

SUSANNA KUBE, PETER DEUFLHARD

**Errata on  
“Robust Perron Cluster Analysis in  
Conformation Dynamics”**

Errata on  
**Robust Perron Cluster Analysis in  
Conformation Dynamics**  
by Peter Deuffhard and Marcus Weber

Susanna Kube      Peter Deuffhard

December 5, 2006

**Abstract**

In the recent paper [3] as well as in the preceding ZIB report ZR-03-19 [2], one of the authors presented an  $O(\epsilon^2)$  perturbation result for the eigenvectors of a generalized symmetric stochastic matrix, where the parameter  $\epsilon$  characterized the departure of the perturbed matrix from a completely decomposable Markov chain. Due to some erroneous interchange of indices, the proof has turned out to be incorrect. Here, we give the corrected results.

## Introduction

In the recent papers [2, 3], a new cluster algorithm for nearly completely decomposable Markov chains had been introduced, the Robust Perron Cluster Cluster Analysis (named PCCA+). Upon characterizing the departure from completely decomposable chains by  $\epsilon$ , the robustness of the method had been justified by an  $O(\epsilon^2)$  perturbation result for the eigenvectors of a generalized symmetric stochastic matrix. Unfortunately, the proof of that result has turned out to be incorrect, due to some unlucky interchange of indices. However, it can still be shown that the  $O(\epsilon)$  bound for metastability, which was originally derived for a hard  $\{0, 1\}$ -clustering (PCCA), carries over to PCCA+. In the following, we indicate where the error in the previous proof had occurred and how the results must be corrected. In order to facilitate the orientation, we kept the names of the sections from [2, 3].

## 1 Perron cluster eigenproblem

**Nearly uncoupled Markov chains.** Let us shortly repeat the assumptions which preceded the proof of the perturbation result.

Denote the number of nearly uncoupled Markov chains by  $k$ . In this case the transition matrix  $\tilde{T}$  will be block diagonally dominant after suitable permutation. As a perturbation of the  $k$ -fold Perron root  $\lambda = 1$ , a Perron cluster of eigenvalues

$$\tilde{\lambda}_1 = 1, \quad \tilde{\lambda}_2 = 1 - O(\epsilon), \quad \dots \quad \tilde{\lambda}_k = 1 - O(\epsilon),$$

will arise, where  $\epsilon > 0$  denotes some perturbation parameter, which we here scale as

$$\epsilon = 1 - \tilde{\lambda}_2. \quad (1.1)$$

Let formal  $\epsilon$ -expansions be introduced for the stochastic matrix as

$$\tilde{T}(\epsilon) = T + \epsilon T^{(1)} + O(\epsilon^2), \quad (1.2)$$

and for the Perron cluster eigenvectors  $\tilde{X} = [\tilde{X}_1, \dots, \tilde{X}_k] \in \mathbb{R}^{N \times k}$  as

$$\tilde{X}_i(\epsilon) = X_i + \epsilon X_i^{(1)} + O(\epsilon^2). \quad (1.3)$$

In [1], the result

$$X_i^{(1)} = \underbrace{\sum_{j=1}^k b_{ji} X_j}_{(I)} + \underbrace{\sum_{j=k+1}^N \frac{1}{1 - \lambda_j} \Pi_j T^{(1)} X_i}_{(II)} \quad (1.4)$$

has been obtained using projections  $\Pi_j$  as defined in the book of Kato [5]. Obviously, the term (I) represents just shifts of the locally constant levels to be

associated with the almost invariant sets. In [2, 3], the term (II) had wrongly been dropped. Thus, the corresponding Lemma therein (1.1 in [2], 2.1 in [3]), which claimed that  $X^{(1)} = \chi B$ , must be abandoned.

In order to see the mistake, let us shortly revisit the lines at the “proof”. Starting from

$$\tilde{T}(\epsilon)\tilde{X}_i(\epsilon) = \tilde{\lambda}_i(\epsilon)\tilde{X}_i(\epsilon), \quad i = 1, \dots, N$$

and inserting the  $\epsilon$ -expansions leads to

$$TX_i = X_i \quad \text{for } i = 1, \dots, k, \quad (1.5)$$

and

$$T^{(1)}X_i = (I - T)X_i^{(1)} - X_i\delta\lambda_i, \quad i = 1, \dots, k. \quad (1.6)$$

Hence, for  $j = k + 1, \dots, N$  and  $i = 1, \dots, k$ , one obtains

$$\begin{aligned} \langle X_j, T^{(1)}X_i \rangle_\pi &= \langle X_j, (I - T)X_i^{(1)} - \delta\lambda_i X_i \rangle_\pi \\ &= \langle (I - T)X_j, X_i^{(1)} \rangle_\pi - \delta\lambda_i \langle X_j, X_i \rangle_\pi. \end{aligned}$$

The last term above vanishes due to the  $\pi$ -orthogonality of the unperturbed eigenvectors. Using  $TX_j = \lambda_j X_j$  results in

$$\langle X_j, T^{(1)}X_i \rangle_\pi = (1 - \lambda_j) \langle X_j, X_i^{(1)} \rangle_\pi.$$

The last line is the place where the error occurred in the proof. Due to an index permutation it was assumed that  $TX_j = X_j$  which is only satisfied for  $j = 1, \dots, k$  but not for  $j = k + 1, \dots, N$ .

With the above result, equation (9) from [1] for  $i = 1, \dots, k$  reads correctly

$$X_i(\epsilon) = \sum_{j=1}^k \tilde{\alpha}_{ij} \chi_j + \epsilon \sum_{j=k+1}^N \frac{1}{1 - \lambda_j} \Pi_j T^{(1)}X_i + O(\epsilon^2) \quad (1.7)$$

and can thus be rewritten as

$$X_i(\epsilon) = \sum_{j=1}^k \tilde{\alpha}_{ij} \chi_j + \epsilon \sum_{j=k+1}^N \langle X_j, X_i^{(1)} \rangle_\pi X_j + O(\epsilon^2). \quad (1.8)$$

In general, the terms  $\langle X_j, X_i^{(1)} \rangle_\pi X_j$  are of order  $O(1)$  due to normalization  $\|X_j(\epsilon)\| = 1$ . However, if the perturbation matrix  $T^{(1)}$  has a special structure, the result can be improved. This is verified by the following corollary.

**Corollary 1.1** *Under the modeling assumption that  $T^{(1)}$  inherits the nearly completely decomposable structure of  $T$ , i.e.*

$$T^{(1)} = \kappa T + O(\epsilon) \quad (1.9)$$

with some constant parameter  $\kappa$ , (1.7) simplifies to

$$X_i(\epsilon) = \sum_{j=1}^k \tilde{\alpha}_{ij} \chi_j + O(\epsilon^2).$$

**Proof:** Replacing  $T^{(1)}$  with (1.9) leads to

$$\epsilon \Pi_j T^{(1)} X_i = \epsilon \kappa \Pi_j T X_i + O(\epsilon^2) = \epsilon \kappa \Pi_j X_i + O(\epsilon^2) = O(\epsilon^2), \quad i = 1, \dots, k,$$

due to the orthogonality of the unperturbed eigenvectors. Insertion into the second summand of (1.7) yields the result.  $\square$

Thus, the  $O(\epsilon^2)$  perturbation result for the eigenvectors would still be valid. Equivalently to (1.9), one can write

$$(T(i, j) - T^{(1)}(i, j))/T(i, j) = (1 - \kappa) + O(\epsilon), \quad \forall i, j = 1, \dots, N.$$

Thus,  $T^{(1)}$  inherits the structure of  $T$  if the element-wise relative error has nearly the same size for all elements of  $T$ .

**Robustness of the PCCA approach.** Although it cannot be shown that the constant level pattern of the eigenvectors are perturbed in only  $O(\epsilon^2)$ , the newly proposed algorithm PCCA+ is nevertheless more robust than the older method based on the sign structure. This is due to the fact that it avoids the generic “dirty zero” problem by allowing the occurrence of transition states.

## 2 Almost characteristic functions

As pointed out in [3], Huisinga and Schmidt [4] had already shown an  $O(\epsilon)$  lower bound for the measure of metastability in the case of a strict  $\{0, 1\}$ -clustering. The following theorem shows that the result carries over to the framework of almost characteristic functions. However, in contrast to the earlier proposition in [2, 3], this bound cannot be improved. The proof is essentially the same as before, but only uses

$$\tilde{\pi} = \pi + O(\epsilon).$$

**Theorem 2.1** *Let  $\tilde{\Lambda} = \text{diag}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_k)$ . Assume that for a feasible set of almost characteristic functions  $\tilde{\chi} = \tilde{X} \tilde{\Lambda}$  the inequality*

$$\Theta = \|\tilde{\mathcal{A}}^{-1} \tilde{\Lambda} \tilde{\mathcal{A}} - I_k\|_1 < 1 \quad (2.1)$$

*is satisfied. Then metastability can be bounded in terms of the perturbation parameter  $\epsilon$  via*

$$\sum_{i=1}^k \tilde{\lambda}_i - O(\epsilon) \leq \sum_{i=1}^k w_{ii} < \sum_{i=1}^k \tilde{\lambda}_i. \quad (2.2)$$

The proof is based on the representation of the coupling matrix

$$\tilde{W} = \tilde{D}^{-2} \langle \tilde{\chi}, \tilde{T} \tilde{\chi} \rangle_{\pi}. \quad (2.3)$$

**Proof:** Upon reformulating

$$\langle \tilde{\chi}, \tilde{T} \tilde{\chi} \rangle_{\pi} = \tilde{\mathcal{A}}^T \langle \tilde{X}, \tilde{T} \tilde{X} \rangle_{\pi} \tilde{\Lambda} = \tilde{\mathcal{A}}^T \langle \tilde{X}, \tilde{X} \tilde{\Lambda} \rangle_{\pi} \tilde{\Lambda} = \tilde{\mathcal{A}}^T \langle \tilde{X}, \tilde{X} \rangle_{\pi} \tilde{\Lambda} \tilde{\Lambda}$$

and using  $\pi$ -orthogonality of the eigenvectors  $\tilde{X}$  we arrive at

$$\tilde{W} = \tilde{D}^{-2} \tilde{A}^T \tilde{\Lambda} \tilde{A}. \quad (2.4)$$

Moreover, based on the relation

$$\langle \tilde{\chi}, \tilde{\chi} \rangle_\pi = \tilde{A}^T \langle \tilde{X}, \tilde{X} \rangle_\pi \tilde{A} = \tilde{A}^T \tilde{A},$$

we may derive the alternative expression

$$\tilde{W} = \underbrace{\tilde{D}^{-2} \langle \tilde{\chi}, \tilde{\chi} \rangle_\pi}_S \underbrace{\tilde{A}^{-1} \tilde{\Lambda} \tilde{A}}_M. \quad (2.5)$$

By a short calculation, the above matrix  $S = (S_{ij})$  can be shown to be stochastic, which implies that  $\sum_{j=1}^k S_{ij} = 1$ . The matrix  $M = (M_{ij})$  is obviously spectrally similar to the eigenvalue matrix  $\tilde{\Lambda}$ . By assumption,  $M$  satisfies the condition  $\Theta = \|M - I_k\|_1 < 1$ , where  $\|\cdot\|_1$  denotes the maximum column sum norm; this implies that  $M_{ii} > M_{ji}$  for  $j \neq i$ .

With these properties, the upper bound in (2.2) can be directly verified as follows:

$$\sum_{i=1}^k w_{ii} = \sum_{i=1}^k \sum_{j=1}^k S_{ij} M_{ji} < \sum_{i=1}^k \left( \sum_{j=1}^k S_{ij} \right) M_{ii} = \sum_{i=1}^k M_{ii} = \sum_{i=1}^k \tilde{\lambda}_i.$$

In order to verify the lower bound in (2.2), let  $D^2 = \text{diag}(\pi_1, \dots, \pi_k)$  and observe that  $D^2 = \langle \chi, \chi \rangle_\pi$ . With this preparation, we derive the perturbation pattern of the matrix  $S$  as:

$$S = \tilde{D}^{-2} \langle \tilde{\chi}, \tilde{\chi} \rangle_\pi = (D^2 + O(\epsilon))^{-1} (\langle \chi, \chi \rangle_\pi + O(\epsilon)) = I_k + O(\epsilon).$$

Insertion into the expression (2.5) then immediately yields

$$\text{tr}(\tilde{W}) = \text{tr}(SM) = \text{tr}(M) + O(\epsilon) = \text{tr}(\tilde{\Lambda}) + O(\epsilon).$$

This result applies for both the upper and the lower bound of the metastability and therefore confirms (2.2) in particular, which completes the proof.  $\square$

Note that  $\tilde{\Lambda}(\epsilon)|_{\epsilon=0} = I_k$  so that for sufficiently small perturbation parameter the above inequality (2.1) will be satisfied.

## Conclusion

In contrast to the recent papers [2, 3], the  $O(\epsilon^2)$  perturbation result for eigenvectors has to be replaced by  $O(\epsilon)$  bounds. Fortunately, the cluster algorithm PCCA+ maintains its validity because it is not based on this result. However, we still do not have a satisfactory theoretical explanation for the simplex structure of the perturbed eigenvectors. This question will be in the focus of future work.

## References

- [1] P. Deuffhard, W. Huisinga, A. Fischer, and Ch. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Lin. Alg. Appl.*, 315:39–59, 2000.
- [2] P. Deuffhard and M. Weber. Robust Perron Cluster Analysis in Conformation Dynamics. ZIB-Report 03-19, Zuse Institute Berlin, 2003.
- [3] P. Deuffhard and M. Weber. Robust Perron Cluster Analysis in Conformation Dynamics. In M. Dellnitz, S. Kirkland, M. Neumann, and C. Schütte, editors, *Lin. Alg. App. – Special Issue on Matrices and Mathematical Biology*, volume 398C, pages 161–184. Elsevier, 2005.
- [4] W. Huisinga and B. Schmidt. Metastability and dominant eigenvalues of transfer operators. In B. Leimkuhler, C. Chipot, R. Elber, A. Laaksonen, A. Mark, T. Schlick, Ch. Schütte, and R. Skeel, editors, *New Algorithms for Macromolecular Simulation*, volume 49 of *Lecture Notes in Computational Science and Engineering*, pages 167–181, Berlin, 2006. Springer-Verlag.
- [5] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, Berlin Heidelberg, 1984.