Grenzwertsätze für Differentialgleichungen mit zufälligen Kopplungen

Thomas Wennekers
Universität Ulm
Fakultät für Informatik
Abteilung Neuroinformatik
D-89069 Ulm, Germany

11. September 1997

Zusammenfassung

Die Analyse des globalen zeitlichen Verhaltens der Aktivität in großen neuronalen Netzen erfordert die Reduktion der hoch-dimensionalen Netzwerkgleichungen auf kleine, mathematisch leichter zugängliche Gleichungssysteme. In dieser Arbeit wird eine Klasse stochastisch gekoppelter Integro-Differentialgleichungen betrachtet, die eine Verallgemeinerung häufig verwendeter zeitkontinuierlicher Modelle neuronaler Netze mit graduellen, sigmoiden Neuronen darstellt. Es werden Bedingungen an die Dynamik der Einzelneurone und die statistischen Eigenschaften der Kopplungen angegeben, unter denen das betrachtete System asymptotisch in der Systemgröße und der Zeit exakt auf ein niedrigdimensionales System reduziert werden kann. Hierbei können die Einzelzellen mehrkomponentig, schwach nichtlinear und zeitabhängig sein; Kopplungen werden in Form von Responsekernen mit statistischen Parametern erfaßt.

1 Einleitung

Die Untersuchung der Vorgänge in biologischen neuronalen Netzen erfordert gewöhnlich die Modellierung sehr großer Netzwerke gekoppelter, meist nichtlinearer Einzelzellen. Eine detaillierte mathematische Analyse oder gar die Angabe vollständiger Lösungen der das Netzwerk beschreibenden Gleichungen, wird daher nur in Ausnahmefällen möglich sein. Aus diesem Grund beschränkt man sich häufig auf die Betrachtung des "typischen" oder "mittleren" Verhaltens, zum Beispiel dem zeitlichen Verlauf des Ensemble-Mittelwertes der Zellaktivität. Hierzu ist für die jeweils konkret betrachtete globale Größe aus dem vollständigen System der Netzwerkgleichungen ein niedrigdimensionales System abzuleiten. Bei dieser Reduktion ist man meist auf weitgehend heuristische Methoden angewiesen.[3, 16, 13, 17]

Einige der auf diese Weise gefundenen Gleichungen für das Verhalten der Netzwerke im Mittel lassen sich jedoch mathematisch strikt in überraschend allgemeiner Weise begründen. Dies ist Gegenstand dieser Arbeit.

Hierbei betrachten wir Netzwerke, die aus einer beliebigen endlichen Anzahl von Gruppen (Pools) jeweils identischer Zellen bestehen. Die Theorie ist anwendbar auf graduelle Neuronen, wobei ein relativ kompliziertes Einzelneuronmodell verwendet wird, dessen innerer Zustand mehrdimensional und (schwach) nichtlinear sein kann, so daß neben dem üblicherweise in ähnlichen Modellen anzutreffenden einfachen Tiefpaßverhalten noch weitere, die Dynamik von Einzelzellen bestimmende Einflußgrößen im Modell berücksichtigt werden können. Die Zellen in verschiedenen solcher Pools können unterschiedliche Eigenschaften haben und etwa verschiedene Zellklassen beschreiben (Pyramidenzellen, Sternzellen, etc.).[19] Innerhalb jeder Gruppe und zwischen verschiedenen Gruppen sollen die Zellen stochastisch verknüpft sein, im einfachsten Fall durch zufällig generierte Kopplungskonstanten oder synaptische Gewichte. Wir wollen aber auch die Möglichkeit zeitlich transienter Wechselwirkungen zulassen, die biologisch plausibler ist und beispielsweise durch die Faltung der Eingangssignale mit "synaptischen Responsefunktionen" modelliert werden kann.[13] Hier sind dann die Parameter der Responsefunktionen Zufallsvariable. Die statistischen Eigenschaften der Kopplungsstärken und Responsefunktionen zwischen je zwei Pools dürfen sich voneinander unterscheiden (exzitatorische vs. inhibitorische Zellen, schnelle vs. langsame Synapsen; Verzögerungen, Anstiegs- und Abfallszeiten der Responsefunktionen usw.). Es zeigt sich, daß dann alle Zellen in den einzelnen Gruppen (unter einigen weiteren Voraussetzungen) asymptotisch in der Zeit und der Netzwerkgröße (d.h. mit wachsender Anzahl Zellen pro Pool) synchronisieren müssen, so daß zur Beschreibung des Netzwerkverhaltens schließlich ein niedrigdimensionales Gleichungssystem ausreicht, welches jeden Pool durch ein repräsentatives Neuron ersetzt. Dies stellt ein Grenzwertgesetz der großen Zahlen für zufällig gekoppelte Differentialgleichungen dar.

2 Netzwerk-Modell

Geman hat in [12] Differentialgleichungen mit konstanten Koeffizienten von der Form

$$\dot{x}_i = -x_i + \frac{1}{N}\sum_{j=1}^{N} C_{ij} f(x_j),\ x_i(0) = x_{i0} \in \mathbb{R};\ i = 1, 2, \cdots N$$

betrachtet, wobei die Koeffizienten C_{ij} unabhängig und identische verteilte Zufallsvariablen waren. Modelle dieser oder ähnlicher Form (im allgemeinen mit strukturierteren Kopplungsmatrizen) sind häufig zu finden bei der Modellierung neuronaler Netze.[16, 18, 25, 28] Unter der Annahme, daß die charakteristische Funktion $Ee^{i\lambda C_{11}}$ der Verteilung der Koeffizienten analytisch in $\lambda = 0$ ist (was insbesondere impliziert, daß alle Momente $EC_{11}^k;\ k = 1, 2\cdots$ existieren) und die Funktion f beschränkt und Lipschitzstetig ist, konnte gezeigt werden, daß asymptotisch für $N \to \infty$ und $t \to \infty$ alle Variablen $x_i(t),\ i = 1, 2, \cdots$ identisch werden und ihr zeitliches Verhalten außerdem einer eindimensionalen Differentialgleichung gehorcht.[12] Werden P Systeme (Pools) gleichzeitig betrachtet, die untereinander stochastisch miteinander gekoppelt sind, reduziert sich das System asymptotisch auf eine P-dimensionale Differentialgleichung bestehend aus einer Gleichung pro Pool, die das volle System, asymptotisch exakt, beschreibt. Der Beweis in [12] macht wesentlich Gebrauch von einem ebenfalls von Geman in [11] bewiesenen Resultat, daß nämlich, falls $EC_{11} = 0,\ EC_{11}^2 = \sigma^2$ und $EC_{11}^k \leq k^{\alpha k}$ für ein geeignetes α und $k > 2$, der obere Limes $\overline{\lim}_{N\to\infty} \left\|\frac{1}{\sqrt{N}}(C_{ij})_{1\leq i,j\leq N}\right\|_2$ fast sicher gleich 2σ ist. (Bem.: Mittlerweile ist gezeigt, daß hierfür bereits $EC_{11}^4 < \infty$ notwendig und hinreichend ist; siehe [4] und Referenzen darin).

Wir erweitern das Theorem von Geman [12] in verschiedene Richtungen. Theorem 8 stellt eine Version bereit, bei der an die Stelle der eindimensionalen, linearen Gleichung der ungekoppelten

Untersysteme, $\dot{x}_i = -x_i,\ x_i \in \mathbb{R}$, die höherdimensionale, explizit zeitabhängige und nichtlineare Differentialgleichung $\dot{x}_i = A(t)x_i(t) + H(x_i(t),t) + I(t),\ x_i \in \mathbb{R}^n,\ 1 \leq n < \infty$ tritt. Weiterhin werden die Voraussetzungen an die Verteilung der Kopplungskoeffizienten von der Existenz aller Momente auf lediglich $EC_{11}^6 < \infty$ abgeschwächt. Die Verwendung des recht schwierig zu beweisenden Theorems $\overline{\lim_{N\to\infty}} \left\| \frac{1}{\sqrt{N}}(C_{ij})_{1\leq i,j\leq N} \right\| = 2\sigma\ f.s.$ wird in unserem Beweis umgangen. Theorem 9 berücksichtigt anschließend zusätzlich synaptische Responsefunktionen mit zufälligen Parametern, d.h. die Zufallsvariablen C_{ij} werden durch Zufallsfunktionen $C_{ij}(t)$ ersetzt.

Wir betrachten die Netzwerkgleichungen

$$\begin{aligned} \frac{dx_i}{dt}(t) &= A(t)x_i(t) + H(x_i(t),t) + I(t) + \frac{1}{N}\sum_{j=1}^{N}\int_0^t C(t-s,\omega_{ij})F(x_j(s),s)ds \\ x_i(0) &= x_{i0} \in \mathbb{R}^n, \qquad i = 1\cdots N\ . \end{aligned} \tag{1}$$

Hier beschreibt x_i den internen Zustand eines Einzelelementes/Neurons. Dieser Zustand kann mehrkomponentig sein; durch Verwendung mehrerer Variable können beispielsweise Adaptation, Fazilitation oder andere aktivitätsabhängige Effekte erfaßt werden. Die Anzahl innerer Freiheitsgrade wird durch $n \geq 1$ gegeben, N ist die Anzahl der Elemente im Netzwerk. Die Dynamik eines isolierten Einzelelementes in (1) wird festgelegt durch die Differentialgleichung

$$\dot{x}_i(t) = A(t)x_i(t) + H(x_i(t),t) + I(t)\ . \tag{2}$$

Demgegenüber definiert der Summenterm in (1) die Wechselwirkung der Elemente untereinander. Wir nehmen an, daß alle Elemente identische Eigenschaften haben, aber untereinander zufällig gekoppelt sind.

In (2) repräsentiert die Funktion $I(t):\ \mathbb{R}_+ \to \mathbb{R}^n$ externe Einflüsse auf x_i. $A(t):\ \mathbb{R}_+ \to \mathbb{R}^{n\times n}$ ist eine Matrixfunktion und erfaßt den linearen Anteil der Dynamik eines Einzelelementes, wogegen $H(x,t):\ \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$ eine geeignete nichtlineare Funktion ist. In der Literatur findet man häufig den Fall $n = 1,\ H \equiv 0$ und $A(t) \equiv 1/\tau$, mit der Membranzeitkonstanten τ.[16, 18, 28] Gleichung (2) beschreibt mit diesen Einschränkungen als Spezialfall einen einfachen Tiefpaß. Indem man $n > 1,\ H \equiv 0$ und für $A(t)$ eine konstante Matrix A wählt, kann man zunächst weitere zeitverschiebungsinvariante, lineare Systeme erfassen. Weiterhin gestattet eine Zeitabhängigkeit von A, daß sich Eigenschaften der Neurone auch zeitlich ändern können und Hinzunahme von $H(x,t)$ schließlich läßt auch die Untersuchung gewisser nichtlinearer Eigenschaften von Neuronen zu.

Die Ausgabe einer Zelle wird durch eine Funktion $F(x,t):\mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^m$ beschrieben. Diese Funktion kann außer vom inneren Zustand x auch von der Zeit t abhängen und soll für alle Elemente des Netzwerks dieselbe sein. Man beachte, daß jede Komponente von x Einfluß auf die Ausgabe nehmen kann. Ebenso soll eine Zelle im Prinzip auf jede Komponente einer Zielzelle einwirken können, wobei dieser Einfluß insbesondere funktionell unterschiedlich für verschiedene Komponenten der Zielzelle sein kann. Beispielsweise mögen gewisse Prozesse in der Zielzelle Aktivierungsschwellen unterliegen, andere aber nicht (z.B. AMPA/NMDA-Rezeptoren). Aus diesem Grund ist F als m-komponentige Funktion angesetzt. Die Ausgabe kann dann nicht mehr einfach als Feuerrate der Zelle interpretiert werden. Will man darauf beharren, daß trotzdem im wesentlichen mittels Feuerraten Signale zwischen Zellen ausgetauscht werden, mag man die Komponenten von F als Funktionen der Feuerrate auffassen (die wiederum vom internen Zustand x abhängt).

Im Wechselwirkungsterm von Gleichung (1) tauchen neben $F(x,t)$ weiterhin Funktionen $C(t,\omega_{ij}):\ \mathbb{R}_+ \times \Omega \to \mathbb{R}^{n\times m}$ auf. Diese Funktionen repräsentieren die synaptischen Responsefunktionen, wobei $\omega_{ij} \in \Omega$ etwa einen Satz von Parametern wie Anstiegs- und Abfallszeiten, Übertragungstotzeit und Amplitude von postsynaptischen Potentialen darstellt und C ein durch ω_{ij} parametrisierter Funktionsprototyp ist. Für $\omega_{ij} = J_{ij}$ und $C(t, \omega_{ij}) = \omega_{ij}\delta(t)$ können die Integrale im Kopplungsterm in (1) aufgelöst werden und der Ausdruck reduziert sich auf die häufiger verwendete gewichtete Summe mit konstanten Koeffizienten; allerdings sind diese Koeffizienten für $m, n > 1$ genau genommen bereits $n \times m$-Matrizen. Zur Verallgemeinerung des Theorems von Geman werden wir annehmen, daß die Funktionen $C(t, \omega_{ij})$ Zufallsfunktionen sind, d.h. wir nehmen an, daß die ω_{ij} Elemente eines Wahrscheinlichkeitsraumes $(\Omega, \mathcal{F}, P)$ sind. C kann dann aufgefaßt werden als eine Funktion, die jedem Element $\omega \in \Omega$ eine Zeitfunktion $C(t, \omega)$ zuordnet.

Ohne weitere Einschränkungen an A und H ist (2) äquivalent zu der völlig allgemeinen Differentialgleichung $\dot{x}_i(t) = f(x_i(t), t)$ für eine beliebige Funktion f. In dieser Allgemeinheit kann man kaum erwarten Aussagen über das typische Verhalten des gekoppelten Systems (1) machen zu können. Im folgenden müssen wir daher weitere Forderungen an A und H stellen, die eine Reduktion von (1) auf ein niedrigdimensionales System ermöglichen. Die Annahmen hier sind recht rigoros und laufen - falls A und H asymptotisch konstant sind - darauf hinaus, daß das isolierte Einzelsystem (2) selbst keine autonomen, z.B. oszillatorische oder möglicherweise sogar chaotische, Lösungen besitzt, sondern nur asymptotisch stabile Fixpunkte. Interessanterweise gibt es in der Literatur Hinweise darauf, daß Reduktionstheoreme der hier behandelten Form auch für Systeme mit gewissen oszillatorischen oder sogar chaotischen Einzelelementen gelten könnten.[22, 20, 23] Die zitierten Arbeiten sind jedoch bisher numerischer Natur und betrachten auch lediglich den Fall der sogenannten 'Meanfield'-Kopplung, bei der alle Kopplungen identisch (nicht-zufällig) sind. Es ist weiterhin klar, daß sicher nicht beliebige 'Meanfield'-gekoppelte Systeme synchronisieren. Beispielsweise haben Pikovsky et al. erst kürzlich ein Phänomen "schwacher Synchronisation" in Systemen identisch gekoppelter chaotischer Oszillatoren gefunden, bei dem Phasenkohärenz nur in einem gewissen mittleren Sinn eintritt und die Amplituden der Schwingungen einzelner Oszillatoren chaotisch und nahezu unkorreliert sind.[24] Ein Beweis von Bedingungen, die Synchronisation (strikt oder schwach) solch allgemeiner, komplexer Systeme gewährleisten, wäre sehr nützlich, kann aber mit den nachfolgend verwendeten Methoden vermutlich nicht geführt werden.

Es sei bemerkt, daß wir die Diskussion soweit nur für eine einzige Gruppe identischer, zufällig gekoppelter Zellen geführt haben. Eine Verallgemeinerung von (2) und (1) auf endlich viele (z.B. P) Gruppen ist offensichtlich. Die Eigenschaften der Zellen in jedem Pool müssen dann als identisch angenommen werden, können sich aber zwischen Pools unterscheiden. Außerdem kann man eine beliebige Kopplungsstruktur der Pools untereinander annehmen. Die statistischen Eigenschaften der Kopplungen zwischen je zwei Pools müssen weiterhin identisch sein, können sich aber von Paar zu Paar unterscheiden. In (1) sind dann alle Größen noch durch ihre Zugehörigkeit zu dem jeweiligen Pool zu kennzeichnen, und es wird außer über j noch über alle Pools mit denen Verbindungen bestehen summiert (vgl. (14) und (23)).

3 Mathematische Hilfsmittel

Für die Analyse des asymptotischen Verhaltens des Systems (1) benötigen wir eine Reihe von Hilfsätzen, einerseits aus der Theorie der Differentialgleichungen [1], andererseits aus der Wahrscheinlichkeitstheorie [9, 6].

3.1 Differentialgleichungen und Matrixmaß

Wir beginnen mit einer fundamentalen Ungleichung, deren Beweis man zum Beispiel in [10] findet.

Lemma 1 *(Bellman-Gronwall) Sei (i) f,g,k:* $\mathbb{R}_+ \to \mathbb{R}$ *lokal integrabel; (ii)* $g \geq 0, k \geq 0$ *(iii) gk lokal integrabel auf* $\mathbb{R}_+$*. Falls dann für eine Funktion* $u : \mathbb{R}_+ \to \mathbb{R}$ *gilt:*

$$u(t) \leq f(t) + g(t) \int_0^t k(t')u(t')dt', \quad \forall t \in \mathbb{R}_+ \ ,$$

so gilt auch

$$u(t) \leq f(t) + g(t) \int_0^t k(t')f(t') \left[\exp \int_{t'}^t k(t'')g(t'')dt''\right] dt', \quad \forall t \in \mathbb{R}_+ \ .$$

Man beachte, daß in der zweiten Ungleichung u nicht mehr auf der rechten Seite enthalten ist, wohl aber in der ersten.
Wir betrachten nun die lineare, inhomogene Differentialgleichung, deren Theorie wohlbekannt ist (siehe etwa [1]):

$$\dot{x}(t) = A(t)x(t) + b(t), \ x(0) = x_0 \in \mathbb{R}^n \ . \tag{3}$$

Hier (wie im folgenden) nehmen wir an, daß A und b(t) beschränkt und stetig sein mögen. Ist $n = 1$, lautet die eindeutige Lösung des Anfangswertproblems (3)

$$x(t) = e^{\int_0^t A(t')dt'} x_0 + \int_0^t e^{\int_{t'}^t A(t'')dt''} b(t')dt', \ t \geq 0 \ .$$

Für $n > 1$ bleibt die Form der Lösung dieselbe, falls $A(t)$ und $\int_0^t A(t')dt'$ für jedes $t \in \mathbb{R}_+$ kommutieren, also speziell, wenn A eine beliebige konstante Matrix ist. Andernfalls ist die allgemeine Form der Lösung komplizierter. Schranken für das Wachstum von Lösungen können mit Hilfe einer Matrixnorm $||A||$ von A gewonnen werden. So gilt für die Euklidische Norm von Lösungen der zu der inhomogenen Gleichung (3) gehörigen homogenen Differentialgleichung

$$\dot{x}(t) = A(t)x(t), \ x(0) = x_0 \ , \tag{4}$$

etwa

$$||x(t)||_2 \leq ||x(0)||_2 \exp \int_0^t ||A(t')||_2 dt' \ .$$

Die Matrixnorm ist stets größer oder gleich Null und gleich Null nur für $A = 0$, so daß in allen interessierenden Fällen lediglich eine exponentiell in t anwachsende Schranke erhalten wird. Abschätzungen dieser Form sind für unsere Zwecke zu grob. Zur Beschränkung von Lösungen von (3) und letztlich solcher des Systems (1) verwenden wir daher das Konzept des *Matrixmaßes*.

Definition 2 *(Matrixmaß) Sei* $|\cdot|$ *eine Norm auf* $\mathbb{C}^n$ *und* $||\cdot||$ *die hiervon induzierte Matrixnorm auf* $\mathbb{C}^{n\times n}$*. Sei I die identische Abbildung in* $\mathbb{C}^{n\times n}$*. Als* Matrixmaß $\mu(A)$ *einer Matrix* $A \in \mathbb{C}^{n\times n}$ *definieren wir:*

$$\mu(A) := \lim_{h\to 0} \frac{(||I + hA|| - 1)}{h} \ .$$

Offensichtlich hängt μ von der gewählten Norm ab. Bei Bedarf deuten wir dies durch eine zusätzlichen Index an, etwa μ_2 für das von der Euklidischen Norm induzierte Matrixmaß. Das Matrixmaß ist im übrigen kein Maß im Sinne der Maßtheorie; beispielsweise kann es auch negative Werte annehmen. Eine Liste mit Eigenschaften des Matrixmaßes und deren Beweis findet sich in [10]. Wir benötigen lediglich

Lemma 3 *(Eigenschaften des Matrixmaßes) Sei $x \in \mathbb{C}^n, A \in \mathbb{C}^{n\times n}$, $|\cdot|$, $||\cdot||$ und μ wie in Definition 2. Dann gelten*

$$
\begin{array}{ll}
(a) & -||A|| \leq -\mu(-A) \leq \mu(A) \leq ||A|| \\
(b) & -\mu(-A)|x| \leq |Ax| \ \ \textit{und} \ -\mu(A)|x| \leq |Ax| \leq ||A|| \cdot |x| \\
(c) & -\mu(-A) \leq Re\lambda_i(A) \leq \mu(A), \ \forall i \in \{1,2,\ldots,n\} \\
(d) & \mu_2(A) = \max_i \lambda_i(\frac{A+A^*}{2})
\end{array}
$$

Hier bedeuten A^* die zu A adjungierte Matrix, $\lambda_i(A), 1 \leq i \leq n$, die (möglicherweise komplexen) Eigenwerte von A, und $Re\lambda_i$ den Realteil des i-ten Eigenwertes. Mit Hilfe von (a) und (b) sind schärfere Abschätzungen des Wachstumsverhaltens der Lösungen von Differentialgleichungen möglich, als mit der Norm allein (siehe nächstes Lemma). Da μ kleiner als Null sein kann, kann oft sogar asymptotische Stabilität nachgewiesen werden, denn (c) zeigt, daß μ das Spektrum der Matrix A auf einen Streifen in der komplexen Zahlenebene einschränkt, der durch Realteile im Intervall $[-\mu(-A), \mu(A)]$ bestimmt ist. (d) gibt an, wie μ für die Euklidische Norm berechnet oder gegebenenfalls abgeschätzt werden kann. Da $A + A^*$ symmetrisch ist, sind seine Eigenwerte alle reell und oft leichter (z.B. numerisch) zu bestimmen als die Eigenwerte von A selbst. Außerdem zeigt (d), daß die Abschätzung (c) für symmetrische, reelle Matrizen scharf wird, da dann $(A + A^T)/2 = A$ ist und alle Eigenwerte symmetrischer Matrizen reell sind.

Im weiteren werden Vektor- und Matrixnormen einheitlich mit $||\cdot||$ bezeichnet, wobei stets die Euklidische Norm bzw. die von ihr induzierte Operatornorm gemeint ist. Außerdem sei $\mu = \mu_2$.

Lemma 4 *(Lineare Differentialgleichungen und Matrixmaß)*

a) Für jede Lösung der homogenen Gleichung (4) *gilt:*

$$||x_0|| \exp - \int_0^t \mu(-A(t'))dt' \leq ||x(t)|| \leq ||x_0|| \exp \int_0^t \mu(A(t'))dt', \ t \geq 0.$$

b) Jede Fundamentalmatrix $\Psi(t)$ der homogenen Gleichung (4) *erfüllt die Abschätzung:*

$$||\Psi(t)|| \leq ||\Psi(s)|| \exp \int_s^t \mu(A(t'))dt', \ 0 \leq s \leq t$$

c) Jede Lösung der inhomogenen Gleichung (3) *wird beschränkt durch*

$$||x(t)|| \leq ||x_0|| \exp \left[\int_0^t \mu(A(t'))dt'\right] + \int_0^t \exp \left[\int_{t'}^t \mu(A(t''))dt''\right] ||b(t')||dt', \ t \geq 0$$

Beweis: Teil (a) wird in [10] bewiesen. Ein alternativer Beweis ist folgender: Aus (4) folgt für beliebiges $\epsilon > 0$

$$\frac{dx(t)}{dt} = A(t)x(t) = \frac{I + \epsilon A(t) - I}{\epsilon} x(t) \ .$$

Multiplikation von links mit $x^T(t)$ führt auf

$$x^T \frac{dx}{dt} = \frac{1}{2}\frac{d}{dt}||x(t)||^2 = \frac{x^T(I+\epsilon A(t))x - ||x(t)||^2}{\epsilon} \le \frac{||I+\epsilon A(t)||-1}{\epsilon}||x(t)||^2 \; .$$

Da dies für beliebiges $\epsilon > 0$ gilt, so auch im Limes $\epsilon \to 0$. Es folgt

$$\frac{1}{2}\frac{d}{dt}||x(t)||^2 \le \mu(A(t))||x(t)||^2 \; .$$

Integration dieser eindimensionalen Differentialgleichung für $||x(t)||^2$ führt schließlich auf die rechte Ungleichung in (a); die linke kann ähnlich gezeigt werden.
Teil (b) ist eine direkte Folge von (a), da eine Fundamentalmatrix spaltenweise aus Lösungen des homogenen Systems (4) besteht und man in (a) statt Null eine beliebige Anfangszeit s wählen kann.
Zum Beweis von (c) sei $\Psi(t)$ eine Fundamentalmatrix der homogenen Gleichung $\dot{x} = A(t)x$. Dann ist die allgemeine Lösung des inhomogenen Systems (3)

$$x(t) = \Psi(t)\Psi^{-1}(0)x(0) + \int_0^t \Psi(t)\Psi^{-1}(t')b(t')dt' \; .$$

Oder mit $U(t,t') := \Psi(t)\Psi^{-1}(t')$

$$||x(t)|| \le ||U(t,0)|| \cdot ||x(0)|| + \int_0^t ||U(t,t')|| \cdot ||b(t')||dt'$$

Die Matrix $U(t,t')$ ist selbst wieder eine Fundamentalmatrix der Gleichung $\dot{x} = A(t)x$, nun mit der Anfangsbedingung $U(t',t') = I$. Anwendung von (b) auf $U(t,t')$ liefert $||U(t,t')|| \le ||U(t',t')|| \exp\int_{t'}^t \mu(A(t''))dt''$, $0 \le t' \le t$, woraus sich mit $||U(t',t')|| = ||I|| = 1$ die Behauptung (c) ergibt. □

Nachfolgend treten häufig Funktionen der Form

$$g(t,t') := e^{\int_{t'}^t a(t'')dt''}, \quad h_1(t) := g(t,0), \quad h_2(t) := \int_0^t g(t,t')dt' \tag{5}$$

mit bestimmten skalaren Funktionen $a(t)$ auf (z.B. $a(t) = \mu(A(t))$). Der nächste Hilfsatz faßt einige einfache Eigenschaften dieser Funktionen zusammen, von denen wiederholt Gebrauch gemacht wird.

Lemma 5 *Sei $a(t) : \mathbb{R}_+ \to \mathbb{R}$ lokal integrabel, $g(t,t'), h_1(t)$, $0 \le t' \le t$, definiert wie in (5). Weiterhin mögen Konstanten $c \ge 0$ und $a_0 < 0$ existieren, so daß*

$$|\int_0^t a(t')dt' - a_0 t| \le c, \; \forall t \in \mathbb{R}_+ \; . \tag{6}$$

Dann gelten:

$$\begin{array}{llll} (a) & 0 \le h_1(t) \le e^c e^{a_0 t}, \; t \ge 0; & (b) & \lim_{t\to\infty} h_1(t) = 0 \\ (c) & 0 \le g(t,t') \le e^{2c} e^{a_0(t-t')}, \; 0 \le t' \le t; & (d) & \lim_{t\to\infty} g(t,t') = 0, \; t' \in \mathbb{R}_+ \\ (e) & 0 \le \int_0^t g(t,t')dt' \le \frac{e^{2c}}{|a_0|}, \; t \ge 0; & (f) & \lim_{t\to\infty} \int_0^t g(t,t')dt' = \frac{e^{2c}}{|a_0|} \end{array}$$

Falls außerdem $b(t) : \mathbb{R}_+ \to \mathbb{R}$ stetig und beschränkt und der Limes $\overline{\lim}_{t\to\infty} b(t) = \hat{b} < \infty$ ist, gilt weiterhin

$$(g) \; \overline{\lim_{t\to\infty}} \int_0^t g(t,t')b(t')dt' \le \frac{e^{2c}}{|a_0|} \cdot \overline{\lim_{t\to\infty}} b(t) \tag{7}$$

Beweis: Gleichung (6) ist äquivalent zu

$$-c + a_0 t \le \int_0^t a(t')dt' \le c + a_0 t, \quad t \ge 0 \ ,$$

so daß

$$0 \le h_1(t) = e^{\int_0^t a(t')dt'} \le e^{c+a_0 t} \ , \ \forall t \in \mathbb{R}_+ \ ,$$

woraus (a) und (b) folgen, da nach Vorausetzung $a_0 < 0$ ist. (c) und (d) ergeben sich aus:

$$0 \le g(t,t') = e^{\int_{t'}^t a(t'')dt''} = e^{\int_0^t a(t'')dt''} e^{-\int_0^{t'} a(t'')dt''} \le e^{c+a_0 t} e^{c-a_0 t'} = e^{2c} e^{a_0(t-t')} \ .$$

Integration dieser Ungleichungen über t' liefert schließlich (e) und (f). Zum Beweis von (g) beachte man, daß nach Definition des oberen Limes für jedes $\epsilon > 0$ ein T existiert, so daß $b(t)$ für $t > T$ fast überall kleiner als $\hat{b} + \epsilon$ ist. Dann ist

$$\overline{\lim_{t\to\infty}} \int_0^t g(t,t')b(t')dt' \le \overline{\lim_{t\to\infty}} \int_0^T g(t,t') \sup_{0 \le t'' \le T} b(t'')dt' + \overline{\lim_{t\to\infty}} \int_T^t g(t,t')(\hat{b}+\epsilon)dt'$$

$$\le \sup_{0\le t''\le T} b(t'') \cdot \overline{\lim_{t\to\infty}} \int_0^T g(t,t')dt' + (\hat{b}+\epsilon) \overline{\lim_{t\to\infty}} \int_T^t g(t,t')dt' \le \frac{e^{2c}}{|a_0|} \cdot (\hat{b}+\epsilon) \ .$$

Die letzte Ungleichung verwendet (c) und da ϵ beliebig ist, folgt (g) im Limes $\epsilon \to 0$. □

Bemerkung: Aus (6) ergibt sich unmittelbar

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t a(t')dt' = a_0 \ . \tag{8}$$

a_0 ist daher offensichtlich das Langzeitmittel des Integrals der Funktion $a(t)$. Definiert man $\phi(t) := \int_0^t a(t')dt' - a_0 t$, so beschreibt $\phi(t)$ gerade die Abweichungen vom mittleren asymptotischen Verhalten von $\int_0^t a(t')dt'$, das proportional zu $a_0 t$ ist. Diese Abweichungen werden durch Bedingung (6) auf $|\phi(t)| < c$ absolut beschränkt. Identifiziert man weiterhin $a(t)$ mit $\mu(A(t))$ in Lemma 4, zeigen die Abschätzungen (a-g), daß die Bedingung (6) zur asymptotischen Stabilität von Lösungen der homogenen linearen Differentialgleichung führt (Lemma 4(a)) bzw. zur Beschränktheit von Lösungen der inhomogenen Gleichung (4c). In Anschluß an Lemma 3 haben wir bemerkt, daß $\mu(A(t))$ die Realteile der Eigenwerte der Matrix A von oben beschränkt; Bedingung (6) kann also auch dahingehend interpretiert werden, daß man fordert, daß asymptotisch *im Mittel* alle Realteile kleiner als Null sind. Realteile strikt kleiner als Null implizieren asymptotisch exponentielle Stabilität. Die Form von (6) läßt aber auch zu, daß a(t) asymptotisch in t noch zeitabhängig sein und sogar positiv werden kann. Dann können Lösungen vorübergehend auch wieder anwachsen. In den späteren Theoremen wird gezeigt, daß asymptotisch die Einzelelemente des Systems (1) synchronisieren. Wenn $a(t)$ bzw. $\mu(A(t))$ positiv werden kann, bedeutet dies, daß die Konvergenz nicht monoton ist, sondern die Einzeltrajektorien phasenweise auch wieder divergieren können und sich nur asymptotisch Synchronität einstellt.

Der Zusammenhang des Matrixmaßes $\mu(A)$ mit den Eigenwerten der Matrix A legt nahe, an Stelle von $a(t) = \mu(A(t))$ direkt den Realteil des Eigenwerts mit größtem Realteil von A zu verwenden, $a(t) = Re\hat{\lambda}(A(t))$. Die Bedingung (8) fordert dann gerade, daß der sogenannte *größte Lyapunovexponent* $\lim_{t\to\infty} \frac{1}{t} \int^t Re\hat{\lambda}(A(t'))dt'$ (siehe z.B. [14]) kleiner als Null ist, was in der Theorie dynamischer

Systeme ein häufig verwendetes Kriterium für Stabilität ist. Die in dieser Arbeit entwickelte Theorie kann vermutlich auch weitgehend mit Lyapunovexponenten formuliert werden. Dies erfordert an manchen Stellen Modifikationen, z.B. darf in Lemma 4 nicht ohne weiteres $\mu(A(t))$ durch $Re\hat{\lambda}(A(t))$ ersetzt werden. Letzteres wäre korrekt für konstante, symmetrische Matrizen A, ist aber im allgemeinen nicht richtig. Im Fall einer symmetrischen Matrix liefert das Matrixmaß μ_2 im übrigen auch genau den größten Realteil der Eigenwerte, so daß nichts gewonnen wird. Bei allgemeinen Matrixfunktionen können die Lyapunovexponenten schärfer sein, sind aber auch aufwendiger zu berechnen. Häufig ergeben sie sich jedoch, quasi als Nebenprodukt, bei der numerischen Analyse von dynamischen Systemen.

3.2 Statistische Grenzwertsätze

Lemma 6 *Sei $w_{ij},\ i = 1, 2, \cdots,\ j = 1, 2, \cdots$ eine Menge unabhängig und identisch verteilter, reeller Zufallsvariablen mit Erwartungswert $Ew_{11} = 0$ und $Ew_{11}^2 = \sigma^2 < \infty$. Sei $W^{(N)} := (w_{ij})_{1\leq i,j,\leq N}$ und $1^{(N)} := (1, 1, \cdots, 1)^T \in \mathbb{R}^N$. Dann gelten*

(a) *Falls $Ew_{11}^6 < \infty$: $\lim\limits_{N\to\infty} \sup\limits_{1\leq i\leq N} |\frac{1}{N}\sum_{j=1}^N w_{ij}| = 0$ f.s.*

(b) *Falls $Ew_{11}^4 < \infty$: $\lim\limits_{N\to\infty} ||\frac{W^{(N)}}{N}||_2 = 0$ f.s.*

(c) *Falls $Ew_{11}^4 < \infty$: $\lim\limits_{N\to\infty} ||\frac{W^{(N)}}{N} \cdot 1^{(N)}||_2 = \sigma$ f.s.*

Bemerkung: Die Aussage $\lim\limits_{N\to\infty} |\frac{1}{N}\sum_{j=1}^N w_{ij}| = 0$ *f.s.* (vgl. (a)) ist gerade das gewöhnliche starke Gesetz der großen Zahlen, das schon gültig ist, wenn die w_{ij} nur unabhängige, identisch verteilte, reelle Zufallsvariablen mit endlichem Erwartungswert sind (Satz von Kolmogoroff, vgl. Satz 37.2 in [6]). Aussage (a) des Lemmas besagt dann, daß die Konvergenz nicht nur für eine Einzelsumme zutrifft, sondern für abzählbar unendlich viele Summen, indiziert durch i, gleichmäßig in i erfolgt.
Beweis: Für (a) gilt aufgrund der Symmetrie in i:

$$P_N := P\left\{\sup_{1\leq i\leq N} |\frac{1}{N}\sum_{j=1}^N w_{ij}| \geq \epsilon\right\} \leq NP\left\{|\frac{1}{N}\sum_{j=1}^N w_{1j}| \geq \epsilon\right\} \leq \frac{N}{\epsilon^6} E\left[\left(\frac{1}{N}\sum_{j=1}^N w_{1j}\right)^6\right] ,$$

die letzte Abschätzung unter Verwendung der Tschebyscheff-Markoffschen Ungleichung (z.B. [6]). Nun ist

$$E\left[\left(\frac{1}{N}\sum_{j=1}^N w_{1j}\right)^6\right] = \frac{1}{N^6} \sum_{j_1, j_2, \cdots, j_6} E\left[w_{1j_1} w_{1j_2} \cdots w_{1j_6}\right] \ . \tag{9}$$

In der Summe laufen alle Indizes von 1 bis N. Summanden sind stets Null, wenn ein Index ungepaart auftritt. Beiträge ungleich Null haben die Form $(j_1, j_2, j_3, j_4, j_5, j_6) = (i, i, i, i, i, i), (i, i, j, j, j, j), (i, i, i, j, j, j)$ oder (i, i, j, j, k, k) für $1 \leq i, j, k \leq N,\ i, j, k$ paarweise verschieden und beliebige Permutationen hiervon. Terme der ersten Form erfordern $Ew_{11}^6 < \infty$ damit die Summe (9) endlich bleibt. Weil aus $E|w|^p < \infty$ für ein $p \in \mathbb{N}$ folgt, daß auch $E|w|^q < \infty$ für alle $q \leq p$ ist, sind die restlichen Terme ebenfalls endlich. Von der ersten Form gibt es N Summanden, von der zweiten und dritten $O(N^2)$ und von der vierten $O(N^3)$. Asymptotisch in N sind also nur Terme von der vierten Form von Bedeutung, so daß $P_N \leq C/N^2$ für alle N und eine geeignete von N unabhängige Konstante C ist. Nun ist die Summe $\sum_{N=1}^{\infty} P_N \leq \frac{C}{\epsilon^6} \sum_N \frac{1}{N^2} < \infty$, so daß das Borel-Cantelli-Lemma ([6, 9]) die Behauptung (a) impliziert.

Zum Beweis von (b) beachte man, daß $||\frac{W^{(N)}}{N}||_2^2 = \hat{\lambda}(\frac{W^{(N)}W^{(N)^T}}{N^2})$, wobei $\hat{\lambda}(A)$ die Spektralnorm der Matrix A bezeichne. Da die Spektralnorm kleiner oder gleich jeder induzierten Operatornorm ist, können wir sie (z.B.) durch die Zeilensummennorm abschätzen:

$$\left|\left|\frac{W^{(N)}}{N}\right|\right|_2^2 = \hat{\lambda}\left(\left(\frac{\sum_{k=1}^N w_{ik}w_{jk}}{N^2}\right)_{i,j}\right) \leq \sup_{1\leq i\leq N}\left[\sum_{j=1}^N \left|\frac{1}{N^2}\sum_{k=1}^N w_{ik}w_{jk}\right|\right]$$

$$\leq \sup_{1\leq i\leq N}\left|\frac{1}{N^2}\sum_k w_{ik}^2\right| + \sup_{1\leq i\leq N}\left[\sum_{j=1,j\neq i}^N \left|\frac{1}{N^2}\sum_k w_{ik}w_{jk}\right|\right] \tag{10}$$

Wir zeigen, daß beide Ausdrücke im Limes N gegen Unendlich verschwinden. Im ersten Term ist $Ew_{ik}^2 = \sigma^2$ und

$$\sup_{1\leq i\leq N}\left|\frac{1}{N^2}\sum_{k=1}^N w_{ik}^2\right| \leq \sup_{1\leq i\leq N}\left|\frac{1}{N^2}\sum_k (w_{ik}^2-\sigma^2)\right| + \frac{\sigma^2}{N} \leq \frac{1}{N}\sum_{i=1}^N\left|\frac{1}{N}\sum_{k=1}^N (w_{ik}^2-\sigma^2)\right| + \frac{\sigma^2}{N}$$

Definiere $\xi_{ik} := w_{ik}^2 - \sigma^2,\ k,i = 1,2,\cdots;\ z_i^{(N)} := |\frac{1}{N}\sum_{k=1}^N \xi_{ik}|,\ i = 1,2,\cdots$ und $S_N = \frac{1}{N}\sum_{i=1}^N z_i^{(N)}$. Offensichtlich sind die $\xi_{ik},\ k,i = 1,2,\cdots$ und $z_i^{(N)},\ i = 1,2,\cdots$ jeweils unabhängige, identisch verteilte Zufallsvariablen. Die $z_i^{(N)}$ hängen aber von N ab, so daß zum Beweis der Konvergenz von S_N gegen Null das gewöhnliche Gesetz der großen Zahlen nicht angewendet werden kann. Wir verfahren daher wie in Teil (a): Zunächst ist $E(z_1^{(N)})^2 = E\xi_{11}^2/N$ und dies kleiner als unendlich, falls $Ew_{11}^4 < \infty$ ist, was nach Voraussetzung zutrifft. Weiterhin verschwindet $E(z_1^{(N)})^2$ im Limes $N \to \infty$ und somit auch $Ez_1^{(N)}$ (wegen $(Ez)^2 \leq Ez^2$). Sodann ist $S_N = Ez_1^{(N)} + \frac{1}{N}\sum_{i=1}^N (z_i^{(N)} - Ez_1^{(N)})$, wobei, wie eben gezeigt, $Ez_1^{(N)}$ im Limes $N \to \infty$ verschwindet. Für die Summe gilt

$$P(|\frac{1}{N}\sum_{i=1}^N (z_i^{(N)} - Ez_1^{(N)})| > \epsilon) \leq \frac{1}{\epsilon^2}E\left|\frac{1}{N}\sum_{i=1}^N (z_i^{(N)} - Ez_1^{(N)})\right|^2$$

$$= \frac{1}{N\epsilon^2}E\left(z_1^{(N)} - Ez_1^{(N)}\right)^2 \leq \frac{1}{N\epsilon^2}E(z_1^{(N)})^2 = \frac{1}{N^2\epsilon^2}E\xi_{11}^2\ .$$

Dies ist offensichtlich summierbar, so daß S_N und damit auch der erste Term in (10) asymptotisch fast sicher verschwinden. Zur Behandlung des zweiten Terms in (10) definiere $\xi_{ij}^{(N)} = \left|\frac{1}{N}\sum_{k=1}^N w_{ik}w_{jk}\right|,\ i,j = 1,2,\cdots,i \neq j$. Dann ist $E\xi_{12}^2 = \frac{\sigma^4}{N}$, so daß $E\xi_{12}$ und $E\xi_{12}^2$ asymptotisch in N verschwinden. Weiter läßt sich der zeite Term in (10) abschätzen durch:

$$\sup_{1\leq i\leq N}\left[\sum_{j=1,j\neq i}^N\left|\frac{1}{N^2}\sum_k w_{ik}w_{jk}\right|\right] \leq E\xi_{12}^{(N)} + \sup_{1\leq i\leq N}\left|\frac{1}{N}\sum_{j=1,j\neq i}^N\left(\xi_{ij}^{(N)} - E\xi_{12}^{(N)}\right)\right|\ .$$

Der erste Term, $E\xi_{12}^{(N)}$, verschwindet im Limes $N \to \infty$. Im zweiten beachte man, daß das Supremum sich über Zufallsvariablen erstreckt, die zwar nicht unabhängig sind, aber identisch verteilt. Aus Symmetriegründen können wir daher analog zum Beweis von Teil (a) verfahren, das Supremum

eliminieren und unter Verwendung vierter Momente (statt sechster wie in (a)) zeigen, daß auch der zweite Term fast sicher verschwindet. Hierzu ist $Ew_{11}^4 < \infty$ ausreichend. Die Details führen wir nicht weiter aus. Insgesamt konvergieren dann beide Terme in (10) fast sicher gegen Null, woraus Behauptung (b) folgt.
Für (c) schreibe

$$\left\| \frac{W^{(N)}}{N} \cdot 1^{(N)} \right\|_2^2 = \left\| \left(\frac{1}{N} \sum_{j=1}^{N} w_{kj} \right)_{1 \leq k \leq N} \right\|_2^2 = \frac{1}{N} \sum_{k=1}^{N} \left(\frac{1}{\sqrt{N}} \sum_{j=1}^{N} w_{kj} \right)^2 .$$

Wieder hängen die Ausdrücke in der Summe über k von N ab, so daß das herkömmliche Gesetz der großen Zahlen nicht anwendbar ist. Weiteres Umformen liefert

$$\cdots = \frac{1}{N^2} \sum_{k,j=1}^{N} w_{kj}^2 + \frac{2}{N} \sum_{k=1}^{N} \left(\frac{1}{N} \sum_{i<j} w_{ki} w_{kj} \right) =: \frac{1}{N^2} \sum_{k,j=1}^{N} w_{kj}^2 + \frac{2}{N} \sum_{k=1}^{N} \xi_k^{(N)} .$$

Hier geht die Doppelsumme über w_{kj}^2 fast sicher gegen σ^2 aufgrund des gewöhnlichen Gesetzes der großen Zahlen. Vom zweiten Term zeigt man mit den Methoden wie in (a) und (b) unter Verwendung der Momente vierter Ordnung, daß er (f.s.) verschwindet, falls $Ew_{11}^4 < \infty$ ist. Auf die Darstellung der Details verzichten wir wieder. □

Lemma 6 stellt Gesetze der großen Zahl für reellwertige Zufallsvariablen zur Verfügung. Falls in der Differentialgleichung (1) die Responsefunktionen als Diracfunktionen angesetzt werden, genauer falls dort die Funktionen $C(t-s, \omega_{ij})$ von der Form $\omega_{ij}\delta(t-s)$ sind, reduziert sich das System von einer Integrodifferentialgleichung auf eine gewöhnliche Differentialgleichung mit konstanten, aber zufälligen, Koeffizienten ω_{ij}. Mit Hilfe von Lemma 6 kann dann gezeigt werden, daß asymptotisch in der Netzwerkgröße N, die Komponenten der Lösungsfunktion $x_i(t)$ identisch werden. Dies wird in Theorem 8 im nächsten Unterabschnitt präzisiert.

Sollen realistischere Responsefunktionen berücksichtigt werden, müssen die Aussagen des Lemmas 6 erweitert werden. Im einfachsten Fall ist $C(t, \omega_{ij}) = \omega_{ij} g(t)$ mit einer festen Funktion g. Hier ist eine Verallgemeinerung noch nicht nötig. Wenn die $\omega_{ij},\ i,j = 1,2,\cdots$, die Bedingungen des Lemmas 6 erfüllen, gelten die Aussagen (a)-(c) des Lemmas offensichtlich gleichmäßig in t, falls man dort w_{ij} durch $\omega_{ij} g(t)$ ersetzt. Erst wenn die Responsefunktionen sich außer in ihrer Amplitude auch in ihrem individuellen Zeitverlauf voneinander unterscheiden können, treten weitere Probleme auf. Die Responsefunktionen sind dann Zufallsfunktionen oder stochastische Prozesse und keine einfachen, reellwertigen Zufallsvariablen mehr. Gleichmäßige Konvergenz (oder Konvergenz in einem anderen Sinne) ist dann nicht mehr ohne weiteres gewährleistet. Lemma 7 formuliert die für unsere Zwecke notwendigen Verallgemeinerungen von Lemma 6.

Lemma 7 *Sei* $\{w_{ij}(t)\}_{t \in \mathbb{R}_+},\ i = 1,2,\cdots,\ j = 1,2,\cdots$ *eine Familie unabhängiger, identisch verteilter stochastischer Prozesse auf einem gemeinsamen Wahrscheinlichkeitsraum* $(\Omega, \mathcal{F}, P)$ *mit* $Ew_{11}(t) = 0,\ Ew_{11}^2(t) = \sigma^2(t)$ *und* $Ew_{11}(t)^6 < \infty$ *für alle* $t \in \mathbb{R}_+$. *Für jedes* $\omega \in \Omega$ *sei* $w_{11}(t, \omega)$ *global Lipschitzstetig auf* $\mathbb{R}_+$. *Ferner sei* $\{l_{ij}\},\ i = 1,2,\cdots,\ j = 1,2,\cdots$ *eine Menge unabhängig und identisch verteilter, reeller Zufallsvariablen auf demselben Wahrscheinlichkeitsraum* $(\Omega, \mathcal{F}, P)$ *mit* $El_{11}^6 < \infty$, *derart daß für alle* $i = 1,2,\cdots,\ j = 1,2,\cdots$ *und fast alle* $\omega \in \Omega$, $l_{ij}(\omega)$ *Lipschitzkonstante für* $w_{ij}(t, \omega)$ *ist. Definiere* $W^{(N)}(t) := (w_{ij}(t))_{1 \leq i,j, \leq N},\ t \in \mathbb{R}_+$ *und* $1^{(N)} := (1, 1, \cdots, 1)^T \in \mathbb{R}^N$. *Dann gelten für* $0 \leq T < \infty$

(b) $\lim\limits_{N\to\infty} \sup\limits_{0\le t\le T} \left\| \dfrac{W^{(N)}(t)}{N} \right\|_2 = 0 \ f.s.$ **(a)** $\lim\limits_{N\to\infty} \sup\limits_{0\le t\le T} \sup\limits_{1\le i\le N} \left| \dfrac{1}{N} \sum\limits_{j=1}^{N} w_{ij}(t) \right| = 0 \ f.s.$

(c) $\lim\limits_{N\to\infty} \displaystyle\int_0^T \left| \left\| \frac{W^{(N)}(t)}{N} \cdot 1^{(N)} \right\|_2 - \sigma(t) \right| dt = 0 \ f.s.$

Beweis: Zunächst definiere

$$\Phi_1^{(N)}(t) := \left\| \frac{W^{(N)}(t)}{N} \right\|_2, \qquad \Phi_2^{(N)}(t) := \sup_{1\le i\le N} \left| \frac{1}{N} \sum_{j=1}^{N} w_{ij}(t) \right|,$$

$$\Phi_3^{(N)}(t) := \left\| \frac{W^{(N)}(t)}{N} \cdot 1^{(N)} \right\|_2$$

Nach Voraussetzung sind die Bedingungen des Lemmas 6 für beliebiges $t \in \mathbb{R}_+$ erfüllt; (a) und (b) gelten daher punktweise für $t \in \mathbb{R}_+$. Zu zeigen ist, daß die Konvergenz auf $[0,T]$ gleichmäßig erfolgt. Hierzu ist nachzuweisen, daß für fast alle $\omega \in \Omega$ die Folgen $\{\Phi_i^{(N)}(t,\omega)\}_N,\ i = 1,2$ auf $[0,T]$ gleichgradig stetig sind und daher in $C[0,T],\ 0 \le T < \infty$ gleichmäßig konvergieren. Aufgrund der punktweisen Konvergenz von $\Phi_i^{(N)}(t,\omega), i = 1,2$ gegen Null und der Stetigkeit der Grenzfunktion folgt dann, daß die Konvergenz gleichmäßig gegen Null erfolgt, außer vielleicht auf einer Nullmenge in $\mathcal{F}$.

(a)

$$\left(\Phi_1^{(N)}(t) - \Phi_1^{(N)}(s)\right)^2 \le \left\| \frac{W^{(N)}(t) - W^{(N)}(s)}{N} \right\|^2 \le \frac{1}{N^2} \sum_{1\le k,i\le N} (w_{ki}(t) - w_{ki}(s))^2$$

$$\le \frac{1}{N^2} \sum_{1\le k,i\le N} l_{ki}^2 \cdot |t-s|^2 \ \stackrel{N\to\infty}{\longrightarrow} \ El_{11}^2 |t-s|^2 \ \ f.s. \quad \forall s,t \in \mathbb{R}_+ \ .$$

Der Grenzübergang benutzt das gewöhnliche Gesetz der großen Zahlen. Man beachte, daß die Konvergenz unabhängig von der speziellen Wahl von t und s ist. Aufgrund der fast sicheren Konvergenz existiert daher ein N_0, so daß für alle $N > N_0$ und für fast alle $\omega \in \Omega$ die Zahl $2 \cdot \sqrt{El_{11}^2}$ Lipschitzkonstante für $\Phi_1^{(N)}(t,\omega)$ auf ganz $\mathbb{R}_+$ ist. Für $N \le N_0$ sind die $\Phi_1^{(N)}(t,\omega)$ nach der obigen Abschätzung ebenfalls Lipschitzstetig, möglicherweise jedoch für eine größere Konstante als $2 \cdot \sqrt{El_{11}^2}$. Endlich viele gleichmäßig stetige Funktionen sind jedoch immer gleichgradig stetig, so daß die gesamte Menge $\{\Phi_1^{(N)}(t,\omega)\}_{N=1,2,\ldots}$ fast sicher gleichgradig stetig ist, woraus sich nach dem oben Gesagten (a) ergibt.

(b)

$$\left|\Phi_2^{(N)}(t) - \Phi_2^{(N)}(s)\right| \le \sup_{1\le i\le N} \left| \frac{1}{N} \sum_{j=1}^{N} w_{ij}(t) - w_{ij}(s) \right| \le \sup_{1\le i\le N} \frac{1}{N} \sum_{j=1}^{N} |w_{ij}(t) - w_{ij}(s)|$$

$$\le \left(\sup_{1\le i\le N} \frac{1}{N} \sum_{j=1}^{N} l_{ij} \right) |t-s| \le \left(\sup_{1\le i\le N} |\frac{1}{N} \sum_{j=1}^{N} l_{ij} - El_{11}| + El_{11} \right) |t-s|$$

$$\Rightarrow \quad \lim_{N\to\infty} \left|\Phi_2^{(N)}(t) - \Phi_2^{(N)}(s)\right| \le El_{11}|t-s| \ \ f.s. \quad \forall s,t \in \mathbb{R}_+ \ .$$

Die Konvergenz erfolgt hier wegen Lemma 6(a). Der Rest der Argumentation verläuft wie im Fall (a).

(c) Für $\Phi_3^{(N)}(t)$ kann man schreiben

$$\left(\Phi_3^{(N)}(t)\right)^2 = \frac{1}{N^2}\sum_{1\le k,i\le N} w_{ki}^2(t) + \frac{1}{N^2}\sum_k\sum_{i\ne j} w_{ki}(t)w_{kj}(t)$$

Vom ersten Term weist man leicht nach, daß er für $N\to\infty$ gleichmäßig in t fast sicher gegen $\sigma^2(t)$ konvergiert. Der zweite geht wegen Lemma 6c punktweise gegen Null. Vermutlich ist die Konvergenz gegen Null auch gleichmäßig. Dies ist aber schwierig nachzuweisen, weshalb wir nur die schwächere Aussage (c) beweisen. Hierzu approximieren wir die Funktionen $w_{ij}(t,\omega)$ zunächst durch stückweise lineare Näherungen: sei $T>0$ und $h>0$ beliebig aber fest, S eine abzählbare, dichte Teilmenge von $[0,T]$, $T_m=\{t_1,t_2\cdots t_m\}\subset S$ mit $0=t_1<t_2<\cdots<t_m=T$ und $t_{k+1}-t_k\le h,\ k=1,2\cdots,m-1$. Für $\omega\in\Omega,\ t\in[0,T]$ und $i,j=1,2,\cdots$ definiere

$$\widetilde{w}_{ij}(t,\omega) := w_{ij}(t_l,\omega)+(t-t_l)\frac{w_{ij}(t_{l+1},\omega)-w_{ij}(t_l,\omega)}{t_{l+1}-t_l}\ ;\ t_l\le t\le t_{l+1}\ .$$

Die Funktionen $\widetilde{w}_{ij}(t,\omega)$ sind Lipschitzstetig mit denselben Konstanten, $l_{ij}(\omega)$ wie $w_{ij}(t,\omega)$. Weiterhin sind die Differenzen $\Delta_{ij}(t,\omega):=w_{ij}(t,\omega)-\widetilde{w}_{ij}(t,\omega)$ überall betragsmäßig kleiner als $l_{ij}(\omega)\cdot h$. Definiere $\widetilde{\Phi}_3^{(N)}(t) := \left\|\frac{\left(\widetilde{w}_{ij}(t,\omega)\right)_{ij}\cdot 1^{(N)}}{N}\right\|$ und $\widetilde{\sigma}^2(t):=E\widetilde{w}_{ij}^2(t)$. Dann ist $\lim_{N\to\infty}\widetilde{\Phi}_3^{(N)}(t)=\widetilde{\sigma}(t)$ $f.s.$ Die Konvergenz gilt punktweise in t wegen Lemma 6, aber auch gleichmäßig, da $\widetilde{\Phi}_3^{(N)}(t)$ nur von den höchstens abzählbar vielen Punkten in T_m abhängt und auf T_m fast sicher konvergiert. Demnach ist

$$\lim_{N\to\infty}\sup_{0\le t\le T}|\widetilde{\Phi}_3^{(N)}(t)-\widetilde{\sigma}(t)| = 0 f.s\ , \tag{11}$$

so daß für stückweise lineare Funktionen $\widetilde{w}_{ij}(t,\omega)$ Aussage (c) des Lemmas auch in der stärkeren Form der gleichmäßigen Konvergenz analog zu den Teilaussagen (a) und (b) zutrifft. Da h beliebig ist, trifft dies sogar für beliebig genaue Approximationen der ursprünglichen Funktionen $w_{ij}(t,\omega)$ zu. Obwohl im Limes $h\to 0$ die Funktionen $\widetilde{\sigma}(t)$ und $\widetilde{\Phi}_3^{(N)}(t)$ gleichmäßig gegen $\sigma(t)$ beziehungsweise $\Phi_3^{(N)}(t)$ konvergieren, folgt hieraus allerdings nicht, daß auch $\Phi_3^{(N)}(t)$ gleichmäßig fast sicher gegen $\sigma(t)$ konvergiert, denn es könnte im Prinzip für jedes $\omega\in\Omega$ eine Nullmenge $A\subset[0,T]$ geben, auf der sich die Approximationsfehler $\Delta_{ij}(t,\omega)$ zufällig derart akkumulieren, daß sie im Limes $N\to\infty$ nicht verschwinden.

Wir definieren daher $\Delta^{(N)}(t,\omega):=(\Delta_{ij}(t,\omega))_{1\le i,j\le N}$ und betrachten

$$\int_0^T\left\|\frac{\Delta^{(N)}(t)\cdot 1}{N}\right\|^2 dt = \frac{1}{N^2}\sum_{1\le k,i\le N}\int_0^T\Delta_{ki}^2(t)dt + \frac{1}{N^2}\sum_k\sum_{i\ne j}\int_0^T\Delta_{ki}(t)\Delta_{kj}(t)dt \tag{12}$$

Im ersten Term auf der rechten Seite ist $\lim_{N\to\infty}\frac{1}{N^2}\sum_{1\le k,i\le N}\int_0^T\Delta_{ki}^2(t)dt \le El_{11}^2h^2T$ f.s., wegen $\Delta_{ki}^2(t)\le l_{ki}^2h^2,\ \forall t\in[0,T]$. Im zweiten Term definiere $z_k^{(N)} := \frac{1}{N}\sum_{i\ne j}\int_0^T\Delta_{ki}(t)\Delta_{kj}(t)dt$. Die $z_k^{(N)},\ k=1,2\cdots N$ sind unabhängige, identisch verteilte Zufallsvariable. Wir zeigen mit der schon im Beweis von Lemma 6 wiederholt verwendeten Methode, daß $\frac{1}{N}\sum_{k=1}^N z_k^{(N)}$ fast sicher gegen Null geht. Hierzu rechnet man leicht nach, daß $Ez_k^{(N)} = \frac{1}{N}\sum_{1\le i\ne j\le N}E\int_0^T\Delta_{ki}(t)\Delta_{kj}(t)dt = \frac{1}{N}\sum_{1\le i\ne j\le N}\int_0^T E\Delta_{ki}(t)\Delta_{kj}(t)dt = 0$ und ähnlich $E(z_k^{(N)})^2\le c_1(El_{11}^2h^2T)^2<\infty$, sowie $E(z_k^{(N)})^4<$

$c_2(El_{11}^2h^2T)^4 < \infty$ für geeignete, von N unabhängige, Konstanten c_1, c_2 gilt. Mit Hilfe dieser Abschätzungen folgt weiter, daß $E\left|\frac{1}{N}\sum_{k=1}^N z_k^{(N)}\right|^4$ von der Ordnung $\mathcal{O}(1/N^2)$ ist. Da dies summierbar ist, ergibt sich $\lim\limits_{N\to\infty}\frac{1}{N}\sum_{k=1}^N z_k^{(N)} = 0$ $f.s$ aus dem Borel-Cantelli-Lemma. Der Approximationsfehler in Gleichung (12) ist somit beschränkt durch

$$\lim_{N\to\infty}\int_0^T \left\|\frac{\Delta^{(N)}(t)\cdot 1}{N}\right\|^2 dt \le El_{11}^2h^2T \;\; f.s.\,. \tag{13}$$

Schließlich ist

$$\begin{aligned}&\int_0^T |\Phi_3^{(N)}(t)-\sigma(t)|dt \\ &\quad\le \int_0^T |\Phi_3^{(N)}(t)-\tilde{\Phi}_3^{(N)}(t)|dt + \int_0^T |\tilde{\Phi}_3^{(N)}(t)-\tilde{\sigma}(t)|dt + \int_0^T |\tilde{\sigma}(t)-\sigma(t)|dt\end{aligned}$$

für alle $h > 0$. (Die Abhängigkeit der Funktionen $\tilde{\Phi}_3^{(N)}(t)$ und $\tilde{\sigma}(t)$ von h ist nicht explizit gekennzeichnet!). Gleichung (11) zeigt, daß der zweite Term aymptotisch in N fast sicher verschwindet. Das dritte Integral verschwindet für $h \to 0$ da $\tilde{\sigma}(t)$ gleichmäßig gegen σ konvergiert. Der erste Ausdruck kann mit Hilfe der Cauchy-Schwarz'schen Ungleichung weiter abgeschätzt werden durch

$$\int_0^T |\Phi_3^{(N)}(t)-\tilde{\Phi}_3^{(N)}(t)|dt \le \int_0^T \left\|\frac{\Delta^{(N)}(t)\cdot 1}{N}\right\| dt \le \sqrt{T}\cdot\left(\int_0^T \left\|\frac{\Delta^{(N)}(t)\cdot 1}{N}\right\|^2 dt\right)^{1/2}.$$

Wegen (13) ist dies im Limes $N\to\infty$ kleiner oder gleich $T\sqrt{El_{11}^2}h$, was wiederum im Limes $h \to 0$ verschwindet. Insgesamt ergibt sich

$$\lim_{h\to 0}\lim_{N\to\infty}\int_0^t |\Phi_3^{(N)}(t')-\sigma(t')|dt' = 0 \;\; f.s.\,,$$

also die Behauptung (c) des Lemmas. □

4 Grenzwertsätze

Nach den Vorbereitungen des letzten Abschnitts können wir nun Grenzwertsätze für Differentialgleichungen mit zufälligen Kopplungen formulieren.

Theorem 8 *Wir betrachten das Anfangswertproblem*

$$\begin{aligned}\frac{dx_i^\alpha}{dt}(t) &= A^\alpha(t)x_i^\alpha(t) + H^\alpha(x_i^\alpha(t),t) + I^\alpha(t) + \sum_{\beta=1}^P \frac{1}{N^\beta}\sum_{j=1}^{N^\beta} C_{ij}^{\alpha\beta}F^\beta(x_j^\beta(t),t) \\ x_i^\alpha(0) &= x_{i0}^\alpha \in \mathbb{R}^{n^\alpha} \;\; \alpha = 1,2,...P,\; i = 1,2...N_\alpha\end{aligned} \tag{14}$$

Für alle $\alpha = 1,2,\cdots P$ sei

(i) *$N_\alpha(N)$ eine durch N indizierte Folge in $\mathbb{N}$, derart daß $\lim\limits_{N\to\infty}\frac{N_\alpha(N)}{N} = c_\alpha$ für gegebene Konstanten $c_\alpha > 0$,*

(ii) $A^\alpha(t) : \mathbb{R}_+ \to \mathbb{R}^{n^\alpha \times n^\alpha}$ *und* $I^\alpha(t) : \mathbb{R}_+ \to \mathbb{R}^{n^\alpha}$ *beschränkt und stetig,*

(iii) $F^\alpha(x,t) : \mathbb{R}^{n^\alpha} \times \mathbb{R}_+ \to \mathbb{R}^{m^\alpha}$ *beschränkt, Lipschitzstetig in x gleichmäßig in t und stetig in t,*

(iv) $H^\alpha(x,t) : \mathbb{R}^{n^\alpha} \times \mathbb{R}_+ \to \mathbb{R}^{n^\alpha}$ *Lipschitzstetig in x gleichmäßig in t mit Konstante L_h^α und stetig in t,*

(v) $|\int_0^t \mu(A^\alpha(s))ds - a_0^\alpha \cdot t| \le c^\alpha$ *für alle $t \in \mathbb{R}_+$ und geeignete reelle Zahlen $c^\alpha \ge 0$ und $a_0^\alpha < -L_h^\alpha \le 0$.*

(vi) *Ferner seien für alle $1 \le \alpha, \beta \le P$, $1 \le k \le n^\alpha$, $1 \le l \le m^\beta$ die Mengen $\left\{C_{ij}^{kl,\alpha\beta}\right\}_{i,j=1,2,\cdots}$ paarweise unabhängige abzählbar unendliche Mengen unabhängig und identisch verteilter Zufallsvariablen mit $E\left[\left(C_{11}^{kl,\alpha\beta}\right)^6\right] < \infty$ und $\left(\sigma^{kl,\alpha\beta}\right)^2 := E\left[\left(C_{11}^{kl,\alpha\beta}\right)^2\right]$.*

Dann gelten

A1 $\lim\limits_{N\to\infty} \overline{\lim\limits_{t\to\infty}} \sup\limits_{1\le\alpha\le P} \sup\limits_{1\le i,j\le N_\alpha} \left|\left|x_i^\alpha(t) - x_j^\alpha(t)\right|\right| = 0$ *f.s.*

A2
$$\lim_{N\to\infty} \overline{\lim_{t\to\infty}} \sup_{1\le\alpha\le P} \sup_{1\le i\le N_\alpha} \left|\left|\dot{x}_i^\alpha(t) - \left\{A^\alpha(t)x_1^\alpha(t) + H^\alpha(x_1^\alpha(t),t) + I^\alpha(t) + \right.\right.\right.$$
$$\left.\left.\left. + \sum_{\beta=1}^{P} EC_{11}^{\alpha\beta} F^\beta(x_1^\beta(t),t)\right\}\right|\right| = 0 \text{ f.s.}$$

Bemerkung 1: In (14) und A2 sind die $C_{ij}^{\alpha\beta}$ ($C_{11}^{\alpha\beta}$ in A2) bereits $n^\alpha \times m^\beta$-Matrizen mit den Komponenten $C_{ij}^{kl,\alpha\beta}$, $1 \le k \le n^\alpha$, $1 \le l \le m^\beta$.
Bemerkung 2: Die Behauptung A1 besagt, daß sich asymptotisch in N und t alle Elemente jedes Einzelpools identisch verhalten. Elemente verschiedener Pools können sich dabei durchaus voneinander unterscheiden. Man kann daher jeden Pool durch ein einziges repräsentatives Element ersetzen, zum Beispiel das erste, $i = 1$. A2 gibt dann weiter eine Differentialgleichung dafür an, *wie* das asymptotische Verhalten genau aussieht. Bei P untereinander gekoppelten Pools reduziert man so das hochdimensionale, vollständige System (14) auf ein Ersatzsystem aus lediglich P gekoppelten repräsentativen Elementen. Man beachte aber, das jedes dieser Elemente seinerseits eine Differentialgleichung der Ordnung n_α, $\alpha = 1 \cdots P$ ist. Die Komponenten eines Elementes werden selbstverständlich im allgemeinen nicht identisch werden. Es sollte aber klar sein, daß aus A1 und A2 auch die aymptotische Gleichheit derselben Komponenten *verschiedener* Elemente eines Pools folgt; letzteres auf Grund der Abschätzung $\sup_{1\le k\le n} |v_k| \le ||v||$ für ein beliebiges $v \in \mathbb{R}^n$.
Bemerkung 3: Bedingung (i) des Theorems stellt sicher, daß bei mehreren Subsystemen/Pools ($P > 1$) die Anzahl der Elemente pro Pool N_α asymptotisch in N in festen Verhältnissen wächst. Weiterhin stellen die Stetigkeitsannahmen in den Bedingungen (ii)-(iv) (unter anderem) sicher, daß Lösungen des Anfangswertproblems existieren und eindeutig sind. Bedingung (v) ist notwendig, damit die Komponentenfunktionen $x_i^\alpha(t)$ asymptotisch in t konvergieren (vgl. Lemma 4). Bedingung (vi) schließlich stellt Forderungen an die Wahrscheinlichkeitsverteilungen der Kopplungskonstanten und steht in direktem Bezug zu Lemma 6.
Notation: Die erste Bemerkung zeigt, daß die Notation aufgrund der verschiedenen Indizes recht mühsam ist. In der Formulierung des Theorems bezeichnen griechische Indizes unterschiedliche Subsysteme/Pools; Indexe i und j Elemente innerhalb eines Pools, und k,l indizieren Komponenten

eines einzelnen Elementes. Den Beweis führen wir nur für $P = 1$, da dieser Fall schon alles wesentliche enthält; die griechischen Indexe können daher im weiteren weggelassen werden.

Als weitere Bezeichnungen werden verwendet: $x_i^k(t)$, $k = 1 \cdots n$, für die Komponenten eines Elements $x_i(t)$; $x(t) := (x_1(t), \cdots, x_N(t))^T$ für den Vektor aller Elemente; $F(x_i, t) = (f^1(x_i, t), \cdots, f^m(x_i, t))^T = (f^1(x_i^1, \cdots, x_i^n, t), \cdots, f^m(x_i^1, \cdots, x_i^n, t))^T$; $\mathbf{F}(x, t) := (F(x_1, t), \cdots, F(x_N, t))^T$; dasselbe für die Funktion H. β sei eine Schranke für die Norm von F und L_f eine Lipschitzkonstante für F. Weiterhin definieren wir $w_{ij}^{kl} := C_{ij}^{kl} - EC_{11}^{kl}$; $W^{kl} := (w_{ij}^{kl})_{ij}$; $W_{ij} := (w_{ij}^{kl})_{kl}$; $W := (W^{kl})_{kl}$. Die Zufallsmatrizen W^{kl} erfüllen offensichtlich die Voraussetzungen von Lemma 6.

Der Ablauf des Beweis folgt im wesentlichen dem von Geman in [12]. Aufgrund zahlreicher Modifikationen im Hinblick auf die vorgenommenen Verallgemeinerungen führen wir ihn hier trotzdem aus.

Beweis: Wie schon vorab bemerkt setzen wir $P = 1$. Sei dann $\Psi(t)$ ein Fundamentalsystem der Differentialgleichung $\dot{x}(t) = A(t)x(t)$ und $U(t, t') := \Psi(t)\Psi^{-1}(t')$. Variation der Konstanten ergibt als (formale) Lösung von (14) für $i = 1, 2 \cdots N$:

$$x_i(t) = U(t,0)x_i(0) + \int_0^t U(t,t') \left\{ I(t') + H(x_i(t'), t') + \frac{1}{N} \sum_{j=1}^N C_{ij} F(x_j(t'), t') \right\} dt' \ . \tag{15}$$

Definiere

$$u_i(t) := \int_0^t U(t,t') \left\{ I(t') + H(u_i(t'), t') + \frac{1}{N} \sum_{j=1}^N EC_{11} F(x_j(t'), t') \right\} dt' \ , \tag{16}$$

und $u(t) = (u_1(t), \cdots, u_N(t))^T$ und beachte, daß die Lösungen $u_i(t)$, $i = 1, \cdots, N$ aufgrund der Stetigkeitsannahmen (ii-iv) existieren, eindeutig sind und insbesondere alle $u_i(t)$ identisch sind. Definiere weiterhin $a(t) := \mu(A(t))$ und $h_1(t), h_2(t)$ und $g(t, t')$ wie in (5). Dann kann $||U(t, t')||$ gemäß Lemma 5 abgeschätzt werden durch $||U(t, t')|| \leq g(t, t')$, $0 \leq t' \leq t < \infty$.

Der Beweis von A1 geht nun in zwei Schritten vor: zunächst wird gezeigt, daß die Norm $||x(t) - u(t)||$ beschränkt bleibt, wenn N gegen unendlich geht. Dies wiederum wird im zweiten Schritt benutzt um zu zeigen, daß tatsächlich sogar alle Komponenten $x_i(t)$ von $x(t)$ asymptotisch identisch werden.

Es ist

$$\begin{aligned} ||x(t) - u(t)|| &\leq ||x(0)||\, h_1(t) + \int_0^t g(t,t')\, ||H(x(t'), t') - H(u(t'), t')||\, dt' + \\ &+ \int_0^t g(t,t') \left\{ \left|\left| \frac{W}{N} \left(\mathbf{F}(x(t'), t') - \mathbf{F}(x(t'), t') \right) \right|\right| + \left|\left| \frac{W}{N} \mathbf{F}(u(t'), t') \right|\right| \right\} dt' \\ &\leq \ ||x(0)||\, h_1(t) + \int_0^t g(t,t') \left\{ \left(L_h + L_f \left|\left| \tfrac{W}{N} \right|\right| \right) ||x - u(t')|| + \left|\left| \frac{W \cdot \mathbf{F}(u(t'), t')}{N} \right|\right| \right\} dt' \end{aligned} \tag{17}$$

wobei die Dreiecksungleichung und die Lipschitzstetigkeit von H und F benutzt wurden. Die Ausdrücke $\left|\left| \frac{W}{N} \right|\right|$ und $\left|\left| \frac{W}{N} \mathbf{F} \right|\right|$ können weiter abgeschätzt werden

$$\left|\left| \frac{W \cdot \mathbf{F}(u(t), t)}{N} \right|\right| = \left|\left| \left(\frac{1}{N} \sum_{l=1}^m \sum_{j=1}^N w_{ij}^{kl} f^l(u_j(t), t) \right)_{1 \leq i \leq N; 1 \leq k \leq n} \right|\right|$$

$$= \left\| \left(\frac{1}{N} \sum_{l=1}^{m} f^l(u_1(t),t) \sum_{j=1}^{N} w_{ij}^{kl} \right)_{i,k} \right\| \leq \beta mn \sup_{k,l} \left\| \frac{W^{kl} \cdot 1}{N} \right\| \; .$$

Die zweite Gleichheit gilt auf Grund der Identität der $u_i(t),\ i = 1, 2, \cdots$. Analog ergibt sich $\left\|\frac{W}{N}\right\| \leq mn \sup_{k,l} \left\|\frac{W^{kl}}{N}\right\|$. Mit den Definitionen

$$a_N \;:=\; L_h + L_f mn \cdot \sup_{k,l} \left\| \frac{W^{kl}}{N} \right\| \tag{18}$$

$$b_N(t) \;:=\; ||x(0)||\, h_1(t) + \beta mn \cdot \sup_{k,l} \left\| \frac{W^{kl} \cdot 1}{N} \right\| h_2(t) \tag{19}$$

folgt dann

$$||x(t) - u(t)|| \leq b_N(t) + a_N \int_0^t g(t,t')\, ||x(t') - u(t')||\, dt'$$

und die Anwendung des Bellman-Gronwall-Lemmas (Seite 5) hierauf führt weiter auf

$$\begin{aligned} ||x(t) - u(t)|| &\leq b_N(t) + a_N \int_0^t g(t,t') b_N(t') \exp\left[a_N(t-t')\right] dt' \\ &= \; b_N(t) + a_N \int_0^t b_N(t') \exp\left[\int_{t'}^t (\mu(A(t'')) + a_N) dt''\right] dt' \; . \end{aligned} \tag{20}$$

Die Definition von $b_N(t)$, Bedingung (v) des Theorems und Lemma 5a&e implizieren nun, daß $b_N(t)$ beschränkt ist durch

$$b_N(t) \leq ||x(0)||\, e^{c+a_0 t} + \beta mn \cdot \sup_{k,l} \left\| \frac{W^{kl} \cdot 1}{N} \right\| \frac{e^{2c}}{|a_0|} \; , \tag{21}$$

insbesondere existiert der Limes $t \to \infty$, da a_0 nach Annahme kleiner als Null ist. Weiterhin konvergiert a_N aufgrund der Definition (18), Bedingung (vi) und Lemma 6b für $N \to \infty$ fast sicher gegen L_h. Daß heißt, es gibt für jedes $\epsilon > 0$ ein N_0, so daß für all $N > N_0$ gilt $P(|a_N - L_h| \leq \epsilon) = 1$. Die Funktion $\tilde{a}_N(t) := a(t) + a_N = \mu(A(t)) + a_N$ in (20) erfüllt daher und wegen Bedingung (v) für genügend großes N mit Wahrscheinlichkeit Eins die Voraussetzungen des Lemmas 5, so daß das Integral in (20) beschränkt für alle $t \in \mathbb{R}_+$ ist. Insbesondere bleibt es auch im Limes $t \to \infty$ endlich. Es gibt daher eine von N unabhängige (o.B.d.A in t stetige) Funktion $c(t, ||x(0)||)$, derart, daß für genügend großes N fast sicher $||x(t) - u(t)|| \leq c(t, ||x(0)||)$ für alle $t \in \mathbb{R}_+$ erfüllt ist. Die Schranke hängt im allgemeinen von den Anfangsbedingungen ab, Gleichung (20), Lemma 5b&g zeigen jedoch, daß sie so gewählt werden kann, daß der Limes $\lim_{t\to\infty} c(t, ||x_0||) = c_\infty < \infty$ unabhängig von x_0 ist.

Nun betrachten wir eine Komponente $x_i(t) - u_i(t)$ von $x(t) - u(t)$:

$$\begin{aligned} ||x_i(t) - u_i(t)|| \;&\leq\; ||x_i(0)||\, h_1(t) + \int_0^t g(t,t') L_h\, ||x_i(t') - u_i(t')||\, dt' + \\ &+\; \int_0^t g(t,t') \left\{ \left\| \frac{1}{N} \sum_{j=1}^{N} W_{ij} \left(F(x_j(t'),t') - F(u_j(t'),t')\right) \right\| + \right. \\ &+\; \left. \left\| \frac{1}{N} \sum_{j=1}^{N} W_{ij} F(u_j(t'),t') \right\| \right\} dt' \; . \end{aligned}$$

Im letzten Ausdruck ist $\left\|\frac{1}{N}\sum_{j=1}^{N} W_{ij}F(u_j(t'),t')\right\| \leq \beta mn \cdot \sup_{k,l}\left|\frac{1}{N}\sum_{j=1}^{N} w_{ij}^{kl}\right|$ und im dritten Term gilt

$$\left\|\frac{1}{N}\sum_{j=1}^{N} W_{ij}\left(F(x_j(t'),t') - F(u_j(t'),t')\right)\right\| \leq \left\|\frac{W}{N}\left(\mathbf{F}(x(t'),t') - \mathbf{F}(u(t'),t')\right)\right\|$$
$$\leq L_f\left\|\frac{W}{N}\right\| ||x(t)-u(t)|| \leq L_f mn \cdot \sup_{k,l}\left\|\frac{W^{kl}}{N}\right\| c(t,||x_0||) \ ,$$

wobei in der letzten Ungleichung, die Schranke $||x(t)-u(t)|| \leq c(t,||x_0||)$ aus dem ersten Teil des Beweises eingesetzt wurde. Damit ist

$$\begin{aligned}
\sup_{1\leq i\leq N} ||x_i(t)-u_i(t)|| &\leq \sup_i ||x_i(0)||\, h_1(t) + \beta mn \cdot \sup_{k,l}\sup_i\left|\frac{1}{N}\sum_{j=1}^{N} w_{ij}^{kl}\right| h_2(t) \\
&+L_f mn \cdot \sup_{k,l}\left\|\frac{W^{kl}}{N}\right\| \int_0^t g(t,t')c(t',||x_0||)dt' + L_h\int_0^t g(t,t')\sup_i ||x_i(t')-u_i(t')||\, dt' \\
&=: b_N'(t) + L_h\int_0^t g(t,t')\sup_i ||x_i(t')-u_i(t')||\, dt' \\
&\leq b_N'(t) + L_h\int_0^t b_N'(t')\exp\left[\int_{t'}^t (\mu(A(t''))+L_h)dt''\right] dt' \ . \qquad (22)
\end{aligned}$$

Die dritte Zeile definiert hier $b_N'(t)$ und die vierte benutzt die Definition von $g(t,t')$, sowie noch einmal das Bellman-Gronwall-Lemma um $\sup_i ||x_i(t')-u_i(t')||$ auf der rechten Seite der Abschätzung zu eliminieren. Wie schon für $b_N(t)$ zeigt man, daß auch $b_N'(t)$ im Limes $N\to\infty$ fast sicher für alle $t\in\mathbb{R}_+$ bschränkt bleibt und Lemma 5g auf (22) angewendet werden kann. Es folgt

$$\overline{\lim_{t\to\infty}}\sup_{1\leq i\leq N} ||x_i(t)-u_i(t)|| \leq \left\{1+L_h\frac{e^{2c}}{|a_0+L_h|}\right\}\cdot \overline{\lim_{t\to\infty}}\, b_N'(t) \ .$$

Im Ausdruck für $b_N'(t)$ konvergiert nun der erste Term gegen Null, da $h_1(t)$ asymptotisch in t verschwindet. Im zweiten Term von $b_N'(t)$ konvergiert $h_2(t)$ gegen $\exp(2c)/|a_0|$ (Lemma 5f), bleibt also endlich. Andererseits konvergiert $\sup_{k,l}\sup_i\left|\frac{1}{N}\sum_{j=1}^{N} w_{ij}^{kl}\right|$ aufgrund von Lemma 6a für $N\to\infty$ fast sicher gegen Null, so daß der zweite Term von $b_N'(t)$ dann verschwindet. Ebenso verschwindet auch der dritte Term, da $\lim_{N\to\infty}\sup_{k,l}\left\|\frac{W^{kl}}{N}\right\| = 0$ f.s. und außerdem wegen Lemma 5g und der oben angegebenen Eigenschaften von $c(t,x)$ der Limes $\lim_{t\to\infty}\int_0^t g(t,t')c(t',||x_0||)dt' \leq \exp(2c)\cdot c_\infty/|a_0| < \infty$ ist. Also haben wir insgesamt

$$\lim_{N\to\infty}\overline{\lim_{t\to\infty}}\sup_{1\leq i\leq N} ||x_i(t)-u_i(t)|| = 0 \quad f.s. \ ,$$

und weil $||x_i-x_j|| \leq ||x_i-u_i|| + ||x_j-u_j||$ ist, folgt hieraus die Behauptung A1.
Zum Beweis von Behauptung A2 betrachte

$$\begin{aligned}
&\sup_{1\leq i\leq N} ||\dot{x}_i(t) - \{A(t)x_1(t) + H(x_1(t),t) + I(t) + EC_{11}F(x_1(t),t)\}|| \\
&= \sup_i\left\|A(t)(x_i-x_1) + (H(x_i,t)-H(x_1,t)) + \frac{1}{N}\sum_{j=1}^{N} C_{ij}F(x_j,t) - EC_{11}F(x_1,t)\right\|
\end{aligned}$$

$$
\begin{aligned}
&\leq (||A(t)|| + L_h) \sup_i ||x_i - x_1|| + \sup_i \left|\left| \frac{1}{N} \sum_{j=1}^{N} W_{ij} F(x_1(t), t) \right|\right| \\
&\quad + \sup_i \left|\left| \frac{1}{N} \sum_{j=1}^{N} C_{ij} \left(F(x_j(t'), t') - F(x_1(t), t) \right) \right|\right| \\
&\leq \beta m n \sup_{k,l} \sup_i \left| \frac{1}{N} \sum_{j=1}^{N} w_{ij}^{kl} \right| + \left(||A(t)|| + L_h + L_f \sup_i \frac{1}{N} \sum_{j=1}^{N} ||C_{ij}|| \right) \sup_i ||x_i - x_1||
\end{aligned}
$$

Hier konvergiert der erste Term wegen Lemma 6a im Limes $N \to \infty$ fast sicher gegen Null. Im zweiten Term verschwindet $\sup_i ||x_i - x_1||$ aufgrund der gerade bewiesenen Teilaussage A1 des Theorems im Limes $N, t \to \infty$. Der geklammerte Ausdruck bleibt gleichzeitig endlich, da $A(t)$ nach Voraussetzung beschränkt ist und $\sup_i \left(\frac{1}{N} \sum_{j=1}^{N} ||C_{ij}|| \right) \leq E\,||C_{11}|| + \sup_i \left| \frac{1}{N} \sum_{j=1}^{N} (||C_{ij}|| - E\,||C_{11}||) \right| \longrightarrow E\,||C_{11}||$ f.s. durch Anwendung von Lemma 6a auf die Zufallsvariablen $||C_{ij}|| - E\,||C_{11}||$. □

Eine Erweiterung von Theorem 8 im Hinblick auf die Einführung von zufälligen Responsefunktionen ist recht offensichtlich. Wir formulieren nachfolgend eine Variante, in der angenommen wird, daß die Responsefunktionen nur auf einem endlichen Intervall $[0, T]$, $0 \leq T < \infty$, von Null verschieden sind. In Anschluß an Theorem 9 diskutieren wir den Fall exponentiell beschränkter Responsefunktionen.

Theorem 9 *Betrachtet sei das Anfangswertproblem*

$$
\begin{aligned}
\frac{dx_i^\alpha}{dt}(t) &= A^\alpha(t) x_i^\alpha(t) + H^\alpha(x_i^\alpha(t), t) + I^\alpha(t) + \\
&\quad + \sum_{\beta=1}^{P} \frac{1}{N^\beta} \sum_{j=1}^{N^\beta} \int_0^t C_{ij}^{\alpha\beta}(t-s) F^\beta(x_j^\beta(s), s) ds \qquad (23) \\
x_i^\alpha(0) &= x_{i0}^\alpha \in \mathbb{R}^{n^\alpha} \;\; \alpha = 1, 2, \ldots P, \;\; i = 1, 2 \ldots N_\alpha
\end{aligned}
$$

Es seien die Voraussetzungen (i) bis (v) von Theorem 8 erfüllt und an Stelle von (vi):

(vi') *Für alle $1 \leq \alpha, \beta \leq P$, $1 \leq k \leq n^\alpha$, $1 \leq l \leq m^\beta$ seien die Mengen $\left\{ C_{ij}^{kl,\alpha\beta}(t) \right\}_{i,j=1,2,\cdots}$ paarweise unabhängige Familien jeweils unabhängiger und identisch verteilter stochastischer Prozesse auf $[0, T]$, $T < \infty$. $(\Omega, \mathcal{F}, P)$ bezeichne den zugrunde gelegten Wahrscheinlichkeitsraum. Für jedes α, β, k, l sei $E \left| C_{11}^{kl,\alpha\beta}(t) \right|^6 < \infty$ für alle $t \in [0, T]$ und $C_{ij}^{kl,\alpha\beta}(t) = 0$, $i, j = 1, 2 \cdots$ für $t > T$. Weiterhin sollen für alle $\omega \in \Omega$ die $C_{ij}^{kl,\alpha\beta}(t, \omega)$, $i, j = 1, 2 \cdots$ Lipschitzstetig auf $[0, T]$ sein mit Konstanten $l_{ij}^{kl,\alpha\beta}(\omega)$. Die $l_{ij}^{kl,\alpha\beta}(\omega)$, $i, j = 1, 2 \cdots$ seien unabhängig und identisch verteilte Zufallsvariablen auf $(\Omega, \mathcal{F}, P)$ mit $E \left| l_{11}^{kl,\alpha\beta} \right|^6 < \infty$.*

Unter den Voraussetzungen (i)-(v) und (vi') gilt die Behauptung A1 aus Theorem 8 ebenfalls für das System (23) und an die Stelle von A2 tritt

A2' $$\lim_{N \to \infty} \overline{\lim_{t \to \infty}} \sup_{1 \leq \alpha \leq P} \sup_{1 \leq i \leq N_\alpha} \left|\left| \dot{x}_i^\alpha(t) - \left\{ A^\alpha(t) x_1^\alpha(t) + H^\alpha(x_1^\alpha(t), t) + I^\alpha(t) + \right.\right.\right.$$
$$\left.\left.\left. + \sum_{\beta=1}^{P} \int_0^t E C_{11}^{\alpha\beta}(t-s) F^\beta(x_1^\beta(s), s) ds \right\} \right|\right| = 0 \; f.s.$$

Beweis: Der Beweis verläuft im wesentlichen wie der von Theorem 8. Wir werden im folgenden daher nur einige Stellen skizzieren, an denen Unterschiede bestehen. Es reicht den Fall $P = 1$ zu betrachten. Außerdem sei $n = m = 1$; höherdimensionale Subsysteme können wie in Theorem 8 behandelt werden.

Zunächst zeigt man leicht, daß mit $C_{ij}(t),\ i,j = 1,2\cdots$ auch $EC_{11}(t)$ und $w_{ij}(t) := C_{ij}(t) - EC_{11}(t),\ i,j = 1,2\cdots$ Lipschitzstetig sind. Weiterhin ist offensichtlich $E|w_{11}(t)|^k < \infty$ für alle $t \in \mathbb{R}_+$ und $k \leq 6$. Die zentrierten Prozesse $w_{ij}(t)$ erfüllen demnach die Voraussetzungen des Lemmas 7. Aufgrund der Stetigkeit der $w_{ij}(t)$ sind $E|w_{11}(t)|$ und $\sigma(t) := Ew_{11}^2(t)$ integrabel, so daß $\int_0^T E|w_{11}(t)|dt$ und $\int_0^T \sigma(t)dt$ endlich sind.

Wie im Beweis von Theorem 8 kann man eine formale Lösung für $x_i(t),\ i = 1,2\cdots$ angeben und identische Funktionen $u_i(t)$, in denen die Prozesse $C_{ij}(t)$ durch ihren Erwartungswert $EC_{11}(t)$ ersetzt werden (vgl. Gleichungen (15) und (16)). Definiere $v(t) := ||x(t) - u(t)||$. An die Stelle der Abschätzung (17) tritt dann

$$\begin{aligned} v(t) \;\; &\leq \;\; ||x(0)||\, h_1(t) + \beta \int_0^t g(t,t') \int_0^{t'} \left|\left| \frac{W(t'-t'')\cdot 1}{N} \right|\right| dt''dt' + \\ &\quad + L_h \int_0^t g(t,t')v(t')dt' + L_f \int_0^t g(t,t') \int_0^{t'} \left|\left| \frac{W(t'-t'')}{N} \right|\right| v(t'')dt''dt' \end{aligned} \tag{24}$$

Der erste Term auf der rechten Seite verschwindet asymptotisch in t. Der zweite bleibt für alle N und t fast sicher beschränkt wegen Lemma 7c und der Beschränktheit von $\int_0^T \sigma(t)dt$. Im vierten Term beachte man, daß dort $v(t)$ innerhalb des Integrals über t'' steht und daher das Bellman-Gronwall-Lemma nicht direkt angewendet werden kann. Man kann jedoch die Integration über t' und t'' vertauschen. Beachtet man dabei, daß $g(t,t') \leq h_1(t)e^c e^{a_0 t'},\ 0 \leq t' \leq t < \infty$, ist, $\left|\left|\frac{W(t)}{N}\right|\right|$ außerhalb von $[0,T]$ verschwindet und auf $[0,T]$ wegen Lemma 7a asymptotisch in N fast sicher gleichmäßig kleiner als jedes $\epsilon > 0$ wird, kann der vierte Term in (24) für genügend große N mit Wahrscheinlichkeit 1 abgeschätzt werden durch

$$L_f \int_0^t g(t,t') \int_0^{t'} \left|\left| \frac{W(t'-t'')}{N} \right|\right| v(t'')dt''dt' \leq \epsilon \underbrace{L_f \frac{e^c e^{-a_0 T}}{|a_0|}}_{=:a'} \int_0^t h_1(t) e^{-a_0 t''} v(t'')dt'' \; . \tag{25}$$

Definiere $b_N(t)$ als Summe der ersten beiden Terme auf der rechten Seite von (24) und $h_-(t) := \exp\left(-\int_0^{t'} \mu(A(t''))dt''\right)$. Dann ist $g(t,t') = h_1(t)h_-(t')$ im dritten Term von (24) und $v(t)$ kann weiter abgeschätzt werden durch

$$v(t) \leq b_N(t) + h_1(t) \int_0^t \left(L_h h_-(t') + \epsilon a' e^{a_0 t'} \right) v(t')dt' \; .$$

Hier kann nun wieder das Bellman-Gronwall-Lemma angewendet werden, und gezeigt werden, daß für N gegen unendlich, respektive ϵ gegen Null, $v(t) = ||x(t) - u(t)||$ fast sicher beschränkt bleibt durch eine geeignete Funktion $c(t)$, wie schon im ersten Teil des Beweises der Aussage A1 von Theorem 8. Die Abschätzungen im zweiten Teil von Aussage A1 für $\sup_{1\leq i\leq N} ||x_i(t) - u_i(t)||$ und im Beweis von Ausage A2 kann man analog übertragen, was wir hier jedoch nicht weiter ausführen. □

Bemerkung 1: Eine wichtige Klasse von Responsefunktionen, die Theorem 9 nicht umfaßt, sind *exponentiell beschränkte Funktionen.* Diese können zum Beispiel in der Form charakterisiert werden, daß T in Theorem 9 als unendlich angesetzt wird und angenommen wird, daß unabhängige, identisch verteilte Zufallsvariablen $b_{ij} \geq 0$ existieren, sowie ein $\gamma > 0$ derart, daß $|C_{ij}(t)| \leq b_{ij}e^{-\gamma t}$ für alle $i, j = 1, 2, \cdots$ und $t \in \mathbb{R}_+$ gilt. Die im Theorem 9 eingeschlossenen Responsefunktionen sind offensichtlich eine Teilmenge exponentiell beschränkter Funktionen. $EC_{11}(t)$ und $w_{ij}(t) = C_{ij}(t) - EC_{11}(t)$ sind ebenfalls exponentiell beschränkt und Lemma 6 zeigt direkt, daß, wenn $E|b_{11}|^6 < \infty$ ist, auch $\overline{\lim\limits_{N\to\infty}} \left\|\frac{W(t)}{N}\right\| \leq \overline{\lim\limits_{N\to\infty}} 2\left\|\frac{(b_{ij})}{N}\right\| e^{-\gamma t} \leq 2Eb_{11}e^{-\gamma t}$ $(f.s.)$ und $\sup_i |\frac{1}{N}\sum_j w_{ij}(t)| \leq 2Eb_{11}e^{-\gamma t}$ $(f.s.)$ gelten. Daher kann man für jedes $\epsilon > 0$ ein N_0 und ein T_0 findet, so daß die beiden Ausdrücke für alle $t > T_0$ und $N > N_0$ (f.s.) kleiner als ϵ sind. Dasselbe kann man auf $[0, T_0]$ wegen Lemma 7 ebenfalls erreichen (evtl. muß hierzu N_0 vergrößert werden). Die in Lemma 7 (a) und (b) behauptete gleichmäßige Konvergenz, ist also für exponentiell beschränkte Prozesse auf ganz $\mathbb{R}_+$ erfüllt. Weiterhin ist auch $\sigma(t)$ exponentiell beschränkt, so daß $\int_0^\infty \sigma(t)dt$ endlich ist und in Lemma 7c daher ebenfalls T durch Unendlich ersetzt werden kann.

Die oben skizzierte Erweiterung von Lemma 7 auf exponentiell beschränkte Prozesse ist noch nicht ausreichend um auch Theorem 9 entsprechend zu verallgemeinern. Schätzt man in der Gleichung (25) zum Beispiel $\left\|\frac{W(t'-t'')}{N}\right\|$ gleichmäßig durch ein beliebiges $\epsilon > 0$ ab, evaluiert die Integrale in (24) und wendet das Bellmann-Gronwall-Lemma an, so stellt man fest, daß man $v(t) = ||x(t) - u(t)||$ im Limes $t \to \infty$ so nicht beschränken kann. Um dies zu erreichen muß explizit das Verhalten von $\left\|\frac{W(t)}{N}\right\|$ für $t \to \infty$ bekannt sein, z.B. in der Form $\left\|\frac{W(t)}{N}\right\| \leq \epsilon \cdot e^{-\gamma t}$ für alle $t \geq 0$. Allgemeine Bedingungen, wann letzteres gilt, sind unbekannt. Eine Möglichkeit ist, zu fordern, daß nicht die Prozesse $C_{ij}(t)$, sondern statt dessen $C_{ij}(t)e^{+\gamma t}$ die Lipschitzbedingung in Bedingung (vi') von Theorem 9 erfüllen. (Man beachte, daß $|C_{ij}(t)|e^{\gamma t}$ beschränkt durch b_{ij} ist.) Solange $\gamma > |a_0|$ ist, kann man Theorem 9 dann wieder ohne weitere Einschränkungen beweisen. Ist andererseits $\gamma < |a_0|$ zeigt sich bei der Abschätzung der Integrale von der Form (25), daß man Bedingung (v) im Theorem noch etwas abschwächen muß, insofern als $a_0 < -L_h e^{2c}$ zu gelten hat. Ob man hierauf gegebenenfalls verzichten kann, ist nicht klar. Unter der Annahme, daß in Verallgemeinerung von $\overline{\lim\limits_{N\to\infty}} \left\|\frac{W}{\sqrt{N}}\right\| = 2\sigma$ (Geman 1980, [11]) der Limes $\overline{\lim\limits_{N\to\infty}} \int_0^\infty \left\|\frac{W(t)}{\sqrt{N}}\right\| dt$ fast sicher beschränkt ist, kann man den Beweis modifizieren, so daß man die weitere Einschränkung $a_0 < -L_h e^{2c}$ nicht vornehmen muß. Die angegebene Verallgemeinerung von Gemans Theorem für die Norm von Zufallsmatrizen, ist sehr wahrscheinlich richtig, aber unbewiesen.

Bemerkung 2: *Alphafunktionen* mit stochastischen Parametern als Responsefunktionen stellen eine spezielle Wahl exponentiell beschränkter Zufallsfunktionen dar. Alphafunktionen lassen sich als Spezialfall von Lösungen der linearen Differentialgleichung $\dot{C}(t) = AC(t)$, $C(0) = C_0$, für geeignete $n \in \mathbb{N}$, $n < \infty$, $A \in \mathbb{R}^{n\times n}$ und $C_0 \in \mathbb{R}^n$ darstellen. Sind A und/oder C_0 stochastisch erzeugt, so sind die Lösungen dieser Differentialgleichung Zufallsfunktionen. Falls nun der Realteil des größten Eigenwertes der (Zufalls-)Matrix A mit Wahrscheinlichkeit 1 kleiner oder gleich einer Konstanten $-\gamma < 0$ ist, gibt es (ebenfalls mit Wahrscheinlichkeit 1) eine Konstante M, $0 < M < \infty$, die von der speziellen Realisierung von A nicht abhängt, derart daß $||C(t)|| \leq M\,||C_0||\,e^{-\gamma t}$ ist (siehe [5], Seite 204). Da C(t) differenzierbar ist, ist es Lipschitzstetig und wegen $\left\|\dot{C}(t)\right\| \leq ||A||\ \ ||C(t)|| \leq ||A||\,M\,||C_0||\,e^{-\gamma t}$ sind die Funktionen $C(t)e^{\gamma t}$ fast sicher gleichmäßig Lipschitzstetig auf $\mathbb{R}_+$ mit Lipschitzkonstante $l = ||A||\,M\,||C_0||$. Wenn die sechsten Momente der Verteilung von l existieren, erfüllen auf diese Weise generierte (unanhängig und identisch verteilte) Zufallsfunktionen also die in Bemerkung 1 formulierten Bedingungen, und somit ebenfalls Theorem 9.

Bemerkung 3: Indem die Theoreme 8 oder 9 ganze Pools von Zellen auf eine einzige, äquivalente Zelle reduzieren, scheint es, als sei die Möglichkeit der Betrachtung jeglicher Informationsverarbeitungsprozesse in Netzwerken ausgeschlossen und nur die Analyse biophysikalisch-dynamischer Eigenschaften der Modelle möglich. Am Beispiel des *Assoziativspeichers mit endlich vielen Mustern* zeigen wir nun, daß bei geeigneter Wahl der Pools doch auch gewisse Aspekte der Informationsverarbeitung erfaßt werden können (vgl. [13]). Hierzu nehmen wir an, daß in einem Netzwerk aus N Zellen M zufällig generierte, binäre Muster $\xi^\alpha,\ \alpha = 1, \cdots M$, gemäß der inkrementellen Hebb'schen Lernregel gespeichert sein mögen: $C_{ij} = \sum_{\alpha=1}^{M} \xi_i^\alpha \xi_j^\alpha$. Die $\xi_i^\alpha,\ \alpha = 1 \cdots M,\ i = 1, \cdots N$ seien unabhängig und identisch verteilte binäre Zufallsvariablen, mit $P(\xi_1^1 = 1) = p \in [0, 1]$. Auf die genaue Spezifikation der Einzelzellen usw. kommt es im weiteren nicht an.

Neuron i kann in jedem Muster Null oder Eins sein. Insgesamt gibt es 2^M Möglichkeiten, die von der Form x $= (\xi_i^1, \cdots, \xi_i^M) \in \{0, 1\}^M$ sind. Die Menge der N Neuronen läßt sich nach x $\in \{0, 1\}^M$ in Äquivalenzklassen zerlegen, die durch [x], x $\in \{0, 1\}^M$ bezeichnet sein sollen. Jede Äquivalenzklasse kann als Pool im Sinne der Theoreme 8 oder 9 aufgefaßt werden: falls in x $\in \{0, 1\}^M$ insgesamt k Komponenten Eins sind, $|x| = k$, so ist die Wahrscheinlichkeit, daß ein Neuron zur Klasse [x] gehört gleich $c_x = c_k = p^k (1 - p)^{M-k}$. Offensichtlich enthält eine Klasse [x] dann asymptotisch in N mit Wahrscheinlichkeit Eins einen Bruchteil von c_x Zellen, so daß Bedingung (i) der Theoreme (fast sicher) erfüllt ist. Weiterhin sieht man leicht ein, daß sich die Synapsenvektoren je zweier Zellen derselben Klasse [x] nicht voneinander unterscheiden, d.h. $C_{ij} = C_{kj}, j = 1, \cdots N;\ i, k \in [x]$. Hierzu braucht man nur in der Lernregel zu beachten, daß $\xi_i^\alpha = \xi_k^\alpha$ für alle $\alpha = 1, \cdots M$ ist, wenn i und k aus derselben Klasse sind. Weiterhin sind dann sogar alle Synapsen irgendeiner Zelle $i \in [x_1]$ zu allen Zellen einer anderen Klasse $j \in [x_2]$ identisch, nämlich gleich der Anzahl der koinzidierenden Einsen in x_1 und x_2. Die Kopplunsmatrizen zwischen Neuronen der jeweiligen Klassen haben demnach identische Einträge, was als Spezialfall einer entarteten Verteilung in Bedingung (v) von Theorem 8 natürlich enthalten ist.

5 Diskussion

Zusammenfassend haben wir gezeigt, daß sich hochdimensionale modulare neuronale Netze mit stochastischen Kopplungen mathematisch exakt auf niedrigdimensionale Systeme reduzieren lassen. Die niedrigdimensionalen Gleichungen können zur leichteren Analyse des asymptotischen Verhaltens des vollen Systems benutzt werden. Im allgemeinen können die reduzierten Systeme immer noch sehr kompliziert sein, wenn die Einzelelemente höherdimensional sind und eventuell auch mehrere/viele Pools vorhanden sind. Außerdem können niedrigdimensionale Differentialgleichungssysteme bereits sehr kompliziertes Lösungsverhalten zeigen. In der Referenz [26] haben wir das exemplarisch an einem zeitdiskreten Spezialfall von Theorem 8 gezeigt. Das dort betrachtete Netzwerk besteht aus zwei Pools wechselseitig stochastisch gekoppelter Zellen, von denen eine Gruppe exzitatorische und die andere inhibitorische Neuronen enthält. Die reduzierten Gleichungen sind zweidimensional und wurde von Pasemann und Nelle[21] detailliert numerisch untersucht. Es zeigt sich, daß in diesem System sehr vielfältiges zeitliches Verhalten möglich ist, das von einfachen Fixpunkten, über Oszillationen und Quasiperiodizität, bis hin zu deterministischem Chaos reicht. All diese Verhaltensmodi kann man in der hochdimensionalen Variante des Systems in Ref. [26] reproduzieren und Theorem 8 sagt ganz konkret aus, daß man asymptotisch in der Systemgröße und der Zeit auch nur genau dieses

Verhalten bekommt, nicht weniger und nicht mehr. Dies bedeutet, daß die Analyse kleiner neuonaler Netze durchaus sinnvolle Resultate liefern kann, die sich auf Netzwerke realistischer Größenordnung verallgemeinern lassen.

Eine Grundannahme der Betrachtungen dieser Arbeit besteht in der Verwendung gradueller Neurone. Simulationen und heuristische Argumente legen nahe, daß Theoreme analog zu den hier bewiesenen mindestens für stochastische Schwellenneurone in derselben Form gelten (Aussagen A1,A2&A2' der Theoreme 8 und 9). Ein rigoroser Beweis dieser Behauptung steht noch aus. Simulationen des oben bereits angesprochenen chaotischen Systems in Referenz [26] mit stochastischen Schwellenneuronen weisen jedoch asymptotisch dieselbe komplexe Bifurkationsstruktur auf, wie die entsprechenden graduellen Systeme bzw. deren reduzierte Gleichungen. Hierdurch wird die Behauptung zumindest sehr unterstützt. Für deterministische Schwellenneurone in diskreter Zeit (McCulloch-Pitts-Neurone) mit stochastischen Kopplungen hat Amari verwandte Resultate gezeigt.[2] Ob und wann auch kompliziertere Neuronenmodelle, mit gegebenenfalls sogar oszillatorischem oder chaotischen Zeitverhalten bereits in der isolierten Einzelzelle, Mittelungsprozeduren erlauben, und von welcher mathematischen Form deren mittlere Dynamik ist, ist im allgemeinen unbekannt. Hierunter fallen Integrate-and-Fire-Neurone und natürlich die Hodgkin-Huxley-Gleichungen, sowie alle davon abgeleiteten Modelle.

Eine weitere Grundannahme sind stochastische Kopplungen zwischen Zellen. Es gibt experimentelle Hinweise, die nahelegen, daß die Verbindungstopologie kortikaler Kopplungen tatsächlich in einem hohen Maß stochastisch ist.[8, 15] Trotzdem handelt es sich bei dieser Annahme um eine starke Vereinfachung, mindestens insofern funktionelle Korrelationen zwischen Synapsen, wie sie beispielsweise durch Hebb'sches Lernen entstehen können, im wesentlichen außer Betracht gelassen werden. Strukturelle Korrelationen, die etwa auf unterschiedliche Zellklassen oder auf grobe anatomische Verschaltungsstrukturen zurückzuführen sein können, lassen sich dagegen mittels geeigneter Definitionen von Pools erfassen. Da weiterhin nicht alle Zelltypen lernfähig sind, mag die "mikroskopische" Unordnung für Kopplungen zwischen solchen Zellklassen funktionell erhalten bleiben, so daß dann Massenwirkungsargumente anwendbar sein sollten. Dies mag zum Beispiel auf die häufig anzutreffende Vorstellung zutreffen, nach der hemmenden Zellen weitgehend Aufgaben der Aktivitätskontrolle zukommt und weniger detaillierte Informationsverarbeitungsprozesse.

Die grobe Verschaltungsstruktur der betrachteten Zellpools kann in konkreten Anwendungen nahezu beliebig komplex sein und insoweit an ganz unterschiedliche biologische Fragestellungen angepaßt werden. So können Pools etwa verschiedene Zellklassen repräsentieren, oder die Schichtenstruktur des Kortex oder die kolumnäre Struktur, die man in vielen Arealen findet. Ebenso kann ein einzelner Pool aufgefaßt werden als ein ganzes, nicht weiter strukturiertes Areal oder ein Kern, so daß Modelle auf einem gröberen anatomischen Niveau ebenfalls eingeschlossen sind. Zusammen mit den weiteren Freiheiten bei der Wahl der Einzelzelldynamik und den synaptischen Kopplungsfunktionen bieten sich kaum zu überschauende Anwendungsmöglichkeiten.

Auch in Bezug auf künstliche neuronale Netze hat die hier vorgestellte Theorie Relevanz, wenn auch eher indirekte. Klassische Modelle und Anwendungen der Neuroinformatik (Assoziativspeicher, Backpropagation, etc.) bestehen meist nur aus weniger als einigen tausend künstlichen Neuronen. Angesichts der großen Zahl von Zellen bereits in einer einzigen Kolumne eines primären visuellen Areals ($\approx 10^5$) und der relativ einfachen Aufgabe, die sie gemeinsam ausführen und die anscheinend kaum über eine Kantendetektion hinausgeht, erscheint es naheliegend, die Neuronen eines künstlichen neuronalen Netzes bereits als Repräsentanten großer Gruppen von Zellen im Sinne dieser Arbeit zu interpretieren.

Ein Beispiel in dem Theorem 8, 9 oder eine geeignete Variante direkt anwendbar ist, stellt gemäß Bemerkung 3 nach Theorem 9 der Assoziativspeicher dar. Allerdings müssen bei M gespeicherten Mustern bereits 2^M Pools in Erwägung gezogen werden, was in der Praxis eine starke Einschschränkung ist. Das Beispiel zeigt aber, daß man Mittelungsprozeduren offensichtlich *auch* benutzen kann, wenn Aspekte kortikaler Informationsverarbeitung qualitativ untersucht werden sollen. Interessante Fragen in dieser Richtung betreffen zum Beispiel Interaktionen (weniger) gleichzeitig aktivierter neuronaler Assemblies, wie wir sie mit Hilfe von Simulationen großer Netzwerke bereits in einem einzelnen [27] oder verteilt über mehrere Areale [7] untersucht haben.

Danksagung

Diese Arbeit wurde im Rahmen des Schwerpunktprogramms "Physiologie und Theorie neuronaler Netze" von der Deutschen Forschungsgemeinschaft finanziell gefördert (DFG, Pa 268/8-1).

Literatur

[1] Amann, H.: Gewöhnliche Differentialgleichungen. 2.Auflage. deGruyter, Berlin, New York, 1995.

[2] Amari, S.I.; Joshida, K.; Kanatani K.I.: A mathematical foundation for statistical neurodynamics. SIAM J. Appl. Math. 33:95-126, 1977.

[3] Amit,D. J.: Modeling Brain Function. Cambridge University Press, 1988.

[4] Bai, Z.D.; Yin, Y.Q.: Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. The Annals of Probability 21(3):1275–1294, 1993.

[5] Barucha-Reid, A.T.: Random Integral Equations. Academic Press, New York, 1972.

[6] Bauer, H.: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. de Gruyter & Co, Berlin, 1968.

[7] Bibbig, A.; Wennekers, T.; Palm, G.: A Neural Network Model of the Cortico-Hippocampal Interplay and the Representation of Contexts. Behavioural Brain Research 66: 169-175, 1995.

[8] Braitenberg, V.; Schüz, A.: Anatomy of the cortex. Springer, Berlin, 1991

[9] Breiman, L. Probability. Addison-Wesley Publishing Company; Reading, Massachusetts, 1968.

[10] Desoer, C.A.; Vidyasagar, M.: Feedback Systems: Input-Output Properties. Academic Press, New York, 1975.

[11] Geman, S.: A limit theorem for the norm of random matrices. Ann.of Probab. 8:252–261, 1980.

[12] Geman, S.: Almost sure stable oscillations in a large system of randomly coupled equations. SIAM J.Appl.Math. 42:695–703, 1982.

[13] Gerstner, W. and v.Hemmen, J.L.: Associative memory in a network of 'spiking' neurons. Network 3, 139–64, 1993.

[14] Guckenheimer, J.; Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, New York, 1983.

[15] Hellwig, B.; Schüz, A.; and Aertsen, A.: Synapses on axon collaterals of pyramidal cells are spaced at random intervals: a Golgi study in the mouse cerebral cortex. Biol.Cybern.71:1-12,1994.

[16] Hertz, J., Krogh,A., and Palmer R.G.: Introduction to the theory of neural computation. Addison Wesley, 1991.

[17] Horn, D. and Usher, M.: Neural Networks with dynamical thresholds. Phys.Rev.A **40**, 1036–44, 1989.

[18] Hopfield, J.J.: Neural Networks and Physical Systems with Emmergent Collective Computational Properties. PNAS, USA 79, 2554–8, 1982.

[19] Kandel,E. R. ;Schwartz,J. H. : Principles of Neural Science. Elsevier, New York, 1981.

[20] Kaneko, K.: Globally Coupled Chaos Violates the Law of Large Numbers but Not the Central-Limit Theorem. Phys.Rev.Lett. 65:1391ff,1990.

[21] Pasemann, F. and Nelle, E.: Elements of nonconvergent neurodynamics. In: Anderson, S.; Anderson, A; and Ottoson, U.: Dynamical Systems - Theory and Application, World Scientific, Singapore, 1993, pp 167-201.

[22] Pecora, L.M. and Caroll, T.L.: Synchronization in Chaotic Systems, Phys.Rev. Lett. 64:821–824,1990.

[23] Pikovsky, A.S. and Kurths,J.: Do Globally Coupled Maps Really Violate the Law of Large Numbers?, Phys.Rev.Lett. 72:1644–1664,1994.

[24] Pikovsky, A.S.; Rosenblum, M.G.; and Kurths J. : Synchronization in a population of globally coupled oscillators. Europhysics Letters 34(3), 160-170, 1996.

[25] Schillen, T.B.; König, P.: Binding by temporal structure in multiple feature domains of an oscillatory neural network. Biol.Cybern. 70: 397-405, 1994.

[26] Wennekers, T. and Pasemann F.: Synchronous Chaos in High-Dimensional Modular Neural Networks. International Journal of Bifurcation and Chaos 6:2055–2067, 1996.

[27] Wennekers, T., Sommer, F.T., and Palm, G.: Iterative Retrieval in Associative Memories by Threshold Control of Different Neural Models. In: H.J. Herrmann and D.E.Wolf and E.Pöppel (eds.) Supercomputing in Brain Research: From Tomography to Neural Networks. World Scientific Publishing, Singapore, 1995, pp 301-319.

[28] Wilson, H.R. and Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. Biophysical Journal 12,1–24, 1972.