. θ must be non-negative for $d > 2$. Copulas of the Frank family are shown in figure 2 for different parameters.

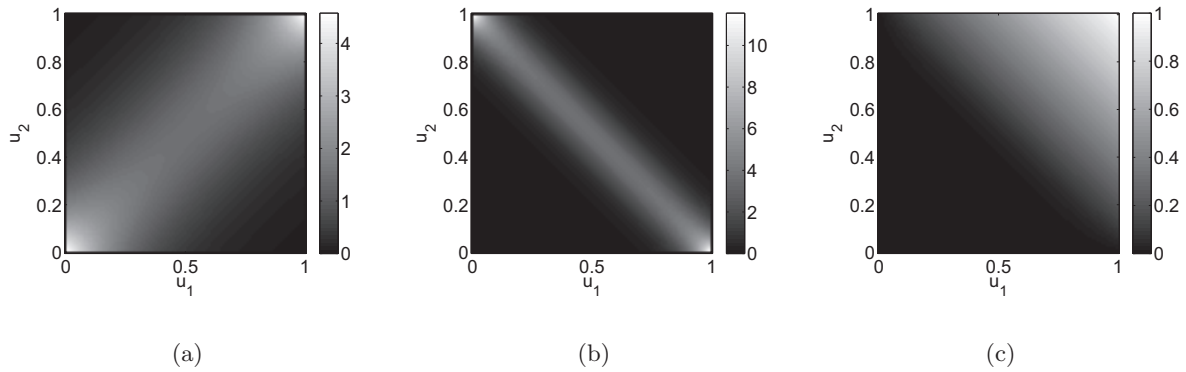


Figure 2. Distribution functions of selected bivariate Frank copulas. (a): Frank copula probability density function for $\theta = 5$. (b): Frank copula probability density function for $\theta = -15$. (c): Frank copula cumulative distribution function for $\theta = -15$.

Another copula family is the Farlie-Gumbel-Morgenstern (FGM) copula family [15]. This family has $2^d - d - 1$ parameters that vary the pairwise and higher-order interactions. Its CDF is given by

$$C_{\alpha}^{\text{FGM}}(\mathbf{u}) = \left(1 + \sum_{k=2}^d \sum_{1 \leq j_1 < \dots < j_k \leq d} \alpha_{j_1 j_2 \dots j_k} \prod_{i=1}^k (1 - u_{j_i}) \right) \prod_{i=1}^d u_i$$

subject to

$$\forall \varepsilon_1, \varepsilon_2, \dots, \varepsilon_d \in \{-1, 1\} : \quad 1 + \sum_{k=2}^d \sum_{1 \leq j_1 < \dots < j_k \leq d} \alpha_{j_1 j_2 \dots j_k} \prod_{i=1}^k \varepsilon_{j_i} \geq 0.$$

These constraints turn out to be very tight. Due to the constraints for α , the correlations that the FGM copula can model are small in terms of their absolute value.

For a given realization $\mathbf{r}$ which can represent the counts of two neurons, we can set $u_i = F_i(r_i)$ and $F_X(\mathbf{r}) = C_\theta(\mathbf{u})$, where F_i can be arbitrary univariate CDF's. Thereby, we can generate a multivariate distribution with specific marginals F_i and a dependence structure determined by C .

Copulas allow us to have different discrete marginal distributions [10, 11]. Typically, the Poisson distribution is a good approximation to spike count variations of single neurons [17]. For this distribution the CDF's of the marginals take the form

$$F_i^{\text{Pois}}(r; \lambda_i) = \sum_{k=0}^{\lfloor r \rfloor} \frac{\lambda_i^k}{k!} e^{-\lambda_i},$$

where λ_i is the mean spike count of neuron i for a given bin size. We will also use the negative binomial distribution as a generalization of the Poisson distribution:

$$F_i^{\text{NBin}}(r; \lambda_i, v_i) = \sum_{k=0}^{\lfloor r \rfloor} \frac{\lambda_i^k}{k!} \frac{1}{(1 + \frac{\lambda_i}{v_i})^{v_i}} \frac{\Gamma(v_i + k)}{\Gamma(v_i) \Gamma(v_i + \lambda_i)^k},$$

where Γ is the gamma function. The additional parameter v_i controls the degree of overdispersion: the smaller the value of v_i , the greater the Fano factor. The variance is given by $\lambda_i + \frac{\lambda_i^2}{v_i}$. As v_i approaches infinity, the negative binomial distribution converges to the Poisson distribution.

Now we have three parametric distributions for the spike count marginals at the disposal: the Poisson distribution with variance equal to the mean, the negative binomial distribution with variance greater than the mean, and the binomial distribution with variance less than the mean. Moreover, we can use the empirical marginals to construct a semi-parametric model with a parametric dependence function and non-parametric marginals. The empirical cumulative marginal of a sample $x_1, \dots, x_n$ is given by $\hat{F}_i(r_i) = \frac{1}{n} \sum_{j=1}^n \mathbf{I}(x_j \leq r_i)$, where $\mathbf{I}$ is the indicator function.

Likelihoods of discrete vectors can be computed by applying the inclusion-exclusion principle of Poincaré and Sylvester [10]. The probability of a realization $\mathbf{r}$ is given by

$$P_{\mathbf{X}}(\mathbf{r}) = \sum_{k=0}^d (-1)^k \sum_{\substack{\mathbf{m} \in \{0,1\}^d, \\ \sum m_i = k}} F_{\mathbf{X}}(r_1 - m_1, \dots, r_d - m_d). \quad (2)$$

Thus, we can compute the probability mass of a realization $\mathbf{r}$ using only the CDF of $\mathbf{X}$.

For spike count data at hand, we can apply the line-search algorithm for constrained nonlinear optimization [18] to estimate the copula parameters using equation 2 as the objective function.

2.4. Frank mixture copula family

The correlation strengths that can be modeled by the FGM family are small due to the tight constraints for the parameters. We will now introduce a new copula family C_θ^{ho} that overcomes

this limitation. We define

$$C_{\theta}^{\text{ho}}(\mathbf{u}) \equiv \sum_{k=2}^d \sum_{1 \leq j_1 < \dots < j_k \leq d} z_{j_1 \dots j_k} C_{\theta_{j_1 \dots j_k}}^{j_1 \dots j_k}(\mathbf{u}), \quad (3)$$

where

$$C_{\theta_{j_1 \dots j_k}}^{j_1 \dots j_k}(\mathbf{u}) \equiv C_{\theta}^{\text{Frank}}(u_{j_1}, \dots, u_{j_k}) \prod_{\substack{i \in \{1, \dots, d\} \\ \setminus \{j_1 \dots j_k\}}} u_i. \quad (4)$$

The family has $2^d - d - 1$ mixture components, each consisting of a Frank copula of different order. These Frank copulas are augmented by independent elements such that each mixture component has dimension d . Hence, each of these mixture components has a single scalar parameter that controls the respective pairwise or higher-order correlation.

The model complexity can be reduced by omitting mixture components that are not necessary or that cannot be estimated reliably. E.g. a model can easily be constructed that has only pairwise and third order parameters.

The Frank mixture family is constructed as a mixture of copulas and thus is a mixture of distributions. A natural choice for an inference algorithm is therefore the expectation-maximization algorithm [19, 20]. We have latent variables z_i that need to be estimated in addition to the Frank copula parameters. The expectation-maximization algorithm has two steps that are repeated until convergence of the parameter values. In the expectation step, the weights z_i can be updated according to

$$z_i^{t+1} = \frac{1}{n} \sum_{s=1}^n \frac{z_i^t P_{C_{\theta_i}^i}(\mathbf{r}_s | \alpha_i^t)}{\sum_j z_j^t P_{C_{\theta_j}^j}(\mathbf{r}_s | \alpha_j^t)},$$

where n is the number of trials and $P_{C_{\theta_i}^i}$ is the probability mass function of the model based on the copula $C_{\theta_i}^i$ (see equation 2.) In the maximization step the copula parameters α_i^{t+1} are optimized for fixed values of z_i^{t+1} . Both steps are repeated until parameter values converge.

The Frank mixture copula can be used to investigate the importance of generalized correlation coefficients on information theoretic quantities. There is no generally accepted definition of higher-order correlation coefficients. There are several candidates, however, one of them being the higher-order cumulants, which generalizes the pairwise covariance [21]. Another candidate is based on a generalization of the pairwise normalized product expectation. We adopt the definition given in [7]:

$$\rho_{i_1, \dots, i_k} \equiv \frac{\langle (r_{i_1} - \langle r_{i_1} \rangle) \dots (r_{i_k} - \langle r_{i_k} \rangle) \rangle}{\langle r_{i_1} \rangle \dots \langle r_{i_k} \rangle},$$

where $\langle \cdot \rangle$ denotes the expectation operator. For the Frank mixture family the impact of a copula parameter $\theta_{i_1, \dots, i_k}$ on this higher-order correlation is given by

$$\begin{aligned} & \left| \frac{\langle (r_{i_1} - \langle r_{i_1} \rangle) \dots (r_{i_k} - \langle r_{i_k} \rangle) \rangle - \langle (r_{i_1} - \langle r_{i_1} \rangle) \dots (r_{i_k} - \langle r_{i_k} \rangle) \rangle_{\theta_{i_1, \dots, i_k}=0}}{\langle r_{i_1} \rangle \dots \langle r_{i_k} \rangle} \right| \\ &= \left| \frac{\sum_{x_1, \dots, x_d} (x_{i_1} - \lambda_{i_1}) \dots (x_{i_k} - \lambda_{i_k}) \psi_{i_1, \dots, i_k} \prod_{\substack{n \in \{1, \dots, d\} \\ \setminus \{i_1 \dots i_k\}}} F_n(x_n - m_n)}{\lambda_{i_1} \dots \lambda_{i_k}} \right|, \end{aligned}$$

where

$$\psi_{i_1, \dots, i_k} = \sum_{s=0}^d (-1)^s \sum_{\substack{\mathbf{m} \in \{0,1\}^d \\ \sum m_i = s}} \left(C_{\theta_{i_1, \dots, i_k}}^{\text{Frank}}(F_{i_1}(x_{i_1} - m_{i_1}), \dots, F_{i_k}(x_{i_k} - m_{i_k})) - \prod_{j=1}^k F_{i_j}(x_{i_j} - m_{i_j}) \right)$$

Here we assume that the weight vector $\mathbf{z}$ is the canonical unit vector with all components set to zero except the $i_1, \dots, i_k$ component. For this weight vector the impact of the copula parameter is maximal. Note that the impact on this higher-order correlation depends on $\theta_{i_1 \dots i_k}$ only.

2.5. Shannon entropy

The Shannon entropy [22] of dependent spike counts $\mathbf{X}$ is a measure of the information that a decoder is missing when it does not know the value $\mathbf{x}$ of $\mathbf{X}$. It is given by

$$H(\mathbf{X}) = \langle I(\mathbf{X}) \rangle = \sum_{\mathbf{x} \in \mathbb{N}^d} P_{\mathbf{X}}(\mathbf{x}) I(\mathbf{x}),$$

where $I(\mathbf{x}) = -\log_2(P_{\mathbf{X}}(\mathbf{x}))$ is the self-information of the realization $\mathbf{x}$.

We can estimate the expectation by drawing samples $\mathbf{x}_i$ with probability $P_{\mathbf{X}}(\mathbf{x}_i)$. The term $\frac{1}{k} \sum_{i=1}^k I(\mathbf{x}_i)$ will then converge to the entropy almost surely, as k approaches infinity [23]. For k sufficiently large the confidence interval of this term is approximately normal distributed.

3. Results

In the following, we will explore the effects of violations of the Ising model assumptions on the Shannon information and the goodness of fit and compare it to the performance of the copula-based models.

3.1. The cost of ignoring short-term non-stationarity

The Ising distribution yields a binomial distribution for single neuron spike counts between 0 and T due to the independence assumption between successive bins of the binary variables. Thus, statistics of successive bins are assumed to be stationary. For the binomial distribution with parameters T and p the variance $Tp(1-p)$ is always less than the mean Tp . This is at odds with many experimental findings of spike count statistics in the cortex. More frequently, spike count variances are observed that are equal to or greater than the mean [10, 12, 2].

The negative binomial distribution was shown to provide a good parametric model of counts with overdispersion [10]. If an Ising model and therefore a binomial distribution is assumed instead of Poisson or negative binomial distributions, then the entropy estimations based on the inappropriate distribution will deviate from the true entropy. The deviation should increase for increasing variance, since the distributions become more distinct.

We quantify the normalized deviation of the entropy between two distributions P and Q by:

$$D^{P,Q} \equiv \frac{H_P - H_Q}{H_P} = 1 - \frac{H_Q}{H_P},$$

where H_X is the entropy of distribution X . For the binomial and the Poisson distribution this yields

$$D^{\text{Bin,Pois}} = 1 + \frac{e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} (x \log_2(\lambda) - \lambda - \log_2(x!))}{\phi},$$

where

$$\phi = \sum_{x=0}^T \binom{T}{x} \left(\frac{\lambda}{T}\right)^x \left(1 - \frac{\lambda}{T}\right)^{T-x} \left(x \log_2\left(\frac{\lambda}{T}\right) + (T-x) \log_2\left(1 - \frac{\lambda}{T}\right) + \log_2\left(\binom{T}{x}\right)\right).$$

For the binomial and negative binomial distribution we get

$$D^{\text{Bin, NBin}} = 1 + \left(1 + \frac{\lambda}{v}\right)^{-v} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \frac{(v+x-1)!}{(v-1)!(v+\lambda)^x} \cdot \frac{(-\log_2(x!) + x \log_2 \lambda + \log_2((v+x-1)!) - \log_2((v-1)!) - x \log_2(v+\lambda) - v \log_2(1 + \frac{\lambda}{v}))}{\phi}.$$

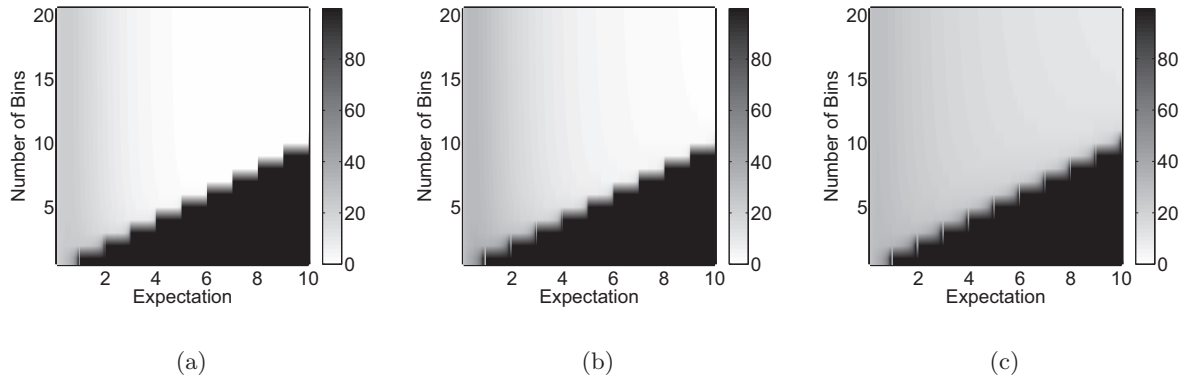


Figure 3. Normalized deviation of entropy in percent between the binomial distribution and different single neuron spike count distributions. The mean and the number of successive independent bins T of the Ising distribution were varied on the axes. The binomial distribution is not defined for parameter combinations in the lower right corners, because for the binomial distribution the mean cannot exceed the number of bins T . (a): Entropy deviations to the Poisson distribution. (b): Entropy deviations to the negative binomial distribution with overdispersion parameter $v = 5$. (c): Entropy deviations to the negative binomial distribution with $v = 1$.

We investigated these deviations of the entropy for different spike rates and bin sizes. The results are shown in figure 3. The variance of the distribution that was used for the comparison increased from (a) to (c). In concert with our expectation the entropy deviation also increased from (a) to (c). Moreover, the deviation increased for decreasing mean. This can be explained in light of the asymptotic distributions: for increasing mean all the distributions approach the normal distribution. The entropy deviation therefore vanishes for increasing means.

3.2. Modeling higher-order correlations

The FGM family can be used to construct multivariate spike count distributions with separate parameters for all pairwise and higher-order correlations. However, the strengths of the correlations that can be modeled using this family are small. To overcome this limitation and to improve the fit further, we introduced the Frank mixture family in this paper (see section 2.4).

We compared the Frank mixture family (see section 2.4) to the FGM family (see section 2.3) on synthetic data generated by a network of leaky integrate-and-fire neurons (see section 2.1). The results are cross-validated, i.e. evaluated on a test set separate of the data that were used to fit the parameters.

The comparison on the test set is shown in figure 4. Since we are interested in the impact of the dependence structure we subtracted the log likelihood that is due to the marginals, i.e. the log likelihood of the product distribution with negative binomial marginals. The remaining residues are due to the copula family. Figure 4 (a) shows the log likelihoods of the residues

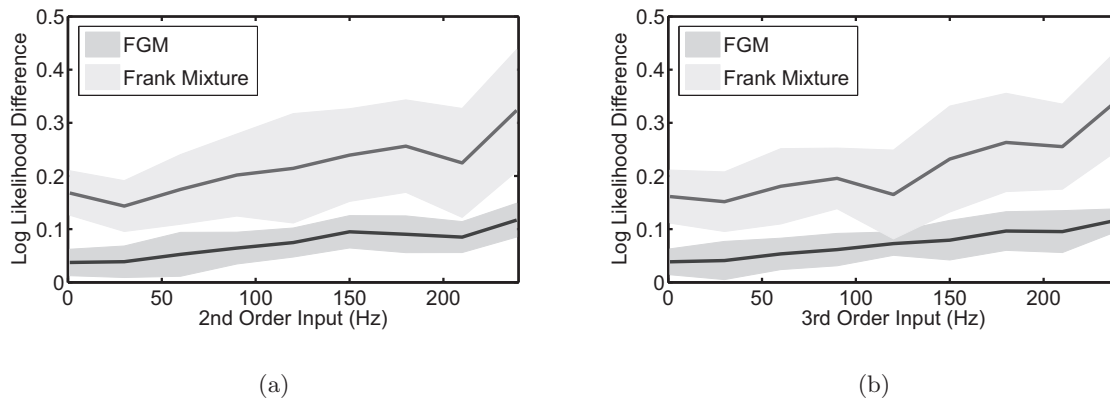


Figure 4. Copula model fits to the data generated by the leaky integrate-and-fire network presented in figure 1 and evaluated on a separate test set. The log likelihood increase is shown for the FGM copula and the Frank mixture copula with negative binomial marginals compared to the independence model with negative binomial marginals. The total excitatory input rate of each neuron was fixed at 840 Hz while the total inhibitory input rate was fixed at 210 Hz. The correlations were increased by increasing the rates of populations B1-B3 and C (see figure 1) while keeping the total input rates constant. The network was simulated for 10 s for the training set and for the test set in each parameter combination. Spikes were counted in 100 ms bins. (a): Log likelihood increase for different pairwise interaction strengths. Each point is averaged over all third order interactions. The shaded area is the standard deviation over third order interactions. (b): Log likelihood increase for different third order interaction strengths. Each point is averaged over all pairwise interactions. The shaded area is the standard deviation over pairwise interactions.

as a function of pairwise interaction strengths in the network and averaged over different third order interaction strengths. The likelihood of the Frank mixture copula family is clearly greater than the likelihood of the FGM family. The likelihood of the fit is increased by the new family in every single point. Figure 4 (b) depicts the log likelihoods of the residues as a function of the third order interaction strength in the network averaged over different pairwise interaction strengths. Again, the likelihood of the proposed Frank mixture family is superior.

On average, the log likelihood is increased by 214 % in terms of the residues. Here it becomes apparent that the Frank mixture family can capture the structure of pairwise and higher-order correlations much better than the FGM family.

4. Discussion

We introduced the Frank mixture family as a new copula family for modeling strong higher-order correlations. As a copula it can be applied to construct multivariate spike count distributions with biologically appropriate marginal distributions such as the negative binomial distribution. We compared the model to the Ising model, a distribution of binary spike trains which has recently gained a lot of attention in neuroscience. Moreover, we compared the model to the FGM copula family, another family that is capable of modeling higher-order correlations. Based on synthetic data generated by a network of leaky integrate-and-fire neurons we found that the fit and the entropy estimation with regard to the marginals and the dependence structure can be improved considerably by applying the Frank mixture family together with appropriate marginal distributions.

The Frank mixture distribution does not overcome the principle problem of models with detailed higher-order correlations, namely the curse of dimensionality. The problem of estimating all higher-order correlations is infeasible even for moderate numbers of neurons. First, the amount of data that is necessary for this estimation is too large and cannot be collected. Second, it is infeasible to compute the likelihood function for larger numbers of neurons. Nevertheless, the complexity of the Frank mixture family can easily be reduced. Since it has separate mixture components for the different orders of correlations, components that can be estimated reliably can be selected. For instance, a model can be constructed that takes correlations up to order three into account but does not have separate parameters for correlations of order greater than three. Directions for future research include the validation of population coding principles that were obtained on the assumption of the Ising model.

Acknowledgments

This work was supported by BMBF grant 01GQ0410.

References

- [1] Bair W, Zohary E and Newsome W T 2001 *J. Neurosci.* **21** 1676–97
- [2] Kohn A and Smith M A 2005 *J. Neurosci.* **25** 3661–73
- [3] Averbeck B B and Lee D 2003 *J. Neurosci.* **23** 7630–41
- [4] Panzeri S, Pola G, Petroni F, Young M P and Petersen R S 2002 *Biosystems* **67** 177–85
- [5] Nirenberg S, Carcieri S M, Jacobs A L and Latham P E 2001 *Nature* **411** 698–701
- [6] Sompolinsky H, Yoon H, Kang K and Shamir M 2001 *Phys. Rev. E* **64** 051904
- [7] Roudi Y, Nirenberg S and Latham P E 2009 *PLoS Comput. Biol.* **5** e1000380+
- [8] Schneidman E, II M J B, Segev R and Bialek W 2006 *Nature* **440** 1007–12
- [9] Shlens J, Field G D, Gauthier J L, Grivich M I, Petrusca D, Sher E, Litke A M and Chichilnisky E J 2006 *J. Neurosci.* **26** 8254–66
- [10] Onken A, Grünewälder S, Munk M and Obermayer K 2009 *Advances in Neural Information Processing Systems* vol 21 ed Koller D, Schuurmans D, Bengio Y and Bottou L (Cambridge: MIT press) pp 1233–40
- [11] Berkes P, Wood F and Pillow J 2009 *Advances in Neural Information Processing Systems* vol 21 ed Koller D, Schuurmans D, Bengio Y and Bottou L (Cambridge: MIT press) pp 129–36
- [12] Osborne L C, Bialek W and Lisberger S G 2004 *J. Neurosci.* **24** 3210–22
- [13] Dayan P and Abbott L F 2001 *Theoretical Neuroscience* (Cambridge: MIT Press)
- [14] Ising E 1925 *Z. Phys.* 253–8
- [15] Nelsen R B 2006 *An Introduction to Copulas* 2nd ed (New York: Springer)
- [16] Frank M J 1979 *Aequationes Math.* **19** 194–226
- [17] Tolhurst D J, Movshon J A and Dean A F 1982 *Vision Res.* **23** 775–85
- [18] Waltz R A, Morales J L, Nosedal J and Orban D 2006 *Math. Program.* **107** 391–408
- [19] Dempster A P, Laird N M and Rubin D B 1977 *J. R. Stat. Soc.* **39** 1–38
- [20] Hu L 2006 *Appl. Finan. Econ.* **16** 717–29
- [21] Stratonovich R L 1967 *Probl. Inf. Transm.* **3**(2) 1–17
- [22] Shannon C E 1948 *AT&T Tech. J.* **27** 379–423
- [23] Robert C P and Casella G 2004 *Monte Carlo Statistical Methods* 2nd ed (New York: Springer)