Chinese Journal of Physiology 53(6): 373-381, 2010
DOI: 10.4077/CJP.2010.AMM036

Invited Paper

The Copula Approach to Characterizing Dependence Structure in Neural Populations

Rick L. Jenison

Department of Psychology,
University of Wisconsin, Madison, Wisconsin 53706, USA

Abstract

The question as to the role that correlated activity plays in the coding of information in the brain continues to be one of the most important in neuroscience. One approach to understanding this role is to formally model the ensemble responses as multivariate probability distributions. We have previously introduced alternatives to linear assumptions of multivariate Gaussian dependence for spike timing in neural ensembles using the probabilistic copula approach. In probability theory the copula "couples" marginal distributions to form flexible multivariate distribution functions for characterizing ensemble behavior. The parametric copula can be factored out of the joint probability density, and as such is independent and isolated from the marginal densities. This greatly simplifies the analysis, and allows a direct examination of the shape of the dependence independent of the marginals. The shape of the copula function goes beyond describing the dependence with a single summarizing statistic. In this review, we illustrate the construction of the copula, and how it contributes to the analysis of information conveyed by populations of neurons.

Key Words: copula, neural dependence, multi-information, Fisher information

Introduction

One of the most important questions in computational neuroscience is what role correlation between neurons plays in the neural coding of sensory information. Specifically does correlated or dependent activity convey more information than neurons firing independently (1, 10, 27). Nirenberg and Latham (24) have argued that should spike train correlations exist, and are important, then they should play a significant role in the coding of information. A fundamental problem is how to separate information carried in correlations from that carried otherwise. Historically Pearson correlation has been employed as a way of characterizing pair-wise dependence, but we know that the correlation coefficient is only valid under the assumption of a multivariate elliptical (*e.g.*, Gaussian) distribution. The correlation coefficient is also problematic in that it is not invariant under monotonic transformations, so a nonlinear change in scale, like a log transform, will change the correlation coefficient. In probability theory a copula "links" arbitrary marginal distributions to form flexible multivariate distribution functions (17, 22). The appeal of the copula is not only that it eliminates the implied reliance on the linear multivariate Gaussian density, but more importantly, it factors the dependence structure out of the full multivariate joint probability density affording a tractable mathematical construction. This elegant construction was derived by Sklar (30), but was largely ignored for many decades until a recent resurgence in the computational finance literature (6) and computational neuroscience (15, 25). There have been significant advances in the construction of univariate (single neuron) probabilistic models of spike trains (4, 5, 8, 18), but our focus has been the probabilistic linkage between neurons through single neuron models to that of full ensemble multivariate probabilistic constructions.

In this paper we review the copula formalism and how we have applied the probabilistic construction to the analysis of sound source information available

Corresponding author: Rick L. Jenison, Ph.D., Department of Psychology, University of Wisconsin-Madison, 1202 W. Johnson St., Madison, WI 53706, USA. Fax: 608-263-5929, E-mail: rjenison@wisc.edu
Received: January 15, 2010; Accepted: January 18, 2010.
 ISSN : 0304-4920. http://www.cps.org.tw

in populations of neurons in primary auditory cortex of cat.

Methods

Copula Probability Theory

The most commonly used measure for dependence, the Pearson product-moment correlation coefficient, was developed on the basis of normal marginals and addresses only linear dependence (19). However, an innovative approach, the so-called copula method, provides the ability to couple arbitrary marginal densities (17, 22, 30). One appealing aspect of the copula is that we eliminate the implied reliance on the multivariate Gaussian when using the correlation coefficient. Perhaps more importantly, we can conveniently factor the dependence structure from independent marginal density.

An N-dimensional copula is a function from $[0,1]^N$ to [0,1]. Sklar's theorem (17, 22, 30) states that any continuous multivariate distribution can be expressed as the copula function $\mathbf{C}(u_1, u_2, ..., u_N)$ evaluated at each of the marginal densities. By the probability integral transform, each marginal $u_i = F_i(x_i)$ has a uniform distribution on [0,1] where $F_i(x_i)$ is the cumulative probability distribution of $p_i(x_i)$ for the random variables X_i. The full joint probability density can be defined as

$$p(x_1, x_2, ..., x_N) = \prod_{i=1}^{N} p_i(x_i) \times \mathrm{c}(u_1, u_2, ..., u_N), \quad [1]$$

where $p_i(x_i)$ is each marginal density and coupling is provided by $\mathrm{c}(u_1, u_2, ..., u_N) = \partial^N \mathbf{C}(u_1, u_2, ..., u_N)/\partial u_1 \partial u_2 ... \partial u_N$, which is itself a probability density. When the random variables are independent, the copula density $\mathrm{c}(u_1, u_2, ..., u_N)$ is uniform and therefore identically equal to one.

The construction of a semi-parametric copula from sampled random variables is shown in Fig. 1. The example illustrates a simultaneous random draw from two different continuous parametric marginal distributions, and the projection through the respective cumulative probability distributions (CDF). This mapping is known as the probability integral transform. For example, the illustrated values for $F_1(x_1)$ and $F_2(x_2)$ in this case are both 0.5. This vector [0.5, 0.5] defines a point in the so-called empirical copula. In principle the empirical marginal distribution would approach a uniform distribution in the limit of Monte Carlo sampling. The dependence structure would therefore be defined by the distribution bounded by the unit square (or unit hypercube as the number of dimensions increase beyond two). Given that the marginal distributions are all equivalently uniform, the copula structure, in principle, is a multivariate

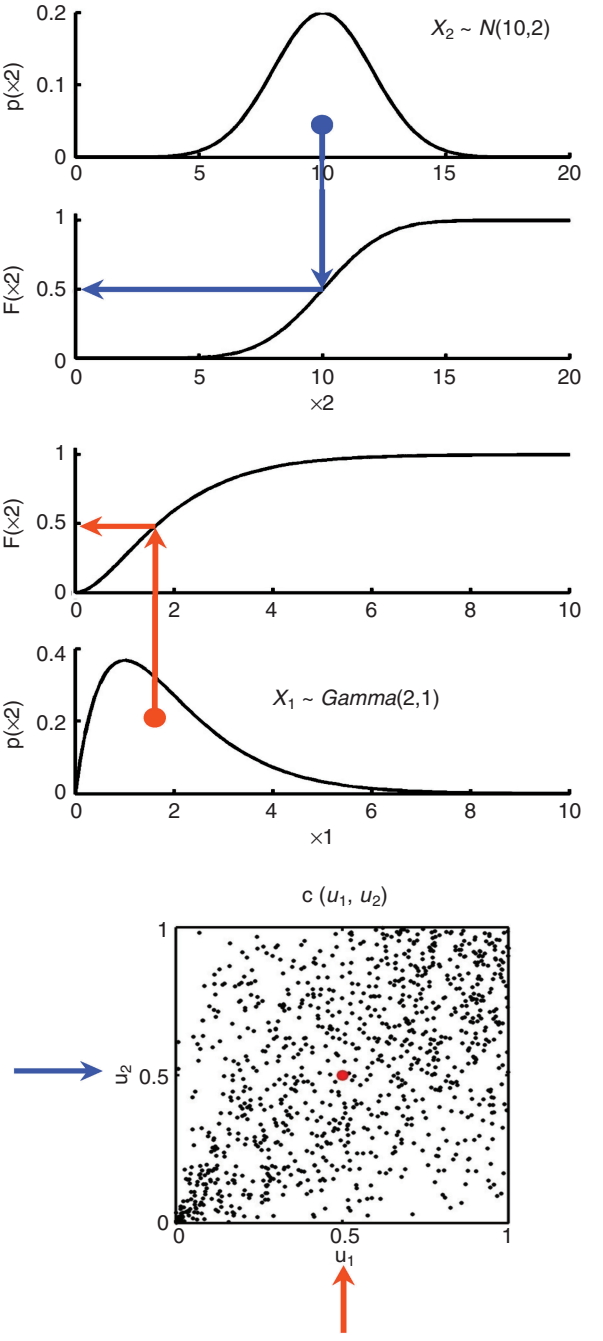

Fig. 1. Schematic of the mapping from sampled random variables x_1 (red) and x_2 (blue) into the empirical copula using the probability integral transform.

density with uniform random variables. Therefore, the process of fitting parameters often proceeds in two phases, first estimating the parameters of the marginal densities, and consequential CDFs, and then estimating the parameters of the copula.

Archimedean Copulas

The Archimedean copula is an important class

of parametric copulas because of its tractability of construction and analysis. The Archimedean family is characterized by a generating function, and its inverse, which provide a very nice method for scaling the copula structure to higher dimensions. For example the Ali-Mikhail-Haq (AMH) (3) copula is an Archimedean copula that can be constructed by an additive generator function φ and its inverse φ^{-1}, if it exists, using the following form

$$\mathbf{C}_{Archimedean}[u_1, u_2, ..., u_N; \alpha] = \varphi^{-1}\left[\sum_{i=1}^{N} \varphi[u_i; \alpha]\right]. \quad [2]$$

The generator function for the AMH copula is $\varphi[u; \alpha] = \log\left[\frac{1-\alpha(1-u)}{u}\right]$, which has an inverse $\varphi^{-1}[y; \alpha] = \frac{1-\alpha}{e^y - \alpha}$, and yields

$$\mathbf{C}_{AMH}[u_1, u_2, ..., u_N; \alpha] = \frac{\alpha - 1}{\alpha - \sum_{i=1}^{N} \frac{1-\alpha+\alpha u_i}{u_i}} \quad [3]$$

where $u_i = F_i(x_i) = \int_{-\infty}^{x_i} p_i(t)dt$ is the marginal cumulative distribution. α is the measure of dependence where $0 \le \alpha < 1$. An Archimedean N-dimensional copula is a proper distribution if and only if the inverse generator function φ^{-1} is completely monotonic (29).

Nelsen (22) has listed 22 known Archimedean copulas that afford different forms of shape and symmetry. The model selection of a specific neural Archimedean copula comes down to a decision based on prior physiological assumptions, together with some application of information criteria, such as Akaike Information Criteria (AIC) (2). The AIC penalizes the negative log maximum likelihood of the estimated model by the number of parameters in the model [-2 log (maximum likelihood) + 2 (number of parameters)]. A smaller relative AIC represents a better model fit while taking into account the complexity of the model.

Neural Dependence Shape

Our recent application of the copula was the analysis of the dependence shape and multi-information of simultaneous first-spike latency recordings in the primary auditory cortex (AI) of the cat (15). First-spike latency is defined as the time from the onset of a physical stimulus until the occurrence of the first action potential (spike) recorded (11). These studies examined the joint activity in response to the presentation of transient sounds in space and the resulting acuity from ensemble (population) decoding by maximum likelihood estimators (12, 14, 28). An empirical joint distribution typical of pairs of single-units recorded from field AI is shown in Fig. 2A.

The joint probability density contours (gray) reflect maximum likelihood fit inverse-Gaussian (IG) marginal probability densities (14) coupled by the AMH copula. As described earlier, the AMH copula is parameterized by a dependence measure α, which for the general multivariate case ranges between 0 and 1. The fit shown in Fig. 2 yields $\alpha = 0.95$ with an AIC equal to 3642. To compare alternative fits for this data set, the AIC for a multivariate independent IG fit was 3728, and the multivariate Gaussian was 3766. As noted above, the smaller AIC represents a better fit taking into account the number of parameters. This figure shows two common characteristics of the joint response of auditory cortical first-spike latencies. First, the marginal distributions are positively skewed, unlike the Gaussian distribution, and second, the dependence structure is not elliptical, which is also inconsistent with a multivariate Gaussian density. The inset shows comparisons of this copula-based joint density to the corresponding multivariate Gaussian fit (red). The estimated AMH copula density $c[u_1, u_2]$ is shown in Fig. 2B, which can be viewed as a function that modulates the product density to form the joint probability density function as defined by Equation 1. In this case the AMH copula density narrows in the lower tails (near zero) and broadens in the upper tails (near one). This is a characteristic of the AMH copula, and characterizes well our common observation of first-spike latency for cortical neurons – the strength of association between ensemble neurons is greater for shorter first-spike latencies compared to responses of longer latencies. When a pair of neurons fire in response to a particular location in space, the shorter latencies (which reflect a stronger response) are more consistent within the ensemble – longer latencies (weaker responses) tend to be less associated. This analysis suggests that not only is a multivariate Gaussian density no longer a prerequisite assumption in delineating the structure of dependence among neurons, but employing a multivariate Gaussian density with this data would actually provide misleading information, particularly for the behavior of the tail regions that greatly influence estimation. An important outcome of this approach is that a decoding analysis based on simultaneous recordings from cortical neurons can now take advantage of parametric probability models using the proper dependence structure and marginals. Discussion of the impact on neural coding of tighter association for stronger responses will be discussed later.

A survey of several candidate copula shapes are illustrated in Fig. 3C-H corresponding to the Clayton (7), Frank, and Gaussian copula (see Nelsen (22) for a review) – all with equivalent Kendall's tau (τ).

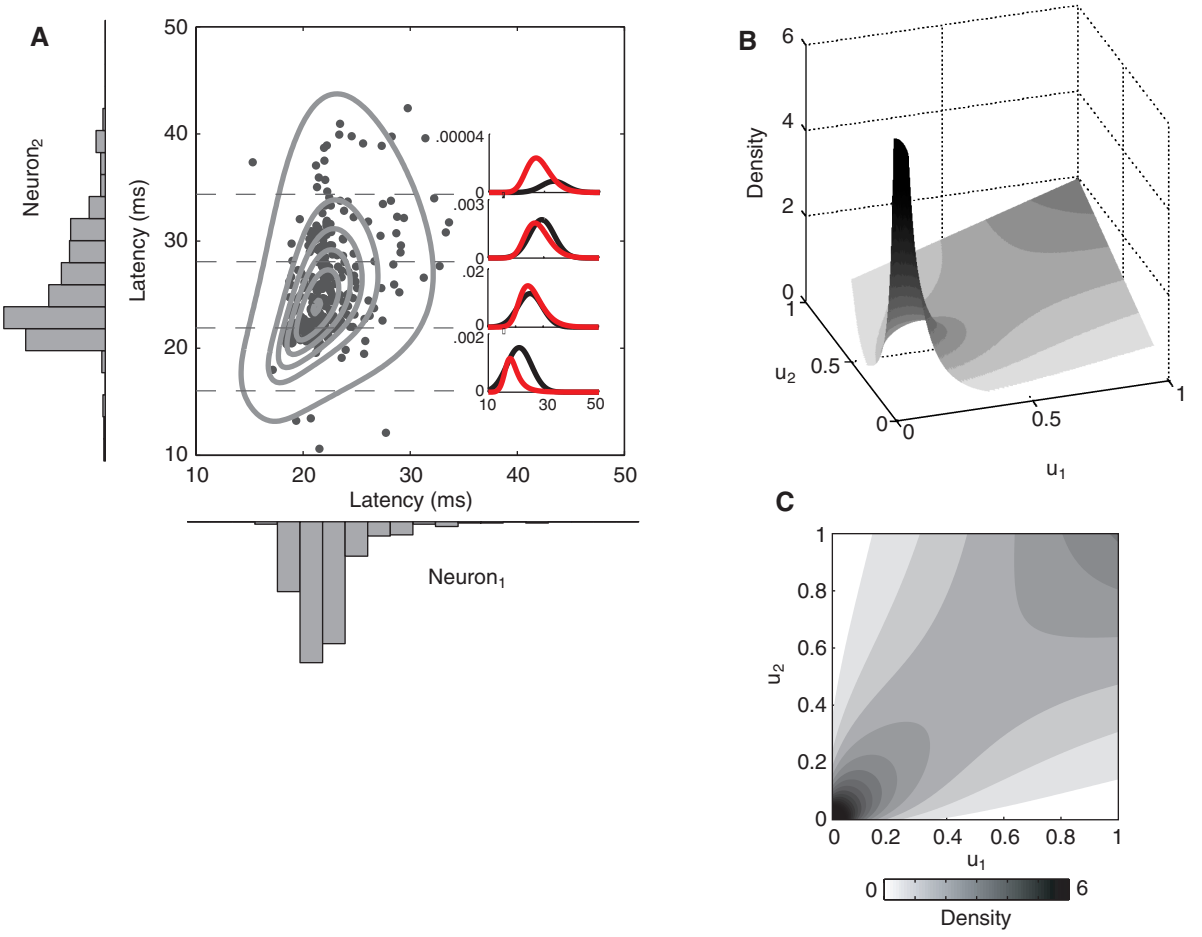

Fig. 2. (A) Joint conditional probability density of first-spike latency recorded from two cortical neurons in response to a given sound source in space. Gray contours correspond to estimated iso-densities based on an Archimedean copula. Inset shows conditional cross sections of the joint density based on the copula (gray) and estimated multivariate Gaussian (black). Marginal histograms are shown (B) surface view of copula density c[u_1, u_2] estimated from data shown in panel A. (C) 2D view of iso-density contours.

Kendall's tau is a measure of rank dependence commonly computed nonparametrically, but does have an analytical relationship with respect to most parametric copulas. Regarding classes of copulas, the AMH, Clayton and Frank copulas are Archimedean, however the Gaussian is not. The Gaussian copula is a multivariate density that captures the elliptical dependence shape, but as is the case with all copulas, is capable of binding arbitrary marginal distributions. It has parameters that would need to be estimated in the same fashion as any parametric copula, typically through maximum likelihood estimation. The standard multivariate Gaussian density of course assumes that the marginals are Gaussian. The Gaussian copula is mathematically less tractable than the Archimedean family of copulas.

Although the AMH copula is somewhat more tractable to work with analytically, it is not as flexible as the Clayton copula in terms of the possible range of dependence. Both of these copulas are asymmetric, with the strongest dependence in the lower tails. The Frank copula (9), also an Archimedean copula, is radially symmetric, a characteristic it shares with the Gaussian copula (Fig. 3, G and H).

Implications for the Neural Code

Mutual Information

The parametric copula can be factored out of the joint probability density, and as such is independent and isolated from the marginal densities. This greatly simplifies the analysis of information, and allows us to examine the shape of the dependence independent of the marginals. In addition, we can easily summarize the copula using well-known information theoretic techniques. The copula has finite support confined to the unit hypercube, which as will be shown, can be exploited for direct estimation of differential entropy and information.

Shannon information can provide a measure of

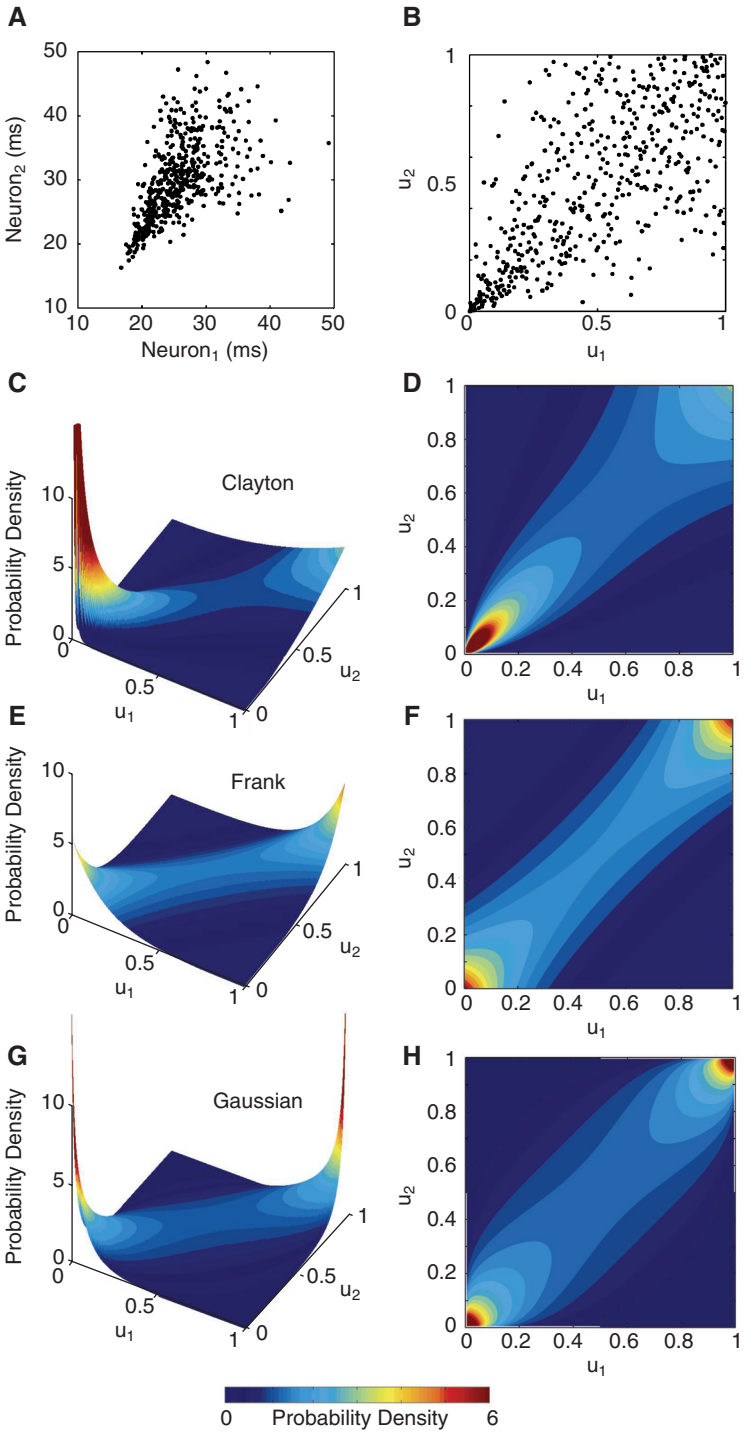

Fig. 3. (A) Scatter plot of first-spike latency of two cortical neurons in response to a single sound source in space. (B) Empirical copula from probability integral transform of the marginals fit to Inverse Gaussian univariate density. Density surface plots for: (C, D) the Clayton copula $\theta = 2$ (E, F) the Frank copula $\theta = 5.8$ (G, H) the Gaussian copula $\rho = .71$. Kendall's $\tau = 0.5$ is the same for all three parametric copulas.

the information available for estimating a physical stimulus given a multivariate ensemble of N neurons. Mutual information between responses is typically based on Shannon's entropy. Let joint entropy be defined as $H[x_1, x_2, ..., x_N | \theta]$, where x_i is the response of the i^{th} neuron and θ is some state or parameter set. Using the copula density, it is straightforward to show that the joint entropy can be split into two terms: one is the sum of entropies due to the independent components of the conditional density, and the other

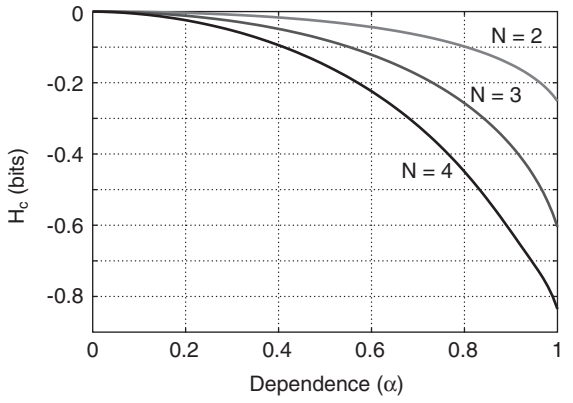

Fig. 4. Copula entropy (negative multi-information) as a function of ensemble size and degree of dependence for AMH copula.

is the quantity due strictly to the dependence structure defined by the copula (See Jenison and Reale (15)). This can be expressed as

$$H[x_1, x_2, ..., x_N|\,\theta] = \sum_{i=1}^{N} H[x_i|\,\theta] + H_c[u_1, u_2, ..., u_N|\,\theta] \quad [4]$$

where the copula entropy is

$$H_c[x_1, x_2, ..., x_N|\,\theta] = -\int_{[0,1]^N} c[u_1, u_2, ..., u_N|\,\theta]\log c[u_1, u_2, ..., u_N|\,\theta]du\,. \quad [5]$$

It follows as a consequence of Equation 4, that copula entropy must be mathematically equivalent to the negative of the mutual information (or multi-information for N > 2) between neurons, and can be computed directly from the dependence structure. Fig. 4 shows how the copula entropy of the AMH copula changes as a function of degree of dependence and the number of neurons in a small ensemble. The entropy of the copula is at its maximum at zero dependence (*i.e.*, uniform), and the copula entropy declines as a function of dependence and ensemble size. So although the decline is rather modest for a pair of neurons, as suggested by Nirenberg and colleagues (23), it accelerates as the ensemble size increases. The importance of factoring entropy in this fashion is that it is often intractable to parametrically model the full multivariate probability density when it is non-Gaussian, and as a consequence, mutual or multiinformation.

Fisher Information

The copula approach eliminates the need for Gaussian assumptions of multivariate neural dependence, and this has important consequences for evaluating neural coding schemes. Our earlier work focused on first-spike latency coding of auditory space by AI neurons that assumed only Gaussian dependence (10). To further demonstrate the importance of accounting for the shape of dependence, simulations were performed using estimated models of cortical receptive field functions. The geometry of log-likelihood functions based on multivariate probability models can quantify how informative a neural code can be for localizing a sound source in space (14). First-spike latency data recorded from 60 neurons (Fig. 5A) were each fit with single spherical basis functions that characterized each individual spatial receptive field as a function of azimuth θ and elevation ϕ of a sound source (see Jenison and colleagues (13, 14, 16) for discussion of spherical approximations). The spherical basis was defined as

$$rf_i(\theta, \phi) = w_i \exp\{\kappa_i(\sin\phi\sin\omega_i\cos(\theta - \xi_i) + \cos\phi\cos\omega_i)\} \quad [6]$$

with the best response (shortest first-spike latency) defined by ξ_i and ω_i for the i^{th} neuron. The probability of observing the i^{th} first-spike latency value x_i given a direction in space was modeled as an inverse Gaussian (IG) density

$$p(x_i|\,\theta, \phi) = \sqrt{\frac{\lambda}{2\pi x_i^3}} \exp\left\{\frac{-\lambda[x_i - r f_i(\theta, \phi)]^2}{2x_i r f_i(\theta, \phi)^2}\right\} \quad [7]$$

The population or ensemble log-likelihood function, including the copula-based dependence structure then follows (26)

$$\log L(\theta, \phi) = \sum_{i=1}^{N} \log p(x_i|\,\theta, \phi) + \log c(u_1, u_2, ..., u_M) \quad [8]$$

For this case, the copula was not conditioned on sound direction, however this could be certainly be added. This construction provided the framework to model our recorded first-spike latency data first as univariate distributions, and then Monte Carlo simulate different degrees of dependence in the population for a selected sound source direction. The log-likelihood function provided a first step in computing a maximum likelihood estimate for sound direction given an ensemble of observations. The performance of four different decoders were evaluated under various assumptions, some ideal while others were not, under differing degrees of dependence defined by α. Conditions were ideal in the sense that the observer or decoder properly assumed the correct dependence structure. The **first**

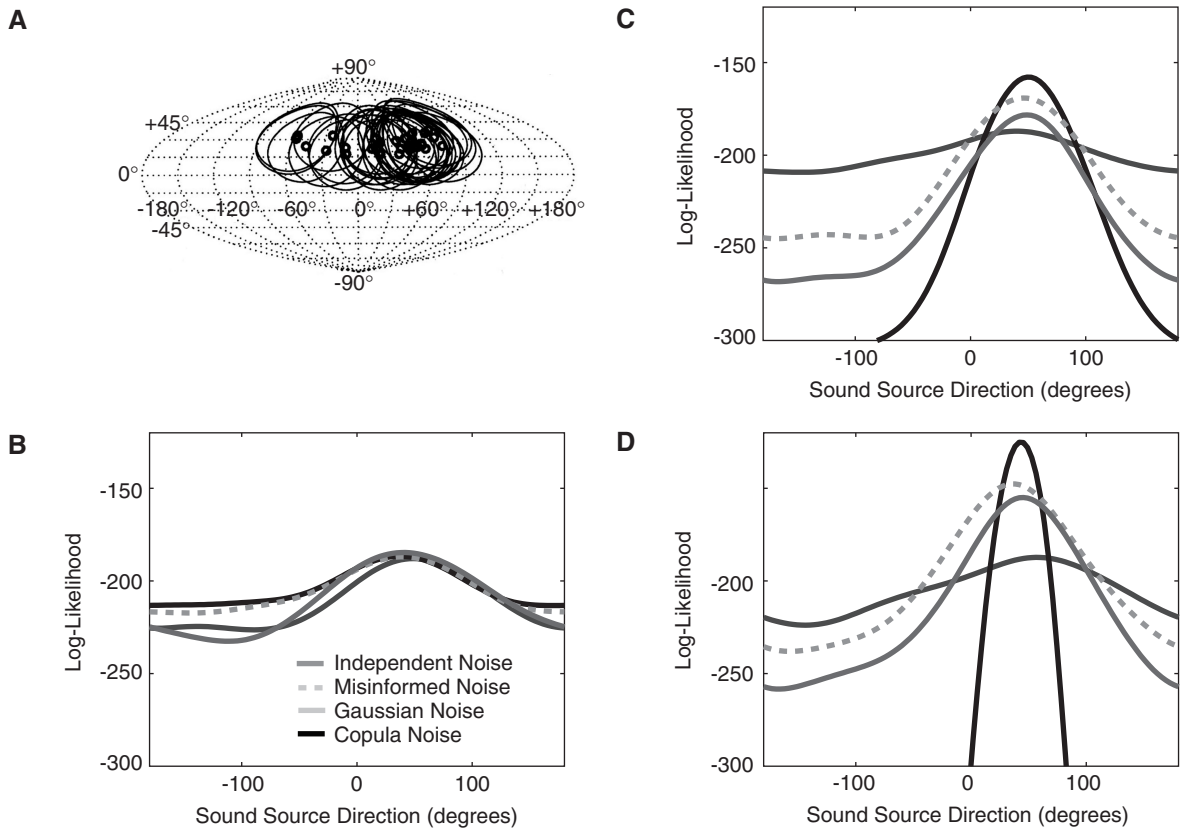

Fig. 5. Log-Likelihood functions based on an ensemble of 60 simulated AI neurons in response (time to first-spike) to a single sound source (transient click) at 45 degrees under different assumptions of neural noise dependence: Independent noise (dark gray), Misinformed noise (dotted light gray), Gaussian noise (light gray) and Clayton copula noise (black) (A) receptive fields for 60 neurons recorded from auditory cortex of the cat used in the simulation. Inverse Gaussian distributed noise is added to the receptive field probabilistic model with different degrees of dependence. Three conditions are shown using parameters having the same Kendall tau. In the misinformed Gaussian condition, the decoder assumes the noise is multivariate Gaussian when in fact it is multivariate Clayton (B) $\theta = 0$, $\rho = 0$, (C) $\theta = 1$, $\rho = 0.5$, (D) $\theta = 2$, $\rho = 0.707$.

condition introduced independent IG noise across the ensemble of modeled neurons, and assumed independence in decoding the direction of the sound source. The **second** condition introduced Clayton dependent noise across the ensemble, however decoded under the assumption of a standard multivariate Gaussian with the same Kendall tau as the simulated Clayton copula (*i.e.*, misinformed). For example, a Clayton $\alpha = 1.0$ and a multivariate Gaussian $\rho = 0.5$ have the same Kendall tau, $\tau = .33$. However, they differ in the shape of the dependence structure. The **third** condition introduced multivariate Gaussian noise across the ensemble, with comparable mean and covariance parameters to the IG-Clayton, and decoded assuming multivariate Gaussian noise. Finally, the **forth** condition introduced Clayton dependent noise across the ensemble, and decoded assuming Clayton dependent noise, again an ideal condition. The ensemble of responses fit to spherical bases are shown as individual contoured receptive fields in Fig. 5A. The spherical projection corresponds to the full auditory space of the cat, with coordinate [0, 0] corresponding to the nose of the animal, and [180, 0] the back of the head.

Results

Example population log-likelihood functions are shown in Fig. 5, B-D given a 60-neuron cortical model localizing a single transient sound located at 45 degrees in azimuth (θ) and 0 degrees in elevation (ϕ). In all four conditions, the maximum of the log-likelihood function could be identified near the true value of the stimulus. However, what is of interest is the curvature about the maximum given the ensemble of observations, commonly referred to as observed Fisher information and defined as

$$I(\theta) = -\frac{\partial^2 \log L(\theta, \phi)}{\partial \theta^2} \quad [9]$$

An overly simplified intuition for Fisher information is that shallower (smaller) the curvature, the more

Table 1. Estimated standard errors from observed Fisher Information under different conditions of neural response dependence. The Clayton copula parameter α and the corresponding correlation coefficient ρ are equivalent in rank correlation (Kendall tau)

α	ρ	Independent	Misinformed Gaussian	Informed Gaussian	Clayton Copula
0	0	9.36	10.26	10.22	10.03
1	0.5	11.76	6.62	5.75	4.45
2	0.707	10.83	5.79	5.34	2.31

uncertainty there is in identifying the maximum likelihood. We explored the impact on the log-likelihood functions and the estimated observed Fisher Information. As a measure of acuity, Fisher Information was inverted to estimate the standard error of the sound direction estimate.

$$s.e. = \sqrt{I(\theta)^{-1}} \qquad [10]$$

This statistic provides a straightforward interpretation using common units to the maximum likelihood estimate of sound direction given the Monte Carlo first-spike latencies. Table 1 shows the measure of acuity as a function of the degree of dependence (equated to the same Kendall tau for α and ρ) and shape of dependence (independent, Clayton and Gaussian). Under all conditions, equated across mean and covariance statistics, the Clayton copula dependence for any parameter value greater than zero yields the most informative estimate when compared to independent, misinformed, and multivariate Gaussian dependent noise. All decoders perform comparably when the dependence parameter is zero and the copula density is consequently uniform, just as we would expect.

These results support the following conclusion – the shape of the noise dependence is an important contributing factor in addition to the magnitude of dependence or correlation. The explanation for why the copula model performs best remains to be explored, but it likely entails the asymmetric structure being more closely matched to certain characteristics of the ensemble marginal distributions. These characteristics may reveal strategies the brain exploits to maximize information transmission, and would be missed has the model assumed only symmetric elliptical (Gaussian) models of dependence.

Discussion

The focus of this review was to show the utility of the copula construction for the analysis of statistical dependence within neural ensembles. In particular we have demonstrated how the copula can contribute to the computation of multivariate information (Shannon and Fisher). These specific examples dealt only with continuous random variables as measures of first-spike latency. The study of copulas has had burgeoning growth in the fields of statistics, finance and risk analysis, however there are a number of limitations that emerge when the analyses are extended to discrete random variables and stochastic processes. These challenges have recently been detailed by Mikosch (20, 21), and are worth consideration when moving in these new directions. The need for new multivariate methods is dire in the field of computational neuroscience. As recording methods evolve to record simultaneously from hundreds, if not thousands, of neurons, the dimensionality of the analysis quickly becomes intractable, a manifestation of the *curse of dimensionality*. Although the copula does not abate the curse, it does constrain the factored multivariate density accounting for the dependence structure to the unit hypercube. This may afford advantages approximating integrals required to compute copula entropy, and equivalently multi-information, in high dimensionality.

Acknowledgments

This work was supported in part by NIH DC04290.

References

1. Abbott, L.F. and Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* 11: 91-101, 1999.
2. Akaike, H. A new look at the statistical model identification. *IEEE T. Automat. Contr.* 19: 723, 1974.
3. Ali, M.M., Mikhail, N.N. and Haq, M.S. Class of bivariate distributions including bivariate logistic. *J. Multivariate Anal.* 8: 405-412, 1978.
4. Brown, E.N., Barbieri, R., Eden, U.T., Frank, L.M. and Feng, J. Likelihood methods for neural spike train analysis. Computational Neuroscience: A Comprehensive Approach. London: Chapman and Hall CRC Press, pp. 253-283, 2004.
5. Brown, E.N., Frank, L.M., Tang, D.D., Quirk, M.C. and Wilson, M.A. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* 18: 7411-7425, 1998.
6. Cherubini, G., Luciano, E. and Vecchiato, W. Copula methods in finance. Chichester: John Wiley & Sons, 2004.

7. Clayton, D.G. Model for association in bivariate life tables and its application in epidemiological-studies of familial tendency in chronic disease incidence. *Biometrika* 65: 141-151, 1978.
8. Eden, U.T., Frank, L.M., Barbieri, R., Solo, V. and Brown, E.N. Dynamic analysis of neural encoding by point process adaptive filtering. *Neural Comput.* 16: 971-998, 2004.
9. Frank, M. On the simultaneous associativity of F(x,y) and x + y - F(x,y). *Aequations Math.* 19: 194-226, 1979.
10. Jenison, R.L. Correlated cortical populations can enhance sound localization performance. *J. Acoust. Soc. Am.* 107: 414-421, 2000.
11. Jenison, R.L. Decoding first-spike latency: A likelihood approach. *Neurocomputing* 38: 239-248, 2001.
12. Jenison, R.L. Models of direction estimation with spherical-function approximated cortical receptive fields. In: Poon, P.W. and Brugge, J.F. (eds). Central Auditory Processing and Neural Modeling. New York: Plenum, pp. 161-174, 1998.
13. Jenison, R.L., Schnupp, J.W.H., Reale, R.A. and Brugge, J.F. Auditory space-time receptive field dynamics revealed by spherical white-noise analysis. *J. Neurosci.* 21: 4408-4415, 2001.
14. Jenison, R.L. and Reale, R.A. Likelihood approaches to sensory coding in auditory cortex. *Network* 14: 83-102, 2003.
15. Jenison, R.L. and Reale, R.A. The shape of neural dependence. *Neural Comput.* 16: 665-672, 2004.
16. Jenison, R.L., Reale, R.A., Hind, J.E. and Brugge, J.F. Modeling of auditory spatial receptive fields with spherical approximation functions. *J. Neurophysiol.* 80: 2645-2656, 1998.
17. Joe, H. Multivariate Models and Dependence Concepts. London: Chapman & Hall, 1997.
18. Kass, R.E. and Ventura, V. A spike-train probability model. *Neural Comput.* 13: 1713-1720, 2001.
19. Mari, D.D. and Kotz, S. Correlation and dependence. London: Imperial College Press, 2001.
20. Mikosch, T. Copulas: Tales and facts. *Extremes* 9: 3-20, 2006.
21. Mikosch, T. Copulas: Tales and facts - rejoinder. *Extremes* 9: 55-62, 2006.
22. Nelsen, R.B. An Introduction to Copulas. New York: Springer-Verlag, 1999.
23. Nirenberg, S., Carcieri, S.M., Jacobs, A.L. and Latham, P.E. Retinal ganglion cells act largely as independent encoders. *Nature* 411: 698-701, 2001.
24. Nirenberg, S. and Latham, P.E. Decoding neuronal spike trains: How important are correlations? *Proc. Natl. Acad. Sci. U.S.A.* 100: 7348-7353, 2003.
25. Onken, A., Grunewolder, S., Munk, M.H.J. and Obermayer, K. Analyzing short-term noise dependencies of spike-counts in macaque prefrontal cortex using copulas and the flashlight transformation. *PLoS. Comput. Biol.* 5: e1000577, 2009.
26. Pawitan, Y. In all likelihood : statistical modelling and inference using likelihood. Oxford New York: Clarendon Press; Oxford University Press, xiii, p. 528, 2001.
27. Quiroga, R.Q. and Panzeri, S. Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* 10: 173-185, 2009.
28. Reale, R.A., Jenison, R.L. and Brugge, J.F. Directional sensitivity of neurons in the primary auditory (AI) cortex: Effects of sound-source intensity level. *J. Neurophysiol.* 89: 1024-1038, 2003.
29. Schweizer, B. and Sklar, A. Probabilistic metric spaces. New York: North Holland, xvi, p. 275, 1983.
30. Sklar, A. Fonctions de répartition a n dimensions et leur marges. *Publ. Int. Stat. Univ. Paris*, pp. 229-231, 1959.