LETTER Communicated by Carlos Brody

Can Spike Coordination Be Differentiated from Rate Covariation?

Benjamin Staude
staude@brain.riken.jp
Computational Neuroscience Group, RIKEN Brain Science Institute, Wako-Shi 351-0198, Japan

Stefan Rotter
stefan.rotter@biologie.uni-freiburg.de
Institute for Frontier Areas of Psychology and Mental Health, 79098 Freiburg, Germany, and Bernstein Center for Computational Neuroscience, 79104 Freiburg, Germany

Sonja Grün
gruen@brain.riken.jp
Computational Neuroscience Group, RIKEN Brain Science Institute, Wako-Shi 351-0198, Japan, and Bernstein Center for Computational Neuroscience, 10099 Berlin, Germany

There has been a long and lively debate on whether rate covariance and temporal coordination of spikes, regarded as potential origins for correlations in cortical spike signals, fulfill different roles in the cortical code. In this context, studies that report spike coordination have often been criticized for ignoring fast nonstationarities, which would result in wrongly assigned spike coordination. The underlying hypothesis of this critique is that spike coordination is essentially identical to rate covariation, only on a shorter timescale. This study investigates the validity of this critique. We provide a decomposition for the cross-correlation function of doubly stochastic point processes, where each of the components corresponds precisely to the concepts of dependence under investigation. This allows us to correct the correlation function for rate effects, which implies that spike coordination and rate covariation are statistically separable concepts of dependence. Furthermore, we present direct and intuitive model implementations of the discussed concepts and illustrate that their difference is not a matter of timescale. Analysis of data generated by our models and analytical description of the relevant estimators reveals, however, that spike coordination dramatically influences the accuracy of rate covariance estimation. As a consequence, extreme parameter combinations can lead to situations where the concept of dependence cannot be identified empirically. However, for a wide range of parameters, the concept of

dependence underlying a given data set can be identified regardless of its timescale.

1 Introduction

The code underlying neuronal computation is heavily discussed in the neuroscience community (e.g., Hebb, 1949; Barlow, 1972; Abeles, 1982; Eggermont, 1990; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997; Shadlen & Movshon, 1999; Singer, 1999; deCharms & Zador, 2000). Meanwhile, there is agreement that neuronal information processing is not performed by single neurons, but rather that populations of neurons form the building blocks of cortical information processing. However, there is no consensus on the code the brain uses to compute and transmit information among such populations. Two opposing theories have been suggested. The first postulates a coding scheme in which populations of neurons contribute with specific rate levels to the encoding of a certain stimulus or a specific behavior (e.g., Georgopoulos, Kalaska, Caminiti, & Massey, 1982; Georgopoulos, Schwartz, & Kettner, 1986; Roelfsema, Lamme, & Spekreijse, 2004). According to the second theory, information processing is carried out by neuronal assemblies (Hebb, 1949), identified by the temporal coordination of the spiking activities across the involved neurons (e.g., von der Malsburg, 1981; Gerstein, Bedenbaugh, & Aertsen, 1989; Abeles, 1991; Singer et al., 1997; Lestienne, 2001). In recent years, the coexistence of both coding strategies has also been discussed (e.g., Riehle, Grün, Diesmann, & Aertsen, 1997; Masuda & Aihara, 2003).

To approach the level of population or ensemble coding, spiking activity of groups of neurons needs to be observed simultaneously. Relevant experimental technologies are now standard equipment in electrophysiological laboratories, and methods to infer the code inherent in parallel spike trains are being developed. Indeed, experimental support was found for all of the discussed coding schemes and in various cortical areas (e.g., Georgopoulos et al., 1986; Vaadia, Aertsen, & Nelken, 1995; Riehle et al., 1997; DeWeese, & Zador, 2006; Nicolelis & Riberio, 2006). However, this did not resolve the controversial discussion, but rather led its scope to other aspects. The question arose whether the applied analysis tools were appropriate in the sense that they accounted for all statistical properties of the data (e.g., Brody, 1999a, 1999b; Oram, Wiener, Lestienne, & Richmond, 1999; Pauluis & Baker, 2000; Musial, Baker, Gerstein, King, & Keating, 2002). In particular, studies that support the assembly hypothesis by reporting (temporal) spike coordination have been criticized for ignoring fast nonstationarities, which would result in wrongly assigned spike coordination. The underlying hypothesis of this critique seems to be that spike coordination and rate covariation are conceptually and statistically similar—the difference being only a matter of timescales (e.g., Rieke et al., 1997).

The aim of this letter is to investigate the validity of this critique. In particular, we are interested in two questions:

Q1: Is the difference between rate covariation and spike coordination really a matter of timescales, or are there more fundamental statistical differences?

Q2: If such a difference exists, can we use this knowledge in data analysis to identify the concept of dependence underlying experimental spike trains?

We approach these questions on different levels of abstraction. The decomposition of the raw cross-correlation function of doubly stochastic point processes (see section 2) reveals that rate covariation and spike coordination are mathematically distinct concepts of dependence. Section 3 presents a model of correlated doubly stochastic Poisson processes that introduces concrete and intuitive implementations of rate covariation and spike coordination in pairs of spike trains. We show that the corrected correlation function of our model implementation depends on only the parameters that control the spike coordination. It turns out in particular that the difference between the two concepts of dependence is not a matter of their timescale (Q1).

The applicability of these theoretical results for data analysis is investigated by estimating the corrected correlation function from three example data sets (section 4). This shows the influence of spike coordination on the estimated corrected correlation function, and reveals that the concept of dependence underlying a given data set can be identified for a large range of parameters. The relevant parameter range is further explored in the discussion section (section 5), together with a treatment of the effects of suboptimal rate estimation. Furthermore, we present extensions of the model implementations to generate more than two parallel processes.

2 The Cross-Correlation Function of Doubly Stochastic Point Processes

We investigate rate covariation and spike coordination in the framework of doubly stochastic point processes. A point process $x(t)$ is called doubly stochastic if its rate profile $l(t)$ is a random signal itself (see section 3 for an example). In this section, we derive a decomposition of the cross-correlation function of a pair of such processes.

The raw cross-correlation function of two time series (not necessarily point processes) $x_1(t)$ and $x_2(t)$ is defined as (e.g., Papoulis, 1991)

$$\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau) := \mathrm{E}[x_1(t)x_2(t-\tau)], \tag{2.1}$$

where $\mathrm{E}[\cdot]$ denotes the ensemble average. As this study investigates only covariance-stationary processes, we omitted the time argument t on the

left-hand side of the definition. If the $x_i(t)$ are doubly stochastic point processes, the ensemble average $\mathrm{E}[\cdot]$ in equation 2.1 consists of the conditional expectation $\mathrm{E}^x[\cdot]$ over realizations of the processes given their rates $l_i(t)$, and the expectation $\mathrm{E}^l[\cdot]$ over realizations of the rate profiles:

$$\mathrm{E}[x_1(t)x_2(t-\tau)] = \mathrm{E}^l[\mathrm{E}^x[x_1(t)x_2(t-\tau)|l_1(t), l_2(t-\tau)]]. \tag{2.2}$$

By definition, the expectation of the process $x_i(t)$ is its rate, meaning that conditioned on the rate, the expectation of $x_i(t)$ is fixed: $\mathrm{E}^x[x_i(t)|l_i(t) = \lambda_i^t] = \lambda_i^t$. Hence, further conditioning of the conditioned process $x_i(t)|l_i(t)$ does not alter its expectation. In particular, we have

$$\mathrm{E}^x[x_i(t)|l_i(t), l_j(t-\tau)] = \mathrm{E}^x[x_i(t)|l_i(t)] = l_i(t), \tag{2.3}$$

even if $l_i(t)$ and $l_j(t-\tau)$ should be correlated. Using the above identity and the definition of the covariance, we add and subtract the product $l_1(t)l_2(t-\tau)$ from the inner expectation $\mathrm{E}^x[\cdot]$ in equation 2.2 and obtain

$$\begin{aligned}
&\mathrm{E}^x[x_1(t)x_2(t-\tau)|l_1(t), l_2(t-\tau)] \\
&\quad = l_1(t)l_2(t-\tau) + \mathrm{E}^x[x_1(t)x_2(t-\tau)|l_1(t), l_2(t-\tau)] \\
&\qquad -\mathrm{E}^x[x_1(t)|l_1(t), l_2(t-\tau)] \cdot \mathrm{E}^x[x_2(t)|l_1(t), l_2(t-\tau)] \\
&\quad = l_1(t)l_2(t-\tau) + \mathrm{Cov}^x[x_1(t), x_2(t-\tau)|l_1(t), l_2(t-\tau)].
\end{aligned}$$

Taking expectation over the rate profiles $\mathrm{E}^l[\cdot]$ and using the definition in equation 2.1 yields the raw cross-correlation function of doubly stochastic point processes:

$$\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau) = \mathrm{CCF}^{\mathrm{raw}}_{l_1,l_2}(\tau) + \mathrm{E}^l\big[\mathrm{Cov}^x[x_1(t)x_2(t-\tau)|l_1(t), l_2(t-\tau)]\big]. \tag{2.4}$$

Equation 2.4 shows that the raw correlation function of doubly stochastic point processes has two separate contributions: the cross-correlation function of the rate profiles (first term) and the mean conditional covariance of the processes given the rates (second term). These separate contributions coincide precisely with the concepts of dependence under investigation. By definition, the cross-correlation function of the rate profiles (first term of equation 2.4) measures the dependence of the rate profiles, and thus captures the contribution of rate covariation on $\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}$. The second term is the covariance of the processes, conditioned on the values of the rate. Because it is conditioned on the rates, this covariance solely measures how much the processes covary if rate effects are neglected. This in turn is typically associated with spike coordination: the alignment of spike times beyond

rate effects. Thus, the second term of equation 2.4 measures the influence of spike coordination on the cross-correlation function.

The decomposition achieved in equation 2.4 shows that spike coordination and rate covariation are mathematically distinguishable components of the cross-correlation function. As such, equation 2.4 serves to decide whether an analytically tractable model generates spike coordination or rate covariation: if it contributes via the first term of equation 2.4, we face pure rate covariation; if it contributes via the second term, we face pure spike coordination; if both terms are influenced, we face a mixture of the two.

The above differentiation between rate covariation and spike coordination relies on the analytical expression for the cross-correlation function. To employ the derived results for data analysis, we need to identify the component that shaped the raw correlation function of a given data set. The approach taken here is to define a predictor that can be used to extract the contribution of spike coordination on $\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau)$. Equation 2.4 suggests correcting for rate contributions by defining a predictor as the cross-correlation function of the rate profiles:

$$\mathrm{Pred}(\tau) := \mathrm{CCF}^{\mathrm{raw}}_{l_1,l_2}(\tau). \tag{2.5}$$

Subtracting this predictor from the raw correlation function $\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}$ yields a corrected correlation function that is independent of rate impact and potential rate covariance:

$$\mathrm{CCF}_{x_1,x_2}(\tau) := \mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau) - \mathrm{Pred}(\tau) \tag{2.6}$$

$$\stackrel{(4)}{=} \mathrm{E}[\mathrm{Cov}[x_1(t), x_2(t-\tau)|l_1(t), l_2(t-\tau)]]. \tag{2.7}$$

The idea of correcting the cross-correlation function for rate effects by subtracting the correlation function of the rate profiles is not new (e.g., Perkel, Gerstein, & Moore, 1967b; Aertsen, Gerstein, Habib, & Palm, 1989). The novelty of this study lies in the model-free analytic identification of the separate contributions to the raw correlation function (see equation 2.4) that allowed us to obtain the explicit expression of the corrected correlation function—the derivation of equation 2.7 from equation 2.6.

3 Model Implementations

This section presents concrete model implementations of rate covariation and spike coordination. Both implementations are based on the same type of point processes for the individual spike trains; however, dependence between these processes is implemented differently: rate covariation is modeled by means of covarying rate profiles, and spike coordination is

implemented by insertion of near-coincident events into background processes (e.g., Holgate, 1964; Grün, Diesmann, Grammont, Riehle, & Aertsen, 1999; Grün, Diesmann, & Aertsen, 2002a; Kuhn, Aertsen, & Rotter, 2003; Daley & Vere-Jones, 2005, and references therein; Ehm, Staude, & Rotter, 2007). Before defining the implementations of the correlation models, we introduce the model of the single processes.

3.1 Single Processes. We model spiking activity of individual neurons by doubly stochastic Poisson processes $x_i(t)$, where $i = 1, 2$ is the process ID. The underlying rate profiles $l_i(t) := \mathrm{E}[x_i(t)]$ are modeled as random signals that are piecewise constant in windows of given length J_l, such that trajectories can be written as

$$l_i(t) = \sum_{k=0}^{M-1} \lambda_i^k \Pi_{I_k}(t - \eta_i), \tag{3.1}$$

where $\Pi_{I_k}(t)$ is the indicator function on the interval $I_k = [kJ_l, (k+1)J_l)$, i.e., $\Pi_{I_k}(t) = 1 \Leftrightarrow t \in I_k$ and zero otherwise. M is the number of time windows of length J_l within the duration of the sample, and η_i is a uniformly distributed variable on $[0, J_l)$ that determines the position of the first rate window with respect to $t = 0$. For each i, the rate levels λ_i^k in the subsequent windows are identical and independently distributed random variables with a common "rate distribution" f_{Λ_i}. As this study concentrates on second-order statistics (variances and correlations), our results depend on only the mean $\lambda_i := \mathrm{E}[\Lambda_i]$ and variance $\sigma_i^2 := \mathrm{Var}[\Lambda_i]$; the particular shape of the rate distribution does not influence the results. In concrete implementations, rate distributions are modeled by gamma distributions to obtain a continuum of only positive rate values (see the distributions on the right of Figures 1A, 1B, and 1D for illustration). Constant rate profiles are generated by delta-like distributions with variance $\sigma_i^2 = 0$ (see Figure 1C).

The duration J_l during which the rate profiles are piecewise constant determines the timescale of the rate fluctuations. We consider here durations that correspond to typical correlation timescales found in experimental data, that is, in the range of tens of milliseconds (e.g., Grün, Diesmann, Grammont, Riehle, & Aertsen, 1999; Nowak, Munk, Nelson, James, & Bullier, 1995; Kohn & Smith, 2005).

To avoid confusion in terminology, we need to make some comments on stationarity. The neurobiological literature tends to identify stationary neuronal firing with "processes of constant rate." However, this identification is not justified in the framework of doubly stochastic point processes. In the model described, the statistical properties of the rate profiles are determined by the fixed rate distribution f_{Λ_i}, which does not vary in time. As a consequence, the ensemble statistics of the process $x_i(t)$ does not depend on the time argument t either. In other words, the doubly stochastic

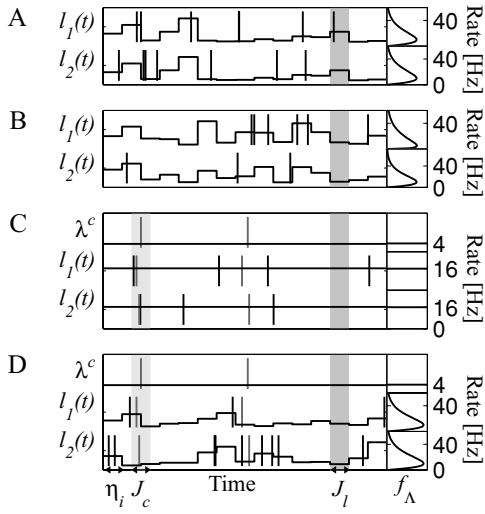

Figure 1: Implementation of rate covariation (A, B) and spike coordination (C, D). Rate profiles ($l_i(t)$ and λ^c, horizontal lines) are illustrated together with spike event times (vertical lines). Rate distributions f_Λ from which the rate levels are drawn are shown on the right. The duration of the piecewise stationary rate segments J_l (timescale of rate (co)-variation) and the timescale of spike coordination J_c are marked by the width of the shaded areas (dark and light gray, respectively, in *C* and *D*); the temporal offset of the first rate window η_i is indicated by a double arrow (D). (A) Identical, time-varying rate profiles for the two processes. (B) Different marginal rate distributions, coupled by a gaussian copula (not shown; correlation parameter $\rho = 0.8$), generating covarying but not identical rate profiles. In *C* and *D*, additional top panels show the coincidence process and its constant rate level ($\lambda^c = 4$ Hz, black line). Its events (dark gray vertical lines) are copied within windows of length J_c into the two background processes that have constant (C) or time-varying but independent (D) rate profiles.

point process $x_i(t)$ is stationary, despite the fact that individual realizations can have time-varying rate profiles. In the following, we therefore avoid the term *stationary* and talk of "constant" or "time-varying" rate profiles, respectively.

3.2 Rate Covariation. Rate covariation between processes is modeled by means of rate profiles (see equation 3.1 for definition) that covary to a predefined degree. To generate covarying rate profiles, we fix the length J_l of the windows of constant rate and couple the temporal offsets η_i for the simultaneous rate profiles (i.e., set $\eta := \eta_1 - \eta_2 = 0$), such that jumps in both rate profiles occur simultaneously in time (see Figures 1A, 1B, and 1D).

Since individual rate profiles are generated by drawing values λ_i^k from the rate distribution f_{Λ_i} for each of the M subsequent windows (see equation 3.1), covariation between the rate profiles is achieved by introducing a dependence between the rate values λ_1^k and λ_2^k.

In the simplest case, the rate levels for both processes are identical in each window, $\lambda_1^k = \lambda_2^k$ for all k, producing identical rate profiles for both processes (see Figure 1A). To generate statistically not identical yet dependent processes, the rate values need not be identical but statistically dependent (see Figure 1B). This can be achieved by coupling the prescribed individual rate distributions f_{Λ_1} and f_{Λ_2} (that control the statistics of the λ_i^k) to a joint, bivariate rate distribution f_{Λ_1,Λ_2} by means of copulas (Nelsen, 1999; Jenison, & Rearle, 2004). A copula is a tool to generate a multivariate distribution from given marginals that allows controling the strength of dependence without altering the prescribed marginals. For the statistics under consideration in this study, the particular choice of copula is not relevant: only the covariance of the marginal rate variables of the individual processes $\gamma_{12} := \mathrm{Cov}[\Lambda_1, \Lambda_2]$ matters (compare equation 3.6). In the concrete example of Figure 1B, the different marginal rate distributions are modeled as gamma distributions with different parameters, coupled by a gaussian copula with correlation parameter $\rho = 0.8$ as in Brunel, Pieczynski, and Derrode (2005). Evidently time-varying yet independent rate profiles are generated by setting $\sigma_i^2 > 0$ and $\gamma_{12} = 0$.

Once a bivariate rate distribution is constructed, our implementation of rate (co)-variation generates pairs of processes by the following paradigm:

1. Construct two rate profiles $l_1(t), l_2(t)$ by drawing values for subsequent windows from the bivariate rate distribution f_{Λ_1,Λ_2}.
2. Realize simultaneous processes $x_1(t)$ and $x_2(t)$ as inhomogeneous Poisson processes with rate profiles $l_1(t)$ and $l_2(t)$, respectively.

The parameters of our implementation of rate covariation are (1) the timescale of joint rate fluctuation J_l, which is also the timescale of dependence between the processes; (2) the means λ_i and variances σ_i^2 of the marginal rate distributions f_{Λ_i} to control the statistics of the single processes; and (3) the covariance γ_{12} of the marginal rate distributions to describe the dependence between the rate profiles.

3.3 Spike Coordination. A different concept of dependence between neurons assumes temporal coordination of spike times, independent of rate effects (Aertsen et al., 1989; Grün et al., 2002b). We implement this concept by injecting near-coincident events into background processes (see Figures 1C and 1D).

Insertion of precise coincidences produces dependence on an arbitrarily short time scale, as opposed to the timescale of dependence J_l in our implementation of rate covariation. To avoid the implementations to be differentiable merely by the timescales of dependence, we allow a temporal

jitter J_c of the inserted events (Grün et al., 1999). The construction is as follows:

1. Realize a Poisson process $x^c(t)$ for the coincidences with a constant coincidence rate of λ^c.
2. Realize Poissonian background processes $x_i^b(t)$ as introduced in the previous section.
3. Copy each spike event of the coincidence process $x^c(t)$, jittered within a window of length J_c around the original event, to the first background process $x_1^b(t)$, and repeat the same procedure for insertion into the second background process.

In the above implementation, individual processes $x_i(t)$ are superpositions of the background process and the jittered coincidence process. Jittering of event times does not affect the Poisson property of the coincidence process (Snyder & Miller, 1991, theorem 3.2.1), and the background process is Poisson by construction. Hence, with both constituents being Poisson processes, the individual processes of our model are Poisson processes (Snyder & Miller, 1991, exercise 2.3.8), and their rate profiles are the sum of the (constant) coincidence rate and the (possibly time-varying) background rates:

$$l_i(t) = l_i^b(t) + \lambda^c. \tag{3.2}$$

Note that the insertion of near-coincident spikes does not pose any constraints on the rates of the background processes $l_i^b(t)$: these can be constant, can vary independently in time, or can even covary. Thus, both implementations are special cases of a more general model. From this model, pure rate covariation is achieved by covarying rates without injecting coincidences ($\gamma_{12} > 0$, $\lambda^c = 0$) and pure spike coordination by combining independent background rates with a positive coincidence rate ($\gamma_{12} = 0$, $\lambda^c > 0$). Evidently arbitrary superpositions of both concepts can also be generated.

3.4 Cross-Correlation. To compute $\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau)$ for the implementation of pure rate covariation (cf. section 3.2), recall that the two processes are independent realizations of their (covarying) respective rate profiles. Hence, conditional on the rates, the processes are independent, and the second term of equation 2.4 vanishes. The computation of the first term, the cross-correlation function of the rate profiles, is rather technical and postponed to the appendix (see equation A.4 with $\eta = 0$). It yields

$$\mathrm{CCF}^{\mathrm{raw}}_{l_1,l_2}(\tau) = \gamma_{12} f(\tau; J_l) + \lambda_1\lambda_2. \tag{3.3}$$

We made use of the triangular function

$$f(\tau; J) := \begin{cases} \dfrac{J - |\tau|}{J} & \text{if } |\tau| \le J \\ 0 & \text{if } |\tau| > J \end{cases},$$

which has a maximal height of 1 and a width of $2J$. As mentioned above, the specific choice of the bivariate rate distribution f_{Λ_1,Λ_2} does not influence the resulting correlation function (see equation 3.3), but only their marginal mean λ_i and covariance γ_{12} are relevant.

To compute the correlation function of our implementation of spike coordination, recall that the injection of spikes alters the rates of the individual processes only by adding the coincident rate to the background rates (see equation 3.2). Since the coincident rate is constant in time, potential (temporal) covariations of the background rates are not influenced by the insertion procedure. Therefore, only the second term of equation 2.4 is affected, showing that the injection of coincident spikes actually implements spike coordination in the sense we discussed after equation 2.4.

For precise coincidences ($J_c = 0$), the second term of equation 2.4 is given by

$$\mathrm{E}[\mathrm{Cov}[x_1(t), x_2(t-\tau)|l_1(t), l_2(t-\tau), J_c = 0]] = \lambda^c \delta(\tau) \tag{3.4}$$

(see Kuhn et al., 2003, appendix A). The peak of mass λ^c at $\tau = 0$ is a result of the precise coincidences that have been injected with rate λ^c. As an effect of the allowed temporal jitter in the copying procedure, not all injected coincidences occur with zero lag ($\tau = 0$), as temporal differences between the injected spikes might occur. Since the times of the individual injected spikes are distributed uniformly in windows of length of J_c around the original spike, the density of their (temporal) difference is the correlation of the two uniform densities, which is a triangle of mass one and base length $2J_c$: $\frac{1}{J_c} f(\tau, J_c)$. The mass of the central peak in equation 3.4 is not changed by the jittering procedure, and we obtain

$$\mathrm{E}[\mathrm{Cov}[x_1(t), x_2(t-\tau) \mid l_1(t), l_2(t-\tau)]] = \frac{\lambda^c}{J_c} f(\tau; J_c). \tag{3.5}$$

By summing the expressions of equations 3.3 and 3.5, we obtain the raw correlation function of the complete model:

$$\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau) = \gamma_{12} f(\tau; J_l) + \lambda_1 \lambda_2 + \frac{\lambda^c}{J_c} f(\tau; J_c). \tag{3.6}$$

Equation 3.6 reveals that the implementations of both, rate covariation and spike coordination alike, contribute to the raw correlation function by triangular peaks (first and last terms, respectively). The heights and widths of these peaks are given by the parameters that control the strength of dependence in the respective implementations. While pure rate covariation without injected coincidences ($\lambda^c = 0$) produces a triangular peak of height γ_{12} and width $2J_l$, spike coordination with independent background ($\gamma_{12} = 0$) produces a triangular peak of height $\frac{\lambda^c}{J_c}$ and width $2J_c$.

Mixtures of the two extreme cases can produce more complicated correlation functions. For instance, if rate covariation is slow with respect to spike coordination ($J_c < J_l$), $\mathrm{CCF}^{\mathrm{raw}}_{x_1,x_2}(\tau)$ has a broad, rate-governed peak of total width $2J_l$, on top of which is a sharp, central spike coordination, or synchrony peak, of width $2J_c$. Cross-correlation functions from experimental data are typically interpreted in that way (Nelson, Salin, Munk, Arzi, & Bullier, 1992; Singer, 1999; Roelfsema et al., 2004). However, the timescales of our implementations can be interchanged arbitrarily (i.e., $J_c > J_l$), such that theoretically, a sharp rate covariance peak can sit on a broad peak of spike coordination.

The above arguments demonstrate that the timescale of dependence, indicated by the width of the raw correlation function, does not allow us to determine the underlying concept of dependence. On the contrary, equation 3.6 explicitly reveals parameter settings for which raw correlation functions are identical even though dependence between the processes is modeled differently. We conclude that the raw correlation function does not provide any information whatsoever about the underlying concept of dependence.

The corrected correlation function of our model is obtained by replacing the right-hand side of equation 2.7 by equation 3.5

$$\mathrm{CCF}_{x_1,x_2}(\tau) = \frac{\lambda^c}{J_c} f(\tau; J_c). \tag{3.7}$$

As opposed to the raw correlation function, the corrected correlation function quantifies the intuitive differences between the presented implementations. Rate covariation ($\lambda^c = 0$) results in a flat corrected correlation function, and spike coordination ($\lambda^c > 0$) results in a central peak. Since this result is independent of the rates of the single processes, and in particular of the timescale J_l of their (co)-variation, the different concepts lead to different results even if both have the same timescale of dependence. We therefore answer our first question, Q1, with no: the difference between spike coordination and rate covariation is not a matter of timescales. How far these results can be used for data analysis (Q2) is the topic of the next section.

Before approaching this question, let us derive the autocorrelation functions of the individual processes in our model implementation. The general

autocorrelation function of doubly stochastic processes can be derived straightforwardly from equation 2.4:

$$\begin{aligned}\mathrm{ACF}^{\mathrm{raw}}_{x_i}(\tau) &:= \mathrm{CCF}^{\mathrm{raw}}_{x_i,x_i}(\tau) = \mathrm{ACF}^{\mathrm{raw}}_{l_i}(\tau) \\ &\quad + \mathrm{E}[\mathrm{Cov}[x_i(t), x_i(t-\tau) \mid l_i(t), l_i(t-\tau)]]. \end{aligned} \tag{3.8}$$

The Poisson process we used in our implementation has zero memory, implying that $x_i(t)$ and $x_i(t-\tau)$ are uncorrelated for $\tau \neq 0$. Hence the second term in equation 3.8 consists of an isolated delta peak of mass λ_i at $\tau = 0$ (see, e.g., Papoulis, 1991, equation 9–109). The autocorrelation function of the rate profiles (first term in equation 3.8) results from equation A.4 (see the appendix). We obtain

$$\mathrm{ACF}^{\mathrm{raw}}_{x_i}(\tau) = \sigma_i^2 f(\tau; J_l) + \lambda_i^2 + \lambda_i \delta(\tau). \tag{3.9}$$

Equation 3.9 shows that the structure in the autocorrelation function must originate from time-varying rate profiles (nonvanishing first term). For our model in particular, the first term of equation 3.9 is a triangular peak of maximal height σ_i^2. Given that the rate varies in time ($\sigma_i^2 > 0$), this term implies a broad central peak that extends from $\tau = -J_l$ to $\tau = J_l$. Hence, the timescale of (potential) rate variation is given by half the width of the central peak of the autocorrelation function.

4 Data Analysis

We demonstrated that the raw correlation function does not allow identifying the underlying concepts of dependence. However, the corrected correlation function (see equations 2.6 and 2.7) extracts only the contribution of spike coordination on the raw correlation function and thus should identify the concept of dependence underlying a given data set.

The key ingredient to obtain the corrected correlation function is the predictor, defined as the cross-correlation function of the rate profiles (see equation 2.5). As opposed to our model implementation, rate profiles and their correlation functions of experimentally recorded neurons are not known, however; they have to be estimated from the data. To investigate to what extent problems in rate estimation influence the ability to extract the amount of spike coordination in experimental spike trains, we apply the presented cross-correlation methods to data that are generated based on our model.

4.1 The Data. We analyze three different data sets: Set1, Set2, and Set3. The first two data sets are examples of pure cases: Set1 contains only rate covariation but no spike coordination (see Figure 1A), while in Set2,

Table 1: Parameter Settings for the Analyzed Data Sets.

Data Set and Parameters	Set1 (Black)	Set2 (Dark Gray)	Set3 (Light Gray)
Firing rate (Hz)	$\lambda_i = 20$	$\lambda_i = 20$	$\lambda_i = 20$
Background rate (Hz)	$\lambda_i^b = \lambda_i = 20$	$\lambda_i^b = 16$	$\lambda_i^b = 16$
Coincidence rate (Hz)	$\lambda^c = 0$	$\lambda^c = 4$	$\lambda^c = 4$
Temporal scale (s)	$J_l = 0.02$	$J_c = 0.02$	$J_c = J_l = 0.02$
Rate variance (Hz^2)	$\sigma_i^2 = 200$	$\sigma_i^2 = 0$	$\sigma_i^2 = 200$
Rate covariance (Hz^2)	$\gamma_{12} = 200$	$\gamma_{12} = 0$	$\gamma_{12} = 0$

coincidences were injected into background processes with constant rates (see Figure 1C). The third data set (Set3) represents a mixture of both concepts. Here, coincidences are injected into background processes with time-varying yet independent rate profiles (see Figure 1D). We chose the parameters such that (1) all three data sets have identical raw correlation functions, and (2) the data sets with time-varying rates (Set1 and Set3) have identical autocorrelation functions.

The concrete parameter values for the data sets are chosen to be in a biologically plausible range: the mean rate of all individual processes is $\lambda_i = 20$ Hz, the coincidence rate for Set2 and Set3 is $\lambda^c = 4$ Hz (Grün et al., 1999), and the timescales of dependence and rate fluctuations are $J_c = J_l = 20$ ms (Nowak et al., 1995). The remaining parameter values are chosen to fulfill constraints (1) and (2) given above (the concrete values given in Table 1 result from equations 3.6 and 3.9). Each data set consists of a single realization, simulated for 1000 seconds. The temporal offset of the first rate window is set to zero ($\eta_1 = \eta_2 = 0$). As performed for experimental data, spike trains are discretized to binary sequences $\hat{x}_i(s)$ by binning (bin size $h = 1$ ms) and clipping. Observe that the time variable s of the binary sequences $\hat{x}_i(s)$ is a unitless integer that denotes the bin number as opposed to the previous section, where t was given in units of time.

To imitate experimental situations most realistically, we conduct our analysis of the three data sets according to the following protocol: begin by estimating the raw correlation function ($\mathrm{CCF}^{\mathrm{raw}}_{x_1x_2}$, equation 2.1), proceed by estimating the appropriate predictor ($\mathrm{Pred} = \mathrm{CCF}^{\mathrm{raw}}_{l_1l_2}$, equation 2.5), and conclude by interpreting the corrected correlation function ($\mathrm{CCF}_{x_1x_2} = \mathrm{CCF}^{\mathrm{raw}}_{x_1x_2} - \mathrm{Pred}$, equation 2.6).

4.2 Empirical Correlation Functions. The theoretical correlation function of the previous section was defined via the expectation over realizations $\mathrm{E}[\cdot]$ of the time-shifted product of the processes (see equation 2.1). Since our data sets consist of one realization (trial) only, expectation over realizations cannot be calculated. However, the correlation functions of our implementations do not depend on the time argument t (see equation 3.6), hence, we can reduce the variance of the estimator by averaging the product of the time-shifted processes over time. We thus estimate the raw correlation

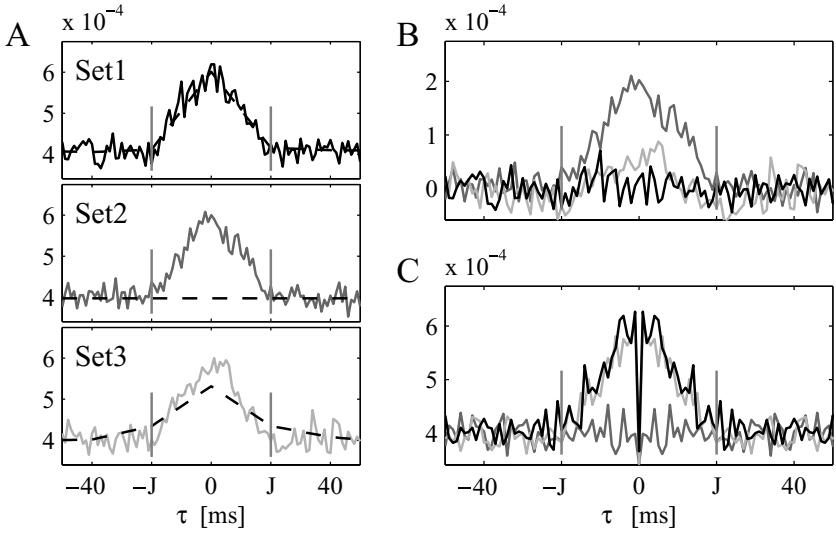

Figure 2: Cross- and autocorrelation functions of Set1 (black), Set2 (dark gray), and Set3 (light gray). (A) Raw cross-correlation functions (solid lines) of pure rate covariation (Set1, top), coincidences injected into background with constant rate (Set2, middle), and coincidences injected into background with time-varying, independent rates (Set3, bottom), together with the respective empirical predictors (dashed). (B) Corrected correlation functions of the three data sets. (C) Autocorrelation functions of the three data sets (peak at $\tau = 0$ removed). Theoretical timescales of dependence and rate variations ($J = J_c = J_l = 20$ ms) are shown as vertical gray lines at $\pm J$.

function $\mathrm{CCF}^{\mathrm{raw}}_{x_1 x_2}$ of equation 2.1 by the time-averaged, unbiased empirical correlation function,

$$\widehat{\mathrm{CCF}}^{\mathrm{raw}}_{\hat{x}_1,\hat{x}_2}(\tau) = \frac{1}{N - |\tau|} \sum_{s=1}^{N} \hat{x}_1(s)\hat{x}_2(s-\tau) \tag{4.1}$$

(Perkel et al., 1967b; also known as the Bartlett spectral estimate in the frequency domain: Bartlett, 1963). N is the total number of bins, and τ is in units of the bin size h. In order for equation 4.1 to be well defined, spike sequences are zero-padded such that $\hat{x}_i(s) = 0$ for $s < 1$ or $s > N$.

As expected from the analytical results of the previous section, the chosen parameterization produces identical raw empirical correlation functions for all three data sets (see Figure 2A). To extract the contribution of spike coordination and thus identify the underlying concept of dependence, the corrected correlation function, equation 2.7, has to be estimated by subtracting

the predictor, that is, the cross-correlation functions of the rate profiles, equation 2.5. This in turn requires estimating the firing rates from the data.

4.3 Empirical Predictor and Corrected Correlation Function. Correct estimation of firing rates from neuronal spike data without the possibility of average over repeated realizations is known to be difficult and prone to errors (e.g., Johnson, 1996; Nawrot, Aertsen, & Rotter, 1999). Erroneous rate estimation, however, is known to complicate interpretation of the resulting corrected correlation function (e.g., Brody, 1999a, 1999b). To avoid these complications, we use prior knowledge of the rate profiles that underlie our samples: that these are constant or at least constant in exclusive windows of fixed but unknown length J_l, and that the two rate profiles have no temporal offset ($\eta_i = 0$).

We decide whether the rate profiles are constant or time varying on the basis of the autocorrelation function of the single processes. Depending on the outcome, the rates profiles are estimated according to one of the two strategies A and B presented below. Since the empirical correlation function of equation 4.1 estimates the probability of coincident spikes in bins of length h (as opposed to the joint density defined in equation 2.1, which was in units of Hz^2), we will describe estimated rate profiles also in units of probabilities of spikes per bin (rather than in units of Hz as in the previous sections).

4.3.1 Strategy A: Average over Time. If the rates can be assumed to be constant, rate estimation by averaging over time is the method of choice. The constant rate estimate is defined by

$$\hat{\lambda}_i = \frac{1}{N} \sum_{s=1}^{N} \hat{x}_i(s). \tag{4.2}$$

The cross-correlation function of two constant signals is simply given by their product; hence, the empirical predictor, equation 2.5, for constant rates reads

$$\widehat{\mathrm{Pred}}_{\mathrm{const}}(\tau) = \hat{\lambda}_1 \cdot \hat{\lambda}_2.$$

The above strategy was chosen for the data of Set2. Its autocorrelation functions (see Figure 2C, dark gray line) are flat and thus indicate that the rates are constant. The resulting time-averaged predictor $\widehat{\mathrm{Pred}}_{\mathrm{const}}$ is a straight line (see Figure 2A, middle, dashed line), which removes only a constant offset when subtracted from the raw correlation function. Consequently, the central peak of width $2J_l$ remains unaltered in the corrected correlation function of Set2 (see Figure 2B, dark gray line), as theoretically predicted for spike coordination (equation 3.7).

4.3.2 Strategy B: Estimate Rate in Discrete Windows. If the firing rate varies in time, it has to be estimated in a time-resolved manner. The appropriate method here is to estimate the rate in windows that correspond to the piecewise constant firing rates (see section 3). The broad central peaks of the autocorrelation functions (see Figure 2C, black and light gray lines) indicate time-varying rate profiles in Set1 and Set3. Using equation 3.9, we estimate the length of the piecewise constant rate windows $\widehat{J}_l$ by half the width of the autocorrelation functions' central peaks.

Estimates of the rate values in subsequent windows of length $\widehat{J}_l$ are given via the spike counts $\hat{X}_i^k = \sum_{s=k\widehat{J}_l+1}^{(k+1)\widehat{J}_l} \hat{x}_i(s)$, where we used the fact that the rate profiles have no temporal offset ($\eta_i = 0$). Concatenating the spike counts per window divided by the window size $\widehat{J}_l$ yields the estimated rate profiles,

$$\hat{I}_i(s) = \sum_{k=0}^{M-1} \frac{\hat{X}_i^k}{\widehat{J}_l} \Pi_{I_k}(s), \tag{4.3}$$

where $\Pi_{I_k}(s)$ denotes the discrete indicator function on I_k (compare equation 3.1). The estimator of the predictor for data with time-varying rate profiles is the unbiased empirical cross-correlation function of these signals (compare equation 4.1):

$$\widehat{\mathrm{Pred}}_{\mathrm{timevar}}(\tau) = \widehat{\mathrm{CCF}}^{\mathrm{raw}}_{\hat{I}_1,\hat{I}_2}(\tau) = \frac{1}{N-|\tau|} \sum_{s=1}^{N} \hat{I}_1(s)\hat{I}_2(s-\tau).$$

The resulting predictors of data sets Set1 and Set3 (dashed lines in Figure 2A, top and bottom) show broad central peaks and appear similar to the raw cross-correlation functions (solid lines in the respective plots). The predictor of Set1 follows the raw correlation function perfectly; hence, when subtracted from the raw correlation function, it removes both the offset and the central peak (see Figure 2B, black line). These data were generated by rate covariance only ($\lambda^c = 0$); hence, the flat corrected correlation function is predicted by the analytical results (see equation 3.7).

For Set3, the predictor is not perfectly identical with the raw correlation function (see Figure 2A, bottom). However, subtraction still reduces the central peak substantially. As a consequence, the corrected correlation function appears almost flat (see Figure 2B, light gray line). This flatness of the corrected correlation function is in sharp contrast to the analytical prediction, as dependence between the processes of Set3 is modeled by spike coordination only ($\lambda^c = 4$ Hz, $\gamma_{12} = 0$). Equation 3.7 explicitly predicts the same peaky corrected correlation function as for Set2: the only difference between these data sets is the variance of the background rates σ_i^2 (compare Table 1), which by equation 3.7 does not affect the corrected correlation

function. However, the empirical corrected correlation function of Set2 shows the predicted central peak, while that of Set3 is almost flat (see Figure 2B, dark and light gray lines, respectively).

4.4 Analytical Versus Empirical Findings for Set3. The seemingly contradictory results for the analytical and empirical corrected cross-correlation function of Set3 (spike coordination in conjunction with independently varying background rates) have their origin in the shape of the empirical predictor. The question is: Why does the empirical predictor have a broad central peak even though the rate profiles are independent? We clarify this question by deriving the estimator for the empirical predictor analytically.

The estimated rate profiles $\hat{I}_i(s)$ of equation 4.3 are piecewise constant signals whose values $\frac{X_i^k}{J_l}$ in subsequent windows are independent. A similar computation as for the rate profiles of equation 3.3 yields their theoretical raw cross-correlation function (see the appendix)

$$\mathrm{Pred}_{\mathrm{timevar}}(\tau) = \mathrm{CCF}^{\mathrm{raw}}_{\hat{I}_1,\hat{I}_2}(\tau) = \frac{1}{J_l^2}\mathrm{Cov}[X_1, X_2] f(\tau; J_l) + \lambda_1\lambda_2. \tag{4.4}$$

The height of the triangular peak of $\mathrm{Pred}_{\mathrm{timevar}}(\tau)$ is determined by the covariance of the counting variables X_i (converted to the units of the rates by scaling with $\frac{1}{J_l^2}$). The counting variables, in turn, are parameter-dependent variables. A similar calculation that led to equation 2.4 shows that their covariance is the sum of the rate covariance plus the covariance of the realizations:

$$\mathrm{Cov}[X_1, X_2] = J_l^2\gamma_{12} + \mathrm{E}[\mathrm{Cov}[X_1, X_2 \mid \lambda_1, \lambda_2]]. \tag{4.5}$$

The jointly injected spikes of Set3 produce a positive covariance of the realizations (second term in equation 4.5). Under the assumption that jittered injected coincidences do not fall into neighboring counting windows, we have (Kuhn et al., 2003, appendix A)

$$\mathrm{E}[\mathrm{Cov}[X_1, X_2 \mid \lambda_1, \lambda_2]] = \lambda^c J_l. \tag{4.6}$$

Combining equations 4.4 to 4.6 finally yields the estimator of the empirical predictor for our implementations when rates are estimated by spike counts in windows of length J_l:

$$\mathrm{Pred}_{\mathrm{timevar}}(\tau) = \gamma_{12} f(\tau; J_l) + \lambda_1\lambda_2 + \frac{\lambda^c}{J_l} f(\tau; J_l). \tag{4.7}$$

As opposed to the theoretical predictor in equation 3.1, its estimator in the time-varying case depends on the coincidence rate λ^c. In particular,

$\mathrm{Pred}_{\mathrm{timevar}}(\tau)$ correctly estimates the correlation function of the rates only if the coincidence rate vanishes: if $\lambda^c = 0$. Even more so, this estimator coincides precisely with the analytical raw correlation function of the processes themselves, equation 3.6. As a consequence, subtracting this empirical predictor removes the central peak of the correlation function regardless of the mechanism that originally generated it. As a consequence, the corrected correlation functions of Set1 and Set3 are flat (see Figure 2B, black and light gray lines). The small remaining central peak in Set3 is an effect of near-coincident spikes being copied into neighboring counting windows (Grün et al., 1999), which we neglected in the above derivations.

In summary, if rates are estimated in small windows, effects of spike coordination influence the estimated predictor. Erroneous estimation of the predictor in turn might interfere with our ability to extract the contribution of spike coordination to the raw correlation function. The parameters of Set3 were tuned so that the size of the windows for rate estimation coincides with the timescale of dependence. In this case, the spike coordination could not be inferred, and the concept of dependence underlying the data set could not be identified. Despite these estimation problems, the concept of dependence underlying a given data set can be identified for a large range of parameter combinations, which will be discussed extensively in section 5.1.

5 Discussion

The aim of this study was to investigate differences between spike coordination and rate covariation. We approached this question by decomposing the cross-correlation function of doubly stochastic point processes (see section 2) and by analyzing the presented model implementation of the two concepts of dependence analytically (see section 3) and empirically (see section 4). We revealed that spike coordination and rate covariation contribute distinct terms to the raw cross-correlation function (see equation 2.4), showing the formal separability of the two concepts of dependence. This difference was underscored by the intuitive model implementations of the proposed concepts, showing in particular that the difference between the two concepts is not a matter of timescales. To identify the concept of dependence underlying a given data set, we used a corrected correlation function (defined by subtracting the cross-correlation function of the rate profiles from the raw correlation function) that extracts the contribution of spike coordination. In pure cases, where the data contained either spike coordination in constant background rates (Set1) or only rate covariation (Set2), the underlying mechanisms could be easily identified. This empirical result was obtained in a case where the two concepts of dependence acted on the same timescale ($J_l = J_c$), thereby underscoring the fact that the difference between the concepts is not a matter of timescales. However, we were able to construct a case where spike coordination as the underlying concept could not be identified (Set3).

The parameter settings in the above example cases were deliberately chosen as extreme cases to illustrate the main effects. However, in experimental situations, we may well face arbitrary superpositions of the two concepts, corresponding to a much larger parameter range. The next section investigates the influence of such effects on the detectability of the underlying concept of dependence, followed by a discussion of alternative rate estimators. As shown in this study, a thorough calibration of the applied analysis tool on the basis of well-defined models is of fundamental importance to avoid misinterpretation of results. To account for the increasing interest in analysis tools for more than two parallel spike trains (Brown, Kaas, & Mitra, 2004), section 5.3 points out ways to extend the proposed model classes to generate more than two simultaneous processes.

5.1 Identification of Spike Coordination. Common experimental practice would interpret the corrected correlation functions of both Set1 and Set3 as indications for rate covariation. Our results show that this interpretation might not be justified, as the underlying concept of Set3 was spike coordination. Furthermore, section 4.4 revealed that the failure in detecting spike coordination as the underlying mechanism of Set3 is not due to the finiteness of the sample size or relatively weak spike coordination; the problem arises from the overestimation of rate covariation induced by spike coordination. However, for a wide range of parameters, spike coordination can in fact be identified as the underlying concept.

In the problematic data set presented in section 4 (Set3), the timescales of the two concepts were identical: $J_l = J_c$. Let us consider the case $J_l > J_c$, as rate covariation is typically associated with slow, stimulus-induced effects (Eggermont, 1990). In this case, the width of the central peak of the autocorrelation function suggests that the rate should be estimated in windows of length J_l (as described in section 4.3). With this estimator of the rate profiles, we can use the results of section 4.4 to predict the corrected empirical correlation

$$\mathrm{CCF}_{\hat{x}_1,\hat{x}_2}(\tau) = \mathrm{CCF}^{\mathrm{raw}}_{\hat{x}_1,\hat{x}_2}(\tau) - \mathrm{Pred}_{\mathrm{timevar}}(\tau) \tag{5.1}$$

$$\overset{3.6,4.7}{=} \frac{\lambda^c}{J_c} f(\tau; J_c) - \frac{\lambda^c}{J_l} f(\tau; J_l). \tag{5.2}$$

The right-hand side of equation 5.2 subtracts a triangle of width J_l and height $\frac{\lambda^c}{J_l}$ from a triangle of width J_c and height $\frac{\lambda^c}{J_c}$. If $J_l > J_c$, the subtracted triangle is broader and lower than the first one, resulting in a corrected correlation function in which a sharp central peak remains. Since equation 5.2 is independent of the parameters of rate covariation, this holds even if the underlying rate profiles are not independent but covary. Hence, a peak reflecting spike coordination is left, albeit not necessarily in its full

strength. Thus, if spike coordination is faster than rate (co-)variation, the former is reliably identified as the underlying concept, despite the fact that its strength might be underestimated.

Further evidence for spike coordination is given in data sets where the height of the peak of the cross-correlation function exceeds that of each autocorrelation function (e.g., Perkel et al., 1967b). Assuming only rate covariance, the former is determined by the covariance of the rate variables $\mathrm{Cov}[\Lambda_1, \Lambda_2]$, equation 3.6, while the latter is determined by the variances of the marginals $\mathrm{Var}[\Lambda_i]$, equation 3.9. Using the general inequality $\mathrm{Cov}[\Lambda_1, \Lambda_2] \leq \mathrm{Var}[\Lambda_i]$, we can conclude that in cases of pure rate covariation, the central peak of the cross-correlation function is always smaller than that of either of the autocorrelation functions. Hence, if the peaks of the autocorrelation functions of a given data set are both lower than that of the cross-correlation function, the latter cannot be explained by rate covariation alone, and spike coordination has to be assumed.

The reasoning in both situations can be applied less quantitatively even if the underlying model class is not known. In the first scenario, the dynamics of the rate profile (inferred from the width of the autocorrelation function) is slower than that of the interaction (inferred from the width of the cross-correlation function). This means that the processes interact faster than the dynamics of the individual rates would allow. In the second scenario, the individual processes vary less than they covary. In both cases, the joint statistics cannot be explained by the statistics of the individual processes alone. A mechanism that acts in addition to rate effects has to be postulated: spike coordination.

5.2 Rate Estimation. A major ingredient of the proposed analysis is the predictor, defined as the correlation function of the firing rate profiles. In particular, correct rate estimation has proven to be crucial for differentiating among the discussed concepts of dependence. Indications for time-varying firing rates may be given by the autocorrelation function or the coefficient of variation/Fano factor (e.g., Johnson, 1996; Nawrot et al., 2007), where the former also provides the timescale of rate variation if the model class of the underlying point processes is known. Without any knowledge of the underlying point process model, however, the shape of the autocorrelation function can at most indicate that rates vary in time. In this case, sliding window rate estimations (Grün et al., 2002b) or other convolution methods (e.g., Nawrot et al., 1999; Dayan & Abbott, 2001) are still the standard choice for single trial data, although novel techniques are being developed (e.g., Kass, Ventura, & Brown, 2005). As opposed to the optimal rate estimator we used for our data sets, such approaches would wash out the boundaries of the rate levels, thereby reducing the accuracy of the rate estimation. As a consequence, the predictor, given by the cross-correlation of the estimated rates, would have a broader, less pronounced central peak.

Another method typically used for estimating the firing rates is the peristimulus time histogram (PSTH; Perkel, Gerstein, & Moore, 1967a; Shimazaki & Shinomoto, 2007). In this approach, rates are estimated by averaging across trials, that is, from repetitions of the same experiment, which assumes the same response in time and across trials. Violation of these assumptions leads to a poor rate estimate, which results in an inaccurate prediction of the rate contributions to the CCF^{raw} (Brody, 1999a, 1999b; Grün, Riehle, & Diesmann, 2003). The data analyzed in section 4 consist of one trial only, in which case the PSTH corresponds to the single-trial spike-count-based rate estimator employed for the data sets with time-varying rate profiles (Set2 and Set3). If we had considered multiple trials instead, the doubly stochastic character of the model would imply a new rate profile in every trial. This, however, violates the assumptions of across-trial stationarity made by the PSTH, rendering this method of rate estimation not applicable for the proposed models.

In summary, without prior knowledge of the underlying model class, rate estimation will most likely be suboptimal. The resulting predictor leaves a residual central peak in the corrected correlation function, which is likely to be wrongly interpreted as spike coordination (Aertsen et al., 1989; Brody, 1999a, 1999b; Pauluis & Baker, 2000; Grün et al., 2003).

5.3 Models of $N \geq 2$ Spike Trains. The models presented in this letter are direct and intuitive implementations of the discussed coding schemes. As this study uses the cross-correlation function to analyze the separability of the concepts, the models were formulated to generate only pairs of simultaneous processes. However, extensions of both models to generate populations of $N \geq 2$ processes can be formulated straightforwardly.

Our implementation of spike coordination (insertion of coincident spikes) is a special case of a compound Poisson processes (e.g., Snyder & Miller, 1991; Daley, & Vere-Jones, 2005, and references therein) and is easily extendable to generate $N \geq 2$ parallel processes. While a single coincidence process suffices to parameterize the correlation structure in a pair of processes, the parameter space increases exponentially with increasing number of processes. This is due to the fact that a single coincidence process is required to describe the joint firing of every subgroup $A \subset \{1, \ldots, N\}$ of the population. The whole population is thus a superposition of all the $2^N - 1$ processes that determine the joint firing of the subgroups. This approach can be traced back to Holgate (1964) and is detailed elsewhere (Kuhn et al., 2003; Staude, Rotter, & Grün, 2005, 2007; Ehm, Staude, & Rotter, 2007).

A straightforward extension of our implementation of rate covariation assumes that all N processes share the same rate profile (Staude et al., 2005). This generates a homogeneous population, where all processes have identical mean rate and all pairs have the same correlation function. For the inhomogeneous population for $N = 2$ in Figure 1B, the marginal rate

distributions f_{Λ_1} and f_{Λ_2} were coupled to a bivariate distribution f_{Λ_1,Λ_2} by means of a gaussian copula. In principle, the copula approach generalizes to the case $N > 2$, where a copula can be used to couple the marginal rate distributions f_{Λ_i} $(i = 1, \ldots N)$ to form a multivariate distribution $f_\Lambda = f_{\Lambda_1,\ldots,\Lambda_N}$. In our case, we wish to prescribe the mean vector $\lambda = (\mathrm{E}[\Lambda_1], \ldots, \mathrm{E}[\Lambda_N])$ and the covariances $\gamma_{ij} = \mathrm{Cov}[\Lambda_i, \Lambda_j]$ of the resulting multivariate distribution. The task would be to find the copula C that couples the marginal distributions such that the resulting multivariate distribution function realizes the above constraints. However, how feasible this approach is in practice, given nongaussian marginal rate distributions, remains to be investigated.

Most likely, cortical activity in vivo can be described faithfully only by mixtures of the above model classes. In order to calibrate correlation analysis tools for large populations of neurons, convenient parameterizations of such mixtures need to be addressed in future work.

5.4 Conclusion. The terms *rate covariation* and *spike coordination* have produced considerable confusion in the debate on the cortical code. Our intuition is that a good deal of this confusion can be attributed to the imprecise use of the terminology. This study provided unambiguous, model-free definitions that emerged from the decomposition of the cross-correlation function, showed that the difference between the concepts of dependence is not a matter of timescales, and provided guidelines to identify the concept underlying experimental data. We hope that the clarification in terminology facilitates future discussions to focus on biologically more relevant questions, for example, how the concepts of dependence discussed here are implemented in the nervous system.

Appendix: Cross-Correlation Function of Piecewise Constant Time Series

This appendix computes the raw cross-correlation function

$$\mathrm{CCF}^{\mathrm{raw}}_{l_1,l_2}(\tau) = \mathrm{E}[l_1(t)l_2(t-\tau)]$$

for piecewise constant random time series $l_i(t)$ with trajectories

$$l_i(t) = \sum_{k=0}^{M-1} \lambda_i^k \Pi_{I_k}(t - \eta_i), \quad i = 1, 2 \tag{A.1}$$

(for a description, see the text following equation 3.1). We assume that the offset of the first process η_1 is newly drawn in each realization, while the shift in the temporal offsets $\eta := \eta_1 - \eta_2$ is fixed.

To compute the correlation function of two such signals, consider marked clock processes $\xi_i(t)$ with trajectories

$$\xi_i(t) = \sum_k \delta(t - kJ_l - \eta_i)\lambda_i^k,$$

that have nonzero weights only on a grid of period J_l. Filtering such a process with a boxcar $b(t)$ of length J_l reproduces the rate profiles of equation A.1:

$$l_i(t) = (\xi_i * b)(t).$$

By the associativity of the convolution product, the cross-correlation function of filtered point processes is the convolution of the autocorrelation function of the filter with the cross-correlation function of the point processes:

$$\mathrm{CCF}^{\mathrm{raw}}_{l_1,l_2}(\tau) = \mathrm{ACF}^{\mathrm{raw}}_b * \mathrm{CCF}^{\mathrm{raw}}_{\xi_1,\xi_2}(\tau). \tag{A.2}$$

Since the filter $b(t)$ is a boxcar, its autocorrelation function (first term of equation A.2) is a triangle of base length $2J_l$ and height 1:

$$\mathrm{ACF}^{\mathrm{raw}}_b(\tau) = f(\tau; J_l),$$

with

$$f(\tau; J) := \begin{cases} \dfrac{J - |\tau|}{J} & \text{if } |\tau| \le J \\ 0 & \text{if } |\tau| > J \end{cases}.$$

Since

$$\begin{array}{r} t - kJ_l - \eta_1 = 0 \\ t - jJ_l - \eta_2 - \tau = 0 \end{array} \iff \begin{array}{r} t - kJ_l - \eta_1 = 0 \\ \tau - (k-j)J_l - (\eta_1 - \eta_2) = 0, \end{array} \tag{A.3}$$

we can compute the time-shifted product of the clock processes as

$$\begin{aligned} \xi_1(t)\xi_2(t-\tau) &= \sum_{k,j} \lambda_1^k \lambda_2^j \delta(t - kJ_l - \eta_1)\delta(t - jJ_l - \eta_2 - \tau) \\ &\overset{(A.3)}{=} \sum_{k,j} \lambda_1^k \lambda_2^j \delta(t - kJ_l - \eta_1)\delta(\tau - (k-j)J_l - \eta) \\ &= \sum_k \lambda_1^k \lambda_2^k \delta(t - kJ_l - \eta_1)\delta(\tau - \eta) \\ &\quad + \sum_{k \neq j} \lambda_1^k \lambda_2^i \delta(t - kJ_l - \eta_1)\delta(\tau - (k-j)J_l - \eta). \end{aligned}$$

Now $\mathrm{E}[\xi_i(t)] = \frac{1}{J_l}\int_0^{J_l} \mathrm{E}^{\Lambda_i}[\xi_i(t)]d\eta_1$, and since $\frac{1}{J_l}\int_0^{J_l}\sum_k \delta(t - kJ_l - \eta_1)d\eta_1 = 1$ and λ_1^k is independent from λ_2^j for $k \neq j$, we obtain

$$\begin{aligned}\mathrm{CCF}^{\mathrm{raw}}_{\xi_1,\xi_2}(\tau) &= \frac{1}{J_l}\int_0^{J_l} \mathrm{E}^{\Lambda}[\xi_1(t)\xi_2(t-\tau)]\,d\eta_1 \\ &= \mathrm{E}[\Lambda_1\Lambda_2]\delta(\tau-\eta) + \sum_{k\neq j}\mathrm{E}[\Lambda_1]\mathrm{E}[\Lambda_2]\delta(\tau-(k-j)J_l-\eta) \\ &= \mathrm{Cov}[\Lambda_1,\Lambda_2]\delta(\tau-\eta) + \mathrm{E}[\Lambda_1]\mathrm{E}[\Lambda_2]\sum_i \delta(\tau - iJ_l - \eta).\end{aligned}$$

Hence,

$$\begin{aligned}\mathrm{CCF}^{\mathrm{raw}}_{l_1,l_2}(\tau) &= \mathrm{ACF}^{\mathrm{raw}}_b * \mathrm{CCF}^{\mathrm{raw}}_{\xi_1,\xi_2}(\tau) \\ &= \mathrm{Cov}[\Lambda_1,\Lambda_2]f(\tau-\eta;J_l) + \mathrm{E}[\Lambda_1]\mathrm{E}[\Lambda_2]\sum_i f(\tau - iJ_l - \eta;J_l) \\ &= \gamma_{12}f(\tau-\eta;J_l) + \lambda_1\lambda_2, \qquad \text{(A.4)}\end{aligned}$$

with $\gamma_{12} := \mathrm{Cov}[\Lambda_1,\Lambda_2]$ and $\lambda_i := \mathrm{E}[\Lambda_i]$

Acknowledgments

This work was funded by NaFöG Berlin, the German Ministry for Education and Research (BMBF grants 01GQ01413 and 01GQ0420), and the Stifterverband für die Deutsche Wissenschaft.

References

Abeles, M. (1982). Role of cortical neuron: Integrator or coincidence detector? *Israel J. Med. Sci., 18*, 83–92.

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.

Aertsen, A. M. H. J., Gerstein, G. L., Habib, M. K., & Palm, G. (1989). Dynamics of neuronal firing correlation: Modulation of "effective connectivity." *J. Neurophysiol., 61*(5), 900–917.

Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, *1*, 371–394.

Bartlett, M. S. (1963). The spectral analysis of point processes. *Proc. R. Stat Soc B, 24*, 264–280.

Brody, C. D. (1999a). Correlations without synchrony. *Neural Comput., 11*, 1537–1551.

Brody, C. D. (1999b). Disambiguating different covariation types. *Neural Comput., 11*, 1527–1535.

Brown, E. N., Kaas, R. E., & Mitra, P. P. (2004). Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nat. Neurosci., 7*(5), 456–461.

Brunel, N., Pieczynski, W., & Derrode, S. (2005). Copulas in vectorial hidden markov chains for multicomponent image segmentation. In *Proc. IEEE Intern. Conf. Acoustic., Speech and Signal Processing 2005 (ICASSP)* (vol. 2, pp. 717–720). Piscataway, NJ: IEEE Press.

Daley, D. J., & Vere-Jones, D. (2005). *An introduction to the theory of point processes, vol.1: Elementary theory and methods* (2nd ed.). New York: Springer.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.

deCharms, R. C., & Zador, A. (2000). Neural representation and cortical code. *Annu. Rev. Neurosci., 23*, 613–647.

DeWeese, M. R., & Zador, A. M. (2006). Non-gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *J. Neurosci., 26*(47), 12206–12218.

Eggermont, J. J. (1990). *The correlative brain*. Heidelberg: Springer.

Ehm, W., Staude, B., & Rotter, S. (2007). Decomposition of neuronal assembly activity via empirical de-Poissonization. *Electronic Journal of Statistics, 1*, 473–495.

Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., & Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci., 2*, 1527–1537.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science, 233*, 1416–1419.

Gerstein, G. L., Bedenbaugh, P., & Aertsen, A. (1989). Neuronal assemblies. *IEEE Trans. Biomed. Eng., 36*, 4–14.

Grün, S., Diesmann, M., & Aertsen, A. (2002a). "Unitary events" in multiple single-neuron spiking activity. I. Detection and significance. *Neural Comput., 14*(1), 43–80.

Grün, S., Diesmann, M., & Aertsen, A. (2002b). "Unitary events" in multiple single-neuron spiking activity. II. Non-stationary data. *Neural Comput., 14*(1), 81–119.

Grün, S., Diesmann, M., Grammont, F., Riehle, A., & Aertsen, A. (1999). Detecting unitary events without discretization of time. *J. Neurosci. Methods, 94*(1), 67–79.

Grün, S., Riehle, A., & Diesmann, M. (2003). Effect of cross-trial nonstationarity on joint-spike events. *Biol. Cybern., 88*(5), 335–351.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Holgate, P. (1964). Estimation for the bivariate poisson distribution. *Biometrika, 51*, 241–245.

Jenison, R. J., & Rearle, R. A. (2004). The shape of neuronal dependence. *Neural Comput., 16*, 665–672.

Johnson, D. H. (1996). Point process models of single-neuron discharges. *J. Comput. Neurosci., 3*, 275–299.

Kass, R., Ventura, V., & Brown, E. (2005). Statistical issues in the analysis of neuronal data. *J. Neurophysiol., 1*(94), 8–25.

Kohn, A., & Smith, M. A. (2005). Stimulus dependence of neuronal correlations in primary visual cortex of the macaque. *J. Neurosci., 25*(14), 3661–3673.

Kuhn, A., Aertsen, A., & Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput., 1*(15), 67–101.

Lestienne, R. (2001). Spike timing, synchronization and information processing on the sensory side of the central nervous system. *Prog. Neurobiol., 65*, 545–591.

Masuda, N., & Aihara, K. (2003). Duality of rate coding and temporal coding in multilayered feedforward networks. *Neural Comput., 15*, 103–125.

Musial, P., Baker, S., Gerstein, G., King, E., & Keating, J. (2002). Signal-to-noise ratio improvement in multiple electrode recording. *J. Neurosci. Methods, 1*(115), 29–43.

Nawrot, M. P., Aertsen, A., & Rotter, S. (1999). Single-trial estimation of neuronal firing rates: From single-neuron spike trains to population activity. *J. Neurosci. Methods, 94*(1), 81–92.

Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., & Rotter, S. (2007). Measurement of variability dynamics in cortical spike trains. *J. Neurosci. Methods,* doi: 10.1016/j.jneumeth.2007.10.013.

Nelsen, R. B. (1999). *An introduction to copulas*. Berlin: Springer.

Nelson, J. I., Salin, P. A., Munk, M. H. J., Arzi, M., & Bullier, J. (1992). Spatial and temporal coherence in cortico-cortical connections: A cross-correlation study in areas 17 and 18 in the cat. *VisNeurosci., 9*, 21–37.

Nicolelis, M., & Riberio, S. (2006). Seeking the neural code. *Sci. Am., 295*(6), 70–77.

Nowak, L. G., Munk, M. H., Nelson, J. I., James, A., & Bullier, J. (1995). Structural basis of cortical synchronization. I. Three types of interhemispheric coupling. *J. Neurophysiol., 74*(6), 2379–2400.

Oram, M. W., Wiener, M. C., Lestienne, R., & Richmond, B. J. (1999). Stochastic nature of precisely timed spike patterns in visual system neuronal responses. *J. Neurophysiol., 81*, 3021–3033.

Papoulis, A. (1991). *Probability, random variables, and stochastic processes* (3rd ed.). New York: McGraw-Hill.

Pauluis, Q., & Baker, S. N. (2000). An accurate measure of the instantaneous discharge probability, with application to unitary joint-event analysis. *Neural Comput., 12*(3), 647–669.

Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967a). Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophys. J., 7*(4), 391–418.

Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967b). Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J., 7*(4), 419–440.

Riehle, A., Grün, S., Diesmann, M., & Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science, 278*, 1950–1953.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press.

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *NatNeurosci., 7*(9), 982–991.

Shadlen, M. N., & Movshon, A. J. (1999). Synchrony unbound: A critical evaluation of the temporal binding hypothesis. *Neuron, 24*, 67–77.

Shimazaki, H., & Shinomoto, S. (2007). A method for selecting the bin size of a time histogram. *Neural Comput., 19*(6), 1503–1527.

Singer, W. (1999). Neural synchrony: A versatile code for the definition of relations. *Neuron., 24*, 49–65.

Singer, W., Engel, A. K., Kreiter, A. K., Munk, M. H. J., Neuenschwander, S., & Roelfsema, P. R. (1997). Neuronal assemblies: Necessity, signature and detectability. *Trends. Cog. Sci., 1*(7), 252–261.

Snyder, D. L., & Miller, M. I. (1991). *Random point processes in time and space*. Berlin: Springer.

Staude, B., Rotter, S., & Grün, S. (2005). Correlated point processes and their statistical properties. In *XXV. Dynamic Days Europe, Europhysics Conference Series*, Volume 29E. Program No. P. 2.120.

Staude, B., Rotter, S., & Grün, S. (2007). Detecting the existence of higher-order correlations in multiple single-unit spike trains. In *Neuroscience Meeting Planner, San Diego, CA: Society for Neuroscience, 2007.* Program No. 103.9.

Vaadia, E., Aertsen, A., & Nelken, I. (1995). "Dynamics of neuronal interactions" cannot be explained by "neuronal transients." *Proc. Biol. Sci., 261*(1362), 407–410.

von der Malsburg, C. (1981). *The correlation theory of brain function* (Internal rep. 81-2). Göttingen: Max-Planck-Institute for Biophysical Chemistry.

Received June 15, 2007; accepted December 9, 2007.