



Statistische Verfahren zur Analyse hochdimensionaler neuronaler Messreihen in Bezug auf neurokognitive Vorgänge

Daniel Durstewitz und Emili Balaguer-Ballester

Zusammenfassung

Die jüngsten Fortschritte bei multiplen Einzelzellmessungen und Verfahren der optischen Bildgebung erlauben uns inzwischen, die Aktivität von bis zu Hunderten von Neuronen gleichzeitig zu registrieren. Das Ergebnis solcher Messungen sind hochdimensionale, multivariate Zeitreihen, die noch nie da gewesene Möglichkeiten bieten, Einblicke in Details raum-zeitlicher neuronaler Dynamiken zu erhalten, die kognitiven Prozessen zugrunde liegen. Beispielsweise können sie den Weg bahnen für reliable Einzel-Trial-Analysen, für die Erforschung der Rolle Korrelationen höherer Ordnung bei der neuronalen Codierung, der Mechanismen, die der neuronalen Ensemble-Formierung zugrunde liegen, oder allgemeiner für Übergänge zwischen Attraktorzuständen, die mit kognitiven Prozessen einhergehen. Um dieses Informationspotenzial multivariater neuronaler Zeitreihen auszuschöpfen, werden oftmals fortgeschrittene statistische Methoden benötigt, die über das üblicherweise verwendete Repertoire hinausgehen. In diesem Artikel diskutieren wir anhand spezifischer experimenteller Beispiele einige dieser Methoden zur Visualisierung von Struktur in hochdimensionalen Datensätzen, zur statistischen Inferenz über solche Strukturen, zur Einzel-Trial-Analyse von neuronalen Zeitreihen und zur Rekonstruktion dynamischer Eigenschaften neuronaler Systeme, die sich nur durch simultane Ableitung vieler Zellen gewinnen lassen.

Abstract

Statistical Approaches for Reconstructing Neuro-Cognitive Dynamics from High-Dimensional Neural Recordings.

Recent advances in multiple single-unit recording and optical imaging techniques now routinely allow to observe the activity from tens to hundreds of neurons simultaneously. The result are high-dimensional multivariate time series which offer an unprecedented range of possibilities for gaining insight into the detailed spatio-temporal neural dynamics underlying cognition. For instance, they may pave the way for reliable single-trial analyses, for investigating the role of higher-order correlations in neural coding, the mechanisms of neural ensemble formation, or more generally of transitions among attractor states accompanying cognitive processes. At the same time, exploiting the information in these multivariate time series may require more sophisticated statistical methods beyond the commonly employed repertoire. Here we review, along some specific experimental examples, some of these methods for visualizing structure in high-dimensional data sets, for statistical inference about the apparent structure, for single-trial analysis of neural time series, and for reconstructing some of the dynamical properties of neural systems that can only be inferred from simultaneous recordings.

Keywords: neural dynamics; statistics; machine learning; prefrontal cortex; multiple single-unit recordings

Einleitung

Die Lebenswissenschaften im Allgemeinen, und die Neurowissenschaften im Besonderen, produzieren Datensätze von stetig zunehmendem Umfang und zunehmender

Komplexität. Simultane Ableitungen von bis zu Hunderten von Neuronen durch optische Bildgebungsverfahren oder mehrere Tetroden werden inzwischen zur Routine, und nicht-invasive bildgebende Verfahren wie die funktionelle Magnetresonanztomografie

(fMRI) erfassen gar die Aktivität von bis zu Tausenden von Einheiten als Funktion der Zeit. Statistisch gesehen sind all diese Datensätze Beispiele für multivariate Zeitreihen (Abbildung 1), in denen ein Vektor (oder eine zusätzlich topologische Beziehungen widerspiegelnde Matrix) von Aktivitätswerten als Funktion der Zeit betrachtet wird. Diese Techniken, wie etwa fMRI und Elektroenzephalografie (EEG), werden immer häufiger auch miteinander kombiniert, was zu Datensätzen führt, die nicht nur multivariat sondern auch multimodal sind. Üblicherweise werden diese multivariaten neuronalen Ableitungen auch nicht isoliert betrachtet, sondern mit einer Fülle von simultan aufgenommenen behavioralen, genetischen, molekularen oder anderen Informationen verglichen.

Diese häufig enorm großen und komplexen Datensätze stellen eine schwierige Herausforderung für Verfahren der Datenanalyse und statistischen Inferenz dar, womit hier Verfahren gemeint sind, mit deren Hilfe sich relevante Informationen aus den multivariaten Zeitreihen extrahieren und mit Wahrscheinlichkeitsangaben versehen lassen (d.h. Rückschlüsse auf die einer Stichprobe zugrunde liegende Population ziehen lassen). Natürlich lassen sich diese multivariaten Zeitreihen auch weiterhin mit traditionelleren Methoden analysieren, z.B. die Betrachtung stimulusabhängiger Feuerratenänderungen oder paarweiser Kreuzkorrelationen zwischen Neuronen. Aus dieser Sicht bedeuten die Fortschritte in den Bereichen der Multizellableitungen und Bildgebung vor allem, dass sehr viel mehr Daten in kürzerer Zeit gewonnen werden können, die dann in gleicher Weise wie z.B. traditionelle Einzelzellableitungen analysiert werden. Die simultane Erfassung der Aktivität von Hunderten neuronaler Elemente bietet jedoch völlig neue und spannende Möglichkeiten, die für ihre volle Erschließung fortgeschrittene mathematische Techniken benötigen, von denen einige bereits seit Jahrzehnten zum Repertoire der multivariaten Statistik gehören. Wir werden hier einige dieser Methoden zusammenfassen mit Fokus auf multiple Einzelzellableitungen von sich verhaltenden Tieren. Die diskutierten Methoden sind jedoch sehr viel universeller anwendbar.

Visualisierung multivariater Datensätze: Reduktion der Dimensionalität und neuronale Zustandsräume

Das Ziel vieler multivariater Datenanalyseverfahren kann als die Aufdeckung raumzeitlicher Muster und Strukturen gesehen werden, die dann in dem durch das Experiment vorgegebenen kognitiven oder verhal-

tensbezogenen Kontext interpretiert werden. Ein erster Schritt kann die Visualisierung der multivariaten Aktivität sein, um ein Gefühl dafür zu entwickeln, wie sich die neuronale Dynamik mit der Zeit entwickelt (Abbildung 1) oder wie sie von verschiedenen Aufgaben oder Verhaltensbedingungen abhängt (Abbildung 3). Eine übliche Vorgehensweise bei der Visualisierung ist die Reduktion der ursprünglich hohen Datendimensionalität auf nur zwei oder drei Dimensionen, indem Redundanzen in den Daten ausgenutzt werden (Abbildung 2). Andere Methoden versuchen die Daten in Form von Prototypen oder Klassen zusammenzufassen, z.B. verschiedene clusteranalytische Verfahren. Die statistische Literatur ist reich an diesbezüglichen mathematischen Techniken (z.B. Krzanowski 2000, Hastie et al. 2009), von denen einige bereits seit mehr als hundert Jahren existieren. Abbildung 2 veranschaulicht die Funktionsprinzipien einiger der am häufigsten eingesetzten Techniken. Die wahrscheinlich am häufigsten verwendete Methode ist die Hauptkomponentenanalyse [Principal Component Analysis (PCA), Abbildung 2A], ein lineares Verfahren, mit dem die Daten in einen niederdimensionalen Raum projiziert werden, der am meisten von der ursprünglichen Variabilität erhält. Genauer gesagt führt die PCA eine Rotation der Achsen des ursprünglichen Raumes durch, und zwar so, dass sie a) in Richtung der höchsten Datenvarianz ausgerichtet werden und b) orthogonal zueinander sind, d.h. also gewissermaßen „unabhängig“ voneinander (unkorreliert nach Abzug der Mittelwerte), sodass jede neue Achse einen „neuen Aspekt“ der Daten reflektiert. Dabei hofft man, mit nur einigen wenigen Dimensionen den Hauptteil der Datenvariation erfassen zu können, sodass Dimensionen, entlang derer die Daten nur wenig streuen, ohne großen Informationsverlust verworfen werden können (Abbildung 2A). Eine (oberflächlich) verwandte Technik, die besonders in der Psychologie lange Zeit beliebt war, ist die Faktoranalyse (FA). Auf den ersten Blick scheint die FA der PCA sehr ähnlich und wird daher häufig mit dieser verwechselt, denn ähnlich wie die PCA versucht die FA die Daten durch das Ausnutzen von Redundanzen als Linearkombination weniger unkorrelierter Faktoren zu beschreiben. Es gibt jedoch einen äußerst wichtigen Unterschied: Während die PCA versucht die Dimensionen in Richtung maximaler Varianz auszurichten, basiert die FA auf einem statistischen Modell mit explizitem Fehlerterm (die Daten werden als Resultat unkorrelierter latenter Faktoren und dieses Fehlerterms angesehen) und versucht Dimensionen zu finden, die *Korrelationen* zwischen den ursprünglichen Variablen

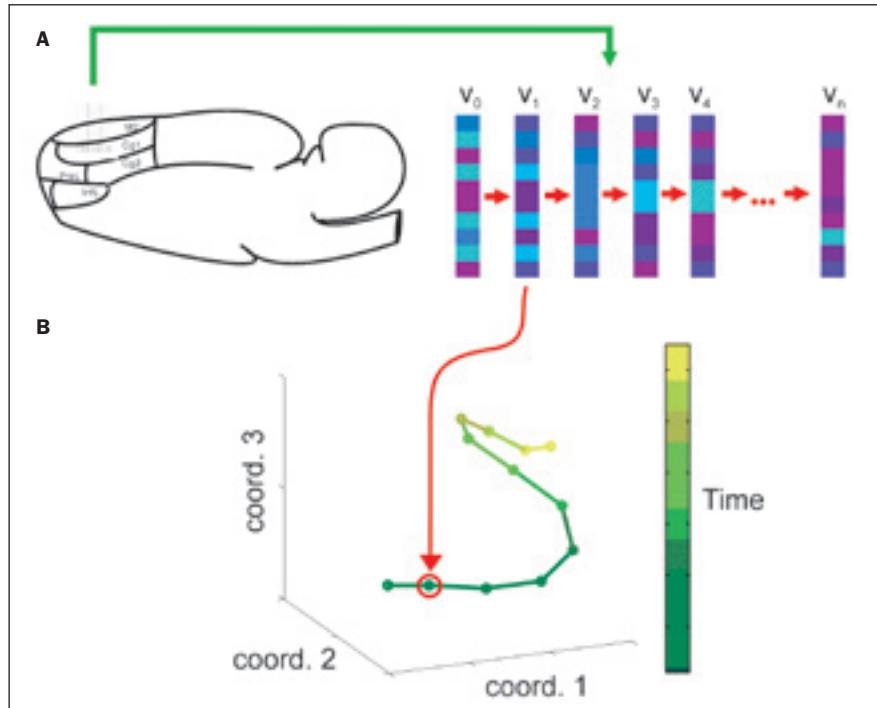


Abb. 1: Illustration von aus Multielektrodenableitungen gewonnenen multivariaten (Vektor-) Zeitreihen und deren Projektion in einen 3-dimensionalen Raum (unterer Teil) mit dem Ziel der Datenreduktion und Visualisierung (siehe auch Abb. 2). Teil A, links, nachgedruckt mit freundlicher Genehmigung aus Lapish et al. (2008), *Proceedings of the National Academy of Sciences USA* (Copyright: National Academy of Sciences of the USA).

möglichst gut erfassen. Daher können die durch PCA und FA gefundenen Lösungen sehr unterschiedlich sein (s. Fig. 4 in Yu et al. 2009). Sowohl PCA als auch FA wurden bereits für die Visualisierung multipler Einzelzellableitungen eingesetzt (z.B. Mazor und Laurent 2005; Yu et al. 2009).

Eine grundsätzlich andere Technik der Dimensionalitätsreduktion verbirgt sich hinter einer Klasse von Verfahren, die unter dem Begriff multidimensionale Skalierung (MDS) bekannt sind. Das Ziel bei der MDS ist, die Daten so in einen viel niedrigerdimensionalen Raum zu projizieren, dass die *Abstände* oder *Unähnlichkeiten* zwischen den ursprünglichen Datenpunkten bestmöglich erhalten bleiben (Abbildung 2B; bzw. im Fall nicht-metrischer MDS ihre ordinalen Relationen). Dies funktioniert im Allgemeinen gut, wenn die Daten im Wesentlichen in einem näherungsweise linearen Unterraum liegen oder bei Erweiterungen wie Isomap (siehe unten) auf eine niederdimensionale Mannigfaltigkeit beschränkt bleiben. In Hinblick auf die Visualisierung neuronaler Dynamiken haben diese Techniken eine Reihe deutlicher Vorteile gegenüber der PCA: Distanzen zwischen Punkten im PCA-Raum können stark von den Distanzen im ursprünglichen hochdimensionalen Raum abweichen, da bei

der Projektion auf die varianzmaximierenden Dimensionen Vektorkomponenten komplett ignoriert werden, die rechtwinklig zu diesen Dimensionen stehen. Daher ist bei der PCA nicht klar, wie die geometrischen Strukturen im ursprünglichen Raum mit denen im reduzierten PCA-Raum zusammenhängen, während MDS die ursprüngliche Geometrie der Datenwolke bestmöglich zu erhalten sucht. Außerdem kann die PCA zu ungünstigen Visualisierungen führen, wenn die Datenvarianz sehr ungleichmäßig entlang der Achsen verteilt ist, während bei der MDS jede neue Dimension gewissermaßen gleichberechtigt ist. Diese Vorteile werden allerdings damit erkaufte, dass MDS als iteratives Optimierungsverfahren einen potenziell hohen Speicher- und Rechenzeitbedarf hat und die Gefahr besteht, dass nur suboptimale (lokal optimale) Lösungen gefunden werden. Im Gegensatz dazu kann bei der PCA der neue Satz von Achsen schnell und explizit durch die Lösung eines einfachen Eigenwertproblems berechnet werden. (Eine Variante der metrischen MDS, die sogenannte „klassische MDS“, liefert jedoch zur PCA äquivalente Lösungen im Falle euklidischer Distanzen). Wie oben angedeutet, gibt es auch „nicht-metrische“ Varianten der MDS, bei denen niedrigdimensionale Räume allein aufgrund

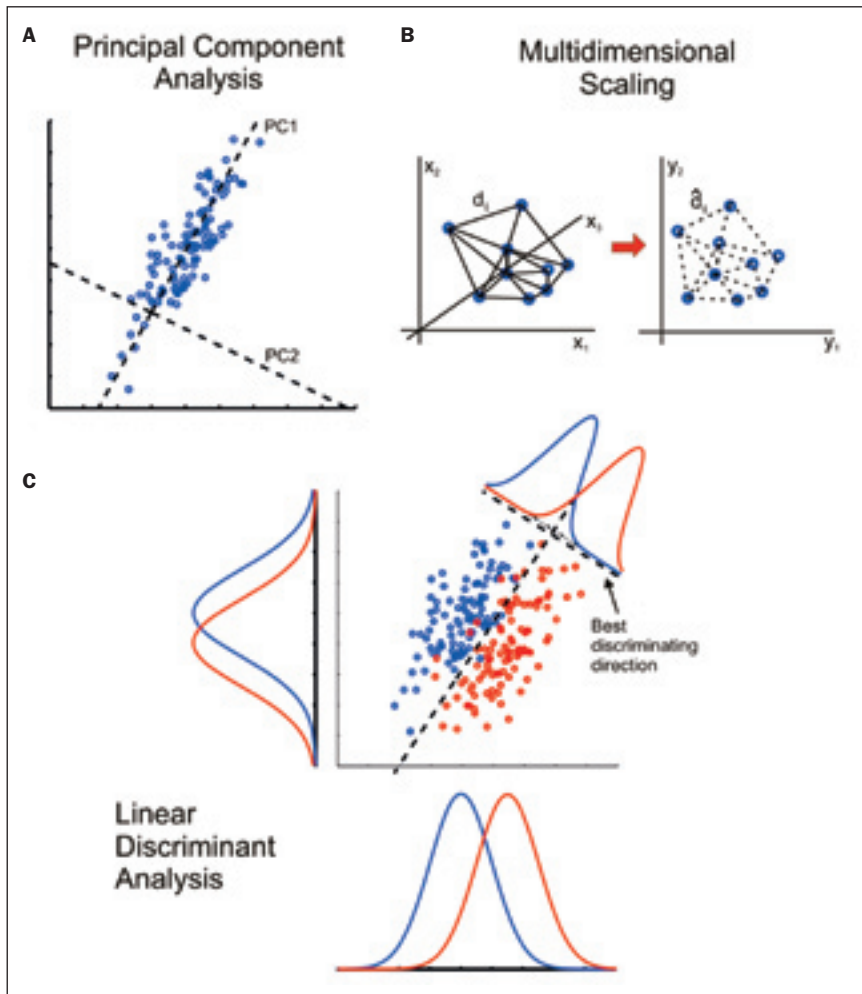


Abb. 2: Methoden zur Dimensionalitätsreduktion. A) Die Hauptkomponentenanalyse (principal component analysis, PCA) rotiert die Achsen des ursprünglichen Raumes so, dass sie in Richtung der größten Datenvarianz ausgerichtet und orthogonal zueinander sind. Wie illustriert, wird durch die erste Hauptkomponente (PC) der größte Teil der Varianz erfasst, durch die zweite PC der zweitgrößte Anteil der Varianz, senkrecht zur ersten PC streuend, und so weiter. Dabei hofft man, dass einige wenige Dimensionen ausreichen, um den größten Teil der Datenvarianz abzubilden. Zum Beispiel könnte PC2 im dargestellten Beispiel entfallen, da die Datenpunkte im Wesentlichen entlang der PC1 streuen. B) Multidimensionale Skalierung (MDS) versucht eine niedrigdimensionalere Einbettung der Daten in der Form zu finden, dass Distanzen im ursprünglichen Raum (gekennzeichnet durch verbundene Linien) in der Einbettung so weit wie möglich erhalten bleiben. C) Fishers Diskriminanzanalyse findet einen Satz neuer Achsen (auch kanonische Variablen genannt), entlang derer die Unterschiede zwischen vordefinierten Gruppen am stärksten hervortreten. Für g Gruppen ergeben sich maximal $g-1$ solcher Achsen, die nicht notwendigerweise orthogonal zu einander stehen müssen. Die Punkte in diesem Beispiel wurden aus zweidimensionalen Gaussverteilungen gezogen. Die abgebildeten Dichtefunktionen zeigen, dass die zwei Klassen von Punkten entlang der ursprünglichen x - und y -Achsen stark überlappen, aber entlang der durch die FDA bestimmten Richtung recht gut separiert sind.

von Ranginformationen zwischen Objektpaaren rekonstruiert werden. Das auf der MDS basierende kürzlich entwickelte Isomap-Verfahren (Tenenbaum et al. 2000) versucht, eine niedrigdimensionale Mannigfaltigkeit, auf der die Daten mutmaßlich liegen, zu rekonstruieren, indem die Distanzen zwischen

Punkten als die auf dieser Mannigfaltigkeit definiert werden (geodätische oder kürzester Pfad-Distanz). Ein ähnliches Ziel verfolgt die Methode des lokalen Einbettens (Locally-Linear-Embedding, LLE; Roweis und Saul 2000), eine andere kürzlich vorgeschlagene Technik zur Dimensionalitätsreduktion.

So wie PCA und FA sind auch MDS (Lapish et al. 2008), Isomap (Compte et al. 2003) und LLE (Broome und Laurent 2006) bereits verschiedentlich zur Visualisierung neuronaler Aktivität verwendet worden. Abbildung 3 zeigt ein Beispiel aus unserer eigenen Arbeit, bei der gleichzeitige Ableitungen von 10 bis 40 Neuronen aus dem anterioren Cingulate-Cortex (ACC) von Ratten durch metrische MDS in einen dreidimensionalen Raum eingebettet wurden, sodass jeder Punkt in diesem Raum die simultane Aktivität der abgeleiteten Population zu einem gegebenen Zeitpunkt repräsentiert. Die Daten stammen aus Ableitungen während der in Abbildung 3 dargestellten Arbeitsgedächtnis- und Entscheidungsaufgabe: Während einer Trainingsphase hatten die Ratten Zugang zu vier von acht mit Futterbelohnungen versehenen Armen in einem Radiallabyrinth, während die vier anderen (zufällig ausgewählten) Arme physikalisch blockiert wurden. Nach Futteraufnahme in allen vier zugänglichen Armen wurde die Ratte in dem zuletzt besuchten Arm für eine Minute oder länger eingeschlossen (die Delay-Phase). Nach dieser Zeit wurden alle Barrieren entfernt, sodass die Ratte nun in der Testphase auch zu den verbliebenen vier noch mit Futter versehenen Armen Zugang hatte. Um die Aufgabe effizient zu lösen, muss das Tier also während der Delay-Phase Informationen über die Labyrintharme, die schon besucht wurden (oder noch zu besuchen sind), aufrechterhalten und diese Informationen in der Testphase verwerten, um bei jedem Armeingang die richtige Wahl zu treffen. Optimale Performanz in dieser Aufgabe hängt stark von verschiedenen Arealen des Frontalkortex (einschl. des ACC) ab. In Bezug auf ihre kognitiven Anforderungen kann diese Aufgabe in verschiedene Perioden unterteilt werden: Entscheidungsphasen (Abbildung 3A), in denen die Tiere ihre Entscheidungen über einen Armeintritt fällen, Belohnungsphasen, in denen die Tiere Zugriff auf das Futter haben, generell die Trainings- und die Testphase, in der sich die Tiere jeweils von einem Arm zum nächsten bewegen, sowie die Delay-Phase. Da die Gedächtnisanforderungen während der Trainings- und Testperiode unterschiedlich sind, kann man weiterhin vermuten, dass sich die neuronale Dynamik bei Entscheidungen und Belohnungen während dieser beiden Perioden unterschiedlich entwickelt. Um den Zusammenhang zwischen der neuronalen Populationsaktivität im MDS-Raum und den verschiedenen kognitiv definierten Aufgabenphasen zu illustrieren, wurden die Punkte im MDS-Raum entsprechend der kognitiven Phasen farblich markiert (Abbildung 3B). Wie die Abbildung zeigt,

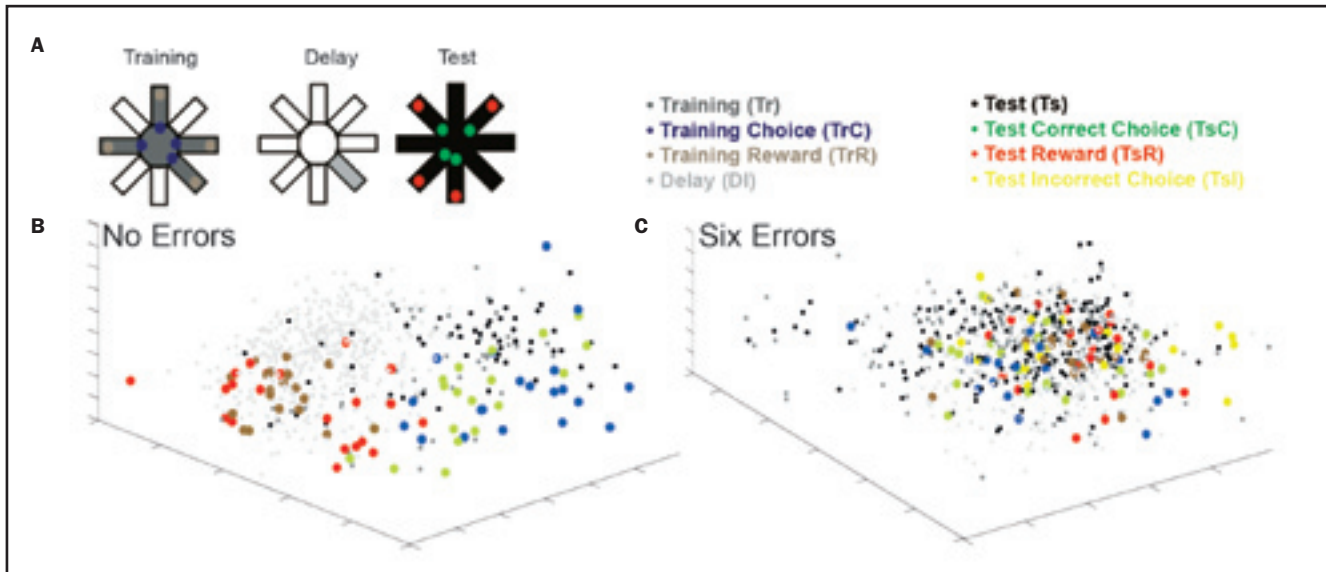


Abb. 3: Populationscodierung verschiedener kognitiv definierter Aufgabenphasen in einer Arbeitsgedächtnis- und Entscheidungsaufgabe. A) Illustration der Labyrinthaufgabe mit ihren verschiedenen, entsprechend der Legende farblich markierten Aufgabenphasen (definiert aufgrund ihrer unterschiedlichen kognitiven Anforderungen). B) und C) Darstellung der via MDS (vgl. Abb. 2B) aus multiplen Einzelzellableitungen gewonnen neuronalen Zustandsräume. Links ist der Zustandsraum für ein Tier während der fehlerfreien Ausführung der Aufgabe abgebildet, während rechts der Zustandsraum für das gleiche Tier für einen Trial rekonstruiert wurde, in dem es insgesamt sechsmal fälschlicherweise in einen bereits vorher besuchten Arm eintritt. Jeder farbcodierte Punkt stellt den Zustand des abgeleiteten Netzwerkes (d.h. des Populationsvektors der instantanen Feuerraten) in einem anderen 200 ms-Zeitfenster während der Aufgabe dar. Die Punkte wurden entsprechend der in A definierten Aufgabenphasen farblich codiert. Die verschiedenen Aufgabenphasen entsprechenden Populationsmuster lassen sich relativ gut trennen während der fehlerfreien (links) aber nicht der stark fehlerbehafteten (rechts) Durchführung der Aufgabe. Nachdruck mit freundlicher Genehmigung aus Lapish et al. (2008), *Proceedings of the National Academy of Sciences USA* (Copyright: National Academy of Sciences of the USA).

scheinen die verschiedenen oben definierten Aufgabenphasen im MDS-Raum zu separieren, d.h. verschiedene Phasen gehen mit verschiedenen Mustern von Populationsaktivität einher und bilden daher separate Cluster im MDS-Raum. Diese Organisation der Populationsaktivität im MDS-Raum bricht jedoch zusammen, wenn die Tiere viele Fehler begehen, d.h. bereits erkundete Arme erneut besuchen (Abbildung 3C). Dies legt nahe, dass die mit den Aufgabenphasen korrelierte Separierung der Populationsaktivität im MDS-Raum funktionell bedeutsam für die Durchführung dieser Aufgabe ist.

Abbildung 2C veranschaulicht noch ein weiteres Verfahren für die niedrigdimensionale Darstellung neuronaler Populationsaktivität, das verwendet werden kann, wenn der Fokus auf der Visualisierung von Unterschieden zwischen Populationsmustern liegt, die mit wohldefinierten Ereignissen einer Aufgabe assoziiert sind. Fishers Diskriminanzanalyse (FDA) ist der PCA vom Grundgedanken her ähnlich, jedoch mit dem Unterschied, dass hier die Achsen so rotiert werden, dass die mittleren Unterschiede zwischen Punktemengen, die zu unterschiedlichen Klassen gehören, maximiert werden, während gleichzeitig die Variabilität innerhalb der Klassen minimiert wird

(Abbildung 2C, siehe auch Abbildung 5B). Mit anderen Worten, sowohl FDA als auch PCA sind lineare Transformationen des ursprünglichen Datenraumes, jedoch versucht FDA im ursprünglichen Raum Richtungen zu finden, entlang derer zwei oder mehr a-priori-definierte Klassen optimal separiert werden, anstatt die Varianz entlang dieser Richtungen zu maximieren. In Abbildung 5B wurde diese Näherung verwendet, um Unterschiede in der neuronalen Aktivität zu visualisieren, die mit zwei verschiedenen Verhaltensregeln und zwei Stimulusbedingungen einhergehen. Diese Daten wurden während eines Set-Shifting-Paradigmas aufgenommen, das weiter unten im Detail erklärt wird.

Statistische Inferenz bezüglich der Struktur in neuronalen Zustandsräumen

Die oben beschriebenen Methoden sind hauptsächlich Werkzeuge zur Visualisierung und Komplexitätsreduktion, obwohl sie auch im Kontext genuiner Probleme der quantitativen statistischen Analyse zur Anwendung kommen können. Eine große Frage in Bezug auf die visualisierten Muster und Dynamiken ist, ob die z.B. im MDS-Raum beobachteten Eigenschaften, wie etwa die Separierung neuronaler Aktivität in Abhängigkeit von

der Aufgabenphase, tatsächlich signifikante Eigenschaften der Populationsdynamik darstellen oder nur zufälliger Natur sind. Die „klassische“ Statistik bietet sowohl exakte als auch asymptotische Tests an, wobei sich letztere zumeist auf den zentralen Grenzwertsatz stützen, der besagt, dass die Verteilung einer Summe von Zufallsvariablen gegen eine Normalverteilung konvergiert, wenn die Anzahl der Beobachtungen (der Stichprobenumfang) gegen Unendlich geht. Viele der üblichen statistischen Tests gehen von der Annahme aus, dass die Beobachtungen unter der Nullhypothese unabhängig voneinander und identisch verteilt sind (d.h. derselben Verteilung entstammen). In allen oben genannten Beispielen handelt es sich aber um *Zeitreihen*, die von biologischen bzw. biophysikalischen Systemen erzeugt wurden. Solche Zeitreihen sind naturgemäß autokorreliert, d.h. zeitlich unmittelbar aufeinander folgende Datenpunkte haben eine größere Ähnlichkeit zueinander als Messpunkte mit einem großen zeitlichen Abstand, womit die Annahme der statistischen Unabhängigkeit verletzt wäre. Unter bestimmten Bedingungen ist es auch in solchen Fällen möglich, asymptotische Tests zu konstruieren. Wenn allerdings über die Art und Verteilungseigenschaften der



der Zeitreihe zugrundeliegenden Dynamik nur wenig bekannt ist, bieten Methoden des parametrischen und nicht-parametrischen Bootstrapping eine mächtige Alternative. Diese Methoden sind im Allgemeinen recht einfach zu implementieren und universell einsetzbar, allerdings z.T. auch ausgesprochen rechenintensiv.

Abbildung 4 illustriert das Bootstrapping-Verfahren anhand der Separierung von z.B. Aufgabenphasen wie in Abb. 3 gezeigt. Um die durch die MDS-Darstellung suggerierte Separierung von Aufgabenphasen auf Signifikanz zu testen, wurde zunächst ein Klassifikationsfehler wie folgt definiert: Für jedes Paar von Ereignissen oder Phasen während der Aufgabe wurde eine Hyperebene so in den ursprünglichen (hochdimensionalen) Raum der multiplen Ableitungen gelegt, dass die beiden Phasen optimal separiert wurden (Abbildung 4A). Optimalität wurde in diesem Fall anhand des oben eingeführten Fisherschen linearen Diskriminanzkriteriums definiert (Abbildung 4A), d.h. die Trennebene wird so im Raum orientiert, dass Unterschiede zwischen den Klassen maximiert und innerhalb der Klassen minimiert werden. Unter Normalverteilungsannahmen minimiert dieses Vorgehen den

Klassifikationsfehler (Abbildung 4A). Es gibt noch andere Möglichkeiten, Optimalität zu definieren, z.B. indem man die Hyperebene so orientiert, dass die Abstände zu den nächstliegenden Datenpunkten maximiert werden. Dies führt dann zu den maximum margin classifiers, die in den sogenannten Support Vector Machines (SVM) eingesetzt werden. Auch lassen sich natürlich nichtlineare (z.B. quadratische) Hyperflächen bestimmen, die Klassen von Datenpunkten optimal trennen. Allgemein führt dies in das weite Feld der (überwachten) Klassifikationsverfahren, von denen viele, unter anderem auch die erwähnten SVM, in jüngster Zeit eingesetzt wurden, um mentale Zustände von Probanden aus fMRI-Messungen abzulesen oder Verhaltensentscheidungen vorherzusagen (siehe z.B. Haynes und Rees 2006). Im Falle der in Abbildung 3 gezeigten Daten wurde allerdings nur ein einfacher linearer Klassifikator wie oben definiert verwendet, da es hier primär um die Definition einer geeigneten Teststatistik als Maß der Separierung zwischen Aufgabenphasen ging, und nicht per se um optimale Klassifikation oder Vorhersagen. Ein linearer Klassifikator auf Basis des Fisherschen Diskriminanzkriteriums hat zudem den Vorteil, dass er einfach zu berechnen und zu verstehen ist und keinerlei untersucherabhängiges manuelles Parametertuning erfordert.

Als Prüfstatistik wurde nun der Anteil falsch klassifizierter Punkte verwendet, d.h. die relative Häufigkeit von Punkten, die auf die falsche Seite der Trennebene fallen (Abbildung 4A). Zum Test der Nullhypothese, dass keine überzufällige Trennung der Verhaltensphasen vorliegt, kann die Verteilung dieser Statistik durch Bootstrapping aus den Daten gewonnen werden, d.h. die beobachteten Daten definieren eine empirische Verteilung, aus der neue Stichproben gezogen werden können. Im einfachsten Falle sind die Daten unabhängig und identisch verteilt, sodass man die beobachteten Datenpunkte einfach zufällig einer der beiden zu vergleichenden Aufgabenphasen zuordnen kann (ohne die Gesamtzahl der Punkte in jeder Klasse zu ändern; Abbildung 4A, rechts), wonach die Statistik (der Klassifikationsfehler) neu berechnet wird. Da wir es aber mit Zeitreihen zu tun haben, könnte eine Clusterung im neuronalen Aktivitätsraum auch einfach durch die zeitlichen Relationen zwischen den Datenpunkten bedingt sein, statt durch die Zugehörigkeit zu einer bestimmten kognitiven Kategorie. Daher sollten die originären Autokorrelationen auch in den Bootstrap-Daten bis zu einem gewissen Grade erhalten bleiben, etwa bis zu einer Korrelationslänge von der Dauer einer einzelnen Aufgabenphase. Dies kann erreicht werden, indem man statt einzelner Datenpunkte

Block Permutation Bootstraps

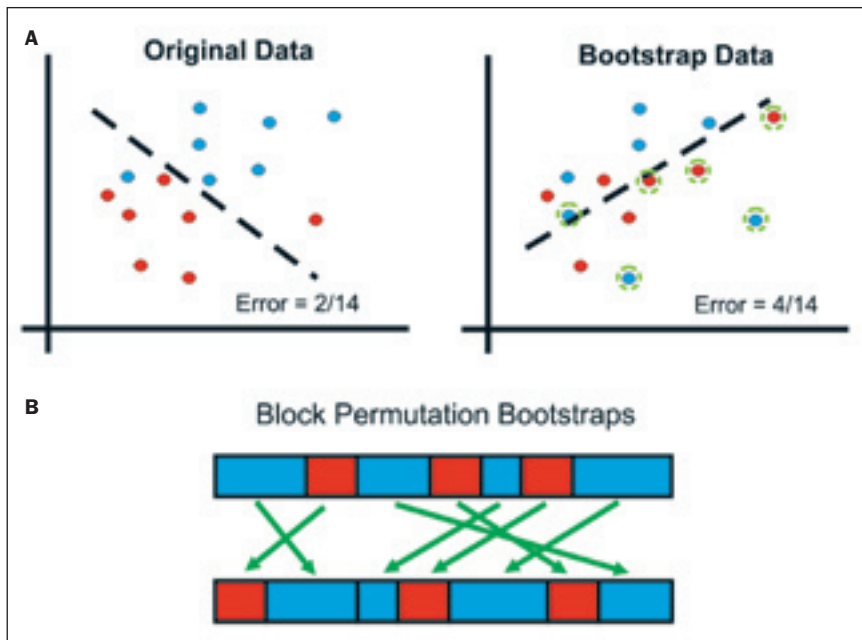
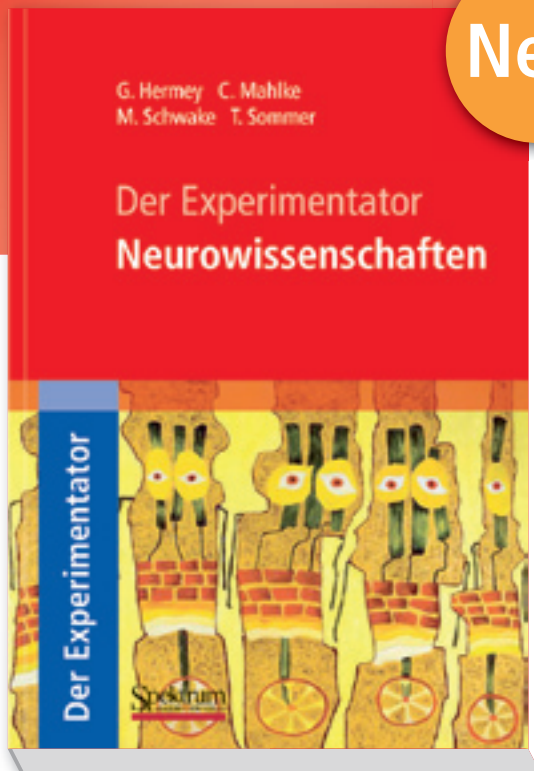


Abb. 4: Statistiken und Bootstrap-Verfahren zur Bestimmung der Signifikanz der Trennung der Populationsaktivität in Abb. 3B in verschiedene, den Aufgabenphasen entsprechende Punktwolken. A) Links: Für jedes Paar von Aufgabenphasen lässt sich ein Klassifikationsfehler (KF) als relativer Anteil der von einem linearen Klassifikator (einer optimalen Trennlinie, -ebene oder hyperebene; vgl. Abb. 2C) falsch zugeordneten Punkte berechnen. Optimalität lässt sich hier in Bezug auf verschiedene Kriterien definieren, z.B. unter Normalverteilungsannahmen. Der Anteil der Datenpunkte, der auf die falsche Seite der Trennebene fällt, kann als Maß für die Trennbarkeit der Aufgabenphasen durch die neuronale Populationsaktivität aufgefasst werden. **Rechts:** Im einfachsten Fall lassen sich Bootstrap-Daten so erzeugen, dass man die Datenpunkte einer der beiden Aufgabenphasen zufällig zuordnet, wobei man jedoch die in den Originaldaten für die beiden Aufgabenphasen gegebenen Gesamthäufigkeiten unverändert lässt. Wiederholt man diese zufälligen Zuordnungen N -mal, erhält man N Bootstrap-Schätzer KF^* für den Klassifikationsfehler unter der Nullhypothese (H_0), dass jegliche Clusterung im neuronalen Zustandsraum rein zufällig bedingt ist. Die H_0 -Verteilung wird also durch wiederholtes Ziehen aus den Originaldaten erzeugt, daher der Ausdruck Bootstrapping. **B) Die oben beschriebene** Prozedur vernachlässigt jedoch mögliche Autokorrelationen in den Daten, die in nahezu jeder Zeitreihe realer Systeme enthalten sind. Eine angemessenere Nullhypothese wäre daher, dass Gruppenbildungen im Zustandsraum durch die zeitlichen Autokorrelationen bedingt sind (aufgrund derer zeitlich benachbarte Punkte auch im Zustandsraum näher beieinander liegen), und nicht durch die Aufgabenphasenzugehörigkeiten. Die Bootstrap-Daten müssen daher Autokorrelationen bis zur Länge eines Aufgabenabschnittes erhalten. Dies kann dadurch erreicht werden, dass man ganze Blöcke aufeinanderfolgender Datenpunkte vertauscht, die jeweils einem ganzen Aufgabenabschnitt entsprechen.

Der Experimentator: Neurowissenschaften

Neu!

- Standardmethoden der Neurowissenschaften
- Fokus auf der Analyse des Vertebratengehirns
- Mit vielen Tipps und Tricks zur Bewältigung des Laboralltags



Guido Hermy / Claudia Mahlke / Michael Schwake / Tobias Sommer

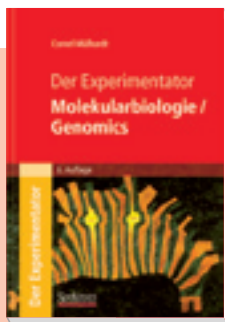
Der Experimentator: Neurowissenschaften

1. Aufl. 2010, 268 S. 100 Abb., kart.
€ (D) 32,95 / € (A) 33,88 / CHF 44,50
ISBN 978-3-8274-2368-9
Erscheint: September 2010

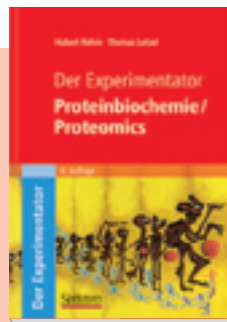
Dieser neue Band in der erfolgreichen EXPERIMENTATOR-Reihe soll dem angehenden Neurowissenschaftler einen Überblick über Fragestellungen und Methoden der neurowissenschaftlichen Forschung geben. Deshalb beschreiben die Autoren gut etablierte Standardmethoden und geben Einblicke in die aktuellen Trends und Entwicklungen, die die moderne neurowissenschaftliche Forschung vorantreiben. Der Fokus des Buches liegt auf der Erklärung von grundsätzlichen Mechanismen und Versuchsprinzipien. Zudem weist es auf viele „kleine“ Tricks des Laboralltags hin, die dem EXPERIMENTATOR das Leben erheblich erleichtern können.

Inhaltlich ist der Band auf die Analyse des Vertebratengehirns fokussiert, da es die Möglichkeit bietet, komplexe neuronale Vorgänge zu untersuchen, die z.B. für das Lernen, aber auch für die Analyse neuronaler Erkrankungen von Bedeutung sind. Methodisch wird dabei den Bogen gespannt von molekularen, proteinbiochemischen, zellbiologischen und elektrophysiologischen Ansätzen, über die Etablierung transgener Mausmodelle und deren Analyse (z.B. in verhaltensbiologischen Studien) bis hin zu nicht-invasiven Imaging-Methoden, die zur Untersuchung des menschlichen Gehirns einsetzbar sind.

Weitere Experimentatoren:



Cornelia Mülhardt
Der Experimentator: Molekularbiologie / Genomics
6. Aufl. 2009, 324 S., kart.
€ (D) 32,95 / € (A) 33,88 / CHF 44,50
ISBN 978-3-8274-2036-7



Hubert Rehm / Thomas Letzel
Der Experimentator: Proteinbiochemie / Proteomics
6. Aufl. 2010, 388 S., kart.
€ (D) 32,95 / € (A) 33,88 / CHF 44,50
ISBN 978-3-8274-2312-2



Sabine Schmitz
Der Experimentator: Zellkultur
2. Aufl. 2009, 289 S., kart.
€ (D) 32,95 / € (A) 33,88 / CHF 44,50
ISBN 978-3-8274-2108-1



Werner Luttmann et al.
Der Experimentator: Immunologie
3. Aufl. 2009, 314 S., kart.
€ (D) 32,95 / € (A) 33,88 / CHF 44,50
ISBN 978-3-8274-2026-8

Erhältlich in jeder Buchhandlung oder direkt beim Verlag:

- unter www.spektrum-verlag.de
- telefonisch: + 49 6221 345-0
- per Post: Springer Verlag Heidelberg
- per E-Mail: SDC-bookorder@springer.com
- per Fax: + 49 6221 345-4229
- Kundenservice Bücher • Haberstrasse 7 • D- 69126 Heidelberg

Alle Preise zzgl. Versandkosten (D: € 3,50 / A: € 3,90 / CH: SFR 6,20, jeweils pro Lieferung). Sämtliche Preise inkl. Mehrwertsteuer. Preise unter Vorbehalt. Der € (A)-Preis ist uns vom dortigen Importeur als Mindestpreis genannt worden. Der sFr-Preis ist eine unverbindliche Preisempfehlung.

Spektrum
AKADEMISCHER VERLAG



ganze Blöcke zeitlich konsekutiver Punkte, die einzelnen Aufgabenphasen entsprechen, permutiert (Abbildung 4B). Es gibt eine Reihe verschiedener Varianten dieser Blockpermutationsverfahren für Zeitreihen, sowie das weiter unten näher beschriebene Verfahren der Phasenrandomisierung, bei dem sowohl die Verteilung der Originaldaten als auch das gesamte Frequenzspektrum (und damit alle Autokorrelationen) erhalten bleiben (in den durch die Abtastrate und die endliche Länge der Zeitreihe gesetzten Grenzen). Zum Test der Nullhypothese wird nun eine Zahl N (z.B. 1000) von Bootstrap-Datensätzen erzeugt und die Statistik für jeden dieser Datensätze neu berechnet. Der üblichen statistischen Logik folgend lässt sich dann ein einfacher nicht-parametrischer Test konstruieren, der die Klassentrennung auf einem Signifikanzniveau von z.B. $p=.05$ prüft, indem ausgezählt wird, wie viele der Bootstrap-Datensätze einen kleineren Klassifikationsfehler aufweisen als die Originaldaten. Im vorliegenden Beispiel konnte so gezeigt werden, dass die neuronale Populationsaktivität zwischen fast allen Aufgabenabschnitten signifikant differenziert (nach Anpassung des Signifikanzniveaus durch die sogenannte Holm-Bonferroni-Prozedur für multiple Tests). Des Weiteren konnten signifikante Unterschiede im Grad der Strukturierung des neuronalen Zustandsraumes nachgewiesen werden zwischen Trials, in denen die Tiere nur wenige Fehler machten, und solchen, in denen es viele waren (Lapish et al. 2008).

Einzel-Trial-Analysen der neuronalen Dynamik

Mithilfe simultaner Ableitungen vieler Neurone lassen sich Fragestellungen angehen, die sich mit Daten von einzelnen oder nur wenigen Zellen nur schwer oder gar nicht beantworten lassen. Beispiele dafür sind die Beiträge neuronaler Korrelationen höherer Ordnung zum neuronalen Code (Ohiorhenuan et al. 2010) oder die zeitliche Dynamik neuronaler Ensembles, wie etwa die Übergänge zwischen verschiedenen Ensemblezuständen. Ein weiterer großer Vorteil multipler Einzelzellmessungen besteht darin, dass sie die Analyse einzelner Durchläufe (Trials) ermöglichen: Die Varianzreduktion und erhöhte statistische Power, die normalerweise durch Mittelung über mehrere Trials hinweg erreicht wird, kann zumindest teilweise durch geeignete Kombinationen der vielen gleichzeitig abgeleiteten Neurone während eines einzigen Trials ausgeglichen werden (z.B. ist die Diskriminationsleistung einer abgeleiteten Population in der Regel weit größer als die der besten einzelnen Zelle; s. z.B. Durstewitz et

al. 2010). Mit diesem Ansatz kann eine ganze Reihe neuer bzw. lange unbeantworteter Fragen angegangen werden, die auf die Daten aus einzelnen Trials angewiesen sind, wie etwa die Frage nach den Auswirkungen neuronaler Trial-zu-Trial-Variabilität auf Kognition und Verhalten (siehe z.B. Yu et al. 2009).

Vor Kurzem haben wir diese Strategie genutzt, um neuronale Dynamiken im präfrontalen Kortex der Ratte über einzelne Trials hinweg zu verfolgen, während die Tiere lernten, von einer alten auf eine neue Verhaltensregel umzuschalten (Durstewitz et al. 2010). Die Tiere wurden zunächst auf eine einfache Unterscheidungsregel trainiert, bei der sie einen von zwei möglichen Hebeln gemäß eines visuellen Hinweisreizes betätigen mussten (visuelle Regel). Ab einem bestimmten Trial wurde dann die Regel plötzlich und für die Tiere unerwartet so geändert, dass nun völlig unabhängig vom Hinweisreiz stets die Betätigung nur des einen oder des anderen Hebels (auf der rechten bzw. linken Seite) zur Belohnung führte (räumliche Regel, Abbildung 5A). Diese Art von Regelwechselaufgabe ist auf einen intakten präfrontalen Kortex angewiesen, wie bereits in früheren Studien gezeigt wurde. Die gesamte Aufgabe wurde so konstruiert, dass sowohl eine ausreichende Anzahl an Trials mit der ersten (visuellen) Regel als auch das Umlernen auf die zweite (räumliche) Regel innerhalb eines Tages erfolgen können, sodass während des gesamten Regelwechsels aller Wahrscheinlichkeit nach dieselbe Neuronenpopulation erfasst wird. Wie in Abbildung 5B zu sehen ist, sind die beiden Regeln nach Erlernen in deutlich unterschiedlichen Bereichen des neuronalen Zustandsraums repräsentiert. Unser Hauptinteresse galt nun der Frage, in welcher Form sich die neuronale Populationsaktivität beim Umschalten von der alten auf die neue Regel ändert, insbesondere ob dies in gradueller Weise (über die gesamte Lernphase hinweg) oder eher abrupt geschehen würde. Detaillierte statistische Analysen von Lernkurven in Verhaltensexperimenten in der Gruppe von Randy Gallistel an der Rutgers Universität, New Brunswick, haben vor Kurzem gezeigt, dass in verschiedenen klassischen und operanten Konditionierungsaufgaben sehr plötzliche Veränderungen in den Verhaltensleistungen auftreten können. Dies widerspricht der weit verbreiteten Ansicht, dass die meisten Lernprozesse in Tieren relativ langsam ablaufen und sich Assoziationen zwischen Stimuli, Reaktionen und Umwelt-Feedback in einem eher inkrementellen Prozess durch graduelle Veränderungen von Synapsenstärken bilden.

In Einklang mit den Ergebnissen der Gallistel-Gruppe fanden wir ebenfalls scharf

definierte Wendepunkte in der Verhaltensleistung, d.h. es gab häufig keine weichen allmählichen Übergänge zwischen einem zufälligen Reaktionsniveau und Beherrschung der neuen Regel, sondern eher sprunghafte Veränderungen innerhalb einiger weniger Trials. Um die Art der Übergänge auf neuronaler Ebene zu quantifizieren, wurden für jeden Trial die Abstände der neuronalen Trajektorie zu den beiden regelrepräsentierenden stationären Zuständen berechnet (Abbildung 5B). Bei einem graduellen Lernprozess würde man erwarten, dass sich der neuronale Zustand nur langsam von der einen Regelrepräsentation zur anderen bewegt, d.h. dass die Distanz zur visuellen Regel langsam anwächst, während die zur räumlichen Regel sinkt. Distanzen im neuronalen Zustandsraum wurden dabei als Mahalanobis-Distanzen definiert, die man sich als auf die Datenstreuung innerhalb der Gruppen normalisierte euklidische Abstände zwischen zwei Gruppenmittelwerten vorstellen kann (die Mahalanobis-Distanz entspricht in der Tat der euklidischen Distanz, nachdem alle Variablen auf eine Varianz von Eins normiert und de-korreliert wurden). Anders als der euklidische Abstand berücksichtigt die Mahalanobis-Distanz also gewissermaßen auch die statistische Unsicherheit über die genaue Position der neuronalen Zustände, oder auch das Faktum, dass stationäre (z.B. regelrepräsentierende) Zustände schon allein aufgrund der Systemdynamik (unabhängig von statistischer Unsicherheit) geometrisch ausgedehnte Objekte sein können. Zur Mahalanobis-Distanz tragen am stärksten diejenigen Richtungen im Zustandsraum bei, entlang derer die mittleren Abstände zwischen relevanten Punktgruppen am größten und die Variation innerhalb dieser Gruppen am geringsten ist. Dies erinnert an das weiter oben eingeführte Fishersche Diskriminationskriterium, und tatsächlich sind Klassifikationen auf Basis von Mahalanobis-Distanzen (mit gepoolter Kovarianzmatrix) formal äquivalent zu einer linearen Diskriminanzanalyse (unterschiedliche a-priori Wahrscheinlichkeiten nicht berücksichtigend).

Abbildung 5C zeigt die Trial-Zeitreihe der Differenz S der Mahalanobis-Distanzen der jeweils aktuellen neuronalen Zustände zu den stationären (nach Lernerfolg) regelrepräsentierenden Zuständen (d.h. $S_i = D_i^{\text{visuell}} - D_i^{\text{räumlich}}$). Wie man der Abbildung entnehmen kann, zeigt dieses Maß keinen graduellen Anstieg, der sich über die gesamte Lernphase erstreckt, sondern eine relativ sprunghafte Veränderung bei etwa Trial 74. Diese Aussage lässt sich auf verschiedene Weise statistisch formalisieren. Eine dieser Formalisie-

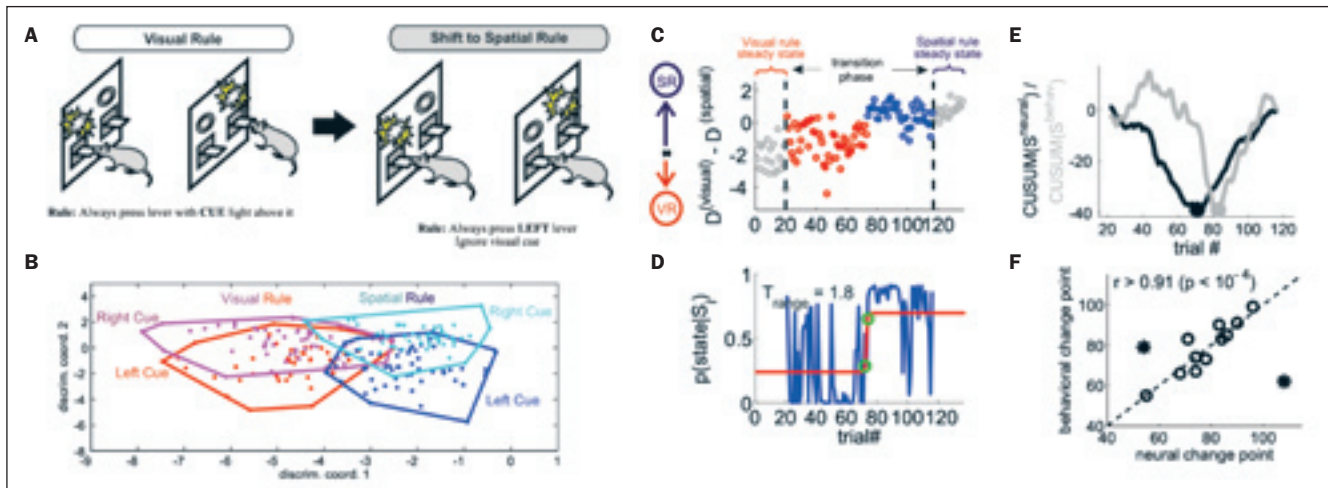


Abb. 5: Einzel-Trial-Analyse von Übergängen zwischen den neuronalen Repräsentationen verschiedener Regeln während einer Umlernaufgabe. A) Illustration des Versuchsaufbaus: Die Ratten führen zunächst 20 Trials nach einer zuvor gelernten Regel durch, in der ein visueller Hinweisreiz den korrekten Hebel indiziert. Nach 20 Trials auf dieser visuellen Regel wurden die Belohnungskontingenzen so geändert, dass völlig unabhängig vom Hinweisreiz immer derselbe Hebel (d.h. immer links oder rechts) gedrückt werden musste (räumliche Regel). B) Differenzierung der beiden Verhaltensregeln (rötliche und bläuliche Punkte) und der beiden Hinweisreize (hellere oder dunklere Punkte) durch die neuronale Populationsaktivität während des stationären (gut gelernten) Zustandes in der Regelwechselaufgabe. Gezeigt ist eine zweidimensionale Projektion des höherdimensionalen Raums der instantanen Feuerraten, die mithilfe der FDA (vgl. Abb. 2C) erzeugt wurde, um Unterschiede zwischen den Regeln und Hinweisreizen möglichst deutlich sichtbar zu machen. Die vier Gruppen umschließenden Linien sind die konvexen Hüllen der jeweiligen Punktwolken, stellen also die maximale Datenausdehnung in jeder Gruppe dar. C) Differenz zwischen den Mahalanobis-Distanzen zwischen dem aktuellen neuronalen Zustand und den beiden regelrepräsentierenden stationären Zuständen, aufgetragen als Funktion der Anzahl von Trials. Die die beiden stationären Zustände definierenden Bereiche sind durch vertikale gestrichelte Linien gekennzeichnet. Mithilfe eines Hidden Markov-Modells (HMM) lassen sich zwei verschiedene Zustände identifizieren (rote und blaue Kreise), die durch einen recht scharfen Übergang nahe Trial 74 getrennt sind. D) Bedingte Wahrscheinlichkeit für das Vorliegen des der räumlichen Regel entsprechenden Zustandes (wie über das HMM identifiziert) als Funktion der Trialanzahl. Um die Steilheit des Übergangs zu bestimmen, wurde eine logistische (Sigmoid-) Funktion an diese Wahrscheinlichkeiten angepasst. Der Bereich der x-Achse (d.h. die Anzahl an Trials), die dem y-Achsenabschnitt entspricht, in dem die angepasste Funktion von 10% auf 90% ihres Maximalwertes klettert (markiert durch die grünen Kreise), wurde als Teststatistik (T_{range}) verwendet. E) Eine andere Möglichkeit diese Zeitreihe zu repräsentieren besteht darin, die Abweichung der Distanzendifferenz vom Mittelwert zu kumulieren (siehe Text) und als Funktion der Trialanzahl aufzutragen (schwarze Kurve). Die Kumulation verringert die Variabilität in der Zeitreihe und Übergangspunkte (schwarzer Kreis) in der ursprünglichen Zeitreihe sind leicht als Extrema der kumulierten Zeitreihe zu erkennen. Die graue Kurve (skaliert auf den Minimalwert der neuronalen Kurve) zeigt dieselbe Darstellung für die Verhaltensleistung, und der graue Kreis markiert den entsprechenden Übergangspunkt. F) In 11 der 13 erhobenen Datensätze stimmten die Zeitpunkte der Übergänge in den neuronalen und den Verhaltensdaten bemerkenswert gut überein. Die beiden Ausnahmen (durch Sterne markiert) stammen von Datensätzen, die eine insgesamt geringe Variabilität in der neuronalen Zeitreihe aufwiesen, sodass die Übergangspunkte nicht gut lokalisiert werden konnten, bzw. stammen von einem Tier mit schlechter Verhaltensleistung. Nachdruck mit freundlicher Genehmigung aus Durstewitz et al. (2010), *Neuron* (Copyright: Cell Press, Elsevier).

rungen gründet auf der Annahme, dass die Zeitreihe tatsächlich von einer an sich nicht direkt beobachtbaren Abfolge neuronaler Zustände erzeugt wird. D.h., dass sich das neuronale System in jedem Trial in einem (an sich unbeobachteten) Zustand K befindet, in dem die direkt beobachteten Werte S mit einer bedingten Wahrscheinlichkeit $p(S|K)$ auftreten. Nimmt man an, dass der aktuelle Zustand K_i nur vom vorhergehenden Zustand K_{i-1} abhängt, spricht man von einem Hidden Markov-Modell (HMM) erster Ordnung. Ein HMM ist eindeutig bestimmt durch die Matrix der Übergangswahrscheinlichkeiten zwischen den Zuständen K und den bedingten Wahrscheinlichkeitsverteilungen $p(S|K)$. Solche Modelle werden gewöhnlich mithilfe des Viterbi-Algorithmus an die

Daten angepasst. Es gibt bereits mehrere Beispiele, in denen HMM zur Bestimmung von Zustandsübergängen in multiplen Einzelzellableitungen verwendet wurden (s. z.B. Jones et al. 2007). Die Farbcodierung in Abbildung 5C entspricht den Zuständen, die auf diese Weise identifiziert wurden. Änderungen des Farbcodes entsprechen also Wendepunkten in der neuronalen Dynamik in der Sequenz der Trials. Die Geschwindigkeit des Übergangs in der neuronalen Dynamik lässt sich nun durch Betrachtung der bedingten Wahrscheinlichkeiten $p(K_i = \text{'räumliche Regel'} | S_i)$ genau quantifizieren, d.h. der Wahrscheinlichkeiten dafür, dass sich das System im die neue räumliche Regel repräsentierenden Zustand K_i befindet, gegeben die beobachtete neuronale Distanzendifferenz S_i . Bei mindestens der

Hälfte der 13 untersuchten Tiere stieg diese Wahrscheinlichkeit innerhalb nur weniger Trials sprunghaft an (Abbildung 5D). Teststatistisch ließ sich dies absichern, indem die Geschwindigkeit der Zustandsübergänge in den Originaldaten mit Bootstrap-Daten verglichen wurde, in denen die Übergangspunkte zufällig entlang der Zeitreihe verschoben wurden. Die durch die HMM-Analyse gefundenen Übergangspunkte sind also tatsächlich besondere, singuläre Punkte, die nicht durch zufällige Fluktuationen erklärt werden können, wie sie auch an anderen Stellen entlang der Zeitreihe auftreten können.

Diese Schlussfolgerung wurde noch durch ein weiteres statistisches Verfahren untermauert, das unter dem Namen Change-Point-Analyse (CP) bekannt ist. Bei diesem



Verfahren wird die Teststatistik meist aus den kumulierten Differenzen der Werte S_i zum Mittelwert der Zeitreihe gebildet. Eine solche Statistik,

$$CUSUM(S_j) = \sum_{j \leq i} (S_j - \langle S \rangle)$$

ist in Abbildung 5E als Funktion der Trials abgebildet. In dieser Darstellung zeigen sich Wendepunkte (change-points) als Minima oder Maxima der Kurve: Solange S_i unterhalb des Mittelwerts bleibt, fällt die Kurve ab, während sie stetig ansteigt für $S_i > \langle S \rangle$. Nach Abzug langsamer, sich über die gesamte Zeit erstreckender Änderungen von der ursprünglichen Zeitreihe (Detrending), sollte dieses Maß um Null herum fluktuieren, sofern die neuronale Dynamik keine Zustandsübergänge aufweist, wohingegen abrupte Übergänge von geringen zu hohen S_i -Werten zu deutlichen Minima in den Kurven führen sollten. Daher kann man die größte Abweichung der Kurve von Null,

$$T_{CP} = \max_i \left| \sum_{j \leq i} (S_j - \langle S \rangle) / M \right|,$$

als Indikator für die Existenz von Wendepunkten verwenden. Um die Signifikanz der so gewonnenen T_{CP} Werte zu testen und Konfidenzintervalle für die genaue Position der Wendepunkte (d.h. ihre zeitliche Präzision) zu ermitteln, kann man phasenrandomisierte Bootstrap-Daten verwenden, wie schon weiter oben kurz eingeführt. Diese Bootstrap-Daten erhalten die Verteilung und, wichtiger noch, auch die Autokorrelationen der ursprünglichen Zeitreihe (bzw., äquivalent dazu, ihr Frequenzspektrum). Diese Bootstrap-Daten enthalten also alle linearen Komponenten und Fluktuationen der Originaldaten, nicht jedoch sprunghafte und über die Zeit konsistente Veränderungen im Mittelwert, wie sie für plötzliche Übergänge in der neuronalen Dynamik charakteristisch sind. Tatsächlich wies mehr als die Hälfte der originären Zeitreihen gegenüber den Bootstrap-Daten signifikant erhöhte T_{CP} -Werte auf, und der Wendepunkt konnte ferner in 5/13 Datensätzen mit 95%-iger Wahrscheinlichkeit innerhalb eines Intervalls von nur 6 Trials lokalisiert werden. Außerdem waren die in der neuronalen Dynamik identifizierten Wendepunkte zeitlich eng korreliert mit den im Verhalten identifizierten (Abbildung 5F). Insgesamt weisen diese Einzel-Trial-Analysen sowohl auf neuronaler als auch auf Verhaltensebene also darauf hin, dass die Tiere ihre Verhaltensstrategie relativ plötzlich an die neue Regel anpassen, fast so als hätten sie einen plötzlichen Moment der Einsicht gehabt.

Zusammenfassend zeigen diese Beispiele, wie gleichzeitige Ableitungen von vielen Neuronen zusammen mit geeigneten statistischen Analyseverfahren genutzt werden können, um Phänomenen auf der Einzeltrial-Ebene auf den Grund zu gehen, die ansonsten im Rauschen verborgen bleiben würden.

Ausblick: Rekonstruktion von neuronalen Trajektorien und gerichteter Aktivität im Zustandsraum

Abbildung 3B zeigt, wie sich die neuronale Populationsaktivität gemäß verschiedener kognitiv definierter Abschnitte einer Verhaltensaufgabe im neuronalen Zustandsraum gruppiert. Aus multiplen Einzelzellaufzeichnungen lassen sich aber noch sehr viel mehr Informationen über die Systemdynamik extrahieren als in dieser Abbildung dargestellt. Zum einen ist es aus der Sicht der Theorie dynamischer Systeme von großer Bedeutung zu wissen, wie genau sich die Aktivität zwischen verschiedenen Zuständen im Zustandsraum *zeitlich entwickelt*, d.h., wie genau die neuronalen Trajektorien verlaufen, die verschiedene Aktivitätsmuster miteinander verbinden. Mazor und Laurent (2005) zum Beispiel benutzten einen solchen Ansatz um zu zeigen, dass sehr viel mehr Information über einen sensorischen (olfaktorischen) Stimulus in der *transienten* neuronalen Dynamik enthalten ist als in den stimulus-bedingten stationären Zuständen, in die das System letztlich konvergiert. Genauer gesagt waren die mit verschiedenen olfaktorischen Stimuli assoziierten Trajektorien im neuronalen Zustandsraum in der transienten Phase weiter voneinander entfernt (besser separiert) als in den Fixpunkten (stationären Zuständen), die nach etwa zwei bis drei Sekunden Reizpräsentation erreicht wurden.

Ein weiteres zentrales Konzept in den theoretischen Neurowissenschaften ist die Idee, dass neuronale Berechnungsoperationen als Bewegung neuronaler Trajektorien zwischen verschiedenen Attraktorzuständen implementiert sind. Attraktorzustände kann man sich als stabile Muster neuronaler Aktivität denken, auf die die neuronale Dynamik mit der Zeit zukonvergiert, an denen sie für einige Zeit persistiert und die robust gegenüber (kleinen) Störungen sind. Die aktive Aufrechterhaltung von Informationen über Stimuli oder Reaktionen bei Arbeitsgedächtnisaufgaben oder die räumlichen Repräsentationen im Hippocampus zum Beispiel wurden oft mithilfe neuronaler Attraktoren erklärt. Kognitive Prozesse wie z.B. der Abruf einer zuvor gelernten Gedächtnissequenz können in diesem Rahmen als ein sequenzielles Ablaufen dieser Attraktorzustände verstanden werden. Simultane Einzelzellaufzeichnungen in Verbindung mit

sophistizierten statistischen Methoden für die Rekonstruktion von Zustandsräumen können solche Attraktorzustände visualisieren und dingfest machen. Ein konzeptuelles Problem dabei ist jedoch, dass die neuronale Dynamik hinreichend ‚entfaltet‘ sein muss, um die für Attraktorzustände definitorische Eigenschaft der Aktivitätskonvergenz nachweisen zu können (Abbildung 6): Da selbst simultane Messungen von 20-100 Neuronen letztlich nur einen sehr kleinen Ausschnitt des gesamten an der Aktivitätsdynamik beteiligten kortikalen Netzwerkes erfassen, werden die neuronalen Trajektorien des zugrunde liegenden (gesamten) dynamischen Systems in einen sehr viel niedrigerdimensionalen Raum projiziert (dem experimentell erfassbaren), in dem die Eindeutigkeit über die Richtungen der Aktivitätsänderungen verloren gehen kann. D.h., in diesem experimentell gemessenen, geringdimensionalen Ausschnitt des gesamten Zustandsraumes sind die Trajektorien häufig so vermengt und ineinander gefaltet, dass sie sich oft (annähernd, in Relation zum Rauschen) schneiden und es unmöglich wird, einen konsistenten Fluss der Aktivität in eine der Richtungen auszumachen (Abbildung 6). So wird ein möglicherweise klar konvergierender Aktivitätsfluss im eigentlichen hochdimensionalen neuronalen Raum in der geringdimensionalen experimentellen Projektion zu einem so ungeordneten Durcheinander von Flussrichtungen, dass keine attrahierenden Zustände mehr feststellbar sind. Eine mögliche Lösung für dieses Problem besteht darin, den geringdimensionalen experimentell erfassten Raum wieder in einen sehr viel höherdimensionalen Raum zu transferieren, indem aus den gemessenen Größen (neuronalen Aktivitätswerten) neue Variablen konstruiert werden, die dann als zusätzliche Dimensionen dem ursprünglich erfassten Raum hinzugefügt werden (Abbildung 6). In diesem Sinne haben wir wie in einer kürzlichen Studie (Balaguer-Ballester et al., in Vorbereitung) den ursprünglich abgeleiteten Aktivitätsraum um zeitversetzte Versionen $v_i(t-\tau)$ der ursprünglichen Variablen $v_i(t)$, sowie durch Interaktionsterme höherer Ordnung wie etwa $v_i(t-\tau_i) \times v_j(t-\tau_j)^2$, erweitert. Auf diese Weise wurde es möglich, den Fluss, also die Richtung der neuronalen Aktivitätsänderungen, soweit zu ‚entfalten‘, dass die kognitiv relevanten Attraktorzustände vollständig aufgelöst werden konnten (Abbildung 6). Ein Problem hierbei ist allerdings, dass die rekonstruierten Zustandsräume so extrem hochdimensional werden (bis zu mehreren Tausend Dimensionen), dass es spezieller Algorithmen bedarf (sogenannte Kernel-Methoden, siehe z.B. Hastie et al. 2009), um in diesen Räumen mathematische Operationen



Spektrum Sachbücher

Bildungsfutter und Lesespaß in einem!

Spektrum
AKADEMISCHER VERLAG
Sachbuch

► Denk- und Diskussionsstoff zur Bildungsreform



Neu!

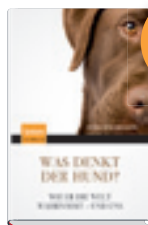
1. Aufl. 2010, 276 S., 56 Abb., geb.
€ (D) 19,95 / € (A) 20,51 / CHF 27,-
ISBN 978-3-8274-2677-2

Manfred Spitzer **Medizin für die Bildung**

Was kann die Bildung von der Medizin lernen? – Sehr viel! So die überraschende Antwort dieses Buchs. Denn es gibt für die Bildung ebenso einen Weg aus der Krise wie für den entzündeten Blinddarm: indem man die richtigen Diagnosen stellt und die richtigen Therapien sorgfältig erforscht. Damit unsere Kinder gut durch die Schule kommen, sollten wir nicht auf politische Reformen hoffen, sondern auf das Wissen über Lernen und Lernerfolg setzen.

In diesem provozierenden und zugleich ermutigenden Buch geht es um nichts weniger als eine neue Sicht von Bildung. Nicht Leistungsziele und Wissenskanons sind gefragt, sondern die Förderung von Neugier und der Lust am Lernen, die auch die Kraft zum nachhaltigen Üben stiftet. An ganz konkreten Beispielen wird aufgezeigt, wie Bildung – von Geburt an – funktioniert.

► Was wissen Hunde? Wie denken sie?



Neu!

1. Aufl. 2010, 390 S.
24 Abb., geb. m. SU
€ (D) 29,95 /
€ (A) 30,79 / CHF 40,50
ISBN 978-3-8274-2459-4

Alexandra Horowitz **Was denkt der Hund?**

Dieses Buch wirft einen unbefangenen Blick auf die Welt der Hunde – aus deren Perspektive. Als Kognitionspsychologin und Hundeliebhaberin ist die Autorin begierig zu erfahren, was ihr Hund und andere Hunde denken, wissen und fühlen und wie sie die Welt erfahren. Klar, verständlich und amüsant führt sie den Leser in die perzeptuellen und kognitiven Fähigkeiten der Hunde ein.

► Warum Menschen an übernatürliche Phänomene und Magie glauben



Neu!

1. Aufl. 2010, 361 S.
22 Abb., geb. m. SU
€ (D) 24,95 /
€ (A) 25,65 / CHF 33,50
ISBN 978-3-8274-2543-0
Erscheint: Oktober 2010

Bruce M. Hood **übernatürlich? natürlich!**

Die Mehrheit der Weltbevölkerung ist religiös oder glaubt an übernatürliche Phänomene. Viele von uns drücken jemandem die Daumen, klopfen auf Holz oder meiden schwarze Katzen. Woher rührt dieses übernatürliche Gedankengut? Werden wir von unseren Eltern, von Kirchen und Medien indoktriniert, oder entstehen diese Glaubensinhalte auf andere Weise? In *übernatürlich? natürlich!* gewährt uns der mehrfach ausgezeichnete Kognitionspsychologe Bruce M. Hood tiefe Einblicke in die Wissenschaft des Glaubens an das Übernatürliche.

Das neue Buch von Gerhard Roth

► Was unterscheidet den Menschen von anderen intelligenten Tieren?



Neu!

1. Aufl. 2010, 434 S.
54 Abb., geb. m. SU
€ (D) 24,95 /
€ (A) 25,65 / CHF 33,50
ISBN 978-3-8274-2147-0

Gerhard Roth

Wie einzigartig ist der Mensch?

Wie dieses Buch zeigt, haben sich Wahrnehmungs- und Erkenntnisleistungen, Intelligenz, Geist und Bewusstsein im Laufe der Evolution allmählich entwickelt – in deutlicher Parallelität und gegenseitiger Bedingtheit. Gerhard Roth beleuchtet insbesondere die neuen empirischen Einsichten aus dem Bereich der Kognitionsforschung an Tieren und am Menschen sowie der Hirnforschung. Sie haben die Idee einer Ko-Evolution von Sinnesorganen, Nervensystemen und Gehirnen einerseits und des Wahrnehmungs- und Erkenntnisvermögens sowie schließlich des Entstehens von Geist-Bewusstsein überhaupt erst plausibel gemacht.



Gerhard Roth, Direktor am Institut für Hirnforschung (Uni Bremen), zählt zu den bekanntesten Neurowissenschaftlern Deutschlands. Als erfolgreicher Sachbuchautor ist er regelmäßig Gast von Fernseh- und Rundfunksendungen.

Erhältlich in jeder Buchhandlung oder direkt beim Verlag:

► unter www.spektrum-verlag.de
► per E-Mail: SDC-bookorder@springer.com

► telefonisch: + 49 6221 345-0
► per Fax: + 49 6221 345-4229

► per Post: Springer Verlag Heidelberg
Kundenservice Bücher • Haberstrasse 7 • D- 69126 Heidelberg

Alle Preise zzgl. Versandkosten (D: € 3,50 / A: € 3,90 / CH: SFR 6,20, jeweils pro Lieferung). Sämtliche Preise inkl. Mehrwertsteuer. Preise unter Vorbehalt.
Der € (A)-Preis ist uns vom dortigen Importeur als Mindestpreis genannt worden. Der sFr-Preis ist eine unverbindliche Preisempfehlung.

Spektrum
AKADEMISCHER VERLAG



sinnvoll durchführen zu können. Solche Ansätze sind jedoch im Bereich der nichtlinearen Dynamik (sogenannte Einbettungstheoreme, Sauer 1991) und statistischen Lerntheorie gut etabliert. Derartige Ansätze ermöglichen einen weitaus detaillierteren Blick auf die neuronale Dynamik und damit auch die neuronale Implementation kognitiver Prozesse.

Literatur

- Balaguer-Ballester, E., Lapish, C., Seamans, J.K. und Durstewitz, D.: Attracting Dynamics of Prefrontal Cortex Ensembles During Memory-Guided Decision-Making. In preparation.
- Broome, B.M., Jayaraman, V. und Laurent, G. (2006): Encoding and decoding of overlapping odor sequences. *Neuron* 51: 467-482.
- Compte, A., Constantinidis, C., Tegnér, J., Raghavachari, S., Chafee, M.V. et al. (2003): Temporally Irregular Mnemonic Persistent Activity in Prefrontal Neurons of Monkeys During a Delayed Response Task. *J Neurophysiol* 90: 3441-3454.
- Durstewitz, D., Vitoz, N.M., Floresco, S.B. und Seamans, J.K. (2010): Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* 66: 438-448.
- Hastie, T., Tibshirani, R. und Friedman, J. (2009): *Elements of Statistical Learning: Data Mining, Inference and Prediction* (Second Edition, New York: Springer-Verlag).
- Haynes, J.D. und Rees, G. (2006): Decoding mental states from brain activity in humans. *Nat Rev Neurosci.* 7: 523-534.
- Jones, L.M., Fontanini, A., Sadacca, B.F., Miller, P. und Katz, D.B. (2007): Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proc. Natl. Acad. Sci. USA.* 104: 18772-18777.
- Krzyszowski, W.J. (2000): *Principles of Multivariate Analysis: A User's Perspective* (Oxford: Oxford University Press).
- Lapish, C.L., Durstewitz, D., Chandler, L.J. und Seamans, J.K. (2008): Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proc. Natl. Acad. Sci. USA.* 105: 12010-12015.
- Mazor, O. und Laurent, G. (2005): Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron* 48: 661-673.
- Ohirohuan, E., Mechler, F., Purpura, K.P., Schmid, A.M., Hu, Q. und Victor, J. (2010): Sparse coding and high-order correlations in fine-scale cortical networks. *Nature* 466: 617-621.
- Roweis, S.T. und Saul, L.K. (2000): Nonlinear dimensionality reduction by locally linear embedding. *Science* 290: 2323-2326.
- Sauer, T., Yorke, J. und Casdagli, M. (1991): Embedology. *J. Stat. Phys.* 65: 579-616.
- Tenenbaum, J.B., de Silva, V. und Langford, J.C. (2000): Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290: 2319-2323.
- Yu, B.M., Cunningham, J.P., Santhanam, G., Ryu, S.I., Shenoy, K.V. und Sahani, M. (2009): Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *J. Neurophysiol.* 102: 614-635.

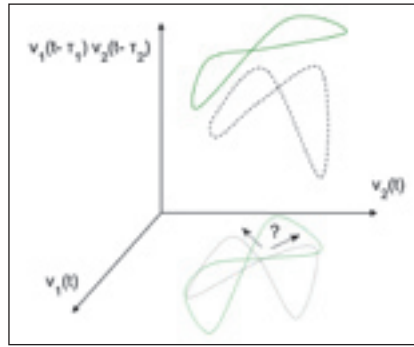


Abb. 6: Entfaltung von Trajektorien neuronaler Populationsaktivität durch Erweiterung des experimentell gegebenen Zustandsraumes. Das Schema zeigt zwei neuronale Trajektorien, die im unteren Teil der Abbildung auf eine Ebene projiziert sind, die von den instantanen Feuerraten zweier abgeleiteter Neurone aufgespannt wird. Die aus einem eigentlich höherdimensionalen System stammenden Trajektorien schneiden sich in der zweidimensionalen Projektion häufig, sodass an mehreren Punkten die zeitliche Entwicklung des Systems nicht eindeutig bestimmt ist (angedeutet durch die beiden Pfeile und das Fragezeichen). Es ist daher nicht möglich, etwa die Konvergenz der Trajektorien auf einen möglichen Attraktorzustand zu untersuchen. Wenn man jedoch eine dritte Achse hinzufügt, die dem Produkt aus zwei zeitversetzten Werten der Originalvariablen, $v_1(t-\tau_1)$ und $v_2(t-\tau_2)$, entspricht, werden die Trajektorien so voneinander getrennt („entfaltet“; oberer Teil des Schemas), dass die zeitliche Entwicklung rekonstruiert werden kann. Auf diese Weise können fehlende Dimensionen des experimentell bestimmten Zustandsraumes durch Kombinationen gemessener Variablen ersetzt werden.

Danksagung

Diese Arbeit wurde mit Mitteln der Deutschen Forschungsgemeinschaft an DD (Du 354/5-1 & 6-1) und des Bundesministeriums für Bildung und Forschung (BMBF, 01GQ1003B) im Rahmen der Bernstein Center for Computational Neuroscience gefördert. Wir danken unseren kanadischen Kooperationspartnern und Ko-Autoren der Originalpublikationen, Jeremy Seamans, Christopher Lapish, Nicole Vitoz und Stan Floresco, herzlich für die Bereitstellung der experimentellen Daten für die hier besprochenen Analysen. Ferner danken wir Loreen Hertäg und Joachim Hass für die freundliche Übersetzung der englischen Version dieses Artikels.

Kurzbiografien

Daniel Durstewitz studierte Psychologie mit Schwerpunkt Statistik und Nebenfach Mathematik an der Technischen Universität Berlin (1989 bis 1994). Von 1994 bis 1998 war er Doktorand im Graduiertenprogramm „Kognition, Gehirn und neuronale Netze“ an der Ruhr-Universität Bochum. Von 1998 bis 2000 war er Postdoktorand (research associate) am Salk Institute for Biological Studies (Computational Neurobiology Lab) in La Jolla, USA. Von 2001 bis 2005 arbeitete er als Juniorgruppenleiter (Computational Neuroscience Lab) im Rahmen des Emmy-Noether-Programms der DFG am Institut für Kognitive Neurowissenschaften der Ruhr-Universität Bochum. Von 2005 bis 2008 war er Reader for Computational Neuroscience am Centre for Theoretical & Computational Neuroscience an der Universität Plymouth in Großbritannien. Seit Oktober 2008 hat er im Rahmen des Heisenberg-Programms eine Gruppenleiterstelle am Zentralinstitut für Seelische Gesundheit in Mannheim inne und ist Koordinator des Bernsteinzentrums für Computational Neuroscience Heidelberg-Mannheim.

Emili Balaguer-Ballester studierte von 1992 bis 1997 Physik (Theoretische Physik) an der Universität Valencia in Spanien. Von 1998 bis 2001 war er Doktorand an der Universität Valencia in der Abteilung für Elektroingenieurwesen (Physik) und Pflanzenbiologie (Fakultät für Biologie), wo er an der Analyse und Verlaufsvorhersage atmosphärischer Schadstoffe arbeitete. Von 2001 bis 2005 war er als wissenschaftlicher Mitarbeiter in der Abteilung für Künstliche Intelligenz und Maschinelles Lernen bei der Firma TISSAT (ein in Valencia, Spanien, ansässiges Software-Unternehmen) beschäftigt (Datenbankauswertung). Von 2005 bis 2008 war er Postdoktorand am Centre for Theoretical and Computational Neuroscience an der Universität Plymouth, Großbritannien. Seit Dezember 2008 ist er Senior Postdoktorand am Zentralinstitut für Seelische Gesundheit.

Korrespondenzadresse

Dr. Daniel Durstewitz
 Bernstein Center for Computational Neuroscience Heidelberg-Mannheim
 Zentralinstitut für Seelische Gesundheit
 J 5
 68159 Mannheim
 Tel.: +49 621 1703 3401
 Fax: +49 621 1703 2005
 E-Mail: daniel.durstewitz@zi-mannheim.de