A contraction based, singular perturbation approach to near-decomposability in complex systems

Gabriel D. Bousquet and Jean-Jacques E. Slotine

Nonlinear Systems Laboratory,
Massachusetts Institute of Technology,
Cambridge, MA 02139, USA

Abstract

We revisit the classical concept of near-decomposability in complex systems, introduced by Herbert Simon in his foundational article The Architecture of Complexity, by developing an explicit quantitative analysis based on singular perturbations and nonlinear contraction theory. Complex systems are often modular and hierarchic, and a central question is whether the whole system behaves approximately as the “sum of its parts”, or whether feedbacks between modules modify qualitatively the modules behavior, and perhaps also generate instabilities. We show that, when the individual nonlinear modules are contracting (i.e., forget their initial conditions exponentially), a critical separation of timescales exists between the dynamics of the modules and that of the macro system, below which it behaves approximately as the stable sum of its parts. Our analysis is fully nonlinear and provides explicit conditions and error bounds, thus both quantifying and qualifying existing results on near-decomposability.
The question of modularity and near-decomposability, first popularized by Herbert Simon in his seminal paper *The Architecture of Complexity* [11], has been central to the study of complex systems [1, 15]. In Physics, modularity and hierarchy have long been addressed through separation of timescales approaches [16], QSSA [2, 10] and perturbation theory. With the growing interest in network sciences, in particular systems biology, synthetic biology [5, 6] and neurosciences [3, 14], where the timescale separation between layers can be quite small, quantitative analysis tools for nonlinear near-decomposable systems are needed.

Here we revisit the question of near-decomposability from a quantitative viewpoint by exploiting recent results in singular perturbation theory and nonlinear contraction analysis [4, 8], and provide unifying tools to analyze the robustness and stability of nonlinear near-decomposable systems. We will base our discussion on a canonical example loosely inspired from Simon while referring to the more mathematical development of the Appendix where the theorems are rigorously presented.

Consider the dynamic system representing the evolution of temperatures in a perfectly insulated building. The building is composed of \( n = 2 \) floors, each floor has \( m = 2 \) rooms. For the simplicity of exposition, we assume that all rooms have equal heat capacity. The temperature in each room is denoted \( x_{ij} \) where \( i \) is the floor number and \( j \) is the room number. The thermal conductance between rooms on the same floor is much higher than between rooms on separate floors. Overall, the dynamic equation for the evolution of the room temperatures might be

\[
\begin{align*}
\dot{x}_{11} &= f(x_{12} - x_{11}) + \epsilon g_1(x_{21} - x_{11}) \\
\dot{x}_{12} &= -f(x_{12} - x_{11}) + \epsilon g_2(x_{22} - x_{12}) \\
\dot{x}_{21} &= f(x_{22} - x_{21}) - \epsilon g_1(x_{21} - x_{11}) \\
\dot{x}_{22} &= -f(x_{22} - x_{21}) - \epsilon g_2(x_{22} - x_{12})
\end{align*}
\]

(1)

where \( f, g_1, g_2 \) represent possibly nonlinear conductances. \( f \) captures the coupling between rooms on the same floor while \( g_i \) represent the coupling across floors. When the parameter \( \epsilon \) is small, cross-floor coupling is weak and the system is nearly-decomposable.

The intuition suggests that in case of weak cross-floor coupling, on each floor the individual room temperatures, often referred to as microstates or fast variables, should approximately equalize on a fast timescale. In singular perturbation theory, the fast variables are said to converge to the slow manifold. Later on only the average temperatures, also called
macrostates or slow variables, are needed to approximately predict the system’s behavior.

It must come at no surprise that stability plays a key role in the present study. The analysis is based on the property of contraction (Definition 1), namely that identical systems started with any two different initial conditions tend towards each other exponentially. Many complex systems, be they physical systems, energy networks or information networks, are not stable stricto sensu because of the existence of invariants: in our particular example, the total energy $\sum_{ij} x_{ij}$ is conserved and prevents trajectories from converging to each other. While some of our results may still hold for non-contracting systems (Remark 1), it is often useful to reformulate the problem without invariants. In our example, the change of variables $\delta_i = \frac{1}{2} (x_{i2} - x_{i1})$ and $\Delta = \frac{1}{2} ((x_{22} + x_{21}) - (x_{12} + x_{11}))$ both leverages on the invariant to reduce the system’s dimensionality and transforms system (1) into the standard singular perturbation form (8) (c.f. Remark 5):

$$\begin{align*}
\dot{\delta}_1 &= -f(\delta_1) + \frac{\epsilon}{2} (g_2 (\Delta + \delta_2 - \delta_1) - g_1 (\Delta + \delta_1 - \delta_2)) \\
\dot{\delta}_2 &= -f(\delta_2) + \frac{\epsilon}{2} (g_1 (\Delta + \delta_1 - \delta_2) - g_2 (\Delta + \delta_2 - \delta_1)) \\
\dot{\Delta} &= -\epsilon (g_1 (\Delta + \delta_1 - \delta_2) + g_2 (\Delta + \delta_2 - \delta_1))
\end{align*}$$

(2)

Note that $\delta_i$ and $\Delta$ are the “barycentric” variables for the subsystems and slow system, respectively.

Modular systems such as this one give rise to two questions: is the overall system stable? Does is approximately behave in a modular way, i.e. as the sum of its parts? In our example, the modular behavior is as follows: the temperature differences between rooms on the same floor $\delta_i$ converge on a fast timescale to $O(\epsilon)$. On a slow timescale, a good approximation $\bar{\Delta}$ for the temperature difference between floors evolves following the third line of Equation (2) with $\delta_i = 0$. Proposition 1 provides criteria guaranteeing stability of the approximate system. The main contribution of this paper, Theorem 1 provides nonlinear criteria based on contraction and finite-gain L-stability guaranteeing that the singular perturbation approximation holds while providing explicit error bounds for $\delta_i$ and $\Delta - \bar{\Delta}$. Theorem 1 relies heavily on Lemma 2. Hypothesis 2 provides an explicit estimate for the minimum timescale separation between the micro and macrostates. This can be used recursively to estimate the timescales in a cascade of modular systems (Remark 8).

If $f, g_i$ are passive, i.e. verify $xf(x) > 0$ for all $x \neq 0$, traditional Lyapunov Analysis is sufficient to prove the system’s stability (but not necessarily its approximate dynamics).
Consider however the following situation. Imagine that a small air-conditioning system has
been installed between rooms 12 and 22, with the intent to bring the two rooms to equal
temperature. Assume now that a mistake has been made and that the thermostats of the
two rooms have been switched up. The reversed air-conditioning system can be seen as a
negative conductance and $g_2$ might now take the value $g_2 = -kg_1$ with $k > 0$. The building
is not passive anymore. For small $\epsilon$ the singular perturbation approximation typically holds
but for large $\epsilon$ the approximation breaks down and the system may become unstable.

As a numerical example, take $f = g_1 = x \mapsto x + \frac{1}{2} \sin x$ and $k = 1/2$. Hypothesis $2$
is verified in a constant metric with $\beta_f = 1/2, \beta_g = 1/4, d_f = 3/4, \alpha_{f,x} = \sqrt{2}/4, \alpha_{f,y} =
3/4, d_g = 1, \alpha_{g,x} = \sqrt{2}/2$. Remark $8$ indicates $\epsilon_c = \sqrt{2}/7$. In Figures $1, 2$ and $3$ the
time evolution of the system for random initial conditions in $x_{ij} \in [-5, 5]$ are plotted for
$\epsilon/\epsilon_c = 0.5, 2.5, 5$. For $\epsilon = 0.5\epsilon_c$, the system behaves as expected. For $\epsilon = 2.5\epsilon_c$, there
are several equilibria. This is unforeseen by the singular perturbation approximation, in
accordance with our results which do not guarantee contraction beyond $\epsilon_c$. For $\epsilon = 5\epsilon_c$, the
overall system is often unstable.

The present example is modular in that the fast dynamics $\delta_i$ are weakly coupled with each
other: in Equation $8$, $\tilde{f}$ is factorizable (Remark $5$). It is also important to note (Remarks $6$
and $7$) that where Simon’s describes the influence of the microstates onto the macrostates
and across modules as done in an aggregate way, it has to be understood in the sense of an
aggregation in time $\bar{x}_{ii}$, not an a priori ensemble aggregation of the microstates.

APPENDIX

General robustness results

In the following, we assume that all functions are at least continuous and more largely,
have sufficient smoothness. The state $x$ is in $\mathbb{R}^p$ equipped with a norm $|\cdot|$. We define the
shorthand $\|\cdot\| = \sup_{\mathbb{R}^p} |\cdot|$. The notation $A \prec B$ (resp. $A \preceq B$) indicates that $B - A$ is
positive definite (resp. positive semi-definite).

**Definition 1** (Contracting systems). Consider the system $\dot{x} = f(x, t)$. It is said to be
contracting if all trajectories converge exponentially towards each other $[8]$. A sufficient
condition for contraction is that there exist $\beta > 0$, called the contraction rate, and a metric
\( M(x, t) = \Theta^T \Theta \succ 0 \) such that
\[
F = \dot{\Theta} \Theta^{-1} + \Theta \nabla \Theta^{-1} \preceq -\beta I.
\]

We assume that the condition number of \( \Theta(x, t) \) is bounded and call \( \chi \) its maximum. With a slight abuse of language, we will use the term “metric” for both \( \Theta \) and \( M \).

A system \( \dot{x} = f(x, y, t) \) is partially contracting if it is contracting in \( \Theta(x, t) \) for any \( y(t) \).

We will make extensive use of the following robustness property of contracting systems.

**Lemma 1** (Robustness with bounded disturbance). Consider the two related dynamic systems
\[
\begin{align*}
\dot{x} &= f(x, t) + d(x, t) \\
\dot{x}_0 &= f(x_0, t)
\end{align*}
\]
with \( f \beta, \chi \)-contracting and \( d \) bounded. Let \( R(t) = |x(t) - x_0(t)| \). Then
\[
R(t) \leq \chi R(0) \exp(-\beta t) + \|d\|\chi/\beta
\]

**Proof.** \( R \) can be expressed in its integral form
\[
R = \left| \int dx \right| = \int |dx|
\]
Using the local transformation \( dx = \Theta^{-1} \delta z \) let \( Q(t) = \int |\delta z(t)| \). Using the triangular equality, the following relation is obtained
\[
\|\Theta\|^{-1} Q \leq R \leq \|\Theta^{-1}\|Q.
\]  \hspace{1cm} (3)

The differential dynamics \( \dot{\delta z} = \nabla f dx + \delta d \) induces \( \dot{\delta z} = F \delta z + \Theta \delta d \) and hence
\[
\dot{Q} \leq -\beta Q + \|\Theta d\|.
\]  \hspace{1cm} (4)

From this, \( Q \leq Q(0) \exp(-\beta t) + \|\Theta\|\|d\|/\beta \). Utilizing the left-hand side of Equation (3) on \( Q(0) \) and the right-hand side on \( Q \) brings the result.

**Lemma 2** (Robustness with finite-gain L-stable disturbance). Consider a system
\[
\dot{x} = f_0(x, t) + d(x, t)
\]  \hspace{1cm} (5)
where \( f_0 \) is \( \beta, \chi \)-contracting and \( d(x,t) \) is finite-gain L-stable: \(|d(x,t)| \leq K_0 + K_x|x|\). Assume also that \( K_x \leq \beta/\chi \) and that the system \( \dot{x}_0 = f_0(x_0,t) \) has a forward bounded solution \( \|x_0(t)\| \leq x_{00} \). Then

\[
|x(t) - x_0| \leq |x(0) - x_0(0)| \chi e^{-(\beta - \chi K_x)t} + \frac{K_1 + K_x x_{00}}{\beta - \chi K_x}
\]

**Proof.** The triangular inequality is applied to \( x = (x - x_0) + x_0 \) to get \( |x| \leq R + |x_0| \). With the bound on \( d \) we obtain

\[
|\Theta d| \leq |\Theta| K_x R + |\Theta|(K_1 + K_x x_{00})
\]

The proof is completed by reinjecting this relation into equation 4 and using the right hand site of equation 3 to dominate \( R \)

\[
\dot{Q} + (\beta - K_x \chi)Q \leq |\Theta|(K_1 + K_x x_{00})
\]

in a similar fashion to the previous lemma.

---

**Robustness of singularly perturbed systems**

**Proposition 1.** Assume that Hypothesis 1 below holds. Then the system

\[
\begin{align*}
\dot{x} &= \bar{f}\left(\bar{x}, y, u(\epsilon t)\right) \\
\dot{y} &= \epsilon \bar{g}(y, u(\epsilon t))
\end{align*}
\]

is contracting.

**Hypothesis 1.** The \( y \) dynamics is contracting with metric \( \Theta_y \). The system \( \bar{f}(\bar{x}, y, u(\epsilon t)) \) is partially contracting in \( \bar{x} \) in the metric \( \Theta_x \). \( B = \frac{1}{2}\Theta_x \nabla_y \bar{f} \Theta_y^{-1} \) is uniformly bounded. The symmetric part of the generalized Jacobian for the \( y \) dynamics \( F_y^s \) is bounded below.

**Proof.** The system is hierarchic \[12\]. With the state \((\bar{x}, y)\) it is contracting in the metric \( \Theta = \text{diag}(\epsilon^{1/2} \nu \Theta_x, \Theta_y) \) for sufficiently small \( \nu \). Indeed, the symmetric part of the generalized Jacobian is

\[
\begin{pmatrix}
F_x^s & \epsilon^{1/2} \nu B \\
* & \epsilon F_y^s
\end{pmatrix}
\]

Using Schurr’s complement, it is negative definite iff \( F_x^s - \nu^2 B^T \left(F_y^s\right)^{-1} B < 0 \), which holds for \( \nu \) sufficiently small, since by hypothesis \( B \) and \( \left(F_y^s\right)^{-1} \) are bounded.
Lemma 3. Consider the two systems:

\[ \begin{align*}
\dot{x} &= \hat{f}(\bar{x}, y, u(\epsilon t)) + \epsilon \delta f(\bar{x}, y, u(\epsilon t)) \\
\dot{y} &= \epsilon \bar{g}(y, u(\epsilon t)) + \epsilon \delta g(\bar{x}, y, u(\epsilon t))
\end{align*} \] (8)

and

\[ \dot{\bar{y}} = \epsilon \bar{g}(\bar{y}, u(\epsilon t)) \] (9)

with initial condition \((O(\epsilon), y_0 + \Delta)\) and \(y_0\) respectively. Assume that Hypothesis \(\mathcal{H}^2\) stated below holds.

Then, for all \(t, \bar{x} = O(\epsilon)\). After decay of the transient \(e^{-\epsilon \beta_y t}, y - \bar{y} = \bar{y} = O(\epsilon)\).

Hypothesis 2. \(\hat{f}\) is partially \((\chi_f, \beta_f)\)-contracting with respect to \(\bar{x}\) and \(\bar{x} = 0\) is an equilibrium for all \(y, u(\epsilon t)\). \(\bar{g}(y, u(\epsilon t))\) is \((\chi_g, \beta_g)\)-contracting. The solution of system \(\mathcal{H}^2\) is bounded \(|\bar{y}(t)| \leq \bar{M}\). \(u\) is bounded. The perturbations are bounded in the following way: \(|\delta f| \leq d_f + \alpha_{f,x}|\bar{x}| + \alpha_{f,y}|y|\) and \(|\delta g| \leq d_g + \alpha_{g,x}|\bar{x}|\). Furthermore, \(\epsilon \frac{\chi_f}{\beta_f} (\alpha_{f,x} + \frac{\chi_g}{\beta_g} \alpha_{f,y} \alpha_{g,x}) < 1\).

Proof. There is a maximal solution on \(t \in \mathbb{R}^+\). Let \(M_\bar{x}(t) = \max_{\tau \leq t} |\bar{x}(\tau)|\) and similarly define \(M_y(t)\) and \(M\). 0 is solution of the unperturbed \(\dot{\bar{x}} = \hat{f}(\bar{x}, y, u(\epsilon t))\). The robustness property applied to the two lines of Equation \(\mathcal{H}^2\) gives on \([0, t]\)

\[ \begin{align*}
|\bar{x}(t)| &\leq \chi_f |\bar{x}_0| + \frac{\epsilon \chi_f}{\beta_f} (d_f + \alpha_{f,x} M_\bar{x}(t) + \alpha_{f,y} M_y(t)) \\
|\bar{y}(t)| &\leq \chi_g |\Delta| + \frac{\chi_g}{\beta_g} (d_g + \alpha_{g,x} M\bar{x}(t))
\end{align*} \] (10)

Note that \(M_y(t) \leq \bar{M} + M\bar{y}(t)\). Taking the supremum of the expression and reorganizing we get after some algebra

\[ M_\bar{x}(t) \leq \frac{\chi_f |\bar{x}_0| + \frac{\epsilon \chi_f}{\beta_f} (d_f + \alpha_{f,y} (\bar{M} + \chi_g (|\Delta| + d_g / \beta_g)))}{1 - \epsilon \frac{\chi_f}{\beta_f} (\alpha_{f,x} + \frac{\chi_g}{\beta_g} \alpha_{f,y} \alpha_{g,x})} \] (11)

\(M_\bar{x}(t)\) is increasing and bounded, therefore it converges to a finite limit \(M_\bar{x}\). The boundedness of \(\bar{x}\) induces that of \(\bar{y}\) and \(y\). \(\bar{x}_0 = O(\epsilon)\) implies \(M_\bar{x} = O(\epsilon)\). Finally, the robustness property with bounded disturbances can be applied to the \(y\) dynamics: \(|\bar{y}| \leq \chi_g |\Delta| e^{-\epsilon \beta_y t} + \frac{\chi_g}{\beta_g} M_\bar{x}\) which becomes \(O(\epsilon)\) after the transient.

Lemma 4. Consider the system \(\mathcal{H}^3\) with initial conditions \((\bar{x}_0, y_0)\). Assume that Hypothesis \(\mathcal{H}^3\) holds. Then at \(t = \log(1/\epsilon)/\beta_f\), the system has evolved to \((O(\epsilon), y_0 + O(\epsilon t))\).
Proof. By continuity of \( \bar{g} \) and boundedness of \( u \), there is \( d_{y_0}, \alpha_{y_0} > 0 \) and a neighborhood \( B_{y_0} = \{ y : |y - y_0| \leq c_{y_0} \} \) such that \( |\bar{g}(y, u)| \leq d_{y_0} + \alpha_{y_0}|y - y_0| \) (take for instance \( d_{y_0} = \max_u \bar{g}(y_0, u) \) and \( \alpha_{y_0} = 2 \max_u |\frac{\partial \bar{g}}{\partial y}(y_0, u)| \)).

The robustness relation for \( \tilde{x} \) and integration of the \( y \) dynamics gives:

\[
|\tilde{x}(t)| \leq \chi_f|\tilde{x}_0| + \frac{\epsilon \chi_f}{\beta_f} (d_f + \alpha_{f,x} M_{\tilde{x}}(t) + \alpha_{f,y} M_y(t))
\]

\[
|y(t) - y_0| \leq \epsilon \int_0^t |\bar{g}(y(\tau), u)| + \alpha_{g,x} M_{\tilde{x}}(\tau) d\tau
\]  

(12)

By continuity, \( y \) remains in the neighborhood \( B_{y_0} \) at least for some time. Assume that \( y \) does leaves the neighborhood at some point and let \( t_c \) be the first time when the neighborhood boundary is crossed. Then,

\[
c_{y_0} = \left| \int_0^{t_c} \dot{y} d\tau \right|
\]

\[
\leq \epsilon t_c \left( d_{y_0} + \alpha_{y_0} c_{y_0} + \alpha_{g,x} \frac{\chi_f|\tilde{x}_0| + \epsilon \chi_f/\beta_f (d_f + \alpha_{f,y} (|y_0| + c_{y_0}))}{1 - \epsilon \chi_f \alpha_{f,x}/\beta_f} \right)
\]

\[
= \epsilon t_c (K_1 + O(\epsilon))
\]

where \( K_1 \) is a constant gathering the terms of the second line. Below some critical \( \epsilon_1 \), in the last expression the \( O(\epsilon) \) can be bounded by some other constant \( K_2 = \lim \sup_{\epsilon \to 0} O(\epsilon) \). As a consequence, \( y \) remains in \( B_{y_0} \) at least until \( t_c \geq t_0/\epsilon \) with \( t_0 = c_{y_0}/K \).

We can now bound

\[
|y(t) - y_0| \leq \epsilon t K, \quad t \leq t_0/\epsilon
\]

If \( y \) never leaves the \( B_{y_0} \) neighborhood, the relation also holds. As a consequence, \( y \) can be bounded by a constant quantity for a time arbitrary large \( t \sim 1/\epsilon: M_y(t_0/\epsilon) \leq |y_0| + t_0 K \). Such a bound still holds if \( y \) never leaves \( B_0 \). We can now apply once more the robustness property to the \( \tilde{x} \) dynamics in order to obtain the tighter bound

\[
|\tilde{x}| \leq \chi_f|\tilde{x}_0|e^{-\beta_f t} + O(\epsilon), \quad t \leq t_0/\epsilon
\]

At \( t = \log(1/\epsilon)/\beta_f \), the first terms becomes \( O(\epsilon) \). Which proves the result. Note that for \( \epsilon \) small enough, this time is smaller than \( t_0/\epsilon \).
Theorem 1. Consider the systems (8) and (9). Assume that Hypothesis 2 holds. After a transient \( \left( \frac{1}{\beta_f} + \frac{1}{\beta_g} \right) \log \frac{1}{\epsilon} \), system (8) remains in a \( O(\epsilon) \) neighborhood of \((0, \bar{y})\). An explicit bound for the neighborhood is shown in the proof of Lemma 3.

Proof. This is the direct consequence of subsequently applying the Lemmas 3 and 4. \( \square \)

Remark 1. The contraction property of \( \bar{g} \) is utilized to obtain the bound in \( \tilde{y} \) in Lemma 3. It is sufficient to assume that \( y \) is bounded, perhaps using general passivity properties of the full system, to obtain the bound on \( \tilde{x} \). Note that the boundedness of \( y \) is guaranteed if \( \dot{y} = \bar{g}(y,u) \) is contracting and has a bounded solution.

Remark 2. It is straightforward to obtain explicit bounds in place of the \( O(\epsilon) \).

Remark 3. Hypothesis 2 is not particularly restrictive. In particular it holds for linear systems and finite-gain L-stable systems. In contrast, the hypotheses for Lemma 1 of [4] require in essence \( \delta f \) to be bounded, while we only require it to be finite-gain L-stable in \( \tilde{x} \).

Remark 4. In general, a transformation is required to make the condition \( \tilde{f}(0,y,u) = 0 \) in Hypothesis 2 hold. Consider a system in the form

\[
\begin{align*}
\dot{x} &= f(x,y,u(\epsilon t)) + \epsilon g_x(x,y,t) \\
\dot{y} &= \epsilon g_y(x,y,t)
\end{align*}
\] (13)

with \( f(0,y,u) \) not necessarily 0. Assume however that \( f \) is partially contracting in \( x \) in metric \( \Theta_x \) with rate and condition numbers \( \beta_x, \chi_x \).

Assume that \( f(x,y,u) = 0 \) has a solution \( \bar{x}(y,u) \). The partial contraction of \( f \) ensures that it is unique. It represents the equilibrium value of \( x \) after suppressing the forcing \( \epsilon g_x \) and freezing \( y,u \). After the change of variable \( \tilde{x} = x - \bar{x} \), system (13) is transformed into system (8) with

\[
\begin{align*}
\tilde{f}(\tilde{x},y,u) &= f(\bar{x}(y,u) + \tilde{x},y,u) \\
\delta f &= -\frac{\partial \bar{x}}{\partial y} (\bar{g}_x(\tilde{x},y,u) + \epsilon \delta g) - \frac{\partial \bar{x}}{\partial u} u' \\
\bar{g}(y,u) &= g(y,\bar{x}(y,u),u) \\
\delta g &= \bar{g}_y(\tilde{x} + \bar{x},y,u,t) - \bar{g}(y,u,t)
\end{align*}
\] (14)

Remark 5. Modular systems may take the form \( x = (x_1, \ldots, x_n) \) such that the first line of Equation (8) can be factored into

\[
\dot{x}_i = f_{x_i}(x_i,y,u_i) + \epsilon g_{x_i}(x,y,u)
\] (15)
The key here is that for each subsystem $i$, $\tilde{f}_i$ depends only on the local microstates $x_i$ and can be analyzed separately; $\tilde{f}$ is partially contracting in the metric $\Theta_x = \text{diag}(\Theta_{x_1}, \ldots, \Theta_{x_n})$ with $\beta_x = \min \beta_{x_i}$ and $\chi_x = \max \chi_{x_i}$.

**Remark 6.** The influence of the individual $x_i$ a priori propagates into the macrostates and other modules through their quasi-steady values $\bar{x}_i$ (see lines 1 and 3 of Equation (14)). It might be that the $x_i$ only have a group influence if the coupling is not only weak, as is assumed in this work, but also low rank (equivalently of higher order). The latter property is about the graph connectivity [9], and is quite different from the property of connection strength discussed here.

**Remark 7.** It can be shown that the important aspect of the micro/macrostate separation is that the microstates converge on average to $\bar{x}_i$. The qualitative behavior still holds if $x_i \to \bar{x}_i + \delta x_i$ with the time average $\bar{\delta x}_i$ small on a fast timescale. In other words our analysis in the singular perturbation framework extends to the averaging framework.

**Remark 8.** In Hypothesis 2 it is required that $\epsilon < \frac{\chi_f}{\beta_f} \left( \alpha_{f,x} + \frac{\chi_g}{\beta_g} \alpha_{f,y} \alpha_{g,x} \right) = \epsilon_c$. As the proof shows, this property is paramount in guaranteeing the boundedness of $\bar{x}$. This constraint is related to small-gain stability. It gives a quantitative estimate of the timescale separation necessary for the micro/macrostate separation to hold.

If a modular system is constituted of a cascade of modules, we expect the separation between the timescale $\tau_\mu$ of the smallest system and that of the largest system $\tau_M$, to be $\tau_\mu = (\epsilon_{c_\mu} \ldots \epsilon_{c_M}) \tau_M$. The estimate holds even for nonlinear systems.

[1] Katy Börner, Soma Sanyal, and Alessandro Vespignani. Network science. *Annual Review of Information Science & Technology*, 41:537–607, 2007.

[2] J. R. Bowen, A. Acrivos, and A. K. Oppenheim. Singular perturbation refinement to quasi-steady state approximation in chemical kinetics. *Chemical Engineering Science*, 18(3):177–188, 1963.

[3] Rishidev Chaudhuri, Kenneth Knoblauch, Marie-Alice Gariel, Henry Kennedy, and Xiao-Jing Wang. A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron*, 88(2):419–431, 2015.
[4] Domitilla Del Vecchio and Jean-Jacques E. Slotine. A Contraction Theory Approach to Singularly Perturbed Systems. *IEEE Transactions on Automatic Control*, 58(3):752–757, 2013.

[5] Andras Gyorgy and Domitilla Del Vecchio. Modular Composition of Gene Transcription Networks. *PLOS computational biology*, 10(3), 2014.

[6] Shridhar Jayanthi and Domitilla Del Vecchio. Retroactivity attenuation in bio-molecular systems based on timescale separation. *IEEE Transactions on Automatic Control*, 56(4):748–761, 2011.

[7] Hassan K Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2002.

[8] Winfried Lohmiller and Jean-Jacques E. Slotine. On Contraction Analysis for Nonlinear Systems. *Automatica*, 34(6):683–696, 1999.

[9] Mark E. J. Newmann. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[10] Lee A. Segel and Marshall Slemrod. The quasi-steady-state assumption: a case study in perturbation. *SIAM Review*, 31(3):446–477, 1989.

[11] Herbert A Simon. The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106(6):467–482, 1962.

[12] Jean-Jacques E. Slotine. Modular Stability Tools for Distributed Computation and Control. *International Journal of Adaptive Control and Signal Processing*, 17(6):397–416, aug 2003.

[13] Jean-Jacques E. Slotine and Weiping Li. *Applied nonlinear control*. Prentice hall New Jersey, 1991.

[14] Jean-jacques E. Slotine and W. Lohmiller. Modularity, evolution, and the binding problem: a view from stability theory. *Neural Networks*, 14(2):137–145, 2001.

[15] Olaf Sporns and Richard F Betzel. Modular Brain Networks. *Annual Review of Psychology*, 67(1), 2015.

[16] Steven H Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press, 2014.
FIG. 1.
FIG. 2.
\[ \epsilon = 5.0 \epsilon_c \]

**FIG. 3.**