

**D**  
Dipartimento

**S**  
Scienze

**E**  
Economiche

# Working Paper

Department  
of Economics

Ca' Foscari University of  
Venice

Michele Bernasconi  
Christine Choirat  
Raffaello Seri

A re-examination of the  
algebraic properties of  
the AHP as a ratio-scaling  
technique



## **A re-examination of the algebraic properties of the AHP as a ratio-scaling technique**

**Michele Bernasconi**

*Università Ca' Foscari Venezia*

**Christine Choirat**

*Universidad de Navarra*

**Raffaello Seri**

*Università dell'Insubria*

This Draft: October 2009

### **Abstract**

The Analytic Hierarchy Process (AHP) ratio-scaling approach is re-examined in view of the recent developments in mathematical psychology based on the so-called separable representations. The study highlights the distortions in the estimates based on the maximum eigenvalue method used in the AHP distinguishing the contributions due to random noises from the effects due to the nonlinearity of the subjective weighting function of separable representations. The analysis is based on the second order expansion of the Perron eigenvector and Perron eigenvalue in reciprocally symmetric matrices with perturbations. The asymptotic distributions of the Perron eigenvector and Perron eigenvalue are derived and related to the eigenvalue-based index of cardinal consistency used in the AHP. The results show the limits of using the latter index as a rule to assess the quality of the estimates of a ratio scale. The AHP method to estimate the ratio scales is compared with the classical ratio magnitude approach used in psychophysics.

### **Keywords**

Separable representations, reciprocally symmetric matrices, consistency indexes.

### **JEL Codes**

C44

*Address for correspondence:*

**Michele Bernasconi**

Department of Economics  
Ca' Foscari University of Venice  
Cannaregio 873, Fondamenta S.Giobbe  
30121 Venezia - Italy  
Phone: (+39) 041 2349191  
Fax: (+39) 041 2349176  
e-mail: [bernasconi@unive.it](mailto:bernasconi@unive.it)

*This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.*

The Working Paper Series  
is available only on line  
([www.dse.unive.it/publicazioni](http://www.dse.unive.it/publicazioni))  
For editorial correspondence, please contact:  
[wp.dse@unive.it](mailto:wp.dse@unive.it)

Department of Economics  
Ca' Foscari University of Venice  
Cannaregio 873, Fondamenta San Giobbe  
30121 Venice Italy  
Fax: +39 041 2349210

# 1 Introduction

After an era in which mathematical psychologists considered subjective measurement methods to be fundamentally unsound, in the last fifteen years or so several progresses have been made to comprehend and make explicit the structural assumptions needed to represent direct measurement data. Much attention has been devoted to ratio-scaling techniques.

In this paper we study the relationships between the recent developments in mathematical psychology and a ratio-scaling procedure widely used in management decisions, the Analytic Hierarchy Process (AHP).

As it is well known, ratio-scaling procedures have been introduced in the behavioral sciences in the middle of the last century by psychophysicist Stanley S. Stevens (1951, 1957). In the simplest ratio-scaling experiment, also known as magnitude estimation or magnitude estimation with a standard, an individual is asked to compare a set of stimuli  $(x_1, \dots, x_n)$  with a baseline stimulus  $x_0$ .<sup>1</sup> Each comparison yields a response value  $\alpha_{i0}$ , with  $i = 1, \dots, n$ . Stevens assumed that the values  $\alpha_{10}, \dots, \alpha_{n0}$  could directly represent a ratio scale, in the sense that he conjectured the existence of a ratio estimation function of the form  $\alpha_{i0} = \left(\frac{x_i}{x_0}\right)^k$ , which corresponds to his famous psychophysical law that equal physical ratios produce equal psychological ratios.

Stevens' model has always been highly criticized by mathematical psychologists because it lacks of normative and descriptive justifications (Mitchell 1999, chapter 4). Recently, however, mainly due to the work of Louis Narens (1996, 2002, 2006) and Duncan Luce (2002, 2004), axiomatic developments of subjective measurement approaches have been developed with stronger theoretical foundations. The new models belong to a class of so-called separable representations, which establish the following relationships between the stimuli and the responses of a ratio estimation exercise:

$$W(\alpha_{i0}) = \frac{\psi(x_i)}{\psi(x_0)}. \quad (1)$$

In equation (1),  $\psi$  is called the *psychophysical* function and  $W$  the *subjective weighting* function. The two functions indicate that two independent transformations may occur in a ratio estimation: one of the stimuli intensities (embodied in  $\psi$ ), and the other of numbers (entailed in  $W$ ). Support for separable forms has been found in a series of recent experiments which have been conducted to test some of their underlying properties and which have estimated different functional specifications of  $\psi$  and  $W$  (see, among others, Ellermeier and Faulhammer 2000, Zimmer 2005, Steingrimsson and Luce 2005a, 2005b,

---

<sup>1</sup>The term magnitude estimation was introduced by Stevens and Galanter (1957). (See also Stevens, 1975.) A dual scaling procedure, also widely used in psychophysics, is known as ratio production, in which an individual is asked to produce a stimulus  $s_i$  which appears  $p_i$  times more intense than a reference stimulus  $s_0$ . (See Luce 2004, and Steingrimsson and Luce 2006, for the axiomatic treatments of the different cases of ratio magnitude estimation and ratio production).

2006, 2007, Bernasconi, Choirat and Seri 2008).

While generalizing Stevens' power law model, a characteristic of the separable forms is that they maintain the deterministic approach of the former, in the sense that "these theories are about idealized situations and do not involve considerations of errors" (Narens 1996, p. 109). Several mathematical psychologists have acknowledged that this is a limitation (e.g. Luce 1997, p. 81).

A different approach which emphasizes the effects of errors in ratio-scaling procedures is the Analytic Hierarchy Process. The AHP is a ratio-scaling technique developed by management scientist Thomas Saaty (1977, 1980, 1986). Two main differences characterize Saaty's approach with regard to the most simple ratio-scaling experiment. First of all, the AHP allows for inconsistencies in the form of random errors in the subjective ratio judgments. Secondly, because subjective assessments can be affected by random noise, in order to improve the validity of the ratio-scaling procedure, in the AHP a decision maker is asked to fill out an entire  $(n \times n)$  matrix of ratio judgments  $\mathbf{A} = [\alpha_{ij}]$ , where each  $\alpha_{ij}$  is taken to measure on a subjective ratio scale the relative dominance of stimulus  $i$  over stimulus  $j$ , from the vector of stimuli  $(x_1, \dots, x_n)$ .

Combining the developments in separable forms with the errors-allowing approach of the AHP, in Bernasconi, Choirat, Seri (2009) we propose the following representation for the ratio assessments of an AHP response matrix  $\mathbf{A}$ :

$$\alpha_{ij} = W^{-1} \left( \frac{\psi(x_i)}{\psi(x_j)} \right) \cdot e_{ij} \quad (2)$$

where  $W^{-1}(\cdot)$  is the inverse of a subjective weighting function from separable representations;  $\psi(x_1), \dots, \psi(x_n)$  are the psychological perceptions of the stimuli intensities corresponding to priority weights  $u_i = \frac{\psi(x_i)}{\sum \psi(x_k)}$  (for  $i = 1, \dots, n$ ) which, coherently with standard AHP, are normalized to sum up to unity, i.e.  $\sum u_i = 1$ ; and where  $e_{ij}$  are the multiplicative error terms introduced in the AHP to account for the inconsistencies in subjective ratio judgments typically observed in practice.

Equation (2) represents a generalization of classical AHP in the sense that in Saaty's standard approach  $W^{-1}$  is the identity. For such a case, the AHP also proposes to use as the best approximation of the vector of the priority weights  $(u_1, \dots, u_n)'$  the right Perron eigenvector  $\mathbf{u} = \mathbf{u}(\mathbf{A})$  of the response matrix  $\mathbf{A}$ .<sup>2</sup> It is in particular well-known by the Perron-Frobenius Theorem that the Perron eigenvector  $\mathbf{u}$  is the unique solution of the system of equations

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}, \quad \sum u_i = 1$$

where  $\lambda$  denotes the Perron root (maximum eigenvalue) of  $\mathbf{A}$ . Moreover, when  $e_{ij} = 1$  all  $i, j = 1, \dots, n$  (and  $W^{-1}(\cdot)$  remains the identity), so that any row  $(\alpha_{i1}, \dots, \alpha_{in})$  of  $\mathbf{A}$  can be

---

<sup>2</sup>We follow the habit of calling *principal* or *Perron eigenvalue* the largest eigenvalue of a matrix with positive entries (that is real and unique by the Perron-Frobenius Theorem), and *Perron eigenvector* the eigenvector associated with it.

obtained from any other row by the relation  $\alpha_{ik} = \alpha_{jk}/\alpha_{ji}$  (the rank  $r(\mathbf{A}) = 1$ ), it is known that the maximum-eigenvalue method delivers the correct priority weights  $u_i \propto \psi(x_i)$  for all  $i$ , with the maximum eigenvalue being at its minimum  $\lambda = n$ . Saaty's argument is to use the same approach even when there are some inconsistencies in the data due to the fact that the error terms  $e_{ij}$  are only close to 1. Saaty's method does not pay attention to the stochastic structure of the data. Therefore, statistical approaches to estimate the priority weights  $(u_1, \dots, u_n)'$  which pay more attention to the stochastic structure of the data have also been proposed for classical AHP, which include the logarithmic least squares method as the most standard alternative (see de Jong 1984, Crawford and Williams 1985, Genest and Rivest 1994).

In Bernasconi, Choirat and Seri (2009) we develop a detailed analysis of representation (2) and propose a method to estimate the priority weights  $(u_1, \dots, u_n)'$  even when  $W^{-1}(\cdot)$  is nonlinear, which can be viewed as an approximate polynomial generalization of the logarithmic least squares method. Moreover, by applying the new estimation method to the data of a ratio estimation experiment, further evidence is provided about the importance of the nonlinearity  $W^{-1}(\cdot)$  to generate inconsistencies in the subjective measurement data.

In this paper we deal with a different issue, concerning the effects of the separable forms and of the nonlinearity of the subjective weighting function  $W(\cdot)$  for the mathematical behavior of Saaty's maximum-eigenvalue method. Even if one does not regard the maximum eigenvalue method as the best method to obtain the priority vector  $(u_1, \dots, u_n)'$  from matrix  $\mathbf{A}$ , looking at the mathematical properties of the maximum-eigenvalue method when representation (2) holds with  $W^{-1}(\cdot)$  nonlinear is still quite relevant. When there are no distortions, the principal eigenvector method is obviously the most natural method to recover the priority weights. So, it is quite important to know how the natural benchmark behaves when there are inconsistencies in the response data. Among other things, it is in particular important to know how large the distance of the responses data from a ratio scale is. Such an issue has been studied by Genest and Rivest (1994) for the case when the distortions in the response matrix  $\mathbf{A}$  are due to the error terms  $e_{ij}$ 's. We extend their analysis to the case in which inconsistencies in the response data may also be due to the subjective weighting function  $W^{-1}(\cdot)$ .

We provide several results. First of all, we measure the extent to which the priority weights obtained by the principal eigenvector with the AHP depart from a ratio scale, distinguishing the contributions of the error terms  $e_{ij}$  from the effects due to the nonlinearity of  $W^{-1}(\cdot)$ . We approach the problem using the theory of matrix differentials developed in Magnus and Neudecker (1999). We take the second order approximation of the priority vector  $(u_1, \dots, u_n)'$  and of the Perron eigenvalue  $\lambda$  around their ideal values holding in a system in which there are no distortions. Our analysis shows that when the stimuli range is small the effects of the deterministic distortions due to  $W^{-1}(\cdot)$  and to the stochastic terms are comparable; whereas when the stimuli are very different the deterministic distortions are much larger than those due to the stochastic errors. We see this

result providing precise normative justification for a well-known axiom which Saaty (1986) introduced heuristically in the AHP, the so called “homogeneity” axiom, which requires the stimuli used in the AHP to be in a range of comparability.

We also study the approximate distribution of Saaty’s maximum eigenvalue  $\lambda$  and eigenvector  $\mathbf{u}$  when equation (2) holds. For the case in which  $W$  is linear, Genest and Rivest (1994) have shown that  $\lambda$  is upward biased with respect to  $\lambda_0 = n$  with an approximate  $\chi^2$  distribution; and they have shown how to relate this statistically-based measure of consistency with the index of cardinal consistency proposed by Saaty for the AHP and based on the quantity  $\mu = (\lambda - n)/(n - 1)$ . In particular Genest and Rivest have shown how the so-called 10% cut-off rule used in the AHP<sup>3</sup> can be considered equivalent to a statistical test of the hypothesis that the variability due to the error terms  $e_{ij}$ ’s in the response data does not exceed some given threshold. We extend their results and show that when  $W$  is allowed to be different from the identity, then  $\lambda$  has approximately a normal distribution or a noncentral  $\chi^2$  distribution whose noncentrality parameter depends on the strength of the distortion induced by  $W$ . We show how in this case the use of the 10% cut-off rule can lead to severely undersized tests, in particular, when the deterministic perturbations due to  $W$  are larger than those caused by the error terms  $e_{ij}$ ’s. We derive the asymptotic distribution of  $\mathbf{u}$  and show that it is always normal with a bias that depends on the subjective weighting function  $W$  and never on the error terms.

Finally, we compare the performance of the AHP with the classical ratio magnitude estimation used in psychophysics. As alluded above, one justification of Saaty to develop the AHP was the intuition that, since the AHP provides an entire matrix of subjective ratio assessments with all the possible pairwise-comparisons between the stimuli  $(x_1, \dots, x_n)$ , it could improve the validity of a procedure like ratio magnitude estimation in which each element of  $(x_1, \dots, x_n)$  is considered only once in comparison to a reference stimulus  $x_0$ . We show that when the principal eigenvector method is applied to a matrix in which the response data obey equation (2), the above intuition is not always valid, but it depends on how the reference stimulus  $x_0$  is actually chosen and whether the homogeneity axiom is actually respected in the AHP. In particular, we show that when homogeneity is respected and all the distortions in the response data are due to random noise, then the AHP is always better than the ratio magnitude estimation to obtain an estimate of vector  $(u_1, \dots, u_n)'$ ; whereas, when homogeneity is not respected and there are psychophysical distortions in the data, the ratio magnitude estimation can be better. We give examples when this could be the case.

The paper is organized as follows. We start in Section 2 with a formal presentation of a model for the  $\alpha_{ij}$ ’s based on equation (2). In Section 3 we derive three main theorems for the second order expansions of the Perron eigenvector and Perron eigenvalue in reciprocally

---

<sup>3</sup>The 10% cut-off rule considers the degree of cardinal inconsistency in a response matrix  $\mathbf{A}$  too large (and hence rejects the data) when its consistency index  $\mu$  is larger than one tenth of the average index  $\mu_0$ , computed from the average of a randomly generated reciprocally symmetric matrix of the same size as  $\mathbf{A}$ .

symmetric matrices with perturbations. In Section 4 these theorems are applied to the AHP to study the algebraic properties of Saaty’s method and to derive the asymptotic distributions of interest. In Section 5 we conduct the comparison between the AHP and ratio magnitude estimation. Section 6 concludes.

## 2 A model for the $\alpha_{ij}$ ’s

As indicated in the Introduction, inspired by the recent separable representations in mathematical psychology, we consider the following model for the  $\alpha_{ij}$ ’s of a response matrix  $\mathbf{A}$  obtained in the AHP:

$$\alpha_{ij} = W^{-1} \left( \frac{\psi(x_i)}{\psi(x_j)} \right) \cdot e_{ij}. \quad (2)$$

In classical AHP the equality  $a_{ij} \cdot a_{ji} = 1$  is assumed. In order to maintain this equality we first of all assume (as it is standard in the AHP) that the error term is given by  $e_{ij} = e^{\sigma \nu_{ij}}$  where  $\sigma$  is a scale parameter that affects the behavior of  $\nu_{ij}$ ,  $\nu_{ii} = 0$  and  $\nu_{ij} = -\nu_{ji}$ . Moreover, for most of this paper we also assume that the subjective weighting function  $W(\cdot)$  is a (monotonic) reciprocally symmetric function, namely  $W(1/\cdot) = 1/W(\cdot)$ . We emphasize that this assumption is implied by several derivations of separable representations, including a specification developed by Luce (2001, 2002), similar to one which Prelec (1998) proposed in the context of utility theory for risky gambles. A more recent specification proposed by Luce (2004) (and further discussed by Aczél and Luce 2007), does not instead impose either symmetry or  $W(1) = 1$ . A part of the analysis below will also cover this more general case.<sup>4</sup>

The theoretical model that is supposed to hold in classical AHP is  $\alpha_{0,ij} = \frac{\psi(x_i)}{\psi(x_j)}$ . In this case, we say that a property of separable forms which Narens (1996) called “multiplicativity” holds; this is equivalent to “consistency” in the sense of Saaty. A matrix  $\mathbf{A}_0$  filled with these elements yields:

$$\begin{aligned} \mathbf{u}_0 &= \left[ \frac{\psi(x_i)}{\sum_{k=1}^n \psi(x_k)} \right], \\ \lambda_0 &= n. \end{aligned}$$

We remark that  $\mathbf{u}_0$  is not an approximation of  $\left[ \frac{x_i}{\sum_{k=1}^n x_k} \right]$ . This is because in the AHP one is not interested in obtaining objective ratio measurements of the stimuli  $x_i$  (which in many cases may not even be possible, for example when the stimuli do not come from a known scale); but one is interested in obtaining subjective ratio measurements of opinions, beliefs, preferences. In fact, in the latter case, one could also refer to the psychophysical

---

<sup>4</sup>On the empirical evidence, several direct tests conducted in psychophysical experiments on loudness production, which include Steingrímsson and Luce (2007) and Zimmer (2005) have rejected the behavioral hypothesis underlying the specification with  $W(1) = 1$  and have accepted one with  $W(1) \neq 1$ . Indirect tests based on the inference of separable forms in an experiment measuring the distance-ratio between Italian cities in Bernasconi, Choirat and Seri (2008) have not rejected  $W(1) = 1$ .

function  $\psi(\cdot)$  as “utility” function (e.g. Luce 2002, p. 523).

Our objective is to approximate  $\mathbf{u}$  and  $\lambda$ , solutions of the system:

$$\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}, \quad (3)$$

as a function of the response matrix  $\mathbf{A}$  of equation (2), when  $\mathbf{A}$  can be considered a perturbation of  $\mathbf{A}_0$  in the following sense.

We suppose that both the perturbation caused by the application of  $W^{-1}$  and by the multiplication by  $e^{\sigma\nu_{ij}}$  are asymptotically negligible (the parameters governing the asymptotic behavior will be introduced later on), so that:

$$\alpha_{ij} = \alpha_{0,ij} \cdot \exp \left\{ \ln \left[ \frac{\psi(x_j)}{\psi(x_i)} \cdot W^{-1} \left( \frac{\psi(x_i)}{\psi(x_j)} \right) \right] + \sigma\nu_{ij} \right\} = \alpha_{0,ij} \cdot e^{d\varepsilon_{ij}} \quad (4)$$

where we define:

$$d\varepsilon_{ij} = \ln \left[ \frac{\psi(x_j)}{\psi(x_i)} \cdot W^{-1} \left( \frac{\psi(x_i)}{\psi(x_j)} \right) \right] + \sigma\nu_{ij}$$

for  $j > i$ . The reason for the use of the differential symbol  $d$  will be clear in the following: we will indeed suppose that  $\alpha_{ij}$  is a small perturbation of  $\alpha_{0,ij}$  so that  $d\varepsilon_{ij}$  is an infinitesimal quantity. In order to respect the property of reciprocal symmetry, we need  $d\varepsilon_{ij} = -d\varepsilon_{ji}$ .

Consider the following approximation, used in Bernasconi, Choirat and Seri (2008), in which the function  $W^{-1}$  is first log-transformed to  $w^{-1}$ :

$$\begin{aligned} \ln W[\exp(\cdot)] &= w(\cdot) \\ w^{-1}(\cdot) &= \ln W^{-1}[\exp(\cdot)] \end{aligned}$$

and  $w^{-1}$  is given by a polynomial in its argument (whose degree  $L$  can even be infinite):

$$w^{-1}(x) = \sum_{\ell=0}^L \phi_{\ell} \cdot x^{\ell}$$

with  $\phi_0 = 0$ ,  $\phi_1 = 1$  and  $\phi_{2n} = 0$  for  $n \in \mathbb{N}$ . This is generally possible: according to the Weierstrass Approximation Theorem, any continuous function on a compact domain can be approximated to any desired degree of accuracy by a polynomial in its arguments. Therefore:

$$\begin{aligned} W^{-1}(x) &= \exp \left\{ \sum_{\ell=0}^L \phi_{\ell} \cdot [\ln(x)]^{\ell} \right\} \\ &= x \cdot \exp \left\{ \sum_{\ell=2}^L \phi_{\ell} \cdot [\ln(x)]^{\ell} \right\} \end{aligned}$$



so that when  $\|\phi_\ell\|_\infty = \max_{2 \leq \ell \leq L} |\phi_\ell| \downarrow 0$ ,  $W^{-1}(x) \rightarrow x$ .

Our infinitesimal error term is:

$$\begin{aligned} d\varepsilon_{ij} &= \ln \left[ \frac{\psi(x_j)}{\psi(x_i)} \cdot W^{-1} \left( \frac{\psi(x_i)}{\psi(x_j)} \right) \right] + \sigma \nu_{ij} \\ &= \sum_{\ell=2}^L \phi_\ell \cdot [\ln(\psi(x_i)/\psi(x_j))]^\ell + \sigma \nu_{ij} \end{aligned} \quad (5)$$

Under the hypotheses that  $\|\phi_\ell\|_\infty = \max_{2 \leq \ell \leq L} |\phi_\ell| \downarrow 0$  and  $\sigma \downarrow 0$ ,  $d\varepsilon_{ij}$  is asymptotically negligible.

In the remainder of the paper we will study the relationships between the response matrix  $\mathbf{A}$  and the ideal matrix  $\mathbf{A}_0$ , constructed as specified above. In the next Section, we show some useful facts about the differentials of reciprocally symmetric matrices, using the theory of matrix differentials in the sense of Magnus and Neudecker (1999). We start from some general results, which are then narrowed down to the case most interesting for the AHP.

### 3 Perturbations of Reciprocally Symmetric Matrices with applications to the AHP

Consider a matrix  $\mathbf{A}$  and a reciprocally symmetric matrix  $\mathbf{A}_0$  such that  $\mathbf{A}$  can be considered a perturbation of  $\mathbf{A}_0$ . We write therefore  $\mathbf{A} \simeq \mathbf{A}_0 + d\mathbf{A} + \frac{1}{2}d^2\mathbf{A}$  where  $d\mathbf{A}$  and  $d^2\mathbf{A}$  are matrix differentials (therefore asymptotically negligible) in the sense of Magnus and Neudecker (1999). We want to study the behavior of the Perron eigenvalue  $\lambda$  and the right Perron eigenvector  $\mathbf{u} = \mathbf{u}(\mathbf{A})$  of  $\mathbf{A}$  in system (3), taken as perturbations of the corresponding quantities  $\lambda_0$  and  $\mathbf{u}_0 = \mathbf{u}(\mathbf{A}_0)$ , of the ideal system:

$$\mathbf{A}_0 \cdot \mathbf{u}_0 = \lambda_0 \cdot \mathbf{u}_0. \quad (6)$$

In particular, we want to obtain  $d\mathbf{u}$ ,  $d^2\mathbf{u}$ ,  $d\lambda$ ,  $d^2\lambda$  in the second order approximations of the right Perron eigenvector  $\mathbf{u} \simeq \mathbf{u}_0 + d\mathbf{u} + \frac{1}{2}d^2\mathbf{u}$  of  $\mathbf{A}$  and of the Perron eigenvalue  $\lambda \simeq \lambda_0 + d\lambda + \frac{1}{2}d^2\lambda$  of system (3).

The following notation will be used throughout the Section. For a  $n$ -vector  $\mathbf{a}$  let  $\bar{\mathbf{a}}$  be the  $n$ -vector defined by  $\bar{\mathbf{a}} = [\bar{a}_i] = [a_i^{-1}]$ .  $\mathbf{u}_n$  is a  $n$ -vector composed of ones.  $\mathbf{I}_n$  is the  $(n, n)$ -identity matrix.  $\mathbf{U}_n$  is a  $(n, n)$ -matrix composed of ones.  $\mathbf{e}_i$  is a vector of zeros with a one in the  $i$ -th position. Let:

$$\mathbf{K}_{nn} \triangleq \sum_{i=1}^n \sum_{j=1}^n \mathbf{e}_i \mathbf{e}_j^\top \otimes \mathbf{e}_j \mathbf{e}_i^\top,$$

where  $\otimes$  denotes the Kronecker product, be the commutation matrix (Magnus and Neudecker

1999, p. 46). The notation  $\overline{\ln \mathbf{A}}$ ,  $\overline{\exp \mathbf{A}}$  and  $\mathbf{A}^{\odot \ell}$  denote the element-wise application of natural logarithm, exponential and power function (of degree  $\ell$ ) to a matrix  $\mathbf{A}$ . On the other hand,  $\mathbf{A}^\ell$  denotes the ordinary product of the matrix  $\mathbf{A}$  by itself, repeated  $\ell$  times.  $\mathbf{A}^+$  is the Moore-Penrose inverse of the matrix  $\mathbf{A}$ .

Our first result yields the Perron eigenvalue and the right Perron eigenvector when  $\mathbf{A}_0$  is a general reciprocally symmetric matrix, and the perturbation  $d\mathbf{A} + \frac{1}{2}d^2\mathbf{A} + o(\|d^2\mathbf{A}\|)$  does not necessarily yield a reciprocally symmetric matrix  $\mathbf{A}$ .

**Theorem 3.1.** *Consider the eigenvalue problems  $\mathbf{A}_0 \cdot \mathbf{u}_0 = \lambda_0 \cdot \mathbf{u}_0$  and  $\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}$  where  $\mathbf{A} = \mathbf{A}_0 + d\mathbf{A} + \frac{1}{2}d^2\mathbf{A} + o(\|d^2\mathbf{A}\|)$  and  $d\mathbf{A}$  and  $d^2\mathbf{A}$  are the matrix differentials in the sense of Magnus and Neudecker (1999).  $\lambda_0$  is a simple eigenvalue with right eigenvector  $\mathbf{u}_0$  and left eigenvector  $\mathbf{v}_0$ . The right eigenvectors  $\mathbf{u}_0$  and  $\mathbf{u}$  are normalized as  $\mathbf{u}_0^\top \mathbf{u}_n = 1$  and  $\mathbf{u}^\top \mathbf{u}_n = 1$ . The following expansions hold:*

$$\begin{aligned}\lambda(\mathbf{A}) &= n + d\lambda + \frac{1}{2}d^2\lambda + o(d^2\lambda), \\ \mathbf{u}(\mathbf{A}) &= \mathbf{u}_0 + d\mathbf{u} + \frac{1}{2}d^2\mathbf{u} + o(\|d^2\mathbf{u}\|),\end{aligned}$$

where:

$$\begin{aligned}d\lambda &= \frac{\mathbf{v}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0}{\mathbf{v}_0^\top \mathbf{u}_0} \\ d^2\lambda &= \frac{\mathbf{v}_0^\top \mathbf{u}_0 \cdot \mathbf{v}_0^\top d^2\mathbf{A} \mathbf{u}_0 + 2\mathbf{v}_0^\top (\mathbf{u}_0 \mathbf{v}_0^\top d\mathbf{A} - d\mathbf{A} \mathbf{u}_0 \cdot \mathbf{v}_0^\top) \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left( \mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\top}{\mathbf{v}_0^\top \mathbf{u}_0} \right) (d\mathbf{A}) \mathbf{u}_0}{(\mathbf{v}_0^\top \cdot \mathbf{u}_0)^2} \\ d\mathbf{u} &= \left( \mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n^\top \right) \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left( \mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0^\top}{\mathbf{v}_0^\top \mathbf{u}_0} \right) (d\mathbf{A}) \mathbf{u}_0 \\ d^2\mathbf{u} &= (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (d^2\mathbf{A} - d^2\lambda \mathbf{I}_n) \cdot \mathbf{u}_0 + 2(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (d\mathbf{A} - d\lambda \mathbf{I}_n) \cdot d\mathbf{u}.\end{aligned}$$

Theorem 3.1 is very general, since the only restriction is that  $\mathbf{A}_0$  must be a reciprocally symmetric matrix. Next theorem holds when matrix  $\mathbf{A}_0$  is also perfectly consistent, but  $\mathbf{A}$  needs not be reciprocally symmetric.  $\mathbf{A}_0$  is perfectly consistent if there exists a  $n$ -vector  $\mathbf{u}$  such that:

$$\mathbf{A}_0 = \mathbf{u} \bar{\mathbf{u}}^\top,$$

with  $\mathbf{u}_n^\top \cdot \mathbf{u} = 1$ . As concerns the right Perron eigenvector  $\mathbf{u}_0 = \mathbf{u}(\mathbf{A}_0)$ , the solution is given by  $\mathbf{u}_0 = \mathbf{u}$  and  $\lambda_0 = n$ , since:

$$\begin{aligned}\mathbf{u} \cdot \left[ \bar{\mathbf{u}}^\top \mathbf{u} \right] &= \lambda_0 \cdot \mathbf{u} \\ n \cdot \mathbf{u} &= \lambda_0 \cdot \mathbf{u}.\end{aligned}$$

The fact that  $\lambda_0$  is the Perron eigenvalue (and  $\mathbf{u}_0$  the correspondent Perron eigenvector) can be shown remarking that  $\mathbf{A}_0$  has one eigenvalue equal to  $n$  and  $(n-1)$  eigenvalues

equal to 0.

**Theorem 3.2.** *When  $\mathbf{A}_0$  is a reciprocally symmetric consistent matrix, the following expansion holds:*

$$\begin{aligned}\lambda(\mathbf{A}) &= n + d\lambda + \frac{1}{2}d^2\lambda + o(d^2\lambda), \\ \mathbf{u}(\mathbf{A}) &= \mathbf{u}_0 + d\mathbf{u} + \frac{1}{2}d^2\mathbf{u} + o(\|d^2\mathbf{u}\|),\end{aligned}$$

where:

$$\begin{aligned}d\lambda &= \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0}{n} \\ d^2\lambda &= 2 \cdot \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot (n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top) \cdot d\mathbf{A} \cdot \mathbf{u}_0}{n^3} + \frac{\bar{\mathbf{u}}_0^\top \cdot d^2\mathbf{A} \cdot \mathbf{u}_0}{n} \\ d\mathbf{u} &= \frac{1}{n} \cdot (\mathbf{I}_n - \mathbf{u}_0 \cdot \mathbf{u}_n^\top) \cdot d\mathbf{A} \cdot \mathbf{u}_0 \\ d^2\mathbf{u} &= (n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top)^+ (d^2\mathbf{A} - d^2\lambda\mathbf{I}_n) \cdot \mathbf{u}_0 \\ &\quad + 2(n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top)^+ (d\mathbf{A} - d\lambda\mathbf{I}_n) \cdot d\mathbf{u}.\end{aligned}$$

With reference to the model for the  $\alpha_{ij}$ 's discussed in Section 2, Theorem 3.2 may for example apply when  $\mathbf{A}_0$  is filled with elements  $\alpha_{0,ij} = \frac{\psi(x_i)}{\psi(x_j)}$ , and the hypothesis on the errors  $e_{ij} = e^{\sigma\nu_{ij}}$  with  $\nu_{ii} = 0$  and  $\nu_{ij} = -\nu_{ji}$  is satisfied, but the subjective weighting function  $W(\cdot)$  is not necessarily reciprocally symmetric. Such a case could in particular be relevant for the general separable forms axiomatized by Luce (2004). When  $W(\cdot)$  is also reciprocally symmetric with  $W(1/\cdot) = 1/W(\cdot)$ , the expansions for the Perron eigenvalue and Perron eigenvector can be further specialized. In particular, in what follows we take  $d\mathbf{A}$  and  $d^2\mathbf{A}$ , the perturbations of  $\mathbf{A}$ , to be determined according to equation (4) of Section 2, namely  $\alpha_{ij} = \alpha_{0,ij} \cdot e^{d\varepsilon_{ij}} \approx \alpha_{0,ij} \cdot \left(1 + d\varepsilon_{ij} + \frac{1}{2}(d\varepsilon_{ij})^2\right)$ . Defining the skew-symmetric matrix:

$$[d\mathbf{E}]_{ij} = \begin{cases} d\varepsilon_{ij} & \text{if } j > i \\ -d\varepsilon_{ij} & \text{if } i > j \end{cases}$$

we have  $\mathbf{A} = \mathbf{A}_0 \odot \overline{\exp}(d\mathbf{E})$ ,  $d\mathbf{A} = \mathbf{A}_0 \odot d\mathbf{E}$  and  $d^2\mathbf{A} = \mathbf{A}_0 \odot d\mathbf{E} \odot d\mathbf{E}$ , where  $\odot$  is the Hadamard or Schur or element-wise product of matrices (see Magnus and Neudecker 1999, p. 45). We obtain the following theorem.

**Theorem 3.3.** *Consider the eigenvalue problems  $\mathbf{A}_0 \cdot \mathbf{u}_0 = \lambda_0 \cdot \mathbf{u}_0$  and  $\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}$  where  $\mathbf{A} = \mathbf{A}_0 \odot \overline{\exp}(d\mathbf{E})$ . Let  $\mathbf{A}_0$  be a reciprocally symmetric and consistent matrix and  $d\mathbf{E}$  be*

a skew-symmetric matrix of perturbations. In this case:

$$\begin{aligned}
d\lambda &= 0 \\
d^2\lambda &= \frac{\mathbf{u}_n^\top \cdot [2(\mathbf{dE} \cdot \mathbf{dE}) + n(\mathbf{dE} \odot \mathbf{dE})] \cdot \mathbf{u}_n}{n^2} \\
d\mathbf{u} &= \frac{1}{n} \cdot (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \cdot \mathbf{dE} \cdot \mathbf{u}_n \\
d^2\mathbf{u} &= \left(n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top\right)^+ (\mathbf{A}_0 \odot \mathbf{dE} \odot \mathbf{dE} - d^2\lambda \cdot \mathbf{I}_n) \cdot \mathbf{u}_0 \\
&\quad + 2 \left(n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top\right)^+ (\mathbf{A}_0 \odot \mathbf{dE} - d\lambda \cdot \mathbf{I}_n) \cdot d\mathbf{u}.
\end{aligned} \tag{7}$$

Theorem 3.3 is central for analyzing Saaty's eigenvalue method whenever there are small distortions in the responses which preserve the property of reciprocal symmetry of the response matrix  $\mathbf{A}$ , because it gives a measure of the perturbation which depends on the whole matrix and not only on the individual elements  $\alpha_{ij}$  (as for example in Saaty 1977, 1980, 1986).

In the following we will show how Theorem 3.3 can be used to study the effect of systematic biases and individual variabilities on  $\mathbf{u}$  and  $\lambda$ . The analysis to be conducted is in this sense a generalization of Genest and Rivest (1994) whose study of the asymptotic distributions of  $\lambda$  and  $\mathbf{u}$  is restricted to the case in which  $W$  is equal to the identity. We first discuss the properties of Saaty eigenvector-eigenvalue method when the distortions are due both to the deterministic component entailed by the subjective weighting function  $W$  and to the stochastic components carried by the  $\varepsilon_{ij}$ 's. Then we analyze the limiting behavior of  $\lambda$  and  $\mathbf{u}$  under various hypotheses on the relative strength of the random errors  $\varepsilon_{ij}$ 's and of the deterministic perturbations induced by  $W$ .

## 4 Properties of Saaty's Method

We provide here an analysis of the general contribution of deterministic and stochastic components to generate inconsistencies in the data and their impact on Saaty's eigenvalue method. To do so we rewrite the expressions for  $d\mathbf{u}$ ,  $d\lambda$ ,  $d^2\lambda$  from Theorem 3.3 using equation (5) of Section 2. In particular, rewriting equation (5) as:

$$d\mathbf{E} = \sum_{\ell=2}^L \phi_\ell \cdot [\overline{\ln \mathbf{A}_0}]^{\odot \ell} + \sigma \mathbf{N} + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma)$$

where  $\mathbf{N} \triangleq [\nu_{ij}]$  is the matrix of (scaled) random errors, the relevant terms of Theorem 3.3 become:<sup>5</sup>

$$\begin{aligned}
d\lambda &= 0, \\
d^2\lambda &= \left[ \sum_{\ell=2}^L \phi_\ell [(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0})]^{\odot \ell} + \sigma \text{vec}(\mathbf{N}) \right]^\top \\
&\quad \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \left[ \sum_{\ell=2}^L \phi_\ell [(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0})]^{\odot \ell} + \sigma \text{vec}(\mathbf{N}) \right] \\
&\quad + o(\|\phi_\ell\|_\infty^2) + o_{\mathbb{P}}(\sigma^2), \\
d\mathbf{u} &= \frac{\sum_{\ell=2}^L \phi_\ell}{n} \cdot \left( \mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \right) \cdot [(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0})]^{\odot \ell} \\
&\quad + \frac{\sigma}{n} \cdot \left( \mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \right) \cdot \text{vec}(\mathbf{N}) + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma). \tag{8}
\end{aligned}$$

We remark that the expression for  $d\mathbf{u}$  derived above shows that  $\mathbf{u} - \mathbf{u}_0 \simeq d\mathbf{u}$  can be separated in two additive parts, a deterministic one depending only on the distortions due to  $W$  (as expressed through the coefficients  $\phi_2, \dots, \phi_L$ ) and a stochastic one depending only on the random errors (as measured by the standard deviation  $\sigma$ ). This is particularly appealing since it allows us to formulate some prescriptive devices concerning the relative contributions of the deterministic and the stochastic components as functions of the values taken by  $n$  and  $\mathbf{u}_0$ .

In order to evaluate the relative contributions of the deterministic and stochastic components of equation (8), we measure the first one through the “bias”  $\|\mathbb{E}\mathbf{u} - \mathbf{u}_0\| \simeq \|\mathbb{E}d\mathbf{u}\|$  and the second one through the “variance”  $\sqrt{\mathbb{E}\|\mathbf{u} - \mathbb{E}\mathbf{u}\|^2} \simeq \sqrt{\mathbb{E}\|d\mathbf{u} - \mathbb{E}d\mathbf{u}\|^2}$ . Notice that in general these quantities depend on  $n$  and  $\mathbf{u}_0$ . In particular, when we suppose that  $\phi_3 \neq 0$  while  $\phi_\ell = 0$  for  $\ell \geq 4$ , the quantities  $\frac{\|\mathbb{E}d\mathbf{u}\|}{\phi_3}$  and  $\frac{\sqrt{\mathbb{E}\|d\mathbf{u} - \mathbb{E}d\mathbf{u}\|^2}}{\sigma}$  depend only on  $n$  and  $\mathbf{u}_0$ :

$$\frac{\|\mathbb{E}d\mathbf{u}\|}{\phi_3} = \frac{1}{n} \cdot \left( \mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \right) \cdot [(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0})]^{\odot 3}$$

and:

$$\frac{\sqrt{\mathbb{E}\|d\mathbf{u} - \mathbb{E}d\mathbf{u}\|^2}}{\sigma} = \frac{1}{n^{1/2}} \cdot \sqrt{\sum_{i=1}^n u_{0,i}^2 + \left( \sum_{i=1}^n u_{0,i}^2 \right)^2 - 2 \sum_{i=1}^n u_{0,i}^3}.$$

The graphs in Fig. 1 show the two quantities for different choices of  $n$  and of  $\mathbf{u}_0$  when  $\phi_3 \neq 0$ . With the names “Constant”, “Logarithmic”, “Square root”, “Linear”, “Square” and “Exponential” we denote, in this order, the vectors with  $u_{0,j} \propto 1$ ,  $u_{0,j} \propto \ln(1+j)$ ,

---

<sup>5</sup>Since, as will be apparent later on, the results depend only marginally on  $d^2\mathbf{u}$ , we do not need to rewrite it.

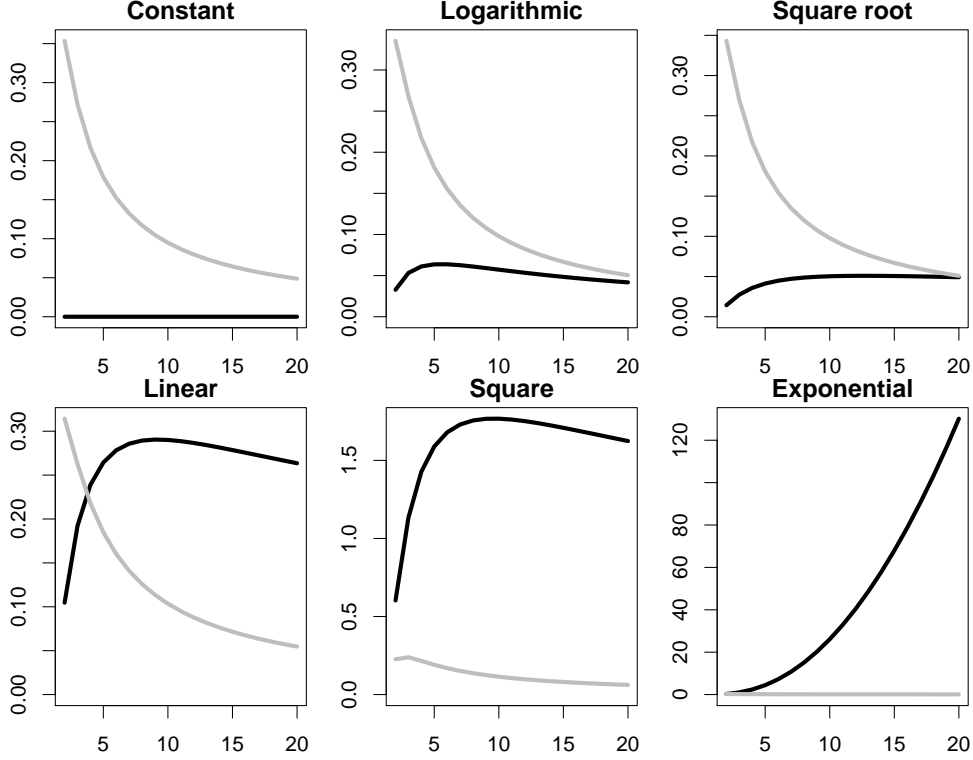


FIGURE 1:  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$  (black) and  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  (grey) for different  $n$  and  $\mathbf{u}_0$ .

$u_{0,j} \propto j^{1/2}$ ,  $u_{0,j} \propto j$ ,  $u_{0,j} \propto j^2$  and  $u_{0,j} \propto \exp(j-1)$ . These vectors represent situations of increasing dispersion of the real values of the weights. In this sense they relate to the so-called homogeneity axiom of the AHP. In particular, Saaty has always argued that “homogeneity is essential for comparing similar things, as the mind tend to make large errors in comparing widely disparate elements. For example we cannot compare a grain of sand with an orange according to size” (Saaty 1986, p. 846).

The graphs in the figure clarify the nature of this heuristic argument and the role of the cognitive or deterministic distortions due to the subjective weighting function  $W$  in it. The thick black lines in the graphs represent  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$  as a function of  $n$  while the thick grey lines represent  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  as a function of  $n$ . The index on the  $x$  axis starts at 2 since for  $n = 1$  both measures are identically 0. The first graphs in the figure are more likely to respect Saaty’s homogeneity requirement while the last ones are more prone not to respect it. The graphs show that if  $\phi_3 = \sigma$ , the effect of deterministic distortions (multiplied by  $\phi_3$ ) are larger than the effect of stochastic ones (multiplied by  $\sigma$ ) when the elements of  $\mathbf{u}_0$  are very different, while the effects of stochastic distortions are larger when the elements are quite similar.

## 4.1 Asymptotic distributions

Now, we derive the asymptotic distributions of Saaty's eigenvector  $\mathbf{u}$  and eigenvalue  $\lambda$  as the error standard deviation  $\sigma \downarrow 0$  and as  $W^{-1}(x) \rightarrow x$  (as measured by the coefficients of the polynomial through  $\|\phi_\ell\|_\infty \downarrow 0$ ). The results on the eigenvalue are expressed in terms of Saaty's consistency index  $\mu = (\lambda - n)/(n - 1)$  to facilitate their interpretation and the comparison with previous literature. The results are summarized in the following Theorem. It shows that the asymptotic distributions depend on the speed of convergence to 0 of  $\sigma$  and  $\|\phi_\ell\|_\infty$ .

**Theorem 4.1.** *Suppose that  $\max\{\|\phi_\ell\|_\infty, \sigma\} \downarrow 0$ . Consider the errors  $\nu_{ij}$  with  $i < j$ : suppose that they are (asymptotically) independent of each other and  $\sigma^{-1}\nu_{ij} \rightarrow_{\mathcal{D}} \mathcal{N}(0, 1)$ . Define  $\boldsymbol{\mu}_\phi \triangleq \sum_{\ell=2}^L \phi_\ell \cdot [(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\mathbf{I}}\mathbf{u}_0)]^{\odot \ell}$ ,  $\boldsymbol{\mu} \triangleq \lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty}$  and  $\mathbf{U}_0 \triangleq \text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top$ .*

a) *If  $\sigma/\|\phi_\ell\|_\infty \rightarrow +\infty$ :*

i)

$$\sigma^{-1}(\mathbf{u} - \mathbf{u}_0) \rightarrow_{\mathcal{D}} \mathcal{N}\left(\mathbf{0}, \frac{1}{n^2} \cdot (\mathbf{U}_n \otimes \mathbf{U}_0^2)\right);$$

ii)

$$\sigma^{-2}\mu \rightarrow_{\mathcal{D}} \frac{1}{n(n-1)} \cdot \chi^2\left(\frac{(n-1)(n-2)}{2}\right).$$

b) *If  $\sigma/\|\phi_\ell\|_\infty \rightarrow c$ :*

i)

$$\sigma^{-1}(\mathbf{u} - \mathbf{u}_0) \rightarrow_{\mathcal{D}} \mathcal{N}\left(\frac{1}{cn} \cdot (\mathbf{u}_n^\top \otimes \mathbf{U}_0) \cdot \boldsymbol{\mu}, \frac{1}{n^2} \cdot (\mathbf{U}_n \otimes \mathbf{U}_0^2)\right);$$

ii)

$$\sigma^{-2}\mu \rightarrow_{\mathcal{D}} \frac{1}{n(n-1)} \cdot \chi^2\left(\frac{(n-1)(n-2)}{2}; \delta\right)$$

where:

$$\delta \triangleq \frac{1}{c^2} \cdot \boldsymbol{\mu}^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \boldsymbol{\mu}.$$

c) *If  $\sigma/\|\phi_\ell\|_\infty \rightarrow 0$ :*

i)

$$\begin{aligned} \lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{1}{\|\phi_\ell\|_\infty} \cdot (\mathbf{u} - \mathbf{u}_0) &= \frac{1}{n} \cdot (\mathbf{u}_n^\top \otimes \mathbf{U}_0) \cdot \boldsymbol{\mu} \\ \sigma^{-1}\left(\mathbf{u} - \mathbf{u}_0 - \frac{1}{n} \cdot (\mathbf{u}_n^\top \otimes \mathbf{U}_0) \cdot \boldsymbol{\mu}_\phi\right) &\rightarrow_{\mathcal{D}} \mathcal{N}\left(0; \frac{1}{n^2} \cdot \mathbf{U}_n \otimes \mathbf{U}_0^2\right); \end{aligned}$$

ii)

$$\begin{aligned}
\lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{1}{\|\phi_\ell\|_\infty^2} \cdot \mu &= \frac{1}{n-1} \cdot \boldsymbol{\mu}^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \boldsymbol{\mu} \\
\frac{1}{\sigma \cdot \|\phi_\ell\|_\infty} \left( \mu - \frac{1}{n-1} \cdot \boldsymbol{\mu}_\phi^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \boldsymbol{\mu}_\phi \right) \\
&\rightarrow_{\mathcal{D}} \mathcal{N} \left( 0; \frac{4}{(n-1)^2} \cdot \boldsymbol{\mu}^\top \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^3} \right. \right. \\
&\quad \left. \left. + \frac{2(\mathbf{U}_n \otimes \mathbf{I}_n)\mathbf{K}_{nn}}{n^3} + \frac{(\mathbf{I}_{n^2} - \mathbf{K}_{nn})}{n^2} - \frac{4\mathbf{U}_{n^2}}{n^4} \right\} \boldsymbol{\mu} \right).
\end{aligned}$$

Theorem 4.1 distinguishes three cases. In the first case in which  $\sigma/\|\phi_\ell\|_\infty \rightarrow +\infty$ , the leading term is given by the stochastic perturbation. This case encompasses Genest and Rivest (1994) in which  $W(\cdot)$  is the identity and the asymptotic distributions of  $\mathbf{u}$  and  $\mu$  are respectively the normal and the  $\chi^2$  distribution (see Genest and Rivest 1994, p. 490). In particular, take the approximation  $\chi^2(\alpha) \approx n(n-1)\mu_0/10\sigma_0^2$ , where  $\chi^2(\alpha)$  stands for the  $100(1-\alpha)\%$  quantile of the  $\chi^2$  with  $p = (n-1)(n-2)/2$  degrees of freedom and  $\mu_0$  is Saaty's eigenvalue-based random index for matrices of size  $n$  whose entries chosen at random within an admissible range. For this case Genest and Rivest (1994, p. 490) have shown that the so-called Saaty's 10% cut-off rule of declaring incoherent a response matrix  $\mathbf{A}$  is in fact equivalent to a  $\chi^2$ -test at significance level  $\alpha$  of the hypothesis  $H_0 : \sigma^2 \leq \sigma_0^2$ , that the background noise in the response data does not exceed some threshold level  $\sigma_0$ .

This equivalence, however, does no longer hold for the other two cases considered in Theorem 4.1, where the leading term in the approximations is no longer given by the stochastic perturbation.

The second case covers in particular the situation in which the deterministic and the stochastic perturbations are comparable and shows that the distribution of  $\mu$  can be different from the  $\chi^2$  distribution obtained above. In particular this shows that the interpretation of Saaty's 10% cut-off rule as a test of the hypothesis  $H_0 : \sigma \leq \sigma_0^2$  advanced in Genest and Rivest (1994) can lead to distortions when  $\delta > 0$ . Indeed, even if the interpretation is perfectly legitimate when the stochastic perturbation is the leading one, when a deterministic perturbation is present, this may lead to severely undersized tests.

The third and last case applies when the deterministic distortion induced by  $W$  is even larger than the stochastic one.<sup>6</sup> Here most of the variability in  $\mathbf{u}$  and  $\mu$  is due to systematic distortions and the asymptotic distribution of  $\mu$  is so shifted to the right that it behaves as a normal distribution.

---

<sup>6</sup>By the way, it is interesting to notice that the empirical analysis in Bernasconi, Choirat and Seri (2009) shows that the cases in which the stochastic terms are the leading terms in the perturbations are indeed the least likely in practice.



## 5 Ratio Magnitude Estimation

An analogue of Theorem 3.3, and in particular of equation (7) and its application to the AHP (8), can be obtained when we work on a single row of the AHP response matrix compared with a reference  $x_0$ . This is a standard Stevens' ratio magnitude estimation experiment. We ask to compare the stimuli  $(x_1, \dots, x_n)$  with a baseline stimulus  $x_0$ . Each pairwise comparison yields the value  $\alpha_{i0} = W^{-1} \left( \frac{\psi(x_i)}{\psi(x_0)} \right) \cdot e^{\sigma \nu_{i0}}$ ; here  $d\varepsilon_{i0} = \ln \left[ \frac{\psi(x_0)}{\psi(x_i)} \cdot W^{-1} \left( \frac{\psi(x_i)}{\psi(x_0)} \right) \right] + \sigma \nu_{i0}$ . Recall that one reason for the AHP to use the pairwise comparison matrix is exactly the fact that increasing the number of comparisons among the  $n$  items of a given set increases the amount of information and should generate better estimates.<sup>7</sup>

In this case the following theorem holds.

**Theorem 5.1.** *Let  $\boldsymbol{\nu} = \begin{bmatrix} \nu_{10} & \nu_{20} & \dots & \nu_{n0} \end{bmatrix}^\top$ . Under the above-described assumptions:*

$$\begin{aligned} d\mathbf{u} &= \left( \text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top \right) \cdot \begin{bmatrix} d\varepsilon_{10} & d\varepsilon_{20} & \dots & d\varepsilon_{n0} \end{bmatrix}^\top \\ &= \left( \text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top \right) \cdot \left\{ \sum_{\ell=2}^L \phi_\ell \cdot [\overline{\ln \mathbf{u}_0} - \ln u_{0,0}]^{\odot \ell} + \sigma \boldsymbol{\nu} + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma) \right\}. \end{aligned}$$

In order to compare the vector  $\mathbf{u}$  estimated through Saaty's eigenvector method or directly obtained through ratio magnitude estimation, we compute also for the present method the bias  $\frac{\|\mathbb{E}d\mathbf{u}\|}{\phi_3}$  and the variance  $\frac{\sqrt{\mathbb{E}\|d\mathbf{u} - \mathbb{E}d\mathbf{u}\|^2}}{\sigma}$ . We have:

$$\frac{\|\mathbb{E}d\mathbf{u}\|}{\phi_3} = \left( \text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top \right) \cdot [\overline{\ln \mathbf{u}_0} - \ln u_{0,0}]^{\odot 3}.$$

We notice that in this case the term depends on  $n$ , on  $\mathbf{u}_0$  but also on the reference point  $u_{0,0} = \psi(x_0)$ . On the other hand the variance term is easily computable:

$$\frac{\sqrt{\mathbb{E}\|d\mathbf{u} - \mathbb{E}d\mathbf{u}\|^2}}{\sigma} = \sqrt{\sum_{i=1}^n u_{0,i}^2 + \left( \sum_{i=1}^n u_{0,i}^2 \right)^2 - 2 \sum_{i=1}^n u_{0,i}^3}.$$

This shows that the AHP reduces the variance of the distortion by a factor  $1/\sqrt{n}$ . As concerns the bias, on the other hand, the situation is less clear and it is uncertain which method is better. This is in a sense surprising since it would seem that more information would allow for better estimates of  $\mathbf{u}$ ; however, the reason is that when  $W^{-1}$  is far from the identity, the principal eigenvector method misuses the additional information to an

<sup>7</sup>Obviously, notice that this is different from increasing the number  $n$  of the items being compared in a response matrix, which could generate just the opposite effect of increasing the level of inconsistency in the data. The latter observation was anticipated by Saaty (1977) and has been commented extensively by Genest and Rivest (1994, p. 494-495) on the basis of results equivalent to those referred in the first case of Theorem 4.1 above.

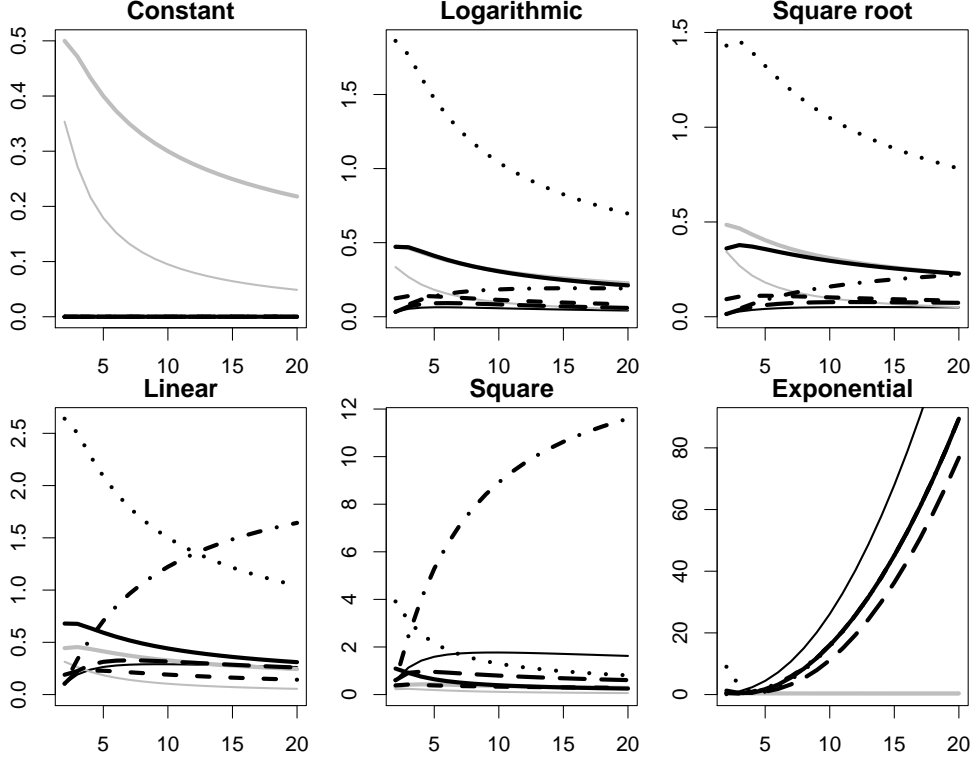


FIGURE 2: Comparisons of  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$  and  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  in AHP versus Ratio Magnitude estimation for different  $n$  and  $\mathbf{u}_0$ .

extent which could produce even more bias than that occurring due to the lower amount of information provided by a ratio magnitude estimation.

The graphs in Fig. 2 exemplify the problem. They show the quantities  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$  and  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  for several values of  $u_{0,0}$ ; as in the previous Figure, the names “Constant”, “Logarithmic”, “Square root”, “Linear”, “Square” and “Exponential” denote in this order the vectors with  $u_{0,j} \propto 1$ ,  $u_{0,j} \propto \ln(1+j)$ ,  $u_{0,j} \propto j^{1/2}$ ,  $u_{0,j} \propto j$ ,  $u_{0,j} \propto j^2$  and  $u_{0,j} \propto \exp(j-1)$ , while the index  $n$  is to be read on the horizontal axis. Here too,  $n$  starts from 2. The thin grey line shows  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  for Saaty’s method (the same as in the previous Figure) while the thick grey line shows  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  for a ratio magnitude experiment (independent of  $u_{0,0}$ ). The thin black line shows  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$  for Saaty’s method, while  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$  for a ratio magnitude experiment is displayed in the thick black lines: the dashed line has  $u_{0,0} = (\prod_{i=1}^n u_{0,i})^{1/2n}$ , the solid line has  $u_{0,0} = (\prod_{i=1}^n u_{0,i})^{1/n}$ , the dotted line has  $u_{0,0} = (\prod_{i=1}^n u_{0,i})^{2/n}$ , the dash dot line has  $u_{0,0} = \min_i u_{0,i}$ , the long-dashed line has  $u_{0,0} = \max_i u_{0,i}$ . It appears that  $\frac{\sqrt{\mathbb{E}\|\mathbf{d}\mathbf{u}-\mathbb{E}\mathbf{d}\mathbf{u}\|^2}}{\sigma}$  is always smaller for Saaty’s method, and the ratio increases with  $n$ . As concerns  $\frac{\|\mathbb{E}\mathbf{d}\mathbf{u}\|}{\phi_3}$ , the situation is more complex. As expected, values of  $u_{0,0}$  far away from  $(\prod_{i=1}^n u_{0,i})^{1/n}$  (or any other measure of centrality of  $\mathbf{u}_0$ ) give larger values. The ratio magnitude estimation experiment using  $u_{0,0} = \max_i u_{0,i}$  yields better results than the one with  $u_{0,0} = \min_i u_{0,i}$ . Moreover Saaty’s method yields smaller

values of  $\frac{\|\mathbb{R}d\mathbf{u}\|}{\phi_3}$  for  $\mathbf{u}_0$  respecting the homogeneity requirement, while when the values in  $\mathbf{u}_0$  are very different Saaty's method is worse than simple ratio magnitude estimation. This has two consequences: first of all, it stresses the relevance of homogeneity. Second, it subverts a widely believed idea: according to common sense and to the identification of random noise as the only source of distortion in the AHP, it is usually thought that Saaty's method is better than ratio magnitude estimation since it is based on a larger number of evaluations; this is false whenever homogeneity is not respected and psychophysical distortions are present in the data, while it is always true when all the distortion is due to random noise.

## 6 Conclusions

Recent developments in mathematical psychology, supported by various experimental tests, have shown that ratio-scaling methods in which individuals use number names to express proportions in which they perceive pairs of stimuli cannot be treated as scientific ratios.

The AHP is a ratio-scaling procedure widely used in management decisions. It is a more articulated method than the classical ratio magnitude estimation used in psychophysics. In the AHP an entire matrix of subjective ratio assessments is obtained and the maximum eigenvalue method is used to extract from the response matrix the single maximum eigenvector, the Perron eigenvector, which is then treated as the ratio scale of interest.

In this paper we have used recent developments in mathematical psychology based on the so-called separable forms, to study the type of distortions which can arise in the AHP when the maximum eigenvalue method is used. The analysis has emphasized the difference between the distortions due to random noise from the systematic or cognitive distortions embodied in the separable representations. The cognitive distortions highlight the importance of the so-called homogeneity axiom of the AHP to keep under control the bias arising in the estimate of the ratio scale.

We have also studied the asymptotic distributions of the maximum eigenvalue and maximum eigenvector under separable representations and have shown the limit of using the eigenvalue-based index of cardinal consistency of classical AHP as a rule to assess the quality of the estimate of the ratio scale.

The analysis has also shown that in some cases, when the cognitive distortions in the data are larger than those due to random noise and homogeneity is not fully respected, the classical ratio magnitude estimation used in psychophysics can be a better ratio-scaling procedure than the AHP, despite the greater amount of comparisons and information obtained by the latter method.

## References

- Aczél, J. and Luce, R.D. (2007). A behavioral condition for Prelec's weighting function on the positive line without assuming  $W(1) = 1$ . *Journal of Mathematical Psychology*, *51*, 126–129.
- Bernasconi, M., Choirat, C., and Seri, R. (2008). Measurement by subjective estimation: Testing for separable representations. *Journal of Mathematical Psychology*, *52*, 184–201.
- Bernasconi, M., Choirat, C., and Seri, R. (2009). The Analytic Hierarchy Process and the Theory of Measurement, working paper.
- Crawford, G.B. and Williams, C. (1985). A note on the analysis of subjective judgment matrices. The geometric mean procedure for estimating the scale of a judgment matrix. *Journal of Mathematical Psychology*, *29*, 387–405.
- de Jong, P. (1984). A statistical approach to Saaty's scaling method for priorities. *Journal of Mathematical Psychology*, *28*, 467–478.
- Ellermeier, W. and Faulhammer, G. (2000). Empirical evaluation of axioms fundamental to Stevens' ratio-scaling approach. I. Loudness production. *Perception & Psychophysics*, *62*, 1505–1511.
- Genest, C. and Rivest, L.-P. (1994). A statistical look at Saaty's method of estimating pairwise preferences expressed on a ratio scale. *Journal of Mathematical Psychology*, *38*, 477–496.
- Luce, R.D. (1997). Several unresolved problems of mathematical psychology. *Journal of Mathematical Psychology*, *41*, 79–97.
- Luce, R.D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, *109*, 520–532.
- Luce, R.D. (2004). Symmetric and asymmetric matching of joint presentations. *Psychological Review*, *111*, 446–454.
- Magnus, J.R., and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley Series in Probability and Statistics. Chichester: John Wiley and Sons Limited.
- Michell, J. (1999). *Measurement in Psychology: a Critical History of a Methodological Concept*. Cambridge, UK: Cambridge University Press.
- Narens, L. (1996). A theory of ratio magnitude estimation. *Journal of Mathematical Psychology*, *40*, 109–788.

- Narens, L. (2002). The irony of measurement by subjective estimations. *Journal of Mathematical Psychology*, *46*, 769–129.
- Narens, L. (2006). Symmetry, direct measurement, and Torgerson’s conjecture. *Journal of Mathematical Psychology*, *50*, 290–301.
- Newey, W.K., and McFadden, D. (1999). Large sample estimation and hypothesis testing. In D. McFadden and R. Engle (eds.), *Handbook of Econometrics*, (Vol. 4). Amsterdam: Elsevier, North-Holland.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.
- Saaty, T.L. (1977). A scaling method for priorities in a hierarchichal structure. *Journal of Mathematical Psychology*, *15*, 234–281.
- Saaty, T.L. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.
- Saaty, T.L. (1986). Axiomatic foundation of the analytic hierarchy process. *Management Science*, *32*, 841–855.
- Saaty, T.L., and Aczél, J. (1983). A procedures for synthesizing ratio judgments. *Journal of Mathematical Psychology*, *27*, 93–102.
- Saaty, T.L., and Vargas, L.G. (1984). Inconsistency and rank preservation. *Journal of Mathematical Psychology*, *28*, 205–214.
- Steingrimsson, R., and Luce, R.D. (2005a). Evaluating a model of global psychophysical judgments-I: behavioral properties of summations and productions. *Journal of Mathematical Psychology*, *49*, 290–307.
- Steingrimsson, R., and Luce, R.D. (2005b). Evaluating a model of global psychophysical judgments-II: behavioral properties linking summations and productions. *Journal of Mathematical Psychology*, *49*, 308–319.
- Steingrimsson, R., and Luce, R.D. (2006). Empirical evaluation of a model of global psychophysical judgments: III. A form for the psychophysical function and perceptual filtering. *Journal of Mathematical Psychology*, *50*, 15–29.
- Steingrimsson, R., and Luce, R.D. (2007). Empirical evaluation of a model of global psychophysical judgments: IV. Forms for the weighting function. *Journal of Mathematical Psychology*, *51*, 29–44.
- Stevens, S.S. (1951). *Handbook of Experimental Psychology*. New York: Wiley.
- Stevens, S.S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153–181.

Stevens, S.S. (1975). *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. New York: Wiley.

Stevens, S.S. and Galanter, E. (1957). Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology*, 54, 377–411.

Tan, W.Y. (1977). On the distribution of quadratic forms in normal random variables. *Canadian Journal of Statistics*, 5, 241–250.

Zimmer, K., (2005). Examining the validity of numerical ratios in loudness fractionation. *Perception & Psychophysics*, 67, 569–579.

## A Appendix - Proofs of propositions

**Proof of Theorem 3.1.** First of all we prove analyticity of  $\lambda$  and  $\mathbf{u}$ . Our proof follows the scheme of Theorem 7 on p. 158 in Magnus and Neudecker (1999, in the following MN) but is more complicated because of the nonstandard normalization of the eigenvector and of the nonsymmetry of  $(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)$ . Consider the vector function  $f : \mathbb{R}^{n+1} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n+1}$  defined by:

$$f(\mathbf{u}, \lambda; \mathbf{A}) = \begin{bmatrix} (\lambda \mathbf{I}_n - \mathbf{A}) \cdot \mathbf{u} \\ \mathbf{u}_n^\top \cdot \mathbf{u} - 1 \end{bmatrix}$$

(remark the difference in the second line with respect to the proof in MN).  $f$  is  $\infty$  times differentiable on  $\mathbb{R}^{n+1} \times \mathbb{R}^{n \times n}$  and  $f(\mathbf{u}_0, \lambda_0; \mathbf{A}_0) = \mathbf{0}$ . The matrix  $(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)$  has reduced rank  $n - 1$  since  $\lambda_0$  is a simple eigenvalue and we can apply Theorem 4 on p. 43 and Theorem 3 on p. 41 in MN:

$$\begin{aligned} \begin{vmatrix} \lambda_0 \mathbf{I}_n - \mathbf{A}_0 & \mathbf{u}_0 \\ \mathbf{u}_n^\top & 0 \end{vmatrix} &= -\mathbf{u}_n^\top (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^\sharp \mathbf{u}_0 \\ &= -\mathbf{u}_n^\top \left[ \mu (\lambda_0 \mathbf{I}_n - \mathbf{A}_0) \frac{\mathbf{u}_0 \mathbf{v}_0^\top}{\mathbf{v}_0^\top \mathbf{u}_0} \right] \mathbf{u}_0 \\ &= -\mu (\lambda_0 \mathbf{I}_n - \mathbf{A}_0) \neq 0 \end{aligned}$$

where  $\mathbf{B}^\sharp$  is the adjoint matrix of  $\mathbf{B}$  defined on p. 40 of MN, and  $\mu(\mathbf{B})$  is the product of the non-zero eigenvalues of  $\mathbf{B}$ . This implies that the conditions of the Implicit Function Theorem (Theorem A.3 in the Appendix of Chapter 7 in MN) are satisfied and there exists a neighborhood  $N(\mathbf{A}_0) \subset \mathbb{R}^{n \times n}$  of  $\mathbf{A}_0$ , a unique real-valued function  $\lambda : N(\mathbf{A}_0) \rightarrow \mathbb{R}$  and a unique (up to the sign) vector function  $\mathbf{u} : N(\mathbf{A}_0) \rightarrow \mathbb{R}^n$  such that:

1.  $\lambda$  and  $\mathbf{u}$  are  $\infty$  times differentiable on  $N(\mathbf{A}_0)$ ;
2.  $\lambda(\mathbf{A}_0) = \lambda_0$  and  $\mathbf{u}(\mathbf{A}_0) = \mathbf{u}_0$ ;
3.  $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$ ,  $\mathbf{u}_n^\top \mathbf{u} = 1$  for every  $\mathbf{A} \in N(\mathbf{A}_0)$ .

Therefore, the following expansion holds:

$$\begin{aligned}\lambda(\mathbf{A}) &= \lambda(\mathbf{A}_0) + d\lambda + \frac{1}{2}d^2\lambda + o(d^2\lambda), \\ \mathbf{u}(\mathbf{A}) &= \mathbf{u}(\mathbf{A}_0) + d\mathbf{u} + o(\|d\mathbf{u}\|).\end{aligned}$$

Differentiating  $\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}$  around the point  $(\mathbf{A}, \lambda, \mathbf{u}) = (\mathbf{A}_0, \lambda_0, \mathbf{u}_0)$  we get:

$$d\mathbf{A} \cdot \mathbf{u}_0 + \mathbf{A}_0 \cdot d\mathbf{u} = d\lambda \cdot \mathbf{u}_0 + \lambda_0 \cdot d\mathbf{u} \quad (9)$$

and premultiplying this by  $\mathbf{v}_0^\top$  we get:

$$\begin{aligned}\mathbf{v}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0 + \mathbf{v}_0^\top \cdot \mathbf{A}_0 \cdot d\mathbf{u} &= d\lambda \cdot \mathbf{v}_0^\top \mathbf{u}_0 + \lambda_0 \cdot \mathbf{v}_0^\top \cdot d\mathbf{u} \\ \mathbf{v}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0 &= d\lambda \cdot \mathbf{v}_0^\top \mathbf{u}_0 \\ d\lambda &= \frac{\mathbf{v}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0}{\mathbf{v}_0^\top \mathbf{u}_0}.\end{aligned} \quad (10)$$

Now, we take the second differential of  $\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}$  around the point  $(\mathbf{A}, \lambda, \mathbf{u}) = (\mathbf{A}_0, \lambda_0, \mathbf{u}_0)$ :

$$d^2\mathbf{A} \cdot \mathbf{u}_0 + 2d\mathbf{A} \cdot d\mathbf{u} + \mathbf{A}_0 \cdot d^2\mathbf{u} = 2d\lambda \cdot d\mathbf{u} + d^2\lambda \cdot \mathbf{u}_0 + \lambda_0 \cdot d^2\mathbf{u} \quad (11)$$

and premultiplying it by  $\mathbf{v}_0^\top$  we get:

$$\begin{aligned}\mathbf{v}_0^\top \cdot d^2\mathbf{A} \cdot \mathbf{u}_0 + 2\mathbf{v}_0^\top \cdot d\mathbf{A} \cdot d\mathbf{u} + \mathbf{v}_0^\top \cdot \mathbf{A}_0 \cdot d^2\mathbf{u} &= 2d\lambda \cdot \mathbf{v}_0^\top \cdot d\mathbf{u} + d^2\lambda \cdot \mathbf{v}_0^\top \cdot \mathbf{u}_0 + \lambda_0 \cdot \mathbf{v}_0^\top \cdot d^2\mathbf{u} \\ \mathbf{v}_0^\top \cdot d^2\mathbf{A} \cdot \mathbf{u}_0 + 2\mathbf{v}_0^\top \cdot d\mathbf{A} \cdot d\mathbf{u} &= 2d\lambda \cdot \mathbf{v}_0^\top \cdot d\mathbf{u} + d^2\lambda \cdot \mathbf{v}_0^\top \cdot \mathbf{u}_0\end{aligned}$$

and:

$$\begin{aligned}d^2\lambda &= \frac{\mathbf{v}_0^\top \cdot d^2\mathbf{A} \cdot \mathbf{u}_0 + 2\mathbf{v}_0^\top \cdot d\mathbf{A} \cdot d\mathbf{u} - 2d\lambda \cdot \mathbf{v}_0^\top \cdot d\mathbf{u}}{\mathbf{v}_0^\top \cdot \mathbf{u}_0} \\ &= \frac{\mathbf{v}_0^\top \mathbf{u}_0 \cdot \mathbf{v}_0^\top d^2\mathbf{A} \mathbf{u}_0 + 2\mathbf{v}_0^\top \mathbf{u}_0 \cdot \mathbf{v}_0^\top d\mathbf{A} d\mathbf{u} - 2\mathbf{v}_0^\top d\mathbf{A} \mathbf{u}_0 \cdot \mathbf{v}_0^\top d\mathbf{u}}{(\mathbf{v}_0^\top \cdot \mathbf{u}_0)^2}\end{aligned} \quad (12)$$

where  $d\mathbf{u}$  will be obtained in the following.

We start from  $\mathbf{A}_0 \cdot \mathbf{u}_0 = \lambda_0 \cdot \mathbf{u}_0$  and we define as  $\mathbf{u}_0$  the vector normalized as  $\mathbf{u}'_0 \mathbf{u}_n = 1$  and as  $\tilde{\mathbf{u}}_0$  the vector normalized as  $\tilde{\mathbf{u}}'_0 \tilde{\mathbf{u}}_0 = 1$ . We have  $\mathbf{u}_0 = \tilde{\mathbf{u}}_0 / \tilde{\mathbf{u}}'_0 \mathbf{u}_n$ . In the same way, we consider  $\mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{u}$  and we define as  $\mathbf{u}$  the vector normalized as  $\mathbf{u}' \mathbf{u}_n = 1$  and as  $\tilde{\mathbf{u}}$  the vector normalized as  $\tilde{\mathbf{u}}' \tilde{\mathbf{u}}_0 = 1$ . We have  $\mathbf{u} = \tilde{\mathbf{u}} / \tilde{\mathbf{u}}' \mathbf{u}_n$ . On the other hand

$\tilde{\mathbf{u}} = \tilde{\mathbf{u}}_0 + d\tilde{\mathbf{u}} + o(\|d\tilde{\mathbf{u}}\|)$  and:

$$\begin{aligned}
\mathbf{u} &= \frac{\tilde{\mathbf{u}}}{\tilde{\mathbf{u}}' \mathbf{u}_n} = \frac{\tilde{\mathbf{u}}_0 + d\tilde{\mathbf{u}} + o(\|d\tilde{\mathbf{u}}\|)}{(\tilde{\mathbf{u}}_0 + d\tilde{\mathbf{u}} + o(\|d\tilde{\mathbf{u}}\|))' \mathbf{u}_n} \\
&= \frac{\tilde{\mathbf{u}}_0 + d\tilde{\mathbf{u}} + o(\|d\tilde{\mathbf{u}}\|)}{\tilde{\mathbf{u}}_0' \mathbf{u}_n \cdot \left(1 + \frac{d\tilde{\mathbf{u}}' \mathbf{u}_n}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} + o\left(\frac{\|d\tilde{\mathbf{u}}\|}{\tilde{\mathbf{u}}_0' \mathbf{u}_n}\right)\right)} \\
&= \left(\frac{\tilde{\mathbf{u}}_0}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} + \frac{d\tilde{\mathbf{u}}}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} + o\left(\frac{\|d\tilde{\mathbf{u}}\|}{\tilde{\mathbf{u}}_0' \mathbf{u}_n}\right)\right) \cdot \left(1 - \frac{d\tilde{\mathbf{u}}' \mathbf{u}_n}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} + o\left(\frac{\|d\tilde{\mathbf{u}}\|}{\tilde{\mathbf{u}}_0' \mathbf{u}_n}\right)\right) \\
&= \frac{\tilde{\mathbf{u}}_0}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} - \frac{\tilde{\mathbf{u}}_0}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} \cdot \frac{d\tilde{\mathbf{u}}' \mathbf{u}_n}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} + \frac{d\tilde{\mathbf{u}}}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} + o(\|d\tilde{\mathbf{u}}\|) \\
&= \mathbf{u}_0 + \frac{1}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} \cdot (\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n') \cdot d\tilde{\mathbf{u}} + o(\|d\tilde{\mathbf{u}}\|)
\end{aligned}$$

From the result in Theorem 8 on page 161 in MN, we have:

$$d\tilde{\mathbf{u}} = (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left( \mathbf{I}_n - \frac{\tilde{\mathbf{u}}_0 \tilde{\mathbf{v}}_0'}{\tilde{\mathbf{v}}_0' \tilde{\mathbf{u}}_0} \right) (d\mathbf{A}) \tilde{\mathbf{u}}_0.$$

Using the proportionality between  $\tilde{\mathbf{u}}_0$  and  $\mathbf{u}_0$  on one hand, and  $\tilde{\mathbf{v}}_0$  and  $\mathbf{v}_0$  on the other hand, we can write:

$$\begin{aligned}
\mathbf{u} &= \mathbf{u}_0 + \frac{1}{\tilde{\mathbf{u}}_0' \mathbf{u}_n} \cdot (\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n') \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left( \mathbf{I}_n - \frac{\tilde{\mathbf{u}}_0 \tilde{\mathbf{v}}_0'}{\tilde{\mathbf{v}}_0' \tilde{\mathbf{u}}_0} \right) (d\mathbf{A}) \tilde{\mathbf{u}}_0 + o(\|d\tilde{\mathbf{u}}\|) \\
&= \mathbf{u}_0 + (\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n') \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left( \mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0'}{\mathbf{v}_0' \mathbf{u}_0} \right) (d\mathbf{A}) \mathbf{u}_0 + o(\|d\tilde{\mathbf{u}}\|)
\end{aligned}$$

From this, we get:

$$\begin{aligned}
d^2\lambda &= \frac{\mathbf{v}_0' \mathbf{u}_0 \cdot \mathbf{v}_0' d^2 \mathbf{A} \mathbf{u}_0}{(\mathbf{v}_0' \cdot \mathbf{u}_0)^2} \\
&\quad + \frac{2\mathbf{v}_0' (\mathbf{u}_0 \mathbf{v}_0' d\mathbf{A} - d\mathbf{A} \mathbf{u}_0 \cdot \mathbf{v}_0') \cdot (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ \left( \mathbf{I}_n - \frac{\mathbf{u}_0 \mathbf{v}_0'}{\mathbf{v}_0' \mathbf{u}_0} \right) (d\mathbf{A}) \mathbf{u}_0}{(\mathbf{v}_0' \cdot \mathbf{u}_0)^2}
\end{aligned}$$

From (11) we get:

$$(\lambda_0 \mathbf{I}_n - \mathbf{A}_0) \cdot d^2 \mathbf{u} = (d^2 \mathbf{A} - d^2 \lambda \mathbf{I}_n) \cdot \mathbf{u}_0 + 2(d\mathbf{A} - d\lambda \mathbf{I}_n) \cdot d\mathbf{u}.$$

Premultiplying the equality with  $(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+$ , we apply the arguments on p. 160 in the proof of Theorem 7 on p. 158 of MN and we get:

$$\begin{aligned}
d^2 \mathbf{u} &= (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (\lambda_0 \mathbf{I}_n - \mathbf{A}_0) \cdot d^2 \mathbf{u} \\
&= (\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (d^2 \mathbf{A} - d^2 \lambda \mathbf{I}_n) \cdot \mathbf{u}_0 + 2(\lambda_0 \mathbf{I}_n - \mathbf{A}_0)^+ (d\mathbf{A} - d\lambda \mathbf{I}_n) \cdot d\mathbf{u}
\end{aligned}$$



**Proof of Theorem 3.2.** Replacing  $\mathbf{u}_0$  and  $\mathbf{v}_0$  with  $\mathbf{u}_0$  and  $\bar{\mathbf{u}}_0$ , we get:

$$\begin{aligned} d\lambda &= \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0}{\bar{\mathbf{u}}_0^\top \mathbf{u}_0} \\ d^2\lambda &= \frac{\bar{\mathbf{u}}_0^\top d^2\mathbf{A} \mathbf{u}_0}{n} + \frac{2\bar{\mathbf{u}}_0^\top (d\mathbf{A}) (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top) \cdot (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top)^+ (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top) (d\mathbf{A}) \mathbf{u}_0}{n^3} \\ d\mathbf{u} &= \frac{1}{n} \cdot (\mathbf{I}_n - \mathbf{u}_0 \mathbf{u}_n^\top) \cdot (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top)^+ (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top) (d\mathbf{A}) \mathbf{u}_0 \end{aligned}$$

The result follows from the properties of the Moore-Penrose inverse.

$$\begin{aligned} d\lambda &= \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0}{n} \\ d^2\lambda &= 2 \cdot \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot (n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top) \cdot d\mathbf{A} \cdot \mathbf{u}_0}{n^3} + \frac{\bar{\mathbf{u}}_0^\top \cdot d^2\mathbf{A} \cdot \mathbf{u}_0}{n} \\ d\mathbf{u} &= \frac{1}{n} \cdot (\mathbf{I}_n - \mathbf{u}_0 \cdot \mathbf{u}_n^\top) \cdot d\mathbf{A} \cdot \mathbf{u}_0 \\ d^2\mathbf{u} &= (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top)^+ \\ &\quad \left( d^2\mathbf{A} - 2 \cdot \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot (n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top) \cdot d\mathbf{A} \cdot \mathbf{u}_0}{n^3} \mathbf{I}_n - \frac{\bar{\mathbf{u}}_0^\top \cdot d^2\mathbf{A} \cdot \mathbf{u}_0}{n} \mathbf{I}_n \right) \cdot \mathbf{u}_0 \\ &\quad + 2 (n\mathbf{I}_n - \mathbf{u}_0 \bar{\mathbf{u}}_0^\top)^+ \left( d\mathbf{A} - \frac{\bar{\mathbf{u}}_0^\top \cdot d\mathbf{A} \cdot \mathbf{u}_0}{n} \mathbf{I}_n \right) \cdot \frac{1}{n} \cdot (\mathbf{I}_n - \mathbf{u}_0 \cdot \mathbf{u}_n^\top) \cdot d\mathbf{A} \cdot \mathbf{u}_0. \end{aligned}$$

**Proof of Theorem 3.3.** First of all, applying an element-wise expansion to  $\mathbf{A} = \mathbf{A}_0 \odot \overline{\exp}(d\mathbf{E})$  remark that  $\mathbf{A} = \mathbf{A}_0 + d\mathbf{A} + \frac{1}{2}d^2\mathbf{A} + o(\|d^2\mathbf{A}\|)$  with  $d\mathbf{A} = \mathbf{A}_0 \odot d\mathbf{E}$  and  $d^2\mathbf{A} = \mathbf{A}_0 \odot d\mathbf{E} \odot d\mathbf{E}$ . The differentials  $d\mathbf{A}$  and  $d^2\mathbf{A}$  can then be written as:

$$\begin{aligned} d\mathbf{A} &= \text{diag}[\mathbf{u}_0] \cdot d\mathbf{E} \cdot \text{diag}[\bar{\mathbf{u}}_0] \\ d^2\mathbf{A} &= \text{diag}[\mathbf{u}_0] \cdot (d\mathbf{E} \odot d\mathbf{E}) \cdot \text{diag}[\bar{\mathbf{u}}_0]. \end{aligned}$$

This gives:

$$\begin{aligned} d\lambda &= \frac{\mathbf{u}_n^\top \cdot d\mathbf{E} \cdot \mathbf{u}_n}{n} = 0 \\ d^2\lambda &= \frac{\mathbf{u}_n^\top \cdot [2(d\mathbf{E} \cdot d\mathbf{E}) + n(d\mathbf{E} \odot d\mathbf{E})] \cdot \mathbf{u}_n}{n^2} \\ d\mathbf{u} &= \frac{1}{n} \cdot (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \cdot d\mathbf{E} \cdot \mathbf{u}_n \\ d^2\mathbf{u} &= (n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top)^+ \\ &\quad \left( \text{diag}[\mathbf{u}_0] \cdot (d\mathbf{E} \odot d\mathbf{E}) \cdot \mathbf{u}_n - \frac{\mathbf{u}_n^\top \cdot [2(d\mathbf{E} \cdot d\mathbf{E}) + n(d\mathbf{E} \odot d\mathbf{E})] \cdot \mathbf{u}_n}{n^2} \cdot \mathbf{u}_0 \right) \\ &\quad + \frac{2}{n} \cdot (n\mathbf{I}_n - \mathbf{u}_0 \cdot \bar{\mathbf{u}}_0^\top)^+ \text{diag}[\mathbf{u}_0] \cdot d\mathbf{E} \cdot (\mathbf{I}_n - \mathbf{u}_n \cdot \mathbf{u}_0^\top) \cdot d\mathbf{E} \cdot \mathbf{u}_n. \end{aligned}$$

**Proof of Theorem 4.1.** First of all, we remark that  $\mathbf{u} = \mathbf{u}_0 + d\mathbf{u} + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma)$  and  $\lambda = n + \frac{1}{2}d^2\lambda + o(\|\phi_\ell\|_\infty^2) + o_{\mathbb{P}}(\sigma^2)$ . In terms of Saaty's consistency index (i.e.  $\mu = \frac{\lambda-n}{n-1}$ ), we have  $\mu = \frac{d^2\lambda}{2(n-1)} + o(\|\phi_\ell\|_\infty^2) + o_{\mathbb{P}}(\sigma^2)$ . Therefore, using the commutation matrix, we derive some alternative formulation of the results of Theorem 3.3:

$$\begin{aligned}
d^2\lambda &= \frac{\mathbf{u}_n^\top \cdot 2(\mathbf{dE} \cdot \mathbf{dE}) \cdot \mathbf{u}_n}{n^2} + \frac{\mathbf{u}_n^\top \cdot (\mathbf{dE} \odot \mathbf{dE}) \cdot \mathbf{u}_n}{n} \\
&= \frac{2\text{vec}(\mathbf{dE}^\top \cdot \mathbf{u}_n)^\top \cdot \text{vec}(\mathbf{dE} \cdot \mathbf{u}_n)}{n^2} + \frac{\text{tr}(\mathbf{dE} \cdot \mathbf{dE}^\top)}{n} \\
&= \frac{2((\mathbf{u}_n^\top \otimes \mathbf{I}_n) \cdot \text{vec}(\mathbf{dE}^\top))^\top \cdot (\mathbf{u}_n^\top \otimes \mathbf{I}_n) \cdot \text{vec}(\mathbf{dE})}{n^2} + \frac{\text{vec}(\mathbf{dE})^\top \cdot \text{vec}(\mathbf{dE})}{n} \\
&= \text{vec}(\mathbf{dE})^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \text{vec}(\mathbf{dE}) \\
d\mathbf{u} &= \frac{1}{n} \cdot \left( \mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \right) \cdot \text{vec}(\mathbf{dE}).
\end{aligned}$$

We recall that  $\mathbf{dE} = \sum_{\ell=2}^L \phi_\ell \cdot [\overline{\ln \mathbf{A}_0}]^{\odot \ell} + \sigma \mathbf{N} + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma)$ . Moreover:

$$\overline{\ln \mathbf{A}_0} = \overline{\ln}(\mathbf{u}_0 \mathbf{u}_0^\top) = (\overline{\ln \mathbf{u}_0}) \cdot \mathbf{u}_n^\top - \mathbf{u}_n \cdot (\overline{\ln \mathbf{u}_0})^\top$$

and taking the vec's on both sides:

$$\begin{aligned}
\text{vec}(\overline{\ln \mathbf{A}_0}) &= \text{vec} \left[ (\overline{\ln \mathbf{u}_0}) \cdot \mathbf{u}_n^\top - \mathbf{u}_n \cdot (\overline{\ln \mathbf{u}_0})^\top \right] \\
&= \text{vec} \left[ (\overline{\ln \mathbf{u}_0}) \cdot \mathbf{u}_n^\top \right] - \text{vec} \left[ \mathbf{u}_n \cdot (\overline{\ln \mathbf{u}_0})^\top \right] \\
&= \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0}) - (\overline{\ln \mathbf{u}_0}) \otimes \mathbf{u}_n \\
&= (\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0}).
\end{aligned}$$

Therefore:

$$\begin{aligned}
\text{vec}(\mathbf{dE}) &= \sum_{\ell=2}^L \phi_\ell \cdot [\text{vec}(\overline{\ln \mathbf{A}_0})]^{\odot \ell} + \sigma \text{vec}(\mathbf{N}) + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma) \\
&= \sum_{\ell=2}^L \phi_\ell \cdot [(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{u}_n \otimes (\overline{\ln \mathbf{u}_0})]^{\odot \ell} + \sigma \text{vec}(\mathbf{N}) + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma) \\
&= \boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N}) + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma).
\end{aligned}$$

This leads us to the following expressions that will be used extensively in the following:

$$\begin{aligned}
\mathbf{u} &= \mathbf{u}_0 + \frac{1}{n} \cdot \left( \mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top) \right) \cdot [\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N})] + o(\|\phi_\ell\|_\infty) + o_{\mathbb{P}}(\sigma) \\
\mu &= \frac{1}{2(n-1)} \cdot [\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N})]^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot [\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N})] \\
&\quad + o_{\mathbb{P}}(\|\phi_\ell\|_\infty \cdot \sigma).
\end{aligned}$$

We need expressions for  $\mathbb{V}[\text{vec}(\mathbf{N})]$  and  $\mathbb{V}\left(\frac{1}{n} \cdot (\mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top)) \cdot \text{vec}(\mathbf{N})\right)$ . It can be shown that:

$$\mathbb{V}(\text{vec}(\mathbf{N})) = (\mathbf{I}_{n^2} - \mathbf{K}_{nn})$$

and:

$$\mathbb{V}\left(\frac{1}{n} \cdot (\mathbf{u}_n^\top \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top)) \cdot \text{vec}(\mathbf{N})\right) = \frac{1}{n^2} \cdot \left(\mathbf{U}_n \otimes (\text{diag}[\mathbf{u}_0] - \mathbf{u}_0 \cdot \mathbf{u}_0^\top)^2\right).$$

Now we pass to the proof of the statements.

a) The proof of the results in this case are exactly the same as those of point b) when  $c = +\infty$ .

b) We start from  $\sigma^{-1} \cdot (\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N}))$  and we remark that it tends to the following normal distribution:

$$\sigma^{-1} \cdot (\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N})) = \frac{\|\phi_\ell\|_\infty}{\sigma} \cdot \frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty} + \text{vec}(\mathbf{N}) \rightarrow_{\mathcal{D}} \frac{1}{c} \cdot \boldsymbol{\mu} + \mathcal{N}(\mathbf{0}, (\mathbf{I}_{n^2} - \mathbf{K}_{nn})).$$

From this the asymptotic result for  $\mathbf{u}$  follows. As concerns the result for  $\boldsymbol{\mu}$ , we use Theorem 3.1 in Tan (1977), identifying his  $\mu$  with  $\frac{1}{c} \cdot \lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty}$ , his  $A$  with  $\left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\}$  and his  $V$  with  $(\mathbf{I}_{n^2} - \mathbf{K}_{nn})$ . The eigenvalues of  $V \cdot A = (\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\}$  can be found reasoning as follows. We want to show that  $\mathbf{A}_1 \triangleq \frac{n}{2} \cdot (\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\}$  has  $\frac{(n-1)(n-2)}{2}$  eigenvalues equal to 1. We use the equality  $(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot (\mathbf{I}_{n^2} - \mathbf{K}_{nn}) = 2(\mathbf{I}_{n^2} - \mathbf{K}_{nn})$  and we remark that the eigenvalues of  $\mathbf{A}_1$  are the same as the ones of the symmetric matrix  $\mathbf{A}_2 \triangleq \frac{n}{4} \cdot (\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot (\mathbf{I}_{n^2} - \mathbf{K}_{nn})$ . Now, by exploiting the relations  $(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot (\mathbf{I}_{n^2} - \mathbf{K}_{nn}) = 2(\mathbf{I}_{n^2} - \mathbf{K}_{nn})$ ,  $(\mathbf{I}_{n^2} - \mathbf{K}_{nn}) \cdot \mathbf{K}_{nn} = (\mathbf{K}_{nn} - \mathbf{I}_{n^2})$  and  $\mathbf{K}_{nn} \mathbf{U}_{n^2} = \mathbf{U}_{n^2}$ , it is possible to show through some extremely tedious algebra, available from the authors upon request, that  $\mathbf{A}_2$  is idempotent. The number of nonzero eigenvalues is therefore given by the trace of matrix  $\mathbf{A}_1$ , that can be shown to be equal to  $\frac{(n-2)(n-1)}{2}$ . Since all  $\chi^2$  random variables appearing in the linear combination of Theorem 3.1 in Tan (1977) have the same weight, the asymptotic distribution is a noncentral chi-squared distribution whose noncentrality parameter can be computed as:

$$\delta \triangleq \frac{1}{c^2} \cdot \left[ \lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty} \right]^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \left[ \lim_{\|\phi_\ell\|_\infty \downarrow 0} \frac{\boldsymbol{\mu}_\phi}{\|\phi_\ell\|_\infty} \right].$$

c) Statement i) is simple. Statement ii) is obtained opening the square  $[\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N})]^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot [\boldsymbol{\mu}_\phi + \sigma \text{vec}(\mathbf{N})]$ , remarking that the leading term under  $\sigma/\|\phi_\ell\|_\infty \rightarrow 0$  is  $\boldsymbol{\mu}_\phi^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \boldsymbol{\mu}_\phi$  and that the second order term (the one that determines the asymptotic distribution) is  $2\boldsymbol{\mu}_\phi^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \sigma \text{vec}(\mathbf{N}) = O_{\mathbb{P}}(\|\phi_\ell\|_\infty \cdot \sigma)$  while

$\sigma^2 \text{vec}(\mathbf{N})^\top \cdot \left\{ \frac{2\mathbf{K}_{nn}(\mathbf{U}_n \otimes \mathbf{I}_n)}{n^2} + \frac{\mathbf{I}_{n^2}}{n} \right\} \cdot \text{vec}(\mathbf{N})$  is asymptotically negligible.