ANALYSIS OF THE CSP REDUCTION METHOD FOR CHEMICAL KINETICS

Antonios Zagaris,[1] Hans G. Kaper,[2] Tasso J. Kaper[1]

[1] Department of Mathematics and Center for BioDynamics
Boston University, Boston, Massachusetts, USA

[2] Mathematics and Computer Science Division,
Argonne National Laboratory, Argonne, Illinois, USA

Abstract. This article is concerned with the asymptotic accuracy of the Computational Singular Perturbation (CSP) method developed by Lam and Goussis to reduce the dimensionality of a system of chemical kinetics equations. The method exploits the presence of disparate time scales to model the dynamics by an evolution equation on a lower-dimensional slow manifold. In this article it is shown that the successive applications of the CSP algorithm generate, order by order, the asymptotic expansion of a slow manifold. The results are illustrated on the Michaelis–Menten–Henri equations of enzyme kinetics.

1 Introduction and Summary of Results

Reduction methods decrease the size and complexity of systems of kinetic equations. They are effective when a small number of variables can be singled out as evolving on a "slow manifold" and the remaining (fast) variables somehow follow from the slow variables. In such cases, the system of kinetic equations can be reduced to a much smaller system for the evolution of only the slow variables, and the fast variables can be determined simply by table look-ups or by direct computation. Over the years, a large number of reduction methods have been proposed and implemented in computer codes; references can be found in our earlier article [11], and additional references are [1, 6, 20].

The focus of Ref. [11] was on the Intrinsic Low-Dimensional Manifold (ILDM) method due to Maas and Pope [17] and an iterative method proposed by Fraser [5] and further developed by Roussel and Fraser [25]. In this article, the focus is on the Computational Singular Perturbation (CSP) method

1

developed by Lam and Goussis [7, 8, 12, 13, 14, 15, 16, 18, 19, 26].

A chemical kinetic equation is an ordinary differential equation (ODE),

$$\frac{dx}{dt} = g(x), \tag{1.1}$$

for a vector x of species concentrations; g is a smooth vector field, and t is time. Reduction methods are effective when the variables fall into two classes, fast and slow, as is the case when the Jacobian of the vector field has a spectral gap. For the analysis, it is convenient to identify the spectral gap with the inverse of a small parameter ε, but we emphasize that this restriction is not necessary for the applicability of the CSP method. The characteristic time scales for the fast and slow species are given by the "fast" time t and the "slow" time $\tau = \varepsilon t$, respectively. We assume that the entries of x are ordered in such a way that the first m components evolve on the slow time scale and the remaining n components on the fast time scale. Then the vector field g has the form

$$g = \left(\begin{array}{c} \varepsilon g_1 \\ g_2 \end{array} \right) = \left(\begin{array}{cc} I_m & 0 \\ 0 & I_n \end{array} \right) \left(\begin{array}{c} \varepsilon g_1 \\ g_2 \end{array} \right), \tag{1.2}$$

where I_m and I_n are the identity matrices in $\mathbf{R}^m$ and $\mathbf{R}^n$, respectively, and the system (1.1) is a *fast–slow system* of ODEs. Both g_1 and g_2 may depend on ε, but the entries of these vectors as well as their partial derivatives are all $\mathcal{O}(1)$ as $\varepsilon \downarrow 0$, uniformly in x.

Geometric singular perturbation theory (GSPT) [4, 10] provides a natural framework for the analysis of fast–slow systems of ODEs. If such a system has a slow manifold, $\mathcal{M}_0$, in the limit as $\varepsilon \downarrow 0$ and this manifold is normally hyperbolic, then GSPT identifies a (usually nonunique) slow manifold $\mathcal{M}_\varepsilon$ for ε sufficiently small. GSPT also gives a complete geometric and analytical description of all solutions near $\mathcal{M}_\varepsilon$, including how trajectories approach $\mathcal{M}_\varepsilon$. The goal of any reduction method is to find $\mathcal{M}_\varepsilon$, if it exists.

Typically, the vector field g is written in a form suggested by chemical kinetics, namely, as a weighted sum of the stoichiometric vectors, the weights being the associated reaction rates. But this representation is in no way unique. In fact, Eq. (1.2) shows an equivalent representation of g as a weighted sum of the standard basis vectors of $\mathbf{R}^{m+n}$, the weights being the coordinates $\varepsilon g_1, \dots , \varepsilon g_m, g_{m+1}, \dots , g_{m+n}$. The objective of the CSP method is to express g in yet another basis, one that is tuned to the dynamics of the system, where the fast and slow coordinates (amplitudes) evolve independently of each other. The CSP method achieves this objective constructively by successive approximation. Starting with a more or less arbitrary initial basis, one derives the evolution equations for the fast and slow amplitudes and

updates the basis iteratively in such a way that the evolution equations for the updated fast and slow amplitudes decouple to increasingly higher order in the small parameter ε. Each iteration consists of two steps. The first step deals with the dependence of the fast amplitudes on the slow amplitudes, the second step with the dependence of the slow amplitudes on the fast amplitudes.

After each iteration, one identifies the *CSP manifold* (CSPM) as the locus of points where the then-current fast amplitudes vanish. The CSPM is an approximation to the slow manifold $\mathcal{M}_\varepsilon$. The question is: How good is the approximation? In this paper, we analyze the general class of fast-slow systems of ODEs (1.1)–(1.2) and show (Theorem 3.1) that the CSP method generates term by term the asymptotic expansion of the slow manifold $\mathcal{M}_\varepsilon$. After q iterations ($q = 0, 1, 2, \ldots$), the asymptotic expansions of the CSPM and $\mathcal{M}_\varepsilon$ agree up to and including terms of $\mathcal{O}(\varepsilon^q)$; they differ in general at $\mathcal{O}(\varepsilon^{q+1})$. Also, the qth application of the CSP algorithm leaves the terms at $\mathcal{O}(1)$ through $\mathcal{O}(\varepsilon^{q-1})$ invariant. (This observation is important because the lower-order terms have already been determined correctly in the preceding applications.) We illustrate Theorem 3.1 with an example from the Michaelis-Menten-Henri mechanism of enzyme kinetics [3, 9, 22, 23, 24]. Similar results (for $q = 1, 2$) have been obtained by Valorani, Goussis, and Najm [27] for a model equation due to Davis and Skodje [2].

Our proof proceeds via an intermediate result for a one-step CSP method. The one-step CSP method is the same as the full two-step CSP method but involves only the first step. It yields a sequence of slow manifolds, just like the full CSP method, whose asymptotic behavior as $\varepsilon \downarrow 0$ can be compared with that of the slow manifold $\mathcal{M}_\varepsilon$. The result (Theorem 4.1) is that q applications of the one-step CSP algorithm yield an approximate slow manifold that agrees asymptotically with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^q)$. In other words, the one-step CSP method is as accurate as the full CSP method; and to prove the main result for the full CSP method, one needs only to show that the second step does not affect the lower-order terms in the asymptotic expansion of the CSPM. Although the second step of the CSP method does not play a role in the approximation of $\mathcal{M}_\varepsilon$, it does play a constructive role in the approximation of the dynamics near $\mathcal{M}_\varepsilon$, as we shall demonstrate in the special case of the Michaelis–Menten–Henri equations.

In [11], we showed that the ILDM method yields an approximate slow manifold that is asymptotically accurate up to and including terms of $\mathcal{O}(\varepsilon)$, with an error of $\mathcal{O}(\varepsilon^2)$ proportional to the curvature of $\mathcal{M}_0$. The CSP method, on the other hand, can generate an approximate slow manifold that is asymptotically accurate up to any order. The difference can be traced to two facts, namely, the choice of the fundamental operator governing the dynamics of

the system and the retention of the variation of the Jacobian over the manifold $\mathcal{M}_0$. While the ILDM method is designed to transform the Jacobian of the vector field into triangular form (and often also into diagonal form), the CSP method is an algorithm to diagonalize the (nonlinear) Lie bracket involving the vector field to successively higher orders in ε. The Jacobian is a linear approximation, so the ILDM method never gets beyond a linear approximation. The variation of the Jacobian over $\mathcal{M}_0$ introduces an extra term in the Lie bracket. By retaining it, the CSP method preserves the nonlinear character of the operator governing the dynamics of the system.

This article is organized as follows. In Section 2, we recall the Fenichel theory of GSPT and give the asymptotic expansion of the slow manifold $\mathcal{M}_\varepsilon$. In Section 3, we describe the full CSP method for fast–slow systems and state Theorem 3.1. The one-step CSP method is introduced in Section 4. The approximation result for the slow manifold is given in Theorem 4.1; its proof occupies most of Section 4 and uses two lemmas that are given in the Appendix. In Section 5, we return to the full CSP method and prove Theorem 3.1. In Section 6, we illustrate the CSP method and the results of this paper on a planar system of equations for the Michaelis–Menten–Henri mechanism of enzyme kinetics. Section 7 is devoted to a discussion of the relation between the CSP and ILDM methods.

2 Fast-Slow Systems of ODEs

Collecting the slow variables in a single (column) vector y and the fast variables in a (column) vector z, we rewrite Eqs. (1.1)–(1.2) as a fast-slow system,

$$
\begin{aligned}
y' &= \varepsilon g_1(y, z, \varepsilon), & (2.1)\\
z' &= g_2(y, z, \varepsilon). & (2.2)
\end{aligned}
$$

(A prime $'$ denotes differentiation with respect to t.) The long-term dynamics of this system are more naturally studied on the time scale of the slow variable $\tau = \varepsilon t$, where the system of Eqs. (2.1)–(2.2) assumes the form

$$
\begin{aligned}
\dot{y} &= g_1(y, z, \varepsilon), & (2.3)\\
\varepsilon \dot{z} &= g_2(y, z, \varepsilon). & (2.4)
\end{aligned}
$$

(A dot $\cdot$ denotes differentiation with respect to τ.)

In the limit $\varepsilon \downarrow 0$, Eq. (2.4) reduces formally to the algebraic equation $g_2(y, z, 0) = 0$. We assume that there exists a compact domain $K \in \mathbf{R}^m$ and

4

a smooth single-valued function h_0 on K such that

$$g_2(y, h_0(y), 0) = 0, \quad y \in K. \tag{2.5}$$

Then the long-time dynamics of the system (2.1)–(2.2) are confined to the *reduced slow manifold* $\mathcal{M}_0$,

$$\mathcal{M}_0 = \{(y, z) \in \mathbf{R}^{m+n} : z = h_0(y), \ y \in K\}. \tag{2.6}$$

We assume, furthermore, that the real parts of the eigenvalues of the matrix $D_z g_2(y, h_0(y), 0)$ are all negative, so $\mathcal{M}_0$ is asymptotically stable. Then the Fenichel theory [4], which applies more generally to normally hyperbolic invariant manifolds, guarantees that $\mathcal{M}_0$ persists as a slow manifold, so for all sufficiently small ε there exists a *slow manifold, $\mathcal{M}_\varepsilon$* , that is invariant under the dynamics of the system of Eqs. (2.1)–(2.2). Moreover, $\mathcal{M}_\varepsilon$ has the same dimension as $\mathcal{M}_0$ and lies near $\mathcal{M}_0$, all nearby solutions relax exponentially fast to $\mathcal{M}_\varepsilon$, and the long-term dynamics of the system (2.1)–(2.2) are governed by an equation on $\mathcal{M}_\varepsilon$. The manifold $\mathcal{M}_\varepsilon$ is not unique; typically, there is a family of slow manifolds, all exponentially close ($\mathcal{O}(\mathrm{e}^{-c/\varepsilon})$ for some $c > 0$). The following theorem is essentially a restatement of [10, Theorem 2].

Theorem 2.1 *For all sufficiently small ε, there is a function h_ε such that the graph*

$$\mathcal{M}_\varepsilon = \{(y, z) : z = h_\varepsilon(y), \ y \in K\} \tag{2.7}$$

is locally invariant under the dynamics of Eqs. (2.1)–(2.2). The function h_ε admits an asymptotic expansion as $\varepsilon \downarrow 0$,

$$h_\varepsilon(y) = h_0(y) + \varepsilon h_1(y) + \varepsilon^2 h_2(y) + \cdots, \tag{2.8}$$

and $h_\varepsilon \in C^r(K)$ for any finite r. The long-term dynamics of the system of Eqs. (2.1)–(2.2) are governed by the equation

$$\dot{y} = g_1(y, h_\varepsilon(y), \varepsilon) \tag{2.9}$$

on $\mathcal{M}_\varepsilon$, where $\dot{} = d/d\tau$ with $\tau = \varepsilon t$.

The coefficients $h_1, h_2, \ldots$ are found from the *invariance equation*,

$$g_2(y, h_\varepsilon(y), \varepsilon) - \varepsilon(Dh_\varepsilon)(y)g_1(y, h_\varepsilon(y), \varepsilon) = 0, \quad y \in K, \tag{2.10}$$

in the following manner. (The invariance equation follows immediately from the chain rule, $z' = Dh_\varepsilon(y)y'$, and Eqs. (2.1)–(2.2).) Each of the functions $g_1(\cdot, h_\varepsilon, \varepsilon)$ and $g_2(\cdot, h_\varepsilon, \varepsilon)$ admits a Taylor expansion near $\varepsilon = 0$,

$$g_1(\cdot, h_\varepsilon, \varepsilon) = \sum_{q=0}^{\infty} g_{1,q}\varepsilon^q, \quad g_2(\cdot, h_\varepsilon, \varepsilon) = \sum_{q=0}^{\infty} g_{2,q}\varepsilon^q, \tag{2.11}$$

5

with coefficients

$$
g_{1,q} = \sum_{k=0}^{q-1} \sum_{j=1}^{q-k} \frac{1}{k!j!} (D_z^j D_\varepsilon^k g_1)_0 \sum_{|i|=q-k} (h_{i_1}, ..., h_{i_j}) + \frac{1}{q!} (D_\varepsilon^q g_1)_0, \quad (2.12)
$$

$$
g_{2,q} = \sum_{k=0}^{q-1} \sum_{j=1}^{q-k} \frac{1}{k!j!} (D_z^j D_\varepsilon^k g_2)_0 \sum_{|i|=q-k} (h_{i_1}, ..., h_{i_j}) + \frac{1}{q!} (D_\varepsilon^q g_2)_0. \quad (2.13)
$$

The notation $(\,\cdot\,)_0$ indicates that the quantity inside the parentheses is evaluated on $\mathcal{M}_0$—that is, at $(y, h_0(y), 0)$. Note that $(D_z^j D_\varepsilon^k g)$ is a multilinear operator, which maps a j-form to a vector. The inner sum in Eqs. (2.12) and (2.13) is taken over all multiindices $i = (i_1, \dots, i_j)$ of j positive integers i_1 through i_j subject to the constraint $|i| = i_1 + \cdots + i_j = q - k$. The expressions (2.12) and (2.13) hold for all q if it is understood that a sum is empty whenever its lower bound exceeds its upper bound. Substituting the expansions (2.12) and (2.13) into the invariance equation, Eq. (2.10), and setting the coefficient of ε^q equal to zero, we obtain an infinite set of equations,

$$
g_{2,q} - \sum_{\ell=0}^{q-1} (Dh_\ell) g_{1,q-1-\ell} = 0, \quad q = 0, 1, \dots . \quad (2.14)
$$

The first few equations are

$$
g_{2,0} = 0, \tag{2.15}
$$
$$
(D_z g_2)_0 h_1 + (D_\varepsilon g_2)_0 - (Dh_0) g_{1,0} = 0, \tag{2.16}
$$
$$
(D_z g_2)_0 h_2 + \tfrac{1}{2} (D_z^2 g_2)_0 (h_1, h_1) + (D_z D_\varepsilon g_2)_0 h_1 + \tfrac{1}{2} (D_\varepsilon^2 g_2)_0
$$
$$
- (Dh_1) g_{1,0} - (Dh_0) ((D_z g_1)_0 h_1 - (D_\varepsilon g_1)_0) = 0. \tag{2.17}
$$

Equation (2.15) is satisfied identically, Eq. (2.16) yields the coefficient h_1, Eq. (2.17) the coefficient h_2, and so on.

Remark. The assumption that the chemical species can be divided into fast and slow species, as in Eqs. (2.1)–(2.2), is made for convenience. Our analysis can also be applied to general chemical systems where each species may be involved in both fast and slow reactions and for which there is a slow manifold.

3 The CSP Method for Fast–Slow Systems

In Eq. (1.2), the vector field g is represented in terms of the standard orthonormal basis. It is useful to examine the representation of g in terms of other bases, especially bases whose entries depend on x.

Let A be an $(m+n) \times (m+n)$ matrix whose entries may depend on x and whose columns form a basis for the space $\mathbf{R}^{m+n}$ for each x. The vector field g may be expressed in terms of this (variable) basis A as

$$g = Af, \tag{3.1}$$

where f is the vector of the coordinates (amplitudes) of g. (When the columns of A are the stoichiometric vectors, the amplitudes are the reaction rates.) Since A is invertible, $B = A^{-1}$, and

$$f = Bg. \tag{3.2}$$

The amplitudes can be split into two classes, $f = \begin{pmatrix} f^1 \\ f^2 \end{pmatrix}$, where f^1 is an n-vector representing the fast amplitudes and f^2 an m-vector representing the slow amplitudes. The splitting suggests that we consider a decomposition of A, namely, $A = (A_1, A_2)$, where A_1 is $(m+n) \times n$ and A_2 is $(m+n) \times m$, and a corresponding decomposition of B, namely, $B = \begin{pmatrix} B^1 \\ B^2 \end{pmatrix}$, where B^1 is $n \times (m+n)$, and B^2 is $m \times (m+n)$. Thus, $f^1 = B^1 g$ and $f^2 = B^2 g$. Also, the identity $AB = I$ on $\mathbf{R}^{m+n}$ implies that $A_1 B^1 + A_2 B^2 = I$, while the identity $BA = I$ on $\mathbf{R}^{m+n}$ implies that $B^1 A_1 = I_n$ and $B^2 A_1 = 0$ on $\mathbf{R}^n$ and $B^2 A_2 = I_m$ and $B^1 A_2 = 0$ on $\mathbf{R}^m$.

The fast and slow amplitudes evolve in time. Differentiating Eq. (3.2) along solutions of the system (1.1), we obtain

$$\frac{df}{dt} = B\frac{dg}{dt} + \frac{dB}{dt}g = B(Dg)g + \frac{dB}{dt}g,$$

where Dg is the Jacobian of g. Hence, f satisfies the nonlinear ODE

$$\frac{df}{dt} = \Lambda f, \tag{3.3}$$

where Λ, the generator of the dynamics for the amplitudes, is given by

$$\Lambda = B(Dg)A + \frac{dB}{dt}A. \tag{3.4}$$

Since $BA = I$ and I is time invariant, A, B, and their time derivatives satisfy the identity

$$(dB/dt)A + B(dA/dt) = 0 \tag{3.5}$$

at all times. Hence, the definition (3.4) is equivalent to

$$\Lambda = B(Dg)A - B\frac{dA}{dt}, \tag{3.6}$$

7

where $dA/dt = (DA)g$. For completeness, we note that the identity (3.5) implies that $((DB)Af)A + B((DA)Af) = 0$.

In general, the operator Λ is not diagonal, and the equations governing the evolution of f^1 and f^2 are coupled. An ideal basis A is one in which Λ is block-diagonalized, so that the ODEs for f^1 and f^2 decouple. The CSP method approaches this ideal by successive refinements of the basis matrices A and B. The algorithm starts from a constant matrix $A^{(0)}$,

$$A^{(0)} = \left(A_1^{(0)}, A_2^{(0)} \right) = \begin{pmatrix} A_{11}^{(0)} & A_{12}^{(0)} \\ A_{21}^{(0)} & A_{22}^{(0)} \end{pmatrix}. \tag{3.7}$$

Here, $A_{11}^{(0)}$ is an $m \times n$ matrix, $A_{22}^{(0)}$ an $n \times m$ matrix, and the off-diagonal blocks $A_{12}^{(0)}$ and $A_{21}^{(0)}$ are full-rank square matrices of order m and n, respectively. A common choice is $A_{11}^{(0)} = 0$, so every column vector of $A_1^{(0)}$ lies in the fast subspace. We follow this convention and assume, henceforth, that $A_{11}^{(0)} = 0$,

$$A^{(0)} = \left(A_1^{(0)}, A_2^{(0)} \right) = \begin{pmatrix} 0 & A_{12}^{(0)} \\ A_{21}^{(0)} & A_{22}^{(0)} \end{pmatrix}. \tag{3.8}$$

A more general choice of $A^{(0)}$ is discussed below, after Theorem 3.1. The inverse of $A^{(0)}$ is

$$
\begin{aligned}
B_{(0)} &= (A^{(0)})^{-1} = \begin{pmatrix} B_{(0)}^1 \\ B_{(0)}^2 \end{pmatrix} = \begin{pmatrix} B_{(0)}^{11} & B_{(0)}^{12} \\ B_{(0)}^{21} & 0 \end{pmatrix} \\
&= \begin{pmatrix} -(A_{21}^{(0)})^{-1} A_{22}^{(0)} (A_{12}^{(0)})^{-1} & (A_{21}^{(0)})^{-1} \\ (A_{12}^{(0)})^{-1} & 0 \end{pmatrix}.
\end{aligned}
\tag{3.9}
$$

The algorithm then proceeds iteratively. For $q = 0, 1, \ldots$, one first defines the matrix $\Lambda_{(q)}$ in accordance with Eq. (3.6),

$$\Lambda_{(q)} = B_{(q)}(Dg)A^{(q)} - B_{(q)} \frac{dA^{(q)}}{dt} = \begin{pmatrix} \Lambda_{(q)}^{11} & \Lambda_{(q)}^{12} \\ \Lambda_{(q)}^{21} & \Lambda_{(q)}^{22} \end{pmatrix}, \tag{3.10}$$

and matrices $U_{(q)}$ and $L_{(q)}$,

$$U_{(q)} = \begin{pmatrix} 0 & (\Lambda_{(q)}^{11})^{-1} \Lambda_{(q)}^{12} \\ 0 & 0 \end{pmatrix}, \quad L_{(q)} = \begin{pmatrix} 0 & 0 \\ \Lambda_{(q)}^{21} (\Lambda_{(q)}^{11})^{-1} & 0 \end{pmatrix}. \tag{3.11}$$

Then one updates $A^{(q)}$ and $B_{(q)}$ according to the formulas

$$
\begin{aligned}
A^{(q+1)} &= A^{(q)}(I - U_{(q)})(I + L_{(q)}), \tag{3.12} \\
B_{(q+1)} &= (I - L_{(q)})(I + U_{(q)})B_{(q)}, \tag{3.13}
\end{aligned}
$$

and returns to Eq. (3.10) for the next iteration.

At each iteration, one imposes the *CSP condition,*

$$B_{(q)}^1 g = 0, \quad q = 0, 1, \ldots, \tag{3.14}$$

to identify those points where the fast reaction rates vanish with respect to the then-current basis. For $q = 0$, $B_{(0)}^1$ is constant and given by Eq. (3.9); for $q = 1, 2, \ldots$, the CSP condition takes the form

$$B_{(q)}^1(y, \psi_{(q-1)}(y, \varepsilon), \varepsilon) g(y, z, \varepsilon) = 0, \quad q = 1, 2, \ldots. \tag{3.15}$$

If, for any q, the CSP condition is satisfied by a function $z = \psi_{(q)}(y, \varepsilon)$, then

$$\mathcal{K}_\varepsilon^{(q)} = \{(y, z) : z = \psi_{(q)}(y, \varepsilon), \ y \in K\}, \quad q = 0, 1, \ldots \tag{3.16}$$

is defined as the *CSP manifold* (CSPM) of order q.

Theorem 3.1 *The CSP manifold $\mathcal{K}_\varepsilon^{(q)}$ agrees asymptotically with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^q)$ for $q = 0, 1, \ldots,$*

$$\psi_{(q)}(\cdot, \varepsilon) = \sum_{j=0}^{q} \varepsilon^j h_j + \mathcal{O}(\varepsilon^{q+1}), \quad \varepsilon \downarrow 0. \tag{3.17}$$

Our proof of Theorem 3.1 proceeds via an intermediate result, which is of independent interest. We introduce a "truncated" CSP method, where we apply, at each iteration, only the first of the two steps of the full CSP method and skip the second step. This *one-step CSP method* reduces the matrix Λ to lower block-triangular form. We show that, after q iterations, the one-step CSP method generates a manifold $\tilde{\mathcal{K}}_\varepsilon^{(q)}$, whose asymptotic expansion agrees with that of $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^q)$ (Theorem 4.1). In other words, the one-step CSP method is as accurate as the full CSP method is claimed to be in Theorem 3.1. We then return to the full CSP method and carry out an asymptotic analysis of the modifications introduced by the second step. This second step reduces Λ further to block-diagonal form. We show that, at the qth iteration, the second step affects only terms of $\mathcal{O}(\varepsilon^{q+1})$ and higher. Hence, $\mathcal{K}_\varepsilon^{(q)}$ approximates $\mathcal{M}_\varepsilon$ as accurately as $\tilde{\mathcal{K}}_\varepsilon^{(q)}$, and Theorem 3.1 follows.

Theorem 3.1 extends readily to the case where the eigenvectors of the Jacobian Dg are used, instead of the stoichiometric vectors, to form the initial basis $A^{(0)}$. In that case, the slow subspace of the leading-order Jacobian coincides with the tangent space $\mathcal{T}_p \mathcal{M}_0$ at any point $p \in \mathcal{M}_0$, so the columns of

9

$A_2^{(0)}$ are tangent to $\mathcal{M}_0$ to leading order. In turn, this implies that the rows of $B_{(0)}^1(p)$ span the orthogonal complement of the tangent space, also to leading order. As a result, the initial CSPM, the solution of $B_{(0)}^1 g = 0$, coincides with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon)$, which is one order higher than is the case when $A^{(0)}$ is given by Eq. (3.8). Moreover, for each $q = 1, 2, \ldots$, the proof of Theorem 4.1 generalizes directly to this case. The asymptotic expansion of $\tilde{\psi}_{(q)}$ coincides with that of h_ε up to and including terms of $\mathcal{O}(\varepsilon^{q+1})$, which is one order higher than is the case when $A^{(0)}$ is given by Eq. (3.8).

Remark. Lam and Goussis, in their presentation [12] of the CSP method, perform the update (3.12) and (3.13) in two steps. The first step corresponds to the postmultiplication of $A^{(q)}$ with $I - U_{(q)}$ and premultiplication of $B_{(q)}$ with $I + U_{(q)}$, the second step to the subsequent postmultiplication of $A^{(q)}(I - U_{(q)})$ with $I + L_{(q)}$ and premultiplication of $(I + U_{(q)})B_{(q)}$ with $I - L_{(q)}$. The nonzero entries of $U_{(q)}$ and $L_{(q)}$ are chosen so that Λ is block-diagonalized to successively higher order in ε.

Remark. The definition (3.6) implies that Λ is the product of B with the Lie bracket of A (considered column by column) and g,

$$\Lambda = B[A, g] = B([A_{\cdot,1}, g], \ldots, [A_{\cdot,m+n}, g]). \tag{3.18}$$

The Lie bracket of two vector fields a and g is $[a, g] = (Dg)a - (Da)g$ [21].

Remark. It is useful to state how Λ transforms to understand its properties as an operator. If $\hat{A} = AC$ and $\hat{B} = C^{-1}B$, where C is an invertible square matrix representing a coordinate transformation in $\mathbf{R}^{m+n}$, then

$$
\begin{aligned}
\hat{\Lambda} &= \hat{B}(Dg)\hat{A} - \hat{B}\frac{d\hat{A}}{dt} = C^{-1}B(Dg)AC - C^{-1}B\frac{d(AC)}{dt} \\
&= C^{-1}B(Dg)AC - C^{-1}B\left(\frac{dA}{dt}C + A\frac{dC}{dt}\right) \\
&= C^{-1}\Lambda C - C^{-1}\frac{dC}{dt}, \tag{3.19}
\end{aligned}
$$

where $dC/dt = (DC)g$. The presence of the term $C^{-1}dC/dt$ in Eq. (3.19) shows that $\hat{\Lambda}$ and Λ are not similar unless C is constant.

4   The One-Step CSP Method

The goal of the one-step CSP method is to reduce the matrix Λ to lower block-triangular form—that is, to push the matrix Λ^{12} to increasingly higher order in ε. The method is identical to the full CSP method except for the updating of the matrices A and B. One starts from the same bases, $\tilde{A}^{(0)} = A^{(0)}$ and $\tilde{B}_{(0)} = B_{(0)}$, and, instead of Eqs. (3.12) and (3.13), uses the one-step expressions

$$
\begin{aligned}
\tilde{A}^{(q+1)} &= \tilde{A}^{(q)}(I - \tilde{U}_{(q)}), & (4.1)\\
\tilde{B}_{(q+1)} &= (I + \tilde{U}_{(q)})\tilde{B}_{(q)}, & (4.2)
\end{aligned}
$$

where the matrix $\tilde{U}_{(q)}$ is defined as in Eq. (3.11) with Λ replaced by $\tilde{\Lambda}$. (A tilde ˜ distinguishes a quantity from its counterpart in the full CSP method.)

The update rule for $\tilde{\Lambda}$ follows immediately from Eq. (3.19),

$$
\tilde{\Lambda}_{(q+1)} = (I + \tilde{U}_{(q)})\tilde{\Lambda}_{(q)}(I - \tilde{U}_{(q)}) + (I + \tilde{U}_{(q)})\frac{d\tilde{U}_{(q)}}{dt}. \tag{4.3}
$$

(Note that the identities $\tilde{A}^{(0)} = A^{(0)}$ and $\tilde{B}_{(0)} = B_{(0)}$ imply that $\tilde{\Lambda}_{(0)} = \Lambda_{(0)}$.) The matrix $\tilde{U}_{(q)}$ and its time derivative have the same block structure; only the upper right block is nonzero, so $\tilde{U}_{(q)}d\tilde{U}_{(q)}/dt = 0$, and Eq. (4.3) reduces to

$$
\tilde{\Lambda}_{(q+1)} = (I + \tilde{U}_{(q)})\tilde{\Lambda}_{(q)}(I - \tilde{U}_{(q)}) + \frac{d\tilde{U}_{(q)}}{dt}. \tag{4.4}
$$

In terms of the constituent blocks, we have

$$
\begin{aligned}
\tilde{\Lambda}^{11}_{(q+1)} &= \tilde{\Lambda}^{11}_{(q)} + \tilde{U}_{(q)}\tilde{\Lambda}^{21}_{(q)}, & (4.5)\\
\tilde{\Lambda}^{12}_{(q+1)} &= \tilde{U}_{(q)}\tilde{\Lambda}^{22}_{(q)} - \tilde{U}_{(q)}\tilde{\Lambda}^{21}_{(q)}\tilde{U}_{(q)} + \frac{d\tilde{U}_{(q)}}{dt}, & (4.6)\\
\tilde{\Lambda}^{21}_{(q+1)} &= \tilde{\Lambda}^{21}_{(q)}, & (4.7)\\
\tilde{\Lambda}^{22}_{(q+1)} &= \tilde{\Lambda}^{22}_{(q)} - \tilde{\Lambda}^{21}_{(q)}\tilde{U}_{(q)}, & (4.8)
\end{aligned}
$$

where we have used Eq. (3.11) to simplify Eq. (4.6). Note that we freely use $\tilde{U}_{(q)}$ to denote both the full update matrix and its restriction to the subspace $\mathbf{R}^m$; the latter is represented by the matrix $(\tilde{\Lambda}^{11}_{(q)})^{-1}\tilde{\Lambda}^{12}_{(q)}$. The appropriate interpretation is clear from the context.

The one-step CSP method generates a sequence of manifolds,

$$
\tilde{\mathcal{K}}^{(q)}_{\varepsilon} = \{(y,z): z = \tilde{\psi}_{(q)}(y,\varepsilon),\ y \in K\}, \quad q = 0,1,\dots, \tag{4.9}
$$

just like the full CSP method; cf. Eq. (3.16). The functions $\tilde{\psi}_{(q)}$ are defined by the conditions

$$\tilde{B}^1_{(q)}g = 0, \quad q = 0, 1, \dots, \tag{4.10}$$

where $\tilde{B}^1_{(q)}$ is obtained from Eq. (4.2).

Theorem 4.1 *The manifold $\tilde{\mathcal{K}}^{(q)}_\varepsilon$ agrees asymptotically with $\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^q)$ for $q = 0, 1, \dots,$*

$$\tilde{\psi}_{(q)}(\cdot, \varepsilon) = \sum_{j=0}^{q} \varepsilon^j h_j + \mathcal{O}(\varepsilon^{q+1}), \quad \varepsilon \downarrow 0. \tag{4.11}$$

The proof of the theorem is by induction on q.

4.1 The Induction Hypothesis

The central idea of the proof of Theorem 4.1 is to express the CSP condition (4.10) in a form that resembles that of the invariance equation (2.10) and then to derive the conditions under which the left and right members of the two equations are the same at each order.

We begin by expressing the quantities $\tilde{A}^{(q+1)}$, $\tilde{B}_{(q+1)}$, and $\tilde{\Lambda}_{(q+1)}$ in terms of the original quantities $A^{(0)}$, $B_{(0)}$, and $\Lambda_{(0)}$. Applying the definition (4.1) recursively, we find

$$\tilde{A}^{(q+1)} = A^{(0)} \prod_{j=0}^{q} (I - \tilde{U}_{(j)}).$$

Since each $\tilde{U}_{(j)}$ is nilpotent, it follows that

$$\tilde{A}^{(q+1)} = A^{(0)}(I - \tilde{P}_{(q)}), \tag{4.12}$$

where

$$\tilde{P}_{(q)} = \sum_{j=0}^{q} \tilde{U}_{(j)} = \begin{pmatrix} 0 & \sum_{\ell=0}^{q}(\tilde{\Lambda}^{11}_{(\ell)})^{-1}\tilde{\Lambda}^{12}_{(\ell)} \\ 0 & 0 \end{pmatrix}. \tag{4.13}$$

Similarly,

$$\tilde{B}_{(q+1)} = (I + \tilde{P}_{(q)})B_{(0)}. \tag{4.14}$$

Substituting Eqs. (4.12) and (4.14) into the transformation formula (3.19), and recalling that $\tilde{\Lambda}_{(0)} = \Lambda_{(0)}$ and $\tilde{P}_{(q)}d\tilde{P}_{(q)}/dt = 0$, we find

$$\tilde{\Lambda}_{(q+1)} = (I + \tilde{P}_{(q)})\Lambda_{(0)}(I - \tilde{P}_{(q)}) + \frac{d\tilde{P}_{(q)}}{dt}. \tag{4.15}$$

We use these expressions to rewrite Eq. (4.10). Since $B_{(0)}^{22} = 0$, the equation becomes

$$B_{(0)}^{12} g_2 + \varepsilon \left[\tilde{P}_{(q-1)} B_{(0)}^{21} + B_{(0)}^{11} \right] g_1 = 0$$

or, since $B_{(0)}^{12} = (A_{21}^{(0)})^{-1}$,

$$g_2 + \varepsilon A_{21}^{(0)} \left[\tilde{P}_{(q-1)} B_{(0)}^{21} + B_{(0)}^{11} \right] g_1 = 0. \tag{4.16}$$

The last equation has the same form as the invariance equation (2.10). The solution of Eq. (2.10) is $z = h_\varepsilon(y)$, which defines $\mathcal{M}_\varepsilon$, while the solution of Eq. (4.16) is $z = \tilde{\psi}_{(q)}(y, \varepsilon)$, which defines $\tilde{\mathcal{K}}_\varepsilon^{(q)}$.

We analyze the CSP condition (4.16) order by order, up to and including the terms of $\mathcal{O}(\varepsilon^q)$. We recall that the components of the vector field $g(y, z, \varepsilon)$ are evaluated at $z = \tilde{\psi}_{(q)}(y, \varepsilon)$, the matrix $\tilde{P}_{(q-1)}$ is evaluated at $z = \tilde{\psi}_{(q-1)}(y, \varepsilon)$, and the blocks of $A^{(0)}$ and $B_{(0)}$ are constant. Substituting the asymptotic expansion of $\tilde{\psi}_{(q)}$,

$$\tilde{\psi}_{(q)}(y, \varepsilon) = \sum_{j=0}^{\infty} \varepsilon^j \tilde{\psi}_{(q,j)}(y), \quad \varepsilon \downarrow 0, \tag{4.17}$$

into Eq. (4.16) and setting the coefficients of $1, \varepsilon, \ldots, \varepsilon^q$ equal to zero, we obtain a set of equations,

$$g_{2,j} + A_{21}^{(0)} \left[\tilde{P}_{(q-1,0)} B_{(0)}^{21} + B_{(0)}^{11} \right] g_{1,j-1} + \sum_{\ell=1}^{j-1} A_{21}^{(0)} \tilde{P}_{(q-1,\ell)} B_{(0)}^{21} g_{1,j-\ell-1} = 0, \tag{4.18}$$

for $j = 0, 1, \ldots, q$. Here, $\tilde{P}_{(q-1,\ell)}$ is the coefficient of the $\mathcal{O}(\varepsilon^\ell)$ term in the asymptotic expansion of $\tilde{P}_{(q-1)}$.

Equation (4.18) defines $\tilde{\psi}_{(q,j)}$ for $j = 0, 1, \ldots, q$. The leading-order ($j = 0$) equation in the system (4.18) is the same for all q,

$$g_2(y, \tilde{\psi}_{(q,0)}(y), 0) = 0, \quad q = 0, 1, \ldots . \tag{4.19}$$

This is also the equation defining h_0. Its solution need not be unique, but we can identify each $\tilde{\psi}_{(q,0)}$ with h_0,

$$\tilde{\psi}_{(q,0)}(y) = h_0(y), \quad q = 0, 1, \ldots . \tag{4.20}$$

Then also $\tilde{\psi}_{(q)}(\cdot, 0) = h_0$ for $q = 0, 1, \ldots$, so to leading order each manifold $\tilde{\mathcal{K}}_\varepsilon^{(q)}$ coincides with $\mathcal{M}_0$.

13

We wish to show that $\tilde{\psi}_{(q,j)} = h_j$ also for $j = 1, 2, \ldots, q$. To this end, we compare Eqs. (2.14) and (4.18). For a fixed j, the two equations match if

$$A_{21}^{(0)} \left[\tilde{P}_{(q-1,0)} B_{(0)}^{21} + B_{(0)}^{11} \right] = -Dh_0, \qquad (4.21)$$

$$A_{21}^{(0)} \tilde{P}_{(q-1,\ell)} B_{(0)}^{21} = -Dh_\ell, \quad \ell = 1, \ldots, j-1. \qquad (4.22)$$

Conversely, if Eqs. (4.21) and (4.22) hold, then $\tilde{\psi}_{(q,j)} = h_j$. Notice that Eqs. (4.21) and (4.22) are independent of j; hence, they are nested, in the sense that, when j is increased by one, the equations for lower values of j remain the same. Thus, it suffices to prove Eqs. (4.21) and (4.22) for $j = q$. The proof is by induction on q, where the induction hypothesis is

$$\tilde{U}_{(q-1)}(\cdot, \tilde{\psi}_{(q-1)}, \varepsilon) = \mathcal{O}(\varepsilon^{q-1}), \qquad (4.23)$$

$$A_{21}^{(0)} \left[\tilde{P}_{(q-1)}(\cdot, \tilde{\psi}_{(q-1)}, \varepsilon) B_{(0)}^{21} + B_{(0)}^{11} \right] = -\sum_{j=0}^{q-1} \varepsilon^j Dh_j + \mathcal{O}(\varepsilon^q), \quad (4.24)$$

$$\tilde{\psi}_{(q)}(\cdot, \varepsilon) = \sum_{j=0}^{q} \varepsilon^j h_j + \mathcal{O}(\varepsilon^{q+1}). \qquad (4.25)$$

The validity of these equations for $q = 1$ is shown in Section 4.2. The induction step is carried out in Section 4.3.

4.2 Proof of Theorem 4.1 for q = 1

We fix $q = 1$ and consider the $\mathcal{O}(\varepsilon)$ terms of Eq. (4.16),

$$(D_z g_2)_0 \tilde{\psi}_{(1,1)} + (D_\varepsilon g_2)_0 + A_{21}^{(0)} \left[\tilde{P}_{(0,0)} B_{(0)}^{21} + B_{(0)}^{11} \right] g_{1,0} = 0. \qquad (4.26)$$

The first and second terms in this equation are exactly the same as those in the equation for h_1, see (2.16). Therefore, we need only to show that the third term equals $-(Dh_0)g_{1,0}$ in order to prove the theorem for $q = 1$.

According to the definitions (4.13) and (3.11) with $q = 0$, we have

$$\tilde{P}_{(0)} = \tilde{U}_{(0)} = (\tilde{\Lambda}_{(0)}^{11})^{-1} \tilde{\Lambda}_{(0)}^{12} = (\Lambda_{(0)}^{11})^{-1} \Lambda_{(0)}^{12}, \qquad (4.27)$$

where $\Lambda_{(0)} = B_{(0)}(Dg)A^{(0)}$, according to the definition in Eq. (3.10). Now, $\Lambda_{(0)}$ admits an asymptotic expansion, $\Lambda_{(0)} = \sum_{j=0}^{\infty} \varepsilon^j \Lambda_{(0,j)}$, and each of the coefficient matrices $\Lambda_{(0,j)}$ consists of four blocks,

$$\Lambda_{(0,j)}^{11} = \left[B_{(0)}^{12}(D_z g_2)_j + B_{(0)}^{11}(D_z g_1)_{j-1} \right] A_{21}^{(0)}, \qquad (4.28)$$

14

$$\Lambda_{(0,j)}^{12} = B_{(0)}^{12} \left[(D_y g_2)_j A_{12}^{(0)} + (D_z g_2)_j A_{22}^{(0)} \right]$$
$$+ B_{(0)}^{11} \left[(D_y g_1)_{j-1} A_{12}^{(0)} + (D_z g_1)_{j-1} A_{22}^{(0)} \right], \tag{4.29}$$

$$\Lambda_{(0,j)}^{21} = B_{(0)}^{21} (D_z g_1)_{j-1} A_{21}^{(0)}, \tag{4.30}$$

$$\Lambda_{(0,j)}^{22} = B_{(0)}^{21} \left[(D_y g_1)_{j-1} A_{12}^{(0)} + (D_z g_1)_{j-1} A_{22}^{(0)} \right]. \tag{4.31}$$

The notation $(\,\cdot\,)_j$ indicates the jth term in the asymptotic expansion of the quantity inside the parentheses, and it is understood that such a term is absent if the subscript is negative.

A direct evaluation shows that the blocks $\Lambda_{(0,0)}^{11}$ and $\Lambda_{(0,0)}^{12}$ are nonzero. Therefore, $\Lambda_{(0)}^{11}$ and $\Lambda_{(0)}^{12}$ are both $\mathcal{O}(1)$, and

$$\tilde{P}_{(0,0)} = \tilde{U}_{(0,0)} = (\Lambda_{(0,0)}^{11})^{-1} \Lambda_{(0,0)}^{12} = B_{(0)}^{12} \left[(D_z g_2)_0^{-1} (D_y g_2)_0 A_{12}^{(0)} + A_{22}^{(0)} \right]. \tag{4.32}$$

Here, all the quantities are evaluated on $\mathcal{M}_0$, where the identity

$$(D_z g_2)_0^{-1} (D_y g_2)_0 = -D h_0 \tag{4.33}$$

holds. Hence, Eq. (4.32) implies

$$\tilde{P}_{(0,0)} = \tilde{U}_{(0,0)} = B_{(0)}^{12} \left(A_{22}^{(0)} - (D h_0) A_{12}^{(0)} \right). \tag{4.34}$$

Finally, substituting this expression for $\tilde{P}_{(0,0)}$ into Eq. (4.26) and using the identity $A_{21}^{(0)} B_{(0)}^{11} = -A_{22}^{(0)} B_{(0)}^{21}$, we obtain

$$(D_z g_2)_0 \tilde{\psi}_{(1,1)} + (D_\varepsilon g_2)_0 - (D h_0) g_{1,0} = 0. \tag{4.35}$$

This equation for $\tilde{\psi}_{(1,1)}$ is the same as Eq. (2.16) for h_1; hence, $\tilde{\psi}_{(1,1)} = h_1$ and $\tilde{\psi}_{(1)} = h_0 + \varepsilon h_1 + \mathcal{O}(\varepsilon^2)$. This proves the theorem for $q = 1$.

4.3   Proof of Theorem 4.1 for q = 2, 3, . . .

We prove that Eqs. (4.23)–(4.25) hold for $q + 1$, assuming that they hold for $0, 1, \ldots, q$. By our discussion of Eqs. (4.21) and (4.22), Eq. (4.25) follows immediately from Eq. (4.24), so we need only to consider Eqs. (4.23) and (4.24).

4.3.1   Establishing Eq. (4.23)

We first consider Eq. (4.23). The induction hypothesis gives the estimate $\tilde{U}_{(i)}(\,\cdot\,, \tilde{\psi}_{(i)}, \varepsilon) = \mathcal{O}(\varepsilon^i)$ for $i = 0, 1, \ldots, q - 1$. Also, $\tilde{\psi}_{(q)} = \tilde{\psi}_{(i)} + \mathcal{O}(\varepsilon^{i+1})$ for

15

$i = 0, 1, \dots, q-1$. Hence, $\tilde{U}_{(i)}(\,\cdot\,, \check{\psi}_{(q)}, \varepsilon) = \tilde{U}_{(i)}(\,\cdot\,, \check{\psi}_{(i)}, \varepsilon) + \mathcal{O}(\varepsilon^{i+1})$, from which it follows that

$$\tilde{U}_{(i)}(\,\cdot\,, \check{\psi}_{(q)}, \varepsilon) = \mathcal{O}(\varepsilon^i), \qquad i = 0, 1, \dots, q-1. \tag{4.36}$$

In particular, $\tilde{U}_{(0)}(\,\cdot\,, \check{\psi}_{(q)}, \varepsilon) = \mathcal{O}(1)$, so

$$\tilde{P}_{(q-1)} = \sum_{\ell=0}^{q-1} \tilde{U}_{(\ell)} = \mathcal{O}(1) \text{ on } \tilde{\mathcal{K}}_\varepsilon^{(q)}.$$

This asymptotic estimate can be used to derive asymptotic expansions of the blocks of $\Lambda_{(q)}$. We begin with $\tilde{\Lambda}_{(q)}^{11}$. From Eq. (4.15), we have

$$\tilde{\Lambda}_{(q)}^{11} = \Lambda_{(0)}^{11} + \tilde{P}_{(q-1)} \Lambda_{(0)}^{21}. \tag{4.37}$$

Since $\Lambda_{(0)}^{21} = \mathcal{O}(\varepsilon)$ by Eq. (4.30), we see immediately that

$$\tilde{\Lambda}_{(q)}^{11} = \Lambda_{(0,0)}^{11} + \mathcal{O}(\varepsilon). \tag{4.38}$$

Next, we examine the block $\tilde{\Lambda}_{(q)}^{12}$. From Eq. (4.6), we have

$$\tilde{\Lambda}_{(q)}^{12} = \tilde{U}_{(q-1)} \tilde{\Lambda}_{(q-1)}^{22} - \tilde{U}_{(q-1)} \tilde{\Lambda}_{(q-1)}^{21} \tilde{U}_{(q-1)} + \frac{d\tilde{U}_{(q-1)}}{dt}.$$

First, $\tilde{U}_{(q-1)}(\,\cdot\,, \check{\psi}_{(q)}, \varepsilon) = \mathcal{O}(\varepsilon^{q-1})$ by Eq. (4.36). Also, $\tilde{\Lambda}_{(q-1)}^{21} = \Lambda_{(0)}^{21} = \mathcal{O}(\varepsilon)$ on $\tilde{\mathcal{K}}_\varepsilon^{(q)}$ by Eqs. (4.15) and (4.30). Moreover, $\tilde{\Lambda}_{(q-1)}^{22} = \Lambda_{(0)}^{22} - \Lambda_{(0)}^{21} \tilde{P}_{(q-2)} = \mathcal{O}(\varepsilon)$ by Eqs. (4.15) and (4.31). Finally, by applying Lemma A.2 with $V = \tilde{U}_{(q-1)}$, we find that $d\tilde{U}_{(q-1)}/dt$ is $\mathcal{O}(\varepsilon^q)$. Putting these estimates together, we obtain the estimate

$$\tilde{\Lambda}_{(q)}^{12} = \varepsilon^q \tilde{\Lambda}_{(q,q)}^{12} + \mathcal{O}(\varepsilon^{q+1}), \tag{4.39}$$

where we grouped all of the $\mathcal{O}(\varepsilon^q)$ terms into $\varepsilon^q \tilde{\Lambda}_{(q,q)}^{12}$. By combining the definition (3.11) with Eqs. (4.38) and (4.39), we derive the desired estimate, $\tilde{U}_{(q)} = (\tilde{\Lambda}_{(q)}^{11})^{-1} \tilde{\Lambda}_{(q)}^{12} = \mathcal{O}(\varepsilon^q)$.

Remark. While the estimates of $\tilde{\Lambda}_{(q)}^{21}$ and $\tilde{\Lambda}_{(q)}^{22}$ are not needed here, they will be needed in Section 5. First, $\tilde{\Lambda}_{(q)}^{21} = \Lambda_{(0)}^{21} = \mathcal{O}(\varepsilon)$ on $\tilde{\mathcal{K}}_\varepsilon^{(q)}$, by Eqs. (4.15) and (4.30). Then, $\tilde{\Lambda}_{(q)}^{22} = \Lambda_{(0)}^{22} - \Lambda_{(0)}^{21} \tilde{P}_{(q-1)}$ by Eq. (4.15). Now, $\Lambda_{(0)}^{22} = \mathcal{O}(\varepsilon)$ by Eq. (4.31), and thus the discussion for the size of $\tilde{\Lambda}_{(q)}^{11}$ also yields that $\tilde{\Lambda}_{(q)}^{22} = \mathcal{O}(\varepsilon)$. Putting the estimates of this section together, we obtain

$$\tilde{\Lambda}_{(q)}(\,\cdot\,, \check{\psi}_{(q)}, \varepsilon) = \begin{pmatrix} \Lambda_{(0,0)}^{11} + \mathcal{O}(\varepsilon) & \varepsilon^q \tilde{\Lambda}_{(q,q)}^{12} + \mathcal{O}(\varepsilon^{q+1}) \\ \varepsilon \Lambda_{(0,1)}^{21} + \mathcal{O}(\varepsilon^2) & \varepsilon \tilde{\Lambda}_{(q,1)}^{22} + \mathcal{O}(\varepsilon^2) \end{pmatrix}. \tag{4.40}$$

16

4.3.2 Establishing Eq. (4.24)

Next, we consider Eq. (4.24). The induction hypothesis gives the estimate $A_{21}^{(0)}[\tilde{P}_{(i)}(\,\cdot\,,\tilde{\psi}_{(i)},\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}] = -\sum_{j=0}^{i} \varepsilon^j Dh_j + \mathcal{O}(\varepsilon^{i+1})$ for $i = 0, 1, \ldots, q-1$. Our goal is to show that this equation also holds for $i = q$. We first show that the terms up to and including $\mathcal{O}(\varepsilon^{q-1})$ in both members of the equation agree for $i = q$. Then we analyze the terms of $\mathcal{O}(\varepsilon^q)$.

By the induction hypothesis, we have the asymptotic expansion

$$A_{21}^{(0)}\left[\tilde{P}_{(q-1)}(\,\cdot\,,\tilde{\psi}_{(q-1)},\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}\right] = -\sum_{j=0}^{q-1} \varepsilon^j Dh_j + \mathcal{O}(\varepsilon^q). \quad (4.41)$$

Also by the induction hypothesis, $\tilde{\psi}_{(q)} = \tilde{\psi}_{(q-1)} + \mathcal{O}(\varepsilon^q)$. Hence,

$$A_{21}^{(0)}\left[\tilde{P}_{(q-1)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}\right] = -\sum_{j=0}^{q-1} \varepsilon^j Dh_j + \mathcal{O}(\varepsilon^q). \quad (4.42)$$

The definition (4.13) of $\tilde{P}_{(q)}$ yields the update formula

$$\tilde{P}_{(q)} = \tilde{P}_{(q-1)} + \tilde{U}_{(q)}. \quad (4.43)$$

We already showed that $\tilde{U}_{(q)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon) = \mathcal{O}(\varepsilon^q)$, so Eq. (4.43) implies that the asymptotic expansions of $\tilde{P}_{(q)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon)$ and $\tilde{P}_{(q-1)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon)$ agree up to and including terms of $\mathcal{O}(\varepsilon^{q-1})$. The same, then, holds for the asymptotic expansions of $A_{21}^{(0)}[\tilde{P}_{(q)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon)B_{(0)}^{21}+B_{(0)}^{11}]$ and $A_{21}^{(0)}[\tilde{P}_{(q-1)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon)B_{(0)}^{21}+B_{(0)}^{11}]$. Therefore,

$$A_{21}^{(0)}\left[\tilde{P}_{(q)}(\,\cdot\,,\tilde{\psi}_{(q)},\varepsilon)B_{(0)}^{21} + B_{(0)}^{11}\right] = -\sum_{j=0}^{q-1} \varepsilon^j Dh_j + \mathcal{O}(\varepsilon^q). \quad (4.44)$$

In other words,

$$A_{21}^{(0)}\left[\tilde{P}_{(q,0)}B_{(0)}^{21} + B_{(0)}^{11}\right] = -Dh_0, \quad (4.45)$$

$$A_{21}^{(0)}\tilde{P}_{(q,j)}B_{(0)}^{21} = -Dh_j, \text{ for all } j = 1,\ldots,q-1, \quad (4.46)$$

which establishes Eq. (4.24) for all terms up to and including $\mathcal{O}(\varepsilon^{q-1})$.

It remains to show that the terms of $\mathcal{O}(\varepsilon^q)$ in both members of Eq. (4.24) agree, that is,

$$A_{21}^{(0)}\tilde{P}_{(q,q)}B_{(0)}^{21} = -Dh_q. \quad (4.47)$$

17

We achieve this by deriving an explicit formula for $A_{21}^{(0)} \tilde{P}_{(q,q)} B_{(0)}^{21}$ and comparing it to that for Dh_q, which is given in the Appendix (Lemma A.1). We proceed in two steps. In step one, we express $A_{21}^{(0)} \tilde{P}_{(q,q)} B_{(0)}^{21}$ in terms of $\tilde{P}_{(q-1,0)}, \ldots, \tilde{P}_{(q-1,q-1)}$. Then, in step two, we obtain the explicit formula for $A_{21}^{(0)} \tilde{P}_{(q,q)} B_{(0)}^{21}$ in terms of the vector field and of Dh_i, $i = 0, 1, \ldots, q - 1$.

Step 1. Recall the update formula (4.43), $\tilde{P}_{(q)} = \tilde{P}_{(q-1)} + \tilde{U}_{(q)}$. Using the definition (3.11) of $\tilde{U}_{(q)}$ and the explicit formula (4.15) for $\tilde{\Lambda}_{(q)}$, we can express $\tilde{U}_{(q)}$ in terms of $\Lambda_{(0)}$ and $\tilde{P}_{(q-1)}$. In particular, Eq. (4.40) implies that $\tilde{U}_{(q,q)} = (\Lambda_{(0,0)}^{11})^{-1} \tilde{\Lambda}_{(q,q)}^{12}$. Also, Eq. (4.15) gives

$$\tilde{\Lambda}_{(q)}^{12} = \Lambda_{(0)}^{12} - \Lambda_{(0)}^{11} \tilde{P}_{(q-1)} + \tilde{P}_{(q-1)} \Lambda_{(0)}^{22} - \tilde{P}_{(q-1)} \Lambda_{(0)}^{21} \tilde{P}_{(q-1)} + \frac{d\tilde{P}_{(q-1)}}{dt}. \qquad (4.48)$$

It follows that

$$
\begin{aligned}
\tilde{U}_{(q,q)} &= (\Lambda_{(0,0)}^{11})^{-1} \left[\Lambda_{(0,q)}^{12} - \left(\Lambda_{(0)}^{11} \tilde{P}_{(q-1)} \right)_q + \left(\tilde{P}_{(q-1)} \Lambda_{(0)}^{22} \right)_q \right. \\
&\qquad \left. - \left(\tilde{P}_{(q-1)} \Lambda_{(0)}^{21} \tilde{P}_{(q-1)} \right)_q + \left(\frac{d\tilde{P}_{(q-1)}}{dt} \right)_q \right],
\end{aligned}
\qquad (4.49)
$$

where we recall the notational convention that $(\cdot)_q$ stands for the coefficient of the $\mathcal{O}(\varepsilon^q)$ term in the asymptotic expansion of the quantity in parentheses. Using Lemma A.2 with $V = \tilde{P}_{(q-1)}$ and the fact that $\Lambda_{(0,0)}^{22}$ and $\Lambda_{(0,0)}^{21}$ are both zero, we rewrite Eq. (4.49) as

$$\tilde{U}_{(q,q)} = (\Lambda_{(0,0)}^{11})^{-1} \left[J_1 + (J_2 - \Lambda_{(0,0)}^{11} \tilde{P}_{(q-1,q)}) + J_3 + J_4 + J_5 \right], \qquad (4.50)$$

where

$$J_1 = \Lambda_{(0,q)}^{12}, \qquad J_2 = -\sum_{\ell=0}^{q-1} \Lambda_{(0,q-\ell)}^{11} \tilde{P}_{(q-1,\ell)}, \qquad J_3 = \sum_{\ell=0}^{q-1} \tilde{P}_{(q-1,\ell)} \Lambda_{(0,q-\ell)}^{22},$$

$$J_4 = -\sum_{i=0}^{q-1} \sum_{j=0}^{q-1-i} \tilde{P}_{(q-1,j)} \Lambda_{(0,q-i-j)}^{21} \tilde{P}_{(q-1,i)}, \qquad J_5 = \sum_{\ell=0}^{q-1} \frac{d\tilde{P}_{(q-1,\ell)}}{dy} g_{1,q-1-\ell}. \qquad (4.51)$$

Substituting the expression (4.50) into the update formula (4.43) for $\tilde{P}_{(q)}$, we find

$$A_{21}^{(0)} \tilde{P}_{(q,q)} B_{(0)}^{21} = A_{21}^{(0)} (\Lambda_{(0,0)}^{11})^{-1} \left[J_1 + J_2 + J_3 + J_4 + J_5 \right] B_{(0)}^{21}, \qquad (4.52)$$

Step 2. We rewrite the terms $J_1, \ldots, J_5$ by means of the induction hypothesis and the explicit formulas (4.28)–(4.31) for the blocks of $\Lambda_{(0)}$.

Equation (4.28) and the identity $A_{21}^{(0)} B_{(0)}^{12} = I_n$ imply that

$$A_{21}^{(0)} \left(\Lambda_{(0,0)}^{11} \right)^{-1} = \left((D_z g_2)_0 \right)^{-1} A_{21}^{(0)}. \tag{4.53}$$

Here, $(D_z g_2)_0$ stands for the leading order term in the asymptotic expansion of $(D_z g_2)(\cdot, \tilde{\psi}_{(q)}, \varepsilon)$. Since $\tilde{\psi}_{(q)}$ and h_ε agree up to and including $\mathcal{O}(\varepsilon^q)$ terms by assumption, the asymptotic expansions of $(D_z g_2)(\cdot, \tilde{\psi}_{(q)}, \varepsilon)$ and $(D_z g_2)(\cdot, h_\varepsilon, \varepsilon)$ also agree up to and including $\mathcal{O}(\varepsilon^q)$ terms. For the remainder of this section, it does not matter whether quantities are evaluated on $\tilde{\mathcal{K}}_\varepsilon^{(q)}$ or on $\mathcal{M}_\varepsilon$, since only the coefficients of ε^q or lower appear in our formulas. Accordingly, we make no distinction between the asymptotic expansions of a quantity evaluated on the two manifolds.

Using Eq. (4.29) and the identities $B_{(0)}^{12} = (A_{21}^{(0)})^{-1}$, $B_{(0)}^{21} = (A_{12}^{(0)})^{-1}$, and $B_{(0)}^{11} = -B_{(0)}^{12} A_{22}^{(0)} B_{(0)}^{21}$, we find

$$A_{21}^{(0)} J_1 B_{(0)}^{21} = (D_y g_2)_q + (D_z g_2)_q A_{22}^{(0)} B_{(0)}^{21}$$
$$- A_{22}^{(0)} B_{(0)}^{21} (D_y g_1)_{q-1} - A_{22}^{(0)} B_{(0)}^{21} (D_z g_1)_{q-1} A_{22}^{(0)} B_{(0)}^{21}. \tag{4.54}$$

Next, substituting for $A_{21}^{(0)} \tilde{P}_{(q-1,\ell)} B_{(0)}^{21}$ from the induction hypothesis (4.24), we obtain

$$A_{21}^{(0)} J_2 B_{(0)}^{21} = \sum_{\ell=0}^{q-1} (D_z g_2)_{q-\ell} Dh_\ell - (D_z g_2)_q A_{22}^{(0)} B_{(0)}^{21}$$
$$- \sum_{\ell=0}^{q-1} A_{22}^{(0)} B_{(0)}^{21} (D_z g_1)_{q-1-\ell} Dh_\ell + A_{22}^{(0)} B_{(0)}^{21} (D_z g_1)_{q-1} A_{22}^{(0)} B_{(0)}^{21}. \tag{4.55}$$

Then, using Eq. (4.31) and the assumptions of the lemma, we find

$$A_{21}^{(0)} J_3 B_{(0)}^{21} = -\sum_{\ell=0}^{q-1} Dh_\ell (D_y g_1)_{q-1-\ell} - \sum_{\ell=0}^{q-1} Dh_\ell (D_z g_1)_{q-1-\ell} A_{22}^{(0)} B_{(0)}^{21}$$
$$+ A_{22}^{(0)} B_{(0)}^{21} (D_y g_1)_{q-1} + A_{22}^{(0)} B_{(0)}^{21} (D_z g_1)_{q-1} A_{22}^{(0)} B_{(0)}^{21}. \tag{4.56}$$

In the same vein, we use the induction hypothesis on J_4,

$$A_{21}^{(0)} J_4 B_{(0)}^{21} = -\sum_{i=0}^{q-1} \sum_{j=0}^{q-1-i} Dh_j (D_z g_1)_{q-1-i-j} Dh_i + \sum_{i=0}^{q-1} A_{22}^{(0)} B_{(0)}^{21} (D_z g_1)_{q-1-i} Dh_i$$
$$+ \sum_{j=0}^{q-1} Dh_j (D_z g_1)_{q-1-j} A_{22}^{(0)} B_{(0)}^{21} - A_{22}^{(0)} B_{(0)}^{21} (D_z g_1)_{q-1} A_{22}^{(0)} B_{(0)}^{21}. \tag{4.57}$$

The terms in Eq. (4.52) containing $A_{22}^{(0)}$ sum to zero, which may be seen as follows. The second and fourth terms in (4.54) cancel against the second and fourth terms in (4.55); the third term in (4.54) cancels against the third term in (4.56); the third term in (4.55) cancels against the second term in (4.57); and the second and fourth terms in (4.56) cancel against the third and fourth terms in (4.57). These cancellations were to be expected because the approximation should be independent of the choice of $A^{(0)}$.

Carrying out the same type of calculation as above, we obtain

$$A_{21}^{(0)} J_5 B_{(0)}^{21} = -\sum_{\ell=0}^{q-1} D^2 h_\ell g_{1,q-1-\ell}, \tag{4.58}$$

where we have used the symmetry of the bilinear form $D^2 h_\ell$.

Equations (4.53)–(4.58), together with the observed cancellations, yield

$$A_{21}^{(0)} \tilde{P}_{(q,q)} B_{(0)}^{21}$$
$$= ((D_z g_2)_0)^{-1} \left[(D_y g_2)_q + \sum_{\ell=0}^{q-1} (D_z g_2)_{q-\ell} D h_\ell - \sum_{\ell=0}^{q-1} D^2 h_\ell g_{1,q-1-\ell} \right.$$
$$\left. - \sum_{\ell=0}^{q-1} D h_\ell (D_y g_1)_{q-1-\ell} - \sum_{i=0}^{q-1} \sum_{j=0}^{q-1-i} D h_j (D_z g_1)_{q-1-i-j} D h_i \right]. \tag{4.59}$$

A term-by-term comparison with the expression for $-Dh_q$ given in the Appendix, Eq. (A.3), shows that $A_{21}^{(0)} \tilde{P}_{(q,q)} A_{(0)}^{21} = -Dh_q$. Thus, the proof of Theorem 4.1 is complete.

Remark. In general, the error term is nontrivial, as can already be seen at $q = 0$. The equation determining $\tilde{\psi}_{(0,1)}$ is

$$(D_z g_2)_0 \tilde{\psi}_{(0,1)} + (D_\varepsilon g_2)_0 - A_{22}^{(0)} B_{(0)}^{21} g_{1,0} = 0. \tag{4.60}$$

This equation is not the same as Eq. (2.16), which determines h_1. Where Eq. (2.16) has the term Dh_0, Eq. (4.60) has the term $A_{22}^{(0)} B_{(0)}^{21}$. When the slow manifold is nonlinear, Dh_0 depends on y, whereas $A_{22}^{(0)} B_{(0)}^{21}$ is a constant matrix. Therefore, in general $\tilde{\psi}_{(0,1)} \neq h_1$, and the strongest claim we can make is $\tilde{\psi}_{(0)} = h_0 + \mathcal{O}(\varepsilon)$. A similar argument applies to higher values of q.

5 Analysis of the Full CSP Method

We now return to the full CSP method and prove Theorem 3.1. Since the full CSP method and the one-step CSP method start from the same basis, the conditions (3.14) and (4.10) are the same for $q = 0$,

$$B^1_{(0)}g = 0. \tag{5.1}$$

Therefore, we can choose $\psi_{(0)} = \tilde{\psi}_{(0)} = h_0$.

5.1 Proof of Theorem 3.1 for q = 1

In this section, we carry out the first iteration of the full CSP method and determine the resulting approximation $\mathcal{K}^{(1)}_\varepsilon$ of the slow manifold. We then compare $\mathcal{K}^{(1)}_\varepsilon$ and $\tilde{\mathcal{K}}^{(1)}_\varepsilon$.

The update quantities $U_{(0)}$ and $L_{(0)}$ follow from the definition (3.11),

$$U_{(0)} = (\Lambda^{11}_{(0)})^{-1}\Lambda^{12}_{(0)}, \qquad L_{(0)} = \Lambda^{21}_{(0)}(\Lambda^{11}_{(0)})^{-1}. \tag{5.2}$$

(We recall that we use the same notation $U_{(0)}$ and $L_{(0)}$ for the full matrix and the nonzero block.) In particular, Eqs. (5.2) and (4.27) imply that $U_{(0)} = \tilde{U}_{(0)}$. Next, we update the matrix $B_{(0)}$. Following Eq. (3.13), we find

$$B_{(1)} = (I - L_{(0)})(I + U_{(0)})B_{(0)}. \tag{5.3}$$

The upper and lower row blocks of $B_{(1)}$ are

$$B^1_{(1)} = B^1_{(0)} + U_{(0)}B^2_{(0)}, \tag{5.4}$$
$$B^2_{(1)} = (I - L_{(0)}U_{(0)})B^2_{(0)} - L_{(0)}B^1_{(0)}. \tag{5.5}$$

Since $\tilde{P}_{(0)} = \tilde{U}_{(0)} = U_{(0)}$ and $\psi_{(0)} = \tilde{\psi}_{(0)}$, Eqs. (4.14) and (5.4) imply that

$$B^1_{(1)} = \tilde{B}^1_{(1)}, \tag{5.6}$$

so after the first iteration the CSP condition is the same as for the one-step method. Therefore, $\psi_{(1)} = \tilde{\psi}_{(1)}$ and, by Theorem 4.1,

$$\psi_{(1)} = h_0 + \varepsilon h_1 + \mathcal{O}(\varepsilon^2). \tag{5.7}$$

This proves Theorem 3.1 for $q = 1$.

21

5.2 The Induction Hypothesis

So far, we have established the identities $B_{(0)}^1 = \tilde{B}_{(0)}^1$ and $B_{(1)}^1 = \tilde{B}_{(1)}^1$, from which we could conclude that $\mathcal{K}_\varepsilon^{(0)} = \tilde{\mathcal{K}}_\varepsilon^{(0)}$ and $\mathcal{K}_\varepsilon^{(1)} = \tilde{\mathcal{K}}_\varepsilon^{(1)}$. In general, though, it is not true that $B_{(q)}^1 = \tilde{B}_{(q)}^1$ for higher values of q, as we now demonstrate.

In the one-step CSP method, Eq. (4.14) yields

$$\tilde{B}_{(2)}^1 = B_{(0)}^1 + (\tilde{U}_{(0)} + \tilde{U}_{(1)})B_{(0)}^2.$$

By contrast, in the full CSP method, we obtain from Eq. (3.13)

$$B_{(2)}^1 = \left(I_n - U_{(1)}L_{(0)}\right)B_{(0)}^1 + \left(U_{(0)} + U_{(1)} - U_{(1)}L_{(0)}U_{(0)}\right)B_{(0)}^2. \tag{5.8}$$

The rows of $B_{(0)}^1$ and $B_{(0)}^2$ are linearly independent, as can be seen from Eq. (3.9), so the presence of the premultiplier of $B_{(0)}^1$ in the expression (5.8) implies that $B_{(2)}^1 \neq \tilde{B}_{(2)}^1$. A similar argument shows that $B_{(q)}^1 \neq \tilde{B}_{(q)}^1$ for $q = 2, 3, \dots$. Consequently, the proof of Theorem 3.1 for $q = 1$ given in Section 5.1 does not generalize to higher values of q.

The matrix $\tilde{B}_{(q)}^1$ has an important property. Using Eq. (4.14), we write

$$\tilde{B}_{(q)}^1 = \left(\tilde{P}_{(q-1)}B_{(0)}^{21} + B_{(0)}^{11}, B_{(0)}^{12}\right) = B_{(0)}^{12}\left(A_{21}^{(0)}\left[\tilde{P}_{(q-1)}B_{(0)}^{21} + B_{(0)}^{11}\right], I_n\right).$$

Given the induction hypothesis (4.24), we rewrite this expression once more,

$$\tilde{B}_{(q)}^1 = B_{(0)}^{12}\left(-\sum_{j=0}^{q-1} \varepsilon^j Dh_j + \mathcal{O}(\varepsilon^q), I_n\right). \tag{5.9}$$

Take any $y \in K$, and let the points $\tilde{Q} \in \tilde{\mathcal{K}}_\varepsilon^{(q-1)}$, $Q \in \mathcal{K}_\varepsilon^{(q-1)}$, and $Q' \in \mathcal{M}_\varepsilon$ be defined by

$$\tilde{Q} = (y, \tilde{\psi}_{(q-1)}(y, \varepsilon)), \quad Q = (y, \psi_{(q-1)}(y, \varepsilon)), \quad Q' = (y, h_\varepsilon(y)).$$

The n row vectors of the matrix $(-Dh_\varepsilon(y), I_n)$ form an exact basis for $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$, the space normal to $\mathcal{M}_\varepsilon$ at Q'. Therefore, by Eq. (5.9), $\tilde{B}_{(q)}^1(\tilde{Q})$ is a linear combination of the basis vectors of $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$, up to and including terms of $\mathcal{O}(\varepsilon^{q-1})$, via the invertible matrix $B_{(0)}^{12}$. Hence, the columns of $\tilde{B}_{(q)}^1(\tilde{Q})$ form a basis for $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^{q-1})$. This property of $\tilde{B}_{(q)}^1(\tilde{Q})$ was central to the proof of Theorem 4.1. We seek to prove a similar result for the rows of $B_{(q)}^1(Q)$.

The rows of $B_{(q)}(Q)$ can be written as linear combinations of the rows of $\tilde{B}_{(q)}(\tilde{Q})$,

$$B_{(q)}(Q) = T_{(q)}(y, \varepsilon)\tilde{B}_{(q)}(\tilde{Q}), \tag{5.10}$$

because $\tilde{B}_{(q)}(\tilde{Q})$ is invertible (see Eq. (4.14)). In terms of the constituent blocks,

$$B^1_{(q)} = T^{11}_{(q)}\tilde{B}^1_{(q)} + T^{12}_{(q)}\tilde{B}^2_{(q)}, \tag{5.11}$$
$$B^2_{(q)} = T^{21}_{(q)}\tilde{B}^1_{(q)} + T^{22}_{(q)}\tilde{B}^2_{(q)}. \tag{5.12}$$

Equation (5.11) shows that the requirement that the rows of $B^1_{(q)}(Q)$ span $\mathcal{N}_{Q'}\mathcal{M}_\varepsilon$ up to and including terms of $\mathcal{O}(\varepsilon^{q-1})$ is equivalent to the conditions

$$T^{11}_{(q)}(y,\varepsilon) = \mathcal{O}(1) \text{ and invertible}, \quad T^{12}_{(q)}(y,\varepsilon) = \mathcal{O}(\varepsilon^q). \tag{5.13}$$

Assume for the moment that these conditions are satisfied. Then the CSP condition (3.14) after the qth iteration can be recast as

$$\left[T^{11}_{(q)}(y,\varepsilon)\tilde{B}^1_{(q)}(y,\psi_{(q-1)}(y,\varepsilon),\varepsilon) + T^{12}_{(q)}(y,\varepsilon)\tilde{B}^2_{(q)}(y,\psi_{(q-1)}(y,\varepsilon),\varepsilon)\right]g(y,z,\varepsilon) = 0$$

or, since $T^{11}_{(q)}(y,\varepsilon)$ is invertible,

$$\tilde{B}^1_{(q)}g + \left(T^{11}_{(q)}\right)^{-1}T^{12}_{(q)}\tilde{B}^2_{(q)}g = 0. \tag{5.14}$$

The second term is at least of $\mathcal{O}(\varepsilon^q)$, by the second assumption in Eq. (5.13), so the terms of $\mathcal{O}(\varepsilon^j)$ in Eqs. (4.10) and (5.14) are equal for $j = 0, 1, \ldots, q-1$. At $\mathcal{O}(\varepsilon^q)$, the two equations differ by the term $(T^{11}_{(q,0)})^{-1}T^{12}_{(q,q)}\tilde{B}^2_{(q,0)}g(y,\psi_{(q,0)},\varepsilon)$. Since the $\mathcal{O}(1)$ terms of the two equations agree, it follows that $\psi_{(q,0)} = \tilde{\psi}_{(q,0)} = h_0$ and, therefore, $g(y,\psi_{(q,0)},\varepsilon) = 0$. Hence, Eqs. (4.16) and (5.14) agree up to and including terms of $\mathcal{O}(\varepsilon^q)$, so Eq. (5.14) produces the asymptotic expansion of the slow manifold up to and including terms of $\mathcal{O}(\varepsilon^q)$, by Theorem 4.1.

To complete the proof of Theorem 3.1, we need to verify the conditions (5.13) for $q = 2, 3, \ldots$, which we do by induction on q. The induction hypothesis is

$$T_{(q)}(\cdot,\psi_{(q-1)},\varepsilon) = \begin{pmatrix} I_n + \mathcal{O}(\varepsilon^2) & \varepsilon^q T^{12}_{(q,q)} + \mathcal{O}(\varepsilon^{q+1}) \\ \varepsilon T^{21}_{(q,1)} + \mathcal{O}(\varepsilon^2) & I_m + \mathcal{O}(\varepsilon^2) \end{pmatrix}, \tag{5.15}$$

$$\psi_{(q)}(\cdot,\varepsilon) = \sum_{j=0}^{q} \varepsilon^j h_j + \mathcal{O}(\varepsilon^{q+1}). \tag{5.16}$$

5.3 Proof of Theorem 3.1 for q = 2, 3, …

In this section, we carry out the induction step of the proof. We assume that Eqs. (5.15) and (5.16) hold for $0, 1, \ldots, q$ and prove that they also hold for

$q + 1$. It suffices to establish Eq. (5.15); Eq. (5.16) follows immediately from Eq. (5.15) and our discussion of the CSP condition (5.14).

Before carrying out the induction step, we derive an update formula for $T_{(q)}$. Using Eq. (5.10) with q replaced by $q + 1$, we obtain

$$T_{(q+1)} = B_{(q+1)}\tilde{A}^{(q+1)}. \tag{5.17}$$

(Here, we used the identity $(\tilde{B}_{(q+1)})^{-1} = \tilde{A}^{(q+1)}$.) Next, we use the update formulas (3.13) and (4.1) for $B_{(q+1)}$ and $\tilde{A}^{(q+1)}$, respectively, to rewrite Eq. (5.17),

$$T_{(q+1)} = \left(I - L_{(q)}\right)\left(I + U_{(q)}\right)T_{(q)}(I - \hat{U}_{(q)}). \tag{5.18}$$

Equation (5.10) also relates $A^{(q+1)}$ to $\tilde{A}^{(q+1)}$,

$$A^{(q+1)} = \tilde{A}^{(q+1)}\left(T_{(q+1)}\right)^{-1}. \tag{5.19}$$

Taking $C = \left(T_{(q)}\right)^{-1}$ in Eq. (3.19), we express $\Lambda_{(q)}$ in terms of $\tilde{\Lambda}_{(q)}$,

$$\Lambda_{(q)} = T_{(q)}\tilde{\Lambda}_{(q)}\left(T_{(q)}\right)^{-1} - T_{(q)}\frac{d\left(T_{(q)}\right)^{-1}}{dt}$$

or, equivalently,

$$\Lambda_{(q)} = T_{(q)}\tilde{\Lambda}_{(q)}\left(T_{(q)}\right)^{-1} + \left(T_{(q)}\right)^{-1}\frac{dT_{(q)}}{dt}. \tag{5.20}$$

Next, we estimate the blocks of the matrices in Eq. (5.20). The estimate of $T_{(q)}$ is given in the induction hypothesis (5.15); its inverse satisfies a similar estimate,

$$\left(T_{(q)}\right)^{-1}(\,\cdot\,,\psi_{(q)},\varepsilon) = \begin{pmatrix} I_n + \mathcal{O}(\varepsilon^2) & -\varepsilon^q T^{12}_{(q,q)} + \mathcal{O}(\varepsilon^{q+1}) \\ -\varepsilon T^{21}_{(q,1)} + \mathcal{O}(\varepsilon^2) & I_m + \mathcal{O}(\varepsilon^2) \end{pmatrix}. \tag{5.21}$$

Also, the induction hypothesis (5.16) and Theorem 4.1 guarantee that $\psi_{(q)} = \tilde{\psi}_{(q)} + \mathcal{O}(\varepsilon^{q+1})$, so the expansions of $\tilde{\Lambda}_{(q)}(y, \tilde{\psi}_{(q)}, \varepsilon)$ and $\tilde{\Lambda}_{(q)}(y, \psi_{(q)}, \varepsilon)$ are equal up to and including terms of $\mathcal{O}(\varepsilon^q)$. It follows from Eq. (4.40) that

$$\tilde{\Lambda}_{(q)}(\,\cdot\,,\psi_{(q)},\varepsilon) = \begin{pmatrix} \Lambda^{11}_{(0,0)} + \mathcal{O}(\varepsilon) & \varepsilon^q\tilde{\Lambda}^{12}_{(q,q)} + \mathcal{O}(\varepsilon^{q+1}) \\ \varepsilon\Lambda^{21}_{(0,1)} + \mathcal{O}(\varepsilon^2) & \varepsilon\tilde{\Lambda}^{22}_{(q,1)} + \mathcal{O}(\varepsilon^2) \end{pmatrix}. \tag{5.22}$$

Taking $V = T_{(q)}$ in Lemma A.2, we conclude from Eq. (5.15) that

$$DT_{(q)}g = \begin{pmatrix} \mathcal{O}(\varepsilon^3) & \mathcal{O}(\varepsilon^{q+1}) \\ \mathcal{O}(\varepsilon^2) & \mathcal{O}(\varepsilon^3) \end{pmatrix}. \tag{5.23}$$

The desired estimate of $\Lambda_{(q)}$ now follows immediately from Eqs. (5.15), (5.21), (5.22), and (5.23),

$$\Lambda_{(q)}^{11} = \Lambda_{(0,0)}^{11} + \mathcal{O}(\varepsilon), \tag{5.24}$$

$$\Lambda_{(q)}^{12} = \varepsilon^q[\tilde{\Lambda}_{(q,q)}^{12} - \Lambda_{(0,0)}^{11}T_{(q,q)}^{12}] + \mathcal{O}(\varepsilon^{q+1}), \tag{5.25}$$

$$\Lambda_{(q)}^{21} = \varepsilon[\Lambda_{(0,1)}^{21} + T_{(q,1)}^{21}\Lambda_{(0,0)}^{11}] + \mathcal{O}(\varepsilon^2), \tag{5.26}$$

$$\Lambda_{(q)}^{22} = \varepsilon\tilde{\Lambda}_{(q,1)}^{22} + \mathcal{O}(\varepsilon^2). \tag{5.27}$$

The definition (3.11) and Eqs. (5.24) and (5.25) imply that $U_{(q)} = \mathcal{O}(\varepsilon^q)$, with the leading-order coefficient given by

$$U_{(q,q)} = \left(\Lambda_{(q,0)}^{11}\right)^{-1}\Lambda_{(q,q)}^{12} = \tilde{U}_{(q,q)} - T_{(q,q)}^{12}. \tag{5.28}$$

Furthermore, the definition (3.11) and Eqs. (5.24) and (5.26) imply that

$$L_{(q)} = \Lambda_{(q)}^{21}\left(\Lambda_{(q)}^{11}\right)^{-1} = \mathcal{O}(\varepsilon). \tag{5.29}$$

Finally, we observe that, to leading order, the blocks of $T_{(q)}(\cdot, \psi_{(q)}, \varepsilon)$ are all equal to the corresponding blocks of $T_{(q)}(\cdot, \psi_{(q-1)}, \varepsilon)$. The latter are given by the induction hypothesis (5.15).

We are now ready to estimate the size of the blocks of $T_{(q+1)}(\cdot, \psi_{(q)}, \varepsilon)$.

The update formula (5.18) gives $T_{(q+1)}^{11} = T_{(q)}^{11} + U_{(q)}T_{(q)}^{21}$. According to the induction hypothesis, $T_{(q)}^{11} = I_n + \mathcal{O}(\varepsilon^2)$ and $T_{(q)}^{21} = \mathcal{O}(\varepsilon)$. Furthermore, $U_{(q)} = \mathcal{O}(\varepsilon^q)$, by Eq. (5.28). Thus, $T_{(q+1)}^{11} = I_n + \mathcal{O}(\varepsilon^2)$, as desired.

The update formula (5.18) also gives $T_{(q+1)}^{12} = T_{(q)}^{12} - T_{(q)}^{11}\tilde{U}_{(q)} + U_{(q)}T_{(q)}^{22} - U_{(q)}T_{(q)}^{21}\tilde{U}_{(q)}$. According to the induction hypothesis, $T_{(q)}^{12} = \mathcal{O}(\varepsilon^q)$, $T_{(q)}^{11} = I_n + \mathcal{O}(\varepsilon^2)$, $T_{(q)}^{21} = \mathcal{O}(\varepsilon)$, and $T_{(q)}^{22} = I_m + \mathcal{O}(\varepsilon^2)$. Furthermore, $U_{(q)} = \mathcal{O}(\varepsilon^q)$, by Eq. (5.28), and $\tilde{U}_{(q)} = \mathcal{O}(\varepsilon^q)$, by Eq. (4.23). Thus, the terms in the formula for $T_{(q+1)}^{12}$ are all at least $\mathcal{O}(\varepsilon^q)$. The same is then true for $T_{(q+1)}^{12}$. We will now show that $T_{(q+1)}^{12}$ is, in fact, at least $\mathcal{O}(\varepsilon^{q+1})$ by showing that $T_{(q+1,q)}^{12} = 0$. To leading order, the update formula for $T_{(q+1)}^{12}$ is

$$T_{(q+1,q)}^{12} = T_{(q,q)}^{12} - \tilde{U}_{(q,q)} + U_{(q,q)}. \tag{5.30}$$

Equation (5.28) implies that the right member of (5.30) vanishes. Therefore, $T_{(q+1,q)} = 0$, as desired. We emphasize again that the choice of $U_{(q)}$ is central to the working of the CSP method.

Next, the update formula (5.18) gives $T_{(q+1)}^{21} = T_{(q)}^{21} - L_{(q)}U_{(q)}T_{(q)}^{21} - L_{(q)}T_{(q)}^{11}$. According to the induction hypothesis, $T_{(q)}^{21} = \mathcal{O}(\varepsilon)$ and $T_{(q)}^{11} = I_n + \mathcal{O}(\varepsilon^2)$.

Furthermore, $U_{(q)} = \mathcal{O}(\varepsilon^q)$ and $L_{(q)} = \mathcal{O}(\varepsilon)$, by Eqs. (5.28) and (5.29). Thus, the terms in the update formula for $T^{21}_{(q+1)}$ are all at least $\mathcal{O}(\varepsilon)$. Hence, $T^{21}_{(q)}$ is also at least $\mathcal{O}(\varepsilon)$, as desired.

Lastly, the update formula (5.18) gives $T^{22}_{(q+1)} = T^{22}_{(q)} - L_{(q)}T^{12}_{(q)} - T^{21}_{(q)}\tilde{U}_{(q)} + L_{(q)}T^{11}_{(q)}\tilde{U}_{(q)} - L_{(q)}U_{(q)}T^{22}_{(q)} + L_{(q)}U_{(q)}T^{21}_{(q)}\tilde{U}_{(q)}$. According to the induction hypothesis, $T^{22}_{(q)} = I_m + \mathcal{O}(\varepsilon^2)$. The remaining terms have already been shown to be at least $\mathcal{O}(\varepsilon^2)$. Hence, $T^{22}_{(q+1)} = I_m + \mathcal{O}(\varepsilon^2)$.

The proof of Theorem 3.1 is complete.

6 The Michaelis–Menten–Henri Reaction

In this section, we apply the CSP method to the Michaelis–Menten–Henri (MMH) mechanism of enzyme kinetics to illustrate Theorems 3.1 and 4.1. We consider the planar system of ODEs for a slow variable s and a fast variable c,

$$s' = \varepsilon(-s + (s + \kappa - \lambda)c), \tag{6.1}$$
$$c' = s - (s + \kappa)c. \tag{6.2}$$

The parameters satisfy the inequalities $0 < \varepsilon \ll 1$ and $\kappa > \lambda > 0$. Only nonnegative values of s and c are relevant. The system of Eqs. (6.1)–(6.2) is of the form (2.1)–(2.2) with $m = 1$, $n = 1$, $y = s$, $z = c$, $g_1 = -s + (s + \kappa - \lambda)c$, and $g_2 = s - (s + \kappa)c$.

In the limit as $\varepsilon \downarrow 0$, the dynamics of the MMH equations are confined to the reduced slow manifold

$$\mathcal{M}_0 = \{(c, s) : c = \frac{s}{s + \kappa}, s \geq 0\}. \tag{6.3}$$

The manifold $\mathcal{M}_0$ is normally hyperbolic, so according to Theorem 2.1 there exists, for all sufficiently small ε, a slow manifold $\mathcal{M}_\varepsilon$ that is $\mathcal{O}(\varepsilon)$ close to $\mathcal{M}_0$ on any compact set. Moreover, $\mathcal{M}_\varepsilon$ is the graph of a function h_ε,

$$\mathcal{M}_\varepsilon = \{(c, s) : c = h_\varepsilon(s), s \geq 0\}, \tag{6.4}$$

and h_ε admits an asymptotic expansion $h_\varepsilon = h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \cdots$. The coefficients are found from the invariance equation,

$$s - (s + \kappa)h_\varepsilon(s) = \varepsilon h_\varepsilon'(s)(-s + (s + \kappa - \lambda)h_\varepsilon(s)). \tag{6.5}$$

The first few coefficients are

$$h_0(s) = \frac{s}{s + \kappa}, \quad h_1(s) = \frac{\kappa\lambda s}{(s + \kappa)^4}, \quad h_2(s) = \frac{\kappa\lambda s(2\kappa\lambda - 3\lambda s - \kappa s - \kappa^2)}{(s + \kappa)^7}. \tag{6.6}$$

6.1 Application of the One-Step CSP Method

Both the one-step and two-step CSP methods start from the same initial basis. We choose the stoichiometric vectors as the basis vectors, so

$$A^{(0)} = (A_1^{(0)}, A_2^{(0)}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B_{(0)} = \begin{pmatrix} B_{(0)}^1 \\ B_{(0)}^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (6.7)$$

The CSP condition $B_{(0)}^1 g = 0$ is satisfied if $c = h_0(s)$, so the CSP manifolds $\tilde{\mathcal{K}}_\varepsilon^{(0)}$ and $\mathcal{K}_\varepsilon^{(0)}$ coincide with $\mathcal{M}_0$. With this choice of initial basis, we have

$$\Lambda_{(0)} = B_{(0)}(Dg)A^{(0)} = \begin{pmatrix} -(s+\kappa) & -(c-1) \\ \varepsilon(s+\kappa-\lambda) & \varepsilon(c-1) \end{pmatrix}. \quad (6.8)$$

First iteration. At any point (s, c), we have

$$\tilde{A}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{c-1}{s+\kappa} \end{pmatrix}, \quad \tilde{B}_{(1)} = \begin{pmatrix} \frac{c-1}{s+\kappa} & 1 \\ 1 & 0 \end{pmatrix}. \quad (6.9)$$

On $\tilde{\mathcal{K}}_\varepsilon^{(0)}$, these expressions reduce to

$$\tilde{A}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & \frac{\kappa}{(s+\kappa)^2} \end{pmatrix}, \quad \tilde{B}_{(1)} = \begin{pmatrix} \frac{-\kappa}{(s+\kappa)^2} & 1 \\ 1 & 0 \end{pmatrix}. \quad (6.10)$$

The CSP condition,

$$\tilde{B}_{(1)}^1 g = s - (s+\kappa)c - \varepsilon\frac{\kappa(-s + (s+\kappa-\lambda)c)}{(s+\kappa)^2} = 0, \quad (6.11)$$

is satisfied if

$$c = \frac{s}{s+\kappa} + \varepsilon\frac{\kappa\lambda s}{(s+\kappa)^4} - \varepsilon^2\frac{\kappa^2\lambda s(s+\kappa-\lambda)}{(s+\kappa)^7} + \mathcal{O}(\varepsilon^3). \quad (6.12)$$

Comparing this result with Eq. (6.6), we see that the asymptotic expansions of $\tilde{\mathcal{K}}_\varepsilon^{(1)}$ and $\mathcal{M}_\varepsilon$ coincide up to and including $\mathcal{O}(\varepsilon)$ terms, in accordance with Theorem 4.1 for $q = 1$; however, the $\mathcal{O}(\varepsilon^2)$ terms differ at this stage.

Second iteration. The blocks of $\tilde{\Lambda}_{(1)}$ are

$$\tilde{\Lambda}_{(1)}^{11} = -(s+\kappa) + \varepsilon\frac{(s+\kappa-\lambda)(c-1)}{s+\kappa}, \quad (6.13)$$

$$\tilde{\Lambda}_{(1)}^{12} = \frac{s}{s+\kappa} - c + \varepsilon\frac{(c-1)[\lambda(c-1) - (-s + (s+\kappa-\lambda)c)]}{(s+\kappa)^2}, \quad (6.14)$$

$$\tilde{\Lambda}_{(1)}^{21} = \varepsilon(s+\kappa-\lambda), \quad \tilde{\Lambda}_{(1)}^{22} = \varepsilon\frac{\lambda(c-1)}{s+\kappa}. \quad (6.15)$$

On $\tilde{\mathcal{K}}_\varepsilon^{(1)}$, the blocks reduce to

$$\tilde{\Lambda}_{(1)}^{11} = -(s+\kappa) - \varepsilon\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^2} + \varepsilon^2\frac{\kappa\lambda s(s+\kappa-\lambda)}{(s+\kappa)^5}, \qquad (6.16)$$

$$\tilde{\Lambda}_{(1)}^{12} = \varepsilon\frac{\kappa\lambda(\kappa-2s)}{(s+\kappa)^4} + \varepsilon^2\frac{\kappa\lambda s(2\kappa(s+\kappa-2\lambda)+\lambda s)}{(s+\kappa)^7}, \qquad (6.17)$$

$$\tilde{\Lambda}_{(1)}^{21} = \varepsilon(s+\kappa-\lambda), \qquad \tilde{\Lambda}_{(1)}^{22} = -\varepsilon\frac{\kappa\lambda}{(s+\kappa)^2} + \varepsilon^2\frac{\kappa\lambda^2 s}{(s+\kappa)^5}. \quad (6.18)$$

The second update is

$$\tilde{A}_1^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad (6.19)$$

$$\tilde{A}_2^{(2)} = \begin{pmatrix} 1 \\ \frac{\kappa}{(s+\kappa)^2} \end{pmatrix} + \varepsilon\begin{pmatrix} 0 \\ \frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5} \end{pmatrix}$$
$$+ \varepsilon^2\begin{pmatrix} 0 \\ \frac{\kappa\lambda[\kappa(5s-\kappa)(s+\kappa-\lambda)+\lambda s(s-2\kappa)]}{(s+\kappa)^8} \end{pmatrix} + \mathcal{O}(\varepsilon^3), \qquad (6.20)$$

$$\tilde{B}_{(2)}^1 = \left(-\frac{\kappa}{(s+\kappa)^2}, 1\right) + \varepsilon\left(-\frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5}, 0\right)$$
$$+ \varepsilon^2\left(-\frac{\kappa\lambda[\kappa(5s-\kappa)(s+\kappa-\lambda)+\lambda s(s-2\kappa)]}{(s+\kappa)^8}, 0\right)$$
$$+ \mathcal{O}(\varepsilon^3), \qquad (6.21)$$

$$\tilde{B}_{(2)}^2 = (1,0). \qquad (6.22)$$

The CSP condition

$$\tilde{B}_{(2)}^1 g = s - (s+\kappa)c - \varepsilon\frac{\kappa(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^2}$$
$$+ \varepsilon^2\kappa\lambda\frac{(3s-\kappa)(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^5} + \mathcal{O}(\varepsilon^3)$$
$$= 0, \qquad (6.23)$$

is satisfied if

$$c = \frac{s}{s+\kappa} + \varepsilon\frac{\kappa\lambda s}{(s+\kappa)^4} + \varepsilon^2\frac{\kappa\lambda s(2\kappa\lambda-3\lambda s-\kappa s-\kappa^2)}{(s+\kappa)^7} + \mathcal{O}(\varepsilon^3). \quad (6.24)$$

Comparing this result with Eq. (6.6), we see that the asymptotic expansions of $\tilde{\mathcal{K}}_\varepsilon^{(2)}$ and $\mathcal{M}_\varepsilon$ coincide up to and including $\mathcal{O}(\varepsilon^2)$ terms, in accordance with Theorem 4.1 for $q=2$.

6.2 Application of the Full CSP Method

First iteration. At any point (s, c), we have

$$A_1^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \varepsilon \frac{s + \kappa - \lambda}{s + \kappa} \begin{pmatrix} 1 \\ -\frac{c-1}{s+\kappa} \end{pmatrix}, \qquad A_2^{(1)} = \begin{pmatrix} 1 \\ -\frac{c-1}{s+\kappa} \end{pmatrix}, \qquad (6.25)$$

$$B_{(1)}^1 = \begin{pmatrix} \frac{c-1}{s+\kappa}, 1 \end{pmatrix}, \qquad B_{(1)}^2 = (1, 0) + \varepsilon \frac{s + \kappa - \lambda}{s + \kappa} \begin{pmatrix} \frac{c-1}{s+\kappa}, 1 \end{pmatrix}. \qquad (6.26)$$

On $\mathcal{K}_\varepsilon^{(0)}$, these quantities reduce to

$$A_1^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \varepsilon \frac{s + \kappa - \lambda}{s + \kappa} \begin{pmatrix} 1 \\ \frac{\kappa}{(s+\kappa)^2} \end{pmatrix}, \qquad A_2^{(1)} = \begin{pmatrix} 1 \\ \frac{\kappa}{(s+\kappa)^2} \end{pmatrix}, \qquad (6.27)$$

$$B_{(1)}^1 = \begin{pmatrix} -\frac{\kappa}{(s+\kappa)^2}, 1 \end{pmatrix}, \quad B_{(1)}^2 = (1, 0) + \varepsilon \frac{s + \kappa - \lambda}{s + \kappa} \begin{pmatrix} -\frac{\kappa}{(s+\kappa)^2}, 1 \end{pmatrix}. \qquad (6.28)$$

The matrix relating $B_{(1)}$ to its one-step counterpart $\tilde{B}_{(1)}$ is

$$T_{(1)} = \begin{pmatrix} 1 & 0 \\ \varepsilon \frac{s+\kappa-\lambda}{s+\kappa} & 1 \end{pmatrix}, \qquad (6.29)$$

so $T_{(1)}$ is indeed of the form (5.15) on $\mathcal{K}_\varepsilon^{(0)}$.

Equations (6.10) and (6.28) imply that $B_{(1)} = \tilde{B}_{(1)}$, so the CSP condition yields $\psi_{(1)} = \tilde{\psi}_{(1)}$. Thus, after one iteration, the full CSP method also finds the expansion of $\mathcal{M}_\varepsilon$ up to and including $\mathcal{O}(\varepsilon)$ terms.

Second iteration. The blocks of $\Lambda_{(1)}$ are

$$\Lambda_{(1)}^{11} = -(s + \kappa) + \varepsilon \frac{(s + \kappa - \lambda)}{s + \kappa} \left[(c - 1) + (c - \frac{s}{s + \kappa}) \right]$$
$$+ \varepsilon^2 \frac{(c - 1)(s + \kappa - \lambda)}{(s + \kappa)^3} [-\lambda(c - 1) + (-s + (s + \kappa - \lambda)c)], \qquad (6.30)$$

$$\Lambda_{(1)}^{12} = \frac{s}{s + \kappa} - c + \varepsilon \frac{c - 1}{(s + \kappa)^2} [\lambda(c - 1) - (-s + (s + \kappa - \lambda)c)], \qquad (6.31)$$

$$\Lambda_{(1)}^{21} = \frac{\varepsilon^2}{(s + \kappa)^2} \left[(c - 1)(s + \kappa - \lambda)(s + \kappa - 2\lambda) \right.$$
$$\left. + \lambda(-s + (s + \kappa - \lambda)c) + (s + \kappa - \lambda)^2 \left(c - \frac{s}{s + \kappa} \right) \right], \qquad (6.32)$$

$$\Lambda_{(1)}^{22} = \frac{\varepsilon}{s+\kappa}\left[\lambda(c-1)+(s+\kappa-\lambda)(\frac{s}{s+\kappa}-c)\right]$$
$$+\varepsilon^2\frac{(c-1)(s+\kappa-\lambda)}{(s+\kappa)^3}\left[\lambda(c-1)-(-s+(s+\kappa-\lambda)c)\right], \quad (6.33)$$

with remainders of $\mathcal{O}(\varepsilon^3)$. On $\mathcal{K}_\varepsilon^{(1)}$, the blocks reduce to

$$\Lambda_{(1)}^{11} = -(s+\kappa)-\varepsilon\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^2}+\varepsilon^2\frac{\kappa\lambda(s+\kappa-\lambda)(3s-\kappa)}{(s+\kappa)^5}, \quad (6.34)$$

$$\Lambda_{(1)}^{12} = \varepsilon\frac{\kappa\lambda(\kappa-2s)}{(s+\kappa)^4}+\varepsilon^2\frac{\kappa\lambda s(2\kappa(s+\kappa-2\lambda)+\lambda s)}{(s+\kappa)^7}, \quad (6.35)$$

$$\Lambda_{(1)}^{21} = -\varepsilon^2\frac{\kappa(s+\kappa-\lambda)(s+\kappa-2\lambda)+\lambda^2 s}{(s+\kappa)^3}, \quad (6.36)$$

$$\Lambda_{(1)}^{22} = -\varepsilon\frac{\kappa\lambda}{(s+\kappa)^2}-\varepsilon^2\frac{\kappa\lambda((2s-\kappa)(s+\kappa-\lambda)-\lambda s)}{(s+\kappa)^5}, \quad (6.37)$$

with errors of $\mathcal{O}(\varepsilon^3)$. The result of the second iteration is

$$A_{11}^{(2)} = -\varepsilon\frac{s+\kappa-\lambda}{s+\kappa}+\varepsilon^2\frac{\kappa(s+\kappa-2\lambda)(s+\kappa-\lambda)+\lambda^2 s}{(s+\kappa)^4}, \quad (6.38)$$

$$A_{12}^{(2)} = 1+\varepsilon^2\frac{\kappa\lambda(2s-\kappa)(s+\kappa-\lambda)}{(s+\kappa)^6}, \quad (6.39)$$

$$A_{21}^{(2)} = 1-\varepsilon\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^3}$$
$$+\varepsilon^2\frac{(s+\kappa-\lambda)(\kappa^2(s+\kappa-2\lambda)+\kappa\lambda s)+\kappa\lambda^2 s}{(s+\kappa)^6}, \quad (6.40)$$

$$A_{22}^{(2)} = \frac{\kappa}{(s+\kappa)^2}+\varepsilon\frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5}$$
$$+\varepsilon^2\frac{\kappa^2\lambda(7s-2\kappa)(s+\kappa-\lambda)+\kappa\lambda^2 s(s-2\kappa)}{(s+\kappa)^8}, \quad (6.41)$$

$$B_{(2)}^{11} = \frac{-\kappa}{(s+\kappa)^2}-\varepsilon\frac{\kappa\lambda(\kappa-3s)}{(s+\kappa)^5}$$
$$-\varepsilon^2\frac{\kappa^2\lambda(7s-2\kappa)(s+\kappa-\lambda)+\kappa\lambda^2 s(s-2\kappa)}{(s+\kappa)^8}, \quad (6.42)$$

$$B_{(2)}^{12} = 1+\varepsilon^2\frac{\kappa\lambda(2s-\kappa)(s+\kappa-\lambda)}{(s+\kappa)^6}, \quad (6.43)$$

$$B_{(2)}^{21} = 1-\varepsilon\frac{\kappa(s+\kappa-\lambda)}{(s+\kappa)^3}$$
$$+\varepsilon^2\frac{(s+\kappa-\lambda)(\kappa^2(s+\kappa-2\lambda)+\kappa\lambda s)+\kappa\lambda^2 s}{(s+\kappa)^6}, \quad (6.44)$$

$$B_{(2)}^{22} = \varepsilon\frac{s+\kappa-\lambda}{s+\kappa} - \varepsilon^2\frac{\kappa(s+\kappa-\lambda)(s+\kappa-2\lambda)+\lambda^2 s}{(s+\kappa)^4}, \qquad (6.45)$$

up to and including terms of $\mathcal{O}(\varepsilon^2)$. Also, on $\mathcal{K}_\varepsilon^{(1)}$,

$$
\begin{aligned}
T_{(2)}^{11} &= 1 + \varepsilon^2\frac{\kappa\lambda(s+\kappa-\lambda)(2s-\kappa)}{(s+\kappa)^6}, \\
T_{(2)}^{12} &= 0, \\
T_{(2)}^{21} &= \varepsilon\frac{(s+\kappa-\lambda)}{s+\kappa} - \varepsilon^2\frac{\kappa(s+\kappa-\lambda)(s+\kappa-2\lambda)+\lambda^2 s}{(s+\kappa)^4}, \\
T_{(2)}^{22} &= 1 - \varepsilon^2\frac{\kappa\lambda(2s-\kappa)(s+\kappa-\lambda)}{(s+\kappa)^8},
\end{aligned}
$$

with remainders of $\mathcal{O}(\varepsilon^3)$. Thus, $T_{(2)}$ is indeed of the form (5.15) on $\mathcal{K}_\varepsilon^{(1)}$.

The CSP condition

$$
\begin{aligned}
B_{(2)}^1 g &= s - (s+\kappa)c - \varepsilon\frac{\kappa(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^2} \\
&\quad + \varepsilon^2\kappa\lambda\left(\frac{(3s-\kappa)(-s+(s+\kappa-\lambda)c)}{(s+\kappa)^5}\right. \\
&\quad\qquad\left. + \frac{(2s-\kappa)(s+\kappa-\lambda)(s-(s+\kappa)c)}{(s+\kappa)^6}\right) \\
&\quad + \mathcal{O}(\varepsilon^3) \\
&= 0, \qquad\qquad (6.46)
\end{aligned}
$$

is satisfied if

$$c = \frac{s}{s+\kappa} + \varepsilon\frac{\kappa\lambda s}{(s+\kappa)^4} + \varepsilon^2\frac{\kappa\lambda s(2\kappa\lambda - 3\lambda s - \kappa s - \kappa^2)}{(s+\kappa)^7} + \mathcal{O}(\varepsilon^3). (6.47)$$

Therefore, after two iterations, the full CSP method finds the expansion of $\mathcal{M}_\varepsilon$ up to and including $\mathcal{O}(\varepsilon^2)$ terms.

6.3   The Second Step and the Fast Fibers of $\mathcal{M}_\varepsilon$

The preceding analysis of the full CSP method shows that, at the qth iteration, the second step alters only the terms of $\mathcal{O}(\varepsilon^{q+1})$, leaving the terms of $\mathcal{O}(1)$ through $\mathcal{O}(\varepsilon^q)$ invariant. Here, we observe that the second step also plays a constructive role for the dynamics near the slow manifold. As can be seen in the MMH example, the second step yields the asymptotic expansions of

the tangent spaces of the fast fibers at their basepoints up to and including terms of $\mathcal{O}(\varepsilon^{q+1})$, at least for $q = 0, 1$, and 2. This additional information is contained in the columns of $A_1^{(q)}$. We remark here that this property is not shared by the one-step CSP method, since the columns of $\tilde{A}_1^{(q)}$ remain tangent to the fast fibers at their basepoints only to leading order after each iteration. Details about the fast fibers and their tangent spaces will be presented in a future publication.

7   Relation between CSPM and ILDM

The CSP iteration procedure is designed to diagonalize the Lie bracket $[\,\cdot\,, g]$. At each iteration, the then-current basis is updated in such a way that $[\,\cdot\,, g]$ is block-diagonalized to the next-higher order in ε. Thus, each iteration improves the quality of the basis of the orthogonal complement of the tangent space. The CSPM is defined as the locus of points where the vector field is orthogonal to that orthogonal complement.

The ILDM method works, instead, with the Jacobian, Dg, of Eqs.(2.3)–(2.4). A Schur decomposition transforms Dg into upper triangular form,

$$Dg = QNQ', \quad N = \begin{pmatrix} N_s & N_{sf} \\ 0 & N_f \end{pmatrix}, \tag{7.1}$$

where $Q = (Q_s \; Q_f)$ is unitary. The eigenvalues of Dg appear on the diagonal of N in descending order of their real parts, from least negative in the upper left to most negative in the lower right. The first m Schur vectors (the columns of Q_s) form an orthogonal basis of the slow subspace and the remaining n Schur vectors (the columns of Q_f) an orthogonal basis of the orthogonal complement of the slow subspace. The vector field g is entirely in the slow subspace if it is orthogonal to this orthogonal complement—that is, if

$$Q_f' g = 0. \tag{7.2}$$

This equation defines the ILDM; see [11, Section 3].

As we showed in [11], the ILDM is only a first-order approximation to $\mathcal{M}_\varepsilon$. The error is always $\mathcal{O}(\varepsilon^2)$ unless $\mathcal{M}_0$ is linear. The error can be traced back to the choice of the operator. The tangent space is a left-invariant subspace of the Jacobian only to leading order, so putting Dg in upper triangular form yields the orthogonal complement only to leading order. Since the linearized system is only an approximation of the original ODEs (2.1)–(2.2), this choice does not produce an exact result unless g is linear. The success of the CSP

method in approximating the slow manifold is due to the fact that the ODEs for the amplitudes f are equivalent to the ODEs (2.1)–(2.2). That is, the full nonlinearity is retained.

The time-derivative term in the definition (3.6) must be included in the evaluation of Λ; otherwise, the accuracy of the CSP method is compromised. In fact, such an omission results in implementing the ILDM rather than the CSP method, which may be seen as follows. With our initial choice of a point-independent basis $A^{(0)}$, the matrix $\Lambda_{(0)}$ is similar to Dg; see Eq. (3.10). The omission of the term $(dB_{(q)}/dt)A^{(q)}$ in the calculation of $\Lambda_{(q)}$, for $q = 1, 2, \dots$, would lead to the formula $\Lambda_{(q)} = (I + \tilde{P}_{(q)})B_{(0)}(Dg)A^{(0)}(I - \tilde{P}_{(q)})$, which would imply that $\Lambda_{(q)}$ is similar to Dg. Therefore, the one-step CSP method would put Dg, rather than Λ, in lower-triangular form, just like the ILDM method. After the second iteration, one would make an error (proportional to the curvature of $\mathcal{M}_0$) at $\mathcal{O}(\varepsilon^2)$, which subsequent iterations would not remove. The MMH example in Section 6 illustrates these observations.

ACKNOWLEDGMENTS

A   Auxiliary Lemmas

Lemma A.1 *The quantity Dh_q is given by the formula*

$$
\begin{aligned}
Dh_q &= -((D_z g_2)_0)^{-1}\left[(D_y g_2)_q + \sum_{i=0}^{q-1}(D_z g_2)_{q-i}\, Dh_i - \sum_{\ell=0}^{q-1} D^2 h_\ell g_{1, q-\ell-1}\right.\\
&\left. - \sum_{\ell=0}^{q-1}(Dh_\ell)(D_y g_1)_{q-1-\ell} - \sum_{i=0}^{q-1}\sum_{\ell=0}^{q-1-i} Dh_\ell (D_z g_1)_{q-1-i-\ell} Dh_i\right]. \ (\mathrm{A.3})
\end{aligned}
$$

33

Proof. The coefficient h_q is found from the $\mathcal{O}(\varepsilon^q)$ terms in the invariance equation (2.10),

$$g_{2,q} = \sum_{\ell=0}^{q-1} (Dh_\ell) g_{1,q-\ell-1}. \tag{A.4}$$

Taking the total derivative with respect to y of both sides of (A.4), we find

$$\frac{d}{dy} g_{2,q} = \sum_{\ell=0}^{q-1} (D^2 h_\ell) g_{1,q-\ell-1} + \sum_{\ell=0}^{q-1} (Dh_\ell) \frac{d}{dy} g_{1,q-\ell-1}. \tag{A.5}$$

The operations of taking the total derivative with respect to y and expanding with respect to ε commute, because the Fenichel theory guarantees C^r smoothness in ε and y for each r. Therefore,

$$\frac{d}{dy} g_{2,q} = \left(\frac{dg_2}{dy}\right)_q = (D_y g_2)_q + \sum_{i=0}^{q} (D_z g_2)_{q-i} (Dh_i), \tag{A.6}$$

$$\frac{d}{dy} g_{1,q-1-\ell} = \left(\frac{dg_1}{dy}\right)_{q-1-\ell} = (D_y g_1)_{q-1-\ell} + \sum_{i=0}^{q-1-\ell} (D_z g_1)_{q-1-\ell-i} (Dh_i). \tag{A.7}$$

Substituting Eqs. (A.6) and (A.7) into Eq. (A.5), we obtain

$$(D_y g_2)_q + \sum_{i=0}^{q} (D_z g_2)_{q-i} \, Dh_i = \sum_{\ell=0}^{q-1} (D^2 h_\ell) g_{1,q-\ell-1} + \sum_{\ell=0}^{q-1} (Dh_\ell)(D_y g_1)_{q-1-\ell}$$

$$+ \sum_{\ell=0}^{q-1} \sum_{i=0}^{q-1-\ell} (Dh_\ell)(D_z g_1)_{q-1-\ell-i}(Dh_i). \tag{A.8}$$

Separating the $i = q$ term in the sum of the left member, changing the order of summation in the last sum of the right member, and solving for Dh_q, we obtain Eq. (A.3). ∎

Lemma A.2 *Let V be a matrix-valued function of y, z, and ε that, together with its first-order derivatives, is smooth and $\mathcal{O}(1)$ as $\varepsilon \downarrow 0$. If $z = \psi_{(q)}(y, \varepsilon)$ and*

$$V(\cdot, \psi_{(q)}, \varepsilon) = \sum_{\ell=0}^{q} \varepsilon^\ell V_\ell + \mathcal{O}(\varepsilon^{q+1}), \quad g_1(\cdot, \psi_{(q)}, \varepsilon) = \sum_{\ell=0}^{q} \varepsilon^\ell g_{1,\ell} + \mathcal{O}(\varepsilon^{q+1}), \tag{A.9}$$

then,

$$\frac{dV}{dt}(\cdot, \psi_{(q)}, \varepsilon) = \sum_{i=0}^{q} \varepsilon^{i+1} \sum_{\ell=0}^{i} \frac{dV_\ell}{dy} g_{1,i-\ell} + \mathcal{O}(\varepsilon^{q+1}). \tag{A.10}$$

34

Proof. A direct computation gives

$$\frac{dV}{dt} = (DV)g = \varepsilon(D_yV)g_1 + (D_zV)g_2, \tag{A.11}$$

where all the terms are evaluated at $(y, \psi_{(q)}(y, \varepsilon), \varepsilon)$. Since $\psi_{(q)}$ approximates the slow manifold up to and including $\mathcal{O}(\varepsilon^q)$ terms,

$$g_1(\cdot, \psi_{(q)}, \varepsilon) = g_1(\cdot, h_\varepsilon, \varepsilon) + \mathcal{O}(\varepsilon^{q+1}), \tag{A.12}$$

$$g_2(\cdot, \psi_{(q)}, \varepsilon) = g_2(\cdot, h_\varepsilon, \varepsilon) + \mathcal{O}(\varepsilon^{q+1}), \tag{A.13}$$

and also

$$D\psi_{(q)} = Dh_\varepsilon + \mathcal{O}(\varepsilon^{q+1}). \tag{A.14}$$

Using Eqs. (2.10), (A.12), and (A.14), we rewrite Eq. (A.13) as

$$g_2(\cdot, \psi_{(q)}, \varepsilon) = \varepsilon(D\psi_{(q)})g_1(\cdot, \psi_{(q)}, \varepsilon) + \mathcal{O}(\varepsilon^{q+1}). \tag{A.15}$$

Equation (A.15) is an equation for $\mathcal{K}_\varepsilon^{(q)}$. We recast it so the right member involves a total derivative with respect to y,

$$(DV)g = \varepsilon\left(D_yV + D_zVD\psi_{(q)}\right)g_1 + \mathcal{O}(\varepsilon^{q+1}) = \varepsilon\frac{dV}{dy}g_1 + \mathcal{O}(\varepsilon^{q+1}) \tag{A.16}$$

or, expanding in powers of ε,

$$(DV)g = \sum_{i=0}^{q} \varepsilon^{i+1} \sum_{\ell=0}^{i} \left(\frac{dV}{dy}\right)_\ell g_{1,i-\ell} + \mathcal{O}(\varepsilon^{q+1}). \tag{A.17}$$

The operations of taking the total derivative with respect to y and expanding with respect to ε commute, so $(dV/dy)_\ell = dV_\ell/dy$ and Eq. (A.10) follows. ∎

References

[1] X. Cabré, E. Fontich, R. de la Llave, The parameterization method for invariant manifolds I: Manifolds associated to non-resonant subspaces, preprint (2002)

[2] M. J. Davis and R. T. Skodje, Geometric investigation of low-dimensional manifolds in systems approaching equilibrium, *J. Chem. Phys.* **111** (1999) 859–874

[3] A. Fehrst, *Enzyme Structure and Mechanisms*, 2nd ed., W. F. Freeman, New York, 1975

[4] N. Fenichel, Geometric singular perturbation theory for ordinary differential equations, *J. Diff. Eq.* **31** (1979) 53–98

[5] S. J. Fraser, The steady state and equilibrium approximations: A general picture, *J. Chem. Phys.* **88** (1988) 4732–4738

[6] A. N. Gorban and I. V. Karlin, Method of invariant manifolds for chemical kinetics, *arXiv:cond-mat/0207231* (9 Jul 2002)

[7] D. A. Goussis and S. H. Lam, A study of homogeneous methanol oxidation kinetics using CSP, in: *Twenty-Fourth Symposium (International) on Combustion, The University of Sydney, Sydney, Australia, July 5–10, 1992*, The Combustion Institute, Pittsburgh, 1992, pp. 113–120

[8] M. Hadjinicolaou and D. A. Goussis, Asymptotic solutions of stiff PDEs with the CSP method: The reaction diffusion equation, *SIAM J. Sci. Comput.* **20** (1999) 781–810

[9] F. Heineken, H. Tsuchiya, and R. Aris, On the mathematical status of the pseudo-steady-state hypothesis of biochemical kinetics, *Math. Biosci.* **1** (1967) 95–113

[10] C. K. R. T. Jones, Geometric singular perturbation theory, in: *Dynamical Systems, Montecatini Terme*, L. Arnold, Lecture Notes in Mathematics, **1609**, Springer-Verlag, Berlin, 1994, pp. 44–118

[11] H. G. Kaper and T. J. Kaper, Asymptotic analysis of two reduction methods for systems of chemical reactions, *Physica D* **165** (2002), 66–93

[12] S. H. Lam, Using CSP to understand complex chemical kinetics, *Combust. Sci. Tech.* **89** (1993) 375–404

[13] S. H. Lam and D. A. Goussis, Understanding complex chemical kinetics with computational singular perturbation, in *Twenty-Second Symposium (International) on Combustion, The University of Washington, Seattle, Washington, August 14–19, 1988,* The Combustion Institute, Pittsburgh, 1988, pp. 931–941

[14] S. H. Lam and D. A. Goussis, Conventional asymptotics and computational singular perturbation theory for simplified kinetics modeling, in *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, M. Smooke, ed., Lecture Notes in Physics **384**, Springer-Verlag, New York, 1991, Chapter 10

[15] S. H. Lam and D. A. Goussis, The CSP method for simplifying kinetics, *Internat. J. Chem. Kin.* **26** (1994) 461–486

[16] T. F. Lu, Y. G. Ju, and C. K. Law, Complex CSP for chemistry reduction and analysis, *Combustion and Flame* **126** (2001) 1445–1455

[17] U. Maas and S. B. Pope, Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space, *Combustion and Flame* **88** (1992) 239–264

[18] A. Massias, D. Diamantis, E. Mastorakos, and D. Goussis, Global reduced mechanisms for methane and hydrogen combustion with nitric oxide formation constructed with CSP data, *Combust. Theory Modelling* **3** (1999) 233–257

[19] A. Massias and D. A. Goussis, On the manifold of stiff reaction-diffusion PDE's: The effects of diffusion, preprint (2001)

[20] M. Massot, Singular perturbation analysis for the reduction of complex chemistry in gaseous mixtures using the entropic structure, *Discr. Cont. Dyn. Systems–Series B*, **2** (2002) 433–456

[21] P. J. Olver, *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics, **107**, Springer-Verlag, New York, 1986

[22] R. E. O'Malley, Jr., *Singular Perturbation Methods for Ordinary Differential Equations*, Springer-Verlag, New York, 1991

[23] B. O. Palsson, On the dynamics of the irreversible Michaelis–Menten reaction mechanism, *Chem. Eng. Sci.* **42** (1987) 447–458

[24] B. O. Palsson and E. N. Lightfoot, Mathematical modelling of dynamics and control in metabolic networks. I. On Michaelis–Menten kinetics, *J. Theor. Bio.* **111** (1984) 273–302

[25] M. R. Roussel and S. J. Fraser, Geometry of the steady-state approximation: perturbation and accelerated convergence methods, *J. Chem. Phys.* **93** (1990) 1072–1081

[26] M. Valorani and D. A. Goussis, Explicit time-scale splitting algorithm for stiff problems: Auto-ignition of gaseous mixtures behind a steady shock, *J. Comp. Phys.* **169** (2001) 44–79

[27] M. Valorani, D. A. Goussis, and H. Najm, personal communication (2002).

Corresponding author:

Hans G. Kaper
Division of Mathematical Sciences
National Science Foundation
4201 Wilson Boulevard, Suite 1025
Arlington, VA 22230

Authors' e-mail addresses:

azagaris@math.bu.edu
kaper@mcs.anl.gov; hkaper@nsf.gov
tasso@math.bu.edu